Biometals (2010) 23:1–17
DOI 10.1007/s10534-009-9262-z

EfeO-cupredoxins: major new members of the cupredoxin superfamily with roles in bacterial iron transport

Mohan B. Rajasekaran · Sanjay Nilapwar · Simon C. Andrews · Kimberly A. Watson

Received: 15 May 2009 / Accepted: 10 August 2009 / Published online: 23 August 2009

Abstract The EfeUOB system of *Escherichia coli* is a tripartite, low pH, ferrous iron transporter. It resembles the high-affinity iron transporter (Ftr1p-Fet3p) of yeast in that EfeU is homologous to Ftr1p, an integral-membrane iron-permease. However, EfeUOB lacks an equivalent of the Fet3p component—the multicopper oxidase with three cupredoxin-like domains. EfeO and EfeB are periplasmic but their precise roles are unclear. EfeO consists primarily of a C-terminal peptidase-M75 domain with a conserved 'HxxE' motif potentially involved in metal binding. The smaller N-terminal domain (EfeO-N) is predicted to be cupredoxin (Cup) like, suggesting a previously unrecognised similarity between EfeO and Fet3p. Our structural modelling of the *E. coli* EfeO Cup domain identifies two potential metal-binding sites. Site I is predicted to bind Cu^{2+} using three conserved residues (C41 and 103, and E66) and M101. Of these, only one (C103) is conserved in classical cupredoxins where it also acts as a Cu ligand. Site II most probably binds Fe^{3+} and consists of four well conserved surface Glu residues. Phylogenetic analysis indicates that the EfeO-Cup domains form a novel Cup family, designated the 'EfeO-Cup' family. Structural modelling of two other representative EfeO-Cup domains indicates that different subfamilies employ distinct ligand sets at their proposed metal-binding sites. The ~100 *efeO* homologues in the bacterial sequence databases are all associated with various iron-transport related genes indicating a common role for EfeO-Cup proteins in iron transport, supporting a new copper-iron connection in biology.

Keywords Iron transport · Homology modelling · Cupredoxin · Electron transfer · GRID prediction · Metal binding site · DUF451 · Peptidase-M75

Electronic supplementary material The online version of this article (doi:10.1007/s10534-009-9262-z) contains supplementary material, which is available to authorized users.

M. B. Rajasekaran · S. C. Andrews (✉) · K. A. Watson
School of Biological Sciences, University of Reading, Whiteknights, Reading RG6 6AJ, UK
e-mail: s.c.andrews@reading.ac.uk

M. B. Rajasekaran · K. A. Watson
Structural Biology Unit at The Biocentre, University of Reading, Reading RG6 6AJ, UK

S. Nilapwar
Manchester Interdisciplinary Biocentre, University of Manchester, Manchester M1 7DN, UK

Introduction

Many major biological processes depend upon iron including oxygen transport and storage, nitrogen fixation, photosynthesis, DNA biosynthesis and the citric acid cycle (Andrews et al. 2003). The predominant environmental form of iron is the poorly soluble ferric form (10^{-18} M at pH 7.0). Relatively soluble (0.1 M at pH 7.0) ferrous iron is mainly restricted to anaerobic/microaerobic niches or low pH

environments. Unsurprisingly, the low solubility of ferric iron leads to problems of poor bioavailability. In addition, iron can cause toxicity in the presence of oxygen due to its tendency to engage in Fenton chemistry generating highly-reactive free-radical species. Thus, under aerobic conditions, iron causes problems of both availability and toxicity (Andrews et al. 2003; Carrondo 2003; Chiancone et al. 2004). Consequently, organisms have developed various strategies that allow them to acquire sufficient iron for growth whilst avoiding any deleterious effects caused by iron toxicity. Bacteria often employ a range of iron transporters, with specificities for alternative sources of iron, to enable them to overcome low iron availability. Within bacteria, the most common type of iron uptake system appears to involve siderophore-dependent transporters. These utilise extracellular chelators (siderophores) as ferric-iron solubilising agents. For Gram-negative bacteria, the resulting ferri-siderophore complexes are translocated across the outer-membrane (OM) by a TonB-dependent OM receptor (TBDR) (Andrews et al. 2003; Braun et al. 1998; Stintzi et al. 2000). The complexes are then bound by corresponding periplasmic-binding proteins and delivered to inner-membrane ABC-transporters. Host iron-containing compounds [transferrin, lactoferrin, haemoglobin, haem, haem-haemopexin] (Andrews et al. 2003; Bracken et al. 1999; Wandersman and Delepelaire 2004) can also be directly acquired by similar transporters involving binding of the host iron complexes to specific TBDRs. In contrast to ferric iron, ferrous iron can be directly imported using FeoAB transporters, which are thought to function under anaerobic-microaerophilic conditions (Cartron et al. 2006; Hantke 1987). In addition, metal-type ABC transporters like (Kehres et al. 2002) [e.g. SitABCD of *Salmonella typhimurium*] and NRAMP-like transporters (Kehres et al. 2000; Makui et al. 2000) [e.g. MntH of *Escherichia coli*] also translocate Fe^{2+}, although these systems appear to have greater affinity for Mn^{2+} than Fe^{2+}. ZupT in *E. coli* also has affinity for Fe^{2+}, along with a range of other divalent metals (Grass et al. 2002).

Recently, a new type of ferrous-iron transporter, designated EfeUOB, has been identified in *E. coli.* It consists of three protein components and acts mainly at low pH (Cao et al. 2007). Homologous systems are present in many other bacteria including *Bacillus subtilis* and *Neisseria meningitides* (Baichoo et al. 2002; Grifantini et al. 2003; Ollinger et al. 2006). EfeUOB resembles the high-affinity Fe^{2+} uptake system (Fet3p-Ftr1p) found in fungi (such as *Saccharomyces cerevisiae*) and some algae (Askwith and Kaplan 1997, 1998; De Freitas et al. 2003; Fang and Wang 2002; Fu et al. 2004; Paronetto et al. 2001) in that EfeU is homologous to Ftr1p, and both systems are ferrous-iron transporters. Ftr1p is a polytopic, cytoplasmic-membrane located, ferric permease containing seven transmembrane helices (TMH), belonging to the oxidase dependent Fe^{2+} transporter family (Debut et al. 2006). It possesses two 'RExxE' motifs, one in TMH 1 and the other in TMH 4, both of which are required for iron-uptake function (Severance et al. 2004; Stearman et al. 1996). These are thought to form an iron-binding site, with the conserved Glu residues acting as the direct ligands. Fet3p is a multicopper ferroxidase and is absolutely required for high-affinity Fe^{2+} uptake. It has three cupredoxin-like domains binding four copper ions and belongs to the multi-copper oxidase family (De Silva et al. 1995). This protein is anchored to the outer face of the cytoplasmic (or inner) membrane and closely associates with Ftr1p. The mechanism of Fet3p-Ftr1p iron transport involves reduction of environmental ferric ion to ferrous ion by means of the ferric/cupric reductases, Fre1p/Fre2p. The ferrous iron thus generated is bound by Fet3p and re-oxidised, using oxygen as the oxidant, to the ferric form which is then transferred to Ftr1p for translocation across the cytoplasmic membrane into the cytosol.

EfeUOB is encoded by a three gene operon, *efeUOB*, that is Fe^{2+}-Fur repressed and alkali-repressed (acid induced) by CpxAR. EfeU, like Ftr1p, is predicted to be located within the cytoplasmic membrane and to contain seven TMHs. EfeU possesses 27% amino acid sequence identity with Ftr1p and the two 'RExxE' motifs are conserved. EfeB is a periplasmic haem-containing peroxidase-like protein (Cartron et al. 2007) that is exported into the periplasm by the Tat pathway (Sturm et al. 2006). EfeO is also periplasmic and contains a predicted N-terminal signal sequence suggesting export to the periplasm in a Sec-dependent manner. EfeO consists of a C-terminal peptidase-M75 (M75) domain ($\sim$225 residues) and an N-terminal domain of $\sim$100 residues that appears to resemble the copper-containing cupredoxins.

In this study, the amino acid sequences of cupredoxins (Cup) and domains resembling the EfeO N-terminal domain (the EfeO-Cup domain) were compared and their phylogenetic relationship analysed. Results suggest that EfeO-Cup-like domains represent a distinct new bacterial group within the cupredoxin superfamily and form two sub-groups: Cup-M75 proteins and solo-Cup proteins. Structural modelling of representative EfeO-Cup-like domains indicates either one or two novel metal-binding sites (distinct from the Cu-site of classical cupredoxins) employing different sets of conserved ligands in each case, with one site favouring Cu^{2+} and the other preferring Fe^{3+}. These findings provide a new insight into the biochemical properties of EfeO and related proteins and suggest a potential iron uptake role as iron-binding and/or electron-transfer components.

Materials and methods

Secondary structure prediction, domain topology and multiple sequence alignment

The amino acid sequence of EfeO used in this study is from *E. coli* K-12 and is available at Uniprot Knowledgebase (Bairoch et al. 2005) (UniprotKB/SwissProt accession no: P0AB24). The SignalP server (Bendtsen et al. 2004) was used to predict the signal peptide cleavage site in EfeO. GeconcT (Ciria et al. 2004), was used to identify genes neighbouring *efeO* and *efeO* homologues. Domain predictions were obtained using various databases and servers: Superfamily (Gough et al. 2001), INTERPRO (Mulder et al. 2007), Pfam (Finn et al. 2006), SCOP (Murzin et al. 1995) and DomFOLD (Jones et al. 2005; Marsden et al. 2002; Ward et al. 2004) [incorporating DomSSEA, DISOPRED, mGenTHREADER and nFOLD]. The presence of low complexity regions in EfeO amino-acid sequence was indicated by NCBI-BLAST (Wheeler et al. 2006) and the Pfam database. The prediction of disordered regions in EfeO was performed by PrDOS (Ishida and Kinoshita 2007) and RONN (Yang et al. 2005). EfeO homologues (125 sequences) were identified using EBI-WU-BLAST2 (Altschul et al. 1990) using the default settings. Multiple-sequence alignments for either the Cup or peptidase-M75 domains (101 sequences containing Cup domain, and 92 sequences containing peptidase-M75 domain) were achieved using EBI-CLUSTALW (Thompson et al. 1994). Secondary-structure predictions were obtained using the PSIPRED (Bryson et al. 2005) and Jpred servers (Cuff et al. 1998). Multiple-sequence alignments obtained from EBI-CLUSTALW were analysed using GENEDOC (Nicholas et al. 1997) to generate improved alignments satisfying secondary structure element locations. Phylogenetic analysis was carried out using MEGA4 version software (Tamura et al. 2007) and the multiple-sequence alignments obtained above. The Neighbour-joining method with default settings was employed for the construction of trees.

