April 2019

Investigating the Role of DNA Methylation in Mycobacterium smegmatis

Samantha Randall
Worcester Polytechnic Institute

Investigating the Role of DNA Methylation in *Mycobacterium smegmatis*

A Major Qualifying Project

Submitted to the Faculty of

Worcester Polytechnic Institute

In partial fulfilment of the requirements for the

Degree of Bachelor of Science

In

Biology and Biotechnology

By

Samantha Randall

Date: 4/25/19

Project Advisor: Scarlet Shell, PhD

ABSTRACT

Mycobacterium tuberculosis is a pathogenic bacterium of great medical relevance, causing disease in millions of people annually and living dormant in many more. During infection, *M. tuberculosis* is exposed to harsh environmental conditions that require it to utilize adaptive mechanisms to survive. One such mechanism is DNA methylation. In *M. tuberculosis*, DNA methylation has been observed to influence gene expression and fitness in hypoxia. To further understand the role of DNA methylation in mycobacteria, we knocked down the expression of the methyltransferase MamA in *Mycobacterium smegmatis*. Expression of *mamA* was repressed using the CRISPR interference knockdown system, which utilizes a modified version of CRISPR-Cas9 to block transcription of *mamA*. The cells with reduced expression of *mamA* demonstrated growth cessation, elongated cell bodies, reduced distribution of DNA, and abnormal septation within cells. These results suggest that MamA could play a role in DNA replication and/or division. Further investigation into the state and localization of *M. smegmatis*' DNA during depletion of *mamA* may provide more information on how MamA influences cell growth and division.

INTRODUCTION

Tuberculosis, a disease caused by the pathogen *Mycobacterium tuberculosis* (Mtb), is a globally relevant disease that is one of the largest causes of death worldwide. In 2017, there were approximately 10 million new cases of tuberculosis and 1.6 million deaths caused by the disease (WHO, 2018). Additionally, nearly one quarter of the world's population is latently infected with Mtb but does not have active tuberculosis disease (WHO, 2018). In order to effectively fight Mtb infections and develop new treatments for drug-resistant strains, it is vital to understand the basic biological functions of this bacteria. For example, cell cycle regulation is an important function in the growth and division of bacteria. In organisms such as *Escherichia coli* and *Caulobacter crescentus*, this process is fairly well characterized, but in Mtb much of the process is unknown (Skerker & Laub, 2004; Thanky, Young, & Robertson, 2007).

DNA methylation plays an important role in cell cycle regulation in some organisms. DNA methylation is the addition of methyl groups to adenine or cytosine bases of DNA via enzymes called DNA methyltransferases. There are three major types of methylated bases: N^6-methyladenine, N^4-methylcytosine, and 5-methylcytosine (Sanchez-Romero, Cota, & Casadesus, 2015). The methyl groups are located in the major groove of the double helix, lowering the thermodynamic stability of DNA, and can alter DNA-protein interactions (Engel & von Hippel, 1978; Sanchez-Romero, Cota, & Casadesus, 2015). DNA methylation has a large variety of uses, including protection against foreign DNA as part of restriction-modification systems, regulation of gene expression, mismatch-repair, initiation of chromosome replication, and host-pathogen interactions (Marinus & Casadesus, 2009; Sanchez-Romero, Cota, & Casadesus, 2015). Two of the most characterized DNA methyltransferases in the bacterial world are Dam from *E. coli* and CcrM from *C. crescentus*.

Dam (DNA adenine methyltransferase) is an N^6-adenine methyltransferase which acts on hemimethylated DNA that has just gone through DNA replication (Marinus & Casadesus, 2009; Adhikari & Curtis, 2016). Dam is cited to be regulated primarily through transcription. While Dam is not essential to *E. coli* survival, knockout of Dam results in increased spontaneous mutation rate and reduced coordination during DNA replication initiation (Adhikari & Curtis, 2016). Both the *oriC* and *dnaA* promoter of *E. coli* are enriched with Dam methylation sites, and

initiation of replication is dependent upon the methylation state of these sites. When these sites are hemimethylated, the *oriC* and *dnaA* are sequestered by SeqA, whereas when DNA is fully methylated increased transcription of *dnaA* and replication initiation can occur (Slater et al, 1995). In addition to aiding in DNA-damage repair and initiation of chromosome replication, Dam has been shown to influence the regulation of gene expression (Marinus & Casadesus, 2009). Dam has also been discovered in other gammaproteobacteria such as *Salmonella enterica*, *Vibrio cholerae,* and *Yersinia pseudotuberculosis* where they have roles in viability, virulence, and chromosome replication (Marinus & Casadesus, 2009; Julio et al., 2001; Demarre, Chattoraj, & Burkholder, 2010).

CcrM, or cell-cycle regulated methyltransferase, is a N^6-adenine methyltransferase found in *C. crescentus* and other alphaproteobacteria (Gonzalez, Kozdon, McAdams, Shapiro, & Collier, 2014). CcrM is highly regulated at both the transcriptional and post-translational levels, and activity of the methyltransferase is limited to the predivisional stage of the cell cycle just at the end of chromosome replication (Kahng & Shapiro, 2001; Gonzalez et al., 2014). In *C. crescentus*, CcrM is involved in controlling the expression of various genes including *ctrA, gyrA*, *gyrB*, and *ftsZ*, which are important to cell cycle regulation and cell division (Brilli et al., 2010; Kozdon et al, 2013; Gonzalez et al., 2014). CcrM was regarded as having influence over replication due to the methylation sites near the *oriC* and through regulating *ctrA* and *dnaA* (Collier, McAdams, & Shaprio, 2007; Kozdon et al, 2013). Recently, however, some studies have found that CcrM does not affect transcription of *dnaA* and that methylation via CcrM is not required for initiation of replication (Felletti, Omnus, & Jonas, 2018; Gonzalez et al., 2014). While it is widely debated whether CcrM influences initiation of replication, it does regulate important genes that are involved in the regulation of cell cycle progression and cell division.

In Mtb there are three reported adenine methyltransferases: MamA, MamB, and HsdM (Zhu et al., 2016). MamA is the major N^6-adenine methyltransferase in the Euro-American lineage of Mtb (Shell et al., 2013). Its recognition motif is CTGGAG, which appears in the genome 1816 times (Shell et al., 2013). Knockout of *mamA* in Mtb has been shown to alter the expression of some genes and reduce fitness in hypoxia (Shell et al., 2013). When mapping transcriptional start sites, it was found that there was a consistent overlap between the MamA

methylation motif and the -10 sigma factor binding site of the genes affected by *mamA* knockout. Homologs of MamA can be found in other relatives of Mtb, including *Mycobacterium smegmatis* (*M. smegmatis*), which is the organism of study within this report. While the MamA encoded by *M. smegmatis* is homologous to MamA and in Mtb, there are notable differences. These differences include that MamA is the only functional methyltransferase in *M. smegmatis*, the recognition motif in *M. smegmatis* is CTCGAG, and that MamA is essential in *M. smegmatis* but not in Mtb.

