bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A hyperthermoactive-Cas9 editing tool reveals the role of a unique arsenite methyltransferase in the arsenic resistance system of *Thermus thermophilus* HB27

Giovanni Gallo[ad#], Ioannis Mougiakos[be#], Mauricio Bianco[b], Miriam Carbonaro[a], Andrea Carpentieri[c], Anna Illiano[f], Pietro Pucci[f], Simonetta Bartolucci[a], John van der Oost[b*], Gabriella Fiorentino[adg]*

[a] Department of Biology, University of Naples Federico II, Via Cinthia 21, 80126 Napoli, Italy

[b] Laboratory of Microbiology, Wageningen University & Research, Stippeneng 4, 6708WE Wageningen, The Netherlands

[c] Department of Chemical Sciences, University of Naples Federico II, Via Cinthia 21, 80126 Napoli, Italy

[d] Consiglio Nazionale delle Ricerche CNR, Institute of Polymers, Composites and Biomaterials (IPCB), via Campi Flegrei, 34, 80078 Pozzuoli, Na, Italy

[e] Helmholtz Institute for RNA-based Infection Research (HIRI), Helmholtz-Centre for Infection Research (HZI), 97080 Würzburg, Germany

[f] CEINGE-Biotecnologie avanzate, Via Comunale Margherita, 484-538, 80131 Napoli NA

[g]Task Force on Microbiome Studies, University of Naples Federico II, Naples, Italy

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[#] These authors contributed equally to this work

*Corresponding authors

e-mail: fiogabri@unina.it, john.vanderoost@wur.nl

Running title: New players in thermophilic arsenic resistance system

Keywords: Arsenic resistance; thermoresistance; extremophiles; *Thermus thermophilus*; CRISPR-Cas9 genome editing; genetic tool; bioreporter.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

Arsenic detoxification systems can be found in a wide range of organisms, from bacteria to man. In a previous study, we discovered an arsenic-responsive transcriptional regulator in the thermophilic bacterium *Thermus thermophilus* HB27 (*Tt*SmtB). Here, we characterize the arsenic resistance system of *T. thermophilus* in more detail. We employed *Tt*SmtB-based pull-down assays with protein extracts from cultures treated with arsenate and arsenite to obtain an S-adenosyl-Lmethionine (SAM)-dependent arsenite methyltransferase (*Tt*ArsM). *In vivo* and *in vitro* analyses were performed to shed light on this new component of the arsenic resistance network and its peculiar catalytic mechanism. Heterologous expression of *TtarsM* in *Escherichia coli* resulted in arsenite detoxification at mesophilic temperatures. Although *Tt*ArsM does not contain a canonical arsenite binding site, the purified protein does catalyze SAM-dependent arsenite methylation with formation of monomethylarsenite (MMAs) and dimethylarsenite (DMAs). In addition, *in vitro* analyses confirmed the unique interaction between *Tt*ArsM and *Tt*SmtB. Next, a highly efficient ThermoCas9-based genome-editing tool was developed to delete the *Tt*ArsM-encoding gene on the *T. thermophilus* genome and to confirm its involvement in the arsenite detoxification system. Finally, the *TtarsX* efflux pump gene in the *T. thermophilus* Δ*TtarsM* genome was substituted by a gene encoding a stabilized

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

yellow fluorescent protein (sYFP) to create a sensitive genome-based bioreporter system for the detection of arsenic ions.

Importance

We here describe the discovery of an unknown protein by using a proteomic approach with a functionally related protein as bait. Remarkably, we successfully obtained a novel type of enzyme through the interaction with a transcription regulator, controlling the expression of this enzyme. Employing this strategy, we isolated *Tt*ArsM, the first thermophilic prokaryotic arsenite methyltransferase, as a new enzyme of the arsenic resistance mechanism in *T. thermophilus* HB27. The atypical arsenite binding site of *Tt*ArsM categorizes the enzyme as the first member of a new arsenite methyltransferase type, exclusively present in the *Thermus* genus. The enzyme methylates arsenite producing MMAs and DMAs. Furthermore, we developed an hyperthermophilic Cas9-based genome-editing tool, active up to 65°C. The tool allowed us to perform highly efficient, marker-free modifications (either gene deletion or insertion) in the *T. thermophilus* genome. With these modifications, we confirmed the critical role of *Tt*ArsM in the arsenite detoxification system and developed a sensitive *whole cell* bioreporter for arsenic ions. We anticipate that the developed tool can be easily adapted for editing the genomes of other thermophilic bacteria, significantly

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

boosting fundamental and metabolic engineering in hyperthermophilic microorganisms.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction

Arsenic is the most abundant environmental toxic element which enters the biosphere mainly from geochemical and (to a lesser extent) anthropogenic sources such as herbicides, growth promoters for livestock, and industrial activities (1). Arsenic has two relevant oxidation states, trivalent arsenite As(III) and pentavalent arsenate As(V). Methylated arsenicals include mono- (MAs), di- (DMAs), and tri- (TMAs) methylated forms. In general, trivalent states are more toxic than the pentavalent ones, and TMAs is more toxic than inorganic arsenite. Although arsenic is not beneficial for life, it can enter cells through transporters such as aquaglyceroporins. Hence, arsenic detoxification systems can be found in a wide range of organisms, from bacteria to humans. Arsenic resistance genes (*ars*) include genes encoding efflux transporters, redox enzymes, methyltransferases, transcriptional repressors and biosynthetic pathways for arsenosugars and arsenolipids (2, 3). The identification and characterization of these pathways have attracted the attention of fundamental, evolutionary, and biotechnological research (4–6).

Microorganisms have been exposed to arsenic since the origin of life and consequently have evolved arsenic resistance systems, encoded by genes generally clustered in operons (7). The organization and number of the operons of arsenic resistance genes are highly variable between different species (8, 9), reflecting differences in the level of arsenic resistance. The key players of arsenic

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

detoxification are: (i) arsenate reductases (ArsC) that reduce intracellular arsenate to arsenite (10), (ii) efflux permeases responsible for arsenite transport outside the cell (11), and (iii) transcriptional repressors that are generally metalloregulatory proteins of the ArsR/SmtB family (12). In addition, arsenite can be methylated by arsenite *S*-adenosylmethionine methyltransferases (ArsM) into MMAs, DMAs and TMAs (13), after which they can passively leave the cell or be extruded by methylarsenite-specific efflux permease (ArsP) (14).

Based on recent molecular clock analyses, it has been concluded that arsenite efflux and arsenite methylation represented the core of microbial arsenic resistance systems before the rise of atmospheric oxygen (15). In such primordial anoxic environments, methyl-arsenicals could also have the function as antibiotics against competitor microbes; after the rise of atmospheric oxygen, the ArsM enzymes did become primary components of the arsenic detoxification machinery; nevertheless, in some microorganisms, they maintained their antibiotic activity (16, 17).

Hydrothermal hot springs, which can be considered environments with conditions similar to niches of primordial Earth, may contain high amounts of arsenic. In these cases, these hot springs are niches for arsenic-tolerant microorganisms, which play a critical role in the global arsenic biogeochemical cycle (18). Although the resistance mechanisms to inorganic arsenic have been studied in many microorganisms (19), the contribution of organo-arsenical

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

biotransformation in extreme environments is still at a stage of infancy. In this regard, only two algal thermoactive ArsM enzymes have been characterized to date (20).

The thermophilic bacterium *Thermus thermophilus* HB27, originally isolated from a volcanic hot spring in Japan (21), has an unusual genetic organization of its machinery to cope with arsenic toxicity. The currently identified arsenic resistance genes are randomly scattered in its genome (22), complicating the identification of all the genes involved. In previous studies, we elucidated some components of the arsenic resistance system of *T. thermophilus* HB27. We identified and characterized *Tt*SmtB, the metalloregulatory transcriptional repressor that is responsible for the regulation of the arsenic detoxification system. *Tt*SmtB recognizes and firmly binds to operator sequences in the promoter regions of the arsenite efflux gene (*TtarsX*) (23) and the arsenate reductase gene (*TtarsC*) (24, 25), efficiently repressing their transcription in the absence of arsenic ions. *Tt*SmtB and *Tt*ArsX are also involved in cadmium sensing and export, respectively (23). In two *T. thermophilus* HB27 deletion mutant strains (Δ*TtarsX* and Δ*TtsmtB*), tolerance to arsenate, arsenite and cadmium was significantly reduced compared to the wild type strain. Although these analyses confirmed the involvement of *Tt*ArsX and *Tt*SmtB in the promiscuous resistance mechanism, the mutant strains could still grow at concentrations of arsenic up to 3 mM (22, 23). Notably, the genome of *T.*

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

thermophilus HB27 is not predicted to express arsenite methyltransferases or arsenite oxidases, suggesting the existence of unidentified component(s) of the *T. thermophilus* HB27 arsenic resistance system that cannot be predicted by *in silico* approaches, highlighting the need to employ an alternative experimental identification method.

Since members of the ArsR/SmtB family are a group of homodimeric proteins with a common HTH-winged helix DNA binding domain and heterogeneous metal-binding domain architectures and interaction modes (26), we hypothesized that *Tt*SmtB could even form protein interactions with unknown, functionally-related protein partners.

In this study, using an integrated proteomic, biochemical and genetic approach, we provide a gain of insight into the arsenic resistance system of *T. thermophilus* HB27. We report the discovery of the first *T. thermophilus* HB27 arsenite methyltransferase, *Tt*ArsM. Moreover, we describe the development of a highly efficient, markerless Cas9-based genome-editing tool at temperatures up to 65°C. Using this ThermoCas9 system, we demonstrated the *in vivo* involvement of *Tt*ArsM in arsenite detoxification. The newly developed genome editing tool was further validated constructing a very sensitive *whole cell* bioreporter system in which the *TtarsX* efflux transporter gene was substituted by a gene encoding a thermo-adapted superfolder yellow fluorescent protein (*syfp*) (27).

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Results

Exploring the protein-protein interactions of *Tt*SmtB

A combined comparative and functional proteomic approach was employed to identify putative *Tt*SmtB interacting proteins with a role in arsenite metabolism/detoxification. Purified recombinant His-tagged *Tt*SmtB was bound to a Ni^{2+}/NTA resin for protein pull-down assays using *T. thermophilus* HB27 cell-free extracts (CFE) from cultures exposed either to arsenite or arsenate, or untreated CFE cultures used as control. SDS-PAGE separation of the pulled-down proteins eluted with 0.5 M imidazole, followed by LC-electrospray ionization (ESI)-MS/MS (28) and comparative analysis of the acquired data resulted in the identification of 51 cytosolic proteins that interact with *Tt*SmtB (Table S1). Only five of these proteins are simultaneously present in CFE from cultures exposed to arsenite and arsenate but not in control CFE from the non-exposed cultures. Amongst these proteins, TTC0109 (Accession No. AAS80457, UniProt code Q72LF0) was predicted to be involved in the arsenic detoxification system (based on homology to annotated ArsR family transcriptional regulators) and to contain a C-terminal SAM-dependent methyltransferase domain (based on homology to annotated methyltransferase domain-containing proteins), suggesting a role in arsenic methylation. To date, there are no annotated arsenite SAM-dependent methyltransferases in the genome of *T. thermophilus* HB27 or the genomes of other thermophilic bacteria. Hence, we selected TTC0109 for

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

further investigation as a potential novel arsenite methyltransferase.

Bioinformatic analysis of TTC0109

BlastP analysis of TTC0109 translated sequence with sequenced microbial genomes and evolutionary analysis conducted with MEGA X, demonstrated that TTC0109 is highly conserved among the members of the *Thermus* genus (Fig. 1 A). Moreover, multiple sequence alignment of TTC0109 with characterized prokaryotic arsenite methyltransferases (Fig. 1 B) showed that all the aligned proteins contain a typical Rossman fold (29). This fold contains a GxGxG motif in a loop region, which presumably interacts with the carboxypropyl moiety of SAM, and a highly conserved aspartic acid residue at the end of the β2 strand which forms hydrogen bonds with the ribose hydroxyls of the cofactor (30). In the case of TTC0109, the predicted GxGxG motif is composed of G114, T115, G116, T117, G118 residues and the conserved aspartic residue is D135 (29, 30). On the other hand, the alignment shows that TTC0109 greatly differ from characterized ArsM proteins in the remaining sequence (Fig. 1 B); nonetheless, if TTC0109 is an arsenite methyltransferase, it would be evolutionarily distant from other archaeal and bacterial arsenite methyltransferases, as shown in the phylogenetic tree (Fig. 1 A), and therefore it could belong to a new type of arsenite methyltransferase. Additionally, all the known arsenite methyltransferases, including those in the alignment, possess at least two, usually

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

three, cysteines that are responsible for the binding of arsenite and its subsequent methylation (16, 31). TTC0109 contains a single cysteine residue at position 77, which is perfectly conserved in all the sequences analyzed (Fig. 1 B); hence, TTC0109 could be an arsenite methyltransferase with a distinct reaction mechanism.

Reasoning that the TTC0109 structure could provide more information regarding the function of TTC0109, we generated a structural model of the protein and performed molecular docking with arsenite and SAM (Fig. S1). The obtained model predicts that TTC0109 forms homodimers via its N-terminal moiety; molecular docking highlighted that arsenite could be coordinated by two histidines, H40 and H179, while C77 interacts with the methyl group of SAM (Fig. S1 A).

Although H40 and H179 residues are not conserved in characterized ArsM proteins, they are maintained at an identical position in the translated genomes of all *Thermus* species (five of them are shown in Fig. 1B); moreover, H40 is encompassed in a sequence motif (34-YRVFPT**H**SE-42) that shares 45% identity (underlined) with a sequence motif (101-YRLADR**H**VE-109) at the C-terminus of the *Tt*SmtB metal-binding site (32), strengthening the hypothesis that H40 could be involved in arsenite binding, and suggesting an evolutionary connection between the two proteins.

We proceeded with the generation of structural models of mutant proteins in

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

which the amino acids H40 or H179 were replaced with alanine residues producing TTC0109 H40A, TTC0109 H179A and C77 replaced with a serine residue producing TTC0109 C77S (Fig. S1 B, C, D). The predicted models showed that the substitution of either H40 or H179 with an alanine residue altered the 3D structure of TTC0109, whereas the effect of the cysteine to serine substitution had only a minimal effect (Fig. S1 C, D), supporting the hypothesis that this residue could have a functional role.

