Molecular Model of the Solution Structure for the Paramagnetic Four-Iron Ferredoxin from the Hyperthermophilic Archaeon *Thermococcus litoralis*†

Peng-Liang Wang,‡ Antonio Donaire,‡,§ Zhi H. Zhou,|| Michael W. W. Adams,|| and Gerd N. La Mar*,‡

Department of Chemistry, University of California, Davis, California 95616, and Department of Biochemistry and Molecular Biology, Center for Metalloenzyme Studies, University of Georgia, Athens, Georgia 30602

Received April 2, 1996; Revised Manuscript Received July 3, 1996⊗

ABSTRACT: A molecular model for the three-dimensional solution structure of the paramagnetic, four-iron ferredoxin (Fd) from the hyperthermophilic archaeon *Thermococcus litoralis* (*Tl*) has been constructed on the basis of the reported ^{1}H NMR spectral parameters [Donaire, A. (1996) *J. Biomol. NMR 7*, 35−47]. The conventional use of long mixing time NOESY cross-peak intensity, backbone angles, and hydrogen-bonding constraints for building the structure was augmented by short mixing time NOESY, steady-state NOE, paramagnetic relaxation constraints, and the angular dependence of the ligated Cys H_β contact shifts. Distance geometry was used to generate various initial structures, and these structures were refined with the simulated annealing protocol. The family of structures with inconsequential violations exhibited low RMS deviations for the backbone except for a few residues in the immediate cluster vicinity and traces out a secondary structure very similar to those of the structurally characterized single cubane cluster Fds. The ability to describe the cluster environment depended on the use of numerous paramagnetic relaxation constraints which resulted in even the cluster loop residues exhibiting well-defined orientations, with the exception of one residue (Ile11) whose ^{1}H signals have not been located. Comparison of the structure of *Tl* Fd to those of mesophilic ferredoxins reveals that *Tl* Fd possesses the same secondary structural elements, two β-sheets, two helices, and four turns, with the exception that the β-sheet involving the termini incorporates a third strand in *Tl* Fd. Several minor structural adjustments in *Tl* Fd relative to other Fds, in addition to the third strand for β-sheet, include the incorporation of the termini into the β-sheet, a likely salt bridge from the side chain of the third β-strand to the N-terminus, and a more hydrophobic and compact interaction between the large β-sheet and the long helix. It is likely that each of these modifications, among others not yet well-defined (*i.e.*, surface salt bridges), contributes to the extraordinary thermostability of *Tl* Fd.

The so-called hyperthermophiles are a recently discovered group of microorganisms that have the remarkable property of growing optimally at 80 °C and above (Stetter et al., 1990; Blöchl et al., 1995). Virtually all of them are classified as archaea (or archaebacteria) rather than as bacteria (or eubacteria: Woese et al., 1990). Proteins obtained from these organisms are of current interest both for their biotechnological potential and for insights into the molecular mechanisms that enable them to function under such extreme conditions (Adams et al., 1995). However, although numerous hyperthermophilic proteins are now available in purified forms (Adams, 1993; Schönheit & Schäfer, 1995), there is a paucity of detailed structural information on them. At present, three-dimensional structures, determined by X-ray crystallographic or solution NMR studies, have been reported for the rubredoxin (Day et al., 1992; Blake et al., 1992a), aldehyde ferredoxin oxidoreductase (Chan et al., 1992), and glutamate dehydrogenase (Yip et al., 1995) from the archaeon *Pyrococcus furiosus* (optimal temperature for growth, T_{opt} = 100 °C), for DNA-binding proteins (Edmonson et al., 1995; Starich et al., 1996; Baumann et al., 1995) from the archaea *Sulfolobus* sp. (T_{opt} = ∼80 °C) and *Methanothermus fervidus* (T_{opt} = 83 °C), and for glyceraldehyde-3-phosphate dehydrogenase (Korndorfer et al., 1995) from the bacterium *T. maritima* (T_{opt} 80 °C). The main conclusions from such analyses are that proteins from hyperthermophiles do not have any "new" interactions and do not possess structural motifs significantly different from those found in their mesophilic counterparts. Instead, it appears as if significantly enhanced thermostability is achieved by minor changes in secondary and tertiary contacts that appear to be widely distributed over the protein. Thus, protein hyperthermostability has been variously attributed, at least in part, to the presence of additional salt bridges and hydrogen bonds, increased volume to surface ratios and increased compactness relative to mesophilic counterparts, and it is more a global

† This research was supported by grants from the National Science Foundation, DMB91-04018, MCB96-00759 (G.N.L.), and MCB94-05783 (M.W.W.A.), and from the National Institutes of Health GM45597 (M.W.W.A.).

* Address correspondence to this author.

‡ University of California.

§ Present address: University of Florence, via Gino Capponi 7, I-50121, Firenze, Italy.

|| University of Georgia.

⊗ Abstract published in *Advance ACS Abstracts,* August 15, 1996.

[1] Abbreviations: *Pf*, *Pyrococcus furiosus*; *Tl*, *Thermococcus litoralis*; *Tm*, *Thermatoga maritima*; *Dg*, *Desulfovibrio gigas*; *Da*, *Desulfovibrio africanus*; *Bt*, *Bacillus thermoproteolyticus*; NOESY, two-dimensional nuclear Overhauser spectroscopy; TOCSY, two-dimensional total correlation spectroscopy; MD, molecular dynamics; SA, simulated annealing; RMD, restrained molecular dynamics; Hipip, high-potential iron−sulfur protein; Fd, ferredoxin; Rd, rubredoxin; RMS, root mean square.

S0006-2960(96)00783-0 CCC: $12.00

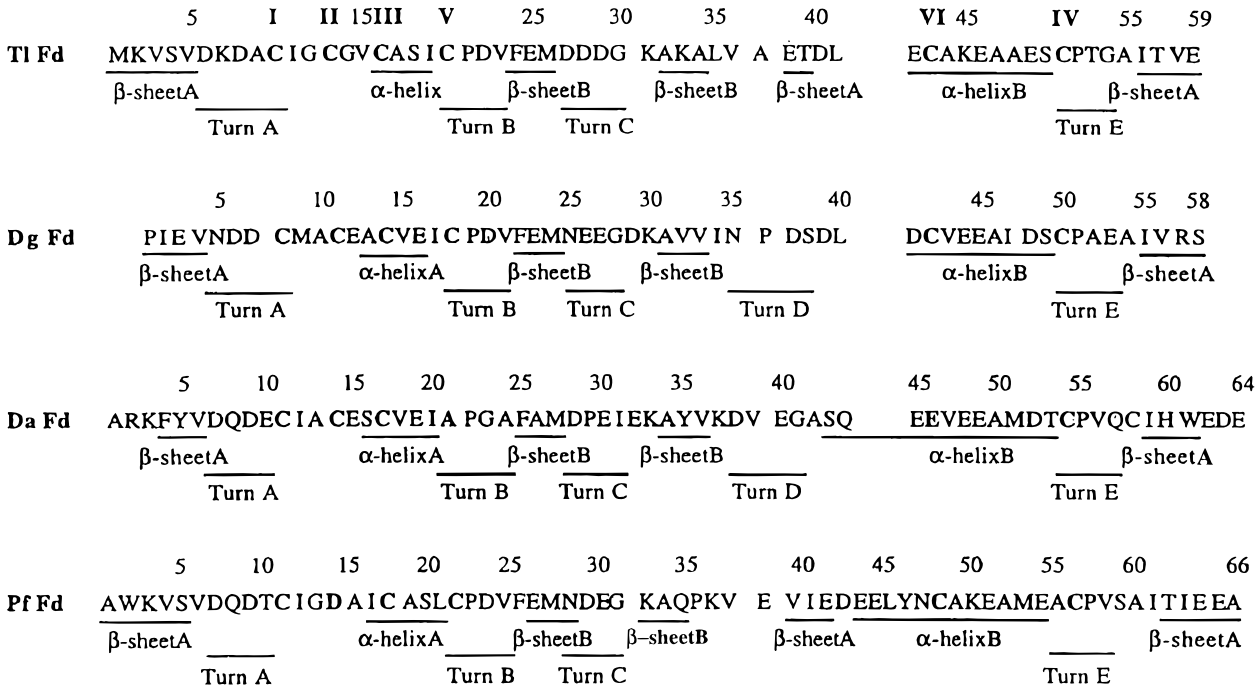

FIGURE 1: Comparison of the amino acid sequences of *Tl* Fd with those of the mesophile *Dg* and *Da* Fds and another hyperthermophile, *Pf* Fd. The alignments are made on the basis of the four cluster ligands (Cys also labeled CysI−IV, except for *Pf*, which has Asp at position II), Cys involved in a disulfide bridge (CysV and CysVI), and structural homology. The positions of secondary structural elements are indicated below the sequences.

