Distribution Agreement

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this dissertation. I retain all ownership rights to the copyright of the dissertation. I also retain the right to use in future works (such as articles or books) all or part of this dissertation.

Signature:

_____ _____

Colleen Reynoso Date

Development and Application of Synthetic Riboswitches as Tools to Study Bacterial Pathogenesis

By
Colleen Reynoso
Doctor of Philosophy

Chemistry

_____
Justin P. Gallivan, Ph.D.
Advisor

_____
Vincent P. Conticello, Ph.D.
Committee Member

_____
Dale E. Edmondson, Ph.D.
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the Graduate School

_____
Date

Development and Application of Synthetic Riboswitches as Tools to Study Bacterial Pathogenesis

By

Colleen Reynoso
B.S. summa cum laude, University of West Georgia, 2006

Advisor: Justin P. Gallivan, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the Graduate School of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Chemistry
2012

Abstract

Development and Application of Synthetic Riboswitches as Tools to Study Bacterial Pathogenesis

By Colleen Reynoso

In the laboratory, we require tools that allow conditional regulation of gene expression to facilitate the study of genes with unknown or little understood function. The ideal tool would provide translation of external stimuli to an internal alteration in gene expression for any gene to be studied. Furthermore, these tools should be transferrable between different species of bacteria to allow manipulation of expression in less understood prokaryotic species. Nature has the advantage of billions of years of evolution to determine efficient mechanisms for gene regulation. We have looked to nature for inspiration and adapted natural systems to serve alternative purposes in the laboratory. In this thesis we describe the development of synthetic riboswitches as orthogonal tools to conditionally express bacterial genes of interest. We will explore the concept of isolating functional riboswitches from a genetically tractable species of bacterial and transporting these tools with limited alteration to less genetically tractable species. Each chapter focuses on a different bacterial species with relevance to different areas of microbiology and biotechnology and presents what we have learned regarding the portability of synthetic riboswitches.

Development and Application of Synthetic Riboswitches as Tools to Study Bacterial Pathogenesis

By

Colleen Reynoso
B.S. summa cum laude, University of West Georgia, 2006

Advisor: Justin P. Gallivan

A dissertation submitted to the Faculty of the Graduate School of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Chemistry
2011

Acknowledgments

I would like to offer my gratitude to my advisor, Professor Justin Gallivan. He gave me the push I needed to keep going even when it looked like the end. I will forever be thankful to him that he gave me the chance to become the scientist I am today. I also extend my sincere thanks to my committee members, Professors Vince Conticello and Dale Edmondson, who granted support when needed and much needed critique as well. My deepest gratitude also goes to Professors June Scott and David Weiss who granted me time and space to perform my work. They also gave me guidance when I was delving into completely new territory. I thank the many other professors and teachers who have nurtured my love for science from the beginning until now.

I also have to thank my lab mates and colleagues from the Gallivan lab as well as the Weiss and Scott labs. You have offered advice, listened to frustrations and generally made grad school fun. Much thanks goes to Shana Topp who believed in me even when I did not. Many more thanks to those who have read and re-read and read all over again my many drafts over the years. Thank you Joy Sinha, Dennis Mishler and Adrianne Edwards. I appreciate your willingness to endure so much torture. Thanks to Daniel Stabley who helped me pretend I knew what I was doing as far as microscopy is concerned. I also send my thanks and congratulations to Crystal Jones who was in the same boat as me and by now will be Dr. Jones. Thanks as well to Dr. Julia Bugrysheva and Dr. Barb Froehlich for much guidance. Additionally, I would like to extend my thanks to the Chemistry Department staff members who keep everything running. I would especially like to thank Ann Dasher and the amazing stockroom staff for hard work and dedication and for solving so many problems day after day. Thank you to my

fellow graduate students and undergraduates and postdocs who have touched my life. Without each and every one of you I know grad school would not have been nearly as enjoyable. I cannot list everyone, because I am sure I would forget someone. I have always been terrible with names. I must give a very special thanks to my wonderful husband Wilson Reynoso, who did not have it easy having a wife in graduate school. There were many late nights alone. There were many nights waiting for me and there were plenty of fast food dinners that were eaten and chores to be shared. Thank you for being a remarkably patient, understanding and helpful man. I also have to thank my family and my parents most of all. You have always supported but never pushed. You let me find my own way and guided me when I felt lost and you always nurtured my thirst for knowledge. I also dedicate this work to my grandmother, who raised me and loved me. She passed just before my defense, but I know she would be proud. Most of all I must thank God. He granted me guidance, strength and determination. Everyday I see the evidence of His hand in my work and I thank Him for his blessings.

Contents

List of Figures

List of Tables

Chapter 1 – Introduction

1.1 Genetic Control in Response to Environmental Stimuli

An important requirement for survival is the ability of an organism to sense its environment and be able to respond and adapt to environmental changes. Nature has evolved many efficient methods of receiving environmental stimuli to affect a corresponding outcome.[1] This is one manner by which an organism can activate biochemical pathways necessary for the given environment and thereby avoid unnecessary energy expenditure.[2] It is especially essential for small, single-celled organisms with limited energy stores to be able to efficiently balance internal metabolite concentrations to maintain optimal functionality without wasting energy.[3] For example, in the bacterium *Neisseria gonorrhoeae* when concentrations of iron (Fe), essential for microbial growth, drop below a certain threshold, gene expression is altered to allow the cell to harvest iron from the surrounding environment. However, once Fe concentrations return to optimal levels, expression of iron-harvesting proteins is turned off to prevent unnecessary energy consumption and prevent iron concentrations from reaching toxic levels.[4]

In some cases, this regulatory role is the responsibility of proteins that have evolved the ability to sense and bind to important metabolites. This binding event can either directly affect gene expression, or trigger a cascade of signals that results in a change of expression for genes involved in the synthesis or degradation of the binding metabolite.[5] The *lac* operon encodes proteins necessary for metabolism (*lacZ*) and import (*lacY*) of lactose as well as an upstream repressor protein *lacI* that senses the presence of lactose and effects expression of the metabolizing proteins. In the absence of

lactose the repressor freely binds to the operator portion of the operon and prevents transcription of the downstream proteins. When lactose is present in high concentrations an isomer of lactose, allolactose, forms and binds to the repressor causing a conformational change that prevents the repressor from binding to the operator site.[6] With transcription no longer inhibited, the cell can transcribe and translate the genes necessary for lactose import and for metabolizing lactose to glucose, the preferred energy source. Because glucose is the preferred energy source, there also exists a regulatory cascading pathway by which activation of lactose metabolizing enzymes is inhibited by the presence of sufficient glucose concentrations. Cyclic AMP (cAMP) is a prevalent signal molecule in prokaryotic cells. Within the *lac* operon, a cAMP binding site aids in association of RNA polymerase (RNAP) to the *lac* promoter. When glucose is present in high concentrations the concentration of cyclic AMP (cAMP) decreases and binding of RNAP is decreased, resulting in repressed expression of *lacZ* and *lacY*.[7]

The interaction of the *lac* operon with allolactose is an example of a genetic inducer, however, nature has also developed repression mechanisms as exhibited by the *trp* operon which inhibits tryptophan-producing enzymes when concentrations of the metabolite are high.[8] This operon encodes five genes necessary in the biosynthesis of the amino acid tryptophan as well as a repressor (*trpR*) and an operator. Like the *lac* operon, the *trp* repressor binds to a metabolite (tryptophan) to elicit a conformational change in the repressor. However, instead of inactivating the repressor upon ligand binding, as demonstrated by the *lac* repressor, the *trp* repressor is activated by binding tryptophan. It then binds to the operator to halt transcription. In addition to this mechanism, the *trp* operon also harbors a *trp* attenuator leader sequence (*trpL*) of approximately 160

nucleotides found upstream of the operator.[9] This leader sequence codes for a small peptide containing several tryptophan repeats.[10] When the concentration of tryptophan in the cell is high is a proportionately high concentration of tryptophan-charged tRNAs available for translation of the leader sequence. However, when supplies of tryptophan are low, availability of tryptophan-tRNA is also low, causing the ribosome to stall on the attenuator sequence. Stalling of the ribosome allows secondary structure of the mRNA of the attenuator to form a transcriptional terminator, thereby halting transcription of the *trp* operon.[11]

1.2 Sensing the Environment with RNA-Small Molecule Interactions: Evidence of the RNA World

The mechanism of the mRNA attenuator alludes to the ability of mRNA structural conformation to affect gene expression and is reminiscent of a form of gene regulation thought to be the predecessor of protein-mediated genetic control. Riboswitches are non-coding regions of mRNA, usually in the 5' untranslated region (5'-UTR), that can control expression of downstream genes through direct binding of a specific ligand without the necessity of protein cofactors.[12] Riboswitches consist of two functional domains, an aptamer domain that binds to a specific ligand and an expression platform that affects gene expression through conformational change.[13] Because of their prevalence across multiple phyla[14], their simple mechanism of action as well as their non-reliance on proteins, they are thought to be remnants of the RNA world.[15] The RNA world hypothesis states that RNA may pre-date DNA and proteins as life molecules.[16]

Riboswitches, like many regulatory proteins, bind to metabolites crucial for normal cell function. Remarkably, although mRNA possesses only 4 possible canonical

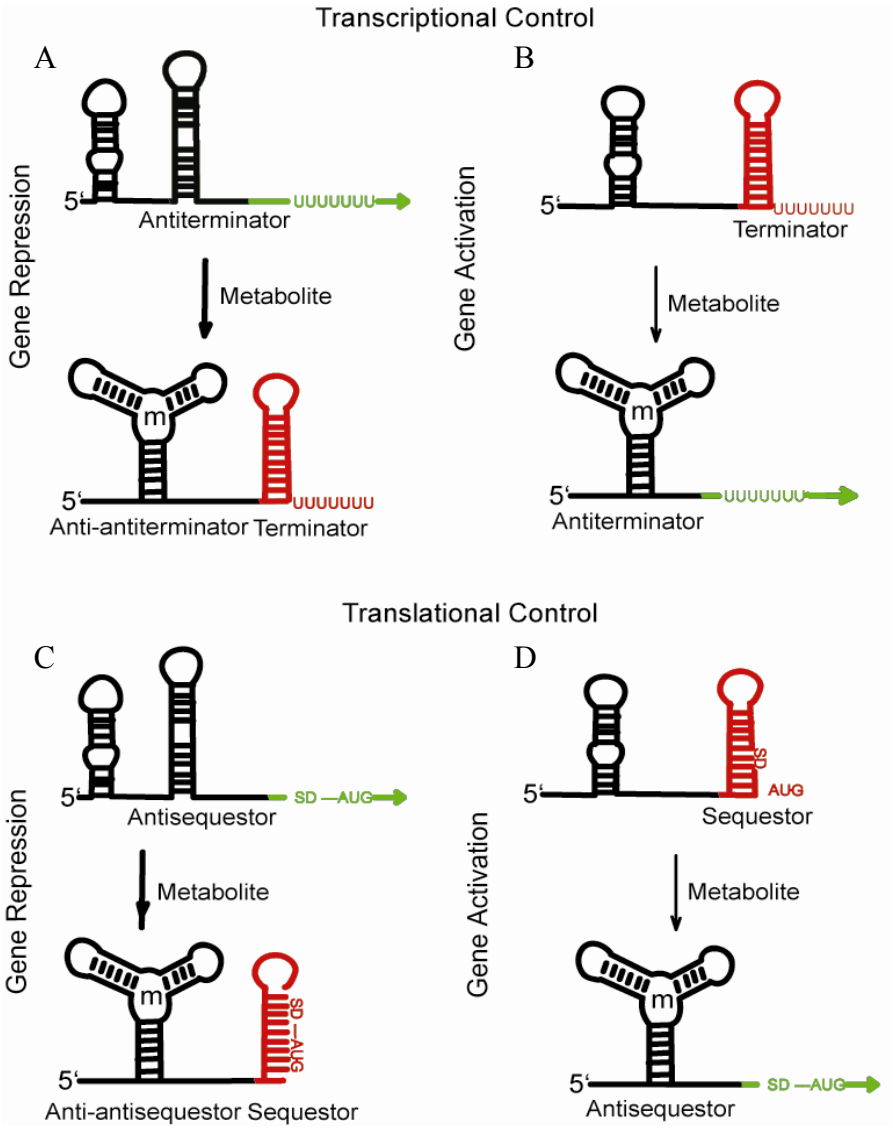

Figure 1.1 - Mechanisms of Natural Prokaryotic Riboswitches. Prokaryotic riboswitches follow 4 typical modes of genetic control. Ligand binding to the aptamer domain causes a conformational change in the expression platform that either (A) forms a terminator, halting transcription, (B) destabilizes a terminator, turning transcription on, (C) sequesters the ribosome binding site (RBS), halting translation or (D) reveals the RBS allowing translation initiation. (Adapted from Nudler and Mironov, 2004.[12])

identities per base (versus 20 per amino acid position in proteins) the mRNA aptamers of riboswitches are able to bind tightly and selectively to a wide variety of molecules including amino acids and their derivatives, nucleobases and their derivatives and metal

[12] Figure adapted from Nudler, E. and Mironov, A. S. *Trends Biochem. Sci.* **2004**, 29, 11-17 with permission from Elsevier.

ions.[17] There are four primary mechanisms by which riboswitches control gene expression. All mechanisms involve binding of a specific ligand and a subsequent conformational change (Figure 1.1). Riboswitches that activate transcription, such as the *Bacillus subtilis* adenine-sensitive riboswitch, adopt a terminator sequence in the absence of ligand, but upon binding of adenine the terminator secondary structure does not form due to secondary structure modulation, thereby allowing transcription of a purine efflux pump (Figure 1.1 B).[18] The guanine riboswitch of *B. subtilis*, on the other hand, presents the opposite mechanism, wherein the absence of the ligand guanine does not hinder transcription of purine biosynthesis genes. However, when levels of guanine accumulate inside the bacterial cell, production of guanine is no longer necessary. Guanine binds to the riboswitch, stabilizing a terminator conformation and deactivating transcription of purine biosynthesis genes (Figure 1.1 A).[19] For translational riboswitches the presence of ligand either stabilizes sequestering of the ribosome binding site (RBS), halting translation as demonstrated by the flavin mononucleotide (FMN) riboswitch in *E. coli* (Figure 1.1 C),[12] or stabilizes the unsequestered RBS conformation allowing translation (Figure 1.1 D). Since their initial discovery, over 15 variants of natural riboswitches have been found.[13, 20]

1.3 Engineered Gene Regulation

The study of genes and their functions has benefitted from the use of genetic control mechanisms. Many of these control elements began as natural systems that were adapted as genetic tools. Several regulatory promoters from *E. coli* have been modified to work in a broad range of bacterial hosts.[21, 22] One such promoter is the P_{BAD} promoter from the *E. coli* L-arabinose catabolic gene. The *ara* operon consists of three regulatory

binding sites, the regulatory *araC* gene and three catabolic genes *araB*, *araA* and *araD*.[23] The AraC protein regulates transcription of itself as well as the catabolic triad of genes.[24] AraC is only expressed when concentrations of AraC are low. In cases of high AraC concentration it binds to a regulatory region of the *ara* operon called the $araO_I$ site, an event that prevents transcription of the *araC* gene.[25] Additionally, AraC represses the transcription of arabinose-specific genes during conditions of no arabinose through formation of a dimer that binds to two positions, the *araO2* site and the *araI* site. Interestingly, unlike the other regulatory systems previously mentioned here, AraC not only functions as an inhibitor, but also as an initiator.[26] When arabinose is present, the AraC dimer changes conformation to bind only the *araI* site, effectively allowing transcription of *araB*, *araA* and *araD*.[26] The P_{BAD} promoter, as well as the araC gene coding for the regulatory AraC protein, have been isolated and fused with genes of interest to aid in the study of those genes both in *E. coli* [27, 28] and other Gram-negative bacteria.[27, 29] There also exists a Gram-positive equivalent *ara* operon from *B. subtilis*.[30]

While use of the *ara* operon has garnered important information, the system is still induced by arabinose, which is naturally metabolized and could cause pleiotropic effects on the bacterial cell. The *lac* promoter and repressor, like the *ara* system, have also been ectopically expressed and used as an inducible system in several bacterial species such as *E. coli* and *B. subtilis*.[22] However, since the native inducer, allolactose, is metabolized by bacteria and can interfere with natural biochemical pathways, an analogous inducer has been employed for orthogonal control. This inducer is Isopropyl β-D-1-thiogalactopyranoside (IPTG) and is not metabolized within bacterial cells, making it an ideal alternative to lactose/allolactose.[31]

However useful it may be to induce gene expression using inducer analogs, the inducing metabolites are limited to a natural or analog ligand. This can severely limit the number of genetic tools available.[32] Artificial molecular switches that control the function of any protein of interest using a small molecule specified by the researcher would allow dynamic, orthogonal control of gene expression.[33] There are several approaches to engineering molecular protein switches. One method is to alter the binding pocket of a pre-existing regulatory protein to recognize an orthogonal molecule.[32] An allosteric site could also be introduced to a protein that could be engineered to affect enzyme activity.[33] Additionally, binding domains can be added to both the ligand and protein that bind to a mediator molecule to force localization in the presence of the mediator molecule causing an alteration the active state of a protein.[33] No matter what the approach, engineering of protein switches must meet certain requirements. The engineered regulator must bind both the target (whether gene or protein) and the new ligand tightly and selectively, and allosteric binding of the effector ligand must translate to an altered functional outcome for the target. In short, a molecular protein switch must be able to link small molecule binding with alteration of protein function.[33] As an example of protein switch engineering, a maltose binding protein (MBP) was engineered to alter ligand specificity from maltose to sucrose through an iterative evolution process.[32] Briefly, MBP was associated to a protein conferring ampicillin resistance and the binding pocket of MBP was randomized. The library was screened in the presence of sucrose, but the absence of maltose and surviving clones were cultured from selective media.[32] While the method was successful in changing the specificity of MBP from maltose to sucrose, the former ligand and new ligand are still very closely related. In

addition, this method requires the existence of a pre-existing scaffold to act as a template. However, what if a switch recognizing a completely novel ligand was desired and no pre-cursor proteins exist?

Unlike natural riboswitches, synthetic riboswitches do not require binding to a natural metabolite of bacteria, which allows for orthogonal genetic regulation that minimizes interruption of normal cellular function. The development of a synthetic

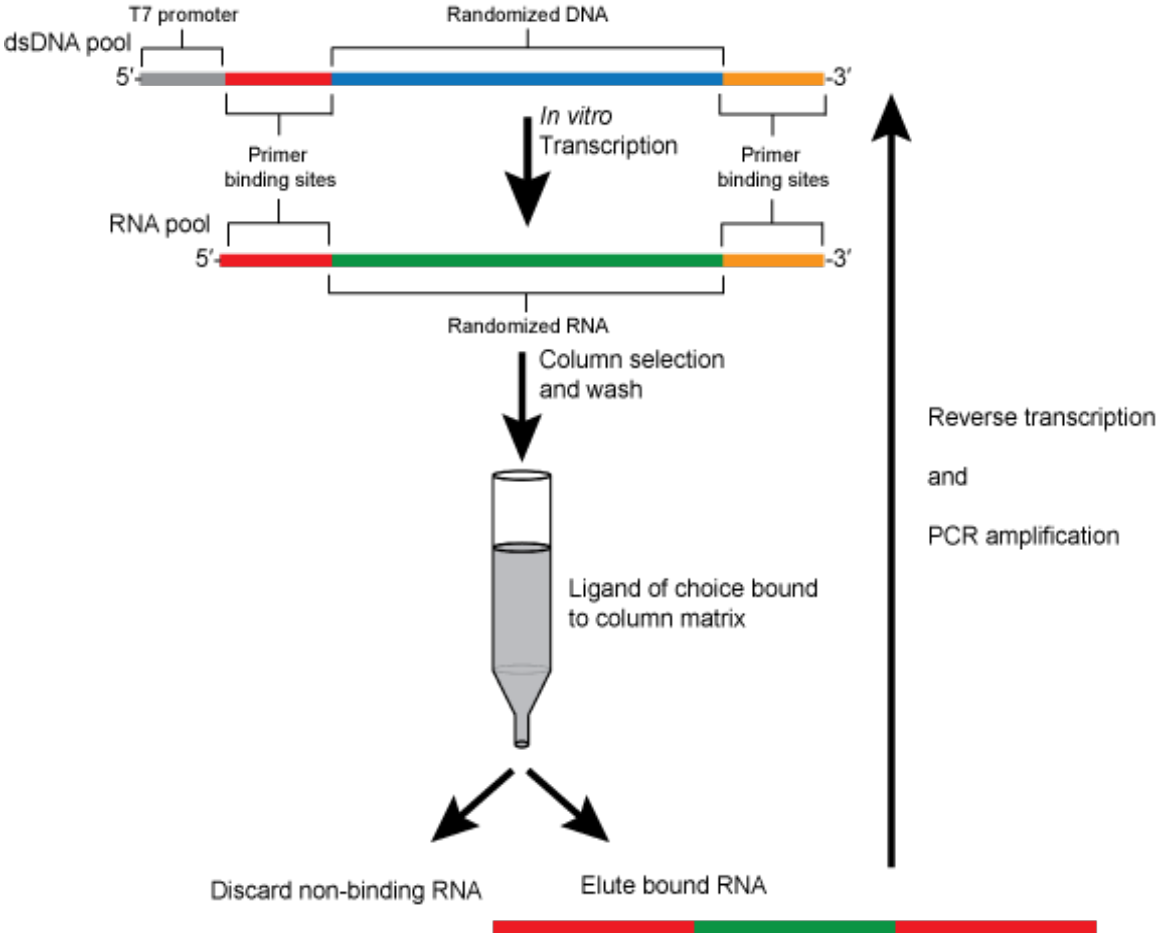

Figure 1.2 – SELEX Aptamer Selection Procedure. A pool of randomized DNA (approximately 100 base pairs) is reverse transcribed *in vitro*. The resulting RNA pool is applied to a column containing immobilized ligand. The column is washed and non-binding RNAs are discarded. Bound RNA is eluted with a high concentration of ligand. Eluted RNA is reverse transcribed to obtain a cDNA library that is PCR amplified. The procedure is repeated until significant enrichment is achieved.

riboswitch is, theoretically, less complicated than engineering a protein switch. When engineering a synthetic riboswitch, the requirements are that the mRNA binds tightly to the ligand in a selective manner and that the binding event causes a conformational change that directly affects gene expression. While it is possible to engineer new riboswitches that respond to novel ligands starting from a pre-existing natural riboswitch[34], it is also possible to develop synthetic riboswitches *de novo*. As mentioned previously, there are two main parts to a riboswitch, the first of which is an aptamer, which binds to a specific target molecule. In principle, using a process called SELEX (Systematic Evolution of Ligands by EXponential enrichment)[35] an RNA aptamer that recognizes any target molecule can be found,[36, 37] with no need for a pre-existing template, as is the case in protein engineering. This iterative process (Figure 1.2) begins with a chemically synthesized pool of randomized DNA flanked by a defined 3' and 5' region. Initiating this procedure with completely randomized sequences (excepting constant primer binding sites) aids in analyzing a large, unbiased representation of sequence space with sample sizes on the order of $\sim 10^{15}$. These sequences are then transcribed to RNA *in vitro*, with the aid of a T7 promoter placed upstream of the randomized sequence. The resulting RNA pool is applied to an affinity column derivatized with the desired ligand. Elimination of non-specific binders can be achieved through a negative selection step, wherein the random pool is washed over a column without derivatized ligand. Any non-binding RNAs are collected then applied to the ligand-bound column. RNA molecules exhibiting affinity for the ligand will associate with the ligand immobilized on the column. Those that do not bind are discarded. All steps are performed *in vitro*, eliminating restraints on library complexity due to

transformation efficiency.[37] The RNAs are then reverse transcribed to obtain the corresponding cDNA. These are re-amplified by PCR, transcribed and passed through the affinity column again. This is repeated until significant enrichment is achieved (see Figure 4), typically 5-10 rounds or more.[37]

To ensure selected aptamers are specific to the desired molecule, counter selection steps may be added to the procedure to eliminate aptamers to structurally related molecules. In this additional step the column is washed with high concentrations of a structurally related molecule. Any aptamers recognizing the ligand analog will dissociate from the column and be discarded with the flow-through. After removal of non-discriminatory aptamers, the column is once again washed with a high concentration of target ligand to elute a pool of aptamers enriched with high-specificity binders.

In 1994, four years after the development of SELEX, Jenison and colleagues reported the isolation of an RNA aptamer highly selective for the small molecule theophylline.[38] Natural riboswitch aptamers are highly specific and exhibit a K_d range on the order of ~5 nM to ~200 μM.[19, 39-41] The mTCT8-4 theophylline aptamer binds tightly to theophylline with a K_d of 100 nM, which is comparable to natural aptamers. The theophylline aptamer is also extremely selective, exhibiting an affinity 10,000 times greater for theophylline over caffeine, a structurally similar xanthine molecule differing by only a single methyl group at the N-7 position.[38]

Werstuck and Green hypothesized that the binding of a small molecule to the 5'-UTR of mRNA could alter gene expression through a change in secondary structure.[42] Previous reports demonstrated that an increase in secondary structure upstream of the RBS correlated to a marked decrease in translation initiation.[43] With those observations

in mind they theorized that addition of the aptamer's target molecule would cause an increase in secondary structure of the mRNA leading to a decrease in translation. Indeed, they were able to repress expression of reporter genes in both *E. coli* and Chinese hamster ovary (CHO) cells in a ligand-dependent dose-responsive manner.[42] This was the first reported instance of engineering a genetic switch from an RNA aptamer that binds to a non-endogenous molecule.

Subsequent work has led to the use of RNA-small molecule interactions to regulate gene expression in response to several classes of molecules, including dyes,[42, 44] antibiotics,[42, 45, 46] and other small molecules.[47, 48] While simply placing an aptamer upstream of an expression platform has led to the discovery of functional riboswitches[42, 48] the development of more robust synthetic riboswitches can be achieved through screening of library sequences.[48-50]

Once of the first high-throughput methods to engineer novel riboswitches was proposed by our group in 2004. We performed a proof-of-principle experiment where response to a desired ligand was linked to bacterial cell survival.[48] We placed the theophylline aptamer[38] and a C27A mutant aptamer (known to change the specificity of the theophylline aptamer to that of a 3-methylxanthine aptamer)[48, 51] upstream of a chloramphenicol acetyl transferase gene (conferring chloramphenicol resistance). Plasmids harboring the parent riboswitch were diluted with plasmids containing the C27A mutated riboswitch in varying ratios. Even in the most dilute case ($1:10^6$ parent:C27A) we were able to isolate the theophylline sensitive riboswitch when transformed cultures were plated onto chloramphenicol supplemented agar media.[48]

We later developed a screen analyzing β-galactosidase activity to select for

functional riboswitches from a randomized library.[52] More information on this method will be provided in Chapter 3. We also developed a genetic selection to identify functional riboswitches using cell motility.[53] In this method the theophylline aptamer, followed by a randomized region, is cloned upstream of the *cheZ* gene that enables cell motility in *E. coli* cells.[54] If a member of the randomized library acts as a functional riboswitch, the presence of theophylline will activate *cheZ* and allow cell motility. In the absence of theophylline, the cell containing a functional riboswitch would remain non-motile.[53] Both negative and positive selections are used to find clones that are non-motile in the absence of theophylline but motile in the presence of theophylline. One disadvantage of this selection system is inherent to CheZ. If CheZ is not present in sufficient amounts, the cells will be non-motile, as expected. However, if CheZ is present in very high concentrations cells can become non-motile due to becoming embedded in the agar media.[55] Therefore, this selection method would theoretically preclude the discovery of strongly activating riboswitches.

To remedy this drawback we wished to develop a screen that would permit the discovery of riboswitches with robust activation coupled with low basal expression in the absence of ligand.[49] In this method we search for dynamic, functional riboswitches using fluorescence activated cell sorting (FACS), which can analyze large randomized libraries with sizes on the order of ~10^8 members.[56] To identify robust theophylline sensitive riboswitches we placed the theophylline aptamer upstream of a randomized region that included the RBS followed the gene for a red fluorescent protein, DsRed. By including the RBS in the library we allowed the selection to optimize the RBS as well as the spacing between the RBS and start codon. To date, the riboswitch reported from this

screen is the most robust of the natural or synthetic riboswitches published.[49]

Other labs have also developed screens and selections for synthetic riboswitches. Weigand and Suess developed a screen in which a pool of aptamers, as opposed to a single aptamer, for the ligand neomycin was placed upstream of a *gfp* reporter gene. They then performed a selection in the yeast *Saccharomyces cerevisiae* for the expression of GFP in the absence of neomycin. The colonies exhibiting GFP expression in the absence of neomycin were then screened for neomycin-dependent inhibition of GFP expression.[45] Fowler and Li used FACS to screen for transcriptional riboswitches in *E. coli,* similar to the method of Lynch and Gallivan described above.[57] Weiland and Hartig randomized a region between the theophylline aptamer and a ribozyme that self-cleaves downstream mRNA. Screening of this library led to the discovery of constructs that activate gene expression 10-fold.[58]

Over 50 years has passed since Jacob and Monod first discovered that a small molecule can effect gene expression. Since then many developments in synthetic and molecular biology have allowed the use of natural and synthetic genetic control mechanisms to regulate gene expression in model organisms. The focus of this thesis will be the identification and development of riboswitches in various species of bacteria. While *E. coli* and *B. subtilis* are the most common model species for Gram-negative and Gram-positive prokaryotes respectively, there are many other medically relevant species that are as yet genetically intractable. After proving that riboswitches can indeed be used across a broad spectrum of prokaryotic species, we further prove the utility of orthogonal synthetic riboswitches for use in studying important genes in pathogenic bacteria.

