Structural and dynamic characterization of an unfolded state of poplar apo-plastocyanin formed under nondenaturing conditions

YAWEN BAI,¹ JOHN CHUNG, H. JANE DYSON, AND PETER E. WRIGHT

Department of Molecular Biology and Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037, USA

(RECEIVED January 2, 2001; ACCEPTED March 2, 2001)

Abstract

Plastocyanin is a predominantly β -sheet protein containing a type I copper center. The conformational ensemble of a denatured state of apo-plastocyanin formed in solution under conditions of low salt and neutral pH has been investigated by multidimensional heteronuclear NMR spectroscopy. Chemical shift assignments were obtained by using three-dimensional triple-resonance NMR experiments to trace through-bond heteronuclear connectivities along the backbone and side chains. The $^3J_{\text{HN,H}\alpha}$ coupling constants, ^{15}N -edited proton–proton nuclear Overhauser effects (NOEs), and ^{15}N relaxation parameters were also measured for the purpose of structural and dynamic characterization. Most of the residues corresponding to β -strands in the folded protein exhibit small upfield shifts of the $^{13}\text{C}^\alpha$ and ^{13}CO resonances relative to random coil values, suggesting a slight preference for backbone dihedral angles in the β region of (ϕ, ψ) space. This is further supported by the presence of strong sequential $d_{\alpha\text{N}}(i, i + 1)$ NOEs throughout the sequence. The few $d_{\text{NN}}(i, i + 1)$ proton NOEs that are observed are mostly in regions that form loops in the native plastocyanin structure. No medium or long-range NOEs were observed. A short sequence, between residues 59 and 63, was found to populate a nonnative helical conformation in the unfolded state, as indicated by the shift of the $^{13}\text{C}^\alpha$, ^{13}CO , and $^1\text{H}^\alpha$ resonances relative to random coil values and by the decreased values of the $^3J_{\text{HN,H}\alpha}$ coupling constants. The ^{15}N relaxation parameters indicate restriction of motions on a nanosecond timescale in this region. Intriguingly, this helical conformation is present in a sequence that is close to but not in the same location as the single short helix in the native folded protein. The results are consistent with earlier NMR studies of peptide fragments of plastocyanin and confirm that the regions of the sequence that form β -strands in the native protein spontaneously populate the β -region of (ϕ, ψ) space under folding conditions, even in the absence of stabilizing tertiary interactions. We conclude that the state of apo-plastocyanin present under nondenaturing conditions is a noncompact unfolded state with some evidence of nativelike and nonnative local structuring that may be initiation sites for folding of the protein.

Keywords: NMR; denatured state; protein folding; β -sheet protein

Understanding the mechanism of protein folding requires detailed structural characterization of all of the species

formed during the folding reaction, including the native, partially folded, and fully denatured states. Although the structure of the folded state can be determined by X-ray crystallography and NMR spectroscopy, the structures of proteins in partially folded and completely unfolded states are difficult to characterize because of their heterogeneity and conformational flexibility. Although NMR could, in principle, provide much information, NMR studies of unfolded states have until recently been hampered by the lack of spectral dispersion, especially for ^1H resonances, which

Reprint requests to: H. Jane Dyson or Peter E. Wright, Department of Molecular Biology and Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037, USA; e-mail: wright@scripps.edu or dyson@scripps.edu; fax: (858) 784-9822.

¹Present address: Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, Maryland.

Article and publication are at www.proteinscience.org/cgi/doi/10.1110/ps.00601.

makes assignment difficult. However, the ability to produce proteins labeled uniformly with ^{15}N and/or ^{13}C , coupled with the development of multidimensional heteronuclear NMR methods, now allows detailed characterization of the structure and dynamics of unfolded states. This was first demonstrated by Neri et al. (1992) for phage 434 repressor and by Logan et al. (1994) for denatured FK506 binding protein. Since then, NMR studies of unfolded states of several proteins have been reported (for reviews, see Shortle 1996; Dyson and Wright 1998). However, with the exception of a staphylococcal nuclease fragment (Alexandrescu et al. 1994) and an SH3 domain of *Drosophila* drk (Zhang and Forman-Kay 1995), the reported NMR studies of unfolded proteins have been performed under strongly denaturing conditions. Denaturant does appear to influence the conformational ensemble (Logan et al. 1994; Zhang and Forman-Kay 1995). A careful examination is therefore necessary of the conformational preferences and dynamics of defined systems under nondenaturing conditions, that is, at neutral pH and room temperature in the absence of denaturant.

Plastocyanin is a small type I or "blue" copper protein found in all green plants and in many algae. X-ray structures have been determined for plastocyanin from several species; these show that the protein folds into a Greek key β -barrel or β -sandwich topology with only a single turn of helix (Guss and Freeman 1983). The X-ray structure of poplar apo-plastocyanin, crystallized from concentrated ammonium sulfate solution, is very similar to that of the holoprotein (Garrett et al., 1984), and, according to the NMR spectrum, the apoprotein is fully folded into the native conformation in solutions containing appreciable concentrations of salt. However, at low salt concentrations the spectra possess all of the characteristics of an unfolded protein, even at neutral pH (Koide et al. 1993). Thus apo-plastocyanin represents an ideal protein for studying the structure and dynamics of the polypeptide chain of a full-length protein in an unfolded state obtained under nondenaturing conditions and in the absence of destabilizing mutations.

Results

Backbone assignment of unfolded apo-plastocyanin

The ^1H - ^{15}N heteronuclear single-quantum coherence (HSQC) spectrum of apo-plastocyanin under low-salt conditions at neutral pH is shown in Figure 1. The resonance dispersion is characteristic of an unfolded protein, but the ^{15}N - ^1H cross peaks are in general sufficiently well resolved to allow sequential assignment by triple resonance methods. A total of 92 of the expected 93 amides can be assigned; the exception is His 87, for which no N-H cross peak is observed under the conditions of the experiment. Backbone C^α assignments were completed by using HNCA, CBCA(CO)NH, and ^{13}C HSQC spectra, and H^α assign-

ments were complete except for residues 2, 83, and 86–88. The resonance assignments have been deposited in the BioMagResBank, under accession number 4922.