rather than a local structural phenomena. This complicates analyses of individual interactions, and a dissection of contributions to "hyperthermostability" is obviously more readily achieved with small rather than large proteins.

As model hyperthermophilic proteins we are using two small iron−sulfur electron transfer proteins, 4Fe ferredoxin (Fd)[1] and rubredoxin (Rd), from the archaea *Pyrococcus furiosus* (*Pf*) and *Thermococcus litoralis* (*Tl*; T_{opt} 88 °C). Rds (∼50 residues) contain a single iron ligated by four cysteinates, whereas 4Fe Fds (∼60 residues) contain a single cubane $[Fe_4S_4]^{+1,+2}$ cluster ligated by four cysteinates, *i.e.*, *Tl* Fd, or three cysteinates and one aspartate, *i.e.*, *Pf* Fd (Blake et al., 1991, Busse et al., 1992; Calzolai et al., 1995). These hyperthermophilic proteins serve as ideal examples for structural studies as a large number of sequences and several crystal structures are available for both protein types from various mesophilic organisms. So far the crystal structure of native *Pf* Rd (Blake et al., 1992a), the solution NMR structure of Zn-substituted *Pf* Rd (Day et al., 1992), and the solution NMR secondary structures of native *Pf* 3Fe Fd (Teng et al., 1994) and *Tl* 4Fe Fd (Donaire et al., 1996) have been reported. Even for these two very small proteins, there appear variations in the degree to which the structures differ between the hyperthermophilic and the mesophilic forms. For example, some secondary structural elements of *Pf* Fd are significantly extended or shifted relative to those of mesophilic Fds (Teng et al., 1994), whereas the tertiary structure of *Pf* Rd is virtually superimposable on its mesophilic counterparts (Blake et al., 1992a,b; Day et al., 1992). This suggests that the nature of the interactions that increase stability may not be universal but depend on the particular folding topology of a given class of proteins.

Definitive structural studies of a hyperthermostable Fd have been hampered both by the instability of crystals to X-ray irradiation and by the intrinsic cluster paramagnetism which complicates data collection and analysis relative to diamagnetic systems in solution NMR studies. The broad and effective relaxed protons in the cluster vicinity are difficult to detect and assign by 2D NMR, and the resulting diminution or loss of NOESY cross-peak intensity near the cluster obliterates structural constraints needed to build a robust molecular model (Teng et al., 1994). However, appropriate tailoring of 1D/2D NMR methods to strongly relaxed resonances has led to the identification of all but one residue from *Tl* 4Fe Fd (Donaire et al., 1996). These preliminary 2D NMR results on *Tl* 4Fe Fd (Donaire et al., 1996) have identified secondary structural elements and some tertiary contacts that demonstrate structural homology to the three crystalographically characterized single clusters from mesophilic Fds, *Desulfovibrio gigas* (*Dg*; Kissinger et al., 1991), *Desulfovibrio africanus* (*Da*; Sery et al., 1994), and *Bacillus thermoproteolyticus* (*Bt*; Fukuyama et al., 1988); the sequences of *Da* and *Dg* Fd, together with that of *Tl* Fd, are aligned in Figure 1.

In this report we present a model for the three-dimensional solution molecular structure of oxidized *Tl* 4Fe Fd. This is based on distance geometry and restrained molecular dynamics using conventional long mixing time NOESY cross-peak intensity, backbone bond angles for protons remote from the cluster, and hydrogen bonding constraints from previously identified slowly exchanging labile protons (Donaire et al., 1996). In the vicinity of the cluster, short mixing time NOESY cross-peak intensity, steady-state NOEs, the ligated Cys $C_\beta H$ contact shift dependence on χ_2 (Noodleman et al., 1995; Bertini et al., 1995), and cluster-induced relaxation (Donaire et al., 1996) provide key structural constraints. The resulting structure is well-defined even in the vicinity of the cluster and confirms strong structural homology to mesophilic single-cluster Fds but also reveals some structural alternations that may contribute to the hyperthermostability of this protein.

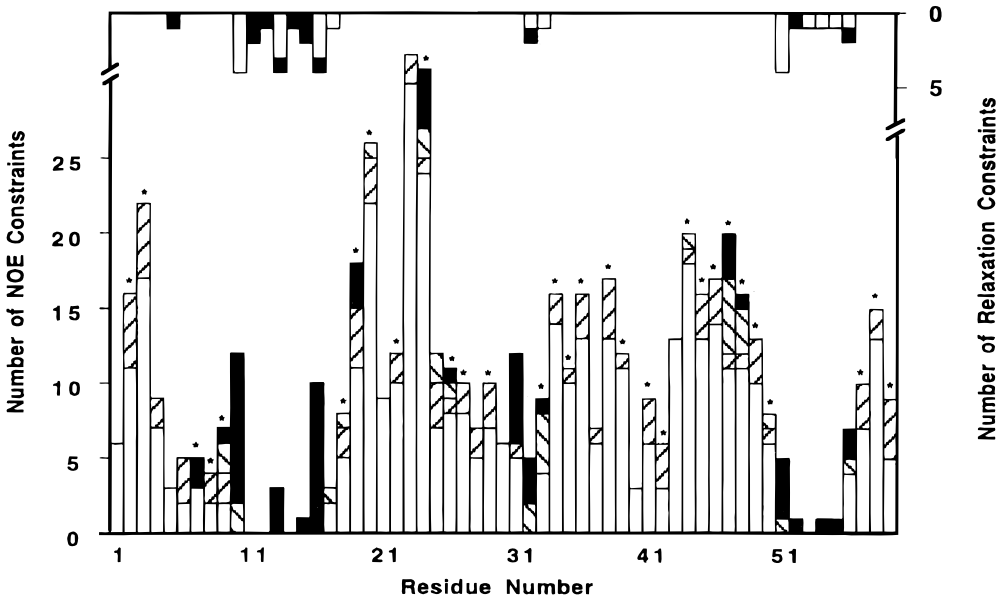

FIGURE 2: Schematic representation of the structural constraints used to generate the molecular model of 4Fe *Tl* Fd. On the lower portion are shown the number of NOE constraints (left margin) from inter- (□) and intra-residue (///) long (300 ms) mixing time NOESY, short mixing time (50 ms) NOESY (\\\), and steady-state NOEs (■). The residues with estimates for $^3J(NHC_\alpha H)$ are shown by asterisks. In the upper portion are shown the number of relaxation constraints (right margin) with both upper and lower R_{Fe} bounds (□) and with only upper bounds to R_{Fe} (■).

MATERIALS AND METHODS

Source of Protein and NMR Spectroscopy. Growth of *T. litoralis* (DSM 5473), purification of the 4Fe Fd, NMR analyses, and the spectral parameters from 2D NMR data have all been described in detail previously (Donaire et al., 1994, 1996).

