Regulation of Gene Expression by 5' UTRs in *Mycobacterium smegmatis*

Advisors:

Scarlet Shell, Ph.D. and Louis Roberts, Ph.D.

Written by:

Carlie Logan Alfaro and Katherine Nugai

A Major Qualifying Project

WORCESTER POLYTECHNIC INSTITUTE

Submitted to the Faculty of the Worcester Polytechnic Institute in partial fulfillment of the requirements for the Degree of Bachelor of Science in Biology and Biotechnology.

ABSTRACT

Tuberculosis, caused by *Mycobacterium tuberculosis*, is a highly infectious disease. TB is unique in that it can survive under extreme environmental stress in the human body, including hypoxia. A possible survival mechanism is the influence of 5' UTRs on gene regulation. To test this, a series of constructs were built to include a strong promoter and a fluorescent reporter gene (mCherry). One construct introduced a degradation tag, while another removed the 5' UTR associated with the native promoter. The constructs were assessed by fluorescence microscopy and colony morphology to evaluate mCherry expression. Findings revealed that the degradation tag was too efficient and the UTR was unnecessary. The final construct contains the MOP promoter and the mCherry reporter gene and can be used in future studies to test the impact of various 5' UTRs on gene expression.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

We would like to thank Professor Scarlet Shell and Professor Lou Roberts as well as Carla Martini, Diego Vargas, and Ying Zhou for their guidance and support. We would also like to recognize Jeremy Rock and Sarah Fortune for their annotations of the pJEB MOP mCherry promoter and UTR regions. Lastly, thank you to the WPI Biology and Biotechnology Department for expanding the coursework to include the BB356X course, without which this project would not have been possible.

INTRODUCTION

Tuberculosis (TB) is an infectious disease that is estimated to affect more than one-third of the world's population (CDC, 2017). *Mycobacterium tuberculosis* (*Mtb*), the causative agent of TB, is extremely contagious; when infected individuals cough, speak, or sing, the sputum in their lungs is disrupted and the infectious particles are transmitted through the air (WHO, 2017). Current treatment options for TB include regular check-ups and the use of specialized antibiotics, such as isoniazid and rifampin, over a period of six to nine months (CDC, 2017). Although antibiotics can be effective in the treatment of TB, this treatment method requires strict patient compliance with the daily regimen of antibiotics (CDC, 2017). If the patient were to miss doses or abandon the treatment altogether, the mycobacteria can become resistant to the specific mechanisms of action of these drugs, making future cases of TB harder to treat. Although TB resistance typically develops due to misuse of pharmaceutical therapies, *Mtb* does have innate capabilities that allow it to survive in the body during periods of environmental stress. Once in the human body, *Mtb* must overcome harsh conditions to survive. The human immune system uses different mechanisms of stress to reduce the likelihood of pathogenic survival. For example, macrophages act on the intracellular *Mtb* by decreasing pH and nutrient content, while the formation of granulomas in the lungs indirectly lead to decreased levels of oxygen, otherwise known as hypoxia.

Mtb has been shown to be more resistant to antibiotic exposure when under hypoxic stress (Rao et al., 2008). For example, when under hypoxic conditions rifampicin was 50-fold less effective at killing *Mtb* than when under normoxic conditions (Rao et al., 2008). The exact mechanism(s) by which *Mtb* evades the innate and adaptive immune systems as well as antibiotic

treatments are not well understood. However, it is believed that when *Mtb* experiences hypoxic conditions, it can downregulate its metabolic activity to maintain survival (Rustad et al., 2009). When *Mtb* responds to the lowered availability of oxygen and exists in the body at a decreased metabolic level, it is considered to be dormant or quiescent. Bacteria in this state of dormancy are less impacted by antibiotics, complicating both treatment and prevention of transmission because infected individuals can remain infectious. If dormancy could be prevented or inhibited, the survival capabilities of *Mtb* would be reduced, therefore making it easier to identify and treat TB. For example, in a recent paper published by Lamprecht et al., they explored the unique ability of *Mtb* to evade the host and antibiotic assaults by altering its electron transport chain (ETC), stabilizing its ATP levels enough to survive. This represents a potential mechanism that could be targeted when developing new treatment methods (Lamprecht et al., 2016).

Mtb is believed to evade natural and artificial stressors of infection by altering immune cell fate, enhancing granuloma formation, developing antibiotic resistance, and is believed to have several regulatory regions that optimize the stress response (Peddireddy et al., 2017). Gene regulation allows for phenotypic changes such as DNA repair, nutrient scavenging, growth cessation, metabolic reprogramming, and the expression of heat shock proteins; these help the mycobacterium persist (Galagan et al., 2013). For example, in an attempt to survive the harsh conditions of the lungs, *Mtb* will utilize its the lipid content of its cell wall as a primary nutrient source (Galagan et al., 2013). A total of 190 predicted gene expression regulators are encoded within the *Mtb* genome; most of these are thought to be used for transcriptional regulation (Manganelli et al., 2004).

While transcriptional regulation of gene expression may be the most commonly known regulation method, is not the only type of regulation that exists. Another important class of gene regulation is through 5' untranslated regions (UTR). These regions were originally thought to have minimal function, and it is only recently that 5' UTRs (particularly those of greater length) have been shown to have gene expression capabilities (Mignone et al., 2002). Although the complete functions of UTRs are not entirely understood, there are several mechanisms that have been identified through which 5' UTRs regulate transcription and translation. Previous analyses of the secondary structure of bacterial and mycobacterial 5' UTRs support the theory that these lengthy sequences can take on a hairpin-like structure that, under certain conditions, enhances or represses translation (Araujo et al., 2012).

One mechanism through which 5' UTRs regulate gene expression is *transcriptional attenuation,* which can be defined as any process that makes use of transcription termination or transcription pausing to moderate downstream gene expression (Gollnick & Babitzke, 2002). This mechanism can operate through several possible methods with the two most important being secondary and tertiary folding of the mRNA and protein binding to the mRNA. Folding regulates translation by deterring binding of the ribosomal preinitiation complex to the mRNA transcript through pause structures such as hairpin loops (Araujo, 2012). Protein binding on the 5' UTR allows for specific proteins bind to the mRNA, thus repressing the translation of the downstream gene (Araujo, 2012).

Another type of regulation that uses both secondary structure and protein binding mechanisms is a *riboswitch*, whose specific structure is controlled by the binding of a ligand. Typically, the riboswitch region has two parts, the first of which is a structure called an aptamer,

whose role is to bind a ligand (Edwards, 2010). The second part, called the expression platform, contains a long hairpin loop which is responsible for the actual gene regulation. When the appropriate ligand binds to the aptamer, it creates a conformational change within the expression platform, shortening the hairpin loop, thus regulating the downstream gene and allowing translation to proceed (Edwards, 2010).

