Hudson, P., John, M., Crawford, R., Haralambidis, J., Scanlon, D., Gorman, J., Tregear, G., Shine, J., & Niall, H. (1984) *EMBO J. 3*, 2333–2339.
Issacs, N., James, R., Niall, H., Bryant-Greenwood, G., Dodson, G., Evans, A., & North, A. C. T. (1978) *Nature (London) 271*, 278–281.
Jeffrey, P. D., & Coates, J. H. (1966) *Biochemistry 5*, 3820–3824.
Johnson, M. L., Correia, J. J., Yphantis, D. A., & Halvorson, H. R. (1981) *Biophys. J. 36*, 575–588.
MacLennan, A., Green, R. C., Bryant-Greenwood, G. D., Greenwood, F. C., & Seamark, R. F. (1981) *Obstet. Gynecol. 58*, 601–604.
MacLennan, A. H., Green, R. C., Grant, P., & Nicolson, R. (1986) *Obstet. Gynecol. 68*, 598–601.
Nelder, J. A., & Mead, R. (1965) *Comput. J. 7*, 308–313.
Niall, H. D., M. John, R. James, S. Kwok, R. Mercado, G. Bryant-Greenwood, R. A. Bradshaw, M. G., & Boime, I. (1980) *Insulin: Chemistry, Structure, and Function of Insulin and Related Hormones*, de Gruyter, Berlin.
Provencher, S. W., & Glockner, J. (1981) *Biochemistry 20*, 33–37.
Rawitch, A. B., Moore, W. V., & Frieden, E. H. (1980) *Int. J. Biochem. 11*, 357–362.
Roark, D. E., & Yphantis, D. Y. (1971) *Biochemistry 10*, 3241–3249.
Schwabe, C., & Braddon, E. N. (1976) *Biochem. Biophys. Res. Commun. 68*, 1126–1132.
Schwabe, C., & Harmon, S. J. (1978) *BBRC 84*, 374–380.
Sherwood, C. D., & O'Byrne, E. M. (1974) *Arch. Biochem. Biophys. 160*, 185–196.
Sherwood, O. D. (1988) *The Physiology of Reproduction*, Raven Press, New York.
Shire, S. J. (1983) *Biochemistry 22*, 2664–2671.
Shire, S. J., Foster, D. L., & Rinderknecht, E. (1988) *Biophys. J. 53*, 72a.
Steinetz, B. G., Beach, V. L., & Kroc, R. L. (1959) *Recent Progress in the Endricrinology of Reproduction*, Academic Press, New York.
Strickland, E. H. (1974) *CRC Crit. Rev. Biochem. 2*, 113–175.
Strickland, E. H., & Mercola, D. (1976) *Biochemistry 15*, 3875–3884.
Strickland, E. H., Kay, E., & Shannon, L. M. (1970) *J. Biol. Chem. 245*, 1233–1238.
Stults, J. T., Bourell, J. H., Canova-Davis, E., Ling, V. T., Laramee, G. R., Winslow, J. W., Griffin, P. R., Rinderknecht, E., & Vandlen, R. L. (1990) *Biomed. Environ. Mass Spectrom. 19*, 655–664.
Teller, D. C. (1972) *Methods Enzymol. 27*, 346–441.
Yphantis, D. A. (1964) *Biochemistry 3*, 297–317.
Wetlaufer, D. B. (1962) *Adv. Protein Chem. 17*, 303–390.

^1H and ^{15}N NMR Resonance Assignments and Preliminary Structural Characterization of *Escherichia coli* Apocytochrome b_{562}[†]

Yiqing Feng,[‡] A. Joshua Wand,*[,‡,§,∥] and Stephen G. Sligar*[,‡]
Department of Biochemistry, University of Illinois at Urbana-Champaign, Illinois 61801, and Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111
Received April 18, 1991; Revised Manuscript Received June 12, 1991

ABSTRACT: The ^1H and ^{15}N resonances of uniformly enriched apocytochrome b_{562} (106 residues) have been assigned. The assignment work began with the identification of the majority of H^N–H^α–H^β subspin systems in two-dimensional DQF-COSY and TOCSY spectra of unlabeled protein in D_2O and in 95% H_2O/5% D_2O buffer. Intraresidue and interresidue NOE connectivities were then searched for in two-dimensional homonuclear NOESY spectra recorded on unlabeled protein and in the three-dimensional NOESY-HMQC spectrum recorded on uniformly ^{15}N-enriched protein. Those data, combined with the main-chain-directed assignment strategy (MCD), led to the assignment of the main-chain and many side-chain resonances of 103 of the 106 residues. Qualitatively, the helical conformation is found to be the dominant secondary structure in apocytochrome b_{562} as it is in holocytochrome b_{562}. The helical segments in apocytochrome b_{562} overlap extensively with the helical regions defined in the crystal structure of ferricytochrome b_{562}. In addition, a number of tertiary NOEs have been identified which indicate that the global fold of the apoprotein at least partially resembles the four-helix bundle of the holoprotein. The results presented here, together with the evidence obtained with other methods [Feng and Sligar (1991) *Biochemistry* (submitted)], support the notion that the interior of the protein is fluid and may correspond to a molten globule state.

Cytochrome b_{562} found in *Escherichia coli* is a small soluble protein consisting of 106 amino acid residues (Itagaki & Hager, 1966; Lederer et al., 1981). It is classified as an electron transfer protein, although its biological partners in *E. coli* remain elusive. X-ray crystallographic studies of ferricytochrome b_{562} indicate that the molecule is composed of four nearly antiparallel helices packed to form a bundle (Mathews et al., 1979; Lederer et al., 1981). A type b heme is ligated to the polypeptide chain through methionine 7 on the N-terminal helix and histidine 102 on the C-terminal helix. This relatively simple structure makes cytochrome b_{562} an attractive model for protein-folding studies. An important task in protein-folding studies is to determine the structure of partly

[†]Supported by NIH research Grants GM 31756 (S.G.S.) and GM 35940 (A.J.W.), by NIH Instrument Grant RR-03385, by NIH Grants CA-06927 and RR-05539, by grants from the Pew Memorial Trust and the F. Ripple Foundation, and by an appropriation from the Commonwealth of Pennsylvania awarded to the Institute for Cancer Research.
* To whom correspondence should be addressed.
[‡] University of Illinois at Urbana-Champaign.
[§] Fox Chase Cancer Center.
[∥] Present address: Department of Biochemistry, University of Illinois at Urbana-Champaign, IL 61801.

