Proton Resonance Assignments of Horse Ferricytochrome *c*†

Yiqing Feng, Heinrich Roder,* and S. Walter Englander
Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104
A. Joshua Wand and Deena L. Di Stefano
Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111
Received July 13, 1988; Revised Manuscript Received September 20, 1988

ABSTRACT: Two-dimensional nuclear magnetic resonance spectroscopy (2D NMR) was used to obtain extensive resonance assignments in the ^{1}H NMR spectrum of horse ferricytochrome *c*. Assignments were made for the main-chain and C_β protons of 102 residues (all except Pro-44 and Gly-84) and the majority of side-chain protons. As starting points for the assignment of the oxidized protein, a limited set of protons was initially assigned by use of 2D NMR magnetization transfer methods to correlate resonances in the oxidized form with assigned resonances in the reduced form [Wand, A. J., Di Stefano, D. L., Feng, Y., Roder, H., & Englander, S. W. (1989) *Biochemistry* (preceding paper in this issue)]. Given the complexity of the spectrum due to the size of this protein (104 residues) and its paramagentic center, the initial search for side-chain spin systems in *J*-correlated spectra was successful only for the simplest side chains, but the majority of NH–C_αH–C_βH subspin systems (NAB sets) could be identified at this stage. The subsequent search for sequential NOE connectivities focused on NAB sets, with use of previously assigned residues to place NOE-connected segments within the amino acid sequence. Selective proton labeling of either the slowly or the rapidly exchanging amide sites was used to simplify the spectra, and systematic work at two temperatures was used to resolve ambiguities in the 2D NMR spectra. These approaches, together with the use of magnetization transfer methods to correlate reduced and oxidized cytochrome *c* spectra, provide multiple cross-checks to verify assignments.

Much effort has been devoted to studies of structure–function relationships in cytochrome *c* [e.g., Margoliash and Boshard (1983), Bechtold et al. (1986), Mayo et al. (1986), Pielak et al. (1987), and Dickerson et al. (1976)]. This wealth of information together with the small size and interesting functional interactions of cytochrome *c* makes it an attractive protein for NMR[1] studies. A major obstacle in any detailed NMR investigation of a macromolecule is the assignment of the resonances. Since the early studies of McDonald and Philips (1969) and Redfield and Gupta (1971), ^{1}H NMR assignments have been reported for parts of some 40 of the 104 amino acid residues of horse ferricytochrome *c*, relying mainly on one-dimensional NMR techniques [Williams et al. (1985) and references cited therein]. The majority of reported assignments are for side-chain resonances, and many of these were based on the assumption that the solution structure of horse cytochrome *c* is similar to the crystal structure of the tuna protein.

This paper presents comprehensive proton resonance assignments of horse ferricytochrome *c* obtained by two-dimensional NMR methods, without reliance on crystallographic information. This work was begun during the course of assignment studies on reduced cytochrome *c*, described earlier (Wand & Englander, 1985, 1986; Wand et al., 1989). The ^{1}H NMR spectra of reduced and oxidized cytochrome *c* show little resemblance, primarily due to the presence of the paramagnetic iron in the oxidized form. Therefore, few assignments can be made simply by comparing the spectra of the two oxidation states, and we were faced with an essentially independent assignment problem. The set of assignments reported here includes the main-chain protons of 102 residues (all except Pro-44 and Gly-84), as well as numerous complete side-chain spin systems. One-dimensional NMR work done in parallel, which led to additional assignments for paramagnetically shifted resonances, will be presented elsewhere.[2] Together with the results for reduced cytochrome *c* (Wand et al., 1989), the present data now provide extensive proton resonance assignments for the protein in two different functional forms. The confidence in both sets of assignments, obtained independently in each oxidation state, is enhanced by the capability to correlate the two data sets by means of two-dimensional magnetization transfer methods.[3]

These assignments provide the basis for proton-resolved studies of properties such as the solution structure, internal dynamics, and hydrogen exchange behavior, and how these properties are affected by the redox transition and other perturbations. The present proton assignments will furthermore open new possibilities in our ongoing investigation of the cytochrome *c* folding process (Roder, 1988; Roder et al., 1988).

Materials and Methods

Horse heart cytochrome *c* (type VI) was obtained from Sigma Chemical Co. To prepare NMR samples, the protein was dissolved in H_2O or D_2O buffer, a small amount of po-

† This work was supported by NIH Research Grants GM-31847 and DK-11295 (S.W.E.), GM-35926 (H.R.), and GM-35940 (A.J.W.), by Instrumentation Grants RR-2497 (NIH) and DMB 84-13986 (NSF), by an award from Marie Z. Cole Montrose, by a grant from the Pew Memorial Trust, by NIH Grants CA-06927 and RR-05539, and by an appropriation from the Commonwealth of Pennsylvania awarded to the Institute for Cancer Research.

[1] Abbreviations: COSY, *J*-correlated spectroscopy; 2D, two dimensional; DQF, double quantum filtered; MCD, main chain directed; NAB, NH–C_αH–C_βH *J*-coupled spin system; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, NOE-correlated spectroscopy; TOCSY, total correlation spectroscopy.

[2] Y. Feng and H. Roder, manuscript in preparation.

[3] Y. Feng, H. Roder, A. J. Wand, and S. W. Englander, manuscript in preparation.

0006-2960/89/0428-0195$01.50/0

Table I: Summary of 2D NMR Experiments

sample	solvent	T (°C)	experiment	mixing time (ms)	resolution (Hz/pt)
nonexchanged	90% H_2O/10% D_2O	20	DQF-COSY		4
		20	NOESY	120	8
		20	TOCSY	32	8
		20	TOCSY	46	8
		20	TOCSY	57	8
		30	COSY		8
		30	NOESY	120	8
		30	TOCSY	30	8
partially exchanged out	D_2O	20	COSY		8
		20	NOESY	120	8
		20	TOCSY	59	8
partially exchanged in	90% H_2O/10% D_2O	20	COSY		8
		20	NOESY	120	8

tassium ferricyanide was added to ensure full oxidation, and the samples were filtered through Sephadex G-25 columns equilibrated with the desired buffer. The final NMR samples contained 8–10 mM protein at pH or pD 5.7 (uncorrected reading, using a combination glass electrode) in 0.1 M sodium phosphate and 0.15 M NaCl.

Oxidized cytochrome *c* samples were prepared in three different ways to select subsets of labile protons according to their hydrogen exchange rates. (i) Nonexchanged. Cytochrome *c* was dissolved in H_2O buffer and passed through a Sephadex column equilibrated with 90% H_2O/10% D_2O buffer. Under these conditions essentially all amide NH resonances are observed. (ii) Partially exchanged out. Cytochrome *c* was dissolved in H_2O buffer and transferred into D_2O by Sephadex gel filtration. Only slowly exchanging labile protons with exchange half-times greater than about 3 h are present. (iii) Partially exchanged in. The protein was dissolved in D_2O buffer and incubated at 60 °C for 3 h in order to exchange all the labile protons. It was then passed through a Sephadex column equilibrated with 90% H_2O/10% D_2O buffer. Under these conditions, labile protons with exchange half-times less than about 15 h are selected. Partially labeled samples were stored at 4 °C up to 12 h prior to 2D NMR experiments.

All ^1H NMR experiments were performed on a Bruker AM 500 spectrometer. Magnitude-mode COSY (Aue et al., 1976; Nagayama et al., 1980; Bax & Freeman, 1981), phase-sensitive DQF-COSY (Shaka & Freeman, 1983; Rance et al., 1984), NOESY (Macura & Ernst, 1980; Kumar et al., 1980), and TOCSY (Braunschweiler & Ernst, 1983; Bax & Davis, 1985) spectra were recorded for oxidized cytochrome *c* samples under conditions indicated in Table I. A spectral width of 8333.3 Hz in both dimensions and a repetition period of about 1.3 s (including acquisition) were used for all experiments. Phase-sensitive spectra were recorded with time proportional phase incrementation (Redfield & Kuntz, 1975; Marion & Wüthrich, 1983) to provide quadrature detection in the incremented time domain. The H_2O or residual HDO signal was suppressed by selective saturation at all times except during data acquisition. A home-built auxiliary decoupler was used for solvent suppression in TOCSY experiments (Dykstra & Wand, 1988).

