crystallization communications

Preliminary X-ray diffraction analysis of YcdB from *Escherichia coli*: a novel haem-containing and Tat-secreted periplasmic protein with a potential role in iron transport

Michaël L. Cartron,[a] **Sue A. Mitchell,**[b] **Mark R. Woodhall,**[a]‡ **Simon C. Andrews**[a] **and Kimberly A. Watson**[b]*

[a]School of Biological Sciences, University of Reading, England, and [b]Structural Biology Unit, The BioCentre, University of Reading, England

‡ Current address: Institute of Molecular Medicine, University Department of Paediatrics, University of Oxford, John Radcliffe Hospital, Headington OX3 9DU, England.

Correspondence e-mail: k.a.watson@reading.ac.uk

YcdB is a periplasmic haem-containing protein from *Escherichia coli* that has a potential role in iron transport. It is currently the only reported haem-containing Tat-secreted substrate. Here, the overexpression, purification, crystallization and structure determination at 2.0 Å resolution are reported for the apo form of the protein. The apo-YcdB structure resembles those of members of the haem-dependent peroxidase family and thus confirms that YcdB is also a member of this family. Haem-soaking experiments with preformed apo-YcdB crystals have been optimized to successfully generate haem-containing YcdB crystals that diffract to 2.9 Å. Completion of model building and structure refinement are under way.

1. Introduction

Although the vast majority of organisms are absolutely dependent on iron, acquisition of sufficient quantities of this metal is often challenging owing to its poor solubility and potential toxicity (Andrews *et al.*, 2003). Competition for iron resources is particularly acute for pathogenic bacteria, which must overcome the iron-withdrawal strategies of their hosts in order to satisfy their iron requirements (Litwin & Calderwood, 1993). Bacteria generally overcome iron restriction by utilizing a number of alternative iron transporters that enable them to scavenge iron from various sources. The most common type of iron-uptake strategy employed by bacteria involves the secretion and/or utilization of high-affinity iron chelators, known as siderophores, which bind and solubilize ferric iron. In Gram-negative bacteria, ferri-siderophore complexes are taken up *via* an outer-membrane (OM) receptor and binding-protein-dependent ABC inner-membrane system requiring the energy-transducing TonB–ExbBD complex for translocation across the OM. Related systems are used by pathogens to directly obtain iron from host iron-containing proteins. Often, bacteria contain multiple ferric complex-uptake systems enabling them to piratize siderophores produced by other organisms. Bacteria can also acquire the soluble ferrous form of iron *via* the FeoAB pathway, NRAMP-like transporters (*e.g.* MntH) or through metal-type ABC transporters (*e.g.* SitABCD) (Andrews *et al.*, 2003).

Recent evidence suggests that bacteria possess another major route for iron uptake that resembles the high-affinity iron-transport system of yeast and other fungi. In yeast, this pathway consists of three plasma-membrane components: Frt1p, a polytopic ferrous permease, Fet3p, an Frtp1-associated multicopper oxidase that oxidizes ferrous iron to the ferric form during the Ftr1p-mediated uptake process, and Fre1p, one of several related flavin- and haem-containing NADPH-dependent ferric reductases that reduces extracellular ferric iron to the ferrous form prior to transport *via* Ftr1p (Kosman, 2003). In bacteria, genes encoding homologues of Ftr1p are common, being found in proteobacteria, actinobacteria and firmicutes, and are often adjacent to other iron-related genes (Woodhall, 2005; Debut *et al.*, 2006). In *Escherichia coli*, the Ftr1p homologue (YcdN) is part of an iron and Fur (ferrous uptake regulator) repressed operon, *ycdNOB* (McHugh *et al.*, 2003). In *Bacillus subtilis* and *Neisseria meningitides*, the *ycdNOB*-like operons, *ywbLMN* and NMB0034-36, respectively, are likewise iron-repressed (Baichoo *et*

