Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray analysis of the catalytic subunit of the ATP-dependent arsenite pump encoded by the *Escherichia coli* plasmid R773

Tongqing Zhou, Barry P. Rosen and Domenico L. Gatti*

Department of Biochemistry and Molecular Biology, Wayne State University School of Medicine, 540 E. Canfield Avenue, Detroit, MI 48201, USA

Correspondence e-mail: mimo@david.med.wayne.edu

Received 16 October 1998
Accepted 5 January 1999

The arsenical resistance (*ars*) operon of the *Escherichia coli* plasmid R773 encodes a system for the active extrusion from cells of the toxic oxyanions arsenite ($As^{III}O_2^-$) and antimonite ($Sb^{III}O_2^-$) *via* an ATP-driven pump. The *arsA* and *arsB* genes of the operon encode the catalytic subunit (ATPase) and the membrane subunit of the pump, respectively. The *arsC* gene codes for a reductase that converts arsenate ($As^{V}O_4^{3-}$) to arsenite, thus extending bacterial resistance to the pentavalent state of arsenic. Crystals diffracting beyond 2.0 Å were obtained for the catalytic subunit of the pump (ArsA). These crystals belong to space group *I*222, with unit-cell parameters $a \simeq 73$, $b \simeq 76$, $c \simeq 223$ Å. A single molecule of ArsA, composed of two homologous halves, occupies the asymmetric unit of the *I*222 crystals with a predicted solvent content of 46%. Self-rotation function analysis suggests, however, that ArsA adopts a molecular packing corresponding to point group 422. One possible explanation of this result is that the two homologous halves of ArsA are related by a twofold axis of local symmetry and that the two halves of a 'pseudo'-tetramer are related by a crystallographic twofold axis.

1. Introduction

The detoxifying system encoded by the *ars* operon of plasmid R773 in *E. coli* confers resistance against oxyanions of arsenic (arsenite and arsenate) and antimony (antimonite) (Chen *et al.*, 1986; Xu *et al.*, 1998). The *ars* operon codes for two regulatory (ArsR and ArsD) and three structural (ArsA, ArsB and ArsC) proteins. The *arsA* and *arsB* gene products form a membrane-bound anion pump. ArsC reduces arsenate to arsenite, which is then extruded by the Ars pump.

ArsA is a 63 kDa peripheral membrane protein which catalyzes As^{III}/Sb^{III}-stimulated ATP hydrolysis (Xu *et al.*, 1998). An interesting feature of ArsA is that the amino- and carboxyl-terminal halves of the protein show 20% identity with each other. These two homologous halves (named A1 and A2) each contain a nucleotide-binding site consensus sequence (P-loop; Kaur & Rosen, 1994). Notably, both the A1 and A2 sites bind nucleotides independently (Karkaria & Rosen, 1991), and mutations in either site can eliminate the ATPase activity of the protein (Kaur & Rosen, 1992*a*). Genetic and biochemical studies have shown that both the A1 and A2 nucleotide-binding sites are required for actions associated with arsenic resistance, namely arsenite transport and ATPase activity (Kaur & Rosen, 1992*b*). Genetic complementation and biochemical reconstitution experiments also suggest that the formation of a catalytic site requires interaction of the A1 and A2 sites (Kaur & Rosen, 1993). The ATPase activity of ArsA is activated allosterically by As^{III} and Sb^{III} (Rosen *et al.*, 1995). On the basis of experiments involving chemical modification with sulfhydryl reagents and site-directed mutagenesis of the cysteine residues, it has been proposed that this allosteric activation is mediated by the formation of a three-coordinate complex between Sb^{III} or As^{III} and the thiolates of cysteines 113, 172 and 422, with the three thiol groups forming the base of a pyramid (Bhattacharjee *et al.*, 1995). A similar coordination geometry is observed in the complex of dithiothreitol and As^{III} (Cruse & James, 1972).

ArsB is a 45 kDa integral membrane protein to which ArsA binds to form the ArsA–ArsB anion pump (Dey, Dou & Rosen, 1994). The function of the ArsA–ArsB complex has been studied both *in vivo* and *in vitro* (Dey, Dou, Tisa *et al.*, 1994; Dey & Rosen, 1995). These studies have shown that the ArsA–ArsB pump extrudes AsO_2^- in a manner which is independent of electrochemical energy from respiration (Dey & Rosen, 1995). However, in mutants which are defective for ArsA (the ATPase moiety), arsenite extrusion *via* ArsB appears to be dependent on the membrane electrochemical gradient (Dey & Rosen, 1995). Thus, ArsB has the ability to function either as a subunit of the ArsA–ArsB ATP-dependent

Table 1
Crystal parameters and data-collection statistics.