TTC0109 is a novel arsenite methyltransferase

A recombinant His-tagged version of TTC0109 was produced and purified to homogeneity from *E. coli* BL21-CodonPlus (DE3)-RIL cells transformed with pET30b(+)*/TtarsM* (predicted mass: 29.1 kDa) (Fig. S2 A). Gel filtration chromatography analysis agreed with the *in silico* predicted dimeric configuration of the protein, showing that the homodimer has a mass of approximately 64.5 kDa (Fig. S2 B). To determine whether TTC0109 had arsenite methyltransferase activity, a coupled spectrophotometric enzymatic assay based on the formation of S-adenosylhomocysteine (SAH) from SAM after transfer of methyl-group(s) on the substrate was employed (33, 34). In this case, the acceptor of methyl groups was As(III). SAH is degraded by SAH nucleosidase into S-ribosylhomocysteine and adenine; adenine deaminase acts on adenine producing hypoxanthine, which is converted into urate and hydrogen

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

peroxide (H_2O_2) by xanthine oxidase. The rate of production of H_2O_2 is measured by an increase in absorbance at 510_{nm} with the help of the colorimetric reagent 3,5-dichloro-2-hydroxybenzensulfonic acid (DHBS). Then, arsenite methyltransferase activity was assayed following the increase in absorbance at 510_{nm} (35). Preliminary assays were set up to assess the thermal stability of the different components, and consequently the optimal assay temperature; afterwards, the saturating concentrations of SAM and arsenite were determined. Therefore, the optimal assay conditions resulted: 50°C, 200 μM arsenite, 800 μM SAM and 3.1 μM of TTC0109. Under these conditions, the specific arsenite methyltransferase activity of TTC0109 was 4.5 mU/mg (Fig. 2 A). For this reason, hereafter, the TTC0109 protein will be denoted as *Tt*ArsM.

In order to characterize which products are formed upon As(III) methylation by *Tt*ArsM, we incubated 10 μM of *Tt*ArsM with As(III), GSH and SAM, at 65°C for 24 h; the mixture was then solubilized and analyzed by GC-MS. The results obtained are shown in Figure 2 C. Two sharp peaks at 2.92 min and 2.96 min attributed to MMAs and DMAs, respectively, are only visible in the enzyme mixture, as no peaks were recorded for the control sample (Fig. 2 B). The fragmentation spectra of MMAs and DMAs are reported in Fig 2 D and E. Peak areas corresponding to MMAs and DMAs were manually integrated, and the results are summarized in Figure S2 C.

The results of *in vitro* assays using purified *Tt*ArsM protein confirmed the ability

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

of the protein to methylate As(III) producing mono and dimethylated arsenic, the latter being the primary product. The oxidation state of the products could not be determined because the reactions were terminated with H_2O_2, which oxidized all arsenicals to pentavalent states.

Since the *in silico* predictions of *Tt*ArsM led to the hypothesis that C77, H40 and H179 were catalytic amino acids, three mutated versions of *TtarsM*, namely *TtarsM* C77S, *TtarsM* H40A and *TtarsM* H179A, were constructed, expressed in *E. coli* BL21-CodonPlus (DE3)-RIL cells and the corresponding *Tt*ArsM mutants were purified (Fig S3). Although the expression levels of the three mutant proteins are comparable (Fig S3), it was not possible to perform *in vitro* characterization of purified *Tt*ArsM H40A and *Tt*ArsM H179A, which precipitated in solution after purification; this phenomenon is probably due to protein instability, thus indicating the importance of these amino acids for *Tt*ArsM structure. On the other hand, the purification of soluble *Tt*ArsM C77S protein was possible, albeit with a lower yield compared to the wild type *Tt*ArsM. Nonetheless, using the previously mentioned coupled assay, this mutant enzyme did not show any *in vitro* arsenite methyltransferase activity, confirming that C77 residue plays a role in *Tt*ArsM activity (Fig. 2 A). These *in vitro* results demonstrated that *Tt*ArsM has an arsenite methyltransferase activity and its distinct active site suggests a novel reaction mechanism when compared to other characterized arsenite methyltransferases (13, 36, 37).

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

*Tt*SmtB interacts with *Tt*ArsM and binds to its promoter

This is the first study to report the protein-protein interaction of an ArsR/SmtB transcriptional regulator with a member of the arsenic detoxification system as identified by pull-down and mass spectrometry. For this reason, we decided to confirm the physical interaction between *Tt*SmtB and *Tt*ArsM and to investigate the effect of different metals on *Tt*SmtB:*Tt*ArsM interaction. A Co-IP assay was carried out upon incubation of purified *Tt*ArsM and *Tt*SmtB either in the presence or in the absence of arsenite, arsenate, cadmium and antimony. The first three ions are *Tt*SmtB effectors as their interaction weakens the binding to target promoters, while antimony has no effect on DNA recognition (32). Immunoprecipitation with anti-*Tt*SmtB antibodies, followed by detection of the His-tagged *Tt*ArsM by anti-His-tag antibodies, showed that the two proteins interact and form a complex in the absence of arsenite and arsenate, confirming the existence of physical interaction between them (Fig. 3). No band was detected when the immunoprecipitation was carried out with the unrelated control protein *Tt*GalA (38). Increasing arsenate and arsenite concentrations negatively affected the stability of *Tt*SmtB:*Tt*ArsM complex; in fact, densitometric analysis of the Western-blot revealed up to a 3-fold decrease in the intensity of the band corresponding to the complex (at 1:100 protein:arsenic ratio) in the presence of both arsenate (Fig. 3 A) and arsenite (Fig. 3 B). Interestingly, the presence of cadmium had the opposite effect, enhancing by up to 2-fold the band intensity

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(Fig. 3 C), suggesting that the interaction of this metal with the complex occurs with a different mechanism. Finally, the presence of antimony had a negligible effect on complex stability, in agreement with previous data showing that this metal ion is not an effector for *Tt*SmtB (32) (Fig. 3 D).

Since *Tt*SmtB is the transcriptional repressor of the genes involved in arsenic and cadmium resistance in *T. thermophilus* HB27 (22, 23), we hypothesized that it could also regulate *TtarsM* transcription. Sequence analysis of *TtarsM* promoter (p_{arsM}), a 108 bp-long region upstream of *TtarsM* and encompassing the translation start codon revealed the presence of an inverted repeat region [GAAC(N14)CTTG] between positions -6 and -27 upstream of the start codon. The sequence overlaps -10 and -35 putative basal promoter region, is 100% identical to the *TtarsX* operator recognized by *Tt*SmtB and matches the consensus binding sites of ArsR/SmtB proteins (39). Hence, we performed EMSA to investigate the capacity of purified *Tt*SmtB to bind to the promoter region of *TtarsM*. *Tt*SmtB binds to p_{arsM} in a concentration-dependent manner, as shown by the gradual formation of lower mobility complexes and the gradual decrease of residual unbound DNA (Fig. 4 A, lanes 2-5); at 10 μM protein, the complex hardly enters the gel, suggesting the formation of multiple dimers associated with target DNA (Fig. 4 A lane 6). This observation suggests that by interacting with the regulatory region *Tt*SmtB controls *TtarsM* transcription in a way comparable to that already reported for other arsenic resistance genes, i.e. the arsenate

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

reductase, the arsenite/cadmium efflux transporter and itself (22, 23).

Since the existence of a physical interaction between *Tt*SmtB and *Tt*ArsM was established, we asked whether *Tt*ArsM influenced *Tt*SmtB interaction with $p_{arsM.}$ Therefore, we pre-incubated 3 μM of both proteins before performing an EMSA in the same conditions described above. Interestingly, when the two proteins are co-incubated, shifted bands can be observed (Fig 4 B lane 3, complex 3) corresponding to complexes of higher molecular weight in comparison to those generated or not by *Tt*SmtB or *Tt*ArsM, respectively (Fig 4 B lane 2 complex 2 and lane 4); this analysis indicates that *Tt*SmtB:*Tt*ArsM multimeric complexes bind to the promoter and suggests that *Tt*SmtB:*Tt*ArsM protein-protein interaction may function in either transcriptional and post-transcriptional control. Notably, very few studies in bacteria report protein-protein interactions of transcriptional regulators with the product of the genes they regulate (40, 41).

In vivo activity of _Tt_ArsM and its catalytic mutants in _E. coli_

Aiming to explore the role of *Tt*ArsM in arsenite resistance *in vivo*, we challenged *E. coli* BL21-CodonPlus (DE3)-RIL strains transformed with plasmids expressing *Tt*ArsM and its catalytic mutants (*Tt*ArsM C77S, *Tt*ArsM H40A and *Tt*ArsM H179A) to grow in the presence of arsenite. Each recombinant strain was grown in the presence of different arsenite concentrations for 24-hours to determine the minimal inhibitory concentration (MIC) towards the metal ion.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

*Tt*ArsM-expressing strain appeared to be more resistant to arsenite than the control strain (MIC 6 mM and 4.5 mM, respectively). Additionally, the strains expressing mutated *Tt*ArsM were inhibited by the presence of arsenite to the same extent as the control strain (Fig. 5). This shows that the heterologous expression of *Tt*ArsM in *E. coli* increases arsenite resistance even at mesophilic temperatures, indicating the role of *Tt*ArsM in arsenite detoxification. Moreover, the result obtained with the mutant strains demonstrates the role of C77, H40 and H179 in the catalytic function of *Tt*ArsM.

Developing a hyperthermoactive Cas9 editing tool

We further aimed to investigate *in vivo* the contribution of *Tt*ArsM to the arsenite detoxification mechanism via the deletion of the *TtarsM* gene from the *T. thermophilus* HB27 genome. Nonetheless, the currently available genome editing tool for *T. thermophilus* is time-consuming, not marker-free and not always efficient (42). For this purpose, we reasoned to develop a marker-free, plasmid-based, homologous recombination (HR) Cas9 counter-selection (CS) genome editing tool for *T. thermophilus* employing ThermoCas9, a thermotolerant and thermoactive Cas9 orthologue (43).

We initially evaluated the targeting efficiency of ThermoCas9 in *T. thermophilus* HB27. Therefore, a set of 3 vectors was constructed, namely pMK-ThermoCas9-NT, pMK-ThermoCas9-sp1 and pMK-ThermoCas9-sp2 (Fig. 6 A), by cloning

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

into the pMK18 vector (44) (i) the codon-harmonized version of the *thermocas9* gene under the transcriptional control of the constitutive *nqo* promoter (42), and (ii) the sgRNA expressing module under the transcriptional control of the constitutive 16S rRNA promoter, either with a non-targeting/control spacer (NT: 5'- CTAGATCCGCAGTAACCCCATGG-3') or with spacers that target the *TtarsM* gene (sp1: 5'-GGGCGTTGGTGATGTGGGCCCTC-3' and sp2: 5'-CCACCTCCTCCTCCCGGTAAGGC-3'). The 3 vectors were used to transform *T. thermophilus* HB27, along with pMK-Pnqo-syfp vector (27) as transformation control. The cells were allowed to recover at 70°C, before being plated on selective agar plates and incubated overnight at 60°C, due to the sensitivity of pMK18 at temperatures above 65°C. The transformation efficiencies of pMK-ThermoCas9-sp1 and pMK-ThermoCas9-sp2 targeting vectors were significantly reduced compared to the transformation efficiency with the pMK-ThermoCas9-NT non-targeting vector (Fig. 6 B). Moreover, the transformation efficiency with the pMK-Pnqo-syfp vector was only slightly higher compared to the transformation efficiency with the pMK-ThermoCas9-NT vector (Fig. 6 B), which could be attributed to the significant size difference between the two vectors (7552 bp and 13554 bp, respectively). This result indicates that ThermoCas9 is expressed in *T. thermophilus* HB27 cells in an active and not toxic form, motivating the development of a ThermoCas9-based genome-editing tool.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

We set out to develop and test the efficiency of an HR ThermoCas9 based CS genome editing tool in *T. thermophilus* HB27. For this purpose, we introduced an HR template for the deletion of the *TtarsM* gene into the 3 previously described ThermoCas9 vectors. The HR template was composed of the fused 1kb upstream and downstream flanking regions of the *TtarsM* gene (Fig. 6 C). The three resulting editing vectors, namely pMK-ThermoCas9-HR-NT, pMK-ThermoCas9-HR-sp1 and pMK-ThermoCas9-HR-sp2, were transformed into *T. thermophilus* HB27 cells, recovered at 70°C, and grown on selective agar plates overnight at 60°C. Colony PCR with genome-specific primers was subsequently employed to screen several colonies for each transformation (Fig. S4, Table S2). None of the colonies from the pMK-ThermoCas9-NT transformation were clean Δ*TtarsM* mutants (0/10 colonies), and only a small number of colonies were mixed wild type/Δ*TtarsM* mutants (2/10 colonies) (Fig. S4 A, Table S2). On the other hand, almost all the screened colonies from the pMK-ThermoCas9-sp1 transformation were clean Δ*TtarsM* mutants (19/19 colonies) (Fig. S4 B, Table S2); most of the screened colonies from the pMK-ThermoCas9-sp2 transformation were clean Δ*TtarsM* mutants (13/18 colonies), and the remaining were mixed wild type/Δ*TtarsM* mutants (5/18 colonies) (Fig. S4 C, Table S2); the latter result suggests that less efficient ThermoCas9 targeting is obtained when employing spacer 2. Subsequently, DNA sequencing on randomly selected clean Δ*TtarsM* mutant colonies was performed to verify the correctness of the

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

genome editing (Fig. S4 D, Table S2).

Aiming to test the temperature limit of the developed ThermoCas9-based genome-editing tool, we repeated the editing experiment increasing the plating temperature to 65°C, corresponding to the temperature limit of the pMK18 backbone for propagation. Under these conditions, the number of colonies formed upon transformation with the pMK-ThermoCas9-sp1 and pMK-ThermoCas9-sp2 was of 3 and 5 vectors, respectively, much lower compared to the corresponding numbers when the plating temperature was 60°C (Table S2). This can be ascribed to the high ThermoCas9 targeting activity at 65°C (43) and decreased vector stability at 65°C. Nonetheless, the DNA sequence of all the screened colonies confirmed that they were clean Δ*TtarsM* mutants, demonstrating the high efficiency of the developed tool at 65˚C (Fig. 6 D).

Finally, assuming that the curing of the editing plasmid from a Δ*TtarsM* mutant strain would facilitate additional editing steps, we randomly selected a Δ*TtarsM* mutant colony for inoculation in liquid, antibiotic-free TM medium, for two culturing rounds at 65°C and then plated the cultures on TM agar plates with and without antibiotic. Multiple colonies were found on the antibiotic-free plate and no colonies on the plate supplemented with the antibiotic, demonstrating that the cells were cured from the edited vector. Seven of these colonies were randomly selected and the absence of the plasmid confirmed by colony PCR using *thermoCas9*-specific primers (Fig. S4 E).

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Therefore, a marker-less HR-ThermoCas9-based CS genome editing tool was developed for *T. thermophilus* HB27, highly efficient at temperatures up to 65°C. Using this tool, a *T. thermophilus* Δ*TtarsM* strain was constructed in less than 10 days (including the plasmid curing process), expanding the repertoire of available genetic tools for this microorganism and considerably accelerating the required time for editing its genome. To the best of our knowledge, this is the highest temperature reported for a CRISPR-Cas9 based genome-editing tool to date.

