Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Andrey M. Grishin,[a] Eunice Ajamian,[a] Linhua Zhang[b] and Miroslaw Cygler[a,b]*

[a]Department of Biochemistry, McGill University, Montréal, Québec, Canada, and [b]Biotechnology Research Institute, NRC, Montréal, Québec, Canada

Correspondence e-mail: mirek@bri.nrc.ca

Received 26 May 2010
Accepted 11 July 2010

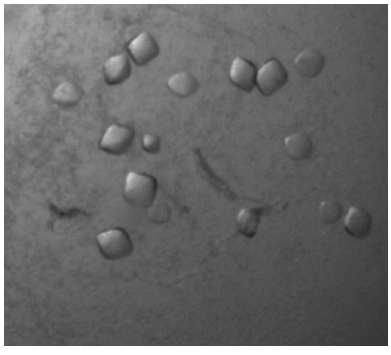
Crystallization and preliminary X-ray analysis of PaaAC, the main component of the hydroxylase of the *Escherichia coli* phenylacetyl-coenzyme A oxygenase complex

The *Escherichia coli paa* operon encodes enzymes of the phenylacetic acid-utilization pathway that metabolizes phenylacetate in the form of a coenzyme A (CoA) derivative. The phenylacetyl-coenzyme A oxygenase complex, which has been postulated to contain five components designated PaaABCDE, catalyzes ring hydroxylation of phenylacetyl-CoA. The PaaAC subcomplex shows low sequence similarity to other bacterial multicomponent monooxygenases (BMMs) and forms a separate branch on the phylogenetic tree. PaaAC, which catalyzes the hydroxylation reaction, was purified and crystallized in the absence of a bound ligand as well as in complexes with CoA, 3-hydroxybutyryl-CoA, benzoyl-CoA and the true substrate phenylacetyl-CoA. Crystals of the ligand-free enzyme belonged to space group $P2_12_12_1$ and diffracted to 2.65 Å resolution, whereas complexes with CoA and its derivatives crystallized in space group $P4_12_12$ and diffracted to ∼2.0 Å resolution. PaaAC represents the first crystallized BMM hydroxylase that utilizes a CoA-linked substrate.

1. Introduction

Escherichia coli is capable of aerobically metabolizing aromatic compounds (Díaz *et al.*, 2001). In particular, this bacterium contains a pathway for the degradation of phenylacetic acid (PA), which is metabolized into intermediates for the TCA cycle. The degradation of PA proceeds through a CoA derivative, which is an unusual route for the aerobic catabolism of aromatic compounds (Ferrández *et al.*, 1998). In some bacteria this pathway serves as the core of a complex metabolic network (catabolon) that is capable of the degradation of 2-phenylethylamine, styrene, phenylacetaldehyde, phenylacetyl amides, *n*-phenylalkanoic acids with an even number of C atoms and phenylacetyl esters (Luengo *et al.*, 2001). The range of chemistries performed by this catabolon shows promise for biotechnological applications such as the enzymatic synthesis of penicillins, whole-cell processes of styrene biodegradation, the production of new bioplastics and the production of 2-hydroxyphenylacetate (Luengo *et al.*, 2001). *E. coli* utilizes only a subset of the catabolon routes related to the utilization of PA and 2-phenylethylamine (Díaz *et al.*, 2001).

The genes coding for PA-metabolizing enzymes are organized into three transcription units within the *paa* gene cluster: *paaZ*, *paaABCDEFGHIJK* and *paaXY* (Ferrández *et al.*, 1998). The first two transcription units encode enzymes utilized in this pathway. The third unit encodes a negative regulator of *paa* genes, PaaX, while the role of PaaY is unknown. The proposed route for PA metabolism is presented in Fig. 1 (Ismail *et al.*, 2003). The first step in the pathway is the activation of PA by phenylacetate-CoA ligase (PaaK), yielding phenylacetyl-CoA (PA-CoA). The aromatic ring of the latter is subsequently aerobically hydroxylated by the putative phenylacetyl-CoA oxygenase complex (PaaABCDE), followed by ring opening by



