Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural and functional relationships in the hybrid cluster protein family: structure of the anaerobically purified hybrid cluster protein from *Desulfovibrio vulgaris* at 1.35 Å resolution

David Aragão,[a,b] Edward P. Mitchell,[b] Carlos F. Frazão,[a] Maria Arménia Carrondo[a] and Peter F. Lindley[a]*

[a]Instituto de Tecnologia Química e Biológica, Avenida da Repúlica, Apartado 127, 2781-901 Oeiras, Portugal, and [b]European Synchrotron Radiation Facility, 6 Rue Jules Horowitz, BP 220, 38043 Grenoble CEDEX, France

Correspondence e-mail: lindley.p@gmail.com

Received 6 February 2008
Accepted 4 April 2008

PDB Reference: hybrid cluster protein, 1w9m, r1w9msf.

The hybrid cluster protein (HCP) from the sulfate-reducing bacterium *Desulfovibrio vulgaris* strain Hildenborough has been isolated and crystallized anaerobically. The phase problem was solved for a $P2_12_12_1$ crystal form using multiple-wavelength anomalous diffraction data collected in the vicinity of the Fe *K* absorption edge. Although the overall protein structure is essentially the same as that previously obtained, it shows that the nature of the hybrid cluster has particular differences when isolated and crystallized in the absence of oxygen and this provides insight into the structural features associated with changes in the oxidation state. A comparison between HCPs and carbon monoxide dehydrogenases (CoDs) shows that they possess a similar fold and that the dehydrogenases have a related cluster at the equivalent HCP hybrid cluster position. This helps to understand the nature of the hybrid cluster and to predict a dimeric structure for class 3 HCPs, which lack the N-terminal region.

1. Introduction

Hybrid cluster proteins (HCPs) have molecular weights of around 60 kDa corresponding to some 550 residues and are present in a wide variety of organisms from the Bacteria, Archaea and Eukarya kingdoms. HCPs contain two types of Fe–S clusters: one variable and present as either [2Fe–2S] or [4Fe–4S] and another so-called hybrid cluster (nominally [4Fe–2S–3O]). The latter is the main feature of the HCP structure and was initially thought to be a [6Fe–6S] moiety in a 'prismane' configuration (Hagen, 1989). Subsequent studies, including an X-ray structure determination at 1.72 Å resolution, indicated a [4Fe–2S–3*X*] moiety and suggested that it may have resulted from aerobic oxidation (Arendsen *et al.*, 1998). Furthermore, not all of the ligands *X* for this cluster could reliably be assigned. Further crystallographic studies led to higher resolution (1.25 Å) structures that were obtained using protein prepared and crystallized anaerobically for *Desulfovibrio desulfuricans* and aerobically for *D. vulgaris* (Macedo *et al.*, 2002). Subsequently, dithionite-reduced structures for both bacteria were also determined (Aragão *et al.*, 2003). Despite all these crystallographic structures, detailed spectroscopic analyses and many other studies, the precise function(s) of HCP is still subject to discussion. Two main hypotheses for its function have been proposed. In the first hypothesis it is thought to play a role in protection against

reactive nitrogen species, while the second involves a role in oxygen-related oxidative-protection mechanisms. The first hypothesis gained ground with the publication of three reports: (i) the conversion of the HCP-related NiFeS carbon monoxide dehydrogenase (CoD) to a hydroxylamine reductase by amino-acid mutation or nickel to iron replacement at the CoD active site (Heo *et al.*, 2002), (ii) the report of hydroxylamine reductase activity for *Escherichia coli* HCP (Wolfe *et al.*, 2002) and (iii) increased anaerobic tolerance to hydroxylamine of *E. coli* cells overproducing *Rhodobacter capsulatus* HCP (Cabello *et al.*, 2004). However, the very high optimum pH (Cabello *et al.*, 2004) and high K_m at neutral pH (Aragão *et al.*, 2003) determined for the hydroxylamine reductase activity do not give complete confidence in the suggestion that the true physiological function of this protein is as a hydroxylamine reductase. More recently, it has been reported that HCP transcription (in *E. coli*) is regulated in response to nitric oxide or to the regulatory protein NsrR, which is involved in reactive nitrogen stress response (Filenko *et al.*, 2005, 2007; Flatley *et al.*, 2005). These findings are reinforced by other reports pointing to increased HCP transcription in response to nitrogen oxides in *Salmonella enterica* (Kim *et al.*, 2003), *Shewanella oneidensis* (Beliaev *et al.*, 2005) and *D. vulgaris* (He *et al.*, 2006). In light of these facts, protection against a reactive nitrogen compound other then hydroxylamine cannot be ruled out. The second hypothesis for HCP function has been linked to an oxidative stress response (Briolat & Reysset, 2002). Almeida *et al.* (2006) have reported HCP induction by hydrogen peroxide, with regulation by the redox-sensitive transcriptional activator OxyR, and peroxidase activity for *E. coli* and *D. desulfuricans* HCPs using ascorbate as the electron donor. Despite a K_m within the values measured for typical peroxidases, the low V_{max} obtained for HCPs from both *E. coli* and *D. desulfuricans* still poses doubts about the true physiological function of this protein. A search for the physiological electron donor and for other substrate oxidative species with or without nitrogen may lead to the true function.

HCPs share sequence and structure homology with anaerobic CoDs but not with aerobic CoDs, which use a molybdopterin cofactor. CoD is an enzyme responsible for the reduction of CO_2 to CO. Anaerobic CoDs can be a single monomer as in *Carboxydothermus hydrogenoformans* (Dobbek *et al.*, 2001) or a homodimer as in *Rhodospirillum rubrum* (Drennan *et al.*, 2001). In both cases there are three Fe–S clusters, although the monomeric enzymes have a hybrid [Ni–4Fe–4S] cluster (also called the C-cluster), a cubane [4Fe–4S] cluster (B-cluster) and a simple [2Fe–2S] cluster (D-cluster), whilst the dimer has the [2Fe–2S] cluster replaced by another [4Fe–4S] cubane with ligands from both monomers composing the dimer. In addition, the carbon monoxide dehydrogenase dimer can also form $\alpha_2\beta_2$ heterotetramers as in *Moorella thermoacetica* (Doukov *et al.*, 2002; Darnault *et al.*, 2003). The latter is a bifunctional enzyme with catalysis occurring at the C- and A-clusters. In the C-cluster (corresponding to the hybrid cluster in HCP), which is located in the β-subunit, CO is generated by reducing CO_2, whilst in the A-cluster, which is located in the α-subunit, CO is combined with CoA and a methyl group to form acetyl-CoA (Doukov *et al.*, 2002).

