Title

Identification of an archaeal mercury regulon by chromatin immunoprecipitation

Journal

Microbiology, 161(12)

ISSN

1350-0872

Authors

Rudrappa, Deepak
Yao, Andrew I
White, Derrick
et al.

Publication Date

2015-12-01

DOI

10.1099/mic.0.000189

Peer reviewed

Microbiology (2015), **161**, 2423–2433 DOI 10.1099/mic.0.000189

Identification of an archaeal mercury regulon by chromatin immunoprecipitation

Deepak Rudrappa,[1] Andrew I. Yao,[3] Derrick White,[1] Benjamin J. Pavlik,[2] Raghuveer Singh,[1] Marc T. Facciotti[3] and Paul Blum[1]

Correspondence
Paul Blum
pblum1@unl.edu

[1]School of Biological Sciences, University of Nebraska-Lincoln, Lincoln, Nebraska, USA

[2]Department of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, Nebraska, USA

[3]Department of Biomedical Engineering and Genome Center, University of California-Davis, Davis, California, USA

Mercury is a heavy metal and toxic to all forms of life. Metal exposure can invoke a response to improve survival. In archaea, several components of a mercury response system have been identified, but it is not known whether metal transport is a member of this system. To identify such missing components, a peptide-tagged MerR transcription factor was used to localize enriched chromosome regions by chromosome immunoprecipitation combined with DNA sequence analysis. Such regions could serve as secondary regulatory binding sites to control the expression of additional genes associated with mercury detoxification. Among the 31 highly enriched loci, a subset of five was pursued as potential candidates based on their current annotations. Quantitative reverse transcription-PCR analysis of these regions with and without mercury treatment in WT and mutant strains lacking *merR* indicated significant regulatory responses under these conditions. Of these, a Family 5 extracellular solute-binding protein and the MarR transcription factor shown previously to control responses to oxidation were most strongly affected. Inactivation of the solute-binding protein by gene disruption increased the resistance of mutant cells to mercury challenge. Inductively coupled plasma-MS analysis of the mutant cell line following metal challenge indicated there was less intracellular mercury compared with the isogenic WT strain. Together, these regulated genes comprise new members of the archaeal MerR regulon and reveal a cascade of transcriptional control not previously demonstrated in this model organism.

Received 28 January 2015
Revised 20 July 2015
Accepted 23 September 2015

INTRODUCTION

Microbes that inhabit naturally occurring metal-rich niches provide an opportunity to identify and study novel metal resistance mechanisms. The order *Sulfolobales* comprises diverse thermoacidophilic microbes, including species that inhabit hot metal-saturated locations (Simbahan *et al.*, 2005). Studies on mercury resistance in *Sulfolobus solfataricus* established the existence and critical regulatory features of an archaeal mercury resistance (*mer*) operon (Dixit *et al.*, 2004; Schelert *et al.*, 2004, 2006). The *S. solfataricus mer* locus encodes four genes where *merH*, *merA* and *merI* are arranged in one transcription unit, and *merR* is divergently transcribed upstream of *merH* (Fig. 1). Protein phylogenetic analysis and gene disruption studies indicated that *merA* encodes a mercuric reductase required for reduction of the mercuric ion Hg(II) to its elemental Hg(0) form (Schelert *et al.*, 2004) despite the lack of an active tyrosine residue (Simbahan *et al.*, 2005) in the putative active site. *merI* (122 aa) is located 3′ to *merA* and is separated by a 142 bp intergenic region. It is transcribed constitutively from its own promoter (*merIp*) and by read-through transcription initiating upstream at *merHp*. However, gene disruption studies have so far excluded a role for MerI in mercury resistance or *mer* regulation (Schelert *et al.*, 2006). The *merH* gene was first identified because of its Cys-Xaa$_{19–22}$-Cys-Xaa$_3$-Cys (CxCxC) motif. Together with its location immediately adjacent to mercuric reductase (MerA), *merH* was implicated as playing a role in cytoplasmic trafficking of mercury (Ettema *et al.*, 2003). This role was subsequently confirmed through

Abbreviations: ABC, ATP-binding cassette; ChIP, chromatin immunoprecipitation; HA, haemagglutinin; ICP, inductively coupled plasma; q, quantitative; RT, reverse transcription; WCE, whole-cell extract.

ChIP-Seq data have been deposited at the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/projects/geo/index.cgi), accession number 476 GSE60461.

Four supplementary tables are available with the online Supplementary Material.

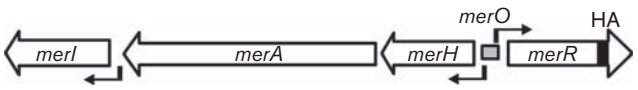

Fig. 1. Components of the *S. solfataricus mer* operon and MerR epitope tagging. Arrows above and below the ORFs indicate transcriptional start sites. The grey box indicates the *merR* binding site (*merO*). A plasmid carrying a modified *merR* sequence with a C-terminal HA epitope sequence (black box) was flanked by sequences identical to the ends of the *merR* region. The plasmid sequence was integrated into the chromosome by markerless exchange to replace the natural homologous region.

experimental studies (Schelert *et al.*, 2013). Finally, the *S. solfataricus* MerR transcription factor regulates *merHAI* transcription in a metal-dependent fashion. A *merR* disruption mutant exhibited increased Hg(II) resistance and constitutive synthesis of the *merA* transcript. This indicates that archaeal MerR negatively regulates *merA* transcription (Schelert *et al.*, 2004). Site-specific mutations in the DNA-binding site of MerR created *in vivo* positioned the binding site immediately 5′ of the predicted *merHp* TATA box (Schelert *et al.*, 2006). Electrophoretic mobility shift assays demonstrated that MerR/*merHp* DNA complex formation was template-specific and dependent on the presence of the binding site, but that binding was insensitive to Hg(II) addition as well as site-specific binding site mutations that relieved *in vivo merHp* repression (Schelert *et al.*, 2006).

In bacteria, genes for mercuric ion transport and cleavage of the carbon–mercury bond are often clustered with those for mercuric ion reduction and transcriptional regulatory control (Wilson *et al.*, 2000). For example, MerT and/or MerC facilitate Hg(II) transport across the plasma membrane. MerC protein possesses four putative α-helical transmembrane (spanning) domains that have been shown to bind and promote influx of Hg(II) ions (Sahlman *et al.*, 1997). MerP is a periplasmic Hg(II)-binding protein synthesized with a cleavable N-terminal leader. Once bound to the dithiol on either the first transmembrane helix of MerT or MerC, which are predicted to be accessible to the periplasmic environment, Hg(II) is directed into the cytoplasm by an unknown mechanism (Sone *et al.*, 2013). The MerF protein catalyses the uptake of Hg(II) for subsequent reduction by mercuric reductase (Barkay *et al.*, 2003). Although there have been no experiments that directly show MerP can donate Hg(II) to either transporter, it is unlikely that MerP only binds Hg(II) simply to protect the periplasmic space (Sone *et al.*, 2013). Despite progress in the understanding of bacterial mercury import, it remains unknown if or how archaea catalyse mercury uptake to reduce extracellular concentrations. Here, efforts using chromosome immunoprecipitation combined with DNA sequence analysis (ChIP-Seq) were used to elucidate such genes that might encode analogous functions. Using a genetics approach, these results led to the identification of the mercury regulon in archaea, including members with transport activity.

METHODS AND METHODS

Archaeal strains and cultivation. Archaeal strains, plasmids and primers are indicated (Table S1, available in the online Supplementary Material). *S. solfataricus* strain PBL2025 and its derivatives were grown at 80 °C with aeration in batch culture as described previously (Allen, 1959; Rolfsmeier & Blum, 1995; Worthington *et al.*, 2003) in Allen's basal salts (Allen, 1959) as modified by Brock *et al.* (1972) at a pH of 3.0. Liquid media were supplemented with either 0.2 % lactose (w/v), 0.2 % sucrose (w/v) or 0.2 % (w/v) tryptone as carbon and energy sources as indicated. Growth was monitored at OD_{540} using a Cary 50 spectrophotometer (Varian). When investigating the effect of mercuric ion, cells were treated with mercuric chloride (Sigma) from a freshly made 10 mM stock.

