Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The 2.6 Å resolution structure of *Rhodobacter capsulatus* bacterioferritin with metal-free dinuclear site and heme iron in a crystallographic 'special position'

D. Cobessi,[a] L.-S. Huang,[a] M. Ban,[a] N. G. Pon,[a] F. Daldal[b] and E. A. Berry[a]*

[a]Physical Biosciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley CA 94720, USA, and [b]Department of Biology, Plant Science Institute, University of Pennsylvania, Philadelphia, PA 19104-6018, USA

Correspondence e-mail: eaberry@lbl.gov

Received 26 April 2001
Accepted 15 October 2001

PDB Reference: bacterioferritin, 1jgc.

Bacterioferritin from *Rhodobacter capsulatus* was crystallized and its structure was solved at 2.6 Å resolution. This first structure of a bacterioferritin from a photosynthetic organism is a spherical particle of 24 subunits displaying 432 point-group symmetry like ferritin and bacterioferritin from *Escherichia coli*. Crystallized in the *I*422 space group, its structural analysis reveals for the first time the non-symmetric heme molecule located on a twofold crystallographic symmetry axis. Other hemes of the protomer are situated on twofold noncrystallographic axes. Apparently, both types of sites bind heme in two orientations, leading to an average structure consisting of a symmetric 50:50 mixture, thus satisfying the crystallographic and noncrystallographic symmetry of the crystal. Five water molecules are situated close to the heme, which is bound in a hydrophobic pocket and axially coordinated by two crystallographic or noncrystallographically related methionine residues. Its ferroxidase center, in which Fe^{II} is oxidized to Fe^{III}, is empty or fractionally occupied by a metal ion. Two positions are observed for the coordinating Glu18 side chain instead of one in the *E. coli* enzyme in which the site is occupied. This result suggests that the orientation of the Glu18 side chain could be constrained by this interaction.

1. Introduction

Ferritins and bacterioferritins (Bfr) are enzymes found in all organisms: eukarya, bacteria and archaea. These spherical particles of an inner 8 nm diameter are iron-storage proteins that oxidize Fe^{2+} to Fe^{3+} (ferroxidase activity). Fe^{3+} is next stored as an inorganic complex in the cavity of the spherical shell. Ferritins and bacterioferritins are protomers of 24 subunits that display 432 point-group symmetry. Two protein chains called H and L are present in the protomer in vertebrate enzymes, whereas only chain H (with the ferroxidase site) is present in bacterioferritin. In eukarya, ferritin stores Fe^{3+}, which is less toxic than Fe^{2+}, therefore reducing its availability to react with the superoxide ion, a reaction leading to the production of reactive hydroxyl radical. In bacteria, Bfr acts as an anion detoxifier (Harrison & Arosio, 1996). Another difference between ferritins and Bfr is the presence in the bacterioferritins of an iron protoporphyrin IX or coproporphyrin whose role still remains unknown (Andrews *et al.*, 1995; Romao *et al.*, 2000). In the structure from the *E. coli* enzyme, the Fe atom of the heme group is axially coordinated to the S atom of a methionine residue from each of two symmetrically related subunits (Frolow *et al.*, 1994). Although the percentage sequence identity between these proteins is low (around

20%), the overall fold of the subunit is similar and consists of a bundle of four antiparallel α-helices enclosing the ferroxidase center (also called the dinuclear center) that binds two metal ions (Frolow *et al.*, 1994). The ferroxidase center differs between the ferritins and bacterioferritins by two residues thought to be involved in the iron binding (Yang *et al.*, 2000).

At this time, the three-dimensional structure of bacterioferritin is known from only one species, *E. coli*. This structure is available at 2.9 Å resolution in two different space groups (Frolow *et al.*, 1994; Dautant *et al.*, 1998). In both of these structures, the 12 heme groups bound to the 24 subunits of the protomer are situated on twofold non-crystallographic axes and the ferroxidase center is occupied by two manganese ions from the crystallization medium.

In the present study, we report the structure of the *R. capsulatus* bacterioferritin solved at 2.6 Å resolution in the *I*422 space group and compare it with that from *E. coli*. The structure was solved by the MIR method independently from previous structures. The final steps of protein purification, crystallization and data collection were carried out in the presence of the metal chelator EDTA and the dinuclear site is empty or fractionally occupied. This Bfr structure is the first from a photosynthetic organism and also constitutes the first example of heme, a non-symmetric molecule, situated on a crystallographic axis.

2. Materials and methods

2.1. Purification and crystallization

Crystals of *R. capsulatus* bacterioferritin were obtained while attempting to crystallize the cytochrome cbb_3 cytochrome oxidase of this organism. A strain overproducing the cytochrome bc_1 complex (pMTS1/MT-RBC1) was grown on MPYE medium as described by Gray *et al.* (1994). Membranes were obtained by passing the cells in 50 m*M* potassium phosphate pH 7.5 through a French pressure cell twice at 82.7 MPa pressure. After removing unbroken cells by centrifuging at 11 000*g* for 30 min, the membranes were pelleted (2 h at 310 000*g*), washed once in the same buffer and then stored frozen until use. Membranes were thawed, resuspended to 10 g l^{-1} protein concentration (by the Lowry procedure; Lowry *et al.*, 1951) in 50 m*M* potassium phosphate pH 7.5 containing enough NaCl and dodecyl maltoside (DM) to give final concentrations of 260 m*M* and 10 g l^{-1}, respectively. After stirring for 30 min, the mixture was centrifuged for 30 min at 12 000*g* to remove undissolved material and the supernatant was passed through a 2.5 × 100 cm column of DEAE-Sepharose CL6B to remove the cytochrome bc_1 complex and other strongly anionic proteins. The unbound flowthrough was diluted twofold with 25 m*M* potassium phosphate pH 7.5, 0.1 g l^{-1} DM and reapplied to DEAE Sepharose 6B. This time, the column was washed with 25 m*M* potassium phosphate, 130 m*M* NaCl, 0.1 g l^{-1} DM and eluted with 50 m*M* potassium phosphate, 260 m*M* NaCl, 0.1 g l^{-1} DM. Colored peaks were identified by taking spectra of several fractions before and after adding dithionite and the peak containing oxidase (high-potential cytochrome *b* reduced before dithionite, cytochrome *c* reduced by dithionite) was pooled and applied to a hydroxyapatite column, washed with 50 m*M* potassium phosphate, 0.1 g l^{-1} DM and eluted with 200 m*M* potassium phosphate, 0.5 g l^{-1} DM.

A final gel-filtration chromatography was performed on Sephacryl S-300 in 20 m*M* K-MOPS 7.5, 100 m*M* NaCl, 0.1 g l^{-1} DM and 0.5 m*M* EDTA. The colored peak containing oxidase was pooled and concentrated by ultrafiltration through Amicon (Pharmacia) YM-100 membranes. For crystallization, 5 or 10 µl of the concentrated protein was supplemented with 20 g l^{-1} octyl glucoside, mixed with an equal volume of 0.1 *M* NaOAc pH 4.6, 4% PEG 4K (precipitant #37 of the 'Screen Lite' screening kit from Hampton Research) and allowed to equilibrate by vapor diffusion with the same precipitant. Because the final chromatography buffer in which the protein was obtained contained 0.5 m*M* EDTA and the precipitant did not contain metal ions other than Na^+ and K^+, it can be assumed that the concentration of divalent and trivalent metal ions is quite low.

