Haem and non-haem iron sites in *Escherichia coli* bacterioferritin: spectroscopic and model building studies

Myles R. CHEESMAN,* Nick E. LE BRUN,* Fahmi H. A. KADIR,* Andrew J. THOMSON,*§ Geoffrey R. MOORE,* Simon C. ANDREWS,† John R. GUEST,† Pauline M. HARRISON,† John M. A. SMITH†‡ and Stephen J. YEWDALL†

*Centre for Metalloprotein Spectroscopy and Biology, School of Chemical Sciences, University of East Anglia, Norwich NR4 7TJ, U.K., and †The Krebs Institute, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2UH, U.K.

The bacterioferritin (BFR) of *Escherichia coli* is an iron-storage protein containing 24 identical subunits and between three and 11 protohaem IX groups per molecule. Titration with additional haem gave a maximum loading of 12–14 haems per molecule. The e.p.r. spectra and magnetic c.d. spectra of the protein-bound haem show it to be low-spin Fe(III), and coordinated by two methionine residues as previously reported for BFRs isolated from *Pseudomonas aeruginosa* and *Azotobacter vinelandii* [Cheesman, Thomson, Greenwood, Moore and Kadir, Nature (London) (1990) **346**, 771–773]. A recent sequence alignment indicated that BFR may be structurally related to ferritin. The molecular model proposed for *E. coli* BFR has a four-α-helix-bundle subunit conformation and a quaternary structure similar to those of mammalian ferritins. In this model there are two types of hydrophobic pocket within which two methionine residues are correctly disposed to bind haem. The e.p.r. spectra also reveal a monomeric non-haem Fe(III) species with spin, $S = 5/2$. On the basis of sequence comparisons, a ferroxidase centre has recently been proposed to be present in BFR [Andrews, Smith, Yewdall, Guest and Harrison (1991) FEBS Lett. **293**, 164–168] and the possibility that this Fe(III) ion may reside at or near the ferroxidase centre is discussed.

INTRODUCTION

Bacterioferritin (BFR), the iron-storage protein of obligate and facultative aerobic prokaryotes, has been isolated from a wide range of bacteria including *Escherichia coli* (Yariv et al., 1981), *Azotobacter vinelandii* (Stiefel and Watt, 1979; Jiudi et al., 1980), *Azotobacter chroococcum* (Chen and Crichton, 1982) and *Pseudomonas aeruginosa* (Moore et al., 1986). Electron microscopy (Yariv et al., 1981), X-ray diffraction (Smith et al., 1989), amino-acid-sequence alignment (Andrews et al., 1991a) and secondary-structure prediction (Andrews et al., 1989a) for *E. coli* BFR (*Ec*-BFR) indicate that the three-dimensional structure of BFR resembles that of ferritin, the structurally defined iron-storage molecule of eukaryotes (Ford et al., 1984; Lawson et al., 1991). Thus, *Ec*-BFR consists of an iron core of up to 8.0 nm diameter surrounded by a spherical protein shell formed by the symmetrical association of 24 equivalent protein subunits. The bacterial protein differs from animal ferritin, as isolated, in containing protohaem IX. Each haem group is coordinated by two axial ligands generating low-spin Fe(II) and Fe(III) configurations. The number of haem groups found in native BFRs varies from three to 12 per 24 subunits. The highest ratio of haem to protein found in freshly isolated *Ec*-BFR is one haem per two protein subunits (Yariv et al., 1981; Smith et al., 1989). Although native *P. aeruginosa* BFR (*Pa*-BFR) contains five to nine haems per 24 subunits (Kadir and Moore, 1990) the protein will bind up to 24 haems per 24-mer *in vitro*, irrespective of whether the iron core is present or not (Kadir and Moore, 1990). The axial ligands to the ferrihaem group of *Pa*-BFR have been established as the thioether side-chains of two methionine residues by near-infrared (n.i.r.) magnetic circular dichroism (m.c.d.) spectroscopy in combination with e.p.r. spectroscopy (Cheesman et al., 1990). The same axial ligation was also established for haem groups in BFRs of *A. vinelandii* (*Av*-BFR) and *Rhodobacter sphaeroides* (*Rs*-BFR) by comparing the e.p.r. *g*-values and the n.i.r.-m.c.d. spectra with those of *Pa*-BFR (Cheesman *et al.*, 1990) and appropriate model compounds (McKnight et al., 1991).

The formation of the iron core of ferritin involves oxidation of Fe (II) by dioxygen and hydrolytic polymerization of the resulting Fe(III). At pH values of 6 or below, iron oxidation and hydrolysis proceed at a significant rate only in the presence of ferritin. This has given rise to the proposal that there is an Fe(II) binding and oxidation site, the ferroxidase centre, within the protein shell. Mammalian ferritins are co-polymers of two types of chain known as H and L. Work on recombinant ferritins has shown that ferroxidase activity is associated with H-chains only and that several residues, namely Glu-27, Tyr-34, Glu-61, Glu-62, Glu-107, His-65, are essential for full ferroxidase activity (Lawson et al., 1989; Bauminger et al., 1991; Treffry et al., 1992; E. R. Bauminger, P. M. Harrison, D. Hechel, N. W. Hodson, I. Nowik and A. Treffry, unpublished work). X-ray structural analysis of human H-chain ferritin reveals two metal-binding sites about 0.31 nm (3.1 Å) apart at the ferroxidase centre (Lawson et al., 1991). It has recently been proposed that binding of a pair of Fe(II) ions at this centre enables two-electron transfer to a dioxygen molecule bound to one of the Fe(II) ions and that this accounts for the fast oxidation (Treffry et al., 1992). The six residues mentioned, plus Gln-141, are conserved in all ferritin H-chains, this latter residue binding dioxygen in the proposed mechanism of catalysis (Treffry et al., 1992). In the recent sequence alignment of Andrews et al. (1991a) five of these residues are conserved in *Ec*-BFR: Glu-18, Tyr-25, Glu-51, His-54 and Glu-94 which correspond to Glu-27, Tyr-34, Glu-62, His-

Abbreviations used: BFR, bacterioferritin; *Ec*-BFR, *Pa*-BFR, *Av*-BFR and *Rs*-BFR, BFRs from *Escherichia coli, Pseudomonas aeruginosa, Azotobacter vinelandii* and *Rhodobacter sphaeroides* respectively; n.i.r., near infrared; m.c.d., magnetic circular dichroism; CT, charge-transfer.

‡ Deceased 14 June 1992.

§ To whom correspondence should be addressed.

65 and Glu-107 respectively in ferritin. The remaining residues Asp-50 and Glu-127 (of *Ec*-BFR) replace Glu-61 and Gln-141 respectively.

The BFR gene (*bfr*) of *E. coli* has been cloned, sequenced and overexpressed in *E. coli* (Andrews et al., 1989a,b). Details are now reported of the low-temperature e.p.r. and n.i.r.-m.c.d. spectra of this protein, both as isolated and after addition of extra haem. The location of haem-binding sites is discussed on the basis of these results together with a molecular model derived from its proposed structural similarity with human H-chain ferritin (Andrews et al., 1989a, 1991a; Smith et al., 1989). E.p.r. studies of the overexpressed protein also show evidence for a monomeric non-haem iron centre, the location of which is discussed.

EXPERIMENTAL

Isolation and characterization of *Ec*-BFR

The BFR was purified from the over-producing *E. coli* strain JRG2033 as previously reported (Andrews et al., 1991b).

The absorption coefficient, ϵ, for the oxidized Soret band was determined using the pyridine haemochromogen method (Falk, 1964) as 1.09×10^5 $M^{-1} \cdot cm^{-1}$ at 418 nm. A similar value was obtained by integration of the haem $g_z = 2.88$ in the e.p.r. spectrum of two different *Ec*-BFR preparations. This ϵ value compares with that of 1.37×10^5 $M^{-1} \cdot cm^{-1}$ reported by Yariv et al. (1981).

Intrinsic haem loading was calculated on the basis of protein concentration of the BFR solution and the absorption coefficient determined for the haem Soret band. The protein concentration was determined by the method of Lowry and, in one case, by amino-acid analysis. The maximum haem loading was investigated through a series of haem titration experiments, analogous to the procedure used by Kadir and Moore (1990) in their study of haem binding to *Pa*-BFR.

Spectroscopic measurements

Electronic absorption spectra were recorded on a Hitachi U-2000 or Hitachi 557 spectrophotometer.

