




# Differential RNA Sequencing Implicates Sulfide as the Master Regulator of $S^0$ Metabolism in *Chlorobaculum tepidum* and Other Green Sulfur Bacteria

Jacob M. Hilzinger,<sup>a,b</sup> Vidhyavathi Raman,<sup>a,b\*</sup> Kevin E. Shuman,<sup>b,c\*</sup>  Brian J. Eddie,<sup>a,b\*</sup>  Thomas E. Hanson<sup>a,b,c</sup>

<sup>a</sup>School of Marine Science and Policy, University of Delaware, Newark, Delaware, USA

<sup>b</sup>Delaware Biotechnology Institute, University of Delaware, Newark, Delaware, USA

<sup>c</sup>Department of Biological Sciences, University of Delaware, Newark, Delaware, USA

**ABSTRACT** The green sulfur bacteria (*Chlorobiaceae*) are anaerobes that use electrons from reduced sulfur compounds (sulfide,  $S^0$ , and thiosulfate) as electron donors for photoautotrophic growth. *Chlorobaculum tepidum*, the model system for the *Chlorobiaceae*, both produces and consumes extracellular  $S^0$  globules depending on the availability of sulfide in the environment. These physiological changes imply significant changes in gene regulation, which has been observed when sulfide is added to *Cba. tepidum* growing on thiosulfate. However, the underlying mechanisms driving these gene expression changes, i.e., the specific regulators and promoter elements involved, have not yet been defined. Here, differential RNA sequencing (dRNA-seq) was used to globally identify transcript start sites (TSS) that were present during growth on sulfide, biogenic  $S^0$ , and thiosulfate as sole electron donors. TSS positions were used in combination with RNA-seq data from cultures growing on these same electron donors to identify both basal promoter elements and motifs associated with electron donor-dependent transcriptional regulation. These motifs were conserved across homologous *Chlorobiaceae* promoters. Two lines of evidence suggest that sulfide-mediated repression is the dominant regulatory mode in *Cba. tepidum*. First, motifs associated with genes regulated by sulfide overlap key basal promoter elements. Second, deletion of the *Cba. tepidum* 1277 (*CT1277*) gene, encoding a putative regulatory protein, leads to constitutive overexpression of the sulfide:quinone oxidoreductase *CT1087* in the absence of sulfide. The results suggest that sulfide is the master regulator of sulfur metabolism in *Cba. tepidum* and the *Chlorobiaceae*. Finally, the identification of basal promoter elements with differing strengths will further the development of synthetic biology in *Cba. tepidum* and perhaps other *Chlorobiaceae*.

**IMPORTANCE** Elemental sulfur is a key intermediate in biogeochemical sulfur cycling. The photoautotrophic green sulfur bacterium *Chlorobaculum tepidum* either produces or consumes elemental sulfur depending on the availability of sulfide in the environment. Our results reveal transcriptional dynamics of *Chlorobaculum tepidum* on elemental sulfur and increase our understanding of the mechanisms of transcriptional regulation governing growth on different reduced sulfur compounds. This report identifies genes and sequence motifs that likely play significant roles in the production and consumption of elemental sulfur. Beyond this focused impact, this report paves the way for the development of synthetic biology in *Chlorobaculum tepidum* and other *Chlorobiaceae* by providing a comprehensive identification of promoter elements for control of gene expression, a key element of strain engineering.

**KEYWORDS** dRNA-seq, *Chlorobaculum tepidum*, *Chlorobiaceae*, sulfur metabolism, energy metabolism, transcriptional regulation

**Received** 6 September 2017 **Accepted** 14 November 2017

**Accepted manuscript posted online** 17 November 2017

**Citation** Hilzinger JM, Raman V, Shuman KE, Eddie BJ, Hanson TE. 2018. Differential RNA sequencing implicates sulfide as the master regulator of  $S^0$  metabolism in *Chlorobaculum tepidum* and other green sulfur bacteria. *Appl Environ Microbiol* 84:e01966-17. <https://doi.org/10.1128/AEM.01966-17>.

**Editor** Harold L. Drake, University of Bayreuth

**Copyright** © 2018 American Society for Microbiology. All Rights Reserved.

Address correspondence to Thomas E. Hanson, [tehanson@udel.edu](mailto:tehanson@udel.edu).

\* Present address: Vidhyavathi Raman, Noble Research Institute, Ardmore, Oklahoma, USA; Kevin E. Shuman, Wesley College, Dover, Delaware, USA; Brian J. Eddie, United States Naval Research Laboratory, Washington, DC, USA.

J.M.H. and V.R. contributed equally to this article.

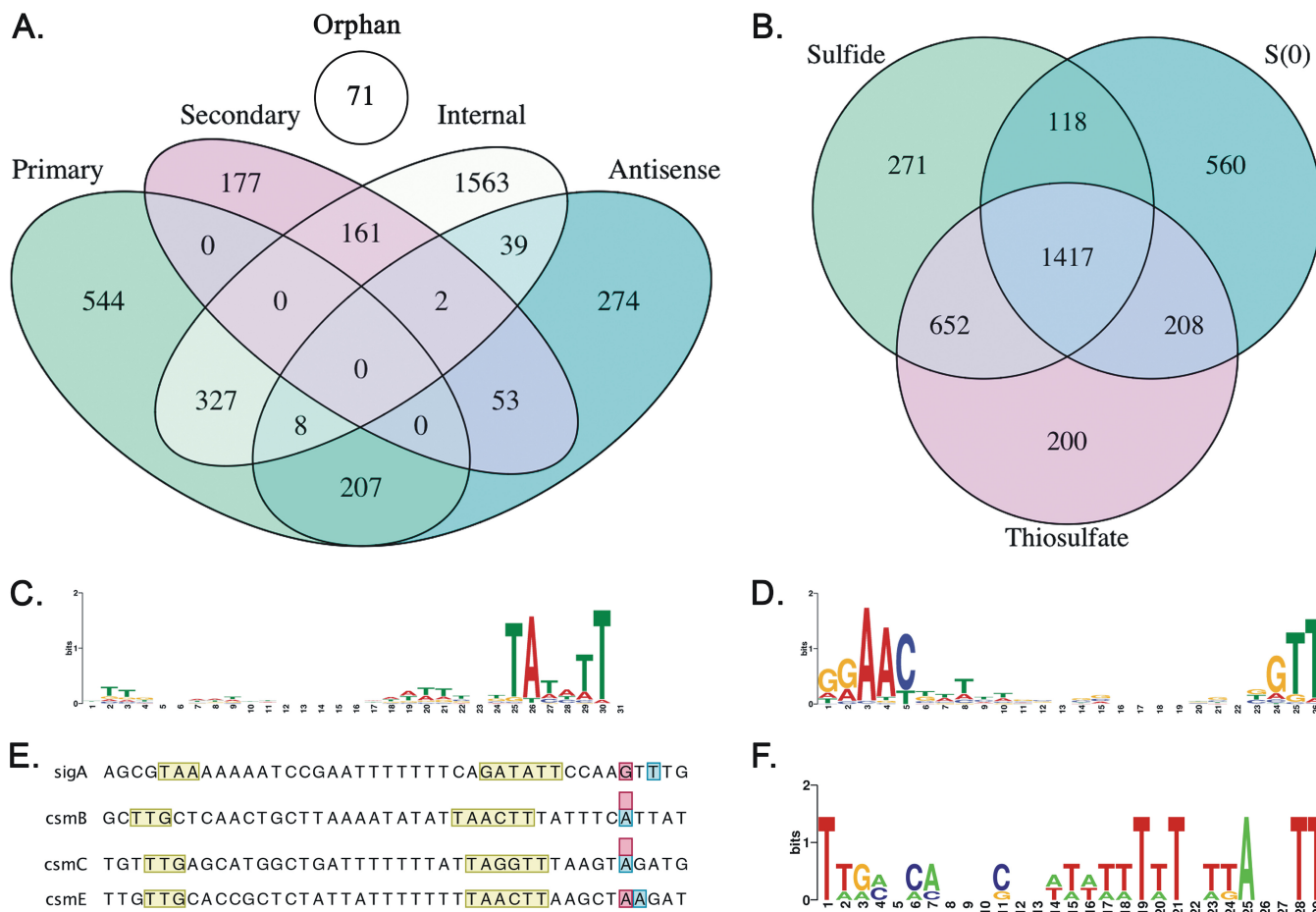
The green sulfur bacteria (*Chlorobiaceae*) are a family of anaerobic photoautotrophic sulfur oxidizers. *Chlorobaculum tepidum*, the model system for this family, both produces and consumes extracellular  $S^0$  globules depending on the availability of sulfide in the environment (1). Recently, Hanson et al. showed that biogenic  $S^0$  globules could serve as the sole photosynthetic electron donor for growth of *Cba. tepidum* and that cell- $S^0$  contact was required for that growth (2). Expanding upon that work, Marnocha et al. showed that cell- $S^0$  contact was dynamic, with large populations of unattached *Cba. tepidum* cells growing at rates similar to those of attached cells (3). As polysulfides were detected in supernatants from cultures producing and consuming  $S^0$ , Marnocha et al. proposed a model whereby polysulfides act as soluble intermediates in the formation and degradation of  $S^0$  globules and could feed unattached cells (3). Those two studies have significantly increased our understanding of how *Cba. tepidum* interacts with  $S^0$ . However, many issues remain, in particular, how  $S^0$  metabolism is regulated and what proteins play a role in cell- $S^0$  attachment.

The analysis of the *Cba. tepidum* genome showed that it encoded few recognizable transcriptional regulators, leading the authors to conclude that *Cba. tepidum* employs little transcriptional regulation (4). However, Eddie and Hanson observed a complex transcriptional response following the addition of sulfide to *Cba. tepidum* growing on thiosulfate (5). The transcript abundance of a two-gene operon (*Cba. tepidum* 1276 [CT1276]-CT1277) was significantly increased in the presence of sulfide, with the protein encoded by CT1277 belonging to the helix-turn-helix xenobiotic response element (HTH XRE) protein superfamily, indicating that transcriptional regulators with limited functional annotation may contribute to gene expression regulation in *Cba. tepidum* (5). While that study identified numerous transcriptionally regulated genes for the transition from thiosulfate to sulfide as a primary electron donor, the characteristics of the *Cba. tepidum* transcriptome during growth on  $S^0$  as a sole electron donor have not been documented.

