Three-dimensional structure of the high-potential iron-sulfur protein isolated from the purple phototrophic bacterium Rhodocyclus tenuis determined and refined at 1.5 Å resolution

ARTICLE *in* JOURNAL OF MOLECULAR BIOLOGY · DECEMBER 1992

Impact Factor: 4.33 · DOI: 10.1016/0022-2836(92)90849-F · Source: PubMed

J. Mol. Biol. (1992) 228, 672–686

Three-dimensional Structure of the High-potential Iron–Sulfur Protein Isolated from the Purple Phototrophic Bacterium *Rhodocyclus tenuis* Determined and Refined at 1·5 Å Resolution

Ivan Rayment

Institute for Enzyme Research, Graduate School, and Department of Biochemistry
College of Agricultural and Life Sciences
University of Wisconsin, Madison, WI 53705, U.S.A.

Gary Wesenberg

Institute for Enzyme Research, Graduate School
University of Wisconsin, Madison, WI 53705, U.S.A.

Terry E. Meyer, Michael A. Cusanovich

Department of Biochemistry
University of Arizona, Tucson, AZ 85721, U.S.A.

and Hazel M. Holden†

Institute for Enzyme Research, Graduate School, and Department of Biochemistry
University of Wisconsin, Madison, WI 53705, U.S.A.

(Received 15 May 1992; accepted 29 July 1992)

The molecular structure of the high-potential iron–sulfur protein (HiPIP) isolated from the phototrophic bacterium, *Rhodocyclus tenuis*, has been solved and refined to a nominal resolution of 1·5 Å with a crystallographic *R*-factor of 17·3% for all measured X-ray data from 30 Å to 1·5 Å. It is the smallest of the HiPIP structures studied thus far with 62 amino acid residues. Crystals used in the investigation belonged to the space group $P2_1$ with unit cell dimensions of $a = 36{\cdot}7$ Å, $b = 52{\cdot}6$ Å, $c = 27{\cdot}6$ Å and $\beta = 90{\cdot}8°$ and contained two molecules per asymmetric unit. The structure was solved by a combination of multiple isomorphous replacement with two heavy-atom derivatives, anomalous scattering from the iron–sulfur cluster, symmetry averaging and solvent flattening. The folding motif for this HiPIP is characterized by one small α-helix, six Type I turns, an approximate Type II turn and one Type I′ turn. As in other HiPIPs, the iron–sulfur cluster is co-ordinated by four cysteinyl ligands and exhibits a cubane-like motif. These cysteinyl ligands are all located in Type I turns. The hydrogen bonding around the metal cluster in the *R. tenuis* protein is similar to the patterns observed in the *Chromatium vinosum* and *Ectothiorhodospira halophila* HiPIPs. Several of the amino acid residues invariant in the previously determined *C. vinosum* and *E. halophila* structures are not retained in the *R. tenuis* molecule. There are 13 solvent molecules structurally conserved between the two *R. tenuis* HiPIP molecules in the asymmetric unit, some of which are important for stabilizing surface loops. Interestingly, while it is assumed that this HiPIP functions as a monomer in solution, the two molecules in the asymmetric unit pack as a dimer and are related to each other by an approximate twofold rotation axis.

Keywords: protein structure; X-ray crystallography; redox proteins; iron–sulfur centers; electron transport

† Author to whom all correspondence should be addressed.

0022–2836/92/220672–15 $08.00/0

1. Introduction

High-potential iron–sulfur proteins (HiPIPs†) are a class of [4Fe–4S] ferredoxins commonly isolated from purple phototrophic bacteria (Bartsch, 1978). These typically small proteins undergo a reversible one-electron transfer reaction at a characteristically high oxidation–reduction midpoint potential of between +50 and +450 mV and are paramagnetic in the oxidized state (Meyer *et al.*, 1983). As a class of redox proteins, the HiPIPs show considerable variation in their amino acid sequences, net overall charges, and oxidation–reduction potentials (Bartsch, 1978; Tedro *et al.*, 1985*a*,*b*; Meyer *et al.*, 1983). While the HiPIPs are presumably involved in electron transport, their exact metabolic role in the bacterium is unclear at the present time. However, it has been recently demonstrated that a HiPIP from *Thiobacillus ferro-oxidans* functions as an iron-oxidizing enzyme (Kusano *et al.*, 1992).

The [4Fe–4S] cluster, as seen in the HiPIPs, is widely distributed throughout nature and occurs in such important enzymes as aconitase, NADH dehydrogenase, sulfite reductase and nitrogenase (for a review, see Beinert, 1990). Due to this widespread occurrence of the iron–sulfur cluster, the HiPIPs have enjoyed renewed scientific interest within the last several years and have been subjected to various forms of biophysical analyses including X-ray, Mössbauer and n.m.r. spectroscopy.

The first three-dimensional structure of a HiPIP to be analyzed by X-ray crystallography was that isolated from *Chromatium vinosum* (Carter *et al.*, 1971, 1974*b*; Freer *et al.*, 1975). It is the largest in the class, with 85 amino acid residues. The folding motif for the *C. vinosum* HiPIP was shown to be rather simple, consisting of a series of reverse turns that envelop the metal prosthetic group and provide the four cysteinyl ligands to the [4Fe–4S] cluster. Subsequent to this X-ray investigation, it was shown that the bacterial ferredoxin from *Peptococcus aerogenes* also contained [4Fe–4S] centers (Adman *et al.*, 1973, 1976). In contrast to the HiPIPs, however, the bacterial ferredoxins have two such clusters per polypeptide chain, show oxidation–reduction potentials near −400 mV and are paramagnetic in the reduced state. The redox potential of the two equivalent clusters in the *P. aerogenes* ferredoxin is −427 mV (Stombaugh *et al.*, 1976), whereas the potential for the *C. vinosum* HiPIP is +360 mV (Meyer *et al.*, 1983).

When the molecular structures of the reduced *C. vinosum* HiPIP and the oxidized *P. aerogenes* ferredoxin were subsequently compared, the iron–sulfur clusters appeared to be indistinguishable (Carter *et al.*, 1972). In an attempt to reconcile this with the drastic differences in redox potentials and magnetic properties demonstrated by these two proteins, Carter *et al.* (1972) proposed an elegant theory referred to as the "three-state hypothesis". According to this model, three oxidation states are available to the metal cluster with formal charges of −1, −2 and −3. The HiPIPs transfer electrons between the −1 and −2 oxidation states, whereas the bacterial ferredoxins utilize the −2 and −3 oxidation states. In the oxidized form of the HiPIP, there are formally three Fe(III) ions and one Fe(II) ion, whereas in the reduced form of the protein there are two Fe(III) and two Fe(II) ions. From this three-state theory, it was also suggested that the extent of hydrogen bonding to the metal cluster plays a crucial role in the redox potential modulation of these iron–sulfur proteins (Carter *et al.*, 1974*a*; *Adman et al.*, 1975; Carter, 1977). In the *C. vinosum* HiPIP, there are five potential hydrogen bonds to the iron–sulfur center, while in the *P. aerogenes* low potential ferredoxin, there are eight to each cluster.

It has become apparent within the last year, however, that the role of hydrogen bonding in redox modulation of iron–sulfur clusters may not be as important as once thought (Backes *et al.*, 1991). The recent structure determination of the *Ectothiorhodospira halophila* HiPIP in our laboratory has demonstrated that, in this 71-amino acid residue protein, which has a midpoint potential of +120 mV, the hydrogen bonding around the [4Fe–4S] cluster is nearly identical to the pattern observed in the *C. vinosum* HiPIP, despite the low amino acid sequence identity between the two proteins and the presence of several insertions and deletions (Breiter *et al.*, 1991). Consequently, the 240 mV difference in redox potentials between these two HiPIPs cannot be attributed to hydrogen bonding patterns alone, but rather to other factors, such as the number and positions of hydrophobic amino acid residues surrounding the cluster, the solvent exposure of the iron–sulfur center and the electrostatic field of the protein.

In an effort to understand those factors responsible for modulating the redox potentials of the high-potential ferredoxins, we have been studying, by X-ray crystallographic techniques, the molecular structure of another HiPIP, that isolated from *Rhodocyclus tenuis*. This particular HiPIP was selected for study because it is one of the smallest members in its class, with 62 amino acid residues, is a basic protein (both the *C. vinosum* and *E. halophila* HiPIPs are acidic) and has a midpoint oxidation–reduction potential of approximately 30 mV less than the *C. vinosum* HiPIP (Meyer *et al.*, 1983; Tedro *et al.*, 1985*a*). Moreover, based on amino acid sequence alignments, the *R. tenuis* HiPIP is the most divergent of its class in which a number of presumed "invariant" residues are not retained (Tedro *et al.*, 1985*a*). Specifically, the only amino acid residues that are conserved, excluding the four cysteinyl cluster ligands, are Tyr12 and Gly53. Here we describe the initial structure determination and refinement of the *R. tenuis* protein to a nominal resolution of 1·5 Å (1 Å = 10^{-1} nm) and compare its overall structural motif to those

† Abbreviations used: HiPIP, high-potential iron–sulfur protein; n.m.r. nuclear magnetic resonance; EPR, electron paramagnetic resonance.

observed in the *C. vinosum* and *E. halophila* HiPIPs. X-ray co-ordinates for this HiPIP have been deposited in the Brookhaven Protein Data Bank or may be obtained immediately *via* HOLDEN@VMS.MACC.WISC.EDU(INTERNET) or HOLDEN@WISCMACC(BITNET).