2. Materials and methods

2.1. Purification of *D. vulgaris* HCP

All purification, crystallization and associated steps were performed at 277 K in an anaerobic chamber (model A-2463) under a nitrogen atmosphere in the presence of less than 3 p.p.m. oxygen. Throughout the procedures, all possible attempts were made to avoid oxygen contamination of the samples. HCP was overexpressed using *D. vulgaris* cells harbouring the recombinant plasmid pJSP104. Cells were resuspended in a degassed solution of 10 m*M* Tris–HCl pH 7.6 and disrupted in an oxygen-free French press previously flushed with argon. Crude cell extract was obtained by ultracentrifugation at 40 000 rev min^{-1} for 5 h, followed by centrifugation at 18 000 rev min^{-1} for 1 h. The supernatant was applied onto a Q-Sepharose anion-exchange column (Pharmacia XK 26 × 40 Fast-Flow) pre-equilibrated with buffer *A* (20 m*M* Tris–HCl pH 7.6). The column was washed with 400 ml of the same buffer and a gradient to 100% buffer *B* (20 m*M* Tris–HCl pH 7.6, 400 m*M* NaCl) was applied (1.6 l at 4 ml min^{-1}). HCP eluted together with other contaminants at approximately 120 m*M* NaCl. The eluted fractions were immediately concentrated and desalted by ultrafiltration on an Amicon YM-30 membrane before application onto the next column, a ceramic hydroxyapatite HTP column (Pharmacia XK 3 × 28) previously equilibrated with buffer *C* (5 m*M* KH_2PO_4 pH 7.6). This column was washed with 500 ml buffer *C* and a linear gradient to buffer *D* (200 m*M* KH_2PO_4 pH 7.6) was then applied (500 ml at 0.8 ml min^{-1}). HCP eluted together with contaminants at approximately 40 m*M* KH_2PO_4. These fractions were desalted, transferred into buffer *A* and applied onto a Mono-Q column (Pharmacia Mono-Q). A linear gradient to 50% buffer *E* (20 m*M* Tris–HCl pH 7.6, 200 m*M* NaCl) was applied (120 ml at 4 ml min^{-1}). HCP eluted at approximately 50 m*M* NaCl. HCP-containing fractions were concentrated, dialysed against buffer *E* and then applied onto a Superdex 75 column (Pharmacia XK 16). The column was eluted with the same buffer at a flow rate of 0.3 ml min^{-1}. The purity was verified by a 15% SDS–PAGE gel and UV–Vis spectra. A total of 27.5 mg pure HCP protein was concentrated to 25 mg ml^{-1} and stored in buffer *F* (50 m*M* Tris–HCl pH 7.5, 50 m*M* NaCl).

UV–Vis spectra were recorded during the purification process, although it was difficult to assess the redox state of the HCP since there were other species and proteins present in the solution. However, the UV–Vis spectra of the protein after the final step of purification showed that the protein was not in the fully oxidized state. Moreover, if an aliquot of the protein was left to oxidize in air, its UV–Vis spectra changed to that expected for the oxidized state. EPR measurements were not taken and UV–Vis alone cannot distinguish the intermediate states. Thus, the precise oxidation state of the protein could

Table 1
Data-collection statistics for the as-isolated *D. vulgaris* HCP.

Values in parentheses are for the highest resolution shell.

	Inflection point	Peak	Remote	Highest resolution
ESRF beamline	ID29	ID29	ID29	ID14-2
Resolution limits (Å)	38.20–2.20 (2.32–2.20)	38.20–2.20 (2.32–2.20)	47.40–2.20 (2.32–2.20)	67.40–1.35 (1.36–1.35)
Wavelength (Å)	1.742	1.736	1.284	0.933
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters				
a (Å)	64.3	64.3	64.1	64.27
b (Å)	66.8	66.8	67.5	66.84
c (Å)	134.8	134.9	135.0	134.76
No. of images	224	224	224	535
Oscillation range (°)	0.5	0.5	0.5	0.2
Total reflections	111349 (8684)	112006 (9145)	110922 (9242)	710927
Unique reflections	26707 (2546)	26842 (2657)	26296 (2501)	127993
Multiplicity	4.1 (3.0)	4.1 (3.1)	4.1 (3.1)	5.5 (1.7)
$R_{\rm merge}$†	0.039 (0.062)	0.040 (0.062)	0.033 (0.039)	0.059 (0.35)
Average $I/\sigma(I)$	13.9 (9.3)	13.4 (9.0)	17.2 (14.9)	12.2 (3.5)
Completeness (%)	98.8 (98.8)	98.8 (98.8)	96.9 (92.4)	100.0 (100.0)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

not be established, except to note that it was not that of the fully oxidized protein.

2.2. Crystallization of *D. vulgaris* HCP

Initial crystallization trials based on the original conditions obtained for aerobic HCP (Arendsen *et al.*, 1998) were set up immediately using fresh (less then 5 d old) protein. All crystallization trials were performed with protein at a concentration of 25 mg ml^{-1} at 278 K. All buffers were previously degassed. Screens were made with 0.1 *M* MES buffer (pH range 6.0–6.5), PEG 4000 (concentration range 15–30%) and drops of protein plus precipitant as follows: 1.5 + 2.5 µl, 2 + 2 µl, 2 + 3 µl and 3 + 2 µl. The best crystals were obtained using 1.5 µl protein plus 2.5 µl 22% PEG 4000 and 0.1 *M* MES buffer pH 6.5. The crystals took about eight weeks to grow to maximum dimensions of 0.18 × 0.18 × 0.05 mm. A mother-liquor solution containing 15% glycerol was used as cryoprotectant and crystal mounting was undertaken within the anaerobic chamber. Crystals were transferred to the solution containing the cryoprotectant and then mounted on a loop prior to freezing. A door at the bottom of the anaerobic chamber was opened directly adjacent to a Dewar vessel containing liquid nitrogen and the crystals were frozen, typically in less than 15 s. Although this procedure was not strictly anaerobic, it is unlikely that the crystals were significantly affected by exposure to air.

2.3. Data collection

X-ray diffraction data were collected on ESRF beamline ID29 using X-rays with wavelengths of 1.742, 1.736 and 1.284 Å, corresponding to the inflection point and peak of the Fe *K* absorption edge and a high-energy remote point, respectively. For the three data sets, images of 0.5° oscillation and 5 s exposure for a total of 112° were taken, which resulted in diffraction patterns with reflections to better then 2.2 Å resolution. Data images were integrated with *MOSFLM* (Leslie, 2006) and scaled, merged and converted to structure factors with *SCALA* (Evans, 2006) and *TRUNCATE* (Collaborative Computational Project, Number 4, 1994). For structure refinement, data were collected using X-rays of 0.933 Å wavelength on ESRF beamline ID14-2 with 0.2° oscillation and 5 s exposure for a total of 107°, giving a 1.35 Å resolution data set. Data images were integrated with *DENZO* and reduced and scaled with *SCALEPACK* (Otwinowski & Minor, 1997). Data-collection statistics are given in Table 1.

The crystals belong to space group $P2_12_12_1$, with unit-cell parameters a = 64.27, b = 66.84, c = 134.76 Å, and accommodate one molecule of molecular weight approximately 60 kDa in the asymmetric unit, corresponding to a $V_{\rm M}$ of 2.07 Å^3 Da^{-1} (Matthews, 1968) and a solvent content of 41%.

2.4. Phase determination

The structure was determined by the multiwavelength anomalous diffraction (MAD) method using data sets collected to 2.2 Å resolution at the peak, inflection-point and remote wavelengths of the Fe *K* absorption edge. The anomalous signal of the Fe atoms within the structure was sufficient for automated interpretation of anomalous Patterson maps using *SOLVE* (Terwilliger, 1994*a*,*b*; Terwilliger & Berendzen, 1999), which located six of the eight irons expected to be present in HCP. The locations of the remaining two irons were determined using the corresponding anomalous difference map calculated with the phases obtained from the six-iron substructure (Fig. 1*b*). The eight iron positions were then recycled into *SOLVE* for further phase improvement.

