AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology®

RESEARCH ARTICLE

Identification of the Preferred DNA-Binding Sequence and Transcription Regulatory Network for the Thermophilic Zinc Uptake Regulator TTHA1292

John K. Barrows,[a] Alaina B. Westee,[a] Arianna Y. Parrish,[a] Michael W. Van Dyke[a]

[a]Department of Chemistry and Biochemistry, Kennesaw State University, Kennesaw, Georgia, USA

ABSTRACT D-block metal cations are essential for most biological processes; however, excessive metal exposure can be deleterious to the survival of microorganisms. To tightly control heavy metal regulation, prokaryotic organisms have developed several mechanisms to sense and adapt to changes in intracellular and extracellular metal concentrations. The ferric uptake regulator superfamily of transcription factors associates with DNA when complexed with a regulatory metal cofactor and often represses the transcription of genes involved in metal transport, thus providing a genomic response to an environmental stressor. Although extensively studied in mesothermic organisms, there is little information describing ferric uptake regulator homologs in thermophiles. In this study, we biochemically characterize the ferric uptake regulator homolog TTHA1292 in the extreme thermophile *Thermus thermophilus* HB8. We identify the preferred DNA-binding sequence of TTHA1292 using the combinatorial approach, restriction endonuclease, protection, selection, and amplification (REPSA). We map this sequence to the *Thermus thermophilus* HB8 genome and identify the TTHA1292 transcription regulatory network, which includes the zinc ABC transporter subunit genes *TTHA0596* and *TTHA0453/4*. We formally implicate TTHA1292 as a zinc uptake regulator and show that zinc coordination is critical for the multimerization of TTHA1292 dimers on DNA *in vitro* and transcription repression *in vivo*.

IMPORTANCE Discovering how organisms sense and adapt to their environments is paramount to understanding biology. Thermophilic organisms have adapted to survive at elevated temperatures (>50°C); however, our understanding of how these organisms adapt to changes in their environment is limited. In this study, we identify a zinc uptake regulator in the extreme thermophile *Thermus thermophilus* HB8 that provides a genomic response to fluctuations in zinc availability. These results provide insights into thermophile biology, as well as the zinc uptake regulator family of proteins.

KEYWORDS metalloregulation, *Thermus thermophilus*, transcriptional regulation, zinc uptake regulator

For prokaryotic organisms, regulation of protein expression occurs predominantly at the level of transcription initiation. In order to tightly control the genome, organisms rely on the activity of transcription regulatory proteins, also referred to as transcription factors, to facilitate RNA polymerase initiation or repression. Single-celled organisms can have hundreds of transcription factor proteins, many of which have never been studied. Discovering the transcription regulatory networks for these unknown transcription factors can provide valuable insight into the biology of their host organism.

Many transcription regulatory proteins transmit environmental cues to a genomic response through interactions with cofactors. A classic example of this regulatory mechanism is the ferric uptake regulator (FUR) family. Traditional FUR proteins bind intracellular d-block metal cations, as the d-orbitals are critical for metal-protein coordination. The

Editor Patricia A. Champion, University of Notre Dame

Address correspondence to Michael W. Van Dyke, mvandyk2@kennesaw.edu.

The authors declare no conflict of interest.

Received 12 August 2022
Accepted 29 September 2022

coordinated amino acids (typically histidine and cysteine), as well as surrounding amino acids, help dictate metal specificity (1). Metal coordination by FUR proteins results in a conformational change and subsequent increase in DNA-binding affinity (2). DNA-bound FUR proteins often repress the expression of genes involved in the transport of their regulatory metal, although there are several examples of FUR-mediated transcription activation (3–5). The regulation of these genes under metal-saturated conditions, as well as their deregulation under metal-deficient conditions, allows organisms to adapt and respond to changes in their environment.

FUR family members are commonly named based on the regulatory heavy metal they bind, such as iron (Fur), zinc (Zur), manganese (Mur), or nickel (Nur) (6). They can also be named after the genes they regulate, such as the iron-binding peroxidase stress regulator, PerR. Structurally, FUR proteins are identified by a highly conserved histidine-rich patch preceding a CxxC domain, with some family members having a second CxxC domain at their C terminus (7). These domains are important in coordinating metal binding as well as protein dimerization. A winged-helix DNA-binding domain is often found at the N terminus of this protein family, and FUR members usually bind DNA as a homodimer or multimer of homodimers (8). The preferred DNA-binding sequence for FUR proteins varies between organisms, but most contain a 15- to 20-bp, A/T-rich, pseudopalindromic region, as exemplified by the *Escherichia coli* Fur consensus DNA-binding sequence: GATAATGATAATCATTATC (9).

The extreme thermophile *Thermus thermophilus* HB8 has been used extensively as a model organism in scientific laboratories and has sparked considerable interest in industrial biotechnology. This organism is also the model for RIKEN's Structural-Biological Whole Cell Project, which aims to understand all cellular biology at the atomic level through protein structure analysis. Although heavy metal regulation is well documented in mesothermic organisms, there are few data detailing how *Thermus thermophilus* senses and responds to environmental metals (10). Indeed, there are limited biochemical studies available for FUR homologs from any thermophilic organism (11, 12) or *Deinococcota* bacterium (13). Notably, *Thermus thermophilus* HB8 contains three FUR homologs, encoded by *TTHA0255*, *TTHA0344*, and *TTHA1292*. Currently, the proteins encoded by these three genes do not have any structural data available and only limited biological studies (12). Identifying the transcription regulatory networks of these proteins will broaden our understanding of how thermophilic organisms respond to environmental stressors.

In this study, we discover the preferred DNA-binding sequences for the *Thermus thermophilus* HB8 FUR homolog, TTHA1292, using the nonbiased, iterative selection approach restriction endonuclease, protection, selection and amplification (REPSA). We characterize the TTHA1292 consensus DNA-binding motif, TGAKAWYNNRWTMTCA, map this sequence to the *Thermus thermophilus* HB8 genome, and validate TTHA1292-binding to candidate gene promoters *in vitro*. We further uncover the transcription regulatory network for TTHA1292 *in vivo* and establish TTHA1292 as a zinc uptake regulator.

RESULTS

Identification of preferred DNA-binding sequences for TTHA1292. To identify DNA-binding sequences for the *TTHA1292* gene product, we expressed and purified TTHA1292 from *E. coli* (see Fig. S1 in the supplemental material). Using a random, 24-bp-long selection template (14), we performed several rounds of REPSA with 20 nM TTHA1292 (Fig. 1A). In the initial round of REPSA, TTHA1292 was incubated with random DNA sequences and then treated with the type IIS restriction endonuclease (IISRE), BpmI. TTHA1292-bound DNA sequences were protected from BpmI cleavage, amplified by PCR, and subsequently used for the next round (Fig. S2). After 4 rounds of REPSA, we observed a population of DNA species that exhibited noticeable cleavage protection (Fig. 1A). The presence of a protected DNA species suggests the selection of TTHA1292-specific DNA-binding sequences. To verify binding of TTHA1292 to these REPSA-selected DNA sequences, we performed an electromobility shift assay (EMSA)

