Characterizing the function of the SigM regulon in mycobacterium smegmatis

Sarah Montgomery
University at Albany, State University of New York, montgomerysarah8@gmail.com

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Part of the Biology Commons

Recommended Citation
https://scholarsarchive.library.albany.edu/legacy-etd/2529

This Master's Thesis is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive.
Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
Characterizing the function of the SigM regulon in *Mycobacterium smegmatis*

by

Sarah A. Montgomery

A Thesis
Submitted to the University at Albany, State University of New York
In Partial Fulfillment of
the Requirements for the Degree of
Master of Science

School of Public Health
Department of Biomedical Sciences
Spring 2020
ABSTRACT

Secretion systems, which are essential in bacteria, secrete proteins that maintain cellular homeostasis, promote growth, reduce the impact of environmental stressors, and promote virulence in pathogens. In addition to universally conserved Sec and Tat secretion systems, mycobacteria utilize type VII, or ESX, secretion systems (T7SS), to translocate proteins across their mycolic acid-rich cell wall.

The goal of this project is characterization of esx-4, one of five possible ESX systems encoded by mycobacteria. Esx-4 is the hypothesized progenitor of all other T7SS and, thus, makes a good model for studying T7SS. All proteins encoded by esx-4 are required for distributive conjugal transfer (DCT), a unique form of horizontal gene transfer first observed in Mycobacterium smegmatis.

Esx-4 is co-regulated with additional genes by Sigma factor M (SigM). In addition to esx-4, several other SigM-regulated genes were found to be necessary for DCT. However, the entire SigM regulon has not been characterized using DCT. Therefore, we determined whether several remaining SigM-regulated genes are required for DCT, and through our analyses, we discovered that the SigM regulon protein MKD8_6925 is required for DCT.

We hypothesized that SigM-induced proteins encoded outside the core esx-4 locus are substrates of ESX-4. Therefore, we also examined ties between ESX-4 and the additional proteins regulated by SigM using mycobacterial protein fragment complementation (M-PFC), a form of 2-hybrid analysis. This assay was used to establish whether SigM regulon proteins form protein-protein interactions (PPI) that could be relevant to their secretion through ESX-4. Using M-PFC, we identified protein-protein interactions between MKD8_1686 and MKD8_1685 as well as EsxU and EsxT.
T7SS are required by several mycobacterial pathogens for virulence, and studying a model T7SS will inform our understanding of mechanisms behind T7SS-mediated pathogenesis in mycobacteria. This could, in turn, aid in the development of new prophylactics and therapeutics that could quell the spread of pathogens that are hard to treat, including mycobacteria.
ACKNOWLEDGEMENTS

I would like to thank my mentors, Dr. Keith Derbyshire and Dr. Todd Gray, whose energy and passion for science and learning have inspired and motivated me throughout completion of this thesis project. I want to thank them for mentoring me, encouraging curiosity, preparing me to move confidently into a scientific career, and for their patience. Additionally, I would like to recognize Melissa Stone, Ryan Clark, and Jill Canestrari, all members of the Gray/Derbyshire lab, for their countless contributions, sharing of ideas that influenced the outcomes of this project, and for making research incredibly fun. I would like to also thank Sophie Boerboom who was instrumental in conducting M-PFC experiments as an REU student.

I would like to thank Dr. Kathleen McDonough and Dr. Pallavi Ghosh for serving on my thesis committee and providing help wherever they could. Also, I would like to thank Dr. Matthew Champion who provided insight on many aspects of this project and carried out mass spectrometry experiments on behalf of our lab.

Finally, I would like to recognize and thank Dr. Valerie Bolivar, Dr. Magdia De Jesus, and Dr. Janice Pata for their continued academic and emotional support throughout my time at UAlbany.
## TABLE OF CONTENTS

Abstract .................................................................................................................................................. ii

Acknowledgements ............................................................................................................................... iv

List of Figures ....................................................................................................................................... vi

List of Tables ......................................................................................................................................... vii

- **Chapter 1: Introduction** ..................................................................................................................... 1
- **Chapter 2: Results** ............................................................................................................................. 22
- **Chapter 3: Discussion** ....................................................................................................................... 50
- **Chapter 4: Methods** ......................................................................................................................... 70

References .............................................................................................................................................. 83
LIST OF FIGURES

Chapter 1

Figure 1.1. Characteristics of bacterial secretion systems .................................................. 16
Figure 1.2. Features of mycobacterial T7SS ................................................................. 18
Figure 1.3. ESX-4: Hypothesized structure and requirement during DCT ....................... 19
Figure 1.4. The SigM regulon, DCT, and colony morphology .......................................... 20

Chapter 2

Figure 2.1. Features of MKD8_1560 and MKD8_15750 .................................................. 35
Figure 2.2. Features of MKD8_1687 operon ................................................................. 36
Figure 2.3. Features of MKD8_6925 ................................................................................. 37
Figure 2.4. Recombineering approach utilized to generate MKD8 knockouts .................. 38
Figure 2.5. SigM-regulated genes and their requirement during DCT ............................... 40
Figure 2.6. EGTA and MKD8 viability ................................................................................ 41
Figure 2.7. The effects of EGTA on DCT ......................................................................... 42
Figure 2.8. The effects of calcium supplementation on DCT ............................................ 43
Figure 2.9. EGTA affects expression of SigM-regulated genes during co-culture ............. 44
Figure 2.10. Plasmids used for M-PFC ............................................................................ 46
Figure 2.11. General principles of M-PFC ....................................................................... 47
Figure 2.12. M-PFC results .............................................................................................. 49

Chapter 3

Figure 3.1. Strategy to determine the substrates of ESX-4 ................................................ 66
Figure 3.2. Preliminary MS results reveal possible substrates of ESX-4 ............................ 68
LIST OF TABLES

Chapter 1

Table 1.1. Features of the SigM regulon................................................................. 21

Chapter 2

Table 2.1. Summary of M-PFC results................................................................. 48

Chapter 3

Table 3.1. Strains used for co-dependent secretion analysis................................. 65

Chapter 4

Table 4.1. Antibiotic concentrations used for bacterial cultures........................... 78

Table 4.2. Oligonucleotides used for research...................................................... 79
Chapter 1. Introduction

Secretion systems are vital for bacterial survival and cellular homeostasis. All bacteria possess secretion systems known as Sec (general secretion) and Tat (twin arginine transport) that are located in the cytosolic membrane (Green, E., & Mecsas, J., 2016). These systems export protein substrates, which contain specific N-terminal signal sequences, from the cytosol across the cytoplasmic membrane. In Gram-positive bacteria protein substrates are transported to the extracellular matrix (ECM), whereas the periplasm is the destination of Sec and Tat substrates in Gram-negative bacteria (Green, E., et al., 2016). These systems, described in more detail below, transport the majority of proteins produced by bacteria and are vital to homeostasis, survival, and pathogenesis in bacteria.

Sec System
Sec was the first secretion system described in bacteria and requires three distinct components necessary for the translocation of unfolded proteins across the inner bacterial membrane (IM), which are ubiquitous and essential to bacterial survival. Sec secretion systems are composed of protein targeting components (SecB/SecA), an ATPase motor protein (SecA), and an integral translocation channel (SecYEG) (Figure 1.1A) (Wickner, W., et al., 1991). Proteins secreted via Sec are directed to the apparatus through conserved, hydrophobic N-terminal signal sequences that are recognized by SecB (Reviewed in Green, E. R., & Mescas, J., 2016). SecB prevents protein folding and, along with SecA, shuttles proteins to SecYEG for translocation. After transport into the periplasm a protease cleaves the SecB recognition site from the substrate, which initiates protein folding. Sec signal sequences also play a role in the final destination of the protein, which can be the IM (integral protein) or ECM (secreted substrate). Sec is involved in the general well-being of bacteria but is also involved in the virulence of some pathogens.
**Tat System**

In contrast to Sec, which transports proteins directly from the ribosome in their unfolded state, Tat apparatuses export substrates in their folded state (Robinson, C., & Bolhuis, A. 2004). Tat systems are comprised of proteins that bind to the substrate being transported (TatB/TatC) and the channel through which the substrate is being secreted (TatA) (Reviewed in Green, E. R., & Mescas, J., 2016). As opposed to Sec, this pathway is named for the presence of two, “twin” arginines in the N-terminal domain of Tat-dependent substrates, which serves as the Tat signal sequence. The Tat signal sequence is recognized by chaperones TatB and TatC, which both help to guide substrates to TatA, the porin through which substrates are secreted (Figure 1.1B).

In Gram-negative bacteria, Tat substrates either remain in the periplasm or are exported to the ECM through other secretion systems, which will be discussed later in this chapter. Like Sec, Tat is required for cellular homeostasis, growth, and virulence. It was found that pathogenic *Escherichia coli* strain O157:H7 requires Tat secretion for virulence (Pradel, M., et al., 2003).

Gram-negative bacteria require additional secretion apparatuses to transport proteins from the periplasm across the outer membrane (OM). These apparatuses are referred to as type I through type IX secretion systems (T1SS – T9SS) and each export a unique set of substrates that play a role in bacterial survival, pathogenesis, or both (Green, E. et al., 2016). T2SS, T5SS, T8SS, and T9SS are reliant upon Sec and Tat for secretion of substrates into the periplasm, as they only span the OM (Figure 1.1C). In contrast, the remaining secretion systems work independently of Sec and Tat as they span both membranes (Figure 1.1C). In some cases, these systems can span an additional membrane belonging to a host cell, which aids in the direct delivery of virulence factors to the cytoplasm (Figure 1.1C).

**Sec- and Tat-dependent Secretion System – T2SS**

T2SS are embedded in the OM, which makes them dependent upon Sec or Tat delivery of substrates to the periplasm (Figure 1.1C). T2SS are present in most Gram-negative organisms
and its substrates are typically enzymes that contribute to a wide range of biological processes such as the processing of complex carbohydrates (Korotkov, K., et al., 2012). T2SS are comprised of 12-15 different genes that are generally co-transcribed in an operon (Sandkvist M., 2001). These genes encode core T2SS components which include a pseudopilus, OM complex, and an ATPase required for active transport (Recently reviewed in Korotkov, K. V., and Sandkvist, M., 2019). T2SS are required for the secretion of cholera toxin, one of the main virulence factors of *Vibrio cholerae* (Sandkvist, M., 2001).

**T5SS**

T5SS substrates are secreted via Sec to the periplasm, where they then self-translocate across the OM. This unique process is possible because T5SS substrates contain a domain capable of insertion into the OM to form a β-barrel (Figure 1.1C) (Leyton DL., et al., 2012). This insertion creates a channel that is used for secretion of the N-terminal portion of the substrate (autotransport) or another substrate entirely (two-partner secretion). Many T5SS substrates are virulence factors that exhibit toxin activity. *Pseudomonas aeruginosa* protein EstA is autotransported from the cell and contributes to changes in OM composition and biofilm formation (Wilhelm, S., et al. 2007).

**T8SS**

This system, also known as the curli secretion system, is dedicated to the transport of curli pili. Curli pili are fibrous structures involved in colonization of host tissues by pathogenic bacteria, establishing cell-cell contact between neighboring bacteria, and promoting biofilm formation (Barnhart, M. M., & Chapman, M. R., 2006). Proteins that assemble curli pili are secreted to the periplasm by Sec where they are then translocated through the OM by T8SS. This secretion event results in the assembly of a pilus on the OM. *Salmonella* spp. utilize T8SS to promote adhesion of bacteria to the human gut epithelium, biofilm formation, and eventually host invasion (Austin, J. W., et al., 2006).
**T9SS**

This system, unlike others, is only found in the *Bacteroidetes* phylum, which includes anaerobic, Gram-negative bacilli found in soil, water, and the human gastrointestinal tract (Reviewed in Lasica, A. M., et al., 2017). T9SS have been implicated in virulence of *Porphyromonas gingivalis*, an oral pathogen of humans. These bacteria use T9SS to secrete enzymes called gingipains that damage eukaryotic cells and cause disruptions in innate immunity (Potempa., et al., 2003). T9SS have also been found to secrete proteins that facilitate gliding motility in bacteria that utilize this type of movement (Lasica, A. M., et al., 2017). Substrates can be secreted to the periplasm by Sec or, through a C-terminal signal sequence, directly by T9SS.

**Sec and Tat-independent Secretion Systems – T1SS**

T1SS, which resemble cellular transporters of other small molecules, secrete proteins in a one-step fashion that does not require Sec or Tat (Thomas, S., et al., 2014) (Figure 1.1B). T1SS are comprised of an ABC transporter required for translocation of protein substrates across the inner membrane, a transmembrane protein that connects the cytosolic portion of the system to a porin in the OM, which is the final core structural component of T1SS (Thomas, S., et al., 2014). These core structural components work to export substrates that are targeted to the system by C-terminal signal sequences as opposed to N-terminal sequences found in Sec and Tat-dependent effectors (Thomas, S., et al., 2014). *Serratia marcescens* uses T1S to secrete HasA, a hemophore that sequesters heme and leads to breakdown of the host circulatory system (Létoffé, S., et al., 1996).

**T3SS**

These secretion systems span the IM and OM of many Gram-negative bacteria and can extend into the cytosol of a host for injection of virulence factors, leading to T3SS being described as “injectisomes” (Figure 1.1C). This one-step injection process requires coordination between highly conserved core proteins and additional associated proteins for effective secretion. These
systems can be chromosomally encoded or expressed on a plasmid. T3SS substrates contain non-cleaveable signal sequences and perturb the cytoskeleton of eukaryotic host cells which leads to enhanced bacterial invasion (Reviewed in Coburn, B., et al., 2007).

**T4SS**

Although T4SS primarily secrete protein substrates, DNA can be transported through T4SS if it is linked to the substrate being secreted. T4SS are used in DNA conjugation to transport plasmid DNA directly from donor cells into recipients. These transfer events result in the passage of extrachromosomal plasmids that often confer resistance to antibiotics or other fitness advantages to recipient cells. For example, *Neisseria gonorrhoeae* utilizes T4SS to facilitate the transfer of DNA relevant to virulence, which are incorporated into the genome of the recipient bacterium via homologous recombination (Hamilton, H. L., and Dillard, J. P., 2006).

**T6SS**

These secretion systems are found in almost 25% of sequenced bacterial genomes (Boyer, et al., 2009). These large complexes, comprised of up to 20 different structural proteins, transfer substrates into recipient cells in a contact-dependent manner (Figure 1.1C) (Reviewed in Russell, A. B., et al., 2014). T6SS have certain structures homologous to those in phage leading researchers to the hypothesis that T6SS may have arisen partially from phage and that they function similar to T3SS, which also exhibit phage-like features (Russell, A. B., et al., 2014). *Yersinia pestis* uses T6SS to secrete substrates important for its virulence. These substrates aid in adaptation to the eukaryotic host environment and niche establishment among commensal organisms (Yang, X., et al., 2018).

**T7SS**

Mycobacteria, which encode T7SS, are classified as Gram-positive, but possess a diderm membrane structure similar to that of Gram-negative bacteria. The cell wall of mycobacteria is
highly complex and contains a high concentration of mycolic acids which, together with other lipids and polysaccharides, create a lipid-rich bilayer or outer membrane “OM” known as the mycomembrane (MM). This structure protects the cell from the negative impacts of environmental and host stresses such as extremes in pH, the presence of radicals, or exposure to antibiotics. However, while the mycobacterial OM protects these organisms, it also acts as a barrier between these bacteria and their environment. Therefore, mycobacterial secretory apparatuses are extremely important for maintaining nutrient homeostasis, promoting communication between bacteria, and, in pathogens, virulence.

