# Solution $^1$H NMR Determination of Secondary Structure for the Three-Iron Form of Ferredoxin from the Hyperthermophilic Archaeon *Pyrococcus furiosus*†

Quincy Teng,‡ Zhi Hai Zhou,§ Eugene T. Smith,§ Scott C. Busse,‡ James B. Howard,∥ Michael W. W. Adams,§ and Gerd N. La Mar*,‡

*Department of Chemistry, University of California, Davis, California 95616, Department of Biochemistry, University of Georgia, Athens, Georgia 30602, and Department of Biochemistry, University of Minnesota, Minneapolis, Minnesota 55455*

*Received November 23, 1993; Revised Manuscript Received March 15, 1994*®

ABSTRACT: Two-dimensional $^1$H NMR data have been used to make sequence-specific assignments and define the secondary structure of the three-iron form of the oxidized ferredoxin, Fd, from the hyperthermophilic archaeon *Pyrococcus furiosus*, *Pf*. Signals for at least some protons were located for 65 of the 66 amino acids in the sequence, in spite of the paramagnetic ($S = 1/2$) ground state, but not all could be assigned. Unassigned and missing signals could be qualitatively correlated with the expected proximity of the protons to the paramagnetic cluster. The secondary structure was deduced from qualitative analysis of the 2D nuclear Overhauser effect, which identified two antiparallel β-sheets, one triple-stranded including Ala1–Ser5, Val39–Glu41, and Thr62–Ala66, and one double-stranded consisting of Glu26–Asn28 and Lys32–Glu34, as well as an α-helix involving Glu43–Glu54. Three tight type I turns are located at residues Asp7–Thr10, Pro22–Phe25, and Asp29–Gly31. Comparison with the crystal structure of *Desulfovibrio gigas*, *Dg*, Fd (Kissinger *et al.*, 1991) reveals a very similar folding topology, although several secondary structural elements are extended in *Pf* relative to *Dg* Fd. Thus the β-sheet involving the two termini is expanded to include the two terminal residues and incorporates a third strand from the internal loop that is lengthened by several insertions in *Pf* relative to *Dg* Fd. The double-stranded β-sheet in the interior of *Pf* Fd is lengthened slightly due to a much tighter type I turn between the two strands. The helix near the C-terminus is three residues longer in *Pf* than in *Dg* Fd, as well as being shifted toward the N-terminus. The disulfide link between the two nonligating Cys residues (Cys21 and Cys48) is conserved in *Pf* Fd, but the link near the C-terminus is in the middle of the long α-helix in *Pf* Fd, instead of at the N-terminus of the helix as in *Dg* Fd. The extensions of the β-sheets and α-helix increase the number of main-chain hydrogen bonds in *Pf* Fd by approximately 8 relative to those in *Dg* Fd and likely contribute to its remarkable thermostability (it is unaffected by anaerobic incubation at 95 °C for 24 h). Qualitative studies of a sample in $^2H_2O$ reveal some 20 slowly exchanging labile protons that are located in the various identified secondary structural elements, except for two protons whose relaxation properties dictate that they arise from residues whose peptide NHs participate in hydrogen bonding interaction with the coordinated sulfur atoms of the cluster.

During the past decade, microorganisms have been isolated from sulfide-rich marine environments with the remarkable property of thriving at temperatures near or above 100 °C. Virtually all of these so-called "hyperthermophiles" are classified as Archaea (formerly Archaebacteria) (Stetter *et al.*, 1990; Adams *et al.*, 1992; Adams, 1993a), and many are strictly anaerobic, sulfur-reducing organisms that may be related to early life forms (Woese *et al.*, 1990). However, we know little about the potentially unique biochemistry that affords growth under such extreme conditions. In particular, we are relatively unsophisticated in understanding the factors that stabilize the functional three-dimensional structures of proteins at these elevated temperatures. The best studied of the hyperthermophilic archaea is *Pyrococcus furiosus*, *Pf*,[1] which grows optimally at 100 °C (Fiala & Stetter, 1986), and several of its enzymes and proteins have been purified (Adams, 1993, 1994). These include non-heme iron-containing redox proteins such as rubredoxin, Rd (Blake *et al.*, 1991), which has a single iron ligated by four cysteinyl residues, and ferredoxin, Fd (Aono *et al.*, 1989; Park *et al.*, 1991), which contains a single $[Fe_4S_4]^{2+,1+}$ cluster. Besides its stability, *Pf* Fd is distinguished from most other Fds by coordination to the cluster of three rather than four cysteines and by its facile conversion *in vitro* to a form containing a $[Fe_3S_4]^{1+,0}$ cluster (Conover *et al.*, 1990). These extremely thermostable monomeric proteins comprise only about 60 residues and hence provide ideal examples for high-resolution structural studies. Moreover, the crystal structures of several mesophilic Fds (Adman *et al.*, 1977; Watenpaugh *et al.*, 1979; Kissinger *et al.*, 1991; Fukuyama *et al.*, 1988, 1989) and Rds (Adman *et al.*, 1973) have been solved at atomic-level resolution, which provides valuable comparisons for elucidating secondary and tertiary structural elements responsible for hyperthermostability.

† This research was supported by grants from the National Science Foundation, DMB91-04018 (G.N.L.), DMB91-05150 (M.W.W.A.), and DMB91-20515 (J.B.H.), and the National Institutes of Health, GM 45597 (M.W.W.A.).

* Address correspondence to this author at Department of Chemistry, University of California, Davis, CA 95616.

‡ University of California.

§ University of Georgia.

∥ University of Minnesota.

® Abstract published in *Advance ACS Abstracts,* April 15, 1994.

[1] Abbreviations: NMR, nuclear magnetic resonance; 2D, two dimensional; DSS, 2,2-dimethyl,2-silapentane-5-sulfonate; Fd, ferredoxin; Rb, rubredoxin; Hipip, high-potential iron–sulfur protein; *Pf*, *Pyrococcus furiosus*; *Dg*, *Desulfovibrio gigas*; *Bt*, *Bacillus thermoproteolyticus*; TOCSY, total correlation spectroscopy; DQF-COSY, double quantum filtered correlation spectroscopy; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser spectroscopy.

0006-2960/94/0433-6316$04.50/0 

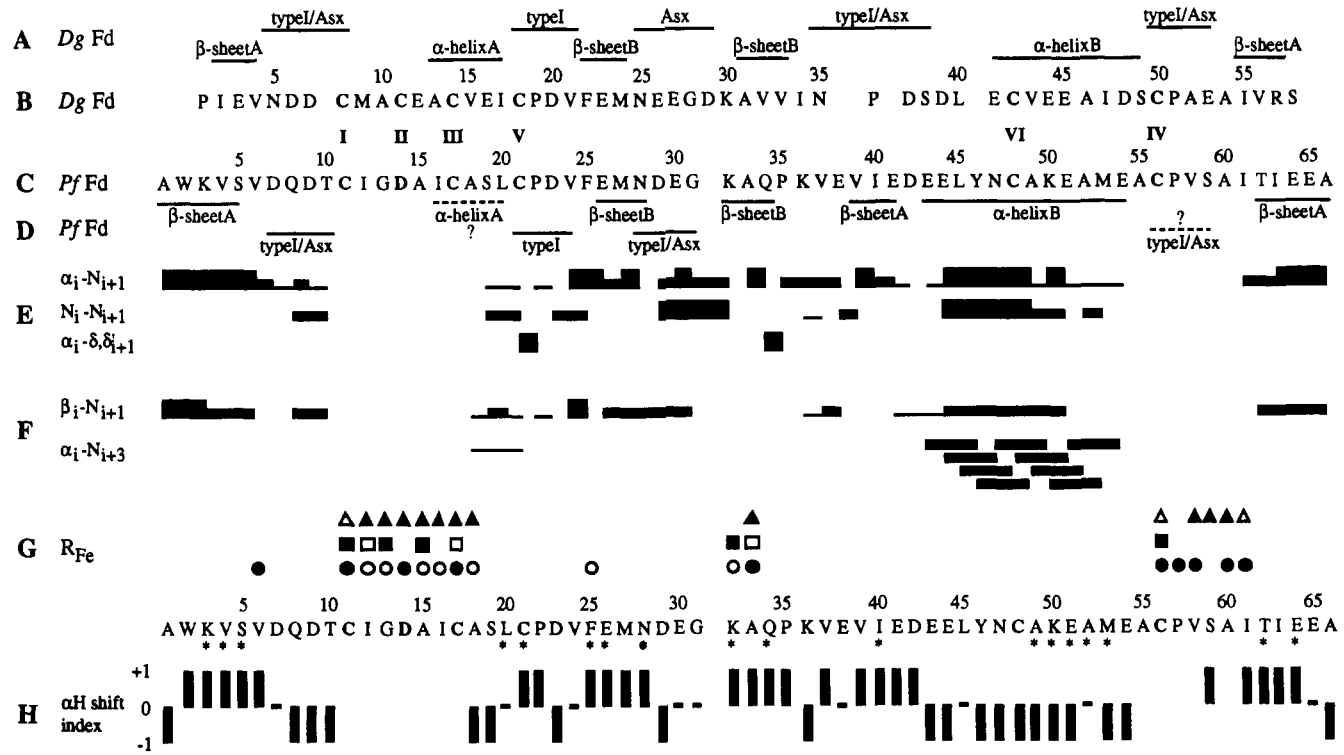


FIGURE 1: (A) Secondary structural elements identified in the crystal structure of *Dg* Fd. (B, C) Amino acid sequences of *Pf* and *Dg* Fd. The alignment is made on the basis of the position of invariant Cys residues; these are labeled I–IV for the ligating ligands in the 4Fe forms (note that position II in *Pf* is Asp, not Cys), and Cys V and VI are linked by a disulfide bridge. (D) Secondary structural elements and their deduced *Pf* Fd locations. (E, F) NOEs used in making the sequence-specific assignments. (G) The portion of a residue (○ side chain; □, αH; △, NH) that is close to a cluster iron (<5.0 Å, closed symbol; <6.0 Å open symbol) based on the *Dg* Fd crystal structure (Kissinger *et al.*, 1991). (H) The chemical shift index (+1, ≥0.1 ppm upfield shift; 0, <0.1 ppm shift; −1, ≥0.1 ppm downfield shift) for αH in the protein relative to isolated residues (Wishart *et al.*, 1991). The location of slowly exchanging peptide NHs is indicated by asterisks in the sequence above the shift index.