Structural modelling of the Cup domain of EfeO and its select homologues

A search for suitable structural templates was performed using local similarity search algorithms (NCBI-BLAST, EBI-WU-BLAST2) against the Protein Data Base (PDB) (Berman et al. 2000) and fold recognition tools [Phyre (Kelley et al. 2000), mGenTHREADER (Bryson et al. 2005)] using the default settings. Amicyanin (Carrell et al. 2004) from *Paracoccus denitrificans* (PDB id: 1SFD) was the nearest structural homologue for EfeO according to mGENTHREADER (Table S3). Based on the multiple sequence alignment (Fig. 5a) and the mGENTHREADER result (Table S3), modelling of the EfeO homologues: Jan and Buc2 from domain I & V organisation (*Janibacter* sp. and *Burkholderia cepacia*, Table S1), respectively, was performed. For all models, amicyanin was used as the template, in order to maintain the sequence and structural alignment found between these homologues. The comparative modelling program, MODELLER (Marti-Renom et al. 2000), was used to construct all three theoretical models. The three models herein are referred to as Cup_{ecoli}, Cup_{jani} and Cup_{burk}, respectively. The validation programs PROCHECK (Laskowski et al. 1993), and MODFOLD (ModFOLDclust v 1.1) (McGuffin 2007; McGuffin 2008) were used for assessing the overall stereochemical quality of the models and selecting the final model in each case. The overall structural superimpositions and local superimpositions of metal binding sites between pairs of structures were performed using the program LSQKAB as provided by CCP4 (Collaborative Computational Project 1994).

Metal-binding site prediction for the Cup domain of EfeO and selected homologues by GRID

The program GRID (Goodford 1985) was used to predict potential metal binding sites for amicyanin (PDB id: 1SFD), as a control, the Cup domain of EfeO from *E. coli* (Cup_{ecoli}), and the Cup domain from *B. cepacia* (Cup_{burk}) and *Janibacter* sp. (Cup_{jani}). GRID predicts energetically favourable interactions between a probe (i.e. a small biologically active species such as water, carbohydrates, metals) and a protein. The results are represented in the form of 3-dimensional (3D) energy contours, showing the location of energetically favourable binding sites. The interaction energy E_{XYZ} of the probe at each xyz position on the GRID is calculated as a summation of energy components, specifically van der Waals, hydrogen bond, electrostatic interactions and a hydrophobic term as defined by $E_{xyz} = \Sigma E_{lj} + \Sigma E_{el} + \Sigma E_{hb}$. Contours at negative energy levels indicate favourable regions of attraction representing the most likely points of interaction between a probe (in this case, metal) and the protein. In this work, interactions between each of the Cup domains, Cup_{ecoli}, Cup_{jani} and Cup_{burk}, and their structural homologue amicyanin (PDB id: 1SFD) against relevant metal cation probes (Fe^{2+}, Cu^{2+}, Zn^{2+}, Mg^{2+}, Fe^{3+}) were carried out by GRID. All HETATM records (i.e. all non-protein atoms such as water, Cu^{2+}, sulphate, etc.) were first removed from the PDB file for amicyanin. There were no HETATM records in the homology models for Cup_{ecoli}, Cup_{jani} and Cup_{burk}. The program GRIN, the first step in the GRID calculation, was used to prepare each of the atomic coordinate files by removing hydrogens (in the case of the homology models) adding counterions (Na^+ was used in each case) to neutralise the overall charge on the protein taking care not to include counterions near any potential binding sites, and including the associated energy variables to individual atom types as defined in GRIN. The move directive in GRID was set to -1 to allow the counterions to move in response to the probe. The GRID calculations were performed in two stages; firstly, over the entire protein using a grid spacing (NPLA) of 1 Å, and secondly, over those regions highlighted by this first run which indicated the most promising regions for metal binding using a grid spacing of 0.5 Å. GRID allows assignment of the metal cation probes (JTYPE) to reflect the fact that these metals do not participate in hydrogen bonding and to emphasise the relative hardness or softness of the metal thus affecting the type of interaction. For example, a metal that interacts favourably with nitrogen and sulphur ligands but not oxygen would be assigned JTYPE 113. In GRID, metal cation probes by default are assigned JTYPE 110. For amicyanin and site I of Cup_{ecoli}, Cup_{jani} and Cup_{burk}, all metal probes were assigned JTYPE 113, to reflect the increased number of sulfhydryl ligands found at site I for each of these proteins. For site II in Cup_{ecoli}, all metals probes were assigned a JYTPE of 111, favouring interaction with oxygen ligands. All other GRID parameters used were default. The program MINIM, provided as part of the GRID package, was used to convert the GRID output to a readable format suitable for input to PYMOL (Delano 2002) for graphical viewing and interpretation.

Circular dichroism (CD) spectroscopy of $EfeO_{ecoli}$

The CD measurements for the mature $EfeO_{ecoli}$ protein were carried out using a JASCO J-810 spectropolarimeter. CD spectra of $EfeO_{ecoli}$ at 3.2 mg/ml in 30 mM MES buffer (pH 6.0) were measured at 25°C with a 0.01 cm pathlength cell in the far UV (190–260 nm) region, at a bandwidth of 1 nm using an average of four scans. The raw spectral data, ellipticity (θ), were expressed as mean residue ellipticity, $[\theta]_{mrw\lambda}$ (deg cm^2 $dmol^{-1}$), according to the equation:

$$[\theta]_{mrw\lambda} = MRW \times \theta_\lambda / 10 \times d \times c$$

where 'MRW' (mean residue weight for peptide bond) is described as M/N − 1 and M is the molecular weight of the mature $EfeO_{ecoli}$ (M = 38,313 Da), 'N' is the total number of amino acids (N = 349), 'θ_λ' is the observed ellipticity in degrees at each wavelength, 'd' is the path length of the cuvette (0.01 cm), 'c' is the concentration of the protein in (3.2 mg/ml, in this case). The raw CD spectral data for $EfeO_{ecoli}$ were submitted to the DICHROWEB (Whitmore and Wallace 2004) server for qualitative and quantitative secondary structure analysis. The dataset SP175, (Lees et al. 2006) consisting of 72 reference proteins from high resolution crystal structures (average 1.9 Å resolution) of type β and mixed α/β, was used as the

reference set for the analysis with an aim to reduce the spectral variability in the β-sheets of proteins.

Results and discussion

Domain topologies

The domain analysis results from the Superfamily server, Pfam (PF04302), DomFOLD and the INTERPRO database (IPR008972) indicate the presence of a cupredoxin (Cup) like domain at the N-terminus of the *E. coli* EfeO protein (residues 8–116) and a 'peptidase-M75' domain (M75, residues 135–372; Pfam PF09375, also known as 'DUF451') in the C-terminal region. Signal P predicted an N-terminal signal sequence for EfeO with the most likely cleavage site between residues 26 and 27 (ANA-AD). Together, the above predictions strongly suggest a two-domain organisation for EfeO of *E. coli* K-12 ($EfeO_{ecoli}$): an N-terminal Cup-like domain (Cup_{ecoli}) and a C-terminal M75 domain ($M75_{ecoli}$) separated by a flexible linker (Fig. 1 and Supplementary Information).

WU-BLAST2 analysis of the non-redundant database using the EBI site revealed 125 proteins displaying significant (*E*-value threshold 10) amino-acid sequence similarity (>20% identity) to $EfeO_{ecoli}$, all of which are from bacteria and possess either an M75 domain or a Cup domain. Of these, five domain organisations (I–V) are apparent (Fig. 1). The most common organisation (I) is that observed for *E. coli* consisting of an N-terminal Cup domain plus a C-terminal M75 domain (65 examples). All such proteins appear to possess an appropriately positioned signal sequence and so all are expected to be secreted. The second most common organisation (V) consists of an isolated Cup domain (without any C-terminal M75 domain, 34 examples) which implies that the N-terminal Cup domain can function independently of the M75 domain. The third most common domain organisation (II) comprises an isolated M75 domain (apparently secreted) without any Cup domain (18 examples). This indicates that the Cup domain may not be essential for the iron uptake process mediated by EfeO-like proteins. The fourth most frequent domain arrangement (III) contains an undefined N-terminal domain (~130 residues, designated 'domain X', found in *B. subtilis* and *Listeria* species) linked to a C-terminal M75 domain (seven examples) and so also lacks a recognisable Cup domain, again indicating that the Cup domain is not essential for EfeO function. Note that the *B. subtilis efe*-like locus, *ywbLMN*, has been shown to have a role in iron uptake (Baichoo et al. 2002; Ollinger et al. 2006) yet the protein corresponding to EfeO (YwbM), lacks the Cup domain. A signal sequence is predicted for the *B. subtilis* protein indicating these class III EfeO proteins are secreted. Finally, there is just one example of the class IV domain organisation where an N-terminal Ftr1p/EfeU-like domain is followed by a central Cup domain which is in turn followed by a C-terminal M75 domain giving an 'EfeU-EfeO fusion' protein. This fusion

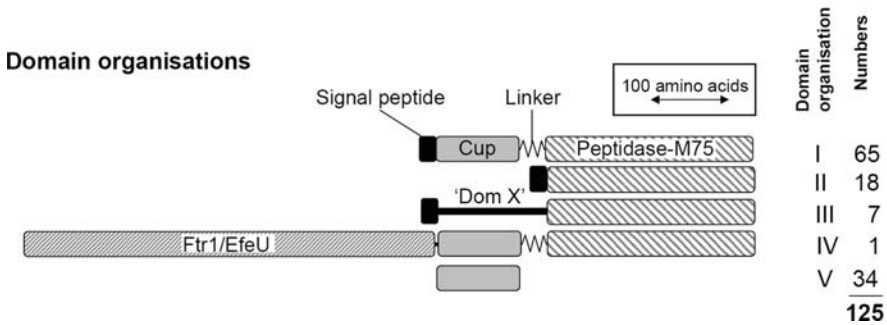

Fig. 1 Domain topology of EfeO. Five domain topologies (I–V) are apparent for EfeO homologues. The unknown domain found in the *Listeria* and *B. subtilis* proteins is indicated as '*Dom X*' (Domain X). The putative flexible linker region is indicated (not to scale). The SignalP-predicted signal peptide cleavage site for EfeO of *E. coli* K-12 is between residues 26 and 27. The predicted N-terminal Cup-like domain of $EfeO_{ecoli}$ extends from residue 30 to 116 followed by the C-terminal M75 domain (135–372). The low complexity region between the two domains (117–134) is also predicted by Pfam and the disorder prediction tools RONN and PrDOS

suggests that EfeU and EfeO interact during the iron uptake process.

