Journal of Biomolecular NMR, 7 (1996) 35–47
ESCOM

J-Bio NMR 319

^{1}H NMR investigation of the secondary structure, tertiary contacts and cluster environment of the four-iron ferredoxin from the hyperthermophilic archaeon *Thermococcus litoralis*

Antonio Donaire[a,*], Zhi-Hao Zhou[b], Michael W.W. Adams[b] and Gerd N. La Mar[a,**]

[a]*Department of Chemistry, University of California, Davis, CA 95616, U.S.A.*
[b]*Department of Biochemistry and Molecular Biology and Center for Metalloenzymes Studies, University of Georgia, Athens, GA 30602, U.S.A.*

Received 31 July 1995
Accepted 7 September 1995

Keywords: Hyperthermophilicity; Ferredoxin; *Thermococcus litoralis*; Contact shifts; Fe-S cluster; Paramagnetic relaxation

Summary

The solution molecular structure of the four-iron ferredoxin (Fd) from the hyperthermophilic archaeon *Thermococcus litoralis* (*Tl*) has been investigated by ^{1}H NMR spectroscopy. TOCSY and NOESY experiments in H_2O, tailored to detect both weakly and strongly relaxed resonances, together with steady-state NOEs in both H_2O and D_2O, allowed the identification of 58 of the 59 residues, with one residue near the paramagnetic center undetected. It is shown that the contact shifted and strongly relaxed signals for all four cysteines ligated to the paramagnetic cluster can be assigned by standard backbone connectivities that do not require any assumptions about the tertiary structure. Secondary structural elements identified in *Tl* Fd are a three-stranded antiparallel β-strand involving the termini of the protein, a double β-strand (also antiparallel), two α-helices and four turns. The existence of a disulfide bridge between the nonligated cysteines is also proposed. Dipolar contacts observed in the NOESY maps and by steady-state NOEs between the ligated cysteines and the 'diamagnetic' protein matrix indicate that the overall folding pattern of *Tl* Fd is very similar to that of the 3Fe ferredoxin from the mesophilic bacterium *Desulfovibrio gigas* [Kissinger et al. (1991) *J. Mol. Biol.*, **219**, 693–723]. The influence of the paramagnetism of the cluster on the relaxation properties of the proton signals of nonligated residues near the cluster, as well as on the ligated cysteines, correlates well with the proximity to the cluster iron(s), as predicted from the crystal structures for homologous protons of other single-cluster ferredoxins. Finally, the potential role of the various identified structural factors in contributing to the hyperthermostability of this protein is discussed.

Introduction

The bacterial-type ferredoxins (Fds) are small (~60 residues) electron-transfer proteins that possess a cubane iron–sulfur cluster $[Fe_4S_4]^{+1,+2}$ or $[Fe_3S_4]^{0,+1}$; see Fig. 1 (Cammack et al., 1977; Beinert, 1990; Howard and Rees, 1991; Cammack, 1992). Ferredoxins possessing two $[Fe_4S_4]^{+1,+2}$ clusters are the most common, but a variety of proteins possessing a single 4Fe or 3Fe cluster, or one of each, have been characterized. The Fds, like the similar cubane cluster high-potential iron–sulfur proteins (HiPiPs), exhibit the properties of delocalized valence for the transferred electron, which is shared by a specific pair of the three- or four-iron in the cluster (Cammack et al., 1977; Cammack, 1992). Establishing the structural basis for the location of the reducing electron in cubane Fds is central to understanding the mechanism of control of redox potential in these proteins. Single-crystal X-ray structures

*Present address: Department of Chemistry, University of Florence, Via G. Capponi 7, I-50121 Florence, Italy.
**To whom correspondence should be addressed.

Abbreviations: Fd, ferredoxin; HiPiP, high-potential iron–sulfur proteins; *Dg*, *Desulfovibrio gigas*; *Av*, *Azotobacter vinelandii*; *Pf*, *Pyrococcus furiosus*; *Tl*, *Thermococcus litoralis*; *Pa*, *Peptostreptococcus asaccharolyticus*; *Bt*, *Bacillus thermoproteolyticus*; *Cp*, *Clostridium pasteurianum*; *Ca*, *Clostridium acidi urici*; *Da*, *Desulfovibrio africanus*; *Tm*, *Thermatoga maritima*; NOE, nuclear Overhauser effect; NOESY, 2D NOE spectroscopy; MCOSY, 2D magnitude correlation spectroscopy; TOCSY, total correlation spectroscopy.

have been reported for cubane 2×4Fe Fds from *Peptostreptococcus asaccharolyticus* (*Pa*; formerly *Peptococcus aerogenes*) (Adman et al., 1976; Backes et al., 1991), and *Clostridium acidiurici* (*Ca*) (Duie et al., 1994), 4Fe+3Fe Fd from *Azotobacter vinelandii* (*Av*) (Stout, 1989), 4Fe Fds from *Bacillus thermoproteolyticus* (*Bt*) (Fukayama et al., 1988) and *Desulfovibrio africanus* (*Da*) (Sery et al., 1994), and 3Fe Fd from *Desulfovibrio gigas* (*Dg*) (Kissinger et al., 1991). The cysteine-ligating consensus sequences **I**–**VI** and **I'**–**VI'** in the two clusters of the common bacterial Fds, Cys^{I}-X_2-Cys^{II}-X_2-Cys^{III}-X_n-Cys^{IV}-X_m-$Cys^{I'}$-X_2-$Cys^{II'}$-X_2-$Cys^{III'}$-X_n-Cys^{IV}, exhibit a pseudo-twofold symmetry, and this is reflected in the crystal structure of *Pa* and *Ca* Fds (Adman et al., 1976; Backes et al., 1991; Duie et al., 1994). The single-cluster Fds result from deletion of two or more of the cysteines in the **I'**–**IV'** sequence from the second cluster, but largely retain the overall folding topology, with the environment of the second cluster substituted by a helix (Fukayama et al., 1989) and, in certain cases, a disulfide bridge formed between two of the cysteines retained from the second cluster (Kissinger et al., 1991). The relative geometries of the ligating cysteines relative to the cluster in the consensus sequence are essentially conserved in the structurally characterized Fd.

The single-cluster Fds from hyperthermophilic archaea such as *Pyrococcus furiosus* (*Pf*) (Conover et al., 1990) and *Thermococcus litoralis* (*Tl*) (Mukund and Adams, 1993) possess extreme thermostability (negligible denaturation over 24 h at 95 °C), in spite of the fact that they exhibit redox properties and sequence homology remarkably similar to those of significantly less thermostable Fds (Busse et al., 1992). Both *Pf* and *Tl* Fd allow facile and reversible interconversion between the native 4Fe Fd and a 3Fe form in which the second cluster ligand (**II**) in the consensus sequence is not ligated (Fig. 1). The detailed molecular structures of these Fds should not only shed light on several of their interesting spectroscopic and functional properties, but also provide insight into the basis of their remarkable thermostability.

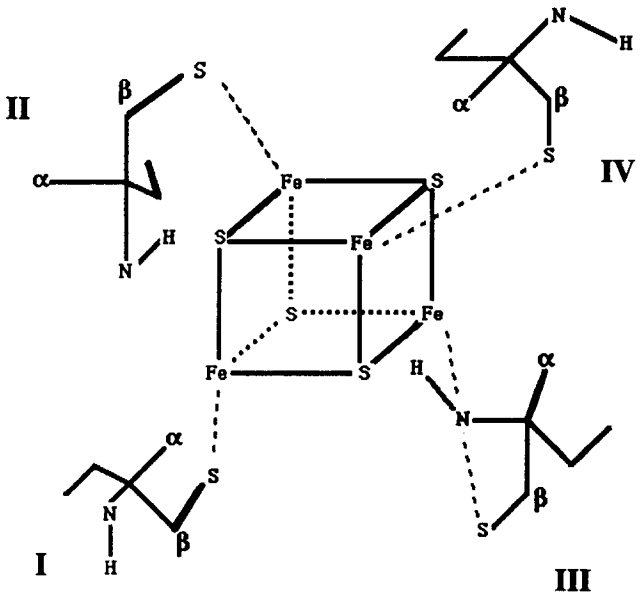

Fig. 1. Schematic representation of the cubane cluster in bacterial-type ferredoxins, showing the relative orientation of the backbone protons of the four cysteine ligands with respect to the cluster; the amide protons of Cys **II** and **III** and the C_αHs of Cys **I** and **IV** point towards the cluster, while the C_αHs of Cys **II** and **III** and the NHs of Cys **I** and **IV** are directed away from the cluster.

In the absence of any suitable crystals, we have pursued ^{1}H NMR characterization of the molecular and electronic structures of *Pf* and *Tl* Fds in both the 3Fe and 4Fe forms (Busse et al., 1992; Donaire et al., 1994; Teng et al., 1994; Gorst et al., 1995a). The molecules are sufficiently small to allow high-resolution NMR structural studies. However, the paramagnetic ground states ($[Fe_4S_4]^{+1}$, $[Fe_3S_4]^{0,+1}$) and the extensive population of paramagnetic excited states ($[Fe_4S_4]^{+2}$) at ambient temperatures (Poe et al., 1970; Cammack, 1992; Luchinat and Ciurli, 1993) induce effective relaxation, which severely complicates, but does not necessarily obviate, the use of appropriately tailored 1D and 2D NMR methods. Initial studies focused on the oxidized 3Fe form (S = 1/2) of *Pf* Fd, for which conventional 2D NMR spectroscopy defined the majority of the secondary structure (Teng et al., 1994). Subsequent NMR studies, optimized for detecting the paramagnetically influenced cluster environment including the ligated cysteines, led to the proposal that, in contrast to previous approaches (Bertini et al., 1994; Donaire et al., 1994), the needed unambiguous assignments of the ligated cysteines can be obtained without any assumption regarding the tertiary structure of the protein (Gorst et al., 1995a). While the overall folding topology, relative to the reference 3Fe *Dg* Fd, was maintained in 3Fe *Pf* Fd, the protein exhibited extension of several secondary structural motifs, which likely contribute to the enhanced thermostability. Thus, the surface β-sheet incorporated a third strand and extended to include the two terminal residues, the turn between the two strands of the internal β-sheet was tightened, and the major α-helix was slightly extended in length but shifted significantly away from the carboxy terminus relative to *Dg* and *Bt* Fds.

