Identification of Targeting Factors and Substrates of the Mycobacterial SecA2 Protein Export System

Meghan Elizabeth Feltcher

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Microbiology and Immunology, School of Medicine.

Chapel Hill
2013

Approved by:

Miriam Braunstein, PhD

Peggy Cotter, PhD

Thomas Kawula, PhD

Virginia Miller, PhD

Anthony Richardson, PhD

ABSTRACT

MEGHAN ELIZABETH FELTCHER: Identification of Targeting Elements and Substrates of the Mycobacterial SecA2 Protein Export System
(Under the direction of Miriam Braunstein)

Tuberculosis disease is a major global health crisis, as nearly one-third of the world's population is infected with *Mycobacterium tuberculosis*, resulting in 1.4 million deaths annually. The essential general Sec export pathway is the most widely conserved system for exporting proteins in bacteria. Central to Sec export is the SecA ATPase, which powers translocation of unfolded preproteins containing Sec signal peptides through the SecYEG membrane channel. Mycobacteria have two non-redundant SecA homologs: SecA1 and SecA2. While the essential SecA1 handles housekeeping export, the nonessential SecA2 exports a subset of proteins and is required for *M. tuberculosis* virulence. SecA2 is thought to function in concert with the SecA1/SecYEG machinery, but this relationship is poorly understood. Using two SecA2 substrates of the model organism *M. smegmatis*, we demonstrated that there does not appear to be SecA2-specific signal peptides, and instead it is the mature portion of the preprotein that dictates SecA2-dependent export. We also demonstrated that export of a SecA2 substrate is influenced by presence of both SecA2 and a functional twin-arginine translocation (Tat) pathway. Because the Tat system only accommodates proteins that fold in the cytoplasm, this result suggests that mycobacterial SecA2 substrates might be amenable to cytoplasmic folding. Furthermore, there may be a preprotein pool shared between the mycobacterial SecA2

and Tat export systems, suggesting that some mycobacterial Sec signal peptides might be compatible for export by the Tat pathway. We also performed a proteomic analysis of the *M. tuberculosis* Δ*secA2* mutant cell wall, which led to the identification of several candidate SecA2-dependent exported effectors that could explain the attenuation of the Δ*secA2* mutant. Additionally, our proteomic analysis revealed that export of several predicted Tat substrates is reduced in the *M. tuberculosis* Δ*secA2* mutant. We propose that SecA2 functions in export by adapting proteins with cytoplasmic folding tendencies, a property shared with Tat substrates, for export via SecYEG. Thus, our work has led to a model for mycobacterial SecA2 export where SecA2 would serve an integral role in connecting Sec- and Tat-mediated export.

ACKNOWLEDGMENTS

I would like to thank the following people that contributed to the work described in this dissertation. Lauren S. Ligon constructed and confirmed by Southern blotting both the *M. smegmatis* Δ*secA2* Δ*blaS* double mutant used in Chapter 2 and the Δ*secA2* Δ*tatC* double mutant used in Chapter 4. Lauren also tested the azide sensitivity of the Δ*secA2* Δ*tatC* mutant. Dr. Henry S. Gibbons assisted in construction of some of the signal sequence swap chimeric proteins used in Chapter 2. Dr. Harsha P. Gunawardena performed the LC-MS/MS work described in Chapter 3. Dr. Ellen Young constructed a *M. tuberculosis tatC* plasmid from which the pMF253 plasmid used in Chapter 4 was derived.

I would also like to thank Carol Teschke and members of the Braunstein lab for critical reading of Chapter 2. I also appreciate the antibodies generously provided by Dr. Douglas Young , Dr. Michael Niederweis, and Dr. Murty V. V. S. Madiraju. This work was supported by NIH AI05454.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF ABBREVIATIONS AND SYMBOLS

Δ	deletion
2D	Two-dimensional
ABC	ATP binding cassette
ADP	adenosine diphosphate
AIDS	Acquired Immune Deficiency Syndrome
ATP	adenosine triphosphate
BCG	bacillus Calmette-Guerin
CW	cell wall fraction
E.	*Escherichia*
ESX	ESAT secretion system
HA	hemagglutinin
HIV	Human Immunodeficiency Virus
hyg	hygromycin resistance gene
K	lysine
kan	kanamycin resistance gene
kDa	kilodalton
L.	*Listeria*
M.	*Mycobacterium*
Mtb	*Mycobacterium tuberculosis*
MDR	Multi-drug resistant
MEM	membrane fraction
mg	milligram

min	minute
ml	milliliter
mM	millimolar
MS	mass spectra
ng	nanogram
N-terminal	amino terminal
OD_{600}	optical density at 600 nanometers
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
R	arginine
S.	*Streptococcus*
SBP	solute-binding protein
SDS	sodium dodecyl sulfate
Sec	secretion
SOL	soluble fraction
SRP	signal recognition particle
Tat	Twin-arginine translocation
TB	tuberculosis
WCL	whole cell lysate

CHAPTER 1

INTRODUCTION[1]

Mycobacterium tuberculosis, the causative agent of tuberculosis (TB), is a serious global health threat accounting for nearly two million deaths every year (169). There is no effective mycobacterial vaccine available and increased prevalence of drug-resistant *M. tuberculosis* strains seriously undermine current therapies (148). *M. tuberculosis* is spread through aerosols released from infected individuals and inhaled bacilli are engulfed by alveolar macrophages inside the lung. Instead of being killed, *M. tuberculosis* survives within the phagosomal compartment of the macrophage, blocks phagosome maturation, and replicates intracellularly (129). Another significant feature of *M. tuberculosis* pathogenesis is the ability of the bacteria to persist long-term in a latent state in the host and later reactivate to cause disease, particularly in individuals who become immune-compromised.

The ability of *M. tuberculosis* to reside in a latent state has contributed to an ongoing HIV-TB co-epidemic, where TB is the leading killer of those infected with HIV (169). The ability of *M. tuberculosis* to survive and replicate within macrophages is

[1]Adapted for this dissertation from:
Feltcher, M. E., and M. Braunstein. 2012. Emerging themes in SecA2-mediated protein export. Nature Reviews Microbiology 10:779-789.
Feltcher, M. E., J. T. Sullivan, and M. Braunstein. 2010. Protein export systems of *Mycobacterium tuberculosis*: novel targets for drug development? Future Microbiology 5:1581-1597.

essential for *M. tuberculosis* pathogenesis, but our understanding of this process is incomplete. A better understanding of *M. tuberculosis* biology will facilitate the discovery of novel mycobacterial targets for new TB treatment strategies.The research described in this thesis is directed at elucidating the mechanisms of the SecA2-dependent protein export pathway that contributes to the success of *M. tuberculosis* as an intracellular pathogen.

Bacterial protein export

Over 20% of bacterial proteins have functions outside the cytoplasm and are exported to their proper locations by protein export systems (137). All bacteria, including mycobacteria, have systems for exporting these specific proteins out of the cytoplasm and into the cell envelope or extracellular environment, where they have roles in cell wall synthesis, nutrient acquisition, and other vital physiological processes (23, 30, 99). For bacterial pathogens like *M. tuberculosis*, some exported proteins are virulence factors that are required for modulation of the host environment to promote bacterial survival and growth.

Most exported proteins are translocated across the bacterial cytoplasmic membrane by the conserved general secretion (Sec) pathway or twin-arginine translocation (Tat) pathway (35, 97). The Sec pathway is present and essential for viability in all bacteria, while Tat systems are not universal but found in many Gram positive and Gram negative species (34). Mycobacteria possess functional Sec and Tat pathways, which are both essential for in vitro growth of *M. tuberculosis* (16, 131, 135). In addition to the conserved Sec and Tat systems, *M. tuberculosis* also has specialized protein export systems that are dedicated to exporting specific subsets of proteins. These

systems include five ESX pathways (ESX-1 through ESX-5) and the SecA2 export pathway (1, 48). Both ESX-1 and the SecA2 system are essential for *M. tuberculosis* virulence, presumably because they export effectors that modulate the host immune response or promote bacterial growth in the intracellular environment. The research described in this thesis is aimed at elucidating the mechanisms of SecA2-dependent export and how SecA2-dependent exported effectors promote intracellular *M. tuberculosis* growth.

The general Sec export pathway

Although best studied in *Escherichia coli,* all bacteria have a general secretion (Sec) pathway, which performs the bulk of protein export [for extensive reviews of Sec export, see references (97, 114)]. The Sec pathway translocates unfolded proteins through a heterotrimeric protein complex composed of the SecY, SecE, and SecG proteins (19). SecY forms the channel through which the unfolded proteins pass the cytoplasmic membrane (159). SecE is thought to stabilize an open SecY conformation necessary for translocation while SecG increases export efficiency (37, 63, 101, 102, 153, 155).

The SecYEG channel is used in two types of Sec export: post-translational and co-translational. In post-translational Sec export, proteins translocate completely across the cytoplasmic membrane through SecYEG with energy provided by the cytoplasmic SecA motor protein. In co-translational export, SecYEG works with the signal recognition particle (SRP) to insert integral membrane proteins into the cytoplasmic membrane (174) or, in some cases, translocate proteins across the cytoplasmic membrane (142). For integral membrane proteins, SRP recognizes transmembrane domains as they

emerge from the ribosome during translation and targets them as ribosome-mRNA-nascent protein complexes to FtsY for delivery to SecYEG (158). A lateral gate in SecY is thought to then open and allow passage of transmembrane domains into the membrane (41).

In post-translational Sec export, the proteins destined for translocation across the cytoplasmic membrane are synthesized as preproteins that are distinguished from the larger pool of cytoplasmic proteins by the presence of N-terminal Sec signal peptides. Sec signal peptides have a positively charged N-terminus, hydrophobic core, and polar C-terminus containing a signal peptidase cleavage site (164). In addition to the signal peptide, another requirement for Sec export is that the mature portion of the preprotein remains unfolded for competent passage through SecY. Some proteins are recognized and kept unfolded by cytoplasmic chaperones, such as SecB (6, 51), although other preproteins may be unfolded as they are translocated (3, 104).

A central component of the post-translational Sec pathway is the cytosolic SecA ATPase motor protein (77), which has a vital role in targeting and powering preprotein transport through SecYEG (19, 40). Since the discovery of SecA in 1981, Sec export has been the focus of extensive study (108). A combination of genetic, structural and biochemical studies including *in vitro* reconstitution have led to a relatively sophisticated understanding of the Sec pathway and SecA function in particular (Figure 1.1). SecA binds preproteins along a cleft that includes the SecA preprotein-crosslinking (PPXD) domain, and targets them to the SecYEG complex in the membrane through interactions with SecY and membrane phospholipids (4, 18, 53, 74, 93, 96, 113, 160). After delivery to SecYEG, the signal peptide of the preprotein inserts into SecY to stabilize an open

SecY channel conformation (58, 67, 82, 147). SecA then undergoes cycles of conformational changes coupled to ATP-binding and hydrolysis to drive preproteins through the SecY channel (40, 57, 161). Several models have been proposed to explain how SecA powers preprotein insertion through SecY (77). Nonetheless, most of these models propose that portions of SecA, including the IRA-1 (intramolecular regulator of ATP hydrolysis 1) two-helix finger, insert into SecY to promote forward preprotein motion through the channel (42, 168). During or shortly after translocation through the SecYEG channel, the signal peptide is removed. This cleavage event takes place on the periplasmic side of the membrane by one of two possible peptidases: the Type I signal peptidase (LepB) or the lipoprotein Type II signal peptidase (LspA) (110). After signal peptide cleavage, the protein folds into a mature conformation.

SecA, SecY, and SecE all have essential roles in Sec export and are consequently essential for cell viability (95, 107). The SecG component of SecYEG is not essential but increases translocation efficiency, possibly by stabilizing the SecY/E complex or assisting the conformational changes of SecA (100-102). Other non-essential membrane-bound proteins that increase Sec export efficiency are SecD, SecF and YajC (38, 100, 117).

The Sec pathway of mycobacteria

Mycobacteria have homologs of all the Sec export factors reviewed above except the SecB chaperone (Figure 1.1). However, detailed studies have focused on only a few components of the mycobacterial Sec pathway. The *M. tuberculosis* Type I signal peptidase LepB is essential for in vitro growth, which is expected given the essential role

of the general Sec pathway (109). The *M. tuberculosis* Type II signal peptidase (LspA) removes signal peptides from lipoproteins (133), but this lipoprotein processing is not required for in vitro growth. However, an *lspA* mutant of *M. tuberculosis* is attenuated in both macrophages and mice, illustrating the importance of functional lipoproteins to *M. tuberculosis* virulence (120, 133). The other components of the mycobacterial Sec pathway to receive attention are the SecA proteins. Mycobacteria are unusual in having two homologs of SecA: SecA1 and SecA2. SecA1 is the "housekeeping" SecA protein of mycobacteria responsible for exporting the majority of proteins while SecA2 is an accessory SecA, which is discussed in detail below. As is the case for housekeeping SecA proteins of other bacteria, SecA1 of mycobacteria is essential. The *secA1* gene cannot be deleted from *M. tuberculosis* or the nonpathogenic *M. smegmatis* unless an exogenous copy of *secA1* is provided (16, 61, 135). The contribution of SecA1 to protein export in mycobacteria can be examined by conditional silencing of *secA1* in *M. smegmatis* (61). Under the control of a tetracycline repressor, SecA1 depletion leads to growth inhibition and decreased Sec export as evidenced by reduced export of the cell wall porin, MspA (61, 125).

SecA2 Systems

For years, it was thought that all bacterial species had a single SecA protein (39). However, we now know that a number of bacteria possess two SecA homologs: SecA1 and SecA2. The first example of a second SecA was revealed by the *Mycobacterium tuberculosis* genome sequencing project (16, 29). It is now recognized that two SecA proteins exist in all mycobacteria and a diverse, but small, set of Gram positive bacteria

including *Listeria*, *Staphylococcus*, and some *Streptococcus* species (48, 124). In bacteria with two SecAs, the two proteins are not interchangeable and each SecA has unique functions (10, 16). SecA1 is the name given to the SecA with higher sequence similarity to the well-studied SecA proteins of *Escherichia coli* and *Bacillus subtilis*. SecA1 is essential and is responsible for canonical Sec export, as described above (16, 24, 43, 61, 125). Unlike SecA1, SecA2 is responsible for exporting a smaller set of proteins and often dispensable. Notably, proteins exported by SecA2 are linked to virulence in many bacterial pathogens including *M. tuberculosis* (17, 76, 150), *Streptococcus gordonii* (175), *Streptococcus parasanguinis* (172), *Staphylococcus aureus* (146), and *Listeria monocytogenes* (80).

Currently, there are two types of SecA2 systems known to exist. Some bacteria with a SecA2 also have an accessory SecY2 protein. As a consequence, these SecA2–SecY2 systems appear to function largely independent of the canonical Sec machinery to export a set of proteins that are highly glycosylated and incompatible with the canonical SecA1/SecYEG (12, 26). There are also SecA2-only systems, so named because they lack a SecY2 or an obvious accessory membrane channel. SecA2-only systems likely function as part of the canonical Sec pathway, utilizing SecYEG (43, 125) . Furthermore, the repertoire of proteins exported by SecA2-only systems is more diverse than that of SecA2–SecY2 systems.

While several published crystal structures for SecA (SecA1) proteins exist (134), including those of *M. tuberculosis* (141), *B. subtilis* (69) and *E. coli* (112), there is no structure available for any SecA2 protein. However, sequence alignments and structural modeling predict that most domains, including PPXD and IRA-1 mentioned above, are

conserved between SecA (SecA1) and SecA2 (Figure 1.2). In addition, all SecA2 proteins have two nucleotide-binding domains (NBD1 and NBD2) which together constitute the DEAD (Asp-Glu-Ala-Asp)-like motor domain. The motor domain contains two ATP-binding Walker boxes and is responsible for ATP hydrolysis, suggesting that SecA2 proteins are functional ATPases (69, 73, 143). In fact, SecA2 from *S. gordonii* and *M. tuberculosis* have demonstrated endogenous ATPase activity *in vitro* (10, 68). Furthermore, SecA2 ATPase activity is shown to be required for accessory SecA2 protein export in mycobacteria and *C. difficile* (43, 68, 125).

Even though SecA2 proteins of SecA2–SecY2 and SecA2-only systems likely function differently, it is interesting that all SecA2 proteins are smaller than their SecA1 counterparts due to a carboxyl-terminal domain (CTD) truncation, although the boundary of this truncation varies (Figure 1.2). In *E. coli*, portions of the SecA CTD binds phospholipids, SecB, and Zn+ (18, 45-47, 162, 180). One area of the CTD missing in all SecA2 proteins is the C-terminal linker (CTL), which lies within the preprotein-binding cleft and in *E. coli*, has been shown to influence substrate binding (53). In addition to the CTD truncation, the helical wing domain (HWD) is absent in the mycobacterial SecA2 and truncated in other SecA2 proteins. However, even in the canonical SecA1, the function of the HWD is not clear. The significance of the CTD and HWD truncations in SecA2 proteins awaits further studies.

SecA2–SecY2 systems

Bacteria with SecA2–SecY2 systems share a conserved *secA2–secY2* genomic locus, composed of a suite of similarly arranged homologous genes (Figure 1.3a). In

addition to genes encoding SecA2 and SecY2, each locus contains a gene that encodes a large serine-rich glycosylated protein that is exported by the SecA2–SecY2 system, as well as glycosylation factors that modify this substrate, and additional export machinery with unknown functions. The SecA2–SecY2 systems that are found in a subset of Gram positive species include pathogenic *Streptococcus gordonii* (9), *Streptococcus agalactiae* (91), *Streptococcus parasanguinis* (27), *Streptococcus pneumoniae* (105), and *Staphylococcus aureus* (145), although it should be noted that not all streptococcal species possess SecA2–SecY2 systems (124).

The current model of SecA2–SecY2 export suggests that these specialized systems exist to export the large serine-rich protein encoded in the *secA2–secY2* locus. The serine-rich substrates have cleavable N-terminal signal peptides that are unusually long and mature domains that are heavily glycosylated (124). Examples of experimentally confirmed SecA2–SecY2-exported substrates include GspB of *S. gordonii* (9), Fap1 of *S. parasanguinis* (27), Srr1 of *S. agalactiae* (91), and SraP of *S. aureus* (145). These substrates have roles related to bacterial adhesion to host tissues and/or biofilm formation (9, 91, 132, 146, 171, 172). Consequently many of these exported glycoproteins, and presumably their respective SecA2–SecY2 systems, are required for virulence (91, 105, 146, 175). In *S. parasanguinis*, the FimA adhesin is a second protein whose export is reported to depend on SecA2 (27). However, FimA is not a serine-rich glycoprotein and the *fimA* gene is not at the *S. parasanguinis* s*ecA2–secY2* locus. It is currently unknown whether FimA is a true SecA2–SecY2 substrate.

SecA2 is not essential for growth in bacteria with SecA2–SecY2 systems; however, SecA2 is absolutely required for export of their respective serine-rich

substrates, suggesting a lack of functional redundancy with canonical SecA1. Mutations in *secA2* abolish export of the serine-rich glycosylated substrates of *S. gordonii* (9), *S. parasanguinis* (27), *S. agalactiae* (91), and *S. aureus* (145). Another notable Gram positive pathogen with a putative SecA2–SecY2 pathway is *Bacillus anthracis*. However, the *B. anthracis secA2* is phylogenetically more distant from those of *Streptococcus* and *Staphylococcus* (124), which is also reflected by the dissimilar organization of the *B. anthracis secA2* locus (Figure 1.3a). *B. anthracis* SecA2 is required for optimal export of two glycosylated proteins, Sap and EA1, but it appears that SecY2 is not required for this export (98). Furthermore, the Bacillus SecA2–SecY2 system lacks the additional export machinery found in the other SecA2–SecY2 systems, which suggests this system may function differently than those of streptococcal and staphylococcal species.

Below, we discuss other genes in the *secA2–secY2* loci that have also been analyzed for roles in glycosylation and/or export.

Glycosylation factors of SecA2–SecY2 systems

In export-defective SecA2–SecY2 mutants, the serine-rich substrate retained in the cytoplasm is glycosylated, indicating that the protein is modified by cytoplasmic glycosylation factors prior to export (7, 27). There are two core glycosyltransferases conserved in all SecA2–SecY2 systems, GtfA and GtfB (Gtf1 and Gtf2) (Figure 1.3a). Some SecA2–SecY2 systems include additional glycosylation factors that further modify the substrate prior to export, including Gly and Nss of *S. gordonii* (151), Nss (Gtf3) and GalT1-2 of *S. parasanguinis* (173, 182), and the GtfC-GtfH proteins of *S. agalactiae* (91).

Export machinery of SecA2–SecY2 systems

In *S. gordonii* (9) and *S. aureus* (145), *secY2* mutations result in a loss of substrate export that is equivalent to the export defect exhibited by *secA2* mutations, demonstrating that SecY2 is essential for accessory SecA2–SecY2 export in these systems. However, in *S. parasanguinis* (170), deletion of *secY2* has only a modest effect on Fap1 export and the residual exported Fap1 species is incorrectly glycosylated (170). This result suggests that in the absence of SecY2 and full glycosylation, Fap1 export defaults to the canonical SecA1/SecYEG pathway. This result also suggests that in the *S. parasanguinis* SecA2–SecY2 system, export and glycosylation of Fap1 are coupled (as discussed further below).

There are additional proteins encoded by *SecA2–SecY2* loci that are referred to as accessory secretion proteins (Asps) in *S. gordonii* or glycosylation accessory proteins (Gaps) in *S. parasanguinis*. All SecA2–SecY2 systems include the Asp1, Asp2, and Asp3 proteins (Gap1-3). While Asp1 and Asp3 are primarily cytosolic proteins, Asp2 may be membrane localized (140, 178). Some organisms, including *S. gordonii*, have the additional Asp4 and Asp5 (Figure 1.3). Asp4 and Asp5 are both required for GspB export in *S. gordonii* (152) and are predicted integral membrane proteins with sequence homology to *B. subtilis* SecE (17% identity) and SecG (15% identity), respectively. This sequence homology is intriguing, albeit limited, and suggests that Asp4 and Asp5 may be accessory components of a SecY2 membrane channel. With regard to the role(s) of Asps 1-3, findings in different bacterial systems are not in complete agreement. In both *S. aureus* and *S. gordonii*, Asps 1-3 are clearly required for export of the respective serine-rich glycoproteins, SraP and GspB (145, 151). By contrast, deletion of either *gap1* or

gap3 in *S. parasanguinis* (the *asp1* and *asp3* homologs) has only a modest effect on export of Fap1, and the residual exported Fap1 protein has altered glycosylation (83, 115). On the basis of this result, the Asp1 and Asp3 homologs of *S. parasanguinis* were named Gap1 and Gap3 to reflect a proposed function in glycosylation. However, a *secY2* deletion in *S. parasanguinis* results in a phenotype similar to that of the *gap1* and *gap*3 mutants – export of an aberrantly glycosylated Fap1 (83, 115, 170). It seems highly unlikely that SecY2 would have a direct role in protein glycosylation. Additionally, the *S. gordonii* and *S. aureus* studies of Asp1-3 in SecA2–SecY2 export are compelling. It is possible that the discrepancy in *S. parasanguinis* results is because Fap1 glycosylation and export are highly coupled processes. In this case, export defects of *gap1*, *gap3* and *secY2* mutants (83, 115, 170) would indirectly affect Fap1 glycosylation to such an extent that the resulting altered Fap1 species is then compatible for export by the canonical SecA1/SecYEG system. However, at this time, a more direct role for Gaps in glycosylation cannot be ruled out. In fact, it was recently demonstrated that in addition to a role in GspB export, *S. gordonii* Asp2 is required for N-acetylglucosamine deposition on GspB (139). While currently there is no clear understanding of the function of any of the Asp proteins, a network of interactions between Asp1, Asp2, Asp3 and SecA2 has been mapped in *S. gordonii* (140) and *S. parasanguinis* (83, 181, 183). Asp3 may be a central scaffolding protein in this network as it interacts with multiple members of the SecA2–SecY2 system in *S. gordonii* including Asp1, Asp2, SecA2, and itself (140). Interestingly, the *S. gordonii* Asp2 and Asp3 proteins also bind the GspB substrate prior to its glycosylation, which suggests a possible function for these proteins in delivering the

substrate to an export/glycosylation complex (179). In *S. parasanguinis*, interactions between Gap1 (Asp1), Gap3 (Asp3), and SecA2 have also been identified (83, 181, 183).

Targeting proteins to the SecA2–SecY2 pathway

The serine-rich glycoproteins exported by SecA2–SecY2 systems have features that not only prevent their routing to the canonical SecA1/SecYEG pathway but promote their targeting to the SecA2–SecY2 pathway (Figure 1.4). The characteristic glycosylation of these exported substrates is one such element (7, 27, 91). In addition to being important for protein stability (12, 91), glycosylation of these proteins blocks their export by the canonical Sec pathway in both *S. gordonii* and *S. parasanguinis* (12, 26). For example, in the absence of *secA2*, the canonical SecA1/SecYEG pathway can export a stable, truncated GspB variant that is non-glycosylated. However, a glycosylated GspB protein cannot utilize the canonical Sec pathway and instead requires SecA2–SecY2 for export (12).

As mentioned above, SecA2–SecY2 serine-rich proteins are glycosylated in the cytoplasm prior to export. This is in contrast to many other glycosylated Sec substrates in bacteria (103, 163), as well as the analogous eukaryotic Sec pathway where glycosylation occurs only after proteins are translocated from the cytosol into the endoplasmic reticulum lumen (138). However, there is evidence that some bacterial proteins in addition to SecA2–SecY2 substrates share the unusual property of being glycosylated prior to Sec export (25, 28, 59), such as the HMW1 adhesin of *Haemophilus influenza* (59). But, there is likely a limit to the modifications that the SecA1/SecYEG system can handle, as the level and/or type of glycosylation of SecA2–SecY2 substrates appear incompatible with canonical Sec export. It will be interesting in the future to determine

the degree and structure of glycosylation modifications that are incompatible with the SecA1/SecYEG.

In addition to glycosylation, there are other features of SecA2–SecY2 substrates that dictate export by the accessory SecA2 pathway. The distinctive long signal peptides of GspB, Fap1, and presumably other SecA2–SecY2 substrates, are absolutely required for export. Furthermore, three glycine residues in the hydrophobic core of the GspB signal peptide promote SecA2–SecY2 -dependent export (8). However, these same glycine residues also act along with glycosylation to block export by the canonical SecA1 (8). The mechanisms by which these glycine residues act in preprotein targeting are currently unknown. Interestingly, these glycine residues are conserved in the signal peptides of most SecA2–SecY2 substrates.

Finally, there is also a region of approximately 20 amino acids at the start of the mature domain of GspB that is required for targeting this protein to the SecA2–SecY2 system (11). This accessory Sec transport (AST) domain can interact with SecA2 (but not SecA1), which suggests that in *S. gordonii* the AST is required for GspB docking to the accessory SecA2 system (13). It was initially hypothesized that the AST domain may interact directly with SecY2 to stabilize an open channel conformation. But it is possible that stabilization of the SecY2 channel is promoted by a high affinity interaction occurring between SecA2 and SecY2. This SecA2-SecY2 interaction only occurs when the preprotein binds to SecA2 via the AST domain in the mature domain of the substrate (11, 13). Currently, it is not known if the AST domain is a conserved feature of all SecA2–SecY2 substrates. However, the first 34 amino acids of the Fap1 mature domain are also required for SecA2-dependent export (26).

Model for SecA2–SecY2 export

The current model of SecA2–SecY2 export is as follows (Figure 1.5). The distinctive glycosylation of the serine-rich proteins of SecA2–SecY2 systems is incompatible with export via the canonical SecA1/SecYEG pathway and demands a specialized export system. In a signal peptide-dependent manner, the SecA2–SecY2 preproteins are targeted to the SecA2–SecY2 machinery (8, 12, 26). Features of the mature domain, such as the AST (11, 13), may also be involved in targeting. In addition, Asp2 and Asp3 could contribute to translocase-targeting by binding the unmodified substrate (179).

Analogous to canonical Sec export, SecA2 likely uses cycles of ATP hydrolysis to drive glycosylated preproteins through the SecY2 channel. In some bacteria, Asp4 and Asp5 may function like SecE and SecG, whereas SecA2–SecY2 systems lacking Asp4 and Asp5 could utilize the canonical SecE or SecG for export. In fact, there is some genetic evidence that SecY2 and SecG may function together in *S. aureus* (144). Also akin to canonical Sec export, experiments using a slow-folding model protein suggest that SecA2–SecY2 preproteins must remain unfolded for passage through the SecY2 channel (11).

SecA2-only systems

Compared to the SecA2–SecY2 systems, considerably less is known regarding the mechanisms of export by SecA2-only systems. SecA2-only systems are characterized by the presence of a non-redundant SecA homolog that is functionally distinct from the canonical SecA1 (SecA2) but lack a SecY2 homolog or obvious accessory membrane

channel (10, 16). The emerging model is that the SecA2 proteins of these systems work with the canonical SecA1/SecYEG translocase. Unlike *secA2–secY2* loci, there is no conservation of gene content or organization at the *secA2* genomic region for SecA2-only systems (Figure 1.3). In addition, there is a greater variety in the types of proteins exported by SecA2-only systems when compared to the category of glycosylated serine-rich proteins exported by SecA2–SecY2 systems of *Streptococcus* and *Staphylococcus*. SecA2-only systems exist in all mycobacteria, including the human pathogen *M. tuberculosis* (16), as well as some Gram positive bacteria such as *L. monocytogenes* (81), *Corynebacterium glutamicum* (24), and *Clostridium difficile* (43).

In mycobacteria (16, 17, 130) and *Listeria* (20, 62, 81, 85, 90), SecA2 is not essential for growth in liquid media but *secA2* mutants are defective in the export of specific proteins. However, *secA2* mutants of both *M. tuberculosis* and *L. monocytogenes* are attenuated for growth in infection models, indicating the importance of the respective SecA2 systems for exporting virulence factors (17, 76, 80). Additionally, *secA2* mutants of both *M. tuberculosis* and *L. monocytogenes* elicit aberrant immune responses during infection, which has led to the use of these mutants in vaccination studies (65, 66, 76, 119). By contrast, SecA2 is essential for growth of *Corynebacterium glutamicum* (24), and *Clostridium difficile* (43).

SecA2-only exported substrates

Proteins exported by SecA2-only systems have been identified in the model organism *M. smegmatis* (a nonpathogenic mycobacterial species), *M. tuberculosis*, *L. monocytogenes*, and *C. difficile*. Some of these proteins have predicted N-terminal Sec signal peptides and some do not. In *M. smegmatis* there are two lipoproteins (Ms1704 and

Ms1712) with predicted N-terminal Sec signal peptides that are exported to the cell wall in a SecA2-dependent manner (54, 125). Ms1704 and Ms1712 are homologous solute-binding proteins, specifically predicted to bind sugars as part of two ABC-type transporters (54). However, it is important to note that not all mycobacterial lipoproteins require SecA2 for export (54).

In *M. tuberculosis* and *L. monocytogenes*, several proteins are reduced in exported fractions of *secA2* mutant bacteria analyzed by 2D-PAGE, 3 and 17 respectively (17, 80). Of these proteins, only a few have been studied further and confirmed to be SecA2-dependent. In *M. tuberculosis*, one of these proteins is superoxide dismutase SodA, which notably lacks a predicted cleavable Sec signal peptide.

In *L. monocytogenes* the p60 autolysin, which is a cell wall amidase that cleaves peptidoglycan, is a confirmed SecA2 substrate (80). The gene for p60 is positioned adjacent to *secA2* in the genome (Figure 1.3); although, other SecA2-dependent proteins of *Listeria* are encoded elsewhere. An additional peptidoglycan-hydrolyzing autolysin NamA (MurA) of *Listeria* is also SecA2-dependent (80, 85, 90). However, unlike p60, NamA lacks a typical Sec signal peptide. MnSod superoxide dismutase is another protein lacking a predicted Sec signal peptide that is exported in a SecA2-dependent manner in *L. monocytogenes* (2). This particular finding parallels the SecA2-dependence of SodA export in *M. tuberculosis* and suggests that other similarities may exist between the mycobacteria and *Listeria* SecA2-only systems. Also among the list of SecA2-dependent proteins identified in *L. monocytogenes* are five predicted lipoproteins with Sec signal sequences, three of which are predicted solute-binding proteins (80, 122). One of these

solute-binding proteins is predicted to bind a sugar, similar to the Ms1704 and Ms1712 substrates of *M. smegmatis* (54, 80).

In *C. difficile* the S-layer protein SlpA, which constitutes a proteinaceous lattice structure surrounding the *Clostridium* cell, has been identified as being exported in a SecA2-dependent fashion (44) (22). SlpA is a member of a larger family of 29 clostridial cell wall proteins (Cwp) that are implicated in host-pathogen interactions (21, 44, 71, 166). CwpV is another protein shown to require SecA2 for export (43), suggesting that additional members of this Cwp family may be SecA2-dependent as well. In *C. difficile*, the *secA2* gene is adjacent to *slpA* and the larger *secA2* genomic region includes genes encoding 12 Cwps (Figure 1.3) (43). However, the gene encoding CwpV is notably located elsewhere in the genome (43). In *C. difficile*, both of the demonstrated SecA2-dependent proteins (SlpA and CwpV) contain predicted N-terminal Sec signal peptides (44).

Export machinery of SecA2-only systems

SecA2-only systems lack an obvious alternative membrane channel and accessory export factors. An attractive idea is that SecA2 works with the canonical SecA1/SecYEG machinery either through cooperation with SecA1 or by sharing SecYEG. In support of this model, depletion of the essential SecA1 protein in *M. smegmatis* abolishes export of the SecA2 substrate Ms1712 (125). The simplest interpretation of this experiment is that mycobacterial SecA2 export requires the canonical SecA1. However, it remains possible that SecA1 depletion in this experiment has an indirect effect on SecA2 export.

In mycobacteria and *C. difficile*, studies using ATPase-defective dominant negative SecA2 proteins are also consistent with a model where SecA2 works with the

canonical Sec machinery (43, 125). Dominant negative proteins often exert their effect by forming nonfunctional complexes with their normal binding partners. In mycobacteria, over-expression of the dominant negative SecA2 inhibits growth (125). This result implies an interaction between SecA2 and proteins important to an essential process, with the essential SecA1/SecYEG machinery being a leading candidate. In *C. difficile*, expression of the corresponding dominant negative SecA2 also inhibits growth and over shorter time frames is shown to impact protein export (43). Importantly, over-expression of a dominant negative SecA1 in *C. difficile* reduces export of SecA2 substrates, possibly by blocking accessibility of the SecA2 substrates to the SecYEG channel (43). However, unlike in mycobacteria, depletion of SecA1 in *C. difficile* does not influence export of SecA2 substrates (43) suggesting that in *Clostridium* SecA2 works with SecYEG but not SecA1.

There has not been a similar investigation for a relationship between SecA2 and SecA1/SecYEG in *Listeria*, but recently it was shown that secretion of the SecA2-dependent proteins p60 and NamA depends on the DivIVA protein (62). The DivIVA protein is involved in localizing proteins to the cell poles and septa of bacteria (15). Interestingly, GFP fusions to DivIVA, SecA2, p60, and NamA all localize to the septum in *Listeria* (62). Thus, it is possible that the SecA2-only system is specifically localized and DivIVA is required to either establish that localization pattern or deliver the SecA2-dependent proteins to the SecA2 machinery.

Targeting proteins to the SecA2-only pathway

The features defining exported substrates of SecA2-only systems have not yet received significant attention. One goal of this thesis research was to determine the

SecA2-targeting factors of export in the mycobacterial SecA2-only system. The two *M. smegmatis* lipoproteins that require SecA2 for export contain typical N-terminal Sec signal peptides. While these signal peptides are required for export (54), they are not specific for targeting these proteins to SecA2 (49). Therefore, there appears to be one or more features of the mature domains that determine SecA2-dependent export in mycobacteria. The mature domain of these SecA2-dependent substrates could possess post-translational modifications, as seen with the SecA2–SecY2 substrates, although there is currently no evidence for this possibility.

Another unresolved issue is whether preproteins without recognizable Sec signal peptides, like the Sod proteins of *M. tuberculosis* and *L. monocytogenes*, are true SecA2 substrates or if the export of these proteins is indirectly affected by SecA2. For example, SecA2 could export a currently unknown protein with a signal peptide that is itself required for export of proteins like Sod. Still, it is also plausible that proteins without signal peptides are recognized by SecA2 and exported directly through the SecYEG channel. In support of this model, SodA of *Rhizobium leguminosarum* is exported in a SecA-dependent manner despite lacking a recognizable signal peptide (75). However, at this time an indirect role for the Sec system in exporting SodA of *R. leguminosarum* cannot be ruled out (17, 43).

Model for SecA2-only Export

For SecA2-only systems, genetic studies support a model where SecA2 utilizes SecYEG to assist in Sec export of a certain class of preproteins (43, 125) (Figure 1.6). These SecA2 substrates must have features, which are not well understood at this time,

that make them incompatible for Sec export without the assistance of SecA2. SecA2 of both mycobacteria and *Clostridium* primarily localize to the cytoplasm, while much of SecA1 is found at the membrane in these bacteria (43, 125). Thus, SecA2 could possibly function in the cytoplasm to recognize and target for export a specific subset of proteins that are otherwise overlooked or incompatible with the canonical SecA1. Alternatively, SecA2 could serve as an alternate motor protein that is necessary for translocation of certain proteins through SecYEG. In either case, SecA2 ATPase activity is required for export.

Contribution of SecA2 export to M. tuberculosis *pathogenesis*

As previously mentioned, the SecA2 system is important to virulence. The *secA2* mutant of *M. tuberculosis* is attenuated for growth during the acute phase of mouse infection when it is thought that *M. tuberculosis* is actively proliferating in macrophages (17). Consistent with this mouse phenotype, the *secA2* mutant is also unable to grow in bone-marrow derived murine macrophages (17, 150). Because SodA is an antioxidant, the identification of SodA as a SecA2-dependent protein suggested that the role of the SecA2 system might be to protect *M. tuberculosis* against reactive oxygen intermediates (ROI) produced by macrophages. However, the *M. tuberculosis secA2* mutant is attenuated for growth in macrophages even if they are derived from *phox*$^{-/-}$ mice, which are unable to elicit an oxidative burst (76). While these results with *phox*$^{-/-}$ macrophages do not rule out a role for the SecA2 system in resisting oxidative stress, it does reveal another role for SecA2 export beyond detoxification of ROI. This result also implies that proteins other than SodA are exported by the SecA2 system of *M. tuberculosis*.

In further support of this possibility, it has been shown that SecA2 is required to block phagosome maturation, which is a vital step in allowing *M. tuberculosis* to grow inside host macrophages during infection (150). The *secA2* mutant more readily localizes to acidified phagosomal compartments than wild type *M. tuberculosis*, and chemical inhibitors of phagosome acidification rescue growth of the *secA2* mutant (150). These results suggest that SecA2 is required for exporting one or more protein effectors that block phagosome acidification. As very little is currently known about proteins involved in *M. tuberculosis* phagosomal maturation arrest, elucidation of SecA2-dependent exported proteins that contribute to this process would be extremely informative.

The SecA2-dependent export pathway of *M. tuberculosis* also has affects on host immune responses. Macrophages infected with the *M. tuberculosis* Δ*secA2* mutant produce higher levels of pro-inflammatory cytokines and exhibit more apoptosis than wild type infected macrophages (66, 76). The Δ*secA2* mutant of *M. tuberculosis* also elicits better protective immunity in mice and guinea pigs to *M. tuberculosis* challenge than vaccination with the *M. bovis* Bacille Calmette-Guérin (BCG) vaccine (66). The identification of SecA2-dependent exported proteins that contribute to these processes would also be valuable.

The Tat Export Pathway

Like the general Sec pathway, the twin-arginine translocation (Tat) pathway exports proteins with N-terminal signal peptides beyond the cytoplasmic membrane [for an extensive review of Tat export, see references (97) and (111)]. The Tat pathway is found in many Gram positive and Gram negative bacteria but, unlike the Sec pathway, is

not present in all bacteria (34). Also, the Tat system differs from the Sec system because it only exports proteins that are pre-folded in the cytoplasm. Substrates requiring export by the Tat system include those that must acquire metal co-factor insertion prior to translocation out of the cytoplasm (156), as well as hetero-oligomeric protein complexes that form before export (136). It has also been proposed that Tat export systems are important for handling export of proteins that are difficult to keep unfolded in the cytoplasm prior to export. This rationale might explain why halophilic archaea, which live in high salt environments that interfere with periplasmic protein folding, rely extensively on the Tat system for protein export of proteins folded in the cytoplasm (14, 127).

The N-terminal signal peptides of Tat preproteins are similar to Sec signal peptides (79). However, a distinguishing feature of Tat signal peptides is the presence of a pair of arginine residues that are contained within the Tat motif, R-R-X-Φ-Φ, where Φ is a hydrophobic residue. The twin-arginine pair is nearly invariant and replacement of both arginines with lysine residues abolishes Tat-dependent export (149).

The Tat export machinery consists of two core components: the TatA and TatC integral membrane proteins (126). Most Gram positive bacteria possess this minimal Tat translocase, while Gram negative and some Gram positive species (including mycobacteria and other actinomycete species) possess an additional protein, TatB (111). TatB is similar to TatA in amino acid sequence but functionally distinct. The mechanisms of Tat export are less understood than those of Sec export, but there is a growing understanding of the process (Figure 1.7) (111).

The current model is that a Tat signal peptide first targets a folded preprotein to the TatBC complex in the cytoplasmic membrane (52, 78, 87, 123, 154). With energy supplied by the proton motive force, TatA is then recruited to the TatBC complex and forms a homo-oligomeric translocase channel (94, 176). There is evidence that the size of the TatA pore can vary, which may explain how the pore can handle folded proteins of different shapes and sizes (56, 167). The preprotein is then translocated across the cytoplasmic membrane through the TatA channel and the signal peptide is removed by a Type I signal peptidase (84). Type II signal peptidases may also act on Tat precursors since some Tat signal peptides contain lipobox motifs. For example, in *Haloferax volcanii* a lipoprotein is exported by the Tat pathway (55). There is also a category of Tat substrates that become integrally-embedded in the membrane (64).

A folded conformation prior to export is not only a characteristic of Tat substrates, but is actually a requirement for Tat export. Proteins are only exported by the Tat system when conditions are favorable for cytoplasmic folding (33). Therefore, in addition to the Tat signal peptide there are features of the mature domain of Tat substrates that promote folding and thereby dictate Tat export (157). Some Tat preproteins require substrate-specific chaperones to facilitate substrate folding prior to export, such as the HyaE and HybE chaperones which assemble hydrogenase components in *E. coli* (36, 70). Because not all Tat substrates require assistance of substrate-specific chaperones, general cytoplasmic chaperones may facilitate Tat export of other substrates. The DnaK general chaperone has been shown to interact with Tat substrates and is required for export of the multi-copper oxidase, CueO Tat substrate (60, 116).

The Tat pathway is present and linked to virulence in a number of bacterial pathogens (31). In *Pseudomonas aeruginosa*, the Tat pathway exports multiple virulence factors, and a mutant defective in Tat export is attenuated in the rat model of infection (106). Two of the Tat-dependent virulence factors in *P. aeruginosa* are secreted phospholipase C enzymes (106). In *Legionella pneumophila*, Tat export is required for replication in the amoeba host as well as in macrophages (32, 128). Furthermore, the *L. pneumophila* phospholipase C enzyme also requires the Tat pathway for export (128).

The Mycobacterial Tat Pathway

The Tat pathway is functional in both *M. tuberculosis* and *M. smegmatis*. Both species contain genes encoding TatA, TatB, and TatC. Tat export is essential for growth of *M. tuberculosis*, at least under standard laboratory conditions, as shown by the inability to delete *tatA*, *tatB*, or *tatC* unless exogenous copies of the *tat* genes are provided (131). However, deletion mutants of *tatA*, *tatB*, and *tatC* can be made in *M. smegmatis*. These mutants have growth defects in vitro; nonetheless, *M. smegmatis tat* mutants can be utilized to study Tat export in mycobacteria (88, 118).

Another phenotype of *M. smegmatis tat* mutants is increased sensitivity to β-lactam antibiotics (88, 118). Because β-lactamases need to be exported to the cell wall in order to degrade β-lactams, the hypersensitivity of these *tat* mutants can be attributed to reduced export of the β-lactamase, BlaS. *M. smegmatis* BlaS has a predicted Tat signal peptide and BlaS is not exported by a *tat* mutant (88, 118). By expressing the *M. tuberculosis* β-lactamase (BlaC) in wild type *M. smegmatis* and *tat* mutants, the Tat dependence of BlaC is also established (88). Furthermore, when the RR dipeptide of the

BlaC signal peptide is changed to KK, BlaC export in *M. smegmatis* is abolished, indicating that the twin-arginine motif is required for Tat export in mycobacteria as expected (88).

In silico analysis of the *M. tuberculosis* genome using several Tat prediction programs predicts a total of 108 proteins with Tat signal peptides (86, 89). Some of these predicted Tat substrates have demonstrated or suggested roles in *M. tuberculosis* pathogenesis or essential physiologic processes. However, relying exclusively on Tat prediction programs to identify Tat substrates is risky. The current Tat prediction programs are built on Tat consensus sequences defined in bacteria other than mycobacteria. There is also little overlap in the predictions of the currently available Tat prediction programs (89). Furthermore, there is an increasing list of unusual Tat exported proteins that lack a cleavable signal peptide with a twin arginine motif, and are missed by the current programs (50).

To overcome some of the bias of Tat prediction programs, a genetic reporter approach can be used to indentify Tat exported proteins of *M. tuberculosis* (89). This approach utilizes a ‘BlaC reporter lacking its endogenous Tat signal peptide and a β-lactam-sensitive *blaC* or *blaS* mutant mycobacteria background. When the signal peptide from a Tat substrate is fused to ‘BlaC, the resulting BlaC fusion protein can be exported and confer β-lactam resistance, reporting on Tat export. Importantly, the ‘BlaC reporter only works when exported by the Tat pathway and not by the Sec pathway (88). Using the ‘BlaC reporter, eighteen *M. tuberculosis* proteins are shown to have functional Tat signal peptides (5, 86, 89). The list of proteins with proven Tat signal peptides includes two Phospholipase C proteins, PlcA and PlcB. Phospholipase C is necessary for the full

virulence of *M. tuberculosis* in mice, providing strong support for the Tat pathway contributing to *M. tuberculosis* pathogenesis (121). Another protein identified as having a functional Tat signal peptide is Rv2525c, which is suggested to have a role in infection by the demonstration of increased virulence of a *M. tuberculosis rv2525c* mutant in macrophages and mice (89, 131).

Summary

Our understanding of the mechanisms of canonical Sec export has reached an impressive level of detail (35), but by comparison our understanding of SecA2 export is limited (48). In addition to mechanistic unknowns, our understanding of how the SecA2 export pathway contributes to *M. tuberculosis* remains incomplete, as we do not fully appreciate the repertoire of proteins exported by SecA2 in this pathogen. In the following chapters, we describe experiments to determine what features define mycobacterial SecA2 substrates. We also examine a potential relationship between preproteins export by the SecA2 and Tat export systems in mycobacteria. Finally, we identify new SecA2-dependent exported proteins of *M. tuberculosis*.

In Chapter 2, we describe experiments utilizing two *M. smegmatis* SecA2 substrates, Ms1704 and Ms1712, to determine what features of these preproteins target them to SecA2 for export. Interestingly, we find that that the mature domains dictate SecA2-dependent export. But unlike the substrates of SecA2–SecY2 systems, there does not appear to be a SecA2-targeting AST domain or post-translational modification that influences preprotein routing to SecA2. Furthermore, mycobacterial SecA2 substrates

may share folding characteristics of Tat substrates, as the mature domain a SecA2 preprotein is compatible for export by the Tat pathway.

In Chapter 3, we describe a semi-quantitative comparative mass spectrometry study aimed at defining the SecA2-dependent exported proteome of *M. tuberculosis*. We identified several proteins underrepresented in the cell wall of the *M. tuberculosis* Δ*secA2* mutant, as expected for proteins that require SecA2 for their export. Interestingly, we observed that the majority of predicted solute-binding proteins of *M. tuberculosis* (the same family of proteins represented by the two known SecA2 substrates of *M. smegmatis*) are dependent on SecA2. We also identified several predicted Tat substrates under-represented in the *M. tuberculosis* Δ*secA2* mutant. Because the Tat pathway only exports proteins that are folded, this supports a model where SecA2 assists in export of a subset of proteins with cytoplasmic folding tendencies. Further, it suggests that the SecA2 and Tat systems of mycobacteria may share a common pool of preprotein substrates. Our proteomic analysis also resulted in several candidate SecA2-dependent exported effectors that could help explain the requirement of SecA2 for *M. tuberculosis* virulence.

In Chapter 4, we demonstrate that export of Ms1704 is not only dependent on SecA2 but additionally influenced by the Tat pathway of *M. smegmatis*. Or data suggest portions of the cytoplasmic pool of Ms1704 preprotein can be exported by SecA2/SecYEG while another portion can be exported independently by the Tat system. These observations suggest that there is a preprotein pool between the mycobacterial SecA2 and Tat export systems. Furthermore, the data presented in Chapter 4 reinforce a

model where the defining feature of mycobacterial SecA2 substrates is cytoplasmic folding, a property shared with preproteins of the Tat export system.

Taken together, our results reinforce the fundamental differences observed between SecA2–SecY2 and SecA2-only systems. Unlike SecA2–SecY2 systems, which appear dedicated to exporting unfolded glycosylated proteins, our data suggest that the mycobacterial SecA2-only system is adapted to handle preproteins with unique folding characteristics. Possible roles for the mycobacterial SecA2 in exporting proteins with a tendency to fold prior to export include keeping preproteins unfolded prior to or during export. Alternatively, SecA2 could provide the energy required to translocate such challenging substrates through the SecYEG channel. Our results are significant because they suggest that in addition to Tat export, SecA2-dependent export is another strategy bacteria use to solve the problem of exporting folding proteins across the cytoplasmic membrane.

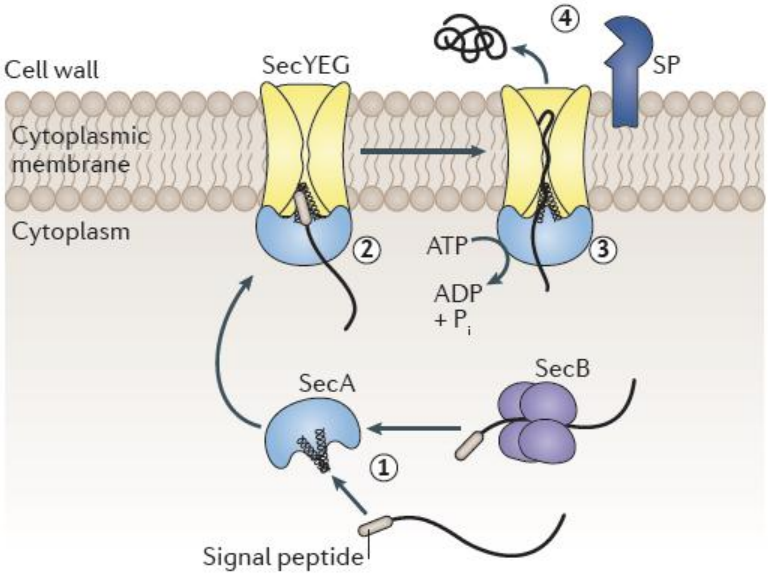

Figure 1.1 General Sec export
Post-translational Sec export is powered by the essential SecA ATPase. SecA can be divided into two main structural domains: a motor domain that drives ATP hydrolysis and a specificity domain that interacts with the preprotein destined for export. **Step 1**: Preproteins synthesized with N-terminal Sec signal peptides (grey) are bound by cytoplasmic SecA along a cleft between the two domains. Cytoplasmic chaperones, such as SecB, aid in keeping some preproteins unfolded prior to export and can directly deliver these preproteins to SecA. **Step 2:** SecA delivers the preprotein to a membrane-spanning complex composed of SecY, SecE, and SecG. Here, the signal peptide inserts into SecY to help keep an open channel conformation. **Step 3:** SecA goes through rounds of conformational changes coupled to ATP hydrolysis to promote forward movement of the unfolded preprotein through the SecY channel. During these cycles, it is proposed that SecA inserts into the SecY channel, specifically by the IRA-1 two-helix finger (two helices). **Step 4:** During or shortly after translocation the signal peptide is removed by periplasmic signal peptidases (SP) and the protein then adopts its mature, folded conformation. The SecD, SecF, and YajC transmembrane proteins contribute to the efficiency of Sec export but are not shown. After translocation, the exported protein can remain associated with the cell envelope or be fully secreted into the extracellular environment.

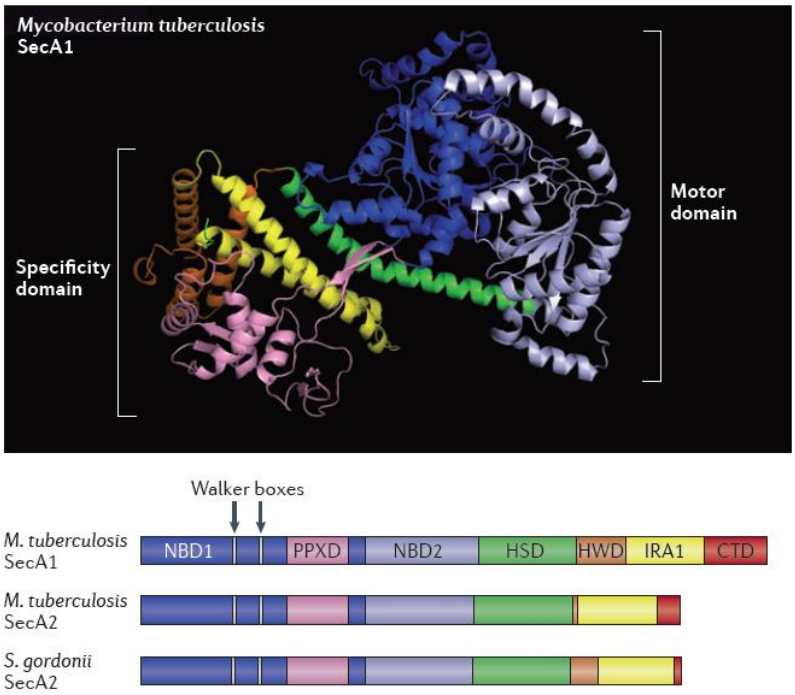

Figure 1.2 Domain organization of SecA2 proteins
Sequence alignments and structural modeling suggest that most functional domains are conserved between SecA1 and SecA2 proteins. The crystal structure of the *M. tuberculosis* SecA1 protein depicted here represents a typical SecA1 protein with the corresponding colored domains outlined below (141). *The C-terminal domain (CTD) was not resolved in the *M. tuberculosis* crystal structure but is shown in the domain graphic. For comparison to SecA1, the predicted domain organization of *M. tuberculosis* SecA2 and *Streptococcus gordonii* SecA2 are included. SecA1 can be divided into two main structural domains, which are both composed of several subdomains (77, 134). The DEAD (Asp-Glu-Ala-Asp)-like motor domain is responsible for the ATP hydrolysis (69, 73, 143) and consists of two nucleotide-binding folds: NBD1 and NBD2. NBD1 contains the two ATP-binding Walker boxes (92, 177). The helical scaffold domain (HSD) connects the motor domain with the rest of the specificity domain, to allow coupling of preprotein-binding with ATP hydrolysis (72, 165). Interactions between SecA1 and the preprotein map along a hydrophobic cleft formed by NBD1, the preprotein crosslinking domain (PPXD) (53, 96, 113), and portions of the HSD linker. Within the HSD region is a two-helix finger known as IRA-1, which is thought to assist inserting of the preprotein into the SecY channel during translocation. Conservation of the motor domain between SecA1 and SecA2 proteins confirms the observation that SecA2 proteins are ATPases (10, 68). There are differences between SecA1 and SecA2 in the specificity domain that could affect substrate specificity and/or function. All SecA2 proteins are smaller than their SecA1 counterparts due to a truncation of the CTD and helical wing domain (HWD). Although not depicted in the *M. tuberculosis* SecA1 structure, the CTD was resolved in an *E. coli* SecA structure (53) and the C-terminal linker (CTL) within the CTD lies along the hydrophobic preprotein-binding cleft.