Two regions within 5' UTRs have been identified that may modulate gene expression; uORFs (upstream open reading frame) and upstream AUGs. Upstream AUGs may act as decoys to the true ORF/AUG, allowing for the down-regulation or cease of translation, while the uORFs may reduce the initiation of translation through the production of cis-acting peptides which stall the ribosome at the end of the uORF (Araujo, 2012). One important example of these mechanisms can be found in *Escherichia coli*, where the uORF of *ptrB* mRNA impacted the expression of *ptrB* CDS due to the reliance on the 5'-terminal AUG, even though the 5'-terminal uORF was not effectively translated (Beck & Janssen, 2017).

Mycobacterium smegmatis (*Msmg*) can be considered a model organism when it comes to the study of tuberculosis (Akinola et al., 2013). Both organisms are from the Mycobacterium genus and share about 983 orthologous proteins (Akinola et al., 2013; Reyrat & Kahn, 2001). Key differences between *Msmg* and *Mtb* exist that make *Msmg* the superior organism for certain types of laboratory experimentation. One of the main reasons *Msmg* is used in place of *Mtb* is that it is much less virulent; *Msmg* typically lives in the soil and has no known pathogenic effects in humans, such that it is BioSafety Level 1 while *Mtb* is BioSafety Level 3 (Wallace, 1988; WHO, 2012; Singh & Reyrat, 2009). The other key reason for using *Msmg* is that it can grow much more quickly than *Mtb*, with a doubling time for *Msmg* being about three hours compared to *Mtb's* 24

hour or longer doubling time; thus, experimentation can be sped up dramatically (Wallace, 1988; Cox, 2004).

As mentioned previously, *Msmg* is an ideal model organism for *Mtb* because they share many homologous genes, including a homolog of the *Mtb* gene Rv3241c, known as MSMG_1878. In a paper written by Trauner et al. in 2012 they found that MSMG_1878 appears to play a role in adaptation of *Msmg* to hypoxic conditions. Although the specific functions of MSMG_1878 and *Mtb* Rv3241c are not fully understood, the genes do exhibit homology, both through nucleic acid sequence similarity and the identity of the translated protein. We therefore hypothesize that they have similar if not the same regulatory functions and capabilities when the organism is exposed to stress conditions.

MSMG_1878 is speculated to be associated with the production of a stress responding protein (Li et al., 2015). It is hypothesized that this protein acts similarly to the hibernation promoting factor (HPF) in *E. coli* to allow for transition into a stationary or dormant phase, by inhibiting translation through the induction of dimerization of the ribosome (Li et al., 2015). This mechanism down-regulates translation of proteins to reduce cellular energy needs, therefore allowing the bacteria to enter a dormant state and increase the overall vitality of *Mtb* in the harsh host environment. The 5' UTR of the Rv3241c gene is suggestive of regulation due to its long length (228 nt) and the presence of secondary structure that may allow for variation in ribosome binding and subsequent alterations in regulation of gene expression.

The primary goal of this project is to develop a universal plasmid construct with a strong promoter and fluorescent reporter gene into which MSMG_1878 or any other proposed 5' UTR could be added into to be used in *Mycobacterium smegmatis,* with the intent to evaluate the impact

of hypoxic stress on this organism. To carry this out we created several DNA constructs. Each contain a medium strength (moderate affinity for RNA polymerase) mycobacterium optimized promoter (MOP) and a red fluorescent protein (mCherry) gene as the reporter gene. One construct also contained a protein degradation tag (Promoters, 2002; Chabot et al, 2007). The aforementioned constructs were used to determine whether the 5' UTR of the MOP promoter and the degradation tag were necessary for the optimization of the final construct. We wanted to remove the 5' UTR normally associated with the MOP promoter (UTR_{MOP}) because it may alter the amount of expression observed (potentially up-regulatory effect or down-regulatory effect). We chose to add a degradation tag to combat the strength of the MOP promoter and normalize the amount of expression observed by increasing turnover of protein. Findings suggested that the degradation tag was too efficient at degrading protein and that the UTR_{MOP} was unnecessary, and a potential hindrance, for the function of the MOP promoter. The final construct, which we present as a platform for future testing of 5' UTRs of interest, contains only the MOP promoter segment and the mCherry reporter gene.

MATERIALS AND METHODS

CONSTRUCT DESIGN

A critical component of this project was the development of testable experimental constructs. To create these constructs, we started with a plasmid already in use in the Shell laboratory, pJEB MOP mCherry. This plasmid was modified from pJEB402 (Lee 1991) by Tracy Rosebrock and Sarah Fortune. It was constructed to include a mycobacterium optimized promoter, made up of a promoter sequence and native 5' UTR region and an mCherry (red) fluorescent reporter gene. The construct also included the genetic elements required for integration into the L5 site by a site-specific phage recombinase and a genetic sequence that confers kanamycin resistance. In order to test the native roles of the MOP promoter and MOP 5' UTR, we created two separate constructs, one lacking the promoter and one lacking the MOP 5' UTR. Lastly, we evaluated the addition of a degradation tag to the construct to reduce the protein half-life to a "medium" brightness to allow change to be detectable via fluorescent microscopy. The work plan that was followed to construct, validate and analyze these constructs as seen below in Figure 1. The specific materials and methods utilized will be described in order according to this diagram, starting with primer design.

