Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# Preparation and crystallization of a *Bacillus subtilis* arsenate reductase

**Zhi Guan,[a] Lars Hederstedt,[b] Jin-ping Li[c] and Xiao-Dong Su[a]***

[a]Department of Molecular Biophysics, Kemicentrum, PO Box 124, SE-221 00 Lund, Sweden, [b]Department of Microbiology, Lund University, Sölvegatan 12, SE-223 62 Lund, Sweden, and [c]Department of Medical Biochemistry and Microbiology, University of Uppsala, Biomedical Center, Box 582, SE-751 23 Uppsala, Sweden

Correspondence e-mail: xiao-dong.su@mbfys.lu.se

Received 19 July 2001
Accepted 24 August 2001

Arsenate reductase (AR) in *B. subtilis* is encoded by the chromosomal *arsC* gene. Together with *arsB* and *arsR*, *arsC* participates in detoxification processes for the arsenate and arsenite ions. Full-length arsenate reductase without any modification has been expressed in *Escherichia coli* and purified in a soluble form. The recombinant protein has been crystallized at 277 K using polyethyleneglycol (PEG) or poly(ethyleneglycol) methyl ether (PME) as the main precipitant. At least two forms of crystals large enough for data collection have been obtained from wild-type protein under different conditions. An orthorhombic crystal diffracted to beyond 2.2 Å with space group $P2_12_12_1$ and unit-cell parameters $a$ = 51.22, $b$ = 91.62, $c$ = 101.93 Å. A near-complete data set has been collected to 2.5 Å. The application of the flash-annealing technique was crucial for high resolution during the data collection. The SeMet-substituted AR has also been produced and crystallized under very similar conditions as the wild type, but the unit-cell parameters are very different. The crystals of the SeMet protein diffracted to higher resolution than those of the wild type.

## 1. Introduction

Arsenic compounds are widely distributed in the environment and are toxic to all life forms. It is thus important for free-living single-cell organisms to develop versatile and efficient ways to get rid of toxic arsenic compounds. In bacteria, at least two ways exist to detoxify arsenic compounds. In a few organisms, such as *Alcaligenes faecalis*, the more toxic arsenite is methylated or oxidized to less toxic compounds (Anderson *et al.*, 1992; Ellis *et al.*, 2001). Most bacteria, however, have transport systems to export arsenic compounds out of the cells (Rensing *et al.*, 1999).

Resistance to arsenate, arsenite and antimony (III) ions is governed by an operon of five genes, *arsABCDR*, in most Gram-negative bacteria, *e.g. E. coli* (Rosen *et al.*, 1992), and three genes, *arsBCR*, in Gram-positive bacteria such as *Bacillus subtilis* (Sato & Kobayashi, 1998) and *Staphylococcus aureus* (Silver *et al.*, 1993). The *arsC* gene encodes arsenate reductase (AR), which reduces intracellular less toxic $As^{V}$ (arsenate) to more toxic $As^{III}$ (arsenite). This may seem irrational; however, only arsenite ions can be rapidly exported out of the cells. Fig. 1 shows a multiple sequence alignment and a phylogenic tree of eight ARs from both Gram-negative and Gram-positive bacteria. It is apparent that these proteins can be divided into two subfamilies corresponding to the enzymes in Gram-negative and Gram-positive bacteria. The ARs from *B. subtilis* and *E. coli* share only about 10% sequence identity in this alignment. The catalytic mechanism of the AR of the two families is also distinct. All ARs feature the same cysteine in the active site (indicated by the double arrow at residue 12 in Fig. 1) but the redox partner is different. In *E. coli*, the glutathione–glutaredoxin system is used by AR (Gladysheva *et al.*, 1994; Shi *et al.*, 1999), while in *S. aureus* the thioredoxin system is utilized (Ji *et al.*, 1994; Ji & Silver, 1992). The use of different redox partners may be reflected in the fact that the AR of Gram-positive bacteria contains a pair of cysteines (Cys82 and Cys89 in both *S. aureus* and *B. subtilis*) as indicated by single arrows in Fig. 1. These two cysteines are shown to be essential for the activity of *S. aureus* AR (Messens *et al.*, 1999); however, such a cysteine pair is absent in *E. coli* AR.

