Methylation of two-component response regulator MtrA in mycobacteria negatively modulates 1 its DNA binding and transcriptional activation. 2

Anshika Singhal, Richa Virmani, Saba Naz, Gunjan Arora, Mohita Gaur, Parijat Kundu, Andaleeb Sajid, Richa Misra, Ankita Dabla, Suresh Kumar, et al.

Methylation of two-component response regulator MtrA in mycobacteria negatively modulates its DNA binding and transcriptional activation.

Anshika Singhal[2,#], Richa Virmani[1,2,#], Saba Naz[1,3], Gunjan Arora[2], Mohita Gaur[1,2], Parijat Kundu[2], Andaleeb Sajid[2], Richa Misra[2], Ankita Dabla[3], Suresh Kumar[3], Jacob Nellissery[4], Virginie Molle[5], Ulf Gerth[6], Anand Swaroop[4], Kirti Sharma[7], Vinay K Nandicoori[3*], and Yogendra Singh[1*]

[1]Department of Zoology, University of Delhi, Delhi 110007, India.

[2]CSIR- Institute of Genomics and Integrative Biology, Delhi 110007, India.

[3]National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110067, India.

[4]Neurobiology-Neurodegeneration and Repair Laboratory, National Eye Institute, National Institutes of Health, Bethesda, MD 20892, USA.

[5]DIMNP, CNRS, University of Montpellier, Montpellier, France.

[6]Institute of Microbiology, Ernst-Moritz-Arndt-University Greifswald, D-17487 Greifswald, Germany.

[7]Department of Proteomics and Signal Transduction, Max-Planck-Institute of Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany.

[#]Contributed equally to this work.

[*]Corresponding author. E-mail ysinghdu@gmail.com & vinaykn@nii.ac.in

Abstract

Posttranslational modifications such as phosphorylation, nitrosylation, and pupylation modulate multiple cellular processes in *Mycobacterium tuberculosis*. While protein methylation at lysine and arginine residues is widespread in eukaryotes, to date only two methylated proteins in *Mtb* have been identified. Here we report the identification of methylation at lysine and/or arginine residues in nine mycobacterial proteins. Among the proteins identified, we chose MtrA, an essential response regulator of a two-component signaling system, which gets methylated on multiple lysine and arginine residues to examine the functional consequences of methylation. While methylation of K207 confers a marginal decrease in the DNA binding ability of MtrA, methylation of R122 or K204 significantly reduces the interaction with the DNA. Overexpression of S-adenosyl homocysteine hydrolase (SahH), an enzyme that modulates the levels of S-adenosyl methionine in mycobacteria decreases the extent of MtrA methylation. Most importantly, we show that decreased MtrA methylation results in transcriptional activation of *mtrA* and *sahH* promoters. Collectively, we identify novel methylated proteins, expand the list of modifications in mycobacteria by adding arginine methylation, and show that methylation regulates MtrA activity. We propose that protein methylation could be a more prevalent modification in mycobacterial proteins.

INTRODUCTION

Mycobacterium tuberculosis (Mtb), the causative agent of tuberculosis, is responsible for nearly one million deaths annually around the globe [1]. It resides dormant in the host for decades without detection and when the immune system wanes, it proliferates and causes active disease. The adeptness of mycobacteria to hijack the host cell can be attributed to the fine-tuning of signaling pathways. Post-translational modifications (PTMs) including serine/threonine phosphorylation, nitrosylation, and pupylation (addition of prokaryotic ubiquitin-like protein) play an important role in regulating mycobacterial physiology and virulence [2-6]. While there are few specific examples of how these modifications affect the function of a protein, more mechanistic insight is required to delineate their regulatory roles. In addition to these modifications, proteins can be post-translationally modified by the addition of methyl groups, catalyzed by S-adenosyl methionine (SAM) dependent methyltransferases [7], at the ε-amino group of lysine, guanidino group of arginine, or oxygen in the carboxylate side chain of glutamate [8-10]. Glutamate methylation of methyl-accepting chemotaxis proteins play a biologically conserved role in chemotaxis and provide rotational directionality to bacteria [11].

In eukaryotes, methylation of histone proteins at specific lysine residues regulates chromatin architecture and transcription, and aberrant methylation is associated with aging and cancer [12]. Arginine methylation is the most extensively studied protein modification in eukaryotes and its role in DNA repair, RNA metabolism, and transcriptional repair is well established [13]. Guanidino group of arginine is involved in the interaction with DNA; the addition of methyl group directly affects the activity of proteins. Methylation of Sam68 (an adapter protein for Src kinases during mitosis) at arginine residue restrains it's binding to Src homology 3 (SH3) domain of phospholipase Cγ-1 and methylation at arginine and lysine residues of CHD1 (chromo-helicase/ATPase DNA-binding protein 1) results in a significant decrease in its binding affinity to DNA [8]. Several non-histone proteins, mainly transcription factors and histone- or chromatin-associated proteins are also regulated by methylation [12, 14].

In bacteria, however, our understanding of the functional role of lysine or arginine methylation is limited [9]. Lysine methylation is associated with bacterial cell motility of *Synechocystis* sp. and with host colonization and disease initiation by *Pseudomonas aeruginosa* [15]. A recent proteomics study has identified abundant lysine and arginine methylation in *Escherichia coli* [16]. In *Mtb*, lysine residues of Heparin-Binding Hemagglutinin Adhesin (HBHA) and Histone-like Protein (HupB) have been shown to undergo methylation but there are no reports of arginine methylation. HBHA and HupB are both critical for infection by *Mtb* [17] and their methylation imparts protease resistance and thus increased stability, suggesting a role for methylation in disease pathogenesis [18]. Methylation reactions are catalyzed by S-adenosyl methionine (SAM)- dependent methyltransferases where S-adenosyl homocysteine (SAH) and consequently homocysteine (Hcy) are generated as by-products. Methyltransferase reactions are dependent on the presence of balanced amounts of SAM and SAH as they are prone to SAH-mediated inhibition. Under normal conditions, SAH levels are regulated using SahH-mediated reversible hydrolysis

of SAH to Hcy. We have previously shown that perturbation of levels of *Mtb* SahH impacts metabolic levels of Hcy and may affect SAH, a potent inhibitor of methyltransferases [19].

In this manuscript, we set out to determine the prevalence of methylation in *Mtb* proteins. Nine among the 72 proteins tested were found to be methylated either on lysine or arginine residues. To determine the functional consequences of methylation, we chose MtrA; an essential response regulator of the MtrB-MtrA two-component system (TCS) that regulates cell cycle progression. We show that methylation perturbs MtrA DNA binding activity leading to modulation of its own expression. We also reveal that SahH, an enzyme that is required for SAM synthesis, modulates MtrA methylation. Taken together, we propose that methylation of lysine and arginine residues is an important additional regulatory modification in *Mtb*.

Materials and Methods

Bacterial strains and growth conditions.

E. coli strains DH5α (Novagen) and BL21-DE3 (Stratagene) were used for cloning and expression of recombinant proteins, respectively. *M. smegmatis mc²155* (*Msm*) and *Mtb* H37Rv were maintained in Middlebrook 7H9 broth (Difco, BD) containing 10% ADC (Albumin/Dextrose/Catalase) and 0.05% Tween-80 (Merck, USA), supplemented with 25 μg/ml kanamycin or 50 μg/ml apramycin when required. For assessing the effect of homocysteine (Hcy) on bacterial growth, *Msm* cells were grown in Sauton's minimal medium supplemented with 0 - 0.8 mM DL-homocysteine (Sigma-Aldrich) at an initial A_{600} of 0.01. Absorbance was measured up to 36 h and colony-forming units (CFUs) were enumerated at 25 h. Reagents and chemicals were purchased from Sigma-Aldrich unless otherwise mentioned.

Generation of plasmid constructs.

We selected 180 protein-coding genes from *Mtb* genome representing a random set across various functional classes (Fig S1). Genes involved in regulation and information processing were over-represented in the list, while conserved hypotheticals and PE/PPE genes (encoding proteins containing Proline-Glutamate or Proline-Proline-Glutamate motifs) were under-represented. We did not select any gene from the categories stable RNAs, insertion sequences and phages, and those with unknown function. The generation of recombinant plasmids using the shuttle vector pVV16 was explained previously [20]. The recombinant clones (2 μg) were transformed individually in *Msm*. *Mtb* H37Rv genomic DNA was used to amplify *mtrA* (*rv3246c*; 687 bp) using forward and reverse primers containing NdeI and HindIII restriction sites. Digested PCR product was cloned into either pVV16 or the *E. coli* expression vector pET28a and recombinants were selected on kanamycin. *E. coli* K12 genomic DNA was used to amplify *envZ* (1353 bp) and cloned into pMAL-c2x at BamHI and HindIII restriction sites, recombinants were selected on ampicillin. Site-specific mutants of pVV16-*mtrA* and pET28a-*mtrA* were generated using QuikChange® XL Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's protocol. *Mtb sahH* (*rv3248c*; 1488 bp) was cloned in pVV16 vector at NdeI and HindIII restriction sites. All constructs were confirmed by restriction enzyme digestion and DNA sequencing (Invitrogen). Information about primers and plasmids used in this study is compiled in tables 1 and 2, respectively.

For analyzing the effect of SahH on MtrA methylation, the genes encoding these proteins were co-expressed in *Msm*. *Mtb sahH* was cloned in mycobacterial integrative vector pSET152 [21]. For this, pVV-*sahH* was digested with HindIII and the ends were made blunt. A second digestion with XbaI yielded 1.86 kb fragment containing *sahH* under a heat shock gene promoter (*hsp60*). This fragment was ligated to pSET152 pre-digested with XbaI and EcoRV. pSET-*sahH* and pSET152 (2 μg each) were electroporated in *Msm* competent cells and apramycin resistant transformants were selected. pVV16-*mtrA* plasmid (2 μg) was then electroporated in competent *Msm* cells harboring either pSET152 or pSET152-

sahH and the positive clones were selected on apramycin and kanamycin. His_6-MtrA was purified from cells containing both pSET-*sahH* and pVV16-*mtrA* and used for Western blotting.

