# Hybrid-Cluster Protein (HCP) from *Desulfovibrio vulgaris* (Hildenborough) at 1.6 Å Resolution[†,‡]

S. J. Cooper,[§] C. D. Garner,[§] W. R. Hagen,[⊥] P. F. Lindley,[#] and S. Bailey*[,||]

*School of Chemistry, Nottingham University, University Park, Nottingham, NG7 2RD, U.K., CLRC Daresbury Laboratory, Daresbury, Warrington, Cheshire, WA4 4AD, U.K., Delft University of Technology, Kluyver Department of Biotechnology, Julianalaan 67, 2628 BC Delft, The Netherlands, and European Synchrotron Radiation Facility, BP 220, F-38043 Grenoble, France*

*Received June 28, 2000; Revised Manuscript Received September 18, 2000*

ABSTRACT: The three-dimensional structure of the hybrid cluster protein from *Desulfovibrio vulgaris* (Hildenborough) has been determined at 1.6 Å resolution using synchrotron X-ray radiation. The protein can be divided into three domains: an N-terminal mainly α-helical domain and two similar domains comprising a central $\beta$-sheet flanked by α-helices. The protein contains two 4Fe clusters with an edge-to-edge distance of 10.9 Å. Four cysteine residues at the N-terminus of the protein are ligands to the iron atoms of a conventional [4Fe-4S] cubane cluster. The second cluster has an unusual asymmetric structure and has been named the hybrid cluster to reflect the variety of protein ligands, namely two $\mu$-sulfido bridges, two $\mu_2$-oxo bridges, and a further disordered bridging ligand. Anomalous differences in data collected at 1.488 Å and close to the iron edge at 1.743 Å have been used to confirm the identity of the metal and sulfur atoms. The hybrid cluster is buried in the center of the protein, but is accessible through a large hydrophobic cavity that runs the length of domain 3. Hydrophobic channels have previously been identified as access routes to the active centers in redox enzymes with gaseous substrates. The hybrid cluster is also accessible by a hydrophilic channel. The [4Fe-4S] cubane cluster is close to an indentation on the surface of the protein and can also be approached on the opposite side by a long solvent channel. At the present time, neither the significance of these channels nor, indeed, the function of the hybrid cluster protein is known.

The hybrid-cluster protein (HCP,[1] formerly named "prismane protein") contains two clusters, one of which is a hybrid of an iron-sulfide substructure and an iron-oxo substructure, and the other a 4Fe-4S cubane cluster (*1*). The name "prismane protein" was abandoned when it became clear from the initial crystallographic study that the HCP does not contain the postulated 6Fe cluster. Other authors have also referred to HCP as fuscoredoxin (*2*, *3*). HCP was initially purified from a species of the strictly anaerobic sulfite reducing bacteria *Desulfovibrio: D. vulgaris* (Hildenborough), (*4*, *5*) and *D. desulfuricans* (ATCC 27774) (*6*, *7*). This soluble cytoplasmic protein has been extensively studied because of the unusual properties of its redox-active iron−sulfur clusters.

The initial EPR spectroscopic analysis indicated that HCP had unprecedented redox chemistry and a 6Fe-6S cluster was proposed, with four stable oxidation states in a potential window of 0.5 V. These initial intriguing findings led to further biochemical, structural, and spectroscopic work. The structural work identified that the protein contained two 4Fe clusters: a cubane 4Fe-4S cluster and a novel hybrid cluster. The EXAFS, MCD, EPR, and Mössbauer data were reinterpreted and found to be readily compatible with the crystal structure (*1*). The hybrid cluster has four oxidation states, ranging from the most oxidized state containing four $Fe^{3+}$ to the fully reduced state with three $Fe^{2+}$ and one $Fe^{3+}$. It behaves magnetically as a single entity resulting from the exchange coupling of four high-spin $Fe^{3+}/Fe^{2+}$ ions, with system spins $S = 0$, $^9/_2$, 0 (and 4), $^1/_2$, respectively, for the four redox states. The EPR signal of the cubane cluster is very unusual and difficult to detect from the native HCP. EPR measurements on a *D. desulfuricans* mutant protein (Cys428Ser, equivalent to Cys434 in the *D. vulgaris* protein) lacking the hybrid cluster gave a signal which has only previously been observed for the $[4Fe\text{-}4S]^{1+}$ cubane core in synthetic compounds and is referred to as the "spin-admixed state" (*8*). Thus, the HCP protein appears to be the first biological example of a spin admixed iron−sulfur cluster.

Despite the wealth of spectroscopic and structural information on HCP, the biological function of this protein is enigmatic. However, genes encoding hybrid cluster proteins are widespread among bacterial and archaeal species. ORFs with a very high similarity to the *Desulfovibrio hcp* gene have recently been detected in more than 15 bacterial and

[†] We gratefully acknowledge the support of the Biotechnology and Biological Research Council for providing a fellowship (to S.J.C.) for this work, Grant reference 34/B08005.

[‡] Protein Data Bank deposition codes are 1E1D and 1E2U.

* To whom correspondence should be addressed.

[§] The University of Nottingham.

[⊥] Delft University of Technology.

[||] CLRC Daresbury Laboratory.

[#] European Synchrotron Radiation Facility.

[1] Abbreviations: CODH, carbon monoxide dehydrogenase; EPR, electron paramagnetic resonance; EXAFS, extended X-ray absorption fine structure, HCP, hybrid-cluster protein; MCD, magnetic circular dichroism; MES, 2-[*N*-morpholino]ethanesulfonic acid; MIR, multiple isomorphous replacement; ORF, open reading frame; PEG, poly(ethylene glycol).

10.1021/bi001483m CCC: $19.00 
Published on Web 11/11/2000

archaeal species, ranging from 40 to 66% identity (*8*). All residues that serve as ligands to the hybrid cluster of *D. vulgaris* HCP are strictly conserved in the polypeptides encoded by these similar ORFs. Furthermore, all these polypeptides have an N-terminal cluster of cysteine residues that serve as ligands for the [4Fe-4S] cluster in the *D. vulgaris* protein (*1*). HCP genes have not been found within the genomes of *Bacillus subtilis*, *Pyrococcus horikoshii*, *Archeoglobus fulgidus*, and in several other bacterial species with small genomes.

Differences in the primary structures allow the distinction of three classes of putative HCPs (*8*). Genes for class 1 HCPs have been found only in strictly anaerobic bacteria and in the methanogenic archaeon *Methanococcus jannaschii*. A distinctive feature of class 1 HCPs is the spacing of the N-terminal cysteine ligands for the cubane cluster: Cys-$X_2$-Cys-$X_{7-8}$-Cys-$X_5$-Cys. This is different from typical cubane binding motifs where the final ligand is donated from the C-terminal end of the molecule (*9*). Genes for class 2 HCPs are found in facultative anaerobic Gram-negative bacteria and the spacing of the N-terminal cysteines is Cys-$X_2$-Cys-$X_{11}$-Cys-$X_6$-Cys. Class 3 HCPs are found in (hyper)-thermophilic bacteria and archaea. Class 3 proteins have the same spacing of the N-terminal cysteines as class 1 proteins but have a deletion of 116 amino acids downstream of the N-terminal cysteine cluster. This deletion corresponds to one of the two three-helix bundles in domain 1 of *D. vulgaris* HCP. Figure 1 shows the alignment of the primary structure of representatives of the three classes and the conservation of ligands to both Fe−S clusters.

The only relevant sequence similarity of HCPs with other protein sequences in any of the databases is with CO-oxidizing subunits of carbon monoxide dehydrogenases (CODHs) of several species (*8*). Some 25−30% of residues are conserved between the CO-oxidizing subunits of CODH and HCP domains 2 and 3. These include most of the hybrid cluster ligands. Although CODHs also contain nickel, the sequence conservation may indicate some similarity in global folding and the structure of a catalytic Fe−S cluster.

In facultative anaerobes, the gene encoding HCP occurs within an operon along with a gene encoding an NADH oxidoreductase (*hcr*); the *hcp* gene precedes the *hcr* gene. In the presence of NADH, the isolated *E. coli* NADH oxidoreductase catalyses the reduction of the *E. coli* hybrid-cluster protein. Midpoint potentials allow transfer of electrons from NADH, via the [2Fe-2S] cluster in NADH oxidoreductase and 4Fe-4S cluster of HCP, to the Fe−S−O hybrid-cluster of HCP (*8*). Thus NADH oxidoreductase is probably the natural electron donor for the hybrid cluster protein.

The induction of HCP-accumulation in *E. coli* and *M. morganii* under anaerobic conditions in the presence of nitrate or nitrite suggests that HCP might be involved in anaerobic nitrate/nitrite respiration (*8*). The only other redox enzymes whose synthesis is positively regulated by nitrate and/or nitrite and anaerobic growth are nitrate reductases, nitrite reductases, and formate dehydrogenase-N (*10*). However, no physiological effect has been demonstrated in *E. coli* or *M. morganii* mutant strains in which the *hcp* gene has been inactivated.