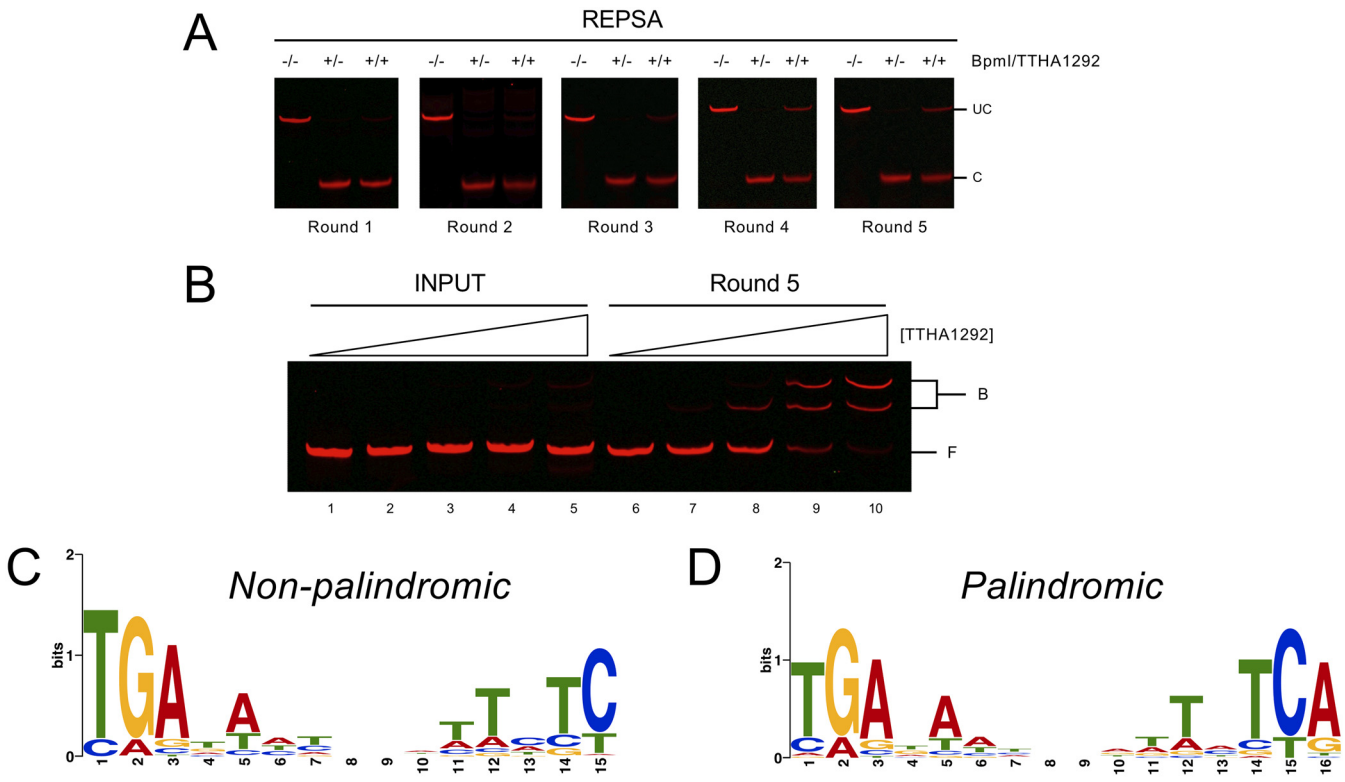
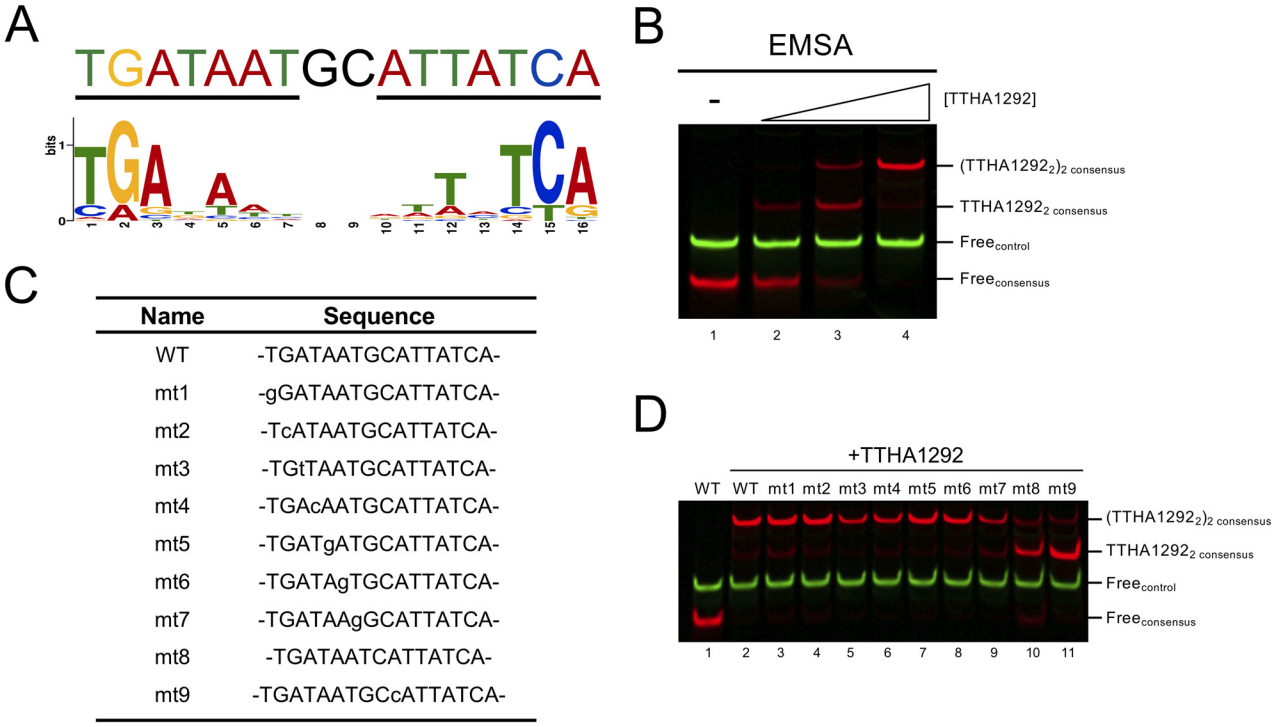


FIG 1 TTHA1292 REPSA selection. (A) Selection templates (73-bp long) containing a random stretch of 24 nucleotides were incubated with TTHA1292 or buffer control and challenged with the type IIS restriction endonuclease BpmI. Samples from reactions containing both TTHA1292 and BpmI were amplified by PCR and used as the input for the next round of REPSA. Uncut (UC) and cut (C) DNA bands are denoted. (B) Selection templates (input) or DNAs amplified after round 5 of REPSA were incubated with 0, 1.25, 6.25, 20, or 80 nM TTHA1292. DNA complexes were separated by native PAGE and visualized using a LI-COR Odyssey imager. Free (F) and TTHA1292-bound (B) DNA complexes are denoted. (C and D) Sequencing reads of round 5 DNAs from REPSA were input into MEME software without (C) or with (D) a palindrome restriction. The position weight matrix for the most significant motif is presented.

with our selection template containing a 24-bp randomized cassette (input) and round 5 DNAs. We observed TTHA1292-dependent shifts in round 5 DNAs and little to no shifts in input DNAs (Fig. 1B). This result confirms that TTHA1292 specifically associates with our REPSA-selected DNA sequences.

To identify TTHA1292 DNA-binding sequences, DNAs from round 5 of REPSA (Fig. 1A) were sequenced using an Ion PGM semiconductor sequencer yielding 3,109,998 bases with an incorrect base-calling quality score, ≥Q20, of 2,546,700 and 63,117 reads of 49 bp average length. Sequencing1.java refinement (14) reduced this to 13,041 sequences. Total and refined sequence data sets are available upon request. The resulting sequences were entered into Multiple Em for Motif Elicitation (MEME) software to identify significant DNA motifs (15). MEME analysis was conducted for nonpalindromic and palindromic sequences, and sequence logos of each position weighted matrix are presented in Fig. 1C and D. Both sequences share a similar 15-nucleotide motif, with the palindromic motif containing one extra nucleotide. The nonpalindromic sequence motif (Fig. 1C) had a statistical significance of 9.3E-469 and was present in 74% of input sequences. The palindromic sequence motif (Fig. 1D) had a statistical significance of 7.8E-338 and was present in 46% of input sequences.

Experimental validation and mutational analysis of TTHA1292 consensus sequence. Since FUR homologs often bind palindromic regions, we derived the consensus sequence TGATAATGCATTATCA from the palindromic MEME motif presented in Fig. 1D. This consensus sequence can be described as two 7-bp inverted repeats separated by 2 nucleotides (Fig. 2A). To analyze the binding of TTHA1292 to our consensus sequence, we performed an EMSA with increasing concentrations of TTHA1292 (Fig. 2B). We observed substantial shifts of our consensus sequence (shown in red) in the presence of TTHA1292, while a control DNA sequence (shown in green) was



Name	Sequence
WT	-TGATAATGCATTATCA-
mt1	-gGATAATGCATTATCA-
mt2	-TcATAATGCATTATCA-
mt3	-TGtTAATGCATTATCA-
mt4	-TGAcAATGCATTATCA-
mt5	-TGATgATGCATTATCA-
mt6	-TGATAgTGCATTATCA-
mt7	-TGATAAgGCATTATCA-
mt8	-TGATAATCATTATCA-
mt9	-TGATAATGCcATTATCA-

FIG 2 Identification and validation of TTHA1292 consensus DNA-binding motif. (A) Wild-type TTHA1292 consensus DNA-binding sequence derived from the palindromic MEME output (position weighted matrix presented below). (B) DNAs containing the TTHA1292 consensus sequence in panel A (5′ IRDye-700 labeled; red) and control DNAs containing a sequence with no homology (5′ IRDye-800 labeled; green) were incubated with 0, 5, 20, or 80 nM TTHA1292. DNA complexes were separated by native PAGE and visualized using a LI-COR Odyssey imager. Free DNA and TTHA1292-DNA complexes are denoted. The total length of consensus sequence DNA is 65 bp, while the total length of the control DNA is 80 bp. (C) Nucleotide sequences of point mutations within the TTHA1292 consensus sequence. (D) DNAs containing the sequences depicted in panel C (5′ IRDye-700 labeled; red) and control DNAs containing a sequence with no homology (5′ IRDye-800 labeled; green) were incubated with 80 nM TTHA1292. DNA complexes were separated by native PAGE and visualized using a LI-COR Odyssey imager.

unaffected. We saw two distinct shifted species in the presence of TTHA1292 (Fig. 2B, lanes 2 to 4), suggesting that TTHA1292 can multimerize when binding to its consensus sequence (4, 9). Based on an adaptation of a Ferguson plot (16), we estimated the faster-mobility species observed by EMSA to be approximately 67 kDa and the slower-mobility species to be approximately 93 kDa (Fig. S3). These values are consistent with the predicted molecular weights of dimeric (~65 kDa, denoted as TTHA1292$_2$) and tetrameric [~92 kDa, denoted as (TTHA1292$_2$)$_2$] TTHA1292 complexes with DNA. Collectively, these results highlight REPSA as a useful approach in determining preferred DNA-binding sequences and show that TTHA1292 can bind its consensus sequence as a multimerization of homodimers.