The major precipitation was fine and amorphous, but after several weeks tiny (<0.1 mm) red crystals were observed that continued to grow slowly. In some cases we obtained small cube-shaped crystals; in other cases they were cross-shaped and dendritic. Because of their small size, it was not deemed practical to analyze the crystals by SDS–PAGE to ascertain their polypeptide content.

Table 1
Statistics of X-ray diffraction data collection for the native crystal.

Values in parentheses refer to data in the highest resolution shell.

Mosaicity	0.57
Resolution (Å)	21.5–2.6 (2.64–2.59)
No. of reflections	260696
Unique reflections	22679
Completeness (%)	99.4 (94)
R_{sym}† (%)	7.7 (33.1)

† $R_{sym} = \sum |I_i - \langle I \rangle| / \sum \langle I \rangle$, where i is the ith measurement and $\langle I \rangle$ is the weighted mean of I.

2.2. Data collection, phasing and model refinement

Crystals were transferred *via* mixtures to a cryoprotectant solution containing 25% glycerol, 12% PEG 4K, 10 m*M* KMES 6.7, 3 m*M* azide and 3 g l^{-1} OG for 10–30 min and then frozen in liquid nitrogen for data collection. X-ray fluorescence spectra taken at SSRL beamline 1-5 (H. D. Bellamy and E. A. Berry, unpublished results) showed the presence of iron and absence of copper. Data were collected at SSRL on beamline 7-1 using a wavelength of 1.08 Å. The best crystals diffracted to 2.6 Å.

The reciprocal-space lattice was highly symmetrical, with primitive cell parameters of $a = 122.89$, $b = 122.02$, $c = 122.38$ Å, $\alpha = 109.06$, $\beta = 110.19$, $\gamma = 109.25^\circ$. The diffraction could be accurately indexed on a trigonal *I*-centered cubic (or tetragonal or orthorhombic) or *F*-centered orthorhombic lattice. Merging symmetry-related reflections ruled out the possibility of rhombohedral or cubic symmetry and confirmed tetragonal

Table 2
Phasing statistics for the heavy-atom derivatives.

R_{cullis} is defined here as R_{cullis} = (⟨phase-integrated lack of closure⟩/⟨$|F_{PH} - F_P|$⟩) and phasing power (PP) as PP = (⟨$|F_{Pcalc}|$/phase-integrated lack of closure⟩) from *SHARP* in the shell 4.6–5.8 Å for acentric reflections. F_P and F_{PH} represent the structure factors for the protein and derivative, respectively.

			'Isomorphous'		'Anomalous'	
Compound	Sites	Resolution (Å)	R_{cullis}	PP	R_{cullis}	PP
Native	0	2.6				
Me_3Pb	6	2.8	0.87	1.64	0.80	1.55
$OAcHg_2$	4	2.6	0.77	1.74	0.92	1.06
PCMB	2	2.8	0.99	0.16	1.00	0.17
Mersalyl	0					

Table 3
Figure-of-merit statistics for MIR phases from *SHARP*.

FOM, figure of merit for the quality of the centroid structure factor; *N*, number of reflections.

	Acentric reflections		Centric reflections	
d_{min}–d_{max} (Å)	*N*	FOM	*N*	FOM
37.94–7.20	835	0.84047	318	0.55480
7.20–5.14	1597	0.78699	325	0.66542
5.14–4.21	2107	0.64435	324	0.46472
4.21–3.65	2519	0.50549	329	0.45110
3.65–3.27	2873	0.40877	321	0.41385
3.27–2.98	3186	0.37257	325	0.43030
2.98–2.76	3485	0.31457	330	0.33725
2.76–2.59	3565	0.21986	301	0.20582
Overall	20167	0.43790	2573	0.44213

Table 4
Refinement and model statistics.

Values indicated in parentheses correspond to the highest resolution shell.

Resolution range (Å)	21.5–2.6 (2.69–2.60)
No. of reflections used for R_{cryst} calculation	21429
No. of reflections used for R_{free} calculation	1097
Data cutoff $F/\sigma(F)$	0.0
R_{cryst}† (%)	22.5 (31.4)
R_{free} (%)	24.2 (39.8)
No. of non-H protein atoms	3,804
No. of water molecules	21
Mean *B* factor, protein main-chain atoms (Å^2)	25.2
Mean *B* factor, protein side-chain atoms (Å^2)	27.1
Mean *B* factor, solvent atoms (Å^2)	23.1
Ramachandran plot	
Residues in most favored regions (%)	96.9
Residues in additionally allowed regions (%)	3.1
R.m.s.d. from ideal geometry	
Bond length (Å)	0.008
Bond angle (°)	1.2

† $R_{cryst} = \sum ||F_{obs}| - |F_{calc}|| / \sum |F_{obs}|$. R_{free} is the same as R_{cryst} but calculated for 4.9% of the data omitted from the refinement.

symmetry and the space group *I*422 (Table 1), with unit-cell parameters a = 142.37, b = 142.37, c = 140.88 Å.

Initially, it was assumed that the crystals contained the cytochrome cbb_3 oxidase. Molecular replacement using subunit 1 of the *Paracoccus denitrificans* cytochrome aa_3 cytochrome oxidase (PDB code 1ar1) was unsuccessful, so further crystals were soaked in 5 m*M* solutions of several heavy-atom reagents in the same cryoprotectant solution. Full data sets were collected for four of these; these data sets and the best native data set were used for MIR phasing with the program *SOLVE* (Terwilliger & Berendzen, 1999). *SOLVE* found useable heavy-atom sites in three of the derivatives: those containing trimethyl lead, PCMB and dimercury acetate. The resulting phases showed that the asymmetric unit contained several hemes and at least ten α-helices of about the right length to be transmembraneous. The sites identified by *SOLVE* were further refined using *SHARP* (La Fortelle *et al.*, 1997). Statistics for the isomorphous phasing are presented in Tables 2 and 3. The sites in the PCMB derivative refined to very low occupancy and did not contribute significantly to phasing. The density-modification program *SOLOMON* (Abrahams & Leslie, 1996) was run using the script automatically generated by *SHARP* to further improve the phases. Maps generated from the phases from *SHARP* were not significantly better than those from *SOLVE*; however, the density modification was very effective and the final map was quite interpretable.