In order to establish which of these expanded and/or modified secondary structure elements detected in *Pf* Fd are essential contributors to the hyperthermostability of these Fds, we extend here our solution ^{1}H NMR structural studies to *Tl* Fd. The *Pf* and *Tl* Fds, like *Dg* Fd, each possess cysteines **V** and **VI**, which can (and in *Dg* and *Pf* Fd do) (Kissinger et al., 1991; Teng et al., 1994) participate in a disulfide bridge. This disulfide bridge has been shown to be redox active in *Dg* and *Pf* Fd (Macedo et al., 1994; Gorst et al., 1995b), with each of the four redox states of *Pf* Fd identified. In contrast, under the same solution conditions, 4Fe *Tl* Fd does not appear to

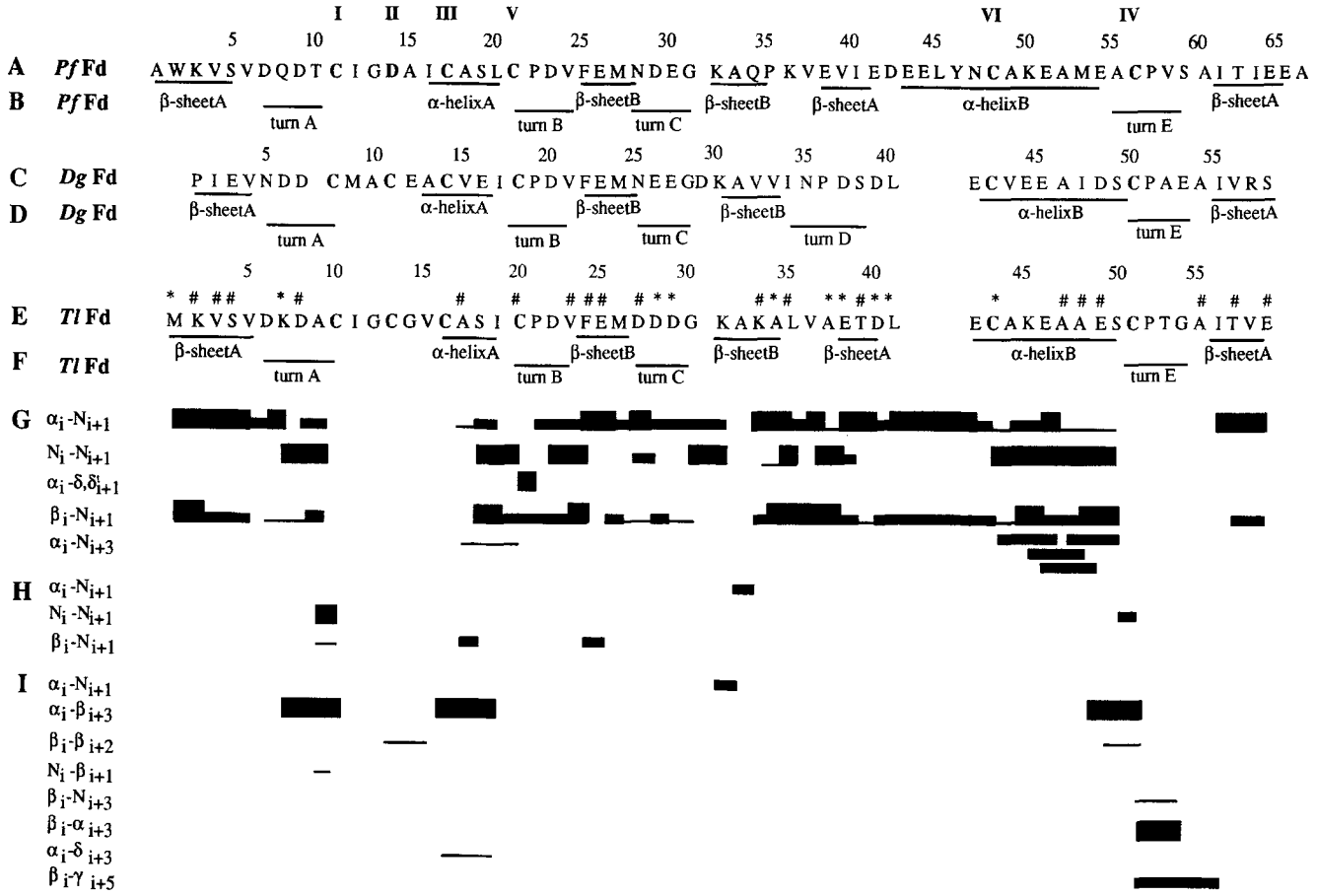

Fig. 2. Sequence and secondary structure elements of: (A,B) *Pf* 3Fe Fd (Teng et al., 1994); (C,D) *Dg* 3Fe Fd (Kissinger et al., 1991); and (E,F) *Tl* 4Fe Fd (this report). The basis for alignments are Cys **I**–**VI**. (G) Sequential and medium-range connectivities observed in the long-mixing-time (300 ms) NOESY. (H) Sequential and medium-range connectivities only observed in the short-mixing-time (50 ms) NOESY experiment. (I) Sequential and medium-range connectivities obtained from the steady-state NOE experiments. Peptide NHs that exhibit saturation transfer (very rapid exchange) or that are present in D_2O after 10 h (very slow exchange) are marked by * and # above the residue symbol, respectively.

involve cysteines **V** and **VI** in its redox cycle (Donaire et al., 1994). The sequence of *Tl* Fd is compared with those of *Pf* Fd and our reference protein, *Dg* Fd, in Fig. 2. *Tl* Fd, unlike *Pf* Fd, retains its full complement of four ligating cysteines (also labeled **I**–**IV** in Fig. 1, where *Pf* has Asp^{14} at position **II**). The sequence homology between *Tl* and *Dg* Fds is similar to that of *Pf* and *Dg* Fds, which have approximately 42% identity, while *Tl* and *Pf* Fds share 61% identity. Previous ^1H NMR analysis of the relaxation properties of hyperfine-shifted cysteine resonances in both oxidized 3Fe and 4Fe *Tl* Fds, however, has demonstrated that the intrinsic cluster paramagnetic relaxation influences are weaker in the 4Fe than in the 3Fe oxidized Fd (Busse et al., 1992; Donaire et al., 1994). Hence the prospects for describing the molecular structure by ^1H NMR are improved for the 4Fe relative to the 3Fe form of the protein, and the present study concerns only the 4Fe *Tl* Fd.

Our goals for this report are to assign as completely as possible the ^1H NMR spectrum of *Tl* 4Fe Fd, including the paramagnetically influenced signals for the ligated cysteines, to describe the secondary structural elements and cluster geometry, and to identify the tertiary contacts among the various secondary structural elements and/or the cluster that may give clues to the origin of the thermostability of the protein. The assignment strategy will include conventional 2D NMR techniques as applied to diamagnetic proteins (Wüthrich, 1986), as well as 2D and 1D methods tailored to detect paramagnetically relaxed resonances (La Mar and de Ropp, 1993; Gorst et al., 1995a). The four-cysteine ligation for native *Tl* Fd will also constitute a test of our proposed scheme (Gorst et al., 1995a) for sequence-specific identification of the sequence origin of hyperfine-shifted ligated cysteine resonances without assuming tertiary homology to structurally characterized Fds.

Materials and Methods

Protein preparation

Culturing of *T. litoralis* and the purification of its ferredoxin were carried out as previously described (Mukund and Adams, 1993). The samples for NMR spectroscopy were equilibrated with 50 mM sodium phos-

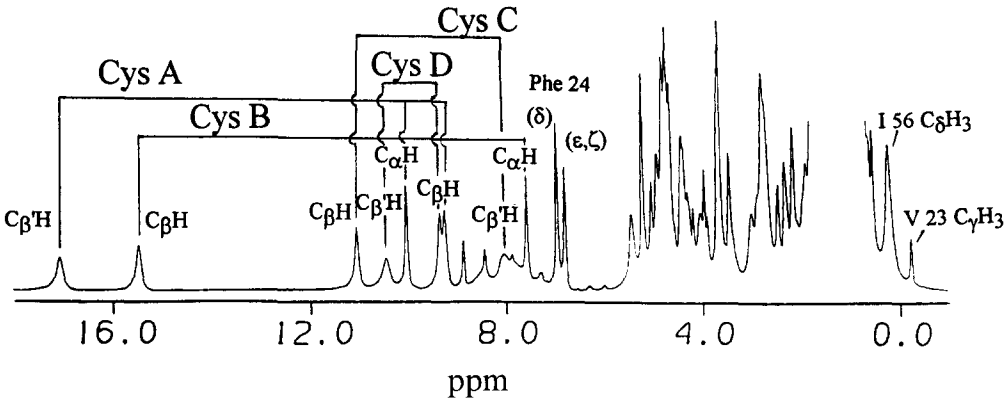

Fig. 3. 500 MHz ^{1}H NMR super-WEFT spectrum (repetition rate 4.4 s^{-1}) of *Tl* 4Fe Fd in D_2O (8 mM, D_2O, phosphate 50 mM, pH 7.5). The contact-shifted cysteine protons labeled $C_\beta H$, $C_{\beta'}H$ (close to iron), $C_\alpha H$ for Cys A–D were previously located (Donaire et al., 1994) by 2D correlation and steady-state NOE data; the remaining protons for Cys B ($C_\beta H$), Cys C ($C_\alpha H$) and Cys D ($C_\alpha H$) are under the diamagnetic envelope.

phate buffer in H_2O, pH 7.6, concentrated to 6–9 mM in an Amicon ultrafiltration device and diluted 10% with D_2O.