Position weighted matrices derived from MEME provide information on the likelihood of a nucleotide being at a certain position based on input sequences. However, this output does not formally address the relative importance of each nucleotide in promoting a protein-DNA interaction. To experimentally test this, we introduced single-point mutations within our consensus sequence (Fig. 2C) and assayed binding by EMSA (Fig. 2D). We assayed mutants within one inverted repeat (mutant 1 [mt1] to mt7), as well as deletion and insertion mutants within the region separating the inverted repeats (mt8 and mt9, respectively). Mutants within the inverted repeat region were created using the least common nucleotide at that position found in the MEME motif. Surprisingly, mutating single nucleotides within an inverted repeat region did not dramatically affect the ability of TTHA1292 to form multimeric complexes on DNA (Fig. 2D, lanes 3 to 9). Conversely, adding or removing a nucleotide in the region between the inverted repeats resulted in a clear reduction in (TTHA1292$_2$)$_2$ binding (Fig. 2D, lanes 10 to 11). However, we still observed dimeric TTHA1292 binding with

TABLE 1 Potential genomic TTHA1292-binding sites identified by FIMO[a]

Chromosome	Start	End	*P* value	*Q* value	Sequence	Distance to TSS (bp)	Gene name	Op
1	562950	562965	2.95e-10	0.000624	TGATATTGGATTATCA	−9	*TTHA0596*	S
						−66	*TTHA0597*	S
1	183117	183132	5.31e-08	0.0562	TAAGATCTTTTTATCA	−92	*TTHA0185*	1/2
						−1	*TTHA0186*	S
pTT27	113443	113458	8.57e-07	0.314	TGACTAATAATGATCA	−134	*TTHB125*	1/2
1	1000549	1000564	1.04e-06	0.314	TGAGCATTTTTACTCA	−6	*TTHA1055*	S
						−95	*TTHA1056*	S
1	1233778	1233793	1.26e-06	0.334	TGAAAACGCGTTATCG	−35	*TTHA1292*	S
1	39655	39670	1.47e-06	0.345	TGAGATTTCCCTTTCA	+1	*TTHA0038*	2/2
1	425230	425245	3.84e-06	0.813	TGAAAATGCGCTATCG	−9	*TTHA0453*	1/2
pTT27	72914	72929	4.78e-06	0.920	TAAGAAAGCGCTTTCA	+5	*TTHB082*	1/2
pTT27	173876	173891	7.03e-06	1	AGAGTAGGTTTTATCA	−6	*TTHB180*	2/2

[a]*P* value, the probability of a random sequence of similar length matching the input motif with an as good as or better score; *q* value, false-discovery rate if the identified sequence is accepted as significant; TSS, translation start site; Op, gene position within the operon (operons were predicted using the MicrobesOnline Operon Predictions database); S, single transcription unit.

these mutants, suggesting that the spacing between the 7- bp inverted repeat regions of our consensus sequence is critical for coordinating the multimerization of TTHA1292 homodimers on DNA. Further binding analyses with sequences containing only one inverted repeat resulted in little to no binding of TTHA1292, showing that both inverted repeat regions are critical for the DNA recognition of TTHA1292 (Fig. S4). Notably, we only observed $TTHA1292_2$ and $(TTHA1292_2)_2$ species for all DNAs and TTHA1292 concentrations tested by EMSA, consistent with only the homodimer species having DNA-binding activity (7). Moreover, the identification of DNAs that lack $(TTHA1292_2)_2$ binding suggests that TTHA1292 homodimers form tetrameric species on DNA rather than in solution.

Identification and *in vitro* validation of genomic TTHA1292-binding sequences. To identify potential TTHA1292-binding sequences in the *Thermus thermophilus* HB8 genome, we input the position weighted matrix of the palindromic consensus motif identified by MEME (Fig. 1D) into Find Individual Motif Occurrences (FIMO) software (17). The complete output of this search is presented in Data Set S1. To identify potential TTHA1292-regulated genes, we determined FIMO-identified sequences with a *P* value of $<1 \times$ 10E-5 that were found within −200/+10 bp of a translation start site, the most common region to find prokaryotic transcription factors. These sequences, along with the gene(s) whose promoters they are found in, are presented in Table 1. Some sequences were found within promoters of divergent genes, for which both genes are annotated.

To test whether TTHA1292 binds to FIMO-identified sequences, we created DNA fragments containing each genomic sequence shown in Table 1 and analyzed TTHA1292-binding by EMSA (Fig. 3A and B). We found several promoter sequences that showed clear evidence

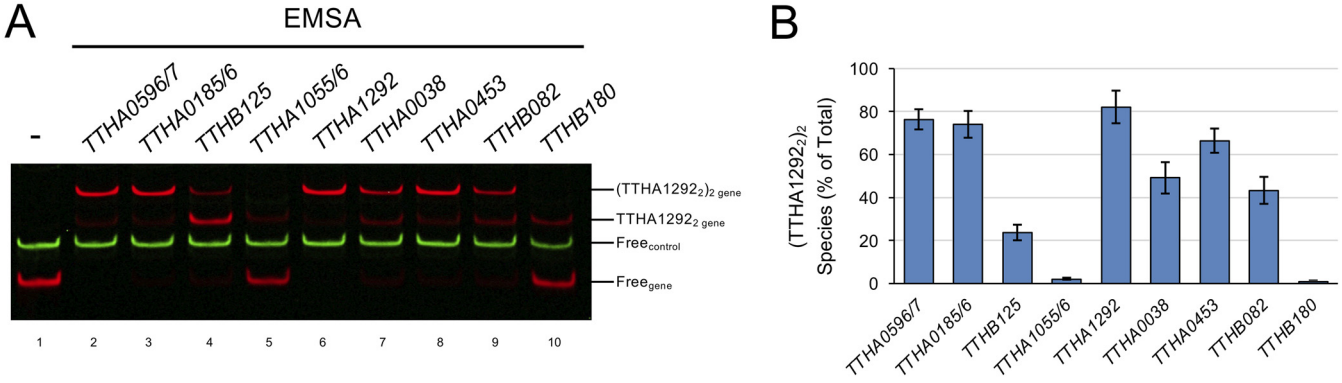


FIG 3 Binding analysis of promoter sequences from potential TTHA1292-regulated genes. (A) DNAs containing FIMO-identified sequences within the promoters of the specified genes (5′ IRDye-700 labeled; red) and control DNAs containing a sequence with no homology (5′ IRDye-800 labeled; green) were incubated with 80 nM TTHA1292. DNA complexes were separated by native PAGE and visualized using a LI-COR Odyssey imager. (B) The intensity of the $(TTHA1292_2)_2$-complexed DNA band in panel A was quantified and normalized to total 5′ IRDye-700 labeled DNAs within each reaction sample. Error bars represent ±1 standard deviation between independent experiments (n = 2).

TABLE 2 Gene functions of *in vitro* validated TTHA1292-binding sequences[a]

Operon	Gene	Role
S	*TTHA0596*	Zinc transport system substrate-binding protein
S	*TTHA0597*	Conserved hypothetical protein
1	*TTHA0185*	Pyruvate dehydrogenase E1 component
2	*TTHA0184*	Pyruvate dehydrogenase E2 component
S	*TTHA0186*	LysR family transcriptional regulator
1	*TTHB125*	Chromosome partitioning ATPase, ParA family
2	*TTHB124*	Chromosome partitioning protein, ParB family
S	*TTHA1292*	Fur family transcriptional regulator
2	*TTHA0038*	Hypothetical protein
1	*TTHA0453*	Zinc transport system ATP-binding protein
2	*TTHA0454*	Zinc transport system permease protein
1	*TTHB082*	Multiple sugar transport system substrate-binding protein
2	*TTHB083*	Multiple sugar transport system permease protein

[a]Operon, gene position within the operon (operons were predicted using the MicrobesOnline Operon Predictions database); S, single transcription unit; role, functional role of the encoded protein predicted by GenBank.

of tetrameric TTHA1292-binding (*TTHA0596/7*, *TTHA0185/6*, *TTHA1292*, *TTHA0453*), others that showed some evidence of tetramerization (*TTHB125*, *TTHA0038*, *TTHB082*), and others that showed little evidence of TTHA1292 binding (*TTHA1055/6* and *TTHB180*). Collectively, these results validate the binding of TTHA1292 to several genomic sequences identified by FIMO and suggest that TTHA1292 regulates the expression of these genes.