Pertinent data acquisition parameters used for the 2D NMR experiments are listed in Table I. For all spectra 64 transients were collected per t_1 increment. The number of t_1 increments collected was 650 for the DQF-COSY spectrum and 512 for all other experiments. The data were processed on a MicroVax II, using the program FTNMR (courtesy of Dr. D. Hare, Hare Research, Woodinville, WA). Magnitude-mode COSY and phase-sensitive DQF-COSY spectra were processed with nonshifted sine multiplication and exponential weighting (2–3-Hz line broadening) in both dimensions. For NOESY and TOCSY spectra, phase-shifted sine multiplication (20–30°) was employed, followed by polynomial base-line correction in both frequency dimensions. The final digital resolution indicated in Table I was obtained by zero filling of the t_1 time-domain data only.

Sample preparation and experimental design for chemical exchange spectroscopy, used to correlate assignments in reduced and oxidized forms, will be described in detail elsewhere.[3]

Results

Assignment Strategy. Since the sequential assignment approach (Wüthrich et al., 1982; Wüthrich, 1986) is difficult to apply to a protein the size of cytochrome *c*, we concentrated on the main-chain resonances throughout this work. An initial set of main-chain assignments in oxidized cytochrome *c* was based on available assignments in the reduced protein (Wand et al., 1989), making use of 2D magnetization transfer methods (Jeener et al., 1979; Kumar et al., 1980) to correlate corresponding resonances in mixed samples of oxidized and reduced cytochrome *c* undergoing rapid electron exchange (Redfield & Gupta, 1971). It soon became clear that this approach could yield only a limited number of unambiguous assignments, namely, for those resonances that are resolved in at least one oxidation state and experience relatively large redox shifts. These served as stepping stones for a systematic study of the NOE connectivities for the oxidized protein. Chemical exchange spectroscopy, including a recently developed relayed magnetization transfer method (Feng & Roder, 1988), became important again at a later stage, where it was used to cross-check assignments obtained in each oxidation state.[3]

The subsequent effort to assign the proton NMR spectrum of oxidized cytochrome *c* was guided by three general strategies. One involved the simplification of the fingerprint region of the spectrum by use of kinetic proton labeling (Wagner et al., 1981; Calhoun et al., 1985). Three types of samples were prepared (see Materials and Methods) to obtain proton labeling of all amide sites, the slowly exchanging amide sites, or the rapidly exchanging amide sites. A second strategy relied on the sensitivity of ^1H chemical shifts to temperature to help resolve ambiguities in 2D NMR spectra. These two approaches were also successfully used in the assignment of hen egg white lysozyme (Redfield & Dobson, 1988). Assignments based on cross correlation with the reduced protein and the extension of side-chain assignments for the oxidized form reported by other workers [e.g., Keller and Wüthrich (1978), Moore and Williams (1980a,b) and Williams et al. (1985) and references cited therein] provided reference points along the sequence during the implementation of the third strategy, the application of the main chain directed (MCD) assignment

procedure (Englander & Wand, 1987; Di Stefano & Wand, 1987). As with the assignment of reduced horse cytochrome *c* (Wand et al., 1989), the MCD approach was used in order to overcome our inability to adequately identify a sufficient number of amino acid side-chain spin systems to carry out a formal sequential assignment (Wüthrich, 1986).

Preliminary Spin System Recognition. As recommended in the sequential assignment strategy (Wüthrich et al., 1982), a search for side-chain spin systems in *J*-correlated spectra was carried out initially. Due to the complexity of the spectrum, only a limited number of simple spin systems could be identified at this stage. Furthermore, the presence of the paramagnetic iron can lead to atypical chemical shifts and short T_2 relaxation times, which makes spin system identification more difficult.

Residues that could be identified played a confirmatory role during the search for the main-chain NOE connectivities. Nine of the 12 glycines could be recognized directly in *J*-correlated spectra due to their characteristic NH–$C_\alpha H_2$ coupling pattern. It is interesting to note that all of the 11 glycines ultimately found are present and can be recognized by inspection in the spectrum of the partially exchanged-in sample (Figure 4A). Five of the six alanine spin systems were identified in COSY or DQF-COSY spectra and confirmed by relayed NH–$C_\beta H$ connectivities in TOCSY spectra. The sixth alanine was obscured by a cross-peak arising from the side chain of a threonine (see Discussion) and by a near degeneracy with the $C_\alpha H$–$C_\beta H$ cross peak of another residue. The search for valines and threonines was disappointing. Horse cytochrome *c* contains three valines. It was ultimately found that only one of the three valines showed long-range TOCSY connectivities from NH to $C_\beta H$ and to one $C_\gamma H_3$. Another valine could be identified in the TOCSY spectra by NH–$C_\beta H$ and $C_\alpha H$–$C_\gamma H_3$,$C_{\gamma'} H_3$ cross-peaks. The third spin system ultimately assigned to valine did not exhibit long-range correlations for any of the mixing times used in the TOCSY experiments. Of the 10 threonines, 4 were identified by recognition of their complete spin systems in COSY and TOCSY spectra. The remaining threonines did not produce NH–$C_\beta H$ cross-peaks in the TOCSY experiments. Some of the spin systems identified at this stage are shown in Figure 1.

Main-Chain Assignment. Sequence-specific main-chain assignments for oxidized cytochrome *c* are identified in Figure 2, which shows the fingerprint region (NH–$C_\alpha H$ cross-peaks) of the DQF-COSY spectrum. The number of cross-peaks in this spectral region is close to the expected number, in spite of the fact that some resonances experience considerable paramagnetic broadening. A complete list of proton assignments for ferricytochrome *c* obtained in this and earlier work is given in Table II.

The main-chain assignments were based primarily on the sequential NOE connectivities between NH–$C_\alpha H$–$C_\beta H$ spin systems (NAB sets) summarized in Figure 3. Among the 104 residues in horse cytochrome *c*, all except Pro-44 and Gly-84 were assigned. A total of 15 NAB sets could be directly identified in 2D spectra by extension of side-chain assignments reported previously [Williams et al. (1985) and references cited therein], and 13 NAB sets were obtained by cross correlation with the reduced protein.

Residues used as landmarks along the sequence are discussed below; connectivities that defined the intervening segments can be found readily in the NOE connectivity map in Figure 3. For the purpose of coherence, steps in the assignment are described along the polypeptide sequence although the actual assignment process did not proceed in this order.

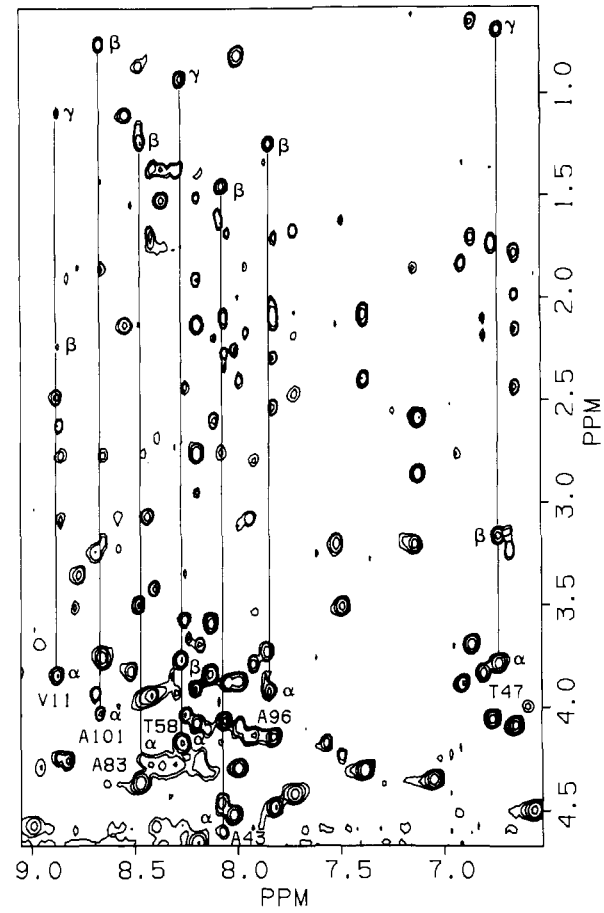

FIGURE 1: TOCSY spectrum of nonexchanged horse ferricytochrome *c* in 90% H_2O/10% D_2O at 20 °C (0.1 M sodium phosphate, 0.15 M NaCl, pH 5.7), obtained with a mixing time of 57 ms. Examples of side-chain spin systems identified at the first stage of assignment are shown.

