Investigating the Post-Transcriptional Effects of the *sigA* 5' UTR on Gene Expression

---

A Major Qualifying Project Report

Submitted to

WORCESTER POLYTECHNIC INSTITUTE,

In Partial Fulfillment of the Requirements for the Degree of Bachelor of Science

In

Biology & Biotechnology

Authored By:

Tien Nguyen

April 25th, 2019

Approved by:

Scarlet S. Shell, PhD

Biology and Biotechnology

ABSTRACT

Mycobacterium tuberculosis (MTB) causes tuberculosis, which sickens over 10 million people per year. Despite harsh environmental stresses inside the human host, MTB is able to survive through adaptation and regulation of its gene expression. MTB accomplishes this in part by regulating its mRNA stability. In *Escherichia coli*, 5' Untranslated Regions (UTRs) have been shown to affect mRNA stability; however, this has yet to be shown in mycobacteria. In both MTB and the non-pathogenic model *Mycobacterium smegmatis*, the essential sigma factor, SigA, has an unstable transcript with a relatively short half-life. We hypothesized that *sigA*'s long 5' UTR caused this instability. To test this, we constructed fluorescent reporters and demonstrated that the *sigA* 5' UTR has a modest effect on expression in *M. smegmatis*, and this effect appears to be mediated by altered translation efficiency. Surprisingly, the first 54 nts of the *sigA* coding sequence substantially decreased expression, and this effect appeared to be attributable to reduced transcription and/or reduced mRNA stability.

INTRODUCTION

Although no longer prevalent in developed countries, tuberculosis (TB) continues to rank as one of the leading causes of death worldwide by a single infectious agent, topping both HIV and AIDS (World Health Organization, 2018). In 2017 alone, TB claimed an estimated 1.6 million lives and infected roughly 10 million individuals according to the World Health Organization (World Health Organization, 2018). Various antibiotics and one vaccine have been approved to combat the ongoing war with *Mycobacterium tuberculosis*, although many challenges impede TB eradication. One major obstacle is the constant evolution of the pathogen, resulting in antibiotic resistant strains of the bacteria. Drug resistant *Mycobacterium tuberculosis* (MTB) is much harder, and in some cases, impossible to treat with currently used treatments. Ongoing research focuses on developing new and more efficient antibiotics for MTB to shorten and simplify the current 6 to 9 month intensive multidrug regimen. However, the biology of MTB is incompletely understood, limiting the efficiency and rate of drug discovery. Research on essential genes will provide a deeper understanding how MTB behaves in its environment and how its gene expression alters to increase survivability within a human host.

Sigma factor alpha (SigA) is the primary sigma factor within all mycobacteria and its gene, *sigA,* is essential for growth (Gomez *et al*. 2002, Manganelli *et al*. 2004, and Wagmeester *et al*. 2005). Gomez *et al*. demonstrated that disruption of *sigA* in *Mycobacterium smegmatis*, a non-pathogenic model for MTB, was impossible without an additional copy of the gene at the chromosomal site. In the same study, *sigA* expression and protein levels were shown to be consistently high throughout the growth of a culture (Gomez *et al.* 2002). Other studies have shown that the point mutation R522H in *sigA* resulted in attenuation in *Mycobacterium bovis* or a decrease in antibiotic resistance (Collins *et al.* 1995, Steyn *et al.* 2002, and Burian *et al.* 2013). As the primary housekeeping sigma factor, SigA plays a major role in initiating transcription of most genes within the organism, making it a research target to understand how gene expression is regulated under various environmental stresses. Additionally, *sigA* expression has been shown to be upregulated within human macrophages, which suggest that

SigA increases the expression of genes required for growth within a host (Wu *et al.* 2004, Volpe *et al.* 2006, and Wu *et al.* 2009).

Due to its important roles in the pathogen, several studies investigated the stability of the *sigA* mRNA transcript. One study reported that the *sigA* mRNA transcript was extremely stable with a half-life greater than 40 mins under non-aerated conditions (Hu *et al.* 1999). However, a recent global mRNA stability analysis in MTB demonstrated that *sigA* is one of the most unstable transcripts in the organism with a half life of 6 minutes under aerated conditions (Rustad *et al*. 2013). Similarly, it has been reported that in *M. smegmatis*, *sigA*'s transcript is unstable with a half life of 0.6 minutes (Shell *et al*. 2019). Understanding mRNA stability is important as it contributes to regulation of gene expression within an organism. This regulation changes depending on environmental stresses to increase the organism's growth and adaptation to its surroundings (Nilsson *et al*. 1984, Höner zu Bentrup *et al*. 2001, and Hui *et al*. 2014). Adaptation under environmental stresses can allow bacteria, such as MTB, to obtain resistance to antibiotics and survive within the host (Russell *et al*. 2010).

Regulation of mRNA stability will affect its abundance, which in turn affects how much protein is made. Increased stability is often assumed to result in increased translation and expression (Balestrino *et al*. 2010). Regulation of mRNA half-life to alter mRNA abundance was observed to adapt to changing growth conditions (Rustad *et al*. 2013 and Esquerré *et al*. 2014). However, several studies have demonstrated that there is an inverse correlation between mRNA stability and mRNA abundance in various bacteria (Bernstein et al. 2002, Redon et al. 2005, Rustad *et al*. 2013, and Nouaille *et al*. 2017). The relationship between mRNA abundance, half-life, and expression levels is therefore complicated.

Other mRNA stability studies have shown that stability is influenced and affected by a variety of factors. At the transcript level, stability is regulated through degradation by endoribonucleases, exoribonucleases, as well as interactions with RNA binding proteins and molecules such as sRNAs that can both promote and prevent degradation (Anderson *et al.* 2009, Rustad *et al*. 2013, and Mackie, 2013). In some cases mRNA stability is directly affected by 5' untranslated regions (UTRs) through the promotion or protection from RNases (Belasco *et al*. 1986; Bechhofer *et al*. 1987, and Cho *et al*. 1988). In *Escherichia coli,* multiple studies have shown that the 5' UTR of *ompA*, a gene with one of the most stable transcripts in the organism, significantly increased the stability of the unstable *bla* transcript when substituted for the native *bla* 5' UTR in a chimeric transcript (Belasco *et al*. 1986, Emory *et al*. 1990, and Chen *et al*. 1991). Removal of the 5' UTR of *ompA* decreased its mRNA half life by 5-fold, but stability was restored when the 5' UTR was replaced with known non-native but stabilizing 5' UTRs (Emory *et al*. 1990). The primary stabilizing agents in these 5' UTRs were shown to be a nonspecific stem loop and a sequence that contained the ribosomal binding site, which together seemed to protect the transcript from degradation (Emory *et al*. 1990, Chen *et al*. 1991, Bouvet *et al*. 1992, and Arnold *et al*. 1998). Secondary structure formation in the 5' UTR has also been shown to play a major role in stabilizing transcripts of *ermC* and *pufBA* in *Bacillus subtilis* and *Rhodobacter capsulatus* respectively (Bechhofer and Dubnau, 1987; Heck *et al*. 1996). However, further study of different 5' UTRs revealed that these untranslated sequences can utilize different mechanisms besides secondary structure formation to regulate mRNA stability (Bouvet and Belasco, 1992, Agaisse and Lereclus, 1996).

The sequence of the 5' UTR itself has been shown to determine stability through changing the efficiency of degradation as well as binding to the 30S ribosomal subunit (Bouvet and Belasco, 1992, Agaisse and Lereclus, 1996). One study demonstrated that the target-binding efficiency of RNase E, the primary RNase in *E. coli*, is directly controlled by unpaired nucleotide spacing within the 5' UTR (Bouvet and Belasco; 1992 and Emory *et al.* 1992). In addition to providing binding sites for RNases, the 5' UTR is susceptible to binding to different proteins or sRNAs to prevent or promote translation or affect degradation through other mechanisms (Anderson and Dunman; 2009, Chappell; 2015). In *E. coli*, CsrA, an RNA binding protein, was identified as a global regulator that specifically binds to 5' UTRs to promote degradation (Liu and Romeo; 1997). Also in *E. coli*, the sRNA DrsA has been repeatedly shown to both increase and decrease mRNA stability through binding to the 5' UTR with the help of host factor I protein (Lease and Belfort; 2000 and Moll *et al.* 2003). Agaisse and Lereclus (1996) discovered a "perfect" Shine Dalgarno sequence, referred to as STAB-SD, that is found in many 5' UTRs of gram-positive bacteria and significantly enhances mRNA stability (Agaisse and Lereclus, 1996). This stabilization was a result through enhanced binding of the 3' tail of the 16S rRNA to STAB-SD and STAB-SD's deletion from the 5' UTR prevented stabilization but did not affect translation initiation (Agaisse and Lereclus, 1996). Although effects of 5' UTRs on mRNA stability and the specific mechanisms involved have been widely studied in different organisms, there have been few studies to investigate the presence of this phenomenon in mycobacteria (Unniraman *et al.* 2001).

