Functional and structural characterization of AntR, an Sb(III) responsive transcriptional repressor

Thiruselvam Viswanathan[1#], Jian Chen[1#], Minghan Wu[2], Lijin An[2] Palani Kandavelu[3], Banumathi Sankaran[4], Manohar Radhakrishnan[1#], Mingshun Li[2*] and Barry P Rosen[1*]

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

[2]State Key Laboratory of Agricultural Microbiology, College of Life Science and Technology, Huazhong Agricultural University, Wuhan, Hubei 430070, PR China

[3]SER-CAT and the Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA

[4]Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley Laboratory, Berkeley Center for Structural Biology, Berkeley, CA 94720, USA

# These authors contributed equally to this work.

* Corresponding authors: **Barry P Rosen,** Department of Cellular Biology and Pharmacology, Herbert Wertheim College of Medicine, Florida International University, Miami, FL 33199, United States, E-mail: brosen@fiu.edu, Tel: (305)-348-0657, ORCID https://orcid.org/0000-0002-5230-4271, and **Mingshun Li**, State Key Laboratory of Agricultural Microbiology, College of Life Science and Technology, Huazhong Agricultural University, Wuhan, Hubei 430070, PR China, Tel: 13477014086, E-mail: mshli7125@mail.hzau.edu.cn.

Key words: Antimony, AntR, ArsR family, transcriptional repressor, gene regulation. *Comamonas testosterone*

Abbreviations: Arsenite**,** As(III); antimonite, Sb(III); methylarsenite, MAs(III); methylarsenate, MAs(V) or MSMA; dimethylarsenate, DMAs(V); Tris 2-carboxyethyl)phosphine, TCEP; sodium dodecyl sulfate poly acrylamide gel electrophoresis, SDS PAGE; isothermal titration calorimetry, ITC; HPLC, high performance liquid chromatography; ICP-MS, inductively coupled mass spectroscopy; electrophoresis mobility shift assay, EMSA.

RUNNING TITLE: AntR, an Sb(III) responsive transcriptional repressor

Summary

The *ant* operon of the antimony-mining bacterium *Comamonas testosterone* JL40 confers resistance to Sb(III). The operon is transcriptionally regulated by the product of the first gene in the operon, *antR*. AntR is a member of ArsR/SmtB family of metal/metalloid-responsive repressors resistance. We purified and characterized *C. testosterone* AntR and demonstrated that it responds to metalloids in the order Sb(III)>methylarsenite (MAs(III)>>As(III)). The protein was crystallized, and the structure solved at 2.1 Å resolution. The homodimeric structure of AntR adopts a classical ArsR/SmtB topology architecture. The protein has five cysteine residues, of which $Cys103_a$ from one monomer and $Cys113_b$ from the other monomer, are proposed to form one Sb(III) binding site, and $Cys113_a$ and $Cys103_b$ forming a second binding site. This is the first report of the structure and binding properties of a transcriptional repressor with high selectivity for environmental antimony.

Introduction

Antimony is a toxic element that is considered a priority environmental pollutant by the United States Environmental Protection Agency (EPA) and the European Union (EU, 1976, USEPA, 1979). Exposure to antimony leads to a number of human disorders, including cancer, liver, cardiovascular and respiratory diseases. Antimony is found in two oxidation states, highly toxic antimonite (Sb(III)) and much less toxic antimonate (Sb(V)). Antimony is located below arsenic on the periodic table, and both metalloids have similar chemical and toxicological properties (Li *et al.*, 2016).

Bacterial resistance to arsenic is conferred by *ars* operons. These genes also usually confer antimony as well. Recently an antimony-specific *antRCA* operon was identified in *Comamonas testosterone*, which was isolated from an antimony mine (An *et al.*, 2021). The *antA* gene encodes an Sb(III)-translocating P_{1B}-type ATPase that extrudes the toxic metalloid from the cells, giving resistance. AntC is a small protein that is proposed to be an Sb(III) chaperone for the AntA efflux pump.

AntR is a negative repressor that regulates the *ant* operon. It responds to Sb(III) and MAs(III) but only weakly to As(III) and not at all to Sb(V) or As(V). AntR (124 amino acid residues, 13.5 kDa, accession number WP034375793.1) is a member of the ArsR/SmtB family of metal(oid)-responsive transcriptional repressors (Busenlehner *et al.*, 2003, Chen & Rosen, 2014). These transcriptional repressors respond to a variety of toxic metal(oid) ions, including four different types of ArsR repressors (As(III), Sb(III), Bi(III)) (Chen *et al.*, 2017, Ordóñez *et al.*, 2008, Qin *et al.*, 2007, San Francisco *et al.*, 1990), SmtB (Zn(II)) (Huckle *et al.*, 1993), CadC (Cd(II)) (Ye *et al.*, 2005, Ji & Silver, 1992), and CmtR (Cd(II)) (Cavet *et al.*, 2003). In the absence of

metal ions, the homodimeric repressor binds to the promoter, preventing transcription. When substrate binds, the repressor dissociates from the promoter, allowing gene expression (Busenlehner *et al.*, 2003, Chen & Rosen, 2014). Members of the ArsR/SmtB family share common features such as structural topology, dimerization and functional activity. AntR homologs form a separate clade from the four known types of ArsR repressors (An *et al.*, 2021). Members of the four ArsR groups have evolved four different types of As(III)/Sb(III) binding sites. Each type of binding site is composed of two or three cysteine residues, but the location of those residues differs among the groups, reflecting convergent evolution of As(III) binding sites. In some ArsRs the binding site residues are within the same monomer such as the plasmid R773 ArsR (Shi *et al.*, 1994) and *Corynebacterium glutamicum* CgArsR (Ordóñez *et al.*, 2008). In others both chains contribute cysteine residues to the binding site that are in close spatial proximity such as *Acidothiobacillus ferrooxidans* AfArsR (Qin *et al.*, 2007).

