

## 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](mailto:dkn@caltech.edu)

### **This PDF file includes:**

Supplementary materials and methods  
Figs. S1 to S10

## 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 AGAGACCATGGGCCATCACCATCACCATCAGACTACGACATCCCGACTACCGAAAA 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<sub>6</sub>-TEV vector. The *arrAB* ORF, without the TAT sequence, was PCR amplified from ANA-3 genomic DNA using primers ATTTTGTTTAACTTTAAGAAGGAGATATACCATGAAGAAAGAGAATCAAGTCAACTT GGG and GTGATGGTGTGATGGTGTGATGGCCAGCGCAATCCCCTCGACAATAGG. The pET15b-His<sub>6</sub>-TEV vector was linearized by PCR amplification with primers GGATCCGGCTGCTAACAAAGC and GCCCTGGAAGTACAGGTTTTTCG, and the *arrAB* fragment was joined using Gibson assembly (2) to create pET15b-His<sub>6</sub>-TEV-*arrAB*. The TAT sequence was then PCR amplified from ANA-3 genomic DNA using primers ATTTTGTTTAACTTTAAGAAGGAGATATACCATGAAGAAAGAGAATCAAGTCAACTT GGG and GTGATGGTGTGATGGTGTGATGGCCAGCGCAATCCCCTCGACAATAGG, the pET15b-His<sub>6</sub>-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<sub>6</sub>-TEV-*arrAB*. For expression in *Shewanella*, this final construct was amplified using primers TTTTTGGGCTAGCGAATTCAGGAGTGTGGAAAATGAAGAAAGAGAATCAAGTCAA CTTG and CGCCAAAACAGCCAAGCTTTTAATAAGCGGTTTTAACACCAAAAAC, the vector pBAD18-kan (3) was linearized using PCR with primers GAATTTCGCTAGCCCAAAAAAACGG and AAGCTTGGCTGTTTTGGCG, and the fragments were joined using Gibson assembly. The completed expression vector was transformed into *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<sub>2</sub>HPO<sub>4</sub>, and 2.2 g/L KH<sub>2</sub>PO<sub>4</sub>) with 4 mL/L glycerol, 5 mM MgSO<sub>4</sub>, 200 μM Na<sub>2</sub>MoO<sub>4</sub>, 0.2 g/L ferric ammonium citrate, and 35 μg/mL kanamycin. At an OD<sub>600</sub> 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  $-80^{\circ}\text{C}$  until purification.

For purification, the cell pellets were thawed at room temperature and placed on ice. All subsequent steps were performed at  $4^{\circ}\text{C}$ . 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,  $\text{MgCl}_2$  and  $\text{CaCl}_2$  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  $50,000\times g$ . 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  $(\text{NH}_4)_2\text{SO}_4$  was slowly added, and after 15 min the precipitated material was removed by centrifugation for 15 min at  $5000\times g$ . To ensure a homogenous protein redox state, arsenate was added to a final concentration of 10 mM. The protein was then passed through a  $0.45\ \mu\text{m}$  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  $(\text{NH}_4)_2\text{SO}_4$  (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  $\mu\text{L}$  aliquots in 200- $\mu\text{L}$  PCR tubes, flash frozen in liquid nitrogen, and stored at  $-80^{\circ}\text{C}$ .

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. S1A). 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  $\mu\text{L}$  of screen solution with 0.2  $\mu\text{L}$  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  $\mu\text{L}$  of protein solution (20 mg/ml in 50 mM HEPES, pH 7.5) with 2  $\mu\text{L}$  of crystallization solution (30% PEG 2000 MME, 300 mM KSCN, 100 mM HEPES, pH 7.5) and covering the drop with 50  $\mu\text{L}$  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 ( $MV^{+}$ ,  $\epsilon_{605} = 13,700 \text{ M}^{-1} \text{ cm}^{-1}$  (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  $MV^{+}$  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  $MV^{+}$  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),  $MV^{+}$ , 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

$$V = V_{\max} \frac{[S]}{K_m + [S]}$$

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  $MV^{+}$ ), and  $K_m$  is the Michaelis constant. The global fit for ping-pong kinetics and competitive inhibition used the model