While two-dimensional NMR methods are now capable of defining in detail the solution molecular structure of such small proteins (Wüthrich, 1986), the ubiquitous presence of paramagnetic iron atoms in Rd and Fd (Münck & Kent, 1986) leads to serious line broadening that can result in the loss of the signals that define the folding topology. This problem is particularly acute for the strongly paramagnetic Rd, where a structural study of *Pf* Rd became feasible only after substituting the paramagnetic iron with a diamagnetic zinc ion (Blake *et al.*, 1991). This substitution afforded a detailed solution structure (Blake *et al.*, 1992a) that was shown to be essentially the same as that deduced by X-ray crystallography (Day *et al.*, 1992) for the native iron-containing protein (Blake *et al.*, 1992b). To date it has not proven possible to substitute completely diamagnetic metal ions in a cubane iron–sulfur cluster. However, the 3Fe and 4Fe Fds are less strongly paramagnetic than Rd due to strong antiferromagnetic coupling among the iron atoms (Münck & Kent, 1986) and hence lead to less severe line broadening for residues near the cluster. Thus a $^1$H NMR approach to the solution structure determination of a 3Fe or 4Fe Fd appears practical. Such assignments are needed to determine the structural basis for differentiating the magnetic coupling among the three iron atoms (Münck & Kent, 1986; Busse *et al.*, 1992). 2D $^1$H NMR assignments have been reported (Nettesheim *et al.*, 1992; Gaillard *et al.*, 1992) for a similar iron–sulfur cluster protein, high-potential iron–sulfur protein, HiPiP, and partial assignments have been reported for a flexible 4Fe oxidized Fd from *Desulfovibrio desulfuricans* (Norway) (Marion & Guerlesquin, 1989).

We report herein an NMR study of the folding topology of the oxidized 3Fe form of *Pf* Fd. Preliminary $^1$H NMR studies were carried out in $^2H_2O$ solution, which yielded the assignment of the three aromatic side chains and the hyperfine shifted signals of Cys17, as well as the location of the strongly hyperfine shifted signals for the other two coordinated Cys (Busse *et al.*, 1992). Two of the aromatic residues, Trp2 and Tyr46, were shown to have strongly interacting side chains. Our goal in the present study was to complete the sequence-specific assignment of the 3Fe Fd$^{ox}$ to the degree possible in this paramagnetic derivative, determine its secondary structure, and explore qualitatively the dynamic stability of these structural elements as revealed by labile proton exchange. The sequence of *Pf* Fd shown in Figure 1C, as well as those of Fds from other hyperthermophilic archaea (Zhou *et al.*, 1994), exhibits some homology to that of less thermostable bacterial single-cluster Fds, such as FdI from *Desulfovibrio gigas*, *Dg*, with the sequence shown in Figure 1B (41% identify with *Pf* Fd). Moreover, since *Dg* FdI is the only 3Fe Fd for which there exists a high-resolution crystal structure (Kissinger *et al.*, 1991), we use *Dg* Fd as our reference protein for comparison with *Pf* Fd. The four iron-ligating Cys residues are also identified by I–IV in Figure 1, panels B and C (*Pf* Fd has Asp14 at position II), and those participating in the disulfide bridge are designated V and VI and are used as the major basis for alignment in Figure 1, panels B and C.

## MATERIALS AND METHODS

*Sample Preparation. Pyrococcus furiosus* (*Pf*) cells were grown (Bryant & Adams, 1989), and their ferredoxin (Fd) was isolated, purified (Aono *et al.*, 1989), and converted to the 3Fe form (Conover *et al.*, 1990) as described previously. 2D NMR samples were 8 mM in protein dissolved in 90% $^1H_2O$/10% $^2H_2O$, 50 mM in phosphate, at pH 7.0 and 8.0. The study of the slowly exchanging peptide NHs was carried out on a 6 mM sample which was exchanged into $^2H_2O$, 50

mM in phosphate, in an Amicon ultrafiltration cell, and its $^1H$ NMR spectrum was recorded. This sample was kept at 30 °C at pH 7.0 for 310 h (~2 wk) and then stored for ~7 months at 5 °C at the same pH, after which its $^1H$ NMR spectrum was recorded again.

*NMR Spectroscopy.* 2D $^1H$ NMR experiments were performed on a GE Omega PSG-500 spectrometer operating at 500 MHz. Phase-sensitive $^1H$ DQF-COSY (Rance *et al.*, 1983), TOCSY (Bax & Davis, 1985; Griesinger *et al.*, 1988), and NOESY (Jeener *et al.*, 1979; Macura & Ernest, 1980) spectra were recorded at 30 and 40 °C, with sweep widths of 7017 Hz in both dimensions. Typically, a total of 2048 complex data points were collected with 512 $t_1$ increments and 96 transients for each $t_1$ increment. Quadrature detection in the $t_1$ dimension was obtained using hypercomplex mode (States *et al.*, 1982). The water signal was suppressed by low-power selective irradiation during the predelay period for all experiments, followed by a SCUBA sequence (Brown *et al.*, 1988) to allow for magnetization recovery of resonances close to the water frequency. For 2D TOCSY experiments, a spin-lock field of 11 kHz was used during the MLEV-17 mixing time of 60 ms, which includes two 1-ms trim pulses. NOESY spectra were recorded with 150-, 250-, and 350-ms mixing times.

NMR data were processed on either a Silicon Graphics Iris/4D workstation or SUN Sparc stations using the Biosym Felix 2.1 program. Thirty-degree-shifted sine-bell-squared functions were applied in both dimensions for NOESY and TOCSY data; DQF-COSY data were processed with 10°-shifted sine-bell-squared functions. The data sets were zero-filled to a final matrix size of 2048 × 2048 real data points prior to Fourier transformation. Chemical shifts were referenced to 2,2-dimethyl-2-silapentane-5-sulfonate, DSS, via the $^1H_2O$ resonance frequency at 4.71 ppm at 30 °C and 4.61 ppm at 40 °C.

## RESULTS

The assignment of spin systems was accomplished by a standard two-stage procedure (Wüthrich, 1986). Spin systems of different amino acids have been categorized and identified on the basis of their characteristic patterns of chemical shifts in DQF-COSY and TOCSY spectra. The coupling patterns of spin systems were examined at both the NH and $C_\alpha H$ regions with TOCSY spectra, and primary connectivities were established by DQF-COSY. Particular emphasis was placed on identifying partial spin systems where the loss of peaks can be traced to line broadening due to the paramagnetic cluster. Sequence-specific assignment of the above defined spin systems was achieved by building up peptide segments using the $\alpha_i$–$N_{i+1}$, $N_i$–$N_{i+1}$, and $\beta_i$–$N_{i+1}$ connectivities in NOESY spectra. For prolines, which do not have a backbone amide proton, the connectivities are established by the $\alpha_{i-1}$–$\delta,\delta_i'$ and $\alpha_i$–$N_{i+1}$ NOEs. Interstrand NOEs are carefully examined on the basis of the protein sequence and the identification of spin systems. Spectra collected at pH 7.0 and 8.0 at 30 °C and at pH 7.0 at 40 °C are used to circumvent several pairwise degeneracies in the NH region; the chemical shift was found to be the same at the two pH values.

The amide region of the 500-MHz NMR spectrum of oxidized 3Fe *Pf* Fd in 90% $^1H_2O$:10% $^2H_2O$, 50 mM in phosphate, pH 7.0, is illustrated in Figure 2A. The full intensity of the resonance at 9.23 ppm is recovered when a soft-pulse sequence (Redfield *et al.*, 1975) is used without exciting the solvent line (not shown). The resonances previously assigned to the nonlabile protons of the aromatic

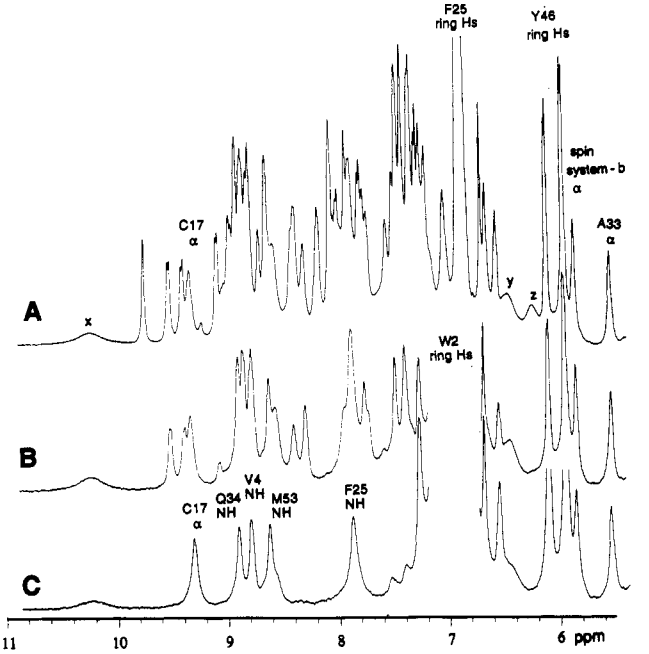


FIGURE 2: (A) Peptide proton region of a 500-MHz $^1H$ NMR spectrum of 6 mM oxidized *Pf* 3Fe Fd in 90% $^1H_2O$/10% $^2H_2O$, 50 mM in phosphate, at 30 °C, pH 7.0. The assigned nonlabile proton signals are labeled. The three broad partially resolved resonances labeled x, y, and z arise from strongly relaxed exchangeable protons. (B) The same region of the spectrum of a sample of *Pf* 3Fe Fd upon exchanging $^1H_2O$ for 100% $^2H_2O$ at 30 °C, pH 7.0. Approximately 20 labile slowly exchanging proton signals are observed, of which 18 participate in the 2D maps; also present with full intensity are the broad signals x and y. (C) The same sample as in panel B after ~2 weeks at 30 °C and ~7 months at 5 °C, pH 7.0. Note that four labile protons still appear with essentially full intensity, while spins x and y are retained with ~0.5 intensity.

side chains of Trp2, Phe25, and Tyr46 are labeled, as is the hyperfine shifted $C_\alpha H$ signal of the coordinated Cys17 (III) (Busse *et al.*, 1992). The other five previously reported hyperfine shifted Cys $C_\beta H$ signals in the low-field window (11–28 ppm) are not shown. Three broad, paramagnetically relaxed labile protons which do not participate in any 2D spectra are partially resolved and are labeled x, y and z, with $T_1$ values of 3, 8, and 16 ms, respectively. The only other nonlabile protons in this window are two weakly paramagnetically relaxed resonances at 5.85 and 5.58 ppm ($T_1$ ~ 50 ± 15 ms). The assignment of relatively narrow resonances to spin systems and to their sequence positions is detailed below.