Strain constructions. Strain construction procedures were as described previously (Maezato *et al.*, 2011; Sowers *et al.*, 2007), unless otherwise noted. DNA was electroporated into strain PBL2025 (Table S1) and its derivatives. Recombinants were enriched and individuals isolated, screened and processed as described (Maezato *et al.*, 2011; Sowers *et al.*, 2007). PCR, restriction analysis and DNA sequencing were used to genotype alleles. Overlap extension PCR (Higuchi *et al.*, 1988) was used to create site-specific mutations and DNA fusions. The *merR*-HA (haemagglutinin epitope-tagged *merR*) mutant (PBL2138) was constructed by markerless exchange as described previously (Schelert *et al.*, 2006) using pBN1035 (Fig. 1). PCR was used to generate fragments identical in sequence to both ends of the *merR* region targeted for deletion, using primer p1 combined with p2 and primer p3 combined with p4 (Table S1). These amplicons were annealed at their overlapping region and amplified to produce a single amplicon. The *merR*-HA product was then cloned at the *SalI*/*Sph*I site of pBN1035 to generate pBN1291, sequenced to verify its composition and integrated into the chromosome by markerless exchange. The insertion of the HA epitope at the C-terminal end of the chromosomally encoded *merR* gene was verified by PCR and DNA sequencing in all strains. The *Ssol_0429*::*lacS* disruption strain was constructed by inactivating the *Ssol_0429* gene through insertion of the *lacS* gene using pBL1313 to generate pBL1310 (Table S1). This construct was sequenced to verify its composition and integrated into the chromosome as described previously (Maezato *et al.*, 2011) to generate PBL2139. The insertion of the disrupted copy of *Ssol_0429* was verified by PCR and DNA sequencing (Table S1).

Sample growth and treatment. For each ChIP, samples were prepared from 50 ml cultures. When challenged with mercuric chloride [Hg(II)], samples were prepared from 250 ml cultures. When there was no mercury challenge, cells were harvested during mid-exponential phase at OD_{540} 0.5. For mercury treatment, 0.5 μM Hg(II) was added at OD_{540} 0.1, which corresponds to 10^8 cells ml^{-1}, and cells harvested after 1 h. Cells were centrifuged and immediately fixed using 1 % (v/v) formaldehyde for 10 min. Fixation was terminated through the addition of glycine to a final concentration of 125 mM. Cell pellets were washed twice with 1× PBS, pH 7.0, and then frozen at −80 °C until further use.

Mercury resistance and analysis of intracellular concentrations. Strains were grown with aeration in defined sucrose minimal medium. At a cell density of 10^8 cells ml^{-1}, 0.5 μM mercuric chloride was added to each culture. Cultures of strains with no added mercury were included as controls. Growth was monitored by measuring the OD_{540} and growth curves were plotted. To determine the intracellular concentration of mercury, inductively coupled plasma (ICP)-MS analysis was carried out as described previously (Schelert *et al.*, 2013). Mercury-treated cells corresponding to $\sim 10^8$ cells ml^{-1} were then removed at the indicated times and cells were harvested by centrifugation at 10 000 ***g*** for 5 min followed by two successive washes using distilled water to remove free mercury. Cell pellets were frozen at −20 °C for subsequent analysis by ICP-MS.

Cell pellets were extracted using 50 µl concentrated nitric acid and the resulting extracts were analysed by ICP-MS using an Agilent ICP-MS 7500cx instrument. A certified mercury reference standard was used for sample normalization. All samples were run as triplicates and the values are indicated as mean values.

Immunoprecipitation, Western blotting and ChIP-Seq. Cell pellets were resuspended in 1.6 ml of lysis buffer [50 mM HEPES, 140 mM NaCl, 1 mM EDTA, 1 % (v/v) Triton X-100, 0.1 % (w/v) sodium deoxycholate, pH 7.5] containing protease inhibitors (Roche). Resuspended pellets were sonicated using a Bioruptor (Diagenode) until DNA fragment size reached a mean value of ~500 bp (two 7.5 min cycles, 30 s on/30 s off, high power setting followed by an additional 5 min cycle 30 s on/30 s off, high power setting). The cell lysate was combined with 1 µg anti-HA antibody (Abcam) and 50 µl Protein A-conjugated magnetic beads (JSR Micro) blocked with 5 mg BSA ml^{-1} in PBS and incubated overnight at 4 °C. Magnosphere beads were washed twice with the lysis buffer, twice with 1 ml lysis buffer supplemented with 500 mM NaCl, twice with 1 ml wash buffer [10 mM Tris, 250 mM LiCl, 0.5 % NP-40 (v/v), 0.5 % sodium deoxycholate (w/v), 1 mM EDTA, pH 8.0] and once with 1 ml Tris-EDTA (TE) buffer. Enriched ChIP DNA/transcription factor complexes were eluted by the addition of 50 µl elution buffer [50 mM Tris, 10 mM EDTA, 1 % SDS (w/v), pH 8.0] and incubation at 65 °C for 10 min. Cross-links were reversed by incubating in TE/SDS (10 mM Tris, 1 mM EDTA, 1 % SDS) overnight at 65 °C. RNA was digested using RNase A (Fermentas) at 37 °C for 1 h and a DNA sample was subsequently prepared for Illumina single-read sequencing.

Individual ChIP samples were blunt-ended with T4 DNA polymerase (NEB), DNA polymerase I, Large (Klenow) Fragment (NEB) and T4 polynucleotide kinase (NEB) at 20 °C for 30 min. Blunt-ended DNA was 3′ A-tailed with 3′→5′ exo-Klenow fragment (NEB) for 30 min at 37 °C. Adapters containing 6 bp barcodes were ligated to the prepared ChIP DNA samples for 15 min at room temperature with T4 DNA ligase (Enzymics). Barcode oligonucleotide sequences are provided in Table S1. A background control of whole-cell extract (WCE) genomic DNA from each sample was prepared as described above. ChIP DNA and whole-cell extract DNA samples were then used as template for an 18-cycle PCR amplification using Illumina single-read-compatible prAc and prBc primers (Table S1). PCR-amplified products were quantified and visualized with a high-sensitivity DNA chip (Agilent) on a Bioanalyzer (Agilent). ChIP and background libraries were pooled in equimolar concentrations.

For Western blot, PBL2138 cell pellets were immunoprecipitated using the same ChIP methods described above. Immunoprecipitated samples were run on a 4–12 %, 1.5 mm polyacrylamide gel (Invitrogen) in MOPS buffer (Invitrogen). Protein was then transferred onto a 0.2 nm-pore size PVDF membrane (Invitrogen) at 30 V for 1 h in transfer buffer [25 mM Tris, 192 mM glycine, 10 % (v/v) methanol, pH 8.4]. The PVDF membrane was blocked in 0.5 % (w/v) casein overnight and subsequently probed with horseradish peroxidase-conjugated anti-HA antibody (Abcam). The blot was incubated with GE ECL Plus reagents (Amersham) according to the manufacturer's suggestions and exposed to light-sensitive film.

Quantitative (q) PCR was performed on a Bio-Rad Chromo 4 Real-Time Detector (Bio-Rad) using KAPA SYBR FAST Universal 2 × qPCR master mix (Kapa Biosystems) with 5 µM of both forward and reverse primers according to the supplied protocol [95 °C for 5 min, (95 °C for 30 s, 60 °C for 45 s) × 50 cycles]. Primer sets for enriched regions and negative regions were designed using a known MerR-binding site and sites unexpected to undergo MerR binding (Table S1). Prior to qPCR, primers were tested for specificity/single product formation using *Taq* polymerase (Qiagen), 10 µM each primer and a concentration equivalent to the template used for qPCR of *S. solfataricus* 98/2 WCE [98 °C for 10 min (98 °C for 45 s, 55 °C for 1 min, 72 °C for 30 s) × 26 cycles, 72 °C for 5 min]. Fold enrichment above background was calculated as 2 to the power of cycle threshold difference between a non-enriched region and the expected enriched site. WCE extract, ChIP samples and amplified libraries were all used as template for a qPCR.