NMR spectroscopy

All ^{1}H NMR spectra were collected at 500 MHz on a General Electric Ω500 spectrometer. Chemical shift values were referenced to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) through the solvent signal. ^{1}H NMR experiments were collected in four different ways, i.e., the normal single pulse with H_2O presaturation (0.33 s^{-1} repetition rate) to detect slowly relaxing protons, the super-WEFT pulse sequence (Inubushi and Becker, 1983) with a repetition rate of 3 s^{-1} and a delay time of 160 ms to observe strongly relaxed protons, and the 'jump and return' pulse sequence (Sklenář and Bax, 1987), with and without presaturation of the water, to detect rapidly exchanging protons (the repetition rate was 0.33 s^{-1} and the delay between pulses was 165 μs, obtaining a maximum of the excitation profile at 3.00 ppm from the carrier frequency). The spectral width was 15.04 kHz in all cases. Nonselective T_1 values of the amide protons were estimated from the delay time at which the intensity of the signal reached a null value in a series of inversion-recovery experiments. These values were only estimated because of the overlap present in the amide region.

Steady-state 1D NOE experiments in H_2O were performed by using the super-WEFT pulse sequence (La Mar and de Ropp, 1993; Gorst et al., 1995a). Data were acquired by interleaving a block of scans with saturation on-resonance with an equal block of scans with saturation off-resonance. TOCSY spectra (Bax and Davis, 1985) with slow repetition rates (0.6 s^{-1}) and long mixing times (100 ms) were performed at three different temperatures (10, 30 and 40 °C); the WALTZ-16 pulse sequence (Shaka et al., 1983), with a soft pulse of 20–25 μs, was used as spin-lock. Presaturation on the solvent signal was achieved with the SCUBA pulse sequence (Cavanagh and Rance, 1990). The data were collected over a spectral window of 7.017 kHz with 512 t_1 values, each consisting of 2048 t_2 points. A clean-TOCSY with a faster repetition rate (2 s^{-1}) and a shorter mixing time (25 ms) was also performed at 40 °C to detect scalar connectivities corresponding to broad signals. Other conditions were the same as in the long-mixing-time TOCSY experiments. Standard NOESY experiments (Jeener et al., 1975) with long mixing time (300 ms) were performed at three different temperatures (10, 30 and 40 °C). A NOESY map with a mixing time of 50 ms and a more rapid repetition rate (2 s^{-1}) was also collected at 40 °C to detect dipolar connectivities among effectively relaxed signals (Gorst et al., 1995b). The other acquisition parameters were the same as for the TOCSY experiments. For processing, the first two data points were calculated by linear prediction in both dimensions. Prior to Fourier transformation all data points were apodized by a sine-squared function, shifted 45° and 30° in the acquisition and evolution dimensions, respectively, and zero-filled to obtain a 2048 × 2048 data point matrix. All 2D data processing was performed on a Silicon Graphics data station using the Biosym FELIX software (San Diego, CA).

Results

1D NMR experiments

The complete ^{1}H NMR spectrum of *Tl* 4Fe Fd^{ox} is displayed in Fig. 3; the previously identified hyperfine-shifted and relaxed nonlabile proton resonances are labeled $C_\beta H$, $C_{\beta'}H$ for the geminal proton further from and closer to the iron, respectively, and $C_\alpha H$ for the four ligated cysteines located previously (which are labeled Cys A–D in Fig. 3) (Donaire et al., 1994). The T_1 values for hyperfine-shifted and/or paramagnetically relaxed resolved or partially resolved resonances, together with the chemical shifts, are listed in Table 1. The amide region (6–10.2 ppm) of the 40 °C ^{1}H NMR spectrum in H_2O, collected with solvent presaturation under slow repetition conditions (0.33 s^{-1}), is expanded in Fig. 4A. Figures 4B and C

display the 'jump and return' ^{1}H NMR spectra of the same sample without and with presaturation of the solvent, respectively. As expected from the excitation profile of this pulse sequence, having a maximum at 7.71 ppm, the cysteine $C_\alpha H$ peak at 10.1 ppm displays only 20% of the intensity shown in Fig. 4A. Labeled signals in Fig. 4B exhibit reduced intensity or disappear when the H_2O resonance is saturated, indicating that these amide protons exchange relatively rapidly with solvent. In contrast, the same region of the ^{1}H NMR spectrum of a sample exchanged into D_2O is shown in Fig. 4D: the remaining signals correspond to nonlabile protons of two ligated cysteines and the Phe^{24} ring, and to slowly exchangeable amide protons, which are labeled as assigned below. Figure 4E displays the amide region of a sample kept in D_2O for 1.5 years at 5 °C. It is essentially the same as that obtained after heating the sample of Fig. 4D at 90 °C for 1 h.

The presence of rapidly relaxed proton signals in the amide/aromatic spectral window is more clearly detected by comparing the normal slow repetition rate spectrum in Fig. 5A with the super-WEFT spectrum in Fig. 5B or the partially relaxed trace from an inversion-recovery experiment in Fig. 5C. In addition to the previously characterized effectively relaxed signals from ligated cysteines and the Phe^{24} ring, three strongly relaxed ($T_1 < 20$ ms) peaks a, d and e and two intermediate relaxed ($T_1 \sim 50$ ms) labile proton peaks, b and c, are observed. Four of the relaxed labile proton signals labeled in Fig. 5C are assigned by a combination of 1D and 2D methods (see below). The nonselective T_1 values for labile protons are listed in Table 1.

Standard sequence-specific assignments

In the long-mixing-time TOCSY (100 ms), 45 spin systems can be located (two of these, valine and isoleucine, only partially; see below) with their NHs and another two complete residues with missing NHs (a proline and the terminal Met^1; not shown). In the short-mixing-time TOCSY (25 ms), two additional glycines can be located (one with missing NH) as well as one complete

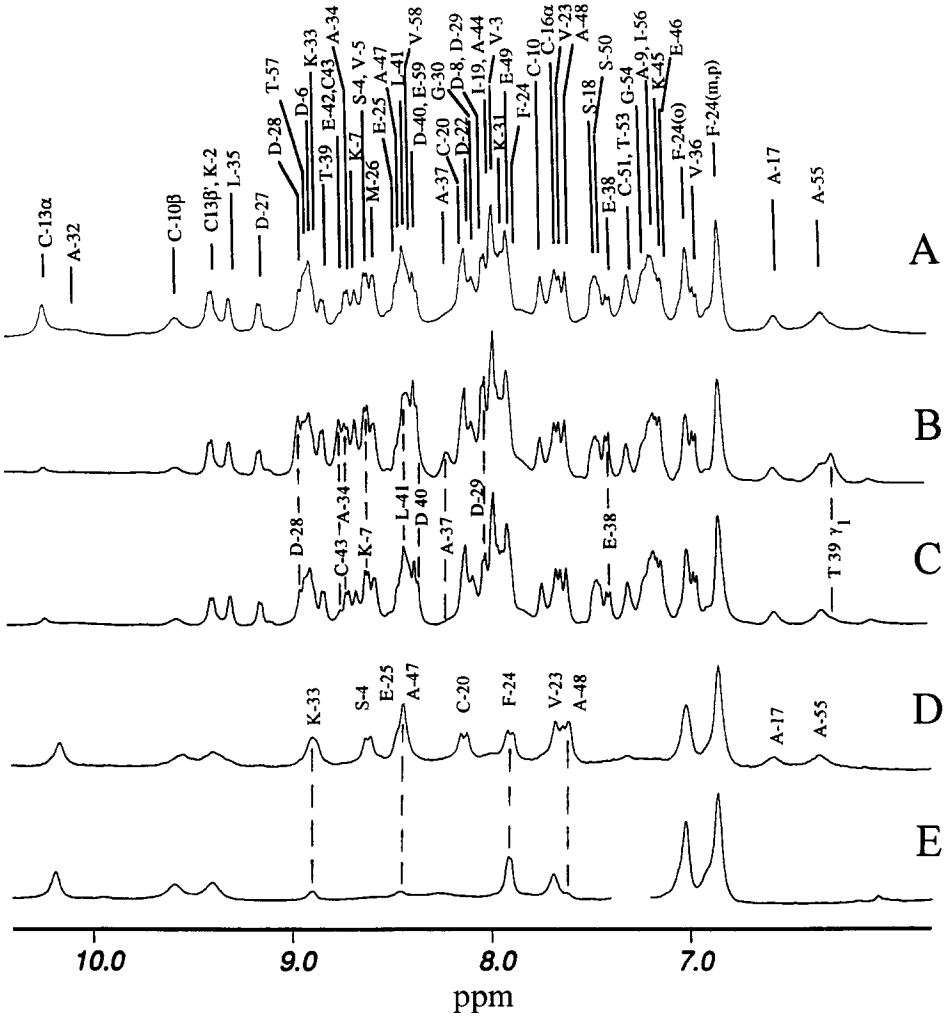

Fig. 4. Amide region of the ^{1}H NMR spectrum of *Tl* 4Fe^{ox} Fd (phosphate 50 mM, pH 7.6, 40 °C) under the following conditions: (A) in H_2O, using the standard presaturation pulse sequence on the water signal (recycle time 3 s); (B) in H_2O, using the 'jump and return' pulse sequence (Sklenář and Bax, 1987) with the carrier at 7.71 ppm; (C) same as in (B) but with presaturation of the water signal during the relaxation delay; (D) a recently exchanged sample in D_2O; and (E) a sample kept in D_2O at 5 °C for 1.5 years.