DNA methylation is poorly understood in mycobacteria, and with the discovery that MamA can impact gene regulation and survival in hypoxia, we sought to learn more about MamA. By studying MamA in *M. smegmatis*, we can expand our understanding of how DNA methylation can affect cell growth and division. In this study we use an inducible CRISPR interference (CRISPRi) system to knockdown MamA within *M. smegmatis* to elucidate the role DNA methylation plays in cell growth. Through growth curve determinations and microscopy, we found that knockdown of MamA results in growth cessation and morphological defects such as cell body elongation. Additionally, knockdown cells show a reduced distribution of DNA within the cell, abnormal septation patterns, and production of small cells lacking DNA.

MATERIALS AND METHODS

Strain Construction

M. smegmatis strains were derived from strain mc^{2}155. Strains were made using either wild type mc^{2}155. A total of five different constructs were created for the experiments performed (Table 1). All constructs were created using NEB HiFi Assembly. The plasmid pSS229 was derived from pJR962 (Rock *et al.*, 2017) and contains a tet repressor, kanamycin resistance marker, and attP sequence and phage integrase for integration into the L5 site, dCas9, and an sgRNA, in which the targeting region is 20 nts of the *mamA* (msmeg_3213) coding sequence. The plasmid pSS317 was derived from pSS047 (Morris, Marinelli, Jacobs-Sera, Hendrix & Hatfull, 2008) and contains the full *mamA* coding sequence with two point mutations at nucleotides 96, C to G, and 99, C to T. All plasmids and inserts were amplified using Q5 PCR (New England Biolabs), and details of PCR conditions are outlined in PCR Variations section. HiFi assembly was performed by adding 5 µL of HiFi assembly mix with 5 µL of equal parts (ranging 30-50 ng) of each fragment for assembly for a total volume of 10 µL. The assembly mixture was incubated at 50℃ for 30 minutes before being used for *E. coli* transformation.

Table 1: Plasmid Constructs

Plasmid number	Description	Reference
pSS229	MamA Knockdown construct. Contains dCas9, sgRNA components, tet repressor, and part of *mamA* coding sequence. Kan marked and integrates in the L5 site.	This work
pSS317	MamA rescue construct. Contains *mamA* coding sequence with two-point mutations at nucleotides 96, C to G, and 99, C to T. Hyg marked and integrates in the Giles site.	This work
pJR962	Nonspecific CRISPRi construct. Contains dCas9, nonspecific sgRNA component, tet repressor. Kan marked and integrates in the L5 site.	Rock *et al*. 2017

Culturing Conditions

All bacteria were cultured in Middlebrook 7H9 media supplemented with ADC (Albumin Dextrose Catalase, final concentrations 0.5 g/L bovine serum albumin fraction V, 20 g/L dextrose, 8.5 g/L sodium chloride, and 30 mg/L catalase), 0.2% glycerol and 0.05% Tween 80. Cells were cultured in shaking incubators set to 37 degrees C with a shaking speed of 200 rpm. Strains containing pSS229 were cultured with 50 μg/mL of kanamycin, and strains containing pSS317 were cultured with 250 μg/mL of hygromycin.

Transformation of *E. coli* and *M. smegmatis*

Constructs were transformed into NEB-5-alpha Competent *E. coli* cells (New England Biolabs). Competent cells were thawed on ice and split into two tubes, 5 μL for control and the remaining was used for the transformation. 10 μL of assembly product was added to the transformation tube of cells. Cells were chilled on ice before heat shocking at 42℃. 900 μL of LB media was then added to the tubes and cells were incubated at 37℃ for one hour at a shaking speed of 200 rpm. 25 μL of cells were then spread onto selection LB plates, then remaining cells were centrifuged for 30 seconds at 1300 rpm before discarding all but 50 μL of the supernatant. Cells were then resuspended and 25 μL were spread onto selection plates. Plates incubated overnight at 37℃.

M. smegmatis competent cells were made from cultures grown to an OD_{600nm} of between 0.8 and 1. Cells were placed on ice, then centrifuged for 10 minutes at 4℃ at 3900 rpm. Supernatant was removed and 5 mL of 10% glycerol was added and the pellet was resuspended. Once resuspended, 45 mL of 10% glycerol was added and cultures were centrifuged with the conditions stated above. Repeated process for a total of three glycerol washes. Supernatant was removed and pellet was resuspended in 100 μL 10% glycerol. 30 μL aliquots were snap frozen in liquid nitrogen and stored at -80℃. For transformation, competent cells were thawed on ice, and cells were distributed into tubes, one with 5 μL for a control and the remaining 25 μL for the transformation. 200 ng of plasmid was added to the cells, and the cells were electroporated using the Bio-Rad MicroPulser Electroporator. Cells were then rescued from the cuvette using 500 μL of 7H9, then incubated in 37℃ with a shaking speed of

200 rpm for 2.5 hours. 25 µL of cells were then spread onto selection 7H10 plates, then remaining cells were centrifuged for 30 seconds at 1300 rpm before discarding all but 50 µL of the supernatant. Cells were then resuspended and 25 µL were spread onto selection plates. Plates incubated 3 days at 37℃.

PCR Variations

For all cloning experiments, Q5 PCR was used to amplifying DNA (New England Biolabs). There are two major variations of Q5 PCR used in this project: regular and touchdown. Example of both can be seen in Table 2, 3, and 4. All Q5 PCR reactions follow the materials and thermocycler programs outlined in Tables 2, 3, and 4, unless otherwise stated. During regular Q5 PCR, thermocycler steps two through four are repeated for 35 total cycles. Annealing temperature (step 3) was determined by the Tm of primers 1 and 2. The time for elongation (step 4) is determined by adding 30 secs per kb in length of the template. The backbones for plasmids pSS229 and pSS317 were generated using Touchdown PCR as outlined in Table 3, while all inserts for these plasmids were generated using regular Q5 PCR as shown in Table 4. Oligo primers used in the PCR reactions associated with the works within this report are given in Table 6.

