Extremophiles (2003) 7:123–130
DOI 10.1007/s00792-002-0303-6

ORIGINAL PAPER

Thomas M. Gihring · Philip L. Bond
Stephen C. Peters · Jillian F. Banfield

Arsenic resistance in the archaeon *"Ferroplasma acidarmanus"*: new insights into the structure and evolution of the *ars* genes

Received: 2 March 2002 / Accepted: 25 October 2002 / Published online: 16 January 2003

Abstract Arsenic resistance in the acidophilic iron-oxidizing archaeon "*Ferroplasma acidarmanus*" was investigated. *F. acidarmanus* is native to arsenic-rich environments, and culturing experiments confirm a high level of resistance to both arsenite and arsenate. Analyses of the complete genome revealed protein-encoding regions related to known arsenic-resistance genes. Genes encoding for ArsR (arsenite-sensitive regulator) and ArsB (arsenite-efflux pump) homologues were found located on a single operon. A gene encoding for an ArsA relative (anion-translocating ATPase) located apart from the *arsRB* operon was also identified. Arsenate-resistance genes encoding for proteins homologous to the arsenate reductase ArsC and the phosphate-specific transporter Pst were not found, indicating that additional unknown arsenic-resistance genes exist for arsenate tolerance. Phylogenetic analyses of ArsA-related proteins suggest separate evolutionary lines for these proteins and offer new insights into the formation of the *arsA* gene. The ArsB-homologous protein of *F. acidarmanus* had a high degree of similarity to known ArsB proteins. An evolutionary analysis of ArsB homologues across a number of species indicated a clear relationship in close agreement with 16S rRNA evolutionary lines. These results support a hypothesis of arsenic resistance developing early in the evolution of life.

Keywords Arsenic · Resistance · *ars* · *Ferroplasma acidarmanus* · Acidophile · Evolution · Phylogeny · Genomics

Communicated by W.D. Grant

T.M. Gihring (✉)
Department of Geology and Geophysics,
The University of Wisconsin Madison, Madison,
Wisconsin, USA
E-mail: thomas.gihring@pnl.gov
Tel.: +1-509-3721254
Fax: +1-509-3769650

P.L. Bond
School of Biological Sciences,
University of East Anglia, Norwich, UK

S.C. Peters
Department of Geological Sciences,
The University of Michigan, Ann Arbor, Michigan, USA

J.F. Banfield
Department of Geology and Geophysics,
The University of Wisconsin Madison,
Madison, Wisconsin, USA

J.F. Banfield
Department of Earth and Planetary Sciences,
University of California Berkeley, Berkeley, California, USA

Present addresses: T.M. Gihring
Pacific Northwest National Laboratory, MS P7-50,
P.O. Box 999, Richland, WA 99352, USA

S.C. Peters
Department of Earth and Environmental Sciences,
Lehigh University, Bethlehem, PA, USA

Introduction

"*Ferroplasma acidarmanus*" is an extremely acidophilic iron-oxidizing archaeon isolated from a metal-rich acid mine drainage environment at Iron Mountain, California (Edwards et al. 2000). Oxidative sulfide mineral dissolution at this site is thought to be accelerated by iron-oxidizing microorganisms, including *F. acidarmanus*, contributing to the generation of acid mine drainage and the release of heavy metals (McGuire et al. 2001). Metal concentrations in the solutions colonized by *F. acidarmanus* at Iron Mountain are remarkably high, with iron, zinc, copper, and cadmium levels typically exceeding 28 g l^{-1}, 2.5 g l^{-1}, 380 mg l^{-1}, and 250 mg l^{-1}, respectively (Edwards et al. 1998). Measurements of total dissolved arsenic range from 53 to 56 mg l^{-1} (Alpers et al. 1994; Edwards et al. 1998), approximately three orders of magnitude greater than the current United States Environmental Protection Agency drinking water limit (United States Environmental Protection Agency 2001).

The mode and degree of arsenic toxicity are dependent largely on its form and oxidation state. Inorganic aqueous arsenite has a strong affinity for sulfhydryl groups of proteins and is considered many times more toxic than inorganic aqueous arsenate (Saha et al. 1999). Because of its similar size and electrochemical characteristics, inorganic arsenate can substitute for inorganic phosphate during ATP synthesis and other cellular processes (Saha et al. 1999). Arsenite may pass into the cell directly through the membrane or by way of inorganic carrier proteins as an unionized species (Cervantes et al. 1994). Arsenate enters a cell via transmembrane phosphate transport proteins (Cervantes et al. 1994). To counteract the toxic effects of arsenic, microbial resistance factors are required to actively exclude or expel arsenic from the cell.

Chromosomal and plasmid-based arsenic-resistance genes of a number of bacteria (see Silver 1996 for review) and one archaeon (Ng et al. 1998) have been described. The well-studied *Escherichia coli* plasmid R773 contains an operon for the *arsRDABC* genes, while other plasmid and chromosomal systems have only the *arsRBC* genes (Diorio et al. 1995; Silver 1996; Butcher et al. 2000). All published operons of experimentally proven arsenic-resistant microorganisms have at a minimum the *arsRBC* genes (Rosen 1999). The gene *arsR* encodes for an arsenite-responsive transcriptional repressor that controls basal levels of *ars* gene expression (Wu and Rosen 1991), while upper levels of *ars* operon expression are controlled by the *arsD* gene product (Chen and Rosen 1997; Rensing et al. 1999). ArsA is an ATPase consisting of two homologous halves designated as the A1 and A2 loops (Chen et al. 1986). ArsA is allosterically activated by arsenite and functions as a catalytic subunit of ArsB, a membrane-located arsenite transporter (Rosen 1999). ArsB can function with or without the catalytic ArsA ATPase, although resistance is reportedly enhanced by the ArsA unit (Cervantes et al. 1994). Resistance to arsenate is conferred by ArsC, a reductase for the conversion of arsenate to the substrate of the efflux pump (Gladysheva et al. 1994).

