Proton Resonance Assignments of Horse Ferrocytochrome c[†]

A. Joshua Wand* and Deena L. Di Stefano
Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111
Yiqing Feng, Heinrich Roder, and S. Walter Englander*
Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104
Received May 10, 1988; Revised Manuscript Received July 21, 1988

ABSTRACT: Two-dimensional nuclear magnetic resonance (NMR) spectroscopy was used to assign the proton resonances of horse ferrocytochrome c. Assignments were based on the main chain directed (MCD) and sequential assignment procedures. The fundamental units of the MCD approach, the main-chain NH–C_αH–C_βH J-coupled subspin systems of each amino acid residue (NAB sets), were defined by analysis of direct and relayed coherence transfer spectra. Recognition of main-chain NOE connectivity patterns specified in the MCD algorithm then allowed NAB sets to be aligned in their proper juxtaposition within secondary structural units. The units of secondary structure were placed within the polypeptide sequence of identification of a small number of side-chain J-coupled spin systems, found by direct recognition in 2D spectra of some J-coupled spin systems and by pairwise comparisons of the J-correlated spectra of six homologous cytochromes c having a small number of known amino acid differences. The placement of a given segment in this way defines the amino acid identity of all its NAB sets. This foreknowledge allowed the vast majority of the side-chain resonances to be discerned in J-correlated spectra. Extensive confirmation of the assignments derives internally from multiple main-chain NOE connectivities and their consistency following temperature-induced changes of the chemical shifts of NOE-correlated protons. The observed patterns of main-chain NOEs provide some structural information and suggest small but potentially significant differences between the solution structure observed by NMR and that defined earlier in crystallographic studies at 2.8-Å resolution.

This paper presents extensive resonance assignments for the protons in reduced horse heart cytochrome c, obtained by two-dimensional nuclear magnetic resonance (2D NMR)[1] methods. The present work was undertaken to provide a basis for a variety of protein chemical studies on cytochrome c, a particularly interesting protein that has provided a focus for intensive structure–function studies [see, e.g., Margoliash and Boshard (1983), Bechtold et al. (1986), Mayo et al. (1986), and Pielak et al. (1987)]. Given pertinent resonance assignments, direct NMR measurements on structure and internal motions can be performed on the protein in the solution state. In addition we are especially interested in hydrogen-exchange studies, which may provide detailed information on protein folding, internal stability and dynamics, and structure–function relationships (Wand et al., 1986; Englander et al., 1987). Such studies depend upon the assignment of the backbone amide NH, which can then allow H-exchange observations at a proton-resolved level at 100 probe points throughout the cytochrome c protein.

Assignment of the proton resonances in a protein the size of cytochrome c (104 residues) represents a considerable challenge to currently available NMR technology and strategies. Our efforts to deal with these problems led us to the main chain directed (MCD) assignment strategy, which was instrumental in providing the assignments.

Materials and Methods

Samples of cytochromes c from various species were obtained from Sigma Chemical Co. in the highest available grade. Most NMR spectra shown were obtained on a Bruker AM 500 NMR spectrometer located at the Institute for Cancer Research. Early spectra were obtained on Bruker WM 500 NMR spectrometers at Yale University and the University of Washington. Magnitude-mode J-correlated spectra (COSY and RCT COSY) were obtained as previously described (Wand & Englander, 1986). Phase-sensitive NOESY (Macura & Ernst, 1980), DQF COSY (Shaka & Freeman, 1983; Rance et al, 1984), and TOCSY (Braunschweiler & Ernst, 1983; Bax & Davis, 1985) spectra employed time-proportional phase incrementation to provide quadrature detection during the incremented time domain (Redfield & Kuntz, 1975; Marion & Wüthrich, 1983) as described more completely elsewhere (Di Stefano & Wand, 1987). All spectra presented were obtained on samples 5–8 mM in cytochrome c in 50 mM potassium phosphate buffer at a pH meter reading of 5.75 (uncorrected for the isotope effect). Samples were deoxygenated with nitrogen and reduced with solid sodium dithionite or ascorbate. Solvent suppression when required was achieved by direct saturation at all times except during acquisition and benefited from an electromagnetic shield kindly provided by R. Dykstra (Dykstra, 1987).

Results

Assignment Strategy. As with much prior 2D NMR work on proteins, our initial work on the assignment of horse ferrocytochrome c (Wand & Englander, 1985) followed the sequential assignment strategy (Wüthrich et al., 1982; Billeter et al., 1982, Wüthrich, 1983), which focuses initially on the difficult side-chain region of J-correlated spectra. A great deal

[†] This work was supported by NIH Research Grants GM-35940 (A.J.W.), GM-35926 (H.R.), and GM-31847 and AM-11295 (S.W.E.), by Instrumentation Grants RR-02497 (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.
* Address correspondence to these authors.

[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; ppm, parts per million; ppb, parts per billion; RCT, relayed coherence transfer; TOCSY, total correlation spectroscopy.

0006-2960/89/0428-0186$01.50/0

of work has been expended on efforts to assign side-chain and heme proton resonances in cytochrome *c* [e.g., Moore and Williams (1980a–f), Boswell et al. (1982), Eley et al. (1982a,b), Robinson et al. (1983), Moore et al. (1985), Williams et al. (1985) and references cited therein, Redfield and Gupta (1971), and Keller and Wüthrich (1978, 1981)]. During the course of the studies presented here, the basic principles of the main chain directed assignment algorithm became apparent (Englander & Wand, 1987; Di Stefano & Wand, 1987) and were used during the later stages of the assignment.

The data set described earlier (Wand & Englander, 1985, 1986), collected over a 2-year period on three different spectrometers, contained numerous subtle but significant inconsistencies. Consequently, a completely new data set was obtained on a single sample on the same spectrometer. A most helpful step in this work was the use of spectra obtained at two different temperatures, namely, 20 and 40 °C, which relieved much of the ambiguity encountered at any single temperature since many amide NH resonances, especially those H-bonded to water, display some chemical shift sensitivity to temperature. The assignments listed here (Table I) are based wholly on these data obtained for the reduced protein. Ongoing work with oxidized cytochrome *c* helped in some cases to resolve ambiguities in spectra for the reduced protein and ultimately provided extensive confirmatory cross-checks of the assignments presented here. Resonance assignments for oxidized cytochrome *c* and correlations obtained between the reduced and oxidized proteins in magnetization transfer experiments will be reported elsewhere.

Preliminary Side-Chain Identification. Prolonged attempts to identify amino acid side-chain spin systems proved disappointing. Although the great majority of *J*-coupled NH–C_αH–C_βH subspin systems could be identified by use of COSY and RCT COSY spectra, few complex side-chain spin systems could be well-defined by any combination of *J*-correlated spectra acquired for these studies.

In a continuing effort to satisfy the first requirement of the sequential assignment procedure, the definition of amino acid side-chain spin systems, we turned to the identification of spin systems of variant amino acids in six homologous cytochromes *c* by focusing on the small differences in their COSY spectra. This followed much work by Williams and co-workers [e.g., Robinson et al. (1983)], who have identified portions of a number of residues from the differences in one-dimensional spectra of homologous cytochromes *c*. The complete side-chain spin systems of several of these residues could be easily completely in 2D spectra (Wand & Englander, 1985). Detailed comparisons of COSY spectra for horse, cow, rabbit, chicken, dog, and tuna ferrocytochromes *c* further led to the unambiguous identification of 10 residues of the horse protein (Table I). However, the relatively small number of side-chain spin systems that could be identified in this way proved insufficient to support a formal sequential assignment.