Structural investigations have been carried out on AR for many years. For example, diffracting crystals have been obtained of *E. coli* AR (de Mel *et al.*, 1994; Rosen *et al.*, 1991), but the structure is still not available to the public (there are two PDB deposits currently on hold with entry codes 1j9d and 1j9b). NMR studies of *E. coli* AR seem to have encountered difficulties in obtaining a refined three-dimensional structure (Stevens *et al.*, 1999). Since Gram-positive ARs belong to a different subfamily to the Gram-negative ARs, we have started to work on *B. subtilis* AR,

aiming towards solving its structure and understanding its catalytic mechanism. It will also be very interesting to compare the structures of Gram-negative and Gram-positive arsenate reductases.

## 2. Experimental

### 2.1. Protein cloning, expression and purification

The *arsC* gene was cloned using the PCR technique with *B. subtilis* strain 168 genomic DNA and the primers 5′-CAT G**CC ATG G**AG AAT AAA ATC ATT TAC-3′ and 5′-ATA GTT TA**G CGG CCG C**TT ATTTCCCTG TTT CAG CA-3′. The PCR product contains *Nco*I and *Not*I sites (in bold) for cloning into the pET28a (Novagen) expression vector. The *arsC* open reading frame sequence starts with ATG GAG, which fits perfectly to an *Nco*I site. The ligation of the digested PCR product containing *arsC* into the pET28a, plasmid amplification and transformation were performed by conventional methods. *E. coli* strain BL21 (DE3) was used for the protein-production work. The cloned DNA was sequenced to confirm that no mutation had resulted during PCR and cloning.

For protein expression, the *E. coli* cells were cultured overnight in 10 ml Luria–Bertani broth (LB) at 310 K. The overnight culture was then transferred to 1 l of LB medium. When the $OD_{595}$ reached 0.6–0.8, isopropyl-1-thio-$\beta$-D-galactopyranoside (IPTG) was added to a final concentration of 0.3 m*M* to induce ArsC synthesis. The cells were allowed to grow in the culture for an additional 3 h after induction at 310 K. Cells were harvested by centrifugation at 7000*g* for 15 min at 277 K. Cell pellets were pooled and stored at 189 K.

For protein purification, the cell pellet was suspended in 20 ml cold buffer containing 0.1 *M* Tris–HCl pH 8.2, 5 m*M* EDTA, 0.1%(*v*/*v*) 2-mercaptoethanol and 0.5 m*M* phenylmethylsulfonylfluoride (PMSF). The cell suspension was then sonicated by applying 80 W pulses (Sonics, USA). Cell debris was removed by centrifugation for 20 min at 17 000*g* at 277 K. After centrifugation, about 5 ml of the supernatant was directly loaded onto a Superdex 200 preparative-grade column (Amersham Pharmacia Biotech, Sweden) pre-equilibrated in 20 m*M* Tris–HCl pH 8.2 containing 0.1 *M* NaCl. AR-containing fractions were identified by SDS–PAGE and pooled from multiple runs. The pooled material was diluted at least threefold in water and then loaded onto a Mono Q column (Pharmacia) equilibrated with buffer *A* (20 m*M* Tris–HCl pH 8.2). The column was eluted with a 20 column volume linear gradient of buffer *B* (buffer *A* with 1.0 *M* NaCl). The AR-containing fractions were pooled and concentrated to about 10–20 mg ml$^{-1}$ using a Centricon-10 device (Amicon, USA). The FPLC columns used in the purification steps were mounted on a BioCAD SPRINT system (PerSeptive Biosystems, Inc., USA).

The SeMet-substituted AR was produced according to Van Duyne *et al.* (1993) in *E. coli* strain BL21 (DE3) with slight modifications. Purification of the selenium-containing AR is similar to the purification of the native protein.

### 2.2. Crystallization

The initial crystallization conditions were screened by the sparse-matrix method (Jancarik & Kim, 1991) using the Hampton Screen Kit 1 and Screen Kit 2 (Hampton Research, USA). All crystallization experiments were performed at 277 K by the hanging-drop vapour-diffusion method in 24-well VDX plates (Hampton Research, USA). In each trial, a hanging drop of 1 µl of protein solution mixed with 1 µl of precipitant solution was equilibrated against a reservoir containing 500 µl of precipitant solution.