While arsenic resistance has been studied extensively in a small number of bacterial species, little is known about the arsenic tolerance mechanisms of archaea. *F. acidarmanus*, native to an arsenic-rich environment, is likely to be resistant to high levels of arsenic and to utilize a specific system to achieve arsenic resistance. We have studied arsenic tolerance in this archaeon and identified genes possibly relevant to arsenic resistance. The recent full genomic sequencing of *F. acidarmanus*, as well as many other prokaryotes, has provided for a comparison of arsenic-resistance genes across a wide range of species and a new look at the evolution of arsenic resistance.

Materials and methods

Growth experiments

Aqueous arsenic stock solutions (3,750 mg l^{-1} As) were prepared from an arsenious acid solution (LabChem Inc., Pittsburgh, Pa.) or NaH_2AsO_4 (J. T. Baker, Phillipsburg, N.J.). Arsenic solutions were acidified to pH 1.2 using H_2SO_4 and were sterilized by filtration. *Ferroplasma acidarmanus* was maintained at 37 °C in pH 1.2 basal medium [800 mg l^{-1} $(NH_4)_2SO_4$, 400 mg l^{-1} KH_2PO_4, 160 mg l^{-1} $MgSO_4 \cdot 7H_2O$, 1 μg l^{-1} $CoCl_2 \cdot 6H_20$, 1 μg l^{-1} $CuCl_2 \cdot 2H_2O$, 1 μg l^{-1} $MnCl_2 \cdot 4H_2O$, 1 μg l^{-1} $ZnCl_2$, and 1 μg l^{-1} $CaCl_2 \cdot 2H_2O$] augmented with 0.02% (w/v) yeast extract, 20 g l^{-1} $FeSO_4$, and 100 mg l^{-1} arsenite or arsenate. For growth experiments, cells conditioned in the above medium were washed twice in fresh medium prior to experiment inoculation. To test for growth in the presence of arsenic, *F. acidarmanus* was cultured with additions of 0, 1, 100, and 1,000 mg l^{-1} arsenic as arsenite or arsenate. Samples were taken daily for measurements of cell density and arsenic speciation. Abiotic control experiments were also incubated under the same conditions using sterile media with 100 mg l^{-1} arsenite or arsenate. Cell densities were measured using a Petroff-Hausser counting chamber.

Total arsenic was determined using the method of Klaue and Blum (1999) for continuous-flow online hydride generation with magnetic sector ICP-MS detection (Finnigan ELEMENT, Bremen, Germany). Arsenic speciation was determined using the same method coupled with ultraviolet oxidation (Dagnac et al. 1999; Peters 2001; Wei et al. 2001). Experimental samples were filtered (0.22 μm) and diluted with 1% HNO_3 (Seastar Chemicals, Sidney, B.C.) for hydride generation analysis. To verify that the sample matrix of 20 g l^{-1} $FeSO_4$ would not interfere with arsenic measurements, samples of the matrix were spiked with 10 μg l^{-1} each of arsenite and arsenate. Spike recoveries ranged from 94% to 101%. Analysis of reference standards NIST-1643d (NIST, Gaithersburg, Md.), CRM-RSA, and CRM-RSB (HPS, Charleston S.C.) agreed to within 5% of published values.

Analysis of *ars*-related operons

Preliminary *F. acidarmanus* genome sequence information was obtained from the Department of Energy Joint Genome Institute (http://www.jgi.doe.gov/index.html) and the University of Wisconsin-Madison *Escherichia coli* Genome Center. Open reading frames (ORFs) identified as having similarity to known *ars* genes based on translated protein BLAST (basic local alignment search tool; Altschul et al. 1997) searches were downloaded for further analysis. Plasmids containing *ars* gene relatives were not identified.

For the evaluation of *F. acidarmanus* arsenic-resistance gene homology, related amino acid sequences were acquired from the GenBank database. Amino acid sequences used in the ArsR comparison (GenBank protein accession numbers are in parentheses) were *Aquifex aeolicus* (G70420), *Acidithiobacillus ferrooxidans* (AAF69241), *Bacillus halodurans* (BAB06719), *Bacillus subtilis* (CAB15384), *E. coli* plasmid R46 (AAB09624), *E. coli* plasmid R773 (BVECAR), *E. coli* K12 (AAC76526), *Halobacterium* sp. NRC1 plasmid pNRC100 (T08342), *Klebsiella oxytoca* plasmid pMH12 (AAF89638), *Staphylococcus aureus* plasmid pI258 (P30338), *Staphylococcus xylosus* plasmid pSX267 (AAA27587), *Sinorhizobium* sp. As4 (AAD51845), *Thermoplasma acidophilum* (CAC12237), and *Yersinia enterocolitica* plasmid pYV (AAB42205). *T. acidophilum* is the closest relative to *F. acidarmanus* for which the genomic sequence was available.

Sequences used in the ArsA comparison included the following: *Acidiphilum multivorum* plasmid pKW301 (BAA24822), *Aquifex aeolicus* (O66908 and O66674), *Bacillus halodurans* (BAB05514), *Chlorobium vibrioforme* (Q46465), *E. coli* plasmid R46 (AAB09626), *E. coli* plasmid R773 (AAA21094), *Halobacterium* sp. NRC1 (AAG18929), *Halobacterium* sp. NRC1 plasmid pNRC100 (AAC82907), *K. oxytoca* plasmid pMH12 (AAF89640), *Methanococcus jannaschii* (AAB99142), *Methanothermobacter thermoautotrophicum* (AAB85986), *Sinorhizobium* sp. As4 (AAD51849), *Synechocystis* sp. PCC 6803 (Q55794), *T. acidophilum* (CAC11579), and *T. volcanium* (NP_111575).

The ArsB-related amino acid sequences used in this study were from *Acidiphilium multivorum* plasmid pKW301 (BAA24823), *Acidithiobacillus ferrooxidans* (AAF69238), *B. halodurans* (BAB06718), *B. subtilis* (BAA06969), *E. coli* plasmid R46

(AAB09627), *E. coli* plasmid R773 (AAA21095), *E. coli* K12 (CAA56362), *K. oxytoca* plasmid pMH12 (AAF89641), *Pseudomonas aeruginosa* (G83361), *Serratia marcescens* (CAB88405), *S. aureus* (BAB43886), *Staphylococcus aureus* plasmid pI258 (AAA25637), *Staphylococcus xylosus* plasmid pSX267 (AAA27588), *T. acidophilum* (CAC11316), *T. volcanium* (NP_110804), and *Yersinia enterocolitica* plasmid pYV (P74985). Putative ArsB protein sequences in the GenBank database having less than 30% amino acid identity with the R773 ArsB sequence were not included in this study.

