High-Resolution Structure of the Soluble, Respiratory-Type Rieske Protein from *Thermus thermophilus*: Analysis and Comparison[†]

Laura M. Hunsicker-Wang,[‡,§] Andreas Heine,[§,||] Ying Chen,[‡] Eugene P. Luna,[‡] Thomas Todaro,[‡,⊥] Yan Ming Zhang,[§,#] Pamela A. Williams,[§,▽] Duncan E. McRee,[§,○] Judy Hirst,[@] C. David Stout,*[,§] and James A. Fee*[,‡,§]

Division of Biology, the University of California at San Diego, 9500 Gilman Drive, La Jolla, California 92093, Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, and Medical Research Council Dunn Human Nutrition Unit, Wellcome Trust/MRC Building, Hills Road, Cambridge CB2 2XY, United Kingdom

Received February 17, 2003; Revised Manuscript Received April 23, 2003

ABSTRACT: The structure of the soluble Rieske protein from *Thermus thermophilus* has been determined at a resolution of 1.3 Å at pH 8.5 using multiwavelength anomalous dispersion (MAD) techniques. This is the first report of a Rieske protein from a menaquinone-utilizing organism. The structure shows an overall fold similar to previously reported Rieske proteins. A novel feature of this crystal form appears to be a shared hydrogen between the His-134 imidazole ring ligated to Fe2 of the [2Fe−2S] cluster and its symmetry partner, His-134′, one being formally an imidazolate anion, Fe2-(His-134)Nϵ^{-}···H-Nϵ'-(His-134′)-Fe2′, in which crystallographic C_2 axes pass equidistant between Nϵ···Nϵ' and normal to the line defined by Nϵ···Nϵ'. This provides evidence for a stable, oxidized cluster with a His^{-} ligand and lends support to a previously proposed mechanism of coupled proton and electron transfer. A detailed comparison of the *Thermus* Rieske protein with six other Rieske and Rieske-type proteins indicates: (a) The cluster binding domain is tightly conserved. (b) The 3-D structure of the 10 β-strand fold is conserved, even among the most divergent proteins. (c) There is an approximately linear relation between acid-pH redox potential and number of H-bonds to the cluster. (d) These proteins have two faces, one points into the larger complex (bc_1, b_6f, or other), is involved in the proton coupled electron transfer function, and is highly conserved. The second is oriented toward the solvent and shows wide variation in charge, sequence, length, hydrophobicity, and secondary elements in the loops that connect the β-sheets.

The Rieske protein is an electron transport protein that contains a [2Fe−2S] rhombus ligated by two cysteine thiolates (to Fe1) and by the Nδ atoms of two histidine imidazole rings (to Fe2). Rieske and Rieske-type[1] proteins play essential roles in several biological processes including respiration, photosynthesis, biodegradation of aromatic compounds, and detoxification (*1*, *2*). Here, we report on a high-resolution structure of the water-soluble, Rieske protein fragment encoded by a gene from the menaquinone-utilizing, thermophilic-eubacterium, *Thermus thermophilus* (*Tt*)[1] that has been expressed in *Escherichia coli*.

In mitochondria and many bacteria, electrons move from NADH, succinate, or other electron donors to O_2 through four respiratory complexes: NADH dehydrogenase, succinic dehydrogenase, the central bc_1 complex, and a terminal oxidase (*3*). The O_2 is reduced to water, and the free energy released from this process is conserved, in turn driving ATP synthesis (*3*). The central complex in the respiration process contains cytochromes *b* and c_1 or *f*, along with the Rieske iron−sulfur protein. It performs a two electron oxidation of quinol via a bifurcated reaction. One electron is transferred to cytochrome *c* through the high potential chain, the Rieske protein and cytochrome c_1, while the second electron is transferred to heme b_L, a low potential heme, and from there to heme b_H and cycled back to the quinone pool. These electron transfers occur concomitantly with less well-understood H^+ transfers (*4*−*6*).

On the basis of sequence similarity, Rieske proteins have been grouped into three clades: mitochondrial/bacterial, chloroplast/plastid, and outliers, which are the divergent sequences of Rieske proteins (*7*, *8*). Interestingly, the

[†] Supported by NIH Grants GM35342 (J.A.F.) and Program Project GM48495 (C.D.S.).

* Address correspondence to these authors. E-mail: (J.A.F.) jafee@scripps.edu or (C.D.S.) dave@scripps.edu.

[‡] University of California at San Diego.

[§] The Scripps Research Institute.

[||] Present address: Philipps-Universität Marburg, Institut für Pharmazeutische Chemie, Marbacher Weg 6, 35032 Marburg, Germany.

[⊥] Present address: Diversa Corp., 4955 Directors Place, San Diego, CA 92121.

[#] Present address: Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306-4380.

[▽] Present address: Astex Technology Ltd., 250 Cambridge Science Park, Milton Rd., Cambridge, CB4 0WE, UK.

[○] Present address: Syrrx Corp., 10450 Science Center Dr., Suite 100, San Diego, CA 92121-1125.

[@] Medical Research Council Dunn Human Nutrition Unit.

[1] In this paper, Rieske protein will refer to the Rieske protein from bc_1 and b_6f complexes and the Rieske portion from the arsenite oxidase fragment. Rieske-type proteins will refer to the dioxygenases. Abbreviations: AO, arsenite oxidase; BO, bovine Rieske protein; BP, biphenyl oxygenase; ND, naphthalene dioxygenase; SF, SoxF or *Sulfolobus* Rieske protein; SP, spinach Rieske protein; TT, *Thermus* Rieske protein; esd, estimated standard deviations of structural parameters; DFT, Density Functional Theory; MEAD, macroscopic electrostatics with atomic detail. SI, Supporting Information consists of 10 brief sections, I−X.

10.1021/bi0342719 CCC: $25.00
Published on Web 05/24/2003

sequence similarity of the *Thermus* Rieske with all other [2Fe−2S] Rieske proteins is very low; thus, it is considered an outlier. Other divergent Rieske proteins include the archeal Rieske proteins, the Rieske fragment from arsenite oxidase, and the Rieske-type ferredoxins that act as dioxygenases (*2*, *7*, *9*−*11*) (cf. also http://www.biochem.mu-luebeck.de/public/groups/schmidt_group/schmidt_group.html for a more extensive phylogenetic tree).

Crystal structures of Rieske and Rieske-type proteins from different sources have been published previously at 1.1−1.8 Å resolution and include examples from each of these categories (*2*, *8*−*12*). As previously noted (*2*), there is low sequence similarity between the proteins, yet they all show high structural similarity, especially in the cluster binding domain. The similarity of the *Tt* Rieske to these other Rieske proteins will be discussed in detail below

The structures of whole bc_1 complexes from bovine (*5*, *13*) and chicken (*14*) mitochondria as well as yeast (*15*, *16*) have also been reported at lower resolution. In different crystal forms of the bc_1 complex, the Rieske protein appears in several positions relative to the other redox centers. Analysis of these structures has led to the hypothesis that the Rieske protein shuttles electrons as it physically moves between the redox centers (refs *5* and *14* and see below). The suggested positions within the complex are near the quinol binding site (Q_o site) and near the propionate group of the cytochrome c_1 heme (the c_1-position), which are some 22 Å apart, and an intermediate position (*5*, *14*−*16*). Indeed, recent spectroscopic studies with Rieske protein in its reduced state provide evidence that one of the cluster Nϵ atoms interacts with stigmatelin bound in the Q_o site (*17*).

The redox potentials of the Rieske and Rieske-type proteins vary over a wide range, from −150 to +400 mV (*1*). The Rieske proteins from *bc* complexes that utilize plastoquinol or ubiquinol exhibit the more positive potentials at acid-pH (>+250 mV), whereas the menaquinol-utilizing complexes, including that found in *T. thermophilus*, show lower redox potentials (∼ +150 mV). All of the bc_1 and b_6f Rieske proteins also exhibit redox-linked ionizations; the redox potential is pH-dependent. The dioxygenases have the lowest potentials of the Rieske-type proteins, found at ∼ −150 mV and do not show pH-dependent reduction potentials in the physiological pH region (*1*). It has been reasonably assumed that coupled e^-/H^+ binding observed in electrochemical studies involves deprotonation of the Nϵ-bound H atoms of the Fe-bound, histidine imidazole groups; this idea is given additional support by the present work.

This is the first reported structure of a Rieske protein from an organism that utilizes menaquinol, and the structure has been refined to 1.3 Å resolution. Although several structures of Rieske and Rieske-type proteins have been previously reported, our analysis compares, in detail, extant structures of these proteins. The differences among this set of proteins and how they affect the macroscopic properties of the cluster, such as reduction potential, are discussed.

MATERIALS AND METHODS

Rieske Protein Expression. Gatti et al. (*7*) have discussed the difficulty of sequencing the high G + C containing fragment of *Thermus* DNA encoding *rpt*, the Rieske protein gene. The goal was to express a gene encoding the sequence of the earlier reported soluble protein, which began with the sequence SLRPREE and ended with TWRV*. Several attempts to PCR amplify this segment of the DNA failed.[2] Therefore, a strategy was developed to clone *rpt* without the use of PCR. Referring to Figure 2 of Gatti et al. (*7*), the *rpt* containing DNA was restricted with *BstE*II, a single site ∼20 bp downstream of the desired Ser-38 codon and at the *BamH*I site that is ∼300 bp downstream of the *rpt* stop codon. The expression vector pET17b was similarly cut at its *Nde*I and *BamH*I sites that occur in the multiple cloning site (Novagen, Madison, WI). Both the *BstE*II/*BamH*I *rpt* segment and the *Nde*I/*BamH*I segment of pET17b were isolated from 0.8% agarose gels and purified using a Qiagen II kit (Qiagen, Valencia, CA). The strategy then calls for a three-fragment ligation in which *rpt* is inserted into the multiple cloning site of pET17b by using an appropriate double-strand linker having a 5′-*BstE*II overhang and a 3′-*Nde*I overhang. Accordingly, two synthetic, complementary oligonucleotides were annealed into the double strand linker that encodes ATG of the initiating methionine followed by SLRPREE.

```
   NdeI
5′ TATGAGCCTCAGGCCCCGGGAGGAG 3′
3′   ACTCGGAGTCCGGGGCCCTCCTCCACTG 5′
                                BstEII
```

Finally, the *Nde*I and *BamH* I restricted pET17b expression vector, *BstE*II and *BamH*I restricted *rpt* gene fragment, and annealed linker were ligated together. *E. coli* strain DH10 was transformed to ampicillin resistance with the ligation mixture, and several colonies were isolated for analysis. The correct plasmid construct was verified by DNA sequencing on an Applied Biosystems Prism DNA sequencing system at the UCSD Department of Pathology. In contrast to the previously reported sequence (*7*), the DNA sequence indicates that codon 173 should encode for arginine rather than lysine, and this is confirmed by the structure.

Rieske Protein Purification and Characterization. Bacteria were grown in LB broth with 50 μg/mL Ampicillin to an OD_{600} of ∼1.0. The culture was induced with 0.4 mM IPTG and shaken for 6−8 h. Cells were collected by centrifugation and resuspended in a volume of 25 mL of 10 mM Tris HCl, pH 8.0 for each liter of bacterial culture. Lysis was achieved using 4 mg/mL lysozyme, 40 units/mL DNase, 3 units/mL RNase, 2 mM phenylmethylsulfonylfluoride, and 0.1% Triton X-100 by incubating for 60 min at 30 °C. The lysis mixture was centrifuged at 12 000*g* for 15 min at 4 °C, and the supernatant was brought to 55% v/w ammonium sulfate. This solution was allowed to stir for 15 min at room temperature and centrifuged as above. The pellet was discarded, and the supernatant was adjusted to 70% v/w ammonium sulfate and allowed to stir at room temperature for 15 min. After centrifugation as above, the supernatant was discarded, and the pellet was resuspended in 50 mM Tris HCl, pH 8.0, in a volume of ∼10 mL per liter of original bacterial culture. The protein was dialyzed against 50 mM Tris HCl, pH 8.0 and 0.1 mM EDTA, with two changes.