_Tt_ArsM mutant is more sensitive to arsenite

To compare the arsenic resistance of Δ*TtarsM* to that of wild type *T. thermophilus* HB27, both strains were grown in TM liquid medium with different arsenite and arsenate concentrations for 24-hours (Fig. 7). As expected, the arsenite resistance of Δ*TtarsM* was significantly lower than that of the wild type strain, with the corresponding MIC values being 18 mM and 40 mM, respectively (Fig. 7 A). Moreover, the resistance of the Δ*TtarsM* strain to arsenate was comparable to the wild type strain (42 mM and 44 mM, respectively) (Fig. 7 B), in agreement with its role in arsenite resistance. This result confirmed that the thermoactive arsenite methyltransferase *Tt*ArsM, is involved in arsenite detoxification and is a novel component of the arsenic resistance machinery.

Developing a sensitive arsenic bioreporter

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

In a previous study, we demonstrated that *Tt*ArsX is the arsenic efflux membrane protein of *T. thermophilus* HB27 and reported that a Δ*TtarsX* mutant strain is more sensitive to arsenate and arsenite (23). In this study, we wanted to ascertain whether a strain lacking both *TtarsM* and *TtarsX* would be even more sensitive to arsenic ions than the single mutants Δ*TtarsX* and Δ*TtarsM* strains, and therefore could represent an even better bioreporter strain for arsenic detection.

For this purpose, the HR ThermoCas9 based CS editing tool was employed to exchange *TtarsX* gene in the genome of the Δ*TtarsM* strain with the *syfp* reporter gene (27), setting the expression of the encoded thermotolerant yellow fluorescence protein (sYFP) under the control of the arsenic responsive *TtarsX* promoter (p_{arsX}). The employed editing vector, denoted as pMK-ThermoCas9-HR-*syfp*, contained a spacer that targets the *TtarsX* gene (5'-TTTCGACGGAGGAGGCCTTGGCC-3') and an HR-template composed of the 1kb upstream flanking genomic region of *TtarsX* followed by *syfp* and the 1kb downstream flanking genomic region of *TtarsX*. Ten colonies grown after transformation of pMK-ThermoCas9-HR-*syfp* vector into *T. thermophilus* Δ*TtarsM* cells were screened by colony PCR with genome-specific primers and sequenced; eight of them were clean *T. thermophilus* HB27 Δ*TtarsM*-Δ*TtarsX* (*syfp*) knock-in mutants (Fig. S5), also proving that the developed tool was highly efficient for gene insertions and substitutions.

The double mutant strain was challenged with different arsenite and arsenate

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

concentrations in TM liquid medium. As shown in Figure 7A and 7B, arsenite resistance is strikingly lower (0.5 mM) compared to that of the wild type (40 mM), Δ*TtarsM* (18 mM) and Δ*TtarsX* strains (3 mM) (23). Interestingly, the Δ*TtarsM*-Δ*TtarsX* (*syfp*) strain showed also lower resistance to arsenate compared to the single mutant *ΔTtarsX* (2 mM and 3 mM, respectively) (22). To evaluate the sensitivity to arsenate and arsenite of the *whole cell* bioreporter system, exponentially growing cultures of *T. thermophilus* HB27 Δ*TtarsM*-Δ*TtarsX* (*syfp*) were treated with increasing concentrations of arsenite and arsenate and the intensity of the emitted fluorescence was compared (Fig. 8). The background fluorescence of the Δ*TtarsM*-Δ*TtarsX* (*syfp*) strain was low, indicating that the system is repressed in the absence of metal ions. Moreover, the developed bioreporter system was able to detect arsenite and arsenate concentrations as low as 0.5 µM (Fig. 8). This performance substantially overtakes the detection limit of the previously developed arsenite and arsenate bioreporter system, which was based on the *T. thermophilus* Δ*TtarsX* strain and plasmid-based expression of the β-galactosidase (23).

Discussion

In this study, we aimed to identify novel proteins involved in the arsenic resistance system of *T. thermophilus* HB27 and employed *Tt*SmtB as a starting point looking beyond its transcriptional regulation activity. As *Tt*SmtB contains

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

a protein interaction domain (32), we set out to identify putative *Tt*SmtB interacting proteins with a role in arsenic metabolism/detoxification, following an immunoprecipitation and comparative proteomics approach. This strategy led to the discovery of *Tt*ArsM, the first prokaryotic thermoactive arsenite SAM-dependent methyltransferase, evolutionarily distant from other known arsenite methyltransferases.

The original structure and activity mechanism of *Tt*ArsM was explored *in silico*. Like other arsenite methyltransferases known to date, *Tt*ArsM contains a C-terminal SAM-dependent methyltransferase domain and an N-terminal domain comprising only one (instead of the usually three) conserved cysteine; the other two catalytic cysteines required for arsenite coordination might be replaced by two histidines identified by docking analysis. Interestingly, H40 is part of an arsenite binding domain typical of many ArsR metalloregulatory proteins (45) and H179 is located close to the cysteine (C177) of the arsenite methyltransferase of the extremophilic alga *Cyanidioschyzon merolae* (46). Moreover, the conservation of these three amino acids among the putative arsenite methyltransferases in the *Thermus* genus supports their possible role in catalysis and suggests an adaptation in this group of microorganisms.

Indeed, the role of the cysteine and histidine residues in the structure-function of *Tt*ArsM predicted *in silico* was demonstrated by site-directed mutagenesis as the heterologous expression in *E. coli* of *Tt*ArsM mutants conferred lower arsenite

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

resistance than *Tt*ArsM.

To date, only for a few arsenite methyltransferases, the reaction mechanism has been reported (13, 20, 36, 37, 47). Those described possess three or at least two cysteine residues present in their catalytic site that can methylate the arsenic sequentially in its trivalent form through alternating reduction and oxidative methylation reactions; noteworthy, different enzymes produce mono- di- and tri-methylated forms of arsenic in diverse amounts, highlighting a biochemical diversity in the arsenite methylation mechanism (36). The newly discovered *Tt*ArsM, highly conserved within *Thermus* genus, possesses only one cysteine, and can methylate As(III) mainly into DMAs and a smaller amount of MMAs as determined by GC-MS analysis of the products of the *in vitro* assay. To the best of our knowledge, this is the first arsenite methyltransferase functioning with a single cysteine in the active site.

The discovered transcriptional and post-translational interaction of *Tt*ArsM with the transcriptional regulator *Tt*SmtB was investigated in more detail. It was demonstrated that *Tt*SmtB binds to the promoter region of *TtarsM* and that this binding is stabilised by the *Tt*SmtB:*Tt*ArsM complex. Moreover, Co-Ip experiments confirmed the interaction of *Tt*SmtB with *Tt*ArsM and showed a reverse correlation between the stability of the complex and arsenic concentration. Presumably, the complex enhances the repression of *TtarsM* transcription in the absence of arsenic ions through a novel mechanism. Hence,

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

through this analysis, we shed light on a novel kind of interaction, rarely described for bacteria, in which the transcriptional repressor of a gene interacts with the protein product of the gene that it regulates (41, 48).

An example of an enzyme that can modulate the transcriptional activity of regulators by protein-protein interaction is reported in the cysteine metabolism of *Bacillus subtilis* where the stable complex formed by CymR (the master regulator of the system) and CysK (*O*-acetyl-L-serine-thiol-lyase) represses the transcription of the genes involved in the cysteine pathway (including *cysK* gene itself) when cysteine concentration is low. The advantage of this regulatory mechanism is that it employs enzymes that can specifically recognize their substrates or allosteric effectors; thus, enzymes and/or transcriptional regulators can act simultaneously as intracellular molecular sensors and participate to their own transcriptional regulation (40).

The role of *Tt*ArsM in *T. thermophilus* HB27 arsenic detoxification was also demonstrated *via* the construction and characterization of a *ΔTtarsM* mutant strain. For this purpose, a marker-free, homologous recombination and ThermoCas9 based counter-selection genome editing tool was developed, which was highly efficient and active at temperatures up to 65°C. Our tool equals the highest reported temperature for a Cas9-based editing tool to date (49, 50). The characterization of the *T. thermophilus ΔTtarsM* strain confirmed its expected higher sensitivity to arsenite, but not arsenate, compared to the wild type strain.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

To better define the role of *Tt*ArsM in the context of the already characterized components of the arsenic resistance system, a double mutant was constructed upon exchanging the *TtarsX* efflux pump gene in the *T. thermophilus* Δ*TtarsM* genome with the gene encoding the yellow fluorescent protein; the double mutant resulted much more sensitive to arsenite and arsenate treatment. Hence, it was demonstrated that *Tt*ArsM and *Tt*ArsX are critical players of the arsenite detoxification system. This is the first example of a successfully insertion of a heterologous gene on the *T. thermophilus* genome by genome editing. The double mutant strain was also considered as a sensitive bioreporter for the development of a *whole cell* biosensor system. Indeed, it was able to detect arsenite and arsenate concentrations as low as 0.5 µM, showing 40 times higher sensitivity when compared to the previously developed *T. thermophilus* HB27 *ΔTtarsX*-plasmid-based biosensor (23).

In conclusion, this study explores a unique strategy to identify novel enzymes and/or regulative networks in non-model bacteria and expands the repertoire of genetic systems for hyperthermophiles. In addition, this work has resulted in a gain of insight into the arsenite/arsenate detoxification mechanism, particularly in that of *T. thermophilus*. On top of that, this has allowed us to develop a highly robust and sensitive biosensor.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Materials and Methods

T. thermophilus HB27 cell-free extract preparation

T. thermophilus HB27 cultures were grown aerobically at 70°C in TM medium, as previously described (24). Once the cultures reached 0.5 OD_{600nm}, they were treated either with 8 mM $NaAsO_2$ or with 12 mM NaH_2AsO_4 (Sigma) [the used concentrations were below the previously reported MIC values for arsenate and arsenite (22, 23)] or they remained untreated. Samples were harvested from each culture, either immediately after treatment or 60 min post-treatment. The samples were centrifuged, the precipitates were resuspended in phosphate buffer (20 mM Na_3PO_4 pH 7.5) supplemented with protease inhibitor cocktail (Thermo Scientific) and the resuspended cells were lysed by sonication (10 cycles: 30''on/30''off, power 40%, Misonix® Sonicator Ultrasonic Processor XL). The lysates were centrifuged and the cell-free extracts (CFE) used for pull-down assays.

Purification of recombinant *Tt*SmtB, Immobilized Metal Affinity Chromatography (IMAC) and pull-down

C-terminal His-tagged *Tt*SmtB was purified from *E. coli* BL21-CodonPlus (DE3)-RIL cells transformed with the pET28/*Ttsmt*B vector, as previously described (22). Purified C-terminal His-tagged *Tt*SmtB (2 mg), was incubated with 200 µL of Ni^{2+}-NTA resin (Sigma-Aldrich) equilibrated in 20 mM Na_3PO_4,

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

0.5 M NaCl, 20 mM imidazole, pH 7.5 for 16 h at 4°C and then washed three times with the same buffer to remove unbound proteins. *T. thermophilus* HB27 CFE, treated with arsenite, or treated with arsenate, or not treated were incubated with the functionalized resin (Ni^{2+}-NTA/*Tt*SmtB) for 16 h at 4°C under stirring conditions. Subsequently, the resin was extensively washed and the interacting proteins were eluted with 20 mM Na_3PO_4, 0.5 M NaCl, 0.5 M imidazole; pH 7.5. As negative controls, samples of Ni^{2+}-NTA resin not functionalized with *Tt*SmtB were incubated with the same *T. thermophilus* HB27 CFE.

***In situ* hydrolysis and LC-MS/MS analysis**

The fractions eluted from the pull-down process were analyzed by 15 % SDS-PAGE and *in situ* hydrolyzed for mass spectrometry analysis. Specifically, mono-dimensional SDS-PAGE gel was coloured with Coomassie Brilliant Blue; the revealed bands were cut and de-stained with 100 µL of 0.1 M ammonium bicarbonate (AMBIC) and 130 µL of acetonitrile (ACN). Each band was hydrolyzed *in situ* with 0.1 µg/µL trypsin in 10 mM AMBIC, incubated at first for 1.5 h at 4°C and then for an additional 16 h at 37°C. The hydrolysis reactions were stopped by adding acetonitrile and 0.1% formic acid; then, the samples were filtered and dried in a Savant vacuum centrifuge before being analyzed by LC-MS/MS mass spectrometry. In detail, before analysis, the samples were dissolved in 10 µL of 0.1% formic acid, and 5 µL were directly loaded into the instrument.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Reverse-phase capillary liquid chromatography (HPLC 1200 system experiments), followed by MS analysis, was performed using a binary pump system connected to a nano-spray source of the mass spectrometer (28, 51). The latter is represented by a hybrid Q-TOF spectrometer (MS CHIP 6520 QTOF) equipped with a chip (Agilent Technologies).

***In silico* analysis**

Analysis of the LC-MS/MS data, using Mascot software (http://www.matrixscience.com/search_form_select.html) allowed the identification of putative *Tt*SmtB-interacting proteins. Among these proteins, TTC0109 (UniProt code Q72LF0), herein named *Tt*ArsM, was further analyzed using the UniProt database (http://www.uniprot.org); homologous proteins and conserved domains were identified by performing a Blast analysis (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

Phylogenetic tree of archaeal and bacterial arsenite methyltransferases, SAM-dependent methyltransferases and methyltransferase domain-containing proteins, including *Tt*ArsM from *T. thermophilus* HB27, was conducted in MEGA X (52). The amino acidic sequences used for the construction of the phylogenetic tree are *Tt*ArsM from *T. thermophilus* HB27, arsenite methyltransferase from *Rhodopseudomonas palustris* (13), *Methanosarcina acetivorans*, *Clostridium* sp BMX (53), *Halobacterium salinarum* (54), *Pseudomonas alcaligenes* (36),

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Cyanidioschyzon merolae (46); a SAM-dependent methyltransferase from 5 members of the *Thermus* genus (*islandicus*, *caldilimi*, *antranikianii*, *oshimai* and *brockianus*); a SAM-dependent methyltransferase from *Mesorhizobium amorphae*, *Anaerolineae bacterium*; a methyltransferase domain-containing protein from *Sanguinobacter* sp. Phylogenetic reconstruction was accomplished using the maximum likelihood statistical method.

The alignment of *Tt*ArsM to its templates was based on a multiple sequence alignment, performed with the program Clustal Omega (55); the amino acidic sequences used for the construction of the alignment of functionally characterized archaeal and bacterial arsenite methyltransferases are: *Tt*ArsM from *T. thermophilus* HB27, 5 members of the *Thermus* genus (*islandicus*, *caldilimi*, *antranikianii*, *oshimai* and *brockianus*), *Clostridium* sp BMX, *R. palustris*, *M. acetivorans*, *H. salinarum*, *P. alcaligenes* and *Cyanidioschyzon merolae* (46).