Sequencing, analysis, peak identification and characterization. ChIP and WCE DNA libraries were multiplexed and sequenced on the Illumina MiSeq Platform yielding 7.2 million 75 bp reads. Individual samples were reconstituted by sorting each read by barcode and each sample was quality trimmed (minimum Phred quality, 20; minimum length, 25 bp) using the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). Sequencing primer and adaptor contamination were filtered using the TagDust package (Lassmann *et al.*, 2009). A reference genome for *S. solfataricus* 98/2 (GenBank accession number CP001800.1) was created using Bowtie (Langmead *et al.*, 2009). For each sample, quality-filtered reads were mapped using Bowtie (Langmead *et al.*, 2009) to the reference genome, and SAM format sequence files were converted to sorted BAM files using the SAMtools package (Li *et al.*, 2009) for peak picking and visualization.

Putative protein–DNA binding events were detected using Pique, a microbe-focused and freely available peak calling application (available at https://github.com/ryneches/pique, version tag: halo_egw; R. Y. Neches and others, unpublished). ChIP-Seq coverage data and candidate peaks were visualized and manually curated using the Gaggle Genome Browser (Bare *et al.*, 2010).

High-confidence peaks were selected manually based on peak size and peak shape as described previously (Wilbanks *et al.*, 2012). High-confidence peaks were associated with National Center for

Table 1. Identification of *merR*-like target binding sites by pairwise alignment and relationship to transcriptional start sites

ORF	Annotation	Sequence/motif	Integrated peak Hg−/Hg+	Peak ratio Hg−/Hg+	*P*-value
0502	*merR*	TCTTTCTATGAACATATGTTCATATGAAAGA	57.6//4.4	13.1	Schelert *et al.* (2006)
0429	Family 5 extracellular solute-binding protein	TTATAGTAATAGCCTATGAATGGGTTATAT	7.0//2.1	3.3	1.04×10^{-5}
1340	Multiple antibiotic resistance *marR*	AATAGCTGATGCGATAATGAATAGCGAAAT	12.7//2.7	4.7	5.79×10^{-5}
0152	Transport protein	TAAGGAGCGGGGATTATGAATACTCGTTCA	1.2//1.8	0.7	NA
0648	AAA ATPase	TCAAAGTAAGCCCAGCCCAATTTATTCGCTA	0.7//1.1	0.6	NA

NA, Not applicable.

Biotechnology Information-predicted ORFs using the Gaggle Genome Browser. Peaks were quantified by integrating the number of reads within each peak window. To account for the variation in mapped reads within a ChIP-Seq sample and differences in sequencing efficiency between ChIP-Seq samples, local enrichment ratios were calculated for each peak by taking the ratio of integrated reads within a peak window (w) over the integrated reads from two w/2-size windows spaced ~100 bp from the original peak window. To compare the quantified ChIP-Seq peaks between the sets without mercury (Hg−) and with mercury (Hg+), ratios between Hg− and Hg+ local enrichment ratios were calculated.

Identification of *merR* binding signature. Putative *merR* binding sequences were obtained by analysing 150 bp centred on each Pique-predicted peak centre. These sequences were input into MotifCatcher (Seitzer *et al.*, 2012) running MEME (Bailey & Elkan, 1994) with the following parameters: minimum motif width, 9; maximum motif width, 20; reverse complement, no; random seeds, 100; seed size, 4 (all other parameters were unchanged from the default settings). MotifCatcher recovered two independent putative binding motifs that were each subsequently scanned against the *S. solfataricus* 98/2 genome using MAST (Bailey & Gribskov, 1998) and visualized in the Gaggle Genome Browser (Bare *et al.*, 2010). For motif analyses a third-order background file was generated for the *S. solfataricus* 98/2 genome using the FASTA-get-markov script (Bailey & Elkan, 1994). As one of the motifs contained a strong ATG element, the potential overlap between motif position and putative ATG start codons was examined. Gene annotations from the National Center for Biotechnology Information were used to select 20 bp windows around predicted 'ATG' start codons. ATG-containing motifs (12 bp) that were found to reside completely within these 20 bp windows were considered to be associated with predicted 'ATG' start codons. All data parsing and sequence analysis were performed in R (Team, 2008).

Quantitative reverse transcription (qRT)-PCR analysis. Testing for functional regulatory associations between selected putative MerR-HA binding sites identified by ChIP-Seq and their candidate target genes was accomplished by qRT-PCR (Table 1) as described (Maezato *et al.*, 2011). Total cellular RNA was prepared as described previously (Bini *et al.*, 2002). All samples were obtained from cells in the exponential growth phase. Parallel RT-PCR amplifications were used to evaluate RNA levels from target genes relative to those of the reference gene TFBI (*Ssol_1927*), as described previously (Maezato *et al.*, 2011). RNA was treated to remove DNA by the addition of 1 U DNase I (Fermentas) (μg total RNA)$^{-1}$ at room temperature for 15 min and then neutralized with 2 ml 25 mM EDTA and incubated at 70 °C for 10 min. cDNA synthesis used 20 pmol PCR antisense primer, 20 mM dNTPs mix (Fermentas) and 200 U Moloney murine leukemia virus reverse transcriptase (NEB) for 60 min at 37 °C.

RESULTS

Construction and validation of transcription factor epitope-tagged cell lines

S. solfataricus strain PBL2138 was constructed with an integrated copy of the MerR transcription factor fused at its C-terminal end to a HA-tag (MerR-HA) that replaced the native *merR* allele (Fig. 1). Transcription of the recombinant *merR* was under its native promoter, *merRp* (Schelert *et al.*, 2006). Selection of the peptide tag was based on prior studies demonstrating the compatibility of the HA epitope with formaldehyde-based ChIP (Wilbanks *et al.*, 2012). Strong mercury-induced expression of mercury resistance in PBL2138 showed that the peptide-tagged *merR* allele did not compromise MerR function as a mercury-responsive transcription factor.

Immunoprecipitation and ChIP-Seq analysis

To verify that immunoprecipitations were purifying MerR-HA and associated DNA, a fraction of the immunoprecipitated material was analysed by Western blot. Immunoprecipitated MerR-HA from formaldehyde cross-linked and formaldehyde-free cultures of the PBL2138 strain were compared (Fig. 2a). A single band migrating at the molecular mass corresponding to the MerR-HA monomer was evident in the formaldehyde-free control. This band faded slightly and a broad band of high molecular mass appeared in the formaldehyde cross-linked sample, indicating that MerR-HA and associated DNA were being co-precipitated. DNA from the remainder of the immunoprecipitated sample (that which was not used for the Western blot) was purified and used for qPCR analysis to verify that the known MerR-binding site (*merHp*) was enriched in the MerR-HA immunoprecipitations (Fig. 2b). Two sets of primers that amplified a product in the region of *merHAp* served as a test for targeted enrichment against two sets of negative control reactions that amplified regions of putative *amyA* and *malA* promoters (*amyAp* and *malAp*, regions known not to bind MerR). qPCR showed 470.35 ± 29.21-fold enrichment of *merHp* over *amyAp* and *malAp* in cross-linked samples, and 1.12 ± 0.17-fold enrichment in WT controls, indicating that DNA accompanying MerR-HA in the immunoprecipitations was enriched for MerR-binding sites. ChIP samples were also prepared for high-throughput sequencing on the Illumina MiSeq platform. A summary of the resulting sequence data is reported in Table S2.

ChIP-Seq data were processed as described in Methods, and a manually curated list of putative peaks was called for both Hg+ and Hg− mercury samples, which had eight and 186 putative peaks, respectively. As expected, this peak list included one very large peak located at the canonical *merAp* site (Fig. 3a). Whilst many of other identified peaks likely represented real binding events captured by the experiment (Fig. 3b, c), it remains challenging, without further experimental data, to determine which of these events represent functionally relevant protein–DNA associations and which might be functionally spurious. To limit the subsequent investigation, a list of 31 high-confidence peaks was created by filtering for peaks with a maximum peak height >500 reads/base (seven times the sequencing depth) (Table S3). Filtered peaks were confirmed visually in the Gaggle Genome Browser.