Here, RNA sequencing (RNA-seq) and differential RNA-seq (dRNA-seq) were used to characterize the transcriptome of *Cba. tepidum* growing on biogenic  $S^0$  and to globally identify transcript start sites (TSS) active during growth on sulfide, biogenic  $S^0$ , or thiosulfate as the sole electron donor. RNA-seq data suggest that the most dynamic changes in transcript abundance occur in response to sulfide and that the majority of genes differentially expressed in response to sulfide are downregulated. TSS positions were used to identify putative promoter elements and, in combination with RNA-seq data, were used to identify DNA motifs that were conserved across homologous *Chlorobiaceae* promoters. The positions of the discovered motifs relative to TSS and basal promoter elements suggest that, in agreement with the RNA-seq data, repression is the dominant mode of transcriptional regulation for genes involved in sulfur metabolism. In support of this observation, deletion of CT1277 from the *Cba. tepidum* genome led to overexpression of the sulfide:quinone oxidoreductase (SQR) CT1087 (6) during growth on thiosulfate, suggesting that the CT1277 gene product acts as a repressor.

## RESULTS

**The transcriptome of *Cba. tepidum* grown with  $S^0$  as the sole electron donor.** RNA-seq analysis of transcriptomes in *Cba. tepidum* cultures growing on thiosulfate and biogenic  $S^0$  revealed similar expression profiles for growth on the two substrates. A total of 28,183,709 reads were uniquely aligned to the genome for the thiosulfate library, while the  $S^0$  library had 37,350,690 uniquely aligned reads. Despite the tight correlation of expression results between thiosulfate and  $S^0$ , 120 genes were differentially expressed (see Materials and Methods), with 55 genes being downregulated on  $S^0$  relative to thiosulfate and 65 genes being upregulated on  $S^0$  (see Table S1 in the supplemental material). Comparing the  $S^0$  transcriptome with the previously published sulfide transcriptome (5) identified 106 differentially expressed genes (see Materials and Methods), with 35 genes upregulated on sulfide relative to  $S^0$  and 71 genes downregulated on sulfide (Table S1).



**FIG 1** (A and B) Venn diagrams for TSS data organized by TSS classification (A) or by growth condition (B). (C and D) Core promoter motifs identified by analyzing 50 bp upstream of all TSS: the RpoD motif (C) and an ECF factor-like motif (D). (E) TSS identified by dRNA-seq here (pink) closely match those previously reported (blue) for the *sigA*, *csmB*, *csmC*, and *csmE* genes and are preceded by RpoD motifs (yellow); see the text for references. (F) Extended –10 motif for *Cba. tepidum* RpoD promoters.

The magnitude of transcript abundance changes was examined in detail across the substrates. Log<sub>2</sub> fold change values for genes with increased expression on sulfide relative to thiosulfate and S<sup>0</sup> (15.2 and 15.7, respectively) were both much larger than that of S<sup>0</sup> relative to thiosulfate (5.74). Similarly, the fold change values for genes with decreased expression on sulfide relative to thiosulfate and S<sup>0</sup> were considerably lower than that of S<sup>0</sup> relative to thiosulfate (–9.31 and –9.85 versus –2.30). This suggests that the most dynamic changes in transcript abundance occur in response to sulfide. These changes are clearly observed in genes encoding key components of the sulfur oxidation machinery of *Cba. tepidum* (7). For example, the thiosulfate oxidation (Sox) and S<sup>0</sup> oxidation (Dsr) operons were strongly downregulated on sulfide (5) but were not differentially expressed between growth on S<sup>0</sup> and growth on thiosulfate (Table S1).

**Global identification of TSS and basal promoter elements in *Cba. tepidum*.** The dRNA-seq protocol used in this study produced 6.1 to 17 million reads per replicate (Table S2), with 2.0 to 6.7 million reads uniquely aligned to the genome for each replicate. Analysis of these data identified a total of 3,426 putative TSS across the results of growth on three electron donors in *Cba. tepidum*: 1,086 primary TSS (pTSS), 393 secondary TSS (sTSS), 583 antisense TSS (asTSS), 2,100 internal TSS (iTSS), and 71 orphan TSS (oTSS) (Fig. 1A; see also Table S3). Although the absolute numbers of putative TSS (2,458 for sulfide, 2,303 for S<sup>0</sup>, and 2,477 for thiosulfate) were similar for all of the conditions, there were condition-specific differences between the predicted TSS (Fig. 1B). However, there was little correlation between the condition-dependent TSS

and changes in transcript abundance. A total of 1,417 of the predicted TSS (718 pTSS, 102 sTSS, 613 iTSS, 338 asTSS, and 33 oTSS) were found under all three conditions.

Two motifs were identified when a 50-bp sequence upstream of all TSS was analyzed by the MEME software package (Fig. 1C and D) (8). The most abundant motif (Fig. 1C) closely resembles consensus RpoD binding sequences, with the highest similarity to those of members of the *Bacteroidetes*, specifically, *Flavobacterium* spp. (TTG-N<sub>17-23</sub>-TANNTTTG) (9). This would be expected, as RpoD protein sequences from *Bacteroidetes* and *Cba. tepidum* form clades together and away from other RpoD proteins (10). However, the  $-7$  consensus sequence, TA[ATC][ATC][AT]T, is different than those of other published RpoD consensus binding sites (9), making the *Cba. tepidum* RpoD consensus binding sequence unique among the published consensus sequences reported to date. The RpoD motif was associated with 1,227 TSS (36% of all TSS) and with 64% of pTSS. Taken together, the data suggest that the primary sigma factor of *Cba. tepidum* is RpoD encoded by *CT1551* (*sigA*).

The second motif (Fig. 1D) resembles  $\sigma^{70}$  extracytoplasmic function (ECF) subfamily binding sequences, with the distinct AAC motif in the  $-35$  consensus region (11). This motif closely resembles the promoter consensus sequences of  $\sigma^E$  (GG[AG][AC]C-N<sub>18</sub>-[CG]GTTg) and  $\sigma^H$  ([CG]GGAAC-N<sub>17</sub>-[CG]GTT[CG]) from *Mycobacterium tuberculosis* (12, 13) and that of  $\sigma^R$  (GGAAT-N<sub>18</sub>-GTT) from *Streptomyces coelicolor* (14). As the *Cba. tepidum* genome encodes three  $\sigma^{70}$  ECF factors (CT0278, CT0502, and CT0648), this motif may represent a consensus motif that all three ECF factors bind to, or it may represent the consensus sequence of a number of highly similar and yet distinct motifs that bind the ECF factors with different affinities. For example,  $\sigma^X$  and  $\sigma^W$  promoter sequences of *Bacillus subtilis* are highly similar, with some promoters binding both proteins whereas others bind one but not the other (15). Thus, this motif likely binds one or more ECF factors in *Cba. tepidum*. This ECF-like motif was associated with 182 TSS (5.3% of all TSS) and with 7.9% of pTSS.

Aside from RpoD and the three ECF sigma factors, *Cba. tepidum* encodes one additional sigma factor, CT1193 (*rpoN*), a  $\sigma^{54}$  factor that controls expression of nitrogen-regulated genes, including the *nif* operon that encodes proteins for N<sub>2</sub> fixation in diverse bacteria. Phylogenetic footprinting of the *nifH* promoter across the *Chlorobiaceae* identified RpoN and NtrC binding motifs (data not shown) (16, 17). Using the FIMO search tool (18) in the MEME suite with the RpoN motif as a query returned one additional pTSS (with a false-discovery-rate [*q*] value of <0.05) preceding gene *CT0644*, which encodes an HSP20 family protein.

Four transcript start sites previously identified by primer extension analysis (*sigA*, *csmA*, *csmB*, and *csmE*) were captured in these data (Fig. 1E) (19, 20). While the *sigA* TSS predicted in this study was two base pairs upstream of the TSS identified by primer extension analysis and the *csmE* TSS was one base pair upstream, the *csmB* and *csmC* TSS were identical between the studies. All four promoters appear to be RpoD dependent, with the consensus motif, generated by WebLogo (21), closely resembling the RpoD motif identified by MEME with an extended  $-10$  region (Fig. 1C and F). Chung et al. (20) showed that the *csmA* genes were cotranscribed as a single unit from the *csmC* TSS. Additionally, the authors found a higher-abundance *csmA* transcript of approximately 350 bp, which they characterized as the processing product of the *csmA* transcript. However, we found an RpoD-dependent pTSS and a sTSS 57 and 67 bp upstream of the *csmA* start codon, respectively (Fig. S1). The pTSS occurs 5 bp upstream of the predicted 5' end of the *csmA* transcript predicted by Chung et al. (20). These TSS can explain the observations of Chung et al. without the need to invoke transcript processing.

**Putative sulfide operator sequence 1 (PSOS-1) is associated with *sqr* and putative regulatory genes.** *Cba. tepidum* displays a robust transcriptional response depending on the reduced sulfur compound utilized for growth. Binning genes by expression pattern, e.g., all genes with increased transcript abundance on sulfide relative to thiosulfate, and then analyzing regions upstream of the associated pTSS did not produce any significant motifs associated with the promoter regions (data