Crystal type	ArsA native	Selenomet-ArsA	ArsA + Et_3PbOAc
a (Å)	73.47	73.77	74.156
b (Å)	75.71	76.08	76.272
c (Å)	222.67	222.89	222.89
Space group	$I222$	$I222$	$P222$
Data-collection device†	R-AXIS IV	R-AXIS IV	R-AXIS IV
Number of crystals	1	1	1
Resolution (Å)	50–2.3	50–2.6	50–2.8
Number of observations	372828	267755	356579
Number of unique reflections	28277	19743	31866
Completeness (%)	99.86	99.76	99.42
$\langle I/\sigma I\rangle$ (highest resolution shell)	2.65 (2.38–2.30 Å)	4.29 (2.69–2.60 Å)	4.12 (2.90–2.80 Å)
Reflections with $I/\sigma I < 2.0$ (%)	17.3	13.4	13.8
R_{sym}‡	0.118	0.085	0.084

† Data sets were collected with crystals flash-frozen at 100 K. ‡ $R_{sym} = \sum_h \sum_i |I(h,i) - \langle I(h)\rangle| / \sum_h \sum_i I(h,i)$, where $\langle(h)\rangle$ is the mean intensity of symmetry-equivalent reflection (h).

pump or independently as a passive translocator (Dey & Rosen, 1995).

2. Results and discussion

2.1. Protein purification and crystallization

Initial attempts to crystallize ArsA were hindered by the heterogeneity of the purified protein, which often contained minor amounts of smaller peptides in the range 40–50 kDa. Western blot analysis provided evidence that the smaller contaminants were fragments of ArsA originating either from proteolytic degradation or from premature termination of translation. In order to discriminate between these two possibilities, N-terminal and C-terminal His-tagged derivatives of ArsA were engineered and purified according to a previously described protocol (Zhou & Rosen, 1997). Both N-terminal and C-terminal His-tagged ArsA bind to a ProBond Ni^{2+} affinity resin and can be eluted in nearly pure form. The smaller ArsA peptide fragments were absent from the preparation of the C-terminal His-tagged protein, but were observed in affinity-column eluates of ArsA tagged at the N-terminus. These results suggest that ArsA can undergo proteolytic degradation at the C-terminus and that the C-terminal His-tag probably protects against proteolysis. The concentrated eluate of C-terminal His-tagged ArsA from the ProBond Ni^{2+} affinity column was then applied to a Sephacryl S-200 column equilibrated with a buffer containing 250 mM NaCl, 1 mM EDTA, 1 mM DTT, 25 mM Bis–Tris–propane (pH 8.0) and 20%(v/v) glycerol. Pooled fractions containing >97% pure ArsA were brought to a protein concentration of 25–30 mg ml^{-1} and stored in liquid nitrogen in small aliquots.

For crystallization, a single frozen aliquot (100 µl) of the ArsA protein was thawed and passed through a Sephadex G-25 spin column (1 ml bed volume) equilibrated with 10 mM Bis–Tris–propane (pH 8.0). The column was spun at 1800 rev min^{-1} in a Sorvall TC6 centrifuge with swing-out rotor for 2 min to remove excess buffer. 75–100 µl of the protein solution was then loaded and the column was spun again for 2 min at the same speed. The eluate was collected and used immediately for crystallization. C-terminal His-tagged ArsA was crystallized by vapor diffusion in hanging drops at 303 K. Drops were prepared by mixing equal amounts of a solution containing 20 mg ml^{-1} ArsA protein, 2 mM ADP, 1 mM $MgCl_2$, 2 mM $NaSbO_2$, 1 mM $CdCl_2$, 10 mM Bis–Tris–propane (pH 8.0) with reservoir solution containing 4%(w/v) PEG 3000. Crystals appeared after two weeks and continued growing for an additional two weeks. At this point the crystallization trays were slowly equilibrated at progressively lower temperatures over a period of 2 d until the crystals became stable at room temperature. ArsA crystals were then harvested from the crystallization tray in a holding solution consisting of 20%(w/v) PEG 3000, 5 mM ADP, 2.5 mM $MgCl_2$, 4 mM $NaSbO_2$, 2 mM $CdCl_2$, 100 mM Bis–Tris–propane (pH 8.0) and 15%(v/v) glycerol as cryoprotectant. After a minimum of 30 min incubation in the cryoprotectant solution, ArsA crystals could be flash frozen under a nitrogen stream at 100 K. Under these conditions diffraction was observed to at least 1.9 Å with a conventional X-ray source.