Distance Constraints from NOEs. The majority of the distance constraints were obtained from 300 ms mixing time NOESY spectra at 30 and 40 °C which reflect interproton distances remote from the cluster (Donaire et al., 1996). Cross-peak intensities were calibrated as strong, medium, and weak to correspond to interproton distance ranges of <2.5, 2.5−3.5, and 3.5−5.0 Å, respectively. In cases where stereospecific assignments were not available, distances were adjusted accordingly to the pseudoatom position. 295 constraints [85 intra-residue, 112 sequential, 38 intermediate-range (i+2, i+4), and 60 long-range] are obtained from the 300 ms map. The proximity to the cluster of numerous protons, as well as their resultant enhanced relaxation, will suppress NOESY peak development involving such protons. Hence, the distance constraints for cross-peaks involving protons which the resulting molecular structure placed <6 Å from an iron were recalibrated interactively by comparing their cross-peak intensity in a $\tau_m = 50$ ms NOESY map. Eleven such cross-peaks were located in the 50 ms NOESY map and compared to the intensity of a strong cross-peak involving protons remote from the cluster (see supporting information). Lastly, six cross-peaks were observed solely in the 50 ms NOESY map, and their intensities were calibrated with the same reference peak in this map.

In the immediate vicinity of the cluster, in particular for spatial constraints to ligated residues, distance estimates were drawn from steady-state NOEs (two intra-residue, three sequential, eight medium-range, and 17 long-range) for which distance estimates were obtained from the magnitude of the NOEs and the estimates of the T_1 for the detected NOE as given by eq 1 (see below).

$$\eta_{ij} = \sigma_{ij} T_{ij} \tag{1}$$

The distribution of the various constraints are summarized at the bottom of Figure 2. The individual constraints are listed in supporting information.

Paramagnetic Constraints. Protons near the cluster with $R_{Fe} > 4.5$ Å, $T_1 > 10$ ms, are relaxed by the cluster paramagnetism according to the relation (Banci et al., 1991; Gorst et al., 1995a; Donaire et al., 1996)

$$T_{1i}^{-1} = D \sum_{q=1}^{4} R_{Fe_q - i}^{-6} \tag{2}$$

where $R_{Fe_q - i}$ is the distance of proton i from Fe_q in the 4Fe cluster. Since the $[Fe_4S_4]^{+2}$ cluster possesses four magnetically equivalent $Fe^{+2.5}$ (Middleton et al., 1978), D may be assumed as a constant (Donaire et al., 1996). The value of $D = 5 \times 10^5$ Å^6/s in eq 2 is estimated from correlation of observed T_1 values of *Tl* 4Fe Fd with distance obtained from crystal structure of three homologous Fd (Donaire et al., 1996) and adjusted as needed on the basis of the structure deduced in this report. R_{Fe-i} estimated from T_1 values were allowed ranges of ±10% (corresponding to ±60% uncertainty in D or due to differences in spin density among the iron). For protons with $R_{Fe-i} < 4.5$ Å, even minor changes in geometry near the cluster lead to very large changes in R_{Fe-i}. Hence all protons with $T_1 < 10$ ms were simply restricted to $R_{Fe} < 4.5$ Å. Backbone protons which could not be detected even by short mixing time NOESY (50 ms) or TOCSY (10 ms) spectra were assigned $T_1 < 10$ ms and $R_{Fe} < 4$ Å on the basis of detection of TOCSY cross-peaks for Ile54 $C_\delta H_3$-$C_\gamma H$ for which $C_\delta H_3$ exhibits $T_1 = 10$ ms (Donaire et al., 1994), and $C_\gamma H$ is predicted to be much more weakly relaxed. The relaxation analysis resulted in 18 constraints with both upper and lower bounds to R_{Fe} and 14 constraints with only upper bounds to R_{Fe}. The distribution of relaxation constraints among the residues in *Tl* Fd is

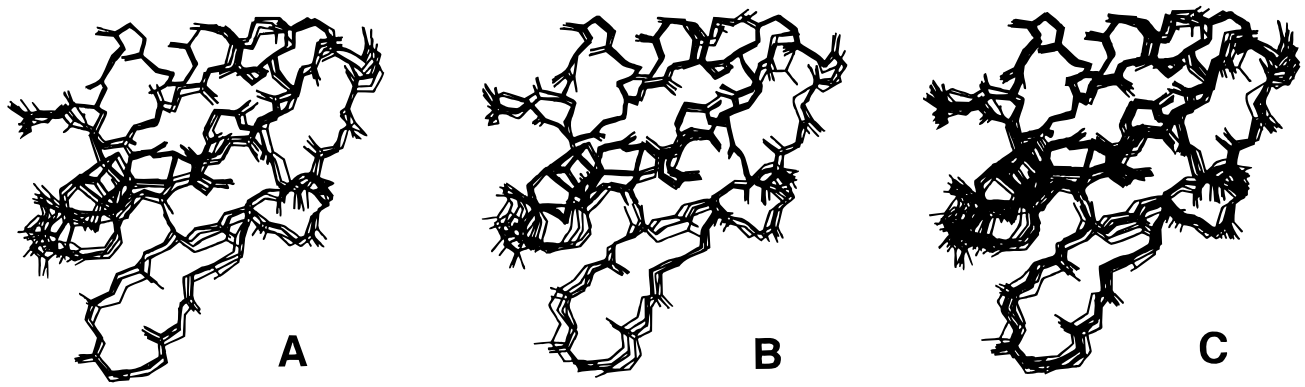

FIGURE 3: Superposition of backbones for accepted structures of *Tl* 4Fe Fd from the structural calculations. (A) Five structures obtained with Cys χ_2 as constraints, (B) five structures obtained without using Cys χ_2 as constraints, and (C) superposition of the two sets of five structures for A and B above.

included at the top of Figure 2; the individual constraints are listed in supporting information.

The Fe−S−C_β−H angle, θ (related to χ_2), for the four ligated Cys C_βHs [previously assigned stereospecifically by the predicted and observed differential paramagnetic relaxation behavior (Busse et al., 1991; Davy et al., 1995; Donaire et al., 1994)] are related to the contact shift, $\delta_{con}(C_\beta H)$ via eq 3.

$$\delta_{con}(C_\beta H) = a \sin^2 \theta + b \cos \theta + c \qquad (3)$$

with constants $a = 11.5$, $b = -2.9$, $c = 3.7$ (in ppm) (Bertini et al., 1994; Noodleman et al., 1995). The uncertainty in θ (χ_2) value deduced from the reported $\delta_{con}(C_\beta H)$ was estimated as $\pm 10°$.

Other Constraints. A disulfide bond is formed between Cys20 and Cys43 on the basis of both sulfhydryl titration [see Gorst et al. (1995b)] and local NOESY cross-peak patterns. Hydrogen bond constraints were introduced for the very slowly exchanging peptide NH identified previously (Donaire et al., 1996) by examination of NOESY cross-peak pattern to indicate the likely acceptor oxygen; distance ranges involving these likely hydrogen-bound NH···O were set at 1.9−2.5 Å. Backbone $^3J(NH_\alpha)$ coupling constraint estimates were from the DQF-COSY and TOCSY data.