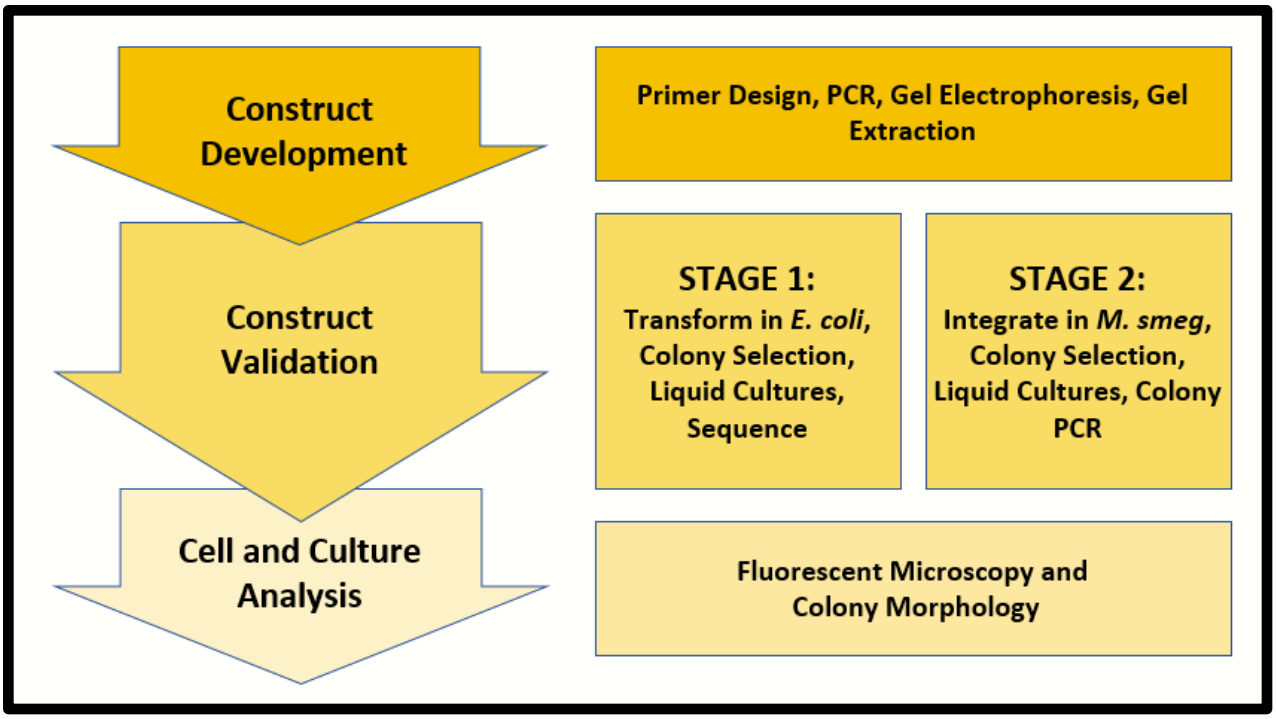

Figure 1. Project workflow

Construct Development

Primer Design

Each set of primers, was optimized to be no more than 24 nucleotides in length and to have the highest GC content possible within the desired attachment region. The majority of the primer sets were developed with the annealing tail on just one primer so that the other set could be "universal" and utilized for future use with the more permanent elements of the construct. The primer names, functions and anticipated length for amplification/deletion can be seen below in Table 1.

Table 1. Primer names, functions and amplified/deleted lengths

Primer Pair	Shell Lab Name	Function	Length Amplified or Deleted (bp)
N/A	SSS103, SSS142	Confirm Integration at the Left Junction	~400
N/A	SSS104, SSS144	Confirm integration at the Right Junction	~500
delUTR forward 15 nt overlap, CA_KN_delUTR reverse	SSS1418, SSS1419	Delete native MOP 5' UTR	43
delMOP forward 15 nt overlap, CA_KN_delMOP reverse	SSS1420, SSS1421	Delete MOP promoter	55
MOPmCherry+degtag forward, MOPmCherry+degtag reverse	SSS1422, SSS1423	Add GFP-LAA degradation tag	33

PCR Amplification and Dpn1 Digestion

PCR was done in 50 µL reactions with 25 µL of Q5 High-Fidelity 2X Master Mix (New England Biolabs) and 2.5 µL of the forward and reverse primers were used at 10 µM. 50 ng of template DNA (pJEB MOP mCherry) was added as well as nuclease free water to bring the reaction up to volume.

Following the completion of PCR, the samples were cleaned up using a Dpn1 digestion at a total volume of 60 uL. 6uL of 10X CutSmart Buffer and 4 µL of Dpn1 are added to the 50 µL

PCR reaction in a PCR tube and are incubated in a thermocycler for 30 minutes at 37 ℃ and then 20 minutes at 80 ℃.

ΔMOP

This PCR was performed in accordance with the Shell Lab methodology for Touchdown PCR and are seen in Figure 2 below. The temperature range for Touchdown is defined by five degrees higher and lower than the ideal Tm's for the primers in use (which already displayed a five-degree spread in this case) and the anneal temperature should decrease in increments of 1-2 degrees with each subsequent cycle. Annealing temperatures were determined by the NEB Tm calculator (https://tmcalculator.neb.com/#!/main).

Thermocycling Conditions:

Table12

	A	B	C
1	**STEP**	**TEMP**	**TIME**
2	Initial Denaturation	98°C	30 seconds
3	15 Cycles	98°C 70-55 (dropping 1 degree each cycle) 72°C	10 seconds 30 seconds 3 min.
4	25 cycles	98°C 50°C 72°C	10 seconds 30 seconds 3 min.
5	Final Extension	72°C	2 minutes
6	Hold	10°C	

Figure 2. Thermocycling conditions for creating ΔMOP

ΔUTR AND +DEGTAG

These PCR's were performed according to the PCR cycling protocol seen below. Annealing temperatures were determined by the NEB Tm calculator (https://tmcalculator.neb.com/#!/main). The main variations between these PCR reactions were the primers being used to remove or add different portions of sequence. The primers used for these constructs can be seen in Table 1.

Thermocycling Conditions:

Table6

	A	B	C
1	**STEP**	**TEMP**	**TIME**
2	Initial Denaturation	98°C	30 seconds
3	25–35 Cycles	98°C 51°C 72°C	10 seconds 30 seconds 3 min.
4	Final Extension	72°C	2 minutes
5	Hold	10°C	

Figure 3. Thermocycling conditions for creating ΔUTR and +degtag

GEL ELECTROPHORESIS FOR PURIFICATION

Gel electrophoresis was performed to clean up the PCR and broadly evaluate the efficacy of the PCR according to approximate band size. Gels were created using agarose (Fisher Bioreagents), 0.6% for smaller bands and 1% for larger bands, and 1X TAE made from stock 10X

TAE. SYBR Safe nucleic acid gel stain (Thermo Fischer Scientific) was added to the gel in order to visualize the DNA. When loading the samples, the DNA was mixed with purple gel loading dye (New England Bio Labs) and no more than 10 µL of this solution was added to the well. The gel electrophoresis was run for 30 minutes at 100 volts. The gel was imaged using Image Lab Software, and the DNA fragment sizes were determined using either a 1 kb, 2log, or 100 bp DNA ladder (New England Bio Labs) depending on the fragment sizes expected.

PCR Cleanup and Gel Extraction Clean-up

To recover high quality DNA from a gel or following a PCR the PCR Cleanup and Gel Extraction Clean-up kit from Macherey-Nagel was used. The kit was used according to the manufacturer's instructions, except for in the elution step, where the incubation and elution was performed with DI water (30 µL) instead of elution buffer. The concentration of the DNA was then evaluated using a Nanodrop machine.