Table 2: Q5 PCR materials

Component	Amount (µL)
Q5 buffer	10
10 mM each dNTP mix	1
10 µM Primer 1	2.5
10 µM Primer 2	2.5
Q5 Polymerase	0.5
Q5 GC Enhancer	10
Template DNA	1
dH_20	22.5
Total	50

Table 3: Q5 Touchdown PCR Thermocycler Program

Temperature (℃)	Time
98	2 min
98	20 sec
65	30 sec
72	2 min 30 sec
Repeat Steps 2-4 for 10 cycles, lowering annealing temp by 1℃ every cycle	
98	20 sec
55	30 sec
72	2 min 30 sec
Repeat steps 5-7 for 25 cycles	
72	2 min
10	forever

Table 4: Q5 Regular Thermocycler Program

Temperature (℃)	Time
98	30 sec
98	20 sec
55	30 sec
72	2 min
Repeat steps 2-4 for 35 cycles	
72	5 min
10	forever

Table 5: Specific Annealing Temperatures

Plasmid	Annealing Temperatures (℃)
pSS229	Backbone = 72 - 55 Insert = 68 - 55

Table 6: Oligo Primers for PCR reactions

Oligo Name	Function	Sequence
SSS1149	First forward primer for DNA assembly containing sgRNA for PAM AGAAG in tail for repression of *mamA* using CRISPRi	GTGGTGATCGCGTCGGGGGTGTTTTTGTACTCGAAAGAAGC
SSS1150	First reverse primer for DNA assembly for repression of *mamA* using CRISPRi, binds to plasmid backbone	AGTGCTTCTGCTTCGGCATGGTGCGCGACAGGAAGCGGAA
SSS1151	Second forward primer for DNA assembly for repression of *mamA* using CRISPRi, binds to plasmid backbone	CATGCCGAAGCAGAAGCACT
SSS1152	Second reverse primer for DNA assembly containing sgRNA for PAM AGAAG in tail for repression of *mamA* using CRISPRi	ACCCCCGACGCGATCACCACTCCCAGATTATATCTATCACTGATA
SSS1253	Forward primer for amplifying promoter upstream of *M. smegmatis mamA* operon and cloning into HindIII cut pSS047	TAATTAACCATGGAGCGAGAACGGGCGTCATCACCGGATT
SSS1254	Reverse primer for amplifying promoter upstream of *M. smegmatis mamA* operon	CAGCAGATCGGCGTCGAAGCATGTTCATTGTCCCTGATCAG
SSS1255	Forward primer for amplifying *M. smegmatis mamA* with a 42 nt 5' UTR	TGATCAGGGACAATGAACATGCTTCGACGCCGATCTGCT
SSS1256	Reverse primer for amplifying *M. smegmatis mamA* for cloning into XbaI cut pSS047	CGATCCATATGACTAGTAGATCCTTGGCCTTGGTGGAGGGTTCA
SSS1257	Reverse primer for amplifying upstream part of *M. smegmatis mamA* making mutations in PAM to prevent CRISPRi.	GTGATCGCGTCGGGGGTGAAAAACGCACCCCGCGCCTTGCGA
SSS1258	Forward primer for amplifying downstream part of *M. smegmatis mamA* making mutations in PAM to prevent CRISPRi.	GATCGCAAGGCGCGGGGTGCGTTTTTCACCCCCGACGCGATCA

ATc Induction and Growth Curve Determination

Growth curves were determined in two different ways, by manual sampling and automated OD recording in Epoch 2 plate reader. MamA knockdown and overexpression strains were grown to an OD_{600nm} of between 0.5 and 0.8, then diluted to an OD of 0.01. To the diluted cultures ATc was added to a final concentration of 200 ng/mL for positive ATc cultures, while no-ATc controls received a similar volume of water. During the manual growth curve determinations, cultures were incubated at 37℃ with a shaking speed of 200 rpm. Every 2.5 hours the OD_{600nm} was sampled and recorded and later plotted to create a growth curve. During growth curve determinations using a plate reader, samples were diluted to an OD of 0.01 and 200 µL of each culture was placed in a 96 well plate, with each sample being plated in duplicate. All other wells were filled with 200 µL of 7H9. The samples were incubated at 37℃ and were shaking using a single orbital pattern at 807 cpm. The OD_{600nm} was measured every 10 minutes for 24 hours.

Fixing Cells with Paraformaldehyde

To prepare cells for microscopy, they were first fixed with paraformaldehyde. Cultures were grown for 12 hours either in the presence or absence of ATc, resulting in an OD_{600} of 0.05 to 0.15. Cultures were pelleted by centrifugation for 2 minutes at 1300 rpm and the supernatant was removed. Pellets were resuspended in 500 µL of 2% paraformaldehyde and incubated at room temperature for 30 minutes. Cells were then pelleted and resuspended in 900 µL of PBS + 0.1% Tween. This was repeated for a total of two washes. Pellets were finally resuspended in 50 µL of PBS-T.

Brightfield and Fluorescence Microscopy

Cultures were grown for 12 hours either in the presence or absence of ATc and fixed with paraformaldehyde. Cells were then stained with either Syto 24 or FM 4-64fx. For DNA visualization, cells were stained with 2 µM of Syto 24 for 10 minutes before washing and resuspending in 50 µL of TBS-T. For FM staining, cells were stained with 8 µL of 100 ug/mL FM

4-64fx for 10 minutes before washing twice with PBS-T. 2 µL of cells were then mixed with 6 µL of mounting media (20mM Tris pH 8, 0.5% N-propyl gallate, and 50% glycerol in milliQ water), and pipetted onto an agar pad on a slide. Once prepared, cells were viewed using the Zeiss AX10 microscope with ApoTome with a magnification of 400X. All images used the same display settings for consistency of brightness of fluorescence.

Viability Assay

Cultures of MamA knockdown and nonspecific CRISPRi cells were normalized and diluted to an OD of 0.01 as described above. ATc was added to +ATc cultures to a concentration of 200 ng/mL. Samples of the cultures were taken every three hours for a total of 24 hours. Samples were then serially diluted and spot plated in duplicate for dilutions between 10^{-1} and 10^{-8} at each timepoint. Spots were grown on antibiotic free plates and incubated for 2 days before counting colonies.

Data Analysis

Analysis of microscopy photos was performed using ImageJ (NIH) and GraphPad Prism 7.04 (GraphPad Software). Statistical analysis was performed using one-way ANOVA, Dunn's multiple comparison tests, and non-parametric tests.