1.4 References

(1) Boor, K. J. *PLoS Biol.* **2006**, 4, e23.

(2) Stoebel, D. M., Dean, A. M. and Dykhuizen, D. E. *Genetics*. **2008**, 178, 1653-1660.

(3) DeBusk, R. *Nutr. Clin. Pract.* **2010**, 25, 627-633.

(4) McKenna, W. R., Mickelsen, P. A., Sparling, P. F. and Dyer, D. W. *Infect. Immun.* **1988**, 56, 785-791.

(5) Jacob, F. and Monod, J. *J. Mol. Biol.* **1961**, 3, 318-356.

(6) Jobe, A. and Bourgeois, S. *J. Mol. Biol.* **1972**, 69, 397-408.

(7) Pastan, I. and Perlman, R. L. *Proc. Natl. Acad. Sci.* **1968**, 61, 1336-1342.

(8) Monod, J. and Cohen-Bazire, G. *C. R. Hebd. Seances Acad. Sci.* **1953**, 236, 530-532.

(9) Jackson, E. N. and Yaetofsky, C. *J. Mol. Biol.* **1973**, 76, 89-101.

(10) Morse, D. E. and Morse, A. N. C. *J. Mol. Biol.* **1976**, 103, 209-226.

(11) Lee, F. and Yanofsky, C. *Proc. Natl. Acad. Sci.* **1977**, 74, 4365-4369.

(12) Nudler, E. and Mironov, A. S. *Trends Biochem. Sci.* **2004**, 29, 11-17.

(13) Winkler, W. C. and Breaker, R. R. *Annu. Rev. Microbiol.* **2005**, 59, 487-517.

(14) Abreu-Goodger, C., Ontiveros-Palacios, N., Ciria, R. and Merino, E. *Trends Genet.* **2004**, 20, 475-479.

(15) Sudarsan, N., Barrick, J. E. and Breaker, R. R. *RNA*. **2003**, 9, 644-647.

(16) Joyce, G. F. *Nature*. **2002**, 418, 214-221.

(17) Bastet, L., Dubé, A., Massé, E. and Lafontaine, D. A. *Mol. Microbiol.* **2011**, 80, 1148-1154.

(18) Lemay, J.-F. o., Penedo, J. C., Tremblay, R., Lilley, D. M. J. and Lafontaine, D. A. *Chem. Biol.* **2006**, 13, 857-868.

(19) Mandal, M., Boese, B., Barrick, J. E., Winkler, W. C. and Breaker, R. R. *Cell*. **2003**, 113, 577-586.

(20) Barrick, J. E. and Breaker, R. R. *Genome Biol.* **2007**, 8, R239.

(21) Davison, J. *BioTechniques*. **2002**, 32, 386-388, 390, 392-384, passim.

(22) Yansura, D. G. and Henner, D. J. *Proc. Natl. Acad. Sci.* **1984**, 81, 439-443.

(23) Helling, R. B. and Weinberg, R. *Genetics*. **1963**, 48, 1397-1410.

(24) Doyle, M. E., Brown, C., Hogg, R. W. and Helling, R. B. *J. Bacteriol.* **1972**, 110, 56-65.

(25) Lee, D. H., Huo, L. and Schleif, R. *J. Mol. Biol.* **1992**, 224, 335-341.

(26) Englesberg, E., Irr, J., Power, J. and Lee, N. *J. Bacteriol.* **1965**, 90, 946-957.

(27) Sukchawalit, R., Vattanaviboon, P., Sallabhan, R. and Mongkolsuk, S. *FEMS Microbiol. Lett.* **1999**, 181, 217-223.

(28) Guzman, L. M., Belin, D., Carson, M. J. and Beckwith, J. *J. Bacteriol.* **1995**, 177, 4121-4130.

(29) Newman, J. R. and Fuqua, C. *Gene*. **1999**, 227, 197-203.

(30) Sá-Nogueira, I., Nogueira, T. V., Soares, S. and de Lencastre, H. *Microbiology*. **1997**, 143, 957-969.

(31) Hansen, L. H., Knudsen, S. and Sørensen, S. J. *Curr. Microbiol.* **1998**, 36, 341-347.

(32) Guntas, G., Mansell, T. J., Kim, J. R. and Ostermeier, M. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, 102, 11224-11229.

(33) Buskirk, A. R. and Liu, D. R. *Chem. Biol.* **2005**, 12, 151-161.

(34) Dixon, N., Duncan, J. N., Geerlings, T., Dunstan, M. S., McCarthy, J. E., Leys, D. and Micklefield, J. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, 107, 2830-2835.

(35) Tuerk, C. and Gold, L. *Science*. **1990**, 249, 505-510.

(36) Ellington, A. D., Jack W. Szostak. *Nature*. **1990**, 346, 818-822.

(37) Wilson, D. S. and Szostak, J. W. *Annu. Rev. Biochem.* **1999**, 68, 611-647.

(38) Jenison, R. D., Gill, S. C., Pardi, A. and Polisky, B. *Science*. **1994**, 263, 1425-1429.

(39) Wickiser, J. K., Winkler, W. C., Breaker, R. R. and Crothers, D. M. *Mol. Cell*. **2005**, 18, 49-60.

(40) Winkler, W. C., Nahvi, A., Sudarsan, N., Barrick, J. E. and Breaker, R. R. *Nat. Struct. Mol. Biol.* **2003**, 10, 701-707.

(41) Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A. and Breaker, R. R. *Nature*. **2004**, 428, 281-286.

(42) Werstuck, G. and Green, M. R. *Science*. **1998**, 282, 296-298.

(43) de Smit, M. H. and van Duin, J. *Proc. Natl. Acad. Sci.* **1990**, 87, 7668-7672.

(44) Grate, D. and Wilson, C. *Bioorg. Med. Chem.* **2001**, 9, 2565-2570.

(45) Weigand, J. E., Sanchez, M., Gunnesch, E. B., Zeiher, S., Schroeder, R. and Suess, B. *RNA*. **2008**, 14, 89-97.

(46) Hanson, S., Berthelot, K., Fink, B., McCarthy, J. E. and Suess, B. *Mol. Microbiol.* **2003**, 49, 1627-1637.

(47) Harvey, I., Garneau, P. and Pelletier, J. *RNA*. **2002**, 8, 452-463.

(48) Desai, S. K. and Gallivan, J. P. *J. Am. Chem. Soc.* **2004**, 126, 13247-13254.

(49) Lynch, S. A. and Gallivan, J. P. *Nucleic Acids Res.* **2009**, 37, 184-192.

(50) Weigand, J. E., Sanchez, M., Gunnesch, E.-B., Zeiher, S., Schroeder, R. and Suess, B. *RNA*. **2008**, 14, 89-97.

(51) Soukup, G. A., Emilsson, G. A. M. and Breaker, R. R. *J. Mol. Biol.* **2000**, 298, 623-632.

(52) Lynch, S. A., Desai, S. K., Sajja, H. K. and Gallivan, J. P. *Chem. Biol.* **2007**, 14, 173-184.

(53) Topp, S. and Gallivan, J. P. *ChemBioChem*. **2008**, 9, 210-213.

(54) Wolfe, A. J., Conley, M. P., Kramer, T. J. and Berg, H. C. *J. Bacteriol.* **1987**, 169, 1878-1885.

(55) Wolfe, A. J. and Berg, H. C. *Proc. Natl. Acad. Sci.* **1989**, 86, 6973-6977.

(56) Link, A. J., Jeong, K. J. and Georgiou, G. *Nat Rev Micro*. **2007**, 5, 680-688.

(57) Fowler, C. C., Brown, E. D. and Li, Y. *ChemBioChem*. **2008**, 9, 1906-1911.

(58) Wieland, M. and Hartig, J. S. *Angewandte Chemie International Edition*. **2008**, 47, 2604-2607.

Chapter 2 – Observing Assembly of a Magnetosome protein in *Magnetospirillum magneticum* Using Synthetic Riboswitches

2.1 Introduction

Model organisms aid in the elucidation of biochemical mechanisms, genetics and evolution. This biological information is often transferable to related organisms.[1] The ideal model organism possesses short generation times, relatively easy genetic manipulation, small genomes and low cost.[1] For prokaryotic organisms *E. coli* and *B. subtilis* are the most established models for Gram-negative and Gram-positive bacteria respectively.[2-4] While model organisms can provide insight into basic gene function, they cannot provide detailed information on genes unique to other bacterial species. Unfortunately, many of these non-model organisms are also less understood and have fewer tools available for genetic study. Though an array of genetic manipulation devices are available for model organisms such as *E. coli* or *B. subtilis*, these tools are often not portable to other species.[2-4] Transport of these systems from one organism to another has been reported,[5, 6] however, challenges including promoter compatibility, codon usage and protein folding prevent widespread portability from one species to another. The development of a genetic manipulation system that functions independent of protein cofactors and a can be driven by native promoters would circumvent these obstacles and facilitate closer examination of genes exclusive to less genetically tractable organisms. Translational synthetic riboswitches offer these advantages with the additional benefit of operating orthogonally to native biochemical pathways through induction by a non-endogenous small molecule.

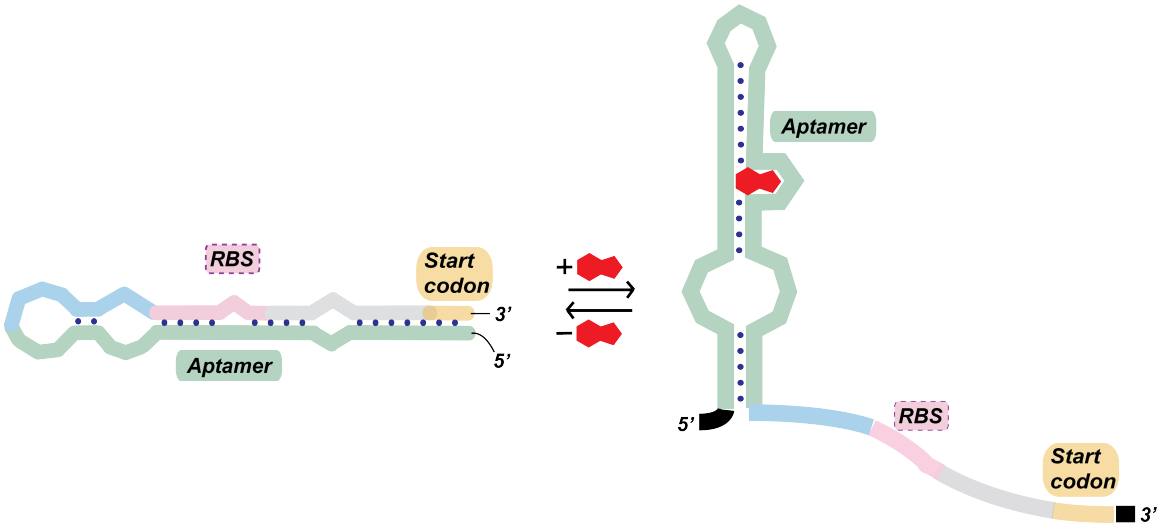

Figure 2.1 – Mechanism of translational theophylline-sensitive synthetic riboswitch. (Left) The aptamer (green) is shown forming a hairpin structure through interactions with the expression platform of the mRNA (pink to yellow region). The RBS (pink) is sequestered in this conformation and it not available for ribosome binding, thereby preventing translation. (Right) In the presence of theophylline, a conformation in stabilized wherein the RBS is no longer sequestered and is available for ribosome binding.

For these reasons, we hypothesized that our synthetic riboswitches isolated from *E. coli* would be easily portable to other organisms. Our lab has developed a number of theophylline-sensitive synthetic riboswitches[7-10] that regulate gene expression at the translational level by sequestering the Shine-Dalgarno sequence, commonly know as the ribosome binding site (RBS). When theophylline is present, a conformation is stabilized that presents the RBS, allowing binding of the ribosome (Figure 2.1).[8] Because of well-conserved translational machinery across bacterial phylogeny, this mechanism should be compatible with most, if not all, species of bacteria.[11] Additionally, the translational mechanism operates independently of promoter and can therefore be paired with native host promoters.[12]

We selected the aquatic Gram-negative α-proteobacterium *Magnetospirillum magneticum*[13] to test this new hypothesis of riboswitch portability. At the commencement of this study *M. magneticum* had no reported tools for the conditional manipulation of gene expression. *M. magneticum* is magnetotactic and orients itself to the geomagnetic field using iron magnetite (Fe_3O_4) crystals within prokaryotic organelles called magnetosomes (Figure 2.2).[13-16] Magnetosomes form chains along the vertical axis of the bacterium forming what is essentially a compass needle.[17] The mode of locomotion for *Magnetospirillum* is not true magnetotaxis. *Magnetospirillum spp.* are passively oriented along geomagnetic lines (by the internal compass) but do not travel up or down a magnetic gradient. They are instead propelled by flagella either parallel or anti-parallel to magnetic field lines[18] with direction governed by decreasing concentrations of oxygen, a process called aerotaxis.[19] Experiments have shown that magnetotactic bacteria, when constrained to thin tubing, will preferentially travel down oxygen gradients to the ideal microaerobic environment at the oxic-anoxic barrier.[20]

The magnetosomes, while governing interesting behavior, are important for several other reasons. Organelles, once thought to be an exclusively eukaryotic attribute, have been described in several species of bacteria.[21-23] Magnetosomes are one of the most well known prokaryotic organelles and share many common features of eukaryotic organelles. These organelles originate as invaginations of the outer cell membrane and are thought to become independent vesicles (though this is under debate)[24] aligned by a network of cytoskeletal actin-like filaments.[17] The magnetosome lipid bilayer encompasses a single magnetite crystal and possesses a set of soluble and transmembrane proteins unique from the cell membrane.[25-27] Studies have revealed that magnetosomes

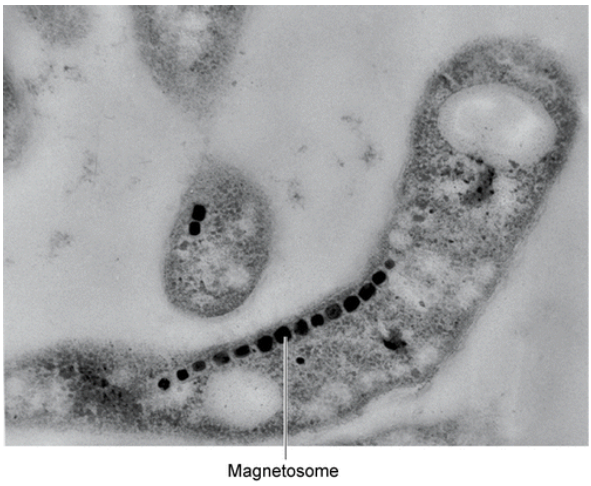

Figure 2.2 – Magnetosomes within *M. magneticum*. Transmission Electron Microscopy image of an ultrathin section of *M. magneticum* AMB-1 grown under iron-rich conditions. Magnetosomes are seen as black dots lined up along the axis of the bacterial cell. Reprinted from Komeili et al. **2007**,[17].

form independently of the formation of magnetite crystals.[28] These and other results suggest that assembly of a complete magnetosome requires invagination of the outer membrane, import of appropriate proteins and iron to these vesicles, and signals to begin magnetite biogenesis. These cell biological characteristics make the magnetosome an ideal model for prokaryotic organelle assembly and development.[28]

Though the genes necessary for the structure and function of magnetosomes have been identified, the roles those genes play are still not well understood.[29] It is known that the actin-like cytoskeletal protein MamK is necessary for alignment of the magnetosomes within the bacterial cell.[17] However, prior to alignment of magnetosomes it is hypothesized that another protein, MamA, is responsible for activating the invaginated portions of the cell membrane to form mature magnetosomes.[28] Inducible expression of these genes could aid in elucidation of the specific mechanisms of function.

The process by which *M. magneticum* forms magnetite crystals is also not well understood. Magnetite biogenesis is of particular interest for commercial purposes. The magnetite crystals formed by *M. magneticum* are highly pure and identical in size and

[17] Figure reprinted from Komeili, A. *Annu. Rev. Biochem.* **2007**, 76, 351-366. with permission from Annual Reviews.

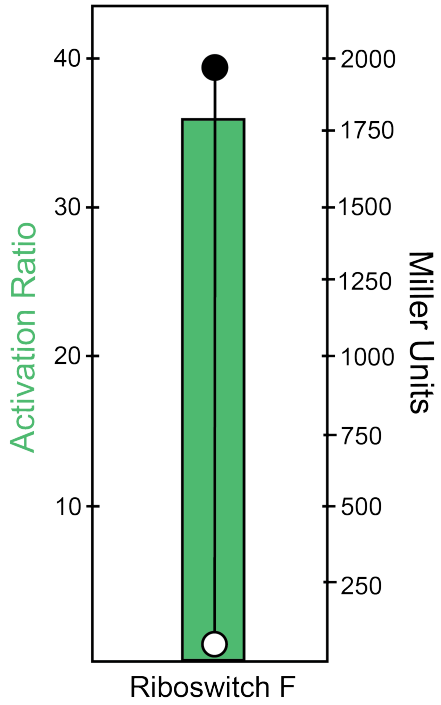

Figure 2.3 – Measure of reporter gene activity induced by riboswitch F in *E. coli*. Strains encoding the indicated riboswitches controlling *gusA* were grown in the presence (●) or absence (○) of 1 mM theophylline to mid-log phase. Cells were harvested, lysed and assayed for β-galactosidase activity. The left axis indicates the Miller Units, a measure of β-galactosidase activity, and the right axis shows the activation ratio (green bar) which is calculated by dividing the Miller units in the presense of theophylline by the units in the absense of theophylline. Standard error of miller units lies within the area of circles.

morphology.[18] Manufacturing processes to produce nanocrystals of the same purity and uniformity require high temperature and pressure, whereas *M. magneticum* produces those same crystals under mild conditions.[30] Bacterially-produced magnetite nanocrystals have been used in mRNA recovery,[31] as DNA carriers for bacterial transformation[32] and as scaffolds for display of other biomolecules.[33] The size of the crystals is putatively regulated by a group of proteins, MamGFDC.[34] When mutants deficient of these proteins were grown in the presence of iron, magnetosomes of irregular shape and size were formed.[34] Wild type crystals were restored upon *trans*-complementation of any combination of three of these genes. Interestingly, when all four genes are complemented, crystals larger than wild type form.[34] Also involved in crystal formation is a putative self-aggregating protein Mms6 shown to catalyze the formation of uniform magnetite crystals when added *in vitro*, however, its effects *in vivo* have yet to be determined.[35] The ability to conditionally express the genes coding for these proteins

in vivo would aid in probing the functions and mechanisms of MamGFDC and Mms6 and provide insight into the biogenesis of magnetite.

3.2 Results and Discussion

To transport our riboswitches to *M. magneticum* we used a *Magnetospirillum*-*E. coli* shuttle vector, pAK22,[17] based on the broad-host range vector pBBR1MCS-2[36] featuring a P_{tac}[37] promoter and kanamycin resistance marker. We cloned our riboswitches upstream of a *mamK-gfp* C-terminal translational fusion. While there is published data suggesting that ferric iron quenches GFP signal,[38] previous reports of GFP fusions to magnetosome proteins demonstrated that iron concentrations ranging from 0 to 10 mM do not noticeably affect the fluorescence of GFP within magnetotactic bacteria.[39] GFP was used to visualize the linear assembly of MamK, an actin-like cytoskeletal protein thought to be responsible for the alignment of magnetosomes within the *Magnetospirillum* cell.[17] Mutants lacking *mamK* exhibit magnetosomes scattered throughout the cytosol and constitutive expression of MamK restores alignment of the magnetosomes.[17] However, due to a lack of reported expression systems for the induction of gene expression in *M. magneticum*, the *mamK-gfp* fusion could not be conditionally expressed in order to observe gradual phenotypic changes as magnetosomes organize.

A previously described riboswitch[8] (riboswitch F), exhibiting an activation ratio in *E. coli* of 35-fold (Figure 2.3), was introduced to an *M. magneticum* AMB-1 Δ*mamK* strain by conjugation[40] and selected for DNA transfer by growth in media supplemented with kanamycin. Cells grown in the absence of theophylline overnight were supplemented with 1 mM theophylline and imaged at several time points post induction

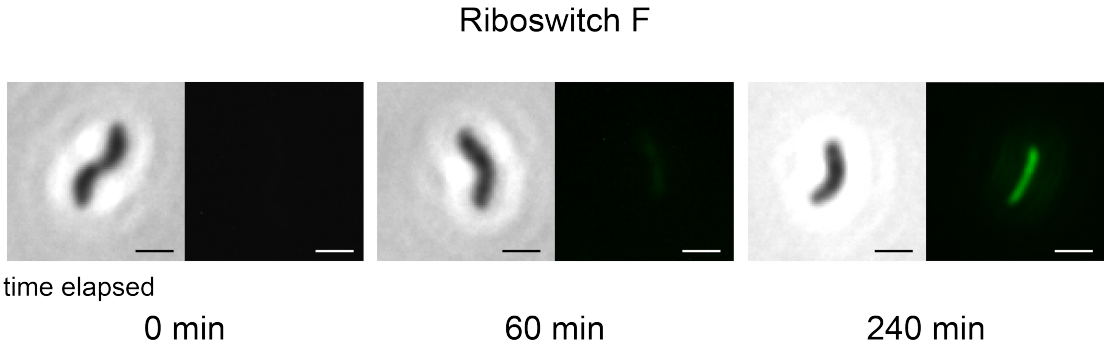

Figure 2.4 – Riboswitch-mediated induction of MamK-GFP. *M. magneticum* cells harboring a P_{tac} promoter driving transcription of riboswitch F upstream of mamK-gfp were grown in the absence of theophylline overnight. Theophylline (1mM) was added to the culture and imaging by epifluorescence was performed at time 0, 2 hours post-induction and 4 hours post-induction. The phase contrast (black and white) and fluorescent images are both presented.

to observe the expression of the MamK-GFP fusion. No fluorescence was detected from 0-40 minutes post induction. As early as 45 minutes, faint fluorescence was observed with the expected pattern of MamK-GFP assembly.[17] By 2 hours post induction the pattern observed was indistinguishable from an overnight induction (data not shown). Epifluorescence images of 0, 2 and 4 hours post induction are shown in Figure 2.4 for riboswitch F. The images clearly demonstrate the appearance of GFP fluorescence induced over time. Magnetosomes align with the vertical axis of *M. magneticum* cells and are stabilized by networks of MamK.[17] The fluorescence pattern observed in Figure 2.4 mimics this vertical alignment. Notably, the image at 4 hours post induction appears to be comparable to images obtained from constitutive expression of MamK-GFP.[17, 41]

2.3 Conclusions

These encouraging results led our lab to import other riboswitches into *M. magneticum* with positive results.[12] As will be discussed in later chapters, we have

found that RBS strength and spacing from the start codon have effects on the

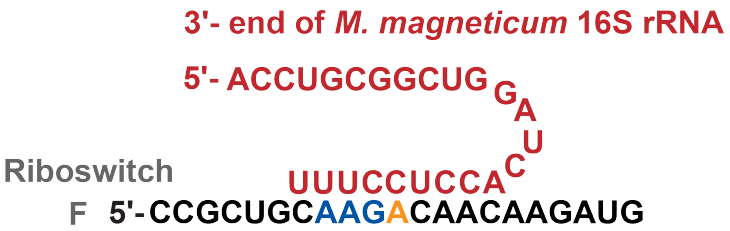

Figure 2.5 – Alignment of riboswitch F RBS with 16S rRNA of *M. magneticum*. The 3' end of the 16S rRNA of *M. magneticum* is shown in red. The RBS of riboswitch F is shown in blue with mismatched bases in gold.

functionality of our synthetic riboswitches. We have also found that riboswitches exhibit higher basal expression in the absence of theophylline when operating at 30 °C (the growth temperature of *M. magneticum*) versus 37 °C (SI of Ref #12).[12] The RBS of riboswitch F carries only 3-base complementarity to the anti-Shine-Dalgarno sequence (Figure 2.5), which resides at the 3' end of the 16S rRNA and is responsible for ribosomal recognition of the RBS on mRNA. Studies have shown that a single mismatch in complementarity to the anti-Shine-Dalgarno can adversely affect gene expression.[42] We believe that the weak RBS of riboswitch F aids in counteracting the effects of increased basal expression caused by operation at 30 °C.

This initial confirmation of our hypothesis regarding the portability of synthetic riboswitches directed us to further explore the possibility of introducing synthetic riboswitches into less genetically tractable species of bacteria. In the following chapters we will discuss our successes and failures in this endeavor and apply what we have learned to the development of synthetic riboswitches as an orthogonal tool for the conditional expression of bacterial genes.

2.4 Experimental

General Considerations. Synthetic oligonucleotides were purchased from IDT (Coralville, IA). All sequences were verified using DNA sequencing (MWG Operon, Huntsville, AL or in-house sequencing). Purifications of plasmid DNA, PCR products, and enzyme digestions were performed using kits from Qiagen (Germantown, MD). All enzymes were purchased from NEB (Ipswich, MA). Reagents, such as theophylline and kanamycin were purchased from Sigma-Aldrich (St. Louis, MO). All cloning was performed in *E. coli* TOP10 F' cells (Invitrogen, Carlsbad, CA). All *E. coli* cultures were in LB media (EMD Bioscience, Merck KGaA, Darmstadt, Germany). *M. magneticum* strain AMB-1 experiments were performed by Dr. Arash Komeili or Dr. Dorothée Murat in the laboratory of Dr. Arash Komeili at the University of California, Berkeley. Cultures of *M. magneticum* were grown in MG medium supplemented (per liter) with 5 mL Wolfe's mineral solution (prepared without iron), 0.68 g potassium phosphate, 0.12 g sodium nitrate, 0.07 g sodium acetate, 0.035 g ascorbic acid, 0.37 g tartaric acid, 0.37 g succinic acid and 0.05 g sodium thiosulfate (pH 6.9). Following autoclaving Wolfe's vitamin solution was added to 1X final concentration. Iron-rich cultures were supplemented with 30-50 μM ferric malate.

Genetic manipulations. The C-terminal GFP fusions to *mamK* were produced as previously reported.[17] The *Magnetospirillum-E. coli* shuttle plasmid, pAK22, was a generous gift from Dr. Arash Komeili (University of California, Berkeley). Riboswitches were amplified from previously described plasmids[8] and assembled by overlap PCR to the mamK-gfp fusion amplified from plasmid pAK22. Vector and insert were digested with EcoRI and XhoI and ligated with T4 DNA ligase after dephosphorylation of vector with Calf intestinal phosphatase (CIP). Plasmid DNA was introduced to *M. magneticum*

by conjugation as previously described[40] and selected for plasmid uptake with kanamycin (10 μg/mL).

Fluorescence microscopy.

The strains were grown at 30 ºC in MG liquid media supplemented with 10 μg/mL kanamycin under microaerobic conditions (<10% oxygen). A 24-hour culture in exponential phase (OD_{400} = 0.1) was diluted into 10 mL cultures at an initial OD_{400} of 0.05 in fresh media in the absence or presence of 1 mM theophylline. Cultures were grown in <10% oxygen and 100 μL of cells were spun down at different time points after inoculation (0, 1 and 4 hours). Aliquots were spotted on agarose pads prepared with 1% agarose in growth media, and imaged under phase contrast and fluorescence microscopy as previously described[41] using a Nikon Eclipse 80i microscope equipped with a Qimaging RETIGA 2000R Fast 1394 camera. All experiments in *M. magneticum* were performed by Dr. Arash Komeili or Dr. Dorothée Murat.

2.5 References

(1) Hedges, S. B. *Nat. Rev. Genet.* **2002**, 3, 838-849.

(2) Riley, M., Abe, T., Arnaud, M. B., Berlyn, M. K. B., Blattner, F. R., Chaudhuri, R. R., Glasner, J. D., Horiuchi, T., Keseler, I. M., Kosuge, T., Mori, H., Perna, N. T., Plunkett, G., Rudd, K. E., Serres, M. H., Thomas, G. H., Thomson, N. R., Wishart, D. and Wanner, B. L. *Nucleic Acids Res.* **2005**, 34, 1-9.

(3) Völker, U. and Hecker, M. *Cell. Microbiol.* **2005**, 7, 1077-1085.

(4) Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V., Bertero, M. G., Bessieres, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S. C., Bron, S., Brouillet, S., Bruschi, C. V., Caldwell, B., Capuano, V., Carter, N. M., Choi, S. K., Codani, J. J., Connerton, I. F., Cummings, N. J., Daniel, R. A., Denizot, F., Devine, K. M., Dusterhoft, A., Ehrlich, S. D., Emmerson, P. T., Entian, K. D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, N., Ghim, S. Y., Glaser, P., Goffeau, A., Golightly, E. J., Grandi, G., Guiseppi, G., Guy, B. J., Haga, K., Haiech, J., Harwood, C. R., Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M. F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Klaerr-Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P., Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S. M., Levine, A., Liu, H., Masuda, S., Mauel, C., Medigue, C., Medina, N., Mellado, R. P., Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S. H., Parro, V., Pohl, T. M., Portetelle, D., Porwollik, S., Prescott, A. M., Presecan, E., Pujic, P.,

Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M., Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadaie, Y., Sato, T., Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J., Sekowska, A., Seror, S. J., Serror, P., Shin, B. S., Soldo, B., Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K., Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpstra, P., Tognoni, A., Tosato, V., Uchiyama, S., Vandenbol, M., Vannier, F., Vassarotti, A., Viari, A., Wambutt, R., Wedler, E., Wedler, H., Weitzenegger, T., Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K., Yoshida, K., Yoshikawa, H. F., Zumstein, E., Yoshikawa, H. and Danchin, A. *Nature*. **1997**, 390, 249-256.