Here we describe the structural and functional properties of the *C. testosterone* AntR. The results of binding assays performed with isothermal titration calorimetry (ITC) and the fluorescence quenching of a single tryptophan derivative demonstrate that AntR is highly selective for Sb(III) and MAs(III) over As(III). AntR has five cysteine residues. From the results of mutagenesis, Cys103 and Cys113 appear to be involved in Sb(III)/MAs(III) binding. AntR was crystallized, and the apo structure solved at 2.1 Å resolution. It is a homodimer with a ArsR/SmtB topology architecture. The structure shows that Cys103 from one monomer is opposed to Cys113 from the other monomer. We proposed that these two cysteine residues form a two-coordinate binding site for Sb(III)/MAs(III), with two binding sites in the

homodimer. This report of the structure and binding properties of a transcriptional repressor with high selectivity for environmental antimony sheds light on how life has adapted to the presence of this toxic metalloid.

Results

AntR is a new member of the ArsR/SmtB family of metal(loid)-responsive transcription factors

Recently an antimony resistance operon was identified in the chromosome of *C. testosterone* JL40 from an antimony mine in Lengshuijiang, Hunan Province, China (Li *et al.*, 2013). The *antRCA* operon encodes an Sb(III) efflux system that confers resistance by pumping the metalloid from the cells. The first gene in the operon encodes the AntR repressor that controls expression of itself and the *antCA* genes by sensing environmental Sb(III) (An *et al.*, 2021). From phylogenetic analysis, AntR belongs to the ArsR/SmtB family of negative repressors (Xu & Rosen, 1999, Busenlehner *et al.*, 2003). Members of this family are homodimers that bind transition metals, heavy metals or metalloids, including As(III), Sb(III), Cd(II), Pb(II), Zn(II), Co(II), and Ni(II). Binding dissociates the repressor from the promoter DNA, resulting in transcription of the resistance genes. The regulatory sites are often three- or four-coordinate metal binding sites composed of cysteine thiolates and histidine imidazole nitrogens. Reflecting shared evolution, they have a common backbone structure. Unexpectedly, however, the binding sites are found in quite different locations in various structures. We have hypothesized that those sites are the result of parallel evolution and arose in response to different environmental pressures. For example, the three-cordinate As(III) binding site in the *E. coli* and plasmid R773 ArsRs is more closely related to the four coordinate Cd(II)/Pb(II)/Zn(II)

binding site of the CadC repressor of the *S. aureus* pI258 plasmid *cadCA* operon (Ye *et al.*, 2005) than to other ArsR repressors such as CgArsR or AfArsR (Ordóñez *et al.*, 2008, Qin *et al.*, 2007, Prabaharan *et al.*, 2019). AntR appears to more closely related to the SmtB Zn(II)-responsive repressor than to ArsRs (Fig. 1), so the location of the Sb(III) binding site could not be predicted from the structure of ArsR repressors.

AntR is a Sb(III)/MAs(III)-selective repressor

To determine the inducer specificity of AntR, a two-plasmid biosensor was constructed that utilizes the gene for *antR* under control of the arabinose promoter in one plasmid (pBAD-*CtantR*), and the *ant* promoter controlling *gfp* expression in the other (pACYC184-*ParsO-gfp*) (Fig. S1) using the system described previously (Chen *et al.*, 2012). In cells of *E. coli* AW3110*Δars* bearing both plasmids, addition of arabinose drives expression of AntR, which binds to the *antO* promoter and inhibits the expression of *gfp* from pACYC184-P*arsO-gfp* (Fig. 2). In the absence of arabinose, *gfp* is constitutively expressed. Induction by addition of 0.2% arabinose, represses *gfp* expression, and the cells are not fluorescent. The biosensor with wild type AntR exhibited little *gfp* expression upon addition of 5 µM As(III). Addition of 5 µM Sb(III) or MAs(III) induced *gfp* expression. Pentavalent arsenicals (5 µM of As(V), MAs(V) or DMAs(V)) were not inducers (data not shown).

Role of cysteine residues in in vivo response to metalloids

ArsR repressors typically have two- or three-coordinate binding sites per monomer (Moinier *et al.*, 2014). AntR has five cysteine residues, Cys12, Cys58, Cys91, Cys103 and Cys113. Only Cys103 and Cys113 are conserved in the 10 closest AntR homologs, so we focused on these two cysteine residues (Fig. S2). To examine their

involvement in Sb(III) sensing, two mutant repressors were constructed, C103S and C113S. Each mutant gene was put into the biosensor plasmids, and their response to metalloids examined. The response of biosensor with wild AntR to metalloids was proportional to the concentration of Sb(III) and MAs(III) with little or no response to MAs(V) or As(III) or As(V) (Fig. 3A). Both mutants had reduced responses to Sb(III) and MAs(III) compared with the wild type (Fig. 3C and D). These results are consistent with a role of both Cys103 and Cys113 in binding and sensing of Sb(III) and MAs(III).