The initial identification of the N-terminal helical segment came from a redox exchange study. Two cross-peaks in the oxidized cytochrome *c* spectrum were found to correlate with Lys-13 NH–$C_\alpha H$ and Cys-14 NH–$C_\alpha H$ in the reduced cytochrome *c* spectrum (Wand & Englander, 1986). The Lys-13 and Cys-14 neighbor relationship was established by main-chain NOE connectivities in the NOESY spectrum of oxidized cytochrome *c* (Figure 3). The assignments were confirmed by the NOE from Cys-14 NH to the bridge 2-methine[2] and by the extension of the NOE connectivities to an alanine previously assigned to Ala-15 (Williams et al., 1985). Moving from Lys-13 toward the N-terminus, the observation of NOE connectivities typical for helical secondary structure (Billeter et al., 1982) allowed unambiguous identification of the NAB sets for Gln-12 to Lys-5. The connectivities terminated at Glu-4 because of the degeneracy of Glu-4 NH and Lys-5 NH, a situation that also occurred at this point in the reduced protein (Wand et al., 1989). This ambiguity was later overcome by the identification of a segment containing the only unassigned valine, Val-3. The NH–NH connectivities for the N-terminal segment are shown in Figure 4.

Unambiguous NOE connectivities could be established from the N-terminal helix through to Glu-21 NH. Although no NH–$C_\alpha H$ cross-peak was found for this residue in spectra obtained at 20 °C, the NH–$C_\alpha H$ cross-peak was present at 30 °C, suggesting that at the lower temperature the Glu-21 $C_\alpha H$ resonates near the solvent signal (4.83 ppm). An NH–NH NOE between two glycines revealed Gly-23 and Gly-24, the only nearest-neighbor pair of glycines in the sequence.

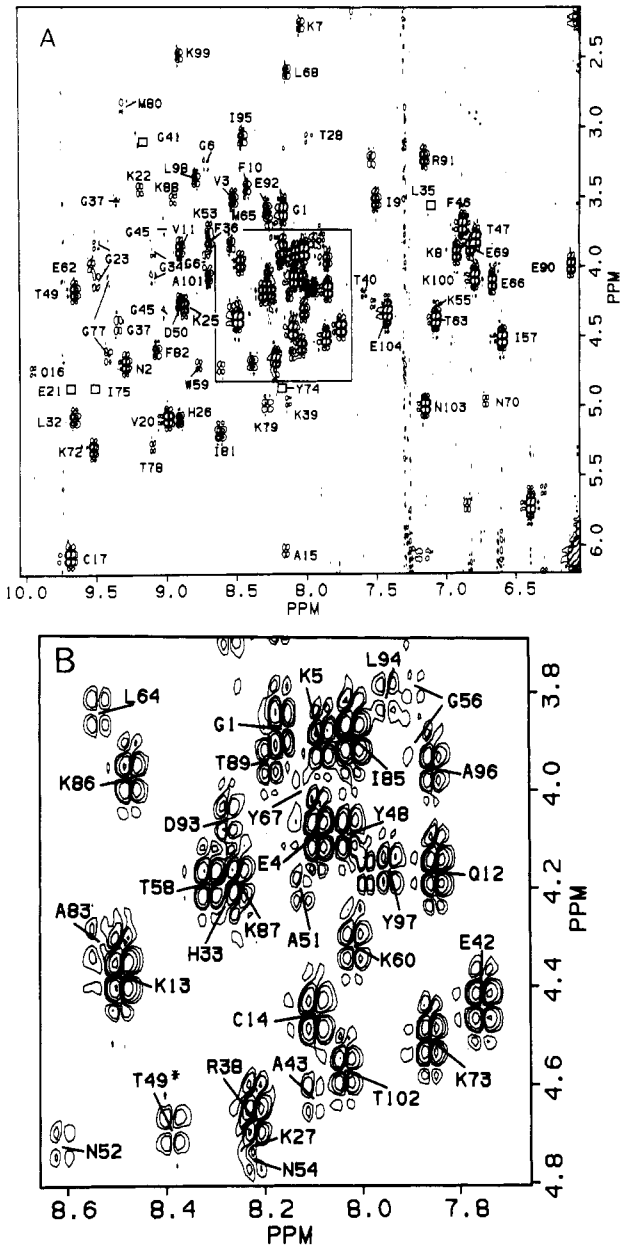

FIGURE 2: (A) Fingerprint region of a DQF-COSY spectrum of nonexchanged horse ferricytochrome *c* at 20 °C (sample condition as in Figure 1). Assigned cross-peaks are indicated. The NH–C_αH cross-peaks of His-18, Thr-19, Asn-31, and Glu-61 are outside the region shown. The squares represent weak cross-peaks found only in other COSY spectra or cross-peaks missing at 20 °C due to the proximity of the C_αH resonance to the solvent signal (see text). An expanded plot of the spectral region in the box is shown in panel B. The asterisk in panel B indicates the correlation of Thr-49 OH and C_βH.

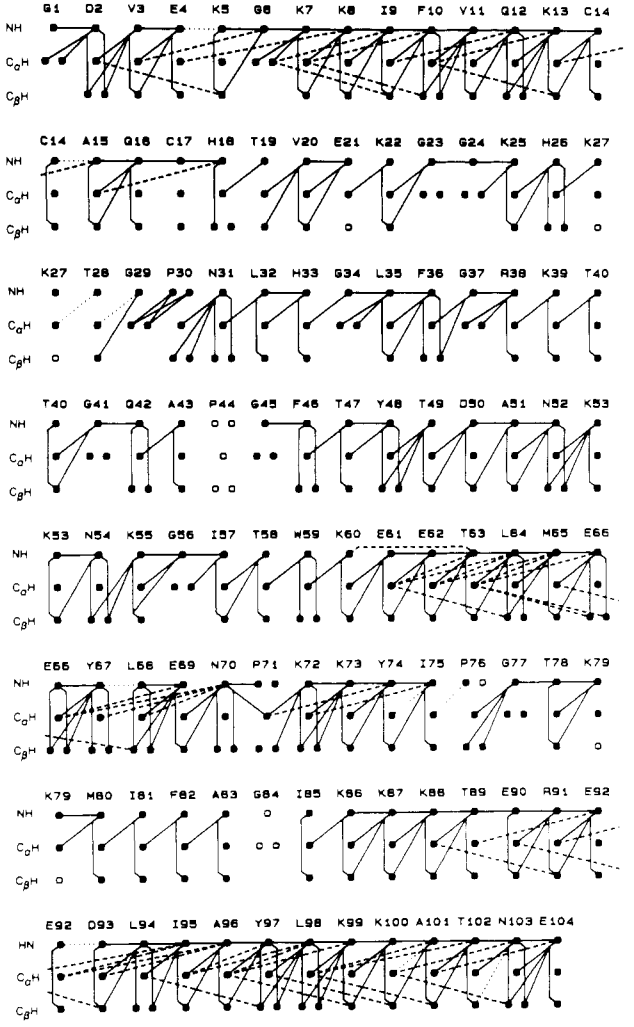

FIGURE 3: NOE connectivity diagram for backbone protons and β-protons. Solid lines indicate NH–C_βH NOEs within individual residues and NOEs between neighboring residues. Dashed lines indicate medium-range NOEs within segments of secondary structure; only unambiguous NOEs are included. Dotted lines indicate connectivities that are ambiguous due to either degeneracy or proximity of a resonance to the solvent signal. For prolines, the C_δ protons are placed at the NH position, and C_γ protons are not shown. The map summarizes data obtained at 20 and 30 °C.

Because Lys-27 C_αH is close to the solvent signal and Thr-28 C_αH and C_βH are degenerate, the assignment for Thr-28 NH–C_αH remains tentative. Partial assignments for Gly-29 and Pro-30, obtained by one-dimensional NMR methods,[2] were completed in the 2D NMR studies. The Pro-30 assignment was further confirmed by sequential NOE connectivities up to Leu-32, one of the residues verified by redox exchange correlations with the reduced protein.