Following from the preliminary crystallographic studies (*1*) the present paper gives a more comprehensive description of the structure of the protein. Previous interpretations regarding the prosthetic groups are confirmed using multiple wavelength anomalous difference data. The present structural model has several distinct features that should have a bearing on the still elusive biological function(s) of this unique protein.

## MATERIALS AND METHODS

*Data Collection and Structure Solution.* The HCP was isolated from a *D. vulgaris* clone overproducing the protein from recombinant plasmid pJS104 (*5*, *11*). Crystals were grown at 277 K either in hanging drops or under paraffin oil (*12*). The precipitant solution contained 0.1 M MES, pH 5.9, 0.2 M magnesium acetate, and 25−30% PEG 8000. The precipitant solution was mixed with an equal volume (3−5 $\mu$L) of a 12 mg/mL solution of the protein in 5 mM Tris HCl, pH 8.0. For hanging drop conditions, a reservoir solution of equal volumes (200 $\mu$L) of precipitant and water was used. Initially, a brown precipitant formed within the drop from which straw-colored rectangular crystals appeared after 4−5 days. The structure was determined by MIR methods, using four heavy-atom derivatives, viz. uranyl acetate, *p*-mercuribenzoic acid, trimethyl lead acetate, and potassium tetrachloroplatinate. Data for the room temperature native protein and four heavy atom derivatives were collected on stations 9.5 or 9.6 at the synchrotron radiation source (SRS), CLRC Daresbury Laboratory (U.K.), see Table 1, and processed using the program MOSFLM (*13*). Heavy-atom sites were located and refined, and phases were calculated using programs from the CCP4 suite (*14*); phasing details are presented in Table 2. Solvent flattening was performed using the density modification program DM (*15*). The initial model was built using these experimental phases followed by several rounds of phase recombination.

Subsequently, data were collected from a crystal frozen at 100 K on station 7.2 at the SRS. Prior to cryocooling, crystals were transferred to a solution of 0.05 M MES, 0.1 M Mg-acetate, 15% PEG 8000, and 20% glycerol. Owing to the physical limitations of the station it was possible to collect data only to 1.59 Å resolution, although strong diffraction extended to the edge of the 30 cm Mar image plate detector. Details of both the frozen and the room-temperature native data are given in Table 1.

*Refinement.* The initial model was refined using the program RESTRAIN (*16*) against the room temperature (RT) data to 1.72 Å resolution (Table 2). For the final cycles of refinement, atoms of the two iron clusters were allowed to refine anisotropically. This RT model was used as the starting model for refinement against 1.6 Å data from a frozen crystal, utilizing the program REFMAC (*17*) (refinement mode: maximum likelihood, conjugate gradient). The initial step was a rigid body refinement with the best result obtained using each domain as a separate rigid body. The initial *R*-factor was reduced from 38.2 to 35.0% ($R_{free}$ = 38.2−34.5%). All protein atoms were used, but the irons, sulfurs, and oxygens of the two clusters were excluded allowing assignment of the cluster independently of the original model. Further cycles of restrained positional and isotropic *B*-factor refinement were carried out and models were examined on a graphics system using the program O (*18*).

The metal−ligand bond distances were unrestrained throughout the refinement and no van der Waals restraints were applied to the iron atoms. As with the original

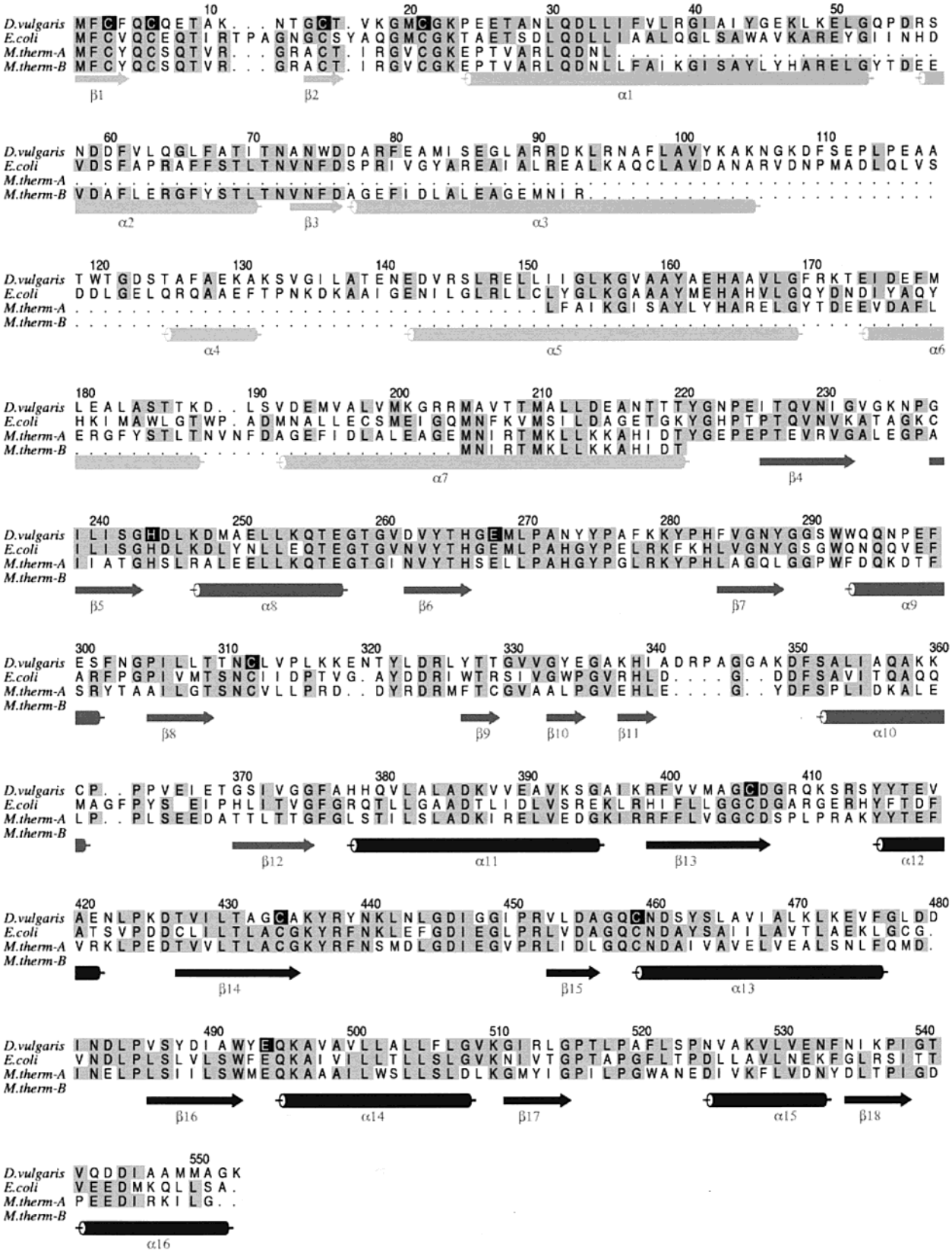


FIGURE 1: Sequence alignment of representative sequences from each of the three classes of HCP. *D. vulgaris* (*11*) represents class 1, *E. coli* (*35*) represents class 2 and *M. thermo* (*Methanobacterium thermoautotrophicum*) (*36*) represents class 3. A and B are alternative alignments for domain 1 of class 3. Conserved residues are shaded in light gray. Amino acid ligands to the cubane and hybrid clusters, which are strictly conserved in other HCPs, are printed reversed. Secondary structure elements are shaded from light to dark according to their organization in domains and numbered sequentially starting from the $NH_2$-terminus. Figure prepared using ALSCRIPT (*37*).

refinement, the cluster atoms were allowed to refine anisotropically for the last few cycles. Residue 406 was refined as a thiocysteine. Solvent atoms were allocated based on $F_o - F_c$ density peaks greater than $4\sigma$, no closer than 2.2 Å, and no further than 3.4 Å from a suitable hydrogen bonding partner. All further references to the HCP structure are to

Table 1: Data Collection and Structure Refinement[a]

| | room temperature | frozen |
|---|---|---|
| space group | $P2_12_12_1$ | $P2_12_12_1$ |
| resolution (Å) | 1.72 | 1.59 |
| wavelength (Å) | 0.870 | 1.488 |
| cell dimensions | | |
| *a* (Å) | 64.05 | 63.63 |
| *b* (Å) | 65.07 | 64.68 |
| *c* (Å) | 154.07 | 152.29 |
| total no. of (*hkl*) | 70 418 | 77 446 |
| avg multiplicity | 2.8 (2.5)[b] | 2.5 (2.4)[c] |
| $R_{merge}$[d] | 0.05 (0.16)[b] | 0.07 (0.10)[c] |
| avg $I/(\sigma I)$ | 10.5 (3.7)[b] | 5.6 (6.5)[c] |
| % completeness | 97.2 (81.3)[b] | 91.3 (87.7)[c] |
| refinement details | | |
| program | RESTRAIN | REFMAC |
| resolution range (Å) | 12.00−1.72 | 20.00−1.60 |
| no. protein atoms | 4218 | 4195 |
| no. cluster atoms | 17 | 17 |
| no. solvent molecules | 417 | 630 |
| no. reflections used | 70 157 | 72 244 |
| % data used in $R_{free}$ | 4.9 | 5.3 |
| final $R$ ($R_{free}$) | 17.1 (20.8) | 16.0 (18.3) |
| deviations of geometry from ideal values (target $\sigma$) | | |
| bond lengths | 0.011 (0.020) | 0.007 (0.020) |
| bond angles | 0.011 (0.040) | 0.021 (0.040) |
| peptide planarity | 0.007 (0.010) | 0.018 (0.030) |
| other planes | 0.007 (0.010) | 0.008 (0.020) |

[a] Details of the room temperature and frozen crystal data sets. [b] Final resolution bin 1.77−1.72 Å. [c] Final resolution bin 1.68−1.59 Å. [d] $R_{merge}$ is $\frac{\sum|I - \langle I\rangle|}{\sum\langle I\rangle}$. [e] $R$ is $\frac{\sum|F_{obs} - F_{calc}|}{\sum F_{obs}}$, $R_{free}$ is $R$ but calculated over a small % of data that were excluded from the refinement process.