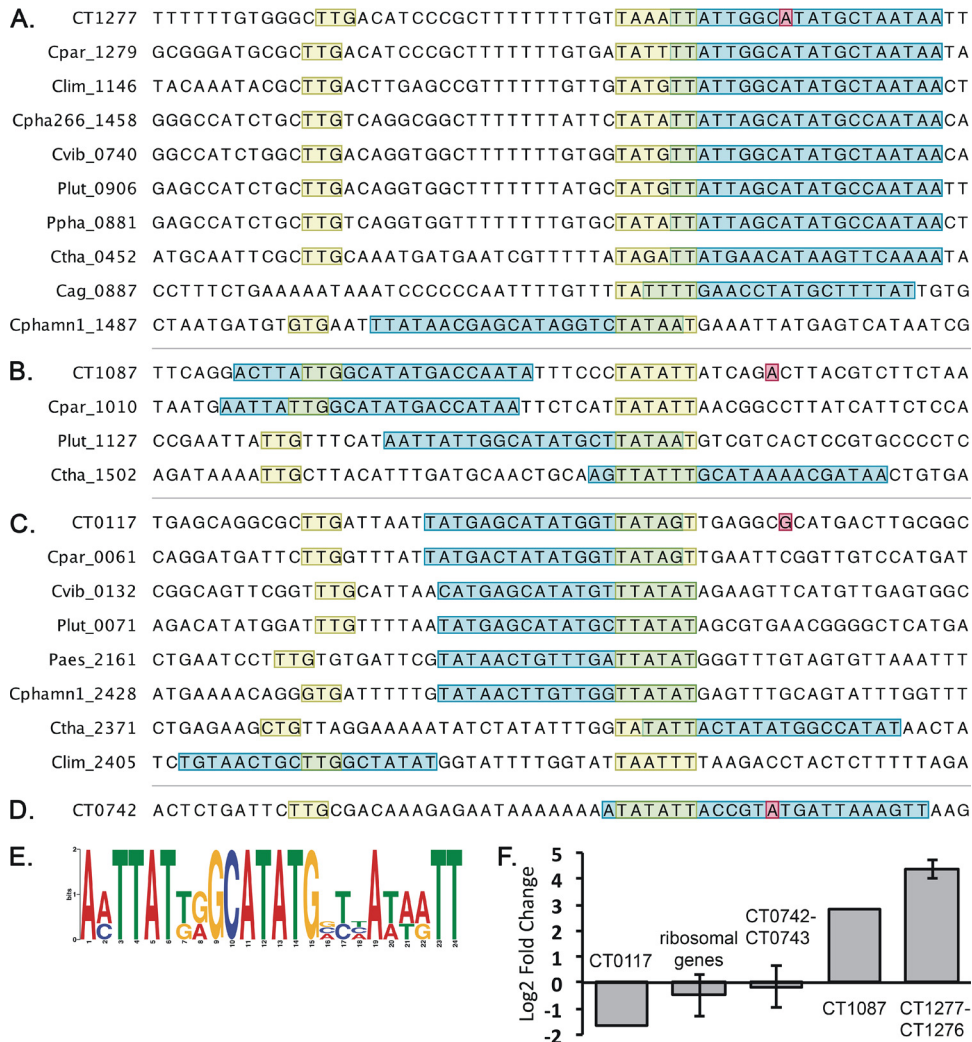
**TABLE 1** *Chlorobiaceae* genomes used for phylogenetic footprinting

Genome	Accession no.
<i>Chlorobaculum tepidum</i> TLS	<a href="#">NC_002932</a>
<i>Chlorobaculum parvum</i> NCIB 8327	<a href="#">NC_011027</a>
<i>Chlorobium chlorochromatii</i> CaD3	<a href="#">NC_007514</a>
<i>Chlorobium ferrooxidans</i> DSM 13031	<a href="#">NZ_AASE01000000</a>
<i>Chlorobium phaeobacteroides</i> DSM 266	<a href="#">NC_008639</a>
<i>Chlorobium phaeobacteroides</i> BS1	<a href="#">NC_010831</a>
<i>Chlorobium phaeovibrioides</i> DSM 265	<a href="#">NC_009337</a>
<i>Chlorobium limicola</i> DSM 245	<a href="#">NC_010803</a>
<i>Chlorobium luteolum</i> DSM 273	<a href="#">NC_007512</a>
<i>Chloroherpeton thalassium</i> ATCC 35110	<a href="#">NC_011026</a>
<i>Prosthecochloris aestuarii</i> DSM 271	<a href="#">NC_011059</a>
<i>Pelodictyon phaeoclathratiforme</i> BU-1	<a href="#">NC_011060</a>

not shown). Therefore, we turned to conservation across *Chlorobiaceae* genomes, i.e., phylogenetic footprinting, to identify putative regulatory motifs associated with sulfur-regulated genes that are shared between *Cba. tepidum* and other *Chlorobiaceae* (Table 1).

Phylogenetic footprinting identified an unknown motif (putative sulfide operator sequence 1 [PSOS-1]) in the promoter of *CT1277* (Fig. 2A), encoding a putative transcriptional regulator that was highly upregulated on sulfide compared to thiosulfate (5) and  $S^0$  (Fig. 2). PSOS-1 was found associated with the pTSS for the two bona fide SQRs, *CT0117* and *CT1087*, via FIMO searches (6, 18). *CT0117* is a low-sulfide-concentration-adapted SQR, while *CT1087* is adapted for high sulfide concentrations (6). Phylogenetic footprinting recovered PSOS-1 motifs associated with *sqr* genes that encode sulfide:quinone oxidoreductases across the *Chlorobiaceae*, including all *CT1087* homologues (Fig. 2B) and a subset of *CT0117* homologues (Fig. 2C). A consensus motif for *Cba. tepidum* PSOS-1 sites was constructed from *CT1277*, *CT0117*, and *CT1087* sites (Fig. 2E) and was used as the seed in a FIMO search against the *Cba. tepidum* genome. In addition to returning the input loci, PSOS-1 was predicted to occur near 156 TSS. The only other TSS with a *q* value of  $<0.05$  was a pTSS for *CT0742* (Fig. 2D), a putative transmembrane protein in the TauE-like family of anion transporters. In *Neptuniibacter caesariensis*, TauE has been proposed as a sulfite transporter (22). Based on our TSS data, it appears that *CT0743* is cotranscribed with *CT0742*. The *CT0743* gene product is a hypothetical protein containing a domain of unknown function. *Chlorobium ferrooxidans* is a member of the *Chlorobiaceae* that cannot oxidize reduced sulfur compounds and instead oxidizes iron (23). PSOS-1 was searched against the *Chl. ferrooxidans* genome via FIMO, and no significant PSOS-1 positions were returned.

In *Cba. tepidum*, the putative *CT1277* regulator appears to be expressed from an RpoD pTSS (Fig. 2A). In the *CT1277* promoter, the PSOS-1 motif overlapped the +1 site and the first two base pairs of the -6 box. The positioning of PSOS-1 relative to the -6 box was conserved for all genomes analyzed except *Chlorobium phaeobacteroides* BS1, where it was found largely in the spacer sequences between the -6 and -33 boxes, with partial overlap of the -6 box (Fig. 2A). In *Chlorobium chlorochromatii*, the PSOS-1 motif was discovered in the promoter of *Cag\_0887*, a hypothetical protein encoded upstream of the *CT1277* homologue *Cag\_0886*; *Cag\_0887* to *Cag\_0885* are likely transcribed as a single unit. The *CT1087* and *CT0117* PSOS-1 motifs were positioned near the pTSS for each gene (Fig. 2B and C). The positioning of PSOS-1 within homologous *CT1087* promoters was variable, with PSOS-1 overlapping the -33 box of two of the promoters and the -6 box of the other two. PSOS-1 overlapped the -6 box and extended upstream into the spacer sequence of the six homologous *CT0117* promoters. However, in the *Chloroherpeton thalassium* promoter, PSOS-1 overlapped the -6 box and extended downstream, and in the *Chlorobium limicola* promoter, it overlapped the -33 box. Given that the PSOS-1 motif overlaps the RpoD binding motif or the +1 site or both, the data suggest that this motif binds a negative repressor, as occlusion of the RpoD binding site, or the +1 site, would likely interfere with transcription initiation. The

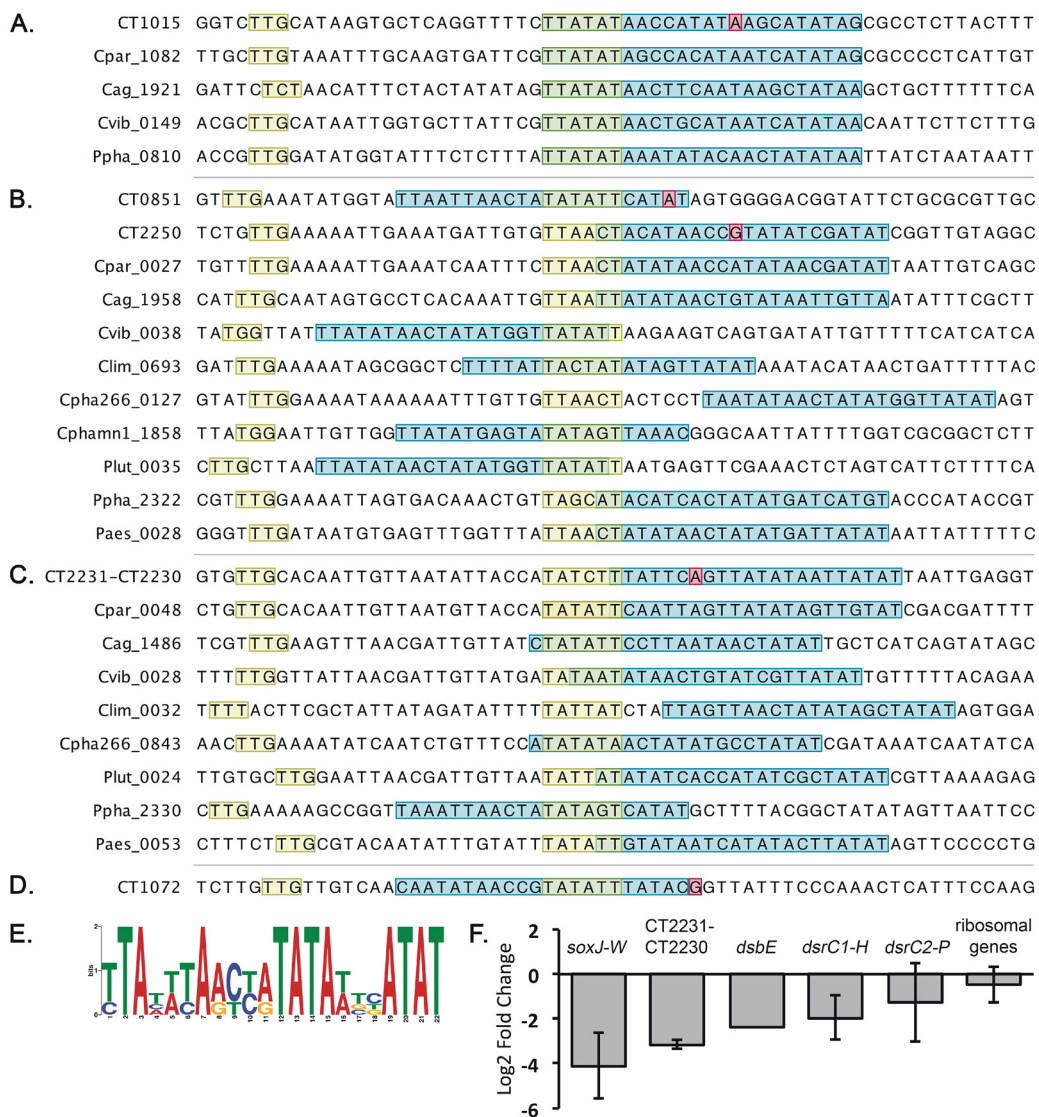


**FIG 2** Identification of putative sulfide operator sequence 1 (PSOS-1). Promoter regions for orthologs of *CT1277* (A), *CT1087* (B), and *CT0117* (C) were extracted from all *Chlorobiaceae* genomes. RpoD (yellow) and PSOS-1 (blue) motifs were discovered by promoter analysis. TSS mapped in *Cba. tepidum* are shown in pink. (D) The *Cba. tepidum* consensus motif was searched against the *Cba. tepidum* genome, which returned an additional TSS associated with PSOS-1. (E) The PSOS-1 consensus motif for the three *Cba. tepidum* sites. (F) Log<sub>2</sub> fold changes in transcript abundance on sulfide relative to S<sup>0</sup> of genes associated with PSOS-1 with ribosomal protein genes as a comparator as described by Eddie and Hanson (5).