We hypothesized that the 5' UTR of *sigA* destabilizes its transcript. To identify any potential destabilizing effects, we constructed various reporter plasmids to test the effects of the *sigA* 5' UTR in *M. smegmatis*. Our results demonstrated a significant decrease in expression when we replaced the associated 5' UTR of the strong P_{myc1}tetO promoter with the *sigA* 5' UTR and the first 54 nts of the *sigA* coding sequence. However, we discovered that it was the first 54 nts of the *sigA* coding sequence, not the *sigA* 5' UTR itself, that was primarily responsible for the decrease. Instead, the *sigA* 5' UTR played a more moderate role in the decrese by altering translational efficiency.

METHODS

Bacterial strains and culture conditions

NEB 5-alpha competent *E. coli* (New England Biolabs) were used throughout the study for *E. coli* transformation of various reporter plasmid constructs. Transformed *E. coli* was grown on LB Agar and cultured in LB Broth containing 200 µg/mL of Hygromycin at 37°C. *E. coli* liquid cultures were placed slanted in a 200 rpm shaking incubator for aeration. Due to Hygromycin's sensitivity to light, all plates and liquid cultures were wrapped in aluminum foil.

After confirmation of correct reporter plasmids, SS-M_0346, a non-clumping strain of *M. smegmatis* through the deletion of msmeg_2952, was used for all experiments (Yang *et al.* 2017). *M. smegmatis* was grown on Middlebrook 7H10 plates suplemented with 10X Albumin Dextrose Catalase, final concentrations 5 g/L bovine serum albumin fraction V, 2 g/L dextrose, 0.85 g/L sodium chloride, and 3 mg/L catalase), and 0.5% glycerol. All plates and cultures were grown with 250 µg/mL of Hygromycin at 37°C. Liquid cultures were grown in a 200 rpm shaking

incubator to OD_{600} of ~0.8 for all experiments. Middlebrook 7H9 supplemented with ADC (Albumin Dextrose Catalase, final concentrations 5 g/L bovine serum albumin fraction V, 2 g/L dextrose, 0.85 g/L sodium chloride, and 3 mg/L catalase), 0.2% glycerol and 0.05% Tween 80.

Table 1. *M. smegmatis* strains and plasmids used.

Plasmid Name	Strain Name	Plasmid characteristics	Reference
pSS303	SS-M_0486	P_{myc1} *tetO* Promoter + P_{myc1} 5' UTR + *yfp*	Ehrt et. al. 2005
pSS309	SS-M_0489	P_{myc1} *tetO* Promoter + *sigA* 5' UTR + *yfp*	This work
pSS310	SS-M_0493	P_{myc1} *tetO* Promoter + No 5' UTR + *yfp*	This work
pSS314	SS-M_0497	ΔP_{myc1} *tetO* Promoter + *yfp*	This work
pSS316	SS-M_0521	P_{myc1} *tetO* Promoter + Δ1GTG *sigA* + *yfp*	This work
pSS335	SS-M_0524	P_{myc1} *tetO* Promoter + Δ2GTG *sigA* + *yfp*	This work
pSS359	SS-M_0623	P_{myc1} *tetO* Promoter + first 54 nts of *sigA* + *yfp*	This work
pSS360	SS-M_0626	P_{myc1} *tetO* Promoter + No 5' UTR + first 54 nts of *sigA* + *yfp*	This work
pSS365	SS-M_0629	ΔP_{myc1}*tetO* Promoter + first 54 nts of *sigA* + *yfp*	This work
pSS384	SS-M_0636	ΔP_{myc1}*tetO* Promoter + *sigA* 5' UTR + *yfp*	This work
pSS385	SS-M_0639	ΔP_{myc1}*tetO* Promoter + P_{myc1} 5' UTR + first 54 nts of *sigA* + *yfp*	This work

Primer Design

All primers were designed in Benchling. Primers were designed to be approximately 18 nts to 25 nts in length with a melting temperatures within 5°C of each other, as determined by the New England Biolabs T_M Calculator. In addition to the annealing sequence itself, primers designed for HiFi DNA Assembly included an additional 25 nts tail complementary to the respective backbone for HiFi Assembly (New England Biolabs). All primers were ordered from Eton Bioscience Inc and are identified and described in table 2.

Table 2. Primers sequences and uses.

Primer Name	Primer Sequence	Primer Use
JR273	GACTACACCAAGGGCTACAAG	Forward *sigA* qPCR Primer
JR273	TTGATCACCTCGACCATG	Reverse *sigA* qPCR Primer
SSS247	CAATACGCAAACCGCCTCTC	Forward Primer to amplify pSS242 insert to insert into pSS047 backbone
SSS829	AGCAGAAGAACGGCATCAA	Reverse primer for checking PCR and sequencing
SSS833	GATAGCACTGAGAGCCTGTT	Forward *yfp* qPCR primers set 3
SSS834	CTGAACTTGTGGCCGTTTAC	Reverse *yfp* qPCR primers set 3 also used for checking PCR and sequencing
SSS1013	GCGTTGGCCGATTCATTAATTAGTG CTTGTGGTGGCATCC	Forward checking PCR and sequencing
HiFi pUCIDT rev	TAACTACGTCGACATCGATACTCGCT GGTCCAGAACTGAT	Reverse checking PCR and sequencing
SSS1211	CGAAGAGCCGGTGAAGCGCACCGC TGCCAGCGATAGCACTGAGAG	Forward Primer to amplify backbone of pSS242 with a tail that is complementary to *sigA* 5' UTR for HiFi assembly
SSS1212	CGCTATTCGGTAGGCGGCCGCCGC CCAGAGCCTATCTATCAC	Reverse Primer to amplify backbone of pSS242 with a tail that is complementary to *sigA* 5' UTR for HiFi assembly
SSS1213	GGCGGCCGCCTACCGAATAG	Forward Primer for *sigA* (Msmeg_2758) 5' UTR amplification
SSS1214	AGCGGTGCGCTTCACCGGCTCTT	Reverse Primer for *sigA* (Msmeg_2758) 5' UTR amplification
SSS1228	ATGCCTGGCAGTCGATCGTA	Reverse Primer to amplify pSS242 insert to insert into pSS047 backbone
SSS1229	ACGATCGACTGCCAGGCATCGTCGA CTCTAGAGGATCTACT	Forward Primer to amplify pSS047 backbone with complementary tail to pSS242 insert
SSS1230	GAGAGGCGGTTTGCGTATTGCTGCA GGCATGCAAGCTT	Reverse Primer to amplify pSS047 backbone with complementary tail to pSS242 insert
SSS1376	TAGGGCGTTGCCTCAATCG	Forward Primer to remove destabilization tag of pSS242 insert in pSS265
SSS1377	GGCAACGCCCTAGTGATGGTGATGG TGATGAC	Reverse Primer to remove destabilization tag of pSS242 insert in pSS265 w/ tail for pSS265 annealing
SSS1400	GTGATAGATAGGCTCTGGGCGGCG GCCGCCTACCGAATA	Forward Primer for *sigA* 5' UTR w/ tail to insert into pSS295
SSS1401	TCAGTGCTATCGCTGGCCATAGCGG TGCGCTTCACCGGCT	Reverse Primer for *sigA* 5' UTR w/ tail to

		insert into pSS295
SSS1402	GCCAGCGATAGCACTGAG	Forward Primer for pSS295 for *sigA* 5' UTR Insertion
SSS1403	GCCCAGAGCCTATCTATCA	Reverse Primer for pSS295 for *sigA* 5' UTR Insertion
SSS1410	CATGGCCAGCGATAGCAC	Forward Primer to open pSS293 to insert CT56 sequence
SSS1411	TCCTGGATGACCTCTTTTCC	Reverse Primer to open pSS293 to insert CT56 sequence
SSS1412	GGAAAAGAGGTCATCCAGGAAGAAA TATTGGATCGTCGGC	Forward Primer for CT56 insert w/ tail complementary to pSS293
SSS1413	CAGTGCTATCGCTGGCCATGATGTA TATCTCCTTCTTAAT	Reverse Primer for CT56 insert w/ tail complementary to pSS293
SSS1442	AGATAGGCTCTGGGAATGGCCAGCG ATAGCACT	Forward Primer to amplify pSS303 and remove P_{myc1}*tetO* Promoter 5'UTR w/ tail
SSS1443	TCCCAGAGCCTATCTATCAC	Reverse Primer to amplify pSS303 and remove P_{myc1}*tetO* Promoter 5'UTR
SSS1444	GCCAGCGATAGCACTGAG	Forward Primer to open pSS303 for *sigA* 5' UTR insertion
SSS1445	TCCCAGAGCCTATCTATCAC	Reverse Primer to open pSS303 for *sigA* 5' UTR insertion
SSS1446	GTGATAGATAGGCTCTGGGAGGCG GCCGCCTACCGAATAG	Forward Primer to amplify *sigA* 5'UTR with tails complementary to pSS303
SSS1447	CTCTCAGTGCTATCGCTGGCAGCGG TGCGCTTCACCGGCTCTT	Reverse Primer to amplify *sigA* 5'UTR with tails complementary to pSS303
SSS1482	ACGAGCGGGAGAACTATGGCCAGC GATAGCACT	Forward Primer for 53 nts deletion in P_{myc1}*tetO* Promoter w/ tail
SSS1483	AGTTCTCCCGCTCGTCAG	Reverse Primer for 53 nts deletion in P_{myc1}*tetO* promoter
SSS1484	CGAAAGGGTGTACGTCGCAGCGACA AAGGCAAGC	Forward Primer for mutagenesis on first GTG of *sigA* w/ tail
SSS1485	ACGTACACCCTTTCGGTC	Reverse Primer for mutagenesis on first GTG of *sigA*
SSS1486	AACCGAAGAGCCGGTCAAGCGCAC CGCTGCCAGC	Forward Primer for mutagenesis on second GTG of *sigA* w/ tail
SSS1487	ACCGGCTCTTCGGTTGCC	Reverse Primer for mutagenesis on second GTG of *sigA*
SSS1534	CTCTGACGAGCGGGAGAACTGGCG GCCGCCTACCGAATA	Forward Primer for *sigA* 5' UTR insert w/ tail for promoterless construct
SSS1535	CTCTCAGTGCTATCGCTGGCAGCGG TGCGCTTCACCGG	Reverse Primer for *sigA* 5' UTR insert w/ tail for promoterless construct