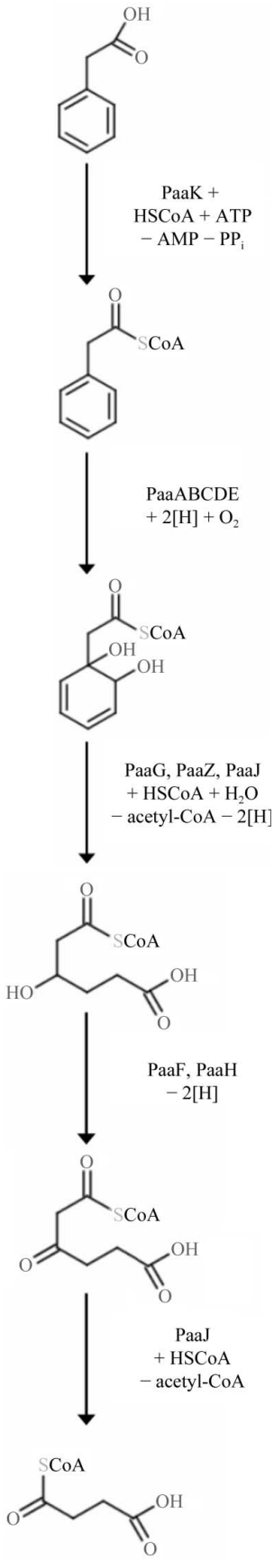


Figure 1
Proposed steps of the aerobic metabolic pathway of phenylacetate in *E. coli* (Ismail *et al.*, 2003).

PaaJ, PaaZ and PaaG to yield β-hydroxyadipyl-CoA. This compound is oxidized by PaaF and PaaH to β-ketoadipyl-CoA, which is cleaved to the convenient metabolites succinyl-CoA and acetyl-CoA by β-ketoadipyl-CoA thiolase (PaaJ) (Nogales *et al.*, 2007).

The phenylacetyl-coenzyme A oxygenase complex PaaABCDE has been classified within the bacterial multicomponent monooxygenases (BMMs; Fernández *et al.*, 2006). BMMs are divided into six major groups (Notomista *et al.*, 2003; Fernández *et al.*, 2006) and include soluble methane monooxygenase (MMO; Rosenzweig *et al.*, 1993, 1997), phenol hydroxylase (Cafaro *et al.*, 2004; Sazinsky *et al.*, 2006), toluene/*o*-xylene monooxygenase (ToMO; McCormick *et al.*, 2006; Notomista *et al.*, 2009) and toluene 4-monooxygenase (T4MO; Moe *et al.*, 2006; Bailey *et al.*, 2008). Phylogenetic analysis has shown that phenylacetyl-CoA oxygenase is the prototype of a new class within the large BMM family (Fernández *et al.*, 2006).

By analogy with other monooxygenases, PaaA, PaaC and PaaD have been predicted to form the dimeric hydroxylase $(ACD)_2$ core of the complex (Fernández *et al.*, 2006). Based on sequence homology, the PaaA subunit was proposed to contain a di-iron centre coordinated by four negatively charged residues and two histidines and to perform the oxygenase reaction, PaaC, which shows a low level of sequence identity to PaaA but is devoid of the catalytic centre, was proposed to play a structural role in complex formation, PaaD was thought to be the small subunit of the hydroxylase, PaaE has been assigned the role of the reductase component of the complex, which delivers electrons from NADH or NADPH through FAD or FMN and an iron–sulfur cluster to the iron in the active centre of PaaA, and PaaB has been proposed to play a regulatory role in the complex (Fernández *et al.*, 2006).

We have expressed PaaA and PaaC, the main components of the hydroxylase. We showed that PaaA expressed alone is insoluble, but that it can be obtained in a soluble form when co-expressed with PaaC, with which it forms a stable complex. We have obtained well diffracting crystals of the PaaAC complex with bound ligands, including CoA, 3-hydroxybutyryl-CoA, benzoyl-CoA and phenylacetyl-CoA, which represents the true substrate. We have also crystallized the ligand-free form of PaaAC.

2. Materials and methods

2.1. Cloning, expression and purification

The *paaA* and *paaC* genes were PCR-amplified from *E. coli* K-12 genomic DNA and cloned into a modified pET15b vector (Novagen) designated pFO4, a pRDSFDuet-1 vector (Novagen) derivative designated pZL72 and a pCDFDuet-1 vector derivative designated pZL71 (Novagen). In all cases *E. coli* BL21 (DE3) was used for protein expression.

An overnight inoculum of the transformed *E. coli* expression strain was diluted 200-fold into fresh TB medium supplemented with appropriate antibiotics and cells were allowed to grow at 310 K until the absorbance at 600 nm reached 0.6. Protein expression was induced by adding 0.1 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) and the culture was grown overnight at 293 K. Cells were harvested by centrifugation and disrupted by sonication and the soluble protein fraction was obtained after centrifugation at 20 000*g* for 30 min at 277 K.