Accession numbers for 16S rDNA sequences were: *Acidiphilium multivorum* (AB006712), *Acidithiobacillus ferrooxidans* (AF362022), *B. halodurans* (AB043971), *B. subtilis* (NC_000964), *E. coli* (NC_000913), *K. oxytoca* (AF390083), *P. aeruginosa* (NC_002516), *P. putida* (AJ308313), *Serratia marcescens* (AF286873), *Staphylococcus aureus* (NC_002745), *Staphylococcus xylosus* (D83374), *T. acidophilum* (M38637), *T. volcanium* (NC_002689), and *Y. enterocolitica* (AF366378).

ArsR, ArsA, and ArsB amino acid sequence alignments were performed using the maximal linkage clustering method of the Pattern Induced Multiple Alignment algorithm (Smith and Smith 1990, 1992) on the Baylor College of Medicine Search Launcher website (http://searchlauncher.bcm.tmc.edu). Neighbor-joining phylograms for protein sequences were constructed by comparison of mean character differences using PAUP* version 4.0b4a. 16S rDNA sequences were aligned using arbEDIT4 and GDE within the arb software package. A neighbor-joining 16S rDNA tree was constructed using PAUP* with the Jukes-Cantor correction and distances measures.

Results and discussion

Growth rates of *Ferroplasma acidarmanus* were not reduced by the presence of 1, 100, and 1,000 mg l^{-1} arsenate relative to growth in the absence of arsenate (Fig. 1a). Cultures challenged with 1, 100, and 1,000 mg l^{-1} arsenite also showed no significant change in exponential growth rates compared to cultures with no arsenic (Fig. 1b). Measurements of As(III)/(V) speciation in the experiments indicate that no oxidation or reduction of arsenic occurred (data not shown). Abiotic experiments confirm that both arsenite and arsenate were stable under the culturing conditions and were not being recycled by abiotic oxidation or reduction. Therefore, an As-oxidation pathway for the conversion of arsenite to the less toxic form arsenate is not used by *F. acidarmanus* as a tolerance mechanism. In addition, the inability to reduce arsenic is consistent with the lack of an *arsC*-homologous gene for an arsenate reductase component of the *ars* operon. Despite an apparent inability to convert intracellular arsenate to the substrate of the efflux pump, sensitivity to arsenate was not evident. It is possible that *F. acidarmanus* employs phosphate-specific transporters homologous to the *Escherichia coli* Pst protein, which reduces non-specific uptake of arsenate (Cervantes et al. 1994). Although genes related to the *E. coli* phosphate-specific transporter were not identified in the *F. acidarmanus* genome, an analogous system may be present. At this time, the mechanism of arsenate resistance by *F. acidarmanus* remains unknown.

Analysis of the *F. acidarmanus* genomic sequence revealed three ORFs related to known arsenic resistance genes. A region with two adjacent ORFs encoding for putative ArsR and ArsB homologues was identified. An ORF coding for an ArsA-related protein was also observed. The *arsA*-related gene was not situated near the *arsRB* genes, and genes homologous to *arsC* and *arsD* were not identified.

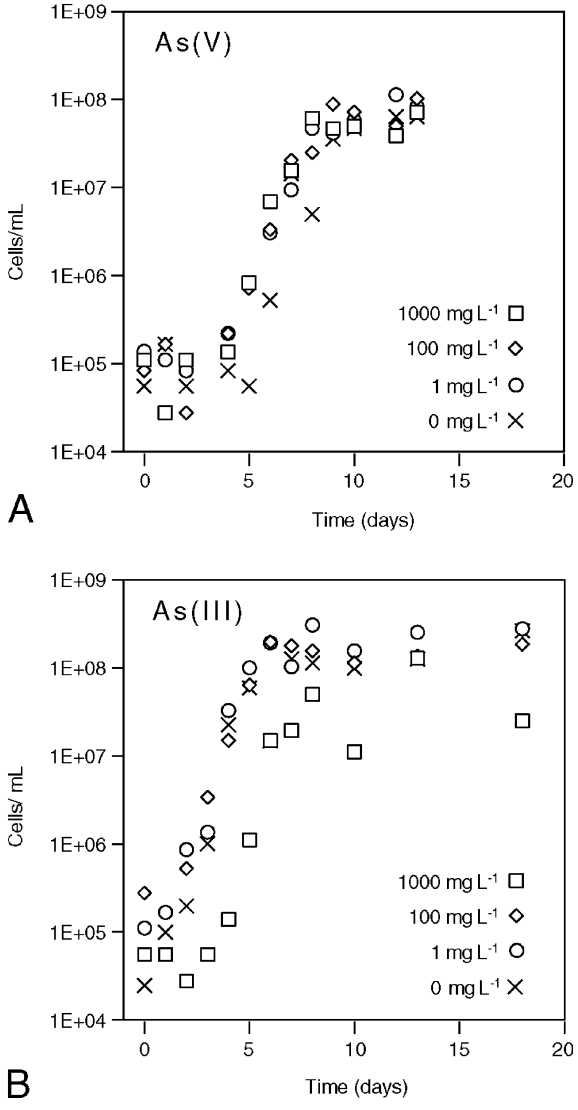


Fig. 1a, b Cell densities during growth of *F. acidarmanus*. **a** Growth in the presence of arsenate. **b** Growth in the presence of arsenite

The *F. acidarmanus* gene homologous to *arsR* putatively encodes for a protein of 118 amino acids. An alignment of ArsR-related proteins in Fig. 2 indicates that the ArsR-related protein is comparable in size and composition to other ArsR homologues. The *F. acidarmanus* ArsR protein sequence has 25.0% identity with the *E. coli* plasmid R773 ArsR protein and only 15.6% identity with the *Thermoplasma acidiphilum* ArsR-related protein. Cysteine (Cys) residues at the 32, 34, and 37 positions of the R773 ArsR (as indicated in Fig. 2) have been shown to be required for the regulatory function of this protein (Shi et al. 1996; Rosen 1999).