Table 2: Phasing Statistics of the Four Heavy Atom Derivatives[a]

| | pMBA | U(AC) | TMLA | $KPtCl_4$ |
|---|---|---|---|---|
| concentration (mM) | 5 | 5 | 5 | 5 |
| soak time | 7 days | 34 h | 7 days | 7 days |
| wavelength | 1.0 | 0.87 | 1.488 | 0.87 |
| resolution | 2.80 | 2.82 | 2.5 | 2.5 |
| acentric reflections | | | | |
| no. of reflections | 11 069 | 8933 | 18 739 | 18 176 |
| phasing power[b] | 0.80 | 1.44 | 1.05 | 0.93 |
| Cullis R[c] | 0.91 | 0.75 | 0.84 | 0.63 |
| centric reflections | | | | |
| no. reflections | 1638 | 1226 | 2198 | 1433 |
| phasing power[b] | 0.64 | 1.03 | 0.85 | 0.63 |
| Cullis R[c] | 0.87 | 0.73 | 0.80 | 0.84 |

[a] UAC, uranyl acetate; pMBA, *p*-mecuribenzoic acid; TMLA, trimethyl lead acetate; $KPtCl_4$, potassium tetrachloroplatinate. [b] Phasing power is $F_H$/lack of closure. [c] Cullis R is $\frac{\sum|F_{PH} \pm F_H| - F_P}{\sum F_{PH} - F_P}$.

this model refined with the frozen crystal data to 1.6 Å resolution.

Coordinates and structure factor tables have been deposited in the Protein Data Bank (PDB) for both the room temperature and cryo-cooled structures; the deposition codes are 1E1D and 1E2U, respectively.

## RESULTS AND DISCUSSION

*Anomalous Dispersion Data.* The unusual nature and the mixed ligands of the hybrid cluster made it advisable to seek experimental confirmation that the metal atoms were all irons. The data from a frozen HCP crystal were collected at $\lambda = 1.488$ Å where iron has a significant anomalous dispersion signal compared to other transition metals (values of $f''$ are 3.026e, 0.522e, 0.637e, 0.478e, and 0.553e for iron, sulfur, zinc, nickel, and copper, respectively). The anomalous dispersion signal from the iron was sufficient to produce clear peaks on an anomalous difference Fourier synthesis, calculated using phases ($\phi + 90°$) from the room-temperature model with the cluster atoms removed to reduce phase bias. The top eight peaks of this Fourier synthesis are all of similar heights and correspond to the eight iron atoms assigned in the model as shown in Figure 2 (panels a and b). The next highest peaks identify some of the sulfur ligands in the clusters and other sulfur atoms within the structure, presented in Table 3.

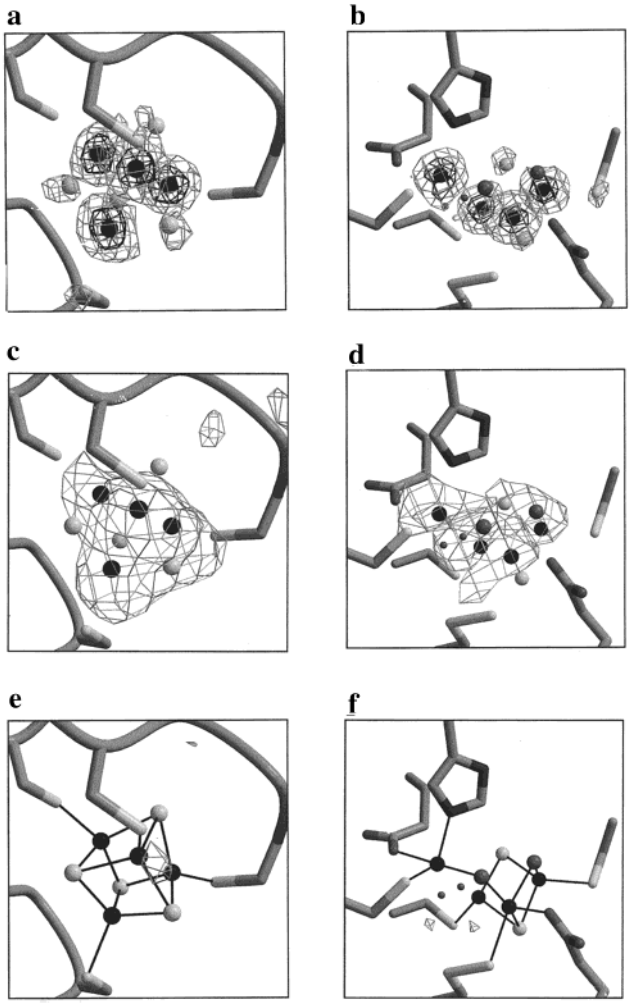


FIGURE 2: Anomalous difference Fourier syntheses. Syntheses a and b were computed using data collected with $\lambda = 1.488$ Å and are contoured at $4\sigma$, shown as light gray chicken wire, and at $15\sigma$ shown in black; (a) the cubane cluster and (b) the hybrid cluster. Iron, sulfur and oxygen atoms are depicted as black, light gray and dark gray spheres, respectively. Syntheses c and d are anomalous difference Fourier syntheses using data collected to 3 Å resolution with $\lambda = 1.743$ Å. They are contoured at $3\sigma$, represented by gray chicken wire. The cubane and hybrid clusters are shown in panels c and d, respectively. Syntheses e and f are analogous to c and (d) but using data collected with $\lambda = 1.745$ Å and contoured to $3\sigma$, represented by gray chicken wire. This figure was prepared with the programs BOBSCRIPT (*38*) and Raster3D (*39*).

For further verification, data were collected to 3 Å resolution, on beam line, BM14, at the ESRF Grenoble, at the iron edge ($\lambda_1 = 1.743$ Å) where $f''_{Fe} = 4.0$e and below the edge ($\lambda_2 = 1.745$ Å) where $f''_{Fe} = 0.5$e. For other elements the $f''$ signals at $\lambda_1$ and $\lambda_2$, respectively, are 0.699e and 0.700e for S, 0.846e and 0.847e for Zn, 0.735e and 0.737e

Table 3: Peak Heights and Corresponding Atoms for the 20 Highest Peaks in the Anomalous Difference Fourier Synthesis Calculated Using Coefficients $|F^+ - F^-|$, and Phases $\phi + 90$, Collected to 1.59 Å Resolution at $\lambda = 1.488$ Å[a]

| peak number | peak height ($\sigma$) | corresponding atom | $B$ value (Å$^2$) |
|---|---|---|---|
| 1 | 31.0 | Fe2 | 8.3 |
| 2 | 30.0 | Fe4 | 8.7 |
| 3 | 29.5 | Fe3 | 8.8 |
| 4 | 27.3 | Fe1 | 7.7 |
| 5 | 25.4 | Fe7 | 11.9 |
| 6 | 24.4 | Fe5 | 10.0 |
| 7 | 24.8 | Fe6 | 10.4 |
| 8 | 23.1 | Fe8 | 11.9 |
| 9 | 7.8 | S4 | 4.3 |
| 10 | 6.1 | Cys21 S$\gamma$ | 5.2 |
| 11 | 6.1 | S5 | 7.3 |
| 12 | 6.0 | S6 | 7.4 |
| 13 | 5.6 | S3 | 4.7 |
| 14 | 5.5 | S2 | 5.7 |
| 15 | 5.4 | Met210 S$\delta$ | 8.2 |
| 16 | 5.3 | Cys312 S$\gamma$ | 6.9 |
| 17 | 5.0 | Met403 S$\delta$ | 10.0 |
| 18 | 5.0 | S1 | 5.3 |
| 19 | 5.0 | Met269 S$\delta$ | 5.6 |
| 20 | 4.9 | Cys3 S$\gamma$ | 5.5 |
| | 4.0 | Thiocys406 S$\gamma$ | 7.2 |
| | 3.2 | Thiocys406 S$\delta$ (S7) | 9.9 |

[a] The hybrid cluster atoms comprise Fe5 to Fe8 and S5 to S7. The cubane cluster atoms comprise Fe1 to Fe4 and S1 to S4. The peaks for atoms S7 and S$\gamma$ of Cys 406 (the two sulfurs forming the thiocysteine) were not among the 20 highest peaks in the map but are shown here for completeness.

for Cu, and 0.636e and 0.637e for Ni. Anomalous difference Fourier syntheses using data from these two wavelengths show strong peaks for iron cluster atoms at $\lambda_1$ and virtually no peaks at $\lambda_2$, as illustrated in Figure 2 (panels c−f).