### 2.3. Data collection and processing

Crystals of AR were picked up with a nylon CryoLoop (Hampton Research, USA) directly from mother liquor and were flash-frozen in a cold nitrogen stream (Oxford Cryosystems Cryostream Cooler). Diffraction data were collected at the crystallographic beamline BL711 at the MAX-II synchrotron in Lund, Sweden (Cerenius *et al.*, 2000) at 100 K by the oscillation method using a MAR345 image-plate detector (X-ray Research GbmH). The so-called *in situ* flash-annealing method (Yeh & Hol, 1998) (that is, stopping the cold nitrogen stream for several seconds to allow the frozen crystal to thaw completely and then freezing the crystal quickly again) was used to impove the diffraction quality of the crystals. The wavelength for the data collection was 1.032 Å for wild-type AR; the crystal-to-detector distance was 250 mm, with an oscillation range of 1° per image. Data were processed using the *DENZO* and *SCALEPACK* packages (Otwinowski & Minor, 1996).

## 3. Results and discussion

### 3.1. Protein preparation

The *B. subtilis* AR is composed of 139 amino-acid residues with a calculated $M_r$ of

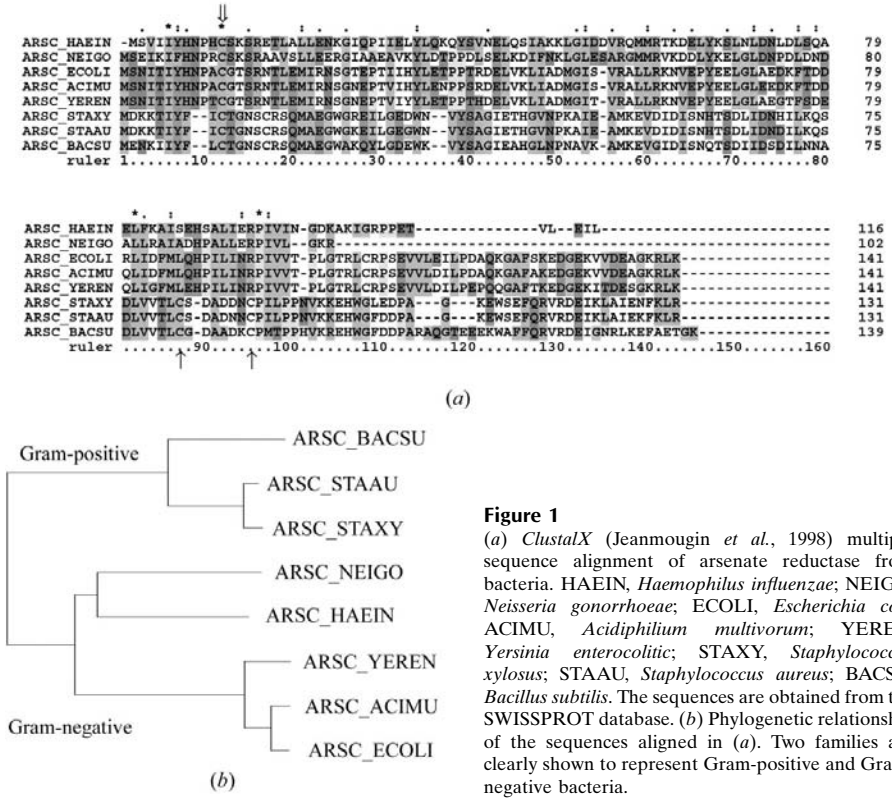


**Figure 1**
(*a*) *ClustalX* (Jeanmougin *et al.*, 1998) multiple sequence alignment of arsenate reductase from bacteria. HAEIN, *Haemophilus influenzae*; NEIGO, *Neisseria gonorrhoeae*; ECOLI, *Escherichia coli*; ACIMU, *Acidiphilium multivorum*; YEREN, *Yersinia enterocolitic*; STAXY, *Staphylococcus xylosus*; STAAU, *Staphylococcus aureus*; BACSU, *Bacillus subtilis*. The sequences are obtained from the SWISSPROT database. (*b*) Phylogenetic relationship of the sequences aligned in (*a*). Two families are clearly shown to represent Gram-positive and Gram-negative bacteria.