(5) Geissendörfer, M. and Hillen, W. *Appl. Microbiol. Biotechnol.* **1990**, 33, 657-663.

(6) Yansura, D. G. and Henner, D. J. *Proc. Natl. Acad. Sci.* **1984**, 81, 439-443.

(7) Desai, S. K. and Gallivan, J. P. *J. Am. Chem. Soc.* **2004**, 126, 13247-13254.

(8) Lynch, S. A., Desai, S. K., Sajja, H. K. and Gallivan, J. P. *Chem. Biol.* **2007**, 14, 173-184.

(9) Lynch, S. A. and Gallivan, J. P. *Nucleic Acids Res.* **2009**, 37, 184-192.

(10) Topp, S. and Gallivan, J. P. *ChemBioChem*. **2008**, 9, 210-213.

(11) Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B. S. and Stormo, G. *Annu. Rev. Microbiol.* **1981**, 35, 365-403.

(12) Topp, S., Reynoso, C. M. K., Seeliger, J. C., Goldlust, I. S., Desai, S. K., Murat, D., Shen, A., Puri, A. W., Komeili, A., Bertozzi, C. R., Scott, J. R. and Gallivan, J. P. *Appl. Environ. Microbiol.* **2010**, 76, 7881-7884.

(13) Frankel, R. B., Blakemore, R. P. and Wolfe, R. S. *Science*. **1979**, 203, 1355-1356.

(14) Sakaguchi, T., Burgess, J. G. and Matsunaga, T. *Nature*. **1993**, 365, 47-49.

(15) Mann, S., Sparks, N. H. C., Frankel, R. B., Bazylinski, D. A. and Jannasch, H. W. *Nature*. **1990**, 343, 258-261.

(16) Balkwill, D. L., Maratea, D. and Blakemore, R. P. *J. Bacteriol.* **1980**, 141, 1399-1408.

(17) Komeili, A., Li, Z., Newman, D. K. and Jensen, G. J. *Science*. **2006**, 311, 242-245.

(18) Bazylinski, D. A. and Frankel, R. B. *Nat. Rev. Micro.* **2004**, 2, 217-230.

(19) Taylor, B. L., Zhulin, I. B. and Johnson, M. S. *Annu. Rev. Microbiol.* **1999**, 53, 103-128.

(20) Frankel, R. B., Bazylinski, D. A., Johnson, M. S. and Taylor, B. L. *Biophys. J.* **1997**, 73, 994-1000.

(21) Seufferheld, M., Vieira, M. C. F., Ruiz, F. A., Rodrigues, C. O., Moreno, S. N. J. and Docampo, R. *J. Biol. Chem.* **2003**, 278, 29971-29978.

(22) Fuerst, J. A. *Annu. Rev. Microbiol.* **2005**, 59, 299-328.

(23) Zemer, G. *Cell*. **2005**, 120, 577-586.

(24) Komeili, A. *Annu. Rev. Biochem.* **2007**, 76, 351-366.

(25) Gorby, Y. A., Beveridge, T. J. and Blakemore, R. P. *J. Bacteriol.* **1988**, 170, 834-841.

(26) Okuda, Y., Denda, K. and Fukumori, Y. *Gene*. **1996**, 171, 99-102.

(27) Matsunaga, T., Tsujimura, N., Okamura, Y. and Takeyama, H. *Biochem. Biophys. Res. Commun.* **2000**, 268, 932-937.

(28) Komeili, A., Vali, H., Beveridge, T. J. and Newman, D. K. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, 101, 3839-3844.

(29) Jogler, C. and Schuler, D. *Annu. Rev. Microbiol.* **2009**, 63, 501-521.

(30) Matsunaga, T. and Sakaguchi, T. *J. Biosci. Bioeng.* **2000**, 90, 1-13.

(31) Sode, K., Kudo, S., Sakaguchi, T., Nakamura, N. and Matsunaga, T. *Biotechnol. Tech.* **1993**, 7, 688-694.

(32) Takeyama, H., Yamazawa, A., Nakamura, C. and Matsunaga, T. *Biotechnol. Tech.* **1995**, 9, 355-360.

(33) Yoshino, T., Maeda, Y. and Matsunaga, T. *Recent Pat. Biotechnol.* **2010**, 4, 214-225.

(34) Scheffel, A., Gärdes, A., Grünberg, K., Wanner, G. and Schüler, D. *J. Bacteriol.* **2008**, 190, 377-386.

(35) Arakaki, A., Webb, J. and Matsunaga, T. *J. Biol. Chem.* **2003**, 278, 8745-8750.

(36) Kovach, M. E., Phillips, R. W., Elzer, P. H., Roop, R. M., 2nd and Peterson, K. M. *BioTechniques*. **1994**, 16, 800-802.

(37) de Boer, H. A., Comstock, L. J. and Vasser, M. *Proc. Natl. Acad. Sci.* **1983**, 80, 21-25.

(38) Takeda, S., Kamiya, N. and Nagamune, T. *Anal. Biochem.* **2003**, 317, 116-119.

(39) Komeili, A., Vali, H., Beveridge, T. J. and Newman, D. K. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, 101, 3839-3844.

(40) Wahyudi, A. T., Takeyama, H. and Matsunaga, T. *Appl. Biochem. Biotechnol.* **2001**, 91-93, 147-154.

(41) Murat, D., Quinlan, A., Vali, H. and Komeili, A. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, 107, 5593-5598.

(42) Jacob, W. F., Santer, M. and Dahlberg, A. E. *Proc. Natl. Acad. Sci.* **1987**, 84, 4757-4761.

Chapter 3 - Screening for Synthetic Riboswitches in *Acinetobacter baylyi*

3.1 Introduction

With our initial success of transferring our synthetic riboswitches from *E. coli* to *M. magneticum* (Chapter 2), we applied the same principle to a species more closely related to *E. coli*. *Acinetobacter baylyi*, like *E. coli*, is a Gram-negative γ-proteobacteria. At the time of this study no inducible gene expression systems had been reported for *A. baylyi*, although one has been reported recently.[1] Natural riboswitches are found in many bacterial species,[2] including *A. baylyi,*[3] suggesting that synthetic riboswitches could function within the organism as well. Recognizing *A. baylyi* and *E. coli* are phylogenetically related Gram-negative bacteria, and encouraged by our success with *M. magneticum* (Chapter 2), we posited that synthetic riboswitches developed in *E. coli* could be directly transported to *A. baylyi* with little to no modification. However, preliminary experiments suggested that our hypothesis was incorrect.[4] We attempted to clone previously characterized riboswitches screened from *E. coli* directly to *A. baylyi* upstream of a *gusA* (β-glucuronidase) reporter gene, but found the transferred riboswitches functioned poorly.[4] It was later published that the origin of replication for the plasmid used (pWH1266[5]) seems to affect foreign promoter function.[1] However, since we were unaware of this knowledge at the time, we speculated that to find a functional riboswitch in *A. baylyi* we must screen for a functional riboswitch *de novo* from a randomized pool of candidates. Chapter 1 presented several approaches to screening for functional synthetic riboswitches. This chapter will focus on a previously reported method[6] applied to finding a theophylline-sensitive riboswitch for the organism *Acinetobacter baylyi*.

While *E. coli* and *A. baylyi* are closely related, *A. baylyi* (strain ADP1), unlike *E. coli*, is naturally competent, with the ability to uptake DNA (linear or plasmid) from the environment during the mid-exponential phase of growth.[7, 8] Cells can either uptake and maintain plasmid DNA or undergo homologous recombination to integrate DNA into a pre-existing plasmid or the genome.[9] Proponents for the use of *A. baylyi* as a model organism argue that it is more genetically tractable than *E. coli*, while growing on the same time scale with the same, or in fact, a wider range of growth conditions.[7, 10] At the time of this study, the main hindrance for *A. baylyi* being widely used as a model organism was the dearth of genetic tools available, save for a gene deletion library.[11] *A. baylyi* is also closely related to the organism *Acinetobacter baumannii*, an opportunistic human pathogen often responsible for multi-drug resistant infections among hospital patients.[12, 13] Developing tools for *A. baylyi* could prove useful in *A. baumannii* as well.

When initial attempts at transporting riboswitches from *E. coli* to *A. baylyi* failed,[4] we screened for functioning riboswitches in *A. baylyi* using a previously reported screen.[6] In this method, we placed the theophylline aptamer, mTCT8-4,[14] under transcriptional control of the P_{tac}[15] promoter upstream of an expression platform containing *gusA* (β-glucuronidase) preceded by a randomized linker region. We chose an eight base randomized library followed by a GGG-motif upstream of a conserved linker and the start codon of the expression platform (Figure 3.1). The randomization of eight nucleotide bases yields a theoretical library size of 65,000 members. This library size falls within the capacity of high-throughput screening methods, which can accommodate analysis of ~10^5 clones per day.[16] Our random library was screened for theophylline-

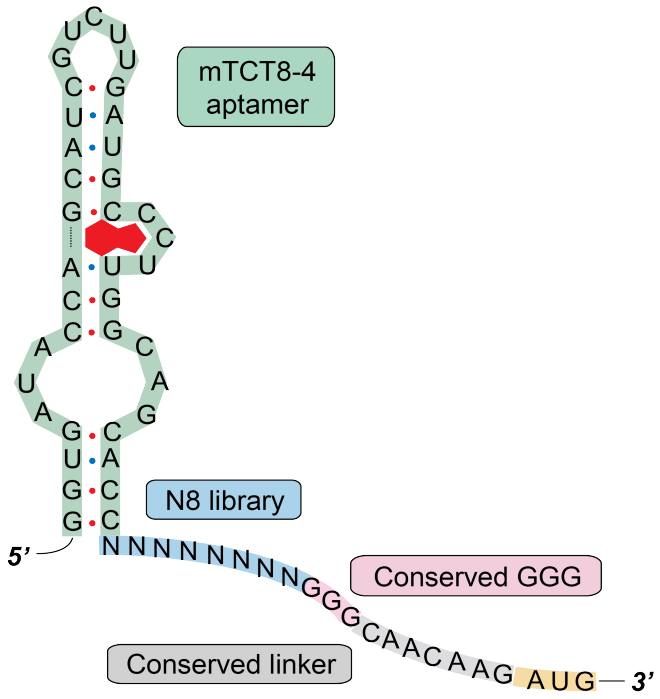

Figure 3.1 – Diagram of the location of the N8 library in the 5' UTR. The mTCT8-4 theophylline aptamer is shown in light green. The conserved GGG portion of the RBS is shown in pink. The start codon of *gusA* is shown in yellow and the N8 library is depicted in blue. Between the conserved GGG and start codon is a conserved linker found in most of our theophylline-induced riboswitches.

dependent gene expression using a blue-white screen. A blue-white screen utilizes the ability of β-glucuronidase to cleave the colorless compound, 5-bromo-4-chloro-indolyl-β-D-glucuronide (X-gluc), into 5-bromo-4-chloro-3-hydroxyindole, which spontaneously forms the blue compound, 5,5'-dibromo-4,4'-dichloro-indigo.[17,18] Therefore, if β-glucuronidase is present, x-gluc will be hydrolyzed to form blue colonies.[19, 20] Clones that were selected for ligand-dependent expression were isolated and further characterized.

3.2 Results and Discussion

Though we were previously unsuccessful in directly transporting our synthetic riboswitches from *E. coli* to *A. baylyi*, we still hypothesized that synthetic riboswitches

would function in *A. baylyi* if screened *de novo*. We constructed a library of eight randomized nucleotide bases upstream of a conserved linker sequence (Figure 3.1), producing a theoretical library size of 65,000 clones. Previously, we found that optimal spacing between the aptamer and RBS is 4-8 bases,[21] a fact that encouraged our design of an N8 library. The conserved linker sequence downstream of the randomized region was maintained due to experimental evidence suggesting its importance for function of our theophylline-inducible riboswitches. In fact, when the CAACAAG linker was mutated

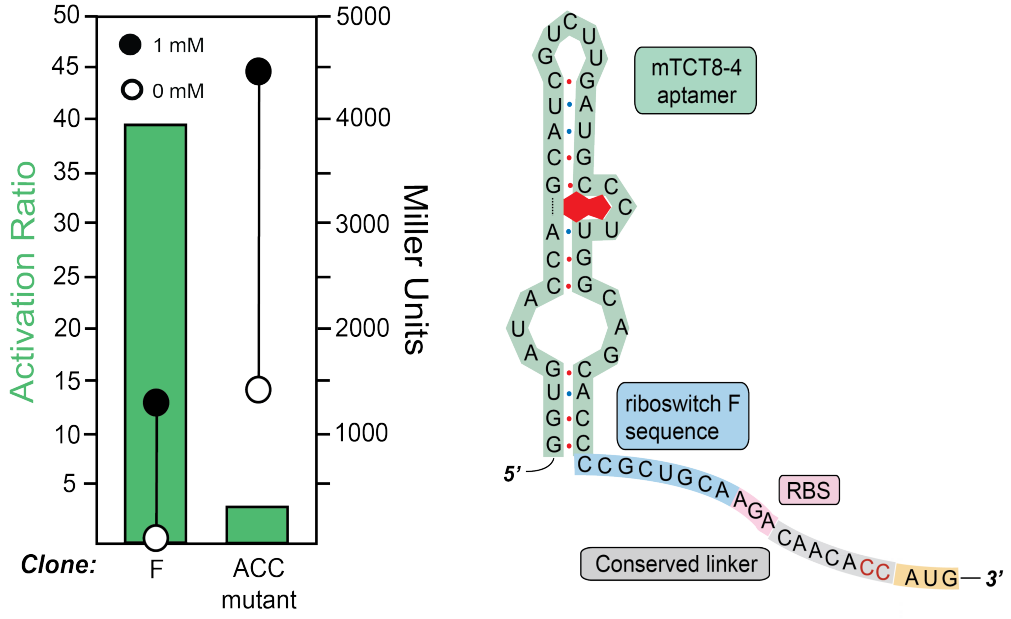

Figure 3.2 – Effect of ACC mutation on linker region of theophylline-inducible riboswitch.
(Left)Strains encoding the indicated riboswitches controlling *gusA* were grown in the presence (●) or absence (○) of 1 mM theophylline to mid-log phase. Cells were harvested, lysed and assayed for β-galactosidase activity. The right axis indicates the Miller Units, a measure of β-galactosidase activity, and the left axis shows the activation ratio (green bar) which is calculated by dividing the Miller units in the presense of theophylline by the units in the absense of theophylline. Error bars lie within the area of circles. (Right) The mTCT8-4 theophylline aptamer is shown in light green. The sequence of riboswitch F is shown in blue and the RBS in pink. The AG to CC mutation is denoted in red font.

to CAACACC, riboswitch F[6] exhibited an increase in the basal expression in the absence of theophylline increased. This increase in background expression resulted in the fold activation decreasing from 40-fold to only 5-fold (Figure 3.2). Indeed, later studies in our lab suggest that the CAACAAG linker may create a bias towards gene repression in the absence of theophylline.[22]

As previously mentioned, successful candidate riboswitches are determined using a blue-white screen. In a blue-white screen, we used X-gluc, which is a colorless compound. When β-glucuronidase is expressed the X-gluc molecule is hydrolyzed leading to the formation of a blue compound (Figure 3.3). In this screen we plate *A. baylyi* that has been transformed with the N8 GGG library onto selective media without ligand but containing X-gluc. Under these conditions, candidate riboswitches should not exhibit expression of β-glucuronidase. Therefore, clones are selected that are white to light blue. Any clones that are blue in the absence of ligand show high basal expression of β-glucuronidase in the absence of theophylline (always on) and are

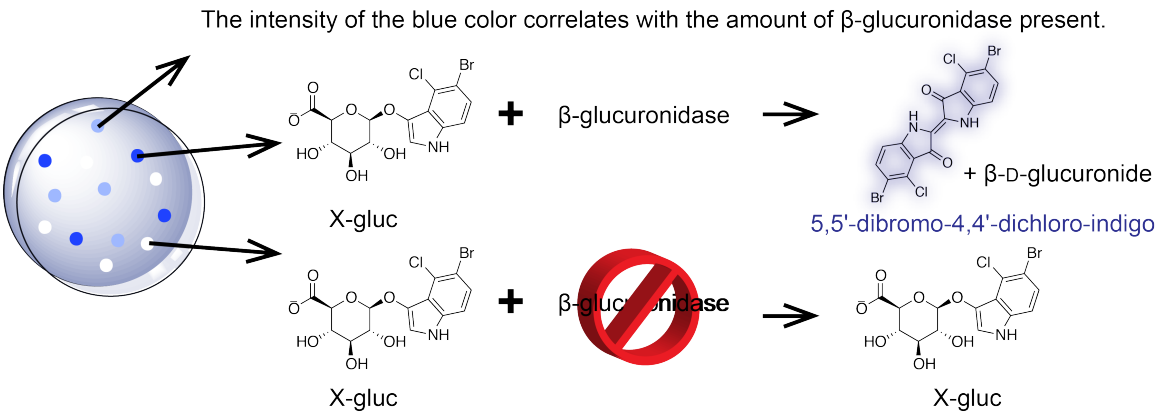

Figure 3.3 – Schematic of the blue-white screen. In the presence of β-glucuronidase, X-gluc will be hydrolyzed into glucuronic acid and 5-bromo-4-chloro-3*H*-indol-3-ol (not shown), which then dimerizes to form the blue compound 5,5'-dibromo-4,4'-dichloro-indigo (top). However, when no β-glucuronidase is present, X-gluc will remain intact and therefore produce no color.

therefore not ideal riboswitch candidates. In this initial selection process it is possible to select a candidate that is always off, which means that β-glucuronidase expression will not be induced even in the presence of theophylline. These clones will be eliminated in the next step of the screening process.

Clones that fulfill the requirements of the initial screen were cultured to perform a high-throughput β-glucuronidase assay adapted from Lynch et al.[6] Each library transformation produces 2,000-3,000 colonies. White to light blue colonies from this transformation were inoculated into a 96-well microtiter plate containing selective LB media without ligand. These cultures were grown overnight and subsequently used to inoculate two sets of fresh 96-well plate cultures: one containing 1 mM theophylline and the other containing no theophylline. These cultures were allowed to grow till mid log

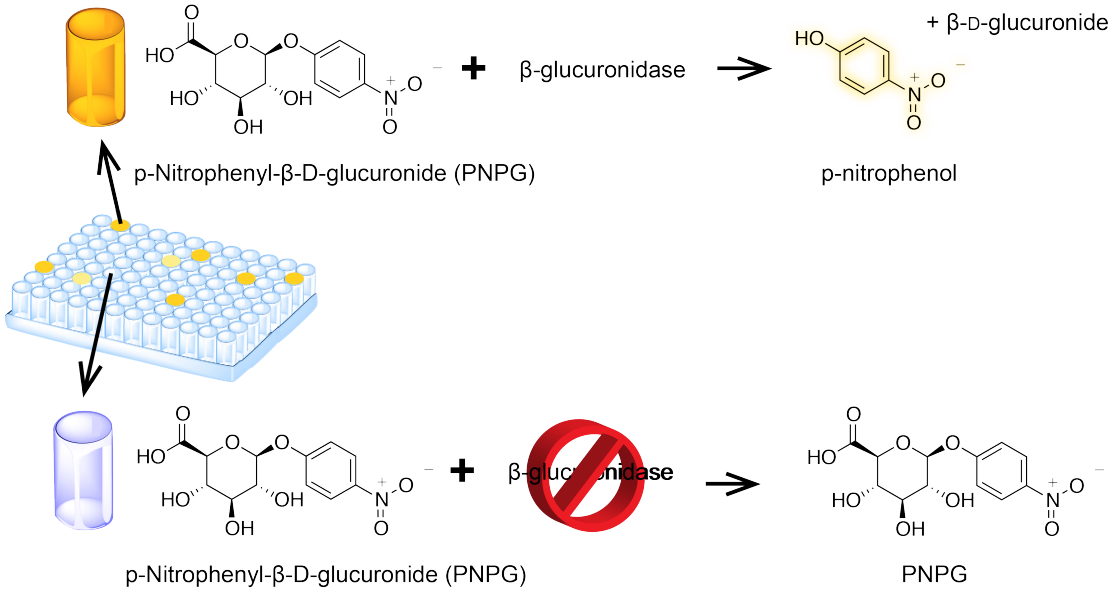

Figure 3.4 – Hydrolysis of PNPG by β-glucuronidase. PNPG (p-Nitrophenyl-β-D-glucuronide is a clear compound. However, in the presence of β-glucuronidase (top scheme) it is hydrolyzed to form a yellow compound p-nitrophenol. If there is no β-glucuronidase present, PNPG remains intact and colorless.

phase and assayed in a modified protocol of Miller's method.[17] Assays were performed using a multichannel pipettor that allows simultaneous addition of reagent to several wells. This assay relies on another substrate of β-glucuronidase, p-nitrophenyl-β-D-glucuronide, which is hydrolyzed to form a yellow compound in the presence of β-glucuronidase (Figure 3.4). Functional riboswitches are expected to form yellow solutions in the presence of theophylline due to induced expression of β-glucuronidase, but remain colorless in the absence of theophylline. Any clones exhibiting this behavior are selected and further characterized in triplicate.

After screening approximately 30,000 clones we identified seven variants with an activation ratio of 2 or greater. Of these variants, 5 of exhibited induction of 5-fold or greater. Figure 3.5 shows the Miller Units and Activation Ratios of the variants discovered from the N8 GGG library. The two best clones, N8.1 and N8.2 had been isolated from a previous N11 library that spurred the creation of the N8 GGG library.[4] The only novel variant that emerged from the N8 GGG library is clone N8.4, which exhibits a comparable activation ratio to clones N8.1 and N8.2 but does not exceed 500 Miller Units in the presence of theophylline, and therefore may not be useful for many applications. These switches were tested with the *lacZ* reporter gene and were found to exhibit similar attributes.[4]

Our initial observation that riboswitches could not be transported from *E. coli* to *A. baylyi* was later found to be erroneous. As mentioned previously, the origin of replication of the original plasmid (pWH1266[5]) used appears to have negative effects on promoter function.[1] We soon became aware of a broad-host range plasmid, pBAV1K found to be stable in a number of bacteria, including *A. baylyi* and *E. coli*.[23] Indeed, the

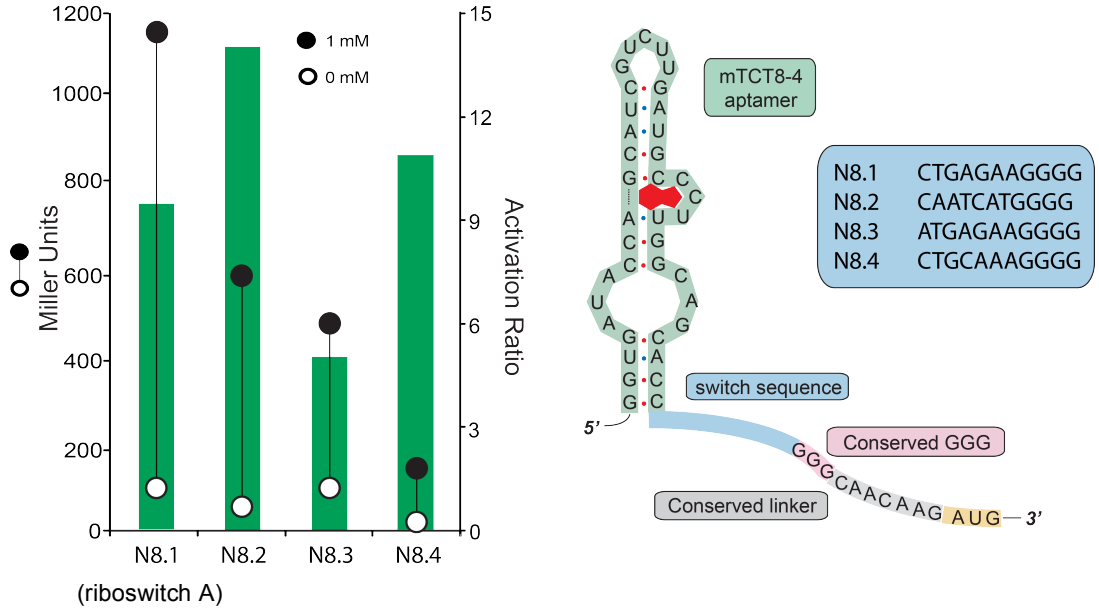

Figure 3.5 – β-Glucuronidase Activities for Riboswitches Identified from the *A. baylyi* library screens. (Left) Strains encoding the indicated riboswitches controlling *gusA* were grown in the presence (●) or absence (●) of 1 mM theophylline to mid-log phase. Cells were harvested, lysed and assayed for β-glucuronidase activity. The left axis indicates the Miller Units, a measure of β-glucuronidase activity, and the right axis shows the activation ratio (green bar) which is calculated by dividing the Miller units in the presense of theophylline by the units in the absense of theophylline. Standard deviations of triplicate experiments lie within the area of circles. (Right) The mTCT8-4 theophylline aptamer is shown in light green. The sequence of each library clone is shown in the blue box corresponding to the blue portion of the riboswitch scheme.

work reported here was part of a larger project that tested riboswitch function using this plasmid in 5 different bacterial species across 3 phyla.[24] Interestingly, the N8.1 riboswitch (from now on referred to as riboswitch A) exhibited new characteristics after being cloned into the pBAV1K plasmid. Namely, the activation ratio improved to approximately 25-fold (Figure 3.7). However, expression of β-galactosidase in the presence of 1 mM theophylline was reduced from about 1,100 Miller Units to less than 250 Miller Units. This difference in reporter activity is most likely due to the use of the T5 promoter, P_{T5}, versus P_{tac}. The proposed mechanism of action for riboswitch A is

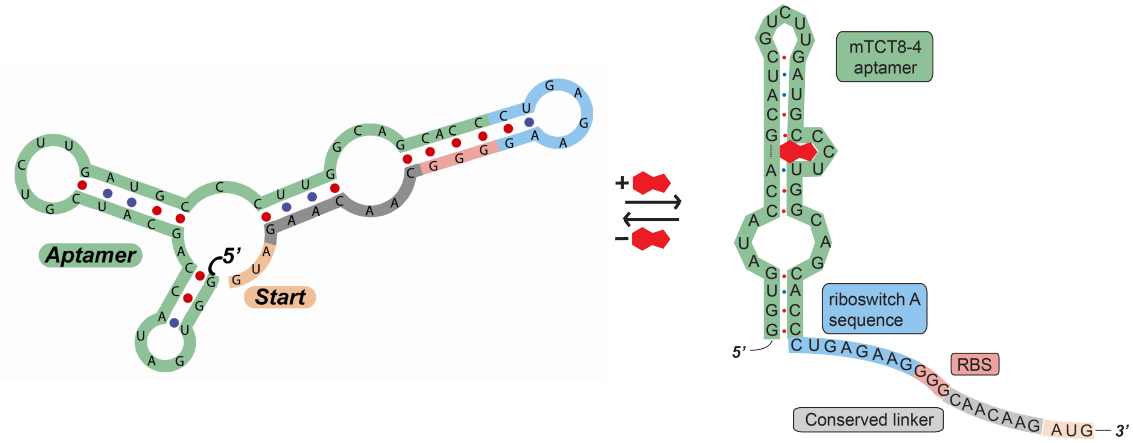

Figure 3.6 – Proposed mechanism of *A. baylyi* riboswitch isolated from library screens. (Top)Proposed Riboswitch RNA folding off clone N8.1 (riboswitch A) in the absence (left) and presence (right) of theophylline ( ). Without theophylline (left), the RBS is sequestered, preventing translation of gene. In the presence of theophylline (right), the RBS is presented to the ribosome, allowing translation.

similar to our previous riboswitches and can be seen in Figure 3.6. Folding predictions were made using *mFold*.[25]

The functionality of several other riboswitches was also examined in *A. baylyi* including a riboswitch exhibiting greater than 100-fold induction isolated via FACS (riboswitch D),[26] a riboswitch switch isolated from an N4 library in *E. coli* exhibiting 25-fold activation (riboswitch B),[22] a riboswitch originally isolated for *B. subtilis* demonstrationg fold-activation of 60 (riboswitch C),[22] and a modified version of riboswitch D which was rationally designed to carry a strong ribosome binding site (riboswitch E) and will be discussed in a later chapter. While all 5 of these riboswitches showed greater than 10-fold induction in *E. coli*, only 3 of them exceeded a 10-fold activation ratio in *A. baylyi* (Figure 3.7).