RESULTS

Knockdown Strategy

In order to study the effects of MamA (msmeg_3213) depletion in *M. smegmatis* cells, we used the CRISPR interference (CRISPRi) knockdown strategy (Qi *et al*, 2013, Rock *et al*, 2017). This strategy uses a catalytically inactive form of CRISPR-Cas9 to block transcription of the target gene. The CRISPRi consists of two major components, the dCas9 protein and a single guide RNA that mediates binding with the bacterial DNA (Qi *et al*, 2013). In this work, we used a dCas9 optimized for mycobacteria described in Rock *et al* (2017). When choosing a sequence to act as the guide RNA, we used a 20 nt long sequence (GTGGTGATCGCGTCGGGGGT) about 100 nucleotides downstream from the start of the coding sequence of MamA. This sequence was associated with the PAM AGAAG, which is cited to bind dCas9 with a repression efficiency of 216.7-fold (Rock *et al*, 2017). The CRISPRi system was placed under a Tet repression system in which expression of dCas9 occurs when in the presence of ATc.

Depletion of MamA causes Growth Cessation and Reduced Viability

The first step to elucidating a possible phenotype of MamA depletion was to observe the growth of the MamA knockdown (KD) strains in the presence or absence of ATc. As a control, a strain of *M. smegmatis* containing a nonspecific CRISPRi construct was grown in parallel. This nonspecific CRISPRi contains a sgRNA sequence that does not target any specific part of the *M. smegmatis* genome. Biological replicate cultures of the MamA KD strains and the nonspecific CRISPRi strains were grown in culture to an OD between 0.5-0.8 and then diluted to an OD of 0.01. These cultures were then incubated either in the presence or absence of ATc for a total of 24 hours in a plate reader, with OD readings taken every eight minutes (Figure 1). For the first 10 hours of growth, both +ATc and -ATc cultures of each strain appeared to grow at a similar pace with similar OD measurements. After this point, the MamA KD +ATc cultures experienced a slowing of growth and eventually cultures stopped increasing in OD. This plateau of cell growth occurred at about an OD of 0.1. This cessation of growth in MamA depleted cells suggests that MamA might play an important role in cell growth pathways.

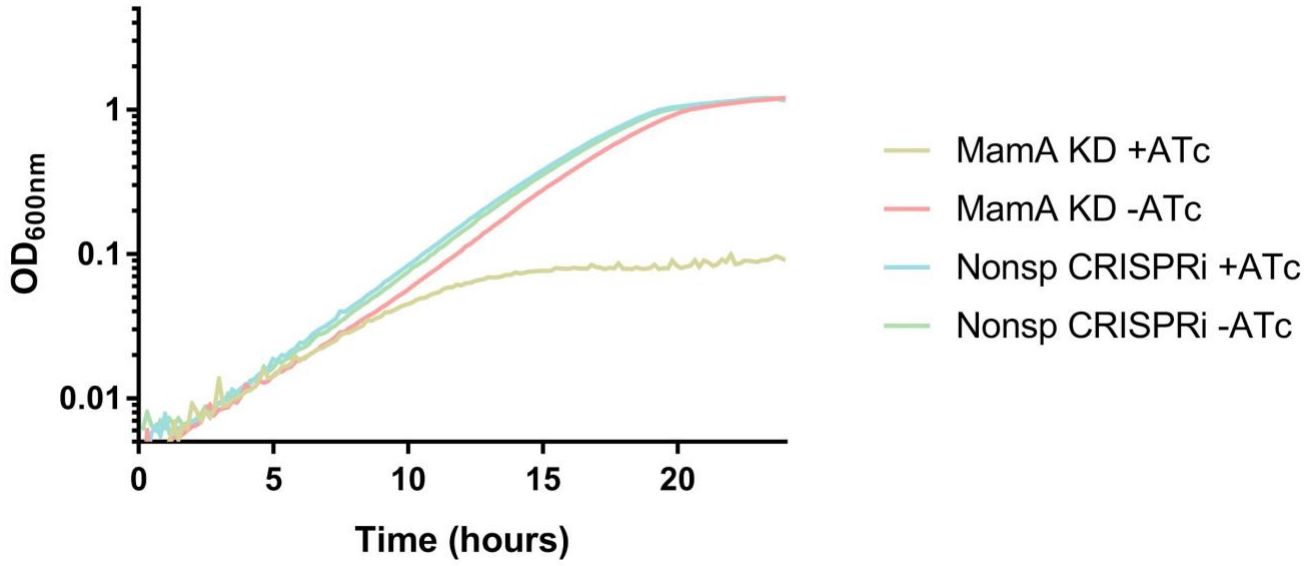

Figure 1: MamA is essential for growth in *M. smegmatis.*

Growth curve of *M. smegmatis* samples in the presence or absence of ATc. Three biological replicates of MamA KD, two biological replicates of nonspecific CRISPRi (Nonsp CRISPRi), and two technical replicates of each were used and values were averaged. These data are representative of two independent experiments.

The gene encoded downstream of *mamA* on the *M. smegmatis* chromosome, msmeg_3214, is located on the same strand, partially overlaps *mamA*, and lacks a separate transcription start site (Martini, Zhou, Sun & Shell, 2019). CRISPRi represses transcription of genes that are co-transcribed downstream of the target. In order to confirm that the phenotype caused by CRISPRi knockdown of MamA was a direct result of the depletion of MamA and not related to a depletion of msmeg_3214, we created a rescue vector that does not bind dCas9. This rescue vector contains a second copy of MamA, but the PAM of the target sequence was altered by a synonymous point mutation predicted to reduce binding with dCas9. This vector was then integrated into the genome of *M. smegmatis* that also contained the CRISPRi knockdown construct for MamA depletion. When grown in the presence of ATc, the cultures

with the rescue vector were able to grow more similarly to those without ATc and nonspecific CRISPRi controls, confirming that the growth cessation observed in the MamA KD strain is indeed attributable primarily to depletion of MamA (Fig 2).