TABLE 1
^{1}H NMR SPECTRAL PARAMETERS FOR OXIDIZED *THERMOCOCCUS LITORALIS* 4Fe FERREDOXIN[a]

Residue	NH [T_1][b]	C_αH [T_1][b]	C_αH shift index	C_βH [T_1][b]	Others
Met[1]	–	4.25	–1	1.94, 2.15	2.45, 2.62
Lys[2]	9.33	4.90	1	1.91, 1.79	1.68 (δ), 1.57 (γ), 1.46 (γ), 2.97 (ε)
Val[3]	7.97	5.01	1	1.75	0.73, 0.70
Ser[4]	8.64	4.75	1	3.82	
Val[5]	8.67	4.72	1	2.26	1.19
Asp[6]	8.91	4.85	0	2.92, 2.58	
Lys[7]	8.73	3.75	–1	1.92, 1.81	1.75 (δ), 1.55 (γ), 3.02, 3.03 (ε)
Asp[8]	8.02	4.36	–1	2.61, 2.76	
Ala[9]	7.19	4.38	0	1.24	
Cys[10] (D)	7.72 [55]	3.5		10.50 [6.1], 9.40 [14.3]	
Ile[11]	d	d			
Gly[12 e]	7.28	4.37, 3.93			
	d	4.20, 3.58			
Cys[13] (A)	d	10.08 [30]		17.07 [7.5], 9.30 [14.1]	
Gly[14 e]	d	4.20, 3.58			
	7.28	4.37, 3.92			
Val[15 c]	d	5.27	1	2.56	1.06, 1.00
Cys[16] (B)	d	7.66 [28]		15.49 [11], 5.0 [≤10]	
Ala[17]	6.59 [15]	4.28	0	1.43	
Ser[18]	7.49	4.40	0	3.99	
Ile[19]	8.01	4.01	0	1.68	1.56 (γ), 1.13 (γ), 0.88 ($C_\gamma H_3$), 0.64 ($C_\delta H_3$)
Cys[20]	8.15	5.51	1	3.46, 2.71	
Pro[21]	–	5.13	1	2.65, 2.06	1.89 (γ), 3.77 (δ), 3.40 (δ)
Asp[22]	8.17	4.46	–1	2.25, 2.59	
Val[23]	7.68	3.84	–1	1.29	0.64, –0.13
Phe[24]	7.92	5.52	1	2.85, 2.24	7.03 (δ), 6.88 (ε,ζ)
Glu[25]	8.48	4.63	1	1.72, 1.99	2.19
Met[26]	8.62	4.96	1	2.13, 1.89	2.81, 2.37
Asp[27]	9.18	4.85	0	3.50, 2.57	
Asp[28]	8.97	4.36	–1	2.64, 2.74	
Asp[29]	8.03	4.50	–1	2.91, 2.61	
Gly[30]	8.09	4.10, 3.67	–1		
Lys[31]	7.96	4.69	1	2.03	1.44, 1.10 (γ), 1.75, 1.63 (δ), 2.96 (ε)
Ala[32]	10.07 [4.3]	5.27			
Lys[33]	8.89 [85]	4.73	1	1.84, 1.72	1.37, 1.06 (γ), 1.58 (δ), 2.74 (ε)
Ala[34]	8.78	5.00	1	1.33	
Leu[35]	9.33	3.97	–1	1.46	1.30 (γ), 0.73 (δ), 0.69 (δ)
Val[36]	6.98	4.49	1	2.28	0.92, 0.64
Ala[37]	8.25	4.56	1	1.48	
Glu[38]	7.43	5.33	1	1.91, 1.78	2.07, 2.15
Thr[39]	8.85	4.53	1	3.97	0.82 (CH_3), 6.34 (OH)
Asp[40]	8.41	4.92	1	2.60, 2.76	
Leu[41]	8.45	4.16	0	1.71	1.65(γ), 0.88(δ), 0.83(δ)
Glu[42]	8.79	3.98	–1	2.38	2.03
Cys[43]	8.79	4.61	–1	2.38,	2.03
Ala[44]	7.99	4.11	–1	1.39	
Lys[45]	7.19	3.97	–1	2.01, 1.84	1.44 (γ), 1.70 (δ), 2.98 (ε)
Glu[46]	7.15	3.97	–1	2.02, 1.90	2.26
Ala[47]	8.45	3.56	–1	1.06	
Ala[48]	7.64	3.74	–1	1.46	
Glu[49]	7.94	4.10	–1	2.06, 2.28	2.44
Ser[50]	7.47	4.49	0	3.74, 3.56	
Cys[51] (C)	7.28 [45]	4.9		11.10 [12.3], 8.1 [5.1]	
Pro[52]		5.11	1	2.88, 2.45	2.77 (δ), 3.09 (γ)

[a] Chemical shifts in ppm referenced to DSS, in H_2O solution, pH 7.5 at 30 °C, except where indicated.
[b] Nonselective T_1 values, in ms (±15%) are given in square brackets where peaks are partially or completely resolved.
[c] Chemical shift at 40 °C.
[d] Not located.
[e] The NH for one of the two residues Gly[12] and Gly[14] is not located and the two glycines cannot be differentiated at this time.

TABLE 1
(continued)

Residue	NH $[T_1]^b$	$C_\alpha H$ $[T_1]^b$	$C_\alpha H$ shift index	$C_\beta H$ $[T_1]^b$	Others
Thr^{53} [c]	7.28 [30]	4.04		3.95	0.98 (γ), 6.46 (OH)
Gly^{54} [c]	7.22 [20]	4.32, 3.75			
Ala^{55}	6.33 [10]	4.03	−1	2.66	1.20
Ile^{56}	7.18 [20]	4.85	1	1.49	2.73 (γ), 2.21 (γ), 0.86 ($C_\gamma H_3$), 0.33 ($C_\delta H_3$) [10]
Thr^{57}	8.96	4.54	1	3.95	1.05
Val^{58}	8.46	5.23	1	1.86	0.85, 0.77
Glu^{59}	8.38	4.36	0	1.86, 1.96	2.11

alanine, one valine without NH, a partial proline, and a complete threonine. These bond correlation experiments account for all but two of the 55 residues not ligated to the cluster, the two missing residues being an alanine and an isoleucine. The specific assignment of the residues not affected by the paramagnetism of the cluster were performed according to the standard methodology (Wüthrich, 1986). The backbone NOESY connectivities (Fig. 2G) between sequential amino acids are clearly observed for the segments 1–9, 17–31, 33–50 and 56–59. The sequential connections for Ala^{37} and Cys^{43} were only observed at 10 °C, since the amide protons of these two residues exhibit significant saturation with the solvent at higher temperatures (see Fig. 4C). Although Val^5 and Ile^{56} are only partially identified in the TOCSY map due to relaxation effects, both are uniquely assigned by their connectivities

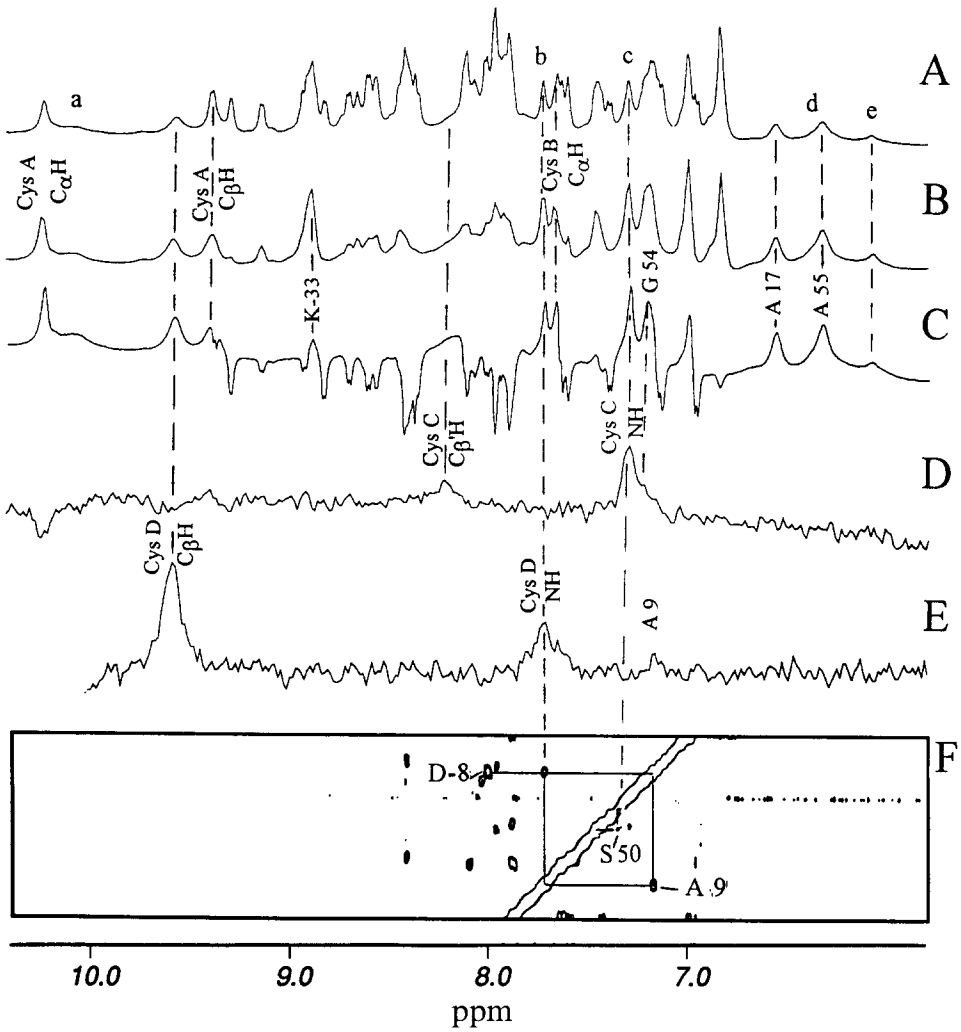

Fig. 5. Amide region of the ^1H NMR spectrum of *Tl* 4Fe Fd in H_2O recorded under the following conditions: (A) same spectrum as in Fig. 4A; (B) a super-WEFT spectrum (3 s^{-1}, 160 ms relaxation delay); (C) partially relaxed spectrum with a delay of 75 ms and a repetition rate of 2 s^{-1}; (D,E) steady-state NOE difference traces upon saturation of $C_\beta H$ of Cys C (D) or $C_{\beta'} H$ of Cys D (E); and (F) portion of the NOESY spectrum with 50 ms mixing time, exhibiting backbone connectivities from NHs of Cys C and Cys D to NHs of Ser^{50} and Ala^9, respectively.