*Spin Systems Identification. Pf* Fd comprises 66 residues with the primary sequence shown in Figure 1C (Zhou *et al.*, 1994). These residues can be classified by composition for the purpose of NMR as follows: 22 residues which normally have unique chemical shift patterns (2 Gly, 9 Ala, 2 Thr, 6 Val, and 3 Pro); 19 residues whose $C_\alpha H$ is part of an AMX spin system (2 Asn, 6 Asp, 5 Cys, 3 Ser, 1 Phe, 1 Trp, and 1 Tyr); and 25 residues with longer chains (2 Gln, 10 Glu, 4 Lys, 2 Met, 2 Leu, and 5 Ile), of which the latter two types include methyls in their spin topology. Inspection of the fingerprint region of the TOCSY (and DQF-COSY) spectrum, which is shown in Figure 3, reveals 54 peptide NH-containing residues, for some of which only a portion of the side chain is observed. Further inspection of the $C_\alpha H$ region of the TOCSY spectrum locates an additional eight partial or whole spin systems for which a NH could not be detected, including a resolved single proton at 5.58 ppm which failed to exhibit any scalar correlations.

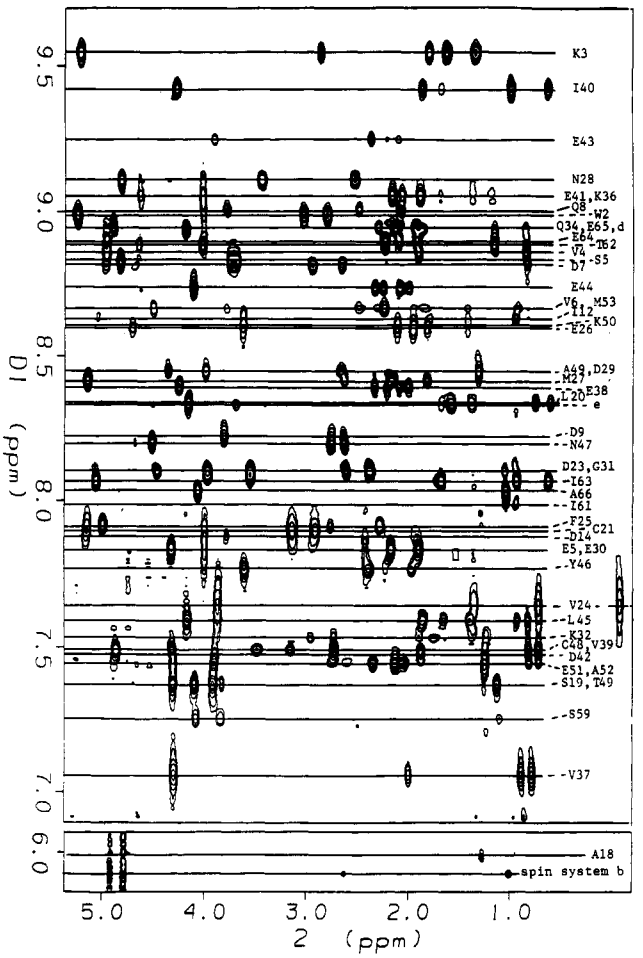


FIGURE 3: Fingerprint region of the 30 °C TOCSY spectrum (60-ms mixing time) showing the 54 spin systems detected through their NH peaks, and one (spin system b) detected only at its αH frequency; the several pairwise degeneracies are resolved at 40 °C.

The complete spin systems with unique chemical shift patterns readily located are both Gly (one without NH), six of nine Ala (two without NHs), both Thr, and four of six Val. In addition, two $AMX_3$ spin systems, labeled a and b, are located whose chemical shifts indicate a portion of a Val or Ile (the two Thr have been identified above by both spin topology and chemical shifts). The $C_\alpha H$ region clearly located two Pro spin systems. The fingerprint region yields 16 AMX spin system. The three remaining of the 19 AMX systems are associated with the previously located three Cys ligands coordinated to the cluster (Busse *et al.*, 1992). The amide region of the TOCSY and DQF-COSY spectra reveals four $NH_2$ groups, two of which identify the two Asn by the characteristic NOESY cross peak between $NH_2$ and $C_\beta Hs$. NOESY cross peaks between the $C_\beta Hs$ of two AMX systems and the previously assigned aromatic protons of Trp2 and Tyr46 (Busse *et al.*, 1991) lead to the complete assignment of two of the three aromatic residues. The Phe25 ring, however, exhibits cross peaks to the $C_\beta Hs$ of two AMX spin systems, dictating the need to locate these two AMX systems sequence specifically (see below). Three Ser are located by their unique chemical shift pattern. The remaining nine detected AMX systems must account collectively for Phe25, the two non-coordinating Cys, and six Asp.

Twenty-three of the 25 long-chain residues are readily located: both Gln by their characteristic NOESY cross peaks from $NH_2$ groups to the $C_\gamma Hs$, both Leu by their terminal isopropyl fragment, all four Lys, two complete and one partial of the five Ile (the two $AMX_3$ systems a and b defined above are also candidates for Ile; spin system b possesses a resolved, moderately relaxed ($T_1$ = 50 ± 15 ms) $C_\alpha H$ at 5.85 ppm). The remaining 12 spin systems collectively arise from the 10 Glu, the two Met, and the remaining two Ile. Three NH–$C_\alpha H$ fragments (labeled c, d, and e) are located in the fingerprint region but fail to exhibit additional scalar cross peaks to $C_\alpha H$ and, hence, give no clue as to the type of spin system; another NH–$C_\alpha H$ also possesses a methyl and is labeled spin system *f*. The resolved proton at 5.58 ppm which exhibits no scalar correlation is designated spin system g. A partial spin system including α, β, and γ protons (labeled spin system h) likely arises from the remaining Pro. When we include the three coordinated Cys ligands whose hyperfine shifted and strongly relaxed signals have been reported previously (Busse *et al.*, 1992), there is available to date some NMR data on 65 of the 66 residues in the protein. The question remains whether the partial spin systems can be located in the sequence so as to identify the nature of the residues.

*Sequence-Specific Assignments.* The assignments were initiated at the various unique residues identified above, *i.e.*, Trp2 and Tyr46, and at Gly and Asn. Some of the backbone NOESY cross peak patterns that allowed the assignments are depicted in Figures 4 and 5, and the relevant cross peaks are summarized in Figure 1, panels E and F. The segment Ala1–Thr10 was assigned on the basis of the standard NOESY cross peak pattern for the backbone. TOCSY failed to detect or resolve one methyl for Val4 (identified above as $AMX_3$ spin system a) and exhibited only very weak cross peaks to the methyls of Val6 (Figure 3). No signals could be uniquely identified by these 2D NMR methods for the residues in the segment Cys11(I)–Cys17(III), which includes all but one of the cluster-ligating residues (see below). The hyperfine shifted and strongly relaxed signals for Cys17(III) had been previously assigned by steady-state NOEs (Busse *et al.*, 1992). The sequence after Cys17(III), including Ala18–Glu34, is the most highly conserved, and the *Dg* Fd crystal structure was used to guide the assignments to Lys32. Gly 31 having been assigned, the other Gly without a detectable NH and a weak TOCSY peak can be assigned to the Gly11 near the cluster. It is observed that the ring protons of Phe25 exhibit much stronger NOESY cross peaks to $C_\beta Hs$ of Cys21(V) than to those of Phe25 (not shown). The moderately relaxed proton at 5.58 ppm ($T_1$ = 50 ± 15 ms), designated spin system g, which failed to exhibit scalar cross peaks, is uniquely identified as $C_\alpha H$ of Ala33. Continuing with Ala33 $C_\alpha H$, the assignments are extended to Glu54 by standard connectivities depicted in Figure 1, panels E and F, and the NH–$C_\alpha H$ spin system c is identified as a portion of Glu54. The C-terminal fragment Ile61–Ala 66 is readily assigned by strong backbone NOEs and identifies spin system f as a part of Ile61. The segment Ala55–Ala60, which includes the remaining cluster ligand, Cys56(IV), could not be identified by these 2D NMR (see below) methods. The chemical shifts for all residues are listed in Table 1.