0006-2960/91/0430-7711$02.50/0

folded intermediates, which will shed light on the folding process and the forces that stabilize proteins. Such intermediates are often difficult to study due to their transient nature. An increasing number of reports have emerged recently showing that partly folded structures can be trapped by modifying the covalent structure of proteins (Creighton, 1977), changing the solvent conditions (Baum et al., 1989; Jeng et al., 1990), altering the primary sequence of proteins (Oas & Kim, 1988; Hughson & Baldwin, 1989), or removing a ligand or a prosthetic group (Ikeguchi et al., 1986; Moore & Lecomte, 1990; Cocco & Lecomte, 1990; Hughson et al., 1990).

As an initial effort in our investigation of the folding and stability of cytochrome b_{562}, we have focused on the role of the heme in maintaining the integrity of the native structure. Structures of similar systems, such as apohemoglobin [e.g., Yip et al. (1972)], apomyoglobin [e.g., Griko et al. (1988) and Cocco and Lecomte (1990)], and apocytochrome b_5 [e.g., Huntley and Strittmatter (1972) and Moore and Lecomte (1990)], have been investigated with a variety of techniques. Several lines of evidence suggest the existence of significant, albeit relatively unstable, structural elements in these apoproteins. In the case of cytochrome b_{562}, we have shown that the apoprotein maintains a relatively high helical content (ca. 60% by circular dichroism) and unfolds under denaturing conditions cooperatively (Feng & Sligar, 1991). As is the case with other apoproteins, the stability of apocytochrome b_{562} is decreased relative to the holoprotein. To date, no detailed structural descriptions of these apoproteins have been available. Consequently, it remains unclear whether or how the residual structure of the apoprotein is related to the structure of the holoprotein. Here we present a high-resolution multidimensional NMR spectroscopic study of the structure of apocytochrome b_{562}. The assignments of ^1H and ^{15}N resonances of apocytochrome b_{562}, which provide the foundation reqauired to determine the solution structure of this protein, are reported. In addition, it is shown that the type and distribution of secondary structure elements remain very similar in the apoprotein as in the holoprotein. Finally, evidence is also presented that supports a "molten globule" model for the interior of apocytochrome b_{562}.

Materials and Methods

Protein Preparation. Unlabeled cytochrome b_{562} was obtained and purified from *E. coli* cells (TB-1 harboring pNS207) as previously described (Nikkila et al., 1991) except that cytochrome b_{562} was released from periplasmic space by chloroform treatment (Ames et al., 1984). The uniformly ^{15}N-labeled cytochrome b_{562} was produced from *E. coli* BL21(DE3) harboring pNS207 grown in a medium containing (per liter): 0.204 g of $MgCl_2$, 0.6 mg of $FeSO_4$, 5.46 g of KH_2PO_4, 11.4 g of K_2HPO_4, 1.5 g of NaCl, 4.0 g of glucose, 3.35 mg of δ-aminolevulinic acid, 0.2 mg of vitamin B_1, and 1.0 g of $(^{15}NH_4)_2SO_4$ (Cambridge Isotope Laboratories). About 12 mg of purified protein per liter was obtained with the BL21(DE3)/pNS207 expression system in the above minimal medium. Apocytochrome b_{562} was prepared by the methyl ethyl ketone extraction method (Teale, 1959) followed by extensive dialysis and concentration.

NMR Spectroscopy. NMR samples were approximately 1.0–4.0 mM in either 95% H_2O/5% D_2O or D_2O buffer containing 100 mM potassium phosphate and 20 μM sodium azide at pH 5.1–5.2 (uncorrected reading). Two-dimensional ^1H–^1H DQF-COSY[1] (Shaka & Freeman, 1983; Rance et al., 1984), TOCSY (mixing time 40–70 ms) (Braunschweiler & Ernst, 1983; Bax & Davis, 1985), and NOESY (mixing time 120–140 ms) (Macura & Ernst, 1980; Kumar et al., 1980) spectra of the unlabeled protein were recorded on a Bruker AM500 spectrometer at 20 °C and on a Bruker AM600 spectrometer at 15 and 25 °C. For the uniformly ^{15}N-labeled sample, two-dimensional ^1H–^{15}N HMQC (Muller, 1979; Bax et al., 1983), HMQC-TOCSY (Gronenborn et al., 1989), ^{15}N-decoupled ^1H–^1H DQF-COSY, and three-dimensional NOESY-HMQC spectra (Kay et al., 1989) were recorded on Bruker AM600 at 20 °C. Spectral widths used were 6024–7042 Hz in the ^1H dimension and 2347 Hz in the ^{15}N dimension. A recycle time of at least 1.4 s including acquisition was used for all experiments. Time-proportional phase incrementation was used to obtain quadrature detection in all incremented time domains (Redfield & Kuntz, 1975; Marion & Wüthrich, 1983). Solvent suppression was achieved by direct on-resonance presaturation in all experiments. The data were processed with FTNMR and FELIX (Hare Research, Woodinville, WA).

Typically, 512–700 t_1 increments with 1024 complex points in t_2 were collected for the 2D homonuclear experiments. The data were zero-filled once in t_2 and twice in t_1 to achieve a digital resolution of 2.9–3.4 Hz/point. For NOESY and TOCSY data, a Gaussian-Lorentzian filter function was used in t_2 and a shifted sine function was used in t_1. For DQF-COSY spectra, unshifted sine and exponential digital filters were used in both dimensions.

Typically, 700 t_1 and 300 t_1 increments were collected in 2D HMQC and HMQC-TOCSY experiments, respectively. A delay of 5.3 ms was used as nominal $1/(2J_{NH})$ time period. A Gaussian–Lorentzian apodization function was used in t_2, and a shifted sine function was used in t_1. The spectra have a final digital resolution of 6.9 Hz/point in the ^1H dimension and 2.3 Hz/point in the ^{15}N dimension (zero-filled). Three-dimensional NOESY-HMQC spectra were obtained essentially as described by Kay et al. (1989) except that time-proportional phase incrementation was used to provide quadrature detection during the incremented time domains, ^{15}N nuclei were decoupled from protons during the ^1H incremented time domain by an inversion pulse rather than by composite decoupling, and solvent elimination was achieved by direct rather than by off-resonance presaturation. The primary data set of the 3D spectrum show here consisted of 128 points in the incremented ^1H dimension and 86 points in the ^{15}N dimension with 256 complex points collected during acquisition. The data was processed to 256 × 128 × 256 giving a final resolution of 27.5 Hz/point in both ^1H dimensions and 18.3 Hz/point in the ^{15}N dimension.