2.5. Model construction and refinement

The phases obtained were extended in steps to 1.35 Å with *DM* (Cowtan, 1994; Cowtan & Main, 1998). Initial model building was performed automatically with *ARP*/*wARP* (Lamzin & Wilson, 1993; Perrakis *et al.*, 2001). After 50 cycles of refinement and ten cycles of model building, 525 residues of a possible total of 553 were constructed with a global connectivity index of 0.91. At the end of the automatic model building, the refinement converged to 28.5% and 30.2% for the *R* factor and $R_{\rm free}$, respectively. The amino-acid side chains were placed with the side-dock script from the *ARP*/*wARP* suite. After this procedure, $\sigma_{\rm A}$-weighted (Read, 1986) $2F_{\rm o} - F_{\rm c}$ and $F_{\rm o} - F_{\rm c}$ electron-density maps were inspected with *XTALVIEW* (McRee, 1992) and the missing sections were built. The iron–sulfur clusters were added with the help of the iron anomalous map to position the iron ions. The model obtained was then used to perform refinement with *SHELXL*

(Sheldrick, 2008). Positional and isotropic *B*-factor parameters were initially refined for each atom to a resolution limit of 1.35 Å. Alternative conformations for side chains were added where suggested by $|F_o| - |F_c|$ maps and finally accepted if the occupancy did not refine to below 0.25. 12 parameters for an overall anisotropic *B* factor were added (parameter HOPE in *SHELXL*) and when the refinement stabilized these were fixed. After this stage of refinement, solvent molecules were added to the model based on standard geometrical and chemical restraints; they were subsequently deleted if they were not visible at the 0.8 r.m.s. level in σ_A-weighted $2|F_o| - |F_c|$ electron-density maps. Refinement using anisotropic *B* factors on the Fe atoms initially and then extended to all atoms led to a decrease of 2.0% in *R* and 1.2% in R_{free} (see Table 2). After several cycles of refinement and manual rebuilding where necessary, the refinement converged to a final *R* factor of 12.8% (115 017 reflections) and an R_{free} of 17.7% (6422 reflections).

Table 2
R factor progress though refinement.

	R_{free} (%)	*R* factor (%)
ARP/wARP (no irons) + side chains	25.2	23.5
Add 8 Fe + 6 S atoms	24.8	23.0
SHELXL (HOPE) + 200 water molecules	23.2	19.6
Add 789 water molecules + 11 alternative side chains	19.4	15.1
SHELXL (all *B* factors anisotropic)	18.2	13.1
Final values	17.7	12.8

2.6. Three-dimensional structure comparison

All the HCP and the carbon monoxide dehydrogenase coordinate files deposited in the Protein Data Bank were aligned with the as-isolated *D. vulgaris* structure presented here using the secondary-structure matching algorithm (*SSM*; Krissinel & Henrick, 2004) implemented in *Coot* (Emsley & Cowtan, 2004). This initial fit was further refined with *MODELLER* (Sali & Blundell, 1993), in which residues are aligned by their C^{α} coordinates and considered to be structurally homologous if within a threshold of 3.5 Å. Two residues of different structures are considered to be structurally

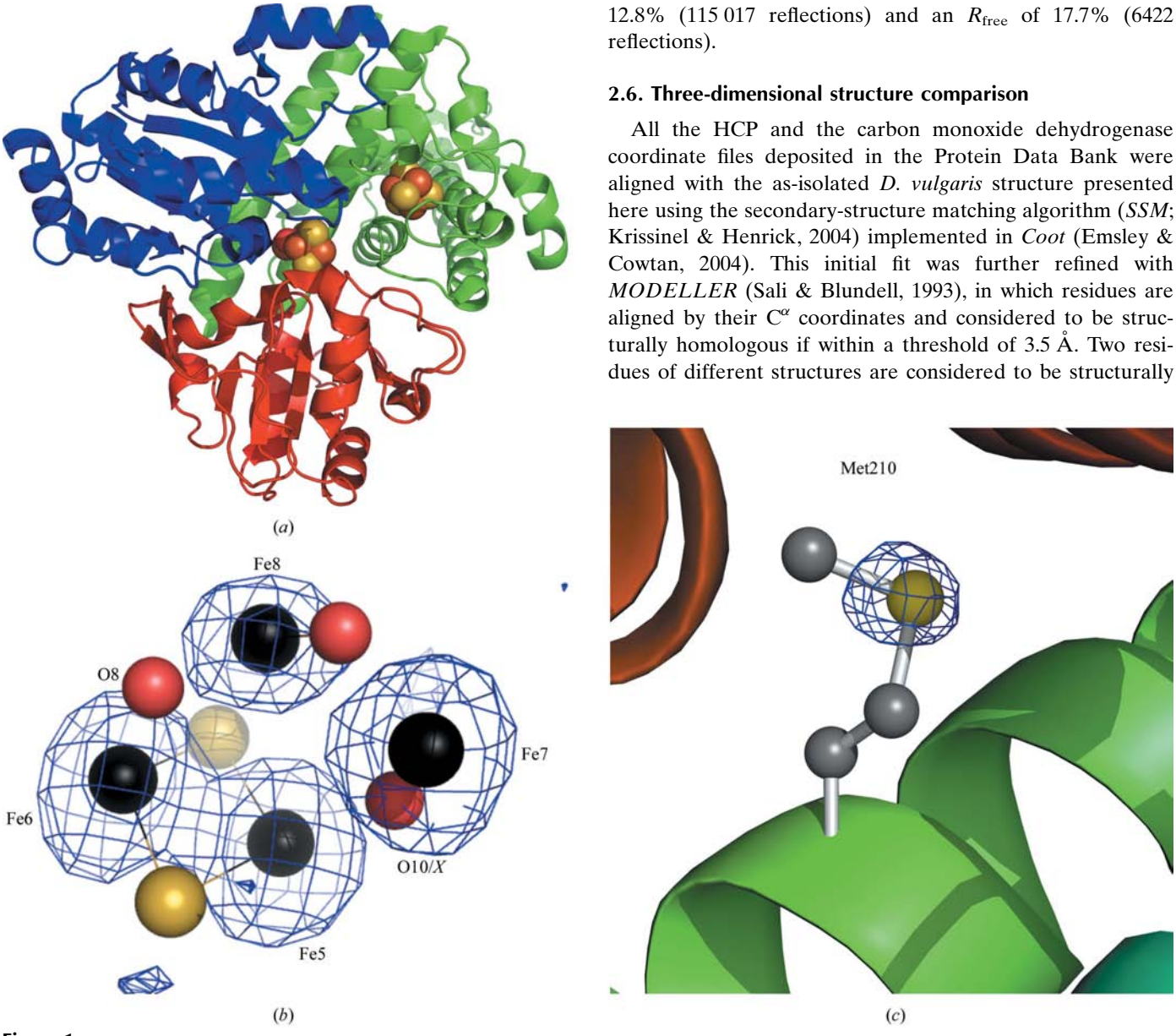

Figure 1
The overall topology of HCP. (*a*) Ribbon representation of as-isolated *D. vulgaris* HCP. The molecule is divided into three domains, numbered 1–3, which are coloured green, red and blue, respectively. The hybrid cluster lies at the centre of the three domains, whilst the cubane cluster is situated at the N-terminus in domain 1. (*b*) An anomalous difference synthesis using phases derived from the MAD experiment and contoured at the 6σ level in the vicinity of the hybrid cluster. The four Fe atoms are clearly identified. (*c*) An anomalous difference synthesis using phase information from the final refinement cycles in the vicinity of Met210, showing the sulfur moiety of this residue; the contour level is 3σ.

identical if they are of the same residue type and are structurally homologous.

3. Results and discussion

3.1. Quality of the structure

The stereochemistry of the final model was analysed with *PROCHECK* (Laskowski *et al.*, 1993), giving an overall *G* factor of 0.09 (see Table 3 for detailed refinement statistics). The Ramachandran plot (Ramakrishnan & Ramachandran, 1965) shows 93.7% (448) and 6.1% (19) of the relevant amino acids in the most favourable regions and additional allowed regions, respectively. As reported previously, Asn303 (Macedo *et al.*, 2002; Aragão *et al.*, 2003) lies in the generously allowed region but is supported by well defined electron density.