to Asp[6] and Thr[57], respectively. The backbone connectivities that allow the assignments are summarized graphically in Fig. 2G. The chemical shifts for the assigned residues are listed in Table 1. The chemical shift indices (Wishart et al., 1991; +1 for ≥0.1 ppm downfield; 0 for a <0.1 ppm shift; −1 for a ≥0.1 ppm shift upfield for $C_\alpha H$) in *Tl* Fd relative to an isolated residue are also included.

The three unique spin systems (valine, threonine and proline) observed in the 25 ms but not in the 100 ms TOCSY map fail to exhibit backbone cross peaks to other residues in both the 50 and 300 ms NOESY maps; they can be tentatively assigned to Val[15], Thr[53] and Pro[52], the only unassigned residues of each of these amino acids. The two remaining glycine resonances, observed only in the 25 ms TOCSY map (one without NH), must arise from Gly[12] and Gly[14], but fail to exhibit NOESY cross peaks to differentiate them. The alanine resonance detected solely in the 25 ms TOCSY map must arise from the remaining Ala[32] or Ala[55]. Direct evidence for Ala[32] in the 2D maps are the 50 ms (but not 300 ms) NOESY cross peaks from both Lys[33] NH and Met[26] $C_\alpha H$ to a signal at 5.27 ppm typical for a $C_\alpha H$ (confirmed by 1D NOEs and the secondary structure, see below). The alanine detected only in the 25 ms TOCSY map is therefore assigned to the only remaining residue, Ala[55]. Previous 1D and 2D experiments had identified $C_\beta Hs$ and $C_\alpha H$ for four cysteines labeled Cys A–D; see Fig. 3. Hence the 2D experiments identify the majority of all protons for 57 of the 59 residues, and one signal for Ala[32]. The available 2D data at both long and short mixing times failed to provide any information on only one residue, Ile[11].

However, two strong NOESY cross peaks are observed in the amide region at 50 ms (but much weaker or absent at 300 ms) mixing time, involving labile protons not identified above, one from Ala[9] NH to 7.73 ppm (peak b in Fig. 5C) and one from Ser[50] NH to 7.28 ppm (peak c in Fig. 5C). The origin of the labile proton peaks b and c is established by 1D NOEs (see below).

Steady-state NOEs

Irradiation of $C_{\beta'}H$ of Cys D in H_2O results in a strong NOE to $C_\beta H$ of Cys D (previously described in D_2O (Donaire et al., 1994)), as well as a moderate (~−3%) NOE to labile proton peak b and a weak NOE to Ala[9] NH, as shown in Fig. 5E. The proximity to Cys D $C_\beta H$, the T_1 of ~50 ms and, most importantly, the NOESY cross peak to Ala[9] NH in the 50 ms NOESY map shown in Fig. 5F, identify peak b as the NH of Cys D, and, therefore, Cys D as arising from Cys[10] (**I**). The $C_\alpha H$ resonance of Cys[10] was not detected. The remaining strong NOE to Cys[10] $C_\beta H$ observed in D_2O (Donaire et al., 1994) is identified as arising from Lys[31] $C_\alpha H$. Saturation of $C_\beta H$ of Cys C in H_2O results in the NOE to the strongly relaxed $C_{\beta'}H$ of Cys C, as reported previously in D_2O, as well as a strong NOE (~−3%) to labile proton c, and a weaker NOE to Gly[54] NH, as shown in Fig. 5D. The magnitude of the NOE from the Cys C $C_\beta H$ and the relaxation properties of peak c establish that it arises from NH of Cys C. The NOESY cross peak between peak c and the Ser[50] NH in the short-mixing-time NOESY spectrum (Fig. 5F) uniquely establishes that Cys C arises from Cys[51] (**IV**). This is confirmed by the NOE from $C_{\beta'}H$ of Cys C to the NH of Gly[54] (Fig. 5D), as expected from the crystal structure of *Dg* Fd (see below).

Saturation of $C_\beta H$ and $C_\alpha H$ of Cys B in H_2O yields NOEs to Ile[19] NH and $C_\beta H$, respectively, which identify

TABLE 2
TERTIARY NOEs BETWEEN RESIDUES ON IDENTIFIED SECONDARY STRUCTURAL ELEMENTS IN OXIDIZED *Tl* 4Fe FERREDOXIN[a]

Structural element	Residue	Residue	Structural element
α-Helix B	Cys[43] $C_\beta Hs$ [Cys[42]][b]	- [Ile[17]] Ile[19] $C_\delta H_3$ (m)	Helix A
	Ala[44] $C_\beta H_3$ [Val[43]]	- [Ile[2]] Val[3] $C_\gamma H_3$ (m)	
	Lys[45] $C_\delta H$ (m), $C_\varepsilon H$ (m) [Glu[44]]	- [Val[57]] Val[58] $C_\gamma H_3$ (m)	β-Sheet A
	Ala[47] $C_\beta H_3$ (m) [Ala[46]]; Ala[48] $C_\beta H_3$ (m) [Ile[47]]	- [Ile[55]] Ile[56] $C_\gamma H_3$ (m)	
Hydrophobic core	Val[23] $C_\gamma H_3$ [Val[21]]	- [Ile[2]] Val[3] $C_\alpha H$ (m), $C_\gamma H_3$ (s);	
	Val[23] $C_\gamma H_3$ [Val[21]]	- Thr[39][c] $C_\beta H$ (m), $C_\gamma H_3$ (s)[b]	
	Phe[24] ring [Phe[22]]	- [Val[43]] Ala[44] $C_\beta H_3$ (s); [Ala[46]] Ala[47] $C_\beta H_3$ (m)	α-Helix B
	Val[23] $C_\gamma H_3$ (m) [Val[21]]; Phe[24] ring (w) [Phe[22]]	- [Ala[32]] Ala[34] $C_\alpha H$	β-Strand B
	Phe[24] ring [Phe[22]]	- [Cys[14]] Cys[16] $C_\beta H$ (m)	α-Helix A
Cluster	Cys[16] $C_\beta H$ [Cys[14]]	- [Phe[23]] Phe[24] ring (m), $C_\beta Hs$	Hydrophobic core
	Cys[51] $C_\beta H$ [Cys[50]]	- [Ile[55]] Ile[56] $C_\gamma H_3$ (m), $C_\delta H_3$ (w)	β-Sheet A
	Cys[16] $C_\alpha H$(s), $C_\beta H$(s) [Cys[14]]	- [Ala[46]] Ala[47] $C_\beta H_3$ (s)	α-Helix B
	Cys[10] $C_\beta H$ [Cys[8]]	- [Lys[30]] Lys[31] $C_\alpha H$ (s), $C_\gamma H$ (s) $C_\delta H$ (m)	β-Sheet B

[a] NOESY cross peak intensity (except for Cys[10], Cys[16] and Cys[51] and Ile[56] $C_\delta H$, where steady-state NOEs are used) with strong (s), medium (m) and weak (w) intensity.
[b] Residues in *Dg* Fd on the same secondary structural element that are expected to exhibit similar interchain NOEs are given in square brackets.
[c] The third strand in β-sheet A is not present in *Dg* Fd.