Expression and purification of recombinant proteins.

For expression and purification of proteins from *Msm*, recombinant clones (2 μg) in pVV16 vector were electroporated and recombinants were selected on kanamycin. Expressed proteins were purified as described before [20]. Briefly, *Msm* cells expressing recombinant proteins were cultured individually in 200 ml of 7H9 medium and grown till mid-log phase (A_{600} ~ 0.8). The cells were harvested and lysed by sonication in lysis buffer (1X Phosphate-buffered saline (PBS), 1 mM phenylmethylsulfonyl fluoride (PMSF, Sigma-Aldrich), and protease inhibitor cocktail [Roche]). The lysates were centrifuged at 14,000 rpm at 4°C for 30 min, and the resulting supernatants containing His_6-tagged proteins were incubated with Ni^{2+}-NTA resin (Qiagen). The resin was washed thoroughly with a wash buffer (1X PBS, 1 mM PMSF, 20 mM imidazole, and 10% glycerol) and proteins were eluted in the elution buffer (1X PBS, 1 mM PMSF, 200 mM imidazole, and 10% glycerol).

For purifying His_6-MtrA from *Mtb*, pVV16-*mtrA* construct (2 μg) was electroporated and recombinants were selected on kanamycin. Recombinant cells were cultured in 200 ml of 7H9 media and grown till mid-log phase (A_{600} ~ 0.8). The cells were harvested and lysed by bead beating in lysis buffer (1X Tris-buffered saline (TBS), 1 mM PMSF, 100 μg/ml lysozyme, and protease inhibitor cocktail) using 0.1 mm zirconium beads (Biospec). The lysate was centrifuged and the resulting supernatant containing His_6-MtrA was incubated with Co^{2+} superflow resin (Thermo Scientific). The resin was washed thoroughly with wash buffer (1X TBS, 1 mM PMSF, and 10 mM imidazole) and the protein was eluted in the elution buffer (1X PBS, 1 mM PMSF, and 300 mM imidazole).

For protein expression in *E. coli*, pET28a- or pMAL-c2x-based plasmids (100 ng) were transformed, and proteins were overexpressed in *E. coli* BL21 (DE3). The recombinant His_6-tagged proteins were purified using Ni^{2+}-NTA affinity chromatography (Qiagen) and MBP (Maltose binding protein)-tagged EnvZ was purified using Amylose resin (NEB) as per the manufacturer's instructions. The concentration of purified proteins was estimated by Bradford assay (Bio-Rad). Purified proteins were resolved on SDS-PAGE and analyzed by staining with coomassie brilliant blue R250 (Sigma-Aldrich).

To analyze the effect of Hcy on MtrA methylation, *Msm* cells harboring pVV16-*mtrA* were grown in Sauton's medium containing 0.4 mM Hcy. Cells were grown up to A_{600} ~ 0.8, harvested, and subjected to Ni^{2+}-NTA chromatography for purification of His_6-MtrA, as described above. The proteins were later analyzed by Western blotting. For normalization, His_6-MtrA was purified from cells grown without the addition of Hcy.

Western blot analysis

Purified proteins were resolved by SDS-PAGE and transferred to nitrocellulose membrane (Millipore). Membranes were blocked with 3% bovine serum albumin (Sigma-Aldrich) in 1X PBS containing 0.1% Tween-20 (Sigma-Aldrich) (1X $PBST_{20}$) overnight at 4°C. After blocking, the blots were washed thrice with 1X $PBST_{20}$ followed by incubation with antibodies for 1 h at room temperature.

Methyllysine immunoblots were developed by two separate antibodies from different manufacturers and one representative blot is shown in figure 1a. The antibodies were: anti-methyllysine antibody from Abcam (ab23366; 1:10,000 dilution) and pan anti-mono, dimethyllysine antibody from PTM Biolabs (PTM-602; 1:2000 dilution). Other antibodies and their dilutions used were: anti-MtrA antibody (1:15,000 dilution; generated in the lab), anti-acetyllysine antibody (Cell Signaling; 1:3,000 dilution), pan anti-succinyllysine antibody (PTM Biolabs; 1:2,000 dilution), HRP-conjugated anti-His_6 antibody (Abcam; 1:20,000 dilution), and HRP-conjugated anti-mouse and anti-rabbit IgG antibodies (Bangalore Genei; 1:20,000 dilution). According to the manufacturer, the methyllysine antibodies used here can detect mono- or dimethyllysine with no cross-reactivity to acetyllysine. Antibodies to acetyllysine and succinyllysine have been successfully used in our previous study [20]. Antibodies against recombinant *Mtb* His_6-MtrA were raised in Rabbits with the help of Dr. A. K. Goel (DRDE, Gwalior, India). The specificity of anti-MtrA antibody was validated by using a preparation of purified MtrA protein that had been confirmed by mass spectrometry. Also, these antibodies identified a single protein band corresponding to MtrA when whole cell lysate preparations of *Mtb* were probed. Anti Ef-Tu antibodies were used as previously mentioned [22]. Immunoblots were developed using Immobilon™ western chemiluminescent HRP substrate kit (Millipore) according to the manufacturer's instructions. Quantification of immunoblots was performed using ImageJ software [23].

Mass spectrometry

Recombinant *Mtb* proteins purified from *Msm* were resolved on 12% SDS-PAGE and stained with coomassie brilliant blue R250. The stained bands were sliced from the polyacrylamide gel and subjected to in-gel reduction, carbamidomethylation, and an overnight tryptic digestion at 37°C. Alternatively, protein samples were subjected to chloroform-methanol precipitation and pooled before in-solution digestion and single-shot analysis. Mass spectrometry to identify protein methylation was essentially performed as described [24]. Peptides were separated on a 50 cm reversed-phase column (75 mm inner diameter, packed in-house with ReproSil-Pur C18-AQ 1.9 mm resin [Dr. Maisch GmbH]) over a 60- or 120-min gradient using the Proxeon Ultra EASY-nLC system. The LC system was directly coupled online with a Q Exactive HF instrument (Thermo Fisher Scientific) via a nano-electrospray source. Full scans were acquired in the Orbitrap mass analyzer with resolution 60,000 at 200 m/z. For the full scans, 3E6 ions were accumulated within a maximum injection time of 120 ms and detected in the Orbitrap analyzer. The ten most intense ions were sequentially isolated to a target value of 1e5 with a maximum injection time of 120 ms and fragmented by HCD in the collision cell (normalized collision energy of 25%) and detected in the Orbitrap analyzer at 30,000 resolution. Raw mass spectrometric data were analyzed in the MaxQuant environment v.1.5.3.31 and employed Andromeda for database search [25]. The MS/MS spectra were matched against the H37Rv proteome. Enzyme specificity was set to trypsin, and the search included cysteine carbamidomethylation as a fixed modification and methylation of lysine and arginine (+14.015650 Da) as variable modifications. Based on optimized parameters for PTM identification and localization [24], the search engine score was set to a minimum cutoff of 40 for

the identification of methylated peptides. Additional annotations on low and high scoring peptides were performed by the 'expert system' for computer-assisted annotation of MS/MS spectra. Up to two missed cleavages were allowed for protease digestion, and peptides had to be fully tryptic. Downstream bioinformatics analysis was done in the Perseus software environment, which is part of MaxQuant. For MtrA mutants, we employed the matching between runs algorithm [26, 27] in MaxQuant to alleviate the stochasticity of shotgun proteomics, which consists of transferring identifications of MS1 features between samples based on accurate mass and retention time values. Identification of lysine acetylation and succinylation was performed as described earlier [20].

In silico analysis

Gene names, protein names, protein subcellular localization, and molecular functions were extracted from Mycobrowser (https://mycobrowser.epfl.ch/) and UniProt databases (http://www.uniprot.org/). Protein functional categories were obtained as described earlier by Lew *et al.* [28]. Gene essentiality data was procured from Mycobrowser and from previous studies documenting gene essentiality during *in vitro* growth, infection, or growth on cholesterol-containing media [29-32]. MtrA crystal structure was obtained from Protein Data Bank (PDB ID: 2GWR) [33] and viewed using UCSF Chimera [34].

Putative promoter regions of *Mtb sahH*, or *Msm sahH* were predicted using BPROM (http://softberry.com). Probable MtrA-binding sites were searched at these promoter regions using the LASAGNA online tool (https://biogrid-lasagna.engr.uconn.edu) [35].

In vitro kinase assay

In vitro kinase assay was performed by a protocol described earlier [36]. Briefly, *E. coli* purified His_6-MtrA and mutants (5 µg each) were incubated with MBP-EnvZ kinase (2 µg) in the kinase buffer (50 mM Tris-Cl [pH 7.4], 50 mM KCl, 20 mM $MgCl_2$, 10 mM $CaCl_2$, and 1 mM DTT) and [γ-^{32}P]ATP (BRIT, Hyderabad, India) at 37°C for 30 min. The reaction was stopped by adding 2X Laemmli buffer and proteins were resolved on 12% SDS-PAGE followed by autoradiography using Personal Molecular Imager (PMI, Bio-Rad).

Electrophoretic Mobility Shift Assay (EMSA)

DNA region encompassing *Mtb oriC* (205 bp) [37], putative *Mtb sahH* promoter ($sahH_{Mt}$-Pr, 199 bp), or putative *Msm sahH* promoter ($sahH_{Ms}$-Pr, 201 bp) were PCR amplified, and purified products were end-labeled with [γ-^{32}P]ATP using T4 Polynucleotide Kinase (Roche) as per manufacturer's instructions. Varying amounts of His_6-MtrA and its site-specific mutants were phosphorylated using 2 µg MBP-tagged *E. coli* EnvZ in the kinase buffer and 1 mM ATP at 37°C for 30 min. Phosphorylated MtrA and MtrA mutants (10-100 µM) were incubated with the labeled DNA probes at 4°C for 30 min in a buffer containing 50 mM Tris-HCl (pH 8.0), 50 mM NaCl, 10 mM $MgCl_2$, 1 mM ATP, and 5% glycerol in a total volume of 20 µl. Reaction samples were resolved using 5% non-denaturing polyacrylamide gel in

0.5X Tris/Borate/EDTA buffer. Gels were dried and subjected to autoradiography in Personal Molecular Imager (BioRad).