Binding of metalloids and role of cysteine residues in in vitro response to metalloids

AntR was purified by NiNTA affinity chromatography, and binding of metalloids was assayed by two assays that apply different physical principles. First, the intrinsic tryptophan fluorescence of the AntR Y96W derivative was quenched by addition of metalloids. This method has been used to examine binding specificity in proteins that bind As(III) and Sb(III) such as the ArsA ATPase (Zhou & Rosen, 1997, Ruan *et al.*, 2006), the ArsD As(III)/Sb(III) metallochaperone (Yang *et al.*, 2010) and the ArsM As(III) S-adenosylmethionine methyltransferase (Marapakala *et al.*, 2012). From titration of the fluorescence quenching with metalloid, the apparent K_d values for both Sb(III) and MAs(III) were approximately 5 μM (Fig. 4). The estimated K_d for As(III) of approximately 160 μM is at the limit of the method.

Second, the binding parameters of purified AntR for Sb(III), MAs(III) and As(III) determined by ITC (Fig. 5). The apparent K_d values for Sb(III) and MAs(III) were observed to be 57 μM and 77 μM, respectively. The K_d for As(III) was calculated as

440 μM, nearly an order of magnitude lower affinity. It should be pointed out that the 10-fold differences between apparent values with the two assays are probably due to the 10-fold different protein concentrations used in each, so these apparent binding constants should not be considered actual K_d values but are useful for comparative analysis of different ligands. However, the results from both assays clearly demonstrate that AntR is an Sb(III)/MAs(III)-selective repressor with low affinity for As(III).

Binding of AntR to the ant promoter

AntR binding to the operator/promoter site and the response to inducers were examined by EMSA. The 193 bp upstream sequence of the *antR* gene containing the operator/promoter sequence was amplified by PCR using a 6-FAM-labeled primer and purified AntR was used for the EMSA assays. The mobility of the probe was gradually retarded when increasing amounts of AntR were added, reflecting binding of the repressor to its cognate promoter (Fig. 6A). The effect of metal(loids) on AntR binding was assayed (Fig. 6B-F). The mobility shift decreased with increasing concentration of Sb(III). Neither pentavalent metalloids nor other metals evoked a response. These results are consistent with the results of binding experiments demonstrating that ArsR is selective for Sb(III).

Structural architecture of AntR

Members of the ArsR/SmtB family are homodimers that adopt a winged helix structural topology. The AntR homodimer crystallized in the $P12_11$ space group with cell dimensions of a=51.7 Å, b=58.5 Å and c=53.5 Å. The structure refined with acceptable stereochemistry and a final R-factor of 0.22 (R_{free} =0.26) at 2.1 Å

resolution (PDB ID 6UVU). The details of data collection and refinement are given in Table 1. The structure consists of chain A from residues 11 to 124 and chain B from residues 11-121 (Fig. 7). N terminal residues (1-10) are not visible in either chain. The overall structure is composed of five α-helices and two anti-parallel β-strands connected through coils (α1-α2-α3-α4-β1-β2-α5). All five cysteine residues are visible in both chains. The two conserved cysteines, Cys103 and Cys113, are visible in helix α5 at the dimer interface. The sulfur atom of $Cys103_a$ is 10.9 Å from that of $Cys113_b$, and $Cys113_a$ sulfur atom is 11.2 Å from the sulfur atom of $Cys103_b$. Combining the mutagenesis results described above with these structural data are consistent with Cys103 and Cys113 from each monomer forming two 2-coordinate Sb(III)/MAs(III) binding sites between the two dimers.

Discussion

In this report we characterized the AntR transcriptional repressor that controls expression of the *antRCA* operon of *Comamonas testosterone* JL40, which was isolated from an antimony mine (Li *et al.*, 2013). The results of biosensor and binding assays clearly demonstrate that the repressor is highly selective for Sb(III), as well as MAs(III), which has been described as a primordial antibiotic (Chen *et al.*, 2019). AntR is a member of the ArsR/SmtB family of metal(loid)-responsive transcriptional regulatory proteins. Most ArsR repressors also bind and respond to Sb(III). In contrast, AntR has little response to As(III). What governs the differential selectivity of the AntR and ArsR repressors? AntR is more closely related to SmtB than to any characterized ArsR (Fig. 1B). Since SmtB is a Zn(II)-responsive repressor, there is no reason to expect that AntR would bind As(III) (Fig. 6C) any better than the non-inducer Zn(II) (Fig. 6F). Typical ArsR repressors bind As(III) with

