Submitted to Environmental Microbiology January 19, 2022

Revised version submitted 18 February, 2022

An ArsRC fusion protein enhances arsenate sensing and detoxification

Jian Chen, Adriana E. Galván, Venkadesh Sarkarai Nadar, Masafumi Yoshinaga and Barry P. Rosen[1]*

[1]Department of Cellular Biology and Pharmacology, Herbert Wertheim College of Medicine, Florida International University, Miami, Florida 33199, United States

*Correspondence: Barry P. Rosen, Florida International University Herbert Wertheim College of Medicine, 11200 SW 8th Street, Miami, FL 33199 Tel: (+1) 305-348-0657, Fax: (+1) 305-348-0651, Email: brosen@fiu.edu, ORCID 0000-0002-5230-4271.

Abbreviations: *ars*, arsenical resistance operons; arsenite, As(III); arsenate, As(V); 1As3PGA, 1-arseno-3-phosphoglycerate; HPLC, high pressure liquid chromatography; ICP-MS, inductively coupled plasma mass spectroscopy.

Running Title: An ArsRC fusion for arsenate sensing and detoxification

25 Keywords: arsenite, arsenate reductase, repressor, fusion protein, channeling, arsenic
26 biosensor
27

Originality-Significance Statement

Arsenic biotransformations and detoxification allow for life in the presence of this pervasive environmental toxin. Arsenate is the most prevalent species in an oxic environment, but selective arsenate biosensing has not been identified until this study, Here we report identification of a fusion gene between an arsenite sensor and arsenate reductase that enhances the ability of the cells to sense arsenate.

Summary

Arsenical resistance (*ars*) operons encode genes for arsenic resistance and biotransformation. The majority are composed of individual genes, but fusion of *ars* genes are not uncommon, although it is not clear if the fused gene products are functional. Here we report identifieation of a four-gene *ars* operon from *Paracoccus* sp. SY that has two *arsR-arsC* gene fusions. ArsRC1 and ArsRC2 are related proteins that consist of an N-terminal ArsR arsenite (As(III))-responsive repressor with a C-terminal ArsC arsenate reductase. The other two genes in the operon are *gapdh* and *arsJ.* GAPDH*,* glyceraldehyde 3-phosphate dehydrogenase, forms 1-arseno-3-phosphoglycerate (1As3PGA) from 3-phosphoglyceraldehade and arsenate (As(V)), ArsJ is an efflux permease for 1As3PGA that dissociates into extracellular As(V) and 3-phosphoglycerate (3PGA). The net effect is As(V) extrusion and resistance. ArsRs are usually selective for As(III) and do not respond to As(V). However, the substrates and products of this operon are pentavalent, which would not be inducers of the operon. We propose that ArsRC fusions overcome this limitation by channeling the ArsC product into the ArsR binding site without diffusion through the cytosol, a *de facto* mechanism for As(V) induction. This novel mechanism for arsenate sensing can confer an evolutionary advantage for detoxification of inorganic arsenate.

Introduction

Arsenic is a naturally occurring metalloid that is widely distributed in the Earths crust. Natural and anthropogenic activities increase arsenic contamination in water, air, food and soil (Zhu et al., 2014). Consequently, microbes have evolved various strategies for coping with environmental arsenic, including transport and biotransformations such as redox and methylation cycles (Mathivanan et al., 2021). The genes for these pathways are usually clustered in *ars* operons (San Francisco et al., 1990). Most *ars* operons are controlled by ArsR arsenite (As(III))-responsive transcriptional repressors (Wu and Rosen, 1993). In most *ars* operons, intracellular trivalent As(III) binds to ArsR, producing a conformation change that dissociates the repressor from the *ars* promoter and transcription of the *ars* operon genes (Rosen, 2002). Arsenic and sulfur were abundant in the anoxic Archean environments, and are still found in volcanoes, hot springs, fumaroles and geothermal wells (Herath et al., 2016). Consequently resistance pathways for arsenic transport and biotransformation are believed to have emerged early in the evolution of life on Earth, when reduced trivalent As(III) would have predominated (Chen et al., 2020). In addition to ArsR, primordial *ars* operons probably encoded an As(III) efflux permease such as Acr3 and later ArsB. After the Great Oxidation Event (GOE) about 2.4 Bya, atmospheric oxygen increased, raising the redox potential of the Earth's surface and oceans (Oremland and Stolz, 2003). Intensive oxidative weathering of arsenic-bearing minerals fundamentally changed the arsenic species in rivers and oceans from trivalent As(III) to pentavalent arsenate (As(V)). Since Acr3 and ArsB transport As(III) and not As(V), early life evolved arsenate reductase enzymes to reduce As(V) to As(III), the substrate of the efflux systems. Arsenate reductases arose by convergent evolution at least four times(Mukhopadhyay et al., 2002). Most of the reductase genes appeared after the GOE, but the birth date of one, termed ArsC2, is predicted to predate the GOE (Chen et al., 2020). This enzyme, typified by *Staphylococcus aureus* plasmid pI258 ArsC, is a small protein (13-15 kDa) that efficiently

78 catalyze reduction of As(V) to As(III) prior to extrusion (Ji and Silver, 1992). ArsC catalyzes
79 reduction of As(V) coupled to thioredoxin, thioredoxin reductase and NADPH (Ji et al., 1994).
80 While arsenate reductases convert arsenate into arsenite, the inducer of the *ars* operon,
81 transfer of arsenite from the reductase active site to the ArsR binding site is diffusion-limited,
82 which would slow the induction process. Under most environmental conditions there would be
83 sufficient As(III) for induction of As(III) efflux permeases, so this might not be a major issue for
84 many species of bacteria.

85 However, there is a class of *ars* operons found in many bacteria that have pentavalent
86 substrates and products, with no trivalent arsenicals available for induction of *arsR*. In these
87 operons there are two genes, *gapdh* and *arsJ*, that encode a coupled pathway for arsenate
88 resistance. The *gapdh* gene codes for the enzyme glyceraldehyde 3-phosphate dehydrogenase
89 (GAPDH). Most cells have a related GAPDH enzyme that plays a critical role in glycolysis. It
90 catalyzes condensation of phosphate with glyceraldehyde-3-phosphate (G3P) to produce 1,3-
91 bisphosphoglycerate (1,3-DPG), an intermediate in the production of pyruvate and ATP.
92 Arsenate is an uncoupler of energy production primarily because it substitutes for phosphate in
93 the GAPDH reaction, forming the unstable organoarsenical 1-arseno-3-phosphoglycerate
94 (1As3PGA), which hydrolyzes to G3P and arsenate with a half-life of 2 sec (Byers et al., 1979).
95 This is a primary cause of arsenate toxicity. Bacteria couple a *gapdh* gene with *arsJ*, which
96 encodes a very efficient efflux permease for 1As3PGA (Chen et al., 2016). When ArsJ extrudes
97 1As3PGA, the unstable organoarsenical rapidly hydrolyzes to form extracellular As(V), resulting
98 in As(V) detoxification. In effect, the couple of GAPDH and ArsJ form an arsenate efflux system.
99 The *gapdh-arsJ* genes are controlled by an ArsR repressor, but, as pointed out, the reaction
100 does not include trivalent arsenicals, so an ArsC enzyme is required to produce the inducer by
101 reduction of As(V).

To overcome the diffusion limitation of separate ArsC and ArsR reactions, some bacteria have created fusions between the two genes, producing chimeric ArsRC proteins that are predicted to channel the product of the ArsC reaction into the ArsR As(III) binding site. As more and more bacterial genomes are sequenced, atypical *ars* fusion genes have been identified in various *ars* operons (Ben Fekih et al., 2018). Combining *arsR* and *arsC* genes is one of the more commonly observed gene fusions and is found in many bacteria. An *arsRC* fusion gene was first reported in arsenic resistant strain *Microbacterium* sp. A33 (Achour-Rokbani et al., 2010). Micobacterial *arsRC* was shown to regulate *ars* gene expression and be induced by As(V), As(III) and Sb(III). However, the arsenate reductase activity of the ArsRC fusion protein was not described, and its physiological function was not investigated further.

