Structural and mechanistic analysis of the arsenate respiratory reductase provides insight into environmental arsenic transformations*

Nathaniel R. Glasser[a], Paul. H. Oyala[b], Thomas H. Osborne[c], Joanne M. Santini[c], and Dianne K. Newman[a,d,1]

[a]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125; [b]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125; [c]Institute of Structural and Molecular Biology, Division of Biosciences, University College London, London, WC1E 6BT, United Kingdom; and [d]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125

Abstract

Arsenate respiration by bacteria was discovered over two decades ago and is catalyzed by diverse organisms using the well-conserved Arr enzyme complex. Until now, the mechanisms underpinning this metabolism have been relatively opaque. Here we report the first structure of an Arr complex (solved by X-ray crystallography to 1.6 Å resolution), which was enabled by an improved Arr expression method in the genetically-tractable arsenate respirer *Shewanella* sp. ANA-3. We also obtained structures bound with the substrate arsenate (1.8 Å), the product arsenite (1.8 Å), and the natural inhibitor phosphate (1.7 Å). The structures reveal a conserved active-site motif that distinguishes Arr ((R/K)GRY) from the closely-related arsenite respiratory oxidase (Arx) complex (XGRGWG). Arr activity assays using methyl viologen as the electron donor and arsenate as the electron acceptor display two-site ping-pong kinetics. A Mo(V) intermediate was detected with electron paramagnetic resonance spectroscopy, which is typical for proteins with a molybdopterin guanine dinucleotide cofactor. Arr is an extraordinarily fast enzyme that approaches the diffusion limit (K_m = 44.6 ± 1.6 μM, k_{cat} = 9,810 ± 220 s^{-1}) and phosphate is a competitive inhibitor of arsenate reduction (K_i = 325 ± 12 μM). These observations, combined with knowledge of typical sedimentary arsenate and phosphate concentrations and known rates of arsenate desorption from minerals in the presence of phosphate, suggest that: (1) arsenate desorption limits microbiologically-

* The authors declare that they have no conflicts of interest with the contents of this article. The atomic coordinates and structure factors (codes 6CZ7, 6CZ8, 6CZ9, and 6CZA) have been deposited in the Protein Data Bank (http://wwpdb.org/).

[1] To whom correspondence may be addressed: California Institute of Technology, 1200 E California Blvd., Mail Code 147-75, Pasadena, CA 91125. Tel.: 626-395-3543; Fax: 626-395-4135; E-mail: dkn@caltech.edu.

induced arsenate reductive mobilization and (2) phosphate enhances arsenic mobility by stimulating arsenate desorption rather than by inhibiting it at the enzymatic level.

Significance

Microbial arsenate respiration enhances the mobility of arsenic and contributes to the poisoning of tens of millions of people worldwide. Our ability to quantitatively predict how microbial activities shape arsenic geochemistry depends on a detailed understanding of how the enzymes that catalyze arsenate reduction work under environmentally relevant conditions. The structural and kinetic findings of the Arr enzyme complex reported here both help rationalize its cellular localization and allow us to predict that the rate of arsenate release from minerals likely constrains its activity in sedimentary environments. Moreover, this work illustrates that engineering environmental bacteria to over-express their native proteins can be straightforward, a strategy that may advance the study of enzymes that are challenging to express in traditional hosts.

Introduction

Arsenic mobilization into drinking water is a health threat affecting tens of millions of people worldwide (1). This threat often originates from the underlying mineralogy and microbiology of water reservoirs (2). In oxidizing environments, the predominant water-soluble form of arsenic is arsenate (found as $H_2AsO_4^-$ and $HAsO_4^{2-}$ with a pK_a of 6.94; we use arsenate to reference both species interchangeably), which adsorbs onto minerals and does not rapidly contaminate water reservoirs (2). However, if the water becomes polluted with organic material, microbes can consume the dissolved oxygen, leading to a decline in reduction potential and the conversion of arsenate to arsenite (found as H_3AsO_3) (3). Compared to arsenate, arsenite more readily leeches from mineral surfaces into water (4). Further exacerbating the problem, many sediments contain bacteria and/or archaea with the capacity to generate energy by respiring arsenate, reducing it to arsenite (5–9), and thereby stimulating the release of arsenic into the

aqueous phase (10, 11). Consequently, understanding the microbiology and biochemistry of arsenate respiration is of great interest to public health.

A microbe's capacity for arsenate respiration corresponds with the genes encoding the arsenate respiratory reductase (Arr) complex (12). The Arr complex is a member of the DMSO reductase family of molybdoenzymes and comprises two proteins, a large subunit ArrA and a smaller subunit ArrB (12, 13). The ArrA subunit is predicted to contain a Mo atom coordinated by a molybdopterin guanine dinucleotide cofactor (Mo-*bis*MGD), as well as a single [4Fe–4S] cluster (13, 14). The ArrB subunit co-purifies with ArrA (13, 15, 16) and is predicted to contain several [4Fe–4S] and/or [3Fe–4S] clusters (12, 14). The ArrA subunit contains a predicted N-terminal twin-arginine translocation (TAT) signal (12, 14), suggesting the assembled ArrAB complex is transferred post-translationally to the periplasm, and its periplasmic localization has been confirmed experimentally (15, 16). By analogy to the well-studied DMSO and nitrate reductase complexes (17, 18), the reaction mechanism likely involves oxygen atom transfer from arsenate to the Mo atom followed by two 1-electron reductions of the Mo atom (Scheme 1). In some organisms, ArrAB likely associates with a third subunit ArrC, a heme-containing quinol oxidase that is presumably the electron donor for the complex (17, 19). In species lacking ArrC, such as *Shewanella* sp. ANA-3 (20), ArrAB instead acquires electrons from the tetraheme cytochrome CymA (19).

Although arsenate respiration is a specialized microbial metabolism, small cytosolic arsenate reductases (ArsC and Acr2p) are present in numerous microorganisms (3). In contrast to Arr, which enables arsenate respiration using electrons from the quinone pool, ArsC and Acr2p likely evolved to detoxify cytoplasmic arsenate (21) and work through a disulfide redox cascade with glutaredoxin or thioredoxin (21–23). Emphasizing their separate physiological roles, *Shewanella* sp. ANA-3 does not require the Ars detoxification system for arsenate respiration (20). While Arr is significantly more effective than Ars at reducing arsenate in equilibrium with iron minerals (24), both Arr and Ars have been detected in arsenic-contaminated soils (6) and both can contribute to solubilized arsenate reduction (20, 25, 26). Given the

comparatively limited biochemical characterizations of Arr, it is difficult to disentangle the relative contribution of these two systems to arsenic mobilization. A major challenge for mechanistic enzymological studies of Arr has been obtaining enough protein for spectroscopic studies. The Arr complex expresses poorly in the heterologous host *Escherichia coli* (16), and ordinarily native induction and purification generates large volumes of arsenic-laden waste (13, 15). In this work, we successfully overcame both obstacles by overexpressing the Arr in its native host, *Shewanella* sp. ANA-3. With an abundance of protein, we optimized Arr crystal growth and used X-ray crystallography to solve the first structure of an Arr complex. We also used enzyme kinetics and EPR spectroscopy to validate the presumed overall reaction cycle. Together, these mechanistic studies advance our ability to predict the extent to which microbial arsenate respiration will mobilize arsenic in the environment.

Results

Overview of the ArrAB complex

We obtained initial Arr crystal hits using protein from the heterologous host *E. coli* (16). To purify more protein for crystal optimization, we developed an overexpression vector for the native host *Shewanella* sp. ANA-3 (see SI Materials and Methods), providing ~100-fold greater yields compared to *E. coli* (up to 5 mg of protein per liter of culture, Fig. S1 *A* and *B*). Subsequent crystal optimization (Fig. S1*C*) enabled us to solve the X-ray crystal structure of ArrAB, bound to the substrate arsenate and product arsenite, and to the inhibitor phosphate. Data collection and refinement statistics are shown in Table 1. Purified Arr crystallized as the ArrAB heterodimer with two non-crystallographic symmetry (NCS) copies of ArrAB per asymmetric unit. The single ArrAB heterodimer is consistent with the size of the complex observed during size-exclusion chromatography (13, 15, 16), and so it likely represents the physiological complex rather than the crystallographic dimer of heterodimers. The structures do not include the N-terminal TAT sequence as it was removed to facilitate expression and purification. For consistency, the residue numbers refer to positions in the full open reading frame including the TAT sequence with the N-terminal methionine starting at position 1.

The overall structure and arrangement of the complex (Fig. 1*A*) resembles that of other Mo-*bis*MGD protein complexes including nitrate reductase (NarGH) (27), formate dehydrogenase-N (FdnGHI) (28), polysulfide reductase (PsrABC) (29), and perchlorate reductase (PcrAB) (30). All five Fe-S metal clusters are [4Fe−4S] and all twenty iron atoms are coordinated by cysteine side chains. The ArrA and ArrB subunits are bound together by extensive surface contact covering approximately 3400 $Å^2$. Reflecting their close phylogeny (12, 31), the backbones of ArrAB and PsrAB are almost completely superimposable including their [4Fe–4S] clusters and Mo-*bis*MGD cofactors (Fig S2*A*). The active sites of ArrA and PsrA are distinguished by a unique set of residues that occupy distinct positions in the protein fold (Fig. S2 *B* and *C*).

ArrA and the arsenate binding site

The ArrA subunit contains the 4-domain structure of related Mo-*bis*MGD proteins (17, 18) with a funnel leading from solvent to the active site. The substrate funnel is lined mostly with basic and aromatic residues (Fig. S3). The Mo atom coordination geometry is distorted octahedral, in contrast to the trigonal prismatic geometry in PsrA (29), and it includes the dithiolene of both MGD cofactors with the cysteine side chain of Cys193 in an apical position (Fig. 1*B*). The electron density map suggests an oxygen atom, observed in one NCS copy and refined to 0.55 occupancy, completes the coordination geometry in the equatorial plane. The weak density and Mo–O bond distance of 2.3 Å likely indicates reduction of an oxo group (Mo=O) in the synchrotron X-ray beam (18). As in PsrA (29), an arginine side chain (Arg165) provides a hydrogen bond to the MoO group, but it is present in a different loop in ArrA and donates with $N^{\eta 2}$ instead of N^{ε} (Fig. 1*B* and Fig. S2*B*). The [4Fe–4S] cluster in ArrA (FS0) is coordinated by cysteines 61, 64, 68, and 96, in contrast to NarG and PsrA which contain a coordinating histidine (27, 30).

Co-crystallization yielded structures bound to arsenate (Fig. 2 *A* and *B*). Occupancy was highest at pH 6.0, lower at pH 7.5. and unobservable at pH 8.5, suggesting ArrA might bind the $H_2AsO_4^-$ state of arsenate. Alternatively, these binding differences might reflect changes in the protein's protonation or subtle changes in structure. Unexpectedly, we observed arsenate in two alternate, mutually exclusive

conformations, each with an occupancy of approximately 0.5. This model is supported by two discrete features in the anomalous difference map collected at a wavelength of 1.0 Å, representing the locations of the central As atoms (Fig 2 *A* and *B*). Both anomalous peaks disappeared at a wavelength of 1.1 Å, consistent with the As K-edge at 1.0448 Å. In the conformation closest to the Mo atom, arsenate forms hydrogen bonds with Arg165, His189, Tyr166, Tyr210, and several water molecules (Fig. 2*A*) with an As–Mo distance of 5.2 Å. In the second conformation, two waters substitute for arsenate oxygen atoms from the first conformation, and arsenate is hydrogen-bonded with Tyr166, Ser190, Lys198, and multiple water molecules (Fig. 2*B*) with an As–Mo distance of 7.3 Å. The alternative sites might reflect differences between the oxidized Mo(VI)O and the X-ray reduced Mo(V)OH (or Mo(IV) + H_2O) forms of the protein. It is important to note that we are unable to resolve individual hydrogen atoms at the resolution of our data (*ca.* 1.8 Å), and the proposed hydrogen bonds are hypothesized based on bond distances and angles.

We also obtained structures bound to arsenite, the product of arsenate respiration. Much like the first arsenate binding conformation, arsenite is coordinated by residues Arg165, Tyr166, His189, and Tyr210, but arsenite is shifted closer to the Mo atom with a Mo–As distance of 3.8 Å (Fig. 2*C*). In this position, arsenite forms a 2.3 Å Mo–O bond. While the proximity of arsenite to Cys193 resembles that observed in the ArsC enzymes, which form a covalent Cys-S–As intermediate with a bond distance of 2.2 Å (32), the crystallographic Arr Cys-S–As distance is too long to represent a covalent bond (3.0 Å as refined, or 2.8 Å to the center of the anomalous As peak). This structure might represent a non-physiological form of the enzyme, as it likely illustrates arsenite-binding to Mo(IV) or Mo(V) rather than to Mo(VI) as expected in the overall reaction cycle (Scheme 1). Similar structures have been observed in arsenite-inhibited aldehyde reductase (33, 34) and proposed as a dead-end pathway for the arsenite oxidase Aio (35).

Co-crystallization with phosphate also yielded structures bound to phosphate at pH 6.0, but not at pH 7.5. The structure resembles that of arsenate in the second conformation, with coordination by Tyr166,

Ser190, Lys198, and several water molecules (Fig. 2*D*). We did not observe phosphate binding at the other arsenate site, suggesting the two sites have different arsenate specificities.

We found no differences in the protein side chains between arsenate and phosphate binding. However, we did observe a subtle change in dithiolene bond angles of the P-MGD cofactor upon arsenite binding. In the substrate-free and arsenate-bound structures, the S-C-C-S group is nearly planar (Fig. S4), consistent with sp^2 hybridization of the dithiolene carbon atoms. In the arsenite-bound structure, there is a small but notable deviation from planarity (Fig. S4) that suggests partial reduction of the dithiolene carbons. As Arr is bidirectional *in vitro* (36), arsenite might be the electron donor for this reduction, and the shift might indicate that the dithiolene is electronically coupled to the Mo atom.