After dialysis, the protein was run though a 500 mL gravity column containing DE52 resin previously equilibrated with

[2] Recent advances in manipulating high G + C DNA have, however, removed some of these difficulties (e.g., the use of GC-Melt, BD Biosciences Clontech, Palo Alto, CA, in PCR and sequencing reactions).

Table 1: Data Collection and MAD Phasing

	MAD	high-resolution
unit cell parameters (Å)	$a = 58.72$ $b = 79.13$ $c = 136.78$	a = 58.57 $b = 78.27$ $c = 136.79$
space group $C222_1$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$
SSRL beamline	9−2	11−1

	peak	first remote	inflection (edge)	high resolution
wavelength (Å)	1.7373	1.3776	1.7415	0.965
resolution	50−2.0	50−1.8	50−2.9	50−1.3
(last shell)	(2.08−2)	(1.85−1.8)	(2.06−2.0)	(1.37−1.3)
total observations	31 662	415280	310598	319817
unique observations	26171	30284	20110	76407
redundancy	11.9	13.7	11.9	4.2
completeness %	99.2 (92.8)	100.0 (100.0)	98.9 (90.1)	98.8 (93.3)
I/σ_I	16.2 (3.8)	15.7 (3.5)	15.3 (2.9)	6.8 (1.9)
R_{symm} (I)	0.104 (0.413)	0.127 (0.509)	0.112 (0.480)	0.057 (0.431)
		Mad Phasing		
f'/f'', e$^-$	−5.7/4.2	−0.5/3.0	−7.44/3.59	

25 mM Tris HCl, pH 8.0 and washed with 1 L of the same buffer. Protein was eluted using the same buffer containing 1 M NaCl; the brown-colored fractions contained the Rieske protein and were combined and dialyzed against 2 L of 25 mM Tris HCl, pH 8.0. The protein was further purified using a 5 mL HiTrap Q HP FPLC column equilibrated with 100 mL of 20 mM Tris, pH 8.0 (buffer A). The protein was eluted using a gradient of 0% B for 1−10 min, followed by 0−30% B for 10−25 min, and 30−100% B for 25−70 min at 1 mL per minute. Buffer B is 0.5 M NaCl, 20 mM Tris, pH 8.0. The protein eluted at approximately 80% B, and the brown fractions contained the Rieske protein.

The final step of purification was gel filtration on a HiLoad 16/60 Superdex 75 FPLC column. The protein solution was concentrated, and aliquots of ∼2 mL were loaded onto the gel filtration column previously equilibrated with 0.1 M NaCl, 25 mM Tris, pH 8.0. The column was run at 1 mL/min for 95 min after which two major fractions appeared. The second fraction was brown and contained the Rieske protein. The protein was analyzed for purity using a ratio of the absorbance at 456−280 nm (A_{456}/A_{280}), which was consistently near 0.23 (*18*). Sodium dodecyl sulfate−polyacrylamide gel electrophoresis exhibited a single band that may be composed of more than one polypeptide; there were no evident non-Rieske protein impurities.

Electrospray mass spectrometric analysis of samples that had been exchanged into 10 mM NH_4^+ acetate in water revealed majority *m*/*z* peaks at 2993 (+6) and 2566 (+7), corresponding to a mass of 17 952 Da and a single ion at 2643 (+7) that corresponds to a neutral mass of 18 496 ± 2 Da, which can be rationalized in terms of the available information; see SI-I in the Supporting Information for a discussion of these data. Optical absorption spectra and helium temperature EPR spectra of the dithionite reduced material are indistinguishable from those reported earlier for the native protein (not shown, refs *19* and *20*)).

Crystallization Conditions and Structure Determination. Large crystals of the Rieske protein were formed in different 19−21 μL sitting drops that contained 1−2 μL of 20 mg/mL protein combined with 8 μL of well solution, which contained 25% PEG 4000, 100 mM Tris, pH 8.5, 0.1−0.4 M $MgCl_2$, and 10−11 μL of H_2O. Other drops contained 4 μL of protein, 12 μL of well (same as above), and 4 μL of H_2O. All drops also had 1 μL of HECAMEG (Hampton Research) as a detergent. They were equilibrated against 0.5 mL of well solution. The crystal tray was kept in a 24 °C incubator for ∼24−48 h and then moved to an 8 °C incubator after which crystals appeared in 1−2 days as rectangular rods up to 1 mm in length—usually with frayed ends. Two solutions containing 25% PEG 4000, 100 mM Tris, pH 8.5, 0.1−0.4 M $MgCl_2$, and 10 or 20% ethylene glycol served as cryo-buffers. The crystals were frozen using a 10:20 dip method where a crystal was removed from its mother liquor with a nylon loop, dipped into the 10% ethylene glycol cryo-buffer, followed by a quick dip in the 20% ethylene glycol cryo-buffer, and immediately frozen in liquid N_2; they are quite unstable in the 20% ethylene glycol cryo-protectant, being destroyed within seconds. The crystals belong to the $C222_1$ space group (number 20) and have two molecules in the asymmetric unit. Crystals grown under these conditions have unit cell parameters: $a = 58.6 \pm 0.1$ Å, $b = 78.3 \pm 0.2$ Å, $c = 136.8 \pm 0.4$ Å, $\alpha = \beta = \gamma = 90°$ and have a solvent content of 41.8%.

This space group has eight symmetry operators of which two are nonintersecting but orthogonal 2-fold axes: a 2-fold axis parallel to *a* through the origin and a 2-fold axes parallel to *b* at 0 on *a* and 1/4 on *c*. The independent copies of the Rieske protein are packed with respect to the crystallographic 2-fold axes in a manner that places the [Fe−S] cluster ligand (His134), of each copy, adjacent to a 2-fold axis. Hence, the His134−His134′ interaction observed in the crystals (described below) is both 2-fold symmetric and present for each copy of the protein in the asymmetric unit.

The Rieske protein structure was determined from a three-wavelength MAD experiment using the anomalous scattering of Fe (Table 1). Highly redundant data were collected at SSRL beamline 9−2 and processed using MOSFLM in the CCP4 suite of programs (*21*). Data quality was evaluated using the program XPREP (http://www.bruker-axs.de/index.html?page=/products/scd/xprep.php). For the peak wavelength, the anomalous signal/noise was significant (>1.5) to 2.2 Å resolution. The anomalous correlation coefficients between datasets were greater than 30 for data to 2.4 Å resolution. Positions of the four expected Fe peaks (two Fe/

Table 2: Refined Model Statistics

residue range (molecules A and B)	46−201
no. of non-H atoms (in asymmetric unit)	3023
no. of water molecules	351
no. of Fe atoms	4
no. of Mg ions	3
R (%) ($F_o > 4\sigma F_o$/all data)	13.7/15.8
R_{Free}(%) ($F_o > 4\sigma F_o$/all data)	18.9/21.2
rms bond length deviations (Å)	0.12
rms angle deviations (deg)	2.28

Average B[a] Values (Å^2)		
water molecules	34.7	
	molecule A	molecule B
protein atoms	24.0	23.2
main chain	21.2	19.8
side chain	27.2	27.3
metal cluster	14.7	13.5
Ramachandran plot		
residues in most favored regions (%)	91.6	90.8
residues in additional allowed regions (%)	8.4	9.2

[a] Isotropic B factors.

molecule and two molecules/asymmetric unit) were determined at 2.5 Å using SHELXD (*22*), and the shortest distance allowed between atoms was 2.5 Å. The correlation coefficient for the best solution was 67.4 and $R = 33.8\%$ for data to 2.5 Å resolution. Phases were calculated from these Fe positions using the program SHELXE (*23*). A final contrast of 0.4 and connectivity of 0.93 for all data were higher than for the inverted sites (contrast 0.33, connectivity 0.90). Density modification was performed in DM (*24*) using solvent flattening and histogram matching. The resulting electron density showed clear solvent boundaries (Figure 1a).

For rapid tracing of the protein main chain, the program ARP/wARP version 5.1 (*25*) in combination with the CCP4 program package (*21*) was used on the MAD data that were combined with a 1.3 Å data set collected on a second crystal (Table 1). The auto main chain tracing routine produced seven polypeptide fragments (294 residues, 85%) and a connectivity index of 0.95 for data to 1.3 Å resolution. It was possible to clearly distinguish the two independent monomers in the asymmetric unit. Chain direction and side-chain assignment could be determined unambiguously by manual inspection of the electron density in O (*26*) and in Xtalview (*27*). Side chains were placed into electron density using the automated procedure implemented in ARP/wARP (*25*), and any side chains not built by the program were inserted manually using Xtalview (*27*) (Figure 1b).

The structure was refined using SHELXL (*28*) to 1.3 Å resolution (Table 2). One cycle of rigid body refinement at 3 Å was followed by 10 rounds of refinement cycles using data to 1.3 Å; each round of refinement was comprised of 20 steps of conjugate-gradient least squares refinement with standard restraints on bond distances, angles, and B values, but no constraints were imposed on the [2Fe−2S] cluster. Anisotropic displacement parameters were introduced for heavy atoms after four rounds of refinement and applied to all non-hydrogen atoms after eight rounds. Riding hydrogens were added, and anisotropic displacement parameters were applied to all non-hydrogen atoms in the final two rounds. A single cycle of block, full-matrix, least-squares refinement, using the L. S. and BLOC 1 commands of SHELXL (*28*), was performed on the final structure to calculate esd (estimated standard deviations) for cluster distances. The quality of the model was evaluated using PROCHECK (*29*) (Table 2). More than 90% of the φ/ψ angles are in the most favored regions with the remainder being in allowed regions (Table 2). PDB deposition number is 1NYK.

Comparison of 3-D Structures. To identify 3-D structures similar to the TT protein within the PDB database, the DALI server (www.ebi.ac.uk/dali/) produced a list of six Rieske and Rieske-type proteins. The protein ranking[3] with the highest structural similarity to the *Thermus* Rieske (TT) was the Rieske-type protein from the arsenite oxidase (AO) fragment isolated from *Alcaligenes faecalis* (*2*) ($Z = 14.5$) (PDB 168J) followed by SoxF (SF) from *Sulfolobus acidocaldarius* ($Z = 11.4$) (*11*) (PDB 1JM1). Biphenyl dioxygenase (BP) ($Z = 11.3$) (*10*) (PDB 1FQT), the Rieske protein from spinach chloroplasts (SP) ($Z = 10.8$) (*8*) (PDB 1FRS), the Rieske protein from bovine heart (BO) mitochondria ($Z = 10$) (*12*) (PDB 1RIE), and naphthalene dioxygenase ($Z = 5.8$) (*9*) (PDB 1NDO) followed. Structure-based amino acid alignment of these proteins was carried out using the program SEQUOIA; overall alignment score for the seven PDB files was 5.56×10^3 with 99 equivalances (*30*) (cf. http://bruns.homeip.net/~bruns/sequoia.html).

Other programs used in this work were GETAREA1.1, which calculates the exposed area of each amino acid side chain (http://www.scsb.utmb.edu/cgi-bin/get_a_form.tcl (*31*)), and MOLIMAGE (D. E. McRee, Molecular Images Software, San Diego, CA).

RESULTS

The structure of the *T. thermophilus* Rieske protein at pH 8.5 has been determined by MAD techniques using the intrinsic Fe atoms of the [2Fe−2S] cluster. The two molecules in the asymmetric unit are only slightly different. (Some details are presented in the Supporting Information, section II (SI-II).) The electron density was very well-defined for both molecules, and model structures containing 156 residues were built for each; the N- and C-terminal eight to nine residues are disordered. The model was then refined against a 1.3 Å data set (Table 2, Figure 1a−c).