To understand the biological function of each potential TTHA1292-regulated gene, we identified the proposed role of each gene product, as well as the role of any downstream gene products if the potential TTHA1292-regulated gene is within an operon (18, 19) (Table 2). The identified gene roles varied in biological function, including transcriptional regulators, transport systems, chromosome partitioning proteins, and proteins of unknown function. We identified *TTHA1292* as a potential TTHA1292-regulated gene, and autoregulation is common among other FUR homologs (20). Notably, several potential TTHA1292-regulated genes were involved in zinc uptake (*TTHA0596* and *TTHA0453/4*), consistent with traditional FUR proteins that regulate genes involved in heavy metal transport. The specificity to genes involved in zinc regulation suggests that TTHA1292 is a zinc uptake regulator (Zur).

DNA-binding activity of TTHA1292 is regulated by zinc. FUR homologs classically bind heavy metal cations to localize to DNA. To determine if the DNA-binding activity of TTHA1292 is affected by heavy metals, we performed REPSA in the presence of $MnCl_2$, which is often used as a surrogate for Fe^{2+} (12, 21). We observed cleavage-resistant species by round 4 that specifically interacted with TTHA1292 even in the absence of $MnCl_2$ (Fig. S5A and B). Round 4 DNAs were sequenced, and motif identification by MEME resulted in similar motifs as REPSA when performed without $MnCl_2$ (Fig. S5C and D). This result suggests that our *E. coli*-purified TTHA1292 does not need additional metal cofactors to bind DNA. However, our purified TTHA1292 protein could already be bound to a regulatory metal. To test this, we performed an EMSA with TTHA1292 and its consensus DNA-binding sequence in the presence of the metal chelator, EDTA. As seen in Fig. 4A, increasing amounts of EDTA led to a dramatic decrease in tetrameric TTHA1292-bound DNA. However, the addition of EDTA did not appear to affect the formation of dimeric TTHA1292-bound DNA complexes. Interestingly, when titrating the concentration of TTHA1292, we saw that the tetrameric TTHA1292-bound DNA species was specifically affected by EDTA, rather than the dimeric TTHA1292-bound DNA species or free DNA bands (Fig. 4B). The addition of EDTA will compete for regulatory metals bound to FUR homodimers; however, homodimers will retain structural zinc ions that are required for dimerization even in the presence of EDTA (22, 23). This result suggests that TTHA1292 can bind DNA in the absence of a regulatory metal and that a regulatory metal is critical for multimerization of TTHA1292 dimers on DNA. To analyze this phenomenon further, we performed biolayer interferometry using the TTHA1292 consensus DNA-binding

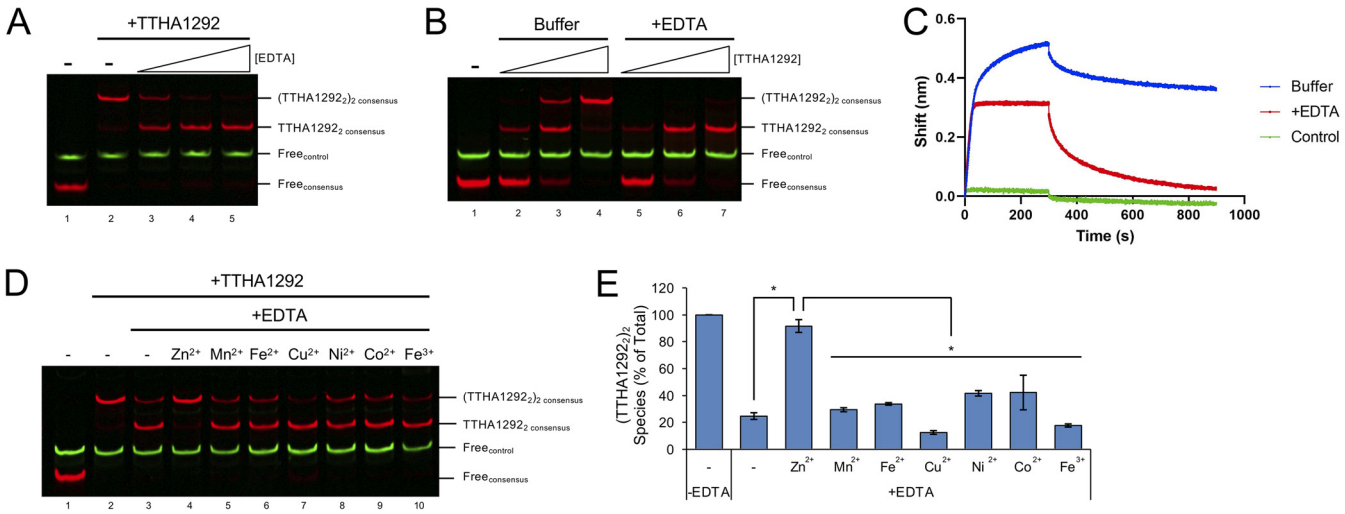


FIG 4 The DNA-binding activity of TTHA1292 is regulated by zinc. (A) DNAs containing the TTHA1292 consensus sequence (5′ IRDye-700 labeled; red) and control DNAs containing a sequence with no homology (5′ IRDye-800 labeled; green) were incubated with 60 nM TTHA1292 and 0, 0.75, 1.5, or 3 mM EDTA. DNA complexes were separated by native PAGE and visualized using a LI-COR Odyssey imager. (B) DNAs containing the TTHA1292 consensus sequence (5′ IRDye-700 labeled; red) and control DNAs containing a sequence with no homology (5′ IRDye-800 labeled; green) were incubated with 0, 5, 20, or 80 nM TTHA1292 in the presence of 3 mM EDTA or buffer control. DNA complexes were separated by native PAGE and visualized using a LI-COR Odyssey imager. (C) Representative graph showing the association (time 0 to 300 s) and dissociation (time 300 to 900 s) steps of biolayer interferometry. Reactions were conducted with the TTHA1292 consensus sequence (blue and red traces) or a control sequence with no homology (green) in the presence of 100 nM TTHA1292. Where indicated, reactions were supplemented with 3 mM EDTA. (D) DNAs containing the TTHA1292 consensus sequence (5′ IRDye-700 labeled; red) and control DNAs containing a sequence with no homology (5′ IRDye-800 labeled; green) were incubated with either 80 nM TTHA1292, 3 mM EDTA, 100 μM of the indicated metal cation or a buffer control. DNA complexes were separated by native PAGE and visualized using a LI-COR Odyssey imager. (E) The intensity of the (TTHA1292$_2$)$_2$-complexed DNA band in panel D was quantified and normalized to total 5′ IRDye-700 labeled DNAs within each reaction sample. Error bars represent $\pm$1 standard deviation between independent experiments (n = 2). Student's two-tailed t test; *, $P < 0.05$.

sequence and a control sequence with no homology. As expected, we observed little to no interaction with TTHA1292 and the control DNA sequence (Fig. 4C, green trace). In the absence of EDTA, we see a clear association of TTHA1292 and its consensus sequence (Fig. 4C, blue trace). The association curve mimics a biphasic interaction, with a high rate of association within the first ~50 s followed by a slower rate of association between 100 and 300 s. We hypothesize that the faster rate of association depicts dimeric TTHA1292-binding, which is followed by *de novo* formation of tetrameric TTHA1292 complexes on DNA (slower rate). Consistent with this interpretation, the presence of EDTA specifically affects the slower association step (Fig. 4C, red trace). Furthermore, we saw a higher rate of dissociation (initiated at time [T] = 300 s) in the presence of EDTA, suggesting that the EDTA-treated dimeric complex has a weaker affinity to DNA than the tetrameric complex observed in reactions not containing EDTA.

To determine which regulatory metal TTHA1292 uses to tetramerize on DNA, we performed an EMSA in the presence of EDTA and several different metal cations. As with previous experiments, we observed a substantial decrease in (TTHA1292$_2$)$_2$ in the presence of EDTA (Fig. 4D, compare lanes 2 and 3). Notably, we observed a nearly complete rescue of tetrameric TTHA1292-bound DNA by supplementing reactions with zinc, while all other metals tested resulted in little to no effect on tetrameric TTHA1292-binding. (Fig. 4D and E). We observe a slight reduction in (TTHA1292$_2$)$_2$ with the addition of Cu^{2+} and Fe^{3+} cations, which may be attributed to these ions' ability to bind DNA (24, 25). These results show that TTHA1292 utilizes zinc to promote tetrameric protein complexes on DNA and further implicates TTHA1292 as a *bona fide* Zur protein.