Our experience with the assignment of the N-terminal helix (Wand & Englander, 1986) suggested an alternative approach that does not require the early assignment of side chains but initially defines each residue in terms of its main-chain protons and attempts to link these by the use of predefined closed loops of NOE connectivities.

Main Chain Directed Assignments. The main chain directed (MCD) assignment procedure focuses initially on *J*-correlated information that is most readily obtained in COSY-type spectra, specifically the amide NH–C_αH–C_βH subspin system (Englander & Wand, 1987). With reduced cytochrome *c*, examination of phase-sensitive DQF COSY (Figure 1), magnitude-mode COSY, and RCT COSY spectra (Wand & Englander, 1986) allowed the unambiguous definition of 83 complete amide NH–C_αH–C_βH (NAB) *J*-coupled subspin systems. A further five NAB sets had an ambiguity at the β-proton level. Nine of the 12 NH–$C_\alpha H_2$ glycine spin systems could be directly identified. This accounts for 96 of the 100 main-chain sets (NAB or Na) in cytochrome *c*. The four prolines, the remaining three glycines, and a final NAB set representing Lys-72 were found later in the analysis.

The MCD search algorithm is based on combinations of short distances that produce ^{1}H–^{1}H NOEs between NAB sets in common secondary structural elements. Searches are performed, one at a time, for predefined cyclic NOE patterns characteristic of each kind of protein secondary structure. The closed-loop character of the patterns imposes an essential self-checking property on the search.

NOESY spectra for ferrocytochrome *c* (120- and 140-ms mixing time) were searched for the specific combination of NOE interactions defined in the closed-loop helical MCD pattern formed by $NH_i \leftrightarrow C_\beta H_i \leftrightarrow NH_{i+1} \leftrightarrow NH_i$ NOEs [Englander & Wand, 1987; see also Billeter et al. (1982) and Wüthrich et al. (1984)], with the high fidelity double-helical MCD pattern (two joined single loops) as a starting point (Englander & Wand, 1987). A total of six MCD-defined helical segments linking 51 residues were defined. Two of these apparently separate segments comprises a single continuous helix (N-terminal helix) joined by a glycine, which fails to display the required $d_{\beta N}$ NOE. In subsequent searches, no antiparallel or parallel β-sheet patterns (Englander & Wand, 1987) involving residues not already incorporated into MCD-defined helices were found.

Placement of Helical Segments. Five of the six helical segments were placed within the primary sequence of the protein by reference to the relative positions of amino acid spin systems found in *J*-correlated spectra. The long N-terminal helical segment was identified by three known residues (Wand & Englander, 1986). These are the Ala-15 spin system, obtained by comparison of *J*-correlated spectra of homologous cytochromes *c*, the Cys-14 to heme bridge (Wand & Englander, 1985), and a defined glycine at position 6. The C-terminal segment was placed by the complete identification of Asn-103 (Figure 2) and Thr-89 (Wand & Englander, 1985; Figure 1) and by the appearance of an alanine at position 96. The main-chain NOESY cross-peaks that define this segment were shown elsewhere [Figure 4 of Englander and Wand (1987)]. Analogous results led to the placement of an MCD-defined helical segment spanning residues 49–54. One segment was associated with residues 62–70 by the identification of Glu-62 (by species comparison) and Thr-63 (by spin system identification) and by alignment provided by intermediate-range interactions (Figure 4). The NOE connectivities in this region found in the MCD search are shown in Figure 3.

Each helical segment was then examined for additional interresidue NOEs. The main-chain NOEs found are summarized in Figure 4. The sixth helical segment, formally involving three residues, could not be identified at this point in the analysis.

Definition of Irregular Secondary Structure. The search for helical MCD patterns accounted for over half the amino acid NAB sets. The NOE cross-peaks connecting these NH's with β-CH's and especially α-CH's within these helical segments (Figure 4) were sorted out and eliminated from further consideration. The NOESY spectrum must be simplified in

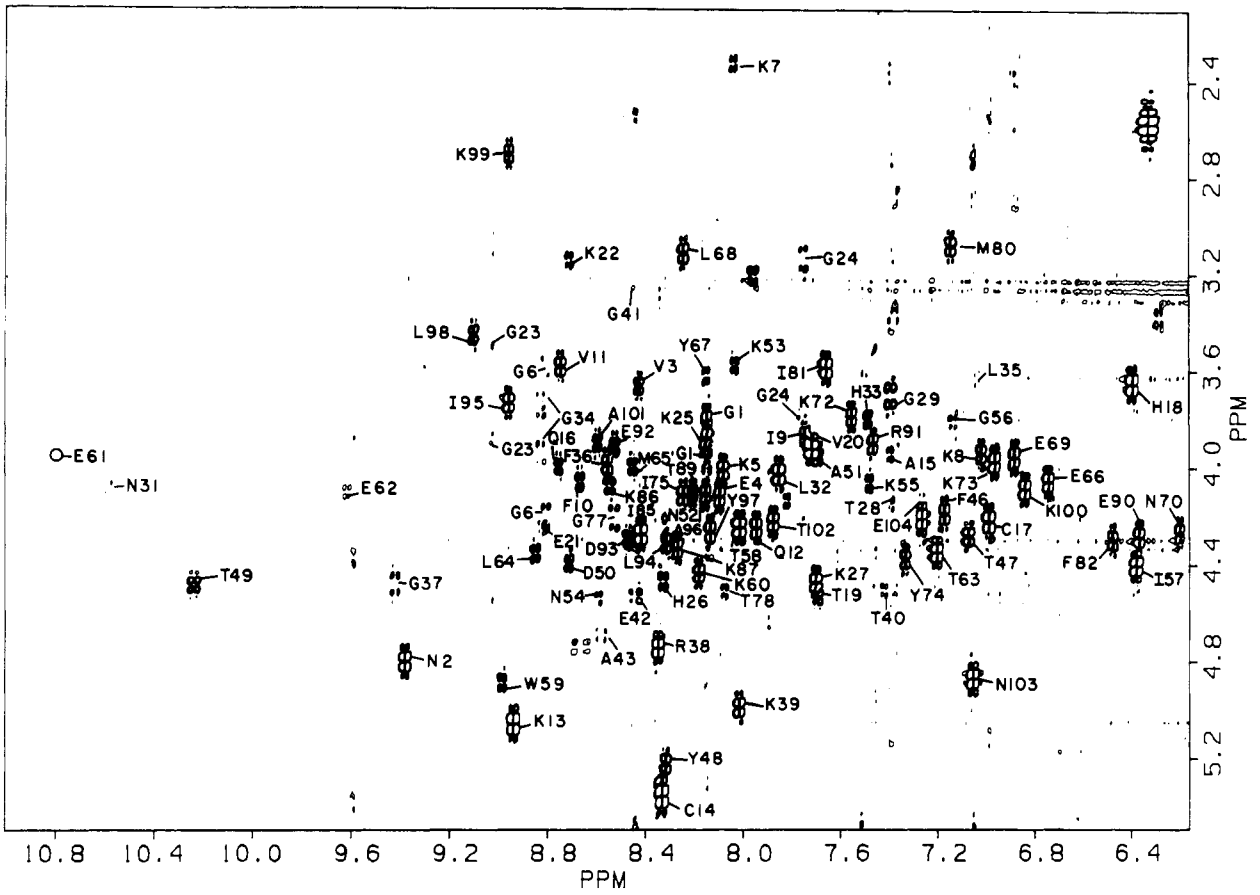

FIGURE 1: Fingerprint region of a DQF COSY of horse ferrocytochrome *c* obtained in 90% H_2O/10% D_2O in 50 mM K_3PO_4 buffer at pH 5.75 and 40 °C. Assigned amide NH–C_αH proton cross-peaks are indicated.