3.2. Structure

3.2.1. Overall structure. The protein has three domains: the N-terminal domain 1 comprising two bundles of three α-helices arranged almost perpendicular to one another and two similar domains, 2 and 3, comprising a central β-sheet flanked by α-helices in a typical $\alpha\beta$ Rossmann fold-like geometry (Fig. 1*a*). The protein contains two iron–sulfur clusters that are approximately 11 Å apart. A conventional [4Fe–4S] cubane cluster binds through its Fe ions to four cysteine residues at the N-terminus of the protein and a

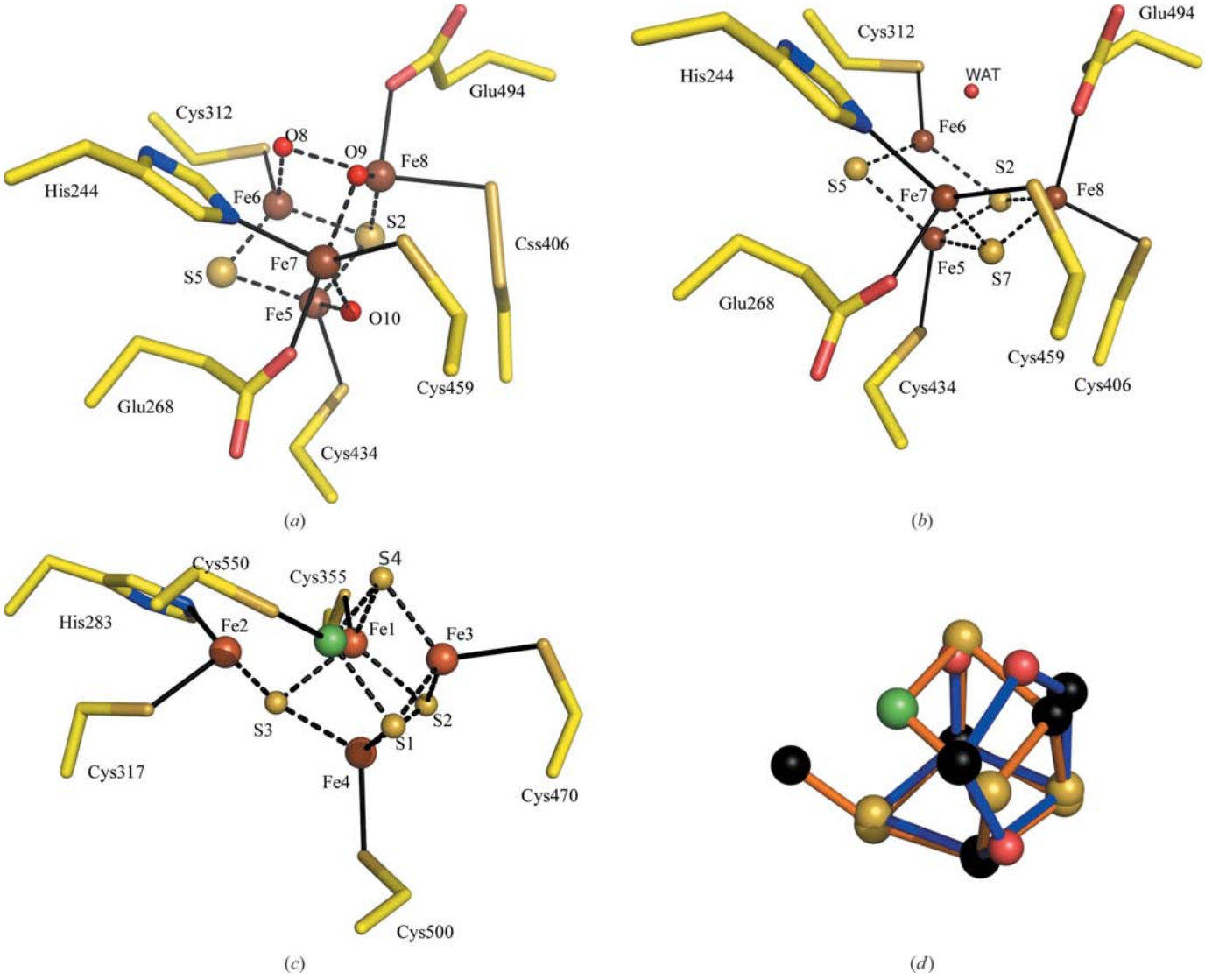

Figure 2
The hybrid clusters in HCP and the CoD β-subunit. (*a*) Model representation of the hybrid cluster in the as-isolated oxidized form of *D. vulgaris* HCP prepared anaerobically (this work; PDB code 1w9m). Bonds associated with the metal cluster are shown as dotted lines, whereas metal–residue ligand interactions are drawn as full lines; only residue side chains are shown for clarity. (*b*) Model representation of the hybrid cluster in the *D. vulgaris* HCP reduced form (PDB code 1oa1). (*c*) Model representation of the C-cluster in *M. thermoacetica* CoD (PDB code 1mjg). (*d*) Superposition of the anaerobically purified HCP hybrid cluster (bonds in blue) and CoD C-cluster (bonds in orange). The ligands binding Fe5 (Cys434), Fe6 (Cys312) and Fe8 (Cys406) in HCP are structurally identical to the ligands binding Fe4 (Cys500), Fe1 (Cys355) and Fe3 (Cys470) in the *M. thermoacetica* CoD C-cluster (see also Table 4 and Fig. 4). Atom colouring: nickel, green; nitrogen, blue; oxygen, red; sulphur, yellow; iron brown in (*a*), (*b*) and (*c*) and black in (*d*).

Table 3
Final refinement statistics.

PDB code	1w9m
Amino acids	553
Protein atoms	4259
Solvent atoms	789
Other atoms	1 × S_4Fe_4, 1 × $S_2Fe_2O_3$
Resolution limits (Å)	67.4–1.35
R factor (No. of observations)	0.128 (115017)
R_{free} (No. of observations)	0.177 (6422)
Highest resolution shell (Å)	1.36–1.35
Cruickshank's DPI (Å)	0.048
Average B_{iso} (Å^2)	
All atoms	16.3
Protein atoms	13.6
Main chain	11.8
Side chains	15.4
Solvent	30.8
Residues with dual conformations	Ile35, Asp60, Thr173, Glu174, Asp189, Val314, Lys338, Pro363, Glu366, Asn533, Ile536
R.m.s. distance deviations	
Bond lengths (Å)	0.012
Angle distances (Å)	0.027
Distances from restraint planes (Å)	0.026

second unusual and so-called hybrid cluster binds another four cysteine residues located near the geometrical centre of the protein and at the interface of the three domains. The hybrid cluster [4Fe–2S–3O] is composed of a square-like base with two Fe atoms (Fe5 and Fe6) and two S atoms (S2 and S5), with the irons bound to cysteine residues Cys434 and Cys312, and a cap with two Fe atoms (Fe7 and Fe8) and three O atoms (O8, O9 and O10), with Fe7 bound to His244, Glu268 and Cys459 and Fe8 bound to Cys406 through a persulfide bond and Glu494 (Fig. 2*a*). To confirm the nature of the atoms in the cluster, a phased anomalous difference map was calculated using data taken at the peak of the Fe *K* absorption edge. As expected, all the Fe atoms and the two S atoms S2 and S5 were visible in a phased anomalous difference map, whilst the O atoms O8, O9 and O10 were not (the oxygen anomalous signal is negligible at this wavelength). Moreover, these O atoms refined to similar *B*-factor values to the atoms to which they are bound. In the initial *D. vulgaris* structure (Arendsen *et al.*, 1998), which was prepared aerobically, O10 was modelled as a single O atom. Subsequently, Cooper *et al.* (2000) proposed an O atom disordered between two distinct positions around the coordinates of O10. Finally, structures at 1.25 Å resolution (Macedo *et al.*, 2002) showed an elongated electron-density peak between Fe5 and Fe7 for aerobically prepared *D. vulgaris*, whilst the anaerobically prepared *D. desulfuricans* structure showed a single position for this O atom. It was previously reported that the major difference between *D. desulfuricans* HCP and *D. vulgaris* HCP lay in the O10 moiety bound to Fe5 and Fe7 (Macedo *et al.*, 2002; Aragão *et al.*, 2003). The careful anaerobic protein preparation and crystallization presented here clearly shows that the hybrid cluster of *D. vulgaris* HCP is not an oxidation artefact arising from aerobic purification and crystallization; the density for the O10 position of the hybrid cluster is readily modelled by a single O atom. The hybrid centre can have four oxidation states, ranging from the most oxidized state (+6) to the fully reduced state (+3) with three Fe^{II} ions and one Fe^{III} ion (Macedo *et al.*, 2002). Upon successive one-electron reduction and starting from the fully oxidized protein, four redox states are detected for the hybrid cluster. It was shown that the anaerobically purified *D. desulfuricans* HCP is EPR silent and it was therefore suggested to have its hybrid cluster in the +4 oxidation state (Macedo *et al.*, 2002) with a single oxygen position for O10. *D. vulgaris* HCP also shows a single position for this oxygen corresponding to the +4 oxidation state. The hybrid cluster for the aerobically purified *D. vulgaris* HCP has an O atom that is disordered in two positions and appears to be in the +5 oxidation state (Arendsen *et al.*, 1998; Macedo *et al.*, 2002). Thus, the nature of the electron density at the O10