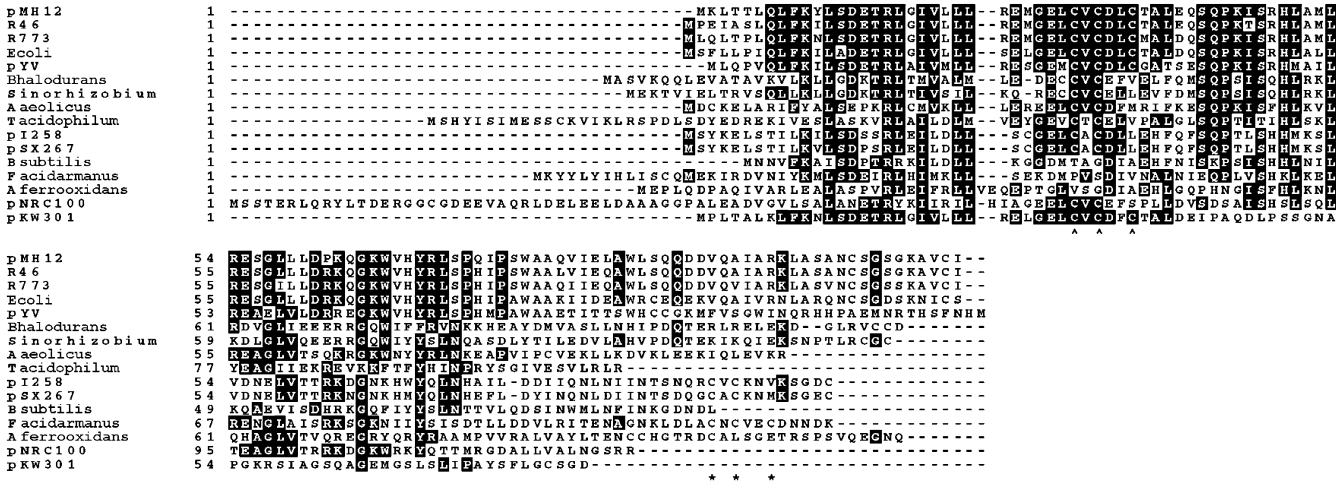


Fig. 2 Alignment of ArsR-homologous protein sequences. A *caret* indicates the position of arsenite-binding cysteine residues within the R773 ArsR protein sequence. An *asterisk* specifies the cysteine residues of the *F. acidarmanus* ArsR-homologous protein. Positions with >30% consensus are *shaded*

The protein ArsR is notable in requiring the binding of arsenite during the regulatory response. Binding of arsenite to sulfhydryl groups of cysteine often results in structural changes in proteins and is one of the primary biochemical factors in arsenite toxicity. However, the conformational change caused through binding of arsenite to the cysteine residues of ArsR (as studied in R773) allows the dissociation of the regulator protein from the promoter region, permitting subsequent transcription of the *ars* genes (Wu and Rosen 1991; Rosen 1999). The putative *arsR* gene product of *F. acidarmanus* does not possess Cys residues, required for binding of arsenite, in the same locations as the Cys residues of ArsR, as encoded on the R773 plasmid (Fig. 2). However, cysteine residues at the carboxyl-terminal end of the protein located in an arrangement similar to the R773 Cys residues may hypothetically function as arsenite-binding sites and allow the ArsR homologue to function as an arsenite-sensitive regulator.

The *arsA*-related gene has a translation product corresponding to a protein of 386 amino acids. Putative ArsA proteins indexed in the GenBank database were aligned together with the deduced *F. acidarmanus* ArsA-related protein. The alignment, shown in Fig. 3, has been annotated to indicate the position of the A1 and A2 loops (Chen et al. 1986), the A1-A2 linker (Li and Rosen 2000), ATP-binding regions (Li et al. 1996), DTAP motifs (12-residue sequences highly conserved in ArsA; Zhou and Rosen 1997), and the arsenite-binding cysteine residues (Bhattacharjee et al. 1995). These regions were previously described for the *E. coli* plasmid R46 and R773 ArsA proteins.