Mycobacteria encode Sec, SecA2, and Tat secretion systems (Champion, P. A. and Cox, J. S., 2007). However, these systems only export proteins from the cytoplasm to the periplasm between the IM and MM, meaning that another secretory apparatus is required for secretion of non-Sec and -Tat protein substrates. Mycobacteria were predicted to have a unique type of secretory apparatus when researchers discovered secreted proteins from *M. tuberculosis* that lacked signal sequences required for secretion by Sec (Pallen, M., 2002). These proteins, ESAT-6 (early secreted antigenic target) and CFP-10 (culture filtrate protein), renamed EsxA and EsxB respectively, were found to be secreted by ESX-1, one of five possible T7SS found in mycobacteria (Figure 1.2A)(Andersen, P., et al., 1995, Berthet, F., et al., 1998, Pittus, N., et al., 2001). Orthologous systems named Type-VIIb secretion systems have been identified in pathogens such as *Staphylococcus aureus* and *Listeria monocytogenes*, making T7SS and T7-like systems all that more important to study. By studying T7SS, targets for more potent therapeutics or prophylactics against bacteria that use T7SS to mediate virulence may be identified.

Mycobacteria encode up to five ESX systems, which are paralogues and perform non-redundant functions (Figure 1.2B). Paralogous ESX systems evolved from the universally conserved *esx*-4 locus, the topic of this thesis. It’s important to note that due to its conservation
and progenitor status, \textit{esx-4} is a good model for dissecting the common mechanisms of secretion via T7SS and the role(s) of their substrates.

\textit{ESX Conserved Components (Ecc)}

These proteins, as their name implies, are conserved in all ESX systems. EccA, which is not encoded by \textit{esx-4}, is not essential for secretion but is hypothesized to act as a chaperone for protein substrates (Reviewed in Phan, T. H., & Houben, E. N. G., 2018). Additionally, EccA of \textit{Mycobacterium marinum} appears to interact with mycolic acids and promote their synthesis (Joshi, S. A., et al., 2012). Based on structural analyses of ESX-5 at the membrane, which have provided a model for T7SS structure, EccB is required for ESX function as it complexes with two other secretion proteins, EccC and EccD, to form the membrane channel that substrates are transported through (Figure 1.2C) (Houben, E. N., et al., 2012). EccC contains 3 ATPase domains which are located in the cytoplasm and are thought to drive secretion across the IM (Figure 1.2C) (Reviewed in Houben, E. N. G., et al., 2014). In addition to powering T7SS, the ATPase domains of EccC are also critical for substrate recognition and translocation (Callahan, B., et al., 2010, Rosenberg, OS., et al.,2015). EccD is a highly hydrophobic protein with 11 transmembrane domains and makes up a large portion of the secretion channel (Figure 1.2C) (Houben, E. N. G., et al., 2012). Another Ecc protein known as EccE is required for secretion, but is not encoded in \textit{esx-4} (Figure 1.3A). EccE localizes to the IM with EccB, EccC, and EccD and is essential for secretion via all other ESX systems (Houben, E. N. G., et al., 2012). Mycosins, described below, are also essential for secretion via all ESX systems, but are not referred to as “Ecc” proteins.

\textit{Mycosins}

Mycosins are proteases encoded by each \textit{esx} locus, but their cleavage targets are largely unknown (Figure 1.2C). The mycosin of ESX-1, MycP1, was found to bind and cleave EspB, a substrate of ESX-1 (Ohol, Y. M., et al., 2010). This implicates mycosins as potential ESX-
substrate-processing proteins. However, other ESX substrates, including WXG100 proteins, are not cleaved during secretion into the periplasm. Thus, mycosins appear to process highly specific targets but may play a role in T7S that has not yet been described.

**WXG100 Proteins**

Each esx locus contains an operonic pair of genes that encode small, 100 amino acid (aa) secreted substrates defined by the presence of a tryptophan, variable, glycine motif (WXG). These proteins are similar in structure and contain two anti-parallel alpha-helices and a turn facilitated by the WXG motif. EsxA and EsxB are the current models for WXG100 protein structure (Figure 1.2A). Each ESX system encodes a unique pair of WXG100 and may secrete other non-esx encoded WXG proteins known as ESX secretion-associated proteins (Esp) (Fortune, S. M., et al., 2005). The functionality of these substrates is critical as they determine the biological role(s) of the secretion system itself. WXG100 substrates have been shown to exhibit a mutually co-dependent mechanism of secretion (Fortune, S. M., et al., 2005). WXG100 proteins are trafficked to ESX systems through a conserved YxxxD/E motif at the C-terminus of the substrate. Recent research shows that secretion of WXG T7SS substrates through all T7SS, except for ESX-4 (T7SS not explored), is mediated by a conserved nucleotide-binding pocket located in the third ATPase domain of EccC (Wang, S., et al., 2019).

**ESX-1 – History and Function**

The ESX-1 secretion system has been heavily studied and is essential for virulence of *M. tuberculosis*. In the early 1900s, Albert Calmette and Camille Guérin isolated bacillus Calmette-Guérin (BCG), a strain of *Mycobacterium bovis*, a close relative of *M. tuberculosis*, through continuous passaging in hopes of identifying an attenuated vaccine candidate (Calmette, A., et al. 1927). The resulting strain was found to be attenuated *in vivo* and was capable of protecting vaccine recipients subsequently infected with *M. tuberculosis* (Calmette, A., et al. 1927). BCG has since been used as the standard live-attenuated vaccine for *M. tuberculosis*. While it has
helped many, this vaccine exhibits a broad efficacy, meaning there is a chance that immunized, and of course non-immunized, individuals are still susceptible to infection. We do not have a comprehensive understanding of the roles T7SS play during infection, which highlights the need to expand research on mycobacterial T7SS. There are other non-ESX-1 T7SS that may harbor targets for the development of more efficacious vaccines and therapeutics.

Without advanced genetic techniques, it was unknown why BCG exhibited attenuation until the late 20th century. Researchers in the late 1990s compared the BCG and *M. tuberculosis* genomes and, through these analyses, discovered regions of difference (RD) between the two strains (Mahairas, G., et al. 1996). Of these regions, RD1 proved to be the primary attenuating mutation found in all strains of BCG (Pym, A. S., et al. 2002). This was confirmed when the RD1 region was complemented in BCG and led to significant recovery of virulence (Pym, AS., et al. 2002). Soon after, the entire esx-1 locus was soon identified as the attenuating factor, which opened a new field of research into the function(s) of ESX-1 substrates and their role(s) in pathogenesis.

The functions of ESX-1 substrates are diverse and include degradation of the phagosomal membrane that envelopes *M. tuberculosis* upon entry into macrophages during infection (de Jonge, M. I. et al., 2007). This leads to its escape from the phagosome and the establishment of infection in the cytosol of infected cells. *M. marinum*, a close relative of *M. tuberculosis* that causes soft tissue infections in humans and fish, also exhibits an ESX-1 substrate-dependent escape from phagosomes (Conrad, WH., et al. 2017). In *M. marinum*, ESX-1 substrates EsxA and EsxB contribute to hemolysis of human red blood cells (Gao, L., et al., 2004). Although ESX-1 is a major virulence-associated system, there are pathogenic mycobacteria that do not contain this locus, such as *Mycobacterium abscessus*, another relative of *M. tuberculosis* that causes soft tissue infections in humans. Therefore, there are other T7SS required for pathogenesis in *M. abscessus* and again highlights the need to further study T7SS and their
potential as targets for therapeutics. ESX-1 in the non-pathogen *M. smegmatis* is required for a form of horizontal gene transfer termed Distributive Conjugal Transfer (DCT), which will be described later in this chapter (Coros, A. et al., 2008). This demonstrates the complexity of T7SS and highlights the importance of understanding each of these systems in pathogens and non-pathogens alike.

**ESX-2**

Little is known about *esx-2* as active secretion of substrates via this system has not been described. However, there is evidence that this system may be active or, at least, transcribed because EspR, a known regulator of ESX-1 related virulence genes, binds to regions within *esx-2* (Blasco, B., et al. 2012). Additionally, a transposon mutagenesis screen of *M. tuberculosis* that identified strains with decreased fitness in human dendritic cells, identified insertions in *esx-2* (Mendum, T. A., et al. 2015). Beyond these data, *esx-2* remains largely unexplored.

**ESX-3**

This secretion system is conserved across mycobacteria and is critical for iron and zinc uptake (Serafini, A., et al., 2009, Serafini, A., et al., 2013). This system, which is required for growth of *M. tuberculosis*, was found to be regulated by two different proteins in *M. tuberculosis* including the zinc-uptake regulator (Zur) and the iron-dependent transcriptional repressor (IdeR) (Maciag, A., et al. 2007, Rodriguez, G. M., et al. 2002). Only IdeR is responsible for the regulation of *esx-3* in *M. smegmatis* (Maciag, A., et al. 2009). ESX-3 maintains iron levels in these two mycobacteria by the secretion of mycobactin, a siderophore responsible for binding iron that can re-enter the cell through conserved, non-esx pathways (Siegrist M. S., et al. 2009). Growth defects are observed in *M. tuberculosis* lacking *esx-3*, which can be rescued by supplementation of mycobactin J, further supporting the role of mycobactins and ESX-3 in iron homeostasis (Tufariello, J. M., et al. 2016).
ESX-5

This T7SS is only found in slow-growing mycobacteria, which includes many pathogens. ESX-5 was shown to be required for growth of mycobacteria including *M. marinum* and *M. bovis* by modulating cell wall permeability (Ates, L. S., et al. 2015). ESX-5 was not essential for growth of these bacteria when their membranes were artificially made more porous by expression of the porin MspA, demonstrating the direct effect of ESX-5 on the mycomembrane (Bottai, D., et al., 2012). ESX-5 may play additional roles in nutrient homeostasis due to its upregulation in phosphate-depleted conditions (Elliott, S. R. and Tischler, AD. 2017). Key structural insights about T7SS have also been obtained from the resolved structure of ESX-5 at the mycobacterial inner membrane (Beckham, K. S. H., et al. 2017).

In addition to secreting WXG substrates, it's estimated that roughly 95% of PPE proteins are secreted via ESX-5 due to their interaction with the known ESX-5 chaperone EspG (Abdallah, A. M., et al., 2009). PE and PPE proteins are non-WXG T7SS substrates that contain a conserved N-terminal proline (P) – gulatamate (E) (PE) or PPE domain (Reviewed in Brennan, M., 2017). Proteins in these families are necessary for different cellular processes including maintaining cell-wall integrity (Brennan, M., 2017). PPE10, specifically, was shown to directly mediate maintenance of cell wall integrity in an *esx*-5-dependent manner (Ates, L. S., et al. 2016).

*The focus of this thesis is characterization of esx-4 and its predicted substrates, which will be described in detail below. To describe the defined functions of ESX-4 it is important to first review DCT, the first described process that requires a functional ESX-4 apparatus (Gray, TA., et al. 2016).*
**DCT: Links to T7SS**

Mycobacteria can mediate horizontal gene transfer (HGT) by a novel conjugative process called DCT and this process is reliant upon T7SS. Bacteria can acquire new genetic information, which can influence their evolution and differentiation, by HGT (transformation, transduction, or conjugation) (Sun, D., 2018). The most predominate form of bacterial conjugation is mediated by plasmids. Plasmid-mediated conjugation occurs from a site on the plasmid called the origin of transfer (oriT) (De La Cruz, F., et al., 2009). Single-stranded DNA is then transferred from a donor to a recipient cell through a channel created by a T4SS (Reviewed in Ilangovan, et al., 2015). DCT in mycobacteria occurs by a completely different mechanism (Derbyshire, K. M. & Gray, T. A., 2014). Most notably, it involves transfer of chromosomal DNA, not plasmid DNA, from donor to recipient cells, it does not require oriT sites or a type IV secretion apparatus, and results in transconjugants with mosaic genomes (Derbyshire, K. M. & Gray, T. A., 2014).

DCT has been described in *M. smegmatis* and *Mycobacterium canetti*, a progenitor of *M. tuberculosis* capable of causing tuberculosis in humans (Boritsch, E. C. et al., 2016). While there may only be experimental evidence of DCT in *M. smegmatis* and *M. canetti*, sequence analyses have identified events consistent with HGT between other mycobacteria. These events include gene clustering in *M. tuberculosis* and the presence of mosaic chromosomes in clinical isolates of *M. abscessus* (Sapriel, G., et al., 2016). This unique form of HGT requires type VII secretion systems ESX-1 and ESX-4. Recipient strains of *M. smegmatis* lacking esx-1 are defective in transfer (Coros, A. et al., 2008, Gray, T. A. et al., 2016). However, donor strains lacking esx-1 exhibit a hyperconjugative phenotype, in which the transfer of chromosomal DNA is more efficient than that seen between two wild-type parents (Flint, JL., et al., 2004)

**ESX-4**

ESX-4 is the simplest of the T7SS described in mycobacteria; as it contains a core locus of only 7 genes (Figure 1.3A). This secretion system does not include EccA, EccE, or EspG
components, which removes a layer of complexity from understanding this system as a whole (Figure 1.3A). As previously mentioned, this secretion system is considered the universally conserved progenitor of other esx systems. The first cellular process that requires esx-4 was DCT (Gray, T. A. et al., 2016). While an esx-4 null donor is a competent conjugal partner, recipients harboring mutations in any of the 7 genes of esx-4 are unable to participate in transfer (Gray, T. A. et al., 2016) (Figure 1.3B). Individual esx-4 mutations in recipients were individually complemented and resulted in recovery of transfer (Gray, TA., et al., 2016). We do not know the precise role of ESX-4 during conjugation, but it has provided us with a functional assay that we can use to study this T7SS.

**ESX-4 and Pathogenesis**

In *M. abscessus* esx-4 is required for survival in macrophages and is, thus, required for virulence (Laencina, L., et al., 2018). This discovery is especially important for our field of research because *M. abscessus* does not encode esx-1, which implies esx-4 could play a major role in pathogenesis in other mycobacteria. In support of this, one study has shown that patients infected with *M. tuberculosis* contained antibodies against EsxU and EsxT, the WXG100 pair of proteins encoded within esx-4 (Pandey, H., et al., 2018).

**Regulation of esx-4**

When trying to determine the regulator of esx-4, we noted that several microarray analyses identified esxU and esxT as being upregulated in *M. tuberculosis* constitutively expressing an alternative sigma factor, sigM (Agarwal, N., et al., 2007, Rustad, T. R., et al., 2014). Sigma factors are proteins that recruit RNA polymerase (RNAP) to the promoter region of genes and are required for transcription to occur (Paget, M. S., 2015). Sigma factors are often referred to as “house-keeping” or “alternative” sigma factors due to differing sequence preferences. Levels of sigM expression were found to be minimally elevated in response to heat shock and during stationary phase growth in *M. tuberculosis*, but SigM was not found to be required for virulence
(Agarwal, N., et al., 2007). In *M. smegmatis*, *sigM* is one of 26 putatively defined sigma factors and our laboratory became interested in whether DCT could lead to the activation of SigM.

We first investigated whether *sigM* (*MKD8_6931*) is required for DCT. Deletion of *sigM* in donor cells did not lead to defects in DCT, but recipients lacking *sigM* were completely defective in transfer (Figure 1.4A). DCT was restored when *MKD8_6931* or *sigM* orthologs *Rv3911* and *Mab_4938* were complemented in *sigM* null recipients (Figure 1.4A, unpublished observations). We reasoned that SigM likely regulates other genes required for DCT and overexpressed *sigM* in recipient cells to define the SigM regulon via RNA-seq analysis.

A 21-gene SigM regulon was identified and includes the entire *esx-4* locus as well as other genes that encode proteins with T7SS substrate features (Table 1). All transcripts in the SigM regulon are preceded by a conserved SigM-binding motif (Figure 1.4C) (Clark, R. R., et al., 2018). This binding motif can be found upstream of orthologs of *MKD8_1560* and *MKD8_6925*, the only conserved non-*esx-4* genes in the SigM regulon. *MKD8_1560* and *MKD8_6925* were also identified as SigM-regulated genes in *M. tuberculosis* (Agarwal, N., et al., 2007, Rustad, T. R., et al., 2014). To explore the functions of SigM-regulated genes we used DCT in *M. smegmatis* as a functional assay.