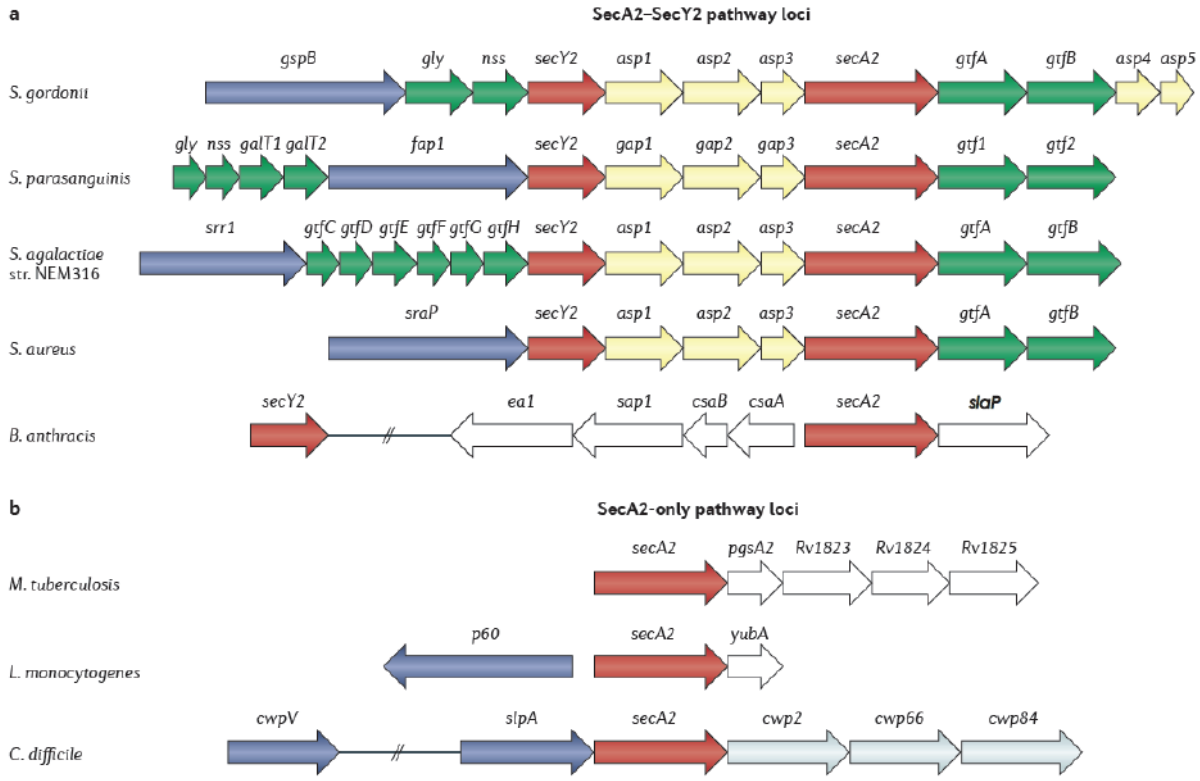

Figure 1.3 Organization of *secA2* genomic loci
a. SecA2–SecY2 systems are found in a diverse set of *Streptococcus* and *Staphylococcus* spp. (including *Streptococcus gordonii*, *Streptococcus parasanguinis*, *Streptococcus agalactiae* and *Staphylococcus aureus*), in which they function in the biogenesis of surface glycoproteins (the genes encoding which are shown in blue). All of these loci contain the core *secA2* and *secY2* genes (shown in red). Some *Bacillus* spp., including *Bacillus anthracis*, have putative SecA2–SecY2 systems, although the exported substrates of these systems are unknown, and *secA2* and *secY2* are separated. Genes encoding putative export machinery components are shown in yellow, and those encoding the glycosylation machinery are shown in green. **b.** Examples of SecA2-only systems are found in *Mycobacterium tuberculosis*, *Listeria monocytogenes* and *Clostridium difficile*. The *secA2* loci of SecA2-only systems are not conserved and export a diverse set of substrates. In some cases, the genes encoding the exported substrates (blue) are found at the *secA2* locus. However, this is not always the case, and substrate-encoding genes that are located elsewhere in the genome are not depicted. In addition to *slpA* (encoding the protein that constitutes the S-layer (surface layer) of bacteria and is a SecA2-only substrate), the *C. difficile secA2* locus contains genes encoding 11 cell wall proteins (Cwps) that are putative SecA2 substrates, of which three are shown in light blue. *asp*, accessory Sec system protein gene; *gap*, glycosylation-associated protein gene; *sraP*, serine-rich adhesin for platelets.

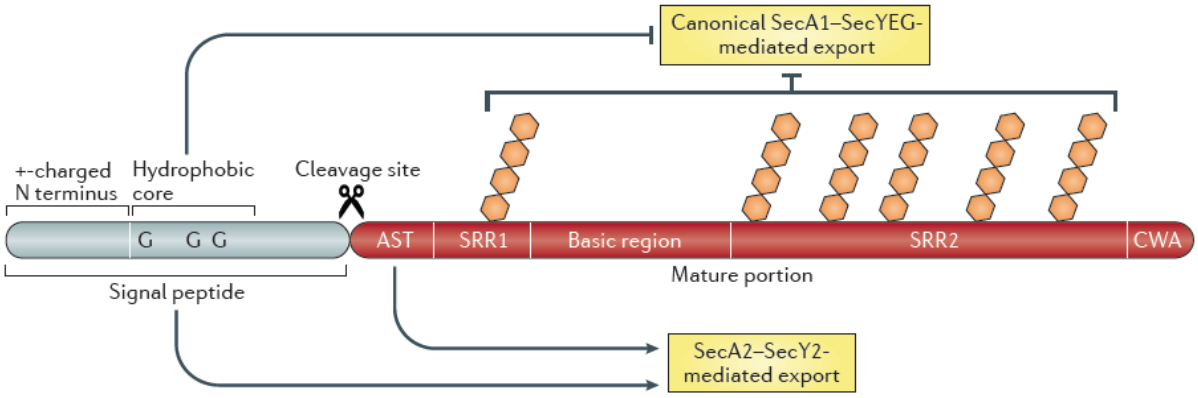

Figure 1.4 SecA2/SecY2-targeting features
SecA2–SecY2 preproteins have features both for targeting to the SecA2–SecY2 machinery for glycosylation and/or export, and for blocking export by the canonical Sec pathway (involving SecA1–SecYEG). Most of these targeting features have been defined using the *Streptococcus gordonii* protein GspB, shown here. The mature portion of GspB can be divided into several domains. Two serine-rich-repeat domains (SRR1 and SRR2) are glycosylated in the cytoplasm before export, and this post-translational modification blocks GspB export by SecA1–SecYEG. GspB, like all SecA2–SecY2 preproteins, has an unusually long amino-terminal signal peptide that is required for export. The signal peptide has the same tripartite structure as Sec signal peptides, including a positively charged N terminus, a hydrophobic core and a cleavage domain. In the hydrophobic core of the signal peptide are three glycine residues that are required for GspB export though SecA2–SecY2, but these same glycines also inhibit export by SecA1–SecYEG through unknown mechanisms. The accessory Sec transport (AST) domain is adjacent to the signal peptide and is required for export by SecA2–SecY2. Also depicted in the mature portion are the basic region and the carboxy-terminal cell wall-anchoring (CWA) domain.

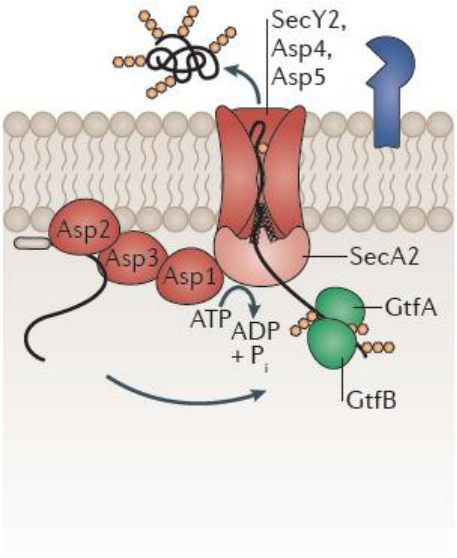

Figure 1.5 Model of SecA2–SecY2 Export
The biogenesis of surface glycoproteins by SecA2–SecY2 systems involves both glycosylation factors (green) and export machinery (red) that are distinct from the canonical Sec machinery. Serine-rich proteins are synthesized with N-terminal signal peptides. The accessory Sec system proteins (Asps) promote SecA2–SecY2-mediated export by unknown mechanisms, but could target preproteins to the translocase and/or serve as a scaffold for the export complex. Asp4 and Asp5 are putative accessory components of the SecY2 channel, but they are not present in all SecA2–SecY2 systems. SecA2–SecY2-mediated export and glycosylation are likely to be coupled processes. As glycosyl groups (orange hexagons) are added to the preprotein by cytoplasmic glycosylation factors, including the core GtfA and GtfB glycosyltransferases, SecA2 energizes transport of the unfolded preprotein through a channel formed by SecY2. SecA2–SecY2 substrates also contain a carboxy-terminal cell wall-anchoring domain, which targets the exported protein to the cell wall after SecA2–SecY2-mediated export through the membrane.

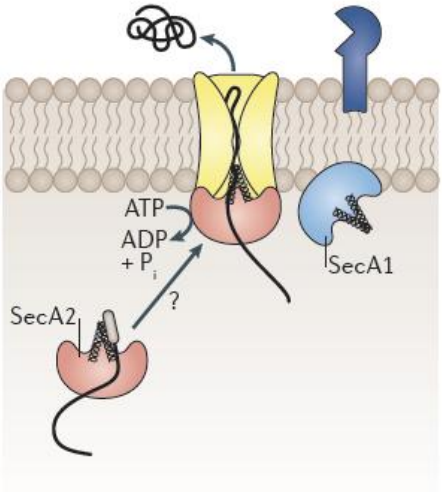

Figure 1.6 Model of SecA2-only Export
SecA2-only systems lack a SecY2 channel; therefore, SecA2 probably uses the canonical SecYEG channel for export. SecA2 could target a certain subset of preproteins to the Sec translocase that are otherwise overlooked or incompatible with SecA1. The proteins that require SecA2-only systems for transport include proteins with N-terminal Sec signal peptides (pictured here) and also proteins lacking signal peptides. SecA2 might deliver some preproteins to SecYEG and/or associate with SecYEG as needed to export substrates. The ATPase activity of SecA2 probably drives the export of preproteins through the SecYEG channel, possibly in conjunction with SecA1.

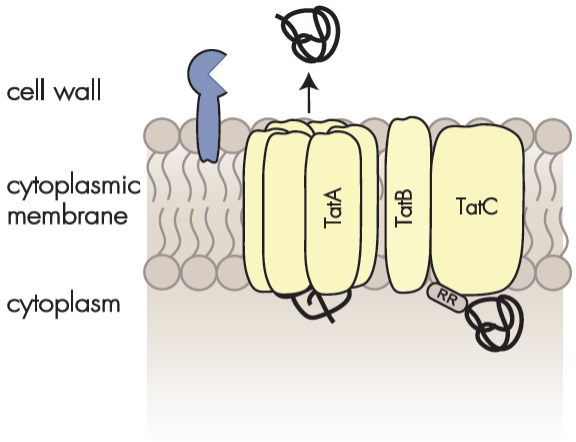

Figure 1.7 The Twin-arginine translocation (Tat) Pathway
The conserved Tat pathway exports folded preproteins containing N-terminal signal peptides with a twin-arginine (RR) motif. Tat preproteins are recognized by the TatBC complex in the membrane and then translocated across a homo-oligomeric TatA channel. As with Sec export, the signal peptide is removed by a membrane-bound signal peptidase (SP).

References

1. **Abdallah, A. M., N. C. Gey van Pittius, P. A. Champion, J. Cox, J. Luirink, C. M. Vandenbroucke-Grauls, B. J. Appelmelk, and W. Bitter.** 2007. Type VII secretion--mycobacteria show the way. Nat Rev Microbiol **5:**883-891.

2. **Archambaud, C., M. A. Nahori, J. Pizarro-Cerda, P. Cossart, and O. Dussurget.** 2006. Control of *Listeria* superoxide dismutase by phosphorylation. J Biol Chem **281:**31812-31822.

3. **Arkowitz, R. A., J. C. Joly, and W. Wickner.** 1993. Translocation can drive the unfolding of a preprotein domain. Embo J **12:**243-253.

4. **Auclair, S. M., J. P. Moses, M. Musial-Siwek, D. A. Kendall, D. B. Oliver, and I. Mukerji.** 2010. Mapping of the signal peptide-binding domain of *Escherichia coli* SecA using Forster resonance energy transfer. Biochemistry **49:**782-792.

5. **Bashiri, G., E. F. Perkowski, A. P. Turner, M. E. Feltcher, M. Braunstein, and E. N. Baker.** 2012. Tat-dependent translocation of an F420-binding protein of *Mycobacterium tuberculosis*. PLoS One **7:**e45003.

6. **Bechtluft, P., N. Nouwen, S. J. Tans, and A. J. Driessen.** 2010. SecB--a chaperone dedicated to protein translocation. Mol Biosyst **6:**620-627.

7. **Bensing, B. A., B. W. Gibson, and P. M. Sullam.** 2004. The *Streptococcus gordonii* platelet binding protein GspB undergoes glycosylation independently of export. J Bacteriol **186:**638-645.

8. **Bensing, B. A., I. R. Siboo, and P. M. Sullam.** 2007. Glycine residues in the hydrophobic core of the GspB signal sequence route export toward the accessory Sec pathway. J Bacteriol **189:**3846-3854.

9. **Bensing, B. A., and P. M. Sullam.** 2002. An accessory sec locus of *Streptococcus gordonii* is required for export of the surface protein GspB and for normal levels of binding to human platelets. Mol Microbiol **44:**1081-1094.

10. **Bensing, B. A., and P. M. Sullam.** 2009. Characterization of *Streptococcus gordonii* SecA2 as a paralogue of SecA. J Bacteriol **191:**3482-3491.

11. **Bensing, B. A., and P. M. Sullam.** 2010. Transport of preproteins by the accessory Sec system requires a specific domain adjacent to the signal peptide. J Bacteriol **192:**4223-4232.

12. **Bensing, B. A., D. Takamatsu, and P. M. Sullam.** 2005. Determinants of the streptococcal surface glycoprotein GspB that facilitate export by the accessory Sec system. Mol Microbiol **58:**1468-1481.

13. **Bensing, B. A., Y. T. Yen, R. Seepersaud, and P. M. Sullam.** 2012. A Specific interaction between SecA2 and a region of the preprotein adjacent to the signal peptide occurs during transport via the accessory Sec system. J Biol Chem **287:**24438-24447.

14. **Bolhuis, A.** 2002. Protein transport in the halophilic archaeon Halobacterium sp. NRC-1: a major role for the twin-arginine translocation pathway? Microbiology **148:**3335-3346.

15. **Bramkamp, M., and S. van Baarle.** 2009. Division site selection in rod-shaped bacteria. Curr Opin Microbiol **12:**683-688.

16. **Braunstein, M., A. M. Brown, S. Kurtz, and W. R. Jacobs, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. J Bacteriol **183:**6979-6990.

17. **Braunstein, M., B. Espinosa, J. Chan, J. T. Belisle, and W. R. J. Jacobs.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium tuberculosis*. Mol Microbiol **48:**453-464.

18. **Breukink, E., N. Nouwen, A. van Raalte, S. Mizushima, J. Tommassen, and B. de Kruijff.** 1995. The C terminus of SecA is involved in both lipid binding and SecB binding. J Biol Chem **270:**7902-7907.

19. **Brundage, L., J. P. Hendrick, E. Schiebel, A. J. Driessen, and W. Wickner.** 1990. The purified *E. coli* integral membrane protein SecY/E is sufficient for reconstitution of SecA-dependent precursor protein translocation. Cell **62:**649-657.

20. **Burkholder, K. M., K. P. Kim, K. K. Mishra, S. Medina, B. K. Hahm, H. Kim, and A. K. Bhunia.** 2009. Expression of LAP, a SecA2-dependent secretory protein, is induced under anaerobic environment. Microbes Infect **11:**859-867.

21. **Calabi, E., F. Calabi, A. D. Phillips, and N. F. Fairweather.** 2002. Binding of *Clostridium difficile* surface layer proteins to gastrointestinal tissues. Infect Immun **70:**5770-5778.

22. **Calabi, E., S. Ward, B. Wren, T. Paxton, M. Panico, H. Morris, A. Dell, G. Dougan, and N. Fairweather.** 2001. Molecular characterization of the surface layer proteins from *Clostridium difficile*. Mol Microbiol **40:**1187-1199.

23. **Calva, E., and R. Oropeza.** 2006. Two-component signal transduction systems, environmental signals, and virulence. Microb Ecol **51:**166-176.

24. **Caspers, M., and R. Freudl.** 2008. *Corynebacterium glutamicum* possesses two *secA* homologous genes that are essential for viability. Arch Microbiol.

25. **Charbonneau, M. E., and M. Mourez.** 2008. The *Escherichia coli* AIDA-I autotransporter undergoes cytoplasmic glycosylation independently of export. Res Microbiol **159:**537-544.

26. **Chen, Q., B. Sun, H. Wu, Z. Peng, and P. M. Fives-Taylor.** 2007. Differential roles of individual domains in selection of secretion route of a *Streptococcus parasanguinis* serine-rich adhesin, Fap1. J Bacteriol **189:**7610-7617.

27. **Chen, Q., H. Wu, and P. M. Fives-Taylor.** 2004. Investigating the role of *secA2* in secretion and glycosylation of a fimbrial adhesin in *Streptococcus parasanguis* FW213. Mol Microbiol **53:**843-856.

28. **Choi, K. J., S. Grass, S. Paek, J. W. St Geme, 3rd, and H. J. Yeo.** 2010. The *Actinobacillus pleuropneumoniae* HMW1C-like glycosyltransferase mediates N-linked glycosylation of the *Haemophilus influenzae* HMW1 adhesin. PLoS One **5:**e15888.

29. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M. A.**

Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead, and B. G. Barrell. 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393:**537-544.

30. **Crick, D. E., L. Quadri, and P. J. Brennan.** 2008. Biochemistry of the cell envelope of *Mycobacterium tuberculosis*, p. 1-20. *In* K. S.H.E and E. J. Rubin (ed.), Handbook of Tuberculosis: Molecular Biology and Biochemistry. WILEY-VCH Verlag GmbH & Co.

31. **De Buck, E., E. Lammertyn, and J. Anne.** 2008. The importance of the twin-arginine translocation pathway for bacterial virulence. Trends Microbiol **16:**442-453.

32. **De Buck, E., L. Maes, E. Meyen, L. Van Mellaert, N. Geukens, J. Anne, and E. Lammertyn.** 2005. *Legionella pneumophila* Philadelphia-1 tatB and tatC affect intracellular replication and biofilm formation. Biochem Biophys Res Commun **331:**1413-1420.

33. **DeLisa, M. P., D. Tullman, and G. Georgiou.** 2003. Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. Proc Natl Acad Sci U S A **100:**6115-6120.

34. **Dilks, K., R. W. Rose, E. Hartmann, and M. Pohlschroder.** 2003. Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. J Bacteriol **185:**1478-1483.

35. **du Plessis, D. J., N. Nouwen, and A. J. Driessen.** 2011. The Sec translocase. Biochim Biophys Acta **1808:**851-865.

36. **Dubini, A., and F. Sargent.** 2003. Assembly of Tat-dependent [NiFe] hydrogenases: identification of precursor-binding accessory proteins. FEBS Lett **549:**141-146.

37. **Duong, F., and W. Wickner.** 1999. The PrlA and PrlG phenotypes are caused by a loosened association among the translocase SecYEG subunits. EMBO J **18:**3263-3270.

38. **Duong, F., and W. Wickner.** 1997. The SecDFyajC domain of preprotein translocase controls preprotein movement by regulating SecA membrane cycling. EMBO J **16:**4871-4879.

39. **Economou, A.** 1999. Following the leader: bacterial protein export through the Sec pathway. Trends Microbiol **7:**315-320.

40. **Economou, A., and W. Wickner.** 1994. SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. Cell **78:**835-843.

41. **Egea, P. F., and R. M. Stroud.** 2010. Lateral opening of a translocon upon entry of protein suggests the mechanism of insertion into membranes. Proc Natl Acad Sci U S A **107:**17182-17187.

42. **Erlandson, K. J., S. B. Miller, Y. Nam, A. R. Osborne, J. Zimmer, and T. A. Rapoport.** 2008. A role for the two-helix finger of the SecA ATPase in protein translocation. Nature **455:**984-987.

43. **Fagan, R. P., and N. F. Fairweather.** 2011. *Clostridium difficile* has two parallel and essential Sec secretion systems. J Biol Chem **286:**27483-27493.

44. **Fagan, R. P., C. Janoir, A. Collignon, P. Mastrantonio, I. R. Poxton, and N. F. Fairweather.** 2011. A proposed nomenclature for cell wall proteins of *Clostridium difficile*. J Med Microbiol **60:**1225-1228.

45. **Fekkes, P., J. G. de Wit, A. Boorsma, R. H. Friesen, and A. J. Driessen.** 1999. Zinc stabilizes the SecB binding site of SecA. Biochemistry **38:**5111-5116.

46. **Fekkes, P., J. G. de Wit, J. P. van der Wolk, H. H. Kimsey, C. A. Kumamoto, and A. J. Driessen.** 1998. Preprotein transfer to the *Escherichia coli* translocase requires the co-operative binding of SecB and the signal sequence to SecA. Mol Microbiol **29:**1179-1190.

47. **Fekkes, P., C. van der Does, and A. J. Driessen.** 1997. The molecular chaperone SecB is released from the carboxy-terminus of SecA during initiation of precursor protein translocation. EMBO J **16:**6105-6113.

48. **Feltcher, M. E., and M. Braunstein.** 2012. Emerging themes in SecA2-mediated protein export. Nat Rev Microbiol **10:**779-789.

49. **Feltcher, M. E., H. S. Gibbons, L. S. Ligon, and M. Braunstein.** 2013. Protein export by the mycobacterial SecA2 system Is determined by the preprotein mature domain. J Bacteriol **195:**672-681.

50. **Ferrandez, Y., and G. Condemine.** 2008. Novel mechanism of outer membrane targeting of proteins in Gram-negative bacteria. Mol Microbiol **69:**1349-1357.

51. **Fisher, A. C., and M. P. DeLisa.** 2004. A little help from my friends: quality control of presecretory proteins in bacteria. J Bacteriol **186:**7467-7473.

52. **Frobel, J., P. Rose, F. Lausberg, A. S. Blummel, R. Freudl, and M. Muller.** 2012. Transmembrane insertion of twin-arginine signal peptides is driven by TatC and regulated by TatB. Nat Commun **3:**1311.

53. **Gelis, I., A. M. Bonvin, D. Keramisanou, M. Koukaki, G. Gouridis, S. Karamanou, A. Economou, and C. G. Kalodimos.** 2007. Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. Cell **131:**756-769.

54. **Gibbons, H. S., F. Wolschendorf, M. Abshire, M. Niederweis, and M. Braunstein.** 2007. Identification of two *Mycobacterium smegmatis* lipoproteins exported by a SecA2-dependent pathway. J Bacteriol **189:**5090-5100.

55. **Gimenez, M. I., K. Dilks, and M. Pohlschroder.** 2007. *Haloferax volcanii* twin-arginine translocation substates include secreted soluble, C-terminally anchored and lipoproteins. Mol Microbiol **66:**1597-1606.

56. **Gohlke, U., L. Pullan, C. A. McDevitt, I. Porcelli, E. de Leeuw, T. Palmer, H. R. Saibil, and B. C. Berks.** 2005. The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. Proc Natl Acad Sci U S A **102:**10482-10486.

57. **Gold, V. A., S. Whitehouse, A. Robson, and I. Collinson.** 2013. The dynamic action of SecA during the initiation of protein translocation. Biochem J **449:**695-705.

58. **Gouridis, G., S. Karamanou, I. Gelis, C. G. Kalodimos, and A. Economou.** 2009. Signal peptides are allosteric activators of the protein translocase. Nature **462:**363-367.

59. **Grass, S., A. Z. Buscher, W. E. Swords, M. A. Apicella, S. J. Barenkamp, N. Ozchlewski, and J. W. St Geme, 3rd.** 2003. The *Haemophilus influenzae* HMW1 adhesin is glycosylated in a process that requires HMW1C and phosphoglucomutase, an enzyme involved in lipooligosaccharide biosynthesis. Mol Microbiol **48:**737-751.

60. **Graubner, W., A. Schierhorn, and T. Bruser.** 2007. DnaK plays a pivotal role in Tat targeting of CueO and functions beside SlyD as a general Tat signal binding chaperone. J Biol Chem **282:**7116-7124.

61. **Guo, X. V., M. Monteleone, M. Klotzsche, A. Kamionka, W. Hillen, M. Braunstein, S. Ehrt, and D. Schnappinger.** 2007. Silencing *Mycobacterium smegmatis* by using tetracycline repressors. J Bacteriol **189:**4614-4623.

62. **Halbedel, S., B. Hahn, R. A. Daniel, and A. Flieger.** 2012. DivIVA affects secretion of virulence-related autolysins in *Listeria monocytogenes*. Mol Microbiol **83:**821-839.

63. **Harris, C. R., and T. J. Silhavy.** 1999. Mapping an interface of SecY (PrlA) and SecE (PrlG) by using synthetic phenotypes and in vivo cross-linking. J Bacteriol **181:**3438-3444.

64. **Hatzixanthis, K., T. Palmer, and F. Sargent.** 2003. A subset of bacterial inner membrane proteins integrated by the twin-arginine translocase. Mol Microbiol **49:**1377-1390.

65. **Hinchey, J., B. Y. Jeon, H. Alley, B. Chen, M. Goldberg, S. Derrick, S. Morris, W. R. Jacobs, Jr., S. A. Porcelli, and S. Lee.** 2011. Lysine auxotrophy combined with deletion of the *secA2* gene results in a safe and highly immunogenic candidate live attenuated vaccine for tuberculosis. PLoS One **6:**e15857.

66. **Hinchey, J., S. Lee, B. Y. Jeon, R. J. Basaraba, M. M. Venkataswamy, B. Chen, J. Chan, M. Braunstein, I. M. Orme, S. C. Derrick, S. L. Morris, W. R. Jacobs, Jr., and S. A. Porcelli.** 2007. Enhanced priming of adaptive immunity by a proapoptotic mutant of Mycobacterium tuberculosis. J Clin Invest **117:**2279-2288.

67. **Hizlan, D., A. Robson, S. Whitehouse, V. A. Gold, J. Vonck, D. Mills, W. Kuhlbrandt, and I. Collinson.** 2012. Structure of the SecY complex unlocked by a preprotein mimic. Cell Rep **1:**21-28.

68. **Hou, J. M., N. G. D'Lima, N. W. Rigel, H. S. Gibbons, J. R. McCann, M. Braunstein, and C. M. Teschke.** 2008. ATPase activity of *Mycobacterium tuberculosis* SecA1 and SecA2 proteins and its importance for SecA2 function in macrophages. J Bacteriol **190:**4880-4887.

69. **Hunt, J. F., S. Weinkauf, L. Henry, J. J. Fak, P. McNicholas, D. B. Oliver, and J. Deisenhofer.** 2002. Nucleotide control of interdomain interactions in the conformational reaction cycle of SecA. Science **297:**2018-2026.

70. **Jack, R. L., G. Buchanan, A. Dubini, K. Hatzixanthis, T. Palmer, and F. Sargent.** 2004. Coordinating assembly and export of complex bacterial proteins. Embo J **23:**3962-3972.

71. **Janoir, C., S. Pechine, C. Grosdidier, and A. Collignon.** 2007. Cwp84, a surface-associated protein of *Clostridium difficile*, is a cysteine protease with degrading activity on extracellular matrix proteins. J Bacteriol **189:**7174-7180.

72. **Karamanou, S., E. Vrontou, G. Sianidis, C. Baud, T. Roos, A. Kuhn, A. S. Politou, and A. Economou.** 1999. A molecular switch in SecA protein couples ATP hydrolysis to protein translocation. Mol Microbiol **34:**1133-1145.

73. **Keramisanou, D., N. Biris, I. Gelis, G. Sianidis, S. Karamanou, A. Economou, and C. G. Kalodimos.** 2006. Disorder-order folding transitions underlie catalysis in the helicase motor of SecA. Nat Struct Mol Biol **13:**594-602.

74. **Kimura, E., M. Akita, S. Matsuyama, and S. Mizushima.** 1991. Determination of a region in SecA that interacts with presecretory proteins in *Escherichia coli*. J Biol Chem **266:**6600-6606.

75. **Krehenbrink, M., A. Edwards, and J. A. Downie.** 2011. The superoxide dismutase SodA is targeted to the periplasm in a SecA-dependent manner by a novel mechanism. Mol Microbiol **82:**164-179.

76. **Kurtz, S., K. P. McKinnon, M. S. Runge, J. P. Ting, and M. Braunstein.** 2006. The SecA2 secretion factor of *Mycobacterium tuberculosis* promotes

growth in macrophages and inhibits the host immune response. Infect Immun **74:**6855-6864.

77. **Kusters, I., and A. J. Driessen.** 2011. SecA, a remarkable nanomachine. Cell Mol Life Sci **68:**2053-2066.

78. **Lausberg, F., S. Fleckenstein, P. Kreutzenbeck, J. Frobel, P. Rose, M. Muller, and R. Freudl.** 2012. Genetic evidence for a tight cooperation of TatB and TatC during productive recognition of twin-arginine (Tat) signal peptides in *Escherichia coli*. PLoS One **7:**e39867.

79. **Lee, P. A., D. Tullman-Ercek, and G. Georgiou.** 2006. The Bacterial Twin-Arginine Translocation Pathway. Annu Rev Microbiol **60:**373-395.

80. **Lenz, L. L., S. Mohammadi, A. Geissler, and D. A. Portnoy.** 2003. SecA2-dependent secretion of autolytic enzymes promotes *Listeria monocytogenes* pathogenesis. Proc Natl Acad Sci U S A **100:**12432-12437.

81. **Lenz, L. L., and D. A. Portnoy.** 2002. Identification of a second Listeria *secA* gene associated with protein secretion and the rough phenotype. Mol Microbiol **45:**1043-1056.

82. **Li, W., S. Schulman, D. Boyd, K. Erlandson, J. Beckwith, and T. A. Rapoport.** 2007. The plug domain of the SecY protein stabilizes the closed state of the translocation channel and maintains a membrane seal. Molecular Cell **26:**511-521.

83. **Li, Y., Y. Chen, X. Huang, M. Zhou, R. Wu, S. Dong, D. G. Pritchard, P. Fives-Taylor, and H. Wu.** 2008. A conserved domain of previously unknown function in Gap1 mediates protein-protein interaction and is required for biogenesis of a serine-rich streptococcal adhesin. Mol Microbiol **70:**1094-1104.

84. **Luke, I., J. I. Handford, T. Palmer, and F. Sargent.** 2009. Proteolytic processing of *Escherichia coli* twin-arginine signal peptides by LepB. Arch Microbiol **191:**919-925.

85. **Machata, S., T. Hain, M. Rohde, and T. Chakraborty.** 2005. Simultaneous deficiency of both MurA and p60 proteins generates a rough phenotype in Listeria monocytogenes. J Bacteriol **187:**8385-8394.

86. **Marrichi, M., L. Camacho, D. G. Russell, and M. P. DeLisa.** 2008. Genetic toggling of alkaline phosphatase folding reveals signal peptides for all major modes of transport across the inner membrane of bacteria. J Biol Chem **283:**35223-35235.

87. **McDevitt, C. A., G. Buchanan, F. Sargent, T. Palmer, and B. C. Berks.** 2006. Subunit composition and in vivo substrate-binding characteristics of *Escherichia coli* Tat protein complexes expressed at native levels. FEBS J **273:**5656-5668.

88. **McDonough, J. A., K. E. Hacker, A. R. Flores, M. S. Pavelka, Jr., and M. Braunstein.** 2005. The twin-arginine translocation pathway of *Mycobacterium smegmatis* is functional and required for the export of mycobacterial beta-lactamases. J Bacteriol **187:**7667-7679.

89. **McDonough, J. A., J. R. McCann, E. M. Tekippe, J. S. Silverman, N. W. Rigel, and M. Braunstein.** 2008. Identification of functional Tat signal sequences in *Mycobacterium tuberculosis* proteins. J Bacteriol **190:**6428-6438.

90. **Mishra, K. K., M. Mendonca, A. Aroonnual, K. M. Burkholder, and A. K. Bhunia.** 2011. Genetic organization and molecular characterization of *secA2* locus in *Listeria* species. Gene **489:**76-85.

91. **Mistou, M. Y., S. Dramsi, S. Brega, C. Poyart, and P. Trieu-Cuot.** 2009. Molecular dissection of the *secA2* locus of group B Streptococcus reveals that glycosylation of the Srr1 LPXTG protein is required for full virulence. J Bacteriol **191:**4195-4206.

92. **Mitchell, C., and D. Oliver.** 1993. Two distinct ATP-binding domains are needed to promote protein export by *Escherichia coli* SecA ATPase. Mol Microbiol **10:**483-497.

93. **Mori, H., and K. Ito.** 2006. Different modes of SecY-SecA interactions revealed by site-directed *in vivo* photo-cross-linking. Proc Natl Acad Sci U S A **103:**16159-16164.

94. **Mould, R. M., and C. Robinson.** 1991. A proton gradient is required for the transport of two lumenal oxygen-evolving proteins across the thylakoid membrane. J Biol Chem **266:**12189-12193.

95. **Murphy, C. K., E. J. Stewart, and J. Beckwith.** 1995. A double counter-selection system for the study of null alleles of essential genes in *Escherichia coli*. Gene **155:**1-7.

96. **Musial-Siwek, M., S. L. Rusch, and D. A. Kendall.** 2007. Selective photoaffinity labeling identifies the signal peptide binding domain on SecA. J Mol Biol **365:**637-648.

97. **Natale, P., T. Bruser, and A. J. Driessen.** 2008. Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. Biochim Biophys Acta **1778:**1735-1756.

98. **Nguyen-Mau, S. M., S. Y. Oh, V. J. Kern, D. M. Missiakas, and O. Schneewind.** 2012. Secretion genes as determinants of Bacillus anthracis chain length. J Bacteriol **194:**3841-3850.

99. **Niederweis, M.** 2008. Nutrient acquisition by mycobacteria. Microbiology **154:**679-692.

100. **Nishiyama, K., M. Hanada, and H. Tokuda.** 1994. Disruption of the gene encoding p12 (SecG) reveals the direct involvement and important function of SecG in the protein translocation of *Escherichia coli* at low temperature. EMBO J **13:**3272-3277.

101. **Nishiyama, K., S. Mizushima, and H. Tokuda.** 1995. Preferential interaction of Sec-G with Sec-E stabilizes an unstable Sec-E derivative in the *Escherichia coli* cytoplasmic membrane. Biochem Biophys Res Commun **217:**217-223.

102. **Nishiyama, K., T. Suzuki, and H. Tokuda.** 1996. Inversion of the membrane topology of SecG coupled with SecA-dependent preprotein translocation. Cell **85:**71-81.

103. **Nita-Lazar, M., M. Wacker, B. Schegg, S. Amber, and M. Aebi.** 2005. The N-X-S/T consensus sequence is required but not sufficient for bacterial N-linked protein glycosylation. Glycobiology **15:**361-367.

104. **Nouwen, N., G. Berrelkamp, and A. J. Driessen.** 2007. Bacterial sec-translocase unfolds and translocates a class of folded protein domains. J Mol Biol **372:**422-433.

105. **Obert, C., J. Sublett, D. Kaushal, E. Hinojosa, T. Barton, E. I. Tuomanen, and C. J. Orihuela.** 2006. Identification of a candidate *Streptococcus pneumoniae* core genome and regions of diversity correlated with invasive pneumococcal disease. Infect Immun **74:**4766-4777.

106. **Ochsner, U. A., A. Snyder, A. I. Vasil, and M. L. Vasil.** 2002. Effects of the twin-arginine translocase on secretion of virulence factors, stress response, and pathogenesis. Proc Natl Acad Sci U S A **99:**8312-8317.

107. **Oliver, D., C. Kumamoto, M. Quinlan, and J. Beckwith.** 1982. Pleiotropic mutants affecting the secretory apparatus of *Escherichia coli*. Ann Inst Pasteur Mic **A133:**105-110.

108. **Oliver, D. B., and J. Beckwith.** 1981. *E. coli* mutant pleiotropically defective in the export of secreted proteins. Cell **25:**765-772.

109. **Ollinger, J., T. O'Malley, J. Ahn, J. Odingo, and T. Parish.** 2012. Inhibition of the sole type I signal peptidase of Mycobacterium tuberculosis is bactericidal under replicating and nonreplicating conditions. J Bacteriol **194:**2614-2619.

110. **Paetzel, M., A. Karla, N. C. Strynadka, and R. E. Dalbey.** 2002. Signal peptidases. Chem Rev **102:**4549-4580.

111. **Palmer, T., and B. C. Berks.** 2012. The twin-arginine translocation (Tat) protein export pathway. Nat Rev Microbiol **10:**483-496.

112. **Papanikolau, Y., M. Papadovasilaki, R. B. Ravelli, A. A. McCarthy, S. Cusack, A. Economou, and K. Petratos.** 2007. Structure of dimeric SecA, the *Escherichia coli* preprotein translocase motor. J Mol Biol **366:**1545-1557.

113. **Papanikou, E., S. Karamanou, C. Baud, M. Frank, G. Sianidis, D. Keramisanou, C. G. Kalodimos, A. Kuhn, and A. Economou.** 2005. Identification of the preprotein binding domain of SecA. J Biol Chem **280:**43209-43217.

114. **Papanikou, E., S. Karamanou, and A. Economou.** 2007. Bacterial protein secretion through the translocase nanomachine. Nat Rev Microbiol **5:**839-851.

115. **Peng, Z., H. Wu, T. Ruiz, Q. Chen, M. Zhou, B. Sun, and P. Fives-Taylor.** 2008. Role of *gap3* in Fap1 glycosylation, stability, in vitro adhesion, and fimbrial and biofilm formation of *Streptococcus parasanguinis*. Oral Microbiol Immunol **23:**70-78.

116. **Perez-Rodriguez, R., A. C. Fisher, J. D. Perlmutter, M. G. Hicks, A. Chanal, C. L. Santini, L. F. Wu, T. Palmer, and M. P. Delisa.** 2007. An Essential Role for the DnaK Molecular Chaperone in Stabilizing Over-expressed Substrate Proteins of the Bacterial Twin-arginine Translocation Pathway. J Mol Biol.

117. **Pogliano, J. A., and J. Beckwith.** 1994. SecD and SecF facilitate protein export in *Escherichia coli*. EMBO J **13:**554-561.

118. **Posey, J. E., T. M. Shinnick, and F. D. Quinn.** 2006. Characterization of the twin-arginine translocase secretion system of *Mycobacterium smegmatis*. J Bacteriol **188:**1332-1340.

119. **Rahmoun, M., M. Gros, L. Campisi, D. Bassand, A. Lazzari, C. Massiera, E. Narni-Mancinelli, P. Gounon, and G. Lauvau.** 2011. Priming of protective anti-*Listeria monocytogenes* memory CD8+ T cells requires a functional SecA2 secretion system. Infect Immun **79:**2396-2403.

120. **Rampini, S. K., P. Selchow, C. Keller, S. Ehlers, E. C. Bottger, and P. Sander.** 2008. LspA inactivation in *Mycobacterium tuberculosis* results in attenuation without affecting phagosome maturation arrest. Microbiology **154:**2991-3001.

121. **Raynaud, C., C. Guilhot, J. Rauzier, Y. Bordat, V. Pelicic, R. Manganelli, I. Smith, B. Gicquel, and M. Jackson.** 2002. Phospholipases C are involved in the virulence of *Mycobacterium tuberculosis*. Mol Microbiol **45:**203-217.

122. **Renier, S., C. Chambon, D. Viala, C. Chagnot, M. Hebraud, and M. Desvaux.** 2013. Exoproteomic analysis of the SecA2-dependent secretion in *Listeria monocytogenes* EGD-e. J Proteomics **80:**183-195.

123. **Richter, S., and T. Bruser.** 2005. Targeting of unfolded PhoA to the TAT translocon of *Escherichia coli*. J Biol Chem **280:**42723-42730.

124. **Rigel, N. W., and M. Braunstein.** 2008. A new twist on an old pathway--accessory Sec systems. Mol Microbiol **69:**291-302.

125. **Rigel, N. W., H. S. Gibbons, J. R. McCann, J. A. McDonough, S. Kurtz, and M. Braunstein.** 2009. The accessory SecA2 system of mycobacteria requires ATP binding and the canonical SecA1. J Biol Chem **284:**9927-9936.

126. **Rollauer, S. E., M. J. Tarry, J. E. Graham, M. Jaaskelainen, F. Jager, S. Johnson, M. Krehenbrink, S. M. Liu, M. J. Lukey, J. Marcoux, M. A. McDowell, F. Rodriguez, P. Roversi, P. J. Stansfeld, C. V. Robinson, M. S. Sansom, T. Palmer, M. Hogbom, B. C. Berks, and S. M. Lea.** 2012. Structure of the TatC core of the twin-arginine protein transport system. Nature **492:**210-214.

127. **Rose, R. W., T. Bruser, J. C. Kissinger, and M. Pohlschroder.** 2002. Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. Mol Microbiol **45:**943-950.

128. **Rossier, O., and N. P. Cianciotto.** 2005. The Legionella pneumophila tatB gene facilitates secretion of phospholipase C, growth under iron-limiting conditions, and intracellular infection. Infect Immun **73:**2020-2032.

129. **Russell, D. G.** 2008. *Mycobacterium tuberculosis:* life and death in the phagosome, p. 307-322. *In* S. H. Kaufmann and E. J. Rubin (ed.), Handbook of Tuberculosis: Molecular Biology and Biochemistry. Wiley-VCH.

130. **Sadagopal, S., M. Braunstein, C. C. Hager, J. Wei, A. K. Daniel, M. R. Bochan, I. Crozier, N. E. Smith, H. O. Gates, L. Barnett, L. Van Kaer, J. O. Price, T. S. Blackwell, S. A. Kalams, and D. S. Kernodle.** 2009. Reducing the activity and secretion of microbial antioxidants enhances the immunogenicity of BCG. PLoS One **4:**e5531.

131. **Saint-Joanis, B., C. Demangel, M. Jackson, P. Brodin, L. Marsollier, H. Boshoff, and S. T. Cole.** 2006. Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of *Mycobacterium tuberculosis*, increases {beta}-lactam susceptibility and virulence. J Bacteriol **188:**6669-6679.

132. **Samen, U., B. J. Eikmanns, D. J. Reinscheid, and F. Borges.** 2007. The surface protein Srr-1 of *Streptococcus agalactiae* binds human keratin 4 and promotes adherence to epithelial HEp-2 cells. Infect Immun **75:**5405-5414.

133. **Sander, P., M. Rezwan, B. Walker, S. K. Rampini, R. M. Kroppenstedt, S. Ehlers, C. Keller, J. R. Keeble, M. Hagemeier, M. J. Colston, B. Springer,**

and E. C. Bottger. 2004. Lipoprotein processing is required for virulence of Mycobacterium tuberculosis. Mol Microbiol **52:**1543-1552.

134. **Sardis, M. F., and A. Economou.** 2010. SecA: a tale of two protomers. Mol Microbiol **76:**1070-1081.

135. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2001. Comprehensive identification of conditionally essential genes in mycobacteria. Proc Natl Acad Sci U S A **98:**12712-12717.

136. **Sauve, V., S. Bruno, B. C. Berks, and A. M. Hemmings.** 2007. The SoxYZ complex carries sulfur cycle intermediates on a peptide swinging arm. J Biol Chem **282:**23194-23204.

137. **Schneider, G.** 1999. How many potentially secreted proteins are contained in a bacterial genome? Gene **237:**1113-1121.

138. **Schwarz, F., and M. Aebi.** 2011. Mechanisms and principles of N-linked protein glycosylation. Curr Opin Struct Biol **21:**576-582.

139. **Seepersaud, R., B. A. Bensing, Y. T. Yen, and P. M. Sullam.** 2012. The accessory Sec protein Asp2 modulates GlcNAc deposition onto the serine-rich repeat glycoprotein GspB. J Bacteriol **194:**5564-5575.

140. **Seepersaud, R., B. A. Bensing, Y. T. Yen, and P. M. Sullam.** 2010. Asp3 mediates multiple protein-protein interactions within the accessory Sec system of *Streptococcus gordonii*. Mol Microbiol **78:**490-505.

141. **Sharma, V., A. Arockiasamy, D. R. Ronning, C. G. Savva, A. Holzenburg, M. Braunstein, W. R. Jacobs, Jr., and J. C. Sacchettini.** 2003. Crystal structure of *Mycobacterium tuberculosis* SecA, a preprotein translocating ATPase. Proc Natl Acad Sci U S A **100:**2243-2248.

142. **Shimohata, N., Y. Akiyama, and K. Ito.** 2005. Peculiar properties of DsbA in its export across the *Escherichia coli* cytoplasmic membrane. J Bacteriol **187:**3997-4004.

143. **Sianidis, G., S. Karamanou, E. Vrontou, K. Boulias, K. Repanas, N. Kyrpides, A. S. Politou, and A. Economou.** 2001. Cross-talk between catalytic

and regulatory elements in a DEAD motor domain is essential for SecA function. EMBO J **20:**961-970.

144. **Sibbald, M. J., T. Winter, M. M. van der Kooi-Pol, G. Buist, E. Tsompanidou, T. Bosma, T. Schafer, K. Ohlsen, M. Hecker, H. Antelmann, S. Engelmann, and J. M. van Dijl.** 2010. Synthetic effects of secG and secY2 mutations on exoproteome biogenesis in Staphylococcus aureus. J Bacteriol **192:**3788-3800.

145. **Siboo, I. R., D. O. Chaffin, C. E. Rubens, and P. M. Sullam.** 2008. Characterization of the accessory Sec system of *Staphylococcus aureus*. J Bacteriol **190:**6188-6196.

146. **Siboo, I. R., H. F. Chambers, and P. M. Sullam.** 2005. Role of SraP, a Serine-Rich Surface Protein of *Staphylococcus aureus*, in binding to human platelets. Infect Immun **73:**2273-2280.

147. **Smith, M. A., W. M. Clemons, Jr., C. J. DeMars, and A. M. Flower.** 2005. Modeling the effects of prl mutations on the *Escherichia coli* SecY complex. J Bacteriol **187:**6454-6465.

148. **Smith, T., K. A. Wolff, and L. Nguyen.** 2012. Molecular biology of drug resistance in *Mycobacterium tuberculosis*. Curr Top Microbiol Immunol.

149. **Stanley, N. R., T. Palmer, and B. C. Berks.** 2000. The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. J Biol Chem **275:**11591-11596.

150. **Sullivan, J. T., E. F. Young, J. R. McCann, and M. Braunstein.** 2012. The *Mycobacterium tuberculosis* SecA2 system subverts phagosome maturation to promote growth in macrophages. Infect Immun **80:**996-1006.

151. **Takamatsu, D., B. A. Bensing, and P. M. Sullam.** 2004. Genes in the accessory sec locus of *Streptococcus gordonii* have three functionally distinct effects on the expression of the platelet-binding protein GspB. Mol Microbiol **52:**189-203.

152. **Takamatsu, D., B. A. Bensing, and P. M. Sullam.** 2005. Two additional components of the accessory sec system mediating export of the *Streptococcus gordonii* platelet-binding protein GspB. J Bacteriol **187:**3878-3883.

153. **Tam, P. C., A. P. Maillard, K. K. Chan, and F. Duong.** 2005. Investigating the SecY plug movement at the SecYEG translocation channel. EMBO J **24:**3380-3388.

154. **Tarry, M. J., E. Schafer, S. Chen, G. Buchanan, N. P. Greene, S. M. Lea, T. Palmer, H. R. Saibil, and B. C. Berks.** 2009. Structural analysis of substrate binding by the TatBC component of the twin-arginine protein transport system. Proc Natl Acad Sci U S A **106:**13284-13289.

155. **Taura, T., T. Baba, Y. Akiyama, and K. Ito.** 1993. Determinants of the quantity of the stable SecY complex in the *Escherichia coli* cell. J Bacteriol **175:**7771-7775.

156. **Tottey, S., K. J. Waldron, S. J. Firbank, B. Reale, C. Bessant, K. Sato, T. R. Cheek, J. Gray, M. J. Banfield, C. Dennison, and N. J. Robinson.** 2008. Protein-folding location can regulate manganese-binding versus copper- or zinc-binding. Nature **455:**1138-1142.

157. **Tullman-Ercek, D., M. P. DeLisa, Y. Kawarasaki, P. Iranpour, B. Ribnicky, T. Palmer, and G. Georgiou.** 2007. Export pathway selectivity of *Escherichia coli* twin arginine translocation signal peptides. J Biol Chem **282:**8309-8316.

158. **Valent, Q. A., J. W. de Gier, G. von Heijne, D. A. Kendall, C. M. ten Hagen-Jongman, B. Oudega, and J. Luirink.** 1997. Nascent membrane and presecretory proteins synthesized in *Escherichia coli* associate with signal recognition particle and trigger factor. Mol Microbiol **25:**53-64.

159. **Van den Berg, B., W. M. Clemons, Jr., I. Collinson, Y. Modis, E. Hartmann, S. C. Harrison, and T. A. Rapoport.** 2004. X-ray structure of a protein-conducting channel. Nature **427:**36-44.

160. **van der Sluis, E. O., N. Nouwen, J. Koch, J. de Keyzer, C. van der Does, R. Tampe, and A. J. Driessen.** 2006. Identification of two interaction sites in SecY that are important for the functional interaction with SecA. J Mol Biol **361:**839-849.

161. **van der Wolk, J. P., J. G. de Wit, and A. J. Driessen.** 1997. The catalytic cycle of the *Escherichia coli* SecA ATPase comprises two distinct preprotein translocation events. EMBO J **16:**7297-7304.

162. **van Wely, K. H., J. Swaving, R. Freudl, and A. J. Driessen.** 2001. Translocation of proteins across the cell envelope of Gram-positive bacteria. FEMS Microbiol Rev **25:**437-454.

163. **VanderVen, B. C., J. D. Harder, D. C. Crick, and J. T. Belisle.** 2005. Export-mediated assembly of mycobacterial glycoproteins parallels eukaryotic pathways. Science **309:**941-943.

164. **von Heijne, G.** 1990. The signal peptide. J Membr Biol **115:**195-201.

165. **Vrontou, E., and A. Economou.** 2004. Structure and function of SecA, the preprotein translocase nanomotor. Biochim Biophys Acta **1694:**67-80.

166. **Waligora, A. J., C. Hennequin, P. Mullany, P. Bourlioux, A. Collignon, and T. Karjalainen.** 2001. Characterization of a cell surface protein of *Clostridium difficile* with adhesive properties. Infect Immun **69:**2144-2153.

167. **White, G. F., S. M. Schermann, J. Bradley, A. Roberts, N. P. Greene, B. C. Berks, and A. J. Thomson.** 2010. Subunit organization in the TatA complex of the twin arginine protein translocase: a site-directed EPR spin labeling study. J Biol Chem **285:**2294-2301.

168. **Whitehouse, S., V. A. Gold, A. Robson, W. J. Allen, R. B. Sessions, and I. Collinson.** 2012. Mobility of the SecA 2-helix-finger is not essential for polypeptide translocation via the SecYEG complex. J Cell Biol **199:**919-929.

169. **WorldHealthOrganization.** 2012. WHO Information tuberculosis fact sheet.

170. **Wu, H., S. Bu, P. Newell, Q. Chen, and P. Fives-Taylor.** 2007. Two gene determinants are differentially involved in the biogenesis of Fap1 precursors in *Streptococcus parasanguis*. J Bacteriol **189:**1390-1398.

171. **Wu, H., and P. M. Fives-Taylor.** 1999. Identification of dipeptide repeats and a cell wall sorting signal in the fimbriae-associated adhesin, Fap1, of *Streptococcus parasanguis*. Mol Microbiol **34:**1070-1081.

172. **Wu, H., K. P. Mintz, M. Ladha, and P. M. Fives-Taylor.** 1998. Isolation and characterization of Fap1, a fimbriae-associated adhesin of *Streptococcus parasanguis* FW213. Mol Microbiol **28:**487-500.

173. **Wu, H., M. Zeng, and P. Fives-Taylor.** 2007. The glycan moieties and the N-terminal polypeptide backbone of a fimbria-associated adhesin, Fap1, play distinct roles in the biofilm development of *Streptococcus parasanguinis*. Infect Immun **75:**2181-2188.

174. **Xie, K., and R. E. Dalbey.** 2008. Inserting proteins into the bacterial cytoplasmic membrane using the Sec and YidC translocases. Nat Rev Microbiol **6:**234-244.

175. **Xiong, Y. Q., B. A. Bensing, A. S. Bayer, H. F. Chambers, and P. M. Sullam.** 2008. Role of the serine-rich surface glycoprotein GspB of *Streptococcus gordonii* in the pathogenesis of infective endocarditis. Microb Pathog **45:**297-301.

176. **Yahr, T. L., and W. T. Wickner.** 2001. Functional reconstitution of bacterial Tat translocation in vitro. EMBO J **20:**2472-2479.

177. **Ye, J., A. R. Osborne, M. Groll, and T. A. Rapoport.** 2004. RecA-like motor ATPases--lessons from structures. Biochim Biophys Acta **1659:**1-18.

178. **Yen, Y. T., T. A. Cameron, B. A. Bensing, R. Seepersaud, P. C. Zambryski, and P. M. Sullam.** 2013. Differential localization of the streptococcal accessory sec components and implications for substrate export. J Bacteriol **195:**682-695.

179. **Yen, Y. T., R. Seepersaud, B. A. Bensing, and P. M. Sullam.** 2011. Asp2 and Asp3 interact directly with GspB, the export substrate of the *Streptococcus gordonii* accessory Sec System. J Bacteriol **193:**3165-3174.

180. **Zhou, J., and Z. Xu.** 2003. Structural determinants of SecB recognition by SecA in bacterial protein translocation. Nat Struct Biol **10:**942-947.

181. **Zhou, M., H. Zhang, F. Zhu, and H. Wu.** 2011. Canonical SecA associates with an accessory secretory protein complex involved in biogenesis of a streptococcal serine-rich repeat glycoprotein. J Bacteriol **193:**6560-6566.

182. **Zhou, M., F. Zhu, S. Dong, D. G. Pritchard, and H. Wu.** 2010. A novel glucosyltransferase is required for glycosylation of a serine-rich adhesin and biofilm formation by *Streptococcus parasanguinis*. J Biol Chem **285:**12140-12148.

183. **Zhou, M., F. Zhu, Y. Li, H. Zhang, and H. Wu.** 2012. Gap1 functions as a molecular chaperone to stabilize its interactive partner Gap3 during biogenesis of serine-rich repeat bacterial adhesin. Mol Microbiol **83:**866-878.

