Genetic mechanisms contributing to antimicrobial resistance: gene repression in *Mycobacterium*, gene transfer in *Salmonella*, and phage sensitivity in *Escherichia*

A Thesis
Submitted to the Faculty of Graduate Studies and Research
In Partial Fulfilment of the Requirements
For the Degree of

Master of Science

In

Biology

University of Regina

By

Joshua Raj Kotaro Yoneda

Regina, Saskatchewan

July, 2018

Copyright ©2018: J. R. K. Yoneda
Joshua Raj Kotaro Yoneda, candidate for the degree of Master of Science in Biology, has presented a thesis titled, *Genetic Mechanisms Contributing to Antimicrobial Resistance: Gene Repression in Mycobacterium, Gene Transfer in Salmonella, and Phage Sensitivity in Escherichia*, in an oral examination held on July 12, 2018. The following committee members have found the thesis acceptable in form and content, and that the candidate demonstrated satisfactory knowledge of the subject material.

External Examiner: Dr. Tanya Dahms, Department of Chemistry & Biochemistry

Supervisor: Dr. Andrew Cameron, Department of Biology

Committee Member: *Dr. Tzu-Chiao Chao, Department of Biology

Chair of Defense: Dr. Kyle Hodder, Department of Geography & Environmental Studies

*Not present at defense*
Abstract

Bacteria develop resistance to antimicrobial agents through genetic mutation or horizontal acquisition of genetic material from resistant cells. *Mycobacterium*, *Salmonella*, and *Escherichia* are all capable of causing disease, and each demonstrates increasing resistance to conventional antibiotic treatment. Understanding the genetic mechanisms that regulate antibiotic resistance are critical in developing new antimicrobials and improving current treatment regimens. Alternative therapies to antibiotics should also be considered in the fight against multidrug resistant pathogens. Thus, I carried out molecular studies to investigate the regulation of efflux pump expression in *Mycobacterium intracellulare*, regulation mechanisms of natural competence in *Salmonella* Typhimurium, and phage therapy of Enterohemorrhagic *Escherichia coli* serogroup O157:H7.

A recent study showed specific mutations in the uncharacterized locus *mmpT5* increased resistance to anti-TB drug bedaquiline in *M. intracellulare* lung infections. Based on previous work, MmpT5 is a TetR transcription factor hypothesized to repress expression of the downstream Resistance Nodulation Division (RND) efflux pump *mmpSL*. To test this, we assembled multigene constructs on a plasmid with the modular cloning (MoClo) toolkit, which allowed us to control the expression of *mmpT5* with the *lacZ* promoter and monitor *mmpSL* promoter activity with a fusion to *lux*. Consistent with the hypothesis, induction of *mmpT5* reduced expression of *mmpSL*. The nonsynonymous mutations in *mmpT5* first identified in clinical isolates resulted in upregulation of *mmpSL*, suggesting a mechanism for bedaquiline resistance.

The periplasmic protein ComA is essential for active uptake of DNA from the
environment. In *Salmonella* Typhimurium, *comA* is transcriptionally silent, even when the predicted activator Sxy is overexpressed. Identification of a mRNA stem-loop in the *comA* promoter region provides a possible gene repression mechanism. We used the MoClo toolkit to assemble *comA* promoter-luciferase fusions with or without the predicted mRNA secondary structure, and we used site-directed mutagenesis (SDM) to destabilize the predicted mRNA stem-loop. Luciferase reporter assays showed stem-loop removal or destabilization via SDM increased *comA* expression under the control of the *lacZ* promoter. SDM destabilization of the mRNA stem-loop did not increase luminescence of the native *comA*, but induction of *sxy* resulted in a 197-fold increase in luminescence compared to the wild type when the predicted stem-loop was destabilized, presenting a clear explanation for the cryptic nature of *comA* expression in *Salmonella*.

Phage therapy presents an attractive alternative to conventional antibiotic treatment, prompting a study of the sensitivity of Enterohemorrhagic *Escherichia coli* (EHEC) serogroup O157:H7 to phage killing using a tetrazolium reduction assay. Measuring tetrazolium reduction at A$_{485}$ successfully showed which EHEC strains were sensitive to the T4-like typing phage 13 (TP13). Addition of T7-like phage TP9 to TP13 increased the efficacy of phage treatment. TP13 resistance was observed in EHEC strains with the stx2a lysogen. RT-qPCR and plasmid complementation experiments showed that the hypothetical gene *H2* was not responsible for TP13 resistance. Subsequent review of recent literature revealed that the downstream tyrosine kinase *stk* is likely responsible for TP13 resistance through abortive infection. Altogether, this thesis studies three bacterial systems to better understand the genetic bases of antimicrobial resistance, including gene transfer mechanisms, and to find the genetic bases of phage sensitivity.
Acknowledgments

First and foremost, I would like to express my deep gratitude to my outstanding supervisor, Dr. Andrew Cameron, for his continuous support of my Master’s project, and my career goal of becoming a physician. His extraordinary patience, thoughtful guidance, insightful comments, and kind demeanor highlighted my time here in Regina. I sincerely appreciate his confidence in me to complete a diverse array of projects. I will never forget the exceptional opportunities he gave me, such as traveling to Scotland on the Queen Elizabeth II Diamond Jubilee Scholarship to do research at the University of Edinburgh. At a time in my life when I was unsure what career steps to take next, Andrew gave me a pathway to become an engaged and enthusiastic participant of the scientific community. I feel very fortunate to have Dr. Andrew Cameron as my mentor and I hope to keep his teachings and values as I begin my medical studies.

Thank you to Dr. Tzu-Chiao for his insightful recommendations for formatting my thesis. I extend my thanks to Dr. David Alexander and Dr. Jeffrey Chen for their support and helpful comments on my Mycobacterium research.

I would also like to thank my excellent lab colleagues, Dr. Ebtihal Alshabib, Skip Olshefsky, Stefani Kary, Benjamin Perry, Danae Suchan, Nicole Lerminiaux, Ruhul Amin, Dr. Jordyn Bergsveinson and Dr. Keith MacKenzie for their constant support, encouragement, and helpful advice. I feel very grateful for being a part of such a supportive, fun, and intelligent team.

Special thanks to my friends and colleagues in Edinburgh, Dr. Stephen Fitzgerald, Dr. Alison Tidswell, Dr. Sean McAteer, Dr. Nadejda Lupolova, and Dr. David Gally for making my time abroad educational, memorable, and immensely enjoyable.
Abstract ................................................................................................................................. I
Acknowledgments .................................................................................................................. III
Table of Contents ................................................................................................................... IV
List of Tables .......................................................................................................................... VII
List of Figures ........................................................................................................................ VIII
List of Abbreviations ............................................................................................................. X
FOREWORD .......................................................................................................................... XII

Chapter 1: General Introduction ......................................................................................... 1

1.1 INTRODUCTION .............................................................................................................. 1

1.2 Three bacterial systems addressed in this thesis ............................................................. 2
  1.2.1 *Mycobacterium* pathogenicity and antibiotic resistance ........................................... 2
  1.2.1.1 *Mycobacterium* efflux pumps ............................................................................. 3
  1.2.2 Horizontal transfer of antibiotic resistance genes ...................................................... 6
    1.2.2.1 Regulation of genes for natural competence in *Salmonella* ................................ 7
  1.2.3 Golden Gate cloning .................................................................................................. 11
  1.2.4 Alternative antimicrobial agents ............................................................................. 13
    1.2.4.1 EHEC phage therapy ....................................................................................... 14

1.3 Thesis Objective ............................................................................................................. 15

REFERENCES ....................................................................................................................... 16

Chapter 2: Transcriptional Regulator *mmpT5* Represses Efflux Pump Expression in *Mycobacterium intracellular* and Provides a Mechanism for Bedaquiline Resistance. ........................................................................................................................................ 23

ABSTRACT ............................................................................................................................ 24

2.1 INTRODUCTION .............................................................................................................. 25

2.2 MATERIALS AND METHODS ....................................................................................... 27
  2.2.1 Bacterial Strains and Plasmids ................................................................................ 27
  2.2.2 Bioinformatic analysis ............................................................................................ 27
  2.2.3 PCR amplification of *mmpT5* and *mmpSL* ......................................................... 27
  2.2.4 Golden Gate Assembly ......................................................................................... 29
  2.2.5 Construction of plasmids for expression in *M. smegmatis* ................................... 32
  2.2.6 Luciferase-Reporter Assay ................................................................................... 32

2.3 RESULTS ......................................................................................................................... 33
  2.3.1 Sanger sequencing confirms nonsynonymous mutations in *mmpT5* cloned from relapse isolates .......................................................................................................................... 33
  2.3.2 The *mmpSL* promoter is expressed in *E. coli*, but the *mmpT5* promoter is not... 36
Chapter 3: Regulation of \( \text{comA} \) gene expression by Sxy and an mRNA secondary structure in \textit{Salmonella Typhimurium} ........................................54

ABSTRACT ..................................................................................55

3.1 INTRODUCTION .......................................................................56

3.2 MATERIALS AND METHODS ......................................................59

3.2.1 Bacterial Strains and Plasmids .................................................59
3.2.2 PCR amplification of \( \text{sxy} \) and \( \text{comA} \) promoters .......................61
3.2.3 Multigene construct design and assembly .................................61
3.2.4 Quantifying competence gene promoter activity ........................65
3.2.5 Prediction of \( \text{comA} \) mRNA secondary structure .......................65
3.2.6 \( \text{comA} \) site-directed mutagenesis .............................................66

3.3 RESULTS ..................................................................................66

3.3.1 Impairing mRNA secondary structure increases \( \text{comA} \) expression ....66
3.3.3 Sxy-dependent \( \text{comA} \) expression is increased through hypercompetence mutations ........................................................................71

3.4 DISCUSSION ...........................................................................73

3.5 Acknowledgements ....................................................................76

REFERENCES ..............................................................................76

Chapter 4: EHEC phage therapy and characterization of tyrosine-kinase phage exclusion mechanism \(^1\) ..........................................................80

ABSTRACT ..................................................................................81

4.1 INTRODUCTION .......................................................................82

4.2 MATERIALS AND METHODS ......................................................84

4.2.1 Bacterial Strains, Phages, and Plasmids .....................................84
4.2.2 Measurement of EHEC phage sensitivity with tetrazolium assay ....87
4.2.3 Reverse transcription quantitative PCR (RT-qPCR) .........................87
4.2.4 Plasmid complementation of H2 mutant .......................................89

4.3 RESULTS ..................................................................................89

4.3.1 The \( \text{stx2a} \) phage region confers resistance to TP13 .......................90
4.3.2 Addition of TP9 to TP13 prevents relapse growth of \( \Delta \text{stx2a} \) strains 92
4.3.3 Expression profile of \( \text{stx2a} \) phage region .......................................92
4.3.4 Plasmid complementation of the 1599ΔH2 mutant does not restore TP13 resistance........................................................................................................94

4.4 DISCUSSION ................................................................................................................94

4.5 Acknowledgements ....................................................................................................99

REFERENCES .................................................................................................................99
List of Tables

Table 2.1. Bacterial plasmids and strains used in this study.............................................28
Table 2.2. BDQ MIC and mmpT5 mutations of clinical pretreatment and relapse isolates........................................................................................................................................34
Table 2.3. Unsuccessful mmpT5 and mmpSL constructs made to test the proposed hypothesis........................................................................................................................................44
Table 3.1. Bacterial plasmids and strains used in this study.............................................60
Table 3.2. Final level 1 comA constructs, level 2 inducible sxy constructs, and site-directed mutagenesis (SDM) targets..............................................................................................................62
Table 4.1. Bacterial strains, phages, and plasmids used in this study...............................86
List of Figures

Figure 1.1. The efflux pump families found in bacteria.................................................4
Figure 1.2. Uptake mechanism of DNA using Type IV Pilus........................................9
Figure 1.3. Alignment of competence genes from four members of proteobacteria........9
Figure 1.4. Hierarchical assembly of multigene constructs using MoClo toolkit.............12
Figure 2.1. Location and characterization of DNA mutations in mmpT5........................34
Figure 2.2. Alignment of mmpT5 DNA and protein sequences from pretreatment and relapse isolates.................................................................35
Figure 2.3. Bioluminescence assay of mycobacterial promoters..................................37
Figure 2.4. Assembly of experimental constructs using MoClo Toolkit..........................39
Figure 2.5. Bioluminescence assay of 1 plasmid experimental setup.............................41
Figure 2.6. Bioluminescence assay of 2 plasmid experimental setup.............................42
Figure 2.6. Proposed mechanism of BDQ resistance......................................................47
Figure 3.1. Predicted comA mRNA secondary structure and location of site-directed mutations........................................................................................................58
Figure 3.2. Final synthetic multigene constructs used to study sxy regulation and the predicted comA secondary structure.................................................................67
Figure 3.3. Bioluminescence assay of level 1 comA constructs......................................68
Figure 3.4. Bioluminescence assay of level 1 PcomA-TSS site-directed mutagenesis constructs..............................................................................................................70
Figure 3.5. Bioluminescence assay of level 2 comA SDM/inducible Sxy constructs........72
Figure 4.1. The strain 9000 family tree and stx2a phage region....................................85
Figure 4.2. Reduction of tetrazolium dye to measure cellular respiration.......................88
Figure 4.3. Measurement of EHEC growth treated with different dilutions of TP13........91
Figure 4.4. Measurement of EHEC growth treated with TP9 and TP13.........................93
Figure 4.5. RT-qPCR relative expression of $H_1$, $H_2$, and $stk$........................................95
Figure 4.6. Measurement of EHEC growth treated with TP13........................................96
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDQ</td>
<td>Bedaquiline</td>
</tr>
<tr>
<td>CRP</td>
<td>cAMP receptor protein</td>
</tr>
<tr>
<td>dsDNA</td>
<td>Double strand DNA</td>
</tr>
<tr>
<td>EHEC</td>
<td>Enterohemorrhagic <em>Escherichia coli</em></td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-ß-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>MAC</td>
<td><em>Mycobacterium avium</em> complex</td>
</tr>
<tr>
<td>MDR</td>
<td>Multi-drug resistant</td>
</tr>
<tr>
<td>MmpL</td>
<td>Mycobacterial membrane protein Large</td>
</tr>
<tr>
<td>MmpS</td>
<td>Mycobacterial membrane protein Small</td>
</tr>
<tr>
<td>MOI</td>
<td>Multiplicity of infection</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>NTM</td>
<td>Non-tuberculous mycobacteria</td>
</tr>
<tr>
<td>OD&lt;sub&gt;600&lt;/sub&gt;</td>
<td>Optical density at wavelength of 600 nm</td>
</tr>
<tr>
<td>ORF</td>
<td>Open Reading Frame</td>
</tr>
<tr>
<td>ORI</td>
<td>Origin of replication</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>RBS</td>
<td>Ribosome binding site</td>
</tr>
<tr>
<td>RNAP</td>
<td>RNA polymerase</td>
</tr>
<tr>
<td>RND</td>
<td>Resistance Nodulation Division</td>
</tr>
<tr>
<td>RT-qPCR</td>
<td>Reverse transcription quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>SDM</td>
<td>Site-directed mutagenesis</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>Stx</td>
<td>Shiga toxin</td>
</tr>
<tr>
<td>TB</td>
<td>Tuberculosis</td>
</tr>
<tr>
<td>TetR</td>
<td>Tetracycline repressor</td>
</tr>
<tr>
<td>TFs</td>
<td>Transcription factors</td>
</tr>
<tr>
<td>T4P</td>
<td>Type IV pilus</td>
</tr>
<tr>
<td>TP</td>
<td>Typing phage</td>
</tr>
<tr>
<td>TSS</td>
<td>Transcription Start Site</td>
</tr>
<tr>
<td>WT</td>
<td>Wild type</td>
</tr>
<tr>
<td>XDR</td>
<td>Extensively-drug resistant</td>
</tr>
</tbody>
</table>
FOREWORD

This is a manuscript-based thesis divided into four chapters based on three different bacterial systems. Chapter 1 is the general introduction. Chapter 2 describes characterization of the genetic basis of bedaquiline resistance in *Mycobacterium intracellulare*. Chapter 3 describes characterization of gene repression by RNA secondary structure in *Salmonella enterica*. Chapter 4 summarizes the EHEC phage therapy work conducted during a Queen Elizabeth II Diamond Jubilee Scholarship held at the University of Edinburgh (Sept-Dec 2017). All chapters are written by Joshua Yoneda.

Figures 1.2 and 1.3 in the Introduction and Figure 3.1 in Chapter 3 where made by Ebtihal Y Alshabib. In Chapter 3, Danae Suchan conducted the second round of site-directed mutagenesis titled “SDM 2” on the level 1 *comA* constructs.

During my Master’s program, I have contributed to two additional manuscripts not appearing in this thesis:


Alshabib, E.Y. Yoneda, J.R.K., Lerminiaux, N.A., Cameron, A.D.S. (xxxx) The chromosomal preplin peptidase gene *hopD* is activated by CRP and Sxy in *Salmonella enterica*. In preparation for submission to *Microbiology*
Chapter 1: General Introduction

1.1 INTRODUCTION

Antimicrobial resistance is one of humanity’s greatest challenges, predicted to kill more people than cancer by the year 2050 (52). Conventional antibiotics are becoming increasingly ineffective with the rise of multidrug resistant (MDR) pathogens, and researchers are struggling to come up with novel drug compounds. Bacteria have evolved multiple ways to survive, including the activation or repression of gene expression that allows bacteria to fine-tune their physiology for specific environments and respond quickly to stresses such as antibiotics (47). Horizontal gene transfer provides a mechanism for acquiring antibiotic resistant genes from the environment (7). Additionally, natural physiological mechanisms such as efflux pumps prevent the toxic accumulation of clinical drugs (55). It is imperative to understand the genetic bases of these mechanisms because they pinpoint how antibiotic resistance develops and provide potential targets for antimicrobial therapy. We must also explore alternative treatment methods such as phage therapy (3). Increasing the efficacy of phage cocktails against bacterial disease would be a powerful addition to the antimicrobial arsenal.