At this point, it becames clear that there was threefold noncrystallographic symmetry and that the asymmetric unit contained three copies of a unit with only four long helices and 1.5 hemes, with the hemes ligated by the same residue in two different monomers and one of the hemes on a twofold axis ligated by a crystallographically related monomer (see §3). Solvent flattening and threefold NCS symmetry averaging (*RAVE* package; Kleywegt & Jones, 1994) further improved the maps, allowing the chain to be completely traced and many side chains to be tentatively assigned. After several cycles of rebuilding, side-chain assignment or modification and refinement in *CNS_SOLVE* (Brunger *et al.*, 1998), the *R* and free *R* factors (Brünger, 1992) were 29 and 31%, respectively. The assigned sequence was then submitted to a *BLAST* search of the GenBank nonredundant database which showed the closest match to be bacterioferritin from *R. capsulatus* (36% identity). Comparison with the structure of the *E. coli* homolog [PDB entries 1bcf (Frolow *et al.*, 1994) and 1bfr (Dautant *et al.*, 1998)] confirmed that our protein was a bacterioferritin. The sequence was thus corrected (Penfold *et al.*, 1996) and the structure was refined by energy minimization between 21.5 and 2.6 Å resolution, $F/\sigma(F) > 0$ (22 526 independent reflections), using *CNS_SOLVE* and rebuilding using *O* (Jones *et al.*, 1991). Bulk-solvent correction was applied to all data during the refinement. Atoms of the heme group on a twofold crystallographic axis were given occupancy 0.5 and fixed during positional refinement (see §3). A noncrystallographic restraint of 1256 kJ mol^{-1} was also applied between protein atoms of the three molecules in the asymmetric unit. Individual *B*-factor refinement was performed in the last stage of the refinement and 21 water molecules were included using the automated procedure implemented in *CNS_SOLVE* followed by a manual examination using *O*. After inspection of the $F_o - F_c$ electron-density map, two conformations were imposed on the Glu18 side chain, each with an occupancy of 0.5. At the end of the refinement, the *R* and free *R* factors were 22.7 and 24.7%, respectively, for all data (σ_F cutoff = 0)

between 21.5 and 2.6 Å resolution (Table 4). In order to refine the heme on the special position, the data set was reduced again from unmerged intensities using *I*4 symmetry and an asymmetric unit of three dimers enclosing three hemes was refined in this lower symmetry space group. After this refinement converged, the symmetry was again returned to *I*422 and the trimer with 1.5 hemes was further refined. From these coordinates the trimeric structure with proper threefold symmetry (Fig. 1*b*) was then generated and refined using *CNS_SOLVE*, without fixing the atoms of the heme situated on the crystallographic axis, and applying threefold NCS restraints to the protein. Interactions between the overlapping heme molecules were ignored during positional refinement. This resulted in the final submitted structure, with *R* and free *R* factors of 22.5 and 24.2%, respectively. Questions regarding the symmetry of this heme will be discussed below.

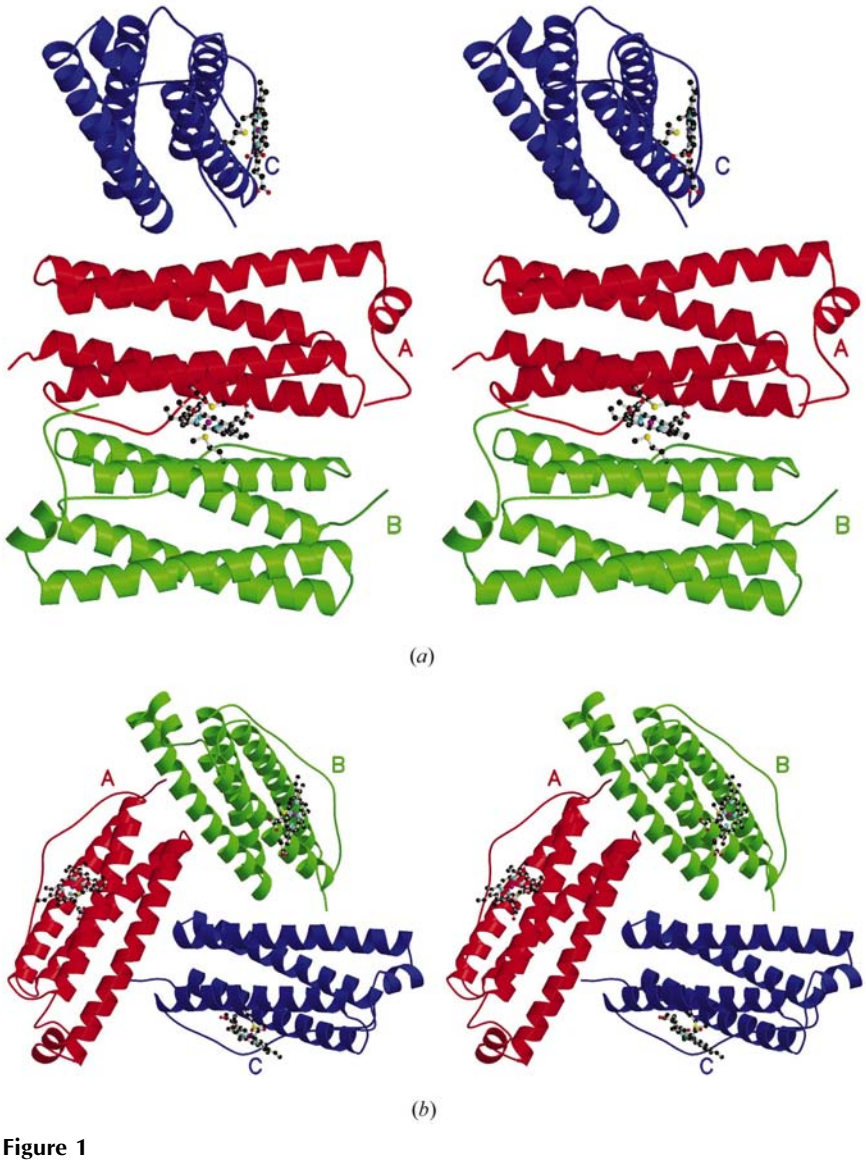

Figure 1
Stereoview of the *R. capsulatus* bacterioferritin subunits of the asymmetric unit schematized as a ribbon. The heme group is represented in ball-and-stick. Monomers related by a noncrystallographic twofold axis are colored red (*A*) and green (*B*). The subunit (*C*) binding to the heme situated on the twofold crystallographic axis is colored blue. (*a*) Trimeric structure including one heme with its enclosing non-crystallographic dimer and one heme on a crystallographic axis (at occupancy 0.5) with one monomer of its enclosing dimer. (*b*) Trimeric structure with proper threefold symmetry, including three heme moieties at 0.5 occupancy. Met52 and heme are represented in ball-and-stick. The figure was generated using *MOLSCRIPT* (Kraulis, 1991) and *Raster3D* (Merritt & Murphy, 1994).

3. Results and discussion

3.1. Crystal contents and structure quality

The unit cell contains two 24-subunit protomers related by the crystallographic centering operator. The asymmetric unit contains three Bfr subunits of 161 residues each (*A*, *B* and *C*), 1.5 hemes and 21 water molecules. The C-terminal residue (Glu161) was not observed in the density and was omitted from the model. Despite the presence of two detergents in the mother liquor and cryoprotectant, no detergent molecule was found in the structure. The r.m.s. deviation values between pairs of the molecules are 0.02 Å for all the atoms and 0.01 Å for the C^{α} atoms. Analysis of the model using *CNS_SOLVE* (Brunger *et al.*, 1998) and *PROCHECK* (Laskowski *et al.*, 1993) shows a good stereochemistry with an r.m.s.d. of 0.008 Å for the bond lengths and 18° for the dihedral angles. No residue is observed in the disallowed regions of the Ramachandran plot (Table 4).

