Charles Packianathan,^a Jitesh K. Pillai,^a Ahmed Riaz,^a Palani Kandavelu,^b Banumathi Sankaran^c and Barry P. Rosen^{a*}

^aDepartment of Cellular Biology and Pharmacology, Herbert Wertheim College of Medicine, Florida International University, Miami, FL 33199, USA, ^bSER-CAT and the Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA, and ^cBerkeley Center for Structural Biology, Lawrence Berkeley Laboratory, 1 Cyclotron Road, BLDG 6R2100, Berkeley, CA 94720, USA

Correspondence e-mail: brosen@fiu.edu

Received 12 July 2014

Accepted 13 August 2014

Crystallization and preliminary X-ray crystallographic studies of CrArsM, an arsenic(III) *S*-adenosylmethionine methyltransferase from *Chlamydomonas reinhardtii*

Arsenic is one of the most toxic environmental substances. Arsenic is ubiquitous in water, soil and food, and ranks first on the Environmental Protection Agency's Superfund Priority List of Hazardous Substances. Arsenic(III) *S*-adenosylmethionine methyltransferases (AS3MT in animals and ArsM in microbes) are key enzymes of arsenic biotransformation, catalyzing the methylation of inorganic arsenite to give methyl, dimethyl and trimethyl products. Arsenic methyltransferases are found in members of every kingdom from bacteria to humans (EC 2.1.1.137). In the human liver, hAS3MT converts inorganic arsenic into more toxic and carcinogenic forms. CrArsM, an ortholog of hAS3MT from the eukaryotic green alga *Chlamydomonas reinhardtii*, was purified by chemically synthesizing the gene and expressing it in *Escherichia coli*. Synthetic purified CrArsM was crystallized in an unliganded form. Crystals were obtained by the hanging-drop vapor-diffusion method. The crystals belonged to space group *R*3:*H*, with unit-cell parameters $a = b = 157.8$, $c = 95.4$ Å, $\gamma = 120^\circ$ and two molecules in the asymmetric unit. Complete data sets were collected and processed to a resolution of 2.40 Å.

1. Introduction

While there are many toxins present in the environment, none is as pervasive as arsenic. In fact, the 2013 Environmental Protection Agency's Comprehensive Environmental Response, Compensation and Liability Act (CERCLA) Priority List of Hazardous Substances ranks arsenic first based on a combination of its frequency, toxicity and potential for human exposure (<http://www.atsdr.cdc.gov/spl/>). The widespread impact of arsenic toxicity can be attributed to the ubiquity of arsenic in the environment (Zhu *et al.*, 2014). Natural sources of arsenic include soil, water and food supplies. Arsenic consumption from foods such as meats, vegetables and fish has been linked to skin, bladder and lung cancer (Oberoi *et al.*, 2014). Genes for arsenic(III) *S*-adenosylmethionine (SAM) methyltransferases are widespread in the genomes of bacteria, archaea, fungi, lower plants and animals, including humans. In microbes, the ArsM (arsenite *S*-adenosylmethyltransferase) orthologs detoxify arsenic (Qin *et al.*, 2006, 2009). In contrast, the health effects of arsenic may be attributed to the methylation of inorganic arsenic by the human ortholog, hAS3MT, which catalyzes the biotransformation of arsenic(III), primarily in the liver, to methylated products that are more toxic and carcinogenic in humans (Thomas & Rosen, 2013; Thomas *et al.*, 2004, 2007).

Structural data from the CmArsM enzyme of the eukaryotic acidothermophilic red alga *Cyanidioschyzon merolae* have shed light on the functional aspects of human hAS3MT (Ajeev *et al.*, 2012). However, a more complete understanding of the relationship of this family of enzymes to arsenic carcinogenesis would be aided by additional crystal structures from other orthologs. Comparison of the structures of orthologs will illuminate which domains and residues may be functionally significant. CrArsM from the eukaryotic green alga *Chlamydomonas reinhardtii* (JX480492.2) catalyzes arsenic methylation and volatilization, leading to arsenic resistance (Chen *et al.*, 2013). The gene for CrArsM, a 379-residue enzyme (41.6 kDa), was synthesized and expressed in *Escherichia coli*. Here, we report the crystallization, X-ray data collection and preliminary crystallographic analysis of CrArsM.