The overall folding of the polypeptide is similar to those reported for related proteins and is dominated by 10 antiparallel β-strands (Figures 2a and 3). The protein consists of two domains. The large, or base, domain is composed of an irregular β-barrel formed by $\beta4-\beta3-\beta2-\beta1-\beta10-\beta9$, where $\beta4$ and $\beta9$ share hydrogen bonds to close the barrel (Figures 2a and 3); the barrel is capped by a loop connecting $\beta1$ and $\beta2$ (residues 60−81). The smaller, cluster-binding domain is composed of $\beta5-\beta8$ and three loops that link $\beta4-\beta5$, $\beta6-\beta7$, and $\beta8-\beta9$. The [2Fe−2S] cluster is ligated by

[3] The Z score is a measure of similarity reported from the search. A score of $Z < 2$ is considered not similar.

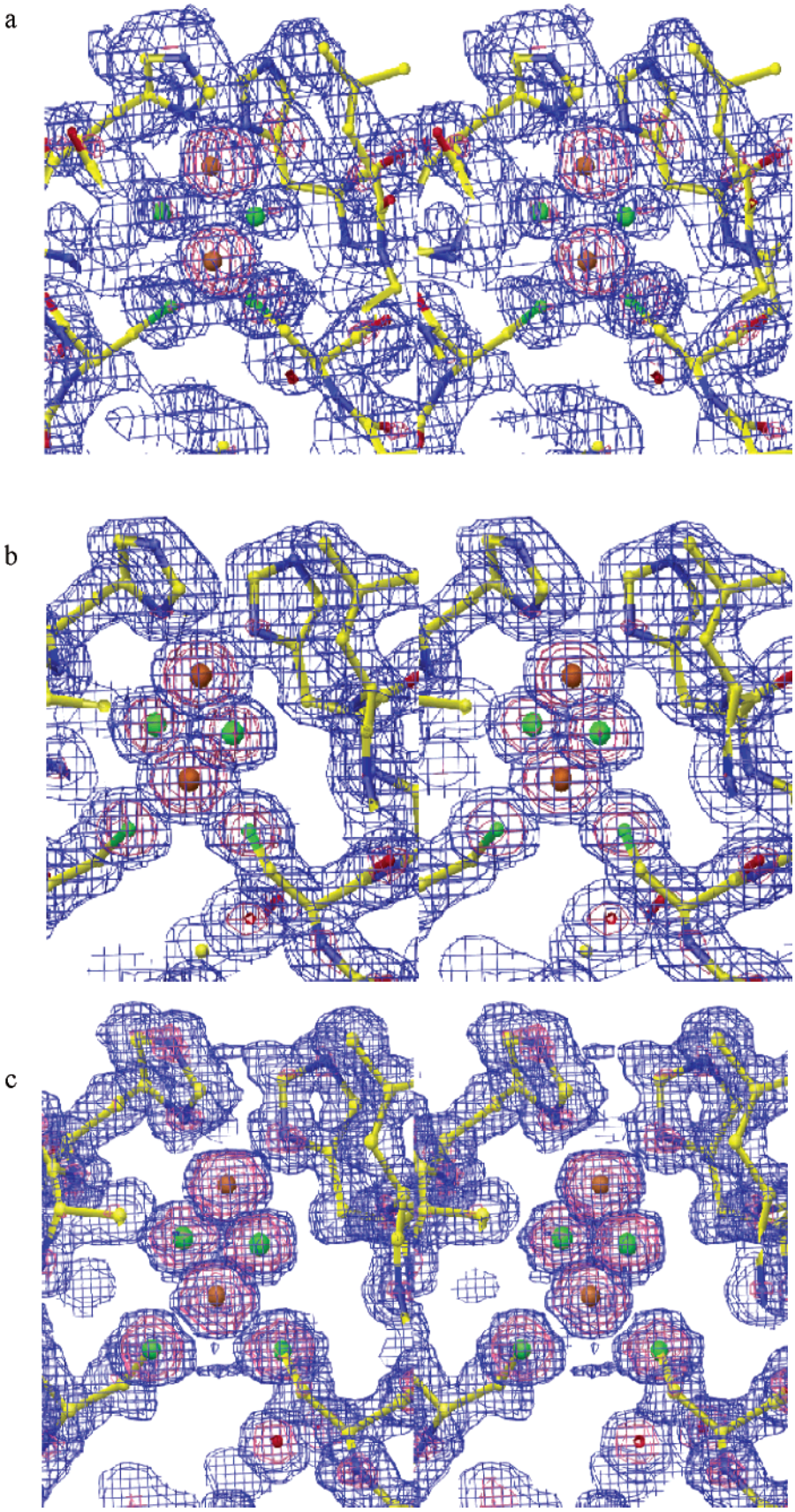

FIGURE 1: Stereoview of the electron density at the [2Fe−2S] cluster. (a) Experimental electron density calculated from MAD phases to 2.5 Å resolution. (b) $2|F_o| - |F_c|$ map of the electron density including data to 1.3 Å after density modification and model building in ARP/wARP. The models in both a and b are from the output of ARP/wARP. (c) Final SigmaA weighted $2|F_o| - |F_c|$ density at 1.3 Å with the final model. In all maps, blue contours correspond to 1σ and red contours to 5σ; model and density are shown from the A protein of the asymmetric unit.

two cysteine Sγ atoms (Cys-132 and Cys-151) and two histidine Nδ (His-134 and His-154) residues (Figures 1 and 2b). Fe1 and its associated S atoms are shielded from solvent, whereas the imidazole rings of the histidine ligands are partially exposed (see below). There is also a disulfide bridge between the Sγ atoms of Cys-137 and Cys-153 on the surface that tethers the two ligand-containing loops together. The β8$-\beta$9 loop, termed the Pro-loop, is rich in proline.

The two subdomains are covalently connected by regions we term the crossover segments. The crossover from the large to the small domain corresponds to the N-terminal region of β4, whereas the end of the Pro-loop to β9 serves as the crossover when re-entering the large domain (Figures 2a and 3); these are highly conserved in 3-D space (see below). In a larger sense, the entire cluster binding domain may be viewed as an extended loop that begins as β4, exits the large domain, and returns to β9 in the large domain. The cluster domain interacts extensively but noncovalently with the large domain through its β5$-\beta$6 and β7$-\beta$8 loops (see below). The β-barrel of the large domain is folded by rolling β9 upward and onto β4, β10 onto β3, and β1 onto β2. In the cluster binding domain, β8 is folded onto β5 and β7 onto β6 (Figure 3). Note that this presentation of the β-strands differs from that of Iwata et al. (*12*).

Metric Features of the Cluster. These, along with standard uncertainties, are presented as averaged values in the schematic of Figure 2b and are compared with the SoxF structure in Table 3. As discussed by Guss et al. (*32*), there are many pitfalls in refining metal-to-ligand atom distances. Full matrix refinement provides an additional level of confidence, especially for higher resolution data; in our case, we obtained esds from a single cycle of block, full-matrix, least-squares refinement (*28*) and found esds of ~0.01 Å for Fe$-$S bonds and ~0.02 Å for Fe$-$N bonds. Similar approaches in the cases of Fe/S clusters (*33*) at 1.35 Å resolution, and the Cu_A center of cytochrome *c* oxidase (*34*) at 1.6 Å resolution gave metal to sulfur esds of ~0.01 and 0.03 Å, respectively. The structures are considered to be accurate within these limits.[4]

All three Rieske structures share a common distortion at Fe1, as indicated by the length of the smaller numbered cysteine residue Sγ to Fe1 bond being greater than the larger numbered cysteine residue Sγ to Fe1 bond by 0.017 ± 0.004 Å, 0.050 ± 0.018 Å, and 0.051 ± 0.015 Å in SoxF, *Thermus* A and B molecules, respectively. The origin of this geometric distortion is not evident.

The Fe$-S^{2-}$ bonds of the SoxF rhombus are all identical within their esds, indicating that the cluster in SoxF is more symmetric than in the *Thermus* molecules. In the A molecule of the *Thermus* structure, Fe1$-$S1 is 0.042 ± 0.02 Å longer than Fe1$-$S2; similarly, the Fe2$-$S2 bond is 0.03 ± 0.02 Å longer than the Fe2$-$S1 bond. This corresponds to a lengthening of the cluster along two opposite sides. In SoxF, the Fe2$-$Nδ (His-142) differs from the Fe2$-$Nδ (His-173) by 0.017 ± 0.011 Å whereas in *Thermus* molecule A, the Fe2$-$Nδ (His-134) bond is 0.064 ± 0.047 Å shorter than the Fe2$-$Nδ (His-154) bond.

[4] The average Fe$-$N bond lengths for the oxidized proteins in Table 3 are longer by a factor of ~1.07 than the average Fe$-$N bond lengths reported by EXAFS for sulredoxin from *Sulfolobus* and anthranilate dioxygenase from *Acinetobacter* sp. SDP1 (*65*). We have no explanation for this discrepancy.

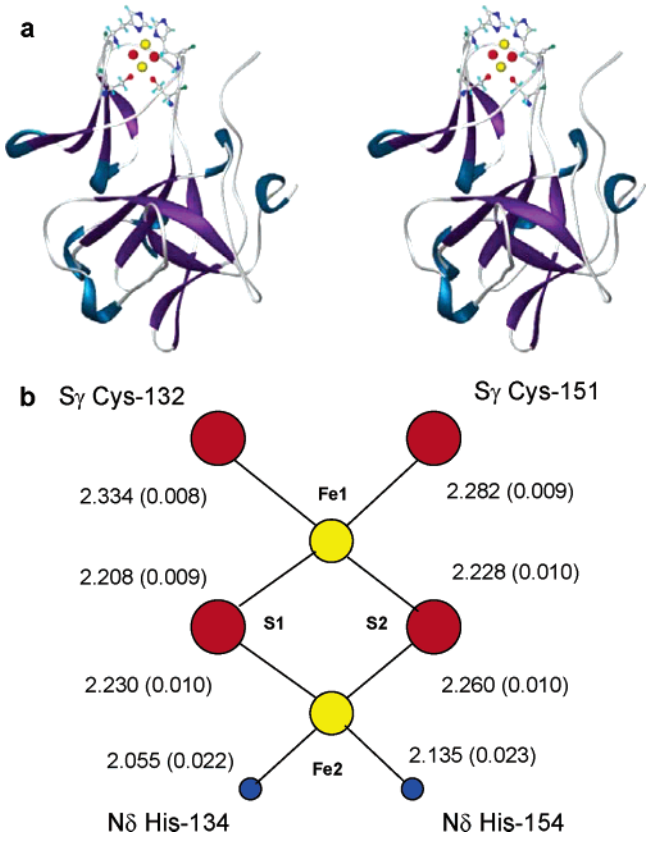

FIGURE 2: (a) Ribbon model of the *T. thermophilus* Rieske protein in wall-eye stereo. β-sheets are shown in purple, α- and 3_{10}-helices are shown in blue, and β-turns are shown in dark gray. The ligands to the [2Fe$-$2S] cluster and the cluster are also shown with S atoms in red, Fe atoms in yellow, and N atoms in blue. (b) Schematic presentation of the cluster structure with average bond lengths from molecules A and B (Table 3) and estimated standard deviations. Fe1, Fe2, CysSγ-132, CysSγ-151, HisNδ-134, and HisNδ-154 form a rough plane, while S1 projects out of the plane and S2 into the plane.