219 a trigonal pyramidal molecular geometry (Prabaharan *et al.*, 2019). Two recent
220 crystal structures of ArsR repressors from *A. ferrooxidans* (AfArsR) and *C.*
221 *glutamicum* (CgArsR), respectively, show two different three-cysteine As(III) sites. In
222 AfArsR As(III) is bound to Cys95, Cys96 and Cys102 from the same monomer of the
223 homodimer, while, in the CgArsR structure, As(III) is bound to Cys15 and Cys16 from
224 one monomer and Cys55 from the other monomer (Prabaharan *et al.*, 2019).
225
226 Those ArsR repressors can also bind Sb(III) and MAs(III), but is a three-coordinate
227 site required to bind those more thiophilic metalloids? An atypical ArsR is the
228 MAs(III)-responsive SpArsR from *Shewanella putrefaciens* 200 that regulates
229 expression of the MAs(III) resistance genes *arsP* and *arsH (Chen et al., 2017)*.
230 SpArsR is induced by MAs(III), with little response to As(III). SpArsR is closely related
231 to AfArsR from *A. ferrooxidans*, in which As(III) is bound to a three-coordinate site
232 composed of residues Cys95, Cys96 and Cys102, which are located at the C-
233 terminus of the repressor (Qin *et al.*, 2007). In SpArsR the residue equivalent to
234 Cys102 is not a cysteine, and MAs(III) is bound to a two-coordinate site composed of
235 Cys101 and Cys102. This illustrates how replacing a single cysteine residue can
236 alter an S_3 site into an S_2 site that changes selectivity from As(III) to MAs(III). SpArsR
237 and AntR both respond to MAs(III) but are not closely related, indicating that this is
238 probably convergent evolution of two different two-coordinate binding sites.
239
240 In AntR there are only two cysteine residues that could reasonably form a binding
241 site, Cys103 and Cys113. This idea is supported by 1) the fact that only these two
242 cysteine residues of the five in AntR are conserved; 2) mutagenesis of either of
243 those residues reduces *in vivo* biosensing; 3) the proximity of those two residues in

the crystal structure. It should be pointed out, however, that the distance between the Cys103 sulfur atom in one monomer and the Cys113 sulfur atom in the other monomer of the apo-AntR structure is approximately 10-11 Å. In CgArsR and AfArsR As(III) is bound with a trigonal pyramidal molecular geometry with As-S distances of approximately 2.25 Å. Typical Sb-S bond lengths fall in the range of 2.4-2.5 Å. To form a similar trigonal pyramidal binding site in AntR in which Sb(III) or MAs(III) is bound to two sulfurs and a hydroxyl replacing a third, the cysteine residues would have to move closer to each other. Even though distances are too long in the apo structure, it is reasonable to propose that binding of Sb(III)/MAs(III) would bring the a6 helices close enough to each other to form the binding site. In fact, this might propagate a conformational change in the repressor that would result in dissociation from the promoter DNA and hence derepression. Since the *ant* operon is found in an antimony-mining organism, it seems reasonable to consider that the physiological inducer for AntR is Sb(III) even though it responds to MAs(III). In summary, this is the first report of the structure and function of a transcriptional regulatory element that controls detoxification of antimony and illustrates how microbes adapted to the presence this highly toxic environmental metalloid.

Experimental procedures

Strains, plasmids, medium and reagents

E. coli Stellar™ (Clontech Laboratories, Mountain View, CA) (*F2, endA1, supE44, thi-1, recA1, relA1, gyrA96 phoA, lacZΔ M15, Δ(lacZYA-argF)U169, Δ(mrrhsdRMSmcrBC*), *ΔmcrA*, k^-) was used for plasmid DNA construction and replication. *E. coli* AW3110(DE3) (*Δars::cam* F2IN(*rrn-rrnE*) (Carlin *et al.*, 1995) in bearing two plasmids, pACYC184-*PantO-gfp* and pBAD-*CtantR* was constructed for

use as an Sb(III)/MAs(III) biosensor (Chen *et al.*, 2012). For most experiments cultures of *E. coli* bearing the indicated plasmids were grown aerobically in lysogeny broth (LB) medium or M9 medium at 37°C supplemented with 100 μg/mL ampicillin or 34 μg/mL chloramphenicol, as required (Sambrook *et al.*, 1989). Bacterial growth was monitored by measuring the absorbance at 600 nm (A_{600nm}).

Unless otherwise indicated, chemicals were purchased from Sigma-Aldrich. MAs(V) was obtained from Thermo Fisher Acros Organics Division (Waltham, MA). MAs(III) was reduced as described (Reay & Asher, 1977). Briefly, 0.2 mM of pentavalent arsenicals were mixed with 27 mM $Na_2S_2O_3$, 66 mM $Na_2S_2O_5$, and 82 mM H_2SO_4, following which the pH was adjusted to 6.0 with NaOH. The reduced products were not thiolated, as determined by simultaneous As and S analysis by high pressure liquid chromatography (HPLC) inductively coupled mass spectroscopy (ICP-MS) using an ELAN DRC-e (PerkinElmer, Waltham, MA). The dynamic reaction cell detects arsenic as $^{91}AsO^+$ and sulfur as $^{48}SO^+$ with high sensitivity.

Plasmid construction

The *antR* gene from *C. testosterone* JL40 was chemically synthesized with 5′-*Nco*I and 3′-*Sal*I sites and codon optimization for expression in *E. coli* GenScript (NJ, USA). The gene was cloned into pBAD/myc-HisA, generating plasmid pBAD-*CtantR* (Fig. S1). The *antR* (WP_034375793) promoter corresponding to the sequence of the genomic DNA of *C. testosteroni* JL40 was chemically synthesized together with a *gfp* gene originated from vector plasmid pGreen (Hellens *et al.*, 2000). The *gfp* reporter was under control of the AntR promoter in vector plasmid pACYC184, generating plasmid pACYC184-*PantO-gfp*. To purify AntR for biochemical characterization and