2. Materials and methods

2.1. Reagents

All chemicals were obtained from Sigma–Aldrich (St Louis, Missouri, USA) unless otherwise mentioned.

2.2. Cloning, expression and purification of CrArsM

The native *CrArsM* gene (accession No. JX480492.2) was expressed and purified as described by Chen *et al.* (2013). However, we were unable to crystallize the purified protein. For this reason, a *CrArsM* gene was designed based on the native gene, with *NcoI* and *XhoI* restriction sites flanking the 5' and 3' ends of the gene, respectively. The gene was chemically synthesized by GenScript USA Inc. (New Jersey) and inserted into vector plasmid pUC57 (Thermo Fisher Scientific, Waltham, Massachusetts, USA). The sequence of the gene was optimized by GenScript using their *OptimumGene* algorithm, which improves expression by analyzing and upgrading the codon-adaptation index (Sharp & Li, 1987), modifying the GC content and disrupting stem-loop structures that affect the ribosome binding and the stability of the mRNA (Supplementary Fig. S1a¹). Native CrArsM has 17 cysteine residues, of which four are conserved in orthologs. Multiple cysteine residues may interfere with crystallization, and alanine substitutions can produce a more tractable protein with improved storage properties (Hari *et al.*, 2010). For this reason, the four conserved cysteine residues at the conserved positions 46, 74, 170 and 220 were retained, while the other 13 cysteine residues were changed to alanine residues. Trp332 and Tyr72 of the native CrArsM protein were mutated to tyrosine and tryptophan residues, respectively, in the synthetic CrArsM for future measurements of ligand binding by changes in intrinsic tryptophan fluorescence. Otherwise, the sequence of the synthetic protein was the same as native CrArsM (Supplementary Fig. S1b). The gene was cloned into vector plasmid pET-29b(+) (EMD Biosciences, California, USA) as an *NcoI/XhoI* digest, generating plasmid pET29-CrArsM in which the *CrArsM* gene is under the control of the T7 promoter and is expressed with a six-histidine tag at the C-terminus.

Cells of *E. coli* strain BL21(DE3) pET29-CrArsM were grown at 37°C in Luria–Bertani medium (Sambrook *et al.*, 1989) containing 50 mg ml⁻¹ kanamycin with shaking at 240 rev min⁻¹. When the absorbance at 600 nm of the culture reached 0.6, 0.1 mM isopropyl β -D-1-thiogalactopyranoside (IPTG; Research Products International, Mount Prospect, Illinois, USA) was added to induce expression of CrArsM. The cells were grown for a further 4 h, harvested by centrifugation (3500g) for 30 min, washed once with buffer A [50 mM MOPS pH 7.5, 0.5 M NaCl containing 20% (w/v) glycerol, 20 mM imidazole and 10 mM β -mercaptoethanol] and suspended in 5 ml buffer A per gram of wet cells. The cells were lysed by two passes through a French pressure cell press at 138 MPa and immediately treated with 2.5 μ l of the protease inhibitor diisopropyl fluorophosphate per gram of wet cells. Membranes and unbroken cells were removed by centrifugation at 150 000g for 1 h (Qin *et al.*, 2006) and the supernatant solution was loaded at a flow rate of 0.5 ml min⁻¹ onto a 5 ml HisTrap HP column (GE Healthcare) pre-equilibrated with buffer A using an ÄKTA FPLC purification system (GE Healthcare). The column was washed with five column volumes of buffer A followed by elution with five column volumes of buffer A containing 0.2 M imidazole. Purified synthetic CrArsM was identified by SDS–PAGE (Laemmli, 1970; Fig. 1, lane 2). The migration of the band corresponded to the mass of CrArsM. Fractions containing

CrArsM were concentrated by centrifugation using a 10 kDa cutoff Amicon Ultrafilter (EMD Millipore Corporation, Billerica, Massachusetts, USA) and were buffer-exchanged with buffer C consisting of 50 mM MOPS pH 7.5, 0.5 M NaCl, 5 mM dithiothreitol (DTT), 2.5 mM ethylenediamine tetraacetate, 10% (w/v) glycerol for storage at –80°C until use. The yield was 23 mg of purified protein from 1 l of cell culture.