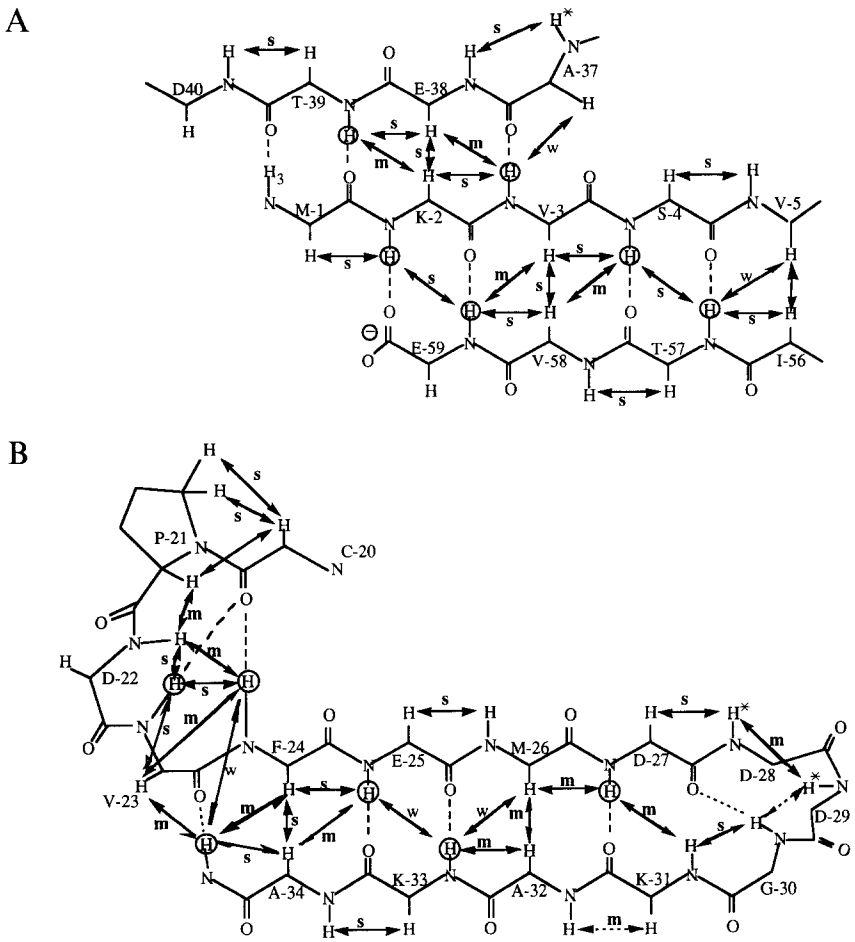

Fig. 6. Schematic representation of: (A) triple-strand antiparallel β-sheet A; connectivities observed in the 2D NOESY map with 300 ms mixing time are shown by arrows; (B) antiparallel β-sheet B; connectivities observed in the long- (300 ms) and short- (50 ms) mixing-time NOESY and by 1D NOE of Ala^{32} NH are shown by arrows. The inferred hydrogen bonds are shown by dashed lines. Dotted lines indicate expected NOEs that are not observed due to degeneracy. Slowly exchanging labile protons are circled.

Cys B as Cys^{16} (**III**) (Fig. 2I). Strong NOEs from Cys B $C_\beta H$ to the Phe^{24} ring and Ala^{47} $C_\beta H_3$, as well as strong NOEs from Cys^{16} (**III**) $C_\alpha H$ to both $C_\alpha H$ and $C_\beta H_3$ of Ala^{47} (in D_2O) are also observed (see Table 2). It was not possible to locate the Cys^{16} NH signal. $C_\beta H$ for Cys A similarly failed to exhibit a NOE to a labile proton, although both $C_\beta Hs$ and the $C_\alpha H$ were identified. Since Cys B, C and D have been established to arise from Cys^{16} (**III**), Cys^{51} (**IV**) and Cys^{10} (**I**), respectively, Cys A must arise from Cys^{13} (**II**). Weak NOEs from Cys^{13} (**II**) $C_\beta H$ to resonances at 2.58 and 2.81 ppm, observed previously in D_2O (Donaire et al., 1994), can be identified here as arising from the relaxed Val^{15} $C_\beta H$ and a Pro^{52} $C_\gamma H$, detected only in short-mixing-time experiments (see above). These contacts would be unique to Cys^{13} (**II**) and directly confirm the assignment.

The upfield broad, rapidly relaxed ($T_1 \sim 10$ ms) methyl peak at 0.33 ppm in Fig. 3 had previously been shown to exhibit NOEs to the Phe^{24} ring and a short-mixing-time (25 ms) TOCSY peak to the resonance at 2.73 ppm, and was suggested to arise from Ile^{56} (Donaire et al., 1994). The present TOCSY and NOESY data extend the available data for this residue to the backbone and confirm its origin as Ile^{56}. The remaining resolved hyperfine-shifted and relaxed signal a at 10.08 ppm arises from a labile proton and fails to exhibit any scalar correlation. However, saturation of peak a in H_2O results in a strong NOE to the $C_\alpha H$ of Lys^{31} (see Fig. 2I) and much weaker NOEs to NH of Lys^{33} and the $C_\alpha Hs$ of Met^{26} and Ala^{32}. These connectivities are unique for the NH of Ala^{32}, and, moreover, place it in a β-strand (see below). The chemical shifts for all assigned signals, and the T_1 estimates for the resolved or partially resolved resonances, are listed in Table 1.

Secondary structure

The characteristic strong α_i-N_{i+1} NOESY cross peaks identify five extended segments, 1–5, 24–27, 31–34, 37–39 and 56–59. The strong α_i-α_j and moderate α_i-N_{j+1} and N_{i+1}-N_{j-1} NOEs identify two antiparallel β-strands. An apparent triple strand (labeled β-sheet A) involves the two termini (1–5 and 56–59) and the short section 37–39, with

the N-terminus as the central strand, as reflected by the NOESY cross peak pattern depicted in Fig. 6A. Slowly exchanging protons are indicated by a circle, and the interstrand hydrogen bonds inferred from the NOEs and slow exchange rates are shown in dashed lines. The α_5-α_{56} NOESY cross peak is not detected due to near degeneracy, but the Thr[57] NH to Ser[4] CO hydrogen bond is supported by the NOESY cross peak pattern for Thr[57] NH to the Val[5] and Ile[56] C_αHs. The C_αH chemical shift indices in Table 1 support the β-sheet structures. The residues Phe[24]–Asp[27] and Lys[31]–Ala[34] exhibit the NOESY cross peaks characteristic for an antiparallel β-strand in the protein interior (β-strand B); the two strands are connected by a tight type I turn (likely Asx, and labeled turn C), characterized by strong to medium intensity N_i-N_{i+1} NOESY cross peaks as depicted schematically in Fig. 6B (the N_{29}-N_{30} cross peak is too close to the diagonal to detect). The C_αH shifts in Table 1 support the β structure. The inferred hydrogen bonds are shown by dashed lines and slowly exchanging NHs are circled. The tight type I (turn B) turn prior to the N-terminus of β-strand B involves Pro[21]–Phe[24], the highly conserved hydrophobic core, and exhibits the same NOESY cross peak pattern as observed for the homologous turn B for Pro[22]–Phe[25] in *Pf* 3Fe Fd (Teng et al., 1994), and as expected for Pro[19]–Phe[22] in the crystal structure from *Dg* 3Fe Fd (Kissinger et al., 1991). This hydrophobic core, in particular segment Val[23]–Phe[24], represents the most stable portion of the protein, as witnessed by the extraordinarily slow exchange rate for NH of Phe[24] (Fig. 4E).

Strong N_i-N_{i+1}, α_i-N_{i+3} and α_i-β_{i+3} (Fig. 2G) NOESY cross peaks are observed that dictate the presence of a helix (labeled helix B in Fig. 2F) comprising Cys[43]–Ser[50]. Overlap prevented detection of some of these characteristic peaks for Glu[42], but the value for the C_αH shift index in Table 1 suggests that Glu[42] may be included in the helix. The observation of several α_i-N_{i+4} NOESY cross peaks (α_{43}-N_{47}, α_{44}-N_{48}, α_{46}-N_{50}) favors an α-helix. A second short helix, including Cys **III** and several of the subsequent residues in the crystal structure of single-cluster cubane Fds, has been shown by NMR to be conserved for *Pf* 3Fe Fd in solution (Teng et al., 1994). The observation of β_i-N_{i+3} and α_i-β_{i+3} NOEs for Cys[16] (**III**)-Ile[19] and Ala[17]-Cys[20] suggests that this helix is retained (labeled helix A in Fig. 2F) in *Tl* 4Fe Fd. Strong paramagnetic influences weaken or render undetectable other expected NOESY cross peaks for this helix and possible dipolar shifts undermine the use of the C_αH shift index (see Discussion).

The two turns present just before Cys **I** (turn A) and just after Cys **IV** (turn E) in *Bt* and *Dg* Fds are also present in *Tl* 4Fe Fd. The strong N_i-N_{i+1} NOESY peaks for residues Asp[6]–Ala[9] are consistent with an Asx type I turn, as observed in *Pf* 3Fe Fd (Teng et al., 1994). Most signals in the segment Cys[51]–Gly[54] are effectively relaxed by the cluster paramagnetism (see below). However, the steady-state NOE from Cys[51] (**IV**) C_βH to both the NH and C_αH of Gly[54] (Fig. 2I) is that expected for the highly conserved type I turn (turn E) involving Cys[51]–Gly[54] and supports its retention in *Tl* 4Fe Fd. The segment between the second strand of β-sheet B and the three-stranded β-sheet A, Leu[35]–Ala[37], is characterized by strong α_i-N_{i+2} but no N_i-N_{i+2} NOESY cross peaks (Fig. 2G), and hence cannot be classified in terms of any conventional turn.

Disulfide bridge

The positions of Cys[20] (**V**) and Cys[43] (**VI**) in *Tl* Fd are homologous to the cysteines that participate in a disulfide bridge in the crystal structure of *Dg* 3Fe Fd (Kissinger et al., 1991) and in the solution structures of both 3Fe and 4Fe *Pf* Fd^{ox} (Teng et al., 1994; Gorst et al., 1995b). The present NMR data provide strong support for the presence of this disulfide bridge through strong and medium intensity NOESY cross peaks between the C_αH and C_βHs of Cys[20] (**V**) and the C_βHs of Cys[43] (**VI**). Other observed dipolar connectivities indicative of the disulfide bond are listed in Table 2.

Tertiary contacts

Major long-range NOESY cross peaks or steady-state NOEs with significant intensity, which identify tertiary contacts among the various secondary structural elements, are listed in Table 2 (with intensities listed as s (strong), m (medium) and w (weak)). The contacts to the cluster-ligated cysteines and to Ile[56] $C_\delta H_3$ were determined from steady-state NOEs.