ArrB and the electron transfer pathway

ArrB shares the same overall fold of related electron transfer subunits in the Mo-*bis*MGD family (27–30). It contains four [4Fe–4S] clusters extending linearly from FS0 in ArrA to the distal FS4 (Fig. 1*A*). The coordinating cysteines differ in arrangement from the original prediction (12) and are as follows: (FS1) 12, 15, 18, 183; (FS2) 22, 164, 167, 179; (FS3) 57, 60, 65, 99; and (FS4) 69, 89, 92, 95. The last cluster (FS4) is separated from solvent by a single layer of the peptide backbone. By comparison to PsrABC, this region likely binds to the membrane protein CymA *in vivo*, and one NCS copy of ArrB forms extensive crystallographic contacts near FS4. The FS4 cluster is located near a protruding ridge distinguished by a series of positively-charged residues including Lys67, Lys94, Lys93, Arg159, Lys158, and Arg113 (Fig. S5), and it is tempting to speculate that these residues interact with CymA.

Arr and Arx have different active sites

The Arx family of arsenite oxidases forms a clade closely related to Arr (37, 38). Both enzymes are bidirectional (36) but differ in the physiological direction of activity. An open question is whether Arr and Arx directionality is determined by the enzyme active site, redox potentials of the metal cofactors, the electron donor or acceptor in the electron transport chain, or a combination thereof (36). The structure of Arr reveals a conserved distinction between Arr and Arx in the active site, suggesting the two complexes

are tuned for directionality. In Arr, the motif beginning at position 163 is (R/K)GRY (Fig. 3 and Fig. S6) and is conserved with few exceptions across several hundred partial ArrA sequences analyzed previously (39). In Arx, the motif is XGRGWG (Fig. 3 and Fig. S6), where X is F, L, T, or Y in the sequence diversity observed to date. It is unclear whether the Arx equivalent of Tyr166 is Gly or Trp. In Arr, Arg163 and Tyr166 form a π-stacking interaction that might affect the pK_a of the tyrosine side chain. Accordingly, one possibility for Arx is that Trp substitutes for Tyr in the active site, forming a hydrogen bond with arsenite or stabilizing the transition state. Because tryptophan is a hydrogen bond donor, this model is difficult to reconcile with the high pK_a of arsenite (9.2). Alternatively, this position might be occupied by Gly, allowing for hydrogen bonding with water or the protein backbone, or possibly disallowing hydrogen bonding to one –OH group of arsenite. Future structural and mutational studies of Arx will resolve the functions of these residues.

Arr exhibits two-site ping-pong kinetics

While confirming the activity of our enzyme preparations, we discovered some technical caveats that placed uncertainty on the originally reported arsenate K_m of 5 μM and V_{max} of 11,111 s^{-1} (16). Firstly, we found that dithionite, used to generate the electron donor methyl viologen radical ($MV^{\bullet+}$) for the colorimetric assay, interfered with the assay at pH 7.0 and below, possibly due to contaminating sulfite in our stock of commercial dithionite (40). Secondly, we found that the affinity for $MV^{\bullet+}$ is surprisingly low compared to other Mo-*bis*MGD reductases, which was not accounted for previously. In light of these considerations, we performed new kinetic experiments using Ti(III) citrate as a substitute for dithionite.

In our hands, the *Shewanella* sp. ANA-3 Arr complex has maximum activity at pH 7.5 (Fig. 4*A*) and above 200 mM NaCl (Fig. 4*B*) (as opposed to pH 7.0 and 150 mM NaCl as previously assayed (16)). The comparatively low affinity for $MV^{\bullet+}$ enabled us to test for the ping-pong kinetics predicted by the proposed reaction cycle (Scheme 1). As expected, measurements of enzyme rate versus arsenate concentration, determined at multiple $MV^{\bullet+}$ concentrations, produces parallel lines on a double-reciprocal plot (Fig. 4*C*), illustrating the classical picture of ping-pong kinetics. A global non-linear fit of a ping-

pong kinetic model to the data yields the following parameters (±1 standard error): arsenate K_m = 44.6 ± 1.6 μM, $MV^{\bullet+}$ K_m = 59.1 ± 2.4 μM, and k_{cat} = 9,810 ± 220 s^{-1} (as arsenate reduced).

Based on the chemical similarity between arsenate and phosphate, and similar binding sites in our crystal structures, we expected phosphate to act as a competitive inhibitor for arsenate reduction. Rate measurements as a function of arsenate, with and without phosphate, confirmed competitive inhibition as illustrated by intersecting lines on a double-reciprocal plot (Fig. 4*D*). In contrast, similar rate measurements as a function of $MV^{\bullet+}$ yielded parallel lines on a double-reciprocal plot (Fig. 4*E*), indicative of uncompetitive inhibition. Using the K_m and k_{cat} values determined above, a global non-linear fit to the inhibition data yields a phosphate K_i of 325 ± 12 μM. The different modes of phosphate inhibition indicate that the two substrates, $MV^{\bullet+}$ and arsenate, bind at distinct locations in the enzyme, consistent with arsenate reduction at the Mo atom in ArrA and protein reduction at FS4 in ArrB.

The Arr reaction cycle produces Mo(V)

A feature common to Mo-*bis*MGD proteins is the generation of a discrete Mo(V) species in the reaction cycle (17, 18). A notable exception is the arsenite oxidase (Aio) complex for which no Mo(V) intermediate has been observed (41, 42). Because Arr and Aio catalyze the reverse reactions of one another, we wondered whether the Arr mechanism would more closely resemble that of Aio than the other characterized Mo-*bis*MGD proteins. In comparison to Mo(IV) and Mo(VI), which are invisible in EPR spectroscopy, Mo(V) produces a characteristic EPR signal. We therefore used EPR to confirm the presence of a Mo(V) species during arsenate reduction. In an effort to capture the species produced during the reaction cycle, we analyzed Arr in a reaction mixture containing dithionite, arsenate, and varying concentrations of methyl viologen.

An X-band CW EPR spectrum collected on the reaction mixture produced an intense EPR signal at 120 K (Fig. 5*A*) attributable to Mo(V). This signal can be distinguished as arising from Mo(V) as opposed to the [4Fe–4S] clusters based on the acquisition temperature. Signals from at least one class of [4Fe–4S] cluster become evident only upon lowering temperature to 20 K (Fig. S7). In control experiments, the reaction

mixture or protein alone yielded no measurable signal. The Mo(V) EPR spectrum (Fig. 5*A*) exhibits a roughly axial symmetry (g = [1.990 1.980 1.957]), with clear low-intensity satellite peaks at the low and high field edges of the spectrum (Fig. 5*B*) due to hyperfine interactions between the unpaired electron and the two magnetic nuclei of molybdenum (^{95}Mo and ^{97}Mo, both I = 3/2; 15.92% and 9.55% natural abundance, respectively). Simulations including these interactions alone could not reproduce the complexity of the experimental spectrum (Fig. S8), requiring the inclusion of an additional hyperfine coupling to produce the fine structure evident in the spectrum.

Simulations of the X-band CW EPR spectrum consistent with the observed data could be achieved with inclusion of either a large, relatively isotropic hyperfine coupling to a single proton ($|^{1}$H $A|$ = [44 49 45] MHz, 99.99% natural abundance, I = 1/2) or with a somewhat smaller hyperfine coupling to arsenic ($|^{75}$As $A|$ = [23 23 32]; 100% natural abundance, I = 3/2) in addition to a relatively rhombic hyperfine coupling to Mo ($|^{95/97}$Mo $A|$ = [122 68 143]). Because CW-EPR is typically insensitive to the relative sign of each of the principle components of the hyperfine tensor, ^{75}As A could be [23, 23, 32], [−23, −23, 32], [23, 23, −32], etc. and so we only note the absolute value here. In the case of the ^{75}As simulation, addition of a nuclear quadrupole interaction was not necessary to simulate the CW spectrum, but an upward limit was able to be placed on this term of $|^{75}$As $P| \leq$ [12, 12, 24]. Both simulation models are consistent with a Mo(V) species that could be generated in the reaction cycle of Arr, with a large isotropic ^{1}H hyperfine coupling consistent with Mo(V)–OH as previously observed in xanthine oxidase (43, 44), or with a ^{75}As hyperfine coupling consistent with an arsenite-bound Mo(V) species as has been observed in arsenite-inhibited xanthine oxidase (45) and aldehyde reductase (33, 34).

To interrogate the identity of the observed Mo(V) species, Q-band (34 GHz) pulse EPR and electron nuclear double resonance (ENDOR) spectroscopies were used to further evaluate the potential magnetic interactions. The higher-frequency Q-band field-swept ESE-EPR spectrum provides a more rigorous constraint on the g-values due to the increase in resolution of these field-dependent parameters in comparison to X-band (Fig. 5*C*). Field-dependent Q-band ^{1}H Davies ENDOR (Fig. 5*D*) revealed the

presence of only relatively small hyperfine couplings to protons, with the largest coupling corresponding to ^{1}H A = [8.5, 5.2, 5.2] MHz. This hyperfine interaction is not large enough to account for the resolved splittings present in the X-band CW spectrum (Fig. S8), and the lack of any features consistent with the large ^{1}H hyperfine coupling necessary to reproduce the CW spectrum eliminates the possibility of any appreciable Mo(V)–OH being present in the experimental samples. The ^{1}H hyperfine coupling observed in the ENDOR likely corresponds to the proximal β-proton of the Mo-ligating Cys193, which should be within approximately 3.2 Å of the Mo(V) center. Attempts to detect signals associated with ^{75}As via ENDOR were unsuccessful, likely owing to a combination of decreased sensitivity due to the lower gyromagnetic ratio of ^{75}As in comparison to ^{1}H (^{75}As γ/^{1}H γ = 0.172), as well as the potential for significant hyperfine anisotropy due to localization of electron density in As p-orbitals, and a large ^{75}As nuclear quadrupole interaction which would cause substantial broadening of the ENDOR spectrum at any given field. These factors would not limit resolution of these interactions in the CW-EPR spectrum and would not be present for ^{1}H, and so we assign the observed species to an arsenite-bound Mo(V).

We observed the Mo(V)–arsenite signal only with higher MV concentrations (1 nM to 1 μM). Based on the rate of Arr (k_{cat} = 9,810^{-1}), these concentrations are likely to allow near completion of the reaction over the time scale of sample preparation (1–5 minutes). Lower MV concentrations (0.1 nM) produced no measurable signal, and we were unable to observe a signal consistent with Mo(V)–OH. We interpret this result to indicate that the putative Mo(V)–OH intermediate does not accumulate to measurable amounts during turnover conditions. By driving the reaction to completion, we instead captured an off-pathway Mo(V)–arsenite conjugate that formed after converting all arsenate into arsenite. Consistent with this interpretation, we observed no Mo(V) signal when the enzyme was incubated with arsenite and MV (Fig. S9), indicating that generating the signal requires reduction (*e.g.* by dithionite) as observed for xanthine oxidase (33, 34) or possibly enzyme turnover. Nonetheless, our detection of the Mo(V)–arsenite conjugate establishes the ability of Arr to adopt a Mo(V) redox state in contrast to Aio, suggesting Arr behaves similarly to other well-characterized Mo-*bis*MGD enzymes.

Discussion

Microbial arsenate reduction is a key driver of arsenic mobilization in carbon-rich anoxic environments, where arsenate is commonly found adsorbed to minerals. To better understand this biogeochemical phenomenon, we determined the X-ray crystal structure of the Arr complex and performed a detailed characterization of its kinetics. Our results help clarify what may limit arsenate respiration in the environment and provide insight into the cell biological strategy taken to achieve it.

Two very different enzyme systems can catalyze arsenate reduction: ArsABC and ArrAB. The cytosolic ArsABC detoxification system reduces solubilized arsenate, but it cannot stimulate the release of arsenic from minerals (26). In contrast, the Arr respiration system is presumed to mobilize mineral-associated arsenic (10, 24), despite the fact that the Arr enzyme complex is localized intracellularly. Arr must therefore access aqueous phase arsenate that dissociates from minerals. Arsenate adsorption and desorption have been modeled as multi-step processes (46). In a fast phase (minutes to hours), arsenate rapidly equilibrates with surface accessible binding sites. In a subsequent slow phase (hours to weeks), arsenate subsequently equilibrates with less accessible sites, possibly representing slow diffusion through the insoluble material, and it is this slower phase that likely controls arsenic mobility (47). The slow phase in model soils proceeds with first-order rate constants of 0.1–0.8 h^{-1} (3×10^{-5}–2×10^{-4} s^{-1}) for adsorption and 0.007–0.035 h^{-1} (2×10^{-7}–1×10^{-5} s^{-1}) for desorption (46). In comparison, our kinetic data reveal that Arr is an exceptionally fast enzyme that approaches the diffusion limit. With a rate constant (k_{cat} = 9,810 s^{-1}) around 8 orders of magnitude higher than that for arsenate adsorption, Arr is poised to easily outcompete mineral surfaces for arsenate. A similar rate has been reported for the Arr complex from *Chrysiogenes arsenatis* (15) (k_{cat} = 14,000 s^{-1}). A slower rate has been reported for *Bacillus selenitireducens* (13) (k_{cat} = 5 s^{-1}), possibly owing to metal lability during purification, assay artifacts caused by dithionite, or a low affinity for $MV^{\bullet+}$ leading to rate underestimation. It is possible that the rate-limiting step of arsenate respiration *in vivo* is oxidation of the electron acceptor (at CymA or ArrC), analogous to how the similarly fast Aio is limited by electron transfer to cytochrome *c* (48), but given the vast differences in their rate constants, it is likely that biological arsenate reduction by ArrAB is orders of

magnitude faster than arsenate adsorption. Together, this kinetic picture is consistent with observations that arsenate respiration proceeds markedly slower with solid-phase arsenate (10, 26), suggesting that desorption is the rate-limiting step of arsenate mobilization in the environment.