Transformation in *E. coli*

Once the constructs were assembled, we transformed them into *E. coli* to quickly and effectively replicate the constructs. The Addgene protocol was followed as written (https://www.addgene.org/protocols/bacterial-transformation/). 1 µL of the aqueous plasmid and 20 µL of competent cells were added together. This solution was incubated on ice for 20 minutes, then heat shocked at 42°C for 45 seconds in a water bath. The tube was then placed back on ice for 2 min, then 250 µL of LB/SOC media was added. The transformed cells were incubated at 37°C for 60 min with agitation. Following this recovery period 50 µL and 200 µL of the solution were pipetted onto kanamycin selective LB plates and spread with glass beads. These plates were incubated overnight at 37 °C.

E. COLI COLONY SELECTION

The *E. coli* colonies that grew following transformation were picked using a sterile pipette tip and swirled in pre-warmed kanamycin selective (50 µg/mL) LB liquid media. The *E. coli* liquid cultures were left to incubate for 24-48 hours at 37 °C, and 200 rpm.

PLASMID DNA PURIFICATION

To recover the DNA from *E. coli* LB culture the Nucleospin Plasmid/Plasmid (No Lid) protocol was used in the Plasmid (No lid) kit from Macherey-Nagel was used. The kit was used according to the manufacturer's instructions, except for in the elution step in which we performed the incubation and elution with DI water (30 µL) instead of elution buffer. The concentration of the DNA was then evaluated using a Nanodrop machine.

SEQUENCING FOR CONSTRUCT CONFIRMATION

Following the DNA purification stage, the DNA samples along with appropriate primers (see Table 1) were sent to Eton Bio Labs to validate and evaluate the sequence of the constructs. These files were uploaded into Benchling and compared to the theoretical sequences of the modified pJEB MOP mCherry plasmid and then were compared and analyzed to identify "clean" sequences that lacked single nucleotide polymorphisms (SNPs) and presented with the predicted addition or deletion.

INTEGRATION IN *M. SMEGMATIS*

The plasmids with verified sequences for the ΔMOP, ΔUTR and +degtag constructs were then integrated into *M. smegmatis* to evaluate the impact of the addition or deletion of these units. To perform this transformation ~100 ng of plasmid DNA was added 10 µL of competent *M. smegmatis* cells and 20 µL of DI water into a 0.2 Electroporation cuvette. The Eppendorf

electroporator was set to 2500 V and 25 μF and the sample was electroporated. The solution was then transferred to a 1.5 mL microcentrifuge tube containing 200 μL of 7H9 media (no antibiotics) and incubated at 37 ℃ and 250 rpm for 2-4 hours. Following this recovery period, 20 μL and 200 μL are plated onto kanamycin selective 7H10 plates and allowed to incubate for 3-5 days at 37 ℃. μL

M. smegmatis Colony Selection

The *M. smegmatis* colonies that grew following transformation were picked using a sterile pipette tip and swirled in 5 mL of pre-warmed kanamycin selective 7H9 media (25 μg/mL) . The liquid cultures were left to incubate for 24-48 hours at 37 ℃, and 200 rpm. These liquid cultures were developed to confirm that the colonies did in fact have kanamycin resistance, therefore suggesting that the plasmids were properly integrated. These liquid cultures were used for both fluorescent microscopy and PCR.

PCR and Gel Electrophoresis for Integration Confirmation

This PCR was used to confirm integration of the constructs at the left and right junctions. The PCR cycling parameters were followed as seen below in Figure 4.

Initial Denaturation- 98 °C- 30 sec

30 cycles- 98, 52, 72= 10, 30, 2.5 min

Final extension- 72 °C- 2 min

Hold- 10 °C

Figure 4. Thermocycling conditions for integration confirmation

The samples were run on a 0.5% gel at 100 volts for 30 min. To confirm integration, the bands were anticipated to be approximately 400 and 500 bps.

Agar Pads for Microscopy

The agar pad protocol was obtained from the laboratory of Christopher Sassetti (University of Massachusetts Medical School). First a glass plate was wrapped with tape; the tape was one layer thick on the top surface of the plate (it overlapped on the bottom). The top side of the glass plate was coated with RainX standard hydrophobic formulation and rubbed with a kimwipe to remove the excess, then allowed to dry. 0.5 grams of agar was combined with 50 mL of water and melted in the microwave for 1 to 2 minutes until dissolved. The melted agar was allowed to cool slightly, then a bead of agar (~75 µL) was dropped onto the RainX coated glass plate. A slide was immediately laid flat on top of the agar bead, flattening it into a pad. The pad was allowed to solidify for 1 to 2 minutes. Finally, the slide was lifted gently off of the glass plate, keeping the pad intact and stuck to the slide. The rest of the agar can be stored at room temperature and re-melted for later use.

Fluorescence Microscopy

Fluorescence microscopy was used as a qualitative measurement of change in fluorescent protein produced by the *M. smegmatis* containing the modified constructs. A Zeiss Fluorescence Microscope with ApoTome attachment (Axio Imager Z1 stand) was used first under white light to identify *M. smegmatis* cells and then under green filtered fluorescent light to capture the amount of fluorescence for that sample. All strains were imaged using the same exposure settings. All viewing was done under the 40x objective with oil on the slide. These images were run through

ZEN 2.1 (blue edition) fluorescent reading computer software, in order to quantify and compare the relative fluorescent units (RFUs) among the different samples.

Results and Discussion

Reporter Construct Design and Construction

Our main intention for this project was to create a testable construct to evaluate the impact of 5' UTRs in *M. smegmatis* under stress conditions. We originally intended to build off the previous work conducted in the Shell lab using a construct that included a verified promoter (rpsA with tetO) and fluorescent reporter (YFP) that integrates at the Giles site. Preliminary results using this construct suggested that integration at the Giles site was not successfully achieved and therefore our team decided to move ahead in developing and verifying a novel construct. The new constructs that were designed and later assembled according to NEB HiFi Assembly protocols were built similarly to include a moderately strong promoter (MOP) and fluorescent reporter gene (mCherry). The specific composition of the constructs we set out to make are seen in Figure 5 below.