Results

H^N–H^α Fingerprint and H^N–H^α–H^β Subspin System Identification. The fingerprint region of the DQF-COSY spectrum of apocytochrome b_{562} is shown in Figure 1. This region contains the backbone H^N–H^α cross peaks as well as cross peaks arising from some labile side-chain hydrogens. A total of 107 cross peaks were identified in this region when the spectra recorded in D_2O buffer and in H_2O buffer at different temperatures were compared. Cytochrome b_{562} contains 106 amino acid residues, including three glycines and

[1] Abbreviations: MCD, main chain directed; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; DQF-COSY, double-quantum-filtered correlation spectroscopy; TOCSY, total correlation spectroscopy; HMQC, heteronuclear multiple-quantum COSY.

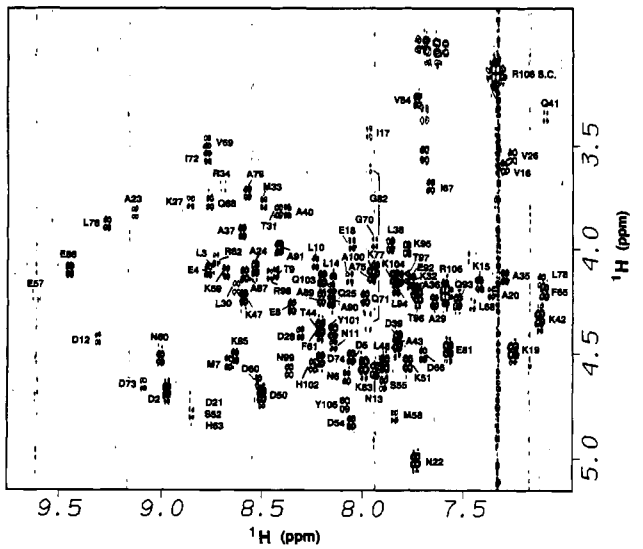

FIGURE 1: Fingerprint region of the DQF-COSY spectrum of unlabeled apocytochrome b_{562} in 95% H_2O/5% D_2O buffer at 20 °C (500-MHz proton frequency). Assigned cross peaks are labeled. Assigned side-chain resonances are indicated by S.C.

four prolines. A total of 104 cross peaks are therefore expected from the protein backbone (the N-terminus amino group is usually unobservable). At the conclusion of the assignment work, all amide proton resonances but those of A1 and E49 were assigned, which account for 100 cross peaks in this region. Most of the remaining cross peaks, arising from labile side-chain hydrogens, were not assigned to specific residues.

The DQF-COSY and TOCSY spectra, recorded under a variety of conditions, allowed the identification of amide H^N–H^α–H^β subspin systems [NAB sets, Englander and Wand (1987) and Wand and Nelson (1991)] for 75 of the 99 residues expected. The NAB sets for the remaining residues, for which the H^α–H^β connectivities remained ambiguous, were completed by reference to the NOESY spectra. Three residues (at positions 21, 52, and 63) presented problems because their H^N–H^α resonances frequencies are similar and also the H^α resonance frequencies coincide with the solvent resonance frequency. These degeneracies were not removed by recording the spectrum at different temperatures (i.e., 15–25 °C). Nevertheless, the presence of these residues could be established by H^α–H^β correlations in the DQF-COSY spectrum recorded in D_2O buffer and by H^N–H^β correlations in the TOCSY spectra recorded in H_2O buffer.

Spin System Identification. Due to the complexity of the spectra of apocytochrome b_{562}, only simple spin systems could be unambiguously identified at the initial stage of the assignment. These include alanines, glycines, threonines, and valines. Cytochrome b_{562} contains 17 alanines, 5 threonines, 4 valines, and 3 glycines. A total of 16 alanines were identified in the TOCSY spectra. All five of the threonine spin systems displayed correlation of the complete spin system to the amide proton in the TOCSY spectra, making their identification straightforward. Three of the four valines displayed cross peaks from the β proton and from both γ methyls to the amide protons in the TOCSY spectra. The side chain of the remaining valine was unambiguously identified in the upfield region of the DQF-COSY and TOCSY spectra, and the amide proton was traced to in the DQF-COSY spectra. The search for glycines was less successful. Among the three glycines, only one displayed a pair of cross peaks characteristic of glycines in the fingerprint region. A second glycine, identified at a later stage, has nearly degenerate α proton resonances.

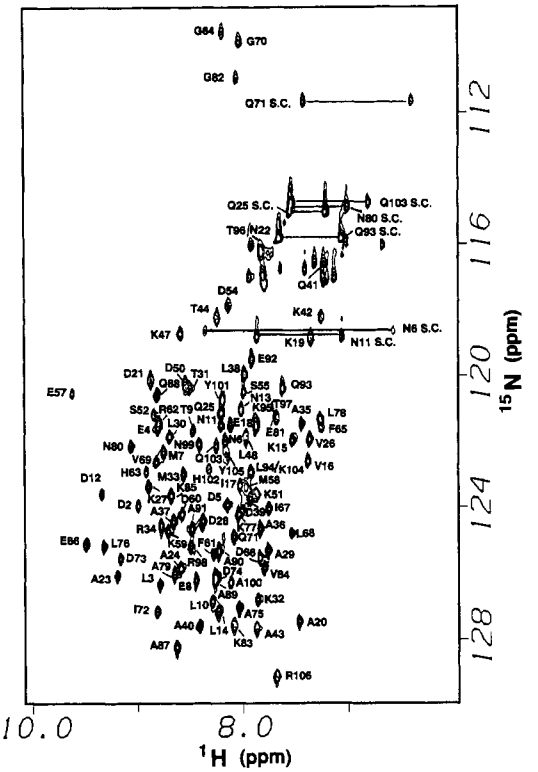

FIGURE 2: Section of the HMQC spectrum of uniformly ^{15}N-labeled apocytochrome b_{562} in H_2O buffer (600-MHz proton frequency). Assigned H^N–^{15}N cross peaks are indicated. For assigned side-chain $^{15}NH_2$ groups, the two peaks with a common ^{15}N chemical shift are connected by the horizontal lines.