RNA isolation and real-time PCR.

RNA isolation and quantitative real-time PCR (qRT-PCR) were performed using the protocols described previously [38] with few modifications. Briefly, log phase *Msm* cells were lysed in TRIzol® (Invitrogen) by bead beating using 0.1 mm zirconium beads. RNA was precipitated using isopropanol, washed with 70% ethanol, and dissolved in nuclease-free water. Before performing cDNA synthesis, RNA was treated with DNase (Ambion) according to the manufacturer's protocol to remove traces of genomic DNA. cDNA was synthesized from total RNA using random primers according to the protocol provided by the supplier (Thermo Scientific), and then used for measuring the expression of *mtrA* or *sahH* with gene-specific primers. qRT-PCR was performed using SYBR Green master mix (Roche) as per previously described protocols [39]. The data obtained were analyzed using the $\Delta\Delta C_T$ method and the relative fold change in expression was calculated. *Msm* housekeeping gene *sigA* (encoding Sigma factor A) or 16S rRNA, was used as a control. The primers were sequence-specific for each gene analyzed, with PCR products between 100 and 200 bp.

For studying the effect of Hcy on *mtrA* or *sahH* expression, *Msm* cells were grown in an increasing concentration of Hcy and gene expression was measured using qRT-PCR. For assessing the effect of SahH on *mtrA* expression, *Msm* cells containing pVV16 or pVV16-*sahH* were used.

RESULTS

Multiple Mtb proteins are methylated on lysine and arginine residues.

In *Mtb*, the proteins involved in metabolism, respiration, and cell wall-related processes form the majority of functional proteome compared with regulatory and signaling proteins (Fig S1a) [28]. Moreover, in *E. coli* the proteins involved in regulation and signaling represent a low copy number group as compared with the proteins involved in translation, protein folding, and other constitutive functions [40]. Therefore, for our study, we chose 180 candidate protein-coding genes belonging to different functional classes (Fig S1a & b and Table S1) with a focus on less prevalent regulatory and signaling proteins. We selected very few genes from "PE and PPE proteins" and "conserved hypotheticals" and did not select any genes from the categories "Stable RNAs", "Insertion sequences and phages", and "Unknown". These genes were cloned into mycobacterial expression vector pVV16 that contains a carboxy-terminal His_6-tag and the constructs were electroporated into *Msm*, a non-pathogenic model organism that provides appropriate cellular milieu close to *Mtb*. We successfully purified 72 out of the 180 proteins under non-denaturing conditions, while others failed to express or purify. Purified proteins were probed with anti-methyllysine to determine their methylation status where Histone protein served as a positive control. The identity of each purified protein was confirmed by re-probing the immunoblot membranes with an anti-His_6 antibody. Apart from identifying the purified target proteins, the anti-His_6 antibody also detected the presence of a consistent protein band corresponding to *Msm* chaperon protein GroEL, which is highly abundant and contains a C-terminal histidine-rich tail (MSMEG_1583) [41], and a few other contaminating proteins (Fig 1a). Among the 72 recombinant proteins, ten proteins were recognized by the anti-methyllysine antibody, suggesting the presence of methylation on lysine residues (Fig 1a). However, we could not detect a distinct methylated band for PykA. To validate the protein identity and the methylation of the ten western blot-positive candidate proteins, we performed high-resolution mass spectrometry. We were successful in detecting 20 methyllysine sites belonging to 7 candidate proteins except LldD1, LldD2, and Tpi (Fig 1b, Fig S2, Table S2). Interestingly, we also detected 18 methylarginine sites in 7 candidate proteins, which included LldD2 and Tpi (Fig 1b, Fig S2, Table S2). Together, we identified 20 methyllysine and 18 methylarginine sites in 9 out of 10 western blot-positive proteins.

MtrA is methylated in Mtb

Subsequently, we set out to investigate the biological significance of methylation on lysine and arginine residues. Towards this, we chose MtrA as the candidate protein, which was methylated on both lysine and arginine residues. MtrB-MtrA is one among the 11 TCS systems present in *Mtb* where MtrB is the sensor histidine kinase and MtrA is the cognate response regulator. High throughput transposon mutagenesis experiments suggested *mtrA* to be an essential gene for *in vitro* growth of *Mtb* [29]. To determine if MtrA is methylated in *Mtb*, pVV16-*mtrA* expression construct was expressed in *Mtb*, and purified His_6-tagged protein was probed with anti-MtrA and anti-methyllysine antibodies. Consistent with the results obtained in *Msm* (Fig 1a), MtrA was found to be methylated in *Mtb* (Fig 2b).

MtrA is a 228 amino acid (aa) long protein with a 102 aa long N-terminal response regulator domain and a 93 aa long C-terminal winged helix-turn-helix DNA binding domain homologous to *E. coli* OmpR (Fig 2a & Fig S3) [42, 43]. Mass spectrometry data showed that MtrA was methylated on six arginine residues and seven lysine residues (Fig 1b). Analysis of these 13 methylated residues showed that R122 is a conserved residue present in the linker region, and K204 and K207 are adjacent to the DNA recognition helix (Fig 2c & Fig S3). Therefore, we examined the roles of R122, K204, and K207 by mutating them individually to methionine residues, the closest structural mimic to dimethyllysine [44]. Wild type and mutant MtrA proteins were expressed in *Msm* and the purified proteins were probed with anti-methyllysine antibody to compare their relative methylation (Fig 3a). Densitometric analysis of blots suggested that mutating R122, K204, or K207 individually resulted in decreased overall methylation levels, albeit the extent of reduction varied. While $MtrA_{K204M}$ and $MtrA_{K207M}$ had considerable decrease in the extent of methylation (95% and 80%, respectively; p value < 0.0001), $MtrA_{R122M}$ mutant only lost marginal methylation (30%) (Fig 3b). The contribution of these three amino acid residues towards total MtrA methylation was further analyzed by a quantitative proteomics-based method to determine the extent of methylation at ten sites in MtrA. A heat-map was generated representing the methylation intensities of identified peptides in MtrA, $MtrA_{R122M}$, $MtrA_{K204M}$, and $MtrA_{K207M}$. The fold change of signal intensities at specific sites in MtrA mutants relative to that in MtrA are shown in Table S3. The comparison shows that mutating any of these residues negatively affects the methylation at other sites; K204M or K207M completely abolish the methylation at four other sites (Fig. 3c). We observed a background signal for amino acid position 207 in the $MtrA_{K207M}$ mutant, which was due to the 'match between run' event rather than a bonafede fide MS/MS signal. Moreover, the signal was only 0.4% compared with that in MtrA-WT signal, suggesting that it most likely represents the noise (Table S3). Similarly, background signal was also observed for amino acid position 122 in $MtrA_{R122M}$ mutant, which may be due to noise but could not be attributed to 'match between run' event. Collectively, the data indicate that multiple methylated residues of MtrA act co-operatively and K204 and K207 are crucial for MtrA methylation.

Methylation of MtrA is critical for DNA binding.

Binding of MtrA to DNA is contingent upon its phosphorylation on D56 residue by the sensor kinase MtrB [42]. Once phosphorylated, MtrA is known to regulate DNA replication by binding to the repeat nucleotide motifs at the origin of replication (*oriC*) [37]. To evaluate the role of methylation of MtrA on its DNA binding ability, we chose the 205 bp long region of *oriC* as a probe to perform EMSA with purified MtrA and its methylation site mutants $MtrA_{R122M}$, $MtrA_{K204M}$, and $MtrA_{K207M}$. To phosphorylate and activate MtrA, we utilized EnvZ, a homolog of MtrB in *E. coli* that has been used in several previous studies [36, 37, 45]. MtrA, $MtrA_{R122M}$, $MtrA_{K204M}$, and $MtrA_{K207M}$ proteins were incubated with EnvZ in the presence of [γ-^{32}P] ATP and their phosphorylation status was analyzed by autoradiography. As anticipated, EnvZ was found to be autophosphorylated likely on the histidine residue (Fig 4a; upper band). In addition to the autophosphorylated EnvZ, we detected efficient

phosphorylation of MtrA and its site-specific mutants (Fig 4a; lower band). Moreover, phosphorylation of MtrA was found to be similar for wild type and mutant proteins suggesting that EnvZ does not differentiate between these substrates (Fig 4a).

Next, we evaluated the DNA binding activity of wild type MtrA with or without EnvZ incubation by EMSA using radiolabeled *oriC* fragment as the DNA probe. It is apparent from the data that there is no DNA: protein complex formation if either MtrA or EnvZ is absent (Fig 4b). We observed DNA binding only upon incubation of phosphorylated MtrA with radiolabeled *oriC* DNA fragment and the binding efficiency was dependent on the concentration of MtrA (Fig 4b). These results show that, EnvZ efficiently phosphorylates MtrA *in vitro*, and phosphorylated MtrA proficiently interacts with the DNA (Fig 4a & 4b).

Finally, we compared the DNA binding activity of MtrA and MtrA mutants that were phosphorylated by EnvZ (Fig 4c). Equal amounts of phosphorylated MtrA, $MtrA_{R122M}$, $MtrA_{K204M}$, or $MtrA_{K207M}$ were incubated with *oriC* DNA probe and EMSA was performed. While we could detect DNA: protein complex with MtrA, and $MtrA_{K207M}$ proteins; mutants $MtrA_{R122M}$ and $MtrA_{K204M}$ showed marginal or no binding, respectively. Since lysine residues can be modified by other PTMs such as acetylation, we analyze whether MtrA was a target of any of these other modifications. We performed additional mass spectrometric analysis to identify lysine modifications on His_6-$MtrA_{Mtb}$ expressed and purified from *Msm*. Mass spectrometric analysis showed the presence of acetylation and succinylation on MtrA and both modifications were found to be on K207 residue, but not on K204 residue, suggesting that the only modification detected on K204 is methylation (Fig S4). Taken together this data suggest that the methylation of R122 and K204 plays an important role in modulating the interaction of MtrA with DNA.