2. Materials and Methods

(a) *Crystallization and preparation of heavy atom derivatives*

The *R. tenuis* HiPIP (formerly referred to as *Rhodospirillum tenue*) was isolated and purified according to the method described by Bartsch (1978). Large single crystals were grown as previously described, the only modification being the use of macro-seeding to maintain a reproducible supply of crystals (Holden *et al.*, 1986). All the seeding experiments were conducted using the hanging drop method of vapor diffusion at room temperature and in the presence of atmospheric oxygen. Typically, 20 μl of protein were mixed with 20 μl of 2·4 M-ammonium sulfate, buffered with 50 mM-Na^+/K^+ phosphate, 5 mM-NaN_3 (pH 7·0) and these droplets were suspended over wells containing 2·4 M-ammonium sulfate, 50 mM-Na^+/K^+ phosphate, 5 mM-NaN_3 (pH 7·0). After equilibration for 5 days, small seed crystals, washed in 2·0 M-ammonium sulfate, 50 mM-Na^+/K^+ phosphate (pH 7·0), were then added to the droplets. The protein concentration for these crystallization trials was typically 16 mg/ml. Crystallization was generally complete within 3 weeks with many of the crystals achieving dimensions of 1·1 mm × 0·8 mm × 0·2 mm. The crystals belonged to the space group $P2_1$ with unit cell dimensions of $a = 36·7$ Å, $b = 52·6$ Å, $c = 27·6$ Å and $\beta = 90·8°$. There were two molecules in the asymmetric unit. No attempts were made to keep the protein crystals in the reduced state, although due to the high potential of the *R. tenuis* HiPIP, the reduced form is stable in the presence of atmospheric oxygen.

For heavy atom derivative searches, crystals were transferred to a synthetic mother liquor containing 3·0 M-ammonium sulfate, 50 mM-Na^+/K^+ phosphate, 5 mM-NaN_3 (pH 7·0) and various heavy metal reagents. It proved difficult to prepare isomorphous heavy atom derivatives partly because of the small size of the protein, thereby presenting fewer reactive side chains, and partly because of the lability of the iron–sulfur cluster under the experimental conditions. Most of the mercurial and gold compounds investigated completely bleached the crystals. Even trimethyl-lead acetate, a compound that has proven so effective in our laboratory for other structural investigations, failed to produce any observable changes in the X-ray diffraction pattern (Holden & Rayment, 1991). Only two isomorphous heavy atom derivatives were successfully prepared using 5 mM-$K_2Pt(CN)_4$ or 5 mM-K_2PtCl_4 with 500 mM-NaCl.

(b) *X-ray data collection and processing*

Precession photography was employed for monitoring the potential heavy atom binding to the crystalline protein. Once heavy atom derivatives were identified, the three-dimensional X-ray data were collected at room temperature on a Syntex P3 four-circle diffractometer operated at 45 kV and 35 mA using nickel-filtered copper K_α radiation. Each reflection was measured with an ω-scan at a variable scan rate. Two background measurements were made for each reflection at 0·8° on either side of the peak. Peak intensities were calculated as a sum of the 6 highest steps per reflection. After peak intensity integration, the X-ray data were corrected for absorption, background, radiation damage and Lorentz and polarization effects with a set of programs written by Dr Ivan Rayment. The R_{sym} for symmetry-related reflections was approx. 2% on intensity for the native and the 2 derivative crystals. Only 1 crystal was required for each X-ray data set to 2·8 Å resolution. Based on a set of standard reflections, the intensity decrease due to radiation damage was approx. 4% after 70 h of X-ray exposure. The native X-ray data set was extended to 1·8 Å resolution by use of another crystal. All the initial X-ray data were collected with facilities at the University of Arizona. A complete sphere of X-ray data was measured for the native X-ray data set, whereas only a hemisphere of X-ray data was collected for each heavy atom derivative.

Upon moving to the University of Wisconsin, the X-ray data were recollected at 4 °C with a Siemens X1000D area detector system. The X-ray source was nickel-filtered copper K_α radiation from a Rigaku RU200 X-ray generator operated at 50 kV and 30 mA. Two crystals were required for collection of the native X-ray data to 1·5 Å, whereas only 1 crystal was needed for each of the heavy atom derivative data sets to 2·5 Å. These X-ray data were processed with the data reduction software package XDS (Kabsch, 1988) and internally scaled according to the algorithm of Fox & Holmes (1966) as implemented by Dr Phil Evans. Relevant X-ray data collection statistics for the native crystals may be found in Table 1. The native X-ray data set contained 96% of the total theoretical number of observations to 1·5 Å resolution.

Based on scaling statistics, it was apparent that the anomalous scattering signal from the native crystals was more accurately determined in the X-ray data collected by diffractometry. Consequently, the 2 native X-ray data sets collected at the University of Arizona and Wisconsin were subsequently scaled together by a "local" scaling procedure developed in the laboratory by Drs G. Wesenberg, W. Rypniewski and I. Rayment. With this method, the scale for a particular reflection was computed from the neighboring reflections in a volume defined by a sphere or a rectangular prism. The relative contribution of a reflection to a scale factor was weighted according to the distance of the reflection from that which was to be scaled. This local scaling procedure was especially important because the crystals were flat plates and therefore the X-ray data were subject to large absorption errors. Likewise, the heavy atom derivative X-ray data sets were placed on the same scale as the native X-ray data set with this procedure. Only the heavy atom derivative X-ray data collected with the area detector were used in the subsequent analyses. The average isomorphous R-factors to 2·5 Å resolution, based on amplitudes, between the native and heavy atom derivative X-ray data sets were 32% and 26% for the K_2PtCl_4 and $K_2Pt(CN)_4$ derivatives, respectively.

(c) *Computational methods*

A native anomalous Patterson map was first calculated with X-ray data from 30 Å to 5·0 Å in an attempt to locate the positions of the 2 iron–sulfur clusters within the asymmetric unit. There were 2 large peaks on the Harker section at $v = 1/2$ that corresponded to the centers of mass of the 2 metal prosthetic groups. The appropriate cross-vectors for these sites were also immediately obvious in the Patterson map.

Subsequent analyses of the isomorphous difference

Table 1
Intensity statistics for the native crystal

	Resolution range (Å)						
	∞–5·41	3·83	3·13	2·71	2·42	2·21	2·05
Average intensity	847	912	718	377	277	270	236
Average sigma	48·9	55·8	43·8	26·5	19·2	20·8	18·6
No. of measurements	794	1942	4263	5142	5456	5392	4104
R-factor†, ‡	0·034	0·036	0·036	0·040	0·041	0·043	0·044

	Resolution range (Å)						
	1·92	1·81	1·71	1·63	1·57	1·50	1·45
Average intensity	163	120	91	63	46	33	22
Average sigma	13·2	9·8	8·1	6·4	5·3	4·1	4·4
No. of measurements	3755	3480	3461	3012	2365	1504	375
R-factor†, ‡	0·046	0·048	0·056	0·067	0·079	0·088	0·141

† $R\text{-factor} = \sum |I - \bar{I}| / \sum I$.
‡ For these statistics, the Friedel pairs were not merged.

Patterson maps calculated to 5·0 Å resolution for the K_2PtCl_4 and $K_2Pt(CN)_4$ derivatives revealed that each had only 1 major heavy atom binding site. To place the heavy atom derivatives on a common origin, the K_2PtCl_4 site was set at $y = 0$ and a Fourier map was calculated with the structure factor amplitude differences between the native and the $K_2Pt(CN)_4$ derivative. In the space group $P2_1$, electron density maps calculated with a single site derivative contain a superposition of the correct and incorrect structures. Consequently, the difference Fourier map contained 2 peaks corresponding to the 2 enantiomorphs of the $K_2Pt(CN)_4$ binding site. One site was selected and the positions and occupancies of the K_2PtCl_4 and $K_2Pt(CN)_4$ sites were refined by the program HEAVY (Terwilliger & Eisenberg, 1983) using the origin-removed Patterson-function correlation method with X-ray data from 30·0 Å to 5·0 Å resolution (Rossmann, 1960). The refined heavy atom parameters may be found in Table 2.

Both platinum derivatives were used to calculate protein phases, again with the program HEAVY. Relevant phase calculation statistics may be found in Table 3. An electron density map calculated to 5·0 Å resolution and based on these phases and a figure-of-merit weighting scheme showed two large peaks, at a height of approx. 7σ, which were assumed to be the iron–sulfur clusters. The positions of these peaks agreed with the native anomalous Patterson map. Finally, to ensure that the proper heavy atom constellation had been chosen, these platinum derivatives were also used to phase a native anomalous difference Fourier map to a resolution of 5·0 Å. This map showed large positive peaks at the expected positions for the metal clusters.

The crystal form $P2_1$ can be a very difficult space group to work with when each of the heavy atom derivatives contains single sites. The phases will often be dominated by one of the heavy atom derivatives, leading to an electron density map with the incorrect structure partially superimposed on the correct structure. This was, indeed, the case for the *R. tenuis* electron density map, which was exceedingly noisy at 3·0 Å resolution. Since there were 2 independent molecules within the asymmetric unit, however, it was possible to improve the quality of the protein phases by the technique of molecular averaging. The necessary rotational and translational matrices relating the 2 molecules within the asymmetric unit for molecular averaging were determined and refined by the program MUNCHKINS (developed in the laboratory by Drs G. Wesenberg and I. Rayment and described by Rypniewski *et al.*, 1991). In the "globally" averaged, but not solvent-flattened electron density map, it was possible to model the iron–sulfur cluster and many of the amino acid residues within the near vicinity of the prosthetic group. The electron density map was still not of the quality necessary to fit the entire molecule.

To obtain an electron density map into which a molecular model could be unambiguously built, 2 strategies were subsequently adopted to improve the

Table 2
Refined heavy atom parameters

Derivative	Site no.	Relative occupancy	*x*	*y*	*z*	*B* (Å²)	Location
K_2PtCl_4	1	13·05	0·082	0·000	0·193	15	Between Met6 (Mol I) and Met6 (Mol II)
$K_2Pt(CN)_4$	1	16·21	0·899	0·888	0·833	15	At the interface between Mol I and its symmetry-related partner near carbonyl oxygen of Thr2 and side chain of Lys20
	2	6·81	0·371	0·888	0·275	15	Near carbonyl oxygen of Ala4 and side chain of Arg7 (Mol I)

† *x*, *y* and *z* are the fractional atomic co-ordinates. *B* is the thermal factor and was not refined.