Structural Computations. Structural calculations were performed using the Biosym software package (San Diego, CA) operating on a Silicon Graphics Indigo workstation. The cluster geometry and charge parameters used are those for a $[Fe_4S_4]^{+2}$ cluster as reported by Mouesca et al. (1994). Distance geometry (DG) was used to generate a variety of random structures. The structures from DG were refined using the restrained simulated annealing (SA) protocol (DISCOVER 2.95), which substantially reduces the force constants at the beginning and then selectively scales them up until the full values are regained. During the calculations, all peptide bonds were forced to *trans* and C_α chirality to *L*, a nonbonded cutoff distance of 6 Å was used, and the charges and cross terms were not included. The structures from SA calculations were further refined by restrained molecular dynamics (RMD) at 300 K for 10 ps and restrained energy minimization with low force constant constraints; during this period charges were included and distance-dependent dielectric was used. PROCHECK (Laskowski et al., 1993) was employed to check the geometry of the refined structures. Only the structures with reasonable geometry and low constraint violations were retained. The restrained simulated annealing protocol consisted of seven stages. These included initially raising the temperature to 1000 K for 12 ps in three stages during which the NOE force constants were scaled from 0.001 to 20 kcal. Subsequently, the system was cooled to 300 K in two stages. Finally, energy minimization was performed in two steps (100 steps steepest descent, 2000 steps conjugate gradient).

The procedures described above were repeated numerous times. Initially we focused on the local constraints and built the structure of each segment separately. The constraints which always resulted in geometry violations were found to be incorrect and were corrected by checking the original experimental data. Some side chain chiralities were also assigned during this process. Subsequently, the structures of segments were built again with the updated constraints and the constraints were re-evaluated. After local refinement, the whole molecular structure was constructed. Errors in the long-range constraints were identified and corrected in the same way as for the local constraints, and, on the basis of the whole structure, some pseudoatoms were assigned to specific atoms. With the refined constraints, two sets of 50 initial structures each were generated from distance geometry and all of them were refined subsequently by simulated annealing. The two sets differed only in that the $\delta_{con}(C_\beta H)$ constraint in eq 3 was omitted from one set but was included in the other set of 50 initial structures. The residue properties of the refined structures were checked by the PROCHECK program, structures with disallowed geometry were eliminated, and the remaining structures further refined. Finally, a total of ten structures resulted, five from each set of initial 50 structures, which are not only geometry reasonable but also exhibit minimal restraint violations.

RESULTS

Quality of the Structure. The superpositions of the backbones for the five acceptable structures of *Tl* 4Fe-Fd obtained with and without the constraints in eq 3 are shown in Figure 3A and B, respectively. The superposition of all ten acceptable structures is shown in Figure 3C. A stereoview of a ribbon representation of one of these structures is shown in Figure 4A. The pairwise RMS deviations for the backbone and side chain heavy atoms for two sets of five structures shown in Figures 3A and B were found to be similar, so the RMS deviations shown in Figure 5 are based on the total accepted ten structures for the main chain (Figure

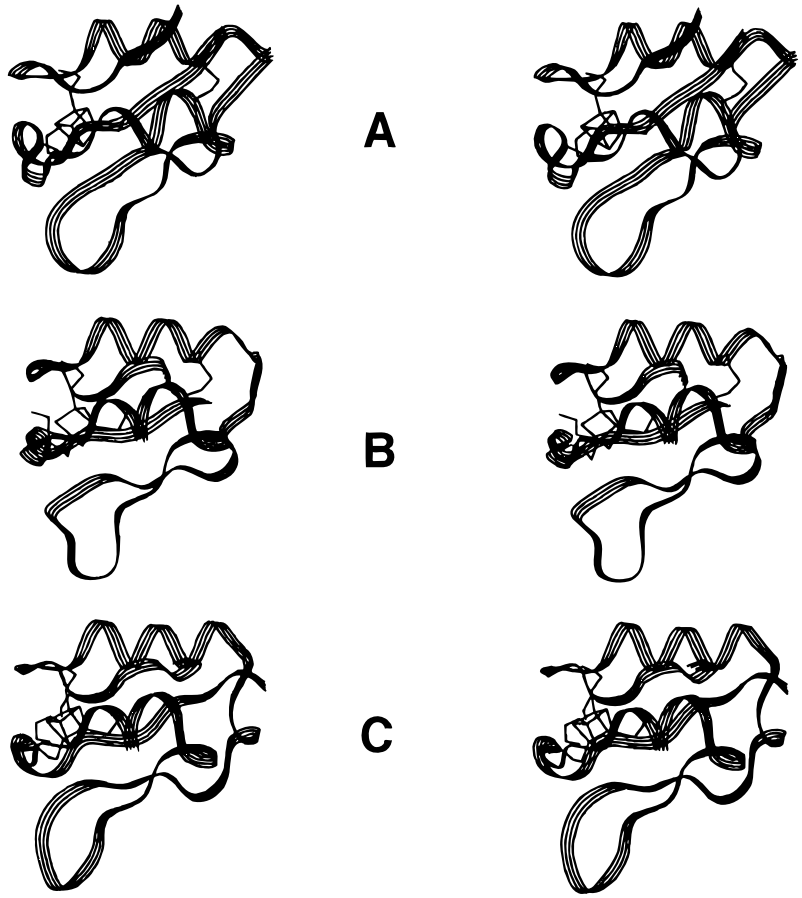

FIGURE 4: Stereoviews of the ribbon representations of the backbone for (A) one of the accepted NMR structures of *Tl* 4Fe Fd and the crystallographical structures of (B) 3Fe *Dg* Fd (Kissinger et al., 1991) and (C) 4Fe *Da* Fd (Sery et al., 1994).

5A) and side chain heavy atoms (Figure 5B), respectively. The only detectable difference in the two sets of five structures determined with, and without, eq 3 constraints, appeared in its orientation of Cys13 (see below). All secondary structural elements and the 4Fe cluster are well-formed in each of the ten structures and establish the presence of the following: two helices (helix-A, residues 16−19; helix-B, residues 42−50), two β-sheets (triple-stranded β-sheet A, residues 1−5, 56−59, and 38−39; and double-stranded β-sheet B, residues 24−26 and 32−34), and four Asx type-turns A−C and E (a double type I turn A, residues 6−9/7−10, and three single turns B, C, and E, residues 20−23, 27−30, and 51−54, respectively). The positions of these structural elements are given in Figure 1. At present, there are two structural elements for which only minimal structural constraints exist: the cluster loop Cys10-Cys16 (Figure 6A) and the turn E following Cys51 (Figure 6B).

The chart of the RMS deviations for the backbone for all ten structures (Figure 5A) reveals less than optimally defined regions for residues near the cluster ligand (residues 11−14 and 52−54). No signals could be located for Ile11, and neither Thr53 backbone proton exhibited detectable NOESY cross-peaks (Donaire et al., 1996). However, in spite of the paucity of data, the secondary structural elements in which these residues reside, the cluster binding loop (Figure 6A) and Asx turn E (Figure 6B), respectively, are reasonably well-formed. The other region of the protein where larger backbone RMS deviations are observed involves β-sheet B residues 29−31. While this β-sheet is well-formed and the superposition of the β-sheet residues by themselves is excellent, the proximity of the cluster leads to the detection of fewer than expected tertiary NOESY cross-peaks between the cluster and β-sheet B, and hence there is a less than optimal definition of the interface between them. The family of the five structures based on the use of eq 3 as a constraint leads to a good correlation between the reported $\delta_{con}(C_\beta H)$ and computed θ values, as shown in Figure 7 (open marker). Omitting the χ_2 constraint leads to a very similar plot except for Cys13 as shown in Figure 7 (closed marker). All ten accepted structures produced very reasonable correlations for eq 2, with $D = (4.0 \pm 0.5) \times 10^5$ Å^6/s (not shown, see supporting information). The resultant structures, moreover, confirm the stereospecific assignment of ligated Cys by the observation of reasonable correlation for both T_1 (not shown; see supporting information) and $\delta_{con}(C_\beta H)$ (Figure 2) in eqs 2 and 3, respectively.