M.c.d. spectra were recorded on a JASCO J-500 D, for the wavelength range 300–1100 nm, and on a home-built circular dichrograph in the range 800–3000 nm. The sample was loaded into a split-coil superconducting solenoid (Oxford Instruments; SM-4) capable of generating a maximum magnetic field of 5 T and giving temperature control at the sample between 1.5 K and 300 K. To ensure that good optical-quality glasses were formed on freezing, glycerol was added to samples at a level of 50% (v/v). These additions left the form of the absorption spectrum unchanged. E.p.r. spectra were recorded on an X-band spectrometer (Bruker ER200D with an ESP 1600 computer system) fitted with a liquid-helium flow cryostat (Oxford Instruments; ESR9). Concentrations of the low-spin ferric haem of BFR were obtained by double integration of the e.p.r. spectrum and comparison with the integrated e.p.r. spectrum of a known concentration of an aqueous copper(II)–EDTA complex using the method of Aåsa and Vänngård (1975).

Molecular modelling

Modelling was based on the multiple amino-acid-sequence alignment and predicted helical elements of BFR, which established the similarity between BFR and ferritin (Andrews et al., 1989a, 1991a). A three-dimensional model was constructed by fitting the aligned sequence to the coordinates of human H-chain ferritin (Lawson, 1990; Lawson et al., 1991) with appropriate modifications due to sequence deletions and/or substitution of 147 of its 182 residues.

The molecular modelling program FRODO (Jones, 1978), implemented on an Evans and Sutherland series 'V' workstation supporting 'crystal eyes' stereo graphics, was used throughout. Residues were placed systematically on the human H-chain ferritin skeletal framework (Lawson et al., 1991) and where possible structural identities or similarities were preserved. Regions of the conformation away from symmetry axes were modelled first but, to investigate the effects of side-chain substitutions on inter-subunit contacts, the whole 24-mer molecule was generated by applying 432 symmetry operators. Special attention was paid to side-chain interactions, e.g. in the three-fold and four-fold channel regions, and to regions in which the sequence alignment suggested residue deletions. Idealized geometries, torsion angles (Ramachandran and Sasisekharan, 1968) and hydrogen-bond distances of 0.285 nm (2.85 Å) were generated throughout.

The initial molecular model was searched for potential haem-binding pockets: namely deep hydrophobic clefts which could be generated with minimum distortion and in which two methionine residues could be positioned so as to form axial haem iron ligands with the correct geometry and Fe–S distances. Two such pockets were found, one lying within the subunit and the other between subunits at the twofold axis interface. Protohaem IX groups could be fitted in both positions with minor side-chain shifts to avoid short contacts. These models were subjected to repeated cycles of geometrically restrained regularization using the method of Hermans and McQueen (1974).

The MOLSCRIPT drawing program (Kraulis, 1991) was used in preparing figures displaying interesting aspects of the molecule.

RESULTS

Spectroscopic characteristics of the over-produced *Ec*-BFR

Haem content and absorption spectra

As isolated from *E. coli* strain JRG 2033, BFR contains a variable amount of haem. Three different samples have been studied which contained 3–4, 8–9 and 10–11 haems per BFR molecule as extracted, see Table 1. The reason for the variable haem loading is unclear. The optical absorption spectrum of native BFR (preparation *Ec*-BFR-1, Table 1) shows peaks for the haem group at 418 nm and 531 nm (Figure 1). At higher gain a well-resolved peak is observed at 734 nm, as is the case for all samples of haem-containing BFR from several sources. The ratio of

Table 1 Haem and non-haem iron content of *Ec*-BFR

Preparation	Haem/protein (as isolated)	Fe/molecule	Total haem/protein after addition
Ec-BFR-1	3–4	24–26	~ 6*
			12–14†
Ec-BFR-2	10–11	67–68	12–14‡
Ec-BFR-3	8–9	31–42	8–9§

* Addition of haemin by titration followed by purification using DE52 anion exchange column.
† Addition of haem by titration. Ratio determined by 'Kadir and Moore' treatment of results.
‡ Based on increase in absorbance at 418 nm after incremental addition of haemin.
§ No significant change in absorbance at 418 nm after the addition of 12 molar excesses of haemin.

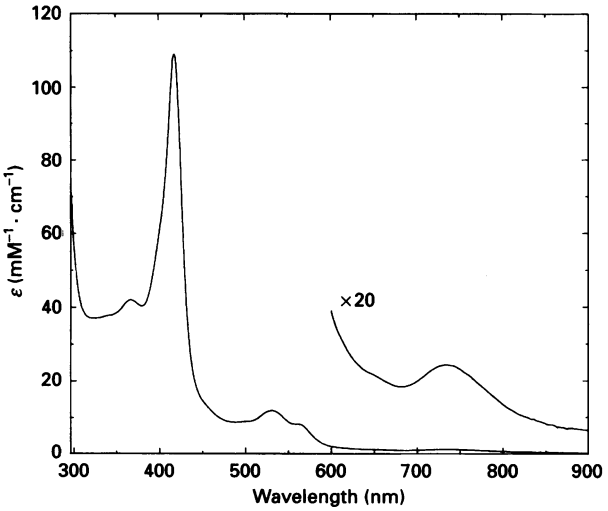

Figure 1 Electronic absorption spectrum of *Ec*-BFR

BFR in a mixture (1:1, v/v) of 25 mM Hepes (in ^2H_2O at p^{2}H 7.8) and glycerol. The spectrum was recorded relative to a buffer/glycerol baseline and was plotted using a haem concentration of 13 μM, as determined by integration of the low-spin ferric haem e.p.r. spectrum.

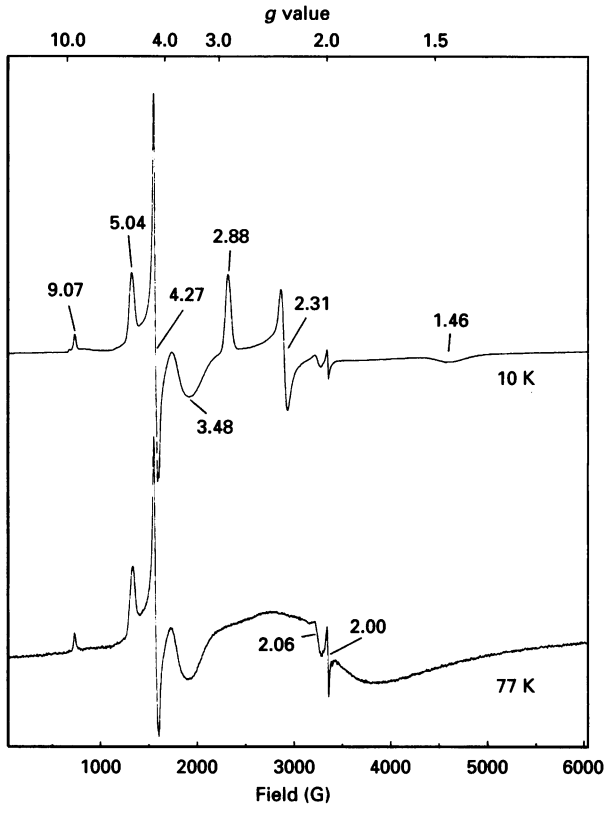

Figure 2 X-band e.p.r. spectrum of *Ec*-BFR

The protein sample was as used for Figure 1 but concentrated to a haem level of 675 μM. Temperatures are indicated. Modulation amplitude was 10 Gauss (1 Gauss = 10^{-4} T), at a frequency of 100 kHz. Microwave frequency 9.39 GHz; power 2.01 mW.

A_{418}/A_{280} is low, indicating either that the protein contains some iron core or that it has a low haem content. Analysis of the non-haem iron content give a value of 24–26 Fe atoms/molecule while the haem content was found to be 3–4 per 24 subunits.

E.p.r. spectra

The e.p.r. spectra of the same protein preparation recorded at several temperatures between 4 K and 77 K (Figure 2) indicate the presence of three distinct paramagnetic iron-containing centres in the protein, namely, the haem moiety, the iron core and monomeric non-haem iron. A sharp signal at $g = 2.0$, which arises from an organic radical, was also seen in all such spectra. The trio of g values at $g = 2.88$, 2.31 and 1.46 are characteristic of low-spin Fe(III) haem. Similar values assigned to the haem moiety have been reported in *Pa*-BFR and *Av*-BFR and *Rs*-BFR (Cheesman et al., 1990). These g values are distinctive for bis-coordinated protohaem IX in proteins and closely match those of a bis-thioether model Fe(III) porphyrin, namely, bis-tetrahydrothiophene octaethyl porphyrin Fe(III), which are $g = 2.94$, 2.30 and 1.43. It was this similarity that first suggested that the haem in BFR might be liganded by methionine. Proof of this assignment is provided by n.i.r.-m.c.d. spectra (McKnight et al., 1991) (see later).