CHAPTER 2

Protein Export by the Mycobacterial SecA2 System is Determined by the Preprotein Mature Domain[1,2]

At the core of the bacterial general secretion (Sec) pathway is the SecA ATPase, which powers translocation of unfolded preproteins containing Sec signal sequences through the SecYEG membrane channel. Mycobacteria have two non-redundant SecA homologs: SecA1 and SecA2. While the essential SecA1 handles "housekeeping" export, the nonessential SecA2 exports a subset of proteins and is required for *Mycobacterium tuberculosis* virulence. Currently, it is not understood how SecA2 contributes to Sec export in mycobacteria. In this study we focused on identifying the features of two SecA2 substrates that target them to SecA2 for export, the Ms1704 and Ms1712 lipoproteins of the model organism *M. smegmatis*. We found that the mature domains of Ms1704 and Ms1712, not the N-terminal signal sequences, confer SecA2-dependent export. We also demonstrated that the lipid modification and the extreme N-terminus of the mature protein do not impart the requirement for SecA2 in export. Because the Tat system only exports folded proteins, this result implies that SecA2 substrates can fold in the

[1]Adapted for this dissertation from:
Feltcher, M. E., H. S. Gibbons, L. S. Ligon, and M. Braunstein. 2013. Protein export by the mycobacterial SecA2 system is determined by the preprotein mature domain. J Bacteriol **195:**672-681.

[2]Current address of contributing author Henry S. Gibbons: Edgewood Chemical Biological Center, BioSciences Division, Aberdeen Proving Ground, MD, 21010

cytoplasm and suggests a putative role of SecA2 in enabling export of such proteins. Thus, the mycobacterial SecA2 system may represent another way that bacteria solve the problem of exporting proteins that can fold in the cytoplasm.

Introduction

Protein export pathways exist in all bacteria and fulfill the vital role of transporting proteins synthesized in the cytoplasm to the bacterial cell envelope or extracellular environment. Many of the proteins exported by these pathways function in essential physiological processes or have important roles in bacterial pathogenesis.

There are two conserved protein export pathways found across bacteria: the general secretion (Sec) and the twin-arginine translocation (Tat) pathways (49, 58). The Sec pathway carries out the bulk of protein export, which makes Sec export an essential process in all bacteria. The core components of the Sec pathway are a cytoplasmic SecA ATPase and a membrane complex composed of SecY, SecE, and SecG proteins (19).

Proteins bound for export by the Sec pathway are synthesized as preproteins with N-terminal signal sequences that are composed of a positively charged N-terminus, a hydrophobic core, and a polar C-terminus containing a cleavage site (76). The signal sequence promotes binding of the preprotein to a cleft of SecA formed by multiple domains, including the preprotein-binding domain and C-terminal domain (3, 32, 40, 54, 59). SecA then targets the preprotein to SecYEG at the membrane, where the signal sequence inserts into a channel formed by SecY (36, 47, 69). SecA also contains two nucleotide-binding domains with ATPase activity. SecA powers translocation of the preprotein across the membrane through the SecY channel using cycles of

conformational changes coupled to ATP-binding and hydrolysis (26, 48). Preproteins must be in an unfolded conformation to be competent for Sec export. Cytosolic chaperones and the process of translocation itself can help keep preproteins in an unfolded and translocation competent state (2, 29, 56). After export, the signal sequence is cleaved from the preprotein by signal peptidases to yield the mature protein (57).

Proteins exported by the Tat pathway are also synthesized as preproteins containing cleavable N-terminal signal sequences (45). Tat signal sequences, however, are distinguished from Sec signal sequences by the presence of a twin-arginine motif that includes a nearly invariant pair of arginine residues (12). Tat exported proteins are targeted to a different membrane channel comprised of the TatA, TatB, and TatC proteins (67). A major distinction between the Sec and Tat pathways is that, while the Sec pathway exports proteins in an unfolded state, the Tat pathway only exports folded preproteins (25, 61).

Mycobacteria, including the human pathogen *Mycobacterium tuberculosis*, and some Gram positive species also have a SecA2 protein export system (17, 28). SecA2 systems are defined by the presence of a second, non-redundant, homolog of the SecA ATPase. In these bacteria, SecA1 is the essential SecA protein responsible for canonical Sec export, while SecA2 is involved in exporting a subset of proteins (17, 28). In many bacterial pathogens containing SecA2 systems, such as *M. tuberculosis*, SecA2 is not essential but required for virulence (8, 18, 22, 46, 53, 68)

Some SecA2 systems are known as SecA2-SecY2 systems because they contain an extra SecY, which is presumed to function as a specialized channel (28). SecA2-SecY2 systems appear dedicated to exporting proteins that are glycosylated in the

cytoplasm prior to export, such as the GspB protein of *Streptococcus gordonii* (8, 22, 53, 68). Other SecA2 systems, such as those found in mycobacteria, are known as SecA2-only systems because they lack an extra SecY or any obvious alternative membrane channel. In mycobacteria, SecA2 likely works with machinery from the general Sec pathway including SecA1 and SecYEG (64).

A past study in *M. smegmatis* identified two proteins that require SecA2 for export to the cell wall: Ms1704 and Ms1712 (33). Both of these proteins are predicted ABC-type sugar-binding components of putative ABC transporters (73). Ms1704 and Ms1712 possess N-terminal Sec signal sequences that contain a lipobox motif (L[A/S][G/A]C), common to exported bacterial lipoproteins (15, 39). Currently, it is not known what feature(s) make a preprotein dependent on SecA2 for export in SecA2-only systems. In this study, we used *M. smegmatis* Ms1704 and Ms1712 to address this fundamental question about SecA2 export in mycobacteria. We tested the role of the signal sequence and mature domain in determining SecA2-dependency by swapping the signal sequences of Ms1704 and Ms1712 with those of preproteins exported by the canonical Sec pathway, independent of SecA2. These studies showed that the mature domain of Ms1704 and Ms1712, not the signal sequence, contains the information that makes these proteins SecA2-dependent. We also showed that the lipid modification and extreme N-terminus of the mature domain is not sufficient to impart the requirement for SecA2 in export. Remarkably, however, when the mature domain of Ms1704 was fused to a Tat signal sequence, the protein was no longer exported in a SecA2-dependent manner; instead, Ms1704 was then exported by the Tat pathway. Because the Tat pathway is uniquely built to export folded preproteins, this suggests that SecA2 substrates

may be prone to folding in the cytoplasm and the role of SecA2 may be to facilitate export of such proteins. Our results also reinforce the differences that exist between the SecA2-SecY2 systems of other bacteria and the SecA2-dependent export system of mycobacteria.

Materials and Methods

Bacterial strains and growth conditions: *Escherichia coli* DH5α was used for all DNA cloning and was grown at 37º C in Luria-Bertani medium (Fisher) supplemented with 40 µg/ml kanamycin or 150 µg/mL hygromycin when appropriate. *M. smegmatis* mc^2155 (wild type) (70) or NR116 (Δ*secA2*) (64) were grown at 30º C in Mueller-Hinton medium (BD Diagnostic Systems) with 0.1% (v/v) Tween 80 and supplemented with 20 µg/mL kanamycin when appropriate. Because the *M. smegmatis* strains PM759 (31) and LL115 are leucine auxotrophs, they were always grown in Luria-Bertani medium (Fisher) supplemented with 80 µg/mL L-Lysine.

Construction of *M. smegmatis* Δ*blaS* Δ*secA2* mutant: Strain LL115 (*ept-1* Δ*lysA4 rpsL6* Δ*blaS1* Δ*secA2)*, the *M. smegmatis* Δ*blaS* Δ*secA2* mutant, was constructed by two-step allelic exchange, as described previously (64), using the Δ*blaS* mutant PM759 strain and the Δ*secA2* suicide plasmid pNR6. Successful deletion of *secA2* was confirmed by Southern blot analysis (data not shown).

Plasmid construction and *M. smegmatis* electroporation: All plasmids and oligonucleotides used in this study are listed in Tables 2.1 and 2.2, respectively. All cloned plasmid inserts were confirmed error-free by DNA sequencing (Eton Biosciences, NC). Plasmids were electroporated into *M. smegmatis* as previously described (16).

Construction of Ms1704 and PhoA signal sequence chimera expression vectors

To create ssPhoA-Ms1704-HA, the signal sequence of Ms1704-HA was first removed by inverse PCR of pHSG58 using primers pHSG51-SSswap-1 and pHSG58-SSswap-2 to yield plasmid pMF122. The signal sequence of *M. smegmatis* PhoA was amplified from pML440 using primers PhoAss-1 and PhoAss-2. The PCR product was cut with BamHI and NdeI and ligated with similarly cut pMF122 to produce the ssPhoA-Ms1704-HA expression plasmid, pMF124. To create ssMs1704-PhoA, the signal sequence of Ms1704 was PCR amplified from pHSG58 using primers 1708ss-19kDa-1 and 1704ss-PhoA-2, while the mature domain of PhoA was amplified from *M. smegmatis* mc^2155 genomic DNA using primers 1708ss-PhoA-3OL and 1708ss-PhoA-4OL. The two PCR products were fused using splice overlap extension PCR with primers 1708ss-19kD-1 and 1708ss-PhoA-4OL. The resultant amplicon was cloned into pCR2.1 to yield pMF125. The ssMs1704-PhoA piece was excised by digestion with NotI and EcoR1, and ligated into similarly cut pMV261 to produce pMF126.

Construction of Ms1712 and PhoA signal sequence chimera expression vectors

To create ssPhoA-1712-HA, the signal sequence of Ms1712-HA was first removed by inverse PCR of pHSG51 using primers 19kDss-1708-1 and 19kDss-1708-2 to generate pHSG61. The signal sequence of *M. smegmatis* PhoA was amplified from pML440 using primers PhoAss-1 and PhoAss-2, cut with BamHI and NdeI, and ligated with similarly cut pHSG61 to produce the ssPhoA-Ms1712-HA expression plasmid, pHSG68. To create ssMs1712-PhoA, the signal sequence of Ms1712 was PCR amplified from pHSG51 using primers 1708ss-19kD-1 and 1708ss-PhoA-3OL, while the mature domain of PhoA was amplified from *M. smegmatis* mc^2155 genomic DNA using primers 1708ss-PhoA-3OL

and 1708ss-PhoA-4OL. The two PCR products were fused using splice overlap extension PCR with primers 1708ss-19kD-1 and 1708ss-PhoA-4OL. The resultant amplicon was cloned into pCR2.1 to yield pHSG70. The ssMs1712-PhoA piece was excised by digestion with NotI and EcoR1, and ligated into similarly cut pMV261 to produce pMF128.

Construction of Ms1712 and 19kD signal sequence chimera expression vectors

To create ss19kD-Ms1712-HA, the signal sequence of 19kD was amplified from *M. tuberculosis* H37Rv genomic DNA with primers 19kDass-1 and 19kDass-2 and cloned into pCR2.1 to yield pMB244. The 19kD signal sequence was amplified from pMB244 using primers 19kDss-1 and 19kDss-2, cut with BamHI and NdeI, and ligated with similarly cut pHSG61 to generate pMF115. To create ssMs1712-19kD, the signal sequence of Ms1712 was PCR amplified from pHSG51 using primers 1708ss-19kD-1 and 1708ss-19kD-2, while the mature domain of 19kD was amplified from *M. tuberculosis* H37Rv genomic DNA using primers 1708ss-19kD-3 and 1708ss-19kD-4. The two PCR products were fused using splice overlap extension PCR with primers 1708ss-19kD-1 and 1708ss-19kD-4. The PCR product was cloned into pCR2.1 to yield pHSG64. The ssMs1712-19kD piece was excised by digestion with NotI and EcoR1, and ligated into similarly cut pMV261 to produce pHSG66.

Construction of Non-lipidated Ms1704CA and Ms1712CA expression vectors

The invariant cysteine in the lipobox motifs of both Ms1704 and Ms1712 carried on pHSG58 and pHSG85 respectively, was changed to alanine using the Quikchange site-directed mutagenesis kit (Stratagene).

Construction of the Ms1704 and Antigen 85B signal sequence chimera expression vector

To create ssAg85B-Ms1704-HA, the signal sequence of *M. tuberculosis* Antigen85B was excised as a BamHI-NdeI fragment from pMM7-2. This fragment was ligated with similarly cut pMF122 containing the Ms1704 mature region to generate pMF127. The region of pMF127 containing the *hsp60* promoter and ss*Ag85B-1704-HA* fusion was excised as a NotI-EcoRV fragment and ligated with similarly cut pMV306 to generate the integrating expression vector, pMF136.

Construction of the ssPhoA-ΔN-Ms1704-HA expression vector

The sequence regions flanking either side of the Ms1704 sequence to be deleted were amplified using primers 19kDHA-F and phoA1704F1SOE-R, and phoA1704F1SOE-F and msmeg1704R primers respectively. The two PCR products were fused using splice overlap extension PCR with primers 19kDHA-F and msmeg1704R. The PCR product was digested with NotI and EcoRV, and ligated into similarly cut pMF124 to produce pMF246.

Construction of Ms6020-HA expression vector

To generate an Ms6020-HA expression vector, *ms6020* was amplified from *M. smegmatis* genomic DNA using primers Msmeg6020-F and Msmeg6020-R and cloned into pCR2.1 to generate pMF169. The fragment containing *ms6020* was excised as a SmaI-HindIII fragment and ligated with MscI/HindIII-cut pJSC77, a pMV261.*kan* derivative containing the HA epitope. The resultant plasmid, pMF170, expresses Ms6020-HA.

Construction of Ms1704-'BlaTEM1 expression vector

To generate an expression plasmid containing full-length Ms1704 fused to the N-terminus of 'BlaTEM1 (lacking its native signal sequence), the C-terminal HA-tag was first removed from Ms1704-HA by amplifying *msmeg1704* from pHSG58 using primers hsp60 SD plus smeg secA2 5' and 1704R-BamHI. This piece containing *msmeg1704* was cut with XmnI and BamHI, then ligated into MscI/BamHI-cut pMV261 to generate pMF168. A fragment containing'BlaTEM1 was excised from pJM114 with BamHI, and ligated into BamHI-cut pMF168 to generate pMF171, containing Ms1704-'BlaTEM1.

Construction of ssPlcB-Ms1704-HA and ssPlcB-19kD chimera expression vectors

To generate an ssPlcB-Ms1704-HA chimera, the signal sequence from *M. tuberculosis plcB* was removed from pMB222 using BamHI and NheI and ligated into similarly cut pMF122 to generate the episomal pMF131. A NotI-EcoRV fragment containing the *hsp60* promoter and the ss*plcB-Ms1704-HA* fusion was then subcloned into similarly cut pMV306 to generate the integrating expression vector, pMF135. To create ssPlcB-19kD, the mature domain of 19kD was amplified without the native signal sequence from pHSG67 using primers 19kDmature-F and Msmeg1712R-2, and then cloned into pCR2.1 to generate pMF161. A BamHI fragment containing Δ*ss19kD* was excised from pMF161 and ligated with similarly cut pMB222 to generate pMF167.

Subcellular fractionation: Subcellular fractions of *M. smegmatis* cells were generated as previously described (33, 65). Briefly, whole cell lysates of *M. smegmatis* were generated by passing cells through a French pressure cell at 20,000 psi. Unlysed cells were removed by centrifugation at 3,000 x g to generate clarified whole cell lysates (WCL). Protein concentrations of clarified lysates were determined by BCA assay using a BSA standard

(Pierce) and equalized between strains for every experiment. One milliliter of each equalized lysate was centrifuged at 27,000 x g for 30 minutes to pellet cell wall material (CW), then 100,000 x g for 2 hours to separate membrane (M) and soluble (SOL) fractions. For all experiments, protein derived from the same amount of starting cells for each fraction was analyzed by immunoblotting.

Culture filtrate protein preparation: *M. smegmatis* was grown in Mueller Hinton media with Tween 80 omitted. Identical *M. smegmatis* cultures grown with Tween 80 were used to monitor bacterial growth to an OD_{600} of 0.4-0.7 upon harvesting. Cells from the cultures where Tween 80 was omitted were separated from culture supernatants by centrifugation at 3,000 x g and lysed by passage through the French press to prepare WCL. Two hundred milliliters of culture supernatants were filtered through a 0.2 µM filter to remove cell material and concentrated using a 10kD membrane (Amicon) in a stirred cell concentrator (Amicon) at 4º C to a volume of approximately 10 mL. The concentrated volume was kept equivalent between strains for each experiment. Protein from 1 mL of concentrated culture supernatant was precipitated overnight at 4º C with 10% trichloroethanoic acid. Protein pellets were washed with acetone, re-suspended in 25 µL 1X SDS-PAGE buffer, and then boiled. Whole cell lysates from the cell pellets were analyzed for protein concentration by BCA assay and equal amounts of protein from each strain loaded on a 1D-PAGE gel. Protein obtained from equal culture supernatant volumes between strains were also loaded on a 1D-PAGE gel.

Immunoblotting: Subcellular fraction material was combined with SDS-PAGE buffer, separated by 1D-PAGE, and transferred to a nitrocellulose membrane (Whatman). The following antibodies were used for immunoblots: anti-HA (Covance) used at 1:10,000,

anti-PhoA (Research Diagnostics International) used at 1:20,000, anti-19kD (provided by Douglas Young, MRC National Institute for Medical Research) used at 1:20,000, anti-BlaTEM1 (QED Biosciences) used at 1:5,000, anti-MspA (provided by Michael Niederweis, University of Alabama at Birmingham) used at 1:20,000. For cytoplasmic and cell lysis controls, either the GroEL1 or GroEL2 chaperone was detected using anti-HIS (Abgent) at 1:10,000, or anti-GroEL2 (HAT5, IT-64, obtained from the World Health Organization collection) at 1:20,000. The anti-HIS antibody recognizes a string of endogenous histidines in GroEL1 (62). Proteins were detected using either anti-rabbit or anti-mouse secondary antibodies conjugated to either horseradish peroxidase (BioRad) or alkaline phosphatase (GE Healthcare). Signal was detected using Western Lightning Chemiluminescent detection reagent (Perkin-Elmer) or ECF reagent (GE Healthcare), as appropriate.

Immunoprecipitation of ssMsPhoA-Ms1704-HA for glycoprotein staining

Wild type *M. smegmatis* cells expressing either ssPhoA-Ms1704-HA or MsPhoA-HA were re-suspended in 200 µL of denaturing lysis buffer (50mM Tris-Cl pH 8.0, 1 mM EDTA. 1 mM DTT, 1% SDS, 1X protease inhibitor cocktail (PIC) [2 µg/mL each of aprotinin, E-64, leupeptin, and pepstatin A, and 100 µg/mL Pefabloc SC], and 1 mM phenylmethanesµLfonylfluoride, PMSF). Whole cell lysates were generated by vortexing cells with glass beads and boiling. Next, 100 µL of the lysates supernatant was incubated overnight with 650 µL immunoprecipitation buffer (50 mM Tris-Cl pH 8.0, 150 mM NaCl, 2% Triton X-100, 0.1 mM EDTA, 1X PIC, 1 mM PMSF) containing 25 µL of anti-HA beads (Sigma).

Periodic Acid Schiff Staining for Glycosylation

Two SDS-PAGE gels were run with 25 µL of immunoprecipitated ssPhoA-Ms1704-HA or PhoA-HA, along with 600 µg Mpt32 protein (BEI Resources), 30 µg PhoS1 protein (BEI Resources), 7.5 ng trypsin (Pierce), and 7.5 ng horse radish peroxidase (Pierce). One gel was Coomassie stained to visualize all proteins, while the other was stained with the periodic acid Schiff stain (Pierce) to visualize glycoproteins only.

Results

The mature domain of Ms1704 imparts the requirement for SecA2 in export

We set out to determine if preproteins exported by the mycobacterial SecA2 system have signal sequences that specifically target them to SecA2 for export or if their distinguishing feature resides in the mature domain. To address this question, we created a series of chimeric fusion proteins in which the signal sequence of a SecA2-dependent protein (Ms1704 or Ms1712) was swapped with the Sec signal sequence of a protein exported by the canonical Sec pathway, independent of SecA2 (Figure 2.1). Export of the chimeras was assessed in the following manner. *M. smegmatis* cells expressing each chimera were lysed to generate whole cell lysates (WCL) that were then subjected to differential ultracentrifugation to separate cell wall (CW), membrane (M), and cytosol-containing soluble (SOL) fractions. These subcellular fractions were then analyzed for localization of the fusion protein by immunoblot analysis with antibodies that recognize the mature domain or, in the case of Ms1704 and Ms1712, recognize a C-terminal HA epitope tag fused to the mature domain. To assess the requirement for SecA2 in protein export we specifically compared export to the cell wall between wild type and Δ*secA2*

mutant cells, although all the fractions are shown on the immunoblots. Immunoblotting for a cytoplasmic GroEL chaperone and the cell wall MspA porin was used as controls for subcellular fraction integrity.

As previously reported, when expressed in wild type *M. smegmatis* the C-terminally HA-tagged Ms1704 protein containing its native signal sequence was exported to the cell wall (CW) (33). However, Ms1704-HA export was severely compromised when expressed in the Δ*secA2* mutant, indicating that SecA2 is required for export of this protein. In comparison, export of the MspA control protein was unchanged by deletion of *secA2*, as was also shown previously (Figure 2.2A). Our first set of chimeras employed an *M. smegmatis* alkaline phosphatase (MsPhoA), which is exported in a SecA2-independent manner (33). In addition, MsPhoA is a proven lipoprotein with a Sec signal sequence that contains a lipobox motif, as is also the case for Ms1704 and Ms1712 (42). When the signal sequence of Ms1704 was replaced with the signal sequence of MsPhoA, the ssMsPhoA-Ms1704-HA chimera was exported to the cell wall in wild type *M. smegmatis* but was not exported to the cell wall of the Δ*secA2* mutant (Figure 2.2A). These results indicated that the MsPhoA signal sequence can support export of the Ms1704 mature domain, but this export requires SecA2. In the converse fusion, when the Ms1704 signal sequence was fused to the mature domain of MsPhoA, the ssMs1704-MsPhoA fusion protein was efficiently exported to the cell wall in both wild type and Δ*secA2* mutant *M. smegmatis*. Thus, export of the mature domain of MsPhoA remained SecA2-independent, despite containing a signal sequence from a SecA2 substrate. This set of chimeras suggested there is something distinct about the mature domain of Ms1704 that dictates SecA2-dependency.

The mature domain of Ms1712 also imparts the requirement for SecA2 in export

To see if the above results extend to the other known *M. smegmatis* SecA2 substrate, we constructed similar chimeras with Ms1712. As previously demonstrated, Ms1712-HA containing its native signal sequence is only exported to the cell wall when SecA2 is present (33). Consistent with results using Ms1704, the Sec signal sequence from MsPhoA supported export of the Ms1712 mature domain to the cell wall in wild type *M. smegmatis*, but this export was reduced in the Δ*secA2* mutant (Figure 2.2B). Further, the converse fusion with the signal sequence of Ms1712 fused to the mature domain of MsPhoA (ssMs1712-MsPhoA) was exported equally in both wild type and the Δ*secA2* mutant.

The *M. tuberculosis* 19kD lipoprotein is another protein previously shown to be exported by *M. smegmatis* in a manner independent of SecA2 (33, 80). Therefore, we also created chimeras between Ms1712 and the 19kD lipoprotein. In line with the results from the MsPhoA fusions, the Sec signal sequence of 19kD supported export of the Ms1712 mature domain to the cell wall in wild type *M. smegmatis* but not in the Δ*secA2* mutant. Furthermore, the signal sequence of Ms1712 could export the mature domain of the 19kD lipoprotein in both wild type and cells lacking SecA2, indicating that ssMs1712-19kD export occurred independently of SecA2 (Figure 2.2B).

The results from these signal sequence swap experiments showed that whenever a chimera contained a mature domain from a SecA2-dependent protein the fusion protein was exported to the cell wall only in the presence of SecA2. In contrast, all chimeras containing the signal sequence of Ms1704 or Ms1712 fused to a mature domain of a

SecA2-independent protein were exported equally well by wild-type and the Δ*secA2* mutant strains. These results indicated that the mature domain of SecA2 substrates contains feature(s) that necessitate a role for SecA2 in export and that SecA2 substrates do not contain SecA2 specific signal sequences.

SecA2-dependent export is independent of lipid modification of substrates

Because both Ms1704 and Ms1712 are lipoproteins we considered the possibility that they have a unique lipid modification that imparts the need for SecA2 in export. To test the influence of lipid modification on SecA2-dependent export, we used site-directed mutagenesis to replace the codon for the invariant cysteine in the lipoboxes of Ms1704 and Ms1712 to one encoding an alanine. This cysteine is the site of diacylglycerol attachment, which is the first step in lipid modification (15). An alanine substitution for this cysteine (CA) will prevent lipid attachment but should not prevent export because it converts the lipoprotein signal sequences of Ms1704 and Ms1712 to standard Sec signal sequences (60).

When the non-lipidated Ms1704CA-HA and Ms1712CA-HA proteins were expressed in wild type *M. smegmatis*, we could detect the proteins in whole cell lysates, but were initially surprised to find that neither protein localized to the cell wall fraction (data not shown). For at least some mycobacterial lipoproteins, the lipid moiety is shown to be required for anchoring to the cell wall and without lipid modification the proteins are fully secreted (74). Indeed, both Ms1704CA-HA and Ms1712CA-HA proteins were secreted into the media, as detected in culture filtrates (CF) from the wild type strain. Immunoblotting of culture filtrates with anti-GroEL antibodies confirmed that the presence of Ms1704CA-HA and Ms1712CA-HA in the media was not due to cell lysis.

When culture filtrates from wild type or Δ*secA2* mutant expressing Ms1704CA-HA or Ms1712CA-HA were compared, we found less protein secreted in the culture filtrates obtained from the Δ*secA2* mutant (Figure 2.3A). There was also evidence of Ms1712CA-HA accumulation in the whole cell lysate of the Δ*secA2* mutant, which included a higher molecular weight band that likely represents precursor protein containing the uncleaved signal sequence. This higher molecular weight form is consistent with a form of Ms1712 that accumulated in the presence of globomycin, an inhibitor of the lipoprotein signal peptidase, seen in a previous study (33). These observations are indicative of an export defect in the Δ*secA2* mutant and they showed that secretion of non-lipidated Ms1704CA-HA and Ms1712CA-HA proteins is dependent on the presence of SecA2.

As another way of testing the importance of the lipid modification to SecA2 export, we replaced the Ms1704 signal sequence with the signal sequence of *M. tuberculosis* Antigen 85B, a well known secreted protein of *M. tuberculosis* with a standard Sec signal sequence (i.e. no lipobox) (78). This chimera did not include the invariant cysteine of the Ms1704 lipobox and was therefore non-lipidated. As with the CA mutant proteins, the ssAg85B-Ms1704-HA chimera expressed by wild type *M. smegmatis* did not localize to the cell wall but was instead secreted into the culture filtrate. In addition, secretion of the ssAg85B-Ms1704-HA protein was impaired in the Δ*secA2* mutant, which demonstrated a requirement for SecA2 in the export of this non-lipidated chimera (Figure 2.3B). Taken together, these results show that SecA2-dependent export of the Ms1704 and Ms1712 mature domains is not a consequence of lipid modification.

The Ms1704 mature domain is not heavily glycosylated

We next considered the possibility that the mature domains of mycobacterial SecA2 substrates such as Ms1704 contain a post-translational modification, other than lipidation, which influences targeting to SecA2. Substrates of SecA2–SecY2 systems are heavily glycosylated prior to export and this modification blocks export by the canonical SecA1/SecYEG in species such as *Streptococcus gordonii* and *Streptococcus parasanguinis* (6, 22, 53). Because of the precedent that glycosylation can influence targeting to SecA2; we tested the mature domain of Ms1704 for evidence of sugar modification. The ssPhoA-Ms1704-HA chimera, containing the Ms1704 mature domain, was chosen for analysis of possible glycosylation because of the high level of expression that was achieved with this construct.

The ssPhoA-Ms1704-HA protein was expressed in wild type *M. smegmatis* and immunoprecipitated using anti-HA beads. We then resolved the protein by 1D SDS-PAGE and stained with either Coomassie or periodic acid–Schiff (PAS) stain. The PAS stain is a generic stain for protein-bound carbohydrates and is often used to detect glycoproteins on polyacrylamide gels (50). While ssPhoA-Ms1704-HA was readily visible by Coomassie staining, the PAS stain failed to detect a corresponding band (Figure 2.4). However, the PAS stain did react with multiple proteins known to be glycosylated including horseradish peroxidase (HRP), and the *M. tuberculosis* PhoS1 and Mpt32 proteins (35). We were also unable to detect PAS staining of trypsin and immunoprecipitated MsPhoA-HA, indicating that our PAS staining was specific for glycoproteins.

These results demonstrate that the mature domain of Ms1704 is not heavily glycosylated. While we cannot completely rule out the possibility that the Ms1704 mature domain is modified post-translationally, our PAS staining suggests that mycobacterial SecA2 substrates are not necessarily glycosylated heavily like the serine-rich glycoproteins of SecA2–SecY2 systems.

Export of the Msmeg6020 ABC-type sugar-binding lipoprotein is independent of SecA2

Ms1704 and Ms1712 share 40% amino acid similarity in their mature domains and both are predicted to be ABC-type sugar-binding proteins. Given the signal sequence chimera results described above, it is interesting to point out that this similarity is limited to the mature domain and does not include the signal sequences of these two proteins. We questioned whether similar structure or function of the mature domain is the reason both these proteins share a SecA2 requirement for export.

To address this possibility, we tested the role of SecA2 in the export of Ms6020, which is one of the other 19 predicted ABC-type sugar-binding proteins of *M. smegmatis* (73). Ms6020 is also a predicted lipoprotein with a lipoprotein type Sec signal sequence (39, 73). The amino acid sequence of the Ms6020 mature domain is 48% similar to the Ms1704 mature domain and 41% similar to the Ms1712 mature domain. By expressing a C-terminal HA-tagged Ms6020 protein in wild type *M. smegmatis* we showed that the protein was indeed exported to the cell wall. However, Ms6020-HA was exported to the cell wall equally well in the Δ*secA2* mutant of *M. smegmatis* (Figure 2.5). This result indicated that the requirement for SecA2 does not extend to all exported ABC-type sugar-binding proteins of mycobacteria.

The extreme N-terminus of the Ms1704 mature domain is not required for SecA2-dependent export

Because our chimera analysis demonstrated that the mature domain of SecA2 substrates imparts the requirement for SecA2, we reasoned the mature domain may contain sequences required for SecA2 targeting. To address this possibility, we constructed versions of Ms1704 with deletions at the N- or C-terminus of the mature domain. To focus in on the contribution of the mature domain, we made the deletions in the ssPhoA-Ms1704 chimera, which we showed above is SecA2-dependent (Figure 2.2A). Unfortunately, all but one of these derivatives failed to produce enough protein for subsequent analysis, probably because the deletions yielded instable proteins. The only construct that could be analyzed had 9 amino acids at the most extreme N-terminus of the Ms1704 mature domain deleted (Figure 2.6A). Export of ssPhoA-ΔN-Ms1704-HA to the cell wall was monitored in both wild type *M. smegmatis* and the Δ*secA2* mutant (Figure 2.6B). The ssPhoA-ΔN-Ms1704-HA protein was localized to the cell wall of the wild type strain, but export was severely reduced in the Δ*secA2* mutant. This result indicated that the most extreme N-terminus is not required for export by the SecA2-dependent pathway of mycobacteria. This result is interesting in light of the fact that the exported GspB protein of the *S. gordonii* SecA2-SecY2 system has a short amino acid sequence named AST (for *a*ccessory *s*ec *t*ransport) located in the extreme N-terminus of its mature domain. This AST domain is required for SecA2-SecY2 targeting/export of GspB (9, 11).

Addition of a 'BlaTEM1 reporter to the C-terminus of full length Ms1704 eliminates the requirement for SecA2 in export

As another way to identify sequences in the mature domain that target the protein to the SecA2-dependent export system, we tested if SecA2-dependency could be transferred to a fused heterologous protein. For this reason, we constructed a fusion protein in which the C-terminus of full length Ms1704 was fused to the truncated β-lactamase reporter ('BlaTEM1) that lacks its native signal sequence. β-lactam antibiotics target the bacterial cell wall, therefore the 'BlaTEM1 reporter must be exported in order to protect against these drugs. We previously showed that 'BlaTEM1 can report on export when fused to Sec exported proteins that are expressed in a β-lactamase mutant (Δ*blaS*) of *M. smegmatis* (51). Expression of the Ms1704-'BlaTEM1 fusion protein in either a *M. smegmatis* Δ*blaS* mutant or in a Δ*blaS* Δ*secA2* double mutant resulted in β-lactam resistance (data not shown). This result indicated that Ms1704 was able to support export of the 'BlaTEM1 reporter and this export occurred in a SecA2-independent manner. To more directly evaluate Ms1704-'BlaTEM1 export we also localized the fusion protein to subcellular fractions by immunoblotting with anti-BlaTEM1 antibodies (Figure 2.7). This analysis showed that the Ms1704-'BlaTEM1 fusion was exported to the cell wall in both wild type and Δ*secA2* mutant *M. smegmatis.*

While this approach also failed to reveal any evidence of SecA2 targeting signals in Ms1704, these results are significant. Unlike all the chimeras described earlier, here the presence of the mature domain of Ms1704 did not impart a requirement for SecA2 in export. This result indicated that fusion to 'BlaTEM1 influences the feature(s) of the Ms1704 mature domain that necessitate SecA2 export. One plausible explanation is that

fusion to ‘BlaTEM1 altered the structure of Ms1704, enabling it to now be exported in the absence of SecA2.

Ms1704 can be exported by the twin-arginine translocation (Tat) pathway, independent of SecA2

The SecA2-independent export of the Ms1704-‘BlaTEM1 fusion suggested that a structural feature of Ms1704 may influence the need for SecA2. Because the canonical Sec pathway exports proteins that are maintained in an unfolded state, we hypothesized that SecA2 substrates may be prone to more rapid folding in the cytoplasm compared to strictly SecA1-dependent preproteins. To this end, we tested the ability of a SecA2 substrate to fold in the cytoplasm by engineering a signal sequence chimera, where the Tat signal sequence of *M. tuberculosis* phospholipase C (PlcB) was fused to the mature domain of Ms1704 (52). This strategy uses export by the twin-arginine translocation (Tat) system as a reporter for cytoplasmic folding because the Tat pathway requires preproteins to fold prior to export (25, 63). The Tat pathway has been utilized previously as an in vivo assay for cytoplasmic folding (30).

As seen with the non-lipidated mutants and ssAg85B-Ms1704-HA chimera discussed above, the ssPlcB-Ms1704-HA chimera did not localize to the cell wall fraction when expressed in wild type *M. smegmatis* but was instead secreted into the culture filtrate (Figure 2.8). This was not surprising given the lack of a lipobox in the PlcB signal sequence. Remarkably, when ssPlcB-Ms1704-HA secretion was compared between wild type and Δ*secA2* mutant the fusion was exported equally well by the two strains. To confirm that this export was occurring through the Tat pathway, we engineered an

ssPlcB-Ms1704-HA chimera where the twin-arginine residues of the PlcB signal sequence were substituted with lysine residues. Such twin arginine substitutions have previously been shown to eliminate Tat export of numerous proteins (1, 20, 71). Further, we previously showed that changing the arginine residues in the PlcB signal sequence abolishes export of *M. tuberculosis* PlcB (52). The resultant protein, ssPlcB(KK)-Ms1704-HA, was not secreted into the culture supernatants of either wild type or Δ*secA2* mutant cells (Figure 2.8B). This loss of secretion demonstrated that the Tat pathway is responsible for the SecA2-independent export of ssPlcB-Ms1704-HA.

Because the Sec pathway exports proteins in an unfolded state, we expected that this Tat-dependent export of the mature domain of Ms1704 was unusual. For this reason, we similarly tested the ability of the PlcB signal sequence to direct export of the Sec signal sequence containing 19kD lipoprotein using an ssPlcB-19kD fusion protein. When the 19kD protein with its native Sec signal sequence was expressed in *M. smegmatis*, the protein was detected in cell wall, membrane, and culture filtrate fractions of wild type *M. smegmatis* (Figure 2.8C-D). In contrast, although the ssPlcB-19kD fusion protein was detected in whole cell lysates and the soluble fraction of wild type *M. smegmatis*, this chimera was not exported to the cell wall, membrane, or culture filtrate (Figure 2.8C-D). Further, multiple migrating species of ssPlcB-19kD were observed in the whole cell lysate and cytosolic containing soluble fractions. The larger of these species is likely unprocessed precursor and the smaller product a result of degradation of non-exported protein retained in the cytosol. The difference in compatibility for Tat export between the mature domains of Ms1704 and 19kD argues for a difference in the tendency of these

representative SecA2-dependent and SecA2-independent substrates to fold in the cytoplasm prior to export.

Discussion

SecA2 systems are emerging as important pathways for the export of virulence factors in a set of Gram positive pathogens and mycobacteria (28). Yet, many mechanistic questions concerning SecA2 export remain, especially in organisms containing SecA2-only systems like mycobacteria. In mycobacteria, both SecA1 and SecA2 proteins are proven ATPases with significant homology to the well-studied *Escherichia coli* SecA (37, 64). However, SecA1 and SecA2 cannot substitute for one another, indicating that each has distinct functions in mycobacterial protein export, and the role of the SecA2 protein is not immediately obvious (17).

In this study, we investigated why certain preproteins require SecA2 for export, despite containing Sec signal sequences that would be expected to be compatible with the canonical SecA1. There is precedent for the amino acid composition of N-terminal signal sequences influencing the export pathway of a given protein: the post-translational SecA pathway, the co-translational SRP-mediated pathway, or the Tat pathway (13, 24, 55). However, our data from the signal sequence swap experiments showed that the signal sequences of two exported *M. smegmatis* SecA2 substrates are not uniquely capable of directing preproteins to SecA2 for export. Instead, the mature domains of preproteins impart SecA2-dependency. This result, however, is not entirely unexpected because, in addition to the signal sequence, it is shown that the *E. coli* SecA interacts with the mature domains of preproteins at multiple steps during translocation (5, 36, 59, 75).

We further showed that the mature domain of a SecA2-dependent preprotein is compatible with export by the Tat pathway. This latter finding is intriguing and consistent with the possibility that a key difference between mature domains of SecA2-dependent preproteins and the larger collection of preproteins exported by the canonical SecA1/SecYEG pathway is a tendency to fold in the cytoplasm prior to export.

Comparison to SecA2-SecY2 exported proteins

While this is the first study exploring the features that define proteins exported by a SecA2-only system, we can compare our results to those published on the glycosylated preproteins exported by SecA2-SecY2 systems. In comparison to the most thoroughly studied SecA2-SecY2 substrate, the *S. gordonii* GspB protein, there are clear differences in the features that determine export by the two classes of SecA2 systems.

With GspB, the signal sequence prevents export by the canonical Sec pathway by way of three glycine residues (7). With the *M. smegmatis* Ms1704 and Ms1712 proteins, the signal sequences are required for export, as truncated derivatives of these proteins lacking the signal sequences are not exported (33). However, unlike the case with the signal sequence of GspB, our chimera analysis demonstrated that signal sequences of Ms1704 and Ms1712 do not contain features that prohibit export by the canonical Sec pathway independent of SecA2.

There are also features in the mature domain of *S. gordonii* GspB that function in preventing export by the canonical Sec pathway. The distinctive glycosylation of the GspB mature domain, which occurs prior to export, also blocks export by the canonical Sec machinery (10). A similar finding is reported for the Fap1 preprotein of the

Streptococcus parasanguinis SecA2-SecY2 system (21). We explored the possibility that the mature domains of mycobacterial SecA2 substrates are glycosylated. However, periodic-acid Schiff staining failed to detect glycosylation of the mature domain of Ms1704 (Fig. S1). Although this analysis cannot completely rule out glycosylation of Ms1704 as a factor in preventing export by the canonical Sec pathway, it makes this possibility seem unlikely. We also asked whether a specific lipid modification of the mature domain could be a defining feature of *M. smegmatis* SecA2 substrates, but our data from testing non-lipidated variants of Ms1704 and Ms1712 argues against this possibility.

Finally, approximately 20 amino acids at the N-terminus of the GspB mature domain (the AST domain) are required for targeting this preprotein to the SecA2-SecY2 system of *S. gordonii* (9). The AST domain is predicted to adopt an alpha-helical conformation and mutations predicted to disrupt this secondary structure reduce GspB export. There are no predicted alpha-helices within the first 20 amino acids of either the Ms1704 or Ms1712 mature domains. Further, we showed that deletion of the extreme N-terminus of the Ms1704 mature domain did not affect SecA2-export. Thus, there does not appear to be an AST-like targeting element in the extreme N-terminus of substrates of the mycobacterial SecA2-only system.

Model for the role of the mature domain in SecA2-dependent export in mycobacteria

How might the mature domain dictate the need for SecA2-dependent export in mycobacteria? With the mycobacterial SecA2-only system proteins are most likely exported via the canonical SecYEG channel. In support of this model, depletion of the

essential *M. smegmatis* SecA1 protein eliminates export of Ms1704, which indicates that both SecA1 and SecA2 are required for export of this SecA2 substrate (64). However, SecA1 and SecA2 are non-redundant homologs, indicating that SecA2 fulfills a unique role in promoting export in mycobacteria (17). Below we discuss three non-exclusive possibilities for how and why the preprotein mature domain may dictate a requirement for SecA2 in their export by mycobacteria (Figure 2.9).

First, the mature domain could possess a post-translational modification that inhibits canonical export by the SecA1/SecYEG pathway and, thereby, necessitates SecA2 function. However, as described above, there is no evidence currently to support this possibility.

Second, the mature domains of SecA2 substrates may have specific targeting signals for SecA2 recognition. While we did not detect such a sequence at the extreme N-terminus of the Ms1704 mature domain, we cannot rule out the possibility that SecA2 targeting signals exist elsewhere. Unfortunately, our efforts to test other regions of Ms1704 by deletion analysis or with chimeric proteins, in which we swapped regions of the SecA2-dependent Ms1704 and SecA2-independent Ms6020 proteins, were unsuccessful due to protein instability problems (data not shown). However, if a SecA2-binding sequence exists, it does not appear sufficient for SecA2-dependent export as demonstrated by the inability of full length Ms1704 to promote SecA2-dependent export of the fused ‘BlaTEM1 protein. In fact, addition of ‘BlaTEM1 to the C-terminus of the Ms1704 mature domain overrode the requirement of SecA2 for export. This latter result indicates that the fusion of the ‘BlaTEM1 reporter to Ms1704 influenced a property of the

Ms1704 mature domain to enable export by the canonical Sec pathway in the absence of SecA2.

The third possibility is that a defining feature of the mature domain of SecA2 substrates is a propensity to fold in the cytoplasm, which would be problematic for canonical Sec export. Studies on Sec export in *E. coli* show that fast-folding preproteins, such as dihydrofolate reductase and β-galactosidase, are incompatible with export though the SecYEG channel (4, 38). Also, slower-folding variants of preproteins result in better translocation efficiency through SecYEG (61). In thinking then about the fusion of ‘BlaTEM1 to the C-terminus of full length Ms1704, it may have abolished the requirement for SecA2 by slowing down folding of the resultant fusion protein.

In further support of this third possibility we found that the mature domain of Ms1704 was exported by the Tat system when it was fused to a Tat signal sequence. The Tat system only exports folded proteins, and will even reject Tat substrates engineered to remain unfolded (25, 30). The Tat compatibility of the Ms1704 mature domain indicates that it can fold in the cytoplasm. Tat export has been used before as a reporter for cytoplasmic folding (25, 30, 63). A limitation of this approach is that it only indicates that some, but not necessarily all, of the Ms1704 preprotein pool exists in a Tat-competent, folded state. Some preproteins exhibit a range of cytoplasmic conformations that allow export by both the Sec and Tat pathways (51). However, when we tested the SecA2-independent 19kD lipoprotein it did not share the compatibility with Tat export. Interestingly, ssPlcB-Ms1704 was exported by the Tat pathway in the presence of SecA2, which suggests that any interactions between SecA2 and the preprotein do not interfere with Tat export.

Our results are consistent with a model where cytoplasmic folding of Ms1704 inhibits export by the canonical SecA1/SecYEG in the absence of SecA2 (Fig. 8). There is evidence in *E. coli* that SecA has some chaperone function and the canonical Sec translocase can assist in unfolding some preproteins through the action of translocation itself (2, 27). Thus, it is possible that the mycobacterial SecA2 fulfills a chaperone-like role by keeping preproteins unfolded prior to or during export. Alternatively, SecA2 could provide the energy required to translocate such challenging substrates through SecY. It is possible that SecA2 cooperates with SecA1 to carry out such functions. This cooperation could be at the level of delivering the preprotein to SecA1 or in working with SecA1 to translocate the protein across the SecY channel. The possibility of SecA1 and SecA2 working together is supported by evidence that the canonical *E. coli* SecA can function as a dimer (66).

In *E. coli* and other Gram-negative bacteria, the SecB chaperone promotes Sec export of preproteins that fold quickly in the cytoplasm (44, 77). It should be noted that mycobacteria possess a "SecB-like" homolog, but current data suggests this protein has a chaperone role unrelated to general Sec export (14). Interestingly, the *E. coli* ABC-type maltose-binding protein (MBP) requires the SecB chaperone in order to remain unfolded for Sec export, while export of the related ribose-binding protein (RBP) is SecB-independent (23, 43). These data demonstrate that proteins with similar functions can have different folding kinetics in the cytoplasm and therefore different requirements for cytoplasmic chaperones. These differences could also explain the distinction in SecA2-dependence of Ms6020 versus the Ms1704 and Ms1712 proteins.

In *M. tuberculosis*, one of the few known SecA2-dependent exported proteins, the superoxide dismutase SodA, lacks an obvious Sec signal sequence (18). While it remains possible that the role of SecA2 in the export of *M. tuberculosis* SodA is indirect, it is interesting to discover in this study that the mature domain, not the signal sequence, harbors the feature(s) that identify *M. smegmatis* proteins as SecA2 substrates. Thus, it is possible that SecA2 of *M. tuberculosis* facilitates the export of proteins like SodA in a similar manner to the way *M. smegmatis* SecA2 functions with its substrates. With these ideas in mind, it is worthwhile to note that in *Rhizobium leguminosarum* export of an SodA protein that lacks a typical Sec signal sequence was recently reported to occur in a SecA-dependent manner (41).

The data presented in this study provides a better foundation for understanding SecA2-dependent protein export. In this study we clearly identify the preprotein mature domain as the distinguishing feature of SecA2 substrates in mycobacteria. By showing that the mature domain of a *M. smegmatis* SecA2 substrate is compatible with the Tat export pathway, this study indicates that preproteins exported by this SecA2-only system are capable of folding in the cytoplasm. Thus, acquisition of a second SecA homolog in bacteria with SecA2-only systems may represent another solution to the problem of exporting proteins that pre-fold in the cytoplasm prior to translocation across the cytoplasmic membrane. In addition, our results further reinforce the distinction between the two classes of SecA2 export systems: the SecA2-SecY2 and SecA2-only pathways.

Table 2.1 Plasmids used in this study

Plasmid	Genotype	Description	Source
pCR2.1	*bla aph* ColE1	TA cloning vector	Invitrogen
pMV261	*aph oriM* P_{hsp60} ColE1	Episomal mycobacterial shuttle vector	(72)
pMV306	*aph int attP* ColE1	Integrating mycobacterial shuttle vector	(72)
pJSC77	*aph oriM* P_{hsp60}-*HA* ColE1	HA tag cloned into pMV261.*kan*	(34)
pHSG58	*aph* P_{hsp60}-*Ms1704-HA oriM* ColE1	Ms1704-HA expression vector, episomal	(33)
pHSG85	*aph* P_{ms1712}-*Ms1712-HA oriM* ColE1	Ms1712-HA expression vector, episomal	(33)
pMF122	*aph* P_{hsp60}-Δ*ssMs1704-HA oriM* ColE1	ΔssMs1704-HA for signal sequence chimera cloning	This work
pML440	*PAL5000 origin hyg,* P_{imyc}-*phoA* ColE1	*M. smegmatis* PhoA expression plasmid	(79)
pMF124	*aph* P_{hsp60}-*ssPhoA-Ms1704-HA oriM* ColE1	ssPhoA-Ms1704-HA expression vector, episomal	This work
pMF125	*bla aph* ColE1	ssMs1704-PhoA in pCR2.1	This work
pMF126	*aph* P_{hsp60}-*ssMs1704-PhoA oriM* ColE1	ssMs1704-PhoA expression vector, episomal	This work
pHSG51	*aph* P_{hsp60}-*Ms1712-HA oriM* ColE1	Ms1712-HA expression vector, episomal	(33)
pHSG61	*aph* P_{hsp60}-Δ*ss*M*s1712-HA oriM* ColE1	ΔssMs1712-HA for signal sequence chimera cloning	This work
pHSG68	*aph* P_{hsp60}-*ssPhoA-Ms1712-HA oriM* ColE1	ssPhoA-Ms1712-HA expression vector, episomal	This work
pHSG70	*bla aph* ColE1	ssMs1712-PhoA in pCR2.1	This work
pMF128	*aph* P_{hsp60}-*ssMs1712-PhoA oriM* ColE1	ssMs1712-PhoA expression vector, episomal	This work
pMB244	*bla aph* ColE1	ss19kD in pCR2.1	This work
pHSG64	*bla aph* ColE1	ssMs1712-19kD in pCR2.1	This work
pHSG66	*aph* P_{hsp60}-*ssMsmeg1712-'19kD- oriM* ColE1	ssMs1712-19kD expression vector, episomal	This work
pMF115	*aph* P_{hsp60}-*ss19kD-Ms1712-HA oriM* ColE1	ss19kD-Ms1712-HA expression vector, episomal	This work
pMF123	*aph* P_{hsp60}-*Ms1704CA-HA oriM* ColE1	Non-lipidated Ms1704CA-HA expression vector, episomal	This work
pMF105	*aph* $P_{msmeg1712}$-*Ms1712CA-HA oriM* ColE1	Non-lipidated Ms1712CA-HA expression vector, episomal	This work
pMM7-2	*aph* P_{hsp60}-*ssAg85B oriM* ColE1	*M. tuberculosis* Ag85B signal sequence in pMV261	{Zeoli, 2001 #660}
pMB222	*aph* P_{hsp60}-ssPlcB *oriM* ColE1	*M. tuberculosis* PlcB signal sequence in pMV261	(52)
pMF127	*aph* P_{hsp60}-*ssAg85B-Ms1704-HA oriM* ColE1	ssAg85B-Ms1704-HA expression vector, episomal	This work
pMF136	*aph* P_{hsp60}-*ssAg85B-Ms1704-HA int attP* ColE1	ssAg85B-Ms1704-HA expression vector, integrating	This work

pMF169	*bla aph* ColE1	Ms6020 in pCR2.1	This work
pMF170	*aph* P_{hsp60}-*Ms6020-HA oriM* ColE1	Ms6020-HA expression vector, episomal	This work
pJM114	*cat oriV ori2*	*E. coli 'blaTEM-1* cloned into pCC1	(51)
pMF168	*aph* P_{hsp60}-*Ms1704 oriM* ColE1	Ms1704 (no HA tag) expression vector, episomal	This work
pMF171	*aph* P_{hsp60}-*Ms1704-'BlaTEM1 oriM* ColE1	Ms1704-'BlaTEM1 expression vector, episomal	This work
pMF131	*aph* P_{hsp60}-*ssPlcB-Ms1704-HA oriM* ColE1	ssPlcB-Ms1704-HA expression vector, episomal	This work
pMF135	*aph* P_{hsp60}-*ssPlcB-Ms1704-HA int attP* ColE1	ssPlcB-Ms1704-HA expression vector, integrating	This work
pMF165	*aph* P_{hsp60}-*ssPlcB(KK)-Ms1704-HA int attP* ColE1	ssPlcB(KK)-Ms1704-HA expression vector, integrating, PlcB signal sequence with RR to KK change	This work
pHSG67	*aph* P_{hsp60}-*19kD oriM* ColE1	19kD expression vector, episomal	This work
pMF161	*bla aph* ColE1	Δss19kD in pCR2.1	This work
pMF167	*aph* P_{hsp60}-*ssPlcB-19kD oriM* ColE1	ssPlcB-19kD expression vector, episomal	This work
pMF246	*aph* P_{hsp60}-*ssPhoA-ΔN-Ms1704-HA oriM* ColE1	ssPhoA-ΔN-Ms1704-HA expression vector, episomal	This work

Table 2.2 Oligonucleotides used in this study

Primer Name	Sequence (5'-3')
pHSG51-SSswap-1	*ATCCATATGGAAGTGATTCCTCCGGATCG*
pHSG58-SSswap-2	*ATCAGGATCCTCCGCCGAAGGTGGCGGCGGTGG*
PhoAss-1	*CCATATGCCTGTCAGTACCTATC*
PhoAss-2	*GGATCCGGTGTCGCTGCTGCAGGCCG*
19kDss-1708-1	*CATATGGAAGTGATTCCTCCGGATCG*
19kDss-1708-2	*AGGATCCAGCGGCAAGGCACCGGGGT*
1704ss-PhoA-2	*GTGTCACCGCTGGACTCCGTTTCGGCGGACCGT*
1708ss-PhoA-3OL	*ACGGAGTCCAGCGGTGAC*
1708ss-PhoA-4OL	*GAATTCTCAGTGCGCGGGTTCCTC*
19kDass-1	*GCATATGAAGCGTGGACTGACGGTCGC*
19kDass-2	*TGGATCCGCTTGAACATCCGGAAAGAC*
19kDss-1	*GCATATGAAGCGTGGACTGACGGTCGC*
19kDss-2	*TGGATCCGCTTGAACATCCGGAAAGAC*
1708ss-19kD-1	*GTGGTCGATACCAAGCCATTTCC*
1708ss-19kD-2	*GTCTCACCGCTTCCTGTAGTTGCCTTGCCGCTGCCGCAG*
1708ss-19kD-3	*ACTACAGGAAGCGGTGAGAC*
1708ss-19kD-4	*GAATTCTTAGGAACAGGTCACCTCGA*
Msmeg6020-F	*TCCCGGGTGAAGCGCACCAGCACCCTG*
Msmeg6020-R	*TAAGCTTCTCCTGCCGCGGAGCACA*
hsp60 SD plus smeg secA2 5'	*AGGATCCATCCGGAGGAATCACTT*
1704R-BamHI	*TGGATCCCTTGACCTGGGTCTCGGTGTAG*
Msmeg1712R	*GGATCCACGTCGACATCGATAAGC*
19kDmature-F	*TCGGATCCAACAAGTCGACTACAGGAAGCGGTGAGAC*
19kDHA-F	*CGCGGCCGCGGTACCAG*
phoA1704F1SOE-R	*TCACCGTCGCTGCTGCAGGCCGTGAG*
phoA1704F1SOE-F	*CTGCAGCAGCGACGGTGACGCCAAGG*
msmeg1704R	*AGGATCCACTACGTCGACGATATCAG*

Figure 2.1 Schematic of signal sequence chimeras
The signal sequence regions of chimeras generated between *M. smegmatis* SecA2 substrates and SecA2-independent lipoproteins are shown. Amino acid sequences from either SecA2 substrate (Ms1704 or Ms1712) are highlighted in grey, while sequences from SecA2-independent proteins (MsPhoA or 19kD) are boxed in white. The predicted lipobox in each signal sequence is highlighted in bold and the predicted cleavage site adjacent to the invariant cysteine (site of lipid modification) at the +1 position of the mature domain is noted with an arrow. Amino acids introduced from cloning are unboxed.