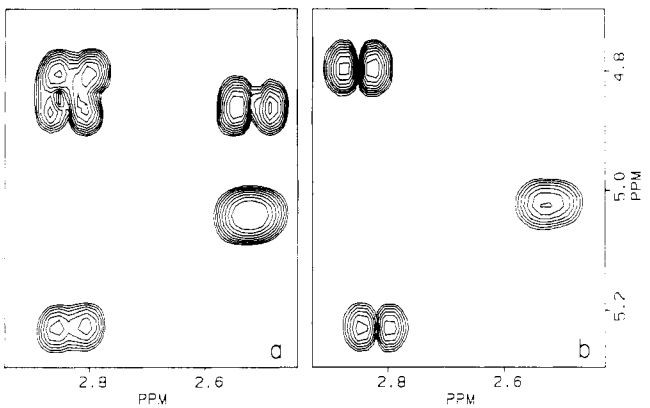

FIGURE 2: Sections of magnitude-mode COSY spectra of (a) horse and (b) chicken ferrocytochromes *c* obtained in D_2O buffer. Other conditions were as described in Figure 1. Horse and chicken cytochromes *c* differ at six positions in the primary sequence. Only one residue change (N103 → S103) involves loss of an AMX_2 spin system. The two $C_\alpha H$–$C_\beta H$ cross-peaks of such a residue, clearly evident in panel a, are absent in panel b and were assigned to N103 in horse ferrocytochrome *c*.

this way in prepartion for the difficult search for disordered chain segments which are characterized only by linear strings of $C_\alpha H_i$–NH_{i+1} NOEs rather than cyclic patterns. In the simplified NOESY spectrum, several stretches of irregular secondary structure could be unambiguously defined, with minimal need for side-chain identification. These groupings include residues 1–5, 55–61, 77–83, 85–88. Residues 22–29 and 31–39 were particularly difficult, primarily due to the sparsity of the interresidue NOEs (Figure 4) and required careful examination of all available information including side-chain definition in *J*-correlated spectra.

Late-Stage Side-Chain Identification. The analysis described produces the NOE linkage map shown in Figure 4. The map provides in advance the identity of the side chain linked to each NAB set. Given this foreknowledge, many confusing long-range correlations in the TOCSY experiment could be reconciled with the limited unambiguous information in the complementary COSY and RCT COSY spectra. Several examples are shown in Figures 5 and 6. The most recently described side-chain assignments of Williams and co-workers (Boswell et al., 1982; Eley et al., 1982a,b; Robinson et al., 1983; Moore et al., 1985) are, for the most part, in agreement with our analysis (Table I).

One particularly interesting set of side-chain–side-chain interactions involving the rings of Phe-46 and Tyr-48 was evident in the NOESY spectra obtained at 20 °C (Figure 7). Both rings are slowly flipping, as indicated by the efficient magnetization transfer between ortho (2, 6) protons and between meta (3, 5) protons of each ring. In addition, NOEs are seen between the ortho protons and the α-proton of Tyr-48 and between the meta proton(s) of Phe-46 and the α-proton of Tyr-48.

The four prolines were provisionally distinguished by NOEs from neighboring amino acids and their spin systems confirmed by analysis of TOCSY spectra. The identification of Pro-71, by NOEs to Asn-70, and of Pro-76, by NOEs to Gly-77, allowed the remaining helical segment mentioned above to be placed at residues 73–75. NOEs to Pro-71 and Lys-73 then pointed to the previously undiscovered NAB set of Lys-72.

The analysis of side-chain NOEs also suggested assignments for several exchangeable side-chain protons. Some of these could be confirmed by *J*-correlated data (e.g., Lys-99 $N_\epsilon H$ by correlation with $C_\epsilon H$), while others remain tentative (e.g.,

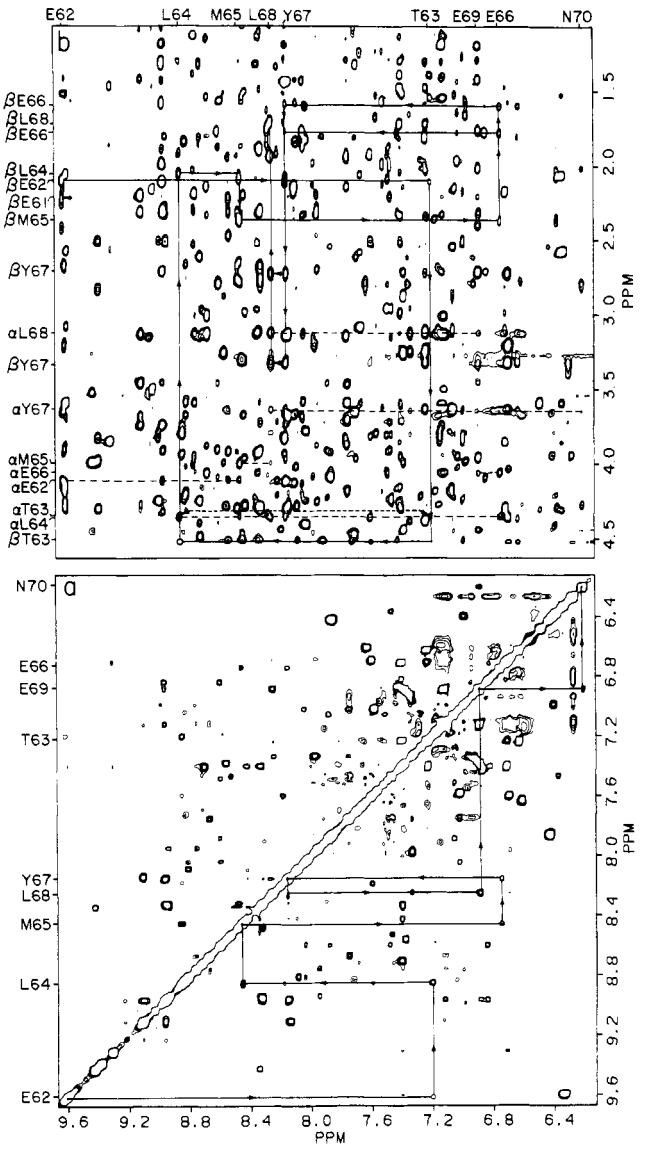

FIGURE 3: Sections of a phase-sensitive NOESY spectrum of horse ferrocytochrome *c* obtained with a mixing time of 120 ms. Other conditions were as described in Figure 1. The basic NOE elements recognized and ordered by the MCD pattern of a helical unit of secondary structure are shown for residues 62–70. Panel a shows the d_{NN} NOE connectivities, which run in parallel to the inter- and intraresidue $d_{\beta N}$ NOEs shown in panel b. Also shown in panel b (dashed lines) are longer range interresidue acid NOEs not used by the current MCD pattern. This figure was prepared from NOESY spectra obtained at 40 °C; some additional NOEs shown in Figure 4 were observed only at 20 °C.

Thr-47 OH). These are listed in Table I.

