Regulation of the ESX-5 Secretion System in *Mycobacterium tuberculosis*

A THESIS
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Sarah Rae Elliott

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Adviser: Anna Tischler, PhD

May 2017

Acknowledgements

Firstly, I need to thank my advisor, Anna Tischler, for being such a wonderful mentor throughout my graduate career. Thank you for creating an environment where I was allowed to be as independent as I wanted to be, while still managing to be available to answer my endless questions. Your high standards have pushed me to become a better writer and scientist, and for that I will be forever grateful.

I feel very lucky to have had the opportunity to work with such wonderful people in the Tischler lab. Thank you to everyone for making our lab environment so much fun! Sarah Namugenyi is quite possibly the kindest, most giving person I've had the pleasure of meeting, thank you for all of your help through the years. You certainly embody the core values of our program. Dylan White, thank you for being such a fun lab partner and bench buddy. I appreciate our conversations about science, beer, and whatever happened to be annoying me at the moment. You are a great scientist and I look forward to seeing what you accomplish in the future. Alisha Aagesen, I always enjoyed our mid-day coffee runs, they really helped get me through the evenings. Everything in the lab was made easier thanks to Alyssa Brokaw, thanks for all of the last-minute reagent prep and your endlessly positive attitude.

I must also thank members of our brother lab. Thanks of course to Tony Baughn, I have always enjoyed our conversations, whether about science or some imprecise language you caught me using. Nick Dillon, thanks for being my main MICaB partner through the years. It was a blast to share lab meetings, lab space and lunch orders with you. Thanks also to Josh Thiede and Yusuke Minato for always being available to discuss scientific ideas, and for distracting me with current events.

A special thanks to the Dunny lab as well. Gary Dunny, thanks for mentoring me since I was a young, naïve undergrad! Thanks as well to Dawn, who knows everything and is always happy to share knowledge and reagents. A huge thanks to Jenny Dale is required. I can't thank you enough, my gene regulation guru, for always taking the time to talk through my hair-brained ideas and giving me thoughtful feedback. I also feel lucky to call you my friend, so I must also thank you for our movie nights and fun outings, which were always memorable.

I feel very lucky to have had such a wonderful, supportive thesis committee. Thank you, Anna Tischler, Sandy Armstrong, Tony Baughn and Gary Dunny. Your

thoughtful input and guidance is very much appreciated. Thank you for your letters of support. Sandy, thank you so much for being my chair. You always took the time to help me talk through the major milestones throughout my graduate career, and I am so thankful for that.

I have met so many wonderful people in the Microbiology, Immunology and Cancer Biology PhD program, too many to mention them all here. Thank you to Emily Thompson and Nick Dillon for being two thirds of our triumphant triumvirate, I couldn't have asked for better partners as fellow student representatives. I feel very lucky to call both of you friends. Big T, thanks for being such a wonderful human being. Your constant support and cheerleading has made this crazy adventure much easier to bear. I have escaped this labyrinth with much more sanity than I expected, largely due to your thoughtful advice. Thanks for climbing mountains with me. Jenn McCurtain, you always liven up any event, thanks for being such a fun person! Ladies, may we conquer every national park together in the future!

Of course, I would be remiss if I didn't acknowledge the wonderful Louise Shand and Megan Ruf. Thanks for keeping the MICaB program, and myself, on track. Thank you for answering all of my questions, and thanks for all of your support over the years. And a big thank you for your guidance during my time as student representative, you made a sometimes overwhelming situation seem manageable.

Thank you so much to my wonderful, supportive, crazy family. Thank you to my parents who have always supported me and cheered me on. Your love and belief in me has kept me on track. Mom, thanks for being such a strong role-model. I certainly owe my strength and stubbornness to you, and both traits have served me well in grad school! Thanks of course to my sisters Renee and Anissa, who always at least pretend to be interested in my work. Thanks so much for always letting me vent my frustrations and share my triumphs. Thanks to my brothers, Rick and Matt. I can always count on you to make me laugh, even when I felt so stressed I thought I might explode. I am also lucky enough to have wonderful in-laws, a rare occurrence from what I'm told. Thank you, Bill and Annette, for cheering me on over the years. Your kindness and support seem never-ending, and for that I am humbled and grateful.

Lastly, but most importantly, I owe more thanks to my husband than I can ever hope to repay, though I promise to try my best. You have been my biggest supporter, my best champion, my loudest cheerleader and my constant sounding board. Thank you so

much for your love and support. You believed in me when I doubted myself. Because of you, I am a better person. Thank you for taking this crazy ride with me. I am so excited for our next adventure!

Dedication

This work is dedicated to my best friend, my partner in crime, my husband, my heart. Thank you for always believing in me. Your endless love and support has made all the difference.

Abstract

Mycobacterium tuberculosis (*Mtb*) is one of the most prolific bacterial pathogens in the history of human disease. Robert Koch discovered that *Mtb* was the causative agent of the disease tuberculosis in 1882, and despite intensive research and major advances, *Mtb* represents a major global health threat today. Worldwide in 2015, there were 10.4 million newly diagnosed active cases and 1.8 million deaths attributed to this infection. Bacterial pathogens often secrete factors to promote survival during infection, and *Mtb* is no exception. *Mtb* has evolved a unique, diderm cell membrane, which contributes to the ability of the bacterium to resist host immune responses. However, this hydrophobic barrier also presents an obstacle for the export of factors critical to success of the organism. Mycobacteria, including *Mtb*, have evolved the Type VII ESX secretion systems to facilitate protein export across the complex membrane. Three ESX systems have been implicated in *Mtb* pathogenesis, ESX-1, -3 and -5. While the regulatory mechanisms and biological functions for both ESX-1 and ESX-3 are well-defined, little was known about ESX-5 aside from a general role in *Mtb* virulence.

The work described in chapter 2 reveals that ESX-5 secretion is directly regulated at the transcriptional level by the Pst/SenX3-RegX3 system in response to inorganic phosphate (P_i) limitation. RegX3, the response regulator, is normally activated when P_i is scarce. Disruption of a transmembrane component of the Pst P_i uptake system, through deletion of *pstA1*, causes constitutive activation of RegX3. We observed overexpression of *esx-5* transcripts and hyper-secretion of the ESX-5 substrates EsxN and PPE41 in the *Mtb* Δ*pstA1* mutant, and this response requires RegX3. In wild-type Erdman (WT) *Mtb*, transcription of *esx-5* genes and secretion of ESX-5 proteins was induced by P_i limitation in a RegX3-dependent manner. Using electrophoretic mobility shift assays (EMSA), we found that RegX3 directly binds to a segment of DNA within the *esx-5* locus, demonstrating that regulation of ESX-5 mediated by the Pst/SenX3-RegX3 system occurs directly.

Experiments outlined in chapter 3 expand on the work reported in chapter 2. Using *in vitro* EMSAs, we defined the RegX3 binding sequence located within the intergenic region between *ppe27* and *pe19* in the *esx-5* locus. RegX3 is a global response regulator, and targeted mutation of the *esx-5* binding site sequence uncouples

the secretion system from the myriad effects mediated by RegX3 throughout the cell. We found that mutating the *esx-5* RegX3 binding site sequence reversed expression of *esx-5* transcripts and secretion of EsxN and PPE41 in WT *Mtb* during P_i limitation. Similarly, *esx-5* overexpression and ESX-5 hyper-secretion was suppressed in the Δ*pstA1* mutant when the RegX3 binding site sequence was mutated or deleted. We then tested the importance of RegX3-mediated regulation of ESX-5 for *Mtb* virulence by infecting C57BL/6 and IrgM1$^{-/-}$ mice with a binding site mutant. Notably, deletion of the *esx-5* RegX3 binding site partially restored virulence to the attenuated Δ*pstA1* mutant. Our findings demonstrate that precise regulation of ESX-5 is critical for full *Mtb* virulence. Further, hyper-secretion of antigenic ESX-5 substrates sensitizes Δ*pstA1* bacteria to host responses, suggesting that one or more of these aberrantly secreted proteins is responsible for attenuation.

We next sought to determine whether aberrant hyper-secretion of one ESX-5 secreted factor, EsxN, sensitizes Δ*pstA1* bacteria to host immune responses. Previous work has shown that the Δ*pstA1* mutant is attenuated in immune competent C57BL/6 mice and immune compromised IrgM1$^{-/-}$ and Nos2$^{-/-}$ mice, while experiments described in chapter 3 demonstrated that attenuation of the Δ*pstA1* mutant in C57BL/6 and IrgM1$^{-/-}$ mice was specifically due to constitutive activation of *esx-5*. Experiments in Chapter 4 evaluate the contribution of EsxN to *Mtb* virulence. Deletion of *esxN* did not reverse Δ*pstA1* mutant sensitivity to reactive oxygen species, acidic pH or cell wall stress *in vitro*. However, WT bacteria were more sensitive to reactive nitrogen stress when *esxN* was deleted, suggesting a role for EsxN in resistance to reactive nitrogen species (RNS). We found that *esxN* does not suppress the Δ*pstA1* mutant virulence defect in C57BL/6 or IrgM1$^{-/-}$ mice. However, deletion of *esxN* in the Δ*pstA1* mutant did partially reverse the replication and virulence defect in Nos2$^{-/-}$ mice, indicating hyper-secretion of EsxN sensitizes Δ*pstA1* bacteria to immune responses other than RNS production in these mice. Aberrant hyper-secretion of EsxN may influence sensitivity to other host responses in the Δ*pstA1* mutant. EsxN seems to be required for survival during RNS stress *in vitro* in WT *Mtb*. Further work will be required to tease apart the potential role EsxN plays in RNS resistance.

The work described in this thesis expands the knowledge of ESX-5 secretion system biology, and *Mtb* protein secretion in general. We have uncovered the

mechanism of regulation, and also revealed a relevant environmental signal, P_i limitation, that activates this system. We have demonstrated that regulation of ESX-5 by the Pst/SenX3-RegX3 system occurs directly by identifying the RegX3 binding site sequence within the *esx-5* locus. Using targeted mutation of the RegX3 binding sequence, we have shown that dysregulation of ESX-5 activity has a detrimental impact on *Mtb* virulence. These findings highlight the importance of proper regulation of the ESX-5 system to *Mtb* pathogenesis. An understanding of the regulatory mechanism and environmental signals that activate ESX-5 during infection provides an important framework for future studies to elucidate the functional role of this system.

Table of Contents

List of Tables

List of figures

Chapter 1

Introduction

History of tuberculosis

Tuberculosis (TB) infections have been plaguing mankind for millennia, but the 'tubercle bacillus' *Mycobacterium tuberculosis (Mtb)* was discovered by Robert Koch as the causative agent barely a century ago. At the time of this momentous discovery in 1882, 1 in 7 people in the world were dying from TB infections [1]. Koch's discovery was the first to conclusively demonstrate a disease was caused by infection with a specific bacterium, and is considered a triumph in the field of microbiology. Hippocrates hypothesized TB disease was hereditary, while others suspected the disease was spontaneous, and even supernatural causes were invoked to explain the origin of this illness [2]. Koch's finding cleared up centuries of speculation surrounding the cause of this mysterious scourge of a disease. Fittingly, Robert Koch was awarded the Nobel Prize in Medicine in 1905 for his discoveries relating to TB [2].

The long and storied history of *Mtb* and resulting TB disease is indelibly intertwined with our own. Understanding the history of *Mtb*, and how this bacterium has evolved alongside humans, will aid our current and future endeavors aimed at eradicating this plague, or at the very least motivate our efforts.

Although mycobacteria have existed for around 150 million years, *Mtb* is a comparatively young species, having evolved only 150,000 years ago, and has existed with humans for at least 15,000 years, since the Neolithic era [3]. While we can trace the evolution of *Mtb* through the centuries using phylogeny, in practice it can be difficult to trace historical cases of TB disease. However, this endeavor is not impossible, and there is an extensive body of literature dedicated to uncovering evidence of TB disease through the ages. This evidence takes many forms, including written accounts, signs of disease depicted in art, and molecular identification of mycobacterial DNA from mummified tissue.

Research into the oldest cases of human *Mtb* infection is greatly aided by the fact that TB disease can be recognized and diagnosed in skeletal remains due to lesions in the bone, especially in the spine. This occurs because *Mtb* can infect any area of the body; these infections are called extra-pulmonary TB, though the lung is a more common site of infection in current times. Skeletal abnormalities, sometimes called Potts deformities, caused by *Mtb* infection have been found in Egyptian mummies, and these types of bone defects are also depicted in their art [4]. The lesions have been verified to

be caused by *Mtb* using molecular techniques from samples obtained from the mummies [5]. There is also evidence for *Mtb* induced Potts deformities in Peruvian mummies, which has likewise been verified from tissue samples using PCR [6]. This evidence demonstrates that *Mtb* must have traveled along with humans during the major migrations which populated the Americas, and speaks to the long, intimate history we have with this organism [6]. The oldest evidence of *Mtb* infecting humans uncovered so far comes from the 9000 year old remains of a Neolithic woman and her child [7]. It is thought that TB became prolific in the Neolithic era because of animal domestication and increased density of human populations, which allowed for efficient dissemination of the disease. Scholars believe there are references to TB disease in the bible, though it is referred to as schachepheth, an ancient Hebrew word meaning 'wasting disease' [8]. In fact, TB has been called many things through the years, including consumption, the King's evil, lupus vulparis, and pthisis, to name a few [9].

General *Mtb* Background

Mtb is a facultative intracellular bacterial pathogen, and causes the pulmonary infection TB. According to the World Health Organization (WHO) global report, there were 10.4 million new active cases of TB, and 1.8 million deaths attributed to this infection in 2015, worldwide. TB was one of the top 10 leading causes of death worldwide in 2015, killing more people than HIV and malaria combined. Additionally, it is estimated that ¼ of the world's population is latently infected with *Mtb* [10]. There is a 5-10% lifetime risk of reactivation, so this latently infected population of around 2 billion people serves as a massive reservoir for continued transmission of *Mtb*.

TB is caused by a group of closely related mycobacterial species, called the *Mtb* complex (MTBC). There are many members of the MTBC, but *M. tuberculosis, M. leprae, M. marinum, M. ulcerans, M. avium* and *M. bovis* are likely the most well-known [11]. While some mycobacterial species are non-pathogenic and can be found in the environment, all members of the MTBC are obligate pathogens. Today, the principal cause of human TB is *Mtb*, and indeed there is no known reservoir for this bacterium outside of humans. *M. bovis* has a broader host range, and is the principal cause of TB in other animals [9].

Treatment of *Mtb*

Duration of TB treatment is significantly longer than for most other bacterial infections. *Mtb* infections are difficult to cure due to the ability of the bacterium to enter a dormant non-replicating state [12]. In this dormant state, *Mtb* is phenotypically resistant to many antibiotics that target replicating bacteria, leading to the requirement for long-term therapy with multiple antibiotics to effectively treat the infection.

Treatment of an active *Mtb* infection requires a cocktail of 4 drugs for 2 months (isoniazid, ethambutol, rifampicin, and pyrazinamide), followed by continuous administration of 2 drugs for another 4 months (isoniazid & rifampicin) [10]. Given that *Mtb* is endemic in many poverty-stricken regions, this regimen is nearly impossible to sustain. Patient non-compliance is unfortunately a common occurrence, and along with other factors leads to the emergence of drug resistant strains of *Mtb* [13]. Treating drug resistant *Mtb* increases the duration of chemotherapy to up to 2 years [14]. The increasing emergence of drug resistant *Mtb* is a massive global health concern.

***Mtb* Immunology**

Immunological life cycle of an Mtb infection

Mtb is transmitted through the aerosol route, and thus the lung is the primary site of infection, although the bacterium has the ability to survive in almost all tissues. The infectious dose is thought to be between 1-10 bacteria, which is incredibly low compared to other bacterial infections [15]. Upon inhalation of *Mtb*, innate immune cells are recruited to the site of infection. *Mtb* bacilli are primarily phagocytosed by resident alveolar macrophages, although other phagocytic cells are involved, including neutrophils and dendritic cells (DCs) [16]. In most bacterial infections, recruitment of innate immune cells is beneficial to the host, since neutrophils and macrophages are quite efficient at killing microbial invaders. However, during *Mtb* infection, recruitment of phagocytic cells is thought to actually benefit the bacterium, providing a protective compartment for replication and immune evasion. Macrophages especially play a dichotomous role during *Mtb* infection: they are critical for the activation of protective responses, including both innate and adaptive immunity, but they simultaneously provide

the major host cell niche for *Mtb* growth and survival. Since phagocytes regularly travel through the blood and lymphatic systems, when infected with *Mtb*, these cells act like a Trojan horse, and allow the bacterium to disseminate throughout the lung, and eventually other body sites [17, 18]. *Mtb* has evolved an arsenal to counteract host defense mechanisms which allow the bacterium to survive within phagocytes, ultimately leading to spread of the infection [19].

However, all is not lost in our fight against *Mtb*, as it is thought that fewer than 10% of infected individuals develop active primary infection, and the majority are able to contain the infection within granulomas [20]. Granulomas, or tubercles, are highly ordered structures employed by the host response to serve as a containment system that both restricts bacterial growth and concentrates immune cells to the site of infection. Basically, if the immune system is unable to clear the infection, it settles for containing it. Resident macrophages engulf the invading bacilli, and in turn initiate an inflammatory response aimed at recruiting backup to help eliminate the threat. One outcome of this chain of events is the formation of the granuloma. A granuloma consists of *Mtb*-infected macrophages at the center, surrounded by foamy macrophages and other innate immune cells, which are in turn surrounded by a layer of *Mtb* specific T-cells. Eventually, a fibrous cuff composed of collagen and other components walls off the whole area, encasing and physically containing the bacilli [21, 22]. At this point, the infection is considered to be latent. Latent *Mtb* infections can persist for decades, and reactivate to cause active disease when the immune system is no longer able to control the bacteria. Factors contributing to reactivation include co-infection with HIV, NFK-β or TNF-α inhibitors, obesity, diabetes, and smoking [19].

Host immune defense mechanisms

The immune system utilizes pattern recognition receptors (PRRs) to detect the presence of invading microbes, and these receptors trigger cytokine responses to control *Mtb* replication. Toll-like receptors (TLRs) are a major class of PRR, and TLR-2 (lipoprotein) and TLR-9 (mycobacterial DNA) are most commonly activated during *Mtb* infection [23, 24]. Additional innate PRRs are also involved, including NOD and NLRP3 receptors, which recognize peptidoglycan and some ESX-1 secreted substrates, respectively [24]. Once *Mtb* has been recognized by host PRRs, many pro-inflammatory

cytokines are produced, including tumor necrosis factor α (TNF-α), interleukin-1 β (IL-1β), IL-6, IL-12, and type I interferons [18, 25]. This pro-inflammatory response leads to T-cell activation, subsequent secretion of IFN-γ by Th1 T-cells, ultimately leading to activated macrophages with enhanced anti-microbial functions [25].

Once a macrophage has been stimulated by IFN-γ, reactive nitrogen species (RNS) are produced by nitric oxide synthase 2 (NOS2) and reactive oxygen species (ROS) are produced by the NAPDH oxidase (NOX2) within the cell [25]. The cytokine TNF-α also stimulates the production of ROS in macrophages [26]. Bacteria engulfed by phagocytes are deposited in a membrane-bound compartment called the phagosome. All RNS and ROS are generated in the host cytosol and must cross the phagosomal membrane where they spontaneously form highly reactive intermediates that destroy bacterial membrane lipids and DNA via oxidation; RNS intermediates additionally target iron-sulfur clusters of bacterial enzymes [25, 27]. Humans carrying mutations in the genes encoding NOS2 or NOX2 have an impaired ability to control *Mtb* and typically develop much more severe TB disease [25].

Another important host mechanism aimed at eliminating invading microbes is the maturation and acidification of the phagosomal compartment. The phagosome undergoes a sequential series of fusion events with endosomes and lysosomes to reach maturity and peak anti-microbial activity [27]. Phagosomal maturation requires a host of factors, including Rab GTPases, phosphoinisitides, kinases, and the GTPase IrgM1 (previously called LRG-47) [25, 27, 28]. Upon successful maturation, phagosomes become acidified to pH 4.5-5 [29]. The lowered pH triggers the acquisition of additional antimicrobial peptides and proteases through fusion with lysosomes containing degradative enzymes [29]. Importantly, loss of IrgM1 causes defects in phagosomal trafficking and maturation, and IrgM1$^{-/-}$ mice are highly susceptible to *Mtb* infection [28].

Controlling host cell death is another important immune defense mechanism utilized to control the replication and spread of invading bacteria. There are multiple pathways governing host cell death, the two most well-studied being apoptosis and necrosis [29]. Apoptosis is an ordered cell death that is typically considered anti-inflammatory because cell membrane integrity is maintained. Apoptosis allows for retention of *Mtb* within apoptotic bodies, which can be efficiently phagocytosed by new macrophages [30]. This process is generally detrimental for *Mtb* survival and dissemination, because the response enhances antigen presentation and T-cell priming

[29]. Necrosis is generally favorable to bacterial survival, since host cell membranes are disrupted, allowing bacterial dissemination [29]. Necrosis is pro-inflammatory, and the resulting recruitment of phagocytic cells is beneficial to *Mtb*, as described previously.

Mtb evasion strategies

Mtb infections have been confounding researchers since Koch demonstrated this bacterium is the causative agent of TB back in 1882 [1]. By all accounts, most humans mount an appropriate immune response to *Mtb* infection; however the response is typically ineffective at killing *Mtb*. This is because *Mtb* has evolved a myriad of ways to counteract host defenses.

One critical way *Mtb* subverts the host response is through its ability to impede and/or evade acidification and maturation of the phagosome, and instead survive and replicate within this typically deadly environment [31, 32]. When acidification of the phagosome is blocked, the phagosomal compartment maintains a pH around 6.2 due to decreased fusion with lysosomes, and *Mtb* is able to survive quite comfortably under these conditions [33]. The mycobacterial proteins MavC and MtbA maintain recruitment of host proteins associated with early stages of phagosomal maturation, such as Rab5, and this precludes host recruitment of proteins involved in later stages of maturation, like Rab7 [27]. Also, certain mycobacterial membrane lipids, such as mannose-capped lipoarabinomannan (manLAM), can interact with the host mannose receptor, to prevent phagosomal fusion with the lysosome [27, 34]. Moreover, *Mtb* also has the ability to escape the phagosome to the cytosol, a process requiring the ESX-1 secretion system [35, 36]. *Mtb* also promotes necrotic cell death as opposed to apoptosis, which favors bacterial survival [36, 37].

Mtb is able to evade clearance through many mechanisms, but one important strategy employed by the bacterium serves to alter the adaptive immune response through manipulation of antigen presentation. Host cells display foreign antigens on their surfaces in the context of MHC I or II molecules, which can then be recognized by CD8+ or CD4+ T-cells, respectively. *Mtb* is able to manipulate antigen transfer by limiting the bacterial antigens presented via the MHC II pathway, which serves to lessen the ability of adaptive immune cells to recognize *Mtb*-infected host cells [38]. The ability of *Mtb* to inhibit phagosomal acidification also alters antigen presentation because if there is no

fusion with the lysosome, antigens are not processed and therefore not displayed [33]. *Mtb* can also downregulate expression of MHC II on host cell surfaces by several mechanisms, including activation of TLR-2 by LpqH [39-41] and manipulation of the intracellular MHC II loading pathway [42]. Reduction in MHC II expression results in decreased recognition of *Mtb* and therefore increased bacterial survival.

Mtb delays the onset of adaptive immunity through several mechanisms. The bacterium hampers the migration of *Mtb*-infected antigen presenting cells to the lymph nodes for T-cell priming [43], and prevents neutrophil apoptosis [44]. However, even with the onset of adaptive immunity, the response seems to only arrest the growth of *Mtb*, and is typically unable to completely clear the infection [19]. There are many contributing factors thought to enable continued survival of *Mtb* after adaptive immunity is initiated, including, but not limited to, reduced antigen presentation through MHC II [39], and downregulation of anti-mycobacterial host responses by regulatory T-cells [45]. All of these factors are exacerbated by the ability of *Mtb* to resist many host-mediated defense mechanisms.

Reactivation of Mtb

There are many factors that influence reactivation of a latent *Mtb* infection. One notable factor involves a decreased CD4+ T-cell response. CD4+ T-cells are essential for the control of *Mtb* infections, and diminished numbers lead to reactivation. Reduced CD4+ T-cell populations can occur due to many factors, though HIV co-infection is a major cause of T-cell depletion [46]. The cytokines IFN-γ and TNF-α are required to maintain integrity of *Mtb* granulomas [47]. Both cytokines are important because they serve to activate macrophages. Another described mechanism of reactivation is neutralization of TNF-α with monoclonal-antibodies to treat some cancers and autoimmune diseases [48]. Several other factors that contribute to *Mtb* reactivation include: diabetes mellitus, smoking, obesity, old age, poor nutrition and stress [19].

Protein export systems

Protein secretion is an essential task that must be efficiently performed by all bacteria. Bacteria are surrounded by cell membranes, either one or two, depending on

the organism, and these barriers serve as protection against the harsh external environment. However, the cell membrane(s) also present an obstacle that must be overcome to successfully secrete protein substrates vital for survival, whether to acquire nutrients or to secrete virulence factors.

In both Gram-positive and Gram-negative organisms, the bulk of protein transport is conducted by the general housekeeping, or Sec, secretion system. The Sec system is present in all bacteria, including *Mtb*, and is essential for viability. The Sec export system is comprised of the SecYEG membrane channel, and SecA, which acts as both the Sec signal peptide receptor and the ATPase that powers transport of proteins through the SecYEG channel [49-52]. Together, SecYEG and SecA are referred to as the Sec translocase [53]. Substrates exported by the Sec system are in an unfolded state for translocation across the inner membrane, and are targeted for transport by an N-terminal signal sequence. Once the unfolded peptide is transported across the inner membrane, the signal sequence is cleaved by signal peptidases bound to the membrane to form the mature protein [54]. In most Gram-negative bacteria, proteins secreted by the Sec system require association with cytosolic chaperones, such as SecB, to keep them in a translocation competent state [55]. SecB also helps target peptides to the inner membrane due to a high affinity for SecA [56]. However, as can be expected given the diversity of microbial life, not all Sec secreted proteins require SecB, and in those cases it is thought other protein chaperones are required [57].

While the Sec system transports secretory proteins across the cytoplasmic membrane, it also functions to insert proteins into the inner membrane. Membrane proteins possess transmembrane domains. When these domains are present in nascent proteins, SecY opens a lateral gate which allows the transmembrane domains of the integral membrane proteins to be passed sideways from the SecYEG channel, ultimately inserting the protein in the inner membrane [58, 59].

Like the Sec system, the twin-arginine translocation (Tat) pathway is also present in both Gram-positive and Gram-negative bacteria including *Mtb*, but unlike the Sec system it is not ubiquitous [60]. Notably, the Tat system is commonly used to export virulence factors in bacterial pathogens. Though the Tat export system is not essential in most organisms, it is essential in *Mtb*, where it is involved in both virulence and drug resistance [61]. TatA, TatB and TatC comprise the Tat export machinery. TatA forms the transport channel, while TatB and TatC form a complex that binds Tat-exported proteins

[62-64]. Unlike the Sec system, there is no ATPase subunit, and proton motive force is used to power Tat protein export [65]. The Tat system transports proteins across the cytoplasmic membrane, which are targeted for export via N-terminal signal peptide sequences which contain the canonical Tat motif, twin-arginine residues [66]. Another notable difference from the Sec system, Tat transports pre-folded proteins. Misfolded proteins are often unable to be transported, suggesting the Tat system also has a proof-reading function [67].

Protein export in *Mtb*

Mtb has evolved a unique and complex diderm cell envelope structure. Though mycobacteria are technically classified as Gram-positive bacteria, they possess an outer membrane containing unusual fatty acids, called mycolic acids, that form a hydrophobic barrier equivalent in function to the Gram-negative outer membrane [68]. While *Mtb* encodes the ubiquitous Sec and Tat secretion systems, as well as the SecA2 accessory system, these systems only span the inner membrane [69, 70]. In Gram-positive bacteria, proteins transported using the Sec or Tat systems can be directly secreted or translocated. Gram-negative bacteria, however, typically employ additional secretion systems, types I-VI, to export proteins across the additional outer membrane [71]. For example, in Gram-negative bacteria, the type II secretion system exports substrates first translocated across the inner membrane by the Sec or Tat systems [72]. Due to the presence of the hydrophobic outer membrane that other Gram-positive bacteria do not possess, mycobacteria require additional mechanisms to export proteins outside the cell from the periplasmic-like space. To that end, mycobacteria have evolved the ESX protein transport systems to facilitate protein export across its complex diderm membrane.

ESX Secretion Systems

The ESX secretion systems are classified as type VII; Gram-negative nomenclature was chosen to highlight the fact that the mycobacterial cell envelope is functionally similar to that of Gram-negative bacteria [73]. There are five ESX systems, denoted ESX-1 to ESX-5, and the numbers indicate the order the systems were

discovered, not the order in which they evolved. The ESX secretion systems were first discovered in mycobacteria, although a small number of other Gram-positive bacteria, such as *Staphylococcus aureus,* also encode rudimentary versions of these systems [74, 75]. The first clue to uncovering the presence of type VII secretion systems in mycobacteria came with the detection of the 6 kDa early secreted antigenic target (ESAT-6, or EsxA) from *Mtb* culture supernatant [76]. The discovery of a secreted protein without an obvious secretion signal hinted towards an as yet unknown transport system, and a comparative genomics screen revealed a virulence-associated region, the *esx-1* locus [77]. The field of type VII secretion is relatively new, since secretion through ESX-1, the first system to be discovered, was only demonstrated in 2003 [68].

Phylogenetic analysis revealed the five ESX systems evolved due to gene duplication events, mediated by plasmids encoding ESX gene clusters, which allowed for their subsequent divergence of function [78, 79]. The ESX systems likely evolved in the following order: ESX-4, -3, -1, -2 and -5 [79]. In *Mtb,* only ESX-1, -3 and -5 have been demonstrated to be functional, while active secretion by the ESX-2 and -4 systems has yet to be detected. It is possible the signals or growth conditions that activate secretion via these two systems are not present in standard laboratory growth medium, or perhaps these loci are no longer functional.

Likely a consequence of how the ESX systems evolved, the genomic organization of each locus follows a similar basic pattern, though the exact number of genes varies. All *esx* loci encode at least one pair of *esx* genes, the canonical ESX secreted substrates [80]. Esx proteins belong to the WxG100 family, which are so named for conserved tryptophan and glycine residues that are separated by a variable amino acid. These proteins are typically around 100 amino acids in length and secreted as a heterodimer [68, 81]. All *esx* loci except ESX-4 encode members of two unique familes of proteins found exclusively in mycobacteria, termed PE and PPE, and will be discussed in more detail in another section. Flanking each locus are genes encoding the components of the secretion machinery itself [80]. All five ESX systems possess the core secretion machinery components EccB, EccC, EccD and MycP; all systems but ESX-4 also encode the EccE component (**Figure 1.1)** [82, 83]. The core membrane complex is localized exclusively within the inner membrane, and is comprised of EccB, EccC, EccD, and EccE [82, 83]. EccB is a transmembrane protein, EccC is an ATPase predicted to provide energy for protein transport, and EccD is another transmembrane

protein which is predicted to form the channel in the inner membrane to facilitate protein secretion [84]. EccE is located on the periphery of the membrane complex and plays an unknown role in secretion [83]. MycP is a membrane-tethered protease that may play a role in processing substrates [84]. However, the protease activity of MycP is dispensable for ESX secretion, and recent evidence suggests this component is essential for stability of the entire membrane complex [85]. It is currently unknown exactly how ESX substrates are secreted outside the cell, and whether it is a one or two step process.

While there are multiple ESX systems, not all species of mycobacteria possess all of the systems. Mycobacteria can be divided into fast and slow growing lineages; the evolution of ESX-5 coincides with this differentiation, so that only slow growing species possess ESX-5 [78]. This is intriguing, since all of the major mycobacterial pathogens belong to the slow growing lineage. Another notable feature about the distribution of ESX systems among the various mycobacterial species is that many have lost the virulence-associated ESX-1, including the pathogenic species *M. ulcerans* and *M. avium*, However, both of these species encode ESX-5. Though ESX-1 is a critical virulence factor in many pathogenic mycobacteria, its loss may be compensated for by ESX-5 [84].

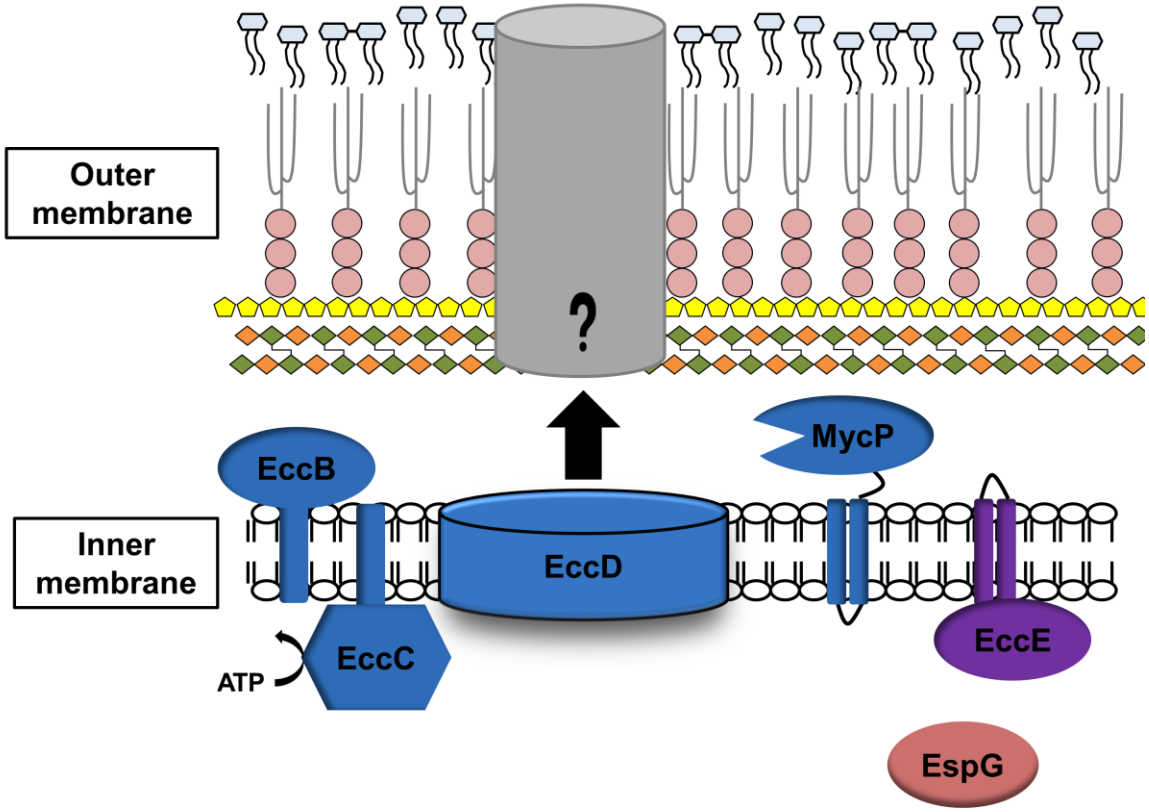

Figure 1.1 Organization of the ESX secretion system machinery. Proteins shown in blue comprise the core secretion machinery components, which are encoded by all ESX systems. All ESX systems, except the archaic ESX-4, also encode EccE and EspG components. Currently, it is unknown if ESX secretion occurs as a 1 or 2 step process, as it is thought EccD, the putative export channel, only spans the inner membrane. The outer membrane consists primarily of mycolic acids interspersed with other lipids and proteins, and is separated from the inner membrane by a layer of peptidoglycan and arabinogalactan.

Known functions of ESX secretion systems in *Mtb*

The ESX-4 system is involved in conjugal DNA transfer in the non-pathogenic species *Mycobacterium smegmatis* [86]. However, secretion through ESX-4 has yet to be detected in *Mtb*, and it is currently unknown if this system is functional [87]. Similarly, the function of the ESX-2 system is unknown, and transposon screens suggest it is not essential for *Mtb* replication, at least *in vitro* [88].

ESX-3 is ubiquitous amongst mycobacterial species, including *Mtb* [88]. ESX-3 functions in the homeostasis of divalent cations through the transport of iron and zinc, and is essential for *in vitro* growth of *Mtb* [88-91]. The ESX-3 system is involved in iron uptake specifically through an iron acquisition pathway using mycobactin, an iron binding siderophore [92]. Strikingly, ESX-3 is also involved in *Mtb* virulence, a function mediated independently of its role in metal homeostasis [93].

ESX-1 was the first ESX system to be identified, and has been the most intensely studied due to its clear role in *Mtb* pathogenesis. This system secretes several factors involved in virulence and modulating the host response. ESX-1 promotes intracellular growth of *Mtb* during the early stages of infection, although there is some debate over exactly how this occurs. One function involves permeabilization of macrophage phagosomes, allowing mycobacterial DNA to enter the host cell cytosol and interact with host receptors to modulate responses to promote bacterial survival [94]. ESX-1 secreted substrates are also involved in the arrest of phagosomal maturation and in escape from the phagosome into host cytosol [36, 95-99]. ESX-1 is also involved in the inhibition of T-cell responses during infection [100]. The ESX-1 locus contains the RD1 region; RD1 is missing in the vaccine strain *M. bovis* BCG, and is thought to be the major factor causing attenuation of this strain [101].

ESX-5 is the most recently evolved of the ESX systems, and is involved in manipulating host responses. ESX-5 secreted factors promote necrotic cell death over apoptosis in macrophages during *Mtb* infection, which favors bacterial survival [102]. The virulence factor LipY, a lipase, is secreted by ESX-5 in *Mtb* [103]. LipY is required for full virulence of *Mtb*, and overexpression of this factor skews host defenses towards the less effective Th2 response, and somehow downregulates the protective Th1 response [104]. Disruption of $EccD_5$, the putative membrane secretion channel, results in severe attenuation in both macrophages and SCID (severe combined immunodeficient) mice, implying that ESX-5 secreted substrates are required for survival in these

environments [105]. ESX-5 secreted substrates are highly immunogenic, and are actively secreted during human infection, which highlights the biological relevance of investigating this system [106, 107]. Although ESX-5 is clearly involved in *Mtb* pathogenesis, both the regulatory mechanism and activating signals were unknown, and will be the main focus of this thesis.

Regulation of ESX-1 and ESX-3 in *Mtb*

Regulatory mechanisms have been characterized for both the ESX-1 and ESX-3 secretion systems. ESX-1 has a well-established role in *Mtb* virulence, as previously described. Most regulators that control ESX-1 secretion act at the *espACD* operon, which encodes the ESX-1 substrates EspA, EspC and EspD that exhibit co-dependent secretion with the canonical ESX-1 substrates EsxA (ESAT-6) and EsxB (CFP-10) [108-110]. EspR, a nucleoid-associated protein, plays a critical role in the regulation of ESX-1 by activating transcription of *espACD* [111-113]. Transcription of *espR* is activated during phagosomal acidification [114] and production of EspR protein peaks during stationary phase growth in culture [113]. ESX-1 regulation also involves a pair of two-component systems; MprAB represses the *espACD* operon [115, 116] while PhoPR positively regulates transcription of many genes in the *esx-1* locus [115]. Several factors are known to activate MprAB, including cell envelope stress and nutrient starvation [117-119] while PhoPR is activated by acidic pH [120]. Finally, EspI negatively regulates secretion of the ESX-1 substrates EsxA, EsxB and EspB in response to low cellular ATP concentration [121].

In contrast, the ESX-3 regulatory mechanism is much more straightforward. ESX-3 is negatively regulated by the repressors IdeR and Zur in response to high levels of iron and/or zinc, respectively [90, 91]. Notably, the regulatory mechanism governing ESX-3 was discovered first, which provided clues for the elucidation of its function.

PE and PPE proteins

With the exception of the evolutionarily archaic ESX-4 system, all ESX gene loci encode two unique families of proteins, termed PE and PPE, which are found exclusively in mycobacteria [122]. These proteins are so named for conserved N-terminal proline-

glutamate (PE) and proline-proline-glutamate (PPE) domains, which are typically about 110 and 180 amino acids in length, respectively, and are secreted as heterodimers consisting of one PE and one PPE protein [123-125]. There are 99 *pe* and 69 *ppe* genes in the sequenced *Mtb* H37Rv genome; other lab strains have varying numbers of these genes [123]. The expansion of *pe* and *ppe* genes seems to have occurred simultaneously with gene duplication events surrounding the ESX systems.

One confounding aspect of Type VII secretion system research is the lack of a universal signal sequence associated with the export of ESX substrates. However, the intricacies of ESX secretion are slowly being revealed, and a signal sequence present in some ESX secreted substrates has been discovered. A conserved YxxxD/E motif is present in some ESX proteins, including PE proteins and the ESX-1 substrates EspB and EsxB, and is required for ESX secretion [126]. However, this secretion signal does not provide specificity for secretion through a particular ESX system [126]. PE and PPE proteins are directed to their cognate ESX secretion systems by EspG protein chaperones that stabilize these proteins in the cytoplasm [124, 125]. EspG binds to a specific hydrophobic motif found on PPE proteins, and this interaction keeps the protein in a soluble state. All ESX systems encode a unique EspG paralog. Each chaperone binds only the cognate PE/PPE pairs, and in turn directs them to a specific ESX system for secretion [127]. The EspG chaperone binds the PPE protein such that the ESX secretion signal located on the PE protein is unobstructed, making interaction with the secretion machinery possible [127].

Although the function of PE and PPE proteins remains largely unknown, several proteins have been shown to play a role in virulence [128, 129]. Many PE and PPE proteins are localized to the cell wall, making these factors perfectly poised to interact with the host, and/or maintain cell membrane integrity during harsh environmental changes encountered during infection [130-132]. Multiple PE and PPE proteins have been found to interact with the host receptor TLR-2, suggesting at least a subset of these proteins modulate immune responses, presumably to promote bacterial survival and/or persistence [133-137]. Further, other PE and PPE proteins, including PE_PGRS62, PE5, PPE15 and PPE2 have been shown to inhibit the production of nitric oxide in the host, a key antimicrobial defense mechanism [134, 135]. PPE41 is actively secreted by the ESX-5 system into the macrophage phagosome during infection by *M. marinum*; though the exact function of PPE41 is unknown, it seems likely to be involved

in manipulation of macrophage function [138]. Nearly 10% of the coding capacity of the *Mtb* genome is dedicated to the *pe* and *ppe* families of genes [139]. Since ESX-5 secretes the vast majority of PE and PPE proteins in the closely related *M. marinum*, the *Mtb* ESX-5 system is also predicted to play a major role secreting these proteins [105, 140]. The *esx-5* encoded *pe* and *ppe* genes are required for *Mtb* survival in macrophages, maintenance of cell wall integrity, and virulence since deletion of these genes impairs these functions [105]. PE and PPE proteins play a major role in *Mtb* pathogenesis by modulating host responses to favor bacterial survival, and the ESX-5 system is required to transport at least a subset of these proteins.

Two component regulatory systems in *Mtb*

Bacteria live and thrive in every environment imaginable. To be successful, bacteria must be able to assess the surrounding environment, interpret the stimuli and make transcriptional changes to appropriately adapt to changing conditions. Two component regulatory systems are one way that bacteria sense and respond to environmental cues to ensure their survival. All bacteria encode two component systems (TCS), and the number of TCS typically increases as the genome size and lifestyle complexity of the bacterium increases [141]. For example, the soil bacterium *Myxococcus xanthus* has over 100 TCS, which reflects the diverse array of social behaviors and complicated life-cycle exhibited by this fascinating microbe [142]. At the opposite end of the spectrum, *Mtb* encodes only 12 TCS, a relatively small number, and is indicative of the evolution of this bacterium as a human pathogen with a primarily intracellular lifestyle [143]. Even with its limited repertoire of TCS, *Mtb* senses and responds to a diverse array of stimuli within the human host, including hypoxia, acid stress, nutrient limitation, membrane stress, ROS and RNS [144].

TCS are typically comprised of a membrane bound sensor histidine kinase (HK) and a DNA binding response regulator (RR). The HK usually senses the input signal from the environment and activates the RR through a phospho-transfer event [143]. The RR is generally a cytoplasmic protein that possesses both a phospho-receiver domain and a DNA binding domain, and once phosphorylated mediates transcriptional changes [143]. The HK and RR are commonly encoded in operons and therefore transcriptionally linked and co-expressed. A given HK usually only interacts with its cognate RR, which

allows for specificity in signaling, though of course there are some exceptions, and cross-talk between *Mtb* TCS has been demonstrated [145, 146]. This type of regulation allows for a precise transcriptional response tailored to the specific environmental signal encountered by the bacterium.

SenX3-RegX3 TCS

SenX3-RegX3 is a typical TCS in *Mtb*, where SenX3 is the membrane bound HK and RegX3 is the cytoplasmic RR [147]. One study suggests that SenX3 directly senses oxygen, but this finding is considered controversial [148]. SenX3 appears to be indirectly activated by phosphate limitation through interaction with a phosphate uptake system [149]. When phosphate limitation is detected, SenX3 autophosphorylates at His167, then phosphorylates Asp52 on RegX3 through phospho-transfer [150]. Although some *Mtb* TCS display cross-talk, the SenX3-RegX3 interaction is specific [145]. The RegX3 regulon includes genes important for inorganic phosphate scavenging, including the *senX3-regX3* operon itself and genes encoding the phosphate uptake system [149].

Many studies have demonstrated that a functional RegX3 is required for *Mtb* virulence, as deletion of *regX3* results in attenuation in both macrophage and murine infection models [147, 149, 151, 152]. Some genes in the RegX3 regulon are involved in phosphate scavenging, so part of the attenuation seen in these various models may be due to an inability of *Mtb* to scavenge sufficient phosphate. However, the work presented in this thesis suggests that there are other reasons *regX3* is required for *Mtb* virulence. RegX3 regulates the virulence-associated ESX-5 secretion system in *Mtb*, and work in this thesis shows that dysregulation of ESX-5 secretion mediated by RegX3 causes attenuation *in vivo*.

Pst and phosphate uptake in *Mtb*

Inorganic phosphate (P_i) is an essential nutrient required by all bacteria, as it is a component of many lipids, sugars, DNA and RNA, and is also an integral part of many cellular processes. Bacteria must acquire P_i from the environment to survive, and typically employ two different uptake systems to accomplish this. The phosphate inorganic transport (Pit) system is a low-affinity, high-velocity uptake system powered by

proton motive force and consisting of a single membrane protein which, in *E. coli*, is constitutively expressed [153, 154]. Recently, the Pit system has been shown to transport the divalent metal cation Zn^{2+}, in complex with P_i, suggesting uptake by this system is not specific for P_i [155]. Given its low affinity for P_i, the Pit system is thought to be important only when phosphate is abundant [154]. In contrast, the phosphate specific transport (Pst) system is a high-affinity, low-velocity ATP binding cassette (ABC) transporter, which is comprised of several subunits (**Figure 1.2)** [156]. PstS binds P_i in the periplasm with high affinity, PstA and PstC are transmembrane proteins that form the transport channel, and PstB is a cytoplasmic ATPase which provides energy for P_i import [157]. The Pst system is thought to be the predominant P_i transporter in *Mtb*, and is active even when P_i is abundant, though it is highly upregulated when P_i is scarce. There are actually two complete high affinity Pst systems encoded in the *Mtb* genome, along with a third PstS substrate binding protein [156]. Importantly RegX3 only controls activation of the *pstS3-pstC2-pstA1* operon [151].