SSS1536	GCCAGCGATAGCACTGAG	Forward Primer to open pSS314
SSS1537	AGTTCTCCCGCTCGTCAG	Reverse Primer to open pSS314
SSS1573	GTGGCAGCGACAAAGGCA	Forward Primer for sigA first 54 nts + 488 nts of YFP to use as the *sigA* insert
SSS1574	TTGGCCTTGATGCCGTTCTT	Reverse Primer for *sigA* first 54 nts + 488 nts of YFP to use as the *sigA* insert
SSS1575	CTTGCCTTTGTCGCTGCCACGATGT ATATCTCCTTCTTAATTAAGCAT	Reverse Primer for new pSS303 construct with tail for *sigA* insert
SSS1576	CTTGCCTTTGTCGCTGCCACTCCCA GAGCCTATCTATCA	Reverse Primer for new pSS310 construct with tail for *sigA* insert
SSS1577	CTTGCCTTTGTCGCTGCCACAGTTCT CCCGCTCGTCAG	Reverse Primer for new pSS314 construct with tail for *sigA* insert
SSS1578	CTCTGACGAGCGGGAGAACTTACCC GTGTGTACGACCA	Forward Primer for cryptic P_{myc1}*tetO* 5' UTR + *sigA* insert w/ tail for pSS314

Plasmid construction

All plasmid constructions utilized a previously made backbone in the Shell Lab, systematically known as pSS047. This construct included a Giles phage integration site and a hygromycin antibiotic resistance marker and was derived from pGH1000A (Morris *et al.* 2008). The *yfp* expression cassette from another Shell Lab plasmid, pSS242, was inserted into the pSS047 backbone. This expression cassette included the reporter *yfp* with a C-terminal His-tag, and was flanked by an upstream terminator, tsynA, and a synthetic downstream bi-directional terminator, tt_{sbi} (Czyz *et al.* 2014 and Huff *et al.* 2010). Additionally, the promoter, P_{myc1}*tetO tetO*, was taken from plasmid CT56 and replaced the original promoter in the expression cassette. Included in this strong promoter was the P_{myc1}*tetO* Promoter itself and its associated 5' UTR (Ehrt *et al*. 2005). The P_{myc1}*tetO* associated 5' UTR was replaced with the *sigA* 5' UTR with the first 54 nts of the *sigA* coding sequence obtained from genomic *M. smegmatis* DNA when indicated.

All backbones and inserts were amplified using Q5 DNA Polymerase (New England Biolabs) as indicated in Table 2. Samples were denatured at 98°C for 30 seconds before going through 32 cycles of repeated denature at 98°C for 20 seconds, annealing at the appropriate temperature determined by New England Biolabs T_M calculator, and elongation at 72°C for 30 seconds per kb. Samples underwent a final elongation at 72°C for 1 minute per kb. Q5 PCR products underwent a Dpn1 digest to cleave methylated plasmid template DNA. 0.5 µL of Dpn1 was added directly to the PCRs and incubated at 37°C for 15 minutes. Dpn1 was heat inactivated at 80°C for 20 minutes. All Q5 PCR products were run on a 1% TAE Quick Dissolve agarose gel (Genesee) containing ethidium bromide. Correctly sized bands were cut out and extracted using the Monarch® DNA Gel Extraction Kit following the manufacturer's instructions (New England Biolabs). All gel extracts were measured for concentration and purity on a Nanodrop (Thermo Fisher).

For HiFi assembly, ~50 ng of vector was used with the appropriate amount of insert required to achieve a 2:1 molar ratio. Using the New England Biolabs Ligation Calculator, the

recommended concentration of insert was determined to meet the requirements of the recommended 2:1 ratio. Vector and insert were diluted appropriately and combined into a PCR tube with a final volume of 2.5 µL. 2.5 µL of 2X HiFi DNA Assembly® Master Mix (New England Biolabs) was added to the tube and incubated at 50°C for 1 hour. The assembled product was transformed into NEB® 5-alpha Competent *E. coli* (New England Biolabs) following the manufacturer's high efficiency protocol. Transformed *E. coli* were plated on LB+ Hygromycin (200 µg/mL) and incubated at 37°C for 24 hours.

Colony checking PCR was done with Taq Polymerase and primers previously created (Table 3) to verify the presence of appropriately sized inserts before sequencing positive clones. Samples were denatured at 95°C for 5 minutes before entering 32 repeated cycles of denaturing at 95°C, annealing at 65°C, and elongation at 65°C. All Checking PCR products were run on 1% TAE agarose gels (Genesee).

The ZR Plasmid Miniprep - Classic (Zymo) kit was used to extract and purify plasmids from transformed *E. coli* according to the manufacturer's instructions. To verify the sequences of purified plasmids, 80 ng/µL of samples were sent out to Quintarabio with sequencing primers that were approximately 100 nts upstream or downstream of the region of interest. Sequences were verified both through visual inspection of chromatographs and sequence alignments in Benchling.

Table 3. Q5 and Checking PCR 1X Master Mix

Q5 PCR	**1X Reaction (µL)**	**Checking PCR**	**1X Reaction (µL)**
Q5 Buffer	10	Taq Buffer	2.5
10 mM each dNTPs	2	10 mM each dNTPs	0.5
10 µM forward primer	2.5	DMSO	1
10 µM reverse primer	2.5	10 µM forward primer	0.5
GC Enhancer	10	10 µM reverse primer	0.5
Ultra Pure H_2O	22.5	Ultra Pure H_2O	19.875
Q5 Polymerase	0.5	Taq Polymerase	0.125
Template DNA	1	Liquid culture	1
Total Volume	24	Total Volume	25

M. smegmatis Competent Cells and Transformation

50 mL cultures of SS-M_0346 were grown to an OD_{600} of ~0.6 to make competent cells for transformation. Cultures were spun down at 4°C at 3,900 rpm and washed with 25 mL of

10% glycerol. 2 additional washes were completed with 10 mL and then 5 mL of 10% glycerol. The cells were resuspended in 150 µL of 10% glycerol and aliquoted into 1.5 mL tubes, each containing 30 µL of competent cells. All tubes were stored at 80°C for future use.

Approximately 200 ng of reporter plasmid was added to a cuvette of thawed competent cells and were electroporated in a MicroPulser Electroporator (Bio-rad). Cells were immediately rescued with 200 µL of Middlebrook 7H9 and incubated in a 37°C shaking incubator for 2 hours. All samples were plated on Middlebrook 7H10 plates with 250 ug/mL of Hygromycin and incubated at 37°C for 3 days.

Paraformaldehyde Fixation for Microscopy and Flow Cytometry

Transformed *M. smegmatis* liquid cultures were grown to an OD_{600} of ~0.8. Pellets from 1.5 mL of culture were resuspended with 50 µL of 2% paraformaldehyde and incubated at room temperature in the fume hood for 30 minutes. Samples underwent two washes of 900 µL of Phosphate Buffered Saline with 1% Tween-20 (PBS-T), spinning at 14,000 rpm for 2 minutes each wash. All samples were resuspended with an appropriate volume of PBS-T to have a final calculated OD_{600} of 15.

Fluorescence Microscopy

To prevent photobleaching, 2 µL of all samples were mixed with 6 µL of mounting media. Melted General Purpose Agarose (Genesee) was prepared on a microscope slide to keep samples in the same plane of focus. Samples were loaded onto the solidified gel and imaged with a Zeiss Axio Imager Z1 with the Zeiss Apotome (Zeiss) under 40X oil immersion. Fields containing multiple identifiable single cells were imaged using a DIC 60 channel and a GFP channel for brightfield and fluorescent images respectively. All image display parameters were set to be identical to the parameters of the appropriate control for specific experiments.