	1e1d	1e2u	1e9v	1gnt	1oa1	1w9m	1upx	1gn9	1gnl	1oa0	1su6	1su7	1su8	1suf	1jqk	1mjg	1oao
1e1d	—	100	100	100	100	100	95	95	95	95	58	58	59	59	58	57	58
1e2u	100	—	100	100	100	100	95	95	95	95	58	58	59	59	58	57	58
1e9v	99	99	—	100	100	100	95	95	95	95	59	59	59	59	59	58	58
1gnt	99	99	100	—	100	100	95	95	95	95	59	59	59	59	59	58	58
1oa1	99	99	100	100	—	100	95	95	95	95	59	59	59	59	59	58	58
1w9m	99	99	100	100	100	—	95	95	95	95	59	59	59	59	59	57	58
1upx	65	65	65	66	66	65	—	99	99	100	59	59	60	60	59	59	59
1gn9	66	66	66	66	66	66	100	—	100	100	60	60	60	60	59	59	59
1gnl	66	66	66	66	66	66	100	100	—	100	60	60	60	60	59	59	59
1oa0	65	65	65	66	66	65	100	100	100	—	60	59	60	59	59	58	58
1su6	10	10	10	10	10	10	10	10	11	10	—	100	100	100	95	97	97
1su7	10	10	10	10	10	10	10	10	11	10	100	—		100	95	97	97
1su8	10	10	10	10	10	10	10	10	11	10	100	100	—	100	95	97	97
1suf	10	10	10	10	10	10	10	10	11	10	100	100	100	—	95	97	97
1jqk	10	10	10	10	10	10	11	11	11	11	57	57	57	57	—	98	98
1mjg	9	9	9	9	9	9	8	8	8	8	49	49	49	49	47	—	100
1oao	9	9	9	9	9	9	8	8	8	8	49	49	49	49	47	100	—

Figure 3
Structure similarity and identity between HCPs and CoD β-subunits. Values above the diagonal represent the percentage of structural similarity between proteins (an amino acid is considered to be structurally homologous to another amino acid when after a C^{α} protein-to-protein fit their C^{α} are 3.5 Å or less from each other). Values below the diagonal represent the percentage of structural homology between proteins (an amino acid is considered to be structurally identical when it is structurally homologous and the same type of amino acid). The *D. vulgaris* HCP structures are shown in dark grey (PDB codes 1e1d, 1e2u, 1e9v, 1gnt, 1oa1 and 1w9m), the *D. desulfuricans* HCP structures are in light grey (PDB codes 1upx, 1gn9, 1gnl and 1oa0), the *C. hydrogenoformans* CoD structures in darker orange (PDB codes 1su6, 1su7, 1su8 and 1suf), the *R. rubrum* CoD structure in orange (PDB code 1jqk) and the *M. thermoacetica* CoD structures in light orange (PDB codes 1mjg and 1oao).

position appears to depend on the oxidation state of the enzyme and this in turn probably depends on the catalytic activity of the enzyme, although this activity has yet to be defined.

3.2.2. Comparison with other structures. It has been reported that both *D. vulgaris* and *D. desulfuricans* strains have carbon monoxide dehydrogenase activity (Meyer & Rohde, 1984) and that CoDs share both fold and hybrid cluster similarity with HCP (Dobbek *et al.*, 2001; Macedo *et al.*, 2002). CoDs are present in three different arrangements: monomeric (*C. hydrogenoformans*), dimeric (*R. rubrum*) and tetrameric bifunctional (*M. thermoacetica*). The tetrameric form is an $\alpha_2\beta_2$ heterotetramer that catalyses two different functions: the β-subunit generates CO from CO_2, while in the α-subunit CO is combined with CoA and a methyl group to form acetyl-CoA (Ragsdale & Wood, 1985; Lu *et al.*, 1990).