***In vivo* and *in vitro* validation of the TTHA1292 transcription regulatory network.** To understand where our FIMO-identified TTHA1292-binding sites are located within the promoters of the genes identified in Table 2, we mapped the TTHA1292-binding sequence as well as the predicted −10 and −35 elements of each gene in Fig. S6. Regulatory sequences of each promoter were predicted using BPROM (26). In several gene promoters, the predicted TTHA1292-binding sequence overlapped −10 elements

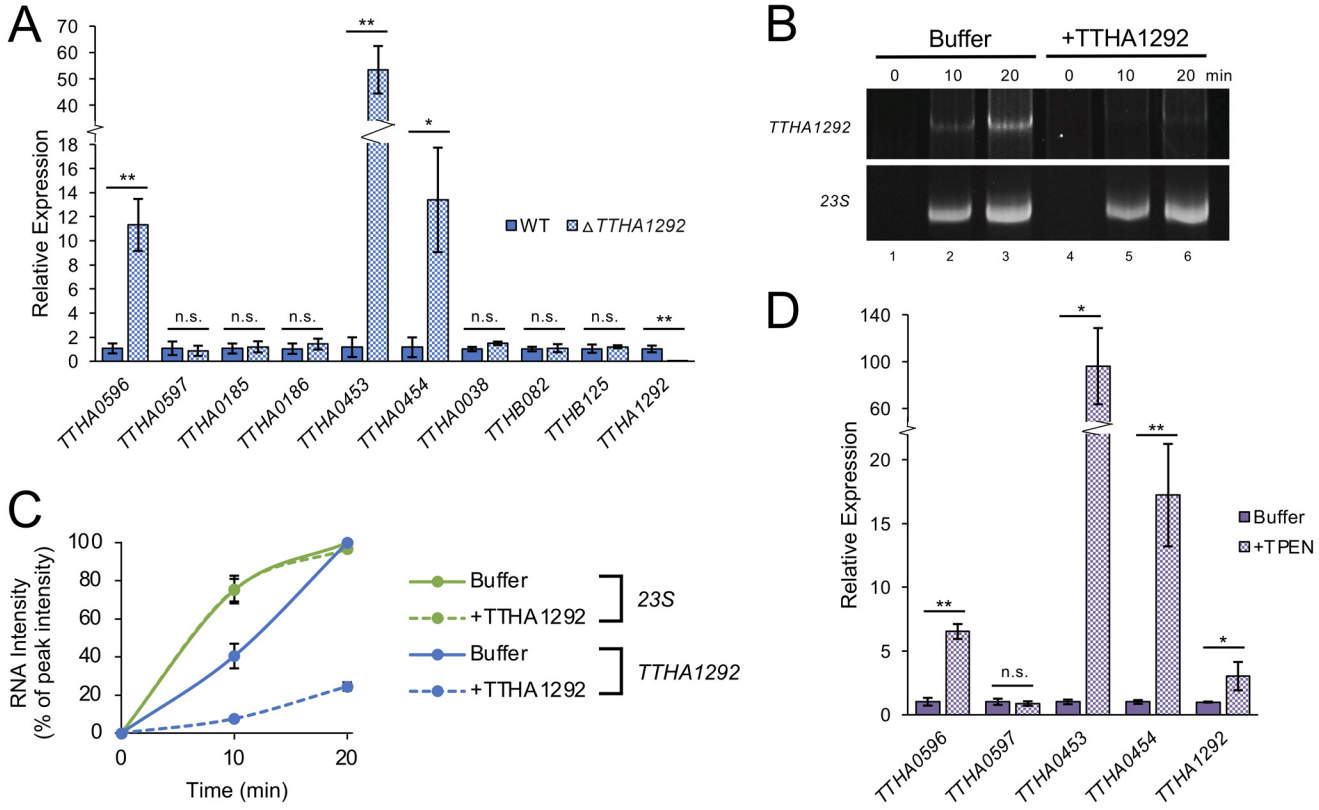


FIG 5 Validation of TTHA1292-mediated gene regulation *in vivo* and *in vitro*. (A) RNA was isolated from wild-type and *TTHA1292*-disrupted (Δ*TTHA1292*) *Thermus thermophilus* HB8 strains during log-phase growth. Gene expression was quantified by reverse transcriptase quantitative PCR (RT-qPCR) and normalized to 16S rRNA expression. Values relative to wild-type expression are presented. Error bars represent ±1 standard deviation between independent experiments ($n = 3$). Student's two-tailed *t* test; *, $P < 0.05$; **, $P < 0.005$; n.s., not significant. (B) At the indicated time points, samples were taken from *in vitro* transcription reactions containing 80 nM RNAP/sigma, 40 nM the respective promoter templates, and 1 μM TTHA1292 or buffer control. Samples were treated with DNase I and separated by native PAGE. RNA was visualized by SYBR gold staining. (C) The intensities of the RNA bands in panel B were quantified and normalized to the peak intensity observed for each promoter template. Error bars represent ±1 standard deviation between independent experiments ($n = 2$). (D) RNA was isolated from *Thermus thermophilus* HB8 cultures treated with 20 μM TPEN or buffer control for 30 min. Gene expression was quantified by RT-qPCR and normalized to 16S rRNA expression. Values relative to wild-type expression are presented. Error bars represent ±1 standard deviation between independent experiments ($n = 3$). Student's two-tailed *t* test; *, *P* value < 0.05; **, *P* value < 0.005.

(*TTHA0596*, *TTHA0185*, *TTHA0186*, *TTHA0038*, and *TTHA0453*), while others had the TTHA1292-binding sequence overlapping the −35 element (*TTHB125* and *TTHA1292*).

To test which genes TTHA1292 controls the expression of *in vivo*, we developed mutant strains of *Thermus thermophilus* HB8 that had a disrupted *TTHA1292* gene (27). We then compared the expression levels of our potential TTHA1292-regulated genes in wild-type and *TTHA1292*-deleted (Δ*TTHA1292*) strains. Notably, since *Thermus thermophilus* HB8 is a polyploid bacterium, it is possible that a portion of the genomic *TTHA1292* copies persists within our Δ*TTHA1292* strains. However, the clear reduction in *TTHA1292* expression in these strains compared to the wild-type strain (Fig. 5A) suggests that a majority of *TTHA1292* is disrupted. Expression of the two transcription units involved in zinc transport (*TTHA0596* and *TTHA0453/4*) was significantly elevated in the Δ*TTHA1292* strain, suggesting TTHA1292 represses these genes *in vivo* (Fig. 5A). Surprisingly, the remaining FIMO-identified genes had little to no change in expression between the wild-type and Δ*TTHA1292* strains. This result shows that although TTHA1292 can bind to the promoter of these genes *in vitro* (Fig. 3), loss of TTHA1292 may not be sufficient to regulate the expression of these genes under the tested growth conditions.

Since Δ*TTHA1292* strains lack the *TTHA1292* sequence, we sought an alternative to determine if TTHA1292 regulates its own promoter activity. To do so, we developed an *in vitro* transcription assay using purified *Thermus thermophilus* HB8 RNA polymerase (RNAP) core enzyme and sigma factor, RpoD (Fig. S7). We analyzed RNA production

from DNA templates containing the promoter regions of either *TTHA1292* or the 23S rRNA gene in the presence or absence of recombinant TTHA1292 protein (Fig. 5B). RNA levels from the *23S* template were unaffected by the addition of TTHA1292; however, expression from the *TTHA1292* template exhibited a substantial decline in the presence of TTHA1292 (Fig. 5B and C). This result implicates TTHA1292 as a negative regulator of its own promoter. Consistent with these findings, the addition of the zinc-specific chelator, TPEN (*N,N,N′,N′-tetrakis*(2-pyridinylmethyl)-1,2-ethanediamine), promoted a significant increase in *TTHA1292* expression *in vivo* (Fig. 5D). The expressions of *TTHA0596*, *TTHA0453*, and *TTHA0454* were also significantly increased in the presence of TPEN, similar to the expression changes observed in Δ*TTHA1292* strains (compare Fig. 5A and D). Collectively, these results reveal novel members of the TTHA1292 transcription regulatory network (*TTHA1292*, *TTHA0596*, and *TTHA0453/4*) and suggest that TTHA1292 regulates these genes in a zinc-dependent mechanism.

DISCUSSION

In this study, we used REPSA to identify a consensus DNA binding motif for the thermophilic FUR homolog TTHA1292 (Fig. 1). Although DNA-binding consensus sequences for FUR proteins have been characterized based on putative binding sequences within target gene promoters, we provide the first utilization of an iterative selection method to identify preferred DNA-binding sequences for a member of the FUR superfamily. The discovered TTHA1292 DNA-binding motif was used to identify genomic DNA sequences (Table 1), and binding to these sequences was validated *in vitro* (Fig. 3). Then, we confirmed members of the TTHA1292 regulon by analyzing expression levels in TTHA1292-mutant *Thermus thermophilus* HB8 *in vivo* (Fig. 5). This workflow promotes directed identification of transcription regulatory networks that is not muddled by confounding effects on gene expression observed in mutant bacterial strains.