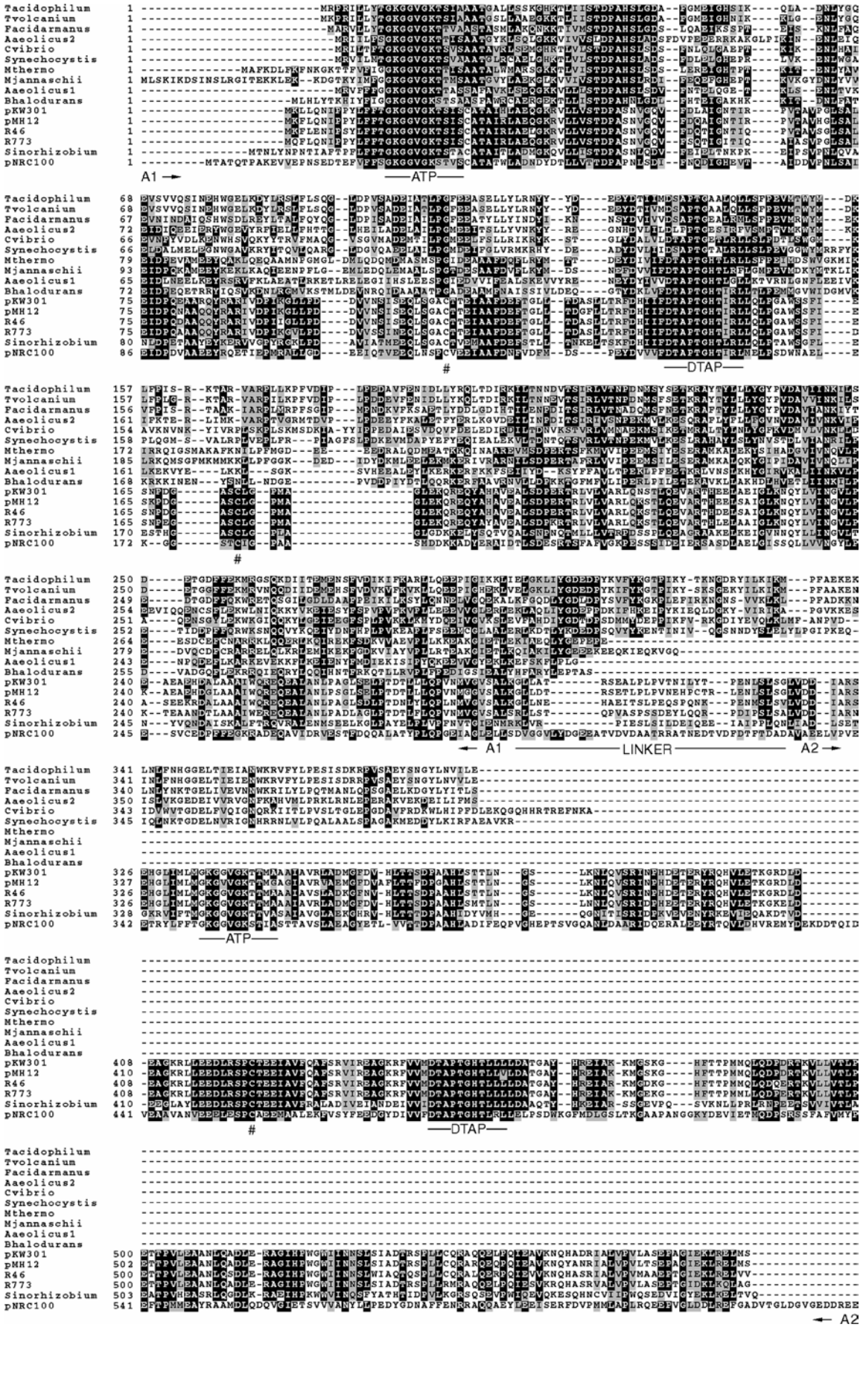
A significant disparity in the number of encoded amino acids of the *arsA* homologues is readily apparent. The ArsA-analogous protein of *F. acidarmanus* is over 200 amino acids shorter than the plasmid-based ArsA proteins of *Acidiphilium multivorum* (pKW301), *Klebsiella oxytoca* (pMH12), *E. coli* (R773 and R46), and *Halobacterium* sp. NC1 (pNRC100) and the chromosome-based protein of *Sinorhizombium* sp. As4. This discrepancy in the length of the protein sequences forms a substantial division between the ArsA-analogous proteins. Sequences in Fig. 3 containing both the A1 and A2 loops (pKW301, pMH12, R773, R46, pNRC100, and *Sinorhizobium*) will be referred to as the ArsA1:A2 proteins, while the remaining sequences (*Aquifex aeolicus* sequences1 and 2, *Bacillus halodurans*, *Chlorobium vibrioforme*, *F. acidarmanus*, *Methanococcus jannaschii*, *Methanothermobacter thermoautotrophicum*, *Synechocystis* sp. PCC 6803, *Thermoplasma acidophilum*, and *T. volcanium*) will be termed ArsA1-analogous proteins.

A region of the *F. acidarmanus* ArsA1-analogous protein corresponding to the A1 loop of R773 includes an analogous ATP-binding region and a DTAP motif (Fig. 3). Following the A1 region, the putative protein extends past the A1-A2 linker and partially into the A2 loop. As seen in Fig. 3, there is essentially no similarity between the full A1:A2 proteins and the partial A1 proteins beyond the A1 loop. Arsenite-binding cysteine residues present in the full ArsA1:A2 protein are not found in the *F. acidarmanus* ArsA1 analogue. Within the A1 region, there is 19.8% amino acid identity between the *F. acidarmanus* and *E. coli* plasmid R773 proteins. The *F. acidarmanus* ArsA1-analogous protein has 54.5% overall amino acid identity with the *T. acidophilum* sequence.

According to the proposed model of Rensing et al. (1999), both the A1 and A2 loops, as well as 3 Cys residues for the binding of arsenite, are required for the function of ArsA. Based on comparisons of amino acid sequences, the *F. acidarmanus* putative protein could not function in the same manner as ArsA and should not be

►

Fig. 3 Alignment of ArsA-related proteins. The locations of significant residues and regions are indicated as identified for the R773 sequence: *ATP* ATP-binding sites, # arsenite-binding Cys residues, *DTAP* DTAP motifs. Positions with >30% consensus are *shaded*


Tacidophilum
Tvolcanium
Facidarmanus
Aaeolicus2
Cvibrio
Synechocystis
Mthermo
Mjannaschii
Aaeolicus1
Bhalodurans
pKW301
pMH12
R46
R773
Sinorhizobium
pNRC100
A1
ATP
DTAP
LINKER
A2

identified as an ArsA protein. A large number of other proteins in the GenBank database classified as ArsA based on amino acid sequence similarity show similar characteristics (see examples in Fig. 3) and also should not be listed as ArsA proteins. While ArsA1 analogues may function as a subunits of a homodimer analogous to the ArsA1:A2 enzyme, the lack of arsenite-binding Cys residues advocates that these proteins are not involved in arsenic resistance.

The region of the *F. acidarmanus* chromosome homologous to *arsB* has a deduced translation product of 429 amino acids. The *F. acidarmanus* putative ArsB has a high degree of similarity with all ArsB proteins including the *T. acidophilum* ArsB protein at 47.4% and the R773 ArsB at 43.3% identical amino acids. This high degree of similarity with known ArsB proteins suggests that *F. acidarmanus* is likely capable of using the ArsB efflux system to achieve arsenite resistance. Based on the above discussion of the ArsA1-analogous protein and the lack of a complete *arsA*-homolgous ORF, the ArsB homologue may not function as a subunit of a primary pump. It has been postulated that in the absence of an ArsA ATPase, ArsB functions as a secondary carrier using the membrane potential to transport the anion from the negative to the positive side of the membrane, expelling arsenite from the cell (Rensing et al. 1999). However, acidophilic microorganisms commonly have a "reversed" membrane potential, and, as noted by previous authors (Butcher et al. 2000), the membrane potential may not be a suitable energy source for the stand-alone ArsB pump. The mechanism of arsenite efflux by arsenic resistant acidophiles is therefore unsolved.