At 77 K the e.p.r. signals from the haem are broadened by an increase in the electro spin relaxation rate. However, a broad derivative-shaped signal centred at a g value of approx. 2.0 appears. E.p.r. experiments on samples of *Pa*-BFR, with and without an iron core, assigned this signal to the iron core (Cheesman et al., 1992). The rather weak intensity of the signal is consistent with the low iron-core content of the recombinant BFR.

The e.p.r. spectrum also shows a series of signals between g values of approx. 3.0 and 10.0 which are characteristic of magnetically isolated, monomeric non-haem Fe(III) in the high spin state, $S = 5/2$. As isolated, many proteins contain a derivative-shaped e.p.r. signal centred at a g value of approx. 4.3. This is due to an impurity of Fe(III) bound at low-symmetry sites within the protein. The e.p.r. signals of *Ec*-BFR are not of this type. Its spectrum in frozen buffer solution (no glycerol) shows a prominent derivative shape at $g = 4.27$ and, at higher gain, a number of other signals between $g = 9.67$ and $g = 5.0$ (Figure 3a). Addition of glycerol to a level of 50% (v/v) caused the replacement of the e.p.r. signals by a set of sharp features at $g = 3.48$, 4.27, 5.04, and $g = 9.07$ and 9.85 [Figures 2 and 3(b)]. There is a broad weak signal peaking at a g value of approx. 8.3 detectable at the lowest temperatures of the measurement, which may be due to a residual amount of one of the species observed in the absence of glycerol. The g values observed in the presence of glycerol can be assigned to a single species of $S = 5/2$ Fe(III) ion.

Under the action of low-symmetry fields from the ligand environment the ground state of high-spin Fe(III) ion is split into three pairs of Kramers doublets, separated in energy by the zero-field splitting term. Each pair is further split in an applied magnetic field and can give rise to three sets of g values. A total of nine g values could, in principle, be observed in the e.p.r. spectrum of monomeric Fe(III) ion. In practice some transitions lie at very high field (low g value) and are too broad to detect. The g values of 3.48, 4.27 and 5.04 can be assigned to the middle pair of the three Kramers doublets. The temperature dependence of the intensity ratios of the signals at g values of 9.85 and 9.07 shows that they arise, respectively, from the lowest and highest Kramers doublets. The set of g values cannot be fitted to a simple spin Hamiltonian. The zero-field splitting must be comparable to, or less than, the energy of the microwave photons, 0.33 cm^{-1}.

These results clearly show that in *Ec*-BFR there are magnetically isolated high-spin non-haem Fe(III) sites with some structural variation. In the presence of glycerol all of these sites are converted into a single species as detected by the e.p.r.

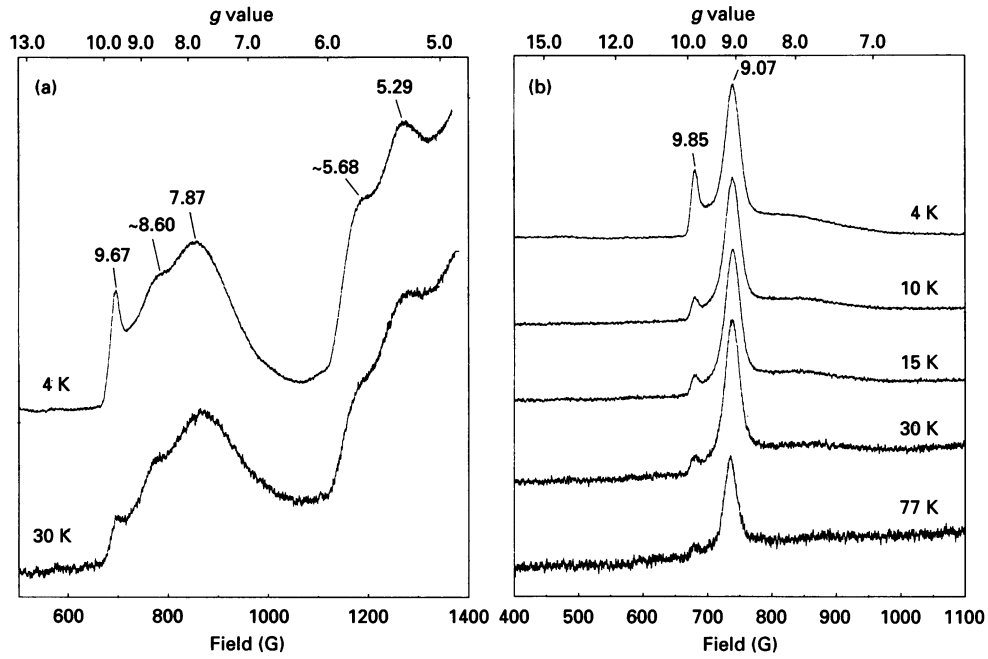

Figure 3 X-band e.p.r. spectrum of *Ec*-BFR

(**a**) BFR (1.35 mM in haem) in 25 mM Hepes (in ^2H_2O at p^{2}H 7.8) and (**b**) plus glycerol (1:1, v/v). Temperatures are indicated. Modulation amplitude was 10 Gauss (1 Gauss = 10^{-4} T), at a frequency of 100 kHz. Microwave frequency 9.39 GHz; power 2.01 mW.

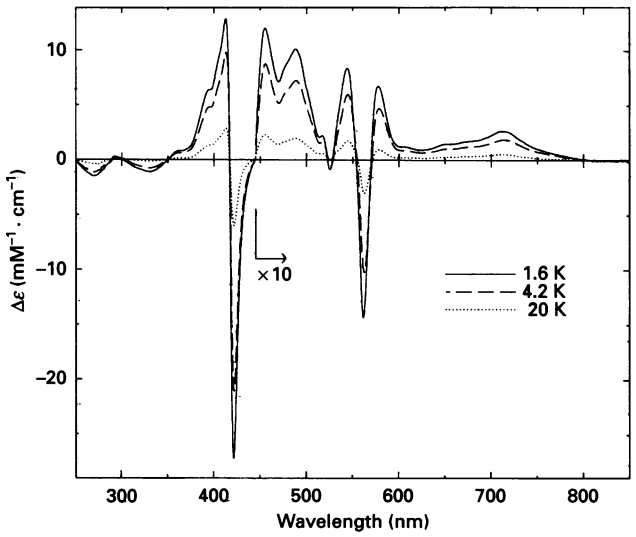

Figure 4 U.v.-visible m.c.d. spectra of *Ec*-BFR

Spectra were recorded with a magnetic field at 5 T and at temperatures as shown. The sample used was as for the e.p.r. spectra of Figure 2.

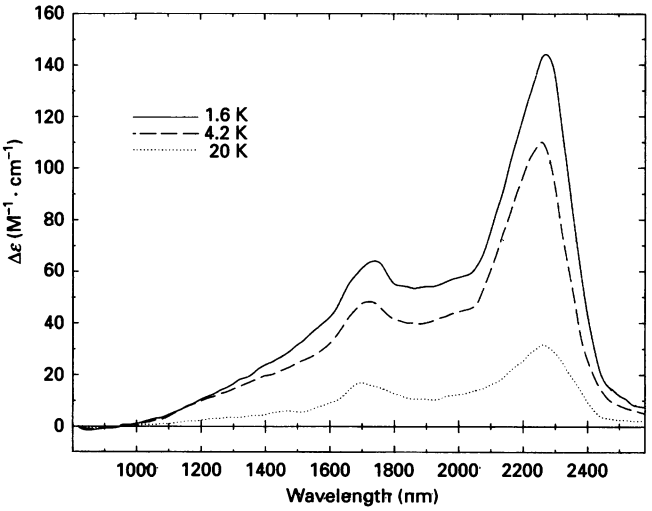

Figure 5 N.i.r.-m.c.d. spectrum of *Ec*-BFR

Instrumental and sample details were as for Figure 4.

spectrum, and must therefore have the same structure. Although varying somewhat in intensity, similar e.p.r. signals have been recorded from samples of *Ec*-BFR isolated from three different cell growths. Quantification of the number of Fe(III) ions is not straightforward given the complexity of the ground state, but the signals are intense and clearly represent more than a minor, adventitious component of Fe(III).