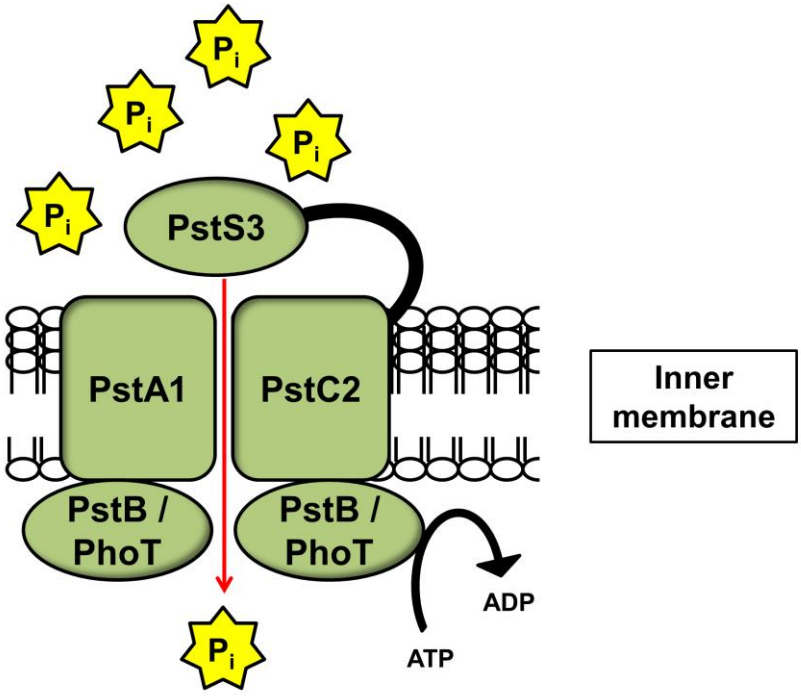

Figure 1.2 The Pst uptake system in *Mtb*. The Pst system is a high-affinity, low-velocity P_i uptake system, which harnesses the power of ATP for transport across the cytoplasmic membrane.

In addition to its role in P_i uptake, the Pst system encoded by the *pstS3-pstC2-pstA1* operon is also involved in P_i-responsive gene regulation through interaction with a TCS. Previous work from our lab has shown that this *Mtb* Pst system behaves similarly to the canonical *E. coli* system (**Figure 1.3**). In *Mtb*, the Pst system senses P_i availability, and when it is abundant inhibits activation of the SenX3-RegX3 TCS, and thus the P_i scavenging response. When P_i becomes limiting, Pst mediated inhibition of the TCS is relieved, leading to activation of RegX3. When RegX3 is activated, it upregulates genes involved in P_i scavenging [149]. Disruption of a transmembrane component of the Pst system, through deletion of *pstA1*, causes constitutive activation of RegX3, regardless of P_i abundance [149]. Notably, the Δ*pstA1*mutant is attenuated in murine infection models, and the attenuation of virulence is due in part to constitutive activation of RegX3 [149]. It is likely that it is ultimately aberrant gene expression specifically mediated by RegX3, which causes attenuation of the Δ*pstA1* mutant. My thesis work has shown that RegX3 controls activation of genes within the virulence-associated *esx-5* locus in response to P_i limitation, and will be discussed in later chapters [158].

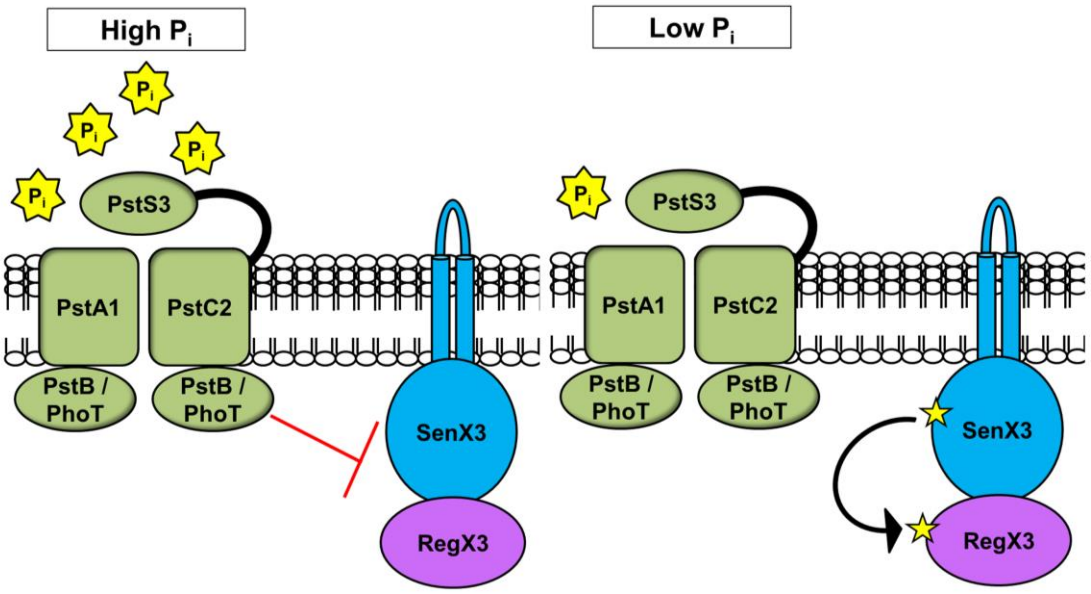

Figure 1.3 Pst/SenX3-RegX3 P_i-responsive system. Together, the Pst phosphate uptake system and the SenX3-RegX3 TCS control P_i-responsive gene regulation. When P_i is abundant, Pst inhibits SenX3/RegX3 activation. When P_i is scarce, Pst mediated inhibition is relieved and the TCS is activated.

Evidence *Mtb* encounters phosphate limitation during infection

It is commonly thought that bacterial pathogens encounter nutrient limitation during the course of infection, and indeed sequestration of nutrients is one key host defense mechanism. Intriguingly, there is a very well-established link between P_i scarcity and bacterial virulence. There is a plethora of evidence that suggests *Mtb* encounters P_i limitation during infection. Several genetic screens have identified genes involved in P_i acquisition as essential for *Mtb* viability and/or virulence in both macrophage and murine infection models.

Components of the Pst system are required for survival and virulence during infection. Deletion of either *pstS1* or *pstS2*, encoding two Pst P_i binding subunits, causes a severe growth defect in both BALB/c and C57BL/6 mice [159]. The entire *pstS3-pstC2-pstA1* operon is also essential for survival in macrophages [160]. Disruption of *pstA1* results in constitutive activation of RegX3 and causes attenuation in several mouse lines [149]. Further, *pstA2* is not required for virulence, but may have a partially redundant function in P_i uptake in *Mtb* [161].

Deletion of the *senX3-regX3* operon in *Mtb* resulted in reduced survival in macrophages and attenuated virulence in both immune compromised SCID mice and immune competent mice [147]. Similarly, a *regX3* transposon insertion mutant had reduced survival in the lungs of mice and guinea pigs [151]. Moreover, data from our own lab demonstrate that a Δ*regX3* mutant has a replication and virulence defect in mice compared to WT *Mtb* [149]. Taken together, these data suggest that P_i stress occurs during infection and highlights the requirement of a functional RegX3 for full *Mtb* virulence.

Evidence supporting P_i as a limiting nutrient *in vivo* lies in the genetic screens that determined genes involved in P_i acquisition are essential for growth in macrophages and in P_i limited conditions, outlined above. One likely site of P_i paucity is within phagosomes of *Mtb* infected macrophages. During the course of an infection, *Mtb* spends a significant amount of time within the macrophage phagosome, and there is evidence that this environment is nutrient poor. Although P_i levels have not been precisely measured in this environment, the amount of elemental phosphorus did decrease in phagosomes of *Mtb* infected macrophages after 24 hours [162, 163]. During the chronic phase of infection *Mtb* resides within foamy macrophages, and the bacterium survives primarily by utilizing host lipids found within this environment [21]. Notably, the

lipid bodies that accumulate in foamy macrophages contain mostly cholesterol, which does not contain P_i.

Thesis summary

Most of my thesis work has focused on elucidating of the mechanism of regulation of the ESX-5 secretion system in *Mtb*. A summary of my findings, and details of the experiments supporting these conclusions, will be found in the following chapters.

Chapter 2 of this thesis outlines the evidence demonstrating that the P_i-responsive Pst/SenX3-RegX3 system regulates ESX-5 activity at the transcriptional level in response to P_i limitation. Work in this chapter shows that the response regulator, RegX3, directly binds to a region within the *esx-5* locus. Based on work described in this chapter, we developed a working model of how the Pst/SenX3-RegX3 and ESX-5 systems interact (**Figure 1.4).** The Pst system inhibits activation of SenX3 when there is abundant P_i outside the cell, and the regulatory system is in effect turned 'off'. However, when P_i becomes scarce, the system turns 'on'. Pst relieves its inhibition of SenX3, which autophosphorylates and then in turn activates RegX3. Once activated, RegX3 directly binds to a region of DNA within the *esx-5* locus, inducing transcription of a subset of genes, ultimately leading to production and secretion of ESX-5 proteins.

Experiments described in chapter 3 build upon the work outlined in chapter 2 by identifying the RegX3 binding site sequence within the *esx-5* locus and identifying an imperfect direct repeat essential for binding within the region. Defining the *esx-5* RegX3 binding site *in vitro* allowed for the construction of binding site mutants to abolish regulation of *esx-5* mediated by RegX3. ESX-5 gene expression and protein secretion were examined in the RegX3 binding site mutants to determine their effect on regulation of this system. These mutants were used to evaluate the importance of Pst/SenX3-RegX3 regulation of ESX-5 *in vivo.* To that end, this chapter also explores how *esx-5* dysregulation impacts *Mtb* virulence in immune-competent C57BL/6 and immune-compromised IrgM1$^{-/-}$ mice through aerosol infection with an *esx-5* RegX3 binding site deletion mutant in the Δ*pstA1* background. Overall, this chapter provides conclusive evidence that ESX-5 is regulated by the Pst/SenX3-RegX3 system, both *in vitro* and *in vivo*, and that precise regulation of ESX-5 secretion is important for *Mtb* virulence.

The work described in chapter 4 examines the effect of the ESX-5 secreted substrate EsxN in *Mtb* pathogenesis. Deletion of *pstA1,* encoding a transmembrane component of the Pst system, causes constitutive activation of RegX3. Notably, data outlined in chapter 3 showed that ESX-5 hyper-secretion is responsible for sensitizing

the Δ*pstA1* mutant to host defenses. To determine whether Δ*pstA1* attenuation could be due to hyper-secretion of one specific antigenic ESX-5 substrate, we deleted *esxN* in the Δ*pstA1* mutant. C57BL/6, IrgM1$^{-/-}$ and Nos2$^{-/-}$ mice were infected with the Δ*pstA1*Δ*esxN* mutant to evaluate if deletion of *esxN* can reverse the Δ*pstA1* virulence defect. While deletion of *esxN* did not improve survival of the Δ*pstA1* mutant in C57BL/6 or IrgM1$^{-/-}$ mice, it did partially restore virulence in Nos2$^{-/-}$ mice. This suggests that EsxN sensitizes Δ*pstA1* bacteria to host defenses in Nos2$^{-/-}$ mice.

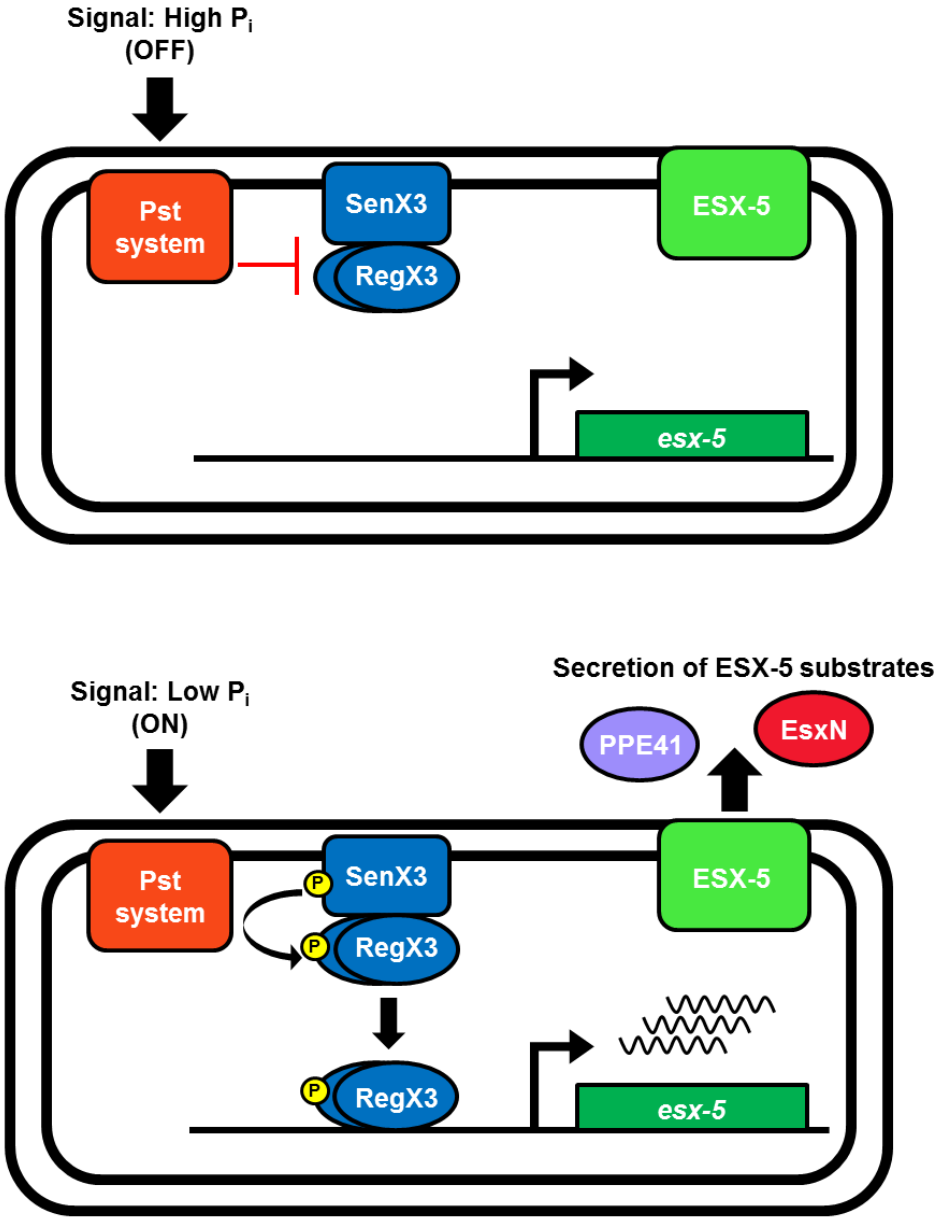

Figure 1.4 Model: Pst/SenX3-RegX3 directly regulates ESX-5 activity. The Pst P_i uptake system regulates P_i-responsive gene regulation through interaction with the TCS SenX3-RegX3. When P_i is abundant, the system is turned off, and the TCS is inhibited. When external P_i becomes limiting, the system is turned on, the TCS is activated, and the response regulator RegX3 mediates transcriptional changes. RegX3 directly regulates ESX-5 transcription, leading to production and secretion of ESX-5 proteins.

Chapter 2

Phosphate starvation: a novel signal that triggers ESX-5 secretion in *Mycobacterium tuberculosis*

This chapter is a reprint, with minor alterations, of a published manuscript

License number 4032091486543

Elliott, SR, Tischler, AD. (2016) Phosphate Starvation: a Novel Signal that Triggers ESX-5 Secretion in *Mycobacterium tuberculosis*. *Molecular Microbiology* 100: (3) 510-526.

Note: The phosphate free qRT-PCR time-course experiment was performed by Dr. Anna Tischler.

Summary

Mycobacterium tuberculosis uses the Type VII ESX secretion systems to transport proteins across its complex cell wall. ESX-5 has been implicated in *M. tuberculosis* virulence, but the regulatory mechanisms controlling ESX-5 secretion were unknown. Here we uncover a link between ESX-5 and the Pst/SenX3-RegX3 system that controls gene expression in response to phosphate availability. The DNA-binding response regulator RegX3 is normally activated by phosphate limitation. Deletion of *pstA1*, which encodes a Pst phosphate uptake system component, causes constitutive activation of RegX3. A Δ*pstA1* mutant exhibited RegX3-dependent over-expression of *esx-5* genes and hyper-secretion of the ESX-5 substrates EsxN and PPE41 when the bacteria were grown in phosphate-rich medium. In wild-type *M. tuberculosis*, phosphate limitation activated *esx-5* transcription and secretion of both EsxN and PPE41, and this response required RegX3. Electrophoretic mobility shift assays revealed that RegX3 binds directly to a promoter within the *esx-5* locus. Remarkably, phosphate limitation also induced secretion of EsxB, an effector of the virulence-associated ESX-1 secretion system, though this induction was RegX3 independent. Our work demonstrates that the Pst/SenX3-RegX3 system directly regulates ESX-5 secretion at the transcriptional level in response to phosphate availability and defines phosphate limitation as an environmental signal that activates ESX-5 secretion.

Introduction

Pathogenic bacteria often use specialized protein secretion systems to promote colonization, replication and survival within the host, and precise regulation of these virulence factors is critical to the success of the organism. All bacteria rely on the general housekeeping Sec secretion system, and many also maintain Tat export pathways to facilitate the transport of proteins vital for cellular processes [164]. *Mycobacterium tuberculosis*, a facultative intracellular pathogen and etiologic agent of the pulmonary infection tuberculosis, also possesses five ESX Type VII secretion systems designated ESX-1 through ESX-5. Each ESX system is comprised of the core secretion machinery components EccB, EccC, EccD, and MycP [165]. EccB is a transmembrane protein, EccC is an ATPase predicted to energize protein translocation,

and MycP is a membrane-tethered protease that may play a role in substrate processing. EccD is also a transmembrane protein, which is predicted to form the inner membrane channel through which ESX substrates are secreted [166]. Esx proteins, the canonical substrates of the ESX systems, are ~100 amino acids in length and are usually encoded in pairs and secreted as heterodimers [166]. ESX systems also secrete two families of proteins unique to mycobacteria that are named PE and PPE proteins based on highly conserved N-terminal domains containing either proline-glutamic acid (PE) or proline-proline-glutamic acid (PPE) sequence motifs [167]. These proteins are also thought to be secreted as heterodimers, consisting of one PE and one PPE protein [168]. The PE and PPE heterodimer requires interaction with its cognate EspG chaperone to be directed to the proper ESX secretion machinery for transport [168-170]. Though the function of most PE and PPE proteins remains unknown, several have recently been implicated in virulence and persistence [171].

ESX-5 is the most recently evolved of the ESX secretion systems [172]. While the exact function of ESX-5 is unknown, this system is found only in the slow-growing pathogenic mycobacteria, including *M. tuberculosis*, and increasing evidence supports a role for ESX-5 in virulence. ESX-5 secretion activates the inflammasome and IL-1β secretion, and induces necrotic cell death, thus promoting bacterial survival [173]. *M. tuberculosis* requires $EccD_5$, the putative ESX-5 transmembrane channel, for full virulence as disruption of $eccD_5$ causes a replication defect within murine macrophages and attenuation in severe combined immune-deficient (SCID) mice [174]. Immune responses to certain PE and PPE epitopes are undetectable when $eccD_5$ is deleted, suggesting that these PE and PPE proteins are transported through ESX-5, and that ESX-5 actively secretes these proteins during infection [175]. The *esx-5* locus encodes several PE and PPE proteins and the Esx protein EsxN. The *esxM* gene adjacent to *esxN* contains a frameshift mutation predicted to truncate EsxM, but there are four homologous *esxM/esxN* gene pairs outside of the *esx-5* locus, and it is probable that one of these alternative EsxM-like proteins pairs with EsxN [166, 174]. The ESX-5 system secretes EsxN and the PE25/PPE41 heterodimer [174], but the mechanism(s) by which ESX-5 substrate secretion is regulated are currently unknown.

Mechanisms regulating activity of the related mycobacterial ESX-3 secretion system have been well characterized. ESX-3 is negatively regulated by the repressors IdeR and Zur in response to high levels of iron and/or zinc, respectively [176, 177].

These observations provided a clue to the function of ESX-3, which is involved in the import of iron, and possibly zinc [178, 179]. Discovering factors that regulate ESX-5 activity may similarly provide insight into its function.

We previously demonstrated that the *M. tuberculosis* Pst/SenX3-RegX3 system controls gene expression in response to extracellular phosphate availability [180]. The Pst (phosphate specific transport) system is an ABC type transporter that imports inorganic phosphate (P_i) across the cytoplasmic membrane using energy from ATP hydrolysis. In some bacterial species, in addition to its role in P_i uptake, the Pst system also plays a role in gene regulation through an interaction with a two-component signal transduction system [181]. The two-component system SenX3-RegX3, comprised of a membrane bound sensor histidine kinase and DNA binding response regulator, respectively, is activated during P_i limitation, and genes involved in P_i scavenging are part of its regulon [180]. Genetic evidence suggests that the Pst system inhibits activation of SenX3-RegX3 when P_i is abundant. Deletion of *pstA1,* which encodes a transmembrane component of the Pst system, causes loss of this negative regulation of SenX3-RegX3, resulting in constitutive activation of RegX3, regardless of extracellular P_i levels [180]. The SenX3-RegX3 two-component system is required for *M. tuberculosis* virulence [180, 182], likely because the bacteria must be able to activate the RegX3 regulon in response to P_i limitation encountered during infection to survive. The regulatory function of PstA1 is also required for *M. tuberculosis* virulence, which suggests that inappropriate constitutive activation of RegX3 and expression of its regulon is also detrimental to survival of the bacterium [180].

Transcriptional profiling experiments identified many genes that were differentially expressed by the Δ*pstA1* mutant compared to wild-type *M. tuberculosis* during growth in P_i-rich conditions. Genes in the *esx-5* locus were over-expressed by the Δ*pstA1* mutant, revealing a potential link between the P_i starvation responsive Pst/SenX3-RegX3 system and ESX-5 activity. Here we demonstrate that ESX-5 protein secretion is activated in response to P_i limitation. Regulation of ESX-5 activity is mediated by the Pst/SenX3-RegX3 system at the level of transcription of a subset of genes that are encoded in the *esx-5* locus. Further, we demonstrate that secretion of the ESX-1 substrate EsxB is also stimulated in response to P_i limitation, though this regulation is independent of the Pst/SenX3-RegX3 system. Our results have provided an

important clue towards understanding the function of the ESX-5 secretion system and relevant environmental signals that may stimulate ESX secretion during infection.

Results

The ΔpstA1 mutant over-expresses esx-5 genes in a RegX3-dependent manner

Previously, we conducted microarray experiments to examine gene expression in a Δ*pstA1* mutant grown in P_i-rich medium [180]. Of the 66 genes that were differentially expressed by Δ*pstA1* bacteria relative to the wild-type (WT) Erdman strain, several were located within the *esx-5* locus **(Figure 2.1A)**. Some *esx-5* genes were significantly over-expressed by the Δ*pstA1* mutant, including *ppe25, pe18, ppe27,* and *esxM* (relative fold change: 9.7, 3.6, 6.8, and 1.6, respectively) [180]. Other genes in the *esx-5* locus were also over-expressed, though the difference in expression level for these genes did not achieve statistical significance: *pe19, espG$_5$, eccD$_5$, esxN* and *mycP$_5$* (relative fold change: 4.09, 2.91, 1.7, 1.5, and 1.56, respectively) [180]. To confirm these results and test whether additional genes in the *esx-5* locus were similarly over-expressed, we conducted quantitative reverse transcription PCR (qRT-PCR) experiments. Consistent with the transcriptional profiling, the *esxM* and *esxN* transcripts were over-expressed approximately 2-fold by the Δ*pstA1* mutant compared to the WT control (Fig. 1B). Additionally, *espG$_5$, eccD$_5$,* and *mycP$_5$* transcripts were significantly over-expressed by the Δ*pstA1* mutant strain relative to WT (fold change = 4, 2.7, and 2; *P* = 0.001, 0.0025, and 0.046, respectively). In contrast, the *eccB$_5$* transcript, which is located in a separate operon on the 5' end of the *esx-5* locus (Fig. 1A), was not significantly over-expressed by the Δ*pstA1* mutant **(Figure 2.1B)**. These data indicate that only the genes located at the 3' end of the *esx-5* locus are aberrantly expressed by the Δ*pstA1* mutant.

We previously demonstrated that aberrant gene expression in the Δ*pstA1* mutant is dependent on the RegX3 DNA-binding response regulator [180]. To determine if RegX3 is responsible for over-expression of *esx-5* genes by the Δ*pstA1* mutant, we examined *esx-5* transcript levels in Δ*regX3,* Δ*pstA1*Δ*regX3* and the respective pND*regX3* complemented strains. Expression of the *esxM, esxN, espG$_5$, eccD$_5$* and *mycP$_5$* transcripts was restored to the WT level in Δ*pstA1*Δ*regX3* bacteria **(Figure 2.1B).** These changes in *esx-5* expression were attributable to deletion of *regX3* because complementation with *regX3* expressed *in trans* (Δ*pstA1*Δ*regX3* pND*regX3*) restored

significant over-expression of the *espG*$_5$, *eccD*$_5$, and *mycP*$_5$ transcripts compared to WT (P = 0.04, 0.01, and 0.02, respectively) **(Figure 2.1B).** These data suggest that *esx-5* genes are over-expressed by Δ*pstA1* bacteria due to constitutive activation of RegX3.

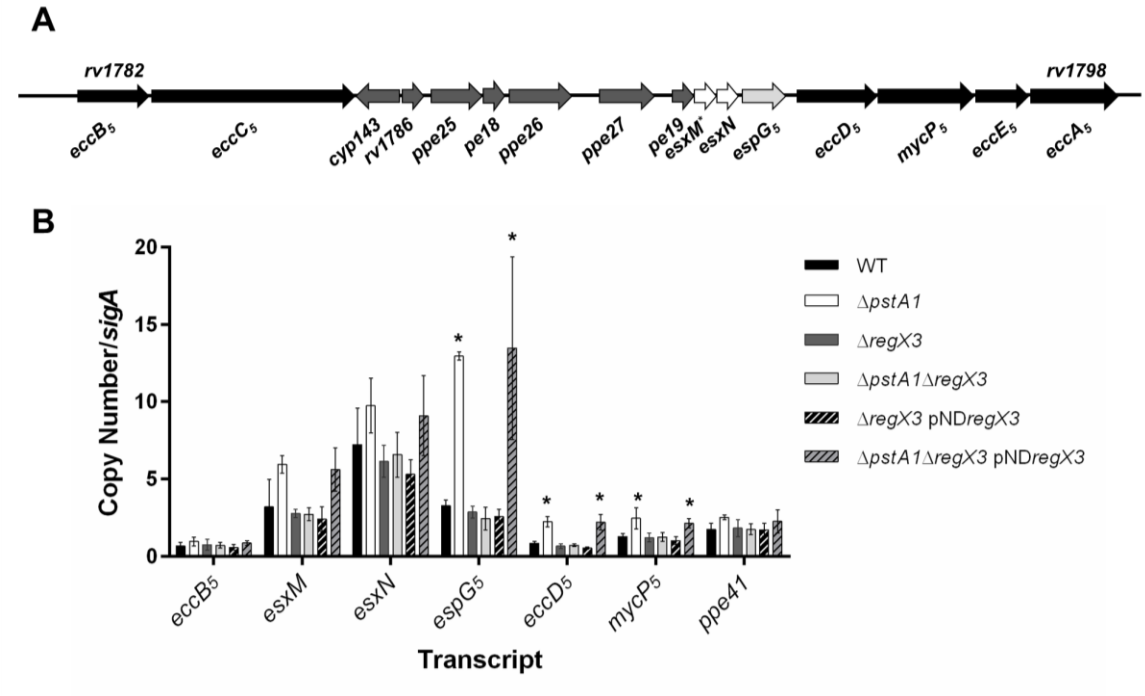

Figure 2.1 Overexpression of *esx-5* genes in the Δ*pstA1* mutant is RegX3-dependent. A. Schematic representation of the *esx-5* locus. Genes encoding ESX-5 conserved components are in black, known ESX-5 substrates are in white, and a known ESX-5 cytoplasmic chaperone is in light gray. Genes with no confirmed function in ESX-5 secretion are in dark gray. The * indicates a gene with a known frame-shift mutation. B. Quantitative RT-PCR analysis of *esx-5* transcription. Wild-type *M. tuberculosis* Erdman (WT), Δ*pstA1*, Δ*regX3*, Δ*pstA1*Δ*regX3,* Δ*regX3* pND*regX3* and Δ*pstA1*Δ*regX3* pND*regX3* were cultured in P_i-rich 7H9 medium to mid-exponential phase and RNA was extracted. Abundance of the *eccB*$_5$, *esxM, esxN, espG*$_5$, *eccD*$_5$, *mycP*$_5$ and *ppe41* transcripts relative to *sigA* was determined by quantitative RT-PCR. Data shown are the means +/- standard deviations of three independent experiments. Asterisks indicate a statistically significant difference in transcript abundance compared to the WT control (P < 0.05).

*Hyper-secretion of ESX-5 substrates by the Δ*pstA1* mutant requires RegX3*

To determine if over-expression of *esx-5* genes by the Δ*pstA1* mutant affects activity of the ESX-5 secretion system, we performed Western blots to monitor the secretion of two known ESX-5 substrates, EsxN and PPE41 (Bottai *et al.*, 2012). EsxN was undetectable in the culture filtrate fractions, but was hyper-secreted by the Δ*pstA1* mutant compared to WT **(Figure 2.2).** Since EsxN secretion was undetectable in the WT control, we quantified the increase in EsxN secretion by performing Western blots on decreasing amounts of the Δ*pstA1* mutant culture filtrate. EsxN secretion became undetectable when 0.5 µg of total culture filtrate protein was loaded, suggesting at least 8-fold induction of EsxN secretion in the Δ*pstA1* mutant compared to WT **(Figure 2.3)**. This induction of EsxN secretion exceeded the roughly 2-fold increase in *esxN* transcription that we detected in the Δ*pstA1* mutant **(Figure 2.1B).** PPE41 was detectable in the cell lysate fraction only upon prolonged exposure of the blot, but was present in higher abundance in the Δ*pstA1* mutant compared to WT **(Figure 2.2).** We also observed approximately 4-fold hyper-secretion of PPE41 by the Δ*pstA1* mutant compared to WT **(Figure 2.1 and 2.2),** though we detected no significant difference in *ppe41* transcript abundance **(Figure 2.1B).** ModD, a protein secreted by the general Sec secretion system, was detected in equivalent amounts in the culture filtrate fractions of WT and Δ*pstA1* bacteria, confirming equivalent protein loading, and demonstrating that the Δ*pstA1* mutation does not globally affect other secretion pathways **(Figure 2.2).** We routinely detect ModD as a doublet in our experiments, which is consistent with previously observed glycosylation of this protein [183]. GroEL2, a cell-associated protein, was undetectable in the culture filtrate, verifying there was no cell lysis **(Figure 2.2).** These data indicate potential post-transcriptional regulation of both EsxN and PPE41 secretion, as in each case the fold change in protein secretion was substantially greater than the change in transcript abundance.

To determine whether the Δ*pstA1* mutation alters expression of the ESX-5 secretion system itself, we examined production of the ESX-5 conserved components $EspG_5$ and $EccB_5$. PPE41 requires an interaction with its cognate chaperone $EspG_5$ for secretion via the ESX-5 system [169]. We observed 4-fold over-production of $EspG_5$ in the cell lysate fraction of Δ*pstA1* bacteria **(Figure 2.2),** which correlated well with the 4-fold over-expression of the *espG*$_5$ transcript **(Figure 2.1B)**. $EccB_5$ is a core component of

the ESX-5 secretion apparatus [184]. We observed 2-fold over-production of $EccB_5$ in the Δ*pstA1* mutant compared to the WT strain **(Figure 2.2 and 2.3),** despite no evidence for transcriptional regulation of the *eccB*$_5$ gene **(Figure 2.1B).** GroEL2, the cell lysate loading control, was detected in equal amounts in WT and Δ*pstA1* bacteria confirming equivalent protein loading **(Figure 2.2).** These data indicate post-transcriptional regulation of $EccB_5$ protein production or stability. Increased production of the ESX-5 secretion machinery and the $EspG_5$ chaperone may contribute to the enhanced secretion of EsxN and PPE41, respectively, that we observe in the Δ*pstA1* mutant.

To determine if the increase in ESX-5 protein production and secretion by the Δ*pstA1* mutant requires RegX3, Western blots were performed on proteins isolated from the Δ*regX3*, Δ*pstA1*Δ*regX3* and the respective pND*regX3* complemented strains. EsxN and PPE41 were not hyper-secreted by the Δ*pstA1*Δ*regX3* mutant **(Figure 2.2).** The cell-associated proteins $EspG_5$ and $EccB_5$ were also produced at WT levels by the Δ*pstA1*Δ*regX3* strain **(Figure 2.2).** Complementation of the *regX3* deletion *in trans* (Δ*pstA1*Δ*regX3* pND*regX3*) restored ESX-5 protein production and secretion to amounts comparable to those observed in the Δ*pstA1* mutant **(Figure 2.2).** The GroEL2 and ModD controls confirmed equivalent protein loading among the WT, Δ*pstA1*Δ*regX3,* and Δ*pstA1*Δ*regX3* pND*regX3* strains and verified that cell lysis did not contribute to proteins present in the culture filtrate **(Figure 2.2).** These results indicate that increased production of ESX-5 secretion system components and hyper-secretion of ESX-5 substrates by the Δ*pstA1* mutant require the DNA-binding response regulator RegX3.

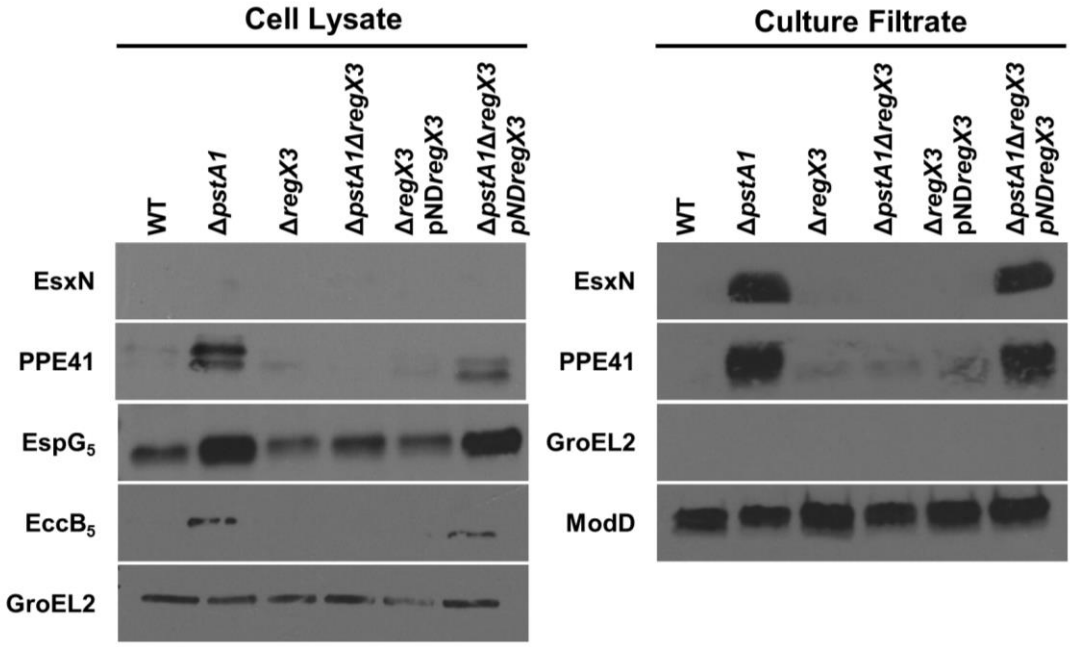

Figure 2.2 Hyper-secretion of ESX-5 substrates by the Δ*pstA1* mutant requires RegX3. Wild-type *M. tuberculosis* Erdman (WT), Δ*pstA1*, Δ*regX3*, Δ*pstA1*Δ*regX3,* Δ*regX3* pND*regX3* and Δ*pstA1*Δ*regX3* pND*regX3* were cultured in Sauton's complete medium without Tween-80 for 5 days. 10 µg of cell lysate (CL) and 4 µg of culture filtrate (CF) proteins were subjected to SDS-PAGE and Western blot analysis. Antibodies used are indicated. Anti-GroEL2 was used as both a loading control for the CL fraction and a cell lysis control in the CF fraction. Anti-ModD was used as a loading control for the CF fraction. Results shown are from a single experiment and are representative of 3 independent experiments.

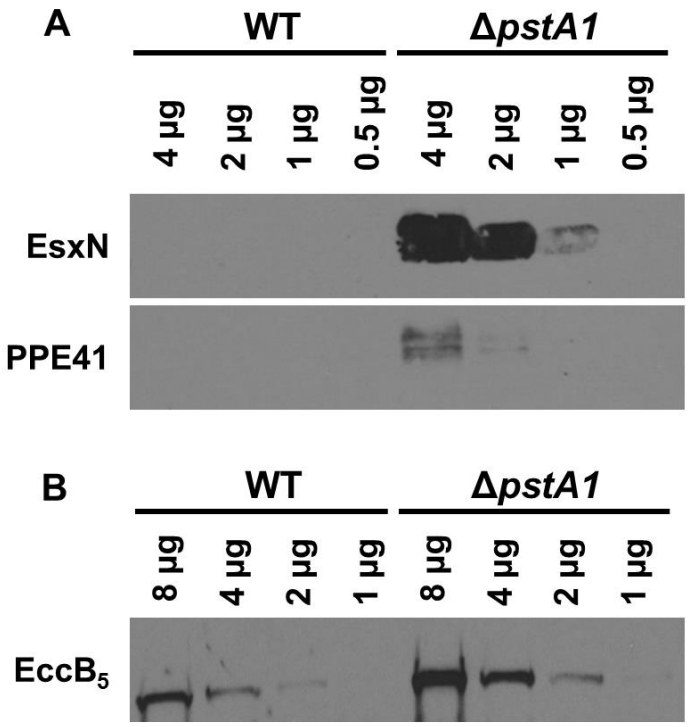

Figure 2.3 Titration of ESX-5 proteins. Wild-type *M. tuberculosis* Erdman (WT) and Δ*pstA1* were cultured in Sauton's complete medium without Tween-80 for 5 days. The indicated amounts of protein were subjected to SDS-PAGE and Western blot analysis to titrate the amount of protein that is hyper-secreted (EsxN, PPE41) or overproduced ($EccB_5$) in the Δ*pstA1* strain compared to WT. Antibodies used are indicated.

esx-5 gene expression is induced by phosphate limitation

Because the SenX3-RegX3 two-component system is activated by P_i limitation [180], we predicted that *esx-5* genes would be induced in a RegX3-dependent manner by P_i-limiting conditions. To test this prediction, we grew WT, Δ*regX3* and Δ*regX3* pND*regX3* strains in P_i-free 7H9 medium, and monitored gene expression over time by qRT-PCR. The *esxN, espG*$_5$, and *eccD*$_5$ transcripts were all significantly induced 2.5- to 4-fold in WT bacteria after 24 hours of P_i limitation (*P*-values: 0.001, 0.03, and 0.008, respectively), and abundance of these transcripts remained significantly elevated for the duration of the experiment **(Figure 2.4 A-C).** There was no induction of either *eccD*$_5$ or *espG*$_5$ transcription in the Δ*regX3* mutant in response to P_i limitation **(Figure 2.4B and C).** The Δ*regX3* mutant did exhibit a 2-fold increase in *esxN* transcription at 24 and 48 hours and this induction was statistically significant (*P* = 0.04 and 0.003, respectively), but was not sustained at the 72 and 96 hour time points **(Figure 2.4A)**. In addition, the level of *esxN* transcript in the ∆*regX3* mutant was significantly lower than the WT control at 24 hours (*P* = 0.025) and all subsequent time points **(Figure 2.4A)**. Complementation of the Δ*regX3* mutation *in trans* restored rapid induction and continuous elevated expression of the *esxN, espG*$_5$, and *eccD*$_5$ genes **(Figure 2.4A-C).** These data demonstrate that activation of *esx-5* gene expression in response to P_i limitation requires *regX3.*

Because *eccB*$_5$ was not significantly over-expressed by the Δ*pstA1* mutant in P_i-rich medium (**Figure 2.1B**, *P* = 0.24), we did not expect this gene to be induced by P_i limitation. We observed a modest 2.2-fold increase in abundance of the *eccB*$_5$ transcript in WT bacteria that was statistically significant (*P* = 0.014), but occurred only after 96 hours of P_i limitation **(Figure 2.4D).** Induction of *eccB*$_5$ transcript was not observed in the ∆*regX3* mutant, but this phenotype could not be complemented *in trans* **(Figure 2.4D).** These data suggest that RegX3-dependent regulation of *esx-5* transcription is restricted to those genes located at the 3' end of the *esx-5* locus. Finally, though it is located outside the *esx-5* locus, we examined transcription of *ppe41*. The level of *ppe41* transcript increased approximately 2-fold in WT bacteria during P_i limitation, but this induction was not statistically significant **(Figure 2.4E).** However, maintenance of this high level of *ppe41* transcription required RegX3 **(Figure 2.4E).** The level of *ppe41* transcript was reduced 4-fold in the ∆*regX3* mutant within 48 hours of P_i limitation, the

decrease was statistically significant (P = 0.0017) and this phenotype could be complemented *in trans* **(Figure 2.4E).** These data indicate that RegX3 contributes to maintenance of *ppe41* transcription during P_i limitation, whether directly or indirectly.

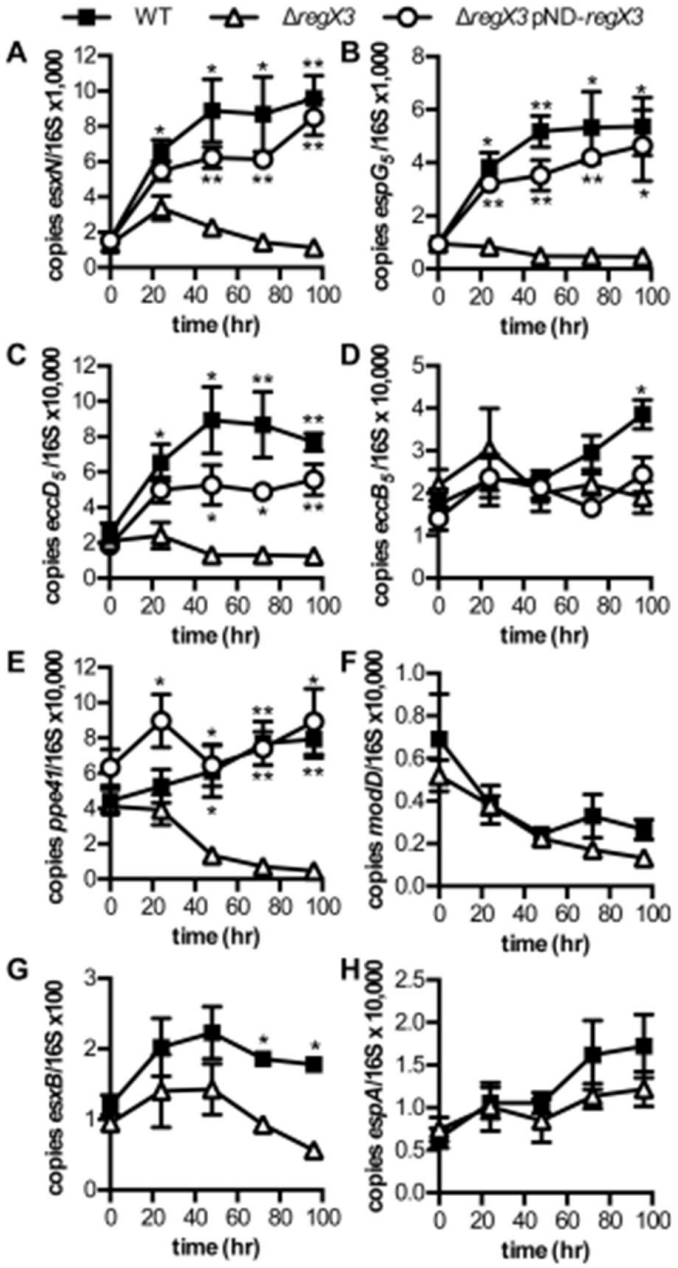

Figure 2.4 Induction of *esx-5* genes by P_i limitation requires RegX3. Wild-type *M. tuberculosis* Erdman (WT), Δ*regX3*, and Δ*regX3* pND*regX3* were cultured in P_i-free 7H9 medium for 96 hours. RNA was extracted at 0, 24, 48, 72 and 96 hours. Abundance of the *esxN, espG$_5$, eccD$_5$, eccB$_5$, ppe41*, *modD*, *esxB*, and *espA* transcripts relative to 16S rRNA was determined by quantitative RT-PCR. Data shown are the means +/- standard deviations of three independent experiments. Asterisks indicate statistically significant differences in transcript abundance between WT and Δ*regX3* or Δ*regX3* pND*regX3* and Δ*regX3*: * $P < 0.05$; ** $P < 0.005$

Secretion of ESX-5 and ESX-1 substrates is induced by phosphate limitation

Since expression of some *esx-5* genes was induced by P_i limitation, we predicted that ESX-5 protein secretion would also be induced by this growth condition. Our initial attempts to monitor ESX-5 protein secretion in P_i-free medium were unsuccessful, likely due to insufficient bacterial growth (data not shown). We therefore conducted experiments to identify a P_i concentration that would be sufficient to sustain growth and yet restrictive enough to activate the P_i starvation-responsive Pst/SenX3-RegX3 system. The analogous two-component system in *E. coli*, PhoRB, is activated by P_i concentrations below 4 µM [181]. To ensure minimal carryover of residual P_i into the cultures used to test protein secretion, we pre-grew the bacteria in Sauton's medium containing only 250 µM P_i, a concentration we determined is sufficient to support a rate of bacterial growth similar to that observed in complete Sauton's medium (data not shown). We then monitored ESX-5 secretion by WT *M. tuberculosis* across a 100-fold range of P_i concentrations in Sauton's medium, including a P_i concentration (2.5 µM) predicted to activate RegX3 and therefore induce ESX-5 secretion.