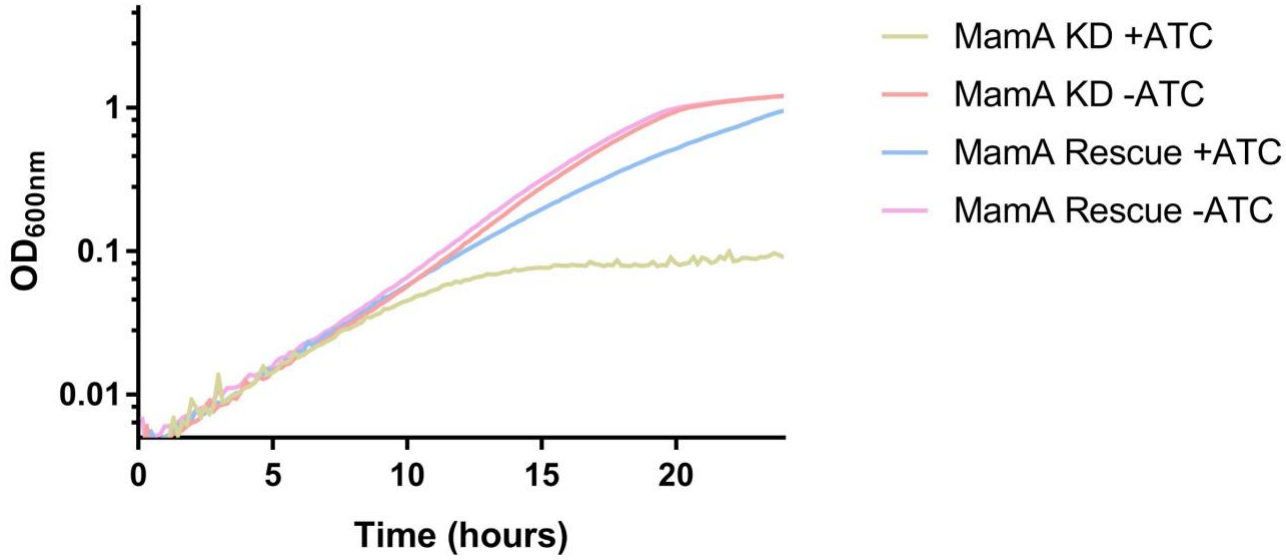

Figure 2: A second copy of *mamA* results in a rescue of growth in the presence of ATc.
Growth curve of *M. smegmatis* samples in the presence or absence of ATc. Three biological replicates of MamA KD, four biological replicates of MamA Rescue, and two technical replicates of each were used and values were averaged.

After establishing that MamA depleted cells cease growing within the first 12 hours of incubation, we tested to determine whether the cells within these cultures were viable. Cultures of MamA KD and nonspecific CRISPRi were grown and diluted as above. Initial samples were taken and serially diluted. Cultures were then incubated for a total of 24 hours in the presence or absence of ATc, and samples were taken every three hours beginning at hour six. Samples were serially diluted and plated on antibiotic-free 7H10 in 5 µL drops, performed in duplicate. The dilution with the highest countable number of colonies was used for determination of colony forming units (CFU's). Representative CFU's were plotted for each sample over time, shown in Figure 3. Data from the 15-hour time point have been excluded

from analysis due to technical error. While control cultures showed an approximately 1,000-fold increase in CFU over the time course, the MamA depleted cultures displayed a nearly 100-fold decrease in CFU from the initial introduction of ATc to the 24-hour time point. Reduction of viable CFU's occurred very early, beginning at hour 6. This suggests that the depletion of MamA in these cells resulted in cell killing.

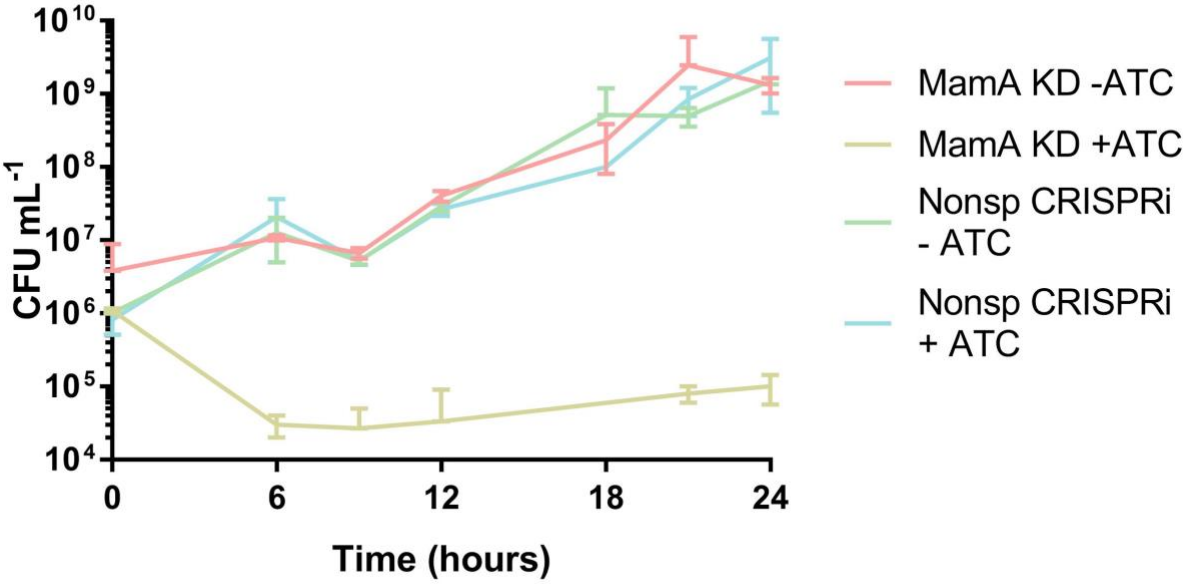

Figure 3: Depletion of MamA results in cell killing.

Calculated CFUs of *M. smegmatis* samples treated either with ATc or without ATc over a 24-hour period.

MamA Depletion Results in Cell Elongation and DNA Devoid Regions

With knowledge of MamA depletion causing growth cessation and reduction of cell viability, we investigated the impact of MamA knockdown on cell morphology. MamA KD, nonspecific CRISPRi, and wild type cells were grown as previously described and incubated in the presence or absence of ATc for 12 hours, then fixed with paraformaldehyde and stained with Syto 24, a green fluorescent DNA stain. When viewed by DIC microscopy, MamA depleted cells exhibited elongated cell bodies compared to those of the control cells. As shown in Figure 4 panels A and B, control cells showed a characteristic rod shape, whereas MamA KD cells were

much longer and filamentous (panel D). Cell lengths were measured and recorded in a blind analysis for each strain type and condition and shown in Figure 4 panel E. While all control strains and treatments had an average length approximately 5.7 µm, MamA KD cells had an average length of 12.8 µm. MamA KD cells also had a larger coefficient of variance, 31.6%, compared to that of control cells, 21.6%

