Preliminary X-ray diffraction analysis of YqjH from Escherichia coli: a putative cytoplasmic ferri-siderophore reductase


It is advisable to refer to the publisher's version if you intend to cite from the work.

To link to this article DOI: http://dx.doi.org/10.1107/S174430910802352X

Publisher: International Union of Crystallography

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.
www.reading.ac.uk/centaur

CentAUR
Central Archive at the University of Reading
Reading’s research outputs online
Preliminary X-ray diffraction analysis of YqjH from 
*Escherichia coli*: a putative cytoplasmic ferri-siderophore reductase

YqjH is a cytoplasmic FAD-containing protein from *Escherichia coli*; based on homology to ViuB of *Vibrio cholerae*, it potentially acts as a ferri-siderophore reductase. This work describes its overexpression, purification, crystallization and structure solution at 3.0 Å resolution. YqjH shares high sequence similarity with a number of known siderophore-interacting proteins and its structure was solved by molecular replacement using the siderophore-interacting protein from *Shewanella putrefaciens* as the search model. The YqjH structure resembles those of other members of the NAD(P)H:flavin oxidoreductase superfamily.

1. Introduction

Iron is required by virtually all living organisms as a micronutrient. The ability of iron to easily switch between two relatively stable redox states (ferric/Fe$^{3+}$ and ferrous/Fe$^{2+}$ forms) makes it ideally suited as a mediator of redox chemistry within biology. However, this redox activity can create problems for organisms exposed to oxygen since iron is capable of catalysing the generation of reactive oxygen species that can cause toxicity and ultimately cell death (Dubrac & Touati, 2000; Touati, 2000). Furthermore, in the presence of oxygen at neutral (or higher) pH, the relatively soluble ferrous form (solubility of 0.1 M at pH 7.0) is rapidly converted to the poorly soluble ferric form (solubility of $10^{-18}$ M at pH 7.0), leading to low iron bioavailability (Andrews *et al.*, 2003). Thus, lack of iron is a common form of malnutrition in nature and most organisms utilize specific iron-acquisition systems to enable them to scavenge it from their environment to counter iron restriction.

In bacteria, the major mechanism employed to overcome the problem of ferric iron insolubility involves the production and secretion of siderophores, which are molecules with a high affinity and specificity for ferric iron. Siderophores are released into the environment, where they chelate and thus solubilize ferric iron. The resulting ferri-siderophore complexes can then be imported into the cell through specific transport systems. The complex is thought to be dissociated in the cytosol, releasing the bound iron, which can then be incorporated into iron-requiring proteins. The proposed mechanism for the release of iron from siderophores involves either degradation of the siderophore or reduction of the bound iron from the ferric form to the ferrous state, resulting in loss of affinity of the iron for the chelator. In *Escherichia coli* K-12, enterobactin (a catecholate), a cyclic trimer of dihydroxybenzoylserine (DHBS), is the sole siderophore produced. Following import of the highly stable ferri-enterobactin complex, the iron is dissociated through the activity of the Fes (ferric enterobactin esterase) protein. Fes acts by converting enterobactin into monomeric DHBS units, generating a ferri-DHBS complex that is more readily dissociated (Raymond *et al.*, 2003). Fes is also reported to possess a distinct reductase activity, although it does not appear to contain a redox centre that could mediate electron transfer (Brickman & McIntosh, 1992; Hollifield & Neilands, 1978).
Two Fes homologues (IroD and IroE) have recently been shown to be involved in the breakdown of a glycosylated form of enterobactin (namely salmochelin) in Salmonella enterica (Lin et al., 2005). E. coli also utilizes FhuF, a cytosolic Fe–S protein that has been reported to be required for the release of iron from ferri-terricrin (an exogenous hydroxamate-type siderophore) complexes via a reduction mechanism (Matzanne et al., 2004). Appropriately, the genes encoding Fes and FhuF are both induced by iron starvation. Thus, E. coli appears to employ both a degradation and reduction mechanism to achieve the release of iron from siderophores, although the two mechanisms utilize distinct components and appear to target very different siderophores.