In molecule B, with the exception of a somewhat longer Fe2$-$S2 bond, the apparently symmetric distortion of the rhombus observed in molecule A is not found. However, the Fe2$-$Nδ (His-134) bond in molecule B is 0.096 ± 0.045 Å shorter than the Fe2$-$Nδ (His-154) bond. These small geometric distortions may require rather little energy (*35*), and their presence is likely due to the different lattice environment around the two His-134 Nδ atoms in this particular lattice. The shortening of one of the Fe2$-$Nδ bond lengths (His-134) is consistent with partial deprotonation of the His-134 Nϵ atom.

Possible Deprotonation of One Iron-Bound Histidine Residue. A unique feature of this particular crystal form is the relation of the [2Fe$-$2S] cluster to its symmetry related cluster. In both the A and the B molecules, one of the ligating histidines, His-134, is hydrogen bonded to its counterpart (His-134′) across a crystallographic 2-fold axis (Figure 4). The N$\epsilon-$Nϵ' distance across the symmetry axis is 2.69 Å for the A molecule and 2.76 Å for the B molecule. We suggest that the shorter Fe2$-$Nδ (His-134) bond results from this imidazole ring being involved in an equilibrium such that each Nϵ atom across the symmetry axis is protonated nominally half of the time and each Fe2$-$Nδ bond is the

Table 3: Bond Distances and Angles of the [2Fe−2S] Clusters of *Thermus* Rieske Protein as Compared with Those of Sox F

	SoxF[a]	*Thermus* molecule A[b]	*Thermus* molecule B[b]
resolution of structure (Å)	1.1	1.3	1.3
pH of crystal	7.0	8.5	8.5
		[2Fe−2S] cluster−Bond distances (esd)[c] Å	
Fe1−S1	2.2672 (0.0021)	2.19_9 (0.0095)[d]	2.21_8 (0.0087)
Fe1−S2	2.2626 (0.0018)	2.23_1 (0.0097)	2.22_5 (0.0093)
Fe2−S1	2.2582 (0.0019)	2.23_2 (0.0097)	2.22_9 (0.0093)
Fe2−S2	2.2589 (0.0020)	2.26_2 (0.0095)	2.25_7 (0.0091)
Fe1−Fe2	2.7191 (0.0013)	2.69_4 (0.0060)	2.68_3 (0.0061)
Fe1−S1−Fe2	73.86 (0.06)	74.8_9 (0.31)	74.2_3 (0.28)
Fe1−S2−Fe2	73.94 (0.06)	73.6_7 (0.30)	73.5_5 (0.28)
S1−Fe1−S2	105.81 (0.07)	106.7_9 (0.33)	106.8_2 (0.33)
S1−Fe2−S2	106.24 (0.07)	104.6_2 (0.33)	105.3_7 (0.33)
		Cluster ligands−Bond distances (esd) Å	
Fe1−CysSγ	2.3484 (0.0017) (Cys-140)	2.32_2 (0.0089) (Cys-132)	2.34_5 (0.0075) (Cys-132)
Fe1−CysSγ	2.3319 (0.0021) (Cys-170)	2.27_0 (0.0095) (Cys-151)	2.29_4 (0.0079) (Cys-151)
Fe2−HisNδ	2.0997 (0.0057) (His-142)	2.05_0 (0.0242) (His-134)	2.06_1 (0.0207) (His-134)
Fe2−HisNδ	2.0827 (0.0051) (His-173)	2.11_4 (0.0225) (His-154)	2.15_7 (0.0239) (His-154)
		Cluster ligands−Bond distances (esd) Å	
Fe1−CysSγ	2.3484 (0.0017) (Cys-140)	2.32_2 (0.0089) (Cys-132)	2.34_5 (0.0075) (Cys-132)
Fe1−CysSγ	2.3319 (0.0021) (Cys-170)	2.27_0 (0.0095) (Cys-151)	2.29_4 (0.0079) (Cys-151)
Fe2−HisNδ	2.0997 (0.0057) (His-142)	2.05_0 (0.0242) (His-134)	2.06_1 (0.0207) (His-134)
Fe2−HisNδ	2.0827 (0.0051) (His-173)	2.11_4 (0.0225) (His-154)	2.15_7 (0.0239) (His-154)
Sγ −Fe1−Sγ	109.73 (0.07) (Cys-140/Cys-170)	110.6_6 (0.31) (Cys-132/(Cys-151)	109.6_8 (0.3) (Cys-132/Cys-151)
Nδ−Fe2−Nδ	92.12 (0.22) (His-142/His-173)	94.2_7 (0.76) (His-134/His-154)	94.7_1 (0.8) (His-134/His-154)

[a] Data for SoxF (PDB 1JM1) was adapted from Bönisch et al. (*11*). [b] Data taken from Shelx least-squares refinement (see text). [c] Estimated standard deviation (esd). [d] The third significant figure was obtained by rounding and is therefore presented as a subscript.

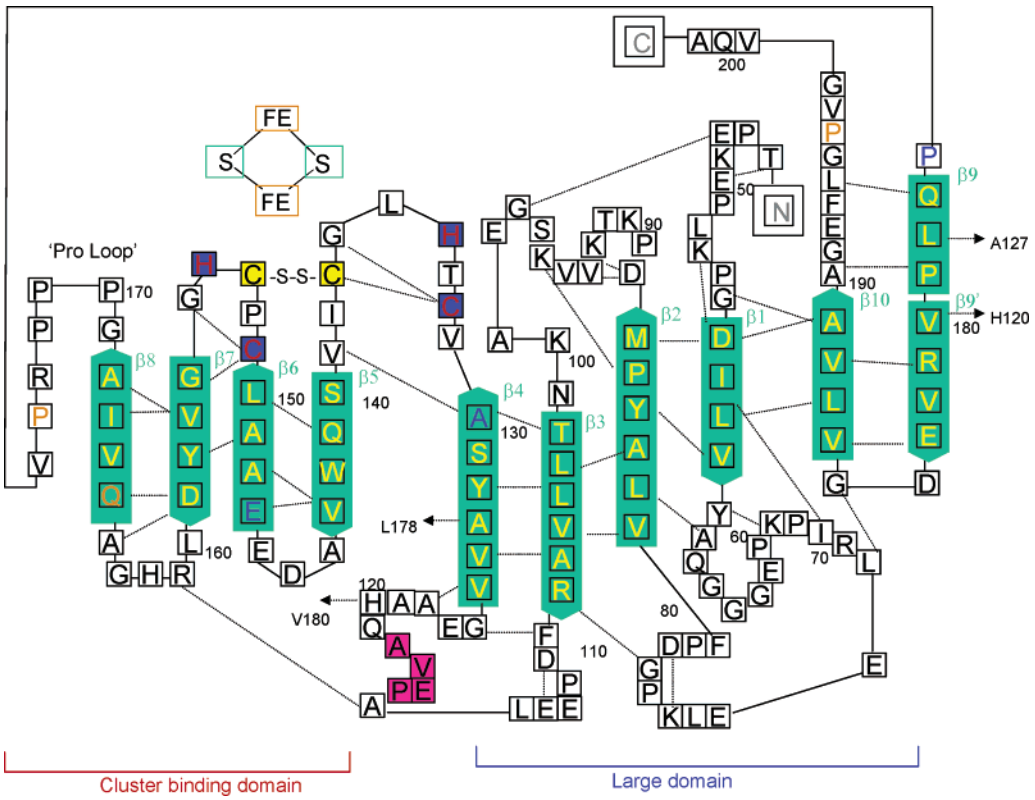

FIGURE 3: Topology map of the *Thermus* Rieske protein. β-sheets are shown as green box arrows and are numbered in green from β1−β10. Every 10th residue is numbered in black. The cluster ligands are indicated as blue boxes with red letters, the residues participating in the disulfide bond are shown in yellow filled boxes, and pink filled boxes indicate α-helical regions. Residues in close 3-D proximity have like colors. Dashed lines indicate hydrogen bonds between main chain polar groups. To make the β-barrel of the large domain 3-D structure, roll β9/β9′ over to β4. For the small domain, lay β5−β8 over β4−β2, where P_{174}, Q_{165}, and P_{196} are near in space.

time and space average of the properties of Fe2−HisH and Fe2−His^-.

$$N\epsilon-H\cdots {}^{-}N\epsilon' \leftrightarrow N\epsilon^{-}H-N\epsilon'$$

Assuming that the Nϵ atom of the other His-ligand, His-154, is protonated, this crystal packing arrangement would indicate that each cluster binds an average of 1.5 protons to the Fe2-bound, histidine-imidazole rings. This hydrogen

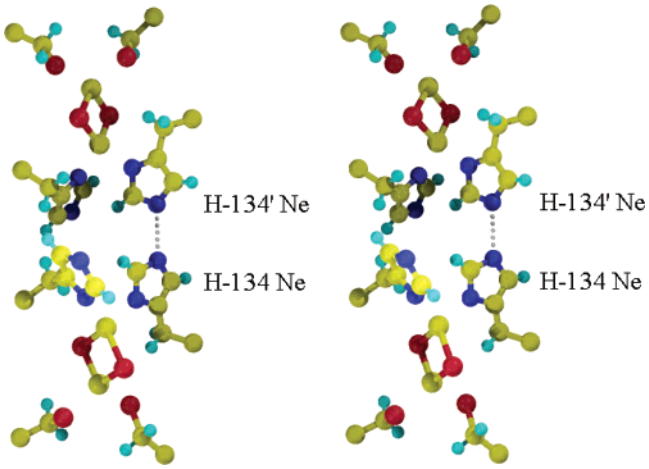

FIGURE 4: Wall-eye stereoview of symmetry related clusters. The four ligands to the [2Fe−2S] cluster are shown in relation to the symmetry related cluster that shares a hydrogen bond with one of the histidine ligands. The distance between the histidine (His-134) and the symmetry related histidine (His-134′) is ~2.7 Å across a crystallographic 2-fold axis. This is observed with both A and B molecules and corresponds to two of the 2-fold axes in the $C222_1$ space group. The cluster shown is from the A protein of the asymmetric unit, and the 2-fold lies in the plane from left to right and bisects the dashed line.

bond, present in both molecules in the asymmetric unit, is evidently important in the crystal lattice because either reduction with dithionite or lowering the pH to 6.5, both of which would result in the uptake of a proton by the cluster (*36*), destroy this feature of the lattice and cause the crystals to rapidly dissolve.

Fitting Thermus Rieske Protein into Bovine bc_1. Although an intact bc_1 complex has not been isolated from *Thermus* membranes, DNA sequence (*7*) and spectral/metabolic (*37*) information indicates that one exists. It is therefore important to see how the Rieske protein would fit into the larger complex. To this end, the *Tt* Rieske protein was aligned with that of bovine Rieske protein as observed in whole bc_1 complex, where the Rieske protein is interacting with cytochrome c_1 (c_1 state). The *Thermus* and bovine electrostatic potential maps display remarkably similar charge distribution, including a hydrophobic patch near the region of the cluster, suggesting that the two proteins perform similar functions in vivo (see Supporting Information SI-III). It is also noteworthy that the highly variable loop that joins β3 and β4 points away from the complex when in this orientation (see below).

Comparison of Overall Folds. A search for similar Rieske protein structures (see Materials and Methods) yielded representatives from each of the major clades, mitochondrial/bacterial, chloroplast/plastid, and outliers. Accordingly, comparisons can be made between each type of Rieske protein. The sequence alignment of these proteins, based on both structural and residue alignment (using the program SEQUOIA, see Materials and Methods), reveals that in comparing the seven Rieske and Rieske-type proteins, only the four ligands to the [2Fe−2S] cluster are conserved (Figure 5). Thus, there is extremely low sequence similarity between all Rieske and Rieske-type proteins. However, within the proteins from bc_1 and b_6f complexes (top four sequences in Figure 5), there are a few more conserved residues (yellow boxes), as has been reported previously (*8*), including several cluster binding loop residues and a disulfide bridge (Figure 5, green arrow). An overlay of these four structures is shown in the Supporting Information (see SI-IV).