Oxidative dissolution of sulfide minerals at Iron Mountain has resulted in acid mine drainage containing elevated concentrations of heavy metals. This process, which is augmented by iron-oxidizing organisms such as *F. acidarmanus*, has created an extreme, toxic environment. Culturing experiments presented here confirm that *F. acidarmanus* is resistant to elevated levels of arsenite and arsenate. The organism must therefore employ measures for arsenic resistance that allow it to flourish in otherwise toxic arsenic-rich settings. This ability to thrive in surroundings with elevated dissolved arsenic as well as extreme acidity, while contributing to the production of these fluids, would allow for enhanced growth in acid mine drainage as well as other settings, such as arsenopyrite bioreactors for gold extraction.

Many authors have postulated that life arose in anoxic, metal-rich waters and that resistance to reduced toxic metal species would have been necessary early in the evolution of life (Rensing et al. 1999; Rosen 1999). It has also been acknowledged that chromosomal metal-resistance genes are likely the precursors of plasmid-based genes (Carlin et al. 1995). Study of *ars*-related genes located on the chromosomes of archaea and extremophilic bacteria may therefore provide a window to the evolution of the arsenic-resistance genes.

Rosen (1999) speculated that the evolution of arsenical resistance genes involved a series of steps resulting in the "present-day" *ars* operon. It was assumed that shortly after the development of the *arsB* gene, *arsR* was acquired to form the *arsRB* operon as is found on the *F. acidarmanus* chromosome. Following the development of this two-gene operon, it was thought that *arsC* was added to provide resistance to both arsenate and arsenite. Finally, the addition of *arsDA* would have conferred an even higher degree of As resistance and resulted in the modern *arsRDABC* operon seen in the *K. oxytoca, Acidiphilum multivorum,* and *E. coli* plasmids as well as the *Sinorhizobium* sp. As4 chromosome.

The present-day *arsA* gene is thought to be the result of gene duplication and fusion followed by recruitment as the catalytic subunit of the ArsB arsenite efflux protein (Chen et al. 1986; Rosen 1999). Although the ArsA-related proteins containing only the A1-loop analogue are unlikely to function in the same manner as ArsA, they may represent the ancestor to the "modern" dimeric *arsA* genes. Prior to the gene duplication event, the coupling of two ArsA1-analogous monomers, such as those in *F. acidarmanus* and *T. acidophilum*, may have resulted in a functional ArsA1:A2 homologous protein.

Phylogenetic analyses of ArsA1- and A2-analogous protein fragments were carried out to provide additional views into the development of this gene. For full ArsA1:A2 proteins, the A1 and A2 loop regions (see Fig. 3) were extracted and aligned as separate sequences. For ArsA1-analogues, sequence information beyond the analogous A1 region was excluded prior to the alignment. The phylogenetic tree generated from protein sequence fragments corresponding to the A1 and A2 loops identifies a division between the partial ArsA1-analogous sequences and the fragments associated with full ArsA1:A2 proteins (Fig. 4). This distinct split indicates two main separate lines in *arsA*-analogue evolution

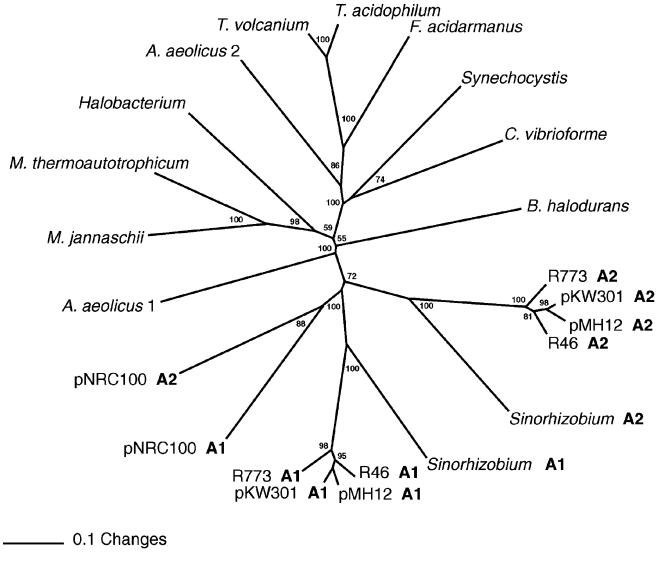


Fig. 4 Neighbor-joining tree showing the phylogenetic relationships of A1 and A2 fragments from ArsA-related and -homologous protein sequences. Percentages from 1,000 replicate bootstrapping analyses are shown near each branching point

corresponding to those genes borne on plasmids and those located on chromosomes (with the exception of *Sinorhizobium* sp. As4). Therefore, *arsA1*-analogous and full *arsA1:A2* genes appear to have differentiated at an early point into the plasmid and chromosomal formats. Duplication of the partial sequences to form the full *arsA1:A2* gene occurred in the plasmid system, while this event did not occur in the chromosomal *arsA1*-analogous genes. ArsA1-analogous sequences are found in both bacteria and archaea. Therefore, the duplication event most likely occurred after the divergence of the bacterial and archaeal domains.

The grouping of all ArsA1-analogous genes into one major clade is significant because it indicates that they are more similar to each other than to either of the A1 or A2 loops of the full ArsA1:A2 proteins. Proteins making up the ArsA1-analog clade, such as those of the archaeal methanogens (*Methanococcus jannaschii* and *Methanothermobacter thermoautotrophicum*) and the archaeal acidophiles (*T. volcanium*, *T. acidophilum*, and *F. acidarmanus*), have phylogenetic relationships consistent with those inferred for 16S rRNA genes, suggesting that their divergence is due to evolutionary descent. The placement of the other ArsA1-analogues such as those from *Aquifex aeolicus* (which has two ArsA1-analogous genes) may have resulted from lateral gene transfer events early in evolution of this gene.