Interestingly, E. coli possesses a third potential ferri-siderophore dissociation protein designated YqjH. The corresponding gene (yqjH) is iron-repressed in a Fur-dependent fashion (McHugh et al., 2003) and the protein is homologous to ViuB of Vibrio cholerae. ViuB is apparently cytosolic (like YqjH) and is required for the utilization of ferri-vibriobactin, a catechol siderophore. In addition, viuB can complement an fes E. coli mutant, strongly suggesting an iron-release role (Butterton & Calderwood, 1994). However, the manner in which ViuB functions remains unclear since the protein has not yet been characterized. Many other YqjH homologues (e.g. BauF and MxcB) are encoded by genes in loci specifying siderophore biosynthesis and transport components (Mihara et al., 2004; Silakowski et al., 2000), which is consistent with a role in iron release.

Here, we report the overproduction, purification, crystallization and preliminary X-ray, UV–visible and mass-spectrometric analysis of YqjH from E. coli. The three-dimensional structure of YqjH will provide valuable insights into the mechanism of iron release from siderophores in bacteria.

2. Materials and methods

2.1. Cloning, expression and purification

The E. coli K-12 yqjH gene was PCR-amplified from genomic DNA using Accuzyme (Bioline) and cloned into the expression vector pET101/D-TOPO (Promega) to generate pET101yqjH-C(His)6. Overexpression of the resultant YqjH-V5 epitope-His6 fusion protein (YqjH-His6, 32.4 kDa, containing 32 additional amino acids at the C-terminus) was performed using E. coli BL21 (DE3) cultured in L broth containing ampicillin (100 µg ml−1) at 298 K and 250 rev min−1. At a culture optical density at 650 nm of 0.5, the cells were induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and cultured for a further 8 h before harvesting by centrifugation for 20 min at 3000g and 277 K.

Cell pellets were resuspended in buffer A (25 mM HEPES pH 7.4, 50 mM imidazole, 0.2 M NaCl) before being lysed at 138 MPa using a French press. YqjH-His6 was then purified from the supernatant by Ni2+-affinity chromatography (NTA-Sepharose, Qiagen) using buffer A as the binding buffer and eluted with a linear gradient of buffer B (buffer A containing 0.5 M imidazole). The protein eluted as a single peak at approximately 200 mM imidazole. The resulting protein was dialysed against buffer C (25 mM HEPES pH 7.4, 20 mM mannitol, 1 mM EDTA, 1 mM DTT) and further purified by anion-exchange chromatography (DEAE Sepharose FF, Pharmacia). YqjH-His6 was eluted with a linear gradient of buffer D (buffer A containing 0.5 M NaCl) and was released from the column with approximately 200 mM NaCl. The purified protein was dialysed against buffer C and samples were stored in buffer C at 253 K. The final protein concentrations used for each subsequent experiment were determined by the method of Bradford (1976) using bovine serum albumin as standard.

2.2. UV–visible spectroscopy

The UV–visible spectra were recorded with the as-isolated YqjH-His6 protein at 1.33 mg ml−1 (43.6 µM) in buffer containing 25 mM HEPES pH 7.4 and 0.2 M NaCl using a Unicam Helios α spectrophotometer equipped with a 1 cm path-length quartz cuvette. The scanning rate was 120 nm min−1 and the bandwidth was 2 nm.

2.3. Mass spectrometry

Mass-spectrometric analysis was performed using an electrospray ionization (ESI) mass spectrometer to determine the molecular weight of the protein and associated cofactor. A sample of YqjH-His6 (45 µl at a concentration of 0.6 mg ml−1 protein) in pure water was diluted with 90 µl of a methanol/water (50:50) solution containing 0.1% formic acid. This mixture was then directly used for infusion into the electrospray source of a Bruker MicroTOF high-performance mass spectrometer. Similar solutions of two control proteins, cytochrome c (12 361 Da) and myoglobin (16 952 Da), were also applied at 0.1 and 1 µM (prepared as above) to confirm instrument performance and for calibration. For each sample analysed, a volume of 50 µl was infused at 2 µl min−1 using a 100 µl syringe in a Harvard syringe pump. Data were acquired over an mz range of 100–5000.
Table 1
Data-collection and processing statistics for YqjH.