294 crystallization, *antR* was excised from pBAD-*CtantR* with *Nco*I and *Xho*I restriction
295 enzymes and subcloned into plasmid pET-29a(+) using the same restriction sites in
296 frame with a C-terminal six-histidine residue tag.
297
298 ***Biosensor assays***
299 Plasmids pBAD-*CtantR* and pACYC184-*PantO-gfp* were co-expressed in *E. coli*
300 AW3110*Δars (Carlin et al., 1995)*, generating a biosensor that could be used to
301 determine the binding affinity of AntR for antimonials and arsenicals, as previous
302 reported (Chen *et al.*, 2012). Transcriptional activity of the biosensor was estimated
303 from metalloid-responsive expression of *gfp*. Cultures of the biosensor were grown
304 to mid-exponential phase in M9 medium at 37 °C with 100 μg ml^{-1} ampicillin and 34
305 μg ml^{-1} chloramphenicol with shaking. Glycerol (0.5%) was added as carbon source
306 and for constitutive expression of gfp. The *CtantR* gene was induced by addition of
307 0.2% arabinose for 5 h. Derepression was produced by simultaneous addition of
308 arabinose and metalloid for 5 h. Cell densities were normalized by dilution or
309 suspension to the same A_{600nm}, and expression of *gfp* was assayed from the
310 fluorescence of cells using a Photon Technology International spectrofluorometer
311 with an excitation wavelength of 470 nm and emission wavelength of 510 nm.
312
313 ***Mutagenesis of the antR gene***
314 The codons for residues Cys58, Cys103 and Cys113 were changed to serine codons,
315 generating single cysteine AntR mutants. The mutations were generated by site-
316 directed mutagenesis using a Quick-Change mutagenesis kit (Stratagene, La Jolla,
317 CA). The forward and reverse mutagenic oligonucleotides used for both strands and
318 the respective changes introduced (underlined) are as follows: C103S forward, 5'-

GCATCGTGGACCCGAGCGTGCTGAGAATG-3' and reverse 5'-CATTCTCAGCACGCTCGGGTCCACGATGC-3'; C113 forward, 5'-GAATGCTCGAACTTGGGCTAAGCCTTATCGAGGAG-3' and reverse 5'-CTCCTCG ATAAGGCTTAGCCCAAGTTCGAGCATTC-3'. Each mutation was confirmed by DNA sequencing (Sequetech, Mountain View, CA).

Protein purification

Cells of *E. coli* BL21(DE3) bearing *CtantR* in plasmid pET-29a(+) were grown in LB medium containing 100 µg/ml kanamycin with shaking at 37 °C. At an A_{600nm} of 0.6, 0.3 mM isopropyl β- d-1-thiogalactopyranosidewas added as an inducer, and the culture was grown for an additional 4 h at 37 °C. The cells were harvested and suspended in 5 ml/g of wet cells in buffer A (50 mM 4-morpholinepropanesulfonic acid, 20 mM imidazole, 0.5 M NaCl, 10 mM 2-mercaptoethanol, 20% glycerol (vol/vol), pH 7.5 and 2 mM Tris 2-carboxyethyl)phosphine (TCEP)). 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/ g wet cell). Membranes and unbroken cells were removed by centrifugation at 150,000 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. The column was washed with more than 25 column volumes of buffer A. CtAntR was eluted with a liner gradient of 0-0.2 M imidazole in buffer A, and purity was analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) on a 15% gel. Protein concentrations were estimated from A_{280nm} (ε=1880 M^{-1} cm^{-1}). CtAntR-containing fractions were divided into small portions, rapidly frozen and stored at -

80°C until use. On gel filtration purified AntR eluted at the position of a dimer, consistent with the dimeric structure of ArsR repressors (Prabaharan *et al.*, 2019)

Phylogenetic analysis

Multiple alignment of AntR with selected members of the ArsR/SmtB family was performed using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/). AntR, SmtB, CadC and ArsR sequences with conserved cysteines were selected for phylogenetic analysis. Acquisition of sequences was performed by searching a list of reference organisms. Phylogenetic analysis was performed to infer the evolutionary relationship among the arsenic repressor of various organisms. The phylogenetic tree was constructed using the Neighbor-Joining method with MEGA 6.0.1 (Saitou & Nei, 1987). The statistical significance of the branch pattern was estimated by conducting a 1000 bootstrap.

Binding affinity for metalloids

The binding affinity of AntR with metalloids was assayed by two methods. The first measured the quenching of intrinsic protein fluorescence of AntR upon binding of metalloids. AntR has no tryptophan residues and low intrinsic fluorescence. A single tryptophan mutant was constructed by changing Tyr96 to a tryptophan residue, and the AntR Y96W derivative was used for metalloid quenching assays. A temperature-controlled Quanta-Master UV−vis QM-4 steady state spectrofluorometer (Photon Technology International, Birmingham, NJ) was used for fluorescence measurements. The fluorescence of 5 μM AntR Y96F and or derivatives was assayed in a buffer consisting of 50 mM MOPS, pH7.8, 0.5 M NaCl and 2 mM TCEP at excitation and emission wavelengths of 295 and 331 nm, respectively. Sb(III) (25,

50, 100, 150, 200 µM) or MAs(III) (2.5, 5, 7.5 10, 15 µM) were added, and quenching of fluorescence was measured. For all assays, the temperature was 23 °C, and the spectrum of the buffer solution alone was subtracted to correct for background fluorescence and Raman scattering. For determination of relative affinities of Sb(III), MAs(III) and As(III), fluorescence spectra were acquired at various concentrations of metalloids, and the affinity for each was calculated according to the method of Rosenthal (Rosenthal, 1967) by plotting (ΔF/ΔFmax)/[L] as a function of ΔF/ΔFmax, where ΔF/ΔFmax is the fractional change in fluorescence at 345 nm, ΔF is the quenching at a particular concentration of arsenical ligand, [L], and ΔFmax is the quenching at the highest (saturating) arsenical concentration.