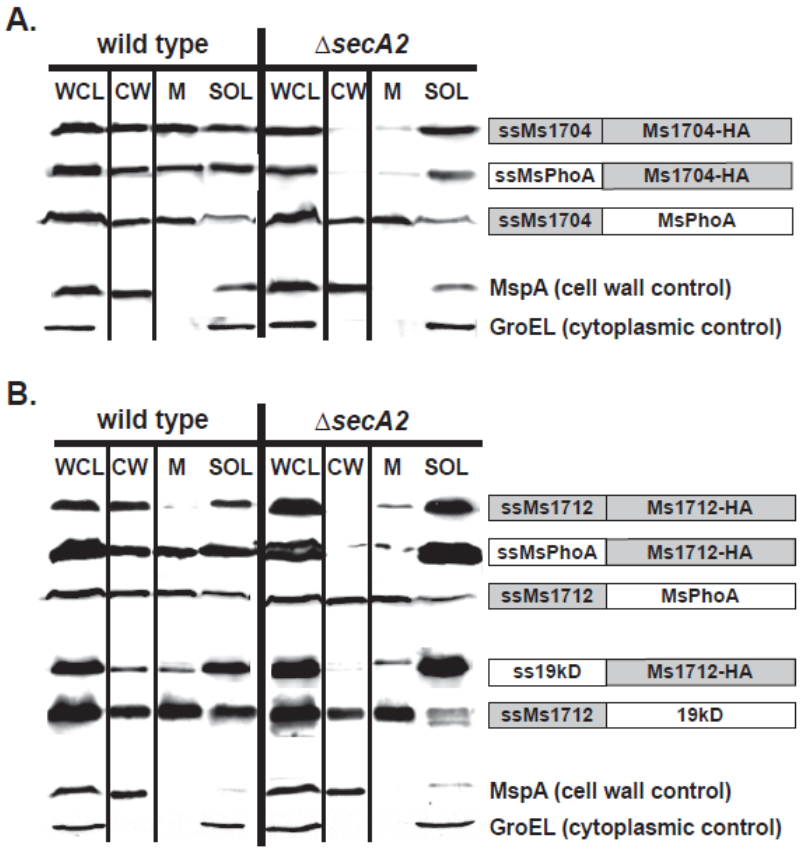

Figure 2.2 The mature domains of Ms1704 and Ms1712 require SecA2 for export to the cell wall.
(A) Equalized whole cell lysates (WCL) generated from wild type and Δ*secA2 M. smegmatis* cells expressing either Ms1704-HA, ssMsPhoA-Ms1704-HA, or ssMs1704-PhoA were subjected to ultracentrifugation to generate subcellular fractions. Fractions were separated by SDS-PAGE, and proteins detected with either anti-HA or anti-PhoA antibodies. The total amount of cell wall (CW), membrane (M), and soluble (SOL) fractions shown is equivalent to the amount of WCL loaded. Native MspA and GroEL were detected as cell wall and cytoplasmic controls respectively. **(B)** Wild type or Δ*secA2 M. smegmatis* cells expressing either Ms1712-HA, ssPhoA-Ms1712-HA, ssMs1712-PhoA, ss19kD-Ms1712-HA, or ssMs1712-19kD were fractionated and material was separated by SDS-PAGE. Immunoblotting was performed as described in (A) with the addition of the anti-19kD antibody when appropriate. The experiment shown is representative of three independent experiments.

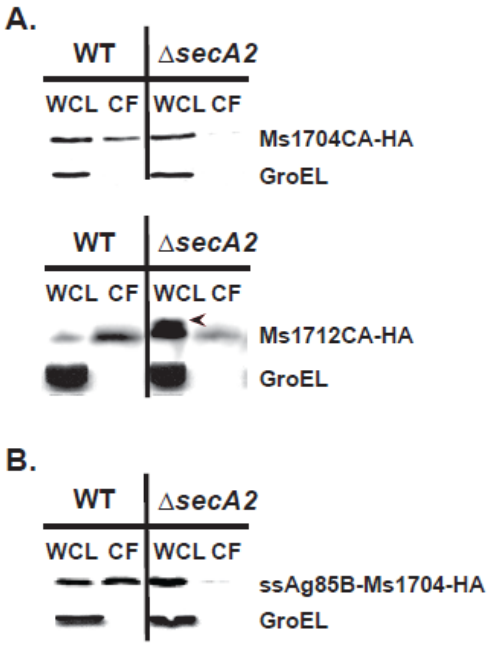

Figure 2.3 SecA2-dependent export occurs regardless of lipidation.
(A) Culture filtrate (CF) proteins from *M. smegmatis* cells expressing either Ms1704CA-HA or Ms1712CA-HA proteins were analyzed by SDS-PAGE and immunoblot using an anti-HA antibody. Whole cell lysate (WCL) and CF material loaded was obtained from an equal number of wild type (WT) or Δ*secA2 M. smegmatis* cells for each experiment. Native GroEL protein was detected as a loading and cell lysis control. Accumulation of presumptive precursor Ms1712CA-HA, containing an uncleaved signal sequence, in the WCL of the Δ*secA2* mutant is noted with an arrow. **(B)** Immunoblot of WCL and CF proteins from wild type and Δ*secA2 M. smegmatis* cells expressing the nonlipidated ssAg85B-Ms1704-HA chimera. The experiments shown are representative of three independent experiments.

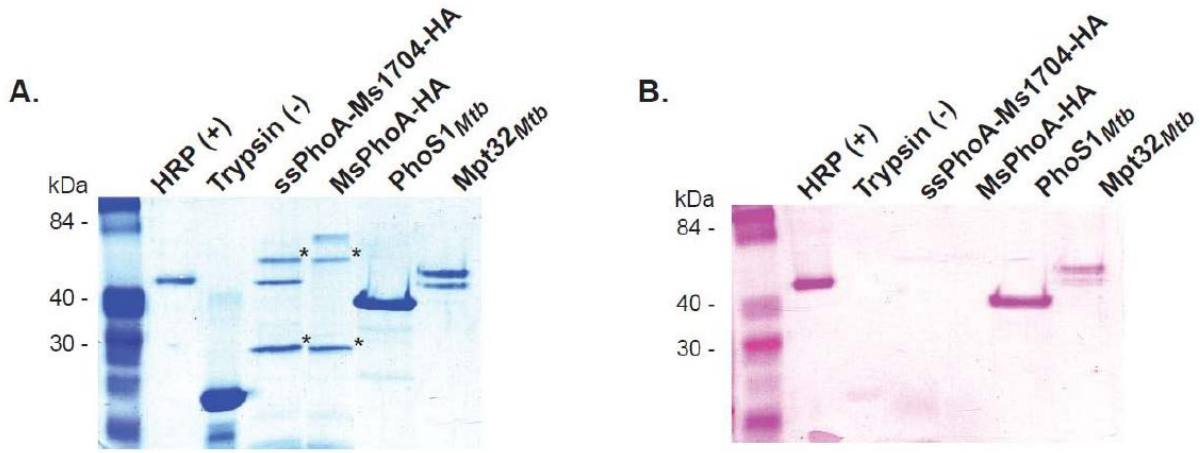

Figure 2.4 Periodic acid-Schiff staining fails to detect glycosylation of the Ms1704 mature domain

(A) Immunoprecipiated ssPhoA-Ms1704-HA or MsPhoA-HA, HRP, trypsin, PhoS1, and Mpt32 proteins were separated on a SDS-PAGE gel and coomassie stained. Asterisks denote the mouse IgG heavy and light chains from the anti-HA beads used for immunoprecipitation. **(B)** An identical SDS-PAGE gel was stained with the Periodic Acid Schiff Stain (PAS) (Pierce) to detect glycoproteins. The only reactive proteins include the positive controls: HRP, PhoS1, and Mpt32. The ssPhoA-Ms1704-HA failed to react with the PAS stain indicating that there is no evidence for glycosylation of the Ms1704 mature domain based on this system.

Figure 2.5 Ms6020-HA is exported to the cell wall independent of SecA2
Equalized whole cell lysates (WCL) generated from wild type and Δ*secA2 M. smegmatis* cells expressing HA-tagged Ms6020 were subjected to ultracentrifugation to subcellular fractions. Ms6020-HA was detected with an anti-HA antibody while native MspA and GroEL were detected as cell wall and cytoplasmic controls respectively. The total amount of cell wall (CW), membrane (M), and soluble (SOL) material shown is equivalent to the amount of WCL loaded. Shown here is a representative of three independent experiments.

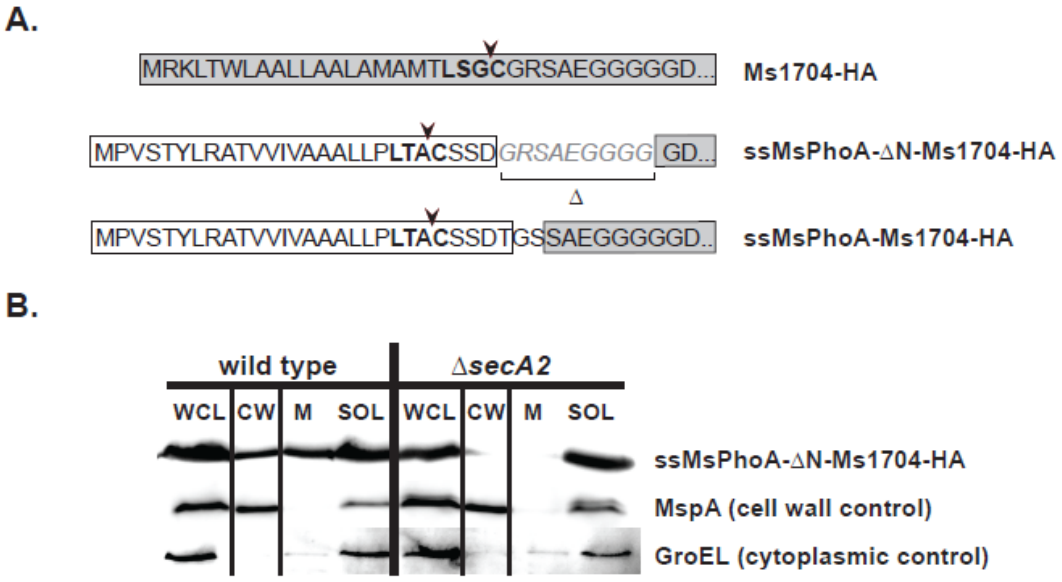

Figure 2.6 The extreme N-terminus of the Ms1704 mature domain is not required for SecA2-mediated export

(A) The first 9 amino acids (italics) were removed from the Ms1704 mature domain, adjacent to the signal peptide cleavage site (arrow), to assess the contribution of this region to SecA2-mediated export. This extreme N-terminus truncation was created in the ssMsPhoA-Ms1704-HA signal sequence chimera, where sequence derived from MsPhoA is boxed in white and sequence from Ms1704 is boxed in grey. **(B)** Equalized whole cell lysates (WCL) generated from wild type and Δ*secA2 M. smegmatis* cells expressing this ssPhoA-ΔN-Ms1704-HA protein were subjected to ultracentrifugation to generate subcellular fractions. Fractions were separated by SDS-PAGE, and proteins detected an anti-HA antibody. The total amount of cell wall (CW), membrane (M), and soluble (SOL) fractions shown is equivalent to the amount of WCL loaded. Native MspA and GroEL were detected as cell wall and cytoplasmic controls respectively. Shown here is a representative of three independent experiments.

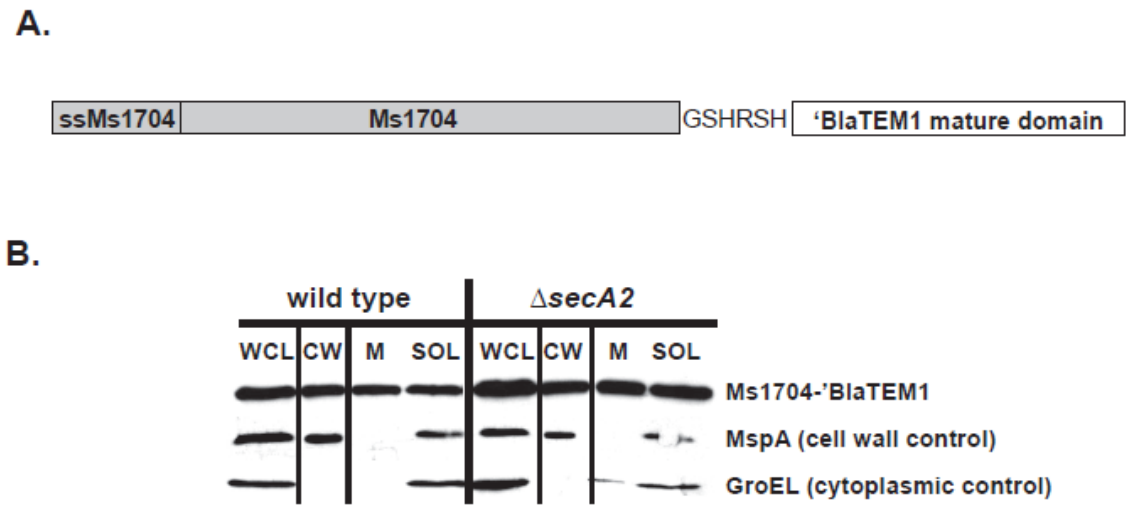

Figure 2.7 An Ms1704-'BlaTEM1 fusion is exported to the cell wall independently of SecA2

(A) The structure of the Ms1704-'BlaTEM1 fusion protein is shown. The signal peptide and entire Ms1704 mature domain constitute the N-terminus of this fusion and are depicted in thegrey rectangle. The *E. coli* 'BlaTEM1 (lacking its native signal sequence) was fused to the C-terminus of Ms1704. Amino acids introduced from cloning are unboxed. **(B)** Equalized whole cell lysates (WCL) generated from wild type and Δ*secA2 M. smegmatis* cells expressing Ms1704-'BlaTEM1 were subjected to ultracentrifugation to generate subcellular fractions. The Ms1704-'BlaTEM1 fusion was detected with an anti-BlaTEM1 antibody while native MspA and GroEL were detected as cell wall and cytoplasmic controls respectively. The total amount of cell wall (CW), membrane (M), and soluble (SOL) fractions shown is equivalent to the amount of WCL loaded. Shown is a representative of two independent experiments.

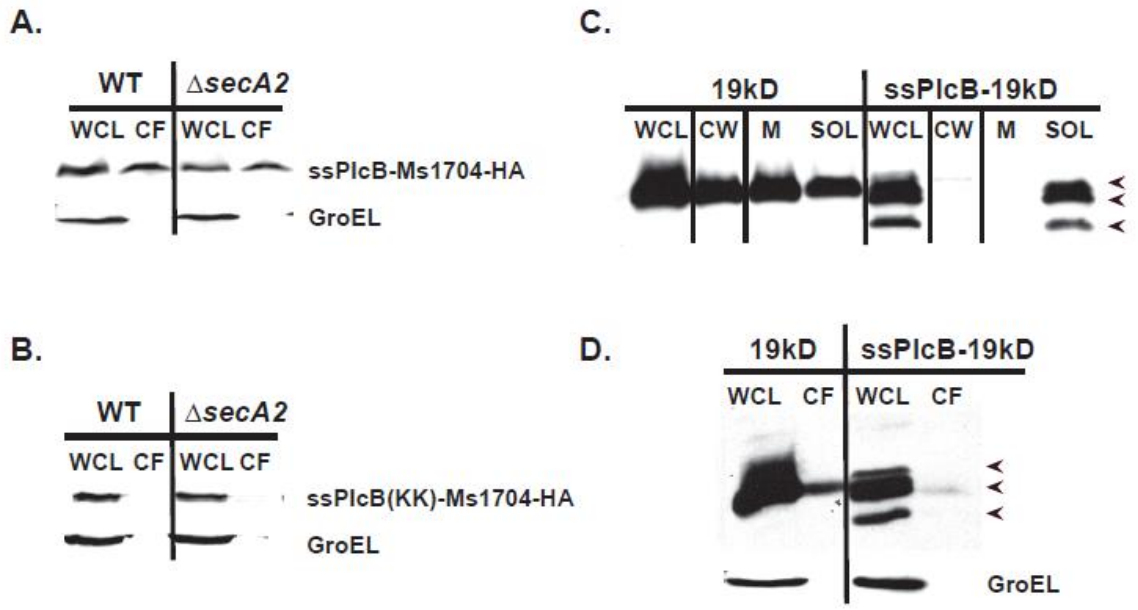

Figure 2.8 Ms1704 is compatible with export by the twin-arginine translocation (Tat) pathway

(A) Concentrated culture filtrate (CF) proteins from *M. smegmatis* cells expressing ssPlcBMs1704-HA were analyzed by SDS-PAGE and immunoblot using an anti-HA antibody. Whole cell lysate (WCL) and CF material loaded was obtained from an equal number of wild type and Δ*secA2 M. smegmatis* cells for each experiment. Native GroEL was detected as a loading and cell lysis control. **(B)** In the signal sequence of the ssPlcB(KK)-Ms1704-HA chimera, the twin arginine residues were changed to lysines to abolish targeting to the Tat machinery. Secretion of this chimera was analyzed as described in (A). **(C)** Export of 19kD containing its native Sec signal sequence and the ssPlcB-19kD chimera were analyzed by SDS-PAGE and immunoblot using an anti-19kD antibody. Equalized whole cell lysates (WCL) generated from wild type *M. smegmatis* cells expressing each protein were subjected to ultracentrifugation to generate cell wall (CW), membrane (M), and soluble (SOL) fractions. **(D)** WCL and culture filtrate (CF) proteins from wild type *M. smegmatis* cells expressing either protein were also analyzed by immunoblot. Wild type 19kD is readily detected in both CW and CF fractions while the ssPlcB-19kD chimera is not exported. The ssPlcB-19kD also exhibits three forms in the WCL and SOL fraction (denoted by arrows), likely representing precursor protein containing an uncleaved signal sequence and cytoplasmic degradation products. All immunoblots are representative of at least two independent experiments

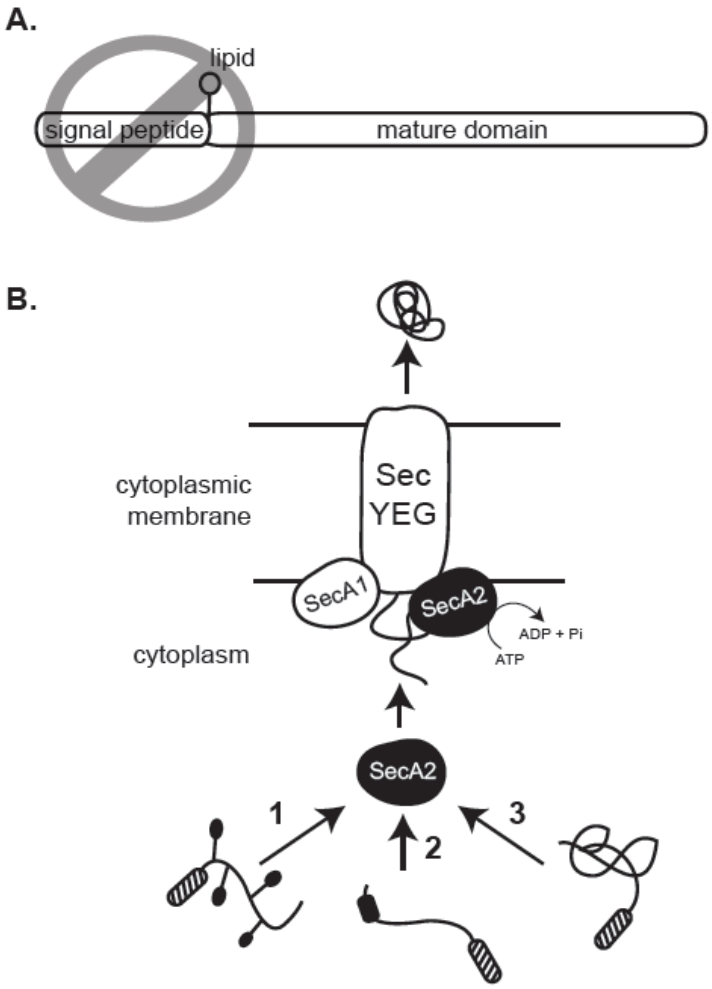

Figure 2.9 Models for mycobacterial SecA2 export
(A) The feature(s) that make a preprotein dependent on SecA2 for export are not contained in the N-terminal signal sequence. Instead, the mature domain of select proteins, such as Ms1704 and Ms1712, impart the requirement for SecA2. Lipidation of the mature domain is not a factor. **(B)** SecA2 is required for the export of certain proteins, likely through the canonical SecYEG channel. One possibility is that glycosylation or another post-translational modification in the mature domain (depicted as attached circles) prevents export by the canonical Sec pathway and/or direct preproteins to SecA2 for export (1). A second possibility is that the mature domain contains an amino acid sequence that directs the preprotein to SecA2 and/or away from SecA1 (depicted as a thick black oval) (2). A third possibility is that the defining feature of SecA2 substrates is a tendency to fold in the cytoplasm (3). Our data is consistent with this last possibility, where SecA2 would then function in maintaining an unfolded preprotein conformation prior to or during export. The signal sequence of preproteins is depicted by the oval with diagonal lines.

References

1. **Alami, M., D. Trescher, L. F. Wu, and M. Muller.** 2002. Separate analysis of twin-arginine translocation (Tat)-specific membrane binding and translocation in *Escherichia coli*. J Biol Chem **277:**20499-20503.

2. **Arkowitz, R. A., J. C. Joly, and W. Wickner.** 1993. Translocation can drive the unfolding of a preprotein domain. Embo J **12:**243-253.

3. **Auclair, S. M., J. P. Moses, M. Musial-Siwek, D. A. Kendall, D. B. Oliver, and I. Mukerji.** 2010. Mapping of the signal peptide-binding domain of *Escherichia coli* SecA using Forster resonance energy transfer. Biochemistry **49:**782-792.

4. **Bassford, P. J., Jr., T. J. Silhavy, and J. R. Beckwith.** 1979. Use of gene fusion to study secretion of maltose-binding protein into *Escherichia coli* periplasm. J Bacteriol **139:**19-31.

5. **Bauer, B. W., and T. A. Rapoport.** 2009. Mapping polypeptide interactions of the SecA ATPase during translocation. Proc Natl Acad Sci U S A **106:**20800-20805.

6. **Bensing, B. A., B. W. Gibson, and P. M. Sullam.** 2004. The *Streptococcus gordonii* platelet binding protein GspB undergoes glycosylation independently of export. J Bacteriol **186:**638-645.

7. **Bensing, B. A., I. R. Siboo, and P. M. Sullam.** 2007. Glycine residues in the hydrophobic core of the GspB signal sequence route export toward the accessory Sec pathway. J Bacteriol **189:**3846-3854.

8. **Bensing, B. A., and P. M. Sullam.** 2002. An accessory sec locus of *Streptococcus gordonii* is required for export of the surface protein GspB and for normal levels of binding to human platelets. Mol Microbiol **44:**1081-1094.

9. **Bensing, B. A., and P. M. Sullam.** 2010. Transport of preproteins by the accessory Sec system requires a specific domain adjacent to the signal peptide. J Bacteriol **192:**4223-4232.

10. **Bensing, B. A., D. Takamatsu, and P. M. Sullam.** 2005. Determinants of the streptococcal surface glycoprotein GspB that facilitate export by the accessory Sec system. Mol Microbiol **58:**1468-1481.

11. **Bensing, B. A., Y. T. Yen, R. Seepersaud, and P. M. Sullam.** 2012. A specific interaction between SecA2 and a region of the preprotein adjacent to the signal peptide occurs during transport via the accessory Sec system. J Biol Chem **287:**24438-24447.

12. **Berks, B. C.** 1996. A common export pathway for proteins binding complex redox cofactors? Mol Microbiol **22:**393-404.

13. **Blaudeck, N., P. Kreutzenbeck, R. Freudl, and G. A. Sprenger.** 2003. Genetic analysis of pathway specificity during posttranslational protein translocation across the *Escherichia coli* plasma membrane. J Bacteriol **185:**2811-2819.

14. **Bordes, P., A. M. Cirinesi, R. Ummels, A. Sala, S. Sakr, W. Bitter, and P. Genevaux.** 2011. SecB-like chaperone controls a toxin-antitoxin stress-responsive system in *Mycobacterium tuberculosis*. Proc Natl Acad Sci U S A **108:**8438-8443.

15. **Braun, V., and H. C. Wu.** 1994. Chapter 14 Lipoproteins, structure, function, biosynthesis and model for protein export, p. 319-341. *In* J. M. Ghuysen and R. Hakenbeck (ed.), New Comp. Biochem., vol. Volume 27. Elsevier.

16. **Braunstein, M., S. S. Bardarov, and W. R. J. Jacobs.** 2002. Genetic methods for deciphering virulence determinants of *Mycobacterium tuberculosis*, p. 67-99. *In* V. L. Clark and P. M. Bavoil (ed.), Methods in Enzymol, vol. 358. Academic Press, London.

17. **Braunstein, M., A. M. Brown, S. Kurtz, and W. R. Jacobs, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. J Bacteriol **183:**6979-6990.

18. **Braunstein, M., B. Espinosa, J. Chan, J. T. Belisle, and W. R. J. Jacobs.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium tuberculosis*. Mol Microbiol **48:**453-464.

19. **Brundage, L., J. P. Hendrick, E. Schiebel, A. J. Driessen, and W. Wickner.** 1990. The purified *E. coli* integral membrane protein SecY/E is sufficient for reconstitution of SecA-dependent precursor protein translocation. Cell **62:**649-657.

20. **Buchanan, G., F. Sargent, B. C. Berks, and T. Palmer.** 2001. A genetic screen for suppressors of *Escherichia coli* Tat signal peptide mutations establishes a critical role for the second arginine within the twin-arginine motif. Arch Microbiol **177:**107-112.

21. **Chen, Q., B. Sun, H. Wu, Z. Peng, and P. M. Fives-Taylor.** 2007. Differential roles of individual domains in selection of secretion route of a *Streptococcus parasanguinis* serine-rich adhesin, Fap1. J Bacteriol **189:**7610-7617.

22. **Chen, Q., H. Wu, and P. M. Fives-Taylor.** 2004. Investigating the role of *secA2* in secretion and glycosylation of a fimbrial adhesin in *Streptococcus parasanguis* FW213. Mol Microbiol **53:**843-856.

23. **Collier, D. N., V. A. Bankaitis, J. B. Weiss, and P. J. Bassford, Jr.** 1988. The antifolding activity of SecB promotes the export of the *E. coli* maltose-binding protein. Cell **53:**273-283.

24. **Cristobal, S., J. W. de Gier, H. Nielsen, and G. von Heijne.** 1999. Competition between Sec- and TAT-dependent protein translocation in *Escherichia coli*. Embo J **18:**2982-2990.

25. **DeLisa, M. P., D. Tullman, and G. Georgiou.** 2003. Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. Proc Natl Acad Sci U S A **100:**6115-6120.

26. **Economou, A., and W. Wickner.** 1994. SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. Cell **78:**835-843.

27. **Eser, M., and M. Ehrmann.** 2003. SecA-dependent quality control of intracellular protein localization. Proc Natl Acad Sci U S A **100:**13231-13234.

28. **Feltcher, M. E., and M. Braunstein.** 2012. Emerging themes in SecA2-mediated protein export. Nat Rev Microbiol **10:**779-789.

29. **Fisher, A. C., and M. P. DeLisa.** 2004. A little help from my friends: quality control of presecretory proteins in bacteria. J Bacteriol **186:**7467-7473.

30. **Fisher, A. C., W. Kim, and M. P. DeLisa.** 2006. Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway. Protein Sci **15:**449-458.

31. **Flores, A. R., L. M. Parsons, and M. S. Pavelka, Jr.** 2005. Genetic analysis of the beta-lactamases of *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* and susceptibility to beta-lactam antibiotics. Microbiology **151:**521-532.

32. **Gelis, I., A. M. Bonvin, D. Keramisanou, M. Koukaki, G. Gouridis, S. Karamanou, A. Economou, and C. G. Kalodimos.** 2007. Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. Cell **131:**756-769.

33. **Gibbons, H. S., F. Wolschendorf, M. Abshire, M. Niederweis, and M. Braunstein.** 2007. Identification of two *Mycobacterium smegmatis* lipoproteins exported by a SecA2-dependent pathway. J Bacteriol **189:**5090-5100.

34. **Glickman, M. S., J. S. Cox, and W. R. Jacobs, Jr.** 2000. A novel mycolic acid cyclopropane synthetase is required for cording, persistence, and virulence of *Mycobacterium tuberculosis*. Mol Cell **5:**717-727.

35. **Gonzalez-Zamorano, M., G. Mendoza-Hernandez, W. Xolalpa, C. Parada, A. J. Vallecillo, F. Bigi, and C. Espitia.** 2009. *Mycobacterium tuberculosis* glycoproteomics based on ConA-lectin affinity capture of mannosylated proteins. J Proteome Res **8:**721-733.

36. **Gouridis, G., S. Karamanou, I. Gelis, C. G. Kalodimos, and A. Economou.** 2009. Signal peptides are allosteric activators of the protein translocase. Nature **462:**363-367.

37. **Hou, J. M., N. G. D'Lima, N. W. Rigel, H. S. Gibbons, J. R. McCann, M. Braunstein, and C. M. Teschke.** 2008. ATPase activity of *Mycobacterium tuberculosis* SecA1 and SecA2 proteins and its importance for SecA2 function in macrophages. J Bacteriol **190:**4880-4887.

38. **Joly, J. C., and W. Wickner.** 1993. The SecA and SecY subunits of translocase are the nearest neighbors of a translocating preprotein, shielding it from phospholipids. EMBO J **12:**255-263.

39. **Juncker, A. S., H. Willenbrock, G. Von Heijne, S. Brunak, H. Nielsen, and A. Krogh.** 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci **12:**1652-1662.

40. **Kimura, E., M. Akita, S. Matsuyama, and S. Mizushima.** 1991. Determination of a region in SecA that interacts with presecretory proteins in *Escherichia coli*. J Biol Chem **266:**6600-6606.

41. **Krehenbrink, M., A. Edwards, and J. A. Downie.** 2011. The superoxide dismutase SodA is targeted to the periplasm in a SecA-dependent manner by a novel mechanism. Mol Microbiol **82:**164-179.

42. **Kriakov, J., S. Lee, and W. R. Jacobs, Jr.** 2003. Identification of a regulated alkaline phosphatase, a cell surface-associated lipoprotein, in Mycobacterium smegmatis. J Bacteriol **185:**4983-4991.

43. **Kumamoto, C. A., and J. Beckwith.** 1985. Evidence for specificity at an early step in protein export in *Escherichia coli*. J Bacteriol **163:**267-274.

44. **Kumamoto, C. A., and J. Beckwith.** 1983. Mutations in a new gene, *secB*, cause defective protein localization in *Escherichia coli*. J Bacteriol **154:**253-260.

45. **Lee, P. A., D. Tullman-Ercek, and G. Georgiou.** 2006. The Bacterial Twin-Arginine Translocation Pathway. Annu Rev Microbiol **60:**373-395.

46. **Lenz, L. L., S. Mohammadi, A. Geissler, and D. A. Portnoy.** 2003. SecA2-dependent secretion of autolytic enzymes promotes *Listeria monocytogenes* pathogenesis. Proc Natl Acad Sci U S A **100:**12432-12437.

47. **Li, W., S. Schulman, D. Boyd, K. Erlandson, J. Beckwith, and T. A. Rapoport.** 2007. The plug domain of the SecY protein stabilizes the closed state of the translocation channel and maintains a membrane seal. Molecular Cell **26:**511-521.

48. **Lill, R., K. Cunningham, L. A. Brundage, K. Ito, D. Oliver, and W. Wickner.** 1989. SecA protein hydrolyzes ATP and is an essential component of the protein translocation ATPase of *Escherichia coli*. EMBO J **8:**961-966.

49. **Lycklama, A. N. J. A., and A. J. Driessen.** 2012. The bacterial Sec-translocase: structure and mechanism. Philos Trans R Soc Lond B Biol Sci **367:**1016-1028.

50. **Matthieu, J. M., and R. H. Quarles.** 1973. Quantitative scanning of glycoproteins on polyacrylamide gels stained with periodic acid-Schiff reagent (PAS). Anal Biochem **55:**313-316.

51. **McCann, J. R., J. A. McDonough, M. S. Pavelka, and M. Braunstein.** 2007. Beta-lactamase can function as a reporter of bacterial protein export during *Mycobacterium tuberculosis* infection of host cells. Microbiology **153:**3350-3359.

52. **McDonough, J. A., K. E. Hacker, A. R. Flores, M. S. Pavelka, Jr., and M. Braunstein.** 2005. The twin-arginine translocation pathway of *Mycobacterium smegmatis* is functional and required for the export of mycobacterial beta-lactamases. J Bacteriol **187:**7667-7679.

53. **Mistou, M. Y., S. Dramsi, S. Brega, C. Poyart, and P. Trieu-Cuot.** 2009. Molecular dissection of the *secA2* locus of group B Streptococcus reveals that glycosylation of the Srr1 LPXTG protein is required for full virulence. J Bacteriol **191:**4195-4206.

54. **Musial-Siwek, M., S. L. Rusch, and D. A. Kendall.** 2007. Selective photoaffinity labeling identifies the signal peptide binding domain on SecA. J Mol Biol **365:**637-648.

55. **Neumann-Haefelin, C., U. Schafer, M. Muller, and H. G. Koch.** 2000. SRP-dependent co-translational targeting and SecA-dependent translocation analyzed as individual steps in the export of a bacterial protein. EMBO J **19:**6419-6426.

56. **Nouwen, N., G. Berrelkamp, and A. J. Driessen.** 2007. Bacterial sec-translocase unfolds and translocates a class of folded protein domains. J Mol Biol **372:**422-433.

57. **Paetzel, M., A. Karla, N. C. Strynadka, and R. E. Dalbey.** 2002. Signal peptidases. Chem Rev **102:**4549-4580.

58. **Palmer, T., and B. C. Berks.** 2012. The twin-arginine translocation (Tat) protein export pathway. Nat Rev Microbiol **10:**483-496.

59. **Papanikou, E., S. Karamanou, C. Baud, M. Frank, G. Sianidis, D. Keramisanou, C. G. Kalodimos, A. Kuhn, and A. Economou.** 2005. Identification of the preprotein binding domain of SecA. J Biol Chem **280:**43209-43217.

60. **Petersen, T. N., S. Brunak, G. von Heijne, and H. Nielsen.** 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods **8:**785-786.

61. **Randall, L. L., and S. J. Hardy.** 1986. Correlation of competence for export with lack of tertiary structure of the mature species: a study in vivo of maltose-binding protein in *E. coli*. Cell **46:**921-928.

62. **Rengarajan, J., E. Murphy, A. Park, C. L. Krone, E. C. Hett, B. R. Bloom, L. H. Glimcher, and E. J. Rubin.** 2008. *Mycobacterium tuberculosis* Rv2224c modulates innate immune responses. Proc. Natl. Acad. Sci. U.S.A. **105:**264-269.

63. **Richter, S., and T. Bruser.** 2005. Targeting of unfolded PhoA to the TAT translocon of *Escherichia coli*. J Biol Chem **280:**42723-42730.

64. **Rigel, N. W., H. S. Gibbons, J. R. McCann, J. A. McDonough, S. Kurtz, and M. Braunstein.** 2009. The accessory SecA2 system of mycobacteria requires ATP binding and the canonical SecA1. J Biol Chem **284:**9927-9936.

65. **Roberts, A. D., J. T. Belisle, A. M. Cooper, and I. M. Orme.** 1998. Specific Models: Murine Model of Tuberculosis, p. 389-417. *In* S. H. Kaufmann and D. Kabeltz (ed.), Methods in Microbiology, vol. 25.

66. **Sardis, M. F., and A. Economou.** 2010. SecA: a tale of two protomers. Mol Microbiol **76:**1070-1081.

67. **Sargent, F., E. G. Bogsch, N. R. Stanley, M. Wexler, C. Robinson, B. C. Berks, and T. Palmer.** 1998. Overlapping functions of components of a bacterial Sec-independent protein export pathway. Embo J **17:**3640-3650.

68. **Siboo, I. R., D. O. Chaffin, C. E. Rubens, and P. M. Sullam.** 2008. Characterization of the accessory Sec system of *Staphylococcus aureus*. J Bacteriol **190:**6188-6196.

69. **Smith, M. A., W. M. Clemons, Jr., C. J. DeMars, and A. M. Flower.** 2005. Modeling the effects of prl mutations on the *Escherichia coli* SecY complex. J Bacteriol **187:**6454-6465.

70. **Snapper, S. B., R. E. Melton, S. Mustafa, T. Kieser, and W. R. Jacobs, Jr.** 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. Mol Microbiol **4:**1911-1919.

71. **Stanley, N. R., T. Palmer, and B. C. Berks.** 2000. The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. J Biol Chem **275:**11591-11596.

72. **Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennett, G. P. Bansal, J. F. Young, M. H. Lee, and G. F. Hatfull.** 1991. New use of BCG for recombinant vaccines. Nature **351:**456-460.

73. **Titgemeyer, F., J. Amon, S. Parche, M. Mahfoud, J. Bail, M. Schlicht, N. Rehm, D. Hillmann, J. Stephan, B. Walter, A. Burkovski, and M. Niederweis.** 2007. A genomic view of sugar transport in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. J Bacteriol **189:**5903-5915.

74. **Tschumi, A., T. Grau, D. Albrecht, M. Rezwan, H. Antelmann, and P. Sander.** 2012. Functional analyses of mycobacterial lipoprotein diacylglyceryl transferase and comparative secretome analysis of a mycobacterial *lgt* mutant. J Bacteriol **194:**3938-3949.

75. **van der Wolk, J. P., J. G. de Wit, and A. J. Driessen.** 1997. The catalytic cycle of the *Escherichia coli* SecA ATPase comprises two distinct preprotein translocation events. EMBO J **16:**7297-7304.

76. **von Heijne, G.** 1990. The signal peptide. J Membr Biol **115:**195-201.

77. **Weiss, J. B., P. H. Ray, and P. J. Bassford, Jr.** 1988. Purified *secB* protein of *Escherichia coli* retards folding and promotes membrane translocation of the maltose-binding protein *in vitro*. Proc Natl Acad Sci U S A **85:**8978-8982.

78. **Wiker, H. G., and M. Harboe.** 1992. The antigen 85 complex: a major secretion product of *Mycobacterium tuberculosis*. Microbiol Rev **56:**648-661.

79. **Wolschendorf, F., M. Mahfoud, and M. Niederweis.** 2007. Porins are required for uptake of phosphates by *Mycobacterium smegmatis*. J Bacteriol **189:**2435-2442.

80. **Young, D. B., and T. R. Garbe.** 1991. Lipoprotein antigens of *Mycobacterium tuberculosis*. Res Microbiol **142:**55-65.

CHAPTER 3

Identification of *Mycobacterium tuberculosis* Proteins that Depend on SecA2 for Cell Wall Localization[1]

Mycobacterium tuberculosis is an intracellular pathogen that is able to survive in host macrophages by blocking phagosome maturation. This immune evasion strategy is a complex process involving multiple exported bacterial protein effectors. One protein export pathway that contributes to interrupting maturation of the *M. tuberculosis* phagosome is the SecA2 system. Mycobacteria have two non-redundant SecA homologs. While the essential SecA1 handles "housekeeping" protein export, the nonessential SecA2 exports a subset of proteins and is required for *M. tuberculosis* virulence. Identification of SecA2-dependent exported proteins will help elucidate the role of SecA2 in *M. tuberculosis* virulence, which is not currently understood. Here, we used a label-free mass spectrometry method called spectral counting to identify *M. tuberculosis* proteins that require SecA2 for export to the cell wall. We identified a total of 1,920 cell wall-associated *M. tuberculosis* proteins. Proteins underrepresented in the Δ*secA2* cell wall included the majority of predicted solute-binding proteins and multiple proteins belonging to two transporter complexes with roles in *M. tuberculosis* virulence:

[1]Contributing authors: Harsha P. Gunawardena and Xian Chen (Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, NC, 27599) and Miriam Braunstein (Department of Microbiology and Immunology, School of Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7290)

Mce1 and Mce4. Several predicted substrates of the twin-arginine translocation (Tat) pathway were also reduced in the cell wall of the Δ*secA2* mutant. Because the Tat pathway only exports folded preproteins, this suggests that the mycobacterial SecA2 and Tat systems might assist in export of a common pool of preprotein substrates that fold in the cytoplasm. Additionally, several proteins were more abundant in the Δ*secA2* mutant, suggesting that they were up-regulated in the mutant strain. Our semi-quantitation analyses generated a list of potential SecA2 substrates, some of which are candidates for *M. tuberculosis* virulence factors.

Introduction

Mycobacterium tuberculosis, the etiological agent of tuberculosis, is one of the most prevalent infectious agents worldwide, infecting an estimated 1/3 of the global population (130). Emergence of drug-resistant strains of *M. tuberculosis* combined with co-infection with HIV has resulted in the highest mortality rate of any infectious bacteria, with over 1 million tuberculosis-related deaths reported annually (130). One of the keys to the success of *M. tuberculosis* as a pathogen is its ability to survive and grow inside host cells as an intracellular pathogen (93).

Upon inhalation into the lung, *M. tuberculosis* bacilli are engulfed by resident alveolar macrophages. These macrophages normally sequester bacteria inside phagosomal compartments that undergo a process of maturation and acidification, eventually leading to the destruction of the invading bacterial cell (23). Many intracellular pathogens, such as *Legionella pneuomophila*, *Salmonella typhimurium*, and *M. tuberculosis*, have evolved mechanisms to circumvent this innate host defense

pathway. Often, such bacterial pathogens export protein effectors into the macrophage environment to promote their intracellular survival and growth (43).

M. tuberculosis survives in macrophages by interrupting phagosomal maturation using a multi-factorial process involving both bacterial surface lipids and exported protein effectors (93). While several *M. tuberculosis* exported effectors have been linked to phagosomal maturation arrest, our understanding of this complex process is incomplete, as the list of bacterial effectors and corresponding host targets continues to grow. One strategy to better understand how *M. tuberculosis* survives and grows in the host is to study protein export pathways that are required for intracellular growth in macrophages. In *M. tuberculosis*, the specialized SecA2 protein export pathway is required for growth in macrophages and, more specifically, for phagosome maturation arrest (120). However, the SecA2-dependent exported proteins required for interrupting phagosomal maturation are unknown.

Mycobacteria contain two conserved protein export systems for transporting proteins from the cytosol to the cell wall and/or the extracellular environment: the universal general secretion (Sec) pathway and the twin-arginine translocation (Tat) pathway (67). In *M. tuberculosis*, both Sec and Tat export are essential for in vitro growth and also contribute to survival in the host (96, 105, 106). The Sec pathway handles the majority of protein export while the Tat pathway is responsible for exporting a more limited set of proteins. The core of the general Sec export pathway is the protein-conducting SecYEG membrane channel and a cytoplasmic ATPase called SecA (13). SecA recognizes proteins synthesized with N-terminal signal peptides and powers their translocation through SecYEG. Mycobacteria, including *M. tuberculosis*, are among a

select group of bacteria that have two non-redundant SecA protein homologs known as SecA1 and SecA2 (24, 99). While SecA1 is essential and required for powering housekeeping export as part of the canonical Sec pathway, SecA2 is dispensable for in vitro growth but necessary for exporting a subset of proteins (24).

In *M. tuberculosis*, SecA2 is required for growth in macrophages and mice, suggesting that at least some SecA2-dependent exported proteins are virulence factors (7, 63). Deletion of *secA2* is linked to multiple phenotypes during macrophage infection with *M. tuberculosis* including increased host cell apoptosis, increased induction of pro-inflammatory cytokines, and an inability to block phagosome maturation (44, 63, 120). Understanding how SecA2 influences these macrophage responses requires a complete understanding of the proteins exported by the SecA2 system.

In a previous study, comparative 2-dimensional gel electrophoresis (2D-GE) analysis of proteins secreted into the culture media by wild type *M. tuberculosis* or the Δ*secA2* mutant was used to identify SecA2-dependent proteins (7). Only a small number of proteins were found to be secreted less into culture media by the Δ*secA2* mutant. One protein was later confirmed to require SecA2 for export: the Fe-superoxide dismutase, SodA (7, 44). Interestingly, SodA lacks a recognizable signal peptide for export. Subsequent studies showed that inadequate SodA secretion is not sufficient to explain the growth defect of the Δ*secA2* mutant in macrophages (63, 120). Therefore, there must be additional SecA2-dependent exported proteins important for intracellular growth of *M. tuberculosis* awaiting identification.

In the non-pathogenic *Mycobacterium smegmatis* model species, two cell wall proteins were identified as dependent on SecA2 for export: Ms1704 and Ms1712 (34).

These proteins are predicted solute-binding lipoprotein (SBP) components of two ABC transporters with putative roles in sugar uptake, and both these proteins contain typical N-terminal Sec signal peptides for export (84, 90). The SecA2 systems of *M. tuberculosis* and *M. smegmatis* are functionally conserved, suggesting that in *M. tuberculosis* there will also be cell wall proteins that require SecA2 and are synthesized with N-terminal signal peptides (100).

Past efforts to identify mycobacterial SecA2 substrates in the cell wall were subject to the technical limitations of 2D gel comparisons, in particular, the difficulties associated with separating hydrophobic cell wall proteins by 2D-GE (7, 8, 34). Large-scale "shot-gun" proteomic analyses offer an attractive alternative to 2D-GE for comparing large complex protein samples, and spectral counting is one increasingly popular label-free method for determining relative protein abundance in such studies. Spectral counting quantitation is based on the observation that the total number of MS spectra assigned to a particular protein correlates with the abundance of that protein (69). Spectral counting has previously been used to identify exported proteins in *Pseudomonas*, *Porphyromonas*, *Shigella*, and *Salmonella* (46, 59, 62, 85).

Here we compared the proteome of the wild type *M. tuberculosis* cell wall to that of a Δ*secA2* mutant using shot-gun proteomics and spectral counting. Our analysis led to the identification of 1,920 *M. tuberculosis* cell wall-associated proteins. Proteins exported to the cell wall by the SecA2-dependent pathway should be underrepresented in the cell wall of the Δ*secA2* mutant versus wild type strain. We identified 124 proteins that met our criteria for being significantly reduced or absent in the Δ*secA2* mutant cell wall.

Among the proteins underrepresented in the *M. tuberculosis* Δ*secA2* mutant were multiple lipoproteins, including the majority of predicted solute-binding proteins, which is significant in revealing similarity with the *M. smegmatis* SecA2 system. The broad impact of *secA2* deletion on export of SBPs in *M. tuberculosis* also suggests a possible role for SecA2 export in nutrient acquisition inside the host. Additionally, multiple proteins belonging to two transporter complexes with roles in *M. tuberculosis* virulence, Mce1 and Mce4, were underrepresented in the Δ*secA2* mutant cell wall. Finally, several predicted Tat substrates were less abundant in the Δ*secA2* mutant, most of which have putative SBP functions. Our list of cell wall-associated proteins affected by *secA2* helps provide insight into the mechanisms of mycobacterial SecA2 export and the contribution of this protein export system to *M. tuberculosis* virulence.

Materials and Methods

***M. tuberculosis* growth conditions for cell wall analysis:** For analysis of the cell wall proteome, *M. tuberculosis* H37Rv (wild type) and $mc^2$3112 (Δ*secA2*) were first grown at 37° C in liquid Middlebrook 7H9 medium (Difco) supplemented with 0.05% Tyloxapol (Sigma), 0.5% glycerol, and 1X ADS [0.5% bovine serum albumin, 0.2% glucose, 0.85% NaCl]. After cultures reached an OD_{600} of 2 or 3, cells were harvested by centrifugation and washed two times with modified 7AGT medium. The washed cells were used to inoculate three independent 100 mL cultures at an OD_{600} of 0.08 using the modified 7AGT medium, resulting in three replicates for each *M. tuberculosis* strain. This modified medium consisted of 7H9 supplemented with 0.1% Tyloxapol, 0.1% glycerol, 0.5% bovine serum albumin, and 1 mM proprionic acid (Sigma). The media was buffered

using 100 mM 2-(4-morpholino)-ethane sulfonic acid (MES) and the pH adjusted to 6.5 (92). Cultures were grown at 37° C to an OD_{600} of approximately 1, and then harvested by centrifugation. Cell pellets were washed with 1X phosphate-buffered saline (PBS) then sterilized by gamma-irradiation for 27 hours in a JL Shephard Mark I 137Cs irradiator (Dept. of Radiobiology, University of North Carolina at Chapel Hill) for removal from BSL-3 containment.

Preparation of subcellular fractions: Cell wall fractions were isolated by centrifugation as previously described (38). Briefly, cells suspended in 1X PBS containing a cocktail of protease inhibitors were lysed by four passages through a french press cell. Unlysed cells were removed by centrifugation at 3,000 x g to generate clarified whole cell lysates (WCLs). Protein concentrations were determined by BCA assay (Pierce) using a BSA standard and equalized among both strains and all replicates. Equalized WCLs were spun at 27,000 x g for 30 minutes to obtain the cell wall fraction. The supernatant was then spun at 100,000 x g for 2 hours to isolate the cytoplasmic membrane fraction from the remaining soluble fraction. Cell wall and membrane fractions were washed once and suspended in PBS. Prior to proteomic analysis, protein concentrations of the cell wall fractions were determined by BCA assay and equalized among all replicates of both strains.

In-gel trypsin digestion of cell wall proteins: Cell wall proteins from the three biological replicates of *M. tuberculosis* H37Rv and $mc^2$3112 were separated on a precast 12% SDS-PAGE gel. For each sample, 34 mg of protein was loaded in an individual lane for separation. Protein bands were visualized by Coomassie Blue R-250 staining (Bio-Rad, Hercules, CA) and the lane for each biological sample was cut into 32 equivalent

gel slices, approximately 2 mm thick. Trypsin digestion of the excised gel slices was performed as previously described with modifications (17). Briefly, each gel slice was cut into smaller pieces and placed in a single well of a 96-well U-bottom polypropylene plate containing HPLC-grade water. After de-staining with multiple acetonitrile (ACN) washes, slices were incubated in 25 mM ammonium bicarbonate (ABC) containing 20 μg/mL of sequencing-grade trypsin overnight at 37° C. The ABC supernatants from overnight trypsin incubation were then transferred to a new 96-well plate. Tryptic peptides were extracted from the gel slices by two washes using 50% ACN then 100% ACN. The ACN washes containing the tryptic peptides were added to the overnight ABC supernatant in the fresh plate. Digested peptides were store at -80° C until lyophilization.

Mass spectrometry analysis: Samples were desalted using PepClean C18 spin columns (Pierce, Rockford, IL) according to manufacturer's directions, and re-suspended in an aqueous solution of 0.1% formic acid. All samples were analyzed by reversed phase LC-MS/MS using a 2D-nanoLC ultra system (Eksigent Inc, Dublin, CA) coupled to an LTQ-Orbitrap XL system with ETD (Thermo Scientific, San Jose, CA). The Eksigent system was configured to trap and elute peptides in 1D mode of operation via a sandwiched injection of ~ 250 fmol of sample. The trapping was performed on a 3 cm long 100 μm i.d. C18 column while elution was performed on a 15 cm long 75 μm i.d., 5 μm, 300Å particle; ProteoPep II integraFrit C18 column (New Objective Inc, Woburn, MA). Analytical separation of all the tryptic peptides was achieved with a linear gradient of 2-40% buffer B over either 120 min at a 200 nL/min flow rate, where buffer A is aqueous solution of 0.1% formic acid and buffer B is a solution of acetonitrile in 0.1% formic acid. Mass spectrometric data acquisition was performed in a data dependent manner on a

hybrid LTQ-Orbitrap mass spectrometer. A full scan mass analysis on an Orbitrap (externally calibrated to a mass accuracy of < 1 ppm, and a user defined resolution of 60 000) was followed by intensity dependent MS/MS of the top 10 most abundant peptide ions in the collision cell of the mass spectrometer. High energy collision activated dissociation (HCD)-MS/MS spectra were obtained for peptide ions that were subjected to 10 m.s. ion activation at a normalized collision energy of 35 eV. The MS/MS acquisition of a precursor m/z was repeated for a 30 s. duration and subsequently excluded for 60 s. Monoisotopic precursor ion selection and charge state screening was enabled for triggering data dependent MS/MS scans.

Peptide identification and assignment of proteins: Mass spectra were processed, and peptide identification was performed using Mascot ver. 2.3 (Matrix Science Inc.) implemented on Proteome Discoverer ver. 1.3 (Thermo-Fisher Scientific). All searches were performed against a NCBI *M. tuberculosis* H37Rv data base with methionine oxidation as a variable modification. Peptides were confidently identified using a target-decoy approach with a false discover rate (FDR) of 0.01. A precursor ion mass tolerance of 200 p.p.m. and a product ion mass tolerance 0.5 Da were used during the search to increase search space and reduce false positive identifications with a maximum of two missed trypsin cleavages. All spectra were post-search filtered and peptides having a mass accuracy of 6 p.p.m or less were reported for all the identified proteins.

Peptide validation and label-free spectral count quantitation: Label-free quantitation was based on spectral counting was performed on the Mascot DAT files searched against a forward *M. tuberculosis* database as-well-as a decoy *M. tuberculosis* database using ProteoIQ: ver 2.3.02 (NuSep Inc., Athens, GA). The decoy database search was used to

independently asses FDR (error rate) of the identified spectra in ProteoIQ. All Mascot protein identifications were subjected to probability-based confidence measurements using an independent implementation of the statistical models commonly known as Peptide and Protein Prophet deployed in ProteoIQ (53, 82). All protein hits were filtered with a protein probability of 0.5 and Mascot identity having a significant score cut-off that was greater than 26. Label-free quantification was performed on the pooled DAT files of three replicate runs for both wild type and Δ*secA2* sample groups. Protein ratios between wild type and Δ*secA2* mutants were obtained using average spectral count for each identified protein. Average spectral counts were normalized by total spectral counts as part of ProteoIQ software. Protein quantification values were reported with F-test statistics with a *p* value cut-off of 0.05. We have also used a 2-fold change in $\log_2$ (H37Rv/Δ*secA2*) ratio as cut-off for changes in protein abundance. Q values were obtained using the QVALUE software developed by Alan Dabney and John Storey (118).

Immunoblotting: Equalized whole cell lysates (WCLs) derived from replicates of wild-type or *secA2* mutant *M. tuberculosis* strains were separated on a 12% SDS-PAGE gel, then transferred to a nitrocellulose membrane (Whatman). Native proteins were detected using the following antibodies: anti-SecA1 at 1:50,000 (39), anti-SecA2 at 1:25,000 (39), or anti-19kD (provided by Douglas Young) at 1:20,000. For immunoblotting of subcellular fractions, native proteins were detected using either anti-MtrA at 1:2000 (29), anti-SigA at 1:15,000 (81), anti-HbHA at 1:1000 (NIH Biodefense and Emerging Infections Research Resources Repository, NIAID, NIH), anti-19kD at 1:20,000 (provided by Douglas Young), or anti-PhoS1 at 1:1,000 (BEI Resources). Anti-mouse and anti-rabbit secondary antibodies conjugated to horse radish peroxidase (BioRad)

were used as appropriate at a concentration of 1:20,000 and detected using Western Lightning Chemiluminescent detection reagent (Perkin-Elmer). For quantitation of the PhoS1 immunoblot, secondary antibody conjugated to alkaline phosphatase (GE Healthcare) was used and detected using or ECF reagent (GE Healthcare). Fluorescence was quantified using ImageJ (NIH).

Results

Our appreciation of the proteins exported by the SecA2 system of *M. tuberculosis* remains incomplete. Previous efforts to identify SecA2-dependent proteins were subject to limitations of 2D-GE analysis. Quantitative "shotgun" proteomics enables a more comprehensive assessment of complex protein samples than 2D-GE. Consequently, to identify additional examples of proteins exported in a SecA2-dependent fashion, we used shotgun proteomics combined with label-free spectral counting to measure the relative abundance of proteins in cell wall fractions of wild-type versus a Δ*secA2* deletion mutant of *M. tuberculosis*.