Genetic context analysis

Seven gene organisations (from 58 loci) were identified for *efeO* homologues, all of which consisted of genes encoding M75 domains and/or Cup-M75 domain fusions (Fig. 2a). The most frequent arrangement is *efeUO$_I$B* (type A; 29 cases found in proteobacteria, actinobacteria and firmicutes), as found for *E. coli*. Given the similar genetic organisation and composition, it is highly likely that all 29 of these loci specify analogous ferrous-iron transporters. The next most common arrangement is *efeO$_{II}$BU* (type B; 13 cases found in firmicutes) where *efeU* is located at the end of the *efe* operon and the *efeO* gene is of class II (i.e. it does not specify a Cup domain). Such systems thus lack any Cup domain indicating, as suggested above, that this domain is not essential for function of Efe transporters. The next most common gene arrangement is *efeUO$_I$BO$_{II}$* (type C; 9 cases all from proteobacteria) which also resembles the type A organisation except for the presence of an additional *efeO* gene, at the end of the operon, which is of class II (lacking the Cup domain). The presence of two *efeO* genes, each specifying a different class of EfeO protein, indicates that EfeO$_I$ and EfeO$_{II}$ may have distinct, but related, functions in transport. The fourth most common organisation is *efeUO$_{III}$B* (type G; 3 cases from actinobacteria) which again resembles the type A arrangement except that the *efeO* gene encodes a class III EfeO protein containing domain X in place of the Cup domain. In addition, there are three other gene organisations that are found only once or twice: *efeO$_I$BU*; *efeUO$_{IV}$B*; and *efeO$_{II}$B*. Interestingly, all seven arrangements include an *efeB* homologue along with *efeO*, and in all but one case there is also an *efeU* equivalent. This indicates that EfeO is functionally

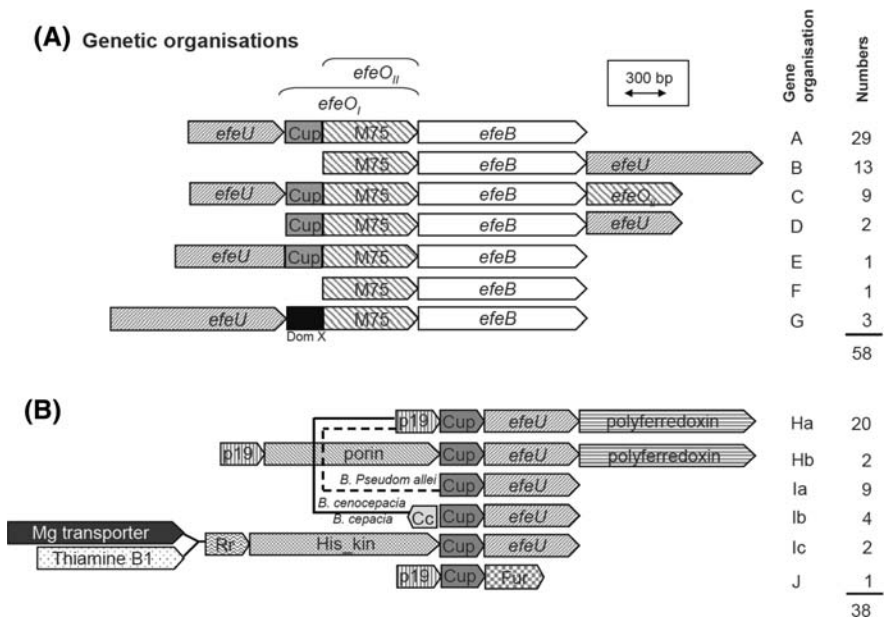

Fig. 2 Schematic representations of the genetic contexts of *efeO*-like loci. Loci containing genes specifying **a** EfeO-like M75 domains (domain organization I–IV; Fig. 1) or **b** solo-Cup domains (of domain V organization; Fig. 1). The EfeO$_{ecoli}$ amino acid sequence was used as input for Gene Context Tool 2 analysis, using default variables to search 580 genomes for sequence similarity (probability score minimum of 1×10^{-10}). A total of 58 matches were reported and seven genetic organizations were observed (A–G). Slight differences in the total number of matches discovered in Figs. 1, 2 reflect differences in cutoff scores applied and the sequence databases utilized. **b** Gene context analysis for the solo-Cup domain. Among the 38 matches obtained, three distinct gene organizations (H–J) were apparent. The presence of two solo-Cup domain gene clusters (of the H and I organizations) in *B. cenocepacia, B. cepacia* and *B. pseudomallei*, is indicated with *thick and dotted black lines*. '*Rr*' and '*His_kin*' indicate genes specifying a histidine kinase and response regulator from the two component signal transduction systems

dependent upon both EfeB and EfeU, and is consistent with a similar biological function for all 58 loci in ferrous iron uptake (Cao et al. 2007).

Gene-context analysis for those genes encoding a solo-Cup domain of the EfeO family (solo-Cup, domain V organisation), all from proteobacteria, revealed 38 gene clusters containing solo-Cup encoding genes. These divide into three major types (H–J; Fig. 2b), all possessing features suggestive of an iron (or metal) transport function and are discussed further in the Supplementary Information.

The peptidase-M75 domain of EfeO

The above analyses indicate that the M75 domain of EfeO is crucial for EfeUOB transporters. Indeed, no loci could be found where *efeU*- and *efeB*-like genes are both present in the absence of an M75-encoding *efeO*-like gene. To date, the only other characterized member of the peptidase-M75 family is the Zn-containing membrane-located metallopeptidase 'imelysin' protein from *Pseudomonas aeruginosa.* Secondary-structure prediction for the M75 domain of EfeO (residues 135–372) using Jpred and PSI-PRED indicates a high α-helix content (55 and 65%, respectively) and no β-strand composition. Furthermore, quantitative estimation based on the far-UV CD spectrum of $EfeO_{ecoli}$ (Fig. S1; residues 27–375) indicates a secondary structure composition of 48% α-helix, 13% β-strand and 39% random coil or turn. Subtraction of the anticipated secondary structure contribution of the EfeO Cup domain (50% β-strand; Fig. 4) from the CD-derived composition suggests that all the β-strands of EfeO are contributed by the Cup domain whereas all the α-helices are within the M75 domain, consistent with the PSIPRED and Jpred predictions.

Alignment of the amino acid sequence of the M75 domain of *E. coli* (K-12) EfeO with that of 91 EfeO-homologues (65 class I, 18 class II, 7 class III and one class IV sequences; Table S1; Fig. 3 and S2) shows that the highly conserved HxxE motif, that acts as a potential Zn-binding site for M75 metallopeptidases, is completely conserved within the M75-domain of EfeO proteins. This suggests that the M-75 domain is likely to be metal binding with the HxxE motif corresponding, at least in part, to a metal-binding site (here designated 'site III'). The above observations would be consistent with a role for the M75 domain of EfeO in specific inter-protein iron transfer, with the peptidase similarity providing protein–protein interaction capacity (perhaps with EfeU) and the HxxE motif acting as part of an Fe^{2+}-binding site enabling delivery of iron to EfeU.

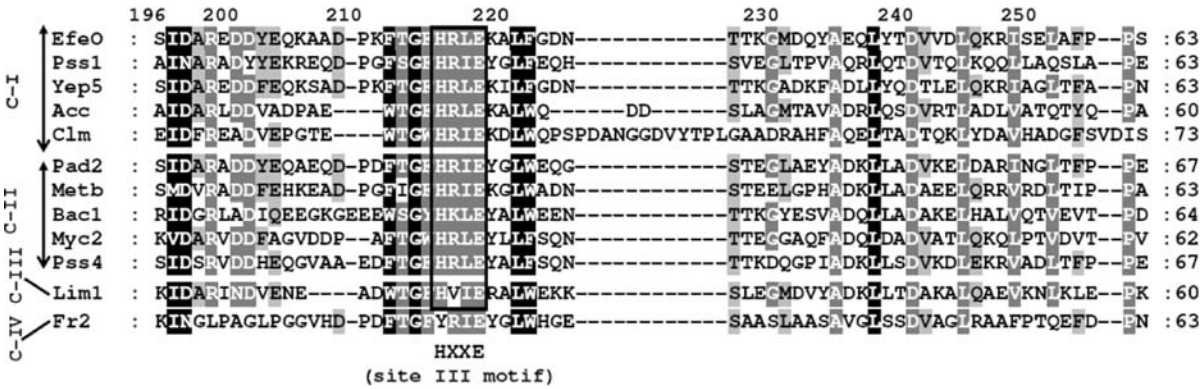

Fig. 3 Multiple-sequence alignment of a region of the M75 domain (residues 196–258) of EfeO from *E. coli* K-12 with selected homologues. The M75 domain region of class I, II, III and IV domain organization members were aligned and manually edited using EBI-CLUSTALW and GENEDOC software, respectively. The HxxE motif (site III), a potential zinc-binding site for the M75 metallopeptidases, is fully conserved. Absolutely conserved residues are shown in *white bold letters* against a *black background* whereas 80 and 60% conserved residues are shown in *white or black bold* type against a *grey background*, respectively. Residue numbers are based on those of $EfeO_{ecoli}$. Representatives from all four domain organizations are shown here for simplicity and the complete alignment with all members is provided in the supplementary material (Fig. S2). A description of each protein is provided in Table S1

The cupredoxin (Cup) domain of EfeO

The Cup domain of the EfeO proteins is not fully conserved, it is present in more than 70% of all cases observed, indicating an important, although not vital, function in ferrous transport. In addition, it is found as a solo domain associated with 38 potential bacterial iron/metal transporters which is again indicative of an important iron (or metal) transport-related function. Cupredoxins (blue-copper proteins) are copper-containing, mono-domain, electron-transfer proteins possessing a 'Greek key' β-barrel topology. They generally contain a single 'type I' (or 'blue-copper binding') copper ion acting as a high-potential (>250 mV) electron carrier. The copper is held in a distorted tetrahedral fashion by one conserved Cys and two conserved His residues, together with one or two variable axial ligands (e.g. Met, Gln, Leu) (Adman 1991; Carrell et al. 2004; Dennison 2005; Messerschmidt et al. 2001). The so called 'multicopper oxidases' (MCO) possess two, three or six cupredoxin domains (Adman et al. 1995; Suzuki et al. 2000; Messerschmidt et al. 1993; Ducros et al. 2001; Zaitseva et al. 1996). Such proteins either function as oxidases or reductases and typically possess four Cu atoms classified into three types: a type I, II and III (Solomon et al. 1996). Since only a single Cup domain is found within EfeO proteins, this domain would appear more closely similar to the cupredoxins suggesting a role in electron transfer.