In the second assay, quantitative binding parameters were determined by isothermal titration calorimetry (ITC). Assays were performed at 30 °C with a stirring speed of 350 rpm using a MicroCal iTC_{200} instrument (GE Healthcare Bio Sciences) in a freshly prepared buffer solution consisting of 50 mM MOPS, 20% glycerol, 0.5 M NaCl, 20 mM imidazole and 1 mM TCEP. Each assay contained 0.3 ml of 70 µM AntR, with sequential addition of 45 µl aliquots of 0.5 mM Sb(III), MAs(III) or As(III) in the same buffer that had been filtered through 0.2 µM filters, as indicated. MAs(III) was freshly prepared by chemical reduction of MAs(V) before use.

Electrophoretic mobility shift assays

AntR binding to the operator/promoter site and the response to inducers were examined by electrophoretic mobility shift assays (EMSA). A DNA fragment was amplified by PCR using primers AntA-EMSA-F and AntA-EMSA-R consisting of the 193 nt upstream of the start of the *antR* gene containing the *ant* operator/promoter

sequence. The forward primer was synthesized with 5’-labeling with the fluorophore 6-fluorescein phosphoramidite (6-FAM) (Integrated DNA Technologies). AntR purified as described above was used for EMSA assays. All reaction mixtures with or without metalloids were incubated at 28 °C for 30 min in a buffer consisting of 20 mM Tris-HCl, pH 7.0, 50 mM NaCl, 1 mM dithiothreitol, 10 mM $MgCl_2$ and 0.1 mg/ml bovine serum albumin. The binding solution was loaded onto a 6 % native PAGE gel. After 3 h at 80 V in Tris-glycine-EDTA (TGE) buffer (12 mM Tris, 0.95 M glycine and 5 mM EDTA, the gels were analyzed using a Fujifilm FLA-5100 fluorescence imaging system.

Crystallization, data collection and structure determination

Purified AntR (15 mg/ml) was used for initial crystallization screens using sitting drop vapor diffusion and hanging drop methods using crystallization screens from Hampton Research, Molecular Dimensions and Emerald BioSystems, Inc. Small plate-like crystals were grown in crystal screen kit condition 41 (0.1 mM HEPES sodium, pH 7.5, 10% v/v 2-propanol, 20% w/v polyethylene glycol 4,000) in sitting drops. Addition of a 1:1 ratio of protein:reservoir solution led to a slight precipitate that cleared by addition of 1 µl of 0.1 M TCEP. Diffraction-quality crystals were obtained with Hampton additive screen condition 47 within a week. High quality crystals were flash cooled in liquid nitrogen for data collection. X-ray diffraction data were collected at the Advanced Light Source (ALS), Lawrence Berkeley National Laboratory, Berkeley, California and the Southeast Regional Collaborative Access Team (SER-CAT) facility at the Advanced Photon Source (Sapsford *et al.*, 2006), Argonne National Laboratory. The structures were determined by molecular replacement (McCoy, 2007). Structural refinement of each data set was performed

with REFMAC5 (Vagin *et al.*, 2004) implemented with the CCP4 suite (Winn *et al.*, 2011). The model and electron density map were visualized using COOT software (Emsley & Cowtan, 2004).

Acknowledgements

This work was supported by the National Natural Science Foundation of China, grant number 31970095 and the Development Program of China, grant number 2016YFD0800702 to M.L., National Natural Science Foundation of China, grant number 41967023 to J.C., and National Institutes of Health (Grant numbers GM055425, GM136211 and ES023779) to bpr. This project utilized the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beam line of the Advanced Photon Source, Argonne National Laboratory. SER-CAT is supported by its member institutions, and equipment grants (S10_RR25528, S10_RR028976 and S10_OD027000) from the National Institutes of Health.Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science and Office of Basic Energy Sciences under contract no. W-31-109-Eng-38. The Berkeley Center for Structural Biology is supported in part by the National Institutes of Health, National Institute of General Medical Sciences, and the Howard Hughes Medical Institute. The ALS-ENABLE beamlines are supported in part by the National Institutes of Health, National Institute of General Medical Sciences, grant P30 GM124169-01 and the Howard Hughes Medical Institute. The Advanced Light Source is a Department of Energy Office of Science User Facility under Contract No. DE-AC02-05CH11231.

Data Availability Statement: Data available on request from the authors

References

An, L., Luo, X., Wu, M., Feng, L., Shi, K., Wang, G., Rosen, B.P., and Li, M. (2021) *Comamonas testosteroni antA* encodes an antimonite-translocating P-type ATPase. *Sci. Total Environ.* **754**: 142393.

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.

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.

Cavet, J.S., Graham, A.I., Meng, W., and Robinson, N.J. (2003) A cadmium-lead sensing ArsR-SmtB repressor with novel sensory sites: Complementary metal-discrimination by NMTR and CMTR in a common cytosol. *J. Biol. Chem.* **278**: 44560-44566.

Chen, J., Nadar, V.S., and Rosen, B.P. (2017) A novel MAs(III)-selective ArsR transcriptional repressor. *Mol. Microbiol.* **106**: 469-478.