Consistent with the results of our qRT-PCR experiments, in P_i-limited medium we observed a 9-fold increase in production of $EspG_5$ **(Figure 2.5).** Indeed, production of $EspG_5$ steadily increases as the P_i concentration decreases, with the highest protein production observed at 2.5 µM P_i **(Figure 2.5).** Similarly, EsxN secretion is induced in P_i-limited medium and is the most evident when bacteria are cultured with 2.5 µM P_i **(Figure 2.5).** Although we did not observe significant changes in transcription of *eccB*$_5$ or *ppe41* in WT bacteria grown in P_i-limited conditions, $EccB_5$ protein production and PPE41 secretion were also induced in response to P_i limitation **(Figure 2.5).** We observed an 8-fold increase in production of $EccB_5$ and induction of PPE41 secretion at 2.5 µM P_i **(Figure 2.5).** These data demonstrate that ESX-5 secretion is activated by P_i limitation. They also suggest that $EccB_5$ production or stability and PPE41 secretion are post-transcriptionally regulated during P_i limitation.

In these experiments, we observed a marked decline in secretion of ModD, our Sec-secreted control, in P_i-limited growth conditions **(Figure 2.5).** At 2.5 µM P_i, ModD secretion was decreased approximately 4-fold. The decrease in ModD secretion during P_i limitation is not completely explained by a modest and statistically insignificant 2-fold decrease in *modD* transcript abundance **(Figure 2.4F).** To determine whether this effect

of P_i limitation on ModD secretion was generalizable to other Sec-secreted proteins, we examined secretion of the antigen 85 complex (Ag85). Secretion of Ag85 was also reduced roughly 3-fold at 2.5 µM P_i, with a concomitant accumulation of Ag85 in the cell lysate fraction **(Figure 2.5).** These data suggest that P_i limitation also causes decreased activity of the general Sec secretion system.

To determine if P_i limitation globally affects *M. tuberculosis* protein secretion, we examined secretion of EsxB (CFP-10), an Esx protein secreted via the ESX-1 system that has a well-established role in *M. tuberculosis* virulence [94, 98, 185-188]. We observed less EsxB in the cell lysate fraction and a concurrent increase in EsxB in the culture filtrate fraction at lower P_i concentrations **(Figure 2.5).** Although EsxB secretion was undetectable in 4 µg of secreted protein from P_i-rich control cultures, when 8 µg of secreted protein were loaded we could readily detect EsxB in all samples. We observed a roughly 7-fold increase in EsxB secretion at 2.5 µM P_i **(Figure 2.5)**. The change in distribution of EsxB cannot be explained by differences in cell lysate protein loading or contamination of the culture filtrate fraction by cell lysis, based on the GroEL2 control **(Figure 2.5).** To determine whether activity of the ESX-1 system is generally increased during P_i limitation, we examined secretion of a second ESX-1 substrate, EspB [189]. We detected EspB secretion at all P_i concentrations examined, with an approximately 2-fold increase in EspB secretion at 2.5 µM P_i in the experiment shown **(Figure 2.5),** though this result was not reproducible. These data indicate that secretion or release of the ESX-1 substrate EsxB, rather than activity of the ESX-1 system, is increased during P_i limitation.

To determine if increased EsxB secretion in P_i limited growth conditions was due to a change in *esxB* expression, we monitored *esxB* transcript abundance. Although there was a nearly 2-fold increase in *esxB* expression during P_i limitation in WT bacteria, this induction was not statistically significant **(Figure 2.4G).** We also examined expression of *espA*, since secretion of EsxB is dependent on expression of the *espACD* operon and several transcriptional regulators are known to act at this locus [190, 191]. The *espA* transcript increased approximately 2.5-fold in abundance after 96 hours of P_i limitation, but this change in *espA* expression was not statistically significant **(Figure 2.4H).** These results suggest that increased secretion of the ESX-1 substrate EsxB secretion in response to P_i limitation occurs at the post-transcriptional level.

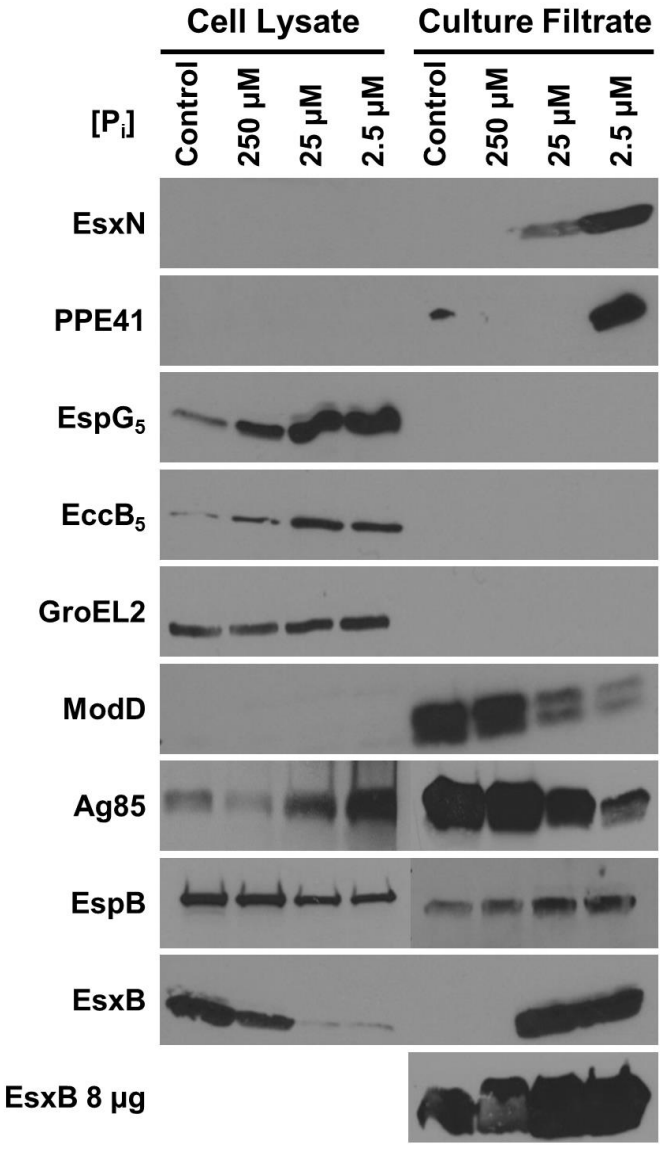

Figure 2.5 P_i limitation induces ESX-5 and ESX-1 protein secretion.
Wild-type *M. tuberculosis* Erdman was grown for 5 days in Sauton's complete medium without Tween-80 (control) or P_i-free Sauton's medium without Tween-80 to which 250, 25, or 2.5 µM KH_2PO_4 was added exogenously. 10 µg of cell lysate (CL) and 4 µg or 8 µg (as indiciated) of culture filtrate (CF) proteins were subjected to SDS-PAGE and Western blot analysis. Antibodies used are indicated. Anti-GroEL2 was used as both a loading control for the CL fraction and a cell lysis control in the CF fraction. Anti-ModD was used as a loading control for the CF fraction. Results shown are from a single experiment and are representative of 2 independent experiments.

Phosphate limitation, not potassium limitation, causes increased secretion of ESX-1 and ESX-5 substrates

In the experiments we conducted to determine if P_i limitation induces ESX-5 protein secretion, we used monopotassium phosphate (KH_2PO_4) as a P_i source. By limiting the source of P_i, we also inadvertently limited the only source of potassium (K^+) in the medium. K^+ is critical for many bacterial processes, including maintaining an electrochemical gradient and regulating intracellular pH [192]. We tested whether the induction of ESX-5 and ESX-1 secretion was due to either K^+ or P_i limitation, or perhaps both. We grew WT *M. tuberculosis* in P_i and K^+ limiting conditions identical to those used in our previous experiments, and in conditions where only P_i or K^+ was the limiting nutrient. Consistent with our previous experiments, both EsxN and EsxB were hyper-secreted in medium that was limiting for both P_i and K^+ **(Figure 2.6).** We observed similar hyper-secretion of both EsxN and EsxB when WT *M. tuberculosis* was grown in medium that was only limiting for P_i **(Figure 2.6)**. In contrast, neither EsxN nor EsxB was hyper-secreted in K^+-limited medium **(Figure 2.6).** GroEL2 was undetectable in the culture filtrate fractions, confirming that the culture filtrates were not contaminated by cell lysis. ModD secretion was unaffected by K^+ limitation, but decreased in both cultures with limiting amounts of P_i **(Figure 2.6).** These data indicate that P_i limitation is a relevant signal that induces secretion of EsxN and EsxB.

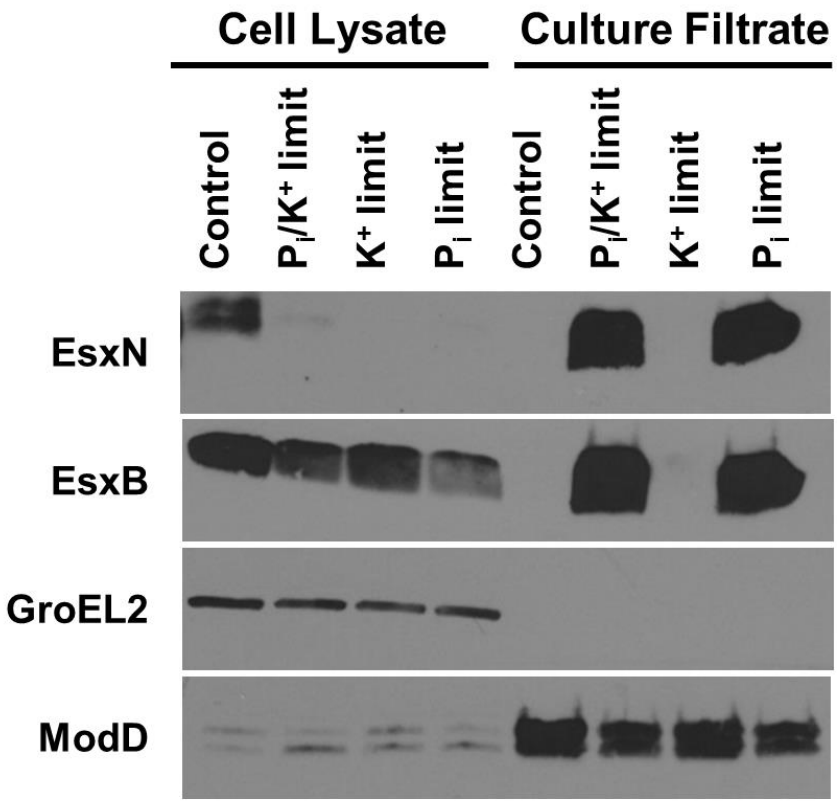

Figure 2.6 Secretion of the ESX-1 substrate EsxB is induced by P_i limitation. Wild-type *M. tuberculosis* Erdman was cultured for 5 days in Sauton's complete medium without Tween-80 (control), or P_i-free Sauton's medium without Tween-80 to which either 2.5 µM KH_2PO_4 (P_i and K^+ limitation), 2.5 µM NaH_2PO_4 and 4 mM KCl (P_i limitation) or 2.5 µM KCl and 4 mM NaH_2PO_4 (K^+ limitation) was added exogenously. 10 µg of cell lysate (CL) and 4 µg of culture filtrate (CF) proteins were subjected to SDS-PAGE and Western blot analysis. Anti-GroEL2 was used as both a loading control for the CL fraction and a cell lysis control in the CF fraction. Anti-ModD was used as a loading control for the CF fraction. Results shown are from a single experiment and are representative of 3 independent experiments.

RegX3 is required for induction of ESX-5 secretion in phosphate limiting conditions

Secretion of the ESX-5 substrates EsxN and PPE41 and the ESX-1 substrate EsxB is induced under P_i-limiting conditions. To determine if this induction is dependent on the response regulator RegX3, we grew WT, Δ*regX3* and Δ*regX3* pND*regX3* strains in medium containing either 4 mM P_i (control) or 2.5 μM P_i. We chose to use the lowest P_i concentration that was tested in our preliminary experiments because we observed the highest levels of EsxN, PPE41 and EsxB protein secretion at this concentration. Hyper-secretion of the ESX-5 substrates EsxN and PPE41 in P_i-limiting conditions was dependent on RegX3. Unlike the WT or Δ*regX3* pND*regX3* complemented strains, the Δ*regX3* strain did not induce EsxN or PPE41 secretion under P_i-limiting conditions **(Figure 2.7).** RegX3 was also required for increased $EspG_5$ and $EccB_5$ protein production under P_i-limiting conditions **(Figure 2.7)**. These data suggest that RegX3-dependent transcriptional activation of a subset of *esx*-5 genes during P_i limitation leads to increased production of ESX-5 secretion system components and activation of ESX-5 substrate secretion.

We also examined secretion of the ESX-1 substrates EsxB and EspB. EsxB secretion was induced in the WT strain under P_i-limiting conditions, but this secretion was independent of RegX3, since we still observed hyper-secretion of EsxB in the Δ*regX3* mutant **(Figure 2.7).** Secretion of a second ESX-1 substrate, EspB, was unchanged by P_i limitation in these experiments **(Figure 2.7)**. These data suggest the existence of an alternative P_i-starvation sensing mechanism that regulates secretion of the ESX-1 substrate EsxB.

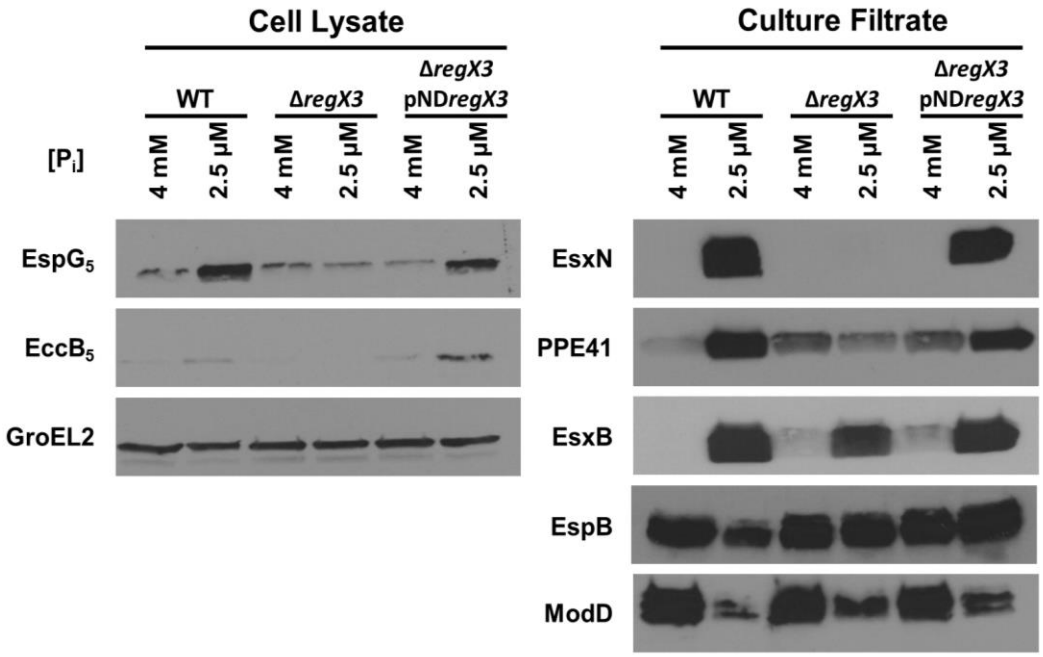

Figure 2.7 RegX3 is required for induction of ESX-5 protein secretion in response to P_i limitation. Wild-type *M. tuberculosis* Erdman (WT), Δ*regX3*, and Δ*regX3* pND*regX3* strains were cultured for 5 days either in Sauton's complete medium without Tween-80 (control) or in P_i-free Sauton's medium without Tween-80 to which 2.5 μM KH_2PO_4 was added exogenously. 10 μg of cell lysate (CL) and 4 μg of culture filtrate (CF) proteins were subjected to SDS-PAGE and Western blot analysis. Anti-GroEL2 was used as both a loading control for the CL fraction and a cell lysis control in the CF fraction (data not shown). Anti-ModD was used as a loading control for the CF fraction. Results shown are representative of 3 independent experiments.

RegX3 directly regulates esx-5 transcription by binding a promoter upstream of pe19

Our results indicate that RegX3 controls ESX-5 secretion at the transcriptional level, though the data do not differentiate between direct or indirect regulation. To determine if RegX3 directly regulates *esx-5* transcription, we sought to identify RegX3 binding sites within the *esx-5* locus. We previously observed that the *pe19* gene, which is located upstream of *esxN*, *espG*$_5$ and *eccD*$_5$ at the 3' end of the *esx-5* locus **(Figure 2.8A)**, is over-expressed by the Δ*pstA1* mutant and induced during P_i limitation in a RegX3-dependent manner [193]. As part of our prior work, we also demonstrated that deletion of the *ppe27-pe19* locus, including the 426 bp intergenic region, caused a severe *in vitro* growth defect that could not be complemented *in trans* [193]. These results suggested that the Δ*ppe27-pe19* mutation is polar on expression of the downstream *esx-5* genes. We therefore hypothesized that the 426 base pair intergenic region between *ppe27* and *pe19* contains a RegX3-dependent promoter that controls transcription of *esx-5* genes **(Figure 2.8A).**

To test this hypothesis, we examined *esx-5* gene expression in strains harboring deletions within the *ppe27-pe19* locus that we previously constructed [193]. Deletion of *ppe27-pe19* including the intergenic region (Δ*ppe27-pe19*) in the Δ*pstA1* background restored expression of *esxN*, *espG*$_5$, and *eccD*$_5$ to WT levels **(Figure 2.8B).** In contrast, *esx-5* transcript levels remained elevated in the Δ*pstA1*Δ*ppe27* and Δ*pstA1*Δ*pe19* strains, which both retain an intact intergenic region **(Figure 2.8B).** These data suggest that the *ppe27-pe19* intergenic region contains regulatory elements that are necessary for the RegX3-dependent over-expression of *esx-5* genes that we observed in the Δ*pstA1* mutant.

To narrow down the approximate location of RegX3-dependent transcription initiation, we performed qRT-PCR using primers spanning the *ppe27-pe19* intergenic region **(Figure 2.8A)**. The *pe19* transcript is over-expressed 4-fold in the Δ*pstA1* mutant compared to both the WT and Δ*pstA1*Δ*regX3* strains (*P* = 0.02 and 0.01, respectively, **Figure 2.8C**). We predicted that amplicons within the intergenic region located 3' of the site of transcription initiation would exhibit a similar expression pattern in these three strains. We found that the Ig4 amplicon was overexpressed 4-fold in the Δ*pstA1* mutant compared to both the WT and Δ*pstA1*Δ*regX3* strains (*P* = 0.03 and 0.02, respectively, **Figure 2.8C**). In contrast, neither the Ig1 nor Ig3 amplicons were significantly over-

expressed by the Δ*pstA1* mutant **(Figure 2.8C**). These data suggest that RegX3-dependent initiation of transcription likely occurs between the Ig3F and Ig4F primers.

To determine if transcripts initiating upstream of Ig4F can extend into the *esx-5* locus, we performed standard RT-PCR. Using primers Ig4F and esxNR3 (**Figure 2.8A**), we detected the predicted 828 bp transcript in the WT, Δ*pstA1* and Δ*pstA1*Δ*regX3* strains grown in P_i-rich medium **(Figure 2.8D**). Although the assay is not quantitative, we consistently observed increased intensity of the PCR product in the Δ*pstA1* mutant relative to the WT control that was RegX3-dependent **(Figure 2.8D**). We also observed increased intensity of the Ig4F-esxNR3 PCR product in samples from the WT strain grown in P_i-limited conditions relative to the P_i-rich control **(Figure 2.8D).** We were unable to amplify a PCR product from cDNA using the Ig3F and esxNR3 primers, providing further evidence that transcription is initiated between Ig3F and Ig4F (data not shown). Importantly, we did not detect PCR products in the no reverse transcriptase controls, indicating the RNA was not contaminated with genomic DNA **(Figure 2.8D**). To establish whether full-length *esxN* and *espG$_5$* are included in the operon initiating 5' of Ig4F, we performed similar RT-PCR experiments using the Ig4F and espG_5R1 primers **(Figure 2.8A).** We detected a 1.8 kb PCR product, confirming that *pe19*, *esxN* and *espG$_5$* are transcribed as an operon **(Figure 2.8D).** Notably, we did not detect this transcript in RNA extracted from the Δ*regX3* mutant grown in P_i-limiting conditions **(Figure 2.8D**). We attempted to determine if this transcript extends into *eccD$_5$* using reverse primers within the *eccD$_5$* gene, but were not able to detect any PCR product (data not shown).

To determine if RegX3 binds directly to the *ppe27-pe19* intergenic region to promote transcription of *esx-5* genes, we purified recombinant His_6-RegX3 and performed electrophoretic mobility shift assays (EMSAs). Two probes were designed within the intergenic region to span the putative RegX3-dependent promoter **(Figure 2.8A).** As a positive control, we generated a probe for a region 5' of *senX3*, to which RegX3 is known to bind [194]. Using a range of His_6-RegX3 concentrations, we found that RegX3 bound to the *senX3* probe and Probe A with similar affinity **(Figure 2.8E).** In contrast, RegX3 was unable to bind Probe B, even at the highest protein concentration tested **(Figure 2.8E).** These data suggest that RegX3 binds directly to a sequence within the *ppe27-pe19* intergenic region located between positions -151 and -27 relative to the PE19 translational start site. To investigate the specificity of this binding, we performed

EMSAs in the presence of competitors. An excess of specific unlabeled competitor resulted in a reversal of the mobility shift for both the *senX3* probe and Probe A (**Figure 2.8F**). Addition of excess non-specific unlabeled competitor did not alter the observed mobility shift for either probe **(Figure 2.8F).** These results suggest that the binding interaction between RegX3 and Probe A is DNA sequence specific. To demonstrate that RegX3, and not a contaminating protein, is responsible for the observed mobility shift, α-His_6 antibodies were added to the binding reactions. We observed a supershift for both probes **(Figure 2.8F)**, indicating the His_6-RegX3 protein specifically binds both the *senX3* promoter and the *ppe27-pe19* intergenic region. These data demonstrate that RegX3 directly controls *esx-5* transcription by binding to a promoter located upstream of *pe19*.

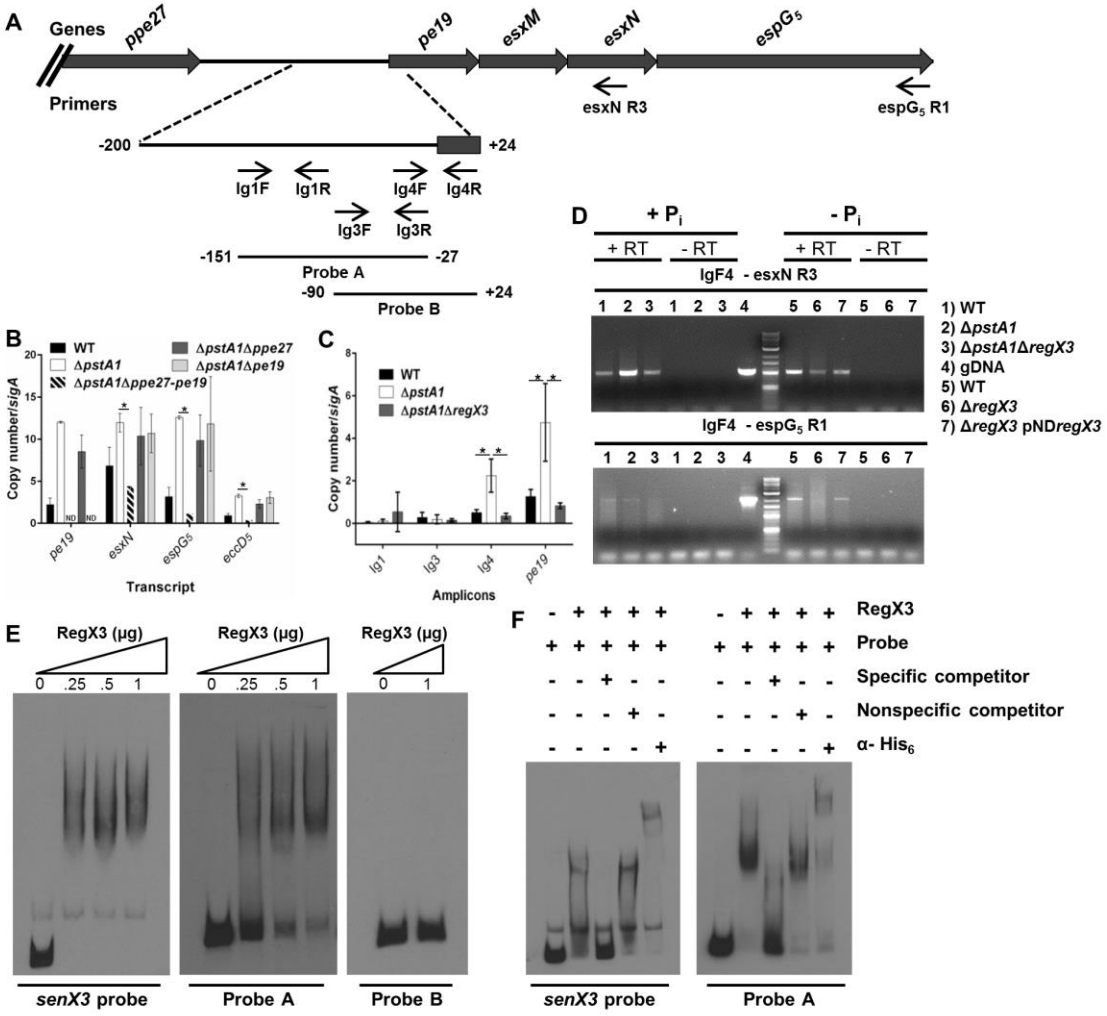

Figure 2.8 Determining the RegX3 binding site within the *esx-5* locus.
A. Schematic representation of the 3' *esx-5* locus. The intergenic region between *ppe27* and *pe19* is enlarged. Locations of relevant primers and probes used for EMSAs are indicated.
B. Quantitative RT-PCR analysis of *esx-5* transcription. Wild-type *M. tuberculosis* Erdman (WT), Δ*pstA1*, Δ*pstA1*Δ*ppe27-pe19,* Δ*pstA1*Δ*ppe27* and Δ*pstA1*Δ*pe19* were cultured in P_i-rich 7H9 medium to mid-exponential phase and RNA was extracted. Abundance of the *pe19, esxN, espG$_5$,* and *eccD$_5$ transcripts* relative to *sigA* was determined by quantitative RT-PCR. Data shown are the means +/- standard deviations of three experiments. Asterisks indicate statistically significant differences in transcript abundance between Δ*pstA1* and Δ*pstA1*Δ*ppe27-pe19* ($P < 0.05$).
C. Quantitative RT-PCR analysis of mRNA levels of amplicons within the *ppe27-pe19* intergenic region and *pe19*. WT, Δ*pstA1* and Δ*pstA1*Δ*regX3* strains were cultured in P_i-rich 7H9 medium to mid-exponential phase and RNA was extracted. Abundance of the amplicons relative to *sigA* was determined by quantitative RT-PCR. Data shown are the means +/- standard deviations of three independent experiments. Asterisks indicate

statistically significant differences in amplicon abundance between Δ*pstA1* and either WT or Δ*pstA1*Δ*regX3* ($P < 0.05$).
D. Identification of an *esx-5* operon by RT-PCR. WT, Δ*pstA1,* and Δ*pstA1*Δ*regX3* were cultured in P_i-rich 7H9 medium to mid-exponential phase (+P_i). WT, Δ*regX3* and Δ*regX3* pND-*regX3* strains were cultured for 24 hours in 7H9 no P_i (-P_i). RNA was extracted, reverse transcribed to cDNA, and PCR amplified using the indicated primers. +RT and -RT denote cDNA synthesis reactions where reverse transcriptase was included or excluded, respectively. gDNA was included as a template for each primer pair as a positive control.
E. EMSA analysis of binding between purified His_6-RegX3 the SenX3 probe (positive control), and Probes A and B probes. 0.5 ng of DIG-labeled probe was incubated with increasing concentrations (0-1 µg) of purified recombinant His_6-RegX3.
F. EMSA analysis of RegX3 binding specificity. DIG-labeled *senX3* probe (positive control) or Probe A was incubated with purified recombinant His_6-RegX3. Unlabeled competitors (both specific or non-specific) or α-His_6 antibodies were added to the binding reactions as indicated by the + symbols.

Discussion

While an intact ESX-5 secretion system is required for full virulence of *M. tuberculosis*, the environmental signal that induces ESX-5 activity was previously unknown. We demonstrate that ESX-5 protein production and secretion are induced during P_i limitation, a nutritional signal that may be encountered by *M. tuberculosis* during infection. Our data support a model in which activation of ESX-5 secretion during P_i limitation is mediated by the P_i starvation responsive Pst/SenX3-RegX3 system that directly regulates transcription of a subset of *esx-5* genes **(Figure 1.4).**

The Pst phosphate uptake system regulates gene expression in response to Pi availability by interacting with the SenX3-RegX3 two-component system [149]. In our model, when external P_i is abundant, the Pst system inhibits activation of SenX3-RegX3, *esx-5* genes are expressed at a basal level, and the ESX-5 system is in effect "off" **(Figure 1.4**). However, when P_i is limiting, inhibition of SenX3-RegX3 is relieved, resulting in activation of *esx-5* transcription and switching ESX-5 secretion "on" **(Figure 1.4**). We provide evidence that RegX3 directly activates transcription of *esx-5* genes in response to low P_i by binding to DNA upstream of *pe19*, leading to increased production of ESX-5 conserved components and increased secretion of the substrates EsxN and PPE41. Notably, deletion of *pstA1* prevents inhibition of SenX3-RegX3 by the Pst system in P_i-rich conditions, resulting in constitutive activation of RegX3. In the Δ*pstA1* mutant, *esx-5* genes are constitutively expressed and ESX-5 is always "on", leading to hyper-secretion of ESX-5 substrates, regardless of external P_i availability.

Our data indicate that RegX3 binds directly to the *ppe27*-pe19 intergenic region to promote *esx-5* transcription. Recent genome-wide transcriptional start site mapping has uncovered two start sites within this intergenic region at -38 and -133 relative to the PE19 translational start [195, 196]. Our qRT-PCR data suggest there is a third transcriptional start site that is active in the Δ*pstA1* mutant and is RegX3-dependent. This putative RegX3-dependent transcriptional start site is located between the 5' end of the Ig3F primer and the 5' end of the Ig4F primer, at -90 and -46 relative to the PE19 translational start, respectively. It is likely that RegX3 binds upstream of this region to promote transcription. We analyzed the Probe A sequence (encompassing a region from -151 to -27), to which RegX3 can bind, for direct and inverted repeats. We found a 5 bp imperfect direct repeat sequence separated by 5 bp (GGTGCcaactGGTGA) located at -

114 to -106 relative to the PE19 translational start site, within the region that we predict is likely to bind RegX3. Using similar analysis, we found a related repeat sequence (GGTGTgctttGGTGC) within the *senX3* probe. These direct repeat sequences may represent sites of RegX3 binding. Our future work will focus on determining the precise RegX3 binding site and mechanism of *esx-5* transcriptional regulation.

Although ESX-5 induction in response to P_i limitation occurs at the transcriptional level by RegX3-dependent activation of a subset of *esx-5* genes, our observations suggest post-transcriptional regulation of some ESX-5 conserved components and substrates. Though we observed over-production of $EccB_5$ during P_i limitation, the *eccB*$_5$ transcript was not induced by this growth condition, suggesting post-transcriptional regulation of $EccB_5$ translation or stability. $EccB_5$ is a member of the ESX-5 membrane complex, along with $EccC_5$, $EccD_5$, and $EccE_5$ [184]. Over-production of $EccB_5$ in P_i-limited conditions may be due to enhanced stability of the protein mediated by increased production of other members of the membrane complex, including $EccD_5$, which is induced at the transcriptional level in these conditions. We also observed a greater increase in EsxN protein secretion compared to the change in *esxN* transcription. It is possible that additional EsxN is secreted due to enhanced production and activity of the ESX-5 core secretion machinery, both in the Δ*pstA1* mutant and during P_i limitation. Finally, we observed increased secretion of the ESX-5 substrate PPE41, though transcription of *ppe41* was not induced during P_i limitation. PPE41 is directed to the ESX-5 system for secretion by $EspG_5$, a protein chaperone that binds to PE and PPE proteins secreted by the ESX-5 system, including PPE41 [168-170]. Increased production of $EspG_5$ during P_i limitation may serve to stabilize PPE41, leading to the increased secretion we detected. Other PE and PPE protein pairs chaperoned by $EspG_5$ may also be hyper-secreted by ESX-5 during P_i limitation. Such post-transcriptional regulation of ESX-5 may allow the bacterium to undergo substantial changes in the ESX-5 secretome mediated by increased transcription of a small subset of genes that are required for stabilization of the system and its substrates.

While P_i limitation induces ESX-5 activity, this condition seems to result in decreased secretion via the general Sec protein secretion system. We detected reduced secretion of both ModD and the antigen 85 complex, which are Sec system substrates, during P_i limitation. The 4-fold decrease in ModD secretion that we observed is only partially explained by the 2-fold decrease in *modD* transcription under P_i-limiting

conditions. In addition, we observed accumulation of the antigen 85 complex in the cell lysate fraction during P_i limitation, suggesting its secretion was prevented. Decreased protein secretion by the Sec system may reflect less available ATP to power Sec secretion. The ESX systems also require ATP for the transport of proteins [166], and perhaps priority for protein secretion during P_i limitation goes to ESX-5. Alternatively, it is possible that there are simply less Sec secreted proteins in the culture filtrate as a fraction of the total amount of secreted proteins due to increased export of ESX-5 substrates and the ESX-1 substrate EsxB.

During our examination of the activating signal driving ESX-5 secretion, we discovered that P_i limitation also triggers secretion of the ESX-1 substrate EsxB (CFP-10). Induction of EsxB secretion is not dependent on RegX3, because enhanced EsxB secretion still occurs in a Δ*regX3* mutant. Since the Δ*regX3* mutant does not exhibit increased ESX-5 activity, hyper-secretion of EsxB is likely not dependent on the ESX-5 secretion system. Increased EsxB secretion in response to P_i limitation appears to occur at the post-transcriptional level, since transcription of *esxB* was not significantly induced during P_i starvation. We did observe a modest 2-fold increase in *espA* transcription during P_i starvation, which may contribute to increased secretion of EsxB, since EspA and EsxB exhibit co-dependent secretion [197]. In contrast, secretion of the ESX-1 substrate EspB was unchanged during P_i limitation, suggesting that P_i limitation does not alter overall activity of ESX-1. The differences in secretion of the ESX-1 substrates EsxB and EspB during P_i limitation could be due to differences in their mechanisms of export, since EspB secretion is independent of both EsxB and EspA [189, 198].

Since the increased EsxB secretion in P_i-limiting conditions is not dependent on RegX3, our data suggest that an additional, as yet unknown, P_i sensing mechanism exists and is responsible for activating EsxB secretion. The ESX-1 transcriptional regulators that have been identified to date, including the nucleoid associated protein EspR and the two-component systems PhoPR and MprAB, have not been implicated in P_i sensing [190, 191, 199, 200]. ESX-1 activity is also controlled by an ATP binding protein EspI, which negatively regulates secretion of the ESX-1 substrates EsxA, EsxB and EspB in response to low cellular ATP concentration [200]. It is possible that downstream physiological changes the bacteria undergo during P_i limitation, perhaps including changes in ATP concentration, are the driving force behind EsxB secretion, as opposed to P_i availability being sensed directly. Hundreds of genes that are not part of

the RegX3 regulon are differentially expressed in response to P_i limitation [201], and it is possible one of these factors is responsible for inducing EsxB secretion under these conditions. Regardless, we discovered that P_i starvation is an environmental signal that triggers secretion of the ESX-1 substrate EsxB, though the precise mechanism by which this occurs will require further investigation.

P_i, an essential nutrient, is a component of many lipids, sugars, and nucleic acids, and is required for many cellular processes, including energy storage and signal transduction. *M. tuberculosis* is a facultative intracellular pathogen which may face conditions with low P_i availability *in vivo.* Within the lung, the bacterium resides inside the phagosomal compartment of infected macrophages, where the environment is harsh and predicted to be nutrient poor [202]. It is possible that P_i is a limiting nutrient within phagosomes, but the P_i concentration in mycobacterial phagosomes has not been measured with any certainty. While the amount of elemental phosphorus within the phagosomes of cultured murine macrophages decreased approximately 1.5-fold after 24 hours of infection with either *M. tuberculosis* or the related virulent species *M. avium* [203, 204], there was no available calibration control to allow determination of the phosphorus concentration. It is therefore unclear if this change in phosphorus availability within the macrophage phagosome would be predicted to starve *M. tuberculosis* for P_i. During the chronic phase of infection, *M. tuberculosis* persists within foamy macrophages inside granulomas, the signature feature of a tuberculosis infection. Foamy macrophages accumulate lipid droplets, which primarily contain cholesterol [205], and *M. tuberculosis* is able to survive by utilizing host lipids found within these lipid bodies [206]. Cholesterol does not contain any phosphorus so foamy macrophages within the granuloma are another niche encountered by *M. tuberculosis* that may have a limiting concentration of P_i. It will be important to determine when and where *M. tuberculosis* encounters P_i limitation during the course of infection to uncover when ESX-5 is active and when its activity is critical for virulence.

We have discovered that P_i starvation is an environmental signal that activates ESX-5 secretion, and the Pst/SenX3-RegX3 system is required for this response. The Pst/SenX3-RegX3 system plays an important role during *M. tuberculosis* infection, considering that both Δ*pstA1* and Δ*regX3* mutants are attenuated in the murine aerosol infection model [180]. Attenuation of these mutants may be explained by inappropriate regulation of ESX-5 secretion. The Δ*pstA1* mutant constitutively hyper-secretes ESX-5

substrates, some of which are known to be highly antigenic [175]. We suspect that the Δ*pstA1* mutant is attenuated *in vivo* due to enhanced susceptibility to host immune responses driven by inappropriate secretion of ESX-5 antigens. Conversely, the Δ*regX3* mutant may be attenuated due to an inability to activate ESX-5 secretion in response to P_i limitation. We liken these dichotomous interpretations to a Goldilocks effect, since either constitutive ESX-5 hyper-secretion or an inability to initiate ESX-5 secretion would be predicted to negatively impact bacterial survival. We propose that *M. tuberculosis* encounters environments with reduced P_i availability during infection, which triggers the bacterium to secrete factors important for survival of these conditions through ESX-5. It is possible that ESX-5 is directly involved in P_i scavenging, similar to the proposed iron and/or zinc scavenging function of ESX-3 [178, 179]. Alternatively, P_i limitation may simply be a signal to *M. tuberculosis* that it is within a nutrient-limited phagosomal environment and must activate ESX-5 and EsxB secretion to manipulate phagosome function. Our future work will further investigate the relationship between P_i starvation and ESX-5 function and the importance of ESX-5 regulation during infection.

Acknowledgements

We thank Dr. Wilbert Bitter for generously providing anti-sera against the $EccB_5$, $EspG_5$, EsxN and PPE41 proteins and Dr. Stewart Cole for providing anti-sera against EspB. The following reagents were obtained through BEI Resources, NIAID, NIH: Polyclonal Anti-*Mycobacterium tuberculosis* CFP10 (Gene Rv3874, EsxB) (antiserum, Rabbit), NR-13801; Monoclonal Anti-*Mycobacterium tuberculosis* GroEL2 (Gene Rv0440), Clone IT-70 (DCA4) (produced *in vitro*), NR-13657; Polyclonal Anti-*Mycobacterium tuberculosis* Mpt32 (Gene Rv1860) (antiserum, Rabbit), NR-13807; and Polyclonal Anti-*Mycobacterium tuberculosis* Antigen 85 complex (FbpA/FbpB/FbpC; Genes Rv3804c, Rv1886c, Rv0129c) (antiserum, Rabbit), NR-13800.

Chapter 3

ESX-5 is directly regulated by the Pst/SenX3-RegX3 system in *Mycobacterium tuberculosis*

Summary

The Type VII secretion system ESX-5 has been implicated in *Mycobacterium tuberculosis* (*Mtb*) virulence, although both the regulatory mechanism and the signal inducing ESX-5 activity were unknown. My previous work discovered that ESX-5 is activated during phosphate (P_i) limitation, an environmental signal *Mtb* likely experiences during infection. ESX-5 activity is directly regulated at the transcriptional level by the Pst/SenX3-RegX3 system, which controls gene expression in response to P_i availability. P_i limitation resulted in RegX3-dependent overexpression of *esx-5* transcripts, as well as overproduction and hyper-secretion of ESX-5 proteins. Experiments described in the previous chapter demonstrated that RegX3 directly binds within the *esx-5* locus, and work outlined in this chapter identified the RegX3 binding site sequence within this region. RegX3 is a global response regulator, and deletion of this factor results in many transcriptional changes, each with potential downstream effects on *Mtb* virulence. Defining the *esx-5* RegX3 binding site sequence allows for subsequent targeted mutation of the sequence to eliminate binding, which ultimately uncouples global effects mediated by RegX3 from the ESX-5 system. Using electrophoretic mobility shift assays (EMSA), a series of unlabeled competitors were screened to determine which sequences were essential for RegX3 binding. The RegX3 binding site sequence was then mutated to evaluate the importance of RegX3-dependent regulation of ESX-5 secretion for *Mtb* virulence. RegX3 binding site mutations reversed both overexpression of *esx-5* genes and hyper-secretion of EsxN. To evaluate the effect of *esx-5* regulation on *Mtb* virulence *in vivo,* C57BL/6 and IrgM1$^{-/-}$ mice were infected via the aerosol route with a binding site null mutant, Δ*pstA1*ΔBS. Interestingly, deletion of the RegX3 binding site restored virulence to the otherwise attenuated Δ*pstA1* mutant strain, suggesting hyper-secretion of ESX-5 substrates sensitizes Δ*pstA1* bacteria to host defenses, and highlighting the importance of precise regulation of *esx-5* in *Mtb* pathogenesis.

Introduction

Our previous findings, outlined in chapter 2, demonstrated that the ESX-5 secretion system is activated in response to P_i limitation, a relevant environmental signal *Mtb* likely encounters during infection. This work also revealed that the ESX-5 system is regulated at the transcriptional level by the phosphate-responsive Pst/SenX3-RegX3 system. Moreover, EMSA experiments showed this regulation was mediated through direct binding of RegX3 within the *esx-5* locus, within the *ppe27-pe19* intergenic region. However, it remained unclear where within the 125 bp sequence RegX3 binds. Identifying the RegX3 binding site sequence, followed by targeted mutation of the region, will allow us to examine the contribution of ESX-5 to *Mtb* pathogenesis independently of global RegX3 effects.

Proper gene expression is critical to the survival of all bacteria. Regulation of transcription is one important mechanism utilized by prokaryotes to maintain precise control of gene expression. Transcription is accomplished in three main steps, initiation, elongation and termination, though regulation of gene expression occurs most commonly at the initiation step [207]. Initiation of transcription requires interaction between RNA polymerase and the DNA sequence of a promoter. The basic architecture of prokaryotic promoters consists of a -35 region, a -10 region, the transcription start point, followed by the ribosomal binding site and finally the start codon [208]. Mycobacterial promoters share some structural similarities to the canonical *E. coli* promoter. Both organisms typically encode a guanine residue at the transcription start site, maintain a conserved -10 region comprised of six bases, and have GC-rich ribosomal binding sites [208]. The -10 region is made of six bases which are variable in mycobacteria [209]. Notably, there is much more variability found within mycobacterial promoter sequences compared to *E. coli*. There is no conserved -35 region in mycobacteria, which is well defined in other bacteria [208]. Importantly, all gene expression is controlled at the level of the individual promoter.

Successful transcriptional initiation requires binding of a sigma factor to a specific promoter sequence. Sigma factors are a critical component of the transcriptional machinery, and position RNA polymerase by recognizing consensus -35 and -10 elements within promoters [210]. Most bacteria utilize one main σ^{70}-type sigma factor which recognizes promoters of constitutively active genes, and employ alternative sigma

factors to control genes activated only under certain conditions [208]. There are 13 sigma factors in *Mtb,* all of which are σ^{70}-type [211]. SigA is the primary sigma factor in *Mtb*, and controls genes involved in general viability, which are constitutively expressed [208]. The other sigma factors are expressed in response to specific stresses sensed by the bacterium [208].

RNA polymerase relies on transcription factors to enhance binding to specific promoter sites, and ultimately confer precision of transcriptional responses to specific signals [210]. Transcription factors are DNA-binding proteins that integrate input signals from the environment to alter transcription, so these proteins play a critical role in gene regulation. Transcription factors can act as either activators or repressors of transcription through binding to promoter-specific DNA sequences. Binding to a promoter and recruiting RNA polymerase results in activation of transcription [210]. Though initiation of transcription can be repressed in a variety of ways, the simplest form of repression occurs when a transcription factor binds to promoter regions and blocks RNA polymerase [210].

RegX3 is a global transcription factor that activates *esx-5* transcription when P_i is limiting. We propose that RegX3 mediated regulation of ESX-5 is important for *Mtb* virulence. Notably, the Δ*pstA1* mutant is attenuated for both replication and virulence in C57BL/6, Nos2$^{-/-}$ and IrgM1$^{-/-}$ mice [149]. However, the underlying cause of attenuation seems to differ between the strains of mice, likely a direct result of the differing immune responses among them. IrgM1$^{/-}$ and Nos2$^{-/-}$ mice are immune-compromised, and are particularly susceptible to intracellular pathogens, since they lack either a critical GTPase which is recruited to the phagosome and is required for its maturation and acidification, or the enzyme producing reactive nitrogen species, respectively [28, 212, 213]. The Δ*pstA1* mutant attenuation in Nos2$^{-/-}$ mice is due to constitutively active RegX3, since deleting *regX3* in the Δ*pstA1* mutant restores virulence to almost WT levels [149]. Further attenuation of the Δ*pstA1* mutant in Nos2$^{-/-}$ mice is partly due to constitutive overexpression of *pe19,* encoded within the *esx-5* locus [193]. But attenuation of the Δ*pstA1* mutant in C57BL/6 and IrgM1$^{-/-}$ mice is not due to overexpression of *pe19* [193]. Deletion of *regX3* in the Δ*pstA1* mutant did not reverse attenuation in C57BL/6 and IrgM1$^{-/-}$ mice, and a *regX3* null mutant is just as attenuated

in these mice as the Δ*pstA1* mutant [149]. Therefore, we could not directly test if attenuation of the Δ*pstA1* mutant is due to RegX3 independent factors by deleting *regX3*.