3. Crystallization

The purified CrArsM was buffer-exchanged into a buffer consisting of 50 mM MOPS pH 7.0, 0.5 M NaCl, 5 mM DTT. Initial crystallization trials were performed by the hanging-drop vapor-diffusion method using crystallization screens from Hampton Research (Aliso Viejo, California, USA), Emerald Bio (Bainbridge Island, Washington, USA) and Jena Biosciences GmbH (Jena, Germany).

At 18 and 20°C, crystallization trials produced crystalline aggregates upon mixing 2.0 μ l each of protein and reservoir (20% PEG 3350, 0.2 M calcium acetate, 0.1 M Tris–HCl pH 7.0) solutions. Reduction of the PEG concentration from 20 to 18% led to the formation of needle-like crystals that were used for seeding experiments. The needles were transferred from the hanging drop to eppendorf tubes containing seed beads (Hampton Research) and 0.1 ml well solution. Sample preparation for seeding was as described previously (Marapakala *et al.*, 2010). Diffraction-quality crystals (Fig. 2) were obtained at room temperature using Linbro 24-well plates (Hampton Research). The wells, sealed with a cover slip, held 0.5 μ l of seed stock (1:100 ratio) in drops consisting of equal volumes of protein and reservoir (10% PEG 8000, 2.0 M NaCl) solutions. The crystals from a single drop were washed several times in 50% glycerol and dissolved in water for analysis by SDS–PAGE (Fig. 1, lane 1).

4. Data collection and processing

Crystals were transferred to a cryoprotectant solution (10% PEG 8000, 2.0 M NaCl, 20% 2-methyl-2,4-pentanediol) and flash-cooled in liquid nitrogen for data collection. Initial crystals were screened at

Figure 1

CrArsM was purified by Ni–NTA chromatography as described in §2.1. The purity of the protein was analyzed by SDS–PAGE stained with Coomassie Blue. Lane 1, a single crystal of CrArsM dissolved in water. Lane 2, purified and concentrated CrArsM from Ni–NTA chromatography. Lane 3, molecular-weight markers (labelled in kDa). The position of CrArsM is indicated by an arrow.

¹ Supporting information has been deposited in the IUCr electronic archive (Reference: RL5080).

the Advanced Light Source, Berkeley Center for Structural Biology. Subsequent data sets were collected at the Southeast Regional Collaborative Access Team (SER-CAT) facility at the Advanced Photon Source (APS), Argonne, Illinois, USA. Data were obtained from 360 image frames with a 1° rotation angle about φ using a MAR 225 CCD detector under standard cryogenic conditions (-173°C) on synchrotron beamline 22-BM with a crystal-to-detector distance of 240 mm. The data sets were indexed, integrated and scaled with the *HKL-2000* suite (Otwinowski & Minor, 1997). The data-processing statistics are shown in Table 1.

5. Results and discussion

Needle-shaped crystals of CrArsM initially diffracted to a resolution in the range 7.0–4.5 Å, but improved with seeding. Diffraction-quality

Figure 2

Crystals of CrArsM grown by hanging-drop vapor diffusion. The crystals were grown at room temperature and belonged to the trigonal space group *R3:H* with approximate dimensions $0.3 \times 0.3 \times 0.2$ mm.

Figure 3

Diffraction images (1° oscillation) were collected from CrArsM crystals at APS (Advanced Photon Source) using a MAR 300 detector at -173°C . The edge of the image corresponds to 2.40 Å resolution.

Table 1

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	22-BM, APS
Wavelength (Å)	1.00
Temperature (K)	100
Detector	MAR 225 CCD
Crystal-to-detector distance (mm)	240
Rotation range per image ($^\circ$)	1
Total rotation range ($^\circ$)	360
Exposure time per image (s)	3.9
Space group	<i>R3:H</i>
<i>a</i> , <i>b</i> , <i>c</i> (Å)	157.8, 157.8, 95.4
α , β , γ ($^\circ$)	90.0, 90.0, 120.0
Mosaicity ($^\circ$)	0.2
Resolution range (Å)	50–2.40 (2.49–2.40)
Total No. of reflections	404464 (40054)
No. of unique reflections	34690 (3483)
Completeness (%)	100 (100)
Multiplicity	11.7 (11.5)
$\langle I/\sigma(I) \rangle$	34.1 (6.8)
$R_{\text{r.i.m.}}$ (%)	8.3 (42.8)
$R_{\text{merge}}^\dagger$ (%)	8.0 (40.9)
Matthews coefficient ($\text{\AA}^3 \text{Da}^{-1}$)	2.73
Overall <i>B</i> factor from Wilson plot ($\text{\AA}^2$)	45.2