Chemical shift deviation from random coil values

Deviations of chemical shifts from random coil values (secondary chemical shifts) are an excellent measure of the conformational propensities of the polypeptide chain. Because secondary shifts are very small for unstructured proteins, it is imperative that appropriate random coil reference shifts are used (Schwarzinger et al. 2000; Yao et al. 2001). Figure 2 illustrates the deviation of $^{13}\text{C}^\alpha$, ^{13}CO , and H^α chemical shifts from random coil values derived from model peptides with sequence Gly-Gly-X-Ala-Gly-Gly (Wishart et al. 1995a). Corrections were made for the effect of a proline on the chemical shift of the preceding residue (Wishart et al. 1995a) and for additional sequence-dependent contributions to the chemical shift (Schwarzinger et al. 2000). The corrections include compensation for the effects of the alanine residue on the chemical shift of residue X in the peptide models (-0.17 ppm for $^{13}\text{C}^\alpha$ and -0.77 ppm for ^{13}CO [Schwarzinger et al. 2001]); in other words, the random coil shifts are normalized to the values expected for the peptide fragment Gly-X-Gly. The corrected $^{13}\text{C}^\alpha$ secondary shifts (Fig. 2A) are close to random coil values for most residues in the polypeptide. However, relatively large downfield secondary $^{13}\text{C}^\alpha$ shifts (~ 1 ppm) are observed for all six amino acids between residues Glu 59 and Asn 64, indicating that the polypeptide backbone significantly populates the α -region of (ϕ , ψ) space (Spera and Bax 1991; Wishart and Sykes 1994). Consistent downfield shifts of the ^{13}CO resonances of residues 59–64 (Fig. 2B) provide further evidence for helical structure. This region forms an extended loop in the native apo-plastocyanin structure (Garrett et al. 1984). In addition, significant downfield secondary shifts are observed for the $^{13}\text{C}^\alpha$ resonances of Ala 52 and Ser 53 and for the ^{13}CO resonances of Asp 51–Ser 53; these residues form part of the short α -helix found in the native protein structure and clearly populate α -region backbone dihedral angles in the unfolded state. A systematic upfield shift of the C^β for residues 52–54 and 59–66 provides confirmatory evidence for preferential population of helical ϕ and ψ angles in these regions.

Many residues exhibit small (0.2 – 0.4 ppm) upfield $^{13}\text{C}^\alpha$ secondary shifts, suggesting a slight preference for backbone dihedral angles in the β -region of (ϕ , ψ) space (Wishart and Sykes 1994). In several regions of the polypeptide (residues 3–5, 12–16, 19–22, 27–30, 39–42, 68–73, 79–83, and 95–98), all of which correspond to the β -strands in the native folded protein, three or more adjacent residues exhibit consistent patterns of upfield secondary shifts (Fig. 2A). The ^{13}CO secondary shifts also suggest a preference toward the β -region of (ϕ , ψ) space in these parts of the polypeptide chain. These regions include three phenylalanines

Fig. 1. 500-MHz ^1H - ^{15}N HSQC spectrum of recombinant unfolded poplar apo-plastocyanin at 35°C in low-salt buffer at near-neutral pH (5mM K_2HPO_4 , 5mM β -mercaptoethanol, pH 6.0).

(Phe14, Phe19, and Phe29) that show a preference for β dihedral angles according to the $^3J_{\text{HN,H}\alpha}$ coupling constants (see following).

The deviations of the $^1\text{H}^\alpha$ chemical shifts from the random coil values determined for the peptides Gly-Gly-X-Gly-Gly (Merutka et al. 1995), corrected for local sequence effects by using the parameters of Schwarzsinger et al. (2001), are shown in Figure 2C. In general, the $^1\text{H}^\alpha$ chemical shifts for unfolded Pc are very close to random coil values; this is not unexpected, because the changes in $^1\text{H}^\alpha$ chemical shifts with secondary structure are much smaller than those for $^{13}\text{C}^\alpha$ resonances. Nevertheless, the $^1\text{H}^\alpha$ secondary shifts, in conjunction with the $^{13}\text{C}^\alpha$ shifts, do provide useful insights into the backbone conformational preferences. Small upfield $^1\text{H}^\alpha$ secondary shifts are observed for residues 59, 60, 62, and 63, suggesting that helical ϕ angles are populated in the unfolded state, consistent with the conclusions drawn from the observed $^{13}\text{C}^\alpha$ and ^{13}CO secondary chemical shifts. Thus, all of the secondary chemical shifts indicate the presence of a short region of nonna-

tive helical structure between residues 59 and 64. In addition, the $^1\text{H}^\alpha$ resonances of Ala 52 and Ser 53 are shifted slightly upfield, confirming a propensity for α -region backbone dihedral angles in this short sequence. Small downfield $^1\text{H}^\alpha$ secondary shifts are evident for residues 3–4, 12–15, 19–22, 28–29, 40–42, 79–84, and 96–98, in agreement with the $^{13}\text{C}^\alpha$ secondary shifts that indicate a slight preference for backbone dihedral angles in the β -region of (ϕ, ψ) space.

The general preference for dihedral angles in the β -region of (ϕ, ψ) space is a characteristic of “random” polypeptide chains (Brant et al. 1967). However, these preferences should not be construed as indicating the presence of well-formed β -structure. This is shown by the $^{13}\text{C}^\beta$ secondary shifts, which in general deviate little from random coil values (data not shown). This is a strong indication that the polypeptide chain is predominantly in a random, highly fluctuational state because a pronounced downfield shift of the C^β resonance is characteristic of ordered β -structure (Spera and Bax 1991).

Fig. 2. (A) Deviation of the $^{13}\text{C}^\alpha$ chemical shifts of unfolded apo-plastocyanin from the sequence-corrected random coil values of Wishart et al. (1995a). (B) Deviation of the $^{13}\text{C}^\text{O}$ chemical shifts of unfolded apo-plastocyanin from the sequence-corrected random coil values of Wishart et al. (1995a). (C) Deviation of the $^1\text{H}^\alpha$ chemical shifts from the sequence-corrected random coil values of Merutka et al. (1995). The secondary structure in the folded apo-plastocyanin is indicated. Details of the sequence corrections are given in the text.

$^3J_{\text{HN,H}\alpha}$ coupling constants

To characterize the distribution of ϕ angles, we measured $^3J_{\text{HN,H}\alpha}$ coupling constants by using the HNHA experiment (Vuister and Bax 1993). The results are shown in Figure 3. Coupling constants for several residues could not be measured accurately because of resonance overlap. In general, the coupling constants differ only slightly from the statistical random coil values (Smith et al. 1996), corrected for local sequence context (Penkett et al. 1997). For residues Glu 60 to Leu 63, $^3J_{\text{HN,H}\alpha}$ values are slightly smaller than for random coil, in agreement with the chemical shift data in indicating a preference for dihedral angles in the α -region of (ϕ, ψ) space. Interestingly, $^3J_{\text{HN,H}\alpha}$ coupling constants for Phe 14, Phe 19, and Phe 29, each of which is located near the center of a β -strand in the native plastocyanin, are significantly larger than the random coil values (Fig. 3). Thus, both $^3J_{\text{HN,H}\alpha}$ and the upfield $^{13}\text{C}^\alpha$ secondary shifts (Fig. 2A) suggest a significant propensity for these phenylalanine residues to populate backbone dihedral angles in the β -region of (ϕ, ψ) space. We note, however, that deviations of $^3J_{\text{HN,H}\alpha}$ from random coil values are very small and it is difficult to discern subtle variations in backbone conformational propensities from coupling constants alone.