In the dioxygenases (Figure 5, bottom two sequences), the two conserved cysteine residues forming the disulfide have been replaced with interacting hydrophobic residues. In addition, phenylalanine replaces a conserved tyrosine (Tyr 158 in *Thermus*) from bc_1 and b_6f proteins (the effect of this is discussed below). An overlay of these two structures, along with that of the arsenite oxidase Rieske fragment, are shown in the Supporting Information (see SI-IV).

Even with the extremely low sequence homology between the seven Rieske proteins, the 3-D structural similarity is high. Comparison of these structures illustrates that, as expected, the cluster binding domains are highly similar, but the large domains differ. Of interest is that all of the β-strands (β1−β10) align very closely. Figure 6 shows the overlay of the seven Rieske and Rieske-type proteins highlighting the basic 10 β-strand motif (shown in maroon) within each of the proteins.

Comparison of Cluster Binding Domains. The details of the [2Fe−2S] binding region within the representative proteins reveals that, while their overall folds are similar, small differences in geometry and solvent exposure exist between the Rieske and the Rieske-type proteins. Differences in these features may be important in determining high-resolution spectroscopic (*38*) properties as well as the redox behavior of each protein.

We used the torsion angle defined as that between the projection of the CE1−ND1 bond of the imidazole ring and the projection of the Fe2−Fe1 vector of the cluster in a plane normal to the ND1−Fe2 bond. This angle, τ, is limited to $-\pi < \tau \leq +\pi$; a positive angle occurs when a clockwise motion of the projection of the CE1−ND1 bond moves it onto the projection of the Fe2−Fe1 vector (cf. p 60 of ref *39*). Among the respiratory Rieske proteins, τ_L, that of the lesser numbered histidine (His_L) imidazole ring, varies from ~177 to 171° and τ_G, that of the greater numbered histidine (His_G) imidazole ring, varies from 98 to 92°. Hence, the planes of the two Fe-ligated imidazole rings are approximately normal to each other (Figure 1 and 2a). In the Rieske-type proteins, these angles deviate significantly from the ranges found in the respiratory proteins (see Supporting Information SI-V).

The exposed area of each Fe-ligated histidine imidazole side chain is greatly different between the His_L and the His_G (calculated using GETAREA1.1, see Materials and Methods). Importantly, neither ring is fully exposed to solvent, a factor that may contribute to their having different apparent pK_a values (*36*). In the respiratory type Rieske proteins, the percentages of solvent exposure of His_L range from 21.1 to 27.6%, while those of His_G range from 42.8 to 46.9%, indicating that, on average, His_G is approximately twice as exposed to solvent as is His_L. This difference is engendered largely by the presence of the neighboring leucine residue (see Figures 1 and 5), the side chain of which abuts the imidazole ring of His_L (see Figure 1). In AO and BP, the differences are substantially smaller between the two histidine imidazole rings (see Supporting Information SI-V), and in the case of BP their levels of solvation are nearly equal.

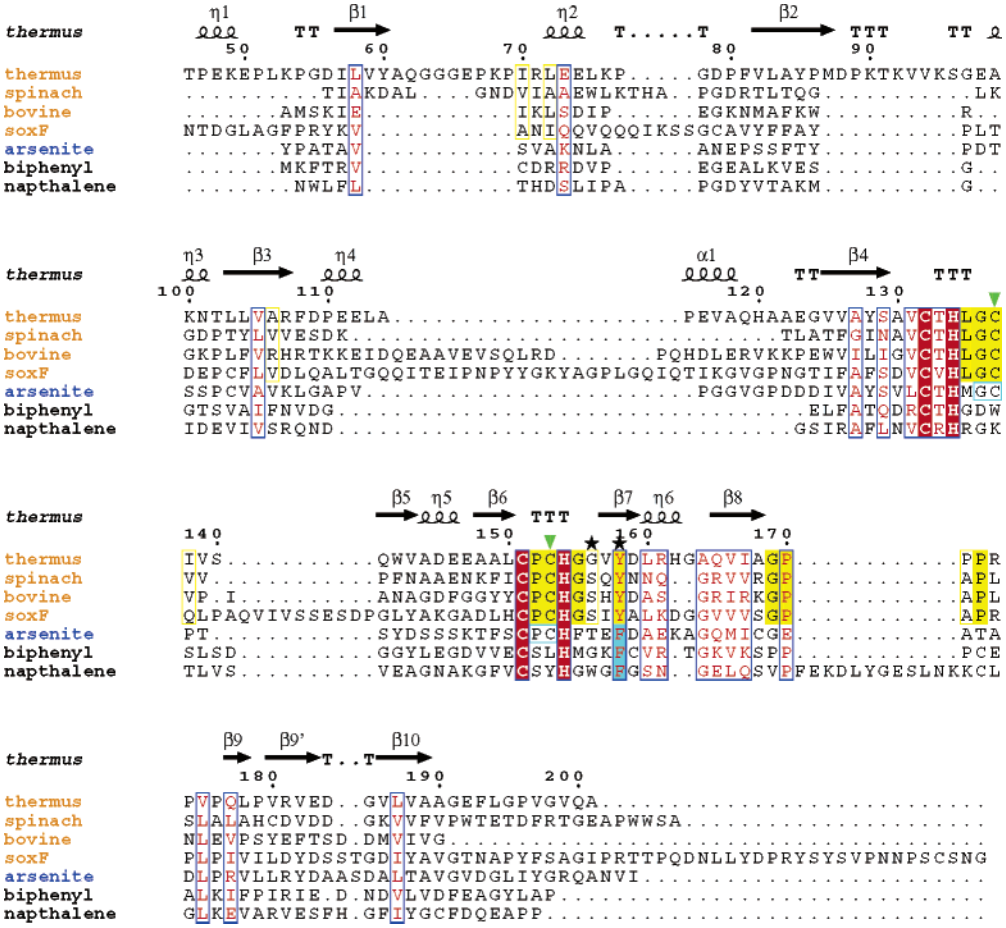

FIGURE 5: Structure-based sequence alignment of seven Rieske-type proteins. Sequence alignment based on both residue alignment and 3-D structural alignment was generated using the program SEQUOIA (*30*). Secondary elements from the *Thermus* Rieske protein are shown on top. The top four sequences (named in orange) are from bc_1 and b_6f complexes (bovine PDB 1RIE; spinach 1RFS; and SoxF 1JM1). Arsenite is the Rieske protein from the arsenite oxidase fragment isolated from *Alcaligenes faecalis* (1G8J). Bottom two sequences are biphenyl (1FQT) and naphthalene (1NDO) dioxygenases. Red boxed residues are strictly conserved residues, and blue boxed amino acids are conservative replacements. For the top four sequences, filled yellow boxes show residues conserved within respiratory-type Rieske proteins, and yellow boxes show conservative variation. Blue filled and open boxes show conserved or nearly conserved residues in Rieske-type proteins, respectively. The green triangles specify the cysteine residues involved in disulfide formation, and the black stars indicate those residues that have been the object of mutational studies that affect reduction potential. The alignment figure was made using ESPript (*63*).

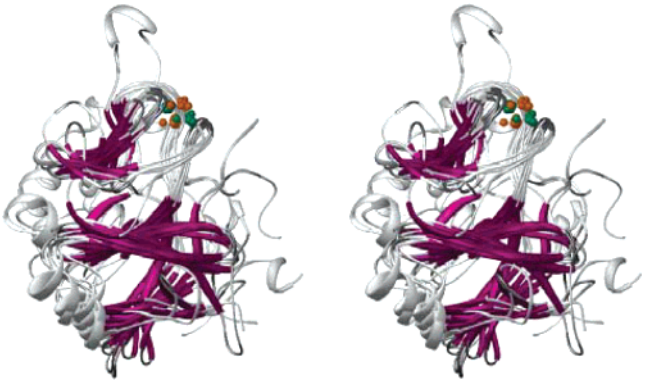

FIGURE 6: Comparison of the overall β-sheet folding of seven Rieske proteins. Wall-eye stereoview of seven overlaid Rieske proteins highlighting how the maroon β-strands of each subdomain align. The positions of the seven [2Fe−2S] clusters are shown. Fe (orange) and S (green). The superposition was carried out by the program SEQUOIA (see text).

Comparison of Topologies. For each of the seven proteins being compared, a topology map was generated to compare the length, hydrophobicity, number of charged amino acids, and secondary structural elements within the β-strands and connecting loops. With the exception of β4, the lengths of the β-strands are fairly consistent across the different proteins. However, the loops show much more diversity. (See Supporting Information SI-VI for presentations of the topology maps for the different proteins examined here.)

Loops. A detailed representation of the loops is given in Figure 7. Within the smaller cluster binding domain, the loop and β-sheet structures are very similar, and in particular, the two loops that contain the cluster ligands (β4 to β5 and β6 to β7). The loop joining β6 to β7 is the most conserved in both length and number of charged amino acids. The Pro-loop (β8 to β9) is also generally conserved, with the exception of naphthalene dioxygenase, which has an extended loop. The larger domain shows greater variation across the different proteins. The loop that connects β3 to β4 appears to have the most diversity. The length of the loop ranges from two amino acid residues in BP to 38 in SoxF and can also include different secondary structure elements (Figure 7). The net charge on this loop varies from −4 in TT to +1 in BO. In the bovine bc_1 complex, this loop is oriented away from the rest of the complex (Figure 1b in Supporting Information), possibly indicating why that face