We selected PaaA co-expressed with His_6-PaaC for further purification. The cells were lysed in a buffer containing 50 m*M* HEPES pH 7.5, 0.4 *M* NaCl, 5%(*v*/*v*) glycerol, 20 m*M* imidazole, 0.5 m*M* benzamidine, 10 μ*M* leupeptin. The protein supernatant was loaded onto an Ni–NTA column pre-equilibrated with lysis buffer, incubated

Table 1
X-ray data-collection and refinement statistics.

Values in parentheses are for the last resolution shell.

Data set	PaaAC + CoA	PaaAC + 3-hydroxybutyryl-CoA	PaaAC + benzoyl-CoA	PaaAC + phenylacetyl-CoA	PaaAC	PaaAC + Fe^{2+}
Space group	$P4_12_12$	$P4_12_12$	$P4_12_12$	$P4_12_12$	$P2_12_12_1$	$P4_12_12$
Unit-cell parameters (Å)						
a	77.5	77.4	77.6	77.6	110.3	78.4
b	77.5	77.4	77.6	77.6	109.1	78.4
c	300.1	301.4	300.7	300.4	305.9	309.9
Wavelength (Å)	0.97949	0.97949	0.97949	0.97949	0.97949	0.97949
Resolution (Å)	50.0–2.15 (2.23–2.15)	50.0–2.03 (2.10–2.03)	50.0–2.06 (2.13–2.06)	50.0–2.25 (2.33–2.25)	50.0–2.64 (2.73–2.64)	50.0–5.8 (5.9–5.8)
Observed *hkl*	268595	599496	339272	300130	707383	19671
Unique *hkl*	45467	60180	58037	44543	108855	2987
Completeness (%)	88.7 (92.6)	99.3 (94.8)	99.5 (96.6)	99.1 (94.4)	99.9 (99.9)	98.4 (99.3)
Redundancy	5.9 (6.0)	10.0 (4.1)	5.8 (4.2)	6.7 (5.9)	6.5 (4.0)	6.6 (7.0)
R_{merge}	0.055 (0.22)	0.063 (0.382)	0.078 (0.389)	0.098 (0.459)	0.08 (0.547)	0.044 (0.422)
$\langle I/\sigma(I)\rangle$	10.8 (3.8)	11.8 (2.1)	8.6 (2.1)	6.5 (2.0)	10.0 (2.0)	39.6 (3.49)
V_M† (Å^3 Da^{-1})	3.49 or 1.75	3.49 or 1.75	3.49 or 1.75	3.49 or 1.75	2.37	3.68 or 1.84
Solvent content (%)	65.5 or 29.5	65.5 or 29.5	65.5 or 29.5	65.5 or 29.5	48.0	66.6 or 33.2
No. of molecules‡	1 or 2	1 or 2	1 or 2	1 or 2	6	1 or 2

† The Matthews coefficient V_M was calculated using 64 800 Da as the molecular mass of PaaAC. ‡ The number of PaaAC complexes in the asymmetric unit.

for 2 h at room temperature and washed with lysis buffer supplemented with 40 m*M* imidazole. Protein was eluted with a buffer consisting of 50 m*M* HEPES pH 7.5, 50 m*M* NaCl, 5%(*v*/*v*) glycerol, 250 m*M* imidazole (Fig. 2). Immediately after elution, the protein buffer was exchanged into buffer consisting of 50 m*M* HEPES pH 7.5, 50 m*M* NaCl, 5 m*M* DTT using a PD-10 desalting column (GE Healthcare) and the protein sample was loaded onto a pre-equilibrated Superose 12 (GE Healthcare) column for size-exclusion chromatography. To remove excess PaaC, the leading fractions (3 and 4; Fig. 3) from the peak corresponding to the PaaAC complex were pooled together and re-chromatographed on the same column.