Growth of wild type and secA2 *mutant* M. tuberculosis *for cell wall proteome analysis*

It is well known that bacteria alter gene expression in response to environmental triggers associated with infection. These transcriptional responses can influence expression of genes encoding proteins destined for export. For example, temperature and aerobic conditions influence expression of some SecA2-dependent proteins in *Listeria* (9, 98). The protein export systems themselves can also be regulated by host conditions. As an example, the Sp-II protein secretion system is activated when *Salmonella typhimurium* is located inside the macrophage (15).

M. tuberculosis also undergoes transcriptional adaptations upon macrophage infection (30, 101, 109). Therefore, to increase the possibility that we detect SecA2-dependent substrates important to virulence we modified standard laboratory growth media to be more reflective of the phagosomal environment encountered by *M. tuberculosis*. Although *M. tuberculosis* arrests the normal process of phagosome maturation, the bacterium is exposed to mildly acidic conditions (27, 101). Therefore, we adjusted conventional mycobacterial 7H9 media to a slightly acidic pH of 6.5 to mimic the phagosomal pH *M. tuberculosis* experiences during infection (119). Among the many *M. tuberculosis* genes induced inside a macrophage are those involved in fatty acid metabolism (109) and it is now clear that *M. tuberculosis* preferentially uses fatty acids for carbon and energy sources during infection (72, 76). For this reason, we also provided the fatty acid propionate as a carbon source (49). Finally, to avoid possible growth problems associated with degradation products of Tween detergent we instead utilized the non-hydrolyzable Tyloxapol as the dispersing agent in liquid cultures (127). Triplicate cultures of *M. tuberculosis* H37Rv (wild type) and $mc^2$3112 (Δ*secA2*) were grown in this modified media then cells harvested during log phase at an OD_{600} of approximately 1.0 for preparation of cell wall fractions.

Preparation of* M. tuberculosis *cell wall proteins for MS analysis

Whole cell lysates were generated from triplicate cultures of wild type or Δ*secA2* mutant, assayed for protein concentration, and equalized among samples prior to subcellular fractionation. The equalized whole cell lysates (WCLs) were further analyzed by immunoblotting to ensure equivalent total protein levels between all strains and

replicates (Figure 3.1A). All lysates had equivalent levels of the cytoplasmic SecA1 protein and the exported 19kD lipoprotein, indicating that total proteins levels were equivalent between strains. As expected, the SecA2 protein was only detected in wild type H37Rv samples and not in samples from the Δ*secA2* deletion mutant.

These equalized whole cell lysates (WCL) were then subjected to differential ultracentrifugation to generate cell wall (CW), cytoplasmic membrane (M), and cytosolic-containing soluble (SOL) fractions. The WCL samples and corresponding subcellular fractions were analyzed for fraction integrity by immunoblotting using antibodies to known *M. tuberculosis* cytoplasmic or cell wall associated proteins. Two cytoplasmic proteins, the housekeeping sigma factor SigA and response regulator MtrA were only detected in the soluble fraction as expected (Figure 3.1B). Two known cell wall proteins, the 19 kD lipoprotein and the heparin-binding haemagglutinin adhesin (HBHA) were only detected in the cell wall fraction as expected. Notably, both of the cell wall proteins were detected at similar levels in the two *M. tuberculosis* strains. The cell wall proteins from each biological replicate were also separated by SDS-PAGE and Coomassie stained for protein visualization (Figure 3.2). Protein bands exhibited equivalent intensity between each strain. Together, the immunoblotting and Coomassie stained gels showed that we had prepared quality cell wall fractions with relatively equal amounts of protein from both wild type *M. tuberculosis* and the Δ*secA2* mutant samples.

M. tuberculosis *cell wall protein LC-MS analysis*

Initial efforts to identify SecA2-dependent cell wall proteins in *M. tuberculosis* used 2D-GE and were unsuccessful (Gibbons, H.S. unpublished). However, the high

insolubility of cell wall proteins can be a significant technical challenge for 2D-GE. As an alternate strategy, here we combined SDS-solubilization and 1D SDS-PAGE separation of cell wall proteins with LC-MS/MS label-free quantitative mass spectrometry to comprehensively identify and compare the cell wall proteomes of wild type and the Δ*secA2* mutant. Following SDS-PAGE of each of the replicate samples, cell wall proteins were excised for the entire lane in 32 slices and subjected to trypsin digestion. Tryptic peptides were then analyzed by LC-MS/MS to generate a complete proteomic profile for each sample.

In total, we identified 1,920 cell wall-associated proteins between the fractions of wild type *M. tuberculosis* and the Δ*secA2* mutant, representing diverse functional categories (Figure 3.3A and Appendix). The largest group of proteins identified (50%) are predicted to be involved in cell wall processes or intermediary metabolism, while the next most abundant group represents conserved hypotheticals (23%) (11, 16). The large representation of these three broad functional protein classes is consistent with what has previously been reported for the *M. tuberculosis* cell wall proteome identified by MS analysis (73).

Of the 1,920 proteins identified, 27% are predicted to contain a signal for export out of the cytoplasm (Figure 3.3B). These include 370 proteins predicted by the TMHMM algorithm to contain one or more transmembrane domains for insertion in the cytoplasmic membrane (116). Additionally, 134 proteins are predicted by SignalP to contain an N-terminal Sec signal peptide, which would target them for export out of the cytoplasm by the general Sec pathway (90). Another 33 proteins are predicted by the TatP program to contain Tat signal peptides for export by the Tat pathway (3). On the

basis of a lipobox motif in the N-terminus, we identified 74 potential lipoproteins, which represent 72% of the total number of lipoproteins predicted in the H37Rv genome (6, 51, 121). Although all these lipoproteins contain potential lipoboxes, some of these lipoproteins have predicted Sec or Tat signal peptides, while others have potential signal peptides not recognized by the SignalP or TatP prediction programs (121). In Gram positive bacteria, lipoproteins are anchored to the inner membrane via their lipid moiety, while in Gram negative bacteria lipoproteins can be anchored in either the inner or outer membrane (47). In mycobacteria it is unclear where lipoproteins localize, either the inner membrane or the outer mycolate-based cell wall layer of the cell envelope, which is the fraction analyzed here (47, 121). Indeed, there are mycobacterial lipoproteins shown to have a peripheral localization (either secreted or associated with in the cell wall), which makes the high number of lipoproteins identified in the cell wall here not unexpected (34, 64, 65, 83, 131).

It may be surprising that 73% of the proteins identified in the cell wall fraction do not contain a predicted transmembrane or signal peptide for export However, past proteomic studies of both the mycobacterial cytoplasmic membrane and cell wall fractions have also identified an abundance of proteins lacking predicted export signals and may not be exported (41, 70, 73). Although care is taken to wash the cell wall fraction, the sensitivity of mass spectrometry inevitably detects some level of cytoplasmic contaminants, particularly the most abundant cytoplasmic proteins. Additionally, it is difficult to separate the cytoplasmic membrane and cell wall fractions. Therefore, some non-exported cytoplasmic proteins may peripherally associate with membrane-bound protein complexes or membrane lipids, contributing to co-fractionation with the cell wall

material. Finally, it is also possible that some exported proteins simply lack a predictable signal peptide for export.

Label-free quantitation and comparison of **M. tuberculosis** *and* **ΔsecA2** *cell wall proteomes*

We next wanted to determine which proteins were differentially localized between the cell wall proteomes of wild type *M. tuberculosis* and the Δ*secA2* mutant, with the primary goal of identifying SecA2-dependent cell wall proteins. We employed a label-free semi-quantitative mass spectrometry method called spectral counting to obtain relative protein amounts between strains. Using the ProteoIQ program (ver. 2.3.02, NuSep), we compared the total number of MS/MS spectra associated with each identified protein between wild type *M. tuberculosis* and the Δ*secA2* mutant. Although care was taken to begin with equivalent amounts of total protein for each strain and biological replicate, we first performed a normalization step across the entire experiment (86). The ProteoIQ program first normalized the spectral count for each protein within a replicate by the total spectral counts obtained for all proteins from that particular replicate. Then, the spectral count for that protein is normalized by the maximum number of total spectral counts obtained for wild type versus the Δ*secA2* mutant.

The normalized spectral counts were then compared between the two *M. tuberculosis* strains using the ProteoIQ software to produce a $\log_2$ (Δ*secA2*/WT) ratio for each protein, where ratios of ± 1 represent a 2-fold change. In total, we obtained relative abundance ratios for 1,780 proteins, most of which (89%) were relatively unchanged between the strains with $\log_2$ (Δ*secA2*/WT) ratios between $\pm$ 1. This is expected as past

2D-GE comparisons of *M. tuberculosis* or *M. smegmatis secA2* mutant fractions revealed few differences when compared to the respective wild type strains (7, 34).

Statistical significance for changes in protein abundance was determined with the ProteoIQ program. ProteoIQ assigns a standard *p* value for each $\log_2$ (Δ*secA2*/WT) ratio based on ANOVA testing, which takes into account the reproducibility of the relative abundance of each protein across all groups. The volcano plot in Figure 3.4 displays the $\log_2$ (Δ*secA2*/WT) ratios and *p* values assigned to each of the 1,780 proteins subjected to semi-quantitative analysis.

Because our analysis resulted in 1,780 pairwise comparisons, we wanted to determine the number of false-positive hits we could expect with a standard $p < 0.05$ cut-off for significance. This is achieved by calculating a corrected *p* value, called the *q* value, which is becoming a common way to determine the number of false positives that can be expected when imposing *p* value cut-offs to large proteomic data sets (118). The *q* value is an extension of a "false discovery rate" and can be described as the rate at which significantly different proteins are actually null, or in other words, the number of false-positives that can be expected given a particular *p* value significance cut-off. For our data set, a *p* value of 0.05 corresponds to a FDR of 10%. In addition to a $p < 0.05$, we also required that proteins exhibit a $\geq$ 2-fold change in abundance in the Δ*secA2* mutant in order to be considered a significant change. Using these criteria, we observed 69 proteins underrepresented in the Δ*secA2* mutant and 131 that were more abundant in the Δ*secA2* mutant by at least 2-fold (Tables 3.1).

We also searched the MS/MS data for proteins that were only identified in either the wild-type or the Δ*secA2* mutant strains. For a protein to be considered strain specific,

we required that it be detected by at least 2 unique peptides in only one of the strains. This resulted in 85 proteins that were unique to the Δ*secA2* mutant and 55 that were unique to the wild type strain (i.e. not detected in the Δ*secA2* mutant) (Tables 3.2 and 3.3). As expected, peptides corresponding to SecA2 (a total of 26 peptides) were identified in the wild type replicates while no SecA2 peptides were detected in the *secA2* mutant samples. It should be noted that some of these "strain-specific" proteins could represent proteins reduced in abundance, but present at a level below our detection in one strain.

***Proteins with differential abundance between wild type* M. tuberculosis *and* ΔsecA2**

Of the total 1,920 proteins identified, there were 124 proteins that were either reduced in the Δ*secA2* mutant cell wall by at least 2-fold with a p value of <0.05, or were uniquely identified in the wild type strain (Figure 3.5A). Conversely, there were 216 proteins either up in the Δ*secA2* mutant by 2-fold ($p < 0.05$) or unique to the Δ*secA2* mutant. Relative protein abundance changes we considered significant ranged from down in the Δ*secA2* mutant by 6-fold to increased in the Δ*secA2* mutant by 16-fold. However, most of the differences in relative protein abundance were not particular large, with an average change of $\pm$ 3 fold.

Of the 124 proteins classified as down or absent in the Δ*secA2* mutant, 32% (40 proteins) contained a signal for export, including 17% (21 proteins) with putative transmembrane domains and 15% (19 proteins) with predicted N-terminal signal peptides (Figure 3.6A). Of the proteins with potential signal peptides, 3% (5 proteins) contained a predicted Tat signal peptide and 12% (14 proteins) had predicted Sec signal peptides (3,

90, 102, 121). Both signal peptide categories included putative lipoproteins (7 Sec and 3 Tat).

In contrast to the 32% of proteins with export signals down in the Δ*secA2* mutant, only 13% of the proteins more abundant or unique to the Δ*secA2* mutant had predicted signals for export (Figure 3.6B). These included 5% with possible transmembrane domains and 8% with predicted signal peptides. Moreover, only three putative lipoproteins were more abundant in the Δ*secA2* mutant, all of which had predicted Sec signal peptides.

Some functional protein classes included multiple examples of both proteins with increased abundance in the Δ*secA2* mutant and proteins with decreased abundance in the mutant, including intermediary metabolism and respiration, conserved hypotheticals, and cell wall processes (Figure 3.5B). However, there were also functional classes showing disparate representation of proteins increased or decreased in abundance, which is discussed further below.

Proteins underrepresented the* ΔsecA2 *cell wall

The 124 proteins reduced in relative abundance or absent in the Δ*secA2* mutant represent candidate substrates for export to the cell wall by the SecA2 system. Among the functional categories represented by the proteins reduced in the Δ*secA2* mutant (Figure 3.5B) were a specific group of lipoproteins: the periplasmic solute-binding proteins (SBPs). *M. tuberculosis* is predicted to have 18 SBPs that have putative functions in binding peptides, sugars, iron, anions, and other substrates for uptake into the bacterial cell via ABC-type membrane transporter complexes (6). Five SBPs were significantly

reduced in the Δ*secA2* mutant cell wall, including those predicted to bind sulfate, glutamine, iron, and peptides (Table 3.1).

Because multiple SBPs were reduced in the Δ*secA2* mutant, and the two known SecA2 substrates of *M. smegmatis* are SBPs, we went back and looked for all predicted SBPs we identified to see how their cell wall presence was affected by deletion of *secA2*. Of the 18 predicted *M. tuberculosis* SBPs, we identified 15 in our proteomic study and 13 of these SBPs followed a trend of being reduced in the Δ*secA2* mutant (Table 3.4). While 5 SBPs were significantly reduced ($p < 0.05$), an additional 8 SBP were present at reduced levels in the Δ*secA2* mutant cell wall although their reduced abundance did not achieve statistical significance. Another interesting observation is that of the 5 SBPs significantly down in the mutant, 3 of these SBPs contain predicted Tat signal peptides. Moreover, of the 5 SBPs with predicted Tat signal peptides in *M. tuberculosis* we identified 4 of these proteins in our proteomic analysis and all 4 were reduced in the mutant, although only 3 had $p < 0.05$.

Another functional category of proteins that were significantly reduced in the Δ*secA2* mutant were members of *mce* transporter systems: Mce1 and Mce4 (Figure 3.4 and Figure 3.5B). The *M. tuberculosis* H37Rv genome contains four homologous *mce* loci, each containing two genes encoding putative membrane-spanning proteins (*yrbE*), and six genes encoding proteins with possible N-terminal transmembrane domains or signal peptides (*mce*) including one predicted lipoprotein (Figure 3.7). Mce1 and Mce4 also contain additional ORFs referred to as mce-associated proteins (*mas*).

Of the 10 ORFs comprising Mce4 we found 7 to be significantly reduced in abundance in the Δ*secA2* mutant by 2-fold ($p < 0.05$). We also detected reduction of an

additional Mce4-associated protein (Mas4B) in the Δ*secA2* mutant. Mas4B showed a $\log_2$ (Δ*secA2*/WT) ratio of -0.6 with a *p* value of 0.007. Among the 12 Mce1 members, eight proteins were reduced in the Δ*secA2* mutant ($p < 0.05$). However, only 3 met our 2-fold change cut-off. Interestingly, the predicted Mce-associated ATPase Mkl was also reduced in the Δ*secA2* mutant cell wall by about 1.8-fold ($p = 0.02$). Mkl does not contain a signal peptide and is predicted to provide energy required for substrate import by all the *M. tuberculosis* Mce transporters (12, 50).

Proteins with increased abundance in the* ΔsecA2 *cell wall

For the 216 proteins increased in relative abundance or uniquely identified in the Δ*secA2* mutant, many have putative functions in intermediary metabolism, cell wall processes, or are conserved hypotheticals of unknown function (Figure 3.5B). There were also a large number of proteins in the categories of information pathways and gene regulation. Most notably, numerous proteins encoded by the DosR dormancy regulon were found to be more abundant in the Δ*secA2* mutant samples. These findings suggest that deletion of *secA2* could result in regulatory responses at the transcriptional level (including the DosR regulon). Increased transcription in the Δ*secA2* mutant would also help explain why most of the proteins increased in the mutant do not contain export signals (Figure 3.6).

The DosR response regulator is part of a two-component response system that is induced upon exposure to hypoxia (a dormancy-associated condition), nitric oxide, or redox stress (61, 112, 128) and activates expression of at least 49 genes. Our proteomic analysis showed that 14 DosR-regulated proteins were more abundant in the Δ*secA2*

mutant (>2-fold, $p < 0.05$) and 7 DosR-regulated proteins were uniquely identified in the Δ*secA2* mutant. Among these DosR-regulated proteins present in higher levels in the Δ*secA2* mutant samples, was the DosR response regulator itself. We subsequently looked for other DosR-regulated proteins in the dataset and found that there were an additional 4 DosR-regulated proteins that followed the trend of being increased in the *secA2* mutant, although they failed to meet our 2-fold, $p < 0.05$ cut-off (Table 3.5). Only 1 of the 49 DosR-regulated proteins was underrepresented in the mutant (Figure 3.5B). Because genes in the DosR regulon are co-regulated at the transcriptional level, it seems likely that the higher levels of DosR-regulated proteins in the *secA2* mutant are a reflection of transcriptional effects.

Immunoblot confirmation of label-free mass spectrometry quantification

We next wanted to confirm some of our label-free quantitation results through immunoblot analysis on the cell wall samples used for our proteomic comparison. This effort was hindered by the lack of available antibodies for proteins showing statistically significant results in our study. We were, however, able to obtain antibodies to PhoS1 which had an almost 3-fold reduced level in the Δ*secA2* mutant versus wild type strain. PhoS1 is one of the solute-binding lipoproteins identified in our study. The PhoS1 protein had a $\log_2$ (Δ*secA2*/WT) ratio of -1.37 with a *p* value of 0.08. Immunoblotting using an anti-PhoS1 antibody revealed that there was less PhoS1 protein in the cell wall of the Δ*secA2* mutant compared to the cell wall of wild type *M. tuberculosis* (Figure 3.8A). Quantitation of the PhoS1 signal in the cell wall fractions showed there was a consistent reduction in PhoS1 levels among the Δ*secA2* replicates of about 1.8 fold, although this

difference was not statistically significant (Figure 3.8B). Nonetheless, this result was consistent with the proteomic ratio we obtained with spectral counting, which also just failed to pass a significance of $p < 0.05$.

Discussion

The SecA2 protein export system is required for *M. tuberculosis* growth inside host macrophages and mice, but our understanding of the exported substrates of the SecA2 system and their role in virulence remains limited. Thus far, efforts to identify mycobacterial SecA2 substrates have relied on 2D-GE comparisons of extra-cytoplasmic fractions, but a surprisingly small number of proteins have been identified as underrepresented in Δ*secA2* deletion mutants (7, 34). Here, we employed a proteomic approach that avoids the use of 2D-GE to generate a more comprehensive list of *M. tuberculosis* proteins that depend on SecA2 for localization to the cell wall.

An advantage to shotgun proteomics combined with approaches like spectral counting is that these techniques can report on small differences in protein levels within large complex protein samples. Because past efforts to identify SecA2-dependent wall proteins using 2D-GE were unsuccessful in *M. tuberculosis* and resulted in only two protein differences in *M. smegmatis*, we anticipated that many SecA2-dependent export defects would be subtle (Gibbons, H.S. unpublished) (34). Indeed, the average difference in protein abundance we observed was only 3-fold. Many of our "strain-specific" protein identifications likely also represent small differences in abundance for proteins just below our level of detection in one strain. One reason for such subtlety in global protein differences can be attributed to the sensitivity of mass spectrometry. It is possible some proteins that are retained in the cytoplasm because of export defects associated with

secA2 deletion could still be identified to some degree in the Δ*secA2* mutant as cytoplasmic contaminates. Many of the protein differences likely reflect partial export defects in the Δ*secA2* mutant, as we confirmed this to be the case for at least the PhoS1 protein by immunoblotting. This highlights another advantage of using such a sensitive method as spectral counting, in that it can report on incomplete export defects that are not apparent from immunoblotting comparisons.

Importantly, although our proteomic analysis of the Δ*secA2* mutant revealed subtle changes in protein abundance, these changes encompassed some specific families of proteins. Specifically, we detected a reduction in solute-binding lipoproteins and Mce transporter proteins in the Δ*secA2* mutant. Thus, our proteomic analysis allowed us to make some predictions concerning the role of SecA2-mediated export in *M. tuberculosis* virulence and also the mechanism of export by SecA2 in mycobacteria.

Contribution of SecA2 export to M. tuberculosis *virulence*

As previously mentioned, comparative 2D-GE analysis of wild type *M. tuberculosis* and the Δ*secA2* mutant was used in a previous study to identify SecA2-dependent proteins secreted into the culture media. But, just a few protein differences were observed and only the superoxide dismutase SodA was later confirmed to be SecA2-dependent for export (7, 44). Macrophages infected with the *M. tuberculosis secA2* mutant exhibit higher levels of apoptosis, and this pro-apoptotic phenotype was linked to the decrease in SodA secretion (44). It was therefore reasoned that increased apoptosis due to defective SodA secretion could explain why the Δ*secA2* mutant fails to grow in macrophages. But restoration of SodA secretion failed to rescue the *M.*

tuberculosis Δ*secA2* growth defect in macrophages (120). It is also possible that SodA secretion is required for detoxifying reactive oxygen species produced by the macrophage, which limits bacterial growth. However, restoration of SodA secretion failed to rescue the *M. tuberculosis* Δ*secA2* growth defect in macrophages (120). Further the *M. tuberculosis* Δ*secA2* mutant remains attenuated for growth in macrophages even if they are derived from *phox*$^{-/-}$ mice, which are unable to elicit an oxidative burst (63). These results suggest there are contributions of SecA2 export to *M. tuberculosis* intracellular growth that are not due to defects in SodA secretion. Therefore, there must be additional SecA2-dependent exported proteins that contribute to virulence.

Our proteomic analysis resulted in the identification of several exported proteins underrepresented in the Δ*secA2* mutant that have known or predicted roles in *M. tuberculosis* virulence. Most of these export defects were incomplete. Thus, it is possible that a defect in export of a single protein may not completely explain the growth attenuation of the Δ*secA2* mutant. But collectively, many of these proteins could contribute to the ability of *M. tuberculosis* to block phagosome maturation and proliferate in macrophages. Below we discuss proteins underrepresented in the Δ*secA2* mutant with predicted roles in blocking phagosomal maturation, promoting growth in macrophages and mice, and acquiring nutrients during infection.

Normally, *M. tuberculosis* blocks phagosome acidification and maturation inside macrophages (93). The *M. tuberculosis* Δ*secA2* mutant is unable to block the process of phagosome acidification and maturation, suggesting that SecA2-dependent exported effectors might be important for targeting host pathways that are critical to these processes (120). Several *M. tuberculosis* secreted proteins are implicated in inhibiting

phagosome maturation including PtpA, SapM, Ndk, LpdC, and PknG (93).Our proteomic study identified two of these proteins: Ndk and PknG. However, Ndk was more abundant in the Δ*secA2* cell wall by over 2-fold ($\log_2$ ratio = 1.8303, p = 0.049) and PknG was unchanged in abundance between the wild type and Δ*secA2* mutant cell wall. However, because even SodA was identified as unchanged in abundance in the Δ*secA2* mutant cell wall by spectral counting, it remains possible that the export defect of SodA and possibly other known secreted effectors are only apparent when looking at proteins fully secreted into the culture medium. Therefore, in the future it will be important to also look at the culture filtrate fraction when looking for SecA2-dependent exported effectors.

LipO (*rv1426c*) is a predicted esterase, which contains a putative Sec signal peptide for export, that was significantly reduced in the Δ*secA2* mutant by about 2-fold. Interestingly, a transposon screen performed in macrophages for *M. tuberculosis* mutants unable to block phagosome-lysosome fusion, identified LipO as a putative effecter of phagosomal maturation arrest (91). While *lipO* inactivation did not translate to strong growth attenuation during macrophage infection in past studies, LipO might be one contributing factor to SecA2-dependent intracellular *M. tuberculosis* growth (91).

Another possible contribution of SecA2 export to *M. tuberculosis* virulence is for the proper assembly of the Mce1 and Mce4 transporters. Mce1 is thought to function as a lipid transporter and numerous independent studies have linked Mce1 with *M. tuberculosis* virulence (21, 31). However, *mce1* deficient *M. tuberculosis* strains have produced phenotypes ranging from attenuation to hyper-virulence, so the exact contribution to Mce1 function during infection is unclear (1, 12, 35, 50, 97, 108, 114, 117). Mce4 is necessary for *M. tuberculosis* growth in mice, a requirement that is linked

to its function in cholesterol import (88, 110). While the composition and mode of activity of Mce transporters is still unknown, it is thought that for each Mce complex the two YrbE proteins form a cytoplasmic membrane-spanning channel, analogous to the permease components of ABC-type transporters (12). The Mce proteins themselves share a conserved domain found in the periplasmic component of a toluene efflux pump of *Pseudomonas*, suggesting that they could form a channel spanning the space between the cytoplasmic membrane and mycobacterial outer membrane (55, 60). The Mce proteins have hydrophobic stretches at their N-termini that represent putative Sec signal peptides and each *mce* locus encodes for one Mce protein with a predicted lipobox motif (77, 90). Some Mce proteins are not predicted by SignalP to contain Sec signal peptides; instead their N-termini could contain a transmembrane domain (90, 116). It's been suggested that Mce proteins are similar to the periplasmic solute-binding proteins of ABC transporters, which suggests a possible function in binding the cognate lipid molecule (12). The proposed similarity between Mce proteins to SBPs is especially notably given the large number of SBPs we identified in this study that were reduced in the Δ*secA2* mutant. But, currently there is no evidence that Mce and SBP share similar functions. Our study showed a decrease in abundance of every Mce component and at least one YrbE permease of both Mce1 and Mce4 in the Δ*secA2* mutant. This is intriguing because it suggests that perhaps SecA2 has a role in exporting many components of these transporters. An equally plausible explanation is that SecA2 assists in the export of a single protein in the complex resulting in instability of the entire complex in the cell wall when export of that protein is impaired in the absence of *secA2*. If SecA2 were to export only one of the Mce proteins, the lipoprotein Mce would be the leading candidate due to

its similarity to SecA2-dependent SBP lipoproteins. A final possibility is that transcription of the *mce* operon could be reduced in the Δ*secA2* mutant. This last possibility could be tested by RT-PCR analysis of *mce* loci in wild-type versus Δ*secA2* mutant strains.

Interestingly, a transposon screen performed in an *M. tuberculosis mce1* mutant background identified a positive genetic interaction between the *mce1* locus and *secA2* (50). In this study, a transposon insertion in *secA2* in wild type *M. tuberculosis* reduced bacterial growth during a mouse infection, which is expected given the established in vivo requirement for SecA2 (7). However, a mutation in *mce1* suppressed the growth attenuation of the *secA2* mutant in mice and vice versa (50). The functional significance of this genetic interaction is unclear at this time. Nonetheless, combined with results from our proteomic analysis of the Δ*secA2* cell wall, it is important to better elucidate the relationship between SecA2 and Mce transporters. It is unlikely that Mce1-mediated transport is a process directly required for blocking phagosomal maturation because an *M. tuberculosis mce1A* mutant retains the ability to block this process (120). But the reduction of lipid import into the Δ*secA2* mutant could produce defects in the mycobacterial cell wall that change the way *M. tuberculosis* interacts with the host, or have implications in nutrient import.

Another interesting finding is that one of the four putative phospholipase C proteins of *M. tuberculosis*, PlcA (*rv2351c*), was underrepresented in the Δ*secA2* mutant by over 4-fold. As phospholipase activity is required for *M. tuberculosis* growth during mouse infection, PlcA is another potential SecA2-dependent virulence factor warranting further study (96).

Finally, a striking observation from our study was the decrease in abundance of multiple solute-binding lipoproteins (SBPs) in the Δ*secA2* mutant cell wall. Because SBPs function as the periplasmic substrate-binding components of ABC-type importers, the reduction in a variety of SBPs in the *secA2* mutant could have a severe impact on the ability of *M. tuberculosis* to import vital nutrients during infection (Table 3.4). The SBPs that exhibited the most significant reduction in the Δ*secA2* mutant are predicted to have roles in uptake of sulfate (SubI), peptides (DppA and OppA), iron (FecB2), and glutamine (GlnH). FecB2 is one of two *M. tuberculosis* SBPs with homology to FecB of *Escherichia coli* suggesting that *M. tuberculosis* may be able to uptake and utilize ferric dicitrate as an iron source (2). However, *M. tuberculosis* has other systems devoted to iron acquisition so it is unclear what exactly the role of the FecB2 protein is, particularly during infection. Inactivation of *subI* completely prevents sulfate uptake, demonstrating that this ABC transporter is the sole locus encoding inorganic sulfur transport in *M. tuberculosis* (129). While a *subI* transposon mutant of *Mycobacterium bovis* did not show growth attenuation during mouse infection, it remains possible that this sulfur transporter is important during *M. tuberculosis* infection (74). DppA and OppA are both demonstrated peptide-binding SBPs (28, 37). The Dpp ABC transporter (*rv3663c-rv3666c*) may serve a vital function during *M. tuberculosis* infection as *dpp* mutants are attenuated for growth in mice (28, 107, 117). Bacterial peptide transporters can have nutrient acquisition functions, but are also used for peptide-mediated signaling that influences protein expression (79). The specific peptide substrate imported by the *M. tuberculosis* Dpp transporter *in vivo* is unknown, but it has been suggested that this particular transporter might serve a signaling function. A *dpp* deficient *M. tuberculosis*

strain has altered expression of several genes encoding surface-exposed PE-family proteins and proteins involved in synthesis of virulence-associated surface lipids (28). However, we did not detect differences in the abundance of PE proteins thought to be influenced by the Dpp transporter (28, 31). Further, a defect in the synthesis of virulence-associated lipids in the Δ*secA2* mutant resulting from an effect on Dpp could certainly contribute to growth attenuation in macrophages, but this possibility has yet to be explored. There has yet to be a link to pathogenesis made for the *M. tuberculosis* Opp transporter (*rv1280c-rv1283c*). However, the homologous peptide transporter of *M. bovis* is implicated in binding and importing a toxic peptide produced by host macrophages, which subsequently reduces host cell apoptosis (19, 37). However, the significance of this is unclear because the increase in apoptosis observed did not correlate to growth attenuation (19). GlnH is a predicted glutamine-binding SBP and whole genome mutagenesis studies suggest it is essential for in vitro growth (107). A *M. tuberculosis* mutant deficient in glutamine synthase is auxotrophic for L-glutamine, and attenuated for growth in macrophages and guinea pigs (124). These results suggest that in the host phagosome, *M. tuberculosis* has limited access to glutamine, which has an essential role in bacterial nitrogen metabolism. It is possible that glutamine import via GlnH may contribute to maintaining intracellular glutamine levels in *M. tuberculosis* during infection.

It is unlikely that any one of the SBPs discussed will fully explain the inability of the Δ*secA2* mutant to grow inside macrophages, since their presumed function would not involve blocking phagosome maturation. However, a combination of subtle export defects in multiple cell wall components with roles in nutrient acquisition or signaling

could result in a fitness disadvantage and decreased virulence for the Δ*secA2* mutant during infection.

Insights into SecA2 Export Mechanisms

Our proteomic comparison showed that many, but not all, predicted lipoproteins were underrepresented in the *M. tuberculosis* Δ*secA2* mutant cell wall (Figures 3.4 and 3.6). Of the 102 predicted *M. tuberculosis* lipoproteins, 52% followed a trend of being reduced in the Δ*secA2* mutant, of which 10 lipoproteins were significantly reduced by at least 2-fold (Figure 3.4) (121). The only known SecA2 substrates in *M. smegmatis* are two lipoproteins yet, SecA2 is not required for export of all lipoproteins in this species, this level of association observed between SecA2 and *M. tuberculosis* lipoproteins was not entirely unexpected (25, 34, 123). Like all bacterial lipoproteins, the putative lipoproteins affected by deletion of *secA2* in *M. tuberculosis* have putative N-terminal signal peptides with lipobox motifs. In Chapter 2, we demonstrated that SecA2 exported proteins of mycobacteria do not have SecA2-specific signal peptides. Instead, the mature domain of the preprotein is what necessitates SecA2 for export. The question then remains, what are the features of particular *M. tuberculosis* lipoproteins that make them SecA2-dependent for export?

The two SecA2 substrates of *M. smegmatis*, Ms1704 and Ms1712 are examples of putative sugar-binding proteins, which are part of the larger solute-binding protein (SBP) family (34, 122). SBPs are the periplasmic components of ABC-type transporters, and are responsible for binding their cognate substrate for import past the cytoplasmic membrane via the ABC permease membrane-channel. Given the precedent set in *M. smegmatis* with

Ms1704 and Ms1712, we were particularly interested in seeing how deletion of *secA2* would influence export of the 17 predicted SBPs in *M. tuberculosis* (6, 121). We identified 15 of these putative SBPs in our proteomic study of the *M. tuberculosis* cell wall, 13 of which were underrepresented in the Δ*secA2* mutant, albeit some do not pass our significance tests (Figure 3.4 and Table 3.4). These results clearly show that, while not all mycobacterial SBPs may require SecA2 for export, SBPs are a particular subset of mycobacterial lipoproteins whose export is influenced by SecA2. Because the *M. smegmatis* Ms1704 and Ms1712 proteins are also SBPs requiring SecA2 for export, these results argue for there being a particular property of SBP proteins that imparts some requirement for SecA2 in their export.

Export of putative Tat substrates are reduced in the M. tuberculosis *Δ*secA2 *mutant*

We know the determining factor for SecA2-dependent export in mycobacteria is the mature domain of the preprotein (Chapter 2). This is somewhat analogous to preproteins of the twin-arginine translocation (Tat) pathway, as the mature domain of Tat substrates must be folded to be compatible for Tat export. More specifically, we hypothesize that the mature domains of mycobacterial SecA2 substrates also exhibit some degree of cytoplasmic folding, which necessitates SecA2 for their export. Surprisingly, 4 of the 13 SBP underrepresented in the *M. tuberculosis* Δ*secA2* mutant have predicted Tat signal peptides, 3 of which are significantly reduced in the mutant cell wall with $p < 0.05$ (Figure 3.6 and Table 3.4). One of these SBPs, Rv2041c, has a signal peptide that was previously proven to be compatible with Tat export (75). It is interesting to note that bacterial SBPs are often predicted Tat substrates, suggesting that structures

shared among SBPs results in similar cytoplasmic folding properties of those preproteins (115). This idea is supported by our data showing that a *M. smegmatis* SecA2 substrate, the Ms1704 SBP, can be exported by the Tat pathway (25).

There are two possible explanations for the reduction of certain putative Tat substrates in the *M. tuberculosis* Δ*secA2* cell wall. The first possibility is that deletion of *secA2* has an indirect affect on Tat export, perhaps by influencing the levels of Tat machinery components. This possibility seems unlikely, because one would predict that an indirect affect of *secA2* deletion on Tat export would affect all Tat substrates, and this was not the case. Of the 33 proteins with predicted Tat signal peptides that were identified in this study, only 5 were significantly reduced in the Δ*secA2* mutant by at least 2-fold (3 of which are SBPs), while 3 were significantly more abundant in the mutant (Figure 3.4).

Therefore, a more plausible explanation is that there is a common set of preproteins whose export is influenced by both the SecA2 and Tat export pathways, and this shared preprotein pool includes the SBPs with predicted Tat signal peptides. The fact that export of these SBPs with Tat signal peptides was reduced, but not eliminated in the Δ*secA2* mutant, is consistent with the idea that these proteins can be exported by two independent systems (either SecA2/Sec or Tat). In support of this possibility, there are reports of preproteins in other bacteria that can utilize either the canonical Sec or Tat pathways for export (56, 94, 125).

A model where export of a particular SBP preprotein is shared (some being exported by Tat while another fraction is exported by SecA2/SecYEG) requires that the signal peptides of some SecA2 SBPs be compatible with both the Sec and Tat machinery.

Four of the SBPs reduced in the *M. tuberculosis* Δ*secA2* mutant have predicted Tat signal peptides with "RR" motifs. However, there are examples of Tat signal peptides with "RR" residues that are functional with the Sec machinery (56, 57). Therefore, it is plausible that a portion of the SBPs with predicted Tat signal peptides are targeting to the canonical Sec machinery where SecA2 assists in their export. Other SBPs that were reduced in the Δ*secA2* mutant cell wall are not predicted to contain Tat signal peptides. Because there are naturally occurring variations to the twin-arginine motif that are still capable of targeting preproteins to Tat (45, 48), these other SBPs could also have signal peptides that are functional with both Sec and Tat machinery. An example of a functional Tat motif variant is "KR," which is found in the Tat signal peptide of the *B. subtilis* TtrB protein (45). Notably, "KR" is similar to the naturally occurring "RK" in the signal peptide of the *M. smegmatis* SecA2 substrate, Ms1704. Additionally, the DppA SBP that was reduced in the *M. tuberculosis* Δ*secA2* mutant has a "RQ" in its signal peptide. There is an example of a similar "RQ" signature within a signal peptide that is functional with the mycobacterial Tat pathway (75).

The partial export defects of many SBPs in the Δ*secA2* mutant, including SBPs with Tat signal peptides, is consistent with a model where the SecA2 and Tat pathways can export a common class of preprotein. This model reaffirms our hypothesis that the mature domains of preproteins impart SecA2-dependency for export, and further supports the idea that SecA2 is required for exporting preproteins that fold in the cytoplasm. If a protein has a signal peptide compatible for either Sec- or Tat-targeting, than it falls in a category of preprotein that potentially could be exported in an unfolded conformation through SecYEG with assistance from SecA2, or exported independently of the Sec

system as a folded preprotein through the TatABC membrane machinery. Our proteomic analysis also indicated that deletion of *secA2* resulted in increased abundance of the TatB component of the Tat machinery, as this protein could only be detected in the Δ*secA2* cell wall. The integral membrane protein TatB interacts with preproteins containing Tat signal peptides after their initial docking with the TatC protein at the membrane (87). It is thought that TatB might help mediate interaction between the preprotein-bound TatC and TatA. Homomultimers of TatA eventually form the membrane-spanning pore for translocation of the substrate across the membrane (80, 133). At this time it is unclear if the increase in TatB protein in our *secA2* mutant reflects a difference in *tatB* gene expression or TatB protein stability. However, if Tat and SecA2 share a common pool of preproteins it would make sense that increased levels of TatB might help compensate for absence of SecA2.

The influence of secA2 *deletion on other protein export machinery*

Currently it is not known what membrane channel SecA2 utilizes in order to promote protein export, but current data suggests the mycobacterial SecA2 works with components of the canonical Sec export system including SecYEG (22, 100). We wanted to know whether deletion of *secA2* would influence the expression or stability of other components of the Sec pathway. This was particularly important given our observation that *secA2* deletion resulted in increased TatB levels. Our spectral counting analysis showed that levels of SecY, SecE, SecD, and SecA1 were unchanged in the Δ*secA2* mutant cell wall compared to wild type.

Interestingly, the *M. tuberculosis* signal recognition particle (SRP) was absent in the Δ*secA2* mutant cell wall. SRP is required for a separate branch of Sec export known as co-translational export, where it works with SecYEG to insert integral membrane proteins into the cytoplasmic membrane as they emerge from the ribosome during translation (132). Because co-translational export is an essential process, there must be SRP present in the Δ*secA2* mutant that was below the level of detection with MS. It will be important to validate the changes in SRP abundance associated with the Δ*secA2* mutant. For now, we can speculate on why deletion of *secA2* reduced the level of SRP detected in our study. There are two possible reasons why the presence of SecA2 could impact SRP levels. First, deletion of *secA2* could indirectly affect the level of membrane-associated SRP because of the shared usage of SecYEG by both SecA2 and SRP. The other possibility is that SecA2 has a direct function related to co-translational export. Co-translational export is thought to occur independently of SecA, but it has also been suggested that there might be some overlap in the function of SecA proteins and SRP (113). While we detected 21 proteins with putative transmembrane domains that were significantly reduced in the Δ*secA2* cell wall, we also detected 11 that were significantly increased. Because there doesn't appear to be a global reduction in integral membrane protein insertion upon deletion of *secA2*, we favor the more indirect explanation for the effect of *secA2* deletion on the reduction of SRP. Nonetheless, this result warrants further investigation for a possible relationship between SRP and SecA2.

Because we hypothesize that SecA2 fulfills a role in mycobacterial export that is related to preproteins that fold in the cytoplasm, we also specifically examined proteins with known or putative chaperone roles related to protein export. The cytoplasmic

chaperones GroEL1 and DnaK can both increase Sec export efficiency in *E. coli*, but neither showed changes in protein level in the Δ*secA2* mutant (26). However, we did detect relative differences in amounts of two proteins with putative functions in protein folding: PpiB and trigger factor (Tf).

PpiB is a predicted peptidyl-prolyl cis–trans isomerase that was decreased by over 2-fold in the Δ*secA2* mutant ($p = 0.04$). Peptidyl-prolyl cis–trans isomerases, also known as cyclophilins, catalyze the cis–trans isomerization of peptide bonds and can also accelerate protein folding (58). An example of a Ppi protein with functions related to protein export is the *Escherichia coli* SurA protein, which assists protein folding in the periplasm following export by the Sec pathway (103). The *M. tuberculosis* PpiB is predicted to be essential and contains a putative transmembrane domain, but a role for this PpiB related to protein export has not yet been established (42). The functional significance of PpiB reduction in the Δ*secA2* mutant is not clear at this time, but it suggests possible changes in the levels of mis-folded proteins in the absence of SecA2. A homologous Ppi protein of *Listeria monocytogenes* (32% identical and 48% similar to *M. tuberculosis* PpiB at the amino acid level) was recently shown to be reduced in abundance in the supernatant of a *secA2* mutant (98). Because *L. monocytogenes* also has a SecA2-only protein export system, this similar link between SecA2 and PpiB is intriguing.

Trigger factor (Tf) is a chaperone that binds ribosomes and interacts with nascent polypeptides as they emerge from the ribosome (126). In our proteomic study, Tf was increased 2-fold in the Δ*secA2* mutant ($p = 0.004$). As a chaperone, Tf also has peptidyl-prolyl cis–trans isomerase activity and assists in folding of cytoplasmic proteins. In *E.*

coli, Tf is not essential, however when Tf is deleted in combination with other cytoplasmic chaperones it results in massive protein aggregation (20). If our model is correct, where SecA2 assists in export of proteins with tendencies to fold in the cytoplasm, then deletion of *secA2* in *M. tuberculosis* could result in increased levels of mis-folded cytoplasmic preproteins. An increase in Tf expression or localization to the membrane in the Δ*secA2* mutant could reflect a response to this cytoplasmic stress.

Proteins increased in the M. tuberculosis Δ**secA2** *mutant reflect possible transcriptional changes*

While the ultimate goal of this study was to identify proteins underrepresented in the Δ*secA2* cell wall, we also identified a large list of proteins that were more abundant in the Δ*secA2* cell wall. Only 13% of the proteins more abundant in the Δ*secA2* mutant have predicted signal peptides or transmembrane domains, while 32% of proteins underrepresented in the Δ*secA2* mutant have these putative export signals. Therefore, many of the proteins increased in the Δ*secA2* mutant cell wall are likely a reflection of regulatory effects that lead to increased expression of cytoplasmic (non-exported) proteins. It is not surprising to detect non-exported proteins in the cell wall as past mass spectrometry studies of the mycobacterial cell wall also result in identification of cytoplasmic proteins (41, 73). These results highlight an advantage of using shotgun proteomic studies for studying protein export pathways, which is that they can report both on protein localization as well as expression differences.

Interestingly, among those proteins over-represented in the Δ*secA2* cell wall were almost half of the proteins that correspond to genes under control of the DosR response

regulator (25 out of the 49 genes of the Dos regulon) (89, 112). Of these 25 proteins, 14 were statistically significant in their increased abundance and 7 were unique to the Δ*secA2* mutant. The DosR system was first identified as a regulon that was induced under hypoxic conditions, a state that is associated with *M. tuberculosis* dormancy (5, 68). Since then, the stresses that induce the DosR response have expanded to include exposure to ethanol, as well as nitric oxide and redox stresses (54, 61, 112, 128). In fact, the DosR regulon can even be induced in *M. tuberculosis* as quickly as 10 minutes after being removed from aeration, such as the conditions experienced during the centrifugation steps upon harvesting cells for proteomic analysis (54). The DosR response system is thought to be important for *M. tuberculosis* survival in the host as it is induced in macrophages, mice, guinea pigs, and possibly during human infection (18, 109, 111, 128). In addition, it has been shown that DosR is required for rapid *M. tuberculosis* growth recovery following anaerobic or nitric oxide-induced non-respiring states in vitro (66). Our proteomic analysis suggests that the Δ*secA2* mutant is more sensitive to induction of the DosR regulon. However, exactly why the DosR regulon appears to be activated in the Δ*secA2* mutant is unclear at this time. At this time, we cannot definitively say if the increased levels of Dos regulated proteins observed reflect transcriptional changes in the Δ*secA2* mutant. This possibility requires RT-PCR to measure transcript levels between wild type and the Δ*secA2* mutant. Given the importance of DosR in *M. tuberculosis* survival in infection models, this possibility certainly merits further testing.

In addition to the DosR response regulator, we also detected changes in the abundance of some alternative sigma factors in the Δ*secA2* mutant. *M. tuberculosis* has 12 alternative sigma factors for inducing transcriptional changes in response to different

environmental conditions (71, 104). Three alternative sigma factors showed differences in abundance between wild type *M. tuberculosis* and the Δ*secA2* mutant, suggesting that there may be additional transcriptional alterations in the Δ*secA2* mutant. The SigF and SigI proteins were not detected in the Δ*secA2* cell wall fraction, while SigD was uniquely identified in the Δ*secA2* mutant. It is unclear if these changes in protein abundance reflect altered activity of these sigma factors. But the absence of SigF and SigI in the Δ*secA2* mutant could reflect localization away from the cytoplasmic membrane (and consequently the cell wall) to the cytoplasm in order to activate gene transcription. On the other hand, the higher levels of SigD in the Δ*secA2* mutant could increase transcription of SigD-regulated genes. While we detected changes in the levels of some proteins expressed from genes within both the SigF and SigD regulons, most of the proteins within these regulatory networks were unchanged or not identified. For corresponding genes under SigF control, nine proteins were more abundant in the Δ*secA2* mutant while three proteins were underrepresented in the mutant (33, 40). Our analysis showed 6 genes under SigD regulation had corresponding proteins more abundant in the Δ*secA2* mutant and two proteins that were less abundant in the mutant (10, 95). At this time, the regulon of SigI has not been defined (78). There is evidence that both SigF and SigD are activiated during host infection (4, 14, 33, 36, 52, 109). In the future, it will be important to test for activation of these sigma response networks in the Δ*secA2* mutant, which could influence survival in the host.

In addition to genes regulated by DosR and the alternative sigma factors described, there were also many proteins with predicted roles in intermediary metabolism and respiration that were more abundant in the Δ*secA2* mutant. Given the large effect of

secA2 deletion on solute-binding proteins, these increased protein abundances could reflect adaptations the Δ*secA2* mutant to changes in nutrient acquisition.

Conclusions

Using a large-scale shotgun proteomic approach combined with semi-quantitative spectral counting, we generated a list of putative SecA2-dependent exported proteins with possible roles in *M. tuberculosis* virulence. The most interesting virulence candidates are components of two Mce lipid importer complexes and an array of solute-binding proteins, which could have possible roles in nutrient acquisition in the host. Going forward, it will be important to validate the SecA2-dependence of exported proteins, particularly those with possible roles in *M. tuberculosis* virulence. It will also be important to establish if the reduction of specific proteins in the Δ*secA2* mutant reflects differences at the level of export as opposed to transcription. Finally, the SecA2-dependent export defects observed will need to be confirmed by complementation to ensure they are due to inactivation of *secA2*. Many of the changes in protein abundance detected in the Δ*secA2* mutant cell wall were subtle, which may present difficulties confirming their SecA2-dependency by immunoblotting. A mass spectrometry technique called selective reaction monitoring (SRM) is becoming a widely used method for confirming hits obtained from large-scale proteomic studies because it is a MS detection method targeted to specific proteins, which improves sensitivity and dynamic range (32). Because our proteomic strategy resulted in many more candidate SecA2-dependent proteins than identified previously by 2D-GE analyses of mycobacterial *secA2* mutants,

future studies should include spectral counting analysis of culture supernatants to find SecA2-dependent secreted proteins.

Results from our proteomic analysis of the Δ*secA2* mutant also strengthen our model whereby SecA2 assists in export of preproteins that have a tendency to fold in the cytoplasm. This model is supported by the observation that multiple predicted Tat substrates were under-represented in the Δ*secA2* mutant. Thus, the mycobacterial SecA2 and Tat systems might assist in export of a common pool of preprotein substrates that tend to fold in the cytoplasm. Deletion of *secA2* also resulted in protein level differences in one component of the Tat pathway (TatB) and two proteins with putative chaperone activities (PpiB and Tf), which could indicate that deletion of *secA2* in *M. tuberculosis* increased cytoplasmic levels of preproteins with tendencies to fold. Prior to this *M. tuberculosis* study, the data from which our model of SecA2 export was built was generated in *M. smegmatis*. Thus, the results presented here are also significant in supporting the model developed in *M. smegmatis*. Taken together, our results from Chapters 2 and 3 suggest that SecA2 may be an alternate solution to solving the problem of getting folded preproteins past the cytoplasmic membrane. In Chapter 4 we return to our model organism, *M. smegmatis*, to specifically test the influence of the Tat export system on export of SecA2 substrates.

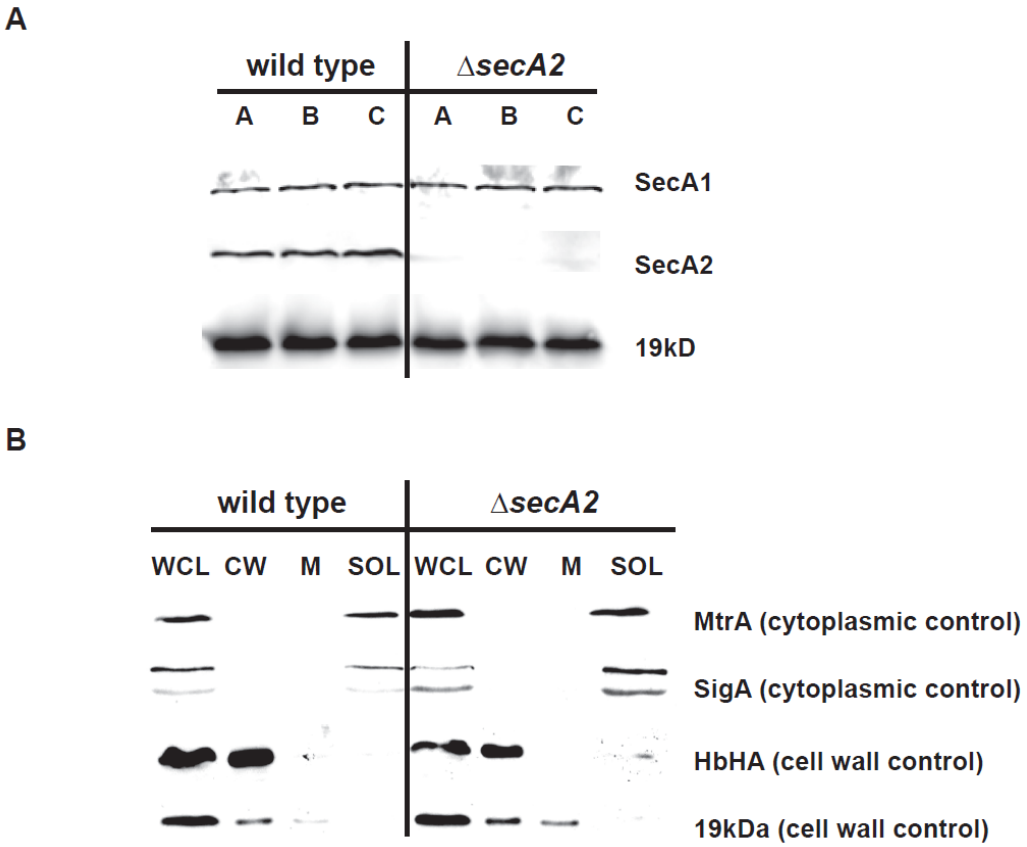

Figure 3.1 Control immunoblots of *M. tuberculosis* lysates and cell wall fractions **(A)** Equalized whole cell lysates (WCL) were generated from triplicate cultures of *M. tuberculosis* H37Rv (wild type) and mc^{2}3112 (Δ*secA2*) and analyzed by immunoblotting using anti-SecA1, anti-SecA2, or anti-19kD antibodies. **(B)** The same equalized WCLs were then subjected to ultracentrifugation to generate subcellular fractions. The total amount of cell wall (CW), membrane (M), and soluble (SOL) fractions shown is equivalent to the amount of WCL loaded. Fractions were separated by SDS-PAGE then transferred to a nitrocellulose membrane for immunoblotting using antibodies against native proteins. The MtrA and SigA proteins were detected as representative cytoplasmic controls, while two cell wall-associated proteins, 19 kDa and HbHA, were detected using anti-HbHA and anti-19kDa antibodies.

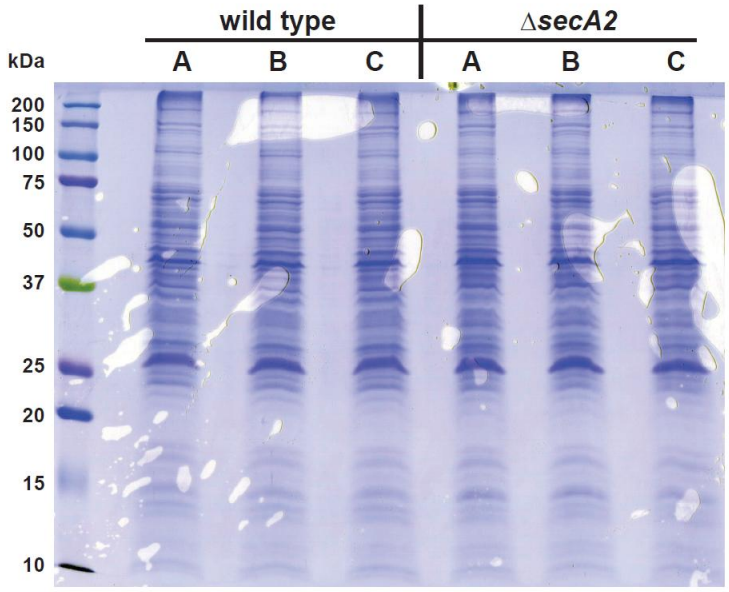

Figure 3.2 SDS-PAGE of *M. tuberculosis* cell wall fractions for proteomic analysis
Cell wall fractions generated from triplicate cultures of *M. tuberculosis* H37Rv (wild type) and $mc^2$3112 (Δ*secA2)* were separated on a 12% SDS-PAGE gel and Coomassie stained for protein visualization. For each of the six samples, protein bands were excised from the entire lane for trypsin digestion. Tryptic peptides were then analyzed by LC-MS/MS for spectral counting comparison between wild type and Δ*secA2*.