3.2. Structure of the *R. capsulatus* Bfr

The *R. capsulatus* Bfr monomer is a bundle of four long antiparallel α-helices (named *A*, *B*, *C* and *D*) and a small α-helix (*E*) in the C-terminal part (Figs. 1*a* and 1*b*). These helices extend from Ala5 to Trp35 (helix *A*), from Ala38 to Phe64 (*B*), from Pro83 to Glu110 (*C*) and from Ile114 to Leu144 (*D*). Helix *E* is composed of residues Pro146–Leu152. The long loop connecting helix *B* to *C* (*L* loop from residues Leu65 to Glu82; Dautant *et al.*, 1998) runs from one end of the bundle to the other. In doing so it loops out over the heme cleft, separating it from the external medium and providing hydrophobic residues 71 and 74 to the lining of the cavity which contains the heme ring and a cluster of five waters described below.

Bacterioferritin from *R. capsulatus* has 49.7% sequence identity with the *E. coli* Bfr, whose structure is known (Frolow *et al.*, 1994). Except for the C-terminal part of the subunit, which is longer and in a different orientation in *R. capsulatus* Bfr, the overall structure is similar. A subunit of the

R. capsulatus bacterioferritin can be superimposed on the *E. coli* monomer (1bcf) with an r.m.s. deviation between C^{α} atoms of 0.78 Å. Dimers sharing a heme molecule (chains *A* and *B*) superimpose on the corresponding *E. coli* dimer with an r.m.s. deviation 0.96 Å.

The choice of the asymmetric unit is somewhat arbitrary, but for the *I*422 *R. capsulatus* crystals it must contain 1/16 of the unit-cell contents, *i.e.* three monomers and 1.5 hemes. For most of the refinement we have used the asymmetric unit shown in Fig. 1(*a*), in which the three monomers are related by rotations of 180° (*A* and *B*), 120° (*A* and *C*) and 90° (*B* and *C*). Molecules *A* and *B* share one heme group in the asymmetric unit, while molecule *C* binds heme with another molecule *C* related by a twofold crystallographic axis.

The protomer described as the biological form of bacterioferritin (Frolow *et al.*, 1994) contains 24 protein subunits and 12 heme molecules, *i.e.* eight asymmetric units of the *I*422 crystal described above. Four threefold symmetry axes and three fourfold axes are present in the protomer. Holes where these axes pass through the shell result in eight channels with threefold symmetry and six channels with fourfold symmetry (Figs. 2*a* and 2*b*). There are six twofold axes (Fig. 2*c*), but instead of passing through channels in the shell they pass through the 12 heme molecules, along the quasi-twofold axis of heme as described by Frolow *et al.* (1994). Correlation between this symmetry of the protomer and crystallographic/noncrystallographic symmetry in the different crystals will be discussed below.

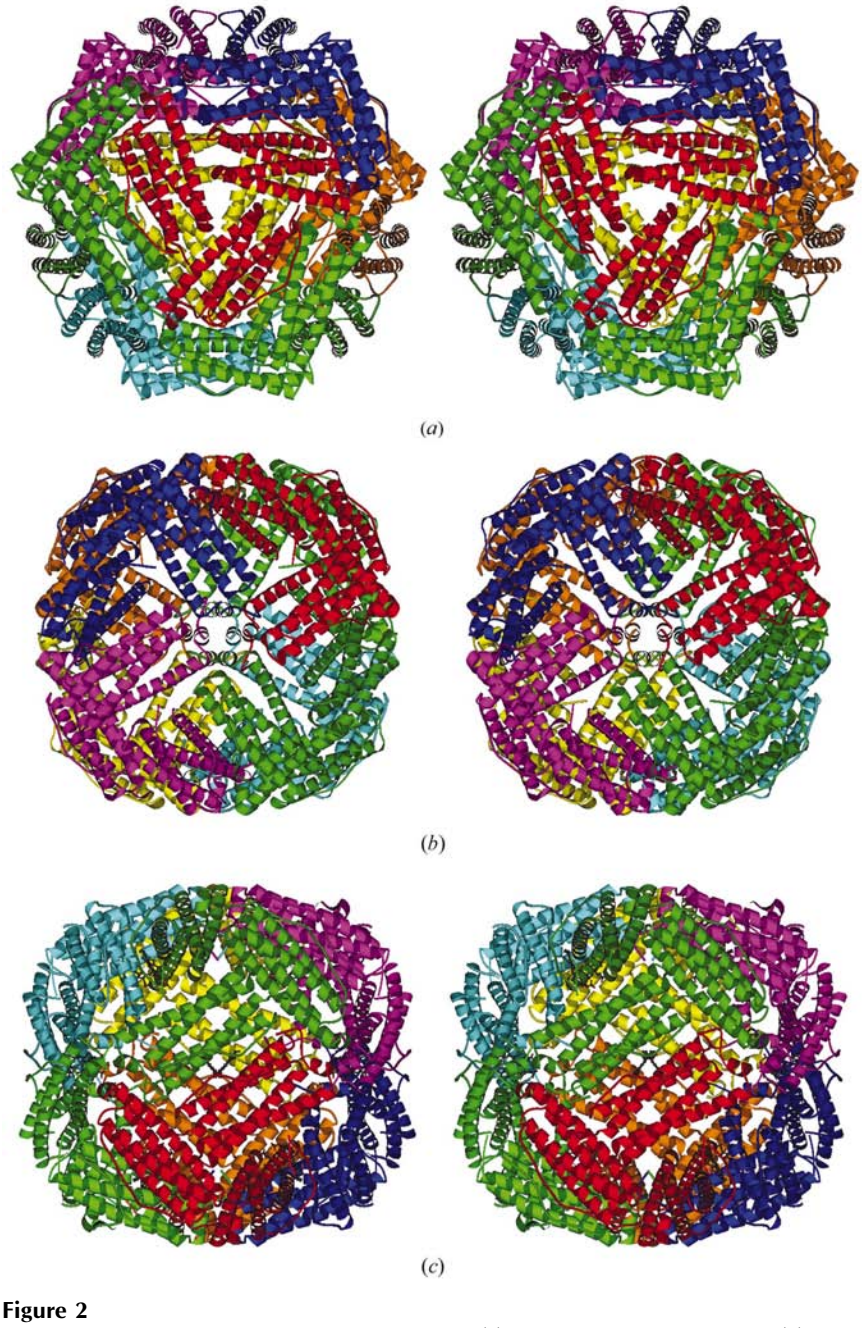

Figure 2
Stereoview of the Bfr protomer from *R. capsulatus* (*a*) along the threefold channel, (*b*) along the fourfold channel situated on the fourfold crystallographic axis and (*c*) along the twofold axis. The figure was generated using *MOLSCRIPT* (Kraulis, 1991).