Our work has shown that *esx-5* transcripts are overexpressed and ESX-5 substrates are hyper-secreted in the Δ*pstA1* mutant in a RegX3-dependent fashion [158]. ESX-5 substrates are highly antigenic, and we suspect that infection with the Δ*pstA1* mutant may be hyper-inflammatory, since ESX-5 secreted substrates induce IL-1β secretion, inflammasome activation and promote necrotic cell death [102]. We noted that Δ*pstA1* bacteria replicate with WT kinetics for the first two weeks of infection in C57BL/6 mice, which coincides with the onset of adaptive immunity, after which replication of the mutant is controlled [149]. We propose that inappropriate constitutive hyper-secretion of ESX-5 substrates may sensitize Δ*pstA1* bacteria to host responses in C57BL/6 and IrgM1$^{-/-}$ mice [107]. However, a *regX3* null strain cannot be used to test this possibility, since Δ*regX3* bacteria are also attenuated in C57BL/6 and IrgM1$^{-/-}$ mice [149]. Hence, a more targeted approach, disrupting only RegX3-mediated activation of *esx-5* transcription, is required to address our hypothesis. Despite several studies aiming to map the binding sites of all 214 transcription factors in *Mtb* using ChIP-seq analysis, a consensus binding site for RegX3 has yet to be uncovered [214-216]. Therefore, to specifically disrupt RegX3-mediated regulation of *esx-5* we sought to decipher the RegX3 binding site. Moreover, an *esx-5* RegX3 binding site mutant in WT *Mtb* would allow us to test whether regulating ESX-5 activity in an important function of RegX3 *in vivo.*

In this work, we identified the RegX3 binding site sequence upstream of *pe19*, within the *esx-5* locus, which includes an imperfect direct repeat that is essential for binding. Despite the variability of mycobacterial promoter regions, known consensus sequences allow for prediction of a putative RegX3 promoter within the *esx-5* locus. Importantly, mutation of the RegX3 binding site suppressed *esx-5* gene expression during P_i limitation in WT *Mtb*, and reversed hyper-secretion of ESX-5 substrates in WT and Δ*pstA1* bacteria. Together, these data conclusively demonstrate that Pst/SenX3-RegX3 mediated regulation of ESX-5 activity occurs in *Mtb*. Further, deletion of the *esx-5* RegX3 binding site restored virulence in the attenuated Δ*pstA1* mutant in IrgM1$^{-/-}$ mice, demonstrating that precise regulation of ESX-5 is critical to the success of *Mtb* during infection.

Results

RegX3 binds within the upstream portion of Probe A′, but not the downstream portion

Previous work demonstrated that ESX-5 is directly regulated by the Pst/SenX3-RegX3 system at the transcriptional level and revealed that RegX3 bound within the *ppe27-pe19* intergenic region in the *esx-5* locus (**Figure 3.1**) [158]. However, the region identified was 125 bp in length, and a typical transcription factor binding site is less than 40 bp [208]. To more identify the *esx-5* RegX3 binding site sequence, we conducted EMSA experiments, using purified recombinant His_6-RegX3 protein and PCR-amplified DNA probes. A list of all primers used to amplify EMSA probes can be found in **Table S5**. **Table 3-1** lists all EMSA probe sequences that were tested and reports the ability of RegX3 to bind the competitor probe.

Table 3-1 EMSA probe amplicon sequences and RegX3 binding data. Bold indicates mutated bases.

Probe Name	Amplicon	RegX3 binding (Y/N)
IgF1/R1/ Probe A	CGCTGTTACTGTGGCGTTACCACAGGTGAATTTGCGGTGCCAACTGGTGAACACTTGCGAA	Y
IgF3/R3 / Down	CGGGTGGCATCGAAATCAACTTGTTGCGTTGCAGTGATCTACTCTCTTGCAGAGAGCCGTTGC	N
Probe 1	AGGTGAATTTGCGGTGCCAACTGGTGAACACTTGCGAACGGGTGGCATCGAAATCAACTTGTTGCGTTGCAGT	Y
Probe 2	TGAATTTGCGGTGCCAACTGGTGAACACTTGCGAACGGGTGGCATCGAAATCAACTTGTTGCGTTGCAGT	N
Probe 3	ACCCGGTCAAACGCTGTTACTGTGGCGTTACCACAGGTGAATTTGCGGTGCCAACTGGTGA	Y
Probe 4	CGCTGTTACTGTGGCGTTACCACAGGTGAATTTGCGGTGCCAACTGG	N
DR1 Probe	CGCTGTTACTGTGGCGTTACCACAGGTGAATTTGC**TTGTA**CAACTGGTGAACACT	N
DR2 Probe	CGCTGTTACTGTGGCGTTACCACAGGTGAATTTGCGGTGCCAACT**TTGTC**ACACT	N
DR3 Probe	ACA**TTGTC**ATTTGCGGTGCCAACTGGTGAACACTTGCGAACGGGTGGCATCGAAATCAACTTGTTGCGTTGCAGT	Y
Spacer - Probe	CGCTGTTACTGTGGCGTTACCACAGGTGAATTTGC**GGTGC**CT**GGTGA**ACACT	N
Spacer + Probe	CGCTGTTACTGTGGCGTTACCACAGGTGAATTTGCGGTGCCAACTTCA**GGTGA**ACACT	N
Ig Probe 2	CGCTGTTACTGTGGCGTTACCACAGGTGAATTTGCGGTGCCAACTGGTGAACACT	Y
Ig Probe 3	GTGGCGTTACCACAGGTGAATTTGCGGTGCCAACTGGTGAACACTTGCGAA	Y
Ig Probe 4	CACAGGTGAATTTGCGGTGCCAACTGGTGAACACTTGCGAACGGGTGGCATCGAAATCAACTTGTTGCGT	Y

Ig Probe 5	TTTGCGGTGCCAACTGGTGAACACTTGCGAACGGGTG GCATCGAAATCAACTTGTTGCGT	N
Ig Probe 6	CAACTGGTGAACACTTGCGAACGGGTGGCATCGAAATC AACTTGTTGCGTTGCAGTGATCTACTCTCTTGCAGAGA GCCGTTGCT	N
Ig Probe 9	ATTTGCGGTGCCAACTGGTGAACACTTGCGAACGGGTG GCATCGAAATCAACTTGTTGCGTTGCAGT	N
Ig Probe 10	CGCTGTTACTGTGGCGTTACCACAGGTGAATTTGCGGT GCCAACTGGTGAACACTTGCGA	Y
Ig Probe 11	CGCTGTTACTGTGGCGTTACCACAGGTGAATTTGCGGT GCCAACTGGTGAACACTTGCG	Y
Ig Probe 12	CGCTGTTACTGTGGCGTTACCACAGGTGAATTTGCGGT GCCAACTGGTGAACACTTGC	Y
Ig Probe 13	CGCTGTTACTGTGGCGTTACCACAGGTGAATTTGCGGT GCCAACTGGTGAACACTTG	Y
Ig Probe 14	CGCTGTTACTGTGGCGTTACCACAGGTGAATTTGCGGT GCCAACTGGTGAACACTT	Y
Ig Probe 15	CGCTGTTACTGTGGCGTTACCACAGGTGAATTTGCGGT GCCAACTGGTGAACACT	Y

Previous work demonstrated the RegX3 binding site was located within bases -151 to -27, relative to the *pe19* start codon [158]. This region corresponds to the sequence of Probe A′. Two smaller unlabeled competitors were designed to divide the Probe A′ sequence in half, named Up (-151 to -90) and Down (-90 to -27) to denote which segment of Probe A′ is being tested (**Figure 3.1A**). Excess unlabeled Up or Down competitor was added to binding reactions with labeled Probe A′ (**Figure 3.1B)**. RegX3 was able to bind to the Up region, since addition of the unlabeled competitor resulted in a competition of the mobility shift to free probe (**Figure 3.1B**). However, RegX3 was unable to bind the Down region, since addition of excess unlabeled competitor did not alter the observed mobility shift. Taken together, these data indicate that RegX3 binds within bases -151 to -90 relative to the *pe19* start codon. This 61 bp segment, containing the RegX3 binding site, will now be referred to as Probe A, and will be used as the DIG-labeled probe in all future experiments.

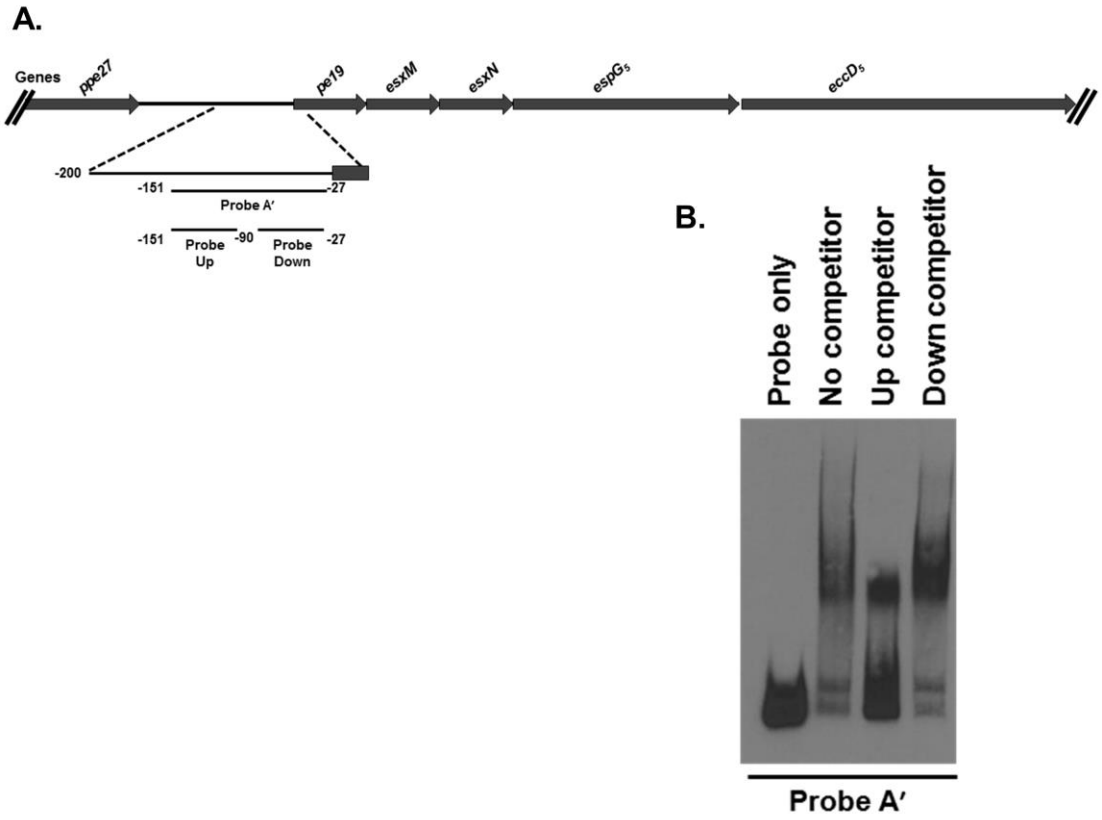

Figure 3.1 RegX3 binds within the 5' half of Probe A′
A) Schematic of the partial ESX-5 gene locus, depicting the genomic location of the probe and competitors.
B) EMSA analysis of binding between purified His_6-RegX3, DIG-labeled Probe A′ and unlabeled Up or Down competitors. 0.5 µg of purified His_6-RegX3 was incubated with 200 ng unlabeled competitors for 15 minutes, then 0.5 ng labeled Probe A′ was added and incubated for 15 minutes.

Identification of essential bases required for RegX3 binding

To further identify the RegX3 binding site sequence, we devised a strategy wherein a series of unlabeled competitors truncating the Probe A sequence at the 5' and 3' ends were screened for the ability of RegX3 to bind. Unlabeled competitors were added in excess in EMSAs. A full list of all sequences tested can be found in **Table 3-1**, though only data from competitors which defined the essential bases required for RegX3 binding are shown (**Figure 3.2A**). An excess of unlabeled Probe 1, which truncates the Probe A reference sequence at the 5' end, reversed the mobility shift, indicating RegX3 can bind this competitor (**Figure 3.2B**). However, RegX3 was unable to bind to unlabeled Probe 2, which truncates an additional three bases at the 5' end, since the mobility shift was unperturbed. This indicates that one or more of the bases missing in Probe 2 are essential for RegX3 binding. Taken together, the 5' end of the RegX3 binding site is located near position -128, relative to the *pe19* start codon. Similarly for the 3' end, excess Probe 3 reversed the mobility shift, indicating RegX3 can bind to this probe (**Figure 3.2B**). Conversely, when excess Probe 4 was added, which eliminates an additional three bases from the 3' end compared to Probe 3, RegX3 binding is abrogated, as mobility of the shift was unaffected. These data demonstrate the three bases missing in Probe 4 compared to Probe 3 are required for RegX3 binding, and thus define the 3' end of the RegX3 binding sequence, which is located at -101, relative to the *pe19* start codon. Altogether, the RegX3 binding site sequence consists of 27 bp, and is located at the -128 to -101 position relative to the *pe19* start codon.

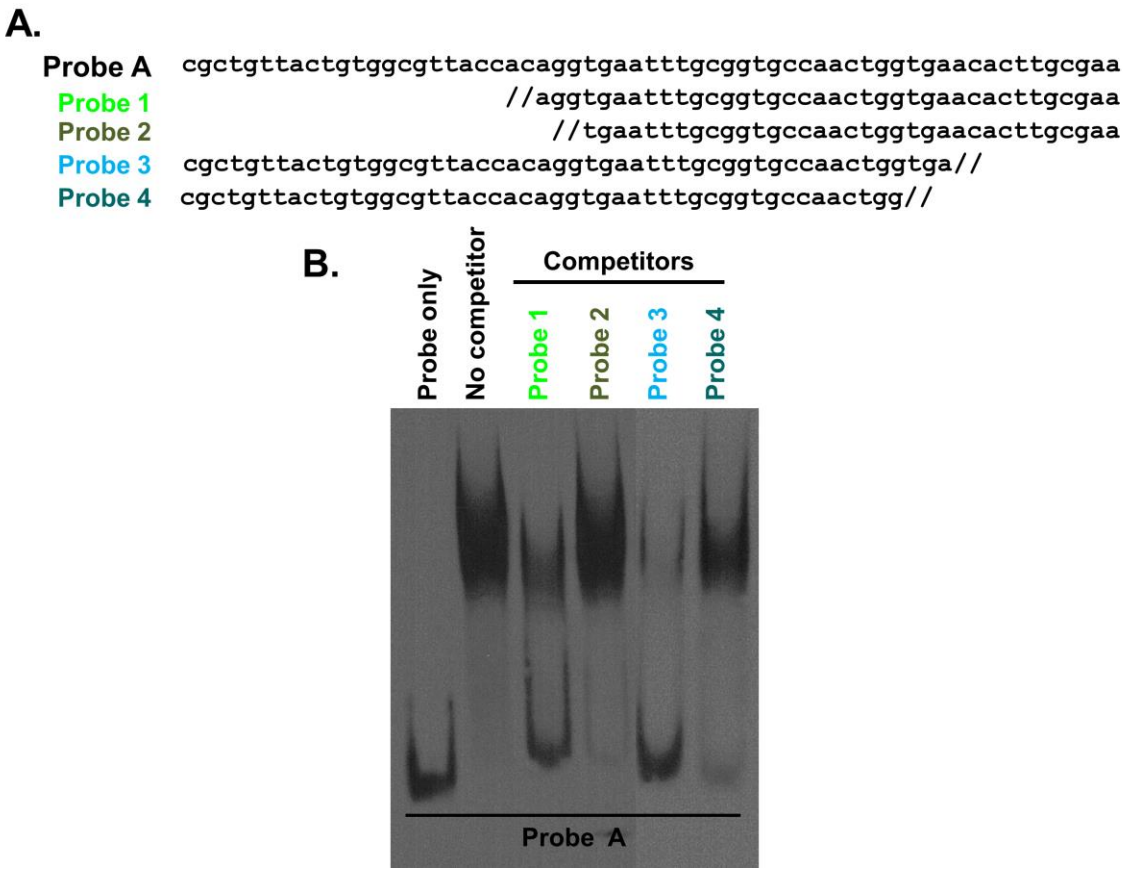

Figure 3.2 Identifying bases essential for RegX3 binding. A) Schematic depicting the sequences of the experimental probes compared to the labeled Probe A sequence. B) EMSA analysis of binding between purified His_6-RegX3, DIG-labeled Probe A and unlabeled competitor Probe 1, 2, 3, or 4. 0.5 µg of purified His_6-RegX3 was incubated with 200 ng unlabeled competitor for 15 minutes, then 0.5 ng labeled Probe A was added and incubated for 15 minutes.

Evaluating essential binding elements within the esx-5 RegX3 binding sequence

Inverted repeats, direct repeats (perfect or imperfect) and palindromes are common sequence features found within promoter regions; these elements are recognized by transcription factors and are required for specific binding [208]. Previously, a putative imperfect direct repeat separated by a 5 bp spacer was identified within the Probe A′ sequence [158]. Upon further examination, a third putative direct repeat (DR3) was found upstream of the first two (DR1 and DR2), though DR3 and DR1 are separated by a 6 bp spacer. Importantly, these elements are all contained within the truncated -128 to -101 region we defined as the RegX3 binding site sequence (**Figure 3.3A**). To determine if any of these elements are required for RegX3 binding, EMSA probes were designed that mutated each putative direct repeat individually. Each of the five bases in the direct repeat sequences were altered using transversion mutations, since exchanging purine bases for pyrimidine bases, and vice versa, is the most dramatic change possible (**Figure 3.3A**). The spacer sequence between DR1 and DR2 was altered by either adding or removing three bases (**Figure 3.3A**). All of the mutated sequences were screened as unlabeled competitors, added in excess, with reversal of the mobility shift as a read-out for the ability of RegX3 to bind the mutated sequence being tested. Addition of the unlabeled DR3 mutated competitor did reverse the mobility shift, indicating RegX3 can still bind this sequence, even when the putative DR3 has been mutated (**Figure 3.3B**). Alternately, when either unlabeled mutated DR1 or DR2 competitors were added, mobility of the gel shift was unaffected, indicating RegX3 cannot bind these mutated sequences, and these bases are required for binding. Taken together, these data indicate that DR1 and DR2 are legitimate sequence elements required for RegX3 binding. Mutation of the putative DR3 sequence is permissive to RegX3 binding, implying the exact sequence in this region is not important for binding. Altering the spacing between DR1 and DR2, either by adding or removing 3 bases, abrogates RegX3 binding, since adding the unlabeled Spacer + and Spacer – competitors did not affect mobility of the shift. This indicates that this 5 bp spacer region is essential for RegX3 binding, at least *in vitro*. These EMSA experiments have demonstrated the presence of an imperfect direct repeat sequence separated by a 5 bp spacer within the RegX3 binding site, all of which are essential for RegX3 binding *in*

vitro.

We have identified a 27 base pair RegX3 binding site sequence, which is within the range of known transcription factor binding regions. Moreover, this binding site is located upstream of other putative canonical promoter elements, including the -35 region, the -10 region, the ribosomal binding site (RBS) and the *pe19* start codon (**Figure 3.4**). The RegX3 binding site sequence identified upstream of *pe19* is consistent with the location of a transcriptional activator, and is in line with our transcriptional data.

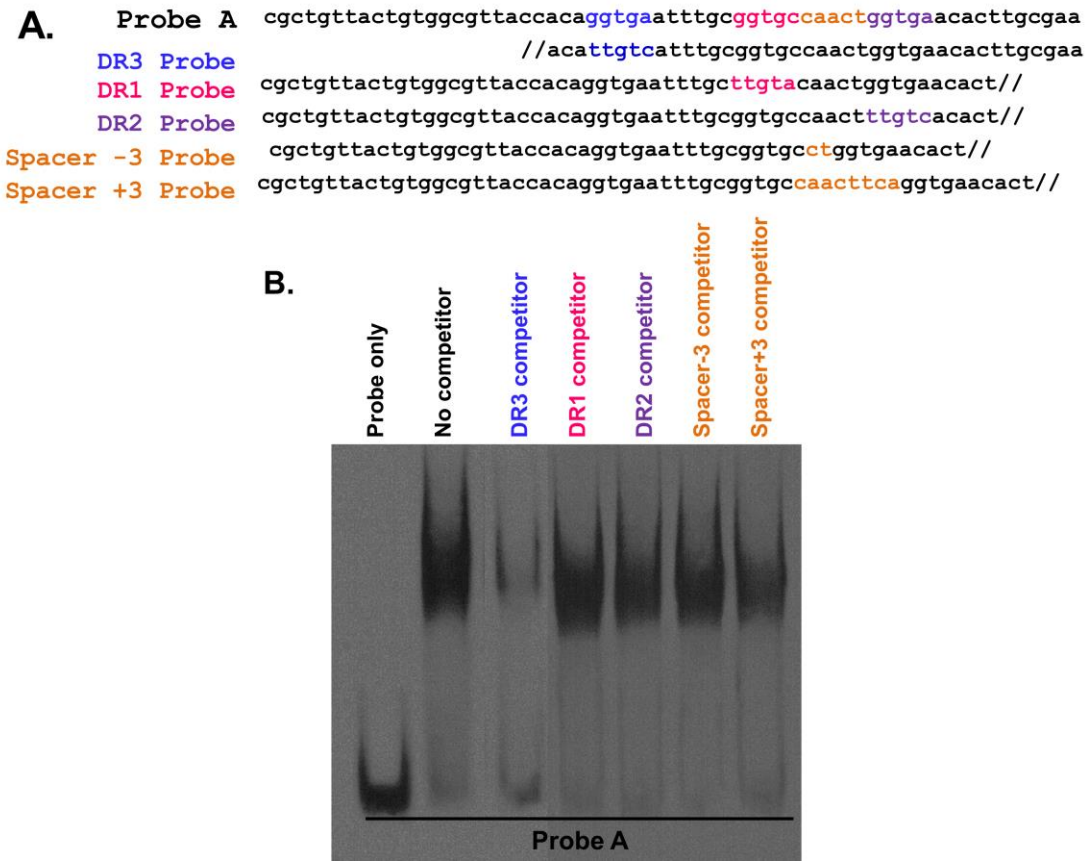

Figure 3.3 Evaluation of essential binding elements within the Probe A sequence. A) Schematic depicting the sequences of the mutated unlabeled competitiors compared to the labeled Probe A sequence. Competitors are color coded to coincide with the EMSA results in panel B.
B) EMSA analysis of binding between purified His_6-RegX3, DIG-labeled Probe A and unlabeled competitors DR3, DR1, DR2, Spacer +3 or Spacer -3. 0.5 µg of purified His_6-RegX3 was incubated with 200 ng unlabeled competitors for 15 minutes, then 0.5 ng labeled Probe A was added and incubated for 15 minutes.

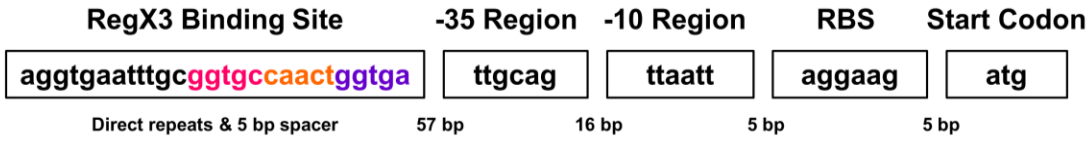

Figure 3.4 Putative RegX3 promoter organization. Proposed organization of the *esx-5* RegX3 promoter region, containing the canonical -35 and -10 regions, as well as the RBS, upstream of the *pe19* translational start site.

Constructing RegX3 binding site sequence mutants in WT and ΔpstA1 Mtb

Now that the RegX3 binding site sequence within the *esx-5* locus has been identified, targeted mutation of this region can be accomplished. Four different binding site mutations were constructed, outlined in **Figure 3.5**. Each direct repeat was mutated individually, using the same transversion mutations that were tested in the EMSA experiments. The spacer region between DR1 and DR2 was mutated by adding three bases. Finally, the entire 27 bp binding site was deleted. Mutations were introduced into both WT and Δ*pstA1* mutant *Mtb* strain backgrounds.

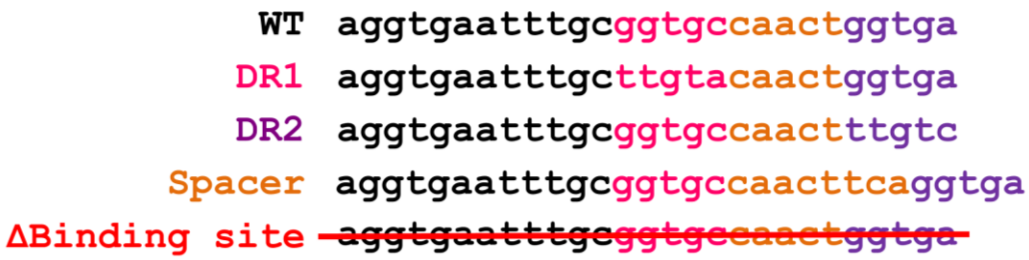

Figure 3.5 RegX3 binding site sequence mutant constructs. A schematic illustrating the sequences within the *esx-5* locus that were mutated in each of the four RegX3 binding site mutants.

RegX3 binding mutations in WT Mtb suppress esx-5 overexpression and ESX-5 protein production and hyper-secretion

We previously demonstrated that ESX-5 activity is triggered by P_i limitation in WT *Mtb*, and this response requires RegX3 [158]. To determine if the RegX3 binding site sequence we identified *in vitro* is also important for regulation of *esx-5* transcription in response to P_i limitation *in vivo,* we conducted qRT-PCR experiments to examine how *esx-5* gene expression is affected when the RegX3 binding site sequence is mutated. To be clear, all references in this work to the RegX3 binding site sequence are located within the intergenic region upstream of *pe19* within the *esx-5* locus. WT, Δ*regX3,* WT_{DR2} and WT_{Spacer} *Mtb* were grown in either P_i-free or P_i-rich medium, as a control. Depending on the mutation introduced, different transcriptional effects were detected. In the P_i-rich controls, *esx-5* transcripts were expressed at a basal level and not induced, consistent

with previous results (**Figure 3.6A**). When grown in P_i-free medium, *pe19, espG*$_5$ and *eccD*$_5$ transcripts were overexpressed in WT *Mtb* compared to the P_i-rich controls, and compared to the Δ*regX3* mutant grown in P_i-free medium (**Figure 3.6B**). This finding is also consistent with previously published results. Notably *espG*$_5$ and *eccD*$_5$ transcripts in the WT_{DR2} mutant were not significantly different from Δ*regX3* (*P*=0.47, 0.2), and abundance of the *pe19* transcripts was also close to the Δ*regX3* mutant level. Conversely, in the WT_{Spacer} mutant transcription of *pe19, espG*$_5$ and *eccD*$_5$ transcripts was significantly different from the Δ*regX3* mutant (*P*= 0.004, 0.039, 0.033), though transcription did not reach WT levels of induction during P_i starvation. Further, *espG*$_5$ and *eccD*$_5$ transcripts were not statistically different from WT during P_i starvation (*P*=0.148, 0.166). Taken together, the WT_{DR2} mutant data conclusively demonstrate that activation of *esx-5* transcription is mediated directly by RegX3 *Mtb*, and establishes that the DR2 sequence is essential for RegX3 binding *in vivo*.

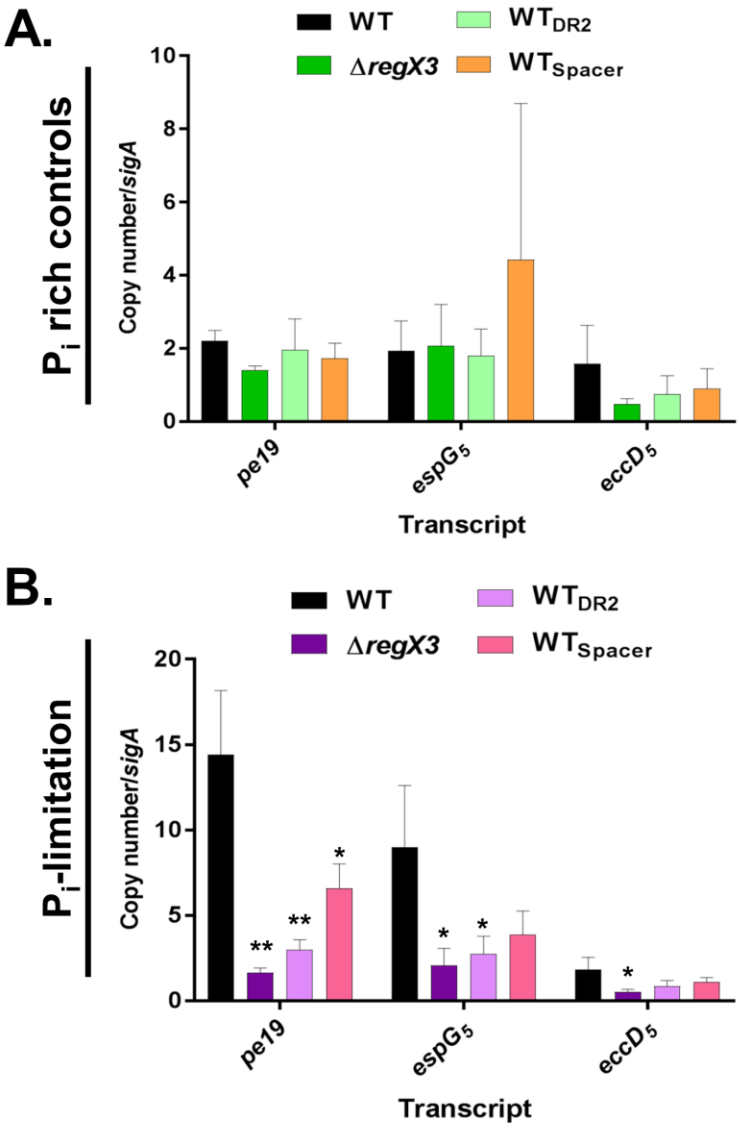

Figure 3.6 RegX3 binding site sequence mutations in WT *Mtb* reverse induction of *esx-5* expression during P_i starvation. Quantitative RT-PCR analysis of *esx-5* transcription. Wild-type *Mtb* Erdman (WT), Δ*regX3,* WT_{DR2} and WT_{Spacer} were cultured in A) P_i-rich 7H9 or B) in P-free 7H9 medium for 24 hours and RNA was extracted. Abundance of *pe19, espG$_5$* and *eccD$_5$* transcripts were determined by qRT-PCR and normalized to the housekeeping gene, *sigA*.. Data shown are the means +/- standard deviations of three independent experiments. Asterisks indicate statistically significant differences from WT (*, *P*<0.05; **, *P*<0.01).

To determine whether the reversal of induction of *esx-5* transcription observed in the RegX3 binding site mutants translates to changes in protein, we monitored production of ESX-5 conserved components and secretion of ESX-5 proteins EsxN and PPE41 during P_i limitation by Western blot analysis. Secretion of EsxN and PPE41 was induced in the WT strain during P_i limitation relative to the Δ*regX3* null strain, as previously reported [158] (**Figure 3.7**). Notably, induction of EsxN and PPE41 secretion was suppressed in the WT_{DR2} mutant during P_i limitation. Indeed PPE41 was undetectable and EsxN was barely detectable. Conversely, in the WT_{Spacer} mutant induction of EsxN and PPE41 secretion was induced by P_i limitation, with levels similar to WT induction. Anti-ModD loading controls ensure equivalent amounts of culture filtrate fraction proteins were added, and anti-GroEL2 controls demonstrate there is no contamination from cell lysate fraction proteins.

Production of the cytosolic protein chaperone, $EspG_5$, and the ESX-5 secretion machinery components $EccB_5$ and $EccD_5$ was induced in WT *Mtb* during P_i limitation relative to the Δ*regX3* mutant, as previously demonstrated [158]. However, induction of $EspG_5$, $EccB_5$ and $EccD_5$ in both the WT_{DR2} and WT_{Spacer} mutants was suppressed relative to WT induction. Taken together, these data demonstrate that mutating the RegX3 binding site within the *esx-5* locus can suppress induction of EsxN and PPE41 secretion during P_i limitation. This is further evidence that RegX3 directly mediates activation of ESX-5 secretion in response to P_i limitation in *Mtb.*

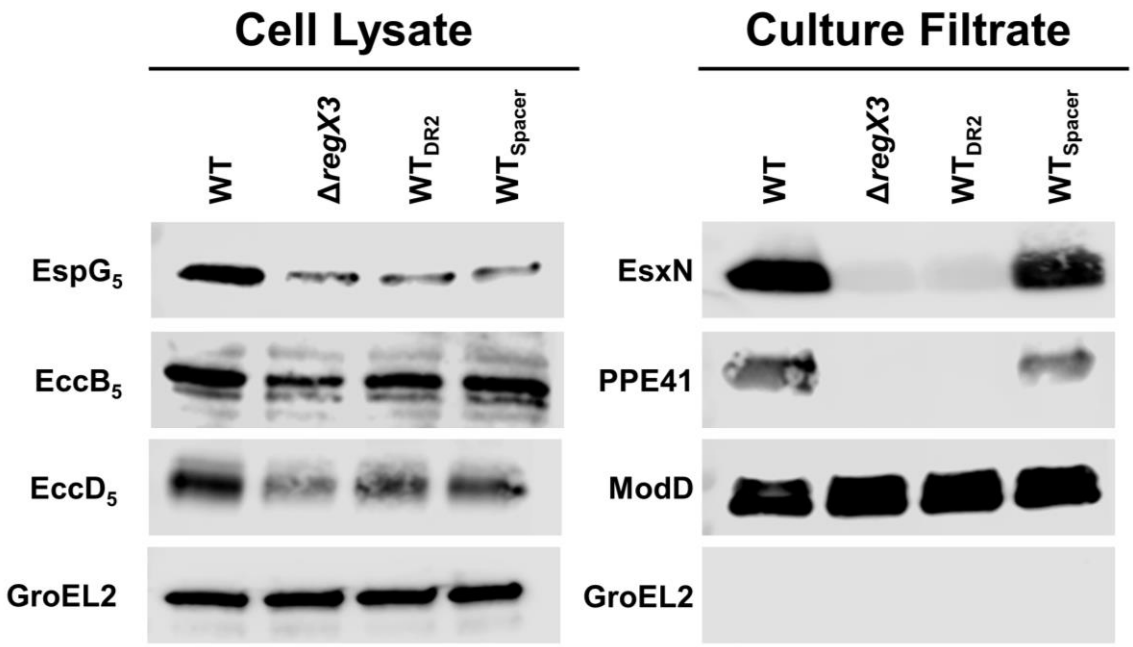

Figure 3.7 Secretion of ESX-5 substrates and production of ESX-5 conserved components during P_i limitation is reversed by RegX3 binding site mutations. Wild-type *Mtb* Erdman (WT), Δ*regX3,* WT_{DR2} and WT_{Spacer} were cultured in P_i-free Sauton's medium without Tween-80 to which 2.5 µm of KH_2PO_4 was added exogenously. 10 µg of cell lysate (CL) and 5 µg of culture filtrate (CF) proteins were subjected to SDS-PAGE and Western blot analysis. Anti-GroEL2 was used as both a loading control for the CL fraction, and a cell lysis control for the CF fraction. Anti-ModD was used as a loading control for the CF fraction. Results are representative of two independent experiments.

RegX3 binding site mutations in the Mtb ΔpstA1 mutant reverse esx-5 overexpression and hyper-secretion of EsxN

Previously, our lab demonstrated that RegX3 is constitutively activated in the *Mtb* Δ*pstA1* mutant [149]. Further, *esx-5* transcripts are overexpressed in the Δ*pstA1* mutant in a RegX3-dependent manner [158]. To determine whether *esx-5* RegX3 binding site mutations suppress RegX3-dependent overexpression of these transcripts, binding site mutations were introduced in the Δ*pstA1* mutant background. Then, qRT-PCR experiments were performed to examine *esx-5* gene expression in Δ*pstA1* RegX3 binding site mutants. As previously reported, the Δ*pstA1* mutant exhibited overexpression of *pe19, espG*$_5$, and *eccD*$_5$ transcripts, and overexpression was dependent on RegX3; *esx-5* transcription was restored to WT levels in the Δ*pstA1*Δ*regX3* mutant [158] (**Figure 3.8**). Overexpression of *pe19, espG*$_5$, and *eccD*$_5$ transcripts were reversed in both the Δ*pstA1*$_{DR2}$ and Δ*pstA1*ΔBS mutants, similar to the Δ*pstA1*Δ*regX3* mutant. Transcripts of *pe19* and *espG*$_5$ in both Δ*pstA1*$_{DR2}$ and Δ*pstA1*ΔBS mutants were not significantly different from Δ*pstA1*Δ*regX3* (*pe19 P*= 0.65, 0.14; *espG*$_5$ *P*= 0.54, 0.12). Transcripts of *pe19, espG*$_5$, and *eccD*$_5$ were detected at intermediate levels in the Δ*pstA1*$_{Spacer}$ mutant. The *pe19* and *espG*$_5$ transcripts in the Δ*pstA1*$_{Spacer}$ mutant were statistically different from Δ*pstA1*Δ*regX3* (*pe19 P*= 0.0006; *espG*$_5$ *P*= 0.013). Transcripts of *pe19, espG*$_5$, and *eccD*$_5$ were detected at intermediate levels in the Δ*pstA1*$_{Spacer}$ mutant. Taken together, these data demonstrate that RegX3 directly mediates overexpression of *esx-5* genes in the Δ*pstA1* mutant of *Mtb*, and that the DR2 sequence is essential for this regulation *in vivo*.

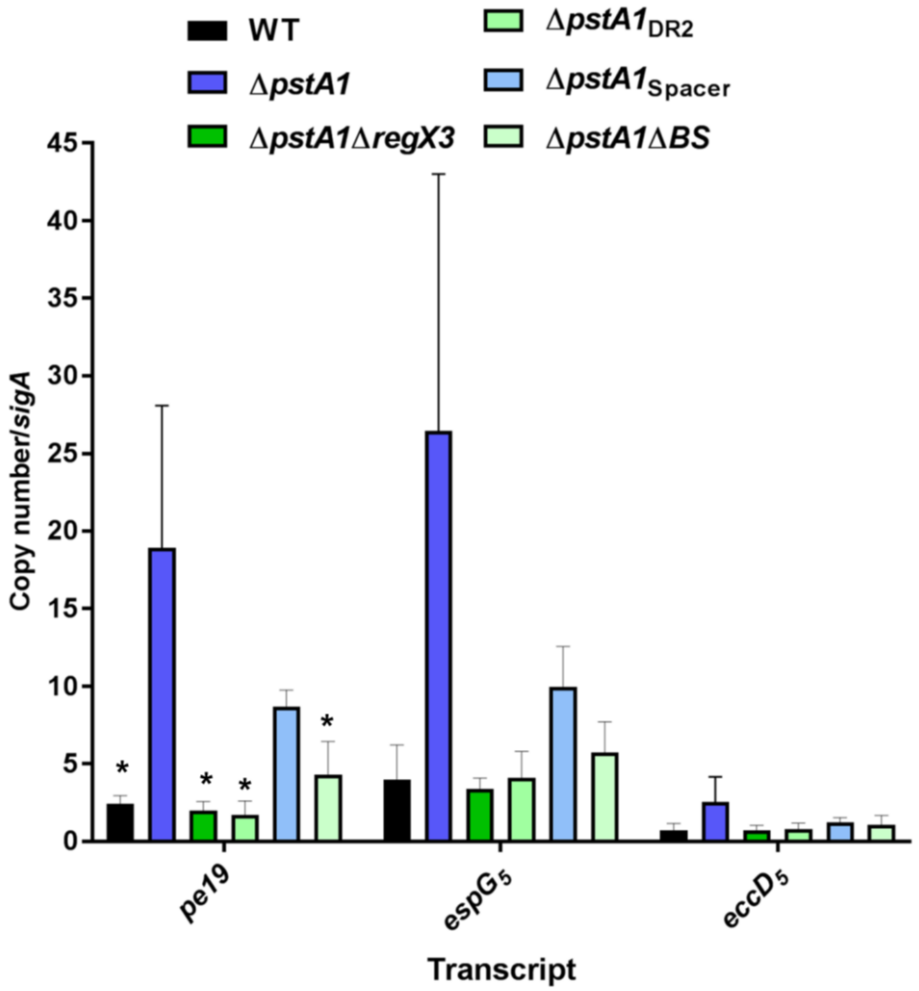

Figure 3.8 RegX3 binding sequence mutations reverse *esx-5* overexpression in Δ*pstA1 Mtb*. Quantitative RT-PCR analysis of *esx-5* transcription. Wild-type *Mtb* Erdman (WT), Δ*pstA1,* Δ*regX3,* Δ*pstA1*Δ*regX3,* Δ*pstA1*$_{DR2}$, Δ*pstA1*$_{Spacer}$ and Δ*pstA1*ΔBS were cultured in 7H9 medium to mid-exponential phase and RNA was extracted. Abundance of *pe19, espG*$_5$, and *eccD*$_5$ transcripts were determined by qRT-PCR and normalized to the housekeeping gene, *sigA*. Data shown are the means +/- standard deviations of three independent experiments. Asterisks indicate statistically significant differences from Δ*pstA1* (*, *P*<0.05).

RegX3 is a global response regulator, which activates and represses many genes outside of the *esx-5* locus [149]. To investigate whether the RegX3 binding site sequence mutants resulted in regulatory perturbations affecting only the *esx-5* locus, or if global RegX3 activity was altered in some way, we examined transcription of two genes that are highly upregulated in the Δ*pstA1* mutant in a RegX3-dependent manner, but that are not associated with *esx-5* [149]. The *udgA* and *mgtA* transcripts were overexpressed in the Δ*pstA1* mutant relative to both WT and Δ*pstA1*Δ*regX3* strains (**Figure 3.9**). Importantly, *udgA* and *mgtA* transcripts were still overexpressed in Δ*pstA1*$_{DR2}$, Δ*pstA1*$_{Spacer}$ and Δ*pstA1*ΔBS bacteria. These data demonstrate that global RegX3 activity is unaffected by mutation of the RegX3 binding site sequence within the *esx-5* locus, and all phenotypes associated with these mutations can be attributed to changes in *esx-5* transcription.

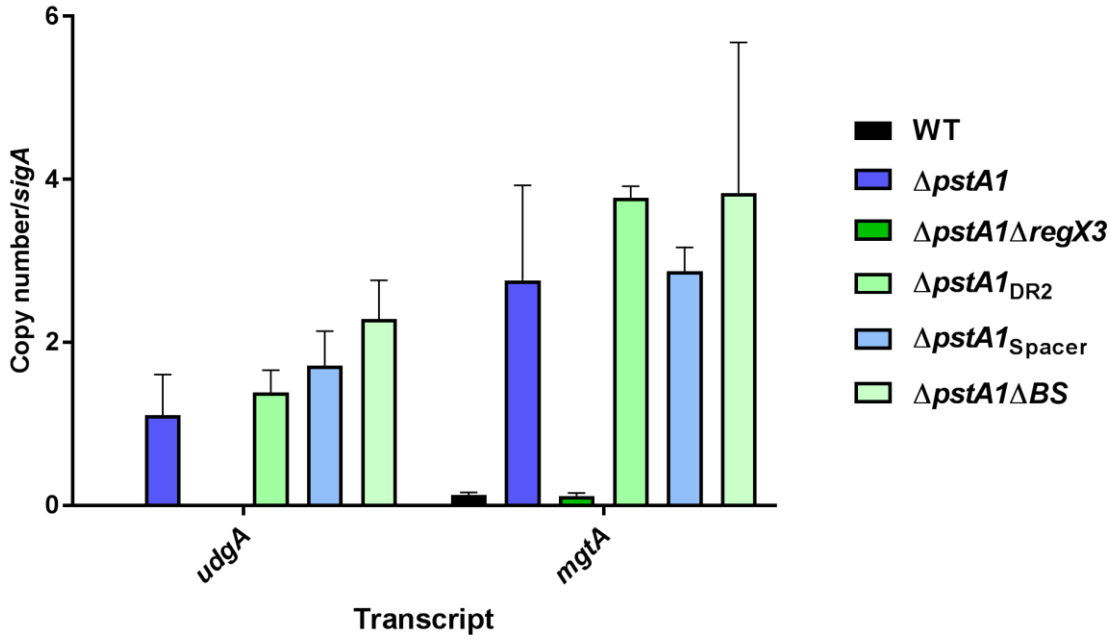

Figure 3.9 RegX3 binding site sequence mutations do not alter global RegX3 activity. Quantitative RT-PCR analysis of *esx-5* transcription. Wild-type *Mtb* Erdman (WT), Δ*pstA1,* Δ*regX3,* Δ*pstA1*Δ*regX3,* Δ*pstA1*$_{DR2}$, Δ*pstA1*$_{Spacer}$ and Δ*pstA1*ΔBS were cultured in 7H9 medium to mid-exponential phase and RNA was extracted. Abundance of *udgA* and *mgtA* transcripts were determined by qRT-PCR and normalized to the housekeeping gene, *sigA*. Data shown are the means +/- standard deviations of two independent experiments.

ESX-5 protein production and hyper-secretion in the Δ*pstA1* RegX3 binding site mutants was also evaluated using Western blot analysis. As with the WT mutants, different RegX3 binding sequence mutations produced a range effects on protein secretion. In line with published data, hyper-secretion of EsxN and PPE41 and overproduction of $EspG_5$, $EccB_5$ and $EccD_5$ was detected in the Δ*pstA1* mutant relative to WT *Mtb*, and the response required RegX3 [158] (**Figure 3.10**). Notably, EsxN hyper-secretion was suppressed in both the Δ$pstA1_{DR2}$ and Δ*pstA1*ΔBS mutants; the protein was undetectable, comparable to the Δ*pstA1* Δ*regX3* mutant. Contrastingly, EsxN secretion was detected in the Δ$pstA1_{Spacer}$ mutant, though the protein was less abundant than the Δ*pstA1* single mutant. Hyper-secretion of PPE41 was suppressed in the Δ$pstA1_{DR2}$ mutant, similar to Δ*pstA1* Δ*regX3* mutant levels. PPE41 was less abundant in the Δ*pstA1*ΔBS mutant compared to Δ*pstA1*. PPE41 secretion in the Δ$pstA1_{Spacer}$ mutant is hyper-secreted almost to the degree of the Δ*pstA1* single mutant. Anti-ModD loading controls ensure equivalent amounts of culture filtrate fraction proteins were added, and anti-GroEL2 controls demonstrate there is no contamination from cell lysate fraction proteins.

There was less $EspG_5$, $EccB_5$ and $EccD_5$ protein detected in all three Δ*pstA1* RegX3 binding sequence mutants compared to the Δ*pstA1* mutant. There appeared to be similar levels of $EspG_5$ and $EccB_5$ in the three Δ*pstA1* RegX3 binding sequence mutants and Δ*pstA1* Δ*regX3*. There may be slightly more $EccD_5$ protein detected in the Δ$pstA1_{Spacer}$ mutant compared to Δ$pstA1_{DR2}$ and Δ*pstA1*ΔBS and the Δ*pstA1* Δ*regX3* control. Anti-GroEL2 loading controls ensure equivalent amounts of cell lysate fraction proteins were added. Taken together, these results indicate that the Δ$pstA1_{DR2}$ and Δ*pstA1*ΔBS RegX3 binding site mutations reverse hyper-secretion of EsxN in the *Mtb* Δ*pstA1* mutant.