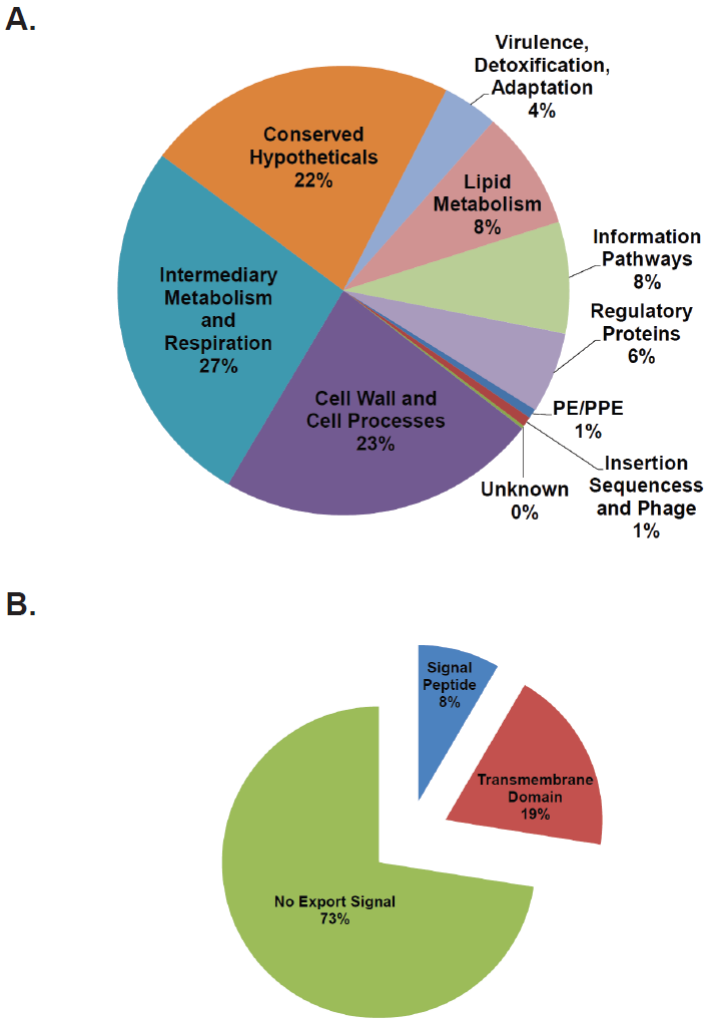

Figure 3.3 Functional categories and export signals of cell wall proteins identified **(A)** A total of 1,920 proteins were identified by mass spectrometry analysis of cell wall protein fractions of *M. tuberculosis* H37Rv (wild type) and mc^{2}3112 (Δ*secA2)*, representing multiple functional categories. **(B)** Amino acid sequences of all proteins identified were analyzed for the presence of N-terminal signal peptides or transmembrane domains. A total of 379 (19%) of proteins identified in the *M. tuberculosis* cell wall contain probable transmembrane domains predicted by the TMHMM 2.0 (116). Additionally, 167 proteins (8%) are predicted to contain signal peptides for export through the conserved Sec and Tat export machinery. These include 134 proteins with putative Sec signal peptides as determined by SignalP and 33 with predicted Tat signal peptides as determined by the TatP and/or TatFind programs (3, 90, 102). Of the 167 proteins with putative export signals, 74 represent predicted lipoproteins as they contain putative lipobox motifs (51).

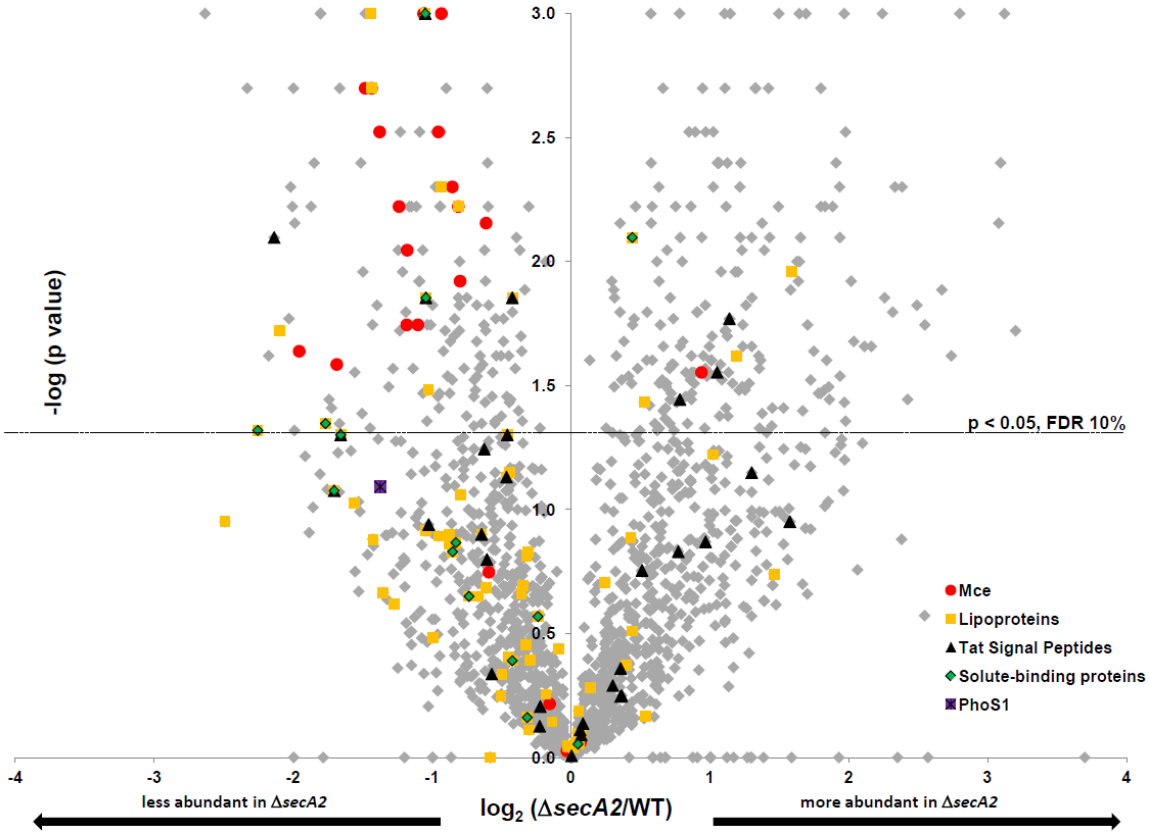

Figure 3.4 Label-free quantitation of *M. tuberculosis* wild type and Δ*secA2* cell wall proteins

To identify changes in protein abundance between the cell wall fractions of wild type *M. tuberculosis* (H37Rv) and the Δ*secA2* mutant, we employed a label-free quantification analysis (LFQ) based on spectral counting. The total number of MS/MS spectra assigned to a particular protein were compared between the two bacterial strains using the ProteoIQ program. We obtained relative abundance ratios for 1,780 proteins (grey diamonds), shown here plotted by Δ*secA2*/WT ratio (log_2 scale) and *p* value (-log scale). Of these shared proteins, 200 were statistically different ($p < 0.05$, ANOVA) in abundance between strains by 2-fold or more, with a corresponding false discovery rate (FDR) of 10% for this *p* value cut-off. Of these proteins, 69 were less abundant in the Δ*secA2* cell wall and 130 were more abundant in the mutant cell wall. The majority of lipoproteins (orange squares) are underrepresented in the mutant, particularly solute-binding lipoproteins (green diamonds). Additionally, all four solute-binding proteins with putative Tat signal peptides identified (black triangles) were less abundant in the Δ*secA2* mutant protein. The PhoS1 lipoprotein (pink asterisk) was underrepresented in the Δ*secA2* mutant ($p = 0.08$) and analyzed by immunoblot to confirm our LFQ findings. Also labeled here are proteins underrepresented in the Δ*secA2* cell wall that correspond to Mce1/Mce4 transporters (red circles).

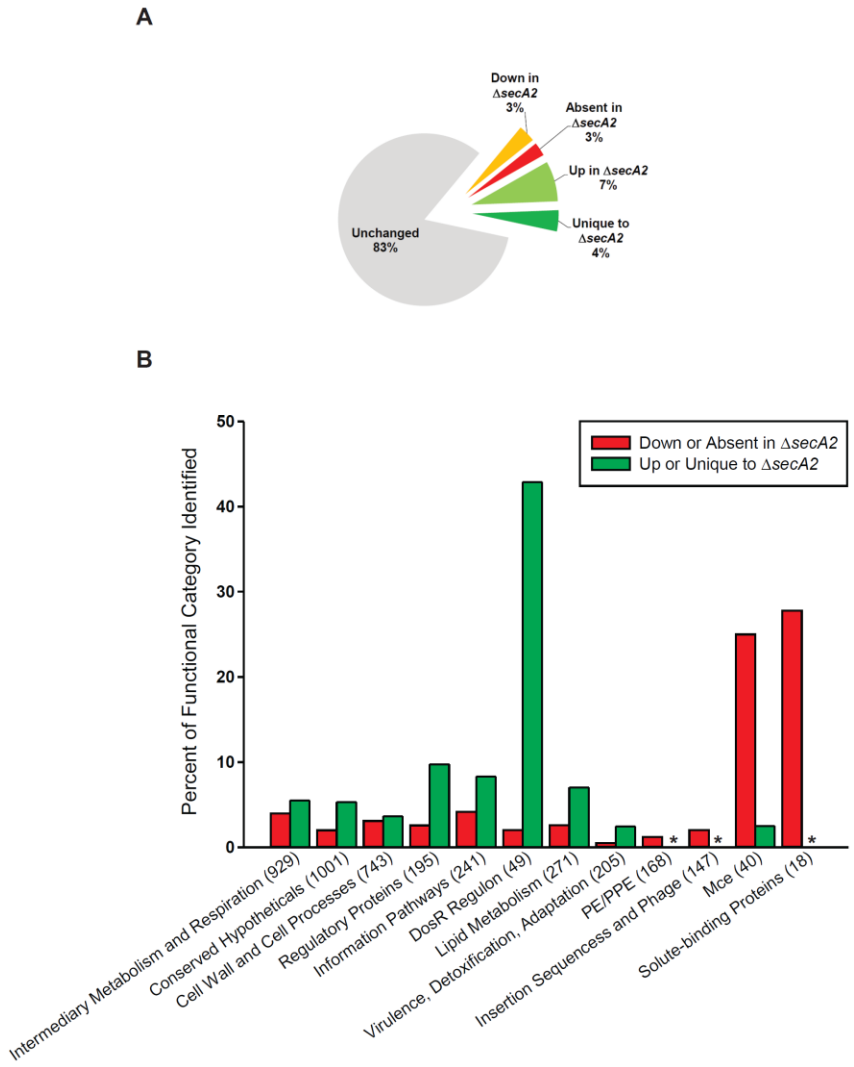

Figure 3.5 Functional categories of proteins with different relative abundances between wild type *M. tuberculosis* and the Δ*secA2* cell wall

(A) Of the 1,920 proteins identified in the cell wall fractions of wild type *M. tuberculosis* (H37Rv) and the Δ*secA2* mutant by mass spectrometry analysis, 83% were unchanged between these strains as determined by semi-quantitative spectral counting analysis. This label-free quantitation also revealed that 3% of proteins were underrepresented in the Δ*secA2* mutant cell wall, while 7% were more abundant in the mutant cell wall. An additional set of proteins were only identified in 1 strain: 3% were absent in the Δ*secA2* mutant cell wall and 4% were only unique to the cell wall of the mutant. **(B)** Proteins with different abundances between the *M. tuberculosis* and the Δ*secA2* mutant cell wall fractions represent multiple functional categories (6, 11, 12, 16, 121). The total number of *M. tuberculosis* H37Rv predicted proteins within each category is in parentheses. An asterisk denotes functional categories for which no proteins were identified as increased in abundance or unique to the Δ*secA2* mutant.

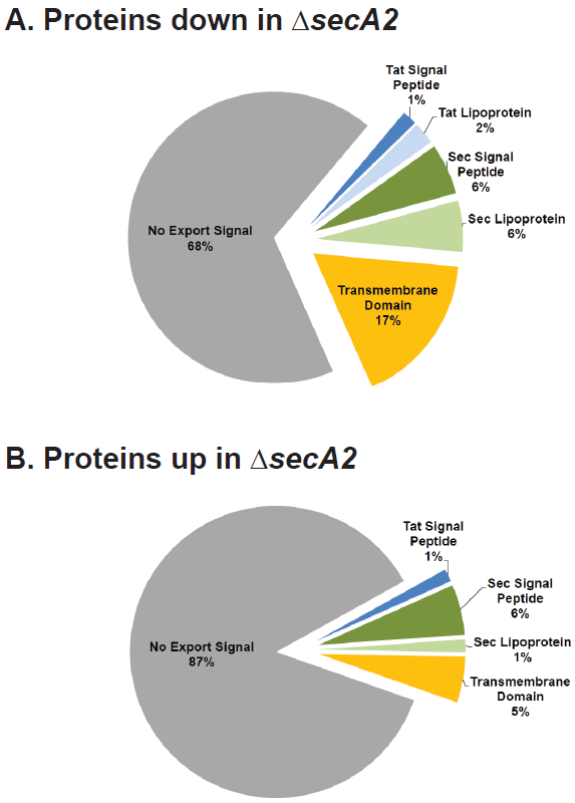

Figure 3.6 Export signals of proteins with differential abundance
(A) Our proteomic analysis of the *M. tuberculosis* cell wall resulted in the identification of 124 proteins either less abundant or absent in the Δ*secA2* mutant cell wall. 32% of these protein contained predicted export signals. **(B)** We also identified 216 proteins that were more abundant or unique to the Δ*secA2* mutant cell wall. Of this group, only 13% are predicted to contain signals for export.

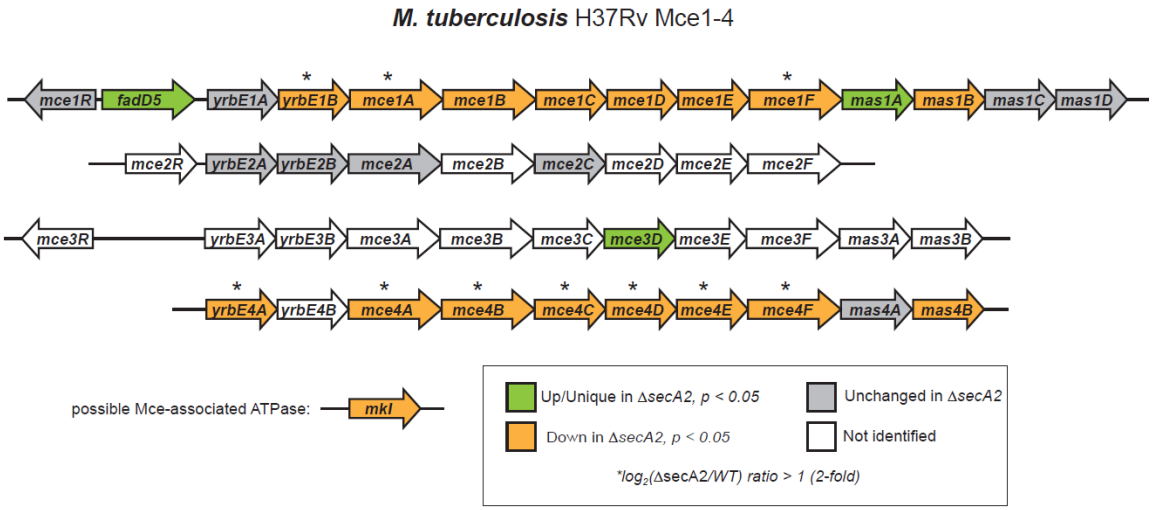

Figure 3.7 Components of two transporters are reduced in the *M. tuberculosis* Δ*secA2* mutant cell wall.
The *M. tuberculosis* H37Rv genome contains four *mce* loci encoding putative transporters. Each *mce* locus contains two *yrbE* ORFs with predicted membrane-spanning domains, six *mce* ORFs with possible signal peptides, and one *mce* regulatory ORF. Some *mce* loci have additional ORFs for "*mce*-associated proteins" (*mas*). Mkl is encoded at a different chromosomal locus than the four *mce* loci and is a predicted ATPase possibly powering Mce transport. Spectral counting comparison of the wild type *M. tuberculosis* cell wall to that of the Δ*secA2* mutant showed eight proteins of Mce1 and Mce4 underrepresented in the mutant cell wall (orange) Proteins with fold changes of -2 or more are denoted with asterisks.

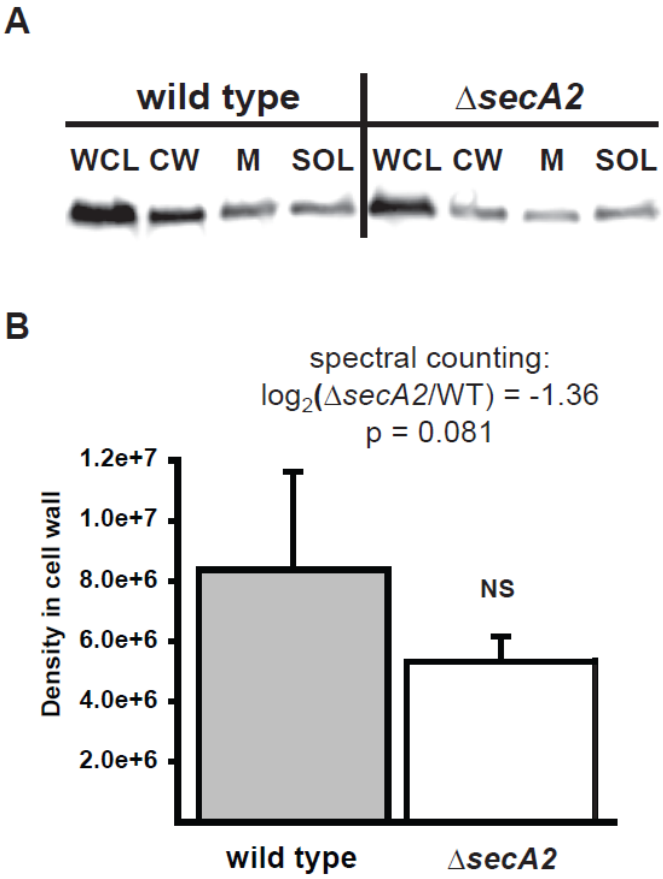

Figure 3.8 Immunoblot conformation of relative PhoS1 protein abundance between the *M. tuberculosis* wild type and Δ*secA2* mutant cell wall

(A) The same *M. tuberculosis* wild type and Δ*secA2* mutant cell wall fractions (CW) used for spectral counting were also analyzed by immunoblot using an anti-PhoS1 antibody. Our label-free quantitation produced a $\log_2$ (Δ*secA2*/WT) ratio for PhoS1 of -1.36 (p = 0.081). Also included are the whole cell lysates (WCL) from which the corresponding cell wall fraction was derived, as well as cytoplasmic membrane (M) and soluble fractions (SOL). Shown here is a representative blot from all three replicates analyzed. **(B)** The mean intensity obtained from the cell wall fraction immunoblotting was determined using ImageJ and reported with +/- the standard error. A paired t-test showed no statistical difference between PhoS1 in the cell wall of wild type and Δ*secA2* mutant *M. tuberculosis*.

Table 3.1 Proteins with 2-fold or greater abundance differences between the *M. tuberculosis* H37Rv and Δ*secA2* cell wall fractions. Proteins with label-free quantification ratios showing at least a 2-fold difference ($\log_2 \pm 1$) with a t-test significance of $p < 0.05$ are shown below. Proteins less abundant in the Δ*secA2* mutant are highlighted in pink while proteins that were more abundant in the Δ*secA2* mutant are highlighted in blue. Proteins with higher average spectral counts (Avg SpC) were most abundant in the *M. tuberculosis* cell wall.

Locus	Gene Name	Description/Function	$\log_2$ Δ*A2*/ WT	*p*	Avg SpC	Sec[a]	Tat[b]	TM[c]	Lipo[d]
Rv2688c		antibiotic ABC transporter ATP-binding	-2.63	0.001	4.8	-	-	X	-
Rv0778	cyp126	cytochrome P450 126	-1.80	0.001	11.6	-	-	-	-
Rv1462		hypothetical protein	-1.80	0.001	2.3	-	-	-	-
Rv3083		monooxygenase	-1.48	0.001	12.1	-	-	-	-
Rv1922		lipoprotein	-1.44	0.001	3.9	X	-	-	X
Rv3499c	mce4A	MCE4-associated	-1.06	0.001	19.8	X	-	X	-
Rv1280c	oppA	ABC transporter oligopeptide-binding protein	-1.05	0.001	22.5	-	X	X	X
Rv2599		hypothetical protein	-2.33	0.002	3.2	X	-	-	-
Rv0728c	serA2	D-3-phosphoglycerate dehydrogenase	-2.00	0.002	5.2	-	-	-	-
Rv3572		hypothetical protein	-1.66	0.002	2.1	X	-	-	-
Rv0001	dnaA	chromosomal replication initiation protein	-1.48	0.002	47.4	-	-	-	-
Rv3497c	mce4C	MCE4-associated	-1.48	0.002	43.5	-	-	X	-
Rv3495c	lprN	MCE4-associated lipoprotein	-1.43	0.002	28.3	X	-	-	X
Rv3501c	yrbE4A	MCE4-associated	-1.38	0.003	11.5	-	-	X	-
Rv0227c		hypothetical protein	-1.23	0.003	66.9	-	-	X	-
Rv3062	ligB	ATP-dependent DNA ligase	-1.09	0.003	4.7	-	-	-	-
Rv0845		two component sensor kinase	-1.85	0.004	6.2	-	-	X	-
Rv1320c		adenylate cyclase	-1.51	0.004	3.9	-	-	X	-
Rv1754c		hypothetical protein	-2.02	0.005	2.7	-	-	X	-
Rv2932	ppsB	phenolpthiocerol synthesis type-I polyketide synthase	-2.00	0.006	2.6	-	-	-	-
Rv1844c	gnd1	6-phosphogluconate dehydrogenase	-1.87	0.006	10.9	-	-	-	-
Rv3498c	mce4B	MCE4-associated	-1.24	0.006	23.0	-	-	X	-
Rv2328	PE23	PE family protein	-1.16	0.006	3.9	-	-	-	-
Rv1380	pyrB	aspartate carbamoyltransferase catalytic subunit	-1.15	0.006	15.6	-	-	-	-
Rv1426c	lipO	esterase	-1.11	0.006	2.5	X	-	X	-
Rv1641	infC	translation initiation factor	-1.99	0.007	1.8	-	-	-	-
Rv2351c	plcA	membrane-associated phospholipase C	-2.14	0.008	3.2	-	X	-	-
Rv2721c		hypothetical protein	-1.24	0.009	106.7	X	-	-	-
Rv0174	mce1F	MCE1-associated	-1.18	0.009	155.4	-	-	X	-
Rv2516c		hypothetical protein	-1.07	0.009	11.6	-	-	-	-
Rv1391	dfp	phosphopantothenoylcysteine decarboxylase/ phosphopantothenate synthase	-1.50	0.011	5.5	-	-	-	-
Rv0379	secE2	protein transport protein	-1.21	0.011	69.5	-	-	-	-
Rv0500A		hypothetical protein	-1.09	0.012	1.2	-	-	-	-
Rv2447c	folC	folylpolyglutamate synthase protein	-1.09	0.012	1.2	-	-	X	-
Rv2975c		hypothetical protein	-1.09	0.012	1.2	-	-	-	-
Rv3704c	gshA	glutamate--cysteine ligase (gamma-glutamylcysteine synthetase)	-1.09	0.012	1.2	-	-	-	-

Rv3764c		two component sensor kinase	-1.09	0.012	1.2	-	-	X	-
Rv0265c	fecB2	iron-transport SBP lipoprotein	-1.04	0.014	13.0	-	X	-	X
Rv0377		LysR family transcriptional regulator	-1.40	0.015	9.7	-	-	-	-
Rv2240c		hypothetical protein	-1.00	0.015	6.8	-	-	-	-
Rv3918c	parA	chromosome partitioning protein	-1.19	0.016	12.4	-	-	-	-
Rv0156	pntAb	NAD(P) transhydrogenase subunit alpha	-2.03	0.017	8.8	-	-	-	-
Rv3059	cyp136	cytochrome P450 136	-1.43	0.018	26.0	-	-	-	-
Rv0168	yrbE1B	MCE1-associated	-1.18	0.018	4.4	-	-	X	-
Rv0169	mce1A	MCE1-associated	-1.10	0.018	80.5	-	-	X	-
Rv2467	pepN	aminopeptidase N	-1.09	0.018	4.2	-	-	-	-
Rv0548c	menB	naphthoate synthase	-1.03	0.018	3.6	-	-	-	-
Rv3194c		hypothetical protein	-1.01	0.018	28.0	X	-	X	-
Rv0419	lpqM	lipoprotein peptidase	-2.10	0.019	11.5	X	-	X	X
Rv2799		hypothetical protein	-1.23	0.019	14.1	-	-	-	-
Rv3496c	mce4D	MCE4-associated	-1.96	0.023	19.3	-	-	X	-
Rv0339c		transcriptional regulator	-2.18	0.024	8.1	-	-	X	-
Rv3494c	mce4F	MCE4-associated	-1.69	0.026	16.5	-	-	X	-
Rv2413c		hypothetical protein	-1.28	0.027	15.1	-	-	-	-
Rv1188		proline dehydrogenase	-1.14	0.027	3.2	-	-	-	-
Rv1596	nadC	nicotinate-nucleotide pyrophosphorylase	-1.56	0.028	12.1	-	-	-	-
Rv2700		secreted alanine rich protein	-1.31	0.032	10.6	-	-	-	-
Rv2784c	lppU	lipoprotein	-1.03	0.033	4.3	X	-	-	X
Rv2822c		hypothetical protein	-1.06	0.034	9.2	-	-	-	-
Rv3609c	folE	GTP cyclohydrolase I	-1.74	0.036	4.8	-	-	-	-
Rv2692	ceoC	TRK system potassium uptake protein	-1.72	0.039	1.3	-	-	-	-
Rv2582	ppiB	peptidyl-prolyl cis-trans isomerase B	-1.50	0.041	26.1	-	-	-	-
Rv1259		hypothetical protein	-1.04	0.044	0.9	-	-	-	-
Rv3666c	dppA	periplasmic dipeptide-binding lipoprotein	-1.77	0.045	5.7	X	-	X	X
Rv3077		hydrolase	-1.67	0.045	9.5	-	-	-	-
Rv1232c		hypothetical protein	-1.45	0.047	8.1	-	-	-	-
Rv2400c	subI	sulfate-binding lipoprotein	-2.25	0.048	2.6	-	-	-	X
Rv0533c	fabH	3-oxoacyl-(acyl carrier protein) synthase III	-1.99	0.048	6.3	-	-	-	-
Rv0411c	glnH	glutamine-binding lipoprotein	-1.66	0.05	6.5	X	X	-	X
Rv0715	rplX	50S ribosomal protein L24	1.41	0.049	7.1	-	-	-	-
Rv2445c	ndk	nucleoside diphosphate kinase	1.83	0.049	4.1	-	-	-	-
Rv1589	bioB	biotin synthase	1.36	0.047	3.2	-	-	-	-
Rv2018		hypothetical protein	1.45	0.046	0.9	-	-	-	-
Rv2683		hypothetical protein	1.98	0.044	3.2	-	-	-	-
Rv0701	rplC	50S ribosomal protein L3	1.07	0.042	10.7	-	-	-	-
Rv2531c		amino acid decarboxylase	1.13	0.042	10.5	-	-	-	-
Rv1904		hypothetical protein	1.38	0.041	5.1	-	-	-	-
Rv2626c		hypothetical protein	1.67	0.04	1.9	-	-	X	-
Rv2928	tesA	thioesterase TESA	1.12	0.039	3.8	-	-	-	-
Rv2915c		hypothetical protein	1.45	0.039	1.4	-	-	-	-
Rv2766c	fabG	3-ketoacyl-(acyl-carrier-protein) reductase	1.53	0.039	4.6	-	-	-	-
Rv0516c		hypothetical protein	1.97	0.037	6.3	-	-	-	-
Rv0802c		hypothetical protein	1.86	0.036	1.0	-	-	-	-
Rv2031c	hspX	heat shock protein	2.42	0.036	36.2	-	-	-	-
Rv3602c	panC	pantoate--beta-alanine ligase	1.78	0.034	2.4	-	-	-	-
Rv3161c		dioxygenase	1.80	0.034	4.1	-	-	-	-
Rv3130c	tgs1	triacylglycerol synthase	1.21	0.033	19.9	-	-	-	-
Rv1957		hypothetical protein	1.25	0.033	3.5	-	-	-	-
Rv2495c	pdhC	branched-chain alpha-keto acid dehydrogenase subunit E2	1.82	0.033	2.4	-	-	-	-
Rv1546		hypothetical protein	1.23	0.032	3.6	-	-	-	-

Rv3134c		hypothetical protein	1.86	0.032	7.5	-	-	-	-
Rv3865		hypothetical protein	1.16	0.031	4.1	-	-	-	-
Rv3147	nuoC	NADH dehydrogenase subunit C	1.28	0.031	15.3	-	-	-	-
Rv1278		hypothetical protein	1.09	0.03	14.6	-	-	-	-
Rv2777c		hypothetical protein	1.73	0.03	2.2	-	-	-	-
Rv2995c	leuB	3-isopropylmalate dehydrogenase	1.94	0.029	2.9	-	-	-	-
Rv1290c		hypothetical protein	1.05	0.028	1.1	-	X	X	-
Rv0328		TetR/AcrR family transcriptional regulator	1.14	0.028	3.2	-	-	-	-
Rv1767		hypothetical protein	1.36	0.027	1.7	-	-	-	-
Rv0861c	ercC3	DNA helicase	1.08	0.026	10.9	-	-	-	-
Rv0880		MarR family transcriptional regulator	1.01	0.025	1.1	-	-	-	-
Rv1593c		hypothetical protein	1.01	0.025	3.3	-	-	-	-
Rv1674c		transcriptional regulator	1.01	0.025	1.1	-	-	-	-
Rv2464c		DNA glycosylase	1.01	0.025	1.1	-	-	-	-
Rv1630	rpsA	30S ribosomal protein S1	1.37	0.025	18.1	-	-	-	-
Rv1501		hypothetical protein	1.84	0.025	1.9	-	-	-	-
Rv1274	lprB	lipoprotein	1.19	0.024	6.4	X	-	-	X
Rv2118c		RNA methyltransferase	1.35	0.024	5.2	-	-	-	-
Rv3099c		hypothetical protein	2.74	0.024	1.6	-	-	-	-
Rv0189c	ilvD	dihydroxy-acid dehydratase	1.75	0.023	7.8	-	-	X	-
Rv0282		hypothetical protein	1.24	0.022	4.6	-	-	-	-
Rv2071c	cobM	precorrin-4 C11-methyltransferase	1.30	0.022	2.4	-	-	-	-
Rv0079		hypothetical protein	2.11	0.022	12.3	-	-	-	-
Rv3224		short chain dehydrogenase	2.16	0.022	4.4	-	-	-	-
Rv1019		TetR family transcriptional regulator	2.04	0.021	2.9	-	-	-	-
Rv0434		hypothetical protein	1.02	0.02	5.4	-	-	-	-
Rv0708	rplP	50S ribosomal protein L16	1.13	0.02	12.7	-	-	-	-
Rv0060		hypothetical protein	1.11	0.019	1.1	-	-	-	-
Rv0110		integral membrane protein	1.11	0.019	1.1	-	-	-	-
Rv0691c		transcriptional regulator	1.11	0.019	1.1	-	-	-	-
Rv1401		hypothetical protein	1.11	0.019	1.1	-	-	X	-
Rv2693c		integral membrane alanine and leucine rich protein	1.11	0.019	1.1	-	-	-	-
Rv3188		hypothetical protein	1.11	0.019	1.1	-	-	-	-
Rv2625c		hypothetical protein	3.20	0.019	9.8	-	-	-	-
Rv2444c	rne	ribonuclease E	1.24	0.018	6.1	-	-	X	-
Rv1738		hypothetical protein	1.70	0.018	6.2	-	-	-	-
Rv2823c		hypothetical protein	2.55	0.018	1.3	-	-	-	-
Rv3804c	fbpA	secreted antigen (Mycolyl transferase 85A)	1.14	0.017	50.0	-	X	X	-
Rv0078		transcriptional regulator	1.15	0.017	8.6	-	-	-	-
Rv3249c		TetR family transcriptional regulator	2.32	0.016	6.1	-	-	-	-
Rv3819		hypothetical protein	1.26	0.015	17.3	-	-	-	-
Rv3617	ephA	epoxide hydrolase	2.49	0.015	2.4	-	-	-	-
Rv0187		O-methyltransferase	1.30	0.014	1.3	-	-	-	-
Rv2534c	efp	elongation factor P	2.26	0.014	2.6	-	-	-	-
Rv1338	murI	glutamate racemase	1.58	0.013	5.0	-	-	-	-
Rv2030c		hypothetical protein	2.67	0.013	12.4	-	-	-	-
Rv2883c	pyrH	uridylate kinase	1.64	0.012	3.0	-	-	-	-
Rv3868		hypothetical protein	2.02	0.012	8.9	-	-	-	-
Rv2940c	mas	multifunctional mycocerosic acid synthase membrane-associated MAS	1.08	0.011	70.4	-	-	-	-
Rv1383	carA	carbamoyl phosphate synthase small subunit	1.18	0.011	4.0	-	-	-	-
Rv0317c	glpQ2	glycerophosphoryl diester phosphodiesterase	1.20	0.011	29.9	-	-	-	-

Rv2251		flavoprotein	1.59	0.011	2.1	-	-	-	X
Rv3271c		integral membrane protein	1.63	0.01	5.7	-	-	X	-
Rv2911	dacB2	D-alanyl-D-alanine carboxypeptidase	1.65	0.01	6.6	X	-	-	-
Rv3766		hypothetical protein	1.23	0.009	1.9	-	-	-	-
Rv2017		transcriptional regulator	1.39	0.009	5.5	-	-	-	-
Rv3799c	accD4	propionyl-CoA carboxylase beta chain	1.22	0.008	6.6	-	-	-	-
Rv3297	nei	endonuclease VIII	1.30	0.008	1.9	-	-	-	-
Rv3051c	nrdE	ribonucleotide-diphosphate reductase subunit alpha	1.43	0.008	19.4	-	-	-	-
Rv3401		hypothetical protein	1.65	0.008	15.2	-	-	-	-
Rv3813c		hypothetical protein	1.94	0.008	4.2	-	-	-	-
Rv2026c		hypothetical protein	1.38	0.007	12.8	-	-	-	-
Rv2748c	ftsK	cell division transmembrane protein	3.08	0.007	7.5	-	-	-	-
Rv2919c	glnB	nitrogen regulatory protein P-II	1.12	0.006	5.6	-	-	-	-
Rv1284		hypothetical protein	1.49	0.006	1.0	-	-	-	-
Rv3715c	recR	recombination protein	1.80	0.006	6.2	-	-	-	-
Rv3592	TB11.2	hypothetical protein	1.83	0.006	2.5	-	-	-	-
Rv0695		hypothetical protein	1.88	0.006	2.3	-	-	-	-
Rv2604c	PdxT	glutamine amidotransferase subunit	1.03	0.005	3.3	-	-	-	-
Rv3236c		integral membrane transport protein	1.22	0.005	3.6	-	-	X	-
Rv1837c	glcB	malate synthase G	1.94	0.005	7.0	-	-	-	-
Rv2070c	cobK	cobalt-precorrin-6x reductase	2.33	0.005	2.2	-	-	-	-
Rv0425c	ctpH	metal cation transporting P-type ATPase	2.38	0.005	9.3	X	-	-	-
Rv0409	ackA	acetate kinase	1.06	0.004	1.3	-	-	-	-
Rv1422		hypothetical protein	1.06	0.004	1.3	-	-	-	-
Rv2397c	cysA1	sulfate-transport ATP-binding protein ABC transporter	1.06	0.004	3.8	-	-	-	-
Rv2462c	tig	trigger factor	1.06	0.004	1.3	-	-	-	-
Rv2736c	recX	recombination regulator	1.06	0.004	1.3	-	-	-	-
Rv3036c	TB22.2	hypothetical protein	1.06	0.004	1.3	X	-	-	-
Rv3600c		pantothenate kinase	1.06	0.004	1.3	-	-	-	-
Rv0102		integral membrane protein	1.07	0.004	1.3	-	-	-	-
Rv2676c		hypothetical protein	1.13	0.004	5.7	-	-	-	-
Rv0999		hypothetical protein	1.22	0.004	14.9	-	-	-	-
Rv2286c		hypothetical protein	1.91	0.004	1.2	-	-	-	-
Rv0860	fadB	fatty oxidation protein	3.09	0.004	17.6	-	-	-	-
Rv3909		hypothetical protein	1.03	0.003	52.5	X	-	-	-
Rv1463		ABC transporter ATP-binding protein	1.98	0.003	2.6	-	-	-	-
Rv3880c		hypothetical protein	1.11	0.002	11.3	-	-	-	-
Rv2989		transcriptional regulator	1.33	0.002	9.5	-	-	-	-
Rv3052c	nrdI	ribonucleotide reductase stimulatory protein	1.33	0.002	18.2	-	-	-	-
Rv0080		hypothetical protein	1.42	0.002	7.4	-	-	-	-
Rv0034		hypothetical protein	1.80	0.002	1.7	-	-	-	-
Rv2337c		hypothetical protein	1.11	0.001	5.1	-	-	-	-
Rv1167c		transcriptional regulator	1.15	0.001	8.6	-	-	-	-
Rv0972c	fadE12	acyl-CoA dehydrogenase	1.50	0.001	2.1	-	-	-	-
Rv0910		hypothetical protein	1.64	0.001	3.3	-	-	-	-
Rv3133c	devR/dosR	two component transcriptional regulatory protein	1.69	0.001	4.6	-	-	-	-
Rv2002	fabG3	20-beta-hydroxysteroid dehydrogenase	1.97	0.001	8.0	-	-	-	-
Rv3676	crp	CRP/FNR family transcriptional regulator	2.24	0.001	15.3	-	-	-	-
Rv2927c		hypothetical protein	2.80	0.001	4.4	-	-	-	-
Rv0570	nrdZ	ribonucleoside-diphosphate	3.12	0.001	6.6	-	-	-	-

		reductase large subunit							
Rv2627c		hypothetical protein	4.04	0.001	12.0	-	-	-	-
Rv1996		hypothetical protein	1.17	0	10.4	-	-	-	-
Rv2971		oxidoreductase	1.27	0	3.7	-	-	-	-
Rv1908c	katG	catalase-peroxidase-peroxynitritase T	1.30	0	8.6	-	-	-	-
Rv1133c	metE	5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase	1.30	0	51.3	-	-	-	-
Rv1489		hypothetical protein	1.93	0	10.3	-	-	X	-
Rv0540		hypothetical protein	2.35	0	2.4	-	-	-	-
Rv3295		TetR family transcriptional regulator	2.57	0	3.8	-	-	-	-

a – An "X" indicates the presence of an N-terminal signal peptide predicted using SignalP v. 4.0 (90)

b – An "X" indicates the presence of an N-terminal Tat signal peptides as predicted using TatP and TatFind

c – An "X" indicates one or more putative transmembrane domains predicted using using TMHMM v. 2.0 (116)

d – An "X" indicates a potential lipoprotein containing a putative lipobox-type signal peptide as predicted by the LipoP v. 1.0 program (51) or using a lipoprotein pattern search against the *M. tuberculosis* H37Rv genome (121)

Table 3.2 Proteins identified only in the *M. tuberculosis* H37Rv (wild type) cell wall by at least 2 peptides.

Locus	Gene Name	Description/Function	Total Peptides	Sec[a]	Tat[b]	TM[c]	Lipo[d]
Rv1821	secA2	preprotein translocase subunit SecA2	26	-	-	-	-
Rv2121c	hisG	ATP phosphoribosyltransferase	5	-	-	-	-
Rv2264c		hypothetical protein	5	-	-	-	-
Rv1995		hypothetical protein	4	-	-	-	-
Rv2092c	helY	ATP-dependent DNA helicase	4	-	-	-	-
Rv0764c	cyp51	cytochrome P450 sterol 14-alpha demethylase	3	-	-	-	-
Rv1189	sigI	RNA polymerase sigma factor	3	-	-	-	-
Rv1878	glnA3	glutamine synthetase	3	-	-	-	-
Rv2064	cobG	cobalamin biosynthesis protein CobG	3	-	-	-	-
Rv2714		hypothetical protein	3	-	-	-	-
Rv2791c		transposase	3	-	-	-	-
Rv2916c	ffh	signal recognition particle protein	3	-	-	-	-
Rv3032		transferase	3	-	-	-	-
Rv3397c	phyA	phytoene synthase	3	-	-	-	-
Rv3518c	cyp142	cytochrome P450 monooxygenase 142	3	-	-	-	-
Rv3860		hypothetical protein	3	-	-	-	-
Rv0082		oxidoreductase	2	-	-	-	-
Rv0154c	fadE2	acyl-CoA dehydrogenase FADE2	2	-	-	-	-
Rv0212c	nadR	transcriptional regulatory protein NadR	2	-	-	-	-
Rv0483	lprQ	lipoprotein	2	-	-	-	X
Rv0486		mannosyltransferase	2	-	-	-	-
Rv0510	hemC	porphobilinogen deaminase	2	-	-	-	-
Rv0680c		transmembrane protein	2	X	-	X	-
Rv0819		hypothetical protein	2	-	-	-	-
Rv1011	ispE	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	2	-	-	-	-
Rv1302	rfe	UDP-GlcNAc transferase	2	-	-	X	-
Rv1330c		nicotinate phosphoribosyltransferase	2	-	-	-	-
Rv1420	uvrC	excinuclease ABC subunit C	2	-	-	-	-
Rv1522c	mmpL12	transmembrane transport protein	2	-	-	X	-
Rv1551	plsB1	glycerol-3-phosphate acyltransferase	2	-	-	-	-
Rv1561		hypothetical protein	2	-	-	-	-
Rv1933c	fadE18	acyl-CoA dehydrogenase	2	-	-	-	-
Rv2067c		hypothetical protein	2	-	-	-	-
Rv2089c	pepE	dipeptidase PepE	2	-	X	-	-
Rv2356c	PPE40	PPE family protein	2	-	-	-	-
Rv2361c		long (C50) chain Z-isoprenyl diphosphate synthase (Z-decaprenyl diphosphate synthase)	2	-	-	-	-
Rv2386c	mbtI	salicylate synthase	2	-	-	-	-
Rv2391	nirA	ferredoxin-dependent nitrite reductase	2	-	-	-	-
Rv2650c		phiRv2 prophage protein	2	-	-	-	-
Rv2733c		hypothetical protein	2	-	-	-	-
Rv2832c	ugpC	sn-glycerol-3-phosphate transport ATP-binding ABC transporter	2	-	-	-	-
Rv2978c		transposase	2	-	-	-	-
Rv3175		amidase	2	-	-	-	-
Rv3201c		ATP-dependent DNA helicase	2	-	-	-	-
Rv3255c	manA	mannose-6-phosphate isomerase	2	-	-	-	-
Rv3286c	sigF	RNA polymerase sigma factor	2	-	-	-	-
Rv3300c		hypothetical protein	2	-	-	-	-
Rv3647c		hypothetical protein	2	-	-	-	-
Rv3665c	dppB	peptide ABC transporter transmembrane protein	2	-	-	X	-
Rv3730c		hypothetical protein	2	-	-	-	-
Rv3737		transmembrane protein	2	-	-	X	-
Rv3754	tyrA	prephenate dehydrogenase	2	-	-	-	-
Rv3797	fadE35	acyl-CoA dehydrogenase FADE35	2	-	-	-	-
Rv3817		phosphotransferase	2	-	-	-	-
Rv3863		hypothetical protein	2	-	-	-	-

a – An "X" indicates the presence of an N-terminal signal peptide predicted using SignalP v. 4.0 (90)
b – An "X" indicates the presence of an N-terminal Tat signal peptides as predicted using TatP and TatFind
c – An "X" indicates one or more putative transmembrane domains predicted using using TMHMM v. 2.0 (116)
d – An "X" indicates a potential lipoprotein containing a putative lipobox-type signal peptide as predicted by the LipoP v. 1.0 program (51) or using a lipoprotein pattern search against the *M. tuberculosis* H37Rv genome (121)

Table 3.3 Proteins identified only in the *M. tuberculosis* mc^{2}3112 (Δ*secA2*) cell wall by at least 2 peptides.

Locus	Gene Name	Description/Function	Total Peptides	Sec[a]	Tat[b]	TM[c]	Lipo[d]
Rv1527c	pks5	polyketide synthase pks5	8	-	-	-	-
Rv3127		hypothetical protein	6	-	-	-	-
Rv3414c	sigD	RNA polymerase sigma factor	6	-	-	-	-
Rv1997	ctpF	metal cation transporter P-type ATPase A	5	-	-	X	-
Rv2905	lppW	alanine rich lipoprotein	5	X	-	-	X
Rv3267		hypothetical protein	5	X	-	X	-
Rv0572c		hypothetical protein	4	-	-	-	-
Rv1477		invasion protein	4	X	-	-	-
Rv1710		hypothetical protein	4	-	-	-	-
Rv2486	echA14	enoyl-CoA hydratase	4	-	-	-	-
Rv2590	fadD9	fatty-acid-CoA ligase	4	-	-	-	-
Rv3822		hypothetical protein	4	-	-	-	-
Rv0063		oxidoreductase	3	X	-	-	-
Rv0711	atsA	arylsulfatase AtsA	3	-	-	-	-
Rv0807		hypothetical protein	3	-	-	-	-
Rv0865	mog	molybdopterin biosynthesis	3	-	-	-	-
Rv1094	desA2	acyl-[acyl-carrier protein] desaturase	3	-	-	-	-
Rv1467c	fadE15	acyl-CoA dehydrogenase	3	-	-	-	-
Rv1813c		hypothetical protein	3	X	X	X	-
Rv1816		transcriptional regulatory protein	3	-	-	-	-
Rv1889c		hypothetical protein	3	-	-	-	-
Rv2308		hypothetical protein	3	-	-	-	-
Rv2417c		hypothetical protein	3	-	-	-	-
Rv2925c	rnc	ribonuclease III	3	-	-	-	-
Rv3203	lipV	lipase LipV	3	-	-	-	-
Rv3263		DNA methylase (modification methylase)	3	-	-	-	-
Rv3396c	guaA	GMP synthase	3	-	-	-	-
Rv3562	fadE31	acyl-CoA dehydrogenase	3	-	-	-	-
Rv3913	trxB2	thioredoxin reductase TRXB2	3	-	-	-	-
Rv0040c	mtc28	secreted proline rich antigen	2	X	-	X	-
Rv0101	nrp	peptide synthetase	2	-	-	-	-
Rv0244c	fadE5	acyl-CoA dehydrogenase	2	-	-	-	-
Rv0263c		hypothetical protein	2	-	-	-	-
Rv0264c		hypothetical protein	2	-	-	-	-
Rv0266c	oplA	5-oxoprolinase	2	-	-	-	-
Rv0302		TetR/ACRR family transcriptional regulator	2	-	-	-	-
Rv0309		hypothetical protein	2	X	-	X	-
Rv0803	purL	phosphoribosylformylglycinamidine synthase II	2	-	-	-	-
Rv0839		hypothetical protein	2	-	-	-	-
Rv0840c	pip	proline iminopeptidase	2	-	-	-	-
Rv0844c	narL	nitrate/nitrite response transcriptional regulatory protein NarL	2	-	-	-	-
Rv0853c	pdc	pyruvate or indole-3-pyruvate decarboxylase pd	2	-	-	-	-
Rv0868c	moaD2	molybdenum cofactor biosynthesis protein D	2	-	-	-	-
Rv1074c	fadA3	acetyl-CoA acetyltransferase	2	-	-	-	-
Rv1076	lipU	lipase LipU	2	-	-	-	-
Rv1117		hypothetical protein	2	-	-	-	-
Rv1219c		transcriptional regulatory protein	2	-	-	-	-
Rv1224	tatB	sec-independent translocase	2	-	-	X	-
Rv1315	murA	UDP-N-acetylglucosamine 1-	2	-	-	-	-

		carboxyvinyltransferase					
Rv1358		transcriptional regulatory protein	2	-	-	-	-
Rv1466		hypothetical protein	2	-	-	-	-
Rv1529	fadD24	acyl-CoA synthetase	2	-	-	-	-
Rv1559	ilvA	threonine dehydratase	2	-	-	-	-
Rv1566c		inv protein	2	X	-	X	-
Rv1724c		hypothetical protein	2	-	-	-	-
Rv1856c		short chain dehydrogenase	2	-	-	-	-
Rv1887		hypothetical protein	2	-	-	-	-
Rv1969	mce3D	MCE-family protein MCE3D	2	X	-	X	-
Rv2003c		hypothetical protein	2	-	-	-	-
Rv2004c		hypothetical protein	2	-	-	-	-
Rv2032	acg	hypothetical protein	2	-	-	-	-
Rv2109c	prcA	proteasome (alpha subunit)	2	-	-	-	-
Rv2166c		cell division protein MraZ	2	-	-	-	-
Rv2178c	aroG	3-deoxy-D-arabino-heptulosonate 7-phosphate synthase	2	-	-	-	-
Rv2211c	gcvT	glycine cleavage/aminomethyltransferase T	2	-	-	-	-
Rv2362c	recO	DNA repair protein RecO	2	-	-	-	-
Rv2454c		2-oxoglutarate ferredoxin oxidoreductase subunit beta	2	-	-	-	-
Rv2501c	accA1	acetyl-/propionyl-coenzyme A carboxylase subunit alpha	2	-	-	-	-
Rv2917		hypothetical protein	2	-	-	-	-
Rv2923c		hypothetical protein	2	-	-	-	-
Rv3210c		hypothetical protein	2	-	-	-	-
Rv3370c	dnaE2	error-prone DNA polymerase	2	-	-	-	-
Rv3432c	gadB	glutamate decarboxylase	2	-	-	-	-
Rv3463		hypothetical protein	2	-	-	-	-
Rv3524		hypothetical protein	2	-	-	X	-
Rv3550	echA20	enoyl-CoA hydratase	2	-	-	-	-
Rv3583c		transcription factor	2	-	-	-	-
Rv3628	ppa	inorganic pyrophosphatase	2	-	-	-	-
Rv3651		hypothetical protein	2	-	-	-	-
Rv3710	leuA	2-isopropylmalate synthase	2	-	-	-	-
Rv3711c	dnaQ	DNA polymerase III subunit epsilon	2	-	-	-	-
Rv3717		hypothetical protein	2	X	-	-	-
Rv3823c	mmpL8	integral membrane transport protein	2	-	-	X	-
Rv3837c		phosphoglycerate mutase	2	-	-	-	-
Rv3858c	gltD	glutamate synthase subunit beta	2	-	-	-	-

a – An "X" indicates the presence of an N-terminal signal peptide predicted using SignalP v. 4.0 (90)

b – An "X" indicates the presence of an N-terminal Tat signal peptides as predicted using TatP and TatFind

c – An "X" indicates one or more putative transmembrane domains predicted using using TMHMM v. 2.0 (116)

d – An "X" indicates a potential lipoprotein containing a putative lipobox-type signal peptide as predicted by the LipoP v. 1.0 program (51) or using a lipoprotein pattern search against the *M. tuberculosis* H37Rv genome (121)

Table 3.4 Solute-binding lipoproteins of *M. tuberculosis*. The *M. tuberculosis* H37Rv genome contains 18 predicted periplasmic solute-binding proteins (SPBs), which are the lipoprotein components of ABC-type transporters. Mass spectrometry analysis resulted in the identification of 15 SBPs. A $\log_2$ (Δ*secA2*/WT) ratio of 1 corresponds to a 2-fold change. Protein changes of at least 2-fold are highlighted red, and *p* values of 0.05 or less are highlighted yellow. Also shown is the average spectral count (Avg SpC) obtained for each protein and the presence of either a Sec or Tat signal peptides (SP).

Name	Locus	$\log_2$ (Δ*secA2*/ WT)	*p*	Avg SpC	SP[a]	Putative Substrate[b]
SubI	Rv2400c	-2.25	0.048	2.6	Sec	sulfate
DppA	Rv3666c	-1.79	0.045	5.7	Sec	peptide
Rv2041c	Rv2041c	-1.69	0.084	5.3	Tat	sugar
GlnH	Rv0411c	-1.64	0.05	6.5	Tat	glutamine
PhoS1	Rv0934	-1.36	0.081	137	Sec	phosphate
OppA	Rv1280c	-1.06	0.001	22.5	Tat	peptide
FecB2	Rv0265c	-1.03	0.014	13.0	Tat	iron (III) dicitrate
Rv2585c	Rv2585c	-0.86	0.148	13.5	Sec	peptide
FecB	Rv3044	-0.84	0.136	22.2	Sec	iron (III) dicitrate
PstS/PstS2	Rv0932c	-0.74	0.224	26.8	Sec	phosphate
ProX	Rv3759c	-0.42	0.407	7.4	Sec	glycine-betaine/L-proline
LpqZ	Rv1244	-0.32	0.69	5.2	Sec	glycine/betaine
LpqY	Rv1235	-0.23	0.27	25.7	Sec	sugar
LpqW	Rv1166	0.06	0.884	21.7	Sec	peptide
PstS3/PhoS2	Rv0928	0.44	0.008	9.7	Sec	phosphate
UspC	Rv2318	-	-	-	Sec	sugar
ModA	Rv1857	-	-	-	Sec	molydenum
UgpB	Rv2833c	-	-	-	Tat	glycerol-3-phosphate

a – Signal peptide predictions for Sec using SignalP v. 4.0 (90) and Tat using TatP (3)and TatFind (102)
b – Predicted substrate specificity (6)

Table 3.5 DosR-regulated proteins of *M. tuberculosis*. The *M. tuberculosis* H37Rv genome contains at least 49 genes within the Dos regulon. Our proteomic analysis resulted in the identification of 26 of the corresponding proteins. Proteins identified are shown here with the relative abundance obtained between wild type and Δ*secA2* mutant. A $\log_2$ (Δ*secA2*/WT) ratio of 1 corresponds to a 2-fold change. Protein changes of at least 2-fold are highlighted green, and *p* values of 0.05 or less are highlighted yellow. Some proteins were either identified in only wild type (WT-only, red) or the Δ*secA2* mutant strain (SecA2-only, dark green).