This thesis addresses antimicrobial resistance by dissecting two gene repression mechanisms in Mycobacterium intracellular and Salmonella enterica serovar Typhimurium, and it investigates a novel method for testing phages against Enterohemorrhagic Escherichia coli (EHEC) serogroup O157:H7. The modular cloning (MoClo) system was used to construct plasmids and assess efflux pump regulation in M. intracellular and regulation of natural competence in S. Typhimurium. The last chapter describes my development and use of a tetrazolium-based assay to assess phage
sensitivity of EHEC strains, which revealed a tyrosine kinase-mediated phage resistance mechanism.

1.2 Three bacterial systems addressed in this thesis

The first system introduced is *Mycobacterium* pathogenicity and efflux pumps (1.2.1 & 1.2.1.1), then horizontal gene transfer and natural competence in *Salmonella* is examined (1.2.2 & 1.2.2.1). This is followed by introducing the modular cloning system (1.2.3) used to study *Mycobacterium* and *Salmonella* gene regulation in this thesis. Finally, alternatives to antibiotics are addressed (1.2.4) and phage therapy of enterohemorrhagic *E. coli* is explored (1.2.4.1).

1.2.1 *Mycobacterium* pathogenicity and antibiotic resistance

The genus *Mycobacterium* contain amongst the most infectious pathogens found in humans with *Mycobacterium tuberculosis* and *Mycobacterium leprae* collectively infecting millions of people each year (75, 76). According to the World Health Organization (WHO), Tuberculosis (TB) is the ninth leading cause of death worldwide with 6.3 million new cases of TB in 2016 (76). Leprosy continues to target Asian and African countries with 216,000 new cases in 2016 (75). In addition, pathogenic *Mycobacterium* are notoriously adept at developing resistance to antibiotics (41). In 2016, there were 490,000 cases of multidrug-resistant TB (MDR-TB), with almost half of these occurring in India, China, and Russia (75). Extensively-drug resistant TB (XDR-TB) is defined by a strain of *M. tuberculosis* that is resistant to isoniazid, rifampin, fluoroquinolones, and at least one of three injectable second-line drugs (i.e., amikacin, kanamycin, or capreomycin) (19). XDR-TB cases have been found in almost every country, with about 6.2% of all MDR-TB cases being classified as XDR-TB (75).
Treating MDR-TB and XDR-TB requires 9 to over 18-month treatment regimens, increasing the chances that patients will fail to adhere to treatment. In addition to improving early detection and developing efficient treatment regimens, researchers are studying the molecular mechanisms of *Mycobacterium* antibiotic resistance. By identifying resistance mechanisms, we can predict the potential for similar resistance mechanisms to evolve in clinical settings, and potentially identify new targets for treatment.

**1.2.1.1 *Mycobacterium* efflux pumps**

Antibiotic resistance in TB has been largely attributed to the expression of a diverse range of multidrug efflux pumps in *M. tuberculosis* (6). Efflux pumps are proteinaceous membrane transporters common to both pathogenic and non-pathogenic bacteria and possesses a multitude of functions, including the export of antibiotics (61). There are five main categories of efflux pumps in bacteria: ATP-binding cassette (ABC), major facilitator super-family (MFS), resistance nodulation division (RND), small multidrug resistance (SMR) and multidrug and toxic-compound extrusion (MATE) (Figure 1.1) (55). All five families of bacterial efflux pumps are associated with multidrug-resistance (MDR) and are classified based on their structural components and energy source. *M. tuberculosis* has one of the highest numbers of efflux pumps relative to its genome size, containing a diverse range of multidrug efflux pumps from the MFS, ABC, SMR, and RND families (53).

The RND family subclass of mycobacterial membrane protein Large (MmpL) transporters have become a recent group of interest because of their active role in antibiotic resistance. The *M. tuberculosis* genome encodes 14 MmpL proteins, which
Figure 1.1. The efflux pump families found in bacteria.

This figure is from (55). Five bacterial efflux pump superfamilies contributing to multidrug resistance (MDR). The five families all use a form of active transport but differ in their protein components and energy source used for transportation. The ABC transporter family is the only group to utilize ATP for transport, while the other four families use a sodium or proton motor force. The mycobacterial membrane protein Large (MmpL) transporters are part of the RND family and use a proton antiport mechanism to export substrates against their concentration gradient.
display a multitude of functions including the transport of trehalose monomycolate, heme import, siderophore export, and sulfolipid transport (73). In *Mycobacterium abscessus* and *Mycobacterium bolletii*, MmpL4a and MmpL4b were reported to transport glycopeptidolipids (GPL) to the bacterial surface, contributing to cell envelope integrity and helping to regulate membrane permeability (9, 48). MmpL5 has been shown to actively export antitubercular drugs, directly contributing to antibiotic resistance (35). Orthologs of MmpL5 have been identified in non-tuberculosis species such as *Mycobacterium intracellulare*, where they confer a similar increase in resistance to anti-TB drugs (1). How MmpL efflux pumps are genetically regulated is still being clarified (13, 35, 56).

All sequenced *Mycobacterium* genomes, except for *M. leprae*, contain genes *mmpS5* and *mmpL5*. The MarR-like transcription factor Rv0678 was found to regulate the *mmpS5-mmpL5* operon in *M. tuberculosis* (56). Mutations in Rv0678 resulted in upregulation of the *mmpL5* efflux pump, conferring cross-resistance to both clofazimine and bedaquiline (35). Bioinformatic analysis of the *mmpS5-mmpL5* operon locus in non-tuberculous *Mycobacterium* species revealed either homologs to Rv0678 (MarR-like transcription factors), or nonhomologous TetR transcription factors adjacent to *mmpSL* (13). Recent studies have shown the trans-acting tetracycline repressors (TetR) play an important role in the regulation of mycobacterial efflux pumps (12, 18, 25). TetR are homodimers with a helix-turn-helix (HTH) motif and were first described conferring resistance to tetracycline (57). TetR has two DNA-binding domains (one per monomer) and a regulatory core responsible for tetracycline recognition and dimerization. When tetracycline binds to a cavity in the outer helices of the regulatory domain, there is a
conformational change that affects the DNA-binding domain so that TetR is no longer able to bind DNA, resulting in expression of efflux pump TetA, and subsequent Tetracycline resistance. In M. smegmatis, the TetR-like transcriptional regulator MSMEG_3765 was found to be induced under acid-nitrosative stress and negatively regulate the upstream efflux pump MSMEG_3760 (54). Orthologs of MSMEG_3760 were identified in M. marinum, and M. avium, but have not yet been characterized. The role of these trans-acting regulatory proteins in non-tuberculous Mycobacterium infections still needs to be clarified as it could provide new targets for drug therapy in antibiotic resistant infections. In Chapter 2, we characterize TetR regulation of the mmpS5-mmpL5 operon in M. intracellulare, which is one of the first studies of a TetR regulator in relation to mmpS5-mmpL5.

1.2.2 Horizontal transfer of antibiotic resistance genes

The horizontal acquisition of antibiotic resistance genes has been well documented in several pathogens and serves as the primary mechanism for spread of antimicrobial resistance (34). Horizontal gene transfer (HGT) is defined as the movement of genetic material between organisms, distinct from the vertical DNA transmission from parent to offspring. The horizontal transfer of genetic material adds a whole new dimension to bacteria pathogenicity, where an antibiotic resistant gene can “jump” between hosts and rapidly spread resistance among multiple unrelated pathogens (21). There are three main mechanisms in which HGT can potentiate antibiotic resistance: conjugation by plasmids, transduction by a bacteriophage, or uptake of exogenous DNA by natural competence (71). Bacterial conjugation is considered the most prevalent mechanism of antibiotic resistance gene dissemination (8); however, transduction and
natural competence must not be overlooked. Natural competence is poorly understood in several pathogenic bacteria yet may play an important role in the acquisition of antibiotic resistant genes. For the purpose of this project, natural competence in *Salmonella* will be further discussed.

### 1.2.2.1 Regulation of genes for natural competence in *Salmonella*

Prokaryotic natural competence is defined by a bacterium’s ability to bind and take up extracellular DNA (49). Internalized DNA is either recombined with chromosomal DNA, potentially increasing genetic diversity, or DNA is completely degraded into nucleotides that are reused for DNA synthesis (23, 26, 58). Each of these physiological functions can increase the evolutionary fitness of the bacteria.

Natural competence was first detected in *Streptococcus pneumoniae* in 1928 by Frederick Griffith; today this phenomenon has been identified in more than 80 bacterial species (11, 33). Prominent examples of naturally competent bacteria include Gram-positive bacteria *Bacillus subtilis* and *S. pneumoniae* and Gram-negative bacteria *Haemophilus influenzae* and *Neisseria* (24, 36, 46, 68). Other well-studied bacteria like *Escherichia coli* and *Salmonella enterica* encode homologs of genes required for natural competence, but to date these species have not been demonstrated to be naturally competent in laboratory or natural settings (16, 65). The mechanism of competence induction, initiation and regulation differ greatly between groups of naturally competent bacteria. For example, natural competence is constitutive in *Helicobacter pylori* and *Neisseria gonorrhoea* but tightly regulated in *H. influenzae* (37, 67). Environmental signals such as nutrient starvation are required to induce *H. influenzae* competence, while
nutrient starvation and bacterial communication via quorum sensing is important for competence induction in *Vibrio cholerae* (59, 69).

Gram-negative and Gram-positive bacteria use a homologous system to transport DNA across the cytoplasmic membrane, and Gram-negative bacteria have evolved an expanded transport mechanism for DNA uptake across the outer membrane (20). The mechanism of DNA-uptake in Gram-negative bacteria is characterized by a Type IV pili (T4P) system (38, 62). T4P are extracellular appendages that are 6-9 nm in diameter and assemble up to several micrometers in length (10). T4P allow bacteria to adhere to surfaces, form bacterial aggregates, and are involved in cellular invasion, bacterial motility, and phage and DNA uptake (38). In naturally competent Gram-negative bacteria, the Type 4 pilus is proposed to elongate into the extracellular environment through polymerization of pilin proteins and bind double-stranded DNA (20, 22, 31, 42). After binding DNA, the pilus depolymerizes and retracts the bound dsDNA into the periplasmic space. Once inside the periplasmic space, DNA-binding proteins direct the DNA to the inner membrane channel Rec2, where one strand is degraded, and the other passes into the cytoplasm. The single-stranded DNA (ssDNA) is either degraded by cytoplasmic endonucleases or used in homologous recombination.

Figure 1.2 outlines the proposed mechanism of T4P DNA uptake and highlights important proteins involved in natural competence (4). Competence genes *pilA* and *comN* code for prepilin proteins, which polymerize to form the Type 4 pilus structure (64) (Figure 1.2). *comA* is a periplasmic protein that is predicted to translocate DNA across the inner membrane (29, 64). *hopD* codes for a Type 4 leader peptidase, which cleaves prepilin into mature pilus subunits at the inner membrane (30, 51). Figure 1.3 shows the
Figure 1.2. Uptake mechanism of DNA using Type IV Pilus.
This figure is from (4). The proposed DNA uptake machinery in S. Typhimurium. PilA and ComN makeup the Type IV pilus (T4P). ComA helps assembly of pilus through the outer membrane pore ComE. Once T4P binds to DNA, PilT retracts pilus into the periplasmic space where DNA is dissociated into two strands. One strand is degraded while the other is passed into the cytoplasm through Rec2. The competence proteins are highlighted in yellow.

Figure 1.3. Alignment of competence genes from four members of proteobacteria.
This figure is from (4). Genes in red are the regulators of the competence genes. *comA* shown in blue is the focus of this thesis. The abbreviations are as follows: Pasteur, *Pasteurellaceae*; Entero, *Enterobacteriaceae*; Vibrio, *Vibrionaceae*; *H.i.*, *Haemophilus influenzae*; *E.c.*, *Escherichia coli*; *S.e.*, *Salmonella enterica*; *V.C.*, *Vibrio cholerae*. 
operon structure and orientation of competence genes on the chromosome of

*Haemophilus influenzae*, *Escherichia coli*, *Salmonella enterica*, and *Vibrio cholerae*.

Positive regulation of competence genes in *Salmonella* Typhimurium is proposed to depend on two regulatory proteins, cAMP receptor protein (CRP) and Sxy (16). CRP is a global regulatory protein that acts primarily as a transcription activator through direct protein-protein interaction with RNAP (14). If the cell is starved for its preferred sugar sources, cAMP levels are elevated, and cAMP binding to CRP induces a conformational change that allows CRP to bind to specific sites in DNA (72). Once bound to DNA, CRP recruits RNAP by interacting with RNAP’s α C-terminal domain, improving the binding affinity of RNAP and the subsequent activation of transcription (63). CRP regulates the expression of more than 490 genes in *E. coli* (32). CRP binding sites can be differentiated into either CRP-N sites or CRP-S sites; CRP-N sites only require CRP for activation whereas CRP-S sites require both CRP and Sxy for transcriptional activation (16).

Competence gene promoters contain CRP-S sites, consistent with a requirement for both CRP and Sxy for expression of competence genes.

The Sxy protein was first identified in *H. influenzae* and was demonstrated to be essential for induction of natural competence (59). It is proposed that Sxy stabilizes the binding of CRP to CRP-S sites by increasing DNA curvature, leading to the recruitment of RNAP and the expression of competence genes (15). Expression of *H. influenzae sxy* is regulated at both the transcriptional level and post-transcriptional level. Transcription is induced under starvation conditions by CRP. In *S. Typhimurium*, it was found that *sxy* expression is activated by CRP, as observed in *H. influenzae* (59), and is repressed by nucleoid-associated proteins, FIS and H-NS, as well as the alternative sigma factor RpoS.
At the post-transcriptional level, the sxy mRNA forms a secondary structure that impedes translation and thus prevents production of the Sxy protein (17). This cis-acting regulatory mechanism is proposed to regulate Sxy levels according to nucleotide availability (66). Using the bioinformatic tool Mfold, we have predicted strong mRNA secondary structures in the promoter regions of S. Typhimurium competence genes sxy and comA (4). These mRNA secondary structures could explain the silent promoter activity of sxy and comA in S. Typhimurium, and possibly reveal why natural competence has not been detected in S. Typhimurium. By weakening these mRNA stem-loops with site-directed mutagenesis, we could possibly upregulate the translation of sxy and comA and make S. Typhimurium naturally competent. In chapter 3, we characterize an mRNA stem-loop in the comA promoter region which negatively regulates comA expression in S. Typhimurium. In addition to weakening the comA mRNA stem-loop, Sxy overexpression is required for comA gene induction.

1.2.3 Golden Gate cloning

Golden Gate cloning is a molecular biology technique that uses Type IIS restriction enzymes and T4 DNA ligase to directionally assemble multiple DNA fragments into a plasmid vector backbone in one reaction (27). Type IIS restriction enzymes recognize non-palindromic asymmetric DNA sites and cut outside of their recognition sequences, generating 4 bp overhangs. Designing a series of different 4 bp overhangs specifies the order of assembly of multiple DNA fragments that are simultaneously ligated into a vector (Figure 1.4B). The desired ligation products lose the Type IIS restriction sites during the Golden Gate cloning reaction, resulting in a ligation efficiency close to 100%.
Figure 1.4. Hierarchical assembly of multigene constructs using MoClo toolkit.

This figure is taken from (74). **A** shows the level 0 modules containing promoters (P), 5’ untranslated regions (U), signal peptides (SP), coding sequences (CDS) and terminators (T). Level 0 modules are assembled into a level 1 transcription unit by the Golden Gate cloning reaction. Level 2 multigene constructs are assembled from multiple level 1 modules using the same Golden Gate cloning reaction. **B** shows the 4 nucleotides flanking each level 0 module, which specify their position in the level 1 transcription unit. **C** shows two examples of the different combinations of level 0 modules available.
The modular cloning (MoClo) system developed by the Sylvestre Marillonnet's lab uses Golden Gate cloning for the hierarchical assembly of DNA fragments (74). The MoClo system comprises of three different modules: level 0, level 1, and level 2 (Figure 1.4A). Level 0 constructs contain one of the five standard transcriptional elements: a promoter, a 5′ untranslated region (UTR), a signal peptide (SP), a coding sequence (CDS), or a terminator. Level 0 modules are directionally assembled into a level 1 module using Golden Gate cloning, forming one transcriptional unit. In a second cloning step, multiple level 1 constructs can be assembled into a level 2 multigene construct. The MoClo system was initially developed for the assembly of eukaryotic multigene constructs, however recent kits have been adapted specifically for *E. coli* synthetic biology (40, 50). Here we adapt the Sylvestre Marillonnet MoClo toolkit to study *S. Typhimurium* and *M. intracellularare* gene regulation. This involved the construction of a level 0 lux*CDABE* reporter module to measure promoter activity, the development of level 1 *com*A promoter constructs to study mRNA secondary structure cis-regulation, and the assembly of level 2 multigene constructs to study trans-regulation of promoter activity by transcription factors MmpT5 and Sxy.

1.2.4 Alternative antimicrobial agents

In a post-antibiotic era, alternatives to antibiotics will be key in staving off catastrophe. The current dissemination of antibiotic resistant genes via HGT in conjunction with bacterial efflux mechanisms have contributed to a new era of multidrug resistant pathogens (34, 55). New antibiotics are desperately needed in both animal agriculture and human medicine. Alternatives to antibiotics are a possible solution; however, more research is needed before they can be effectively implemented.
Current antibiotic alternatives are categorized as either disease prevention or disease treatment (3). Disease prevention includes the use of vaccines and probiotics, whereas disease treatment uses phages, endo- or exolysins, bacteriocins, and predatory bacteria to directly kill pathogens. There are advantages and limitations to each antibiotic alternative; one of the most promising alternatives is phage therapy, which is the focus of Chapter 4.

1.2.4.1 EHEC phage therapy

Enterohemorrhagic *Escherichia coli* (EHEC) serogroup O157:H7 is a food-borne pathogen considered to be a global health threat responsible for many disease outbreaks (5). Approximately 100,000 people are infected with EHEC each year in the U.S.A., primarily through improper food preparation (28). Antibiotic treatment of EHEC is problematic because this can induce Shiga toxin (Stx) production, which causes bloody diarrhea and hemolytic uremic syndrome (60). Combined with the increasing threat of antibiotic resistance, researchers are looking to other treatment options such as phage therapy.