These experiments clearly indicate that the structure contains eight Fe atoms, four in each cluster.

*Quality of the Structure.* The $R_{factor}$ and $R_{free}$ of the final model are 16.3 and 18.6%, respectively. Further refinement statistics are given in Table 2. The model consists of 553 amino acid residues, 17 cluster atoms, and 630 solvent molecules. Residues Asp 60 and Ile152 have two conformations, and there are two cis peptides: Asn311−Cys312 and Gly515−Pro516. There is missing electron density for the side chains of several surface residues: Glu140; Lys254; Lys280; Lys317; Lys411; Asn443; Glu532, and Lys553.

The Ramachandran plot has no residues in the disallowed region, with 94.1% of residues within the core regions, as determined by PROCHECK (*16*). Average thermal coefficients are 12 Å$^2$ for protein atoms, 24 Å$^2$ for solvent atoms, and 8 Å$^2$ for the cluster atoms. The data precision indicator (DPI) as determined in REFMAC (*17*) is 0.08 Å. Figure 3 shows representative electron density of the final ($2mF_o - dF_c$) map and highlights the quality of the model.

*Architecture of Molecule.* The hybrid cluster protein from *D. vulgaris* comprises three domains numbered from N- to C-termini as shown in Figure 4(panels a and b). Domain 1 forms the top of a triangle while domains 2 and 3 form the bottom portion. The molecule is ca. 73 Å from the top of domain 1 to the base of the triangle between domains 2 and 3 and is approximately 60 Å at its widest point and 40 Å deep. A cubane type iron sulfur cluster lies in domain 1 and the hybrid iron cluster is at the interface between the three domains.

Domain 1 (residues 1−221) comprises a metal binding region (residues 1−21) and two 3-helix bundles, Figure 4c and 5b. The polypeptide chain of the metal binding region forms three sides of a box around the cubane cluster. The box is formed from two antiparallel strands ($\beta$1, $\beta$2) linked by a nine-residue loop, which folds around the cluster. A second, long loop links strand $\beta$2 with the first helix bundle, and forms the third side of the box. The four cysteine ligands, Cys3, Cys6, Cys15, and Cys21 are donated by strand $\beta$1, the first loop, strand $\beta$2, and the last loop of the binding motif, respectively. The bulk of domain 1 comprises two orthogonal 3-helix bundles. Each of these bundles contains three antiparallel helices with topology α1−α2−α3 and α5−α6−α7, respectively. The helices within each bundle are linked by short loops with the exception of α2 and α3, which are linked by a six-residue loop containing a short strand $\beta$3. Strand $\beta$3 is part of a two stranded sheet with $\beta$10 in domain 2. The two bundles are approximately orthogonal to one another and linked by a 36-residue (106−142) loop containing a short helical segment α4. The C-terminus of helix 2 points toward the cubane cluster, so that its negative dipole could have a stabilizing effect on the cluster.

The two bundles superimpose on one another with an rms deviation of 1.38 Å for 80 Cα. Eighteen residue types overlay

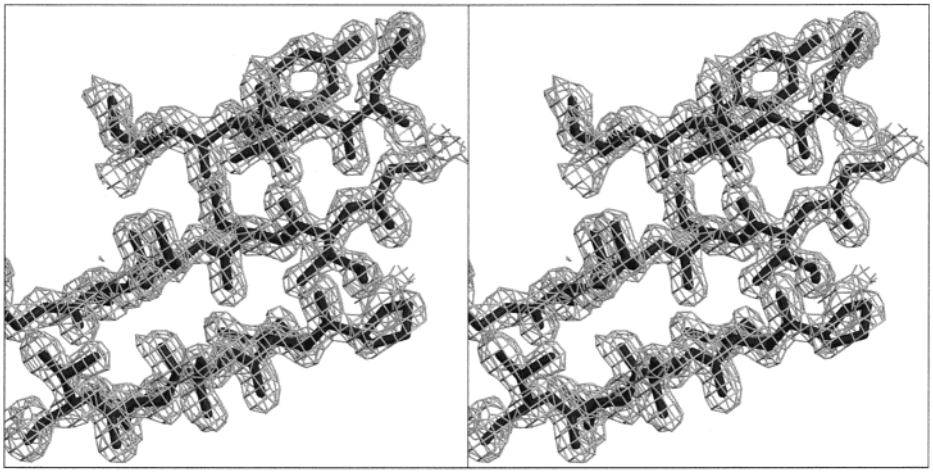

FIGURE 3: Stereoview of a representative portion of the electron density map. The ($2mF_o - dF_c$) map is contoured at 1.5$\sigma$ around strands $\beta$5, $\beta$6, and $\beta$8 from the $\beta$ sheet in domain 2. Atoms in this region have thermal parameters close to the average value, 12 Å$^2$. This figure was prepared with the programs BOBSCRIPT (*38*) and Raster3D (*39*).

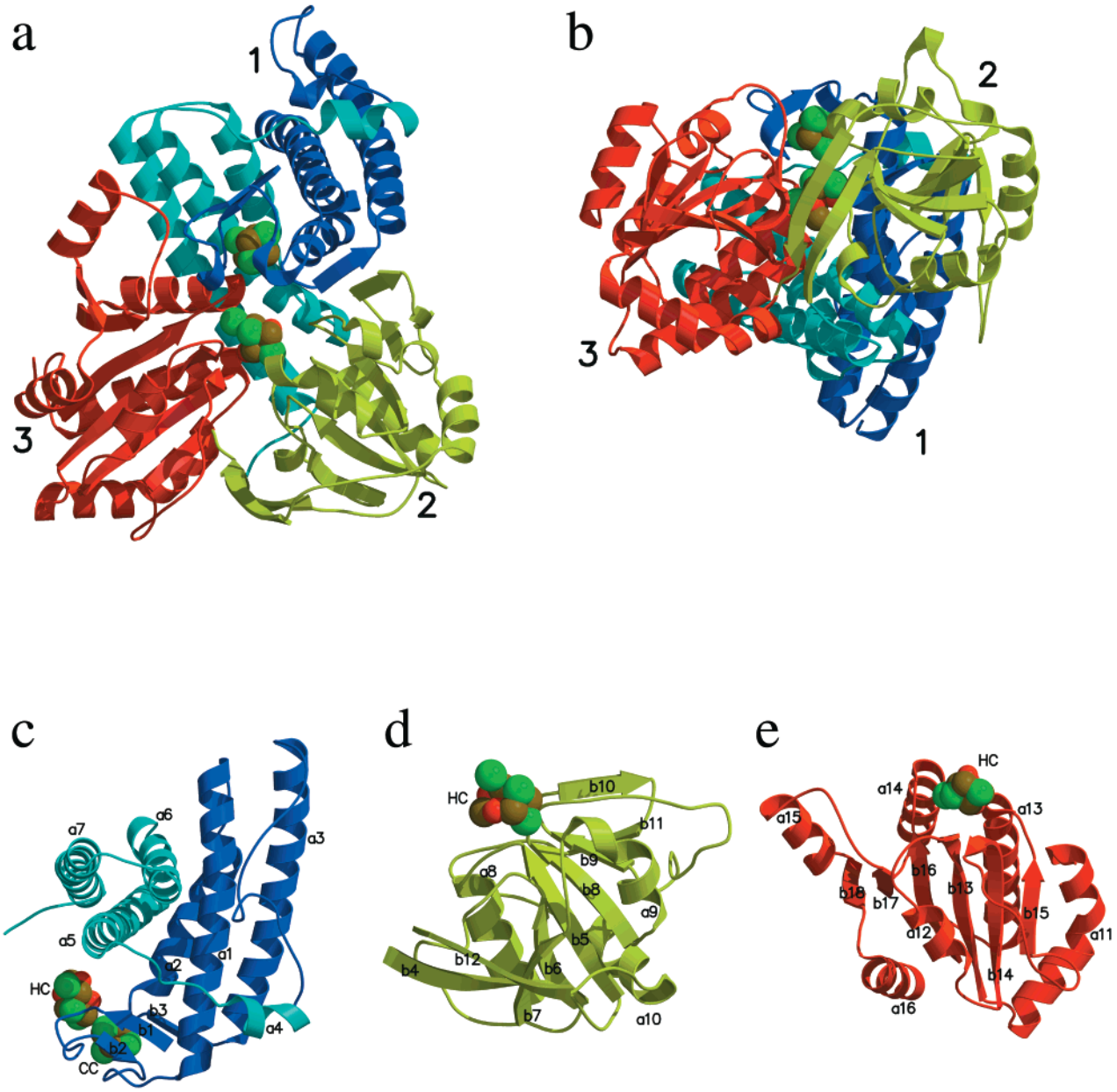


FIGURE 4: Overall fold of the hybrid-cluster protein. (a) A schematic representation of the 3-D structure, domain 1 is depicted in blue (light and dark), domain 2 in yellow and domain 3 in red. The iron, sulfur and oxygen atoms of the clusters are shown as brown, green, and red spheres, respectively. The cubane cluster is the upper of the two clusters. (b) An orthogonal view to (a) looking between domains 2 and 3, the hybrid cluster the lower of the two clusters. (c) Domain 1: the two 3-helix bundles are colored in different shades of blue. Individual secondary structure elements are labeled. (d) Domain 2 and (e) domain 3. This figure was prepared with the programs MOLSCRIPT (*40*) and Raster3D (*39*).