In terms of arsenate acquisition, the similarity between phosphate and arsenate ions poses challenges for arsenate respirers. As a natural phosphate analog, arsenate is improperly incorporated by phosphorylases into arsenate esters with aqueous half-lives up to 100,000-fold shorter than the corresponding phosphate esters (49), leading to an uncoupling of energy generation through ADP-arsenate hydrolysis (50). Arsenate enters bacterial cells through the low-affinity Pit phosphate transporter (51). Using Pit with a V_{max} of 55 nmol P_i (mg cell dry weight)$^{-1}$ min^{-1} (51) devoted solely to arsenate transport, we estimate (using a cell weight of 1 pg) a maximal uptake rate of approximately 5×10^5 cell^{-1} s^{-1}. Although this rate is comparable to the arsenate reduction rate observed in cultures (approximately 1×10^5 cell^{-1} s^{-1}) (10) and this method of arsenic acquisition would not be rate-limiting *per se*, it would flood the cytoplasm with toxic arsenic species. Moreover, we found phosphate to be an effective inhibitor of Arr (K_i = 325 μM) at levels far below the typical intracellular phosphate concentration of 10 mM (52). By localizing Arr to the periplasm, cells solve both problems simultaneously by providing a segregated environment for arsenate respiration. Compared to the intracellular environment, the average phosphate concentration in arsenic-rich ground waters is much lower and ranges from 7 μM to 100 μM (53). Assuming the periplasmic phosphate concentration reflects that of the extracellular milieu, these concentrations are not expected to significantly inhibit arsenate respiration. Additionally, in arsenic-rich environments, bacteria (including *Shewanella* sp. ANA-3, NCBI genome accession NC_008577.1) often contain the high-affinity Pst phosphate transporter (51) that can maintain up to a 4500-fold selectivity for phosphate over arsenate (54). With a K_m of 0.4 μM (51), the Pst system might lower the periplasmic phosphate concentration further below that which would inhibit arsenate respiration. Moreover, even in artificial medium with up to 5 mM phosphate, *arrA* gene expression is unaffected in *Shewanella* sp. ANA-3 (55), and higher phosphate concentrations actually increase growth rates from arsenate respiration, presumably by

315 protecting against the toxicity of arsenic (55). Together, these considerations suggest that environmental
316 phosphate concentrations are not a significant contributing factor for arsenate respiration rates. By
317 competing with arsenate for mineral sorption sites (56), phosphate likely enhances arsenic mobility rather
318 than inhibiting it through biological mechanisms.

319 Consistent with our kinetic results, our crystal structures illustrate that Arr struggles to distinguish
320 between phosphate and arsenate with both ions binding at the same site (Fig. 2 *B* and *D*). Phosphate
321 binding at this site sterically blocks arsenate from accessing the catalytic Mo atom. Interestingly, we also
322 observed arsenate, but not phosphate, at a second overlapping site (Fig. 2*A*) that might be more selective
323 for arsenate. Binding at this site might reflect changes in the protein structure owing to Mo-reduction in
324 the synchrotron X-ray beam, or it might compete with the non-specific site. Phosphate selectivity has
325 been studied in detail using sub-Å resolution crystal structures of PstS (54). This protein achieves
326 selectivity through a single rigid hydrogen bond that is positioned to interact with the specific bond
327 distances and angles of phosphate (54). While the resolution of our data (1.8 Å) and diffuse substrate
328 electron density (Fig. 2) preclude a similar analysis here, we note that arsenate in the specific site forms
329 hydrogen bonds *ca.* 2.6 Å from Tyr210 and Arg165 (Fig. 2*A*), similar to the 2.5 Å distance observed in
330 PstS phosphate selectivity. Both residues are universally conserved in all known ArrA homologs (Fig.
331 S6), suggesting that, like PstS, Arr might use a hydrogen bonding mechanism to enforce arsenate
332 selectivity near the catalytic Mo atom.

333 Regardless of their mechanistic roles, the active site residues revealed by the Arr structure provide a
334 powerful handle for the annotation of ambiguous sequences as Arr or Arx. The sequence similarity
335 between these enzymes makes it difficult to distinguish them without a phylogenetic analysis (38, 39), but
336 doing so is essential to understand the directionality of arsenic redox transformations based on
337 metagenomic data. The (R/K)GRY motif in Arr and the XGRGWG motif in Arx provide a structural
338 basis for this distinction. Although PCR amplification with *arrA* primers has been used to identify Arr-
339 encoding organisms in the environment (24, 39), the similarity between ArrA and ArxA raises questions

about the specificity of this detection. For example, the "Cluster V" group of *arrA* sequences identified by PCR (39) contains the Arx motif, suggesting this group is Arx rather than Arr, even though the amplified sequences appeared more related to *arrA* than to *arxA* (39). Interestingly, two annotated ArxA sequences from *Methanoperedens* (57) deviate from the ArxA motif and contain SGRN(R/Q)G (Fig. S5). These sequences form a sister group to the other ArxA members, suggesting the full sequence diversity of Arr and Arx has yet to be explored, and it is presently ambiguous whether these *Methanoperedens* enzymes operate as Arr or Arx.

Both the crystal structure (Fig. 2*C*) and the EPR spectrum of Arr (Fig. 5) suggest that the Arr Mo atom forms stable complexes with arsenite. Similar structures have been observed in xanthine oxidase (33, 34) for which arsenite is a potent inhibitor (K_d < 0.1 μM for the reduced protein) (58). However, unlike the case of xanthine oxidase, arsenite is a relevant product necessarily encountered by Arr as part of its function. Highlighting this point, arsenite is more effective than arsenate at inducing *arr* gene expression in *Shewanella* sp. ANA-3 (55), possibly because its increased mobility makes it a more reliable indicator for the presence of environmental arsenic (55). Interestingly, although we readily observed a Mo(V)-arsenite EPR signal in Arr (Fig. 5), a similar signal has never been observed for the arsenite oxidase Aio (41, 42), a distantly related molybdoenzyme to Arr and Arx (38, 59). Compared to Arr and Arx, the Aio active site contains primarily basic residues (42, 60) and, unique amongst all known Mo-*bis*MGD proteins, it does not coordinate its Mo atom with the polypeptide chain (18, 42, 60). Moreover, although both Arr and Arx are bidirectional enzymes (36), Aio has only been observed to operate unidirectionally (59), a possible phenotypic manifestation of these biochemical differences. The Mo-arsenite conjugate we observed is likely an off-target pathway that forms in an abundance of arsenite, and the distinct active-site motif of ArxA might discourage unproductive Mo-arsenite complexes in arsenite-rich environments. The more drastic differences observed in Aio might also allow it to operate at higher arsenite concentrations than Arr.

In conclusion, this study has clarified important aspects of microbial arsenate respiration that allow its environmental activity to be better predicted. In addition, the native host overexpression system described here will enable the nature of the relationship between Arr, Arx, and Aio to be probed mechanistically. Future studies of the Arr mechanism will benefit from a side-by-side comparison to Arx, especially as it pertains to the functions of the distinct active-site residues, and we anticipate that expression in *Shewanella* or a similarly facile native host will provide ample amounts of Arx. By establishing the mechanisms used to control the directionality of arsenic redox transformations, such comparisons will provide further insight into arsenic geochemical cycling.

Methods

Please see SI Materials and Methods for information regarding cell culturing, protein expression and purification, crystallography, enzyme activity assays, and electron paramagnetic resonance spectroscopy.

Acknowledgements

This research is rooted in work performed in the laboratory of Francois M. M. Morel (MIT/Princeton). DKN dedicates this paper to him on the occasion of his upcoming retirement and in gratitude for decades of outstanding mentorship. We thank Shu-ou Shan and Lisa Racki for assistance with enzyme kinetics, and we thank Barbara Schoepp-Cothenet for correspondence about the phylogeny of Arr and Arx. The Caltech Molecular Observatory provided essential training, knowledge, and equipment for the crystallography in this work.

The Caltech EPR Facility was supported by NSF-1531940. This research used resources of the Advanced Light Source, which is a DOE Office of Science User Facility under contract no. DE-AC02-05CH11231. The Molecular Observatory is supported by the Gordon and Betty Moore Foundation, the Beckman Institute, and the Sanofi-Aventis Bioengineering Research Program. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515.

The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the National Institutes of Health, National Institute of General Medical Sciences (including P41GM103393). JMS funding was provided by a Biotechnology and Biological Sciences Research Council (BBSRC) grant (BB/N012674/1). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH.

Author Contributions

N.R.G. developed the Arr expression protocol, performed crystallography, and wrote the manuscript with D.K.N. N.R.G. and D.K.N. prepared EPR samples, and P.H.O collected and interpreted the EPR spectra and wrote sections of the manuscript pertaining to EPR. T.H.O. provided the initial protein samples for crystallography. N.R.G. and T.H.O. performed enzyme kinetics. D.K.N. and J.M.S. coordinated the project. All authors reviewed and edited the manuscript.

References

1. Naujokas MF, et al. (2013) The broad scope of health effects from chronic arsenic exposure: update on a worldwide public health problem. *Environ Health Perspect* 121(3):295–302.
2. Oremland RS, Stolz JF (2005) Arsenic, microbes and contaminated aquifers. *Trends Microbiol* 13(2):45–49.
3. Zhu Y-G, Yoshinaga M, Zhao F-J, Rosen BP (2014) Earth Abides Arsenic Biotransformations. *Annu Rev Earth Planet Sci* 42(1):443–467.
4. Appelo CAJ, Van Der Weiden MJJ, Tournassat C, Charlet L (2002) Surface Complexation of Ferrous Iron and Carbonate on Ferrihydrite and the Mobilization of Arsenic. *Environ Sci Technol* 36(14):3096–3103.
5. Kulp TR, et al. (2006) Dissimilatory Arsenate and Sulfate Reduction in Sediments of Two Hypersaline, Arsenic-Rich Soda Lakes: Mono and Searles Lakes, California. *Appl Environ Microbiol* 72(10):6514–6526.
6. Luo J, Bai Y, Liang J, Qu J (2014) Metagenomic Approach Reveals Variation of Microbes with Arsenic and Antimony Metabolism Genes from Highly Contaminated Soil. *PLOS ONE* 9(10):e108185.
7. Song B, Chyun E, Jaffé PR, Ward BB (2009) Molecular methods to detect and monitor dissimilatory arsenate-respiring bacteria (DARB) in sediments. *FEMS Microbiol Ecol* 68(1):108–117.
8. Osborne TH, McArthur JM, Sikdar PK, Santini JM (2015) Isolation of an Arsenate-Respiring Bacterium from a Redox Front in an Arsenic-Polluted Aquifer in West Bengal, Bengal Basin. *Environ Sci Technol* 49(7):4193–4199.
9. Hassan Z, et al. (2015) Diverse arsenic- and iron-cycling microbial communities in arsenic-contaminated aquifers used for drinking water in Bangladesh. *FEMS Microbiol Ecol* 91(4). doi:10.1093/femsec/fiv026.
10. Zobrist J, Dowdle PR, Davis JA, Oremland RS (2000) Mobilization of Arsenite by Dissimilatory Reduction of Adsorbed Arsenate. *Environ Sci Technol* 34(22):4747–4753.
11. Dhar RK, et al. (2011) Microbes Enhance Mobility of Arsenic in Pleistocene Aquifer Sand from Bangladesh. *Environ Sci Technol* 45(7):2648–2654.
12. Saltikov CW, Newman DK (2003) Genetic identification of a respiratory arsenate reductase. *Proc Natl Acad Sci* 100(19):10983–10988.
13. Afkar E, et al. (2003) The respiratory arsenate reductase from *Bacillus selenitireducens* strain MLS10. *FEMS Microbiol Lett* 226(1):107–112.
14. Stolz JF, Basu P, Santini JM, Oremland RS (2006) Arsenic and Selenium in Microbial Metabolism. *Annu Rev Microbiol* 60(1):107–130.
15. Krafft T, Macy JM (1998) Purification and characterization of the respiratory arsenate reductase of *Chrysiogenes arsenatis*. *Eur J Biochem* 255(3):647–653.
16. Malasarn D, Keeffe JR, Newman DK (2008) Characterization of the arsenate respiratory reductase from *Shewanella* sp. strain ANA-3. *J Bacteriol* 190(1):135–142.
17. Grimaldi S, Schoepp-Cothenet B, Ceccaldi P, Guigliarelli B, Magalon A (2013) The prokaryotic Mo/W-bisPGD enzymes family: A catalytic workhorse in bioenergetic. *Biochim Biophys Acta* 1827(8):1048–1085.
18. Hille R, Hall J, Basu P (2014) The Mononuclear Molybdenum Enzymes. *Chem Rev* 114(7):3963–4038.

19. Murphy JN, Saltikov CW (2007) The *cymA* Gene, Encoding a Tetraheme c-Type Cytochrome, Is Required for Arsenate Respiration in *Shewanella* Species. *J Bacteriol* 189(6):2283–2290.
20. Saltikov CW, Cifuentes A, Venkateswaran K, Newman DK (2003) The *ars* detoxification system is advantageous but not required for As(V) respiration by the genetically tractable *Shewanella* species strain ANA-3. *Appl Environ Microbiol* 69(5):2800–2809.
21. Oden KL, Gladysheva TB, Rosen BP (1994) Arsenate reduction mediated by the plasmid-encoded ArsC protein is coupled to glutathione. *Mol Microbiol* 12(2):301–306.
22. Messens J, et al. (2002) All intermediates of the arsenate reductase mechanism, including an intramolecular dynamic disulfide cascade. *Proc Natl Acad Sci* 99(13):8506–8511.
23. Messens J, et al. (2002) Kinetics and active site dynamics of *Staphylococcus aureus* arsenate reductase. *J Biol Inorg Chem* 7(1–2):146–156.
24. Malasarn D, et al. (2004) *arrA* is a reliable marker for As(V) respiration. *Science* 306(5695):455.
25. Macur RE, Jackson CR, Botero LM, Mcdermott TR, Inskeep WP (2004) Bacterial Populations Associated with the Oxidation and Reduction of Arsenic in an Unsaturated Soil. *Environ Sci Technol* 38(1):104–111.
26. Langner HW, Inskeep WP (2000) Microbial Reduction of Arsenate in the Presence of Ferrihydrite. *Environ Sci Technol* 34(15):3131–3136.
27. Jormakka M, Richardson D, Byrne B, Iwata S (2004) Architecture of NarGH reveals a structural classification of Mo-bisMGD enzymes. *Struct Lond Engl 1993* 12(1):95–104.
28. Jormakka M, Törnroth S, Byrne B, Iwata S (2002) Molecular basis of proton motive force generation: structure of formate dehydrogenase-N. *Science* 295(5561):1863–1868.
29. Jormakka M, et al. (2008) Molecular Mechanism of Energy Conservation in Polysulfide Respiration. *Nat Struct Mol Biol* 15:730.
30. Youngblut MD, et al. (2016) Perchlorate Reductase Is Distinguished by Active Site Aromatic Gate Residues. *J Biol Chem* 291:9190–9202.
31. Duval S, Ducluzeau A-L, Nitschke W, Schoepp-Cothenet B (2008) Enzyme phylogenies as markers for the oxidation state of the environment: The case of respiratory arsenate reductase and related enzymes. *BMC Evol Biol* 8(1):206.
32. Martin P, et al. (2001) Insights into the Structure, Solvation, and Mechanism of ArsC Arsenate Reductase, a Novel Arsenic Detoxification Enzyme. *Structure* 9(11):1071–1081.
33. Boer DR, Thapper A, Brondino CD, Romão MJ, Moura JJG (2004) X-ray Crystal Structure and EPR Spectra of "Arsenite-Inhibited" *Desulfovibrio gigas* Aldehyde Dehydrogenase: A Member of the Xanthine Oxidase Family. *J Am Chem Soc* 126(28):8614–8615.
34. Thapper A, Boer DR, Brondino CD, Moura JJG, Romão MJ (2007) Correlating EPR and X-ray structural analysis of arsenite-inhibited forms of aldehyde oxidoreductase. *J Biol Inorg Chem* 12(3):353–366.
35. Warelow TP, Pushie MJ, Cotelesage JJH, Santini JM, George GN (2017) The active site structure and catalytic mechanism of arsenite oxidase. *Sci Rep* 7(1):1757.
36. Richey C, et al. (2009) Respiratory arsenate reductase as a bidirectional enzyme. *Biochem Biophys Res Commun* 382(2):298–302.
37. Zargar K, Hoeft S, Oremland R, Saltikov CW (2010) Identification of a novel arsenite oxidase gene, *arxA*, in the haloalkaliphilic, arsenite-oxidizing bacterium *Alkalilimnicola ehrlichii* strain MLHE-1. *J Bacteriol* 192(14):3755–3762.