Perturbation of metabolic intermediate levels influences MtrA methylation.

Next, we tried to identify mechanisms that regulate protein methylation. Methylation reactions are catalyzed by SAM-dependent methyltransferases where SAH and consequently Hcy are generated as by-products (Fig 5a). We have previously shown that perturbation of levels of *Mtb* SahH impacts metabolic levels of Hcy and may affect SAH, a potent inhibitor of methyltransferases [19]. Interestingly, *Mtb sahH* (encoding SahH), an essential gene, is present in the genomic vicinity of *mtrA* [29]. This led us to hypothesize that SahH-mediated perturbation in the levels of SAH or Hcy may impact methylation of proteins like MtrA.

To test our hypothesis, we first evaluated the effect of increasing Hcy on the growth of *Msm*. Bacteria were grown in minimal growth medium containing varying concentrations of Hcy and their growth was measured. We found that increasing concentration of Hcy negatively affects bacterial growth in a concentration-dependent manner (Fig 5b & 5c). Results suggested that higher than 0.4 mM Hcy resulted in more than a log-fold decrease in *Msm* CFUs during the exponential growth phase. Therefore, we decided to use a sub-lethal concentration of 0.4 mM for further experiments. We analyzed methylation of MtrA purified from *Msm* grown in the presence or absence of 0.4 mM Hcy using immunoblotting. In

line with our hypothesis, the addition of Hcy resulted in a ~70% decrease in methylation of MtrA (Fig 5d and 5e).

Next, we addressed the influence of overexpressing SahH on MtrA methylation. We analyzed methylation of MtrA purified from *Msm* containing an integrated copy of *Mtb* SahH. We observed that overexpression of SahH also resulted in a ~70% decrease in MtrA methylation levels, presumably because of perturbed SAH levels as SAH is a potent inhibitor of methyltransferases (Fig 5f and 5g). Collectively, the data suggest that perturbation of metabolic intermediates negatively modulates MtrA methylation.

MtrA methylation negatively regulates transcriptional activation.

In Fig 4, we showed that the methylation mimetic mutant of K204 (K204M) does not bind with the DNA. As a corollary, methylation of MtrA should negatively modulate MtrA-mediated transcriptional activation whereas a decrease in the methylation should positively modulate transcriptional activation. Results in Fig 5 showed that the addition of Hcy or overexpression of SahH results in decreased methylation of MtrA. Taken together, we theorized that the addition of Hcy or overexpression of SahH would increase the transcriptional activation by MtrA. MtrA is known to bind to its own promoter and regulates its expression [37]. Thus, we monitored the expression level of *mtrA* in the presence of an increasing concentration of Hcy. *Msm* cells were grown in minimal medium supplemented with increasing concentration of Hcy and *mtrA* expression was measured using quantitative real-time PCR (qRT-PCR). In line with our hypothesis, we observed increased transcription of *mtrA* with an increasing concentration of Hcy (Fig 6a). In these qRT-PCR reactions, the expression was normalized with respect to the expression of *sigA*. To reconfirm these results, we performed a new set of qRT-PCR reactions in the presence of 0.4 mM Hcy, except that the expression of *mtrA* was normalized with respect to the expression of 16S rRNA (Fig 6b). The results were in agreement with the data presented in Fig 6a, confirming that the addition of Hcy increases the expression of *mtrA*. Next, we examined the expression levels of *mtrA* upon expression of SahH by utilizing *Msm* harboring pVV16-*sahH* plasmid (Fig 6c). We observed a ~6-fold increase in the transcript levels of *mtrA* in the presence of overexpressed SahH.

Subsequently, we asked if the addition of Hcy impacts the expression of *sahH* and if so, does MtrA binds to the promoter region of *sahH*. To address this question, we evaluated the expression of *Msm sahH* in the presence of an increasing concentration of Hcy (Fig 6d). The results showed a direct correlation between Hcy concentration and expression of $sahH_{Ms}$. Besides *mtrA* promoter regions, we identified MtrA-binding sites in the putative *sahH* promoter region. Thus to examine if MtrA binds to putative $sahH_{Ms}$ ($sahH_{Ms}$-Pr) and $sahH_{Mt}$ ($sahH_{Mt}$-Pr) promoter regions, we performed EMSA with radiolabeled *sahH* promoter regions from *Msm* and *Mtb*, respectively. While only EnvZ or unphosphorylated MtrA does not bind with the DNA, we observed robust binding of phosphorylated MtrA with both $sahH_{Ms}$-Pr and $sahH_{Mt}$-Pr DNA fragments (Fig 6e). Most importantly, $MtrA_{K204M}$ mutant that showed abrogated binding with *oriC* fragment in the previous EMSA experiments (Fig 4c) failed to bind with both $sahH_{Ms}$-Pr and $sahH_{Mt}$-Pr DNA fragments, confirming that methylation of MtrA negatively modulates DNA binding and hence its activity both *in vitro* and *in vivo*.

Discussion

Covalent modification of side chains of multiple amino acids in proteins regulates their activity and function thus controlling cellular processes [46]. In addition to phosphorylation, which has been extensively investigated, multiple additional modifications have been identified with the help of high throughput mass spectrometry or by candidate-specific approaches. In this report, we used a candidate approach to identify proteins that are methylated on lysine residue, and the methylation of the positive candidates was validated by mass spectrometry. Most of the candidate proteins that were chosen for the study belonged to regulatory protein class (Supplementary Fig. 1a & b), followed by intermediary metabolism, and cell wall and cell processes. While a high throughput mass spectrometry approach may have provided a more comprehensive list of methylated proteins, with our approach we detected methylation of proteins in the functional categories that are relatively less prevalent. We have used a similar candidate approach previously to identify novel acylated proteins in *Mtb* [20]. In an independent study, western blot analysis of MtHU (HupB) expressed and purified from *Msm* revealed the presence of acetylation, and the target sites were subsequently identified by mass spectrometry [47]. We identified a total of 10 proteins by western blot and the mass spectrometry analysis showed nine of them to be methylated on lysine residues. Identification of ten western blot positive methylated proteins among the 72 candidates suggest that methylation could be a more frequent modification in mycobacterial proteins and warrant future large-scale analyses of the whole proteome. We propose that the present study be used in parallel with global proteomics-based approaches in order to have an unbiased analysis of both over- and under-represented protein functional categories in the whole proteome.

Analysis of *Mtb* genome suggests the presence of 57 probable methyltransferases- 29 of them may be involved in intermediary metabolism and respiration, 12 of them are probable lipid methyltransferases, 8 could be involved in RNA methylation, and 7 in DNA methylation (https://mycobrowser.epfl.ch/). To date, only three methyltransferases-Rv1988, Rv2966c, and MamA-have been functionally characterized. Rv1988 is a secretory methyltransferase that enters the host nucleus and methylates histone H3 at arginine residues and regulates the expression of genes involved in combating reactive oxygen species [48]. Rv2966c is also a secretory methyltransferase that localizes to the host nucleus and methylates host DNA at cytosine residues [49]. MamA is a DNA N6-adenine methyltransferase that regulates the expression of multiple genes that provides fitness during hypoxia [50]. Methylation of *Mtb* HBHA and HupB by unknown methyltransferase(s) renders them proteolytic resistant [18]. Recently, a host methyltransferase is shown to methylate *Mtb* HupB to confer protection against invading bacilli [51]. Thus far, lysine/arginine methyltransferases that can act on the bacterial protein targets have not been characterized in mycobacteria. Identification of lysine/arginine methylation of many essential *Mtb* proteins indicate mycobacterial methyltransferases might play important role in pathogenesis and physiology of mycobacteria. Elucidating the specificities and mode of substrate recognition of methyltransferase would help in understanding the biological significance of protein methylation.

To elucidate the functional relevance of methylation in mycobacteria, we chose MtrA, an essential response regulator of TCS MtrB-MtrA in *Mtb*. MtrB is a non-essential membrane-bound sensor kinase that transfers a phosphate group to a conserved aspartate residue (D56) in MtrA. MtrA binds to the promoters of *ripA* (encoding peptidoglycan hydrolase), *fbpA* (encoding secreted antigen 85B), *fbpB* (encoding cell wall mycolyl hydrolase), *dnaA* (encoding replication initiator protein) and *oriC* (origin of replication) and regulates cell cycle progression [37, 52]. Although phosphorylation at D56 is the primary regulatory mechanism for MtrA, the protein has also been reported to be pupylated at K207 [53] and acetylated at K110 [54]. We now show that MtrA is modified by lysine/arginine methylation, lysine acetylation, and lysine succinylation. Different lysine modifications occurring on MtrA might play a role in regulating different aspects of MtrA, such as methylation-mediated regulation of DNA-binding activity and pupylation-mediated regulation of protein turnover rate. Methylation of MtrA on arginine and lysine residues was found to negatively regulate its DNA binding function (Fig 4). Arginine methylation regulates several mammalian processes associated with gene expression but is largely unrecognized in bacteria [8]. Proteomics analysis has revealed >25 arginine methylated proteins in *Leptospira interrogans*, but no functional role has been assigned to them [55]. In most of these proteins, lysine methylation occurs in conjunction with arginine methylation on the same protein as is the case with mycobacterial MtrA. On the contrary, all the arginine methylated proteins of *Desulfovibrio vulgaris* do not contain methyllysine [56]. Although a dimethylarginine was spotted on a *D. vulgaris* transcriptional response regulator DVUA0086, its functional role remained obscure. Further revelation of the role of arginine methylation in bacteria is therefore essential.