Discussion

Assignment Methodology. Nearly 120 side-chain and heme resonance assignments for horse ferrocytochrome *c* became available before or during the course of this work [Redfield & Gupta, 1971; Keller & Wüthrich, 1978, 1981; Moore & Williams, 1980a–f; Boswell et al., 1982; Eley et al., 1982a,b; Robinson et al., 1983; Moore et al., 1985; Williams et al. (1985) and references cited therein]. In the early stages of our work (Wand & Englander, 1985), this information was extended by further side-chain assignments obtained by the application of 2D methods and approaches such as species comparisons. Nevertheless, we found it impossible to solve the cytochrome *c* assignment problem by use of the standard sequential assignment strategy in which one seeks to group the protons of individual amino acids into *J*-coupled spin

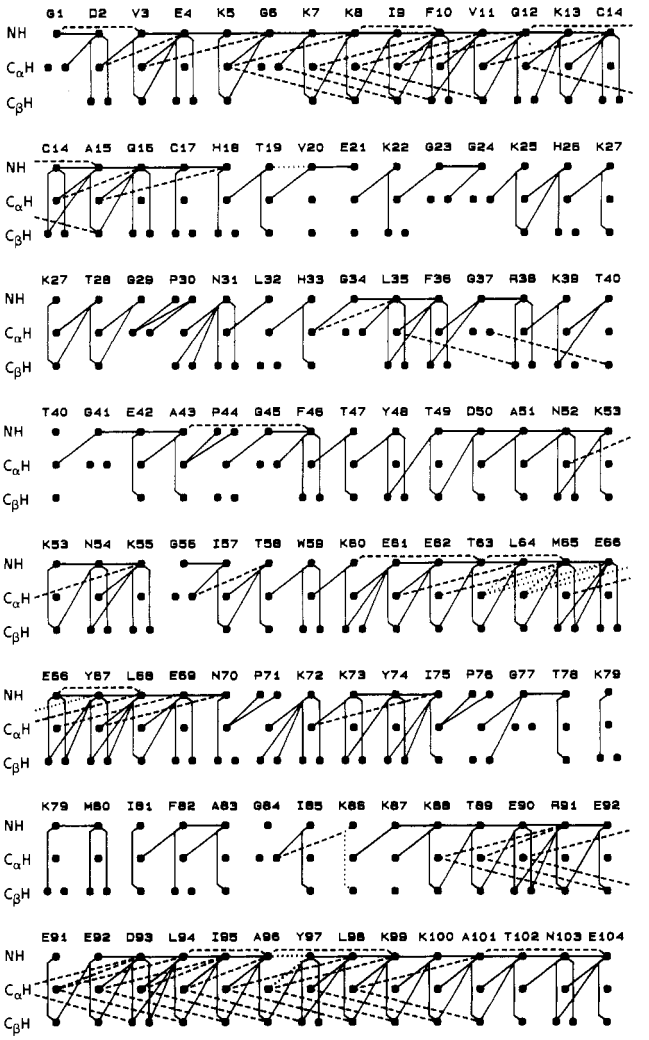

FIGURE 4: NOE connectivity map for backbone and β-carbon protons. Solid lines (—) indicate NOEs within individual residues and between neighboring residues which served as the primary data base for the MCD analysis. Dashed lines (- - -) indicate intermediate-range NOEs within segments of secondary structure. These help to confirm assignments in helical segments and bends. Dotted lines (···) indicate connectivities that are ambiguous due to degeneracy. This connectivity map summarizes data obtained at 20 and 40 °C. For proline, the amide proton position represents NOEs arising from the δ protons. Only NOE cross-peaks having a signal-to-noise ratio greater than 2.5:1 are included.

systems and to recognize the chemical identity of each spin system. To do this successfully, one must consider for any given residue all possible sets of interactions and conditions that may give rise to the great plasticity of cross-peak fine structure and the number and arrangement of cross-peaks on the two-dimensional chemical shift surface. As one turns to larger proteins, problems that further complicate spin system recognition—peak broadening, loss of multiplet structure, resonance overlap—grow more acute.

Experience with the difficulties in defining cytochrome *c* side-chain spin systems in *J*-correlated spectra led us to the formal development of the main chain directed assignment procedure (Englander & Wand, 1987; Di Stefano & Wand, 1987), which requires initially the definition of the simple main-chain NAB spin systems rather than the varied and challenging side-chain systems. The MCD procedure also simplifies and formalizes the NOESY search. At any step in the search one selects a given structure (e.g., helix) and searches only for the predefined companion NOE pattern. All other potential connectivities and patterns are initially ignored.

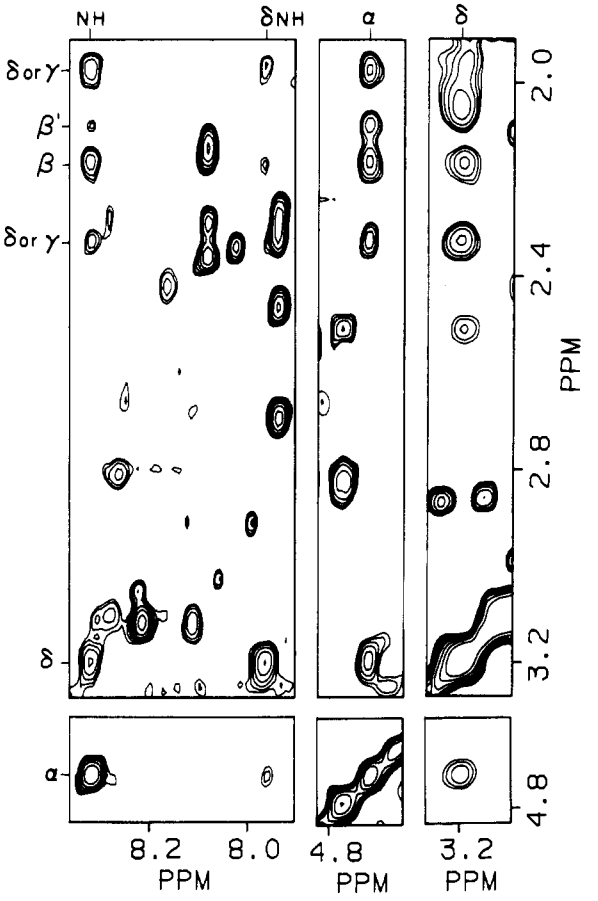

FIGURE 5: Sections of a TOCSY spectrum of horse ferrocytochrome *c*, showing the extensive *J* correlations involving the side-chain spin system of Arg-38. An isotropic mixing time of 75 ms was used. Other conditions were as described in Figure 1. This spin system could be defined only after the identity of the residue was known (thereby allowing the long-range correlations provided by the TOCSY spectrum to be reconciled with the direct correlations provided by DQF COSY spectra).

The great advantage of this can be appreciated by examination of the full main-chain NOE network of a helix (Figure 4). In the presence of considerable ambiguity brought about by chemical shift degeneracy and the absence of reliable rules of simplification, the NOESY spectrum would be extremely difficult to analyze without knowledge of the type of amino acid associated with each of the NOE-connected main-chain protons. This is the point of the sequential assignment procedure, which demands initial side-chain assignments. However, by restricting each search of the NOESY spectrum to a given closed-loop pattern involving only main-chain protons (e.g., $NH_i \leftrightarrow C_\beta H_i \leftrightarrow NH_{i+1} \leftrightarrow NH_i$ in the case of helical secondary structure), the analysis if greatly simplified. Even though protein helices tend to show many NOE interactions in addition to those included in the MCD search pattern (see Figure 4), these are not required for recognition of an ordered collection of residues. The key is to predefine reliable recognition patterns and search for these one at a time, while ignoring the myriad of other possible connectivities. The predefined search patterns should be simple when the NOESY analysis is pursued manually but can be more complex when computer assisted. Such patterns, possessing a cyclic self-checking property, have been proposed and evaluated against a data base of crystallographically derived protein structures (Englander & Wand, 1987; Wand & Nelson, 1988).