Models of *Tt*ArsM were generated through I-TASSER (56) (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) using as input the complete sequence of *Tt*ArsM (C-score = -2.5). The dimeric structure was predicted using the GalaxyWEB tool (http://galaxy.seoklab.org/index.html) (57). The molecular dockings of *Tt*ArsM with arsenite and SAM were generated using the Hex Protein Docking server (58). 100 rigid-body docking solutions were generated per case and the best 10 were refined by energy minimization. The proposed

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

model for the metal ion docked into *Tt*ArsM is the structure with the smallest distance between arsenite-histidine and cysteine-SAM (4.33 Å from H40 and 5.77 Å from H179 in *Tt*ArsM model and 4.40 Å from C77).

Cloning, expression and purification of recombinant *Tt*ArsM and *Tt*ArsM mutants

The pET30b(+)*/TtarsM* vector was constructed for the expression and subsequent purification of the C-terminal His-tagged version of *Tt*ArsM. For the construction of the pET30b(+)*/TtarsM*, *TtarsM* gene was PCR amplified from *T. thermophilus* HB27 genome, using Taq DNA polymerase (Thermo Fisher Scientific) and primers containing the NdeI (*arsMfw,* Table S3) and HindIII (*arsMrv,* Table S3) restriction sites at their 5'-ends. The PCR product was purified, digested with the NdeI and HindIII restriction enzymes (NEB), and ligated (T4 ligase, NEB) into NdeI/HindIII-digested pET30b(+) vector (Novagen). The ligase mixture was transformed into *E. coli* TOP10F' cells were plated on LB agar plates supplemented with 50 µg/mL kanamycin (Sigma-Aldrich). Single colonies were selected and inoculated in LB liquid medium supplemented with 50 µg/mL kanamycin. Plasmid isolation and sequencing were subsequently performed before transforming *E. coli* BL21-CodonPlus (DE3)-RIL cells with pET30b(+)*/TtarsM* vector.

To obtain mutation of *TtarsM* gene sequence at specific sites, the QuickChange

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

II-E Site-Directed Mutagenesis Kit (Agilent Technologies) was employed; pET30b(+)/*TtarsM* was used as a template and amplified with three different mutagenic primer pairs (Table S3) to get pET30b(+)/*TtarsM* C77S, pET30b(+)/*TtarsM* H40A and pET30b(+)/*TtarsM* H179A vectors. The reaction mixtures were transformed into *E. coli* TOP10F' cells were plated on LB agar plates supplemented with kanamycin (50 µg/mL). Single colonies were randomly selected and inoculated in LB liquid medium supplemented with kanamycin (50 µg/mL). Plasmid isolation was subsequently performed and *E. coli* BL21-CodonPlus (DE3)-RIL cells were transformed with sequence-verified pET30b(+)/*TtarsM* C77S, pET30b(+)/*TtarsM* H40A and pET30b(+)/*TtarsM* H179A vectors.

For protein expression of the His-tagged versions of *Tt*ArsM C77S, *Tt*ArsM H40A and *Tt*ArsM H179A catalytic mutants, the recombinant *E. coli* BL21-CodonPlus (DE3)-RIL strains were cultured in LB medium supplemented with kanamycin (50 µg/mL) and chloramphenicol (33 µg/mL). Protein expression was induced via the addition of 1 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) when the cultures reached 0.7 OD_{600nm}. The cultures were further incubated with vigorous shaking at 37°C for 16 h, then centrifuged, resuspended in lysis buffer (20 mM NaP pH 7.4, 50 mM NaCl and 20 mM imidazole) supplemented with protease inhibitor cocktail (Thermo Scientific) and lysed by sonication (10 cycles: 30''on/30''off, power 40%, Misonix® Sonicator

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Ultrasonic Processor XL). The lysates were centrifuged and the supernatants used for the purification on HisTrap HP columns (1 mL; GE Healthcare) connected to an AKTA Explorer system (GE Healthcare). The fractions containing His-tagged *Tt*ArsM proteins were eluted from the columns using a linear gradient of the elution buffer (20 mM NaP pH 7.4, 50 mM NaCl and 500 mM Imidazole). The eluted protein fractions were subjected to SDS-PAGE analysis and the fractions containing purified *Tt*ArsM were pooled and dialyzed for 16 h at 4°C in 20 mM NaP pH 7.4 buffer supplemented with protease inhibitor cocktail (Thermo Scientific). The identity of the purified His-tagged *Tt*ArsM protein was confirmed by mass spectrometry and protein aliquots were stored at -20°C.

_Tt_ArsM quaternary structure assessment

The native molecular mass of *Tt*ArsM was determined by loading 500 µg of the purified protein onto an analytical Superdex PC75 column (3.2 by 30 cm) connected to an AKTA Pure system, in 50 mM Tris-HCl, pH 7.5, 0.2 M KCl buffer. The column was calibrated using a set of gel filtration markers (low range, GE Healthcare), including Ovalbumin (43.0 kDa), Carbonic anhydrase (29.0 kDa), RNase A (13.7 kDa) and Aprotinin (6.5 kDa) as previously described (24).

Methyltransferase activity assay

According to the manufacturer's protocol, the *Tt*ArsM arsenite methyltransferase

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

activity was measured using the SAM510: SAM Methyltransferase Assay Kit (G-Biosciences) with modifications regarding the temperature, the SAM concentration and the reaction time (33–35). The assay relies on the degradation of S-adenosylhomocysteine (SAH) into urate and hydrogen peroxide by a mixture of enzymes (adenosylhomocysteine nucleosidase, adenine deaminase, xanthine oxidase). Then, the reaction of hydrogen peroxide with 4-aminoantipyrine produces 5-dichloro-2-hydroxybenzene sulfonic acid (DHBS) with ε_{mM}=15.0 at 510_{nm}. A typical reaction mixture containing: 200 µM As(III), 800 µM SAM, 3.1 µM of the enzyme, SAM enzyme mixture and SAM colorimetric mix in a final reaction volume of 115 µL, was incubated for 1h at 50°C in a Synergy™ HTX Multi-Mode Microplate Reader (BioTek). The same reaction mixture was tested with 10 µg of *Tt*ArsC or 10 µg of *Tt*SmtB, as negative controls. One unit of arsenite methyltransferase produces 1.0 µmol of DHBS per minute at 50°C under the conditions described above. Preliminary assays were performed to define substrate saturating concentrations, varying the As(III) and SAM concentrations from 50 µM to 300 µM and from 200 µM to 1.2 mM, respectively.

In vitro arsenite methylation

As(III) methylation by *Tt*ArsM was determined in an assay solution containing 10 µM *Tt*ArsM, 250 µM As(III), 6 mM glutathione (GSH) and 1 mM SAM, in

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Na-phsphate 50 mM, pH 7.4 at 65°C for 24 h; the same reaction without *Tt*ArsM was used as the negative control (control sample). The reactions were terminated by the addition of 10% (v/v) H_2O_2. The oxidized solution samples were filtered through 0.22 µm MCE syringe filters and used for GC-MS analysis.

For the detection of the methylated products, GC analyses were performed using Agilent GC 6890, coupled with a 5973 MS detector.

200 µl of the control sample and the enzyme mixture were treated with 800 µL of methanol (Sigma-Aldrich). The supernatants were recovered and dried under vacuum. The samples were solubilized in methanol (10 µl) and 1 µl was analyzed by GC-MS. The column used was an HP5 capillary (30 m × 0.25 mm, 0.25 mM, 5% polisilarilene 95% PDMS). Helium was used as the carrier gas at a rate of 1.0 mL min^{-1}. The GC injector was maintained at 230 °C, while the oven temperature was held at 40 °C for 5 min and then increased to 280 °C at 20 °C/min and held for 5 min for a total separation time of 20 min. The analyzer temperature was kept at 250 °C. The collision energy was set to a value of 70 eV, and fragment ions generated were analyzed mass range 20–450 m/z. The identification of each compound was based on the combination of retention time and fragmentation spectra matching those collected into the NIST 05 Mass Spectral Library. The identification was reliable when the matching values were higher than 700 according to the NIST guidelines (59). The analyses were performed in triplicate for each sample.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Co-Ip assay

Protein-protein interaction between *Tt*SmtB and *Tt*ArsM was *in vitro* verified via co-immunoprecipitation assays that employed recombinant *Tt*SmtB, recombinant His-tagged-*Tt*ArsM, anti-*Tt*SmtB antibodies (GeneCust™) and His-tag antibodies (Sigma-Aldrich). His-tag removal from recombinant *Tt*SmtB was performed as previously described (22).

A typical Co-Ip mixture contained 1 mL of Co-Ip buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10% glycerol and 0.1% Triton X-100), 5 µg of *Tt*SmtB, 5 µg of His-tagged *Tt*ArsM and was incubated at 4°C for 2h in continuous rotation. In some cases, arsenite, arsenate, cadmium and antimony at 1:0, 1:25, 1:50, 1:100 molar ratios pre-incubated with *Tt*SmtB for 10 min at 60°C were added. As controls, *Tt*ArsM (5 µg), *Tt*SmtB (5 µg) and *Tt*GalA (38) were also separately incubated under the same conditions. All the samples were subjected to immunoprecipitation using 2 µL of purified anti-*Tt*SmtB antibodies (2 µg/µL) (GeneCust™) for 3 h at 4°C in continuous rotation, before adding 15 µL of Protein A-Sepharose beads (Sigma-Aldrich) and allowing the incubation to continue for further16 h at 4°C. The formed immunocomplexes were washed with Co-Ip buffer and analyzed by Western Blot on 15% SDS-PAGE, using PVDF membranes (Millipore) and anti-poly-Histidine-Peroxidase antibodies (Sigma-Aldrich) diluted 1:10000, as previously described (60).

The ImageJ software (https://imagej.net/) was used for densitometric analysis of

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the formed bands, setting as maximum value the intensity of the band of the sample containing *Tt*ArsM and *Tt*SmtB in the absence of any metal. Each experiment was performed in three technical and two biological replicates; statistical analysis was performed using one-way ANOVA and significant differences are indicated as * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Electrophoretic mobility shift assay (EMSA)

Electrophoretic mobility shift assays (EMSA) were performed to determine the *in vitro* binding of *Tt*SmtB to the promoter region upstream *TtarsM*. The 108 bp chromosomal region that encompasses the start codon of the *TtarsM* gene and the 105 bp-long region upstream, denoted as p_{arsM}, was PCR amplified with Taq DNA polymerase (Thermo Fisher Scientific) using the p_{arsM} *Fw* and p_{arsM} *Rv* primers (Table S3). EMSA reactions were set up as previously described (23, 61) in the presence of 1µg of poly (dI-dC), 20 ng p_{arsM} and increasing concentrations of *Tt*SmtB (1, 2, 3, 5 and 10 µM, considering *Tt*SmtB as a dimer) using SYBR™ Gold Nucleic Acid Gel Stain for band detection. The EMSA reactions that simultaneously employed *Tt*SmtB and *Tt*ArsM were set up in identical conditions using 3 µM of each protein alone or in combination.

Arsenic tolerance of *E. coli* expressing *Tt*ArsM

The following strains were inoculated in 10 mL of LB pre-cultures, supplemented

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

with kanamycin (50 µg/mL) and chloramphenicol (33 µg/mL): i) *E. coli* BL21-CodonPlus (DE3)-RIL: pET30/*TtarsM*, ii) *E. coli* BL21-CodonPlus (DE3)-RIL: pET30/ *TtarsM C77S*, iii) *E. coli* BL21-CodonPlus (DE3)-RIL: pET30/ *TtarsM H40A*, iv) *E. coli* BL21-CodonPlus (DE3)-RIL: pET30/ *TtarsM H179A* and v) *E. coli* BL21-CodonPlus (DE3)-RIL: pET30 (control). The pre-cultures were incubated at 37°C for 16 h at 180 rpm. Subsequently, 50 mL LB cultures, supplemented with antibiotics, were inoculated with the pre-cultures to initial 0.08 OD_{600nm} and incubated at 37°C and 180 rpm until 0.6 OD_{600nm} (exponential growth). At that point, protein expression was induced with 1mM IPTG and the cultures were incubated at 37°C and 180 rpm for 3 additional hours. From these growing cells, fresh LB cultures were inoculated to 0.05 OD_{600nm} and distributed to 24-well plates (1 mL per well) containing LB medium supplemented with 1 mM IPTG, kanamycin (50 µg/mL), chloramphenicol (33 µg/mL) and increasing concentrations of arsenate and arsenite (from 2.5 mM to 7.0 mM). The minimal inhibitory concentrations (MIC) endpoint for each strain were determined as the lowest concentration of arsenite at which there was the difference between grown and start culture lower than 0.01 OD_{600nm} after 16 hours of incubation at 37°C (22). All the cells up to the MIC value were able to grow if reinoculated in an arsenic-free medium. The reported values are the average of three biological replicates.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ThermoCas9 editing and targeting constructs

The plasmids used for the ThermoCas9-based targeting and editing experiments are listed in Table S4. The vector pMK18 was used as the template for the construction of the ThermoCas9-based targeting and editing plasmids and the employed primers, the DNA templates, and the DNA fragments, which are listed in Table S4. The *thermoCas9* gene was codon harmonized, according to *T. thermophilus* HB27 codon-usage using the Galaxy/Codon harmonizer online tool (62), and it was synthesized (Twist Bioscience) (Table S5). The DNA fragments were designed with appropriate overhangs for NEBuilder HiFi DNA assembly (NEB) and they were obtained through PCR with Q5 Polymerase (NEB). The PCR products were subjected to 1% agarose gel electrophoresis and they were purified using a Zymogen gel DNA recovery kit (Zymo Research). The assembly reactions were transformed to chemically competent *E. coli* DH5α cells (NEB) and the cells were plated on LB agar plates supplemented with kanamycin (50 µg/mL). Single colonies were inoculated in LB medium supplemented with kanamycin (50 µg/mL) for overnight incubation at 37°C. Plasmid material was isolated using the GeneJet plasmid miniprep kit (Thermo Fisher Scientific), sequence verified (GATC-biotech) and 300 ng of each plasmid (pMK-ThermoCas9-NT/sp1/sp2 and pMK-ThermoCas9-HR-NT/sp1/sp2) was transformed to either *T. thermophilus* HB27 cells (22), as indicated per experimental process.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The obtained plasmid, pMK-ThermoCas9-NT, was used as the backbone to construct a new plasmid to obtain the deletion of the *TtarsX* gene (Table S4) and the insertion of the gene coding sYFP (27). The obtained plasmid (pMK-ThermoCas9-HR-syfp) was used to transform *T. thermophilus* HB27 *ΔTtarsM* cells as already described to obtain the strain denoted Δ*TtarsM*-Δ*TtarsX* (*syfp*).

Arsenic tolerance of *T. thermophilus* HB27 wild type and mutant strains

Exponentially growing pre-cultures of *T. thermophilus* HB27 (control), *T. thermophilus* HB27 Δ*TtarsM* and *T. thermophilus* HB27 Δ*TtarsM*-Δ*TtarsX* *(syfp)* were diluted to 0.08 OD_{600nm} in 10 mL TM cultures containing increasing concentrations of arsenite and arsenate (from 0.1 mM to 50 mM). The cultures were incubated aerobically at 70°C for 18 h and the MIC values were determined as the lowest concentrations of arsenite and arsenate that completely inhibited the growth of a strain (22). The reported values are the average of three biological replicates.