Flow Cytometry and Analysis

All samples were diluted to an OD_{600} of 0.015 with Middlebrook 7H9. Controls and samples were run on a BD Accuri C6 flow cytometer following the manufacturer's instructions (data shown in figure 1) or on a BD LSRII B-4 laser flow cytometer (data shown in figure 3 and 5). Samples run on the Accuri C6 were filtered with a 5 µm filter needle to remove clumps. For each sample, 100,000 events were collected and later gated during analysis. However, only 50,000 samples were collected using the BD LSRIIB-4. Appropriate thresholds were made after running Middlebrook 7H9 and water controls to exclude any electronic noise.

FlowJo (FlowJo®) software was used to analyze all flow cytometry data. Forward scatter and side scatter represented relative cell size and complexity of samples and was used to select a gate around the densest population. Within the gate, samples were compared on forward scatter vs fluorescence graphs to identify difference in fluorescence levels. Fluorescence histograms were created to more accurately represent fluorescence peaks between gated samples. Statistical analysis of mean, median, standard deviation, and coefficient of variation were calculated within the software.

RNA Extraction

RNA samples were extracted from 5 mL of liquid nitrogen frozen *M. smegmatis* liquid cultures using the Direct-zol™ RNA extraction and purification kit (Zymo). Thawed liquid cultures were centrifuged at 3,900 rpm at 4°C for 5 minutes. After discarding the supernatant, samples were resuspended in 1 mL of TRIzol™ (Thermo Fisher) and immediately transferred to 100 µm zirconium lysing matrix bead breaking tubes (OPS Diagnostics). Samples were lysed in a FastPrep 5G using 3 cycles of 7 m/s, 30 seconds per cycle, with a 2 min on-ice step between each cycle (MP Biomedicals). 300 µL of chloroform was added to all samples in the fume hood.

To prevent degradation of RNA during extraction, all equipment and bench were wiped with RNase*ZAP*™ (Thermo Fisher). All samples were vortexed for 15 seconds on the highest setting with a subsequent centrifugation at 15,000 rpm at 4°C for 15 minutes for phase separation. 500 µL of the aqueous phase was removed without disturbing the bottom phase and evenly mixed with 500 µL of 100% ethanol. Samples were then transferred to Direct-Zol columns and centrifuged at 15,000 g for 30 seconds at room temperature. All flow through was removed by careful aspiration and discarded. Samples were washed with 400 µL of RNA Wash Buffer (Zymo). After initial wash, samples underwent DNAse Treatment using 80 µL of DNase Master mix which contained 75 µL of DNase Digestion Buffer and 5 µL of DNase 1 and were left to incubate at room temperature for 30 minutes (Zymo). Samples were then washed 400 µL of Direct-Zol Pre-wash Buffer twice before completing a final wash of 700 µL of RNA Wash Buffer with a 2 minute centrifugation (Zymo). Direct-Zol Columns were transferred to a clean 1.5 mL tube and were eluted in 50 µL of RNase-free water. RNA extracts were briefly vortexed for resuspension and measured on a Nanodrop™ (Thermo Fisher). Samples were stored at -80°C.

cDNA Synthesis and Clean Up

All equipment required for cDNA synthesis were decontaminated with RNase*ZAP*™ (Thermo Fisher). RNA samples were thawed and diluted to 600 ng in 5.25 µL. This dilution was completed twice to have reverse-transcriptase (RT) and no RT samples for experiments. 1 µL of a random primer master mix was added to each sample (0.83 µL of 100 mM Tris pH 7.5 and 0.17 µL of 3 mg/mL random primers (NEB)). Samples were incubated at 70°C for 10 minutes and immediately placed in an ice-water bath for 5 minutes. 3.75 µL of RT and no-RT master mixes were added to each respective sample (Table 4). Samples were incubated at 25°C for 10 minutes and at 42°C for 2 hours. cDNA samples were kept at 4°C.

Prior to cDNA clean up, all samples underwent RNA degradation using 5 µL of 0.5 mM EDTA and 5 µL of 1 N NaOH for each sample and a 10 minute incubation at 65°C followed by 4°C. After 1 minute at 4°C, RNA degradation was immediately stopped by adding in 12.5 µL of 1 M Tris-HCl pH 7.5 to each sample. The Qiagen MinElute PCR Purification Kit was used for cDNA clean up (Qiagen) and all centrifugation steps occurred at 17,900 *g* for 30 seconds at room temperature. Samples were mixed with 400 µL of Buffer PB and then transferred to the MinElute columns and spun down. All flow through was discarded after each centrifugation. Columns were washed with 750 µL of Buffer PE three times. A final centrifugation was completed without any buffer to remove any remaining ethanol in the column. MinElute columns

were transferred to new 1.5 mL tubes and eluted with 30 µL of RNase free water. Samples were carefully resuspended and measured on a Nanodrop (Thermo Fisher). All cDNA samples were stored at -20°C.

Table 4. cDNA synthesis RT and no-RT master mixes

Reagent	1x RT (µL)	1x no-RT (µL)
ProtoScript II Buffer (5x)	2	2
dNTPs mix (10 mM each)	0.5	0.5
100 mM DTT	0.5	0.5
RNase Inhibitor, Murine, NEB (40,000 U/mL)	0.25	0.25
ProtoScript® II Reverse Transcriptase, NEB (200,000 units/ml)	0.5	0
H_2O	0	0.5
TOTAL	3.75	3.75

qPCR

All cDNA samples were diluted to 1 ng/ µL in RNase-Free water. Samples were further diluted to bring the final concentration of all samples to 200 pg and loaded 2 µL into a 96 well plate. A qPCR master mix was made using 1 µL of 2.5 µM qPCR primer mix, 5 µL of iTaq SYBR Green Supermix (BioRad), and 2 µL of RNase-Free Water for each sample. Two different primer sets were used, JR273 and JR274 for *sigA* and SSS833 and SSS834 for *yfp*. 8 µL of the qPCR master mix was mixed with each respective sample in the 96 well plate. After mixing, the 96 well plate was sealed and placed in the thermocycler. The plate was incubated at 50°C for 2 minutes followed by 95°C for 10 minutes and 40 cycles of 95°C for 15 seconds and 61°C for 1 minute. Once run was complete, the threshold was changed to 0.2 prior to exporting the data for analysis.

Statistical Analysis

Mean values and median were used for all statistical analysis within GraphPad (Prism; GraphPad Software, San Diego, CA). A One-way ANOVA with a post-hoc Tukey Test (****, p<0.0001) for qPCR data or Kruskal-Wallis with a post-hoc Dunn's Test (****, p<0.0001) for flow cytometry data were used to determine if differences between central tendencies were statistically significant.

RESULTS

Plasmid construction

To test the effects of the *sigA* 5' UTR on expression, 11 *yfp* reporter constructs were designed and transformed into *M. smegmatis*. In the design process, we wanted to include the first few nucleotides of the *sigA* coding sequence with our *sigA* 5' UTR construct. However, we were unable to unambiguously identify the *sigA* 5' UTR/coding sequence boundary due to differing annotations of the *sigA* start codon (Figure 1A). To determine which of the two annotated *sigA* start codons is actually used to initiate translation, we designed two constructs where either the upstream GTG or downstream GTG was mutated to GTC and compared these to a construct containing both GTGs (Figure 1B). Mutation of first GTG resulted in no fluorescence above the autofluorescence of a no-plasmid control (Figure 1C and D). Mutagenesis of the second GTG produced fluorescence levels statistically distinguishable from the no plasmid control, but lower than that of the *sigA* 5' construct with both GTGs unmutated (Figure 1C and D). We concluded that while the second GTG may play a role in expression, the first GTG is required for translation and therefore designated the first GTG as the 'true' start codon of *sigA*. We therefore included 54 nts of the *sigA* coding sequence, beginning with the first GTG, in our *sigA* 5 UTR construct. We introduced the same first 54 nts of the *sigA* coding sequence, referred to as *sigA54*, into all of our constructs for translational consistency (Figure 2). We used these 9 constructs for our downstream experiments to test the effect of the *sigA* 5' UTR and *sigA54* on expression of our YFP reporter.

Insertion of *sigA54* significantly decreases fluorescence levels

Flow cytometry fluorescence quantifiaction demonstrated high fluorescence for our P_{myc1}*tetO* Promoter + P_{myc1}*tetO* associated 5' UTR construct, which was expected as this promoter was reported to be very strong (Ehrt et al. 2005). However, replacement of the P_{myc1}*tetO* associated 5' UTR with the *sigA* 5' UTR + *sigA*54 significantly decreased fluorescence (Figure 3A and B). Surprisingly, when we introduced only *sigA54* into the P_{myc1}*tetO* associated 5' UTR construct, there was also a substantial decrease in fluorescence. However, there was a statistical 2-fold difference in median fluorescence between the P_{myc1}*tetO* associated 5' UTR + sigA54 construct with the *sigA* 5' UTR + *sigA54* construct. Despite the major decrease caused by *sigA54,* both constructs were significantly more fluorescent than promoterless control constructs. We concluded that both the 5' UTR of *sigA* and the first 54 nts of its coding sequence lead to reduced production of YFP protein compared to the P_{myc1}*tetO* associated 5' UTR and native YFP.