In this study we investigated whether the *Paracoccus* sp. SY *arsRC* fusion has a physiological role in arsenate resistance. An NCBI Conserved Domains analysis indicates that the 276-residue ArsRC1 fusion protein (WP_103174508) has two domains connected by a linker region of approximately 12-15 residues (Marchler-Bauer et al., 2017). The first 102 residues form a domain related to the AfArsR repressor (Qin et al., 2007). Residues 115-250 form a domain that is related to the *S. aureus* pI258 ArsC arsenate reductase (Ji et al., 1994). Cells of *Escherichia coli* heterologously expressing the *arsRC1* gene exhibit low levels of As(V) reduction, but transcription of the operon responds to both As(III) and As(V). The gene was divided into two portions corresponding to *arsR* and *arsC.* Importantly, the cells expressing the intact *arsRC1* gene responded to lower concentrations of arsenate than cells expressing just the $arsR_N$ sequence. Purified ArsRC1 exhibited low rates of arsenate reduction compared with the purified ArsC domain, suggesting that the function of the ArsC domain of the chimeric protein is not related to bulk arsenate reduction but rather to reduce sufficient arsenate to induce operon expression. The increase in efficiency exhibited by the *arsRC1* gene and purified protein compared with the separate domains is consistent with channeling of the product of the ArsC

127 domain to the binding site of the ArsR domain. This improvement in responsiveness to As(V)
128 may present an evolutionary advantage to cells expressing the *gapdh-arsJ* arsenate resistance
129 genes.
130
131 **Results**
132 ***Occurrence of ars fusion genes and a unique arsRC1-arsRC2-gapdh-arsJ ars operon in***
133 ***Paracoccus sp. SY***
134 Concurrent with the rapid development of sequencing technologies, the number of
135 sequenced genomes has increased dramatically in the last decade. There are now more than
136 30,000 sequenced bacterial genomes currently publicly available in NCBI database in 2014
137 (Land et al., 2015). Comparative genetic analysis of chromosomal *ars* operons from more than
138 50 bacterial species has revealed numerous fusions of *ars* genes (representative examples are
139 shown in Fig. 1). A gene fusion between a 5'-aquaglyceroporin As(III) channel gene (*aqp*) and a
140 3-'*arsC* gene confers As(V) resistance in *Salinispora tropica* (Wu et al., 2010). In that study the
141 authors proposed that the fusion protein would reduce As(V) to As(III) in contact with the
142 channel, so that efflux could take place without release of toxic As(III) in the cytosol. In
143 *Microbacterium* sp. strain A33, an *arsRC* fusion gene encodes a transcriptional repressor that is
144 responsive to As(V) in addition to As(III) (Achour-Rokbani et al., 2010). The authors
145 hypothesized that the ArsC domain of the fusion protein would catalyze reduction of As(V) to
146 As(III), which would bind to the ArsR domain without diffusion through the cytosol, an efficient
147 channeling mechanism for arsenate sensing.
148 Here we examine the hypothesis that As(III) is channeled from ArsC directly into ArsR
149 using an *arsRC* gene fusion from *Paracoccus* sp. SY. This environmental isolate, from an
150 arsenic-contaminated paddy soil (Zhang et al., 2015), is a chemoautotrophic As(III)-oxidizing
151 bacterium that exhibits high resistance both to inorganic and organic arsenicals (Chen et al.,

2021). It has an arsenite oxidase (*aio*) gene cluster in the chromosome (KP881606) and efficiently oxidizes As(III) under both aerobic and anaerobic conditions using either O_2^- or NO_3^- as electron acceptor. The genes for a phosphate (Pi) uptake system (Pst) (Hudek et al., 2016) are located adjacent to a novel four-gene *ars* operon. The Pst system is highly efficient at transporting phosphate, particularly under low-phosphorus conditions (Rao and Torriani, 1990). Arsenate, as an analog of phosphate, can be readily taken up by *Paracoccus* sp. SY. The downstream *ars* operon is composed of two As(V) detoxification genes, *gapdh* and *arsJ*, which are regulated by two *arsR-arsC* fusion genes (Fig.1). Similar *arsRC-gapdh-arsJ* genes clusters are found in many bacteria (Fig.S1). Since the genes are linked, it seemed reasonable to consider that the ArsRC proteins could be As(V) responsive repressors that regulate expression of GAPDH and ArsJ for As(V) detoxification.

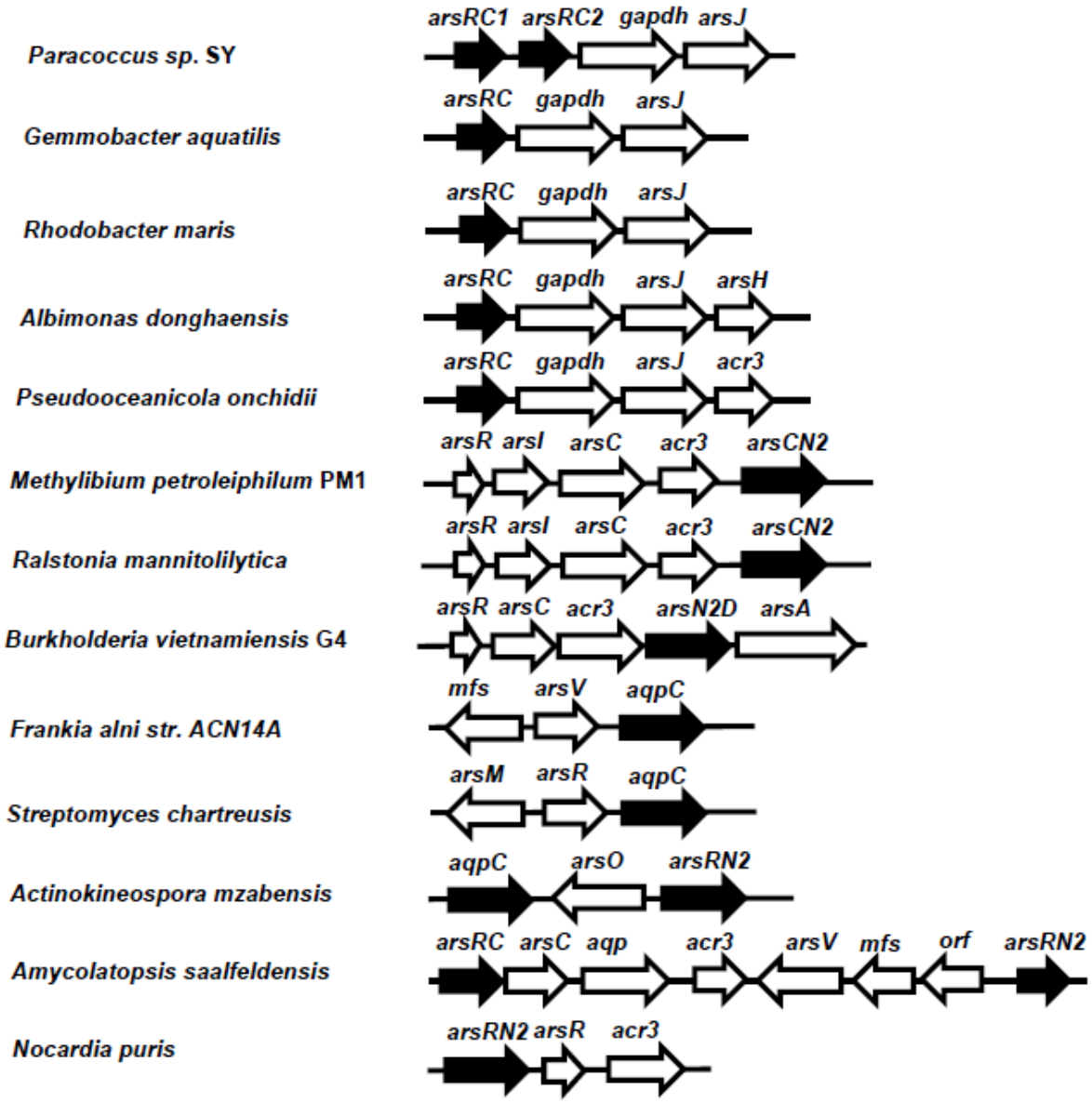
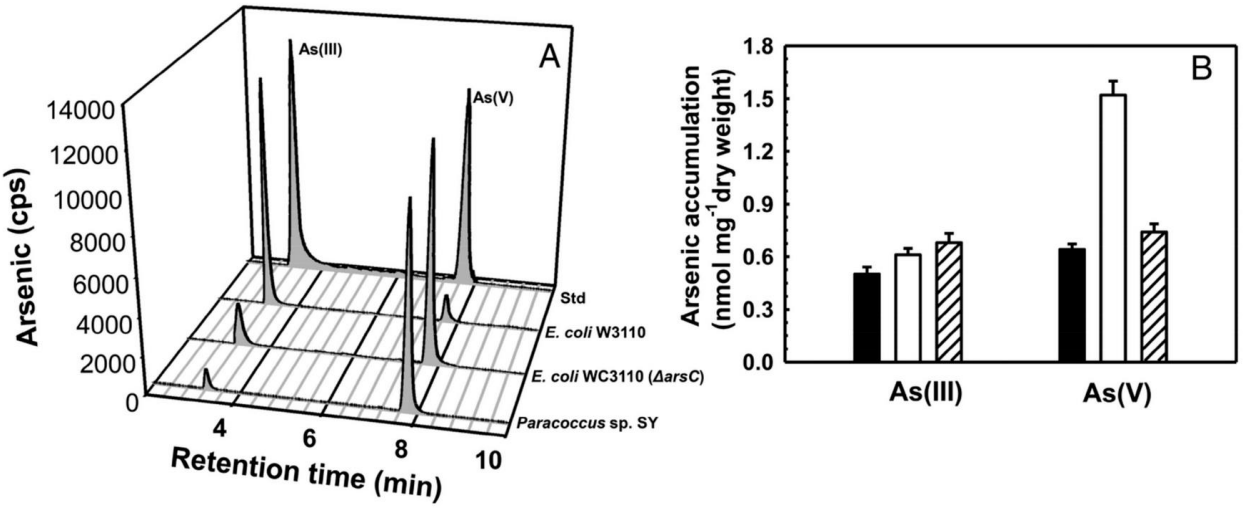