Discovery of a MerR-binding motif

In order to assess the nature of MerR target sites across the *S. solfataricus* 98/2 genome, sequences associated with the curated ChIP-Seq peaks were analysed using MotifCatcher,

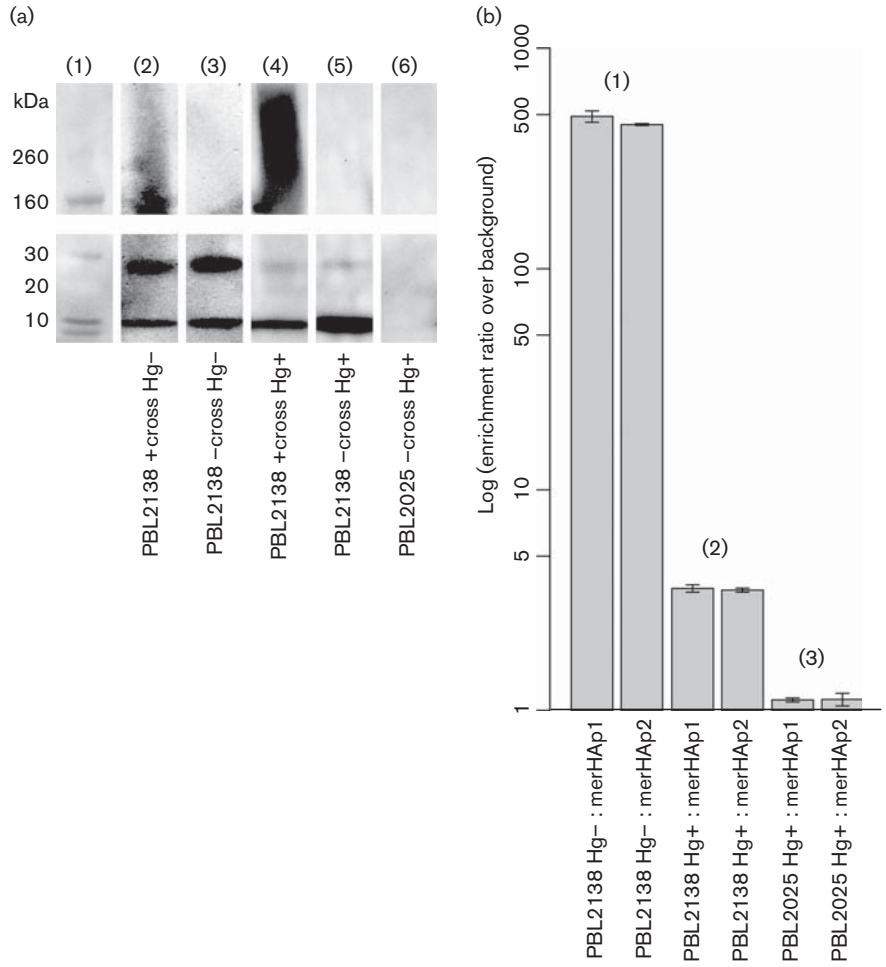

Fig. 2. MerR immunoprecipitation and target site enrichment. (a) Western blot showing using immunoprecipitates from strain PBL2138 with (Hg+) and without (Hg−) mercury addition and in the presence and absence of formaldehyde. Immunoprecipitation used an antibody directed against the HA-tag attached to MerR. Lane 1, protein standards; lane 2, PBL2138 immunoprecipitation in the absence of mercury and cross-linked with formaldehyde; lane 3, PBL2138 immunoprecipitation in the absence of mercury and not cross-linked with formaldehyde; lane 4, PBL2138 immunoprecipitation in the presence of mercury and cross-linked with formaldehyde; lane 5, PBL2138 immunoprecipitation in the presence of mercury and not cross-linked with formaldehyde; lane 6, WT (not HA-tagged) control sample showing that bands in PBL2138 strains correspond to MerR-HA and not a contaminant. (b) qPCR of immunoprecipitated DNA showing ChIP fold enrichment on a logarithmic scale of the *merHAp* site relative to the mean of signal at negative control sites *amyAp* and *malAp*. For each experimental condition two independent primer sets targeting different regions within the *merHAp* site were used to assess enrichment. Three sets of experimental samples are shown: (from left to right) (1) PBL2138 immunoprecipitated DNA from cells grown in the absence of mercury and cross-linked with formaldehyde, (2) PBL2138 immunoprecipitated DNA from cells grown in the presence of mercury and cross-linked with formaldehyde, and (3) PBL2025 (WT, not HA-tagged) immunoprecipitated DNA from cells grown in the presence of mercury and cross-linked with formaldehyde. Data represent mean ± SD of technical triplicate reactions.

a Monte Carlo-based motif searching algorithm (Seitzer *et al.*, 2012). The informatics analysis identified the motif WDKRGMGMAHAWGAAT (MC-Motif 1, Fig. 4b) and the discovery of a less prevalent alternate motif, MRWW-RVTGRRG (MC-Motif 2, Fig. 4b). Additional cycles using MotifCatcher using different subsets of the peak list resulted in repeated identification of MC-Motif 1. These two motifs were mapped to the *S. solfataricus* 98/2 genome using MAST (Bailey & Gribskov, 1998), resulting in 1137 occurrences ($P \leq 9.63 \times 10^{-5}$) of MC-Motif 1 and 1325 occurrences ($P \leq 9.19 \times 10^{-5}$) of MC-Motif 2. The discovery of the strong 'ATG' element in the consensus motif seems to be independent of 'ATG' start codons, as only 1.39 % (29/2091) of the ATG-rich motif sites co-occur within 10 bp of predicted 'ATG' start codons. MC-Motif 1 appears to contain elements similar to the first half site