The side chain heavy atom RMS deviations per residue are shown in Figure 5B; generally well-defined orientations occur for residues with side chains directed toward the protein interior. The majority of the side chains with large RMS deviations are Lys and Glu on the protein surface. For the most part, this results from the side chain termini exhibiting no detectable NOESY cross-peaks (*i.e.*, Lys7, Glu46, Asp27), or the peaks overlap other proton signals at all accessible temperatures (Lys2, Glu49). Lys33 and Glu25 exhibit weak NOESY cross-peaks to Leu35 that suggest a salt bridge between these residues. The inspection of the positions of the surface Lys and Glu in the *Tl* Fd molecular model (not shown; see supporting information) indicates the salt bridges can form between side chains of Lys2 and Glu38,

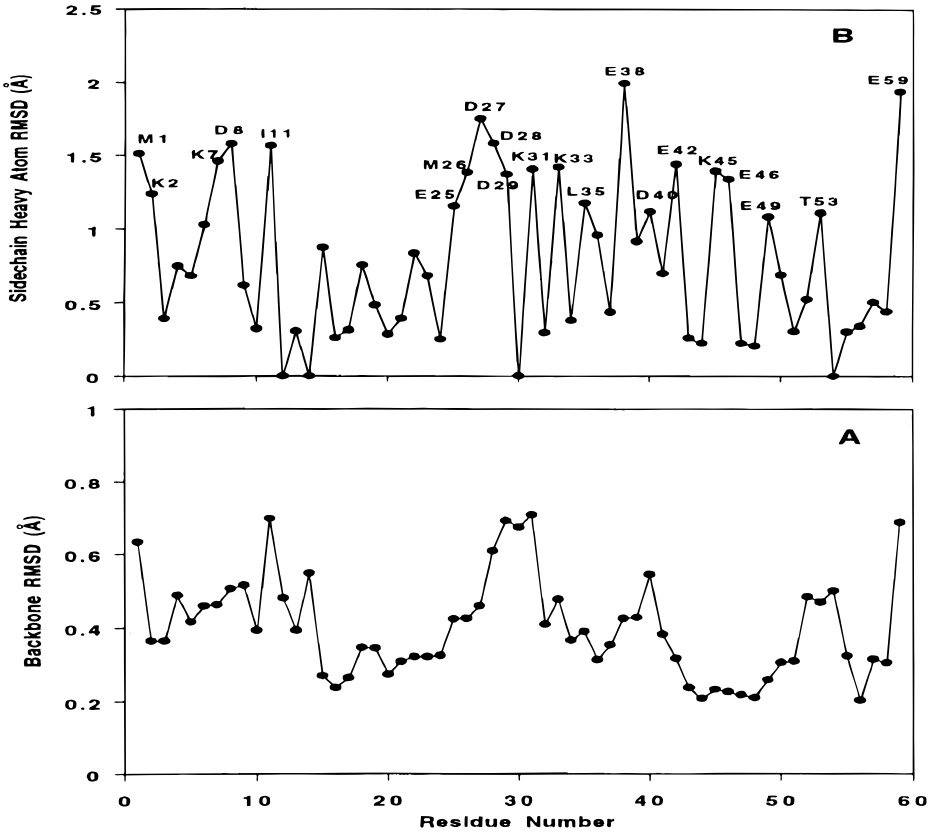

FIGURE 5: Heavy atom pairwise RMS deviations per residue for the ten accepted NMR structures of 4Fe *Tl* Fd for (A) backbone and (B) side chains.

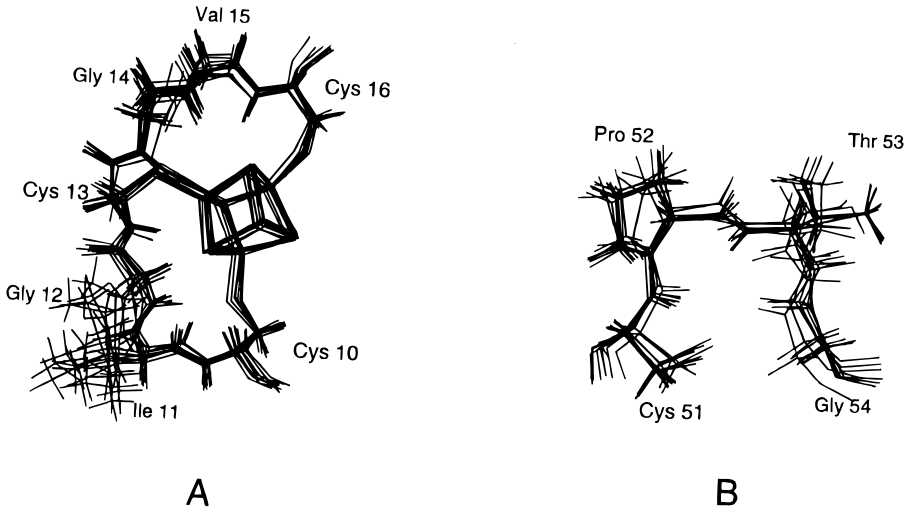

FIGURE 6: Superimposition of the ten accepted structures for portions of the structure with minimal constraints. (A) Cluster binding loop, Cys10-Cys16, and (B) turn E, Cys51-Gly54.

Lys31 and Asp27, and Lys45 and Glu59; the Glu38 side chain, moreover, can participate in salt bridge to both Lys2 and the N-amino terminus. However, the NOESY cross-peaks to confirm these interactions are not detected at this time.

Comparison with Other Single Cubane Cluster Fd. As noted even in the preliminary analysis of NOESY cross-peak patterns alone (Donaire et al., 1996), both the secondary structure and the tertiary interaction are remarkably conserved in comparison to those of crystallographically determined single cubane cluster Fd (Fukuyama et al., 1988; Kissinger et al., 1991; Sery et al., 1994), as illustrated by the comparison in Figure 4 of the ribbon structure of 4Fe *Tl* Fd to that of 3Fe *Dg* Fd and 4Fe *Da* Fd. *Bt* Fd is significantly longer than the other single cubane cluster Fd, with a loop inserted between turn D and helix B and an extension of the C-terminus (Fukuyama et al., 1988), and hence does not provide as direct a comparison as do the *Dg* and *Da* Fds. The *Tl*, *Dg*, and *Da* Fds possess the same two helices, two antiparallel β-sheets, and four turns (A, B, C, E) involving similar residues (Figure 1). The single and most prominent feature that differentiates the secondary structure of *Tl* Fd from that of the other Fds [but not that of hyperthermophilic *Pf* Fd (Teng et al., 1994); see Discussion] is that a turn (turn D) that makes a single backbone contact to the β-sheet involving the two termini in other structurally characterized

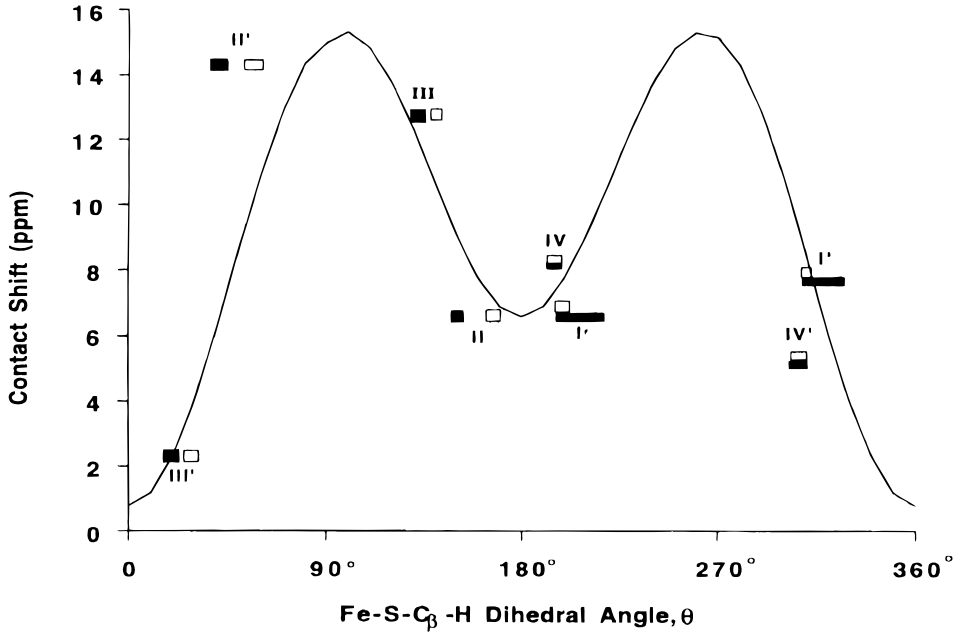

FIGURE 7: Plot of the observed contact shifts for ligated Cys C_βHs as a function of the Fe−S−C_β−H dihedral angle, θ, for the five accepted structures using eq 3 as a constraint (open marker) and the five accepted structures excluding eq 3 as a constraint (closed marker). The horizontal dimension of these boxes represent the range of the angles obtained for various members of the set of five structures. The data points are labeled by Cys in the sequence in Figure 1, *i.e.*, I → IV, with C_βH closer to iron indicated by the prime.