Tracing the main-chain NOE connectivities was straightforward between Leu-32 and Ala-43. At this point the chain was interrupted due to difficulties in assigning Pro-44. The identification in the TOCSY spectrum of threonine spin systems at positions 47 and 58, partially assigned previously by species comparisons (Williams et al., 1985), made it possible to extend the main-chain assignments from Gly-45 to Tyr-67 (Figure 3).

A redox exchange correlation for Leu-68 NH–C_αH identified Leu-68 NH, which is degenerate with Tyr-67 NH. NOE connectivities to the side-chain resonances of Leu-68 (Williams et al., 1985) and Trp-59 (Boswell et al., 1982) further confirmed the assignment. Glu-69 and Asn-70 were readily assigned on the basis of multiple NOE connectivities. The main-chain NOE connectivities for the helical segment from Glu-61 to Glu-69 are indicated in the NOESY spectrum shown in Figure 5.

The recognition of Gly-77 and Thr-78, by the identification of glycine and threonine spin systems and the NH–NH NOE connectivity between them, led to assignment of the segment between Pro-71 and Pro-76. The *J*-connectivities of the two prolines are shown in Figure 6. Although Tyr-74 NH was well-defined, no NH–C_αH cross-peak was found at either 20 or 30 °C, presumably because Tyr-74 C_αH is close to the solvent signal at both temperatures. An extra cross-peak aligned with Tyr-74 NH in the TOCSY experiment was tentatively assigned to Tyr-74 C_βH. The recognition of Gly-77 and Thr-78 also allowed the continuation of the assignment

Table II: Chemical Shifts of Assigned Proton Resonances of Horse Ferricytochrome *c*[a]

residue	NH	$C_\alpha H$	$C_\beta H$	others	residue	NH	$C_\alpha H$	$C_\beta H$	others
Gly-1	8.17	3.88, 3.61			Asn-52	8.61	4.73	3.12, 3.25	
Asp-2	9.29	4.70	2.74, 2.42		Lys-53	8.70	3.72	1.86	$C_\gamma H$ 1.45
Val-3	8.52	3.53	2.15[b]	$C_\gamma H_3$ 1.02,[b] 1.07[b]	Asn-54	8.21	4.76	2.79, 2.99	
Glu-4	8.09	4.10	2.12	$C_\gamma H$ 2.34 (or β'), 2.26	Lys-55	7.05	4.29	2.26	
					Gly-56	7.90	3.79, 3.89		
Lys-5	8.09	3.91	1.70		Ile-57	6.59	4.52	1.94	$C_\gamma H$ 1.24, 0.70; $C_\gamma H_3$ 0.86;[b] $C_\delta H_3$ −0.36[b]
Gly-6	8.72	3.28, 3.97							
Lys-7	8.05	2.28	1.78		Thr-58	8.31	4.20	3.80[b]	$C_\gamma H_3$ 0.96[b]
Lys-8	6.92	3.91	1.86, 1.60	$C_\gamma H$ 1.40	Trp-59	8.77	4.71	3.64, 2.16	N_1H 9.61; C_2H 6.89;[j] C_4H 7.37;[j] C_5H 6.55;[j] C_6H 6.37;[j] C_7H 7.62[j]
Ile-9	7.51	3.54	1.64[b]	$C_\gamma H$ 0.93; $C_\gamma H_3$ 0.40;[b] $C_\delta H3$ 0.76[b]					
Phe-10	8.43	3.44	3.17, 2.75	C_2H 7.54;[c] C_3H 8.41;[c] C_4H 8.79;[c] C_5H 7.66;[c] C_6H 6.92[c]	Lys-60	8.02	4.32	2.03	
					Glu-61	10.61	3.68	2.07	
					Glu-62	9.54	4.00	2.03	
Val-11	8.90	3.88	2.26	$C_\gamma H_3$ 1.12,[b] 1.34[b]	Thr-63	7.08	4.38	4.57[b]	$C_\gamma H$ 1.42[b]
Gln-12	7.85	4.17	2.10, 2.02	$C_\gamma H$ 2.30, 2.56	Leu-64	8.54	3.85	1.57, 0.49	$C_\gamma H$ 1.34;[b] $C_\gamma H_3$ −0.22,[b] −0.59[b]
Lys-13	8.49	4.38	1.19	$C_\gamma H$ 0.88	Met-65	8.28	3.69	2.18	
Cys-14	8.10	4.46	1.69		Glu-66	6.67	4.12	1.80, 2.02	$C_\gamma H$ 2.44, 2.17
Ala-15	8.16	6.03[b]	2.24[b]		Tyr-67	8.13	4.03	2.67, 3.32	
Gln-16	9.95	4.75	2.47, 2.69	$C_\gamma H$ 3.03, 2.88	Leu-68	8.14	2.62	1.02, −0.05	$C_\gamma H$ 0.67;[b] $C_\delta H_3$ −2.78,[b] −0.60[b]
Cys-17	9.69	6.08	7.17		Glu-69	6.83	3.87	2.13	$C_\gamma H$ 2.10 (or β')
His-18	11.00	9.16[d]	9.00,[d] 14.82[d]	C_2H −25.60;[e] C_5H 24.96;[d] $N_\pi H$ 12.85[e]	Asn-70	6.72	4.97	3.25, 3.20	
					Pro-71		5.43	5.10, 5.31	$C_\gamma H$ 2.11, 3.30; $C_\delta H$ 4.34, 3.69
Thr-19	10.75	6.32	5.61[b]	$C_\gamma H_3$ 2.27;[b] OH 9.06[f]	Lys-72	9.52	5.32	2.83, 2.68	$C_\gamma H$ 2.41, 3.64
Val-20	8.99	5.09	2.25	$C_\gamma H_3$ 1.06,[b] 1.11[b]	Lys-73	7.87	4.52	2.14	$C_\gamma H$ 1.73
Glu-21	9.68	4.83			Tyr-74	8.20	4.83	4.16[m]	$C_{2,6}H$ 7.78;[g] $C_{3,5}H$ 6.84[g]
Lys-22	9.19	3.45	1.60	$C_\gamma H$ 0.79 (or β')	Ile-75	9.54	4.83	3.52[i]	$C_\gamma H_3$ 1.37[i]
Gly-23	9.50	4.11, 3.85			Pro-76		5.25	2.14, 2.62	$C_\gamma H$ 2.34, 3.77; $C_\delta H$ 3.87
Gly-24	8.24	3.75, 4.28			Gly-77	9.42	4.64, 4.10		
Lys-25	8.88	4.29	1.95	$C_\gamma H$ 2.09 (or β')	Thr-78	9.10	5.29	5.98[b]	$C_\gamma H_3$ 3.58;[b] OH 9.65[f]
His-26	8.91	5.10	3.12, 2.80	C_2H 7.66;[g] C_5H 6.99[g]	Lys-79	8.28	5.00		
Lys-27	8.23	4.72			Met-80	9.29	2.87	12.85[e]	$C_\gamma H$ −28.70;[e] $C_\epsilon H$ −24.86[k]
Thr-28	7.98[m]	3.08[m]	3.08[b]	$C_\gamma H$ −0.02[b]	Ile-81	8.62	5.20	2.18[b]	$C_\gamma H$ 1.43, 1.65; $C_\gamma H_3$ 1.14;[b] $C_\delta H_3$ 0.98[b]
Gly-29	6.97	−4.42,[e] −0.98[e]							
Pro-30		3.77	−0.32, 1.41	$C_\gamma H$ −0.34; $C_\delta H$ −6.69,[e] −2.42[e]	Phe-82	9.07	4.61	3.71	C_4H 5.89;[j] $C_{3,5}H$ 6.04;[j] $C_{2,6}H$ 6.07[l]
Asn-31	11.61	6.00	2.96, 2.59	NH_2 8.10[f]	Ala-83	8.54	4.34[b]	1.30[b]	
Leu-32	9.66	5.10	2.65	$C_\gamma H$ 2.41 (or β'); $C_\delta H_3$ 2.14, 1.76	Gly-84				
					Ile-85	8.02	3.90	0.83[b]	$C_\gamma H$ 0.43, 0.66; $C_\gamma H_3$ 0.57;[b] $C_\delta H_3$ −0.17[b]
His-33	8.30	4.22	3.36	C_2H 7.84;[h] C_5H 8.75[h]					
Gly-34	9.10	4.09, 3.90			Lys-86	8.47	4.00	1.77	$C_\gamma H$ 1.40 (or β')
Leu-35	7.13	3.65	2.23	$C_\gamma H$ 1.12;[b] $C_\delta H_3$ 0.50,[b] 0.16[b]	Lys-87	8.25	4.19	1.55	
					Lys-88	8.95	3.51	1.80	
Phe-36	8.69	3.81	3.22, 2.81	C_4H 6.78;[g] $C_{3,5}H$ 6.63;[g] $C_{2,6}H$ 7.27[g]	Thr-89	8.19	3.95	4.07[b]	$C_\gamma H_3$ 1.24[b]
					Glu-90	6.10	4.01	1.75	
Gly-37	9.35	4.42, 3.54			Arg-91	7.17	3.23	1.89	
Arg-38	8.22	4.67	2.17	$C_\gamma H$ 1.94 (or β')	Glu-92	8.28	3.59	2.11	$C_\gamma H$ 2.59
Lys-39	8.13	4.94	1.60		Asp-93	8.28	4.07	2.45	
Thr-40	7.59	4.20	4.50[i]	$C_\gamma H_3$ 0.87;[i] OH 7.37[f]	Leu-94	7.95	3.82	1.45, 1.09	CH 0.70;[b] $C_\delta H_3$ −0.29,[b] 0.60[b]
Gly-41	9.19	3.09, 1.27			Ile-95	8.46	3.09	1.67[i]	$C_\gamma H$ 1.30; $C_\gamma H$ 0.86;[b] $C_\delta H3$ 0.45[b]
Gln-42	7.76	4.44	2.51, 1.69	$C_\gamma H$ 2.19, 2.09					
Ala-43	8.10	4.63[b]	1.48[b]		Ala-96	7.86	3.96[b]	1.27[b]	
Pro-44					Tyr-97	7.94	4.17	2.83, 3.56	C_2H 6.41;[c] C_3H 5.73;[c] C_5H 6.86;[c] C_6H 7.16[c]
Gly-45	9.01	4.33, 3.72							
Phe-46	6.88	3.71	1.72, 0.67	C_2H 5.75; C_3H 5.53; C_4H 4.79; C_5H 5.00; C_6H 3.75	Leu-98	8.79	3.38	1.85, 1.06	$C_\gamma H$ 1.66;[b] $C_\delta H_3$ 0.49,[b] −0.13[b]
					Lys-99	8.91	2.51	1.50	
					Lys-100	6.79	4.08	1.78	$C_\gamma H$ 1.37
Thr-47	6.79	3.81	3.18[b]	$C_\gamma H_3$ 0.72[b]	Ala-101	8.70	4.08[b]	0.81[b]	
Tyr-48	8.03	4.10	2.41, 3.13		Thr-102	8.04	4.58	4.97[i]	$C_\gamma H_3$ 1.31[i]
Thr-49	9.65	4.18	4.70[b]	$C_\gamma H_3$ 1.56;[b] OH 8.39	Asn-103	7.15	5.01	2.62, 2.88	
Asp-50	8.91	4.27	2.65		Glu-104	7.42	4.34	2.12	$C_\gamma H$ 2.39 (or β')
Ala-51	8.21	4.19[b]	1.63[b]						