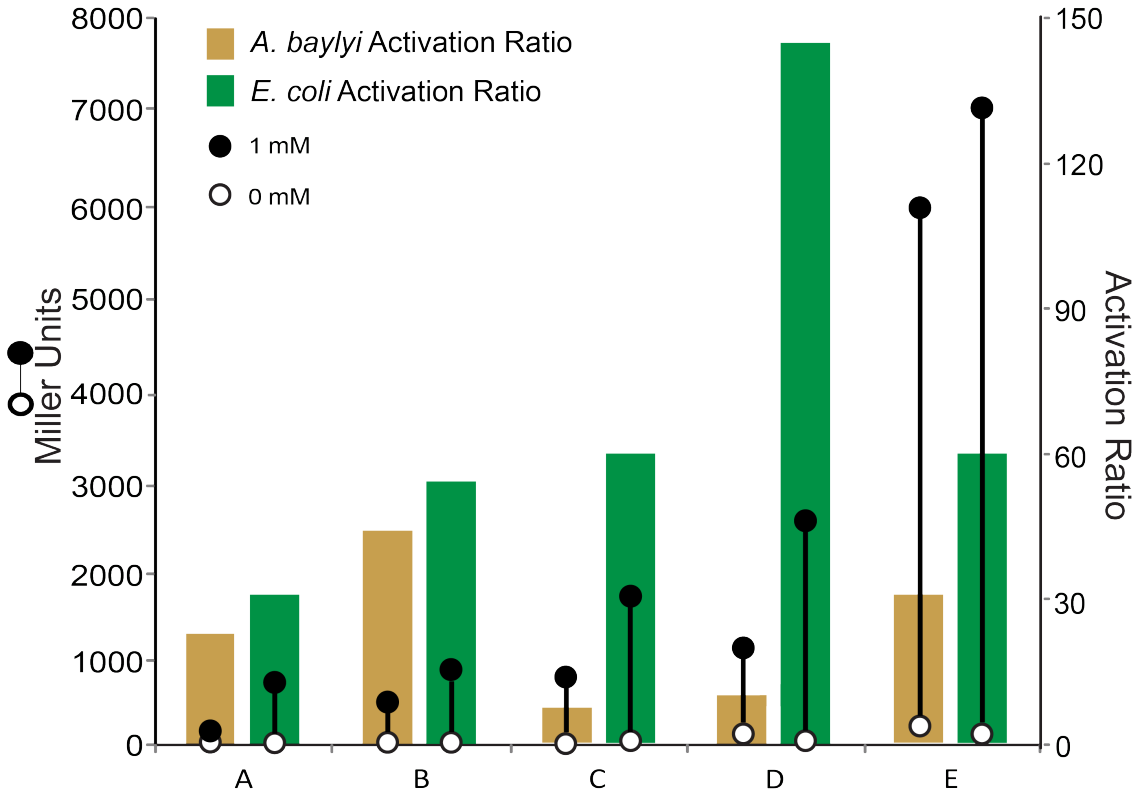

Figure 3.7 – Comparison of "switch pack" riboswitches in *A. baylyi* and *E. coli*. Strains encoding the indicated riboswitches upstream of *lacZ* were grown in the presence (●) or absence (○) of 1 mM theophylline to mid-log phase. Cells were harvested, lysed and assayed for β-galactosidase activity. The left axis indicates the Miller Units, a measure of β-galactosidase activity, and the right axis shows the activation ratio for *A. baylyi* (■) or *E. coli* (■), which is calculated by dividing the Miller units in the presense of theophylline by the units in the absense of theophylline. Dynamic range is indicated by the length of the line connecting the filled and empty circles. Standard deviations of triplicate experiments lie within the area of circles.

Interestingly, though *A. baylyi* and *E. coli* posess similar 3' ends of 16S rRNA (Figure 3.8), there is a large disparity in riboswitch behavior between each organism (Figure 3.7). Riboswitches A and B function best in *A. baylyi*, with Riboswitch E exhibiting high basal expression, but also a high activation ratio. Figure 3.8 illustrates that those same riboswitches (A, B and E) have 5 or more bases of complementarity to the anti-Shine Dalgarno sequence (a sequnce of 16S rRNA that is complementary to the RBS).[27] Remarkably, the the 3′ region of 16S rRNA containing the anti-Shine-Dalgarno

sequences is identical in both *A. baylyi* and *E. coli*. All of the switches function well in *E. coli*, though riboswitch E, which was rationally engineered to have the consensus RBS, AAGGAGGT,[28, 29] presents the highest basal expression levels in the organism. This may be explained by studies that have shown *E. coli* is able to tolerate mismatched complementarity of the RBS to the 16S rRNA, non-ideal spacing between RBS and start codon (ideal range is 7-9 nts) and more secondary structure in the RBS region. Other organisms, such as *B. subtilis,* however, are not as resilient.[30] In fact, in our multi-species studies, riboswitches C and D, which have the weakest RBS, only showed activation ratios greater than 20 in *E. coli* and *Agrobacterium tumefaciens*.[24]

It is also interesting to note that the stronger the RBS the higher basal expression becomes. This is not surprising considering, in general, weakening the RBS, even by a single base, causes a decrease in translation.[31] Therefore, having a stronger RBS should lead to increased translation. It is also possible that having a more complementary RBS allows the ribosome to effectively compete for binding over riboswitch secondary structure.'

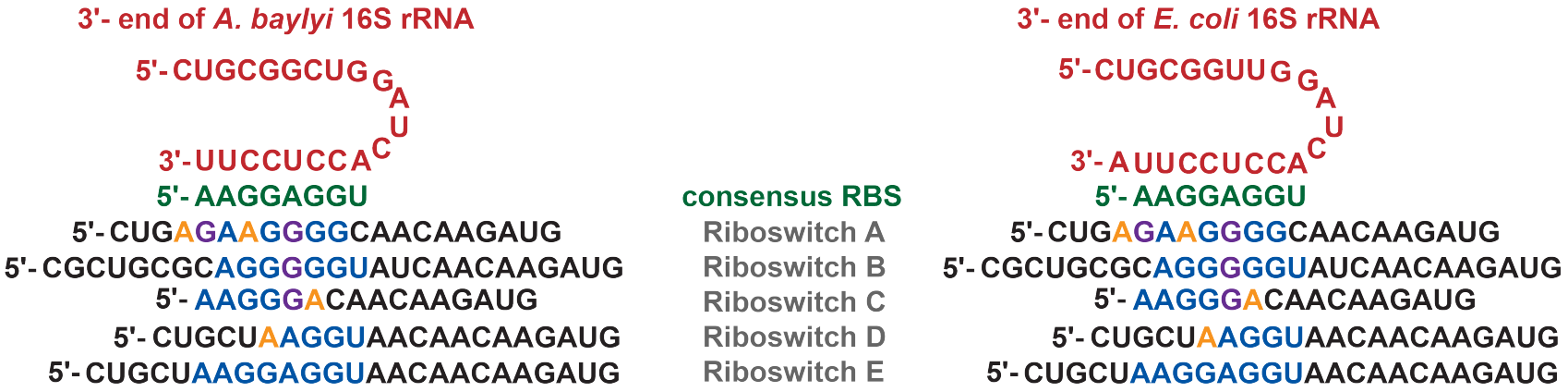

Figure 3.8 - Comparison of 16S rRNA of *A.bayyi* and *E. coli* in context of riboswitch RBS strength. The sequences of the 3' ends of the *A. baylyi* and *E. coli* 16S rRNA are shown in red. The concensus RBS is shown in green. The sequences of the linker and RBS regions (in blue) of each of the indicated riboswitches is shown aligned with the anti-Shine Dalgarno sequences of the 16S rRNA. Mismatched bases of the RBS are shown in gold. Purple denotes a wobble base pair.

3.3 Conclusion

Initially, we sought to determine if the transport of riboswitches from *E. coli* to a related Gram-negative bacterium would be a straightforward task. We posited that the simple translational mechanism would take advantage of well-conserved expression machinery in prokaryotes.[32] The ability to transport ready-made genetic regulation elements from *E. coli* to other prokaryotes could help to overcome the lack of tools found in many less studied bacterial species and avoid *de novo* isolation of riboswitches in less genetically tractable species. Preliminary results seemed to indicate that the process is not so straightforward. However, after optimizing the promoter and plasmid we found that many switches are indeed portable. The main caveat to transportability is the strength of the RBS. The *E. coli* 30S ribosomal subunit is know to be particularly promiscuous with regards to recognizing the RBS.[30] While no studies of this nature have been published regarding *A. baylyi*, evidence from this study suggests that complementarity of at least 5 bases is required for efficient translation in the presence of 1 mM theophylline in the context of our synthetic riboswitches.

To overcome the problems of transporting riboswitches identified in *E. coli* to other organisms, it may be necessary to amend our methods. The riboswitch identified from the *A. baylyi* library was found to function well in *E. coli*. More importantly, this riboswitch was also found to function in other species of bacteria including *A. tumefaciens, A. baumannii*, and, more interestingly, the Gram-positive bacterium *B. subtilis*.[24] This result supports a strategy of identifying riboswitches in a tractable host organism with more stringent RBS requirements in order to isolate portable riboswitches. At the very least, the concept of transporting riboswitches to avoid *de novo* engineering is

one worth exploring. In the next chapter, we will apply what we have learned regarding the strength of the RBS to semi-rationally design a functional riboswitch in *Streptococcus pyogenes*.

3.4 Experimental

General Considerations. Synthetic oligonucleotides were purchased from IDT (Coralville, IA). All sequences were verified using DNA sequencing (MWG Operon, Huntsville, Al or in-house sequencing). Purifications of plasmid DNA, PCR products, and enzyme digestions were performed using kits from Qiagen (Germantown, MD). Reagents, such as Theophylline, *o*-nitrophenyl-β-D-galactopyranoside (ONPG), *p*-nitrophenyl-β-D-glucuronide (PNPG), ampicillin and kanamycin were purchased from Sigma-Aldrich (St. Louis, MO). 5-bromo-4-chloro-indolyl-β-D-glucuronide (X-gluc) was purchased from US Biological (Swampscott, Massachusetts). All experiments were performed in *E. coli* TOP10 F' cells (Invitrogen, Carlsbad, CA) or *E. coli* MDS42 cells (Scarab Genomics). *A. baylyi* ADP1 cells were obtained from ATCC (Manassas, VA). All cultures were in LB media (EMD Bioscience, Merck KGaA, Darmstadt, Germany).

Randomized Libraries. Libraries were constructed using mutagenic oligonucleotides synthesized with eight randomized bases. These oligonucleotide primers annealed to complementary regions of the mTCT8-4 aptamer[14] or regions of the expression platform with degenerate nucleotide overhangs. The outermost 5' primer featured a KpnI restriction enzyme recognition site while the outermost 3' primer featured a HindIII site. Overlap PCRs of the library and the pWH1266-derived plasmid pSKD-AB were restriction digested with KpnI and HindIII (NEB, Ipswich, MA) for at least 1 hour at 37 °C. Vector DNA was further processed with calf intestinal phosphatase (CIP, NEB)

for at least 1 hour at 37 °C then 30 minutes at 50 °C to minimize self-ligation. Ligation reaction were performed with T4 DNA ligase (NEB) at either room temperature for 1 hour or 16 °C overnight. DNA was extracted by butanol precipitation and resuspended in water.

Construction of broad-host range "switch pack". All "switch pack" experiments in *A. baylyi* and *E. coli* were performed using the pBAV1K[23] plasmid featuring a T5 promoter upstream of *lacZ*. Plasmids were constructed using the 5' KpnI restriction site and the 3' SpeI site flanking P_{T5} upstream of a riboswitch controlling *lacZ* expression. DNA fragments were constructed by overlap PCR. The linker sequence of each riboswitch is shown in Figure 2.7.

Transformation of *A. baylyi.* Riboswitch library constructs or "switch pack" constructs were used to electroporate TOP10 F' *E. coli* cells via a previously reported protocol.[33] A successful transformation yields approximately 3,000 to 4,000 grown on selective LB media supplemented with kanamycin (50 μg/mL). All clones are cultured together and grown for ~4 hours with shaking at 37 °C in 50 mL of selective media. The resultant library culture is miniprepped and used for transforming *A. baylyi* as previously reported.[10] Briefly, a single *A. baylyi* colony was picked from an LB agar plate and used to inoculate 5 mL of LB media that was then shaken at 30 °C overnight. A 60 μL aliquot of overnight culture was used to inoculate 900 μL of fresh LB in a glass culture tube and shaken (225 rpm) at 30 °C for 2-3 hrs. A 5 μL aliquot of miniprepped plasmid DNA (~500 μg) was added to the culture, which was allowed to continue shaking at 30 °C for 2.5 hrs. A fraction of this culture (100 μL) was plated on LB agar, supplemented with kanamycin (50 μg/mL) and X-gluc (25 mg dissolved in 4.0 mL of dimethyl formamide,

final concentration 0.008%) but no theophylline, and the cells were incubated at 30 °C for 18 h in a stationary incubator. Plates were incubated at 4 °C after overnight incubation to allow the development of blue color. White to light blue colonies were selected for further characterization.

96-well Plate-Based Screen The whitest colonies from each plate were picked by hand and inoculated into each well of a 96-well microtiter plate (Costar, Sigma-Aldrich) containing 200 μL LB supplemented with 50 μg/mL kanamycin. The plate was sealed with parafilm and incubated overnight at 30 ºC with gentle shaking (180 rpm). On day 2, four 96-well plates (duplicate sets of two) were inoculated using 2 μL of the overnight culture into 200 μL fresh media supplemented with kanamycin (50 μg/mL). One set of plates was also supplemented with 1 mM theophylline. Plates were incubated for approximately 2.5 hrs at 30 ºC with shaking (210 rpm) to an OD600 of 0.085-0.14 as determined by a Biotek microplate reader. These values correspond to an OD600 of 0.3-0.5 with a 1 cm path length cuvette. A high-throughput microtiter plate assay for β-glucuronidase activity (or β-galactosidase activity) was adapted from previously described methods.[17] Instead of permeabolization, cultures were lysed using a 10:1 Pop Culture® (Novagen, Merck KGaA, Darmstadt, Germany): hen egg white lysozyme mixture. After addition of lysing solution, the cultures were mixed by pipetting up and down, and allowed to stand at room temperature for 5 min. Fresh plates were prepared with 132 μL of Z-buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 50 mM β-mercaptoethanol, pH 7.0). To this buffer was added 15 μL of lysed *A. baylyi* culture followed by addition of 29 μL of 4 mg/mL PNPG in 100 mM NaH_2PO_4 (or ONPG if assaying β-galactosidase). PNPG (or ONPG) was allowed to hydrolyze until

the solution became a faint yellow color. The reaction was quenched by the addition of Na_2CO_3 (75 μL of a 1 M solution), which caused the solution to turn a deeper yellow. The reaction time was recorded as the length of time from addition of substrate until addition of quenching reagent. The OD420 for each well was determined using a Biotek microplate reader. Enzyme activity was determined by Miller units, calculated by the formula: $Miller\ Units = \frac{Vol_{total}(in\ \mu L)\times(OD_{420})}{OD_{600}\times time_{rxn(in\ minutes)}\times Vol_{lysate}(in\ \mu L)}$.

Characterization of riboswitches. Activation Ratios are calculated by dividing the Miller Units in the presence of theophylline by the Miller Units in the absence of theophylline. Candidate clones featuring an activation ratio of 5 or higher were inoculated and assayed for β-glucuronidase (or β-galactosidase) activity in higher volume in triplicate by the Miller method.[17]

3.5 References

(1) Murin, C. D., Segal, K., Bryksin, A. and Matsumura, I. *Appl. Environ. Microbiol.* **2012**, 78, 280-283.

(2) Winkler, W. C. and Breaker, R. R. *Annu. Rev. Microbiol.* **2005**, 59, 487-517.

(3) Barrick, J. E. and Breaker, R. R. *Genome Biol.* **2007**, 8, R239.

(4) Desai, S. K. **(2008)** Engineering Control of Gene Expression in Bacteria Using RNA-Small Molecule Interactions. Doctoral Dissertation, Emory University.

(5) Hunger, M., Schmucker, R., Kishan, V. and Hillen, W. *Gene*. **1990**, 87, 45-51.

(6) Lynch, S. A., Desai, S. K., Sajja, H. K. and Gallivan, J. P. *Chem. Biol.* **2007**, 14, 173-184.

(7) Barbe, V., Vallenet, D., Fonknechten, N., Kreimeyer, A., Oztas, S., Labarre, L., Cruveiller, S., Robert, C., Duprat, S., Wincker, P., Ornston, L. N., Weissenbach, J., Marlière, P., Cohen, G. N. and Mèdigue, C. *Nucleic Acids Res.* **2004**, 32, 5766-5779.

(8) Vaneechoutte, M., Young, D. M., Ornston, L. N., De Baere, T., Nemec, A., Van Der Reijden, T., Carr, E., Tjernberg, I. and Dijkshoorn, L. *Appl. Environ. Microbiol.* **2006**, 72, 932-936.

(9) Thomas, C. M. and Nielsen, K. M. *Nat Rev Micro*. **2005**, 3, 711-721.

(10) David Metzgar, J. M. B., Valérie Pezo, John Reader, Volker Döring, Paul Schimmel, Philippe Marlière and Valérie de Crécy-Lagard,. *Nucleic Acids Res.* **2004**, 32, 5780-5790.

(11) de Berardinis, V., Vallenet, D., Castelli, V., Besnard, M., Pinet, A., Cruaud, C., Samair, S., Lechaplais, C., Gyapay, G., Richez, C., Durot, M., Kreimeyer, A., Le Fevre,

F., Schachter, V., Pezo, V., Doring, V., Scarpelli, C., Medigue, C., Cohen, G. N., Marliere, P., Salanoubat, M. and Weissenbach, J. *Mol. Syst. Biol.* **2008**, 4,

(12) Nan‐Yao Lee, M., Hsin‐Chun Lee, MD, Nai‐Ying Ko, RN, PhD, Chia‐Ming Chang, MD, Hsin‐I Shih, MD, Chi‐Jung Wu, MD, Wen‐Chien Ko, MD. *Infect. Control Hosp. Epidemiol.* **2007**, 28, 713-719.

(13) Adams, M. D., Goglin, K., Molyneaux, N., Hujer, K. M., Lavender, H., Jamison, J. J., MacDonald, I. J., Martin, K. M., Russo, T., Campagnari, A. A., Hujer, A. M., Bonomo, R. A. and Gill, S. R. *J. Bacteriol.* **2008**, 190, 8053-8064.

(14) Jenison, R. D., Gill, S. C., Pardi, A. and Polisky, B. *Science*. **1994**, 263, 1425-1429.

(15) de Boer, H. A., Comstock, L. J. and Vasser, M. *Proc. Natl. Acad. Sci.* **1983**, 80, 21-25.

(16) Olsen, M., Iverson, B. and Georgiou, G. *Curr. Opin. Biotechnol.* **2000**, 11, 331-337.

(17) Miller, J. H. *Experiments in Molecular Genetics*. **1972**, 352-355.

(18) Jefferson, R. A., Burgess, S. M. and Hirsh, D. *Proc. Natl. Acad. Sci.* **1986**, 83, 8447-8451.

(19) Platteeuw, C., Simons, G. and de Vos, W. M. *Appl. Environ. Microbiol.* **1994**, 60, 587-593.

(20) Jefferson, R. A. *Nature*. **1989**, 342, 837-838.

(21) Desai, S. K. and Gallivan, J. P. *J. Am. Chem. Soc.* **2004**, 126, 13247-13254.

(22) Topp, S. **(2009)** Biological Engineering with Chemical-Sensing Macromolecular Switches:

I. Discovery and Applications of Small-Molecule Dependent Synthetic Riboswitches

II. A Genetic Toolbox for Creating Reversible Ca2+-Sensitive Biomaterials. Doctoral Dissertation, Emory University.

(23) Bryksin, A. V. and Matsumura, I. *PLoS ONE*. **2010**, 5, e13244.

(24) Topp, S., Reynoso, C. M. K., Seeliger, J. C., Goldlust, I. S., Desai, S. K., Murat, D., Shen, A., Puri, A. W., Komeili, A., Bertozzi, C. R., Scott, J. R. and Gallivan, J. P. *Appl. Environ. Microbiol.* **2010**, 76, 7881-7884.

(25) Zuker, M. *Nucleic Acids Res.* **2003**, 31, 3406-3415.

(26) Lynch, S. A. and Gallivan, J. P. *Nucleic Acids Res.* **2009**, 37, 184-192.

(27) Shine, J. and Dalgarno, L. *Proc. Natl. Acad. Sci.* **1974**, 71, 1342-1346.

(28) McCarthy, J. E. G. and Gualerzi, C. *Trends Genet.* **1990**, 6, 78-85.

(29) Scherer, G. F., Walkinshaw, M. D., Arnott, S. and Morré, D. J. *Nucleic Acids Res.* **1980**, 8, 3895-3908.

(30) Vellanoweth, R. L. and Rabinowitz, J. C. *Mol. Microbiol.* **1992**, 6, 1105-1114.

(31) Jacob, W. F., Santer, M. and Dahlberg, A. E. *Proc. Natl. Acad. Sci.* **1987**, 84, 4757-4761.

(32) Marilyn, K. *Gene*. **1999**, 234, 187-208.

(33) Dower, W. J., Miller, J. F. and Ragsdale, C. W. *Nucleic Acids Res.* **1988**, 16, 6127-6145.

Chapter 4 - Developing Synthetic Riboswitches for use in *Streptococcus pyogenes*

4.1 Introduction

The ability to transport synthetic riboswitches between a variety of Gram-negative, non-pathogenic bacteria has demonstrated that these function as portable regulatory components (Chapters 2 and 3). However, to extend the utility of our synthetic riboswitches to pathogenic organisms that possess few or no tools for conditional gene expression, *Streptococcus pyogenes* was chosen for further study. *S. pyogenes*, also known as Group A Streptococcus (GAS), is a Gram-positive, exclusively human pathogen that grows in culture as pairs or chains of variable length (Figure 4.1a). It is a beta-hemolytic bacterium that, when plated on blood agar, creates a transparent zone of complete hemolysis of red blood cells (Figure 4.1b).[1] *S. pyogenes* causes an array of diseases ranging from superficial to acute or even fatal.[2-4] Some of the more minor infections include streptococcal pharyngitis (Strep throat) and impetigo (suppurative or non-suppurative infection of the skin).[5] Other diseases include Scarlet fever, which can lead to rheumatic fever and heart valve damage,[3] as well as necrotizing fasciitis (flesh-eating disease) among many others.[6] While *S. pyogenes* remains sensitive to penicillin,[7] the failure rate of treatment is as high as 37%[8], and the occurrence of allergic reaction (1-10% of the population)[9] precludes its use in some patients. *S. pyogenes* causes high rates of human morbidity and mortality, with 8,950–11,500 episodes and 1,050–1,850 deaths in the United States each year,[10] culminating in 1 billion dollars of annual direct cost in the United States alone.[11]

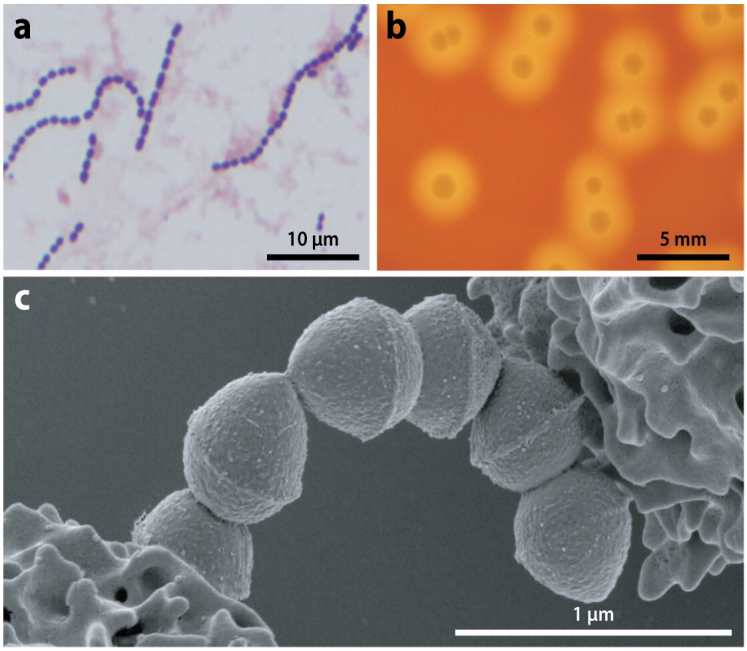

Figure 4.1 – *Streptococcus pyogenes* morphology. (a) Micrograph depicts *S. pyogenes* cells growing in pairs and chains. (b) Photograph of *S. pyogenes* grown on blood agar media. Apparent halos of cleared blood agar are caused by beta (complete) hemolysis. (c) Scanning electron micrograph of GAS organisms interacting with human neutrophils). Reprinted from Olsen and Musser. 2010.[6]

To be capable of easily disseminating throughout the human body and causing a wide spectrum of diseases, *S. pyogenes* must persist and replicate in numerous environments. *S. pyogenes* has evolved sophisticated strategies and complex regulatory mechanisms to recognize and adapt to new surroundings.[4] Many bacteria respond to fluctuating environmental stimuli using two-component signal transduction systems. These systems convert signal recognition to a change in gene expression or triggers cellular response cascades.[12] Two-component systems include a sensor kinase possessing a signal recognition domain. Upon recognition of a specific signal, the sensor kinase transfers a phosphoryl group to its partner response regulator protein. This phosphorylation event activates a downstream domain that triggers the appropriate response.[13] There are 13 putative two-component regulatory systems in *S. pyogenes*.[14] The best characterized of these is the CovR/S (also known as CsrR/S) system that

[6] Figure reprinted from Olsen, R. and Musser, J. *Annual Reviews in Pathology: Mechanisms of Disease*. **2010**, 5, 1-31 with permission from Annual Reviews.

regulates many virulence factors necessary for *S. pyogenes* pathogenesis[14] and controls 15% of the genome.[15] The virulence factors regulated by CovR/S allow the organism to attach to the host (Figure 4.1c), evade the immune system, and break down tissue to propagate to other tissues in the human body.[14, 16, 17] CovR is a transcriptional repressor that not only represses expression of several virulence factors, but also of its own transcription.[17] This elegant system carefully regulates the intracellular concentration of CovR by self-repression of transcription in the presence of high CovR concentration and de-repression in the presence of low CovR concentration. By limiting the intracellular concentration of CovR, the cell is able to rapidly inactivate the repression of important environmental responses and allows for immediate response to even short-lived environmental stimuli.[17] CovS-independent phosphorylation of CovR results in dimerization and inactivation of CovR resulting in derepression of all CovR-repressed genes, including virulence genes.[18] Later studies found that CovS is the sensor kinase that phosphorylates CovR under mild stress conditions,[19] thereby inactivating CovR-mediated repression of virulence genes. Simultaneously, phosphorylated CovR represses other genes, such as *speB*.[20] SpeB degrades proteins that *S. pyogenes* uses to guard against host cell attack of the bacterium.[21] Finally, *covS* mutants are unable to adhere to human host cells.[22] Taken together, these interactions suggest CovS is required for upregulation of virulence genes and down-regulation of genes that would hinder infection.

There are many known and putative virulence factors that contribute to host–pathogen interactions, including those regulated by the CovR/S system. However, efforts to further characterize these genes and develop treatments and preventative measures

against GAS-mediated diseases are impeded by the absence of genetic regulatory tools for conditional gene expression. While many regulatory tools exist for Gram-negative bacteria, these tools are often not functional in Gram-positive bacteria. This is likely due to the more stringent promoter requirements found in Gram-positive bacteria.[23] Regulatory systems within Gram-positive bacteria often require factors present only in the native organism, such as unique sigma factors.[24] At the time of this study, only one tool for conditional gene expression in *S. pyogenes* had been published. The nisin-inducible promoter, P_{nisA}, is from the *Lactococcus lactis* nisin gene cluster. Nisin, an antibacterial produced during fermentation by *L. lactis*, is synthesized via post-translational modification mediated by the gene products of the nisin biosynthetic gene cluster, *nisABTCIPRKFEG*.[25] This gene cluster contains two nisin-dependent promoters, one of which is P_{nisA}.[25] The genes *nisK* and *nisR* belong to the class of two-component regulatory systems similar to CovR/S,[25] where NisR is the response regulator and NisK is a histidine kinase. Together they exhibit nisin-dependent regulation of nisin biosynthesis.[26, 27] This system produced a 10-fold increase in reporter gene activity in *S. pyogenes* after addition of exogenous nisin.[28] However, this inducible system has several disadvantages, including high basal reporter activity in the absence of inducer.[28] Also, expression from P_{nisA} requires the NisR/K two-component system, necessitating expression of *nisR* and *nisK* in the new organism. Additionally, nisin is a bactericidal compound, and lethal concentrations vary among different strains of *S. pyogenes* studied.[28] With these disadvantages, employment of this regulatory system has not been widespread.[29]

One possible solution to overcome the dearth of conditional expression systems in Gram-positive organisms is to use synthetic riboswitches, which have been selected to display robust expression in the presence of inducer and low basal expression in the absence of inducer.[30, 31] Our theophylline-dependent synthetic riboswitches bypass foreign promoter incompatibility by functioning at the translational level and with a simple mechanism that uses highly conserved translational machinery.[30] Unlike most naturally occurring riboswitches, which respond to specific metabolites within the cell, the ligand of our synthetic riboswitch, theophylline, is a non-endogenous molecule. Inducing with an orthogonal ligand will minimize interactions with natural biosynthetic pathways. Additionally, riboswitches do not require accessory proteins[32] and can be encoded on a single plasmid or easily integrated into the chromosome. Natural riboswitches are encoded in many bacterial species including *S. pyogenes*.[33] Therefore, we hypothesize that synthetic riboswitches may be a valuable alternative for developing inducible expression systems for *S. pyogenes*.