For a sizable protein, the assignment problem presents a gigantic puzzle. In attempting to solve it, the operator may make many thousands of tactical moves, some more productive and many less so. The MCD strategy attempts to optimize the efficiency of the solution by suggesting a small subset of all possible moves that one should first try, with the important implication that one should not initially spend time on many other possible moves. The MCD algorithm proposes a beginning game, which uses *J*-correlated spectra, especially COSY and RCT COSY, to obtain the NAB sets and a minimal number of easily solved side-chain spin systems. The NAB sets provide an enhanced target for interresidue NOE connectivities and provide for the cyclic self-checking NOE patterns characteristic of helices, sheets, and bends. The middle game uses the NOESY to establish and place units of secondary structure. The end game returns to *J*-correlated specta to fill in side-chain assignments by taking advantage of the residue identifications established or restricted by the MCD analysis.

The present MCD strategy is a minimalist one. It may well be necessary to take additional steps, especially to deal with spectral regions of high cross-peak density [see Wand and Nelson (1988)]. With reduced cytochrome *c*, we exploited temperature-dependent chemical shifts to help resolve ambiguities in both *J*-correlated and NOESY spectra and used magnetization-transfer methods to obtain confirmatory cross-correlations of assignments in the reduced and oxidized forms of the protein (to be reported elsewhere). Another potentially useful step is to selectively window in subsets of amide NH connectivities by manipulation of ^1H–^2H exchange protocols. The various amide NH's of proteins exchange over a time scale about 10 orders of magnitude in breadth (13 in ferrocytochrome *c*). The ^1H–^2H exchange reaction therefore can effectively provide yet another dimension for resolving amide NH's and their *J*-correlated and NOE cross-peaks. Selective hydrogen isotope incorporation that exploits intrinsic rate differences among different H-exchanging sites has been referred to as kinetic labeling (Wagner et al., 1981; Calhoun & Englander, 1985), and as functional labeling when one utilizes changes in H-exchange rate due to experimental manipulation of the protein structure [e.g., Louie et al. (1988)].

It is interesting to note that cytochrome *c* provides a relatively poor case for the application of the main chain directed assignment procedure. Its regular secondary structure is all helical, with no β-sheet structure, and half its residues are found in irregular secondary structure which provides the lowest fidelity and frequency parameters among all MCD patterns (Englander & Wand, 1987). Further, these characteristics tend to increase spectral congestion due to the relatively small chemical shift disperison of helical amide protons and α-protons, which makes the definition of even the helical MCD patterns more difficult than usual. Nevertheless, good success was obtained (Figure 4).

Structure. Although cytochromes of the *c* type have been widely studied by crystallography [e.g., Takano and Dickerson (1981a,b), Matsuura et al. (1982), and Finzel et al. (1985)], the structure available for horse cytochrome *c* is relatively poor (Dickerson et al., 1971). Consequently, most structural interpretations of studies using horse cytochrome *c* have been based on the higher resolution structures available for the tuna protein (Takano & Dickerson, 1981a,b). The low-resolution crystal structure of horse cytochrome *c* indicates several minor but potentially significant differences from the tuna protein.

The structural information contained in the NOE map presented in Figure 4 provides only a semiquantitative description of the secondary structure of the horse protein in solution. With this qualification in mind, several interesting

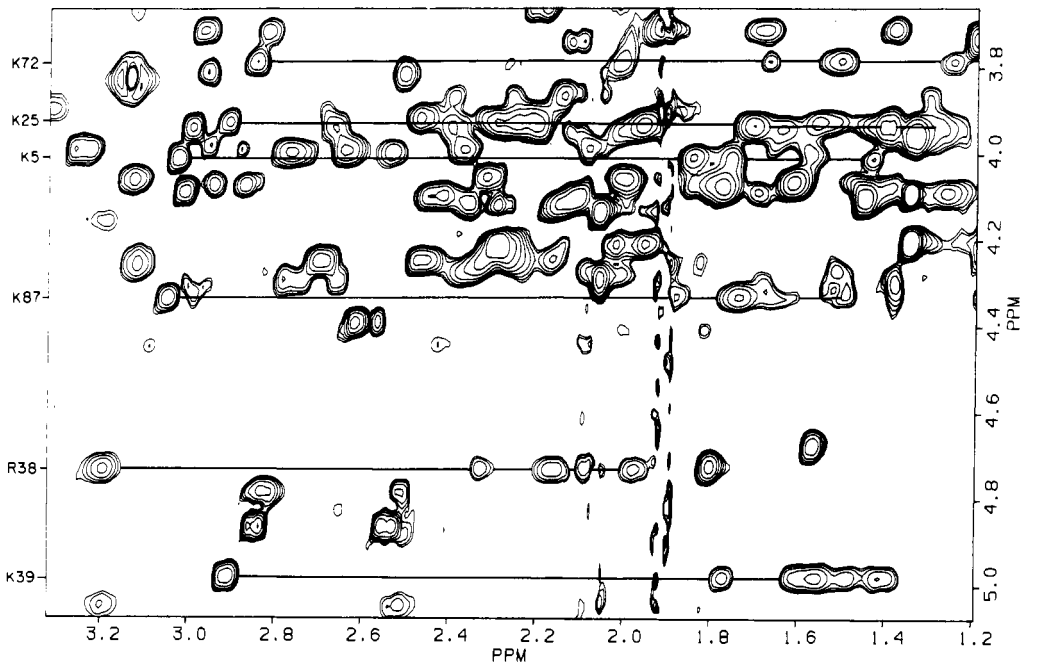

FIGURE 6: Sections of a TOCSY spectrum of horse ferrocytochrome *c*. Conditions are indicated in Figure 5. The extensive *J* correlations involving the side-chain spin systems of lysine-5, -25, -39, -72, and -87 and arginine-38 are shown. These spin systems could be defined only after the identity of each residue was known (thereby allowing the long-range correlations provided by the TOCSY spectrum to be reconciled with the direct correlations provided by DQF COSY spectra).