[a] Chemical shifts (ppm) are referenced to [^{2}H]TSP [sodium 3-(trimethylsilyl)[2,2,3,3-$^2H_4$]propionate] for horse ferricytochrome *c* in 0.1 M sodium phosphate–0.15 M sodium chloride buffer at 20 °C, pH 5.7. [b] In agreement with Williams et al. (1985). [c] In agreement with Arean et al. (1988). [d] In agreement with Satterlee and Moench (1987). [e] Feng and Roder (unpublished results). [f] Based on NOE only. [g] In agreement with Moore and Williams (1980a,b). [h] In agreement with Dobson et al. (1975). [i] In disagreement with Williams et al. (1985). [j] In agreement with Boswell et al. (1982). [k] In agreement with Redfield and Gupta (1971). [l] In agreement with Boswell et al. (1980). [m] Tentative assignment.

up to Ala-83. The connectivity search was terminated here because of the unassigned Gly-84; no candidate for the missing pair of glycine cross-peaks was found in any of the *J*-correlated spectra.

Assignments for the C-terminal helical segment were initiated by correlation of Leu-98 NH–$C_\alpha H$ and Lys-99 NH–$C_\alpha H$ cross-peaks between the reduced and the oxidized cytochrome *c* spectra in the redox exchange experiment. The main-chain protons of the segment are defined by a dense network of NOEs (Figure 3). These assignments were further confirmed by the appearance of two alanine coupling patterns at the expected residue positions (Ala-96 and Ala-101). The NOE patterns for the C-terminal helix are displayed in Figure 7.

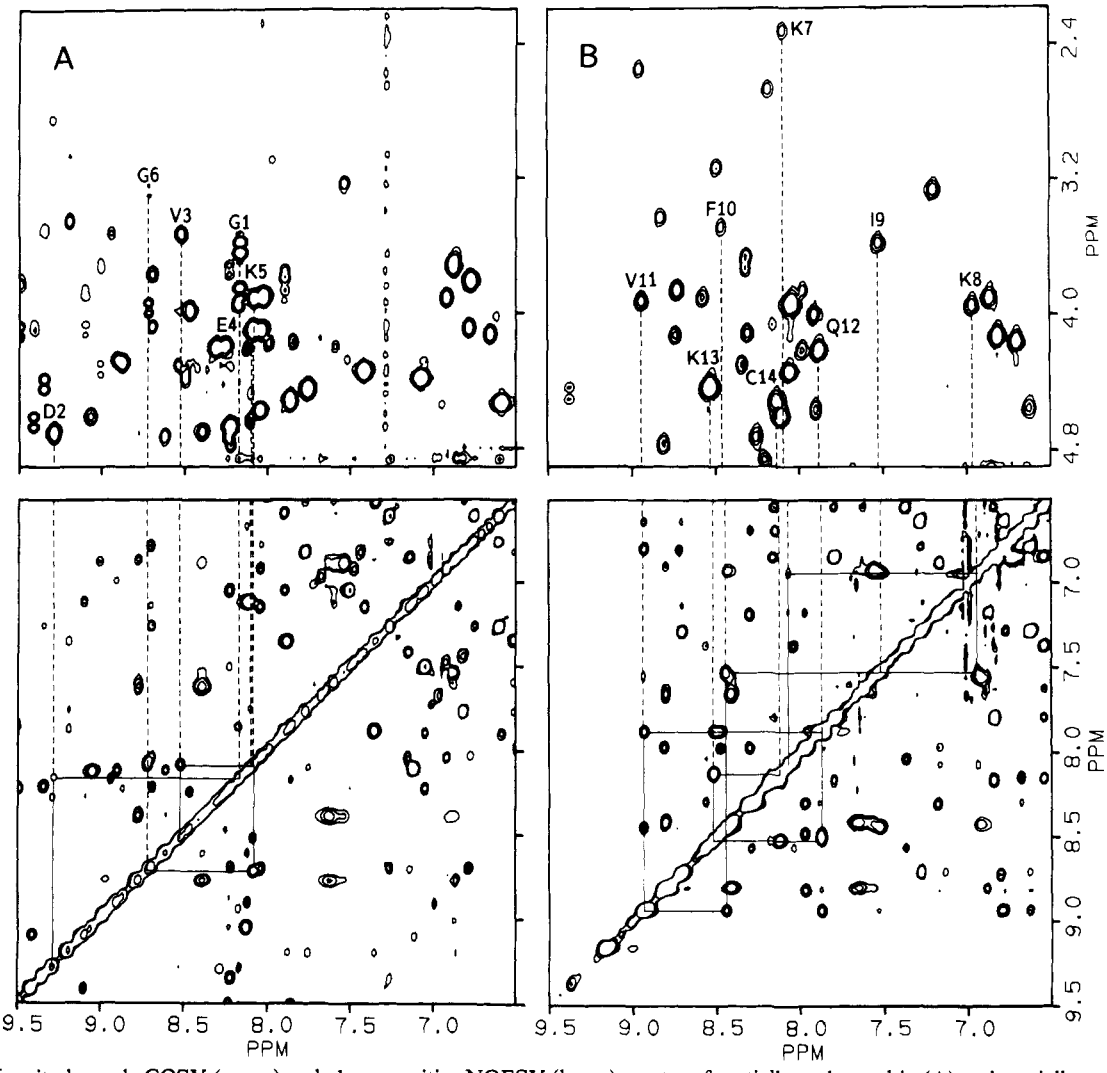

FIGURE 4: Magnitude-mode COSY (upper) and phase-sensitive NOESY (lower) spectra of partially exchanged-in (A) and partially exchanged-out (B) ferricytochrome *c* at 20 °C. Residues 1–14 are labeled in the upper panels, and their sequential d_{NN} connectivities are shown in the lower panels.