Methyltransferase reactions are dependent on the presence of balanced amounts of SAM and SAH as they are prone to SAH-mediated inhibition. Under normal conditions, SAH levels are regulated using SahH-mediated reversible hydrolysis of SAH to Hcy. Hcy supplementation may allow the net flux of this reversible reaction towards SAH synthesis, which can negatively regulate methyltransferases activity. In a similar vein, overexpression of SahH may lead to depletion of SAH, which in turn leads to lower levels of SAM, a substrate for methyltransferases, thus influencing the activity. In agreement with this hypothesis, we observed that the addition of Hcy or overexpression of SahH led to decreased methylation of MtrA (Fig 5). DNA binding experiment suggests that methylation negatively regulates MtrA interaction with DNA (Fig 4) and overexpression of SahH or addition of Hcy decreases MtrA methylation. In accordance, SahH or Hcy were found to increase *mtrA* transcription which may lead to altered expression of genes targeted by MtrA such as *ripA*, *fbpB*, and *dnaA* and regulate cell cycle progression. It is to be noted that SahH has previously been found to be associated with differential DNA and RNA methylation in eukaryotes [57-59], thus pointing towards a more general implication of SahH in regulating one-carbon metabolism. In mycobacteria, one-carbon metabolism pathway involving SAM and methionine biosynthesis has been proposed as a powerful target for anti-mycobacterial agents [60]. Mycobacterial strains deficient in SAM and methionine biosynthesis were found to be remarkably vulnerable in host tissues. Interestingly, such metabolic perturbation was shown to be associated with

altered methylation at DNA and other important metabolites like biotin. In another study, disruption of one-carbon metabolism by antifolate molecules led to efficient killing of *Mtb* [61]. These studies suggest the significance of studying regulators of one-carbon metabolism and highlight SahH as a promising drug target.

In summary, the present study provides a framework for elucidation of protein methylation in mycobacteria. We report the addition of protein arginine methylation to the growing list of regulatory PTMs in mycobacteria and suggest that methylation of MtrA at lysine and arginine residues regulates its activity. This study provides an orchestration of methylation and TCS signaling and therefore illuminates the critical role of methylation in bacterial physiology.

Acknowledgments: The authors acknowledge support from Prof. Matthias Mann (Max Planck Institute of Biochemistry, Germany) for the mass spectrometry for the identification of methylated peptides. The authors thank Christian Hentschker and Döerte Becher (University of Greifswald, Germany) for their help in identifying protein acetylation and succinylation by mass spectrometry.

Funding: This work was supported by the CSIR BSC-0123, BSC-0104, and J.C. Bose fellowship to Y.S.; CSIR research associate fellowship to A.Si.; and CSIR senior research fellowship to R.V. S.N. is funded through a CSIR-Senior research fellowship.

Competing interests: The authors declare that they have no financial or non-financial competing interests.

Data availability: Original mass spectrometry spectra are submitted in Sup Figure 2. All the reagents utilized in the manuscript would be available upon request.

References

1 WHO. (2019) World Health Organization Global Tuberculosis report

2 van Els, C. A., Corbiere, V., Smits, K., van Gaans-van den Brink, J. A., Poelen, M. C., Mascart, F., Meiring, H. D. and Locht, C. (2014) Toward Understanding the Essence of Post-Translational Modifications for the Mycobacterium tuberculosis Immunoproteome. Front Immunol. **5**, 361

3 Canova, M. J. and Molle, V. (2014) Bacterial serine/threonine protein kinases in host-pathogen interactions. J Biol Chem. **289**, 9473-9479

4 Festa, R. A., McAllister, F., Pearce, M. J., Mintseris, J., Burns, K. E., Gygi, S. P. and Darwin, K. H. (2010) Prokaryotic ubiquitin-like protein (Pup) proteome of Mycobacterium tuberculosis [corrected]. PLoS One. **5**, e8589

5 Rhee, K. Y., Erdjument-Bromage, H., Tempst, P. and Nathan, C. F. (2005) S-nitroso proteome of Mycobacterium tuberculosis: Enzymes of intermediary metabolism and antioxidant defense. Proc Natl Acad Sci U S A. **102**, 467-472

6 Sajid, A., Arora, G., Singhal, A., Kalia, V. C. and Singh, Y. (2015) Protein Phosphatases of Pathogenic Bacteria: Role in Physiology and Virulence. Annu Rev Microbiol. **69**, 527-547

7 Schubert, H. L., Blumenthal, R. M. and Cheng, X. (2003) Many paths to methyltransfer: a chronicle of convergence. Trends Biochem Sci. **28**, 329-335

8 Bedford, M. T. and Clarke, S. G. (2009) Protein arginine methylation in mammals: who, what, and why. Mol Cell. **33**, 1-13

9 Lanouette, S., Mongeon, V., Figeys, D. and Couture, J. F. (2014) The functional diversity of protein lysine methylation. Mol Syst Biol. **10**, 724

10 Cain, J. A., Solis, N. and Cordwell, S. J. (2014) Beyond gene expression: the impact of protein post-translational modifications in bacteria. J Proteomics. **97**, 265-286

11 Salah Ud-Din, A. I. M. and Roujeinikova, A. (2017) Methyl-accepting chemotaxis proteins: a core sensing element in prokaryotes and archaea. Cell Mol Life Sci. **74**, 3293-3303

12 Greer, E. L. and Shi, Y. (2012) Histone methylation: a dynamic mark in health, disease and inheritance. Nat Rev Genet. **13**, 343-357

13 Bedford, M. T. (2007) Arginine methylation at a glance. J Cell Sci. **120**, 4243-4246

14 Zhang, X., Wen, H. and Shi, X. (2012) Lysine methylation: beyond histones. Acta Biochim Biophys Sin (Shanghai). **44**, 14-27

15 Barbier, M., Owings, J. P., Martinez-Ramos, I., Damron, F. H., Gomila, R., Blazquez, J., Goldberg, J. B. and Alberti, S. (2013) Lysine trimethylation of EF-Tu mimics platelet-activating factor to initiate Pseudomonas aeruginosa pneumonia. MBio. **4**, e00207-00213

16 Zhang, M., Xu, J. Y., Hu, H., Ye, B. C. and Tan, M. (2018) Systematic Proteomic Analysis of Protein Methylation in Prokaryotes and Eukaryotes Revealed Distinct Substrate Specificity. Proteomics. **18**

17 Pethe, K., Alonso, S., Biet, F., Delogu, G., Brennan, M. J., Locht, C. and Menozzi, F. D. (2001) The heparin-binding haemagglutinin of M. tuberculosis is required for extrapulmonary dissemination. Nature. **412**, 190-194

18 Pethe, K., Bifani, P., Drobecq, H., Sergheraert, C., Debrie, A. S., Locht, C. and Menozzi, F. D. (2002) Mycobacterial heparin-binding hemagglutinin and laminin-binding protein share antigenic methyllysines that confer resistance to proteolysis. Proc Natl Acad Sci U S A. **99**, 10759-10764

19 Singhal, A., Arora, G., Sajid, A., Maji, A., Bhat, A., Virmani, R., Upadhyay, S., Nandicoori, V. K., Sengupta, S. and Singh, Y. (2013) Regulation of homocysteine metabolism by Mycobacterium tuberculosis S-adenosylhomocysteine hydrolase. Sci Rep. **3**, 2264

20 Singhal, A., Arora, G., Virmani, R., Kundu, P., Khanna, T., Sajid, A., Misra, R., Joshi, J., Yadav, V., Samanta, S., Saini, N., Pandey, A. K., Visweswariah, S. S., Hentschker, C., Becher, D., Gerth, U. and Singh, Y. (2015) Systematic Analysis of Mycobacterial Acylation Reveals First Example of Acylation-mediated Regulation of Enzyme Activity of a Bacterial Phosphatase. J Biol Chem. **290**, 26218-26234

21 Santhosh, R. S., Pandian, S. K., Lini, N., Shabaana, A. K., Nagavardhini, A. and Dharmalingam, K. (2005) Cloning of mce1 locus of Mycobacterium leprae in Mycobacterium smegmatis mc2 155 SMR5 and evaluation of expression of mce1 genes in M. smegmatis and M. leprae. FEMS Immunol Med Microbiol. **45**, 291-302

22 Sajid, A., Arora, G., Gupta, M., Singhal, A., Chakraborty, K., Nandicoori, V. K. and Singh, Y. (2011) Interaction of Mycobacterium tuberculosis elongation factor Tu with GTP is regulated by phosphorylation. J Bacteriol. **193**, 5347-5358

23 Schneider, C. A., Rasband, W. S. and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. Nat Methods. **9**, 671-675

24 Sharma, K., D'Souza, R. C., Tyanova, S., Schaab, C., Wisniewski, J. R., Cox, J. and Mann, M. (2014) Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. Cell Rep. **8**, 1583-1594

25 Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V. and Mann, M. (2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. J Proteome Res. **10**, 1794-1805

26 Pavel Sinitcyn, J. D. R., * and Jürgen Cox. (May 4, 2018) Computational Methods for Understanding Mass Spectrometry–Based Shotgun Proteomics Data. Annual Review of Biomedical Data Science. **Vol. 1:207-234**

27 Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J. V. and Mann, M. (2009) A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. Nat Protoc. **4**, 698-705

28 Lew, J. M., Kapopoulou, A., Jones, L. M. and Cole, S. T. (2011) TubercuList--10 years after. Tuberculosis (Edinb). **91**, 1-7

29 DeJesus, M. A., Gerrick, E. R., Xu, W., Park, S. W., Long, J. E., Boutte, C. C., Rubin, E. J., Schnappinger, D., Ehrt, S., Fortune, S. M., Sassetti, C. M. and Ioerger, T. R. (2017) Comprehensive Essentiality Analysis of the Mycobacterium tuberculosis Genome via Saturating Transposon Mutagenesis. MBio. **8**

30 Sassetti, C. M., Boyd, D. H. and Rubin, E. J. (2003) Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol. **48**, 77-84