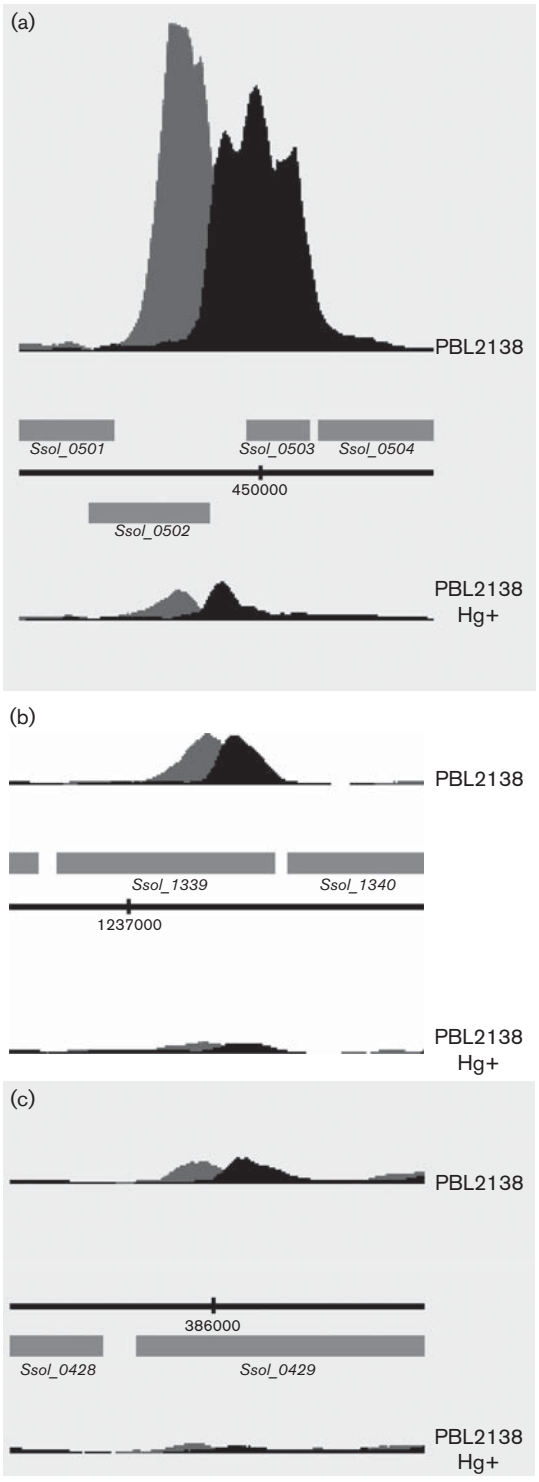

Fig. 3. ChIP-Seq analysis: positional read density derived from the ChIP-Seq data for three representative peaks identified in the Hg+ and Hg− conditions. (a) The canonical peak at *merHAp*. (b) The peak at *Ssol_1340*. (c) The peak at *Ssol_0428*. Two ChIP-Seq tracks are shown for each peak shown, one for each of the Hg+ and Hg− conditions, and are marked accordingly. Grey tracks correspond to forward read density, whilst black tracks correspond to reverse read density. A black horizontal line indicates the genome coordinate. Genomic regions encoding genes are indicated by grey rectangles and labelled accordingly. Genes above the genomic coordinate are transcribed from left to right, whilst genes below are transcribed from right to left. The *y*-axes in all panels are equally scaled.

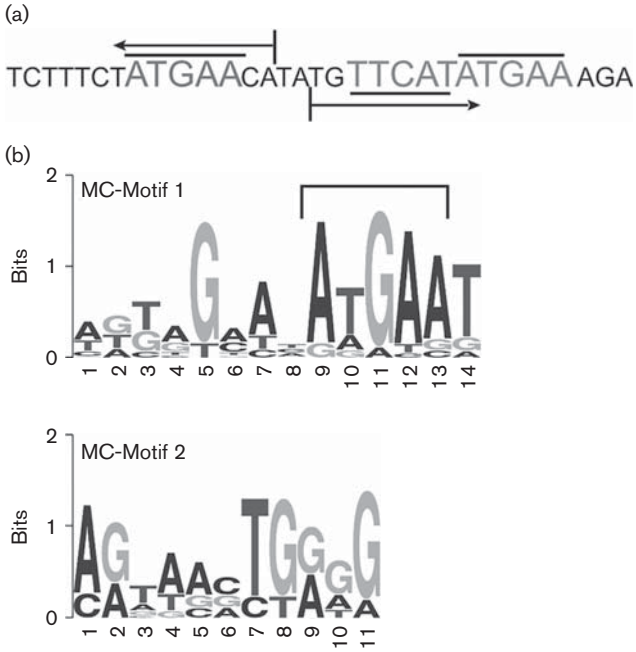

Fig. 4. MerR-associated binding motifs identified by MotifCatcher. (a) The sequence of the canonical MerR-binding site located between divergent *merH* and *merR* coding sequences is shown (horizontal arrows) separated by a 2 nt spacer. The conserved AWGAA elements identified in MC-Motif 1 by MotifCatcher are shown (grey and large font) with overlines and underlines, respectively. (b) Motif logos of MC-Motif 1 and MC-Motif 2. The AWGAA element of MC-Motif 1 is indicated by a bracket.

of the canonical *Ssol_0502*/*Ssol_0503* MerR-binding site (TCTTTCTATGAACATATGTTCATATGAAAGA; Fig. 4a) (Schelert *et al.*, 2006). MC-Motif 2 does not seem to share sequence similarity with this canonical MerR-binding sequence (Fig. 4a) and does not overlap with MC-Motif 1. Rather, when MC-Motif 2 co-occurs with MC-Motif 1 it is typically found ~48 bp downstream.

MC-Motif 1 mapped to 19/31 (61.3 %, $P=2.18\times10^{-59}$ hypergeometric test) high-confidence ChIP-Seq peaks that shared a consensus motif site. Meanwhile MC-Motif 2 occurred in 15/31 (48.4 %, $P=1.44\times10^{-46}$ hypergeometric test) of the ChIP-Seq peaks. The high-confidence ChIP-Seq peaks contain one or both of these motifs 24/31 (77.4 %, *P*-value$=1.94\times10^{-74}$ hypergeometric test) of the time. Interestingly, the motif was not found by MAST at the canonical *merAp* (*Ssol_0502/Ssol_0503*) promoter

site although manual inspection clearly showed the presence of a sequence corresponding to MC-Motif 1. An additional manual inspection of the remaining seven high-confidence peaks, where neither of the two putative MerR motifs was found by MAST, nevertheless revealed that five additional peaks contained a MC-Motif 1-like GNNNATGAA site within a 150 bp window around the peak centre. The remaining two peaks did not contain GNNNATGAA sites within the same 150 bp window. However, of these two unassociated peaks, one contained a GNNNTTGAA site and the other was discovered through a MAST search against the *S. solfataricus* 98/2 genome to contain the canonical *merR* motif.

Peak enrichment was calculated with respect to a local background for the 31 high-confidence peaks with and without mercury treatment (Table S3). In addition, the ratio of background-adjusted peak enrichment between normal (Hg−) and mercury-treated (Hg+) conditions was calculated to assess the effect of mercury on enrichment. A decrease in this (Hg−/Hg+) enrichment ratio would indicate that the addition of mercury decreased the association of MerR-HA with the respective binding site. All of the 31 high-confidence peaks showed a decrease in the (Hg−/Hg+) enrichment ratio, indicating that the affinity of MerR for these sites may have changed (likely decreased) upon binding mercury. This would be consistent with MerR's known role as a mercury-dependent repressor (Schelert *et al.*, 2006). It is also worth noting that whilst the (Hg−/Hg+) enrichment ratio was >1 in all cases (ranging from 1.76 to 8.27), nearly all peaks showed MerR binding at levels above the local background, even in the presence of mercury. This is also consistent with previous observations of MerR function showing that whilst a functional switch occurs in MerR which leads to derepression of regulated genes, this switch need not dissociate the protein completely from its binding site (Schelert *et al.*, 2006). Finally, we note that during the manual data curation, eight peaks were found that showed greater enrichment in the Hg+ track than in the Hg− track. Five of these peaks were also detected in the Hg− track (Table S4). The remaining three peaks were discovered only in the Hg+ track, indicating a potential increase in MerR binding at those sites upon addition of mercury (Table S4). These Hg+ unique peaks were found within the ORFs of *Ssol_0569*, *Ssol_2401* and *Ssol_2788*, and had an increase in enrichment over the Hg− condition (Hg+/Hg−) by 2.66-, 4.44- and 2.75-fold, respectively. Manual analysis of the Hg+ enriched *merR* binding regions revealed only one ATGAA site in the *Ssol_2401*-associated peak.

Identification of an archaeal mercury regulon

To determine whether the binding events were sufficient to mediate MerR-specific or mercury-specific transcriptional regulation, genes of interest identified from ChIP-Seq data were selected for further study by qRT-PCR (Fig. 5). Genes were selected from the larger set based on annotation indicating possible MerR control, responsiveness to mercury stress and metal import, binding or efflux. The influence of mercury and the presence of MerR was examined on two high-priority target genes (*Ssol_0429* and *Ssol_1340*) and three low-priority target genes (*Ssol_0152*, *Ssol_0428* and *Ssol_0648*). The two high-priority genes were found in the high-confidence peak list and contained a copy of MC-Motif 1 (Table 1). The low-priority targets – whilst associated with a detectable peak and potentially of functional relevance based on their annotations – were smaller peaks lacking either MC-Motif 1 or MC-Motif 2 (Table 1). Transcript abundance for these targets was determined by qRT-PCR in cell lines with a functional copy of MerR without or with mercury treatment, or in the absence of MerR without metal treatment (Fig. 5). All the genes tested exhibited significantly increased transcript abundance in response to mercury treatment. Of these, the solute-binding protein (*Ssol_0429*) was the most strongly influenced by mercury challenge. *Ssol_0429*, a substrate-binding component of an uncharacterized ATP-binding cassette (ABC)-type nickel/oligopeptide-like import system, contains the type 2 periplasmic binding fold. The solute-binding protein gene was upregulated 13.9-fold in response to mercury challenge in the MerR-containing strain. Only one gene, the AAA ATPase (*Ssol_0648*), had increased transcript abundance in the absence of MerR without metal treatment, indicating that it was negatively regulated by MerR. All of the low-priority targets, including the putative sialic acid transporter (*Ssol_0152*), the inner membrane-binding protein (*Ssol_0428*) and the AAA ATPase (*Ssol_0648*), exhibited only small but significant responses to mercury challenge. Interestingly, the transcript abundance of both *marR* (*Ssol_1340*) and the sialic acid transporter (*Ssol_0152*) were reduced in the absence of MerR, implicating a positive role for MerR in their expression.