Table 3
Phase calculation statistics

	Resolution range (Å)							
	∞–8·86	5·64	4·42	3·76	3·32	3·01	2·77	2·58
No. of reflections	169	304	387	452	519	568	606	651
Figure of merit	0·60	0·58	0·50	0·43	0·37	0·35	0·37	0·36
Phasing power†								
K_2PtCl_4	2·07	1·94	1·16	0·81	0·93	0·51	0·81	0·61
$K_2Pt(CN)_4$	3·04	2·63	2·00	1·86	1·59	1·78	1·73	1·73

† Phasing power is the ratio of the root-mean-square heavy-atom scattering factor amplitude to the root-mean-square lack of closure error (calculated for acentric reflections).

quality of the protein phases: inclusion of the native anomalous signal and iterative cycles of molecular averaging and solvent flattening. As mentioned above, the anomalous signal in the native X-ray data was sufficient to locate the centers of mass of the 2 iron–sulfur clusters in the asymmetric unit. The native anomalous Patterson function was not of high enough resolution, however, to resolve the positions of the individual atoms in the metal center. Consequently, it was not possible, in the early stages of the investigation, to use the anomalous signal from the native crystals in the phase calculations. Once the positions for the atoms in the metal cluster had been determined from the initial electron density map, the native anomalous signal was subsequently incorporated into the phase calculations by modification of the program HEAVY. This resulted in a significantly improved electron density map.

In addition, since it was expected that the HiPIPs isolated from *E. halophila* and *R. tenuis* would have similar overall molecular shapes, the known *E. halophila* structure was used to create a molecular envelope for the further refinement of the protein phases by iterative molecular averaging and solvent flattening (Bricogne, 1976). The averaging was performed at 3·0 Å resolution for 15 cycles; the initial protein phases, based on the 2 isomorphous derivatives and the native anomalous scattering, were discarded after the 1st cycle. The structure factor weighting algorithm used in the averaging process was of the form:

$$w = \exp -(||F_o| - |F_c||)/|F_o|,$$

where $|F_o|$ was the observed structure factor amplitude and $|F_c|$ was the calculated structure factor amplitude (Rayment, 1983). The final R-factor between the calculated structure factors from the averaged electron density map and the observed X-ray data was 18%.

The quality of the averaged and solvent flattened electron density map to 3·0 Å resolution was sufficiently improved to be able to build a model for approx. 40 out of a total of 62 amino acid residues. This partial model of the "averaged" *R. tenuis* protein was then placed back into the unit cell and subjected to 2 cycles of alternate least squares refinement at 3·0 Å resolution with the computer package TNT (Tronrud *et al.*, 1987) and manual adjustment of the model. The [4Fe–4S] cluster geometry was restrained to that observed for the small molecule structure determination of $Fe_4(NO)_4(\mu_3\text{–}S)_4$ by Chu *et al.*, 1982. At this stage, the R-factor dropped to 34% for all measured X-ray data from 30 Å to 3·0 Å resolution. Although the electron density map, calculated with coefficients of the form $(2F_o - F_c)$ was considerably improved, the positions of amino acid residues 1 to 20 were still not clear for both molecules in the asymmetric unit.

In an attempt to improve the quality of the electron density still further, the protein phases, calculated from the partially refined models in the asymmetric unit, were combined with the heavy atom derivative and native anomalous scattering phases according to the method of Read (1986), who graciously supplied us with the software. These partially refined models were also used to determine new rotational and translational matrices relating one HiPIP to another within the unit cell. Fifteen more cycles of molecular averaging and solvent flattening were performed at 3·0 Å resolution as described above using the molecular envelope determined from the *E. halophila* structure. The resulting averaged electron density map was considerably improved, such that it was possible to trace tentatively all amino acid residues, although certain loop regions were not entirely clear. The complete *R. tenuis* model was subsequently used to generate a new, tighter molecular envelope and the molecular averaging and solvent flattening procedures were performed once again at 3·0 Å resolution for 15 cycles. As before, the initial protein phases were a combination of model, heavy atom derivative and native anomalous phases. The resolution was extended to 2·5 Å in 0·1 Å increments by including the "combined" phases for each new wedge of X-ray data for the 1st cycle, followed by 15 cycles of refinement at the corresponding resolution. The final R-factor between the calculated structure factors from the averaged electron density map and the observed X-ray data was 16% at 2·5 Å resolution.

This final averaged map, calculated to 2·5 Å resolution, was simply spectacular when compared to the original electron density map phased with heavy atom derivatives. It was immediately obvious that amino acid residues 1 and 2 had been incorrectly traced. Corrections to the averaged model were thus made and subsequently placed back into the unit cell for another cycle of least squares refinement. It is important to note that at this stage of the analysis, the amino acid sequence for the *R. tenuis* strain 3761 HiPIP was being used and that several regions of the map seemed somewhat suspect in that the electron density did not agree with the published primary structure. After the next cycle of refinement, the R-factor dropped to 27% for all measured X-ray data from 30 Å to 2·5 Å.

During the following round of manual adjustment of the model, it became obvious that something was wrong. For example, the amino acid residue at 34 was reported to be an alanine for this bacterial strain (Tedro *et al.*, 1979). However, as can be seen from Figure 1, it was clearly a proline residue. At this stage it was decided to calculate an electron density map to a nominal resolution of 1·5 Å in an attempt to resequence the protein. Quite surprisingly, our "X-ray" sequence based on the electron density

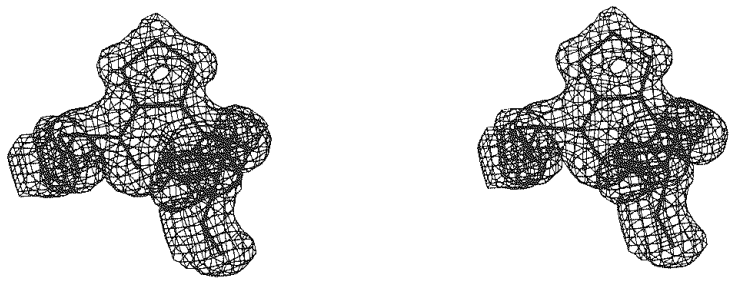

Figure 1. Stereo view of the electron density for amino acid residues Ser33, Pro34 and Thr35. The electron density shown here was contoured at 1σ and calculated with coefficients of the form $(2F_o-F_c)$ to a resolution of 1·5 Å. The original electron density map, calculated with multiple isomorphous replacement phases and subsequent averaging and solvent flattening, was of similar quality. Originally, it was believed that the *R. tenuis* HiPIP from strain 3761 had been crystallized. In strain 3761, however, amino acid residue 34 is an alanine. Clearly, the electron density shown above corresponds to a proline residue. On the basis of the present electron density map calculated to 1·5 Å resolution, it has now been concluded that, in fact, the HiPIP from *R. tenuis* 2761, not strain 3761, was crystallized and that sometime prior to crystallization experiments the strains were inadvertently switched.

map calculated to 1·5 Å, with one ambiguity, matched the amino acid sequence for strain 2761, not strain 3761 (Tedro *et al.*, 1985*a*). The source of the discrepancy is unknown at the present time, but it is possible that these bacterial strains were inadvertently switched prior to our initial crystallization experiments, or that the *R. tenuis* strains are capable of producing similar isozymes (Holden *et al.*, 1986). Recent amino acid sequence analyses of the protein used in this structural investigation agree with the X-ray sequence (van Beeuman, unpublished results).

The *R. tenuis* structure has now been refined to an *R*-factor of 17·3% for all measured X-ray data from 30 Å to 1·5 Å. Small peaks of electron density observed in maps calculated with $(2F_o-F_c)$ coefficients and contoured at a height of 1σ were modeled as solvent molecules if they were within 3·8 Å of potential hydrogen bonding groups. Both the positions and the temperature factors for all atoms were refined and 143 solvent molecules were included in the refinement. Forty-five of these putative solvent molecules had temperature factors below 30 Å^2. The average temperature factor of the solvent was 37·7 Å^2. Plots of the mean main-chain temperature factors *versus* amino acid residue number for each molecule in the asymmetric unit are shown in Fig. 2(a) and (b). Relevant refinement statistics may be found in Table 4. In the last 3 cycles of refinement and model building, the iron–sulfur cluster geometries were not restrained.

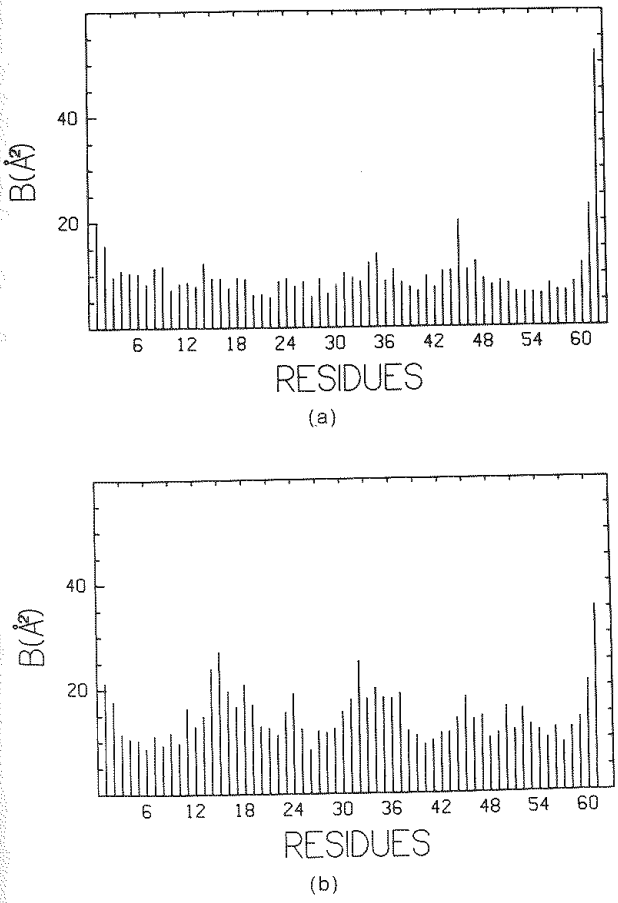

Figure 2. Plot of the mean *B*-value *versus* amino acid residue for all main-chain atoms. (a) The mean *B*-value for all main-chain atoms is shown for Mol I in the asymmetric unit. (b) The mean *B*-value for all main-chain atoms is shown for Mol II in the asymmetric unit. Figure 3 shows the positions of Mol I and Mol II in the unit cell.