Sequence conservation and phylogenetic relationships of the Cup domains of EfeO proteins

A multiple-sequence alignment of the Cup domain of 119 EfeO-like sequences and cupredoxins (Fig. 1) (Table S1; Fig. 4 and S3) shows that the sequences fall into three major groups on the basis of shared conservation, designated Cup-I (Cup domains from class I and IV EfeO proteins), Cup-II (solo domain Cup proteins from class V EfeO proteins) and Cup-III (classical cupredoxin proteins). Unsurprisingly, these three groups form three distinct clads in the corresponding phylogenetic tree (Fig. S4). The defined secondary-structure elements of cupredoxins appear well conserved with few interruptions caused by insertions or deletions. In addition, the predicted secondary structure for the Cup domain of EfeO and its homologues matches the defined structure well (80%), suggesting an accurate alignment has been achieved (Fig. 4 and S3). Thus, all of the proteins included in Fig. 4 (and S3) are likely to adopt a cupredoxin fold.

Only three residues are absolutely conserved (G51, G97 and Y99), which indicates the divergence of these sequences (Fig. S3; see Table 1 for roles played by such residues). Of particular note is the manner in which the patterns of conserved residues acting as defined, or potential, Cu ligands varies between the three distinct groups (Fig. 4). For the Cup-I group (class I and IV EfeO Cup domains), only the Cys (position 103) is conserved among the four Cu ligands (two His and one Cys and one Met) typically employed by cupredoxins. However, an alternative potential ligand (C41) is highly conserved in the Cup-I proteins, but not in the other groups, indicating that this residue may act as a novel metal ligand for members of this group. An EEREN motif (residues 77–81) is also highly conserved in the Cup-I group only, that may also contribute to metal binding. For the Cup-II group, a different pattern of potential metal ligands is observed. Of the four classical type I Cu ligands, only H106 is conserved. However, there is a well conserved Asp residue at the same position as the Cys residue normally employed as a Cu ligand in cupredoxins (position 103) which could therefore act as a metal ligand for Cup-II proteins. In addition, there is an absolutely conserved RKEKV motif (residues 77–81) that replaces the (E/D)EREN motif observed in the Cup-I group and this may also contribute to metal binding through its Glu residue. Both the Cup-I and -II groups share a very well conserved E(W/F)E motif (66–68) that may also act in metal binding. This motif is absent from the classical cupredoxins of the Cup-III group suggesting a specific purpose within the EfeO-like Cup proteins. It is interesting that both the Cup-I and -II proteins possess just one of the four highly conserved Cu ligands employed by the Cup-III proteins, but the identity of the conserved residue is different in each case. This highlights the cupredoxin-like nature the Cup-I and -II proteins but also indicates that they have each evolved in a different manner presumably to perform somewhat different functions.

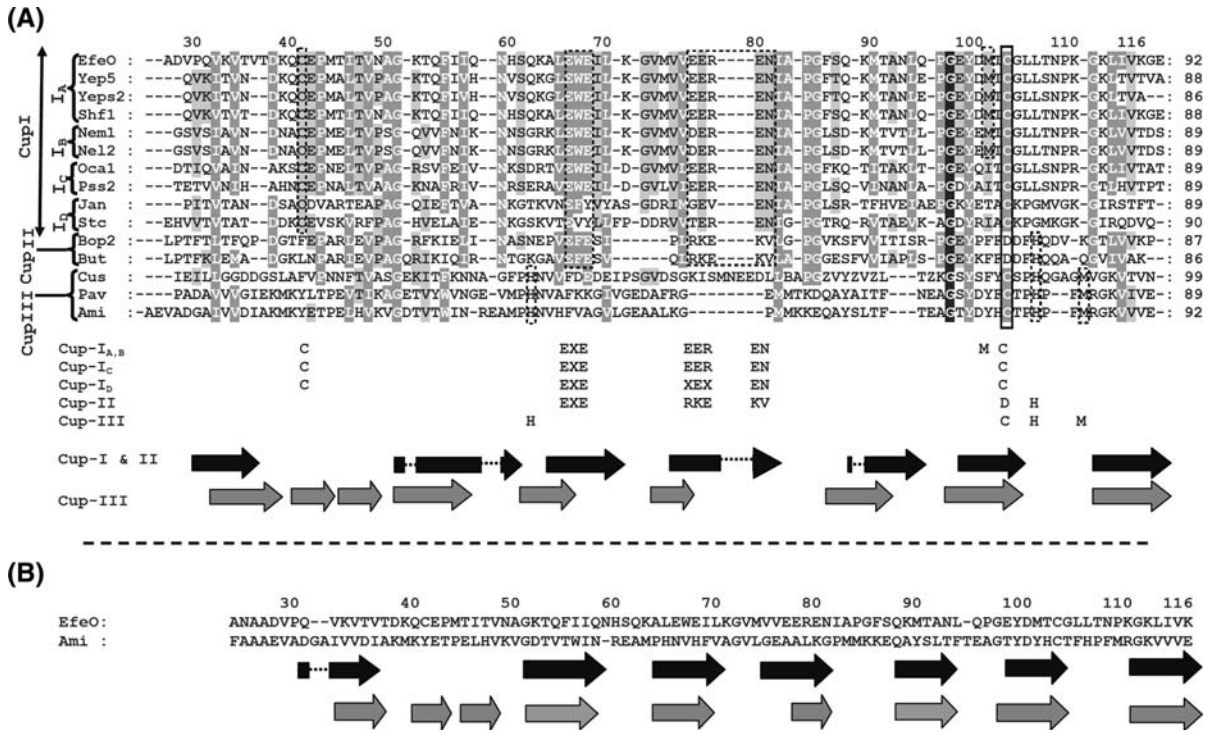

Fig. 4 Multiple-sequence alignment of the Cup domain of EfeO from *E. coli* K-12 with related proteins. **a** The complete alignment (Fig. S3) includes 65 and 34 EfeO-like Cup-I and Cup-II domains, respectively, together with 19 classical (Cup-III) cupredoxins from the 'copper bind' family (PF00127). Representatives from each species for the three families are shown here for simplicity. Alignment of the Cup-I and -II domains was achieved as for Fig. 3. The 'copper bind' cupredoxin alignment was taken from Pfam and then aligned with the Cup-I and -II sequences to generate a final combined alignment. The sequences are organized into three groups (Cup-I, -II and -II) on the basis of sequence similarity and phylogeny, with the Cup-I family further divided into four subfamilies: Cup-I_A, I_B, I_C, I_D (Fig. S4). Conserved known (or potential) metal ligands for each families are indicated in *bold letters at the bottom* of the alignment. The consensus defined secondary structure (β-strands only) for 19 classical cupredoxins (Cup-III proteins) is indicated by *grey arrows*, and that predicted by PSIPRED and Jpred for the Cup domain of $EfeO_{ecoli}$ is shown with *black arrows on the bottom* of the alignment. The conserved motifs and potential metal ligands in the case of Cup I–III families are highlighted with *dotted boxes* whereas the potential metal ligands which are common among these families are in *solid-line boxes*. **b** Pairwise sequence-structure based alignment between Cup_{ecoli} and *P. denitrificans* amicyanin. The pair-wise alignment was generated using mgenTHREADER and this alignment was used to generate the homology model of Cup_{ecoli} (Fig. 5). The β-strands for Cup_{ecoli} and amicyanin are indicated with *black* and *grey arrows*, respectively

Structural modelling of the Cup domain of EfeO and select homologues

Although the above sequence comparison indicates that there are two major types of EfeO-like Cup domain and that each has a unique pattern of potential metal ligands, it remains unclear whether the observed motifs occupy positions within the corresponding tertiary structures that would allow them to form metal-binding sites. To determine whether this is possible, a structural model was generated of the Cup domain of EfeO from *E. coli* K-12. A protein fold recognition method (McGuffin et al. 2001) was used to find structural homologues of Cup_{ecoli} (residues 27–116). Phyre and mGenTHREADER gave highest predicted matches for amicyanin, a known cupredoxin, from *P. denitrificans* (PDB id: 1SFD), and the derived structural alignment gave a sequence identity of 18%. PSIPRED and Jpred evaluation of the alignment of secondary structural elements for EfeO and amicyanin gave an overall agreement of 90%, with slight differences in the third and fourth β-strands of Cup_{ecoli} (III and IV; Fig. 5b). MODELLER was used to construct an amicyanin-based model for Cup_{ecoli} (Table S3). There were no irregularities in the stereochemical parameters according

Table 1 Summary of the GRID predicted metal binding ligands for site I & site II of Cup_{ecoli}, Cup_{burk}, Cup_{jani} and the known copper binding site of amicyanin (PDB id: 1SFD)