Phage therapy is the clinical use of bacteriophages to kill pathogenic bacteria and treat disease. This antimicrobial treatment has been studied for over 50 years and is implemented widely in Eastern Europe (2). Phage therapy is advantageous in treating EHEC because of its ability to kill the bacteria without inducing Shiga toxin production (43, 70). The *E. coli* O157:H7 Shiga toxin is located on a lysogenic phage and induced by the SOS response (39). Damage to chromosomal DNA activates RecA which cleaves LexA and prophage repressors, resulting in bacteriophage and Stx production. It is
hypothesized that lysing EHEC without triggering the SOS response would prevent Stx production and subsequent tissue damage.

Advantages of phage therapy include low economic cost, low inherent toxicity, high effectiveness against antibiotic resistance bacteria, and low environmental impact, however challenges such as narrow host range, bioavailability, and phage resistance still must be addressed (45). Recent research is using “omics” data and machine learning predicts which phages would be effective against pathogenic bacteria (44). This is one possible solution to efficiently interpret the countless number of phage-host interactions and help formulate an effective phage cocktail for clinical use.

1.3 Thesis Objective

This manuscript-based thesis seeks to characterize gene regulation contributing to antibiotic resistance in *M. intracellulare* and natural competence in *S. Typhimurium* using a modular cloning toolkit. Each chapter focuses on one bacterial system and addresses several hypotheses as follows: (1) MmpT5 is a transcriptional repressor of the mmpSL promoter. (2) Nonsynonymous mutations in mmpT5 result in upregulation of mmpSL. (3) The comA mRNA contains a secondary structure that represses expression. (4) Sxy induces expression of comA.

Chapter 2 describes the characterization of the mmpT5 gene in *M. intracellulare*. Synthetic multigene constructs are assembled to show the repressive effect of MmpT5 on the mmpSL promoter. Nonsynonymous mutations in mmpT5 from BDQ-resistant isolates resulted in the eradication of this repressive effect.

Chapter 3 uses synthetic multigene constructs to demonstrate the regulation of *S. Typhimurium* competence gene comA. Directed mutations in the predicted comA mRNA
secondary structure resulted in up to an 8-fold increase in \textit{comA::lux} luminescence. This effect was reversed with compensatory mutations designed to confirm the mRNA secondary structure. Level 2 constructs with promoter-\textit{lux} fusions and inducible \textit{sxy} showed a positive regulation effect of Sxy on competence genes.

The final chapter describes the research conducted between September 2017 and December 2017 at the University of Edinburgh on the Queen Elizabeth II Diamond Jubilee Scholarship and investigated phage sensitivity and resistance of \textit{E. coli} O157:H7. Tetrazolium reduction was used to measure EHEC phage sensitivity and a tyrosine kinase-mediated phage resistance mechanism was characterized.

REFERENCES


   *Bacteriophage*, 1, 111–114.

46. Macfadyen Leah P., Chen David, Vo Hung C., Liao David, Sinotte Ryan and 
   Redfield Rosemary J. (2001) Competence development by *Haemophilus influenzae* 
   is regulated by the availability of nucleic acid precursors. *Molecular Microbiology*, 
   40, 700–707.


50. Moore, S.J., Lai, H.-E., Kelwick, R.J.R., Chee, S.M., Bell, D.J., Polizzi, K.M. and 
   Freemont, P.S. (2016) EcoFlex: A Multifunctional MoClo Kit for *E. coli* Synthetic 


54. Perrone, F., De Siena, B., Muscariello, L., Kendall, S.L., Waddell, S.J. and Sacco, M. 
   (2017) A Novel TetR-Like Transcriptional Regulator Is Induced in Acid-Nitrosative 


56. Radhakrishnan, A., Kumar, N., Wright, C.C., Chou, T.-H., Tringides, M.L., Bolla, J.R., 
   the Transcriptional Regulator Rv0678 of *Mycobacterium tuberculosis*. *J. Biol. 
   Chem.*, 289, 16526–16540.

57. Ramos, J.L., Martínez-Bueno, M., Molina-Henares, A.J., Terán, W., Watanabe, K., 


76. WHO | Tuberculosis *WHO*. http://www.who.int/mediacentre/factsheets/fs104/en/
Chapter 2: Transcriptional Regulator $mmpT5$ Represses Efflux Pump Expression in

*Mycobacterium intracellulare* and Provides a Mechanism for Bedaquiline Resistance.
ABSTRACT

Bedaquiline (BDQ), a diarylquinoline antibiotic that targets ATP synthase, is effective in treating multidrug-resistant tuberculosis (MDR-TB) and non-tuberculous mycobacterial infections. A recent study using BDQ to treat *Mycobacterium intracellulare* lung infections observed 7 out of 13 patients suffering a relapse in infection during treatment (1). Whole-genome sequence comparison between pretreatment and drug-resistant relapse isolates revealed that in each infected individual, *M. intracellulare* had independently evolved resistance, and all resistant isolates carried nonsynonymous mutations in the uncharacterized locus *mmpT5*. The gene *mmpT5* is annotated as a TetR transcriptional regulator, located upstream of the Resistance Nodulation Division (RND) efflux pump *mmpSL*. Previous studies have shown a regulatory role of TetR transcription factors on mycobacterial efflux pumps. In this study, we use Golden Gate cloning to generate multigene constructs with inducible *mmpT5* and a *mmpSL* promoter-luciferase reporter fusion. Optimization of the Golden Gate cloning technique allowed us to efficiently exchange *mmpT5* mutations from different clinical isolates. Luciferase-reporter assays showed that *mmpT5* encodes a transcriptional repressor which represses *mmpSL* efflux pump genes. Bedaquiline-resistant clinical mutations in *mmpT5* resulted in upregulation of the *mmpSL* efflux pump, suggesting a mechanism of bedaquiline resistance.
2.1 INTRODUCTION

The diarylquinoline antibiotic Bedaquiline (BDQ) is currently one of the most promising novel antimycobacterials, developed primarily for the treatment of multidrug-resistant (MDR) and extensively drug-resistant (XDR) Tuberculosis (TB) (9). BDQ inhibits ATP synthesis by binding to the \textit{atpE} encoded subunit c rotor ring complex of mycobacterial ATP synthase preventing AtpE rotation and proton transfer, effectively killing the bacteria through energy deprivation (23). Clinical studies have shown successful treatment of MDR-TB and XDR-TB cases by BDQ in combination therapy, however resistance to BDQ has also been observed (14). Emergence of BDQ-resistant strains in \textit{Mycobacterium tuberculosis} has been attributed to rare mutations in \textit{atpE}, or more common mutations in Rv0678, a MarR family transcriptional regulator that controls the expression of the Resistance Nodulation Division (RND) MmpS5-MmpL5 efflux pump (12). Upregulation of the MmpL5 efflux pump through inactivation of Rv0678 has shown increased resistance to anti-TB drugs Clofazimine, Bedaquiline, and PepQ.

BDQ shows antimicrobial activity against selective non-tuberculous mycobacteria (NTM) such as \textit{Mycobacterium intracellulare}. The activity of BDQ against all NTM species is variable due to the wide heterogeneity of the \textit{atpE} gene in the \textit{Mycobacterium} genus (23). Recent studies have shown BDQ to be moderately effective in treating \textit{Mycobacterium avium} complex (MAC) infection of the lungs. In 2015, Philley et al. (17) used BDQ as salvage therapy for MDR-NTM lung infections. After a 6-month treatment period, half of the patients showed one or more negative sputum cultures and 6 out of 10 patients showed a microbiological relapse. In more recent work, Alexander et al. (1) treated 16 MAC patients with BDQ, which resulted in 7 out of 13 patients suffering a
relapse in infection during treatment. Whole-genome sequence comparison between pretreatment and drug-resistant relapse isolates revealed that in each infected individual, *M. intracellulare* had independently evolved resistance, and all resistant isolates carried nonsynonymous mutations in the uncharacterized locus *mmpT5*. The gene *mmpT5* is annotated as a TetR transcriptional regulator, located upstream of the RND efflux pump *mmpSL*, suggesting a mechanism of Bedaquiline resistance.

RND efflux pumps actively transport many antibiotics out of the cell and have been shown to play a role in BDQ-resistant TB infections. The TetR superfamily of transcriptional regulators have demonstrated a repressive mechanism in the regulation of mycobacterial efflux pumps (16, 19). Bioinformatic analysis shows homology between the *mmpT5-mmpSL* operon of *M. intracellulare* and the drug efflux operon *mmpS5-mmpL5* in *M. tuberculosis* (4). We predict MmpT5 to regulate expression of *mmpSL* by binding to the promoter region, similar to Rv0678. Nonsynonymous mutations in *mmpT5*’s DNA-binding domain or frameshift mutations resulting in a truncated MmpT5 protein could result in the upregulation of *mmpSL* and subsequent BDQ resistance.

Bioluminescence assays of promoter-/*luxCDABE* fusions are considered quick and reliable methods of measuring gene expression. In this study, we adapt the MoClo Toolkit (22) to construct inducible MmpT5 and a *mmpSL* promoter::*lux* fusion in a one and two plasmid system. The optimization of the Golden Gate cloning technique allowed us to efficiently and quickly exchange *mmpT5* mutations from different clinical isolates. Using Bioluminescence assays, we show that the TetR regulator *mmpT5* regulates expression of the RND efflux pump *mmpSL*, and clinical *mmpT5* mutations result in the de-repression of drug efflux, leading to drug resistance. The applicability of the MoClo
Toolkit in this study suggests future roles for synthetic biology in studying prokaryotic trans-gene regulation.

2.2 MATERIALS AND METHODS

2.2.1 Bacterial Strains and Plasmids

The bacterial strain and plasmids used in this study are listed in Table 2.1.

*E. coli* DH5-Alpha was grown in Luria-Bertani (LB) broth at 37°C with shaking.

Spectinomycin (50 μg/mL), Ampicillin (100 μg/mL), Kanamycin (50 μg/mL), Isopropyl β-D-1-thiogalactopyranoside (IPTG 1 mM), and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal 50 μg/mL) were used for selection and screening of correct plasmid constructs. Golden Gate plasmids used were obtained from the MoClo Toolkit (Addgene) deposited by the Sylvestre Marillonnet lab. The software Geneious® was used to visualize *M. intracellulare* DNA sequence, design Golden Gate primers, and validate Golden Gate reactions with the Golden Gate plugin. Golden Gate primers were ordered through Eurofins Genomics. Representative primer sequences used to generate all experimental constructs are presented in Supplementary Table 2.1.

2.2.2 Bioinformatic analysis

The Basic Local Alignment Search Tool (BLAST) was used to identify DNA sequences similar to the *mmpT5* gene.

Alignment of *mmpT5* DNA and protein sequences was carried out using Kalign, which uses the Wu-Manber string-matching algorithm for fast and accurate multiple sequence alignment. The output format was ClustalW and parameters were set at default.

2.2.3 PCR amplification of *mmpT5* and *mmpSL*
Table 2.1. Bacterial plasmids and strains used in this study.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bacterial strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td>High-efficiency competent cells used for Golden Gate cloning and studying <em>M. intracellulare</em> gene expression.</td>
<td>(21)</td>
</tr>
<tr>
<td>DH5-Alpa</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>plCH41308</td>
<td>level 0 landing vector for SC module</td>
<td>(22)</td>
</tr>
<tr>
<td>plCH47732</td>
<td>level 1 landing vector (position 1*)</td>
<td></td>
</tr>
<tr>
<td>plCH47742</td>
<td>level 1 landing vector (position 2*)</td>
<td></td>
</tr>
<tr>
<td>plCH47751</td>
<td>level 1 landing vector (position 3*)</td>
<td></td>
</tr>
<tr>
<td>pAGM4673</td>
<td>level 2 landing vector (pUC19 ORI)</td>
<td></td>
</tr>
<tr>
<td>plCH82094</td>
<td>level 2 landing vector (p15A ORI)</td>
<td></td>
</tr>
<tr>
<td>plCH54022</td>
<td>Dummy gene for position 2* in level 2 vectors</td>
<td></td>
</tr>
<tr>
<td>plCH41766</td>
<td>level 2 end-linker</td>
<td></td>
</tr>
<tr>
<td>pOGG031</td>
<td>LacZ promoter cloned into level 0 vector</td>
<td>This study</td>
</tr>
<tr>
<td>pOGG039</td>
<td>T7 terminator cloned into level 0 vector</td>
<td></td>
</tr>
<tr>
<td>level 0 Lux</td>
<td><em>luxCDABE</em> (5846 bp) inserted into plCH41308, a level 0 landing vector for SC modules</td>
<td>This study</td>
</tr>
<tr>
<td>level 1 LacI</td>
<td><em>lacI</em> promoter and LacI gene inserted into plCH47751</td>
<td></td>
</tr>
</tbody>
</table>

*position refers to orientation of insert when cloned into level 2 landing vector*
Chromosomal DNA extracted from *M. intracellular* pretreatment and BDQ-resistant relapse isolates were provided by David Alexander (1), and used as template for PCR amplification of *mmpT5* and *mmpSL*. PCR amplification of the full *mmpT5* gene was performed using Q5® High-Fidelity DNA Polymerase (New England Biolabs), 1X Q5 reaction buffer, 1X GC enhancer, 10 mM dNTPs, 10 µM of each primer set shown in Table S1, and ~25 ng chromosomal DNA. Reaction mixtures were scaled to 50 µL and incubated for 30 s at 98°C (initial denaturation), followed by 35 cycles of 30 s at 98°C (denaturation), 30 s at 63°C (annealing), and 1 min at 72°C (extension), with a final elongation step at 72°C for 2 minutes. PCR amplification of the *mmpSL* promoter (317 bp upstream of initiation codon) was performed using the same reaction setup and thermocycler settings as described for *mmpT5*.

The *mmpT5* and *mmpSL* PCR products were then isolated on agarose gels and purified using the QIAquick Gel Extraction Kit (Qiagen). The purified DNA was eluted in 30 µl of Elution Buffer and quantified using a Nano Drop® Spectrophotometer ND-1000 (Thermo Scientific™).

2.2.4 Golden Gate Assembly

Golden gate reactions were set up following the MoClo cloning protocol (22). *mmpT5* contains an internal BbsI site, and therefore was amplified into two parts with overlapping primers to remove the restriction site through a silent mutation. For level 0 *mmpT5* cloning, reactions were prepared by pipetting in one tube 40 fmol of the gel-purified *mmpT5* part 1 PCR product, 40 fmol of the gel-purified *mmpT5* part 2 PCR product, 40 fmol of pICH41308, 10U of type IIS restriction enzyme BbsI (New England Biolabs), 20U of high concentration T4 Ligase (New England Biolabs), 1.5 µL of 2
mg/mL Bovine Serum Albumin (BSA), 1.5 µL of 10X T4 DNA Ligase Buffer (New England Biolabs), and Milli-Q® H₂O (Millipore) up to a total volume of 15 µL. This procedure was repeated three times using *mmpT5* from a different clinical isolate (BDQ1-WT, BDQ1-Fs, BDQ2-WT, BDQ2-137G) for each reaction (4 total). The reactions were incubated in a thermocycler for 25 cycles at 37°C for 3 min and 16°C for 4 min followed by 5 min at 50°C and 5 min at 80°C. The restriction-ligation mixtures were added to 100 µL chemically competent DH5α cells, incubated for 30 min on ice and transformed by heat shock. Then 900 µL of SOC media was added to the transformation, and the cells were allowed to recover for 1 hour at 37°C with shaking. Next 200 µL of the transformation were plated on LB plates containing spectinomycin (50 µg/mL) and X-gal/IPTG and incubated overnight at 37°C. Four white recombinant colonies were then selected from the plates and streaked on separate spectinomycin plates. Each of the four isolated recombinant colonies were inoculated in 5 mL LB broth and grown overnight. Plasmids were isolated using the EZ-10 Spin Column Plasmid DNA Miniprep Kit (Biobasic). The level 0 *mmpT5* constructs were confirmed by restriction digest with DraIII-HF (New England Biolabs) and PCR amplification of the *mmpT5* region using the level 0 clone as template.

The level 1 inducible *mmpT5* construct was prepared by pipetting in one tube 40 fmol of the confirmed level 0 *mmpT5* plasmid, 40 fmol of pOGG031, 40 fmol of pOGG039, 40 fmol of pICH47732, 10U of type IIS restriction enzyme BsaI (New England Biolabs), 20U of high concentration T4 Ligase (New England Biolabs), 1.5 µL of 2 mg/mL BSA, 1.5 µL of 10X T4 DNA Ligase Buffer (New England Biolabs), and Milli-Q® H₂O (Millipore) up to a total volume of 15 µL. The level 1 *mmpSL::lux*
construct was prepared by pipetting in one tube 40 fmol of the gel-purified \textit{mmpSL} promoter PCR product, 40 fmol of a level 0 Lux, 40 fmol of pOGG039, 40 fmol of pICH47732, 10U of BsaI (New England Biolabs), 20U of high concentration T4 Ligase (New England Biolabs), 1.5 $\mu$L of 2 mg/mL BSA, 1.5 $\mu$L of 10X T4 DNA Ligase Buffer (New England Biolabs), and Milli-Q® H$_2$O (Millipore) up to a total volume of 15 $\mu$L. The reactions were incubated in a thermocycler at the same settings for level 0 cloning previously mentioned. The restriction-ligation mixtures were transformed into chemically competent DH5$\alpha$ as described above. Transformations were plated on ampicillin plates (100 $\mu$g/mL) with X-gal/IPTG and incubated overnight at 37°C. Recombinant colonies were restreaked and the level 1 plasmids were isolated as described above. The level 1 \textit{mmpT5} and \textit{mmpSL} promoter::\textit{lux} constructs were confirmed by restriction digest with DraIII-HF (New England Biolabs) and PCR amplification of either \textit{mmpT5} or the \textit{mmpSL} promoter region using the level 1 clones as template.