Late-Stage Side-Chain Assignments. At the last stage of assignment, given the known sequential position of each NAB set, it was possible to identify additional side-chain resonances. Groups of resonances belonging to a given side chain were identified in TOCSY spectra (cf. Figure 1), and *J*-correlated spectra were used to trace the coupling pattern whenever possible. In some cases, remaining ambiguities were resolved by consideration of intraresidue NOEs. Under our experimental conditions all isoleucines exhibited long-range *J*-connectivities in the TOCSY spectra from the NH to the $C_\beta H$ and $C_\gamma H_3$, as well as from the $C_\beta H$ through to the $C_\delta H_3$. Most of the leucines showed *J*-connectivities from the NH to the C_β protons and from the $C_\beta H$ through to the $C_\delta H_3$ and $C_{\delta'} H_3$. It would have been difficult to identify these spin systems at the first stage of assignment.

The phenylalanine and tyrosine side-chain spin systems were also most clearly defined in the TOCSY spectrum as demonstrated by the examples shown in Figure 8. Their placement along the amino acid sequence was based on multiple side-chain to main-chain NOEs. For rings that undergo a transition from slow to fast ring flip motion over the experimentally accessible temperature range, the assignment of individual protons within the ring was verified by the observation of the motionally averaged resonances (C_2H merges with C_6H and C_3H with C_5H) in TOCSY spectra recorded at elevated temperature (unpublished results). The majority of the published assignments for aromatic rings (Table II) were confirmed, and new assignments were added for the previously unknown Phe-46 ring spin system. Assignments for Tyr–97 differ from a recent report (Arean et al., 1988) with respect to the order of ring resonances. Our assignment was based on the relative intensity of NOEs from the ring protons to main-chain protons and was shown by magnetization transfer[3] to be consistent with reduced cytochrome *c* (Wand et al., 1989). The remaining two tyrosine rings (Tyr-48 and Tyr-67) remain elusive, possibly because of line broadening due to their close proximity to the paramagnetic heme or to an intermediate flip rate.

It is interesting to note that for slowly flipping aromatic rings (on the NMR time scale) all the ring proton resonances are mutually connected by TOCSY cross-peaks, even for tyrosines where the two sides of the ring are only weakly coupled. The observation that the C_2H–C_6H cross-peak is often comparable in intensity to the C_3H–C_5H cross-peak (Figure 8) suggests that magnetization transfer is due, in part, to chemical exchange effects [cf. Feng and Roder (1988)].

One particularly interesting side chain is that of Thr-49. Its hydroxyl proton is involved in an internal H-bond and gives rise to a sharp resonance at 8.39 ppm (Table II). The Thr-49 side-chain spin system perfectly mimics an alanine; the hydroxyl proton resonates in the amide region, there is a $C_\beta H$–OH cross-peak in the COSY spectrum (Figure 2), and there is an additional OH–$C_\gamma H_3$ cross-peak in the TOCSY spec-

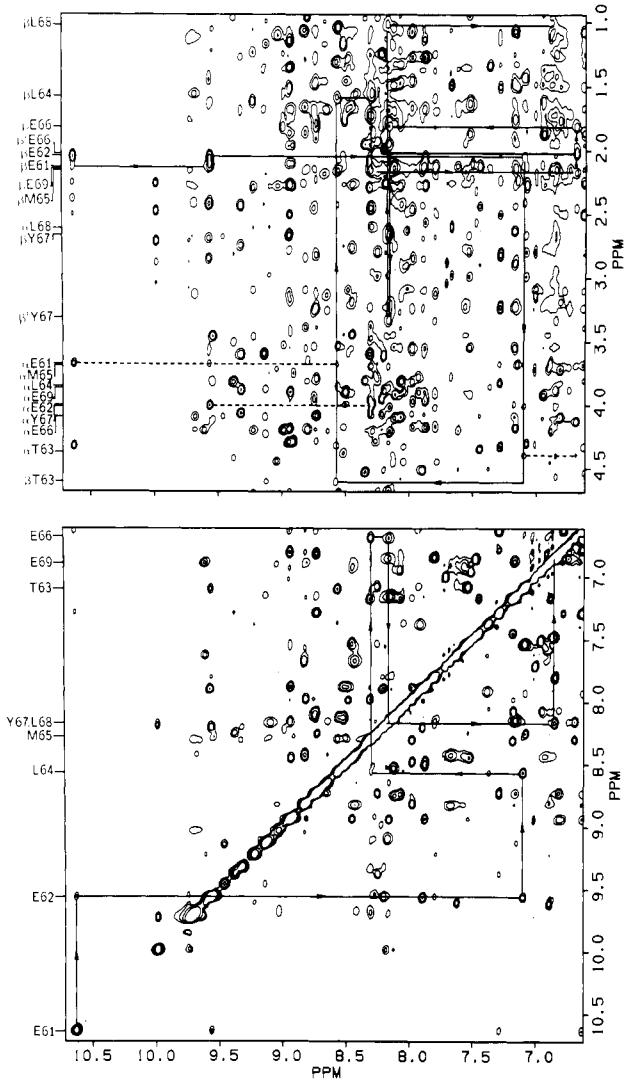

FIGURE 5: Main-chain NOE connectivities for the helical segment between residues 61 and 69, observed in a phase-sensitive NOESY spectrum of nonexchanged ferricytochrome *c* at 20 °C (120-ms mixing time). The d_{NN} NOE connectivities are identified in the lower panel. The upper panel shows $d_{\beta N}$ NOEs (solid line) and medium-range $d_{\alpha N}(i,j)$ NOEs (dashed line).

trum. We initially assigned the OH to an alanine NH, but the fact that two apparent NH protons were correlated with the same methyl resonance at both temperatures and the presence of an extra "NH–C_αH" cross-peak in this structural cluster was suspicious. A self-consistent NOE connectivity pattern then led to the assignment of this cross-peak to C_βH–OH coupling of Thr-49.

Tertiary NOEs. The density of NOE cross-peaks in some regions of the NOESY spectrum is quite staggering (see, e.g., Figure 5). A meaningful interpretation of this wealth of structural information requires systematic analysis by distance geometry methods (A. J. Wand, D. L. Di Stefano, R. A. Beckman, and D. R. Hare, unpublished results). The limited set of tertiary NOE connectivities analyzed in the present work was found to be consistent with the crystallographic model (Takano & Dickerson, 1981). For example, the proximity of the N-terminal and C-terminal helices is indicated by NOEs between Tyr-97 aromatic ring protons and Lys-7 NH and Phe-10 NH; the close approach of the extended chain segments from residue 17 to residue 20 and from residue 23 to residue 27 is reflected in several NOEs between main-chain protons on each strand; the Phe-36 ring displays multiple NOEs to main-chain protons of residues 60, 61 and 99, consistent with