GRID probe	Residues in contact with GRID probe									
	41	62	66	68	77	78	80	101	103	106	111
Cu site: amicyanin											
Cu^{2+}		H							C	H	
[a]Cu^{2+}		H							C	H	M
Site I											
Cup_{ecoli} (Cup-I_A)											
Cu^{2+}	C		E					M	C		
Fe^{2+}	C		E					M	C		
Zn^{2+}	C		E					M	C		
Mg^{2+}	C		E					M	C		
Fe^{3+}			E						C		
Cup_{jani} (Cup-I_D)											
Cu^{2+}	C		E						C		
Fe^{2+}	C		E						C		
Zn^{2+}	C		E						C		
Mg^{2+}	C		E						C		
Fe^{3+}			E						C		
Cup_{burk} (Cup-II)											
Cu^{2+}			E						D	H	
Mg^{2+}			E						D	H	
Fe^{2+}			E						D	H	
Zn^{2+}			E						D	H	
Fe^{3+}			E						D		
Site II											
Cup_{ecoli} (Cup-I_A)											
Fe^{3+}				E	E	E	E				
Fe^{2+}				E	E	E	E				
Cu^{2+}				E	E	E	E				
Zn^{2+}				E	E	E	E				
Mg^{2+}				E	E	E	E				

[a] Crystallographic copper for amicyanin. The preferred metals for the different binding sites in each protein are listed, according to their relative interaction energies for the respective M^{2+} and M^{3+} probes used. Residues in bold are potential Cup-I/II Cu ligands that are conserved in classical cupredoxins (e.g. amicyanin)

to a PROCHECK analysis, and the Ramachandran plot showed 95% of residues in the most favoured regions with no residues in the disallowed regions suggesting a stereochemically accurate model. A pairwise Cα structural superimposition between the Cup_{ecoli} model and amicyanin (1SFD), using the program LSQKAB as provided by CCP4, gave an RMS deviation of 0.7 Å (Fig. 5a).

The Cup_{ecoli} domain model structure

The Cup_{ecoli} domain structural model consists of seven β-strands (strand I, residues 32–36; II, 52–59; III, 65–70; IV, 75–81; V, 88–94; VI, 98–103; VII, 110–116) accounting for 50% of the Cup portion of the $EfeO_{ecoli}$ polypeptide (Fig. 5a). The β-strands are organised into two sheets forming a Greek key β-barrel structure typical of cupredoxins. In accordance with the multiple-sequence alignment (Fig. 4), the model suggests a metal-binding site (designated site I) comprising two conserved Cys residues (residues 41 and 103), the conserved E66 residue, and the partially-conserved M101, all of which appear to be in close structural proximity (Fig. 5b). A comparison between the site I motif of Cup_{ecoli} (C41, E66, M101 and C103) and the copper-binding site of classical cupredoxins (amicyanin, azurin and pseudoazurins) (Carrell et al. 2004; Crane et al. 2001; Inoue et al. 1999) shows both sequence and structural conservation for these proteins of a single Cys residue (Cys103; Fig. 4 alignment, Fig. 5b, c), supporting a role for this Cys in metal binding for Cup_{ecoli}. The two His residues involved in copper binding in cupredoxins, as mentioned above, are not conserved in Cup_{ecoli}. These His residues superimpose with two Leu residues in the Cup_{ecoli} model. However, C41 of Cup_{ecoli} is well conserved in Cup-I sequences and together with C103 forms part of putative metal binding site I in Cup_{ecoli}. The potential role of Cys residues as ligands in metal binding is well known (Cobbett and Goldsbrough 2002; Kojima et al. 1999; Klaassen et al. 1999; Murphy and Taiz 1995; Palmiter 1998; Solioz and Vulpe 1996). The Met residue (M111 in amicyanin) that usually occurs as an axial ligand in cupredoxins is not present in the Cup_{ecoli} model; instead there is a Pro. However, an alternative residue, M101, appears able to act as a distal ligand at site I in the model. The final residue in site I, E66, structurally replaces the Asn residue (not a Cu ligand in classical cupredoxins) typically observed in other cupredoxins. It is possible that the carboxylate (or carbonyl) of E66 of Cup_{ecoli} could perform a role in metal binding at site I.

The Cup_{ecoli} model contains a negatively-charged cluster of residues (E68, 77, 78 and 80) on the protein

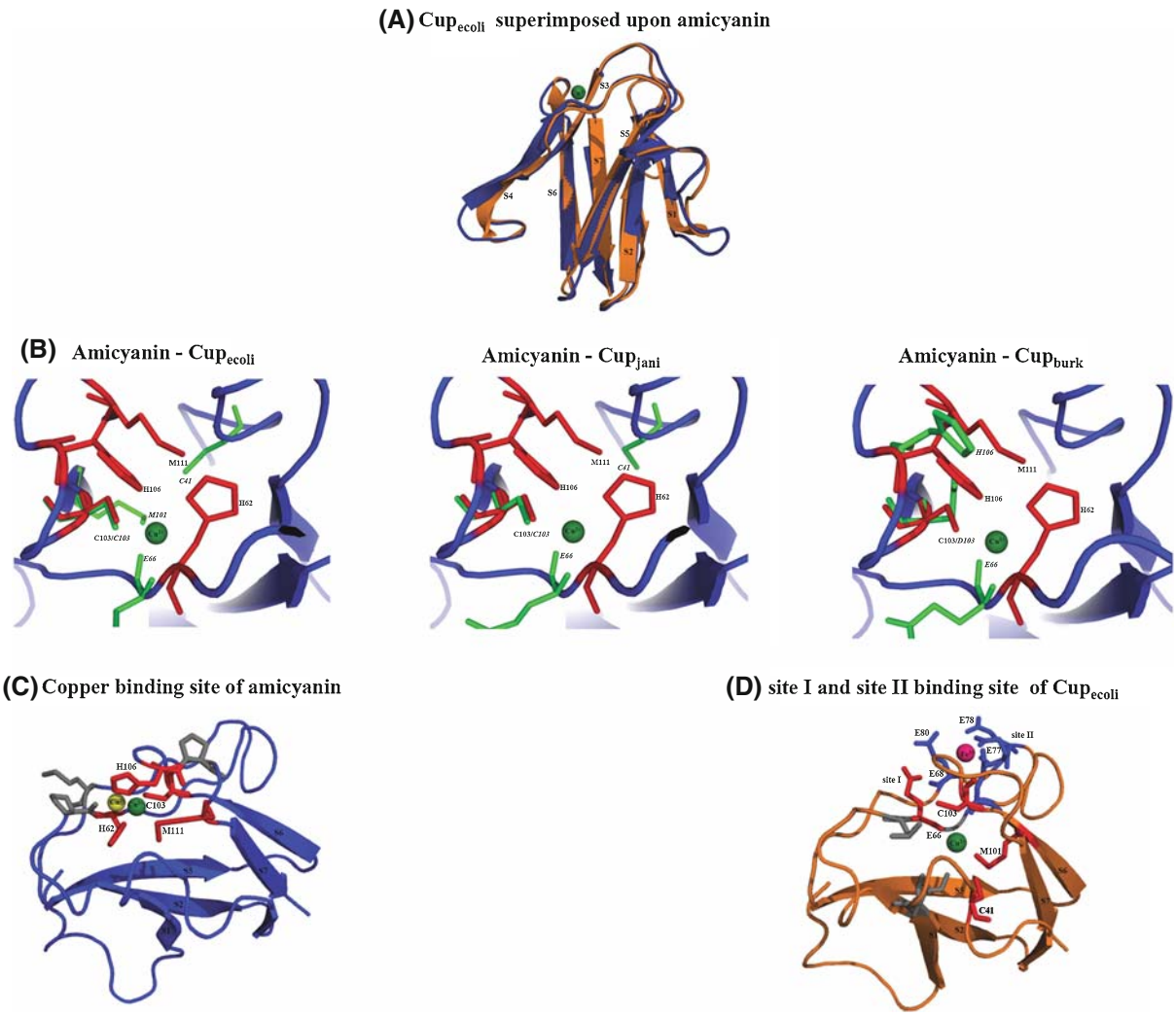

Fig. 5 Cup domain homology models and metal-binding site prediction for the Cup domain of EfeO from *E. coli* K-12. **a** The Cup_{ecoli} domain homology model (*orange*) was generated using amicyanin of *P. denitrificans* (PDB id: 1SFD) as template (*blue*). The β-strands are numbered: S1-S7. **b** Structural comparison of representatives from the Cup-I_A (Cup_{ecoli}), Cup-I_D (Cup_{jani}), Cup-II (Cup_{burk}) and Cup-III (amicyanin) groups, based on the homology models and the crystallographic structure of amicyanin. The comparison demonstrated conservation of two residues (E66 and C103) for Cup_{ecoli} (*green*) and Cup_{jani} (*green*), conservation a His residue (H106) between Cup_{burk} (*green*) and amicyanin (*red*). A conserved Asp residue, found only in the Cup-II family (as represented by Cup_{burk}), structurally superimposes with the otherwise highly conserved C103 in the remaining Cup family members. Residue numbering is according to EfeO and the view is approximately 90° rotated in the *y* axis with respect to the view in A. **c** Close up view of the copper-binding site in the crystallographically determined amicyanin structure. Metal-binding residues (H62, C103, H106, M111) are shown as *red sticks* and surrounding residues, forming a hydrophobic patch (P52, M71, P96) are shown as *grey sticks*. The crystallographic and GRID predicted position for copper are represented as *green* and *yellow spheres*, respectively. **d** Close up view of the predicted metal-binding sites showing specific residues, according to GRID, for the Cup_{ecoli} domain. According to GRID, site I has preference for an M^{2+} (with preference for Cu^{2+}, indicated by a *green sphere*) in tetrahedral geometry (C41, E66, M101 and C103) and site II has preference for Fe^{3+} (indicated by a *magenta sphere*) in square planar geometry (E68, 77, 78 and 80), highlighted with *red and blue sticks* (respectively) with a surrounding hydrophobic patch of residues (V36, I56, L65, W67) shown as *grey sticks*. The views in the case of **c** and **d** are approximately 90° rotated in the *x* axis with respect to the view in **b**. Structural comparisons of amicyanin and the third Cup domain (residues 382–500) of Fet3p (the third Cup domain of Fet3p was selected because, unlike the other two, it has a type 1 copper centre) with Cup_{ecoli}, show an overall similarity in fold but with much closer similarity to the amicyanin (RMSD 0.7 Å, and RMSD 2.8 Å, respectively) reflecting the corresponding sequence identities (18 and 6%, respectively) (Color figure online)

surface, adjacent to the proposed copper binding site. Three of these residues (E68, 78 and 80) are absolutely or highly conserved in the Cup-I group. This negative patch of Glu residues represents a second potential metal-binding motif, designated site II (Fig. 5d; Pastore et al. 2007; Taylor et al. 2005; Fang and Wang 2002). And finally, there is a hydrophobic patch comprised mainly of residues V34, I56, I57, L65 and W67 surrounding and below site I (Fig. 5d). A similar hydrophobic patch (M28, Met51, M71, P52, P94, P96 and F97), surrounding an exposed His residue, is also seen in amicyanin and acts as potential electron transfer site (Durley et al. 1993). Collectively, these highly conserved residues and motifs in the Cup-I proteins may play a significant role in metal binding and electron transfer as part of the EfeUOB-mediated iron uptake process.