**Table 1**
Data-collection statistics.

Values in parentheses are for the last resolution shell.

| | Wild-type crystal | SeMet crystal |
|---|---|---|
| Wavelength (Å) | 1.032 | 1.09 |
| Resolution (Å) | 30.0–2.5 (2.6–2.5) | 30.0–1.6 (1.7–1.6) |
| Completeness (%) | 86.0 (60.2) | 97.3 (97.4) |
| $R_{\rm merge}$† (%) | 8.0 (34.3) | 4.4 (30.0) |
| $I/\sigma(I)$ | 13.5 (2.8) | 19.0 (4.7) |
| Space group | $P2_12_12_1$ | $P2_12_12_1$ |
| Unit-cell parameters (Å) | $a$ = 51.22, $b$ = 91.62, $c$ = 101.93 | $a$ = 50.94, $b$ = 94.31, $c$ = 118.70 |
| No. of unique reflections possible | 17388 | 76215 |
| No. of reflections observed | 164533 | 297895 |
| Tetramer $V_{\rm M}$ (Å$^3$ Da$^{-1}$) | 1.92 | 2.28 |
| Tetramer solvent content (%) | 35.8 | 46.2 |
| Trimer $V_{\rm M}$ (Å$^3$ Da$^{-1}$) | 2.56 | 3.05 |
| Trimer solvent content (%) | 67.9 | 73.1 |

† $R_{\rm merge} = \sum|I_{\rm obs} - I_{\rm avg}|/\sum I_{\rm obs}$, where the summation is over all reflections.

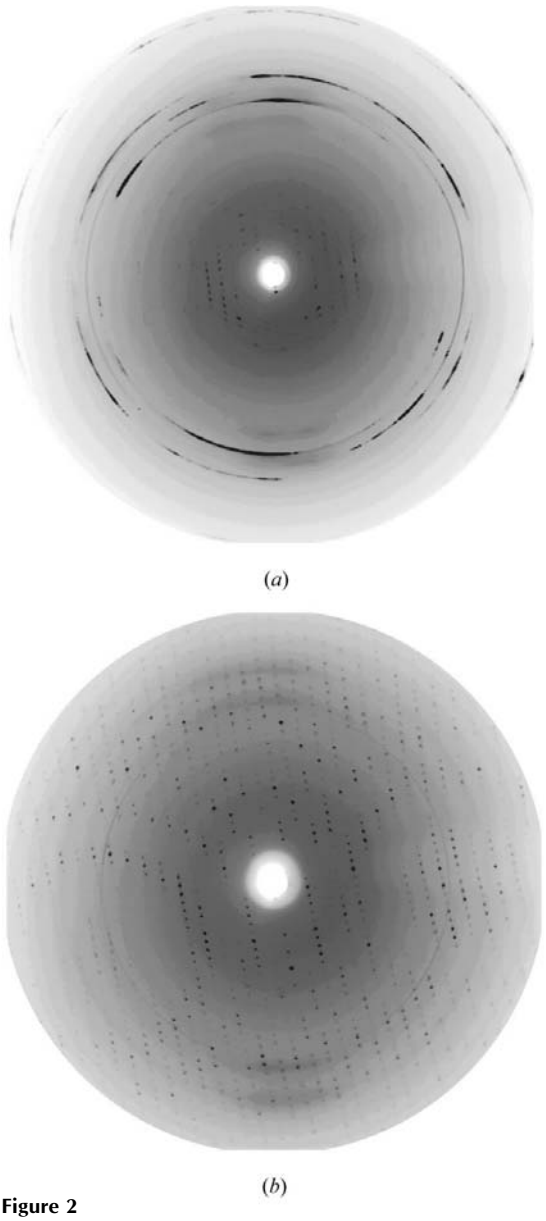

**Figure 2**
Diffraction patterns of the wild-type crystal (*a*) before and (*b*) after flash-annealing.

15 600. About 20 mg of pure *B. subtilis* AR can be obtained from 1 l of culture.

### 3.2. Crystallization

The crystals were first observed in conditions 4 and 20 of the Hampton Screen Kit 1 and then also in condition 13 of the Screen Kit 2. The crystallization conditions (pH, PEG or PME concentration, protein, salt and buffer concentrations) were then optimized by finer screening around these conditions. The best crystals were obtained in drops with pH 4.4–4.6 in 0.1 *M* sodium acetate and 0.2 *M* ammonium sulfate with the presence of 30–35% PME. The crystallization of AR has been difficult since the reproducibility is low with different protein batches. Spontaneous nucleation and growth of crystals were not possible for protein batches stored in a refrigerator for longer than a few days. Macro-seeding procedures using either nylon cryoloop or a piece of hair were tried and could not produce better crystals. The SeMet crystals can be obtained under almost identical conditions to the wild type.