Fig. 1. Distribution of various *ars* fusion genes in *ars* operons. Shown are representative *ars* operons (accession numbers in parentheses) containing *ars* gene fusions (black fill). *Paracoccus* sp. SY (NZ_NWMQ01000040.1), *Gemmobacter aquatilis* (NZ_FOCE01000001.1), *Rhodobacter maris* (NZ_OBMT00000000.1), *Albimonas donghaensis* (NZ_FNMZ00000000.1), *Pseudooceanicola onchidiid* (NZ_SPUW00000000.1), *Methylibium petroleiphilum* PM1 (CP000555), *Ralstonia mannitolilytica* (NZ_CP010799.1), *Burkholderia vietnamiensis* G4 (CP000615), *Paraburkholderia kururiensis* M130 (NZ_ANSK01000005.1), *Frankia alni str.*

ACN14A (CT573213), *Streptomyces chartreusis* (NZ_CP023689.1), *Actinokineospora mzabensis* (NZ_QHCP00000000.1), *Amycolatopsis saalfeldensis* (NZ_FOEF00000000.1), *Nocardia puris* (NZ_QNRE00000000.1),

Arsenate reduction and accumulation in Paracoccus sp. SY

Paracoccus sp. SY is an As(III)-oxidizing bacterium with a Pst phosphate uptake system. For that reason biotransformation and accumulation of arsenate by *Paracoccus* sp. SY were examined and compared with *E. coli* W3110 and its arsenate-hypersensitive derivative WC3110 (*ΔarsC*) (Mukhopadhyay et al., 2000) (Fig. 2). As(V) was nearly entirely reduced to As(III) by *E. coli* W3110, which expresses ArsC. Cells of either the *arsC* deletion strain *E. coli* WC3110 or *Paracoccus* sp. SY exhibited only very low levels of As(V) reduction (Fig. 2A). The small amounts of As(III) produced in *E. coli* WC3110 are probably the result of nonenzymatic reduction by intracellular thiols such as reduced glutathione (Delnomdedieu et al., 1994). It is not clear why *Paracoccus* sp. SY exhibits low rates of As(V) reduction since it has two *arsC* genes (Chen et al., 2021). One possibility is that As(III) is rapidly oxidized by the Aio arsenite oxidase (Zhang et al., 2015). When cultured with As(III), cells of *Paracoccus* sp. SY accumulated significantly more As(III) than *E. coli* strains W3110 or WC3110, which again may be the consequence of As(III) oxidation by the Aio system in *Paracoccus* sp. SY (Fig. 2B). *E. coli* has ArsB, which extrudes As(III) but not As(V), so *E. coli* W3110, which reduces As(V) to As(III), accumulates less total arsenic than *E. coli* WC3110, which cannot reduce As(V) (Fig. 2B). On the other hand, cells of *Paracoccus* sp. SY accumulated lower amounts of As(V) compared with *E. coli* WC3110 due to the synergistic action of GAPDH and ArsJ resulting in As(V) efflux (Fig. 2B).



195
196 **Fig. 2. As(V) reduction and accumulation in *Paracoccus sp.* SY**. **(A) As(V) reduction and**
197 **(B) accumulation by *Paracoccus sp.* SY.** Overnight cultures of *Paracoccus* sp. SY (hatched
198 bar), *E. coli* wild-type W3110 (black bar) and the arsenate-hypersensitive strain
199 WC3110(*ΔarsC*) (white bar) were grown in low phosphate medium containing As(III) or As(V) at
200 30 °C with shaking. Arsenic accumulation was determined by ICP-MS, and species were
201 determined by HPLC-ICP-MS, as described under Experimental Procedures. Data are the mean
202 ± SE (n = 3).
203
204 ***ArsRC1 responds to both As(III) and As(V)***
205 The wide occurrence of *arsRC-gapdh-arsJ* operons suggest that non-canonical *arsRC* genes
206 are potentially co-transcribed with As(V) resistance determinants. In *Paracoccus* sp. SY, *gapdh*
207 and *arsJ* expression are regulated by two ArsRC fusion proteins. ArsRC1 (WP_103174508)
208 shows 49% identity with ArsRC2 (WP_103174507), and we focused on ArsRC1 in this study. It
209 consists of an N-terminal ArsR repressor with a C-terminal As(V) reductase (Fig. 3A). ArsRC1
210 has only 23% identity with ArsRC (CAT03228) from *Microbacterium* sp. strain A33. In the latter,
211 the ArsR domain has a CVC sequence that forms the As(III) binding site in *E. coli* R773 ArsR
212 (Xu et al., 1996). In contrast, the ArsRC1 ArsR domain has sequence similarity with AfArsR

213 from *Acidithiobacillus ferrooxidans*, which has a three-cysteine As(III) binding site in the C-
214 terminal region (Qin et al., 2007). Both ArsRC1 and ArsRC2 have C-terminal domains similar to
215 thioredoxin-dependent ArsC arsenate reductases (Zegers et al., 2001) (Fig. 3A). To determine
216 the substrate specificity and sensitivity of ArsRC1, a two-plasmid biosensor was constructed
217 that utilizes the *arsRC1* gene under control of the *lac* promoter in one plasmid (pETduet-
218 *arsRC1*), and the *ars* promoter controlling *gfp* expression in the other (pACYC184-P_{arsRC1}-*gfp*)
219 (Fig.S2A) using the system described previously (Chen et al., 2012). In cells of *E. coli*
220 AW3110(*Δars*) bearing both plasmids, addition of isopropyl-β-d-1-thiogalactopyranoside (IPTG)
221 drives expression of ArsRC1, which binds to the *ars* promoter and inhibits expression of *gfp*
222 from pACYC184-P_{arsRC1}-*gfp* (Fig. S2B). Unlike characterized ArsRs, which bind only trivalent but
223 not pentavalent arsenicals (Chen and Rosen, 2014), the ArsRC1 biosensor is responsive to
224 As(V) as well as As(III) (Fig. 3B). A construct with only the ArsR domain of ArsRC1 ($arsR_N$) no
225 longer responded to As(V), indicating that the ArsC domain is required for As(V) induction. The
226 ArsRC1 biosensor exhibited a similar concentration dependence for both As(III) and As(V) (Fig.
227 3C and D). The response at low concentrations (<5 μM) indicates that the apparent affinity of
228 the chimeric protein was similarly for both As(III) and As(V) (Fig. 3C). At higher concentrations
229 of As(V), the fluorescence intensity continued to increase with As(III) but not As(V) (Fig.3D),
230 which could be due to saturation of the ArsC domain with substrate.