Structural comparison of the metal-binding sites of Cup-I_A, Cup-I_D, Cup-II and Cup-III proteins

It is interesting to observe that while all four site I ligands (C41, E66, M101, C103) found in the Cup-I_A group proteins (e.g. Cup_{ecoli}) are conserved in the Cup-I_B group, only three out of four site I ligands (all but M101) are conserved in Cup-I_C and Cup-I_D groups. For the solo-domain Cup-II proteins, C41 and M101 are absent and C103 is replaced with an Asp. This leaves E66 as the only conserved potential ligand at site I for the EfeO-Cup family as a whole. However, there are other well-conserved residues that could act as alternative potential metal ligands for the Cup-$I_{C,\ D}$ and -II proteins (Fig. 4). In order to explore the distinctive features of the various EfeO-like Cup domains in more detail, representative proteins from the Cup-$I_{C,\ D}$ (Cup_{jani}) and Cup-II groups (Cup_{burk}) were selected for model building and structural comparison with Cup_{ecoli} and amicyanin (Cup-I_A and -III groups). Both Cup_{jani} and Cup_{burk} produced classical cupredoxin folds with RMSD values of 0.5 and 0.9, respectively, when superimposed upon the amicyanin structure (Table S3). Comparison of the site I region in the three model structures indicates that all have the potential to bind metal but, as expected, would employ slightly different ligand sets. Residues E66 and C103 of Cup_{jani} (Cup-I_D) are well conserved within the Cup-$I_{C\text{-}D}$ groups and align structurally with the corresponding site I residues in the Cup-I_A and -II structures, and so could contribute to metal binding at site I (Fig. 5b). In addition, potential ligand C41, which is well conserved in the Cup-I group, may also act as a metal ligand at site I for Cup_{jani}. For Cup_{burk} (Cup-II), although potential ligand, C103, is replaced by an Asp (conserved in Cup-II proteins), D103 could still act as a metal ligand at this site (Fig. 5b). The absence of potential ligand C41 is compensated by the presence of H106 which is structurally aligned with the same residue in the Cu-binding site of amicyanin. This residue is highly conserved in the Cup-II and -III proteins only. This suggests a plausible, yet distinct, metal binding motif at site I of the Cup-II proteins comprising three residues: E66, D103 and H106 (Fig. 5b).

Structural superimposition of the predicted site II metal-binding residues ($_{EW}E—EE_RE_N$; potential metal ligands non-subscripted) of Cup_{ecoli}, with the equivalent residues in the other two models, Cup_{jani} ($_{EF}Y—GE_VE_N$) and Cup_{burk} ($_{EW}E—RK_EK_V$), shows that two of the four proposed ligands of this site are well conserved in Cup_{jani} (and other Cup-I_D proteins) with one of the other two replaced with Tyr, an alternative potential metal ligand. However, in Cup_{burk} (and other Cup-II protein) site II is poorly conserved with just one of the four Glu ligands in Cup_{ecoli} present, and none of the three alternative residues are likely to be able to bind metal (Fig. 4). This suggests that site II is unlikely to be a metal-binding site in the solo-Cup-II domain proteins, but could be in all of the Cup-I proteins.

GRID prediction of metal-binding sites in Cup-I_A, Cup-I_D, Cup-II and Cup-III proteins

To further assess the validity of these potential metal-binding sites, the program GRID was used to predict possible metal (Cu^{2+}, Zn^{2+}, Fe^{2+}, Fe^{3+}, Mg^{2+}) binding sites for each of the three Cup domain models as well as amicyanin. For amicyanin, the top most favourable interaction site was within 2 Å of the known crystallographic position for Cu^{2+} and showed interaction with three of the four key residues (H62, C103, H106) known to bind copper (Fig. 5c). For Cup_{ecoli}, GRID predicted two favourable metal-binding sites corresponding to sites I and II (Table 1; Fig. 5d). GRID identified site I as a mononuclear divalent-metal (Cu^{2+}, Fe^{2+}, Zn^{2+} or Mg^{2+}) binding site (Fig. 5d) with an energetically favoured interaction with Cu^{2+} employing all four site I ligands,

namely C41, E66, M101 and C103. These residues are in agreement with the potential metal ligands predicted above. In addition, GRID predicted a potential interaction of a trivalent Fe^{3+} at site I, however, only two (C103 and E66) of the four potential ligand residues are predicted to interact with the trivalent metal. This suggests a potential of site I in EfeO to support either a trivalent or divalent metal, but with a preference of Cu^{2+} over Fe^{3+}.

Analogous GRID calculations for Cup_{jani} (Cup-I_D) showed an equal preference for Cu^{2+} and Fe^{2+} at site I, with the three anticipated metal-binding ligands (C41, E66, C103) suggested in the multiple sequence alignment (Fig. 4) interacting with the metal (Table 1). For the Cup_{burk} domain model (Cup-II), the GRID predictions also showed a preference for Cu^{2+} at site I, again showing energetically favourable interactions (Table 1) with the three conserved metal-binding residues (E66, C103 and H106) indicated by the multiple sequence alignment (Fig. 4). As with Cup_{ecoli}, both the Cup-II and -I_D models also show potentially favourable interactions with Fe^{3+} at site I (Table 1) but only two of the three possible ligands are predicted. In addition to the potential Cu^{2+}-binding site I, GRID also predicted a second metal-binding site (matching site II) for Cup_{ecoli} with a preference for Fe^{3+} involving residues E68, E77, E78 and E80 (corresponding to the $_{EW}E—EE_RE_N$ motif, see above). In contrast, there were no analogous GRID interactions predicted for an equivalent site II in either Cup_{burk} ($_{EW}E—RK_EK_V$) or Cup_{jani} ($_{EF}Y—GE_VE_N$).

Conclusion

Sequence comparison coupled with homology modelling suggest that the Cup_{ecoli} domain contains two well-conserved metal-binding sites separated by 12 Å with site I binding Cu^{2+} and site II favouring Fe^{3+}. Potential Cu-binding site I differs from that of the well characterised cupredoxins since Cup_{ecoli} employs ligand-residues that are mostly distinct from those of classical cupredoxins and the site is located in a slightly different region of the structure. The predicted Fe^{3+}-binding site II has no apparent equivalent in the cupredoxins and it is therefore probable that, if this site is genuine, it represents an adaptation to enable the iron-transport associated function of EfeO.

From 125 Cup_{ecoli} homologous, two major types of EfeO-like Cup domain were recognisable on the basis of domain organisation, metal-binding sites, phylogeny and genetic co-localisation. One type (Cup-I) is composed of two domains, an N-terminal Cup domain and a C-terminal peptidase-M75 domain (e.g. EfeO of *E. coli*), and the other (Cup-II) consists of a Cup domain only. All appear to be secreted beyond the cytoplasmic membrane and nearly all are encoded by genes adjacent to *ftr1*/*efeU* homologues suggesting interaction of the Cup domain with the Ftr1p-like ferric permease. All of the Cup-I proteins were additionally associated with an EfeB homologue, although in some cases the EfeO proteins consist of an M75 domain only. The Cup-II solo domain proteins lack the predicted Fe^{3+}-binding site II of the Cup-I proteins, employ a different set of ligands at the proposed Cu-binding site I and in no case are they associated with EfeB-like or M75-domain proteins. Instead, they tend to be associated with a cytosolic polyferredoxin and a periplasmic iron-transporting 'p19' protein. These differences in domain organisation, predicted metal-binding sites and transporter composition indicate a functional distinction between the Cup-I and Cup-II systems. Phylogenetic analysis (Fig. S4) shows that the EfeO-like Cup-I and -II proteins populate two distinct, but related, branches of the cupredoxin superfamily. The tree organisation suggests that the EfeO-Cup proteins form a major new cupredoxin sub-family. As far as we are aware, there is no other well defined case of a cupredoxin domain (excluding MCOs) functioning as part of any type of transporter, yet the data presented here strongly suggest that all the EfeO-like Cup proteins are components of iron (or metal) transporters. Thus, we suggest that the EfeO-like Cup proteins represent a new family of cupredoxins with all known members serving roles in iron transport.