PSOS-1 sequence appears to be unique in that it does not match any characterized motifs present in the CollecTF, Prodoric, and RegTransBase databases (24–27).

Expression levels of the components of the PSOS-1 regulon were variable between growth on sulfide and S<sup>0</sup> (Fig. 2F). *CT1277* and *CT1087* were both significantly upregulated on sulfide relative to thiosulfate and S<sup>0</sup>, while *CT0117* was downregulated on both electron donors relative to sulfide (Fig. 2F) (5). *CT0742* and *CT0743* did not change in expression significantly between the growth conditions and displayed expression profiles similar to those seen with ribosomal genes.

**PSOS-2 is associated with *sox*, *dsr*, and *CT2230* TSS.** As many genes related to thiosulfate oxidation (*sox*) and S<sup>0</sup> oxidation (*dsr* and *dsbE*) were found to be downregulated on sulfide relative to S<sup>0</sup> and thiosulfate (Fig. 3) (5), we searched the promoters of genes downregulated on sulfide for putative regulatory motifs. Only one RpoD TSS was observed in the *sox* operon immediately adjacent to *soxJ* (*CT1015*), supporting the assertion that the *sox* operon is transcribed as a single unit (7). Therefore, sequences upstream of the start codon of the first gene in the *sox* operon across the *Chlorobiaceae* were analyzed for



**FIG 3** Identification of putative sulfide operator sequence 2 (PSOS-2). (A to C) Promoter regions for orthologs of the *sox* operon (A), *dsrC* (B), and CT2230 (C) were extracted from all *Chlorobiaceae* genomes. RpoD (yellow) and PSOS-2 (blue) motifs were discovered by promoter analysis. TSS mapped in *Cba. tepidum* are shown in pink. (D) The *Cba. tepidum* consensus motif was searched against the *Cba. tepidum* genome, which returned additional TSS associated with PSOS-2. (E) The PSOS-2 consensus motif for the four *Cba. tepidum* sites. (F) Log<sub>2</sub> fold change in transcript abundance on sulfide relative to S<sup>0</sup> of genes associated with PSOS-2 with ribosomal protein genes as a comparator as described by Eddie and Hanson (5).

motif discovery. Phylogenetic footprinting identified an unknown motif (putative sulfur operator sequence 2 [PSOS-2]) in the promoters analyzed (Fig. 3A). The *sox* PSOS-2 motif was searched against the *Cba. tepidum* genome. The top two positions occurring near TSS, excluding the pTSS for *soxJ*, were pTSS for *dsrC-1* (CT0851) and CT2231. Phylogenetic footprinting of *dsrC-1* and *dsrC-2* (CT2250) and of orthologs across the *Chlorobiaceae* returned PSOS-2 sites in all *dsrC* promoters across the *Chlorobiaceae* (Fig. 3B). CT2231, encoding a hypothetical protein with no apparent homologues, appears to be cotranscribed with CT2230, encoding a putative outer membrane protein. As no TSS were found between these two genes, and as both are downregulated on sulfide relative to thiosulfate and S<sup>0</sup> (Fig. 3) (5), sequences upstream of CT2230 homologues across the *Chlorobiaceae* and the CT2231 promoter were analyzed for motif discovery. PSOS-2 sites were found in all promoters (Fig. 3C). A consensus motif for PSOS-2 was constructed from *soxJ*, *dsrC-1*, *dsrC-2*, and CT2231 sites (Fig. 3E). In addition to returning the input loci, PSOS-2 was predicted to

occur near 184 additional TSS. The only TSS with a  $q$  value of  $<0.05$  were the pTSS for *dsbE* (CT1072; Fig. 3D) and an aTSS for CT1231, encoding a transposase and annotated as having an internal deletion (data not shown). PSOS-2 was searched via FIMO against the *Chl. ferrooxidans* genome, and no significant positions were identified.

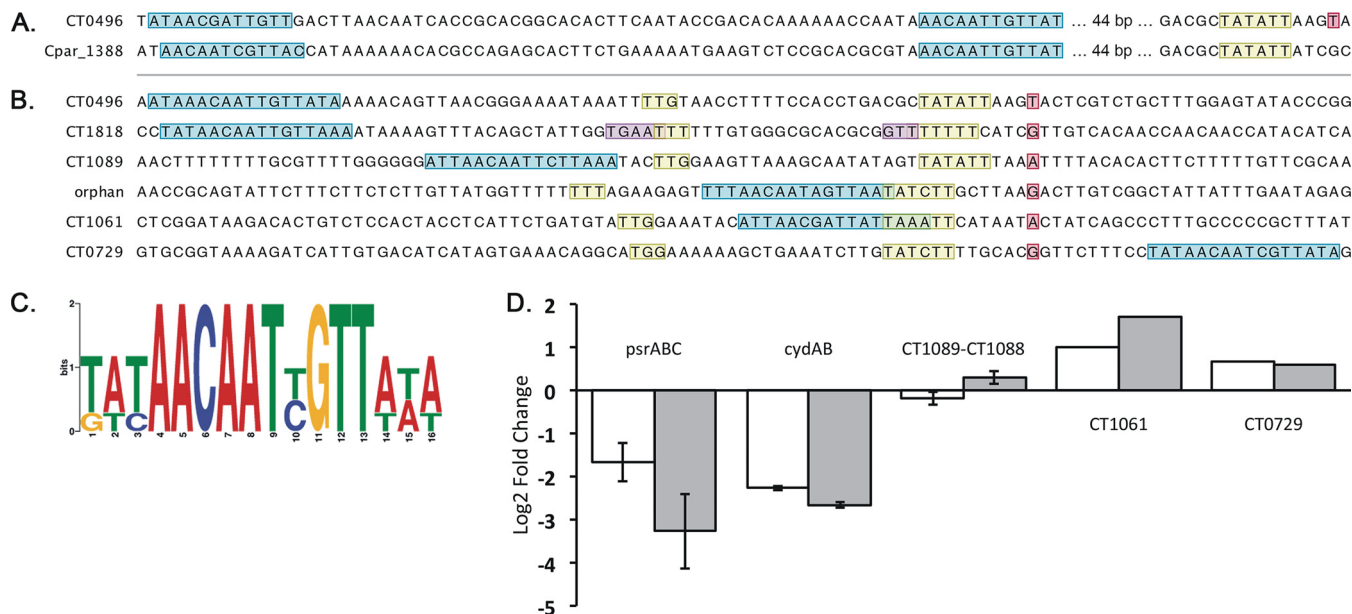
In *Cba. tepidum*, PSOS-2 was found to overlap the +1 site of the pTSS of *soxJ* and the -6 box of the RpoD motif predicted using the bulk TSS data (Fig. 3A). The positioning of PSOS-2 across *sox* promoters of the *Chlorobiaceae* was highly conserved, with all PSOS-2 sites found to overlap the -6 box and to extend downstream. PSOS-2 overlapped the +1 sites of both *dsrC-1* and *dsrC-2* and partially overlapped the -6 box of the RpoD motif (Fig. 3B). The position of PSOS-2 was variable across *dsrC* promoters, and yet it at least partially overlapped the -6 box in all promoters except that of *Chl. phaeobacteroides*, where it was found 6 bp downstream of the -6 box. PSOS-2 was found to overlap the +1 site of CT2231 and the last base pair of the -6 RpoD box (Fig. 3C). Positioning of PSOS-2 was fairly highly conserved across CT2230 promoters, with all but one site at least partially overlapping the -6 box. The PSOS-2 site in the *dsbE* promoter overlapped the -6 box and part of the RpoD spacer sequence (Fig. 3D). The PSOS-2 sequence appears to be unique in that it does not match any characterized motifs present in the CollecTF, Prodoric, and RegTransBase databases (data not shown) (24–27).

The components of the PSOS-2 regulon were downregulated on sulfide relative to  $S^0$  (Fig. 3F). The *sox* operon, CT2231–CT2230, *dsbE*, *dsrC1A1*, and *dsrC2* are significantly downregulated on sulfide relative to  $S^0$  and thiosulfate (Table S1) (5). The first *dsr* cluster, *dsrCABLEFH* (CT0851–CT0857), appears to be a single transcriptional unit under the control of the *dsrC1* RpoD pTSS. The second *dsr* cluster, *dsrNCABLUEFHMTKJOPVW* (CT2251–CT2238), appears to be broken into three transcriptional units: *dsrN*, *dsrC2-P*, and *dsrVW*. *dsrN* and *dsrC2-P* appear to be controlled by independent RpoD pTSS, while *dsrVW* appears to be controlled by an ECF factor pTSS (Table S3). This may explain why *dsrVW* is upregulated on sulfide relative to thiosulfate and  $S^0$  whereas the rest of the cluster is downregulated on sulfide relative to thiosulfate and  $S^0$  (Fig. 3F; Table S1) (5).