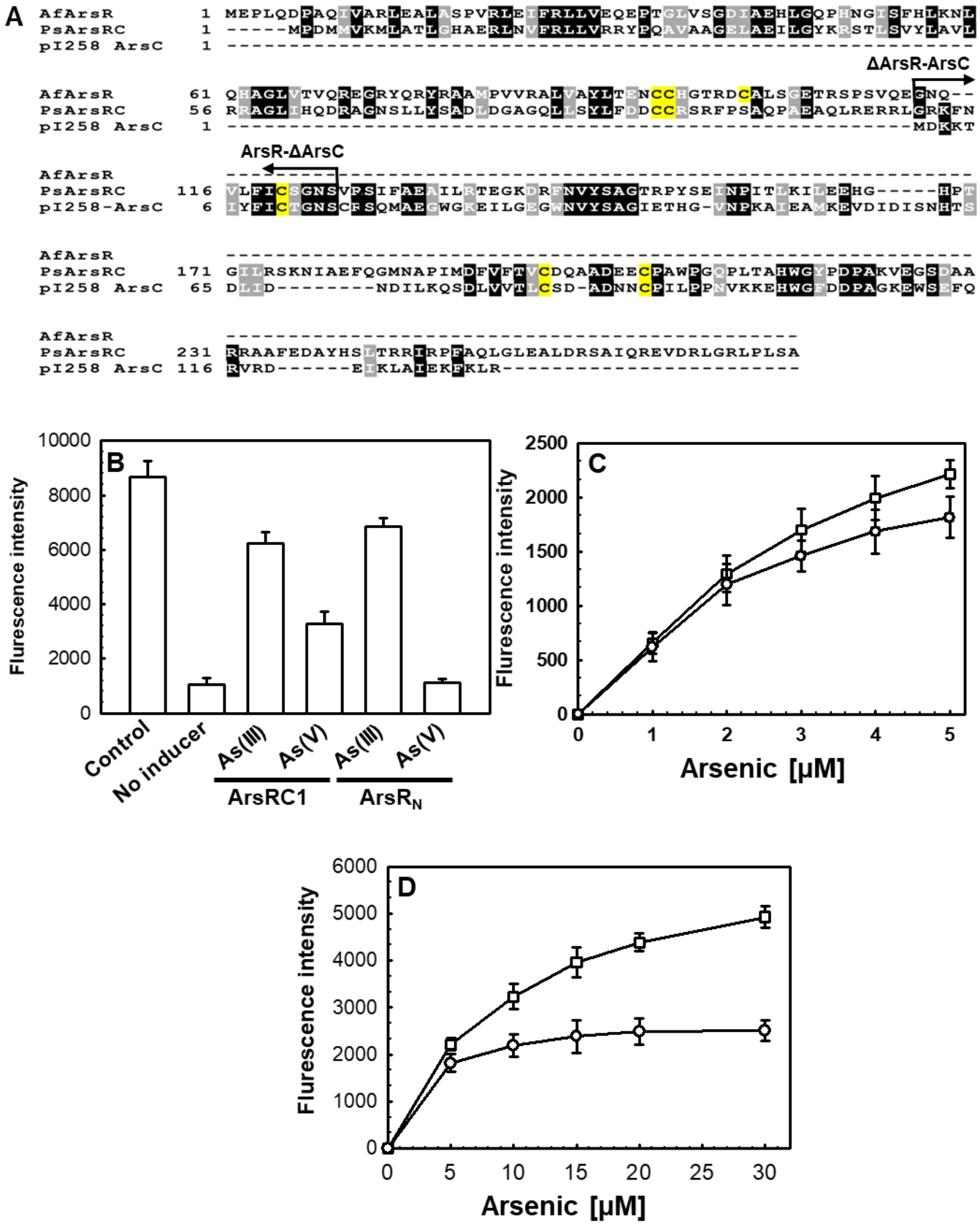

A
AfArsR
PsArsRC
pI258 ArsC
ΔArsR-ArsC
ArsR-ΔArsC
B
Flurescence intensity
Control
No inducer
As(III)
As(V)
As(III)
As(V)
ArsRC1
ArsR$_N$
C
Flurescence intensity
Arsenic [μM]
D
Flurescence intensity
Arsenic [μM]

Fig. 3. The *arsRC1* biosensor responds to both As(III) and As(V). A. **Alignment of ArsRC1 with *A. ferrooxidans* AfArsR and *S. aureus* plasmid pI258 ArsC.** The protein sequence of ArsRC1 from *Paracoccus* sp. SY (WP_103174508) was compared with *A. ferrooxidans* AfArsR (WP_012537435) and *S. aureus* pI258 plasmid ArsC (AYK28273). The locations of N-terminal and C-terminal truncations are labeled by arrows, forming $ArsR_N$ (ArsR-ΔArsC) and $ArsC_C$ (ΔArsR-ArsC), respectively. **B. The biosensor with the *arsRC1* gene responds both As(III) and As(V)**. Conditions for constitutive, repressed or derepressed *gfp* expression are shown (Fig. S2). Cells of *E. coli* strain AW3110(DE3) bearing plasmids encoding full-length ArsRC1 or only the N-terminal domain ($ArsR_N$) *in trans* with reporter plasmid pACYC184-P_{arsRC1}-*gfp* were grown in low phosphate medium for 14 hr with 0.2% glucose as carbon source, 0.3 mM IPTG and 20 μM of either As(III) or As(V). **C and D. Comparison of the response of the bacterial biosensor to low (C) or high (D) concentrations of As(III) (□) or As(V) (O) .** The response of cells expressing the intact *arsRC1* gene is shown at the indicated concentrations of inducer. Cellular expression of the *gfp* reporter gene was assayed as described under *Experimental Procedures*. Fluorescence intensities of cell suspensions were quantified using a Photon Technology International spectrofluorometer with an excitation wavelength of 470 nm and emission wavelength of 510 nm. The data are the mean ± SE (n = 3).

ArsRC1 reduces As(V) to As(III) in vivo and in vitro

To further investigate catalytic properties of ArsRC1, reduction of arsenate by ArsRC1 was assayed in cells and with purified protein. The *arsRC1* gene and its derivatives were cloned individually into vector plasmid pET29 under control of the T7 promoter, creating plasmids pET29-*arsRC1* and pET29-$arsC_C$ (Table S1). *E. coli* plasmid R773 ArsC was used as a positive control (Liu and Rosen, 1997). Plasmids were individually expressed in the arsenate-sensitive *E. coli* strain WC3110 (*ΔarsC*) (Mukhopadhyay et al., 2000). Cells with plasmid pET29-$arsC_{R773}$

and pET29-*arsC*$_C$ exhibited significant resistance to As(V) compared with the same strain with vector plasmid pET29 (Fig. 4A). However, cells expressing pET29-*arsRC1* conferred little resistance to As(V). These results indicate that the ArsC domain of the chimeric protein is active as a separate protein but that the activity is suppressed to levels just sufficient for induction when it is a domain of the ArsRC1 protein.

To test this idea, the ArsRC1 protein and the ArsC$_C$ domain were heterologously expressed in *E. coli,* purified and assayed for arsenate reductase activity using thioredoxin (Trx) as reductant. Purified ArsRC1 converted only 20% of the arsenate into arsenite in 4 h (Fig. 4B). In contrast, the purified ArsC$_C$ domain exhibited four-fold higher enzymatic activity, converting nearly 80% of substrate into product during the same time period. These results are consistent with the results of the biosensor and resistance assays and demonstrate that the enzymatic activity of the ArsC$_C$ domain is suppressed in the ArsRC1 protein. We hypothesize that direct interaction of the two domains reduces the rate of ArsC-catalyzed production of As(III) to allow only enough to be produced to activate the ArsR domain.

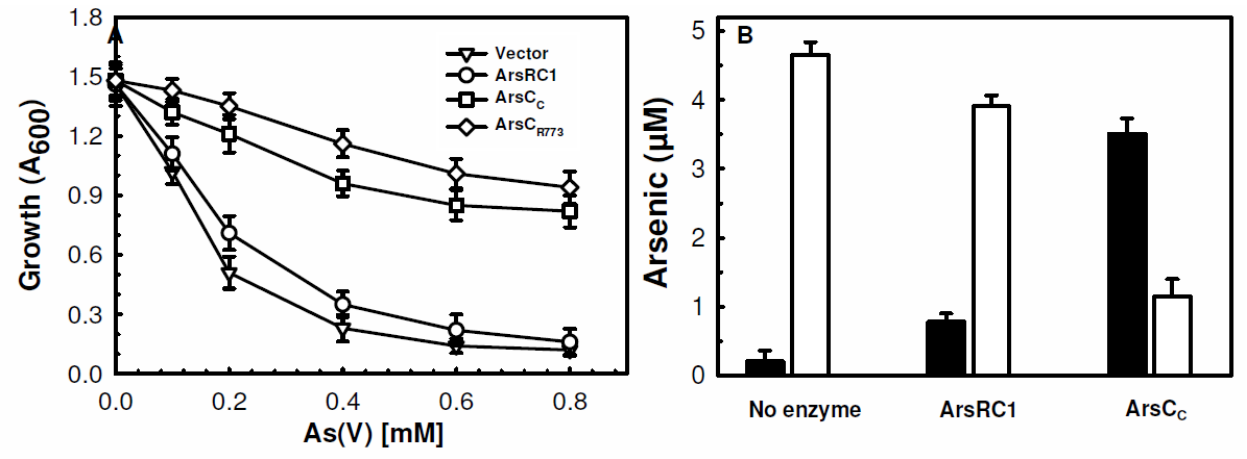
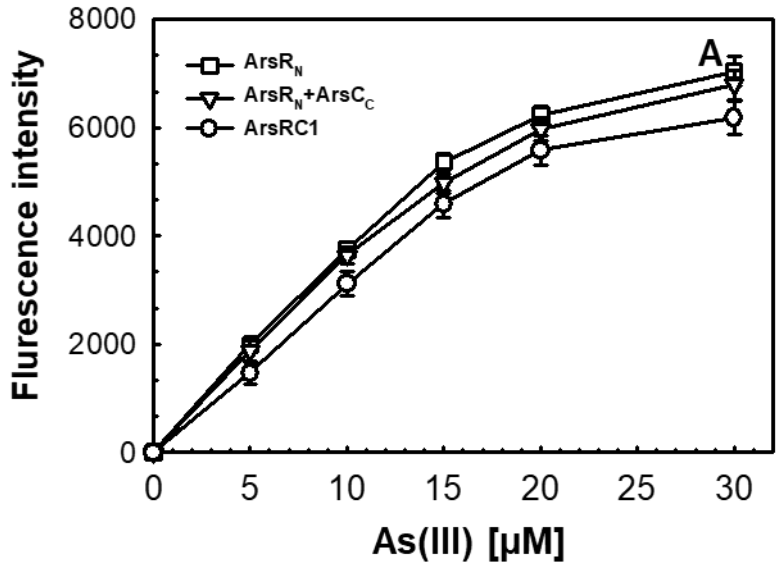