<table>
<thead>
<tr>
<th></th>
<th>Data set 1</th>
<th>Data set 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synchrotron beamline</td>
<td>ESRF ID23-EH1</td>
<td>ESRF ID23-EH1</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.976</td>
<td>0.976</td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁2₁2₁</td>
<td>P2₁2₁2₁</td>
</tr>
<tr>
<td>Unit-cell parameters (Å)</td>
<td>a = 75.23, c = 84.43</td>
<td>a = 80.72, c = 84.59</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>85.0-3.0 (3.16-3.0)</td>
<td>85.0-3.0 (3.16-3.0)</td>
</tr>
<tr>
<td>Rmerge</td>
<td>11.7 (35.5)</td>
<td>10.2 (41.3)</td>
</tr>
<tr>
<td>No. of observations</td>
<td>19303 (2849)</td>
<td>23119 (3427)</td>
</tr>
<tr>
<td>No. of unique reflections</td>
<td>5825 (829)</td>
<td>6696 (956)</td>
</tr>
<tr>
<td>Mean</td>
<td>10.4 (3.1)</td>
<td>10.6 (3.1)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.8 (99.9)</td>
<td>99.9 (100)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>3.3 (3.4)</td>
<td>3.5 (3.6)</td>
</tr>
<tr>
<td>Solvent content (%)</td>
<td>42</td>
<td>49.5</td>
</tr>
<tr>
<td>Molecules per ASU</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

* Rmerge = \sum \sum |I(hk)| - |\langle I(hk)\rangle|/\sum \sum |I(hk)|, where the outer summation is over all unique reflections with multiple observations and the inner summation is over all observations of each reflection.

2.4. Crystallization

Purified YqjH-His<sub>S</sub> was concentrated in buffer C (25 mM HEPES pH 7.4, 20 mM mannitol, 1 mM EDTA, 1 mM DTT) to 8 mg ml<sup>−1</sup> using centrifugal membrane concentrators (Vivaspin; 5 kDa molecular-weight cutoff). Initial crystallization screening was performed manually in 2 + 2 µl drops using the hanging-drop vapour-diffusion method in 24-well Linbro plates against the PACT screen (Molecular Dimensions Ltd) at 291 K.

Initial conditions were found that gave crystal morphologies of either needles or flattened rods (Fig. 1a). The most promising condition [No. 1-1; 0.1 M SPG buffer pH 4.0 and 25% (w/v) PEG 1500] was further optimized. Optimization of the PEG concentration resulted in single rod-shaped crystals of typical dimensions 300 × 20 × 20 µm that grew optimally from 0.1 M SPG buffer pH 4.0 and 18% (w/v) PEG 1500 (Fig. 1b). Crystals from the optimized conditions appeared in less than 24 h and grew to maximum dimensions within 3 d. The crystal quality was observed to degrade over a period of 6–14 d. The optimized crystals, harvested within one week of appearance, were of suitable size and quality for X-ray diffraction analysis.

2.5. Data collection and diffraction analysis

Typically, upon appearance YqjH-His<sub>S</sub> crystals were cryoprotected in mother liquor containing either 25% (w/v) PEG 550 MME or 25% (w/v) ethylene glycol by quick transfer directly from the hanging drop. Diffraction data were collected on an ADSC Q315 CCD detector at 100 K on the macromolecular crystallography beamline station ID23-EH1 (ESRF, Grenoble, France). Data were collected to a resolution of 3.0 Å and were integrated and scaled using the programs MOSFLM v6.2.4 (Leslie, 1992) and SCALA (Evans, 1997), respectively, from the CCP4 program package (Collaborative Computational Project, Number 4, 1994). The YqjH-His<sub>S</sub> crystals exhibited trigonal symmetry, space group P2₁2₁2₁, and had unit-cell parameters a = 75.23, c = 84.43 Å or a = 80.72, c = 84.59 Å when using PEG MME 550 (data set 1, Table 1) or ethylene glycol (data set 2, Table 1) as the cryoprotectant, respectively. Solvent-content analysis using the program MATTHEWS_COEF (Kantardjieff & Rupp, 2003) suggested solvent contents of 42 and 49.5% for the two unit cells, respectively, for one molecule per asymmetric unit in each case. Data-processing statistics for YqjH-His<sub>S</sub>, are shown in Table 1.