**A CRP-like motif (CLM) is associated with *psrABC*, *cydAB*, and CT0729.** The putative polysulfide oxidoreductase complex, *psrABC* (CT0496–CT0494), is significantly upregulated on sulfide compared to  $S^0$  and thiosulfate (Fig. 4) (5). Phylogenetic footprinting identified two occurrences of a small motif (cyclic AMP receptor protein [CRP]-like motif [CLM]) upstream of the RpoD binding site in the *Cba. tepidum* and *Cba. parvum* promoters (Fig. 4A). The *psrABC* CLM was searched against the *Cba. tepidum* genome for positions that occurred within 200 bp upstream or 50 bp downstream of a TSS. Aside from returning the two *psrABC* sites, FIMO returned positions near the pTSS for CT1818 and CT0729 as the only positions with  $q$  values of  $<0.05$ . CT1818 and CT1819 encode *cydAB*, a terminal oxidase cytochrome *bd* complex that confers sulfide-resistant  $O_2$ -dependent respiration in *Escherichia coli* (28). CT0729 encodes a Nudix hydrolase domain protein. A consensus motif was generated from the *psrABC*, *cydAB*, and CT0729 sites and was used to search the *Cba. tepidum* genome (Fig. 4C). Significant positions ( $q$  values =  $<0.05$ ) included pTSS for CT1089 and CT1061 and an oTSS at position 503855 (Fig. 4B). CT1089 and CT1088 encode the two subunits of ATP:citrate lyase. CT1061 is a hypothetical protein with a predicted steriodogenic acute regulatory protein-related lipid transfer (Start) domain found in polyketide cyclases and dehydrogenases. The sequence downstream of the oTSS was searched for open reading frames and for RNA families (29), but nothing of significance was found.

In *psrABC* promoters of *Cba. tepidum* and *Cba. parvum*, the first CLM site occurred 49 bp upstream of the RpoD -6 box, while the second site occurred 114 bp upstream of the *Cba. tepidum* -6 box and 113 bp upstream of the *Cba. parvum* -6 box (Fig. 4A). In *Cba. tepidum*, the RpoD -6 box occurs 4 bp upstream of the +1 site. The CLM site for *cydAB* occurred 57 bp upstream of the TSS, in a position similar to that of the first *psrABC* site, which occurred 58 bp upstream of the *psrABC* TSS (Fig. 4B). The CLM site for CT1089 occurred 35 bp upstream of the TSS and 4 bp upstream of the -33 box,





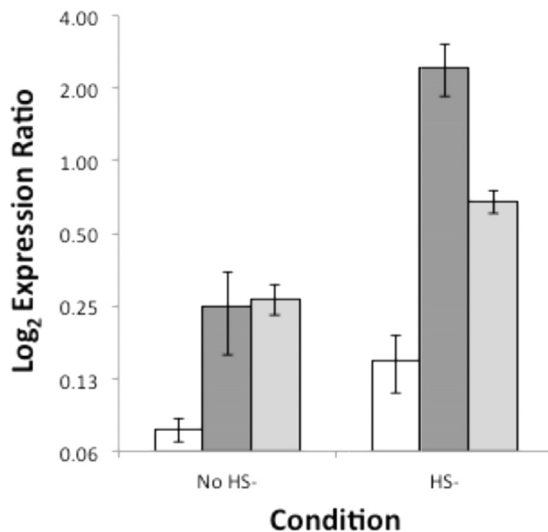
**FIG 4** Identification of the CRP-like motif (CLM). Promoter regions for orthologs of the *psrABC* operon were extracted from all *Chlorobiaceae* genomes. (A) RpoD (yellow), ECF sigma factor (purple), and CLM (blue) motifs were discovered by promoter analysis. TSS mapped in *Cba. tepidum* are shown in pink. (B) The *Cba. tepidum* consensus motif was searched against the *Cba. tepidum* genome, which returned additional TSS associated with CLM. (C) Consensus motif derived from *Cba. tepidum* CLM sites. (D) Log<sub>2</sub> fold change in transcript abundance of genes associated with CLM; S<sup>0</sup> relative to thiosulfate (white) and sulfide relative to S<sup>0</sup> (gray).

while both the *CT1061* and orphan sites overlapped the RpoD spacer sequence and  $-6$  box. The *CT0729* CLM site occurred 10 bp downstream of the *CT0729* TSS.

The CLM motif resembles the cyclic AMP receptor protein (CRP [TGTGA-N<sub>6</sub>-TCACA]) and the fumarate nitrate reduction protein (FNR [TTGAT-N<sub>4</sub>-ATCAA]) consensus binding motifs of *E. coli* without the spacer sequence (30, 31). The spacing of the motif sites (Fig. 4A) echoes that of CRP-activating promoters in *E. coli* with multiple CRP binding sites upstream of the RpoD binding site (30). In support of this, TOMTOM (27) matched the CLM consensus to the PrfA consensus motif from *Listeria monocytogenes*, a CRP-domain transcriptional activator. The PrfA consensus sequence, WTAACAWWTGTAA (32), does not contain the N<sub>4-6</sub> spacer present in *E. coli* CRP/FNR domain binding sites, adding further support for the idea that the *Cba. tepidum* CLM may bind a CRP domain protein. *Cba. tepidum* encodes a single CRP/FNR domain protein, CT1719. However, no gene encoding a CRP/FNR domain protein was found in the *Cba. parvum* genome or in any other *Chlorobiaceae* genome examined.

While the components of the CLM regulon had variable expression on S<sup>0</sup> relative to thiosulfate and sulfide, the directions of expression on S<sup>0</sup> relative to sulfide and thiosulfate were similar for all components; *psrABC* and *cydAB* were both significantly downregulated on S<sup>0</sup> relative to thiosulfate and sulfide, while *CT1061* and *CT0729* were upregulated on S<sup>0</sup> relative to thiosulfate and sulfide (Fig. 4D). The *CT1089-CT1088* operon was not differentially expressed under the different growth conditions.

**CT1277 is required for transcriptional repression of CT1087.** Previous data showed that the *CT1276-CT1277* cassette displayed increased transcript abundance in cells following sulfide addition after growth on thiosulfate (5). Given that *CT1277* belongs to the HTH-XRE family of transcriptional regulators, the *CT1277* gene was deleted from the *Cba. tepidum* genome to assess its role in sulfide-dependent gene regulation. Expression of *CT1087* was monitored by quantitative reverse transcriptase PCR (qRT-PCR) in the wild-type strain and in two independently isolated  $\Delta$ *CT1277* strains (Fig. 5). *CT1087* transcript abundance was significantly elevated (3.5-fold,  $P < 0.03$ ) in both  $\Delta$ *CT1277* strains compared to the wild type during growth on thiosulfate. Transition to growth on sulfide from thiosulfate resulted in a 1.9-fold increase ( $P = 0.04$ )



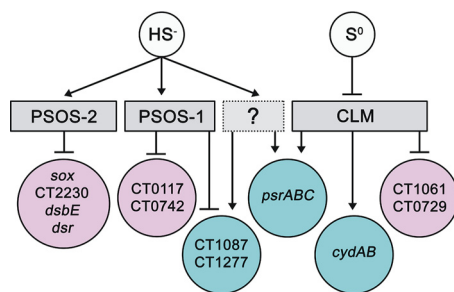
**FIG 5** *CT1277* is a transcriptional repressor of *CT1087*. qRT-PCR of *CT1087* in the *Cba. tepidum* wild-type strain (white), the  $\Delta CT1277.6$  mutant (dark gray), and the  $\Delta CT1277.11$  mutant (light gray) pre-sulfide addition and 40 min post-sulfide addition. Error bars are equal to 1 standard deviation. HS<sup>-</sup>, hydrogen sulfide.

in *CT1087* transcript abundance in the wild-type strain, confirming that its expression is sulfide dependent as previously reported (5, 6). Post-sulfide addition, *CT1087* displayed 16.3-fold ( $P = 0.003$ ) and 4.6-fold ( $P < 0.001$ ) increases in transcript abundance relative to the wild type in strains  $\Delta CT1277.6$  and  $\Delta CT1277.11$ , respectively. Thus, *CT1087* expression was induced much more strongly (9.7-fold;  $P = 0.003$ ) in the absence of *CT1277*. Increased expression of *CT1087* after sulfide addition in the  $\Delta CT1277$  strains suggests the presence of a sulfide-dependent activator. Sequences up to 2,000 bp upstream of the *CT1087*, *CT1277*, and *CT0117* promoters were analyzed for motif discovery, but no candidate activator motif was found. Altered expression of *CT1087* in the  $\Delta CT1277$  strains indicates that *CT1277* negatively regulates the transcription of *CT1087* in the presence of an activator. No significant growth phenotype was observed for the  $\Delta CT1277$  strain compared to the wild-type strain (data not shown).

## DISCUSSION

In this report, we provide global transcript abundance data for *Cba. tepidum* during growth on  $S^0$  as the sole electron donor and use these data to identify genes that were differentially expressed between growth on sulfide and growth on  $S^0$ . Many of these genes encode key components of the sulfur oxidation machinery of *Cba. tepidum*. The most dynamic changes in gene expression with respect to growth on different reduced sulfur compounds were in response to sulfide. We also provide a global transcript start site map for *Cba. tepidum*. dRNA-seq identified 3,426 putative TSS across the results of growth on sulfide, thiosulfate, and  $S^0$ , of which 1,086 were primary TSS. These data also include 71 orphan TSS that may control transcription of functional elements that were missed during genome annotation and provide evidence for antisense transcription in *Cba. tepidum*. Two basal promoter motifs were identified: an RpoD motif and an ECF sigma factor motif. Three putative regulatory motifs were discovered by phylogenetic footprint analysis of orthologous promoters across the *Chlorobiaceae* for genes that were differentially expressed during growth on reduced sulfur compounds in *Cba. tepidum*. Together, the data presented in this report provide a set of predictions for a mechanistic understanding of transcriptional regulation in different sulfur-dependent growth states in *Cba. tepidum* and in the members of the *Chlorobiaceae* as a whole.