Mycobacteria are known to translate leaderless genes that completely lack 5' UTRs. We therefore made leaderless constructs to compare to our various leadered constructs. Leaderless constructs both with and without *sigA54* showed even lower fluorescence than the *sigA* 5' UTR + *sigA54* construct but were higher than the promoterless control. Interestingly, we did not find a substantial difference in expression in the presence or absence of *sigA54* in the leaderless constructs. This suggests that the impact of *sigA54* is context-dependent.

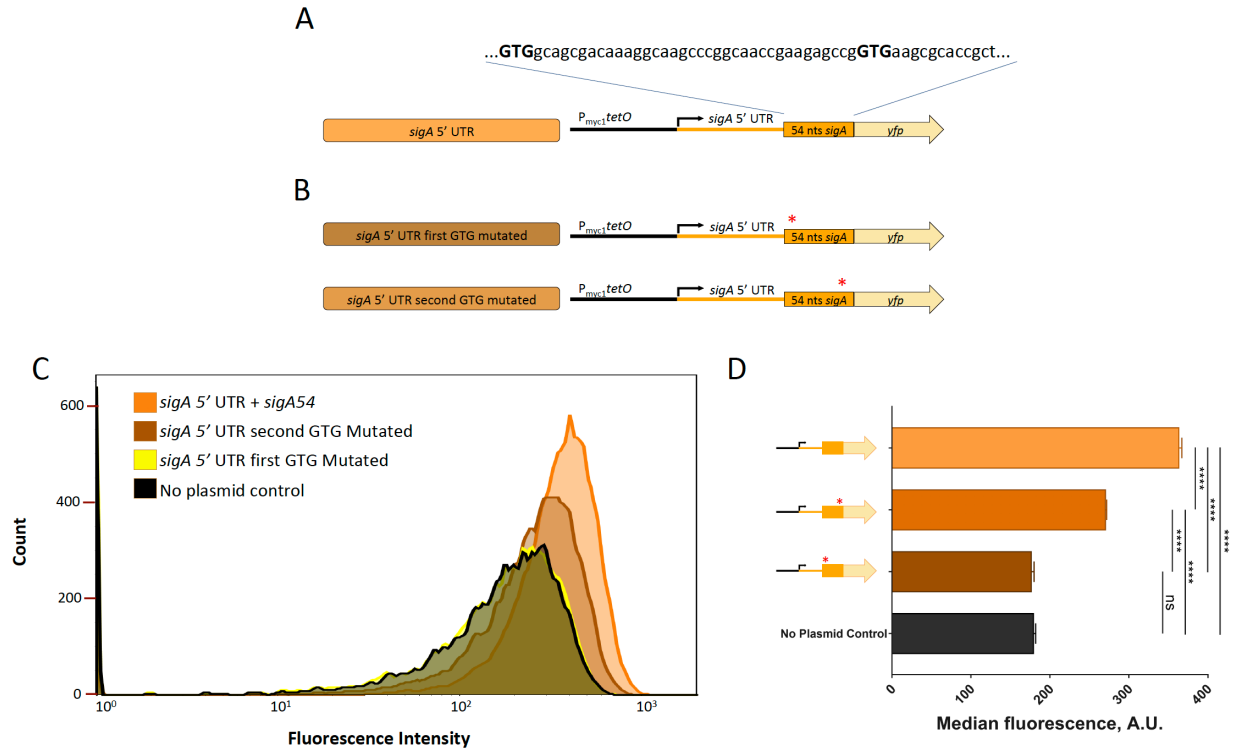

Figure 1. Identifying the *sigA* start codon. A) Mutation strategy. We used primer mutagenesis to mutate either the first or second GTG to GTC. B) Two constructs containing either the first or second GTG mutation. C) Flow cytometry fluorescence quantification. 100,000 events were collected per sample. Shown in light orange is the *sigA* 5' UTR construct. Shown in dark orange is the *sigA* 5' UTR with the second GTG mutated. Shown in yellow is the *sigA* 5' UTR with the first GTG mutated. Shown in black is the no plasmid control. D) Statistical analysis using the median fluorescence of our flow data showed a decrease in fluorescence when the second GTG was mutated and complete abolishment of above-background fluorescence with the first GTG was mutated. (Kruskal-Wallis with a post hoc Dunn's test, **** p<0.0001). Error Bars: Median with 95% CI.

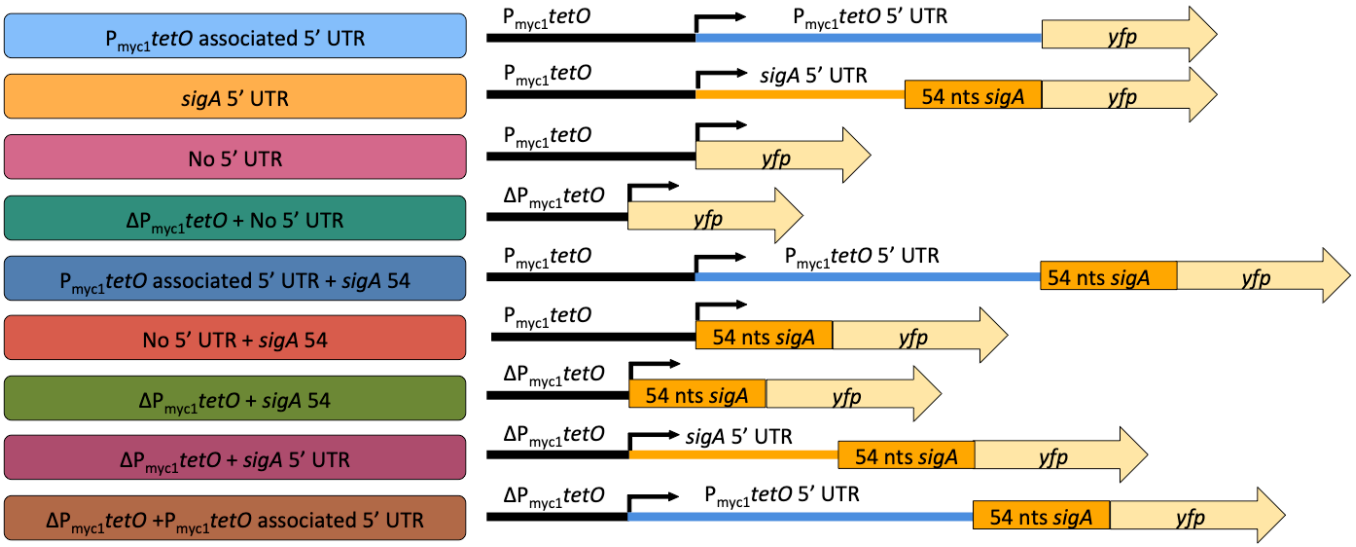

Figure 2. Fluorescent Reporter Constructs. Constructs were designed and transformed into *M. smegmatis* for flow cytometry and qPCR. Constructs with *sigA54* replaced the native start codon of *yfp*. 53 nts of the P$_{myc1}$*tetO* promoter were deleted to create ΔP$_{myc1}$*tetO* promoter, a non-functional promoter