2.2. X-ray diffraction analysis

Native ArsA crystals belong to the orthorhombic crystal system and possess a body-centered unit cell with unit-cell parameters $a \simeq 73$, $b \simeq 76$, $c \simeq 223$ Å. Native data sets have been collected at 100 K. Statistical parameters representative of the quality of one of the sets collected at the resolution limit of 2.3 Å are shown in Table 1.

There are two possible space groups for the unit cell of ArsA crystals: $I222$ and $I2_12_12_1$. The same systematic absences are associated with both space groups, which therefore cannot be distinguished on the basis of the diffraction pattern. However, a distinction can be based on the analysis of the Harker sections. In $I222$, the patterns in the three Harker sections are projections of one another. In $I2_12_12_1$, the corresponding patterns are also projections of one another, but with origins shifted by 1/2, 1/2, 1/2 (Buerger, 1959). The ambiguity with $I2_12_12_1$ was resolved by identifying the position of heavy atoms in isomorphous or anomalous difference Patterson maps (see below).

ArsA was expressed in *E. coli* under conditions which favour incorporation of selenomethionine (Hendrickson *et al.*, 1990) in place of methionine [one monomer of ArsA (64 kDa) contains 13 methionines]. Incorporation of selenomethionine, as evaluated by electrospray mass spectrometry, was greater than 97%. Crystals of seleno-

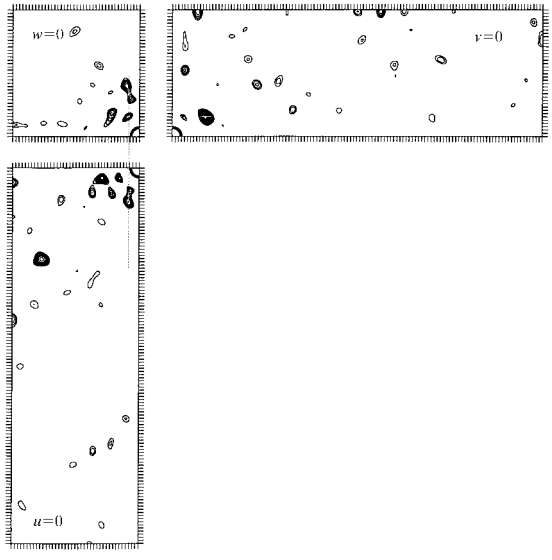

Figure 1
Anomalous difference Patterson from a $P222$ crystal of ArsA derivatized with triethyl lead acetate. Dotted lines cross Harker peaks originating from the main heavy-atom site, as identified and refined using *CNS* (Brünger *et al.*, 1998).

methionine ArsA are isomorphous to crystals of native ArsA. A data set of high quality (Table 1) was collected at the Cu $K\alpha$ wavelength from a selenomethionine crystal. The isomorphous differences between the native and selenomet-ArsA data were used to determine the positions of Se atoms using the program *SOLVE* (Terwilliger, 1998). Using the space group *I*222, *SOLVE* identified 11 potential sites out of the 13 expected, with good phasing statistics ($\langle$FOM$\rangle \simeq 0.50$, $R_{\rm centric} < 0.5$, r.m.s.(FH)/r.m.s.(E) > 1.5 in all resolution shells). When the space group $I2_12_12_1$ was used in the search, *SOLVE* identified only two potential sites with poor phasing statistics (mean FOM < 0.1).

Additional evidence that the true space group of these ArsA crystals is *I*222 arises from the observation that soaking with triethyl lead acetate converts the cell from centered to primitive (Table 1). Conversion between *I*222 and *P*222 (or between $I2_12_12_1$ and $P2_12_12_1$) is possible *via* transformation of a crystallographic axis of symmetry into a local axis of symmetry; cross-conversions (*e.g.* $I2_12_12_1$ to *P*222) by simple transformation of crystallographic axes to local axes (*i.e.* without extensive rearrangement of the molecular packing) are forbidden. A Patterson map generated with the anomalous differences from the triethyl lead acetate data set shows one strong peak at 9σ in the $v = 0$ Harker section and corresponding peaks of smaller height in the other sections (Fig. 1). Peaks of much smaller intensity (originating from pseudo-centering) are observed in the Harker sections at $u = 1/2$, $v = 1/2$, $w = 1/2$, strongly suggesting that the correct space group for this derivatized crystal is *P*222, and that the space group for the native ArsA crystals is *I*222.