Name	Locus	$\log_2$ (Δ*secA2*/WT)	*p*	Avg SpC	Peptides	Description
	Rv0079	2.113	0.022	12.3	-	hypothetical protein
	Rv0080	1.4222	0.002	7.4	-	hypothetical protein
	Rv0082	WT-only	-	-	2	oxidoreductase
nrdZ	Rv0570	3.1218	0.001	6.6	-	ribonucleoside-diphosphate reductase large subunit
	Rv0572c	SecA2-only	-	-	4	hypothetical protein
	Rv1738	1.704	0.018	6.2	-	hypothetical protein
	Rv1812c	0.1784	0.425	9.2	-	dehydrogenase
	Rv1813c	SecA2-only	-	-	3	hypothetical protein
	Rv1996	1.1724	0	10.4	-	hypothetical protein
ctpF	Rv1997	SecA2-only	-	-	5	metal cation transporter P-type ATPase A
	Rv2003c	SecA2-only	-	-	2	hypothetical protein
	Rv2004c	SecA2-only	-	-	2	hypothetical protein
	Rv2005c	0.9686	0.015	21.7	-	hypothetical protein
	Rv2030c	2.6715	0.013	12.4	-	hypothetical protein
hspX	Rv2031c	2.4246	0.036	36.2	-	heat shock protein
acg	Rv2032	SecA2-only	-	-	2	hypothetical protein
TB31.7	Rv2623	3.7	0	7.6	-	hypothetical protein
	Rv2625c	3.202	0.019	9.8	-	hypothetical protein
	Rv2626c	1.6694	0.04	1.9	-	hypothetical protein
	Rv2627c	4.0448	0.001	12.0	-	hypothetical protein
	Rv3127	SecA2-only	-	-	6	hypothetical protein
tgs1	Rv3130c	1.2116	0.033	19.9	-	triacylglycerol synthase
devS	Rv3132c	0.7346	0.274	27.8	-	two component sensor histidine kinase
dosR/devR	Rv3133c	1.692	0.001	4.6	-	two component transcriptional regulatory protein
	Rv3134c	1.8568	0.032	7.5	-	hypothetical protein
bfrB	Rv3841	0.5666	0.038	21.7	-	bacterioferritin
	Rv0081	-	-	-	-	probable transcriptional regulatory protein
	Rv0083	-	-	-	-	probable oxidoreductase
	Rv0569	-	-	-	-	hypothetical protein
	Rv0571c	-	-	-	-	hypothetical protein
	Rv0574c	-	-	-	-	hypothetical protein
	Rv1733c	-	-	-	-	probable conserved transmembrane protein
	Rv1734c	-	-	-	-	hypothetical protein
narX	Rv1736c	-	-	-	-	probable nitrate reductase
narK2	Rv1737c	-	-	-	-	possible nitrate/nitrite transporter
otsB1	Rv2006	-	-	-	-	probable trehalose-6-phosphate phosphatase
fdxA	Rv2007	-	-	-	-	ferredoxin
	Rv2028c	-	-	-	-	Universal stress protein family protein
pfkB	Rv2029c	-	-	-	-	6-phosphofructokinase
	Rv2624c	-	-	-	-	Universal stress protein family protein
	Rv2628	-	-	-	-	hypothetical protein
	Rv2629	-	-	-	-	hypothetical protein
	Rv2630	-	-	-	-	hypothetical protein
	Rv2631	-	-	-	-	hypothetical protein
vapB22	Rv2830c	-	-	-	-	possible antitoxin
	Rv3126c	-	-	-	-	hypothetical protein

	Rv3128c	-	-	-	-	hypothetical protein
	Rv3129	-	-	-	-	hypothetical protein
	Rv3131	-	-	-	-	hypothetical protein

References

1. **Arruda, S., G. Bomfim, R. Knights, T. Huima-Byron, and L. W. Riley.** 1993. Cloning of an *M. tuberculosis* DNA fragment associated with entry and survival inside cells. Science **261:**1454-1457.

2. **Banerjee, S., A. Farhana, N. Z. Ehtesham, and S. E. Hasnain.** 2011. Iron acquisition, assimilation and regulation in mycobacteria. Infect Genet Evol **11:**825-838.

3. **Bendtsen, J. D., H. Nielsen, D. Widdick, T. Palmer, and S. Brunak.** 2005. Prediction of twin-arginine signal peptides. BMC Bioinformatics **6:**167.

4. **Betts, J. C., P. T. Lukey, L. C. Robb, R. A. McAdam, and K. Duncan.** 2002. Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. Mol Microbiol **43:**717-731.

5. **Boon, C., R. Li, R. Qi, and T. Dick.** 2001. Proteins of *Mycobacterium bovis* BCG induced in the Wayne dormancy model. J Bacteriol **183:**2672-2676.

6. **Braibant, M., P. Gilot, and J. Content.** 2000. The ATP binding cassette (ABC) transport systems of *Mycobacterium tuberculosis*. FEMS Microbiol Rev **24:**449-467.

7. **Braunstein, M., B. Espinosa, J. Chan, J. T. Belisle, and W. R. J. Jacobs.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium tuberculosis*. Mol Microbiol **48:**453-464.

8. **Bunai, K., and K. Yamane.** 2005. Effectiveness and limitation of two-dimensional gel electrophoresis in bacterial membrane protein proteomics and perspectives. J Chromatogr B Analyt Technol Biomed Life Sci **815:**227-236.

9. **Burkholder, K. M., K. P. Kim, K. K. Mishra, S. Medina, B. K. Hahm, H. Kim, and A. K. Bhunia.** 2009. Expression of LAP, a SecA2-dependent secretory protein, is induced under anaerobic environment. Microbes Infect **11:**859-867.

10. **Calamita, H., C. Ko, S. Tyagi, T. Yoshimatsu, N. E. Morrison, and W. R. Bishai.** 2005. The *Mycobacterium tuberculosis* SigD sigma factor controls the

expression of ribosome-associated gene products in stationary phase and is required for full virulence. Cell Microbiol **7:**233-244.

11. **Camus, J. C., M. J. Pryor, C. Medigue, and S. T. Cole.** 2002. Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. Microbiology **148:**2967-2973.

12. **Casali, N., and L. W. Riley.** 2007. A phylogenomic analysis of the Actinomycetales *mce* operons. BMC Genomics **8:**60.

13. **Chatzi, K. E., M. F. Sardis, S. Karamanou, and A. Economou.** 2013. Breaking on through to the other side: protein export through the bacterial Sec system. Biochem J **449:**25-37.

14. **Chen, P., R. E. Ruiz, Q. Li, R. F. Silver, and W. R. Bishai.** 2000. Construction and characterization of a *Mycobacterium tuberculosis* mutant lacking the alternate sigma factor gene, sigF. Infect Immun **68:**5575-5580.

15. **Cirillo, D. M., R. H. Valdivia, D. M. Monack, and S. Falkow.** 1998. Macrophage-dependent induction of the *Salmonella* pathogenicity island 2 type III secretion system and its role in intracellular survival. Mol Microbiol **30:**175-188.

16. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M. A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead, and B. G. Barrell.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393:**537-544.

17. **Collins, M. O., L. Yu, and J. S. Choudhary.** 2008. Analysis protein complexes by 1D-SDS-PAGE and tandem mass spectrometry.

18. **Commandeur, S., M. Y. Lin, K. E. van Meijgaarden, A. H. Friggen, K. L. Franken, J. W. Drijfhout, G. E. Korsvold, F. Oftung, A. Geluk, and T. H. Ottenhoff.** 2011. Double- and monofunctional CD4(+) and CD8(+) T-cell responses to *Mycobacterium tuberculosis* DosR antigens and peptides in long-term latently infected individuals. Eur J Immunol **41:**2925-2936.

19. **Dasgupta, A., K. Sureka, D. Mitra, B. Saha, S. Sanyal, A. K. Das, P. Chakrabarti, M. Jackson, B. Gicquel, M. Kundu, and J. Basu.** 2010. An oligopeptide transporter of *Mycobacterium tuberculosi*s regulates cytokine release and apoptosis of infected macrophages. PLoS One **5:**e12225.

20. **Deuerling, E., A. Schulze-Specking, T. Tomoyasu, A. Mogk, and B. Bukau.** 1999. Trigger factor and DnaK cooperate in folding of newly synthesized proteins. Nature **400:**693-696.

21. **Dunphy, K. Y., R. H. Senaratne, M. Masuzawa, L. V. Kendall, and L. W. Riley.** 2010. Attenuation of *Mycobacterium tuberculosis* functionally disrupted in a fatty acyl-coenzyme A synthetase gene fadD5. J Infect Dis **201:**1232-1239.

22. **Fagan, R. P., and N. F. Fairweather.** 2011. *Clostridium difficile* has two parallel and essential Sec secretion systems. J Biol Chem **286:**27483-27493.

23. **Fairn, G. D., and S. Grinstein.** 2012. How nascent phagosomes mature to become phagolysosomes. Trends Immunol **33:**397-405.

24. **Feltcher, M. E., and M. Braunstein.** 2012. Emerging themes in SecA2-mediated protein export. Nat Rev Microbiol **10:**779-789.

25. **Feltcher, M. E., H. S. Gibbons, L. S. Ligon, and M. Braunstein.** 2013. Protein export by the mycobacterial SecA2 system Is determined by the preprotein mature domain. J Bacteriol **195:**672-681.

26. **Fisher, A. C., and M. P. DeLisa.** 2004. A little help from my friends: quality control of presecretory proteins in bacteria. J Bacteriol **186:**7467-7473.

27. **Fisher, M. A., B. B. Plikaytis, and T. M. Shinnick.** 2002. Microarray analysis of the *Mycobacterium tuberculosis* transcriptional response to the acidic conditions found in phagosomes. J Bacteriol **184:**4025-4032.

28. **Flores-Valdez, M. A., R. P. Morris, F. Laval, M. Daffe, and G. K. Schoolnik.** 2009. *Mycobacterium tuberculosis* modulates its cell surface via an oligopeptide permease (Opp) transport system. Faseb J **23:**4091-4104.

29. **Fol, M., A. Chauhan, N. K. Nair, E. Maloney, M. Moomey, C. Jagannath, M. V. Madiraju, and M. Rajagopalan.** 2006. Modulation of *Mycobacterium*

tuberculosis proliferation by MtrA, an essential two-component response regulator. Mol Microbiol **60:**643-657.

30. **Fontan, P., V. Aris, S. Ghanny, P. Soteropoulos, and I. Smith.** 2008. Global transcriptional profile of *Mycobacterium tuberculosis* during THP-1 human macrophage infection. Infect Immun **76:**717-725.

31. **Forrellad, M. A., L. I. Klepp, A. Gioffre, J. Sabio y Garcia, H. R. Morbidoni, M. de la Paz Santangelo, A. A. Cataldi, and F. Bigi.** 2013. Virulence factors of the *Mycobacterium tuberculosis* complex. Virulence **4:**3-66.

32. **Gallien, S., E. Duriez, and B. Domon.** 2011. Selected reaction monitoring applied to proteomics. J Mass Spectrom **46:**298-312.

33. **Geiman, D. E., D. Kaushal, C. Ko, S. Tyagi, Y. C. Manabe, B. G. Schroeder, R. D. Fleischmann, N. E. Morrison, P. J. Converse, P. Chen, and W. R. Bishai.** 2004. Attenuation of late-stage disease in mice infected by the *Mycobacterium tuberculosis* mutant lacking the SigF alternate sigma factor and identification of SigF-dependent genes by microarray analysis. Infect Immun **72:**1733-1745.

34. **Gibbons, H. S., F. Wolschendorf, M. Abshire, M. Niederweis, and M. Braunstein.** 2007. Identification of two *Mycobacterium smegmatis* lipoproteins exported by a SecA2-dependent pathway. J Bacteriol **189:**5090-5100.

35. **Gioffre, A., E. Infante, D. Aguilar, M. P. Santangelo, L. Klepp, A. Amadio, V. Meikle, I. Etchechoury, M. I. Romano, A. Cataldi, R. P. Hernandez, and F. Bigi.** 2005. Mutation in mce operons attenuates *Mycobacterium tuberculosis* virulence. Microbes Infect **7:**325-334.

36. **Graham, J. E., and J. E. Clark-Curtiss.** 1999. Identification of *Mycobacterium tuberculosis* RNAs synthesized in response to phagocytosis by human macrophages by selective capture of transcribed sequences (SCOTS). Proc Natl Acad Sci U S A **96:**11554-11559.

37. **Green, R. M., A. Seth, and N. D. Connell.** 2000. A peptide permease mutant of *Mycobacterium bovis* BCG resistant to the toxic peptides glutathione and S-nitrosoglutathione. Infect Immun **68:**429-436.

38. **Gu, S., J. Chen, K. M. Dobos, E. M. Bradbury, J. T. Belisle, and X. Chen.** 2003. Comprehensive Proteomic Profiling of the Membrane Constituents of a *Mycobacterium tuberculosis* Strain. Mol Cell Proteomics **2:**1284-1296.

39. **Guo, X. V., M. Monteleone, M. Klotzsche, A. Kamionka, W. Hillen, M. Braunstein, S. Ehrt, and D. Schnappinger.** 2007. Silencing *Mycobacterium smegmatis* by using tetracycline repressors. J Bacteriol **189:**4614-4623.

40. **Hartkoorn, R. C., C. Sala, S. Uplekar, P. Busso, J. Rougemont, and S. T. Cole.** 2012. Genome-wide definition of the SigF regulon in *Mycobacterium tuberculosis*. J Bacteriol **194:**2001-2009.

41. **He, Z., and J. De Buck.** 2010. Cell wall proteome analysis of *Mycobacterium smegmatis* strain MC2 155. BMC Microbiol **10:**121.

42. **Henriksson, L. M., P. Johansson, T. Unge, and S. L. Mowbray.** 2004. X-ray structure of peptidyl-prolyl cis-trans isomerase A from *Mycobacterium tuberculosis*. Eur J Biochem **271:**4107-4113.

43. **Hilbi, H., and A. Haas.** 2012. Secretive bacterial pathogens and the secretory pathway. Traffic **13:**1187-1197.

44. **Hinchey, J., S. Lee, B. Y. Jeon, R. J. Basaraba, M. M. Venkataswamy, B. Chen, J. Chan, M. Braunstein, I. M. Orme, S. C. Derrick, S. L. Morris, W. R. Jacobs, Jr., and S. A. Porcelli.** 2007. Enhanced priming of adaptive immunity by a proapoptotic mutant of Mycobacterium tuberculosis. J Clin Invest **117:**2279-2288.

45. **Hinsley, A. P., N. R. Stanley, T. Palmer, and B. C. Berks.** 2001. A naturally occurring bacterial Tat signal peptide lacking one of the 'invariant' arginine residues of the consensus targeting motif. FEBS Lett **497:**45-49.

46. **Hood, R. D., P. Singh, F. Hsu, T. Guvener, M. A. Carl, R. R. Trinidad, J. M. Silverman, B. B. Ohlson, K. G. Hicks, R. L. Plemel, M. Li, S. Schwarz, W. Y. Wang, A. J. Merz, D. R. Goodlett, and J. D. Mougous.** 2010. A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. Cell Host Microbe **7:**25-37.

47. **Hutchings, M. I., T. Palmer, D. J. Harrington, and I. C. Sutcliffe.** 2009. Lipoprotein biogenesis in Gram-positive bacteria: knowing when to hold 'em, knowing when to fold 'em. Trends Microbiol **17:**13-21.

48. **Ignatova, Z., C. Hornle, A. Nurk, and V. Kasche.** 2002. Unusual signal peptide directs penicillin amidase from *Escherichia coli* to the Tat translocation machinery. Biochem Biophys Res Commun **291:**146-149.

49. **Jain, M., C. J. Petzold, M. W. Schelle, M. D. Leavell, J. D. Mougous, C. R. Bertozzi, J. A. Leary, and J. S. Cox.** 2007. Lipidomics reveals control of *Mycobacterium tuberculosis* virulence lipids via metabolic coupling. Proc Natl Acad Sci U S A **104:**5133-5138.

50. **Joshi, S. M., A. K. Pandey, N. Capite, S. M. Fortune, E. J. Rubin, and C. M. Sassetti.** 2006. Characterization of mycobacterial virulence genes through genetic interaction mapping. Proc Natl Acad Sci U S A **103:**11760-11765.

51. **Juncker, A. S., H. Willenbrock, G. Von Heijne, S. Brunak, H. Nielsen, and A. Krogh.** 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci **12:**1652-1662.

52. **Karls, R. K., J. Guarner, D. N. McMurray, K. A. Birkness, and F. D. Quinn.** 2006. Examination of *Mycobacterium tuberculosis* sigma factor mutants using low-dose aerosol infection of guinea pigs suggests a role for SigC in pathogenesis. Microbiology **152:**1591-1600.

53. **Keller, A., A. I. Nesvizhskii, E. Kolker, and R. Aebersold.** 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem **74:**5383-5392.

54. **Kendall, S. L., F. Movahedzadeh, S. C. Rison, L. Wernisch, T. Parish, K. Duncan, J. C. Betts, and N. G. Stoker.** 2004. The *Mycobacterium tuberculosis* dosRS two-component system is induced by multiple stresses. Tuberculosis (Edinb) **84:**247-255.

55. **Kim, K., S. Lee, K. Lee, and D. Lim.** 1998. Isolation and characterization of toluene-sensitive mutants from the toluene-resistant bacterium *Pseudomonas putida* GM73. J Bacteriol **180:**3692-3696.

56. **Kolkman, M. A., R. van der Ploeg, M. Bertels, M. van Dijk, J. van der Laan, J. M. van Dijl, and E. Ferrari.** 2008. The twin-arginine signal peptide of *Bacillus subtilis* YwbN can direct either Tat- or Sec-dependent secretion of different cargo proteins: secretion of active subtilisin via the *B. subtilis* Tat pathway. Appl Environ Microbiol **74:**7507-7513.

57. **Kouwen, T. R., R. van der Ploeg, H. Antelmann, M. Hecker, G. Homuth, U. Mader, and J. M. van Dijl.** 2009. Overflow of a hyper-produced secretory protein from the *Bacillus* Sec pathway into the Tat pathway for protein secretion as revealed by proteogenomics. Proteomics **9:**1018-1032.

58. **Kromina, K. A., A. N. Ignatov, and I. A. Abdeeva.** 2008. Role of peptidyl-prolyl-cis/trans-isomerases in pathologic processes. Biochemistry (Moscow) Supplement Series A: Membrane and Cell Biology **2:**195-202.

59. **Kuboniwa, M., E. L. Hendrickson, Q. Xia, T. Wang, H. Xie, M. Hackett, and R. J. Lamont.** 2009. Proteomics of *Porphyromonas gingivalis* within a model oral microbial community. BMC Microbiol **9:**98.

60. **Kumar, A., A. Chandolia, U. Chaudhry, V. Brahmachari, and M. Bose.** 2005. Comparison of mammalian cell entry operons of mycobacteria: in silico analysis and expression profiling. FEMS Immunol Med Microbiol **43:**185-195.

61. **Kumar, A., J. C. Toledo, R. P. Patel, J. R. Lancaster, Jr., and A. J. Steyn.** 2007. *Mycobacterium tuberculosis* DosS is a redox sensor and DosT is a hypoxia sensor. Proc Natl Acad Sci U S A **104:**11568-11573.

62. **Kuntumalla, S., Q. Zhang, J. C. Braisted, R. D. Fleischmann, S. N. Peterson, A. Donohue-Rolfe, S. Tzipori, and R. Pieper.** 2011. In vivo versus in vitro protein abundance analysis of *Shigella dysenteriae* type 1 reveals changes in the expression of proteins involved in virulence, stress and energy metabolism. BMC Microbiol **11:**147.

63. **Kurtz, S., K. P. McKinnon, M. S. Runge, J. P. Ting, and M. Braunstein.** 2006. The SecA2 secretion factor of *Mycobacterium tuberculosis* promotes growth in macrophages and inhibits the host immune response. Infect Immun **74:**6855-6864.

64. **Lefevre, P., M. Braibant, L. de Wit, M. Kalai, D. Roeper, J. Grotzinger, J. P. Delville, P. Peirs, J. Ooms, K. Huygen, and J. Content.** 1997. Three different putative phosphate transport receptors are encoded by the *Mycobacterium*

tuberculosis genome and are present at the surface of Mycobacterium bovis BCG. J Bacteriol **179:**2900-2906.

65. **Lefevre, P., O. Denis, L. De Wit, A. Tanghe, P. Vandenbussche, J. Content, and K. Huygen.** 2000. Cloning of the gene encoding a 22-kilodalton cell surface antigen of *Mycobacterium bovis* BCG and analysis of its potential for DNA vaccination against tuberculosis. Infect Immun **68:**1040-1047.

66. **Leistikow, R. L., R. A. Morton, I. L. Bartek, I. Frimpong, K. Wagner, and M. I. Voskuil.** 2010. The *Mycobacterium tuberculosis* DosR regulon assists in metabolic homeostasis and enables rapid recovery from nonrespiring dormancy. J Bacteriol **192:**1662-1670.

67. **Ligon, L. S., J. D. Hayden, and M. Braunstein.** 2012. The ins and outs of *Mycobacterium tuberculosis* protein export. Tuberculosis (Edinb) **92:**121-132.

68. **Lim, A., M. Eleuterio, B. Hutter, B. Murugasu-Oei, and T. Dick.** 1999. Oxygen depletion-induced dormancy in *Mycobacterium bovis* BCG. J Bacteriol **181:**2252-2256.

69. **Liu, H., R. G. Sadygov, and J. R. Yates, 3rd.** 2004. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem **76:**4193-4201.

70. **Malen, H., G. A. De Souza, S. Pathak, T. Softeland, and H. G. Wiker.** 2011. Comparison of membrane proteins of *Mycobacterium tuberculosis* H37Rv and H37Ra strains. BMC Microbiol **11:**18.

71. **Manganelli, R., E. Dubnau, S. Tyagi, F. R. Kramer, and I. Smith.** 1999. Differential expression of 10 sigma factor genes in *Mycobacterium tuberculosis*. Mol Microbiol **31:**715-724.

72. **Marrero, J., K. Y. Rhee, D. Schnappinger, K. Pethe, and S. Ehrt.** 2010. Gluconeogenic carbon flow of tricarboxylic acid cycle intermediates is critical for *Mycobacterium tuberculosis* to establish and maintain infection. Proc Natl Acad Sci U S A **107:**9819-9824.

73. **Mawuenyega, K. G., C. V. Forst, K. M. Dobos, J. T. Belisle, J. Chen, E. M. Bradbury, A. R. Bradbury, and X. Chen.** 2005. *Mycobacterium tuberculosis*

functional network analysis by global subcellular protein profiling. Mol Biol Cell **16:**396-404.

74. **McAdam, R. A., T. R. Weisbrod, J. Martin, J. D. Scuderi, A. M. Brown, J. D. Cirillo, B. R. Bloom, and W. R. Jacobs, Jr.** 1995. In vivo growth characteristics of leucine and methionine auxotrophic mutants of *Mycobacterium bovis* BCG generated by transposon mutagenesis. Infect Immun **63:**1004-1012.

75. **McDonough, J. A., J. R. McCann, E. M. Tekippe, J. S. Silverman, N. W. Rigel, and M. Braunstein.** 2008. Identification of functional Tat signal sequences in *Mycobacterium tuberculosis* proteins. J Bacteriol **190:**6428-6438.

76. **McKinney, J. D., K. Honer zu Bentrup, E. J. Munoz-Elias, A. Miczak, B. Chen, W. T. Chan, D. Swenson, J. C. Sacchettini, W. R. Jacobs, Jr., and D. G. Russell.** 2000. Persistence of *Mycobacterium tuberculosis* in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. Nature **406:**735-738.

77. **Mitra, D., B. Saha, D. Das, H. G. Wiker, and A. K. Das.** 2005. Correlating sequential homology of Mce1A, Mce2A, Mce3A and Mce4A with their possible functions in mammalian cell entry of *Mycobacterium tuberculosis* performing homology modeling. Tuberculosis (Edinb) **85:**337-345.

78. **Monahan, I. M., J. Betts, D. K. Banerjee, and P. D. Butcher.** 2001. Differential expression of mycobacterial proteins following phagocytosis by macrophages. Microbiology **147:**459-471.

79. **Monnet, V.** 2003. Bacterial oligopeptide-binding proteins. Cell Mol Life Sci **60:**2100-2114.

80. **Mould, R. M., and C. Robinson.** 1991. A proton gradient is required for the transport of two lumenal oxygen-evolving proteins across the thylakoid membrane. J Biol Chem **266:**12189-12193.

81. **Nair, N., R. Dziedzic, R. Greendyke, S. Muniruzzaman, M. Rajagopalan, and M. V. Madiraju.** 2009. Synchronous replication initiation in novel *Mycobacterium tuberculosis* dnaA cold-sensitive mutants. Mol Microbiol **71:**291-304.

82. **Nesvizhskii, A. I., A. Keller, E. Kolker, and R. Aebersold.** 2003. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem **75:**4646-4658.

83. **Neyrolles, O., K. Gould, M. P. Gares, S. Brett, R. Janssen, P. O'Gaora, J. L. Herrmann, M. C. Prevost, E. Perret, J. E. Thole, and D. Young.** 2001. Lipoprotein access to MHC class I presentation during infection of murine macrophages with live mycobacteria. J Immunol **166:**447-457.

84. **Niederweis, M.** 2008. Nutrient acquisition by mycobacteria. Microbiology **154:**679-692.

85. **Niemann, G. S., R. N. Brown, J. K. Gustin, A. Stufkens, A. S. Shaikh-Kidwai, J. Li, J. E. McDermott, H. M. Brewer, A. Schepmoes, R. D. Smith, J. N. Adkins, and F. Heffron.** 2011. Discovery of novel secreted virulence factors from *Salmonella enterica* serovar Typhimurium by proteomic analysis of culture supernatants. Infect Immun **79:**33-43.

86. **Old, W. M., K. Meyer-Arendt, L. Aveline-Wolf, K. G. Pierce, A. Mendoza, J. R. Sevinsky, K. A. Resing, and N. G. Ahn.** 2005. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. Mol Cell Proteomics **4:**1487-1502.

87. **Palmer, T., and B. C. Berks.** 2012. The twin-arginine translocation (Tat) protein export pathway. Nat Rev Microbiol **10:**483-496.

88. **Pandey, A. K., and C. M. Sassetti.** 2008. Mycobacterial persistence requires the utilization of host cholesterol. Proc Natl Acad Sci U S A **105:**4376-4380.

89. **Park, H. D., K. M. Guinn, M. I. Harrell, R. Liao, M. I. Voskuil, M. Tompa, G. K. Schoolnik, and D. R. Sherman.** 2003. Rv3133c/dosR is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis*. Mol Microbiol **48:**833-843.

90. **Petersen, T. N., S. Brunak, G. von Heijne, and H. Nielsen.** 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods **8:**785-786.

91. **Pethe, K., D. L. Swenson, S. Alonso, J. Anderson, C. Wang, and D. G. Russell.** 2004. Isolation of *Mycobacterium tuberculosis* mutants defective in the arrest of phagosome maturation. Proc Natl Acad Sci U S A **101:**13642-13647.

92. **Piddington, D. L., A. Kashkouli, and N. A. Buchmeier.** 2000. Growth of *Mycobacterium tuberculosis* in a defined medium is very restricted by acid pH and Mg(2+) levels. Infect Immun **68:**4518-4522.

93. **Poirier, V., and Y. Av-Gay.** 2012. *Mycobacterium tuberculosis* modulators of the macrophage's cellular events. Microbes Infect **14:**1211-1219.

94. **Pradel, N., C. L. Santini, C. Y. Ye, L. Fevat, F. Gerard, M. Alami, and L. F. Wu.** 2003. Influence of *tat* mutations on the ribose-binding protein translocation in *Escherichia coli*. Biochem Biophys Res Commun **306:**786-791.

95. **Raman, S., R. Hazra, C. C. Dascher, and R. N. Husson.** 2004. Transcription regulation by the *Mycobacterium tuberculosis* alternative sigma factor SigD and its role in virulence. J Bacteriol **186:**6605-6616.

96. **Raynaud, C., C. Guilhot, J. Rauzier, Y. Bordat, V. Pelicic, R. Manganelli, I. Smith, B. Gicquel, and M. Jackson.** 2002. Phospholipases C are involved in the virulence of *Mycobacterium tuberculosis*. Mol Microbiol **45:**203-217.

97. **Rengarajan, J., B. R. Bloom, and E. J. Rubin.** 2005. Genome-wide requirements for *Mycobacterium tuberculosis* adaptation and survival in macrophages. Proc Natl Acad Sci U S A **102:**8327-8332.

98. **Renier, S., C. Chambon, D. Viala, C. Chagnot, M. Hebraud, and M. Desvaux.** 2013. Exoproteomic analysis of the SecA2-dependent secretion in *Listeria monocytogenes* EGD-e. J Proteomics **80:**183-195.

99. **Rigel, N. W., and M. Braunstein.** 2008. A new twist on an old pathway--accessory Sec systems. Mol Microbiol **69:**291-302.

100. **Rigel, N. W., H. S. Gibbons, J. R. McCann, J. A. McDonough, S. Kurtz, and M. Braunstein.** 2009. The accessory SecA2 system of mycobacteria requires ATP binding and the canonical SecA1. J Biol Chem **284:**9927-9936.

101. **Rohde, K. H., R. B. Abramovitch, and D. G. Russell.** 2007. *Mycobacterium tuberculosis* invasion of macrophages: linking bacterial gene expression to environmental cues. Cell Host Microbe **2:**352-364.

102. **Rose, R. W., T. Bruser, J. C. Kissinger, and M. Pohlschroder.** 2002. Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. Mol Microbiol **45:**943-950.

103. **Rouviere, P. E., and C. A. Gross.** 1996. SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins. Genes Dev **10:**3170-3182.

104. **Sachdeva, P., R. Misra, A. K. Tyagi, and Y. Singh.** 2010. The sigma factors of *Mycobacterium tuberculosis*: regulation of the regulators. FEBS J **277:**605-626.

105. **Saint-Joanis, B., C. Demangel, M. Jackson, P. Brodin, L. Marsollier, H. Boshoff, and S. T. Cole.** 2006. Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of *Mycobacterium tuberculosis*, increases {beta}-lactam susceptibility and virulence. J Bacteriol **188:**6669-6679.

106. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2001. Comprehensive identification of conditionally essential genes in mycobacteria. Proc Natl Acad Sci U S A **98:**12712-12717.

107. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol **48:**77-84.

108. **Sassetti, C. M., and E. J. Rubin.** 2003. Genetic requirements for mycobacterial survival during infection. Proc Natl Acad Sci U S A **100:**12989-12994.

109. **Schnappinger, D., S. Ehrt, M. I. Voskuil, Y. Liu, J. A. Mangan, I. M. Monahan, G. Dolganov, B. Efron, P. D. Butcher, C. Nathan, and G. K. Schoolnik.** 2003. Transcriptional Adaptation of *Mycobacterium tuberculosis* within Macrophages: Insights into the Phagosomal Environment. J Exp Med **198:**693-704.

110. **Senaratne, R. H., B. Sidders, P. Sequeira, G. Saunders, K. Dunphy, O. Marjanovic, J. R. Reader, P. Lima, S. Chan, S. Kendall, J. McFadden, and L.**

W. Riley. 2008. *Mycobacterium tuberculosis* strains disrupted in mce3 and mce4 operons are attenuated in mice. J Med Microbiol **57:**164-170.

111. **Sharma, D., A. Bose, H. Shakila, T. K. Das, J. S. Tyagi, and V. D. Ramanathan.** 2006. Expression of mycobacterial cell division protein, FtsZ, and dormancy proteins, DevR and Acr, within lung granulomas throughout guinea pig infection. FEMS Immunol Med Microbiol **48:**329-336.

112. **Sherman, D. R., M. Voskuil, D. Schnappinger, R. Liao, M. I. Harrell, and G. K. Schoolnik.** 2001. Regulation of the *Mycobacterium tuberculosis* hypoxic response gene encoding alpha -crystallin. Proc Natl Acad Sci U S A **98:**7534-7539.

113. **Shimohata, N., Y. Akiyama, and K. Ito.** 2005. Peculiar properties of DsbA in its export across the *Escherichia coli* cytoplasmic membrane. J Bacteriol **187:**3997-4004.

114. **Shimono, N., L. Morici, N. Casali, S. Cantrell, B. Sidders, S. Ehrt, and L. W. Riley.** 2003. Hypervirulent mutant of *Mycobacterium tuberculosis* resulting from disruption of the mce1 operon. Proc Natl Acad Sci U S A **100:**15918-15923.

115. **Shruthi, H., M. M. Babu, and K. Sankaran.** 2010. TAT-pathway-dependent lipoproteins as a niche-based adaptation in prokaryotes. J Mol Evol **70:**359-370.

116. **Sonnhammer, E. L., G. von Heijne, and A. Krogh.** 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc Int Conf Intell Syst Mol Biol **6:**175-182.

117. **Stewart, G. R., J. Patel, B. D. Robertson, A. Rae, and D. B. Young.** 2005. Mycobacterial mutants with defective control of phagosomal acidification. PLoS Pathog **1:**269-278.

118. **Storey, J. D., and R. Tibshirani.** 2003. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A **100:**9440-9445.

119. **Sturgill-Koszycki, S., P. H. Schlesinger, P. Chakraborty, P. L. Haddix, H. L. Collins, A. K. Fok, R. D. Allen, S. L. Gluck, J. Heuser, and D. G. Russell.** 1994. Lack of acidification in *Mycobacterium phagosomes* produced by exclusion of the vesicular proton-ATPase. Science **263:**678-681.

120. **Sullivan, J. T., E. F. Young, J. R. McCann, and M. Braunstein.** 2012. The *Mycobacterium tuberculosis* SecA2 system subverts phagosome maturation to promote growth in macrophages. Infect Immun **80:**996-1006.

121. **Sutcliffe, I. C., and D. J. Harrington.** 2004. Lipoproteins of *Mycobacterium tuberculosis*: an abundant and functionally diverse class of cell envelope components. FEMS Microbiol Rev **28:**645-659.

122. **Titgemeyer, F., J. Amon, S. Parche, M. Mahfoud, J. Bail, M. Schlicht, N. Rehm, D. Hillmann, J. Stephan, B. Walter, A. Burkovski, and M. Niederweis.** 2007. A genomic view of sugar transport in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. J Bacteriol **189:**5903-5915.

123. **Tschumi, A., T. Grau, D. Albrecht, M. Rezwan, H. Antelmann, and P. Sander.** 2012. Functional analyses of mycobacterial lipoprotein diacylglyceryl transferase and comparative secretome analysis of a mycobacterial *lgt* mutant. J Bacteriol **194:**3938-3949.

124. **Tullius, M. V., G. Harth, and M. A. Horwitz.** 2003. Glutamine synthetase GlnA1 is essential for growth of *Mycobacterium tuberculosis* in human THP-1 macrophages and guinea pigs. Infect Immun **71:**3927-3936.

125. **Tullman-Ercek, D., M. P. DeLisa, Y. Kawarasaki, P. Iranpour, B. Ribnicky, T. Palmer, and G. Georgiou.** 2007. Export pathway selectivity of *Escherichia coli* twin arginine translocation signal peptides. J Biol Chem **282:**8309-8316.

126. **Valent, Q. A., D. A. Kendall, S. High, R. Kusters, B. Oudega, and J. Luirink.** 1995. Early events in preprotein recognition in *E. coli*: interaction of SRP and trigger factor with nascent polypeptides. EMBO J **14:**5494-5505.

127. **Vandal, O. H., L. M. Pierini, D. Schnappinger, C. F. Nathan, and S. Ehrt.** 2008. A membrane protein preserves intrabacterial pH in intraphagosomal *Mycobacterium tuberculosis*. Nat Med **14:**849-854.

128. **Voskuil, M. I., D. Schnappinger, K. C. Visconti, M. I. Harrell, G. M. Dolganov, D. R. Sherman, and G. K. Schoolnik.** 2003. Inhibition of respiration by nitric oxide induces a *Mycobacterium tuberculosis* dormancy program. J Exp Med **198:**705-713.

129. **Wooff, E., S. L. Michell, S. V. Gordon, M. A. Chambers, S. Bardarov, W. R. Jacobs, Jr., R. G. Hewinson, and P. R. Wheeler.** 2002. Functional genomics reveals the sole sulphate transporter of the *Mycobacterium tuberculosis complex* and its relevance to the acquisition of sulphur in vivo. Mol Microbiol **43:**653-663.

130. **WorldHealthOrganization.** 2012. WHO Information tuberculosis fact sheet.

131. **Wu, C. H., J. J. Tsai-Wu, Y. T. Huang, C. Y. Lin, G. G. Lioua, and F. J. Lee.** 1998. Identification and subcellular localization of a novel Cu,Zn superoxide dismutase of *Mycobacterium tuberculosis*. FEBS Lett **439:**192-196.

132. **Xie, K., and R. E. Dalbey.** 2008. Inserting proteins into the bacterial cytoplasmic membrane using the Sec and YidC translocases. Nat Rev Microbiol **6:**234-244.

133. **Yahr, T. L., and W. T. Wickner.** 2001. Functional reconstitution of bacterial Tat translocation in vitro. EMBO J **20:**2472-2479.

CHAPTER 4

Export of a Mycobacterial SecA2 Substrate is Influenced by the Twin-arginine Translocation Pathway[1]

The majority of bacterial protein export occurs via the conserved Sec export pathway, which is powered by the essential SecA ATPase. SecA recognizes preproteins with signal peptides and powers their translocation through the SecYEG membrane channel. Mycobacteria and some Gram positive bacteria have two non-redundant SecA homologs. While SecA1 is the essential ATPase required for housekeeping Sec export, SecA2 is required for export of a smaller set of proteins. Some SecA2 proteins utilize an accessory membrane channel called SecY2. SecA2-only systems, which include those in mycobacteria, lack SecY2 and are thought to function in concert with the SecA1/SecYEG machinery. Previous studies in *M. smegmatis* demonstrate that features of the mature domain of a preprotein determine if SecA2 is required for export. In addition, prior work shows the mature domain of the SecA2-dependent Ms1704 protein is compatible with export by the Tat pathway. Here, we show that export of Ms1704 containing its native signal peptide is not only SecA2-dependent, as we observed before, but it is also influenced by the Tat pathway. Because the Tat system only exports folded

[1]Contributing authors: Lauren S. Ligon, Ellen Young, and Miriam Braunstein (Department of Microbiology and Immunology, School of Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7290)

proteins across the cytoplasmic membrane, these results support our hypothesis that the defining feature of the mature domains of SecA2-dependent preproteins of mycobacteria is a tendency to fold in the cytoplasm. Our data also suggest that some Ms1704 export occurs via SecA2/SecYEG, while additional Ms1704 export occurs in a Sec-independent manner using the Tat machinery. Thus, SecA2 may serve an important role in exporting some preproteins with cytoplasmic folding tendencies through SecYEG.

Introduction

Protein export is a fundamental aspect of bacterial physiology, allowing proteins to be transported past the cytoplasmic membrane so they can fulfill necessary roles in cell envelope maintenance, nutrient acquisition, or other vital functions (8, 11, 48). Bacterial pathogens, including *Mycobacterium tuberculosis*, also rely on extra-cytoplasmic proteins for modulating the host environment to promote bacterial survival and growth (71). Mycobacterial species possess two conserved pathways for powering protein export: the general secretion (Sec) and twin-arginine translocation (Tat) systems (39).

General Sec export is universal and responsible for the majority of housekeeping protein export, making Sec systems essential in all bacteria including mycobacteria (9). Proteins destined for Sec export, which are known as preproteins, are synthesized with N-terminal signal peptides, which contain a positively charged N-terminus, hydrophobic core, and C-terminal cleavage region (28). Upon recognition by the cytoplasmic SecA protein, preproteins are delivered to the membrane-embedded SecYEG protein complex (9). In addition to SecA-targeting, the Sec signal peptide also has an important role in interacting with SecY to stabilize an open SecY channel conformation (24, 31, 38, 64).

SecA is an ATPase that undergoes cycles of ATP hydrolysis to drive translocation of the unfolded preprotein through the SecYEG channel (15). Following or during export, periplasmic signal peptidases cleave the signal peptide and the exported protein then folds to assume a mature confirmation (50).

Unlike Sec export systems, the Tat pathway is not universal, but widespread among bacteria (13, 46, 49). Tat export is essential in *M. tuberculosis* and *tat* mutants are not viable. The *Mycobacterium smegmatis* Tat pathway, however, is not essential and *tat* mutants exist in the non-pathogenic model species (43, 53, 61). In contrast to the Sec system that only exports unfolded preproteins, the Tat pathway is built to accommodate proteins that fold in the cytoplasm, and will even reject unfolded substrates (12, 20). Tat preproteins contain N-terminal signal peptides with tripartite structures similar to those of Sec preproteins, but Tat signal peptides are characterized by the presence of the conserved twin-arginine motif: S/T-R-R-X-F-L-K, or more generically as R-R-X-Φ-Φ, where ϕ represents a hydrophobic residue (3). The Tat signal peptide first targets a folded preprotein to the TatBC protein complex in the cytoplasmic membrane (37, 42, 56, 70). TatA monomers are then recruited to the TatBC-preprotein complex with energy supplied by the proton motive force (23, 45, 79, 81). The preprotein is translocated across the membrane through the membrane-spanning TatA pore and the signal peptide is removed by periplasmic signal peptidases (22, 40).

In addition to the conserved Sec and Tat protein export pathways, *M. tuberculosis* also has specialized protein export pathways that have important roles during infection. The SecA2-dependent protein export system is one such specialized pathway (39). SecA2 systems are found in all mycobacteria and some Gram positive species and are

characterized by the presence of a second non-redundant SecA homolog, which is called SecA2 to distinguish it from the canonical SecA described above (57). While the canonical SecA, which is called SecA1 in these systems, has an essential role in powering general Sec export, SecA2 is dispensable for in vitro growth but necessary for export of a subset of proteins (6, 27, 58). The SecA2 proteins are functionally conserved between pathogenic *M. tuberculosis* and the non-pathogenic *M. smegmatis*; consequently, *M. smegmatis* is often used as a model to study mycobacterial SecA2 export (58).

There are two types of SecA2 systems: SecA2–SecY2 systems include an accessory SecY2 membrane channel, while SecA2-only systems lack an obvious alternative protein-conducting channel (18). It is well established that SecA2–SecY2 systems are dedicated to exporting proteins that are first glycosylated in the cytoplasm, a post-translational modification that interferes with export by the canonical SecA1/SecYEG machinery (2, 10). SecA2-only systems, like those in mycobacteria, are thought to utilize one or more components of the canonical pathway to promote protein export (including SecA1 and SecYEG), but this relationship is not well understood (17, 58). In contrast to SecA2–SecY2 systems, the characteristics that direct preproteins to SecA2 in SecA2-only systems is only beginning to be explored, but it is clear that the same targeting rules of SecA2–SecY2 preproteins do not apply to substrates of SecA2-only systems.

In *M. smegmatis* there are two known SecA2 substrates: Ms1704 and Ms1712 (21). Both are lipoproteins and predicted to function as the solute-binding components of two ABC transporters (48). It has been demonstrated that although Ms1704 and Ms1712 contain Sec signal peptides required for export to the cell wall, these signal peptides are

not specific for targeting to SecA2 (19). Instead, there are one or more features of the mature domains of these mycobacterial SecA2 preproteins that necessitate the requirement for SecA2 for export. The feature(s) of the mature domain of mycobacterial substrates that defines SecA2-dependency is unlikely to include glycosylation as seen with substrates of the SecA2–SecY2 systems (19). Interestingly, the Ms1704 mature domain can be efficiently exported by the Tat system (19). Because the Tat pathway only exports folded proteins, it has been used in the past as a reporter for cytoplasmic folding (12, 20, 56). The Tat compatibility of the mature domain of Ms1704 suggests that SecA2 substrates can fold in the cytoplasm and this folding propensity may be the feature that distinguishes SecA2 substrates from SecA2-independent Sec preproteins.

In our proteomic study described in Chapter 3 we discovered that a number of predicted *M. tuberculosis* Tat substrates (presumably preproteins that fold in the cytoplasm) depend on SecA2 for maximum export. These results raised the possibility of there being a subset of preproteins that are compatible with either the SecA2 or Tat pathways for transport across the cytoplasmic membrane. Many of the predicted Tat substrates that were influenced by SecA2 in *M. tuberculosis* are members of the solute-binding protein (SBP) family (4, 69). This was a striking discovery because the two known *M. smegmatis* SecA2 substrates, Ms1704 and Ms1712, are also SBP proteins (21, 48). Moreover, a genomic survey of predicted Tat substrates identifies numerous SBPs, across a diverse set of organisms, suggesting that it is common for SBPs to fold in the cytoplasm to be Tat exported (63).

Here, we test the influence of the Tat machinery on export of a SecA2 substrate in *M. smegmatis*. Surprisingly, inactivation of the *tatC* component of the Tat pathway

negatively impacted export of the SecA2-dependent Ms1704 substrate. Our results suggest there may be a common subset of preproteins that are compatible for export by either the SecA2 or Tat systems. Our data also reaffirms our developing model where SecA2 fulfills a role in facilitating export of certain folded preproteins via the SecA1/SecYEG pathway.

Materials and Methods

Bacterial strains and growth conditions: *Escherichia coli* DH5α was used for all DNA cloning and grown at 37º C in Luria-Bertani medium (Fisher) supplemented with 40 µg/ml kanamycin or 150 µg/mL hygromycin when appropriate. *M. smegmatis* mc^2155 (wild type) (65), NR116 (Δ*secA2*) (58), JM567 (Δ*tatC*) (43), or LL154 (Δ*tatC* Δ*secA2)* were grown at 30º C in Mueller-Hinton medium (BD Diagnostic Systems) with 0.1% (v/v) Tween 80 and supplemented with 20 µg/mL kanamycin and 50 µg/mL hygromycin when appropriate.

Construction of *M. smegmatis* Δ*tatC* Δ*secA2* mutant: Strain LL154, the *M. smegmatis* Δ*tatC* Δ*secA2* mutant, was constructed by two-step allelic exchange, as described previously (58), using the Δ*tatC* mutant JM567 strain and the Δ*secA2* suicide plasmid pNR6. Successful deletion of *secA2* was confirmed by Southern blot analysis (data not shown).

Plasmid construction and *M. smegmatis* electroporation: All plasmids and oligonucleotides used in this study are listed in Tables 4.1 and 4.2, respectively. All cloned plasmid inserts were confirmed error-free by DNA sequencing (Eton Biosciences, NC). Plasmids were electroporated into *M. smegmatis* as previously described (5).

Creation of TatC expression plasmid. To create a plasmid expressing *M. tuberculosis* TatC under control of the *hsp60* promoter, *tatC* (*Rv2093c*) was amplified from *M. tuberculosis* H37Rv genomic DNA with primers ForwMtbTacCwXm and RevMtbTatCwPst and cloned into pCR2.1 to yield pEY3.The TatC piece was then excised by digestion with PstI and XmnI and ligated into similarly cut pMV261 to yield plasmid pEY04. The P_{hsp60}-$tatC_{Mtb}$ piece was then moved from EY04 into the integrating plasmid pMV361 to yield pMF253.

Subcellular fractionation: Subcellular fractions of *M. smegmatis* cells were generated as previously described (21, 59). Briefly, whole cell lysates of *M. smegmatis* were generated by passing cells 4 times through a French pressure cell at 20,000 psi. Unlysed cells were removed by centrifugation at 3,000 x g to generate clarified whole cell lysates (WCL). Protein concentrations of WCLs were determined by BCA assay using a BSA standard (Pierce) and equalized between all strains. One milliliter of each equalized lysate was centrifuged at 27,000 x g for 30 minutes to pellet cell wall material (CW), then 100,000 x g for 2 hours to separate membrane (M) and soluble (SOL) fractions. The soluble fraction contains the cytosolic components.

Immunoblotting: Whole cell lysate (WCL) and subcellular fraction proteins (CW, MEM, and SOL) derived from the same amount of starting cells between strains for each experiment were analyzed by immunoblotting. Protein samples were combined with 1X SDS-PAGE buffer, boiled, and separated on a 12% SDS acrylamide gel. Proteins were transferred to a nitrocellulose membrane and detected using either anti-HA (Covance) used at 1:10,000, anti-PhoA (Research Diagnostics International) used at 1:20,000, anti-19kD (provided by Douglas Young, MRC National Institute for Medical Research) used

at 1:20,000, or anti-MspA (provided by Michael Niederweis, University of Alabama at Birmingham) used at 1:20,000. Proteins were detected using either anti-rabbit or anti-mouse secondary antibodies conjugated to alkaline phosphatase as appropriate (GE Healthcare). Signal was detected using ECF reagent (GE Healthcare).

Results

Export of the SecA2-dependent Ms1704 is influenced by the Tat pathway

The fact that multiple predicted Tat substrates of *M. tuberculosis* are affected by deletion of *secA2* suggested that there might be a common set of preproteins compatible with both SecA2 and Tat export pathways (Chapter 3). In support of this idea, the mature domain of the Ms1704 SecA2 substrate is compatible with Tat export if engineered to possess a Tat signal peptide (Chapter 2) (19). Therefore, we wanted to determine if a functional Tat pathway influenced export of a SecA2-dependent protein containing its native signal peptide. For this purpose, we used a previously characterized *M. smegmatis* Δ*tatC* deletion mutant, which is deficient for Tat-dependent export (43). We first expressed a C-terminal HA-tagged version of the SecA2 substrate Ms1704 in wild type *M. smegmatis*, the Δ*secA2* deletion mutant, or the Δ*tatC* deletion mutant. Cultures were grown to an approximate OD_{600} of 0.5-1.0 at 30° C and then cells harvested by centrifugation. Because the Δ*tatC* mutant grows slower, starting Δ*tatC* mutant cultures were inoculated with approximately twice the number of cells as the wild type and Δ*secA2* mutant strains, and harvested at the same time as the other two strains. Export of Ms1704-HA to the cell wall was analyzed as follows. For each strain equalized whole cells lysates (WCLs) were generated and subjected to differential ultracentrifugation to

isolate cell wall (CW), cytoplasmic membrane (M) and the cytosol-containing soluble fractions. Whole cell lysates and all corresponding subcellular factions were separated by SDS-PAGE, then proteins were transferred to a membrane for immunoblotting using an anti-HA antibody.

As previously published, Ms1704-HA was readily exported to the cell wall in wild type *M. smegmatis*, but this export was compromised upon deletion of *secA2* (Figure 4.1A) (19, 21, 58). Surprisingly, when we expressed Ms1704-HA in the Δ*tatC* mutant, export to the cell wall was completely abolished. However, the total amount of Ms1704 was also reduced in the whole cell lysate. The reduced level of Ms1704 in the whole cell lysate was reminiscent of the protein instability that is observed with certain Tat substrates in the absence of export (12, 62, 75). Therefore, the reduction of Ms1704-HA in the Δ*tatC* cell wall is likely a reflection of an export defect. Export of the control cell wall porin, MspA, was unaffected by deletion of either *secA2* or *tatC* as expected, and indicated that comparable levels of total protein were present in the cell wall samples of all three strains.

Export of the SecA2-independent Ms6020 is influenced by the Tat pathway

Previously, we showed that export of another solute-binding protein, Ms6020, occurred independently of SecA2, despite relatively high amino acid similarity to *M. smegmatis* SecA2 substrates (48% similar to the Ms1704 mature domain and 41% similar to the Ms1712 mature domain) (19). In light of our observation that Ms1704 was influenced by the Tat pathway we set out to test if the SecA2-independent Ms6020 was similarly affected by deletion of *tatC*. As previously demonstrated, export of Ms6020-HA

to the cell wall occurred in wild type *M. smegmatis* but was unaffected by deletion of *secA2* (Figure 4.1B) (19). Similar to Ms1704-HA, the level of Ms6020-HA was also severely compromised in the Δ*tatC* mutant cell wall, consistent with there being an export defect in this strain. Also like Ms1704-HA, total Ms6020-HA levels were reduced in the Δ*tatC* whole cell lysate. Again, endogenous MspA levels were equivalent in all three strains expressing Ms6020-HA, indicating that total cell wall protein levels remained similar.

The Tat pathway effect is specific to export of Ms1704 and Ms6020

For our cell wall localization experiments, both Ms1704-HA and Ms6020-HA were expressed under control of a constitutive promoter on a multi-copy plasmid. We were concerned that the subcellular fractionation phenotypes observed for Ms1704 and Ms6020 in the Δ*tatC* mutant might be artifacts from over-expressing these proteins in this mutant strain background. Therefore, we also tested the impact of *tatC* deletion on two other exported SecA2-independent substrates expressed from the same promoter and multi-copy plasmid. The 19kDa lipoprotein is an *M. tuberculosis* protein previously shown to be exported to the cell wall independent of SecA2 when expressed in *M. smegmatis* (21). As shown in Figure 4.1C, export of the 19kDa lipoprotein to the cell wall occurred equally well in wild type and the Δ*secA2* mutant as expected. Furthermore, 19kD was expressed well in the Δ*tatC* mutant and 19kD export to the cell wall was unchanged in this strain compared to wild type and the Δ*secA2* mutant. We also tested the effect of *tatC* deletion on export of the $PhoA_{Msmeg}$ alkaline phosphatase, which is another mycobacterial lipoprotein exported independently of SecA2 (21). $PhoA_{Msmeg}$ export to the

cell wall was unaffected by deletion of *secA2* as expected. Additionally $PhoA_{Msmeg}$ export readily occurred in the *M. smegmatis* Δ*tatC* mutant.

Taken together, these results demonstrate that export of two *M. smegmatis* solute-binding proteins (SBPs), one that requires SecA2 for export (Ms1704) and one that is SecA2-independent (Ms6020), is compromised in a Δ*tatC* mutant. Total protein levels of both these SBPs were also reduced in the Δ*tatC* whole cell lysate, which is likely reflective of protein instability in the absence of export. Furthermore, these Δ*tatC*-dependent phenotypes are not an artifact of over-expression.

***Export of Ms1704 is completely abolished by simultaneous deletion of* secA2 *and* tatC**

The Ms1704-HA export defect associated with deletion of *secA2* is not always complete. It is common to see some residual Ms1704-HA exported to the *M. smegmatis* cell wall in the Δ*secA2* mutant (Figure 4.2A and unpublished observations). Given the results above, we considered the possibility that the Tat export system accounts for the residual Ms1704-HA export seen in the Δ*secA2* mutant. Conversely, Δ*tatC* deletion does not always result in 100% abrogation of Ms1704-HA localization to the cell wall (Figure 4.2A, long exposure). If export of Ms1704 is shared between the SecA2 and Tat export pathways, then absolutely no Ms1704-HA export should occur in a Δ*secA2* Δ*tatC* double mutant.

To test this, we generated a *M. smegmatis* Δ*secA2* Δ*tatC* double mutant by deleting *secA2* in the Δ*tatC* mutant background using allelic exchange as previously described (58). This Δ*secA2* Δ*tatC* double mutant retained previously published phenotypes of the Δ*tatC* mutant strain used, including both a slow in vitro growth

phenotype and sensitivity to beta-lactam antibiotics (data not shown) (43). Additionally, deletion of *secA2* resulted in acquisition of the previously published increased azide sensitivity of the *M. smegmatis* Δ*secA2* deletion mutant (58).

We next expressed Ms1704-HA and compared its export in the following four *M. smegmatis* strains: wild type, the Δ*secA2* mutant, the Δ*tatC* mutant, and the newly constructed Δ*secA2* Δ*tatC* double mutant (Figure 4.2A). As expected Ms1704-HA was exported to the cell wall in wild type *M. smegmatis*, and this export was reduced in the Δ*secA2* mutant. Again, we also observed an Ms1704-HA cell wall localization defect and reduced levels of Ms1704-HA in the whole cell lysate of the Δ*tatC* mutant. The phenotype of the Δ*secA2* Δ*tatC* double mutant was particularly interesting for two reasons. First, deletion of *secA2* rescued the reduced level of Ms1704-HA protein in the whole cell lysate associated with absence of *tatC*. Second, the Δ*secA2* Δ*tatC* double mutant did, in fact, exhibit a complete export defect of Ms1704-HA to the cell wall.

These results suggest that Ms1704-HA export is influenced by both SecA2 and TatC in *M. smegmatis*, and a strain lacking both export machinery components cannot export any Ms1704-HA to the cell wall. A striking observation is that the presumed instability of Ms1704-HA in the Δ*tatC* mutant is rescued by deletion of *secA2*.

*Expression of SecA2 partially complements the Ms1704 export defect of Δ***secA2** *Δ***tatC** *double mutant*

The Ms1704-HA export defect in the *M. smegmatis* Δ*secA2* mutant is a phenotype that has previously been complemented by introduction of a *secA2* expressing plasmid (21, 58). However, we wanted to determine if the Ms1704-HA export defect in the newly

created Δ*secA2* Δ*tatC* double mutant strain could also be complemented by exogenous expression of SecA2. In addition, we wanted to test the possibility that over-expression of SecA2 might suppress the Ms1704-HA export defect of the Δ*tatC* mutant.