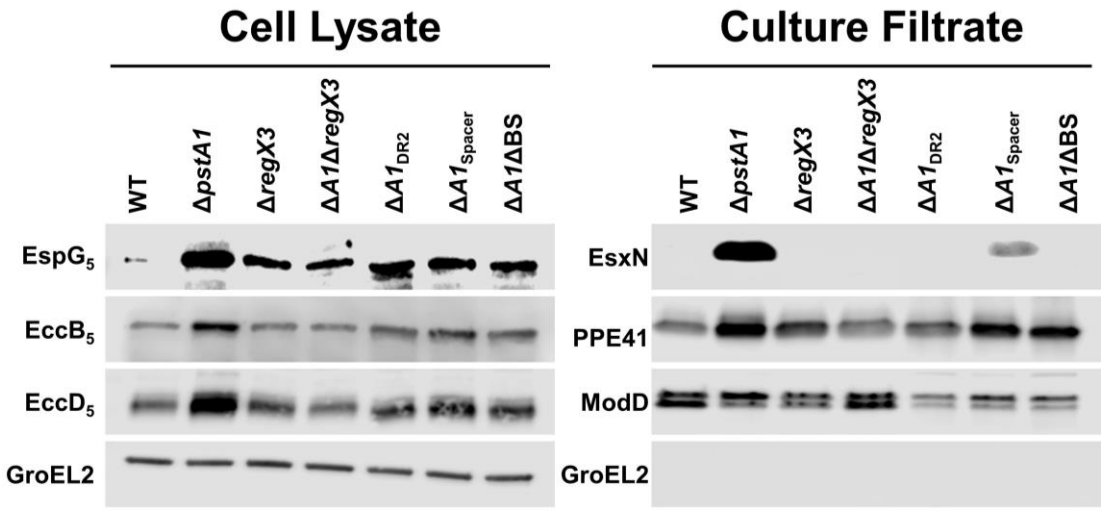

Figure 3.10 Hyper-secretion of EsxN is suppressed by RegX3 binding site sequence mutations in Δ*pstA1 Mtb*. Wild-type *Mtb* Erdman (WT), Δ*pstA1,* Δ*regX3,* Δ*pstA1*Δ*regX3,* Δ*pstA1*$_{DR2}$, Δ*pstA1*$_{Spacer}$ and Δ*pstA1*ΔBS were cultured in Sauton's complete medium without Tween-80 for 5 days. 10 µg of cell lysate (CL) and 5 µg of culture filtrate (CF) proteins were subjected to SDS-PAGE and Western blot analysis. Anti-GroEL2 was used as both a loading control for the CL fraction, and a cell lysis control for the CF fraction. Anti-ModD was used as a loading control for the CF fraction. Results are representative of two independent experiments.

Deletion of the esx-5 RegX3 binding site in ΔpstA1 Mtb partially restores virulence in vivo

To test whether suppression of ESX-5 secretion can reverse Δ*pstA1* attenuation, C57BL/6 and IrgM1$^{-/-}$ mice were infected via the aerosol route with ~100 CFU of WT, Δ*pstA1,* or Δ*pstA1*ΔBS *Mtb*. All IrgM1$^{-/-}$ mice succumbed to infection with WT *Mtb* by 4 weeks post infection, and bacterial loads reached ~10^9 in the lungs (**Figure 3.11A**). Replication of the Δ*pstA1* mutant was controlled by IrgM1$^{-/-}$ mice at 2 weeks post infection, consistent with previous results [149]. In contrast, deletion of the entire *esx-5* RegX3 binding site sequence in the Δ*pstA1*ΔBS mutant partially reversed the attenuation of the Δ*pstA1* single mutant. At 4 weeks post infection, bacterial CFU in the lungs of Δ*pstA1* and Δ*pstA1*ΔBS were similar. However, by 6 weeks post infection, the Δ*pstA1*ΔBS mutant reaches nearly the same bacterial burden in the lungs as WT *Mtb,* albeit with delayed kinetics. Notably, one mouse infected with Δ*pstA1*ΔBS died the day before the final time point (data not shown), and based on the bacterial burden and lung pathology, the mouse likely succumbed to the *Mtb* infection. This indicates how much more virulent the Δ*pstA1*ΔBS strain is compared to Δ*pstA1,* since previous experiments have determined IrgM1$^{-/-}$ mice do not succumb to infection with the Δ*pstA1* mutant even after 14 weeks [149].

Similarly, in C57BL/6 mice, deletion of the *esx-5* RegX3 binding site sequence partially restored virulence in the Δ*pstA1* mutant (**Figure 3.11B**). The Δ*pstA1*ΔBS bacteria are controlled at the same burden in the lungs as the Δ*pstA1* mutant until week 4 post infection, recapitulating the phenotype of this mutant in IrgM1$^{-/-}$ mice. However, by week 8 post infection, the Δ*pstA1*ΔBS mutant reached similar numbers of CFU as the WT strain (P= 0.65). The Δ*pstA1*ΔBS mutant lung CFU were significantly elevated at the 8 week time point compared to Δ*pstA1* bacteria. Taken together, these data demonstrate that deletion of the RegX3 binding site sequence in the *esx-5* locus partially suppressed the virulence defect of the Δ*pstA1* mutant. Further, these data suggest that aberrant hyper-secretion of ESX-5 substrates is at least partially responsible for causing attenuation of the Δ*pstA1* mutant in both C57BL/6 and IrgM1$^{-/-}$ mice.

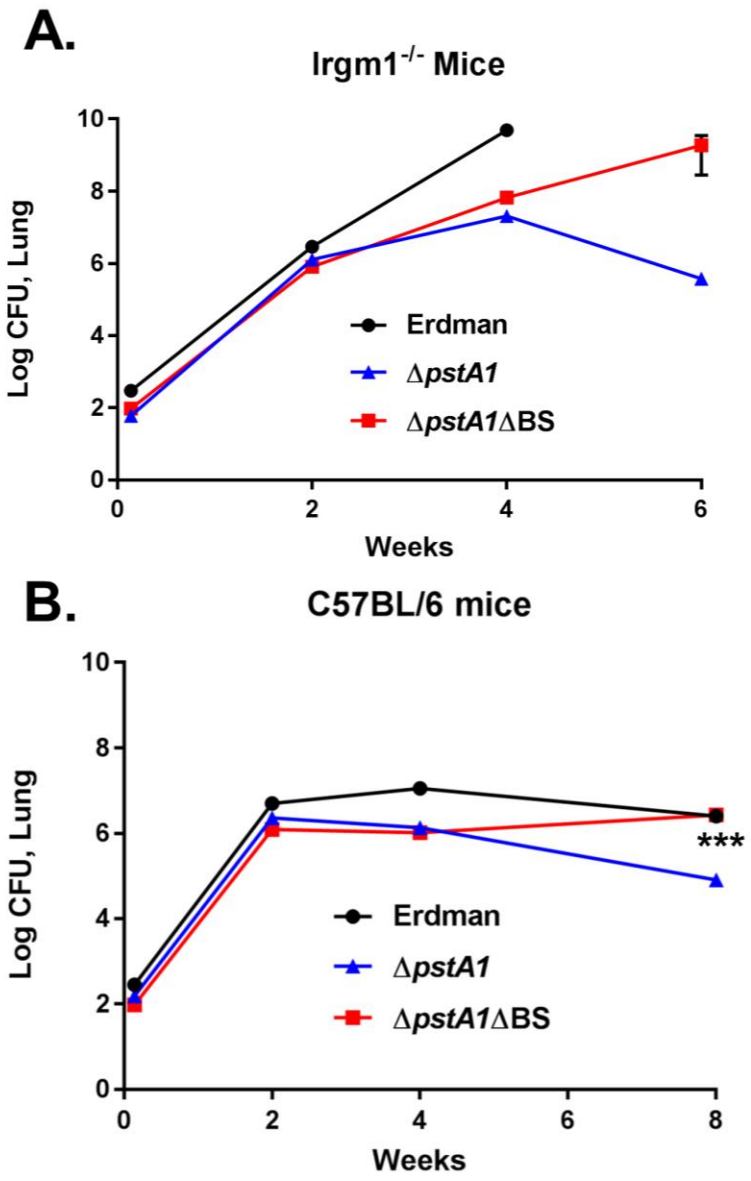

Figure 3.11 Deletion of the *esx-5* RegX3 binding site sequence in the Δ*pstA1* mutant restores virulence in C57BL/6 and IrgM1$^{-/-}$ mice. A) C57BL/6 and B) IrgM1$^{-/-}$ mice were infected via the aerosol route with ~100 CFU of wild-type Erdman, Δ*pstA1* or Δ*pstA1*ΔBS *Mtb.* Groups of infected mice (*n*=4) were sacrificed at the indicated time points. Bacterial CFU were enumerated by plating lung homogenates on 7H10 agar and incubating at 37°C for ~3-4 weeks. Symbols represent means; error bars indicate standard error of the means. Asterisks indicate statistically differences from Δ*pstA1* (*** = *P* <0.0001).

Discussion

Previously, our work revealed that the virulence-associated ESX-5 secretion system is regulated at the transcriptional level by the Pst/SenX3-RegX3 system. Further, we identified a novel environmental signal, P_i limitation, which induces activity of ESX-5. Work outlined in this chapter demonstrates that RegX3-mediated dysregulation of ESX-5 secretion causes attenuation *in vivo*. We propose that precise regulation of ESX-5 is a critical function of the Pst/SenX3-RegX3 system during infection.

While our previous work demonstrated that regulation mediated through Pst/SenX3-RegX3 is direct, the RegX3 binding site sequence upstream of *pe19* remained unknown. We have since identified the RegX3 binding site sequence within the *esx-5* locus to positions -128 to -101, upstream of the *pe19* start codon. While there is quite a range of variability among consensus sequences of mycobacterial promoter regions, *in silico* analysis allows us to predict canonical -35, -10 and RBS regions downstream of the RegX3 binding site sequence and upstream of the *pe19* start codon (**Figure 3.4**) [208]. Our previous work indicates that the *esx-5* RegX3 dependent transcriptional start site is located around -45 to -90 relative to the PE19 translation start site [158]; other studies mapping transcriptional start sites uncovered two potential sites near this region, at -38 and -133, relative to the PE19 translation start [195, 196]. This putative *esx-5* RegX3 promoter arrangement simply demonstrates that the RegX3 binding site sequence identified in this work is upstream of the predicted transcriptional start site, which is consistent with where transcriptional activators are located.

While identifying the RegX3 binding site sequence within the *esx-5* locus, we uncovered an imperfect direct repeat sequence (**GGTGC**caact**GGTGA**) located at the 3' end of the binding site that is essential for RegX3 binding. EMSA studies show that mutation of either direct repeat abrogates RegX3 binding *in vitro*. RegX3 requires an intact imperfect direct repeat sequence for binding within *Mtb*, since mutation of the 3' repeat in this sequence altered *esx-5* transcription. These data demonstrate the importance of the imperfect direct repeat binding elements within *esx-5* for RegX3 binding.

Mutating the RegX3 binding site sequence in WT *Mtb* represses induction of *esx-5* transcription and secretion of the ESX-5 substrates EsxN and PPE41 during P_i limitation. Additionally, production of certain ESX-5 secretion machinery components, along with

the ESX-5 cellular protein chaperone, was also repressed by mutation of the RegX3 binding site sequence in WT *Mtb* during P_i limitation. Consistent with these results, overexpression of *esx-5* transcripts was also suppressed in Δ*pstA1* RegX3 binding site mutants. Notably, hyper-secretion of EsxN and PPE41 was reversed in Δ*pstA1 Mtb* when the RegX3 binding site was mutated or deleted. Together, these results conclusively demonstrate that RegX3 directly binds within the *esx-5* locus to regulate transcription in response to P_i limitation.

RegX3 mediated regulation occurs at the transcriptional level, but previous work from our lab has shown that *eccB*$_5$ and *ppe41* transcripts are not induced during P_i limitation, although the protein levels were increased [158]. Because of this, we do not expect RegX3 binding site mutants to alter *eccB*$_5$ or *ppe41* transcription, though we have not yet tested this directly. However, we did observe changes in $EccB_5$ or PPE41protein when the RegX3 binding site sequence was mutated. The decreased secretion of PPE41 observed in the RegX3 binding site mutants could be due to a concurrent decrease in $EspG_5$ chaperone production. We also observed decreased production of the ESX-5 conserved components $EccB_5$ and $EccD_5$ in the RegX3 binding sequence mutants, and this may destabilize the entire secretion complex, which could also affect levels of PPE41 secretion.

Our lab has previously demonstrated that Δ*pstA1* is attenuated in several mouse strains, including C57BL/6 and IrgM1$^{-/-}$ mice. My work has shown that highly antigenic ESX-5 substrates are constitutively hyper-secreted in the Δ*pstA1* mutant [158, 175]. We proposed that aberrant hyper-secretion of antigenic ESX-5 substrates was somehow sensitizing Δ*pstA1* bacteria to host immune defenses. Notably, we found that deleting the *esx-5* RegX3 binding site sequence in the Δ*pstA1* mutant partially restored virulence in C57BL/6 and IrgM1$^{-/-}$ mice. This supports a model where inappropriate ESX-5 hyper-secretion mediated by RegX3 contributes to Δ*pstA1* attenuation. While deletion of the RegX3 binding site sequence does restore virulence to the Δ*pstA1* mutant, bacterial CFU in the lungs do not quite reach WT levels. This failure to fully restore virulence suggests that an inability to induce *esx-5* expression during infection is also detrimental to bacterial survival. Alternatively, it is possible that other factors in the RegX3 regulon could be contributing to attenuation of the Δ*pstA1* mutant. The Δ*regX3* mutant also has a virulence defect *in vivo*, demonstrating that some component of the RegX3 regulon is

required for full *Mtb* virulence [149]. Our work showing that deletion of the *esx-5* RegX3 binding site sequence reverses attenuation of Δ*pstA1* bacteria *in vivo* supports that precise regulation of ESX-5 is required for full *Mtb* pathogenesis. In the future, we will test an *esx-5* RegX3 binding site mutant in WT *Mtb* to determine whether the residual attenuation in the Δ*pstA1*ΔBS mutant is due to failure to induce ESX-5 activity. Cloning an *esx-5* RegX3 binding site deletion mutation in WT bacteria has been attempted, though screening for a deletion mutant has been unsuccessful.

RegX3 is a global response regulator, and activates and represses many genes outside of the *esx-5* locus. Mutating the DR2 sequence in WT *Mtb* eliminated RegX3 binding and repressed induction of *esx-5* transcription during P_i limitation. This mutant uncouples ESX-5 regulation from global cellular events mediated by RegX3. The WT_{DR2} mutant will be an important tool that will allow us to investigate the importance of ESX-5 secretion for *Mtb* pathogenesis.

Although altering the spacing between the direct repeat sequences abrogated RegX3 binding *in vitro*, this same mutation did not completely eliminate RegX3 binding *in vivo*. Mutation of the RegX3 binding site sequence spacer region did not repress induction of *esx-5* gene expression in WT *Mtb* during P_i limitation or suppress *esx-5* overexpression in the Δ*pstA1* mutant. Similarly, altering the spacer region did not eliminate induction of EsxN or PPE41 secretion in response to P_i limitation in WT *Mtb* or suppress constitutive hyper-secretion of these proteins in Δ*pstA1* bacteria. Clearly, RegX3 maintains some ability to bind within *esx-5* when the spacer region between DR1 and DR2 has been mutated, though the interaction is impaired and never reaches fully induced levels of either transcript or protein. Perhaps the DR3 region is able to compensate for mutation of the other direct repeat sequence *in vivo*, though it was found to be dispensable *in vitro*. It is possible that the DR3 sequence is more important *in vivo* than it is *in vitro*. Yet another possibility is that the DR2 region alone is sufficient to allow activation of *esx-5* transcription *in vivo*, and so alteration of the spacer fails to eliminate transcription because DR2 is intact and RegX3 can still bind this region. Binding events inside a cell are inherently messier and more complicated than *in vitro*, and it is possible that other factors compete for binding within this region *in vivo*.

We have identified the RegX3 binding site sequence upstream of *pe19* within the *esx-5* locus, and subsequent targeted mutation of the region demonstrated that these

sequences are required for induction of *esx-5* transcription during P_i limitation. Further, our work demonstrates that regulation of ESX-5 mediated by Pst/SenX3-RegX3 is important for *Mtb* pathogenesis, since deletion of the *esx-5* RegX3 binding site sequence reversed the virulence defect of the Δ*pstA1* mutant during infection in C57BL/6 and IrgM1$^{-/-}$ mice. Our future experiments will explore the effect of *esx-5* dysregulation in WT *Mtb*, by infecting C57BL/6 and IrgM1$^{-/-}$ mice with a RegX3 binding site mutant in the WT background. Further, characterization of any potential differences between the inflammatory immune responses to WT, Δ*pstA1* and WT RegX3 binding site mutants *in vivo* will provide further clues to the effect of ESX-5 dysregulation on *Mtb* pathogenesis.

Acknowledgements

The anti-sera against EsxN, PPE41, $EspG_5$, $EccB_5$ and $EccD_5$ were generated by Dylan White. The following reagents were obtained through BEI Resources, NIAID, NIH: monoclonal anti-*Mycobacterium tuberculosis* GroEL2 (Gene Rv0440) and polyclonal anti-*Mycobacterium tuberculosis* Mpt32 (Gene Rv1860) (antiserum, rabbit), NR-13807.

Chapter 4

Evaluating the role of EsxN in *Mycobacterium tuberculosis* pathogenesis

Summary

ESX-5, a specialized Type VII protein secretion system found in *Mycobacterium tuberculosis* (*Mtb*), is activated in response to inorganic phosphate (P_i) limitation by the Pst/SenX3-RegX3 system. Further, ESX-5 has been implicated in *Mtb* pathogenesis through modulation of the host response. Work described in chapter 3 has demonstrated that dysregulation of ESX-5 through constitutive transcriptional activation causes attenuation *in vivo*. A Δ*pstA1* mutant is attenuated in several mouse lines due to constitutive activation of RegX3. ESX-5 proteins are overproduced and hyper-secreted in the Δ*pstA1* mutant. Given that Esx proteins are highly antigenic, we proposed that aberrant hyper-secretion of EsxN, a protein with unknown function secreted by ESX-5, sensitized Δ*pstA1* bacteria to host responses. We generated *esxN* deletion mutants in both WT and Δ*pstA1* backgrounds, and assessed survival of these strains in a variety of relevant *in vitro* stress conditions. While *esxN* was dispensable for most stresses tested, it was required to fully resist reactive nitrogen species *in vitro.* Strikingly, deleting *esxN* in Δ*pstA1* bacteria partially restored virulence in $Nos2^{-/-}$, but not C57BL/6 or $IrgM1^{-/-}$ mice, indicating that aberrant *esxN* expression in the Δ*pstA1* mutant is partially responsible for its attenuation. This work also suggests that reactive nitrogen species resistance is a possible function of EsxN.

Introduction

Our lab has previously demonstrated that antigenic ESX-5 substrates, including EsxN, are hyper-secreted in a Δ*pstA1* mutant in a RegX3-dependet manner [158]. The Δ*pstA1* mutant is attenuated for replication and virulence in C57BL/6, IrgM1$^{-/-}$ and Nos2$^{-/-}$ mice [149]. Attenuation of the Δ*pstA1* mutant in Nos2$^{-/-}$ mice is caused, in part, by constitutive activation of RegX3, since deletion of *regX3* reversed the virulence defect [149]. Experiments in chapter 3 using targeted *esx-5* RegX3 binding site mutations has shown that attenuation of the Δ*pstA1* mutant in C57BL/6 and IrgM1$^{-/-}$ mice was in fact due to aberrant activation of ESX-5 by RegX3. Work described in this chapter sought to understand whether hyper-secretion of one ESX-5 protein, EsxN, is involved in sensitization of the Δ*pstA1* mutant to host responses *in vivo.*

EsxN is a highly immunogenic protein that is actively secreted during human infection, though its function is unknown [106, 107]. Esx proteins are typically co-transcribed and encoded adjacently, and are secreted as heterodimers [81]. The putative partner to EsxN is EsxM, which is encoded directly upstream, but *esxM* contains a frameshift mutation that is predicted to truncate the protein [105]. However, there are four homologous *esxN/esxM* pairs encoded outside of the *esx-5* locus, and it is probable one of these EsxM-like paralogs pairs with EsxN [105]. The four *esxN/M* paralogs *esxI/J, esxK/L, esxO/P* and *esxV/W* share a 93-98% amino acid similarity to each other and to *esxM/N* [107, 217]. Together, all five *esxN* paralogs are referred to as the Mtb9.9 family, and all gene pairs are actively secreted during the course of human disease [217]. Due to the high sequence similarity of these paralogs, others have found that anti-EsxN antibodies are able to detect the other EsxN-like proteins [105]. While the function of EsxN is unknown, it may interact with the host to promote bacterial survival given that it is expressed by *Mtb* during infection.

Mtb is exposed to many stresses *in vivo,* including cell wall stress, acidification, ROS and RNS [46]. Many of these stresses are encountered with macrophage phagosomes, where the bacterium spends much of its time during infection. To survive these conditions, *Mtb* secretes many factors to modulate host immune responses and promote bacterial survival.

Many Esx proteins have been shown to manipulate host responses *in vivo*. ESX-1

secreted factors are key players in surviving host-mediated stresses. ESX-1 substrates, including EsxA and EsxB, promote bacterial survival *in vivo* through arrest of phagosomal maturation and permeabilization of the compartment [36, 94, 99, 218]. The ESX-3 secreted EsxG and EsxH proteins are involved in *Mtb* virulence, and deletion of these factors results in attenuation *in vivo* [93]. EsxH directly binds to the host endosomal sorting complex required for transport (ESCRT) machinery and impairs antigen presentation, ultimately leading to decreased priming and activity of T-cells [219]. ESX-1, -3 and -5 are the three ESX systems implicated in *Mtb* virulence. Since Esx proteins secreted from ESX-1 and -3 are essential for bacterial survival *in vivo* due to their roles in host modulation, it is possible that ESX-5 secreted Esx proteins are playing a similar role in counteracting immune responses.

Work in this chapter examines the impact of EsxN in *Mtb* pathogenesis, and investigates whether EsxN hyper-secretion plays a role in attenuation of Δ*pstA1 in vivo.* While EsxN is dispensable for survival during acid pH, ROS and cell wall stress conditions, our studies show that EsxN may be involved in RNS resistance in WT *Mtb*. Deletion of *esxN* does not reverse sensitivity of the Δ*pstA1* mutant to acid pH, ROS, RNS or cell wall stress. Further, deletion of *esxN* does not reverse attenuation of the Δ*pstA1* mutant in either C57BL/6 or IrgM1$^{-/-}$ mice. Intriguingly, the Δ*pstA1*Δ*esxN* mutant exhibits partial restoration of virulence in Nos2$^{-/-}$ mice*.* Together, our data demonstrate that hyper-secretion of EsxN sensitizes Δ*pstA1* bacteria to host responses in Nos2$^{-/-}$ mice.

Results

Constructing esxN deletion mutants in WT and ΔpstA1 Mtb

To investigate whether Δ*pstA1* attenuation *in vivo* is due to inappropriate hyper-secretion of antigenic ESX-5 substrates and further study the impact of ESX-5 secretion on *Mtb* pathogenesis, we deleted *esxN* in both WT and Δ*pstA1* backgrounds. *esxN* is a 300 bp gene encoded within the *esx-5* locus (**Figure 4.1A**). Secretion of EsxN is induced during P_i limitation in WT *Mtb,* and is hyper-secreted in Δ*pstA1* bacteria due to constitutive activation of RegX3. We examined *esxN* gene expression in the deletion

mutants to verify the deletion. Though *esxN* transcript abundance was reduced in both the Δ*esxN* and Δ*pstA1*Δ*esxN* strains, *esxN* transcripts were still detected. This is likely due to the other *esxN* paralogs encoded outside of the *esx-5* locus (**Figure 4.1B**). The decreased abundance of *esxN* transcripts observed in the *esxN* deletion strains supports that the *esx-5* encoded *esxN* has been deleted, though is not conclusive alone. Our results suggest that the qRT-PCR primers are not specific to *esxN*, but can also detect one or more *esxN* paralogs.

To determine if deletion of *esxN* caused any polar effects in the region, we also examined transcripts of the downstream gene $espG_5$ (**Figure 4.1B**). The $espG_5$ transcripts were overexpressed in the Δ*pstA1* relative to WT, and is consistent with published data [158]. Deletion of *esxN* in the Δ*pstA1* mutant did not have an effect on $espG_5$ transcript abundance. Similarly, $espG_5$ transcript abundance was unchanged when *esxN* was deleted in the WT background. Taken together, these data demonstrate deletion of *esxN* does not cause polar effects on the downstream gene.

We verified deletion of *esxN* in both WT and Δ*pstA1* bacteria using standard rt-PCR (**Figure 4.1D**). Using a forward primer located within *esxM*, upstream of *esxN*, and a reverse primer within the downstream $espG_5$ gene, we detected PCR products that were 300 bp smaller in both the Δ*esxN* and Δ*pstA1*Δ*esxN* strains compared to WT and Δ*pstA1* +RT controls, indicating that *esxN* had been deleted in those strains. There was no amplification of any product detected in the –RT negative control lanes, indicating there was no contamination with genomic DNA.

We also examined EsxN secretion in the *esxN* deletion strains. EsxN was hyper-secreted in the Δ*pstA1* mutant and was undetectable in WT, as previously published [158]. However, EsxN secretion was still observed in the Δ*pstA1*Δ*esxN* mutant, though at a decreased level compared to the Δ*pstA1* single mutant (**Figure 4.1C**). There was no observable EsxN in the Δ*esxN* mutant, similar to WT. Since we still detect secretion of EsxN in the Δ*pstA1*Δ*esxN* mutant, this implies that the anti-EsxN antibody can detect the other EsxN paralogs.

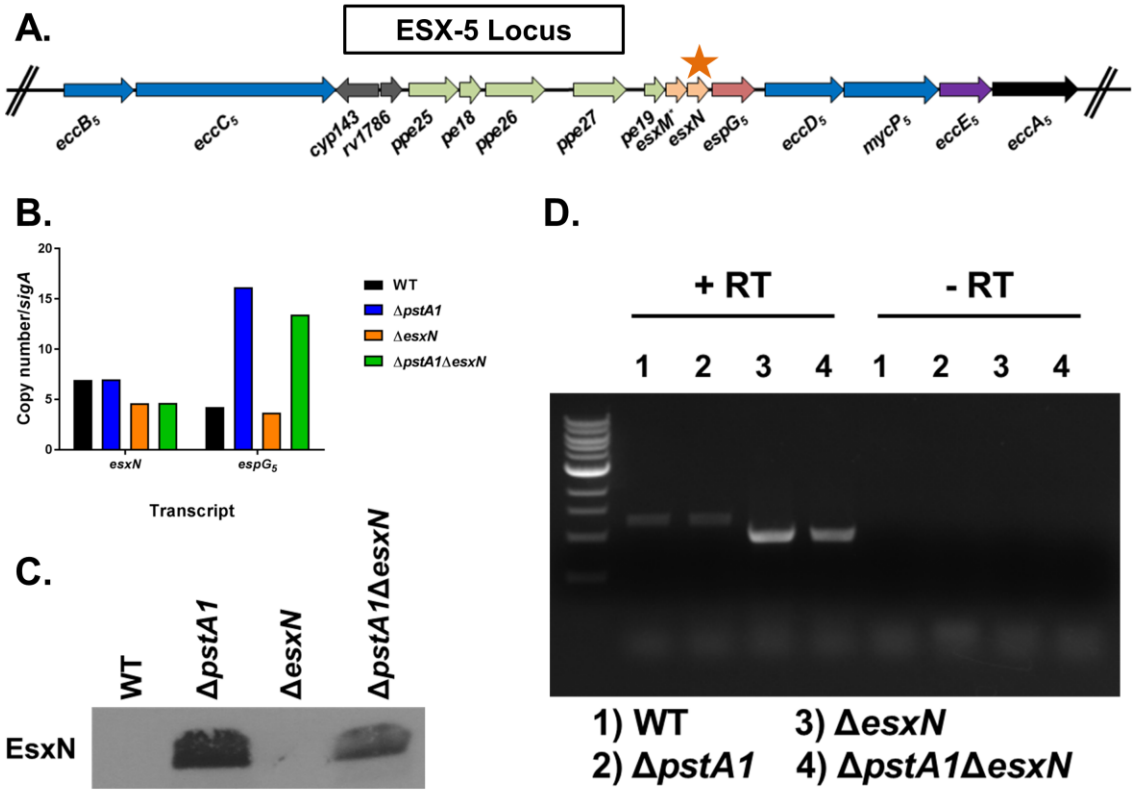

Figure 4.1 Validating *esxN* deletion mutants in WT and Δ*pstA1 Mtb*.
A) Schematic of the *esx-5* locus in *Mtb*. The orange star indicates the location of *esxN* within the locus. The * by *esxM* indicates there is a premature stop codon in this gene, resulting in a truncated protein.
B) Quantitative RT-PCR analysis of *esx-5* transcription. Wild-type *Mtb* Erdman (WT), Δ*pstA1*, Δ*esxN* and Δ*pstA1*Δ*esxN* were cultured in 7H9 medium to mid-exponential phase and RNA was extracted. Abundance of *esxN* and $espG_5$ transcripts were quantified using qRT-PCR and normalized to the housekeeping gene, *sigA*.
C) Wild-type *Mtb* Erdman (WT), Δ*pstA1*, Δ*esxN* and Δ*pstA1*Δ*esxN* were cultured in Sauton's complete medium without Tween-80 for 5 days. 5 µg of culture filtrate proteins were subjected to SDS-PAGE and Western blot analysis.
D) Standard rt-PCR to demonstrate deletion of *esxN*. Wild-type *Mtb* Erdman (WT), Δ*pstA1*, Δ*esxN* and Δ*pstA1*Δ*esxN* were cultured in 7H9 medium to mid-exponential phase and RNA was extracted. Primers within *esxM* and $espG_5$ were used to amplify a region surrounding *esxN*. Products without *esxN* will be 300 bp smaller than products amplifying this region. +/- RT samples were included.
Experiments shown here were performed one time.

Assessing the role of EsxN in counteracting stresses encountered during infection

Previous work in our lab has revealed the Δ*pstA1* single mutant is sensitive to many *in vitro* stresses, including cell wall stress, ROS and RNS [149]. During infection, the host immune response employs these same stresses as defense mechanisms against *Mtb* [46]. To determine if the enhanced stress sensitivity of the Δ*pstA1* mutant is due to aberrant ESX-5 hyper-secretion, and to evaluate if EsxN is required to resist mechanisms of host immunity in WT *Mtb*, we examined survival of *esxN* deletion mutants in a variety of *in vitro* stress conditions, designed to mimic environments encountered by the bacterium *in vivo*. When bacteria were subjected to cell wall stress, through the addition of the membrane disrupting SDS [220], the Δ*pstA1* mutant was more sensitive than WT bacteria, as previously reported [149]. Deletion of *esxN* did not reverse sensitivity of the Δ*pstA1* mutant to cell walls stress (**Figure 4.2A**). Further, Δ*esxN* bacteria were as resistant as WT to this stress condition. Taken together, these data demonstrate EsxN does not play a role in resistance to SDS-mediated cell wall stress, and does not cause the sensitivity observed in the Δ*pstA1* mutant.

We next examined bacterial survival during exposure to the reactive oxygen species hydrogen peroxide (H_2O_2). Consistent with previous data, the Δ*pstA1* mutant was more sensitive than WT bacteria to ROS stress [149]. The sensitivity was not due to *esxN,* because the Δ*pstA1*Δ*esxN* mutant was equally sensitive (**Figure 4.2B**). Deletion of *esxN* did not alter WT resistance to ROS stress. These data show that EsxN is not involved in resistance to ROS stress, and also does not contribute to sensitivity of the Δ*pstA1* mutant to this condition.

We also investigated how bacterial survival was affected by acid pH. The Δ*pstA1* mutant was more sensitive to acid pH than WT bacteria (**Figure 4.2C**). Deleting *esxN* in the Δ*pstA1* mutant did not increase survival under this condition. Deletion of *esxN* in WT bacteria did result in a small decline in survival, but as this experiment was conducted only one time, it is unknown whether this difference is reproducible or significant.

Finally, we examined how deletion of *esxN* affects bacterial survival when exposed to reactive nitrogen species. Again, consistent with published data, the Δ*pstA1* mutant exhibited reduced survival compared to WT when treated with the RNS donor diethylenetriamine-NONOate (DETA-NO) [149]. Deletion of *esxN* did not impact the

sensitivity of the Δ*pstA1* mutant to RNS stress. However, we observed a survival defect in the Δ*esxN* mutant under RNS stress (**Figure 4.2D**). The Δ*esxN* mutant was almost as sensitive as the Δ*pstA1* mutant. Strikingly, complementation of *esxN in trans* in the Δ*esxN* mutant restored survival to WT levels, indicating the sensitivity was due to deletion of *esxN.* Taken together, these data suggest that *esxN* may be playing a role in resistance to RNS stress in WT *Mtb*.

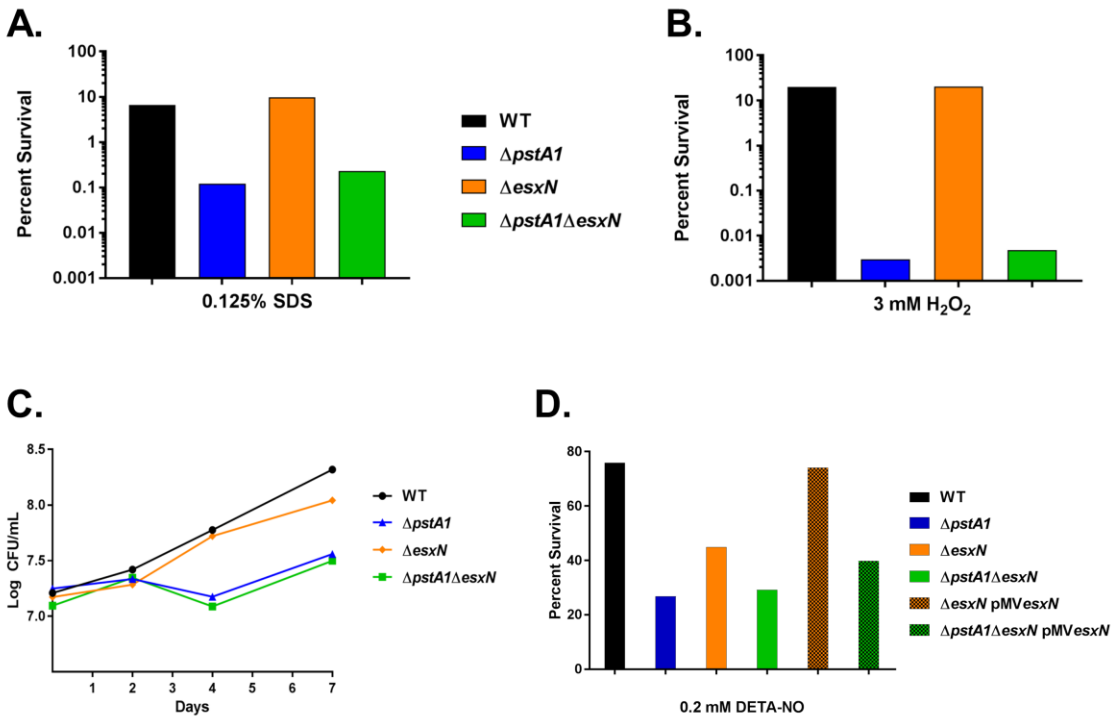

Figure 4.2 Assessing the role of EsxN in resisting *in vitro* stresses. WT *Mtb* Erdman (WT), Δ*pstA1,* Δ*esxN,* and Δ*pstA1*Δ*esxN* were grown to mid-exponential phase in 7H9 complete medium and then subjected to the indicated stresses. Bacterial CFU were enumerated by plating serial diluted cultures on 7H10 agar and incubating for 3-4 weeks at 37°C. Percent survival was calculated as (CFU after stress)/(CFU before stress) x 100.
A) Cell wall stress. Sodium dodecyl sulfate (SDS) was added to the cultures at a final concentration of 0.125%. CFU were enumerated at 0 and 24 hrs.
B) Reactive oxygen species stress. H_2O_2 was added to the cultures at a final concentration of 3mM. CFU were enumerated at 0 and 24 hrs.
C) Acid pH stress. Cultures were grown in acidified 7H9 medium (pH 5.5) and grown for 7 days, shaking at 37°C. CFU were enumerated at day 0, 2, 4 and 7.
D) Reactive nitrogen species stress. The reactive oxide donor DETA-NO was added to the cultures at a final concentration of 0.2 mM at time 0, 24 and 48 hours. CFU were enumerated at 0 and 72 hours. The complemented strains Δ*esxN* pMV*esxN* and Δ*pstA1*Δ*esxN* pMV*esxN* were also evaluated in this experiment.
Experiments shown were performed one time.

Deletion of esxN does not affect production or secretion of other ESX-5 proteins

Since some ESX-1 substrates exhibit co-dependent secretion, we examined if deletion of *esxN* altered secretion of PPE41 through ESX-5 [108]. Consistent with published results, EsxN and PPE41 were hyper-secreted by the Δ*pstA1* mutant compared to WT (**Figure 4.3**) [158]. The ESX-5 protein chaperone $EspG_5$ and secretion machinery component $EccB_5$ in the Δ*pstA1* mutant were overproduced compared to WT, as previously demonstrated. Deletion of *esxN* in WT *Mtb* had no effect on ESX-5 protein production or secretion, since EsxN, PPE41, $EspG_5$ and $EccB_5$ protein levels were unchanged in the Δ*esxN* mutant compared to WT (**Figure 4.3**). Similarly, all ESX-5 proteins examined were detected at similar levels in the Δ*pstA1*Δ*esxN* double mutant compared to the Δ*pstA1* mutant. Notably, deletion of *esxN* had not effect on PPE41 secretion, implying that EsxN and PPE41 are not mutually dependent for secretion. Anti-ModD and anti-GroEL2 loading controls demonstrate equivalent amounts of culture filtrate and cell lysate fraction proteins were added, respectively. Taken together, these data demonstrate that ESX-5 secreted proteins are not mutually dependent on each other for export, and deletion of *esxN* does not affect production of ESX-5 cellular components.

We attempted to complement the Δ*pstA1*Δ*esxN* deletion by providing a WT copy of *esxN in trans* in a single copy integrating plasmid under the HSP60 promoter, pMV*esxN*. However, complementation failed to restore secretion of EsxN to the level observed in the Δ*pstA1* mutant. We have previously shown that *esx-5* transcripts are constitutively overexpressed in Δ*pstA1* [158]. This complementation vector appeared incapable of adequately complementing EsxN levels to those seen in the Δ*pstA1* mutant, suggesting that the HSP60 promoter is not as active as the *esx-5* RegX3 promoter.

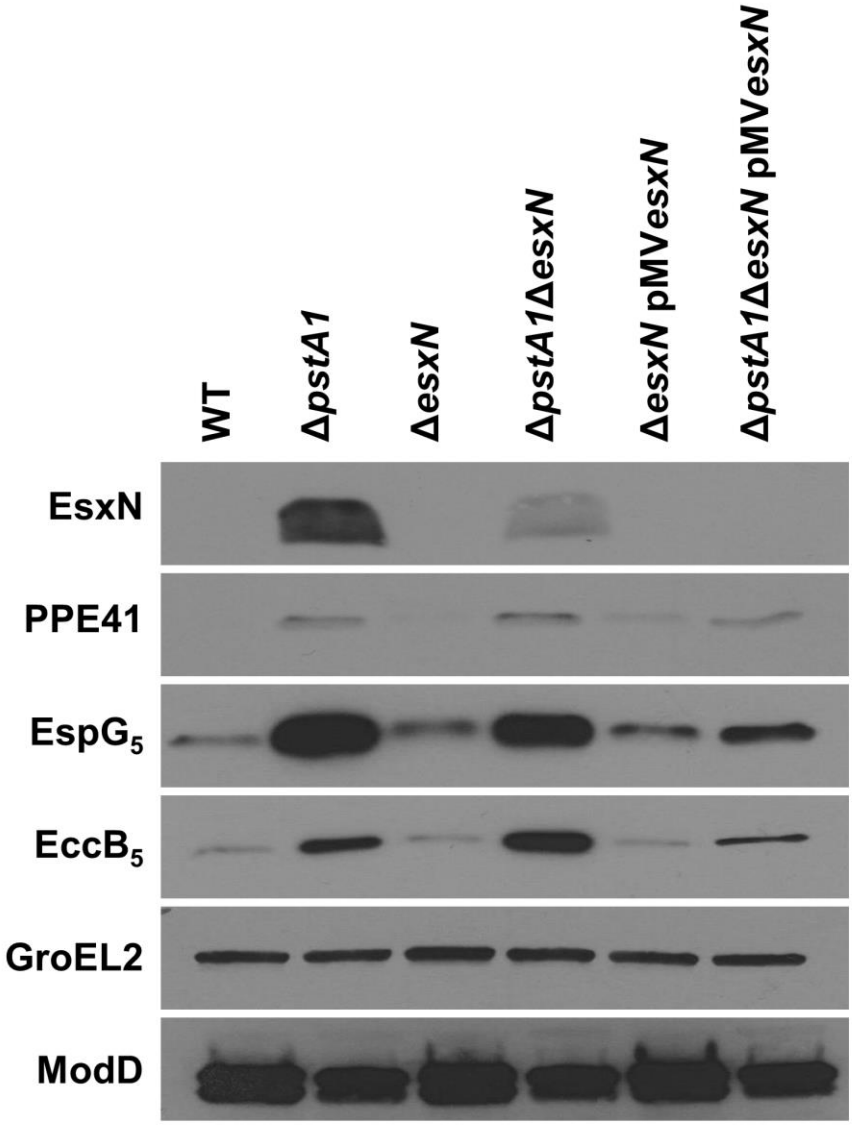

Figure 4.3 Deletion of *esxN* does not impact production of ESX-5 components or secretion of ESX-5 proteins. Wild-type *Mtb* Erdman (WT), Δ*pstA1,* Δ*esxN,* Δ*pstA1*Δ*esxN,* Δ*esxN* pMV*esxN* and Δ*pstA1*Δ*esxN* pMV*esxN* were cultured in Sauton's complete medium without Tween-80 for 5 days. 10 µg of cell lysate ($EspG_5$, $EccB_5$) and 5 µg of culture filtrate (EsxN, PPE41) proteins were subjected to SDS-PAGE and Western blot analysis. Anti-GroEL2 was used as a loading control for the CL fraction and anti-ModD was used as a loading control for the CF fraction. Results are representative of two independent experiments.

Tet-inducible complementation of the esxN deletion in ΔesxN and ΔpstA1ΔesxN Mtb

Since our initial strategy to complement the *esxN* deletion strains could not be accomplished using a single-copy integrating vector, we generated alternative complemented strains. We utilized a Tet-inducible construct (pTIC10a) to enable inducible expression of *esxN* to levels comparable to those in the Δ*pstA1* mutant. To verify the Tet-inducible *esxN* strains, we cultured Δ*esxN* pTIC*esxN* and Δ*pstA1*Δ*esxN* pTIC*esxN* in a range of anhydrous tetracycline (aTc) concentrations to determine the amount of aTc required to restore a level of *esxN* transcription similar to the Δ*pstA1* mutant (**Figure 4.4**). However, our initial attempt to quantify *esxN* in the complemented strains was frustrated by the fact that our primers also detect the other *esxN* paralogs encoded outside of the *esx-5* locus (**Figure 4.4A**). There was no discernable pattern to *esxN* expression, and transcript abundance did not increase along with aTc concentration. Since the *esxN* paralogs share a high sequence identity, it is extremely difficult to design *esxN* specific primers. Instead, we designed a set of primers wherein one primer annealed to the pTIC10a vector, and the other within *esxN*, and in this way we hoped to detect *esxN* transcripts generated from the pTIC*esxN* Tet-inducible vector. Indeed, with the vector-specific primer set we observed an increase in *esxN* transcript abundance as aTc concentration increased (**Figure 4.4B**). All further experiments for *esxN* complementation utilized the Δ*esxN* pTIC*esxN* and Δ*pstA1*Δ*esxN* pTIC*esxN* strains.

We also examined EsxN secretion in the Δ*pstA1*Δ*esxN* pTIC*esxN* complemented strain, to determine whether secretion reaches the Δ*pstA1* mutant levels using this construct. We cultured the Δ*pstA1*Δ*esxN* pTIC*esxN* bacteria without aTc, as an uninduced control to assess leakiness of *esxN* expression, and in either 50 or 400 ng/ml of aTc, representing the lowest and highest range typically used for induction. EsxN secretion in the Δ*pstA1*Δ*esxN* pTIC*esxN* complemented strain was restored to the Δ*pstA1* mutant levels with the addition of 50 ng/ml aTc (**Figure 4.5**). Secretion of EsxN was greatly reduced when Δ*pstA1*Δ*esxN* pTIC*esxN* was uninduced, indicating this construct has tight control of *esxN* expression. There was less EsxN detected when the complemented strain was induced with 400 ng/ml aTc, indicating this level of aTc has a negative impact on the bacteria. Indeed, there is a maximal aTc concentration to induce

gene expression without impacting bacterial growth, and previous studies have shown that 250-500 ng/ml aTc resulted in decreased protein production [221]. Based on these results, it is possible Δ*pstA1* bacteria have a lower aTc threshold than WT *Mtb*, though we have not tested this directly.