2.2. Crystallization

The purified PaaA–His_6-PaaC complex (PaaAC) was concentrated to 8 mg ml^{-1} using an Ultracel-30k centrifugal filter and mixed with 5 m*M* D,L-3-hydroxybutyryl-CoA. Crystallization screening was performed by sitting-drop vapour diffusion at 293 K by mixing 0.4 µl PaaAC solution with 0.3 µl reservoir solution using a Hydra Plus One crystallization robot (Matrix Technologies, Hudson, New Hampshire, USA). Two crystallization conditions, A5 and H10, were identified from initial screening with the Protein Complex suite (Qiagen). Optimized crystals were grown using hanging-drop vapour diffusion in 24-well Linbro plates using equal volumes (1 µl) of protein and one of two different reservoir solutions: (i) 0.1 *M* PIPES pH 6.5, 15%(*v*/*v*) PEG 550 MME or (ii) 0.1 *M* PIPES pH 6.5, 5%(*v*/*v*) 2-propanol, 5%(*v*/*v*) PEG 550 MME. The crystals were 100–200 µm in size in all three dimensions. Crystals from both conditions were cryoprotected by transfer to reservoir solution supplemented with 20%(*v*/*v*) glycerol, picked up in a nylon loop and flash-cooled in an N_2 cold stream at 93 K. Cocrystallization with 5 m*M* CoA, benzoyl-CoA and

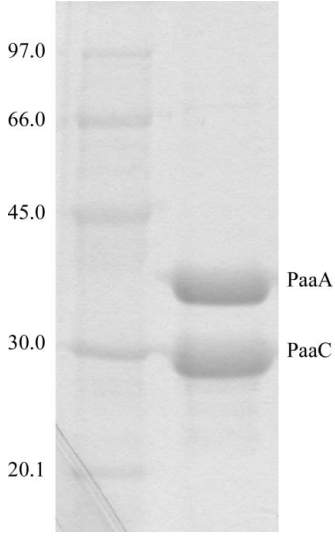


Figure 2
SDS–PAGE analysis of the sample eluted from Ni–NTA agarose. Lane 1, molecular-weight markers (labelled on the left in kDa); lane 2, proteins eluted from the column: the upper band corresponds to PaaA and the lower band corresponds to His_6-PaaC.

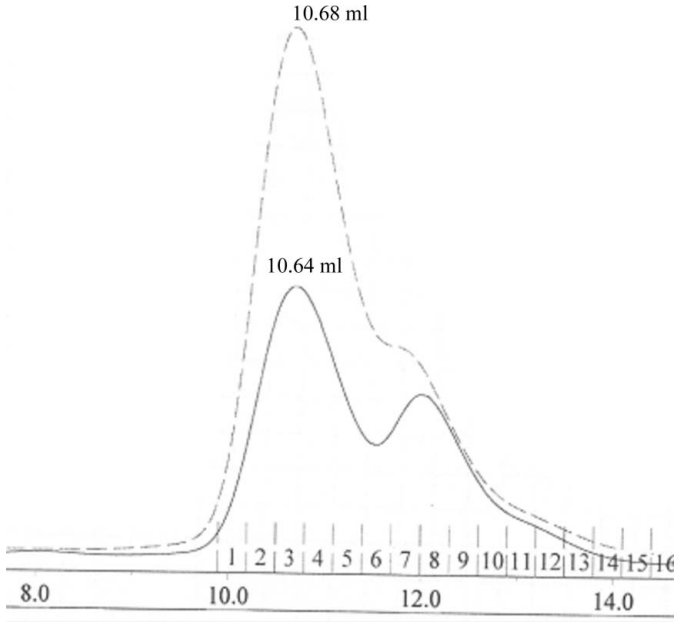


Figure 3
Molecular exclusion chromatography profile on Superose 12 of the sample eluted from the Ni–NTA column. The fraction number and elution volume are shown on the *x* axis. The solid line corresponds to the initial chromatography of the protein eluted from Ni–NTA agarose. The main peak contains PaaAC and the shoulder contains PaaC. The dotted line corresponds to the second chromatography run with the front of the fraction of the main peak and shows good separation of PaaAC and PaaC. Only the fractions corresponding to the front of the peak were pooled and submitted to crystallization.

phenylacetyl-CoA was used to obtain the complexes of PaaAC with these ligands under the same crystallization conditions as described above.

Crystallization conditions for free PaaAC (no substrate bound) were found while attempting to optimize the protein buffer using an in-house extended pH/buffer screen (Jancarik *et al.*, 2004). Crystals of PaaAC were obtained by hanging-drop vapour diffusion from 100 m*M* ADA pH 5.5. The cryosolvent for these crystals consisted of 100 m*M* ADA pH 5.5, 25%(*v*/*v*) MPD. These crystals were 50 µm in size in all three dimensions.

2.3. X-ray data collection

X-ray diffraction data were collected on the CMCF1 beamline at the Canadian Light Source using a MAR 300 CCD detector. Data integration and scaling was performed with *HKL*-2000 (Otwinowski & Minor, 1997). The diffraction limits, crystal space groups and unit-cell parameters of the obtained crystals are presented in Table 1.