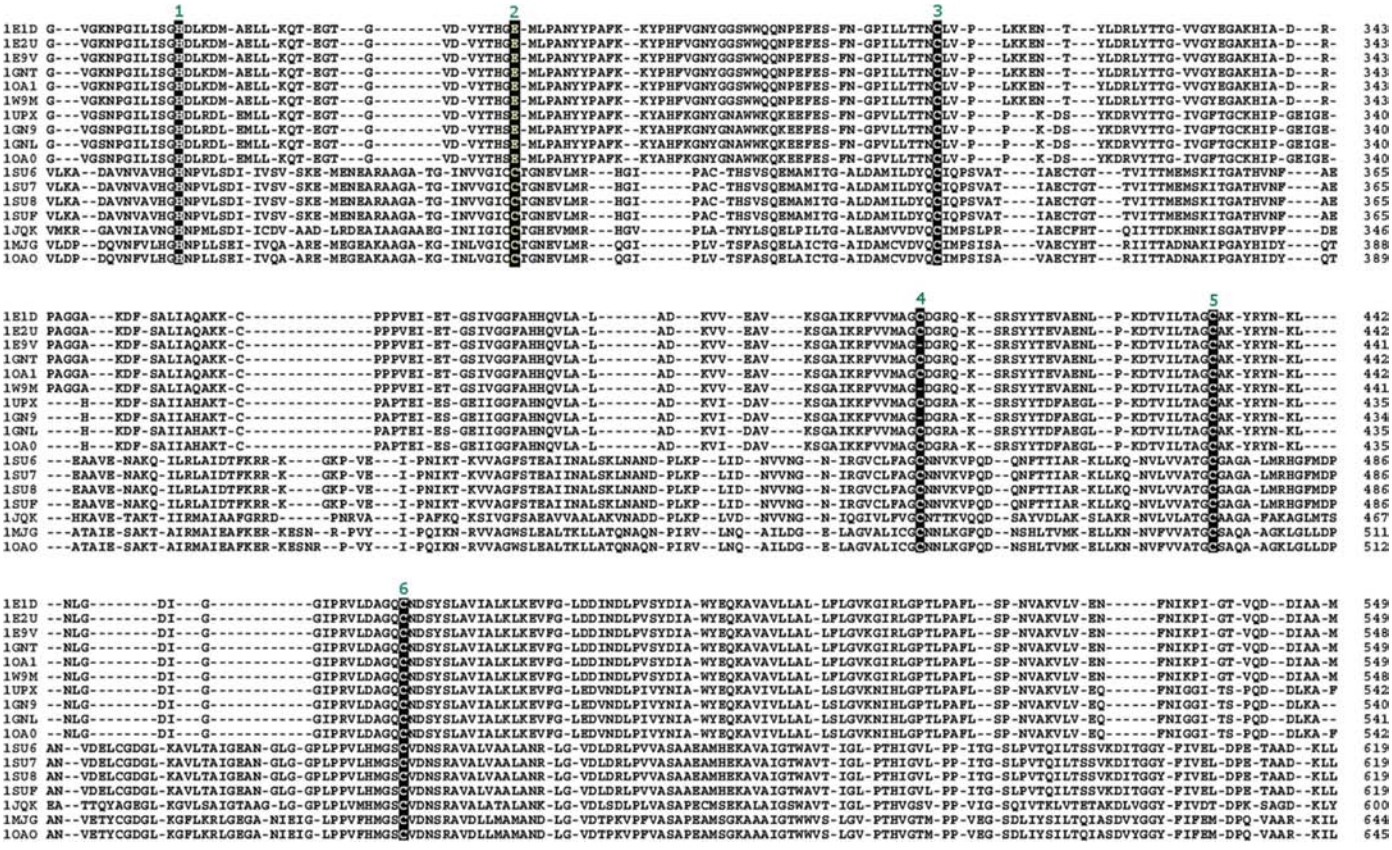

Figure 4
Three-dimensional structure alignment. Alignment of HCP with monomeric and CoD β-subunit structures after superposition with *MODELLER* (Sali & Blundell, 1993). Residues boxed in black are ligands to the hybrid cluster and the C-cluster in HCP and CoD, respectively (see Table 4 for residue numbering). HCP PDB codes 1e1d, 1e2u, 1e9v, 1gnt, 1oa1 and 1w9m are from *D. vulgaris*, and 1upx, 1gn9, 1gnl are from *D. desulfuricans*; CoD PDB codes 1su6, 1su7, 1su8 and 1suf are from *C. hydrogenoformans*, 1jqk is from *R. rubrum*, and 1mjg and 1oao are from *M. thermoacetica*.

	1e1d	1e2u	1e9v	1gnt	1oa1	1w9m	1upx	1gn9	1gnl	1oa0	1su6	1su7	1su8	1suf	1jqk	1mjg	1oao
1e1d	—	0	0.3	0.3	0.4	0.4	0.9	0.8	0.8	0.8	1.9	1.9	1.9	1.9	1.9	1.9	1.9
1e2u	—	—	0.3	0.3	0.4	0.4	0.9	0.8	0.8	0.8	1.9	1.9	1.9	1.9	1.9	1.9	1.9
1e9v	—	—	—	0.1	0.3	0.2	0.8	0.8	0.8	0.8	1.9	1.9	1.9	1.9	1.9	1.9	1.9
1gnt	—	—	—	—	0.3	0.3	0.8	0.8	0.8	0.8	1.9	1.9	1.9	1.9	1.9	1.9	1.9
1oa1	—	—	—	—	—	0.2	0.7	0.7	0.7	0.7	1.9	1.9	1.9	1.9	1.9	1.9	1.9
1w9m	—	—	—	—	—	—	0.8	0.8	0.7	0.8	1.9	1.9	1.9	1.9	1.9	1.9	1.9
1upx	—	—	—	—	—	—	—	0.2	0.1	0.2	1.9	1.9	1.9	1.9	1.9	1.9	1.9
1gn9	—	—	—	—	—	—	—	—	0.2	0.2	1.9	1.9	1.9	1.9	1.9	1.9	0.0
1gnl	—	—	—	—	—	—	—	—	—	0.2	1.9	1.9	1.9	1.9	1.9	1.9	1.9
1oa0	—	—	—	—	—	—	—	—	—	—	1.9	1.9	1.9	1.9	1.9	1.9	1.9
1su6	—	—	—	—	—	—	—	—	—	—	—	0.3	0.1	0.1	0.8	0.8	0.8
1su7	—	—	—	—	—	—	—	—	—	—	—	—	0.2	0.3	0.8	0.8	0.8
1su8	—	—	—	—	—	—	—	—	—	—	—	—	—	0.2	0.8	0.8	0.8
1suf	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0.8	0.8	0.8
1jqk	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0.9	1.0
1mjg	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0.3
1oao	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—

Figure 5
Root-mean-square differences (Å) between homologous C^{α} atoms in HCPs and CoD β-subunits. The same colour code is used as in Fig. 3.

Table 4
Residues acting as ligands to the hybrid clusters in HCP and CoD β-subunit.

	HCP		CoD		
	D. vulgaris	*D. desulfuricans*	*C. hydrogenoformans*	*R. rubrum*	*M. thermoacetica*
1	His244	His240	His261	His265	His283
2	Glu268	Glu264	Cys295	Cys300	Cys317
3	Cys312	Cys308	Cys333	Cys338	Cys355
4	Cys406	Cys399	Cys446	Cys451	Cys470
5	Cys434	Cys427	Cys476	Cys481	Cys500
6	Cys459	Cys452	Cys526	Cys531	Cys550

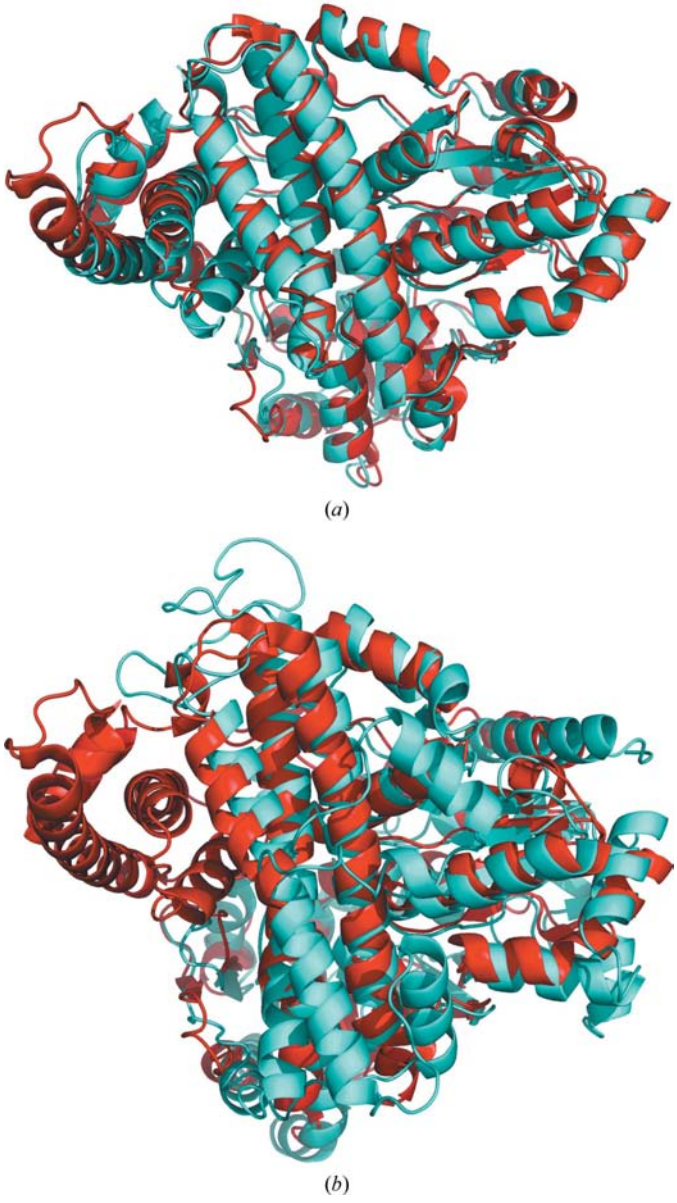

Figure 6
Superposition of HCP and CoD β-subunits. (*a*) Superposition of as-isolated *D. vulgaris* (red ribbon) and *D. desulfuricans* HCP (cyan coil; PDB code 1upx) structures. (*b*) Superposition of as-isolated *D. vulgaris* HCP (red ribbon) and the β-domain of a carbon monoxide dehydrogenase structure (cyan ribbon; PDB code 1su7), showing their overall similar fold, but with the first three-helix bundle of the CoD missing (left-hand side of the figure).