38. Zargar K, et al. (2012) ArxA, a new clade of arsenite oxidase within the DMSO reductase family of molybdenum oxidoreductases. *Environ Microbiol* 14(7):1635–1645.
39. Mirza BS, Sorensen DL, Dupont RR, McLean JE (2017) New Arsenate Reductase Gene (*arrA*) PCR Primers for Diversity Assessment and Quantification in Environmental Samples. *Appl Environ Microbiol* 83(4):e02725-16.
40. McKenna CE, Gutheil WG, Song W (1991) A method for preparing analytically pure sodium dithionite. Dithionite quality and observed nitrogenase-specific activities. *Biochim Biophys Acta* 1075(1):109–117.
41. Anderson GL, Williams J, Hille R (1992) The purification and characterization of arsenite oxidase from Alcaligenes faecalis, a molybdenum-containing hydroxylase. *J Biol Chem* 267(33):23674–23682.
42. Warelow TP, et al. (2013) The respiratory arsenite oxidase: structure and the role of residues surrounding the rieske cluster. *PloS One* 8(8):e72535.
43. Bray RC, Knowles PF, Pick FM, Vänngård T (1968) Electron-spin-resonance evidence for interaction of protons with Mo(V) in reduced forms of xanthine oxidase. *Biochem J* 107(4):601–602.
44. Gutteridge S, Tanner SJ, Bray RC (1978) The molybdenum centre of native xanthine oxidase. Evidence for proton transfer from substrates to the centre and for existence of an anion-binding site. *Biochem J* 175(3):869–878.
45. George GN, Bray RC (1983) Reaction of arsenite ions with the molybdenum center of milk xanthine oxidase. *Biochemistry (Mosc)* 22(5):1013–1021.
46. Zhang H, Selim HM (2005) Kinetics of Arsenate Adsorption−Desorption in Soils. *Environ Sci Technol* 39(16):6101–6108.
47. Badruzzaman M, Westerhoff P, Knappe DRU (2004) Intraparticle diffusion and adsorption of arsenate onto granular ferric hydroxide (GFH). *Water Res* 38(18):4002–4012.
48. Watson C, et al. (2017) Electron transfer through arsenite oxidase: Insights into Rieske interaction with cytochrome c. *Biochim Biophys Acta* 1858(10):865–872.
49. Long JW, Ray WJ (1973) Kinetics and thermodynamics of the formation of glucose arsenate. Reaction of glucose arsenate with phosphoglucomutase. *Biochemistry (Mosc)* 12(20):3932–3937.
50. Moore SA, Moennich DM, Gresser MJ (1983) Synthesis and hydrolysis of ADP-arsenate by beef heart submitochondrial particles. *J Biol Chem* 258(10):6266–6271.
51. Willsky GR, Malamy MH (1980) Characterization of two genetically separable inorganic phosphate transport systems in *Escherichia coli*. *J Bacteriol* 144(1):356–365.
52. Rosing J, Slater EC (1972) The value of ΔG° for the hydrolysis of ATP. *Biochim Biophys Acta* 267(2):275–290.
53. Mahin T, Ngai T, Murcott S, Mondal M (2008) *Importance of evaluating phosphate levels in tubewells in high arsenic areas of Asia*.
54. Elias M, et al. (2012) The molecular basis of phosphate discrimination in arsenate-rich environments. *Nature* 491(7422):134–137.
55. Saltikov CW, Wildman RA, Newman DK (2005) Expression Dynamics of Arsenic Respiration and Detoxification in *Shewanella* sp. Strain ANA-3. *J Bacteriol* 187(21):7390–7396.
56. Hongshao Z, Stanforth R (2001) Competitive Adsorption of Phosphate and Arsenate on Goethite. *Environ Sci Technol* 35(24):4753–4757.

57. Oremland RS, Saltikov CW, Stolz JF, Hollibaugh JT (2017) Autotrophic microbial arsenotrophy in arsenic-rich soda lakes. *FEMS Microbiol Lett* 364(15). doi:10.1093/femsle/fnx146.
58. Stewart RC, Hille R, Massey V (1984) Characterization of arsenite-complexed xanthine oxidase at room temperature. Spectral properties and pH-dependent redox behavior of the molybdenum-arsenite center. *J Biol Chem* 259(23):14426–14436.
59. van Lis R, Nitschke W, Duval S, Schoepp-Cothenet B (2013) Arsenics as bioenergetic substrates. *Biochim Biophys Acta* 1827(2):176–188.
60. Ellis PJ, Conrads T, Hille R, Kuhn P (2001) Crystal Structure of the 100 kDa Arsenite Oxidase from *Alcaligenes faecalis* in Two Crystal Forms at 1.64 Å and 2.03 Å. *Structure* 9(2):125–132.

Tables

Table 1

Crystallography data collection and refinement statistics. Values in parentheses are for the highest resolution shell.

	Arr 6CZ7	**Arr with Arsenate 6CZ8**	**Arr with Arsenite 6CZ9**	**Arr with Phosphate 6CZA**
Cryo protection				
Buffer	HEPES, pH 7.5	MES, pH 6.0	HEPES, pH 7.5	MES, pH 6.0
Data collection				
Space group	C2	C2	C2	C2
Unit cell				
a, *b*, *c* (Å)	231.26, 85.64, 146.68	233.162, 86.344, 148.221	232.512, 86.306, 147.367	232.397, 86.366, 147.429
α, β, γ (°)	90, 127.84, 90	90, 127.675, 90	90, 127.72, 90	90, 127.608, 90
Resolution range (Å)	46.91–1.62 (1.678–1.62)	39.1–1.78 (1.844–1.78)	39.37–1.80 (1.864–1.80)	37.70–1.71 (1.771–1.71)
Unique reflections	286352 (28526)	222359 (22113)	209823 (20962)	245023 (23784)
R_{pim} (%)	3.1 (32.0)	3.4 (30.3)	3.2 (31.3)	3.3 (31.1)
$\langle I/\sigma(I)\rangle$	17.77 (2.40)	15.94 (2.35)	17.49 (2.27)	15.71 (2.17)
Completeness (%)	99.93 (99.99)	99.85 (99.84)	98.35 (98.89)	98.33 (95.72)
Multiplicity	6.8 (6.6)	6.8 (7.0)	6.9 (7.0)	6.9 (6.6)
Wilson *B*-factor (Å^2)	18.10	23.40	23.77	20.40
Refinement				
R_{work}/R_{free}	0.1374/0.1547	0.1372/0.1681	0.1422/0.1694	0.1317/0.1535
No. of non-H atoms				
Protein	16452	16429	16320	16448
Water	2435	2241	2288	2378
RMSD				
Bond lengths (Å)	0.007	0.010	0.006	0.009
Bond angles (°)	1.27	1.23	1.16	1.29
Average *B*-factors (Å^2)				
Protein	25.85	30.22	31.21	26.76
Ligands	23.04	27.30	28.64	23.78
Water	35.34	38.37	39.37	36.77
Ramachandran (%)				
Favored	96.88	96.73	96.78	96.92
Allowed	2.88	2.98	2.98	2.79
Outliers	0.24	0.29	0.24	0.29

Table 2

EPR parameters for the Mo(V) signal observed in Arr under reaction conditions. The hyperfine and nuclear quadrupole coupling tensors are colinear with the *g*-tensor.

axis	*g*	$\lvert^{95/97}\text{Mo } A\rvert$ (MHz)	$\lvert^{1}\text{H } A\rvert$ (MHz)	$\lvert^{75}\text{As } A\rvert$ (MHz)	$\lvert^{75}\text{As } P\rvert$ (MHz)
1	1.990	122	8.5	23	≤ 12
2	1.980	68	5.2	23	≤ 24
3	1.957	143	5.2	32	≤ 24

Figures and Schemes

Scheme 1

Proposed reaction cycle for Arr based on the mechanism of other Mo-*bis*MGD proteins. The reduced form of Arr contains Mo(IV) coordinated by two MGD cofactors and a cysteine side chain (top left). Upon binding to the protein, arsenate coordinates to the Mo atom (top right). An oxygen atom from arsenate is transferred to the Mo atom to form Mo(VI) and arsenite (bottom right). The cofactor is regenerated by sequential one-electron reductions, first to the EPR active Mo(V) species (bottom left) and then to Mo(IV) (top left). Although we have illustrated it here as $H_2AsO_4^-$, the protonation state of bound arsenate is unknown and might instead be $HAsO_4^{2-}$. Also unknown is whether the final water dissociates as illustrated (bottom left to top left) or if is displaced when arsenate binds (top left to top right).

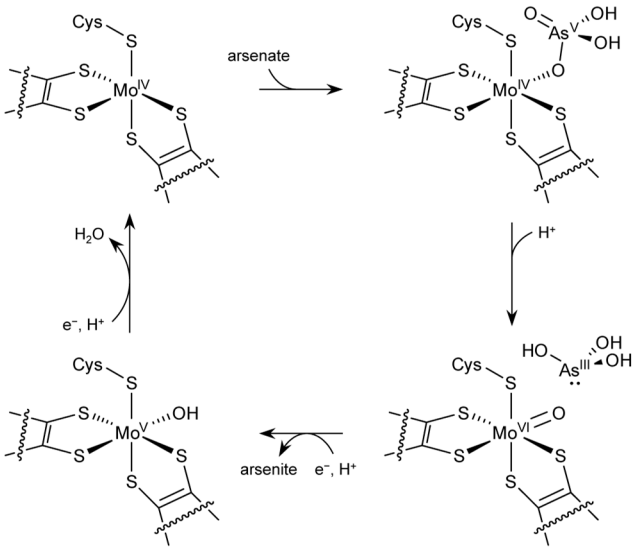


Figure 1

Overview of the ArrAB heterodimer from *Shewanella* sp. ANA-3.

(A) View of the ArrA (yellow) and ArrB (blue) subunits with cofactors highlighted. The [4Fe–4S] clusters (FS0–FS4) and Mo atom are shown as spheres (Fe in orange, S in yellow, and Mo in purple). The MGD cofactors are shown in red labeled Q and P according to the DMSO reductase nomenclature.
(B) Close-up view of the Mo-binding site in ArrA, focusing on the coordination of Mo (purple sphere) by water (red sphere), Cys193, and two MGD cofactors (C in gray, N in blue, O in red, S in yellow, and P in orange). Dashed lines illustrate coordination or hydrogen bonding.

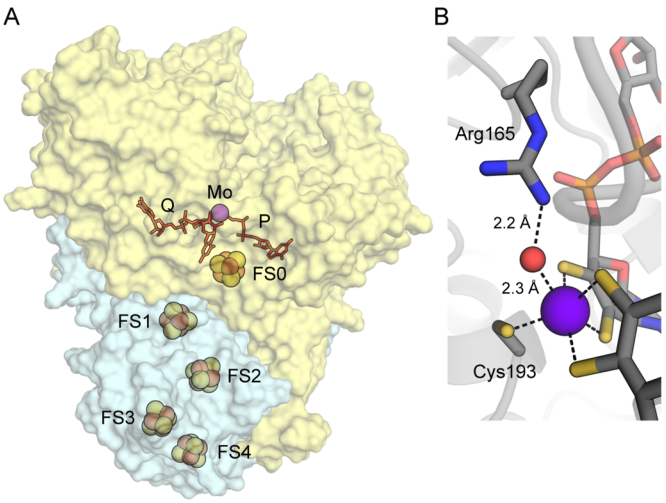


Figure 2

Close-up view of the arsenate-binding site in ArrA (C in gray, N in blue, O in red, S in yellow, P in orange, Mo in purple, As in light purple). The 2*mF*o – F_c map around the relevant portions is shown as a white surface contoured to 1.5σ. For clarity, the map is shown only around ligands, the Mo-*bis*MGD cofactor, and the Mo-coordinating cysteine. The anomalous map collected at a wavelength of 1 Å is shown around arsenic atoms contoured to 6σ. Dashed lines indicate possible hydrogen bonds or coordination to the Mo atom.

(A) Arsenate bound in the conformation nearest the Mo atom.
(B) Arsenate bound in a second conformation farther from the Mo atom. The two water molecules in teal substitute for oxygen atoms from the first conformation shown in *A*.
(C) Arsenite bound and coordinated to the Mo atom.
(D) Phosphate bound to the active site in a conformation resembling the arsenate conformation from *B*.