Several non-*esx-4* core genes, defined as SigM-regulated transcripts not found between *MKD8_1533* and *MKD8_1539*, were knocked out in recipients and were found to be required for DCT (Table 1.1). Notably, *MKD8_1560* is required for DCT and contains a WXG100 domain as well as a T7SS signal sequence within its C-terminus (Figure 2.1C). The features of MKD8_1560, added to the fact that it is encoded near *esx-4* (~19.5 kilobases downstream of *MKD8_1539*), gives support to the hypothesis that MKD8_1560 is secreted via ESX-4 (Kapopoulou A., et al., 2011). In addition to *MKD8_1560*, *MKD8_1686* was required for transfer and also contains a WXG motif that may be relevant to the secretion of MKD8_1686. Other
SigM-regulated genes encode proteins that have T7SS substrate-like features are discussed at length later in this thesis (Table 1.1).

**ESX-4 and Colony Morphology**

When determining ties between SigM and esx-4 expression, a change in colony morphology was observed in recipients overexpressing sigM. Recipients normally exhibit a smooth, shiny, and round colony morphology, but a rough, dry and irregular shaped morphotype is observed when sigM is constitutively expressed (Figure 1.4B) (Clark R. R., et al., 2018). However, recipients maintain their WT colony morphology when sigM is constitutively expressed in esx-4 mutant recipients (Clark R. R., et al., 2018). This implies that ESX-4 is secreting a substrate that affects the cell-wall composition of *M. smegmatis*. This phenotype is also observed when sigM orthologs from *M. tuberculosis* and *M. abscessus* are overexpressed in recipients lacking sigM, which suggests regulation and function(s) of esx-4 in *M. smegmatis* are conserved (Figure 1.4B). In addition to inducing a smooth to rough morphotype switch, overexpression of sigM allows recipients to form a biofilm (Clark, R. R., et al., 2018). Biofilm formation is critical to the pathogenesis of many bacteria, including those within the mycobacteria genus. Formation of biofilms is critical for persistence of *M. tuberculosis* and makes treatment options for tuberculosis more complex. These data provide us with several other phenotypes that could potentially be used for studying the function of ESX-4, but only DCT was used as a functional assay in this thesis.
Figure 1.1. Secretion systems are used to transport proteins across the bacterial membrane(s). A. The General Secretion (Sec) system localizes to and transport proteins across the cytoplasmic membrane of bacteria. SecB recognizes substrates and transports them to the IM where they are secreted through Sec. SecB and SecA facilitate the recognition
and passage of substrates through SecYEG, the channel through which substrates are translocated. Image obtained from Green, ER., & Mescas, J., 2016. **B. The Twin Arginine Secretion System (Tat) transports proteins across the cytoplasmic membrane of bacteria.** TatB and TatC bind to substrates that are then secreted through TatA. Image obtained from Green, E. R., & Mescas, J., 2016. **C. Gram-negative bacteria use different secretion systems to export proteins to the extracellular area or into host cells.** Image obtained from Depluverez, S., et al., 2016.
Figure 1.2. Mycobacteria use T7SS to export proteins in a Sec-/Tat-independent manner.

Figure 1.3 The core esx-4 locus only contains 7 genes and all are required for DCT. A.

Figure 1.4. SigM is required by recipients for DCT and causes an esx-4-dependent switch in colony morphology when overexpressed. A. Recipients lacking sigM are unable to participate in transfer. This phenotype is rescued upon reintroduction of sigM. The sigM ortholog in M. tuberculosis can also rescue transfer in recipients lacking sigM. Image obtained from Clark, R. R., et al., 2018. B. Constitutive expression of SigM, or its orthologs from M. tuberculosis (Rv3911) and M. abscessus (Mab_4938), causes a donor-like rough colony morphology in normally smooth, mucoid recipient cells (unpublished observations). C. SigM contacts a conserved binding motif, represented by largest letters in diagram, upstream of each of its regulated transcripts. Image obtained from Clark, R. R., et al., 2018. All SigM-regulated transcripts are listed in Table 1.
Table 1.1. SigM regulon genes and their predicted functions. Shaded genes belong to the esx-4 core locus. Conservation was determined by presence of orthologs in *M. tuberculosis*, *M. abscessus*, and *M. marinum* (Clark, R. R., et al. 2018).

<table>
<thead>
<tr>
<th>Gene</th>
<th>No. aa</th>
<th>Conserved?</th>
<th>Req’d for DCT?</th>
<th>Annotation and predicted function</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKD8_0914</td>
<td>290</td>
<td>×</td>
<td>-</td>
<td>DUF559 family, endonuclease?</td>
</tr>
<tr>
<td>MKD8_1535</td>
<td>441</td>
<td>✓</td>
<td>+</td>
<td>EccD4, pore protein</td>
</tr>
<tr>
<td>MKD8_1533</td>
<td>436</td>
<td>✓</td>
<td>+</td>
<td>MycP4, mycosin protease</td>
</tr>
<tr>
<td>MKD8_1536</td>
<td>1,211</td>
<td>✓</td>
<td>+</td>
<td>EccC4, ATPase</td>
</tr>
<tr>
<td>MKD8_1537</td>
<td>368</td>
<td>✓</td>
<td>+</td>
<td>Rv3446c family, transmembrane protein?</td>
</tr>
<tr>
<td>MKD8_1538</td>
<td>103</td>
<td>✓</td>
<td>+</td>
<td>EsxU, ESX-4 secreted substrate</td>
</tr>
<tr>
<td>MKD8_1539</td>
<td>98</td>
<td>✓</td>
<td>+</td>
<td>EsxT ESX-4 secreted substrate</td>
</tr>
<tr>
<td>MKD8_15690</td>
<td>283</td>
<td>×</td>
<td>-</td>
<td>DUF559 family</td>
</tr>
<tr>
<td>MKD8_1560</td>
<td>102</td>
<td>✓</td>
<td>+</td>
<td>WXG100/espC family, ESX-4 secreted substrate</td>
</tr>
<tr>
<td>MKD8_15750</td>
<td>459</td>
<td>×</td>
<td>?</td>
<td>Hypothetical, WXG100 motif, potential secretion partner for 1560</td>
</tr>
<tr>
<td>MKD8_16105</td>
<td>11</td>
<td>×</td>
<td>-</td>
<td>Hypothetical smORF</td>
</tr>
<tr>
<td>MKD8_1687</td>
<td>96</td>
<td>×</td>
<td>?</td>
<td>WXG100/espC family, ESX-4 secreted substrate</td>
</tr>
<tr>
<td>MKD8_1686</td>
<td>355</td>
<td>×</td>
<td>+</td>
<td>NlpC-family, endopeptidase activity</td>
</tr>
<tr>
<td>MKD8_1685</td>
<td>55</td>
<td>×</td>
<td>-</td>
<td>DUF2580 family</td>
</tr>
<tr>
<td>MKD8_21385</td>
<td>24</td>
<td>×</td>
<td>?</td>
<td>Hypothetical smORF</td>
</tr>
<tr>
<td>MKD8_2151</td>
<td>453</td>
<td>×</td>
<td>?</td>
<td>Transposase IstA protein</td>
</tr>
<tr>
<td>MKD8_21370</td>
<td>56</td>
<td>×</td>
<td>?</td>
<td>Transposition helper protein</td>
</tr>
<tr>
<td>MKD8_53960</td>
<td>89</td>
<td>×</td>
<td>?</td>
<td>Pseudogene</td>
</tr>
<tr>
<td>MKD8_53950</td>
<td>146</td>
<td>×</td>
<td>-</td>
<td>DUF559 family</td>
</tr>
<tr>
<td>MKD8_5844</td>
<td>299</td>
<td>×</td>
<td>-</td>
<td>DUF559 family, DNA binding and endonuclease protein?</td>
</tr>
<tr>
<td>MKD8_6925</td>
<td>165</td>
<td>✓</td>
<td>?</td>
<td>Potential calcium binding protein</td>
</tr>
</tbody>
</table>
Chapter 2. Results

Introduction
The results in this chapter are divided into several sections. The SigM regulon, the features of SigM-regulated proteins, and the genes required for DCT are described in the first section of this chapter. Protein-protein interaction results are detailed in the second part of this chapter. Each results section is preceded by information highlighting the rationale behind the experiments performed.

As previously mentioned, T7SS play a role in the virulence of several pathogenic mycobacteria. This makes understanding these secretion systems a priority as they could provide a new target for therapeutic intervention. To describe the biological relevance of a secretion system, one must characterize the function of its parts. By determining a role for T7SS substrates, we are able to better determine the biological purposes of these secretion systems. For reasons described below, there is reason to hypothesize that the SigM regulon encodes several, if not many, substrates of ESX-4. Therefore, we can study these non-esx-4 encoded proteins to further understand ESX-4 itself. We believe ESX-4 is a good model for studying T7SS because it is ancestral, small in comparison to other ESX apparatuses, and is universally conserved across mycobacteria.

Since ESX-4 plays a role in DCT, we decided to use DCT as a functional assay to also characterize the SigM regulon because it is likely that ESX-4 substrates, rather than the ESX-4 apparatus itself, are playing a role during DCT. Several genes in the SigM regulon were characterized prior to this thesis, but this analysis was not complete. Therefore, the goals of this thesis include determining which remaining SigM regulon genes are important for transfer, as well as establishing whether proteins encoded within the SigM regulon interact that could be relevant to their function or their secretion via ESX-4.
Part I: Initial screening of SigM regulon mutants for a role in DCT

An initial screen of recipient SigM-regulon mutants led to the identification of 8 genes required for DCT (Clark, R. R., et al., 2018). Of those genes identified, six are esx-4 core locus genes (MKD8_1533-1539) (Table 1.1). Deletion of any core esx-4 locus gene abolished transfer, a phenotype that was rescued when these mutations were complemented with the wild-type gene (Clark, R. R., et al., 2018). It is likely that ESX-4 secretes several substrates involved in DCT and the requirement for all core esx-4 genes for DCT demonstrates how important ESX-4 secretion is for transfer. These substrates could have cell-wall remodeling properties, may contact and activate donor cells, or might be enzymes with nuclease or hydrolase activity that could facilitate DCT. Below I will discuss the key findings from this published work before describing the genes most relevant to this thesis.

Initial screening of transfer-defective recipients identified MKD8_1560 and MKD8_1686 as two non-esx-4 core locus genes required for DCT. MKD8_1560 is situated near (~19.5 kb downstream) esx-4 on the chromosome and encodes a small (102 aa) protein with a predicted helix-turn-helix turn similar to that caused by a WXG motif (Figure 2.1). MKD8_1560 also contains a C-terminal signal sequence (YAVAD), which matches the canonical YXXXD/E signal sequence found on known T7SS substrates (Figure 2.1C/D). A function for MKD8_1560 has not yet been identified, but the requirement for MKD8_1560 during DCT strengthens our hypothesis that ESX-4 is secreting substrates important for transfer and provides a starting point for further defining the role of MKD8_1560.

MKD8_1686 is predicted to be a part of the NPL/P60-family of proteins (Table 1.1, Kapopoulou A., et al., 2011, Kelley, LA., et al., 2015). This family of proteins exhibit cell wall hydrolase activity, which involves the enzymatic breakdown of peptidoglycan (PG) in the bacterial membrane mediated by a conserved N-terminal catalytic cysteine residue (Anantharaman, V. and Aravind, L., 2003). Phyre2 protein modeling software suggests, with high confidence, that MKD8_1686 contains a hydrolase domain with a single cysteine residue (Figure 2.2C) (Kelley,
L. A., et al., 2015). While this hydrolase domain is not in the N-terminal region of MKD8_1686, the cysteine conserved in MKD8_1686 is near the N-terminal end of this predicted hydrolase domain (Figure 2.2C). Thus, a reasonable hypothesis is that MKD8_1686 could cleave either donor or recipient PG to facilitate transfer of substrates or DNA between conjugal partners.

**Described below are the SigM-regulated genes characterized during this thesis.** I am focused on 8 proteins, some of which are likely secreted via ESX-4 based on structural predictions and others that have no distinct structural features that aid in functional predictions. Regardless, DCT and 2-hybrid analysis are excellent assays to begin dissecting a biological role for these proteins.

**MKD8_1560 and MKD8_15750: A Tale of Two Genes**

While we already know that MKD8_1560 is required for DCT, it is co-transcribed with MKD8_15750, which was not previously investigated (Figure 2.1A). According to protein modeling software, MKD8_15750 contains a high-confidence EsxB-like fold from amino acid residues 51-127 out of its 459 aa structure (Kelley, LA., et al., 2015). Furthermore, MKD8_15750 contains a PE (proline-glutamic acid) motif directly adjacent to its start codon, which is another hallmark feature of known T7SS substrates, found primarily in substrates secreted via ESX-5 (Figure 2.1B) (Abdallah, M. A., et al., 2009). Due to its co-transcription with MKD8_1560, we hypothesize that MKD8_15750 is also involved in DCT. Furthermore, because MKD8_15750 contains T7SS substrate-like features, we hypothesize that MKD8_15750 will exhibit a co-dependent secretion through ESX-4 with MKD8_1560 and, thus, will interact.

**MKD8_1687, MKD8_1686, and MKD8_1685: Three’s Company**

MKD8_1687 is the first gene in an operon that also contains MKD8_1686 and MKD8_1685 (Figure 2.2A). All three genes are transcribed from the SigM promoter, located upstream of MKD8_1687. While MKD8_1686 was found to be required for DCT, MKD8_1685 was not required and MKD8_1687 was not characterized. MKD8_1685 encodes a small protein (55 aa)
with no known function. This leaves \textit{MKD8\_1687} as the final gene in this operon not previously characterized by Clark et al. 2018.

\textit{MKD8\_1687} encodes a 96 aa protein that contains a C-terminal T7SS signal sequence that likely mediates its secretion through ESX-4 (Figure 2.2B). We hypothesize that MKD8\_1687 could serve as the secretion partner for either MKD8\_1686, MKD8\_1685, or both. We expect to determine direct contact between MKD8\_1687 and MKD8\_1686 for secretion because they are encoded together and, unlike MKD8\_1685, they both contain common features of T7SS substrate. For this reason, we also hypothesize that MKD8\_1687 will be required for DCT because of its potential role in the secretion of MKD8\_1686.

\textbf{MKD8\_6925: A possible calcium binding protein?}

Despite having no assigned function, \textit{MKD8\_6925} was identified as a SigM-regulated transcript in \textit{M. tuberculosis} via microarray and in \textit{M. smegmatis} via RNAsequencing, highlighting the conservation of the SigM regulon (Agarwal, N., et al., 2006, Rustad, T. R., et al., 2014). Orthologs of MKD8\_6925 are conserved in other pathogenic mycobacteria besides \textit{M. tuberculosis} and are preceded by a conserved SigM binding motif (Figure 2.3A). This makes studying genes like \textit{MKD8\_6925} very valuable in a model system such as \textit{M. smegmatis} because we can make predictions about the function of its orthologs in pathogenic mycobacteria based on its function in our model.

We do not have a prediction for what proteins MKD8\_6925 could contact for secretion, but we rationalized that other SigM regulon proteins are potential candidates, so it was screened against several possible secretion partners using M-PFC.

MKD8\_6925 contains a surprisingly large number (8) of canonical EF-hand motifs, defined as runs of aa including Asp (D), X, D, X, D, Glu (G), X, X, D, where X represents any amino acid (Figure 2.3B). EF-hand motifs bind calcium and are found in a wide range of eukaryotic proteins.
involved in calcium sensing and signal transduction (Reviewed in Lewit-Bentley, A. and Rety, S., 2000). Furthermore, Phyre2 predicts with high confidence that MKD8_6925 has a thrombospondin-like structure (Figure 2.3C) (Kelley, L. A., et al., 2015). Thrombospondins are eukaryotic calcium-binding proteins responsible for platelet recruitment and inflammatory responses during wound healing (Reviewed in Adams, J., et al., 2011).

While there is extensive literature about the functions of calcium-binding proteins in eukaryotes, the same is not true for prokaryotes. There is evidence to suggest bacteria do indeed encode calcium-binding proteins and, similar to eukaryotes, may even use calcium as a signaling molecule, but this is still a poorly developed field of research (Reviewed in Dominguez, DC., 2004). Uncovering a role for MKD8_6925 in DCT would strengthen the hypothesis that calcium-binding proteins may play key roles during a range of cellular processes in bacteria.