The fourfold channels are hydrophobic in *R. capsulatus* Bfr, being lined with four leucine residues (Leu151) (Fig. 3). This channel is also hydrophobic in the mammalian ferritins (Harrison & Arosio, 1996); however, in the *E. coli* enzyme it was reported as hydrophilic (Harrison & Arosio, 1996), being defined by the Asn148 and Gln151 side chains (Dautant *et al.*, 1998). In *R. capsulatus* Bfr, these are replaced by glycine and leucine, respectively, resulting in a non-polar channel. No positive electron-density peak corresponding to an ion or water molecule, as described in the *E. coli* Bfr (Dautant *et al.*, 1998), was observed at the entrance of this channel in the *R. capsulatus* enzyme.

The threefold channels, through which Fe^{II} ions are probably scavenged in ferritins and bacterioferritin (Chasteen & Harrison, 1999), are hydrophilic as in mammalian ferritins and *E. coli* bacterioferritin (Harrison & Arosio, 1996). The charged side chains Arg105, Lys117, Glu121 and Asn118, which define this channel in the *R. capsulatus* protein (Fig. 4), are not highly conserved, but at least residues 117 and 118 are nearly always charged or polar, giving this channel a hydrophylic lining. The threefold channel can bind different ions: Ca^{2+}, Cd^{2+}, Zn^{2+} and Tb^{3+} (Harrison & Arosio, 1996).

The two structures of *E. coli* bacterioferritin are from crystals grown in the presence of heavy-metal ions and both have metal bound at two adjacent sites comprising the 'dinuclear site' at which Fe^{2+} is believed to be oxidized to Fe^{3+}. In the deposited

structures these are modeled as two manganese ions, probably mimicking the iron bound to the enzyme, and they are stabilized by glutamate and histidine residues (Glu18, Glu51, Glu94, Glu127, His54 and His130). The present crystals were grown and used without heavy metals and in the presence of the chelator EDTA; the dinuclear site is either vacant or occupied to a small fractional extent. Analysis of the $F_o - F_c$ electron-density map reveals two positions for the Glu18 side chain (Fig. 5*a*). These two positions are also observed in the $3F_o - 2F_c$ omit map calculated omitting this residue and the atoms around it in a sphere radius of 3.5 Å (Fig. 5*b*). Analysis of the $F_o - F_c$ electron-density map led us to attribute an occupancy of 0.5 to each position of the Glu18 side chain. Also, the *B* factors and occupancy refinements starting with values around 26 Å^2 and 1.0 for either positions of the side chain lead to values around 50 Å^2 and 0.5, respectively, for the Glu18 carboxylate group, reflecting the two positions or a partial disorder. The single Glu18 side-chain orientation in the *E. coli* structure may be constrained by the presence of metal ions in the dinuclear center.

A positive electron-density peak in the $F_o - F_c$ map higher than 3.5σ in the dinuclear center suggests that the dinuclear center is partially occupied by a metal ion in *R. capsulatus* enzyme crystals, even though no salt containing manganese ions was used to crystallize the protein. Distances from this peak to the O atoms of each carboxylate group range between 4.06 and 1.95 Å. The closer contacts are too short for hydrogen bonds and thus not consistent with the peak being a water molecule. This peak was therefore interpreted as an iron ion with a low occupancy. Assignment of occupancy 0.33 and *B* factor 40.0 Å^2 in the three molecules of the asymmetric unit resulted in disappearance of the positive $F_o - F_c$ peak. Refinement of the *B* factors and the occupancies for these three Fe ions led to *B* factors between 50.5 and 57.5 Å^2 and occupancies between 0.28 and 0.44. This putative iron ion is in a position similar to that of Mn202 in the *E. coli* structure 1bfr and Mn601 in 1bcf, being ligated by Glu18, Glu51, His54 and Glu127. The presence of this iron ion at fractional occupancy could be responsible for the two positions observed for the Glu18 side chain.

A tetrahedron of four water molecules in the vicinity of the heme was found as described for the *E. coli* enzyme (Frolow *et al.*, 1994) and a fifth water molecule was observed situated at the center of the tetrahedron. The average *B* factor for these five water molecules is 17.5 Å^2. They are hydrogen bonded to the Ser23 hydroxyl group and to the carbonyl group of Leu19 and Leu71 in two monomers, as well as to each other. No functional role has so far been attributed to these water molecules.

There are 12 hemes in the protomer, located between pairs of monomers. The Fe atom in each heme is axially coordinated by the S atoms of the Met52 side chains from two molecules related by the noncrystallographic or crystallographic symmetry. Axial ligation by methionine is also seen in class I type *c* cytochromes; however, the chirality of the methionine S atom here is the opposite of that in *c* cytochromes. This can be seen by comparing the structures visually or from the 'improper' angle calculated for S^{δ}, C^{γ}, C^{ε}, Fe which is -39° in cytochrome *c* (PDB code 1ycc) but $+19^{\circ}$ in the *R. capsulatus* Bfr. This value is also positive in the *E. coli* structures, being $+22^{\circ}$ for 1bfr and $+52^{\circ}$ for 1bcf.

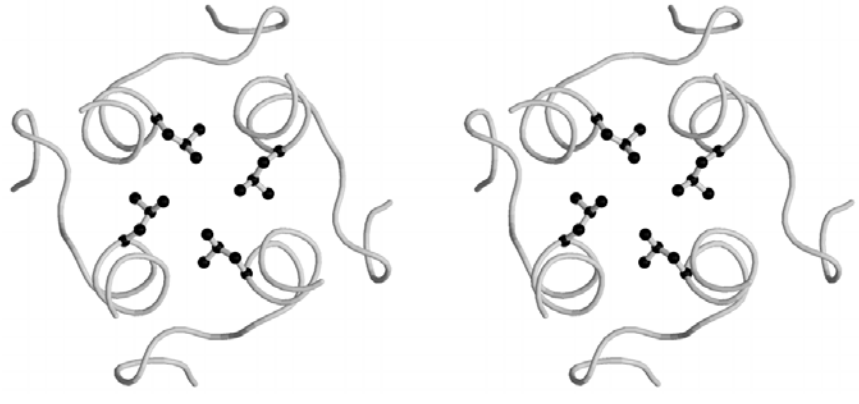

Figure 3
Stereoview of the fourfold channel along the fourfold crystallographic axis. The Leu151 residues situated in the channel are displayed in ball-and-stick. The figure was generated using *MOLSCRIPT* (Kraulis, 1991) and *Raster3D* (Merritt & Murphy, 1994).