Within the full ArsA1:A2 clade there is a deeply rooted split between the A1 and A2 loop fragments. The A1 regions of the *Sinorhizobium*, R46, pKW301, pMH12, and R773 ArsA proteins are all most similar to other A1 fragments and not to their own corresponding A2 fragments. Had the A1 sequences of individual genes been more similar to their A2 counterparts, one would argue that the duplication of A1 to form A1:A2 was a recent event. However, the deeply rooted division into A1 and A2 clades observed here indicates an early duplication event.

Phylogenetic analyses were compared for ArsB protein sequences and each organisms' respective 16S rDNA sequence. The ArsB phylogenetic tree, shown in Fig. 5, has a remarkably high degree of similarity to the 16S rDNA analysis represented in Fig. 6. The chromosomal and R46 plasmid-based ArsB sequences of *E. coli* are identical and lie in the same position on the ArsB tree. These sequences form a tight cluster with the plasmid-based ArsB sequences of *Acidiphilum multivorum* (pKW301), *K. oxytoca* (pMH12), *E. coli* (R773), and *Yersinia enterocolitica* (pYV1). This close grouping of *E. coli*, *K. oxytoca*, and *Y. enterocolitica* is found in both the ArsB and 16S rDNA trees. The topology of the branches containing the *Staphylococcus* and *Bacillus* sequences are also conserved in the ArsB and 16S rDNA analyses. Likewise, the organisms *Pseudomonas*

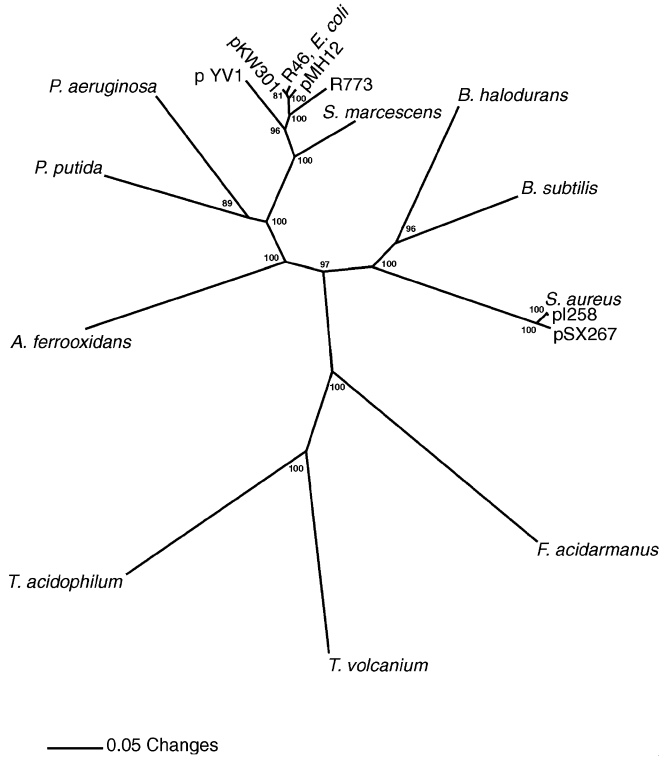


Fig. 5 Neighbor-joining tree based on ArsB protein phylogenetic analyses. Percentages from 1,000 replicate bootstrapping analyses are shown near each branching point

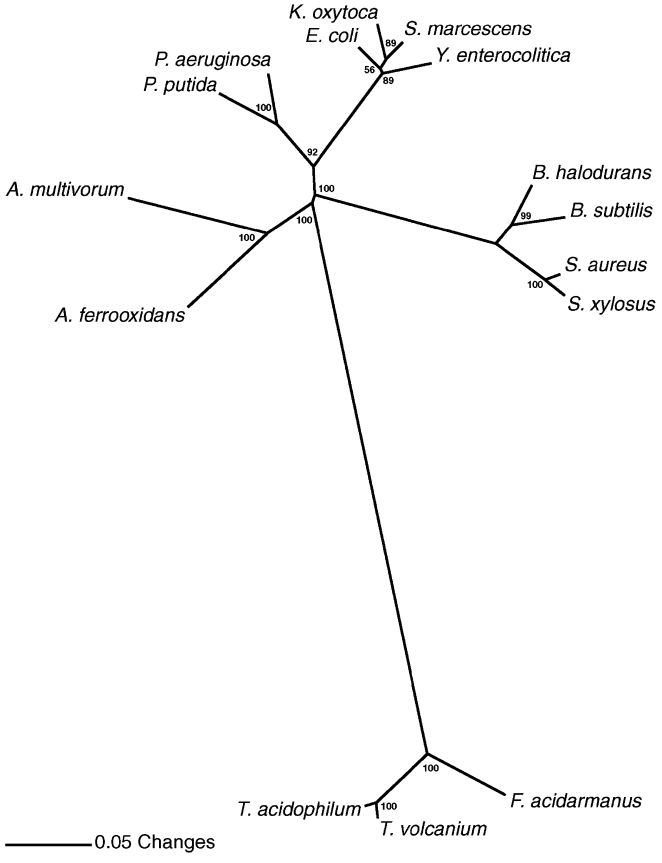


Fig. 6 Neighbor-joining phylogenetic tree for 16S rDNA sequences. Percentages from 1,000 replicate bootstrapping analyses are shown near each branching point

aeruginosa, *P. putida*, *Acidithiobacillus ferrooxidans*, *Thermoplasma acidiphilum*, *T. volcanium*, and *F. acidarmanus* lie in similar positions in both phylogenetic trees. The sole inconsistency is the position of *Acidiphilum multivorum* in the ArsB tree, which is likely the result of lateral gene transfer.

Conserved tree topologies strongly suggest a parallel evolutionary history for the *arsB* and 16S rRNA genes. Small subunit ribosomal RNA sequences, commonly used to determine relative phylogenetic positions of organisms, are well-conserved and ancient genes. Therefore, based on the analyses presented here, ArsB-encoding genes are also likely to have developed early in the evolution of life and are related through ancestry.