**MKD8_16105, MKD8_21385, MKD8_53960: Genes of Unknown Function**

Several SigM-regulated genes studied in this thesis are too small to make functional predictions from their encoded protein products. MKD8_16105 and MKD8_21385 are MKD8_53960 are all recipient-specific transcripts regulated by SigM and encode small proteins, 11 and 24 aa respectively, with unknown functions. MKD8_53960 is annotated as a pseudogene but was included this analysis to ensure it does not play a role during DCT.

**Genes not assayed**

MKD8_2151 is a predicted transposase and MKD8_21370 is predicted to act as a transposition helper protein. Transposases target DNA for transposition by binding inverted repeats found at either end of the transposon. Donor DNA segments transferred to the recipient during DCT are not flanked by inverted repeats, which implies DCT is not mediated by transposition of DNA. Therefore, we do not hypothesize that genes MKD8_2151 and MKD8_21370 are required for DCT and, thus, have excluded them from the following analysis.
Section 1 Results: The SigM Regulon and DCT

Generating and Verifying SigM Regulon Knockouts

PCR primers (A, B, C, D) were designed to amplify arms homologous to the 5’ and 3’ flanking regions of the gene of interest via PCR (Figure 2.4A). Primers C and D were designed to include a site homologous to the zeocin resistance gene (zeo') in their sequence (UPS). This allows for the incorporation of UPS sites into the amplified arm products. These UPS sites allow for fusion to occur between the arm products and zeo' (Figure 2.4A/B). These linear substrates were electroporated into recombineering-competent recipients, to allow precise replacement of the target gene (Figure 2.4C). Transformants were grown on TSA-containing zeocin to select for strains that have undergone homologous recombination, replacing the natively encoded gene of interest with that encoding zeo'.

Mutants were verified by screening strains via PCR for the precise replacement of the target gene with zeo' (Figure 2.4D). A PCR was performed using primers A and D. Primers A and D anneal to the 5’ and 3’ regions flanking the gene of interest or, in a mutant strain, zeo'. The sizes of the products generated in WT and mutant recipients by A and D identified precise replacements. An additional PCR was performed, which uses a zeo' specific primer with D to further confirm the mutation.

Recipient mutants of MKD8_15750, MKD8_16105, MKD8_21385, MKD8_53960, and MKD8_6925 were generated. However, a recipient knockout of MKD8_1687 was not obtained. Difficulty knocking out MKD8_1687 could have occurred for several reasons. MKD8_1687 could be essential, which would not allow for the recovery of viable mutants. It is also possible that the recombineering cells used in an effort to obtain a mutant of MKD8_1687 were not optimally competent. Therefore, MKD8_1687 was only characterized here using M-PFC, which is described later in this chapter.

SigM Regulated Genes Not Required for DCT
To determine the effects of deleting SigM-regulated genes on DCT, WT donors were co-cultured in equal volume on TSA with recipients lacking individual SigM-regulated genes. Co-cultures were recovered from TSA using a sterile loop, serially diluted in TSBT, and plated on selective media to recover transconjugants as well as parental donors and recipients.

Recipients lacking MKD8_21385, MKD8_16105, MKD8_53960, or MKD8_15750 were competent mating partners (Figure 2.5A). Mutations in the above recipient genes did not cause a decrease in transfer frequencies between donors and recipients. Therefore, these genes do not play a role during DCT.

**SigM Regulated Genes Required for Transfer**

Out of the genes assayed, only recipients lacking a functional copy of MKD8_6925 were defective for transfer (Figure 2.5B). To ensure this loss of transfer was MKD8_6925 specific, a plasmid containing a WT copy of MKD8_6925 was introduced into ΔMKD8_6925 recipients (Figure 2.5B). Re-introduction of MKD8_6925 into mutants restored transfer between donors and recipients to wild-type levels, identifying MKD8_6925 as another SigM-regulated gene required for DCT. Due to the high abundance of EF-hand calcium binding motifs in MKD8_6925, we hypothesized calcium may be important for DCT.

**The Role of Calcium During Growth and DCT**

To dissect the potential role(s) of MKD8_6925 and calcium during DCT, we first cultured recipients in Sauton's media with or without ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA), a chelator with a high affinity for calcium ions. We first determined whether EGTA caused generalized growth defects caused by lack of calcium in the growth medium. WT recipients were inoculated into Sauton's medium alone or Sauton's medium containing 1 mM EGTA, at an optical density (OD) of 0.05. Cultures were grown at 37°C for several days. Aliquots were taken from each culture at various time points and a growth curve was generated based on OD values.
When grown in Sauton’s with 1 mM EGTA, recipients exhibited a significant growth defect during log-phase growth (Figure 2.6). However, after 50 hours WT recipients cultured in Sauton’s with EGTA grew to roughly the same OD as their counterpart grown in Sauton’s with no added EGTA, which implies recipients only exhibit growth defects during log-phase and these growth defects should not influence the outcome of DCT assays.

Following growth curve analysis, we next determined whether calcium is important for DCT. WT donors and recipients were co-cultured on TSA and TSA containing 1 mM EGTA. In contrast to the positive control (no EGTA), transconjugants were not detected in the presence of EGTA (Figure 2.7A). Importantly, recovery of donor and recipient cells from TSA containing EGTA was not lower than that from TSA alone. Unaltered parental recovery, along with our growth curve analysis, implies this defect in transfer is not caused by growth inhibition. Addition of exogenous calcium to TSA during mating did not enhance WT transfer efficiency (Figure 2.7B). Exogenous calcium was also not enough to rescue transfer between WT donors and recipients lacking MKD8_6925 (Figure 2.8). To determine if the defects in transfer described above are calcium specific, WT donors and recipients were co-cultured on TSA plates containing 1 mM N,N,N′,N′-Tetrakis(2-pyridinylmethyl)-1,2-ethanediamine (TPEN), a different ion chelator with an affinity for zinc (Figure 2.7B). The addition of TPEN had no effect on WT transfer frequencies, implying that the phenotype we have observed is calcium specific.

The Effects of EGTA on Recipient Gene Expression During DCT

To further investigate the role of calcium during DCT, quantitative real-time PCR (qRT-PCR) was used to quantitate gene expression levels during calcium-depleted co-culture conditions. WT donors and recipients were co-cultured on TSA as well as TSA containing 1 mM EGTA. A monoculture containing only recipients was also cultured to serve as a control for expression; SigM-regulated transcripts are not expressed in monocultured recipients. MKD8 monoculture serves as a baseline for SigM regulon expression that our experimental conditions (co-culture with or without EGTA) can be compared to. Cultures were collected from TSA using a sterile
loop and RNA was harvested from each sample. cDNA was generated from each sample and qRT-PCR was performed. We first examined expression levels of SigM-regulated transcripts because they are highly expressed in co-culture but not in monoculture.

Expression of *esxUT* was elevated roughly 500-fold (Figure 2.9A). However, in the presence of EGTA, expression of *esxUT* did not increase; i.e. *esxUT* failed to respond to co-culture (Figure 2.9A). A similar trend in expression was observed for *MKD8_5844* and *MKD8_6925*. In co-culture, *MKD8_5844* expression levels were elevated, on average, 55-fold when compared to MKD8 monoculture but were not elevated during co-culture in the presence of EGTA (Figure 2.9A). Expression of *MKD8_6925* was also increased during WT co-culture (~35-fold) and was not induced during co-culture on TSA containing EGTA (Figure 2.9A). This suggests that EGTA inhibits SigM sensing of cell-cell contact and thus prevents induction of its regulon.

To determine if this reduction in expression is SigM specific, expression levels of *MKD8_5565* were measured in standard and calcium-depleted co-culture conditions. *MKD8_5565* is a highly expressed donor-specific transcript during DCT and its transcription is independent of SigM. Thus, determining *MKD8_5565* transcript levels would not only elucidate whether calcium is important for SigM signaling, but would also determine if a lack of gene expression of normally expressed transcripts during DCT with EGTA is a donor- or recipient-specific phenotype. qRT-PCR primers were designed to contain single nucleotide polymorphisms (SNPs) that were able to differentiate between donor and recipient *MKD8_5565* transcript levels.

Our results indicate that *MKD8_5565* expression is significantly elevated in donor cells during co-culture on TSA and when EGTA is present during co-culture, the induction of donor-specific *MKD8_5565* is still observed (Figure 2.9B). These results point to a recipient-specific requirement for calcium during DCT and that is sensed by anti-SigM for release of SigM or utilized by SigM for induction of its regulon.

**Section II: Protein-Protein Interactions within the SigM Regulon**
SigM-regulon-encoded proteins were screened by mycobacterial protein fragment complementation (M-PFC) to determine protein interactions that might provide insight on their function. M-PFC is a variation of 2-hybrid analysis and allows for the identification of protein-protein linkages using mycobacteria as a background for analysis instead of yeast or *E. coli*, which creates a more native environment for studying interactions between mycobacterial proteins.

In general, if two proteins X and Y are co-transcribed, and protein X has a T7SS signal sequence, it is very likely that proteins X and Y are co-secreted in a contact-dependent manner through a T7SS. Co-dependent secretion through ESX-1 has been demonstrated for EsxA, which requires EsxB, as well as other ESX-associated proteins (Esp), for secretion (Fortune, S. M., et al., 2005). Based on the presence of T7SS substrate-like structural motifs, we were able to identify proteins within the SigM regulon that may form complexes required for their secretion via ESX-4 (Table 2.1). We hypothesized that like EsxA and EsxB, EsxU and EsxT would contact one another for the reasons described below. We also looked for complex formation between SigM-regulated proteins listed in Table 2.1 and EsxU or EsxT.

Beyond substrate-substrate complexes, substrates have been found to also contact and modulate the function of core ESX secretion machinery. EccC1 of ESX-1 forms an interaction with EsxB, but can also interact with EsxB alone (Callahan, B., et al., 2010, Stanley, S. A., et al., 2003). The physical interaction between EsxB and EccC1 results in the activation of the ATPase domains of EccC1 (Rosenberg, O. S., et al., 2015). Physical interactions between T7SS substrates and ESX secretion machinery have been described for ESX-1; this is the first time for testing ESX-4. Therefore, we used M-PFC to determine if any candidate SigM-regulated proteins contact MKD8_1537, which is described below (Table 2.1).
MKD8_1537: A possible chaperone for ESX-4 substrates?

MKD8_1537 does not have an assigned function but is encoded in the esx-4 core locus and is predicted to localize to the IM with other ESX-4 machinery components (Figure 1.3A). Other ESX systems contain structural proteins that form physical interactions with substrates, and it is possible that MKD8_1537 may also form physical connections with SigM-regulated proteins that are secreted through ESX-4. We do not have a predicted function for MKD8_1537, but through M-PFC we can determine whether MKD8_1537 potentially acts as a chaperone for ESX-4, which would involve binding and potentially processing of ESX-4 substrates by MKD8_1537 during secretion.

Mycobacterial Protein Fragment Complementation (M-PFC)

M-PFC, as previously mentioned, is a form of 2-hybrid assay that can determine whether two proteins are in complex and we used to address SigM-regulon protein-protein interactions. This assay relies on pUAB300 and pUAB400, which encode the F[1] and F[2,3] domains of murine dihydrofolate reductase (m-DHFR) respectively, followed by a flexible glycine linker (Figure 2.10) (Singh, A., et al., 2006). Genes are cloned downstream of this linker and the resulting plasmid encodes a fusion of the m-DHFR domain to the protein of interest (Figure 2.10B). If two physically interacting proteins are tagged, the two flexible halves of m-DHFR will assemble to reconstitute a functional m-DHFR enzyme (Figure 2.11) (Singh, A., et al., 2006). DHFR is a modular enzyme responsible for the production of nucleotides and, in bacteria (b-DHFR), is a target for the antibiotic trimethoprim (TRIM) (Singh, A., et al., 2006). This assay utilizes trimethoprim to inhibit DHFR (b-DHFR), which results in cell death. However, bacteria containing interacting proteins that are tagged with m-DHFR can survive in the presence of TRIM because TRIM has a far lower affinity for m-DHFR than b-DHFR. Thus, if prey and bait proteins associate, they facilitate reconstitute of m-DHFR and confer trim resistance to the host bacterium.
The positive control for this assay demonstrates the ability of this system to effectively identify protein interactions. pUAB100 and pUAB200 are relatives of pUAB300 and pUAB400 that encode separate domains of the modular yeast transcriptional activator *gcn4* (Singh, A., et al., 2006). These two domains contact one another to activate transcription in yeast and also promote the association of m-DHFR to confer trim resistance.

**Section 2 Results: Protein-Protein Interactions within the SigM Regulon**

SigM regulated genes were cloned into pUAB300 and pUAB400 and confirmed via DNA sequencing. Plasmids were co-electroporated into donor (mc²155) cells, which have been used in our lab for this assay before (Callahan, B., et al., 2009). Transformants were selected on TSA containing hygromycin (hyg) and kanamycin (kan) to isolate donor cells that successfully acquired both pUAB300 and pUAB400 derivatives. Double transformants were grown overnight in TSBT and diluted into fresh TSBT to an OD of 0.5-0.7. Cells were incubated at 37°C until an OD of 1.0-1.3 was reached. Cultures were serially diluted and the neat, -1, -2, -3, and -4 dilutions were spotted on TSA to determine relative viability and on 7H10 media containing 25 ug/mL TRIM to identify bacteria containing interacting proteins. We tested varying concentrations of TRIM for this assay and decided to use 25 ug/mL TRIM. This concentration of TRIM prevented growth of donor cells expressing empty vector, while allowing growth of the positive control.

**Positive M-PFC Results**

When cloned into either pUAB300 and pUAB400 and co-transformed into donor cells, *esxU* and *esxT* permitted growth on TRIM consistent with this pair of proteins interacting (Table 2.1, Figure 2.12). We reasoned that because related substrates EsxA and EsxB form interactions with Esp proteins, EsxU and EsxT might also act as secretion partners for other SigM-regulated proteins. However, when individually co-transformed with pUAB vectors containing non-*esx*-4
encoded SigM-regulated proteins, neither EsxU or EsxT formed complexes with other SigM regulon proteins (Table 2.1).

In addition, MKD8_1685 and MKD8_1686 were also found to form a physical interaction that allowed for reconstitution of m-DHFR and growth of co-transformed donor cells on media containing TRIM (Table 1.1, Figure 2.12).

**Negative M-PFC Results**

Surprisingly, when cloned into either pUAB300 or pUAB400, MKD8_1560 and MKD8_15750 did not form a complex as we had expected, nor did MKD8_1687 interact with MKD8_1686 or MKD8_1685 (Table 2.1). It is possible that MKD8_1560 and MKD8_1687 require additional substrates to form interactions with the proteins we predicted they would form complexes with. While we do not know what substrates might mediate an interaction between MKD8_1560 and MKD8_15750, it is possible that MKD8_1687 interacts with MKD8_1686 and MKD8_1685 together but not individually. To test this, *MKD8_1686 and MKD8_1685* were cloned together into one pUAB vector. This cloning strategy generated an m-DHFR tagged fusion of MKD8_1686 and MKD8_1685 that was screened against 1687 using M-PFC. This combination did not result in a positive M-PFC result, meaning that MKD8_1687 does not form a complex with MKD8_1686, MKD8_1685, or both together under our experimental conditions.