Bioreporter activity measurement

Overnight cultures of *T. thermophilus* HB27 Δ*TtarsM*-Δ*TtarsX* (*syfp*) strain were diluted to 0.08 OD_{600nm} in TM medium and then grown aerobically at 70°C until 0.5 OD_{600nm}. The cultures were divided into samples of 5 mL each and subsequently supplemented with increasing concentrations of arsenite and

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

arsenate (0.5 μM to 4 μM). After 1-hour of incubation at 70°C, 200 μL of each cell sample were removed and centrifuged for 5' at 6000 rpm. The pellets were washed twice with equal volumes of PBS 1X and resuspended with equal volumes of PBS 1X before being distributed into a 96-well plate. sYFP fluorescence intensity of each sample was measured employing a Synergy™ HTX Multi-Mode Microplate Reader (BioTek), using excitation and emission wavelengths of 458 nm and 540 nm, respectively. The measured fluorescence intensities were normalized for the optical density of each sample at 600 nm. The measured fluorescence was reported as fluorescence relative expression, assuming that the fluorescence value of not treated cells (control) was 1.

Each experiment was performed in three technical and biological replicates. Statistical analysis was performed using one-way ANOVA; significant differences are indicated as: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Acknowledgements

The authors kindly acknowledge Prof. Josè Berenguer from Universidad Autònoma de Madrid for kindly providing the pMK-Pnqo-syfp vector. Prof. Lucia Cavalca and Prof. Gian Attilio Sacchi from the University of Milan also acknowledged for donating disodium methyl arsonate hexahydrate (MMAs) (ChemService) and dimethylarsinic acid (DMAs) (Merck Life Science S.r.l.). The authors also acknowledge Prof. Angela Amoresano (Department of

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Chemical Sciences, University of Naples Federico II) for helpful discussion. Acknowledgements go to grant ERA-NET Cofund MarTERA, who funded the project: "FLAshMoB: Functional Amyloid Chimera for Marine Biosensing". This project has received funding from the European Research Council (ERC), under the European Union's Horizon 2020 research and innovation programme, grant agreement No. 834279, and from the Netherlands Organization for Scientific Research (NWO, Spinoza grant SPI 93-537) to JvdO.

References:

1. Mukhopadhyay R, Rosen BP, Phung LT, Silver S. 2002. Microbial arsenic: From geocycles to genes and enzymes. FEMS Microbiol Rev.
2. Chen Z, Wang Y, Jiang X, Fu D, Xia D, Wang H, Dong G, Li Q. 2017. Dual roles of AQDS as electron shuttles for microbes and dissolved organic matter involved in arsenic and iron mobilization in the arsenic-rich sediment. Sci Total Environ https://doi.org/10.1016/j.scitotenv.2016.09.006.
3. Yang H-C, Rosen BP. 2016. New mechanisms of bacterial arsenic resistance. Biomed J 39:5–13.
4. Gallo G, Puopolo R, Carbonaro M, Maresca E, Fiorentino G. 2021. Extremophiles, a Nifty Tool to Face Environmental Pollution: From Exploitation of Metabolism to Genome Engineering. Int J Environ Res

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Public Health 18:5228.

5. Politi J, Spadavecchia J, Fiorentino G, Antonucci I, Casale S, De Stefano L. 2015. Interaction of Thermus thermophilus ArsC enzyme and gold nanoparticles naked-eye assays speciation between As(III) and As(V). Nanotechnology 26:435703.

6. Politi J, Spadavecchia J, Fiorentino G, Antonucci I, De Stefano L. 2016. Arsenate reductase from *Thermus thermophilus* conjugated to polyethylene glycol-stabilized gold nanospheres allow trace sensing and speciation of arsenic ions. J R Soc Interface 13:20160629.

7. Rosen BP. 2002. Biochemistry of arsenic detoxification. FEBS Lett 529:86–92.

8. Liu Z, Rensing C, Rosen BP. 2011. Toxicity: resistance pathways for metalloids and toxic metals. Encycl Inorg Bioinorg Chem 1–13.

9. Aulitto M, Gallo G, Puopolo R, Mormone A, Limauro D, Contursi P, Piochi M, Bartolucci S, Fiorentino G. 2021. Genomic Insight of Alicyclobacillus mali FL18 Isolated From an Arsenic-Rich Hot Spring. Front Microbiol 12.

10. Mukhopadhyay R, Rosen BP. 2002. Arsenate reductases in prokaryotes and eukaryotes. Environ Health Perspect 110:745–748.

11. Chen C-M, Misra TK, Silver S, Rosen BP. 1986. Nucleotide sequence of the structural genes for an anion pump. The plasmid-encoded arsenical

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

resistance operon. J Biol Chem 261:15030–15038.

12. Wu J, Rosen BP. 1991. The ArsR protein is a trans-acting regulatory protein. Mol Microbiol 5:1331–1336.

13. Qin J, Rosen BP, Zhang Y, Wang G, Franke S, Rensing C. 2006. Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase. Proc Natl Acad Sci 103:2075–2080.

14. Chen J, Madegowda M, Bhattacharjee H, Rosen BP. 2015. ArsP: a methylarsenite efflux permease. Mol Microbiol 98:625–635.

15. Chen S-C, Sun G-X, Yan Y, Konstantinidis KT, Zhang S-Y, Deng Y, Li X-M, Cui H-L, Musat F, Popp D. 2020. The Great Oxidation Event expanded the genetic repertoire of arsenic metabolism and cycling. Proc Natl Acad Sci 117:10414–10421.

16. Chen S-C, Sun G-X, Rosen BP, Zhang S-Y, Deng Y, Zhu B-K, Rensing C, Zhu Y-G. 2017. Recurrent horizontal transfer of arsenite methyltransferase genes facilitated adaptation of life to arsenic. Sci Rep 7:1–11.

17. Chen J, Rosen BP. 2020. The arsenic methylation cycle: how microbial communities adapted methylarsenicals for use as weapons in the continuing war for dominance. Front Environ Sci.

18. Gihring TM, Druschel GK, McCleskey RB, Hamers RJ, Banfield JF.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

2001. Rapid arsenite oxidation by Thermus aquaticus and Thermus thermophilus: field and laboratory investigations. Environ Sci Technol 35:3857–3862.

19. Saltikov CW, Olson BH. 2002. Homology of Escherichia coli R773 arsA, arsB, and arsC genes in arsenic-resistant bacteria isolated from raw sewage and arsenic-enriched creek waters. Appl Environ Microbiol https://doi.org/10.1128/AEM.68.1.280-288.2002.

20. Qin J, Lehr CR, Yuan C, Le XC, McDermott TR, Rosen BP. 2009. Biotransformation of arsenic by a Yellowstone thermoacidophilic eukaryotic alga. Proc Natl Acad Sci 106:5213–5217.

21. Henne A, Brüggemann H, Raasch C, Wiezer A, Hartsch T, Liesegang H, Johann A, Lienard T, Gohl O, Martinez-Arias R, Jacobi C, Starkuviene V, Schlenczeck S, Dencker S, Huber R, Klenk HP, Kramer W, Merkl R, Gottschalk G, Fritz HJ. 2004. The genome sequence of the extreme thermophile Thermus thermophilus. Nat Biotechnol https://doi.org/10.1038/nbt956.

22. Antonucci I, Gallo G, Limauro D, Contursi P, Ribeiro AL, Blesa A, Berenguer J, Bartolucci S, Fiorentino G. 2017. An ArsR/SmtB family member regulates arsenic resistance genes unusually arranged in Thermus thermophilus HB27. Microb Biotechnol 10.

23. Antonucci I, Gallo G, Limauro D, Contursi P, Ribeiro AL, Blesa A,

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Berenguer J, Bartolucci S, Fiorentino G. 2018. Characterization of a promiscuous cadmium and arsenic resistance mechanism in Thermus thermophilus HB27 and potential application of a novel bioreporter system. Microb Cell Fact 17.

24. Del Giudice I, Limauro D, Pedone E, Bartolucci S, Fiorentino G. 2013. A novel arsenate reductase from the bacterium Thermus thermophilus HB27: its role in arsenic detoxification. Biochim Biophys Acta (BBA)-Proteins Proteomics 1834:2071–2079.

25. Puopolo R, Sorrentino I, Gallo G, Piscitelli A, Giardina P, Le Goff A, Fiorentino G. 2021. Self-assembling thermostable chimeras as new platform for arsenic biosensing. Sci Rep 11:2991.

26. Busenlehner LS, Pennella MA, Giedroc DP. 2003. The SmtB/ArsR family of metalloregulatory transcriptional repressors: structural insights into prokaryotic metal resistance. FEMS Microbiol Rev 27:131–143.

27. Blesa A, Berenguer J. 2015. Contribution of vesicle-protected extracellular DNA to horizontal gene transfer in Thermus spp. Int Microbiol 18:177–187.

28. Carpentieri A, Gamberi T, Modesti A, Amoresano A, Colombini B, Nocella M, Bagni MA, Fiaschi T, Barolo L, Gulisano M. 2016. Profiling carbonylated proteins in heart and skeletal muscle mitochondria from trained and untrained mice. J Proteome Res 15:3666–3678.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

29. Rao ST, Rossmann MG. 1973. Comparison of super-secondary structures in proteins. J Mol Biol 76:241–256.

30. Schubert HL, Blumenthal RM, Cheng X. 2003. Many paths to methyltransfer: a chronicle of convergence. Trends Biochem Sci 28:329–335.

31. Huang K, Xu Y, Packianathan C, Gao F, Chen C, Zhang J, Shen Q, Rosen BP, Zhao F. 2018. Arsenic methylation by a novel ArsM As (III) S-adenosylmethionine methyltransferase that requires only two conserved cysteine residues. Mol Microbiol 107:265–276.

32. Gallo G, Antonucci I, Pirone L, Amoresano A, Contursi P, Limauro D, Pedone E, Bartolucci S, Fiorentino G. 2019. A physicochemical investigation on the metal binding properties of TtSmtB, a thermophilic member of the ArsR/SmtB transcription factor family. Int J Biol Macromol 138.

33. Singh J, Vijay S, Mansuri R, Rawal R, Kadian K, Sahoo GC, Kumar M, Sharma A. 2019. Computational and experimental elucidation of Plasmodium falciparum phosphoethanolamine methyltransferase inhibitors: Pivotal drug target. PLoS One 14:e0221032.

34. Patil NA, Basu B, Deobagkar DD, Apte SK, Deobagkar DN. 2017. Putative DNA modification methylase DR_C0020 of Deinococcus radiodurans is an atypical SAM dependent C-5 cytosine DNA methylase.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Biochim Biophys Acta (BBA)-General Subj 1861:593–602.

35. Dou L, Yan F, Pang J, Zheng D, Li D, Gao L, Wang L, Xu Y, Shi J, Wang Q. 2019. Protein lysine 43 methylation by EZH1 promotes AML1-ETO transcriptional repression in leukemia. Nat Commun 10:1–15.

36. Zhang J, Cao T, Tang Z, Shen Q, Rosen BP, Zhao F-J. 2015. Arsenic methylation and volatilization by arsenite S-adenosylmethionine methyltransferase in Pseudomonas alcaligenes NBRC14159. Appl Environ Microbiol 81:2852–2860.

37. Wang P-P, Sun G-X, Zhu Y-G. 2014. Identification and characterization of arsenite methyltransferase from an archaeon, Methanosarcina acetivorans C2A. Environ Sci Technol 48:12706–12713.

38. Aulitto M, Fusco S, Fiorentino G, Limauro D, Pedone E, Bartolucci S, Contursi P. 2017. Thermus thermophilus as source of thermozymes for biotechnological applications: homologous expression and biochemical characterization of an α-galactosidase. Microb Cell Fact 16:1–10.

39. Osman D, Cavet JS. 2010. Bacterial metal-sensing proteins exemplified by ArsR–SmtB family repressors. Nat Prod Rep 27:668–680.

40. Tanous C, Soutourina O, Raynal B, Hullo M-F, Mervelet P, Gilles A-M, Noirot P, Danchin A, England P, Martin-Verstraete I. 2008. The CymR regulator in complex with the enzyme CysK controls cysteine metabolism in Bacillus subtilis. J Biol Chem 283:35551–35560.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

41. Fisher SH, Wray L V. 2008. Bacillus subtilis glutamine synthetase regulates its own synthesis by acting as a chaperone to stabilize GlnR-DNA complexes. PNAS 1014–1019.

42. Cava F, Laptenko O, Borukhov S, Chahlafi Z, Blas-Galindo E, Gómez-Puertas P, Berenguer J. 2007. Control of the respiratory metabolism of Thermus thermophilus by the nitrate respiration conjugative element NCE. Mol Microbiol 64:630–646.

43. Mougiakos I, Mohanraju P, Bosma EF, Vrouwe V, Bou MF, Naduthodi MIS, Gussak A, Brinkman RBL, Van Kranenburg R, Van Der Oost J. 2017. Characterizing a thermostable Cas9 for bacterial genome editing and silencing. Nat Commun 8:1–11.

44. Blesa A, Baquedano I, Quintáns NG, Mata CP, Castón JR, Berenguer J. 2017. The transjugation machinery of Thermus thermophilus: Identification of TdtA, an ATPase involved in DNA donation. PLoS Genet 13:e1006669.

45. Roy R, Samanta S, Patra S, Mahato NK, Saha RP. 2018. In silico identification and characterization of sensory motifs in the transcriptional regulators of the ArsR-SmtB family. Metallomics 10:1476–1500.

46. Ajees AA, Marapakala K, Packianathan C, Sankaran B, Rosen BP. 2012. Structure of an As (III) S-adenosylmethionine methyltransferase: insights into the mechanism of arsenic biotransformation. Biochemistry 51:5476–

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

5485.

47. Dheeman DS, Packianathan C, Pillai JK, Rosen BP. 2014. Pathway of human AS3MT arsenic methylation. Chem Res Toxicol 27:1979–1989.

48. Fusco S, Aulitto M, Iacobucci I, Crocamo G, Pucci P, Bartolucci S, Monti M, Contursi P. 2020. The interaction between the F55 virus-encoded transcription regulator and the RadA host recombinase reveals a common strategy in Archaea and Bacteria to sense the UV-induced damage to the host DNA. Biochim Biophys Acta - Gene Regul Mech 1863.

49. Le Y, Fu Y, Sun J. 2020. Genome editing of the anaerobic thermophile Thermoanaerobacter ethanolicus using thermostable Cas9. Appl Environ Microbiol 87:e01773-20.