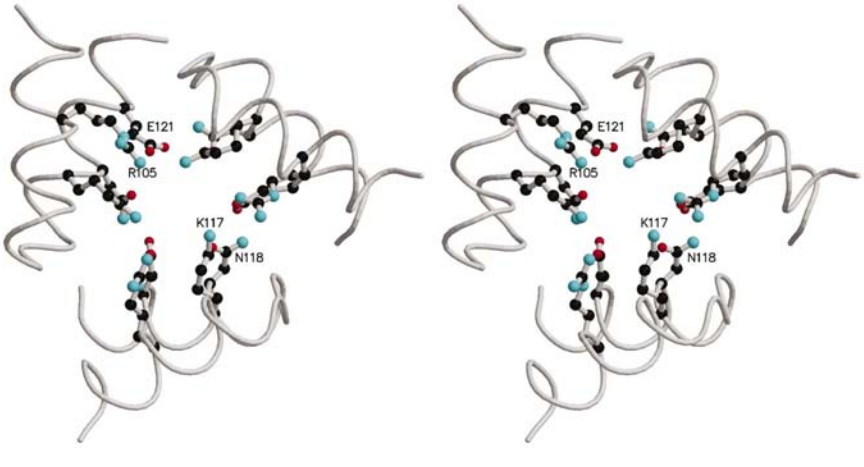

Figure 4
Stereoview of the threefold channel. Arg105, Lys117, Asn118 and Glu119 are represented in ball-and-stick. The figure was generated using *MOLSCRIPT* (Kraulis, 1991) and rendered using *Raster3D* (Merritt & Murphy, 1994).

The tetrapyrrole ring of the heme molecule is situated in a hydrophobic pocket constituted of Leu19, Ile22, Trp26, Ile49, Met52 and Leu71 of both monomers. The ring is perpendicular to the surface, with the edge bearing the vinyl groups (*BC* edge) directed outwards and the edge bearing the propionate groups (*AD* edge) directed inwards towards the internal cavity of the protomer. The vinyl/methyl groups contact the cluster of five water molecules, which is in turn covered by protein of the *L* linker loops, separating the water cluster and the heme from the external medium. The propionates actually extend into the solvent-filled volume inside the shell. This contributes to the highly polar and charged nature of the inner surface, which also has many side chains of Asp, Glu and Arg residues protruding into the cavity. The side chain of Arg45 in the two adjacent monomers approaches the two heme propionates. As described below, the use of two symmetry-related hemes at half occupancy in each site allowed the refinement program to fit two conformations for each propionate (Figs. 6*a* and 6*b*). In one conformation the propionate makes an ion pair with Arg45. The propionates are rather less ordered than the tetrapyrrole ring or the Arg45 side chains and probably have two or more conformations, one making the ion pair directly with Arg45 and others hydrogen bonding only with bulk water.

3.3. Interaction of crystallographic and protomeric symmetry and heme quasi-symmetry

The symmetry of the bacterioferritin protomer can readily be visualized as the symmetry of a cube (point group 432), with three fourfold axes (connecting midpoints of three pairs of opposite faces of the cube), four threefold axes (connecting four pairs of opposing corners) and six twofolds (connecting midpoints of opposite edges). This symmetry results in 24 equivalent positions, implying that the 'asymmetric unit' of the structure consists of one monomer and half a heme moiety. The 12 hemes are located on the six twofold axes, so that each is shared between two identical monomers. The heme molecule is not twofold symmetric, but has a quasi-twofold axis as described in Frolow *et al.* (1994). The asymmetry involves the methyl and vinyl substituents on rings *B* and *C* of the tetrapyrrole: rotation about the quasi-twofold brings the methyl of ring *C* onto the vinyl of ring *B* and *vice versa*. In bacterioferritin the quasi-twofold axis is aligned with the twofold axes of the protomer. Thus, the slight asymmetry of the heme breaks the symmetry of the protein structure, which would otherwise have proper 432 symmetry. To put it another way, the asymmetric heme is binding in a perfectly symmetrical binding site. Looking at a single binding site, it must be assumed that the heme binds equally well in both orientations. It seems unlikely that the slight asymmetry of the heme in one binding site would introduce significant asymmetry in the next binding site, so we believe the orientations of heme in the 12 binding sites are all independent and random. If so, then on average, as seen by crystallography, the heme consists of two hemes in opposite orientations, each with occupancy 0.5. As discussed below, all the crystallographic evidence we could obtain was consistent with such a model.

E. coli bacterioferritin has been crystallized in space groups $P4_22_12$ and $P2_1$. The $P4_22_12$ form (Frolow *et al.*, 1994) has four protomers in the unit cell and a half protomer (12 monomers and six hemes) in the asymmetric unit. This space group has twofold axes parallel to *c* and diagonally in the *ab* plane. The twofolds parallel to *c* and all the crystallographic screw axes pass between protomers. The twofolds in the *ab* plane pass through fourfold axes of the protomer. Thus, the protomeric twofolds on which the hemes lie are all noncrystallographic twofold axes. In the $P2_1$ crystal form (Dautant *et al.*, 1998) all the symmetry of the protomer is noncrystallographic (the space group has only a twofold screw axis), so again the heme is not in a special position in terms of crystallographic symmetry.

In the *I*422 crystals from *R. capsulatus* Bfr described here, there is a much closer matching of crystallographic and protomeric symmetry. The origin of the unit cell is at the center of a protomer (within an arbitrary choice of origin), with the fourfold crystallographic axis along *c* passing through a protomeric fourfold (defining the axial direction of the protomer in this lattice). The space group has twofold axes along *a* and *b* which pass through two equatorial twofolds of the protomer, putting the four hemes on these twofolds in special crystallographic positions. The space group also has twofolds in the *ab* plane along 110 and $1\bar{1}0$; these pass through the two equatorial fourfold axes of the protomer. The only protomeric symmetry axes that do not fall on a crystallographic symmetry axis are the threefold axes. The translation (centering) operators then generate the second protomer at the center of the cell.

The asymmetric unit consists of three protein chains and 3/2 heme molecules. For the initial refinement, we chose a heme on a noncrystallographic axis with its enclosing dimer (*AB*) plus an adjacent monomer *C* with its heme on a crystallographic axis, with that heme occupancy set to 0.5 (Fig. 1*a*). The heme on the noncrystallographic axis was refined positionally, while the atoms of the heme on the crystallographic axis had to be fixed during positional refinement owing to their special position and overlap with symmetry-related heme atoms. In order to positionally refine the heme on the crystallographic axis, we refined a hexameric asymmetric unit (consisting of the trimeric asymmetric unit of Fig. 1(*a*) together with the trimer related by the twofold axis passing through the heme on monomer *C*) against the data set reduced in space group *I*4, which lacks this twofold axis. By using data reduced in space group *I*4 from unmerged intensities, we were also able to test whether the heme on the crystallographic axis conforms to the higher symmetry of *I*422, which would imply equal binding of heme in two orientations. Ordinarily, the space group *I*422 would be chosen over *I*4 based on the similar $R_{\rm sym}$ values (7.7 and 7.4%) obtained in the two space groups. The asymmetry involves only two methyl groups, which might not contribute enough to the scattering to give a high $R_{\rm sym}$ in *I*422; however, asymmetric binding of the heme could only be explained by an asymmetry in the protein shell which presumably would result in a high $R_{\rm sym}$ in *I*422. Nonetheless, to test this question, refinement was carried out with this heme in either of its two

orientations and we looked at positively and negatively contoured $F_o - F_c$ maps. If the heme is actually predominantly in one orientation, then when the wrong orientation is used there should be strong positive peaks on the two vinyl side chains (because they are actually methyls) and negative peaks at the two methyl groups which are actually vinyl. When the correct orientation is used, positive and negative peaks should be much less significant. This was not observed; the fit was actually about the same with either orientation. The analysis of the $3F_o - 2F_c$ omit map calculated in the *I*4 space group (Fig. 6*a*) did not clearly favor one orientation, although it was not perfectly symmetrical. So, we conclude that two orientations of the heme situated on the crystallographic axis are present (as modeled in Fig. 6*a*) and the high-symmetry space group (*I*422) is justified.