M.c.d. spectra

M.c.d. spectra have been recorded over the wavelength range 250–2600 nm in a 5 T field and at several temperatures between 1.4 K and 20 K (Figures 4 and 5). All the features in the spectrum are due to the haem moiety of *Ec*-BFR. The spectrum is quantitatively indistinguishable from the m.c.d. spectra of the haem groups in *Pa*-BFR and *Av*-BFR (Cheesman et al., 1992). The optical transitions lying between 250 and 600 nm, are due primarily to the π electrons of the porphyrin ring.

The broad bands which peak at 713 nm and 2270 nm are the distinctive features of the m.c.d. spectra of BFR (Cheesman et al., 1990). The band at 2270 nm is assigned through evidence from a variety of sources as the porphyrin $\pi(a_{1u})$-to-Fe(III) d_{yz}charge-transfer (CT) transition (McKnight et al., 1991). The

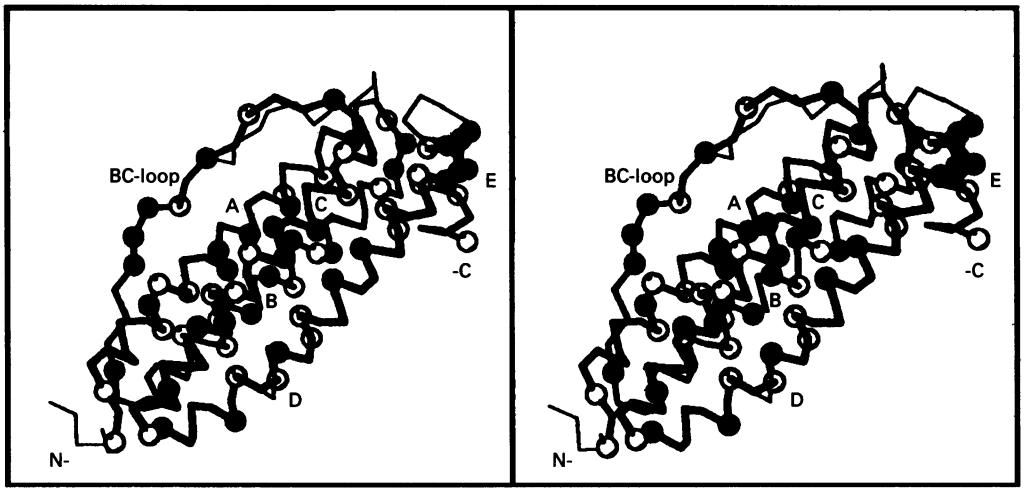

Figure 6 Comparison between subunit conformations of BFR and H-chain of human ferritin

Proposed BFR subunit conformation (thick lines) was superimposed on that of human H-chain ferritin (thin lines). Positions of residues (α-carbons) that are conserved in both molecules are shown as closed circles and semi-conserved residues as open circles. The central part of the four-α-helix bundle is highly conserved.

width of the band envelope and the subsidiary structure to shorter wavelength arise from co-excitation of vibrational quanta of the porphyrin ring, which are coupled to the CT transition. For example, the energy difference between the peak at 2270 nm and the first well-resolved side-band at 1700 nm is 1480 cm^{-1}, which is typical of the aromatic stretch of the haem ring.

The assignment of the band at 713 nm is not so well established. This peak has been detected in the m.c.d. spectra of *Pa*-BFR, *Rs*-BFR and *Av*-BFR, as well as in that of *Ec*-BFR. The band is unaffected by removal of the iron core. The m.c.d. magnetization characteristics of this band at 1.6 K, 4.2 K and 10 K are indistinguishable, within experimental error, from that of the peak at 2270 nm. The shape of the m.c.d. magnetization curve of low-spin Fe(III) haem depends on the ground-state g values and also on the polarization of the optical transition. The close agreement between the magnetization curves at 2270 nm and 713 nm suggests that the two transitions both have the same polarizations. It is established, both from polarized single-crystal spectroscopy and from the assignment of the m.c.d. spectrum, that the CT band $a_{1u} \rightarrow d_{yz}$ is polarized in the plane of the haem ring (Makinen and Churg, 1983). Therefore the band at 713 nm must also be polarized in the haem-ring plane.

Three possible origins of the transition at 713 nm are considered. The first is a d–d transition of low-spin Fe(III), the second is a CT transition from the sulphur orbitals of the axial methionine ligands, and the final possibility is that the transition is a porphyrin-to-Fe(III) CT transition. The absorbance coefficient of the absorption band is approx. 900 $M^{-1} \cdot cm^{-1}$, too intense to be assigned to a d–d transition. The polarization of the band eliminates the assignment to a CT transition from the pair of axial sulphur ligands, since such a transition would be strongly polarized perpendicular to the haem plane. This leaves assignment as a porphyrin(π)-to-Fe(d) CT transition. The obvious candidate is the $\pi(a_{2u}) \rightarrow Fe(III)d_{yz}$ transition from the next highest molecular orbital of the porphyrin ring, namely, $a_{2u}(\pi)$. Both transitions $\pi(a_{1u})$ and $\pi(a_{2u})$ to Fe(III) have been identified in Fe(III) 5,10,15,20-tetraphenylporphyrinate complexes (McKnight et al., 1991) with an energy separation of 2700 cm^{-1}. In the case of Fe(III) protohaem complexes only one transition $\pi(a_{1u}) \rightarrow Fe(III)d_{yz}$ has been detected. However, the low ligand field strength provided by two methionine ligands shifts the $a_{1u} \rightarrow d_{yz}$ transition to 2270 nm ($\equiv$ 4400 cm^{-1}). The energy of the $a_{2u} \rightarrow d_{yz}$ transition should also be shifted to low energy. In bis-ligated complexes with axial ligands of stronger field the $\pi(a_{2u}) \rightarrow Fe(III)d_{yz}$ transition is presumed to lie under the intense π–π^* transitions of the visible region of the porphyrin ring, and hence is not detectable.

Stoichiometry of haem binding to *Ec*-BFR

Ec-BFR extracted from different batches of cells showed considerable variation in haem content, from 3–4 to 10–11 haems per protein (Table 1). However, attempts were made to increase the haem content of *Ec*-BFR by addition of haemin, as was done in the case of *Pa*-BFR by Kadir and Moore (1990). Difference absorption spectra in the visible region were recorded of *Ec*-BFR after each addition of haemin chloride solution. The binding curve shows an initial increase in absorption at 418 nm, which is the peak wavelength of the Soret band of *Ec*-BFR. However, subsequent additions of haemin chloride solution led to proportionately lower increases in absorbance. Several different titrations have been carried out with batches of *Ec*-BFR both low (3–4 per protein) and high (10–11 per protein) in initial haem. Similar binding curves were obtained. These curves could not be analysed in terms of a simple binding process.

Ec-BFR, to which haemin chloride had been added by titration, was passed down a chromatography column to free the protein from any surface-bound haemin. The absorption spectra of the resulting solutions were indistinguishable from those of the starting material except for an increase in the haem/protein ratio. The maximum ratio obtained was 12–14 haems per molecule.

E.p.r. spectra of the haem-loaded samples show that the added and intrinsic haem are spectroscopically indistinguishable, giving the typical low-spin Fe(III) haem g values at 2.88, 2.31, 1.46. The amount of high-spin haem, estimated from the feature at a g value of approx. 6.0, has increased. Spin integration of the low-spin haem resonance at $g = 2.88$ shows a total loading of no more than 12 haems per BFR molecule. These binding experiments were all performed in the absence of glycerol.

