J. Mol. Biol. (1997) **272**, 9–20

JMB

COMMUNICATION

Comprehensive NOE Characterization of a Partially Folded Large Fragment of Staphylococcal Nuclease Δ131Δ, Using NMR Methods with Improved Resolution

Ouwen Zhang[1,2], Lewis E. Kay[2], David Shortle[3] and Julie D. Forman-Kay[1]*

[1]*Biochemistry Research Division, Hospital for Sick Children, 555 University Avenue, Toronto, Ontario Canada, M5G 1X8 and Department of Biochemistry University of Toronto, Toronto Ontario, Canada, M5S 1A8*

[2]*Protein Engineering Network Centres of Excellence and Departments of Medical Genetics, Biochemistry and Chemistry, University of Toronto, Toronto, Ontario Canada, M5S 1A8*

[3]*Department of Biological Chemistry, Johns Hopkins University School of Medicine Baltimore, MD, 21205, USA*

**Corresponding author*

Comprehensive NOE results from detailed structural characterization of a 131 residue partially folded fragment of staphylococcal nuclease (Δ131Δ) made possible by NMR methods with improved resolution are presented. The resulting NOE patterns reflect sampling of both α and β regions of φ, ψ conformational space, yet demonstrate significant preferences for both native-like and non-native-like turn and potentially helical conformations. Together with data from studies of the unfolded state of the drkN SH3 domain, NOE patterns observed for partially folded or unfolded proteins are summarized. It is surprising that few long-range NOEs were observed in Δ131Δ. The two longest-range NOEs are both native-like; one of these, an $(i,i+5)$ NOE, provides evidence for a Schellman capping motif for helix termination. Many aliphatic-aliphatic and aliphatic-amide NOEs, which are not normally observed in folded proteins, were detected. We have ruled out significant contributions from spin-diffusion for a number of these NOEs and suggest that one source may be sampling of non-prolyl *cis* peptide bond configurations in the disordered state of Δ131Δ.

Keywords: protein NMR; unfolded or partially folded states; *cis* peptide bond; Schellman motif; NOE patterns

Complete elucidation of protein folding mechanisms requires structural characterization of all protein states along folding pathways, i.e. unfolded states, folding intermediates and folded states. Structural information for unfolded or partially folded states of proteins (hereafter also referred to as disordered states), however, is still scarce compared to that for folded proteins. Disordered states of proteins are essentially ensembles of various conformations with the possibility of preferences for particular structural features. It has been shown that structural properties of denatured states generated under different denaturing conditions can vary with the conditions used (for reviews, see Dobson, 1992; Shortle, 1993). It is thus not clear what differences exist between unfolded states generated by denaturing agents and those populated under conditions which promote folding. Understanding the nature of these later states is important, since they are the biologically significant species recognized by chaperones, protein translocation machinery and proteasomes (Shortle, 1993). In addition, many proteins in the absence of their specific biological targets have been shown to display regions of conformational disorder (Kriwacki *et al.*, 1996; Shen *et al.*, 1996; Weinreb *et al.*, 1996). Therefore, studying unfolded proteins under conditions close to physiological instead of under strong denaturing conditions is likely to be

Present address: O. Zhang, Department of Molecular Biology, the Scripps Research Institute, La Jolla, CA 92037, USA.

Abbreviations used: Δ131Δ, a 131 residue fragment of staphylococcal nuclease; drkN SH3, the N-terminal SH3 domain of drk; HSQC, heteronuclear single quantum coherence; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; ppm, parts per million; 3-D, three-dimensional; TOCSY, total correlation spectroscopy; SNase, staphylococcal nuclease.

0022–2836/97/360009–12 $25.00/0/mb971219

more directly relevant to disordered states present *in vivo*.

A model system for the study of partially folded states of proteins under non-denaturing conditions is Δ131Δ, a 131 residue fragment of staphylococcal nuclease (SNase) with deletions of residues 4 to 12 and 141 to 149 of the wild-type protein (Alexandrescu *et al.*, 1994). This fragment has been the focus of experiments aimed at understanding the folding pathway of staphylococcal nuclease (Shortle *et al.*, 1996). Elucidating the structural preferences in the Δ131Δ partially folded state of SNase will provide valuable information to address this goal. In addition, this information will significantly enlarge the still extremely small database of structural data for unfolded or partially folded proteins. Another system for investigating unfolded states under folding conditions is the N-terminal SH3 domain of the *Drosophila* adapter protein drk. We have previously demonstrated the equilibrium between the folded and unfolded states of this SH3 domain under conditions close to physiological (Zhang *et al.*, 1994). Detailed structural characterization of this unfolded state and a denatured state in 2 M guanidine hydrochloride of this SH3 domain have been described elsewhere (Zhang & Forman-Kay, 1995, 1997). Nevertheless, brief summaries of those results will be presented here with the structural preferences observed in Δ131Δ.

Although there is significant conformational heterogeneity in these systems, NOEs can be observed and used to characterize even a small population of preferentially structured conformers due to the steep distance dependence of the NOE. Unambiguous assignment of NOEs is important for characterizing disordered states since they often demonstrate preferential sampling of particular conformations and these small populations may be important for directing the folding pathway. In this communication, the NOE patterns observed in Δ131Δ are summarized. Implications of the observed NOEs for residual structure in Δ131Δ and the drkN SH3 domain are briefly discussed. Particularly interesting is evidence for the presence of a Schellman motif in Δ131Δ and the observation of a significant number of sequential aliphatic-aliphatic and amide-aliphatic NOEs. These types of NOEs are generally not observed in folded proteins and possibly reflect the sampling of non-prolyl *cis* peptide bond configurations.

NOE-based methods with higher resolution

Multi-dimensional NMR techniques have greatly increased the size of proteins amenable to structural determination as well as the precision of the resulting structures. However, detailed structural analysis of unfolded or partially folded proteins is still a formidable task due to the extremely poor chemical shift dispersion of aliphatic ^{1}H and ^{13}C nuclei and the dynamic nature of these states, which will attenuate the rather weak NOE interactions existing in the small population of preferentially structured conformers (Shortle, 1996). The realization that even in these disordered states the backbone ^{15}N and $^{13}C'$ (carbonyl) nuclei still maintain dispersion comparable to ^{15}N and $^{13}C'$ nuclei in folded proteins and the often favorable relaxation properties in unfolded protein states have stimulated us to develop a suite of triple-resonance NOESY-based experiments to obtain unambiguous structural information for characterization of unfolded or partially folded states (Zhang *et al.*, 1997).

The newly developed NMR experiments can provide three different types of NOE connectivities, including NOEs between proximal aliphatic protons, between aliphatic and NH protons and between pairs of NH protons. In the case of aliphatic protons, the amide proton and nitrogen chemical shifts of the following residue are often recorded instead of those of the directly involved nuclei in order to exploit the dispersion of the amide nitrogen resonances. A detailed description of each of these pulse sequences is presented elsewhere (Zhang *et al.*, 1997). Here we demonstrate the utility of one particular experiment, the C_i-NOESY-$N_{j+1}H_{j+1}$ experiment, in which NOEs between carbon-bound protons of residue i and α protons of residue j are observed. Obtaining specific NOE information involving α protons is extremely useful for identifying residual secondary structure in partially folded proteins since $d_{\alpha\beta}(i,i+3)$ and $d_{\alpha\alpha}(i,j)$ NOEs are characteristic of α-helix and β-sheets, respectively. This experiment is briefly illustrated in Figure 1(a). The magnetization initiates on any aliphatic or aromatic proton, is transferred to the attached carbon where chemical shift is recorded and back to the proton, then is transferred *via* the NOE to an H^{α} proton of residue j and finally is transferred to the amide ^{15}N and NH of the next residue (residue $j+1$) whose chemical shifts are also recorded. Note that the recording of the shifts for the amide ^{15}N and NH of the next residue is preferred to recording those of the H^{α} and C^{α} in order to maximize resolution based on the much greater dispersion of the amide ^{15}N resonances than C^{α} resonances.

Figure 1(b) illustrates a strip plot from the C_i-NOESY-$N_{j+1}H_{j+1}$ experiment (200 ms mixing time), at the NH (F_3) and ^{15}N (F_2) chemical shifts of Phe61, showing NOEs to the Ala60 H^{α} of Δ131Δ. The auto peak labeled 'd' in the Figure corresponds to magnetization which resides on Ala60 H^{α} before and after the NOE mixing period. All other peaks correspond to NOE transfer of magnetization originating from H^{β} protons of Ala60 or aliphatic protons of other residues. The $d_{\alpha\beta}(i,i+3)$ NOE between the H^{β} of Lys63 and the H^{α} of Ala60 is of interest since it is diagnostic of at least the transient presence of a helical turn. Note that this region of the molecule is helical in wild-type SNase. Particularly worth noting are the sequential aliphatic-aliphatic NOEs between Ser59-Ala60 and Ala60-Phe61 (see below for detailed discussion).