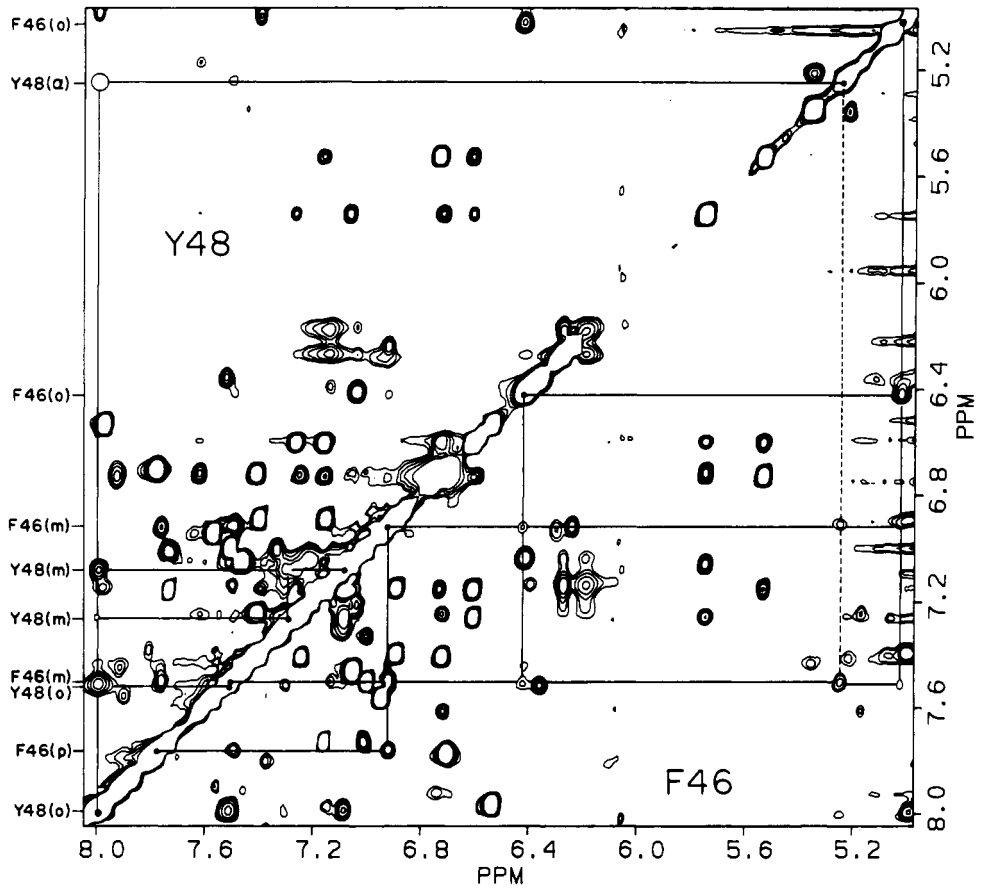

FIGURE 7: Expansions of a NOESY spectrum of horse ferrocytochrome *c* obtained at 20 °C in D_2O buffer. Other conditions were as defined in Figure 1. A mixing time of 120 ms was used. Indicated are cross-peaks of Phe-46 and Tyr-48 arising from intraresidue NOEs and/or magnetization transfer due to ring flipping on the slow-exchange time scale. The dashed line indicates an interresidue NOE between the meta proton of Phe-46 and the $C_\alpha H$ proton of Tyr-48.

structural observations can be made.

The N-terminal α-helix, as in the crystal structure, appears to encompass residues 5–15 with some distorted helical character being maintained to residue 3. Residues 61–70 have been described in the crystal structure of horse ferricytochrome *c* as three sequential 3_{10} bends (Dickerson et al., 1971). The pattern of main-chain NOEs (Figure 4) is more consistent with a more regular α-helical conformation from residue 61–70

Table I: Chemical Shifts of Assigned Proton Resonances of Horse Ferrocytochrome *c*[a]

residue	chemical shift (ppm)					
	NH	$C_\alpha H$	$C_\beta H$	other		
Gly-1	8.15 (8.30)	3.97, 3.83		(2.07) *N*-acetyl		
Asp-2	9.37 (9.43)	4.80	2.82, 2.51			
Val-3	8.42 (8.61)	3.66	2.19[c]	1.09[c] $C_\gamma H_3$		
Glu-4	8.09 (8.17)	4.12	2.11, 2.18	2.36, 2.29 $C_\gamma H$		
Lys-5	8.08 (8.18)	4.01	1.84	1.59, 1.44 $C_\gamma H$	1.76 $C_\delta H$	3.02 $C_\epsilon H$
Gly-6	8.80 (8.89)	4.20, 3.58				
Lys-7	8.03 (8.14)	2.33	1.78	1.62, 1.42 $C_\gamma H$	0.90 $C_\delta H$	
Lys-8	7.01 (7.01)	3.94	1.96	1.43 $C_\gamma H$	1.70 $C_\delta H$	2.99 $C_\epsilon H$
Ile-9	7.74 (7.72)	3.87	2.14	1.30 $C_\gamma H$	1.17 $C_\gamma H_3$	
Phe-10	8.66 (8.72)	4.06	3.12, 3.00	7.17, 7.18 C_2H,C_6H	6.98, 6.12 C_3H,C_5H	6.21 C_4H
Val-11	8.74 (8.82)	3.58	2.19[c]	1.14,[c] 0.97[c] $C_\gamma H_3$		
Gln-12	7.94 (7.98)	4.24	2.35, 2.28	2.71, 2.46 $C_\gamma H$		
Lys-13	8.93 (9.01)	5.05	2.52	1.80 $C_\gamma H$	2.05, 2.32 $C_\delta H$	3.20 $C_\epsilon H$
Cys-14	8.32 (8.39)	5.35	1.14, 1.88			
Ala-15[b]	7.39 (7.44)	3.94[c]	1.40[c]			
Gln-16	8.75 (8.87)	3.99	2.29, 2.06	2.74, 2.52 $C_\gamma H$		
Cys-17	6.99 (7.03)	4.23	0.61, 1.53			
His-18	6.39 (6.41)	3.65	0.79, 1.08	0.41[e] C_2H	0.13[e] C_5H	9.61 $N_\epsilon H$
Thr-19	7.69 (7.86)	4.53	4.46[c]	1.10[c] $C_\gamma H_3$		
Val-20	7.71 (7.79)	3.92	1.56[c]	0.41,[c] 0.04[c] $C_\gamma H_3$		
Glu-21	8.82 (8.98)	4.23	2.22			
Lys-22	8.71 (8.90)	3.15	1.34 (0.58)			
Gly-23	9.04 (9.24)	3.55, 3.86				
Gly-24	7.75 (7.81)	3.13, 3.82				
Lys-25	8.15 (8.33)	3.92	1.73, 1.51	1.44 $C_\gamma H$	1.28 $C_\delta H$	2.90 $C_\epsilon H$
His-26	8.33 (8.53)	4.46	2.89, 3.10	7.51[d] C_2H	7.05[d] C_5H	
Lys-27	7.70 (7.79)	4.47	1.14			
Thr-28[b]	7.39 (7.91)	4.13	4.30[c]	2.06[c] $C_\gamma H_3$		
Gly-29	7.39 (7.45)	3.70, −0.17				
Pro-30		3.61	0.40, 1.26	0.57, 1.45 $C_\gamma H$	1.44,[h] 1.01 $C_\delta H$	
Asn-31	10.58 (10.66)	4.09	1.99, 2.12			
Leu-32	7.85 (7.89)	4.03	1.46, (1.10)	0.53[c] $C_\gamma H$	−0.76,[c] −0.61[c] $C_\delta H_3$	
His-33	7.49 (7.56)	3.80	2.95	8.37[d] C_2H	7.46[d] C_5H	
Gly-34	8.82 (8.93)	3.93, 3.74				
Leu-35	7.03 (7.05)	3.62	2.15, 1.52	1.28[c] $C_\gamma H$	0.69,[c] 0.73[c] $C_\delta H_3$	
Phe-36	8.55 (8.71)	3.99	3.24, 2.87	7.37[d] C_2H,C_6H	6.88[d] C_3H,C_5H	7.12[d] C_4H
Gly-37	9.41 (9.53)	3.51, 4.48				
Arg-38	8.33 (8.37)	4.74	2.17, 2.09	2.32, 1.97 $C_\gamma H$	3.20 $C_\delta H$	7.97 $N_\delta H$ 6.68 $N_\omega H$
Lys-39	8.01 (8.15)	4.99	1.61, 1.78	1.56, 1.51 $C_\gamma H$	1.47, 1.42 $C_\delta H$	2.91 $C_\epsilon H$
Thr-40	7.42 (7.39)	4.50	4.46[c]	0.89[c] $C_\gamma H_3$		
Gly-41	8.43 (8.62)	3.29, 1.56				
Gln-42	8.45 (8.54)	4.53	1.88			
Ala-43	8.57 (8.65)	4.68[c]	1.57[c]			
Pro-44		4.11	1.95 (2.42)	2.14, 2.26 $C_\gamma H$	3.87, 4.20 $C_\delta H$	
Gly-45	(9.18)	4.37, 3.83				
Phe-46	7.17 (7.19)	4.17	2.35, 1.12	(6.42), (5.01) C_2H,C_6H	(6.91),[e] (7.49)[e] C_3H,C_5H	(7.75)[e] C_4H
Thr-47[b]	7.07 (7.32)	4.29	3.77[c]	1.12[c] $C_\gamma H_3$		9.50[g] $O_\delta H$
Tyr-48	8.31 (8.53)	5.23	2.81, 3.72	7.99, 7.51 C_2H,C_6H	7.09, 7.29 C_3H,C_5H	8.69[g] $O_\beta H$
Thr-49	10.23 (10.31)	4.48	4.73[c]	1.80[c] $C_\gamma H_3$		
Asp-50	8.71 (8.85)	4.39	2.61			
Ala-51	7.68 (7.86)	3.96[c]	1.33[c]			
Asn-52	8.32 (8.46)	4.23	3.12, (2.98)			
Lys-53	8.02 (8.10)	3.58	1.88	1.56 $C_\gamma H$		
Asn-54	8.59 (8.75)	4.54	2.77, 2.79			
Lys-55	7.48 (7.49)	4.06	2.09, 1.91			
Gly-56	7.13 (7.14)	3.78, 3.81				
Ile-57	6.38 (6.39)	4.40	1.83[c]	0.89,[f] 0.50[f] $C_\gamma H$	0.73[c] $C_\gamma H_3$	−0.53[c] $C_\delta H_3$
Thr-58[b]	8.01 (8.17)	4.25	3.77	0.99 $C_\gamma H_3$		
Trp-59	8.98 (9.02)	4.88	3.81, 2.51	7.59[f] C_4H 6.68[f] C_5H	5.73[f] C_6H 7.05[f] C_7H	10.03 NH 6.99[f] C_2H
Lys-60[b]	8.17 (8.21)	4.43	2.11, 2.45	1.53, 1.67 $C_\gamma H$	1.75, 1.86 $C_\delta H$	3.01, 3.08 $C_\epsilon H$
Glu-61	10.80 (10.85)	3.94	2.25 (2.70)			
Glu-62[b]	9.61 (9.69)	4.10	2.11	2.43 $C_\gamma H$		
Thr-63	7.20 (7.23)	4.35	4.52[c]	1.39[c] $C_\gamma H_3$	5.17[g] $O_\beta H$	
Leu-64	8.84 (8.89)	4.35	2.06	1.91[c] $C_\gamma H$	0.74, 0.51[c] $C_\delta H_3$	
Met-65	8.45 (8.57)	3.99	2.37, 2.09	2.76, 2.64	2.11[d] $C_\epsilon H_3$	
Glu-66	6.74 (6.75)	4.06	1.60, 1.77	2.00, 2.30 $C_\gamma H$		
Tyr-67	8.15 (8.27)	3.62	3.31, 2.72			
Leu-68	8.24 (8.34)	3.12	1.71	1.94[c] $C_\gamma H$	0.36,[c] 1.07[c] $C_\delta H_3$	
Glu-69	6.87 (6.91)	3.95	2.00, 1.76	2.20, 2.11 $C_\gamma H$		
Asn-70	6.20 (6.24)	4.27	2.82, 2.80			
Pro-71		3.92	1.42, 0.92	0.12, 0.74 $C_\gamma H$	3.08, 2.82 $C_\delta H$	
Lys-72	7.52 (7.73)	3.79	1.54, 1.68	1.25 $C_\gamma H$	2.84 $C_\epsilon H$	
Lys-73	6.96 (7.01)	3.97	1.66, 1.57			