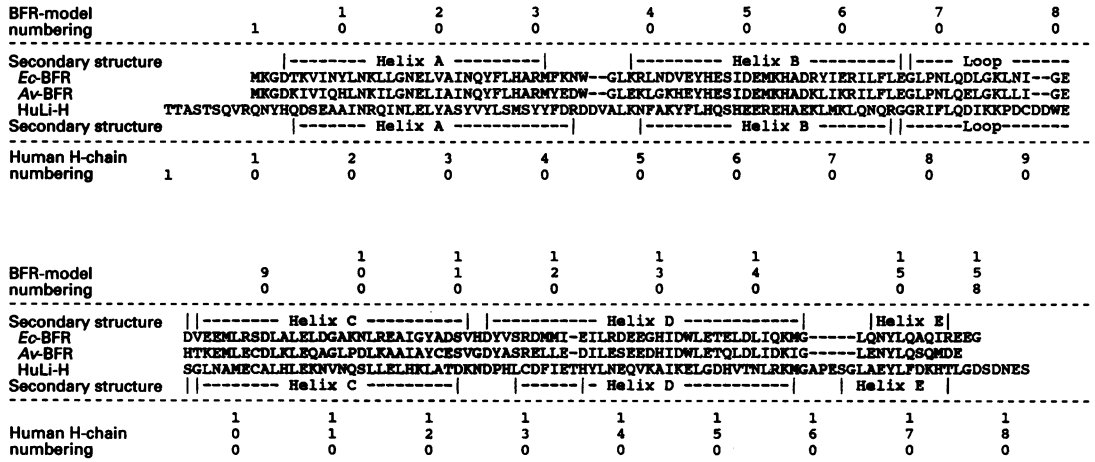

Figure 7 Amino-acid sequences of *Ec*-BFR and *Av*-BFR aligned with that of human H-chain apoferritin

Alignment is made according to the model of Figure 6.

A three-dimensional model structure for the *Ec*-BFR

Construction of the computer graphics model of BFR required only minor modifications of both the human H-chain framework and of the original sequence alignment, as outlined below. Although many side-chain substitutions had to be introduced, a feasible structure could be built for BFR with good geometry and without short intra- or inter-subunit contacts. Features of the model are described below, with special emphasis on those that are pertinent to the spectroscopic studies.

The polypeptide-chain fold

The proposed BFR subunit fold, Figure 6, mainly comprises a four-α-helix bundle which is 5.2 nm (52 Å) long and 2.2 nm (22 Å) wide, with a short fifth helix, as in ferritin. Amino acids within helices are as follows: A (residues 4–31), B (38–66), C (83–111), D (113–145), E (147–155) and the remaining residues lie within the following non-helical segments: N-terminus (residues 1–3), AB-turn (32–37), BC-loop (67–82), CD-turn (112), DE-turn (146) and C-terminus (156–158). Figure 6 shows the main chain of the BFR model superposed on the known conformation of human H-chain (Lawson et al., 1991) and indicates that many side-chains within the bundle are either conserved or conservatively replaced.

Some sequence deletions (as shown in Figure 7 and described below) were required for optimum sequence alignment and conservation of structure.

(i) The N-terminus. This is shorter by nine residues than in human H-chain and a number of stabilizing interactions are therefore lost. However, Lys-2 bonds to the main-chain carbonyl of residue 112 in the CD-turn and the main chain carbonyl of Met-1 is linked through a hydrogen bond to Arg-102 of a symmetry-related subunit.

(ii) The AB-turn. The AB-turn is reduced from four residues to two. Trp-35 and Gly-36 are conserved in all known BFR sequences (Andrews et al., 1991a; Grossman et al., 1992). In the model Gly-36 allows the adoption of a tight turn, while the side-chain of Trp-35 appears to block one end of the four-helix bundle which may be important for the formation of a haem pocket (see below).

(iii) The BC-loop. In human H-chain ferritin the loop contains a kink at Pro-88. In BFR this residue is replaced by Leu-77, enabling the BC-loop to adopt a more extended conformation and thus to follow a more direct course.

(iv) A single residue deletion in the D-helix. The D-helix in human H-chain ferritin is bent owing to the incorporation of an additional residue (His-136) in one turn of the helix. In BFR the equivalent residue is absent and the D-helix is modelled with a straight helix. This seems to present no problems with respect to the maintenance of either the subunit conformation or inter-subunit packing.

(v) The DE-turn, E-helix and C-terminus. The alignment of the E-helix has been displaced by one residue compared with the computer prediction of Andrews et al. (1991a) to optimize similarity of E-helices (see Figure 7). The penultimate residue of the D-helix is Met-144, which is in the same place as the equivalent Met-158 of human H-chain. The following four residues then adopt a tighter fold than the nine residues of human H-chain but the side-chain position of Leu-146 leading into the E-helix mimics that of Leu-165 in human H-chain ferritin. The model thus enables the E-helices of *Ec*-BFR to lie in register with those of human H-chain ferritin. Leu-150 is now equivalent to Leu-169 of human H-chain and Ile-154 takes the place of His-173. Tyr-149 can form a hydrogen bond to the main-chain N atom of Arg-155 in a neighbouring subunit, as it does in human H-chain ferritin. The C-terminus is curtailed by five residues (most of the nine residues extending from the E-helix of human H-chain are not clearly defined in the electron density map).

Quaternary structure and inter-subunit channels

In the modelled BFR structure the application of 432 symmetry produces a roughly spherical shell allowing numerous inter-subunit interactions and the generation of fourfold and threefold channels.

(i) Fourfold channels. The positioning of the E-helices described above allows the formation of hydrophobic channels lined by four leucine residues at the outside and centre (residues 146 and 150 respectively) and isoleucine residues (154) at the inside surfaces (Figure 8a). These channels closely resemble those of horse and rat L-chain ferritins where the fourfold channel has twelve leucine residues (Ford et al., 1984; Lawson, 1990). It is unlikely that metal ions would bind in these channels.

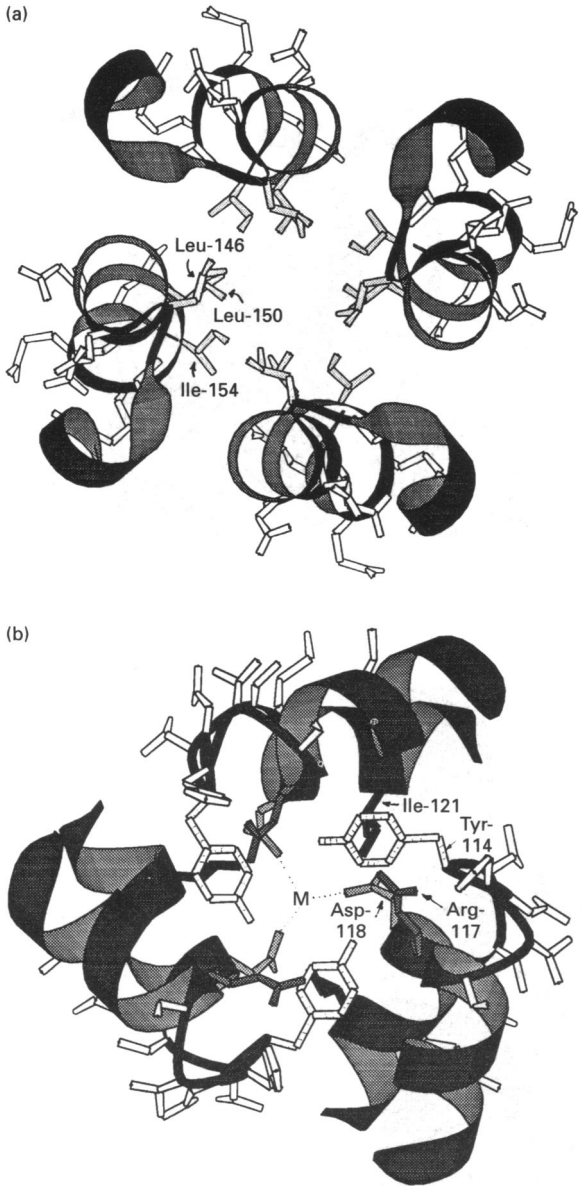

Figure 8 The fourfold and threefold channels in the *Ec*-BFR model

(**a**) Fourfold channel; (**b**) threefold channel. In (**b**) a possible metal site, M, has three symmetry-related aspartate ligands (Asp-118). In *Av*-BFR these aspartate residues are replaced by glutamate residues.

(ii) Threefold channels. As modelled these channels are very different from those of mammalian ferritins, in which the highly conserved Asp-131 and Glu-134 are in a position to bind metal ions (Ford et al., 1984). In *Ec*-BFR the glutamatic acid residues are replaced by isoleucines (residue 121) which would not act as metal ligands, but three Asp-118 residues (the equivalent of Asp-131) could be positioned to provide a trigonal arrangement of carboxyl oxygens at about 0.22 nm (2.2 Å) from a central metal ion, Figure 8(b).