$^\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl) \rangle$ is the average intensity over symmetry-equivalent measurements.

crystals formed overnight with dimensions $0.3 \times 0.3 \times 0.2$ mm. The crystals were stable for at least one week, after which they began to dissolve. Complete data sets were collected to 2.40 Å resolution from a single crystal of CrArsM (Fig. 3). The crystals belonged to space group *R3:H*, with unit-cell parameters $a = b = 157.8$, $c = 95.4$ Å, $\gamma = 120^\circ$. Evaluation of crystal-packing parameters indicated that the lattice can accommodate two molecules in the asymmetric unit, with a Matthews coefficient of $2.73 \text{ \AA}^3 \text{Da}^{-1}$ and a solvent content of 54.9%. The sequence similarity between CmArsM and CrArsM is 44.0%, so we attempted to use the native X-ray crystal structure of CmArsM (PDB entry 4fs8; Ajees *et al.*, 2012) to solve the structure of CrArsM by molecular replacement, without success. To solve the phase problem, attempts to obtain crystals of CrArsM co-crystallized or soaked with heavy metals are in progress.

This work was supported by NIH grant R37 GM55425. This project utilized the Southeast Regional Collaborative Access Team (SER-CAT) 22-BM beamline of the Advanced Photon Source, Argonne National Laboratory. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under contract No.W-31-109-Eng-38. The Berkeley Center for Structural Biology is supported in part by the National Institutes of Health, National Institute of General Medical Sciences and the Howard Hughes Medical Institute. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences of the US Department of Energy under Contract No. DE-AC02-05CH11231.

References

- Ajees, A. A., Marapakala, K., Packianathan, C., Sankaran, B. & Rosen, B. P. (2012). *Biochemistry*, **51**, 5476–5485.
- Chen, J., Qin, J., Zhu, Y.-G., de Lorenzo, V. & Rosen, B. P. (2013). *Appl. Environ. Microbiol.* **79**, 4493–4495.
- Hari, S. B., Byeon, C., Lavinder, J. J. & Magliery, T. J. (2010). *Protein Sci.* **19**, 670–679.
- Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
- Marapakala, K., Ajees, A. A., Qin, J., Sankaran, B. & Rosen, B. P. (2010). *Acta Cryst.* **F66**, 1050–1052.
- Oberoi, S., Barchowsky, A. & Wu, F. (2014). *Cancer Epidemiol. Biomarkers Prev.* **23**, 1187–1194.

- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Qin, J., Lehr, C. R., Yuan, C., Le, X. C., McDermott, T. R. & Rosen, B. P. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 5213–5217.
- Qin, J., Rosen, B. P., Zhang, Y., Wang, G., Franke, S. & Rensing, C. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 2075–2080.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning. A Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press.
- Sharp, P. M. & Li, W.-H. (1987). *Nucleic Acids Res.* **15**, 1281–1295.
- Thomas, D. J., Li, J., Waters, S. B., Xing, W., Adair, B. M., Drobná, Z., Devesa, V. & Styblo, M. (2007). *Exp. Biol. Med. (Maywood)*, **232**, 3–13.
- Thomas, D. J. & Rosen, B. P. (2013). *Encyclopedia of Metalloproteins*, edited by R. H. Kretsinger, V. N. Uversky & E. A. Permyakov, pp. 140–145. New York: Springer.
- Thomas, D. J., Waters, S. B. & Styblo, M. (2004). *Toxicol. Appl. Pharmacol.* **198**, 319–326.
- Zhu, Y.-G., Yoshinaga, M., Zhao, F.-J. & Rosen, B. P. (2014). *Annu. Rev. Earth Planet. Sci.* **42**, 443–467.