50. Adalsteinsson BT, Kristjansdottir T, Merre W, Helleux A, Dusaucy J, Tourigny M, Fridjonsson O, Hreggvidsson GO. 2021. Efficient genome editing of an extreme thermophile, Thermus thermophilus, using a thermostable Cas9 variant. Sci Rep 11:1–15.

51. Monti M, Orru S, Pagnozzi D, Pucci P. 2005. Functional proteomics. Clin Chim acta 357:140–150.

52. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

53. Wang P-P, Bao P, Sun G-X. 2015. Identification and catalytic residues of the arsenite methyltransferase from a sulfate-reducing bacterium, Clostridium sp. BXM. FEMS Microbiol Lett 362:1–8.

54. Wang G, Kennedy SP, Fasiludeen S, Rensing C, DasSarma S. 2004. Arsenic resistance in Halobacterium sp. strain NRC-1 examined by using an improved gene knockout system. J Bacteriol 186:3187–3194.

55. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539.

56. Zhang Y. 2008. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9:40.

57. HeeShin W. 2014. Prediction of protein structure and interaction by GALAXY protein modeling programs. Biodesign 2:1–11.

58. Macindoe G, Mavridis L, Venkatraman V, Devignes M-D, Ritchie DW. 2010. HexServer: an FFT-based protein docking server powered by graphics processors. Nucleic Acids Res 38:W445–W449.

59. Stein SE. 1994. Estimating probabilities of correct identification from results of mass spectral library searches. J Am Soc Mass Spectrom 5:316–323.

60. Fiorentino G, Ronca R, Cannio R, Rossi M, Bartolucci S. 2007. MarR-

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

like transcriptional regulator involved in detoxification of aromatic compounds in Sulfolobus solfataricus. J Bacteriol 189:7351–7360.

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

62. Claassens NJ, Siliakus MF, Spaans SK, Creutzburg SCA, Nijsse B, Schaap PJ, Quax TEF, Van Der Oost J. 2017. Improving heterologous membrane protein production in Escherichia coli by combining transcriptional tuning and codon usage algorithms. PLoS One 12:e0184355.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure captions:

Fig. 1 A) Phylogenetic tree of archaeal and bacterial arsenite methyltransferases, SAM-dependent methyltransferases and methyltransferase domain containing proteins. The aminoacidic sequences used for the construction of the phylogenetic tree are: *Tt*ArsM from *T. thermophilus* HB27; SAM-dependent methyltransferase from five members of the *Thermus* genus (*islandicus, caldilimi, antranikianii, oshimai* and *brockianus*), *Anaerolineae bacterium* and *Mesorhizobium amorphae*; arsenite methyltransferase from *Clostridium* sp BMX (53), *Methanosarcina acetivorans, Rhodopseudomonas palustris* (13), *Pseudomonas alcaligenes* (36) *Halobacterium salinarum* (54) *Cyanidioschyzon merolae* (45) and a methyltransferase domain-containing protein from *Sanguinobacter* sp.

B) Multiple sequence alignment of hypothetical and functionally characterized arsenite methyltransferases (ArsM) with *T. thermophilus* HB27 *Tt*ArsM. The partial alignment includes 5 members of the *Thermus* genus (*islandicus, caldilimi, antranikianii, oshimai* and *brockianus*) (98% identity to *Tt*ArsM, sequence aligned from amino acid 40 to 89), *Clostridium* sp BMX (53) (28% identity to *Tt*ArsM, from 22 to 193), *R. palustris* (13) (32% identity to *Tt*ArsM, from 23 to 168), *M. acetivorans* (37) (29% identity to *Tt*ArsM, from 19 to 190), *P. alcaligenes* (36) (31% identity to *Tt*ArsM, from 149 to 320), *Cyanidioschyzon merolae* (45) (27.7% identity to *Tt*ArsM, from 29 to 214) and

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

H. salinarum (54) (25% identity to *Tt*ArsM, from 58 to 294). Red arrow indicates the catalytic cysteine, conserved in characterized ArsM and *Tt*ArsM; green arrows indicate two catalytic cysteines conserved in characterized ArsM, but not in *Tt*ArsM; blue arrow indicates the conserved aspartic acid; the SAM binding domain, which is part of the typical Rossman fold, is underlined in blue. The two histidines of *Tt*ArsM predicted to interact with arsenite are indicated by black arrows.

Fig. 2 *In vitro* assessment of methyltransferase activity. A) Arsenite methylation assays. The enzyme coupled colorimetric assay was carried out in continuous in the presence of 3.1 µM of recombinant *Tt*ArsM (black curve) or *Tt*ArsM C77S (red curve), 800 µM SAM and 200 µM As(III) at 50°C. The absorbance of the reaction mixture was recorded every minute for a total of 1 hour. The graph represents the average of three independent experiments, each performed in triplicate. **B) Products of As(III) methylation by purified *Tt*ArsM.** Arsenic species were analyzed by GC-MS. Each assay contained 10 µM *Tt*ArsM, 250 µM As(III), 6 mM GSH and 1 mM SAM, incubated at 65°C for 24 h. In **B)** and **C)**, the chromatograms recorded between 2 and 6 min for negative control and the enzyme mixture. **D** and **E** report the fragmentation spectra for MMAs and DMAs.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fig. 3 *Tt*SmtB:*Tt*ArsM interaction in the presence of heavy metal ions. CoIp of *Tt*SmtB:*Tt*ArsM complex with increasing concentrations of **(A)** arsenate, **(B)** arsenite, **(C)** cadmium and **(D)** antimony. Complexes were immunoprecipitated with anti-*Tt*SmtB antibodies and revealed through Western-blot using anti-His antibodies against the His-tag of *Tt*ArsM. Below: densitometric analysis of blots of *Tt*SmtB:*Tt*ArsM complex. The intensity of the unchallenged complex was used as a reference. Average values from three biological replicates are shown, with error bars representing standard deviations. Statistical analysis was performed using one-way ANOVA; significant differences are indicated as: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Fig. 4 *Tt*SmtB:p_{arsM} interaction. (A) Interaction of p_{arsM} in the presence of increasing concentration of *Tt*SmtB. Lane 1: negative control; Lane 2: 1 µM *Tt*SmtB; Lane 3: 2 µM *Tt*SmtB; Lane 4: 3 µM *Tt*SmtB; Lane 5: 5 µM *Tt*SmtB; Lane 6: 10 µM *Tt*SmtB. **(B)** EMSA with p_{arsM}, in presence of 3 µM *Tt*SmtB and 3 µM *Tt*ArsM.

Fig. 5 Growth of *E. coli* BL21 strains expressing *Tt*ArsM and its mutants, in the presence of different arsenite concentrations, measured 24 hours post-inoculation. The strains are: *E. coli* BL21/pET30b (black), *E. coli* BL21/pET30/*Ttars*M (red), *E. coli* BL21/pET30/ *Ttars*M H40A (green), *E. coli* BL21/pET30/*Ttars*M H179A (blue) and *E. coli* BL21/pET30/*TtarsM* C77S

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(orange). Average values from three biological replicates are shown, with error bars representing standard deviations.

Fig. 6 ThermoCas9-based genome engineering in *T. thermophilus* HB27. **A)** pMK-ThermoCas9-sp1/2/NT targeting vectors. **B)** Graphical representation of the ThermoCas9 targeting assay results (CFUs), for assessing the ThermoCas9 toxicity and targeting efficiency in *T. thermophilus* HB27. Average values from three biological replicates are shown, with error bars representing standard deviations. **C)** pMK-ThermoCas9-HR-sp1/2 editing vectors, employed for the genomic deletion of the *TtarsM* gene. **D)** Agarose gel electrophoresis showing the resulting products from genome-specific colony PCRs on *T. thermophilus* colonies formed from the ThermoCas9-based *TtarsM* deletion process. A wild type colony was subjected to the same PCR and the related product is shown here as negative control for *TtarsM* deletion. The expected sizes of the PCR amplification products that correspond to the wild type and Δ*TtarsM* genotypes are indicated with black arrows.

Fig. 7 Growth of *T. thermophilus* HB27 (black), *T. thermophilus* HB27 Δ*TtarsM* (red) and *T. thermophilus* HB27 Δ*TtarsM*-Δ*TtarsX* (*syfp*) (blue) in TM medium in the presence of different concentrations of **A)** arsenite and B**)** arsenate measured 24 hours after inoculation. Average values from three biological replicates are shown, with error bars representing standard deviations.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fig. 8 Bioreporter activity. *T. thermophilus* HB27 Δ*TtarsM*-Δ*TtarsX* (*syfp*) bioreporter strain challenged with increasing concentrations of **A)** arsenite and **B)** arsenate. Average values from three biological replicates are shown, with error bars representing standard deviations. Statistical analysis was performed using one-way ANOVA; significant differences are indicated as: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplemental Tables:

TABLE S1. List of *Tt*SmtB cytosolic interactors.

TABLE S2. Results from genome-specific colony PCRs on *T. thermophilus* HB27 colonies formed from the ThermoCas9-based *TtarsM* deletion experiments.

TABLE S3. List of the primers used in this work.

TABLE S4. List of the PCR products used for the HiFi assembly reactions to construct the ThermoCas9-based targeting and editing plasmids. The primers and templates used for the PCR reactions are also included in this list.

TABLE S5. The sequence of the synthesized *thermoCas9* gene, which is codon-harmonized according to the codon usage of *T. thermophilus* HB27.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplemental Figure captions:

Fig. S1 TTC0109 3D model (wild type and mutants) and docking with arsenite and SAM (the arsenic atom is the purple sphere, the oxygen atoms are the red spheres, SAM is orange). The H40 and H179 residues coordinating arsenite are coloured blue. The C77 residue is coloured black**; B)** TTC0109 C77S; **C)** TTC0109 H40A; **D)** TTC0109 H179A, blue the mutated position.

Fig. S2 Purification of recombinant *Tt*ArsM. A) SDS-PAGE analysis Lane M, protein marker. Lane NI, protein extract from non-induced cells. Lane I, protein extract from induced cells. Lane P, pure protein after His-trap chromatography. **B)** Size exclusion chromatogram of *Tt*ArsM; in the box the calibration curve. The chromatogram shows a peak corresponding to the *Tt*ArsM dimeric form. **C)** Histograms represented the average peak areas for MMAs and DMAs in *Tt*ArsM enzymatic reaction sample. These peak area values correspond to a concentration lower than the LOQ assessed to 1 mg/L by using standard molecules. The CV% values obtained were lower than 15 %.

Fig. S3 Purification of recombinant *Tt*ArsM mutants. SDS-analysis. Lane M, protein marker. Lane NI, protein extract from non-induced cells. Lane I, protein extract from induced cells. Lane W, unbound proteins after His-trap chromatography. Lane P, pure protein after His-trap chromatography. **A)** Recombinant *Tt*ArsM C77S. **B)** Recombinant *Tt*ArsM H40A. **C)** Recombinant *Tt*ArsM H179A.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fig. S4 Agarose gel electrophoresis from genome-specific colony PCRs on *T. thermophilus* HB27 colonies formed from the ThermoCas9-based *TtarsM* deletion process upon transformation with **A)** the pMK-ThermoCas9-HR-NT control-editing vector, **B)** the pMK-ThermoCas9-HR-sp1 editing vector, and **C)** the pMK-ThermoCas9-HR-sp2 editing vector. Wild type colonies were subjected to the same PCR and the related products are shown as negative controls for *TtarsM* deletion. The expected sizes of the PCR amplification products that correspond to the wild type and Δ*TtarsM* genotypes are indicated with arrows. **D)** Sequencing chromatogram of the PCR amplified *TtarsM* genomic region from a randomly selected colony, previously PCR screened as Δ*TtarsM*. **E)** Agarose gel electrophoresis after the curing process of the editing vectors shows the absence of products using pMK18-specific primers for colony PCRs on *T. thermophilus* HB27 Δ*TtarsM* colonies. A wild type colony containing the pMK18 vector was subjected to the same PCR and the related product is shown here as a negative control of the curing process.

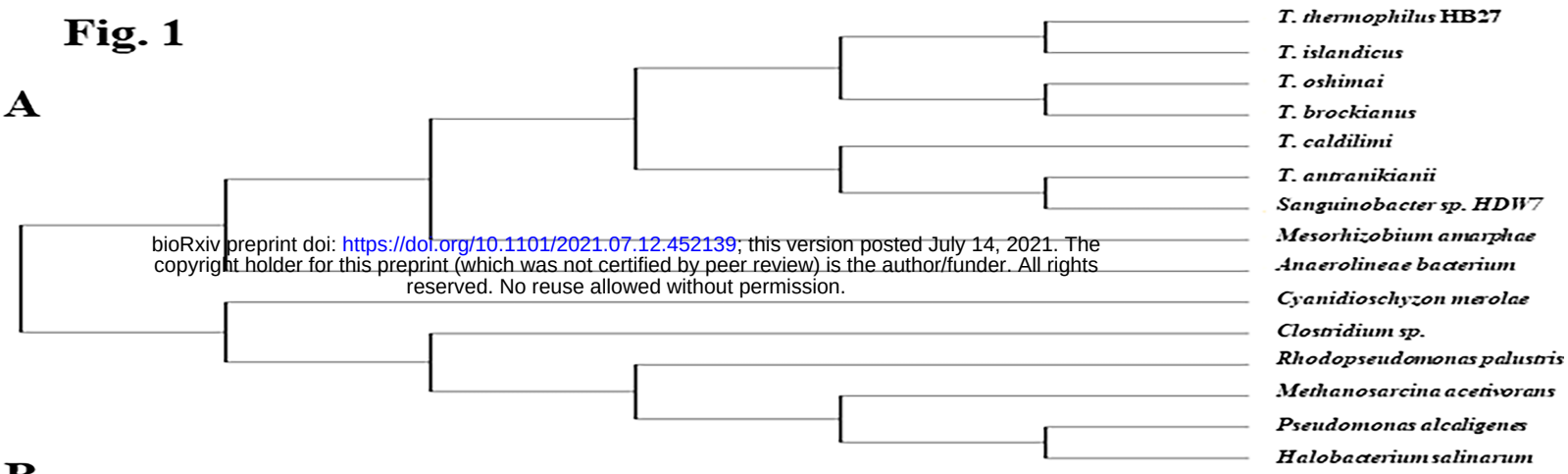
Fig. S5 Agarose gel electrophoresis showing the resulting products from genome-specific colony PCRs on *T. thermophilus* HB27 Δ*TtarsM*-Δ*TtarsX (syfp)* colonies formed from the ThermoCas9-based substitution process of the *TtarsX* gene by the *syfp* gene. A wild type colony was subjected to the same PCR and the related product is shown as a negative control for *TtarsX* substitution. The expected sizes of the PCR amplification products that correspond to the wild

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

type and Δ*TtarsX* (*syfp*) genotypes are indicated with black arrows. Eight out of the ten screened clones were *T. thermophilus* Δ*TtarsM*-Δ*TtarsX* (*syfp*) knock-in mutants.