Final level 2 constructs were cloned into either pAGM4673 for the 1 plasmid setup or pICH82094 for the 2 plasmid setup. The level 2 construct for the 1 plasmid experimental setup was prepared by pipetting in one tube 40 fmol of the confirmed level 1 Plac.\textit{mmpT5}.T7, 40 fmol of \textit{mmpSL} promoter::\textit{lux}.T7, 40 fmol of level 1 LacI, 40 fmol of pICH41766, 40 fmol of pAGM4673, 10U of BbsI (New England Biolabs), 20U of high concentration T4 Ligase (New England Biolabs), 1.5 $\mu$L of 2 mg/mL BSA, 1.5 $\mu$L of 10X T4 DNA Ligase Buffer (New England Biolabs), and Milli-Q® H$_2$O (Millipore) up to a total volume of 15 $\mu$L. The level 2 construct for the 2 plasmid experimental setup was prepared by pipetting in one tube 40 fmol of the confirmed level 1 Plac.\textit{mmpT5}.T7, 40 fmol of pICH54022, 40 fmol of level 1 LacI, 40 fmol of pICH41766, 40 fmol of
pICH82094, 10U of BbsI (New England Biolabs), 10U of BsaI (New England Biolabs), 20U of high concentration T4 Ligase (New England Biolabs), 1.5 µL of 2 mg/mL BSA, 1.5 µL of 10X T4 DNA Ligase Buffer (New England Biolabs), and Milli-Q® H₂O (Millipore) up to a total volume of 15 µL. The reactions were incubated in a thermocycler at the same settings for level 0 cloning previously mentioned. The restriction-ligation mixtures were transformed into chemically competent DH5α as described above. Transformations were plated on kanamycin plates (50 µg/mL) with X-gal/IPTG and incubated overnight at 37°C. Recombinant colonies were restreaked and the level 2 plasmids were isolated as described above. Final level 2 constructs were sent for Sanger sequencing (Eurofins Genomics) for confirmation.

2.2.5 Construction of plasmids for expression in *M. smegmatis*

The *mmpT5-mmpSL* operon from the four clinical isolates was cloned into mycobacterial shuttle vector pMD31 for expression in *M. smegmatis*. PCR amplification of a 4.45 kb product was done using the isolates as template DNA and primers containing *BamHI* sites. The gel-purified PCR product was digested with *BamHI* and cloned into *BamHI*-cut pMD31. Successful clones were electroporated into *M. smegmatis*.

2.2.6 Luciferase-Reporter Assay

Measurement of *mmpSL* promoter activity was accomplished by bioluminescence and OD₆₀₀ readings using a Synergy HT Multi-Detection Microplate Reader (Bio-Tek). DH5α cells with correct level 2 constructs were inoculated in 5 mL LB broth (Lennox) and incubated at 37°C overnight with shaking at 200 rpm. Overnight cultures were diluted 1:1000 (v/v) in LB broth (Lennox) with Kanamycin (50 µg/mL) and IPTG (1 mM) to induce expression of *mmpT5*. The diluted cells containing the constructs were
plated in triplicates in a 96-well plate at 250 µL per well with 50 µL mineral oil overlaid on top. Bioluminescence was measured for 48 hours at 10 min intervals. For bioluminescence measurements, Synergy HT Multi-Detection Microplate Reader (BioTek) sensitivity gain setting was 100, with no excitation and no emission filter in place. Data was collected using Gen 5 2.07 (BioTek Instruments), exported to Microsoft Excel, and plotted using Graphpad Prism 7. Bioluminescence readings were normalized to cell density by dividing by the corresponding OD$_{600}$ values.

2.3 RESULTS

2.3.1 Sanger sequencing confirms nonsynonymous mutations in $mmpT5$ cloned from relapse isolates

*M. intracellulare* relapse isolates recovered from patients undergoing BDQ treatment had elevated BDQ MIC values and contained nonsynonymous mutations in the uncharacterized gene $mmpT5$. Table 2.2 summarizes the mutations and corresponding BDQ MICs in the pretreatment and relapse isolates. BDQ1-WT and BDQ2-WT represent the pretreatment isolates from patient 1 and 2 respectively. BDQ1-Fs and BDQ2-137G represent the relapse isolates from patient 1 and 2 respectively. Bioinformatic analysis of the $mmpT5$-$mmpSL$ operon showed homology to the drug efflux operon $mmpS5$-$mmpL5$ in *M. tuberculosis*, suggesting a possible mechanism of BDQ resistance. Thus, we hypothesized that the wild type $mmpT5$ represses $mmpSL$ expression, and nonsynonymous $mmpT5$ mutations negatively affect repression resulting in $mmpSL$ upregulation.

Sanger sequencing of final multigene constructs confirmed $mmpT5$ mutations in *M. intracellulare* relapse isolates BDQ1-Fs and BDQ2-137G. Figure 2.1 illustrates the
Table 2.2. BDQ MIC from (1) and mmpT5 mutations of clinical pretreatment and relapse isolates.

<table>
<thead>
<tr>
<th>Patient</th>
<th>Isolate</th>
<th>Treatment status (mo)</th>
<th>BDQ MIC (µg/mL)</th>
<th>mmpT5 Mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>BDQ1-WT</td>
<td>Pretreatment</td>
<td>0.004</td>
<td>WT</td>
</tr>
<tr>
<td>1</td>
<td>BDQ1-Fs</td>
<td>On BDQ (5)</td>
<td>0.008</td>
<td>c.194_195insA</td>
</tr>
<tr>
<td>2</td>
<td>BDQ2-WT</td>
<td>Pretreatment</td>
<td>0.004</td>
<td>WT</td>
</tr>
<tr>
<td>2</td>
<td>BDQ2-137G</td>
<td>On BDQ (9)</td>
<td>0.03</td>
<td>c.137T&gt;G</td>
</tr>
</tbody>
</table>

Note: MIC: minimum inhibitory concentration.

Figure 2.1. Location and characterization of DNA mutations in mmpT5

MmpT5 is 609 bp long and predicted to be part of the TetR superfamily of transcriptional regulators. Sanger sequencing of assembled multigene constructs confirmed mutations Alexander et al. (1) found in M. intracellular relapse isolates. Isolate BDQ1-Fs shows a frameshift mutation resulting from the insertion of an Adenine nucleotide at position 195 in the mmpT5 coding sequence. The frameshift mutation produced a stop codon 6 bp downstream of the insertion point. Isolate BDQ2-137G showed a missense mutation 137 bp into mmpT5, changing the amino acid from valine to glycine.
Figure 2.2. Alignment of \textit{mmpT5} DNA and protein sequences from pretreatment and relapse isolates.

Multiple sequence alignment of isolates BDQ1-WT, BDQ1-Fs, BDQ2-WT, and BDQ2-137G with Kalign® highlights the nonsynonymous mutations in the \textit{mmpT5} DNA and peptide sequence. The start codon is highlighted in green and the stop codon is highlighted in red. Synonymous mutations between isolates are highlighted in yellow and nonsynonymous mutations are highlighted in red. \textbf{A} shows the alignment of the \textit{mmpT5} nucleotide sequence. \textbf{B} alignment of the \textit{mmpT5} peptide sequence.
location and characterization of the nonsynonymous \textit{mmpT5} mutations in isolates BDQ1-Fs and BDQ2-137G. Isolate BDQ1-Fs showed a frameshift mutation resulting from the insertion of an Adenine nucleotide at position 194 in the \textit{mmpT5} coding sequence. The frameshift mutation produced a stop codon 6 bp downstream of the insertion point. Isolate BDQ2-137G showed a missense mutation 137 bp into \textit{mmpT5}, changing the amino acid from valine to glycine. Alexander et al. (1) predicted this mutation to be in the DNA binding domain of MmpT5. The alignment of \textit{mmpT5} DNA and protein sequences from clinical isolates BDQ1-WT, BDQ1-Fs, BDQ2-WT, and BDQ2-137G is shown in Figure 2. BDQ1-WT and BDQ2-WT differ at four nucleotide positions in the \textit{mmpT5} coding region, nevertheless both \textit{mmpT5} alleles code for the same amino acid sequence.

\textbf{2.3.2 The \textit{mmpSL} promoter is expressed in \textit{E. coli}, but the \textit{mmpT5} promoter is not}

The challenging nature of \textit{M. intracellulare} lab work incentivized the development of a heterologous gene expression system in \textit{E. coli}. In order for this system to work, we first tested if the mycobacterial promoters of \textit{mmpT5} and \textit{mmpSL} are active in \textit{E. coli}. Thus, we constructed \textit{mmpT5}:\textit{lux} and \textit{mmpSL}:\textit{lux} promoter fusions to test the activity of these mycobacterial promoters in \textit{E. coli}. The promoter sequences of \textit{mmpT5} and \textit{mmpSL} were the same between pretreatment and relapse isolates within patients and therefore only BDQ1-WT and BDQ2-WT were tested. The \textit{mmpSL} promoter and 165 bp of the ORF were fused to the \textit{luxCDABE} operon to measure expression activity. Part of the \textit{mmpSL} ORF was included in the \textit{lux} fusion because of the possibility of MmpT5 binding within the coding sequence. The \textit{mmpSL} promoter was active in \textit{E. coli} (Figure 2.3B), enabling the use of an \textit{E. coli}-based system to study \textit{mmpSL} regulation. Unfortunately, the \textit{mmpT5} promoter did not generate detectable luminescence in \textit{E. coli}
Figure 2.3. Bioluminescence assay of mycobacterial promoters.
Promoter activity of *mmpT5* and *mmpSL* was monitored at 37°C for 48 hours. Luminescence was measured on a BioTek Synergy HT reader according to the manufacturer’s instructions. Promoter activity is expressed in arbitrary units and values are normalized to the number of cells. Data are represented by mean ± SD (n = 3). BDQ1-WT and BDQ2-WT represent the pretreatment isolates used for DNA amplification from patient 1 and 2 respectively. A shows the luminescence activity of the *mmpT5* promoter fused to *luxCDABE* in *E. coli*. B shows the luminescence activity of the *mmpSL* promoter fused to *luxCDABE* in *E. coli*. 
(Figure 2.3A), raising the concern that *E. coli* cannot use the mycobacterial promoter to generate MmpT5. To ensure expression, *mmpT5* was placed under the control of the inducible *lacZYA* promoter so that *mmpT5* expression could be induced by the addition of IPTG to bacterial cultures.

### 2.3.3 Assembly of multigene reporter constructs using the MoClo toolkit

The MoClo toolkit allowed the systematic assembly of multigene constructs to test the hypothesis that MmpT5 is a repressor of *mmpSL* expression. Final multigene constructs contained the *mmpSL* promoter fused to the *lux* reporter and inducible *mmpT5* on one or two plasmids. We hypothesized that induction of *mmpT5* from pretreatment isolates would reduce *mmpSL* promoter activity, and induction of *mmpT5* from relapse isolates would have no effect on *mmpSL* promoter activity.

DNA samples from the four clinical isolates in Table 2.2 were used as template to amplify *mmpT5* and the *mmpSL* promoter. *mmpT5* contained an internal BbsI site, which was domesticated by designing overlapping primers that introduce a single silent nucleotide substitution in the recognition site. After removal of all internal BbsI and BsaI restriction sites in the cloned fragments, final constructs were assembled hierarchically as illustrated in Figure 2.4.

The PCR products were immortalized into level 0 vectors, creating a library of interchangeable modules. The *lacZYA* promoter was added upstream of the *mmpT5* module to control expression of *mmpT5*. T7 terminator sequences were added to the end of the *mmpT5* and *mmpSL::lux* constructs to prevent RNA polymerase (RNAP) read through. A *lacIq* construct was added to the level 2 clones to ensure the *lacZYA* promoter controlling *mmpT5* expression was fully repressed. The *mmpSL::lux* construct was then
Final gene constructs used in the 1 and 2 plasmid experimental setup were assembled in three steps at three levels. Type IIS cleavage sites are represented as boxes with the four-nucleotide sequence of the cleavage site. Arrow direction of genetic elements indicate orientation in module. PCR products were assembled into level 0 modules; level 0 modules were assembled into level 1 constructs; level 1 were assembled into final level 2 constructs, either for the 1 plasmid or 2 plasmid setups.

**Figure 2.4. Assembly of experimental constructs using MoClo Toolkit.**

Final gene constructs used in the 1 and 2 plasmid experimental setup were assembled in three steps at three levels. Type IIS cleavage sites are represented as boxes with the four-nucleotide sequence of the cleavage site. Arrow direction of genetic elements indicate orientation in module. PCR products were assembled into level 0 modules; level 0 modules were assembled into level 1 constructs; level 1 were assembled into final level 2 constructs, either for the 1 plasmid or 2 plasmid setups.
placed downstream of the inducible \textit{mmpT5} construct in one plasmid, or transformed as its own plasmid with the inducible \textit{mmpT5} construct on another plasmid. The two plasmid setup allowed us to put the inducible \textit{mmpT5} construct on a low copy plasmid and put the \textit{mmpSL::lux} construct on a high copy plasmid. The one plasmid setup controlled for plasmid copy number as all required genes were on one plasmid with the same pUC19-derived origin of replication (ORI).

Bioinformatic analysis annotates MmpT5 as a TetR regulatory protein, however its role in gene regulation has not yet been demonstrated. Here we develop two experimental setups with the MoClo Toolkit to test whether MmpT5 is a repressor of on the downstream \textit{mmpSL} promoter. The final constructs used for the one and two plasmid setups are shown in Figure 2.4 at level 2. Luminescence was measured with a plate reader at the visible light spectrum in an \textit{E. coli} DH5 alpha host under induced and uninduced conditions.

\textbf{2.3.4 MmpT5 acts a transcriptional repressor of the \textit{mmpSL} operon}

In the uninduced state, both the BDQ1-WT and BDQ2-WT \textit{mmpSL} promoters show activity in the \textit{E. coli} expression system (Figure 2.5B and C). Upon induction with IPTG, both BDQ1-WT and BDQ2-WT MmpT5 demonstrated reduced \textit{mmpSL} promoter activity, consistent with the hypothesis that MmpT5 is a direct repressor of the \textit{mmpSL} promoter. Pretreatment isolate \textit{mmpSL} expression appears to be higher in BDQ2-WT than BDQ1- WT (Figure 2.5B and C).

\textbf{2.3.5 Nonsynonymous mutations in \textit{mmpT5} eliminate \textit{mmpSL} promoter repression}

With the role of MmpT5 as a transcriptional repressor confirmed, we next tested the impact of replacing the wildtype \textit{mmpT5} with the mutant forms from relapse isolates.
Figure 2.5. Bioluminescence assay of 1 plasmid experimental setup.
Promoter activity of \textit{mmpSL} was monitored at 37°C with or without induction of MmpT5 using 1 mM IPTG. Luminescence was measured on a BioTek Synergy HT reader according to the manufacturer’s instructions. Promoter activity is expressed in arbitrary units and values are normalized to the number of cells. Data are represented by mean ± SD (n = 3). \textbf{A} illustrates the level 2 construct setup for the 1 plasmid experiment. \textbf{B} and \textbf{C} show the promoter activity of \textit{mmpSL::lux} when \textit{mmpT5} from pretreatment isolates is either induced or uninduced and \textbf{D} and \textbf{E} shows the promoter activity of \textit{mmpSL::lux} when \textit{mmpT5} from relapse isolates is either induced or uninduced.
Figure 2.6. Bioluminescence assay of 2 plasmid experimental setup.
Promoter activity of mmpSL was monitored at 37°C with or without induction of MmpT5 using 1 mM IPTG. Luminescence was measured on a BioTek Synergy HT reader according to the manufacturer’s instructions. Promoter activity is expressed in arbitrary units and values are normalized to the number of cells. Data are represented by mean ± SD (n = 3). A illustrates the level 1 and 2 construct setup for the 2 plasmid experiment. The mmpT5 ORF and mmpSL promoter were isolated from either BDQ1-WT or BDQ2-WT. B and C shows the promoter activity of mmpSL::lux when mmpT5 from pretreatment isolates is either induced or uninduced and D and E shows the promoter activity of mmpSL::lux when mmpT5 from relapse isolates is either induced or uninduced.
mmpSL promoter expression was largely unaffected when BDQ2-137G MmpT5 was induced (Figure 2.5E). It appears the nonsynonymous mmpT5 mutations in BDQ2-137G adversely affect MmpT5’s ability to repress the mmpSL promoter. In other words, we observe similar mmpSL promoter activities when the wildtype mmpT5 is either not expressed or is expressed in a mutant form. Surprisingly, when the wildtype form was replaced with mutant BDQ1-Fs form in the single plasmid setup, baseline transcription of mmpSL was lost. Figure 2.5D and E shows the bioluminescence assay of the 1 plasmid experimental setup, where inducible MmpT5 from BDQ1-Fs or BDQ2-137G was cloned upstream of their respective mmpSL::lux constructs. For the 1 plasmid construct, induction of MmpT5-137G resulted in the default expression of mmpSL, but not for MmpT5-Fs. The promoter activity of mmpSL from the BDQ1-Fs isolate was silent in both the induced and uninduced conditions. This result prompted the transition to the 2 plasmid setup, illustrated in Figure 2.4. The repressive effect of MmpT5 on mmpSL promoter activity was replicated to a stronger degree in the 2 plasmid setup (Figure 2.6B and C). Induction of either BDQ1-WT or BDQ2-WT MmpT5 showed almost complete repression of mmpSL::lux activity. Induction of BDQ1-Fs and BDQ2-137G MmpT5 in the 2 plasmid setup both resulted in no effect on mmpSL::lux activity (Figure 2.6D and E). The silent mmpSL promoter activity observed in the 1 plasmid setup with BDQ-Fs MmpT5 was not replicated in the 2 plasmid setup, suggesting a possible issue with the single plasmid multigene construct.