TSS that are not associated with the RpoD motif or the ECF factor motif may not bind  $\sigma$  factors and therefore may not be bona fide TSS. In support of these possibilities, sequences ( $\pm 50$  bp) surrounding TSS associated with  $\sigma$  factor motifs were found to have significantly higher AT content than those sequences surrounding TSS without  $\sigma$



**FIG 6** Simplified model of the oxidative sulfur metabolism regulatory network in *Cba. tepidum* with nodes representing metabolites ( $\text{HS}^-$  and  $\text{S}^0$ ), putative regulatory signals (gray), genes downregulated on sulfide relative to  $\text{S}^0$  (pink), and genes upregulated on sulfide ( $\text{HS}^-$ ) relative to  $\text{S}^0$  (blue). Arrows represent activation, and bars represent repression.

factor motifs (data not shown). Alternatively, these TSS may be associated with divergent  $\sigma$  factor binding sites that are not represented by the motifs discovered in this study (Fig. 1). Of the 2,049 TSS that do not have  $\sigma$  factor motifs, 1,733 (84.6%) are iTSS and 319 (15.6%) are pTSS, whereas of the 1,377 TSS with  $\sigma$  factor motifs, 367 (26.7%) are iTSS and 767 (55.7%) are pTSS. iTSS are enriched in TSS with no associated  $\sigma$  factor motif, suggesting that most iTSS in this study are not bona fide TSS. This is in contrast to a recent study in *E. coli* that found that most iTSS were associated with  $\sigma$  factor motifs (33). Indeed, the fraction of iTSS detected in this study, 61% of all TSS, is higher than that detected in other organisms: 37% in *Escherichia coli*, 36% in *Campylobacter jejuni*, and 18% in *Helicobacter pylori* (33–35). pTSS and sTSS occur at similar percentages across these organisms, while *Cba. tepidum* has fewer asTSS (17%) than others: 43% in *E. coli*, 48% in *C. jejuni*, and 41% in *H. pylori*. These variations may reflect differences in how more closely related organisms in the *Gammaproteobacteria* and *Epsilonproteobacteria* regulate transcription relative to the *Chlorobiaceae*. Future experiments will determine if the high percentage of iTSS in *Cba. tepidum* is characteristic of all *Chlorobiaceae* and if there is any transcriptional activity associated with iTSS-associated sequences that lack a recognizable sigma factor motif in *Cba. tepidum*.

A putative operator sequence, PSOS-1, was discovered in the promoters of the two SQRs CT0117 and CT1087, as well as in the promoters of the putative regulatory protein CT1277 and the TauE-domain protein CT0742 (Fig. 2). While CT1087 and CT1277 are highly upregulated on sulfide relative to  $\text{S}^0$  (Fig. 2) and thiosulfate (5), CT0117 is downregulated, and CT0742 and CT0743 are not differentially expressed. The expression data corresponding to growth on sulfide versus growth on  $\text{S}^0$  corroborate the results of previous studies that have observed downregulation of CT0117, and upregulation of CT1087 and CT1277, in response to high levels of sulfide (5, 6). Here, CT1277 was shown to negatively regulate CT1087 expression on thiosulfate and sulfide (Fig. 5). Therefore, PSOS-1 may represent the CT1277 binding motif, which would suggest that CT1277 is subject to autorepression. If PSOS-1 binds CT1277 and functions as an operator sequence, then promoters associated with PSOS-1 should be downregulated in response to sulfide due to increased CT1277 expression, as increased repressor concentrations would be expected to lead to increased repression of those loci; this is the case for CT0117 (Fig. 2). However, CT1087 expression increased in response to sulfide in the  $\Delta\text{CT1277}$  background, suggesting the presence of a sulfide-dependent activator (Fig. 5). As CT1277 is associated with PSOS-1, and shows an expression profile similar to that of CT1087, the results suggest that the activator acting on CT1087 also activates CT1277. Thus, whatever element activates CT1087 and CT1277 expression should be absent from the CT0117 promoter. This leads to a model where sulfide activates a sulfide-dependent activator and indirectly activates the PSOS-1-binding protein (CT1277) via the inferred activator (Fig. 6). CT1277 represses transcription of CT0117, CT1087, and CT1277, while the sulfide-dependent activator activates transcription of CT1087 and CT1277 but not that of CT0117. This may occur to fine tune expression of CT1277 and CT1087.

The proposed models for CT1277 function are based on its association with the members of the HTH\_XRE protein superfamily (NCBI accession no. [c122854](#)) that are thought to be DNA binding transcriptional regulators. The notion that CT1277 and orthologs bind DNA was supported by a secondary structure prediction that identified three helical regions located at residue 36 to residue 46, residue 53 to residue 58, and residue 63 to residue 82 (see Fig. S2 in the supplemental material). These regions are annotated in multiple databases as homologous to sigma factor DNA binding regions, i.e., InterPro homologous family IPR013324-RNA polymerase sigma factor, region 3/4, and TIGR02937-sigma70-ECF (RNA polymerase sigma factor, sigma-70 family). DNA binding of these regions is conferred by a three-helix bundle as predicted for CT1277. The protein sequence alignment of *Chlorobiaceae* CT1277 orthologs also revealed six conserved cysteine residues (Fig. S2) (36). Four cysteine residues (Cys-105, Cys-108, Cys-124, and Cys-127) occur in -CXXC- motifs often associated with FeS cluster binding, while Cys-7 and Cys-151 are isolated. We hypothesize that these conserved cysteine residues may interact with sulfide, possibly affecting the formation or stability of an FeS cluster, giving the protein a mechanism for sulfide- or redox-mediated activation.

For several of the key components of sulfide oxidation that are downregulated in response to sulfide, a putative operator sequence, PSOS-2, was discovered overlapping the TSS and/or the RpoD -6 box (Fig. 3). The positioning of this motif and the expression patterns of the associated transcriptional units strongly suggest that this motif functions as an operator sequence. As these genes are downregulated in response to sulfide relative to thiosulfate (5), and appear to also be downregulated in response to sulfide on  $S^0$  (Fig. 3), PSOS-2 may bind a sulfide-dependent repressor (Fig. 6). CT2230 is predicted to be a membrane transporter in the FadL family of outer membrane proteins and has been characterized in long-chain fatty acid transport in *E. coli* (37). As CT2230 is involved in the transport of hydrophobic molecules, and is predicted to be part of a sulfide-repressed regulon, it may be involved in transport of polysulfide or hydrophobic sulfur chains across the outer membrane.

The *psrABC* promoter was found to contain two CLM sites, with single CLM sites occurring near the TSS of *cydAB*, *CT1089-CT1088*, *CT1061*, and *CT0729* (Fig. 4). In *E. coli*, CRP can function as both a repressor and an activator depending on its positioning relative to the TSS and sigma factor binding site (30). *psrABC* and *cydAB* displayed similar expression profiles in that they were downregulated on  $S^0$  relative to thiosulfate and sulfide (Fig. 4D). However, *CT1089-CT1088* did not change significantly in expression between growth conditions, and *CT1061* and *CT0729* are both upregulated on  $S^0$  relative to sulfide. These expression patterns can be explained if the location of the CLM site is taken into account (Fig. 4A and B) and if  $S^0$  inhibits the activity of the CLM-binding protein. The CLM sequence likely functions as an activator for *psrABC* and *cydAB*, while it likely functions as a repressor for *CT1061*, the oTSS, and *CT0729* in that it interferes with sigma factor binding or transcription elongation. The CLM site may not have a large effect on *CT1089-CT1088* and could function as either a repressor or an activator (30). Thus, if  $S^0$  inhibits the function of the CLM, it would repress *psrABC* and *cydAB*, while activating *CT1061* and *CT0729* (Fig. 6).

Between the transcript abundance data (this study and reference 5) and the data from the three putative regulatory motifs discovered in sulfur-regulated genes, it appears that sulfide likely acts as a master regulator, activating proteins that bind PSOS-1, where CT1277 is the most likely candidate binding protein, PSOS-2, and the inferred but unidentified sulfide-dependent activator acting on *CT1087* (Fig. 6). If  $S^0$  inhibits the activity of the CLM binding protein, where CT1719 is the most likely candidate binding protein, then the expression pattern of CLM-associated genes can be explained. As the three motifs identified here were found to be conserved between multiple *Chlorobiaceae* genomes, it suggests that the regulatory functions carried out by these motifs are conserved across genomes. Thus, we have identified genes important for growth on  $S^0$  relative to sulfide and thiosulfate, globally mapped TSS that were active during growth on these electron donors, identified basal promoter motifs, and discovered putative sulfur regulatory motifs. From these data, a model has emerged

whereby sulfide is a master regulator of the metabolic state of *Cba. tepidum*, inducing the repression of genes important for thiosulfate and  $S^0$  oxidation and the activation of several key components for sulfide oxidation.

## MATERIALS AND METHODS

**RNA sequencing.** Cultures were grown under conditions of  $20 \mu\text{mol photon m}^{-2} \text{s}^{-1}$  photosynthetically active radiation (PAR) in Pf-7 medium (6) at  $47^\circ\text{C}$  with thiosulfate as the sole electron donor or at  $42^\circ\text{C}$  with biogenic  $S^0$  as the sole electron donor. RNA was extracted using a NucleoSpin RNA kit (Macherey-Nagel), and rRNA was depleted using a MicroExpress kit (Ambion) and treated with a TURBO DNA-free kit (Ambion) to remove residual genomic DNA (gDNA). RNA-seq libraries were constructed using a NuGEN Ovation kit (NuGEN) to convert 25 ng RNA to double-stranded cDNA, with subsequent fragmentation to approximately 100-bp fragments performed using an S2 Adaptive Focused Acoustic Disruptor (Covaris). Fragment size was confirmed by agarose gel electrophoresis on a 2% gel. A NuGEN Encore kit (NuGEN) was used to ligate adapters suitable for Illumina sequencing to the ends of these fragments. Libraries were sequenced on an Illumina HiSeq 1000 sequencer at the University of Delaware Sequencing and Genotyping Center.

**RNA-seq analysis.** Reads were aligned to the *Cba. tepidum* genome using the Eland pipeline (Illumina). Custom Perl scripts (available upon request) were used to calculate the number of sequences that mapped to each annotated gene, noncoding RNA, and intergenic region of more than 50 bp, correcting for the length of the region. Data were normalized using quantile normalization as previously described (5). Levels of expression that differed for each pair of libraries were calculated using DESeq (38).

We previously reported the transcriptional response of *Cba. tepidum* to sulfide addition after growth on thiosulfate (5). The sulfide and  $S^0$  expression libraries were constructed with different kits and were sequenced at different times, so, in order to compare expression of *Cba. tepidum* on  $S^0$  to that on sulfide, fold changes on  $S^0$  relative to thiosulfate were divided by fold changes on sulfide relative to thiosulfate, giving a fold change value for each gene on sulfide relative to  $S^0$ . JMP Pro (version 12.1.0) was used to create a box-and-whisker plot for the distribution of the fold change of growth on sulfide relative to growth on  $S^0$ . Outliers, or those points outside  $1.5\times$  the interquartile range from the mean, were called as differentially expressed genes (see Table S1 in the supplemental material).