Putative ferroxidase centre flanked by hydrophobic cores

The sequence alignment of Andrews et al. (1991a) indicates that five out of the seven H-chain ferroxidase centre residues are conserved. In the model residues Glu-18, Glu-51, His-54, Tyr-25 and Glu-94 are indeed in the same position as their equivalent residues in ferritin (Glu-27, Glu-62, His-65, Tyr-34 and Glu-107

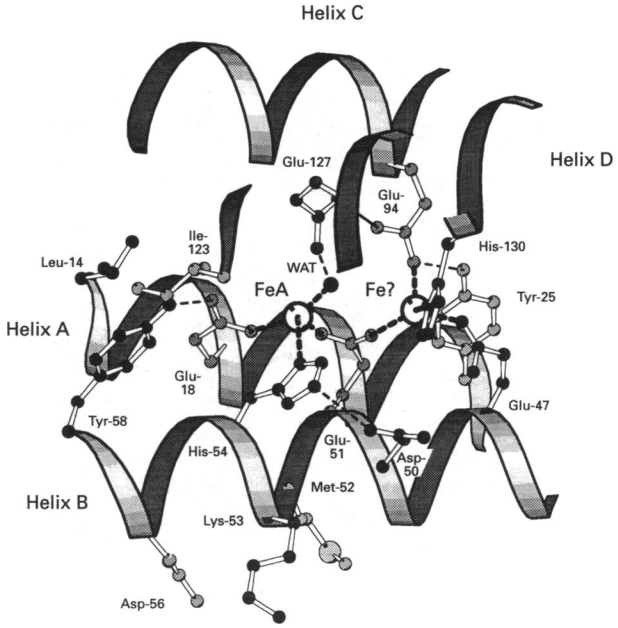

Figure 9 Central 'ferroxidase centre' region of BFR subunit

FeA corresponds to a metal site found in human H-chain ferritin. FeA is also one member of a μ-oxo-bridged Fe(III) dimer proposed to form at the ferroxidase centre (Andrews et al., 1991a; Treffry et al., 1992). Fe? corresponds to a hypothetical metal site not available to human H-chain ferritin. Fe? lies about 0.5 nm (5 Å) from FeA.

respectively). Glu-18 and Glu-51, His-54 and Glu-94 could thus act as metal ligands, Figure 9, and may possibly allow the formation of μ-oxo-bridged Fe(III)-dimers (see Treffry et al., 1992). Tyr-25 is able to hydrogen bond to Glu-94, but the carbonyl group of Asp-50 is unable to move within the bonding distance of either of the Fe sites proposed for human H-chains. Gln-141, the residue thought to stabilize the binding of dioxygen at the ferroxidase centre (Treffry et al., 1992) is replaced by Glu-127 in the BFR model. Other changes are also observed in this region of the structure. Glu-27 of the human H-chain forms a hydrogen bond with Gln-23 (Lawson, 1990), but the corresponding interaction is absent in the BFR model since Gln-23 is replaced by Leu-14. On the other hand a salt bridge between Arg-57 and Glu-122 replaces a much weaker interaction between Lys-68 and His-136. An interesting substitution of a pair of residues is found: Tyr-137 of the human H-chain D-helix makes a hydrophobic contact with Leu-69 of the B-helix. In the BFR model these are replaced by a similar, but inverted pair, namely Ile-123 and Tyr-58, the ring of the latter being only slightly displaced from that of Tyr-137 (a conserved residue in H-chains). Another possibly more significant difference between human H-chain ferritin and the model structure of *Ec*-BFR, arises from the substitution of Ala-144 by His-130 and Gln-58 by Glu-47. Glu-47 and His-130 may be positioned along with either Gln-94 or Glu-51 to provide ligands of an alternative hypothetical metal centre, Figure 9. This would lie close to the ferroxidase centre metal positions of human H-chains, but somewhat displaced towards the C-terminal end of the bundle (see also Figure 10a).

The ferroxidase centres of H-chains and the corresponding residues of L-chains provide a hydrophilic region within the centre of the four-helix bundle flanked by two hydrophobic 'cores' (Lawson, 1990). This is also true of the proposed BFR-subunit fold. At the N-terminal end of the bundle Tyr-10, Leu-11, Leu-14, Tyr-58, Leu-101 and Met-120 form a buried hydrophobic region which replaces that of H-chains comprising respectively Ala-19, Ile-20, Gln-23, Leu-69, Leu-114 and Ile-133.

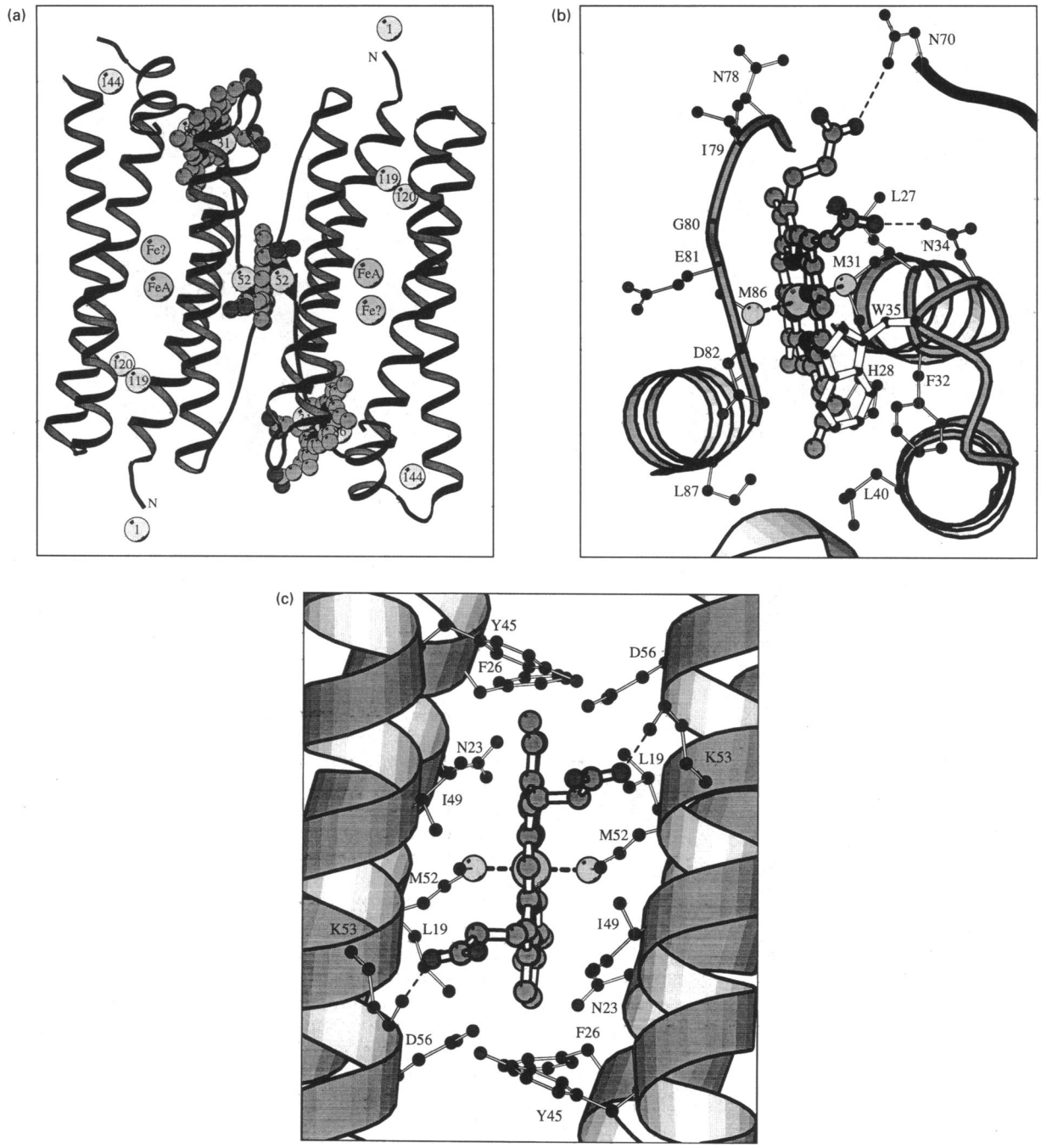

Figure 10 Alternative haem-binding sites in the *Ec*-BFR model

(**a**) shows a pair of subunits with protohaem IX bound at both sites I and II. Site I is intrasubunit and site II intersubunit with haems lying at twofold symmetry axes. It is not envisaged that haem binds at both types of site. (**b**) shows a close-up of site I and (**c**) a close-up of site II. In (**a**) all the methionine residues found in *Ec*-BFR are given as numbered circles. Only methionines 1, 31, 52 and 86 are conserved in *Av*-BFR. (**a**) also shows the positions of Fe atoms binding at or near the putative ferroxidase centre.