Effect of solute-binding protein (*Ssol_0429*) mutation

To further understand the importance of the solute-binding protein (*Ssol_0429*), a mutant cell line of *S. solfataricus* with a disrupted copy of *Ssol_0429* was created by inserting a functional copy of the *lacS* gene. This cell line, along with its WT parent and a *merA* disruption mutant, were then tested for mercury sensitivity. The normal response of the WT strain consists of a lag in growth immediately after metal challenge, with a resumption of growth after 72 h. The *Ssol_0429* mutant strain when challenged with Hg(II) was more resistant than the WT strain (Fig. 6a). As demonstrated previously (Schelert *et al.*, 2006), a *merA* (mercuric reductase) mutant strain was much more sensitive to metal challenge than the WT (Fig. 6a). The reduced sensitivity of the *Ssol_0429* mutant relative to the WT was consistent with a role for this protein as an importer of Hg(II). To further examine this possibility, ICP-MS was used to measure intracellular levels of mercury following metal

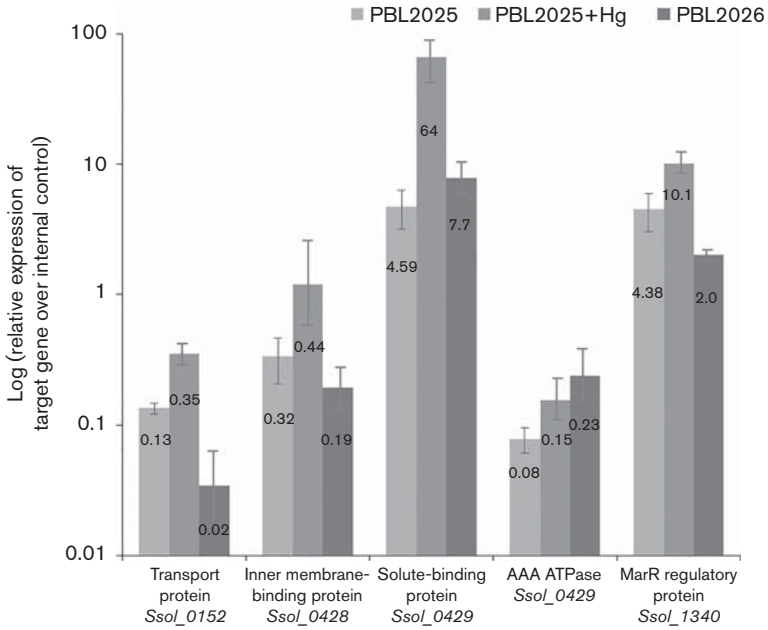

Fig. 5. qRT-PCR analysis to detect the response of selected genes to mercury challenge. Strains: WT, MerR$^+$ (PBL2025) without mercury challenge (light grey) or with mercury challenge (medium grey) and the *merR* mutant MerR$^-$ (PBL2026) without mercury challenge (dark grey). Genes: transport protein (*Ssol_0152*), inner membrane-binding protein (*Ssol_0428*), solute-binding protein (*Ssol_0429*), AAA ATPase (*Ssol_0648*) and MarR regulatory protein (*Ssol_1340*). Transcript abundance was determined by qRT-PCR and values were normalized to the transcript abundance of the generalized transcription factor TFBI (*Ssol_1927*) within each RNA sample obtained from cells in the exponential growth phase as described previously (Maezato *et al.*, 2011). For mercury-challenged samples, RNA was obtained 4 h after the addition of 0.5 μM Hg(II). Technical repeats produced <10 % variation.

treatment by comparing equal amounts of cell biomass. In the WT strain, intracellular mercury levels increased rapidly but briefly within 10 h of metal challenge, returning to baseline 72 h later (Fig. 6b). In contrast, mercury levels remained at low levels in the *Ssol_0429* mutant strain throughout the period of metal challenge (Fig. 6b). As observed previously, the *merA* mutant accumulated the highest levels of mercury during the metal challenge period (Schelert *et al.*, 2013), consistent with the loss of mercuric reductase. As an additional control, the intracellular levels of molybdenum were examined in the same cell extracts during the mercury treatment time course. The abundance of this metal (2100 ± 242 p.p.t.) remained unchanged in all the strains tested. This result indicates that decreased amounts of intracellular mercury arose from specific disruption of *Ssol_0429*. These results indicated that solute-binding protein (*Ssol_0429)* of *S. solfataricus* identified by ChIP-Seq and qRT-PCR analysis plays a role in mercury import.

DISCUSSION

Here, we show the utility of ChIP-Seq combined with qRT-PCR to identify key elements of the archaeal mercury regulon. Of the implicated loci, genetic analysis combined with ICP-MS to measure intracellular metal accumulation revealed a role for *Ssol_0429* as a transporter of mercury (Fig. 6 a, b). *Ssol_0429* encodes a Family 5 extracellular solute-binding protein. Inactivation of this protein by gene disruption generated a mercury-resistant phenotype and greatly reduced the quantity of metal taken up by metal-challenged cells. As these types of uptake transport system are broadly distributed and transport a wide variety of substrates (Tam & Saier, 1993), it remains possible that its natural ligand is some other cationic element, perhaps a nutrient. Mercury may use this route for cellular entry due to structural similarity. In archaea, and particularly in members belonging to the phylum *Crenarchaeota*, this family of solute-binding proteins is highly conserved and shares a high percentage of sequence similarity. This suggests that mercury may enter cells of this lineage using the same path. The conserved domain within this protein represents the substrate-binding domain of an uncharacterized ABC-type nickel/dipeptide/oligopeptide-like transporter (Thomas *et al.*, 2006). The oligopeptide-binding protein OppA and the dipeptide-binding protein DppA show significant sequence similarity to NikA, the progenitor nickel receptor (Palmieri *et al.*, 2006). The structural topology of these conserved domains is most similar to that of the type 2 periplasmic binding proteins, which are responsible for the uptake of a variety of substrates such as phosphate, sulfate,