This preliminary information also suggests that determination of the ArsA

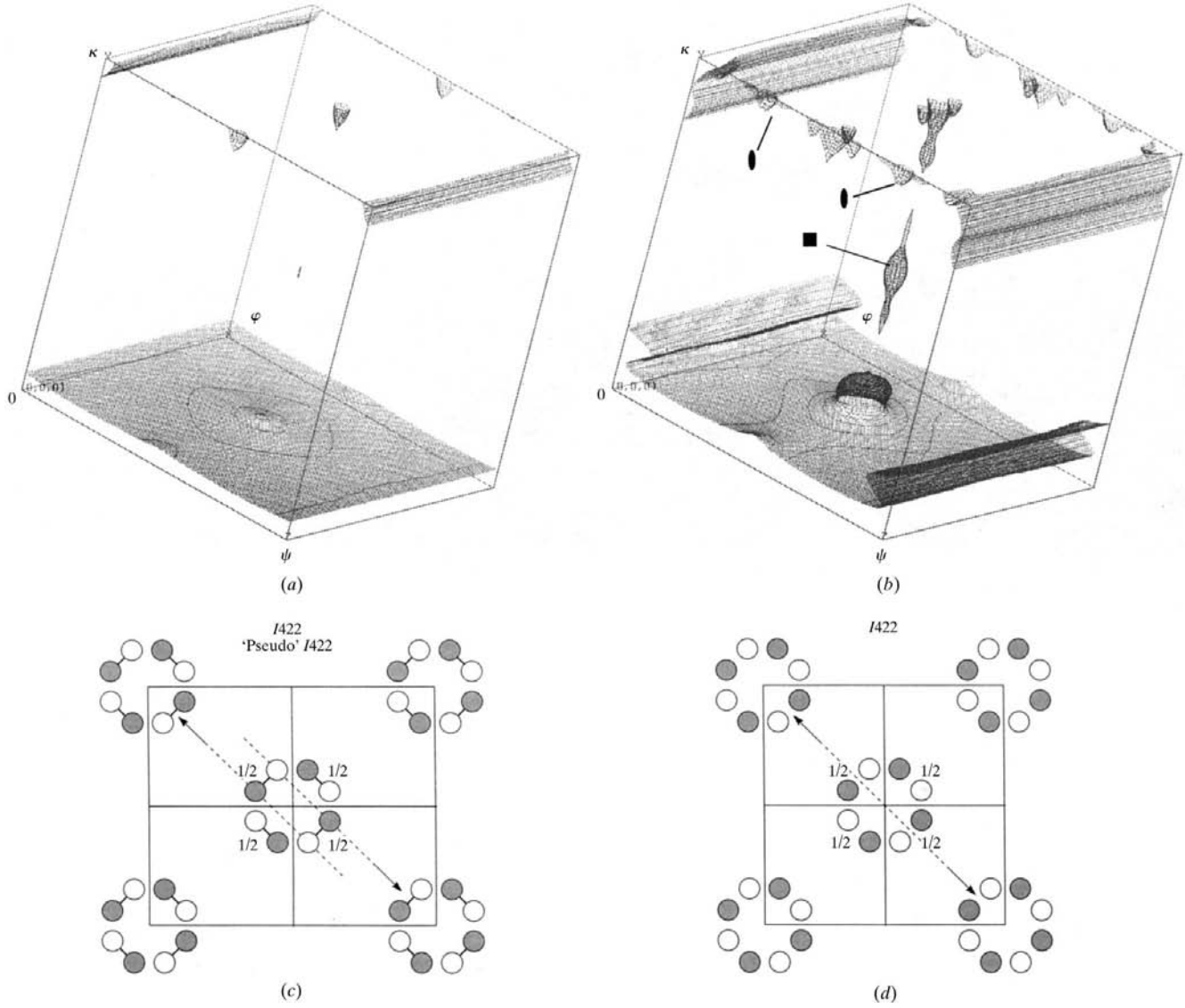

Figure 2
(*a*) Self-rotation function analysis in the range $\varphi = \psi = \kappa = 0$–180° (2$\sigma$ contour): only peaks corresponding to the crystallographic twofold axes are visible. (*b*) As (*a*), but contoured at 1σ to show the presence (marked by arrows) of additional fourfold axis (■) ($\varphi = 90$, $\psi = 90$, $\kappa = 90°$) and twofold axes (◆) ($\varphi = 0$ or 180, $\psi = 45$ or 135, $\kappa = 180°$) corresponding to point group 422. (*c*) *I*222 space-group diagram (view down the *c* axis). A molecule of ArsA is indicated by two circles (one solid and one open) connected by a line, which represent the two homologous domains of the protein related by a twofold axis of symmetry. Owing to the low degree of identity (20%), the symmetry operation relating the two domains is not a true twofold rotation, and the real space group is *I*222 with pseudo-*I*422 symmetry. Also, since the *a* and *b* axes are of different length, the pseudo-twofold axes relating two domains of ArsA are not co-linear in different parts of the unit cell. (*d*) *I*422 space-group diagram.

structure will be possible with MAD phasing (Hendrickson, 1991) of data sets collected at/around the Se *K* edge and the Pb L_{III} edge either from selenomet-ArsA crystals (space group *I*222) or from the triethyl lead acetate derivatized crystals (space group *P*222).