crystallization communications

al., 2002; Grifantini et al., 2003). This iron control, together with the Ftr1p homology, is highly suggestive of an iron-transport role for these YcdNOB-like systems. This supposition has been recently confirmed in *B. subtilis* by the finding that the YwbLMN system is involved in the uptake of free ferric iron (Ollinger et al., 2006). Recent studies on YcdNOB from *E. coli* strain Nissle 1917 suggest a role in ferrous iron transport (Grosse et al., 2006). Interestingly, there is evidence that the *ycdNOB* operon of *E. coli* is induced by low pH (Stancik et al., 2002), conditions which favour ferric iron solubility and ferrous iron stability against oxidation.

The specific roles of the *ycdNOB*-encoded YcdO and YcdB proteins in YcdN-mediated iron uptake are unclear. YcdO is unlike any other characterized protein in the sequence databases, but the corresponding gene is found associated with *ycdN* and *ycdB* homologues in all cases to date. YcdO possesses an N-terminal signal sequence, suggesting that it is exported to the periplasm in a Sec-dependent manner. YcdB is also periplasmic, but unlike YcdO it contains *b*-type haem and is secreted *via* the twin-arginine transport (Tat) pathway. Currently, it is the only known haem-containing Tat substrate (Sturm et al., 2006). The *B. subtilis* YcdB homologue (YwbN) has also been shown to be exported *via* the Tat system (Jongbloed et al., 2004).

Significantly, YcdB exhibits homology to the dye-decolorizing peroxidase, DyP, of *Thanatephorus cucumeris* (previously called *Geotrichum candidum*) Dec 1 (Kim & Shoda, 1999). The crystal structure of DyP has recently been solved, revealing the presence of a single haem *b* moiety within a cytochrome *c* peroxidase fold (Sato et al., 2004). A low-pH-enhanced peroxidase activity has also been reported for YcdB and, consistent with this, the *ycdN* homologue of *N. meningitides* is induced by H_2O_2-mediated redox stress (Grifantini et al., 2003, 2004; Sturm et al., 2006). However, it is unclear how these findings relate to the proposed iron-transport function of YcdNOB.

In an effort to understand the role of these proteins in iron transport and elucidate their mechanism of action, we have embarked on X-ray structural studies of this novel iron-transport system in *E. coli*. This work reports the overexpression, purification, crystallization and preliminary X-ray crystallographic analysis of the apo form (haem-free) of YcdB of *E. coli* K-12 and recent progress towards solving the structure of the haem-containing form.

2. Materials and methods

### 2.1. Cloning, expression and purification

The *ycdB* gene was PCR-amplified from *E. coli* K-12 genomic DNA using Accuzyme (Bioline) and cloned into the overexpression vector pET-46 Ek/LIC (Novagen) generating pET46-His$_6$*ycdB*. The cloning was performed such that the N-terminal Tat signal sequence (residues 1–35) of YcdB was replaced by the His$_6$ tag and entero-kinase cleavage sequence (MAH$_6$VD$_4$KI) encoded by the overexpression vector. Overexpression of '(His)$_6$YcdB' (43 kDa subunit weight) was achieved using BL21/λDE3(pET46-His$_6$*ycdB*) cultured in L-broth with ampicillin (100 µg ml^{-1}) at 250 rev min^{-1} and 310 K and induced with 1 mM isopropyl β-D-thio-galactopyranoside (IPTG) when the culture optical density at 650 nm reached 0.5. The IPTG-induced cells were cultured for a further 16 h, harvested by centrifugation for 20 min at 3000g and 277 K and resuspended in 3 ml (per gram cell weight) buffer *A* (25 mM HEPES pH 7.4, 50 mM imidazole, 0.2 M NaCl) before being lysed at 137 MPa using a French press. (His)$_6$YcdB was then purified from the soluble supernatant by Ni^{2+}-affinity chromatography (HiTrap Affinity resin, Pharmacia) using buffer *A* as the binding buffer and buffer *B* (buffer *A* plus 0.5 M

imidazole) for elution (Fig. 1*a*). Two major (His)$_6$YcdB elution peaks were obtained: the first corresponded to cofactor-containing (His)$_6$YcdB, with the second corresponding to the apo form (Fig. 1*b*). The resulting protein was >95% pure according to SDS–PAGE analysis and was dialysed against 25 mM HEPES pH 7.4 prior to storage at 193 K.