Table I (Continued)

residue	chemical shift (ppm) NH	$C_\alpha H$	$C_\beta H$	other		
Tyr-74	7.33 (7.32)	4.37	3.12, 3.24	7.25[f] C_2H,C_6H	6.60[f] C_3H,C_5H	
Ile-75	8.24 (8.34)	4.11	1.90[c]	1.81, 1.48 $C_\gamma H$	0.64[c] $C_\gamma H_3$	0.99 $C_\gamma H_3$
Pro-76		4.51	2.20, 1.75	1.96 $C_\gamma H$	3.37, 3.17 $C_\delta H$	
Gly-77	8.52 (8.71)	4.21, 3.62				
Thr-78	8.07 (8.16)	4.52	4.27[c]	0.81[c] $C_\gamma H$	8.40[g] $O_\beta H$	
Lys-79	7.90 (7.97)	4.63	(1.96), (2.32)			
Met-80	7.14 (7.14)	3.10	−0.21, −2.59	−3.75, −1.88 $C_\gamma H$	−3.28[e] $C_\epsilon CH_3$	
Ile-81	7.65 (7.97)	3.58	2.07[c]	1.34, 1.04 $C_\gamma H$	0.79[c] $C_\gamma H_3$	0.73 $C_\delta H_3$
Phe-82	6.46 (6.59)	4.30	0.61, 2.18	6.70[d,e] C_2H,C_6H	7.49[d,e] C_3H,C_5H	7.23[d,e] C_4H
Ala-83	8.15 (8.27)	3.62[c]	1.10[c]			
Gly-84		4.27, 2.98				
Ile-85	8.40 (8.48)	4.27	1.51[c]	1.32, 1.72 $C_\gamma H$	1.06[c] $C_\gamma H_3$	1.01 $C_\delta H_3$
Lys-86	8.54 (8.69)	4.07	1.84			
Lys-87	8.26 (8.48)	4.33	1.77, 1.74	1.88, 1.50 $C_\gamma H$	3.05 $C_\epsilon H$	
Lys-88	(9.08)	3.71	1.92			
Thr-89[b]	8.20 (8.30)	4.11	4.21[c]	1.35[c] $C_\gamma H_3$		
Glu-90	6.36 (6.36)	4.28	2.01, 2.08	2.37 $C_\gamma H$		
Arg-91	7.46 (7.51)	3.90	2.20	2.48 $C_\gamma H$	3.31, 3.37 $C_\delta H$	6.29 $N_\epsilon H$, (6.95) $N_\omega H$
Glu-92[b]	8.52 (8.61)	3.91	2.31	2.58 $C_\gamma H$		
Asp-93	8.46 (8.53)	4.30	2.76, 2.66			
Leu-94	8.31 (8.41)	4.31	2.30	1.90[c] $C_\gamma H$	1.55, 1.50[c] $C_\delta H$	
Ile-95	8.95 (9.03)	3.73	2.11	1.98 $C_\gamma H$	1.20[c] $C_\gamma H_3$	0.97 $C_\delta H_3$
Ala-96	8.15 (8.18)	4.11[c]	1.45[c]			
Tyr-97	8.13 (8.20)	4.26	3.12, 3.67	6.54,[c] 7.08[e] C_2H,C_6H	5.46,[e] 6.67[e] C_3H,C_5H	
Leu-98	9.10 (9.14)	3.46	2.32	2.20[c] $C_\gamma H$	1.11, 0.69 $C_\delta H_3$	
Lys-99	8.95 (9.07)	2.71	1.59, 1.32	0.77, 0.44 $C_\gamma H$	3.28 $C_\delta H$	2.56, 3.62 $C_\epsilon H$, 7.07 $N_\epsilon H$
Lys-100[b]	6.83 (6.86)	4.09	1.77	1.25, 1.68 $C_\gamma H$	1.39 $C_\delta H$	3.01 $C_\epsilon H$
Ala-101	8.59 (8.68)	3.90[c]	0.57[c]			
Thr-102	7.87 (7.98)	4.22	4.69[c]	1.03[c] $C_\gamma H_3$		
Asn-103[b]	7.05 (7.08)	4.87	2.84, 2.54			
Glu-104	7.26 (7.25)	4.22	1.95 (2.02)	2.29 $C_\gamma H$		