The β-subunits have the same arrangement as the dimeric CoD structure and are structurally related to HCP. This subunit comprises three domains, one of which contains only one three-helix bundle, while the other two similar domains have a central β-sheet flanked by α-helices (*c.f.* Fig. 1*a*). CoD β-subunits have a cubane-like Fe–S cluster (B-cluster) and a variable Fe–S cluster (D-cluster) bound to the first domain of the two monomers, with a nickel-containing hybrid Fe–S cluster (C-cluster) at the interface of the three domains of each monomer.

The *D. vulgaris* and *D. desulfuricans* HCP structures share over 60% amino-acid identity based on their structural alignment (Figs. 3 and 4[1]) and have a maximum root-mean-square difference of 0.4 Å for structurally homologous C^{α} atoms within *D. vulgaris* structures and 0.8 Å between *D. vulgaris* and *D. desulfuricans* structures (Fig. 5). Fig. 6(*a*) shows a superposition of the current structure (PDB code 1w9m) with that of a *D. desulfuricans* structure (PDB code 1upx). However, HCP and the CoD β-subunits share 10% or less structural amino-acid identity (Figs. 3 and 4) and have a maximum root-mean-square displacement of some 1.9 Å between homologous C^{α} atoms (Fig. 5). Remarkably, however, the overall folds of HCP and CoD β-subunits are similar (Fig. 6*b*) and their superposition shows about 59% of the residues to be structurally homologous (Fig. 3). One of the major differences is that the CoD β-subunits lack one of the three-helix bundles that comprise domain 1 of HCP.

Domain 1 is responsible for the binding of one and two $[Fe_4–S_4]$ cubane clusters in HCP and CoD β-subunits, respectively. Although the CoD monomer lacks the first three-helix bundle that is responsible for the binding of the cubane cluster in HCP, this bundle is replaced in multimeric CoDs (β-subunit of *R. rubrum* CoD, bifunctional CoD of *M. thermoacetica*) by an equivalent domain from a different monomer in order to create a similar arrangement (Fig. 7). Domains 2 and 3 contribute with ligands to the [4Fe–2S–3O] hybrid cluster in HCP and to the [Ni–4Fe–4S] C-cluster in the CoD β-subunit. The CoD C-cluster can be described as a [3Fe–4S] moiety and an [Ni–Fe] subsite. Two of the three facial sulfides of the [3Fe–4S] moiety coordinate the Ni, whilst the third coordinates the unique Fe of the [Ni–Fe] subsite (Doukov *et al.*, 2002; Kim *et al.*, 2004; Volbeda & Fontecilla-Camps, 2005). The fit of the two hybrid clusters after superposition of HCP and the CoD β-subunit shows that six out of the nine atoms overlap (Fig. 2*d*): HCP atoms Fe5, Fe6, Fe8, O8, S5 and S6 (Fe4, Fe1, Fe3, S4, S3 and S2 in *M. thermoacetica* CoD numbering). Fe atoms Fe5, Fe6 and Fe8 (HCP numbering) are bound by structurally identical cysteine residues (Fig. 2, Table 4). The equivalent position to O10 of HCP

[1] Only a portion of the alignment covering the residues binding to the hybrid clusters is shown in Fig. 4. The full alignment has been deposited in the IUCr electronic archive as supplementary material (Reference: HV5104). Services for accessing this material are described at the back of the journal.

is occupied by an S atom in the CoD β-subunit. This is the configuration found in reduced HCP structures (Aragão *et al.*, 2003). CoD structural studies have clearly shown catalysis to occur on the top and mobile part of the C-cluster in the vicinity of Ni and Fe2 (Darnault *et al.*, 2003; Lindahl, 2004; Volbeda & Fontecilla-Camps, 2005). The positions of the Ni and Fe2 atoms in the CoD β-subunit are vacant in HCP structures.

This structural similarity between HCP and CoD β-subunits may hint at some degree of similarity in their catalytic mechanism. In HCP this would therefore involve the substrate binding to the hybrid cluster reduced structure in the vicinity of the water molecule shown in Fig. 2(*b*) and forming hydrogen bonds to the conserved residues Asn311 (Gln in the CoD β-subunit) and Lys496 (Aragão *et al.*, 2003). In the case of a dioxygen-containing substrate, reduction would yield two hydroxyl species, giving rise to the oxygen positions O8 and O9, as in the anaerobically purified oxidized form (Fig. 2*a*). Sequential release of these would lead to a water product at or in the vicinity of the O10 position. The product could then diffuse away from the active site and eventually away from the enzyme through the extensive hydrophobic cavity that gives ready access to the hybrid cluster (Cooper *et al.*, 2000).

It is also intriguing to note that class 3 HCPs, which are found in hyperthermophilic bacteria and archaea (van Den Berg *et al.*, 2000; Bailey *et al.*, 2001), also have a 116-residue deletion downstream of the N-terminal cysteine region, corresponding to one of the three-helix bundles. It is possible that these HCPs adopt a similar structural conformation to that shown in Fig. 7.

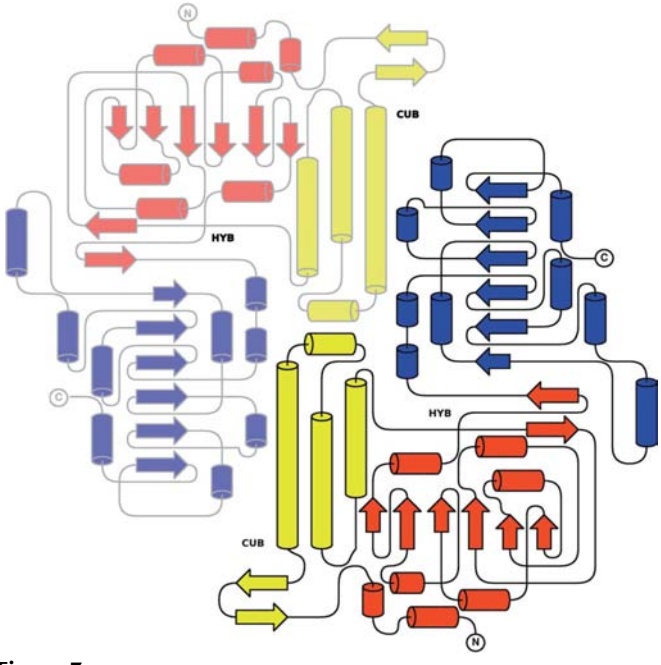

Figure 7
Topology diagram for a dimer of the CoD β-subunit and class 3 HCP. A typical topology diagram is shown for a dimer of the β-subunit of carbon monoxide dehydrogenase, with domains 1, 2 and 3 coloured yellow, red and blue, respectively, for domain identification and comparison with HCPs. Domain 1 of the carbon monoxide dehydrogenase monomer lacks one of the two three-helix bundles that are present in the hybrid cluster proteins. However, dimerization allows the construction of a complete domain 1 shared between two monomers. This may also represent a structural model for class 3 HCPs, which lack this helix bundle. The CoD D-cluster is located at the interface of the two monomers and has been omitted for clarity.

4. Conclusions

Iron MAD phasing allowed the complete structure elucidation of the anaerobically purified HCP protein and its iron–sulfur cluster. Structures of anaerobically purified HCP from both *D. desulfuricans* (Macedo *et al.*, 2002) and *D. vulgaris* (this study) show a single oxygen position for O10, whilst structures determined using aerobically prepared samples have an elongated density at the same position. In some cases this density has been modelled as two half-occupied oxygen positions.