exactly between the bundles, and it is therefore highly probable the two bundles result from a gene duplication event. The sequence corresponding to atoms for one helix bundle is missing from the class 3 hybrid-cluster proteins, suggesting that only one bundle is essential for the function. The 3-helix bundle motif is a versatile structural building block and is found in many proteins, although the orthogonal domain arrangement is unique to HCP. The highest structural homology for the HCP bundle is with similar motifs from the molecular chaperone DnaK (PDB file 1DKZ-A, rms deviation of 3.2 Å for 80 Cα, 15% sequence identity) (*20*) and proteasome activator REGα (PDB file 1AVO-B, rms 2.8 Å for 78 Cα, 10% sequence identity) (*21*) both identified from a search of the Protein Data Bank using the program DALI (*22*).

Domain 2 (residues 222−375) has a mixed α/β structure (Figure 4d). It is entered and exited by a two-strand parallel sheet (β4 and β12). This is distinct from the core of the domain, which is a six-strand parallel twisted sheet of topology: β7-β6-β5-β8-β9-β11. The strands are linked by loops with the exceptions of β5 and β6 and strands β7 and β8, which are linked by helices α8 and α9, respectively. These helices are on opposing sides of the central sheet. Strands β9 and β11 are linked by β10, which is not part of the main sheet, but instead is part of a two stranded parallel sheet with β3 of domain 1. β11 is connected to β12, the last strand in the domain, by α10.

Domain 3 (376−553) is entered by a 20 residue kinked helix α11 and is then organized in a similar way to domain 2, also having a central parallel twisted sheet, of topology

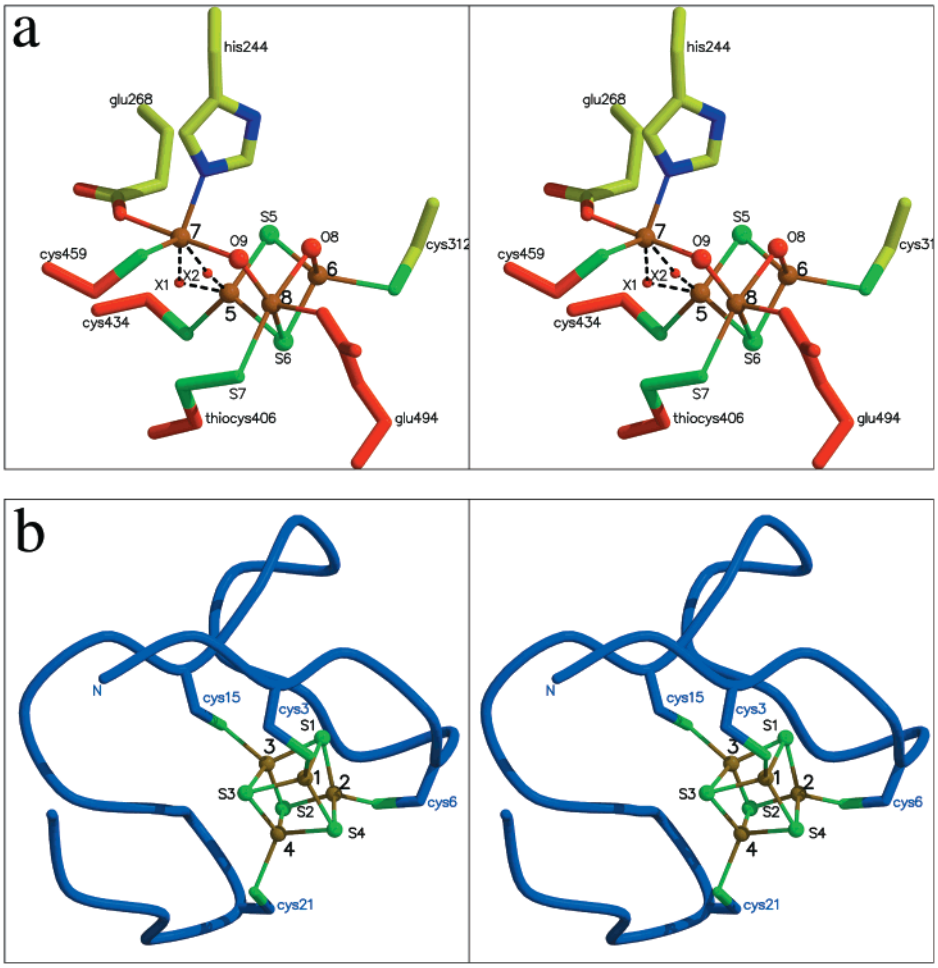


FIGURE 5: Stereoviews of the metal clusters. (a) The hybrid cluster with iron, sulfur and oxygen atoms represented as brown, green and red spheres, respectively. Ambiguous atoms $X_1/X_2$ are shown as smaller spheres and their bonds with dotted lines. The carbon atoms of protein ligands from domain 2 are in yellow and those from domain 3 in red. (b) The cubane cluster bound at the N-terminus of the molecule with iron and sulfur represented in brown and green, respectively. The iron atoms are labeled with large numbers and the oxygens and sulfurs as S*n* and O*n* respectively. This figure was prepared with the programs MOLSCRIPT (*40*) and Raster3D (*39*).

$\beta$15-$\beta$14-$\beta$13-$\beta$16-$\beta$17-$\beta$18. The strands are linked mainly by helices that flank the sheet with α13, α14 and α15 on one side and α12 and α16 on the other.

Domains 2 and 3 can be superimposed although only the central sheets align well. Aligning these strands gives an rms deviation of 1.9 Å for the 48 well-matched Cα atoms. The overall topology of both domains is similar to the flavodoxin doubly wound sheet motif. The hybrid cluster ligands are located at the carboxy ends of the $\beta$ strands, which is in keeping with proteins of this fold having their functional residues in these regions. A search of the PDB using DALI indicates that domain 2 has highest structural homology to the N-terminal receiver domain of *E. coli* nitrate/nitrite response regulator NarL (PDB file 1RNL) (*23*), (rms 3.0 Å for 92 Cα, 17% sequence identity). This protein is part of the Che Y-like protein structural family with a three-layer $\alpha\beta\alpha$ five-stranded, parallel sheet motif with order 21 345 (*24*). It also has good identity with the periplasmic D-ribose binding protein (PDB flie 2DRI, rms 2.9 Å, 96 Cα, 9% identity) (*25*) and a periplasmic glucose/galactose receptor (PDB file 1GCA) (*26*) (rms 3.1 Å, 98 Cα, 6% identity), which both have six-stranded parallel twisted sheets and are chemotaxis receptors. Domain 3 has more helices than domain 2 and has greatest structural homology with N-carbamoylsarcosine amidohydrolase (PDB file 1NBA, rms 3.1 Å, 118 Cα, 8% identity) (*27*), a three-layer $\alpha\beta\alpha$ six-stranded parallel twisted sheet motif. The significance for these similarities for function and or evolution of HCP is unclear.