There remain unassigned by the above methods four complete systems—Ser, Ile, Ala, and AMX—as well as partial spin systems b, d, e, and h, an Ile we designate Ile-x, and an Ala we label Ala-y. In the absence of characteristic NOESY cross peaks to these residues, only some tentative assignments can be offered. The Ser must arise from the only remaining Ser, Ser59, and the AMX spin system (assuming it is complete) likely originates from the only such system unassigned, Asp14. Since there remain two Ile (Ile12 and -16) and three Ala (Ala15, -55, and -60) unidentified, Ile-x and Ala-y are not assignable at this time. The chemical shifts for the unassigned residues are listed at the end of Table 1. *It is noteworthy that*

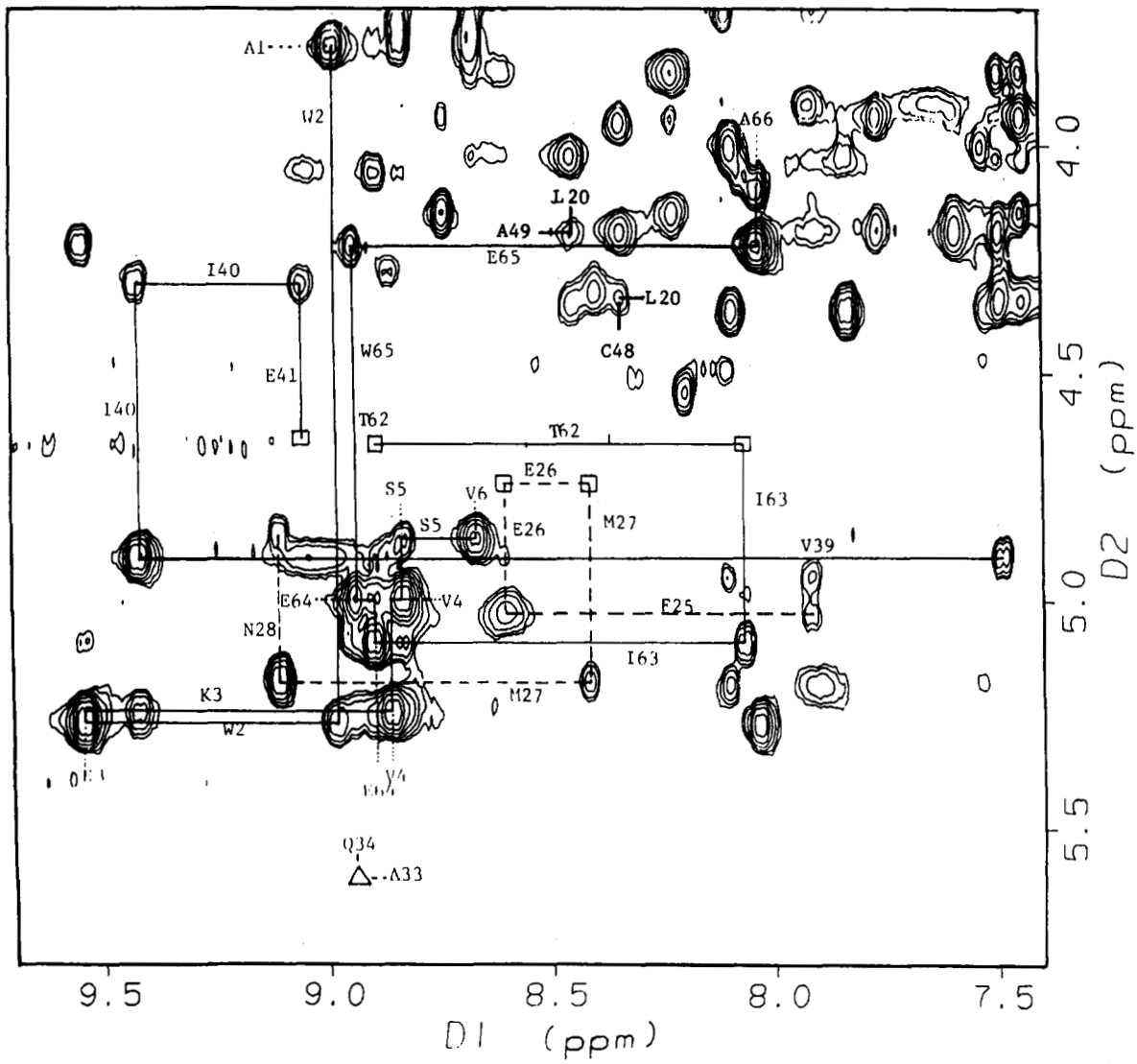


FIGURE 4: Fingerprint portion of the 30 °C NOESY spectrum (250-ms mixing time) showing the $\alpha_i$–$N_{i+1}$ cross peaks that map out the triple antiparallel stranded $\beta$-sheet A (solid lines), the double antiparallel stranded $\beta$-sheet B (dashed lines), and the backbone NOEs involving Leu20, Cys48, and Ala49. The boxes indicate where cross peaks are detected when the water resonances are not saturated during the mixing period. The horizontal and vertical frequencies with labeled lines correspond to the $\alpha$H and NH resonances, respectively. The open triangle indicates the position of the $\alpha_{33}$–$N_{34}$ cross peak observed only for a mixing time of 150 ms.

*all missing residues in the sequence-specific assignment are located in the loop between Cys8(I) and Cys17(III), and near Cys56(IV), where the cluster-ligating residues are included.*

*Secondary Structure.* Five extended peptide segments identified through strong $\alpha_i$–$N_{i+1}$ NOESY cross peaks participate in two antiparallel $\beta$-sheets (Figure 4), as identified through strong interstrand $\alpha_i$–$\alpha_j$, $\alpha_i$–$N_{j+1}$, $N_{i+1}$–$\alpha_j$, $N_{i+1}$–$N_{j-1}$, and $N_{i-1}$–$N_{j+1}$ NOEs, some of which are shown in Figure 5 (also see the supplementary material). $\beta$-Sheet A involves 13 residues on three strands, the N-terminal Ala1–Ser5, the C-terminal Thr62–Ala66, and the inner segment Val39–Glu41, with the N-terminal strand in the middle. The observed NOE connectivities for $\beta$-sheet A are depicted in Figure 6A. The section involving the two termini contains six interstrand hydrogen bonds, as indicated by dashed lines in Figure 6A. The $\alpha_6$–$\alpha_{61}$ NOE is not observed, but the formation of the Thr62 NH hydrogen bond to the carbonyl of Ser5 is supported by both the $\alpha_5$–$N_6$ and $N_5$–$N_{62}$ NOEs. The NOE connectivities indicate the presence of three hydrogen bonds between the N-terminus and the inner segment, as shown in Figure 6A. Similarly, the $\alpha_1$–$\alpha_{41}$ NOE is not observed, but the presence of the Trp2 NH hydrogen bond to the carbonyl of Ile40 can be inferred from the strong $\alpha_1$–$N_2$, $\alpha_{40}$–$N_{41}$, and $N_2$–$N_{40}$ and the weaker $N_2$–$\alpha_{41}$ NOEs. The strong $N_{38}$–$N_{39}$ NOE dictates that only one hydrogen bond is formed between Asp38 and Val4.

$\beta$-Strand B occurs in the inner hydrophobic core and is formed between the extended segments Glu26–Asn28 and Lys32–Gln34, with the interstrand dipolar connectivities and four resulting hydrogen bonds shown in Figure 6B. The expected Lys32–Ala33-NH NOE is not detected because the Ala NH cannot be observed due to paramagnetic effects (see Discussion). The two strands in $\beta$-sheet B are directly connected by a tight turn (turn C), as reflected by continuous $N_i$–$N_{i+1}$ NOEs involving Asp29–Lys32, and a $N_i$–$N_{i+2}$ NOE between Glu30 and Lys32 (Figure 6B). The second strand of $\beta$-sheet B is followed by a loop consisting of Pro35–Glu38; the absence of a tight turn is dictated by only a weak $N_{36}$–$N_{37}$ NOE and a missing $N_{37}$–$N_{38}$ NOE.

A twelve-member helical (helix B) segment is identified for Glu43–Glu54 by the characteristic $N_i$–$N_{i+1}$/$\alpha_i$–$N_{i+3}$/$\alpha_i$–$N_{i+1}$/$\alpha_i$–$\beta_{i+3}$ NOE pattern (Figures 1E and 5). Both $\alpha_i$–$N_{i+2}$ and $a_i$–$N_{i+4}$ NOESY cross peaks are observed for several of the residues, which favor an $\alpha$-helix over a $3_{10}$ helix. The failure to detect any signals from Ala55 precludes determination of whether the helix includes this residue. Other 3Fe/4Fe Fds