No interactions were observed between SigM-regulated proteins and MKD8_1537 (Table 2.1). MKD8_6925, which did not interact with MKD8_1537 (negative result), also did not form a complex with either EsxU or EsxT, the best candidate secretion partners for MKD8_6925. Therefore, the secretion substrates or partners for MKD8_1537 and MKD8_6925 still remain unknown.
Figure 2.1. *MKD8_1560* and *MKD8_15750* are co-transcribed in an operon and have T7 substrate features. A. Operon architecture of *MKD8_1560* and *MKD8_15750* adapted from annotated *M. smegmatis* genome browser (Kapopoulou, A., et al., 2011). B. The amino acid sequence of *MKD8_15750* reveals an N-terminal PE motif as well as a WXG motif. Complete amino acid sequence not shown (459 aa total). C. The genomic region upstream of *MKD8_1560* contains a SigM promoter that is conserved across orthologs of *MKD8_1560*. D. Structure of *MKD8_1560* as modeled by Phyre2 (Kelley, L. A., et al., 2015). The predicted structure of *MKD8_1560* contains a predicted non-WXG helix-turn-helix motif and a C-terminal T7SS signal sequence (YXXXD).
Figure 2.2 Features of MKD8_1687 operon. A. MKD8_1687 is the first gene in an antisense 3-gene operon. B. The amino acid sequence of MKD8_1687 reveals a C-terminal T7SS signal sequence. C. MKD8_1686 contains a predicted hydrolase domain (underlined) from residue 225 to residue 352 (Kelley, L. A., et al., 2015). This region contains a single cysteine residue, colored blue, which in related hydrolases is the catalytic residue (Anantharaman, V. and Aravind, L., 2003).
Figure 2.3. **MKD8_6925** is required by recipients for transfer and contains calcium-binding motifs. **A.** **MKD8_6925** and its SigM binding motif are conserved across *M. tuberculosis* (*Rv3906c*), *M. abscessus* (*Mab_4932c*), and *M. marinum* (*Mmar_5470*) (Clark, R.R., et al. 2018). **B.** **MKD8_6925** contains 8 EF-hand binding motifs (bold and underlined). **C.** Phyre2 modeled structure of **MKD8_6925** reveals a high-confidence thrombospondin-like structure between residues 5-164 (165 aa total) (Kelley, L. A., et al., 2015). Approximate location of all predicted calcium binding motifs within **MKD8_6925** sequence are labelled with an asterisk (*).
Figure 2.4. Genetic approach used to generate knockouts of SigM-regulated genes in recipients. A. A, B, C, and D primers were used to amplify two arm products (AB arm, CD arm) that flank the gene of interest (MKD8_15750 is used as an example). B and C primers have complementarity to a site in the zeocin resistance gene, which adds complementary sites to the gene proximal ends of each substrate. B. Arm products are fused to zeocin resistance gene block in synthesis via overlap expression (SOE) PCR reaction using A' and D' primers. Resulting fragment contains arms sewn to zeocin. C. The resulting knockout substrate is electroporated into recombineering proficient recipient cells, which results in homologous
recombination at the site of the gene being knocked out. D. Potential mutants, which have replaced *MKD8_15750*, are screened via PCR with primers able to differentiate mutants from WT recipients.
Figure 2.5. SigM-regulated genes and DCT. A. SigM-regulated genes not required in recipient cells for DCT. B. Recipient knockouts of MKD8_6925 were defective for DCT. Transfer was restored when a plasmid containing MKD8_6925 was introduced into MKD8_6925 null recipient. Each point represents a single biological replicate. Significance values were calculated using One-way ANOVA in GraphPad Prism V.8.3., ns = not significant (0.05 < p).
Figure 2.6. The effects of EGTA on growth of recipients in Sauton's media. Addition of 1 mM EGTA to Sauton’s medium delayed log-phase, but not stationary-phase growth. Each point represents the average of 3 biological replicates. Significance values were determined using an unpaired student’s T-test in GraphPad Prism V.8.3., *p<0.05.
Figure 2.7. The effect of EGTA on DCT between wild-type MKD10 (donor) and MKD8 (recipient). A. Addition of 1mM EGTA prevents successful transfer between MKD10 and MKD8. B. Addition of CaCl$_2$ or TPEN did not have an effect on WT transfer frequencies. Each point represents a single biological replicate.
Figure 2.8. Adding calcium to TSA during co-culture was not sufficient to rescue transfer between WT donors and recipients lacking MKD8_6925. Each point represents a single biological replicate. A transfer frequency of $1 \times 10^{-9}$ was used to report results from mating pairs that did not yield transconjugants.
Figure 2.9. EGTA affects expression of SigM-regulated genes, but not donor-specific transcripts during DCT. A. qRT-PCR was used to measure transcript levels of SigM-regulated genes. EGTA caused a decrease in expression of *esxUT*, *MSMEG_6925*, and *MSMEG_5844* in co-culture. B. EGTA did not reduce expression of *MKD8_5565* during co-culture. The *MKD8_5565* response is donor-specific and is
induced only during co-culture. Significance values were calculated using One-way ANOVA in GraphPad Prism V.8.3, *p<0.05, ***p<0.001.
Figure 2.10. Genes of interest were cloned into pUAB300 and pUAB400 for M-PFC. A. Plasmid maps of pUAB300 and pUAB400. pUAB300 and pUAB400 encode different fragments of the modular enzyme m-DHFR. These plasmids also encode hygromycin (Hyg') and kanamycin resistance (Kan'), respectively, for selection. Differences between these plasmids are bolded. B. A schematic of a gene fusion generated between *m-DHFR* and *esxU*. This fusion encodes an N-terminal m-DHFR tag (green) linked to EsxU (blue) through a flexible glycine linker (pink). Sizes do not correlate to gene or protein product size. Plasmid maps were generated using APE (Davis, W., 2003).
Figure 2.11. m-DHFR tagged proteins that interact allow domains F[1,2] and F[3] to re-constitute a functional enzyme. Successful reconstitution of m-DHFR, and therefore protein complexes between tagged proteins, is assessed by the growth of bacteria on media containing trimethoprim. EsxU, EsxT, and MKD8_1687 are used as an example. Figure was adapted from Singh, A., et al., 2006.
Table 2.1. Interactions between SigM-regulated proteins as shown through M-PFC. *MKD8_1686* interacts with *MKD8_1685* and *esxU* interacts with *esxT*. Green = PPI, Red = No PPI, Yellow = Not determined.
Figure 2.12. M-PFC analyses reveal several physical interactions between proteins encoded in the SigM regulon. **A.** EsxU and EsxT as well as MKD8_1685 and MKD8_1686 exhibited positive M-PFC results. All strains grow normally on TSA but exhibit differential growth on 7H10 + TRIM. Image is a representative result obtained from one out of three biological replicates.
Chapter 3: Discussion

**Mass Spectrometry Analysis**

One goal of our laboratory is to identify the substrates of ESX-4.

Mass spectrometry (MS) analysis has successfully been used to identify the secretome of ESX-1 and other T7SS and, while executing my thesis research, our laboratory established a project to define the ESX-4 secretome using MS (Champion, M. M., et al., 2014). Defining the substrates of ESX-4 will elucidate how what proteins are secreted through ESX-4 and will also determine if proteins encoded within the SigM regulon are secreted through ESX-4. There may be non-SigM regulon proteins secreted via ESX-4 and MS analysis will help us determine whether this is the case. I was involved in some of this work and we have obtained preliminary data from MS analysis that fills several gaps in the research presented in chapter 2 of this thesis. Therefore, it is important to discuss our preliminary data here before considering the implications of the results presented in this thesis.

There are three questions we want to answer using MS analyses: Are SigM regulon encoded proteins secreted, what are the non-SigM targeted secreted substrates of ESX-4, and what SigM proteins exhibit a co-dependent mechanism of secretion? To answer our first and second questions, \( \text{sigM} \) was overexpressed in a wild-type strain of MKD8 (recipient) and in an \( \text{esx-4} \) null strain of MKD8 (Figure 3.1A). Constitutive expression of \( \text{sigM} \) activates the expression of all ESX-4 structural components (only in WT MKD8) as well as the SigM regulon (in both WT and \( \text{esx-4} \) null MKD8). We can determine whether a particular SigM-induced protein is secreted from the cell by looking for its presence in the culture filtrate. Proteins that are secreted could be transported via ESX-4 or another unknown secretion system (Figure 3.1B). To determine if a protein is secreted in an ESX-4-dependent manner, we can measure protein levels in the
culture filtrate of esx-4 null recipients. If a protein secreted by WT MKD8 is still secreted in the absence of ESX-4, the protein in question is transported by a non-ESX-4 secretion system (Figure 3.1C). Analysis of culture filtrates produced by WT and esx-4 null MKD8 will also identify all substrates differentially secreted in the absence of ESX-4 whether they are encoded within or outside of the SigM regulon. Finally, we can compare protein levels in the culture filtrate of various SigM regulon null strains of MKD8 and WT MKD8 to determine what SigM-induced proteins are co-secreted. For example, if MKD8_1685 is secreted by WT MKD8 but not by MKD8_1686 null MKD8, we can conclude that MKD8_1685 is directly or indirectly dependent upon MKD8_1686 for secretion.

WT and esx-4 null recipients overexpressing sigM were grown in minimal medium to mid-log phase, which allowed for secretion of proteins into the growth medium with minimal cell lysis. Minimizing lysis reduces the release of cytoplasmic proteins from dead cells. Proteins released from the cytoplasm of dead bacteria can complicate our analysis by increasing the number of contaminating lytic proteins which can alter the results of MS, making it seem as though a non-secreted protein is secreted. An overabundance of lytic proteins may also increase background that could mask detection or influence the amount of secreted proteins that is detected. Culture filtrates and cell pellets were harvested and sent for MS analysis by our collaborator, Dr. Matthew Champion, at the University of Notre Dame.

Cell pellets and culture filtrates were processed and quantitated using liquid chromatography followed by tandem mass spectrometry (LC - MS/MS) as described in Bosserman et al., 2017. LC is used to separate the proteins within a sample into fractions that are less complex, which makes quantification by MS less complex (reviewed in Ramachandram, D., et al., 2016). Protein fractions generated by LC are fed directly into a mass spectrometer that uses ionization to fragment each compound into positively charged ions of different sizes. Following proteolytic
fragmentation usually by trypsin, the resulting peptides are passed through a deflector that contains a magnet. Positively charged ions are repelled by this magnet and the amount of deflection exhibited by each fragment passing through the deflector is dependent upon the mass of the fragment measured. The heavier an ion, the less it will travel when repelled by a magnet. A mass analyzer measures the amount of deflection exhibited by an ion and uses this measurement to calculate mass. This is why mass to charge ratios are reported from MS analysis, the charge is determined by the peptide ion fragments, and, when coupled with mass, can be used to determine the parent protein contained within the sample being analyzed.

Each protein in a sample should produce peptide fragments that are detected by MS. Each protein should produce a unique peptide m/z profile. However, some proteins can exhibit similar peptide m/z profiles, which can complicate analysis. Performing tandem MS negates this issue by using a second round of ionization that breaks down compounds into even smaller ions that are measured once again through deflection (Ramachandram, D., et al., 2016). Therefore, secondary MS further separates peptides into even smaller ions for detection, which facilitates defining the composition of a complex solution. The m/z profiles generated by each protein detected by MS can be compared to a database containing m/z ratios derived from known proteins. Using a database for comparison allows us to reconstruct and identify the proteins present in a sample based on the fragmentation and deflection profile produced by each protein present in the sample.

Proteins detected in the culture filtrate, except lytic contaminants, produced by bacteria are those that have been secreted from the cell through a secretion system. Proteins that are only detected in the cell pellet, and not in the culture filtrate, are those that are not secreted, which may include cytoplasmic or membrane-embedded proteins. Some proteins may be detected in both cell pellet and culture filtrate fractions, meaning the protein is both present in the cytoplasm
and is secreted. One limitation of our MS analysis is that characterizing proteins bound to, or embedded in, the cytoplasmic- and myco-membranes is difficult. Membrane-associated proteins are often less soluble than cytoplasmic proteins, which poses a problem for LC and are not accessible to enzymes, namely trypsin, used to prepare samples for MS (Schey, K. L., et al., 2015). Cyanogen bromide has been successfully used to enrich hydrophobic proteins within the membrane prior to MS (Ball, L. E., et al., 1998). We can use targeted MS to detect membrane-associated proteins, which will elucidate characteristics of SigM- and non-SigM-induced proteins, such as hydrophobicity.

MS profiles were generated from cell pellets and culture filtrates of WT recipients as well as recipients lacking esx-4. MS identified the proteins within the filtrate produced by each recipient strain described. The profiles generated from each sample were used to determine what proteins were detected in the culture filtrate of WT recipients but not in the culture filtrate of esx-4 null recipients.

**Preliminary results suggest that ESX-4 secretes SigM regulon encoded proteins.**

We have preliminary results that suggest ESX-4 is responsible for the secretion of several SigM regulon proteins. MKD8_1560, MKD8_1538 (EsxU), and MKD8_5844 were detected in the culture filtrate of WT recipient cells but not in the filtrate produced by recipients lacking ESX-4 (Figure 3.2). MKD8_5844, another SigM induced protein, is secreted by ESX-4. MKD8_5844, a protein containing a domain of unknown function (DUF), was not previously characterized using M-PFC and will be screened against other ESX-4 substrates by members of our laboratory in the future (Table 1.1). MKD8_6925 is found in culture filtrates and is still detectable in the culture filtrate harvested from esx-4 null recipients, but levels are reduced (Figure 3.2). This implies that MKD8_6925, MKD8_1560, EsxU, and MKD8_5844 are substrates of ESX-4. Secretion (MKD8_6398: mycobacterial antigen 85-A) and lysis (MKD8_1654: Isocitrate
dehydrogenase) controls demonstrated that secretion of non-ESX-4 dependent substrates is not altered when ESX-4 is not present and that cytoplasmic proteins are minimally present in culture filtrates which is indicative of cell lysis.

Our results also identified MKD8_1537 and MKD8_15690 as secreted proteins (Figure 3.2). However, these proteins may be secreted in an ESX-4-independent manner. If proteins are detected in the supernatant of cells lacking ESX-4, it is still possible that the protein is not secreted via ESX-4 or that the protein is still a substrate of ESX-4 but can also be secreted through another secretion system.

MKD8_15690 is a SigM regulon encoded protein that contains a domain of unknown function (Table 1.1). The secretion of MKD8_15690 does not appear to be altered in the absence of ESX-4 and, thus, MKD8_15690 is likely secreted by another secretion system. A small amount of MKD8_15690 could be secreted via ESX-4, but the MS analysis described here cannot currently support this conclusion. Identifying other secretion systems responsible for the transport of SigM regulon encoded proteins will help us to further define mechanisms by which SigM regulon encoded proteins are secreted.

MKD8_1537 is part of the esx-4 core locus and it is likely that secreted MKD8_1537 is transported via ESX-4, but our esx-4 locus deletion also removes the MKD8_1537 gene. We could determine if MKD8_1537 is secreted by ESX-4 by analyzing the culture filtrate produced by a recipient lacking an ESX-4 structural protein such as EccC4 instead of removing the entire system from recipient cells (Figure 3.2). If MKD8_1537 is still secreted when ESX-4 is inactive, we can conclude that MKD8_1537 is secreted through a non-ESX-4 secretion system. However, if MKD8_1537 is expressed but not secreted in the inactivation mutant, we will
conclude that ESX-4 is responsible for exporting MKD8_1537. An experiment, described below, will be used to determine if MKD8_1537 is secreted via ESX-4.

As we continue to analyze the data obtained from MS, we will likely identify other substrates of ESX-4. These substrates will be explored using M-PFC to further elucidate protein-protein interactions, and DCT will be used to determine if non-SigM regulon proteins secreted through ESX-4 are required for DCT.

**MS identification of co-dependent secretion.**

Co-dependent mechanisms of secretion are exhibited by other ESX substrates (Fortune, S. M., et al., 2005). We began to characterize protein-protein interactions between SigM regulon proteins using M-PFC screen, which is discussed later in this chapter. M-PFC analyses have elucidated complexes formed between SigM regulon encoded proteins that are likely relevant secretion through ESX-4. MS data will be used to determine whether the protein-protein interactions identified by M-PFC are important for secretion. Therefore, I prepared samples for MS analysis that, when analyzed, will complement the M-PFC results discussed later in this chapter.