*Iron Sulfur Clusters.* The hybrid cluster is unique among the iron sulfur proteins for which structures are available. The atomic distribution, as interpreted from the final electron density syntheses, is shown in Figure 5a and metal−ligand and metal−metal distances are given in Table 4. The original interpretation of the electron density of the hybrid cluster was carried out with great caution, using peak heights in electron density syntheses and the behavior of *B*-factors, in conjunction with chemical considerations, to assign Fe, S, and O atoms. In addition, information is now available from the anomalous difference data that serves to confirm the initial interpretation of the cluster (*1*). The presence of the four Fe atoms was clearly demonstrated by the peak heights in the anomalous difference Fourier synthesis (Table 3), and using this map it was also possible to distinguish the majority of sulfur atoms in the protein including the two sulfur atoms of the thiocysteine moiety. A remaining ambiguity in the cluster is an ellipsoidal region of electron density (maximum height of 8.9$\sigma$) in the difference Fourier synthesis between Fe5 and Fe7 (interpreted as atom X in previous work). This density has now been modeled as an oxygen atom in two

Table 4: Distances (Å) between Atoms within the Hybrid and Cubane Clusters[a]

| | | | |
|---|---|---|---|
| hybrid cluster | | | |
| Fe5−Cys434S$\gamma$ | 2.25 | Fe6−Cys312S$\gamma$ | 2.36 |
| Fe5−S5 | 2.27 | Fe6−S5 | 2.19 |
| Fe5−S6 | 2.28 | Fe6−S6 | 2.25 |
| Fe5−$X_1$ | 2.06 | Fe6−O8 | 2.00 |
| Fe5−$X_2$ | 1.81 | | |
| | | Fe8−Glu494O$\epsilon$2 | 2.02 |
| Fe7−His244N$\epsilon$2 | 2.12 | Fe8−S6 | 2.49 |
| Fe7−Glu268O$\epsilon$2 | 2.14 | Fe8−S7(Thiocys406S$\delta$) | 2.54 |
| Fe7−Cys459S$\gamma$ | 2.40 | Fe8−O9 | 2.00 |
| Fe7−O9 | 2.12 | Fe8−O8 | 2.19 |
| Fe7−$X_1$ | 2.03 | S7(Thiocys406S$\delta$)− | 1.90 |
| Fe7−$X_2$ | 1.90 | Thiocys406S$\gamma$ | |
| iron−iron distances | | | |
| Fe5−Fe6 | 2.68 | Fe6−Fe7 | 4.99 |
| Fe5−Fe7 | 3.68 | Fe6−Fe8 | 3.05 |
| Fe5−Fe8 | 3.15 | Fe7−Fe8 | 3.33 |
| cubane cluster | | | |
| Fe1−CysS$\gamma$3 | 2.30 | Fe2−CysS$\gamma$6 | 2.26 |
| Fe1−S1 | 2.27 | Fe2−S1 | 2.25 |
| Fe1−S3 | 2.28 | Fe2−S2 | 2.29 |
| Fe1−S4 | 2.25 | Fe2−S4 | 2.33 |
| Fe4−CysS$\gamma$21 | 2.33 | Fe3−CysS$\gamma$15 | 2.35 |
| Fe4−S2 | 2.31 | Fe3−S1 | 2.31 |
| Fe4−S3 | 2.25 | Fe3−S2 | 2.26 |
| Fe4−S4 | 2.30 | Fe3−S3 | 2.27 |
| iron−iron distances | | | |
| Fe1−Fe2 | 2.70 | Fe2−Fe3 | 2.70 |
| Fe1−Fe3 | 2.66 | Fe2−Fe4 | 2.64 |
| Fe1−Fe4 | 2.73 | Fe3−Fe4 | 2.69 |

[a] Cluster bond distances were unrestrained and the overall DPI of the coordinates is 0.08 Å (*17*).

alternate positions, $X_1$ and $X_2$, 0.8 Å apart, and of approximately equal occupancy. Published resonance Raman spectroscopic data (*1*, *28*) suggest the possibility of a monobridged Fe−O−Fe center with an Fe−O−Fe angle of 150−180°. This would correlate with an Fe−$X_2$−Fe angle of 163°, whereas the Fe−O8−Fe and Fe−O9−Fe angles are closer to 100°.

The hybrid cluster comprises both O and S bridges between Fe atoms, and the local environments of the metal atoms can be described as tetrahedral for Fe5 and Fe6 and trigonal bipyramidal for Fe7 and Fe8. It should be noted that in the absence of $X_1$/$X_2$ Fe5 and Fe7 have trigonal and square pyramidal coordination, respectively. Fe5 is coordinated to one protein ligand Cys434, to two bridging S atoms (S5 and S6), and to an atom at position $X_1$ or $X_2$. Fe6 is also coordinated to one protein ligand, Cys312, which adopts a cis-peptide conformation, and two bridging S atoms (S5 and S6). Thus, this part of the cluster resembles either a [2Fe2S] moiety or one face of a cubane cluster. However, Fe6 is also coordinated to an oxygen atom (O8) which bridges F6 and Fe8 to give a tetrahedral environment for Fe6. Fe7 is bound to three protein ligands, His244, Glu268, and Cys459; it is also bound to an oxygen atom O9 which bridges Fe7 and Fe8 and to an atom at position $X_1$ or $X_2$. Fe8 is linked directly to only one protein ligand Glu494 but also binds to two bridging O atoms (O8 and O9) and to a sulfur atom S7. S7 forms a persulfide with S$\gamma$ of Cys406, producing a thiocysteine moiety. Fe8 has a fifth ligand, S6 (which also bridges Fe5 and Fe6), thus having a distorted trigonal bipyramidal geometry. The thiocysteine in HCP is a unique iron ligand although thiocysteines have been observed in other proteins that are thought to act as catalytic intermediates in sulfur transfer reactions (*29*−*31*).

There is an additional peak in the difference electron synthesis at the 4$\sigma$ level on the edge of the hybrid cluster. It has been modeled as a solvent molecule, presumably with only partial occupancy, and is 2.4 and 2.9 Å from Fe6 and Fe8, respectively; this solvent molecule can form a hydrogen bond with OE1 of Glu494.

The HCP cubane cluster is a conventional [4Fe-4S] cluster but linked to the protein by an unusual sequential arrangement of four cysteine residues, shown in Figure 5b. The iron atoms Fe1, Fe2, Fe3, and Fe4 are bound via Cys3, Cys6, Cys15, and Cys21, respectively, in a $CysX_2CysX_8CysX_5$-Cys motif. Thus, all four ligands are contributed by residues at the N-terminal region of the molecule, which appears to wrap itself tightly around the cluster. This motif is similar to the typical [4Fe-4S] linkage where the amino acid sequence for cluster binding is $CysX_2CysX_2CysX_nCys$ (*9*), although there is a significant difference in that typically the fourth ligand is donated by the C-terminal end of the sequence. As with other cubane structures each of the iron atoms is also coordinated by three sulfur atoms; bond distances are given in Table 4. The cysteine-binding region is conserved in other HCP sequences although there is some differentiation between the three classes of HCP sequences. However, recent spectroscopy and sequence comparisons indicate that class 2 HCPs may have a [2Fe-2S] cluster instead of the [4Fe-4S] found in the other two classes (*8*).

A comparison of the iron−iron distances in the cubane and hybrid clusters further illustrates the unusual nature of the latter cluster. The cubane structure is compact with all metals being approximately 2.7 Å from one another, while the metal atoms in the hybrid cluster are more separated, with an average iron−iron distance of 3.5 Å, the largest being 5.0 Å.

*Electron Transfer between Clusters.* The edge-to-edge distance between the hybrid cluster and the [4Fe-4S] cluster is 10.9 Å (Fe2 to S6) which is within range for direct electron transfer (*32*). There is no obvious pathway for electron transfer via the protein. Residues between the two clusters are Asn72, Tyr493, Thr71, Cys6, Glu494, and Cys312, the latter three being ligands to the clusters. Additionally, two solvent molecules, one hydrogen bonded to Glu494, are in the space between the clusters. Thr71, Asn72, and Tyr493 are all approximately midway between the two metal clusters. All these residues are conserved except Tyr493, which is replaced by a phenylalanine in class 2 HCPs and a methionine in class 3 HCP sequences.

*Cavities and Channels.* In addition to the overall architecture as described above, there are several other features of interest involving the metal clusters. By far the most intriguing of these is a large, mainly hydrophobic cavity which allows access to the hybrid cluster and specifically to Fe8 and its immediate environment including $X_1$/$X_2$ and the thiocysteine moiety. The bowl of the cavity is formed from residues contributed mainly by domain 3 and involving strands $\beta$13−$\beta$18 inclusive and helices $\alpha$11 and $\alpha$15, see Figure 6. Helices $\alpha$5 and $\alpha$7 from the second three-helix bundle in domain 1 also contribute 12 hydrophobic side chains, suggesting that sequence alignment B in Figure 1 is more likely. Helices $\alpha$13 and $\alpha$14 are within the bowl. In its longest dimension, the cavity is approximately 30−35 Å in length from the backbone of helix $\alpha$11 to $\alpha$15. The only solvent atom in this region is bound to the backbone carbonyl