Fig. 3. (A) $^3J_{\text{HN,H}\alpha}$ coupling constants derived from HNHA three-dimensional experiments. (B) Deviations of experimental $^3J_{\text{HN,H}\alpha}$ coupling constants from context-corrected statistical random coil values (Penkett et al. 1997).

Homonuclear ^1H NOEs

To further characterize the conformational propensities of unfolded apo-plastocyanin, we measured ^1H - ^1H NOEs by using ^{15}N -edited 3D NOESY-HSQC experiments (Bax et al., 1990). Figure 4 summarizes the $d_{\alpha\text{N}}(i, i+1)$ and $d_{\text{NN}}(i, i+1)$ NOE connectivities and their relative intensities. Strong $d_{\alpha\text{N}}(i, i+1)$ NOE connectivities are observed between most residues, whereas the few $d_{\text{NN}}(i, i+1)$ NOEs that are observed are relatively weak. No medium-range or long-range NOEs could be identified unambiguously in the spectra. The strong $d_{\alpha\text{N}}(i, i+1)$ NOEs observed throughout the sequence confirm that all regions of the polypeptide chain sample the β -region of (ϕ, ψ) space. All of the $d_{\text{NN}}(i, i+1)$ NOEs observed are in regions where the $^{13}\text{C}^\alpha$ secondary shifts indicate a preference for α -region dihedral angles; these regions are all from loops and turns in the folded protein. The $d_{\text{NN}}(i, i+1)$ NOEs in the region from residues 59 to 61 are among the strongest, consistent with the conclusion that this region populates the α region of (ϕ, ψ) space in the unfolded state.

^{15}N relaxation measurements

To gain insights into the conformational flexibility of the polypeptide backbone in the low-salt denatured state of apo-plastocyanin, we measured ^{15}N T_1 and T_2 relaxation times and steady state $\{^1\text{H}\}$ - ^{15}N NOEs by using the pulse sequences described by Farrow et al. (1995a). Figure 5 shows the $R_1 (= 1/T_1)$ and $R_2 (= 1/T_2)$ relaxation rates measured

Fig. 4. Relative intensities of the sequential ^1H - ^1H NOEs in unfolded apo-plastocyanin. The thickness of the bars corresponds to the estimated NOE intensity (strong, medium, weak). Asterisks denote NOEs that are not observable because of resonance overlap.

at 500 MHz and the $\{^1\text{H}\}$ - ^{15}N heteronuclear NOE measured at both 500 MHz and 750 MHz. Both T_1 and T_2 are relatively uniform along the sequence, except at the N and C termini, where both relaxation times are increased as a

Fig. 5. (A) ^{15}N R_1 relaxation rates and (B) ^{15}N R_2 relaxation rates for unfolded apo-plastocyanin at 500-MHz spectrometer frequency. (C) $\{^1\text{H}\}$ - ^{15}N steady state NOE data at 500 (O) and 750 (Δ) MHz.

result of increased conformational flexibility. Similar patterns are also observed for the $\{^1\text{H}\}$ - ^{15}N NOEs. It is notable that all of the heteronuclear NOEs are negative at both 500 and 750 MHz, except in the region from residues 59 to 63 at the higher field. The fact that the NOEs are negative at both frequencies indicates large amplitude fluctuations of most of the polypeptide chain on a timescale shorter than 1 nsec. A small degree of motional restriction is evident in the region between residues 59 and 63. Measurements of the diffusion coefficient by using sLED pulse sequences (Altieri et al. 1995) indicate that the diffusion coefficient of unfolded apoPc is significantly greater (~ 1.5 times larger) than that of fully folded apoPc (data not shown). This is as expected for a highly expanded unfolded protein, on the basis of published comparisons of diffusion coefficients and radii of hydration of folded and unfolded proteins (Wilkins et al. 1999).

^{15}N relaxation data for globular proteins are commonly interpreted by using the Lipari-Szabo model-free formalism (Lipari and Szabo 1982a,b). However, model-free analysis is of doubtful validity for an unfolded polypeptide because the motions of the backbone ^{15}N - ^1H vectors are expected to be complex and are not likely to be accurately represented by a spectral density that assumes a single isotropic rotational correlation time. Rather, like polymers, the dynamics of unfolded states of proteins are probably better characterized by a distribution of correlation times. For this reason, we analyzed the relaxation data by using the reduced spectral density mapping method of Farrow et al. (1995a) to determine the values of $J(0)$, $J(\omega_N)$, and $J(0.87\omega_H)$, that is, to sample the spectral density at 0, 50, and 435 MHz. The resulting values of $J(\omega)$ are shown in Figure 6. The most striking features of the spectral densities are the unusually small values of $J(0)$ (< 0.4 nsec. rad^{-1}) throughout the poly-

Fig. 6. Calculated values of $J(0)$ (open circles), $J(50)$ (open triangles), and $J(435)$ (open squares) for unfolded apo-plastocyanin.

peptide, the large values of $J(0.87\omega_H)$ (ranging from 0.03–0.05 nsec. rad^{-1}) relative to those typically found for globular proteins, and the fact that for most residues $J(0)$ is only slightly larger than $J(\omega_N)$. In other words, the spectral density function is flat and extended, indicating that relaxation is dominated by rapid backbone fluctuations occurring on a subnanosecond timescale. The largest values of $J(0)$ are found in the region between Ser 17 and Pro 23, where chemical shifts and coupling constants suggest a propensity toward native-like β -structure. In addition, large values of $J(0)$ and $J(\omega_N)$ occur in the region between Met 57 and Asn 64, where other NMR parameters indicate a substantial population of nonnative helical structure. Thus, some motional restriction is evident in these preferentially structured regions.

Discussion

Structural characterization of unfolded proteins is crucial for understanding the protein folding problem because the unfolded state represents the starting point of protein folding. It has been suggested that areas of residual structure in unfolded states may serve as initiation sites for protein folding (Dyson et al. 1985; Baldwin 1986; Wright et al. 1988). Multidimensional NMR is an especially powerful tool to provide understanding of these issues (Dyson and Wright 1996).

As for folded proteins and peptides, structural information on the nature and population of structured forms can be

deduced for the unfolded state from chemical shifts, coupling constants, and ^1H – ^1H NOEs. The most sensitive NMR parameters for determining secondary structural propensities are the secondary $^{13}\text{C}^\alpha$, ^{13}CO , $^{13}\text{C}^\beta$, and $^1\text{H}^\alpha$ chemical shifts, that is, the deviations of these chemical shifts from random coil values (Yao et al. 1997; Dyson and Wright 1998; Eliezer et al. 2000). Provided suitable random coil reference shifts are available for the solution conditions used, and provided these are corrected for local sequence contributions to the chemical shifts (Schwarzinger et al. 2001), then the $^{13}\text{C}^\alpha$, ^{13}CO , and, to a lesser extent, the $^1\text{H}^\alpha$ secondary shifts can provide a quantitative estimate of the population of secondary structure (Spera and Bax 1991; Wishart et al. 1991; Yao et al. 2001). In principle, the $^3J_{\text{HN,H}\alpha}$ coupling constant can also give information on residual structure in unfolded proteins. In practice, these parameters are not sufficiently sensitive to small differences in backbone structure populations, with the result that coupling constants in conformationally averaged systems are in general not very useful in delineating structure.