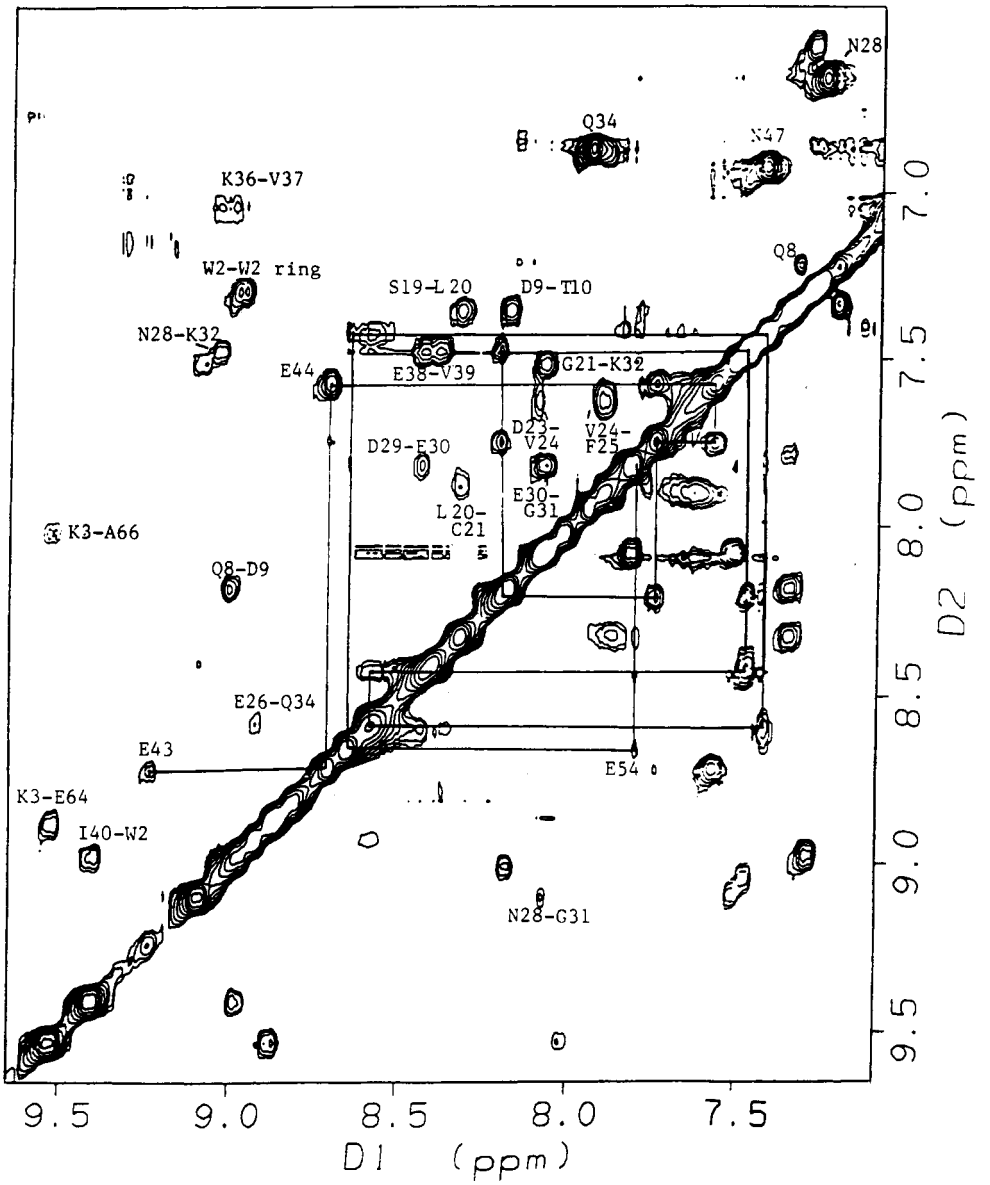


FIGURE 5: Low-field portion of the 30 °C NOESY spectrum (250-ms mixing time) showing NH–NH connectivities. The cross peaks corresponding to α-helix B are connected by solid lines for residues E43–E54. Intrastrand NH–NH cross peaks are labeled by the two contributing residues. Cross peaks labeled by a single-letter code correspond to the $NH_2$ of Asn or Gln.

possess another short helical segment, labeled α-helix A for *Dg* Fd in Figure 1A, that starts one residue before the cluster ligand Cys III and terminates at the residue prior to Cys V, one of the two non-coordinated Cys residues (Fukuyama *et al.*, 1988; Kissinger *et al.*, 1991); i.e., the residues Ile16–Leu20 for *Pf* Fd. However, the signals for the first three of these residues have not been located and/or are too strongly relaxed to give rise to cross peaks in the NOESY map. The cross peaks for Ser19 and Leu20 are consistent with, but not proof for, a helix (see Discussion). A total of three tight turns are directly observed in *Pf* Fd, of which turns B and C are described in Figure 6B. The pattern of NOEs (Figure 1A) for residues Asp7–Thr10 indicate a very rigid type I turn (turn A), with the backbone hydrogen bond from the Thr10 NH to the carbonyl from Asp7. With Asp7 being the initiator of turn A, its side-chain carboxylate likely participates in a hydrogen bond to the NH of Asp9 to form an Asx turn (Baker & Hubbard, 1984; Kissinger *et al.*, 1991). The residues Pro22–Phe25 form a type I tight turn (turn B) which terminates at the beginning of β-sheet B. The turn is well defined by strong $N_{23}$–$N_{24}$, $N_{24}$–$N_{25}$, and $N_{23}$–$N_{25}$ NOEs (Figure 6B), as well as $N_{21}$–$N_{23}$ NOEs (not shown).

The only remaining long-range backbone NOEs observed for *Pf* Fd are $N_{20}$–$\alpha_{48}$, $\alpha_{20}$–$N_{49}$ (Figure 4) involving residues near Cys21(V) and Cys48(VI), $\alpha_{24}$–$N_{36}$, and $\alpha_{24}$–$N_{37}$, between the loop after β-sheet B and turn B (see Discussion).

*Labile Proton Exchange.* The peptide regions of the 500-MHz $^1H$ NMR spectrum of *Pf* $Fd^{ox}$ immediately after exchange from $^1H_2O$ to $^2H_2O$ at pH 7.0 and 30 °C is shown in Figure 2B. This spectrum, together with DQF-COSY and NOESY spectra of this sample (not shown; see supplementary material), identify the slowly exchanging labile protons as arising from the peptide NHs of Lys3, Val4, Ser5, Leu20, Cys21, Phe25, Glu26, Asn28, Lys32, Gln34, Ile40, Ala49, Lys50, Glu51, Ala52, Met53, Thr62, and Glu64, as labeled by asterisks in Figure 1H. After ~2 weeks at 30 °C and ~7 months at 5 °C, pH 7.0, the $^1H$ NMR spectrum (Figure 2C) reveals four narrow labile proton signals with essentially full intensity (Val4, Phe25, Glu34, and Met53) and ~0.5 proton intensity for the broad peaks x and y.

## DISCUSSION

*Influence of Cluster Paramagnetism on Assignments.* The 1D and 2D $^1H$ NMR spectra provide information on at least

Table 1: Chemical Shifts of Amino Acid Residues in the 3Fe Form of Oxidized *Pyrococcus furiosus* Ferredoxin[a]

| residue | NH | αH | βH | others |
|---|---|---|---|---|
| Ala1 | | 3.78 | 1.26 | |
| Trp2 | 8.98 | 5.26 | 3.05, 2.82 | 6.57 (5H), 6.71 (2H), 6.86, 6.89 (6H, 7H), 7.30 (4H) |
| Lys3 | 9.54 | 5.22 | 1.82 | 1.65, 1.37, 2.89 |
| Val4[b] | 8.86 | 4.98 | 1.92 | 0.87 |
| Ser5 | 8.83 | 4.85 | 3.76, 3.72 | |
| Val6 | 8.67 | 4.52 | 2.25 | 1.11, 0.94 |
| Asp7 | 8.80 | 4.74 | 2.96, 2.68 | |
| Gln8 | 8.99 | 3.80 | 1.93, 2.08 | 2.29, 2.51, $^{\epsilon}NH_2$ 7.23; 7.35 |
| Asp9 | 8.18 | 4.54 | 2.77, 2.64 | |
| Thr10 | 7.37 | 4.13 | 3.86 | 1.14 |
| Cys11[c] | | 7.09, 3.94 | 17.14, 19.73 or 11.82, 23.05 | |
| Gly13[d] | | 3.67, 3.91 | | |
| Asp14(?)[e] | 7.87 | 3.81 | 1.95, 2.45 | |
| Cys17[c] | | 9.37 | 5.22, 14.27 | |
| Ala18 | 6.00 | 4.07 | 1.26 | |
| Ser19 | 7.37 | 4.34 | 3.95(2) | |
| Leu20 | 8.34 | 4.19 | 1.72, 1.62 | 1.42, 0.79, 0.64 |
| Cys21 | 7.90 | 5.18 | 3.17, 2.95 | |
| Pro22 | | 4.95 | 2.58, 2.08 | 2.05, 1.93, 3.71, 3.35 |
| Asp23 | 8.09 | 4.47 | 2.63, 2.40 | |
| Val24 | 7.63 | 3.91 | 1.43 | 0.76, –0.05 |
| Phe25 | 7.91 | 5.03 | 2.78, 2.30 | 6.91 (4H, 2, 6H, 3, 5H) |
| Glu26 | 8.59 | 4.76 | 1.97, 1.83 | 2.12 |
| Met27 | 8.40 | 5.15 | 1.83, 2.18 | 2.55 |
| Asn28 | 9.11 | 4.86 | 2.54, 3.48 | $^{\delta}NH_2$ 6.68; 7.28 |
| Asp29 | 8.43 | 4.37 | 2.68, 2.22 | |
| Glu30 | 7.83 | 4.35 | 1.91, 1.95 | 2.19 |
| Gly31 | 8.09 | 3.59, 4.01 | | |
| Lys32 | 7.52 | 4.74 | 1.77 | 1.40, 1.23, 2.99, 2.75 |
| Ala33[f] | | 5.58 | | |
| Gln34 | 8.91 | 4.91 | 1.95, 2.14 | 2.29, 2.33, $^{\epsilon}NH_2$ 6.88; 7.95 |
| Pro35 | | 4.89 | 2.35 | 3.85, 3.74, 1.89 |
| Lys36 | 9.05 | 4.05 | 1.71 | 1.38, 1.21, 2.92 |
| Val37 | 7.06 | 4.32 | 2.01 | 0.89, 0.80 |
| Glu38 | 8.37 | 4.27 | 2.11, 2.02 | 2.35, 2.21 |
| Val39 | 7.48 | 4.90 | 1.26 | 0.84, 0.74 |
| Ile40 | 9.39 | 4.31 | 1.88 | 1.69, 1.02, 0.65 |
| Glu41 | 9.02 | 4.64 | 2.08, 1.91 | 2.18 |
| Asp42 | 7.48 | 4.90 | 2.74, 2.30 | |
| Glu43 | 9.23 | 3.93 | 2.13, 2.22 | 2.40 |
| Glu44 | 8.71 | 4.13 | 2.11, 2.03 | 2.33, 2.27 |
| Leu45 | 7.59 | 4.19 | 1.86, 1.67 | 1.41, 0.95, 0.84 |
| Tyr46 | 7.76 | 3.65 | 2.42, 2.00 | 6.02 (3,5H), 6.17 (2,6H) |
| Asn47 | 8.22 | 3.83 | 2.79, 2.68 | $^{\epsilon}NH_2$ 6.94; 7.44 |
| Cys48 | 7.49 | 4.34 | 3.51, 3.20 | |
| Ala49 | 8.43 | 4.00 | 1.32 | |
| Lys50 | 8.60 | 3.64 | 1.44, 1.20 | 1.05, 1.34, 2.81, 2.68 |
| Glu51 | 7.44 | 3.93 | 2.04, 2.15 | 2.36 |
| Ala52 | 7.42 | 4.34 | 1.26 | |
| Met53 | 8.65 | 3.80 | 2.09, 1.86 | 2.50 |
| Glu54[g] | 7.85 | 4.04 | | |
| Cys56[c] | | 3.94 or 7.09 | 11.82, 23.65 or 17.14, 19.73 | |
| Pro57(?)[e] | | 4.98 | 2.25 | 1.26 |
| Ser59 | 7.06 | 4.72 | 4.11, 3.88 | |
| Ile61[h] | 7.98 | 4.92 | 2.00 | 0.95 |
| Thr62 | 8.88 | 4.70 | 4.06 | 1.18 |
| Ile63 | 8.05 | 5.06 | 1.69 | 0.95, 0.66 |
| Glu64 | 8.88 | 4.98 | 1.94 | 2.29, 2.14 |
| Glu65 | 8.94 | 4.19 | 1.91, 2.08 | 2.15 |
| Ala66 | 8.03 | 4.11 | 1.06 | |
| Ile-*x*[i] | 8.61 | 5.05 | 1.76 | 0.97, 0.63 |
| Ala-*y*[i] | | 4.10 | 1.45 | |
| spin system b[j] | | 5.85 | 2.63, 1.04 | |
| spin system d[j] | 8.89 | 4.50 | | |
| spin system e[j] | 8.31 | 3.71 | | |