A third glycine was observed only in the TOCSY spectra. A pair of glycine-like cross peaks in the fingerprint region was found to arise from an unassigned side chain. Their identity as side-chain resonances was confirmed by the extremely upfield ^{15}N chemical shift (88.5 ppm).

^1H–^{15}N Correlation Spectra of Uniformly ^{15}N-Labeled Protein. The ^{1}H–^{15}N HMQC spectrum of uniformly ^{15}N-labeled apocytochrome b_{562} is presented in Figure 2. Correlations between ^{15}N and directly bonded hydrogens are selected in this experiment. This spectrum, which provides correlation between ^{15}NH pairs, serves as a bridge between the 2D homonuclear spectra and the 3D NOESY-HMQC spectrum. In order to identify the ^{15}N resonance frequency associated with each of the cross peaks in Figure 1, a matching of the cross peaks in the HMQC spectrum (H^N–^{15}N) and those in the fingerprint region of the DQF-COSY spectrum (H^N–H^α) was carried out via two steps. First, an ^{15}N-decoupled DQF-COSY spectrum was recorded on the labeled sample, and the amide proton chemical shifts were compared with those in the HMQC spectrum. This procedure offers the resolution of one dimension only and often leaves ambiguities due to degenerate H^N frequencies. The ambiguities were partly resolved in the second step by appealing to the 2D HMQC-TOCSY spectrum. Under the spin-lock time used here (27.4 ms), the HMQC-TOCSY experiment manifests long-range correlations between H^α–^{15}N pairs, and in some cases between H^β–^{15}N pairs, which allow the matching to be carried out with a second frequency or even a third frequency. In the case of apocytochrome b_{562}, 88 residues were unambiguously identified in the two steps described above. The remaining H^N–^{15}N cross peaks were identified through subsequent examination of the 2D NOESY and 3D NOESY-HMQC spectra.

Three-dimensional NOESY-HMQC spectra were obtained

Table I: *E. coli* Apocytochrome b_{562} Proton and Nitrogen Resonance Assignments[a]

residue	^{15}N	H^N	H^α	H^β	others	residue	^{15}N	H^N	H^α	H^β	others
A1			4.18	1.57		P56			4.00		
D2	124.0	8.98	4.68	3.04, 2.91		E57	120.5	9.62	4.22	2.08	H^γ 2.47
L3	126.4	8.76	4.09	1.92, 1.62	H^γ 1.73; H^δ 1.04, 1.18	M58	123.3	7.84	4.80		
E4	121.6	8.77	4.10	2.05, 2.52		K59	124.7	8.68	4.11	1.97, 1.68	
D5	123.9	8.06	4.52	2.83		D60	124.2	8.52	4.64	2.94	
N6	121.9	8.08	4.61	2.77, 2.61	H^γ 6.37, 8.42; $^{15}N^\gamma$ 118.5	F61	125.4	8.23	4.40	3.29, 3.49	$H^{2,6}$ 7.23; $H^{3,5}$ 7.49; H^4 7.29
M7	122.3	8.67	4.55	2.41		R62	121.3	8.73	4.05	2.19, 2.39	H^γ 2.05
E8	126.3	8.35	4.28	2.33, 2.57		H63	122.9	8.85	4.71	3.57, 3.62	H^2 8.77; H^5 7.49
T9	121.6	8.43	4.12	4.46	H^γ 1.20	G64	109.5	8.10	3.64, 3.30		
L10	126.8	8.24	4.06	2.23, 1.98	H^δ 1.01	F65	121.2	7.10	4.20	2.80, 3.02	$H^{2,6}$ 6.94; $H^{3,5}$ 6.69; H^4 7.04
N11	121.5	8.14	4.46	2.90, 3.02	H^γ 7.90, 7.09; $^{15}N^\gamma$ 118.8	D66	125.5	7.70	4.50	2.76, 2.89	
D12	123.6	9.32	4.43	2.90, 2.75		I67	124.0	7.66	3.69	1.85	mH^γ 0.84; H^γ 1.20, 0.95; H^δ 0.70
N13	120.9	7.94	4.58	2.71, 2.58							
L14	127.1	8.15	4.15	1.87, 1.76		L68	124.8	7.46	4.26	1.72, 1.82	H^δ 0.93
K15	122.0	7.42	4.17	2.07, 1.94		V69	122.6	8.78	3.48	2.25	H^γ 1.12 0.99
V16	122.6	7.30	3.60	2.47	H^γ 1.15, 0.99	G70	109.7	7.94	3.94, 3.96		
I17	123.2	7.97	3.43	2.17	mH^γ 1.01; H^δ 0.89	Q71	124.8	7.99	4.23	2.52, 1.99	H^γ 2.38; H^δ 7.37, 6.26; $^{15}N^\delta$ 111.5
E18	121.5	8.05	3.98	2.34							
K19	118.9	7.26	4.48	1.82, 2.04	H^γ 1.57, 1.71	I72	127.2	8.77	3.56	2.13	mH^γ 1.01; H^δ 0.75
A20	127.5	7.37	4.20	1.52		D73	125.6	9.09	4.65	2.96, 2.65	
D21	120.1	8.83	4.85	2.87, 2.73		D74	126.2	8.21	4.53	2.90, 2.73	
N22	116.4	7.73	5.01	3.04, 2.95		A75	127.1	7.96	4.13	1.54	
A23	126.1	9.13	3.83	1.55		L76	125.2	9.27	3.88	2.04, 1.70	H^γ 1.34; H^δ 0.99, 0.74
A24	125.8	8.53	4.09	1.54		K77	124.2	7.94	4.09	2.05, 1.64	H^γ 1.77
Q25	121.1	8.16	4.19	2.40	H^γ 1.95, 2.67; H^δ 7.47, 7.16; $^{15}N^\delta$ 114.9	L78	121.3	7.12	4.15	2.05	H^δ 0.83
						A79	125.9	8.58	3.73	1.40	
V26	121.9	7.26	3.55	2.22	H^γ 1.11, 0.94	N80	122.1	9.01	4.52	3.04, 2.88	H^γ 7.46, 6.95; $^{15}N^\gamma$ 114.7
K27	123.3	8.86	3.78	2.03, 1.83	H^γ 1.37, 1.34; H^δ 1.69; H^ϵ 2.97	E81	121.3	7.57	4.48	2.26, 2.68	H^γ 2.48
						G82	110.8	7.98	3.61, 4.34		
D28	124.5	8.31	4.40	2.85, 2.73		K83	127.6	7.99	4.55	1.83	H^γ 1.35
A29	125.3	7.64	4.25	1.63		V84	125.8	7.73	3.28	2.10	H^γ 0.92, 0.80
L30	121.8	8.63	4.18	2.14	H^γ 1.35; H^δ 0.74, 0.99	K85	123.6	8.64	4.51	1.94, 1.83	
T31	120.3	8.42	3.82	4.53	H^γ 1.34	E86	125.1	9.45	4.10	2.13, 2.09	H^γ 2.38, 2.58
K32	126.8	7.76	4.15	1.68, 1.78		A87	128.2	8.59	4.12	1.52	
M33	122.9	8.50	3.78	2.45	H^γ 2.21[b]	Q88	120.5	8.77	3.78	2.54	
R34	124.6	8.70	3.70	2.08	H^γ 1.95[b]	A89	126.2	8.21	4.23	1.56	
A35	121.4	7.30	4.13	1.56		A90	125.3	8.16	4.20	1.55	
A36	124.6	7.74	4.20	1.56		A91	124.6	8.42	4.00	1.54	
A37	124.4	8.60	3.92	1.54		E92	119.6	7.80	4.14		
L38	119.9	7.86	3.98	1.88	H^δ 0.94	Q93	120.3	7.53	4.25	2.16, 2.29	H^γ 2.51, 2.57; H^δ 7.61, 6.99; $^{15}N^\delta$ 116.0
D39	123.7	7.82	4.42	2.68, 2.73							
A40	127.6	8.38	3.82	0.94		L94	122.9	7.84	4.19	2.01	
Q41	116.9	7.09	3.36	2.20, 1.79	H^γ 2.48	K95	121.6	7.78	4.00	1.98	
K42	118.1	7.12	4.33	1.91, 2.08	H^γ 1.54, 1.73[b]	T96	116.4	7.72	4.22	4.44	H^γ 1.28
A43	127.7	7.83	4.45	1.61		T97	121.6	7.82	4.12	4.24	H^γ 1.16
T44	118.1	8.21	4.37	3.95	H^γ 1.14	R98	125.1	8.46	4.12	1.97, 1.82	
P45						N99	122.0	8.37	4.58	2.92	
P46			4.35	2.44, 2.03	H^γ 2.16, H^δ 3.81	A100	126.3	8.07	4.14	1.30	
K47	118.6	8.59	4.23	1.94, 2.00		Y101	120.6	8.15	4.39	3.18, 2.92	$H^{2,6}$ 7.16, $H^{3,5}$ 6.74
L48	121.8	7.89	4.55	1.69	H^γ 1.46; H^δ 0.83, 0.72	H102	122.8	8.25	4.56	3.42, 3.38	H^2 8.44, H^5 7.34
E49			3.98			Q103	121.9	8.20	4.14	2.07	H^γ 2.37; H^δ 6.63, 7.44; $^{15}N^\delta$ 114.7
D50	120.2	8.50	4.70	2.83							
K51	123.6	7.77	4.55	2.02		K104	122.9	7.84	4.13	1.49	H^γ 1.21; H^δ 1.58; H^ϵ 2.90
S52	121.1	8.85	4.79	4.23, 4.01		Y105	122.1	8.09	4.74	3.28, 2.85	$H^{2,6}$ 7.21; $H^{3,5}$ 6.91
P53						R106	129.2	7.59	4.21	1.74, 1.61	H^γ 1.81; H^δ 3.15; H^ϵ 7.35; $^{15}N^\epsilon$ 89.1
D54	117.8	8.05	4.83	2.84, 2.66							
S55	120.5	7.90	4.64								