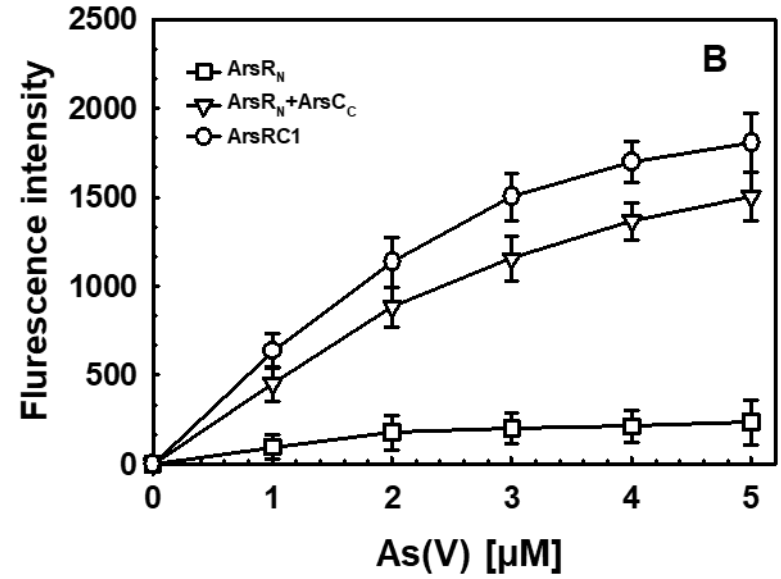
Fig. 4. ArsRC1 arsenate reduction. **A. Expression of ArsRC1 in cells of *E. coli* WC3110.** Overnight cultures bearing vector plasmid pET29a (▽), pET29a-ArsRC1 (○), pET29a-ArsC$_C$ (□) or pET29a-ArsC$_{R773}$ (◊) were diluted 100-fold into fresh low phosphate medium containing

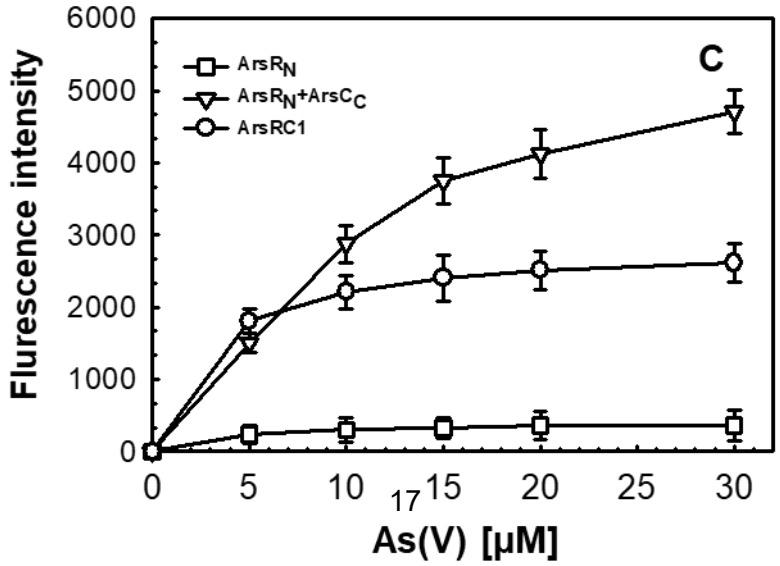
the indicated concentrations of As(V), and expression was induced with 0.3 mM IPTG. Growth was measured after 16 h at 30 °C. Data are the mean ± SE (n = 3). **B. As(V) reduction of As(V) by purified ArsRC1 and the $ArsC_C$ domain**. Reduction of As(V) (white bar) to As(III) (black bar) by purified ArsRC1 and $ArsC_C$ was assayed as described in *Experimental Procedures*. After 4 hr, arsenic species were separated by HPLC and quantified by ICP-MS, as described in *Experimental Procedures*. Data are the mean ± SE (n = 3).

ArsRC1 responds to As(V) with higher efficiency than the $ArsR_N$ domain alone

To compare the arsenic responsiveness of cells expressing the intact *arsRC1* gene to cells co-expressing the separated domains, sequences for the N and C-terminal domains of ArsRC1 were separately amplified, cloned both either individually or together in vector plasmid pETDuet-1 for biosensor assays (Fig. S2A). The results show that the response to As(III) was similar in cells expressing either *arsRC1* or $arsR_N$. Co-expression of $arsR_N$ and $arsC_C$ did not change the responsiveness to As(III) (Fig. 5A). In contrast, the response to As(V) was quite different. Cells expressing *arsRC1* responded to As(V), but cells expressing $arsR_N$ did not (Fig. 5B). However, cells expressing $arsR_N$ together with *arsCc* responded to As(V) (Fig. 5B) and were even more responsive at high concentrations of As(V) (Fig. 5C). These results indicate that $ArsC_C$ efficiently reduces As(V) to As(III), which binds to $ArsR_N$ and induces *gfp* expression in the biosensor cells. Although cells expressing *arsRC1* showed low As(V) reduction and are not sensitive to As(V) at high concentrations, it should be noted that cells expressing *arsRC1* are slightly more responsive to As(V) at low concentrations compared to cells co-expressing the two halves of the gene (Fig. 5B). ArsRC1 controls arsenate resistance genes *gapdh* and *arsJ*, which suggests that the $ArsC_C$ domain is not involved in reduction of bulk As(V) but, instead, controls arsenate resistance by direct interaction with the $ArsR_N$ domain.


A
ArsR$_N$
ArsR$_N$+ArsC$_C$
ArsRC1
Flurescence intensity
As(III) [μM]
0
2000
4000
6000
8000
0
5
10
15
20
25
30


B
ArsR$_N$
ArsR$_N$+ArsC$_C$
ArsRC1
Flurescence intensity
As(V) [μM]
0
500
1000
1500
2000
2500
0
1
2
3
4
5


C
ArsR$_N$
ArsR$_N$+ArsC$_C$
ArsRC1
Flurescence intensity
As(V) [μM]
0
1000
2000
3000
4000
5000
6000
0
5
10
15
20
25
30

300

Fig. 5. Comparison of the response of the biosensor cells expressing the *arsRC1* gene with cells expressing $arsR_N$ with or without $arsC_C$. A. Response of biosensor cells with *arsRC1* constructs to As(III). The results show that the $ArsC_C$ domain is not required for As(III) sensing. **B and C. Response to low (B) or high (C) concentrations of As(V).** Cells expressing *arsRC1* respond to As(V), while those with $arsR_N$ do not. Cells expressing intact *arsRC1* are slightly more responsive to As(V) at low concentrations than cells co-expressing the two halves of the gene. At higher concentrations the response of cells with the *arsRC1* saturates, while cells expressing $arsC_C$ continue to increase. Fluorescence intensities were quantified by spectrofluorometry. The data are the mean ± SE (n = 3).