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

where  $V$  is the observed reaction rate,  $V_{\max}$  is the maximum reaction rate,  $[\text{As}]$  is the arsenate concentration,  $[\text{MV}]$  is the  $\text{MV}^{++}$  concentration,  $K_{m,\text{As}}$  is the Michaelis constant for arsenate,  $K_{m,\text{MV}}$  is the Michaelis constant for  $\text{MV}^{++}$ ,  $[\text{P}]$  is the phosphate concentration, and  $K_{i,\text{P}}$  is the inhibition constant for phosphate. The turnover number  $k_{\text{cat}}$  was determined from  $V_{\max} = k_{\text{cat}} [\text{E}]$ , where  $[\text{E}]$  is the enzyme concentration (0.02 nM). A stoichiometry of 2  $\text{MV}^{++}$  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%  $\text{N}_2$  and 5%  $\text{H}_2$ . A mixture was prepared containing 50 mM HEPES buffer (pH 7.5), 50  $\mu\text{M}$  ArrAB, 10 mM sodium dithionite, 10 mM sodium arsenate, and 0.1 nM to 1  $\mu\text{M}$  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 ( $\pi/2 - \tau - \pi - \text{echo}$ ) 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  $\pi$  pulse length = 32 ns; interpulse delay  $\tau$  = 140 ns; shot repetition time (srt) = 5 ms.

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

In general, the ENDOR spectrum for a given nucleus with spin  $I = 1/2$  ( $^1\text{H}$ ) coupled to the  $S = 1/2$  electron spin exhibits a doublet at frequencies

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

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

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

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

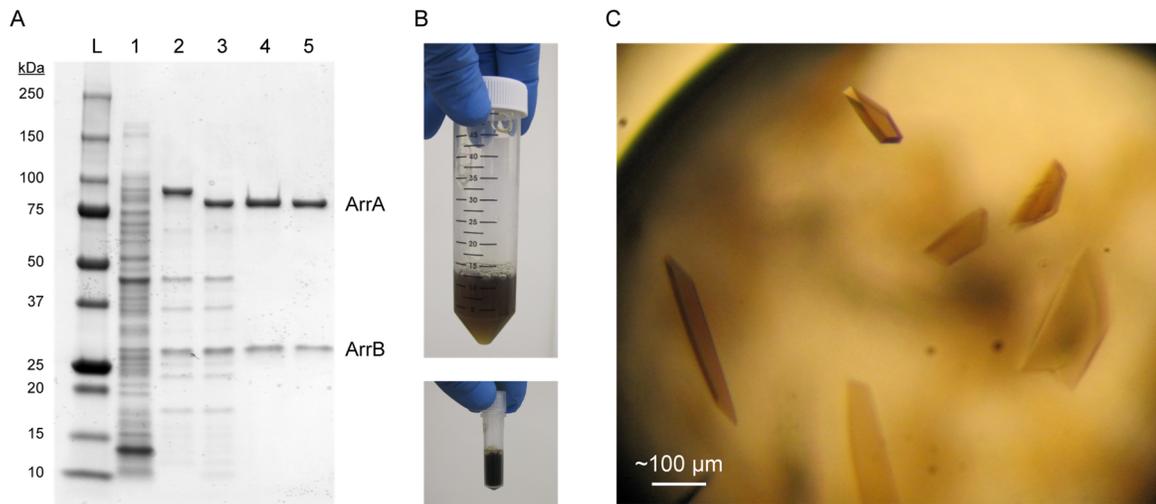
$$\hat{H} = \mu_B \vec{B}_0 g \hat{S} + \mu_N g_N \vec{B}_0 \hat{I} + h \hat{S} \cdot \mathbf{A} \cdot \hat{I} + h \hat{I} \cdot \mathbf{P} \cdot \hat{I} \quad (3)$$

In this expression, the first term corresponds to the electron Zeeman interaction term where  $\mu_B$  is the Bohr magneton,  $g$  is the electron spin  $g$ -value matrix with principle components  $g = [g_{xx} \ g_{yy} \ g_{zz}]$ , and  $\hat{S}$  is the electron spin operator. The second term corresponds to the nuclear Zeeman interaction term where  $\mu_N$  is the nuclear magneton,  $g_N$  is the characteristic nuclear  $g$ -value for each nucleus (e.g.  $^1\text{H}$ ,  $^{75}\text{As}$ ,  $^{95/97}\text{Mo}$ ) and  $\hat{I}$  is the nuclear spin operator. The third term corresponds to the electron-nuclear hyperfine term, where  $\mathbf{A}$  is the hyperfine coupling tensor which can typically be represented as a diagonal matrix with principle components  $\mathbf{A} = [A_{xx} \ A_{yy} \ A_{zz}]$ . 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  $\mathbf{P}$  is the quadrupole coupling tensor. In the principle axis system (PAS),  $\mathbf{P}$  is traceless and parametrized by the quadrupole coupling constant  $e^2 Qq/h$  and the asymmetry parameter  $\eta$  such that:

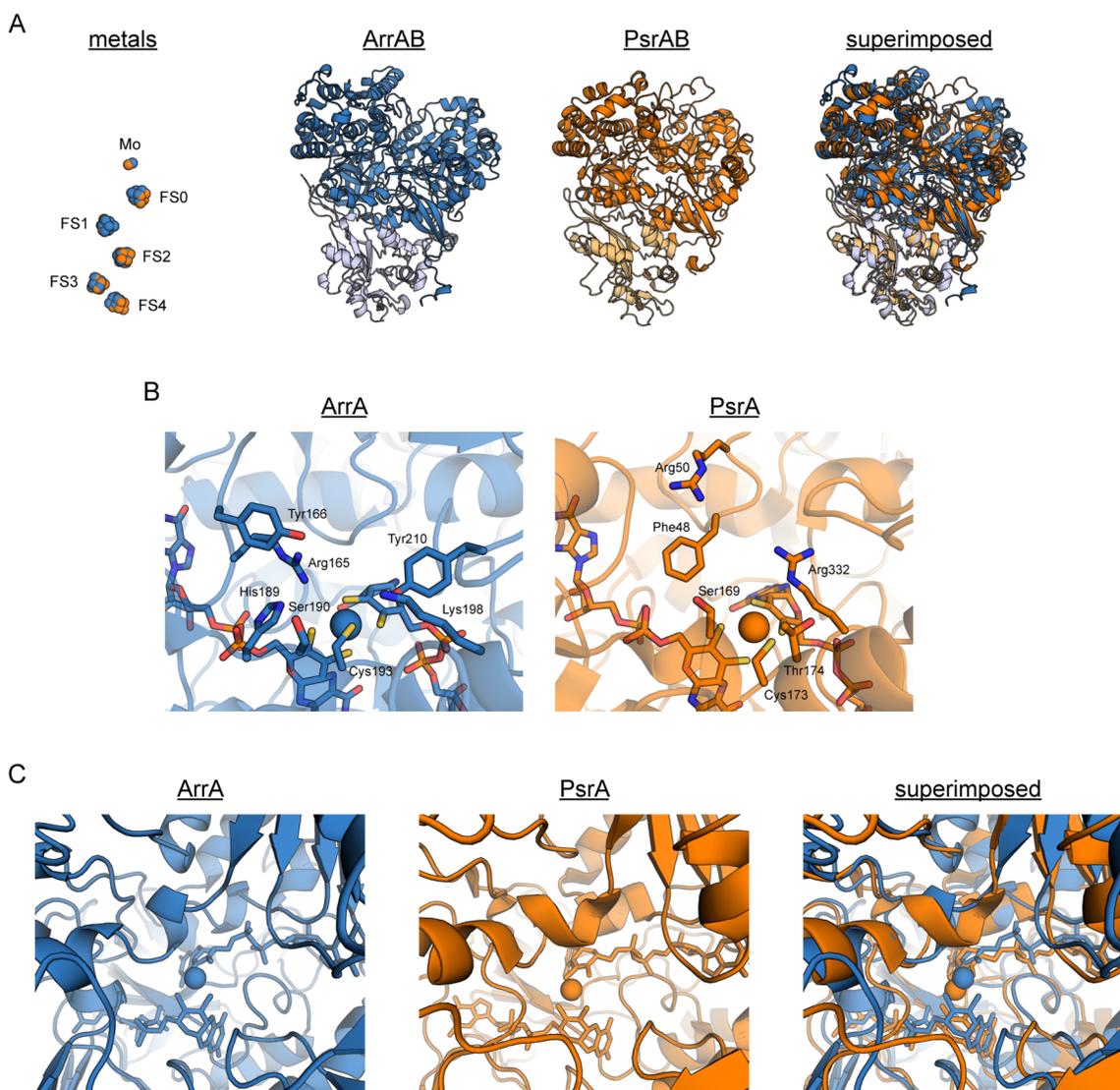
$$\mathbf{P} = \begin{pmatrix} P_{xx} & 0 & 0 \\ 0 & P_{yy} & 0 \\ 0 & 0 & P_{zz} \end{pmatrix} = \frac{e^2 Qq/h}{4I(2I-1)} \begin{pmatrix} -(1-\eta) & 0 & 0 \\ 0 & -(1+\eta) & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad (4)$$

where  $\frac{e^2 Qq}{h} = 2I(2I-1)P_{zz}$  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.