[a] Chemical shifts for apocytochrome b_{562} in 100 mM KPO_4 at 20 °C, pH 5.1–5.2. Proton chemical shifts referenced to internal [^{2}H]TSP [sodium 3-(trimethylsilyl)[2,2,3,3-$^2H_4$]propionate] at 0.00 ppm. Nitrogen-15 chemical shifts are referenced to external $^{15}NH_4Cl$ at 24.93 ppm. Methyl groups are not indicated explicitly except for isoleucine where ambiguity exists. [b] Tentative assignment to a specific proton in the spin system.

during the later stages of the assignment and therefore served to resolve ambiguities arising from degeneracies of H^N resonances in two-dimensional ^1H–^1H NOESY spectra. Although the aliphatic region of the ^1H spectrum of apocytochrome b_{562} shows less dispersion than is expected from a structured protein (see Discussion, Table I), the amide ^1H and ^{15}N spectra remain relatively resolved. Accordingly, the full benefit of using chemical shift correlations of amide H^N NOEs with their associated amide ^{15}N chemical shift is maintained in apocytochrome b_{562}. The additional resolution provided served to reveal many main-chain NOEs involving amide hydrogens that would have otherwise been obscured by degeneracy. Examples are given below.

Sequence-Specific Residue Assignments. The assignment began by searching for groups of NAB sets that displayed specific MCD NOE patterns (Englander & Wand, 1987; Wand & Nelson, 1991). The 16 identified alanines, along with the 5 threonines and 4 valines, were used as markers to place the segments in the primary sequence. Figure 3 is a summary of the main-chain NOE connectivities identified in the 2D NOESY spectra (15, 20, and 25 °C) and/or in the 3D NOESY-HMQC spectrum (20 °C). The residue-specific assignments are listed in Table I. The following is a description of the crucial aspects of the assignment process.

(*A*) *Residues 1–20.* The assignment of this region proceeded by placement of a 17-residue MCD-defined helical