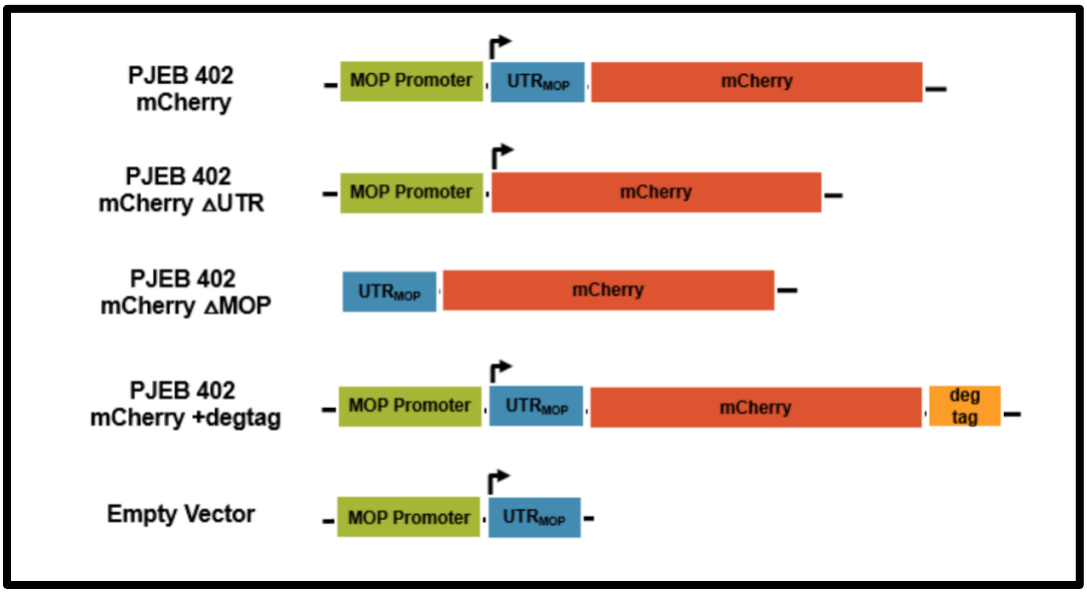

Figure 5. Depiction of the main pJEB MOP mCherry construct and the created ΔMOP, ΔUTR, and +degtag constructs.

The plasmid constructs include several important components; the MOP Promoter (MOP), the MOP associated UTR (UTR_{MOP}), mCherry, and a degradation tag (degtag). MOP has been annotated to consist of 98 nucleotides and consist of two discrete parts: the MOP promoter and MOP associated 5'UTR (Jeremy Rock and Sarah Fortune). The MOP promoter is 43 nucleotides long and the MOP associated 5'UTR is 55 nucleotides long. The fluorescent reporter, mCherry, was chosen for use in this construct as it is highly photostable with a bleaching rate of 96 photons per second and therefore would be adept for evaluation with fluorescent microscopy. The degradation tag consists of 33 nucleotides (11 amino acids, ending in LAA) and was reported to cause faster turnover of a reporter protein in *M. smegmatis* (Blokpoel 2003).

Three constructs were created from the complete pJEB MOP mCherry plasmid (pSS261) using PCR amplification and HiFi assembly (NEB). One construct has the MOP promoter removed (∆MOP), one with the UTR associated with the MOP promoter removed (∆UTR), and one with the addition of a LAA degradation tag (+degtag). Diagrams of the constructs can be seen in Figure 5. The ∆MOP construct was created in order to determine if an upstream or downstream spurious promoter was present within the construct. It was expected that the removal of the MOP promoter would result in no expression of mCherry at either the mRNA or protein level.

The ∆UTR construct was created in order to determine if the UTR_{MOP}, as annotated by Jeremy Rock in Sarah Fortune's lab, was necessary for, or impacted the expression from, the MOP promoter. It was expected that the removal of the UTR_{MOP} would have no effect on mCherry production thus would not be necessary for the final construct. The +degtag construct was created with the goal of reducing mCherry half-life in order to allow detection of expression decreases in non-growing cells where the mCherry protein is not diluted. It was expected that the degradation

tag would increase protein degradation, leading to lower steady-state mCherry expression levels. Correct insertion of the degradation tag and deletion of the MOP promoter and UTR_{MOP} was confirmed through sequencing performed at Eton Bio Science in Worcester, MA (as seen in Figures 6, 7, and 8 below below). The sequence alignments for the all of the aforementioned constructs we reviewed and any constructs with point mutations were discarded. The total length of the +degtag plasmid was 5900 bp, the total length of the ΔMOP plasmid was 5855 bp and the total length of the ΔUTR plasmid was 5867 bp. The length of the full pJEB MOP mCherry plasmid was 5910 bp and the length of the empty vector plasmid was 5199 bp.