3. Results and Discussion

A packing diagram of the four HiPIP molecules within the $P2_1$ unit cell is given in Figure 3. These HiPIPs pack in the crystalline lattice such that there are ten unique hydrogen bonds with lengths

Table 4
Refinement statistics

Resolution limits (Å)	30·0–1·5
Initial *R*-factor (%)†	32·8
Final *R*-factor (%)	17·3
No. of reflections used	16,648
Completeness of X-ray data set (%)‡	96
No. of protein atoms	893
No. of solvent molecules	143
Weighted root-mean-square deviations from ideality	
Bond lengths (Å)	0·014
Bond angle (deg.)	2·063
Planarity (trigonal) (Å)	0·005
Planarity (other planes) (Å)	0·008
Torsion angle (deg.)§	17·415

† *R*-factor $= \sum |F_o-F_c| / \sum |F_o|$.

‡ Percent of measured observations to total theoretical observations.

§ The torsion angles were not restrained during the refinement.

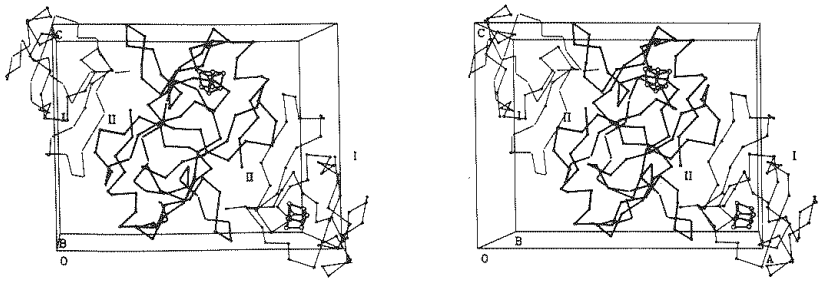

Figure 3. Packing diagram of the *R. tenuis* HiPIP $P2_1$ unit cell. The packing of the 4 HiPIP molecules in the unit cell is shown. The 2 HiPIPs shown with their iron–sulfur clusters correspond to the molecules that were used in the averaging and solvent flattening procedures as described in Materials and Methods. In the text, the HiPIP towards the outer edge of the unit cell is referred to as Mol I, while the HiPIP located in the interior of the unit cell is referred to as Mol II.

less than 3·0 Å formed between symmetry-related molecules. The closest interaction, 2·4 Å, occurs between NZ of Lys8 and a symmetry-related carbonyl oxygen of Lys61. For the purpose of the following discussion, Mol I, as labeled in Figure 3, will refer to the molecule situated at the edge of the unit cell, and Mol II will refer to the molecule located in the unit cell interior. The models for the *R. tenuis* HiPIPs have now been refined to an *R*-factor of 17·3% using all measured X-ray data from 30·0 Å to 1·5 Å resolution. Representative portions of the electron density map, calculated to a nominal resolution of 1·5 Å with coefficients of the form $(2F_o - F_c)$, may be found in Figure 4(a), (b) and (c). As can be seen, the electron density is very well ordered. It is particularly noteworthy that all aro-

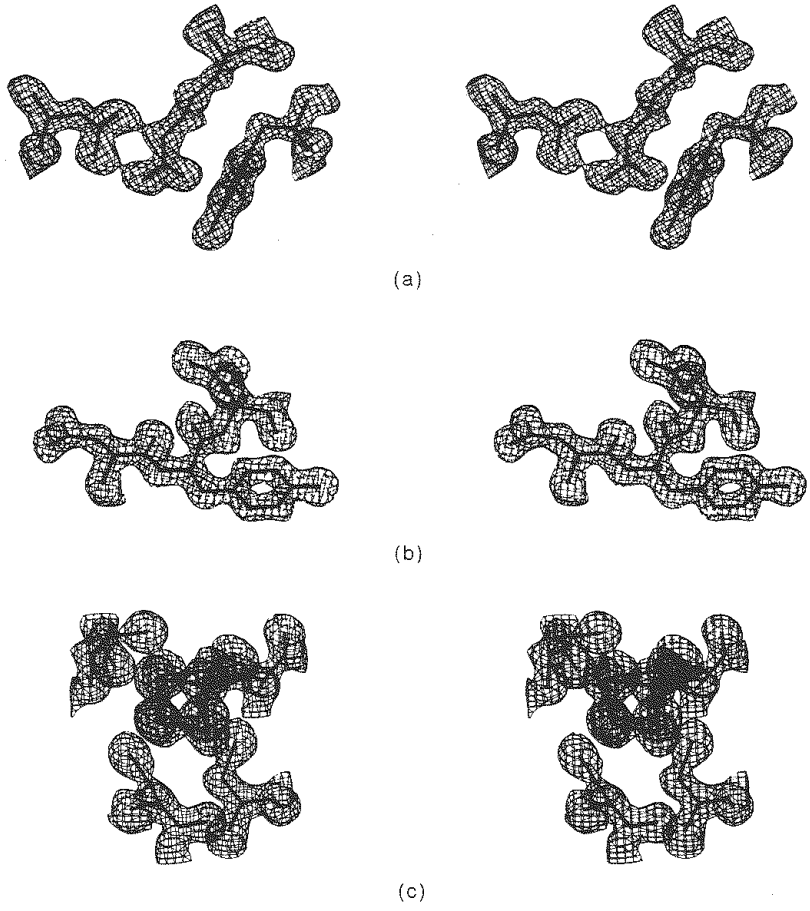

Figure 4. Representative portions of the electron density map calculated to 1·5 Å resolution. Portions of the electron density map shown here were contoured at 1σ and calculated with coefficients of the form $(2F_o - F_c)$. (a) The electron density shown here corresponds to Asp46, Arg7 and Tyr12, Mol II. Note the close interactions between the side chains of Asp46 and Arg7 with OD1 and OD2 of Asp46 being 2·6 Å and 2·9 Å from NE and NH2 of Arg7, respectively. (b) Shown here is the electron density for amino acid residues Asn11, Tyr12 and Gln13, Mol I. Tyr12 is one of two absolutely conserved amino acids found in the *R. tenuis*, *E. halophila* and *C. vinosum* HiPIPs. (c) The electron density for the [4Fe–4S] cluster and the four cysteinyl ligands (Mol I) is shown here.

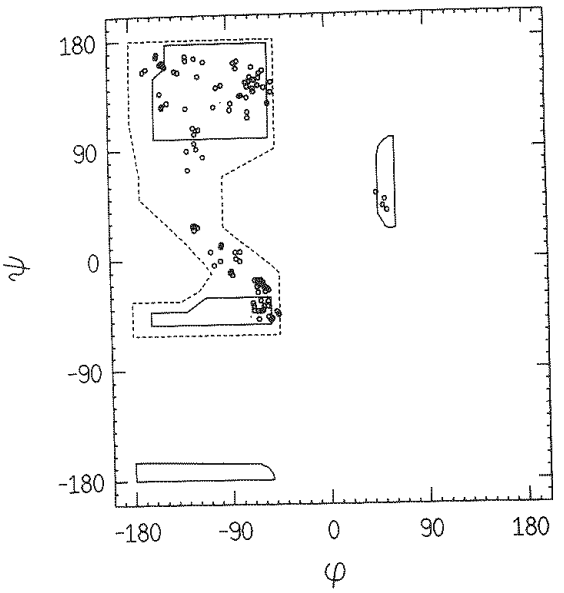

Figure 5. A Ramachandran plot of all non-glycinyl main chain dihedral angles for the two independent HiPIP molecules in the unit cell. Fully allowed ϕ, ψ values are enclosed by solid lines; those only partially allowed are enclosed by dashed lines.

matic side-chain densities display holes in their six-membered rings at a contour level of 1σ. Only the two C-terminal lysine residues for both Mol I and Mol II and the side-chain densities beyond CE for Lys8, Lys17 and Lys41 in Mol I and for Lys17 and Lys21 in Mol II are not well defined.

The electron density was of such quality that it proved possible to resequence the *R. tenuis* HiPIP and to demonstrate that, in fact, we had crystallized a protein corresponding to the amino acid sequence determined for *R. tenuis* strain 2761, not strain 3761 as originally reported (Holden *et al.*, 1986). There is only one region in the electron density map that is in disagreement with the published amino acid sequence for the HiPIP from *R. tenuis* strain 2761 (Tedro *et al.*, 1985*a*). Rather than containing 61 amino acid residues as originally reported, there are 62 amino acid residues in this HiPIP, with an extra lysine residue inserted between Lys20 and Cys21.