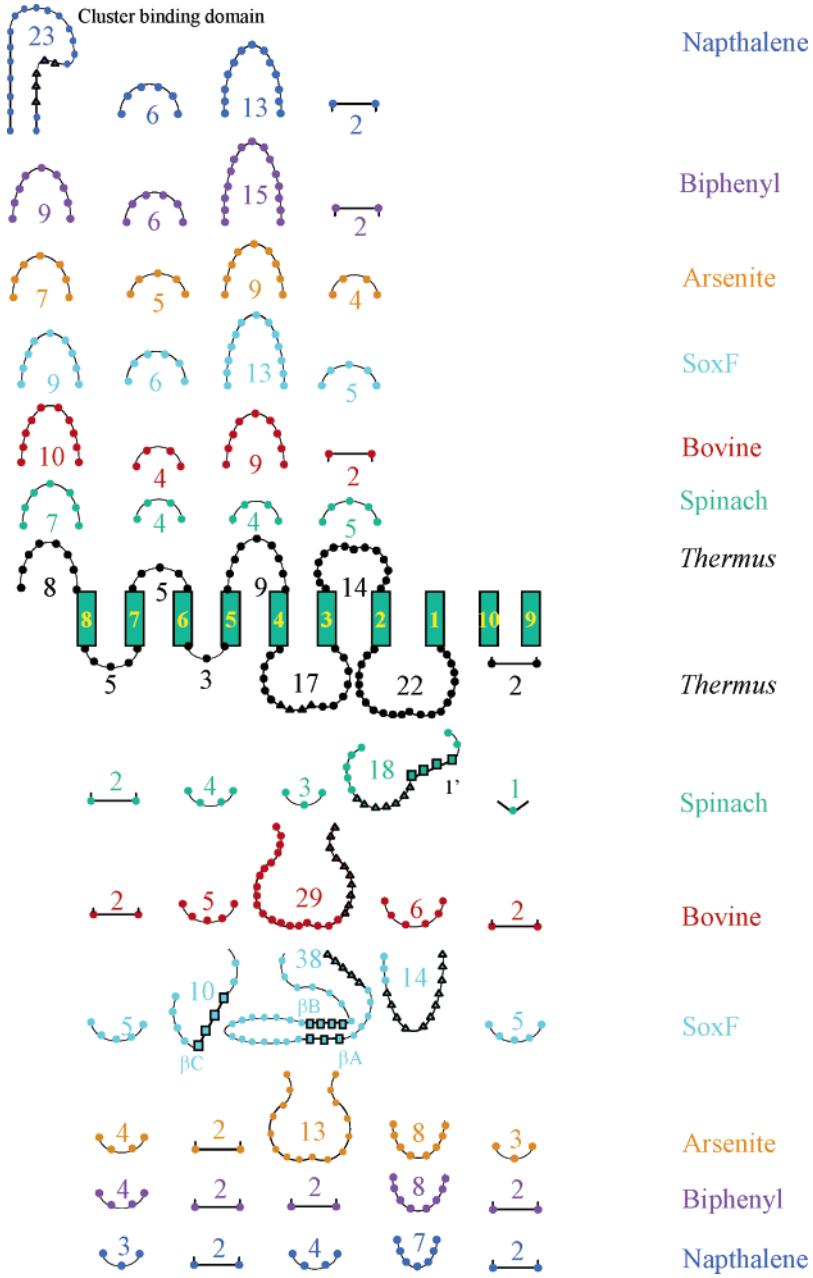

FIGURE 7: Diagram of the different loops connecting the β-strands in seven Rieske proteins. *Thermus* (black), spinach (green), bovine (red), SoxF (cyan), arsenite oxidase (orange), biphenyl dioxygenase (purple), and naphthalene dioxygenase (blue). Squares (■) indicate residues in β-strands, and triangles (▲) show α-helical regions. Loops $\beta4-\beta5$, $\beta6-\beta7$, and $\beta8-\beta9$ link the strands of the cluster binding domain, while loops $\beta4-\beta5$ and $\beta8-\beta9$ link this domain to the large domain. The highly variable $\beta1-\beta2$, $\beta3-\beta4$, and $\beta5-\beta6$ loops are oriented away from the cluster, and in the case of the respiratory proteins, away from the Rieske$-bc_1$ interface.

of the protein is more divergent. The analogous loop joining $\beta3$ to $\beta4$ in the arsenite oxidase Rieske and in naphthalene dioxygenase also points away from the other protein components of their respective complexes (*2*, *9*). It is evident that the portion of the protein that interacts with the larger complex is less variable, whereas the regions that are oriented into the solvent assume divergent charge, length, and geometry (cf. ref *40*).

Domain Interface. The interface between the two domains is largely hydrophobic. This surface appears quite rough, which would influence rotation of the two domains with respect to each other (see Figure 2a). This does not, however, preclude an opening between the domains, as proposed by Iwata et al. (*5*) based on the structures of the whole bc_1 complex.[5] A unique contribution to protein stability within this region may result from an aromatic residue, in most cases from the small domain, within van der Waals distance of another aromatic residue from the large domain. These side chains lie near the center of the hydrophobic patch on the complementary surfaces of the large and small domains and correspond in each protein to Trp-142/Phe-81 in TT, Tyr-188/Phe-97 in BO, Tyr-89/Phe-123 in SP, Phe-130/(Phe-91, Phe-80) in SF, Tyr-56/Phe-76 in AO, and Leu-56/Phe-39 in BD. This interaction is portrayed in the Supporting Information (see SI-VII).

[5] The opening between the c_1 vs the int state in bovine bc_1 is thought to involve trans to cis isomerization of Pro-175 (bovine numbering). In the *Thermus* structure, iso-positional Pro-174 is trans, whereas in the spinach structure, iso-positional Pro-142 is cis.

DISCUSSION

Protonation Status of the Fe2 Ligands. It has been known for some 30 years that the redox potential of the Rieske center is pH dependent (*1*, *41*), and this is now thought to be important in the functioning of the bc_1 complex (see below). In the case of the *Thermus* protein, Hirst and co-workers (*36*) have provided a complete thermodynamic description of redox behavior over the pH range of 3−14. The reduction potential at any pH is given by

$$E_{obs} = E_{alk} - \frac{RT}{F}\ln\left[\left\{\left(1 + \frac{a_{H^+}}{K_{ox2}} + \frac{a_{H^+}^2}{K_{ox1}K_{ox2}}\right)\right\}/\left\{\left(1 + \frac{a_{H^+}}{K_{red2}} + \frac{a_{H^+}^2}{K_{red1}K_{red2}}\right)\right\}\right]$$

where a_{H^+} is the activity of the proton in water, K_{ox1} and K_{ox2} and K_{red1} and K_{red2} are the observed equilibrium constants for the removal of one H^+ and a second H^+ from the Fe2-$(HisH)_2$ elements of the cluster in its oxidized and reduced forms, respectively, and E_{alk} is the pH-independent potential observed at extreme alkalinity. At 20 °C, $pK_{ox1} = 7.85$ and $pK_{ox2} = 9.65$ (*36*) from which one may estimate that under our conditions of crystallization, pH = 8.5, ∼80% of these sites will have lost one H^+ and ∼20% will have lost two protons, corresponding to an average loss of ∼1.1 H^+ per cluster.[6]

We have interpreted our electron density maps in conjunction with two observations of crystal behavior to indicate that each Rieske protein molecule in the unit cell is directly H-bonded to a symmetry related molecule. This extended structure involves apposition of the Nϵ atom of His-134 with the Nϵ atom of His-134 in a symmetry related molecule. In both the A and the B molecules of the asymmetric unit, there is a separation between the 2-fold, symmetry related Nϵ atoms of ∼2.7 Å, which is consistent with a proton being shared between these two atoms. If true, half the clusters in these crystals have lost a proton, a result that would occur if either the actual pH in the crystal is somewhat lower than the buffer pH (8.5 to ∼7.9) or the micro-environment has shifted the pK_{ox1} to higher pH. Such a structure would explain why the crystals dissolve rapidly upon reduction or by lowering the pH to ∼6.5. Both results are explained by protonation of the putative imidazolate of the intermolecular H-bridge.

In the elegant theoretical study of bovine Rieske protein redox behavior by Ullmann et al. (*42*), the Fe−N bond lengths of the Fe−N(His^-) bonds of their gas-phase models were 0.13 Å shorter than those of Fe−N(HisH). In our crystals of *Thermus* Rieske protein, although close to the current error bounds, the apparently shorter (by ∼0.06 to ∼0.09 ± 0.05 Å) Fe2−N_δ (His-134) bond, as compared with the Fe2−N_δ (His-154) bond, is consistent with the Fe2−N_δ (His-134) bond having partial imidazolate character. On the basis of the symmetry constraints and assuming that every His-134, 134′ pair in the crystal shares one proton, only half of the predicted ∼0.1 Å shortening (i.e., ∼0.05 Å) should occur, which is fully consistent with our observations (Table 3). Taken together, these results provide strong evidence that the symmetry related Nϵ atoms of His-134 share a proton.

The combined DFT/MEAD calculations of Ullmann et al. (*42*) with the bovine protein also show that when 0.5 proton (of a possible 2 protons) is removed from the two imidazole rings, the anionic site or proton hole will be distributed on both. This results from the fact that the first observed pK_a is actually composed of two microstate pK_a values that are evaluated in these calculations. The pairs representing the first pK_{obs}^1 ($pK_a^1{}_{12}$ and $pK_a^1{}_{21}$) and the second pK_{obs}^2 ($pK_a^2{}_{12}$ and $pK_a^2{}_{21}$) are very similar, differing by only ∼0.3 p*K* units or ∼0.4 kcal/mol in the protein bound cluster. These theoretical results likely reflect the differences in solvent exposure of the two rings. If so, and because of the near identity of the TT and BO structures in this region, our crystal lattice must preferentially trap the His-134 as the imidazolate ligand in half the molecules, in effect, forcing $pK_a^1{}_{12}$ and $pK_a^1{}_{21}$ to assume different values such that the proton hole becomes localized on His-134 while leaving His-154 protonated. A reasonable estimate for the shift in the microstate pK_a is ∼0.9 kcal/mol, an energy easily covered by the Nϵ−H−Nϵ′ bond. Moreover, it is quite reasonable to suspect that the microstate p*K*s could change within the bc_1 complex during normal function such as to trap a proton on a specific N atom. Such an arrangement of atoms is analogous to that which occurs in the hydrogen maleate anion and may be thought of as a proton trap, as discussed in other contexts by Haines and co-workers (cf. refs *43* and *44*, and references therein).

Rieske Protein Function in bc_1. Crofts and co-workers (*45*−*47*) and Trumpower and co-workers (*48*) have made compelling arguments that the high activation energy step in the bc_1 catalyzed oxidation of quinol by cytochrome *c* is the transfer of one e^- and one H^+ (possibly even as one H°) from quinol bound in the Q_o site to the singly deprotonated form of the oxidized Rieske cluster, residing in its Q_o position. The resulting semiquinone loses its remaining proton to the positive side of the membrane and reduces cytochrome b_L while the reduced, doubly protonated Rieske cluster moves ∼22 Å to its c_1 position where it eventually reduces cytochrome c_1, loses a proton to become the singly deprotonated oxidized form, and returns to its Q_o position. In the crystallographically observed c_1 state of the bc_1 complex (*14*), the Rieske cluster is likely reduced and appears to be stabilized by an intermolecular H-bridge between HNϵ of His-161 (bovine numbering) and a carboxylate O of the heme C propionate side chain, which is within hydrogen bonding distance (*5*, *14*, *45*). In this mechanism, the Q_o position of the Rieske protein would be stabilized by a H-bridge between an OH group of quinol and the localized proton hole on the Nϵ atom of His-161 of the oxidized Rieske cluster. Our results support these types of H-bridging structures. It should be noted that our results do not address the question of the number of quinol molecules bound at the active site, which is the subject of current debate (*1*, *4*, *45*−*47*).

Redox Potentials. Link (*1*) reviewed the extensive literature on the redox properties of Rieske proteins and noted the

[6] Actual evidence for deprotonation of the Fe-bound, histidine-imidazole rings is quite weak, primarily assignment of a pH dependent resonance Raman feature in the *Thermus* protein by Kuila et al. (*18*) and in the sulredoxin of *Sulfolobus* by Iwasaki et al. (*64*) and by DFT calculations by Ullmann et al. (*42*). This is, however, the most chemically reasonable interpretation of the pH dependent redox behavior. Our crystallographic interpretations make a strong addition to the supporting data.

interesting correlation between their acid-pH (defined as pH < pK_{ox1}) redox potentials and their biological function. Rieske proteins involved in plastoquinone mediated photosynthetic respiration have *E* values varying from ~320 to 375 mV, those involved in ubiquinone mediated respiration have *E* varying from ~262 to 312 mV, while those involved in menaquinone mediated respiration have *E* near ~150 mV. The reported potentials for the different quinones ($1/2Q + e^- + H^+ = 1/2QH_2$ as measured chemically) are ~110 mV for plastoquinone (*49*), ~70−90 mV for ubiquinone (*1*, *49*), and ~ −60 mV for menaquinone (*1*). *T. thermophilus* is a menaquinol-utilizing organism (*50*), and its *E* of ~150 mV fits with Link's observation that cells maintain a difference of ≥~200 mV between the Rieske protein and the respective quinone. Indeed, Trumpower and co-workers (*51*) have demonstrated that ubiquinol cytochrome *c* reductase activity decreases as the acid-pH reduction potential of the Rieske protein decreases. Rieske-type proteins have considerably lower potentials, suggesting they function in the dioxygenases to reduce a fairly low-potential electron acceptor.