Again at the C-terminal end there is a cluster of hydrophobic residues. This may provide a haem-binding pocket which is described below (site I). Near to the 'ferroxidase centre' region there are four conserved apolar residues, namely Leu-19, Ala-55, Leu-93 and Ile-131 (respectively Leu-28, Ala-66, Leu-106 and Ile-145 in human H-chains).

Haem-binding sites

If the BFR model is correct then it must offer a stereochemically feasible site for haem binding, with methionines aligned in a position that allows bis-methionine axial ligation (Cheesman et al., 1990). In fact two such sites must be considered: site I, which is fully contained within the subunit and site II, which is an intersubunit site lying at twofold symmetry axes, Figure 10.

Site I lies at the C-terminal end of the four-helix bundle, Figures 10(a) and 10(b). A hydrophobic pocket comprising residues Leu-27, His-28, Met-31, Phe-32, Trp-35, Leu-40, Val-83, Met-86 and Leu-87 contains two methionine residues (31 and 86) positioned opposite one another such that, if a haem group is inserted into the pocket, their two sulphur atoms lie approx.

0.23 nm (2.30 Å) from the haem Fe atom. Trp-35 at the end of the pocket seems to control haem access. Haem binding probably would require the swinging out of this tryptophan. Some movement of the BC-loop and a reorientation of His-28 (which replaces Met-39 of human H-chains) has been made. However, residues can be aligned so as to allow the positioning of haem within the pocket without unduly short contacts. In this site, Figure 10(b), haem occupies a position similar to that of Trp-93 of H-chains. Haem propionates are modelled to interact with Asn-34 and Asn-70.

A distinctive feature of site I is that (given some minor conformational changes) it is accessible from the outside of the molecule. Haem is positioned within the bundle such that its Fe atom and its nearest edge are 1.9 nm (19 Å) and 1.3 nm (13 Å) from the ferroxidase centre respectively.

Site II is shown in Figures 10(a) and 10(c). In contrast to site I, site II would not be accessible to haem titration in intact molecules. It is a symmetrical hydrophobic pocket forming an indentation from the cavity surface at the centre of a long hydrophobic interface between subunits. The pocket is lined by residues Leu-19, Ile-22 and Phe-26 on the A-helices, Tyr-45, Ile-49 and Met-52 on the B-helices and Leu-74 of the BC-loops of the two subunits and the equivalent residues of a twofold axis-related subunit. Met-52 and Met-52′ lie in positions such that their sulphur atoms can make S–Fe bonds of 0.23 nm (2.30 Å). The haem edges are flanked by residues Tyr-45 and Tyr-45′. Propionates can be positioned on the cavity side such that they hydrogen bond to Lys-53. Haem Fe atoms lie at 1.3 nm (13 Å) from ferroxidase centre Fe atoms and these Fe atoms are linked through ten bonds.

In human H-chains, only one residue of the site I haem pocket is conserved (the phenylalanine) and this is also the case for site II (Leu-19/28). Neither of these regions would be haem-binding sites in human H-chain ferritin.

Modified sequence alignment for BFRs and ferritin

Only minor modifications of the computer alignment of Andrews et al. (1991a) were introduced during modelling. The most important of these was a shift by one residue of the 12 C-terminal residues so as to provide fourfold channels more closely resembling those of the ferritins. Figure 7 shows the modified alignment of *Ec*-BFR and human ferritin H-chain corresponding to the model. Figure 7 includes the complete sequence of *Av*-BFR recently published by Grossman et al. (1992). These authors also concluded that the alignment of BFR and ferritin primary sequences meant that BFR would share major structural features with those of the ferritin family. Their alignment of the first 86 residues (in helices A and B and the BC-loop) and of the last residues (E-helices) is identical to that given in Figure 7. However, the alignment suggested by them for residues 94–145 would be incompatible with the ferritin structure, since it would introduce gaps of two, three and one residue(s) within the C-helix and cause the D-helix to be shifted by seven residues. The computer modelling shows that a feasible structure can be built without such major disruption.

DISCUSSION

The spectroscopic results establish that in *Ec*-BFR, the haem iron has bis-methionine ligation, as in BFRs from three other species. This implies strong conservation of the binding site. The e.p.r. spectra for *Ec*-BFR in glycerol also indicate the presence of single high-spin non-haem Fe(III) in a low-symmetry site. This site has not been described previously. A credible structural model must account satisfactorily for these and other results, including the conservation within the BFR family of the amino-acid residues involved in haem- and non-haem-iron-binding sites.

Haem-binding site

Of the methionine residues present in *Ec*- and *Av*-BFRs only four are conserved in both species: Met-1, Met-31, Met-52 and Met-86. Grossman et al. (1992) proposed an intersubunit site similar to that of site II and considered an alternative intersubunit site with Met-1 and Met-86 as ligands. This latter site is not feasible for a ferritin-like model because the methionine residues are incorrectly orientated and cannot approach closely enough to bind haem Fe(III). In the five partial N-terminal sequences (see Andrews et al., 1991a) Met-31 and Met-52 are conserved in all except *Nitrobacter winogradskyi*, which contains leucine in place of Met-31 (Kurokawa et al., 1989), and *Synechocystis* BFR, which contains threonine in place of Met-52 (Laulhère et al., 1992). Since the multiple sequence alignment strongly supports a ferritin-like structure, it is conceivable that one of these residues represents a sequencing error.

Of the nine residues forming the hydrophobic pocket of site I, Figures 10(a) and 10(b), all but two are conserved in *Av*-BFR, in which Phe-32 is replaced by tyrosine (conservatively) and Val-83 by threonine. In other BFRs conservative replacements are found at position 27 (leucine or isoleucine). Phenylalanine, tyrosine, leucine or glutamine are found at position 32. Site II residues Leu-19, Ile-22, Phe-26, Tyr-45, Ile-49 and Met-52 are conserved in *Av*-BFR. However, substitutions are found at positions 22, 26, 45, 49 (and 52), some of which are non-conservative, e.g. *Synechocystis* BFR has Arg-22 and Asp-45 (Laulhère et al., 1992) in place of Ile-22 and Tyr-45 in both *Ec*- and *Av*-BFR.

The maximum stoichiometry of haem to protein found in these studies with *Ec*-BFR (even after addition of extra haem) and also reported for *Av*-BFR (Stiefel and Watt, 1979) is about 12 per 24-mer, consistent with haem binding at site II. The observation that *Pa*-BFR molecules bind 24 haems (Kadir and Moore, 1990) is compatible with site I but not site II. The fit of haem into site I is tight and the need for movement of Trp-35 and other minor structural rearrangements could explain the observed difficulty in obtaining full haem loading. Site I is accessible from the outside of the molecule and this is consistent with haem accessibility in contrast with site II, which could be accessed only after subunit dissociation. Grossman et al. (1992) and more recently Watt et al. (1992) have proposed the intersubunit site as the probable binding location of haem. As this stage it is not possible to choose unequivocally between sites I and II, but the balance of evidence seems to favour site I. The finding that the second iron-storage protein of *E. coli*, a non-haem ferritin, FTN, has methionine at position 52 (Izuhara et al., 1991) but no haem (A. J. Hudson, S. C. Andrews, P. M. Harrison and J. R. Guest, unpublished work), may be taken as indirect support for haem binding at site I in BFR. Attempts to bind haem to FTN to generate a well-defined low-spin site have not been successful (N. E. Le Brun, M. R. Cheesman, G. R. Moore and A. J. Thomson, unpublished work).