2.2. Crystallization

Purified apo-(His)$_6$YcdB was concentrated to 10 mg ml^{-1} in 25 mM HEPES buffer pH 7.4 using centrifugal membrane concentrators (Vivaspin). The final concentration was determined using a Bradford assay. Initial crystallization screening was performed manually using the sitting-drop vapour-diffusion method in 24-well Linbro plates against the following commercial screens at 291 K: Structure Screens I and II, Macrosol I and II, Stura Footprint Screens I and II (all from Molecular Dimensions Ltd) and Wizard I and II (from Emerald BioSystems). The drop size was 2 µl plus 2 µl in all cases.

From the 384 conditions screened, seven hits were found that indicated two types of crystal morphology: feather-like needles (Fig. 2*a*) and small bipyramidal crystals (Fig. 2*b*). Crystals typically
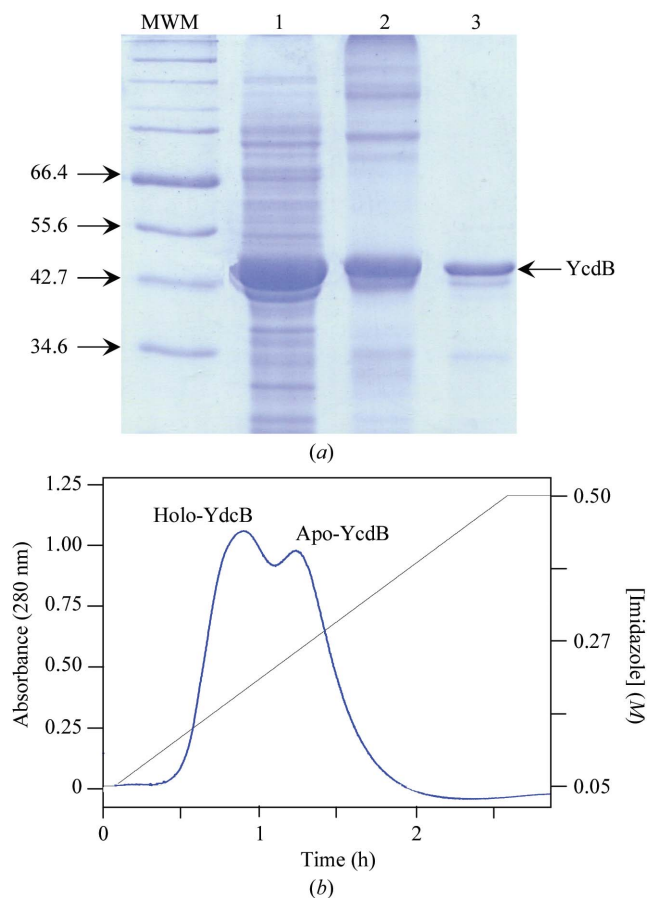
Figure 1
Purification of apo-(His)$_6$YcdB. (*a*) SDS–PAGE (12% acrylamide) of samples at different stages of apo-(His)$_6$YcdB purification from the (His)$_6$YcdB over-expressing strain. MWM, molecular-weight markers (kDa); lane 1, soluble crude extract following cell lysis and centrifugation; lane 2, apo-(His)$_6$YcdB after Ni^{2+}-affinity chromatography; lane 3, apo-(His)$_6$YcdB following anion-exchange chromatography. Approximately 50, 20 and 5 µg of protein were loaded in lanes 1, 2 and 3, respectively. (*b*) Elution profile of overexpressed (His)$_6$YcdB during Ni^{2+}-affinity chromatography. The concentration of imidazole applied to the column is indicated, as are the elution peaks corresponding to the apo and holo forms of (His)$_6$YcdB.