4.2 Results and Discussion

To determine if the use theophylline-dependent synthetic riboswitches is feasible in *S. pyogenes*, theophylline toxicity was measured. We have previously shown that theophylline is toxic to *E. coli* at concentrations exceeding 5 mM.[34] Culture turbidity, measured as Klett units (Klett Manufacturing Co., New York, N.Y.), was recorded each hour after an initial 3-hour incubation in media supplemented with increasing concentrations of theophylline from 0 mM to 10 mM theophylline. Concentrations equal to and exceeding 5 mM theophylline began to affect normal growth rate (Figure 4.2), which is similar to data obtained in *E. coli*.[34] We hypothesized

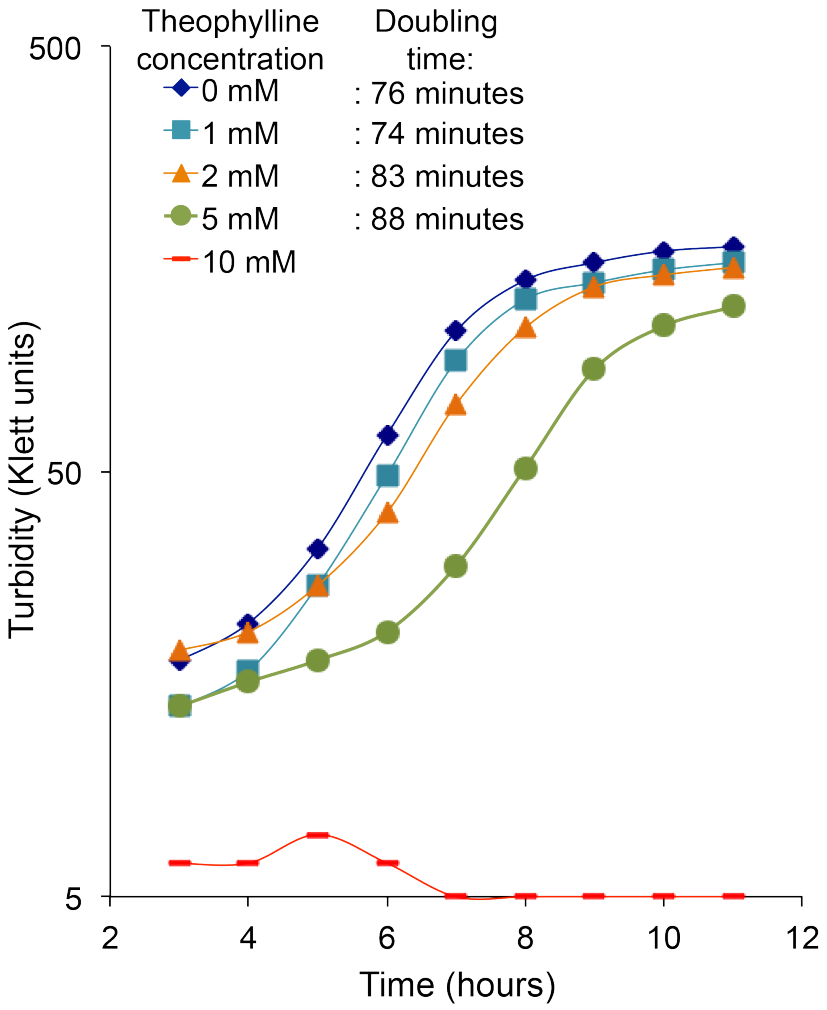

Figure 4.2 – Theophylline toxicity in *S. pyogenes* strain JRS1278. Overnight cultures of *S. pyogenes* strain JRS1278 were diluted into fresh media and grown in 0 mM (◆), 1 mM (■), 2 mM (▲), 5 mM (●) and 10 mM (▬) theophylline. Turbidity was measured in Klett units with a Klett-Summerson photoelectric colorimeter with A filter. Calculated doubling times for each concentration are shown.

that the toxic effects observed were also evidence that theophylline was indeed permeable to the *S. pyogenes* membrane. Due to these similarities, we decided to test riboswitch function in *S. pyogenes* at 1 mM theophylline.

Our lab previously isolated a theophylline-dependent riboswitch (riboswitch D) from *E. coli* with over 100-fold activation (Figure 4.3).[31] We chose this robust riboswitch to begin our studies of riboswitch portability into *S. pyogenes*. Riboswitch D

was cloned into an *E. coli-S. pyogenes* shuttle vector, pEU7742, downstream of a strong promoter, P_{sag}, and upstream of the reporter gene *gusA*, which encodes β-glucuronidase. In *S. pyogenes*, P_{sag} is responsible for initiating transcription the 9-gene *sag* operon.[35] The *sag* operon encodes genes necessary for production of Streptolysin S (SLS), a cytolytic toxin that plays a role in beta-hemolysis.[35] The *sagA* gene encodes the key peptide component of this virulence factor.[36] Notably, the *sag* locus, a known virulence locus, is regulated by the CovR/S system; therefore, a CovR mutant strain (JRS1278, T.C. Barnett and J.R. Scott, unpublished) was used for these studies. Unfortunately, we did not observe any β-glucuronidase activity in the presence or absence of theophylline (data not shown).

There are at least four hypotheses for the absence of reporter activity. One possibility is that the concentration of theophylline used was sufficient to activate detectable riboswitch-mediated gene expression. The concentration of theophylline could be increased to 2 mM without causing a significant decrease in growth rate. This, however, did not serve to improve gene expression. Interestingly, though it is widely accepted that Gram-positive bacteria are more permeable to small molcules due to the absence of an outer membrane,[37] we found in our later studies that Gram-positive bacteria can accommodate higher concentrations of theophylline.[38] However, increasing theophylline conentration did not demonstrate riboswitch-dependent gene expression from riboswitch D (data not shown).

We also suspected that intrinsic characteristics of P_{sag} may affect riboswitch function. The P_{sag} promoter within pEU7742 carries an additional 140 bases encoded in the 5' untranslated region (UTR) downstream of the transcriptional start site. This

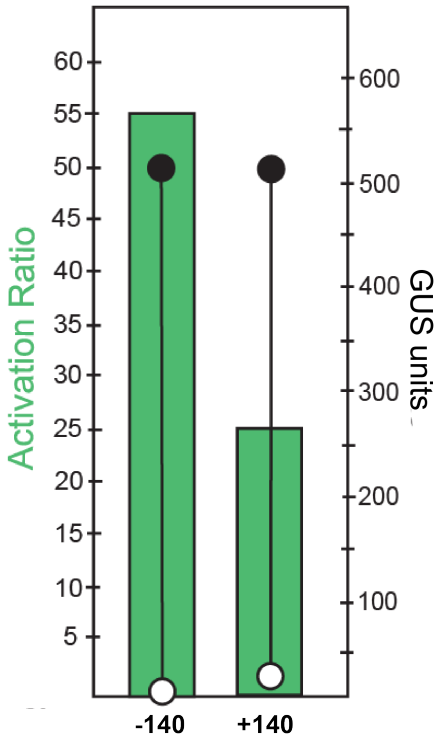

Figure 4.3 - β-Glucuronidase Activities for riboswitches with and without the extended 5'-UTR in *E. coli*. Strains encoding riboswitch D with the extended 5'-UTR (+140) or without (-140) upstream of *gusA* were grown in the presence (●) or absence (○) of 1 mM theophylline to stationary phase. Cells were harvested, lysed and assayed for β-glucuronidase activity. The right axis indicates the GUS Units, a measure of β-glucuronidase activity, and the left axis shows the activation ratio (green bar) which is calculated by dividing the GUS units in the presense of theophylline by the units in the absense of theophylline. Standard deviations of triplicate experiments lie within the area of circles.

extended 5'-UTR could disrupt the mRNA secondary structure required for proper riboswitch function. To test the effects of this extended 5'-UTR, the additional 140 bases was cloned upstream of the aptamer of riboswitch D in the pEU7742 shuttle plasmid and tested in *E. coli*. We found that, in *E. coli*, this extended 5'-UTR did indeed decrease the activation ratio of riboswitch D by causing an increase in basal expression (Figure 4.3). However, when these same constructs were created for riboswitch D and assessed in

S. pyogenes, no reporter activity was observed from either construct, suggesting that the extended 5' UTR is not responsible for the absence of gene activity (data not shown).

Previously, it was reported that transcripts expressed from P_{sag} have increased mRNA stability at 2 hours into the stationary phase of growth.[39] The mRNA exhibited increased half-lives only in stationary and not in exponential phase, suggesting that the mRNA stability is not due to the extended 5'-UTR alone.[39] Increased half-lives were also observed when P_{sag} was used to drive transcription of non-native genes. [40] This evidence suggests that we would observe the highest gene expression from P_{sag} from cells in stationary phase; however, no gene expression was seen in stationary phase. Although these promoter specific factors were not the reason for the inability to induce theophylline-dependent gene expression, all subsequent assays in *S. pyogenes* were performed at 2 hours into stationary phase since previous data indicated the most stable transcripts are produced at this time point.

Finally, we hypothesized that *S. pyogenes* may not recognize the Shine-Dalgarno sequence, or ribosome binding site (RBS), in our theophylline-dependent riboswitch constructs because they were selected for in *E. coli*. The RBS is complementary to the 3'

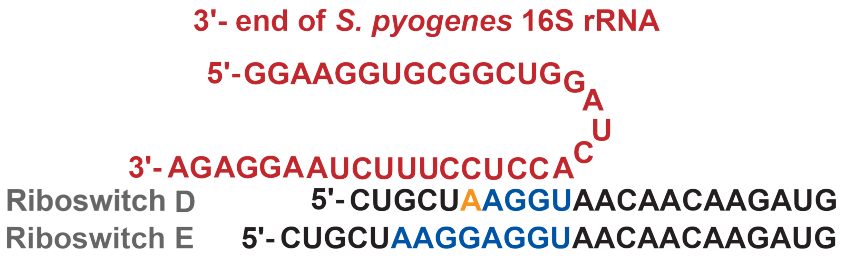

Figure 4.4 - 3' end of *S. pyogenes* 16S rRNA aligned with the RBS of riboswitches D and E. The 16S rRNA of *S. pyogenes* is shown in red. Fragments of riboswitches D and E are shown in black with the RBS highlighted in blue. The gold letters represent mismatched complementarity to the 16S rRNA.

end of the 16S ribosomal RNA (rRNA) and plays a role in 30S ribosomal subunit recognition of the translational start site of open reading frames.[41, 42] The *E. coli* 30S ribosomal subunit is notoriously promiscuous in its recognition of the RBS and is more tolerant of secondary structure in the 5'-UTR.[23] It is also more tolerant to suboptimal spacing between the RBS and the start codon and weak RBS complementarity than the Gram-positive bacterium *B. subtilis*.[43, 44] To optimize recognition of the RBS in our theophylline-dependent ribowitches in Gram-positive bacteria, a semi-rational design approach was taken to engineer riboswitches for *S. pyogenes*. The anti-Shine-Dalgarno sequence of the 16S rRNA of *S. pyogenes* is "ACCUCCUU", so the trinucleotide "AGG" was added to the 5' end of the RBS of riboswitch D to create a complementary sequence to the anti-Shine-Dalgarno, producing riboswitch E (Figure 4.3). An alternative method to engineering identical complementarity to the RBS is mutating the "TAA" sequence of riboswitch D to an "AGG". This method would create a complementary RBS in riboswitch D; however, it would decrease the spacing between the RBS and start codon by 3 nucleotides. Previous reports and analysis of the *S. pyogenes* genome suggest that ideal spacing between the RBS and initation codon is 7-9 nucleotides.[43, 45] Additionally, studies in *B. subtilis* analyzing RBS strength versus spacing length showed that longer spacing yields higher translation up to 13 nucleotides.[43] Taken together, these data suggested that the longer spacing of the insertion construct versus the shorter spacing of the mutation construct may prove more efficient on the translational level. This semi-rationally designed riboswitch was cloned downstream of P_{sag} and upstream of *gusA*.

The resulting riboswitch E construct was grown in increasing concentrations of theophylline (0-9 mM) to obtain a dose-response curve. Indeed, riboswitch E

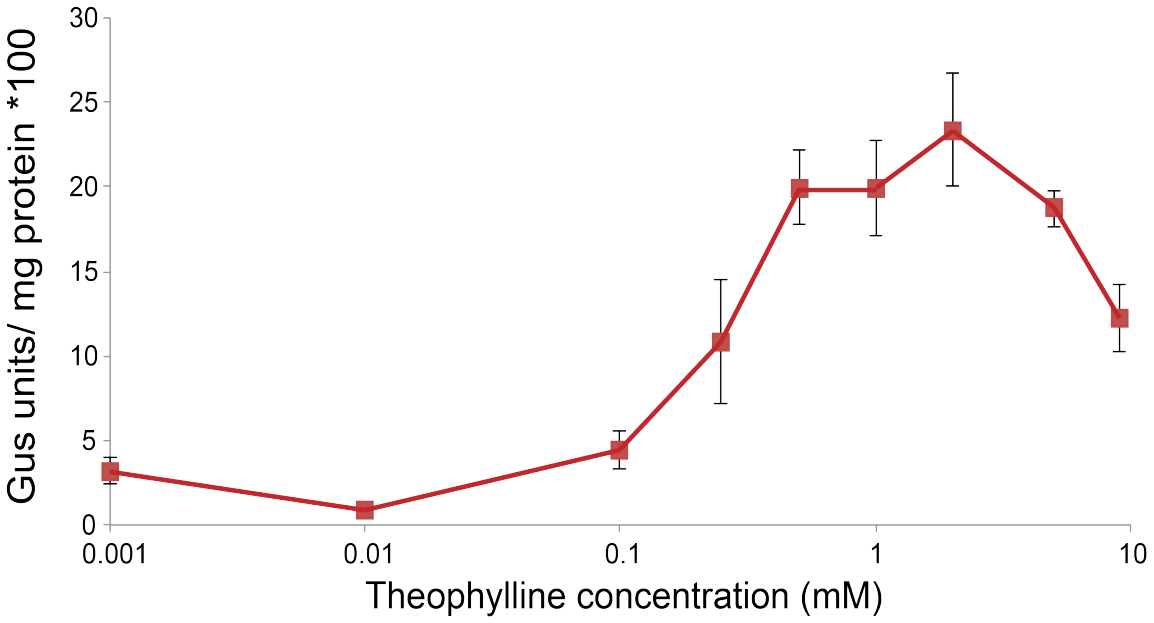

Figure 4.5 – Theophylline-dose-dependent increase of β-glucuronidase activity. Overnight cultures of *S. pyogenes* transformed with riboswitch E upstream of *gusA* were diluted into fresh media containing 0 to 9 mM theophylline. Cultures were grown to stationary phase and assayed for β-glucuronidase activity. The y-axis represents β-glucuronidase activity in gus units/mg total cell protein and the x-axis indicates the concentration of theophylline in mM.

demonstrated a dose-dependent induction of β-glucuronidase activity (Figure 4.5), indicating that a stronger RBS was needed for efficient translation of our theophylline-dependent riboswitch fusion in *S. pyogenes*.

Since the extended P_{sag} 5'-UTR seemed to reduce riboswitch function in *E. coli* (Figure 4.3), we tested the effects it would have in *S. pyogenes*. We tested riboswitch E with and without the 5'-UTR extension in *S. pyogenes*. Surprisingly, the additional 5'-UTR increased β-glucuronidase activity independently of the absence or presence of theophylline (Figure 4.6, left). To verify that the effects of the extended 5'-UTR were riboswitch-independent, constructs lacking the riboswitch with and without the extended 5'-UTR were tested. The positive control lacking the extended 5'-UTR exhibits a

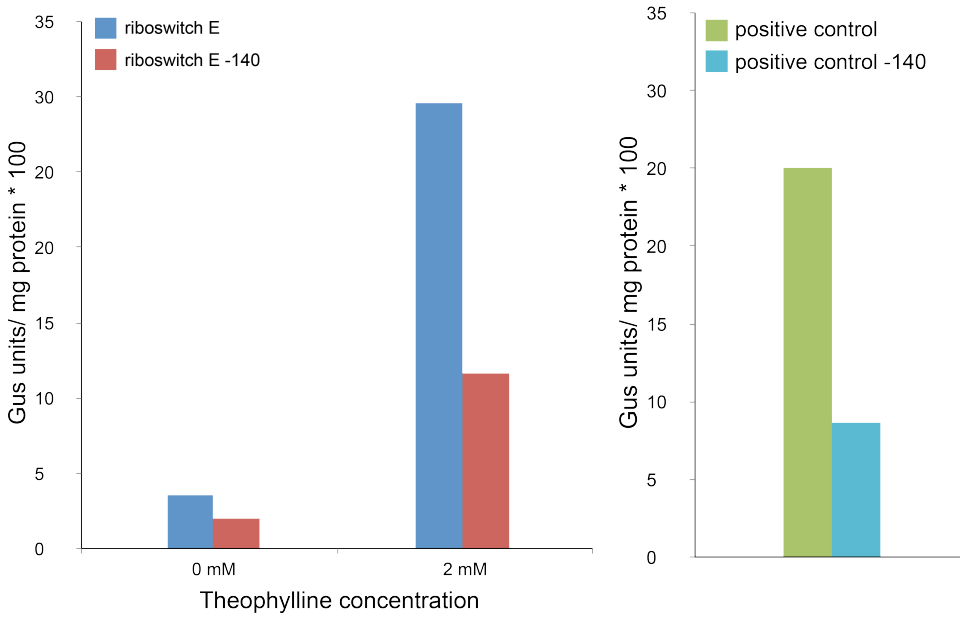

Figure 4.6 – The extended 5'-UTR found downstream of the P_{sag} promoter allows for more robust expression of β- glucuronidase. (Left) Strains of *S. pyogenes* transformed with plasmids harboring riboswitch E with (■) and without (■) the additional 5'-UTR were grown in the presence or absence of 2 mM theophylline and then assayed for β-glucuronidase activity. The y-axis represents β-glucuronidase activity, the x-axis represents theophylline concentration. (Right) Strains of *S. pyogenes* transformed with plasmids harboring constitutively expressed positive control constructs with (■) or without (■) the additional 5'-UTR assayed for β-glucuronidase activity. The y-axis represents β-glucuronidase activity.

significant decrease in β-glucuronidase activity compared to the positive control containing the extended 5'-UTR (Figures 4.6, right). These data may be explained by the finding that transcripts produced from the P_{sag} promoter, including non-native transcripts, have longer half-lives.[40] The 5'-UTR of transcripts often plays a role in mRNA stability by protecting transcripts from ribonuclease activity in both Gram-negative and Gram-positive bacteria.[46, 47] It is possible that without the extended 140 base 5'-UTR, the riboswitch 5'-UTR is more susceptible to degradation by native ribonucleases. However, this does not fully explain the role of the extended 140 base 5'-UTR. Non-native transcripts from P_{sag} are only partially stabilized,[40] indicating that there may be another

feature of the transcript that allows for the more robust expression levels exhibited when the 140 base 5'-UTR is present. We hypothesize that the riboswitch transcripts with the extended 5'-UTR would exhibit at least some amount of increased stability against RNA degradation in stationary phase.

To verify that stationary phase was the ideal time point for riboswitch function analysis in *S. pyogenes*, β-glucuronidase assays were performed on cells grown in increasing concentrations of theophylline (0-9 mM) and harvested during mid-exponential and stationary phases of growth. β-glucuronidase activity was higher in stationary phase compared to mid-exponential phase (Figure 4.7), confirming that the stabilized mRNA transcripts generated from P_{sag} influence growth phase-dependent gene expression.

The protocol traditionally used for β-glucuronidase assays was examined since β-

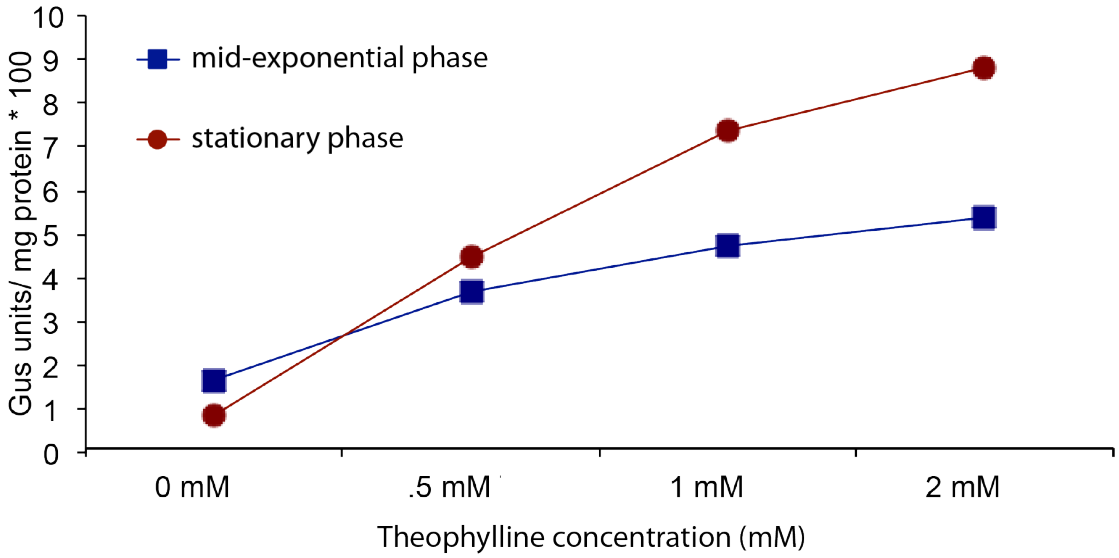

Figure 4.7 – β-glucuronidase activity of cultures harvested at mid exponential and stationary phases of growth. *S. pyogenes* carrying riboswitch E upstream of gusA were grown in the indicated concentrations of theophylline (x-axis). Aliquots were harvested at mid-exponential phase (—■—, Klett unit = ~65) or stationary phase (—●— , ~ 150 Klett units) and assayed for β-glucuronidase activity (y-axis).

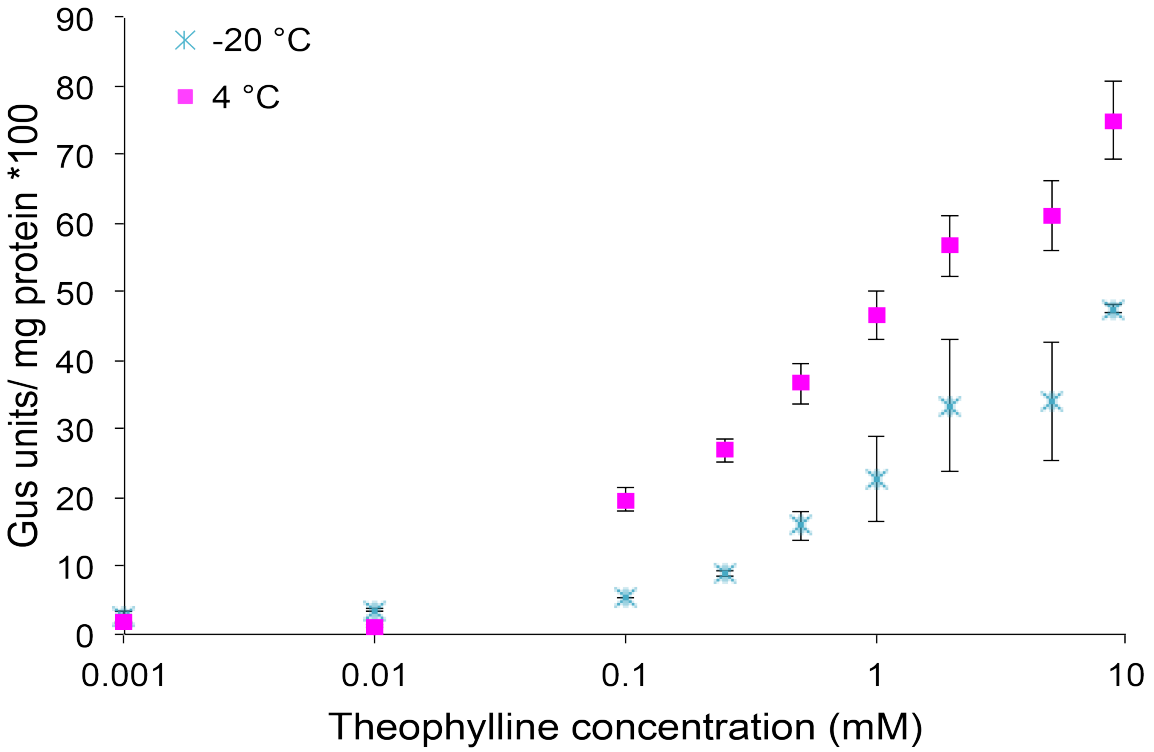

Figure 4.8 – The effects of cell lysate storage conditions on β-glucuronidase activity. *S. pyogenes* strains harboring riboswitch E upstream of gusA were grown in cultures supplemented with the indicated concentration of theophylline (x-axis). Cells were harvested at stationary phase and lysed. Cell lysates were either stored at -20 °C (✳) or 4 °C (■). B-glucuronidase activity was determined as indicated by the y-axis.

glucuronidase activity (Gus units/ mg protein *100) fluctuated greatly from experiment to experiment. Because of the time consuming nature of the protocol, the harvested cell lysates are often stored at -20°C overnight and assayed the next day. To determine whether freezing the lysates affected β-glucuronidase activity, one set of lysates were stored at 4 °C and another set at -20 °C. The lysates stored at -20 °C exhibited less activity than the lysates stored at 4 °C (Figure 4.8), suggesting that storing cell lysates overnight at lower temperatures affects the activity of β-glucoronidase in subsequent enzymatic assays. Given these results and taking into account that the β-glucoronidase activity of cell lysates are compared to the activity of a standard stock of β-glucuronidase

stored at -20 °C, we implemented a new protocol adpated from the method of Miller for β-galactosidase assays.[48] This new protocol reduced our day-to-day variability and produced consistent results. We were able to characterize riboswitch E in *S. pyogenes* as a 60-fold inducible system, which exhibits the lowest basal expression and highest fold activation of any inducible system published for *S. pyogenes* to date (Figure 4.9).[28, 29] Additionally, we used this new protocol to repeat the dose-response experiment for riboswitch E (Figure 4.10).

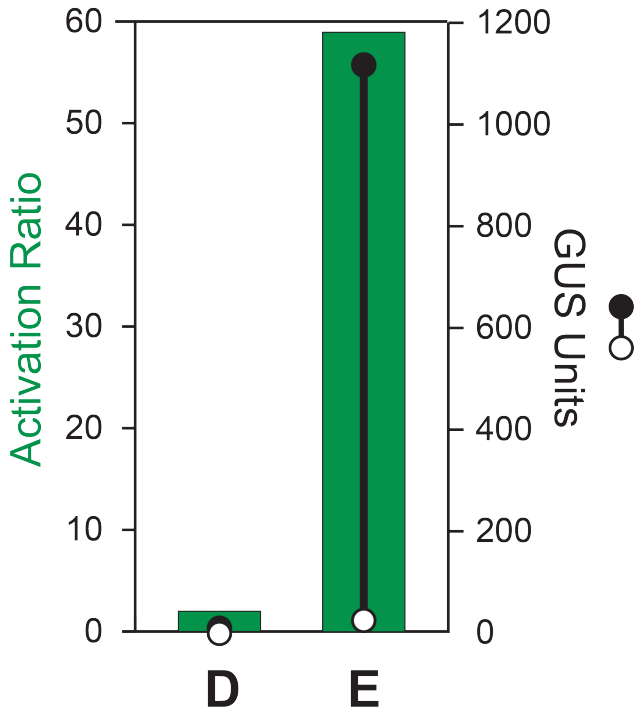

Figure 4.9 - β-Glucuronidase Activities for Riboswitches D and E in *S. pyogenes*. Strains encoding the indicated riboswitches controlling *gusA* were grown in the presence (●) or absence (○) of 2 mM theophylline to stationary phase. Cells were harvested, lysed and assayed for β-glucuronidase activity. The right axis indicates the GUS Units, a measure of β-glucuronidase activity, and the left axis shows the activation ratio (green bar) which is calculated by dividing the Miller units in the presense of theophylline by the units in the absense of theophylline. Standard deviations of triplicate experiments lie within the area of circles.

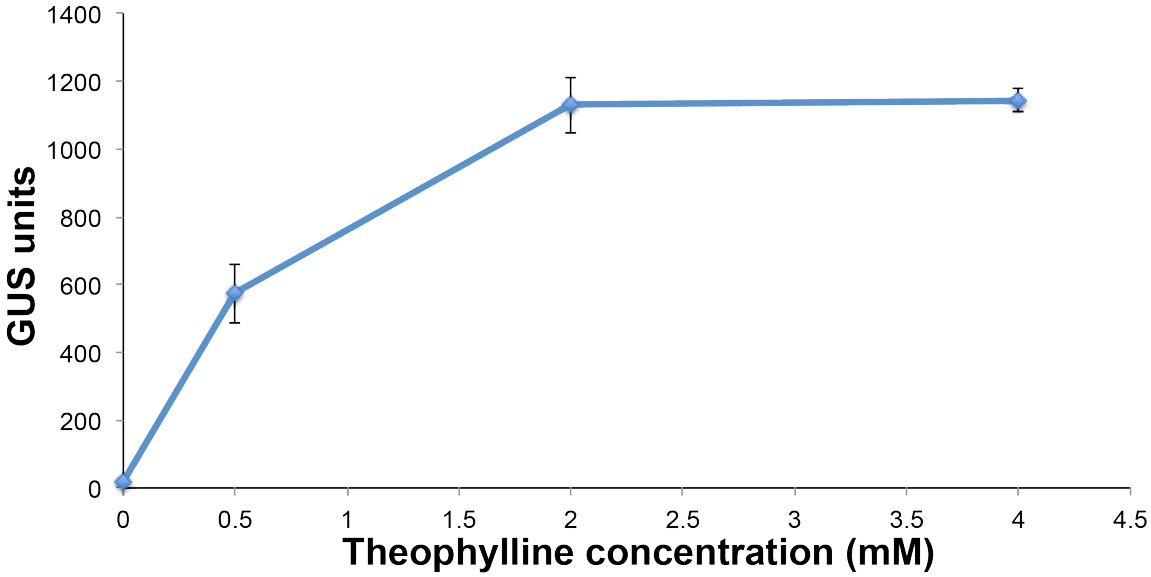

Figure 4.10 – Theophylline-dose-dependent increase of β-glucuronidase activity using improved protocol. Overnight cultures of *S. pyogenes* transformed with riboswitch E upstream of *gusA* were diluted into fresh media containing 0 to 4 mM theophylline. Cultures were grown to stationary phase and assayed for β-glucuronidase activity. The y-axis represents β-glucuronidase activity in GUS units and the x-axis indicates the concentration of theophylline in mM.