We cannot emphasize too strongly the importance of using appropriate random coil reference shifts if small deviations from statistical random coil behavior are to be identified with confidence. In the present case, for unfolded apo-plastocyanin in water solution at pH 6, the appropriate random coil chemical shift references for each residue X are those pertaining to the peptide fragment –Gly-X-Gly–. For $^1\text{H}^\alpha$, reference shifts have been determined for the peptides Gly-Gly-X-Gly-Gly in water solution at pH 5.0 (Merutka et al. 1995). For $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$, and ^{13}CO , reference shifts have been measured by using the peptides Gly-Gly-X-Ala-Gly-Gly (Wishart et al. 1995a); however, these need to be corrected for the effects of the alanine residue on the chemical shifts of the preceding residue X to bring them into the same –Gly-X-Gly– chemical shift reference frame. In addition, it is imperative to correct the $^{13}\text{C}^\alpha$, ^{13}CO , and $^1\text{H}^\alpha$ reference shifts for local sequence-dependent effects arising from the two residues preceding and the two residues following X. These correction factors have been determined by Schwarzinger et al. (2001). The sequence-dependent contributions are large enough to mask the subtle secondary chemical shift variations associated with local backbone conformational preferences in unfolded proteins, unless corrections are applied.

Residual structure in unfolded apo-plastocyanin from NMR

The secondary $^{13}\text{C}^\alpha$ and ^{13}CO chemical shifts (Fig. 2) indicate a small but detectable preference for the β -region of (ϕ, ψ) space in several contiguous regions of the plastocyanin sequence. In other words, in these regions the polypeptide chain samples β backbone dihedral angles to a greater extent than would be expected for a statistical random coil. It

is interesting that all of the regions that favor β dihedrals occur within β -strands in the folded apo-plastocyanin structure (Garrett et al. 1984). Although the secondary shifts are much smaller, the sequence-corrected $^1\text{H}^\alpha$ data are fully consistent with the $^{13}\text{C}^\alpha$ and ^{13}CO chemical shifts. These regions are also associated with the most intense $d_{\alpha\text{N}}(i, i + 1)$ sequential NOEs, and residues Phe 14 and Phe 19 display abnormally large $^3J_{\text{HN, H}\alpha}$ coupling constants, suggesting a strong bias toward the β -region of (ϕ, ψ) space at these residues. The majority of the sequence of unfolded apo-plastocyanin shows little tendency to populate backbone dihedral angles other than the β region of (ϕ, ψ) space. This region of the Ramachandran plot is also most populated in free-flight "random coil" polypeptides (Brant et al. 1967). Some parts of the protein do show measurable tendencies toward local structure. In the region of Phe 19, the $^{13}\text{C}^\alpha$ chemical shifts indicate a significant tendency of the backbone to populate dihedral angles typical of β -sheet, and the unusually high value of the $^3J_{\text{HN, H}\alpha}$ of Phe 19 suggests that the local β -structure is highly populated at this residue.

Downfield $^{13}\text{C}^\alpha$ and ^{13}CO shifts and upfield $^1\text{H}^\alpha$ secondary shifts are observed in many of the regions that form turns or loops in the folded protein (Fig. 2), indicating increased sampling of the α -region of (ϕ, ψ) space. These are also the regions in which sequential $d_{\text{NN}}(i, i + 1)$ NOEs are observed, confirming an increased population of α -region dihedral angles. Nevertheless, in all of these segments, $d_{\alpha\text{N}}(i, i + 1)$ NOEs are also observed, showing that the polypeptide chain samples both the α - and β -regions of (ϕ, ψ) space (Dyson and Wright 1996).

A measurable population of helical structure is present between Glu 59 and Asn 64. The mean population of helix was estimated to be 28% from the magnitude of the $^{13}\text{C}^\alpha$, ^{13}CO , and $^1\text{H}^\alpha$ secondary chemical shifts, which gave helical populations of 28%, 30%, and 24%, respectively. (These populations were estimated by using helical secondary shifts of 2.8 ppm [$^{13}\text{C}^\alpha$], 2.1 ppm [^{13}CO], and 0.3 ppm [$^1\text{H}^\alpha$] [Williamson 1990; Wishart and Sykes 1994]). Interestingly, a helical propensity of 23% is predicted by the peptide-helix prediction algorithm AGADIR (Muñoz and Serrano 1997) at 308K, the temperature of the NMR experiments. Residues 59–64 are not helical in the native protein, but form part of a loop between a short helical segment between residues 52 and 55 and the β -strand between residues 67 and 74. The short Ala 52–Ile 55 helix in the native structure is not significantly populated in unfolded apo-plastocyanin, although the $^{13}\text{C}^\alpha$, ^{13}CO , and $^{13}\text{C}^\beta$ shifts do suggest a small propensity for α -region dihedrals at residue 52 and 53 and may indicate a nucleation site for the short helix in the native folded protein.

Turns and nascent helices are excellent candidates for the initiation of protein folding (Wright et al. 1988). A nonnative helical segment has been observed in the acid-unfolded form of the helical protein apomyoglobin (Eliezer et al.

1998; Yao et al. 2001). The present observation of a population of nonnative helix in a β -sheet protein is a hint that short segments of local structure (helix, turn, or hydrophobic cluster), not necessarily present in the final folded form of the protein, may be a common feature in the initiation of the folding of proteins. The intrinsically plastic nature of such structures, where the equilibrium between folded and unfolded conformers is rapid, would allow distant parts of the chain to come into contact, in furtherance of the folding reaction, but would also allow the nonnative initiating structures to be easily dissolved as the folding reaction progressed.

An important result of the present work is that the intrinsically unfolded form of apo-plastocyanin has a weak propensity to populate the β -region of (ϕ, ψ) space, and that the preference for β -region dihedrals is most pronounced in regions corresponding to β -strands in folded plastocyanin. This is consistent with studies that use overlapping peptide fragments corresponding to the sequence of the French bean protein (Dyson et al. 1992b). Resonance assignments, temperature coefficients, and intramolecular NOEs were used to evaluate the propensity for folded structures in the peptides: evidence was seen for a hydrophobic cluster between the side chains of Ile 27 and Phe 29, and for a weakly populated β -turn conformation between residues Ala 65–Pro–Gly–Glu 68. Apart from these minor folded forms, the majority of the molecule was found to populate unfolded conformations in the β -region of (ϕ, ψ) space. The amino acid sequences of the French bean and poplar proteins are highly homologous, with 78 of 99 residues identical and a further 7 residues substituted conservatively. However, the turn conformation in residues 65–69 of the peptide models was not detected in the full-length unfolded protein. This may be due to the substitution of Pro 66 in the French bean sequence by Lys in the poplar sequence, thus lowering the probability of turn formation in the absence of tertiary interactions: the sequence APGE has a relatively high probability of turn formation, whereas that of AKGE is lower (Chou 1997; Dyson et al. 1988). Conversely, the nonnative helical segment observed in the unfolded protein between residues 59 and 65 was not observed in the peptides, which were designed according to the secondary structural elements in the folded protein; the 59–65 sequence therefore spanned the dividing line between two peptides, Pc11 and Pc13.