Table 1
Data-collection and processing statistics for apo-(His)$_6$YcdB.

Values in parentheses are for the highest resolution shell.

Synchrotron beamline	ESRF ID14-EH2
Wavelength (Å)	0.933
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 83.31$, $b = 86.26$, $c = 119.58$
Resolution range (Å)	50.0–2.0 (2.11–2.0)
R_{merge}†	7.8 (24.1)
No. of observations	627643 (85466)
No. of unique reflections	58729 (8360)
Mean $I/\sigma(I)$‡	26.5 (8.0)
Completeness (%)	99.7 (98.8)
Multiplicity	10.7 (10.2)
Solvent content (%)	41
Molecules per ASU	2

† $R_{merge} = \sum_h \sum_i |I_{h,i} - \langle I_h \rangle| / \sum_h \sum_i |I_{h,i}|$, where the outer summation is over all unique reflections with multiple observations and the inner summation is over all observations of each reflection. ‡ $\sigma(I)$ is the standard deviation of I.

appeared between 14 and 28 d. The most promising lead was Macrosol II No. 14 [1.5 M ammonium sulfate, 2%(v/v) PEG 400, 0.1 M HEPES pH 7.5], which gave single bipyramidal crystals with typical dimensions of $0.1 \times 0.1 \times 0.15$ mm.

Optimization, particularly of the pH, resulted in large single crystals of typical dimensions of $0.2 \times 0.2 \times 0.3$ mm (Fig. 2c). The final conditions for reproducible apo-(His)$_6$YcdB crystal growth were 1.8–2.0 M ammonium sulfate, 2–8%(v/v) PEG 400, 0.1 M HEPES pH 7.0 at 291 K. These crystals were of suitable size and quality for X-ray diffraction analysis. The optimized conditions yielded crystals within 14 d.

Initial attempts to cocrystallize apo-(His)$_6$YcdB in the presence of haem (stock solution of 100 mM bovine haemin, from Sigma–Aldrich, in a mixture of 50 mM arginine and 50 mM glutamine) were unsuccessful. However, the addition of small amounts of solid haemin directly to the hanging drops containing preformed colourless crystals

of apo-(His)$_6$YcdB gave rise to red-coloured crystals after 1–2 d (Fig. 2d).

### 2.3. Data collection and diffraction analysis

Prior to data collection, the apo-(His$_6$)YcdB crystals were cryo-protected in 25% PEG 550 MME by quick transfer directly from the hanging drop. For the haem-soaked apo-(His)$_6$YcdB crystals, optimization of the cryoconditions was critically important in order to eliminate major problems arising from phase separation, rapid crystal decay and icing. The final cryoconditions necessary for successful freezing and data collection of haem-soaked apo-(His)$_6$YcdB crystals were a mixture of 0.75 M ammonium sulfate, 5 mM NaCl, 0.5 M sodium formate, 6.4% PEG 400 and 2.5% glycerol in 0.1 M HEPES pH 7.0.

Intensity data were collected on an ADSC Quantum 4 CCD detector at 100 K on the macromolecular beamline Station ID14-EH2 (ESRF, Grenoble, France). Fig. 3(a) shows a typical diffraction image for apo-(His)$_6$YcdB taken at ESRF. Data were collected to 2.0 Å resolution. Data were integrated and scaled using the programs *MOSFLM* v.6.2.4 (Leslie, 1992) and *SCALA* (Evans, 1997), respectively, from the *CCP*4 program package (Collaborative Computational Project, Number 4, 1994). The apo-(His)$_6$YcdB crystals exhibited orthorhombic symmetry, space group $P2_12_12_1$, with unit-cell parameters $a = 83.31$, $b = 86.26$, $c = 119.58$ Å, and two molecules in the asymmetric unit. Data-processing statistics for apo-(His)$_6$YcdB are shown in Table 1.