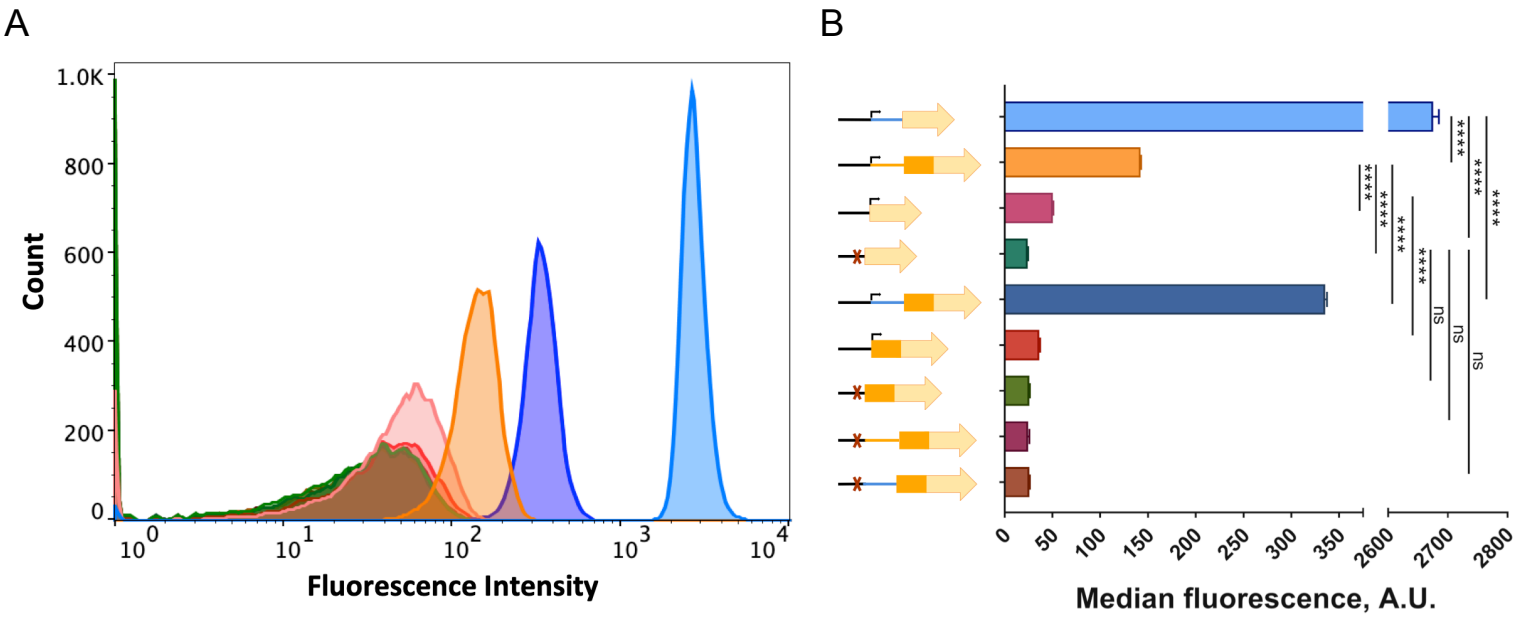

Figure 3. The *sigA* 5' UTR and first 54 nt of the *sigA* coding sequence (*sigA54*) both affect expression. A) Flow cytometry histogram showing fluorescent intensities of constructs. Shown in light blue is the P_{myc1}*tetO* associated 5' UTR. Dark blue is the P_{myc1}*tetO* associated 5' UTR with *sigA*54. Orange is *sigA* 5' UTR with sigA54. Bright red is no 5' UTR. Dark red is no 5' UTR with *sigA*54. Green is promoterless. Dark Green is promoterless with *sigA*54. Drastic decreases in fluorescence intensities are observed when the P_{myc1}*tetO* associated 5' UTR is combined with *sigA*54 or replaced with the *sigA* 5' UTR + *sigA54*. B) Statistical analysis showing the median fluorescence from the flow cytometry data set. Replacment of the the P_{myc1}*tetO* associated 5' UTR with the *sigA* 5' UTR + *sigA54* significantly decreased median fluorescence. Introduction of *sigA*54 into the P_{myc1}*tetO* associated 5' UTR construct also reduced fluorescence. (Kruskal-Wallis with a post hoc Dunn's Test, **** p<0.0001). Error Bars: Median with 95% CI.

Comparing the promoterless controls, which lacked the -35 and -10 sequences, to the no-plasmid control showed no significant difference (data not shown), confirming that residual fluorescence in the promoterless controls is likely autofluorescence rather than spurious expression of YFP. We found no fluorescence when testing both the P_{myc1}*tetO* associated 5' UTR + *sigA54* and the *sigA* 5' UTR + *sigA54* in a promoterless control, confirming that there was no promoter within the inserted sequences. We identified our flow cytometry data set as non-parametric and therefore utilized a Kruskal-Wallis Test and post hoc Dunn's test for our statistical analysis (****, *p<0.0001)*.

Insertion of *sigA54* significantly decreases *yfp* mRNA levels

To determine if differences in YFP protein levels were a consequence of differences in RNA levels, we performed qPCR to determine steady-state abundance of the *yfp* transcript in our strain set. We extracted RNA from all 9 strains with our experimental constructs that were grown to 0.5-0.7 OD, and used these RNA samples for cDNA synthesis and qPCR with previously validated primers for *yfp* as well as *sigA* for an internal control. The *sigA* primers anneal to a downstream portion of the native *sigA* gene that was not present in any of our reporter constructs. Our strong promoter construct with the P_{myc1}*tetO* associated 5' UTR yielded the highest *yfp* mRNA levels (Figure 4). Replacement of the P_{myc1}*tetO* associated 5' UTR with the *sigA* 5' UTR + *sigA54* resulted in a 5-fold reduction of *yfp* expression. Similarly, we observed a 4-fold decrease in expression upon introducing the *sigA54* alone to the P_{myc1}*tetO* associated 5' UTR construct. This demonstrated that the major decrease in *yfp* mRNA abundance was

primarily due to *sigA54* and not the *sigA* 5' UTR. However, comparison between those respective constructs showed slight difference in mRNA levels, although this was not statistically significant (Figure 4).

Constructs lacking a 5' UTR also showed a 5-fold decrease in *yfp* mRNA abundance similar to that of the *sigA* 5' UTR + *sigA54* construct. The differences in mRNA levels between matched leadered and leaderless constructs were substantially smaller than the differences in protein levels, indicating that the leaderless transcripts are translated less efficiently than the leadered transcripts. As expected, the promoterless controls showed no statistical difference in *yfp* mRNA abundance compared to the no plasmid control.

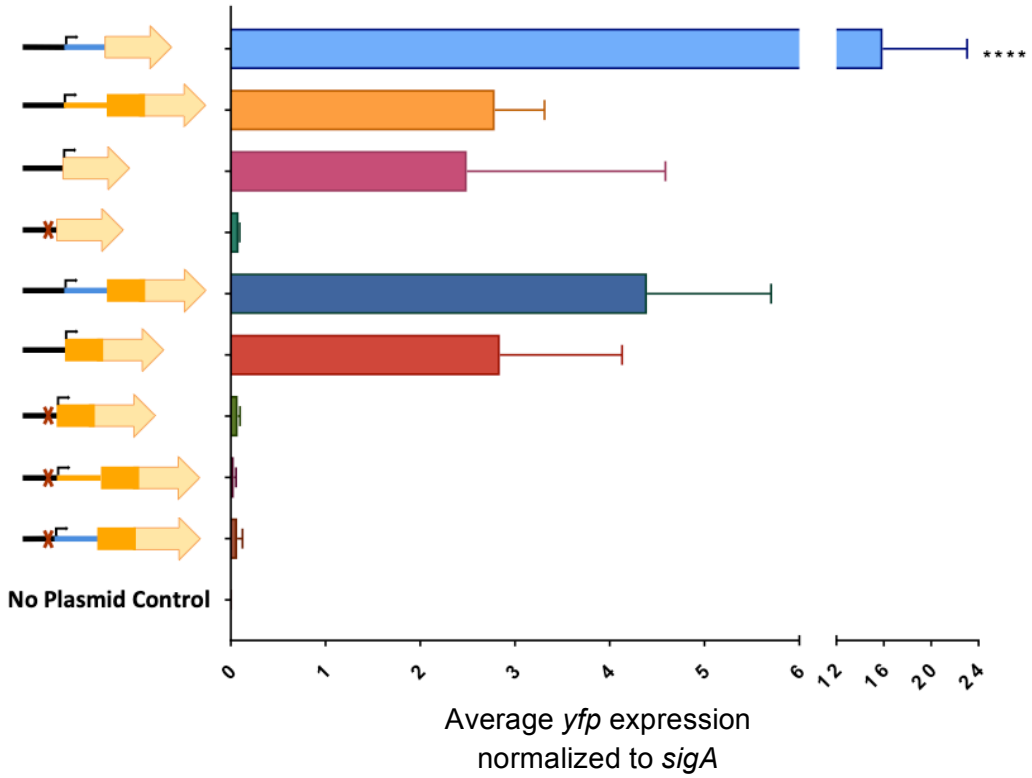

Figure 4. qPCR analysis showed decreases in RNA abundance for *sigA54* and the *sigA* 5' UTR. All respective constructs are shown on the y-axis. The P_{myc1}*tetO* associated 5' UTR construct yielded the highest RNA abundance and expression. Adding in *sigA54* to this construct decreased expression by 4-fold. Replacing the P_{myc1}*tetO* associated 5' UTR with the *sigA* 5' UTR + *sigA54* decreased expression by 5-fold. Only the P_{myc1}*tetO* associated 5' UTR construct was statistically different from all the other constructs.(One-way ANOVA with a post hoc Tukey's Test, **** p<0.0001). Error Bars: Mean with SD.