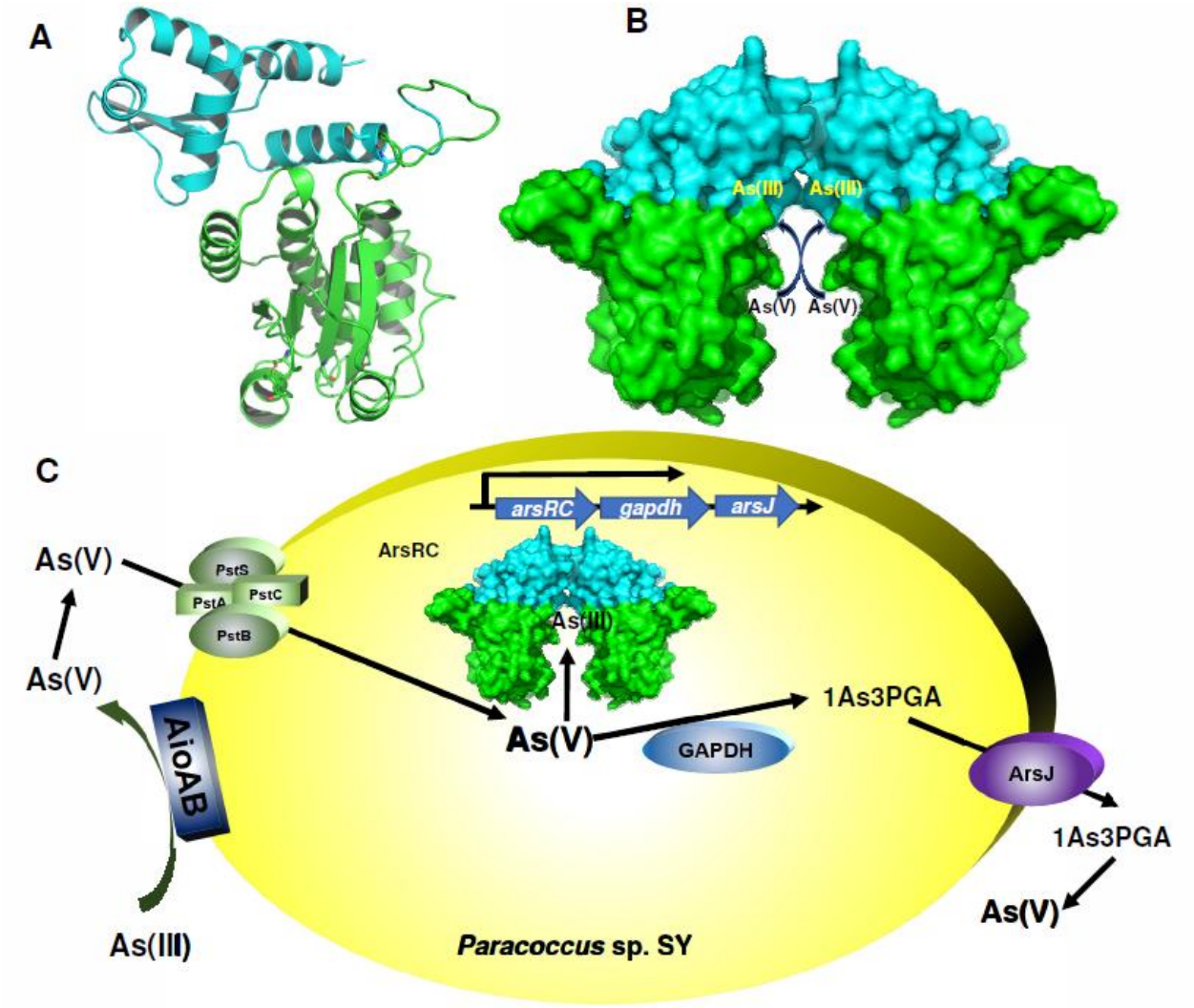
Discussion

In primordial anoxic oceans As(III) would have been the predominant arsenic species. To cope with As(III) toxicity, bacteria evolved arsenite resistance genes controlled by ArsR As(III)-responsive repressors. After the GOE, arsenate dominated in aerobic environments, which reduced the concentrations of the inducer, As(III). In response, microbes expanded their repertoire of *arsC* genes to allow for reduction of and resistance to As(V). Other arsenate resistances evolved such as the synergistic coupling of GAPDH and ArsJ (Chen et al., 2016). Some *gapdh-arsJ* operons are controlled by *arsR-arsC* gene fusions. GAPDH catalyzes formation of 1-arseno-3-phosphoglycerate (1As3PGA), which is extruded from cells by ArsJ, conferring arsenate resistance (Chen et al., 2016). The wide distribution of *arsRC-gapdh-arsJ ars* operons suggests that the ArsRC repressors have a selective role in As(V) detoxification and confer a selective advantage to their host.

An ArsRC fusion was first reported in *Mycobacterium.* sp. Strain A33, which is induced by both As(III) and As(V) (Achour-Rokbani et al., 2010). However, neither the ArsR nor ArsC domains are related to those in *Paracoccus* sp. SYArsRN1, and, in that study the As(V)

reductase activity of the chimeric proteins and its role in As(V) responsiveness was not characterized. *Paracoccus* sp. SY ArsRC1 is an ArsR and ArsC fusion protein with conserved cysteine residues in each domain (Fig. 6A). Like other members of the SmtB/ArsR family of metalloregulatory transcriptional repressors (Busenlehner et al., 2003), ArsRC1 is proposed to be a homodimer, in which the ArsC domain reduces As(V) to As(III) that is channeled directly into the ArsR domain for induction (Fig. 6B).

The *arsRC1* gene conferred little resistance to As(V) compared with the $arsC_C$ portion of the gene expressed separately (Fig. 4A), and purified ArsRC1 reduced four-fold less As(V) than purified $ArsC_C$. This correlates with derepression, where the *arsRC1* gene produced somewhat more *gfp* expression at low As(V) concentrations than co-expression of the individual domains (Fig. 5B). Thus arsenate sensing is more efficient with *arsRC1* than with individual $arsR_N$ and $arsC_C$. Even though the differences were not large in this experiment, small differences in a phenotype can result replacement by a new variant after multiple generations, as was shown for the competitive advantage conferred by the ArsD As(III) chaperone (Lin et. al, 2006). As an autotrophic arsenite-oxidizing microbe, *Paracoccus* sp. SY efficiently oxidizes As(III) to As(V), and the *gapdh-arsJ* operon confers As(V) resistance (Fig. 6C). We propose that the ability to respond to low concentrations of arsenate provides a competitive advantage to *Paracoccus* sp. SY and other bacteria with an *arsRC* fusion.



344

345 **Fig. 6. Model of ArsRC1 in arsenic detoxification in *Paracoccus sp*. SY. A**. A homology
346 model of the ArsRC1 protein was constructed based on the structures of the AfArsR repressor
347 (PDB ID: 6J05 (Prabaharan et al., 2019)) and *B. subtilis* ArsC (PDB ID: 1JL3 (Bennett et al.,
348 2001)). The ArsR and ArsC domains are in cyan and green, respectively. The ArsRC1 model
349 was built using the RaptorX server (Kallberg et al., 2012). **B.** Surface diagram of the ArsRC1
350 structural model. As(V) is proposed to be reduced by the ArsC domain to As(III), and the As(III)
351 product is channeled into to the ArsR domain for derepression. **C**. The combination of As(III)
352 oxidation, As(V) uptake, induction and efflux constitutes an efficient pathway for As(V)
353 detoxification in *Paracoccus* sp. SY. As(III) is proposed to oxidized to As(V) under aerobic or

anaerobic conditions using O_2 or NO_3^- as the respective electron acceptor. As an analog of Pi, As(V) is readily taken up via the Pst phosphate transporter. As(V) is reduced to As(III) by the ArsC domain, which is channeled into the ArsR domain, derepressing transcription of the *gapdh* and *arsJ* As(V) resistance genes.

Experimental Procedures

Strains, medium, reagents and growth conditions

E. coli Stellar™ (Clontech Laboratories, Inc., Mountain View, CA) (*F-, endA1, supE44, thi-1, recA1, relA1, gyrA96, phoA, Φ80d lacZΔ M15, Δ(lacZYA-argF)U169, Δ(mrr-hsdRMS-mcrBC), ΔmcrA, λ*–) was used for plasmid DNA construction and replication. *E. coli* AW3110 (*Δars::cam F- IN(rrn-rrnE*) (Carlin et al., 1995) bearing plasmids, pACYC184-P_{arsRC}-*gfp* and pETduet-*arsRC1* was constructed for use as an As(III) biosensor (Chen et al., 2012). *E. coli* WC3110(DE3) (*ΔarsC*) (Sundaram et al., 2008), which is hypersensitive to As(V), was used for complementation studies. *E. coli* BL21(DE3) (Novagen, Madison, WI) was used for protein expression. *E. coli* and *Paracoccus* sp. SY strains were grown aerobically at either 30 °C or 37 °C in either lysogeny broth (LB) or low phosphate medium (Oden et al., 1994), as indicated, supplemented with 125 μg/ml ampicillin or 34 μg/ml chloramphenicol, as required. Unless otherwise indicated, chemicals were purchased from Sigma-Aldrich.