Backbone dynamics

The ^{15}N relaxation data show that the polypeptide backbone undergoes large amplitude fluctuations on a subnanosecond timescale throughout its length. This is immediately evident from the negative $\{^1\text{H}\}-^{15}\text{N}$ heteronuclear NOEs at both 500 MHz and at 750 MHz. The small values of $J(0)$ and the fact that $J(0)$ is similar in magnitude to $J(\omega_{\text{N}})$ confirm the highly dynamic nature of this unfolded state of apo-plasto-

cyanin. In its folded globular state, apo-plastocyanin is expected to have a rotational correlation time, τ_m of the order of 5 nsec, with $J(0)$ values approaching $2 \text{ nsec} \cdot \text{rad}^{-1}$ [$J(0)_{\text{max}} = 2/5 \tau_m$]. In marked contrast, for the low-salt unfolded state, the maximum $J(0)$ is $0.40 \text{ nsec} \cdot \text{rad}^{-1}$ for Glu 18. The smallest values of $J(0)$ and $J(\omega_N)$ occur for residues near the termini, consistent with increased segmental mobility at the ends of the polypeptide chain. The motional characteristics of unfolded apo-plastocyanin are reminiscent of those of a random coil polypeptide. Molecular dynamics simulations show that the effective motional correlation time of random coil polypeptides varies along the chain, with greatest flexibility near the termini (Hu et al. 1990). A plot of correlation time versus position in the chain has a flattened bell-shaped profile, similar in overall shape to the plots of $J(0)$ and $J(\omega_N)$ versus residue number for unfolded apo-plastocyanin (Fig. 6). Local maxima and minima in the profile indicate influences of the local amino acid sequence on chain flexibility (Hu et al. 1990). For apo-plastocyanin, the plots of $J(0)$ and $J(\omega_N)$ versus residue number (Fig. 6) are relatively featureless from residue 10 through 97, although local maxima can be discerned between residues 17 and 24 in $J(0)$, and between residues 57 and 64 in both $J(0)$ and $J(\omega_N)$. Both regions correspond to segments of the polypeptide that have relatively high propensities to form structure. The $^3J_{\text{HN,H}\alpha}$ coupling constant for Phe 19 is anomalously large, suggesting preference for formation of ordered structure. The restricted motions evident for residues 59–64 are probably associated with the helix formation indicated by the chemical shifts. Although backbone motions on the nanosecond time scale are slightly restricted in these regions, the small values of $J(0)$ indicate that there is still considerable flexibility. In addition, slightly elevated values of $J(\omega_N)$ for residues 52–54 suggest the possibility of motional restriction in the region of the native plastocyanin helix. However, in the segments between residues 4 and 8, 12 and 15, 27 and 30, 40 and 42, and 68 and 69, where the chemical shifts reveal propensity toward population of backbone dihedral angles in the β -region of (ϕ, ψ) space, the polypeptide exhibits a high degree of flexibility on a subnanosecond timescale with little indication of significant motional restriction. We thus conclude that even though the backbone dihedral angles in these segments favor the β -region of (ϕ, ψ) space, the polypeptide chain is highly fluctuational.

^{15}N relaxation studies have been reported for a number of unfolded proteins (Torchia et al. 1975; van Mierlo et al. 1993; Alexandrescu and Shortle 1994; Farrow et al. 1995a; Frank et al. 1995; Buck et al. 1996; Farrow et al. 1997; Eliezer et al. 1998; Meekhof and Freund 1999). Unfolded proteins that possess little ordered structure are characterized by relatively uniform backbone dynamics (Torchia et al. 1975; Frank et al. 1995). The urea-denatured form of the fibronectin type III (TNfn3) domain at 35°C (Meekhof and Freund 1999) has spectral densities that are comparable in

magnitude to those of the low-salt state of apo-plastocyanin. In common with many other denatured proteins (Alexandrescu and Shortle 1994; Buck et al. 1996), the regions of maximal motional restriction in TNfn3 are strongly hydrophobic. However, neither of the regions of apo-plastocyanin where backbone motions are restricted is particularly hydrophobic; in fact, there is no apparent correlation between backbone flexibility and the Kyte-Doolittle hydrophobicity for apo-plastocyanin.

Like apo-plastocyanin, the N-terminal SH3 domain of the protein drk forms an unfolded state in aqueous solution at neutral pH, in the absence of chemical denaturant (Farrow et al. 1995a). However, the drk SH3 domain differs significantly from unfolded apo-plastocyanin in its dynamic behavior. The average value of $J(0)$ for the SH3 domain is much larger than for apo-plastocyanin, even though the polypeptide is smaller (59 versus 99 residues). For some residues, $J(0)$ in the unfolded state is nearly as large as in the native SH3 domain. Thus the spectral densities indicate that the intrinsically unstructured state of the drk SH3 domain is significantly more compact than that formed by apo-plastocyanin, which appears to be rather extended and highly flexible on a subnanosecond time scale.

Potential initiation sites for apo-plastocyanin folding

Unfolded apo-plastocyanin, obtained in low-salt medium at pH 6, provides a unique opportunity to characterize the intrinsic structural propensities of a polypeptide chain under nondenaturing conditions. The most striking observation is that the polypeptide appears to behave as a largely extended and highly fluctuational chain, with no indications of global compaction. The nanosecond timescale motional restriction evident for residues 57–64 indicates some local chain compaction, presumably associated with transient formation of nonnative helical structure. The slight motional restriction observed for residues 17–24 suggests transient local interactions in this region as well. The comparative uniformity of backbone motions in unfolded apo-plastocyanin is in marked contrast to the behavior of acid-unfolded apomyoglobin, in which substantial restriction of motion is observed in many regions of the polypeptide chain (Yao et al. 2001). Marked differences in intrinsic backbone conformational propensities were also observed for peptide fragments derived from plastocyanin and those derived from the helical proteins myoglobin and myohemerythrin (Dyson et al. 1992a,b; Waltho et al. 1993; Reymond et al. 1997). We suggest that these differences in behavior may reflect intrinsic differences in the folding processes of β -sheet and helical proteins. Formation of helical structure is driven by local interactions, and hence local structure formation and local chain compaction is prevalent. However, formation of β -sheet structure requires interactions between distant regions of the polypeptide chain and these are only likely to

occur after global collapse. Nevertheless, many regions of the polypeptide chain do favor dihedral angles in the β -region of $\phi\psi$ space so that the amino acid sequence biases the conformational ensemble toward the secondary structure found in the final folded state. In addition, many of the turn and loop regions of the folded structure display an intrinsic conformational bias toward the α -region of conformational space. Thus, the regions of backbone conformational space sampled by the plastocyanin polypeptide chain are biased by the amino acid sequence, but with the exception of non-native helical structure in one of the loop regions, there is little propensity for significant local structure formation in the absence of stabilizing long-range interactions. Nevertheless, mathematical analysis of a simple model shows that a small bias in conformation such as we observe may be sufficient to reduce the search time for folding to levels reasonably consistent with those observed in nature (Zwanzig et al. 1992).