The main sequence differences between these HCP and CoD proteins lie in the N-terminal domain, where the cubane cluster (B-cluster in CoD) binding region is apparently absent for the CoD β-subunit. However, the structures of dimeric CoD and the β-subunit of tetrameric CoD show that this region is present *via* monomer *B* and that it has the same fold as in HCP. Both HCP and CoD have intriguing and unusual clusters at the interface of their three domains. These clusters, although different, share striking similarities and are linked to the polypeptide chain *via* homologous ligands. The HCP reaction mechanism may also be similar to the CoD reaction mechanism, with substrate binding to the top mobile part of the cluster followed by its reduction and release of an oxygen ion or hydroxyl group. This mechanism would fit both hydroxylamine reduction and hydrogen peroxide reduction as well as other small or linear substrates with the potential to be reduced. In the light of current knowledge, the HCP substrate could be either a reactive nitrogen or oxygen species and further biochemical studies are needed to access and understand the true nature of the substrate of this enzyme in order to clarify its role in either protection against oxidative stress or reactive nitrogen stress.

The European Synchrotron Radiation Facility at Grenoble, France is gratefully acknowledged for the provision of synchrotron facilities; in particular, the staff of ESRF stations ID-14 and ID-29 are acknowledged for their support and assistance in X-ray diffraction data collection. DA acknowledges grant BD/6480/2002/ from the Fundação para a Ciência e Tecnologia (FCT), Portugal. This work was funded in part by the ESRF, FEDER and the FCT, Portugal through grants POCTI/BME/38859/2001, POCTI/BME/44441/2002 and POCTI/BIO/58041/2004. Célia V. Romão and Dora Alves are thanked for their assistance with the anaerobic purification of HCP from *D. vulgaris*.

References

Almeida, C. C., Romão, C. V., Lindley, P. F., Teixeira, M. & Saraiva, L. M. (2006). *J. Biol. Chem.* **281**, 32445–32450.

Aragão, D., Macedo, S., Mitchell, E. P., Romão, C. V., Liu, M. Y., Frazão, C., Saraiva, L. M., Xavier, A. V., LeGall, J., van Dongen, W. M., Hagen, W. R., Teixeira, M., Carrondo, M. A. & Lindley, P. (2003). *J. Inorg. Biol. Chem.* **8**, 540–548.
Arendsen, A. F. *et al.* (1998). *J. Biol. Inorg. Chem.* **3**, 81–95.
Bailey, S., Cooper, S. J., Hagen, W. R., Arendsen, A. F. & Lindley, P. F. (2001). *Handbook of Metalloproteins*, Vol. 1, edited by A. Messerschmidt, R. Huber, T. Poulos & K. Weighardt, pp. 593–601. New York: John Wiley & Sons.
Beliaev, A. S., Klingeman, D. M., Klappenbach, J. A., Wu, L., Romine, M. F., Tiedje, J. M., Nealson, K. H., Fredrickson, J. K. & Zhou, J. (2005). *OMICS*, **6**, 39–60.
Briolat, V. & Reysset, G. (2002). *J. Bacteriol.* **184**, 2333–2343.
Cabello, P., Pino, C., Olmo-Mira, M. F., Castillo, F., Roldan, M. D. & Moreno-Vivian, C. (2004). *J. Biol. Chem.* **279**, 45485–45494.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cooper, S. J., Garner, C. D., Hagen, W. R., Lindley, P. F. & Bailey, S. (2000). *Biochemistry*, **39**, 15044–15054.
Cowtan, K. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 34–38.
Cowtan, K. & Main, P. (1998). *Acta Cryst.* D**54**, 487–493.
Darnault, C., Volbeda, A., Kim, E. J., Legrand, P., Vernede, X., Lindahl, P. A. & Fontecilla-Camps, J. C. (2003). *Nature Struct. Biol.* **10**, 271–279.
Den Berg, W. A. M. van, Hagen, W. R. & van Dongen, W. M. A. M. (2000). *Eur. J. Biochem.* **267**, 666–676.
Dobbek, H., Svetlitchnyi, V., Gremer, L., Huber, R. & Meyer, G. (2001). *Science*, **293**, 1281–1285.
Doukov, T. I., Iverson, T. M., Seravalli, J., Ragsdale, S. W. & Drennan, C. L. (2002). *Science*, **298**, 567–572.
Drennan, C. L., Heo, J., Sintchak, M. D., Schreiter, E. & Ludden, P. W. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 11973–11978.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Filenko, N. A., Browning, D. F. & Cole, J. A. (2005). *Biochem. Soc. Trans.* **33**, 195–197.
Filenko, N., Spiro, S., Browning, D. F., Squire, D., Overton, T. W., Cole, J. & Constantinidou, C. (2007). *J. Bacteriol.* **189**, 4410–4417.
Flatley, J., Barrett, J., Pullan, S. T., Hughes, M. N., Green, J. & Poole, R. K. (2005). *J. Biol. Chem.* **280**, 10065–10072.
Hagen, W. R. (1989). *J. Chem. Soc. Faraday Trans.* **85**, 4083–4090.
He, Q., Huang, K. H., He, Z., Alm, E. J., Fields, M. W., Hazen, T. C., Arkin, A. P., Wall, J. D. & Zhou, J. (2006). *Appl. Environ. Microbiol.* **72**, 4370–4381.
Heo, J., Wolfe, M. T., Staples, M. R. & Ludden, P. W. (2002). *J. Bacteriol.* **184**, 5894–5897.
Kim, C. C., Monack, D. & Falkow, S. (2003). *Infect. Immun.* **71**, 3196–3205.
Kim, E. J., Feng, J., Bramlett, M. R. & Lindahl, P. A. (2004). *Biochemistry*, **43**, 5728–5734.
Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.
Lamzin, V. S. & Wilson, K. S. (1993). *Acta Cryst.* D**49**, 129–147.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Leslie, A. G. W. (2006). *Acta Cryst.* D**62**, 48–57.
Lindahl, P. A. (2004). *J. Biol. Inorg. Chem.* **9**, 516–524.
Lu, W. P., Harder, S. R. & Ragsdale, S. W. (1990). *J. Biol. Chem.* **265**, 3124–3133.
Macedo, S., Mitchell, E. P., Romão, C. V., Cooper, S. J., Coelho, R., Liu, M. Y., Xavier, A. V., LeGall, J., Bailey, S., Garner, C. D., Hagen, W. R., Teixeira, M., Carrondo, M. A. & Lindley, P. (2002). *J. Biol. Inorg. Chem.* **7**, 514–525.
McRee, D. E. (1992). *J. Mol. Graph.* **10**, 44–46.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Meyer, O. & Rohde, M. (1984). *Microbial Growth on C1 Compounds*, edited by R. L. Crawford & R. S. Hanson, pp. 26–33. Washington: American Society for Microbiology.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Perrakis, A., Harkiolaki, M., Wilson, K. S. & Lamzin, V. S. (2001). *Acta Cryst.* D**57**, 1445–1450.
Ragsdale, S. W. & Wood, H. G. (1985). *J. Biol. Chem.* **260**, 3970–3977.
Ramakrishnan, C. & Ramachandran, G. N. (1965). *Biophys. J.* **5**, 909–933.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
Sali, A. & Blundell, T. L. (1993). *J. Mol. Biol.* **234**, 779–815.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Terwilliger, T. C. (1994*a*). *Acta Cryst.* D**50**, 11–16.
Terwilliger, T. C. (1994*b*). *Acta Cryst.* D**50**, 17–23.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Volbeda, A. & Fontecilla-Camps, J. C. (2005). *Dalton Trans.* **21**, 3443–3450.
Wolfe, M. T., Heo, J., Garavelli, J. S. & Ludden, P. W. (2002). *J. Bacteriol.* **184**, 5898–5902.