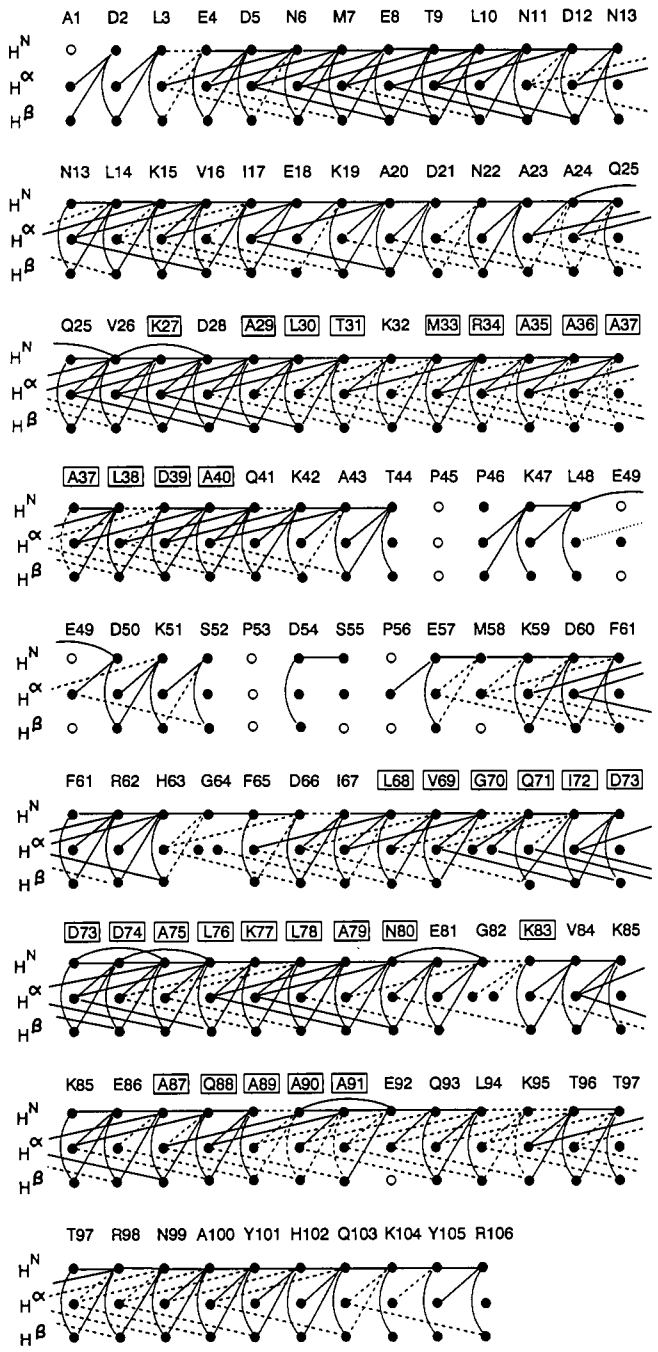

FIGURE 3: Summary of NOE connectivities observed in 2D and 3D spectra for backbone and β-protons at 20 and 25 °C. Open circles indicates unassigned resonances. Solid lines indicate resolved and unambiguously identified NOEs. Dashed lines indicate ambiguous NOEs due to either degeneracy or proximity of the involved resonance to the solvent signal. For prolines, the C_δ protons are in the position of the amide protons. Residues that exhibit relatively slow hydrogen-exchange rates are highlighted by boxes (see text).

segment that bridged threonine 9, valine 16, and alanine 20. The degeneracy of the H^N–H^α resonance shifts of residues 3 and 4 prevented further extension of the assignment to the N-terminus of the protein at this point. Residues 1–3 were assigned at the final stage, when assignment of the protein was largely complete.

(*B*) *Residues 21–44.* The presence of 8 of the 16 identified alanines, in particular the presence of three sequential alanines (Ala-Ala-Ala) in this segment, made it the first segment to be assigned within the primary sequence of apocytochrome b_{562}. The three alanine repeat occurs twice in the primary sequence at residues 35–37 and residues 89–91. The current Ala-Ala-Ala was assigned to residues 35–37 on the basis of the absence of an expected alanine at position 87. The MCD extension from residue 35 to residue 22 was staightforward and was confirmed by alanines at positions 23, 24, and 29, a valine at position 26, and a threonine at position 31. The ambiguity of H^N NOEs arising from the triple degeneracy of D21, S52, and H63 was resolved in the 3D NOESY-HMQC spectrum. The extension from residue 37 to residue 44 was partially hampered by the near degeneracy of the amide protons of leucine 38 and aspartate 39, those of glutamine 41 and lysine 42, and those of threonine 44 and phenylalanine 61. In the first two cases, the difficulties were overcome by observing the medium-range NOEs to these residues (Wand & Nelson, 1991), while the 3D approach proved useful in the third case. The assignment for residues 42–44 was further confirmed by observing Ala-Thr, which is unique in the primary sequence. Figure 4 presents the main-chain NOE connectivities for residues 24–30 in the 3D NOESY-HMQC spectrum.

(*C*) *Residues 46–55.* This segment was the most difficult to assign due to a combination of the lack of extensive NOE connectivities and the lack of easily recognizable spin systems. It was assigned at the late stage of the assignment process when the number of remaining residues was limited. The remaining NAB sets, corresponding to residues 1–3, 47–52, and 54–55, were classified to two categories based on the H^β chemical shift values (Bundi & Wüthrich, 1979). The main-chain NOE connectivities further group them into several segments and allowed the assignment of this region to be completed.

(*D*) *Residues 56–82.* The alignment of this segment within this region of the primary sequence was initiated at residues 75–79, with residues 75 and 79 being alanines. A survey over the primary sequence shows that this is the only Ala-X-X-X-Ala fragment present in apocytochrome b_{562}. The assignment was confirmed by an expected glycine residue at position 82. Extending to the N-terminus revealed glycine 70 with nearly degenerate α proton resonances. The assignment for this glycine was made unambiguous because of the presence of a neighboring valine (at position 69). The remaining glycine at position 64, which is absent in DQF-COSY spectra but present in TOCSY spectra, was revealed through the NOE connectivity search.

(*E*) *Residues 83–106.* Although this segment contains the other Ala-Ala-Ala fragment at positions 89–91, no such fragment could be identified initially. Rather, two fragments Ala-X-Ala and Ala-Ala were found where the third alanine in the first fragment and the first alanine in the second one share similar H^N chemical shifts and essentially identical H^α and H^β chemical shifts. Since the sequence of Ala-Ala at this stage of the assignment can only occur at position 89–90 or 90–91 and Ala-X-Ala can only occur at position 87–89, these two fragments were assigned to residues 87–91, assuming that alanine 89 and 90 are degenerate. The observation of a valine at position 84 confirmed the assignments. The NOE connectivity search toward the C-terminus was stopped at leucine 94 but could be resumed at threonines 96 and 97, the only Thr-Thr in the sequence, and was soon confirmed by alanine 100. Lysine 95 was assigned, by induction, at the last stage of the assignment, assuming the extensive degeneracy. The assignments at the C-terminus were further confirmed by NOE connectivities between the main-chain resonances and the tyrosine side-chain resonances.