2.3.6 Troubleshooting earlier mmpT5 and mmpSL constructs

The results obtained above were not achieved without failure and multiple months of troubleshooting. The approach of studying Mycobacterium gene regulation using
Table 2.3. Unsuccessful *mmpT5* and *mmpSL* constructs made to test the proposed hypothesis.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Result and summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plac.mmpT5.T7</td>
<td>Inducible <em>mmpT5</em> in pICH47732</td>
<td>No effect on <em>mmpSL</em> promoter::<em>lux</em> activity. Starting codon of <em>mmpT5</em> mistakenly omitted in initial construct. Presumably incompatible with <em>mmpSL::lux</em> plasmid in 2 plasmid experiment.</td>
</tr>
<tr>
<td>PUmmpT5</td>
<td>The native <em>mmpT5</em> promoter, <em>mmpT5</em> ORF, and native <em>mmpSL</em> promoter fused to <em>lux</em> in pICH47732</td>
<td>No difference in <em>mmpSL</em> promoter activity observed between pretreatment and relapse isolates. Construct PUmmpT5.Lux.T7 revealed that the native <em>mmpT5</em> promoter is not active in <em>E. coli</em>.</td>
</tr>
<tr>
<td>PUmmpT5.Lux.T7</td>
<td>The native <em>mmpT5</em> promoter and <em>mmpT5</em> ORF in pICH47732</td>
<td>Did not exhibit any effect on <em>mmpSL</em> promoter::<em>lux</em> activity. Construct PUmmpT5.Lux.T7 revealed that the native <em>mmpT5</em> promoter is not active in <em>E. coli</em>.</td>
</tr>
<tr>
<td>PUmmpT5.Lux.T7</td>
<td>The native <em>mmpT5</em> promoter fused to <em>lux</em> in pICH47732</td>
<td>No <em>lux</em> activity observed, indicating the native <em>mmpT5</em> promoter is not active in <em>E. coli</em>.</td>
</tr>
<tr>
<td>1 Plasmid experiment</td>
<td>Inducible MmpT5 and <em>mmpSL::lux</em> in pAGM4673</td>
<td>No change in <em>mmpSL::lux</em> activity after <em>mmpT5</em> induced. First 3 bases of <em>mmpT5</em> mistakenly omitted. <em>mmpSL::lux</em> fusion was out of frame. Only one pretreatment isolate (BDQ2-WT) tested.</td>
</tr>
<tr>
<td>2 Plasmid experiment</td>
<td>Inducible MmpT5 in pICH47732 and <em>mmpSL::lux</em> in pICH89921</td>
<td>The plasmid ORIs were likely incompatible with each other. Only one pretreatment isolate (BDQ2-WT) tested.</td>
</tr>
</tbody>
</table>
synthetic biology tools in an *E. coli* host presented many challenges that had to be addressed. Multiple constructs were designed to test the *mmpT5-mmpSL* regulation hypothesis. Table 2.3 shows all of the multigene constructs made for this project and a brief summary and discussion of why each construct may not have worked. A key result is the absence of activity from the native *mmpT5* promoter in *E. coli*, suggesting that *mmpT5* wasn’t being expressed. The PUmmpT5.Lux.T7 construct showed no activity in *E. coli* and the PUmmpT5PUmmpSL.Lux.T7 constructs from both pretreatment and relapse isolates showed no change in *mmpSL* promoter activity within patient groups. For the first 2 plasmid experiment, the compatibility of both plasmids was not optimized due to lack of information. It is not known if the Ri A4 ORI and the F ORI are compatible with pUC19-derived ORIs. The vector backbones were then switched to known compatible ORIs (p15A and pUC19) for the final 2 plasmid experiment.

### 2.4 DISCUSSION

*Mycobacterium intracellulare* is a prominent nontuberculous mycobacterial (NTM) pathogen with relatively few treatment options available (17). Off-label use of Bedaquiline (BDQ) to treat *M. intracellulare* lung infections showed initial promise, however microbiological relapse from independently evolved mutations soon became an issue (1, 17). Nonsynonymous mutations in the TetR regulatory protein MmpT5 are predicted to play a role in elevated BDQ MIC levels, leading us to characterize the underlying genetic mechanisms. Comparable studies in *Mycobacterium tuberculosis* found that mutations in transcriptional regulator Rv0678 caused upregulation of the efflux pump MmpS5-MmpL5, leading to Bedaquiline and Clofazimine cross-resistance (12). More recently, a study found mutations in the *Mycobacterium abscessus* TetR
repressor \textit{MAB\_4384} resulted in upregulation of the adjacent MmpS5-MmpL5 efflux pump and subsequent resistance to thiacetazone analogues (11). The MmpS5-MmpL5 operon is found in all sequenced genomes of \textit{Mycobacterium} except for \textit{Mycobacterium leprae} (4). However, the adjacent regulator protein differs between species. \textit{M. tuberculosis} MarR-like regulator Rv0678 has no ortholog in \textit{M. intracellulare}, yet the \textit{M. abscessus} TetR regulator \textit{MAB\_43834} is orthologous to \textit{mmpT5} (2). TetR transcriptional regulators are the most abundant family of regulators in mycobacteria and commonly play a role in efflux pump regulation and other biochemical activities such as carbon metabolism, nitrogen metabolism, co-factor metabolism, and cell to cell signalling (3). However, most mycobacterial TetR regulators are uncharacterized in terms of their physiological function. Synthetic biology tools such as the MoClo toolkit (22) provide one way to quickly and efficiently test many different TetR regulators or TetR variants in relation to a promoter of interest. The MoClo toolkit allowed us to assemble inducible TetR regulator constructs of different variants and promoter::\textit{luxCDABE} fusions within three days.

In this study, we construct an inducible MmpT5 construct and \textit{mmpSL} promoter::\textit{lux} fusion to demonstrate the repressive effect of MmpT5 on the \textit{mmpSL} promoter. Induction of the wild type MmpT5 regulatory protein resulted in decreased \textit{mmpSL} promoter luminescence (Figure 2.5B and C, 2.6B and C), suggesting a repressive mechanism. Consistent with results in \textit{M. abscessus} and \textit{M. tuberculosis}, the wild type regulatory protein MmpT5 keeps efflux pump expression turned off in the basal state. Repression of efflux genes in the absence of antibiotics may avoid any fitness costs associated with efflux pump overexpression. Swapping in \textit{mmpT5} variants from relapse
Figure 2.7. Proposed mechanism of BDQ resistance.

Native conformation of *M. intracellulare* mmpT5-mmpSL operon. Nonsynonymous *mmpT5* mutations from relapse isolates reduce or eliminate the binding affinity of MmpT5 to the *mmpSL* promoter. The *mmpSL* RND efflux pump is then upregulated, increasing the BDQ MIC for *M. intracellulare*. 
isolates eliminated the negative regulation of MmpT5, resulting in upregulation of the efflux pump (Figure 2.5D and E, Figure 2.6D and E). Both the DNA binding domain mutation and the frameshift mutation eradicated MmpT5’s ability to bind to the mmpSL promoter. This increased expression of mmpSL is predicted to increase resistance to Bedaquiline via drug efflux as shown in Figure 2.7. The mmpT5-mmpSL clones in pMD31 will be used to test the BDQ resistance phenotype in a mycobacterium system. Future experiments will look into the crystal structure of MmpT5 and how it binds to the intergenic region between mmpT5 and mmpS. Electrophoretic mobility shift assay experiments in M. tuberculosis and M. abscessus indicated both Rv0678 and MAB_4384 bind the intergenic region upstream of mmpS5 to negatively regulate efflux pump expression (18, 20). X-ray crystallography of MAB_4384 revealed a ligand binding domain, which is targeted by TAC derivatives to induce a conformational change and sequester the DNA binding domain (DBD) (20). Ligand-binding is also observed in Rv0678, where 2-stearoylglycerol binds and induces a conformational change to prevent DNA-binding (18).

A ligand binding domain in MmpT5 has not yet been described, however it is likely because TetR regulators generally have a ligand binding site in the C-terminal domain (6). Depending on the pathway that TetR regulates, the ligand binding domain evolves to fit the molecule of choice. Therefore, MmpT5 may bind certain antibiotics such as BDQ when present, changing its conformation and releasing from the mmpSL promoter region. This would upregulate transcription of the efflux pump genes and facilitate drug efflux and antibiotic resistance. In my experimental model however,
mutations in \textit{mmpT5} are proposed to be responsible for BDQ-resistance, rather than a protein-BDQ sensing response.

Important insights were obtained in the development of this novel approach for studying \textit{M. intracellulare} gene regulation. The PUmmpT5.lux construct described in Table 2.3 showed that the native \textit{mmpT5} promoter was not active in \textit{E. coli}. This result explains why no difference in \textit{mmpSL} promoter activity was observed in the PUmmpT5PUmmpSL.lux constructs. Since \textit{mmpT5} was not being expressed in these constructs, no change in \textit{mmpSL} promoter activity would be expected. The lack of \textit{mmpT5} expression in \textit{E. coli} could be due to multiple factors. We included 161 bp upstream of the \textit{mmpT5} ORF in the “native \textit{mmpT5} promoter” constructs. Due to the architecture of mycobacterial promoters (15), it is possible that the full \textit{mmpT5} promoter region was not included in the construct. Another explanation could be that \textit{E. coli} lacks the required sigma factors to express this mycobacterial promoter (5, 10). Several studies have shown that many mycobacterial promoters function poorly in \textit{E. coli} (7, 13). The expression of the \textit{mmpSL} promoter in \textit{E. coli} was an exciting finding and shows promise for studying non-\textit{E. coli} genes in \textit{E. coli}. However, for this system to work, it is important to make sure the promoters of interest are active and the plasmid ORIs are compatible with each other when used in conjunction. Future studies should address the challenges described in this project before designing primers for Golden Gate assembly. Once optimized, the MoClo toolkit is an extraordinarily powerful tool for studying prokaryotic gene regulation.
2.5 Acknowledgements

I would like to express my appreciation to Dr. David Alexander for his research laying the foundation for this project and his excellent advice on my lux reporter experiments. Thank you also to Dr. Jeffrey Chen for providing the mycobacterial shuttle vector pMD31 and conducting the bedaquiline assays in Mycobacterium smegmatis. This work was funded by a Natural Sciences and Engineering Research Council of Canada Discovery grant and a Saskatchewan Health Research Foundation Establishment grant (ADSC).

REFERENCES


Table S2.1. List of oligonucleotide sequences used for cloning.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5′-3′)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mi.mmpT5.P1.F</td>
<td>TTTTGAAGACAAAAATGATGGCCAATCCCCTGGGGCT</td>
<td>This study</td>
</tr>
<tr>
<td>Mi.mmpT5.P1.R</td>
<td>TTTTGAAGACAAAAATTCGCGAGTTCGAGCATGTGAG</td>
<td></td>
</tr>
<tr>
<td>Mi.mmpT5.P2.F</td>
<td>TTTTGAAGACAAAAAACTTCCGCCGCGTGCTCC</td>
<td></td>
</tr>
<tr>
<td>Mi.mmpT5.P2.R</td>
<td>TTTTGAAGACAAAAAGCTCCTTCCGAGATGCTGT</td>
<td></td>
</tr>
<tr>
<td>mmpT5.Native. P1.F</td>
<td>CACTCTGTGTCCTCAGGAGTGACCGGCCCGAACCACCT</td>
<td>This study</td>
</tr>
<tr>
<td>Mi.mmpSL.lux.F</td>
<td>CACTCTGTGTCCTCAGGAGCCTGAAACACCCAGAGCAAA</td>
<td>This study</td>
</tr>
<tr>
<td>Mi.mmpSL.lux.R</td>
<td>CACTTCGTCGTCCTCACATTTCGAGTCGGGGCTGTCAT</td>
<td></td>
</tr>
<tr>
<td>BamHI. PUmmpT5.F</td>
<td>TTTTACCTCCTAGCGCGCCGAAACCCCT</td>
<td>This study</td>
</tr>
<tr>
<td>BamHI. FullmmpSL.R</td>
<td>TTTTACCTCCTAGCGCGCCGAAACCCCT</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 3: Regulation of *comA* gene expression by Sxy and an mRNA secondary structure in *Salmonella Typhimurium*
ABSTRACT

ComA is a periplasmic protein proposed to be essential for the active take up of DNA in Salmonella enterica. Previous work has shown transcription of comA to be undetectable in all experimental conditions in Salmonella Typhimurium, even when the predicted activator Sxy is overexpressed. Identification of a predicted mRNA secondary structure in the comA promoter region provides a possible mechanism for gene repression. An mRNA stem-loop in the comA transcript could either terminate transcription or inhibit translation by sequestering the ribosome binding site. Destabilization or removal of the predicted comA mRNA secondary structure would increase gene expression. Thus, we used synthetic biology tools to assemble comA promoter-luciferase fusions with or without the predicted mRNA secondary structure. To remove the potential dependence on Sxy and CRP, a second set of constructs were assembled where the constitutive lacZ promoter was fused to the comA::lux reporter. Luciferase-reporter assays showed that positioning the lacZ promoter to replace the predicted mRNA stem-loop resulted in a 6-fold increase in luminescence activity. Site-directed mutagenesis to destabilize the predicted comA secondary structure increased luminescence in the lacZ::comA::lux fusions. Compensatory mutations to restore the predicted mRNA structure reduced promoter activity back to wild type levels. Site-directed mutagenesis of the native promoter in a comA::lux fusion did not increase luminescence, suggesting a continued requirement for Sxy and CRP for gene induction. To test this, we assembled a multigene construct containing comA::lux SDM variants and inducible sxy. Induction of sxy resulted in a 197-fold increase in luminescence compared to the wild type when the predicted stem-loop was destabilized.
3.1 INTRODUCTION

Natural competence, the ability of bacteria to actively take up DNA from the environment, serves two possible functions: the uptake of DNA fragments for nutrients, or the acquisition of genetic material for chromosomal recombination (6, 20). DNA uptake requires the coordinated expression of multiple genes to build the apparatuses that bind extracellular DNA and transport it across the cell envelope (11). In most bacteria, natural competence is tightly regulated by transcription factors and cis-regulatory mechanisms, presumably to conserve resources associated with generating the specialized DNA binding and transport apparatuses. Sigma factors, nucleotide pools, mRNA secondary structure, and regulatory proteins all play a role in the regulation of competence gene expression (2, 4, 17). Analyses of the molecular mechanisms that regulate the expression of competence genes provides insight into the cellular and ecological functions of natural competence in bacteria (13).

Among Gram-negative bacteria, Haemophilus influenzae, Helicobacter pylori, Neisseria gonorrhoeae, and Vibrio cholerae are models for the study of natural competence and transformation (15). Currently, there is only indirect evidence of natural competence in the model organism Escherichia coli, and no evidence of natural competence in Salmonella enterica (18). Yet, all Enterobacteriaceae, including E. coli and S. enterica, encode homologs of all the genes required for natural competence in H. influenzae (6). The master regulators of competence, the cAMP Receptor Protein (CRP) and Sxy are also conserved in all Enterobacteriaceae. Despite decades of gene expression studies in E. coli and S. enterica, competence gene homologs remain cryptic.
Non-canonical CRP sites dubbed “CRP-S sites” require both Sxy and CRP for gene induction and have been found in promoter regions of γ-proteobacteria competence regulons (2). In S. Typhimurium, CRP-S sites have been predicted in the promoter regions of comA, comN, hopD, and pilA (2). Recent RT-qPCR analysis conducted by a lab colleague found that overexpression of Salmonella sxy increased gene expression of S. Typhimurium competence genes comN, hopD, and pilA (1); these genes encode components of the DNA binding and uptake apparatus. Curiously, overexpression of Sxy did not induce comA expression. An mRNA secondary structure was previously predicted at the 5’ end of the comA mRNA, which could account for the absence of transcriptional activation (A. Cameron, personal communication). Mfold analysis predicted the secondary structure to be strong, extensive, and to sequester the ribosome binding site in a stem-loop (Figure 3.1). A similar repressive mechanism has been observed in H. influenzae where the translation of the sxy gene is greatly increased through point mutations that weaken the folding at the 5’ end of the sxy mRNA and hypercompetence is induced (3). Inhibiting sxy transcription and translation with a mRNA stem-loop is one possible mechanism to limit DNA uptake under non-inducing conditions and conserve resources. Thus, I set out to test whether directed mutations that weaken the predicted secondary structure in comA followed by compensatory mutations that restore the predicted secondary structure could modulate the expression of this competence gene. Importantly, finding conditions that allow for expression of all competence genes will contribute to characterizing when and how Enterobacteriaceae take up DNA from their environment.

Synthetic biology tools such Golden Gate Cloning have the ability to accelerate
Figure 3.1. Predicted \textit{comA} mRNA secondary structure and location of site-directed mutations.

The mRNA secondary structure predicted by Mfold in the \textit{comA} promoter region. The start codon and ribosome binding site are highlighted in green, and the position of the site-directed mutations are indicated by the red and blue arrows. The red arrows (SDM 1) indicate mutations introduced to destabilize the predicted secondary structure and the blue arrows (SDM 2) are mutations used to restore the base pairs in the SDM 1 mutants. The start of the mRNA transcript is position 1.
the study of prokaryotic gene regulation. Golden Gate cloning is a modular cloning system that allows for the efficient shuffling and assembly of different gene constructs (5). Type IIS enzymes and T4 DNA ligase are used to directionally assemble DNA fragments in a single reaction. Through the Golden Gate cloning method, different bacterial promoters can be efficiently fused to luxCDABE bioreporters, allowing us to directly quantify promoter activity. In addition, inducible promoters such as the Lac promoter can be cloned into our genetic constructs, providing the ability to control when certain genes are turned off and on. In this study, we adapt the MoClo toolkit (22) to generate multiple multigene constructs to study S. Typhimurium gene regulation. The optimization of the Golden Gate cloning technique allowed us to quickly and efficiently test several different promoter variations in conjunction with an inducible sxy construct. The comA promoter and its respective CRP-S sites were fused to a luciferase reporter gene and cloned with IPTG-inducible sxy. Including lacIq in the assembled vector ensured tightly-controlled expression of sxy in S. enterica cells.

Site-directed mutagenesis (SDM) experiments were conducted to test for the presence and regulatory effects of the predicted secondary structure in the comA::lux fusions. Destabilization of the comA mRNA secondary structure via specific point mutations increased expression.