Synthesis of the arsRC1 gene

An *arsRC1* gene encoding ArsRC1 (WP_103174508) from the *Paracoccus* sp. SY genome (NZ_NWMQ01000040.1) was chemically synthesized with 5' *Nco*I and 3' *Sal*I sites and with codon optimization for expression in *E. coli* and subcloned into the *Eco*RV site of vector plasmid pUC57-Kan (GenScript, NJ, USA). The synthetic *arsRC1* gene was cloned as an *Nco*I/*Sal*I double-digested fragment from pUC57-Kan-*arsRC1* into expression vector pETDuet-1

(Millipore Sigma), generating plasmid pETduet-*arsRC1* (Fig. S2). Plasmids used in this study are shown in Table S1. The primers for *arsRC1* derivatives used for amplification and plasmid construction are shown in Table S2 and Fig. S2A. The *arsRC1* promoter corresponding to the sequence of the genomic DNA of *Paracoccus* sp. SY was also chemically synthesized together with a *gfp* gene from vector plasmid pGreen (Hellens et al., 2000). The *gfp* reporter was put under control of the *arsRC1* promoter in vector plasmid pACYC184, generating plasmid pACYC184-P_{arsRC1}-*gfp* (Fig. S2A). For protein purification, the synthetic *arsRC1* gene and its *arsC*$_C$ domain were amplified by PCR using *Nde*I/*Xho*I restriction sites and subcloned into expression vector pET-29a(+) (Novagen) to add six histidine codons at the 3'-end. The forward and reverse primers for PCR were shown in Table S1. All sequences were confirmed by DNA sequencing (Sequetech, Mountain View, CA).

Assay of arsenical biotransformation

Cells of *Paracoccus sp. SY*, *E. coli* W3110 and WC3110 with *arsRC1 or arsC*$_C$ genes were cultured aerobically with shaking in LB medium overnight at 30 °C. The cells were washed once and suspended in low phosphate medium without glucose at a cell density of A_{600nm} = 3.0. Arsenicals were then added at 5 μM, final concentration, to the cell suspensions, which were incubated at 30 °C with shaking for 4 h. Soluble arsenicals were speciated by HPLC coupled to ICP-MS using a BioBasic-18 5 μm C18 300 Å reverse-phase column (250 mm × 4.6 mm; Thermo Fisher Scientific, Waltham, MA) eluted isocratically with a mobile phase consisting of 3 mM malonic acid, 5 mM tetrabutylammonium hydroxide, and 5% methanol (v/v), pH 5.6, with a flow rate of 1 ml min^{-1} at 25 °C.

Arsenical accumulation in Paracoccus sp. SY

For *in vivo* accumulation assays, cells of *Paracoccus* sp. *SY* and *E. coli* were grown in low phosphate medium at 37 °C to A_{600nm} =2. Cells were harvested and suspended in 1/5th volume of a buffer solution consisting of 75 mM HEPES-KOH, pH 7.5, 0.15 M KCl and 1 mM $MgSO_4$. To initiate the transport reaction, arsenicals were added at a final concentration of 20 μM to 1 ml of cell suspension. Portions (0.1 ml) were withdrawn at the indicated times, filtered through nitrocellulose filters (0.2 μm pore diameter; EMD Millipore, Billerica, MA) and washed twice at room temperature with 5 ml of the same buffer. The filters were digested with 0.3 ml of concentrated HNO_3 (68–70%) overnight at room temperature. The dissolved filters were incubated for 10 min at 70 °C, allowed to cool to room temperature and diluted with HPLC-grade water (Sigma-Aldrich, St. Louis) to produce a final HNO_3 concentration of 2%. Arsenic was quantified by ICP-MS. Standard solutions were made in the range of 0.5–50 ppb in 2% nitric acid using an arsenic standard (Ultra Scientific, N. Kingstown, RI).

Resistance assays

For resistance assays in liquid medium, competent cells of WC3110 were transformed with constructs bearing *arsRC1* and *arsC*$_C$. Cells of *Paracoccus* sp. *SY* and *E. coli* were grown overnight with shaking at 30 °C in LB medium to an A_{600nm}=2.0. Overnight cultures were diluted 100-fold in low phosphate medium (Oden et al., 1994) containing various concentrations of either trivalent or pentavalent arsenicals plus 0.3 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated at 30 ˚C with shaking for another 16 h. Growth was estimated from the A_{600nm}.

Biosensor assays

Transcriptional activity of the biosensor with *arsRC1* was estimated from arsenical-responsive expression of *gfp* (Chen and Rosen, 2014). Cultures of the biosensor (*E. coli* strain AW3110 bearing plasmids pETDuet-*arsRC1* or its derivatives and pACYC184-P_{arsRC1}-*gfp*) (Fig. S2A) were grown to mid-exponential phase in low phosphate medium at 37 °C with 100 μg/ml of ampicillin and 34 μg/ml of chloramphenicol with shaking (Fig. S2A). Glucose (0.2%) was added as a carbon source. The *arsRC1* or its derivatives genes were induced by addition of 0.2 mM IPTG for 5 hr. Derepression was produced by simultaneous addition of IPTG and arsenicals for 5 h (Fig. S2B). Cell densities were normalized by dilution or suspension to the same A_{600nm}, and expression of *gfp* was assayed from the fluorescence of cells using a Photon Technology International Spectrofluorometer with an excitation wavelength of 470 nm and emission wavelength of 510 nm.

ArsRC1 purification

E. coli BL21(DE3) cells (Life Technologies) bearing plasmid pET29a-*arsRC1* or *arsC*$_C$ were grown in LB medium containing 50 μg ml^{-1} kanamycin with shaking at 37 °C. Cells at an A_{600nm} of 0.6 were induced by addition of 0.3 mM IPTG, final concentration, and cultured for an additional 4 h. The cells were harvested and suspended in 5 ml per gram of wet cells in buffer A (50 mM 4-morpholinepropanesulfonic acid, 20 mM imidazole, 0.5 M NaCl, 10 mM 2-mercaptoethanol and 20% glycerol (v/v), pH 7.5). The cells were broken by a single passage through a French pressure cell at 20,000 psi and immediately treated with the protease inhibitor diisopropyl fluorophosphate (2.5 μl per gram wet cell). Membranes and unbroken cells were removed by centrifugation at 150,000 x g for 1 h, and the supernatant solution was loaded onto a Ni^{2+}-nitrilotriacetic acid column (Qiagen, Valencia, CA) at a flow rate of 0.5 ml min^{-1}. The column was washed with more than 25 column volumes of buffer A. ArsRC1 or $ArsC_C$ was eluted with buffer A containing 0.2 M imidazole, and the purity was analyzed by sodium dodecyl

sulfate polyacrylamide gel electrophoresis (SDS PAGE). Protein concentrations were estimated from the A_{280nm} (ε = 23500 M^{-1} cm^{-1} for PsArsRC, ε = 16860 M^{-1} cm^{-1} for $PsArsC_C$). Protein fractions were divided into portions, rapidly frozen and stored at −80 °C until use. Thioredoxin (TrxA) and thioredoxin reductase (TR) were prepared as described(Arner and Holmgren, 2000).

As(V) reduction by purified ArsRC1

Enzymatic activity of purified ArsRC1 or $ArsC_C$ was assayed at 37 °C in buffer B (25 mM Bis-Tris propane (pH 7.0)) containing 10 μM TrxA, 3 μM TR, and 0.2 mM NADPH. As(V) (5 μM) was incubated at 37 °C in the presence or absence of 2 μM ArsRC1 or $ArsC_C$. After 4 h, protein was removed by centrifugation using a 3 kDa cutoff Amicon ultrafilter (EMD Millipore). The filtrate was speciated by HPLC-ICP-MS, as described above.

Conflict of Interest

The authors state that they have no conflict of interest to declare.

Acknowledgments

This work was supported by NIH grants R01GM055425 to B.P.R. and NSF grant NSF BIO/MCB grant 1817962 to M.Y.

References

Achour-Rokbani, A., Cordi, A., Poupin, P., Bauda, P., and Billard, P. (2010) Characterization of the *ars* gene cluster from extremely arsenic-resistant *Microbacterium* sp. strain A33. *Appl Environ Microbiol* **76**: 948-955.

Arner, E.S.J., and Holmgren, A. (2000) Physiological functions of thioredoxin and thioredoxin reductase. *Euro J Biochem* **267**: 6102-6109.

Ben Fekih, I., Zhang, C.K., Li, Y.P., Zhao, Y., Alwathnani, H.A., Saquib, Q. et al. (2018) Distribution of arsenic resistance genes in prokaryotes. *Frontiers in Microbiology* **9(**2473), 1-11.