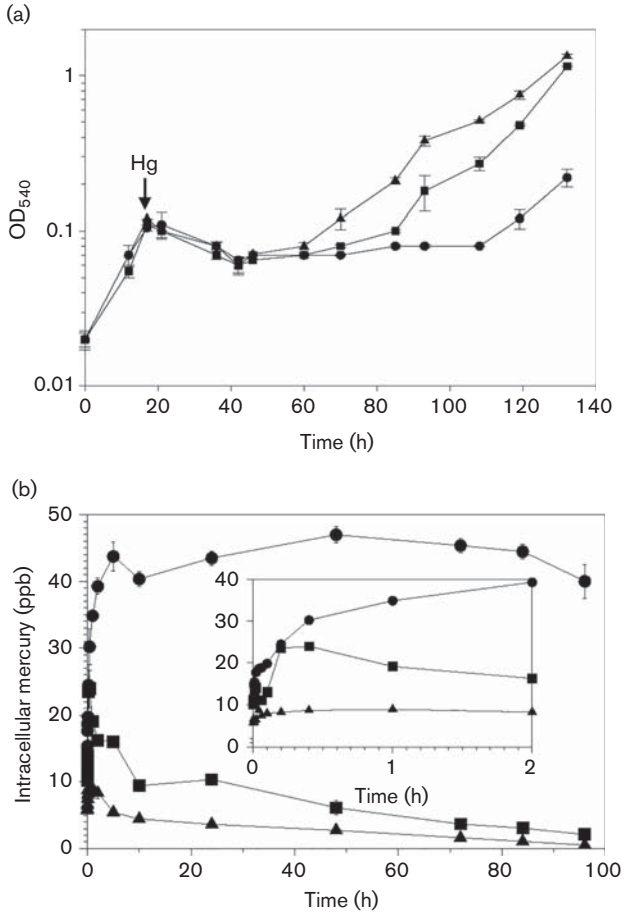

Fig. 6. Response of *S. solfataricus* strains to Hg(II) challenge and intracellular Hg(II) measurement by ICP-MS. (a) All strains were treated with Hg(II) (arrow) at a concentration of 0.5 μM. *Ssol_0429* mutant (▲), WT (■) and *merA* mutant (●). Experiments were performed in duplicate. (b) ICP-MS analysis of cell-associated mercury. Cells were removed at the indicated times after the Hg(II) challenge. *merA* mutant (●), WT (■) and *Ssol_0429* mutant (▲). Zero time in (b) corresponds to 18 h in (a). Inset shows cell-associated mercury during early time points (0–2 h).

polysaccharides, lysine/arginine/ornithine and histidine (Silver & Walderhaug, 1992). *S. solfataricus* contains 37 putative ABC transporters at the genome level, but only a few of these systems have been characterized functionally (Gogliettino *et al.*, 2010). OppA is an outer cell surface-anchored protein and its expression is highly induced in the presence of a source of peptides in the culture medium, suggesting that the genetic loci for sugar and di-/oligopeptide uptake systems could be independent in the *S. solfataricus* genome. (Gogliettino *et al.*, 2010).

The ChIP-Seq (Fig. 3b) and qRT-PCR (Fig. 5) analyses also implicated the transcription factor MarR (*Ssol_1340*) as a member of the mercury regulon. As a regulator of oxidative stress (Alekshun & Levy, 1999; Ariza *et al.*, 1994; Cohen *et al.*, 1993; Ellison & Miller, 2006; Li *et al.*, 2011; Sulavik *et al.*, 1995), its regulated expression may arise from the impact of mercury on free thiols that are titrated by metal sequestration leading to oxidative stress. The MarR transcription factor family shares a common origin in bacteria and archaea as revealed by a conserved N-terminal 40 aa sequence motif (Pérez-Rueda & Collado-Vides, 2001). *Ssol_1340* has been annotated as a member of the MarR family by sequence homology and is implicated in DNA binding due to its composition of divalent metal cation-binding residues (GenBank accession number YP_005643177). Structurally homologous MarR proteins are winged helix–turn–helix dimers that bind palindromic promoter sequences to regulate transcription through exogenous ligand binding (Martin & Rosner, 1995). The regulated genes are functionally linked to antibiotic resistance, oxidative stress response, heat shock resistance, virulence, catabolism of aromatic compounds, and export of disinfectants and organic solvents (Alekshun & Levy, 1999; Ariza *et al.*, 1994; Cohen *et al.*, 1993; Ellison & Miller, 2006; Li *et al.*, 2011; Sulavik *et al.*, 1995). Three other MarR proteins of *Sulfolobus* have now been characterized by crystallography, and their ability to bind DNA is modulated by ligands with various affinities (Kumarevel *et al.*, 2008). The MarR regulator BldR from *S. solfataricus* strain P2 is a transcriptional activator of one of the genes encoding alcohol dehydrogenase (*Sso2536*) and is co-transcribed with a multidrug transporter (*Sso1351*), presumably for the detoxification of aromatic compounds (Di Fiore *et al.*, 2009; Fiorentino *et al.*, 2007). BldR2 shares 35 % sequence homology with BldR and has also been characterized (Fiorentino *et al.*, 2011). The structural homologue ST1710 (StEmrR) from *Solfolobus tokadii* has affinity for ligands of salicylate, carbonyl cyanide *m*-chlorophenylhydrazone and ethidium, which is characteristic of MarR family proteins from *Escherichia coli* (Miyazono *et al.*, 2007; Yu *et al.*, 2009).

In addition to these new members of the MerR regulon, numerous other candidate target genes were identified by this study, but have yet to be pursued in greater depth (Table S3). Previous ChIP-Seq and transcriptome analysis of *Mycobacterium tuberculosis* have clearly demonstrated that novel and functionally relevant binding sites can be identified with high frequency in the relatively large number of smaller ChIP-Seq peaks (Galagan *et al.*, 2013; Gao *et al.*, 2012). Given this precedent, the numerous new candidate targets (particularly those containing instances of MC-Motif 1 and MC-Motif 2; Fig. 4) should be pursued in future studies to test the transcriptional dependence of the associated genes on MerR activity. In addition, a comprehensive approach using genetic strategies could be applied to extend the findings arising from apparent transcriptional dependence and thereby determine the functional importance of these genes *in vivo*.

ACKNOWLEDGEMENTS

This study was supported by funds from the University of Nebraska Cell Development Facility and NSF grant MCB-1517408 to P. B., and

bridge funding from the UC Davis Department of Biomedical Engineering and the UC Davis College of Engineering to M. T. F.

REFERENCES

Alekshun, M. N. & Levy, S. B. (1999). The *mar* regulon: multiple resistance to antibiotics and other toxic chemicals. *Trends Microbiol* **7**, 410–413.

Allen, M. B. (1959). Studies with *Cyanidium caldarium*, an anomalously pigmented chlorophyte. *Arch Mikrobiol* **32**, 270–277.

Ariza, R. R., Cohen, S. P., Bachhawat, N., Levy, S. B. & Demple, B. (1994). Repressor mutations in the *marRAB* operon that activate oxidative stress genes and multiple antibiotic resistance in *Escherichia coli*. *J Bacteriol* **176**, 143–148.

Bailey, T. L. & Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* **2**, 28–36.

Bailey, T. L. & Gribskov, M. (1998). Combining evidence using *p*-values: application to sequence homology searches. *Bioinformatics* **14**, 48–54.

Bare, J. C., Koide, T., Reiss, D. J., Tenenbaum, D. & Baliga, N. S. (2010). Integration and visualization of systems biology data in context of the genome. *BMC Bioinformatics* **11**, 382.

Barkay, T., Miller, S. M. & Summers, A. O. (2003). Bacterial mercury resistance from atoms to ecosystems. *FEMS Microbiol Rev* **27**, 355–384.

Bini, E., Dikshit, V., Dirksen, K., Drozda, M. & Blum, P. (2002). Stability of mRNA in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *RNA* **8**, 1129–1136.

Brock, T. D., Brock, K. M., Belly, R. T. & Weiss, R. L. (1972). *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch Mikrobiol* **84**, 54–68.

Cohen, S. P., Hächler, H. & Levy, S. B. (1993). Genetic and functional analysis of the multiple antibiotic resistance (*mar*) locus in *Escherichia coli*. *J Bacteriol* **175**, 1484–1492.

Di Fiore, A., Fiorentino, G., Vitale, R. M., Ronca, R., Amodeo, P., Pedone, C., Bartolucci, S. & De Simone, G. (2009). Structural analysis of BldR from *Sulfolobus solfataricus* provides insights into the molecular basis of transcriptional activation in Archaea by MarR family proteins. *J Mol Biol* **388**, 559–569.

Dixit, V., Bini, E., Drozda, M. & Blum, P. (2004). Mercury inactivates transcription and the generalized transcription factor TFB in the archaeon *Sulfolobus solfataricus*. *Antimicrob Agents Chemother* **48**, 1993–1999.

Ellison, D. W. & Miller, V. L. (2006). Regulation of virulence by members of the MarR/SlyA family. *Curr Opin Microbiol* **9**, 153–159.