2.6. Phasing

Sequence-orientation searches of the PDB using WU-BLAST2 (http://www.ebi.ac.uk/blast2) revealed the protein that shares the highest amino-acid sequence similarity (30% identity) to YqjH to be the siderophore-interacting protein from *Shewanella putrefaciens* CN-32 (PDB code 2gjj). Sequence alignment using ClustalW (http://www.ebi.ac.uk/clustalw) was used to align the target sequence (YqjH) and the model sequence. Inspection of the sequence alignment, removal of nonconserved loops and mutation of aligned amino acids to the target sequence resulted in a model consisting of residues 15–60, 79–89, 91–167, 171–254 and the FAD cofactor of the YqjH-His<sub>S</sub> sequence. Residue 254 is the C-terminal residue of the naturally produced YqjH. This model was used as a search model for molecular replacement using Phaser (McCoy et al., 2007). Using data collected from a crystal cryoprotected in PEG MME 550 (data set 1) and a crystal cryoprotected in ethylene glycol (data set 2), molecular-replacement solutions were found in space group P3₂12₁ with an LLG of 28 and 59, respectively. After rigid-body refinement in REFMAC (Murshudov et al., 1997), R factor and R<sub>free</sub> values of 46.5 and 46.9 were found for data set 1 and of 50.9 and 50.7 for data set 2, respectively. The resulting electron-density maps were clearly interpretable (Fig. 2). The map clearly showed that both the location and bound conformation of the FAD cofactor was in agreement with the search model. Model building and refinement are in progress (against data set 1) using the programs Coot (Emsley & Cowtan, 2004) and REFMAC, respectively. Crystals of a selenomethionine-derivatized protein have recently been obtained and structure determination for this protein is in progress for independent phase improvement.

3. Results and discussion

3.1. Isolation and spectral characterization

*yqjH* from *E. coli* was cloned from genomic DNA and over-expressed using a T7 promoter-based vector, generating a YqjH-V5 epitope-His<sub>S</sub> fusion protein (YqjH-His<sub>S</sub>). The protein was purified by Ni<sup>2+</sup>-affinity and anion-exchange chromatography and crystallized readily in SPG buffer and PEG 1500, giving crystals that were suitable for X-ray structure determination. The Pfam database indicates that YqjH consists of two domains: an N-terminal FAD-containing domain and a C-terminal ‘siderophore-interacting’

Figure 2
Electron-density map of YqjH in the vicinity of the FAD cofactor at 3.0 Å resolution and contoured at 1σ. The FAD cofactor adopts a bent conformation.
domain. Consistent with this prediction, the isolated YqiH-His<sub>e</sub> protein was found to be bright yellow.

Spectral analysis of YqiH-His<sub>e</sub>, confirmed the presence of an FAD-like species (Fig. 3). The UV–visible spectrum of the as-isolated protein was typical of that of proteins containing FAD in the oxidized form, with absorption maxima at 380 and 456 nm and distinct shoulders associated with the 456 nm peak located at 430 and 481 nm (Table 2). Although the oxidized spectrum is similar to that observed for other FAD-containing proteins, the shoulder at ~430 nm is indistinct in the spectra of FAD proteins in some cases. The 380:456:481 nm absorbance ratios for oxidized YqiH were 0.82:1:0.83. Similar ratios are observed for the equivalent spectral features of known FAD proteins presented in Table 2 (0.79–0.94:1:0.73–0.89).

Upon addition of reductant, the spectrum gradually changed, reaching a fully reduced state that was not further altered by additional aliquots of reductant. The peaks/shoulders at 430, 456 and 481 nm were either greatly lessened upon reduction or eliminated and an additional relatively minor peak appeared at ~530 nm (relative intensity of 0.18 compared with the oxidized spectrum at 456 nm) that was not apparent in the oxidized spectrum. The diminished absorbance of the features centred at 456 nm is a typical characteristic of FAD-containing proteins (Axley et al., 1997), as is the appearance of a novel feature in the 530–540 nm region upon reduction (Gadda & Fitzpatrick, 1998). Thus, the spectral characteristics of YqiH can be considered to be typical of FAD-containing proteins.