[a] Chemical shifts (ppm) are referenced to external (coaxial capillary) sodium 3-(trimethylsilyl)tetradeuteriopropionate and are quoted to 0.01 ppm. For horse ferrocytochrome *c* in 50 mM potassium phosphate in 90% H_2O/10% D_2O, pH 5.75, at 40 °C. Values in parentheses are for 20 °C. [b] Spin system unequivocally identified by species comparison and used for placement of MCD-defined helices within the polypeptide sequence. [c] In agreement with Moore et al. (1985). [d] In agreement with Boswell et al. (1982). [e] In agreement with Robinson et al. (1983). [f] In agreement with Eley et al. (1982a,b). [g] Based on NOE only. [h] In agreement with Santos and Turner (1987).

owing to the noticeable absence of $C_\alpha H_i \leftrightarrow NH_{i+2}$ NOEs characteristic of 3_{10} helices (Wüthrich, 1986).

The C-terminal segment displays NOEs corresponding to an α-helical conformation from Lys-88 to Ala-103. This region, as also suggested by the crystal structure (Dickerson et al., 1971), is the most α-helical in nature; it displays the most comprehensive and regular sets of sequential long-range NOEs (e.g., $C_\alpha H_i \leftrightarrow C_\beta H_{i+3}$). The appearance of strong and self-consistent main-chain NOEs involving residues Thr-102, Asn-103, and Glu-104 suggests a relatively stable but apparently nonhelical conformation through the C-terminus. This is in contrast to the poor definition of this region in the crystalline protein (Dickerson et al., 1971).

For the most part, the remainder of the protein displays NOEs consistent with the horse crystal structure. NOEs characteristic for the central residues of type II turns (Wüthrich et al., 1984) are observed for residues 23–24, 35–36, and 76–77, as found in the crystal (Dickerson et al., 1971). All four prolines show NOEs consistent with a trans configuration, including Pro-44 which is absent in tuna cytochrome *c*.

The temperature sensitivity of amide NH resonances is often a useful marker for hydrogen bonding. The temperature coefficients of amide proton resonances of random-coil peptides fall within the range of −6 to −10 ppb/K (Deslauriers & Smith, 1980). In general, the hydrogen bonding predicted by the observed NOE patterns correlates with the observed reduced temperature dependence of the hydrogen-bonded amide NH resonances.

Several exchangeable side-chain protons have been assigned (Table I). That these protons are observable indicates slow exchange with solvent and therefore their probably participation in hydrogen bonding (Englander & Kallenbach, 1984). Most of these appear H-bonded in the crystal structure of tuna ferrocytochrome *c* (Takano & Dickerson, 1981a). Some are not (e.g., Lys-99 $N_\epsilon H$). To obtain more definitive information concerning these and other side-chain conformations and interactions, rigorous distance geometry studies (Braun, 1987) are required. Such studies are in progress and will be reported elsewhere.

ACKNOWLEDGMENTS

We thank Dennis R. Hare, Gary P. Drobny, David E. Wemmer, and Brian R. Reid (University of Washington, Seattle, WA) and Peter Demou and James H. Prestegard (Yale University, New Haven, CT) for their help in obtaining preliminary data.

Registry No. Ferrocytochrome *c*, 9007-43-6.

REFERENCES

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

Bechtold, T., 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., Eley, C. G. S., Moore, G. R., Robinson, M. N., Williams, G., & Williams, R. J. P. (1982) *Eur. J. Biochem. 124*, 289–294.

Braun, W. (1987) *Q. Rev. Biophys. 19*, 115–157.

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

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

Deslauriers, R., & Smith, I. C. P. (1980) in *Biological Magnetic Resonance* (Berliner, L. J., & Reuben, J., Eds.) Vol. 2, pp 243–344, Plenum, New York.

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

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

Dykstra, R. (1987) *J. Magn. Reson. 72*, 162–167.

Eley, C. G. S., Moore, G. R., Williams, R. J. P., Neupert, W., Boon, P. J., Brinckhof, H., Nivard, R. J. F., & Tesser, G. I. (1982a) *Biochem. J. 205*, 153–165.

Eley, C. G. S., Moore, G. R., Williams, G., & Williams, R. J. P. (1982b) *Eur. J. Biochem. 124*, 295–303.

Englander, S. W., & Kallenbach, N. R. (1984) *Q. Rev. Biophys. 16*, 521–655.

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

Englander, S. W., Roder, H., & Wand, A. J. (1987) in *Protein Structure: Molecular and Electronic Reactivity* (Austin, R., Buhks, E., Chance, B., DeVault, D., Dutton, P. L., Fraunfelder, H., & Goldanski, V. I., Ed.) pp 139–153, Springer-Verlag, Berlin.

Finzel, B. C., Weber, P. C., Hardman, K. D., & Salemme, F. R. (1985) *J. Mol. Biol. 186*, 627–643.

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

Keller, R. M., & Wüthrich, K. (1981) *Biochim. Biophys. Acta 668*, 307–320.

Louie, G., Tran, T., Englander, J. J., & Englander, S. W. (1988) *J. Mol. Biol. 201*, 755–764.

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.

Matsuura, Y., Takano, T., & Dickerson, R. E. (1982) *J. Mol. Biol. 156*, 389–409.

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

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.

Moore, G. R., & Williams, R. J. P. (1980c) *Eur. J. Biochem. 103*, 513–521.

Moore, G. R., & Williams, R. J. P. (1980d) *Eur. J. Biochem. 103*, 523–532.

Moore, G. R., & Williams, R. J. P. (1980e) *Eur. J. Biochem. 103*, 533–541.

Moore, G. R., & Williams, R. J. P. (1980f) *Eur. J. Biochem. 103*, 543–550.

Moore, G. R., Robinson, M. N., Williams, G., & Williams, R. J. P. (1985) *J. Mol. Biol. 183*, 429–446.

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

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. K. (1971) *Cold Spring Harbor Symp. Quant. Biol. 36*, 405–411.

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

Robinson, M. N., Boswell, A. P., Huang, Z.-X., Eley, C. G. S., & Moore, G. R. (1983) *Biochem. J. 213*, 687–700.

Santos, H., & Turner, D. L. (1987) *FEBS Lett. 226*, 179–185.

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

Takano, T., & Dickerson, R. E. (1981a) *J. Mol. Biol. 153*, 79–94.

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

Wagner, G., Anil-Kumar, & Wüthrich, K. (1981) *Eur. J. Biochem. 114*, 375–384.

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., & Nelson, S. J. (1988) *Trans. Am. Crystallogr. Assoc.* (in press).

Wand, A. J., Roder, H., & Englander, S. W. (1986) *Biochemistry 25*, 1107–1114.

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. (1983) *Biopolymers 22*, 131–138.

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

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