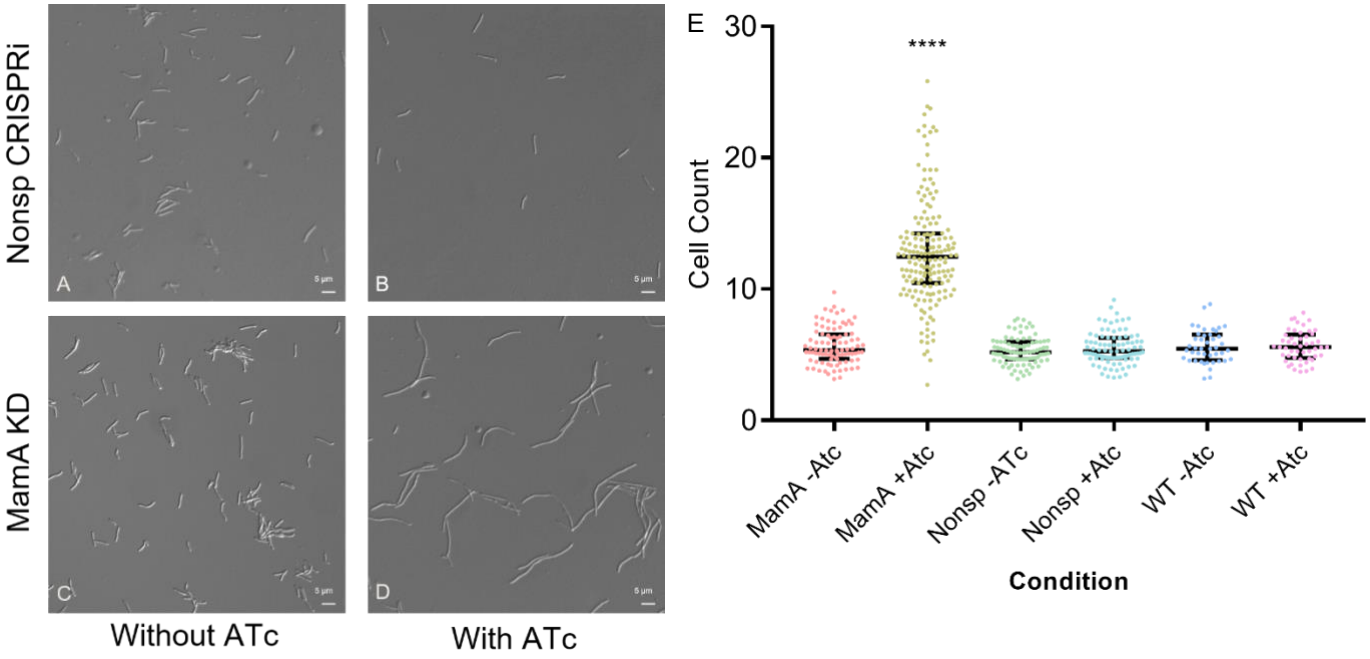

Figure 4: MamA depletion causes broader distribution of cell lengths and higher average cell length.

Microscopy of control nonspecific CRISPRi cells (panels A and B) and the MamA KD cells (panels C and D) samples without treatment and treated with ATc. Dot plot representing cell lengths of *M. smegmatis* samples either in the presence or absence of ATc (panel E). Samples were imaged by DIC microscopy after 12.5 hours of growth from starting OD 0.01. ****p<0.0001, Kruskal-Wallis followed by Dunn's multiple comparisons test.

DNA content and localization were assessed by fluorescence microscopy of Syto 24-stained cells. Representative images are shown in Figure 5 panels A to D. Control cells showed fairly even distribution of DNA throughout the cell length. MamA depleted cells, however, had a

much smaller distribution of DNA within their cells, with fluorescence often focused around the midpoint of the cell leaving large areas devoid of DNA. In order to quantify this difference in DNA distribution, we used the method of calculating DNA occupancy for cells as described in Mann *et al.* 2017. Cells lengths and lengths of fluorescent zones were measured in Image J in a blinded analysis. MamA depleted cells had significantly ($p<0.0001$) lower DNA occupancy than control cells, with averages of 49.1% occupancy, and 79.3% occupancy, respectively (Figure 5 panel E).

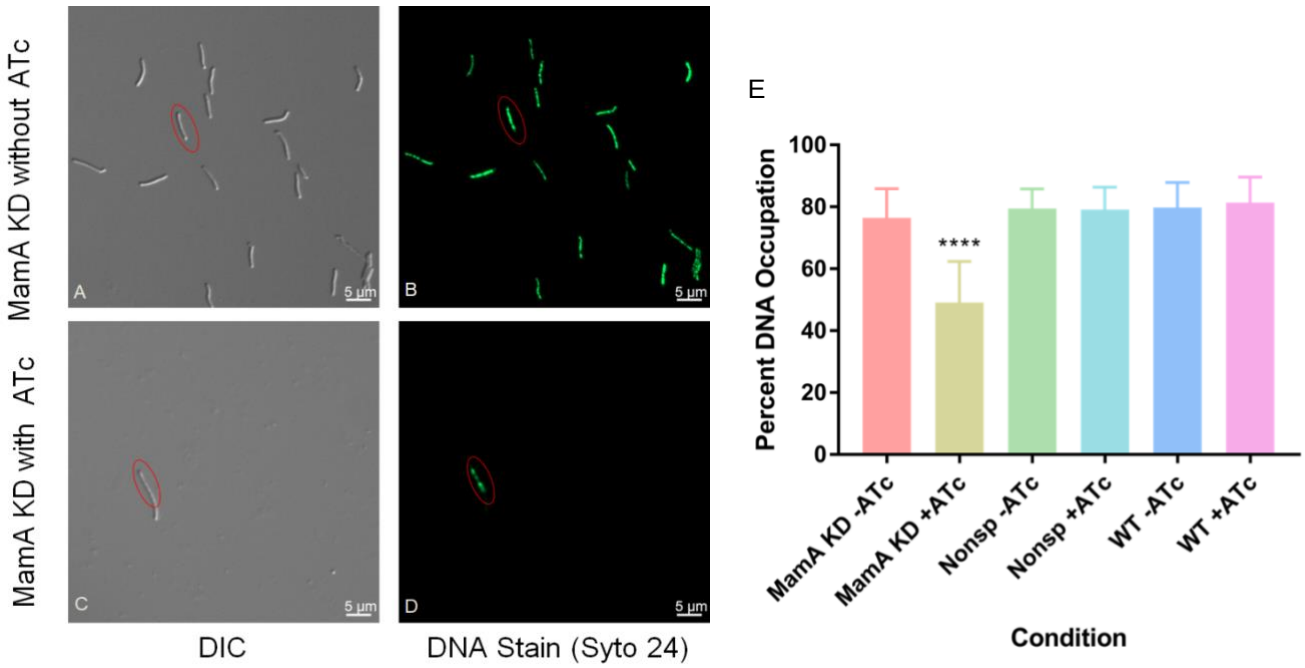

Figure 5: MamA depleted cells have reduced DNA occupation.

Microscopy of *M. smegmatis* samples treated with ATc (panel C and D) and without ATc (panel A and B). Panels B and D show cells stained with Syto 24, which fluoresces when bound to DNA. Red ovals highlight the same area in both brightfield and fluorescent images. Percent DNA occupation of *M. smegmatis* samples either treated or not treated with ATc (panel E). Samples were measured after 12.5 hours of growth. Percent DNA occupation was calculated by dividing length of DNA fluorescence by cell length. Controls consisted of nonspecific CRISPRi construct and wild type. ****$p<0.0001$, Kruskal-Wallis followed by Dunn's multiple comparisons test.