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

Chen, J., Yoshinaga, M., and Rosen, B.P. (2019) The antibiotic action of methylarsenite is an emergent property of microbial communities. *Mol. Microbiol.* **111**: 487-494.

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

467 Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics.
468 *Acta Crystallogr. D. Biol. Crystallogr.* **60**: 2126-2132.
469 Hellens, R.P., Edwards, E.A., Leyland, N.R., Bean, S., and Mullineaux, P.M. (2000)
470 pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated
471 plant transformation. *Plant Mol. Biol.* **42**: 819-832.
472 Huckle, J.W., Morby, A.P., Turner, J.S., and Robinson, N.J. (1993) Isolation of a
473 prokaryotic metallothionein locus and analysis of transcriptional control by
474 trace metal ions. *Mol. Microbiol.* **7**: 177-187.
475 Ji, G., and Silver, S. (1992) Regulation and expression of the arsenic resistance
476 operon from *Staphylococcus aureus* plasmid pI258. *J. Bacteriol.* **174**: 3684-
477 3694.
478 Li, J., Wang, Q., Oremland, R.S., Kulp, T.R., Rensing, C., and Wang, G. (2016)
479 Microbial antimony biogeochemistry: enzymes, regulation, and related
480 metabolic pathways. *Appl. Environ. Microbiol.* **82**: 5482-5495.
481 Li, J., Wang, Q., Zhang, S., Qin, D., and Wang, G. (2013) Phylogenetic and genome
482 analyses of antimony-oxidizing bacteria isolated from antimony mined soil.
483 *Int Biodeter Biodegr* **76**: 76-80.
484 Marapakala, K., Qin, J., and Rosen, B.P. (2012) Identification of catalytic residues in
485 the As(III) S-adenosylmethionine methyltransferase. *Biochemistry* **51**: 944-
486 951.
487 McCoy, A.J. (2007) Solving structures of protein complexes by molecular
488 replacement with Phaser. *Acta Crystallogr. D. Biol. Crystallogr.* **63**: 32-41.
489 Moinier, D., Slyemi, D., Byrne, D., Lignon, S., Lebrun, R., Talla, E., and Bonnefoy, V.
490 (2014) An ArsR/SmtB Family Member Is Involved in the Regulation by Arsenic

of the Arsenite Oxidase Operon in Thiomonas arsenitoxydans. *Appl Environ Microb* **80**: 6413-6426.

Ordóñez, E., Thiyagarajan, S., Cook, J.D., Stemmler, T.L., Gil, J.A., Mateos, L.M., and Rosen, B.P. (2008) Evolution of metal(loid) binding sites in transcriptional regulators. *J. Biol. Chem.* **283**: 25706-25714.

Prabaharan, C., Kandavelu, P., Packianathan, C., Rosen, B.P., and Thiyagarajan, 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.

Reay, P.F., and Asher, C.J. (1977) Preparation and purification of ^{74}As-labeled arsenate and arsenite for use in biological experiments. *Anal. Biochem.* **78**: 557-560.

Rosenthal, H.E. (1967) A graphic method for the determination and presentation of binding parameters in a complex system. *Anal. Biochem.* **20**: 525-532.

Ruan, X., Bhattacharjee, H., and Rosen, B.P. (2006) Cys-113 and Cys-422 form a high affinity metalloid binding site in the ArsA ATPase. *J. Biol. Chem.* **281**: 9925-9934.

Saitou, N., and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406-425.

Sambrook, J., Fritsch, E.F., and Maniatis, T., (1989) *Molecular cloning, a laboratory manual*. Cold Spring Harbor Laboratory, New York.

San Francisco, M.J., Hope, C.L., Owolabi, J.B., Tisa, L.S., and Rosen, B.P. (1990) Identification of the metalloregulatory element of the plasmid-encoded arsenical resistance operon. *Nucleic Acids Res* **18**: 619-624.

Sapsford, K.E., Berti, L., and Medintz, I.L. (2006) Materials for fluorescence resonance energy transfer analysis: beyond traditional donor-acceptor combinations. *Angew. Chem. Int. Ed. Engl.* **45**: 4562-4589.

Shi, W.P., Wu, J.H., and Rosen, B.P. (1994) Identification of a putative metal-ninding site in a new family of metalloregulatory proteins. *J. Biol. Chem.* **269**: 19826-19829.

Vagin, A.A., Steiner, R.A., Lebedev, A.A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G.N. (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D. Biol. Crystallogr.* **60**: 2184-2195.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A., and Wilson, K.S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.* **67**: 235-242.

Xu, C., and Rosen, B.P., (1999) Metalloregulation of soft metal resistance pumps. In: Metals and Genetics. B. Sarkar (ed). New York: Plenum Press, pp. 5-19.

Yang, J., Rawat, S., Stemmler, T.L., and Rosen, B.P. (2010) Arsenic binding and transfer by the ArsD As(III) metallochaperone. *Biochemistry* **49**: 3658-3666.

Ye, J., Kandegedara, A., Martin, P., and Rosen, B.P. (2005) Crystal structure of the *Staphylococcus aureus* pI258 CadC Cd(II)/Pb(II)/Zn(II)-responsive repressor. *J. Bacteriol.* **187**: 4214-4221.

Zhou, T., and Rosen, B.P. (1997) Tryptophan fluorescence reports nucleotide-induced conformational changes in a domain of the ArsA ATPase. *J. Biol. Chem.* **272**: 19731-19737.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher’s web-site.