[a] Shifts in ppm, ±0.01 ppm, from DSS, in 90% $^1H_2O$:10% $^2H_2O$ and 50 mM in phosphate, at 30 °C and pH 7.0–8.0. [b] $AMX_3$ spin system a with a methyl peak that may be lost due to cluster paramagnetism. [c] Taken from Busse *et al.* (1992); the coordinate Cys NHs were not observed. [d] Peptide NH missing due to cluster paramagnetism; see text. [e] Tentative assignment for spin system h; see text. [f] Designated spin system g in text; NH and $C_\beta H_3$ signals lost due to cluster paramagnetism; see text. [g] Designated spin system c in text; the βHs are lost due to cluster paramagnetism. [h] Designated spin system f in text; the remainder of the peaks are lost due to cluster paramagnetism (see text). [i] Residue type established, but not assignable sequence specifically due to cluster paramagnetism. [j] Partial spin systems for which signals are lost due to cluster paramagnetism; the amino acid origins could not be established.

parts of 65 of the 66 residues of the protein in spite of the paramagnetic ($S = 1/2$) ground state of the 3Fe cluster (including the three coordinated Cys signals addressed previously (Busse *et al.*, 1992)); *i.e.*, only one residue is completely undetected to date. However, two identified spin systems, one Ala-y and one Ile-x, cannot be assigned sequence

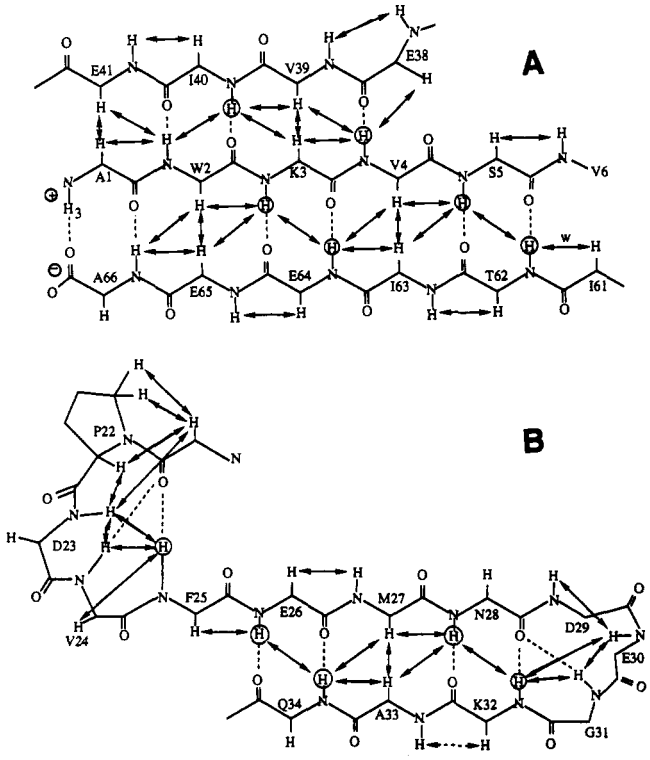


FIGURE 6: Schematic representation of the structures of (A) β-sheet A in *Pf* 3Fe Fd involving the two termini and the inner segment Val39–Glu41 and (B) β-sheet B and turns B and C in *Pf* 3Fe Fd involving residues Pro22–Gln34; the expected NOE $\alpha_{32}$-$N_{33}$ (shown by a dotted arrow) is not observed because paramagnetic relaxation effects do not allow the detection of the Ala33 peptide NH. The pattern of observed NOEs is indicated by double-headed arrows, and the presence of the inferred interstrand hydrogen bonds is shown by dashed lines. The slowly exchanging labile protons are marked by circles.

specifically due to lack of detected NOEs at this time and because there remain three Ala (15, 55, 60) and two Ile (12, 16) unaccounted for in the sequence. Of these 65 located spin systems, all but three could be traced to their unique (albeit only partial in six cases) spin systems. Two of the five residues (spin systems b, d, e, f, and h) which did not provide sufficient detail to provide any clue as to the spin system origin in TOCSY spectra could be identified by their unique positions in sequence-specific assignments, *i.e.*, NH–$C_\alpha$H–$CH_3$ fragment f of Ile61 and fragment g as $C_\alpha$H of Ala33. The chemical shifts for the unassigned residues are listed at the bottom of Table 1.

The residues in the sequence which remain unassigned are close to the paramagnetic cluster, in the short stretch between Cys11(I) and Cys17(III) and near Cys56(IV). The 2D assignment of oxidized and reduced HiPiP, a crystallographically characterized 85-residue 4Fe cluster protein (Nettesheim *et al.*, 1992; Gaillard *et al.*, 1992), yielded similar results, although a larger number of residues remained unassigned. It appears that the distance from the cluster iron at which paramagnetic relaxation, and hence assignments, becomes a problem in cubane Fe-S cluster proteins is ~5 Å. This is in contrast to the more strongly paramagnetic 2Fe cluster plant Fds which, for a similar $S = 1/2$ ground state in the reduced protein, lead to loss of signals within ~7.5 Å of the cluster (Oh & Markley, 1990; Ye *et al.*, 1992). The success in detecting scalar connectivity for the hyperfine shifted and relaxed spin systems of coordinated Cys in 8Fe Fd (Sadek *et al.*, 1993), however, indicates that there is significant promise for eventually mapping all of the spin system in the present molecule. The necessary short mixing time 2D experiments, collected under rapid repetition conditions to improve sensitivity, are under current study (La Mar & de Ropp, 1993).

Inspection of Figure 1G reveals that the missing or difficult to assign residues in *Pf* Fd are generally those residues that homology to *Dg* Fd predicts would have a backbone proton <5 Å from the iron and hence experience severe line broadening and rapid relaxation. The correlation is particularly good for Ala33 in *Pf* Fd (which exhibits only resolved $C_\alpha$H with no spin coupling), whose analog in *Dg* Fd, Ala31, has both NH and $C_\beta H_3$ close to the iron, as well as for Val4 and Ile61, both of whose partial spin systems are consistent with the proximity of their side chains to the cluster. Hence we conclude that the cluster paramagnetism qualitatively accounts for the missing and partial spin system detected by 2D NMR. This correlation of proximity to the iron and missing spin system, moreover, suggests that the geometry of the loop between Cys11(I) and Cys17(III) must be similar in the two proteins.

*Comparison of Secondary Structure with* ***Dg*** **Fd.** Two β-sheets, one double- and the other triple-stranded, one α-helix, and three tight turns are directly identified in *Pf* Fd. A second short helix, designated helix A, encompassing residues Ala13–Ile17 is present in *Dg* Fd (Kissinger *et al.*, 1991). The presence of helix A in *Pf* 3Fe Fd could not be directly established because of the expected and observed close proximity to the cluster of several of the residues (Ile16–Ala18) in the homologous *Pf* Fd sequence Ile16–Leu20. The observed medium-range $\alpha_{18}$–$N_{21}$ NOESY cross peak from the likely central member of helix A (Ala18) to the beginning of the highly conserved type I turn B (Cys21(V)), as well as the previously reported strong steady-state NOE from the hyperfine shifted and strongly relaxed Cys17(III) $C_\beta$H to the Phe25 ring (Busse *et al.*, 1992), is consistent with the expectation for helix A in *Dg* Fd (Kissinger *et al.*, 1991). Hence, we *tentatively* propose the presence of helix A in *Pf* 3Fe Fd.

The relative positions of the various secondary structural elements of *Dg* and *Pf* Fd are compared in Figure 1, panels A and D. The two proteins possess similar folding topology—two β-sheets, two helices (one confirmed, one only tentative; see above), and several tight turns, one of which is in the highly conserved hydrophobic core that contains the invariant Phe—but differ significantly in the detailed nature of some of the secondary structural elements. β-Sheet A in *Dg* Fd also involves the two termini, with Ile2–Val4 pairing with Arg57–Ile55 and forming three interstrand hydrogen bonds. However, in *Dg* Fd, it is a double- rather than a triple-stranded sheet with only one backbone hydrogen bond from the N-terminus Ile62 NH to the loop after β-sheet B (to the carbonyl of Pro36). The N-terminal strand in β-sheet A of *Dg* Fd is followed by a tight Asx turn for which the ligated Cys8(I) provides the hydrogen bond donor to the carboxyl of Asn5 (Kissinger *et al.*, 1991). In *Pf* Fd, a similar Asx turn is initiated by Asp7, whose carbonyl serves as acceptor to the NH of Thr10, one residue prior to Cys11(I). The loop region between Cys11(I) and Cys17(III) can be assumed to be similar to that in *Dg* Fd, since the coordinates for the *Dg* residues qualitatively account for the loss of signals via relaxation effects due to the cluster (see above and Figure 1E,F,G).