Fig. 3(b) shows a diffraction image taken on the in-house source (FR-D Rigaku/MSC, Cu $K\alpha$ radiation) of one of the best diffracting crystals of haem-soaked apo-(His)$_6$YcdB. Diffraction was observed to 2.9 Å resolution. These crystals were also orthorhombic, space group $P2_12_12_1$, and showed minor changes in the unit-cell parameters: $a = 84.88$, $b = 87.76$, $c = 122.99$ Å. Haem-soaked apo-(His)$_6$YcdB

(a) (b) (c) (d)
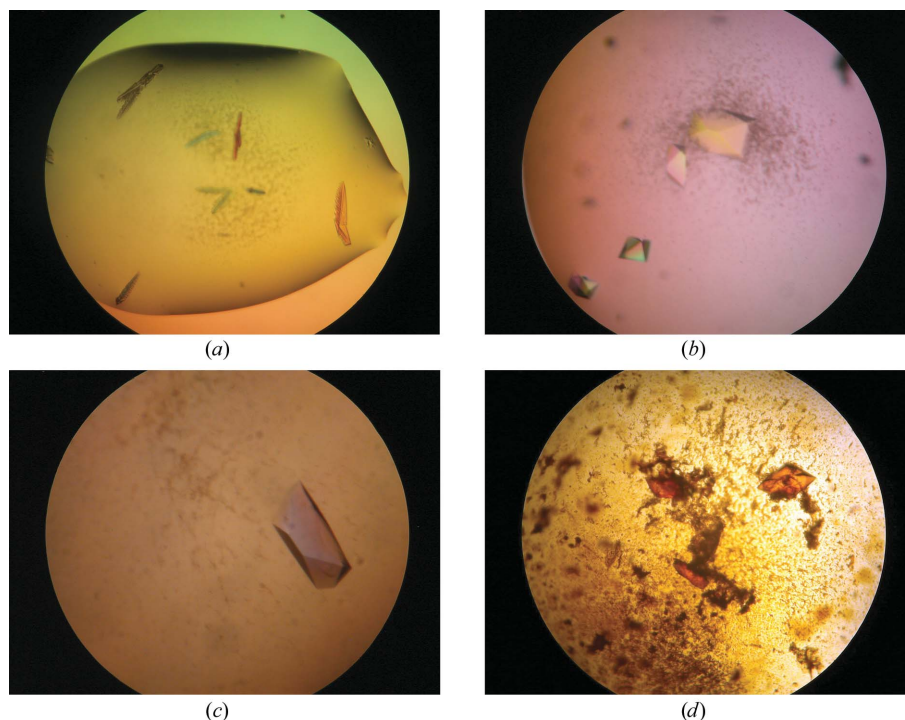
Figure 2
Crystallization of apo-(His)$_6$YcdB from *E. coli*. (*a*) Feather-like needles observed from initial screening, Macrosol I No. 22; (*b*) bipyramidal crystals observed from initial screening, Macrosol