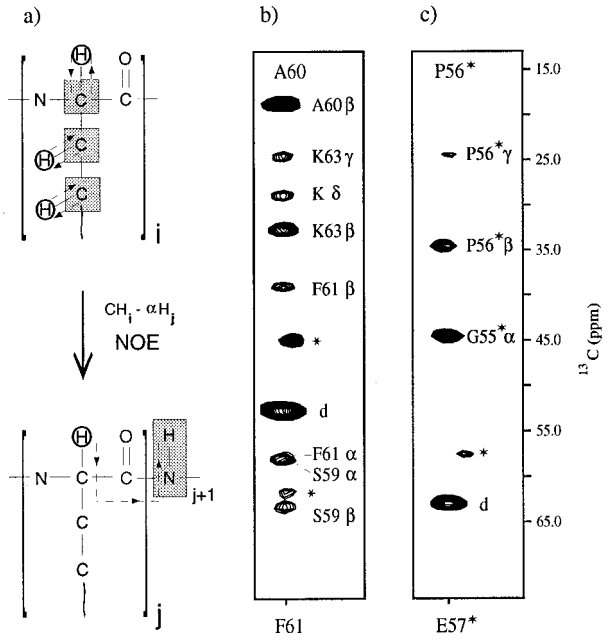

Figure 1. (a) Schematic diagram showing the magnetization transfer pathway for the 3-D C_i-NOESY-$N_{j+1}H_{j+1}$ experiment. Protons in residues *i* and *j* that are involved in the NOE transfer are circled. Nuclei whose chemical shifts are recorded are shown in shaded rectangles. (b) Strip plot from the 3-D C_i-NOESY-$N_{j+1}H_{j+1}$ spectrum at the ^{15}N and NH chemical shifts of Phe61 of Δ131Δ. The ''auto'' peak of Ala60 ($^{13}C^{\alpha}_i$, $^{15}N_{i+1}$) is labeled as 'd'. NOEs from other residues to Ala60 H^{α} are labeled with the residue and side-chain positions. The peaks labeled with an asterisk are from unrelated residues having nearby ^{15}N and NH chemical shifts while the peak labeled with Kδ has not been assigned to a specific residue. (c) Strip plot from the 3-D C_i-NOESY-$N_{j+1}H_{j+1}$ spectrum at the ^{15}N and NH chemical shifts of Glu57* (minor form) of Δ131Δ. The strong sequential NOE from the H^{α} of Gly55* to the H^{α} of Pro56* is labeled as G55*α. Detailed procedures for expression and purification of Δ131Δ, wild-type SNase and the drkN SH3 domain were described in previous publications (Alexandrescu *et al.*, 1994; Zhang *et al.*, 1997). Parameters for all NMR experiments and spectral processing are provided elsewhere (Zhang *et al.*, 1997).

In addition, the C_i-NOESY-$N_{j+1}H_{j+1}$ experiment provides a simple method for distinguishing between *cis* and *trans* peptide bond configurations, provided that the two states are in slow exchange on the NMR timescale. A number of different strategies currently exist for the identification of *cis* or *trans* Xaa-Pro peptide bonds (Chazin *et al.*, 1989; Hinck *et al.*, 1993). The C_i-NOESY-$N_{j+1}H_{j+1}$ experiment is an alternative approach for the identification of H^{α}-H^{α} NOEs diagnostic of *cis* peptide bonds. In spectra of Δ131Δ two sets of peaks exist for Gly55, Pro56, Glu57, Ala58 and Ser59 due to *cis/trans* isomerization of the Gly55-Pro56 peptide bond. Figure 1(c) shows the strip plot at the ^1H (F_3) and ^{15}N (F_2) chemical shifts of Glu57* (minor form), illustrating NOEs (200 ms mixing time) to the α proton of Pro56* of Δ131Δ. A strong NOE is observed connecting the α protons of the Gly55* and Pro56*. It is clear from the data that the minor resonances for Gly55* to Ser59* arise from the *cis* configuration of the Gly55-Pro56 peptide bond. Indeed the $^{13}C^{\gamma}$ chemical shift of the minor form of Pro56* (24.8 ppm) is 2.2 ppm upfield from the shift of the major form (27.0 ppm), which is consistent with the observation that there is generally a 3 ppm upfield shift of the $^{13}C^{\gamma}$ resonance of Pro residues in the *cis* configuration relative to *trans* (Torchia, 1984).

It is interesting to note that in wild-type folded SNase (free form), the Lys116-Pro117 *cis* peptide bond is the major form and *cis/trans* isomerization at this position has been shown to be the slow step during folding (Kuwajima *et al.*, 1991). However, in spectra of Δ131Δ, there is only one set of resonances for Lys116, reflecting either a more rapid isomerization of the Lys116-Pro117 bond than is the case for the Gly55-Pro56 peptide bond or else a dominant *trans* conformer. The $^{13}C^{\gamma}$ chemical shift for Pro117 in Δ131Δ supports the presence of a majority of *trans* conformer, consistent with previous kinetic arguments that the specific stabilization of the Lys116-Pro117 *cis*-peptide bond in SNase occurs only at a late stage in the folding process (Evans *et al.*, 1989).

In a previous study of Δ131Δ, a summary of NOE data obtained from conventional NMR experiments has been presented (Alexandrescu *et al.*, 1994). However, due to severe resonance overlap coupled with the lack of assignments for a number of residues, only a limited number of NOEs could be unambiguously established. Since that publication, complete assignments have been obtained, including the establishment of minor and major forms associated with the Gly55-Pro56 peptide bond isomerization (Supplementary Material). With these assignments and the recording of our newly developed NOE-based experiments, a more complete and less ambiguous list of observed NOE connectivities in this protein has been derived and is presented in Figure 2. Note that cross-peaks for Ala1-Thr41 (except Ser3) of Δ131Δ are not observed in HSQC spectra due to conformational exchange occurring on a μs to ms timescale. In the case of the drkN SH3 domain, residues Gln23 to Leu28 in the unfolded state are broadened beyond detection in ^1H-^{15}N HSQC spectra at 5°C as well. It is likely that the broadening which hinders observation of NOEs for these specific regions of Δ131Δ and the drkN SH3 domain arises from interactions stabilizing conformational preferences on longer time scales than in other regions of these proteins. Thus, characterization of these particularly interesting regions must be accomplished using structural information from sources other than NOE-based experiments. NOEs can, however, be observed for 101 out of 131 residues in Δ131Δ and for 53 of 59 residues in the drkN SH3 domain. A detailed list of NOEs observed in the drkN SH3 domain as well as a discussion of residual structure in the