31 Griffin, J. E., Gawronski, J. D., Dejesus, M. A., Ioerger, T. R., Akerley, B. J. and Sassetti, C. M. (2011) High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. PLoS Pathog. **7**, e1002251

32 Sassetti, C. M. and Rubin, E. J. (2003) Genetic requirements for mycobacterial survival during infection. Proc Natl Acad Sci U S A. **100**, 12989-12994

33 Friedland, N., Mack, T. R., Yu, M., Hung, L. W., Terwilliger, T. C., Waldo, G. S. and Stock, A. M. (2007) Domain orientation in the inactive response regulator Mycobacterium tuberculosis MtrA provides a barrier to activation. Biochemistry. **46**, 6733-6743

34 Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. **25**, 1605-1612

35 Lee, C. and Huang, C. H. (2013) LASAGNA-Search: an integrated web tool for transcription factor binding site search and visualization. Biotechniques. **54**, 141-153

36 Fol, M., Chauhan, A., Nair, N. K., Maloney, E., Moomey, M., Jagannath, C., Madiraju, M. V. and Rajagopalan, M. (2006) Modulation of Mycobacterium tuberculosis proliferation by MtrA, an essential two-component response regulator. Mol Microbiol. **60**, 643-657

37 Rajagopalan, M., Dziedzic, R., Al Zayer, M., Stankowska, D., Ouimet, M. C., Bastedo, D. P., Marczynski, G. T. and Madiraju, M. V. (2010) Mycobacterium tuberculosis origin of replication and the promoter for immunodominant secreted antigen 85B are the targets of MtrA, the essential response regulator. J Biol Chem. **285**, 15816-15827

38 Gupta, M., Sajid, A., Sharma, K., Ghosh, S., Arora, G., Singh, R., Nagaraja, V., Tandon, V. and Singh, Y. (2014) HupB, a nucleoid-associated protein of Mycobacterium tuberculosis, is modified by serine/threonine protein kinases in vivo. J Bacteriol. **196**, 2646-2657

39 Arora, G., Sajid, A., Arulanandh, M. D., Singhal, A., Mattoo, A. R., Pomerantsev, A. P., Leppla, S. H., Maiti, S. and Singh, Y. (2012) Unveiling the novel dual specificity protein kinases in Bacillus anthracis: identification of the first prokaryotic dual specificity tyrosine phosphorylation-regulated kinase (DYRK)-like kinase. J Biol Chem. **287**, 26749-26763

40 Ishihama, Y., Schmidt, T., Rappsilber, J., Mann, M., Hartl, F. U., Kerner, M. J. and Frishman, D. (2008) Protein abundance profiling of the Escherichia coli cytosol. BMC Genomics. **9**, 102

41 Ojha, A., Anand, M., Bhatt, A., Kremer, L., Jacobs, W. R., Jr. and Hatfull, G. F. (2005) GroEL1: a dedicated chaperone involved in mycolic acid biosynthesis during biofilm formation in mycobacteria. Cell. **123**, 861-873

42 Li, Y., Zeng, J. and He, Z. G. (2010) Characterization of a functional C-terminus of the Mycobacterium tuberculosis MtrA responsible for both DNA binding and interaction with its two-component partner protein, MtrB. J Biochem. **148**, 549-556

43 Martinez-Hackert, E. and Stock, A. M. (1997) Structural relationships in the OmpR family of winged-helix transcription factors. J Mol Biol. **269**, 301-312

44 Hyland, E. M., Molina, H., Poorey, K., Jie, C., Xie, Z., Dai, J., Qian, J., Bekiranov, S., Auble, D. T., Pandey, A. and Boeke, J. D. (2011) An evolutionarily 'young' lysine residue in histone H3 attenuates transcriptional output in Saccharomyces cerevisiae. Genes Dev. **25**, 1306-1319

45 Plocinska, R., Purushotham, G., Sarva, K., Vadrevu, I. S., Pandeeti, E. V., Arora, N., Plocinski, P., Madiraju, M. V. and Rajagopalan, M. (2012) Septal localization of the Mycobacterium tuberculosis MtrB sensor kinase promotes MtrA regulon expression. J Biol Chem. **287**, 23887-23899

46 Pejaver, V., Hsu, W. L., Xin, F., Dunker, A. K., Uversky, V. N. and Radivojac, P. (2014) The structural and functional signatures of proteins that undergo multiple events of post-translational modification. Protein Sci. **23**, 1077-1093

47 Ghosh, S., Padmanabhan, B., Anand, C. and Nagaraja, V. (2016) Lysine acetylation of the Mycobacterium tuberculosis HU protein modulates its DNA binding and genome organization. Mol Microbiol. **100**, 577-588

48 Yaseen, I., Kaur, P., Nandicoori, V. K. and Khosla, S. (2015) Mycobacteria modulate host epigenetic machinery by Rv1988 methylation of a non-tail arginine of histone H3. Nat Commun. **6**, 8922

49 Sharma, G., Upadhyay, S., Srilalitha, M., Nandicoori, V. K. and Khosla, S. (2015) The interaction of mycobacterial protein Rv2966c with host chromatin is mediated through non-CpG methylation and histone H3/H4 binding. Nucleic Acids Res. **43**, 3922-3937

50 Shell, S. S., Prestwich, E. G., Baek, S. H., Shah, R. R., Sassetti, C. M., Dedon, P. C. and Fortune, S. M. (2013) DNA methylation impacts gene expression and ensures hypoxic survival of Mycobacterium tuberculosis. PLoS Pathog. **9**, e1003419

51 Yaseen, I., Choudhury, M., Sritharan, M. and Khosla, S. (2018) Histone methyltransferase SUV39H1 participates in host defense by methylating mycobacterial histone-like protein HupB. EMBO J. **37**, 183-200

52 Purushotham, G., Sarva, K. B., Blaszczyk, E., Rajagopalan, M. and Madiraju, M. V. (2015) Mycobacterium tuberculosis oriC sequestration by MtrA response regulator. Mol Microbiol. **98**, 586-604

53 Witze, E. S., Old, W. M., Resing, K. A. and Ahn, N. G. (2007) Mapping protein post-translational modifications with mass spectrometry. Nat Methods. **4**, 798-806

54 Singh, K. K., Bhardwaj, N., Sankhe, G. D., Udaykumar, N., Singh, R., Malhotra, V. and Saini, D. K. (2019) Acetylation of Response Regulator Proteins, TcrX and MtrA in M. tuberculosis Tunes their Phosphotransfer Ability and Modulates Two-Component Signaling Crosstalk. J Mol Biol. **431**, 777-793

55 Cao, X. J., Dai, J., Xu, H., Nie, S., Chang, X., Hu, B. Y., Sheng, Q. H., Wang, L. S., Ning, Z. B., Li, Y. X., Guo, X. K., Zhao, G. P. and Zeng, R. (2010) High-coverage proteome analysis reveals the first insight of protein modification systems in the pathogenic spirochete Leptospira interrogans. Cell Res. **20**, 197-210

56 Gaucher, S. P., Redding, A. M., Mukhopadhyay, A., Keasling, J. D. and Singh, A. K. (2008) Post-translational modifications of Desulfovibrio vulgaris Hildenborough sulfate reduction pathway proteins. J Proteome Res. **7**, 2320-2331

57 Baric, I., Fumic, K., Glenn, B., Cuk, M., Schulze, A., Finkelstein, J. D., James, S. J., Mejaski-Bosnjak, V., Pazanin, L., Pogribny, I. P., Rados, M., Sarnavka, V., Scukanec-Spoljar, M., Allen, R. H., Stabler, S., Uzelac, L., Vugrek, O., Wagner, C., Zeisel, S. and Mudd, S. H. (2004) S-adenosylhomocysteine hydrolase deficiency in a human: a genetic disorder of methionine metabolism. Proc Natl Acad Sci U S A. **101**, 4234-4239

58 Mull, L., Ebbs, M. L. and Bender, J. (2006) A histone methylation-dependent DNA methylation pathway is uniquely impaired by deficiency in Arabidopsis S-adenosylhomocysteine hydrolase. Genetics. **174**, 1161-1171

59 Radomski, N., Kaufmann, C. and Dreyer, C. (1999) Nuclear accumulation of S-adenosylhomocysteine hydrolase in transcriptionally active cells during development of Xenopus laevis. Mol Biol Cell. **10**, 4283-4298

60 Berney, M., Berney-Meyer, L., Wong, K. W., Chen, B., Chen, M., Kim, J., Wang, J., Harris, D., Parkhill, J., Chan, J., Wang, F. and Jacobs, W. R., Jr. (2015) Essential roles of methionine and S-adenosylmethionine in the autarkic lifestyle of Mycobacterium tuberculosis. Proc Natl Acad Sci U S A. **112**, 10008-10013

61 Nixon, M. R., Saionz, K. W., Koo, M. S., Szymonifka, M. J., Jung, H., Roberts, J. P., Nandakumar, M., Kumar, A., Liao, R., Rustad, T., Sacchettini, J. C., Rhee, K. Y., Freundlich, J. S. and Sherman, D. R. (2014) Folate pathway disruption leads to critical disruption of methionine derivatives in Mycobacterium tuberculosis. Chem Biol. **21**, 819-830

Figure legends

Figure 1. Multiple Mtb proteins are methylated on lysine and arginine residues. (a) Ten recombinant proteins containing His_6-tag were purified using Ni^{2+}-NTA beads from *Msm*. Purified proteins were loaded on SDS-PAGE, transferred on nitrocellulose, and probed with anti-methyllysine or anti-His_6 antibody. (b) Table shows the number and location of methylation sites in the recombinant proteins by mass spectrometry. The corresponding supplementary image number of mass spectra is also mentioned.