Frolow *et al.* (1994) considered the possibility that the heme preferentially binds in one orientation about its quasi-twofold axis in the apparently symmetrical binding site. Since all of their hemes were on local and not crystallographic twofolds this could not be excluded and in fact Frolow and coworkers did report some indication of asymmetry in their 2.9 Å structure.

In our case, for the heme situated on the noncrystallographic axis between monomers *A* and *B*, a similar approach to that described previously was used. This time, our original trimeric asymmetric unit (Fig. 1*a*) was refined against data reduced in space group *I*422. Each heme orientation was refined independently with an occupancy of 1.0, leading to similar values for the free *R* and *R* factors. $F_o - F_c$ maps did not show a strong preference for one orientation. Again, the $3F_o - 2F_c$ omit map did show some asymmetry (Fig. 6*b*), but mainly in the position of the propionate side chains which are not well ordered (see above). No significant asymmetry is found in the vinyl and methyl groups on the other side of the ring which are responsible for the intrinsic asymmetry of heme. Thus, we conclude that a model in which all binding sites are occupied by a symmetric 0.5:0.5 mixture of heme in two orientations is more realistic than one in which heme binds with a particular orientation, the space group *I*422 is justified for the crystal and the protomeric point-group symmetry of 432 is appropriate for the biological entity including the heme when the ensemble of a large number of molecules is considered. If there is some non-random pattern to the heme orientation in the protomer, it does not affect crystal contacts and so is lost owing to random packing of the protomer into the lattice.

This situation cannot be represented by our original asymmetric unit depicted in Fig. 1(*a*) unless the heme molecule between monomers *A* and *B* is allowed two orientations at half occupancy, in which case the asymmetric unit of Fig. 1(*b*) would be more intuitive, with three monomers and three hemes at half occupancy, all related by proper threefold noncrystallographic symmetry. This is the asymmetric unit used in our final refinement and submitted to the PDB.

Just as this model accommodates the asymmetry of the methyl and vinyl side chains on the *B* and *C* pyrrole rings in a symmetrical structure, it allows the propionates on the *A* and *D* rings (which are symmetrical in covalent structure) to take on different conformations without violating the crystallographic and noncrystallographic symmetry. The refinement process took advantage of this freedom to refine two different conformations for the propionates on the *A* and *D* ring of both the crystallographic and noncrystallographic hemes (Fig. 6). This does not violate the symmetry of the crystallographically averaged structure, in which the crystallographic or crystallographic plus noncrystallographic symmetry superimposes the *D* ring of one heme with the *A* ring of another. Also, in any heme-

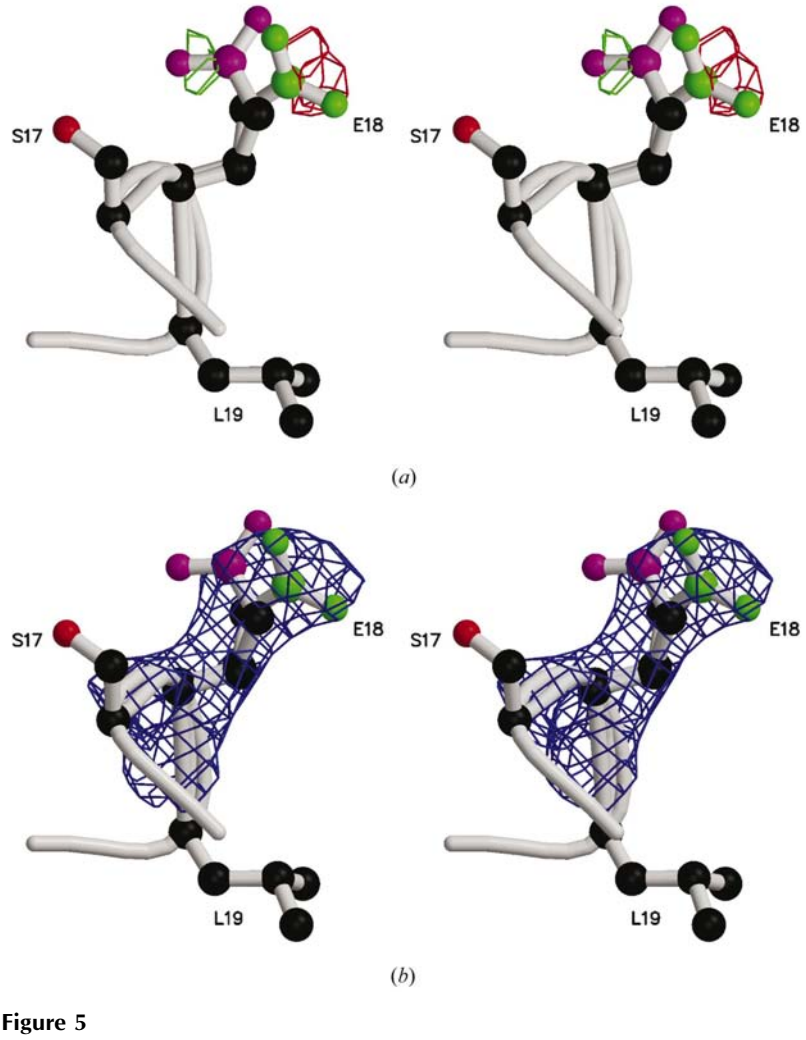

Figure 5
Stereoview of the two positions of the Glu18 side chain (*a*) in an $F_o - F_c$ electron-density map contoured at $+3.5\sigma$ (green) and -3.5σ (red) and (*b*) in a $3F_o - 2F_c$ omit map contoured at 1.2σ and calculated at 2.6 Å resolution. The Ser17, Glu18 and Leu19 side chains are displayed in ball-and-stick. The two positions for the carboxylate group are displayed in different colors. The figure was produced using *MOLSCRIPT* (Kraulis, 1991) and *O* (Jones *et al.*, 1991) and rendered using *Raster3D* (Merritt & Murphy, 1994).

binding site of a single molecule with one heme bound in a particular orientation, it is not expected that the two propionates will be different. Rather, both propionates will take on each conformation for a fraction of the time, resulting in an essentially symmetric structure. This is quite different from the case of the methyl and vinyl substituents on the *B* and *C* rings, which are asymmetric in covalent structure and will of course be asymmetric, showing either one or the other of two orientations, in any single heme-binding site. Only in the averaged structure seen by crystallography is the heme symmetrical with respect to these substituents.

It may be noticed in Fig. 6(*a*) and especially Fig. 6(*b*) that the electron density for the propionates does appear asymmetric, contrary to the above description. We cannot think of any cause for this asymmetry and attribute it to noise in the map, which is contoured at a relatively low level in order to show density for the relatively disordered propionates. In any case, this has no bearing on the symmetric orientation of the heme which is our main point here, as the propionates of heme are symmetric as far as the covalent structure is concerned and so give no information about the orientation.