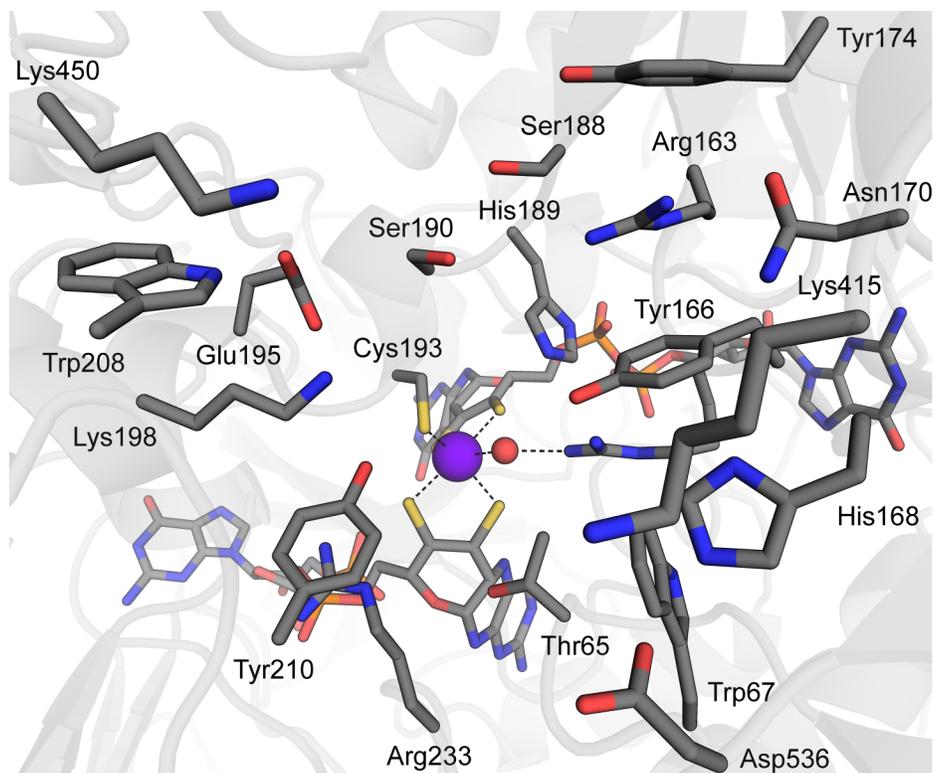
The orientations between the hyperfine and NQI tensor principle axis systems and the  $g$ -matrix reference frame are defined by the Euler rotation angles ( $\alpha$ ,  $\beta$ ,  $\gamma$ ).



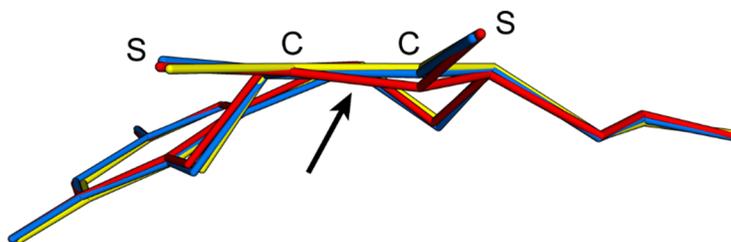
**Fig. S1.** Purification and crystallization of ArrAB from *Shewanella* sp. ANA-3. (A) SDS-polyacrylamide gel stained with Coomassie Blue showing the purification steps of ArrAB. The molecular weights of the ladder bands (L) are shown to the left. The lanes illustrate (1) clarified *Shewanella* lysate 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.