Materials and methods

NMR sample preparation

Poplar apo-plastocyanin was expressed in *Escherichia coli* with $(^{15}\text{NH}_4)_2\text{SO}_4$ and/or ^{13}C -glucose as sole nitrogen and carbon sources. The purification procedure has been described elsewhere (Koide et al. 1993). To prepare NMR samples, we dissolved lyophilized protein in 0.5 mL 6M GdmCl solution, then refolded by dilution with a large volume of buffer at pH 6.0 containing 1.0 M NaCl, 200 μM EDTA, and 5 mM K_2HPO_4 . The solution was further concentrated to 0.5 mL. It was then passed through a spin column that was filled with Sephadex G25 equilibrated with pH 6.0 buffer, containing 5mM K_2HPO_4 , 50 μM EDTA, 5 mM β -mercaptoethanol, and 10% D_2O . After the sample was transferred into the NMR tube, it was flushed with argon and sealed with a rubber cap and parafilm. The final concentrations of ^{15}N -labeled and ^{15}N , ^{13}C -double labeled samples were 2.0 mM and 1.5 mM, respectively. Two samples at 0.5 mM and 0.1 mM protein concentration were also prepared to check the concentration dependence of the spectra.

NMR spectroscopy

Most of the NMR spectra were acquired at 35°C on a Bruker AMX 500 spectrometer with a triple resonance z-shielded gradient probe. A ^{15}N -edited NOESY-HSQC spectrum was recorded on a Bruker AMX 600 spectrometer. Resonance assignments were made by using a suite of triple resonance experiments, including C(CO)NH-TOCSY (Grzesiek et al. 1993; Logan et al. 1993), CBCA(CO)NH (Grzesiek and Bax 1992a), TOCSY-HSQC (Marion et al. 1989), HNCACB (Wittekind and Mueller 1993), HNCA, HNCO (Grzesiek and Bax 1992b), and (HA)CO(CA)NH (Löhner and Rüterjans 1995). $^3J_{\text{HN,H}\alpha}$ coupling constants were measured by using the HNHA pulse sequence (Vuister and Bax 1993). In all of the three-dimensional (3D) experiments, the spectral width in the ^1H acquisition dimension was 4032 Hz, with 1K complex points. The spectral width in the indirect ^1H dimension in the NOESY-HSQC, TOCSY-HSQC, and HNHA experiments was 7800 Hz, with 128 complex points. The spectral width in the ^{15}N dimension was 1506

Hz, and 32 complex points were generally acquired. The ^{13}C spectral width was 3774 Hz, with 56 complex points. Solvent suppression was achieved by the water flip back method (Grzesiek and Bax 1993). All of the 3D experiments used the gradient-selected sensitivity enhanced pulse sequence (Muhandiram and Kay 1994). The offset frequency in the acquisition domain was set to the water signal (Wishart et al. 1995b). The ^1H spectrum was referenced to water at 35°C (4.65 ppm). All heteronuclear dimensions were referenced indirectly on the basis of the water signal. A 45-msec mixing time was used in the C(CO)NH-TOCSY experiment, 80 msec in the ^{15}N edited TOCSY-HSQC, and 200 msec in the ^{15}N NOESY-HSQC experiment. All data were processed by using Felix2.30 or Felix95 (Molecular Simulations, Inc.).

^{15}N T_1 and T_2 relaxation times and steady state $\{^1\text{H}\}$ - ^{15}N NOEs were measured by using the pulse sequences described by Farrow et al. (1995a). The data were acquired at 35°C by using a Bruker AMX 500 spectrometer. Heteronuclear NOE data were also acquired at 750 MHz. For the T_1 measurements, 9 time points were collected with delays of 22, 64, 126, 251, 376, 522, 731, 1148, and 2023 msec; duplicate data were acquired at three delay times to allow estimation of errors. For T_2 , 12 time points with delays of 7, 27, 52, 77, 102, 203, 253, 353, 404, 454, 504, and 1006 msec were acquired. A recycle delay of 5 sec was used for both the T_1 and T_2 measurements. The steady state $\{^1\text{H}\}$ - ^{15}N NOE was measured from duplicate pairs of ^1H - ^{15}N spectra recorded with and without amide proton saturation. Saturation was achieved by using a train of 120° - ^1H pulses applied for a period of 2.3 sec after the 7-sec recycle delay. All spectra at 500 MHz were recorded by using 128 and 256 complex points in the t_1 and t_2 dimensions, respectively, with spectral widths of 1501 Hz (^{15}N) and 4029 Hz (^1H) and with 16 scans per t_1 point.

The amide ^{15}N relaxation data were analyzed by using the reduced spectral density mapping approach (Farrow et al. 1995b; Peng and Wagner 1995). The method assumes that the spectral density function $J(\omega)$ is relatively constant in the high frequency region near $J(\omega_H)$. More explicitly, the approximation $J(\omega_H + \omega_N) \sim J(\omega_H - \omega_N) \sim J(0.87\omega_H)$ is made (Farrow et al., 1995b), allowing calculation of discrete values of the spectral density at three frequencies, $\omega = 0$, $\omega = \omega_N$, and $\omega = 0.87\omega_H$, from the following relationships:

$$1/T_1 = (d^2/4)[3J(\omega_N) + 7J(0.87\omega_H)] + c^2J(\omega_N)$$

$$1/T_2 = (d^2/8)[4J(0) + 3J(\omega_N) + 13J(0.87\omega_H)] + (c^2/6)[3J(\omega_N) + 4J(0)]$$

$$\text{NOE} = 1 + T_1(\gamma^H/\gamma^N)(d^2/4)[5J(0.87\omega_H)]$$

where $d = (\mu_0 h / 8\pi^2) \gamma^H \gamma^N \langle r_{\text{NH}}^{-3} \rangle$, $c = \omega_N(\sigma_{\parallel} - \sigma_{\perp})/\sqrt{3}$, γ is the gyromagnetic ratio, h is Planck's constant, μ_0 is the permeability of a vacuum, $\langle r_{\text{NH}} \rangle$ is the N-H bond length, taken as 1.02 Å, and $\sigma_{\parallel} - \sigma_{\perp}$ is the axially symmetric chemical shift tensor for backbone ^{15}N nuclei, taken as -160 ppm.