3.2 MATERIALS AND METHODS

3.2.1 Bacterial Strains and Plasmids

The bacterial strains and plasmids used in this study are listed in Table 3.1. The E. coli and S. Typhimurium strains were grown in Luria-Bertani (LB) broth at 37°C with shaking. Spectinomycin (50 µg/mL for E. coli), Ampicillin (100 µg/mL for E. coli and
Table 3.1. Bacterial plasmids and strains used in this study.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bacterial strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Escherichia coli</em> DH5-Alpha</td>
<td>High-efficiency competent cells used for Golden Gate cloning</td>
<td>(19)</td>
</tr>
<tr>
<td><em>Salmonella enterica</em> serovar Typhimurium ATCC 14028</td>
<td>Common S. Typhimurium laboratory strain used to study molecular mechanisms of <em>Salmonella</em></td>
<td>(9)</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pICH41295</td>
<td>level 0 landing vector for PU module</td>
<td>(22)</td>
</tr>
<tr>
<td>pICH41308</td>
<td>level 0 landing vector for SC module</td>
<td></td>
</tr>
<tr>
<td>pICH47732</td>
<td>level 1 landing vector (position 1*)</td>
<td></td>
</tr>
<tr>
<td>pICH47742</td>
<td>level 1 landing vector (position 2*)</td>
<td></td>
</tr>
<tr>
<td>pICH47751</td>
<td>level 1 landing vector (position 3*)</td>
<td></td>
</tr>
<tr>
<td>pAGM4673</td>
<td>level 2 landing vector (pUC19 ORI)</td>
<td></td>
</tr>
<tr>
<td>pICH41766</td>
<td>level 2 end-linker</td>
<td></td>
</tr>
<tr>
<td>pOGG031</td>
<td>LacZ promoter cloned into level 0 vector</td>
<td>This study</td>
</tr>
<tr>
<td>pOGG039</td>
<td>T7 terminator cloned into level 0 vector</td>
<td></td>
</tr>
<tr>
<td>Level 0 Lux</td>
<td><em>luxCDABE</em> (5846 bp) inserted into pICH41308, level 0 vector for SC modules</td>
<td>This study</td>
</tr>
<tr>
<td>Level 1 Lac</td>
<td><em>lacI</em> promoter and Lacl gene inserted into pICH47751</td>
<td></td>
</tr>
</tbody>
</table>

*position refers to orientation of insert when cloned into level 2 landing vector
3.2.2 PCR amplification of *sxy* and *comA* promoters

Chromosomal DNA was extracted from *Salmonella enterica* subsp. *enterica* serovar Typhimurium ATCC 14028 and used as a template for PCR amplification of the *sxy* gene and the promoter regions of *comA*. PCR amplification was performed using Q5® High-Fidelity DNA Polymerase (New England Biolabs), 1X Q5 reaction buffer, 10 mM dNTPs, 10 µM of each primer set shown in Table S3.1, and ∼25 ng chromosomal DNA. Reaction mixtures were scaled to 50 µL and incubated for 30 s at 98°C (Initial denaturation), followed by 35 cycles of 30 s at 98°C (denaturation), 30 s at 63°C (annealing), and 1 min at 72°C (extension), with a final elongation step at 72°C for 2 minutes. The PCR products were then isolated on agarose gels and purified using the QIAquick Gel Extraction Kit (Qiagen). The purified DNA was eluted in 30 µl of Elution Buffer and quantified by Nano Drop® Spectrophotometer ND-1000 (Thermo Scientific™).

3.2.3 Multigene construct design and assembly

Synthetic multigene constructs were designed to study the role of Sxy on
Table 3.2. Final level 1 *comA* constructs, level 2 inducible *sxy* constructs, and site-directed mutagenesis (SDM) targets.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Experimental role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Far-PcomA</td>
<td>Contains two CRP-S sites and full mRNA secondary structure fused to <em>luxCDABE</em>.</td>
<td>To assess the native <em>comA</em> promoter activity. (2 CRP-S sites).</td>
</tr>
<tr>
<td>Near-PcomA</td>
<td>Contains one CRP-S sites and full mRNA secondary structure fused to <em>luxCDABE</em>.</td>
<td>To assess the native <em>comA</em> promoter activity. (1 CRP-S site).</td>
</tr>
<tr>
<td>PcomA-TSS</td>
<td>LacZ promoter fused to TSS of <em>comA</em> and contains full mRNA secondary structure fused to <em>luxCDABE</em>.</td>
<td>Positive control to remove potential regulatory protein dependence for <em>comA</em> gene induction and assess the predicted mRNA stem-loop.</td>
</tr>
<tr>
<td>PcomA-RBS</td>
<td>LacZ promoter fused 5bp upstream of the <em>comA</em> RBS, containing only part of the mRNA secondary structure fused to <em>luxCDABE</em>.</td>
<td>Removes potential regulatory protein dependence for <em>comA</em> gene induction and removes the predicted mRNA stem-loop.</td>
</tr>
<tr>
<td>Far-PcomA + sxy</td>
<td><em>comA</em> promoter region with two CRP-S sites fused to <em>luxCDABE</em>. Inducible <em>sxy</em> construct. Constitutive <em>lacI</em> construct.</td>
<td>To deduce if Sxy, the predicted <em>comA</em> mRNA secondary structure, or both are responsible for <em>comA</em> gene expression</td>
</tr>
<tr>
<td>[3.5.9] SDM 1</td>
<td>SDM changing base positions 3, 5, 9 of the <em>comA</em> mRNA transcript to Adenine.</td>
<td>Destabilization of the predicted <em>comA</em> mRNA stem-loop.</td>
</tr>
<tr>
<td>[3.5.9] SDM 2</td>
<td>SDM at base positions 55, 60, 62 of the <em>comA</em> mRNA transcript to Uracil.</td>
<td>Restoration of the predicted <em>comA</em> mRNA stem-loop.</td>
</tr>
<tr>
<td>[5 bp] SDM 2</td>
<td>SDM at base positions 34, 36, 55, 60, 62 of the <em>comA</em> mRNA transcript to Uracil.</td>
<td>Restoration of the predicted <em>comA</em> mRNA stem-loop.</td>
</tr>
</tbody>
</table>

Note: TSS: transcription start site, RBS: ribosome binding site, SDM: site-directed mutagenesis
competence gene expression and the effect of the predicted mRNA secondary structure on \textit{comA} expression. Final constructs and descriptions are described in Table 3.2.

Golden gate reactions were set up following the MoClo cloning protocol (22). For level 0 \textit{comA} cloning, reactions were prepared by pipetting in one tube 40 fmol of the gel-purified \textit{comA} PCR product, 40 fmol of the level 0 landing vector plasmid \textit{pICH41295}, 10U of type IIS restriction enzyme BbsI (New England Biolabs), 20U of high concentration T4 Ligase (New England Biolabs), 1.5 \(\mu\)L of 2 mg/mL Bovine Serum Albumin (BSA), 1.5 \(\mu\)L of 10X T4 DNA Ligase Buffer (New England Biolabs), and Milli-Q® H\(_2\)O (Millipore) up to a total volume of 15 \(\mu\)L. This procedure was repeated three times using a different gel-purified \textit{comA} PCR product for each reaction (4 total). The reactions were incubated in a thermocycler for 25 cycles at 37°C for 3 min and 16°C for 4 min followed by 5 min at 50°C and 5 min at 80°C. The restriction-ligation mixtures were added to 100 \(\mu\)L chemically competent DH5\(\alpha\) cells, incubated for 30 min on ice and transformed by heat shock. Then 900 \(\mu\)L of SOC media was then added to the transformation, and the cells were allowed to recover for 1 hour at 37°C. Next 200 \(\mu\)L of the transformation were plated on LB plates containing spectinomycin (50 \(\mu\)g/mL) and X-gal/IPTG and incubated overnight at 37°C. Four white recombinant colonies were then selected from the plates and streaked on separate spectinomycin plates. Each of the four isolated recombinant colonies were inoculated in 5 mL LB broth and grown overnight. Plasmids were isolated using the EZ-10 Spin Column Plasmid DNA Miniprep Kit (Biobasic). The level 0 \textit{comA} constructs were confirmed by restriction digest with BsaI and PCR amplification of the \textit{comA} promoter region using the level 0 clone as template.
For level 1 comA cloning, reactions were prepared by pipetting in one tube 40 fmol of the confirmed level 0 comA plasmid, 40 fmol of a level 0 luxCDABE bioreporter plasmid, 40 fmol of a level 0 T7 terminator plasmid, 40 fmol of pICH47732, 10U of type IIS restriction enzyme BsaI (New England Biolabs), 20U of high concentration T4 Ligase (New England Biolabs), 1.5 µL of 2 mg/mL BSA, 1.5 µL of 10X T4 DNA Ligase Buffer (New England Biolabs), and Milli-Q® H₂O (Millipore) up to a total volume of 15 µL. The reaction was incubated in a thermocycler at the same settings for level 0 cloning previously mentioned. The restriction-ligation mixtures were transformed into chemically competent DH5α as described above. Transformations were plated on ampicillin plates (100 µg/mL) with X-gal/IPTG and incubated overnight at 37°C. Recombinant colonies were restreaked and the level 1 plasmids were isolated as described above. The level 1 comA constructs were confirmed by restriction digest with BbsI and PCR amplification of the comA promoter region using the level 1 clone as template. The inducible level 1 Sxy construct was made using the above reagents and thermocycler settings with the following parts: 40 fmol of pOGG031, 40 fmol of gel-purified full sxy ORF, 40 fmol of a level 0 T7 terminator plasmid, and 40 fmol of pICH47732.

The level 2 competence gene and Sxy constructs reactions were prepared by with 40 fmol of the level 1 competence promoter::lux plasmid, 40 fmol of the level 1 inducible sxy plasmid, 40 fmol of the level 1 LacI plasmid, 40 fmol of the endlinker pICH41766, 40 fmol of pAGM4673, 10U of type IIS restriction enzyme BbsI (New England Biolabs), 20U of high concentration T4 Ligase (New England Biolabs), 1.5 µL of 2 mg/mL Bovine Serum Albumin (BSA), 1.5 µL of 10X T4 DNA Ligase Buffer (New England Biolabs), and Milli-Q® H₂O (Millipore) up to a total volume of 15 µL. The
reaction was incubated in a thermocycler at the same settings previously described for level 0 cloning. The restriction-ligation mixtures were transformed into chemically competent DH5α as described above. Transformations were plated on Kanamycin plates (50 µg/mL) incubated overnight at 37°C. Final multigene constructs were transformed into Se14028 and confirmed with PCR and Sanger sequencing.

3.2.4 Quantifying competence gene promoter activity

Measurement of competence promoter activity was accomplished by bioluminescence and OD_{600} readings using a Synergy HT Multi-Detection Microplate Reader (Bio-Tek). Se14028 WT cells with level 1 and 2 constructs were inoculated in 5 mL LB broth (Lennox) and incubated at 37°C overnight with shaking at 200 rpm. Overnight cultures were diluted 1:1000 (v/v) in LB broth (Lennox) with Kanamycin (50 µg/mL) or Ampicillin (100 µg/mL) and IPTG (1 mM) to induce expression of the lacZ promoter. The diluted cells containing the constructs were plated in triplicates in a 96-well plate at 250 µL per well with 50 µL mineral oil overlaid on top. Bioluminescence was measured for 48 hours at 10 min intervals. For bioluminescence measurements, Synergy HT Multi-Detection Microplate Reader (Bio-Tek) sensitivity gain setting was 100, with no excitation and no emission filter in place. Data was collected using Gen 5 2.07 (BioTek Instruments), exported to Microsoft Excel, and plotted using Graphpad Prism 7. Bioluminescence readings were normalized to cell density by dividing by the corresponding OD_{600} values.

3.2.5 Prediction of comA mRNA secondary structure

Mfold was used to predict the mRNA secondary structure with the lowest ΔG value located in the comA promoter and coding region. Specific point mutations in the
predicted secondary structure were then analyzed by Mfold to find the greatest $\Delta G$ value, which indicates the most destabilized $comA$ mRNA secondary structure.

3.2.6 $comA$ site-directed mutagenesis

The Q5 Site-Directed Mutagenesis kit (New England Biolabs) was used for all site-directed mutagenesis (SDM) modifications to produce level 1 $comA$ plasmids with the predicted mRNA secondary structure destabilized. An additional round of SDM was done to restore the predicted mRNA secondary structure. Mutagenic primers were designed using the NEBaseChanger tool (New England Biolabs) and primer sequences are presented in Supplementary Table 3.1.

3.3 RESULTS

3.3.1 Impairing mRNA secondary structure increases $comA$ expression

The MoClo toolkit was used to assemble the multigene constructs described in Table 3.2. Figure 3.2 illustrates the layout of each construct used to assess if the predicted mRNA secondary structure in the $comA$ promoter region plays any role in impeding gene expression (Figure 3.2A) and if Sxy regulates the $comA$ promoter (Figure 3.2B). The Near-PcomA and Far-PcomA constructs, with one and two predicted CRP-S sites respectively, contained the native $comA$ promoter fused to the $lux$ reporter. In Se14028, the Far-PcomA and Near-PcomA constructs were transcriptionally silent (Figure 3.3). To remove the potential dependence on Sxy and CRP, and serve as a positive control to see if $comA$ can be induced at all, a second set of constructs were assembled where the $lacZ$ promoter was fused to the $comA::lux$ reporter. The $lacZ$ promoter is constitutively expressed in Se14028, thus removing any requirement of CRP, Sxy, or any other transcription factor for $comA$ induction. The PcomA-TSS construct contains the predicted
Figure 3.2. Final synthetic multigene constructs used to study sxy regulation and the predicted comA secondary structure.

A shows the four experimental constructs used to study the comA secondary structure. Each construct contains either the inducible lac promoter, the predicted mRNA secondary structure, or both, fused to the luxCDABE bioreporter. B shows the layout of the constructs used to study Sxy regulation. Sxy is controlled by the inducible LacZ promoter. Upstream is the comA promoter with its respective CRP-S site fused to the luxCDABE bioreporter. Downstream of the inducible sxy construct is the LacI construct, which keeps the LacZ promoter repressed until induction with IPTG.
Figure 3.3. Bioluminescence assay of level 1 comA constructs.
Promoter activity of level 1 comA constructs in 14028 WT was monitored at 37°C and measured in arbitrary units and values which are normalized to the number of cells. Luminescence was measured on a BioTek Synergy HT reader according to the manufacturer's instructions. Data are represented by mean ± SD (n = 3).
secondary structure whereas the PcomA-RBS construct does not, thus comparing the promoter activity of these two fusions indirectly tests if the secondary structure exists. As illustrated in Figure 3.3, the PcomA-RBS construct showed a 6-fold increase in promoter activity at t = 8 hours compared to PcomA-TSS, consistent with the hypothesis that the mRNA stem-loop inhibits gene expression.

Mfold modelled the most stable mRNA secondary structure in the comA promoter region (Figure 3.1). From this model, we selected base pairs that would maximally weaken the stem-loop structure. GC pairs contain three hydrogen bonds, thus were selected for mutagenesis over the weaker two hydrogen bonds of AU. All of the bases were changed to Adenine to disrupt the predicted stem-loop. SDM was designed to simultaneously mutate multiple bases; mutating base positions 3, 5 and 9 (called “3.5.9”) was predicted to destabilize stems Ia and Ib, while mutating positions 3, 5, 9, 22, and 24 (called “5 bp”) was predicted to destabilize stems Ia, Ib, and II (Figure 3.1).

SDM was used in the Far-PcomA and PcomA-TSS constructs described in Table 3.2. A second round of SDM called “SDM 2” was used in the PcomA-TSS mutants to create compensatory base pairing with the SDM 1 mutations to restore the predicted mRNA secondary structure. This is a strong genetic test of the predicted secondary structure because it confirms whether it is the ability to form base pairs in mRNA, not the coding of information in DNA or RNA sequence, that controls transcription. By measuring comA promoter activity, we can indirectly assess potential destabilization of the mRNA stem loop.

Bioluminescence assays showed an increase in comA promoter activity in both PcomA-TSS SDM constructs (Figure 3.4B). These results ultimately support the
Figure 3.4. Bioluminescence assay of Level 1 PcomA-TSS site-directed mutagenesis constructs.

Promoter activity of level 1 comA site-directed mutagenesis (SDM) constructs in 14028 WT was monitored at 37°C and measured in arbitrary units and values which are normalized to the number of cells. Luminescence was measured on a BioTek Synergy HT reader according to the manufacturer’s instructions. Data are represented by mean ± SD (n = 3). (A) PcomA-TSS [3.5.9] SDM represents the mutation of base positions 3, 5, and 9 of the comA mRNA transcript. (B) PcomA-TSS [5 bp] SDM represents mutation of base positions 3, 5, 9, 22, 24 of the comA mRNA transcript. Second round SDM is designated as SDM 2.
stem-loop model illustrated in Figure 3.1 and are key in understanding the nature of the predicted mRNA secondary structure. Surprisingly, PcomA-TSS [3.5.9] SDM conferred a 2-fold higher lux expression level than PcomA-TSS [5 bp] (Figure 3.4A). The higher luminescence in the PcomA-TSS [3.5.9] SDM construct compared to the PcomA-TSS [5 bp] SDM construct suggests that a stem-loop may still be impeding expression after the [5 bp] SDM 1 mutation. Destabilizing the predicted stem-loop still allows the formation of an alternate secondary structure, which could explain why mutations at positions 22 and 24 in stem II result in a slight repression in the presence of mutation [3.5.9] SDM 1.

The SDM 2 mutants, which replace the original GC base pairs with AU pairs, exhibited a return to wild type comA::lux activity, providing strong evidence that the predicted stem-loop described in Figure 3.1 was restored. In both the PcomA-TSS [3.5.9] and PcomA-TSS [5 bp] constructs, SDM 2 decreased luminescence to the WT level.

3.3.3 Sxy-dependent comA expression is increased through hypercompetence mutations

The prediction of two CRP-S sites in the comA promoter region suggest that Sxy and CRP are required for gene induction. To test this hypothesis, the level 1 Far-PcomA construct illustrated in Figure 3.2A was cloned with an inducible sxy construct into a level 2 multigene construct, shown in Figure 3.2B. Induction of sxy would show if comA expression is dependent on Sxy or not. Additionally, the Far-PcomA [3.5.9], [5 bp] SDM 1 constructs and [5 bp] SDM 2 construct were cloned into the inducible sxy level 2 multigene construct. This experimental setup would help deduce if Sxy, the predicted comA mRNA secondary structure, or both are responsible for comA gene expression. Given the previous results, we predict that Sxy will only induce comA expression in
Figure 3.5. Bioluminescence assay of Level 2 comA SDM/inducible Sxy constructs.

Promoter activity of level 2 comA site-directed mutagenesis (SDM)/inducible Sxy constructs in 14028 WT was monitored at 37°C with (A) or without (B) induction of Sxy using 1 mM IPTG. Luminescence was measured on a BioTek Synergy HT reader according to the manufacturer’s instructions. Promoter activity is expressed in arbitrary units and values are normalized to the number of cells. Data are represented by mean ± SD (n = 3). [3.5.9] SDM represents the mutation of base positions 3, 5, and 9 of the comA mRNA transcript. [5 bp] SDM represents mutation of base positions 3, 5, 9, 22, 24 of the comA mRNA transcript. Second round SDM is designated as SDM 2.
clones with a destabilized mRNA structure.