A proposed mechanism of action for EfeO in the iron transport process is presented in Fig. 6. The model suggests that ferrous iron initially binds to EfeO and is then oxidised prior to transfer to EfeU. The electrons released from the ferroxidation process are delivered initially to the copper at site I, and are then passed on to the haem group of EfeB for disposal by combination with an unknown oxidant (which is unlikely to be molecular oxygen). Two potential iron-binding sites (II and III) are suggested for EfeO, one in the Cup domain and the other in the M75 domain—

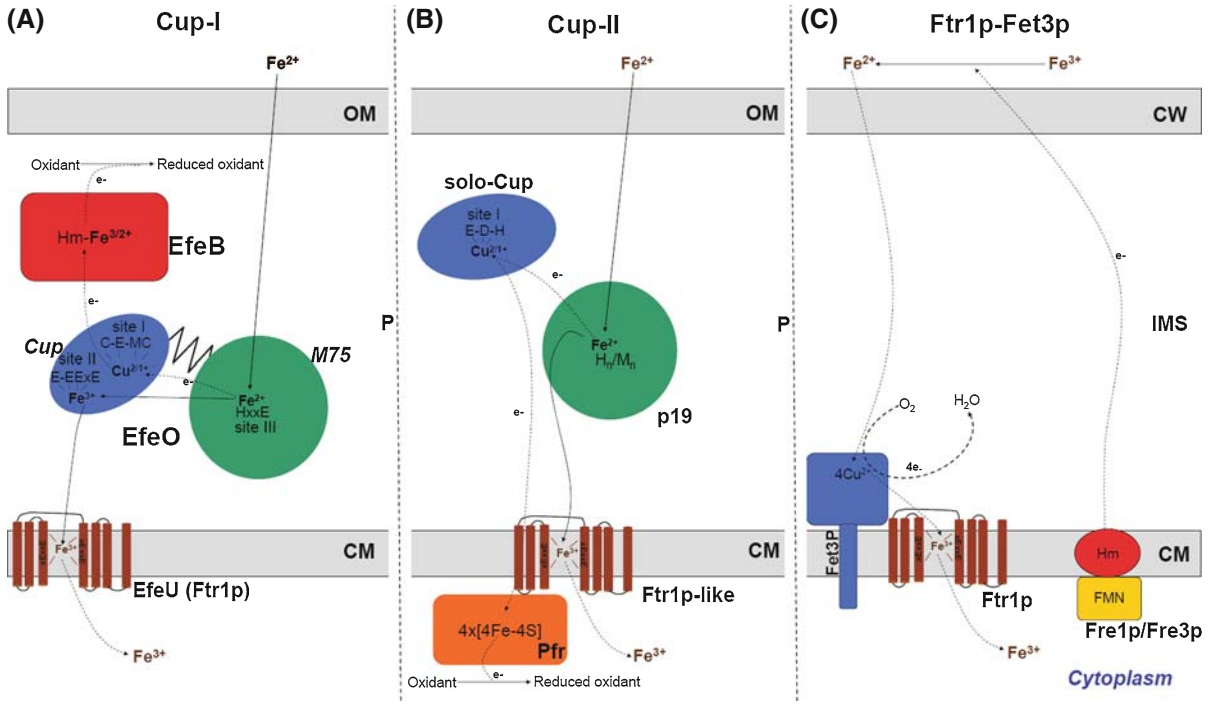

Fig. 6 Schematic representation of three types of Ftr1p/EfeU-dependent ferrous-iron transporter and the suggested mechanisms employed. Three types of defined or putative Ftr1p/EfeU—dependent transporter are shown: **a** the Cup-I type, exemplified by EfeUOB of *E. coli*; **b** the Cup-II type, for which the poorly characterised *C. jejuni* system is the best known example; and **c** the well characterized Ftr1p-Fet3p type, typified by the corresponding system of *S. cerevisiae*. In each case, environmental ferrous ion is the initial substrate and is first engaged within the intermembrane space (periplasm) by a specific ferrous iron binding protein (EfeO-*M75*, p19 or Fet3p). The ferrous iron is then oxidised and the released electrons delivered to an oxidant (such as molecular oxygen) via a Cu site (located in EfeO-*Cup*, solo-Cup or Fet3p). The oxidised iron is then transferred to the Ftr1p ferric-iron permease for translocation across the CM. For the Ftr1p-Fet3p system, ferrous-iron binding and disposal of electrons released through ferrous-iron oxidation occur within a single protein (Fet3p). For the other two systems, it is suggested that these processes are mediated by different components. In the case of the Cup-I system, EfeB (a DyP haem peroxidise) is proposed to function as the reductase using its haem (Hm) group as an intermediate electron acceptor, whereas for the Cup-II system the Pfr (polyferredoxin) is suggested to fulfil this purpose employing its four [4Fe-4S] clusters to accept and pass on electrons to an intracellular acceptor. The extracellular ferric-iron reductases, Fre1p and Fre3p, are indicated for the Frt1p-Fet3p system. Equivalent reductases have not yet been indentified for bacteria. Subcellular locations are indicated: *CM*, cytoplasmic membrane; *IMS/P*, intermembrane space/periplasm; and *OM/CW*, outer membrane/cell wall

this indicates that iron may bind to one site initially, and then move to the other site following oxidation. The transfer of iron between sites III and II could be mediated by the proposed inter-domain mobility of EfeO enabled by the inter-domain flexible linker. This mobility could also assist the proposed site I mediated electron transfer between the M75 domain and EfeB. As there are examples where the Cup domain is absent in EfeUOB-like transporters, it seems likely that the M75 domain would provide the primary ferrous-iron-binding site and that the proposed role of the Cup domain in electron transfer is not an essential requirement. This suggests that if iron does move between sites III and II then it is most likely to transfer in the site III to II direction, possibly following oxidation of the iron bound at site III. Thus, the presence of the Cup domain could enhance the efficiency of the iron transport process by increasing the rate of electron exchange, by acting as an intermediary in the transfer of iron from the M75 domain to EfeU and by increasing the availability of site III for ferrous iron binding. In cases where there are two EfeO proteins (with and without a Cup domain) it is possible that the two

EfeO proteins offer different metal affinities or selectivities.

For the solo EfeO-Cup proteins, the compositions of the corresponding iron transport systems differ considerably to those of the EfeUOB-like systems (Fig. 6). The differences in composition likely reflect the use of a distinct mechanism and/or discrete oxidants for the two types of transporter. It is possible that the p19 protein functionally replaces the M75 domain in ferrous-iron binding and that the polyferredoxin takes the place of EfeB in disposing of electrons, whereas the Cup protein retains its role in electron transfer. However, if this model is true it is unclear how electrons could pass from the periplasmic components to the polyferredoxin, unless the Ftr1p/EfeU homologue is able to mediate this. Clearly, an understanding of the details of these two types of bacterial transporter will require further study.

A major presumption is that the Cup-I and Cup-II transporter systems both utilise ferrous iron as substrate but deliver iron across the cytosolic membrane in the ferric form and so thus engage in ferroxidation. This assumption of a ferroxidation step is based upon the established mechanism employed by the related Ftr1p-Fet3p system and on the apparent inclusion of redox components (cupredoxins, haem peroxidases, polyferredoxins) within the Cup-I and Cup-II systems. However, the purpose of such a oxidation step within the transport processes mediated by these transporters is not clear. It is possible that an oxidation step would provide the driving force to energise the translocation process, or that the inclusion of a redox step enhances transport specificity. Currently, the energy mode employed by the Ftr1p-Fet3p system is unknown so ferroxidation remains a possibility. Indeed, a combination of enhanced specificity along with energy provision would provide a very satisfactory justification for the utilisation of a ferroxidation step within the iron-translocation pathways of the Ftr1p/EfeU-dependent transporters, as depicted in Fig. 6.

Acknowledgments We thank Peter Goodford and Jim Dunwell for helpful discussions in the preparation of the manuscript. We also thank Liam McGuffin for the helpful discussions related with the evaluation of structural models. Authors thank FELIX trust, Biotechnology and Biological Sciences Research (BBSRC) and Lister Institute of Preventive Medicine for providing funding for this study. We thank Sue Mitchell and Nick Spencer (Structural Biology Unit, The Biocentre) for technical support.

References

Adman ET (1991) Copper protein structures. Adv Protein Chem 42:145–197

Adman ET, Godden JW, Turley S (1995) The structure of copper-nitrite reductase from *Achromobacter cycloclastes* at five pH values, with NO_2^- bound and with type II copper depleted. J Biol Chem 270:27458–27474

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403–410

Andrews SC, Robinson AK, Rodriguez-Quinones F (2003) Bacterial iron homeostasis. FEMS Microbiol Rev 27: 215–237

Askwith C, Kaplan J (1997) An oxidase-permease-based iron transport system in *Schizosaccharomyces pombe* and its expression in *Saccharomyces cerevisiae*. J Biol Chem 272:401–405

Askwith C, Kaplan J (1998) Iron and copper transport in yeast and its relevance to human disease. Trends Biochem Sci 23:135–138

Baichoo N, Wang T, Ye R, Helmann JD (2002) Global analysis of the *Bacillus subtilis* Fur regulon and the iron starvation stimulon. Mol Microbiol 45:1613–1629

Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh LS (2005) The universal protein resource (UniProt). Nucleic Acids Res 33:D154–D159

Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340:783–795

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28:235–242

Bracken CS, Baer MT, Abdur-Rashid A, Helms W, Stojiljkovic I (1999) Use of heme-protein complexes by the *Yersinia enterocolitica* HemR receptor: histidine residues are essential for receptor function. J Bacteriol 181:6063–6072

Braun V, Hantke K, Koster W (1998) Bacterial iron transport: mechanisms, genetics, and regulation. In: Sigel A, Sigel H (eds) Metal ions in biological systems. Dekker, New York, pp 67–145

Bryson K, McGuffin LJ, Marsden RL, Ward JJ, Sodhi JS, Jones DT (2005) Protein structure prediction servers at University College London. Nucleic Acids Res 33:W36–W38

Cao J, Woodhall MR, Alvarez J, Cartron ML, Andrews SC (2007) EfeUOB (YcdNOB) is a tripartite, acid-induced and CpxAR-regulated, low-pH Fe^{2+} transporter that is cryptic in *Escherichia coli* K-12 but functional in *E. coli* O157:H7. Mol Microbiol 65:857–875

Carrell CJ, Sun D, Jiang S, Davidson VL, Mathews FS (2004) Structural studies of two mutants of amicyanin from *Paracoccus denitrificans* that stabilize the reduced state of the copper. Biochemistry 43:9372–9380

Carrondo MA (2003) Ferritins, iron uptake and storage from the bacterioferritin viewpoint. EMBO J 22:1959–1968

Cartron ML, Maddocks S, Gillingham P, Craven CJ, Andrews SC (2006) Feo—transport of ferrous iron into bacteria. Biometals 19:143–157

Cartron ML, Mitchell SA, Woodhall MR, Andrews SC, Watson KA (2007) 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. Acta Crystallogr Sect F 63:37–41

Chiancone E, Ceci P, Ilari A, Ribacchi F, Stefanini S (2004) Iron and proteins for iron storage and detoxification. Biometals 17:197–202

Ciria R, Abreu-Goodger C, Morett E, Merino E (2004) GeConT: gene context analysis. Bioinformatics 20:2307–2308