**Differential RNA sequencing.** dRNA-seq libraries were constructed from four independent cultures grown on thiosulfate and on biogenic  $S^0$  and from those grown on thiosulfate, spiked with 1.6 mM sulfide, and harvested 30 min after sulfide addition. *Cba. tepidum* cultures were grown as previously described by Levy et al. (39) at  $20 \mu\text{mol photon m}^{-2} \text{s}^{-1}$  PAR. Replicate 1.5-ml cell pellets were harvested by centrifugation after 20 h of growth (mid-log phase). Cell pellets were flash frozen in liquid nitrogen and were stored at  $-70^\circ\text{C}$ . Cell pellets were thawed on ice, resuspended in  $100 \mu\text{l}$  Tris-EDTA (TE) buffer (pH 8.0) with  $1 \mu\text{l}$  Ready-Lyse lysozyme (Epicentre) ( $20 \text{ KU}/\mu\text{l}$ ), and incubated at room temperature for 30 min. Two replicate cell pellets were combined, and RNA was purified by the use of a NucleoSpin RNA kit (Macherey-Nagel). A  $10\text{-}\mu\text{g}$  volume of RNA was treated with a Turbo DNA-free kit (Thermo Fisher) for gDNA removal and was then concentrated over RNA Clean & Concentrator-25 columns (Zymo Research). rRNA was depleted via the use of a MicroExpress kit (Ambion), and samples were concentrated over Zymo-25 columns. Libraries were normalized by *sigA* copy number (5) prior to terminator 5' phosphate-dependent exonuclease (TEX; Epicentre) treatment. Each replicate was split into two samples for TEX treatment: the first was treated with 1 U TEX in a  $20\text{-}\mu\text{l}$  reaction mixture, while the second was incubated in the same buffer without TEX (35). Reactions were cleaned over RNA Clean & Concentrator-5 columns (Zymo Research). Samples were subsequently treated with 2 U tobacco acid pyrophosphatase (TAP; Epicentre) in a  $50\text{-}\mu\text{l}$  total reaction volume followed by RNA concentration on Zymo-5 columns. 5' adapters from a NEBNext Multiplex Small RNA Library Prep Set for Illumina kit (NEB) were ligated to the treated RNA samples using  $2.5 \mu\text{l}$  of ligation enzyme mix in a  $20\text{-}\mu\text{l}$  total reaction volume with final concentrations of 1 mM ATP and 12.5% polyethylene glycol 8000 (PEG 8000). RNA was subsequently fragmented for 2 min using a NEBNext Magnesium RNA fragmentation module (NEB) and was cleaned using Agencourt RNAClean XP beads (Beckman Coulter). 3' ends were repaired via calf intestinal alkaline phosphatase (NEB) treatment and were cleaned over Zymo-5 columns. 3' adapter ligation, first strand synthesis with indexed primers, and cDNA amplification (15 cycles) were completed using a NEBNext small RNA kit. cDNA reactions were purified via the use of Agencourt AMPure XP beads (Beckman Coulter). BluePippin was used for size selection (150 to 600 bp) prior to Illumina sequencing performed over two lanes (12 libraries pooled per lane) on a HiSeq 2000 sequencer at the University of Delaware Sequencing and Genotyping Center.

**dRNA-seq analysis.** The RNA-seq analysis pipeline READemption (version 0.3.9) was used to align reads to the *Cba. tepidum* genome, and the output was passed to TSSpredator (version 1.06) for TSS identification (34, 40). In order for a TSS to be called as active under a specific condition, it needed to be detected in at least three of four replicates with an enrichment score of  $\geq 2$  and to meet the remaining parameters specified under the "very specific" setting. TSS were classified according to the definitions of Sharma et al. (35). Primary TSS (pTSS) are defined as representing TSS within 300 bp upstream of an annotated gene. If multiple TSS are present in this range, the pTSS is that with the highest enrichment value, and the others are secondary TSS (sTSS). Antisense TSS (asTSS) occur on the antisense strand internal to or within 100 bp downstream of an annotated feature. Internal TSS (iTSS) occur on the sense strand within an annotated gene. Orphan TSS (oTSS) do not fall into any of the other defined classes.

**Motif discovery and analysis.** Basal promoter motifs were identified by analyzing 50 bp upstream of all TSS by the use of the MEME software suite (8). The outputs from multiple MEME runs using various parameter settings were pooled via the use of a custom R script (available upon request) to obtain

estimates for the number of TSS associated with each motif. Promoter regions for orthologs of genes that are strongly regulated by sulfide in *Cba. tepidum* were extracted from all *Chlorobiaceae* genomes (Table 1). Sequences of up to 1,000 bp upstream of the start codon for each orthologous gene set were analyzed by MEME and/or DMINDA (8, 41). Motifs identified as described above were searched against the *Cba. tepidum* genome using FIMO to identify additional occurrences (18).

**Deletion mutagenesis.** *CT1277* was deleted from the *Cba. tepidum* genome using a counterselectable suicide vector that is to be fully described elsewhere (J. M. Hilzinger and T. E. Hanson, unpublished data). Briefly, a PCR product containing flanking DNA upstream and downstream of *CT1277* was cloned into a mobilizable suicide vector with both antibiotic resistance and a counterselectable marker based on a vector used for gene deletions in *Shewanella oneidensis* MR-1 (42). The primers used for amplifying the flanking DNA were as follows: CT1277-5'UTR-F-EcoRV (GATATCTCATCCCACCTATGAGCA), CT1277-5'UTR-R (GTCGAGCGGAATGATGATCTTCATTCTAGATTTTTTGTAGTCAGCAATT), CT1277-3'UTR-F (AATTGCTGACTAAAAAATCTAGATGAAGATCATCATTCCGCTCGAC), and CT1277-3'UTR-R-EcoRV (GATATCGTCGAGTTGTTGAGGATAC).

**Response to sulfide qRT-PCR.** *Cba. tepidum* wild-type and  $\Delta CT1277$  strains were grown in Pf-7 medium at 47°C with 20  $\mu\text{mol photon m}^{-2} \text{s}^{-1}$  PAR. The absence of sulfide from cultures was verified by testing with  $\text{CuCl}_2$  (2). Presulfide biomass was pelleted by centrifugation, flash frozen in liquid nitrogen, and stored at  $-70^\circ\text{C}$ . Sulfide was added to a final concentration of 2 mM, and cultures were incubated at 47°C with 20  $\mu\text{mol photon m}^{-2} \text{s}^{-1}$  PAR for 40 min. Postsulfide biomass was pelleted, flash frozen, and stored at  $-70^\circ\text{C}$ . RNA was extracted and purged of residual genomic DNA as described for dRNA-seq. *sigA* and *CT1087* mRNAs were reverse transcribed into cDNA using the SigA-R-RT and CT1087-R-RT primers (6) and a ProtoScript II cDNA synthesis kit (NEB) in the same reaction mixture. Negative controls lacking reverse transcriptase were performed to detect the presence of gDNA. The expression levels of *sigA* and *CT1087* were determined using RealMasterMix SYBR ROX (5 Prime) and an ABI 7500 Fast real-time PCR system (Applied Biosystems). Genomic DNA standards were used to determine the efficiency of the SigA-RT and CT1087-RT primer sets and to quantify transcript abundance. *CT1087* levels were normalized using *sigA* expression levels.

**Data availability.** Data described in this paper have been deposited in the National Center for Biotechnology Information Sequence Read Archive affiliated with BioSample accession numbers SAMN07413950 for  $\text{S}^0$  and thiosulfate RNA-seq data and SAMN07413841 for all dRNA-seq data.

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/AEM.01966-17>.

**SUPPLEMENTAL FILE 1**, PDF file, 4.2 MB.

**SUPPLEMENTAL FILE 2**, XLSX file, 0.4 MB.

**SUPPLEMENTAL FILE 3**, XLSX file, 0.1 MB.

**SUPPLEMENTAL FILE 4**, XLSX file, 1.2 MB.

## ACKNOWLEDGMENTS

We thank Katie Kalis and Amalie Levy for helpful discussions during data analysis and manuscript preparation and the University of Delaware Sequencing and Genotyping Center staff for advice and help with library construction and sequencing.

This work was supported by NSF grant MCB-1244373 to T.E.H. and an IGERT SBE2 fellowship to J.M.H. This project utilized computational resources at the University of Delaware Center for Bioinformatics and Computational Biology Core Facility funded by Delaware INBRE (NIGMS GM103446), Delaware EPSCoR (NSF EPS-0814251, NSF IIA-1330446), the State of Delaware, and the Delaware Biotechnology Institute.

## REFERENCES

- Chan LK, Weber TS, Morgan-Kiss RM, Hanson TE. 2008. A genomic region required for phototrophic thiosulfate oxidation in the green sulfur bacterium *Chlorobaculum tepidum* (syn. *Chlorobaculum tepidum*). *Microbiology* 154:818–829.
- Hanson TE, Bonsu E, Tuerk A, Marnocha CL, Powell DH, Chan CS. 2016. *Chlorobaculum tepidum* growth on biogenic  $\text{S}^0$  as the sole photosynthetic electron donor. *Environ Microbiol* 18:2856–2867. <https://doi.org/10.1111/1462-2920.12995>.
- Marnocha CL, Levy AT, Powell DH, Hanson TE, Chan CS. 2016. Mechanisms of extracellular  $\text{S}^0$  globule production and degradation in *Chlorobaculum tepidum* via dynamic cell-globule interactions. *Microbiology* 162:1125–1134. <https://doi.org/10.1099/mic.0.000294>.
- Eisen JA, Nelson KE, Paulsen IT, Heidelberg JF, Wu M, Dodson RJ, Deboy R, Gwinn ML, Nelson WC, Haft DH, Hickey EK, Peterson JD, Durkin AS, Kolonay JL, Yang F, Holt I, Umayam LA, Mason T, Brenner M, Shea TP, Parksey D, Nierman WC, Feldblyum TV, Hansen CL, Craven MB, Radune D, Vamathevan J, Khouri H, White O, Gruber TM, Ketchum KA, Venter JC, Tettelin H, Bryant DA, Fraser CM. 2002. The complete genome sequence of *Chlorobaculum tepidum* TLS, a photosynthetic, anaerobic, green-sulfur bacterium. *Proc Natl Acad Sci U S A* 99:9509–9514. <https://doi.org/10.1073/pnas.132181499>.
- Eddie BJ, Hanson TE. 2013. *Chlorobaculum tepidum* TLS displays a complex transcriptional response to sulfide addition. *J Bacteriol* 195:399–408. <https://doi.org/10.1128/JB.01342-12>.
- Chan LK, Morgan-Kiss RM, Hanson TE. 2009. Functional analysis of three sulfide:quinone oxidoreductase homologs in *Chlorobaculum tepidum*. *J Bacteriol* 191:1026–1034. <https://doi.org/10.1128/JB.01154-08>.
- Gregersen LH, Bryant DA, Frigaard N. 2011. Mechanisms and evolution of oxidative sulfur metabolism in green sulfur bacteria. *Front Microbiol* 2:116. <https://doi.org/10.3389/fmicb.2011.00116>.