Fd (Fukuyama et al., 1988; Kissinger et al., 1991; Sery et al., 1994) is replaced by a short segment that forms a third antiparallel strand (beginning residue 38) in β-sheet A (see Figure 1). This difference is readily recognized in the comparison of the ribbon structures of *Tl* Fd and the mesophilic Fds in Figure 4. Turn A in *Tl* Fd is shown here to be a composite of two turns involving residue 6−9 (type I Asx) and 7−10 (type I) involving the ligating Cys10; a similar double turn is found in *Da* Fd, while *Dg* Fd exhibits a simple Asx turn (Figure 4).

The tertiary contacts among the side chains on various secondary structural elements in *Tl* Fd are also very similar to those in the two mesophilic Fds, as shown in Figures 3 and 4. The cluster environment in *Tl* Fd is very similar to that in other Fd, with similar predicted hydrogen bonds to the ligated sulfur atoms. The proposed hydrogen bonds to the sulfur atoms of both the cluster and of ligated cysteines in *Tl* Fd are compared to those in *Da* and *Dg* Fd in Table 1.

Table 1. Hydrogen Bonds to the Bridging Sulfur-Ligated Cysteines in the Cluster

4Fe *Tl* Fd		3Fe *Dg* Fd		4Fe *Da* Fd	
acceptor	donor	acceptor	donor	acceptor	donor
S1[a]	Cys13 N				
S1	Gly14 N	S1	Glu12 N	S1	Glu15 N
S1	Ala32 N				
S2	Ile 11 N	S2	Met9 N	S2	Ile12 N
		S2	Cys11 N		
S4	Cys16 N	S4	Cys14 N	S4	Cys17 N
Cys10 S_γ	Gly12 N	Cys8 Sγ	Ala10 N	Cys11 S_γ	Ala13 N
Cys10 S_γ	Ala32 N	Cys8 Sγ	Ala31 N	Cys11 S_γ	Ala34 N
Cys13 S_γ	Val15 N			Cys14 S_γ	Ser16 N
Cys16 S_γ	Ala17 N			Cys17 S_γ	Val18 N
Cys51 S_γ	Gly54 N	Cys50 S_γ	Glu53 N		
		Cys50 S_γ	Ala52 N		
				Cys54 S_γ	Val56 N
Cys51 S_γ	Ala55 N	Cys50 S_γ	Ala54 N	Cys54 S_γ	Cys58 N

[a] Bridging sulfide labeled as described for 3Fe *Dg* Fd (Kissinger et al., 1991).

DISCUSSION

NMR Structural Approaches. The available NMR data provided sufficient distance constraints to allow the construction of a robust molecular model of *Tl* Fd in spite of its paramagnetism. The conventional ^{1}H−^{1}H NOESY data are sufficient to provide well-formed secondary structural elements much as in similar diamagnetic proteins. This is confirmed by the appearance of the helices and β-sheets in numerous refined structures which exhibited less-defined, and hence unacceptable, local structure near the cluster. Clearly, the number and nature of structural constraints near the cluster are the determining factors in obtaining an accurate molecular model. The paucity of NOE constraints for residues near the cluster is clearly shown in the lower half of Figure 2 and is obvious in the region of the loop including the three ligating Cys10-Cys16, the segment 51−55 near the ligating Cys51, and, to a lesser degree, for residues in β-sheet B near the cluster (Ala32, Lys33). The limited nature of the NOE data is, in small but crucial part, replaced by distance constraints to the cluster iron via T_1 data, as shown in the upper part of Figure 2. The results support an important role for T_1 constraints in model building of paramagnetic Fds, both for the plant-type 2Fe (Oh & Markley, 1990; Pochapsky et al., 1994; Chae et al., 1994; Chae & Markley, 1995; Cheng & Markley, 1995) and for cubane 4Fe/3Fe-type Fds (Teng et al., 1993; Gorst et al., 1995a; Davy et al., 1995; Bertini et al., 1995b; Donaire et al., 1996).

The inclusion of Cys χ_2 (θ) constraints via eq 3 led to significant changes in Cys orientation only for Cys13 for which the value of θ (Figure 7) is closer to the predicted curve than when the constraint is omitted. The larger influence of the constraints in eq 3 for Cys13 relative to the other three ligated Cys is due to the smaller number of constraints available for this residue (see Figure 2). The fit of C_βH contact shifts to eq 3 in Figure 7 for the present *Tl*

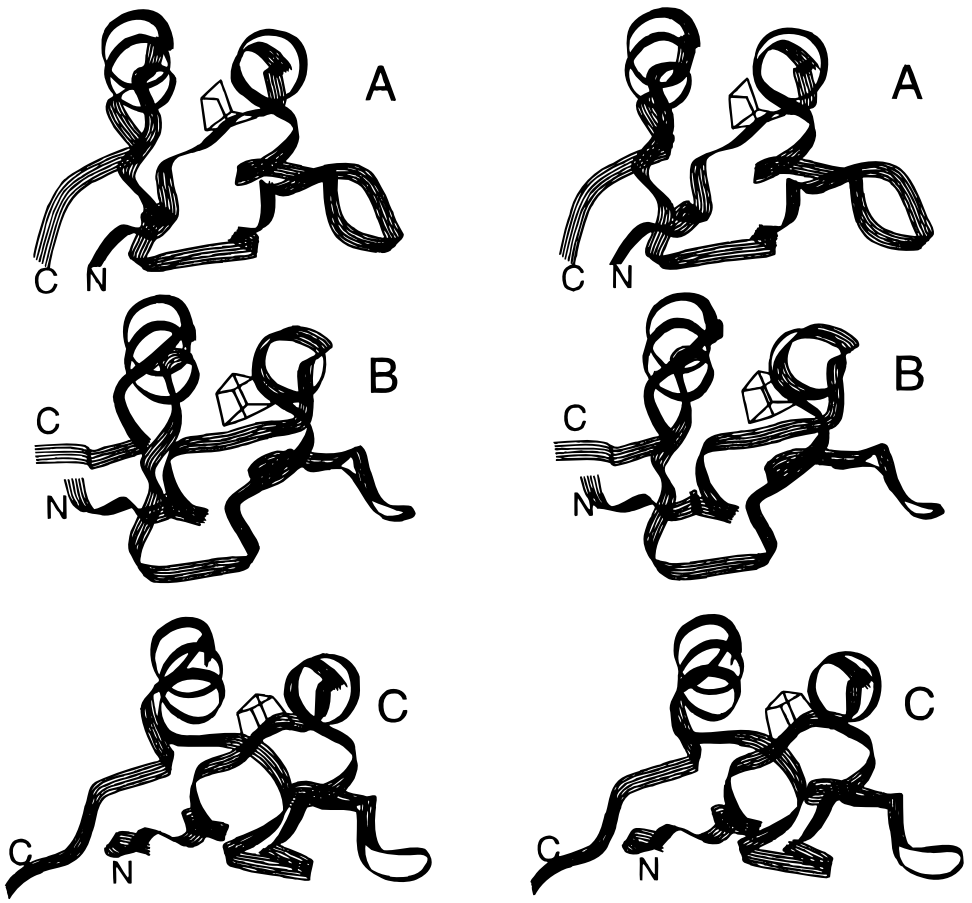

FIGURE 8: Stereoviews of ribbon structures allowing a comparison of the interactions between the N- and C-termini strands of the β-sheet A with α-helix-B in the mesophilic Fd from *Dg* (A) and *Da* (C) with the hyperthermophilic *Tl* Fd (B). For both *Dg* and *Da* Fd, the β-sheet is well-separated from the helix when compared to the position in *Tl* Fd.