Figure 2. MtrA is methylated in Mtb. (a) Pictorial representation of MtrA domain organization showing the location of different methylated sites. (b) His_6-MtrA was overexpressed and purified from *Mtb*. Purified protein was probed with anti-methyllysine antibody (upper panel) and anti-MtrA (lower panel) antibody. Histone was used as a positive control. (c) Structural representation of MtrA (PDB ID: 2GWR). Response regulator domain (blue), linker region (green), and DNA binding domain (pink) are visible with the three crucial methylated sites marked red.

Figure 3. MtrA is methylated at lysine and arginine residues. (a) MtrA and its mutants were over-expressed and purified from *Msm*. Purified proteins were loaded on SDS-PAGE, transferred on nitrocellulose, and probed with anti-methyllysine antibody. Ponceau-stained membrane image is shown in the lower image. (b) Densitometric analysis of the western blot shown in (a). The bar graph depicts intensities obtained after normalization with protein amounts detected by ponceau staining. The intensity of methylated MtrA was considered as 100% and relative methylation intensities of mutants are plotted. Data (mean ± s.d.) are from three individual replicates. **** $P \leq 0.0001$, as determined by two-tailed unpaired Student's *t*-test. (c) Heat map showing the effect of mutation of R122, K204, and K207 residues on methylation intensities at other sites. Each row depicts the residue at which quantitative analysis was performed and each column represents the protein analyzed. Mass spectrometric intensities are color-coded according to the key given below the heat map (log_2 scale).

Figure 4. Role of methylated residues in DNA-binding activity. (a) MtrA WT and mutants were expressed and purified from *E. coli* and equal amounts were phosphorylated by EnvZ in the presence of $\gamma[^{32}P]ATP$. The reactions were resolved on SDS-PAGE, coomassie stained (lower panel), and autoradiographed (upper panel). (b) Radiolabelled *Mtb oriC* DNA probe was synthesized using $\gamma[^{32}P]ATP$ by PCR. DNA binding assay was performed in the presence of unphosphorylated (6 μg, lane 2) or phosphorylated MtrA (6 μg and 2 μg, lanes 4 and 5). As a control, reactions were performed in the absence of MtrA (lane 3) or without any protein (lane 1). The reactions were resolved on native PAGE and gels were autoradiographed. DNA: protein complex and the unbound DNA are shown. (c) Radiolabelled *oriC* probe was incubated in the presence of 0-100 μM phosphorylated MtrA, $MtrA_{K204}$, $MtrA_{K207}$, or $MtrA_{R122}$ proteins. The reactions were resolved on native PAGE and gels were autoradiographed. DNA: protein complex and the unbound DNA are shown.

Figure 5. The perturbation of metabolic intermediate levels influences MtrA methylation. (a) Reaction showing the synthesis of Hcy from SAH catalyzed by SahH. (b & c) *Msm* cells were grown in

the presence of 0-0.8 mM Hcy and growth was measured. Data (mean ± s.d.) are from four individual replicates. (b) A_{600} was plotted as a function of time. (c) Graph shows Log_{10}(CFU/ml) calculated during the exponential growth phase as a function of Hcy concentration. (d & e) MtrA was expressed and purified from *Msm* using Ni^{2+}-NTA chromatography in the absence or presence of 0.4 mM Hcy. (d) Immunoblotting was performed using anti-methyllysine antibody followed by an anti-His_6 antibody. (e) Graph showing the relative methylation of MtrA in the presence of Hcy with respect to methylation of MtrA in the absence of Hcy. Methyllysine intensities were normalized to MtrA protein levels as measured by anti-His_6 immunoblot. Data (mean ± s.d.) are from three individual replicates. (f & g) MtrA was expressed and purified from *Msm* strain that overexpressed SahH using Ni^{2+}-NTA chromatography. (f) Immunoblotting was performed using anti-methyllysine antibody followed by anti-His_6 and anti-MtrA antibodies. (g) Graph showing the methylation of MtrA in the presence of overexpressed SahH relative to the methylation of MtrA in the presence of vector control. Methyllysine intensities were normalized to MtrA protein levels as measured by anti-His_6 immunoblot. Data (mean ± s.d.) are from three individual replicates. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, as determined by two-tailed unpaired Student's *t*-test.

Figure 6. MtrA methylation negatively regulates transcriptional activation. (a-c) *mtrA* expression was analyzed using qRT-PCR in *Msm* cultures grown in the presence of Hcy (a & b) or SahH overexpression (c). Expression level of *mtrA* was analyzed with respect to *sigA* (a & c) or 16S rRNA (b). (d) *sahH* expression was analyzed using qRT-PCR in *Msm* cultures grown in the presence of Hcy with respect to *sigA*. Data (mean ± s.d.) are from six (a) or three (b & c) biological triplicates. * $P \leq 0.05$, ** $P \leq 0.01$, **** $P \leq 0.0001$ as determined by two-tailed unpaired Student's *t*-test (compared to control values). (e) DNA binding assay was performed using putative *sahH* promoter fragments from *Mtb* ($sahH_{Mt}$-Pr) or *Msm* ($sahH_{Ms}$-Pr). MtrA and $MtrA_{K204M}$ were used in unphosphorylated or phosphorylated forms. The reactions were resolved on native PAGE and gels were autoradiographed. DNA:protein complex and the unbound DNA are shown.

727 **Table 1**. List of primers used in this study.

Primer Name	Primer sequence (5′→3′)[b]
MtrA F	GTCCCGATGTGGTGACATATGGACACCATGAGGC (NdeI)
MtrA R	GCATCGTCGCCGGCGAAGCTTCGGAGGTCCGGCCTTG (HindIII)
EnvZ F	ACGGCTCGGATCCATGAGGCGATTGCGCTTC (BamHI)
EnvZ R	CCTTCGCCTCAAGCTTATTTACCCTTCTTTTG (HindIII)
MtrAK204M F	GTCCAGCGTCTGCGGGCCATGGTCGAAAAGGATCCCGAG
MtrAK204M R	CTCGGGATCCTTTTCGACCATGGCCCGCAGACGCTGGAC
MtrAK207M F	CTGCGGGCCAAGGTCGAAATGGATCCCGAGAACCCGACTG
MtrAK207M R	CAGTCGGGTTCTCGGGATCCATTTCGACCTTGGCCCGCAG
MtrAR122M F	GGTGCGGGCGCGGCTGCGCATGAACGACGACGAACCCGCCG
MtrAR122M R	CGGCGGGTTCGTCGTCGTTCATGCGCAGCCGCGCCCGCACC
SahH F	GGATGAAAGCCCATATGACCGGAAATTTGG (NdeI)
SahH R	TGGGCGATTTTGCGTAAGCTTGCGGGTGGGA (HindIII)
OriC F	CACGGCGTGTTCTTCCGAC
OriC R	GTCGGAGTTGTGGATGACGG
$SahH_{Ms}$-pr F	GCGCCTGGCGATGAGCTACG
$SahH_{Ms}$-pr R	GCACACTCATGCCGACAACC
$SahH_{Mt}$-pr F	GCGGCTGTGCTTGAGCTACG
$SahH_{Mt}$-pr R	GCTCACAGGGATCCGAGCG
MtrART F	CCATCGTTCTGCGTGGTGAG
MtrART R	GGTCAGCATGACGATCGGC
SahH-RT F	GCGCCAAGAAGATCAACATC
SahH-RT R	CTCGGACAGCACGATGATC
SigA-RT F	CGTTCCTCGACCTCATCCA
SigA-RT R	GCCCTTGGTGTAGTCGAACTTC
16S-RT F	AATTCGATGCAACGCGAAGA
16S-RT R	GCGGGACTTAACCCAACATC

728 [a]'F' denotes forward primer and 'R' denotes reverse primer.

729 [b]Restriction sites/mutations are underlined and Restriction enzymes are mentioned in parenthesis.

730

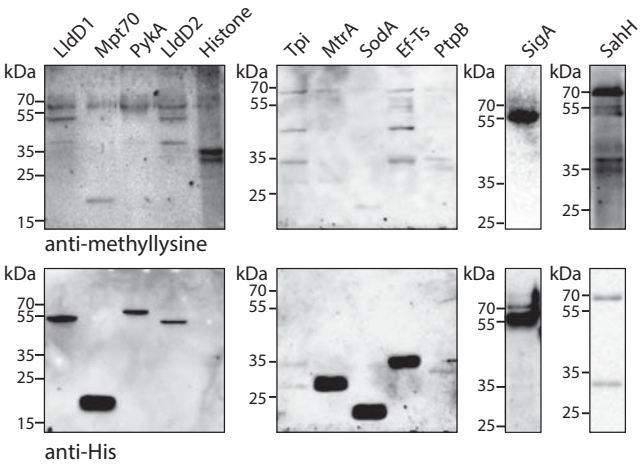
731 **Table 2**: List of plasmids used in this study.

Plasmid construct	Description	Reference or source
pVV16	Mycobacterial expression vector with kanamycin resistance	[22]
pET28a	*E. coli* expression vector with His_6-tag and kanamycin resistance	Novagen
pMAL-c2x	*E. coli* expression vector with MBP-tag and ampicillin resistance	New England BioLabs
pSET152	Mycobacterial integrative vector with apramycin resistance	[21]

732

Figure 1

a.



b.