We biochemically characterized TTHA1292 as a zinc uptake regulator (Fig. 4), providing the first such data for a thermophilic organism. The amino acid make-up of TTHA1292 also provides evidence that this protein is a *bona fide* Zur. Using position-specific iterative BLAST (PSI-BLAST) to search for homology within the Protein Data Bank (PDB) (28), we found that TTHA1292 most closely resembled other Zur proteins (Fig. S8A). When aligning the TTHA1292 amino acid sequence to related Zur proteins, we found that several amino acids experimentally determined to be involved in zinc coordination were conserved in TTHA1292 (Fig. S8B).

Classic Zur proteins repress the transcription of the *znuABC* operon, which encodes an ABC transporter that regulates zinc uptake (29). In *Thermus thermophilus* HB8, the *znuA* homolog, which encodes a zinc transport system substrate-binding protein, is a single transcription unit (*TTHA0596*). The *znuB/C* homologs, which encode a zinc transport system permease protein and ATP-binding protein, respectively, are in an operon (*TTHA0453/4*). Genetic separation of *znuA* and *znuB/C* is common in several other bacteria (30, 31), and several organisms contain the *zur* gene within the *znuB/C* operon, allowing for single regulation of both sets of genes (32, 33). In *Thermus thermophilus* HB8, we identified the *zur* gene (*TTHA1292*) as a single transcription unit separated from each *znu* homolog. We provide *in vitro* and *in vivo* data showing that TTHA1292 represses the transcription of each of these transcription units (*TTHA0596*, *TTHA0453/4*, and *TTHA1292*), thus presenting TTHA1292 as a major regulator of zinc uptake in *Thermus thermophilus* HB8. Although zinc levels are bound to fluctuate in different environments, a hot spring containing *Thermus thermophilus* was found to contain ~0.2 mg/L zinc, suggesting that zinc is relatively abundant in natural *Thermus thermophilus* habitats (34). Indeed, *Thermus thermophilus* growth medium (ATCC medium 697) promoted sufficient repression of *znuABC* genes (Fig. 5A), suggesting that zinc is not a limiting factor under ideal growth conditions. We also identified several other genes with no relation to zinc uptake whose promoters may be occupied by TTHA1292 (Table 2; Fig. 3) but were not affected by disrupting *TTHA1292*. The biological

importance of TTHA1292's affinity to sequences within these genes' promoters may be more complex and warrants further investigation.

Our TTHA1292 DNA-binding sequence has similarities to and differences from established Zur consensus sequences (i.e., Zur boxes). Most Zur boxes are made up of an A/T-rich sequence containing two inverted repeat regions separated by a defined region (35). Our TTHA1292 consensus sequence matches this description; however, the 7-2-7 motif we have identified for TTHA1292 (Fig. 1D) is unique compared to established Zur boxes (35). A previous report suggested that *E. coli* Zur binds DNA as a tetramer by recognizing an RNNNYRNNRYNNYRNNNY motif (R, purine; Y, pyrimidine) (22). TTHA1292 appears to not utilize the same DNA-binding specificity, since mutations within a single inverted repeat did not disrupt tetrameric binding (Fig. 2D). Notably, we show that TTHA1292 can bind as a dimeric species to the 7-1-7 and 7-3-7 motifs (Fig. 2D); however, whether this binding is relevant to its transcription regulation activity is yet to be determined.

Our results suggest that TTHA1292 can bind DNA even in the absence of a regulatory metal (Fig. 4A and B). However, we observed that zinc is critical for the stability of a TTHA1292-DNA interaction (Fig. 4C) and multimerization of TTHA1292 homodimers on DNA (Fig. 4D and E). Both of these factors are likely necessary for TTHA1292's transcription regulatory function (Fig. 5D). This observation is similar to studies conducted with *Streptomyces coelicolor* Zur, where the authors showed zinc-dependent formation of dimeric and tetrameric Zur complexes on DNA (4). Zur proteins contain a structural zinc ion that is retained even in the presence of EDTA and is necessary for protein dimerization (22, 23). Some Zurs contain two additional regulatory zinc-binding sites, while others contain only one regulatory zinc site (22, 36–38). When we align TTHA1292 to homology-related Zurs with experimentally determined crystal structures (including the aforementioned *Streptomyces coelicolor* Zur), we find that the amino acids required for coordinating the structural zinc, as well as one regulatory zinc, are completely conserved (Fig. S8B). However, a second regulatory zinc-binding site has only two of four amino acids conserved. Further studies will be needed to determine if TTHA1292 contains a second regulatory zinc-binding site, and if so, how TTHA1292 coordinates this second regulatory zinc ion. Furthermore, identifying the role of each regulatory zinc site in promoting a protein-DNA interaction may provide insight into the TTHA1292-binding dynamics observed in this study.

Collectively, these data support a model in which dimer coordination dictated by DNA sequence and intracellular zinc is critical for multimerization of TTHA1292 homodimers on DNA and subsequent transcription repression. These results provide new insight into the biology of the FUR superfamily and further our understanding of how thermophilic organisms sense and adapt to heavy metals in their environment.

MATERIALS AND METHODS

Oligonucleotides. All oligonucleotides and primers used in this study were purchased from Integrated DNA Technologies (IDT) and are presented in Table S1. Each oligonucleotide contains identical flanking regions (denoted as ST2L and ST2R) used for PCR amplification. The only exception to this is for EMSA reactions where the control DNA sequence was amplified using ST2R and 5′ IRDye-800-labeled ControlL primers. Where indicated, the ST2R primer was conjugated with 5′ IRDye-700 or 5′ biotin. All PCRs were conducted with New England Biolabs (NEB) *Taq* DNA polymerase with standard *Taq* buffer under reaction conditions specified by the manufacturer.

Expression and purification of TTHA1292. A TTHA1292 expression vector (TEx11D07) was purchased from RIKEN BioResource Research Center (39). *E. coli* strain Rosetta (DE3) competent cells were transformed with the TTHA1292 expression vector. Successfully transformed cells were grown at 37°C and 250 rpm to an optical density (OD) of ~0.5 and induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 4 h. Cells were then pelleted and resuspended in 2× bacterial extract buffer (40 mM Tris-Cl [pH 7.5], 200 mM NaCl, 0.2 mM EDTA, 2 mM dithiothreitol [DTT], and 1 mM phenylmethylsulfonyl fluoride [PMSF]). Next, 10 μg/μL lysozyme was added at 0.02× the total volume and incubated on ice for 10 min. The cell mixture was then sonicated at 3 W, 10 s on/50 s off, for 5 cycles, and then pelleted. Chromatin-bound proteins were extracted from the pellet with 2× bacterial extract buffer containing 1 M NaCl. After centrifugation, the resulting supernatant was collected and heat-treated at 70°C for 15 min, resulting in the denaturation of endogenous *E. coli* proteins. Samples were then centrifuged, and the supernatant was collected, diluted to 50% glycerol, and stored at −20°C. The resulting sample was found to be over 95% pure by SDS-PAGE and Coomassie staining (Fig. S1).

Restriction endonuclease protection, selection, and amplification (REPSA). REPSA was performed with a 24-nucleotide, random selection template as described previously ("ST2R24" in Table S1) (14). To create an input library for REPSA, the 24-nucleotide, random selection template was subjected to 5 rounds of PCR using 5′ IRDye 700-labeled primers and purified using the DNA Clean and Concentrator-5 kit (Zymo Research). Then, 20-μL reactions were performed in 1× Cutsmart buffer (NEB) supplemented with 1 mM DTT containing 2 ng of the selection template and 20 nM TTHA1292. Reactions were incubated at 55°C for 20 min and then incubated at 37°C for 5 min prior to BpmI digestion. 0.4 units of BpmI were used for each digestion. BpmI treatment occurred for 5 min at 37°C, and reactions were stopped by immediately placing the mixture on ice. Reactions containing TTHA1292 and BpmI were subject to PCR using 5′ IRDye 700-labeled primers, purified using the DNA Clean and Concentrator-5 kit (Zymo Research), and used in the subsequent round of REPSA. Samples from reactions containing DNA only, DNA and BpmI, and DNA, BpmI, and TTHA1292 were separated by 10% native PAGE and visualized using a LI-COR Odyssey imager.

DNA sequencing and bioinformatics. REPSA-selected DNAs were sequenced using a Thermo Fisher Ion Personal Genome Machine (PGM) essentially as previously described (14). Ion amplicon libraries were prepared by fusion-cloning REPSA-selected DNAs using the appropriate barcoded primers. These were affixed to individual sequencing particles using Ion PGM Hi-Q reagents and multiplex-sequenced on an Ion PGM using an Ion 314 chip. Sequencing data were initially processed with an Ion Torrent server and output as individual fastq files for each barcoded library. These were further processed by our Sequencing1.java script to remove flanking regions from our sequencing reads, resulting in a library of 24-nucleotide long sequences. Consensus sequences were determined by inputting our refined sequencing libraries into Multiple Em for Motif Elicitation (15) (MEME; https://meme-suite.org/meme/tools/meme; accessed 1 October 2021) v 5.4.1 using default parameters, except for a palindromic filter where indicated. The position weighted matrix of the most significant consensus motif was then scanned to the *Thermus thermophilus* HB8 genome using Find Individual Motif Occurrences (17) (FIMO; https://meme-suite.org/meme/tools/fimo; accessed 1 October 2021) v 5.4.1 using default parameters. Sequences that fell within −200/+10 nucleotides of a gene were identified using the Kyoto Encyclopedia of Genes and Genomes (KEGG) *T. thermophilus* HB8 database (40) (https://www.genome.jp/kegg-bin/show_organism?org=T00220). Predicted biological functions of *Thermus thermophilus* HB8 genes were identified using GenBank (19), and operon predictions were identified using MicrobesOnline (18) (http://www.microbesonline.org/operons/gnc300852.html). Regulatory elements in the promoter regions of candidate TTHA1292-regulated genes were predicted by Softberry BPROM (26) (http://www.softberry.com/berry.phtml?t\opic=bprom&group=programs&subgroup=gfindb).