In addition to characterizing riboswitch D and E, three previously described riboswitches, A (Chapter 2), B^{38}, and C^{38} were also tested in *S. pyogenes*. By sequence analysis of the RBS found in each of these riboswitches (Figure 4.9), We hypothesized that riboswitches A, B and E would function most efficiently in *S. pyogenes* because sequence analysis revealed that the RBS in each of these closely matched. These riboswitches were cloned into the pEU7742 shuttle plasmid downstream of P_{sag} and upstream of *gusA*. Transformants harboring riboswitches A, B, C, D and E in 0 mM and 2 mM theophylline were assayed in parallel at 2 hours into stationary phase. As demonstrated below, riboswitch E exhibited the most robust expression and largest activation ratio (Figure 4.11). As predicted, riboswitches A and B produced moderate

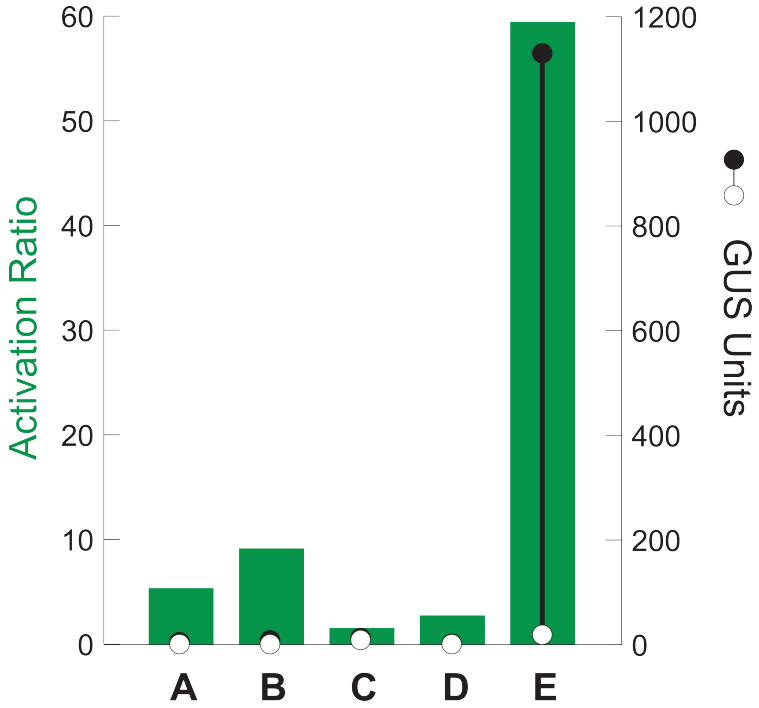

Figure 4.9 - **Figure 4.11 - β-Glucuronidase Activities for several riboswitches in *S. pyogenes*.** Strains encoding the indicated riboswitches controlling *gusA* were grown in the presence (●) or absence (○) of 2 mM theophylline to stationary phase. Cells were harvested, lysed and assayed for β-glucuronidase activity. The right axis indicates the GUS Units, a measure of β-glucuronidase activity, and the left axis shows the activation ratio (green bar) which is calculated by dividing the Miller units in the presense of theophylline by the units in the absense of theophylline. Standard deviations of triplicate experiments lie within the area of circles. Reprinted from Topp.[38]

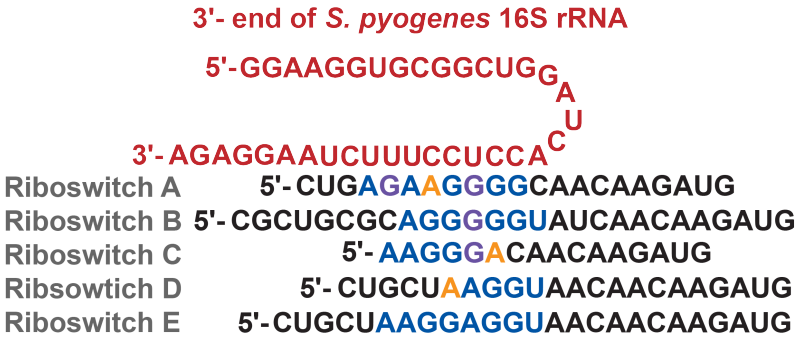

Figure 4.10 – Alignment of *S. pyogenes* 16S rRNA to riboswitch sequences. The 16S rRNA of *S. pyogenes* is shown in red. The sequences of the region 5' to the aptamer of each riboswitch upstream of gusA are shown in black with the RBS highlighted in blue. Gold nucleotides indicate mismatched complementarity to the anti-Shine-Dalgarno sequence within the 16S rRNA. Purple indicates a wobble base pair.

[38] Figure reprinted from Topp et al. *Appl. Environ. Microbiol.* **2010**, 76, 7881-7884. with permission fromAmerican Society for Microbiology.

β-glucuronidase activity.[38] These data confirm that *S. pyogenes* in particular requires high RBS-16S rRNA complementarity in order to achieve strong gene expression.

4.3 Conclusion

We set out to determine whether riboswitches are portable from one species to another, in an effort to avoid development of riboswitches in genetically intractable bacterial species *de novo*. In this chapter we discerned that, while the riboswitch screened from *A. baylyi* (riboswitch A, Chapter 3) functions somewhat in *S. pyogenes*, it is clear that a very strong RBS-16S rRNA interaction must be present to achieve significant induction in this Gram-positive species. We have proven that with a combination of screening methods (Chapter 3) and semi-rational design, we can avoid *de novo* development of synthetic riboswitches. We have demonstrated the ability to modify a pre-existing riboswitch into a functional riboswitch in a phylogenetically unrelated bacterium. Furthermore, the riboswitch developed in this study has been applied to work in *S. pyogenes* to regulate expression of a putative essential gene, *hlp*, that previously could not be studied.[29] Inducible tools available prior to our synthetic riboswitch possessed high basal expression and, therefore, could not repress expression of *hlp* sufficiently enough to demonstrate attenuated growth.[29] However, the use of riboswitch E enabled the authors to demonstrate for the first time that *hlp* is indeed an essential gene in *S. pyogenes*.[29]

While our riboswitches possess some advantages over previously described tools for genetic manipulation, they currently lack the ability to be implemented in studying bacterial pathogenesis during infection. Recently, another inducible system has been shown to work in *S. pyogenes*.[29] Regulation of the tetracycline inducible promoter, P_{tet}, is

mediated by *tetR*, which encodes the *tet* repressor.[49] In the absence of inducer, the *tet* repressor binds to the *tet* operator to inhibit transcription. This system has been adapted from *E. coli* to work in *B. subtilis*, as well as other Gram-positive bacteria, including *Streptococci*.[50-52] A valuable attribute of P_{tet} is the ability to use the system during mammalian cell infection as well as animal infection.[53-55] In the next chapter, we will discuss our efforts to expand the utility of riboswitches to be used to regulate bacterial gene expression during mammalian cell infection.

4.4 Experimental

General Considerations. Synthetic oligonucleotides were purchased from IDT (Coralville, IA). All sequences were verified using DNA sequencing (MWG Operon, Huntsville, Al). Purifications of plasmid DNA, PCR products, and enzyme digestions were performed using kits from Qiagen (Germantown, MD). Reagents, such as Theophylline, *p*-nitrophenyl-β-D-glucuronide (PNPG) and spectinomycin were purchased from Sigma-Aldrich (St. Louis, MO). 5-bromo-4-chloro-indolyl-β-D-glucuronide (X-gluc) was purchased from US Biological (Swampscott, Massachusetts). All cloning was performed in *E. coli* MDS42 cells (Scarab Genomics). *S. pyogenes* strain JRS1278 (M3 serotype) was a kind gift from June Scott (Emory University, Atlanta, GA). All *E. coli* cultures were grown in LB media (EMD Bioscience, Merck KGaA, Darmstadt, Germany). All *S. pyogenes* cultures were grown in Todd-Hewitt Broth supplemented with 0.2% yeast extract (THyB, Difco, Sparks, MD) in a stationary 37 °C water bath.

Toxicity assay. Theophylline solution was made by dissolving 90 mg directly into 50 mL of autoclaved growth media to give a concentration of 10 mM. This solution was then filtered through a 0.2 μm filter prior to use in experiments. The 10 mM solution was

then used to make subsequent dilutions to obtain the desired concentrations of theophylline. GAS strain JRS1278 was grown overnight in THyB. This overnight culture was then used to inoculate pre-warmed THyB supplemented with 0, 1, 2, 5, and 10 mM theophylline. These cultures were grown in Klett flasks in a 37°C water bath with slight shaking. Growth was assayed by optical density with a Klett-Summerson photoelectric colorimeter with the A filter. Cultures were grown for 3 hours before optical density readings were taken every hour for 11 hours. Doubling time was calculated by the following formula: $60\ minutes * (t_y - t_x * \frac{log2}{log\frac{D_y}{D_x}})$, where t_y is the latest time point in exponential phase and t_x is the earliest time point in exponential phase. D_y represents the optical density at time y and D_x the optical density at time x. The formula calculates doubling time in minutes.

Cloning. The parent plasmid for all constructs is pEU7742 (J. V. Bugrysheva and J. R. Scott, unpublished data). A non-standard cloning method was employed to insert the riboswitch downstream of P_{sag} and upstream of *gusA* in the pEU7742 plasmid. This cloning method is termed inverse PCR, where the primers point away from one another.[56] In this method, the sequence to be inserted is integrated into the outward facing primers with no overlaps. PCR is then performed using a high fidelity blunt-end polymerase. Here we use Phusion high-fidelity polymerase (NEB). After completion of PCR the blunt ends are phosphorylated using polynucleotide kinase (PNK), then ligated with T4 DNA ligase to produce a circular plasmid. The plasmid is purified and prepared for transformation.

Transformation. *S. pyogenes* strain JRS1278 was transformed by eletroporation as previously described.[57] JRS1278 is ΔcovR::cat in MGAS315 generated using pJRS1349,

as previously described. Transformatnts were plated onto Thy agar plates supplemented with 2 mM theophylline, X-gluc, and 100 μg/mL spectinomycin. Plates were incubates overnight in a stationary 37 °C incubator with 5% CO_2. Successful transformants were cultured overnight in fresh THyB supplemented with 100 μg/mL spectinomycin in a stationary 37 °C water bath. Glycerol stocks of cultures were made by adding 200 μL of 100% glycerol to 800 μL of overnight culture.

β-glucuronidase assays. Theophylline solution was made as described above. GAS strain JRS1278 containing plasmid pCKR5 was grown overnight in Todd-Hewitt broth containing 100 μg/mL spectinomycin. This overnight culture was then divided in 2.5 mL aliquots into 47.5 mL pre-warmed Todd-Hewitt broth without antibiotic containing theophylline concentrations of 0 mM, 0.001 mM, 0.01 mM, 0.1 mM, 0.25 mM. 0.5 mM, 1 mM, 2 mM, 5 mM, and 9 mM. For a negative control, JRS1278 without plasmid was grown overnight and inoculated as above into 47.5 mL of Todd-Hewitt broth. These inoculated cultures were then grown in a 37°C water bath until 2 hours into stationary phase (Klett = ~150).

Cultures were harvested and set on ice for 10 minutes and subsequently pelleted in a tabletop centrifuge at 3,250 rcf for 10 minutes at 4 ºC. The supernatant was discarded and pellets resuspended and transferred to 1.5 mL eppendorf tubes, pelleted again then stored at -80 ºC. Frozen cell pellets were resuspended in 1 mL of ice-cold Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, and 1 mM $MgSO_4$) then lysed in tubes with a glass bead matrix via vortexing at maximum speed for 10 minutes. Cell debris was pelleted and cell lysates transferred to fresh 1.5 mL tubes. Lysates were assayed by addition of 4-mg/ml solution of *p*-nitrophenyl-β-D-glucuronide in Z buffer and

the OD_{420} was taken every minute for 2 hours (for weaker riboswitches C and D reading was taken every 30 minutes for 18 hours). Gus units in each lysate were determined by comparing sample data to standard curve. Protein concentrations of lysates were determined using a BCA protein assay kit from Thermo. Gus activity was determined by dividing Gus units by protein concentration and multiplying by 100. GUS units were determined by dividing the OD_{420} by protein concentration (in μg/mL), multiplying by 1000 and dividing by the reaction time in minutes

4.5 References

(1) Pichichero, M. E. *Pediatr. Rev.* **1998**, 19, 291-302.

(2) Cunningham, M. W. *Pathogenesis of Group A Streptococcal Infections and Their Sequelae: Hot Topics in Infection and Immunity in Children IV*. **2008**, 29-42.

(3) Krause, R. M. *Clin. Lab. Med.* **2002**, 22, 835-848.

(4) Tart, A. H., Walker, M. J. and Musser, J. M. *Trends Microbiol.* **2007**, 15, 318-325.

(5) Hope-Simpson, R. E. *The Journal of Hygiene*. **1981**, 87, 109-129.

(6) Olsen, R. and Musser, J. *Annual Reviews in Pathology: Mechanisms of Disease*. **2010**, 5, 1-31.

(7) Pichichero, M. E. and Casey, J. R. *Otolaryngology -- Head and Neck Surgery*. **2007**, 137, 851-857.

(8) Kaplan, E. L. and Johnson, D. R. *Pediatrics*. **2001**, 108, 1180-1186.

(9) VanArsdel, P. P. *JAMA: The Journal of the American Medical Association*. **1965**, 191, 238-239.

(10) O'Brien, K. L., Beall, B., Barrett, N. L., Cieslak, P. R., Reingold, A., Farley, M. M., Danila, R., Zell, E. R., Facklam, R., Schwartz, B., Schuchat, A. and Network, A. B. C. S. E. I. P. *Clin. Infect. Dis.* **2002**, 35, 268-276.

(11) Musser, J. M. and DeLeo, F. R. *The American Journal of Pathology*. **2005**, 167, 1461-1472.

(12) Hoch, J. A. *Curr. Opin. Microbiol.* **2000**, 3, 165-170.

(13) Stock, A. M., Robinson, V. L. and Goudreau, P. N. *Annu. Rev. Biochem.* **2000**, 69, 183-215.

(14) Churchward, G. *Mol. Microbiol.* **2007**, 64, 34-41.

(15) Graham, M. R., Smoot, L. M., Migliaccio, C. A. L., Virtaneva, K., Sturdevant, D. E., Porcella, S. F., Federle, M. J., Adams, G. J., Scott, J. R. and Musser, J. M. *Proc. Natl. Acad. Sci.* **2002**, 99, 13855-13860.

(16) Levin, J. C. and Wessels, M. R. *Mol. Microbiol.* **1998**, 30, 209-219.

(17) Federle, M. J., McIver, K. S. and Scott, J. R. *J. Bacteriol.* **1999**, 181, 3649-3657.

(18) Gusa, A. A., Gao, J., Stringer, V., Churchward, G. and Scott, J. R. *J. Bacteriol.* **2006**, 188, 4620-4626.

(19) Dalton, T. L. and Scott, J. R. *J. Bacteriol.* **2004**, 186, 3928-3937.

(20) Treviño, J., Perez, N., Ramirez-Peña, E., Liu, Z., Shelburne, S. A., Musser, J. M. and Sumby, P. *Infect. Immun.* **2009**, 77, 3141-3149.

(21) Lynskey, N. N., Lawrenson, R. A. and Sriskandan, S. *Curr. Opin. Infect. Dis.* **2011**, 24, 196-202 110.1097/QCO.1090b1013e3283458f3283457e.

(22) Sugareva, V., Arlt, R., Fiedler, T., Riani, C., Podbielski, A. and Kreikemeyer, B. *BMC Microbiol.* **2010**, 10, 34.

(23) Moran, C. P., Lang, N., LeGrice, S. F. J., Lee, G., Stephens, M., Sonenshein, A. L., Pero, J. and Losick, R. *Molecular and General Genetics MGG*. **1982**, 186, 339-346.

(24) Haldenwang, W. G. *Microbiol. Rev.* **1995**, 59, 1-30.

(25) Kuipers, O. P., Beerthuyzen, M. M., de Ruyter, P. G. G. A., Luesink, E. J. and de Vos, W. M. *J. Biol. Chem.* **1995**, 270, 27299-27304.

(26) van der Meer, J. R., Polman, J., Beerthuyzen, M. M., Siezen, R. J., Kuipers, O. P. and De Vos, W. M. *J. Bacteriol.* **1993**, 175, 2578-2588.

(27) Engelke, G., Gutowski-Eckel, Z., Kiesau, P., Siegers, K., Hammelmann, M. and Entian, K. D. *Appl. Environ. Microbiol.* **1994**, 60, 814-825.

(28) Eichenbaum, Z., Federle, M. J., Marra, D., de Vos, W. M., Kuipers, O. P., Kleerebezem, M. and Scott, J. R. *Appl. Environ. Microbiol.* **1998**, 64, 2763-2769.

(29) Bugrysheva, J. V., Froehlich, B. J., Freiberg, J. A. and Scott, J. R. *Appl. Environ. Microbiol.* **2011**, 77, 4422-4428.

(30) Lynch, S. A., Desai, S. K., Sajja, H. K. and Gallivan, J. P. *Chem. Biol.* **2007**, 14, 173-184.

(31) Lynch, S. A. and Gallivan, J. P. *Nucleic Acids Res.* **2009**, 37, 184-192.

(32) Wickiser, J. K., Winkler, W. C., Breaker, R. R. and Crothers, D. M. *Mol. Cell*. **2005**, 18, 49-60.

(33) Mandal, M., Boese, B., Barrick, J. E., Winkler, W. C. and Breaker, R. R. *Cell*. **2003**, 113, 577-586.

(34) Desai, S. K. and Gallivan, J. P. *J. Am. Chem. Soc.* **2004**, 126, 13247-13254.

(35) Nizet, V., Beall, B., Bast, D. J., Datta, V., Kilburn, L., Low, D. E. and De Azavedo, J. C. S. *Infect. Immun.* **2000**, 68, 4245-4254.

(36) Betschel, S. D., Borgia, S. M., Barg, N. L., Low, D. E. and De Azavedo, J. C. S. *Infect. Immun.* **1998**, 66, 1671-1679.

(37) Lambert, P. A. *J. Appl. Microbiol.* **2002**, 92 Suppl, 46S-54S.

(38) Topp, S., Reynoso, C. M. K., Seeliger, J. C., Goldlust, I. S., Desai, S. K., Murat, D., Shen, A., Puri, A. W., Komeili, A., Bertozzi, C. R., Scott, J. R. and Gallivan, J. P. *Appl. Environ. Microbiol.* **2010**, 76, 7881-7884.

(39) Barnett, T. C., Bugrysheva, J. V. and Scott, J. R. *J. Bacteriol.* **2007**, 189, 1866-1873.

(40) Bugrysheva, J. V. and Scott, J. R. *RNA Biol.* **2010**, 7, 569-572.

(41) Shine, J. and Dalgarno, L. *Proc. Natl. Acad. Sci.* **1974**, 71, 1342-1346.

(42) Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B. S. and Stormo, G. *Annu. Rev. Microbiol.* **1981**, 35, 365-403.

(43) Vellanoweth, R. L. and Rabinowitz, J. C. *Mol. Microbiol.* **1992**, 6, 1105-1114.

(44) Band, L. and Henner, D. J. *DNA*. **1984**, 3, 17-21.

(45) Chen, H., Bjerknes, M., Kumar, R. and Jay, E. *Nucleic Acids Res.* **1994**, 22, 4953-4957.

(46) Sharp, J. S. and Bechhofer, D. H. *Mol. Microbiol.* **2005**, 57, 484-495.

(47) Arnold, T. E., Yu, J. and Belasco, J. G. *RNA*. **1998**, 4, 319-330.

(48) Miller, J. H. *Experiments in Molecular Genetics*. **1972**, 352-355.

(49) Hillen, W. and Berens, C. *Annu. Rev. Microbiol.* **1994**, 48, 345-369.

(50) Geissendörfer, M. and Hillen, W. *Appl. Microbiol. Biotechnol.* **1990**, 33, 657-663.

(51) Liu, D., Yumoto, H., Murakami, K., Hirota, K., Ono, T., Nagamune, H., Kayama, S., Matsuo, T. and Miyake, Y. *Mol. Microbiol.* **2008**, 68, 1268-1282.

(52) Wang, B. and Kuramitsu, H. K. *Infect. Immun.* **2005**, 73, 3568-3576.

(53) Bateman, B. T., Donegan, N. P., Jarry, T. M., Palma, M. and Cheung, A. L. *Infect. Immun.* **2001**, 69, 7851-7857.

(54) Gandotra, S., Schnappinger, D., Monteleone, M., Hillen, W. and Ehrt, S. *Nat. Med.* **2007**, 13, 1515-1520.

(55) Ji, Y., Marra, A., Rosenberg, M. and Woodnutt, G. *J. Bacteriol.* **1999**, 181, 6585-6590.

(56) Ochman, H., Gerber, A. S. and Hartl, D. L. *Genetics*. **1988**, 120, 621-623.

(57) Caparon, M. G. and Scott, J. R. *Genetic manipulation of pathogenic streptococci*. **1991**, 556-586.

Chapter 5 - Intracellular Study of Genes Involved in *Francisella* Pathogenesis

5.1 Introduction

The use of genetic control elements is fundamental for studying and understanding gene function in many species of bacteria. For common model organisms such as *Escherichia coli* and *Bacillus subtilis,* many genetic tools exist. [1-3] However, the same is not true for less genetically tractable organisms.[4] Recently, we have described a set of broad-host range synthetic riboswitches that allow induction of gene expression in diverse bacterial species (Chapters 2, 3 and 4).[5]

Riboswitches are RNA regulatory elements, usually located in the 5'-untranslated region (UTR) of mRNA. Since riboswitches do not require the presence of accessory proteins and function independently of promoter identity, they can bypass several pitfalls often encountered with exogenous expression of genetic control systems.[6] In addition, riboswitches can be coupled with native promoters to drive expression,[5] avoiding the incompatibility of foreign genetic regulation systems.

Francisella tularensis is a Gram-negative intracellular bacterial pathogen lacking genetic tools like inducible gene expression systems. *F. tularensis* is the causative agent of the zoonotic disease tularemia which is characterized by flu-like symptoms and exhibits a mortality rate of 30% [7] to 60% for the pneumonic form of the disease.[8] *Francisella novicida* is a closely related species to *F. tularensis*, although it is rarely pathogenic to immunocompetent humans. *F. novicida* causes a tularemia-like disease in mice,[9] possesses a genome that is 98% genetically identical to *F. tularensis*[10, 11] and shares many virulence genes with *F. tularensis*. Additionally, *F. novicida,* like *F. tularensis*, exhibits the important virulence trait of replicating within host cells such as

macrophages.[9, 12] It is therefore often used as a model system to study *F. tularensis* biology.

A possible explanation for the difficulty of transporting genetic tools from other organisms may lie partly in the unique α-subunits of the RNA polymerase (RNAP) of *Francisella*. The α-subunit serves as the initiator for RNAP assembly and binds to a supplementary promoter element (the UP-element) upstream of the -35 recognition site [13]. In most bacteria, the RNAP holoenzyme contains a dimer of identical α-subunits [14]. However, in the *Francisella* genus, the α-subunits are coded by two distinct genes, yielding two unique variants of the RNAP α-subunit [15]. Given the fundamental role of the α-subunit in transcriptional regulation [16], the inefficient transcription of foreign promoters in *Francisella* may be attributed to the differences in RNAP. Therefore, it would be beneficial to develop a genetic tool for *Francisella* that does not rely on transcription by a non-native promoter. Riboswitches offer such a tool since they are readily coupled to native promoters [5].

Relatively little is known about the virulence genes required for *Francisella* pathogenesis, and the lack of genetic tools to modulate gene expression has hampered the investigation of gene functions.[17] Attempts to employ pre-existing tools from other organisms in *Francisella* have not been fruitful.[18] To date, only one repressible promoter system (a native system) and no inducible systems exist to study *Francisella* genetics.[17, 19] Riboswitches do not require the presence of accessory proteins to function [6] and native promoters can be coupled with riboswitches to drive expression,[5] avoiding the incompatibility of foreign genetic regulation systems. We therefore set out to determine whether riboswitches could be used to modulate gene expression in *Francisella spp*.

In this study, we demonstrated the efficacy of theophylline-sensitive synthetic riboswitches in regulating gene expression in *F. novicida* and highly virulent *F. tularensis* by using riboswitch-controlled LacZ as a reporter. To further the utility of synthetic riboswitches, we verify that this system can be used to induce reporter genes expressed in the bacteria during intracellular infection of host cells. We also demonstrate the ability to regulate expression of a novel gene, *FTN_0818*, required for intracellular replication [20] and *in vivo* pathogenesis[21, 22] and show that such control can functionally rescue the growth defect of a strain not expressing *FTN_0818* in minimal media. Furthermore, riboswitch-mediated induction of FTN_0818 during macrophage infection restored the ability of a mutant strain to replicate. This is the first demonstration that a synthetic riboswitch can be used to efficiently control an endogenous virulence gene and rescue a virulence trait in an intracellular pathogen. The ability to induce expression of pathogenic genes of intracellular bacteria using synthetic riboswitches provides a valuable tool for more in-depth and careful investigation of gene function within pathogenic organisms and is likely applicable to other intracellular bacterial pathogens.

5.2 Results and Discussion

To determine if any of our previously described riboswitches[5, 23, 24] were functional in *F. novicida*, the riboswitches were coupled to a *lacZ* reporter gene (β-galactosidase) driven by a strong *Francisella* promoter, P_{gro} (the *groEL* promoter [25]) and cloned into the *Francisella-E. coli* shuttle vector, pFNLTP6 [26]. All reporter constructs exhibited at least a 5−fold increase in β-galactosidase activity in the presence of theophylline (Figure 5.1). Not surprisingly, riboswitch E (E-Rs-LacZ), which was rationally designed with the consensus ribosome binding site (Chapter 4), exhibited a

large dynamic range in *F. novicida*, calculated by subtracting expression in the absence of theophylline from the expression in the presence of theophylline (Figure 5.1). Riboswitch E exhibits perfect complementarity to the 16S rRNA of *F. novicida* (Figure 5.2), which leads to high levels of translation. Riboswitch F (F-Rs-LacZ), selected from a library of randomized sequences,[23] exhibited the lowest basal expression levels in the absence of theophylline (Figure 5.1). This can be predicted from the lack of complementarity between riboswitch F and the 16S rRNA of *F. novicida* (Figure 5.2).

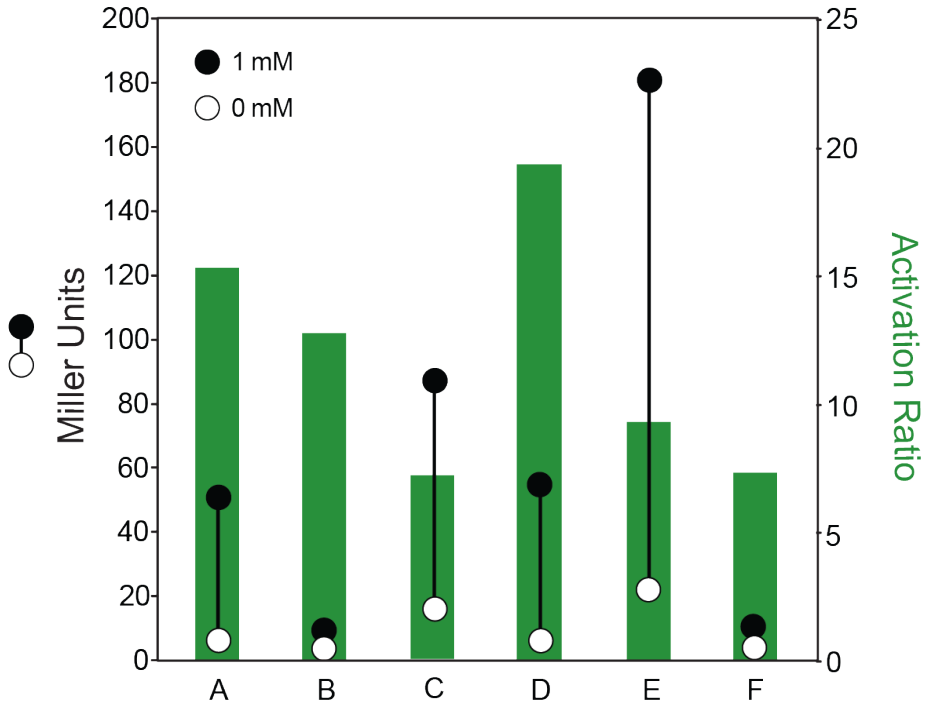

Figure 5.1- Theophylline-dependent synthetic riboswitches controlling *lacZ* in *F. novicida*.
Strains encoding the indicated riboswitches controlling *lacZ* were grown in the presence (●) or absence (○) of 1 mM theophylline to an OD_{600} between 0.7 and 0.8. Cells were harvested, lysed and assayed for β-galactosidase activity. The left axis indicates the Miller Units, a measure of β-galactosidase activity, and the right axis shows the activation ratio (green bar) which is calculated by dividing the Miller units in the presense of theophylline by the units in the absense of theophylline. Standard error of Miller Units lie within the area of circles. The data are representative of three independent experiments performed in triplicate.