3. Results and discussion

3.1. PaaA–PaaC complex formation

PaaA and PaaC cloned into a derivative of pET15b vector could be expressed independently in *E. coli* in high yield, but only PaaC was soluble. Since the hydroxylase components of BMMs are composed of at least α (catalytic) and β (structural) subunits (Notomista *et al.*, 2003), we cloned PaaA into the second molecular cloning site in pRDSFDuet-1 vector (no tag) and co-expressed this protein with N-terminally His-tagged PaaC. Both proteins were expressed in high yield and were soluble.

3.2. Purification and characterization of PaaAC

The co-expressed PaaA and His_6-PaaC were initially loaded onto an Ni–NTA agarose column and both were retained after a thorough wash, indicating the formation of a stable PaaAC complex which could be eluted from the column (Fig. 2). Size-exclusion chromatography on a Superose 12 column showed two partially overlapping peaks with apparent molecular masses of ∼180 and ∼55 kDa. SDS–PAGE showed that the first eluting peak contained the PaaAC complex, while the second peak contained only His_6-PaaC, suggesting that PaaC was expressed in a higher yield exceeding the stoichiometric ratio. To improve the homogeneity of the solution, the leading fractions from the peak containing the PaaAC complex were rerun on a Superose 12 column (Fig. 2) and the central fraction of the protein peak showed a unimodal distribution with an apparent molecular weight of ∼160 kDa as analyzed by dynamic light scattering. Since the calculated molecular weight of the PaaAC complex is ∼65 kDa, there appears to be a heterotetramer or heterohexamer in solution.

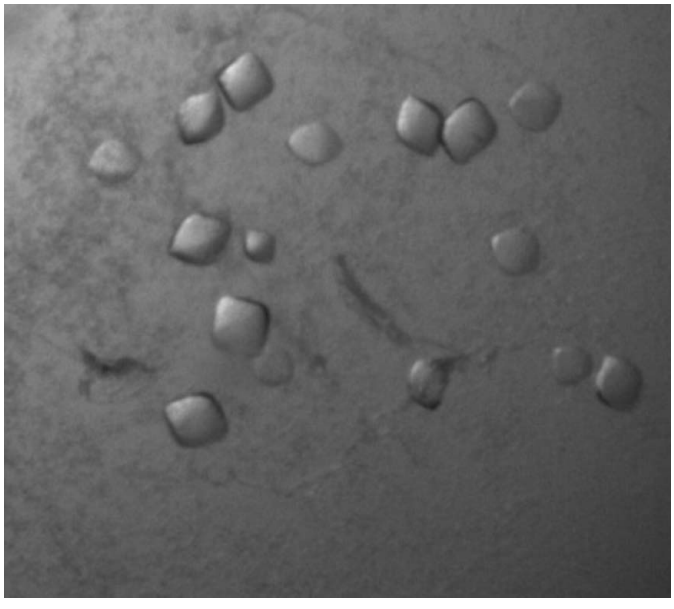


Figure 4
Tetragonal form PaaAC crystals. The shapes of the free PaaAC crystals and of those cocrystallized with substrate or analogues were similar. This figure shows crystals of free PaaAC.

3.3. Crystallization of native PaaAC and its complexes with substrate and substrate analogues

Initial crystals of PaaAC were obtained complexed with D,L-3-hydoxybutyryl-CoA. Crystals appeared within 2–3 d (Fig. 4). SDS–PAGE analysis of redissolved crystals proved that both PaaA and PaaC subunits were present in the crystal (data not shown). Crystals of PaaAC complexed with coenzyme A, benzoyl-CoA and the substrate phenylacetyl-CoA were obtained under similar conditions and were all isomorphous. These crystals diffracted to resolutions in the range 2.03–2.25 Å and a full set of diffraction data was collected on the CMCF1 beamline at the Canadian Light Source (Table 1).