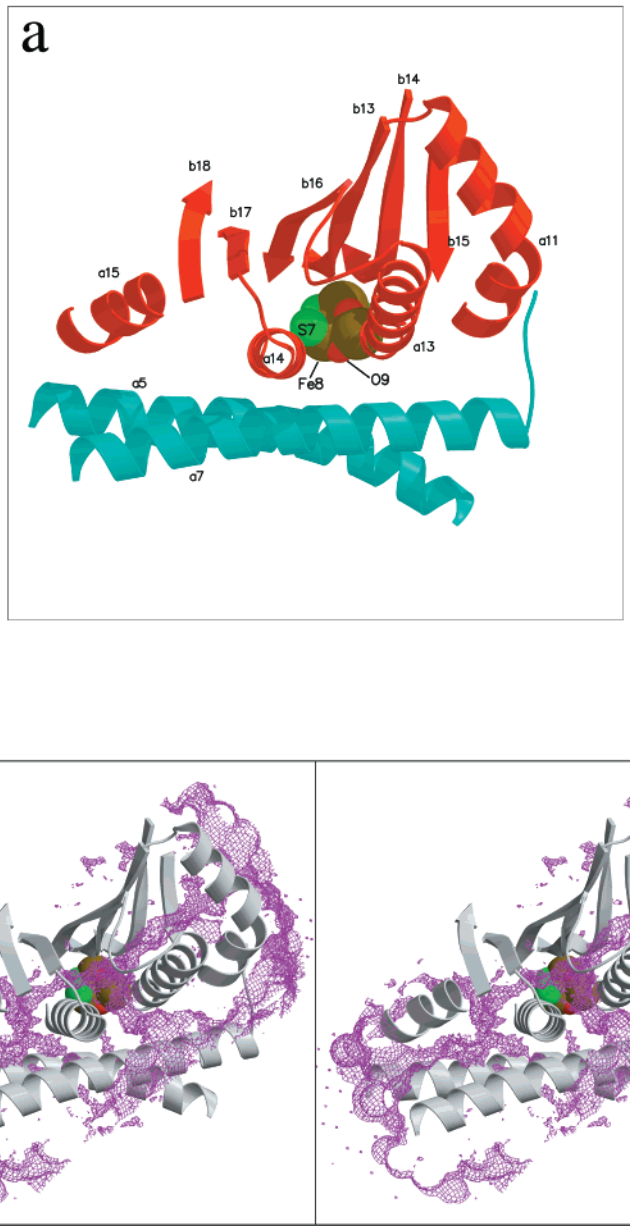


FIGURE 6: Hydrophobic cavity and the hybrid cluster. (a) Shows the structural elements forming the hydrophobic cavity, colored as in Figure 4. This view shows an access route to the hybrid cluster. (b) Stereoview of a 15 Å thick slab of map (in magenta chickenwire) from a cavity calculation performed on the refined model using a 0.8 Å radius probe, the map shows the areas accessible to the probe. The calculation was performed using CAVENV (*33*). This figure was prepared with the programs BOBSCRIPT (*38*), MOLSCRIPT (*40*), and Raster3D (*39*).

group of Phe376, on the loop between $\beta$12 and $\alpha$11. The cavity is lined with 78 side chains, of which 73 are hydrophobic; 47 of these have their hydrophobic nature conserved and 16 are totally conserved within the HCP sequences. There are five polar residues within the cavity, which either have their hydrophilic portions external to the cavity or are hydrogen bonded to backbone atoms.

Strands $\beta$13 and $\beta$14 together with helices $\alpha$13 and $\alpha$14 form the central portion of the cavity and also contribute ligands to the hybrid cluster; thiocys406 ($\beta$13) to Fe8, Cys434 ($\beta$14) to Fe5, Cys459 ($\alpha$13) to Fe7, and Glu494 (the loop between $\beta$16 and $\alpha$14) to Fe8. If the cluster is viewed down the channel shown in Figure 6a, Fe8 and its environment are clearly visible. However, there are other access routes, as shown by a cavity map (Figure 6b), calculated using the program CAVENV (*33*, *14*). These routes are between helix $\alpha$15 at one end of the bowl and helix $\alpha$11 at the other end. All the access routes meet at a central point around the hybrid cluster. The presence of the cavity clearly reinforces the idea that HCP is an enzyme with its activity centered around Fe8 of the hybrid cluster.

The hybrid cluster is also accessible by a solvent channel, channel 1, comprising an ordered network of 12 solvent atoms from the hybrid cluster to the protein surface (Figure 7a). The network is linked to O8 of the cluster, almost opposite the X atoms but again involving Fe8, via Lys496, which is conserved over all three classes of HCP. There are three solvent atoms between Lys496 and Gln458 where the channel divides into two separate paths to the exterior of the protein. One path comprises only two solvent atoms and emerges between helices $\alpha$5 and $\alpha$8. The other path contains seven solvent atoms and is between $\alpha$7 of domain 1, $\beta$4 of domain 2, and $\alpha$11 and $\alpha$13 of domain 3. This pathway has a length of 18 Å from Lys496 to the surface of the protein

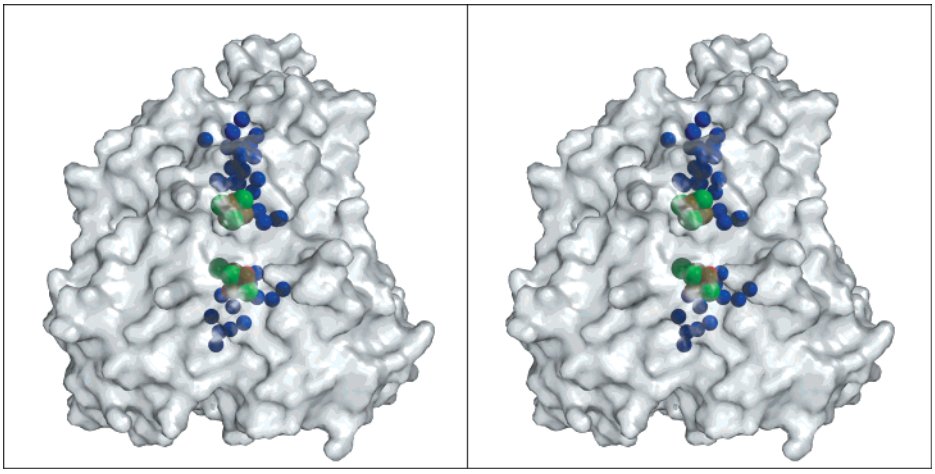

FIGURE 7: Solvent channels. Stereoview of the translucent molecular surface with water molecules in the solvent channels represented by blue spheres and the metal cluster atoms, Fe, S, and O represented as brown, green, and red spheres, respectively. The hybrid cluster is in the center of this view. This figure was prepared using the programs GRASP (*41*), MOLSCRIPT (*40*), and Raster3D (*39*).

and is approximately 6 Å in diameter. The 12 solvent molecules have an average thermal parameter of 9 Å$^2$ and are hydrogen bonded to 11 polar side chains, six of which are conserved across all three HCP sequence classes and three over two classes.

Although the cubane cluster is close to the N-terminus of the molecule and therefore accessible for electron transfer, for example, with NADH oxidoreductase, it is also linked to the surface of the protein through channel 2 which forms a funnel between the two 3-helix bundles of domain 1. The cubane cluster is connected by S4 to a cavity of four solvent atoms, which is situated between the end of helix α2, strand *β*3, and the start of helix α3. The four water molecules are hydrogen bonded to the backbone atoms of α2 and *β*3. The cavity is enclosed within the protein by three side chains, Lys23, Gln31, and Gln5, all of which are conserved over the three classes of HCP. However, these side chains link the cavity to the main bulk of channel 2. The channel then progresses to the surface by way of 18 solvent molecules. At the surface, it is divided into two by the loop between α3 and α4, so that one side emerges between the loop, α1, and the N-terminus of the protein. The other, larger opening is formed by the space between the loop, α5 and α1. Over this opening extend two arginine residues, Arg39 and Arg148; the latter is conserved over classes 2 and 3 but missing in class 3. The channel contains 22 solvent molecules, with an average thermal parameter of 10 Å$^2$ and is approximately 15 Å deep and 6 Å wide. The solvent atoms make the majority of their contacts to the protein via O and N atoms of the protein backbone. However, there are eight polar side chains within the channel, four of which are conserved over all three classes (including the three mentioned above), two have their polar nature conserved, and the other two are missing in class 3 HCPs.

## CONCLUSIONS AND BIOLOGICAL IMPLICATIONS

The crystal structure of HCP from *D. vulgaris* has now been determined to 1.6 Å and shown to contain two 4Fe clusters with a closest approach of 10.9 Å. The cubane cluster is a conventional [4Fe-4S] cubane situated near the exterior of the protein as part of domain 1, while the hybrid cluster is a novel Fe−S−O cluster situated near the interface of all three domains. Both the preparation of the protein and the crystal structure analysis were performed under aerobic conditions and a pertinent question could be whether this has introduced artifacts into the structure and particularly at the active site. Such artifacts may involve oxidation products including the presence of oxygen in the vicinity of Fe8 in the hybrid cluster, but would be unlikely to have any effect on the [2Fe-2S] moiety involving Fe5, Fe6, or, indeed, Fe7, which is coordinated by histidine, cysteine, and glutamic acid residues. However, if possible, it is evident that a study of this HCP protein should also be undertaken under strictly anaerobic conditions.