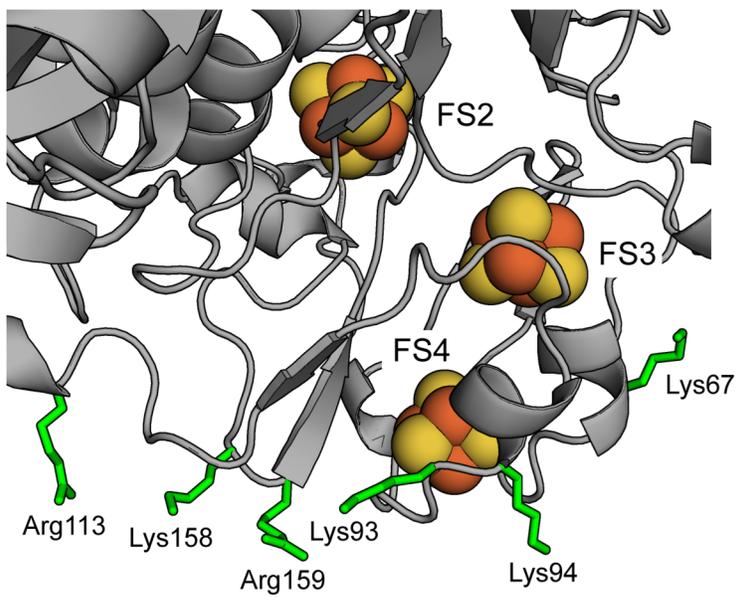
**Fig. S2.** Comparison of ArrAB to its nearest characterized homolog, PsrAB (PDB code 2VPZ) (16). ArrAB is shown in blue and PsrAB is shown in orange. For clarity, the PsrC subunit is omitted because the equivalent subunit for Arr was not determined in this work. (A) Overview of structural similarities 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.



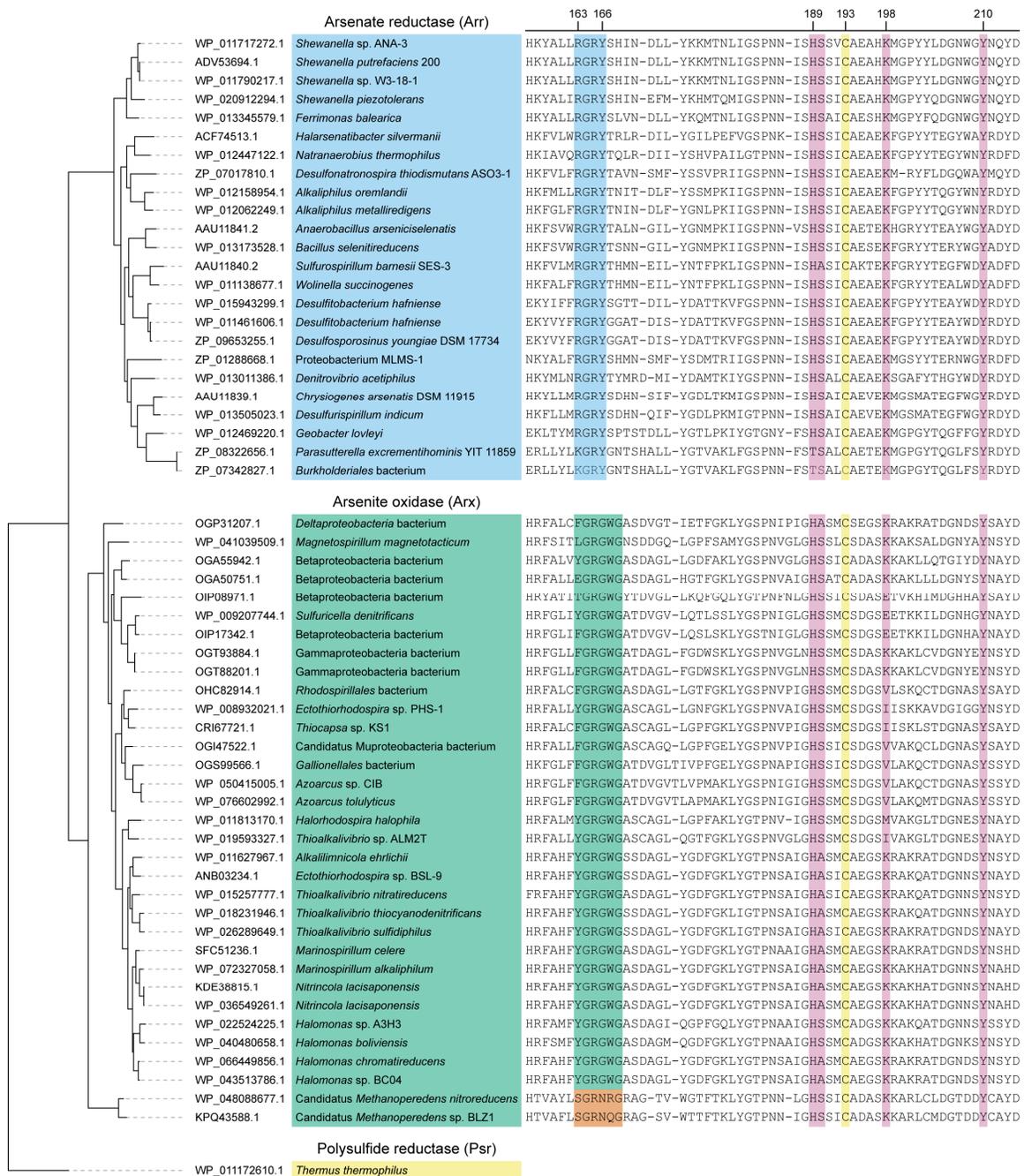
**Fig. S3.** View of residues lining the substrate binding funnel in ArrA. The color scheme is the same as in Fig. 1B (C in gray, N in blue, O in red, S in yellow, and P in orange, Mo in purple).