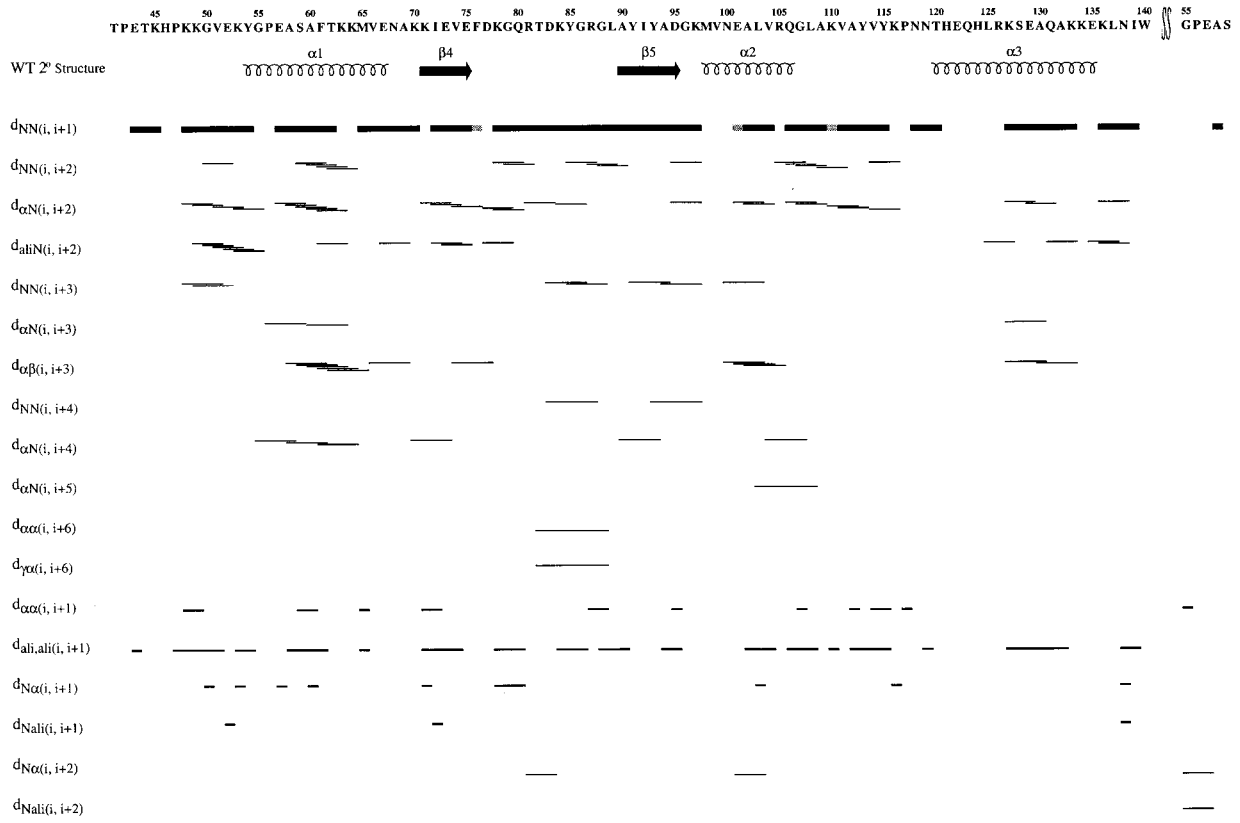

Figure 2. Summary of NOE connectivities observed for Δ131Δ using recently developed pulse schemes. The secondary structure in wild-type staphylococcal nuclease corresponding to the regions which are not line-broadened in Δ131Δ is indicated. Intensities of NOE peaks are not distinguished. The symbol $d_{NN(i,i+1)}$ denotes an NOE between sequential NH protons, $d_{\alpha\alpha(i,i+1)}$, an NOE between sequential H^{α} protons, $d_{aliN(i,i+2)}$ indicates an NOE between a non-specified aliphatic (but not H^{α}) proton of residue i and the NH of residue $i+2$, $d_{ali,ali(i,i+1)}$ indicates an NOE between either two non-specified aliphatic (but not H^{α}) protons or one H^{α} and one non-specified aliphatic (but not H^{α}) proton, and $d_{N\alpha(i,i+1)}$ indicates an NOE between the NH of residue i and the H^{α} of residue $i+1$. Sequential $d_{\alpha N(i+i+1)}$ NOEs are not plotted because they are observed for all the non-broadened residues. The designation for other NOEs, d_{ij}, follows directly from the description above. Shaded bars denote uncertainty in assignment of $d_{NN(i,i+1)}$ NOEs due to the absence of the symmetry-related peaks. NOEs for the final five residues, GPEAS, are for the residues of the minor form arising from the *cis* configuration of the Gly55-Pro56 peptide bond.

unfolded state and implications for the initiation of folding for this molecule is published elsewhere (Zhang & Forman-Kay, 1997).

NOE patterns observed

The resolving power of these new NMR experiments has allowed us to obtain more complete and unambiguous NOE assignments for both protein systems (Δ131Δ and the drkN SH3 domain) than previously possible. Therefore, an analysis of the general NOE connectivity patterns observed for these two disordered protein states should be instructive for illustrating general conformational features existing in protein disordered states. NOE connectivity patterns observed in small linear peptides used as models for studying the earliest events in folding have been extensively reviewed (Dyson & Wright, 1991; Wright *et al.*, 1988). NOE connectivities seen in several unfolded or partially folded proteins have been reported as well (for review, see Shortle, 1996). However, due to resonance overlap and ambiguous assignments, a comprehensive summary of experimentally observed NOE patterns has not previously been possible.

A disordered protein can be described as an ensemble of interconverting conformations. Preferential sampling of particular areas of conformational space for certain regions of the protein is expected based on the chemical and steric properties of amino acids along the sequence. Variations in intensity of observed NOEs are the result of both a population weighting of the interproton distances in the different configurations sampled and differences in dynamics of the internuclear vectors connecting the proton spins giving rise to the NOE. A quantitative analysis of NOE intensities in terms of distances requires, therefore, a consideration of the effects of internal dynamics. In addition, for experiments involving many steps of magnetization transfers (including TOCSY

periods), the NOE intensity is also related to the transfer efficiency for different side-chain positions. Qualitatively, the presence of an NOE connecting two protons indicates that they must be proximal at least some fraction of the time. In contrast, the absence of an NOE between spins does not rule out the possibility that the spins are close in space. The difference between whether an NOE is observed or not may well be as much a function of motional properties as internuclear distances. In this context it is interesting to note that in a recent study of the dynamics of ^{15}N-NH bond vectors in the unfolded state of the drkN SH3 domain (Farrow *et al.*, 1995), a relatively small range for the values of the spectral density function at zero frequency, $J(0)$, was observed as a function of amino acid sequence (less than a factor of 4). These $J(0)$ values, measures of the motional properties of the backbone ^{15}N-NH bond vectors, may be correlated with motions of backbone ^{1}H-^{1}H internuclear vectors that directly modulate intensities of observed NOEs. If this is the case, the small variation in motional properties would imply that the effective r^{-6} averaged distances between backbone protons should in general dominate the intensity differences of NOEs and whether or not they are observed. NOEs involving side-chain protons, however, could be affected much more significantly by dynamic properties. In the present discussion, only a very qualitative interpretation of the observed NOEs is considered.

Sequential $d_{\alpha N}(i,i+1)$ and intra-residue $d_{\alpha N}(i,i)$ distances are less than 3.6 Å and 2.9 Å, respectively, in all possible conformations of the polypeptide chain (Wüthrich, 1986) and thus are generally strong enough to be detected. The $d_{\alpha N}(i,i+1)$ NOEs are stronger than the intra-residue $d_{\alpha N}(i,i)$ NOEs for almost all the residues in the unfolded state of the drkN SH3 domain and Δ131Δ, as is expected for residues predominantly sampling the extended β-region of ϕ, ψ space. The sequential $d_{NN}(i,i+1)$ distance can vary between 2.0 and 4.8 Å depending on the local (ϕ_i, ψ_i) values (Wüthrich, 1986). The relative intensities of the $d_{NN}(i,i+1)$ and $d_{\alpha N}(i,i+1)$ sequential NOEs can provide a measure of the relative population of backbone dihedral angles in the α and β-regions of ϕ, ψ space (Dyson & Wright, 1991; Wright *et al.*, 1988). We observed $d_{NN}(i,i+1)$ NOEs for all non-exchange broadened residues in the unfolded state of the drkN SH3 domain and almost all non-exchange broadened residues in Δ131Δ. The general observation of both $d_{\alpha N}(i,i+1)$ and $d_{NN}(i,i+1)$ NOE connectivities is an indication of conformational sampling of both the α_R and β regions of ϕ, ψ space.

Medium range $(i,i+2)$ and $(i,i+3)$ NOEs indicate at least transient populations of turn-like or helical structures. For Δ131Δ, 27 $d_{\alpha N}(i,i+2)$, three $d_{\alpha N}(i,i+3)$, 16 $d_{NN}(i,i+2)$ and seven $d_{NN}(i,i+3)$ NOEs were observed. In their previous work, Shortle and co-workers observed six $d_{\alpha\beta}(i,i+3)$ NOEs involving (H^{α}-H^{β}) Ala58-Phe61, Ala60-Lys63, Thr62-Met65, Asn100-Leu103, Glu101-Val104 and Ala102-Arg105 using standard 3-D ^{13}C-edited NOE-based techniques (Alexandrescu *et al.*, 1994). Here we observe all of these NOEs including an additional six $d_{\alpha\beta}(i,i+3)$ NOEs in the C_i-NOESY-$N_{j+1}H_{j+1}$ data set (see Figure 2). In both experiments mixing times of 200 ms were employed. We have also recorded a 75 ms NOE data set and observe $d_{\alpha\beta}(i,i+3)$ NOEs (H^{α}-H^{β}) between Ala58-Phe61, Ala60-Lys63, Phe61-Lys64, Thr62-Met65, Val66-Ala69, Asn100-Leu103, Ala102-Arg105 and Ala130-Lys133. These $d_{\alpha\beta}(i,i+3)$ NOEs are characteristic of sampling of helical turn conformations (Wüthrich, 1986). The majority of residues displaying these NOEs reside in portions of the molecule corresponding to helical structure in wild-type SNase. The presence of these NOE peaks provides indirect evidence that native-like helical conformations may occur in folding intermediates.