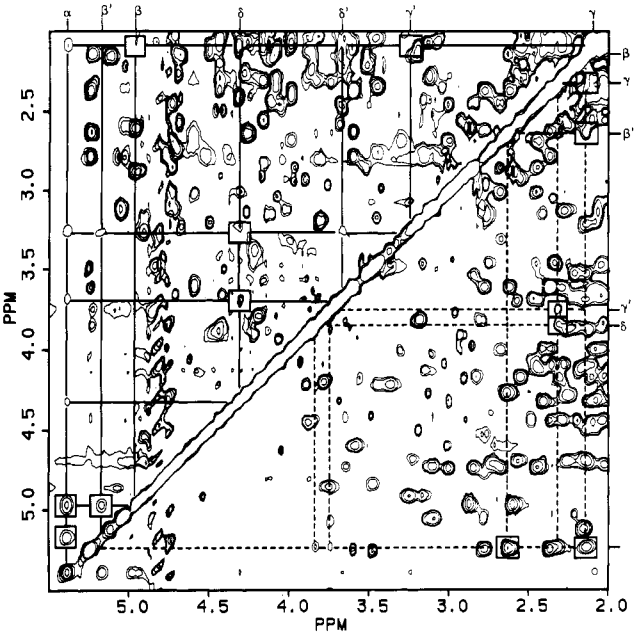

FIGURE 6: Identification of proline spin systems in a TOCSY spectrum of mixing time 59 ms at 20 °C. The *J*-connectivities of Pro-71 (solid line) and Pro-76 (dashed line) are shown. Only one of the $C_\delta H_2$ proton resonances was identified for Pro-76. Boxes indicate where there are cross-peaks in COSY spectra.

its location between the 60s helix and the C-terminal helix. The arrangement of the polypeptide chain around the heme group is reflected in numerous NOE connectivities with heme protons.

Discussion

Cytochrome *c* is among the largest proteins for which the proton resonance assignment puzzle has been solved. Other assigned proteins of comparable size, besides reduced cytochrome *c* (Wand et al., 1989), include hen egg white lysozyme (129 residues; Redfield & Dobson, 1988), thioredoxin (108 residues; LeMaster & Richards, 1988), and plastocyanin (99 residues; Chazin & Wright, 1988). The presence of a paramagnetic center in oxidized cytochrome *c* causes complications not found in other systems. Chemical shift criteria cannot be used to assist assignment (see, e.g., the unusual chemical shifts for the Phe-46 ring protons, Figure 8), and the quality of the spectra is in some cases degraded by paramagnetic broadening. For example, a number of protons that display strong cross-peaks in the diamagnetic reduced protein are difficult to observe in the oxidized form. The line-width problems also made it difficult to obtain useful relayed coherence transfer spectra, and total correlation spectroscopy (TOCSY) was used instead.

As in the previous work (Englander & Wand, 1987; Wand et al., 1989), the success of this assignment project relied heavily on our focus on main-chain protons. The diagram in Figure 3 shows that the main-chain assignments are for the most part based on multiple NOE connectivities between neighboring residues (the average number of nearest-neighbor connectivities per residue is 2.1). Medium-range NOEs, mostly in helical segments, provide additional cross-checks. Also important was the use of kinetic proton labeling, systematic work at different temperatures, and the correlation of reduced and oxidized cytochrome *c* spectra via redox exchange.

Although not all of the main-chain NOEs could be explored due to spectral overlap, the NOE connectivity map shown in Figure 3 provides a qualitative description of the solution structure of oxidized cytochrome *c*. Since no high-resolution

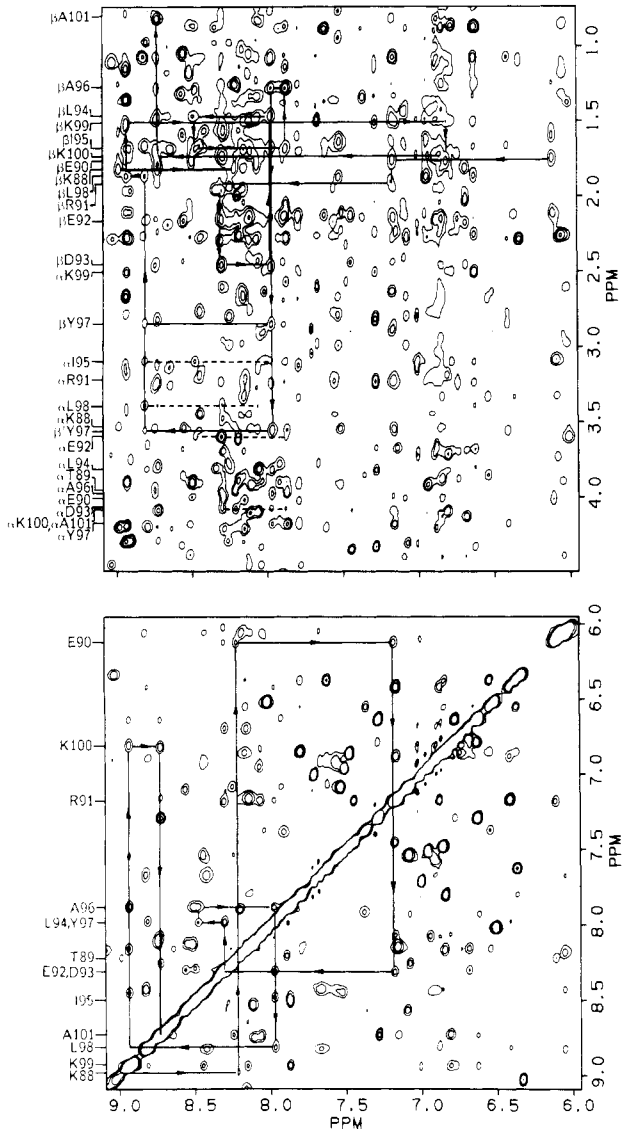

FIGURE 7: Main-chain NOE connectivities for the helical segment between residues 88 and 101 in a phase-sensitive NOESY spectrum of nonexchanged ferricytochrome *c*. Experimental conditions and line types are the same as in Figure 5. The connectivities for Thr-102 through Glu-104 are outside the spectra region shown.

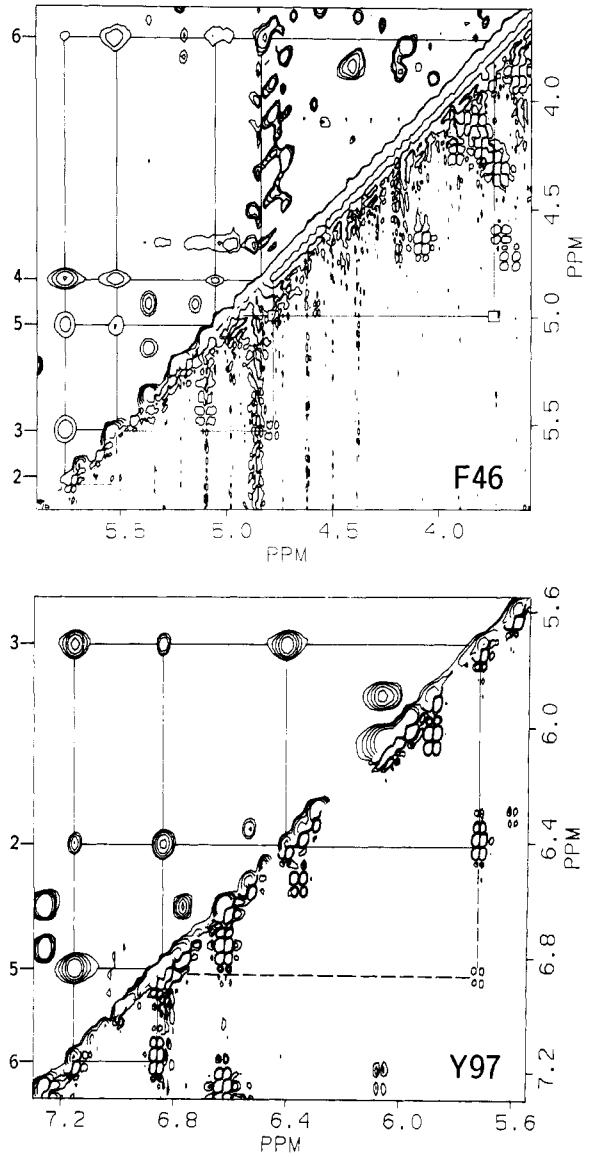

FIGURE 8: Identification of aromatic spin systems in composite DQF-COSY (lower triangle) and TOCSY (upper triangle) spectra of partially exchanged-out ferricytochrome *c* at 20 °C. The *J*-connectivities for the ring protons of Phe-46 and Tyr-97 are indicated. The squares indicate cross-peaks missing due to solvent suppression. The TOCSY spectrum was recorded with a mixing time of 59 ms. RF heating caused a slight temperature increase in the TOCSY experiment, resulting in some chemical shift difference between the two spectra, particularly for paramagnetically shifted resonances. A four-bond coupling is indicated by the dashed line.