Protein Rv number	Protein name	Number of sites detected	Site Locations	Mass Spectrum
Rv0694	LldD1	None	None	None
Rv2875	Mpt70	one	K118	Sup Fig 2a
Rv1872c	LldD2	one	R11	Sup Fig 2b
Rv1438	Tpi	four	R103, R139, R181, R196	Sup Fig 2c-2f
Rv3246c	MtrA	thirteen	R46, R72, R122, R167, R200, R219	Sup Fig 2g-2l
			K85, K105, K108, K110, K144, K204, K207	Sup Fig 2m-2s
Rv3846	SodA	six	R110, K38, K46, K53, K81, K179	Sup Fig 2t-2y
Rv2889c	Ef-Ts	four	R150, R151, R262, K100	Sup Fig 2z-2ac
Rv0153c	PtpB	three	R56, R154, K164	Sup Fig 2ad-2af
Rv2703	SigA	one	K15	Sup Fig 2ag
Rv3248c	SahH	five	R448, K8, K276, K277, K471	Sup Fig 2ah-2al

Figure 2

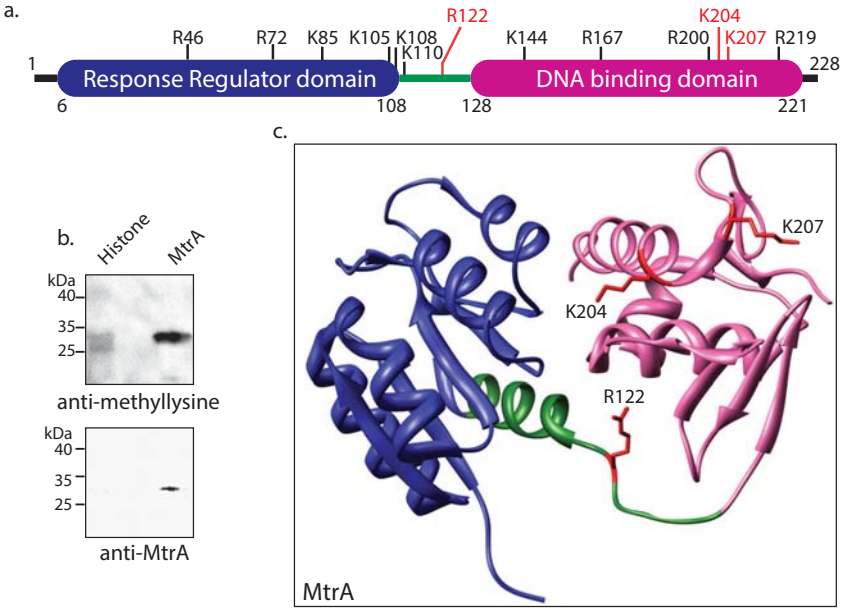


Figure 3

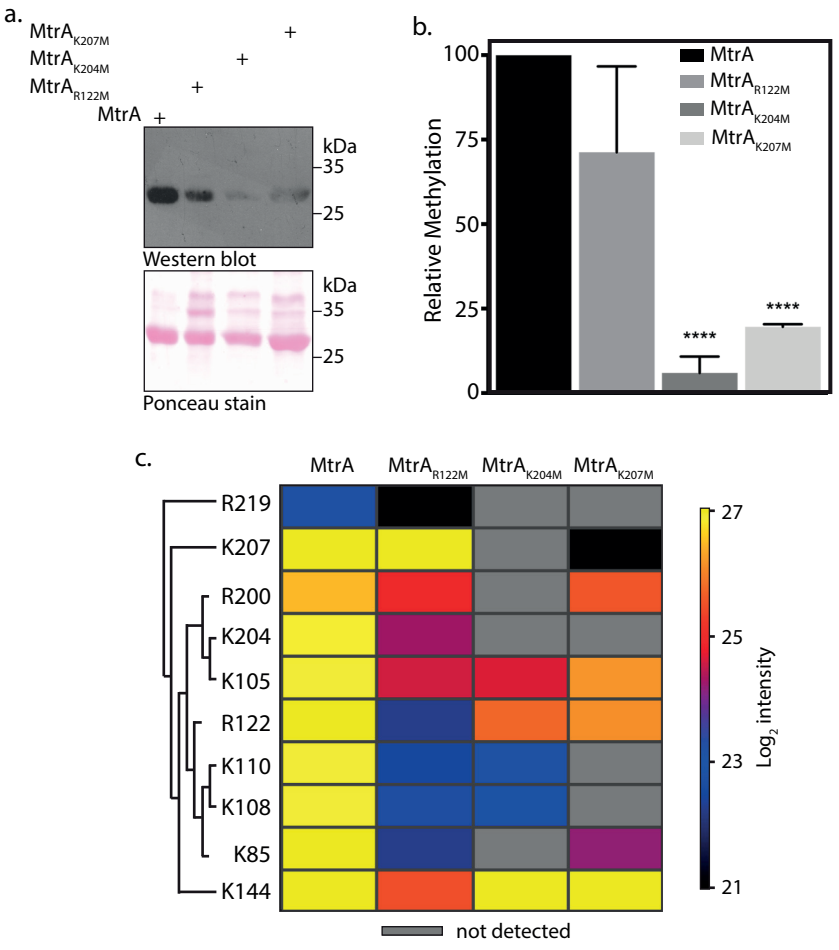


Figure 4

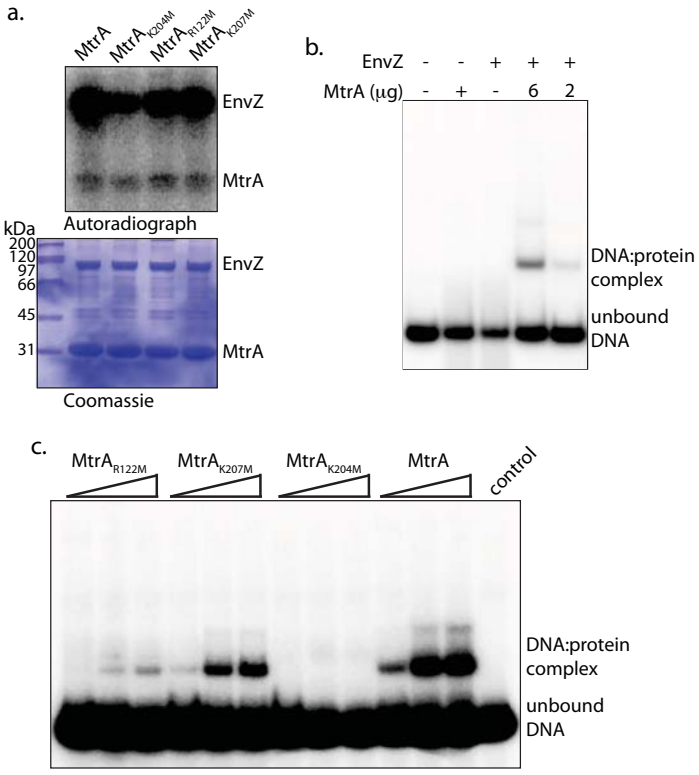


Figure 5

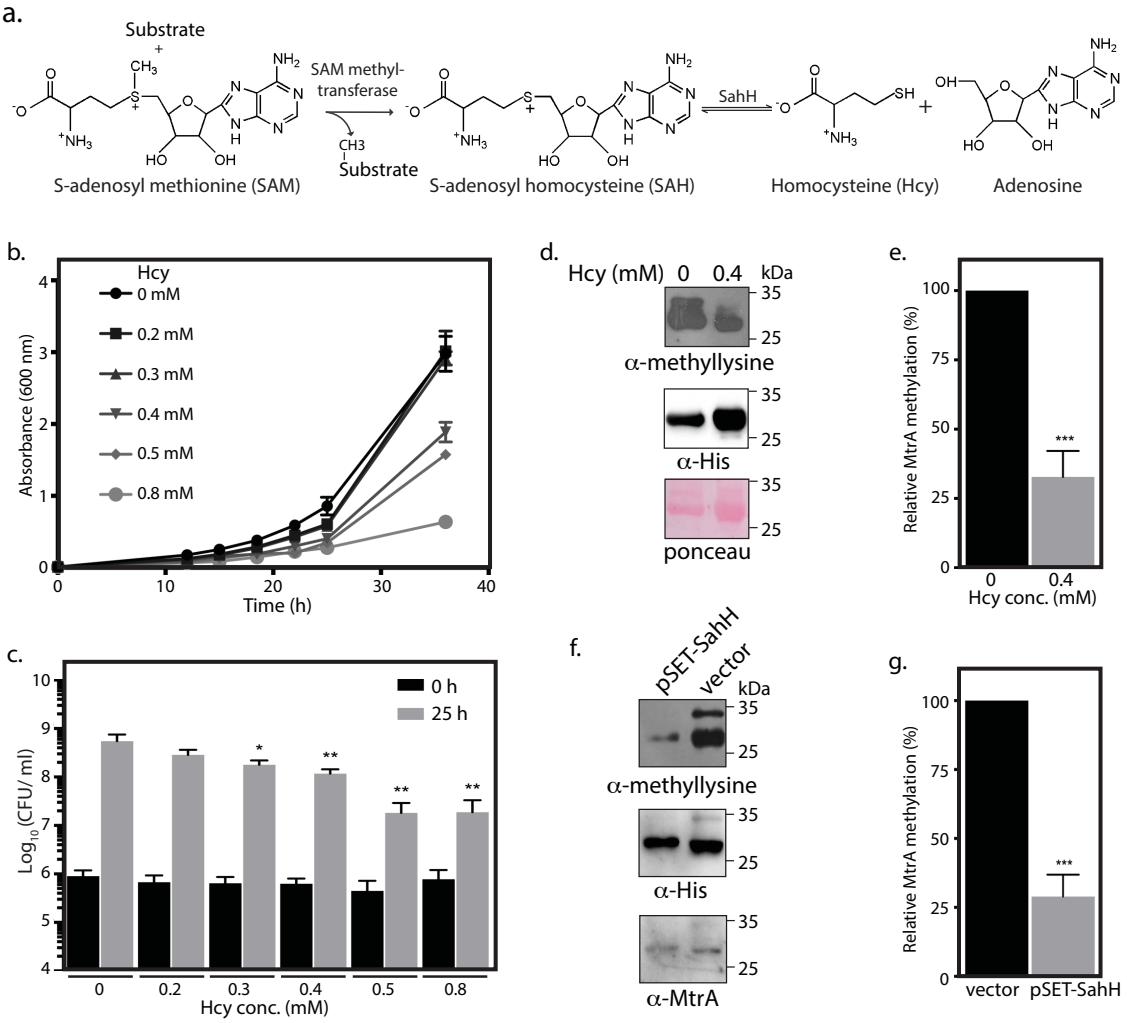


Figure 6