Bioluminescence assay results showed the induction of sxy to have no regulatory effect on the comA::lux WT (Figure 3.5B). This is expected given the predicted mRNA secondary structure, which would negatively regulate expression of comA. The level 2 comA construct with the predicted destabilized mRNA stem-loop showed remarkable results upon Sxy induction, with up to a 197-fold increase in luminescence at t = 8 compared to the wild type comA. The [3.5.9] SDM 1 of comA showed the highest increase in expression, consistent with the PcomA-TSS [3.5.9] SDM results in Figure 3.4A. Curiously, the comA 5 bp SDM 1 showed no increase in expression upon induction of sxy, but the 5 bp SDM 2 exhibited an 18-fold increase in luminescence at t = 8.

3.4 DISCUSSION

In this study, we used synthetic multigene constructs to show positive regulation of comA with Sxy and increased comA expression through targeted destabilization or removal of the mRNA secondary structure. Induction of the Sxy regulator protein increased promoter activity of comA in Se14028 WT. This result is conducive with the identification of a CRP-S site in the comA promoter region. Mutations that destabilized the comA 5′ mRNA stem-loop increased luminescence, indicating a negative regulatory mechanism in S. Typhimurium. Previous work has identified a 5′ mRNA stem-loop in the sxy transcript of S. Typhimurium, similar to Haemophilus influenzae (3). Therefore, Sxy may be expressed at a low level, contributing to the silent promoter activity of comA. Based on these results, we identify two main factors that may contribute to S. Typhimurium’s lack of natural competence: the occurrence of 5′ mRNA secondary
structures in the promoter region of *comA* and *sxy*, and the overexpression of Sxy required to induce *comA* expression, which is part of the DNA-uptake apparatus.

The increased luminescence activity of PcomA-TSS mutants with a destabilized mRNA secondary structure supports the negative regulation model of mRNA stem-loops. Formation of a stem-loop in the nascent mRNA transcript can pause RNAP and abort transcription, or block ribosome access to the ribosome binding site (20). As illustrated by Mfold in Figure 3.1, the *comA* mRNA secondary structure sequesters the RBS, which is predicted to decrease translation efficiency of *comA*. A similar phenomenon has been shown in *H. influenzae*, where the 5′ end of the *sxy* transcript forms a secondary structure that negatively regulates translation of *sxy* (3). Compensatory site-directed mutagenesis (SDM 2) to restore the PcomA-TSS mRNA secondary structure reduced luminescence (Figure 3.4), providing further evidence for a stem-loop regulatory mechanism. SDM of the Far-PcomA construct did not result in increased luminescence (Fig. 3.5A) presumably because Sxy is required for RNAP recruitment. This was confirmed when Sxy induction increased luminescence of Far-PcomA mutants containing destabilizing mutations. Removing mRNA base pairs at positions 3, 5, and 9 resulted in the highest *comA::lux* activity upon Sxy induction, which corresponds to the PcomA-TSS SDM results shown in Figure 3.4. Curiously, Sxy induction of Far-PcomA [5 bp] SDM 2 resulted in higher luminescence than Far-PcomA [5 bp] SDM 1. This result contradicts what we saw with the PcomA-TSS SDM experiments, suggesting another factor at play in the native *comA* promoter. One possibility is the formation of an mRNA tertiary structure that is stabilized by other mutations. Other factors such as the availability of purine nucleotides may contribute to regulating mRNA secondary structure, as observed in *H. influenzae* (17).
Sxy overexpression has been shown to induce hypercompetence in *H. influenzae* (23). In *Enterobacteriaceae* such as *E. coli*, no natural competence was observed upon induction of *sxy* on a plasmid, however a T4P was produced (16). Correspondingly, *S. Typhimurium* showed no sign of natural competence when *sxy* was expressed (1). ComA is required for the extension of the T4P into the extracellular space (12), and therefore may be the missing factor in *Enterobacteriaceae* natural competence. Our results show that impairing the mRNA secondary structure in the *comA* promoter region allows *comA* expression, but that this expression remains Sxy-dependent. Combining the over-expression of *sxy* from a plasmid with chromosomical mutations that ablate the *comA* promoter secondary structure may finally provide the development of natural competence in *S. Typhimurium* in the laboratory.

Future directions include using CRISPR/Cas9 technology to make gene edits to the *S. Typhimurium* chromosome in hopes of inducing natural competence. We are currently using a synthetic biology technique developed by Reisch et al. (14) that combines lambda red and CRISPR/Cas9 to genome edit the *E. coli* chromosome. Two plasmids containing inducible Cas9, an inducible sgRNA, and an inducible λ-Red phage are introduced into the bacteria of interest. The lambda red phage recombination system uses an oligonucleotide carrying the desired mutations to replace the native *comA* sequence in the chromosome. The CRISPR/Cas9 system is programmed to recognize the native (non-mutated) *comA* region and introduce a double-strand DNA break, which kills cells that retain the wild type *comA* sequence. Based on the results of this study, we can accurately pinpoint the bases to change in the chromosomal *comA* promoter region. After successful genome modification, RT-qPCR and protein analysis can be used to detect if
comA is being successfully expressed. Then a natural competence phenotype would be tested in S. Typhimurium to see if DNA-uptake has been induced.

3.5 Acknowledgements

I would like to express my appreciation to Dr. Ebtihal Alshabib for her excellent work in laying the foundation for this project and Danae Suchan for helping with the site-directed mutagenesis experiments. This work was funded by a Natural Sciences and Engineering Research Council of Canada Discovery grant and a Saskatchewan Health Research Foundation Establishment grant (ADSC).

REFERENCES


<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5' -3')</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>FarcomA.F</td>
<td>TTTTGAAGACAAGGAGGCTATCTTCGGTGGCGATAAA</td>
<td>This study</td>
</tr>
<tr>
<td>NearcomA.F</td>
<td>TTTTGAAGACAAGGAGGCTATCTTCGGTGGCGATAAA</td>
<td>This study</td>
</tr>
<tr>
<td>comA.lux.R</td>
<td>TTTTGAAGACACATTTAAAAGCCATAACCGGTCACTCATTTATGAGACACGTCATAGCAGCGAGCCG</td>
<td>This study</td>
</tr>
<tr>
<td>NearcomA.F</td>
<td>TTTTGAAGACAAGGAGGCTATCTTCGGTGGCGATAAA</td>
<td>This study</td>
</tr>
<tr>
<td>comA.lux.R</td>
<td>TTTTGAAGACACATTTAAAAGCCATAACCGGTCACTCATTTATGAGACACGTCATAGCAGCGAGCCG</td>
<td>This study</td>
</tr>
<tr>
<td>PlacTSS.F</td>
<td>TTTTGAAGACAAGGAGGCTATCTTCGGTGGCGATAAA</td>
<td>This study</td>
</tr>
<tr>
<td>PlacTSS.R</td>
<td>TTTTGAAGACAAGGAGGCTATCTTCGGTGGCGATAAA</td>
<td>This study</td>
</tr>
<tr>
<td>comATSS.F</td>
<td>TTTTGAAGACACATTTAAAAGCCATAACCGGTCACTCATTTATGAGACACGTCATAGCAGCGAGCCG</td>
<td>This study</td>
</tr>
<tr>
<td>comATSS.R</td>
<td>TTTTGAAGACACATTTAAAAGCCATAACCGGTCACTCATTTATGAGACACGTCATAGCAGCGAGCCG</td>
<td>This study</td>
</tr>
<tr>
<td>comARBS.R</td>
<td>TTTTGAAGACAACATTTAAAAGCCATAACCGGTCACTCATTTATGAGACACGTCATAGCAGCGAGCCG</td>
<td>This study</td>
</tr>
<tr>
<td>Fullsxy.F</td>
<td>CACTCTGTGTTCTCAAAATGATGAGAGACACTCTCTTTATGAGACACGTCATAGCAGCGAGCCG</td>
<td>This study</td>
</tr>
<tr>
<td>Fullsxy.R</td>
<td>CACTCTGTGTTCTCAAAATGATGAGAGACACTCTCTTTATGAGACACGTCATAGCAGCGAGCCG</td>
<td>This study</td>
</tr>
<tr>
<td>ComA.3.5.9.</td>
<td>AAGCATTGATATACCCGTTAACGCG</td>
<td>This study</td>
</tr>
<tr>
<td>SDM1.F</td>
<td>CTTGTTTGCCAATTTAATCACCACCAACATACGAGCCG</td>
<td>This study</td>
</tr>
<tr>
<td>ComA.3.5.9.</td>
<td>AAGCATTGATATACCCGTTAACGCG</td>
<td>This study</td>
</tr>
<tr>
<td>SDM1.R</td>
<td>CTTGTTTGCCAATTTAATCACCACCAACATACGAGCCG</td>
<td>This study</td>
</tr>
<tr>
<td>ComA.TSS.</td>
<td>TCTGTTCCCACACACATACGAGCCG</td>
<td>This study</td>
</tr>
<tr>
<td>3.5.9.SDM1.R</td>
<td>TCTGTTCCCACACACATACGAGCCG</td>
<td>This study</td>
</tr>
<tr>
<td>ComA.5bp.</td>
<td>TGATATACACATTTGACGGGTATAAGGAGATC</td>
<td>This study</td>
</tr>
<tr>
<td>SDM1.F</td>
<td>TGATATACACATTTGACGGGTATAAGGAGATC</td>
<td>This study</td>
</tr>
<tr>
<td>ComA.5bp.</td>
<td>TGATATACACATTTGACGGGTATAAGGAGATC</td>
<td>This study</td>
</tr>
<tr>
<td>SDM1.R</td>
<td>TGATATACACATTTGACGGGTATAAGGAGATC</td>
<td>This study</td>
</tr>
<tr>
<td>ComA.TSS.</td>
<td>TCTGTTCCCACACACATACGAGCCG</td>
<td>This study</td>
</tr>
<tr>
<td>5bp.SDM1.R</td>
<td>TCTGTTCCCACACACATACGAGCCG</td>
<td>This study</td>
</tr>
<tr>
<td>ComA.3.5.9.</td>
<td>TTTGTTTTTAATGTGAGACCAG</td>
<td>This study</td>
</tr>
<tr>
<td>SDM2.F</td>
<td>TTTGTTTTTAATGTGAGACCAG</td>
<td>This study</td>
</tr>
<tr>
<td>ComA.TSS.</td>
<td>TTTGTTTTTAATGTGAGACCAG</td>
<td>This study</td>
</tr>
<tr>
<td>3.5.9.SDM2.F</td>
<td>TTTGTTTTTAATGTGAGACCAG</td>
<td>This study</td>
</tr>
<tr>
<td>ComA.3.5.9.</td>
<td>TTTGTTTTTAATGTGAGACCAG</td>
<td>This study</td>
</tr>
<tr>
<td>SDM2.F</td>
<td>TTTGTTTTTAATGTGAGACCAG</td>
<td>This study</td>
</tr>
<tr>
<td>ComA.5bp.</td>
<td>TCGACACTTTATGTTTTTATTGAGACCAG</td>
<td>This study</td>
</tr>
<tr>
<td>SDM2.F</td>
<td>TCGACACTTTATGTTTTTATTGAGACCAG</td>
<td>This study</td>
</tr>
<tr>
<td>ComA.TSS.</td>
<td>TCGACACTTTATGTTTTTATTGAGACCAG</td>
<td>This study</td>
</tr>
<tr>
<td>5bp.SDM2.F</td>
<td>TCGACACTTTATGTTTTTATTGAGACCAG</td>
<td>This study</td>
</tr>
<tr>
<td>ComA.5bp.</td>
<td>TCGACACTTTATGTTTTTATTGAGACCAG</td>
<td>This study</td>
</tr>
<tr>
<td>SDM2.R</td>
<td>TCGACACTTTATGTTTTTATTGAGACCAG</td>
<td>This study</td>
</tr>
</tbody>
</table>
Chapter 4: EHEC phage therapy and characterization of tyrosine-kinase phage exclusion mechanism

This chapter represents material and research done by:

Joshua Yoneda, Alison Tidwell, Nadejda Lupolova, Stephen Fitzgerald, and David Gally

Contributions:

JY: wrote the chapter, designed experiments, analyzed the data; conducted tetrazolium assays (Figure 4.2, 4.3, 4.4, 4.5), constructed plasmid, and performed RT-qPCR.

AT: conducted mutational analysis of the stx2a region of strain 1599 (Figure 4.1). Performed tetrazolium toxicity experiment.

NL: conducted bioinformatic analysis to lay the foundation of this project (Figure 4.1).

SF: Assisted in designing and conducting RT-qPCR experiments (Figure 4.4).

DG: Designed experiments and project goals.
Phage therapy of the food-borne pathogen Enterohemorrhagic *Escherichia coli* (EHEC) serogroup O157:H7 is a promising antimicrobial treatment without the drawbacks of conventional antibiotics. Current barriers to phage therapy include the inefficient process of selecting phages for clinical use and addressing the wide range of phage resistance mechanisms. Recent studies using genomic data and machine learning to predict efficacy of phages against bacteria provide a powerful tool for developing effective phage cocktails. However for accurate predictions the machine learning program must be trained first, which requires the validation of predictions and the feedback of laboratory results into the algorithm. In this study we developed a high throughput method for screening phage-bacteria interactions using tetrazolium dye. Measuring tetrazolium reduction at $A_{485}$ successfully showed which EHEC strains were sensitive to the T4-like typing phage 13 (TP13). EHEC strains without the stx2a lysogen were sensitive to TP13 for the first 12 hours. Relapse growth of TP13 sensitive strains after 12 hours could be suppressed with the addition of T7-like phage TP9. Subsequent mutational analysis showed that deletion of H2 in the stx2a lysogen conferred TP13 sensitivity in the EHEC WT strain. However, plasmid complementation of the $\Delta H2$ mutant did not restore TP13 resistance. RT-qPCR analysis showed decreased $stk$ expression in the $\Delta H2$ mutant, suggesting a negative downstream effect of deleting H2. Cross-reference of recent literature revealed the $stk$ gene has already been characterized for its role in phage resistance. Stk is a tyrosine kinase that initiates abortive infection upon phage infection, resulting in mass phosphorylation and subsequent cell death.
4.1 INTRODUCTION

EHEC is a worldwide pathogen infecting both humans and animals. EHEC causes symptoms such as abdominal cramps, bloody diarrhea, and the potentially lethal condition hemolytic uremic syndrome (HUS)(9). Unfortunately, efforts to treat infection can exacerbate the problem because antimicrobial agents such as kanamycin, fosfomycin, and certain quinolones induce production of the EHEC Shiga toxins (Stx) I and II, which can cause HUS once released into the bloodstream (12). Additionally, antibiotics disturb the natural mammalian microbiome and contribute to the increasing problem of pathogen drug resistance (7). New strategies such as phage therapy are being explored as an alternative to traditional antibiotics to the clinical arsenal.

Phage therapy uses bacteriophages to treat pathogenic bacteria and has been practiced in Europe for over the last 50 years (1). Phage therapy has many advantages, such as the incredible diversity and abundance of phages, and the ability of phages to evolve along with their targets (18). Phage therapy has also shown promise in treating antibiotic resistant infections. In multidrug-resistant (MDR) Pseudomonas aeruginosa, phage therapy was observed to restore the bacterium’s sensitivity to conventional antibiotics (3). Phages are also very specific in the bacteria they target, which reduces the chance of side effects from non-specific targeting of nonpathogenic bacteria or human cells (15).

The incredible diversity of phages is mainly driven by adaption to the selective pressures of phage resistance mechanisms (13). Upon infecting a bacterial cell, the phage faces a wide range of antiviral mechanisms, which work to inhibit phage replication directly or stop phage propagation indirectly. Bacteria are constantly evolving tactics to
prevent phage infection, consequently impeding effective phage therapy. Phage cocktails containing multiple different phages is one possible solution for subverting phage resistance mechanisms (2), however choosing the right phages for treatment is a slow and inefficient process (15). Thus, efficiently selecting phages for clinical use and understanding the genetic bases of bacterial phage resistance are important steps for implementing phage therapy effectively.

Recent phage therapy research has been examining ways to efficiently optimize phage cocktails to treat EHEC and prevent the emergence of resistant strains. One strategy is to incorporate machine-learning with genome sequencing to investigate bacteria-phage interactions and predict the phage sensitivities of different EHEC strains (14, 16). However, in order to accomplish this, one must develop a high-throughput assay to assess all the computer predictions of EHEC-phage interactions and feed it back to the machine learning program. One approach is using the OmniLog™ system, which quantifies cellular respiration via tetrazolium reduction, and can measure up to 50 96-well plates at one time. Henry et al. (11) used the OmniLog™ system to measure the effects of phages on *Bacillus anthracis* growth. Successful phage infection and lysis resulted in reduced bacterial growth and a corresponding reduction in color change. The tetrazolium assay provides a method to quickly assess phage efficacy and characterize phage resistance through mutational analysis of bacterial strains.

Furthermore, there are several advantages to measuring tetrazolium reduction instead of optical density (OD) when dealing with phages. The reduction of tetrazolium is proportional to the number of actively growing cells as opposed to OD, which measures both live and dead cells and other debris. Therefore, in an experimental setup that is
predicated on cell lysis, tetrazolium dye is a much better alternative for measuring cell growth as it is a measure of cell viability.

In this project, I helped develop the high-throughput tetrazolium assay to measure EHEC phage sensitivity and characterize phage resistance of one *E. coli* O157:H7 lab strain. Previous plaque assay experiments showed the deletion of the stx2a prophage from the EHEC strain 9000 chromosome resulted in sensitivity to typing phage 13 (Figure 4.1A). Mutational analysis of different genes in the stx2a phage region (Figure 4.1B) found that deletion of the uncharacterized gene H2 confers TP13 sensitivity in strain 9000. Thus, we conducted experiments to investigate the molecular basis of stx2a-mediated phage resistance in EHEC strain 9000. Additionally, we demonstrate an efficient method to measure the efficacy of phage killing of EHEC strains and the advantages of using multiple phages to treat EHEC.