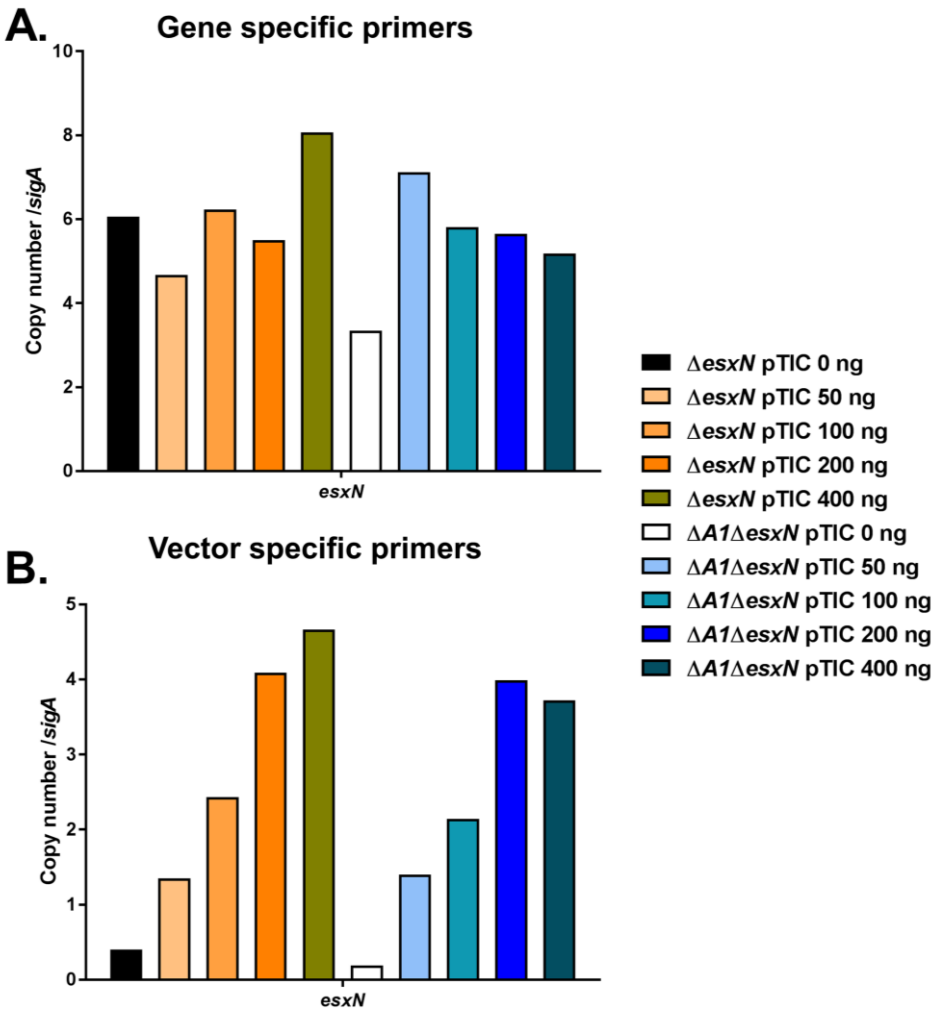

Figure 4.4 Validating Tet-inducible *esxN* complementation in *ΔesxN* and *ΔpstA1ΔesxN Mtb*. Quantitative RT-PCR analysis of *esxN* transcription. Δ*esxN* pTIC*esxN* and Δ*pstA1*Δ*esxN* pTIC*esxN* were cultured in a range of anhydrous tetracycline in 7H9 medium to mid-exponential phase and RNA was extracted. Abundance of *esxN* transcripts were determined using either A) gene-specific or B) vector-specific primers by qRT-PCR and normalized to the housekeeping gene, *sigA*. Data shown are from experiments performed one time.

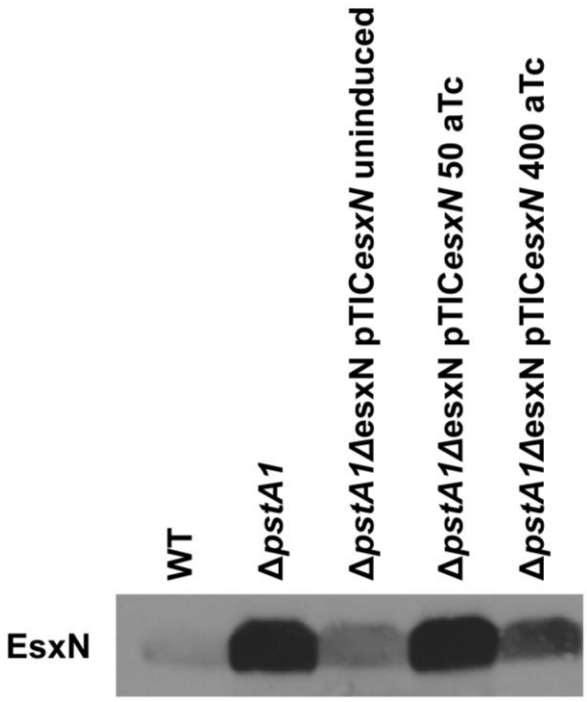

Figure 4.5 EsxN secretion in Δ*pstA1*Δ*esxN* can be fully restored to Δ*pstA1* levels by Δ*pstA1*Δ*esxN* pTIC*esxN* complementation . Wild-type *Mtb* Erdman (WT), Δ*pstA1,* Δ*pstA1*Δ*esxN,* and Δ*pstA1*Δ*esxN* pTIC*esxN* were cultured in Sauton's complete medium without Tween-80 for 5 days. Δ*pstA1*Δ*esxN* pTIC*esxN* was cultured in either 0, 50 or 400 ng/ml aTc. 5 µg of culture filtrate proteins were subjected to SDS-PAGE and Western blot analysis. Results are representative of two independent experiments.

Assessing the contribution of EsxN in Mtb pathogenesis

As previously reported, the Δ*pstA1* mutant has replication and virulence defects in C57BL/6, IrgM1$^{-/-}$ and Nos2$^{-/-}$ mice [149]. Studies described in chapter 3 of this thesis demonstrated that attenuation of Δ*pstA1* in C57BL/6 and IrgM1$^{-/-}$ mice was due, at least in part, to constitutive expression of *esx-5* genes, mediated by RegX3. Work outlined in chapter 2 has shown that EsxN, a highly antigenic ESX-5 secreted protein, is hyper-secreted in the Δ*pstA1* mutant, and chapter 3 demonstrated hyper-secretion is reversed when RegX3 is prevented from activating *esx-5* genes. To investigate if deletion of *esxN* could reverse attenuation of Δ*pstA1* mutant *in vivo*, we infected C57BL/6, IrgM1$^{-/-}$ and Nos2$^{-/-}$ mice via the aerosol route with ~100 CFU of WT, Δ*pstA1*, or Δ*pstA1*Δ*esxN Mtb* (**Figure 4.6**). Consistent with previous results, the Δ*pstA1* mutant was attenuated in C57BL/6, IrgM1$^{-/-}$ and Nos2$^{-/-}$ mice compared to WT *Mtb* [149]. Deletion of *esxN* did not suppress the virulence defect of Δ*pstA1* bacteria in C57BL/6 or IrgM1$^{-/-}$ mice, as the bacterial burdens were almost identical in mice infected with either the Δ*pstA1* mutant or the Δ*pstA1*Δ*esxN* mutant (**Figure 4.6 A, C**). In contrast, there was a significant increase in lung CFU from weeks 2 to 6 in Nos2$^{-/-}$ mice infected with the Δ*pstA1*Δ*esxN* mutant, while CFU of the Δ*pstA1* mutant remained stable during this time (**Figure 4.6B**). Further, there was a significant increase in Δ*pstA1*Δ*esxN* mutant CFU at weeks 4 and 6 compared to the Δ*pstA1* mutant. Therefore, deletion of *esxN* partially reversed the virulence defect of the Δ*pstA1* single mutant specifically in Nos2$^{-/-}$ mice. These data suggest that *esxN* overexpression sensitizes Δ*pstA1* bacteria to host responses in Nos2$^{-/-}$, but not C57BL/6 or IrgM1$^{-/-}$ mice.

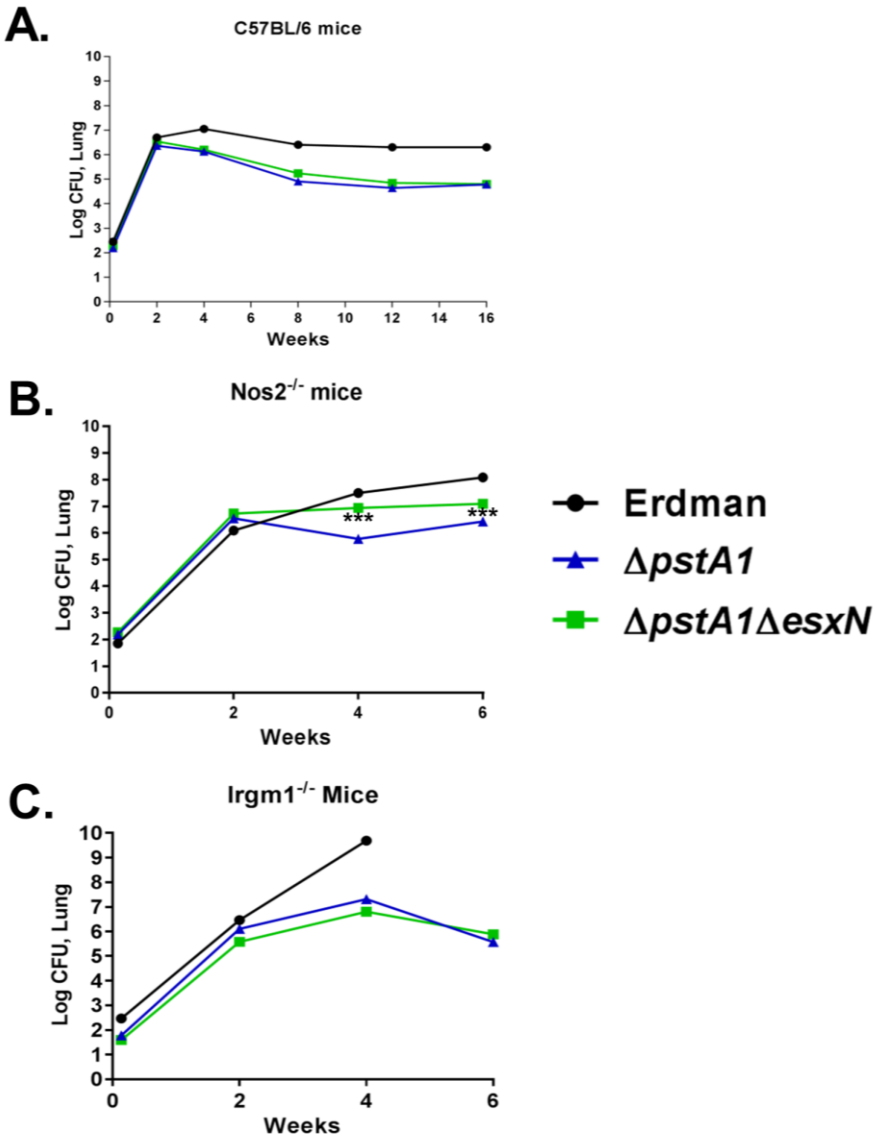

Figure 4.6 Deletion of *esxN* partially reverses Δ*pstA1* attenuation in Nos2$^{-/-}$, but not C57BL/6 or IrgM1$^{-/-}$ mice.

A) C57BL/6, B) Nos2$^{-/-}$, and C) IrgM1$^{-/-}$ mice were infected via the aerosol route with ~100 CFU of wild-type Erdman, Δ*pstA1* or Δ*pstA1*Δ*esxN Mtb.* Groups of infected mice (*n*=4) were sacrificed at the indicated time points. Bacterial CFU were enumerated by plating lung homogenates on 7H10 agar and incubating at 37°C for ~3-4 weeks. Symbols represent means; error bars reflect standard error of the mean. Asterisks indicate statistically significant differences between Δ*pstA1* and Δ*pstA1*Δ*esxN* mutants (*** *P* <0.0001).

Discussion

Previous work from our lab has revealed that the P_i -responsive Pst/SenX3-RegX3 system directly regulates the virulence-associated ESX-5 secretion system at the transcriptional level in response to P_i limitation. We have also demonstrated that RegX3-dependent constitutive overexpression of *esx-5* transcripts causes attenuation of a Δ*pstA1* mutant *in vivo.* Many ESX-5 secreted proteins are known to be highly immunogenic, and we hypothesized that aberrant hyper-secretion of these substrates may contribute to the sensitization of Δ*pstA1* bacteria to host responses observed *in vivo*. To address this possibility, we generated *esxN* deletion mutants in both WT and Δ*pstA1 Mtb.*

We confirmed deletion of *esxN*, although we still detected secretion of EsxN paralogs. Consistent with published results, deletion of *esxN* did not impact secretion of PPE41 [105]. Further, *esxN* deletion did not impact any ESX-5 component examined.

We monitored bacterial replication and persistence in C57BL/6, IrgM1$^{-/-}$ and Nos2$^{-/-}$ mice to assess whether EsxN hyper-secretion contributes to attenuation of Δ*pstA1* bacteria *in vivo.* We found that deletion of *esxN* did not reverse the virulence defect of the Δ*pstA1* mutant in C57BL/6 or IrgM1$^{-/-}$ mice, and the Δ*pstA1* Δ*esxN* mutant was just as attenuated as the Δ*pstA1* single mutant. Intriguingly, the Δ*pstA1*Δ*esxN* mutant partially restored virulence in the Δ*pstA1* mutant in Nos2$^{-/-}$ mice. These data show that EsxN hyper-secretion sensitizes Δ*pstA1* bacteria to some host response in Nos2$^{-/-}$ but not C57BL/6 or IrgM1$^{-/-}$ mice.

We also examined how deletion of *esxN* impacted bacterial survival in a variety of *in vitro* stress conditions aimed to recapitulate relevant stresses encountered by the bacterium during infection. We found that deletion of *esxN* did not impact survival of WT or Δ*pstA1* bacteria when exposed to acid pH, reactive oxygen species or cell wall stress. These findings are in line with published data attributing sensitivity of the Δ*pstA1* mutant to cell wall and ROS stress to overexpression of *pe19* [193]. However, survival was affected when reactive nitrogen species were introduced. Bacterial survival was reduced in both the Δ*esxN* and Δ*pstA1* Δ*esxN* mutants compared to WT when cultures were exposed to RNS. The survival defect was restored in Δ*esxN* when *esxN* was complemented *in trans*. These data suggest that EsxN may play a role in resistance to

RNS in WT *Mtb*.

The work presented here suggests that attenuation of Δ*pstA1* bacteria may occur through different mechanisms in C57BL/6 and IrgM1$^{-/-}$ mice compared to Nos2$^{-/-}$ mice. Deletion of *esxN* was detrimental to survival of WT *Mtb* when exposed to RNS stress *in vitro,* suggesting a role for EsxN in resistance to RNS. Nos2$^{-/-}$ mice cannot produce RNS, and deletion of *esxN* partially reversed the virulence defect of the Δ*pstA1* mutant, an environment in which resistance to RNS is dispensable. The Δ*pstA1*Δ*esxN* mutant was attenuated in C57BL/6 and IrgM1$^{-/-}$ mice, as is the Δ*pstA1* single mutant, though it is possible virulence is affected in different ways. Deletion of *esxN* might cause attenuation in C57BL/6 and IrgM1$^{-/-}$ mice, since these mice can produce RNS. It is therefore impossible to determine if EsxN hyper-secretion contributes to attenuation of the Δ*pstA1* mutant in these mice. However, we can evaluate if EsxN is involved in mediating RNS resistance *in vivo* using the Δ*esxN* mutant. Future experiments would involve infecting C57BL/6, IrgM1$^{-/-}$ and Nos2$^{-/-}$ mice with Δ*esxN* bacteria. If this mutant exhibits a replication or virulence defect in C57BL/6 and IrgM1$^{-/-}$ mice but not in Nos2$^{-/-}$ mice, this would suggest that EsxN is required for survival of host mediated RNS stress.

Acknowledgements

We thank Dr. Anthony Baughn for kindly providing the Tet inducible pTIC10a plasmid, a derivative of pTIC6 [222]. The anti-sera against EsxN, PPE41, $EspG_5$ and $EccB_5$ were generated by Dylan White. The following reagents were obtained through BEI Resources, NIAID, NIH: monoclonal anti-*Mycobacterium tuberculosis* GroEL2 (Gene Rv0440) and polyclonal anti-*Mycobacterium tuberculosis* Mpt32 (Gene Rv1860) (antiserum, rabbit), NR-13807.

Chapter 5

Discussion and future directions

Discussion

Mycobacterium tuberculosis is a facultative intracellular bacterial pathogen and causative agent of the disease tuberculosis. *Mtb* relies on the export of a variety of virulence factors to counteract the onslaught of host immune defenses and promote replication, colonization and persistence during infection. The Type VII ESX specialized protein secretion systems are employed by the bacterium to mediate secretion of many factors critical for survival. ESX-5 is one such system, and has been implicated in virulence, modulation of host responses and nutrient acquisition [105, 223, 224]. Though it is clear that the ESX-5 system is required for *Mtb* pathogenesis, it was unknown how the system was regulated or activated during infection. The work presented in this thesis has defined the regulatory mechanism governing the ESX-5 secretion system, identified a relevant environmental signal activating the system, and examined how dysregulation of the system impacts *Mtb* pathogenesis.

In Chapter 2, our work revealed that the ESX-5 secretion system is directly regulated at the transcriptional level by the P_i -responsive Pst/SenX3-RegX3 system. Further, ESX-5 secretion is triggered by P_i limitation, a relevant nutritional signal that *Mtb* likely encounters during infection [158]. We found that P_i limitation activates transcription of a subset of *esx-5* genes in a RegX3-dependent manner, leading to increased production of ESX-5 core components and secretion of EsxN and PPE41 in WT *Mtb*. Further, we demonstrated that *esx-5* transcripts are aberrantly overexpressed in a Δ*pstA1* mutant due to constitutive activation of RegX3. We also observed constitutive overproduction of ESX-5 components and hyper-secretion of EsxN and PPE41 in the Δ*pstA1* mutant that was dependent on RegX3. During these studies, we also found that P_i limitation triggers secretion of EsxB, an ESX-1 associated virulence factor, though this occurs independently of RegX3. Though this work demonstrated that the response regulator RegX3 directly binds within the *esx-5* locus *in vitro*, the exact binding sequence remained unknown.

Previous work from our lab found that the Δ*pstA1* mutant has a replication and virulence defect in C57BL/6 and IrgM1$^{-/-}$ mice [149]. ESX-5 substrates are highly antigenic, and are actively secreted during infection [106, 107]. We proposed that Δ*pstA1* mutant attenuation may be caused, in part, to aberrant hyper-secretion of

immunogenic ESX-5 factors. However, since a Δ*regX3* strain is also attenuated for replication and virulence in C57BL/6 and IrgM1$^{-/-}$ mice, we could not utilize this strain to test this hypothesis, and a more targeted approach was required. Work described in chapter 3 built on the findings from chapter 2, and identified the RegX3 binding site sequence upstream of *pe19* within the *esx-5* locus using *in vitro* EMSAs. Subsequently, the RegX3 binding site sequence was mutated in both WT and Δ*pstA1 Mtb*. We found that disrupting RegX3 binding within the *esx-5* locus prevented induction of ESX-5 secretion during P_i limitation in WT *Mtb* and reversed constitutive hyper-secretion in the Δ*pstA1* mutant. These results conclusively demonstrate that Pst/SenX3-RegX3 directly regulates ESX-5 activity in *Mtb.* Deletion of the *esx-5* RegX3 binding site sequence in the Δ*pstA1* mutant uncoupled the global effects resulting from *regX3* deletion and assess the impact of ESX-5 independently *in vivo*. To determine whether RegX3-dependent aberrant hyper-secretion of ESX-5 substrates sensitized Δ*pstA1* bacteria to host immune responses, C57BL/6 and IrgM1$^{-/-}$ mice were infected with Δ*pstA1*ΔBS *Mtb*. Strikingly, we found that deletion of the *esx-5* RegX3 binding site sequence partially restored virulence to the attenuated Δ*pstA1* mutant in both mouse strains. We have now determined that aberrant regulation of ESX-5 mediated by RegX3 is responsible, at least in part, for attenuation of the Δ*pstA1* mutant. This illustrates that precise gene regulation is essential for bacterial survival, and shows that dysregulation of ESX-5 is detrimental to *Mtb* virulence.

Experiments outlined in Chapter 4 examined the impact of one ESX-5 secreted factor, EsxN, on *Mtb* pathogenesis. In the previous chapter, we determined that constitutive hyper-secretion of ESX-5 proteins contributes to Δ*pstA1* mutant attenuation *in vivo*. To explore whether the Δ*pstA1* mutant virulence defect could be reversed by elimination of one highly antigenic ESX-5 secreted substrate, we deleted *esxN* in both the WT and Δ*pstA1* mutant *Mtb* strains. In addition to attenuation *in vivo,* the Δ*pstA1* mutant is also sensitive to several stresses *in vitro*. We examined if deletion of *esxN* could reverse these Δ*pstA1* mutant stress sensitivities, and if WT *Mtb* requires *esxN* to resist certain stresses. We found that *esxN* is dispensable for WT survival under cell wall stress, acid pH and ROS conditions. Deletion of *esxN* does not suppress Δ*pstA1* mutant sensitivity to these stresses. However, deletion of *esxN* did sensitize WT bacteria to RNS stress *in vitro*, and resistance was restored when *esxN* was complemented *in*

trans, suggesting EsxN is involved in RNS resistance. We then explored if hyper-secretion of EsxN contributes to Δ*pstA1* attenuation in C57BL/6, IrgM1$^{-/-}$ and Nos2$^{-/-}$ mice. We found that Δ*pstA1*Δ*esxN* bacteria were as attenuated as the Δ*pstA1* mutant in C57BL/6 and IrgM1$^{-/-}$ mice. However, deletion of *esxN* did partially suppress the virulence defect of Δ*pstA1* bacteria in Nos2$^{-/-}$ mice, implying EsxN hyper-secretion sensitizes Δ*pstA1* bacteria to a host response other than RNS production by NOS2. Taken together with the results presented in chapter 3, our results suggest that other factors secreted by ESX-5 cause attenuation of the Δ*pstA1* mutant in C57BL/6 and IrgM1$^{-/-}$ mice.

Overall, the work outlined in this thesis has revealed the regulatory mechanism governing ESX-5 activity, and identified P_i limitation as a novel environmental signal that triggers secretion. This work can be summarized by a model in which the Pst/SenX3-RegX3 system activates *esx-5* transcription when the bacterium encounters a P_i limiting environment (**Figure 5.1**). The phagosomal environment is predicted to be nutrient poor, and several studies have determined that genes involved in P_i uptake and scavenging are essential for *Mtb* survival in macrophages, suggesting P_i is limiting in this environment [88, 159, 160]. P_i limitation acts as a signal to the bacterium to activate genes to acquire more P_i, but perhaps this environmental cue also triggers secretion of factors aimed at modulating host responses within that niche, such as counteracting RNS stress. Overall, this work has broadened our understanding of ESX-5 biology in general, and specifically in the context of *Mtb* pathogenesis.

A more thorough understanding of how ESX-5 is regulated provides clues as to the function of this system, which is still unclear. It has been suggested that ESX-5 factors are involved in nutrient acquisition, since increasing permeability of the outer membrane overcomes the essentiality of ESX-5 secretion system components [223]. This observation is intriguing since we have shown that a nutritional signal, P_i limitation, induces ESX-5 system activity. Indeed, there is precedence for ESX system involvement in nutrient uptake, since ESX-3 is required for iron acquisition [89]. While some ESX-5 secreted factors may be involved in nutrient uptake, evidence presented in this work suggests additional ESX-5 factors play a role in *Mtb* virulence. Looking again to ESX-3, this system has a role in virulence independent from its role in iron uptake [93]. ESX-5 may similarly mediate two independent functions. We propose that P_i limitation is a

critical environmental cue, signaling that secretion of ESX-5 effectors is required to counteract host responses within that niche. As previously described, this may occur while in the macrophage phagosomal compartment, which is predicted to be a P_i-limiting environment. An as yet unidentified ESX-5 factor has been shown to promote necrotic cell death in macrophages, a function that seems unlikely to be involved in nutrient uptake [102]. The work presented in this thesis reveals a potential role for EsxN in resistance to RNS, a host defense mechanism the bacterium surely encounters within the macrophage phagosome. Our work supports the potential for the *Mtb* ESX-5 system to have dual functions in both nutrient acquisition and virulence.

There is a well-known link between bacterial virulence and P_i limitation. Our discovery that ESX-5 activity is regulated by the Pst/SenX3-RegX3 system in response to P_i availability is consistent with a known regulatory network connecting P_i limitation and virulence in other bacteria. Disruption of the Pst system leads to attenuation of virulence in several other bacterial pathogens, likely a result of aberrant gene regulation as opposed to nutrient limitation, and is also consistent with our findings in *Mtb* [225]. For example, inactivation of the Pst system in avian pathogenic *E. coli* leads to increased sensitivity to bactericidal effects of host serum and antimicrobial peptides [226]. Inactivation of the Pst system causes constitutive activation of the PhoB response regulator, the homolog to RegX3 in other bacterial species. Many studies have shown that disruption of regulation mediated by the PhoB response regulator leads to attenuation, though a deficiency in P_i uptake could also be a contributing factor [225]. P_i limitation directly triggers activation of virulence factors in some pathogens. In *Pseudomonas aeruginosa*, P_i starvation induced secretion of the virulence factors pyocyanin and non-hemolytic C phospholipases [227]. We have shown that P_i limitation activates secretion of ESX-5 effectors, some of which are likely involved in *Mtb* virulence, demonstrating our findings fit into the larger context of virulence in response to nutrient limitation found in other bacterial pathogens.

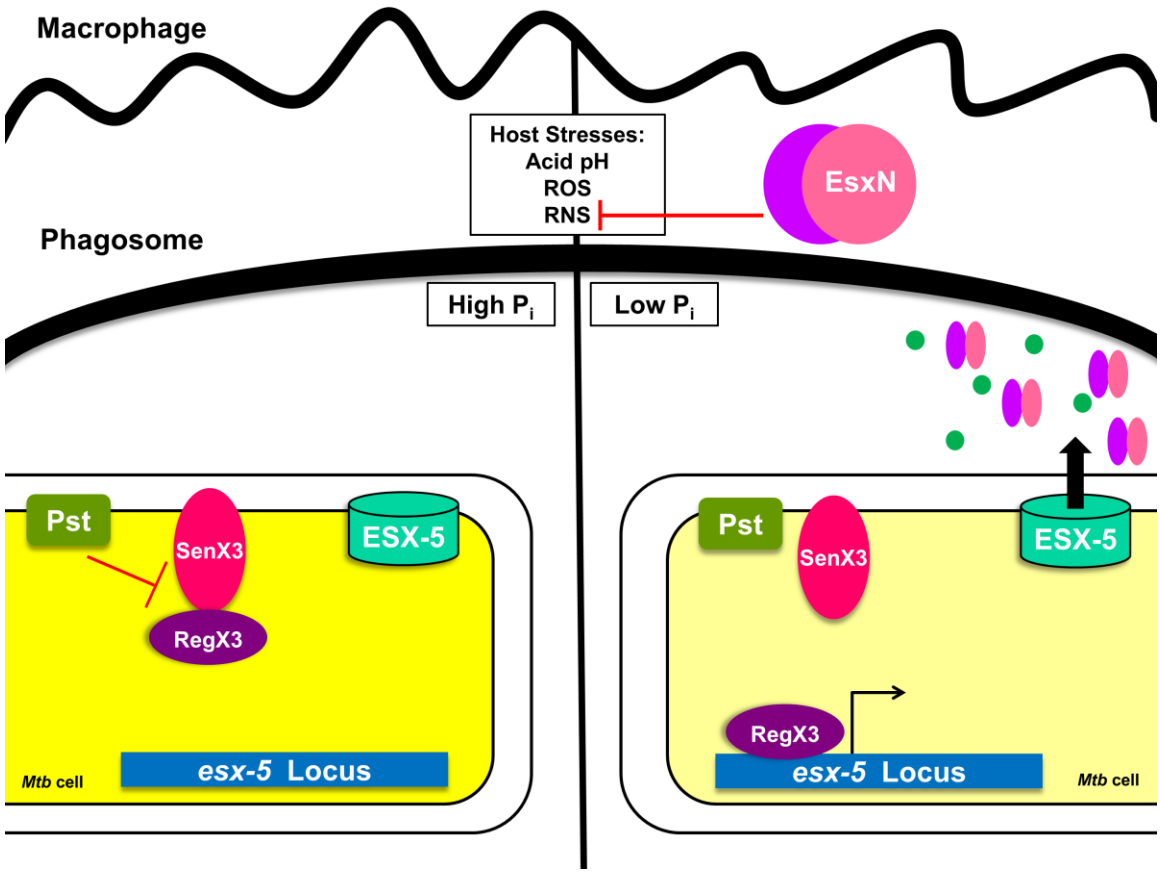

Figure 5.1 Model: P_i limitation triggers ESX-5 secretion to promote *Mtb* survival *in vivo*. The bacterium likely encounters P_i limitation within the phagosomal compartment of macrophages. Initially, *Mtb* may possess sufficient P_i to support cellular processes. However, RegX3 activates *esx-5* transcription when P_i is limiting. ESX-5 secreted effectors promote bacterial survival by mediating nutrient acquisition and modulating the host response. EsxN may be involved with bacterial survival during reactive nitrogen species stress.

Future Directions

Though our work has revealed that ESX-5 is directly regulated by the Pst/SenX3-RegX3 system in response to P_i limitation, there is still much we do not yet know about ESX-5. Work described in this thesis has demonstrated that RegX3 mediated dysregulation of ESX-5 causes attenuation of the Δ*pstA1* mutant in C57BL/6 and IrgM1$^{-/-}$ mice. It will be interesting to infect these strains of mice with an *esx-5* RegX3 binding site mutant in WT *Mtb*. This experiment would allow us to assess how regulation of ESX-5 secretion in response to P_i limitation contributes to *Mtb* pathogenesis, uncoupled from global effects on transcription in a Δ*regX3* mutant.

Another way to uncouple ESX-5 activity from regulation mediated by RegX3 would be to create an *esx-5* inducible *Mtb* strain ($WT_{ESX5\ Tet}$) (outlined in **Appendix IV**). The ability to induce *esx-5* expression using aTc will allow us to examine the ESX-5 system in a number of ways. The $WT_{ESX5\ Tet}$ mutant can be used to infect macrophages, and assess how induction or repression of ESX-5 activity impacts bacterial survival and replication in this environment. With this inducible strain, we can begin to tease apart if ESX-5 is required at different stages of macrophage infection. We have shown that constitutive hyper-secretion of ESX-5 substrates causes attenuation *in vivo*, but given that certain ESX-5 components are essential, it seems likely that this system is important at some time during infection. *Mtb* infections have both an acute and chronic phase. Using the $WT_{ESX5\ Tet}$ mutant, we can infect mice and induce expression of *esx-5* transcripts at certain stages of infection and systematically determine when ESX-5 is required for *Mtb* survival and virulence in an immune competent host.

Although we know that ESX-5 secretion is induced by P_i limitation, it is still unknown when ESX-5 is activated during infection. Utilizing a fluorescent reporter strain as a read-out of *esx-5* transcriptional activation would allow us to determine when the system is activated during infection. Using a reporter strain, macrophages could be infected with *Mtb*, and visualized with a fluorescent microscope to identify when the genes have been activated. It would also be interesting to infect P_i-starved macrophages with this reporter strain, since we would expect this condition to induce *esx-5* expression.

We have shown that Δ*pstA1* mutant attenuation is due to aberrant constitutive *esx-5* gene expression mediated by RegX3. Given that ESX-5 secreted substrates are highly antigenic, and we suspect that infection with Δ*pstA1* bacteria may produce a hyper-

inflammatory response compared to WT *Mtb*. Performing studies aimed at assessing the immunological profile of the Δ*pstA1* mutant compared to WT would start to address this possibility. Examining infected lung tissue through simple hematoxylin and eosin staining (H & E) would allow us to assess if Δ*pstA1* bacteria induce more tissue damage compared to WT *Mtb*. Further, it would be interesting to perform ELISA experiments, looking at pro-inflammatory cytokine profiles to determine if the *Mtb* Δ*pstA1* mutant is hyper-inflammatory in both macrophage and murine models of infection. In particular, it would be informative to evaluate IFN-γ, TNF-α, IL-12, and IL-17, which are pro-inflammatory cytokines associated with *Mtb* infections [228, 229]. Additionally, ESX-5 is required for induction of host IL-1β in *M. marinum*, which leads to inflammasome activation [102]. It would be interesting to quantify this cytokine in the Δ*pstA1* mutant in *Mtb*. Further, examining host cell death mediated by the Δ*pstA1* mutant, either apoptosis or necrosis, will also provide clues as to how this mutant interacts with host cells. Since secretion by ESX-5 is involved in necrotic cell death in *M. marinum [102]*, we predict that the Δ*pstA1* mutant will also cause necrotic host cell death. It is possible that due to constitutive hyper-secretion of ESX-5 substrates, the Δ*pstA1* mutant may even induce more host cell death than WT bacteria.

Further work will need to be accomplished to tease apart whether EsxN is involved in resistance to reactive nitrogen species. In the experiments detailed here, mice were only infected with a Δ*pstA1*Δ*esxN* deletion mutant, and we found that overexpression of *esxN* was actually causing attenuation of Δ*pstA1* bacteria in Nos2$^{-/-}$ mice. But *esxN* deletion did enhance susceptibility to RNS stress in WT *Mtb.* To fully understand the importance of EsxN in *Mtb* pathogenesis, the Δ*esxN* mutant will need to be evaluated *in vivo* as well. If the Δ*esxN* mutant exhibits a virulence defect in C57BL/6 or IrgM1$^{-/-}$ mice, but not Nos2$^{-/-}$ mice, this would suggest that *esxN* is involved in resistance to RNS stress. *Mtb* are exposed to RNS inside macrophages, and using our *esxN* deletion strains during macrophage infection is another way to determine whether this factor is required to resist RNS stress. We could either utilize a strain of macrophages incapable of producing RNS, or add an RNS production inhibitor to RAW264.7 macrophages to investigate the possible role of EsxN plays during RNS stress. Comparing survival of *esxN* deletion mutants in normal macrophages and those unable to produce RNS would directly test whether *esxN* is required for *Mtb* RNS resistance. However, it remains

possible that EsxN is involved in some other form of host modulation, or could be dispensable for *Mtb* virulence altogether. These experiments have the potential to define a novel function for a known ESX-5 secreted effector.

Our work has demonstrated that hyper-secretion of EsxN sensitizes the Δ*pstA1* bacteria to immune responses in Nos2$^{-/-}$ mice, and suggests that other ESX-5 substrates may be sensitizing this mutant in C57BL/6 and IrgM1$^{-/-}$ mice. As previously described, there are 4 EsxN paralogs outside the *esx-5* locus, and these substrates are still secreted through ESX-5 in the *esxN* deletion strains. It is possible that these other EsxN paralogs are contributing to the sensitization of the Δ*pstA1* mutant *in vivo*. To address this possibility, we could delete the other *esxN*-like genes in both WT and Δ*pstA1* mutant *Mtb* strains and assess how this impacts *Mtb* replication and virulence. If the EsxN paralogs sensitize the Δ*pstA1* mutant to host responses, we would expect that deleting all *esxN*-like genes would reverse the virulence defect observed in the Δ*pstA1* single mutant. It is possible some other ESX-5 factor is involved in the attenuation of Δ*pstA1* bacteria. PPE41 is another factor that is hyper-secreted by ESX-5 in the Δ*pstA1* mutant. We have not assessed the impact of this factor in *Mtb* virulence, and it remains possible that PPE41 is contributing to Δ*pstA1* mutant attenuation *in vivo*. Deleting *ppe41* in the Δ*pstA1* mutant would allow us to determine if this factor is involved in sensitizing the Δ*pstA1* mutant in C57BL/6 and IrgM1$^{-/-}$ mice. Identifying ESX-5 factors that cause attenuation in the Δ*pstA1* mutant will provide further insight as to the specific role ESX-5 plays in *Mtb* virulence.

It is also quite possible that an as yet unidentified ESX-5 secreted factor or factors may be responsible for the attenuation of the Δ*pstA1* mutant. Identification of additional ESX-5 secreted substrates would not only aid in the determination of whether they play a role in sensitizing Δ*pstA1* bacteria to host responses, but would also expand our knowledge of ESX-5 biology as a whole. We have conducted a preliminary mass spectrometry experiment aimed to identify ESX-5 secreted substrates using the Δ*pstA1* mutant (**Appendix I**). However, this study lead to the identification of almost 300 proteins, and it is uncertain whether secretion of these proteins are directly linked to ESX-5 without further validation, since hyper-secretion of many non-ESX-5 proteins occurs in this mutant. Additional mass spectrometry experiments would greatly aid in identifying novel ESX-5 substrates, but will require a more targeted approach. Utilizing

Mtb strain(s) that are deficient in secretion of ESX-5 substrates will be required. Any protein missing in an ESX-5 secretion deficient strain compared to WT *Mtb* can in high confidence be assigned to the ESX-5 system. Using iTRAQ mass spectrometry, secreted proteins from WT *Mtb* and ESX-5 secretion deficient strains can be tagged, analyzed and compared, and this approach would allow us to determine which secreted proteins are associated with ESX-5 specifically [230]. iTRAQ has been utilized in the related bacterium *M. bovis*, suggesting this approach would be possible in *Mtb* [231]. Identification of novel ESX-5 secreted substrates would provide clues as to the biological function of this system.

Methods

Bacterial strains and culture conditions

M. tuberculosis Erdman and the derivative Δ*pstA1*, Δ*regX3*, Δ*pstA1*Δ*regX3,* Δ*regX3* pND*regX3* and Δ*pstA1*Δ*regX3* pND*regX3* mutant strains were constructed as previously described [180]. Bacterial cultures were grown at 37°C with aeration in Middlebrook 7H9 liquid medium (Difco) supplemented with albumin-dextrose-saline (ADS), 0.5% glycerol and 0.1% Tween-80, unless otherwise noted. Sauton's medium (3.67 mM KH_2PO_4, 2 mM $MgSO_4$-$7H_2O$, 9.5 mM citric acid, 0.19 mM ammonium iron (III) citrate, 26.64 mM L-asparagine, 6% glycerol, 0.01% $ZnSO_4$, pH 7.4) was used to grow cultures for protein isolation. P_i-free Sauton's medium contains all components of Sauton's complete medium except KH_2PO_4, and is buffered using 50 mM MOPS, pH 7.4. P_i-free 7H9 medium, used to culture bacteria used for RNA isolation during P_i-starvation, was prepared as previously described [149]. Frozen stocks were prepared by growing liquid cultures to mid-exponential phase (OD_{600} 0.8-1.0) in complete 7H9 medium, then adding glycerol to 15% final concentration, and storing 1 ml aliquots at -80°C.

Cloning of RegX3 binding site mutants

Constructs for introducing mutations to the RegX3 binding site were generated in the pJG1100 allelic exchange vector, which contains the *aph* (kanamycin resistance), *hyg* (hygromycin resistance), and *sacB* (sucrose sensitivity) markers. Genome regions ~800 bp 5' and 3' of the RegX3 binding site were PCR amplified from the *Mtb* Erdman genome using the primers in **Table S1**. Forward primers to amplify the 5' region were designed with a PacI restriction site, and the reverse primers to amplify the 3' region were designed with a AscI restriction site. Overlapping PCR was used to amplify the 5' upstream region and 3' downstream region so that there was no restriction site introduced around the binding site. Resulting PCR products were cloned in pCR2.1 (Invitrogen) and sequenced. The entire construct was removed from pCR2.1 by restriction with PacI/AscI, gel purified, and ligated to similarly digested pJG1100.

Cloning of deletion and complement constructs

Constructs for deletion of *esxN* and *espG$_5$* were generated in the pJG1100 allelic exchange vector, which contains the *aph* (kanamycin resistance), *hyg* (hygromycin resistance), and *sacB* (sucrose sensitivity) markers. Genome regions ~800 bp 5' and 3' of the RegX3 binding site were PCR amplified from the *Mtb* Erdman genome using the primers in **Table S1**. Reverse primers to amplify 5' regions and forward primers to amplify 3' regions were designed with an AvrII restriction site in-frame with the start and stop codons. Resulting PCR products were cloned in pCR2.1 (Invitrogen) and sequenced. The 5' and 3' regions were removed from pCR2.1 by restriction with either PacI/AvrII and AvrII/AscI (*esxN*) or PacI/AvrII and AvrII/XhoI (*espG$_5$*). After gel purification, each segment was ligated with pJG1100 after digestion with PacI/AscI or PacI/XhoI, respectively, to generate in-frame deletion constructs.

For complementation of Δ*esxN* and Δ*pstA1*Δ*esxN* deletions, the full length *esxN* gene PCR amplified from the *Mtb* Erdman genome using primers listed in **Table S1**. The resulting PCR product was cloned in pCR2.1 (Invitrogen) and sequenced. For complementation using pMV361, the insert was removed from pCR2.1 by digestion with EcoRI/HindIII and cloned in similarly digested pMV361 to generate pMV*esxN.* For complementation using the Tet inducible pTIC10a vector, the insert was removed from pCR2.1 by digestion with HindIII/EcoRI and cloned in similarly digested pTIC10a to generate pTIC*esxN.*

Mtb strain construction

Mtb strains harboring in-frame unmarked deletions or RegX3 binding site sequence mutations were generated by a two-step homologous recombination allelic exchange method, previously described [193]. Integration of the constructs at the correct chromosomal locus was confirmed by colony PCR on heat-inactivated cell lysates using the primer pairs for detection of the 5' and 3' homologous region integration (**Table S2 & S3**). Clones in which the plasmid was integrated correctly were diluted and plated on 7H10 containing 2% sucrose for counterselection. Sucrose resistant clones were

screened again using colony PCR on heat-inactivated cell lysates using primers for the detection of the deletion or mutation (**Table S2 & S3**). Additionally, RegX3 binding site mutation clones were verified by sequencing, using primers amplifying the region (**Table S2 & S3**).

For complementation of the Δ*esxN* mutant, the Δ*esxN* and Δ*pstA1* Δ*esxN* strains were electroporated with either pMV*esxN* or pTIC*esxN.* Transformants were selected on 7H10 with kanamycin (15 µg/ml). The presence of complementation plasmids was confirmed by PCR on heat-inactivated cell lysates using primers listed in (**Table S2 & S3**).

Cloning and purification of His_6-RegX3

His_6-RegX3 was cloned in pET28b+, which contains a 6-histidine (His_6) tag and a kanamycin resistance cassette for selection. The *regX3* coding sequence was PCR amplified using *M. tuberculosis* Erdman genomic DNA as template (F primer: 5'-tt**catatg**atgaccagtgtgttgattgtggagga-3', *Nde*I restriction site in bold; R primer: 5'-tt**gctagc**ctagccctcgagtttgtagccca-3', *Nhe*I restriction site in bold), cloned in pCR2.1 (Invitrogen) and sequenced. *regX3* was removed from pCR2.1 by restriction with *Nde*I and *Nhe*I, gel purified, and ligated to similarly digested pET28b+. Recombinant N-terminal tagged His_6-RegX3 was expressed in *E. coli* BL21 (DE3). Bacteria were grown at 37˚C with shaking to mid-exponential phase (OD_{600} 0.5) in LB containing 30 µg/ml kanamycin. His_6-RegX3 expression was induced with 0.1 mM IPTG for 3 hours at 37˚C with shaking. Cells were concentrated 100-fold in lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0), incubated with 1 mg/ml lysozyme (30 min on ice) and lysed by sonication. His_6-RegX3 was bound to Ni-NTA agarose (Qiagen) for 1 hour at 4˚C and loaded on a column. The column was washed with lysis buffer containing 20 mM imidazole. His_6-RegX3 was eluted using lysis buffer containing 250 mM imidazole. To remove contaminants that co-purified with His_6-RegX3, the protein was passed through an Amicon Ultra centrifugal filter with a 50 kDa cutoff (Millipore). Purified His_6-RegX3 was dialyzed in PBS and concentrated with Amicon Ultra centrifugal filters with a 10 kDa cutoff (Millipore).

Quantitative RT-PCR

To test gene expression in P_i-rich growth conditions, *M. tuberculosis* bacteria were cultured in complete Middlebrook 7H9 medium to mid-exponential phase (OD_{600} 0.4-0.6). To assess induction of gene expression in response to P_i starvation, cultures were grown in 7H9 to mid-exponential phase (OD_{600} 0.4-0.6), then washed twice and resuspended at OD_{600} 0.2 in P_i-free 7H9. P_i-limited cultures were grown at 37°C with aeration and bacteria were collected at 0, 24, 48, 72 and 96 hours. Bacteria were collected by centrifugation (3700 x *g*,10 min, 4°C). Total RNA was extracted using TRIzol (Invitrogen, CA) with 0.1% polyacryl carrier (Molecular Research Center, Inc) by bead beating with 0.1 mm zirconia beads (BioSpec Products). Equivalent amounts of total RNA were DNase-treated using Turbo DNase (Invitrogen) according to the manufacturer's instructions. Equivalent amounts of total RNA (500 ng) were converted to cDNA using the Transcriptor First Strand cDNA Synthesis Kit (Roche) and random hexamer primers. A no reverse transcriptase control was done for each RNA sample to quantify potential DNA contamination. The reverse transcription cycle parameters were as follows: 10 min at 25°C (annealing of primers), 60 min at 50°C (elongation), and 5 min at 85°C (heat inactivation of reverse transcriptase). Resulting cDNA was stored at -20°C until used for quantitative RT-PCR reactions.

Quantitative PCR primers to amplify an internal region of the genes or intergenic regions of interest (*esxM, esxN, espG$_5$, mycP$_5$, eccD$_5$, ppe41, eccB$_5$, modD, espA, esxB, udgA, mgtA, sigA, 16S rRNA*, *pe19,* and the *ppe27-pe19* intergenic region) were designed with similar annealing temperatures (58-60°C) using either Primer Express software (Applied Biosystems) or ProbeFinder Assay Design software (Roche). Sequences of primers used are listed in **Table S4**. Quantitative RT-PCR reactions were prepared using 2x SYBR Green master mix (Roche), 2.5 µM primer mix and 1 µl cDNA. All reactions were run on a LightCycler 480 (Roche) using the following cycle parameters: 95°C for 10 min; 45 cycles of 95°C for 10s, 60°C for 20s, and 72°C for 20s with data collected once per cycle during the extension phase; and one cycle of 95°C for 5s, 65°C for 1m, 97°C with a ramp rate of 0.11 °C/s for generation of melting curves. Cycle threshold values (C_p,Roche nomenclature) were converted to copy numbers using standard curves for

each gene, and gene copy numbers were normalized to *sigA* (P_i-rich cultures) or 16S rRNA (P_i-limited cultures).

Protein preparation for immunoblots

M. tuberculosis cultures were grown from frozen stocks to mid-exponential phase (OD_{600} 0.4-0.6) in complete Middlebrook 7H9. Bacteria were collected by centrifugation (2800 x *g*, 10 min), resuspended in 7H9 medium at a starting OD_{600} of 0.05-0.1 and grown to mid-exponential phase again. For experiments performed in P_i-rich conditions, bacteria were collected by centrifugation (2800 x *g*, 10 min) and used to inoculate Sauton's medium supplemented with 0.1% Tween-80 at a starting OD_{600} of 0.05-0.1. Cultures were incubated at 37°C with aeration to late exponential phase (OD_{600} 0.8-1.0), bacteria were collected by centrifugation (2800 x *g*, 10 min) and resuspended in Sauton's medium without Tween-80 at a starting OD_{600} of 0.8-1.0. Cultures were incubated at 37°C with aeration for 5 days before protein isolation. For P_i limitation experiments, the bacteria grown in 7H9 as above were collected by centrifugation (2800 x *g*, 10 min) and washed once with P_i-free Sauton's medium. Bacteria were then collected by centrifugation and resuspended in either Sauton's complete medium supplemented with Tween-80 (control) or P_i-free Sauton's medium supplemented with Tween-80 to which 250 µM KH_2PO_4 was added exogenously. Cultures were incubated at 37°C with aeration to late exponential phase (OD_{600} 0.8-1.0). Bacteria were collected by centrifugation (2800 x *g*, 10 min) and resuspended in the final medium used for protein secretion. For experiments conducted in P_i-replete conditions, Sauton's complete medium without Tween-80 was used. For P_i-limitation or P_i-titration samples, P_i-free Sauton's medium without Tween-80 to which 250, 25 or 2.5 µM KH_2PO_4 was added exogenously was used. For experiments to test whether P_i or K^+ was the relevant limiting ion, P_i-free Sauton's medium supplemented with 2.5 µM NaH_2PO_4 and 4 mM KCl (P_i limited) or 2.5 µM KCl and 4 mM NaH_2PO_4 (K^+ limited) was used. Cultures were incubated at 37°C with aeration for 5 days before protein was isolated.