Table 5

List of reverse turns

Amino acid residues	Type	ϕ_2	ψ_2	ϕ_3	ψ_3
Mol I, 17–20	I′	54·1	38·1	85·6	−5·0
Mol I, 22–25	I	−60·5	−19·0	−106·3	12·0
Mol I, 25–28	I	−66·8	−18·6	−87·2	−10·6
Mol I, 33–36	I	−62·9	−16·3	−104·2	−5·4
Mol I, 40–43	I	−64·3	−28·0	−87·8	−10·6
Mol I, 43–46	II	−61·5	147·1	72·6	23·6
Mol I, 50–53	I	−59·7	−25·9	−95·1	6·9
Mol I, 55–58	I	−59·9	−23·3	−80·8	6·2
Mol II, 17–20	I′	44·1	52·0	90·0	−13·2
Mol II, 22–25	I	−62·4	−18·6	−110·9	11·4
Mol II, 25–28	I	−64·9	−21·4	−88·3	−13·3
Mol II, 33–36	I	−52·7	−37·4	−97·2	−1·5
Mol II, 40–43	I	−55·4	−26·2	−107·2	6·6
Mol II, 43–46	II	−66·8	142·1	72·9	11·0
Mol II, 50–53	I	−58·1	−23·8	−101·5	14·7
Mol II, 55–58	I	−62·6	−22·1	−82·3	−0·2

A plot of all main chain non-glycinyl dihedral angles for both molecules in the asymmetric unit is given in Figure 5. The only amino acid residue lying outside of the allowed regions is Asn18 in Mol II with dihedral angles of $\phi = 43{\cdot}9°$ and $\psi = 53{\cdot}0°$. These values, however, are nearly within the allowed areas. Excluding the two C-terminal lysine residues, the α-carbons for the two HiPIP molecules in the asymmetric unit superimpose with a root-mean-square value of 0·36 Å. An α-carbon trace for Mol I is shown in Figure 6. As in other HiPIPs, the molecule shows very little secondary structure in the form of α-helices or β-pleated sheet, but rather 48% of the amino acid residues are involved in tight turns that allow the polypeptide chain to wrap around the metal cluster prosthetic group. As listed in Table 5, the *R. tenuis* HiPIP is composed of six Type I turns and one Type II turn. There is also a reverse turn, delineated by amino acid residues 17 to 20, in which the dihedral angles most nearly approximate those observed for Type I′ turns. In theory, however, Type I′ turns require glycine

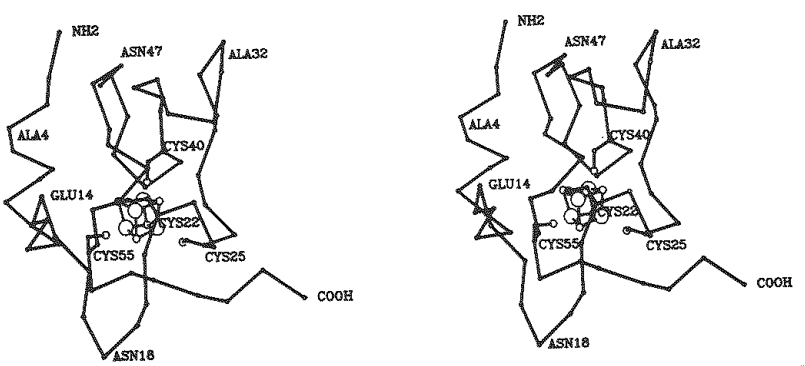

Figure 6. Stereo view of the *R. tenuis* HiPIP backbone. The positions of the α-carbons for Mol I are shown together with the metal cluster and the 4 cysteinyl ligand residues. Amino acid residues are labeled at various positions to aid the reader in following the course of the polypeptide chain. The *R. tenuis* HiPIP is a small protein with overall dimensions of approx. 24 Å × 24 Å × 28 Å.

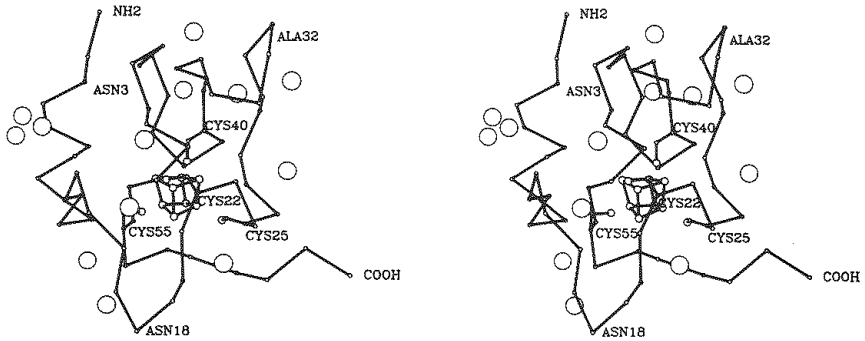

Figure 7. Stereo view of the *R. tenuis* HiPIP with structurally conserved solvent molecules. As represented by the open circles, there are 13 solvent molecules that are located in structurally identical positions in the 2 HiPIP molecules in the asymmetric unit. No solvent molecules are located closer than approx. 7 Å to the metal cluster.

residues at both positions 2 and 3. In the *R. tenuis* HiPIP Type I′ turn, however, positions 2 and 3 are occupied by an asparagine and glycine residue, respectively. The only other major secondary structural element found in the *R. tenuis* HiPIP is a rather short helix composed of amino acid residues 4 to 9. For the most part, the hydrogen bonding capacity of all atoms in the *R. tenuis* molecule capable of such electrostatic interactions is satisfied. The only exceptions to this are the amide nitrogens of Asp14, Thr15 and Ser33 in Mol I, the carbonyl oxygen of Val42 in Mol I, the amide nitrogens of Thr15 and Ser33 in Mol II, the carbonyl oxygens of Tyr12, Ala32, Thr35 and Val42 in Mol II and the side chain NE2 of Gln48, which apparently do not participate in hydrogen bonding interactions.

There are six amino acid residues whose side chains adopt quite different conformations between the two HiPIP molecules in the asymmetric unit (Phe10, Asp14, Gln27, Val42, Ile59 and Val60). All of these perturbations can be attributed to molecular or crystalline contacts. Phe10 and Val42 in Mol I are located in a region of the molecule that forms close interactions with the same area of Mol II. There is not enough room in this region for both phenylalanine residues in Mol I and Mol II to adopt identical conformations. This same situation arises for amino acid residues Gln27, Ile59 and Val60. Again, the regions defined by these residues in Mol I and Mol II are in close contact to one another. With respect to Asp14, it resides in a surface loop. In Mol I, OD1 of Asp14 is within hydrogen bonding distance to the side-chain hydroxyls of Thr35 and Ser33 in Mol II. Also, OG1 of Thr15 in Mol I is within hydrogen bonding distance to the NZ of Lys8, Mol II. In fact, it is this interaction that causes the side-chain density for Lys8 in Mol II to be better defined than in Mol I. These hydrogen bonding interactions, available to Asp14 and Thr15 in Mol I, are not possible in Mol II due to crystal packing. Other than these six amino acid residues, the two HiPIP molecules in the asymmetric unit are very similar and superimpose with a root-mean-square value of 0·51 Å for all atoms, excluding the above-mentioned side chains and the two C-terminal lysines.

There are 13 solvent molecules that are structurally conserved between the two HiPIP molecules in the asymmetric unit, as shown in Figure 7. A list of these solvent molecules and their locations is given in Table 6. They are all located towards the surface of the protein and for the most part serve to bridge either backbone carbonyl oxygens with neighboring backbone amide nitrogens or backbone carbonyl oxygens with other neighboring backbone carbonyl oxygens. As can be seen from Figure 7, there are no solvent molecules located within the near vicinity of the metal cluster binding site. The closest solvent molecule to the metal prosthetic group is positioned approximately 6·9 Å from S2 of the [4Fe–4S] cluster. In each HiPIP molecule in the

Table 6

List of solvent molecules common to both HiPIPs in the asymmetric unit

Solvent name (Mol I)	Solvent name (Mol II)	Groups located within hydrogen bonding distance
03	024	OD1 of Asp56 and carbonyl oxygen of Ala9
08	054	Carbonyl oxygen of Gly52 and amide nitrogen of Ala61
010	057	Carbonyl oxygen of Gly19 and amide nitrogen of Lys21
012	094	Carbonyl oxygen of Cys25 and carbonyl oxygen of Phe28
016	060	Amide nitrogen of Gly31, carbonyl oxygen of Ser33 and carbonyl oxygen of Ala36
017	142	OH of Tyr54 and OD2 of Asp56
018	025	Amide nitrogen of Cys40, carbonyl oxygen of Cys40, carbonyl oxygen of Ile43 and carbonyl oxygen of Asp46
027	139	NH1 of Arg7 and solvent 113
031	138	Carbonyl oxygen of Ala4 and solvent 113
033	068	Amide nitrogen of Gly52
038	062	Carbonyl oxygen of Gly31 and carbonyl oxygen of Ser33
106	119	Carbonyl oxygen of Asn47 and carbonyl oxygen of Ala37
113	140	Solvent 031 and solvent 027

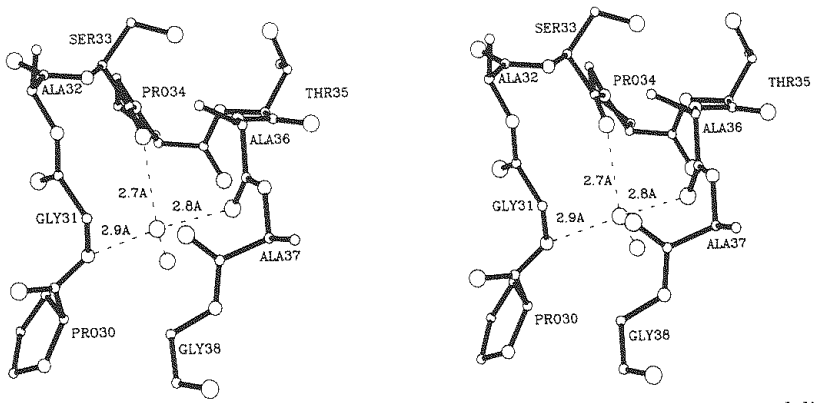

Figure 8. Stereo view of 2 solvent molecules involved in the stabilization of the surface loop delineated by amino acid residues Pro30 to Gly38. The solvent molecules are depicted as open circles. As indicated by the dashed lines, one of these solvent molecules is within hydrogen bonding distance of the backbone nitrogen of Gly31, the carbonyl oxygen of Ser33 and the carbonyl oxygen of Ala36 (Mol I). This solvent is also 2·6 Å from a second water molecule located towards the lower right-hand side.

asymmetric unit, there is a conserved array of three solvent molecules forming a network linking the carbonyl oxygen of Ala4 to NH1 of Arg7. Also, as can be seen in Figure 8, there are two solvent molecules located in the Type I loop region delineated by amino acids Ser33 to Ala36, which may play a role in stabilizing this secondary structural element. One of the most well-ordered solvent molecules in the *R. tenuis* HiPIP is located in the Type I, Type II turn motif defined by amino acid residues Cys40 to Asp46. The *B*-value for this solvent molecule is 11·3 $Å^2$ and 15·5 $Å^2$ for Mol I and Mol II, respectively.