Side-Chain Assignments. Cytochrome b_{562} contains two phenylalanines, two tyrosines, and two histidines. The phenylalanines and tyrosines are all at the fast-flipping limit so

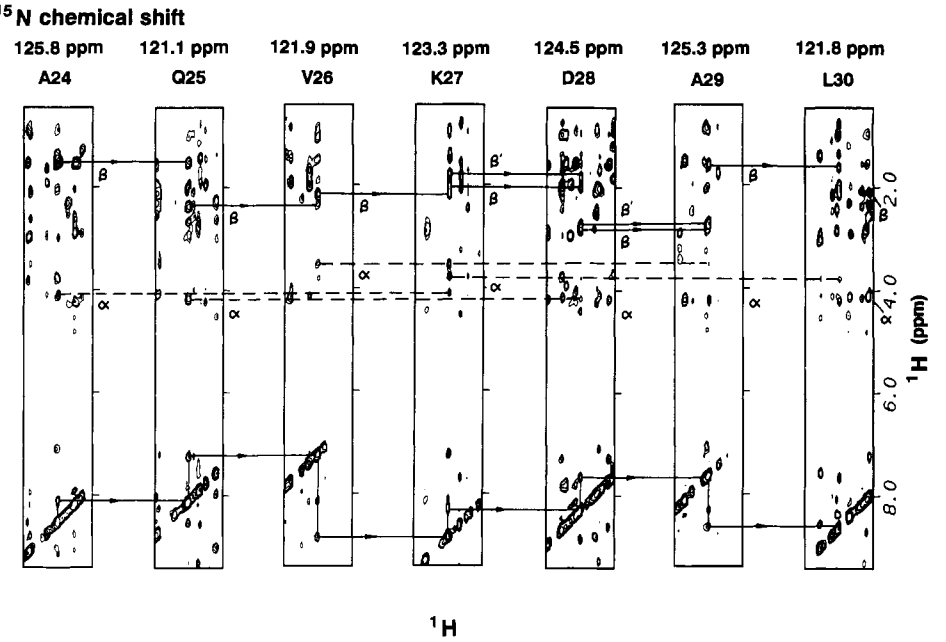

FIGURE 4: Selected 2D ^{1}H–^{1}H NOESY planes from the 3D NOESY-HMQC spectrum (600-MHz proton frequency) of uniformly ^{15}N-enriched apocytochrome b_{562}. Main-chain NOE connectivities for the helical segment between residue 24 and 30 are indicated. Solid lines and dashed lines are used for short-range and medium-range NOEs, respectively.

they present only motion-averaged resonances. The sequence-specific assignment of these aromatic side-chain resonances became straightforward when the main-chain resonances were assigned. Assignments for phenylalanines and tyrosines were based on the intensity of the NOEs between the side chain and the β protons. For both histidines, the NOE connectivities between the β protons and H^5 were observed, and the side chains were assigned accordingly. The aliphatic resonances in apocytochrome b_{562} display only limited chemical shift dispersion (see Discussion). The assignments for aliphatic side chains listed in Table I are based mainly on TOCSY correlations and occasionally on NOE connectivities. In only rare cases can an entire long aliphatic side chain spin system be identified. Each $^{15}N_2$ group in asparagines and glutamines displays two cross peaks in the HMQC spectrum with a common ^{15}N chemical shift. Seven of twelve such groups in apocytochrome b_{562} have been assigned to specific residues on the basis of the NOEs between the $^{15}NH_2$ group and the H^β (asparagines) or H^γ (glutamines) resonances.

Tertiary NOEs. Given the main-chain and many side-chain assignments, a number of long-range NOEs can be identified unambiguously. For example, the aromatic side chain of F65 exhibits NOEs to A37, A40, and Q41, and the methyl of A29 displays NOEs to N13, V16, and I17. The majority of the tertiary NOEs identified thus far are located between the first and the second helices, or between the second and the third helices, indicating the close spatial proximity of these pairs of helices.

Amide Hydrogen Exchange. Hydrogen-exchange behavior of the amide protons serves as a useful tool to probe the structure and dynamics of proteins, in addition to the utility of simplifying spectra. A total of 31 backbone amide protons are present in the 2D ^{1}H–^{1}H spectra of apocytochrome b_{562} freshly transferred into D_2O buffer at pD 5.1 and 20 °C, as a result of their relatively slow hydrogen-exchange rates. These residues are indicated in Figure 3.

Discussion

The essentially complete assignment of the ^{1}H and ^{15}N resonances arising from the backbone of apocytochrome b_{562}, along with the assignment of many side chains, has been described. The application of two- and three-dimensional NMR spectroscopy to this problem has allowed the resonance assignments reported here to be firmly established. Accordingly, a qualitative description of the secondary structure can be given based on the pattern of NOE connectivities documented in Figure 3. The NOE pattern reveals four helical segments in the regions spanned by residues 3–18, 23–43, 59–80, and 84–103. Although many of the medium range $H^\alpha(i)$–$H^N(i+3)$ and $H^\alpha(i)$–$H^\beta(i+3)$ NOEs characteristic of the helical conformation are obscured in residues 91–103, due to the extensive degeneracy of the resonances involved, further support for the presence of a helical-like conformation in this region is provided by the small magnitude of the $^3J_{HN\alpha}$ coupling constants (unpublished data). The placement of helices in apocytochrome b_{562} corresponds closely to that observed in the holoprotein (residues 3–19, 23–40, 56–80, and 84–105; Mathews, personal communication). These results demonstrate that the apoprotein is able to adopt a structure having similar content and placement of secondary structure elements as the holoprotein. The amide hydrogens that exchange relatively slowly with the solvent are found in all the helices but the first one, suggesting these three helices are reasonably stable. In the crystal structure of ferricytochrome b_{562}, the heme makes extensive contacts with the first, the fourth, and, to a lesser extent, the third helix of the four-helix bundle. The presence of a significant number of NOEs between residues of the first and the second helices as well as between residues of the second and the third helices suggests that a significant fraction of the tertiary structure of the holoprotein is retained in the apoprotein.