It should be emphasized that helical conformations in these regions are transient and only marginally stable. Based on deviations of the $^{1}H^{\alpha}$ and $^{13}C^{\alpha}$ chemical shifts from random coil values, we have made rough estimates of the fraction of time helical conformations are sampled for the residues involved in $d_{\alpha\beta}(i,i+3)$ NOEs: 9% ($^{1}H^{\alpha}$)/14% ($^{13}C^{\alpha}$) for Ala58 to Met65, 18% ($^{1}H^{\alpha}$)/27% ($^{13}C^{\alpha}$) for Asn100 to Arg105 and 5% ($^{1}H^{\alpha}$)/10% ($^{13}C^{\alpha}$) for Lys127 to Lys133. We assume that 100% helix would give rise to an upfield shift of 0.38 ppm for $^{1}H^{\alpha}$ and a downfield shift of 2.6 ppm for $^{13}C^{\alpha}$, based on the average secondary shift values for these nuclei in helical regions of proteins relative to their random coil values (Wishart & Sykes, 1994). The values reported for the fraction of time a particular region spends in a helical conformation are based on chemical shifts averaged over all residues in the regions considered. However, the percentage of time an entire region is likely to be in a helical conformation depends on the cooperativity of helix formation. The values reported above assume cooperative formation of helical conformations and are likely upper limits. In the absence of cooperativity, the fraction of time that a given region spends in a helix conformation depends on the product of the helix probability at each site and is very small. The deviations of $^{1}H^{\alpha}$ and $^{13}C^{\alpha}$ chemical shifts from random coil for residues in Δ131Δ which correspond to helical regions in the wild-type SNase but which do not show $d_{\alpha\beta}(i,i+3)$ NOEs are smaller than for regions which are helical in the wild-type protein and do show $d_{\alpha\beta}(i,i+3)$ NOEs. Thus, chemical shifts can also be used for the quantification of secondary structure in flexible or disordered proteins and are an important complement to NOE information such as that presented in Figure 2.

Longer-range NOEs, including six $d_{\alpha\beta}(i,i+4)$, one $d_{\alpha N}(i,i+5)$, one $d_{\alpha\alpha}(i,i+6)$, one $d_{\gamma\alpha}(i,i+6)$, and two $d_{NN}(i,i+4)$ NOEs in Δ131Δ were also detected. Many of the $(i,i+4)$ NOEs could reflect native-like helical or loop conformations but two

$d_{\alpha N}(i,i+4)$ NOEs (Lys70-Val74 and Ala90-Ala94) represent interproton distances which are not consistent with the structure of the wild-type SNase. The observation of the native-like $d_{\alpha N}(i,i+5)$ NOE connecting Leu103 and Leu108 (Zhang *et al.*, 1997) is of particular interest in that it provides evidence of a Schellman capping motif (Schellman, 1980). A statistical analysis of the protein database indicates that the presence of Gly at position *i*, hydrophobic residues at positions $i+1$ and $i-4$ which interact, and a polar or Ala residue at $i-2$ is characteristic of this motif, which identifies helix termination by a Gly residue (Aurora *et al.*, 1994; Viguera & Serrano, 1995). The sequence of Δ131Δ in this region, Leu103 ($i-4$), Arg105 ($i-2$), Gly107 (*i*) and Leu108 ($i+1$), fits the consensus Schellman motif. While studies of peptide fragments have failed to provide evidence for Schellman motifs which are present in intact, folded proteins (Wang & Shortle, 1997; Viguera & Serrano, 1995), this NOE suggests that interactions stabilizing helix termination may exist in folding intermediates of SNase. The $(i,i+6)$ NOEs between Thr82 and Gly88 are also reflective of a native-like loop conformation.

In the case of the unfolded state of the drkN SH3 domain, 19 $d_{\alpha N}(i,i+2)$ and two $d_{\alpha N}(i,i+3)$ NOE connectivities were noted. In addition, we observed 24 $d_{NN}(i,i+2)$, three $d_{NN}(i,i+3)$ and one $d_{NN}(i,i+4)$ NOEs. The significant number of these intermediate-range NOEs, most of which are non-native like, suggests that local interactions leading to preferential sampling of the α_R region of ϕ, ψ space could bring residues that are distant in primary sequence into proximity (Zhang & Forman-Kay, 1997). It is interesting to note the absence of $d_{\alpha\beta}(i,i+3)$ NOEs for the unfolded state of the drkN SH3 domain. Helical structure is not found in the folded state of this molecular either.

We note that the longest-range NOEs assigned in Δ131Δ and in the drkN SH3 domain are $d_{\alpha\alpha}(i,i+6)/d_{\gamma\alpha}(i,i+6)$ and $d_{NN}(i,i+4)$ NOEs, respectively. It is also interesting that the longest-range $(i,i+5$ and $i,i+6)$ NOEs in Δ131Δ reflect native-like interactions, while the observed shorter-range NOEs are diagnostic of turn-like conformations which are often non-native-like. Very few unassigned NOEs are present which could represent longer-range interactions. This is somewhat disappointing, since models of Δ131Δ in particular predict long-range side-chain–side-chain contacts in order to stabilize a relatively compact state (Alexandrescu *et al.*, 1994; Wang & Shortle, 1995) and recent experiments demonstrate that a significant population of molecules have a native-like topology (Gillespie & Shortle, 1997). There are a number of possible causes for our inability to detect these NOEs. NOEs between two residues of the same amino acid type are difficult to distinguish from ''auto'' peaks in the C_i-NOESY-$N_{j+1}H_{j+1}$ experiment (Zhang *et al.*, 1997), due to degeneracy of aliphatic carbon chemical shifts. The TOCSY transfer efficiency (TOCSY transfer is present in some experiments) between the $^{13}C^{\alpha}$ carbon and other carbon atoms along the side-chain often decreases as a function of position from the backbone; sites further from the backbone are most likely to be involved in long-range interactions. The observation of NOEs from transient helical structure populated less than approximately 10% of the time in Δ131Δ provides evidence that interactions involved in relatively low populations can be detected, but it is possible that specific longer-range interactions are populated to a smaller extent. Finally the dynamics governing the interactions between side-chain protons may lead to a significant attenuation of these NOEs. We are currently investigating the relaxation behavior of methyl groups in Δ131Δ in order to address this issue.

Sequential aliphatic-aliphatic and amide-aliphatic NOEs

While long-range NOEs were absent, 16 $d_{\alpha\alpha}(i,i+1)$, 11 $d_{N\alpha}(i,i+1)$, three $d_{N\alpha}(i,i+2)$ and numerous sequential side-chain aliphatic-aliphatic NOEs were observed in the spectra of Δ131Δ. The presence of these NOEs was somewhat surprising, as was their presence to a lesser extent in spectra of the unfolded state of the drkN SH3 domain. The $d_{\alpha\alpha}(i,i+1)$ and $d_{N\alpha}(i,i+1)$ distances are always larger than 4.0 Å in regular secondary structure (including turns) and extended structure where residues are separated by *trans* peptide bonds (Wüthrich, 1986). Therefore, they are seldom observed in structural studies of folded proteins. Considering the sensitivity of the experiments utilized for identification of these NOEs and the sizeable intensities of some of these NOE cross-peaks, it initially seemed unlikely that the NOEs could have arisen from protons more than 4.0 Å apart. Thus, we have also considered the likelihood that these NOEs derive from rapid exchange of *cis* and *trans* peptide bond conformers for both prolyl and non-prolyl peptide bonds. Both of these possibilities, that the NOEs are from distances greater than 4 Å in *trans* conformers or that they are from rapidly exchanging *cis* conformers where the distances are less than 4 Å, are addressed in detail below.