Acknowledgments

We thank Drs. Jian Yao and Xiang Li for assistance with the NMR experiments and Dr. Shohei Koide and Martine Reymond for help and advice with protein purification. This work was supported by grant GM57374 from the National Institutes of Health. Y.B. acknowledges receipt of an NIH postdoctoral fellowship (GM18193).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby

marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Alexandrescu, A.T. and Shortle, D. 1994. Backbone dynamics of a highly disordered 131 residue fragment of staphylococcal nuclease. *J. Mol. Biol.* **242**: 527–546.
- Alexandrescu, A.T., Abeygunawardana, C., and Shortle, D. 1994. Structure and dynamics of a denatured 131-residue fragment of staphylococcal nuclease: A heteronuclear NMR study. *Biochemistry* **33**: 1063–1072.
- Altieri, A.S., Hinton, D.P., and Byrd, R.A. 1995. Association of biomolecular systems via pulsed field gradient NMR self-diffusion measurements. *J. Am. Chem. Soc.* **117**: 7566–7567.
- Baldwin, R.L. 1986. Seeding protein folding. *Trends Biochem. Sci.* **11**: 6–9.
- Bax, A., Ikura, M., Kay, L.E., Torchia, D.A., and Tschudin, R. 1990. Comparison of different modes of two-dimensional reverse-correlation NMR for the study of proteins. *J. Magn. Reson.* **86**: 304–318.
- Brant, D.A., Miller, W.G., and Flory, P.J. 1967. Conformational energy estimates for statistically coiling polypeptide chains. *J. Mol. Biol.* **23**: 47–65.
- Buck, M., Schwalbe, H., and Dobson, C.M. 1996. Main-chain dynamics of a partially folded protein: ^{15}N NMR relaxation measurements of hen egg white lysozyme denatured in trifluoroethanol. *J. Mol. Biol.* **257**: 669–683.
- Chou, K.-C. 1997. Prediction of β -turns. *J. Peptide Res.* **49**: 120–144.
- Dyson, H.J. and Wright, P.E. 1996. Insights into protein folding from NMR. *Annu. Rev. Phys. Chem.* **47**: 369–395.
- . 1998. Equilibrium NMR studies of unfolded and partially folded proteins. *Nat. Struct. Biol.* **5**: 499–503.
- Dyson, H.J., Cross, K.J., Houghten, R.A., Wilson, I.A., Wright, P.E., and Lerner, R.A. 1985. The immunodominant site of a synthetic immunogen has a conformational preference in water for a Type-II reverse turn. *Nature* **318**: 480–483.
- Dyson, H.J., Rance, M., Houghten, R.A., Lerner, R.A., and Wright, P.E. 1988. Folding of immunogenic peptide fragments of proteins in water solution. I. Sequence requirements for the formation of a reverse turn. *J. Mol. Biol.* **201**: 161–200.
- Dyson, H.J., Merutka, G., Waltho, J.P., Lerner, R.A., and Wright, P.E. 1992a. Folding of peptide fragments comprising the complete sequence of proteins: Models for initiation of protein folding. I. Myohemerythrin. *J. Mol. Biol.* **226**: 795–817.
- Dyson, H.J., Sayre, J.R., Merutka, G., Shin, H.-C., Lerner, R.A., and Wright, P.E. 1992b. Folding of peptide fragments comprising the complete sequence of proteins: models for the initiation of protein folding. II. Plastocyanin. *J. Mol. Biol.* **226**: 819–835.
- Eliezer, D., Yao, J., Dyson, H.J., and Wright, P.E. 1998. Structural and dynamic characterization of partially folded states of myoglobin and implications for protein folding. *Nat. Struct. Biol.* **5**: 148–155.
- Eliezer, D., Chung, J., Dyson, H.J., and Wright, P.E. 2000. Native and non-native structure and dynamics in the pH 4 intermediate of apomyoglobin. *Biochemistry* **39**: 2894–2901.
- Farrow, N.A., Zhang, O., Forman-Kay, J.D., and Kay, L.E. 1995a. Comparison of the backbone dynamics of a folded and an unfolded SH3 domain existing in solution in aqueous buffer. *Biochemistry* **34**: 868–878.
- Farrow, N.A., Zhang, O., Szabo, A., Torchia, D.A., and Kay, L.E. 1995b. Spectral density function mapping using ^{15}N relaxation data exclusively. *J. Biomol. NMR* **6**: 153–162.
- Farrow, N.A., Zhang, O., Forman-Kay, J.D., and Kay, L.E. 1997. Characterization of the backbone dynamics of folded and denatured states of an SH3 domain. *Biochemistry* **36**: 2390–2402.
- Frank, M.K., Clore, G.M., and Gronenborn, A.M. 1995. Structural and dynamic characterization of the urea denatured state of the immunoglobulin binding domain of streptococcal protein G by multidimensional heteronuclear NMR spectroscopy. *Protein Sci.* **4**: 2605–2615.
- Garrett, T.P.J., Clingeffer, D.J., Guss, J.M., Rogers, S.J., and Freeman, H.C. 1984. The crystal structure of poplar apo-plastocyanin at 1.8 Å resolution: The geometry of the copper-binding site is created by the polypeptide. *Paper. J. Biol. Chem.* **259**: 2822–2825.
- Grzesiek, S. and Bax, A. 1992a. An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J. Magn. Reson.* **99**: 201–207.
- . 1992b. Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J. Magn. Reson.* **96**: 432–440.
- . 1993. The importance of not saturating H_2O in protein NMR: Application to sensitivity enhancement and NOE measurements. *J. Am. Chem. Soc.* **115**: 12593–12594.
- Grzesiek, S., Anglister, J., and Bax, A. 1993. Correlation of backbone amide and aliphatic side chain resonances in $^{13}\text{C}/^{15}\text{N}$ -enriched proteins by isotropic mixing of ^{13}C magnetization. *J. Magn. Reson. B* **101**: 114–119.
- Guss, J.M. and Freeman, H.C. 1983. Structure of oxidized poplar plastocyanin at 1.6 Å resolution. *J. Mol. Biol.* **169**: 521–563.
- Hu, Y., MacInnis, J.M., Cherayil, B.J., Fleming, G.R., Freed, K.F., and Perico, A. 1990. Polypeptide dynamics: Experimental tests of an optimized Rouse-Zimm type model. *J. Chem. Phys.* **93**: 822–836.
- Koide, S., Dyson, H.J., and Wright, P.E. 1993. Characterization of a folding intermediate of apoplastocyanin trapped by proline isomerization. *Biochemistry* **32**: 12299–12310.
- Lipari, G. and Szabo, A. 1982a. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* **104**: 4546–4559.
- . 1982b. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc.* **104**: 4559–4570.
- Logan, T.M., Olejniczak, E.T., Xu, R.X., and Fesik, S.W. 1993. A general method for assigning NMR spectra of denatured proteins using 3D HC(CO)NH-TOCSY triple resonance experiments. *J. Biomol. NMR* **3**: 225–231.
- Logan, T.M., Thériault, Y., and Fesik, S.W. 1994. Structural characterization of the FK506 binding protein unfolded in urea and guanidine hydrochloride. *J. Mol. Biol.* **236**: 637–648.
- Löhr, F. and Rüterjans, H. 1995. A new triple-resonance experiment for the sequential assignment of backbone resonances in proteins. *J. Biomol. NMR* **6**: 189–197.
- Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M., and Clore, G.M. 1989. Overcoming the overlap problem in the assignment of ^1H NMR spectra of larger proteins by use of three-dimensional heteronuclear ^1H - ^{15}N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: Application to interleukin β . *Biochemistry* **28**: 6150–6156.
- Meekhof, A.E. and Freund, S.M.V. 1999. Probing residual structure and backbone dynamics on the milli- to picosecond timescale in a urea-denatured fibronectin type III domain. *J. Mol. Biol.* **286**: 579–592.
- Merutka, G., Dyson, H.J., and Wright, P.E. 1995. 'Random coil' ^1H chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG. *J. Biomol. NMR* **5**: 14–24.
- Muhandiram, D.R. and Kay, L.E. 1994. Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. *J. Magn. Reson. B* **103**: 203–216.
- Muñoz, V. and Serrano, L. 1997. Development of the multiple sequence approximation within the AGADIR model of α -helix formation: Comparison with Zimm-Bragg and Lifson-Roig formalisms. *Biopolymers* **41**: 495–509.
- Neri, D., Billeter, M., Wider, G., and Wüthrich, K. 1992. NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science* **257**: 1559–1563.
- Peng, J.W. and Wagner, G. 1995. Frequency spectrum of NH bonds in eglin C from spectral density mapping at multiple fields. *Biochemistry* **34**: 16733–16752.
- Penkett, C.J., Redfield, C., Dodd, I., Hubbard, J., McBay, D.L., Mossakowska, D.E., Smith, R.A.G., Dobson, C.M., and Smith, L.J. 1997. NMR analysis of main-chain conformational preferences in an unfolded fibronectin-binding protein. *J. Mol. Biol.* **274**: 152–159.
- Reymond, M.T., Merutka, G., Dyson, H.J., and Wright, P.E. 1997. Folding propensities of peptide fragments of myoglobin. *Protein Sci.* **6**: 706–716.
- Schwarzinger, S., Kroon, G.J.A., Foss, T.R., Wright, P.E., and Dyson, H.J. 2000. Random coil chemical shifts in acidic 8 M urea: Implementation of random coil chemical shift data in NMRView. *J. Biomol. NMR* **18**: 43–48.
- Schwarzinger, S., Kroon, G.J.A., Foss, T.R., Chung, J., Wright, P.E., and Dyson, H.J. 2001. Sequence dependent correction of random coil NMR chemical shifts. *J. Am. Chem. Soc.* (in press).
- Shortle, D. 1996. The denatured state (the other half of the folding equation) and its role in protein stability. *FASEB J.* **10**: 27–34.
- Smith, L.J., Bolin, K.A., Schwalbe, H., MacArthur, M.W., Thornton, J.M., and Dobson, C.M. 1996. Analysis of main chain torsion angles in proteins: Prediction of NMR coupling constants for native and random coil conformations. *J. Mol. Biol.* **255**: 494–506.
- Spera, S. and Bax, A. 1991. Empirical correlation between protein backbone conformation and $\text{C}\alpha$ and $\text{C}\beta$ ^{13}C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc.* **113**: 5490–5492.
- Torchia, D.A., Lyster, J.R.J., and Quattrone, A.J. 1975. Molecular dynamics and structure of the random coil and helical states of the collagen peptide,