This new look at the evolution of arsenic resistance both supports and casts doubt upon previous notions on the development of these genes. Comparisons of ArsB and 16S rDNA evolutionary histories provide the strongest evidence to date demonstrating a very early appearance of the first arsenic-resistance genes. ArsA protein sequence analyses indicate a widespread misidentification of ArsA-related genes in nucleotide databases. New evidence and ideas presented here on the development of the full *arsA* gene and its addition to the *ars* operon will ultimately lead to a more complete picture of the evolutionary history of arsenic resistance. Considering that *F. acidarmanus* is resistant to elevated levels of arsenite and arsenate, it is likely that additional components of the arsenic-resistance pathway remain to be discovered.

Acknowledgements Funding was provided by grants from the National Science Foundation and the United States Environmental Protection Agency.

References

Alpers CN, Nordstrom DK, Thompson JM (1994) Seasonal variations of Zn/Cu ratios in acid mine water from Iron Mountain, California. In: Alpers CN, Blows DW (eds) Environmental geochemistry of sulfide oxidation. American Chemical Society, Washington, DC, pp 324–344

Altschul SF, Madden TL, Schaffer AA, Zhang JH, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Bhattacharjee HJ, Li J, Ksenzenko MY, Rosen BP (1995) Role of cysteinyl residues in metalloactivation of the oxyanion-translocating ArsA ATPase. J Biol Chem 270:11245–11250

Butcher BG, Deane SM, Rawlings DE (2000) The chromosomal arsenic resistance genes of *Thiobacillus ferrooxidans* have an unusual arrangement and confer increased arsenic and antimony resistance to *Escherichia coli*. Appl Environ Microbiol 66:1826–1833

Carlin A, Shi W, Dey S, Rosen BP (1995) The *ars* operon of *Escherichia coli* confers arsenical and antimonal resistance. J Bacteriol 177:981–986

Cervantes C, Ji G, Ramírez JL, Silver S (1994) Resistance to arsenic compounds in microorganisms. FEMS Microbiol Rev 15:355–367

Chen Y, Rosen BP (1997) Metalloregulatory properties of the ArsD repressor. J Biol Chem 272:14257–14262

Chen C, Misra TK, Silver S, Rosen BP (1986) Nucleotide sequence of the structural genes for an anion pump. The plasmid-encoded arsenical resistance operon. J Biol Chem 261:15030–15038

Dagnac T, Padró A, Rubio R, Rauret G (1999) Speciation of arsenic in mussels by the coupled system liquid chromatography—UV irradiation—hydride generation—inductively coupled plasma mass spectrometry. Talanta 48:763–772

Diorio C, Cai J, Marmor J, Shinder R, DuBow MS (1995) An *Escherichia coli* chromosomal ars operon homolog is functional in arsenite detoxification and is conserved in gram-negative bacteria. J Bacteriol 177:2050–2056

Edwards, KJ, Schrenk MO, Hamers R, Banfield JF (1998) Microbial oxidation of pyrite: experiments using microorganisms from an extreme acidic environment. Am Miner 83:1444–1453

Edwards KJ, Bond PL, Gihring TM, Banfield JF (2000) An archaeal iron-oxidizing extreme acidophile important in acid mine drainage. Science 287:1796–1799

Gladysheva TB, Oden KL, Rosen BP (1994) Properties of the arsenate reductase of plasmid R773. Biochemistry 33:7288–7293

Klaue B, Blum JD (1999) Trace analysis of arsenic in drinking water by inductively coupled plasma mass spectrometry: high resolution versus hydride generation. Anal Chem 71:1408–1414

Li J, Rosen BP (2000) The linker peptide of the ArsA ATPase. Mol Microbiol 35:361–367

Li J, Liu S, Rosen BP (1996) Interaction of ATP binding sites in the ArsA ATPase, the catalytic subunit of the Ars pump. J Biol Chem 271:25247–25252

McGuire MM, Edwards KJ, Banfield JF, Hamers RJ (2001) Kinetics, surface chemistry, and structural evolution of microbially mediated sulfide mineral dissolution. Geochim Cosmochim Acta 65:1243–1258

Ng WV, Ciufo SA, Smith TM, Bumgarner RE, Baskin D, Faust J, Hall B, Loretz C, Seto J, Slagel J, Hood L, DasSarma S (1998) Snapshot of a large dynamic replicon in a halophilic archaeon: megaplasmid or minichromosome? Genome Res 8:1131–1141

Peters S (2001) The origins and geochemical behavior of arsenic in a fractured bedrock aquifer, New Hampshire. Ph.D. Dissertation, University of Michigan

Rensing C, Ghosh M, Rosen BP (1999) Families of soft-metal-ion-transporting ATPases. J Bacteriol 181:5891–5897

Rosen BP (1999) Families of arsenic transporters. Trends Microbiol 7:207–212

Saha JC, Dikshit AK, Bandyopadhyay M, Saha KC (1999) A review of arsenic poisoning and its effects on human health. Crit Rev Environ Sci Technol 29:281313

Shi WP, Dong J, Scott RA, Rosen BP (1996) The role of arsenic-thiol interactions in metalloregulation of the *ars* operon. J Biol Chem 271:9291–9297

Silver S (1996) Bacterial resistances to toxic metal ions—a review. Gene 197:9–19

Smith RF, Smith TF (1990) Automatic generation of primary sequence patterns from sets of related protein sequences. Proc Natl Acad Sci USA 87:118–122

Smith RF, Smith TF (1992) Pattern-induced multi-sequence alignment (PIMA) algorithm employing secondary structure-dependent gap penalties for comparative protein modelling. Protein Eng 5:35–41

United States Environmental Protection Agency (2001) National Primary Drinking Water Standards. EPA 816-F-01–007

Wei X, Brockhoff-Schwegel CA, Creed JT (2001) A comparison of urinary arsenic speciation via direct nebulization and on-line photo-oxidation-hydride generation with IC separation and ICP-MS detection. J Anal At Spectrom 16:12–19

Wu J, Rosen BP (1991) The ArsR protein is a *trans*-acting regulatory protein. Mol Microbiol 5:1331

Zhou T, Rosen BP (1997) Tryptophan fourescence reports nuceotide-induced conformational changes in a domain of the ArsA ATPase. J Biol Chem 272:19731–19737