Fd is as good as the fit of the experimental data points used to generate eq 3 (Bertini et al., 1994; Noodleman et al., 1995). The use of Cys χ_2 rather than relaxation constraints resulted in a reasonable model for a two cubane cluster Fd (Bertini et al., 1995). The observation of large contact shifts for H_αs for Cys13 and Cys16 ($S-C_\beta-C_\alpha-H_\alpha$ bond angle, $\sim$150−180°), and small contact shifts for Cys10 and Cys51 ($S-C_\beta-C_\alpha-H_\alpha$ bond angles, $\sim$65°) in *Tl* Fd indicates a Karplus-type $\cos^2 \chi_1$ dependence for the H_α contact shift similar to eq 3 (Busse et al., 1991; Zerbe et al., 1994; Bertini et al., 1994; Donaire et al., 1994; Noodleman et al., 1995) and suggests future use of Cys H_α contact shifts as structural constraints once the parameters correlating contact shift to bond dihedral angle have been calibrated on a series of structurally characterized Fds.

The problems of paramagnetism in determining molecular structure by NMR of (single) cubane cluster Fds are intermediate between those encountered in the plant-type 2Fe Fd (Oh & Markley, 1990; Pochapsky et al., 1994; Chae & Markley, 1995; Chae et al., 1995) and the 4Fe Hipips (Banci et al., 1995; Bertini et al., 1995a). In both classes of Fds, the clusters are on the surface and hence the inevitable paramagnetism obscures crucial links between the remaining secondary structural elements. The significantly stronger paramagnetism in 2Fe than 4Fe Fd obscures more resonances in the former protein (Oh & Markley, 1990; Cheng & Markley, 1995). While the cluster paramagnetism and molecular size of cubane Fds and Hipips are similar (50−60 residues), the cluster in Hipips is buried within the protein, which results in a large number of observable dipolar contacts between each of the cluster ligands and the "diamagnetic" protein matrix (Banci et al., 1995; Bertini et al., 1995a). The surface positions of three of the four cluster ligands in the cubane Fds result in the observation of many fewer dipolar contacts [and none for Cys13(II)] to the "diamagnetic" matrix and hence yield fewer structural constraints for the cluster environment. However, even the limited structural constraints near the cluster do not preclude the generation of reasonable molecular models for cubane cluster Fds (Bertini et al., 1995b). Further improvement of such models can be expected from ^{15}N labeling, since paramagnetic influences on ^{15}N are considerably less than on protons (Chae & Markley, 1995; Cheng & Markley, 1995; Scrofani et al., 1995). Additional ^1H relaxation and direct ^{15}N relaxation data should provide the constraints necessary to define the cluster environment to the same degree now possible for regions remote from the cluster. The labeling of the Lys amino terminus may also provide routes to locating the labile protons whose NOESY contacts may provide constraints to define the surface salt bridges.

Structural Implications for Thermostability. *Tl* Fd provides us with the first complete structure for a hyperthermophilic Fd yet is remarkably similar to that of crystallographically characterized mesophilic Fds, even though the latter exhibit substantially reduced thermal stability [see Busse et al. (1992)]. The most striking differences are that *Tl* Fd exhibits an expanded β-sheet (A) to include a third strand with an additional two backbone H-bonds compared to the mesophilic Fds and a likely side chain salt bridge (Glu38) from this third strand to the N-terminus. The latter

interaction may have the effect of stabilizing the β-strand "zipper" against opening, as was proposed for hyperthermophilic *Pf* Rd (Blake et al., 1992a,b). A triple-stranded β-sheet that incorporates the termini is present in *Pf* Fd (Teng et al., 1994; see Figure 7), but preliminary ^{1}H NMR data support only the two-stranded sheet for the 4Fe Fd from the hyperthermophile, *Thermatoga maritima* (*Tm*; Wildegger et al., 1995). Another element proposed to increase the stability of *Pf* Fd relative to mesophilic Fds was the extension (and shift toward the N-terminus) of helix B, thereby increasing the number of main chain hydrogen bonds (Teng et al., 1994). This helix in *Tl* Fd is virtually identical to that in *Dg* or *Pa* Fd, and the same is apparently the case for *Tm* Fd (Wildegger et al., 1995). The present data for *Tl* Fd provide little direct evidence for surface side chain salt bridges in *Tl* Fd, although one potential surface salt bridge that cannot occur in mesophilic Fds is that between Glu59 on β-sheet A and Lys45 on helix-B.

An additional difference between the *Tl* and the *Dg*, *Da* Fds involves the manner of the packing of the β-sheet A, which involves the termini, against the long helix-B, which is optimally observed in the stereoview shown in Figure 8. In the mesophilic Fds, β-sheet A is well-spaced from helix-B by the location of ionic (*i.e.*, Arg in *Dg* Fd) or very bulky (*i.e.*, Trp in *Da* Fd) side chains in the intervening space; hence the packing is very loose (see Figure 8 and C). This same β-sheet A/helix-B interface in *Tl* Fd is lined with aliphatic side chains, which allow much closer packing of the β-sheet against the helix (Figure 8B); this is strictly observed in each of the family of structures (see supporting information). The closer spacing of the helix and β-sheet presumably leads to a more compact structure that contributes to stability.

Quantitative thermostability data are lacking for all Fds, and, in the case of the hyperthermophiles (Aono et al, 1989; Busse et al, 1992; Blamey et al, 1974), the thermal denaturation is not reversible (Klump et al, 1994). Some differences in stability may be expected from considering the optimum growth temperatures for the hyperthermophilic organisms (*Pf*, 100 °C; *Tl*, 88 °C; *Tm*, 80 °C) that suggest stabilities in the same relative order. Apparent structural stabilizing elements present in *Pf* Fd (incorporation of termini into triple-stranded β-sheet A, long helix-B; Teng et al., 1994), *Tl* Fd (incorporation of termini into triple-stranded β-sheet A, normal helix-B), and *Tm* Fd (incorporation of termini into double-stranded β-sheet A, normal helix-B; Wildegger et al., 1995) correlate with optimum growth temperature. However, our current understanding of the structures of Fds does not point to any single feature that plays a defining role in protein hyperthermostability. Current efforts are directed to improving the present structure model for *Tl* Fd and developing an equally robust model for the solution structure *Pf* Fd. The latter is available in the recombinant form (Heltzel et al., 1996), which enables isotope labeling practical for more definitive NMR structure determination.

ACKNOWLEDGMENT

Valuable discussions with and the experimental assistance of Luigi Calzolai, Carol Gorst, and Quincy Teng are gratefully acknowledged.

SUPPORTING INFORMATION AVAILABLE

Five tables with the individual NMR constraints (300 and 50 ms NOESY; H-bond; NOE; T_1) used to generate the *Tl* Fd molecular model, and three figures (plot of T_1^{-1} vs $\sum^q R_{\mathrm{Fe}_q}^{-6}$, potential surface salt bridges in *Tl* Fd, and stereoviews of the interaction of β-sheet A/helix-B) (12 pages). Ordering information is given on any current masthead page.