Comparisons of the *R. tenuis* backbone structure with the *E. halophila* and the *C. vinosum* HiPIPs are shown in Figure 9(a) and (b), respectively. These superpositions were generated according to the algorithm of Rossmann & Argos (1975). X-ray coordinates for the *C. vinosum* HiPIP were obtained from the Brookhaven Protein Data Bank (Bernstein *et al.*, 1977). The *R. tenuis* protein is considerably smaller than the *E. halophila* and *C. vinosum* HiPIPs and most probably represents close to the near minimum amount of polypeptide chain necessary to surround a [4Fe–4S] cluster and to maintain the proper oxidation states. The small helical region defined by amino acid residues Ala4 to Ala9 in the *R. tenuis* HiPIP is also observed in the *C. vinosum* protein, but not in the *E. halophila* molecule. Consequently, while the *E. halophila* HiPIP lies between the *C. vinosum* and *R. tenuis* HiPIPs with respect to the number of amino acid residues, it is not a structural intermediate between them. Of particular interest is the surface loop composed of amino acid residues Pro30 to Ala36 which, in the *R. tenuis* protein, adopts a quite different conformation as compared to the *E. halophila* and *C. vinosum* HiPIPs. Considering the positions of α carbons only, the *R. tenuis* HiPIP superimposes on the *E. halophila* and the *C. vinosum* proteins with a root-mean-square value of 0·73 Å for 40 equivalent amino acid residues and of 0·87 Å for 45 equivalent amino acid residues, respectively. Thus, the *R. tenuis* protein is approximately 65% structurally equivalent to the *E. halophila* HiPIP and 72% structurally equivalent to the *C. vinosum* molecule. For comparison purposes, the *E. halophila* and *C. vinosum* HiPIPs are 56% equivalent (Breiter *et al.*, 1991).

At the amino acid sequence level, there are only two residues, Tyr12 and Gly53 (*R. tenuis* numbering) in addition to the four cysteinyl ligands, that are invariant among all three HiPIPs whose structures are now known. These amino acid residues are located near one another with the hydroxyl group of Tyr12 being 4·4 Å from the α carbon of Gly53. Because of this close approach, any side chain larger than glycine at position 53 could not be easily accommodated in this region. If, as suggested by Carter *et al.* (1974), Tyr12 plays a critical role in the modulation of oxidation–reduction potentials in these high-potential ferredoxins, then the conservation of a tyrosine residue at this position and orientation necessarily requires a glycine residue at position 53 for proper folding and functioning of the HiPIP molecule. Interestingly, neither Tyr12 nor Gly53 are conserved in the *Rhodospirillum salinarum* iso-2 HiPIP, which is known to be oxidation labile (Meyer *et al.*, 1990; Ambler, personal communication).

In the *C. vinosum* HiPIP, the hydroxyl oxygen of this conserved tyrosine is 3·0 Å from the amide nitrogen of Asn72 and 2·7 Å from the solvent molecule designated H140. This type of hydrogen bonding pattern around the conserved tyrosine side chain is similar in both the *R. tenuis* and *E. halophila* HiPIPs, except that in these proteins, the solvent molecule is structurally replaced with the side chain OD1 donated by Asn14 or Asp14. In the *E. halophila* HiPIP, the hydroxyl group of Tyr12 is 2·9 Å from the amide nitrogen of Lys59 and 2·7 Å from OD1 of Asn14, while in the *R. tenuis* HiPIP (Mol II), the hydroxyl group of Tyr12 is 3·1 Å from the amide nitrogen of Ala50 and 3·1 Å from OD1 of Asp14. The hydrogen bonding pattern around the side chain of Tyr12 in Mol I is slightly

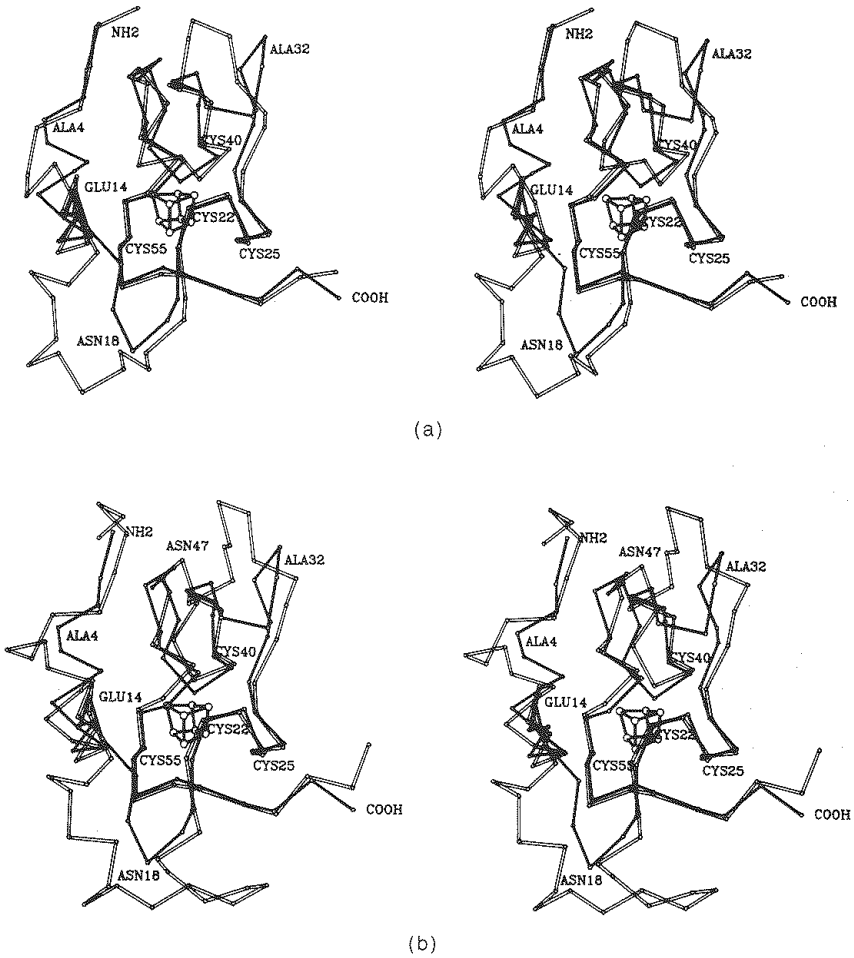

Figure 9. Superposition of the *R. tenuis* HiPIP (Mol I) polypeptide chain backbone onto the *E. halophila* and the *C. vinosum* proteins. These superpositions were generated according to the algorithm of Rossmann & Argos (1975). The *R. tenuis* HiPIP is shown in solid lines; the *E. halophila* and *C. vinosum* molecules are depicted as open bonds. The metal center belongs to the *R. tenuis* protein. Those amino acid residues that are labeled correspond to the *R. tenuis* HiPIP. (a) Shown is the superposition of the *R. tenuis* and *E. halophila* HiPIPs. (b) The three-dimensional overlap of the *R. tenuis* and *C. vinosum* HiPIPs is shown. X-ray co-ordinates for the *C. vinosum* HiPIP were obtained from the Brookhaven Protein Data Bank.

different due to crystal packing as described above. In this case, the side chain of Asp14 is pulled out of the immediate region surrounding Tyr12 because of crystal contacts. As a consequence, the side chain hydroxyl group of Tyr12 is still within hydrogen bonding distance of the amide nitrogen of Ala50, but is no longer close to Asp14. Instead, the tyrosine hydroxyl oxygen is 2·8 Å from a solvent molecule that has filled in the region vacated by Asp14.

A close-up view of the *R. tenuis* [4Fe–4S] cluster with those amino acid residues located approximately 4·0 Å from the metal center is displayed in Figure 10(a). The cluster is covalently attached to the protein *via* cysteinyl ligands Cys22, Cys25, Cys40 and Cys55. Average temperature factors for atoms in the [4Fe–4S] cluster are 7·1 Å^2 and 10·5 Å^2 for Mol I and Mol II, respectively. Bond lengths for the [4Fe–4S] cluster are listed in Table 7. As in other HiPIPs studied thus far, the binding pocket is decidedly hydrophobic with four aromatic residues, Tyr12, Phe28, Tyr54 and Phe58, surrounding one side of the metal center. For comparison purposes, the binding pockets for the metal clusters in the *E. halophila* and *C. vinosum* HiPIPs are shown in Figure 10(b) and (c). As can be seen, there are two

Table 7
List of bond lengths for the two HiPIP [4Fe–4S] clusters in the asymmetric unit

Atoms in the bond		Bond lengths (Å)	
		Mol I	Mol II
Fe1	S1	2·31	2·30
Fe1	S2	2·28	2·33
Fe1	S3	2·16	2·14
Fe2	S1	2·31	2·29
Fe2	S2	2·28	2·34
Fe2	S4	2·18	2·24
Fe3	S2	2·30	2·30
Fe3	S3	2·25	2·26
Fe3	S4	2·33	2·40
Fe4	S1	2·34	2·32
Fe4	S3	2·27	2·25
Fe4	S4	2·27	2·35