II No. 14; (*c*) optimized conditions from (*b*), 2 M ammonium sulfate, 5% PEG 400, 0.1 M HEPES pH 7.0; (*d*) solid haem soaked into preformed apo-(His)$_6$YcdB optimized crystals. The drop size was 2 µl plus 2 µl and the protein concentration was 10 mg ml^{-1} in all cases.
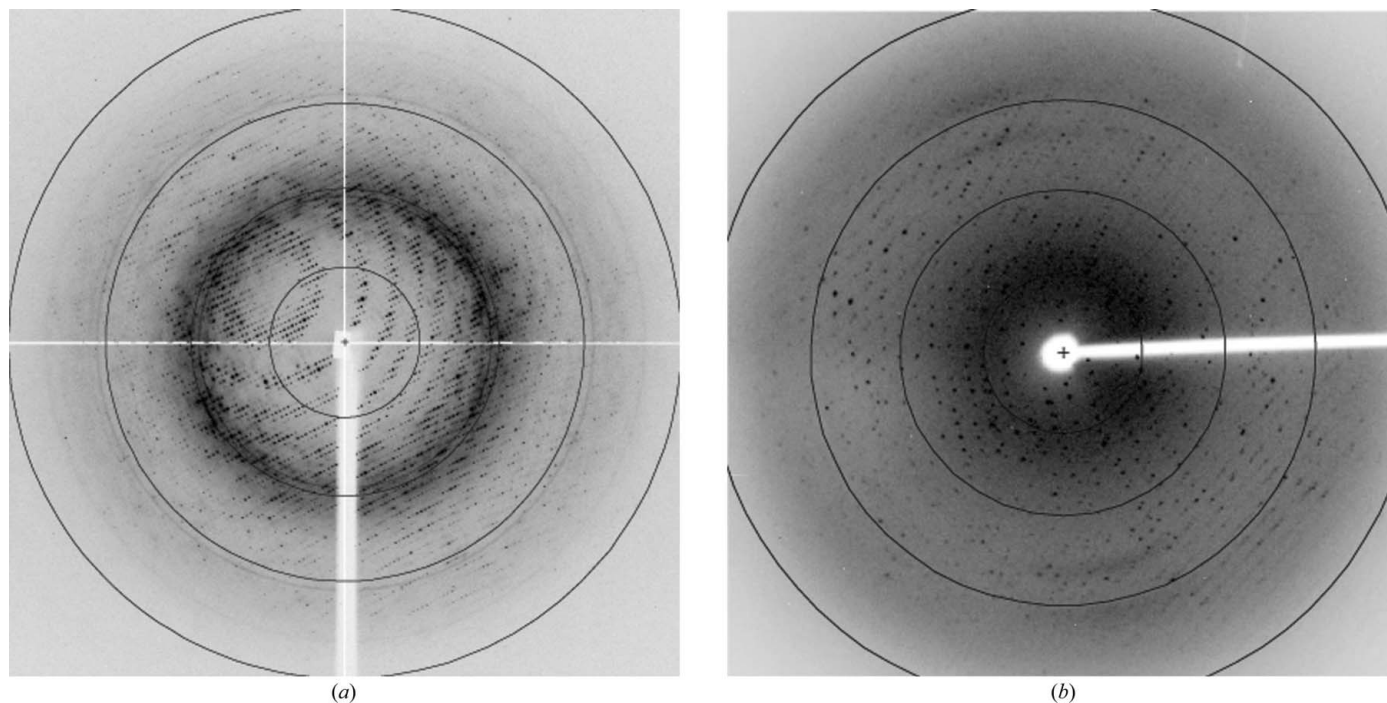
Figure 3
(*a*) Typical diffraction image from apo-(His)$_6$YcdB crystals obtained at ESRF ID14-EH2. (*b*) The best diffraction image from haem-soaked apo-(His)$_6$YcdB crystals obtained using the in-house X-ray source (Rigaku/MSC FR-D, Cu $K\alpha$ radiation).

crystals exhibited high mosaicity (>1°) and anisotropy and showed signs of crystal splitting.