However, before we can attempt to interpret these NOEs, even qualitatively, it is necessary to establish that their appearance is not the result of spin diffusion. In principle, spin diffusion can contribute significantly to the intensity of sequential NOEs. For example, magnetization connecting sequential H^{α} protons may be due to a transfer from H_i^{α} to NH_{i+1} followed by an additional intra-residue NOE from NH_{i+1} to H_{i+1}^{α}. Note that sequential H^{α}-NH distances can be as short as 2.2 Å for residues in an extended conformation, while intra-residue H^{α}-NH distances are ~2.5 Å for amino acids in the α region of ϕ, ψ space (Wüthrich, 1986). Evidence from chemical shifts and other NMR data strongly suggests that this

disordered state rapidly samples both backbone conformations. In addition to recording a 200 ms C_i-NOESY-$N_{j+1}H_{j+1}$ spectrum where 16 symmetry-related pairs of $d_{\alpha\alpha}(i,i+1)$ NOEs were observed, a 75 ms spectrum was also recorded in which six symmetry related pairs of H^α-H^α cross-peaks were observed involving residues Lys49-Gly50, Ser59-Ala60, Ile72-Glu73, Arg87-Gly88, Gly88-Leu89 and Val114-Tyr115. Note that a strong NOE cross-peak between Gly55* and Pro56* (minor form) was also observed. The symmetry-related cross-peak corresponding to magnetization transfer from Pro to Gly cannot be observed in this experiment since Pro does not contain an amide proton.

In an effort to minimize contributions from spin diffusion associated with the NH proton, an NOE experiment was recorded in which a scheme for the selective inversion of NH magnetization was inserted in the middle of the mixing time (150 ms) in the manner suggested by Vincent *et al.* (1996) and Zolnai *et al.* (1995). In this case, four $d_{\alpha\alpha}(i,i+1)$ NOEs were observed corresponding to magnetization transfer from Gly55*-Pro56* (minor form), Ser59-Ala60, Met65-Val66 and Ile72-Glu73. Note that the procedure for NH selective inversion places magnetization in the transverse plane for a time of $1/J_{NH}$, where J_{NH} is the one-bond NH-^{15}N coupling constant. This results in considerable relaxation losses, especially for ^{13}C labeled samples.

We have performed further experiments in which H^α resonances are selectively inverted by band-selective shaped pulses to eliminate spin diffusion pathways through resonances other than H^α. The inversion pulses chosen, 3.88 ms I-burp-2 pulses (Geen & Freeman, 1983), invert magnetization (>98%) over a bandwidth of ±1.0 ppm centered at the carrier, while leaving magnetization resonating outside a window extending ±1.7 ppm from the carrier unaffected (500 MHz ^{1}H frequency). Therefore, NOEs involving protons that resonate upfield of 3 ppm or downfield of 6.5 ppm are suppressed by this approach, eliminating potential spin diffusion pathways contributing to the observation of sequential H^α-H^α NOEs. The efficacy of the scheme is illustrated in Figure 3 which shows a comparison of 2-D spectra without (a) and with (b) two selective inversion pulses during the 200 ms NOE mixing periods. In Figure 3(a), cross-peaks at ^{13}C frequencies less than ≈40 ppm correspond to NOEs linking non-H^α protons with H^α protons. In Figure 3(b) these NOEs have been eliminated; note that cross-peaks in the $^{13}C^\alpha$ region of the carbon spectrum (F_1) are still observed corresponding primarily to magnetization which resides on a particular H^α spin for the duration of the mixing time (diagonal peak) or, much more rarely, magnetization transferred between neighboring H^α protons during the mixing time. A number of cross-peaks arising from H^β protons of Ser and Thr residues are also observed, because of the downfield shifts of the β-protons of these residues. In the resulting 3-D spectra recorded using this procedure, ten $d_{\alpha\alpha}(i,i+1)$ NOEs are observed although surprisingly none of them show symmetry-related peaks. Based on these results, we feel that it is likely that at least some of the observed H^α-H^α NOEs reflect actual short distances rather that artifacts of spin diffusion.

It is important to examine possible spin-diffusion pathways for $d_{N\alpha}(i,i+2)$ and $d_{N\alpha}(i,i+1)$ NOEs as well. The most likely indirect pathway for $d_{N\alpha}(i,i+2)$ NOEs is an intra-residue $d_{N\alpha}(i+2,i+2)$ followed by a $d_{NN}(i,i+2)$ NOE. Similarly, $d_{N\alpha}(i,i+1)$ NOEs may be observed due to intra-residue $d_{N\alpha}(i+1,i+1)$ transfer followed by a $d_{NN}(i,i+1)$ NOE. However, in the case of the two $d_{N\alpha}(i,i+2)$ NOEs (Arg81-Asp83 and Glu101-Leu103) observed from non-proline resi-

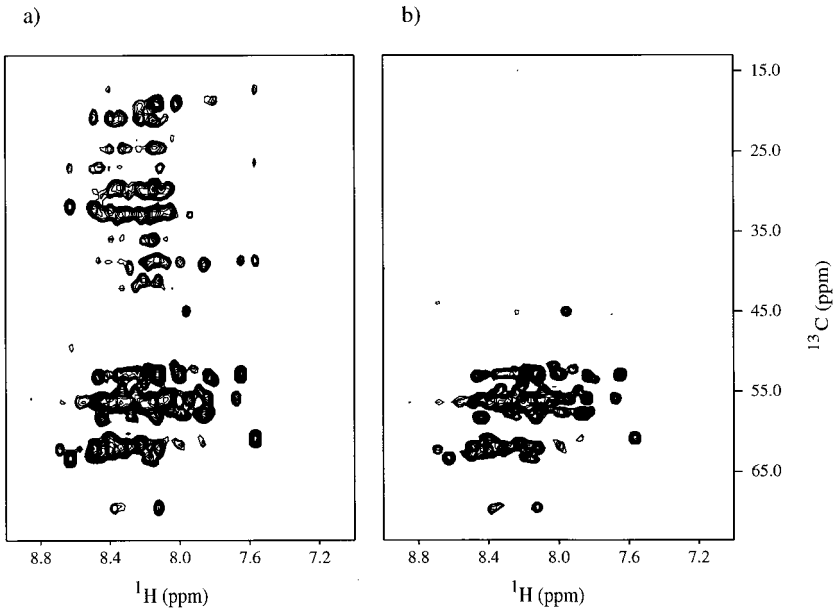

Figure 3. ^{1}H-^{13}C 2-D spectra representing the first ^{15}N increments of the 3-D C_i-NOESY-$N_{j+1}H_{j+1}$ experiments (200 ms mixing time, 500 MHz ^{1}H frequency) without (a) and with (b) selective H^α inversion pulses using am I-burp-2 profile (Green & Freeman, 1983) during the NOE mixing period, plotted at the same contour level. In the case of (b), two 3.88 ms I-burp-2 pulses are employed, with the first applied 50 ms after the start of the mixing period and the second pulse applied 100 ms later. The pulses are centered at 4.71 ppm. All NOEs other than H^α-H^α NOEs, with the exception of NOEs from β protons of Ser and Thr residues due to their down-field chemical shifts, are suppressed by these selective pulses.

dues, $d_{NN}(i,i+2)$ NOEs are not present. The $d_{N\alpha}(i,i+1)$ NOE from Lys71-Ile72 also does not have a corresponding $d_{NN}(i,i+1)$ NOE. Unfortunately, it is difficult to design an experiment to rule out spin diffusion for these NOEs as was done for the $d_{\alpha\alpha}(i,i+1)$ NOEs, since both NH and H^{α} spins would be involved in an actual NOE. However, the absence of the $d_{NN}(i,i+1)$ and $d_{NN}(i,i+2)$ NOEs for the cases described here and the results from the $d_{\alpha\alpha}(i,i+1)$ NOEs strongly argue against spin-diffusion being the sole cause for these $d_{N\alpha}(i,i+2)$ and $d_{N\alpha}(i,i+1)$ NOEs. In this context it is noteworthy that Wang and Shortle have observed at least five $d_{N\alpha}(i,i+1)$ NOEs in the center of a SNase peptide comprising residues 92 to 110 (Wang & Shortle, 1997).