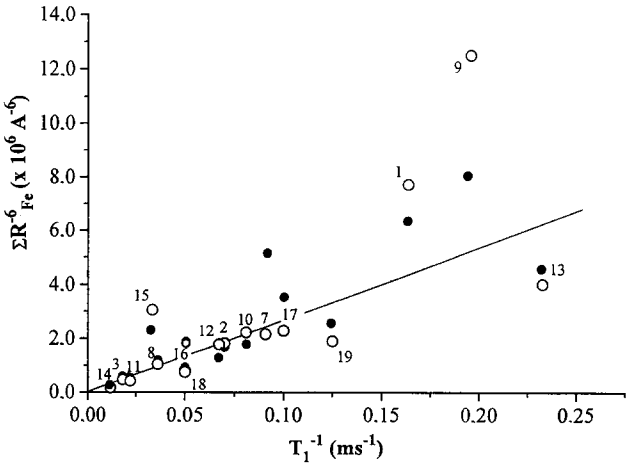

Fig. 7. Plot of T_1^{-1} for paramagnetically influenced resolved and partially resolved protons in *Tl* 4Fe Fd (see Table 1) as a function of $\sum R^{-6}$ (Eq. 1) for the assumed isostructural protons in 3Fe *Dg* Fd (○) and 4Fe *Bt* Fd(●) crystal structures. The protons presented are identified as: 1, Cys[10] $C_{\beta'}$H; 2, Cys[10] C_βH; 3, Cys[10] NH; 4, Cys[13] $C_{\beta'}$H; 5, Cys[13] C_βH; 6, Cys[13] C_αH; 7, Cys[16] C_βH; 8, Cys[16] C_αH; 9, Cys[51] $C_{\beta'}$H; 10, Cys[51] C_βH; 11, Cys[51] NH; 12, Ala[17] NH; 13, Ala[32] NH; 14, Lys[33] NH; 15, Thr[53] NH; 16, Gly[54] NH; 17, Ala[55] NH; 18, Ile[56] NH; 19, Ile[56] $C_\delta H_3$.

Relaxation influences

The effect of paramagnetism by a spin-coupled cluster on nearby protons, H_i, in $[Fe_4S_4]^{+2}$, for which the four iron environments are considered to be essentially identical ($4Fe^{+2.5}$), is given by (Banci et al., 1991; Gorst et al., 1995a):

$$T_{1i}^{-1} = D \sum^{q} R_{Fe}^{-6} \quad (1)$$

where D is a constant, and R_{Fe} is the separation between H_i and Fe-q, where q = **I**–**IV**, the iron atoms ligated to Cys **I**–**IV**. Hence, relative relaxation rates for nonequivalent protons are expected to be proportional to the relative values of $\sum R_{Fe}^{-6}$. Plots of the observed T_1^{-1} for the assigned protons with dominant paramagnetic relaxation influences ($T_1 \leq 200$ ms) for *Tl* 4Fe Fd versus the $\sum R_{Fe}^{-6}$ computed for the same protons on the homologous residue in *Dg* 3Fe Fd (open circles) (Kissinger et al., 1991) and in *Bt* 4Fe Fd (closed circles) (Fukayama et al., 1989) are illustrated in Fig. 7.

Discussion and Conclusions

Assignment strategy

By means of conventional long-mixing-time 2D NMR experiments, we were able to locate and assign sequentially 47 of the 59 residues, with only few proton signals undetected (such as one methyl for Val^5 and the $C_\gamma H$-$C_\delta H_3$ fragment of Ile^{56}). Short-mixing-time 2D experiments provided an additional six complete or major portions of residues, which could be identified by their unique topology. 1D and 2D data, moreover, located two of the signals of an alanine. A combination of short-mixing-time 2D and steady-state NOE data in D_2O had previously located the hyperfine-shifted partial spin systems for four ligated cysteines (Donaire et al., 1994). This leaves only one residue unidentified, namely Ile^{11}, which is expected to have its backbone very near the cluster. However, it is assumed that even the Ile^{11} side chain signals can be located once all 2D cross peaks in the 5–0 ppm window are assigned; such studies are in progress. The expansion over conventional 2D experiments of the information content regarding close-range (backbone) interactions and long-range cluster–protein tertiary contacts by using 2D and 1D experiments tailored to detect strongly relaxed protons is evident in Figs. 2H and I, respectively. The detection of a larger fraction of the total residues in the present *Tl* 4Fe Fd, when compared to those in *Pf* 3Fe Fd (Teng et al., 1994), can be traced to the more favorable electron relaxation properties of the $[Fe_4S_4]^{+2}$ compared to the $[Fe_3S_4]^{+1}$ clusters.

All protons, except those assigned to $C_\beta Hs$ or $C_\alpha H$ of a ligated cysteine, exhibit chemical shifts consistent with essentially 'diamagnetic' environments (see Table 1), although numerous of these signals exhibit strongly enhanced relaxation rates due to their proximity to the cluster. The absence of significant dipolar shifts for even strongly relaxed protons from noncoordinated residues dictates that the cluster magnetic moment exhibits negligible anisotropy, and directly confirms the early postulate of a contact origin for the ligated cysteine hyperfine shifts (Poe et al., 1970).

We have previously proposed (Gorst et al., 1995a) that the apparent highly conserved disposition of the cysteines in the ligation consensus sequence relative to the cubane cluster in bacterial-type Fds will allow systematic identification of their contacted shifted residues using the expected NMR detectability of the $C_\alpha Hs$ of Cys **II** and **III**, and the peptide NHs of Cys **I** and **IV**. Indeed, the weakly relaxed and insignificantly hyperfine-shifted NHs of Cys^{10} (**I**) and Cys^{51} (**IV**), as well as the weakly relaxed but relatively strongly contact-shifted $C_\alpha Hs$ of Cys^{13} (**II**) and Cys^{16} (**III**), are observed for 4Fe *Tl* Fd^{ox}. These contrasting relaxation properties are precisely what is predicted by the twofold pseudosymmetry for the ligated Cys (pairs **I**, **IV** versus **II**, **III**), as shown schematically in Fig. 1. Consistent with the expectation from this cluster geometry, the $C_\alpha Hs$ of Cys^{10} (**I**) and Cys^{51} (**IV**) in *Tl* 4Fe Fd^{ox} are strongly relaxed but have been detected in short-mixing-time TOCSY spectra (Donaire et al., 1994).

Ligated cysteine backbone NOEs to their adjacent residues facilitate their sequence-specific assignments, thereby separating the assignment of these residues from any assumptions about tertiary structure, as has been the case in all previous assignments of ligated cysteines in iron–sulfur proteins. The present definitive cysteine assignments confirm our earlier proposals based on an assumed tertiary structure similar to *Dg* Fd (Donaire et al., 1994). This independent identification of the cysteines sets the stage for the extraction of important data from the hyperfine-shifted residue signals, namely direct information on the tertiary contacts between the cluster and the remainder of the protein. The use of the NOEs from assigned contact-shifted cysteines to elucidate the local structure, rather than the conventional converse, strongly enhances the prospects for complete solution structure determination of bacterial Fds for which crucial structural information for one portion of the protein is obliterated in conventional 2D NMR maps.

Folding topology

Tl 4Fe Fd possesses two antiparallel β-sheets (one three- (A) and one two-stranded (B)), one long helix (B) and another likely short helix (A), and four turns (A–C and E), in addition to the cubane cluster, as shown in Fig. 2F. These same elements, albeit somewhat modified, have been identified in both the crystal and the solution structure of other single cubane cluster Fds, in particular in *Dg* 3Fe Fd (Fig. 2D), *Bt* 4Fe Fd and *Pf* 3Fe Fd (Fig. 2B) (Fukayama et al., 1989; Kissinger et al., 1991; Teng

et al., 1994). The overall structure of *Tl* 4Fe Fd is remarkably similar to that of *Dg* 3Fe Fd, to which it also has the closest sequence homology among the structurally characterized Fds (compare Figs. 2C and E). The secondary structural elements in these two proteins can be compared in Figs. 2D and F. The NOESY pattern for residue Cys^{20} (**V**) to Phe^{24}, which initiates turn B, is consistent with a completely conserved hydrophobic core when compared to the crystal coordinates of *Dg* Fd. The internal β-sheets B are very similar in *Tl* and *Dg* Fds, except that the turn between the two strands is tighter in the former; in fact, β-sheet B in *Tl* Fd more closely resembles that in *Bt* Fd. The major α-helix B in *Tl* and *Dg* Fds consists of nine residues, the second residue of which, Cys **VI**, forms the disulfide bridge in *Tl* and *Dg* Fd; the helix terminates just prior to the cluster-ligated Cys **IV**. Turn E in *Tl* Fd, while incompletely defined by NMR due to relaxation effects, involves very conservative substitution. Turn A, just before the cluster, appears to be initiated one residue sooner (Asp^6) in *Tl* than in *Dg* Fd, and resembles more closely that in *Bt* 4Fe Fd, which also exhibits the apparent insertion prior to Cys **I**. The only qualitatively significant difference between *Tl* and *Dg* Fd is that β-sheet A incorporates a third antiparallel strand in the former, consisting of the internal residues 38–40 that form three hydrogen bonds to the N-terminal strand, as depicted in Fig. 6A. The only similar backbone contact in *Dg* Fd is a single hydrogen bond from Ile^2 to the carboxyl of Pro^{36} at the beginning of a β-turn.