REFERENCES

Adams, M. W. W. (1993) *Annu. Rev. Microbiol. 47*, 627−658.
Adams, M. W. W., Perler, F. B., & Kelly, R. M. (1995) *Bio/Technology 13*, 662−668.
Aono, S., Bryant, F. O., & Adams, M. W. W. (1989) *J. Bacteriol. 171*, 3433−3439.
Banci, L., Bertini, I., & Luchinat, C. (1991) *Nuclear and Electron Relaxation*, pp 143−156, VCH Publishers, Weinheim, FRG.
Banci, L., Bertini, I., Carloni, P., Luchinat, C., & Orioli, P. L. (1992) *J. Am. Chem. Soc. 114*, 10683−10689.
Banci, L., Bertini, I., Eltis, L. D., Felli, I. C., Kastrau, D. H. W., Luchinat, C., Piccioli, M., Pierattelli, R., & Smith, M. (1994) *Eur. J. Biochem. 225*, 715−725.
Banci, L., Bertini, I., Dikiy, A., Kastrau, D. H. W., Luchinat, C., & Sompornpisot, P. (1995) *Biochemistry 34*, 206−219.
Baumann, H., Knapp, S., Karshikoff, A., Ladenstein, R., & Härd, T. (1995) *J. Mol. Biol. 247*, 840−846.
Bertini, I., Capozzi, F., Luchinat, C., Piccioli, M., & Vila, A. J. (1994) *J. Am. Chem. Soc. 116*, 651−660.
Bertini, I., Dikiy, A., Kastrau, D. H. W., Luchinat, C., & Sampornpisut, P. (1995a) *Biochemistry 34*, 9851−9858.
Bertini, I., Donaire, A., Feinberg, B. A., Luchinat, C., Piccioli, M., & Yuan, H. (1995b) *Eur. J. Biochem. 232*, 192−205.
Blake, P. R., Park, J.-B., Zhou, Z. H., Hare, D. R., Adams, M. W. W., & Summers, M. F. (1992a) *Protein Science 1*, 1508−1521.
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.
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.
Blamey, J. M., Mukund, S., & Adams, M. W. W. (1994) *FEMS Microbiol. Lett. 121*, 165−170.
Blöchl, E., Burggraf, S., Fiala, F., Lauerer, G., Huber, G., Huber, R., Rachel, R., Segerer, A., Stetter, K. O., & Völkl, P. (1995) *World J. Microbiol. Biotechnol. 11*, 9−16.
Bruschi, M., Cambillau, C., Bovier-Lapierre, G., Bonicel, J., & Forget, P. (1986) *Biochim. Biophys. Acta 873*, 31−37.
Busse, S. C., La Mar, G. N., & Howard, J. B. (1991) *J. Biol. Chem.* 266, 23714−23723.
Busse, S. A., 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.
Calzolai, L., Gorst, C. M., Zhao, Z.-H., Teng, Q., Adams, M. W. W., & La Mar, G. N. (1995) *Biochemistry 34*, 11373−11384.
Chae, Y. K., & Markley, J. L. (1995) *Biochemistry 34*, 188−193.
Chae, Y. K., Abildgaard, F., Mooberry, E. S., & Markley, J. L. (1994) *Biochemistry 33*, 3287−3295.
Chan, M. K., Mukund, S., Kletzin, A., Adams, M. W. W., & Rees, D. C. (1995) *Science 267*, 1463−1469.
Cheng, H., & Markley, J. L. (1995) *Annu. Rev. Biophys. Biomol. Struct. 24*, 209−237.
Davy, S. L., Osborne, M. J., Breton, J. B., Moore, G. R., Thomson, A. J., Bertini, I., & Luchinat, C. (1995) *FEBS Lett. 363*, 199−204.
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. M., Zhou, Z. H., Adams, M. W. W., & La Mar, G. N. (1994) *J. Am. Chem. Soc. 116*, 8641−6849.
Donaire, A., Zhou, Z. H., Adams, M. W. W., & La Mar, G. N. (1996) *J. Biomol. NMR 7*, 35−47.
Edmondson, S. P., Qiu, L. S., & Shriver, J. W. (1995) *Biochemistry 34*, 13289−13304.

Fukuyama, K, Nagahara, Y., Tsukihara, T., & Katsube, Y. (1988) *J. Mol. Biol. 199*, 183−193.
Gorst, C. M., Yeh, T., Teng, Q., Calzolai, L., Zhou, J. H., Adams, M. W. W., & 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., & La Mar, G. (1995b) *Biochemistry 34*, 8788−8795.
Heltzel, A., Smith, E. T., Zhou, Z. H., Blarney, J. M., & Adams, M. W. W. (1996) *J. Bacteriol. 176*, 4790−4793.
Kissinger, C. R., Sieker, L. C., Adman, E. T., & Jensen, L. H. (1991) *J. Mol. Biol. 219*, 693−715.
Klump, H. H., Adams, M. W. W., & Robb, F. T. (1994) *Pure Appl. Chem. 66*, 485−489.
Korndorfer, I., Steipe, B., Huber, R., Tomschy, A., & Jaenicke, R. (1995) *J. Mol. Biol. 246*, 511−521.
Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
Middleton, P., Dickson, D. P. E., Johnson, C. E., & Rush, J. D. (1978) *Eur. J. Biochem. 88*, 135−141.
Mouesca, J.-M, Chen, J. L., Noodleman, L., Bashford, D., & Case, D. A. (1994) *J. Am. Chem. Soc. 116*, 11898−11914.
Noodleman, L., Chen, J.-L., Case, D. A., Giori, C., Rius, R., Mouesca, J.-M., & Lemotte, B. (1995) in *Nuclear Magnetic Resonance of Paramagnetic Macromolecules* (La Mar, G. N., Ed.) pp 339−367, Kluwer Academic Press, Dordrecht, NL.
Oh, B.-H., & Markley, J. L. (1990) *Biochemistry 29*, 4012−4017.
Pochapsky, T. C., Ye, X. M., Ratnaswamy, G., & Lyons, T. A. (1994) *Biochemistry 33*, 6424−6432.
Schönheit, P., & Schäfer, T. (1995) *World J. Microbiol. Biotechnol. 11*, 26−57.
Scrofani, S. D., Brownlee, R. T. C., Sadek, M., & Wedd, A. G. (1995) *Inorg. Chem. 34*, 3942−3952.
Sery, A., Housset, D., Serre, L., Bonicel, J., Hatchikian, C., Frey, M., & Roth, M. (1994) *Biochemistry 33*, 15408−15417.
Starich, M. R., Sandman, K., Reeve, J. N., & Summers, M. F. (1996) *J. Mol. Biol. 255*, 187−203.
Stetter, K. O., Fiala, G., Huber, G., Huber, R., & Segerer, A. (1990) *FEMS Microbiol. Rev. 75*, 117−124.
Teng, Q., Zhou, Z. H., Smith, E. T., Busse, S. C., Howard, J. B., Adams, M. W. W., & La Mar, G. N. (1994) *Biochemistry 33*, 6316−6326.
Wildegger, G., Bentrop, D., Ejchart, A., Alber, M., Hage, A., Sterner, R., & Rösch, P. (1995) *Eur. J. Biochem. 229*, 658−668.
Woese, C. R., Kandler, O., & Wheelis, M. L. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4576−4579.
Yip, K. S. P., Stillman, T. J., Britton, K. L., Artymiuk, P. J., Baker, P. J., Sedelnikova, S. E., Engel, P. C., Pasquo, A., Chiaraluce, R., Consalvi, V., Scandurra, R., & Rice, D. W. (1995) *Structure 3*, 1147−1158.
Zerbe, O., Poutney, D. L., von Philipsborn, W., & Vasak, M. (1994) *J. Am. Chem. Soc. 16*, 377−378.

BI960783U