At the present time, it can be stated that without the undefined moiety $X_1/X_2$, the hybrid cluster has an open structure with incomplete coordination of the Fe5 and Fe7 sites. This is suggestive that the protein has a capacity to bind and release small species and a possible catalytic role. The density at $X_1/X_2$ was initially thought to accommodate a binuclear species, e.g., CO, CN, or NO. However, refinement of these species, which have a bond length of around 1.2 Å, was not satisfactory. Experiments to replace X, by soaking either CO or CN into HCP crystals, have not been successful. However, a diatomic species, with partial occupancy and/or disorder cannot be eliminated. A chloride ion at this site was also considered, but the electron density is not sufficient for a full occupancy atom of this number of electrons. Published resonance Raman spectroscopic data (*1*, *28*) suggest the presence of a either an Fe═O species or a monobridged Fe−O−Fe center with an Fe−O−Fe angle of 150−180°, which would correlate with the interpretation of $X_1/X_2$ as an oxygen.

Montet et al. (*34*) have suggested that hydrophobic channels in proteins play an instrumental role in both gas storage and gaseous substrate access to active sites. The hybrid cluster is directly adjacent to a large mainly hydrophobic cavity that allows access to it and in particular Fe8 and its surroundings. The unknown moiety $X_1/X_2$ is within the cavity and lends further weight to the idea of it being a substrate, but probably with only a partial occupancy in the crystal structure determined under aerobic conditions. The well-ordered solvent channels to both metal sites provide

substrate access routes, including $H^+$ ions, and further implicate the clusters as having a catalytic role.

The biological function of HCP is currently unknown. The redox properties of HCP have been demonstrated in previous spectroscopic work (*1*), but several features of the structure suggest that the protein could act as an enzyme. There is some evidence that the protein plays a role in nitrate metabolism (*8*). However, the presence of the thiocysteine moiety and its relationship to the sulfur transfer proteins (*29*−*31*) cannot be completely discounted, and it is possible that the protein plays some role in the formation of iron sulfur clusters.

## ACKNOWLEDGMENT

We acknowledge the support provided by the synchrotron radiation sources at Daresbury Laboratory, U.K. (SRS), and in Grenoble, France (ESRF), without which much of this work would not have been possible. We also acknowledge Gordon Leonard for multiwavelength data collection on BM14.

## REFERENCES

1. Arendsen, A. F., Hadden, J., Card, G., McAlpine, A. S., Duke, E. H. M., Lindley, P. F., Kröckel, M., X., T. A., Feiter, M. C., Charnock, J. M., Garner, D. C., Marritt, S. J., Thomson, A. J., Kooter, I. M., Johnson, M. K., van de Berg, W. A. M., van Dongen, W. M. A. M., and Hagen, W. R. (1998) *J. Bioinorg. Chem. 3*, 81−95.
2. Pereira, A. S., Tavares, P., Krebs, C., Huynh, B. H., Rusnak, F., Moura, I., and Moura, J. J. G. (1999) *Biochem. Biophys. Res. Commun. 260*, 209−215.
3. Tavares, P., Pereira, A. S., Krebs, K., Ravi, N., Moura, J. J. G., Moura, I., Huynh, B. H., and Thomson, A. J. (1998) *Biochemistry 37*, 2830−2842.
4. Hagen, W. R., Pierik, A. J., and Veeger, C. (1989) *J. Chem. Soc., Faraday Trans. 1 85*, 4083−4090.
5. Pierik, A. J., Wolbert, R. B. G., Mutsaers, P. H. A., Hagen, W. R., and Veeger, C. (1992) *Eur. J. Biochem. 206*, 697−704.
6. Moura, I., Tavares, P., Moura, J. J. G., Ravi, N., Huynh, B. H., Liu, M. Y., and Legall, J. (1992) *J. Biol. Chem. 267*, 4489−4496.
7. van Den Berg, W. A. M., Stevens, A. A. M., Verhagen, M. F. J. M., van Dongen, W. M. A. M., and Hagen, W. R. (1994) *Biochim. Biophys. Acta 1206*, 240−246.
8. van Den Berg, W. A. M., Hagen, W. R., and van Dongen, W. M. A. M. (2000) *Eur. J. Biochem. 267*, 666−676.
9. Lindley, P. F. (1996) *Rep. Prog. Phys. 59*, 876−933.
10. Gennis, R. B., and Stewart, V. (1996) in *Escherichia coli and Salmonella: cellular and molecular biology I* (Neidhardt, F. C., and Curtiss, R., Eds.). 217−216, ASM Press, Washington, DC.
11. Stokkermans, J. P. W. G., Pierik, A. J., Wolbert, R. B. G., Hagen, W. R., van Dongen, W. M. A. M., and Veeger, C. (1992) *Eur. J. Biochem. 208*, 435−442.
12. Arendsen, A., Bulsink, Y., Hagen, W., Hadden, J., Card, G., McAlpine, A. S., Bailey, S., Zaitsev, V., and Lindley, P. (1996) *Acta Crystallogr., Sect. D 52*, 1211−1213.
13. Leslie, A. G. W. (1992) *Joint CCP4-EACMB Newsletter on Protein Crystallography*, SERC, Daresbury Laboratory, U.K.
14. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr., Sect. D 50*, 760−763.
15. Cowtan, K. D., and Main, P. (1993) *Acta Crystallogr., Sect. D 49*, 148−157.
16. Driessen, H., Haneef, M. I. J., Harris, G. W., Howlin, B., Khan, G., and Moss, D. S. (1989) *J. Appl. Crystallogr. 22*, 510−516.
17. Murshudov, G. N., Vagin, A. A., and Dodson, E. (1997) *Acta Crystallogr., Sect. D 53*, 240−255.
18. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr., Sect. A 47*, 110−119.
19. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
20. Zhu, X. T., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E., and Hendrickson, W. A. (1996) *Science 272*, 1606−1614.
21. Knowlton, J. R., Johnston, S. C. Whitby, F. G., Realini, C., Zhang, Z. G., Rechsteiner, M., and Hill, C. P. (1997) *Nature 390*, 639−643.
22. Holm, L., and Sander, C. (1993) *J. Mol. Biol. 233*, 123−138.
23. Baikalov, I., Schröder, I., Kaczor-Grzeskowiak, M., Grzeskowiak, K., Gunsalus, R. P., and Dickerson, R. E. (1996) *Biochemistry 34*, 11053−11061.
24. Volz, K., and Matsumura, P. (1991) *J. Biol. Chem. 266*, 15511−15519.
25. Mowbray, S. L., and Brent Cole, L. (1992) *J. Mol. Biol. 225*, 155−175.
26. Zou, J. Y., Flocco, M. M., and Mowbray, S. L. (1993) *J. Mol. Biol. 223*, 739−752.
27. Romao, M. J., Turk, D., Gomis-Ruth, F.-X., Huber, R., Schumacher, G., Mollering, M., and Russmann, L. (1992) *J. Mol. Biol. 226*, 1111−1130.
28. De Vocht, M. L., Kooter, I. M., Bulsink, Y. B. M., Hagen, W. R., and Johnson, M. K. (1996) *J. Am. Chem. Soc. 118*, 2766−2767.
29. Ploegman, J. H., Drent, G., Kalk, K. H., and Hol, W. G. J. (1978) *J. Mol. Biol. 123*, 557.
30. Zheng, L. M., White, R. H., Cash, V. L., Jack, R. F., and Dean, D. R. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 2754−2758.
31. Zheng, L. M., White, R. H., Cash, V. L., and Dean, D. R. (1994) *Biochemistry 33*, 4714−4720.
32. Williams, R. J. P. (1997) *J. Bioinorg. Chem. 2*, 373−377.
33. Volbeda, A. (1999) *Les Ecoles Physique et Chimie du Vivant, Analyse de l'organisation tridimensionnelle des proteines 1*, 47−52.
34. Montet, Y., Amara, P., Volbeda, A., Vernede, E., Hatchikian, C., Field, M., Frey, M., and Fontecilla-Camps, J. C. (1997) *Nat. Struct. Biol. 4*, 523−526.
35. Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., ColladoVides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., Shao, Y., Bult, C. J., White, O., Olsen, G. J., Zhou, L. X., Fleischmann, R. D., and Sutton, G. G. (1997) *Science 277*, 1453−1470.
36. Smith, D. R., Doucette-Stamm, L. A., Deloughery, C. L., H. M, D. J., Aldredge, T., Bashirzadeh, R., Blakely, D., Cook, R., Gilbert, K., Harrison, D., Hoang, L., Keagle, P., Lumm, W., Pothier, B., Qiu, D. Y., Spadafora, R., Vicaire, R., Wang, Y., Wierzbowski, J., Gibson, R., Jiwani, N., Caruso, A., Bush, D., Safer, H., Patwell, D., Prabhakar, S., McDougall, S., Shimer, G., Goyal, A., Pietrokovski, S., Church, G. M., Daniels, C. J., Mao, J. I., Rice, P., Nölling, J., and Reeve, J. N. (1997) *J. Bacteriol. 179*, 7135−7155.
37. Barton, G. J. (1993) *Protein Eng. 6*, 37−40.
38. Esnouf, R. M. (1997) *J. Mol. Graphics 15*, 132−136.
39. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−524.
40. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
41. Nicholls, A., Bharadwaj, R., and Honig, B. (1993) *Biophys. J. 64*, A166.

BI001483M