Fig. 1

A



bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

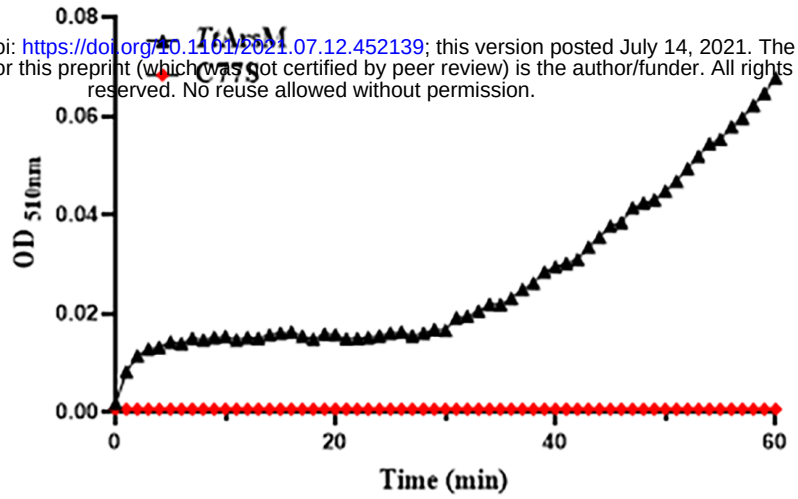
B

```
T. thermophilus HB27      HSEVSGK---EEEVFLTLQRAFQKAAEEGALVMWALLTNACEAGDPFRKEERL  89
T. oshimai                HTELWGE---AEAVFRALRAAYEEAAGEGALVLWALLTNACEAQDPFRRPDRL  88
T. islandicus             HSELWGG---EEEVFRALKEAFARGAEEGALVMWALLTNACEAGDPFRKEDRL  89
T. caldilimi              HSEVSGE---EGAVFQALQAAYLAAAEEGATVMWALFTNACEARDPFRSPERI  89
T. brockianus             HTELSGE---EAVVFGALREAFAKAAEEGALVMWALLTNACEAQDPFRLPERL  89
T. antranikianii          HTELSGE---EEAVFRALEAAFKAAAMEGATVMWALFTNACEAKDPFRRPERL  89
Rhodopseudomonas palustris ----------GASPITSNL-YDAAQEQGLPAEAMLASLGCG------NPTAL  72
Pseudomonas alcaligenes   DGGCCSD---ETEASGSERLGYDADDVASVADG-ADLGLGCG------NPKAF  198
Methanosarcina acetivorans GGGCCGD---LSAADLSRSLGYSEADVQAVPD--ANLGLGCG------NPTAF  68
Halobacterium salinarum   DGGCCSD---ETEASGSERLGYDADDVASVADG-ADLGLGCG------NPKAF  198
Cyanidioschyzon merolae   ----------KLAAAVPESH---RKILADIADEVLEKFYGCG------STLPA  78
Clostridium sp.           SPGCCSDGLSDAADPITGNL-YDESDLQGLDPELIANSFGCG------NPTAL  71

T. thermophilus HB27      RRFPPGEIARKALEGLKAKSVLDIGTGTG----VFAEAFAALGLFVVGLDPRA  138
T. oshimai                RRFNPLGVARKALEGLEAKSVLDIGTGTG----VFAEAFQELGLFTVGLDPRA  137
T. islandicus             RRFPPLEIARKALEGLKAKSALDIGTGTG----VFAEAFAQMGLFTVGLDPRA  138
T. caldilimi              KRFPPSEITRKALEGLGARSALDIGTGTG----VFAEAFARLGLFTVGVDPRS  138
T. brockianus             ERFAPLEVARRALEGLKAKSALDIGTGTG----VFAEAFHGLGLFTVGLDPRA  138
T. antranikianii          KRFPPLAIARKALEGLKAQSVLDIGTGTG----VFAEAFASLGLFTVGVDPRS  138
Rhodopseudomonas palustris AQLSPGET------------VLDLGSGGGIDVLLSARRVGPTGKAY-GLDMTD  112
Pseudomonas alcaligenes   AAMAPGET------------VLDLGSGAGFDCFLAAQEVGPDGHVI-GVDMTP  238
Methanosarcina acetivorans AELKPGDI------------VLDLGSGAGFDSFLAAQRVGSLGKVI-GVDMTQ  108
Halobacterium salinarum   AAMAPGET------------VLDLGSGAGFDCFLAAQEVGPDGHVI-GVDMTP  238
Cyanidioschyzon merolae   DGSLEGAT------------VLDLGCGTGRDVYLASKLVGEHGKVI-GVDMLD  118
Clostridium sp.           MNLNLGEV------------VLDLGSGSGLDVLLSAKRVGPTGKAY-GLDMTD  111

T. thermophilus HB27      DRLEVARAKVK-------------GARFVEGRAEALP------FPDGSFDLAF  172
T. oshimai                DRLEVARAKVP-------------GARFVEGRAEALP------FPDGSFDLAF  171
T. islandicus             DRLEVARSRVQ-------------GARFVEGRAEALP------FPDGSFDLAF  172
T. caldilimi              DRLEVARAKVK-------------KAHFLEAHGENLP------FPAQSFDLAF  172
T. brockianus             DRLEMARAKVK-------------GARFVEGRAESLP------FPDGSFDLAF  172
T. antranikianii          DRLEVARAKVK-------------KARFLEARGESLP------FPDQSFDLAF  172
Rhodopseudomonas palustris EMLALARDNQR--------KAGLDNVEFLKGEIEAIP------LPDHSVDVII  151
Pseudomonas alcaligenes   EMISKARENVA--------KNDAENVEFRLGEIGHLP------VADESVNVVI  277
Methanosarcina acetivorans EMVKKAQDNAR--------KYGYSNVEFRQGDIEALP------LDDRSVDVII  147
Halobacterium salinarum   EMISKARENVA--------KNDAENVEFRLGEIGHLP------VADESVNVVI  277
Cyanidioschyzon merolae   NQLEVARKYVEYHAEKFFGSPSRSNVRFLKGFIENLATAEPEGVPDSSVDIVI  171
Clostridium sp.           EMLAVAKENQR-------KSGIENAEFLKGHIEEIP------LAAKSIDVII  150

T. thermophilus HB27      FGLALH--------HLDPVPALREASRVARRVAVLEWPYREE----------  206
T. oshimai                FGLALH--------HLDPIPALREAARVARRVAVLEWPYREE----------  205
T. islandicus             FGLSLH--------HLDAEKALKEAARVARRVAVLEWPYREE----------  206
T. caldilimi              FGLSLH--------HLDPIPALREAARVARRVVVLEWPFRQE----------  206
T. brockianus             FGLSLH--------HLDPIPALKEAARVARRVAVLEWPYREE----------  206
T. antranikianii          FGLSLH--------HLDPIPALREAARVARRVAVLEWPFRQEVALPEGQTDLR  217
Rhodopseudomonas palustris SNCVINLSGDKDRVLREAFRVLKPGGRFAVSDVVTRGEIPEAL----------  194
Pseudomonas alcaligenes   SNCVVNLAPEKQRVFDDTYRVLRPGGRVAISDVVQTAPFPDDV----------  320
Methanosarcina acetivorans SNCVINLAPDKEKVFREAFRVLKPGGRMYVSDMVLLEDLPEDL----------  190
Halobacterium salinarum   SNCVVNLAPEKQRVFDDTYRVLRPGGRVAISDVVQTAPFPDDV----------  320
Cyanidioschyzon merolae   SNCVCNLSTNKLALFKEIHRVLRDGGELYFSDVYADRRLSEAA----------  214
Clostridium sp.           SNCVINLSGDKDKVLKEAYRVLKPQGRFAVSDIVIKRPLPEKI----------  193
```

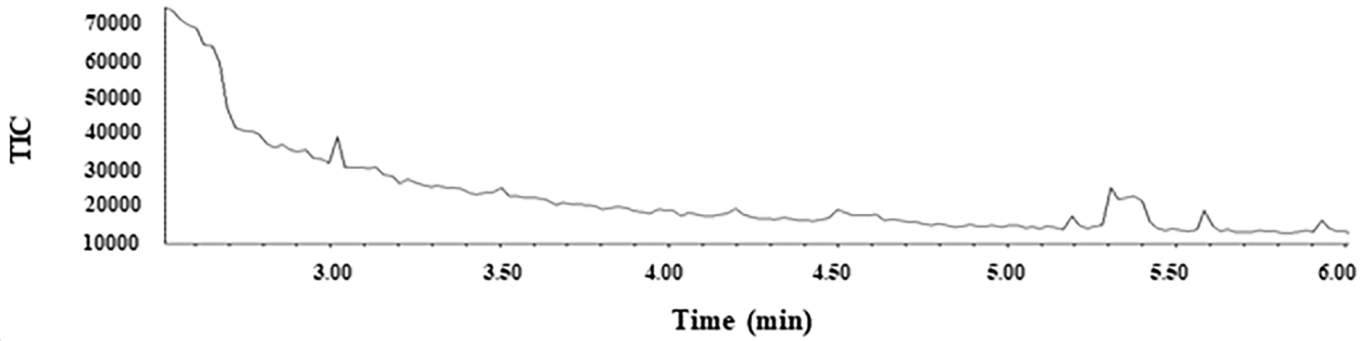
Fig. 2

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

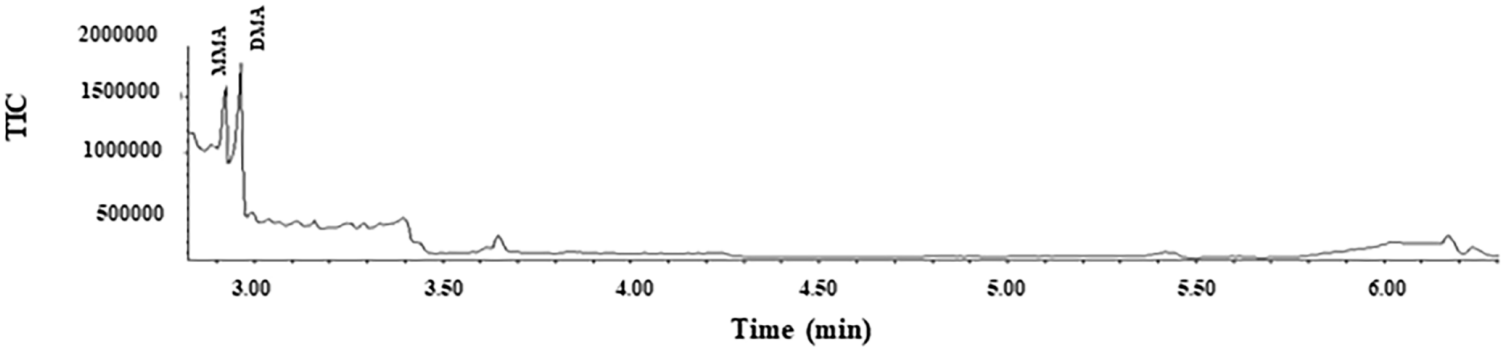
A



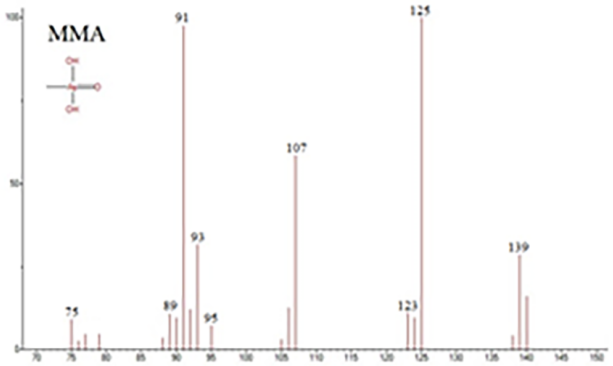
B



C



D



E

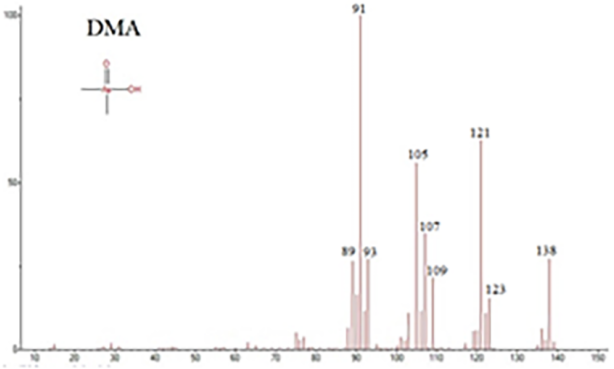
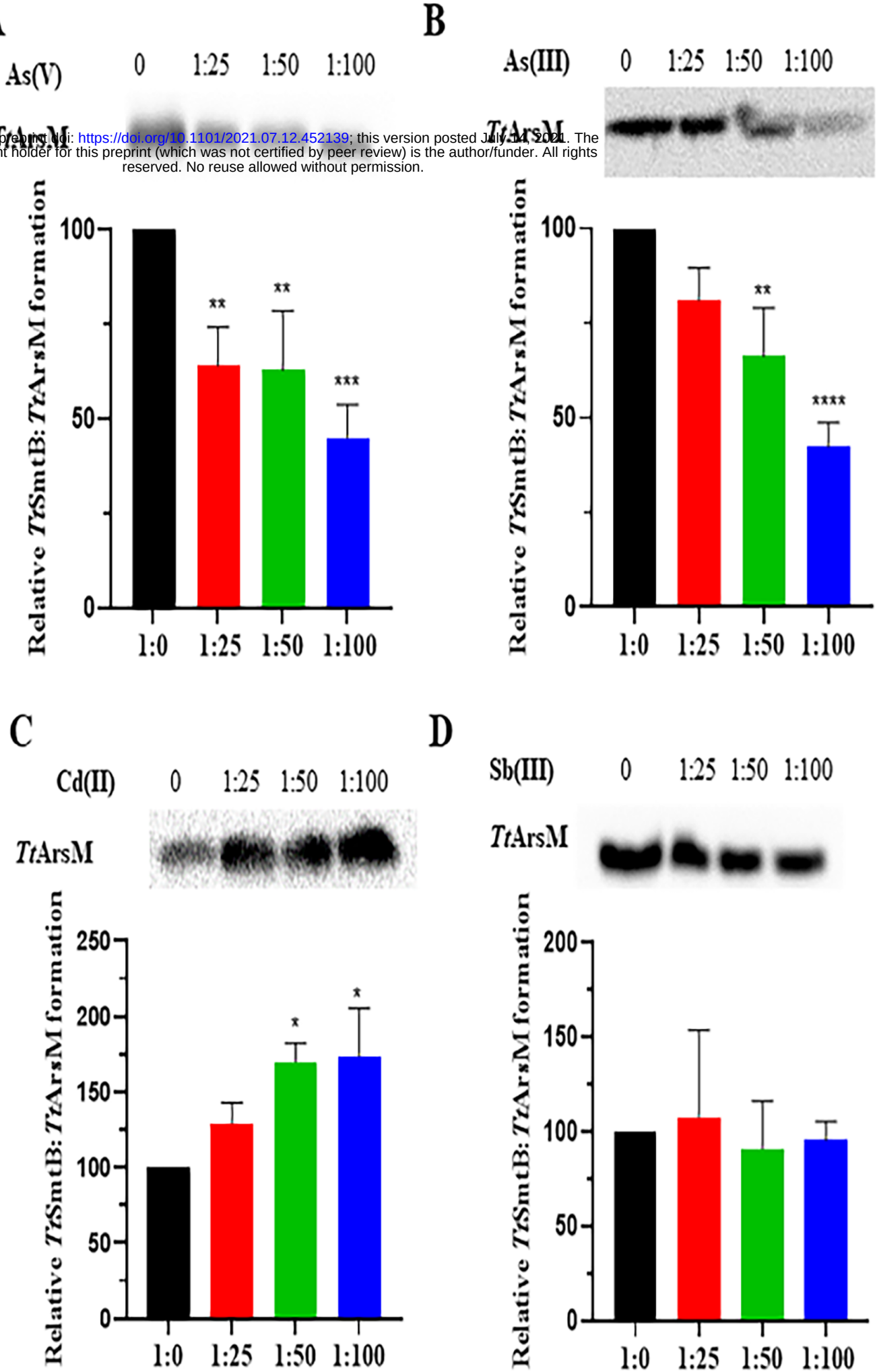