2.3. Non-crystallographic symmetry

Crystals of ArsA are obtained and maintained in the presence of arsenite or antimonite. Under these conditions, ArsA is known to form dimers and/or tetramers (Rosen *et al.*, 1995). However, the unit-cell dimensions of the *I*222 crystals only allow the presence of one molecule of ArsA in the asymmetric unit: under such conditions the expected solvent content is ~46%. Therefore, if a dimer is the biological unit, in the crystal it must adopt a twofold axis of crystallographic symmetry. This conclusion is supported by a self-rotation function analysis carried out using *CNS* (Brünger *et al.*, 1998) over a range of ψ = 0–180, φ = 0–180, κ = 0–180° (displayed as a three-dimensional map contoured at 2σ in Fig. 2*a*), which does not reveal the presence of non-crystallographic axes of symmetry. However, it is noteworthy that ArsA may have additional internal symmetry originating from the two homologous halves of the molecule. This symmetry could appear in the self-rotation function analysis as smaller peaks with respect to the crystallographic symmetry. Indeed, contouring of the self-rotation function map at 1.0σ (Fig. 2*b*) reveals the existence of a fourfold rotation axis of non-crystallographic symmetry oriented along the *c* axis of the unit cell and of twofold axes in the *ab* plane at 45° from the crystallographic twofold axis. This particular orientation of symmetry axes suggests the ArsA adopts a molecular packing corresponding to point group 422 in the crystal. One possible explanation of this result is that the two homologous halves of ArsA are related by a twofold axis of local symmetry and that the two halves of a 'pseudo'-tetramer are related by a crystallographic twofold axis. This situation is explained schematically in Figs. 2(*c*) and 2(*d*); if the symmetry operation relating the two homologous halves of ArsA had not been local, but crystallographic, the space group of ArsA crystals would be *I*422 instead of *I*222.

We would like to thank Drs A. Brunger and R. W. Grosse-Kunstleve for providing us with the developmental release of *CNS* version 0.4. This work was supported by NIH Grants AI43918 (to DG) and GM52216 (to BR).

References

Bhattacharjee, H., Ksenzenko, M. Y., Li, J. & Rosen, B. P. (1995). *J. Biol. Chem.* **270**, 1–6.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Buerger, M. J. (1959). *Vector Space*, pp. 167–168. New York: John Wiley.

Chen, C. M., Misra, T., Silver, S. & Rosen, B. P. (1986). *J. Biol. Chem.* **261**, 15030–15038.

Cruse, W. B. T. & James, M. N. G. (1972). *Acta Cryst.* B**28**, 1325–1331.

Dey, S., Dou, D. & Rosen, B. P. (1994). *J. Biol. Chem.* **269**, 25442–25446.

Dey, S., Dou, D., Tisa, L. S. & Rosen, B. P. (1994). *Arch. Biochem. Biophys.* **311**, 418–424.

Dey, S. & Rosen, B. P. (1995). *J. Bacteriol.* **177**, 385–389.

Hendrickson, W. A. (1991). *Science*, **254**, 51–58.

Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.

Karkaria, C. E. & Rosen, B. P. (1991). *Arch. Biochem. Biophys.* **288**, 107–111.

Kaur, P. & Rosen, B. P. (1992*a*). *J. Biol. Chem.* **267**, 19272–19277.

Kaur, P. & Rosen, B. P. (1992*b*). *Plasmid*, **27**, 29–40.

Kaur, P. & Rosen, B. P. (1993). *J. Bacteriol.* **175**, 351–357.

Kaur, P. & Rosen, B. P. (1994). *Biochem. J.* **33**, 6456–6461.

Rosen, B. P., Bhattacharjee, H. & Shi, W. P. (1995). *J. Bioenerg. Biomembr.* **27**, 85–91.

Terwilliger, T. C. (1998). *SOLVE* 1.0. University of California (Los Alamos National Laboratories) and the US Department of Energy, USA.

Xu, C., Zhou, T., Kuroda, M. & Rosen, B. P. (1998). *J. Biochem.* **123**, 16–23.

Zhou, T. & Rosen, B. P. (1997). *J. Biol. Chem.* **272**, 19731–19737.