Electromobility shift assay (EMSA). Reactions consisted of the indicated concentration of TTHA1292, 1× Cutsmart buffer (NEB), 5 to 10 nM 5′ IRDye 800 or 5′ IRDye 700-labeled DNAs, and 1 mM DTT. Each reaction was incubated at 55°C for 20 min and then diluted with 1× EMSA loading dye (3.3 mM Tris-HCl [pH 7.5], 3.3% glucose, 0.15% Orange G). Samples were then separated by 10% native PAGE and visualized using a LI-COR Odyssey imager.

Biolayer interferometry (BLI). BLI was performed and analyzed essentially as described previously (41). DNA probes utilized for BLI analysis were amplified using 5′ biotin-labeled primers. BLI was performed in the following buffer: 20 mM Tris-Cl (pH 7.5), 300 mM NaCl, 0.01% Tween 20, and 1 mM DTT. Where indicated, reactions were supplemented with 3 mM EDTA. BLI was conducted using the Forte Bio OctetRED96e system and set up as followed: 100 s startup, 900 s biosensor loading, 100 s baseline, 300 s association, and 600 s dissociation. The optical wavelength was recorded every 0.2 s. Representative graphs show the association and dissociation steps.

***Thermus thermophilus* HB8 disruption mutants.** To create Δ*TTHA1292* strains, a TTHA1292-disruption plasmid (TDs07G02) was purchased from RIKEN BioResource Research Center (39). *Thermus thermophilus* HB8 (ATCC 27634, described as wild-type for this study) was grown overnight at 70°C and 180 rpm in TT medium (0.8% polypeptone peptone, 0.4% yeast extract, 0.2% NaCl) supplemented with 0.4 mM $MgCl_2$ and 0.4 mM $CaCl_2$. The overnight culture was diluted 50-fold in the same medium and grown at 70°C and 180 RPM for 2 h. Then, 500 ng of the TTHA1292-disruption plasmid was added to the culture and grown for another 2 h. The culture was then plated onto TT plates (TT medium, 1.5% gellan gum, 1.5 mM $MgCl_2$, and 1.5 mM $CaCl_2$) containing 500 μg/mL kanamycin and incubated at 70°C overnight. Successful Δ*TTHA1292* clones were grown in kanamycin-containing medium and checked by genomic PCR.

***Thermus thermophilus* HB8 RNA isolation and quantitative PCR (qPCR).** Wild-type or Δ*TTHA1292 Thermus thermophilus* HB8 strains were streaked onto TT plates. Colonies were picked and grown overnight in TT medium supplemented with 0.4 mM $MgCl_2$ and 0.4 mM $CaCl_2$ at 70°C and 180 rpm. Overnight cultures were seeded into 30-mL cultures and grown until log phase growth (OD, ~0.5) was reached. Where indicated, cultures were treated with 20 μM TPEN (Cayman Chemical Company) or vehicle control for 30 min. Samples were pelleted at 10,000 × *g* for 1 min, and RNA was isolated using the Quick-RNA fungal/bacterial miniprep kit (Zymo Research) following the manufacturer's protocol. RNA samples were treated with 5 units DNase I (Zymo Research) and further purified using the RNA Clean and Concentrator kit (Zymo Research). cDNA libraries were generated using the First Strand cDNA synthesis kit (ApexBio), and gene expression was quantified by qPCR using the "qPCR" primers in Table S1. Expression levels for each gene were normalized to the expression of the 16S rRNA gene.

***In vitro* transcription.** A plasmid that allows the coexpression of the four subunits of *T. thermophilus* HB8 RNA polymerase (his-tagged β', β, α, ω) was the generous gift of Konstantin Severinov. This plasmid was transformed into the *E. coli* Rosetta 2 (DE3) strain and grown in LB medium with kanamycin and chloramphenicol for 18 h at 37°C to allow the maximal expression of assembled *T. thermophilus* HB8 core RNA polymerase in a soluble form. Bacterial extract preparation and initial purification by heat

treatment were conducted as previously described (42). Further purification was performed by immobilized metal affinity chromatography using Ni^{2+}-charged Cytiva His SpinTrap columns following manufacturer's instructions. A TTHA0532 (Sigma; RpoD) expression vector (TEx09B07) was purchased from RIKEN BioResource Research Center (39) and expressed as described previously (42).

In vitro transcription reactions, as well as any protein dilutions needed for the reaction, were performed in 1× Cutsmart buffer (NEB) supplemented with 1 mM DTT. DNA templates containing ~200 nucleotides (nt) of the promoter region, as well as ~200 nt downstream of the transcription start site for each gene, were amplified from the *Thermus thermophilus* HB8 genome using the "clone" primers in Table S1 and purified using a DNA Clean and Concentrator kit (Zymo Research). Then, 40 nM DNA templates were incubated with either 1 μM TTHA1292 or a buffer control for 5 min at 65°C. *Thermus thermophilus* HB8 RNAP core and RpoD were preincubated in a 1:1 molar ratio for 10 min at 65°C and then added to the reaction at a final concentration of 80 nM. The DNA/RNAP mixture was incubated further for 5 min at 65°C, and then 3 mM ribonucleoside triphosphate (rNTP) was added, denoting the start of the reaction. Reactions continued to be incubated at 65°C, and samples were taken at the indicated time points. Samples were treated with 1 unit of DNase I (Zymo Research) for 10 min at room temperature, diluted with 1× orange gel loading dye (NEB) containing 0.2% SDS, and loaded onto a 6% native PAGE gel. RNA was visualized by SYBR gold staining (Molecular Probes).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, XLSX file, 0.01 MB.

SUPPLEMENTAL FILE 2, PDF file, 1 MB.

ACKNOWLEDGMENTS

This work was supported by the U.S. National Science Foundation (grant 2041202) and the Kennesaw State University Foundation (Foundation Fellow in Biochemistry). This material is based upon work supported by the NSF Postdoctoral Research Fellowships in Biology Program under grant 2208795.