**Samples lacking specific SigM regulon proteins were prepared and sent for MS analyses.**

*sigM* was constitutively expressed in recipients lacking the candidate SigM regulon encoded proteins described in chapter 2 of this thesis (Table 3.1). Cell pellets and culture filtrates were harvested from each knockout strain overexpressing *sigM* and were sent to Notre Dame for the same LC/MS-MS analysis described earlier in this chapter (Bosserman, RE., et al., 2017). In addition to fractions harvested from recipients lacking SigM regulon proteins of interest, I sent cell pellets and culture filtrates produced by recipients with deletions of *MKD8_1535* (EccD4), *MKD8_1536* (EccC4), or *esxUT* for MS analysis (Table 3.1). Knockouts of EccD4 and EccC4
serve as controls for secretion because in the absence of ESX-4 structural components, we should not see secretion of any ESX-4 substrates. Analyzing the secretion profile of esxUT null recipients will distinguish proteins whose secretion is dependent on EsxUT from any that are secreted independently. We will determine if secretion via ESX-4 can occur in the absence of EsxUT and, if secretion can still occur, MS will identify proteins reliant upon EsxU and EsxT for secretion.

By comparing the composition of culture filtrates produced by each knockout, we will determine what proteins are dependent upon a partner for secretion. For example, if MKD8_1686 is required for the secretion of MKD8_1685, we should detect MKD8_1685 in the cell pellet, but not in the culture filtrate produced by recipients lacking MKD8_1686 (Figure 3.1B). In this example, MKD8_1686 could act as a direct secretion partner for MKD8_1685 or could facilitate the formation of a protein complex required for the secretion of MKD8_1685. We will use MS analysis to determine whether other SigM-induced proteins exhibit this same co-dependent mechanism of secretion.

**The SigM Regulon and DCT**

*MKD8_21385, MKD8_MKD8_16105, MKD8_53960, and MKD8_15750 were not required for DCT.* Determining a role for these genes will require the development of additional assays to study their function. It will be important to first determine whether these proteins are secreted via ESX-4 via mass spectrometry analysis. These proteins were not characterized using M-PFC during this analysis, but comprehensive M-PFC data for these proteins may elucidate a role for them as secretion partners for other ESX-4 substrates, especially if MKD8_21385, MKD8_16105, and MKD8_53960 are ESX-4 substrates themselves. However, determining the function(s) of MKD8_21385, MKD8_16105, and MKD8_53960 may be complicated.
MKD8_16105 (11 aa) and MKD8_21385 (24 aa) are small proteins and their size could complicate their characterization. Structural analyses of MKD8_16105 and MKD8_21385 using crystallography may not be successful because obtaining a high enough concentration of crystallized protein for analysis relies on the ability of the protein to form effective crystals, which may not possible with small proteins. Even if we could obtain the structures of MKD8_16105 and MKD8_21385, due to their size, structural analyses may not provide us with much information. We will characterize MKD8_16105 and MKD8_21385 using MS analysis and M-PFC. We will use MS analyses to determine if these two small proteins are secreted and if they are involved in co-dependent secretion. M-PFC will be used to determine if MKD8_16105 or MKD8_21385 form PPI with SigM-induced secretion substrates or secretion machinery.

It was not surprising that MKD8_53960 is dispensable during DCT as MKD8_53960 is a pseudogene. MKD8_53960 will be assayed using M-PFC and we will determine if MKD8_53960 is secreted. If MKD8_53960 is detected in the culture filtrate of MKD8 overexpressing sigM, we can conclude that MKD8_53960 is being transcribed and translated. If MKD8_53960 is fully expressed, there is a chance it could be purified and subjected to crystallography. A function for this MKD8_53960 could potentially be elucidated if any hypothesized functional domains are identified in its structure.

MKD8_15750 is encoded from the same transcript as MKD8_1560, a gene shown by our preliminary work to be required for transfer (Clark, R. R., et al., 2018). Therefore, we hypothesized that MKD8_15750 would also be required for transfer. We were surprised to find that MKD8_15750 is not required for DCT. MS analyses described earlier in this chapter will elucidate whether MKD8_15750 is required for the secretion of MKD8_1560. If MKD8_15750 is required for secretion of MKD8_1560, but not for DCT, there is reason to believe MKD8_1560
carries out its role during DCT intracellularly, which is valuable information that will inform our future analyses.

*While several SigM-regulated genes explored in this thesis were not required for DCT, these data have positive implications. The presence of genes not involved in DCT implies that the SigM regulon is contributing to other cellular functions that are yet to be discovered.*

**MKD8_6925 is a predicted calcium binding protein required for DCT.**

DCT requires cell-cell contact between donors and recipients. This contact could lead to damage in the cell wall of either parent, or open a pore for transfer of donor DNA to the recipient cell. This disturbance of the cell wall could itself be the trigger for release of SigM from its anti-sigma factor. We speculate that cell-cell contact could destabilize the cell membrane and result in an influx of calcium or other ions into the recipient cytosol. Bacteria encode secretion systems that maintain metal homeostasis because excess metals can be toxic to the cell by creating DNA-damaging reactive oxygen species, or by displacing normal metal cofactors required by enzymes for their function (Moraleda-Muñoz, A., et al., 2019). Therefore, the recipient could require proteins to repair damages to the cell wall and to re-establish the ion gradient that exists between the cell and environment. MKD8_6925 could be specifically involved in re-establishing calcium homeostasis while other unknown proteins may be activated to respond to an overabundance of non-calcium ions.

We are not yet certain whether MKD8_6925 carries out its function intracellularly or extracellularly. However, preliminary mass spectrometry results suggest that MKD8_6925 is the most highly secreted substrate of ESX-4 (Figure 3.2, unpublished results). If we are able to conclusively identify MKD8_6925 as a secreted substrate of ESX-4, there is further reason to believe that MKD8_6925 binds to and sequesters excess calcium that has flooded into the
recipient during DCT. Excess calcium would then be removed from the cell upon the secretion of MKD8_6925 through ESX-4, which would aid in re-establishing calcium balance in recipients that have acquired segments of the donor’s chromosome (transconjugants).

*While the data described in this thesis did not lead to assigning MKD8_6925 a precise role during DCT, it did lead to the discoveries that calcium is important for not only log-phase growth of WT recipients but also for DNA transfer between WT donors and recipients.*

**EGTA disrupts signaling between SigM and its regulon during co-culture.**

We have demonstrated that induction of SigM-regulated transcripts during DCT is disrupted when EGTA is present during co-culture. This phenotype is recipient-specific and points to a role for calcium during signaling in DCT. Calcium could play several different roles during signaling between SigM and its regulon. SigM is an ECF sigma factor and, based on model ECF systems, is predicted to be sequestered by its anti-sigma factor at the IM. Calcium could serve as the environmental signal that stimulates the release of SigM. Anti-SigM does not appear to contain a calcium binding motif (DxDxDG), however we experimentally determined if Anti-SigM senses calcium and, thus, releases SigM to upregulate genes involved in calcium homeostasis.

We determined if calcium triggers Anti-SigM by exposing MKD8 to exogenous calcium in monoculture. A reporter assay using luciferase was performed by Melissa Stone, a member of our laboratory, to determine whether exogenous calcium was able to stimulate expression of the SigM regulon. Melissa demonstrated that when exogenous calcium was applied to MKD8 monocultures, expression of the SigM regulon was not increased above WT levels, which suggests that calcium is not sensed by Anti-SigM.
Although calcium does not appear to stimulate Anti-SigM, it is possible that calcium still plays a role in the release of SigM from Anti-SigM. ECF sigma factors are commonly released from their anti-sigma factor by regulated intramembrane proteolysis (Reviewed in Heinrich, J., & Wiegert, T., 2009). This proteolysis requires an enzyme capable of cleaving the connection between a sigma factor and its anti-sigma factor. Many enzymes require metal cofactors for function and, therefore, calcium could be required for SigM signaling if it acts as a co-factor to the protease responsible for release of SigM from the membrane.

*We will perform RNAseq to determine whether the addition of EGTA during co-culture causes global changes in transcription rather than a DCT-specific response. If this were the case, calcium could be acting as a signaling molecule in a pathway related to the re-establishment of homeostasis of the bacterium and is not specific for DCT.*

**Protein-Protein Interactions within the SigM Regulon**

**EsxU and EsxT form a direct protein-protein interaction.**

A number of protein-protein interactions were explored, but only 2 interactions were identified. One positive interaction was observed between EsxU and EsxT. Like EsxAB, EsxUT are almost certainly secreted in a co-dependent fashion that is reliant upon their physical interaction. Moving forward, EsxU and EsxT are the likely the best candidates for screening against other secreted proteins using M-PFC. I believe that beyond an interaction between EsxU and EsxT, EsxU and EsxT, alone or in complex (EsxUT), will interact with other secreted proteins as well as ESX-4 core machinery. Assaying the structural components of ESX-4 for PPI, such as EccC4, will elucidate even more about how ESX-4 assembles at the membrane and the mechanisms by which ESX-4 substrates contact the secretion apparatus for translocation.
If we identify an ESX-4 substrate that contacts MKD8_1536 for secretion, we could use mutagenesis followed by M-PFC to identify the amino acids required for an effective PPI. This could further support the hypothesis that the signal sequence within the C-terminal tail of T7SS substrates is the key feature responsible for trafficking T7SS substrates to the secretion apparatus. Mutagenesis followed by M-PFC could also elucidate other amino acids, not contained within the T7SS signal sequence, that are required for interaction(s) between substrates and secretion machinery. Therefore, the analyses mentioned here could identify other T7SS substrate features that facilitate transport of a substrate through ESX-4.

**MKD8_1686 and MKD8_1685 form a PPI but do not interact with MKD8_1687.**

When screened using M-PFC, we discovered that MKD8_1687 does not interact with MKD8_1686, MKD8_1685. To determine if MKD8_1687 requires both MKD8_1686 and MKD8_1685 to form a protein complex, a fusion protein consisting of a flexible m-DHFR tag at the N-terminal of MKD8_1686 followed by both MKD8_1686 and MKD8_1685 was generated. The MKD8_1686-MKD8_1685 fusion was screened for the formation of a PPI with MKD8_1687 using M-PFC. No PPI appeared to be formed between MKD8_1687 and MKD8_1686, MKD8_1685, or the two combined. This implies that MKD8_1687 might not contact MKD8_1686 and MKD8_1685 as we expected. However, there may be other proteins outside of the MKD8_1687 operon that facilitates a PPI with MKD8_1686 and MKD8_1685. We will use MS analyses to determine other candidate proteins that may be co-secreted and will explore potential PPI using M-PFC.

**MKD8_1560 and MKD8_15750 did not form a protein-protein interaction during M-PFC.**

It was surprising that proteins such as MKD8_15750 and MKD8_1560 did not interact because of their co-transcription and clear T7 substrate-like features. We hypothesized that MKD8_15750 would form a protein-protein interaction for secretion with MKD8_1560 and the
two would be secreted through ESX-4 in a co-dependent fashion. However, all the elements necessary for establishing a stable interaction between MKD8_1560 and MKD8_15750 may not be present under our experimental conditions. These two proteins may require an additional, non-SigM regulon protein partner to form a physical interaction.

**MKD8_6925 and MKD8_1537 did not bind to any proteins screened using M-PFC.** The fact that MKD8_6925 did not form any interactions with our candidate proteins could mean MKD8_6925 has a different secretion partner than EsxU or EsxT, which were the only two proteins screened against MKD8_6925. MKD8_6925 could be secreted by itself, which will be elucidated by determining co-dependent mechanisms of secretion using MS analyses. If MKD8_6925 is secreted alone, we will have identified a new T7SS substrate that does not contain a classic helix-WXG-helix structure and does not contain a C-terminal T7 signal sequence. However, MKD8_6925 could also be secreted through another secretion system such as Sec or Tat. MKD8_6925 does not contain canonical signal sequences for secretion via Sec or Tat, but we could still investigate whether MKD8_6925 could be secreted via Sec or Tat by looking for PPI between MKD8_6925 and Sec/Tat secretion machinery. If MKD8_6925 forms a PPI with general secretion machinery, there would be reason to believe that MKD8_6925 could be translocated by a non-T7SS.

**Controlling the limitations of our M-PFC assay could identify additional PPI.** We used donor cells, which do not exhibit basal expression of the SigM regulon. If complexes between SigM-induced proteins require PPI between more than two SigM-induced proteins, potentially mediated by a chaperone, our analyses could fail to identify certain complexes. To determine if other proteins are involved in the interactions between proteins such as MKD8_1560 and MKD8_15750, as well as others, M-PFC could be performed when \( \text{sigM} \) is
overexpressed. This would provide the m-DHFR tagged substrates with all SigM-regulated proteins and, thus, provide potential protein chaperones needed for interaction.

We could use co-immunoprecipitation (co-IP) to identify protein complexes composed of more than two proteins. For example, we could use cloning to fuse a hemagglutinin tag to a SigM-induced protein such as MKD8_15750. We could then use an anti-hemagglutinin antibody to precipitate MKD8_15750 and any proteins complexed with MKD8_15750 from solution. Finally, MS analysis could be used to identify the proteins complexed with MKD8_15750.

The domains of m-DHFR could also complicate identification of PPI. It is unlikely that the fragments of m-DHFR prevent proteins from associating because they are fused to each protein of interest through a flexible glycine linker that provides space for proteins to form complexes (Singh, A., et al., 2006). However, methods such as isothermal titration calorimetry has been used to identify PPI without the use of protein tags (Pierce, M. M., et al., 1999). Calorimetry relies upon changes in heat caused by the formation of PPI. A PPI can be formed by an exothermic reaction which releases heat or through an endothermic reaction which requires heat. Two proteins are combined and any changes in heat are measured by a calorimeter and changes in heat can be indicative of PPI formation.

While research has suggested several models for ESX secretion these models typically do not include ESX-4, which may secrete proteins a different way (Wang, S., et al., 2019). However, we hypothesize that mechanisms of secretion are the same as ESX-4 is the progenitor T7SS. Although more research is required to definitively determine mechanisms of secretion through ESX-4, we have identified PPI here that will help complement MS data and further our understanding of ESX-4 and T7SS.
Conclusions

Further investigation of the SigM regulon is required if we are to assign function(s) to it. SigM-induced proteins may be fulfilling roles during DCT including hydrolase-mediated cell wall remodeling, or another role not yet determined. There is also a chance that substrates of ESX-4 could decorate the cell surface of recipients that would allow for recognition of recipients by donor cells, but more research is required to substantiate this possibility.

Performing analyses using other mycobacterial species may reveal more information about conserved SigM-regulon genes MKD8_6925, MSMSEG_1560, esxU, and esxT. There is evidence showing ESX-4 plays a role in pathogenesis of M. abscessus; determining a role for conserved SigM genes during infection could lead to insights about the functions of these conserved genes during DCT (Laencina, L et al., 2018). Analyses of T7SS could expand the number of potential targets for therapeutics against drug resistant mycobacteria if T7SS beyond ESX-1 encode antigenic proteins required for pathogenesis that could be targeted with antibody therapy.