Cobbett C, Goldsbrough P (2002) Phytochelatins and metallothioneins: roles in heavy metal detoxification and homeostasis. Annu Rev Plant Biol 53:159–182

Collaborative Computational Project N (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr D B 50:760–763

Crane BR, Di Bilio AJ, Winkler JR, Gray HB (2001) Electron tunneling in single crystals of *Pseudomonas aeruginosa* azurins. J Am Chem Soc 123:11623–11631

Cuff JA, Clamp ME, Siddiqui AS, Finlay M, Barton GJ (1998) JPred: a consensus secondary structure prediction server. Bioinformatics 14:892–893

De Freitas J, Wintz H, Kim JH, Poynton H, Fox T, Vulpe C (2003) Yeast, a model organism for iron and copper metabolism studies. Biometals 16:185–197

De Silva DM, Askwith CC, Eide D, Kaplan J (1995) The FET3 gene product required for high affinity iron transport in yeast is a cell surface ferroxidase. J Biol Chem 270:1098–1101

Debut AJ, Dumay QC, Barabote RD, Saier MH Jr (2006) The iron/lead transporter superfamily of Fe^{2+}/Pb^{2+} uptake systems. J Mol Microbiol Biotechnol 11:1–9

Delano W (2002) The PyMOL molecular graphics system. www.pymol.org. Accessed Apr 2009

Dennison C (2005) Investigating the structure and function of cupredoxins. Coord Chem Rev 249:3025–3054

Ducros V, Brzozowski AM, Wilson KS, Ostergaard P, Schneider P, Svendson A, Davies GJ (2001) Structure of the laccase from *Coprinus cinereus* at 1.68 Å resolution: evidence for different type 2 Cu-depleted isoforms. Acta Crystallogr D 57:333–336

Durley R, Chen L, Lim LW, Mathews FS, Davidson VL (1993) Crystal structure analysis of amicyanin and apoamicyanin from *Paracoccus denitrificans* at 2.0 Å and 1.8 Å resolution. Protein Sci 2:739–752

Fang H-M, Wang Y (2002) Characterization of iron-binding motifs in *Candida albicans* high-affinity iron permease CaFtr1p by site-directed mutagenesis. Biochem J 368:641–647

Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R, Eddy SR, Sonnhammer EL, Bateman A (2006) Pfam: clans, web tools and services. Nucleic Acids Res 34:D247–D251

Fu Y, Lee H, Collins M, Tsai HF, Spellberg B, Edwards JE Jr, Kwon-Chung KJ, Ibrahim AS (2004) Cloning and functional characterization of the *Rhizopus oryzae* high affinity iron permease (rFTR1) gene. FEMS Microbiol Lett 235:169–176

Goodford PJ (1985) A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. J Med Chem 28:849–857

Gough J, Karplus K, Hughey R, Chothia C (2001) Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. J Mol Biol 313:903–919

Grass G, Wong MD, Rosen BP, Smith RL, Rensing C (2002) ZupT is a Zn(II) uptake system in *Escherichia coli*. J Bacteriol 184:864–866

Grifantini R, Sebastian S, Frigimelica E, Draghi M, Bartolini E, Muzzi A, Rappuoli R, Grandi G, Genco CA (2003) Identification of iron-activated and -repressed Fur-dependent genes by transcriptome analysis of *Neisseria meningitidis* group B. Proc Natl Acad Sci USA 100:9542–9547

Hantke K (1987) Ferrous iron transport mutants in *Escherichia coli* K12. FEMS Microbiol Lett 44:53–57

Inoue T, Nishio N, Suzuki S, Kataoka K, Kohzuma T, Kai Y (1999) Crystal structure determinations of oxidized and reduced pseudoazurins from *Achromobacter cycloclastes*. Concerted movement of copper site in redox forms with the rearrangement of hydrogen bond at a remote histidine. J Biol Chem 274:17845–17852

Ishida T, Kinoshita K (2007) PrDOS: prediction of disordered protein regions from amino acid sequence. Nucleic Acids Res 35:W460–W464

Jones DT, Bryson K, Coleman A, McGuffin LJ, Sadowski MI, Sodhi JS, Ward JJ (2005) Prediction of novel and analogous folds using fragment assembly and fold recognition. Proteins 61:143–151

Kehres DG, Zaharik ML, Finlay BB, Maguire ME (2000) The NRAMP proteins of *Salmonella typhimurium* and *Escherichia coli* are selective manganese transporters involved in the response to reactive oxygen. Mol Microbiol 36:1085–1100

Kehres DG, Janakiraman A, Slauch JM, Maguire ME (2002) SitABCD is the alkaline Mn^{2+} transporter of *Salmonella enterica Serovar Typhimurium*. J Bacteriol 184:3159–3166

Kelley LA, MacCallum RM, Sternberg MJ (2000) Enhanced genome annotation using structural profiles in the program 3D-PSSM. J Mol Biol 299:499–520

Klaassen CD, Liu J, Choudhuri S (1999) Metallothionein: an intracellular protein to protect against cadmium toxicity. Annu Rev Pharmacol Toxicol 39:267–294

Kojima T, Binz PA, Kägi JHR (1999) Nomenclature of metallothionein: proposal for a revision. In: Klaassen CD (ed) Metallothionein IV. Birkhauser, Boston, pp 7–13

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291

Lees JG, Miles AJ, Wien F, Wallace BA (2006) A reference database for circular dichroism spectroscopy covering fold and secondary structure space. Bioinformatics 22:1955–1962

Makui H, Roig E, Cole ST, Helmann JD, Gros P, Cellier MF (2000) Identification of the *Escherichia coli* K-12 Nramp orthologue (MntH) as a selective divalent metal ion transporter. Mol Microbiol 35:1065–1078

Marsden RL, McGuffin LJ, Jones DT (2002) Rapid protein domain assignment from amino acid sequence using predicted secondary structure. Protein Sci 11:2814–2824

Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, Sali A (2000) Comparative protein structure modeling of genes and genomes. Annu Rev Biophys Biomol Struct 29:291–325

McGuffin L (2007) Benchmarking consensus model quality assessment for protein fold recognition. BMC Bioinformatics 8:345

McGuffin LJ (2008) The ModFOLD server for the quality assessment of protein structural models. Bioinformatics 24:586–587

McGuffin LJ, Bryson K, Jones DT (2001) What are the baselines for protein fold recognition? Bioinformatics 17:63–72

Messerschmidt A, Luecke H, Huber R (1993) X-ray structures and mechanistic implications of three functional derivatives of ascorbate oxidase from zucchini: reduced, peroxide and azide forms. J Mol Biol 230:997–1014

Messerschmidt A, Huber R, Poulos T, Wieghardt K (2001) Handbook of metalloproteins. Wiley, New York

Mulder NJ, Apweiler R, Attwood TK et al (2007) New developments in the InterPro database. Nucleic Acids Res 35:D224–D228

Murphy A, Taiz L (1995) Comparison of metallothionein gene expression and nonprotein thiols in ten arabidopsis ecotypes (Correlation with copper tolerance). Plant Physiol 109:945–954

Murzin AG, Brenner SE, Hubbard T, Chothia C (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 247:536–540

Nicholas KB, Nicholas HB, DW D (1997) GeneDoc: analysis and visualization of genetic variation. EMBNEW. NEWS 4:14. www.nrbsc.org/gfx/genedoc. Accessed Feb 2009

Ollinger J, Song KB, Antelmann H, Hecker M, Helmann JD (2006) Role of the Fur regulon in iron transport in *Bacillus subtilis*. J Bacteriol 188:3664–3673

Palmiter RD (1998) The elusive function of metallothioneins. Proc Natl Acad Sci USA 95:8428–8430

Paronetto MP, Miele R, Maugliani A, Borro M, Bonaccorsi di Patti MC (2001) Cloning of *Pichia pastoris* Fet3: insights into the high affinity iron uptake system. Arch Biochem Biophys 392:162–167

Pastore C, Franzese M, Sica F, Temussi P, Pastore A (2007) Understanding the binding properties of an unusual metal-binding protein-a study of bacterial frataxin. FEBS J 274:4199–4210

Severance S, Chakraborty S, Kosman DJ (2004) The Ftr1p iron permease in the yeast plasma membrane: orientation, topology and structure–function relationships. Biochem J 380:487–496

Solioz M, Vulpe C (1996) CPx-type ATPases: a class of P-type ATPases that pump heavy metals. Trends Biochem Sci 21:237–241

Solomon EI, Sundaram UM, Machonkin TE (1996) Multicopper oxidases and oxygenases. Chem Rev 96:2563–2606

Stearman R, Yuan DS, Yamaguchi-Iwai Y, Klausner RD, Dancis A (1996) A permease-oxidase complex involved in high-affinity iron uptake in yeast. Science 271:1552–1557

Stintzi A, Barnes C, Xu J, Raymond KN (2000) Microbial iron transport via a siderophore shuttle: a membrane ion transport paradigm. Proc Natl Acad Sci USA 97:10691–10696

Sturm A, Schierhorn A, Lindenstrauss U, Lilie H, Bruser T (2006) YcdB from *Escherichia coli* reveals a novel class of Tat-dependently translocated hemoproteins. J Biol Chem 281:13972–13978

Suzuki S, Kataoka K, Yamaguchi K (2000) Metal coordination and mechanism of multicopper nitrite reductase. Acc Chem Res 33:728–735

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Taylor AB, Stoj CS, Ziegler L, Kosman DJ, Hart PJ (2005) The copper-iron connection in biology: structure of the metallo-oxidase Fet3p. Proc Natl Acad Sci USA 102:15459–15464

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Wandersman C, Delepelaire P (2004) Bacterial iron sources: from siderophores to hemophores. Annu Rev Microbiol 58:611–647

Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. J Mol Biol 337:635–645

Wheeler DL, Barrett T, Benson DA et al (2006) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 34:D173–D180

Whitmore L, Wallace BA (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Res 32:W668–W673

Yang ZR, Thomson R, McNeil P, Esnouf RM (2005) RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. Bioinformatics 21:3369–3376

Zaitseva I, Zaitsev V, Card G, Moshkov K, Bax B, Ralph A, Lindley P (1996) The X-ray structure of human serum ceruloplasmin at 3.1 Å: nature of the copper centres. J Biol Inorg Chem 1:15–23