Bacteria were collected by centrifugation (4700 x *g*, 15 min, 4°C). Culture supernatants were sterilized using 0.45 µm and then 0.22 µm syringe filters (Millipore). Complete EDTA-free protease inhibitor tablets (Roche) were added to each supernatant.

Supernatants were concentrated roughly 25-fold by centrifugation (2400 x *g*, 4°C) using VivaSpin 5 kDa molecular weight cut-off spin columns (Sartorius). Whole cell lysates were prepared by washing the pellet twice in cold PBS, then resuspending in PBS containing Complete EDTA-free protease inhibitors (Roche) and bead beating with 0.1 mm zirconia beads (BioSpec Products). Beads and unlysed material were removed by centrifugation (600 x *g*, 5 min, 4°C). Large cell debris was removed from the lysate by centrifugation (3000 x *g*, 10 min, 4°C). Cell lysates were passaged through a Nanosep MF column with a 0.22 µm filter (Pall Life Sciences) by centrifugation (14000 x *g*, 3 min, 4°C) to remove any remaining intact cells. Total protein concentration in each sample was quantified using the Pierce BCA Protein Concentration Assay kit (Thermo Scientific). Proteins were stored at 4°C for immediate use, or at -80°C with glycerol at 15% final concentration.

Immunoblot analysis

The indicated amounts of culture supernatant or whole cell lysate proteins were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) on Mini-PROTEAN TGX Any kD gels (Bio-Rad). Proteins were transferred to nitrocellulose membranes (Whatman) by electrophoresis and blocked overnight at 4°C in PBS-T (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, 0.05% Tween-20) containing 5% non-fat milk powder. Membranes were washed in PBS-T and probed for 1 hour at room temperature with the primary antisera diluted in PBS-T containing 2.5% non-fat milk powder. Primary antisera were used at the following dilutions: rabbit α-EsxN 1:1000; rabbit α-$EspG_5$ 1:1000; rabbit α-$EccB_5$ 1:5000; rabbit α-$EccD_5$ 1:500; rabbit α-PPE41 1:1000; rabbit α-EsxB 1:10,000; rat α-EspB 1:2000; rabbit α-ModD 1:5000; mouse α-GroEL2 1:10,000; rabbit α-Antigen 85 complex 1:5000. Membranes were washed in PBS-T again, and incubated for 1 hour at room temperature with the appropriate secondary antibody (either goat-anti-rabbit, rabbit-anti-mouse, or rabbit-anti-rat conjugated to HRP, Sigma) diluted 1:20,000 in PBS-T containing 2.5% non-fat milk powder. Membranes were washed again in PBS-T and the reactive bands were detected using SuperSignal West Pico substrate (Thermo Scientific) or Chemiluminescent Peroxidase Substrate (Sigma). Blots were either exposed to film (Blue lite autorad film, GeneMate) and developed using a film processor (Konica, SRX-101A) or visualized

using Studio Imager software. Protein bands from scanned Western blot images were quantified using Image Studio Lite software, version 5.0. Images were imported as jpegs, and rectangular work areas of equivalent size without background correction were used to define regions for quantification of signal intensity.

Standard RT-PCR

To determine the length of transcripts initiating 5' of *pe19*, *M. tuberculosis* bacteria were grown to mid-exponential phase (OD_{600} 0.4-0.6) in complete 7H9 (P_i-rich conditions) or grown to mid-exponential phase in complete 7H9, then washed twice, resuspended at OD_{600} 0.2 in P_i-free 7H9 and cultured for 24 hours (P_i-limiting conditions). RNA was extracted, treated with DNase, and converted to cDNA as described above. Standard PCR reactions were run using 1 μl of cDNA as template, 0.4 μM primers, and Recombinant Taq polymerse (Invitrogen). Cycle parameters were: 95˚C 5 min; 40 cycles of 95˚C 15 sec, 56˚C 15 sec, 72˚C 2.5 min; 72˚C 10 min. The following primer pairs were used: ig4F/esxNR3, ig4F/espG5R1, ig3F/esxNR3 (**Table S4**). PCR products were analyzed by gel electrophoresis.

To validate *esxN* deletion strains, *M. tuberculosis* bacteria were grown to mid-exponential phase (OD_{600} 0.4-0.6) in complete 7H9. RNA was extracted, treated with DNase, and converted to cDNA as described above. Standard PCR reactions were run using 1 μl of cDNA as template, 0.4 μM primers, and Recombinant Taq polymerse (Invitrogen). Cycle parameters were: 95˚C 5 min; 40 cycles of 95˚C 15 sec, 56˚C 15 sec, 72˚C 2.5 min; 72˚C 10 min. The following primer pairs were used: Q91F2/Q94R1 (**Table S4**). PCR products were analyzed by gel electrophoresis.

Electrophoretic mobility shift assays (EMSAs)

Double-stranded DNA probes were PCR amplified using *M. tuberculosis* Erdman genomic DNA as template and appropriate primers (**Table S5**). Probes were labeled using the DIG Gel Shift Kit, 2nd Generation (Roche), following the recommended protocols. Approximately 0.5 ng of DIG-labeled probe was added to binding reactions

containing binding buffer (Roche), poly[d(I-C)], poly L-lysine, and 0.25 – 1 μg purified His_6-RegX3 in 20 ul total volume and incubated at room temperature for 15 min. Where appropriate, a 400-fold excess of unlabeled specific (Probe A) or non-specific (*dnaN*) competitor or anti-His_6 antibodies (THE™ His tag antibody, Genscript) were added to the reaction mixture. Binding reactions that include unlabeled competitor probe were incubated for 15 min at room temperature before adding the DIG-labeled probe, then incubated an additional 15 min. DNA-protein complexes were resolved using 5% native polyacrylamide gels, transferred and UV-crosslinked to nylon membranes (Roche). Membranes were washed with wash buffer (DIG wash and block buffer set, Roche), blocked for 30 min in blocking solution (Roche) and incubated with anti-DIG-AP antibodies (Roche) at a 1:10,000 dilution for 30 min at room temperature. Labeled probes were detected using CDP-Star ready-to-use substrate (Roche). Membranes were exposed to film (Blue lite autrorad film, Genemate) and developed using a film processor (Konica, SRX-101A).

RNS Stress Assay

Bacteria were grown in 7H9 complete medium to mid-exponential phase (OD_{600} 0.5) and diluted to an OD_{600} 0.05 in 7H9 without supplements (no ADS). The cultures were incubated at 37°C shaking, after addition of 0.2 mM NO donor diethylenetriamine-NONOate (DETA-NO, Sigma). DETA-NO was added again at 24 and 48 hours, and was made fresh each day. CFU were enumerated at 0 and 72 hours by plating serial dilutions of culture on 7H10 agar.

Acid Stress Assay

Bacteria were grown in 7H9 complete medium to mid-exponential phase (OD_{600} 0.5) and diluted to an OD_{600} 0.05 in acidified 7H9 medium (pH 5.5). Cultures were incubated at 37°C, shaking for 7 days. CFU were enumerated at 0, 2, 4 and 7 days by plating serially diluted cultures on 7H10 agar.

ROS Stress Assay

Bacteria were grown in 7H9 complete medium to mid-exponential phase (OD_{600} 0.5) and diluted to an OD_{600} 0.05 in fresh 7H9 medium, and incubated at 37°C after addition of 3 mM H_2O_2. CFU were enumerated at 0 and 24 hours by plating serially diluted cultures on 7H10 agar.

Cell Wall Stress Assay

Bacteria were grown in 7H9 complete medium to mid-exponential phase (OD_{600} 0.5) and diluted to an OD_{600} 0.05 in fresh 7H9 medium, and incubated at 37°C after addition of 0.125% SDS. CFU were enumerated at 0 and 24 hours by plating serially diluted cultures on 7H10 agar.

Mouse Infections

Female C57BL/6 and $Nos2^{-/-}$ mice 6-8 weeks of age were purchased from Jackson Laboratories. $IrgM1^{-/-}$ mice were bred under specific-pathogen-free conditions at the University of Minnesota - Research Animal Resources. Mice were infected with ~100 CFU using an inhalation exposure system (GlasCol). The nebulizer was loaded with a bacterial suspension in PBS containing 0.05% Tween-80 at an OD_{600} of 0.0075-0.01. Mice were exposed to the generated aerosol for 20 minutes. Infected mice were euthanized with CO_2 overdose. Bacterial CFU were enumerated by plating serial dilutions of lung homogenates on 7H10 agar + cycohexamide and incubated at 37°C for 3-4 weeks before counting. Animal protocols used in this work were reviewed and approved by University of Minnesota Institutional Animal Care and Use committee. All animal experiments were conducted in accordance with recommendations in the *Guide for the care and use of laboratory animals* of the National Institutes of Health.

In vivo gene expression study

C57BL/6 mice were infected via the aerosol route with WT Erdman *Mtb*. Groups of 4 mice were sacrificed at 1, 2, 3, 4 and 8 weeks post infection. Mouse lungs were divided,

with the left lobe being used for CFU enumeration, and the right four lobes used for NRA extraction. RNA was extracted from the lungs using standard protocols, with the following alterations: tissue was suspended in RNALater (Ambion AM7023), only DEPC water was used, and samples were kept on ice until processed. Bacterial RNA was enriched using Ambion MICROBEnrich kit (AM1901), and qRT-PCR analysis was performed.

Sample prep for mass spectrometry experiment

WT and Δ*pstA1* culture filtrate proteins were prepared as previously described [158], and were subjected to SDS-PAGE. The gel was fixed for 30 minutes using a 40% ethanol/10% acetic acid mixture, then stained using Imperial Protein Stain (cat # 24615, Thermo Scientific) and destained in water. The fixed, stained gel was then transported to the Center for Mass Spectrometry & Proteomics in the University of Minnesota. Proteins were digested using the in-gel trypsin digest method. Bands of interest from Δ*pstA1* samples, corresponding to 75, 50, 37 & 25 kD were excised from the polyacrylamide gel (3 bands of each size were pooled together) and cut into ~2 x 2 mm cubes. Gel slices were washed in 1:1 100mM ammonium bicarbonate:acetonitrile for 15 min at room temperature two times. Gel slices were then incubated with 100% acetonitrile until pieces shrink and turn white and semi-opaque. Proteins were then reduced and alkylated by incubating in 10 mM DTT in 10 mM ammonium bicarbonate for 1 hr at 56 °C, followed by incubation in 55 mM iodoacetamide in 100 mM NH_4HCO_3 for 30 min at room temperature in the dark. Gel slices were washed in 1:1 acetonitrile:100mM ammonium bicarbonate 2 times. Gel slices were incubated with 100% acetonitrile, then rehydrated in digestion buffer at 4 °C (50 mM NH_4HCO_3, 5 mM $CaCl_2$, 12.5 ng/ul trypsin) for 15 minutes on ice. Gel slices were then incubated in 50 mM NH_4HCO_3, 5 mM $CaCl_2$ at 37 °C overnight. Peptides were extracted by low speed centrifugation to pellet the gel slices. 50% acetonitrile, 0.3% formic acid was added to the peptides and incubated for15 minutes, then spun again using low speed centrifugation. 80% acetonitrile and 0.3% formic acid was added to the supernatant and incubated for 15 minutes. Peptides were incubated at -80 °C for 30 minutes, then dried in a speed vac. Samples were stored at -80 °C prior to analysis by mass spectrometry. Samples were run on an Orbitrap Velos

mass spectrometer (ThermoFisher, Watham, MA). Data was analyzed by comparing peptide sequences with a *Mtb* proteomic database using Scaffold V4 software.

Statistical Analysis

Student's unpaired t-test was used to compare wild-type *M. tuberculosis* to mutant strains. *P* values were calculated using GraphPad Prism 6 software. *P* values <0.05 were considered significant.

Table S1. Primers used for strain construction. Sequences in blue indicate overlapping sequence, and sequences in red indicate mutated sequence. Restriction sites used are indicated.

Construct name	Forward primer	Reverse primer	Vector
Δ*esxN* 5' piece	ttttaattaagtcgttcgtgaccacac agc (PacI)	ttcctaggaatcgtcatattggtgttctcct (AvrII)	pJG1100
Δ*esxN* 3' piece	ttcctaggtgggcctaaaactgaac ttcagtcgc (AvrII)	ttggcgcgcctcggcgtggtgaatcgact c (AscI)	pJG1100
esxN comp	ttgaattcgctgcgtacgctttctcac a (EcoRI)	gctgccgcgactgaagttaagcttaa (HindIII)	pMV361
esxN comp	ttaagcttatgacgattaattaccag ttcgggga (HindIII)	ttgaattcttaggcccagctggagcc (EcoRI)	pTIC10a
esx-5 Tet 5'	ttttaattaaagctggccgaatccca atga (PacI)	ttactagttttgcggcccactcgagata (SpeI)	pJG1100
esx-5 Tet 3'	ttcatatgtcgttcgtgaccacacag cc (NdeI)	ttggcgcgccgtgagaaagcgtacgca gctg (AscI)	pJG1100
esx-5 Tet Tet operon	ttactagtatctagaggtgaccaca acgac (anneals upstream of the Tet repressor)	ttcatatgaattcggcgccaagc (downstream of the Tet operator RBS)	pJG1100
Δ*espG5* 5'	tttttaattaaatcggcgtccataaca gcag (PacI)	ttcctaggttgatccatcgctacctcagc (AvrII)	pJG1100
Δ*espG5* 3'	ttcctaggacacacagcagagtat gacgc (AvrII)	tttctcgaggcggtgtagatggtggtcaa (XhoI)	pJG1100
Δ*espG5* 2nd 5'	tttttaattaaatcggcgtccataaca gcag (PacI)	ttcctaggttgatccatcgctacctcagc (AvrII)	pJG1100
Δ*espG5* 2nd 3'	ttcctaggggagtgaagaccgtttt gga (AvrII)**redesigned	tttctcgaggcggtgtagatggtggtcaa (XhoI)	pJG1100
PM 5' F /3' R	ccttaattaaACGACGATCG TTCGTTTGCTC (PacI)	ttggcgcgccCGACGTCCCCGA ACTGGTAATTA (AscI)	pJG1100
BS 5'R1/3'F1	caagtgtGTGGTAACGCCA CAGTAACAG	gttaccacACACTTGCGAACGG GTGG	pJG1100
DR2 5'R1 / 3' F1	agtgtgacaaAGTTGGCAC CGCAAATTCAC	ccaactttgtcACACTTGCGAAC GGGTGG	pJG1100
Spacer 5'R1 / 3'F1	ttcacctgaAGTTGGCACCG CAAATTCAC	ccaacttcaGGTGAACACTTGC GAACGG	pJG1100

Table S2. Primer pairs used for verification of plasmid insertions and gene deletions

Gene deletion or mutation	5' integration	3' integration	Gene deletion or mutation
Δ*esxN*	EsxNF2/R3	EsxNF3/R2	Q91F2/Q94R1
Δ*espG*$_5$	Rv1794 F2/R3	Rv1794F3/R2	Rv1794F2/R2
Δ*pstA1*ΔBS	Tet Int F1/ Deletion HR2 R1	Deletion HR2 F1/ Tet Int R2	Tet Int F1/R2
RegX3 DR2	Tet Int F1/ DR2 HR2 R1	DR2 HR2 F1/ Tet Int R2	Tet Int F1/R2
RegX3 Spacer	Tet Int F1/ Spacer HR2 R1	Spacer HR2 F1/ Tet Int R2	Tet Int F1/R2
esx-5 Tet 5' & 3' regions	Tet Int F1/R1	Tet Int F2/Tet Int R2	Tet Int F1/R2
esx-5 Tet operon	pTIC Tet Op F1/R1	N/A	Tet Int F1/R2

Table S3. Primer sequences used for checking integration and deletion

Name	Purpose	Sequence (5'-3')
EsxN F2	check Δ*esxN*	GAGAGCCGTTGCTGGGATTA
EsxN R3	check Δ*esxN*	CTCAGCGGTCGAGTGCTAGTTA
EsxN F3	check Δ*esxN*	CTAGCGCCGAAAGCCACA
EsxN R2	check Δ*esxN*	GCATCTCCTCCATTGGCACGT
Q91F2	check Δ*esxN*	GGCGAACCTACAGGGTATTGG
Q94R1	check Δ*esxN*	CTCGTTTAGCCAGTCATTGGAA
Rv1794 F2	check Δ$espG_5$	CTCGCACGAACCTGCTGAA
Rv1794 R3	check Δ$espG_5$	ACTTCGGGTTTCACGCCCT
Rv1794 F3	check Δ$espG_5$	GTGCTGCTGTGTCCTGCA
Rv1794 R2	check Δ$espG_5$	CCAGCACCAGCACACCAAT
Tet Int F1	check RegX3 binding mutants / esx-5 Tet	AAGACGCCGTTGCTATGTATGG
Tet Int F2	check esx-5 Tet	ATCCCGGCGTTGATCTGT
Tet Int R1	check esx-5 Tet	CGACAACAAGACCCGTAAATGGT
Tet Int R2	check RegX3 binding mutants/ esx-5 Tet	CCAACTGGGTAATGAACTCCT
DR2 HR2 R1	check RegX3 binding mutants	CCGTTCGCAAGTGTGACAA
DR2 HR2 F1	check RegX3 binding mutants	GAATTTGCGGTGCCAACTTTGTC
Spacer HR2 R1	check RegX3 binding mutants	GTTCGCAAGTGTTCACCTGA
Spacer HR2 F1	check RegX3 binding mutants	TGAATTTGCGGTGCCAACTTCA
Deletion HR2 R1	check RegX3 binding mutants	CCGTTCGCAAGTGTGTGG
Deletion HR2 F1	check RegX3 binding mutants	ACTGTGGCGTTACCACACAC
pTIC Tet Op F1	check Tet operon	ACTAGTATCTAGAGGTGACCACA ACGAC
pTIC Tet Op R1	check Tet operon	GCTTGGCGCCGAATTCATATG

Table S4. Primers used for quantitative RT-PCR

Gene	For primer sequence (5'-3')	Rev primer sequence (5'-3')
eccB5 (F2/R2)	TGACGCCGGAGCAGGTTA	GGCAAAGCGTCGAACAATG
eccD5 (Q95F1/R1)	CCAACGCCCCAGTTTCG	AGCCTCACCGAGCTCTCTGA
espA (F1/R1)	GGCGCCAAGAAAGGTCTC	CCGGGATGTAGGTCAGGTC
espG5 (Q94F1/R1)	CGTTGCGCCTGAGTTACGTT	CTCGTTTAGCCAGTCATTGGAA
esxB (F1/R1)	GGCAGAGATGAAGACCGATG	ACCTGGTCGATCTGGGTTT
esxM (F2/R2)	GGCGGGCCGTTTTGAG	CGCGGACGCCCACAT
esxN (F3/R3)	TCAGGCCATCGTTCGTGAT	TGGAAGTTACGGCCCAACTG
modD (F1/R1)	GCACTCCTCAGCAAAACCAC	ACGATACGGGTGTCATTGG
modD (F2/R2)	CGGTTGGAGGATTCAGCTT	GTGCTGAACCGTAGTCGAAGT
mycP5 (Q96F2/R2)	CGTGCATGCTGCCAACAT	GTGCACGCTGGTCGATGA
ppe41 (F1/R1)	TGTGTGCAGCTGTCTGAGGT	CGTGATGTGCCCATTCATAG
sigA (F5/R5)	CTCAAACAGATCGGCAAGGT	CGCTAAGCTCGGTCATCAG
pe19 (Q91F2/R2)	GGCGAACCTACAGGGTATTGG	GCACTACTCCGGTGGTTGGA
udgA (Q0322F1/R1)	TCAACCCCGACCGTATCGT	CTCGCGGACGGCTACCT
mgtA (Q0557F1/R1)	GCGGTCTACCAAACCGATGT	CGTGCTGTCATCGGAATGC
pckAF1/R1	GTCGATGCCGATGAATGGCG	ACCGGTCGGCAGCTTCTC

Table S5. Primers used in EMSA experiments**.** Bold sequence indicates mutated bases.

Probe Name	Forward primer	Reverse primer	Type of alteration to IgF1/R1 sequence
IgF1/R1 / Up/ Probe A	CGCTGTTACTGTGGCGTTACC	GTGAACACTTGCGAA	None
IgF3/R3 / Down	CGGGTGGCATCGAAATCA	GCAACGGCTCTCTGCAAGA	None
Probe 1	AGGTGAATTTGCGGTGCCAA	ACTGCAACGCAACAAGTTGATT	5' truncation
Probe 2	TGAATTTGCGGTGCCAACTG	ACTGCAACGCAACAAGTTGATT	5' truncation
Probe 3	ACCCGGTCAAACGCTGTTAC	TCACCAGTTGGCACCGC	3' truncation
Probe 4	TGAATTTGCGGTGCCAACTGG	ACTGCAACGCAACAAGTTGATT	3' truncation
DR1 probe	CGCTGTTACTGTGGCGTTACC	AGTGTTCACCAGTTG**TAAA**GCAA	DR1 mutation
DR2 probe	CGCTGTTACTGTGGCGTTACC	AGTGT**GACAA**AGTTGGCACCGCAA	DR2 mutation
Spacer - probe	CGCTGTTACTGTGGCGTTACC	AGTGTTCACCAGGCACCGCAA	-3 bases from spacer
Spacer + probe	CGCTGTTACTGTGGCGTTACC	AGTGTTCACCTGAAGTTGGCACCGCAA	+3 bases in spacer
DR3 probe	ACA**TTGTC**ATTTGCGGTGCC	ACTGCAACGCAACAAGTTGATT	DR3 mutation
Ig Probe 2	CGCTGTTACTGTGGCGTTACC	AGTGTTCACCAGTTGGCAC	3' truncation
Ig Probe 3	GTGGCGTTACCACAGGTGA	GTGAACACTTGCGAA	5' truncation
Ig Probe 4	CACAGGTGAATTTGCGGTGC	ACGCAACAAGTTGATTTCGATGC	5' truncation
Ig Probe 5	TTTGCGGTGCCAACTGGT	ACGCAACAAGTTGATTTCGATGC	5' truncation
Ig Probe 6	CAACTGGTGAACACTTGCGAAC	GCAACGGCTCTCTGCAAGA	5' truncation
Ig Probe 9	ATTTGCGGTGCCAACTGGTG	ACTGCAACGCAACAAGTTGATT	5' truncation
Ig Probe 10	CCAACTGGTGAACACTTGCGA	TCGCAAGTGTTCACCAGTTGG	3' truncation
Ig Probe 11	CGCTGTTACTGTGGCGTTACC	CGCAAGTGTTCACCAGTTGG	3' truncation
Ig Probe 12	CGCTGTTACTGTGGCGTTACC	GCAAGTGTTCACCAGTTGGCA	3' truncation
Ig Probe 13	CGCTGTTACTGTGGCGTTACC	CAAGTGTTCACCAGTTGGCAC	3' truncation
Ig Probe 14	CGCTGTTACTGTGGCGTTACC	AAGTGTTCACCAGTTGGCACC	3' truncation
Ig Probe 15	CGCTGTTACTGTGGCGTTACC	AGTGTTCACCAGTTGGCACC	3' truncation
Ig Probe 18	CGCTGTTACTGTGGCGTTACC	AGTGT**GACAA**AGTTG**TAAA**GCAA	DR1 & DR2 mutation**

References

1. Koch, R., *Die Atiologie der Tuberkulose.* 1882.
2. Daniel, T.M., *The history of tuberculosis.* Respir Med, 2006. **100**(11): p. 1862-70.
3. Gutierrez, M.C., et al., *Ancient origin and gene mosaicism of the progenitor of Mycobacterium tuberculosis.* PLoS Pathog, 2005. **1**(1): p. e5.
4. Cave, A.J.E., *The evidence for the incidence of tuberculosis in ancient Egypt.* British Journal of Tuberculosis, 1939. **33**(142).
5. Crubezy, E., et al., *Identification of Mycobacterium DNA in an Egyptian Pott's disease of 5400 years old.* Anthroplogy, 1998. **321**: p. 941-51.
6. Daniel, T.M., *The origins and precolonial epidemiology of tuberculosis in the Americas: can we figure them out?* INT J TUBERC LUNG DIS, 2000. **4**(5): p. 395-400.
7. Hershkovitz, I., et al., *Tuberculosis origin: The Neolithic scenario.* Tuberculosis (Edinb), 2015. **95 Suppl 1**: p. S122-6.
8. Virginia S. Daniel and a.T.M. Daniel, *Old Testament Biblical References to Tuberculosis.* Clinical Infectious Diseases, 1999.
9. Donoghue, H.D., *Human tuberculosis--an ancient disease, as elucidated by ancient microbial biomolecules.* Microbes Infect, 2009. **11**(14-15): p. 1156-62.
10. WHO, *Global tuberculosis report 2016.* World Health Organization, Geneva, Switzerland, 2016.
11. Fishbein, S., et al., *Phylogeny to function: PE/PPE protein evolution and impact on Mycobacterium tuberculosis pathogenicity.* Mol Microbiol, 2015. **96**(5): p. 901-16.
12. Connolly, L., P. Edelstein, and L. Ramakrishnan, *Why is long-term therapy required to cure tuberculosis?* PloS Med, 2007. **4**(3).
13. Russell, D.G., *Mycobacterium tuberculosis and the intimate discourse of a chronic infection.* Immunological Reviews, 2011. **240**(1): p. 252-268.
14. Cole, S.T., *Inhibiting Mycobacterium tuberculosis within and without.* Philos Trans R Soc Lond B Biol Sci, 2016. **371**(1707).
15. Russell, D.G., *Who puts the tubercle in tuberculosis?* Nat Rev Microbiol, 2007. **5**(1): p. 39-47.
16. Tischler, A.D. and J.D. McKinney, *Contrasting persistence strategies in Salmonella and Mycobacterium.* Curr Opin Microbiol, 2010. **13**(1): p. 93-9.
17. Frieden, T.R., et al., *Tuberculosis.* The Lancet, 2003. **362**(9387): p. 887-899.
18. Guirado, E., L.S. Schlesinger, and G. Kaplan, *Macrophages in tuberculosis: friend or foe.* Semin Immunopathol, 2013. **35**(5): p. 563-83.
19. Ernst, J.D., *The immunological life cycle of tuberculosis.* Nat Rev Immunol, 2012. **12**(8): p. 581-91.
20. Dutta, N.K. and P.C. Karakousis, *Latent tuberculosis infection: myths, models, and molecular mechanisms.* Microbiol Mol Biol Rev, 2014. **78**(3): p. 343-71.
21. Russell, D.G., et al., *Foamy macrophages and the progression of the human tuberculosis granuloma.* Nat Immunol, 2009. **10**(9): p. 943-8.

22. Russell, D.G., et al., *Mycobacterium tuberculosis wears what it eats.* Cell Host Microbe, 2010. **8**(1): p. 68-76.
23. Kumar, H., T. Kawai, and S. Akira, *Toll-like receptors and innate immunity.* Biochem Biophys Res Commun, 2009. **388**(4): p. 621-5.
24. Saiga, H., Y. Shimada, and K. Takeda, *Innate immune effectors in mycobacterial infection.* Clin Dev Immunol, 2011. **2011**: p. 347594.
25. Weiss, G. and U.E. Schaible, *Macrophage defense mechanisms against intracellular bacteria.* Immunol Rev, 2015. **264**: p. 182-203.
26. Cavalcanti, Y.V., et al., *Role of TNF-Alpha, IFN-Gamma, and IL-10 in the Development of Pulmonary Tuberculosis.* Pulm Med, 2012. **2012**: p. 745483.
27. Awuh, J.A. and T.H. Flo, *Molecular basis of mycobacterial survival in macrophages.* Cell Mol Life Sci, 2017. **74**(9): p. 1625-1648.
28. Tiwari, S., et al., *Targeting of the GTPase Irgm1 to the phagosomal membrane via PtdIns(3,4)P(2) and PtdIns(3,4,5)P(3) promotes immunity to mycobacteria.* Nat Immunol, 2009. **10**(8): p. 907-17.
29. Schorey, J.S. and L.S. Schlesinger, *Innate Immune Responses to Tuberculosis.* Microbiol Spectr, 2016. **4**(6).
30. Porcelli, S.A. and W.R. Jacobs, Jr., *Tuberculosis: unsealing the apoptotic envelope.* Nat Immunol, 2008. **9**(10).
31. Clemens, D. and M. Horwitcz, *Characterization of the Mycobacterium tuberculosis phagosome and evidence that phagosomal maturation is inhibited.* J Exp Med, 1995. **181**: p. 257-270.
32. Kyei, G., et al., *Rab14 is critical for maintenance of Mycobacterium tuberculosis phagosome maturation arrest.* EMBO J, 2006. **25**(22): p. 5250-59.
33. Deretic, V., et al., *Mycobacterium tuberculosis inhibition of phagolysosome biogenesis and autophagy as a host defence mechanism.* Cell Microbiol, 2006. **8**(5): p. 719-27.
34. Kang, P.B., et al., *The human macrophage mannose receptor directs Mycobacterium tuberculosis lipoarabinomannan-mediated phagosome biogenesis.* J Exp Med, 2005. **202**(7): p. 987-99.
35. van der Wel, N., et al., *M. tuberculosis and M. leprae translocate from the phagolysosome to the cytosol in myeloid cells.* Cell, 2007. **129**(7): p. 1287-98.
36. Simeone, R., et al., *Phagosomal rupture by Mycobacterium tuberculosis results in toxicity and host cell death.* PLoS Pathog, 2012. **8**(2): p. e1002507.
37. Butler, R.E., et al., *The balance of apoptotic and necrotic cell death in Mycobacterium tuberculosis infected macrophages is not dependent on bacterial virulence.* PLoS One, 2012. **7**(10): p. e47573.
38. Srivastava, S., P.S. Grace, and J.D. Ernst, *Antigen Export Reduces Antigen Presentation and Limits T Cell Control of M. tuberculosis.* Cell Host Microbe, 2016. **19**(1): p. 44-54.
39. Noss, E.H., et al., *Toll-Like Receptor 2-Dependent Inhibition of Macrophage Class II MHC Expression and Antigen Processing by 19-kDa Lipoprotein of Mycobacterium tuberculosis.* The Journal of Immunology, 2001. **167**(2): p. 910-918.

40. Fulton, S.A., et al., *Inhibition of Major Histocompatibility Complex II Expression and Antigen Processing in Murine Alveolar Macrophages by Mycobacterium bovis BCG and the 19-Kilodalton Mycobacterial Lipoprotein.* Infection and Immunity, 2004. **72**(4): p. 2101-2110.
41. Torres, M., et al., *Role of phagosomes and major histocompatibility complex class II (MHC-II) compartment in MHC-II antigen processing of Mycobacterium tuberculosis in human macrophages.* Infect Immun, 2006. **74**(3): p. 1621-30.
42. Sendide, K., et al., *Mycobacterium bovis BCG Attenuates Surface Expression of Mature Class II Molecules through IL-10-Dependent Inhibition of Cathepsin S.* The Journal of Immunology, 2005. **175**(8): p. 5324-5332.
43. Blomgran, R. and J.D. Ernst, *Lung neutrophils facilitate activation of naive antigen-specific CD4+ T cells during Mycobacterium tuberculosis infection.* J Immunol, 2011. **186**(12): p. 7110-9.
44. Blomgran, R., et al., *Mycobacterium tuberculosis inhibits neutrophil apoptosis, leading to delayed activation of naive CD4 T cells.* Cell Host Microbe, 2012. **11**(1): p. 81-90.
45. Scott-Browne, J.P., et al., *Expansion and function of Foxp3-expressing T regulatory cells during tuberculosis.* J Exp Med, 2007. **204**(9): p. 2159-69.
46. Orme, I.M., R.T. Robinson, and A.M. Cooper, *The balance between protective and pathogenic immune responses in the TB-infected lung.* Nat Immunol, 2015. **16**(1): p. 57-63.
47. Flynn, J.L. and J. Chan, *Immunology of tuberculosis.* Ann Rev Immunol, 2001. **19**: p. 93-129.
48. Rayamajhi, M. and E.A. Miao, *Just say NO to NLRP3.* Nat Immunol, 2013. **14**(1): p. 12-4.
49. Cunningham, K. and W. Wickner, *Specific recognition of the leader region of precursor proteins is required for the activation of translocation ATPase of Escherichia coli.* Proc Natl Acad Sci U S A, 1989. **86**: p. 8630-34.
50. Lill, R., R. Downhan, and W. Wickner, *ATPase activity of SecA is regulated by acidic phospholipids, SecY, and the leader and mature domains of precursor proteins.* Cell, 1990. **60**: p. 271-80.
51. van der Wolk, J., J. de Wit, and A.J. Driessen, *The catalytic cycle of the Escherichia coli SecA ATPase comprises two distinct preprotein translocation events.* EMBO J, 1997. **16**(24): p. 7297-7304.
52. Schiebel, E., et al., *Hþ and ATP function at different steps of the catalytic cycle of preprotein translocase.* Cell, 1991. **64**: p. 927-39.
53. Hanada, M., et al., *Reconstitution of an efficient protein translocation machinery comprising SecA and the three membrane proteins, SecY, SecE, and SecG.* J Biol Chem, 1994. **269**(38): p. 23625-23631.
54. Paetzel, M., R.E. Dalbey, and N.C. Strynadka, *Crystal structure of a bacterial signal peptidase apoenzyme: implications for signal peptide binding and the Ser-Lys dyad mechanism.* J Biol Chem, 2002. **277**(11): p. 9512-9.
55. Watanabe, M., *Secb Functions as a Cytosolic Signal Recognition Factor for Protein Export in Escherichia coli.* Cell, 1989. **58**(4): p. 695-705.

56. Hartl, F.U., et al., *The binding cascade of SecB to SecA to SecY/E mediates preprotein targeting to the E. coli plasma membrane.* Cell, 1990. **63**(2): p. 269-79.
57. van der Sluis, E.O. and A.J. Driessen, *Stepwise evolution of the Sec machinery in Proteobacteria.* Trends Microbiol, 2006. **14**(3): p. 105-8.
58. Egea, P. and R.M. Stroud, *Lateral opening of a translocon upon entry of protein suggests the mechanism of insertion into membranes.* Proc Natl Acad Sci U S A, 2010. **107**(40): p. 17182-17187.
59. Dalbey, R.E. and A. Kuhn, *YidC family members are involved in the membrane insertion, lateral integration, folding, and assembly of membrane proteins.* J Cell Biol, 2004. **166**(6): p. 769-74.
60. Dilks, K., et al., *Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey.* J Bacteriol, 2003. **185**(4): p. 1478-83.
61. Saint-Joanis, B., et al., *Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of Mycobacterium tuberculosis, increases beta-lactam susceptibility and virulence.* J Bacteriol, 2006. **188**(18): p. 6669-79.
62. Tarry, M.J., et al., *Structural analysis of substrate binding by the TatBC component of the twin-arginine protein transport system.* Proc Natl Acad Sci U S A, 2009. **106**(32): p. 13284-9.
63. Alami, M., et al., *Differential interactions between a twin-arginine signal peptide and its translocase in Escherichia coli.* Mol Cell, 2003. **12**: p. 937-946.
64. Robinson, C., et al., *Transport and proofreading of proteins by the twin-arginine translocation (Tat) system in bacteria.* Biochim Biophys Acta, 2011. **1808**(3): p. 876-84.
65. Bageshwar, U.K. and S.M. Musser, *Two electrical potential-dependent steps are required for transport by the Escherichia coli Tat machinery.* J Cell Biol, 2007. **179**(1): p. 87-99.
66. Berks, B.C., *A common export pathway for proteins binding complex redox cofactors?* Mol Microbiol, 1996. **22**(3): p. 393-404.
67. DeLisa, M.P., D. Tullman, and G. Georgiou, *Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway.* Proc Natl Acad Sci U S A, 2003. **100**(10): p. 6115-20.
68. Houben, E.N., et al., *Composition of the type VII secretion system membrane complex.* Mol Microbiol, 2012. **86**(2): p. 472-84.
69. Champion, P.A. and J.S. Cox, *Protein secretion systems in Mycobacteria.* Cell Microbiol, 2007. **9**(6): p. 1376-84.
70. Braunstein, M., et al., *Two nonredundant SecA homologues function in mycobacteria.* J Bacteriol, 2001. **183**(24): p. 6979-90.
71. Backert, S. and T.F. Meyer, *Type IV secretion systems and their effectors in bacterial pathogenesis.* Curr Opin Microbiol, 2006. **9**(2): p. 207-17.
72. Gerlach, R.G. and M. Hensel, *Protein secretion systems and adhesins: the molecular armory of Gram-negative pathogens.* Int J Med Microbiol, 2007. **297**(6): p. 401-15.

73. Bitter, W., et al., *Type VII secretion in mycobacteria: classification in line with cell envelope structure.* Trends Microbiol, 2009. **17**(8): p. 337-8.
74. Burts, M., et al., *EsxA and EsxB are secreted by an ESAT-6-like system that is required for the pathogenesis of Staphylococcus aureus infections.* Proc Natl Acad Sci U S A, 2005. **102**(4): p. 1169-1174.
75. Unnikrishnan, M., et al., *The Enigmatic Esx Proteins: Looking Beyond Mycobacteria.* Trends Microbiol, 2017. **25**(3): p. 192-204.
76. Sorensen, A.L., et al., *Purification and characterization of a low-molecular-mass T-cell antigen secreted by Mycobacterium tuberculosis.* Infect Immun, 1995. **63**(5): p. 1710-7.
77. Mahairas, G., et al., *Molecular Analysis of Genetic Differences between Mycobacterium bovis BCG and Virulent M. bovis.* J Bacteriol, 1996. **178**(5): p. 1274-82.
78. Gey Van Pittius, N.C., et al., *The ESAT-6 gene cluster of Mycobacterium tuberculosis and other high G+C Gram-positive bacteria.* Genome Biol, 2001. **2**(10): p. RESEARCH0044.
79. Newton-Foot, M., et al., *The plasmid-mediated evolution of the mycobacterial ESX (Type VII) secretion systems.* BMC Evol Biol, 2016. **16**: p. 62.
80. Bitter, W., et al., *Systematic genetic nomenclature for type VII secretion systems.* PLoS Pathog, 2009. **5**(10): p. e1000507.
81. Pallen, M.J., *The ESAT-6/WXG100 superfamily -- and a new Gram-positive secretion system?* Trends Microbiol, 2002. **10**(5): p. 209-12.
82. Stoop, E.J., W. Bitter, and A.M. van der Sar, *Tubercle bacilli rely on a type VII army for pathogenicity.* Trends Microbiol, 2012. **20**(10): p. 477-84.
83. Beckham, K.S., et al., *Structure of the mycobacterial ESX-5 type VII secretion system membrane complex by single-particle analysis.* Nat Microbiol, 2017. **2**: p. 17047.
84. Houben, E.N., K.V. Korotkov, and W. Bitter, *Take five - Type VII secretion systems of Mycobacteria.* Biochim Biophys Acta, 2014. **1843**(8): p. 1707-16.
85. van Winden, V.J., et al., *Mycosins Are Required for the Stabilization of the ESX-1 and ESX-5 Type VII Secretion Membrane Complexes.* MBio, 2016. **7**(5).
86. Gray, T.A., et al., *Intercellular communication and conjugation are mediated by ESX secretion systems in mycobacteria.* Science, 2016. **354**(6310).
87. Malen, H., et al., *Comprehensive analysis of exported proteins from Mycobacterium tuberculosis H37Rv.* Proteomics, 2007. **7**(10): p. 1702-18.
88. Sassetti, C.M., D.H. Boyd, and E.J. Rubin, *Genes required for mycobacterial growth defined by high density mutagenesis.* Mol Microbiol, 2003. **48**(1): p. 77-84.
89. Serafini, A., et al., *Characterization of a Mycobacterium tuberculosis ESX-3 conditional mutant: essentiality and rescue by iron and zinc.* J Bacteriol, 2009. **191**(20): p. 6340-4.
90. Maciag, A., et al., *Global analysis of the Mycobacterium tuberculosis Zur (FurB) regulon.* J Bacteriol, 2007. **189**(3): p. 730-40.

91. Rodriguez, G.M., et al., *ideR, an Essential Gene in Mycobacterium tuberculosis: Role of IdeR in Iron-Dependent Gene Expression, Iron Metabolism, and Oxidative Stress Response.* Infection and Immunity, 2002. **70**(7): p. 3371-3381.
92. Siegrist, M.S., et al., *Mycobacterial Esx-3 is required for mycobactin-mediated iron acquisition.* Proc Natl Acad Sci U S A, 2009. **106**(44): p. 18792-18797.
93. Tufariello, J.M., et al., *Separable roles for Mycobacterium tuberculosis ESX-3 effectors in iron acquisition and virulence.* Proc Natl Acad Sci U S A, 2016. **113**(3): p. E348-57.
94. Manzanillo, P.S., et al., *Mycobacterium tuberculosis activates the DNA-dependent cytosolic surveillance pathway within macrophages.* Cell Host Microbe, 2012. **11**(5): p. 469-80.
95. Tan, T., et al., *The ESAT-6/CFP-10 secretion system of Mycobacterium marinum modulates phagosome maturation.* Cell Microbiol, 2006. **8**(9): p. 1417-29.
96. Xu, J., et al., *A unique Mycobacterium ESX-1 protein co-secretes with CFP-10/ESAT-6 and is necessary for inhibiting phagosome maturation.* Mol Microbiol, 2007. **66**(3): p. 787-800.
97. MacGurn, J.A. and J.S. Cox, *A genetic screen for Mycobacterium tuberculosis mutants defective for phagosome maturation arrest identifies components of the ESX-1 secretion system.* Infect Immun, 2007. **75**(6): p. 2668-78.
98. de Jonge, M.I., et al., *ESAT-6 from Mycobacterium tuberculosis dissociates from its putative chaperone CFP-10 under acidic conditions and exhibits membrane-lysing activity.* J Bacteriol, 2007. **189**(16): p. 6028-34.
99. Smith, J., et al., *Evidence for pore formation in host cell membranes by ESX-1-secreted ESAT-6 and its role in Mycobacterium marinum escape from the vacuole.* Infect Immun, 2008. **76**(12): p. 5478-87.
100. Samten, B., X. Wang, and P. Barnes, *Mycobacterium tuberculosis ESX-1 system-secreted protein ESAT-6 but not CFP10 inhibits human T-cell immune responses.* Tuberculosis (Edinb), 2009. **89**.
101. Behr, M.A., et al., *Comparative genomics of BCG vaccines by whole-genome DNA microarray.* Science, 1999. **284**(5419): p. 1520-3.
102. Abdallah, A.M., et al., *Mycobacterial secretion systems ESX-1 and ESX-5 play distinct roles in host cell death and inflammasome activation.* J Immunol, 2011. **187**(9): p. 4744-53.
103. Daleke, M.H., et al., *Conserved Pro-Glu (PE) and Pro-Pro-Glu (PPE) protein domains target LipY lipases of pathogenic mycobacteria to the cell surface via the ESX-5 pathway.* J Biol Chem, 2011. **286**(21): p. 19024-34.
104. Singh, V.K., et al., *Increased virulence of Mycobacterium tuberculosis H37Rv overexpressing LipY in a murine model.* Tuberculosis (Edinb), 2014. **94**(3): p. 252-61.
105. Bottai, D., et al., *Disruption of the ESX-5 system of Mycobacterium tuberculosis causes loss of PPE protein secretion, reduction of cell wall integrity and strong attenuation.* Mol Microbiol, 2012. **83**(6): p. 1195-209.

106. Sayes, F., et al., *Strong immunogenicity and cross-reactivity of Mycobacterium tuberculosis ESX-5 type VII secretion: encoded PE-PPE proteins predicts vaccine potential.* Cell Host Microbe, 2012. **11**(4): p. 352-63.
107. Alderson, M.R., et al., *Expression cloning of an immunodominant family of Mycobacterium tuberculosis antigens using human CD4(+) T cells.* J Exp Med, 2000. **191**(3): p. 551-60.
108. Fortune, S.M., et al., *Mutually dependent secretion of proteins required for mycobacterial virulence.* Proc Natl Acad Sci U S A, 2005. **102**(30): p. 10676-81.
109. Millington, K., et al., *Rv3615c is a highly immunodominant RD1 (Region of Difference 1)-dependent secreted antigen specfic for Mycobacterium tuberculosis infection.* Proc Natl Acad Sci U S A, 2011. **108**(14): p. 5730-35.
110. Chen, J.M., et al., *EspD is critical for the virulence-mediating ESX-1 secretion system in Mycobacterium tuberculosis.* J Bacteriol, 2012. **194**(4): p. 884-93.
111. Raghavan, S., et al., *Secreted transcription factor controls Mycobacterium tuberculosis virulence.* Nature, 2008. **454**(7205): p. 717-21.
112. Rosenberg, O.S., et al., *EspR, a key regulator of Mycobacterium tuberculosis virulence, adopts a unique dimeric structure among helix-turn-helix proteins.* Proc Natl Acad Sci U S A, 2011. **108**(33): p. 13450-5.
113. Blasco, B., et al., *Virulence regulator EspR of Mycobacterium tuberculosis is a nucleoid-associated protein.* PLoS Pathog, 2012. **8**(3): p. e1002621.
114. Rohde, K.H., R.B. Abramovitch, and D.G. Russell, *Mycobacterium tuberculosis invasion of macrophages: linking bacterial gene expression to environmental cues.* Cell Host Microbe, 2007. **2**(5): p. 352-64.
115. Gonzalo-Asensio, J., et al., *PhoP: a missing piece in the intricate puzzle of Mycobacterium tuberculosis virulence.* PLoS One, 2008. **3**(10): p. e3496.
116. Pang, X., et al., *MprAB regulates the espA operon in Mycobacterium tuberculosis and modulates ESX-1 function and host cytokine response.* J Bacteriol, 2013. **195**(1): p. 66-75.
117. Betts, J., et al., *Evaluation of a nutrient starvation model of Mycobacterium tuberculosis persistence by gene and protein expression profiling.* Mol Microbiol, 2002. **43**(3): p. 717-731.
118. He, H., et al., *MprAB is a stress-responsive two-component system that directly regulates expression of sigma factors SigB and SigE in Mycobacterium tuberculosis.* J Bacteriol, 2006. **188**(6): p. 2134-43.
119. Pang, X., et al., *Evidence for complex interactions of stress-associated regulons in an mprAB deletion mutant of Mycobacterium tuberculosis.* Microbiology, 2007. **153**(Pt 4): p. 1229-42.
120. Baker, J.J., B.K. Johnson, and R.B. Abramovitch, *Slow growth of Mycobacterium tuberculosis at acidic pH is regulated by phoPR and host-associated carbon sources.* Mol Microbiol, 2014. **94**(1): p. 56-69.
121. Zhang, M., et al., *EspI regulates the ESX-1 secretion system in response to ATP levels in Mycobacterium tuberculosis.* Mol Microbiol, 2014. **93**(5): p. 1057-65.