X-ray crystal structure for horse cytochrome *c* is available, a comparison can be made only with the low-resolution (2.8 Å) horse ferricytochrome *c* (Dickerson et al., 1971) and the high-resolution (1.8 Å) tuna ferricytochrome *c* structures (Takano & Dickerson, 1981). Overall, good agreement is found between the secondary structure in solution (Figure 3) and that in the crystalline state, and the observed long-range NOEs are consistent with the tertiary structure of the crystallographic model.

The NOE map (Figure 3) documents three long stretches of α-helix, characterized by extended sequential d_{NN} connectivities and medium-range $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ NOEs (Billeter et al., 1982; Wüthrich, 1986). The first helix, at the N-terminus of the protein, is well-defined from Val-3 to Cys-14. A second helix extends from Glu-61 to Glu-69. As in the reduced protein (Wand et al., 1989), the pattern of NOEs in this segment is more consistent with a regular α-helix rather than the 3_{10} helix (or series of 3_{10} turns) described in the crystallographic study (Dickerson et al., 1971). The pattern of d_{NN}, $d_{\beta N}$, and medium-range NOEs (Figure 3) indicates a third helix at the C-terminal end, extending from Lys-88 to the C-terminal Glu-104. The NOE pattern found for the stretch from residue 72 to residue 75 provides evidence for an additional short helical segment.

A 2D NMR investigation of the N-terminal and C-terminal helices in tuna cytochrome *c* was reported during the course of this work (Williams, 1986). In tuna cytochrome *c*, the two helices appear to extend from Val-3 to Gln-12 and from Glu-90 to Ser-100. The difference in helix length in the two species is presumably due to amino acid changes. The boundary of the tuna C-terminal helix at Glu-90 (Williams, 1986), compared with Lys-88 in horse, can be rationalized by the substitution of Thr-89 (horse) by glycine, a typical helix breaker. The extension of well-defined structure to the C-terminus of the horse protein suggests a stabilizing role for the additional Glu-104 residue.

In the crystal structure (Dickerson et al., 1971), three type II turns were observed (Glu-21 to Gly-24, Leu-35 to Arg-38, Ile-75 to Thr-78). In the NOESY spectrum, these residues

indeed show NOEs consistent with type II turn conformation (Wüthrich, 1986), though this is not necessarily diagnostic. The trans configuration was found for all three assigned prolines (30, 71, 76), in agreement with the crystallographic results. An apparent NOE from Pro-71 $C_\alpha H$ to Asn-70 NH, weaker than the one to Lys-72 NH, may result from spin diffusion.

The appearance of the $C_\beta H$–OH cross-peak of Thr-49 in the *J*-correlated spectrum is noteworthy. This is the only $C_\beta H$–OH cross-peak found among 10 threonines in horse cytochrome *c*. Simple alcohols in organic solvents display a typical coupling constant of 4–6 Hz for similar groups (Silverstein et al., 1981; Pouchert, 1983) so that the cross-peak may appear in aqueous solvent if the exchange of the hydroxyl proton with water protons is sufficiently slow. A hydroxyl group at position 49 is conserved in cytochrome *c* for all species. In the crystal structure of cytochrome *c*, the OH group is buried and hydrogen bonded to propionate-6, evidently slowing the OH hydrogen exchange rate.

References

Arean, C. O., Moore, G. R., Williams, G., & Williams, R. J. P. (1988) *Eur. J. Biochem. 173*, 607–615.

Aue, W. P., Bartholdi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229–2246.

Bax, A., & Freeman, R. (1981) *J. Magn. Reson. 44*, 542–561.

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

Bechtold, R., Kuehn, C., Lepre, C., & Isied, S. S. (1986) *Nature (London) 322*, 286–288.

Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 321–346.

Boswell, A. P., Moore, G. R., Williams, R. J. P., Chien, J. C. W., & Dickinson, L. C. (1980) *J. Inorg. Biochem. 13*, 347–352.

Boswell, A. P., Eley, C. G. S., Moore, G. R., Robinson, M. N., Williams, G., Williams, R. J. P., Neupert, W. J., & Hennig, B. (1982) *Eur. J. Biochem. 124*, 289–294.

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

Calhoun, D. B., & Englander, S. W. (1985) *Biochemistry 24*, 2095–2100.

Chazin, W. J., & Wright, P. E. (1988) *J. Mol. Biol. 202*, 623–636.

Dickerson, R. E., Takano, T., Eisenberg, D., Kallai, O. B., Samson, L., Cooper, A., & Margoliash, E. (1971) *J. Biol. Chem. 246*, 1511–1535.

Dickerson, R. E., Timkovich, R., & Almassy, R. J. (1976) *J. Mol. Biol. 100*, 473–491.

Di Stefano, D. L., & Wand, A. J. (1987) *Biochemistry 26*, 7272–7281.

Dobson, C. M., Moore, G. R., & Williams, R. J. P. (1975) *FEBS Lett. 51*, 60–65.

Dykstra, R., & Wand, A. J. (1988) *J. Magn. Reson.* (in press).

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

Feng, Y., & Roder, H. (1988) *J. Magn. Reson. 78*, 597–602.

Keller, R. M., & Wüthrich, K. (1978) *Biochim. Biophys. Acta 533*, 195–208.

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

LeMaster, D. M., & Richards, F. M. (1988) *Biochemistry 27*, 142–150.

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

Margoliash, E., & Bosshard, H. R. (1983) *Trends Biochem. Sci. (Pers. Ed.) 8*, 316–320.

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

Mayo, S. L., Ellis, W. R., Crutchley, R. J., & Gray, H. B. (1986) *Science (Washington, D.C.) 233*, 948–951.

McDonald, C. C., & Philips, W. D. (1969) *J. Am. Chem. Soc. 91*, 1513–1521.

Moore, G. R., & Williams, R. J. P. (1980a) *Eur. J. Biochem. 103*, 493–502.

Moore, G. R., & Williams, R. J. P. (1980b) *Eur. J. Biochem. 103*, 503–512.

Nagayama, K., Kumar, A., Wüthrich, K., & Ernst, R. R. (1980) *J. Magn. Reson. 40*, 321–334.

Pielak, G. J., Concar, R. W., Moore, G. R., & Williams, R. J. P. (1987) *Protein Eng. 1*, 83–88.

Pouchert, C. J. (1983) *The Aldrich Library of NMR Spectra*, 2nd ed., Vol. I and II, Aldrich, Milwaukee, WI.

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., & Gupta, R. J. (1971) *Adv. Magn. Reson. 5*, 82–113.

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

Redfield, C., & Dobson, C. M. (1988) *Biochemistry 27*, 122–136.

Roder, H. (1988) *Methods Enzymol.* (in press).

Roder, H., Elöve, G. A., & Englander, S. W. (1988) *Nature (London)* (in press).

Satterlee, J. D., & Moench, S. (1987) *Biophys. J. 52*, 101–107.

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

Silverstein, R. M., Bassler, G. C., & Morrill, T. C. (1981) *Spectrometric Identification of Organic Compounds*, 4th ed., Wiley, New York.

Takano, T., & Dickerson, R. E. (1981) *J. Mol. Biol. 153*, 95–115.

Wand, A. J., & Englander, S. W. (1985) *Biochemistry 24*, 5290–5294.

Wand, A. J., & Englander, S. W. (1986) *Biochemistry 25*, 1100–1106.

Wand, A. J., Di Stefano, D. L., Feng, Y., Roder, H., & Englander, S. W. (1989) *Biochemistry* (preceding paper in this issue).

Williams, G. (1986) *J. Inorg. Biochem. 28*, 373–380.

Williams, G., Moore, G. R., Porteous, R., Robinson, M. N., Soffe, N., & Williams, R. J. P. (1985) *J. Mol. Biol. 183*, 409–428.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Wüthrich, K., Wider, G., Wagner, G., & Braun, W. (1982) *J. Mol. Biol. 155*, 311–319.