Non-haem Fe(III) site

In human H-chain ferritin metal sites have been observed in five positions (Lawson et al., 1991). One of these is on the outside of the molecule. Its ligands are Asp-84 and Gln-82. Gln-82 is not present in native ferritin H-chains and was introduced by

engineering to create intermolecular metal-bridge contacts for crystallographic studies. It is also absent from BFR. A second site, lying in the threefold channels, has three Glu-134 ligands and three Asp-131 bound through water molecules. A site having three symmetry-related Asp-131 residues similar to that found in horse L-chain ferritin (Ford et al., 1984) may be possible for *Ec*-BFR, Figure 8(*b*), although in *Av*-BFR this aspartate is substituted by glutamate. Ligands of a third H-chain metal site, Glu-61 and Glu-64, are replaced by Asp-50 and Lys-53 in BFRs. The remaining two metal sites in human H-chain ferritin (occupied alternatively by Tb^{3+} in crystals containing $TbCl_3$) lie about 0.31 nm (3.1 Å) apart within the subunit, at a position identified as the ferroxidase centre. With the exception of Glu-61, which is replaced by Asp-50, ligands Glu-27, Glu-61, Glu-62, His-65 and Glu-107 (some of which can bind both metals) are conserved in both BFRs (as Glu-18, Glu-51, His-54 and Glu-94). Therefore it is suggested that in BFR these residues may represent a ferroxidase centre capable of binding and oxidizing Fe(II) (Andrews et al., 1991a). In human ferritin it has been proposed (Treffry et al., 1992) that this centre can accommodate a dinuclear iron centre, possibly related to those of haemerythrin and ribonucleotide reductase (Que and True, 1990) and also observed in ferritin by Mössbauer spectroscopy (Bauminger et al., 1989, 1991). Because the two iron atoms are close together and bridged by ligands, they are magnetically coupled and hence e.p.r.-silent in the fully oxidized Fe(III)/Fe(III) state. In the fully reduced Fe(II)/Fe(II), a broad e.p.r. signal at low field may be detectable from an even-spin ground state. The semi-reduced state, Fe(III)/Fe(II), is stable only in certain conditions and generates an unusual e.p.r. signal with all components of the *g*-tensor below 2.0. No such signal has been detected from the BFR samples studied so far. However, occupancy of the ferroxidase site by a single Fe(III) iron would give rise to e.p.r.-detectable non-haem iron signals. Isolated Fe(III) atoms have been observed by Mössbauer spectroscopy as small amounts of Fe(II) are oxidized by apoferritin (Bauminger et al., 1989, 1991). It is uncertain where this Fe(III) is bound but it may have arisen, at least in part, by dissociation of the Fe(III) dimers (Treffry et al., 1992). In ferritin there is evidence from Mössbauer spectroscopy that both single and dimeric Fe(III) sites are labile (Bauminger et al., 1989, 1991). In *Ec*-BFR the isolated Fe(III), although evidently less mobile, does exhibit structural variation. It could be located at the ferroxidase centre or at the alternative neighbouring site, Figure 9. Two of the putative ligands, Glu-47 and His-130, are conserved in *Av*-BFR, although not in ferritin. The e.p.r.-detectable centre interacts with glycerol. This could be due either to direct ligation of the Fe(III) by glycerol or to a structural change of the metal site induced by change of solvent from water to water/glycerol (50:50, v/v).

The e.p.r. spectra of the single non-haem Fe(III) generated in the presence of glycerol is characteristic of iron in a low-symmetry site. Its signal is given by all three batches of *Ec*-BFR prepared to date. These samples vary somewhat in their core content, although it is low, being 30–75 atoms, and the broad core spectra are also seen in all samples examined. Although it is possible that the monomeric non-haem Fe(III) is located on the inside surface of the protein shell, it is difficult to understand why it should not be part of the core if this were the case.

Hence it is proposed that the monomeric iron is located within the shell, probably at or near the putative ferroxidase centre.

In summary, although the validity of the BFR model will ultimately be tested by completion of the X-ray analysis, it provides possible sites for both haem groups and non-haem iron that are compatible with the spectroscopic data and that can be tested with the aid of site-directed mutagenesis.

This work has been supported by the Science and Engineering Research Council, via its Molecular Recognition Initiative and by the Wellcome Trust, both at Norwich and at Sheffield.

REFERENCES

Aåsa, R. and Vänngård, T. (1975) J. Magn. Res. **19**, 308–315
Andrews, S. C., Smith, J. M. A., Guest, J. R. and Harrison, P. M. (1989a) Biochem. Biophys. Res. Commun. **158**, 489–496
Andrews, S. C., Harrison, P. M. and Guest, J. R. (1989b) J. Bacteriol. **171**, 3940–3947
Andrews, S. C., Smith, J. M. A., Yewdall, S. J., Guest, J. S. and Harrison, P. M. (1991a) FEBS Lett. **293**, 164–168
Andrews, S. C., Findlay, J. B. C., Guest, J. R., Harrison, P. M., Keen, J. N. and Smith, J. M. A. (1991b) Biochem. Biophys. Acta **1078**, 111–116
Bauminger, E. R., Harrison, P. M., Nowik, I. and Treffry, A. (1989) Biochemistry **28**, 5486–5493
Bauminger, E. R., Harrison, P. M., Hechel, D., Nowik, I. and Treffry, A. (1991) Biochim. Biophys. Acta **1118**, 48–58
Cheesman, M. R., Thompson, A. J., Greenwood, C., Moore, G. R. and Kadir, F. H. A. (1990) Nature (London) **346**, 771–773
Cheesman, M. R., Kadir, F. H. A., Al-Basseet., J., Al-Massad, F,. Farrar, J. A., Greenwood, C., Thompson, A. J. and Moore, G. R. (1992) Biochem. J. **286**, 361–368
Chen, M. and Crichton, R. R. (1982) Biochim. Biophys. Acta **707**, 1–6
Falk, J. E. (1964) in Porphyrins and Metalloporphyrins. BBA Lib., vol. 2, Elsevier, North Holland
Ford, G. C., Harrison, P. M., Rice, D. W., Smith, J. M. A., Treffry, A., White, J. L. and Yariv, J. (1984) Philos. Trans. R. Soc. London, B **304**, 551–565
Grossman, M. J., Hinton, S. M., Minak-Bernero, V., Slaughter, C. and Stiefel, E. I. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 2419–2423
Hermans, J. and McQueen, J. E. (1974) Acta Crystallogr. Sect. A **30**, 730–739
Izuhara, M., Takamune, K. and Tanaka, R. (1991) Mol. Gen. Genet. **225**, 510–513
Jiudi, L., Jiwen, W., Zepu, Z., Ya, T. and Bei, D. (1980) Sci. Sin. Engl. Ed. **23**, 897–904
Jones, T. A. (1978) J. Appl. Crystallogr. **11**, 268–272
Kadir, F. H. A. and Moore, G. R. (1990) FEBS Lett. **271**, 141–143
Kraulis, P. J. (1991) J. Appl. Crystallogr. **214**, 946–950
Kurokawa, T., Fukomori, Y. and Yamanata, T. (1989) Biochim. Biophys. Acta **976**, 135–139
Laulhère, J. P., Laboure, A. M., van Wuystwinkel, O., Gagnon, J. and Briat, J. F. (1992) Biochem. J. **281**, 785–793
Lawson, D. M. (1990) Ph.D. Thesis, University of Sheffield
Lawson, D. M., Treffry, A., Artymuik, P. J., Harrison, P. M., Yewdall, S. J., Luzzago, A., Cesareni, G., Levi, S. and Arono P. (1989) FEBS Lett. **254**, 207–210
Lawson, D. M., Artymiuk, P. J., Yewdall, S. J., Smith, J. M. G., Thomas, C. D., Shaw, W. V. and Harrison P. M. (1991) Nature (London) **349**, 541–544
Makinen, M. W. and Churg, A. K. (1983) in Iron Porphyrins, Part I (Lever, A. B. P. and Gray, H. B., eds.), ch. 3, Addison–Wesley Publishing Co., Reading
McKnight, J., Cheesman, M. R., Reed, C. A., Orosz, R. D. and Thomson, A. J. (1991) J. Chem. Soc. Dalton Trans. 1887–1894
Moore, G. R., Mann, S. and Bannister, J. V. (1986) J. Inorg. Chem. **28**, 329–336
Que, Jr., L. and True, A. E. (1990) Prog. Inorg. Chem. **38**, 97–200
Ramachandran, G. N. and Sasisekharan, V. (1968) Adv. Protein Chem. **23**, 283–437
Smith, J. M. A., Ford, G. C., Harrison, P. M., Yariv, J. and Kalb, A. J. (1989) J. Mol. Biol. **205**, 465–469
Stiefel, E. I. and Watt, G. D. (1979) Nature (London) **279**, 81–83
Treffry, A., Hirzmann, J., Yewdall, S. J. and Harrison, P. M. (1992) FEBS Lett. **302**, 108–112
Watt, R. J., Frankel, R. B. and Watt, R. B. (1992) Biochemistry **31**, 9673–9679
Yariv, J., Kalb, A. J., Sperling, R., Bauminger, E. R., Cohen, S. G. and Ofer, S. (1981) Biochem. J. **197**, 171–175

Received 5 October 1992/12 November 1992; accepted 24 November 1992