This project has elucidated more information about SigM-regulated genes, particularly and our future analyses will complement the results presented in this thesis. there is still much to be learnt. As previously mentioned, MS analyses will help further define not only the relationship between the SigM regulon and ESX-4, but also the relationship between substrates of ESX-4. These experiments will be important in understanding the underlying mechanisms of T7SS, which could inform future research into other T7SS; esx-4 is conserved and relatively simple, containing only 7 core genes, making it a useful model for more sophisticated ESX systems. Furthermore, the mechanism of ESX-4 secretion will likely be conserved because it's hypothesized that all ESX systems arose from ESX-4.
Table 3.1. Cell pellets and culture filtrates were harvested from recipients lacking candidate SigM-regulated genes to identify any co-dependent mechanisms of secretion through ESX-4. sigM was overexpressed in all strains to activate secretion via ESX-4. Pellet and supernatant fractions were harvested from recipients lacking SigM-regulated proteins not encoded within the core esx-4 locus (shaded blue). Fractions were also harvested from recipients lacking esx-4 core genes to serve as controls for secretion (EccD4, EccC4 mutants) and to explore co-dependent mechanisms of secretion (EsxUT) (shaded green).
Figure 3.1. *sigM* was overexpressed in different recipient strains to identify the substrates of ESX-4 and whether SigM regulon encoded proteins are responsible for the secretion of one another. A. Constitutive expression of *sigM* in recipients activates the SigM regulon and secretion via ESX-4. Secretion systems other than ESX-4 may be active prior to ESX-4 activation B. We will use MS analyses to determine if SigM-regulated proteins are secreted, MKD8_1686 and MKD8_1685 are used as an example. MKD8_1686 and MKD8_1685 may be secreted through ESX-4 or another unknown secretion system. C. By eliminating ESX-4 from recipients, we can determine if MKD8_1686 and MKD8_1685 are
secreted in an ESX-4 dependent manner. D. By knocking out MKD8_1686, we can determine if MKD8_1685 is still secreted in the absence of MKD8_1686 and, therefore, we can use MS to conclude whether MKD8_1685 is co-secreted with MKD8_1686.
Figure 3.2. Preliminary MS analyses of SigM regulon proteins secreted in an ESX-4-dependent manner. Cell pellets (P) and culture filtrates (F) harvested from both WT and esx-4 null MKD8 +/- SigM induction were subjected to LC-MS/MS. The changes in detection of 8 proteins from both P and F generated by each strain as quantified relative to no SigM induction are shown here. Yellow = proteins that were not differentially detected in P of F when sigM was overexpressed. Blue = proteins detected at higher levels upon SigM induction. Black = Not detected. Overexpression of sigM lead to >6-log$_2$ increase for all MS-detectable SigM regulon proteins (MKD8_1537, MKD8_1538, MKD8_15690, MKD8_1560, MKD8_5844, and MKD8_6925). In the absence of ESX-4, the non-ESX-4 SigM-induced proteins were still detected in the cell pellet, however, only
MKD8_15690 was found in F. changes in MKD8_1654 (Isocitrate dehydrogenase) serves as a lysis control, and MKD8_6398 (Antigen 85-A) was used as a secretion control independent of ESX-4. Figure generated by Derbyshire & Gray.
Chapter 4: Materials and Methods

Strains

Strains of *M. smegmatis* used during this research included mc²155 donor cells and MKD8 recipient cells. All mycobacterial protein fragment complementation (M-PFC) assays were completed using mc²155. MKD10, a derivative of mc²155 marked with kanamycin (Kan), was used for conjugation assays. MKD8, which encodes streptomycin (Sm) resistance was the recipient. Chemically competent NEB5a or Stellar Cells (Cat. #C2987l, #636766) were commercially available *E. coli* strains used in all clonings described below.

Media

*M. smegmatis* was cultured using trypticase soy broth (TSB) with added 0.05% tween 80 (T) (TSBT). Trypticase soy agar (TSA) plates were used for all matings and to recover transformations of *M. smegmatis*. Plates containing 7H10 agar, 1X final concentration ADS, and 25 µg/mL trimethoprim were used for M-PFC. MKD8 samples used for mass spectometry were cultured in TSBT + Sm and diluted in 7H9 media containing 0.05% tween80, 0.2% glycerol, and 0.5% asparagines (MM + Asp). Concentrations of antibiotic added to growth media for *E. coli* and *M. smegmatis* are listed in Table A1 and Table A2 respectively.

Homologous recombineering

Creating recombineering competent cells: MKD8 containing pJV53, a plasmid containing an acetamide inducible promoter for genes 60 and 61 from Che9c (JVK and GH), to which sacB was added to facilitate counter selection curing after recombineering, was used (MKD8 + pJV53 sacB). Recombineering-proficient electrocompetent cells were created essentially as described (JVK and GH), but with slight modifications. MKD8 + pJV53 sacB were inoculated into 7H9 medium +T +ADC +kanamycin (kan). The following day, cells were diluted to an OD600 of
0.025 in 50 mL 7H9 -ADC +kan +0.2% succinate. Cultures were grown to an OD600 value of 0.4 (12-16 hours). Acetamide was added to a final concentration of 0.2% to induce gp60-61 expression. Cultures were incubated for an additional 3 hours at 37°C. After incubation, cells were placed on ice for 30 minutes and then centrifuged at 5000 RPM for 10 minutes to collect cells. Cell pellets were washed four times with decreasing volumes of 10% glycerol to obtain a concentrated stock of recombineering cells, which were resuspended in 10% glycerol and used immediately or stored at -80°C until use.

**Generation of knockout constructs:** Gene specific primers were used to amplify genomic DNA surrounding each candidate gene using Q5 polymerase (Table A3). Each product contained tails homologous to a zeocin (zeo) resistance cassette, which allowed for the use of synthesis via overlap extension (SOE) PCR to generate products in which the homologous arms flanked zeocin (zeo). These knockout constructs were electroporated into MKD8 recombineering cells, which allowed for homologous recombination. Cells were recovered, plated on TSA containing zeo, and incubated for 5-7 days at 37°C to isolate mutants. Candidates were screened using colony PCR using primers that flanked either side of the site of recombineering (Table A3). PCR positive mutants were transferred to sterile selective media, incubated for 4-5 days at 37°C, and stored at 4°C until assayed for conjugation.

**Distributive Conjugal Transfer (DCT) Assay**

DCT assays were all performed as described in Clark, RR., et al., 2018. Cultures of conjugal partners, including MKD10 donor cells and MKD8 recipient cells were incubated overnight at 37°C in media containing kan and Sm, respectively. The following day, cultures were diluted to an OD600 below 0.7 and incubated at 37°C until an OD of 1.0-1.2 was reached (2-3 hours of incubation), which ensured actively dividing bacteria were used during our assay. Equal volumes of donor and recipient cells (500 uL) were washed in TSB and resuspended in 50 uL TSB. 50 uL of donor and 50 uL of recipient cells were mixed together and spotted on TSA.
Plates were incubated at 30°C, the optimal temperature for transfer to occur, for 18-24 hours. Co-cultured cell pellets were collected using a sterile loop and transferred to a microcentrifuge tube containing 1 mL TSBT. Cells were then serially diluted (1:9) and 100 uL of dilutions of each mating pellet (N, -1, -5, and -6) were plated for recovery of transconjugants (neat and 10⁻¹ on TSA + Km,Sm), donors (10⁻⁵ and 10⁻⁶ on TSA + Km), and recipients (10⁻⁵ and 10⁻⁶ on TSA + Sm). Plates were incubated for 3-5 days at 37°C and colonies on each plate were enumerated. Transfer frequency was calculated by dividing the number of recovered transconjugants by the average number of recovered recipient cells. The mating frequencies of WT were compared to those of conjugation assays using mutant recipients to determine any defects or increases in transfer levels. Average efficiencies were determined using 3 replicates.

To determine the necessity of calcium during mating, co-cultures containing MKD10 and MKD8 were grown overnight, washed, and resuspended according to the conditions listed above, but were plated on TSA containing 1mM EGTA, a calcium chelator.

Quantitative Reverse Transcription PCR (qRT-PCR)

RNA Isolation: RNA was extracted from 24-hour mating pellets that had been cultured on TSA or TSA 1 mM EGTA. Cocultured cells were scraped from plates using a sterile loop and transferred to a screw-cap microcentrifuge tube and resuspended in 1 mL TSBT. Cells were centrifuged at 13,200 RPM for 1 minute to harvest cells. The supernatant was removed and the pellet was kept on ice and covered in 500 uL zirconia beads (Ambien # AM1925) and 1 mL Trizol (Invitrogen # 15596-026). Samples were homogenized by bead-beating the tubes for 1 minute and 30 seconds at 4°C. In a chemical fume hood, 200 uL of chloroform was added to each homogenate. Tubes were shaken by hand for 15 seconds and allowed to incubate at room temperature for 2-3 minutes. Samples were centrifuged at 12,000xg for 15 minutes at 4°C. The top (aqueous) layer separated from the mixture was removed into a new tube, to avoid the interface containing DNA or protein. Pure isopropanol (500 uL) was added to each tube and
incubated at room temperature for 10 minutes. Samples were centrifuged at 12,000xg for 10 minutes at 4°C and the supernatant was discarded. Pelleted RNA was washed once with 70% ethanol and centrifuged at 7,500xg for 5 minutes at 4°C. The supernatant was carefully separated from RNA pellets. Traces of ethanol were allowed to evaporate at room temperature. RNA pellets were resuspended in 20-150 uL of sterile DEP-C treated Millipore water depending on estimated pellet size. The concentration of RNA in each sample was determined using a nano-drop spectrophotometer (ThermoFisher). Samples were immediately subject to DNase treatment.

**DNAse Treatment**: Five to eight micrograms of extracted RNA was DNase treated to remove genomic DNA from the RNA samples. RNA was mixed with 2 uL of DNase (AM1907), 1 uL RNase Inhibitor (AM2696) and 10X buffer at a final concentration of 1X. Reactions were incubated at 37°C for 30 minutes, then 1 µl of DNase was added to each reaction and incubated for an additional 30 minutes at 37°C. Following incubation, 0.1 volume of DNase inactivation slurry (AM1907) was added to reactions. The amount of DNase inactivation slurry used was based on reaction volume after addition of RNA, DNase, and RNase inhibitor. RNA was precipitated by adding 0.1 volume of 3M sodium acetate, and 2X volume of 100% ethanol. Samples were mixed and incubated at -80°C for at least 20 minutes. The samples were centrifuged at 12,000xg for 15 minutes at 4°C to collect RNA. Supernatants were removed and RNA pellets washed once with 70% ethanol. Dried RNA pellets were resuspended in DEPC water to their original volume (the volume used for DNase treatment). PCR using genomic DNA (gDNA) specific primers was used to ensure the RNA contained no residual gDNA. Treated samples were stored at -20°C until further use.

**cDNA Preparation**: The SuperScriptIII cDNA first strand synthesis kit was used to generate cDNA from DNase treated RNA samples (18080-044). 200 – 500v nanograms (ng) of each RNA
sample were used in a 10 uL annealing reaction containing 50 ug/mL random hexamer primers, dNTPs, and water. Tubes were incubated at 65°C for 5 minutes and then on ice for 1 minute, which allowed for hybridization of the oligonucleotides to the RNA being synthesized into cDNA. A cDNA synthesis mix was made according to Invitrogen protocol #18080-051 and added to pre-incubated RNA and random hexamers. Reaction mixtures were incubated according to the conditions described below:

- 25°C for 10 minutes
- 50°C for 50 minutes
- 70°C for 15 minutes

Reactions were brought up to 100 uL in Millipore water and stored at -20°C until use in qRT-PCR. Note, cDNA used in these experiments was synthesized at the same starting concentration.

qRT-PCR: To measure relative transcript levels in each prepared sample, a master mix including 1 uL of sense primer (10 uM), 1 uL of antisense primer (10 uM), 10 uL SYBR Green pre-made mix (APB #4472908) and 6 uL of ddH2O were added to a 96-well plate (primers listed in Table A3). Plates were assayed using Applied Biosystems 7500 Fast Real-time PCR machine under default conditions and cycle threshold (CT) value of each qRT-PCR reaction was determined using 7500 software version 2.3. Threshold values were used to determine changes in gene expression. cDNAs were amplified in triplicate using each primer set listed in Table 1A. Fold change in gene expression was calculated using data from cDNA amplified in triplicate normalized to rpoB in all experiments. Significant changes in expression were determined using one-way ANOVA analysis in GraphPad Prism 9.
Cloning

**Strains:** Calcium-chloride competent *Escherichia coli*, either NEB 5-alpha (C2987I) or TakaraBio stellar cells (Cat. #636763) were used for all cloning described below.

**In-Fusion cloning:** pSJ25, an L5 integrating plasmid marked with hygromycin resistance, was used for InFusion cloning. Round-away primers listed in table 1 were used to synthesize a linearized vector via PCR using Q5 polymerase. The vector was then treated using DpnI, an enzyme that cleaves methylated DNA, ensuring that template used for InFusion contained no contaminating circularized parental vector. Inserts were generated using primers with at least 15 base pairs (bp) of identity to pSJ25 at the site of linearization and to the genomic region surrounding the target gene. All primers were designed to amplify across the SigM promoter of each cloned gene, which allowed for native expression. Amplified linear pSJ25 and inserts were combined in a 1:3 molar ratio respectively along with In-Fusion HD Enzyme Premix (Takara: 639649) and were incubated following the kit protocol. Reactions were transformed into chemically competent *E. coli* via heat shock. Cells were recovered in SOC and plated on LB + hyg to select for colonies containing clones of pSJ25::insert. Plasmids were isolated using the Qiagen miniprep kit (Qiagen: 27104) and sequences were confirmed via Sanger sequencing.

**Restriction enzyme and ligation cloning:** All M-PFC constructs were generated using pUAB300, an episomal plasmid encoding hyg, resistance and pUAB400, an attP integrating plasmid encoding kan resistance. Each plasmid contained a domain of murine dihydrofolate reductase (m-DHFR) (Singh, A., et al., 2006). pUAB300 was digested at 37°C using BamHI-HF (NEB: R3136) and HindIII-HF (NEB: R3104), while pUAB400 was incubated with MfeI-HF (NEB: R3589) and HindIII-HF, both in cutsmart buffer (NEB: B7204). Gene inserts were amplified using primers containing appropriate restriction sites added onto each insert for cloning into pUAB300 or pUAB400. Primers were designed to ensure inserts were in-frame with the dhfr
domain. Inserts were digested with the appropriate restriction enzymes and were ligated into pUAB300 or pUAB400 following New England Biolabs T4 ligation protocol (NEB: M0202). 10 µL of each ligation reaction was heat shocked into *E. coli* and transformants were selected for on either LB containing hyg or kan. Potential recombinant colonies were cultured overnight for plasmid isolation using the QIAGen Miniprep Kit and protocol. The presence of an insert in either pUAB vector was confirmed via PCR and/or restriction digest. All clones were confirmed by DNA sequencing analysis.

*Mycobacterial Protein Fragment Complementation (M-PFC)*

**Creating electrocompetent mc^2155:** Cells were grown overnight in TSBT and diluted to an OD600 of 0.5-0.7 to ensure actively dividing cells were made competent. Diluted cultures were incubated for 1-2 hours at 37°C and then cells were harvested by centrifugation at 4°C. Cells were washed twice with 10% glycerol and were resuspended in 400-1500 µL of 10% glycerol depending on the starting culture volume. Cells were stored as 200 µL aliquots at -80°C until use.

**M-PFC assay:** pUAB300 and pUAB400 plasmid derivatives were co-electroporated into competent mc^2155 cells. Cells were recovered in 1 mL TSB + tween for 4-5 hours at 37°C and plated on TSA plates containing hyg and kan to select for double transformants. The plasmid transformants were cultured overnight in TSBT. Positive controls included pUAB100 and pUAB200, which contain interacting fragments mediated by the yeast modular transcription factor, GCN4 (Singh, A., *et al.* 2006). Cultures were grown to an OD600 value of 1.0-1.5. One milliliter of each culture was serially diluted 1:10 [usually 1:9] in TSBT and 5 µl of each dilution was spotted on TSA and on 7H10 containing 25 µg/mL trimethoprim, which allows for selection of bacteria containing pUAB vector pairs that encode interacting proteins. Cells were incubated at 37°C for 2 days and growth was recorded.
**Mass Spectrometry (MS) Sample Preparation**

To determine the co-secretion profile of ESX-4, a plasmid overexpressing SigM, which induced expression of *esx*-4, was transformed into MKD8 and knockouts of selected SigM-regulon genes via electroporation (Table 3.1). Cells overexpressing SigM were cultured overnight at 37°C in TSBT until saturation. The following day, cultures were diluted to an OD600 of 0.06 in 10 mL TSBT and incubated for 8-9 hours at 37°C or until an OD600 value of 0.25-0.35 was reached. Cells were collected by centrifugation at 4,000 RPM for 5 minutes and washed once with MM + Asp. Cells were resuspended in 5 mL pre-warmed MM + Asp, and diluted into 25 mL of MM+ Asp, and incubated at 37°C for 3-4 hours. Cells were collected by centrifugation at 4,000 RPM for 5 minutes. Ten mL of each culture supernatant was filtered through 0.22 um sterile syringes to remove particulates and bacteria. Cell pellets were transferred to a microcentrifuge tube. Cells and culture filtrates were stored at -80°C until shipment to our collaborator Dr. Matthew Champion at Notre Dame. Samples were subjected to MS to determine the secretome of each SigM regulon knockout strain.
Table 4.1. Antibiotic concentrations used for solid and liquid bacterial cultures.