2.4. Phasing

The protein for which a structure had been reported that shared the highest amino-acid sequence homology (25% over 183 residues, determined using *WU-Blast*2) was found to be DyP, a dye-colourizing peroxidase from *T. cucumeris*. No coordinates were available for DyP (Sato *et al.*, 2004); however, two distant fungal homologues to DyP were present in the Protein Data Bank: lignin peroxidase (LiP) and manganese peroxidase (MnP). The two fungal peroxidases shared less than 10% overall sequence identity with YcdB. The program *CHAINSAW* (Schwarzenbacher *et al.*, 2004) was used to align the target sequence (YcdB) and the two model sequences from LiP (PDB code 1lga) and MnP (PDB code 1mnp) such that all the atoms common to the three models were retained. The resulting model, consisting of residues 34–291 (of a total of 423 residues) of the (His)$_6$YcdB sequence, was used as the search model for molecular replacement against all data between 50 and 3.0 Å using *MOLREP* (Vagin & Teplyakov, 1997). The program *DM* (Cowtan & Main, 1996) was used for noncrystallographic symmetry (NCS) averaging and solvent flattening of the electron-density map, which clearly showed distinguishable secondary-structural features (Fig. 4). Model building and refinement are in progress using the programs *Coot* (Emsley & Cowtan, 2004) and *REFMAC* (Murshudov *et al.*, 1997), respectively.

3. Conclusion

The gene encoding YcdB from *E. coli* was cloned from genomic DNA and overexpressed using a T7 promoter-based vector and the apo form of (His)$_6$YcdB was purified by Ni^{2+}-affinity chromatography. Crystals of apo-(His)$_6$YcdB were produced and conditions were found for the generation of a potential haem-containing derivative

using preformed apo-(His)$_6$YcdB crystals. Diffraction-quality crystals were obtained in each case through careful optimization of pH and cryoprotecting conditions, which gave reproducible crystals that
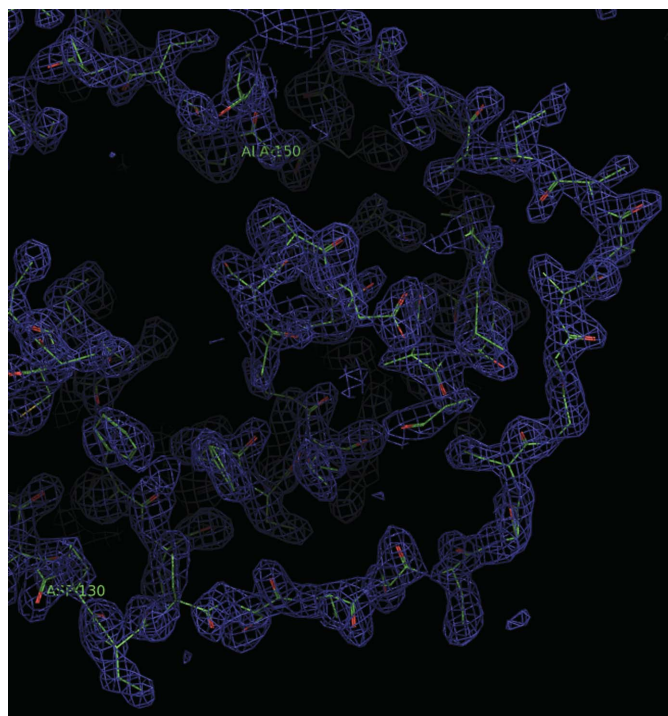
Figure 4
Electron-density map of apo-(His)$_6$YcdB at 2.0 Å resolution contoured at 1σ in the region of residues 130–150, a well conserved surface loop. Model building and refinement are in progress. The initial model covers the region between residues 34 and 291 (of 423 residues).

diffracted to 2.0 Å for apo-(His)$_6$YcdB and more typically to 2.9 Å for the haem-soaked derivative. Much of the success in obtaining suitable diffraction from haem-soaked apo-(His)$_6$YcdB was achieved by careful optimization of the cryoprotectant (0.75 M ammonium sulfate, 5 mM NaCl, 0.5 M sodium formate, 6.4% PEG 400 and 2.5% glycerol in 0.1 M HEPES pH 7.0) in order to eliminate serious problems arising from phase separation and rapid crystal decay. However, numerous haem-soaked apo-(His)$_6$YcdB crystals were tested and many exhibited high mosaicity ($>1°$) and showed signs of anisotropy and crystal splitting, making integration and scaling difficult. Efforts are under way to crystallize holo-(His)$_6$YcdB, since a crucial aspect of this work is to determine the role of this protein in iron transport and as such the location and the surrounding environment of the haem are critically important.