To confirm the identity of the protein isolated, as well as the presence of FAD, the protein was subjected to ESI mass-spectrometric analysis. The spectrum of YqiH-His<sub>e</sub>, gave two main charge-state envelopes as well as two low-mass singly charged ions. The first envelope had a range of +24 to +49, which upon deconvolution gave a peptide is 32 411.52 Da, which matches that corresponding to the first state envelopes as well as two low-mass singly charged ions. The first envelope (32 411 Da), thus confirming the identity of the isolated protein. The difference (786 Da) between the two major species corresponds to the molecular weight of the FAD cofactor (also 786 Da). Therefore, the measured molecular weights are likely to represent the protein in the FAD-bound and FAD-unbound states. In the low-mass region, a singly charged ion at m/z 786 is present which is likely to represent the FAD cofactor. The base peak at m/z 477 presumably represents the sodium salt of an FAD breakdown product comprising FMN. Thus, the MS data confirm the presence of the FAD cofactor.

### Table 2

Comparison of UV-visible spectral parameters for the oxidized forms of YqiH and FAD-containing proteins.

<table>
<thead>
<tr>
<th>YqiH</th>
<th>FNRI</th>
<th>NAO§</th>
<th>UNAGEP Red*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Wavelength (nm)</strong></td>
<td><strong>A</strong></td>
<td><strong>Wavelength (nm)</strong></td>
<td><strong>A</strong></td>
</tr>
<tr>
<td>430</td>
<td>6.31</td>
<td>446</td>
<td>12.25</td>
</tr>
<tr>
<td>481</td>
<td>6.31</td>
<td>593</td>
<td>9.50</td>
</tr>
</tbody>
</table>

† Data derived from Fig. 3. ‡ Ferredoxin-NADP<sup>†</sup> reductase from Anabaena PCC 7119 (Faro et al., 2002). § Nitroalkane oxidase from Fusarium oxysporum (Gadda & Fitzpatrick, 1998). ¶ UDP-N-acetylenopyruvylglucosamine reductase from E. coli (Axley et al., 1997).

### 3.2. Crystallization and structure determination

Crystallization experiments yielded crystals of needle or flattened rod-like morphology. Diffraction-quality crystals were obtained by optimization of the precipitant concentration. It was shown to be possible to cryoprotect these crystals under two different conditions, mother liquor plus 25% (w/v) PEG MME 550 or mother liquor plus 25% (w/v) ethylene glycol, both of which gave crystals that reproducibly diffracted to at least 3.0 Å resolution. However, the unit-cell parameters differed using the two cryoprotectant conditions. Whilst the c unit-cell parameter is essentially unaltered, there is a significant contraction of approximately 5 Å (or 6%) in the a and b unit-cell parameters for the crystals cryoprotected in PEG MME 550 with respect to those cryoprotected in ethylene glycol (Table 1). YqiH shares a high sequence similarity with over 200 proposed siderophore-interacting proteins (Pfam) and using either data set 1 or 2 (Table 1) it was possible to find a molecular-replacement solution using the putative siderophore-interacting protein from S. putrefaciens as the search model (30% identity, 51% similarity). The structure shows a two-domain protein and a search of the SCOP database revealed that the N-terminal domain shares structural similarity with the riboflavin synthase/ferrodoxin reductase FAD-binding domain family and that the C-terminal domain belongs to the ferredoxin reductase-like C-terminal NADP-linked domain family. The FAD-binding site, which is located between the two domains, shows three main features: stabilization of the negative charges of the diphasate group of the FAD via interactions with Arg86 and His106, aromatic stacking interactions between Tyr251, His106 and the adenine moiety of FAD and a second aromatic stacking arrangement between Tyr88, Tyr250 and the isoalloxazine ring. There also appear to be hydrogen-bond interactions between the hydroxyl O atoms of the ribose and the backbone atoms of residues His252 and Lys254 and between the carboxylate O atoms of the isoalloxazine and the backbone atoms of residues Thr89, Asp102 and Phe104. Details of the structure are shown in Fig. 4.
of these interactions will be discussed following structure refinement. Furthermore, once complete, this will enable structural comparisons with related proteins, which may reveal features that differentiate the various family members and help to shed light on the mechanism employed to release iron from siderophores.

We wish to thank Trevor Gibson (BioCentre Facility, University of Reading) for ESI–MS analysis. We also thank the BBSRC for provision of funding to SCA, the Reading Endowment Trust Fund for financial support to MA and the Lister Institute of Preventive Medicine (personal fellowship to KAW). The authors are grateful to the staff at ESRF, particularly the beamline scientists at ID23-EH1, for providing excellent data-collection facilities.

References