REFERENCES

1. Dudev T, Lin YL, Dudev M, Lim C. 2003. First-second shell interactions in metal binding sites in proteins: a PDB survey and DFT/CDM calculations. J Am Chem Soc 125:3168–3180. https://doi.org/10.1021/ja0209722.
2. Troxell B, Hassan HM. 2013. Transcriptional regulation by ferric uptake regulator (Fur) in pathogenic bacteria. Front Cell Infect Microbiol 3:59. https://doi.org/10.3389/fcimb.2013.00059.
3. Pinochet-Barros A, Helmann JD. 2020. Bacillus subtilis Fur is a transcriptional activator for the PerR-repressed. J Bacteriol 202:e00697-19. https://doi.org/10.1128/JB.00697-19.
4. Choi SH, Lee KL, Shin JH, Cho YB, Cha SS, Roe JH. 2017. Zinc-dependent regulation of zinc import and export genes by Zur. Nat Commun 8:15812. https://doi.org/10.1038/ncomms15812.
5. Huang DL, Tang DJ, Liao Q, Li HC, Chen Q, He YQ, Feng JX, Jiang BL, Lu GT, Chen B, Tang JL. 2008. The Zur of Xanthomonas campestris functions as a repressor and an activator of putative zinc homeostasis genes via recognizing two distinct sequences within its target promoters. Nucleic Acids Res 36:4295–4309. https://doi.org/10.1093/nar/gkn328.
6. Sevilla E, Bes MT, Peleato ML, Fillat MF. 2021. Fur-like proteins: beyond the ferric uptake regulator (Fur) paralog. Arch Biochem Biophys 701:108770. https://doi.org/10.1016/j.abb.2021.108770.
7. Fillat MF. 2014. The FUR (ferric uptake regulator) superfamily: diversity and versatility of key transcriptional regulators. Arch Biochem Biophys 546:41–52. https://doi.org/10.1016/j.abb.2014.01.029.
8. Escolar L, Pérez-Martín J, de Lorenzo V. 2000. Evidence of an unusually long operator for the fur repressor in the aerobactin promoter of Escherichia coli. J Biol Chem 275:24709–24714. https://doi.org/10.1074/jbc.M002839200.
9. de Lorenzo V, Wee S, Herrero M, Neilands JB. 1987. Operator sequences of the aerobactin operon of plasmid ColV-K30 binding the ferric uptake regulation (fur) repressor. J Bacteriol 169:2624–2630. https://doi.org/10.1128/jb.169.6.2624-2630.1987.
10. Sakamoto K, Agari Y, Agari K, Kuramitsu S, Shinkai A. 2010. Structural and functional characterization of the transcriptional repressor CsoR from Thermus thermophilus HB8. Microbiology (Reading) 156:1993–2005. https://doi.org/10.1099/mic.0.037382-0.
11. Louvel H, Kanai T, Atomi H, Reeve JN. 2009. The Fur iron regulator-like protein is cryptic in the hyperthermophilic archaeon Thermococcus kodakaraensis. FEMS Microbiol Lett 295:117–128. https://doi.org/10.1111/j.1574-6968.2009.01594.x.
12. Fujino Y, Nagayoshi Y, Iwase M, Yokoyama T, Ohshima T, Doi K. 2016. Silica-induced protein (Sip) in thermophilic bacterium Thermus thermophilus responds to low iron availability. Appl Environ Microbiol 82:3198–3207. https://doi.org/10.1128/AEM.04027-15.
13. Ul Hussain Shah AM, Zhao Y, Wang Y, Yan G, Zhang Q, Wang L, Tian B, Chen H, Hua Y. 2014. A Mur regulator protein in the extremophilic bacterium Deinococcus radiodurans. PLoS One 9:e106341. https://doi.org/10.1371/journal.pone.0106341.
14. Van Dyke MW, Beyer MD, Clay E, Hiam KJ, McMurry JL, Xie Y. 2016. Identification of preferred DNA-binding sites for the Thermus thermophilus transcriptional regulator SbtR by the combinatorial approach REPSA. PLoS One 11:e0159408. https://doi.org/10.1371/journal.pone.0159408.
15. Bailey TL, 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.
16. Orchard K, May GE. 1993. An EMSA-based method for determining the molecular weight of a protein-DNA complex. Nucleic Acids Res 21:3335–3336. https://doi.org/10.1093/nar/21.14.3335.
17. 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.
18. Dehal PS, Joachimiak MP, Price MN, Bates JT, Baumohl JK, Chivian D, Friedland GD, Huang KH, Keller K, Novichkov PS, Dubchak IL, Alm EJ, Arkin AP. 2010. MicrobesOnline: an integrated portal for comparative and functional genomics. Nucleic Acids Res 38:D396–D400. https://doi.org/10.1093/nar/gkp919.
19. Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2016. GenBank. Nucleic Acids Res 44:D67–D72. https://doi.org/10.1093/nar/gkv1276.
20. Carpenter BM, Whitmire JM, Merrell DS. 2009. This is not your mother's repressor: the complex role of fur in pathogenesis. Infect Immun 77:2590–2601. https://doi.org/10.1128/IAI.00116-09.
21. Huang M, Liu M, Liu J, Zhu D, Tang Q, Jia R, Chen S, Zhao X, Yang Q, Wu Y, Zhang S, Huang J, Ou X, Mao S, Gao Q, Sun D, Wang M, Cheng A. 2021. Functional characterization of Fur in iron metabolism, oxidative stress

resistance and virulence of Riemerella anatipestifer. Vet Res 52:48. https://doi.org/10.1186/s13567-021-00919-9.
22. Gilston BA, Wang S, Marcus MD, Canalizo-Hernández MA, Swindell EP, Xue Y, Mondragón A, O'Halloran TV. 2014. Structural and mechanistic basis of zinc regulation across the E. coli Zur regulon. PLoS Biol 12:e1001987. https://doi.org/10.1371/journal.pbio.1001987.
23. Ma Z, Gabriel SE, Helmann JD. 2011. Sequential binding and sensing of Zn(II) by Bacillus subtilis Zur. Nucleic Acids Res 39:9130–9138. https://doi.org/10.1093/nar/gkr625.
24. Sigel RK, Sigel H. 2010. A stability concept for metal ion coordination to single-stranded nucleic acids and affinities of individual sites. Acc Chem Res 43:974–984. https://doi.org/10.1021/ar900197y.
25. Ouameur AA, Arakawa H, Ahmad R, Naoui M, Tajmir-Riahi HA. 2005. A comparative study of Fe(II) and Fe(III) interactions with DNA duplex: major and minor grooves bindings. DNA Cell Biol 24:394–401. https://doi.org/10.1089/dna.2005.24.394.
26. Solovyev V, Salamov A. 2011. Automatic annotation of microbial genomes and metagenomic sequences, p 61–78. *In* Li RW (ed), Metagenomics and its applications in agriculture, biomedicine and environmental studies. Nova Science Publishers, Hauppauge, NY.
27. Hoseki J, Yano T, Koyama Y, Kuramitsu S, Kagamiyama H. 1999. Directed evolution of thermostable kanamycin-resistance gene: a convenient selection marker for Thermus thermophilus. J Biochem 126:951–956. https://doi.org/10.1093/oxfordjournals.jbchem.a022539.
28. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402. https://doi.org/10.1093/nar/25.17.3389.
29. Patzer SI, Hantke K. 1998. The ZnuABC high-affinity zinc uptake system and its regulator Zur in Escherichia coli. Mol Microbiol 28:1199–1210. https://doi.org/10.1046/j.1365-2958.1998.00883.x.
30. Lewis DA, Klesney-Tait J, Lumbley SR, Ward CK, Latimer JL, Ison CA, Hansen EJ. 1999. Identification of the znuA-encoded periplasmic zinc transport protein of Haemophilus ducreyi. Infect Immun 67:5060–5068. https://doi.org/10.1128/IAI.67.10.5060-5068.1999.
31. Desrosiers DC, Bearden SW, Mier I, Abney J, Paulley JT, Fetherston JD, Salazar JC, Radolf JD, Perry RD. 2010. Znu is the predominant zinc importer in Yersinia pestis during in vitro growth but is not essential for virulence. Infect Immun 78:5163–5177. https://doi.org/10.1128/IAI.00732-10.
32. Kandari D, Gopalani M, Gupta M, Joshi H, Bhatnagar S, Bhatnagar R. 2018. Identification, functional characterization, and regulon prediction of the zinc uptake regulator (zur) of Bacillus anthracis: an insight into the zinc homeostasis of the pathogen. (Front Microbiol 9:3314. https://doi.org/10.3389/fmicb.2018.03314.
33. Ellison ML, Farrow JM, Farrow JM, Parrish W, Danell AS, Pesci EC. 2013. The transcriptional regulator Np20 is the zinc uptake regulator in Pseudomonas aeruginosa. PLoS One 8:e75389. https://doi.org/10.1371/journal.pone.0075389.
34. Saxena R, Dhakan DB, Mittal P, Waiker P, Chowdhury A, Ghatak A, Sharma VK. 2016. Metagenomic analysis of hot springs in central India reveals hydrocarbon degrading thermophiles and pathways essential for survival in extreme environments. Front Microbiol 7:2123. https://doi.org/10.3389/fmicb.2016.02123.
35. Mikhaylina A, Ksibe AZ, Scanlan DJ, Blindauer CA. 2018. Bacterial zinc uptake regulator proteins and their regulons. Biochem Soc Trans 46:983–1001. https://doi.org/10.1042/BST20170228.
36. Lucarelli D, Russo S, Garman E, Milano A, Meyer-Klaucke W, Pohl E. 2007. Crystal structure and function of the zinc uptake regulator FurB from Mycobacterium tuberculosis. J Biol Chem 282:9914–9922. https://doi.org/10.1074/jbc.M609974200.
37. Yang X, Wang Y, Liu G, Deng Z, Lin S, Zheng J. 2022. Structural basis of Streptomyces transcription activation by zinc uptake regulator. Nucleic Acids Res 50:8363–8376. https://doi.org/10.1093/nar/gkac627.
38. Shin JH, Helmann JD. 2016. Molecular logic of the Zur-regulated zinc deprivation response in Bacillus subtilis. Nat Commun 7:12612. https://doi.org/10.1038/ncomms12612.
39. Yokoyama S, Hirota H, Kigawa T, Yabuki T, Shirouzu M, Terada T, Ito Y, Matsuo Y, Kuroda Y, Nishimura Y, Kyogoku Y, Miki K, Masui R, Kuramitsu S. 2000. Structural genomics projects in Japan. Nat Struct Biol 7:943–945. https://doi.org/10.1038/80712.
40. Kanehisa M, Goto S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30. https://doi.org/10.1093/nar/28.1.27.
41. Barrows JK, Van Dyke MW. 2022. Biolayer interferometry for DNA-protein interactions. PLoS One 17:e0263322. https://doi.org/10.1371/journal.pone.0263322.
42. Van Dyke M. 2021. Expression and purification of thermostable proteins expressed in *E. coli*. Protocolsio. https://doi.org/10.17504/protocols.io.bq4jmyun.