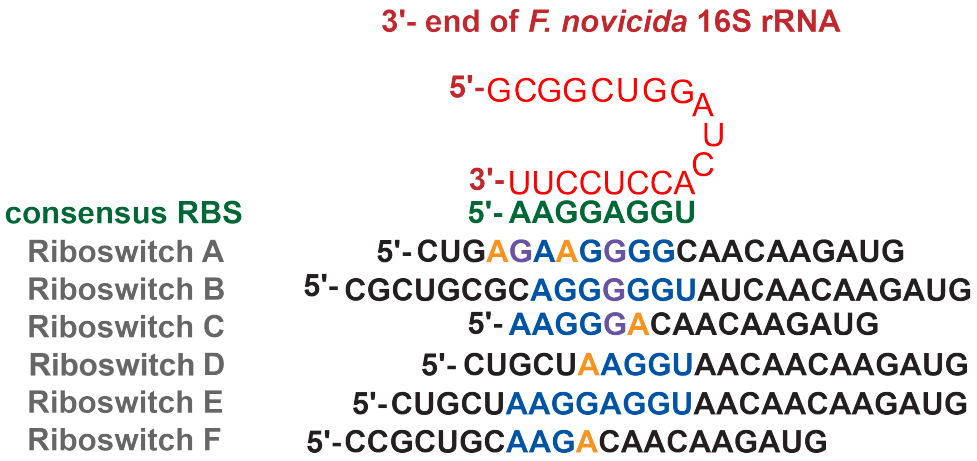

Figure 5.2 – Alignment of *F. novicida* 16S rRNA to riboswitch sequences. The 16S rRNA of *S. pyogenes* is shown in red. The sequences of the region 5' to the aptamer of each riboswitch are shown in black with the RBS highlighted in blue. Gold nucleotides indicate mismatched complementarity to the anti-Shine-Dalgarno sequence within the 16S rRNA. Purple indicates a wobble base pair.

Strains E-Rs-LacZ and F-Rs-LacZ were selected for further study because they exhibited the highest expression in the presence of theophylline and the lowest expression in the absence of theophylline respectively. Each of these strains demonstrated ligand-dependent induction of β-galactosidase activity in a dose-dependent manner (Figure 5.3), which is consistent with previous observations in other organisms. [5, 27, 28] Importantly, induction of gene expression can be obtained with as little as 0.1 mM theophylline (Figure 5.3).

After verifying that the riboswitches function in *F. novicida* in liquid media, we tested the feasibility of regulating a reporter gene (*gfp*) in *F. novicida* while infecting macrophages. We hypothesized that strong expression levels would be required to detect fluorescence. With this in mind riboswitch E, having the highest induced expression levels (E-Rs-LacZ, Figure 5.1), was placed upstream of a *gfp* reporter gene in the pFNLTP6 plasmid under the control of P_{gro}, then transformed into *F. novicida,* to obtain strain E-Rs-GFP. As controls, we used strains constitutively expressing GFP (GFP-pos,

Figure 5.4, Panel B) or a promoterless GFP (GFP-neg, Figure 5.4, Panel A). GFP-neg and E-Rs-GFP in the absence of theophylline (Figure 5.4, Panel C) both represent conditions wherein GFP expression is absent. In Panels A and C the bacterial DNA, as well as the macrophage nuclear DNA, was stained by DAPI (blue), indicating the presence of *F. novicida* cells. Notably, there is no corresponding GFP fluorescence localized to the indicated bacterial cells, confirming that the lack of visible GFP is not

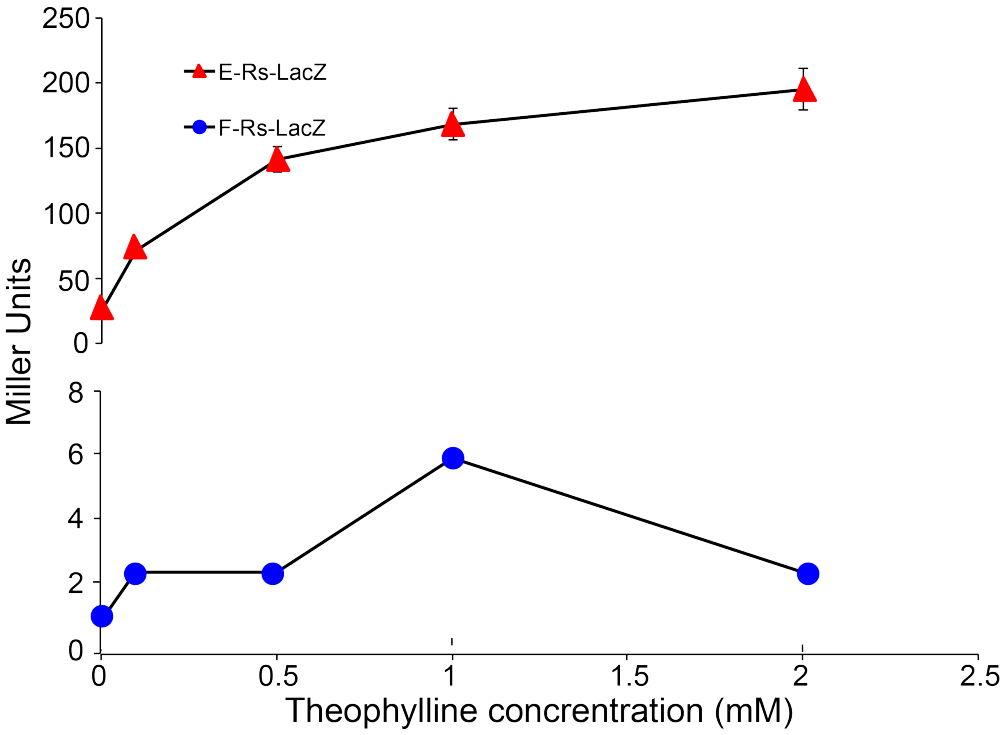

Figure 5.3 - Dose-dependent riboswitch-mediated induction of β-galactosidase
Strains E-Rs-LacZ (▲) and F-Rs-LacZ (●) were subcultured and grown to exponential phase (OD_{600} between 0.7 and 0.8) in media containing 0, 0.1, 0.5, 1 and 2 mM theophylline. β-galactosidase activity was measured in Miller Units. Data are representative of two independent experiments performed in triplicate.

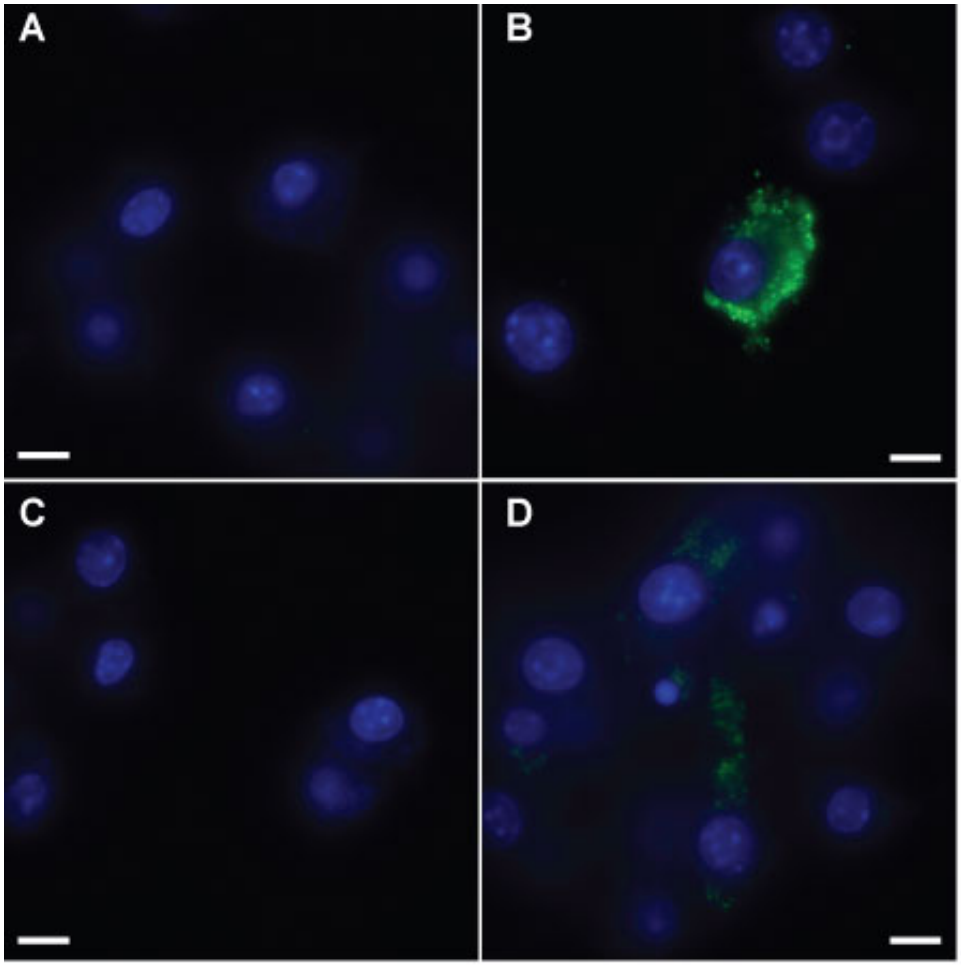

Figure 5.4 - Riboswitch-mediated control of GFP in *F. novicida* during macrophage infection.
RAW264.7 murine macrophages were infected at a multiplicity of infection (MOI) of 50:1 with *F. novicida* harboring *gfp* lacking a promoter (GFP-neg; Panel A), constitutively expressing *gfp* (GFP-pos; Panel B), or the E-Rs-GFP riboswitch contruct (Panels C and D). At 30 minutes post-infection, media without (Panels A-C) or with 1 mM theophylline (Panel D) was added, and macrophages were incubated at 37°C for 24 hours and fixed. GFP (green) and DAPI-stained macrophage nuclei (blue) are shown. Magnification is 100x and scale bar represents 10 μm. Data are representative of three independent experiments in which at least 10 fields of view were analyzed for each condition.

due to absence of bacteria. Importantly, the E-Rs-GFP strain was able to produce visible GFP expression in the presence of 1 mM theophylline, confirming induction of gene expression in intracellular bacteria (Figure 5.4, Panel D). These data demonstrate the

ability of the theophylline-sensitive riboswitch to regulate expression of a bacterial reporter gene during mammalian cell infection.

A possible shortcoming of this system lies in whether or not theophylline is permeable to the macrophage as well as able to access the phagosome. We believe results from the GFP experiments indicate that theophylline is indeed able to enter macrophages. When inducing GFP expression intramacrophage, theophylline is added after extracellular bacterial cells have been removed from the infection media. Therefore, in order for theophylline to induce visible GFP expression as demonstrated in Figure 5.4, theophylline must be able to enter the macrophages and access the intracellular bacterial cells. While this result indicates that theophylline can enter macrophages, it does not verify the permeability of theophylline through the phagosomal membrane. We will discuss this later in the chapter.

Having demonstrated the ability to control an exogenous gene, *gfp*, we proceeded to determine if an endogenous gene could be controlled in a theophylline-dependent manner. *FTN_0818* is required for *F. novicida* replication in murine macrophages (RAW264.7).[20] Unpublished studies also indicate that mutant strains exhibit impaired growth in minimal media (Napier and Weiss, in preparation). This phenotype lends itself to a simple liquid culture assay to determine whether or not FTN_0818 is present in sufficient concentrations to allow wild type growth rates. The *FTN_0818* gene was first placed downstream of riboswitch E (E-Rs-*FTN_0818*) and integrated into the chromosome of a Δ*FTN_0818* strain of *F. novicida* via allelic exchange. The E-Rs-*FTN_0818* strain was grown in minimal media and the OD_{600} was measured every hour for 18 hours to determine growth rate. The results of this experiments showed that even

in the absence of theophylline, basal expression of FTN_0818 was sufficient to elicit a wild type growth curve (Figure 5.5). In an effort to reduce basal expression of *FTN_0818*, it was placed under the control of riboswitch F (F-Rs-*FTN_0818*), which exhibited much lower basal expression in the β-galactosidase assays (Figure 5.6). Indeed, in the absence of theophylline, the growth rate of strain F-Rs-*FTN_0818* was nearly identical to that of Δ*FTN_0818* (Figure 5.6). In the presence of 1 mM theophylline, however, the growth curve of F-Rs-*FTN_0818* resembled that of wild-type *F. novicida* in the presence of 1 mM theophylline (Figure 5.6) indicating restoration of a wild type growth curve. These data suggest that riboswitch F, in the absence of theophylline,

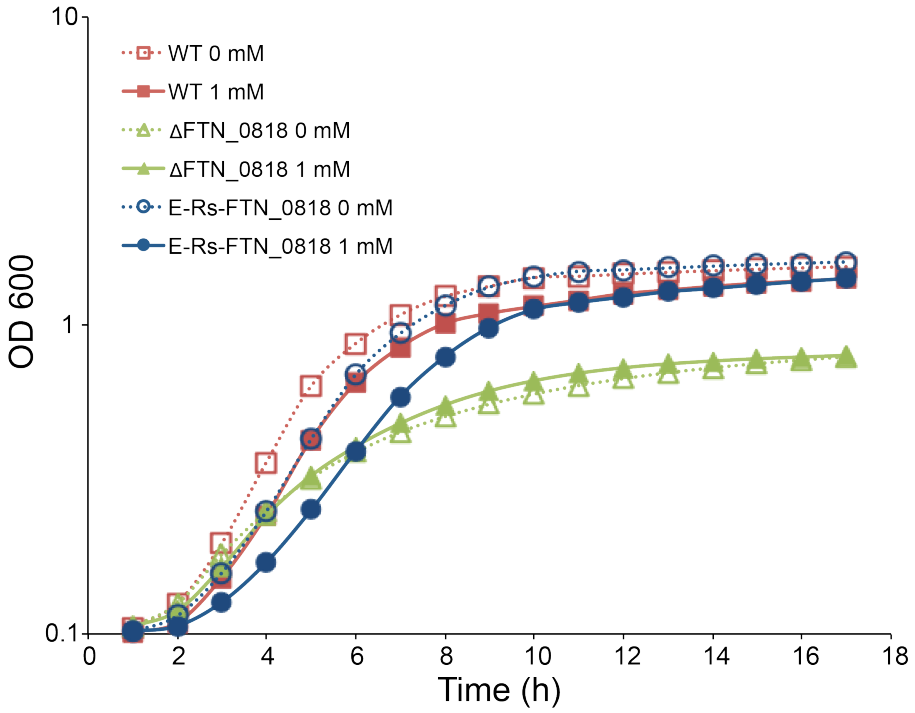

Figure 5.5 - Riboswitch E controlling *FTN_0818* in minimal media.
Growth rates are represented by cell density measured at OD_{600} every hour for 18 hours. Wildtype (,), Δ*FTN_0818* (,) and E-Rs-*FTN_0818* (,)strains were grown overnight (~18 hours) in Tryptic Soy Broth (TSB). Cultures were washed then diluted in Chamberlain's minimal media to OD_{600} = 0.03 in the absence (, ,) or presence (, ,) of 1 mM theophylline. Standard deviation of triplicate samples lie within the area of the symbol.

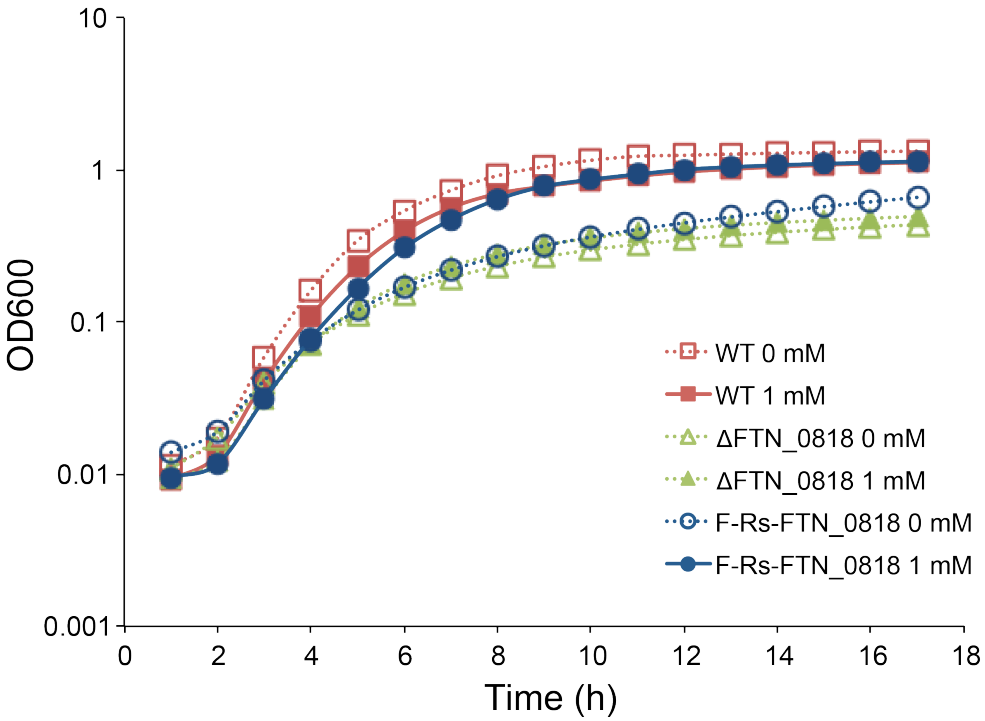

Figure 5.6 - Riboswitch F controlling *FTN_0818* in minimal media. Growth rates are represented by cell density measured at OD_{600} every hour for 18 hours. Wildtype (,), Δ*FTN_0818* (,) and E-Rs-*FTN_0818* (,)strains were grown overnight (~18 hours) in Tryptic Soy Broth (TSB). Cultures were washed then diluted in Chamberlain's minimal media to OD_{600} = 0.03 in the absence (, ,) or presence (, ,) of 1 mM theophylline. Standard deviation of triplicate samples lie within the area of the symbol.

prevents expression of *FTN_0818*, while, in the presence of 1 mM theophylline, riboswitch F permits sufficient expression of *FTN_0818* to elicit a wild type phenotype.

Unpublished results demonstrate that the Δ*FTN_0818* mutant is unable to replicate inside macrophages (Napier and Weiss, in preparation). To determine whether theophylline-induced gene expression of *FTN_0818* can rescue this phenotype during mammalian cell infection, *FTN_0818* was placed under the control of riboswitches E and F and integrated into the chromosome by allelic exchange. The control strains (Δ*FTN_0818* and wild-type) and the riboswitch strains (E-Rs-*FTN_0818* and F-Rs-*FTN*_0818) were used to infect RAW264.7 murine macrophages. Strain F-Rs-*FTN_0818*

exhibited replication of ~5-fold in the absence of theophylline, similar to that of the negative control (Δ*FTN_0818*) and demonstrated increased replication levels with increasing concentrations of theophylline, but replication never exceeded 10-fold (Figure 5.7). The E-Rs-*FTN_0818* strain, however, exhibited only 10-fold replication in the absence of theophylline but replication rates of ~50-fold at 2 mM theophylline, which is

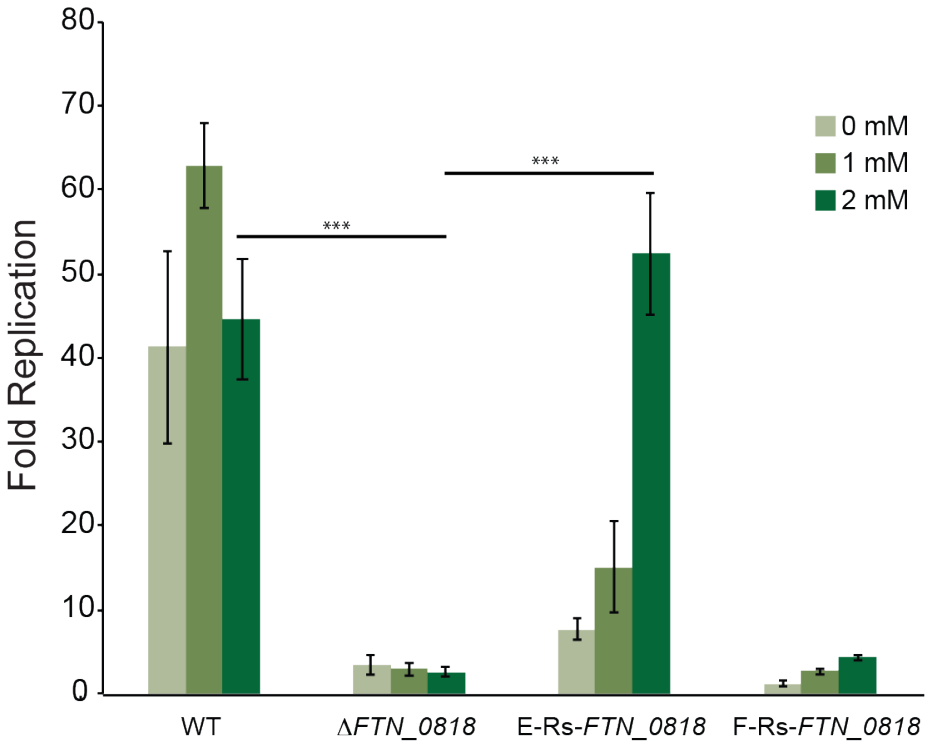

Figure 5.7 - Riboswitch-mediated control of *FTN_0818* facilitates intracellular replication. RAW264.7 murine macrophages were infected with the indicated *F. novicida* strains at an MOI of 20:1 in the presence of 0 (■), 1 (■) or 2 (■) mM theophylline. Macrophages were lysed at 30 minutes and 24 hours postinfection and plated to enumerate intracellular colony forming units (CFUs). Fold replication was calculated by dividing the CFUs at 24 hours post-infection by the number of CFUs at 30 minutes. Bars represent the average fold replication of each strain with error bars depicting the standard deviation. The data are representative of four independent experiments. Asterisks indicate significant difference in fold replication as compared to Δ*FTN_0818* in the presence of 2 mM theophylline. (***, $p<0.005$).

comparable to wild-type replication levels of ~50-fold (Figure 5.7). The results garnered by riboswitch E prove that expression of a native gene can be induced to provide ligand-dependent replication of this intracellular pathogen. These data demonstrate the first reported instance of using synthetic riboswitches to control bacterial expression of a pathogenic gene during mammalian cell infection and provide the first step in characterizing virulence genes intracellularly for *F. novicida*.

Interestingly, when *FTN_0818* gene expression is induced with theophylline in liquid culture, riboswitch F (F-Rs-*FTN_0818*) is sufficient to elicit a wild type phenotype (Figure 5.6). However, in macrophages, the E-Rs-*FTN_0818* strain, possessing a switch with greater dynamic range, is necessary to allow normal replication levels in the presence of theophylline (Figure 5.7). In liquid culture, when riboswitch E is upstream of *FTN_0818*, even in the absence of theophylline the strain exhibits a wild type phenotype (Figure 5.5). Theoretically, the basal expression of *FTN_0818* in the absence of theophylline should be identical within the bacterial cell whether in liquid media or intracellular conditions for each riboswitch. The fact that background expression of *FTN_0818* in strain E-Rs-*FTN_0818* is sufficient to elicit a wild type phenotype in liquid culture but not during infection of the macrophage suggests that a higher concentration of FTN_0818 is required intracellularly versus in liquid culture. *FTN_0818* is hypothesized to be important in biotin synthesis and is required for survival in nutrient-limiting conditions (Napier and Weiss, in preparation). The requirement of additional FTN_0818 may be due to the additional stress imposed upon *F. novicida* while persisting in the harsh environment of the phagosome. However, it is possible that upon engulfment by

the macrophage, the expression profile of *FTN_0818* is altered, which could also alter background levels of expression in the riboswitch strains.

It is possible that the theophylline-dependent induction of *FTN_0818* depicted in Figure 5.7 is due to theophylline being endocytosed with the bacterial cells. Unpublished results indicate that *FTN_0818* is needed early in pathogenesis (Napier and Weiss, in preparation); therefore theophylline was added immediately to the infection media in an effort to induce *FTN_0818* early. Though no reports have been published regarding theophylline permeability across the phagosomal membrane, a related purine molecule, 2-methyl-adenosine has been shown to have intracellular anti-microbial activity towards *Mycobacteria.* [29] Importantly, *Mycobacteria,* after being endocytosed by macrophages, remain and replicate within the phagosome.[30, 31] These combined reports appear to indicate that a purine molecule can indeed cross the phagosomal membrane to affect bacterial cells residing there. However, there is also evidence that *Mycobacteria* cause increased permeability within the resident phagosome,[32] making it unclear whether the phagosomal permeability of 2-methyl-adenosine is dependent on factors exogenous to the phagosome.

It is also possible that theophylline may cross the phagosomal membrane, but be subsequently degraded by xanthine oxidase (XO), an enzyme that catalyzes the oxidation of xanthines, a class of molecules that includes theophylline.[33] However, to date, only microbial XO has been reported to metabolize theophylline, with optimum efficiency at pH 6.5-8.5.[34] The pH within the phagosome is acidic and therefore would not be conducive to bacterial XO-mediated metabolism of theophylline.[35] The pH of the

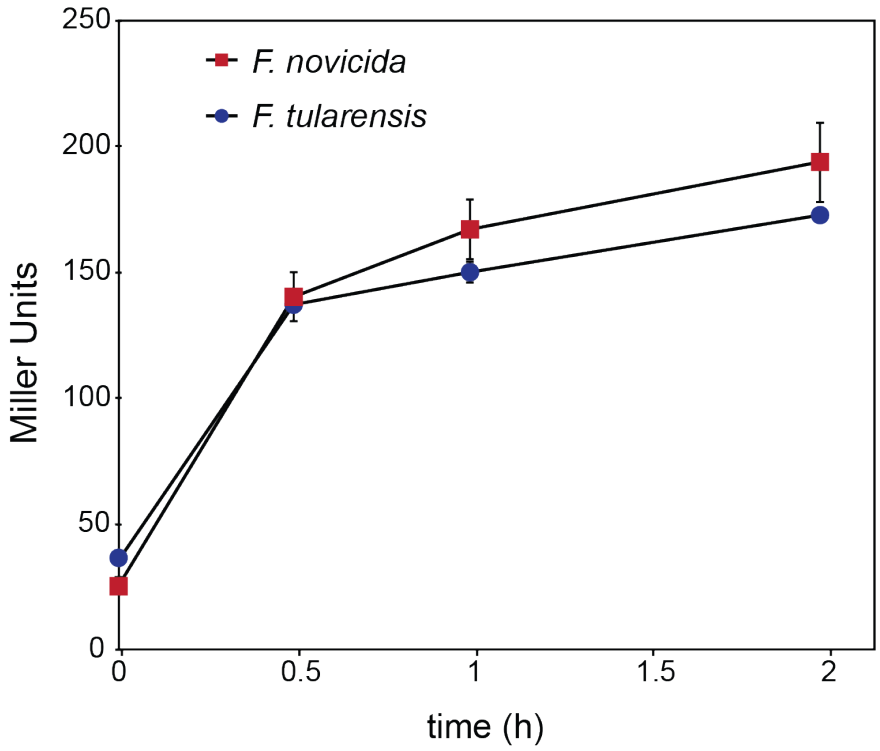

Figure 5.8 - Riboswitch E controlling β-galactosidase activity in *F. novicida* and *F. tularensis*.
Strains E-Rs-LacZ, *F. novicida* (-■-) and E-Rs-LacZ, *F. tularensis* (-●-) were subcultured and grown to exponential phase ($OD_{600} \approx 0.7 – 0.8$) in media containing 0, 0.5, 1 or 2 mM theophylline. β-galactosidase activity was measured in Miller Units. Data are representative of two independent experiments performed in triplicate. Error bars represent standard error of measured Miller Units.

macrophage cytosol is ~7.2 and falls within the range of 6.5-8.5.[36] However, there is still no report of theophylline as a substrate for macrophage XO.

To demonstrate that the functionality of these synthetic riboswitches is not limited to the *F. novicida* species of *Francisella*, we examined whether we could observe similar results in *F. tularensis*, a species pathogenic to humans. A plasmid carrying *lacZ* downstream of riboswitch E, the most robust riboswitch reported here (Figure 5.1), was transformed into *F. tularensis* (SchuS4). As depicted in Figure 5.8, riboswitch E behaves similarly in *F. tularensis* as *F novicida*. In both strains the riboswitch induces gene expression in a dose dependent manner and to similar levels. This result demonstrates

that the utility of these synthetic riboswitches is portable to the BSL-3 strain of *Francisella*.