Though the majority of side-chain spin systems have been completely assigned, a number remain incompletely characterized. This is due to the fact that the chemical shift dispersion observed for aliphatic hydrogens is poor. In general, the chemical shifts of aliphatic side-chain hydrogens of apocytochrome b_{562} tend to cluster much more closely to their values in unstructured peptides than is usually seen in stable globular proteins (e.g., ubiquitin, plastocyanin, thioredoxin, and hen egg white lysozyme). In contrast, both amide nitrogen

and hydrogen spectra are reasonably well resolved and show dispersion typical of a protein the size of apocytochrome b_{562}. In addition, the aromatic rings of Phe 61 and Phe 65, which are tightly packed against the heme prosthetic group in the holoprotein, display completely averaged ortho and meta hydrogens. This is indicative of free rotation of the rings in the interior of the protein. These observations, coupled with the essentially identical secondary structure between the apoprotein and the holoprotein, lead to the suggestion that, in the absence of the heme, the interior of apocytochrome b_{562} is indeed fluid or "molten". This phenomenon presumably arises from the absence of suitable packing constraints, normally provided by the heme prosthetic group. The presence of numerous medium-range periodic NOEs among the backbone amide H^N, H^α, and H^β hydrogens within helical regions of the molecule are consistent with tight and regular helical structure (Wüthrich et al., 1984; Wand & Nelson, 1991) and argue against significant disorder of the backbgone of the protein. Together with a smaller enthalpy change accompanying the unfolding of the apoprotein than the holoprotein, and the slightly larger molecular dimension of the apoprotein than the holoprotein (Feng & Sligar, 1991), apocytochrome b_{562} displays many of the characteristics associated with the so-called "molten globule" or the A-state [reviewed by Ptitsyn (1987) and Kuwajima (1989)]. What is unclear at this time is whether the dynamic averaging of interior side-chain resonances arises from fluctuations in tertiary structure or from a general expansion of the molecle that is sufficient to allow free mobility within the core of the protein. The library of resonance assignments presented here should allow the tertiary structure of the molecule to be determined by distance geometry and related methods and therefore shed light on this question. Such studies are in progress and will be presented elsewhere.

Acknowledgments

We thank Dr. S. Walter Englander for generous access to the Bruker AM500 spectrometer and for helpful discussions. We also thank Sharon Roth and Dr. Diane Schneider for assistance with data collection on the Bruker AM600 spectrometer and Xiurong Qi for comparison of chemical shift distributions.

Supplementary Material Available

Three figures showing spin system identification in ^1H–^1H TOCSY spectra, H^α–^{15}N correlations in HMQC-TOCSY spectra, and ^1H–^1H NOESY spectrum with main-chain NOEs from 9 to 20 indicated (4 pages). Ordering information is given on any current masthead page.

References

Ames, G. F.-L., Prody, C., & Kustu, S. (1984) *J. Bacteriol. 160*, 1181–1183.

Baum, J., Dobson, C. M., Evans, P. A., & Hanley, C. (1989) *Biochemistry 28*, 7–13.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Bax, A., Griffey, R. H., & Hawkins, B. L. (1983) *J. Am. Chem. Soc. 105*, 7188–7190.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.

Bundi, A., & Wüthrich, K. (1979) *Biopolymers 18*, 285–298.

Cocco, M. J., & Lecomte, J. T. J. (1990) *Biochemistry 29*, 11067–11072.

Creighton, T. E. (1977) *J. Mol. Biol. 113*, 275–293.

Dolgikh, D. A., Abaturov, L. V., Bolotina, I. A., Brazhnikov, E. V., Bychkova, V. E., Bushuev, V. N., Gilmanshin, R. I., Lebedev, Y. O., Semisotnov, G. V., Tiktopulo, E. I., & Ptitsyn, O. B. (1985) *Eur. Biophys. J. 13*, 109–121.

Englander, S. W., & Wand, A. J. (1987) *Biochemistry 26*, 5953–5958.

Feng, Y., & Sligar, S. G. (1991) *Biochemistry* (submitted).

Griko, Y. V., Privalov, P. L., Venyaminov, S. Y., & Kutyshenko, V. P. (1988) *J. Mol. Biol. 202*, 127–138.

Gronenborn, A. M., Bax, A., Wingfield, P. T., & Clore, G. M. (1989) *FEBS Lett. 243*, 93–98.

Hughson, F. M., & Baldwin, R. L. (1989) *Biochemistry 28*, 4415–4422.

Hughson, F. M., Wright, P. E., & Baldwin, R. L. (1990) *Science 249*, 1544–1548.

Huntley, T. E., & Strittmatter, P. (1972) *J. Biol. Chem. 247*, 4641–4647.

Ikeguchi, M., Kuwajima, K., & Sugai, S. (1986) *J. Biochem. 99*, 1191–1201.

Itagaki, E., & Hager, L. P. (1966) *J. Biol. Chem. 241*, 3687–3695.

Jeng, M., Englander, S. W., Elove, G. A., Wand, A. J., & Roder, H. (1990) *Biochemistry 29*, 10433–10437.

Kay, L. E., Marion, D., & Bax, A. (1989) *J. Magn. Reson. 84*, 72–84.

Kumar, A., Wüthrich, K., & Ernst, R. R. (1980) *Biochem. Biophys. Res. Commun. 95*, 1–6.

Kuwajima, K. (1989) *Proteins 6*, 87–103.

Lederer, F., Glatigny, A., Bethge, P. H., Bellamy, H. D., & Mathews, F. S. (1981) *J. Mol. Biol. 148*, 427–448.

Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95–117.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Mathews, F. S., Bethge, P. H., & Czerwinski, E. W. (1979) *J. Biol. Chem. 254*, 1699–1706.

Moore, C. D., & Lecomte, J. T. J. (1990) *Biochemistry 29*, 1984–1989.

Muller, L. (1979) *J. Am. Chem. Soc. 101*, 4481–4484.

Nikkila, H., Gennis, R., & Sligar, S. G. (1991) *Eur. J. Biochem.* (in press).

Oas, T. G., & Kim, P. S. (1988) *Nature 336*, 42–48.

Ptitsyn, O. B. (1987) *J. Protein Chem. 6*, 273–293.

Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1984) *Biochem. Biophys. Res. Commun. 117*, 479–485.

Redfield, A. G., & Kuntz, S. D. (1975) *J. Magn. Reson. 19*, 250–254.

Shaka, A. J., & Freeman, R. (1983) *J. Magn. Reson. 51*, 169–173.

Teale, F. W. J. (1959) *Biochim. Biophys. Acta. 35*, 543.

Wand, A. J., & Nelson, S. J. (1991) *Biophys. J. 59*, 1101–1112.

Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715–740.

Yip, Y. K., Waks, M., & Beychok, S. (1972) *J. Biol. Chem. 247*, 7237–7244.