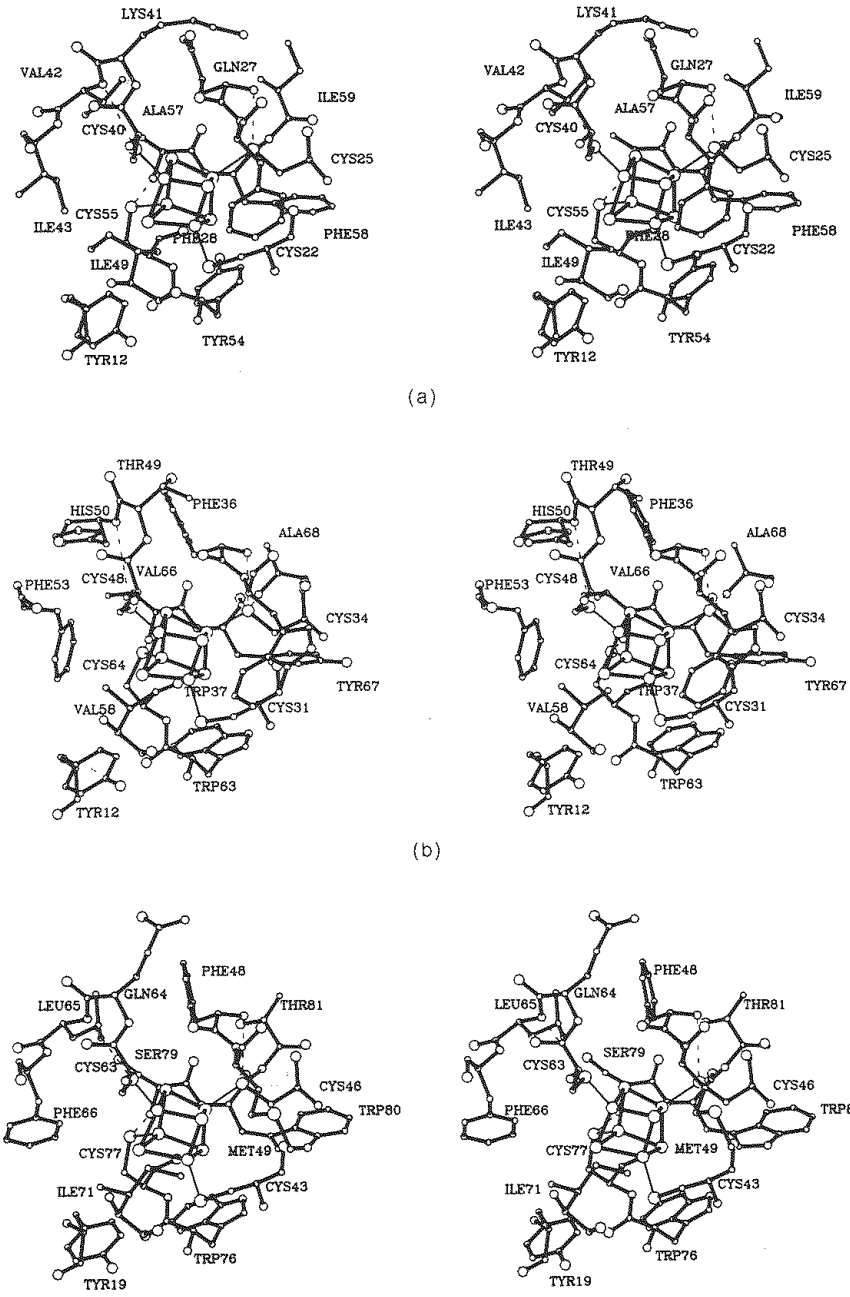

Figure 10. Close-up views of the *R. tenuis*, *E. halophila* and *C. vinosum* [4Fe–4S] cluster binding pockets. Only those amino acid residues within approx. 4·0 Å of the atoms in the metal cluster are shown. Binding pockets for the *R. tenuis*, *E. halophila* and *C. vinosum* HiPIPs are shown in (a), (b) and (c), respectively. Dashed lines indicate potential hydrogen bonds to sulfur atoms in the metal cluster.

quite striking differences between the *E. halophila* and *C. vinosum* HiPIPs and the *R. tenuis* molecule. In the former two proteins, there are phenylalanines at positions 36 and 53 (*E. halophila* numbering), whereas in the *R. tenuis* HiPIP, these amino acids have been replaced with glutamine and isoleucine residues, respectively. Also, in both the *E. halophila* and *C. vinosum* HiPIPs, amino acid residue 63 (*E. halophila* numbering) is a tryptophan, while in the *R. tenuis* HiPIP, it is a tyrosine residue. There are six aromatic residues surrounding the cluster in the *E. halophila* protein, five such residues near the *C. vinosum* cluster and only four located in the binding pocket for the *R. tenuis* HiPIP. All three of the HiPIPs show very similar hydrogen bonding patterns to sulfurs in the redox center. Lengths for these putative hydrogen bonds are given in Table 8.

It had been previously assumed, based on various biophysical characterizations, that HiPIPs function as monomers. Recent isolation of a HiPIP from the purple phototrophic bacterium, *R. salinarum*, however, has shown that at least in this particular halophile, the protein may exist as possibly a tetramer or hexamer (Meyer *et al.*, 1990; Ambler,

Table 8

Potential hydrogen bonds involving sulfur atoms in R. tenuis, E. halophila *and* C. vinosum *HiPIPs*

Atoms involved in H-bonds	Distance (Å)			
	R. tenuis			
	Mol I	Mol II	*E. halophila*†	*C. vinosum*‡
SG of Cys25 and N of Ile59	3·5	3·7	3·6	3·7
SG of Cys25 and N of Gln27	3·4	3·4	3·4	3·7
SG of Cys40 and N of Val42	3·4	3·5	3·8	3·5
SG of Cys55 and N of Ala57	3·5	3·1	3·5	3·5
S1 of the cluster and N of Cys55	3·4	3·5	3·6	3·7

† These distances correspond to the structurally similar amino acid residues in the *E. halophila* HiPIP (Cys34, Ala68, Phe36, Cys48, His50, Cys64, Val66).

‡ Distances were measured from the X-ray co-ordinates obtained from the Brookhaven Protein Data Bank and correspond to the structurally similar amino acid residues in the *C. vinosum* HiPIP (Cys46, Thr81, Phe48, Cys63, Leu65, Cys77, Ser79).

personal communication). Likewise, the EPR studies of Dunham *et al.* (1991) have demonstrated that freezing the *C. vinosum* HiPIP in the presence of 0·1 to 2·0 M-NaCl induces aggregation of the monomers to dimers. These investigations led to a proposed model for a HiPIP dimer in which the iron–sulfur clusters are separated by approximately 14 Å.

Since there are two independent *R. tenuis* HiPIP molecules in the asymmetric unit, it was possible that they might also have aggregated to form a dimeric species. Subsequent careful examination of the crystalline packing arrangement, indeed, revealed that the two HiPIP molecules in the asymmetric unit were related by an approximate local twofold rotation axis. As mentioned earlier, there are several amino acid residues that adopt slightly different conformations in the two independent HiPIPs in the asymmetric unit and four of these amino acid residues (Phe10, Gln27, Ile59 and Val60) are located within the monomer : monomer interface as displayed in Figure 11. The two *R. tenuis* monomers are packed together in a fashion similar to that predicted from the model-building studies with the *C. vinosum* HiPIP (Dunham *et al.*, 1991). As suggested from their insightful model, one of the apices of the [4Fe–4S] cubane (namely S4) is pointed towards the nearby surface of the neighboring HiPIP monomer. For the *R. tenuis* HiPIP dimer, the iron–sulfur clusters are approximately 10·9 Å apart in the dimer as measured from S4. Re-examination of the packing arrangement of the *E. halophila* HiPIP molecules in the asymmetric unit also revealed that this iron–sulfur protein packs as a dimer with a similar subunit : subunit interface (Holden, unpublished results). In the case of the *E. halophila* HiPIP, the iron–sulfur clusters are approximately 13·4 Å apart. Clearly, all three of these HiPIPs are capable of aggregating to form dimers with approximate twofold rotation axes. Whether these dimers are an artifact of crystallization or freezing or, in fact, represent the true

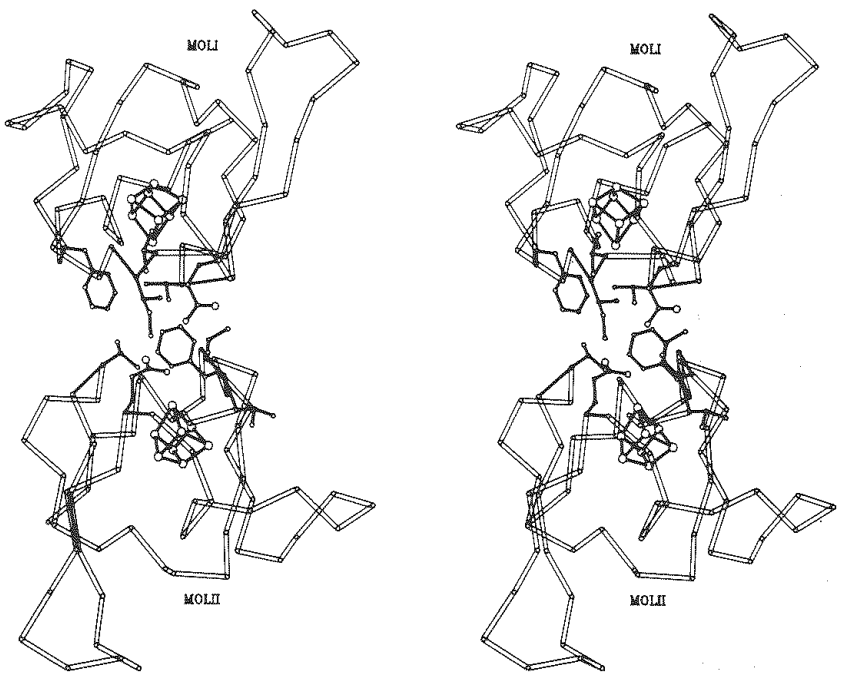

Figure 11. Stereo view of the HiPIP dimer in the asymmetric unit. The *R. tenuis* HiPIPs pack in the asymmetric unit as a dimer. The "local" twofold rotation axis relating Mol I to Mol II is approximately perpendicular to the plane of the paper and runs through the center of the subunit : subunit interface. Those amino acid residues located at the dimer interface and adopting slightly different conformations between the 2 subunits are depicted in filled bonds.

quaternary state of the functional HiPIP, remains to be determined.