Past work has indicated that hydrogen bonding to the cluster may be important in determining the redox potential of these (*1*, *10*) and other electron-transfer proteins (*52*−*54*), the rationale being that H-bonds stabilize electron density on the cluster. Figure 8 shows cluster H-bonding patterns in each of five Rieske and Rieske-type structures. They are arranged top to bottom in order of decreasing, acid-pH reduction potential. Criteria used for assignment of an X−H−S hydrogen bond to the cluster are that the distance between donor, X, and acceptor, S, is ≤3.80 Å and that the X−H−S angle be closer to linear than ~120°; distance and angle information for each protein is provided in the Supporting Information (see SI-VIII). The protein from spinach chloroplasts (SP) has six N−H/cluster interactions arising from nearby peptide NH atoms and two O−H interactions with serine and tyrosine side-chain atoms. Not shown here is the structure of the *Sulfolobus* protein, which has an equal number and type of H-bonds as SP and has an acid-pH midpoint potential of ~400 mV (*55*). The bovine mitochrondrial (BO) Rieske protein has an acid-pH, midpoint potential of ~315 mV, and its cluster is H-bonded by five N−H and two O−H groups. The *Thermus* protein (TT) has an acid-pH potential of ~150 mV with H-bonds from six N−H but only one O−H to the cluster. Here, the serine observed in the higher potential proteins is replaced with an iso-positional glycine causing the loss of an O−H interaction but retaining the N−H interaction. The midpoint potential of the arsenite oxidase (AO) has not been reported; it has only five N−H groups bonding to the cluster. In this structure, threonine has replaced the serine, but it cannot form a strong OH−S1 bond because the O to S distance is >3.80 Å. Finally, the biphenyl dioxygenase (BP) has a midpoint potential of ~ −150 mV (*10*), and it has only three N−H groups bonding to the cluster.

There is clearly a rough correlation between the number of H-bonding interactions and redox potentials, as also noticed in other Fe/S systems (*54*, *56*−*58*). Qualitatively consistent with this are previous results in which the redox potential has been altered by mutation of H-donating residues. For example, the relevant serine to alanine at position 157 in the protein from *Paracoccus denitrificans* (*59*) and at position 183 in the protein from yeast (*51*) results

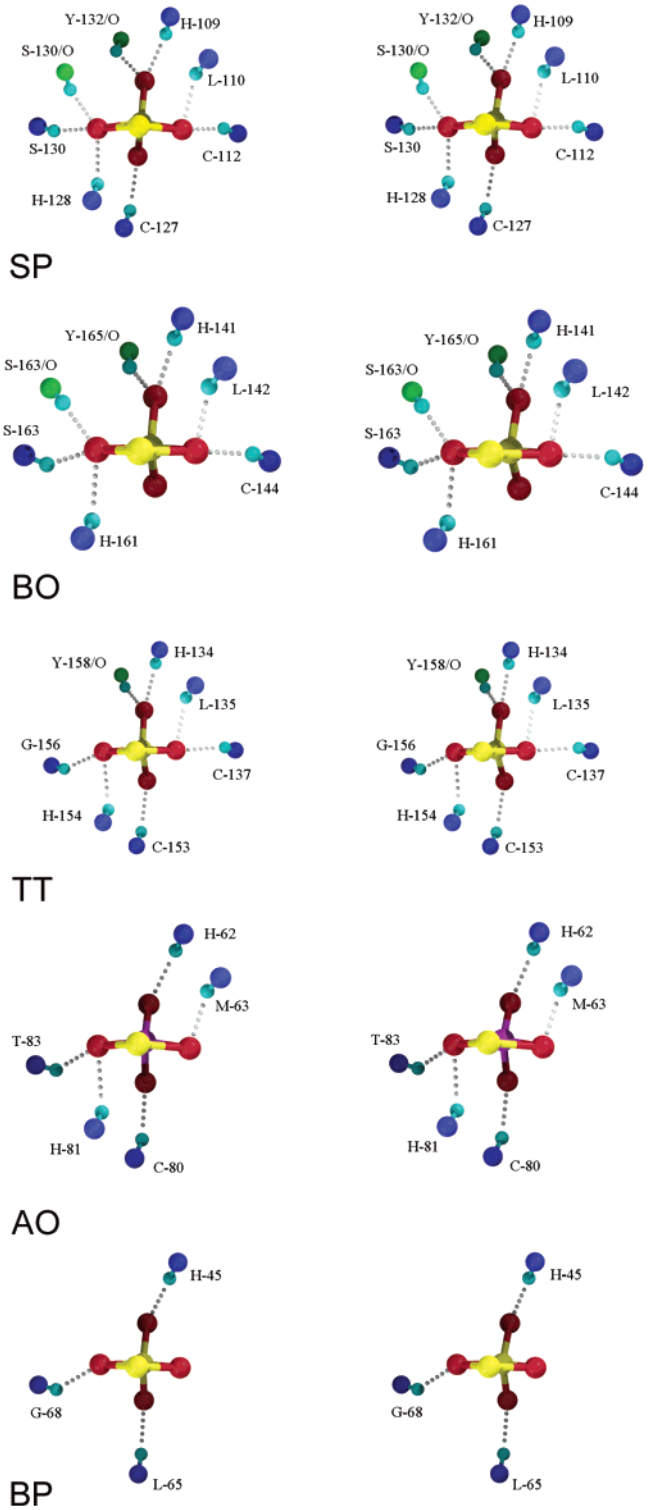

FIGURE 8: Stereo presentation (wall-eye) of cluster H-bonding patterns in Rieske and Rieske-type proteins: Fe, yellow; S, red; O, green; and N, blue. Fe2 is the forward facing Fe atom to which the histidine imidazole rings have been removed, Fe1 is in the rear, S1 is on the left of the cluster, and S2 is on the right. Residue numbers for Sγ (top, bottom) are spinach, SP (107, 125); bovine, BO (139, 158); *Thermus*, TT (132, 151); arsenite oxidase, AO (60, 78); and biphenyl oxygenase, BP (43, 63). Other labels refer to residue numbers in the individual proteins.

in a decrease of 95 and 130 mV, respectively. Notably, the absence of the Ser−OH−S1 H-bond approximately accounts for the lowered reduction potential in menaquinol-utilizing *Tt*. However, there are also quantitative differences between

the individual H-donor interactions among the structures. The N−H−S angles vary from 124 to 165°, the N to S distances vary over the range of 3.10−3.67 Å, and the O to S distances vary from 3.07−3.20 (see Table 2 of the Supporting Information), indicating that the individual H-bonds will have different strengths (see ref *60* for a discussion of H-bond length on redox potential of rubredoxin).

By assuming that each O−H hydrogen bond is nominally twice as strong as each N−H hydrogen bond (cf. ref *61*), a plot of weighted H-bond numbers provides an apparently linear correlation with reduction potential (see Supporting Information SI-IX), the slope of which is ∼70 mV per weighted H-bond. This is consistent with the ΔE observed in the serine to alanine mutants mentioned above (*51, 59*), and it also predicts that the acid-pH, midpoint potential of the arsenite oxidase will be around 0 mV. There is, as should be expected, considerable scatter in such a regression. Notably, removal of the tyrosine OH to Sγ interaction by a Y → F replacement results in a loss of only 44−65 mV in reduction potential (*51, 59*), indicating that single point changes may vary substantially from the ∼70 mV/H-bond average. Indeed, other factors including solvent accessibility have been proposed to affect the redox potential between the Rieske and the Rieske-type proteins (*1, 10*) (i.e., see Table 1 in Supporting Information). Certainly a more rigorous approach would be to calculate the solvation energies of these H-bond arrangements as, for example, described by Bashford and co-workers (*62*).

As all the proteins have a highly similar fold about the cluster, a rationale for why each peptide-to-cluster interaction is lost becomes an important consideration. Loss of the two O−H to cluster S interactions results from residue substitution (as discussed above) and requires no further comment. (Reasonable explanations for the systematic loss of cluster H-bonds in BO and AO are presented in the Supporting Information SI-X.) In general, these changes in the number of hydrogen bonds to the cluster are the result of minor changes in backbone conformations that lead to small movements of the N−H away from the cluster or from rotations of the N−H vector away from the cluster thereby making it a poor candidate to form a hydrogen bond with the cluster.

In the case of BP, however, the loss of the common disulfide bond that tethers the ligand bearing loops together provides an explanation as to why BP lacks H-bond interactions at positions Leu-110, Cys-112, and His-128 (SP numbering). The N−H to S2 interaction between Leu-110 and S2 in SP corresponds to the potential interaction of the N−H of Gly-46 in BP. This H-bonding interaction is lost because the N−H−S2 angle has decreased from 141° in SP to 61° in BP making the N−H point away from the sulfur atom while undergoing very little change in distance from S2 (3.56 Å). The apparent origin of this angle change resides in a greatly different main-chain torsion angle at position Leu-45 in BP.

The N−H−S2 interaction between Cys-112 in SP is potentially present at position Trp-48 in BP, and the His-128 N−H to S1 interaction in SP is potentially at His-66 in BP. However, the peptide N at Trp-48 in BP is 5.49 Å from S2 with the N−H−S2 angle at 92°, and the His-66 N−H to S1 distance has increased to 3.84 Å in BP. The loss of these two interactions in BP are likely caused by two major differences in the BP protein sequence: replacement of the disulfide that bridges the β4,5 and β6,7 loops in each of the other proteins with two bulky, hydrophobic residues, Trp-48 and Leu-65. The distance between the Cα atoms of the cystine in all the other proteins is ∼6.5 Å, whereas in BP the corresponding distance is 7.2 Å, suggesting that the two loops have been pushed away from each other by as much as 0.7 Å. Moreover, BP has a unique salt bridge between Arg-74 and Asp-59 that bridges the β6 and β7 strands and could impact details of the β4,5 and β6,7 loop structure. As noted above, it may also be relevant that BP has a significantly different microenvironment about its partially exposed histidine imidazole rings, which can also affect the observed reduction potential (see Table 1 of the Supporting Information).

There are reasons to suspect that the structural changes that have occurred in BP as compared to SP might also occur in the reduced disulfide form of the TT protein. Recently, Zu et al. (*20*) showed that the disulfide could be reversibly reduced and reoxidized and/or irreversibly reduced and alkylated. In the thiol state, the acid-pH redox potential of the cluster is only diminished by ∼40 mV suggesting that H-bonding to the cluster is largely intact in this state. However, in the SR···RS bis-alkylated state, where the thiol groups have reacted with *N*-β-maleimidopropionate, the acid-pH redox potential has decreased by ∼125 mV. This would be consistent with loss of two peptide N−H interactions with the cluster (see Supporting Information SI-IX) and could occur by way of small structural changes similar to those seen in BP. Thus, an interaction between the bulky alkyl groups in reduced and alkylated TT distorts the peptide structure around the cluster in a manner analogous to that resulting from interaction between the two bulky residues that have replaced the disulfide in BP. One might thereby expect to lose the interaction between the cluster and the peptide N−H groups of Leu-135 and Cys-137 in the TT SR···RS structure.

The absence of the pH dependence on reduction potential of Rieske-type proteins can also be explained by the loss of H-bonds. While fewer hydrogen bonds decrease the stability of the reduced state relative to the oxidized state, as reflected in a lower reduction potential, the presence of fewer hydrogen bonds has the secondary effect of raising the pK_a of the imidazole to imidazolate ionization of the ligands to Fe2 cf. (*20*). Indeed, the analogous pK_{ox1} and pK_{ox2} of the Rieske-type proteins are now known to be shifted out of the physiological range (J. Hirst et al., unpublished).

Evolutionary Implications of Folds and Loops. Close inspection of the available Rieske and Rieske-type protein structures allows for some speculation about how evolution has affected these proteins in their larger complexes. It is likely that evolution will have affected the two subdomains differently (*40*); however, it is clear that the spatial relationship of the two domains has been largely maintained. The cluster binding domain is highly conserved with the only exception being the Pro-loop in naphthalene dioxygenase. The large domain is much more varied and may reflect some of the needs of individual species. For example, in SF, the protein is derived from an acidophile, and the extended loops have been proposed to wrap around the protein as an adaptation to the extreme environment and/or provide interactions in the SoxM super-complex (*11*).