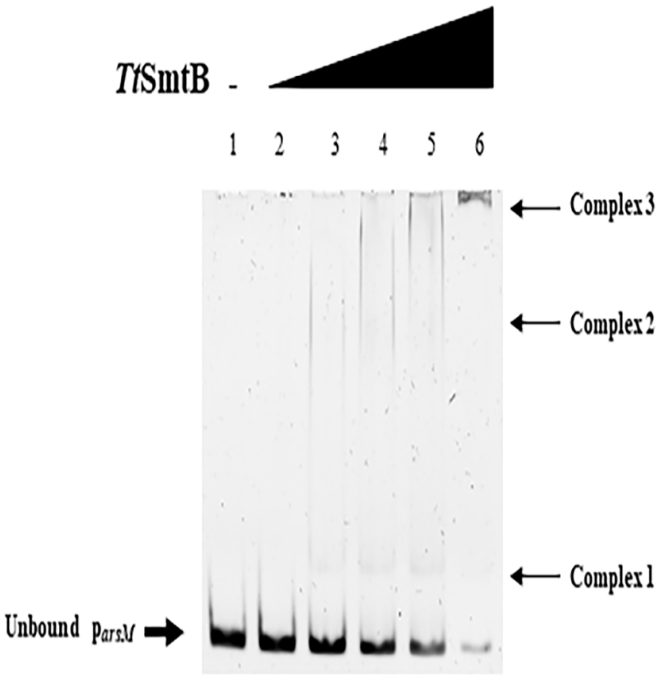
Fig. 3



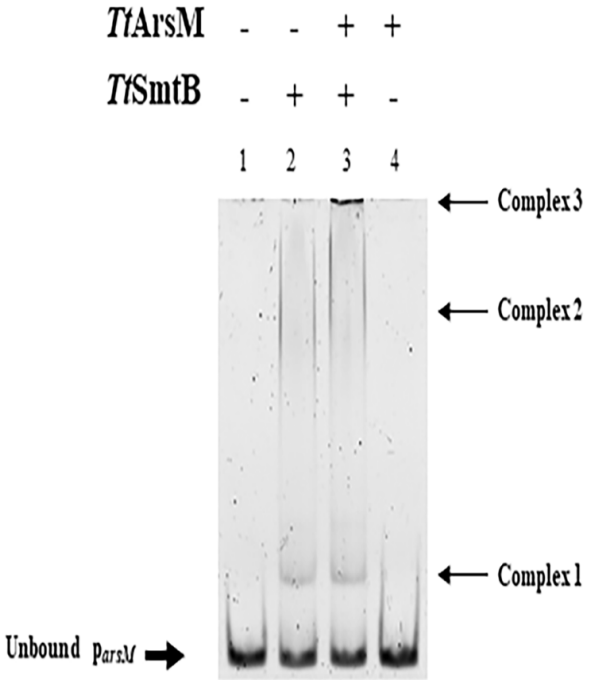
bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fig. 4 A

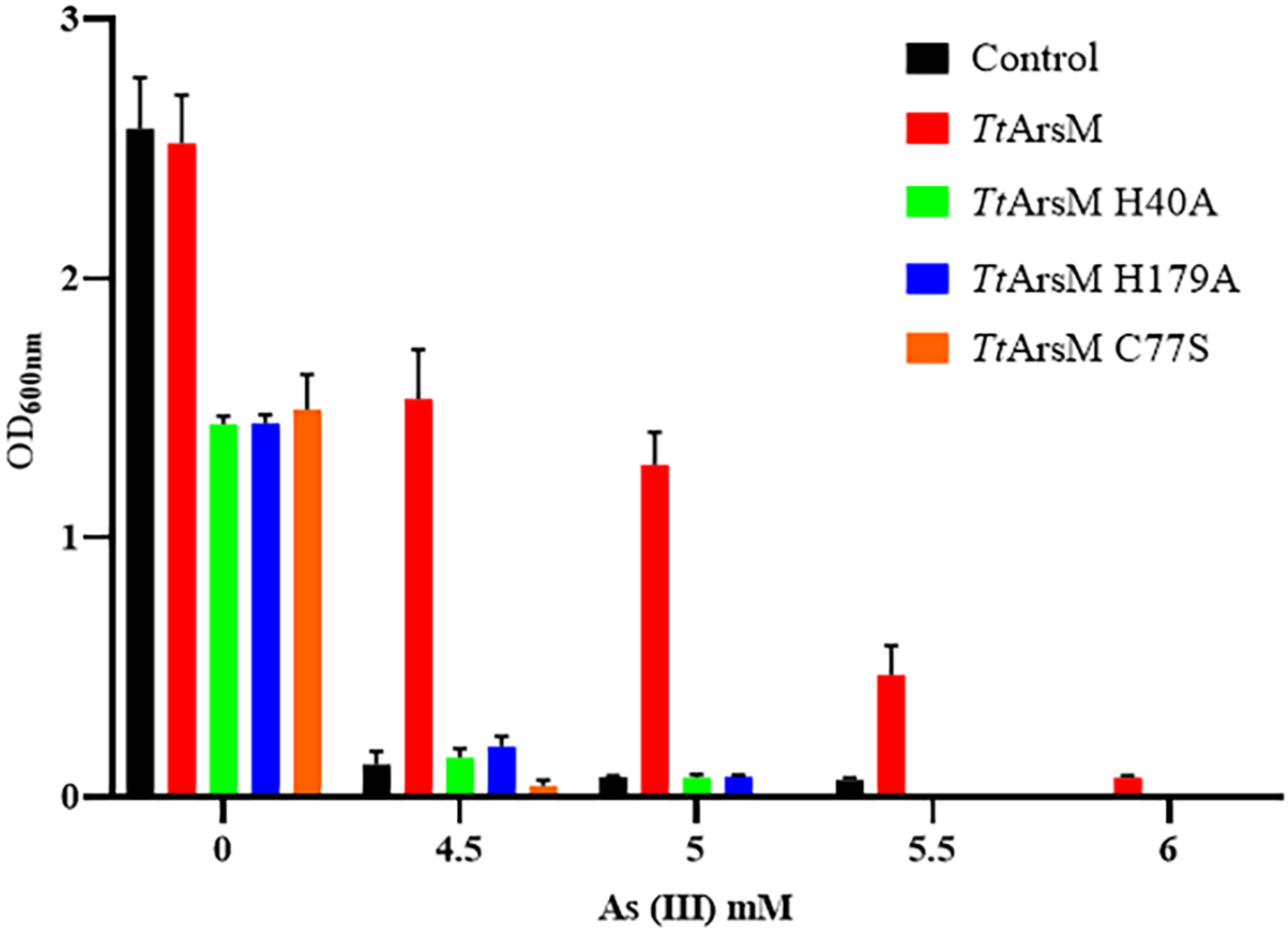


B



bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

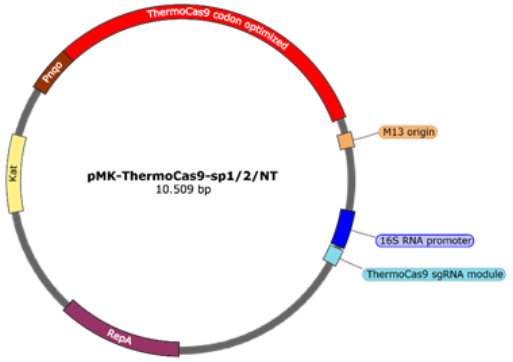
Fig. 5



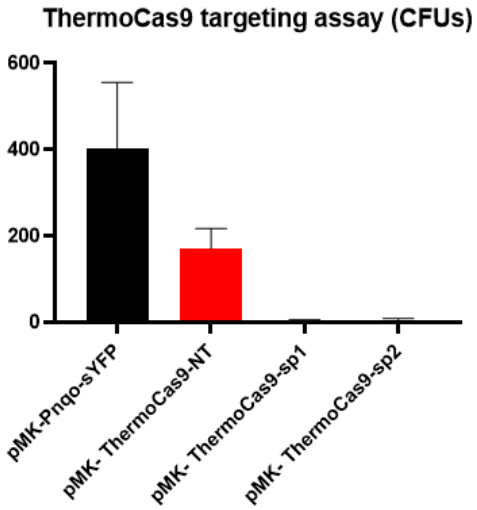
bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fig. 6

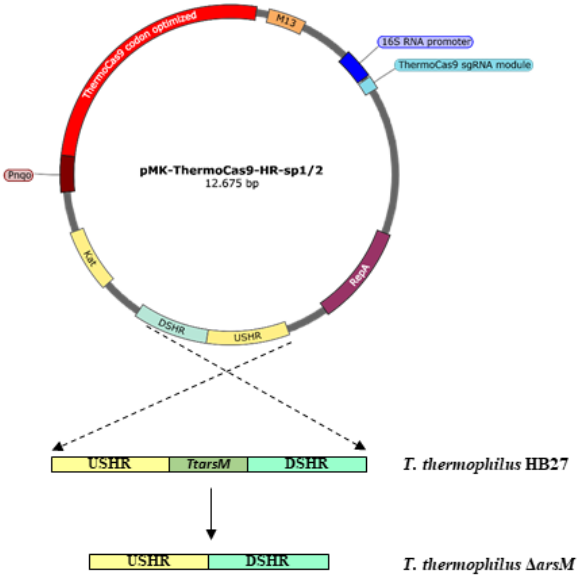
A



B



C



D

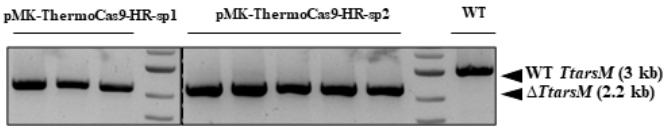
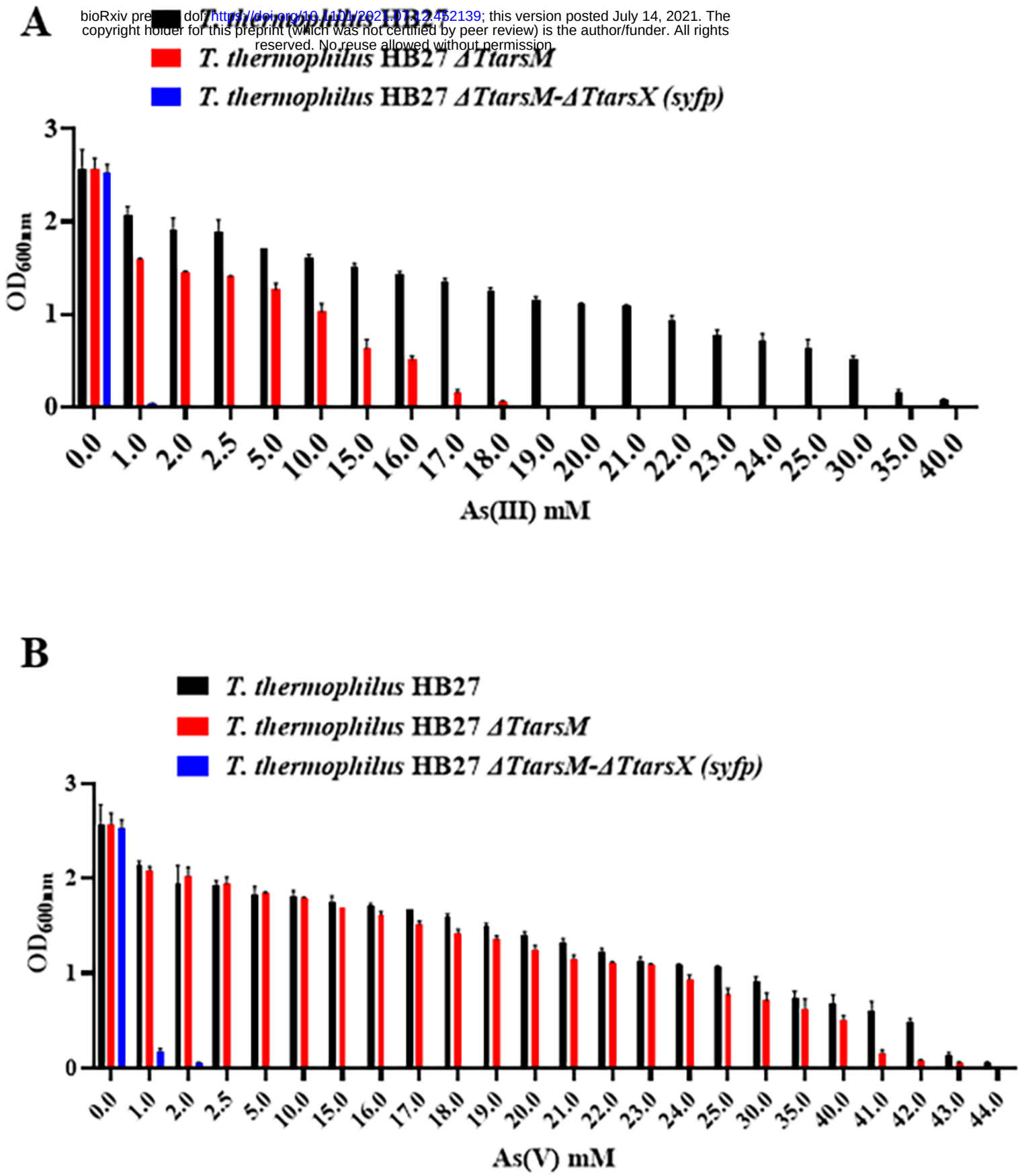


Fig. 7

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: https://doi.org/10.1101/2021.07.12.452139; this version posted July 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fig. 8