Membrane Staining of MamA Depleted Cells Reveal Three Major Phenotypes

To determine if elongated MamA KD cells had septa (chaining) or not (filamenting), MamA KD cells were grown in the presence of ATc for 12 hours and stained with the membrane stain FM 4-64fx, as well as the DNA stain Syto 24. When observing the membranes and DNA of MamA KD cells, we found three distinct categories of cells within the knockdown cells: long cells with septa, long cells without septa, and small, DNA devoid cells. Examples of each phenotype can be seen in Figure below.

The first category, long cells with septa, were grouped due to their long cell length and evidence of septation. The number of septa per cell varied within this category, ranging from one to four septa within a cell. Most of the chaining cells had DNA located within each segment, however many had segments at the poles that do not appear to contain DNA. The second category, long cells without an obvious septum were grouped by the appearance of cell filamentation. These cells tended to have DNA localized near the middle of the cell with large portions of the cells near the poles that were apparently devoid of DNA. The third category included small cells, which were similar in size or smaller than control cells, that had no evidence of DNA within the cell.

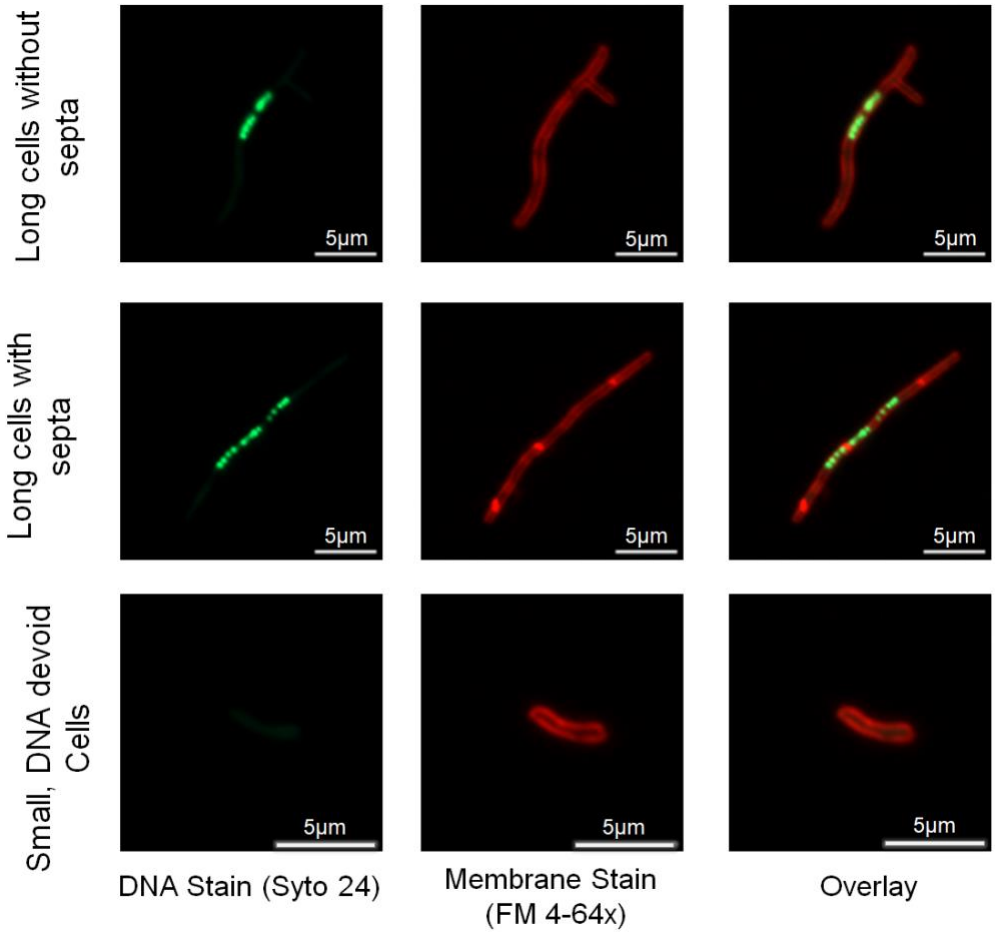

Figure 6: MamA KD cells have three distinct phenotypes.

Microscopy of MamA KD *M. smegmatis* samples treated with ATc for 12 hours. Cells were stained with Syto 24 and FM 4-64fx to visualize DNA and membranes. The images above represent cells of three major phenotypic groups observed. The left-most column shows DNA stained only, the center column shows membrane stained only, and the right-most column shows the merged fluorescent images.

DISCUSSION

MamA is the primary DNA methyltransferase in *M. smegmatis* and appears to play an important role in the normal growth and division of cells. With depletion of MamA, cells cease growing and appear to die after only a few doublings. Viability of MamA knockdown cells drops very quickly after induction with ATc; however, an important consideration of this experiment is the concentration of ATc that is plated with the various cell dilutions. Colony counts for the MamA KD with ATc were taken from the 100-fold dilution and it may be possible that a concentration of 2 ng/mL of ATc may still affect the cells once plated. Further experiments to determine the effects of decreasing levels of ATc on MamA KD cells will be performed to assess the validity of these results.

MamA depleted cells also experience elongation of the cell body after longer periods of time without MamA activity. This elongation suggests the cells cannot properly initiate cell division. There are various possibilities for the inability for cells to complete cell division, including but not limited to inability to initiate DNA replication, stalling and failure to complete DNA replication, and improper septa formation. In order to elucidate the reason for cell elongation, we visualized cellular DNA and cell membranes using DNA dyes. When stained with Syto 24, DNA within MamA depleted cells appeared to have different localization compared to that of control cells. The DNA of these cells is often focused in the center or in clusters in the cell, with large areas devoid of DNA. These areas of DNA devoid regions could suggest decreased amounts of DNA or specific localization and condensing of the DNA, but this distinction cannot be determined through DNA staining alone.

When stained with the membrane dye FM 4-64x, we found three major phenotypes regarding the state the MamA KD cells membranes. Long cells can either filament or chain, and a small population of cells are similar in size to control cells but seem to lack DNA. Filamentous cells tend to have areas devoid of DNA near their poles and chaining cells have at least one or more cells that contain DNA. Interestingly, some chaining cells will have cells within the chain lacking DNA as well. These cells tend to be on the ends, where they may separate from the other chaining cells to become the small DNA devoid cells we see. While it is unclear what causes the elongated cells to form septa or not, our data suggest there is a decoupling between

DNA replication and cell division in MamA depleted cells. Without functioning MamA, *M. smegmatis* cells appear to be growing and attempting to divide without proper replication or segregation of DNA.