- α 1-CB2, as determined by ^{13}C magnetic resonance. *Biochemistry* **14**: 887–900.
- van Mierlo, C.P.M., Darby, N.J., Keeler, J., Neuhaus, D., and Creighton, T.E. 1993. Partially folded conformation of the (30–51) intermediate in the disulphide folding pathway of bovine pancreatic trypsin inhibitor ^1H and ^{15}N resonance assignments and determination of backbone dynamics from ^{15}N relaxation measurements. *J. Mol. Biol.* **229**: 1125–1146.
- Vuister, G.W. and Bax, A. 1993. Quantitative J correlation: A new approach for measuring homonuclear three-bond $J(\text{H}^{\text{N}}\text{H}^{\alpha})$ coupling constants in ^{15}N -enriched proteins. *J. Am. Chem. Soc.* **115**: 7772–7777.
- Waltho, J.P., Feher, V.A., Merutka, G., Dyson, H.J., and Wright, P.E. 1993. Peptide models of protein folding initiation sites. 1. Secondary structure formation by peptides corresponding to the G- and H-helices of myoglobin. *Biochemistry* **32**: 6337–6347.
- Wilkins, D.K., Grimshaw, S.B., Receveur, V., Dobson, C.M., Jones, J.A., and Smith, L.J. 1999. Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry* **38**: 16424–16431.
- Williamson, M.P. 1990. *Biopolymers* **29**: 1423–1431.
- Wishart, D.S. and Sykes, B.D. 1994. Chemical shifts as a tool for structure determination. *Methods Enzymol.* **239**: 363–392.
- Wishart, D.S., Sykes, B.D., and Richards, F.M. 1991. Simple techniques for the quantification of protein secondary structure by ^1H NMR spectroscopy. *FEBS Lett.* **293**: 72–80.
- Wishart, D.S., Bigam, C.G., Holm, A., Hodges, R.S., and Sykes, B.D. 1995a. ^1H , ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids: I. Investigations of nearest neighbor effects. *J. Biomol. NMR* **5**: 67–81.
- Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L., and Sykes, B.D. 1995b. ^1H , ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **6**: 135–140.
- Wittekind, M. and Mueller, L. 1993. HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the α - and β -carbon resonances in proteins. *J. Magn. Reson.* **101**: 201–205.
- Wright, P.E., Dyson, H.J., and Lerner, R.A. 1988. Conformation of peptide fragments of proteins in aqueous solution: Implications for initiation of protein folding. *Biochemistry* **27**: 7167–7175.
- Yao, J., Dyson, H.J., and Wright, P.E. 1997. Chemical shift dispersion and secondary structure prediction in unfolded and partly folded proteins. *FEBS Lett.* **419**: 285–289.
- Yao, J., Chung, J., Eliezer, D., Wright, P.E., and Dyson, H.J. 2001. NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding. *Biochemistry* (in press).
- Zhang, O. and Forman-Kay, J.D. 1995. Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer. *Biochemistry* **34**: 6784–6794.
- Zwanzig, R., Szabo, A., and Bagchi, B. 1992. Levinthal's paradox. *Proc. Natl. Acad. Sci.* **89**: 20–22.