Bennett, M.S., Guan, Z., Laurberg, M., and Su, X.D. (2001) *Bacillus subtilis* arsenate reductase is structurally and functionally similar to low molecular weight protein tyrosine phosphatases. *Proc Natl Acad Sci USA* **98**: 13577-13582.

Busenlehner, L.S., Pennella, M.A., and Giedroc, D.P. (2003) The SmtB/ArsR family of metalloregulatory transcriptional repressors: Structural insights into prokaryotic metal resistance. *FEMS Microbiol Rev* **27**: 131-143.

Byers, L.D., She, H.S., and Alayoff, A. (1979) Interaction of phosphate analogues with glyceraldehyde-3-phosphate dehydrogenase. *Biochemistry* **18**: 2471-2480.

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

Chen, J., and Rosen, B.P. (2014) Biosensors for inorganic and organic arsenicals. *Biosensors (Basel)* **4**: 494-512.

Chen, J., Zhu, Y.G., and Rosen, B.P. (2012) A novel biosensor selective for organoarsenicals. *Appl Environ Microbiol* **78**: 7145-7147.

Chen, J., Yoshinaga, M., Garbinski, L.D., and Rosen, B.P. (2016) Synergistic interaction of glyceraldehydes-3-phosphate dehydrogenase and ArsJ, a novel organoarsenical efflux permease, confers arsenate resistance. *Mol Microbiol* **100**: 945-953.

Chen, J., Zhang, J., Wu, Y.-F., Zhao, F.-J. and Rosen, B.P. (2021), ArsV and ArsW provide synergistic resistance to the antibiotic methylarsenite. *Environ Microbiol*, 23: 7550-7562.

Chen, S.C., Sun, G.X., Yan, Y., Konstantinidis, K.T., Zhang, S.Y., Deng, Y. et al. (2020) The Great Oxidation Event expanded the genetic repertoire of arsenic metabolism and cycling. *Proc Natl Acad Sci USA* **117**: 10414-10421.

500 Delnomdedieu, M., Basti, M.M., Otvos, J.D., and Thomas, D.J. (1994) Reduction and binding of
501 arsenate and dimethylarsinate by glutathione: a magnetic resonance study. *Chem Biol Interact*
502 **90**: 139-155.
503 Fru, E.C., Somogyi, A., El Albani, A., Medjoubi, K., Aubineau, J., Robbins, L.J. et al. (2019) The
504 rise of oxygen-driven arsenic cycling at ca. 2.48 Ga. *Geology* **47**: 243-246.
505 Hellens, R.P., Edwards, E.A., Leyland, N.R., Bean, S., and Mullineaux, P.M. (2000) pGreen: a
506 versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. *Plant*
507 *Mol Biol* **42**: 819-832.
508 Herath, I., Vithanage, M., Bundschuh, J., Maity, J.P., and Bhattacharya, P. (2016) Natural
509 arsenic in global groundwaters: distribution and geochemical triggers for mobilization. *Current*
510 *Pollution Reports* **2**: 68-89.
511 Hudek, L., Premachandra, D., Webster, W.A., and Brau, L. (2016) Role of phosphate transport
512 system component PstB1 in phosphate internalization by *Nostoc punctiforme*. *Appl Environ*
513 *Microbiol* **82**: 6344-6356.
514 Ji, G., and Silver, S. (1992) Reduction of arsenate to arsenite by the ArsC protein of the arsenic
515 resistance operon of *Staphylococcus aureus* plasmid pI258. *Proc Natl Acad Sci U S A* **89**: 9474-
516 9478.
517 Ji, G., Garber, E.A., Armes, L.G., Chen, C.M., Fuchs, J.A., and Silver, S. (1994) Arsenate
518 reductase of *Staphylococcus aureus* plasmid pI258. *Biochemistry* **33**: 7294-7299.
519 Kallberg, M., Wang, H.P., Wang, S., Peng, J., Wang, Z.Y., Lu, H., and Xu, J.B. (2012)
520 Template-based protein structure modeling using the RaptorX web server. *Nature Protocols* **7**:
521 1511-1522.
522 Land, M., Hauser, L., Jun, S.R., Nookaew, I., Leuze, M.R., Ahn, T.H. et al. (2015) Insights from
523 20 years of bacterial genome sequencing. *Funct Integr Genomics* **15**: 141-161.

Lin, Y.F., Walmsley, A.R. and Rosen, B.P. (2006) An arsenic metallochaperone for an arsenic detoxification pump. *Proc Natl Acad Sci USA* **103:**15617-15622.

Liu, J., and Rosen, B.P. (1997) Ligand interactions of the ArsC arsenate reductase. *J Biol Chem* **272**: 21084-21089.

Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C.J., Lu, S. et al. (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* **45**: D200-D203.

Mathivanan, K., Chandirika, J.U., Vinothkanna, A., Yin, H., Liu, X., and Meng, D. (2021) Bacterial adaptive strategies to cope with metal toxicity in the contaminated environment - A review. *Ecotoxicol Environ Saf* **226**: 112863.

Mukhopadhyay, R., Shi, J., and Rosen, B.P. (2000) Purification and characterization of Acr2p, the *Saccharomyces cerevisiae* arsenate reductase. *J Biol Chem* **275**: 21149-21157.

Mukhopadhyay, R., Rosen, B.P., Phung, L.T., and Silver, S. (2002) Microbial arsenic: from geocycles to genes and enzymes. *FEMS Microbiology Reviews* **26**: 311-325.

Oden, K.L., Gladysheva, T.B., and Rosen, B.P. (1994) Arsenate reduction mediated by the plasmid-encoded ArsC protein is coupled to glutathione. *Mol Microbiol* **12**: 301-306.

Oremland, R.S., and Stolz, J.F. (2003) The ecology of arsenic. *Science* **300**: 939-944.

Prabaharan, C., Kandavelu, P., Packianathan, C., Rosen, B.P., and Thiyagarajana, S. (2019) Structures of two ArsR As(III)-responsive transcriptional repressors: Implications for the mechanism of derepression. *J Struct Biol* **207**: 209-217.

Qin, J., Fu, H.L., Ye, J., Bencze, K.Z., Stemmler, T.L., Rawlings, D.E., and Rosen, B.P. (2007) Convergent evolution of a new arsenic binding site in the ArsR/SmtB family of metalloregulators. *J Biol Chem* **282**: 34346-34355.

Rao, N.N., and Torriani, A. (1990) Molecular aspects of phosphate transport in *Escherichia coli*. *Mol Microbiol* **4**: 1083-1090.

549 Rosen, B.P. (2002) Biochemistry of arsenic detoxification. *FEBS Lett* **529**: 86-92.
550 San Francisco, M.J.D., Hope, C.L., Owolabi, J.B., Tisa, L.S., and Rosen, B.P. (1990)
551 Identification of the metalloregulatory element of the plasmid-encoded arsenical resistance
552 operon. *Nucleic Acids Research* **18**: 619-624.
553 Sundaram, S., Rathinasabapathi, B., Ma, L.Q., and Rosen, B.P. (2008) An arsenate-activated
554 glutaredoxin from the arsenic hyperaccumulator fern *Pteris vittata* L. regulates intracellular
555 arsenite. *J Biol Chem* **283**: 6095-6101.
556 Wu, B., Song, J., and Beitz, E. (2010) Novel channel enzyme fusion proteins confer arsenate
557 resistance. *J Biol Chem* **285**: 40081-40087.
558 Wu, J.H., and Rosen, B.P. (1993) Metalloregulated expression of the *ars* operon. *J Biol Chem*
559 **268**: 52-58.
560 Xu, C., Shi, W., and Rosen, B.P. (1996) The chromosomal *arsR* gene of *Escherichia coli*
561 encodes a trans-acting metalloregulatory protein. *J Biol Chem* **271**: 2427-2432.
562 Zegers, I., Martins, J.C., Willem, R., Wyns, L., and Messens, J. (2001) Arsenate reductase from
563 *S. aureus* plasmid pI258 is a phosphatase drafted for redox duty. *Nat Struct Biol* **8**: 843-847.
564 Zhang, J., Zhou, W., Liu, B., He, J., Shen, Q., and Zhao, F.J. (2015) Anaerobic arsenite
565 oxidation by an autotrophic arsenite-oxidizing bacterium from an arsenic-contaminated paddy
566 soil. *Environ Sci Technol* **49**: 5956-5964.
567 Zhu, Y.G., Yoshinaga, M., Zhao, F.J., and Rosen, B.P. (2014) Earth abides arsenic
568 biotransformations. *Annu Rev Earth and Planet Sci* **42**: 443-467.