Attempts to cocrystallize the PaaAC complex with its substrate and a metal ion were undertaken with iron (ferrous ammonium sulfate), cobalt ($CoCl_2$) and manganese ($MnCl_2$). Crystals did not grow at iron molar concentrations higher than the concentration of PaaA or at a cobalt concentration higher than 0.25 of that of PaaA. However, crystals could be grown with 5 m*M* manganese. An X-ray fluorescence scan of these crystals showed no evidence that metals were present in the crystals. We therefore performed soaking experiments of PaaAC crystals in solution containing iron. This resulted in no apparent absorption of the iron by the protein at low iron concentrations or led to a significant loss of resolution when higher iron concentrations were used. We collected diffraction data to 5.8 Å resolution from one such crystal soaked with 10 m*M* ferrous ammonium sulfate. This experiment suggests a rather low affinity of PaaA for iron and that some rearrangement of the protein is induced by iron binding.

Continued crystallization screening of ligand-free PaaAC finally resulted in the identification of suitable conditions that were different from those that supported crystal growth of PaaAC complexed with CoA derivatives. These crystals appeared from solution containing only buffer at pH 5.5 with no other precipitating agent (Table 1).

Patterson self-rotation analysis of the $P4_12_12$ crystal form corresponding to ligand-bound protein only indicated the presence of peaks corresponding to crystallographic symmetry. No other peaks that could indicate the presence of NCS in the asymmetric unit were found. A self-rotation function for the $P2_12_12$ crystal (ligand-free protein) indicated peaks that correspond to $P422$ crystal symmetry. Indeed, the *a* and *b* cell dimensions are very close to the *ac* diagonal in the tetragonal space group. However, this data set could not be merged with good statistics in the tetragonal space group, indicating the presence of only pseudo-tetragonal symmetry. No other significant peaks were found.

Determination of all structures is in progress.

The data collection described in this paper was performed at the Canadian Light Source, which is supported by NSERC, NRC, CIHR and the University of Saskatchewan. This work was supported by

grant GSP-48370 from Canadian Institutes of Health Research (to MC). We thank Allan Matte for helpful suggestions and comments on the manuscript.

References

Bailey, L. J., McCoy, J. G., Phillips, G. N. Jr & Fox, B. G. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 19194–19198.
Cafaro, V., Izzo, V., Scognamiglio, R., Notomista, E., Capasso, P., Casbarra, A., Pucci, P. & Di Donato, A. (2004). *Appl. Environ. Microbiol.* **70**, 2211–2219.
Díaz, E., Ferrández, A., Prieto, M. A. & García, J. L. (2001). *Microbiol. Mol. Biol. Rev.* **65**, 523–569.
Fernández, C., Ferrández, A., Miñambres, B., Díaz, E. & García, J. L. (2006). *Appl. Environ. Microbiol.* **72**, 7422–7426.
Ferrández, A., Miñambres, B., García, B., Olivera, E. R., Luengo, J. M., García, J. L. & Díaz, E. (1998). *J. Biol. Chem.* **273**, 25974–25986.
Ismail, W., El-Said, M. M., Wanner, B. L., Datsenko, K. A., Eisenreich, W., Rohdich, F., Bacher, A. & Fuchs, G. (2003). *Eur. J. Biochem.* **270**, 3047–3054.
Jancarik, J., Pufan, R., Hong, C., Kim, S.-H. & Kim, R. (2004). *Acta Cryst.* D**60**, 1670–1673.
Luengo, J. M., García, J. L. & Olivera, E. R. (2001). *Mol. Microbiol.* **39**, 1434–1442.
McCormick, M. S., Sazinsky, M. H., Condon, K. L. & Lippard, S. J. (2006). *J. Am. Chem. Soc.* **128**, 15108–15110.
Moe, L. A., McMartin, L. A. & Fox, B. G. (2006). *Biochemistry*, **45**, 5478–5485.
Nogales, J., Macchi, R., Franchi, F., Barzaghi, D., Fernández, C., García, J. L., Bertoni, G. & Díaz, E. (2007). *Microbiology*, **153**, 357–365.
Notomista, E., Cafaro, V., Bozza, G. & Di Donato, A. (2009). *Appl. Environ. Microbiol.* **75**, 823–836.
Notomista, E., Lahm, A., Di Donato, A. & Tramontano, A. (2003). *J. Mol. Evol.* **56**, 435–445.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Rosenzweig, A. C., Brandstetter, H., Whittington, D. A., Nordlund, P., Lippard, S. J. & Frederick, C. A. (1997). *Proteins*, **29**, 141–152.
Rosenzweig, A. C., Frederick, C. A., Lippard, S. J. & Nordlund, P. (1993). *Nature (London)*, **366**, 537–543.
Sazinsky, M. H., Dunten, P. W., McCormick, M. S., DiDonato, A. & Lippard, S. J. (2006). *Biochemistry*, **45**, 15392–15404.