Most wild-type crystals diffracted poorly and showed serious disorder and high mosaicity when frozen directly. A rather large crystal with dimensions 0.2 × 0.2 × 0.1 mm was mounted and showed poor diffraction and disorder (Fig. 2*a*); however, the *in situ* flash-annealing method (Yeh & Hol, 1998) improved the diffraction limits drastically. As shown in Fig. 2, the directly frozen crystal showed bad ice rings and only diffracted to about 6 Å, but after flash-annealing the ice rings disappeared completely and the crystal diffracted well beyond 2.3 Å, which is at the edge of the image. A near-complete data set (about 86% completeness to 2.5 Å and about 95% to 3.5 Å) was collected from this crystal. The statistics of this data set are listed in Table 1. The crystal belongs to the space group $P2_12_12_1$, with unit-cell parameters $a$ = 51.22, $b$ = 91.62, $c$ = 101.93 Å. The SeMet crystals diffracted much better than the wild type amd a 1.6 Å data set was collected from a SeMet crystal. Although the crystallization condition and the space group are the same for both wild-type and SeMet crystals, their unit-cell parameters differ greatly (see Table 1 for comparison). The possible quaternary states of the crystallized proteins in these two crystals are also listed; the asymmetric units of both crystals are most likely to contain four or three molecules as estimated by the calculated value of $V_{\rm M}$ (Matthews, 1968) (see Table 1).

Since there is no AR structure yet available, *ab initio* phasing methods have to be used to solve this structure. The MAD (multiple-wavelength anomalous dispersion) method is the first choice to solve the SeMet-containing AR. Heavy-atom derivatization is also under way.

This work was supported by grants from Cancerfonden, Sweden. XDS is supported by the Swedish Foundation for Strategic Research (SSF) and the Structural Biology Network (SBNet). We thank Jitka Vevodova and Matthew Bennett for help with data collection and processing.

## References

Anderson, G. L., Williams, J. & Hille, R. (1992). *J. Biol. Chem.* **267**, 23674–23682.
Cerenius, Y., Ståhl, K., Svensson, L. A., Ursby, T., Oskarsson, Å., Albertsson, J. & Liljas, A. (2000). *J. Synchrotron Rad.* **7**, 203–208.
Ellis, P. J., Conrads, T., Hille, R. & Kuhn, P. (2001). *Structure*, **9**, 125–132.
Gladysheva, T. B., Oden, K. L. & Rosen, B. P. (1994). *Biochemistry*, **33**, 7288–7293.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G. & Gibson, T. J. (1998). *Trends Biochem. Sci.* **23**, 403–405.
Ji, G., Garber, E. A., Armes, L. G., Chen, C. M., Fuchs, J. A. & Silver, S. (1994). *Biochemistry*, **33**, 7294–7299.
Ji, G. & Silver, S. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 9474–9478.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mel, V. S. de, Doyle, M. A., Gladysheva, T. B., Oden, K. L., Martin, P. D., Rosen, B. P. & Edwards, B. F. (1994). *J. Mol. Biol.* **242**, 701–702.
Messens, J., Hayburn, G., Desmyter, A., Laus, G. & Wyns, L. (1999). *Biochemistry*, **38**, 16857–16865.
Otwinowski, Z. & Minor, W. (1996). *Methods Enzymol.* **276**, 307–326.
Rensing, C., Ghosh, M. & Rosen, B. P. (1999). *J. Bacteriol.* **181**, 5891–5897.
Rosen, B. P., Dey, S., Dou, D., Ji, G., Kaur, P., Ksenzenko, M., Silver, S. & Wu, J. (1992). *Ann. NY Acad. Sci.* **671**, 257–272.
Rosen, B. P., Weigel, U., Monticello, R. A. &

Edwards, B. P. (1991). *Arch. Biochem. Biophys.* **284**, 381–385.
Sato, T. & Kobayashi, Y. (1998). *J. Bacteriol.* **180**, 1655–1661.
Shi, J., Vlamis-Gardikas, A., Aslund, F., Holmgren, A. & Rosen, B. P. (1999). *J. Biol. Chem.* **274**, 36039–36042.
Silver, S., Ji, G., Broer, S., Dey, S., Dou, D. & Rosen, B. P. (1993). *Mol Microbiol.* **8**, 637–642.
Stevens, S. Y., Hu, W., Gladysheva, T., Rosen, B. P., Zuiderweg, E. R. & Lee, L. (1999). *Biochemistry*, **38**, 10178–10186.
Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. (1993). *J. Mol. Biol.* **229**, 105–124.
Yeh, J. I. & Hol, W. G. (1998). *Acta Cryst.* D**54**, 479–480.