Confident that some if not most of these cross-peaks reflect true NOEs, we have tried to determine whether they could derive from *trans* peptide bond conformers by searching for such NOEs in data from folded proteins where *trans* conformers are populated almost exclusively. Sequential H^{α} NOEs linking non-proline residues were not observed in a 75 ms NOE spectrum of folded SNase in complex with pdTp and Ca^{2+} (1.5 mM). However, using a 200 ms NOE mixing time, two symmetry-related sequential H^{α}-H^{α} NOEs (Glu67-Asn68 and Ala94-Asp95) and six other H^{α}-H^{α} NOEs without symmetry-related peaks were observed. These results suggest the possibility of spin-diffusion during the 200 ms mixing time for wild-type protein. In order to address the question of whether NOEs arise from conformations present in folded proteins but not usually observable due to the lower protein concentrations used (1-2 mM), we prepared a highly concentrated sample of wild-type SNase (free form, 4.2 mM) comparable in concentration to our sample of Δ131Δ (4.5 mM). We then recorded NOE-based experiments using selective inversion of H^{α} resonances by band-selective shaped pulses during a 200 ms mixing time as discussed above. In this case, two weak sequential H^{α}-H^{α} NOEs were observed (Glu67-Asn68 and Tyr85-Gly86) with no symmetry-related peaks. The sequential H^{α}-H^{α} distances for all peptide bonds (except the single *cis* peptide bond between Lys116-Pro117) in the wild-type SNase (pdb code 1stn; Hynes & Fox, 1991) are within the range of 4.3 to 4.8 Å. The shortest $d_{N\alpha}(i,i+1)$ distance in *trans* peptide bonds in wild-type SNase is 4.2 Å, with many in the range of 4.4 to 4.8 Å while $d_{N\alpha}(i,i+2)$ distances range from 5.0 to 8.8 Å with most between 6 and 7.5 Å. Yet, in previous NMR studies of this protein, no $d_{\alpha\alpha}(i,i+1)$ or $d_{N\alpha}(i,i+1)$ NOEs were reported (Torchia *et al.*, 1989; Wang *et al.*, 1990). In addition, we recorded a N_{i+1}-NOESY-N_jH_j experiment (200 ms mixing time) for identifying NOEs between H^{α} and NH using the 4.2 mM wild-type SNase sample and only one such $d_{N\alpha}(i,i+1)$ NOE (Lys116-Pro117) was observed. Recall again that the peptide bond between Lys116-Pro117 in SNase is *cis* and the distance connecting NH(i) and $H^{\alpha}(i+1)$ in this case is 4.80 Å.

A distance between two protons of less than 5 Å does not necessarily give rise to an NOE interaction. Motional properties of the internuclear vector connecting the two protons can significantly affect the intensity or ability to observe a weak NOE. In this regard, it is interesting to note that the two sequential H^{α}-H^{α} NOEs observed in wild-type SNase do not have the shortest distances (4.56 Å for Glu67-Asn68 and 4.37 or 4.96 Å for Tyr85-Gly86) and are from residues in turn regions where different dynamic behavior may be expected. While the high concentration of the sample and the better resolution of these experiments also contribute to the identification of these sequential NOEs, it seems likely that motional properties play a significant role in determining which of the NOEs are observed. It may also be possible that these turns in wild-type SNase allow *cis/trans* peptide bond isomerization. Both prolyl and non-prolyl peptide bond isomerizations have been shown to occur in folded proteins (Evans *et al.*, 1987; Chazin *et al.*, 1991; Bystroff & Kraut, 1991). The cases described to date by NMR involve distinct resonances for the *cis* and *trans* peptide bonds. It is conceivable, however, that isomerization could occur in the fast exchange limit. This issue is discussed below.

Another example of the lack of observation of sequential H^{α}-H^{α} NOEs is in the NMR structural study of the 56 residue immunoglobulin binding domain of streptococcal protein G (Gronenborn *et al.*, 1991). The observed NOE data and calculated structures are included with the X-PLOR package (Brünger, 1992). All peptide bonds are in the *trans* conformation and the distances between sequential H^{α} protons in this protein are within 4.3 to 5.0 Å. However, no NOEs between any sequential H^{α} protons were observed even though a few weak NOEs were identified for other types of proton pairs having distances of close to 5 Å. In addition, no $d_{N\alpha}(i,i+1)$ NOEs were reported in the NOE table and yet distances are in the range of 4.11 to 5.71 Å with many of them less that 5 Å.

Contrary to these examples where sequential H^{α}-H^{α} NOEs have not been observed, we have identified 19 $d_{\alpha\alpha}(i,i+1)$ NOEs in our recent studies of the fully folded (salt-stabilized) drkN SH3 domain based on a simultaneous C,N-NOESY dataset recorded using a 150 ms mixing time (unpublished results). Some of these NOEs are found in regions of extended β-structure where conformational exchange is extremely unlikely and some in the turn regions. Thus, while these sequential $d_{\alpha\alpha}(i,i+1)$ and $d_{N\alpha}(i,i+1)$ NOEs have rarely been reported, they may be reflective of long (>4.2 Å) distances compatible with *trans* peptide bonds and may be observed in our datasets due to the increased resolution and sensitivity of the experiments.

However, at present we cannot definitively exclude the possibility that these sequential

$d_{\alpha\alpha}(i,i+1)$ and $d_{N\alpha}(i,i+1)$ NOEs could reflect protons in very close proximity in a fraction of substates within the disordered state ensemble sampling *cis* peptide bonds. *Cis* peptide bonds can be stabilized by interactions between neighboring side-chains which pack much more closely than for *trans* peptide conformations. It is of interest that in nearly all of these cases where H^{α}-H^{α} NOEs are observed in Δ131Δ, one if not both of the involved residues are hydrophobic and the stabilization from hydrophobic burial could be substantial, with solvation free energies of residues Val, Leu and Ile between 1.5 and 2 kcal/mol (Eisenberg & McLachlan, 1985).

The proposition that these $d_{\alpha\alpha}(i,i+1)$, $d_{\alpha ali}(i,i+1)$, $d_{N\alpha}(i,i+1)$ and $d_{N\alpha}(i,i+2)$ NOEs may, at least in some cases, arise due to sampling of the *cis* configuration of the peptide bond is intriguing. In order for these NOEs to be attributed to a small fraction of *cis*, two conditions must be met. The first is that the population of *cis* must be large enough to enable NOEs to be observed. Although they are found rarely in native proteins with the *trans* isomer of a non-prolyl peptide bond strongly favored over the *cis* isomer by a factor of 100 to 1000 (Jorgensen & Gao, 1988), *cis* peptide bonds which do not involve proline do exist and can serve important functional or structural roles. In surveying the Protein Data Bank, Stewart *et al.* (1990) have found that only 0.05% of a total of 31,005 non-prolyl peptide bonds are in the *cis* configuration and 6.5% of 1534 prolyl peptide bonds are in the *cis* form. These values are both considerably less than the distribution predicted on the basis of the potential energy difference between the *cis* and *trans* isomeric forms and experimental data on small peptides. The authors argued that this might well be due to the assumption of all *trans* peptide bonds in modelling the backbone in X-ray crystallographic structural analysis, since the occurrence of *cis* peptide bonds increases with the increasing resolution of the structures examined. Conformational preferences in folded proteins may decrease the actual occurrence of *cis* peptide bonds, but it might be expected that peptide bond conformers would be populated in disordered states at a level reflective of the energetic difference between the *cis* and *trans* states.

While *cis* peptide bonds for non-proline residues are energetically unfavored (1:100 to 1:1000; Jorgensen & Gao, 1988) and therefore must only exist in a small fraction of the population, α protons from two adjacent residues in a *cis* peptide bond can be very close in space (~2.0 to 4.0 Å; Wüthrich, 1986). Therefore, NOEs may be observed for even a very low population due to the r^{-6} distance dependence of the NOE. In addition, in *cis* peptide bonds the $d_{N\alpha}(i,i+1)$ distance can be very small, varying from about 2.0 to 5.5 Å. A clear and definitive demonstration of the population of molecules in the *cis* configuration is, however, very difficult.

The second condition that must be met, since only a single set of cross-peaks is observed for each of the residues displaying sequential $d_{\alpha\alpha}(i,i+1)$ and $d_{N\alpha}(i,i+1)$ NOEs, is that exchange between *cis* and *trans* peptide bond conformers should be rapid on the NMR time scale. Fast exchange on the NMR time scale leading to observation of a single set of resonances is dependent on the difference in resonance frequencies between the two states, as well as the rate of exchange. It is quite possible that because of the extra degrees of freedom available to a non-prolyl side-chain relative to proline, *cis* and *trans* states of non-prolyl peptide bonds may well have very similar shifts in unfolded states, leading to fast exchange behavior even with relatively slow transition rates.