Both *Dg* and *Pf* Fd exhibit a characteristic tight type I turn involving the completely conserved Pro-Asp-Val-Phe sequence that connects to one strand of β-sheet B in each protein. The highly conserved structure in this hydrophobic core is confirmed by the detection of weak $\alpha_{18}$–$N_{19}$ and medium $\alpha_{18}$–$\alpha_{22}$ NOEs that are consistent with the 3.8- and 3.2-Å distances from the analogous residues in the crystal structure of *Dg* Fd (Kissinger *et al.*, 1991). The two antiparallel strands of β-sheet

B in each protein are connected by turn C. In *Dg* Fd, turn C is much more open due to additional residues, with the result that only three interstrand hydrogen bonds are formed in β-sheet B. In contrast, *Pf* Fd has a longer β-sheet B (four hydrogen bonds) and a much tighter type I turn (Asx, since the Asn 28 side chain likely makes a hydrogen bond to the NH of Glu 30), as shown in Figure 6B. The observed backbone NOEs from αH of Val24 to the NHs of Lys36 and Val37 are indicative of a loop structure very similar to that for the residue Val21 αH relative to Ile34 NH (3.7 Å) and Asn35 NH (3.2 Å) in the *Dg* Fd crystal structure (Kissinger *et al.*, 1991). It is thus likely that there is a hydrogen bond from the NH of Lys36 to the carbonyl of Val24. Hence the secondary structure between Cys17(III) and the loop after β-strand B is, with the exception of the tighter turn C in *Pf* Fd, highly conserved.

The most significant qualitative difference in secondary structure between the two proteins occurs in the portion between the loop after β-sheet B and the coordinated Cys IV (56 in *Pf*, 50 in *Dg*). In *Dg* Fd, the loop after β-sheet B is followed by a tight Asx turn (turn D) (Asp37–Leu40), followed by a nine-member α-helix B that terminates at the ligating Cys 50(IV). In contrast, the conserved loop after β-sheet B in *Pf* Fd is followed immediately by the extended portion that forms the third strand of β-sheet A, Val39–Glu41, which is then followed by a significantly longer α-helix which starts at Glu43 and runs until at least residue Glu54. Relaxation or dynamic influences preclude assignment of Ala55 and weaken some of the NOEs near the end of the helix. Since relaxation effects are negligible for the residues before Cys50(IV) at the end of helix B in *Dg* Fd, the missing signals for Glu54 and Ala55 suggest that there is less regular structure for the residues before Cys56(IV) in *Pf* Fd. Lastly, the analog of turn D in *Dg* Fd appears to be abolished in *Pf* Fd.

The interstrand NOEs from Cys48 α-H to NH of Leu20 and from Leu20 αH to Ala49 NH are those predicted for the residues Ile17 and Cys42 in *Dg* Fd and, hence, can be taken as very strong indirect evidence of the existence of a disulfide link between Cys21(V) and Cys48(VI) in *Pf* Fd, similar to that between Cys18 and Cys42 in *Dg* Fd (Bruschi, 1979; Kissinger *et al.*, 1991). These results are consistent with chemical modification experiments on the apoprotein of *Pf* Fd (Zhou *et al.*, 1994). It is noted, however, that although both *Dg* and *Pf* Fds possess the extended α-helix B just before ligation of Cys IV, *Dg* Fd has the disulfide bridge between the carboxyl terminus of α-helix A and the N-terminus α-helix B, while *Pf* has the disulfide bridge from the carboxyl terminus of the proposed helix A to the middle of the α-helix B. The conserved location of one side of the disulfide bridge is confirmed by the observation of stronger NOESY cross peaks from the Phe25 ring to $C_\beta$Hs of Cys21 than to the Phe25 $C_\beta$Hs and the NOEs between α of Pro22 and α of Ala18.

The pattern of chemical shifts for αHs has been shown (Wishart *et al.*, 1992) to reflect secondary structural elements in folded proteins. On the basis of the fact that the backbone configuration tends to impart significant upfield and downfield shifts to αHs in α-helices and β-sheets, respectively, residues are given indices of +1, 0, and −1 if the shift of a residue in the protein is changed ≥0.1 ppm downfield, <0.1 ppm, and ≥0.1 ppm upfield from that of the isolated residue. Any four or more sequential +1's indicate a β-strand, while any four sequential −1's reflect an α-helix. This index for the assigned residues in *Pf* 3Fe Fd is depicted graphically in Figure 1H and shows that the correlation with the secondary structure deduced from the NOE patterns is very good for residues relatively remote from the cluster, *i.e.*, α-helix B and β-sheet A. In the

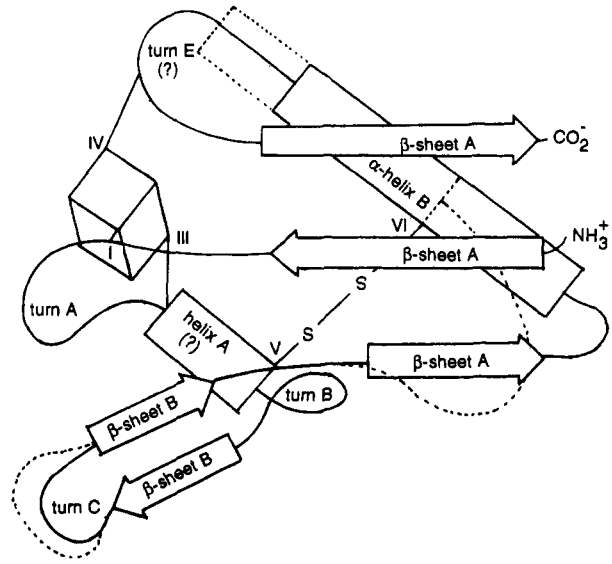


FIGURE 7: Schematic structure of *Pf* Fd (solid lines) depicting the positions of helices, β-sheets, and turns. The region where the *Dg* Fd structure differs from that of *Pf* Fd is shown qualitatively by dotted lines.

vicinity of where the paramagnetism of the cluster strongly influences relaxation, *i.e.*, proposed helix A, either scalar (*i.e.*, ligated Cys17(III); Busse *et al.*, 1992) or minor dipolar shifts, or both, likely contribute and interfere with a simplistic interpretation of the shifts.

From the above analysis we conclude that the folding near the cluster is likely similar in *Pf* and *Dg* Fd and that the structure of the hydrophobic inner core consisting of residues after Cys III through β-sheet B is largely conserved. Moreover, the N- and C-termini in both proteins strongly interact with each other in an antiparallel β-sheet. The position of β-sheet A over the loop between the conserved core and α-helix B, as well as that of α-helix B relative to the disulfide bridge, however, differs significantly in the *Pf* Fd relative to *Dg* Fd (Kissinger *et al.*, 1991). A schematic structure which qualitatively depicts both the similarities and the differences in *Pf* and *Dg* Fd is shown in Figure 7. The secondary structure of 3Fe *Dg* Fd (Kissinger *et al.*, 1991) differs only minimally from that of a 4Fe single-cluster Fd from *Bacillus thermoproteolyticus*, *Bt* (Fukuyama *et al.*, 1988, 1989) except for the absence of the disulfide link in the latter. The present study confirms the earlier prediction that Fds which possess only a single cubane cluster will have a loop and an α-helix in the region where the second cluster is generally found in bacterial Fds (Fukuyama *et al.*, 1988). Moreover, we provide here only the second example (*Dg* Fd is the first; Kissinger *et al.*, 1991) of a single-cluster Fd with a disulfide link between the two retained cysteines from the consensus sequence of the second cluster.

*Thermodynamic/Dynamic Stability.* *Pf* Fd achieves its remarkable thermostability with a sequence and folding topology quite similar to that of much less stable Fds but with extensions of several secondary structural motifs. Thus the N- and C-terminal strands of β-sheet A in *Pf* Fd extend to include the terminal residues Ala1 and Ala66, as was observed in *Pf* Rd (Blake *et al.*, 1991, 1992a,b; Day *et al.*, 1992). Moreover, a third antiparallel strand consisting of three residues, Val39–Glu41, is added from the loop which is longer in *Pf* Fd because of the insertion of three residues between β-sheet B and Cys IV relative to that in *Dg* Fd. This triple-stranded β-sheet A in *Pf* Fd contains nine interstrand hydrogen bonds, compared to only five in the double-stranded sheet A

in *Dg* Fd, which likely contributes to the enhanced stability of the former protein. β-Sheet B also contains one more hydrogen bond in *Pf* than in *Dg* Fd due to the tighter turn C. Lastly, α-helix B is three residues longer in *Pf* than in *Dg* Fd. *Dg* Fd contains five type I β-turns (in Figure 1A), of which three are of the type Asx and each contribute two hydrogen bonds. Three turns are confirmed for *Pf* Fd (Figure 1D). Turn C is significantly tighter in *Pf* than in *Dg* Fd, where the Asn28 carbonyl likely serves as a hydrogen bond acceptor to both Lys32 and Gly31 NHs. On the other hand, turn D of *Dg* Fd is abolished in *Pf* Fd. The last turn E, which follows Cys50(IV) in *Dg* Fd, could not be confirmed in *Pf* Fd because of the paramagnetic relaxation effects on residues Pro57–Ala60, but is likely to exist because of the strong sequence homology for the five residues following Cys IV in the two Fds.

In all, approximately eight additional backbone hydrogen bonds are observed in *Pf* Fd (five in β-sheet A, one in β-sheet B/turn C, and two in α-helix B). Likely important contributions arising from salt bridges and solvent shielding of hydrophobic residues cannot be ascertained until the tertiary structure is better defined. This requires developing methods to detect the relaxed residues near the cluster, a detailed study of side-chain NOE patterns, and computation methods to produce a model. Such studies are in progress.