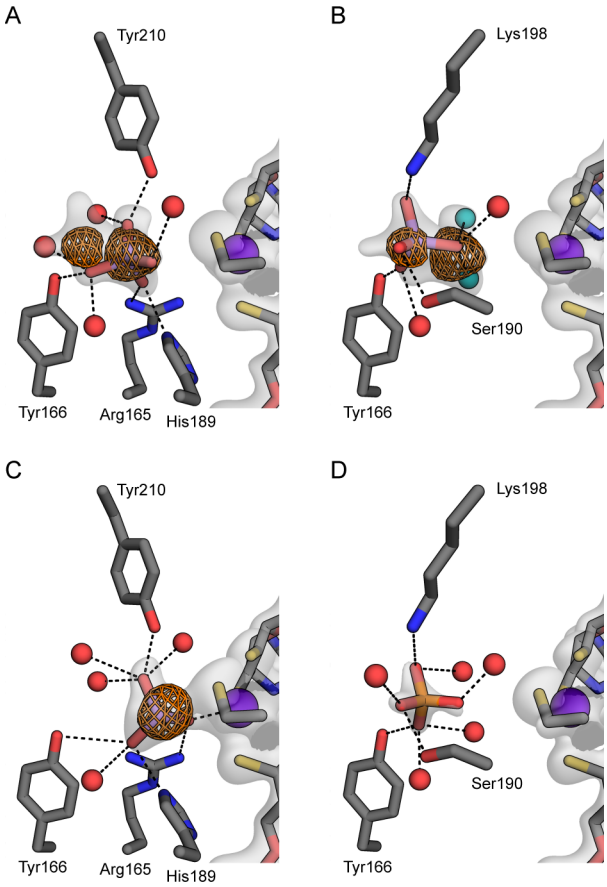


Figure 3

Alignment of representative ArrA and ArxA sequences highlighting the active site residues. A more complete comparison that includes accession numbers is provided in Fig. S6. Annotated ArrA enzymes contain a (R/K)GRY motif (blue), while annotated ArxA sequences contain a XGRGWG motif (green). Other active site residues are generally conserved between ArrA and ArxA (pink), including the Mo-coordinating Cys residue (yellow). Sequence numbers are relative to the ArrA from *Shewanella* sp. ANA-3.

Figure 4

Kinetic characterization of Arr. Enzyme rates are shown as molecules of arsenate reduced per second per molecule of protein. Dashed lines illustrate a kinetic model fit for each curve individually. Solid lines illustrate a single kinetic model fit to all datasets together. For clarity, double-reciprocal plots omit the *y*-axis tick marks and begin at 0 s^{-1}. For panels C–E the assays were performed at pH 7.5 and 300 mM NaCl.

(A) Relative activity of Arr as a function of pH.
(B) Relative activity of Arr as a function of NaCl concentration.
(C) Enzyme rates as a function of arsenate and $MV^{\bullet+}$ concentration. Parallel lines on the double-reciprocal plot (right) indicate ping-pong kinetics.
(D) Enzyme rates as a function of arsenate and phosphate concentration. The $MV^{\bullet+}$ concentration was 51.1 µM. Intersecting lines at the *y*-axis on the double-reciprocal plot (right) indicate competitive inhibition.
(E) Enzyme rates as a function of $MV^{\bullet+}$ and phosphate concentration. The arsenate concentration was 250 µM. Parallel lines on the double-reciprocal plot (right) indicate uncompetitive inhibition.

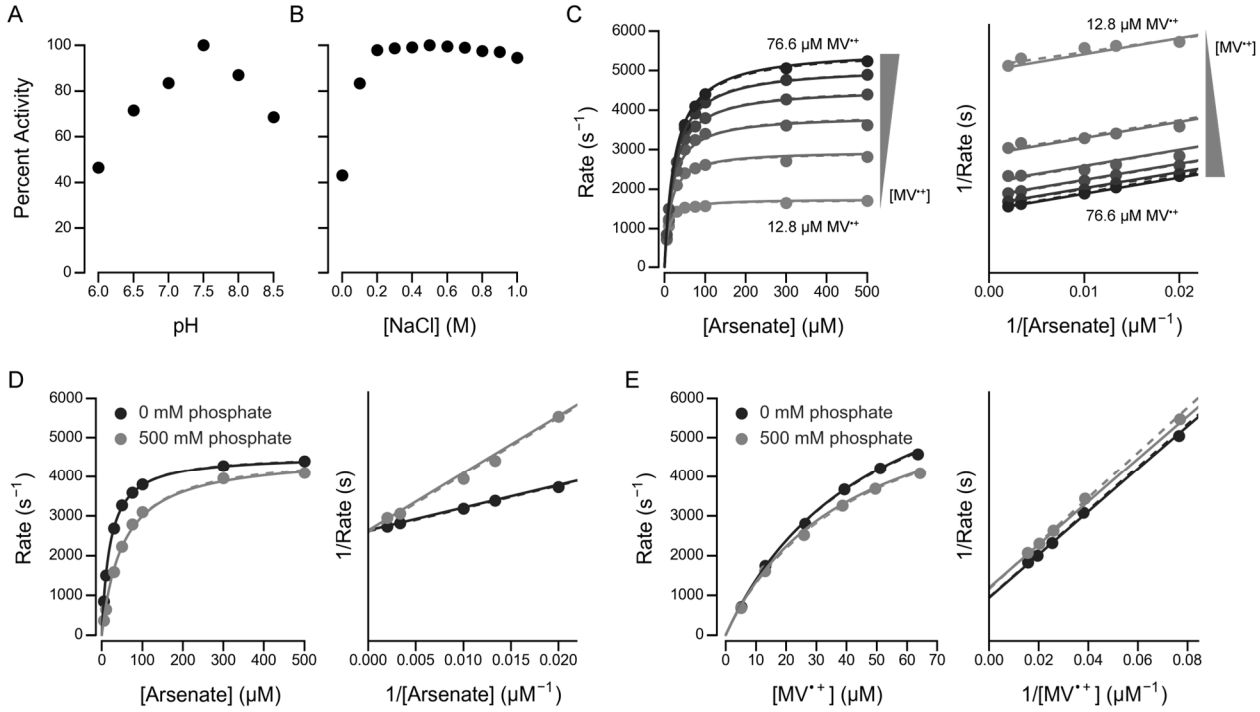
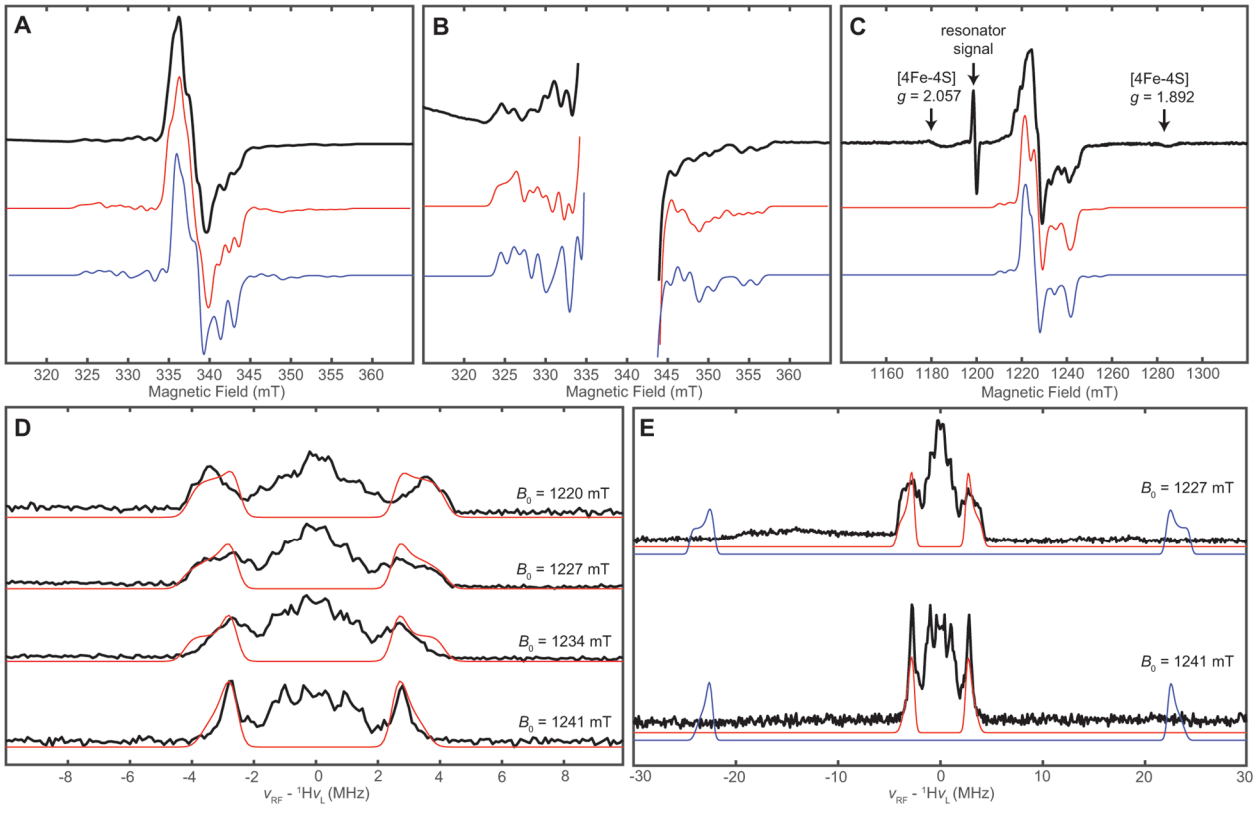


Figure 5

EPR characterization of the Mo(V) species generated in Arr. Samples contained 50 µM protein, 50 mM HEPES buffer (pH 7.5), 10 mM sodium dithionite, 1 µM methyl viologen, and 10 mM sodium arsenate. Simulation parameters are indicated in Table 2.

(A) X-band CW EPR spectrum of Arr acquired at 120 K (black trace) with a simulation of the arsenite-bound Mo(V) species (red trace) compared to a simulation of a putative Mo(V)–OH species (blue trace) with the same simulation parameters except that the coupling to ^{75}As has been replaced with a hyperfine coupling to ^{1}H of A = [44 49 45] MHz. Acquisition parameters: microwave frequency = 9.390 GHz; microwave power = 1 mW; modulation amplitude = 0.4 mT.
(B) Enlarged view of same X-band CW EPR spectrum (black trace) in *A* and simulations of the arsenite-bound Mo(V) species (red trace) and putative Mo(V)–OH species (blue trace) showing the low-intensity features from the hyperfine interaction with $^{95/97}$Mo.
(C) Pseudomodulated Q-band ESE-EPR spectrum of Arr acquired at 12 K (black trace) with a simulation of the arsenite-bound Mo(V) species (red trace) compared to a simulation of a putative Mo(V)–OH species (blue trace) with the same simulation parameters except that the coupling to ^{75}As has been replaced with a hyperfine coupling to ^{1}H of A = [44 49 45] MHz.
(D) Field-dependent Q-band ^{1}H Davies ENDOR of the Mo(V) species in Arr (black trace) with a simulation of the largest observed ^{1}H hyperfine coupling (red trace).
(E) Wide Q-band ^{1}H Davies ENDOR of the Mo(V) species in Arr (black trace) with a simulation of the largest observed ^{1}H hyperfine coupling (red trace) compared to a simulation of the expected ^{1}H ENDOR signals for the putative Mo(V)–OH species with ^{1}H A = [44 49 45] MHz which do not appear in the experimental spectrum to any significant degree.



PNAS
www.pnas.org

Supplementary Information for

Structural and mechanistic analysis of the arsenate respiratory reductase provides insight into environmental arsenic transformations

Nathaniel R. Glasser, Paul. H. Oyala, Thomas H. Osborne, Joanne M. Santini, and Dianne K. Newman

Dianne K. Newman
Email: dkn@caltech.edu

This PDF file includes:

Supplementary materials and methods
Figs. S1 to S9

SI Materials and Methods

Materials, Strains, and Growth Conditions

Crystallography reagents were from Hampton Research. L-arabinose was from Chem-Impex International. Terrific broth (Difco) was from BD Biosciences. Other reagents were from Sigma-Aldrich or Acros Organics and were of ACS grade or better. For routine culturing, Shewanella sp. ANA-3 and E. coli were grown in lysogeny broth (LB) containing 10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl. Solid media contained 15 g/L agar. E. coli was cultured at 37 °C and ANA-3 was cultured at 30 °C. Liquid cultures were incubated in a New Brunswick Innova 44R incubator shaking at 250 rpm (2.54 cm stroke length). In our hands, ANA-3 quickly lost viability on agar plates at 4 °C, and so liquid ANA-3 cultures were usually started directly from a frozen stock. All buffers were adjusted to the appropriate pH using NaOH or HCl.

Cloning

The final Arr expression vector was developed from iterative attempts to improve upon the original Arr expression vector for E. coli (1). The starting point was a pET15b vector (Novagen) modified to include a TEV cleavage site instead of a thrombin site. The primer AGAGACCATGGGCCATCACCATCACCATCACGACTACGACATCCCGACTACCGAAAA CCTGTACTTCCAGGGCATGCCTCGAGCACACA was annealed to its reverse complement and restriction digested with NcoI and XhoI. The resulting fragment was ligated to similarly-digested pET15b to create the pET15b-His_6-TEV vector. The *arrAB* ORF, without the TAT sequence, was PCR amplified from ANA-3 genomic DNA using primers ATTTTGTTTAACTTTAAGAAGGAGATATACCATGAAGAAAGAGAATCAAGTCAACTT GGG and GTGATGGTGATGGTGATGGCCAGCGGCAATCCCCTCGACAATAGG. The pET15b-His_6-TEV vector was linearized by PCR amplification with primers GGATCCGGCTGCTAACAAAGC and GCCCTGGAAGTACAGGTTTTCG, and the *arrAB* fragment was joined using Gibson assembly (2) to create pET15b-His_6-TEV-arrAB. The TAT sequence was then PCR amplified from ANA-3 genomic DNA using primers ATTTTGTTTAACTTTAAGAAGGAGATATACCATGAAGAAAGAGAATCAAGTCAACTT GGG and GTGATGGTGATGGTGATGGCCAGCGGCAATCCCCTCGACAATAGG, the pET15b-His6-TEV-arrAB vector was linearized using PCR with primers GGCCATCACCATCACCATC and GGTATATCTCCTTCTTAAAGTTAAAC, and the two fragments were joined using Gibson assembly to create pET15b-TAT-His_6-TEV-arrAB. For expression in *Shewanella*, this final construct was amplified using primers TTTTTTGGGCTAGCGAATTCAGGAGTGTGGAAAATGAAGAAAGAGAATCAAGTCAA CTTG and CGCCAAAACAGCCAAGCTTTTAATAAGCGGTTTTAACACCAAAAC, the vector pBAD18-kan (3) was linearized using PCR with primers GAATTCGCTAGCCCAAAAAAACGG and AAGCTTGGCTGTTTTGGCG, and the fragments were joined using Gibson assembly. The completed expression vector was transformed into wildtype *Shewanella* sp. ANA-3 by conjugation with *E. coli* (4) to create strain DKN1846. The strain was stored at -80 °C in LB with 15% (v/v) glycerol.