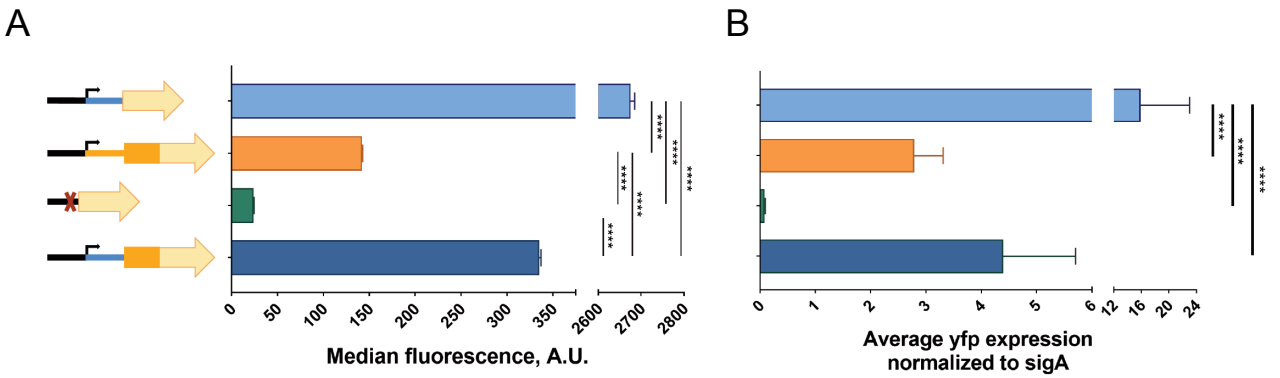

Figure 5. Direct comparison of protein and RNA analysis shows *sigA54* affecting transcription while the *sigA* 5' UTR affects translation. A) Flow cytometry data from a subset of strains focusing on the P_{myc1}*tetO* associated 5' UTR construct, the *sigA* 5' UTR construct, the promtoerless control, and the P_{myc1}*tetO* associated 5' UTR construct with *sigA54.* There was a significant difference in protein levels between the P_{myc1}*tetO* associated 5' UTR construct with *sigA54* and the *sigA* 5' UTR + *sigA54* construct. (Kruskal-Wallis with a post hoc Dunn's Test, **** p<0.0001) Error Bars: Median with 95% CI. *B*) qPCR data from a subset of strains focusing on the same constructs in A. Similar RNA levels are observed levels between the P_{myc1}*tetO* associated 5' UTR construct with *sigA54* and the *sigA* 5' UTR + *sigA54* construct. *sigA54* decreases both RNA levels and protein levels. (One-way ANOVA with a post hoc Tukey's Test, **** p<0.0001). Error bars: Mean with SD

sigA54 affects transcription and/or mRNA stability while the *sigA* 5' UTR affects translation

We compared our protein and mRNA analyses to infer whether changes in protein abundance could be attributed to changes in translation efficiency or changes in transcript abundance. Insertion of the *sigA54* into the P_{myc1}*tetO* associated 5' UTR construct decreased mRNA abundance and protein abundance to similar extents, suggesting that the lower protein levels were a consequence of lower mRNA levels (Figure 5A and 5B). mRNA levels could be lower as a result of reduced transcription or faster mRNA degradation. Furthermore, when specifically comparing P_{myc1}*tetO* associated 5' UTR + *sigA54* to the *sigA* 5' UTR + *sigA54* constructs, we found that although relative RNA levels were similar between the two constructs, protein levels had a statistically significant 2-fold difference (Figure 5A and 5B). This led us to believe that the *sigA* 5' UTR affects translation efficiency.

DISCUSSION

Given the unstable nature of mycobacterial *sigA* transcripts and the presence of unusually long 5' UTRs on the *sigA* transcripts in both *M. tuberculosis* and *M. smegmatis*, we sought to determine if the *sigA* 5' UTR affects expression in *M. smegmatis*. Our preliminary experiments with our first three constructs, the P_{myc1}*tetO* associated 5' UTR construct, the *sigA* 5' UTR + *sigA54* construct, and the leaderless construct, demonstrated that insertion of the *sigA* 5' UTR + *sigA54* significantly decreased fluorescence. However, the cause of this decrease could be explained in various ways. One potential explanation was a difference in translation efficiency. When inserting the *sigA* 5' UTR we also introduced 54 nts of the *sigA* coding sequence fused to our reporter *yfp*, with the goal retaining any potential for secondary structures that could form between the UTR and the early part of the coding sequence. Therefore, our initial results were comparing two slightly different proteins. It is potentially possible that this addition of 54 nts to *yfp* interferes with translation or affects the protein itself due to the

presence of extra amino acids that could affect protein folding or stability. To address this issue, we created additional constructs to replace the native start codon of *yfp* with the first 54 nts of the *sigA* coding sequence as was present in the *sigA* 5' UTR construct (Figure 2). By doing this, we could compare constructs that differed only the presence or absence of specific UTRs while encoding identical proteins.

However, to create these constructs we needed to identify which of two annotated start codons was the true start codon of *sigA*. Creation of two constructs with either the first or second start codon mutated allowed us to determine which of the two was primarily used for translation. Ideally, mutation of one codon would result in complete abolishment of fluorescence while mutation of the other would have no effect at all. However this extreme difference was not observed in our results. Although mutation of the first start codon did result in fluorescence levels non-distinguishable from background, mutation of the second start codon also caused a decrease in fluorescence intensity compared to the unmutated *sigA* 5' UTR construct (Figure 1C and D). This potentially suggests that although the first start codon can be considered the primary start codon, only when both start codons are unmutated and present is *sigA* expressed at its highest efficiency. It is possible that mutating the second start codon affects the efficiency of translation elongation for translation events that initiated at the first start codon, if the GTC codon is less preferred than GTG. In addition to affecting translational efficiency, it is also possible that mutation of the second start codon negatively affects transcript stability. Another potential explanation is that translation is simply initiating at both sites, and recruitment of the translational machinery to the first start codon affects the efficiency of recruitment of the translational machinery to the second start codon. There are various possible explanations for the decrease in fluorescence upon mutating the second start codon, but exploring them was beyond the scope of this study.

After completing quantitative analysis for both fluorescence and expression, we observed a significant decrease in fluorescence in constructs containing either the *sigA* 5' UTR + *sigA54* or *sigA54* without the *sigA* 5' UTR. The *sigA54* sequence itself therefore appeared to decrease expression at the levels of both transcript abundance and protein abundance (Figures 3B and 4). This leads us to hypothesize that the *sigA54* is specifically affecting transcription or mRNA stability but not translation, as its effects on RNA and proteins levels of a similar magnitude (Figure 5A and B). The mechanism in which *sigA54* affects transcript abundance is not known and is an area for future research. We hypothesize that *sigA54* may be either affecting transcription efficiency or transcript half-life, which could be distinguished in future work by measuring the half-life of the *yfp* transcript in our various constructs. Additionally, despite relatively similar RNA levels between the *sigA* 5' UTR + *sigA54* construct and P_{myc1}*tetO* 5' UTR + *sigA54* construct, significantly lower protein levels were observed in the former. This potentially suggests that although RNA abundance is similar, the 5' UTR is affecting translation efficiency resulting in decreased fluorescence. Together, our results demonstrate that the *sigA54* significantly decreases transcript levels while the *sigA* 5' UTR moderately affects expression through regulating translation.

ACKNOWLEDGEMENTS

I would like to thank Professor Shell and everyone else in the Shell Lab for helping and supporting my project. Without them and the Biology and Biotechnology Department at Worcester Polytechnic Institute this project would not have been possible. This project was funded through Professor Shell and her NSF grant: NSF CAREER #1652756 (to SSS)

REFERENCES

Agaisse, H., & Lereclus, D. (1996). STAB-SD: a Shine–Dalgarno sequence in the 5′ untranslated region is a determinant of mRNA stability. *Molecular Microbiology*, *20*(3), 633–643. https://doi.org/10.1046/j.1365-2958.1996.5401046.x

Anderson, K. L., & Dunman, P. M. (2009). Messenger RNA Turnover Processes in Escherichia coli, Bacillus subtilis, and Emerging Studies in Staphylococcus aureus. *International Journal of Microbiology*, *2009*. https://doi.org/10.1155/2009/525491

Arnold, T. E., Yu, J., & Belasco, J. G. (1998). mRNA stabilization by the ompA 5' untranslated region: two protective elements hinder distinct pathways for mRNA degradation. *RNA*, *4*(3), 319–330.

Bechhofer, D. H., & Dubnau, D. (1987). Induced mRNA stability in Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America*, *84*(2), 498–502.