122. van der Woude, A.D., J. Luirink, and W. Bitter, *Getting across the cell envelope: mycobacterial protein secretion.* Curr Top Microbiol Immunol, 2013. **374**: p. 109-34.
123. Cole, S.T., et al., *Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.* Nature, 1998. **393**(6685): p. 537-44.
124. Korotkova, N., et al., *Structure of the Mycobacterium tuberculosis type VII secretion system chaperone EspG5 in complex with PE25-PPE41 dimer.* Mol Microbiol, 2014. **94**(2): p. 367-82.
125. Ekiert, D.C. and J.S. Cox, *Structure of a PE-PPE-EspG complex from Mycobacterium tuberculosis reveals molecular specificity of ESX protein secretion.* Proc Natl Acad Sci U S A, 2014. **111**(41): p. 14758-63.
126. Daleke, M.H., et al., *General secretion signal for the mycobacterial type VII secretion pathway.* Proc Natl Acad Sci U S A, 2012. **109**(28): p. 11342-7.
127. Daleke, M.H., et al., *Specific chaperones for the type VII protein secretion pathway.* J Biol Chem, 2012. **287**(38): p. 31939-47.
128. Gey van Pittius, N.C., et al., *Evolution and expansion of the Mycobacterium tuberculosis PE and PPE multigene families and their association with the duplication of the ESAT-6 (esx) gene cluster regions.* BMC Evol Biol, 2006. **6**: p. 95.
129. Sampson, S.L., *Mycobacterial PE/PPE proteins at the host-pathogen interface.* Clin Dev Immunol, 2011. **2011**: p. 497203.
130. Mazandu, G.K. and N.J. Mulder, *Function prediction and analysis of Mycobacterium tuberculosis hypothetical proteins.* Int J Mol Sci, 2012. **13**(6): p. 7283-302.
131. Sampson, S.L., et al., *Expression, characterization and subcellular localization of the Mycobacterium tuberculosis PPE gene Rv1917c.* Tuberculosis (Edinb), 2001. **81**(5-6): p. 305-17.
132. Delogu, G., et al., *Rv1818c-encoded PE_PGRS protein of Mycobacterium tuberculosis is surface exposed and influences bacterial cell structure.* Mol Microbiol, 2004. **52**(3): p. 725-33.
133. Basu, S., et al., *Execution of macrophage apoptosis by PE_PGRS33 of Mycobacterium tuberculosis is mediated by Toll-like receptor 2-dependent release of tumor necrosis factor-alpha.* J Biol Chem, 2007. **282**(2): p. 1039-50.
134. Tiwari, B.M., et al., *The Mycobacterium tuberculosis PE proteins Rv0285 and Rv1386 modulate innate immunity and mediate bacillary survival in macrophages.* PLoS One, 2012. **7**(12): p. e51686.
135. Bhat, K.H., et al., *PPE2 protein of Mycobacterium tuberculosis may inhibit nitric oxide in activated macrophages.* Ann N Y Acad Sci, 2013. **1283**: p. 97-101.
136. Nair, S., et al., *The PPE18 of Mycobacterium tuberculosis interacts with TLR2 and activates IL-10 induction in macrophage.* J Immunol, 2009. **183**(10): p. 6269-81.
137. Zumbo, A., et al., *Functional dissection of protein domains involved in the immunomodulatory properties of PE_PGRS33 of Mycobacterium tuberculosis.* Pathog Dis, 2013. **69**(3): p. 232-9.

138. Abdallah, A.M., et al., *A specific secretion system mediates PPE41 transport in pathogenic mycobacteria.* Mol Microbiol, 2006. **62**(3): p. 667-79.
139. Poulet, S. and S.T. Cole, *Characterization of the highly abundant polymorphic GC rish repetitive sequence (PGRS) present in Mycobacterium tuberculosis.* Arch Microbiol, 1995. **163**: p. 87-95.
140. Abdallah, A.M., et al., *PPE and PE_PGRS proteins of Mycobacterium marinum are transported via the type VII secretion system ESX-5.* Mol Microbiol, 2009. **73**(3): p. 329-40.
141. Gao, R., T.R. Mack, and A.M. Stock, *Bacterial response regulators: versatile regulatory strategies from common domains.* Trends Biochem Sci, 2007. **32**(5): p. 225-34.
142. Ashby, M.K. and J. Houmard, *Cyanobacterial two-component proteins: structure, diversity, distribution, and evolution.* Microbiol Mol Biol Rev, 2006. **70**(2): p. 472-509.
143. Parish, T., *Two-Component Regulatory Systems of Mycobacteria.* Microbiol Spectr, 2014. **2**(1): p. MGM2-0010-2013.
144. Bretl, D.J., C. Demetriadou, and T.C. Zahrt, *Adaptation to environmental stimuli within the host: two-component signal transduction systems of Mycobacterium tuberculosis.* Microbiol Mol Biol Rev, 2011. **75**(4): p. 566-82.
145. Agrawal, R., et al., *The two-component signalling networks of Mycobacterium tuberculosis display extensive cross-talk in vitro.* Biochem J, 2015. **469**(1): p. 121-34.
146. Stock, A.M., V.L. Robinson, and P.N. Goudreau, *Two-component signal transduction.* Annu Rev Biochem, 2000. **69**: p. 183-215.
147. Parish, T., et al., *The senX3-regX3 two-component regulatory system of Mycobacterium tuberculosis is required for virulence.* Microbiology, 2003. **149**(Pt 6): p. 1423-35.
148. Singh, N. and A. Kumar, *Virulence factor SenX3 is the oxygen-controlled replication switch of Mycobacterium tuberculosis.* Antioxid Redox Signal, 2015. **22**(7): p. 603-13.
149. Tischler, A.D., et al., *Mycobacterium tuberculosis requires phosphate-responsive gene regulation to resist host immunity.* Infect Immun, 2013. **81**(1): p. 317-28.
150. Himpens, S., C. Locht, and P. Supply, *Molecular characterization of the mycobacterial SenX3-RegX3 two-component system: evidence for autoregulation.* Microbiology, 2000. **146 Pt 12**: p. 3091-8.
151. Rifat, D., W.R. Bishai, and P.C. Karakousis, *Phosphate depletion: a novel trigger for Mycobacterium tuberculosis persistence.* J Infect Dis, 2009. **200**(7): p. 1126-35.
152. Carroll, P., M.C. Faray-Kele, and T. Parish, *Identifying vulnerable pathways in Mycobacterium tuberculosis by using a knockdown approach.* Appl Environ Microbiol, 2011. **77**(14): p. 5040-3.
153. van Veen, H., et al., *Mechanism and energetics of secondary phosphate transport system in Acintobacter johnsonii 210A.* J Biol Chem, 1993. **268**(26): p. 19377-19383.

154. Wanner, B.L., *Phosphorus assimilation and control of the phosphate regulon.* American Society for Microbiology, 1996. **Escherichia coli and Salmonella: Cellular and Molecular Biology, vol. 1. American**: p. 1357-1381.
155. Beard, S., et al., *Evidence for the transport of zinc(II) ions via the pit inorganic phosphate transport system in Escherichia coli.* FEMS Microbiol Lett, 2000. **184**(2): p. 231-5.
156. Braibant, M., P. Gilot, and J. Content, *The ATP binding cassette (ABC) transport systems of Mycobacterium tuberculosis.* FEMS Microbiology Reviews, 2000. **24**: p. 449-467.
157. Niederweis, M., *Nutrient acquisition by mycobacteria.* Microbiology, 2008. **154**(Pt 3): p. 679-92.
158. Elliott, S.R. and A.D. Tischler, *Phosphate starvation: a novel signal that triggers ESX-5 secretion in Mycobacterium tuberculosis.* Mol Microbiol, 2016. **100**(3): p. 510-26.
159. Peirs, P., et al., *Mycobacterium tuberculosis with disruption in genes encoding the phosphate binding proteins PstS1 and PstS2 is deficient in phosphate uptake and demonstrates reduced in vivo virulence.* Infect Immun, 2005. **73**(3): p. 1898-902.
160. Rengarajan, J., B.R. Bloom, and E. Rubin, *Genome-wide requirements for Mycobacterim tuberculosis adaptation and survival in macrophages.* Proc Natl Acad Sci U S A, 2005. **102**(23): p. 8327-8332.
161. Tischler, A.D., et al., *Mycobacterium tuberculosis Phosphate Uptake System Component PstA2 Is Not Required for Gene Regulation or Virulence.* PLoS One, 2016. **11**(8): p. e0161467.
162. Wagner, D., et al., *Elemental analysis of the Mycobacterium avium phagosome in Balb/c mouse macrophages.* Biochem Biophys Res Commun, 2006. **344**(4): p. 1346-51.
163. Wagner, D., et al., *Elemental analysis of Mycobacterium avium-, Mycobacterium tuberculosis-, and Mycobacterium smegmatis-containing phagosomes indicates pathogen-induced microenvironments within the host cell's endosomal system.* J Immunol, 2005. **174**(3): p. 1491-1500.
164. Ligon, L.S., J.D. Hayden, and M. Braunstein, *The ins and outs of Mycobacterium tuberculosis protein export.* Tuberculosis (Edinb), 2012. **92**: p. 121-132.
165. Stoop, E.J.M., W. Bitter, and A.M. van der Sar, *Tubercle bacilli rely on a type VII army for pathogenicity.* Trends Microbiol., 2012. **20**: p. 477-484.
166. Houben, E.N.G., K.V. Korotkov, and W. Bitter, *Take five - Type VII secretion systems of Mycobacteria.* Biochim. Biophys. Acta, 2014. **1843**: p. 1707-1716.
167. Cole, S.T., et al., *Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.* Nature, 1998. **393**: p. 537-544.
168. Ekiert, D.C. and J.S. Cox, *Structure of a PE-PPE-EspG complex from Mycobacterium tuberculosis reveals molecular specificity of ESX protein secretion.* Proc. Natl. Acad. Sci. USA, 2014. **111**: p. 14758-14763.
169. Daleke, M.H., et al., *Specific chaperones for the type VII protein secretion pathway.* J. Biol. Chem., 2012. **287**: p. 31939-31947.

170. Korotkova, N., et al., *Structure of the Mycobacterium tuberculosis type VII secretion system chaperone EspG5 in complex with PE25-PPE41 dimer.* Mol. Microbiol., 2014. **94**: p. 367-382.
171. Sampson, S.L., *Mycobacterial PE/PPE proteins at the host-pathogen interface.* Clin. Dev. Immunol., 2011. **2011**: p. 497203.
172. Gey van Pittius, N.C., et al., *The ESAT-6 gene cluster of Mycobacterium tuberculosis and other high G+C gram-positive bacteria.* Genome Biol., 2001. **2**: p. 0044.1-0044.18.
173. Abdallah, A.M., et al., *Mycobacterial secretion systems ESX-1 and ESX-5 play distinct roles in host cell death and inflammasome activation.* J. Immunol., 2011. **187**: p. 4744-4753.
174. Bottai, D., et al., *Disruption of the ESX-5 system of Mycobacterium tuberculosis causes loss of PPE protein secretion, reduction of cell wall integrity and strong attenuation.* Mol. Microbiol., 2012. **83**: p. 1195-1209.
175. Sayes, F., et al., *Strong immunogenicity and cross-reactivity of Mycobacterium tuberculosis ESX-5 type VII secretion-encoded PE-PPE proteins predicts vaccine potential.* Cell Host Microbe, 2012. **11**: p. 352-363.
176. Rodriguez, G.M., et al., *ideR, an essential gene in Mycobacterium tuberculosis: role of IdeR in iron-dependent gene expression, iron metabolism, and oxidative stress response.* Infect. Immun., 2002. **70**: p. 3371-3381.
177. Maciag, A., et al., *Global analysis of the Mycobacterium tuberculosis Zur (FurB) regulon.* J. Bacteriol., 2007. **189**: p. 730-740.
178. Siegrist, M.S., et al., *Mycobacterial ESX-3 is required for mycobactin-mediated iron acquisition.* Proc. Natl. Acad. Sci. USA, 2009. **106**: p. 18792-18797.
179. Serafini, A., et al., *Characterization of a Mycobacterium tuberculosis ESX-3 conditional mutant: essentiality and rescue by iron and zinc.* J. Bacteriol., 2009. **191**: p. 6340-6344.
180. Tischler, A.D., et al., *Mycobacterium tuberculosis requires phosphate-responsive gene regulation to resist host immunity.* Infect. Immun., 2013. **81**: p. 317-328.
181. Lamarche, M.G., et al., *The phosphate regulon and bacterial virulence: a regulatory network connecting phosphate homeostasis and pathogenesis.* FEMS Microbiol. Rev., 2008. **32**: p. 461-473.
182. Parish, T., et al., *The senX3-regX3 two-component regulatory system of Mycobacterium tuberculosis is required for virulence.* Microbiol., 2003. **149**: p. 1423-1435.
183. Dobos, K.M., et al., *Evidence for glycosylation sites on the 45-kilodalton glycoprotein of Mycobacterium tuberculosis.* Infect. Immun., 1995. **63**: p. 2846-2853.
184. Houben, E.N.G., et al., *Composition of the type VII secretion system membrane complex.* Mol. Microbiol., 2012. **86**: p. 472-484.
185. Houben, D., et al., *ESX-1-mediated translocation to the cytosol controls virulence of mycobacteria.* Cell Microbiol, 2012. **14**(8): p. 1287-98.

186. Hsu, T., et al., *The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue.* Proc Natl Acad Sci U S A, 2003. **100**(21): p. 12420-5.
187. Guinn, K.M., et al., *Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of Mycobacterium tuberculosis.* Mol Microbiol, 2004. **51**(2): p. 359-70.
188. Stanley, S.A., et al., *Acute infection and macrophage subversion by Mycobacterium tuberculosis require a specialized secretion system.* Proc Natl Acad Sci U S A, 2003. **100**(22): p. 13001-6.
189. McLaughlin, B., et al., *A mycobacterium ESX-1-secreted virulence factor with unique requirements for export.* PLoS Pathog., 2007. **3**: p. e105.
190. Raghavan, S., et al., *Secreted transcription factor controls Mycobacterium tuberculosis virulence.* Nature, 2008. **454**: p. 717-721.
191. Pang, X., et al., *MprAB regulates the espA operon in Mycobacterium tuberculosis and modulates ESX-1 function and host cytokine response.* J. Bacteriol., 2013. **195**: p. 66-75.
192. Epstein, W., *The roles and regulation of potassium in bacteria.* Prog. Nucleic Acid Res. Mol. Biol., 2003. **75**: p. 293-320.
193. Ramakrishnan, P., et al., *Mycobacterium tuberculosis Resists Stress by Regulating PE19 Expression.* Infect Immun, 2015. **84**(3): p. 735-46.
194. Himpens, S., C. Locht, and P. Supply, *Molecular characterization of the mycobacterial SenX3-RegX3 two-component system: evidence for autoregulation.* Microbiol., 2000. **146**: p. 3091-3098.
195. Cortes, T., et al., *Genome-wide mapping of transcriptional start sites defines an extensive leaderless transcriptome in Mycobacterium tuberculosis.* Cell Rep, 2013. **5**(4): p. 1121-31.
196. Shell, S.S., et al., *Leaderless Transcripts and Small Proteins Are Common Features of the Mycobacterial Translational Landscape.* PLoS Genet, 2015. **11**(11): p. e1005641.
197. Fortune, S.M., et al., *Mutually dependent secretion of proteins required for mycobacterial virulence.* Proc. Natl. Acad. Sci. USA, 2005. **102**: p. 10676-10681.
198. Chen, J.M., et al., *Mycobacterium tuberculosis EspB binds phospholipids and mediates EsxA-independent virulence.* Mol. Microbiol., 2013. **89**: p. 1154-1166.
199. Gonzalo-Asenio, J., et al., *PhoP: a missing piece in the intricate puzzle of Mycobacterium tuberculosis virulence.* PLoS One, 2008. **3**: p. e3496.
200. Zhang, M., et al., *EspI regulates the ESX-1 secretion system in response to ATP levels in Mycobacterium tuberculosis.* Mol. Microbiol., 2014. **93**: p. 1057-1065.
201. Rifat, D., W.R. Bishai, and P.C. Karakousis, *Phosphate depletion: a novel trigger for Mycobacterium tuberculosis persistence.* J. Infect. Dis., 2009. **200**: p. 1126-1135.
202. Russell, D.G., et al., *Mycobacterium tuberculosis wears what it eats.* Cell Host Microbe, 2010. **8**: p. 68-76.
203. Wagner, D., et al., *Elemental analysis of Mycobacterium avium-, Mycobacterium tuberculosis-, and Mycobacterium smegmatis-containing phagosomes indicates*

pathogen-induced microenvironments within the host cell's endosomal system. J. Immunol., 2005. **174**: p. 1491-1500.
204. Wagner, D., et al., *Elemental analysis of the Mycobacterium avium phagosome in Balb/c mouse macrophages.* Biochem. Biophys. Res. Commun., 2006. **344**: p. 1346-1351.
205. Russell, D.G., et al., *Foamy macrophages and the progession of the human tuberculosis granuloma.* Nat. Immunol., 2009. **10**: p. 943-948.
206. Peyron, P., et al., *Foamy macrophages from tuberculous patients' granulomas constitute a nutrient-rich reservoir for M. tuberculosis persistence.* PLoS Pathog., 2008. **4**: p. e1000204.
207. China, A., P. Tare, and V. Nagaraja, *Comparison of promoter-specific events during transcription initiation in mycobacteria.* Microbiology, 2010. **156**(Pt 7): p. 1942-52.
208. Newton-Foot, M. and N.C. Gey van Pittius, *The complex architecture of mycobacterial promoters.* Tuberculosis (Edinb), 2013. **93**(1): p. 60-74.
209. Bashyam, M., et al., *A study of mycobacterial transcriptional apparatus identification of novel features in promoter elements.* J Bacteriol, 1996. **178**: p. 4847-53.
210. Browning, D.F. and S.J. Busby, *The regulation of bacterial transcription initiation.* Nat Rev Microbiol, 2004. **2**(1): p. 57-65.
211. Cole, S.T., *Comparative and functional genomics of the Mycobacterium tuberculosis complex.* Microbiology, 2002. **148**: p. 2919-28.
212. MacMicking, J.D., G.A. Taylor, and J.D. McKinney, *Immune control of tuberculosis by IFN-gamma-inducible LRG-47.* Science, 2003. **302**(5645): p. 654-9.
213. Nathan, C. and M.U. Shiloh, *Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens.* Proc Natl Acad Sci U S A, 2000. **97**(16): p. 8841-8.
214. Rustad, T.R., et al., *Mapping and manipulating the Mycobacterium tuberculosis transcriptome using a transcription factor overexpression-derived regulatory network.* Genome Biol, 2014. **15**(11): p. 502.
215. Minch, K.J., et al., *The DNA-binding network of Mycobacterium tuberculosis.* Nat Commun, 2015. **6**: p. 5829.
216. Turkarslan, S., et al., *A comprehensive map of genome-wide gene regulation in Mycobacterium tuberculosis.* Sci Data, 2015. **2**: p. 150010.
217. Uplekar, S., et al., *Comparative genomics of Esx genes from clinical isolates of Mycobacterium tuberculosis provides evidence for gene conversion and epitope variation.* Infect Immun, 2011. **79**(10): p. 4042-9.
218. Simeone, R., et al., *Cytosolic access of Mycobacterium tuberculosis: critical impact of phagosomal acidification control and demonstration of occurrence in vivo.* PLoS Pathog, 2015. **11**(2): p. e1004650.
219. Portal-Celhay, C., et al., *Mycobacterium tuberculosis EsxH inhibits ESCRT-dependent CD4+ T-cell activation.* Nat Microbiol, 2016. **2**: p. 16232.

220. Manganelli, R., et al., *The Mycobacterium tuberculosis ECF sigma factor sigmaE role in global gene expression and survival in macrophages.* Mol Microbiol, 2001. **41**(2): p. 423-37.
221. Ehrt, S., et al., *Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor.* Nucleic Acids Res, 2005. **33**(2): p. e21.
222. Glover, R.T., et al., *The two-component regulatory system senX3-regX3 regulates phosphate-dependent gene expression in Mycobacterium smegmatis.* J Bacteriol, 2007. **189**(15): p. 5495-503.
223. Ates, L.S., et al., *Essential Role of the ESX-5 Secretion System in Outer Membrane Permeability of Pathogenic Mycobacteria.* PLoS Genet, 2015. **11**(5): p. e1005190.
224. Elliott, S.R. and A.D. Tischler, *Phosphate responsive regulation provides insights for ESX-5 function in Mycobacterium tuberculosis.* Curr Genet, 2016. **62**(4): p. 759-763.
225. Lamarche, M.G., et al., *The phosphate regulon and bacterial virulence: a regulatory network connecting phosphate homeostasis and pathogenesis.* FEMS Microbiol Rev, 2008. **32**(3): p. 461-73.
226. Lamarche, M.G., et al., *Inactivation of the pst system reduces the virulence of an avian pathogenic Escherichia coli O78 strain.* Infect Immun, 2005. **73**(7): p. 4138-45.
227. Chekabab, S.M., J. Harel, and C.M. Dozois, *Interplay between genetic regulation of phosphate homeostasis and bacterial virulence.* Virulence, 2014. **5**(8): p. 786-93.
228. Cooper, A.M. and S.A. Khader, *The role of cytokines in the initiation, expansion, and control of cellular immunity to tuberculosis.* Immunol Rev, 2008. **226**: p. 191-204.
229. Tubo, N.J. and M.K. Jenkins, *CD4+ T Cells: guardians of the phagosome.* Clin Microbiol Rev, 2014. **27**(2): p. 200-13.
230. Wiese, S., et al., *Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research.* Proteomics, 2007. **7**(3): p. 340-50.
231. Li, P., et al., *Comparative Proteomics Analysis of Human Macrophages Infected with Virulent Mycobacterium bovis.* Front Cell Infect Microbiol, 2017. **7**: p. 65.
232. Sassetti, C.M., D.H. Boyd, and E. Rubin, *Genes required for mycobacterial growth defined by high density mutagenesis.* Proc Natl Acad Sci U S A, 2003. **48**(1): p. 77-84.
233. Timm, J., et al., *Differential expression of iron-, carbon-, and oxygen-responsive mycobacterial genes in the lungs of chronically infected mice and tuberculosis patients.* Proc Natl Acad Sci U S A, 2003. **100**(24): p. 14321-6.

Appendix I

Proteomic mass spectrometry identification of novel ESX-5 substrates

The only known *Mtb* ESX-5 secreted substrates are EsxN, PPE41 and LipY [103, 105, 138]. Proteomic studies revealed that ESX-5 secretes the majority of PE and PPE proteins in the related species *M. marinum* [140, 223], though similar studies had not been conducted in *Mtb*. Given the number of proteins identified as substrates of ESX-5 in *M. marinum*, it seemed likely that many more ESX-5 secreted substrates remain to be discovered. Previous work from our lab has demonstrated that the Δ*pstA1* mutant hyper-secretes both EsxN and PPE41 [158]. Proteins hyper-secreted by the Δ*pstA1* mutant compared to WT may represent novel ESX-5 secreted substrates. However, since ESX-5 hyper-secretion is due to constitutive activation of RegX3, some proteins that are hyper-secreted by the Δ*pstA1* mutant may simply be part of the RegX3 regulon and therefore over-produced rather than being associated with ESX-5. Therefore, proteins identified via this proteomic analysis would require validation to confirm that they are secreted by ESX-5.

To discover potential novel ESX-5 secreted substrates, we utilized mass spectrometry to identify proteins hyper-secreted by the Δ*pstA1* mutant. After separation of proteins by SDS-PAGE and staining, we excised protein bands that were in higher abundance in the Δ*pstA1* mutant compared to the WT control (**Figure A.I**). The bands chosen were approximately 75, 50, 45 and 25 kD, The University of Minnesota Center for Mass Spectrometry & Proteomics performed the in-gel trypsin digest and ran the samples on the tandem mass spectrometer.

Using Scaffold software, peptide fragment data obtained from the mass spectrometry experiments were aligned to an *Mtb* proteomic sequence database. Proteins that produced the highest spectral counts, which may correlate to protein abundance in the Δ*pstA1* mutant secreted protein profile, were identified from each protein band (**Table AI.1**). Previously, Dr. Tischler conducted a microarray experiment and determined which genes were upregulated in the Δ*pstA1* mutant compared to WT *Mtb* [149]. We expect some overlap between these data sets, because some proteins hyper-secreted by the Δ*pstA1* mutant may simply be up-regulated at the transcriptional level due to constitutive activation of RegX3. Therefore, data from the mass

spectrometry experiment was cross-referenced with the list of Δ*pstA1* mutant upregulated genes (**Table AI.2**). We found several common hits between the two data sets, which validates at least some of the mass spectrometry findings.

Further work will be necessary to determine which proteins identified using mass spectrometry are truly hyper-secreted by the Δ*pstA1* mutant, and which of those are associated with ESX-5. More quantitative mass spectrometry methods could be used to facilitate this, including iTRAQ, a technique that labels proteins from different conditions, which would allow for the differentiation between proteins secreted by WT and any mutants screened. Further, using the *esx-5* RegX3 binding mutants, where secretion of ESX-5 proteins is not induced, could allow for a more direct comparison of ESX-5 secreted protein profiles between mutants and WT. If antibodies are available against proteins identified in the mass spectrometry study, we could evaluate if it is truly hyper-secreted in the Δ*pstA1* mutant relative to WT *Mtb*. Others in the lab have confirmed that LpqH, PstS1 and GlcB are hyper-secreted by the Δ*pstA1* mutant as proof of principle.

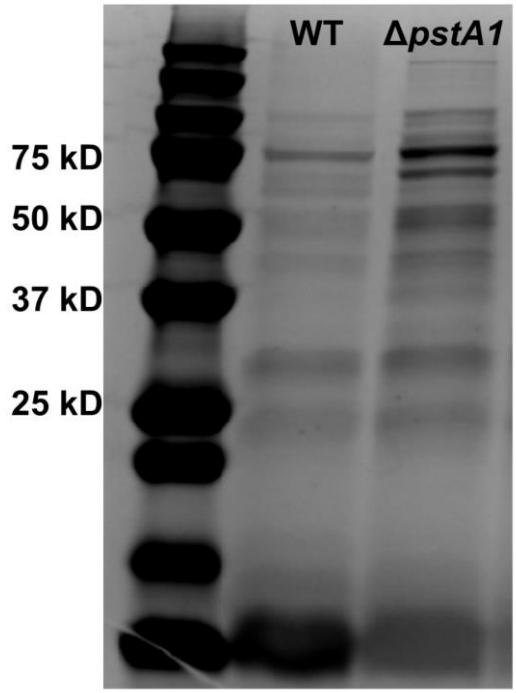

Figure A.I Secreted protein profiles of WT and Δ*pstA1 Mtb*. Wild-type *Mtb* Erdman and Δ*pstA1* were cultured in Sauton's complete medium without Tween-80 for 5 days. 5 µg proteins from the culture filtrate fraction were subjected to SDS-PAGE and Imperial protein staining. This image is representative of the secreted protein profiles.

Table AI.1 Proteins identified in the Δ*pstA1* mutant using mass spectrometry

75 kD Band		
Protein Name	**Protein size**	**Spectra Count**
5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase	85 kD	164
Malate synthase G GlcB	80 kD	106
Probable fatty oxidation protein FadB	76 kD	92
transketolase	76 kD	79
Catalase-peroxidase-peroxynitritase T KatG	81 kD	74
isocitrate dehydrogenase (NADP) Icd2	83 kD	73
isocitrate lyase	85 kD	69
molecular chaperone DnaK	67 kD	62
Probable elongation factor G FusA2 (EF-G)	76 kD	36
2-isopropylmalate synthase	70 kD	32
aminopeptidase	94 kD	30
NADH dehydrogenase subunit G	85 kD	25
prolyl oligopeptidase	74 kD	21
polyphosphate kinase	83 kD	15
phosphate acetyltransferase	73 kD	11
polynucleotide phosphorylase	80 kD	11
maltooligosyltrehalose synthase treX	81 kD	11
arylsulfatase AtsA	86 kD	10
zinc metalloprotease	74 kD	9
pyruvate dehydrogenase E1	103 kD	7
dihydrolipoamide acetyltransferase	57 kD	6
succinate dehydrogenase flavoprotein subunit	71 kD	5
hypothetical protein MRA_3441	87 kD	5
glycogen phosphorylase	96 kD	4
prolyl-tRNA synthetase	63 kD	4
molybdopterin oxidoreductase	83 kD	3
acetyl-/propionyl-coenzyme A carboxylase subunit alpha	71 kD	3
glucanase GlgE	79 kD	2
50 kD Band		
Protein Name	**Protein size**	**Spectra Count**
S-adenosyl-L-homocysteine hydrolase	54 kD	112
glutamine synthetase - GlnA	54 kD	75
dihydrolipoamide dehydrogenase	49 kD	60

cysteine synthase/cystathionine beta-synthase	49 kD	42
glucose-6-phosphate isomerase - pgi	60 kD	40
pyruvate kinase	51 kD	38
inosine 5-monophosphate dehydrogenase	50 kD	38
glutamyl-tRNA synthetase	54 kD	38
acyl-CoA dehydrogenase	59 kD	35
fumarate hydratase	50 kD	31
hypothetical protein MRA_3549	52 kD	28
Probable aldehyde dehydrogenase AldC	48 kD	28
GMP synthase GuaA	56 kD	28
succinic semialdehyde dehydrogenase	54 kD	27
Probable glycerol kinase GlpK (ATP:glycerol 3-phosphotransferase)	56 kD	24
6-phosphogluconate dehydrogenase Gnd1	52 kD	22
nitrite reductase - SirA	63 kD	21
MmsA	54 kD	20
maltokinase	52 kD	18
nitrite reductase	45 kD	17
isocitrate lyase Icl	47 kD	17
glutamate synthase subunit beta	56 kD	16
oxidoreductase	48 kD	15
hypothetical protein MRA_2244	56 kD	15
lipoprotein aminopeptidase LpqL	52 kD	14
Probable arginyl-tRNA synthetase ArgS (ARGRS) (arginine--tRNA ligase)	60 kD	12
phosphoglucomutase	58 kD	12
UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase	55 kD	11
Probable aldehyde dehydrogenase	55 kD	11
hypothetical protein MRA_2334	49 kD	10
isopropylmalate isomerase large subunit	50 kD	9
methionyl-tRNA synthetase	58 kD	9
hypothetical protein MRA_3858	58 kD	9
acetyl-CoA carboxylase carboxyl transferase subunit beta	52 kD	8
fatty-acid--CoA ligase	54 kD	7
hypothetical protein MRA_2335	55 kD	7
propionyl-CoA carboxylase subunit beta	59 kD	7
hypothetical protein MRA_1871	33 kD	7
dihydroxy-acid dehydratase	59 kD	6
acetyl-CoA carboxylase subunit beta	57 kD	6

flavoprotein	56 kD	6
Probable glutamine synthetase GlnA3 (glutamine synthase)	47 kD	3
40 kD Band		
Protein Name	**Protein size**	**Spectra Count**
Probable citrate synthase I GltA2	48 kD	75
acetyl-CoA acetyltransferase	43 kD	63
transaldolase Tal	41 kD	55
3-ketoacyl-ACP reductase	47 kD	47
acetyl-CoA acetyltransferase	40 kD	37
hypothetical protein MRA_2155	48 kD	34
gamma-glutamyl phosphate reductase	44 kD	30
hypothetical protein MRA_3759	47 kD	29
4-aminobutyrate aminotransferase	47 kD	29
Rieske 2Fe-2S family protein	43 kD	26
phosphopyruvate hydratase	45 kD	24
phosphoserine aminotransferase	40 kD	23
cysteine desulfurase Csd	45 kD	22
arginine deiminase	43 kD	22
acetyl-CoA acetyltransferase	42 kD	21
lipoprotein LppF	45 kD	19
hypothetical protein MRA_3202	42 kD	19
ferredoxin reductase	44 kD	18
isocitrate dehydrogenase	46 kD	17
hypothetical protein MRA_2442	44 kD	15
acetylornithine aminotransferase	41 kD	14
cystathionine gamma-synthase	41 kD	13
hypothetical protein	40 kD	13
phosphoribosylaminoimidazole carboxylase ATPase subunit	46 kD	12
3-oxoacyl-ACP synthase	43 kD	12
hydrolase MutT1	35 kD	11
hypothetical protein MRA_0280	41 kD	10
3-hydroxyisobutyryl-CoA hydrolase	36 kD	10
phosphoglycerate kinase	43 kD	10
alanine dehydrogenase	39 kD	10
dihydroorotase	45 kD	9
Methoxy mycolic acid synthase 4 MmaA4 (methyl mycolic acid synthase 4) (MMA4) (hydroxy mycolic acid synthase)	35 kD	9

hypothetical protein MRA_3371	48 kD	9
delta-aminolevulinic acid dehydratase	35 kD	9
O-acetylhomoserine aminocarboxypropyltransferase	47 kD	9
myo-inositol-1-phosphate synthase Ino1	40 kD	8
Probable citrate synthase II CitA	40 kD	8
ornithine carbamoyltransferase	33 kD	7
phosphoribosylamine--glycine ligase	44 kD	7
3-isopropylmalate dehydrogenase	35 kD	7
acyl-CoA dehydrogenase	44 kD	7
cyclopropane-fatty-acyl-phospholipid synthase 2	35 kD	7
phosphate ABC transporter substrate-binding protein PstS3	38 kD	6
lipid-transfer protein	42 kD	6
adhE2 product	38 kD	6
aminopeptidase 2	46 kD	6
bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase	44 kD	6
6-phosphogluconate dehydrogenase	36 kD	6
adenylosuccinate synthetase	47 kD	6
alpha-methylacyl-CoA racemase Mcr	39 kD	5
diaminopimelate decarboxylase	47 kD	4
succinyl-diaminopimelate desuccinylase	37 kD	4
hypothetical protein TBFG_13762	39 kD	3
6-phosphofructokinase	37 kD	3
phosphate ABC transporter substrate-binding protein PstS1	38 kD	3
hypothetical protein MRA_1268	41 kD	3
cytochrome p450 125 CYP125	48 kD	2
phosphoserine phosphatase	43 kD	2
hypothetical protein MRA_0076	43 kD	2

25 kD Band		
Protein Name	**Protein size**	**Spectra Count**
hypothetical protein MRA_1991	25 kD	46
superoxide dismutase - SodA	23 kD	37
enoyl-CoA hydratase	24 kD	27
lipoprotein LppX	24 kD	25
iron-regulated peptidyl-prolyl cis-trans isomerase A	19 kD	21

adenylate kinase	20 kD	21
bacterioferritin	18 kD	19
trk system potassium uptake protein ceoB	24 kD	18
ferritin	20 kD	17
enoyl-CoA hydratase	27 kD	16
ribosome recycling factor	21 kD	15
hypothetical protein MRA_2586	24 kD	15
hypothetical protein MRA_3793	22 kD	15
oxidoreductase	24 kD	14
hypothetical protein MRA_2704	26 kD	14
enoyl-CoA hydratase	27 kD	13
inorganic pyrophosphatase	18 kD	13
acid phosphatase	25 kD	12
thioredoxin protein	23 kD	11
hypothetical protein MRA_0525	25 kD	11
antigen Cfp29	29 kD	11
bifunctional ornithine acetyltransferase/N-acetylglutamate synthase	41 kD	11
AhpD	19 kD	11
dihydrodipicolinate reductase	26 kD	10
hypothetical protein MRA_1839	17 kD	9
phosphoglyceromutase	27 kD	9
two-component system transcriptional regulator	23 kD	9
pyridoxamine 5'-phosphate oxidase	25 kD	9
hypothetical protein MRA_1742	19 kD	9
two component response transcriptional regulatory protein PrrA	25 kD	8
hypothetical protein MRA_1937	17 kD	7
hypothetical protein MRA_3245	31 kD	7
Secreted antigen 85-B FbpB (85B) (antigen 85 complex B) (mycolyl transferase 85B)	35 kD	7
lipoprotein LprA	25 kD	7
hypothetical protein MRA_3741	22 kD	7
hypothetical protein MRA_0796	34 kD	7
lipoprotein LprG	25 kD	7
Secreted antigen 85-a FbpA (mycolyl transferase 85A) (fibronectin-binding protein A) (antigen 85 complex A)	36 kD	7
chorismate mutase	22 kD	6
hypothetical protein MRA_2154	19 kD	6

antigen Mpt70	19 kD	6
lipoprotein LpqH	15 kD	6
hypothetical protein MRA_3068	24 kD	5
hypothetical protein MRA_1322	21 kD	5
ferredoxin FdxC	12 kD	5
Probable lipoprotein LppC	20 kD	5
hypothetical protein TBHG_00776	26 kD	4
secreted mpt51/mpb51 antigen protein FbpD	31 kD	4
hypothetical protein MRA_1386	23 kD	4

Table AI.2 Proteins identified in the Δ*pstA1* mutant in both the mass spectrometry study and previous microarray experiments

Protein	Protein size	Spectra Count	Function
LppF	45 kd	19	Probable conserved lipoprotein
PstS3	38 kd	6	phosphate ABC transporter substrate-binding protein
AhpD	19 kd	11	Involved in oxidative stress response

Appendix II

Failed attempts to generate a Δ$espG_5$ strain

$EspG_5$ is an ESX-5 cytosolic protein chaperone that directs cognate PE and PPE proteins to the ESX-5 secretion machinery for export [127]. We sought to generate an *$espG_5$* deletion strain for several reasons. An *$espG_5$* null strain would be deficient in secretion of any proteins that require this chaperone for export, so would aid our efforts to identify novel ESX-5 secreted substrates. Comparing proteins secreted by the Δ*$espG_5$* strain to WT secreted proteins would allow us to assign any proteins missing in the *$espG_5$* null strain to the ESX-5 system. Another avenue of active investigation in our lab is to understand mechanisms causing Δ*pstA1* mutant attenuation *in vivo*. One hypothesis, which this thesis has supported, is that aberrant ESX-5 hyper-secretion sensitizes Δ*pstA1* bacteria to host defenses. Using a Δ*pstA1*Δ*$espG_5$* mutant, we would be able to assess whether the elimination of a subset of ESX-5 secreted effectors that require the $EspG_5$ chaperone can reverse the virulence defect seen in the Δ*pstA1* mutant.

However, despite years of effort, neither a Δ*$espG_5$* or Δ*pstA1*Δ*$espG_5$* mutant was obtained. The deletion construct was designed to keep the *$espG_5$* start and stop codons intact, to minimize effects on surrounding genes. Though strains in which the allelic exchange vector integrated at the *$espG_5$* locus were readily obtained, all colonies screened for the second recombination event to excise the vector possessed the WT genotype. Through several rounds of transformation, and counter-selection, a total of 333 colonies were screened, to no avail. Upon further examination of the *$espG_5$* genomic region, we found a potential ORF that overlaps with *$espG_5$* at the 3' end that was disrupted by our initial deletion construct. We redesigned the deletion construct to account for the potential alternative start codon of the putative ORF overlapping *$espG_5$*. However, even strains with the newly designed allelic exchange vector integrated at the *$espG_5$* locus were not recovered, and so this attempt was also unsuccessful. Previous screens defining *Mtb* essential genes have produced conflicting results concerning *$espG_5$*, and seemingly depending on how the data are analyzed, the gene is either essential or non-essential [232]. The multitude of failed attempts to delete *$espG_5$* strongly

suggests this gene is essential. Future endeavors to delete this gene, if desired, will require a conditional deletion approach.

Appendix III

Timing of esx-5 expression in vivo

While we have uncovered the regulatory mechanism mediating ESX-5 activity, and described a nutritional signal that activates the system, it is still unknown when *esx-5* is activated during infection. Gaining an understanding of when ESX-5 is activated *in vivo* could help shed light on the biological function of this system. We infected C57BL/6 mice with WT Erdman *Mtb* and monitored *esx-5* gene expression over time. As a control we also monitored expression of *pckA*, which is upregulated in the lungs of mice after 4 weeks of infection, and serves as an internal control [233]. However, our first attempt at this experiment yielded inconclusive results. We normalized *pe19* and $espG_5$ transcripts to either *sigA* or *16S* (**Figure A-III**). Normalization of the transcripts to *16S* yielded overall smaller transcript numbers than *sigA*, though there is no discernable pattern of gene expression of any genes tested. The data shown are the means of transcripts detected in individual mice, and there was a wide range of variability among mice in the same group. While we do observe an increase in *pckA* transcript abundance at later time points, which is consistent with published results, this pattern is only seen when normalized to *sigA* (**Figure A-III)**. When *esx-5* transcripts were normalized to *16S*, they seem to be in higher abundance at the earlier time points, though this trend isn't recapitulated when normalized to *sigA*. Overall, no conclusions can be drawn from this experiment. It is possible we have missed the window of *esx-5* activation using the mouse infection model. Changes in bacterial gene expression can occur within hours of infection. Since we infect with only ~100 CFU, and *Mtb* has a doubling time close to 20 hours, 1 week is the earliest time point that allows for sufficient bacterial RNA required for downstream analysis. Ultimately, determination of the timing of *esx-5* induction *in vivo* will require preliminary work in macrophages, which are more amenable to the restrictions of our model. There is inherently less eukaryotic nucleic acid present when using the macrophage infection model. Further, the timeline for macrophage infections are on the order of days, as opposed to months with murine infection models, which will allow for earlier time points to be examined using macrophage cell lines.

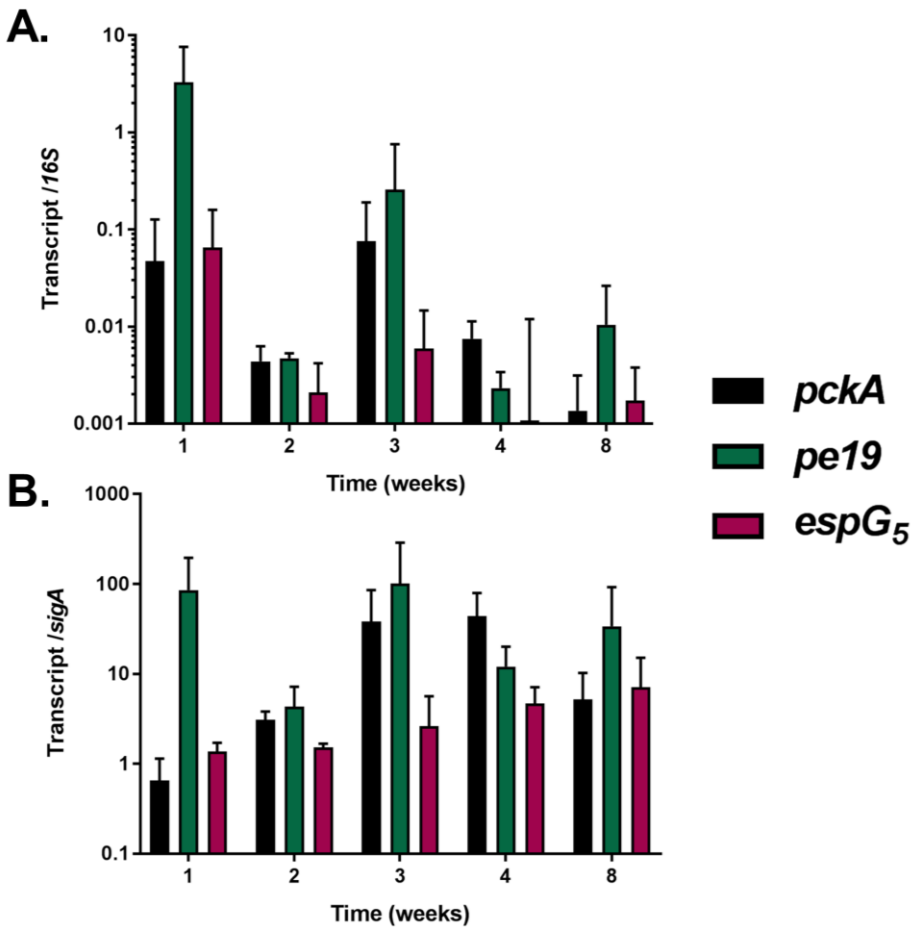

Figure A.III *In vivo esx-5* gene expression in C57BL/6 mice. Quantitative RT-PCR analysis of transcription. RNA was extracted from the lungs of WT Erdman *Mtb* infected C57BL/6 mice. Abundance of *pckA, pe19* and *espG*$_5$ were quantified using qRT-PCR and normalized to either *sigA* or *16S*. Data shown are the means of 4 mice, error bars represent the standard deviation. The data shown here are from a single experiment.

Appendix IV

Generating an esx-5 Tet-inducible strain in WT and ΔpstA1 Mtb

Our lab has provided conclusive evidence that the Pst/SenX3-RegX3 system directly regulates *esx-5* transcription in response to P_i limitation [158]. Much of our work has been conducted in the Δ*pstA1* mutant, which exhibits constitutively activated RegX3. We would like to activate expression of *esx-5* genes and secretion of ESX-5 substrates, independent of other transcriptional changes mediated by RegX3 to assess the role of ESX-5 in *Mtb* virulence. To that end, we have constructed an *esx-5* Tet-inducible strain, in both WT and Δ*pstA1* mutant backgrounds ($WT_{ESX\text{-}5\ Tet}$ and $\Delta pstA1_{ESX\text{-}5\ Tet}$). We have cloned the Tet operator and repressor from a pTIC6 derivative [222], upstream of *pe19*, located within the *esx-5* locus. In the resulting strains, the known RegX3 regulatory regions within the *ppe27-pe19* intergenic region, corresponding to -180 to -1 relative to the *pe19* start codon, were deleted in $WT_{ESX\text{-}5\ Tet}$ and $\Delta pstA1_{ESX\text{-}5\ Tet}$, leaving *esx-5* transcription completely under control of the Tet operon. These strains were constructed by allelic exchange and confirmed by PCR, but additional *in vitro* experiments are necessary to confirm that *esx-5* transcription and ESX-5 secretion can be induced with aTc and repressed by removal of aTc. Given that multiple ESX-5 components are required for viability, aTc was provided at a minimal concentration (50 ng/ml) to support growth. A Tet inducible system was specifically chosen to allow these mutant strains to be utilized in murine infections, since mice can be fed doxycycline to induce expression from the Tet regulated promoter. Using these strains, we can examine how induction of ESX-5 activity during infection impacts *Mtb* pathogenesis. We can also withhold aTc to test the importance of ESX-5 at various stages of infection.