Arr purification

A typical Arr purification used six 2.8-L baffled Fernbach flasks with 1 L of medium each. The flasks were inoculated with 10 mL of an overnight culture of DKN1846. The cells were grown aerobically at 30 °C in terrific broth (containing 12 g/L casein digest, 24 g/L yeast extract, 9.4 g/L K_2HPO_4, and 2.2 g/L KH_2PO_4) with 4 mL/L glycerol, 5 mM $MgSO_4$, 200 µM Na_2MoO_4, 0.2 g/L ferric ammonium citrate, and 35 µg/mL kanamycin. At an OD_{600} of 2.5 (about 3.5 hours), Arr expression was induced by adding 20 mL of 1 M L-arabinose. The cultures were incubated for an additional 4 hours and harvested by centrifugation at 8000×*g* for 15 min. The cell pellets were washed once with cold Ni-binding buffer (50 mM HEPES, 500 mM NaCl, 15 mM

imidazole, pH 7.5), pelleted again, flash frozen in liquid nitrogen and stored at −80 °C until purification.

For purification, the cell pellets were thawed at room temperature and placed on ice. All subsequent steps were performed at 4 °C. The cell pellets were suspended with lysis buffer (50 mM Tris, 300 mM NaCl, 0.5% Triton X-100, pH 7.5) containing one ULTRA protease inhibitor tablet (Roche) per liter of cell culture. To induce lysis, EDTA and lysozyme were added to a concentration of 1 mM and 0.5 mg/ml, respectively. After 1 hour, $MgCl_2$ and $CaCl_2$ were added to 5 mM and the viscous mixture was treated with DNase I (approximately 200 Kuntz units per liter of culture). Once it was no longer viscous (1–2 hours), the lysate was clarified by centrifugation for 30 min at 50,000×*g*. The supernatant was applied to a gravity-flow column of His60 Ni Superflow Resin (Clontech) equilibrated with His-binding buffer with 0.1% Triton X-100. The column contained approximately 4 mL of resin per liter of culture. The column was washed with 5 column volumes of His-binding buffer with 0.1% Triton X-100, followed by 5 column volumes of His-binding buffer (without detergent). (Triton X-100 helped to reduce non-specific binding of *Shewanella* lysate to the Ni resin, but it was not necessary for Arr solubility or stability.) The protein was then eluted with a buffer containing 50 mM HEPES, 500 mM NaCl, and 200 mM imidazole (pH 7.5). The brown eluate was incubated overnight with 1 mM EDTA, 1 mM PMSF, 1 mM TCEP, and 2 mg of TEV protease. The next day, an equal volume of 4 M $(NH_4)_2SO_4$ was slowly added, and after 15 min the precipitated material was removed by centrifugation for 15 min at 5000×*g*. To ensure a homogenous protein redox state, arsenate was added to a final concentration of 10 mM. The protein was then passed through a 0.45 μm filter and applied to a 5-mL HiTrap Phenyl HP column (GE Healthcare) using an Äkta Purifier system. The column was washed with 5 column volumes of 50 mM HEPES, 2 M $(NH_4)_2SO_4$ (pH 7.5) and eluted with a linear gradient over 10 column volumes to 50 mM HEPES (pH 7.5). The brown fractions were pooled. Imidazole (pH 7.5) was added to a concentration of 10 mM and the protein was passed through a 5-mL HisTrap HP column (GE Healthcare) to remove residual uncleaved protein and other contaminants that bind to the Ni resin. EDTA was added to 1 mM (to chelate trace Ni from the Ni resin) and the protein was concentrated to less than 2 mL using an Amicon ultra centrifugal filter (30 kDa cutoff). The protein was finally passed through a HiLoad 16/600 Superdex 200 column (GE Healthcare) equilibrated with 50 mM HEPES (pH 7.5). The brown fractions were pooled and concentrated to 40–60 mg/mL. The protein was divided into 20 μL aliquots in 200-μL PCR tubes, flash frozen in liquid nitrogen, and stored at −80 °C.

The final yield of purified Arr was approximately 5 mg per liter of culture (as measured by a Bradford assay using BSA as the standard), and the product was >95% pure as judged by an SDS-polyacrylamide gel stained with Coomassie Blue (Fig. S1*A*). Metal analysis by ICP-MS (1), normalized to the protein concentration, indicated at least 95% Mo and 80% Fe saturation.

Crystallography

Initial crystal screens were set up in a sitting-drop format using an Art Robins Gryphon Nano liquid-handling robot to mix 0.2 μL of screen solution with 0.2 μL of protein solution (15 mg/ml in 50 mM HEPES, pH 7.5). The screens used were Crystal Screen HT (Hampton), Index HT (Hampton), PEGRx HT (Hampton), JCSG-plus HT-96 (Molecular Dimensions), and Wizard Classic 1 and 2 (Rigaku). Approximately 20 hits were obtained under a wide range of pH, PEG type, and salt type. Crystal optimization on a larger scale led to insurmountable skin growth at the liquid-air interface, and so we used microbatch under oil instead. The optimized crystals were grown by mixing 2 μL of protein solution (20 mg/ml in 50 mM HEPES, pH 7.5) with 2 μL of crystallization solution (30% PEG 2000 MME, 300 mM KSCN, 100 mM HEPES, pH 7.5) and covering the drop with 50 μL of paraffin oil in an MRC Under Oil 96 Well Crystallization Plate (Swissci). Where applicable, the crystallization solution also contained 10 mM arsenate or phosphate (diluted 1:1 with protein to create 5 mM in the final drop); arsenite inhibited crystal growth and was instead included at 5 mM in the cryo protection solution. Microseeding was

essential for reliable crystal growth, and so the crystallization solution also contained a serial dilution of crystals that were crushed by vortexing with a glass bead. The crystal trays were incubated at 20 °C. Light brown crystals were apparent after several hours and reached their maximum size overnight. The plate-like crystals were typically 200–300 μm long and wide and 20–50 μm thick. The crystals were slowly acclimated to a cryoprotection solution (30% PEG 2000 MME, 100 mM KSCN, 1 M Na formate, 50 mM HEPES, pH 7.5, containing 5 mM ligand where appropriate) and flash-frozen by plunging into liquid nitrogen.

The Arr structure was initially solved from a 2.0 Å resolution dataset collected using an in-house MicroMax 007-HF X-ray generator (Rigaku) (wavelength 1.5418 Å) and a RAXIS-IV++ detector (Rigaku). SAD phasing using shelxd (5) placed an iron atom at the center of each [4Fe–4S] cluster. These atoms were used as an initial model for site refinement in Phaser (6), which expanded each cluster into its constituent Fe atoms. A crude model was built using PHENIX Autobuild (7), which was recycled back into Phaser to improve the Fe sites, and the model was rebuilt from the refined phase information using Autobuild. The preliminary model was refined with phenix.refine (7) and iterative model building in Coot (8). A higher resolution, substrate-free dataset was collected at beamline 5.0.2 of the Advanced Light Source. Datasets for substrate binding were collected at beamline 12-2 of the Stanford Synchrotron Radiation Lightsource. The diffraction images were integrated using XDS (9), assigned a space group with POINTLESS (10), merged with AIMLESS (11), and converted to structure factors with CTRUNCATE (12). The structures were solved by molecular replacement with Phaser using the preliminary model. The [4Fe–4S] clusters and Mo-bisMGD cofactor were placed manually in Coot, and the models were refined with phenix.refine.

Arr activity assay and curve fitting

The Arr activity assay colorimetrically monitors oxidation of the methyl viologen radical ($MV^{\bullet+}$, ε_{605} = 13,700 M^{-1} cm^{-1} (13)) coupled to arsenate reduction. The reaction was followed at 605 nm using a Thermo Scientific Evolution 260 Bio spectrometer maintained at 30 °C. All reactions were performed inside an anaerobic chamber (Coy) with an atmosphere of 95% N_2 and 5% H_2. A stock of $MV^{\bullet+}$ was prepared by mixing methyl viologen dichloride and Ti(III) citrate to a final concentration of 1 mM and 0.5 mM, respectively; Ti(III) citrate was prepared by mixing $TiCl_3$ with a 10-fold excess of trisodium citrate. The stock $MV^{\bullet+}$ concentration was determined from the absorbance at 605 nm. A stock solution of 2 nM Arr was prepared in siliconized microcentrifuge tubes containing 50 mM HEPES (pH 7.5) and 0.1% Triton X-100. Diluted protein was prepared fresh and discarded within 1 hour to minimize adsorption to the tube. The reaction was started by adding 10 μL of an arsenate stock solution to 990 μL of reaction mix containing 50 mM buffer, 300 mM NaCl (unless indicated otherwise), $MV^{\bullet+}$, and 0.02 nM Arr. The buffers used were MES (pH 6.0 and 6.5), MOPS (pH 7.0), HEPES (pH 7.5), and Tris (pH 8.0 and 8.5). The reaction rate was determined from a linear fit to the first 5-10 seconds of the reaction (never more than 10% of the total reaction).

Non-linear curve fitting of the kinetic data was performed using the curve fit function of the SciPy optimization package for Python. The local fit for Michaelis–Menten kinetics used the model

$$V = V_{\max} \frac{[\mathrm{S}]}{K_\mathrm{m} + [\mathrm{S}]}$$

where V is the observed reaction rate, V_{max} is the maximum reaction rate under the given conditions, [S] is the varied substrate concentration (arsenate or $MV^{\bullet+}$), and K_m is the Michaelis constant. The global fit for ping-pong kinetics and competitive inhibition used the model

$$V = V_{\max} \frac{[\mathrm{As}][\mathrm{MV}]}{[\mathrm{As}]K_{\mathrm{m,MV}} + [\mathrm{MV}]K_{\mathrm{m,As}}\left(1 + \frac{[\mathrm{P}]}{K_{\mathrm{i,P}}}\right) + [\mathrm{As}][\mathrm{MV}]}$$

where V is the observed reaction rate, V_{max} is the maximum reaction rate, [As] is the arsenate concentration, [MV] is the $MV^{\bullet+}$ concentration, $K_{m,As}$ is the Michaelis constant for arsenate, $K_{m,MV}$ is the Michaelis constant for $MV^{\bullet+}$, [P] is the phosphate concentration, and $K_{i,P}$ is the inhibition constant for phosphate. The turnover number kcat was determined from $V_{max} = k_{cat}$ [E], where [E] is the enzyme concentration (0.02 nM). A stoichiometry of 2 $MV^{\bullet+}$ oxidized per 1 arsenate reduced was used to convert absorbance changes to arsenate reduction.

Electron paramagnetic resonance spectroscopy

Sample preparation: Samples were prepared anaerobically inside an anaerobic chamber (Coy) containing 95% N_2 and 5% H_2. A mixture was prepared containing 50 mM HEPES buffer (pH 7.5), 50 µM ArrAB, 10 mM sodium dithionite, 10 mM sodium arsenate, and 0.1 nM to 1 µM methyl viologen and transferred immediately to an EPR tube. The tube was loosely capped, removed from the anaerobic chamber, and immediately frozen in liquid nitrogen.

CW EPR spectroscopy: X-band (9.4 GHz) CW EPR spectra were acquired using a Bruker EMX spectrometer using Bruker Win-EPR software (ver. 3.0). For spectra acquired at 120 K and 20 K, temperatures were maintained using an Oxford Instruments ESR900 flow cryostat and an ITC-503 temperature controller. Spectra acquired at 77 K were collected using a vacuum-insulated quartz liquid nitrogen immersion dewar inserted into the EPR resonator. Spectra were simulated using the EasySpin simulation toolbox (14) (release 5.2.12) with Matlab R2016b.

Pulse EPR spectroscopy: All pulse Q-band (34 GHz) EPR and electron nuclear double resonance (ENDOR) spectra were acquired using a Bruker ELEXSYS E580 pulse EPR spectrometer equipped with a Bruker D2 resonator. Temperature control was achieved using an ER 4118HV-CF5-L Flexline Cryogen-Free VT cryostat manufactured by ColdEdge equipped with an Oxford Instruments Mercury ITC temperature controller.

Pulse Q-band electron spin-echo detected EPR (ESE-EPR) field-swept spectra were acquired using the 2-pulse "Hahn-echo" sequence ($\pi/2 - \tau - \pi -$ echo) and subsequently, each field swept echo-detected EPR absorption spectrum was modified using a pseudo-modulation function (modulation amplitude = 1.5 mT) to approximate the effect of field modulation and produce the CW-like 1st derivative spectrum (15). Acquisition parameters: temperature = 12 K; microwave frequency = 34.032 GHz; MW π pulse length = 32 ns; interpulse delay τ = 140 ns; shot repetition time (srt) = 5 ms.

Pulse Q-band ENDOR was acquired using the Davies pulse sequence ($\pi - T_{RF} - \pi_{RF} - T_{RF} - \pi/2 - \tau - \pi -$ echo), where T_{RF} is the delay between MW pulses and RF pulses, π_{RF} is the length of the RF pulse and the RF frequency is randomly sampled during each pulse sequence. All 1H ENDOR was acquired using the following acquisition parameters: Temperature = 12 K; microwave frequency = 34.032 GHz; MW π pulse length = 80 ns; interpulse delay τ = 260 ns; RF pulse length = 15 µs; TRF delay = 2 µs; shot repetition time (srt) = 5 ms; RF frequency randomly sampled.