Belasco, J. G., Nilsson, G., von Gabain, A., & Cohen, S. N. (1986). The stability of E. coli gene transcripts is dependent on determinants localized to specific mRNA segments. *Cell*, *46*(2), 245–251. https://doi.org/10.1016/0092-8674(86)90741-5

Bentrup, K. H. zu, & Russell, D. G. (2001). Mycobacterial persistence: adaptation to a changing environment. *Trends in Microbiology*, *9*(12), 597–605. https://doi.org/10.1016/S0966-842X(01)02238-7

Bernstein, J. A., Khodursky, A. B., Lin, P.-H., Lin-Chao, S., & Cohen, S. N. (2002). Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. *Proceedings of the National Academy of*

Sciences of the United States of America, *99*(15), 9697–9702. https://doi.org/10.1073/pnas.112318199

Bouvet, P., & Belasco, J. G. (1992). Control of RNase E-mediated RNA degradation by 5′-terminal base pairing in E. coil. *Nature*, *360*(6403), 488–491. https://doi.org/10.1038/360488a0

Burian, J., Yim, G., Hsing, M., Axerio-Cilies, P., Cherkasov, A., Spiegelman, G. B., & Thompson, C. J. (2013). The mycobacterial antibiotic resistance determinant WhiB7 acts as a transcriptional activator by binding the primary sigma factor SigA (RpoV). *Nucleic Acids Research*, *41*(22), 10062–10076. https://doi.org/10.1093/nar/gkt751

Chappell, J., Watters, K. E., Takahashi, M. K., & Lucks, J. B. (2015). A renaissance in RNA synthetic biology: new mechanisms, applications and tools for the future. *Current Opinion in Chemical Biology*, *28*, 47–56. https://doi.org/10.1016/j.cbpa.2015.05.018

Chen, L. H., Emory, S. A., Bricker, A. L., Bouvet, P., & Belasco, J. G. (1991). Structure and function of a bacterial mRNA stabilizer: analysis of the 5’ untranslated region of ompA mRNA. *Journal of Bacteriology*, *173*(15), 4578–4586. https://doi.org/10.1128/jb.173.15.4578-4586.1991

Cho, K.-O., & Yanofsky, C. (1988). Sequence changes preceding a Shine-Dalgarno region influence trpE mRNA translation and decay. *Journal of Molecular Biology*, *204*(1), 51–60. https://doi.org/10.1016/0022-2836(88)90598-0

Collins, D. M., Kawakami, R. P., de Lisle, G. W., Pascopella, L., Bloom, B. R., & Jacobs, W. R. (1995). Mutation of the principal sigma factor causes loss of virulence in a strain of the Mycobacterium tuberculosis complex. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(17), 8036–8040.

Czyz, A., Mooney, R. A., Iaconi, A., & Landick, R. (2014). Mycobacterial RNA Polymerase Requires a U-Tract at Intrinsic Terminators and Is Aided by NusG at Suboptimal Terminators. *MBio*, *5*(2), e00931-14. https://doi.org/10.1128/mBio.00931-14

Dubnau, E., Fontán, P., Manganelli, R., Soares-Appel, S., & Smith, I. (2002). Mycobacterium tuberculosis Genes Induced during Infection of Human Macrophages†. *Infection and Immunity*, *70*(6), 2787–2795. https://doi.org/10.1128/IAI.70.6.2787-2795.2002

Ehrt, S., Guo, X. V., Hickey, C. M., Ryou, M., Monteleone, M., Riley, L. W., & Schnappinger, D. (2005). Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor. *Nucleic Acids Research*, *33*(2), e21–e21. https://doi.org/10.1093/nar/gni013

Emory, S. A., & Belasco, J. G. (1990). The ompA 5' untranslated RNA segment functions in Escherichia coli as a growth-rate-regulated mRNA stabilizer whose activity is unrelated to translational efficiency. *Journal of Bacteriology*, *172*(8), 4472–4481. https://doi.org/10.1128/jb.172.8.4472-4481.1990

Emory, S. A., Bouvet, P., & Belasco, J. G. (1992). A 5'-terminal stem-loop structure can stabilize mRNA in Escherichia coli. *Genes & Development*, *6*(1), 135–148. https://doi.org/10.1101/gad.6.1.135

Gomez, M., Doukhan, L., Nair, G., & Smith, I. (1998). sigA is an essential gene in Mycobacterium smegmatis. *Molecular Microbiology*, *29*(2), 617–628. https://doi.org/10.1046/j.1365-2958.1998.00960.x

Heck, C., Rothfuchs, R., Jäger, A., Rauhut, R., & Klug, G. (1996). Effect of the pufQ-pufB intercistronic region on puf mRNA stability in Rhodobacter capsulatus. *Molecular Microbiology*, *20*(6), 1165–1178. https://doi.org/10.1111/j.1365-2958.1996.tb02637.x

Hu, Y., & Coates, A. R. M. (1999). Transcription of Two Sigma 70 Homologue Genes, sigA and sigB, in Stationary-Phase Mycobacterium tuberculosis. *Journal of Bacteriology*, *181*(2), 469–476.

Huff, J., Czyz, A., Landick, R., & Niederweis, M. (2010). Taking phage integration to the next level as a genetic tool for mycobacteria. *Gene*, *468*(1), 8–19. https://doi.org/10.1016/j.gene.2010.07.012

Hui, M. P., Foley, P. L., & Belasco, J. G. (2014). Messenger RNA Degradation in Bacterial Cells.

Annual Review of Genetics, *48*, 537–559. https://doi.org/10.1146/annurev-genet-120213-092340

Liu, M. Y., & Romeo, T. (1997). The global regulator CsrA of Escherichia coli is a specific mRNA-binding protein. *Journal of Bacteriology*, *179*(14), 4639–4642.

Mackie, G. A. (2013). RNase E: at the interface of bacterial RNA processing and decay. *Nature Reviews Microbiology*, *11*(1), 45–57. https://doi.org/10.1038/nrmicro2930

Main_text_21Sept2018_v1.1.pdf. (n.d.). Retrieved from http://www.who.int/tb/publications/global_report/Main_text_21Sept2018_v1.1.pdf

Manganelli, R., Proveddi, R., Rodrigue, S., Beaucher, J., Gaudreau, L., & Smith, I. (2004). σ Factors and Global Gene Regulation in Mycobacterium tuberculosis. *Journal of Bacteriology*, *186*(4), 895–902. https://doi.org/10.1128/JB.186.4.895-902.2004

Moll, I., Afonyushkin, T., Vytvytska, O., Kaberdin, V. R., & Bläsi, U. (2003). Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA*, *9*(11), 1308–1314. https://doi.org/10.1261/rna.5850703

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. https://doi.org/10.1128/JB.01657-07

Nilsson, G., Belasco, J. G., Cohen, S. N., & Gabain, A. von. (1984). Growth-rate dependent regulation of mRNA stability in Escherichia coli. *Nature*, *312*(5989), 75–77. https://doi.org/10.1038/312075a0

Nouaille, S., Mondeil, S., Finoux, A.-L., Moulis, C., Girbal, L., & Cocaign-Bousquet, M. (2017). The stability of an mRNA is influenced by its concentration: a potential physical mechanism to regulate gene expression. *Nucleic Acids Research*, *45*(20), 11711–11724. https://doi.org/10.1093/nar/gkx781

Redon, E., Loubière, P., & Cocaign-Bousquet, M. (2005). Role of mRNA Stability during

Genome-wide Adaptation of Lactococcus lactis to Carbon Starvation. *Journal of Biological Chemistry*, *280*(43), 36380–36385. https://doi.org/10.1074/jbc.M506006200

Russell, D. G., Barry, C. E., & Flynn, J. L. (2010). Tuberculosis: What We Don't Know Can, and Does, Hurt Us. *Science (New York, N.Y.)*, *328*(5980), 852–856. https://doi.org/10.1126/science.1184784

Rustad, T. R., Minch, K. J., Brabant, W., Winkler, J. K., Reiss, D. J., Baliga, N. S., & Sherman, D. R. (2013). Global analysis of mRNA stability in Mycobacterium tuberculosis. *Nucleic Acids Research*, *41*(1), 509–517. https://doi.org/10.1093/nar/gks1019

Steyn, A. J. C., Collins, D. M., Hondalus, M. K., Jacobs, W. R., Kawakami, R. P., & Bloom, B. R. (2002). Mycobacterium tuberculosis WhiB3 interacts with RpoV to affect host survival but is dispensable for in vivo growth. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(5), 3147–3152. https://doi.org/10.1073/pnas.052705399

Unniraman, S., Prakash, R., & Nagaraja, V. (2001). Alternate Paradigm for Intrinsic Transcription Termination in Eubacteria. *Journal of Biological Chemistry*, *276*(45), 41850–41855. https://doi.org/10.1074/jbc.M106252200

Volpe, E., Cappelli, G., Grassi, M., Martino, A., Serafino, A., Colizzi, V., … Mariani, F. (2006). Gene expression profiling of human macrophages at late time of infection with Mycobacterium tuberculosis. *Immunology*, *118*(4), 449–460. https://doi.org/10.1111/j.1365-2567.2006.02378.x

Waagmeester, A., Thompson, J., & Reyrat, J.-M. (2005). Identifying sigma factors in Mycobacterium smegmatis by comparative genomic analysis. *Trends in Microbiology*, *13*(11), 505–509. https://doi.org/10.1016/j.tim.2005.08.009

Wu, S., Barnes, P. F., Samten, B., Pang, X., Rodrigue, S., Ghanny, S., … Howard, S. T. (2009). Activation of the eis gene in a W-Beijing strain of Mycobacterium tuberculosis correlates with increased SigA levels and enhanced intracellular growth. *Microbiology*, *155*(Pt 4),

1272–1281. https://doi.org/10.1099/mic.0.024638-0

Wu, S., Howard, S. T., Lakey, D. L., Kipnis, A., Samten, B., Safi, H., … Barnes, P. F. (2004). The principal sigma factor sigA mediates enhanced growth of Mycobacterium tuberculosis in vivo. *Molecular Microbiology*, *51*(6), 1551–1562. https://doi.org/10.1111/j.1365-2958.2003.03922.x