In summary, the molecular structure of the *R. tenuis* HiPIP has been determined on the basis of an electron density map calculated to 1·5 Å resolution. As in other HiPIPs, the polypeptide chain wraps around the metal center by a series of reverse turns. The hydrogen bonding pattern around the cluster is very similar to that observed in other HiPIPs whose three-dimensional structures are known. Consequently, the number of hydrogen bonds to the metal center is not a determinant of oxidation–reduction potentials for this class of ferredoxins. In all three known HiPIP structures, the metal center is not exposed to the solvent and there are no well-ordered solvent molecules within the cluster binding pocket. The *R. tenuis* protein aggregates in the crystalline lattice as a dimer, which may have implications as to the quaternary structure of the protein *in vivo*.

We thank Dr Matt Benning for his help in the preparation of the figures. This research was supported in part by grants from the NIH (GM39082 to H.M.H. and GM21277 to M.A.C.). H.M.H. is an Established Investigator of the American Heart Association.

References

Adman, E. T., Sieker, L. C. & Jensen, L. H. (1973). The structure of a bacterial ferredoxin. *J. Biol. Chem.* **248**, 3987–3996.

Adman, L., Watenpaugh, K. D. & Jensen, L. H. (1975). NH···S hydrogen bonds in *Peptococcus aerogenes* ferredoxin, *Clostridium pasteurianum* rubredoxin, and *Chromatium* high potential iron protein. *Proc. Nat. Acad. Sci., U.S.A.* **72**, 4854–4858.

Adman, E. T., Sieker, L. C. & Jensen, L. H. (1976). Structure of *Peptococcus aerogenes* ferredoxin. *J. Biol. Chem.* **251**, 3801–3806.

Backes, G., Mino, Y., Loehr, T. M., Meyer, T. E., Cusanovich, M. A., Sweeney, W. V., Adman, E. T. & Sanders-Loehr, J. (1991). The environment of Fe_4S_4 clusters in ferredoxins and high-potential iron proteins. New information from X-ray crystallography and resonance Raman spectroscopy. *J. Amer. Chem. Soc.*, **113**, 2055–2064.

Bartsch, R. G. (1978). Purification of $[4Fe\text{–}4S]^{1-2-}$ ferredoxins (high-potential iron–sulfur proteins) from bacteria. *Methods Enzymol.* **53**, 329–340.

Beinert, H. (1990). Recent developments in the field of iron–sulfur proteins. *Fed. Amer. Soc. Expt. Biol.* **4**, 2483–2491.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr, Brice, M. D., Rogers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). The protein data bank: a computer based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.

Breiter, D. R., Meyer, T. E., Rayment, I. & Holden, H. M. (1991). The molecular structure of the high potential iron–sulfur protein isolated from *Ectothiorhodospira halophila* determined at 2·5 Å resolution. *J. Biol. Chem.* **266**, 18660–18667.

Bricogne, G. (1976). Methods and programs for direct space exploitation of geometric redundancies. *Acta Crystallogr. sect. A*, **32**, 832–847.

Carter, C. W., Jr (1977). New stereochemical analogies between iron–sulfur electron transport proteins. *J. Biol. Chem.* **252**, 7802–7811.

Carter, C. W., Jr, Freer, S. T., Xuong, N.-H., Alden, R. A. & Kraut, J. (1971). Structure of the iron–sulfur cluster in the *Chromatium* iron protein at 2·25 Å resolution. *Cold Spring Harbor Symp. Quant. Biol.* **36**, 381–385.

Carter, C. W., Jr, Kraut, J., Freer, S. T., Alden, R. A. Sieker, L. C., Adman, E. & Jensen, L. H. (1972). A comparison of the $Fe_4S_4^*$ clusters in high-potential iron protein and in ferredoxin. *Proc. Nat. Acad. Sci., U.S.A.* **69**, 3526–3529.

Carter, C. W., Jr, Kraut, J., Freer, S. T. & Alden, R. A. (1974*a*). Comparison of oxidation-reduction site geometries in oxidized and reduced *Chromatium* high potential iron protein and oxidized *Peptococcus aerogenes* ferredoxin. *J. Biol. Chem.* **249**, 6339–6346.

Carter, C. W., Jr, Kraut, J., Freer, S. T., Xuong, N.-H., Alden, R. A. & Bartsch, R. G. (1974*b*). Two-angstom crystal structure of oxidized *Chromatium* high potential iron protein. *J. Biol. Chem.* **249**, 4212–4225.

Chu, C. T.-W., Lo, F. Y.-K. & Dahl, L. F. (1982). Synthesis and stereochemical analysis of the $[Fe_4(NO)_4(\mu_3\text{–}S)_4]^n$ series ($n = 0, -1$) which possesses a cubane-like Fe_4S_4 core: direct evidence for the antibonding tetrametal character of the unpaired electron upon a one-electron reduction of a completely bonding tetrahedral metal cluster. *J. Amer. Chem. Soc.* **104**, 3409–3422.

Dunham, W. R., Hagen, W. R., Fee, J. A., Sands, R. H., Dunbar, J. B. & Humblet, C. (1991). An investigation of *Chromatium vinosum* high-potential iron–sulfur protein by EPR and Mössbauer spectroscopy; evidence for a freezing-induced dimerization in NaCl solutions. *Biochim. Biophys. Acta*, **1079**, 253–262.

Fox, G. C. & Homes, K. C. (1966). An alternative method of solving the layer scaling equations of Hamilton, Rollet and Sparks. *Acta Crystallogr.* **20**, 886–891.

Freer, S. T., Alden, R. A., Carter, C. W., Jr & Kraut, J. (1975). Crystallographic structure refinement of *Chromatium* high potential iron protein at two Angstroms resolution. *J. Biol. Chem.* **250**, 46–54.

Holden, H. M. & Rayment, I. (1991). Trimethyllead acetate: a first-choice heavy atom derivative for protein crystallography. *Arch. Biochem. Biophys.* **291**, 187–194.

Holden, H. M., Meyer, T. E., Cusanovich, M. A. & Rayment, I. (1986). Crystallization and preliminary analysis of crystals of high potential iron–sulfur protein from *Rhodospirillum tenue*. *J. Biol. Chem.* **261**, 4219–4220.

Kabsch, W. (1988). Automatic indexing of rotation diffraction patterns. *J. Appl. Crystallogr.* **21**, 67–71.

Kusano, T., Takeshima, T., Sugawara, K., Inoue, C., Shiratori, T., Yano, T., Fukumori, Y. & Yamanaka, T. (1992). Molecular cloning of the gene encoding *Thiobacillus ferrooxidans* iron(II) oxidase. High homology of the gene product with HiPIP. *J. Biol. Chem.* **267**, in the press.

Meyer, T. E., Przysiecki, C. T., Watkins, J. A., Bhattacharyya, A., Simondsen, R. P., Cusanovich, M. A. & Tollin, G. (1983). Correlation between rate constant for reduction and redox potential as a basis for systematic investigation of reaction mechanisms of electron transfer proteins. *Proc. Nat. Acad. Sci., U.S.A.* **80**, 6740–6744.

Meyer, T. E., Fitch, J., Bartsch, R. G., Tollin, D. & Cusanovich, M. A. (1990). Unusual high redox poten-

tial ferredoxins and soluble cytochromes from the moderately halophilic purple phototrophic bacterium *Rhodospirillum salinarum*. *Biochim. Biophys. Acta*, **1017**, 118–124.

Rayment, I. (1983). Molecular replacement method at low resolution: optimum strategy and intrinsic limitations as determined by calculations on icosahedral virus models. *Acta Crystallogr. sect. A*, **39**, 102–116.

Read, R. J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. sect. A*, **42**, 140–149.

Rossmann, M. G. (1960). The accurate determination of the position and shape of heavy-atom replacement groups in proteins. *Acta Crystallogr.* **13**, 221–226.

Rossmann, M. G. & Argos, P. (1975). A comparison of the heme binding pocket in globins and cytochrome b_5. *J. Biol. Chem.* **250**, 7525–7532.

Rypniewski, W. R., Breiter, D. R., Benning, M. M., Wesenberg, G., Oh, B.-H., Markley, J. L., Rayment, I. & Holden, H. M. (1991). Crystallization and structure determination to 2.5 Å resolution of the oxidized [2Fe–2S] ferredoxin isolated from *Anabaena* 7120. *Biochemistry*, **30**, 4126–4131.

Stombaugh, N. A., Sundquist, J. E., Burris, R. H. & Orme-Johnson, W. H. (1976). Oxidation-reduction properties of several low potential iron–sulfur proteins and of methylviologen. *Biochemistry*, **15**, 2633–2641.

Tedro, S. M., Meyer, T. E. & Kamen, M. D. (1979). Primary structure of a high potential four-iron–sulfur ferredoxin from the photosynthetic bacterium *Rhodospirillum tenue*. *J. Biol. Chem.* **254**, 1495–1500.

Tedro, S. M., Meyer, T. E. & Kamen, M. D. (1985*a*). The amino acid sequence of a high-redox-potential ferredoxin from the purple phototrophic bacterium *Rhodospirillum tenue* strain 2761. *Arch. Biochem. Biophys.* **239**, 94–101.

Tedro, S. M., Meyer, T. E. & Kamen, M. D. (1985*b*). Amino acid sequence of high-redox-potential ferredoxin isozymes from the extremely halophilic purple phototrophic bacterium *Ectothiorhodospira halophila*. *Arch. Biochem. Biophys.* **241**, 656–664.

Terwilliger, T. C. & Eisenberg, D. (1983). Unbiased three-dimensional refinement of heavy-atom parameters by correlation of origin-removed Patterson functions. *Acta Crystallogr. sect. A*, **39**, 813–817.

Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. (1987). An efficient general purpose least-squares refinement program for macromolecular structures. *Acta Crystallogr. sect. A*, **43**, 489–501.

Edited by W. A. Hendrickson