In general, the ENDOR spectrum for a given nucleus with spin I = ½ (1H) coupled to the S = ½ electron spin exhibits a doublet at frequencies

$$\nu_{\pm} = \left|\frac{A}{2} \pm \nu_N\right| \quad (1)$$

Where ν_N is the nuclear Larmor frequency and A is the hyperfine coupling. For nuclei with $I \geq 1$ (14N, 2H), an additional splitting of the $\nu_{\pm}$ manifolds is produced by the nuclear quadrupole interaction (P)

$$\nu_{\pm,m_I} = \left| \nu_N \pm \frac{3P(2m_I - 1)}{2} \right| \tag{2}$$

Simulations of all EPR data were achieved using the EasySpin (14) simulation toolbox (release 5.2.12) with Matlab 2016b using the following Hamiltonian:

$$\hat{H} = \mu_B \vec{B}_0 g \hat{S} + \mu_N g_N \vec{B}_0 \hat{I} + h\hat{S} \cdot \boldsymbol{A} \cdot \hat{I} + h\hat{I} \cdot \boldsymbol{P} \cdot \hat{I} \tag{3}$$

In this expression, the first term corresponds to the electron Zeeman interaction term where μ_B is the Bohr magneton, g is the electron spin g-value matrix with principle components g = [gxx gyy gzz], and $\hat{S}$ is the electron spin operator. The second term corresponds to the nuclear Zeeman interaction term where μ_N is the nuclear magneton, g_N is the characteristic nuclear g-value for each nucleus (e.g. 1H, 75As, 95/97Mo) and $\hat{I}$ is the nuclear spin operator. The third term corresponds to the electron-nuclear hyperfine term, where $\boldsymbol{A}$ is the hyperfine coupling tensor which can typically be represented as a diagonal matrix with principle components $\boldsymbol{A}$ = [Axx Ayy Azz]. For nuclei with $I \geq 1$, the final term corresponds to the nuclear quadrupole (NQI) term which arises from the interaction of the nuclear quadrupole moment with the local electric field gradient (efg) at the nucleus, where $\boldsymbol{P}$ is the quadrupole coupling tensor. In the principle axis system (PAS), $\boldsymbol{P}$ is traceless and parametrized by the quadrupole coupling constant e^2Qq/h and the asymmetry parameter η such that:

$$\boldsymbol{P} = \begin{pmatrix} P_{xx} & 0 & 0 \\ 0 & P_{yy} & 0 \\ 0 & 0 & P_{zz} \end{pmatrix} = \frac{e^2Qq/h}{4I(2I-1)} \begin{pmatrix} -(1-\eta) & 0 & 0 \\ 0 & -(1+\eta) & 0 \\ 0 & 0 & 2 \end{pmatrix} \tag{4}$$

where $\frac{e^2Qq}{h} = 2I(2I-1)P_{zz}$ and $\eta = \frac{P_{xx}-P_{yy}}{P_{zz}}$. The asymmetry parameter may have values between 0 and 1, with 0 corresponding to an electric field gradient with axial symmetry and 1 corresponding to a fully rhombic efg.

The orientations between the hyperfine and NQI tensor principle axis systems and the g-matrix reference frame are defined by the Euler rotation angles (α, β, γ).

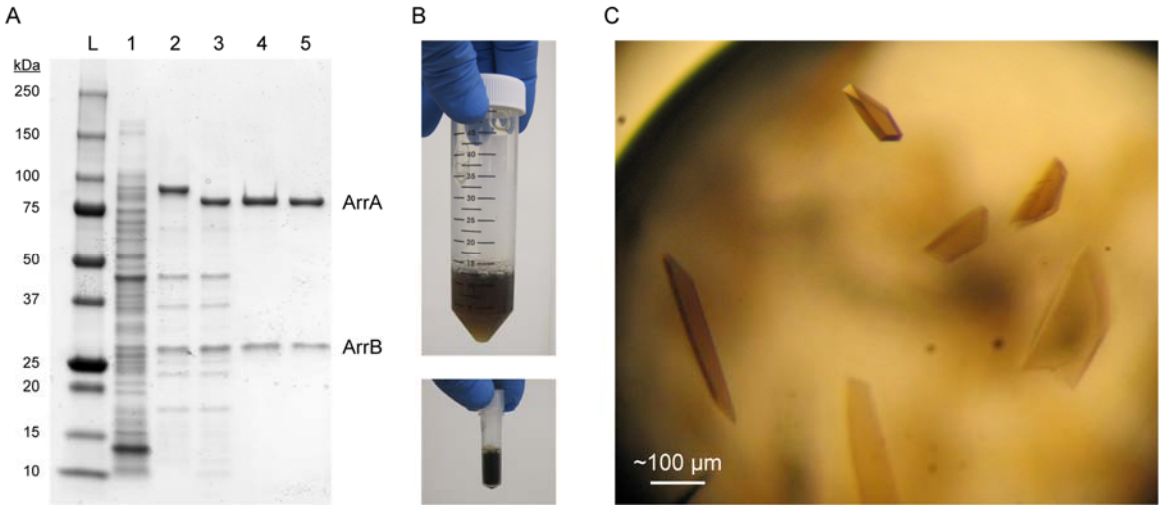


Fig. S1. Purification and crystallization of ArrAB from *Shewanella* sp. ANA-3. (A) SDS-polyacrylamide gel stained with Coomassie Blue showing the purification steps of ArrAB. The molecular weights of the ladder bands (L) are shown to the left. The lanes illustrate (1) clarified Shewanella lysate after L-arabinose induction of ArrAB, (2) after nickel affinity chromatography, (3) after tag cleavage by TEV protease, (4) after hydrophobic interaction chromatography, and (5) after size-exclusion chromatography. (B) Photographs of purified ArrAB. The deep brown color originates from the multiple Fe-S clusters in ArrAB. The top picture was taken immediately after elution from the nickel resin. The bottom picture shows the final product after purification and concentration (60 mg/ml). (C) Representative ArrAB protein crystals formed using the optimized microbatch conditions.

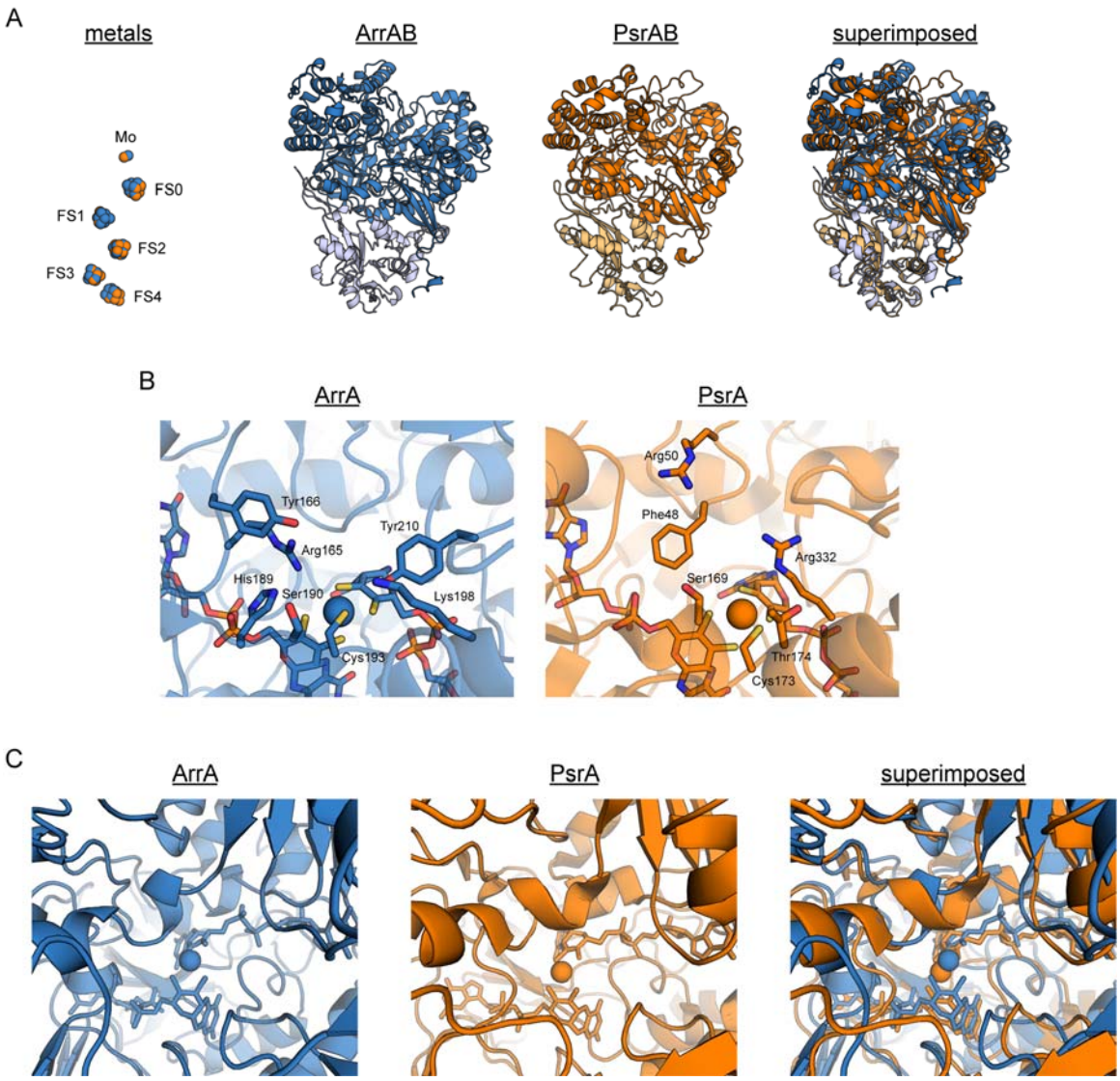


Fig. S2. Comparison of ArrAB to its nearest characterized homolog, PsrAB (PDB code 2VPZ) (16). ArrAB is shown in blue and PsrAB is shown in orange. For clarity, the PsrC subunit is omitted because the equivalent subunit for Arr was not determined in this work. (A) Overview of structural similarities between ArrAB and PsrAB. The metal cofactors (left) are nearly superimposable. The overall topology of the two enzymes is also similar. (B) Close-up of active site residues for ArrA determined in this work compared to those hypothesized for PsrA (16). Apart from Ser190 (ArrA) and Ser169 (PsrA), and the Mo-coordinating cysteine, the active site residues of the two enzymes share no similarity, and they are distinct in both identity and position. (C) View down the substrate binding funnel. Despite a clear distinction in active site residues, the backbones of ArrA and PsrA fold nearly identically. Differences in the active site arise primarily from the orientation and identity of the residue side chains.

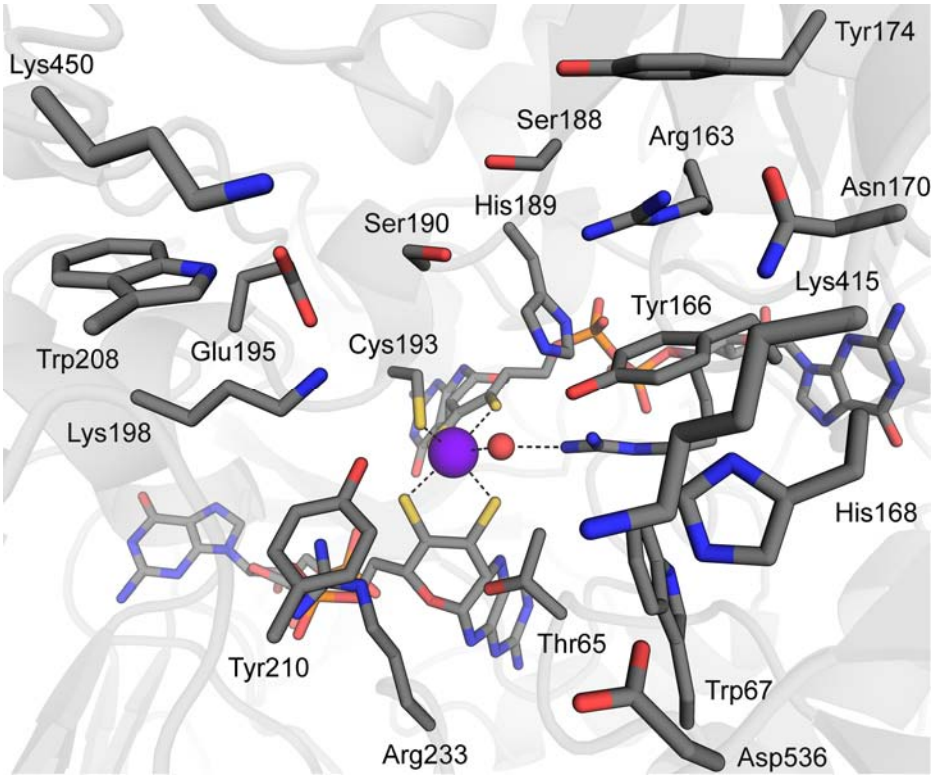


Fig. S3. View of residues lining the substrate binding funnel in ArrA. The color scheme is the same as in Fig. 1B (C in gray, N in blue, O in red, S in yellow, and P in orange, Mo in purple).

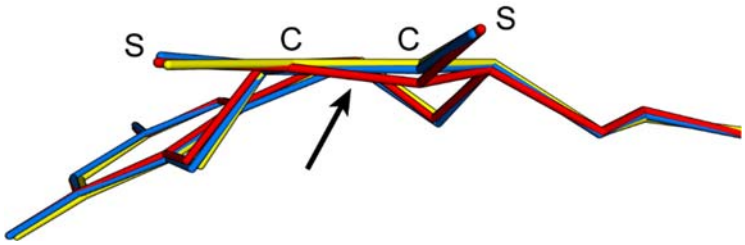


Fig. S4. Comparison of dithiolene bond angles in the P-pterin of ArrA in the absence of substrate (blue), in protein bound to arsenate (yellow), and in protein bound to arsenite (red). The arrow indicates the bond in the arsenite-bound structure that appears to be partially reduced.