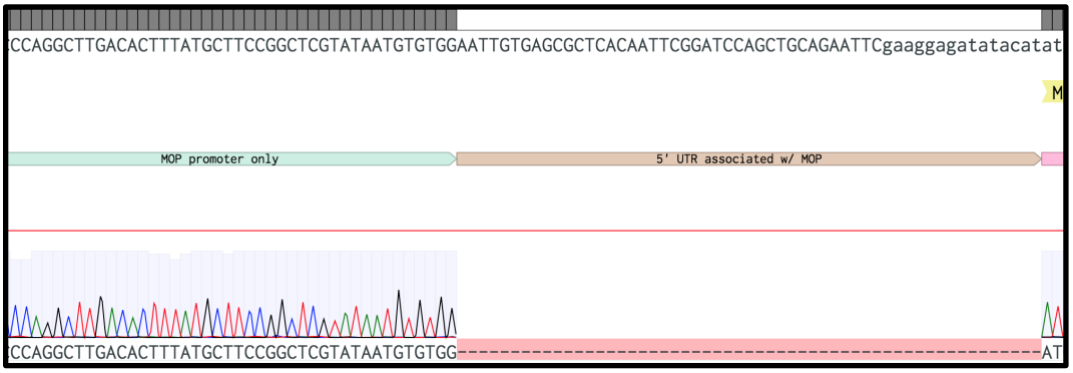

Figure 6. Sequence alignment of the removal of the UTR_{MOP}

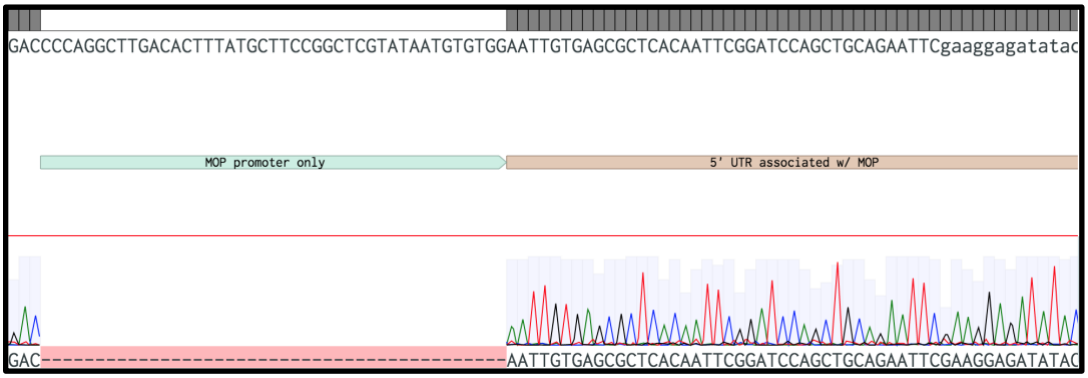

Figure 7. Sequence alignment of the removal of the MOP promoter

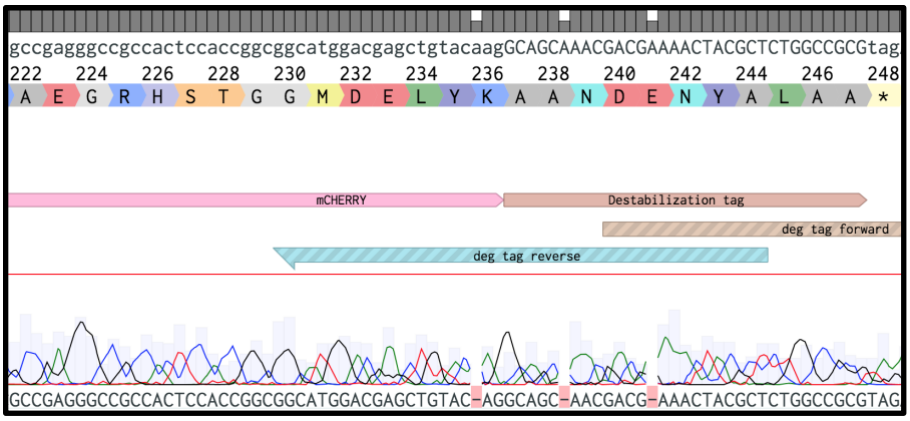

Figure 8. Sequence alignment of the addition of the degradation tag

The sequence-verified plasmids were transformed into *M. smegmatis* by electroporation. Next, integration into the *M. smegmatis* chromosome was confirmed by performing PCR at the junctions of integration. The same thermocycling conditions were used for the +degtag, and ΔMOP, and ΔUTR constructs, with an initial denaturation cycle of 98°C for 30 seconds, then 25-35 cycles of 98°C for 10 seconds, 54°C for 30 seconds and 72°C for 3 minutes, followed by a final extension of 72°C for 2 minutes and a hold at 10°C. These methods yielded successful PCR products for all constructs. As the expected products were about 400 bp and 500 bp, it was shown that there was appropriate integration for nearly all of the samples that were evaluated, seen below in Figure 9.

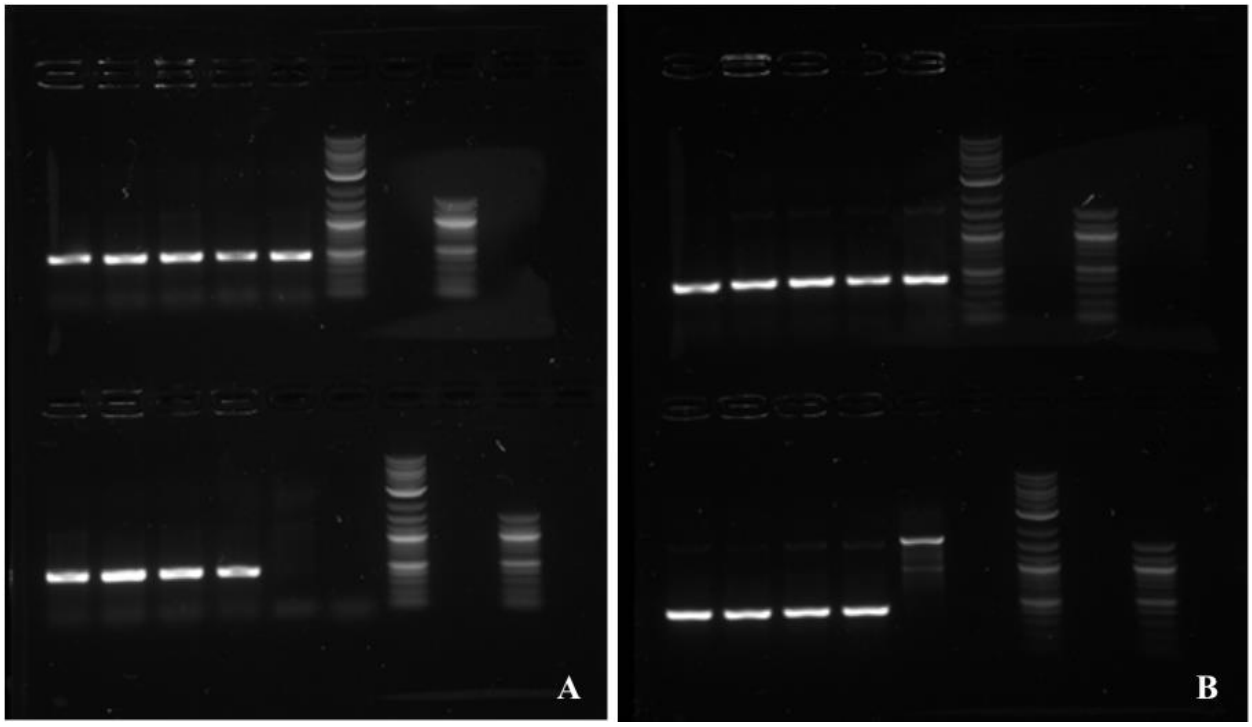

Figure 9. Gels Showing Integration of Constructs into *M. smegmatis*. **[A]** SSS1103/SSS142 primers. TOP LEFT-RIGHT: pJEB 402 (empty vector), pJEB MOP mCherry, pJEB MOP mCherry, ΔMOP, ΔMOP, ladder (2 log), empty lane, ladder (100 bp); BOTTOM LEFT-RIGHT: +degtag, +degtag, ΔUTR, ΔUTR, wild type, water, ladder (2 log), empty lane, ladder (100 bp). **[B]** SSS1104/SSS144 primers. TOP LEFT-RIGHT: pJEB 402 (empty vector), pJEB MOP mCherry, pJEB MOP mCherry, ΔMOP, ΔMOP, ladder (2 log), empty lane, ladder (100 bp); BOTTOM LEFT-RIGHT: +degtag, +degtag, ΔUTR, ΔUTR, wild type, water, ladder (2 log), empty lane, ladder (100 bp).