The long-range interchain contacts for *Tl* Fd, characterized by NMR in terms of interactions between pairs of secondary structural elements, are listed in Table 2; the analogous residues with side chains exhibiting contacts ≤ 3 Å for *Dg* Fd are given in brackets. In each case, the contacts between side chains are those predicted for the homologous residues in the same secondary structural elements in *Dg* Fd. The contacts to the hydrophobic core are particularly strongly conserved. Consistent with its strong sequence homology, the solution structure of *Tl* Fd (Fig. 2F) more closely resembles that of *Dg* Fd (Fig. 2D) than did *Pf* Fd (Fig. 2B). Of particular interest is the observation that α-helix B in *Tl* Fd is essentially identical to that in *Dg* Fd; that for *Pf* Fd was found to be both lengthened and shifted significantly toward the N-terminus when compared to *Dg* Fd (Teng et al., 1994). The extension of the β-sheet to include the two terminal residues and the incorporation of a third strand is common to *Tl* and *Pf* Fd.

The stability of the various secondary structural elements is reflected in the exchange rates for the peptide protons with solvent. The 15 slowly exchanging NHs (Figs. 4D and E) reside in six of the seven (all but the N-terminus) hydrogen bonds in β-sheet A (Fig. 6A), six in β-sheet B and turn C (Fig. 6B), and three in α-helix B. The extreme stability of the hydrophobic core is shown by the very slow exchange of the Phe^{24} NH (Fig. 4E). The relatively rapidly exchanging protons, as observed by saturation transfer from solvent (Figs. 4B and C), occur in regions between secondary structural elements (Asp^{28} and Asp^{29} on the outside of the turn between the two strands of β-strand B) in the structurally undefined segment Ala^{37}, Glu^{38}, Asp^{40} between β-sheet B and the internal strand of β-sheet A (note that the internal Thr^{39} exchanges very slowly because of the H-bond in β-sheet A (Fig. 6A).

The present level of the structural characterization of *Tl* Fd is insufficient to reach conclusions as to the origin of the extreme thermostability of the Fd. However, common motifs found in 3Fe *Pf* Fd and 4Fe *Tl* Fd, but not in mesophilic Fds, are the lengthening of the surface β-sheet to include the terminal residues, and the incorporation of a third strand in this β-sheet, as well as a tightening of the turn between the two strands of the β-sheet near the hydrophobic core. A similar expansion of the β-sheet involving the termini has been observed in rubredoxin from *Pf* (Blake et al., 1992; Day et al., 1992). The secondary structure of the 4Fe Fd from the hyperthermophilic bacterium *Thermotoga maritima*, *Tm*, has been described previously (Wildegger et al., 1995). The *Tm* Fd has 75% sequence homology to *Tl* Fd, but contains a Pro^{37} that precludes the participation of a third strand in β-sheet A. The ~20 °C lower optimal growth temperature for *Tm* compared to *Pf* suggests a role for the expanded β-sheet in contributing to the *Tl* Fd thermostability.

The cluster environment

The NOE patterns from ligated Cys C_βHs to assigned side chains in *Tl* Fd are indicative of tertiary contacts as found in *Dg* Fd (Kissinger et al., 1991). The proximity to the cluster paramagnetism of both side chain and main chain protons provides further support for this conclusion. Thus, the loss of one methyl of Val^5, the methyl of Ala^{32}, and the strong relaxation of the Ile^{56} $C_\delta H_3$, as well as the suppression of most NOESY cross peaks for residues between Cys^{10} (**I**) and Cys^{16} (**III**), and for the C-terminus of Cys^{51} (**IV**), including the failure to detect Ile^{11}, are all consistent with expected ≤ 5 Å distances from a cluster iron to the homologous protons in *Dg* (or *Bt*) Fd.

For the resolved or partially resolved signals whose relaxation is dominated by paramagnetic influences, the estimated relaxation rates correlate reasonably well with the $\sum R^{-6}$ from the cluster iron for presumed isostructural protons in *Dg* or *Bt* Fd, as shown in Fig. 7. The T_1 values of the residues included in these plots are listed in Table 1. It should be noted that data points 11–19, which involve only protons with negligible hyperfine shifts, exhibit a reasonably well defined correlation between T_1^{-1} and $\sum R^{-6}$ for both reference proteins. The absence of systematic deviations (i.e., for protons close to one iron or another) for points 11–19 indicates that the assumed

symmetrical distribution of the spin magnetization over the four irons (which is implicit in Eq. 1) is reasonable. On the other hand, the signals for the contact-shifted cysteines, data points 1–10, exhibit significantly more scatter, but generally about a line with a slope very similar to that obtained from data points 11–19. We interpret these relaxation results as directly supporting the largely conserved cluster tertiary contacts in *Tl* Fd when compared with those of both *Dg* and *Tl* Fds. It is not possible at this time, however, to conclude whether the cluster environment of *Tl* Fd more closely resembles that in *Dg* or in *Bt* Fd. The similar correlation between T_1^{-1} and $\sum R^{-6}$ for both contact-shifted and non-contact shifted protons in Fig. 7 argues that the relaxation of the ligated cysteine signals is determined primarily by metal(s)-centered dipolar relaxation. These results allow optimism regarding the prospects for using such relaxation data as molecular constraints, which, together with the NOESY/NOE constraints presented above, will allow the formulation of a realistic molecular model for cubane Fds. Such modeling studies of *Tl* and *Pf* Fd are in progress.

Acknowledgements

The authors are indebted to Drs. Jeffrey S. de Ropp, Carol M. Gorst and Quincy Teng for experimental assistance and valuable discussions. This research was supported by grants from the National Science Foundation, DMB 91-04018 (G.N.L.) and MCB 94-05783 (M.W.W.A), and from the National Institutes of Health, GM 45597 (M.W.W.A).

References

Adman, E.T., Sieker, L.C. and Jensen, L.H. (1976) *J. Biol. Chem.*, **251**, 3801–3806.

Backes, G., Mino, Y., Loehr, T.M., Meyer, T.E., Cusanovich, M.A., Sweeney, W.V., Adman, E.T. and Sanders-Loehr, J. (1991) *J. Am. Chem. Soc.*, **113**, 2055–2064.

Banci, L., Bertini, I. and Luchinat, C. (1991) *Nuclear and Electronic Relaxation*, VCH Publishers, Weinheim, pp. 143–155.

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

Beinert, H. (1990) *FASEB J.*, **4**, 2483–2491.

Bertini, I., Capozzi, F., Luchinat, C. and Vila, A.J. (1994) *J. Am. Chem. Soc.*, **116**, 651–660.

Blake, P.R., Park, J.-B., Zhou, Z.H., Hare, D.R., Adams, M.W.W. and Summers, M.F. (1992) *Protein Sci.*, **1**, 1508–1521.

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

Cammack, R. (1992) *Adv. Inorg. Chem.*, **38**, 281–322.

Cammack, R., Dickson, D. and Johnson, C. (1977) In *Iron Sulfur Proteins* (Ed., Lovenberg, W.), Academic Press, New York, NY, pp. 283–330.

Cavanagh, J. and Rance, M. (1990) *J. Magn. Reson.*, **88**, 72–85.

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

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

Donaire, A., Gorst, C.M., Zhou, Z.H., Adams, M.W.W. and La Mar, G.N. (1994) *J. Am. Chem. Soc.*, **116**, 6841–6849.

Duie, E.D., Fanchon, E., Vicat, J., Sieker, L.C., Meyer, J. and Moulis, J.-M. (1994) *J. Mol. Biol.*, **243**, 683–695.

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

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

Gorst, C.M., Yeh, T., Teng, Q., Calzolai, L., Zhou, J.H., Adams, M.W.W. and La Mar, G.N. (1995a) *Biochemistry*, **34**, 600–610.

Gorst, C.M., Zhou, Z.-H., Ma, K., Teng, Q., Howard, J.B., Adams, M.W.W. and La Mar, G.N. (1995b) *Biochemistry*, **34**, 8788–8795.

Howard, J.B. and Rees, D.C. (1991) *Adv. Protein Chem.*, **42**, 199–281.

Inubushi, T. and Becker, E.D. (1983) *J. Magn. Reson.*, **51**, 128–133.

Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1975) *J. Chem. Phys.*, **71**, 4546–4553.

Kissinger, C.R., Sieker, L.C., Adman, E.T. and Jensen, L.H. (1991) *J. Mol. Biol.*, **219**, 693–723.

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

Luchinat, C. and Ciurli, S. (1993) In *Biological Magnetic Resonance*, Vol. 12 (Eds, Berliner, L.J. and Reuben, R.), Plenum Press, New York, NY, pp. 357–420.

Macedo, A.L., Moura, I., Surerus, K.K., Papaefthymiou, V., Liu, M.-Y., LeGall, J., Munk, E. and Moura, J.J.G. (1994) *J. Biol. Chem.*, **369**, 8052–8058.

Mukund, S. and Adams, M.W.W. (1993) *J. Biol. Chem.*, **268**, 13592–13600.

Poe, M., Phillips, W.D., McDonald, C.C. and Lovenberg, W. (1970) *Proc. Natl. Acad. Sci. USA*, **65**, 797–804.

Sery, A., Housset, D., Serre, L., Bonicel, J., Hatchikian, C., Frey, M. and Roth, M. (1994) *Biochemistry*, **33**, 15408–15417.

Shaka, A.J., Keeler, J. and Freeman, R. (1983) *J. Magn. Reson.*, **53**, 313–340.

Sklenář, V. and Bax, A. (1987) *J. Magn. Reson.*, **74**, 469–479.

Stout, C.D. (1989) *J. Biol. Chem.*, **263**, 9256–9266.

Teng, Q., Zhou, Z.H., Smith, E.T., Busse, S.C., Howard, J.B., Adams, M.W.W. and La Mar, G.N. (1994) *Biochemistry*, **33**, 6316–6326.

Wildegger, G., Bentrop, D., Ejchart, A., Alber, M., Hage, A., Sterner, R. and Rösch, P. (1995) *Eur. J. Biochem.*, **229**, 658–668.

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

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