Another perspective in the comparison of these proteins is that they have two faces. Analysis of the loops connecting the β-strands shows a clear distinction between the loops on the face of the protein that interact with the larger complex (top of Figure 7) and the face that is oriented into the solvent (bottom of Figure 7). The size of the loops oriented toward the larger complex is more conserved than on the face that is solvent exposed. This result is expected since mutations that interfere, for example, with the Rieske-cytochrome c_1 or cytochrome *b* interaction may be lethal; a mutation on the face of the Rieske protein that interacts with other proteins within the complex would need to be accompanied by a compensatory mutation to not lose function. By contrast, single mutations that are on the solvent exposed face may be tolerated (cf. ref *40*). A possible exception to the idea that evolution has affected the two faces differently is the extended Pro-loop in naphthalene dioxygenase. However, in the larger $\alpha_3\beta_3$ complex in the crystal structure, the extended Pro-loop actually covers the [2Fe−2S] cluster, providing additional interactions to another subunit (*9*).

ACKNOWLEDGMENT

We would like to thank Vandana Sridhar for first crystallizing the *Thermus* Rieske protein, Dr. David Goodin for valuable discussion, Mr. James Zuberi for helping to make some of the figures in this report, and Drs. Anna Marie Hays, Vidyasankar Sundaresan, Lou Noodleman, and Bernard Trumpower for helpful comments on the manuscript. We also thank the staff at the Stanford Synchrotron Radiation Laboratory (SSRL) for their excellent support in the use of the beam lines. This work is based on research conducted at SSRL, which is funded by the Department of Energy, Office of Basic Energy Sciences. SSRL beam lines are supported by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and by the Department of Energy, Office of Biological and Environmental Research. J.A.F. thanks the UCSD Division of Biology for its continuing support of this work.

SUPPORTING INFORMATION AVAILABLE

Ten sections of supplemental text, including two tables and five figures. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Link, T. A. (1999) *Adv. Inorg. Chem. 47*, 83−157.
2. Ellis, P. J., Conrads, T., Hille, R., and Kuhn, P. (2001) *Structure (Cambridge) 9*, 125−32.
3. Trumpower, B. L., and Gennis, R. (1994) *Annu. Rev. Biochem. 63*, 675−716.
4. Link, T. A. (1997) *FEBS Lett. 412*, 257−64.
5. Iwata, S., Lee, J. W., Okada, K., Lee, J. K., Iwata, M., Rasmussen, B., Link, T. A., Ramaswamy, S., and Jap, B. K. (1998) *Science 281*, 64−71.
6. Darrouzet, E., Moser, C. C., Dutton, P. L., and Daldal, F. (2001) *Trends Biochem. Sci. 26*, 445−51.
7. Gatti, D. L., Tarr, G., Fee, J. A., and Ackerman, S. H. (1998) *J. Bioenerg. Biomembr. 30*, 223−33.
8. Carrell, C. J., Zhang, H., Cramer, W. A., and Smith, J. L. (1997) *Structure 5*, 1613−25.
9. Kauppi, B., Lee, K., Carredano, E., Parales, R. E., Gibson, D. T., Eklund, H., and Ramaswamy, S. (1998) *Structure 6*, 571−86.
10. Colbert, C. L., Couture, M. M., Eltis, L. D., and Bolin, J. T. (2000) *Structure Fold Des. 8*, 1267−78.
11. Bonisch, H., Schmidt, C. L., Schafer, G., and Ladenstein, R. (2002) *J. Mol. Biol. 319*, 791−805.
12. Iwata, S., Saynovits, M., Link, T. A., and Michel, H. (1996) *Structure 4*, 567−79.
13. Gao, X., Wen, X., Yu, C., Esser, L., Tsao, S., Quinn, B., Zhang, L., Yu, L., and Xia, D. (2002) *Biochemistry 41*, 11692−702.
14. Zhang, Z., Huang, L., Shulmeister, V. M., Chi, Y. I., Kim, K. K., Hung, L. W., Crofts, A. R., Berry, E. A., and Kim, S. H. (1998) *Nature 392*, 677−84.
15. Lange, C., and Hunte, C. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 2800−5.
16. Hunte, C., Koepke, J., Lange, C., Rossmanith, T., and Michel, H. (2000) *Structure Fold Des. 8*, 669−84.
17. Samoilova, R. I., Kolling, D., Uzawa, T., Iwasaki, T., Crofts, A. R., and Dikanov, S. A. (2002) *J. Biol. Chem. 277*, 4605−8.
18. Kuila, D., Schoonover, J. R., Dyer, R. B., Batie, C. J., Ballou, D. P., Fee, J. A., and Woodruff, W. H. (1992) *Biochim. Biophys. Acta 1140*, 175−83.
19. Fee, J. A., Findling, K. L., Yoshida, T., Hille, R., Tarr, G. E., Hearshen, D. O., Dunham, W. R., Day, E. P., Kent, T. A., and Munck, E. (1984) *J. Biol. Chem. 259*, 124−33.
20. Zu, Y., Fee, J. A., and Hirst, J. (2002) *Biochemistry 41*, 14054−65.
21. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. D50*, 760−3.
22. Schneider, T. R., and Sheldrick, G. M. (2002) *Acta Crystallogr. D58*, 1772−9.
23. Sheldrick, G. M. (2002) *Z. Kristallogr. 217*, 1−7.
24. Cowtan, K. D., and Main, P. (1996) *Acta Crystallogr. D52*, 43−8.
25. Lamzin, V. S., and Wilson, K. S. (1993) *Acta Crystallogr. D49*, 129−49.
26. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjedgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
27. McRee, D. E. (1999) *Practical Protein Crystallography*, 2nd ed., Academic, San Diego, CA.
28. Sheldrick, G. M., and Schneider, T. R. (1997) *Methods Enzymol. 277b*, 319−43.
29. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−91.
30. Bruns, C. M., Hubatsch, I., Ridderstrom, M., Mannervik, B., and Tainer, J. A. (1999) *J. Mol. Biol. 288*, 427−39.
31. Fraczkiewicz, R., and Braun, W. (1998) *J. Comput. Chem. 19*, 319−33.
32. Guss, J. M., Bartunik, H. D., and Freeman, H. C. (1992) *Acta Crystallogr. B48*, 790−811.
33. Stout, C. D., Stura, E. A., and McRee, D. E. (1998) *J. Mol. Biol. 278*, 629−39.
34. Williams, P. A., Blackburn, N. J., Sanders, D., Bellamy, H., Stura, E. A., Fee, J. A., and McRee, D. E. (1999) *Nature Struct. Biol. 6*, 509−16.
35. Fee, J. A., Castagnetto, J. M., Case, D. A., Noodleman, L., Stout, C. D., and Torres, R. A. (2003) *J. Biol. Inorg. Chem. 8*, 519−26.
36. Zu, Y., Fee, J. A., and Hirst, J. (2001) *J. Am. Chem. Soc. 123*, 9906−7.
37. Choc, M. G., Lorence, R. M., Yoshida, T., Dunham, W. R., and Fee, J. A. (1982) in *Flavins and Flavoproteins* (Massey, V., and Williams, C. H., Eds.) pp 746−750, Elsevier North-Holland, Amsterdam.
38. Gurbiel, R. J., Doan, P. E., Gassner, G. T., Macke, T. J., Case, D. A., Ohnishi, T., Fee, J. A., Ballou, D. P., and Hoffman, B. M. (1996) *Biochemistry 35*, 7834−45.
39. Wilson, E. B., Jr., Decius, J. C., and Cross, P. C. (1955) *Molecular Vibrations: The theory of infrared and Raman vibrational spectra*, Dover, New York.
40. Valdar, W. S. J., and Thornton, J. M. (2001) *PROTEINS: Struct., Funct., Genet. 42*, 108−24.
41. Prince, R. C., and Dutton, P. L. (1976) *FEBS Lett. 65*, 117−9.
42. Ullmann, G. M., Noodleman, L., and Case, D. A. (2002) *J. Biol. Inorg. Chem. 7*, 632−9.
43. Haines, T. H. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 160−4.
44. Haines, T. H., and Dencher, N. A. (2002) *FEBS Lett. 528*, 35−9.
45. Crofts, A. R., Barquera, B., Gennis, R. B., Kuras, R., Guergova-Kuras, M., and Berry, E. A. (1999) *Biochemistry 38*, 15807−26.
46. Crofts, A. R., Hong, S., Ugulava, N., Barquera, B., Gennis, R., Guergova-Kuras, M., and Berry, E. A. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 10021−6.
47. Crofts, A. R., Hong, S., Zhang, Z., and Berry, E. A. (1999) *Biochemistry 38*, 15827−39.
48. Snyder, C. H., Gutierrez-Cirlos, E. B., and Trumpower, B. L. (2000) *J. Biol. Chem. 275*, 13535−41.

49. Rich, P. R. (1984) *Biochim. Biophys. Acta 768*, 53−79.
50. McKay, A., Quilter, J., and Jones, C. W. (1982) *Arch. Microbiol. 131*, 43−50.
51. Denke, E., Merbitz-Zahradnik, T., Hatzfeld, O. M., Snyder, C. H., Link, T. A., and Trumpower, B. L. (1998) *J. Biol. Chem. 273*, 9085−93.
52. Carter, C. W., Jr., Kraut, J., Freer, S. T., Alden, R. A., Sieker, L. C., Adman, E. T., and Jensen, L. H. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 3527−9.
53. Adman, E., Watenpaugh, K. D., and Jensen, L. H. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4854−8.
54. Correll, C. A., Ludwig, M. L., Bruns, C. M., and Karplus, P. A. (1993) *Protein Sci. 2*, 2112−33.
55. Anemuller, S., Schmidt, C. L., Schafer, G., Bill, E., Trautwein, A. X., and Teixeira, M. (1994) *Biochem. Biophys. Res. Commun. 202*, 252−7.
56. Low, D. W., and Hill, M. G. (2000) *J. Am. Chem. Soc. 122*, 11039−40.
57. Chen, K., Jung, Y. S., Bonagura, C. A., Tilley, G. J., Prasad, G. S., Sridhar, V., Armstrong, F. A., Stout, C. D., and Burgess, B. K. (2002) *J. Biol. Chem. 277*, 5603−10.
58. Chen, K., Bonagura, C. A., Tilley, G. J., McEvoy, J. P., Jung, Y. S., Armstrong, F. A., Stout, C. D., and Burgess, B. K. (2002) *Nat. Struct. Biol. 9*, 188−92.
59. Schroter, T., Hatzfeld, O. M., Gemeinhardt, S., Korn, M., Friedrich, T., Ludwig, B., and Link, T. A. (1998) *Eur. J. Biochem. 255*, 100−6.
60. Lin, I.-J., Gebel, E. B., Machonkin, T. E., Westler, W. M., and Markley, J. L. (2003) *J. Am. Chem. Soc. 125*, 1464−5.
61. Kollman, P., McKelvey, J., Johansson, A., and Rothenberg, S. (1975) *J. Am. Chem. Soc. 97*, 955−65.
62. Li, J., Nelson, R., Peng, C. Y., Bashford, D., and Noodleman, L. (1998) *J. Phys. Chem. 102*, 6311−24.
63. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) *Bioinformatics 15*, 305−8.
64. Iwasaki, T., Imai, T., Urushiyama, A., and Oshima, T. (1996) *J. Biol. Chem. 271*, 27659−63.
65. Cosper, N. J., Eby, D. M., Kounosu, A., Kurosawa, N., Neidle, E. L., Kurtz, D. M., Jr., Iwasaki, T., and Scott, R. A. (2002) *Protein Sci. 11*, 2969−2973.

BI0342719