<table>
<thead>
<tr>
<th>Antibiotic</th>
<th>Solid Media (ug/mL)</th>
<th>Liquid Media (ug/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Streptomycin</td>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td>Kanamycin</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Hygromycin</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>Apromycin</td>
<td>12.5</td>
<td>12.5</td>
</tr>
<tr>
<td>Zeocin</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>Trimethoprim</td>
<td>25</td>
<td>N/A</td>
</tr>
<tr>
<td>Hygromycin</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>Kanamycin</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>Apromycin</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>Ampicillin</td>
<td>100</td>
<td>N/A</td>
</tr>
</tbody>
</table>

M. smegmatis

E. coli
<table>
<thead>
<tr>
<th>Identifier</th>
<th>Description</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGD5000</td>
<td>MKD8 1687 KO A</td>
<td>ttcggctacgtgctggaaggc</td>
</tr>
<tr>
<td>TGD5001</td>
<td>MKD8 1687 KO A'</td>
<td>gaagcgttctgggagcccgc</td>
</tr>
<tr>
<td>TGD5002</td>
<td>MKD8 1687 KO B</td>
<td>TGGTGAGGGAGATGAGGTCTGAAGcacgcagc</td>
</tr>
<tr>
<td>TGD5003</td>
<td>MKD8 1687 KO C</td>
<td>GTTGAGGTGTAGGCTGCTGAAGgtccgtatgccg ggcgcactg</td>
</tr>
<tr>
<td>TGD5004</td>
<td>MKD8 1687 KO D'</td>
<td>ctfagcatccccgcggatcg</td>
</tr>
<tr>
<td>TGD5005</td>
<td>MKD8 1687 KO D</td>
<td>ttcggctacgtgctggaaggc</td>
</tr>
<tr>
<td>TGD5006</td>
<td>MKD8 16105 KO A</td>
<td>GACGTCGTCGCGAACTTGAC</td>
</tr>
<tr>
<td>TGD5007</td>
<td>MKD8 16105 KO A'</td>
<td>CGGGTACGTCGTGACCTACCG</td>
</tr>
<tr>
<td>TGD5008</td>
<td>MKD8 16105 KO B</td>
<td>TGGTGAGGGAGATGAGGTCTGAAGGTAAGTGT TAGACGCAATCCGAGGG</td>
</tr>
<tr>
<td>TGD5009</td>
<td>MKD8 16105 KO C</td>
<td>GTTGAGGTGTAGGCTGCTGAAGGCAACTGCG CCCCGAGATCCGAGG</td>
</tr>
<tr>
<td>TGD5010</td>
<td>MKD8 16105 KO D'</td>
<td>TGAAGCTGTCGTGACCTATGACA</td>
</tr>
<tr>
<td>TGD5011</td>
<td>MKD8 16105 KO D</td>
<td>CCGGTACGTCGTGACCTACCG</td>
</tr>
<tr>
<td>TGD5012</td>
<td>MKD8 21385 KO A</td>
<td>ACATAGTCCGACGAAACCTCG</td>
</tr>
<tr>
<td>TGD5013</td>
<td>MKD8 21385 KO A'</td>
<td>ATCGCCGGAGATCGTCGTCAACCA</td>
</tr>
<tr>
<td>TGD5014</td>
<td>MKD8 21385 KO B</td>
<td>TGGTGAGGGAGATGAGGTCTGAAGGCTGAAGCTGACCTACCG</td>
</tr>
<tr>
<td>TGD5015</td>
<td>MKD8 21385 KO C</td>
<td>GTTGAGGTGTAGGCTGCTGAAGGCAACTGCG CCCCGAGATCCGAGG</td>
</tr>
<tr>
<td>TGD5016</td>
<td>MKD8 21385 KO D'</td>
<td>TGAAGCTGTCGTGACCTATGACA</td>
</tr>
<tr>
<td>TGD5017</td>
<td>MKD8 21385 KO D</td>
<td>CCGGTACGTCGTGACCTACCG</td>
</tr>
<tr>
<td>TGD5018</td>
<td>MKD8 53960 KO A</td>
<td>CGACGCACCGTACGTTGAGGTC</td>
</tr>
<tr>
<td>TGD5019</td>
<td>MKD8 53960 KO A'</td>
<td>AGCGCTGTCGAGCCAGTCG</td>
</tr>
<tr>
<td>TGD5020</td>
<td>MKD8 53960 KO B</td>
<td>TGGTGAGGGAGATGAGGTCTGAAGGCAATCCGAGGGAGATCCGAGG</td>
</tr>
<tr>
<td>TGD5021</td>
<td>MKD8 53960 KO C</td>
<td>GTTGAGGTGTAGGCTGCTGAAGGCAACTGCG CCCCGAGATCCGAGG</td>
</tr>
<tr>
<td>TGD5022</td>
<td>MKD8 53960 KO D'</td>
<td>TGAAGCTGTCGTGACCTATGACA</td>
</tr>
<tr>
<td>TGD5023</td>
<td>MKD8 53960 KO D</td>
<td>CCGGTACGTCGTGACCTACCG</td>
</tr>
<tr>
<td>TGD5024</td>
<td>MKD8 15750 KO A</td>
<td>GAGGCGGGATACCGAGGTTTC</td>
</tr>
<tr>
<td>TGD5025</td>
<td>MKD8 15750 KO A'</td>
<td>CGTCCGCCAGCCAGGTC</td>
</tr>
<tr>
<td>TGD5026</td>
<td>MKD8 15750 KO B</td>
<td>TGGTGAGGGAGATGAGGTCTGAAGGCAACTGCG CCCCGAGATCCGAGG</td>
</tr>
<tr>
<td>TGD5027</td>
<td>MKD8 15750 KO C</td>
<td>GTTGAGGTGTAGGCTGCTGAAGGCAACTGCG CCCCGAGATCCGAGG</td>
</tr>
<tr>
<td>TGD5028</td>
<td>MKD8 15750 KO D'</td>
<td>TGAAGCTGTCGTGACCTATGACA</td>
</tr>
<tr>
<td>TGD5029</td>
<td>MKD8 15750 KO D</td>
<td>CCGGTACGTCGTGACCTACCG</td>
</tr>
<tr>
<td>TGD4993</td>
<td>MKD8 6925 Fwd C</td>
<td>GTTGAGGTGTAGGCTGCTGAAGGCAACTGATT CGACGCCAGCATCAG</td>
</tr>
<tr>
<td>TGD4994</td>
<td>MKD8 6925 Rev D'</td>
<td>TCGCAAGCTGCTGCTATATCGCACGTGCTGCGGTTC</td>
</tr>
<tr>
<td>TGD4995</td>
<td>MKD8 6925 Fwd F</td>
<td>CATCGAAGTCCACAAATACACCT</td>
</tr>
<tr>
<td>Barcode</td>
<td>Description</td>
<td>Primer Sequence 1</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>------------------</td>
</tr>
<tr>
<td>TGD4996</td>
<td>MKD8 6925 Rev B</td>
<td>TGGTGAGGAGATGAGGTCTGAAGGCACCCGT</td>
</tr>
<tr>
<td>TGD4997</td>
<td>MKD8 6925 Fwd A’</td>
<td>GTGGCCCACATCGGTGTTCGCA</td>
</tr>
<tr>
<td>TGD4998</td>
<td>MKD8 6925 Rev D</td>
<td>CTCGCGGCAAAAGAGGATCT</td>
</tr>
<tr>
<td>TGD4999</td>
<td>MKD8 6925 Fwd A</td>
<td>TCTCGCGGCCAAAGAGGATCT</td>
</tr>
<tr>
<td>TGD5303</td>
<td>2-hybrid 1687 cloning into pUAB300, BamHI site</td>
<td>AATTAA GGATCC</td>
</tr>
<tr>
<td>TGD5304</td>
<td>2-hybrid 1687 cloning into pUAB400, MfeI site</td>
<td>AATTAA CAATTG</td>
</tr>
<tr>
<td>TGD5305</td>
<td>2-hybrid 1687 cloning into pUAB300/400, HindIII site</td>
<td>AATTAA AAGCTT CTACAGGAGCGCCGCCG</td>
</tr>
<tr>
<td>TGD5306</td>
<td>2-hybrid 1686 cloning into pUAB300, BamHI site</td>
<td>AATTAA GGATCC</td>
</tr>
<tr>
<td>TGD5307</td>
<td>2-hybrid 1686 cloning into pUAB400, MfeI site</td>
<td>AATTAA CAATTG</td>
</tr>
<tr>
<td>TGD5308</td>
<td>2-hybrid 1686 cloning into pUAB300/400, HindIII site</td>
<td>AATTAA AAGCTT TCACAGGAGCGCCGCCG</td>
</tr>
<tr>
<td>TGD5309</td>
<td>2-hybrid 1685 cloning into pUAB300, BamHI site</td>
<td>AATTAA GGATCC</td>
</tr>
<tr>
<td>TGD5310</td>
<td>2-hybrid 1685 cloning into pUAB400, MfeI site</td>
<td>AATTAA CAATTG</td>
</tr>
<tr>
<td>TGD5311</td>
<td>2-hybrid 1685 cloning into pUAB300/400, HindIII site</td>
<td>AATTAA AAGCTT TCACAGGAGCGCCGCCG</td>
</tr>
<tr>
<td>TGD5312</td>
<td>2-hybrid 1560 cloning into pUAB300, BamHI site</td>
<td>AATTAA GGATCC</td>
</tr>
<tr>
<td>TGD5313</td>
<td>2-hybrid 1560 cloning into pUAB400, MfeI site</td>
<td>AATTAA CAATTG</td>
</tr>
<tr>
<td>TGD5314</td>
<td>2-hybrid 1560 cloning into pUAB300/400, HindIII site</td>
<td>AATTAA AAGCTT TCACAGGAGCGCCGCCG</td>
</tr>
<tr>
<td>TGD5315</td>
<td>2-hybrid 15750 cloning into pUAB300, BamHI site</td>
<td>AATTAA GGATCC</td>
</tr>
<tr>
<td>TGD5316</td>
<td>2-hybrid 15750 cloning into pUAB400, MfeI site</td>
<td>AATTAA CAATTG</td>
</tr>
<tr>
<td>TGD5317</td>
<td>2-hybrid 15750 cloning into pUAB300/400, HindIII site</td>
<td>AATTAA AAGCTT TCACAGGAGCGCCGCCG</td>
</tr>
<tr>
<td>TGD5318</td>
<td>2-hybrid esxU cloning into pUAB300, BamHI site</td>
<td>AATTAA GGATCC</td>
</tr>
<tr>
<td>TGD5319</td>
<td>2-hybrid esxU cloning into pUAB400, MfeI site</td>
<td>AATTAA CAATTG</td>
</tr>
<tr>
<td>TGD5320</td>
<td>2-hybrid esxU cloning into pUAB300/400, HindIII site</td>
<td>AATTAA AAGCTT TCACAGGAGCGCCGCCG</td>
</tr>
<tr>
<td>TGD5321</td>
<td>2-hybrid esxT cloning into pUAB300, BamHI site</td>
<td>AATTAA GGATCC</td>
</tr>
<tr>
<td>TGD5322</td>
<td>2-hybrid esxT cloning into pUAB400, MfeI site</td>
<td>AATTAA CAATTG</td>
</tr>
<tr>
<td>TGD5323</td>
<td>2-hybrid esxT cloning into pUAB300/400, HindIII site</td>
<td>AATTAA AAGCTT TCACAGGAGCGCCGCCG</td>
</tr>
<tr>
<td>TGD5397</td>
<td>REV Primer to InFuse MKD8_1536 into</td>
<td>acgtGACATCGATAAGCCTTCACGACAGTCGACG</td>
</tr>
</tbody>
</table>
| TGD5398 | FWD Primer to amplify/ligate MKD8\_1537 into pUAB300, BamHI site | AATTAA GGATCC
|         |                                                                   | GTGCAGCTGTCGTGACCGGT
| TGD5399 | FWD Primer to amplify/ligate MKD8\_1537 into pUAB400, MfeI site  | AATTAA CAATTG
|         |                                                                   | GTGCAGCTGTCGTGACCGGT
| TGD5400 | REV Primer to amplify/ligate MKD8\_1537 into pUAB300/400, HindIII site | AATTAA AAGCTT
|         |                                                                   | CTAAGAGGCCGCGTGTCGCGA
| TGD5401 | FWD Primer to amplify/ligate MKD8\_6925 into pUAB300/400, BamHI site | AATTAA GGATCC
|         |                                                                   | ATGGACTACTGCGCTGGGTGAC
| TGD5402 | FWD Primer to amplify/ligate MKD8\_6925 into pUAB300/400, MfeI site | AATTAA CAATTG
|         |                                                                   | ATGGACTACTGCGCTGGGTGAC
| TGD5403 | REV Primer to amplify/ligate MKD8\_6925 into pUAB300/400, HindIII site | AATTAA AAGCTT
|         |                                                                   | CTAAGAGGCCGCGTGTCGCGA
| TGD5415 | 6925 qPCR FWD Primer (use with TGD5431 for qPCR)                  | ATC TCG ACG ACG ACG GTG TGG
| TGD5416 | 6925 qPCR REV Primer DO NOT USE WITH TGD5415 -> TGD5416 heterodimerized w/TGD5415 | cacataacctgagccgtgctg
| TGD5417 | 1534 qPCR FWD Primer                                              | CCGTCTCGGCGGACTGATGC
| TGD5418 | 1534 qPCR REV Primer                                              | G CCCACGTGATCTCGTCTGC
| TGD5431 | 6925 qPCR REV Primer, new sequence --> first primer heterodimerized w/TGD5415 | TAGTAGCTCTCAGGGCCGCGG
| TGD5481 | MKD8\_5565 FWD qRT-PCR primer                                      | Gagccgatgcatcccttctt
| TGD5482 | MKD8\_5565 REV qRT-PCR primer                                      | cttgcccagacgagacgacg
| TGD5483 | MKD8\_5565 FWD qRT-PCR primer                                      | AGGCCATGGACCCCTTCCTC
| TGD5484 | MKD8\_5565 REV qRT-PCR primer                                      | CCTGGCCCAGAGAGAAGT
| TGD5496 | MKD8\_15750 internal sequencing primer for M-PFC vectors           | GAGATCTCCTCGGCCGACAC
<table>
<thead>
<tr>
<th>TGD5497</th>
<th>MKD8_1537 internal sequencing primer for M-PFC vectors</th>
<th>CACCCCTGCACGTGAGTT</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGD5498</td>
<td>MKD8_1686 internal sequencing primer for M-PFC vectors</td>
<td>GACGGCAGCACCAGA</td>
</tr>
<tr>
<td>TGD5497</td>
<td>MKD8_1537 internal sequencing primer for M-PFC vectors</td>
<td>CACCCCTGCACGTGAGTT</td>
</tr>
<tr>
<td>TGD5498</td>
<td>MKD8_1686 internal sequencing primer for M-PFC vectors</td>
<td>GACGGCAGCACCAGA</td>
</tr>
<tr>
<td>TGD5365</td>
<td>FWD Sequencing primer for pUAB300 located in DHFR F[1,2]</td>
<td>ATGATGCCTTAAGACTTATTGAACAA</td>
</tr>
<tr>
<td>TGD5366</td>
<td>FWD Sequencing primer for pUAB400 located in DHFR F[3]</td>
<td>GCCGTCTCCTCTGAGGTCCAG</td>
</tr>
</tbody>
</table>
References