With regard to the haem environment, it has been shown that although the proximal histidine exhibits structural similarity, there are noted differences in the haem environment between the related peroxidases currently known, *i.e.* between DyP and both LiP and MnP. DyP has been reported to contain a four-stranded antiparallel β-sheet distal to the haem ligand, a structural feature which is not present in either LiP or MnP (Sato *et al.*, 2004). The nature of this structural difference is presently unknown. Completion of model building and refinement will reveal the extent of the structural similarity between YcdB from *E. coli* and other structurally defined members of the superfamily of haem-dependent peroxidases.

References

Andrews, S. C., Robinson, A. K. & Rodríguez-Quiñones, F. (2003). *FEMS Microbiol. Rev.* **27**, 215–237.

Baichoo, N., Wang, T., Ye, R. & Helmann, J. D. (2002). *Mol. Microbiol.* **45**, 1613–1629.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Cowtan, K. D. & Main, P. (1996). *Acta Cryst.* D**52**, 43–48.

Debut, A. J., Dumay, Q. C., Barabote, R. D. & Saier, M. H. Jr (2006). *J. Mol. Microbiol. Biotechnol.* **11**, 1–9.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Evans, P. (1997). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **33**, 22–24.

Grifantini, R., Frigimelica, E., Delany, I., Bartolini, E., Giovinazzi, S., Balloni, S., Agarwal, S., Galli, G., Genco, C. & Grandi, G. (2004). *Mol. Microbiol.* **54**, 962–979.

Grifantini, R., Sebastian, S., Frigimelica, E., Draghi, M., Bartolini, E., Muzzi, A., Rappuoli, R., Grandi, G. & Genco, C. A. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 9542–9547.

Grosse, C., Scherer, J., Koch, D., Otto, M., Taudte, N. & Grass, G. (2006). *Mol. Microbiol.* **62**, 120–131.

Jongbloed, J. D., Grieger, U., Antelmann, H., Hecker, M., Nijland, R., Bron, S. & van Dijl, J. M. (2004). *Mol. Microbiol.* **54**, 1319–1325.

Kim, S. J. & Shoda, M. (1999). *Appl. Environ. Microbiol.* **65**, 1029–1035.

Kosman, D. J. (2003). *Mol. Microbiol.* **47**, 1185–1197.

Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.

Litwin, C. M. & Calderwood, S. B. (1993). *Clin. Microbiol. Rev.* **6**, 137–149.

McHugh, J. P., Rodriguez-Quinones, F., Abdul-Tehrani, H., Svistunenko, D. A., Poole, R. K., Cooper, C. E. & Andrews, S. C. (2003). *J. Biol. Chem.* **278**, 29478–29486.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Ollinger, J., Song, K.-B., Antelmann, H., Hecker, M. & Helmann, J. D. (2006). *J. Bacteriol.* **188**, 3664–3673.

Sato, T., Hara, S., Matsui, T., Sazaki, G., Saijo, S., Ganbe, T., Tanaka, N., Sugano, Y. & Shoda, M. (2004). *Acta Cryst.* D**60**, 149–152.

Schwarzenbacher, R., Godzik, A., Grzechnik, S. K. & Jaroszewski, L. (2004). *Acta Cryst.* D**60**, 1229–1236.

Stancik, L. M., Stancik, D. M., Schmidt, B., Barnhart, D. M., Yoncheva, Y. N. & Slonczewski, J. L. (2002). *J. Bacteriol.* **184**, 4246–4258.

Sturm, A., Schierhorn, A., Lindenstrauss, U., Lilie, H. & Bruser, T. (2006). *J. Biol. Chem.* **281**, 13972–13978.

Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.

Woodhall, M. R. (2005). PhD thesis. University of Reading, England.