4.2 MATERIALS AND METHODS

4.2.1 Bacterial Strains, Phages, and Plasmids

The bacterial strains, bacteriophages, and plasmids used in this study are listed in Table 4.1. The *E. coli* strains were cultured in Luria-Bertani (LB) broth at 37°C with shaking. Ampicillin (100 µg/mL for *E. coli*), Isopropyl β-D-1-thiogalactopyranoside (IPTG 1 mM), and 5-bromo-4-chloro-3- indolyl-β-D-galactopyranoside (X-gal 50 µg/mL) were used for selection and screening of correct plasmid constructs. Two phages were used designated as typing phage 9 (TP9) and typing phage (TP13). Phage stocks were prepared by inoculating phage to 50 mL LB containing a phage-sensitive *E. coli* strain. The culture was grown for 8 hours at 37°C with shaking. The broth was then centrifuged at 4000 x g for 7 minutes. The supernatant was filter-sterilized and phage titre
Figure 4.1. The strain 9000 family tree and stx2a phage region.

A shows the family tree of different strain 9000 mutants. Each of the four branches contains a deletion of the full stx2 phage region or just the toxic gene subunits. PT stands for phage type. B shows the stx2a phage region on the strain 9000 chromosome. There are a total of six annotated genes.
### Table 4.1. Bacterial strains, phages, and plasmids used in this study.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bacterial strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Escherichia coli</em> DH5-Alpha</td>
<td>High-efficiency competent cells used for cloning.</td>
<td>(20)</td>
</tr>
<tr>
<td><em>Escherichia coli</em> (EHEC) O157:H7 Strain 9000*</td>
<td>Shiga-toxin producing <em>Escherichia coli</em> O157:H7 strain</td>
<td>(4)</td>
</tr>
<tr>
<td><em>Escherichia coli</em> (EHEC) O157:H7 Strain 1599*</td>
<td><em>Escherichia coli</em> O157:H7 strain 9000 with the stx2a and stx2c toxin genes removed.</td>
<td>(5)</td>
</tr>
<tr>
<td><strong>Phages</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Typing phage 9</td>
<td>T7-like phage</td>
<td>(5)</td>
</tr>
<tr>
<td>Typing phage 13</td>
<td>T4-like lambdoid phage</td>
<td>(5)</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pWKS30</td>
<td>Low copy expression vector</td>
<td>(21)</td>
</tr>
</tbody>
</table>

*Strain 9000 variants not containing the shiga toxin genes were used for experiments in the level 2 biosafety lab.
(pfu/mL) was determined by the double-layer plate method.

4.2.2 Measurement of EHEC phage sensitivity with tetrazolium assay

Bacterial cultures for the tetrazolium growth assay were prepared from overnight cultures diluted to an OD$_{600}$ of 1.0. The optimal concentration of tetrazolium dye for measuring bacterial growth, was determined using a % v/v gradient ranging from 0.0001-1% v/v final tetrazolium concentration. A 96 well microtiter plate was used to measure EHEC phage sensitivity and prepared as follows: 270 µL of LB containing 0.01% v/v tetrazolium dye as added to each well followed by 30 µL of phage. Then 10-fold dilutions of phage were made in subsequent columns with the final column containing just phage and media. The multiplicity of infection (MOI) for the TP13 assays ranged from 0.00001-20. For phage combination assays, the MOI is unknown and 10-fold dilutions are designated as dilutions from the “neat” phage stock. Bacterial culture is then added to each well containing LB and phage to a starting OD$_{600}$ of 0.005. The 96 well plates were incubated in the FLUOstar OPTIMA (BMG Labtech) plate reader for 18-24 hours at 37°C. Absorbance was measured at 485 nm to quantify tetrazolium reduction. The biochemical reaction is shown in Figure 4.2.

4.2.3. Reverse transcription quantitative PCR (RT-qPCR)

Measurement of $H1$, $H2$, and $stk$ gene expression was accomplished by RT-qPCR. RNA samples were taken from a 1/100 dilution culture of 1599, 1599 ΔH1, and 1599 ΔH2 growing at 37°C at the following hour intervals: 0, 2, 3, 4, 5, 6, 7, 8. RNA was stabilized using the RNAProtect Cell Reagent (Qiagen) kit and extracted from each sample using the EZ-10 Spin Column Total RNA Mini-Preps Super Kit (BioBasic). For each sample, 2 µg of total RNA was treated with DNase in a 50 µL reaction using the
Figure 4.2. Reduction of tetrazolium dye to measure cellular respiration.
During cellular respiration, triphenyl tetrazolium chloride (TPH) is reduced to red triphenyl formazan (TPF) by succinate dehydrogenase. Quantification of TPF formation is determined by the spectrophotometric measurement of formazan at 485 nm.
Turbo DNA-free® kit (Ambion), and cDNA templates were synthesized by random priming 200 ng of RNA in a 20 µL reaction using the iScript cDNA Synthesis Kit (BioRad). RT-qPCR was conducted in the StepOne-Plus RT-PCR System (Applied Biosystems) using iTaq SYBR green super mix (BioRad). Standard curves were included in every qPCR run and melt curves were performed at least once with each primer set. Standard curves were generated for each primer set (Figure S4.1) using six serial 10-fold dilutions of chromosomal DNA.

4.2.4. Plasmid complementation of H2 mutant

The 1599 ΔH2 mutant was complemented at gene H2 on the pWKS30 expression vector. Primers for directional cloning into pWKS30 are provided in Table S4.1. H2 was amplified by PCR from strain 1599 chromosomal DNA and purified using the QIAquick PCR Purification Kit (Qiagen). The PCR product primer extensions and pWKS30 vector were double digested by EcoRI and SalI and purified by gel filtration using the QIAquick Gel Extraction Kit (Qiagen). A vector:insert ratio of 3:1 was used and ligation was carried out on the benchtop for 1 hour. The entire ligation mixtures were added to 100 µl chemically competent DH5α cells, incubated for 30 min on ice and transformed by heat shock. Then 900 µl of LB media was then added to the transformation, and the cells were allowed to recover for 1 hour at 37 °C with shaking. Next, 200 µl of the transformation were plated on LB plates containing ampicillin (100 µg/mL) and X-gal/IPTG and incubated overnight at 37 °C. Recombinant colonies were grown overnight, and the plasmids prepared using the QIAprep Spin Miniprep Kit (Qiagen). Correct plasmids were confirmed by restriction digest and PCR.

4.3 RESULTS
4.3.1 The stx2a phage region confers resistance to TP13

To assess the utility of the tetrazolium assay, killing by TP13 was tested at different multiplicity of infections (MOIs) against three different strains of EHEC (1465, 1599, and 1599ΔH2), described in Figure 4.1. Strain 1599 is considered the “wild type” EHEC strain, which lacks the shiga toxin genes and so is safe to work with in a Biosafety Level 2 lab. Strain 1599ΔH2 does not contain the H2 gene in the stx2a region; the whole stx2a prophage region was removed from strain 1465.

When EHEC grows, it will reduce tetrazolium to 1,3,5-triphenylformazan, which produces a red colour detectable by spectroscopy. Titration of tetrazolium was used to determine that 0.01% v/v is the best concentration to measure EHEC growth; higher tetrazolium concentrations inhibited EHEC growth and lower concentrations did not produce an adequate color change for absorbance measurement. The multiplicity of infection (MOI) ranged from 0.0001-20, depending on the available phage stock. As shown in Figure 4.3, the reduction of tetrazolium by strains 1465, 1599, and 1599ΔH2 showed a sigmoidal curve, similar to a typical bacterial growth curve. Addition of TP13 was effective in suppressing growth of strains 1465 and 1599ΔH2 for the first 12 hours at an MOI range of 0.01-1. The late growth of 1465 and 1599ΔH2 after 12 hours is indicative of phage resistant bacteria rising to a detectable number in the populations. Multiple late growth patterns were detected at different MOIs. At the lowest MOI, growth was only partially inhibited, however at the second lowest MOI a plateau is reached. Higher MOIs resulted in a late growth pattern. Measurements were continued to 42 hours, which revealed a similar late growth trend in all MOI conditions for 1465 and 1599ΔH2 (data not shown). Strain 1599 demonstrated resistance to TP13 in all MOI
**Figure 4.3. Measurement of EHEC growth treated with different dilutions of TP13**

*Escherichia coli* (EHEC) O157:H7 strains 1465, 1599, and 1599ΔH2 were cultured at 37°C in LB media with 0.01% v/v tetrazolium dye. Each well contained different MOI dilutions of TP13 ranging from 0-1 and cells were added to a final starting OD$_{600}$ of 0.005. Measurement was carried out for 22 hours and data is represented by $A_{485}$ mean ± SD (n = 3).
conditions, reaching almost the same A_{485} values as for the no phage condition.

4.3.2 Addition of TP9 to TP13 prevents relapse growth of Δstx2a strains

The combination phage therapy of TP9 and TP13 showed impressive results, completely suppressing 1465 and 1599ΔH2 growth at the neat (original phage stock) concentration for 22 hours (Figure 4.4). The resurgent growth observed with TP13 at t = 12 hours was not observed in the phage combination experiment in two cases (Neat and 10^{-1} dilution). The 1599ΔH2 mutant actually showed a decrease in tetrazolium reduction at dilutions 10^{-4} and 10^{-5} after t = 4 hours, which is indicative of lysis from phage infection after cell growth is detected. Lysis after several hours demonstrates continuous action of phage throughout the duration of the experiment. Resistant subpopulations of 1465 and 1599ΔH2 were still observed in the phage cocktail treatment but only at certain dilutions. The minimum number of each TP for complete suppression of EHEC growth was not evaluated in this study.

4.3.3 Expression profile of stx2a phage region

Previous mutational analysis showed deletion of H2 in the stx2a prophage confers TP13 sensitivity to strain 1599, leading us to hypothesize that H2 contributes to phage resistance. The expression profiles of three genes (H1, H2, stk) were measured in strain 1599 to test the hypothesis that they are constitutively expressed. Additionally, stk expression was measured in the 1599ΔH2 mutant to see if the H2 deletion had any downstream effects on gene expression. RT- qPCR revealed that H1, H2, and stk are constitutively expressed in strain 1599 (Figure 4.5A). The highest relative gene expression of H1, H2, and stk was observed at t = 4 hours. Furthermore, stk expression
Figure 4.4. Measurement of EHEC growth treated with TP9 and TP13

*Escherichia coli* (EHEC) O157:H7 strains 1465, and 1599ΔH2 were grown at 37°C in LB media with 0.01% v/v tetrazolium dye. Each well contained different dilutions of TP9 + TP13 ranging from Neat-10^-6 and cells were added to a final starting OD_{600} of 0.005. Measurement was carried out for 22 hours and data is represented by A_{485} mean ± SD (n = 3). The final pfu/mL value of the phage combinations was not measured, which is why dilutions from the stock concentration (Neat) are used.
was detectable but greatly reduced in 1599ΔH2 compared to 1599 (Figure 4.5B),
suggesting that deletion of H2 negatively impacts the expression of downstream gene stk.

4.3.4 Plasmid complementation of the 1599ΔH2 mutant does not restore TP13
resistance

To test whether H2 function and phage resistance could be restored, the H2 gene
was cloned into the pWKS30 expression vector and transformed into 1599ΔH2 for gene
complementation. Given the previous growth curves showing TP13 sensitivity upon H2
deletion (Figure 4.3), we hypothesized plasmid complementation of H2 would restore
TP13 resistance in 1599ΔH2. As shown in Figure 4.6, 1599ΔH2+pH2 follows the same
growth pattern of 1599ΔH2 at the different phage dilutions, suggesting that restoring H2
does not restore TP13 resistance. The previous RT-qPCR results revealed that H2
deletion also decreased stk expression. Analyzing these results together, we have reason
to suspect H2 is not responsible for TP13 resistance. The 1599Δstk mutant was also
tested against TP13 dilutions and showed a similar sensitivity to TP13 as 1599ΔH2.
Therefore, the stk gene may be responsible for TP13 resistance in strain 1599.

4.4 DISCUSSION

Phage therapy has been pursued as a therapeutic alternative to antibiotics for
several decades and brings many advantages compared to the standard arsenal of
antimicrobial agents. In addition to being the most abundant biological entity on the
planet, phages are incredibly diverse and are constantly evolving to effectively engage
bacteria (18). The incredible diversity of phages is also a setback when considering how
to effectively study the amazingly high number of possible phage-bacteria interactions.
Figure 4.5. RT-qPCR relative expression of H1, H2, and stk.
Cells were grown in LB at 37°C with shaking in an incubator and samples were taken at the indicated time intervals. Data are represented by quantity mean ± SD (n = 3). A shows the RT-qPCR measurements of H1, H2, and stk in strain 1599. B shows the RT-qPCR measurement of stk in strain 1599ΔH2.
Figure 4.6. Measurement of EHEC growth treated with TP13.

*Escherichia coli* (EHEC) O157:H7 strains 1599, 1599ΔH2, 1599ΔH2 + pH2, and 1599Δstk were grown at 37°C in LB media with 0.01% v/v tetrazolium dye. The 1599ΔH2 + pH2 strain contained the plasmid H2 complement and was induced with IPTG at 1 mM. Each well contained different dilutions of TP13 ranging from a MOI of 0-20 and cells were added to a final starting OD$_{600}$ of 0.005. Measurement was carried out for 17 hours and data is represented by $A_{485}$ mean ± SD (n = 3).
One possible solution is using machine-learning to predict which phages will be effective against certain bacteria based on their genome sequences (14). However, in order to do this, a considerable amount of information must be provided to train and refine a machine-learning algorithm. In this experiment we used the colorimetric reduction of tetrazolium to quantify phage effectiveness against EHEC strains. Our results show successful measurement of EHEC growth and indicate that the tetrazolium assay can be scaled up using the OmniLog™ system to test thousands of phage-bacteria interactions in a 24-hour period. These results are corroborated by a study of phage killing of B. anthracis using the OmniLog™ system (11).

The phage cocktail containing TP9 and TP13 suppressed growth of strains 1465 and 1599ΔH2 upon addition of the most concentrated phage cocktail (Figure 4.4). Comparison to the resurgent growth observed in the TP13 treatment of 1465 and 1599ΔH2, presents an exciting result demonstrating the additive effect of different phages. TP9 is a T7-like phage and TP13 is a T4-like phage, providing two different mechanisms of assault. In theory, if 1465 develops resistance to TP13 through mutation of an outer membrane protein, it is still sensitive to TP9 which targets a different membrane protein. This phenomenon has been observed in several phage therapy studies (2, 10, 19, 22), and supports the idea of using a diverse phage mixture for bacteria treatment.

The $H2$ gene in the stx2a region of 1599 is not responsible for TP13 resistance because plasmid complementation of the 1599ΔH2 mutant did not restore the resistance phenotype. RT-qPCR results showed the $H2$ knockout mutant has decreased gene expression of adjacent genes $H1$ and $stk$. In addition, previous lab work from our lab
showed H1 complementation did not restore TP13 resistance (A. Tidswell, personal communication). Therefore, it is likely that stk, annotated as a protein kinase, is responsible for TP13 resistance in 1599. This hypothesis is consistent with the recent work by Friedman et al. (8), which identified and characterized the stk gene in the 933W lysogen of *Escherichia coli* strain O157:H7. After comparative analysis, it was determined that the 933W lysogen is the same as the stx2a phage region, and stk is the protein kinase shown in Figure 4.1B. Friedman et al. (8) characterized the stk protein as a tyrosine kinase that protects EHEC from HK97 phage infection. The mechanism of protection is not described, however it appears the end goal is mass protein phosphorylation resulting in cell death, described as abortive infection. An abortive mechanism is consistent with the results shown in Figure 4.6. At a MOI of 20, strain 1599 exhibits a clear decrease in growth, which may result from an abortive infection mechanism being triggered in the vast majority of 1599 cells from the high phage:bacteria ratio. If the majority of the bacteria die, then an obvious delay in growth will ensue. Plasmid complementation of the 1599Δstk mutant would corroborate the evidence provided by Friedman et al. and RT-qPCR measurement of 1599 stk expression in the presence of TP13 would indicate if stk is upregulated or not in the presence of phage. Additionally, mutating TP13 so it is able to infect 1599 would be a starting point to describing the mechanism of stk-mediated abortive infection. By sequencing the mutated TP13 and comparing it to the wild type, we could see which genes are important to work around this defence mechanism.

A similar phage-resistance mechanism has been described in more detail in *Staphylococci*, where a Eukaryotic-like Serine/Threonine Kinase was shown to activate
cell death to block phage propagation (6). Based on this phage exclusion mechanism, it is speculated that EHEC Stk is activated by a phage protein injected into the cytoplasm. Once activated, Stk mass phosphorylates essential cellular pathways involved in global transcription control, cell-cycle control, stress response, DNA topology, DNA repair, and metabolism, leading to cell death and ultimately blocking phage propagation.

4.5 Acknowledgements

I would like to express my appreciation to Dr. David Gally for his helpful advice, Dr. Alison Tidswell for doing the previous mutational experiments, and Dr. Stephen Fitzgerald for his excellent assistance in running the RT-qPCR experiments. This work was funded by the Queen Elizabeth II Diamond Jubilee Scholarship.

REFERENCES


### Table S4.1. List of oligonucleotide sequences used for RT-qPCR and cloning.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5’-3’)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cloning</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sall.H2.F</td>
<td>AAAGTCGACGCGCCTGATTCTAAAAATAGG</td>
<td>This study</td>
</tr>
<tr>
<td>EcoRI.H2.R</td>
<td>AAAGAATTTCTACATCAGCTCTTTGTATGG</td>
<td></td>
</tr>
<tr>
<td><strong>RT-qPCR</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1.qPCR.fwd</td>
<td>GGCAATCCAGGTAAACAAAC</td>
<td>This study</td>
</tr>
<tr>
<td>H1.qPCR.rev</td>
<td>TCATATCCTCGCATCCATAG</td>
<td></td>
</tr>
<tr>
<td>H2.qPCR.fwd</td>
<td>CGACCTTAGACACACCCAAA</td>
<td>This study</td>
</tr>
<tr>
<td>H2.qPCR.rev</td>
<td>TAACCGGTATGGGCAGCTAAT</td>
<td></td>
</tr>
<tr>
<td>stk.qPCR.fwd</td>
<td>AAACACTAACCACGACGACTAC</td>
<td>This study</td>
</tr>
<tr>
<td>stk.qPCR.rev</td>
<td>CATTGAGCAGCATACTGAACTTG</td>
<td></td>
</tr>
</tbody>
</table>