The slowly exchanging labile protons observed in 2D experiments all reside in the identified secondary structural elements, and they are identified by asterisks in Figure 1H and by circles in Figure 6. The retardation of the exchange rate of the peptide NHs relative to that of an isolated residue can be taken as a measure of the dynamic stability of a protein (Gallagher *et al.*, 1992). Of these 18 slowly exchanging protons with relatively narrow signals, the triple β-sheet A contributes six (out of nine hydrogen bonds) and the double β-strand B contributes four (out of four hydrogen bonds), with one from the turn in the hydrophobic core (Phe25), two from α-helix A (Leu20, Cys21), and five arising from the α-helix B (residues 49–54). The distribution of the four slowest exchanging protons is similar, with one each in the central strand of β-sheet A (Val4), the inner β-strand B (Glu34), the tight turn in the hydrophobic core (Phe25), and α-helix B (Met53).

It is noted, however, that several of the slowest exchanging protons (*i.e.*, signals x and y in Figure 2A–C) possess relaxation properties ($T_1$s of 3–8 ms) that must arise from those residues near the iron–sulfur cluster whose peptide NHs serve as hydrogen bond donors to either the coordinated Cys or the bridging sulfide (*i.e.*, <5.0 Å from the irons (Busse *et al.*, 1992)). The *Dg* Fd crystal structure reveals eight residues that participate in such hydrogen bonding (Kissinger *et al.*, 1991), and the similar cluster structure suggests a similar number in *Pf* Fd. Hence the enhanced stability in the identified secondary structural elements also leads to a much more stable cluster environment. A similar retardation of labile proton exchange for residues hydrogen bonding to the ligating atoms was more directly observed in the $^1$H NMR study of Zn-substituted *Pf* Rb, where the diamagnetism allowed the sequence-specific assignment of such signals (Blake *et al.*, 1992b). The extraordinarily slow exchange rates for those peptide NHs that hydrogen bond to the cluster in *Pf* Fd (exchange lifetimes > 300 h at 30 °C, pH 7.0) are in strong contrast to those of the considerably less thermostable 2 × 4Fe Fd from *Clostridium pasteurianum*, where cluster labile proton exchange is complete within a day at 20 °C and pH 7.4 (Orme-Johnson *et al.*, 1983).

The four-residue insertion (Pro35–Glu38) in the region between β-sheet B and Cys56(IV) in *Pf* relative to *Dg* Fd, which likely leads to the repositioning of α-helix B and the formation of the third antiparallel strand β-sheet A, may have a structural role that has less to do with the hyperthermostability of *Pf* Fd than with another of its unusual properties, namely, the apparent participation of the disulfide link in the redox cycle for the 4Fe form of *Pf* Fd (C. Gorst, G. N. La Mar, and M. W. W. Adams, manuscript in preparation). The basis for this speculation is that the Fd from another hyperthermophilic archaeon, *Thermococcus litoralis* (*Tl*), with stability similar to that of *Pf* Fd (Zhou *et al.*, 1994), has a sequence in the region between the likely β-sheet A and Cys IV that is much more similar to that of *Dg* Fd than that of *Pf* Fd, and its disulfide link does not appear to be redox active (Donaire *et al.*, 1994). The relationship between the presently described structure for the oxidized 3Fe *Pf* Fd and the two as yet structurally uncharacterized 4Fe *Pf* Fd forms or that for *Tl* Fd is not yet known. Investigations on the structures of the 4Fe Fds of *Pf* and *Tl* are in progress.

## ACKNOWLEDGMENT

The authors are indebted to Drs. Jun Qin and Liping P. Yu for useful discussions and to Drs. Carol M. Gorst and Tina Yeh for experimental assistance.

## SUPPLEMENTARY MATERIAL AVAILABLE

Two figures (peptide NH and $C_\alpha$H regions of the NOESY spectrum in $^2H_2O$) (2 pages). Ordering information is given on any current masthead page.

## REFERENCES

Adams, M. W. W., & Kelly, R. M., Eds. (1992) *Biocatalysis at Extreme Temperatures: Enzyme Systems Near and Above 100 °C*, ACS Symposium Series 498, American Chemical Society, Washington, DC.

Adams, M. W. W. (1993) *Annu. Rev. Microbiol. 47*, 627–658.

Adams, M. W. W. (1994) *FEMS Microbiol. Rev.* (in press).

Adman, E. T., Sieber, L. C., & Jensen, W. H. (1973) *J. Biol. Chem. 248*, 3987–3996.

Adman, E. T., Sieber, L. C., Jensen, L. H., Bruschi, M., & LeGall, J. (1977) *J. Mol. Biol. 112*, 113–170.

Aono, S., Bryant, F. O., & Adams, M. W. W. (1989) *J. Bacteriol. 171*, 3433–3439l.

Baker, E. N., & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97–179.

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

Blake, P. R., Park, J. B., Bryant, F. O., Aono, S., Magnuson, J. K., Eccleston, E., Howard, J. B., Summers, M. F., & Adams, M. W. W. (1991) *Biochemistry 30*, 10885–10891.

Blake, P. R., Park, J.-B., Zhou, Z. H., Hare, D. R., Adams, M. W. W., & Summers, M. F. (1992a) *Protein Sci. 1*, 1494–1507.

Blake, P. R., Day, M. W., Hsu, B. T., Joshua-Tor, L., Park, J.-B., Zhou, Z. H., Hare, D. R., Adams, M. W. W., Rees, D. C., & Summers, M. F. (1992b) *Protein Sci. 1*, 1522–1525.

Brown, S. C., Weber, P. L., & Mueller, L. (1988) *J. Magn. Reson. 77*, 166–169.

Bruschi, M. (1979) *Biochem. Biophys. Res. Commun. 91*, 623–628.

Bryant, F. O., & Adams, M. W. W. (1989) *J. Biol. Chem. 264*, 5070–5079.

Busse, S. C., La Mar, G. N., Yu, L. P., Howard, J. B., Smith, E. T., Zhou, Z. H., & Adams, M. W. W. (1992) *Biochemistry 31*, 11952–11962.

Carter, C. W., Jr., Kraut, J., Freer, S. T., Xuong, N. N., Alder, R. A., & Bartsch, R. G. (1974) *J. Biol. Chem. 249*, 4212–4225.

Conover, R. C., Kowal, A. T., Fu, W., Park, J.-B., Aono, S., Adams, M. W. W., & Johnson, M. K. (1990) *J. Biol. Chem. 265*, 8533–8541.

Day, M. W.; Hsu, B. T., Joshua-Tor, L., Park, J. -B., Zhou, Z. H., Adams, M. W. W., & Rees, D. C. (1992) *Protein Sci. 1*, 1494–1507.

Donaire, A., Gorst, C. G., Zhou, Z. H., Adams, M. W. W., & La Mar, G. N. (1994) *J. Am. Chem. Soc.* (submitted for publication).

Fiala, G., & Stetter, K. O. (1986) *Arch. Microbiol. 145*, 56–61.

Fukuyama, K., Nagahara, Y., Tsukihara, T., Katsube, Y., Hase, T., & Matsubara, H. (1988) *J. Mol. Biol. 199*, 183–193.

Fukuyama, K., Matsubara, H., Tsukihara, T., & Katsube, Y. (1989) *J. Mol. Biol. 210*, 383–398.

Gaillard, J., Albrand, J.-P., Moulis, J.-M., & Wemmer, D. E. (1992) *Biochemistry 31*, 5632–5639.

Gallagher, W., Tao, F., & Woodward, C. (1992) *Biochemistry 31*, 4673–4680.

Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870–7872.

Jeener, B. H., Meier, P., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.

Kissinger, C. R., Sieker, L. C., Adman, E. T., & Jensen, L. H. (1991) *J. Mol. Biol. 219*, 693–715.

La Mar, G. N., & de Ropp, J. S. (1993) in *Biological Magnetic Resonance* (Berliner, L. J., & Reuben, J. S., Eds.), Vol. 12, pp 1–78, Plenum Press, New York, NY.

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

Marion, D., & Guerlesquin, F. (1989) *Biochem. Biophys. Res. Commun. 159*, 592–598.

Münck, E., & Kent, T. A. (1986) *Hyperfine Interact. 27*, 161–172.

Nettesheim, D. G., Harder, S. R., Feinberg, B. A., & Otvos, J. D. (1992) *Biochemistry 31*, 1234–1243.

Oh, B.-H., & Markley, J. L. (1990) *Biochemistry 29*, 3993–4004.

Orme-Johnson, N. R., Mims, W. B., Orme-Johnson, W. H., Bartsch, R. G., Cusanovich, M. A., & Peisach, J. (1983) *Biochim. Biophys. Acta 748*, 68–72.

Park, J.-B., Fan, C., Hoffman, B. M., & Adams, M. W. W. (1991) *J. Biol. Chem. 266*, 19351–19356.

Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 458–479.

Redfield, A. G., Kunz, S. D., & Ralph, E. K. (1975) J. Magn. Reson. **19**, 114–117.

Sadek, M., Brownlee, R.-T. C., Scrofani, D. B., & Wedd, A. G. (1993) *J. Magn. Reson., Ser. B 101*, 309–314.

States, D. J., Haberkorn, R. A., & Reuben, D. J. (1982) *J. Magn. Reson. 48*, 741–751.

Stetter, K. O. (1982) *Nature 300*, 258–260.

Stetter, K. O., Fiala, G., Huber, G., Huber, R., & Segerer, G. (1990) *FEMS Microbiol. Rev. 75*, 117–124.

Watenpaugh, K. D., Sieker, L. C., & Jensen, L. H. (1979) *J. Mol. Biol. 131*, 509–523.

Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991) *J. Mol. Biol. 222*, 311–333.

Woese, C. R., Kandler, O., & Wheelis, M. L. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4576–4579.

Wüthrich, K. (1980) *NMR of Proteins and Nucleic Acids*, pp 117–174, John Wiley & Sons, New York, NY.

Ye, X. M., Pochapsky, T. C., & Pochapsky, S. S. (1992) *Biochemistry 31*, 1961–1968.

Zhou, Z. H., Blamey, J. M., Park, J. B., Magnuson, J., Howard, J. B., Hankins, H., Johnson, M. K., & Adams, M. W. W. (1994) *J. Biol. Chem.* (submitted for publication).