CONSTRUCT EVALUATION

In order to test the roles of the MOP promoter, MOP associated 5'UTR, and the addition of a degradation tag, fluorescence microscopy was used to determine the relative fluorescence of each construct. The +degtag construct (Figure 10K) did not have above-background fluorescence, suggesting that it caused a significant decrease in the half-life of mCherry. This particular degradation tag had been optimized for use with GFP (Blokpoel 2003), but our experimentation represents the first use of this degradation tag in conjunction with mCherry. Since it appeared to

reduce mCherry levels below the limit of detection, we concluded that this particular degradation tag was not useful in the context of our reporter construct.

The ΔMOP construct (Figure 10H) showed no fluorescence as expected. Due to the removal of the promoter, mCherry was not expressed. The ΔUTR construct (Figure 10E) showed brighter fluorescence than the full pJEB MOP mCherry plasmid. We therefore concluded that UTR_{MOP} was not required for robust transcription from the MOP promoter and robust translation of the resulting transcript. Importantly, this means that the MOP-associated UTR can be replaced with other UTRs of interest in future work to test the impacts of those UTRs on gene expression.

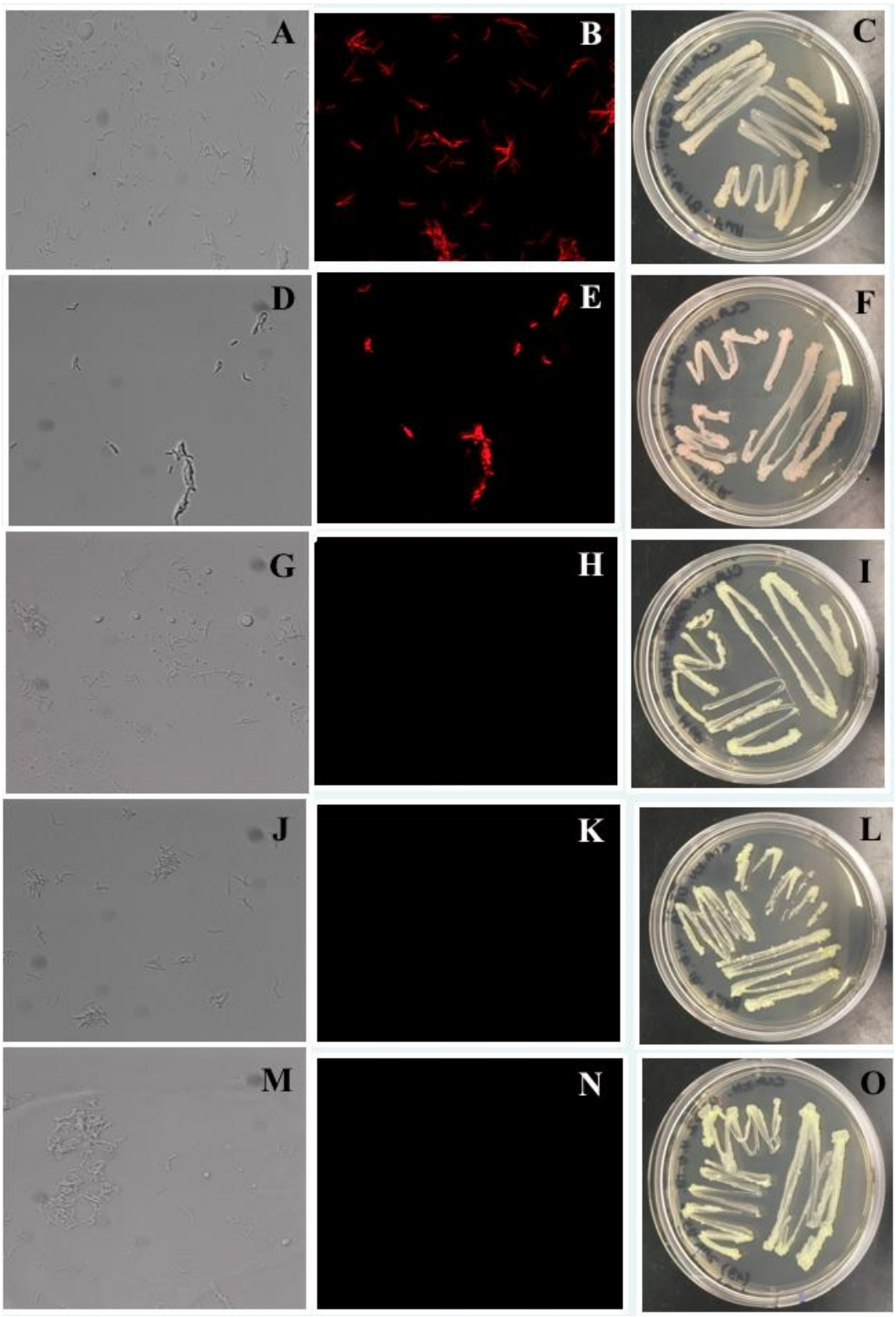

Figure 10. DIC images, fluorescent microscopy, and plate morphology of PJEB 402 mCherry **[A, B, C]**, PJEB 402 mCherry ΔUTR **[D, E, F]**, PJEB 402 mCherry ΔMOP **[G, H, I]**, PJEB 402 mCherry +degtag **[J, K, L]**, Empty Vector **[M, N, O]**

Conclusions, Future Research, and Applications

Our findings via fluorescent microscopy support the conclusion that of the constructs we developed, a construct that includes the MOP promoter and lacks the MOP 5'UTR, as annotated by the Shell lab, yields the greatest relative fluorescence. Future research for this project could include the evaluation of these samples via quantitative PCR to further investigate the cause of this increase and evaluate its correlation with transcript stability.

Although our course of testing was not able to include stress tests, it would be valuable to compare samples under hypoxic and normoxic conditions. Additional areas of future research include the addition of 5'UTRs of interest, such as the MSMG_1878 5'UTR. Once this construct is developed it could be used not only for hypoxic stress testing but also with other stress conditions present in lung granulomas, for example low pH and decreased nutrients. These future explorations should include both fluorescence microscopy and quantitative PCR to evaluate the protein and transcript stability that yields from these constructs under their respective conditions. Further experimentation could also include validation of a different degradation tag that is better suited for use with mCherry.

Applications of this research are directly linked between *M. smegmatis*, a nonpathogenic model organism, and *M. tuberculosis*, its highly pathogenic relative. By evaluating the impact of 5' UTRs associated with the stress response of both organisms there is the potential to gain great insight into how these organisms survive environments of intense stress. By better understanding these mechanisms of survival, scientists could potentially develop methods to inhibit or circumvent these capabilities and reduce the longevity and pathogenicity of the organisms.