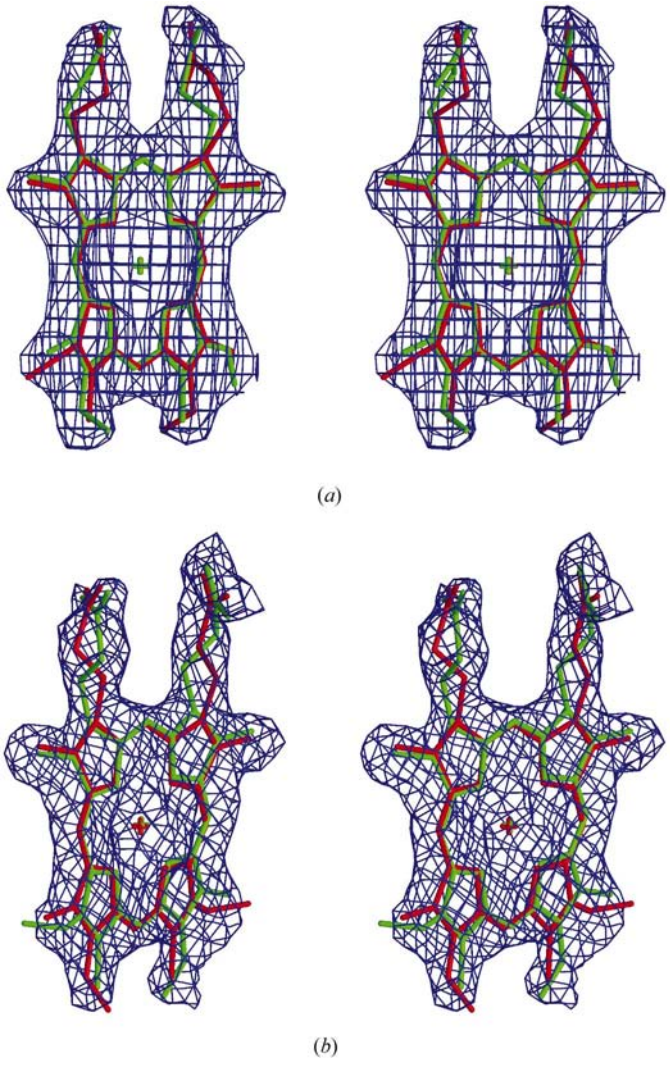

Figure 6
Stereoview of the two positions of the heme situated (*a*) on the twofold crystallographic axis and (*b*) on the twofold noncrystallographic axis in a $3F_o - 2F_c$ omit map calculated at 2.6 Å resolution and contoured at 1.2σ. The two orientations of the heme are shown in green and red. In (*a*) the map was calculated from data reduced in the space group *I*4 to avoid imposing symmetry and in both figures the heme was omitted from the phasing model. The figure was generated using *MOLSCRIPT* (Kraulis, 1991) and *O* (Jones *et al.*, 1991) and rendered using *Raster3D* (Merritt & Murphy, 1994).

4. Conclusions

The structure of *R. capsulatus* bacterioferritin presented here is the first from a photosynthetic organism or from any organism other than *E. coli*. The resolution (2.6 Å) is significantly higher than that of previous bacterioferritin structures. We believe this is the only structure in the database in which the quasi-symmetric protoheme molecule is located on a twofold crystallographic axis. Combining different crystallographic approaches, we conclude that this does not break the symmetry of the space group because the heme is bound in a symmetric mixture of two orientations. Other hemes in the structure, situated on a twofold non-crystallographic axis, probably also exist as a mixture of two orientations and we can also suppose a similar characteristic for the *E. coli* Bfr not shown in previous work.

The dinuclear metal-binding site is empty or fractionally occupied, probably owing to the presence of EDTA in the purification and crystallization solutions. One of the residues making up this site, Glu18, displayed two side-chain orientations in the ferroxidase center, probably resulting from the partial occupancy of the heavy atom which it ligates. The lining of the fourfold channel in the bacterioferritin shell is nonpolar in the *R. capsulatus* enzyme, unlike that of the *E. coli* enzyme.

The fact that bacterioferritin was obtained while purifying the cytochrome cbb_3 oxidase is suggestive of a functional association between these proteins; however, a coincidental copurification cannot be ruled out. This question is currently being investigated.

We would like to acknowledge the contribution of Henry Bellamy and the staff at the Stanford Synchrotron Radiation Laboratory (SSRL) for performing the XAFS scan and assistance with data collection. This work was supported by NIH R01 grants DK44842 and GM62563, and by the US Department of Energy contracts DOE-FG0291ER20052 to FD and DE-AC03-76SF00098 to LBNL. Crystallographic data collection was carried out at SSRL, which is operated by the Department of Energy, Office of Basic Energy Sciences. The SSRL Biotechnology Program is supported by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program and by the Department of Energy, Office of Biological and Environmental Research.

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.

Andrews, S. C., Le Brun, N., Barynin, V., Thomson, A. J., Moore, G. R., Guest, J. R. & Harrison, P. M. (1995). *J. Biol. Chem.* **270**, 23268–23274.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–474.

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Chasteen, N. D. & Harrison, P. M. (1999). *J. Struct. Biol.* **126**, 182–194.

Dautant, A., Meyer, J. B., Yariv, J., Precigoux, G., Sweet, R. M., Kalb, A. J. & Frolow, F. (1998). *Acta Cryst.* D**54**, 16–24.

Frolow, F., Kalb (Gilboa), A. J. & Yariv, J. (1994). *Nature Struct. Biol.* **1**, 453–460.

Gray, K. A., Dutton, P. L. & Daldal, F. (1994). *Biochemistry*, **33**, 723–733.

Harrison, P. M. & Arosio, P. (1996). *Biochim. Biophys. Acta*, **1275**, 161–203.

Jones, T. A., Zhou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kleywegt, G. J. & Jones, T. A. (1994). *Proceedings of the CCP4 Study Weekend. From First Map to Final Model*, edited by S. Bailey, R. Hubbard & D. A. Waller, pp. 59–66. Warrington: Daresbury Laboratory.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

La Fortelle, E. de, Irwin, J. J. & Bricogne, G. (1997). In *Crystallographic Computing 7*, edited by P. E. Bourne & K. D. Watenpaugh. Oxford University Press.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951). *J. Biol. Chem.* **193**, 265–275.

Merritt, E. A. & Murphy, M. E. P. (1994). *Acta Cryst.* D**50**, 869–873.

Penfold, C. N., Ringeling, P. L., Davy, S. L., Moore, G. R., McEwan, A. G. & Spiro, S. (1996). *FEMS Microbiol. Lett.* **139**, 143–148.

Romao, C. V, Louro, R., Timkovich, R., Lubben, M., Liu, M. Y., LeGall, J., Xavier, A. V. & Teixeira, M. (2000). *FEBS Lett.* **480**, 213–216.

Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.

Yang, X., Le Brun, N. E., Thomson, A. J., Moore, G. R. & Chasteen, N. D. (2000). *Biochemistry*, **39**, 4915–4923.