In small peptides, the *trans* prolyl peptide bond is favored over the *cis* prolyl peptide bond by a ratio of approximately 4:1 (since the free energy difference favoring the *trans* form is very small) and NMR parameters for the two states are easily observed. Since the barrier for *trans*- > *cis* isomerization is high as determined from model peptides (~20 kcal/mol), the isomerization rate is slow and generally leads to slow exchange on the NMR time scale. In this regard, it has been shown experimentally that the *cis/trans* isomerization of prolyl peptide bonds can be the rate determining step in a protein folding reaction (Brandts *et al.*, 1975). However, as suggested by Grathwohl & Wüthrich (1981), the increased steric strain from chain elongation may increase the isomerization rate for prolyl peptide bonds in proteins from that seen in model peptide studies. This hypothesis was confirmed by observation of a faster *cis/trans* isomerization rate for the Gly42-Pro43 peptide bond in calbindin, D_{9k}, than in model peptides (Kordel *et al.*, 1990).

It has also been suggested that the *cis/trans* isomerization rate for non-prolyl peptide bonds may be intrinsically faster than for Xaa-Pro bonds (Brandts *et al.*, 1975). In a recent study (Odefey *et al.*, 1995), the *cis* - > *trans* isomerization for the Tyr38-Ala39 peptide bond in a Pro39-Ala mutant of ribonuclease T_1 was found to have a time constant of 730 ms, approximately 60-fold faster than for the isomerization of the Tyr38-Pro39 bond in the wild-type protein. While the *trans* - > *cis* rate measured in this study was significantly slower, it was coupled to refolding and thus may be dominated by formation of protein structure. This work supports the speculation that the isomerization of non-prolyl peptide bonds could be involved in the folding kinetics of proteins (Jennings *et al.*, 1993). Finally, we note that rates measured in model peptides or highly denaturing conditions may not be appropriate in the case of Δ131Δ, particularly since a recent set of experiments (Gillespie & Shortle, 1997) establishes that considerable native-like structure is present in this molecule.

Together, these arguments suggest that both the population and rate conditions for observation of the previously described sequential-sequential and

aliphatic-amide NOEs could be met for at least some of the non-prolyl peptide bonds in Δ131Δ. While one would not expect to observe all of these NOEs in the same conformation, it is often the case that in regions in Δ131Δ where at least one of the NOEs characteristics of *cis* peptide bond conformations is observed, a second or third is present as well (see bottom four rows of Figure 2). Examples include Ala60-Phe61, Ile72-Glu73, Leu103-Val104, and Gly55*-Pro56* peptide bonds, the last of which is known to be in a stable *cis* conformation. While the role of these potential *cis* conformations is unknown, the potential importance of turns in initiating folding and the finding that *cis* peptide bonds in native proteins are found primarily in bends and turns (Stewart *et al*., 1990) is suggestive. In this regard it is of interest to recall the high frequency of turn-like structures in Δ131Δ and in the unfolded states of the drkN SH3 domain (Zhang & Forman-Kay, 1997).

Comparison of NOE patterns with predictions

Models utilizing backbone ϕ, ψ dihedral angle distributions extracted from the Protein Data Bank to describe local conformational preferences in the polypeptide chain have been used to predict NOE patterns in denatured states of proteins (Fiebig *et al*., 1996). Although experimental data from lysozyme denatured in urea at low pH correlate reasonably well with predictions of sequential and medium-range NOEs, it is still uncertain whether the database of conformations accessible to folded proteins (Fiebig *et al*., 1996; Serrano, 1995) includes all the possible conformations available to unfolded or partially folded proteins. Nonetheless, it is of immediate interest to make a general comparison of the published predictions of lysozyme with our results on Δ131Δ and the drkN SH3 domain, that are based on experiments which enable a more complete analysis of sequential and medium-range NOEs than otherwise possible. We are currently comparing predictions of intensities for various classes of sequential and medium-range NOEs for Δ131Δ and the drkN SH3 domain with our experimentally observed NOEs.

Sequential $d_{\alpha N}(i,i+1)$ NOEs are observed for all non-broadened residues in Δ131Δ and the drkN SH3 domain as is predicted, regardless of the local sequence (Fiebig *et al*., 1996). The $d_{\alpha N}(i,i+2)$ NOEs are the most frequently predicted medium-range NOEs in the case of the denatured state of lysozyme. These NOEs are the most abundant in the case of Δ131Δ as well; however, for the drkN SH3 domain $d_{NN}(i,i+2)$ NOEs are more prevalent than $d_{\alpha N}(i,i+2)$ NOEs. The number of medium-range NOEs predicted by Fiebig *et al*. (1996) increases when a cooperativity factor is introduced into the random coil model for the unfolded chain in lysozyme. In the case of both the drkN SH3 domain and Δ131Δ, several clusters of residues account for the majority of the observed medium-range NOEs (see Figure 2). This suggests that the NOEs reflect local cooperativity of folding involving several residues having similar conformational preferences. Unfortunately, sequential and medium-range aliphatic-aliphatic NOEs were not discussed in the study (Fiebig *et al*., 1996), likely due to limited and ambiguous experimental data. The new NMR methods described previously (Zhang *et al*., 1997) should provide an experimental basis for further prediction studies including NOEs which have been difficult to identify. While prediction of NOEs based on models of the unfolded chain derived from distributions of conformational space in folded protein structures in the PDB promises to be useful for interpretation of NOEs in disordered states, our observation of NOEs which are rarely seen in folded proteins argues that a broader database of conformations may be required to explain experimental data from unfolded or partially folded proteins.

In summary, a more comprehensive NOE analysis for Δ131Δ has been described than previously obtained using new NMR experiments which offer higher resolution for disordered states of proteins. The resulting NOE patterns reflect sampling of both α and β regions of ϕ,φ conformational space, yet demonstrate significant preferences for both native-like and non-native-like turn and potentially helical conformations. Surprisingly, few long-range NOEs were observed. The longest-range NOEs involve residues separated by five and six amino acids. We have also demonstrated that at least some of the sequential aliphatic-aliphatic and amide-aliphatic NOEs are reflective of short distances and are not derived from spin diffusion. These NOEs might represent longer inter-proton distances (4 to 5 Å) which are observed due to the better resolution offered by newly developed NMR pulse sequences, the high concentration of protein in our samples and the increased mobility at these sites. Alternatively, in some cases, the NOEs may be due to shorter inter-proton distances (2 to 2.5 Å) in a small population of conformations having *cis* peptide bonds. It is likely that more reports of such NOEs will emerge in studies of other unfolded or partially folded proteins, as experimental methods for NOE detection become more sensitive.

Acknowledgements

This work was supported through grants from the Medical Research Council of Canada (J.D. F.-K. and L.E.K.), the National Sciences and Engineering Research Council of Canada (L.E.K.) and the National Institutes of Health, GM 34171 (D.S.).

References

Alexandrescu, A. T., Abeygunawardana, C. & Shortle, D. (1994). Structure and dynamics of a denatured 131-residue fragment of staphylococcal nuclease: a

heteronuclear NMR study. *Biochemistry,* **33**, 1063–1072.

Aurora, R., Srinivasan, R. & Rose, G. D. (1994). Rules for α-helix termination by glycine. *Science,* **264**, 1126–1130.

Brandts, J. F., Halvorson, H. R. & Brennan, M. (1975). Consideration of the possibility that the slow step in protein denaturation reactions is due to the *cis-trans* isomerism of proline residues. *Biochemistry,* **14**, 4953–4963.

Brünger, A. T. (1992). *X-PLOR: a System for X-ray Crystallography and NMR, version 3.1,* Yale University Press, New Haven, CT, USA.

Bystroff, C. & Kraut, J. (1991). Crystal structure of unliganded *Escherichia coli* dihydrofolate reductase: lingand-induced conformational changes and cooperativity in binding. *Biochemistry,* **30**, 2227–2239.

Chazin, W. J., Kordel, J., Drakenberg, T., Thulin, E., Brodin, P., Grundstrom, T. & Forsen, S. (1989). Proline isomerism leads to multiple folded conformations of calbindin D_{9k}: direct evidence from two-dimensional ^{1}H NMR spectroscopy. *Proc. Natl Acad. Sci. USA,* **86**, 2195–2198.

Dobson, C. M. (1992). Unfolded proteins, compact states and molten globules. *Curr. Opin. Struct. Biol.* **2**, 6–12.

Dyson, H. J. & Wright, P. E. (1991). Defining solution conformations of small linear peptides. *Annu. Rev. Biophys. Biophys. Chem.* **20**, 519–538.

Eisenberg, D. & McLachlan, A. D. (1985). Solvation energy in protein folding and binding. *Nature,* **319**, 199–203.