5.3 Conclusion

In this study, we have demonstrated expansion of the utility of synthetic riboswitches as a novel technique to regulate the gene expression of intracellular bacteria. Previously, the genetic toolbox used to study gene function in *Francisella* consisted of only shuttle plasmids, strong native promoters, suicide vectors and transposon mutagenesis. [18, 25, 26, 37] Here we present a new inducible tool to be used both in liquid culture and intracellularly to characterize gene functions in *Francisella*. To our knowledge, only one regulatory system exists, which uses glucose to repress gene expression in *Francisella.*[17] Although this system holds the advantage of being functional intracellularly, its two main disadvantages are that gene expression is not fully repressed and that the system requires alternate media lacking glucose.[17] Additionally, when adjusting glucose concentrations to control gene expression, there may be pleiotropic affects since an increase in glucose concentration is a common signal for upregulating virulence factors in many bacterial species.[38] Our synthetic riboswitches are induced by a non-endogenous small molecule, theophylline, offering an orthogonal system for gene regulation that minimizes interactions with natural biochemical pathways and indeed, does not appear effect wild type *F. novicida* in any of the experiments shown here.

Identifying and characterizing virulence genes in pathogenic organisms is a priority for public health reasons. Studying these genes in the context of infection is essential to developing new treatments and preventative measures. Our study reveals that

riboswitches are useful tools for the study of pathogenesis in bacteria that have few, if any, genetic regulatory tools. Recently, a theophylline-dependent synthetic riboswitch was used to determine that *hlp*, a gene encoding a potential drug target, is essential in the pathogenic bacterium *Streptococcus pyogenes*.[4] This finding demonstrates the utility of synthetic riboswitches in elucidating the function of essential and pathogenic genes. The ability of riboswitches to induce gene expression during infection of mammalian cells may provide insight into the mechanisms of pathogenesis of *Francisella* and other bacterial species. Seeliger *et al*. demonstrated theophylline-sensitive riboswitch-mediated induction of GFP in *Mycobacterium* cells during a macrophage infection.[39] We further explored the utility of synthetic riboswitches by regulating expression of endogenous genes necessary for pathogenesis during mammalian cell infection. Riboswitches offer a simple method to regulate bacterial gene expression by addition of an inexpensive small molecule directly to media for intracellular or extracellular studies.

Additionally, synthetic riboswitches are useful for conducting time course experiments to understand the function of pathogenic genes. We have designed an experiment wherein we induce expression of FTN_0818 at different time points of infection. Unpublished results suggest that FTN_0818 is essential for escape from the phagosome and is required from time 0 to 1 hour post-infection (Napier and Weiss, in preparation). There is, theoretically, a critical time point after which induction of the gene will no longer elicit a wild type phenotype. We hypothesize that this time point is at time 0 or 30 minutes post-infection. We will induce expression of FTN_0818 prior to infection, at time points early in infection and at time points late in infection. If FTN_0818 is required prior to 1 hour post-infection, then replication will not be rescued

if FTN_0818 is induced at after this critical time point. We expect that pre-induction and induction at time 0 will allow full recovery of the wild type replication phenotype and that induction occurring after time 0 will exhibit an attenuated phenotype. When induction occurs at 1 hour post-infection and later, we expect to see no rescue of replication. These results will allow us to verify if FTN-0818 is indeed required in the first 30 minutes to 1 hour of infection. Experiments to discover this critical time point were not previously possible in *Francisella* due to lack of inducible expression elements.

Controlling gene expression through synthetic riboswitches during infection of mammalian cells is an important capability that may garner further insight into the virulence of pathogenic bacteria such as *Francisella*. However, the application of synthetic riboswitches to animal models of infection would provide a more ideal system in which to study virulence. We hypothesize that it is feasible to regulate bacterial gene expression in live mice using synthetic riboswitches. Theophylline is an FDA approved drug and has been tested in mouse models.[40, 41] By extrapolating previous data correlating dosage to blood serum concentrations,[41] we have determined that it is theoretically possible to achieve blood serum concentrations sufficient for induction of gene expression without exceeding lethal doses of theophylline (Figure 5.9). The LD_{50} (Lethal Dose causing 50% death) of theophylline in mice is 223 mg of theophylline per kg of mouse body weight.[41] This correlates to a theophylline concentration in blood serum of approximately 260 μg theophylline per mL of blood serum. A blood serum level of 260 μg/mL is equivalent to 1.5 mM theophylline. As demonstrated in Figures 5.3 and 5.8, theophylline concentrations ranging from 0.1 mM to 1 mM are sufficient to induce reporter gene expression in *F. novicida* and *F. tularensis*. Given the absence of

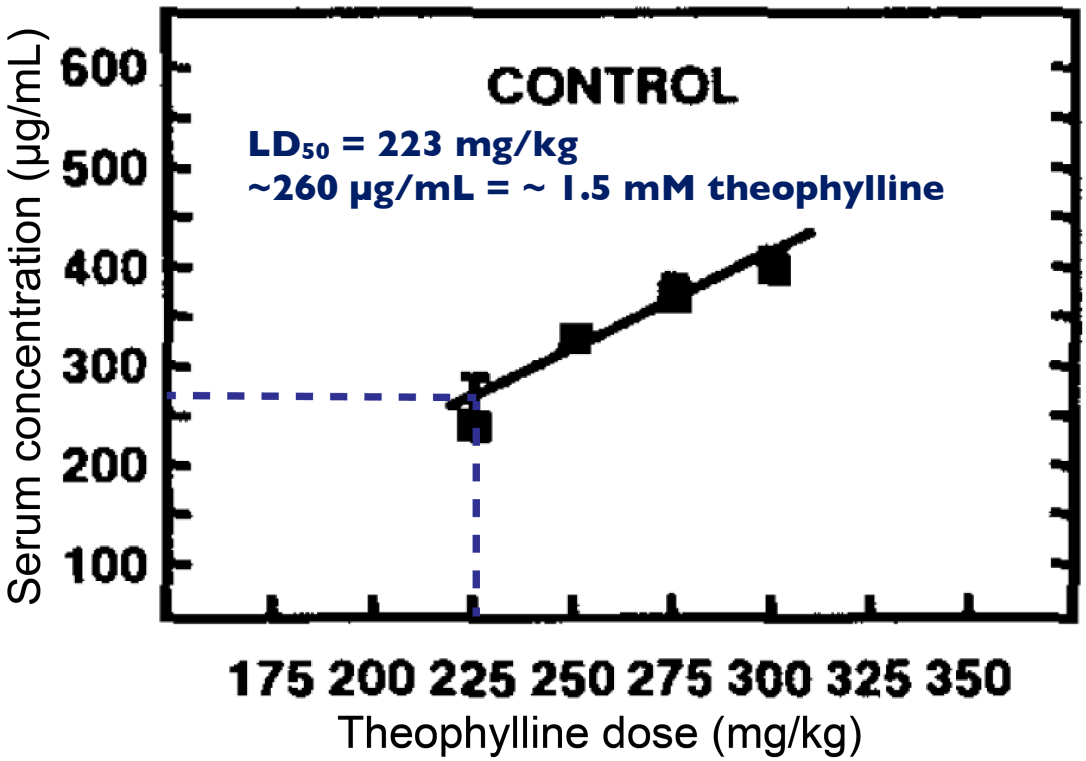

Figure 5.9 – Correlation of theophylline dose to theophylline concentration in blood serum. The x-axis indicates dose representing mg of theophylline given per kg of mouse body weight. The y-axis depicts theophylline concentration in blood serum of sacrificed mice. The blue lines indicate the LD_{50} (223 mg/kg) on the x-axis and the corresponding serum concentration (~260 µg/mL) on the y-axis. Adapted from Blake, et al. 1988.[41]

genetic tools to regulate expression of *Francisella* genes *in vivo*, developing riboswitches to be used in animal models is an avenue that merits further investigation.

5.4 Experimental

Bacterial strains and conditions.

Bacterial strains used in this study are listed in Table 1. Plasmid manipulations were performed using *E. coli* MDS42 (Scarab Genomics, Madison, WI) transformed via

[41] Figure reprinted from Blake, K. V., Massey, K. L., Hendeles, L., Nickerson, D. and Neims, A. *Ann. Emerg. Med.* **1988**, 17, 1024-1028 with permission from Elsevier.

electroporation. All experiments were performed with *F. novicida* strain U112 [12] (a gift from Dr. Denise Monack, Stanford University, Stanford, CA) or related strains and *F. tularensis* strain SchuS4. *F. novicida* overnight cultures were grown on a rolling drum at 37°C in tryptic soy broth (TSB; Difco/BD, Sparks, MD) supplemented with 0.2% L-cysteine (Sigma-Aldrich, St. Louis, MO). Growth assays were performed in Chamberlain's minimal medium as described [42]. For plating on solid media, modified Mueller Hinton (mMH) (Difco/BD) plates supplemented with 0.025% ferric pyrophosphate (Sigma-Aldrich), 0.1% glucose, and 0.1% L-cysteine were used. When appropriate, kanamycin (Kan; Fisher Scientific, Fair Lawn, NJ) was added to media at a concentration of 30 μg ml^{-1}.

Table 1. **Bacterial strains and plasmids used in this study**

Name	**Description**	**Source**
Strains		
U112	Wild type *F. novicida*	D. Monack
Δ*FTN_0818*	U112 with *FTN_0818* removed by allelic exchange	This study
SchuS4	Wild type *F. tularensis*	J. Bina
E-Rs-*FTN_0818*	Δ*FTN_0818* strain complemented with *FTN_0818* downstream of Riboswitch E	This study
F-Rs-*FTN_0818*	Δ*FTN_0818* strain complemented with *FTN_0818* downstream of Riboswitch F	This study
E-Rs-GFP	U112 strain containing plasmid pCKR34	This study
E-Rs-LacZ	U112 strain containing plasmid pCKR47	This study
F-Rs-LacZ	U112 strain containing plasmid pCKR43	This study
GFP-neg	U112 strain containing plasmid pFNLTP6-NP (negative GFP control)	This study
GFP-pos	U112 strain containing plasmid pFNLTP6-gro-gfp (positive GFP control)	This study
Plasmids		
pFNLTP6-gro-gfp	*Francisella – E. coli* shuttle vector with *gro* promoter controlling *gfp*	[26]
pFNLTP6-NP	*Francisella – E. coli* shuttle vector with promoter deleted	This study
pCKR34	pFNLTP6 derivative with Riboswitch E controlling *gfp*	This study
pCKR43	pFNLTP6 derivative with Riboswitch F controlling *lacZ*	This study
pCKR47	pFNLTP6 derivative with Riboswitch E controlling *lacZ*	This study

Cloning and mutagenesis.

Plasmids used in this study are listed in Table 1. Synthetic oligonucleotide primers were purchased from Integrated DNA Technologies (Coralville, IA). DNA polymerase and restriction enzymes were purchased from New England BioLabs (Ipswich, MA). Previously published theophylline synthetic riboswitches [5, 23, 24] were adapted to *F. novicida* through the integration of a *Francisella* promoter. Riboswitch-controlled *lacZ* gene constructs under the control of the strong promoter, P_{gro}, were cloned into an *E. coli-F. novicida* shuttle vector, pFNLTP6-gro-gfp using the PacI and BamHI restriction sites [26] (a gift from Dr. Thomas C. Zahrt, Medical College of Wisconsin, Milwaukee, WI). Riboswitches were amplified from previously reported plasmids [5, 23, 24] and assembled by PCR downstream of the *Francisella* promoter P_{gro} and upstream of *lacZ* flanked by the PacI and BamHI sites. After restriction enzyme digestion and ligation, constructs were transformed into MDS42 *E. coli.* Plasmids were then extracted using a Plasmid Midi-Prep kit (Qiagen, Germantown, MD). *F. novicida* U112 or *F. tularensis* SchuS4 were transformed as previously described [43]. Plasmids for *gfp* were prepared as described above for *lacZ*. Sequences were verified by DNA sequencing (MWG Operon, Huntsville, Al). Riboswitch-mediated complement constructs of *ΔFTN_0818* to be integrated into the *F. novicida* genome were generated by overlap PCR assembly of P_{gro}-driven riboswitches E or F (sequences in Table 1 of SI) upstream of FTN_0818 followed by a Kan-resistance cassette and flanked by sequences homologous to a region of the genome known to accept genomic additions without disrupting normal cell function (nts 818,016-818,037 and 818,649-818,670). Transformation of plasmid DNA or allelic exchange of linear DNA was performed as

previously described [43]. Briefly, a1 mL aliquot of overnight culture was added to 50 mL of TSB supplemented with 0.2% (w/v) cysteine and grown at 37°C with shaking for 2-4 hours until the $OD_{600} \approx 0.8$-1.0. Cells were then harvested by centrifugation and cell pellet was resuspended in 5 mL of room temperature Transformation Buffer [18]. Plasmid or linear DNA (~ 1 μg) was added to 200 μL aliquots of resuspended cells and incubated at 37°C on a rolling drum for 20 min. One mL of TSB + 0.2% (w/v) cysteine was added followed by another 2 hours of incubation at 37°C on a rotary wheel. Cultures were then concentrated to 200 μL by centrifugation at 5,000 x *g* and plated on selective mMH media.

Generation of FTN_0818 mutant.

To generate the clean *FTN_0818* deletion mutant, the regions of the chromosome 5' and 3' to *FTN_0818* were amplified by PCR. A kan-resistance cassette flanked by Flp-FRT recombinase sites was introduced between these flanking regions using overlapping PCR. The sewn PCR construct was gel purified (Qiagen, Valencia, CA) and chemically transformed into competent U112 as previously described (Anthony, 1991). Next, plasmid pLG72 encoding the flippase gene was transformed into the kan-marked *FTN_0818* mutant and clones in which the kan cassette had been deleted were isolated as previously described (Gallagher, 2007). To complement the deletion, constructs were made using overlapping PCR by amplification of the 5' and 3' regions, the wild-type gene, and a kan-cassette for selection and this construct was transformed into the *FTN_0818* clean deletion mutant. All strains were verified by PCR and sequencing (Eurofins EWG Operon, Huntsville, AL).

Bacterial growth curves.

To measure bacterial growth, mutant, wild-type and riboswitch strains (Table 1), were incubated overnight in TSB + 0.2% (w/v) cysteine were centrifuged at 15,294 *x g* for 2 min at room temperature. Cell pellets were resuspended in Chamberlain's minimal media [42] and then subcultured into minimal media with or without theophylline to OD_{600} = 0.03. Subcultures were aliquoted (150 μL) into a 96-well plate and incubated overnight at 37°C with shaking in a plate reader (BioTek Synergy MX, Winooski, VT). OD_{600} was taken every hour for 18 hours in triplicate wells.

β-galactosidase assay.

β-galactosidase assays were performed as previously described [44] with the following modifications: Strains were cultured overnight at 37°C with shaking. Cells were then diluted 1:30 (culture: media) (0 mM theophylline) or 1:25 (1 mM theophylline) into TSB containing kanamycin (30 μg/mL) and 0 or 1 mM theophylline to an $OD_{600} \approx$ 0.7 - 0.8. Theophylline and *o*-nitrophenyl-β-D-galactopyranoside (ONPG) were purchased from Sigma-Aldrich. Dose response assays were performed as described above with the addition of various concentrations of theophylline and dilutions of 1:30 (0 and 0.25 mM theophylline), 1:27 (0.5 mM theophylline) and 1:21 (2 mM theophylline). These volumes were used to compensate for small decreases in growth rate with increasing concentrations of theophylline. All assays performed in *F. tularensis* SchuS4 were performed by Dr. James Bine (University of Pittsburgh).

Macrophage infections.

For infection, RAW264.7 murine macrophages (ATCC, Manassas, VA) were seeded on chambered coverglass (CultureWell, Invitrogen, Carlsbad, CA) with 1 x 10^5 cells per well and incubated overnight in Dulbecco's modified Eagle medium (high glucose, L-glutamine; DMEM; Lonza, Walkersville, MD) supplemented with 10% heat-inactivated fetal calf serum (FCS; HyClone, Logan, UT) at 37°C with 5% CO_2. The media was then removed, and the macrophages were infected with overnight cultures of various GFP strains (Table 1) that were diluted in DMEM/10% FCS to achieve a multiplicity of infection (MOI) of 50:1 colony forming units (CFUs) of bacteria per macrophage. The plates were centrifuged for 15 min at 233 x *g* at room temperature and then incubated for 30 min. The macrophages were then washed twice with DMEM and returned to incubation overnight (18-24 hours). Next, macrophages were rinsed twice with sterile PBS and fixed by incubating in 4% paraformaldehyde (PFA) in PBS for 20 min at room temperature. The cells were again rinsed twice with PBS to wash away PFA. To measure bacterial replication, RAW264.7 macrophages were infected as described above with the following differences: macrophages were seeded on 24-well tissue culture plates at 7.5×10^5 cells/well and infected with an MOI of 20:1 CFUs per macrophage in DMEM/10% FCS containing 100 µg/ml of gentamicin (TekNova, Hollister, CA) with or without 1 mM theophylline. Macrophages were then lysed with 1% saponin (Alfa Aesar, Heysham, Lancs., UK) in phosphate buffered saline (PBS) without calcium and magnesium (Lonza, Walkersville, MD) at the 30 minutes or 24 hours. To quantify the bacterial load within macrophages at each time point, serial dilutions of the macrophage lysates were plated onto modified mMH agar plates and

incubated overnight at 37°C. Finally, Finally fold replication of each strain was calculated (CFU at 24 hr/CFU at 30 minutes).

Microscopy.

Fluorescence microscopy was performed with an Eclipse *Ti* microscope coupled with the Nikon Elements software package (Nikon). Image capture was performed with an Evolve EM CCD (Photometrics) using a CFI Apo 100x (NA = 1.49) objective (Nikon) with an Intensilight epifluorescence source (Nikon). Data acquisition utilized the Nikon Perfect Focus System that allows the capture of Multipoint Z-stack data without loss of focus. Z-stack images in 10 fields of view were taken for each construct using the FITC and DAPI Chroma filter cubes to detect GFP and macrophage nuclei respectively. Z-stacks probed a 5 μm section of the macrophages at 0.4 μm per slice. FITC exposure time was 300 ms while DAPI exposure time was 50 ms. Images of the negative control were used as a reference to eliminate background fluorescence. Settings from the negative control were propagated to all images to allow accurate comparison. False color was applied to aid in analysis.

Statistics.

Macrophage replication data were analyzed for significance using the unpaired student's t-test with 2 degrees of freedom.

5.5 References

(1) Sørensen, H. P. and Mortensen, K. K. *J. Biotechnol.* **2005**, 115, 113-128.

(2) François, B. *Curr. Opin. Biotechnol.* **1999**, 10, 411-421.

(3) Scaife, J., Leach, D. R. F. and Galizzi, A. *Genetics of Bacteria*. **1985**,

(4) Bugrysheva, J. V., Froehlich, B. J., Freiberg, J. A. and Scott, J. R. *Appl. Environ. Microbiol.* **2011**, 77, 4422-4428.

(5) Topp, S., Reynoso, C. M. K., Seeliger, J. C., Goldlust, I. S., Desai, S. K., Murat, D., Shen, A., Puri, A. W., Komeili, A., Bertozzi, C. R., Scott, J. R. and Gallivan, J. P. *Appl. Environ. Microbiol.* **2010**, 76, 7881-7884.

(6) Wickiser, J. K., Winkler, W. C., Breaker, R. R. and Crothers, D. M. *Mol. Cell*. **2005**, 18, 49-60.

(7) Oyston, P. C. F. *J. Med. Microbiol.* **2008**, 57, 921-930.

(8) Saslaw, S., Eigelsbach, H. T., Prior, J. A., Wilson, H. E. and Carhart, S. *Arch. Intern. Med.* **1961**, 107, 702-714.

(9) Lauriano, C. M., Barker, J. R., Yoon, S.-S., Nano, F. E., Arulanandam, B. P., Hassett, D. J. and Klose, K. E. *Proc. Natl. Acad. Sci.* **2004**, 101, 4246-4249.

(10) Larsson, P., Elfsmark, D., Svensson, K., Wikström, P., Forsman, M., Brettin, T., Keim, P. and Johansson, A. *PLoS Pathog.* **2009**, 5, e1000472.

(11) Forsman, M., G. Sandstrom, and A. Sjostedt. *Int. J. Syst. Bacteriol.* **1994**, 44, 38-46.

(12) Baron, G. S. and Nano, F. E. *Mol. Microbiol.* **1998**, 29, 247-259.

(13) Ebright, R. H. and Busby, S. *Curr. Opin. Genet. Dev.* **1995**, 5, 197-203.

(14) Charity, J. C., Costante-Hamm, M. M., Balon, E. L., Boyd, D. H., Rubin, E. J. and Dove, S. L. *PLoS Pathog.* **2007**, 3, e84.

(15) Mukhamedyarov, D., Makarova, K., Severinov, K. and Kuznedelov, K. *BMC Mol. Biol.* **2011**, 12, 50.

(16) Gourse, R. L., Ross, W. and Gaal, T. *Mol. Microbiol.* **2000**, 37, 687-695.

(17) Horzempa, J., Tarwacki, D. M., Carlson, P. E., Robinson, C. M. and Nau, G. J. *Appl. Environ. Microbiol.* **2008**,

(18) Frank, D. W. and Zahrt, T. C. *Ann. N. Y. Acad. Sci.* **2007**, 1105, 67-97.

(19) Grall, N., Livny, J., Waldor, M., Barel, M., Charbit, A. and Meibom, K. L. *Microbiology*. **2009**, 155, 2560-2572.

(20) Llewellyn, A. C., Jones, C. L., Napier, B. A., Bina, J. E. and Weiss, D. S. *PLoS ONE*. **2011**, 6, e24201.

(21) Su, J., Yang, J., Zhao, D., Kawula, T. H., Banas, J. A. and Zhang, J. R. *Infect. Immun.* **2007**, 75, 3089-3101.

(22) Weiss, D. S., Brotcke, A., Henry, T., Margolis, J. J., Chan, K. and Monack, D. M. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, 104, 6037-6042.

(23) Lynch, S. A., Desai, S. K., Sajja, H. K. and Gallivan, J. P. *Chem. Biol.* **2007**, 14, 173-184.

(24) Topp, S. and Gallivan, J. P. *ChemBioChem*. **2008**, 9, 210-213.

(25) M Ericsson, I. G., G Sandström, A Tärnvik and A Sjöstedt. *Infect. Immun.* **1997**, 65, 1824-1829.

(26) Maier, T. M., Havig, A., Casey, M., Nano, F. E., Frank, D. W. and Zahrt, T. C. *Appl. Environ. Microbiol.* **2004**, 70, 7511-7519.

(27) Desai, S. K. and Gallivan, J. P. *J. Am. Chem. Soc.* **2004**, 126, 13247-13254.

(28) Topp, S. and Gallivan, J. P. *J. Am. Chem. Soc.* **2007**, 129, 6807-6811.

(29) Barrow, E. W., Westbrook, L., Bansal, N., Suling, W. J., Maddry, J. A., Parker, W. B. and Barrow, W. W. *The Journal of antimicrobial chemotherapy*. **2003**, 52, 801-808.

(30) Behar, S. M., Divangahi, M. and Remold, H. G. *Nat Rev Micro*. **2010**, 8, 668-674.

(31) Buchmeier, N., Blanc-Potard, A., Ehrt, S., Piddington, D., Riley, L. and Groisman, E. A. *Mol. Microbiol.* **2000**, 35, 1375-1382.

(32) Teitelbaum, R., Cammer, M., Maitland, M. L., Freitag, N. E., Condeelis, J. and Bloom, B. R. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, 96, 15190-15195.

(33) Rinaldo, J. E., Clark, M., Parinello, J. and Shepherd, V. L. *Am. J. Respir. Cell Mol. Biol.* **1994**, 11, 625-630.

(34) Stredansky, M., Pizzariello, A., Miertus, S. and Svorc, J. *Anal. Biochem.* **2000**, 285, 225-229.

(35) Flannagan, R. S., Cosio, G. and Grinstein, S. *Nat. Rev. Micro.* **2009**, 7, 355-366.

(36) Tapper, H. and Sundler, R. *The Biochemical journal*. **1992**, 281 (Pt 1), 239-244.

(37) Zaide, G., Grosfeld, H., Ehrlich, S., Zvi, A., Cohen, O. and Shafferman, A. *Appl. Environ. Microbiol.* **2011**, 77, 1608-1618.

(38) Deutscher, J., Herro, R., Bourand, A., Mijakovic, I. and Poncet, S. *Biochim. Biophys. Acta*. **2005**, 1754, 118-125.

(39) Seeliger, J. C., Topp, S., Sogi, K. M., Previti, M. L., Gallivan, J. P. and Bertozzi, C. R. *PLoS ONE*. **2012**, 7, e29266.

(40) McColl, J. D., Parker, J. M. and Ferguson, J. K. W. *J. Pharmacol. Exp. Ther.* **1956**, 116, 343-350.

(41) Blake, K. V., Massey, K. L., Hendeles, L., Nickerson, D. and Neims, A. *Ann. Emerg. Med.* **1988**, 17, 1024-1028.

(42) Chamberlain, R. E. *Appl. Microbiol.* **1965**, 13, 232-235.

(43) Anthony, L. S. D., Gu, M., Cowley, S. n. C., Leung, W. W. S. and Nano, F. E. *J. Gen. Microbiol.* **1991**, 137, 2697-2703.

(44) Miller, J. H. *Experiments in Molecular Genetics*. **1972**, 352-355.

Chapter 6 – Conclusions

6.1 Summary and Conclusions

Genetic tools for the conditional expression of genes are useful in determining gene function.[1-3] Many species of bacteria lack such tools. While several genetic tools are available for use in model organisms such as *E. coli* and *B. subtilis*,[4, 5] protein-based tools are not always transportable to other species of bacteria. Synthetic riboswitches offer an alternative to protein-based gene regulatory systems and have proven to be transportable from one bacterial species to another.

We have demonstrated the ability to transport synthetic riboswitches isolated from *E. coli* and directly transport them to other Gram-negative species (Chapter 2). We have also demosntrated the ability to isolate functioning synthetic riboswtiches *de novo* from a library expressed in a genetically tractable species, *A. baylyi* (Chapter 3). Analysis of the RBS sequences of riboswitches isolated from *E. coli* versus those isolated from *A. baylyi* indicate the importance of a strong RBS in bacterial species other than *E. coli*. Using this knowledge and the knowlegde that Gram-positive bacteria have more stringent requirements with regards to RBS complementarity, we were able to demonstrate semi-rational design of a synthetic riboswitch that functions in the Gram-positive species *S. pyogenes* (Chapter 4). The use of semi-rational design circumvents the need for *de novo* development of riboswitches, which would be difficult in less genetically tractable species such as *S. pyogenes*.

Comparison of the functionality of our riboswitches in *E. coli* and *S. pyogenes* also led us to the conclusion that the strongest RBS does not necessarily correspond to an ideal riboswitch. While increasing the strength of the RBS will indeed increase overall

gene expression, in Gram-negative bacteria particularly, the strongest RBS also corresponds to high basal gene expression in the absence of theophylline. Therefore, an ideal synthetic riboswitch will possess an RBS that is sufficient to drive efficient protein translation in the presence of ligand while not exhibiting excessively high basal expression in the asence of ligand.

We have also found that synthetic riboswitches with different basal expression levels, activation ratios and dynamic ranges may be ncessary depending on the desired application. As in the case of controlling *FTN_0818* in *F. novicida* cells growing in liquid media, only riboswitch F (exhibiting the lowest basal expression) was able to sufficiently repress FTN_0818 expression in the absence of theophylline (Chapter 5). However, in macrophage experiments, riboswitch E (demonstrating the highest dynamic range) was necessary to elicit rescue of mutant bacteria in the presence of theophylline during macrophage infections (Chapter 5).

To our knowlegde, our theophylline-sensitive synthetic riboswitches are the only inducible gene expression systems available for use in the pathogenic bacteria *F. novicida* and *F. tularensis*. In the Gram-positive pathogenic bacterium *S. pyogenes*, our synthetic riboswitches exihibit the lowest basal expression and highest actviation ratio of any tool available to date. However, the main disadvantage of our synthetic riboswitches in comparison to other inducible systems, such as the tetracyline-inducible system, [1, 2] is that riboswtiches have not been demonstrated to funciton during live animal infection.

We have performed preliminary work to determine whether the use of our theophylline-sensitive synthetic riboswitches is feasible in live mice and have found that

a blood serum theophylline concentration of 1 mM is sublethal. We have demonstrated that in liquid media, half maximal expression of riboswitch E (exhibiting the highest dynamic range in most baterial species) can be achieved with theophylline concentrations of approximately 0.5 mM. We have also found that pre-dosing of mice may be required to obtain blood serum theophylline concentrations sufficient to activate the riboswitch. The need for pre-dosing may preclude the use of our synthetic riboswtiches for certain applications. However, the fact that theophylline is an inexpensive, non-endogenous molecule, and the fact that riboswitches appear to be readliy transportable from one bacterial species to another would merit further investigation of theophylline-sensitive riboswtiches for use in live animal infections. Future work would concentrate on determining ideal methods of dosing, developing infection protocols and optimizing methods to measure successful induction of the target gene.

6.2 References

(1) Bateman, B. T., Donegan, N. P., Jarry, T. M., Palma, M. and Cheung, A. L. *Infect. Immun.* 2001, 69, 7851-7857.

(2) Blokpoel, M. C., Murphy, H. N., O'Toole, R., Wiles, S., Runn, E. S., Stewart, G. R., Young, D. B. and Robertson, B. D. *Nucleic Acids Res.* 2005, 33, e22.

(3) Davison, J. *BioTechniques*. 2002, 32, 386-388, 390, 392-384, passim.

(4) Yansura, D. G. and Henner, D. J. *Proc. Natl. Acad. Sci.* 1984, 81, 439-443.

(5) Sá-Nogueira, I., Nogueira, T. V., Soares, S. and de Lencastre, H. *Microbiology*. 1997, 143, 957-969.