Ettema, T. J., Huynen, M. A., de Vos, W. M. & van der Oost, J. (2003). TRASH: a novel metal-binding domain predicted to be involved in heavy-metal sensing, trafficking and resistance. *Trends Biochem Sci* **28**, 170–173.

Fiorentino, G., Ronca, R., Cannio, R., Rossi, M. & Bartolucci, S. (2007). MarR-like transcriptional regulator involved in detoxification of aromatic compounds in *Sulfolobus solfataricus*. *J Bacteriol* **189**, 7351–7360.

Fiorentino, G., Del Giudice, I., Bartolucci, S., Durante, L., Martino, L. & Del Vecchio, P. (2011). Identification and physicochemical characterization of BldR2 from *Sulfolobus solfataricus*, a novel archaeal member of the MarR transcription factor family. *Biochemistry* **50**, 6607–6621.

Galagan, J. E., Minch, K., Peterson, M., Lyubetskaya, A., Azizi, E., Sweet, L., Gomes, A., Rustad, T., Dolganov, G. & other authors (2013). The *Mycobacterium tuberculosis* regulatory network and hypoxia. *Nature* **499**, 178–183.

Gao, C. H., Yang, M. & He, Z.-G. (2012). Characterization of a novel ArsR-like regulator encoded by Rv2034 in *Mycobacterium tuberculosis*. *PLoS One* **7**, e36255.

Gogliettino, M., Balestrieri, M., Pocsfalvi, G., Fiume, I., Natale, L., Rossi, M. & Palmieri, G. (2010). A highly selective oligopeptide binding protein from the archaeon *Sulfolobus solfataricus*. *J Bacteriol* **192**, 3123–3131.

Higuchi, R., Krummel, B. & Saiki, R. K. (1988). A general method of *in vitro* preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res* **16**, 7351–7367.

Kumarevel, T., Tanaka, T., Nishio, M., Gopinath, S. C. B., Takio, K., Shinkai, A., Kumar, P. K. R. & Yokoyama, S. (2008). Crystal structure of the MarR family regulatory protein, ST1710, from *Sulfolobus tokodaii* strain 7. *J Struct Biol* **161**, 9–17.

Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25.

Lassmann, T., Hayashizaki, Y. & Daub, C. O. (2009). TagDust – a program to eliminate artifacts from next generation sequencing data. *Bioinformatics* **25**, 2839–2840.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. & 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079.

Li, J. S., Bi, Y. T., Dong, C., Yang, J. F. & Liang, W. D. (2011). Transcriptome analysis of adaptive heat shock response of *Streptococcus thermophilus*. *PLoS One* **6**, e25777.

Maezato, Y., Dana, K. & Blum, P. (2011). Engineering thermoacidophilic archaea using linear DNA recombination. *Methods Mol Biol* **765**, 435–445.

Martin, R. G. & Rosner, J. L. (1995). Binding of purified multiple antibiotic-resistance repressor protein (MarR) to *mar* operator sequences. *Proc Natl Acad Sci U S A* **92**, 5456–5460.

Miyazono, K., Tsujimura, M., Kawarabayasi, Y. & Tanokura, M. (2007). Crystal structure of an archaeal homologue of multidrug resistance repressor protein, EmrR, from hyperthermophilic archaea *Sulfolobus tokodaii* strain 7. *Proteins* **67**, 1138–1146.

Palmieri, G., Casbarra, A., Fiume, I., Catara, G., Capasso, A., Marino, G., Onesti, S. & Rossi, M. (2006). Identification of the first archaeal oligopeptide-binding protein from the hyperthermophile *Aeropyrum pernix*. *Extremophiles* **10**, 393–402.

Pérez-Rueda, E. & Collado-Vides, J. (2001). Common history at the origin of the position-function correlation in transcriptional regulators in archaea and bacteria. *J Mol Evol* **53**, 172–179.

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*., Vienna: R Foundation for Statistical Computing.

Rolfsmeier, M. & Blum, P. (1995). Purification and characterization of a maltase from the extremely thermophilic crenarchaeote *Sulfolobus solfataricus*. *J Bacteriol* **177**, 482–485.

Sahlman, L., Wong, W. & Powlowski, J. (1997). A mercuric ion uptake role for the integral inner membrane protein, MerC, involved in bacterial mercuric ion resistance. *J Biol Chem* **272**, 29518–29526.

Schelert, J., Dixit, V., Hoang, V., Simbahan, J., Drozda, M. & Blum, P. (2004). Occurrence and characterization of mercury resistance in the hyperthermophilic archaeon *Sulfolobus solfataricus* by use of gene disruption. *J Bacteriol* **186**, 427–437.

Schelert, J., Drozda, M., Dixit, V., Dillman, A. & Blum, P. (2006). Regulation of mercury resistance in the crenarchaeote *Sulfolobus solfataricus*. *J Bacteriol* **188**, 7141–7150.

Schelert, J., Rudrappa, D., Johnson, T. & Blum, P. (2013). Role of MerH in mercury resistance in the archaeon *Sulfolobus solfataricus*. *Microbiology* **159**, 1198–1208.

Seitzer, P., Wilbanks, E. G., Larsen, D. J. & Facciotti, M. T. (2012). A Monte Carlo-based framework enhances the discovery and interpretation of regulatory sequence motifs. *BMC Bioinformatics* **13**, 317.

Silver, S. & Walderhaug, M. (1992). Gene regulation of plasmid- and chromosome-determined inorganic ion transport in bacteria. *Microbiol Rev* **56**, 195–228.

Simbahan, J., Kurth, E., Schelert, J., Dillman, A., Moriyama, E., Jovanovich, S. & Blum, P. (2005). Community analysis of a mercury hot spring supports occurrence of domain-specific forms of mercuric reductase. *Appl Environ Microbiol* **71**, 8836–8845.

Sone, Y., Nakamura, R., Pan-Hou, H., Itoh, T. & Kiyono, M. (2013). Role of MerC, MerE, MerF, MerT, and/or MerP in resistance to mercurials and the transport of mercurials in *Escherichia coli*. *Biol Pharm Bull* **36**, 1835–1841.

Sowers, K. R., Blum, P. H. & DasSarma, S. (2007). Gene transfer in archaea. In *Methods for General and Molecular Microbiology*, 3rd edn, pp. 800–824. Edited by C. A. Reddy, T. J. Beveridge, J. A. Breznak, G. A. Marzluf & T. M. Schmidt. Washington, DC: American Society for Microbiology.

Sulavik, M. C., Gambino, L. F. & Miller, P. F. (1995). The MarR repressor of the multiple antibiotic resistance (*mar*) operon in *Escherichia coli*: prototypic member of a family of bacterial regulatory proteins involved in sensing phenolic compounds. *Mol Med* **1**, 436–446.

Tam, R. & Saier, M. H. Jr (1993). Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol Rev* **57**, 320–346.

Thomas, G. H., Southworth, T., León-Kempis, M. R., Leech, A. & Kelly, D. J. (2006). Novel ligands for the extracellular solute receptors of two bacterial TRAP transporters. *Microbiology* **152**, 187–198.

Wilbanks, E. G., Larsen, D. J., Neches, R. Y., Yao, A. I., Wu, C.-Y., Kjolby, R. A. S. & Facciotti, M. T. (2012). A workflow for genome-wide mapping of archaeal transcription factors with ChIP-seq. *Nucleic Acids Res* **40**, e74.

Wilson, J. R., Leang, C., Morby, A. P., Hobman, J. L. & Brown, N. L. (2000). MerF is a mercury transport protein: different structures but a common mechanism for mercuric ion transporters? *FEBS Lett* **472**, 78–82.

Worthington, P., Hoang, V., Perez-Pomares, F. & Blum, P. (2003). Targeted disruption of the alpha-amylase gene in the hyper-thermophilic archaeon *Sulfolobus solfataricus*. *J Bacteriol* **185**, 482–488.

Yu, L., Fang, J. & Wei, Y. (2009). Characterization of the ligand and DNA binding properties of a putative archaeal regulator ST1710. *Biochemistry* **48**, 2099–2108.

Edited by: D. Nies