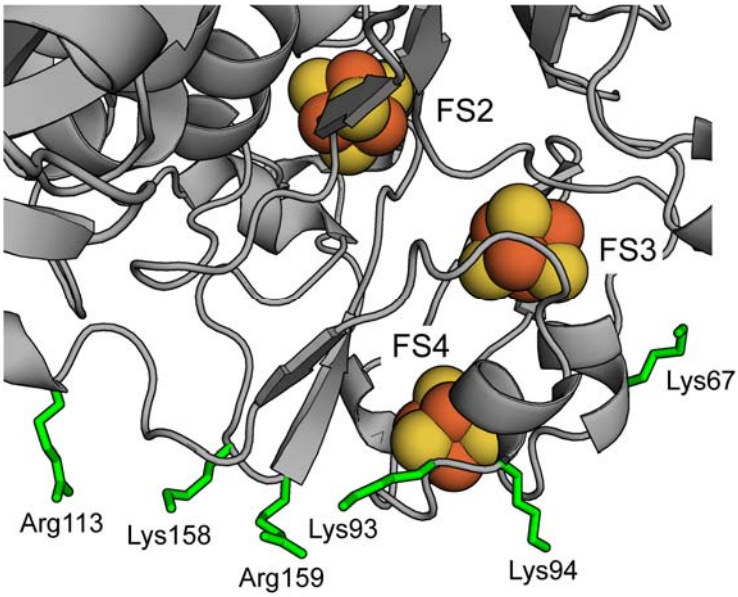


Fig. S5. View of the ridge lined with positively-charged residues near FS4 in ArrB. The protein backbone is shown in gray with the positively-charged side chains shown in green.

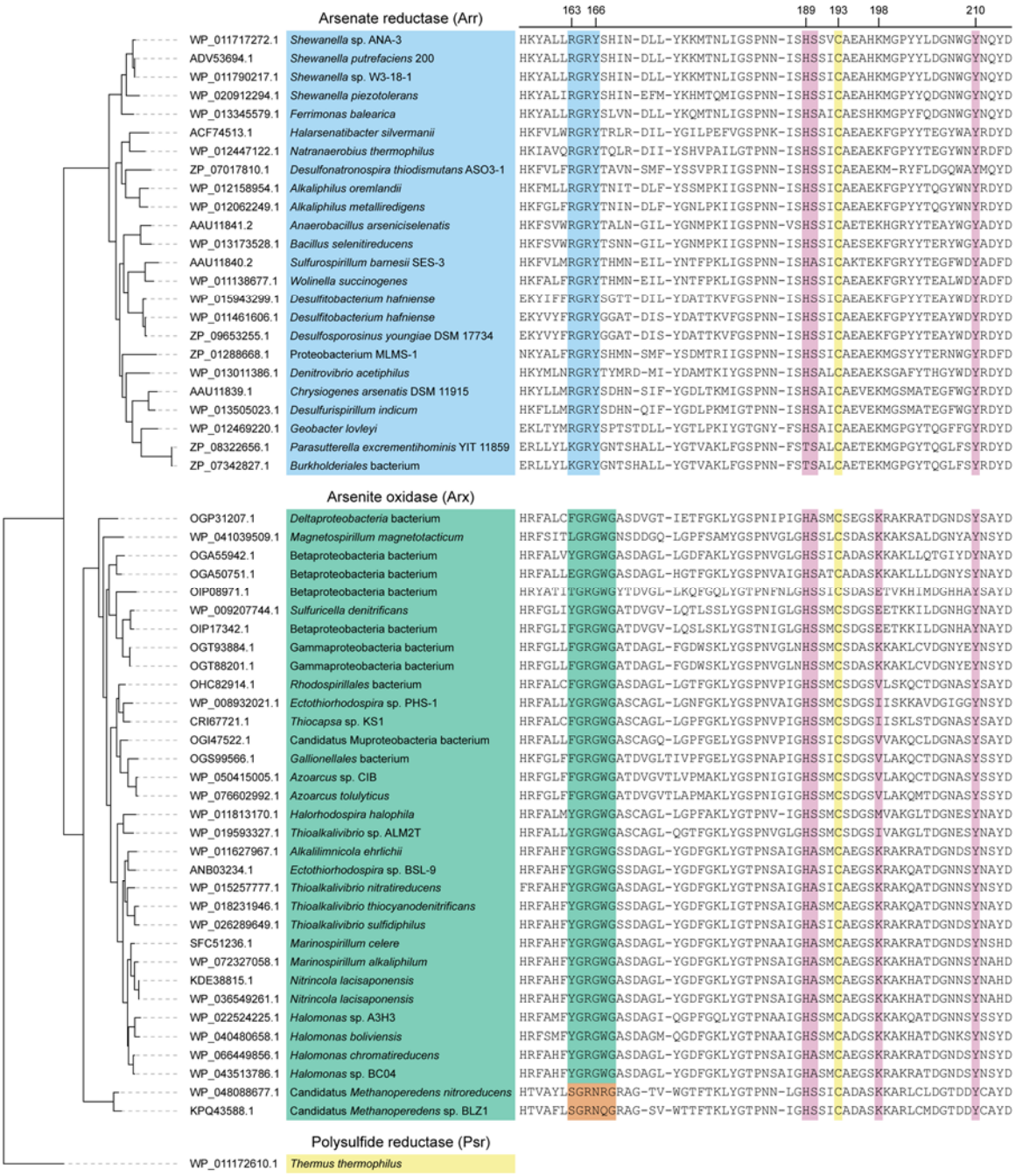


Fig. S6. Alignment of previously analyzed ArrA (17) (blue) and ArxA (18) (green) sequences highlighting the active site residues of ArrA. The sequence of PsrA (16) (yellow) was included as an outgroup. Sequence numbers are relative to the ArrA from *Shewanella* sp. ANA-3. Annotated ArrA enzymes contain a (R/K)GRY motif (blue), while annotated ArxA enzymes contain a XGRGW motif (green). Two sequences annotated as ArxA from *Methanoperedens* (18) deviate from the conserved active-site motif (orange). Other active site residues are generally conserved between ArrA and ArxA (pink), including the Mo-coordinating cysteine (yellow). The sequences were aligned with T-Coffee (19) and the tree was generated using the maximum-likelihood method with RAxML (20). The figure was generated with iTOL (21) and stylized with Adobe Illustrator.

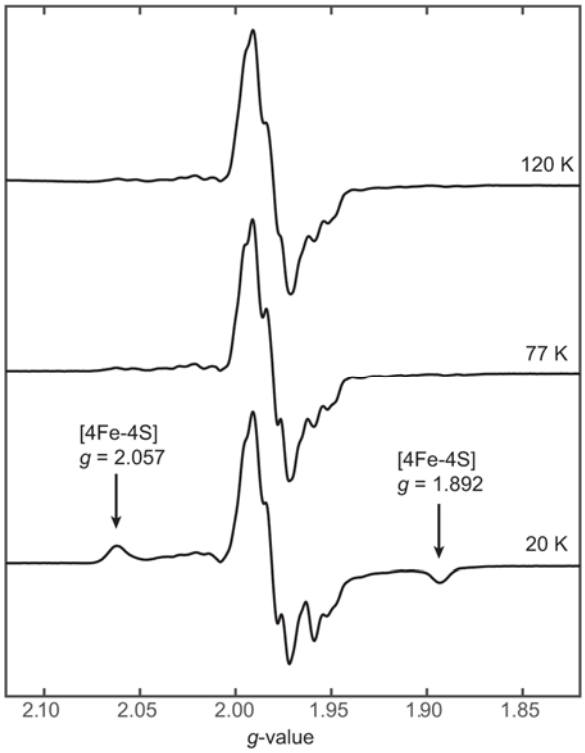


Fig. S7. Temperature dependent X-band CW EPR spectra of Arr. Acquisition parameters: microwave frequency = 9.390 GHz; microwave power = 20 mW (120 K) or 1 mW (77 K, 20 K); modulation amplitude = 0.4 mT.

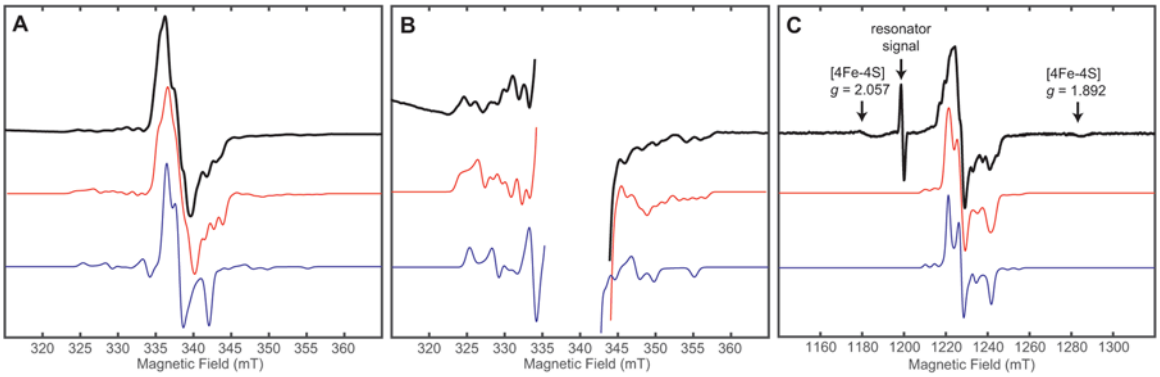


Fig. S8. Comparisons of simulations with and without hyperfine coupling to ^{75}As. Simulation parameters can be found in Table 2 of the main text. (A) X-band CW EPR spectrum collected at 120 K (black trace) with a simulation of the arsenite-bound Mo(V) species (red trace) compared to a simulation excluding coupling to ^{75}As (blue trace). Acquisition parameters: temperature = 120 K; microwave frequency = 9.390 GHz; microwave power = 20 mW; modulation amplitude = 0.4 mT. (B) Enlarged view of same X-band CW EPR spectrum in A (black trace) to show the low-intensity features from the hyperfine interaction with $^{95/97}$Mo, and a simulation of the arsenite-bound Mo(V) species (red trace) compared to a simulation excluding coupling to ^{75}As (blue trace). (C) Pseudomodulated Q-band ESE-EPR spectrum of Arr (black trace) with a simulation of the arsenite-bound Mo(V) species (red trace) compared to simulation excluding coupling to ^{75}As (blue trace).

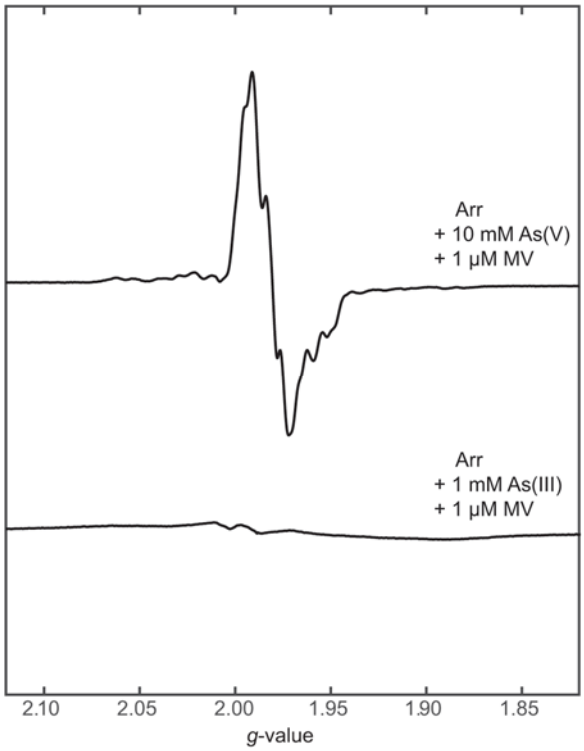


Fig. S9. Comparison of X-band CW EPR spectra of Arr under reaction conditions to Arr incubated with As(III). Acquisition parameters: temperature = 77 K; microwave frequency = 9.390 GHz (reaction sample), 9.393 (As(III) sample); microwave power = 1 mW; modulation amplitude = 0.4 mT.

References

1. Malasarn D, Keeffe JR, Newman DK (2008) Characterization of the arsenate respiratory reductase from *Shewanella* sp. strain ANA-3. *J Bacteriol* 190(1):135–142.
2. Gibson DG, et al. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6(5):343–345.
3. Guzman LM, Belin D, Carson MJ, Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* 177(14):4121–4130.
4. Saltikov CW, Newman DK (2003) Genetic identification of a respiratory arsenate reductase. *PNAS* 100(19):10983–10988.
5. Sheldrick GM (2008) A short history of SHELX. *Acta Cryst A* 64(1):112–122.
6. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(4):658–674.
7. Terwilliger TC, et al. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Cryst D* 64(1):61–69.
8. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Cryst D* 66(4):486–501.
9. Kabsch W (2010) XDS. *Acta Cryst D* 66(2):125–132.
10. Evans P (2006) Scaling and assessment of data quality. *Acta Cryst D* 62(Pt 1):72–82.
11. Evans PR, Murshudov GN (2013) How good are my data and what is the resolution? *Acta Cryst D* 69(Pt 7):1204–1214.
12. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Cryst D* 67(Pt 4):235–242.
13. Watanabe T, Honda K (1982) Measurement of the extinction coefficient of the methyl viologen cation radical and the efficiency of its formation by semiconductor photocatalysis. *J Phys Chem* 86(14):2617–2619.
14. Stoll S, Schweiger A (2006) EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *Journal of Magnetic Resonance* 178(1):42–55.
15. Hyde JS, Pasenkiewicz-Gierula M, Jesmanowicz A, Antholine WE (1990) Pseudo field modulation in EPR spectroscopy. *Appl Magn Reson* 1(3):483.
16. Jormakka M, et al. (2008) Molecular Mechanism of Energy Conservation in Polysulfide Respiration. *Nat Struct Mol Biol* 15:730.
17. van Lis R, Nitschke W, Duval S, Schoepp-Cothenet B (2013) Arsenics as bioenergetic substrates. *Biochim Biophys Acta* 1827(2):176–188.
18. Oremland RS, Saltikov CW, Stolz JF, Hollibaugh JT (2017) Autotrophic microbial arsenotrophy in arsenic-rich soda lakes. *FEMS Microbiol Lett* 364(15). doi:10.1093/femsle/fnx146.
19. Notredame C, Higgins DG, Heringa J (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 302(1):205–217.
20. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312–1313.
21. Letunic I, Bork P (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* 44(W1):W242-245.