One limitation of using stains to visualize DNA is the possibility for inconsistency in membrane permeability. It is possible for some cells, and even specific areas of a cell, to be more permeable to the DNA stain than others. This inconsistency could explain the lack of fluorescence in parts of filamentous and chaining cells, and the complete lack of fluorescence in small DNA devoid cells. In order to validate the presence of these DNA devoid regions, we aim to use fluorescently tagged proteins such as HupB-Venus and SSB-GFP to observe DNA content within cells. With a second method of visualizing the DNA content of MamA depleted cells, we can confirm whether the chaining and small cell phenotypes are actually lacking DNA. Future studies will be necessary to differentiate whether DNA replication is occurring within these cells, and how MamA KD cells are able begin the division process without having properly replicated DNA.

REFERENCES

Adhikari, S., & Curtis, P. D. (2016). DNA methyltransferases and epigenetic regulation in bacteria. *FEMS microbiology reviews*, *40*(5), 575-591.

Brilli, M., Fondi, M., Fani, R., Mengoni, A., Ferri, L., Bazzicalupo, M., & Biondi, E. G. (2010). The diversity and evolution of cell cycle regulation in alpha-proteobacteria: a comparative genomic analysis. *BMC systems biology*, *4*(1), 52.

Collier, J., McAdams, H. H., & Shapiro, L. (2007). A DNA methylation ratchet governs progression through a bacterial cell cycle. *Proceedings of the National Academy of Sciences*, *104*(43), 17111-17116.

Demarre, G., & Chattoraj, D. K. (2010). DNA adenine methylation is required to replicate both Vibrio cholerae chromosomes once per cell cycle. *PLoS genetics*, *6*(5), e1000939.

Engel, J. D., & von Hippel, P. H. (1978). Effects of methylation on the stability of nucleic acid conformations. Studies at the polymer level. *Journal of Biological Chemistry*, *253*(3), 927-934.

Felletti, M., Omnus, D. J., & Jonas, K. (2018). Regulation of the replication initiator DnaA in Caulobacter crescentus. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*.

Gonzalez, D., Kozdon, J. B., McAdams, H. H., Shapiro, L., & Collier, J. (2014). The functions of DNA methylation by CcrM in Caulobacter crescentus: a global approach. *Nucleic acids research*, *42*(6), 3720-3735.

Kahng, L. S., & Shapiro, L. (2001). The CcrM DNA Methyltransferase of *Agrobacterium tumefaciens* Is Essential, and Its Activity Is Cell Cycle Regulated. *Journal of Bacteriology*, *183*(10), 3065–3075. http://doi.org/10.1128/JB.183.10.3065-3075.2001

Kozdon, J. B., Melfi, M. D., Luong, K., Clark, T. A., Boitano, M., Wang, S., ... & Korlach, J. (2013). Global methylation state at base-pair resolution of the Caulobacter genome throughout the cell cycle. *Proceedings of the National Academy of Sciences*, *110*(48), E4658-E4667.

Mann K. M., Huang D.L., Hooppaw A.J., Logsdon M.M., Richardson K, et al. (2017) Rv0004 is a new essential member of the mycobacterial DNA replication machinery. *PLOS Genetics* 13(11): e1007115. https://doi.org/10.1371/journal.pgen.1007115

Marinus, M., & Casadesus, J. (n.d.). Roles of DNA adenine methylation in host–pathogen interactions: mismatch repair, transcriptional regulation, and more. *FEMS Microbiology Reviews*, 33(3), 488–503. doi:10.1111/j.1574-6976.2008.00159.x

Martini, M. C., Zhou, Y., Sun, H., & Shell, S. S. (2019). Defining the transcriptional and post-transcriptional landscapes of Mycobacterium smegmatis in aerobic growth and hypoxia. *Frontiers in Microbiology*, *10*, 591.

Morris, P., Marinelli, L. J., Jacobs-Sera, D., Hendrix, R. W., & Hatfull, G. F. (2008). Genomic characterization of mycobacteriophage Giles: evidence for phage acquisition of host DNA by illegitimate recombination. *Journal of bacteriology*, *190*(6), 2172-2182.

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, *152*(5), 1173-83.

Rock, J. M., Hopkins, F. F., Chavez, A., Diallo, M., Chase, M. R., Gerrick, E. R., Pritchard, J. R., Church, G. M., Rubin, E. J., Sassetti, C. M., Schnappinger, D., … Fortune, S. M. (2017). Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform. *Nature microbiology*, *2*, 16274. doi:10.1038/nmicrobiol.2016.274

Sanchez-Romero, M., Cota, I., & Casadesus, J. (2015). DNA methylation in bacteria: from the methyl group to the methylome. *Current Opinion In Microbiology*, 25, 9–16. https://doi.org/10.1016/j.mib.2015.03.004

Shell S.S., Prestwich E.G., Baek S.H., Shah R.R., Sassetti C.M., et al. (2013) DNA Methylation Impacts Gene Expression and Ensures Hypoxic Survival of *Mycobacterium tuberculosis*. *PLOS Pathogens* 9(7): e1003419. https://doi.org/10.1371/journal.ppat.1003419

Skerker, J., & Laub, M. (2004). Cell-cycle progression and the generation of asymmetry in Caulobacter crescentus. *Nature Reviews Microbiology*, 2(4), 325–337. doi:10.1038/nrmicro864

Slater, S., Wold, S., Lu, M., Boye, E., Skarstad, K., & Kleckner, N. (1995). E. coli SeqA protein binds oriC in two different methyl-modulated reactions appropriate to its roles in DNA replication initiation and origin sequestration. *Cell*, *82*(6), 927-936.

Thanky, N., Young, D., & Robertson, B. (2007). Unusual features of the cell cycle in mycobacteria: Polar-restricted growth and the snapping-model of cell division. *Tuberculosis*, 87(3), 231–236. doi:10.1016/j.tube.2006.10.004

World Health Organization. (2018). *Tuberculosis* [Fact sheet]. Retrieved from: http://www.who.int/news-room/fact-sheets/detail/tuberculosis

Zhu, L., Zhong, J., Jia, X., Liu, G., Kang, Y., Dong, M., … Chen, F. (2016). Precision methylome characterization of *Mycobacterium tuberculosis* complex (MTBC) using PacBio single-molecule real-time (SMRT) technology. *Nucleic Acids Research*, *44*(2), 730–743. http://doi.org/10.1093/nar/gkv1498