Evans, P. A., Dobson, C. M., Kautz, R. A., Hatfull, G. & Fox, R. O. (1987). Proline isomerism in staphylococcal nuclease characterized by NMR and site-directed mutagenesis. *Nature,* **329**, 266–268.

Evans, P. A., Kautz, R. A., Fox, R. O. & Dobson, C. M. (1989). A magnetization-transfer nuclear magnetic resonance study of the folding staphylococcal nuclease. *Biochemistry,* **28**, 362–370.

Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. (1995). Comparison of the backbone dynamics of a folded and an unfolded SH3 domain existing in equilibrium in aqueous buffer. *Biochemistry,* **34**, 868–878.

Fiebig, K. M., Schwalbe, H., Buck, M., Smith, L. J. & Dobson, C. M. (1996). Toward a description of the conformations of denatured states of proteins: comparison of a random coil model with NMR measurements. *J. Phys. Chem.* **100**, 2661–2666.

Geen, H. & Freeman, R. (1983). Band-selective radiofrequency pulses. *J. Magn. Reson.* **93**, 93–141.

Gillespie, J. & Shortle, D. (1997). Characterization of long range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *J. Mol. Biol.* **258**, 170–184.

Grathwohl, C. & Wüthrich, K. (1981). NMR studies of the rate of proline *cis-trans* isomerization in oligopeptides. *Biopolymers,* **20**, 2623–2633.

Gronenborn, A. M., Filpula, D. R., Essig, N. Z., Achari, A., Whitlow, M., Wingfield, P. T. & Clore, G. M. (1991). The immunoglobulin binding domain of streptococcal protein G has a novel and highly stable polypeptide fold. *Science,* **253**, 657–661.

Hinck, A. P., Eberhardt, E. S. & Markley, J. L. (1993). NMR strategy for determining Xaa-Pro peptide bond configuration in proteins: mutants of staphylococcal nuclease with altered configuration at proline-117. *Biochemistry,* **32**, 11810–11818.

Hynes, T. R. & Fox, R. O. (1991). The crystal structure of staphylococcal nuclease refined at 1.7 Å resolution. *Proteins: Struct. Funct. Genet.* **10**, 92–105.

Jennings, P. A., Finn, B. E., Jones, B. E. & Matthews, C. R. (1993). A reexamination of the folding mechanism of dihydrofolate reductase from *Escherichia coli*: vertification and refinement of a 4-channel model. *Biochemistry,* **32**, 3783–3789.

Jorgensen, W. L. & Gao, J. (1988). *Cis-trans* energy difference for the peptide bond in the gas phase and in aqueous solution. *J. Am. Chem. Soc.* **110**, 4212–4216.

Kordel, J., Forsen, S., Drakenberg, T. & Chazin, W. J. (1990). The rate and structural consequences of proline *cis-trans* isomerization in calbindin D_{9k}: NMR studies of the minor (*cis*-Pro43) isoform and the Pro43Gly mutant. *Biochemistry,* **29**, 4400–4409.

Kriwacki, R. W., Hengst, L., Tennant, L., Reed, S. I. & Wright, P. E. (1996). Structural studies of $p21^{Waf1/Cip1/Sdi1}$ in the free and Cdk2-bound state: conformational disorder mediates binding diversity. *Proc. Natl Acad. Sci. USA,* **93**, 11504–11509.

Kuwajima, K., Okayama, N., Yamamoto, K., Ishihara, T. & Sugai, S. (1991). The Pro117 to glycine mutation of staphylococcal nuclease simplifies the unfolding-folding kinetics. *FEBS Letters,* **290**, 135–138.

Odefey, C., Mayr, L. M. & Schmid, F. X. (1995). Non-prolyl *cis-trans* peptide isomerization as a rate-determining step in protein unfolding and refolding. *J. Mol. Biol.* **245**, 60–78.

Schellman, C. (1980). *Protein Folding* (Jaenicke, R., ed.), pp. 53–61, Elsevier/North-Holland, New York.

Serrano, L. (1995). Comparison between the φ distribution of the amino acids in the protein database and NMR data indicates that amino acids have various φ propensities in the random coil conformation. *J. Mol. Biol.* **254**, 322–333.

Shen, F., Triezenberg, S. J., Hensley, P., Porter, D. & Knutson, J. R. (1996). Transcriptional activation domain of the H.Virus protein becomes conformationally constrained upon interaction with basal transcription factors. *J. Biol. Chem.* **271**, 4827–4837.

Shortle, D. (1993). Denatured states of proteins and their roles in folding and stability. *Curr. Opin. Struct. Biol.* **3**, 66–74.

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

Shortle, D., Wang, Y., Gillespie, J. R. & Wrabl, J. O. (1996). Protein folding for realists: a timeless phenomenon. *Protein Sci.* **5**, 991–1000.

Stewart, D. E., Sarkar, A. & Wampler, J. E. (1990). Occurrence and roles of *cis* peptide bonds in protein structures. *J. Mol. Biol.* **214**, 253–260.

Torchia, D. A. (1984). Solid state NMR studies of protein internal dynamics. *Annu. Rev. Biophys. Bioeng.* **13**, 125–144.

Torchia, D. A., Sparks, S. W. & Bax, A. (1989). Staphylococcal nuclease: sequential assignments and solution structure. *Biochemistry,* **28**, 5509–5524.

Viguera, A. R. & Serrano, L. (1995). Experimental analysis of the Schellman motif. *J. Mol. Biol.* **251**, 150–160.

Vincent, S. J. F., Zwahlen, C., Bolton, P. H., Logan, T. M. & Bodenhausen, G. (1996). Measurement of

cross-relaxation between amide protons in ^{15}N-enriched proteins with suppression of spin diffusion. *J. Am. Chem. Soc.* **118**, 3531–3532.
Wang, J., LeMaster, D. M. & Markley, J. L. (1990). Two-dimensional NMR studies of staphylococcal nuclease. 1. Sequence-specific assignments of hydrogen-1 signals and solution structure of the nuclease H124L-thymidine 3′,5′-bisphosphate-Ca^{2+} ternary complex. *Biochemistry,* **29**, 88–101.
Wang, Y. & Shortle, D. (1995). The equilibrium folding pathway of staphylococcal nuclease: identification of the most stable chain-chain interactions by NMR and CD spectroscopy. *Biochemistry,* **34**, 15895–15905.
Wang, Y. & Shortle, D. (1997). Residual helical and turn structure in the denatured state of staphylococcal nuclease: analysis of peptide fragments. *Folding Design,* **2**, 93–100.
Weinreb, P. H., Zhen, W., Poon, A. W., Conway, K. A. & Lansbury, P. T., Jr (1996). NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry,* **35**, 13709–13715.
Wishart, D. S. & Sykes, B. D. (1994). Chemical shifts as a tool for structure determination. *Methods Enzymol.* **239**, 363–392.
Wright, P. E., Dyson, H. J. & Lerner, R. A. (1988). Conformation of peptide fragments of proteins in aqueous solution: implications for initiation of protein folding. *Biochemistry,* **27**, 7167–7175.
Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids,* John Wiley & Sons, New York.
Zhang, O. & Forman-Kay, J. D. (1995). Structural characterization of folded and unfolded states of an SH3 domain in equilibrium of aqueous buffer. *Biochemistry,* **34**, 6784–6794.
Zhang, O. & Forman-Kay, J. D. (1997). NMR studies of the unfolded states of an SH3 domain in aqueous solution and denaturing conditions. *Biochemistry,* **36**, 3959–3970.
Zhang, O., Kay, L. E., Oliver, J. P. & Forman-Kay, J. D. (1994). Backbone ^{1}H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk folded and unfolded states using enhanced-sensitivity pulsed field gradients NMR techniques. *J. Biomol. NMR,* **4**, 845–858.
Zhang, O., Forman-Kay, J. D., Shortle, D. & Kay, L. E. (1997). Triple-resonance NOESY-based experiments with improved spectral solution: applications to structural characterization of unfolded, partially folded and folded proteins. *J. Biomol. NMR,* **9**, 181–200.
Zolnai, Z., Juranic, N., Markley, J. L. & Macura, S. (1995). Magnetization exchange network editing: mathematical principles and experimental demonstration. *Chem. Phys.* **200**, 161–179.

Edited by P. E. Wright

(Received 28 March 1997; received in revised form 3 June 1997; accepted 20 June 1997)

http://www.hbuk.co.uk/jmb

Supplementary material comprising a Table of the ^{15}N, ^{1}HN and ^{13}C chemical shifts for the Δ131Δ mutant of staphylococcal nuclease, is available from JMB OnLine.