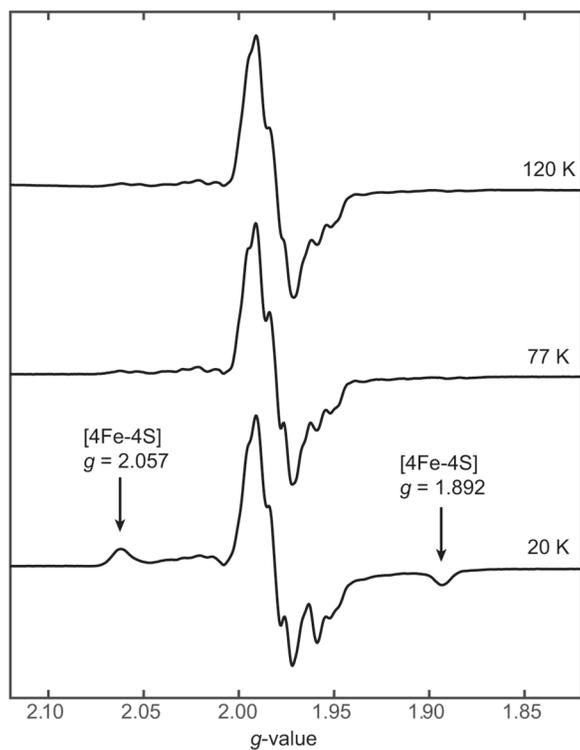
**Fig. S4.** Comparison of dithiolene bond angles in the P-pterin of ArrA in the absence of substrate (blue), in protein bound to arsenate (yellow), and in protein bound to arsenite (red). The arrow indicates the bond in the arsenite-bound structure that appears to be partially reduced.



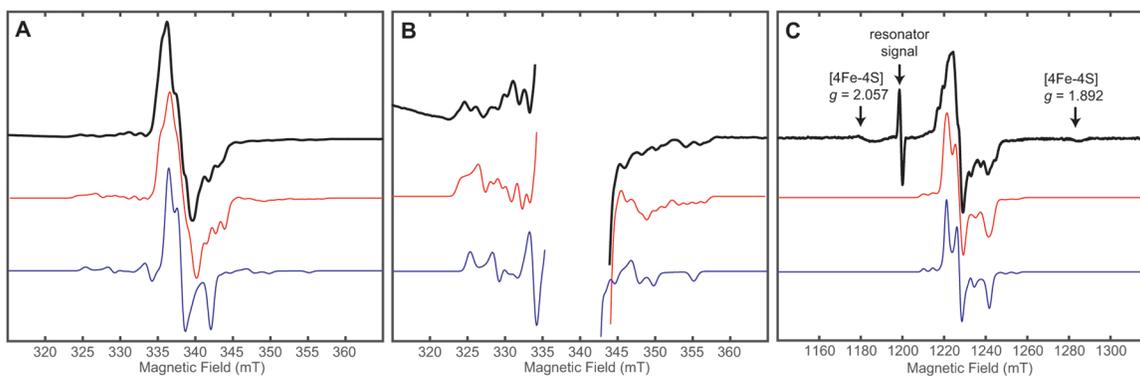
**Fig. S5.** View of the ridge lined with positively-charged residues near FS4 in ArrB. The protein backbone is shown in gray with the positively-charged side chains shown in green.