REFERENCES

Akinola, R. O., Mazandu, G. K., & Mulder, N. J. (2013). A Systems Level Comparison of *Mycobacterium tuberculosis, Mycobacterium leprae* and *Mycobacterium smegmatis* Based on Functional Interaction Network Analysis. *Bacteriology & Parasitology*. doi:10.4172/2155-9597.1000173

Araujo, P. R., Yoon, K., Ko, D., Smith, A. D., Qiao, M., Suresh, U., . . . Penalva, L. O. F. (2012). Before it gets started: Regulating translation at the 5′ UTR. *Comparative and Functional Genomics, 2012*, 1-8. doi:10.1155/2012/475731

Beck, H. J., & Janssen, G. R. (2017). Novel Translation Initiation Regulation Mechanism in Escherichia coli ptrB Mediated by a 5′-Terminal AUG. Journal of Bacteriology, 199(14). doi:10.1128/jb.00091-17

Blokpoel, M. C. J., O'Toole, R., Smeulders, M. J., & Williams, H. D. (2003). Development and application of unstable GFP variants to kinetic studies of mycobacterial gene expression. Journal of Microbiological Methods, 54(2), 203-211. doi:10.1016/S0167-7012(03)00044-7

Chabot, J. R., Pedraza, J. M., Luitel, P., & Oudenaarden, A. V. (2007). Stochastic gene expression out-of-steady-state in the cyanobacterial circadian clock: Supplementary Information. *Nature, 450*(7173), 1249-1252. doi:10.1038/nature06395

Cox, R. A. (2004). Quantitative relationships for specific growth rates and macromolecular compositions of Mycobacterium tuberculosis, Streptomyces coelicolor A3(2) and

Escherichia coli B/r: an integrative theoretical approach. *Microbiology, 150*(5), 1413-1426. doi:10.1099/mic.0.26560-0

Galagan, J. E., Minch, K., Peterson, M., Lyubetskaya, A., Azizi, E., Sweet, L., . . . Schoolnik, G. K. (2013, July 11). The Mycobacterium tuberculosis regulatory network and hypoxia. Retrieved December 05, 2017, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4087036/

Global Tuberculosis Report 2017 (pp. 1-262, Rep.). (2017). Geneva, Switzerland: WHO. Retrieved December 5, 2017, from http://www.searo.who.int/tb/documents/annual_tb_repot_2017/en/

Gollnick, P., & Babitzke, P. (2002). Transcription Attenuation. *Biochimica Et Biophysica Acta (BBA) - Gene Structure and Expression,1577*(2), 240-250. doi:https://doi.org/10.1016/S0167-4781(02)00455-4

Lee, M. H., Pascopella, L., Jacobs, W. R., & Hatfull, G. F. (1991). Site-specific integration of mycobacteriophage L5: integration-proficient vectors for Mycobacterium smegmatis, Mycobacterium tuberculosis, and bacille Calmette-Guérin. *Proceedings of the National Academy of Sciences*, *88*(8), 3111–3115.

Lew JM, Kapopoulou A, Jones LM, Cole ST. TubercuList--10 years after. Tuberculosis. 2011;91:1–7. doi: 10.1016/j.tube.2010.09.008.

Li, X., Sun, Q., Jiang, C., Yang, K., Hung, L., Zhang, J., & Sacchettini, J. (2015). Structure of ribosomal silencing factor bound to mycobacterium tuberculosis ribosome. 23(10), 1858-1865. doi:10.1016/j.str.2015.07.014

Long, Q., Zhou, Q., Ji, L., Wu, J., Wang, W., & Xie, J. (2012). Mycobacterium smegmatis genomic characteristics associated with its saprophyte lifestyle. *Journal of Cellular Biochemistry, 113*(10), 3051-3055. doi:10.1002/jcb.24199

Manganelli, R., Proveddi, R., Rodrigue, S., Beaucher, J., & Gaudreau, A. L. (2004, February 01). σ Factors and Global Gene Regulation in Mycobacterium tuberculosis. Retrieved October 03, 2017, from http://jb.asm.org/content/186/4/895.full

Mignone, F., Gissi, C., Liuni, S., & Pesole, G. (2002). Untranslated regions of mRNAs. Retrieved October 04, 2017, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC139023/

Peddireddy, V., Doddam, S. N., & Ahmed, N. (2017). Mycobacterial dormancy systems and host responses in tuberculosis. *Frontiers in Immunology, 8* doi:10.3389/fimmu.2017.00084

Promoters. University of Texas at El Paso. (2002). Retrieved December 04, 2017, from http://utminers.utep.edu/rwebb/html/promoters.html

Rao S. P., Alonso S., Rand L., Dick T., & Pethe K. (2008). The protonmotive force is required for maintaining ATP homeostasis and viability of hypoxic, nonreplicating *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. U.S.A. 105, 11945–11950 10.1073/pnas.0711697105

Reyrat, J. M., & Kahn, D. (2001). Mycobacterium smegmatis: An absurd model for tuberculosis? Trends in Microbiology, 9(10), 472-474.

Rustad, T. R., Sherrid, A. M., Minch, K. J., & Sherman, D. R. (2009). Hypoxia: A window into mycobacterium tuberculosis latency. Cellular Microbiology, 11(8), 1151-1159. doi:10.1111/j.1462-5822.2009.01325.x

Singh, A. K., & Reyrat, J. M. (2009, August). Laboratory maintenance of Mycobacterium smegmatis. Retrieved December 07, 2017, from https://www.ncbi.nlm.nih.gov/pubmed/19653213

Trauner, A., Lougheed, K. E. A., Bennett, M. H., Hingley-Wilson, S. M., & Williams, H. D. (2012). The dormancy regulator DosR controls ribosome stability in hypoxic mycobacteria.*The Journal of Biological Chemistry, 287*(28), 24053-24063. doi:10.1074/jbc.M112.364851

Tsien, R. Y., Shaner, N. C., & Steinbach, P. A. (2005). A guide to choosing fluorescent proteins. Nature Methods, 2(12), 905-909. doi:10.1038/nmeth819

Wallace, J. R., Nash, D. R., Tsukamura, M., Blacklock, Z. M., & Silcox, V. A. (1988, July). Human disease due to Mycobacterium smegmatis. Retrieved October 7, 2017, from https://www.ncbi.nlm.nih.gov/pubmed/3392420

Westhof, E. (2004). Faculty of 1000 evaluation for Control of gene expression by a natural metabolite-responsive ribozyme. *F1000 - Post-publication peer review of the biomedical literature*. doi:10.3410/f.1018049.206582

WHO. (2012). *TUBERCULOSIS LABORATORY BIOSAFETY MANUAL* (pp. 1-60, Rep.). Italy: GPS Publishing. Retrieved December 7, 2017.