Table 1. Crystal data and refinement statistics. The numbers in parentheses are the data for the highest resolution shell.

Data Collection	
Diffraction source	ALS BEAMLINE 8.2.2
Wavelength (Å)	0.999
Resolution range (Å)	47.11- 2.1 (2.175- 2.1)
Space group	P 1 21 1
a, b, c (Å)	51.7, 58.5, 53.5
α, β, γ (°)	90 118.28 90
Total reflections	32773 (3283)
Unique reflections	16526 (1653)
Multiplicity	2.0 (2.0)
Completeness (%)	98.79 (99.94)
Mean I/sigma(I)	13.87 (4.62)
Wilson B-factor	27.99
R-merge	0.0437 (0.155)
R-meas	0.0618 (0.219)
R-pim	0.0437 (0.155)
CC1/2	0.986 (0.95)
CC*	0.997 (0.99)
Refinement	
Reflections used in refinement	16348 (1652)
Reflections used for R-free	872 (56)
R-work (%)	0.22
R-free (%)	0.26
Number of non-hydrogen atoms	1812
Macromolecules	1701
Water	111
Protein residues	225
RMS(bonds)	0.008
RMS(angles)	1.22
Ramachandran favored (%)	98.64
Ramachandran allowed (%)	1.36

Ramachandran outliers (%)	0.00
Rotamer outliers (%)	7.18
Clash score	4.30
Average B-factor (Å2)	31.66
Macromolecules	31.42
Water	35.27
PDB ID	6UVU

549

550

Figure legends

Figure 1: Multiple sequence alignment of AntR with selected members of the ArsR//SmtB family. A. Multiple alignment (accession numbers in parentheses). *C. testosteroni* JL40 AntR (WP034375793.1); *Synechocystis* sp. PCC 7942 SmtB (CAA45872); *C. glutamicum* CgArsR1 (CAF21518); *A. ferrooxidans* (ACK80311), *S. putrefaciens* SpArsR (ADV53698): plasmid R773 ArsR (CAA34168); and plasmid pI258 CadC (P20047). The multiple alignment was calculated with CLUSTAL W. Cysteine residues are highlighted in yellow. Identical (black highlight) and conservative replacments (grey highlight) are identified. **B. Evolutionary relationships of AntR with other metal (metalloid) repressors from members of other bacterial species.** A neighbor joining phylogenetic tree shows that AntR is more close to SmtB, not arsenic responsive repressor ArsR.

Figure 2. The bacterial biosensor with the *antR gene* responds only to Sb(III) and MAs(III). Conditions for constitutive, repressed or derepressed *gfp* expression. In cells of *E. coli* AW3110 with both plasmids, *antR* is not expressed in the absence of arabinose, and *gfp* expression is constitutive, producing cellular fluorescence. In the presence of arabinose, *antR* is expressed, and *gfp* is repressed, so the cells are not fluorescent. In the presence of both arabinose and arsenical inducer, *gfp* expression is derepressed, and the cells are fluorescent. Expression of the *gfp* reporter gene was assayed as described under Experimental Procedures. Cells of *E. coli* strain AW3110(DE3) bearing wild type *antR* were grown without arabinose, 0.2% arabinose, or 0.2% arabinose and metalloids at the indicated

concentrations. Fluorescence intensities were quantified by spectrofluorometry. The data are the mean ± SE (n = 3).

Figure 3. Role of cysteine residues in binding of Sb(III) and MAs(III) to AntR. Expression of the *gfp* reporter gene was assayed as described under Experimental Procedures. Cells were grown in low phosphate medium for 14 h with 0.5% glycerol as carbon source. Cells of *E. coli* strain AW3110(DE3) bearing plasmids with wild type AntR, C103S or the C113S mutant *in trans* with reporter plasmid pACYC184-ParsP-gfp were grown with 0.2% arabinose and the indicated concentrations of metalloids. **(A),** Comparison of the response of the bacterial biosensor to arsenicals and antimonite. Response mutant biosensors **(B)** C58S; **(C)** C103S; and **(D)** C113S to the indicated inducers indicated. 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).

Figure 4. Metalloid binding affinities by AntR determined by the quenching of the single tryptophan AntR Y96W derivative. The relative affinities for Sb(III), MAs(III) and As(III) were estimated from the quenching of tryptophan fluorescence at the indicated concentrations of metalloids, as described under Materials and Methods.

Figure 5: Protein-ligand interactions determined by isothermal titration calorimetery. ITC was performed with purified AntR as described under materials and methods with **A)** Sb(III); **B)** MAs(III) and **C)** As(III).

Figure 6. AntR negatively regulates expression of the *ant* promoter. (**A**) The FAM-labeled *antR* promoter probe interacts with purified AntR, retarding the electrophoretic mobility. The amounts of DNA probes and AntR are shown above each panel. Binding of substrate releases AntR from the DNA, increasing the electrophoretic mobility. Metal(loids) were (**B**) Sb(III); (**C**) As(III); (**D**), As(V); (**E**) Cu(II); or (**F**), Zn(II) at the indicated concentrations. The amount of DNA probe was 100 ng, the amount of AntR was 1.0 μM.

Figure 7: The structural architecture of AntR. The structure of the AntR homodimer is shown in ribbon representation. Chain A is shown in green, and Chain B is shown in cyan. Structural elements are labeled. Cysteine residues are shown in ball-and-stick. Dashed lines indicate the distances between the sulfur atoms of Cys103 and Cys113.