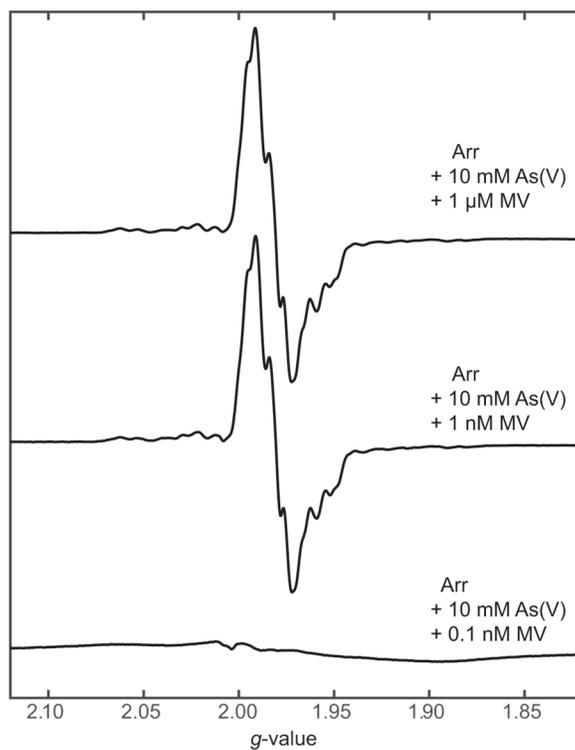
**Fig. S6.** Alignment of previously analyzed ArrA (17) (blue) and ArxA (18) (green) sequences highlighting the active site residues of ArrA. The sequence of PsrA (16) (yellow) was included as an outgroup. Sequence numbers are relative to the ArrA from *Shewanella* sp. ANA-3. Annotated ArrA enzymes contain 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.



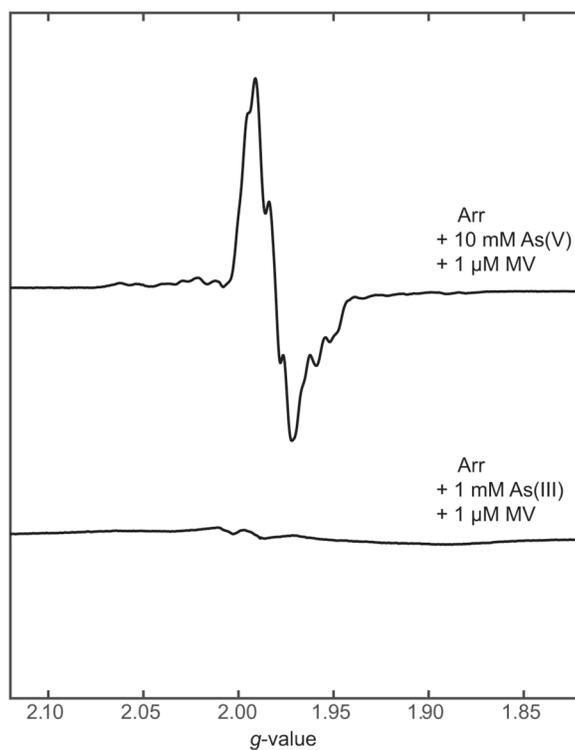
**Fig. S7.** Temperature dependent X-band CW EPR spectra of Arr. Acquisition parameters: microwave frequency = 9.390 GHz; microwave power = 20 mW (120 K) or 1 mW (77 K, 20 K); modulation amplitude = 0.4 mT.



**Fig. S8.** Comparisons of simulations with and without hyperfine coupling to  $^{75}\text{As}$ . 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  $^{75}\text{As}$  (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  $^{95/97}\text{Mo}$ , and a simulation of the arsenite-bound Mo(V) species (red trace) compared to a simulation excluding coupling to  $^{75}\text{As}$  (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  $^{75}\text{As}$  (blue trace).



**Fig. S9.** Comparison of X-band CW EPR spectra of Arr reaction samples with different concentrations of MV. Acquisition parameters: temperature = 77 K; microwave frequency = 9.390 GHz; microwave power = 1 mW; modulation amplitude = 0.4 mT.



**Fig. S10.** 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, Keefe 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.