8. Bailey TL, Williams N, Misleh C, Li WW. 2006. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res* 34(Web Server issue):W369–W373. <https://doi.org/10.1093/nar/gkl198>.
9. Chen S, Bagdasarian M, Kaufman MG, Bates AK, Walker ED. 2007. Mutational analysis of the *ompA* promoter from *Flavobacterium johnsoniae*. *J Bacteriol* 189:5108–5118. <https://doi.org/10.1128/JB.00401-07>.
10. Vingadassalom D, Kolb A, Mayer C, Rybkine T, Collatz E, Podglajen I. 2005. An unusual primary sigma factor in the Bacteroidetes phylum. *Mol Microbiol* 56:888–902. <https://doi.org/10.1111/j.1365-2958.2005.04590.x>.
11. Helmann JD. 2002. The extracytoplasmic function (ECF) sigma factors. *Adv Microb Physiol* 46:47–110. [https://doi.org/10.1016/S0065-2911\(02\)46002-X](https://doi.org/10.1016/S0065-2911(02)46002-X).
12. Manganello R, Voskuil MI, Schoolnik GK, Smith I. 2001. The *Mycobacterium tuberculosis* ECF sigma factor  $\sigma$  E: role in global gene expression and survival in macrophages. *Mol Microbiol* 41:423–437. <https://doi.org/10.1046/j.1365-2958.2001.02525.x>.
13. Manganello R, Voskuil MI, Schoolnik GK, Dubnau E, Gomez M, Smith I. 2002. Role of the extracytoplasmic-function  $\sigma$  Factor  $\sigma$ H in *Mycobacterium tuberculosis* global gene expression. *Mol Microbiol* 45:365–374. <https://doi.org/10.1046/j.1365-2958.2002.03005.x>.
14. Paget MSB, Molle V, Cohen G, Aharonowitz Y, Buttner MJ. 2001. Defining the disulphide stress response in *Streptomyces coelicolor* A3(2): identification of the  $\sigma$ R regulon. *Mol Microbiol* 42:1007–1020. <https://doi.org/10.1046/j.1365-2958.2001.02675.x>.
15. Qiu J, Helmann JD. 2001. The  $\sigma$ 10 region is a key promoter specificity determinant for the *Bacillus subtilis* extracytoplasmic-function  $\sigma$  factors  $\sigma$ X and  $\sigma$ W. *J Bacteriol* 183:1921–1927. <https://doi.org/10.1128/JB.183.6.1921-1927.2001>.
16. Reitzer LJ, Movsas B, Magasanik B. 1989. Activation of *glnA* transcription by nitrogen regulator I (NRI)-phosphate in *Escherichia coli*: Evidence for a long-range physical interaction between NRI-phosphate and RNA polymerase. *J Bacteriol* 171:5512–5522. <https://doi.org/10.1128/jb.171.10.5512-5522.1989>.
17. Wösten MM. 1998. Eubacterial sigma-factors. *FEMS Microbiol Rev* 22: 127–150. <https://doi.org/10.1111/j.1574-6976.1998.tb00364.x>.
18. Grant CE, Bailey TL, Noble WS. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27:1017–1018. <https://doi.org/10.1093/bioinformatics/btr064>.
19. Gruber TM, Bryant DA. 1998. Characterization of the group 1 and group 2 sigma factors of the green sulfur bacterium *Chlorobium tepidum* and the green non-sulfur bacterium *Chloroflexus aurantiacus*. *Arch Microbiol* 170:285–296. <https://doi.org/10.1007/s002030050644>.
20. Chung S, Frank G, Zuber H, Bryant DA. 1994. Genes encoding two chlorosome components from the green sulfur bacteria *Chlorobium vibrioforme* strain 8327D and *Chlorobium tepidum*. *Photosynth Res* 41: 261–275. <https://doi.org/10.1007/BF02184167>.
21. Crooks GE, Hon G, Chandonia J, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Res* 14:1188–1190. <https://doi.org/10.1101/gr.849004>.
22. Weinitschke S, Denger K, Cook AM, Smits THM. 2007. The DUF81 protein TauE in *Cupriavidus necator* H16, a sulfite exporter in the metabolism of C2 sulfonates. *Microbiology* 153:3055–3060. <https://doi.org/10.1099/mic.0.2007/009845-0>.
23. Heising S, Richter L, Ludwig W, Schink B. 1999. *Chlorobium ferrooxidans* sp. nov., a phototrophic green sulfur bacterium that oxidizes ferrous iron in coculture with a “Geospirillum” sp. strain. *Arch Microbiol* 172:116–124. <https://doi.org/10.1007/s002030050748>.
24. Kiliç S, White ER, Sagitova DM, Cornish JP, Erill I. 2014. CollecTF: a database of experimentally validated transcription factor-binding sites in Bacteria. *Nucleic Acids Res* 42(Database issue):D156–D160. <https://doi.org/10.1093/nar/gkt1123>.
25. Grote A, Klein J, Retter I, Haddad I, Behling S, Bunk B, Biegler I, Yarmolinetz S, Jahn D, Münch R. 2009. PRODORIC (release 2009): a database and tool platform for the analysis of gene regulation in prokaryotes. *Nucleic Acids Res* 37(Database issue):D61–D65. <https://doi.org/10.1093/nar/gkn837>.
26. Cipriano MJ, Novichkov PN, Kazakov AE, Rodionov DA, Arkin AP, Gelfand MS, Dubchak I. 2013. RegTransBase—a database of regulatory sequences and interactions based on literature: a resource for investigating transcriptional regulation in prokaryotes. *BMC Genomics* 14:213. <https://doi.org/10.1186/1471-2164-14-213>.
27. Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. 2007. Quantifying similarity between motifs. *Genome Biol* 8:R24. <https://doi.org/10.1186/gb-2007-8-2-r24>.
28. Forte E, Borisov VB, Falabella M, Colaço HG, Tinajero-Trejo M, Poole RK, Vicente JB, Sarti P, Giuffrè A. 2016. The terminal oxidase cytochrome bd promotes sulfide-resistant bacterial respiration and growth. *Sci Rep* 6:23788. <https://doi.org/10.1038/srep23788>.
29. Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR, Floden EW, Gardner PP, Jones TA, Tate J, Finn RD. 2015. Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res* 43:D130–D137. <https://doi.org/10.1093/nar/gku1063>.
30. Kolb A, Busby S, Buc H, Garges S, Adhya S. 1993. Transcriptional regulation by cAMP and its receptors. *Annu Rev Biochem* 62:749–795. <https://doi.org/10.1146/annurev.bi.62.070193.003533>.
31. Spiro S. 1994. The FNR family of transcriptional regulators. *Antonie Van Leeuwenhoek* 66:23–36. <https://doi.org/10.1007/BF00871630>.
32. Novichkov PS, Kazakov AE, Ravcheev DA, Leyn SA, Kovaleva GY, Sultormin RA, Kazanov MD, Riehl W, Arkin AP, Dubchak I, Rodionov DA. 2013. RegPrecise 3.0—a resource for genome-scale exploration of transcriptional regulation in bacteria. *BMC Genomics* 14:745. <https://doi.org/10.1186/1471-2164-14-745>.
33. Thomason MK, Bischler T, Eisenbart SK, Förstner KU, Zhang A, Herbig A, Nieselt K, Sharma CM, Storz G. 2015. Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in *Escherichia coli*. *J Bacteriol* 197:18–28. <https://doi.org/10.1128/JB.02096-14>.
34. Dugar G, Herbig A, Förstner KU, Heidrich N, Reinhardt R, Nieselt K, Sharma CM. 2013. High-resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates. *PLoS Genet* 9:e1003495. <https://doi.org/10.1371/journal.pgen.1003495>.
35. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, Chabas S, Reiche K, Hackermüller J, Reinhardt R, Stadler PF, Vogel J. 2010. The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* 464:250–255. <https://doi.org/10.1038/nature08756>.
36. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>.
37. Nunn WD, Simons RW. 1978. Transport of long-chain fatty acids by *Escherichia coli*: mapping and characterization of mutants in the *fadL* gene. *Proc Natl Acad Sci U S A* 75:3377–3381. <https://doi.org/10.1073/pnas.75.7.3377>.
38. Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biol* 11:R106. <https://doi.org/10.1186/gb-2010-11-10-r106>.
39. Levy AT, Lee KH, Hanson TE. 2016. *Chlorobaculum tepidum* modulates amino acid composition in response to energy availability, as revealed by a systematic exploration of the energy landscape of phototrophic sulfur oxidation. *Appl Environ Microbiol* 82:6431–6439.
40. Förstner KU, Vogel J, Sharma CM. 2014. READemption - a tool for the computational analysis of deep-sequencing-based transcriptome data. *Bioinformatics* 30:3421–3423. <https://doi.org/10.1093/bioinformatics/btu533>.
41. Ma Q, Zhang H, Mao X, Zhou C, Liu B, Chen X, Xu Y. 2014. DMINDA: an integrated Web server for DNA motif identification and analyses. *Nucleic Acids Res* 42(Web Server issue):W12–W19. <https://doi.org/10.1093/nar/gku315>.
42. Burns JL, DiChristina TJ. 2009. Anaerobic respiration of elemental sulfur and thiosulfate by *Shewanella oneidensis* MR-1 requires *psrA*, a homolog of the *phsA* gene of *Salmonella enterica* serovar Typhimurium LT2. *Appl Environ Microbiol* 75:5209–5217. <https://doi.org/10.1128/AEM.00888-09>.