To test this, we transformed an integrating plasmid expressing the *secA2* gene from *M. tuberculosis*, pWRJA2, into either the Δ*secA2* mutant, Δ*tatC* mutant, or the Δ*secA2* Δ*tatC* double mutant. We then expressed Ms1704-HA in all of these strains and performed an immunoblot on subcellular fractions to look for Ms1704-HA export to the cell wall. As expected, exogenous expression of $SecA2_{Mtb}$ restored Ms1704-HA export to the cell wall of the Δ*secA2* mutant (Figure 4.2B). In the *tatC* deletion mutant, the endogenous *secA2* gene is still present, but very little Ms1704-HA export is seen. When SecA2 was over-expressed in this strain (by providing the $secA2_{Mtb}$ gene on pWRJA2), the Ms1704-HA export defect remains the same; we were unable to detect any restoration of Ms1704-HA export. Finally, expression of $SecA2_{Mtb}$ in the Δ*secA2* Δ*tatC* double mutant resulted in partial, but not complete, restoration of Ms1704-HA export (Figure 4.2B). This observation holds true over two independent biological replicates (data not shown).

Taken together, these results show that increased levels of SecA2 protein achieved with pWRJA2, cannot suppress the export defect of Ms1704-HA observed in the Δ*tatC* mutant. In addition, introduction of a $secA2_{Mtb}$ expression plasmid alone into the Δ*secA2* Δ*tatC* double mutant can partially, but not fully, complement the Ms1704-HA export defect of the double mutant. These results are in line with the observation that some portion of Ms1704-HA export is reliant on the Tat pathway.

Expression of TatC complements the Ms1704 export defect of the* ΔsecA2, ΔtatC, *and the Δ*secA2 Δ*tatC mutants.

Next, we wanted to complement the export phenotypes of Ms1704-HA observed in the *M. smegmatis* Δ*tatC* mutant to determine if they are due to inactivation of the *tatC* gene. To test this, we constructed a TatC complementing plasmid. The *tatC* gene from *M. tuberculosis* H37Rv (*rv2093c*) was cloned under the *hsp60* promoter and expressed from an integrating plasmid. This plasmid, pMF253, was shown to produce functional TatC protein by first confirming its ability to complement the Δ*tatC* mutant phenotypes of slow growth and beta-lactam sensitivity (data not shown) (43).

We then transformed pMF253 into either the Δ*secA2* mutant, Δ*tatC* mutant, or Δ*secA2* Δ*tatC* double mutant. We then expressed Ms1704-HA in all of these strains and performed an immunoblot on subcellular fractions to look for Ms1704-HA export to the cell wall. As expected, exogenous expression of the $tatC_{Mtb}$ gene restored Ms1704-HA export in the Δ*tatC* single deletion strain (Figure 4.2C), indicating that the observed Ms1704-HA export defect of this strain was due to deletion of *tatC*. Furthermore, the plasmid expressing $tatC_{Mtb}$ also restored to wild type levels the amount of Ms1704-HA protein in the whole cell lysate. Thus, both the protein stability and export defect of Ms1704-HA in the Δ*tatC* mutant are attributable to the deletion of *tatC*.

The Δ*secA2* deletion mutant still has the endogenous *tatC* gene present in the *M. smegmatis* genome. In the experiment shown in Figure 4.2A, an Ms1704-HA export defect was observed in the Δ*secA2* mutant although residual export of Ms1704-HA to the cell wall was seen. This result is not surprising to us and it is consistent with there being functional TatC that can also promote export of Ms1704 in the Δ*secA2* mutant. A striking

observation was that over-expression of *tatC* achieved by providing the *M. tuberculosis* *tatC* gene on pMF253 restored export of Ms1704-HA in the Δ*secA2* mutant (Figure 4.2C). This restoration of Ms1704-HA export in the Δ*secA2* mutant strain by over-expression of *tatC* was comparable to the Δ*secA2* complemented with $SecA2_{Mtb}$ (Figure 4.2C). Finally, expression of *tatC* from pMF253 restored Ms1704-HA export in the Δ*secA2* Δ*tatC* double mutant to a level comparable to that seen in the wild type strain.

Taken together, our results show that Ms1704-HA export is affected by both deletion of *secA2* and *tatC*. Furthermore, the Ms1704-HA export defect of the Δ*tatC* mutant can be complemented by exogenous expression of *tatC* and over-expression of *tatC* can restore Ms1704-HA export in the Δ*secA2* mutant.

Discussion

In contrast to the better characterized SecA2–SecY2 systems, the mechanisms of SecA2-only export are just beginning to be explored. The current model of export by SecA2-only systems, including those of mycobacteria, is that SecA2 functions with machinery from the canonical Sec pathway including SecA1 and SecYEG (17, 58). A defining feature of mycobacterial SecA2 substrates is their mature domain (19). In Chapter 2, we showed that the mature domain of a SecA2 substrate is compatible for export by the Tat pathway (19). This was demonstrated by fusing a Tat signal peptide to the mature domain of the Ms1704 SecA2 substrate (19). Export of this signal peptide chimera occurred in a Tat-dependent manner. We therefore hypothesized that a tendency to fold in the cytoplasm is the defining feature of the mature domains of SecA2-dependent preproteins in mycobacteria.

In the past, we attributed any residual export of Ms1704 observed in the Δ*secA2* mutant to be a result of export via the canonical SecA1/SecYEG pathway. Here, we provide evidence that the residual export of Ms1704 is instead occurring via the Tat machinery, not the Sec pathway. First, we demonstrated that export of Ms1704 containing its native signal peptide was reduced in a Δ*tatC* mutant. The Δ*tatC* export defect was similar to that of the Δ*secA2* mutant, meaning that export was reduced compared to a wild type strain, but a small amount of residual Ms1704-HA export could sometimes be detected in the cell wall (Figure 4.2A, long exposure). However, in a Δ*secA2* Δ*tatC* double mutant Ms1704-HA export was completely abolished. Furthermore, we showed that expression of *tatC* from a plasmid was able to restore Ms1704-HA export to wild type levels in the Δ*secA2* mutant. Taken together, these results suggest that a portion of Ms1704 can be exported in a SecA2-dependent manner, while additional Ms1704 export can occur via the Tat pathway. These results are significant in suggesting that in mycobacteria, there is a relationship between the SecA2 and Tat export systems at the level of a shared subset of preproteins, which includes Ms1704. Because the Tat pathway only accommodates folded proteins, this reinforces a predicted role of SecA2 in exporting preproteins that are prone to cytoplasmic folding.

Ms1704 can be exported by either the SecA2 or Tat export pathways

Our results suggest there may be a common subset of preproteins, including Ms1704, that are compatible with export by either the SecA2 or Tat systems. This reaffirms a developing model where SecA2 fulfills a role in adapting some preproteins

containing mature domains with folding characteristics of Tat substrates for export by the canonical Sec system (Figure 4.4).

Protein export by the conserved Sec and Tat pathways are generally thought of as independent processes, as Sec export is limited to unfolded proteins and the Tat system only accommodates folded proteins (47). This leads to a simplistic idea that within the cytoplasm there are either unfolded preproteins that are strictly routed to the Sec machinery for export, or folded preproteins that are strictly routed to Tat for export. However, there are examples of preproteins in multiple bacterial species that appear compatible for export by either the Sec or Tat systems. For example, in *Streptomyces lividans* export of multiple proteins with presumed Sec signal peptides, including some solute-binding proteins like Ms1704, are reduced in a *tat* mutant (26). In *Bacillus subtilis*, there are numerous exported proteins containing predicted Tat signal peptides that are secreted in a *tat* mutant, presumably by the Sec system (34). Shared export by the Sec and Tat systems has also been noted in plant chloroplasts, which contain eukaryotic homologs of the Sec and Tat machinery (29). Therefore, the idea that Ms1704 can be exported by both the Sec and Tat pathways in mycobacteria is plausible.

Several lines of evidence suggest that the export of Ms1704 by the SecA2 and Tat systems are independent processes. In other words, SecA2 likely does not function in conjunction with the Tat machinery to promote export in mycobacteria. First, it has been demonstrated that depletion of SecA1 affects export of Ms1704, which is consistent with SecA2 functioning in conjunction with the SecA1/SecYEG machinery (58). Second, the Ms1704 export defects seen in the Δ*secA2* and Δ*tatC* single mutants are not 100%, while Ms1704 export in the Δ*secA2* Δ*tatC* double mutant is completely abolished. The additive

negative effects of deleting *secA2* and *tatC* on Ms1704 export suggests that SecA2 and TatC function independently. Third, Ms6020 export was reduced in the Δ*tatC* mutant, but unaffected by deletion of *secA2*. Additionally, results from our *M. tuberculosis* proteomic study in Chapter 3 showed that export of several predicted Tat substrates was reduced in the Δ*secA2* mutant, but these export defects were incomplete. Similar to the incomplete Ms1704-HA export defect in the *M. smegmatis* Δ*secA2* mutant, the partial export defects of the *M. tuberculosis* Δ*secA2* mutant suggests that SecA2 functions independent from TatC. Thus, while both SecA2 and Tat contribute to Ms1704 export, current data supports a model where these are independent export processes.

Targeting of SecA2 substrates to the Tat machinery

An important piece of data from this study is that it appears signal peptides compatible with both SecA2 and Tat pathways exist in mycobacteria. In *E. coli*, there are examples of strict Tat signal peptides that function exclusively with the Tat machinery, and also examples of "promiscuous" signal peptides that appear to function with both Sec and Tat (74). For promiscuous signal peptides, the deciding factor for whether a protein is exported by the Sec or Tat systems is the nature of the mature domain. Therefore, in addition to the signal peptide, the mature domain is important for determining whether a preprotein can be exported unfolded via Sec or folded via Tat (12). In this Chapter, we show that Ms1704 containing its native signal peptide is exported by both SecA2 and Tat pathways. This suggests that Ms1704 also has a promiscuous signal peptide that allows targeting to both Sec and Tat machinery. Importantly, our results suggest that in addition to a promiscuous signal peptide, the Ms1704 mature domain is also promiscuous in

nature, as it is able to adopt an unfolded conformation for SecA2-dependent export via SecYEG or maintain a folded conformation for Tat export.

The twin-arginine residues in Tat signal peptides are nearly invariant, and substitution of one or both arginine residues usually prevents Tat-dependent export (66). However, there are some examples of Tat signal peptides (both naturally occurring and engineered) that only have one arginine residue yet are compatible for Tat-dependent export (30, 32, 41, 44, 66). These unusual Tat signal peptides usually contain a lysine residue in place of one of the arginine residues. Therefore instead of an arginine pair in the signal peptide there is a "KR" or "RK" pair present. Both Ms1704 and Ms6020 are predicted to contain Sec signal peptides, and no current Tat peptide prediction program picks up either as a potential Tat signal peptide (1, 52, 60). However, Ms1704 contains an "RK" in the N-terminus of its native signal peptide while Ms6020 has a "KR" (Figure 4.3A). In addition, both Ms1704 and Ms6020 have hydrophobic amino acid residues following the "KR" or "RK" pair, which is another signature of the consensus Tat motif, R-R-X-Φ-Φ, where ϕ represents a hydrophobic residue (3, 51). Thus, it is certainly possible that Ms1704 and Ms6020 have signal peptides that are compatible with the Tat system that are overlooked by the current prediction programs.

One approach that has successfully been used in mycobacteria to test if signal peptides are compatible with export by the Tat pathway is to use a beta-lactamase reporter (43, 44). The *M. tuberculosis* BlaC beta-lactamase requires export by the Tat system in order to confer resistance to beta-lactam antibiotics. When the signal peptide from a Tat substrate is fused to 'BlaC (lacking its native Tat signal peptide), the resulting BlaC fusion protein can be exported and confer beta-lactam resistance, reporting on Tat

export. Importantly, the ‘BlaC reporter only works when exported by the Tat pathway and not by the Sec pathway. Efforts are currently underway to use the BlaC reporter system to test the signal peptides of Ms1704 and Ms6020 for Tat compatibility.

As an alternative strategy, we attempted to change the distribution of Ms1704 and Ms6020 preproteins away from the Tat machinery and towards the Sec machinery by switching their signal peptides with a stronger Sec signal peptide, one that is SecA2-independent. We created two signal peptide fusion chimeras where the endogenous signal peptides of Ms1704 and Ms6020 were exchanged with the predicted Sec signal peptide from *M. smegmatis* PhoA. Unlike the signal peptides of Ms1704 and Ms6020, there is no “KR” or “RK” pair in the PhoA signal peptide (Figure 4.3A). Previously, we showed such a fusion of the signal peptide of the *M. smegmatis* PhoA to Ms1704-HA (ssPhoA-Ms1704-HA) is exported in *M. smegmatis* in a SecA2-dependent manner (Chapter 2) (19). We predicted that export of ssPhoA-Ms1704 and ssPhoA-Ms6020 would occur independently of the Tat pathway because the PhoA signal peptide would only target the preproteins to the SecA2/SecYEG machinery. However, we were surprised to find that that these chimeric constructs resulted in slow growth of all *M. smegmatis* strains tested: wild type, the Δ*secA2* mutant, and the Δ*tatC* mutant (data not shown). We suspect that the toxicity issues associated with these signal peptide chimeras are due to blockage of canonical Sec export, which is an essential process (78). If a substrate is folded tightly upon engagement with the SecY channel, this is toxic to the cell, as it blocks essential housekeeping export functions (78). The ability of Ms1704 and Ms6020 to be exported by the Tat system suggests that they fold in the cytoplasm. Therefore, routing these Tat-compatible substrates to the Sec machinery could be detrimental to cell viability. While

these toxicity observations are in line with our current model, the Tat-dependency of the ssPhoA-Ms1704 and ssPhoA-Ms6020 chimeras is still currently being tested.

Another discrepancy that needs to be resolved is why Ms1704 and Ms6020 both have possible Tat-compatible signal peptides, are both Tat-dependent for export, yet only Ms1704 is influenced by SecA2. There are two possibilities to consider that would explain these results. First, the Ms6020 signal peptide might be better at directing a protein to the Tat pathway in comparison to Ms1704, which explains why despite the homology between the Ms1704 and Ms6020 mature domains, Ms6020 export is SecA2-independent. There is a naturally occurring Tat signal peptide in the *B. subtilis* TtrB protein that contains a "KR" motif, similar to Ms6020, and is compatible for Tat export (30). However, changing the "KR" pair of TtrB to "RK" (more similar to that of Ms1704) results in Tat-dependent export that is significantly slowed, but not completely blocked (30). Alternatively, it is possible that the mature domain of Ms6020 folds too rapidly for export by the Sec pathway, even with the aid of SecA2, while the mature domain of Ms1704 tends to remain partially folded and is more amendable to export assistance by SecA2.

A Model for shared export between SecA2 and Tat systems

The two SecA2 substrates of *M. smegmatis*, Ms1704 and Ms1712 are examples of putative sugar-binding proteins, which are part of the larger solute-binding protein (SBP) family (21, 72). Based on Tat signal peptide prediction, SBPs represent a common class of Tat substrates across a diverse set of organisms. There is also experimental evidence that SBP export is reduced by deletion of *tat* components in *E. coli*, *Agrobacterium*

tumefaciens, and *Bacillus subtilis* (14, 34, 54, 63). Thus, it appears that SBPs are a family of proteins whose mature domains tend to fold in the cytoplasm making them compatible for Tat export. Our results add some Tat exported mycobacterial SBPs to the list and, interestingly, show that a Tat compatible mycobacterial SBP (Ms1704) can also be exported in a SecA2-dependent fashion. In addition to Ms1704 and Ms1712, our proteomic study identified 13 putative SBPs underrepresented in the Δ*secA2 M. tuberculosis* cell wall (Chapter 3). Therefore, we find that export of many mycobacterial SBPs also tend to be influenced by SecA2. Many of these *M. tuberculosis* SBPs contain predicted Tat signal peptides with twin-arginine residues (Figure 4.3B). These observations suggest proteins that fold in the cytoplasm, like SBPs, are likely to demand SecA2 for Sec-dependent export and also have the option of export by the Tat pathway. Our developing model could also extend to the SecA2-only system of *Listeria monocytogenes*, where three SBPs also exhibit SecA2-dependent export defects and contain signal peptides with "KK" or "KR" motifs (55).

In our proposed model, there exists a population of preproteins (including Ms1704) that are compatible for export via SecYEG, with assistance of SecA2, or Sec-independent export via the Tat machinery. The purpose of routing preproteins to both export pathways is unknown at this time, but we can speculate on the advantage afforded to the bacterial cell through shared export between the Sec and Tat systems. There might be a benefit to exporting proteins such as Ms1704 in an unfolded conformation through SecYEG with the help of SecA2. Ms1704 is a predicted sugar-binding lipoprotein, and it might benefit the bacterial cell to not allow folding of Ms1704 in the cytoplasm where it might bind a sugar molecule that the bacterial cell could use as an energy source.

Somewhat analogous to this possibility, it has been suggested that it is advantageous for some Tat substrates to fold in the cytoplasm prior to export in order to prevent competitive binding with cofactors in the periplasm (51, 73). Therefore, the location where folding occurs can impact binding to cofactors or substrates.

An equally plausible explanation is that there is a selective pressure to maintain preproteins that utilize both Sec and Tat pathways for export in order to maximize and allow fine-tuning of protein export. This would prevent overwhelming of either export system during times of stress. There is evidence that the balance of Sec versus Tat-dependent export of shared substrates can be altered due to environmental changes. In *Bacillus subtilis*, increases in the salinity of the extra-cellular environment can drive export of some Tat preproteins to the Sec pathway (76, 77). The Sec-Tat balance of export can also be perturbed upon over-expression of proteins (35). In *M. smegmatis*, there might be environmental changes, perhaps changes in salinity or temperature, that perturb the kinetics of protein folding in the cytoplasm and demand adjustment of Sec and Tat usage for protein export. This might help explain why the severity of the SecA2-dependent phenotype of substrates, like Ms1704, varies across experiments. It would be interesting to determine if changes in the demand on the SecA2 system versus the Tat pathway extends to *M. tuberculosis* when inside the stressful environment of the macrophage, as SecA2 export is required for bacterial survival and growth during infection (7, 36). There is a cytoplasmic chaperone of *E. coli* that is activated under conditions of redox stress to specifically deal with increased protein mis-folding (25, 33). While unknown at this time, it is interesting to speculate that redox, pH, temperature, or other environmental changes associated with the phagosomal environment encountered

by *M. tuberculosis* could influence cytoplasmic folding and the demand on SecA2 for export.

As for the exact role of SecA2 in export, it is possible that SecA2 functions in providing the additional energy required for translocation through SecYEG of proteins with a propensity to fold. Alternatively, SecA2 could have a role in maintaining preproteins in an unfolded translocation competent state, prior to export by SecA1 and SecYEG. An interesting observation of these studies is that overall Ms1704 protein levels were reduced in the Δ*tatC* mutant, suggesting that Ms1704 was undergoing degradation in this strain in the absence of Tat export. The reduced levels of Ms1704 in the Δ*tatC* mutant were reversed upon simultaneously deletion of *secA2*. It has been demonstrated that SecA of *E. coli* exhibits a proofreading function, and can assist in the folding of cytoplasmic proteins that are not meant to be exported (16). Somewhat analogous to that type of proofreading function of *E. coli* SecA, the mycobacterial SecA2 protein might lead to degradation of Ms1704-HA when all the protein is exclusively routed to the Sec pathway. The accumulation of folded Ms1704-HA would be especially taxing to the Sec translocase in the *tatC* mutant, which is where we see the most Ms1704 degradation occurring. However, the molecular basis for the SecA2-dependent degradation we observed is unknown at this point.

Conclusion

While the exact function of the mycobacterial SecA2 protein is unknown, our data supports a model where SecA2 is required for allowing preproteins that fold in the cytoplasm to be competently exported by the Sec translocase. Our results also suggest

that there may be a pool of mycobacterial Tat substrates with unconventional Tat signal peptides that are currently not appreciated. Furthermore, these proteins with unappreciated Tat signal peptides could also represent novel SecA2 substrates. Thus, the acquisition of SecA2 appears to be an important adaptation of mycobacteria, possibly to prevent blockage of the essential SecYEG channel by preprotein folding. In addition to the Tat systems, the SecA2 system represents a parallel pathway for getting proteins that fold in the cytoplasm past the membrane.

Table 4.1 Plasmids used in this study

Plasmid	Genotype	Description	Source
pCR2.1	*bla aph* ColE1	TA cloning vector	Invitrogen
pMV261	*aph oriM* P_{hsp60} ColE1	Episomal mycobacterial shuttle vector	(67)
pMV361	*hyg int attP* ColE1	Integrating mycobacterial shuttle vector	(67)
pML440	*PAL5000 origin, ColE1 origin, hyg, pimyc-phoA*	*M. smegmatis* PhoA expression plasmid, episomal	(80)
pHSG58	*aph* P_{hsp60}*-Ms1704-HA oriM* ColE1	Ms1704-HA expression vector, episomal	(21)
pMF170	*aph* P_{hsp60}*-Ms6020-HA oriM* ColE1	Ms6020-HA expression vector, episomal	(19)
pHSG67	*aph* P_{hsp60}*-19kD oriM* ColE1	*M. tuberculosis* 19Kd expression vector, episomal	(21)
pEY03	*bla aph* ColE1	M. tuberculosis TatC in pCR2.1	This work
pEY04	*aph* P_{hsp60}*-tatC oriM* ColE1	*M. tuberculosis* TatC expression vector, episomal	This work
pMF253	*aph* P_{hsp60}*-tatC int attP* ColE1	*M. tuberculosis* TatC expression vector, integrating	This work
pWRJA2	*aph* P_{hsp60}*-Ms1712-HA oriM* ColE1	*M. tuberculosis* SecA2 expression vector, integrating	(68)

Table 4.2 Oligonucleotides used in this study

Primer Name	Sequence (5'-3')
ForwMtb TacCwXm	*GAATACGTTCGTGCGCGCCGCCGGTCTTC*
RevMtbTatCwPst	*CTGAGCTGGCTAGCGGGCCTACG*

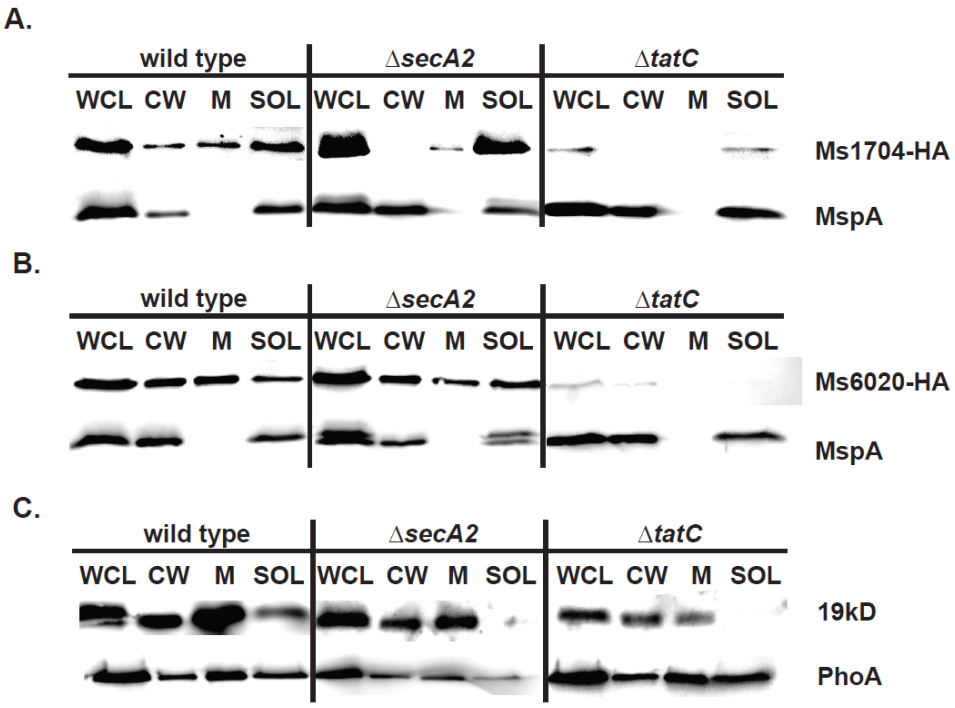

Figure 4.1 Export of Ms1704 and Ms6020 is dependent on TatC
Equalized whole cell lysates (WCL) generated from wild type, Δ*secA2*, and Δ*tatC M. smegmatis* cells expressing either **(A)** Ms1704-HA **(B)** Ms6020-HA **(C)** 19kDa or PhoA were subjected to ultracentrifugation to generate subcellular fractions. Fractions were separated by SDS-PAGE, and proteins detected with either anti-HA antibody, anti-PhoA, or anti-19kDa antibodies as appropriate. The total amount of cell wall (CW), membrane (M), and soluble (SOL) fractions shown is equivalent to the amount of WCL loaded. Native MspA was detected as a cell wall quality control. Parts A and B are representative blots from three biological replicates. Part C represents one biological replicate.

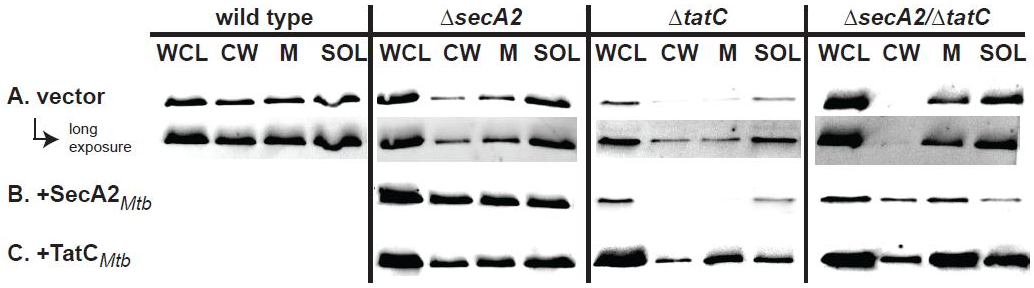

Figure 4.2 Export of Ms1704-HA requires both SecA2 and TatC, and deletion of *secA2* rescues reduced Ms1704 protein levels of *tatC* mutant.
Equalized whole cell lysates (WCL) generated from wild type, Δ*secA2*, Δ*tatC, or* Δ*secA2* Δ*tatC M. smegmatis* cells expressing Ms1704-HA and either **(A)** empty vector (pMV361) **(B)** *M. tuberculosis* SecA2 (pWRJA2) or **(C)** *M. tuberculosis* TatC (pMF253) were subjected to ultracentrifugation to generate subcellular fractions. Fractions were separated by SDS-PAGE, and the localization of Ms1704-HA determined using an anti-HA antibody. The total amount of cell wall (CW) and membrane (M) fractions shown is twice the corresponding amounts WCL and SOL samples loaded. The immunoblot shown is representative of two independent replicates. For part **(A)**: an additional longer exposure of the immunoblots is also included.

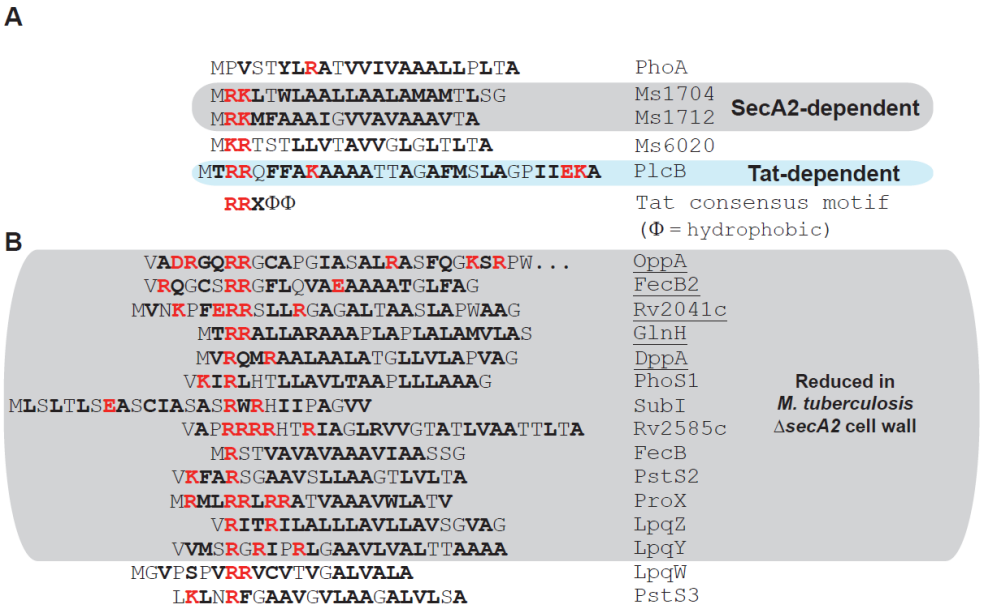

Figure 4.3 Signal peptides of SecA2-dependent SBPs
(A) Shown here are the signal peptides of SecA2-independent *M. smegmatis* PhoA and Ms6020 compared to the signal peptides of SecA2-dependent SBPs Ms1704 and Ms1712. A Tat signal peptide from the *M. tuberculosis* PlcB protein and the Tat consensus motif is also shown for comparison. **(B)** The Tat consensus in all 15 SBPs identified in a proteomic study of the *M. tuberculosis* cell wall (Chapter 3). Proteins that followed a trend of being reduced in the *M. tuberculosis* Δ*secA2* mutant are boxed in grey. *M. tuberculosis* proteins with predicted Tat signal peptides as determined by TatP and/or TatFind are underlined (1, 44, 60). All signal peptides are shown up to the amino acid adjacent to the cleavage site, except OppA where a portion of the signal peptide C-terminus was omitted. All positively charged residues are shown in red, while hydrophobic residues are shown in bold.

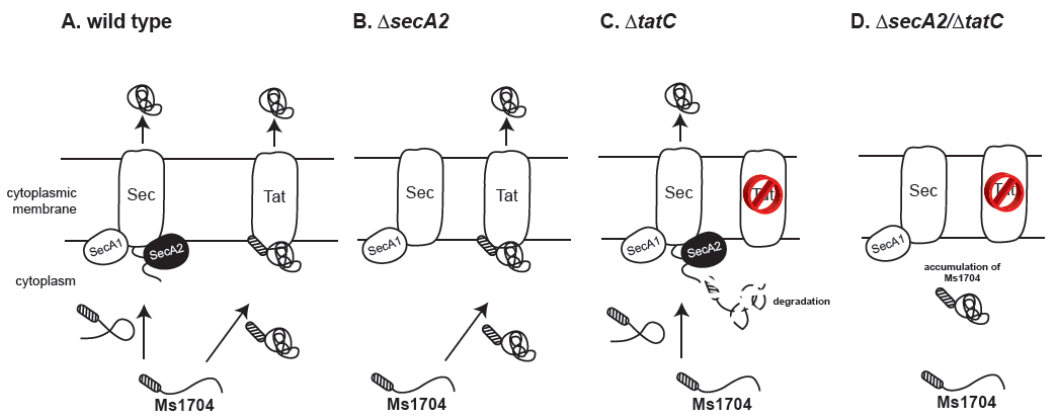

Figure 4.4 Model for shared export of Ms1704 by SecA2 and Tat pathways.
(A) In wild type *M. smegmatis*, Ms1704 has a signal peptide that allows interaction with both the Sec and Tat pathway. Ms1704 must be unfolded for export by the Sec pathway, which requires the assistance of SecA2. Export via Tat occurs with Ms1704 proteins that are fully folded. **(B)** In a Δ*secA2* mutant, all Ms1704 export occurs by the Tat pathway. **(C)** In a Δ*tatC* mutant, all Tat export is abolished (indicated by crossed out Tat translocase). All Ms1704 export occurs through the Sec pathway with assistance from SecA2. We propose that this situation is taxing to the essential Sec translocase leading to degradation of nonexported Ms1704. **(D)** In a Δ*secA2* Δ*tatC* double mutant Ms1704 is not exported at all. Ms1704 preprotein does not encounter machinery from either export system and can accumulate in the cytoplasmic (the instability of Ms1704 that occurs in the Δ*tatC* mutant is avoided).

References

1. **Bendtsen, J. D., H. Nielsen, D. Widdick, T. Palmer, and S. Brunak.** 2005. Prediction of twin-arginine signal peptides. BMC Bioinformatics **6:**167.

2. **Bensing, B. A., D. Takamatsu, and P. M. Sullam.** 2005. Determinants of the streptococcal surface glycoprotein GspB that facilitate export by the accessory Sec system. Mol Microbiol **58:**1468-1481.

3. **Berks, B. C.** 1996. A common export pathway for proteins binding complex redox cofactors? Mol Microbiol **22:**393-404.

4. **Braibant, M., P. Gilot, and J. Content.** 2000. The ATP binding cassette (ABC) transport systems of *Mycobacterium tuberculosis*. FEMS Microbiol Rev **24:**449-467.

5. **Braunstein, M., S. S. Bardarov, and W. R. J. Jacobs.** 2002. Genetic methods for deciphering virulence determinants of *Mycobacterium tuberculosis*, p. 67-99. *In* V. L. Clark and P. M. Bavoil (ed.), Methods in Enzymol, vol. 358. Academic Press, London.

6. **Braunstein, M., A. M. Brown, S. Kurtz, and W. R. Jacobs, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. J Bacteriol **183:**6979-6990.

7. **Braunstein, M., B. Espinosa, J. Chan, J. T. Belisle, and W. R. J. Jacobs.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium tuberculosis*. Mol Microbiol **48:**453-464.

8. **Calva, E., and R. Oropeza.** 2006. Two-component signal transduction systems, environmental signals, and virulence. Microb Ecol **51:**166-176.

9. **Chatzi, K. E., M. F. Sardis, S. Karamanou, and A. Economou.** 2013. Breaking on through to the other side: protein export through the bacterial Sec system. Biochem J **449:**25-37.

10. **Chen, Q., B. Sun, H. Wu, Z. Peng, and P. M. Fives-Taylor.** 2007. Differential roles of individual domains in selection of secretion route of a *Streptococcus parasanguinis* serine-rich adhesin, Fap1. J Bacteriol **189:**7610-7617.

11. **Crick, D. E., L. Quadri, and P. J. Brennan.** 2008. Biochemistry of the cell envelope of *Mycobacterium tuberculosis*, p. 1-20. *In* K. S.H.E and E. J. Rubin (ed.), Handbook of Tuberculosis: Molecular Biology and Biochemistry. WILEY-VCH Verlag GmbH & Co.

12. **DeLisa, M. P., D. Tullman, and G. Georgiou.** 2003. Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. Proc Natl Acad Sci U S A **100:**6115-6120.

13. **Dilks, K., R. W. Rose, E. Hartmann, and M. Pohlschroder.** 2003. Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. J Bacteriol **185:**1478-1483.

14. **Ding, Z., and P. J. Christie.** 2003. Agrobacterium tumefaciens twin-arginine-dependent translocation is important for virulence, flagellation, and chemotaxis but not type IV secretion. J Bacteriol **185:**760-771.

15. **Economou, A., and W. Wickner.** 1994. SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. Cell **78:**835-843.

16. **Eser, M., and M. Ehrmann.** 2003. SecA-dependent quality control of intracellular protein localization. Proc Natl Acad Sci U S A **100:**13231-13234.

17. **Fagan, R. P., and N. F. Fairweather.** 2011. *Clostridium difficile* has two parallel and essential Sec secretion systems. J Biol Chem **286:**27483-27493.

18. **Feltcher, M. E., and M. Braunstein.** 2012. Emerging themes in SecA2-mediated protein export. Nat Rev Microbiol **10:**779-789.

19. **Feltcher, M. E., H. S. Gibbons, L. S. Ligon, and M. Braunstein.** 2013. Protein export by the mycobacterial SecA2 system Is determined by the preprotein mature domain. J Bacteriol **195:**672-681.

20. **Fisher, A. C., W. Kim, and M. P. DeLisa.** 2006. Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway. Protein Sci **15:**449-458.

21. **Gibbons, H. S., F. Wolschendorf, M. Abshire, M. Niederweis, and M. Braunstein.** 2007. Identification of two *Mycobacterium smegmatis* lipoproteins exported by a SecA2-dependent pathway. J Bacteriol **189:**5090-5100.

22. **Gimenez, M. I., K. Dilks, and M. Pohlschroder.** 2007. *Haloferax volcanii* twin-arginine translocation substates include secreted soluble, C-terminally anchored and lipoproteins. Mol Microbiol **66:**1597-1606.

23. **Gohlke, U., L. Pullan, C. A. McDevitt, I. Porcelli, E. de Leeuw, T. Palmer, H. R. Saibil, and B. C. Berks.** 2005. The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. Proc Natl Acad Sci U S A **102:**10482-10486.

24. **Gouridis, G., S. Karamanou, I. Gelis, C. G. Kalodimos, and A. Economou.** 2009. Signal peptides are allosteric activators of the protein translocase. Nature **462:**363-367.

25. **Graf, P. C., and U. Jakob.** 2002. Redox-regulated molecular chaperones. Cell Mol Life Sci **59:**1624-1631.

26. **Guimond, J., and R. Morosoli.** 2008. Identification of *Streptomyces lividans* proteins secreted by the twin-arginine translocation pathway following growth with different carbon sources. Can J Microbiol **54:**549-558.

27. **Guo, X. V., M. Monteleone, M. Klotzsche, A. Kamionka, W. Hillen, M. Braunstein, S. Ehrt, and D. Schnappinger.** 2007. Silencing *Mycobacterium smegmatis* by using tetracycline repressors. J Bacteriol **189:**4614-4623.

28. **Hegde, R. S., and H. D. Bernstein.** 2006. The surprising complexity of signal sequences. Trends Biochem Sci **31:**563-571.

29. **Henry, R., M. Carrigan, M. McCaffrey, X. Ma, and K. Cline.** 1997. Targeting determinants and proposed evolutionary basis for the Sec and the Delta pH protein transport systems in chloroplast thylakoid membranes. J Cell Biol **136:**823-832.

30. **Hinsley, A. P., N. R. Stanley, T. Palmer, and B. C. Berks.** 2001. A naturally occurring bacterial Tat signal peptide lacking one of the 'invariant' arginine residues of the consensus targeting motif. FEBS Lett **497:**45-49.

31. **Hizlan, D., A. Robson, S. Whitehouse, V. A. Gold, J. Vonck, D. Mills, W. Kuhlbrandt, and I. Collinson.** 2012. Structure of the SecY complex unlocked by a preprotein mimic. Cell Rep **1:**21-28.

32. **Ignatova, Z., C. Hornle, A. Nurk, and V. Kasche.** 2002. Unusual signal peptide directs penicillin amidase from *Escherichia coli* to the Tat translocation machinery. Biochem Biophys Res Commun **291:**146-149.

33. **Jakob, U., W. Muse, M. Eser, and J. C. Bardwell.** 1999. Chaperone activity with a redox switch. Cell **96:**341-352.

34. **Jongbloed, J. D., H. Antelmann, M. Hecker, R. Nijland, S. Bron, U. Airaksinen, F. Pries, W. J. Quax, J. M. van Dijl, and P. G. Braun.** 2002. Selective contribution of the twin-arginine translocation pathway to protein secretion in *Bacillus subtilis*. J Biol Chem **277:**44068-44078.

35. **Kouwen, T. R., R. van der Ploeg, H. Antelmann, M. Hecker, G. Homuth, U. Mader, and J. M. van Dijl.** 2009. Overflow of a hyper-produced secretory protein from the *Bacillus* Sec pathway into the Tat pathway for protein secretion as revealed by proteogenomics. Proteomics **9:**1018-1032.

36. **Kurtz, S., K. P. McKinnon, M. S. Runge, J. P. Ting, and M. Braunstein.** 2006. The SecA2 secretion factor of *Mycobacterium tuberculosis* promotes growth in macrophages and inhibits the host immune response. Infect Immun.

37. **Lausberg, F., S. Fleckenstein, P. Kreutzenbeck, J. Frobel, P. Rose, M. Muller, and R. Freudl.** 2012. Genetic evidence for a tight cooperation of TatB and TatC during productive recognition of twin-arginine (Tat) signal peptides in *Escherichia coli*. PLoS One **7:**e39867.

38. **Li, W., S. Schulman, D. Boyd, K. Erlandson, J. Beckwith, and T. A. Rapoport.** 2007. The plug domain of the SecY protein stabilizes the closed state of the translocation channel and maintains a membrane seal. Molecular Cell **26:**511-521.

39. **Ligon, L. S., J. D. Hayden, and M. Braunstein.** 2012. The ins and outs of *Mycobacterium tuberculosis* protein export. Tuberculosis (Edinb) **92:**121-132.

40. **Luke, I., J. I. Handford, T. Palmer, and F. Sargent.** 2009. Proteolytic processing of *Escherichia coli* twin-arginine signal peptides by LepB. Arch Microbiol **191:**919-925.

41. **Marrichi, M., L. Camacho, D. G. Russell, and M. P. DeLisa.** 2008. Genetic toggling of alkaline phosphatase folding reveals signal peptides for all major modes of transport across the inner membrane of bacteria. J Biol Chem **283:**35223-35235.

42. **McDevitt, C. A., G. Buchanan, F. Sargent, T. Palmer, and B. C. Berks.** 2006. Subunit composition and in vivo substrate-binding characteristics of *Escherichia coli* Tat protein complexes expressed at native levels. FEBS J **273:**5656-5668.

43. **McDonough, J. A., K. E. Hacker, A. R. Flores, M. S. Pavelka, Jr., and M. Braunstein.** 2005. The twin-arginine translocation pathway of *Mycobacterium smegmatis* is functional and required for the export of mycobacterial beta-lactamases. J Bacteriol **187:**7667-7679.

44. **McDonough, J. A., J. R. McCann, E. M. Tekippe, J. S. Silverman, N. W. Rigel, and M. Braunstein.** 2008. Identification of functional Tat signal sequences in *Mycobacterium tuberculosis* proteins. J Bacteriol **190:**6428-6438.

45. **Mould, R. M., and C. Robinson.** 1991. A proton gradient is required for the transport of two lumenal oxygen-evolving proteins across the thylakoid membrane. J Biol Chem **266:**12189-12193.

46. **Murphy, C. K., E. J. Stewart, and J. Beckwith.** 1995. A double counter-selection system for the study of null alleles of essential genes in *Escherichia coli*. Gene **155:**1-7.

47. **Natale, P., T. Bruser, and A. J. Driessen.** 2008. Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. Biochim Biophys Acta **1778:**1735-1756.

48. **Niederweis, M.** 2008. Nutrient acquisition by mycobacteria. Microbiology **154:**679-692.

49. **Oliver, D., C. Kumamoto, M. Quinlan, and J. Beckwith.** 1982. Pleiotropic mutants affecting the secretory apparatus of *Escherichia coli*. Ann Inst Pasteur Mic **A133:**105-110.

50. **Paetzel, M., A. Karla, N. C. Strynadka, and R. E. Dalbey.** 2002. Signal peptidases. Chem Rev **102:**4549-4580.

51. **Palmer, T., and B. C. Berks.** 2012. The twin-arginine translocation (Tat) protein export pathway. Nat Rev Microbiol **10:**483-496.

52. **Petersen, T. N., S. Brunak, G. von Heijne, and H. Nielsen.** 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods **8:**785-786.

53. **Posey, J. E., T. M. Shinnick, and F. D. Quinn.** 2006. Characterization of the twin-arginine translocase secretion system of *Mycobacterium smegmatis*. J Bacteriol **188:**1332-1340.

54. **Pradel, N., C. Ye, V. Livrelli, J. Xu, B. Joly, and L. F. Wu.** 2003. Contribution of the twin arginine translocation system to the virulence of enterohemorrhagic *Escherichia coli* O157:H7. Infect Immun **71:**4908-4916.

55. **Renier, S., C. Chambon, D. Viala, C. Chagnot, M. Hebraud, and M. Desvaux.** 2013. Exoproteomic analysis of the SecA2-dependent secretion in *Listeria monocytogenes* EGD-e. J Proteomics **80:**183-195.

56. **Richter, S., and T. Bruser.** 2005. Targeting of unfolded PhoA to the TAT translocon of *Escherichia coli*. J Biol Chem **280:**42723-42730.

57. **Rigel, N. W., and M. Braunstein.** 2008. A new twist on an old pathway--accessory Sec systems. Mol Microbiol **69:**291-302.

58. **Rigel, N. W., H. S. Gibbons, J. R. McCann, J. A. McDonough, S. Kurtz, and M. Braunstein.** 2009. The accessory SecA2 system of mycobacteria requires ATP binding and the canonical SecA1. J Biol Chem **284:**9927-9936.

59. **Roberts, A. D., J. T. Belisle, A. M. Cooper, and I. M. Orme.** 1998. Specific Models: Murine Model of Tuberculosis, p. 389-417. *In* S. H. Kaufmann and D. Kabeltz (ed.), Methods in Microbiology, vol. 25.

60. **Rose, R. W., T. Bruser, J. C. Kissinger, and M. Pohlschroder.** 2002. Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. Mol Microbiol **45:**943-950.

61. **Saint-Joanis, B., C. Demangel, M. Jackson, P. Brodin, L. Marsollier, H. Boshoff, and S. T. Cole.** 2006. Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of *Mycobacterium tuberculosis*, increases {beta}-lactam susceptibility and virulence. J Bacteriol **188:**6669-6679.

62. **Santini, C. L., A. Bernadac, M. Zhang, A. Chanal, B. Ize, C. Blanco, and L. F. Wu.** 2001. Translocation of jellyfish green fluorescent protein via the Tat system of *Escherichia coli* and change of its periplasmic localization in response to osmotic up-shock. J Biol Chem **276:**8159-8164.

63. **Shruthi, H., M. M. Babu, and K. Sankaran.** 2010. TAT-pathway-dependent lipoproteins as a niche-based adaptation in prokaryotes. J Mol Evol **70:**359-370.

64. **Smith, M. A., W. M. Clemons, Jr., C. J. DeMars, and A. M. Flower.** 2005. Modeling the effects of prl mutations on the *Escherichia coli* SecY complex. J Bacteriol **187:**6454-6465.

65. **Snapper, S. B., R. E. Melton, S. Mustafa, T. Kieser, and W. R. Jacobs, Jr.** 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. Mol Microbiol **4:**1911-1919.

66. **Stanley, N. R., T. Palmer, and B. C. Berks.** 2000. The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. J Biol Chem **275:**11591-11596.

67. **Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennett, G. P. Bansal, J. F. Young, M. H. Lee, and G. F. Hatfull.** 1991. New use of BCG for recombinant vaccines. Nature **351:**456-460.

68. **Sullivan, J. T., E. F. Young, J. R. McCann, and M. Braunstein.** 2012. The *Mycobacterium tuberculosis* SecA2 system subverts phagosome maturation to promote growth in macrophages. Infect Immun **80:**996-1006.

69. **Sutcliffe, I. C., and D. J. Harrington.** 2004. Lipoproteins of *Mycobacterium tuberculosis*: an abundant and functionally diverse class of cell envelope components. FEMS Microbiol Rev **28:**645-659.

70. **Tarry, M. J., E. Schafer, S. Chen, G. Buchanan, N. P. Greene, S. M. Lea, T. Palmer, H. R. Saibil, and B. C. Berks.** 2009. Structural analysis of substrate binding by the TatBC component of the twin-arginine protein transport system. Proc Natl Acad Sci U S A **106:**13284-13289.

71. **Thi, E. P., U. Lambertz, and N. E. Reiner.** 2012. Sleeping with the enemy: how intracellular pathogens cope with a macrophage lifestyle. PLoS Pathog **8:**e1002551.

72. **Titgemeyer, F., J. Amon, S. Parche, M. Mahfoud, J. Bail, M. Schlicht, N. Rehm, D. Hillmann, J. Stephan, B. Walter, A. Burkovski, and M. Niederweis.** 2007. A genomic view of sugar transport in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. J Bacteriol **189:**5903-5915.

73. **Tottey, S., K. J. Waldron, S. J. Firbank, B. Reale, C. Bessant, K. Sato, T. R. Cheek, J. Gray, M. J. Banfield, C. Dennison, and N. J. Robinson.** 2008. Protein-folding location can regulate manganese-binding versus copper- or zinc-binding. Nature **455:**1138-1142.

74. **Tullman-Ercek, D., M. P. DeLisa, Y. Kawarasaki, P. Iranpour, B. Ribnicky, T. Palmer, and G. Georgiou.** 2007. Export pathway selectivity of *Escherichia coli* twin arginine translocation signal peptides. J Biol Chem **282:**8309-8316.

75. **Valle, M. T., A. M. Megiovanni, A. Merlo, G. Li Pira, L. Bottone, G. Angelini, L. Bracci, L. Lozzi, K. Huygen, and F. Manca.** 2001. Epitope focus, clonal composition and Th1 phenotype of the human CD4 response to the secretory mycobacterial antigen Ag85. Clin Exp Immunol **123:**226-232.

76. **van der Ploeg, R., U. Mader, G. Homuth, M. Schaffer, E. L. Denham, C. G. Monteferrante, M. Miethke, M. A. Marahiel, C. R. Harwood, T. Winter, M. Hecker, H. Antelmann, and J. M. van Dijl.** 2011. Environmental salinity determines the specificity and need for Tat-dependent secretion of the YwbN protein in *Bacillus subtilis*. PLoS One **6:**e18140.

77. **van der Ploeg, R., C. G. Monteferrante, S. Piersma, J. P. Barnett, T. R. Kouwen, C. Robinson, and J. M. van Dijl.** 2012. High-salinity growth conditions promote Tat-independent secretion of Tat substrates in *Bacillus subtilis*. Appl Environ Microbiol **78:**7733-7744.

78. **van Stelten, J., F. Silva, D. Belin, and T. J. Silhavy.** 2009. Effects of antibiotics and a proto-oncogene homolog on destruction of protein translocator SecY. Science **325:**753-756.

79. **White, G. F., S. M. Schermann, J. Bradley, A. Roberts, N. P. Greene, B. C. Berks, and A. J. Thomson.** 2010. Subunit organization in the TatA complex of the twin arginine protein translocase: a site-directed EPR spin labeling study. J Biol Chem **285:**2294-2301.

80. **Wolschendorf, F., M. Mahfoud, and M. Niederweis.** 2007. Porins are required for uptake of phosphates by *Mycobacterium smegmatis*. J Bacteriol **189:**2435-2442.

81. **Yahr, T. L., and W. T. Wickner.** 2001. Functional reconstitution of bacterial Tat translocation in vitro. EMBO J **20:**2472-2479.

CHAPTER 5

DISCUSSION

Tuberculosis disease is a major global health crisis. Nearly a third of the world's population is infected with *Mycobacterium tuberculosis* and this epidemic results in 1.4 million deaths annually (77). Several factors contribute to the large burden of *M. tuberculosis* infection and mortality including lack of an effective vaccine, co-infection with HIV, and the ability of this bacterium to establish latent infections (77). Despite the existence of anti-mycobacterial therapies, the incidence of multi-drug resistant *M. tuberculosis* strains continues to increase (65). A better understanding of *M. tuberculosis* physiology would aid in the development of new drugs (27).

Protein export systems are one particularly interesting aspect of *M. tuberculosis* biology to explore because exported proteins are ideally positioned to interact with the host during infection (55). Exported proteins also contribute to synthesizing and maintaining the unique architecture of the mycobacterial cell envelope and they are essential for acquiring nutrients from the extra-cellular environment (17, 51). Along with conserved protein export systems of *M. tuberculosis*, there exists the specialized SecA2-dependent protein export system. The SecA2 pathway contributes to the ability of *M. tuberculosis* to survive and grow inside the phagosomal compartments of host macrophages (13, 67). However, we are only beginning to understand the mechanisms behind SecA2 export and how SecA2-dependent exported proteins

contribute to *M. tuberculosis* virulence. The work presented in this dissertation helps to better define the relationship of the specialized SecA2 system to other conserved protein export machinery and it moves us toward a better understanding of the mechanism of SecA2-dependent export. Importantly, the proteomic analysis of the *M. tuberculosis* cell wall fractions that is part of this dissertation generated a list of putative SecA2-dependent exported effectors that not only help refine a mechanistic model of SecA2-mediated export, but could help explain the vital role of SecA2 export during *M. tuberculosis* infection.

The general Sec export pathway is the most widely conserved system for exporting proteins in bacteria, and is essential in all species including mycobacteria (22, 34, 63). The Sec system is comprised of the membrane-spanning SecYEG protein channel and a cytoplasmic SecA motor protein (15). Proteins are synthesized with N-terminal signal peptides for recognition by SecA. SecA then delivers these signal peptide containing preproteins to the cytoplasmic membrane for export. SecA is an ATPase that promotes translocation of the unfolded preprotein through SecYEG using cycles of ATP hydrolysis (23, 50). It is essential that the preprotein maintain an unfolded conformation during passage through the SecYEG channel and this is often achieved by assistance of cytoplasmic chaperones (3, 28). In addition to unfolded Sec exported preproteins there is also a category of preproteins that fold in the cytoplasm prior to their export. Such folded preproteins are translocated across the cytoplasmic membrane by the twin-arginine translocation (Tat) pathway (53). The Tat system only works with folded substrates and will even reject unfolded proteins (19). While the Tat pathway is not universally found in all bacteria, the mechanisms of Tat export appear to be conserved among a large group of

bacteria, including mycobacteria (20, 44, 48). Preproteins destined for Tat export are synthesized with N-terminal signal peptides that are distinguished from Sec signal peptides, specifically in that they contain two nearly invariant arginine residues (9). Tat preproteins first associate with the TatBC integral membrane proteins before TatA monomers are then recruited (30, 43, 47, 59, 69). The resulting TatA homomultimers form membrane-spanning channels of varying sizes in order to accommodate various folded proteins and protein complexes (32, 76).

The specialized SecA2 export systems are characterized by the presence of a second non-redundant SecA ATPase, called SecA2 (25). In bacteria with two SecA proteins, SecA1 fulfills a vital role in canonical Sec export, while SecA2 export is more specialized and limited to a smaller subset of proteins (12, 14, 24, 34, 61). SecA2 systems exist in all mycobacteria and a diverse set of Gram positive species (60). Two general types of SecA2 systems have been described. The SecA2–SecY2 systems appear to function independent of canonical Sec export and are dedicated to exporting a single glycosylated protein substrate (7, 16). The SecA2-only system, which lacks a SecY2 or obvious accessory protein-conducting channel, is the type of SecA2 system found in mycobacteria (12). The developing model of SecA2-only export is that SecA2 functions in concert with machinery from the canonical Sec system, including SecYEG and SecA1 (24, 61). General Sec export has been extensively studied for a long time, resulting in a relatively sophisticated understanding of the Sec export process (22). The idea that some SecA2 proteins evolved as an adaptation to the general Sec machinery is exciting as it suggests that there are capabilities and/or limitations to Sec export that have not been previously appreciated.

The work presented in this thesis focused on investigating the proteins exported by the mycobacterial SecA2 export system. We hypothesized that there exist specific features of SecA2-dependent preproteins that necessitate the involvement of SecA2 in their export. By studying two proteins exported by SecA2 in *Mycobacterium smegmatis*, we demonstrated a distinction between the SecA2-defining elements of substrates exported by SecA2–SecY2 systems and those exported by SecA2-only systems such as in mycobacteria. Our work also revealed a potential relationship between preproteins exported by the mycobacterial SecA2 and Tat pathways. Finally, our efforts to identify SecA2-dependent exported proteins in *M. tuberculosis* revealed new examples of proteins that appear to depend on SecA2 for localization to the cell wall including multiple solute-binding lipoproteins, components of two Mce transporters, and other exported proteins with putative roles in *M. tuberculosis* virulence. Many of these potential SecA2 substrates may account for the important role of the *M. tuberculosis* SecA2 pathway in virulence.

Defining features of mycobacterial SecA2 substrates reside in the mature domain of the protein and are distinct from SecA2-SecY2 systems.

SecA2–SecY2 systems, such as those in *Streptococcus gordonii* and *Streptococcus parasanguinis,* are dedicated to exporting substrates that are first heavily glycosylated in the cytoplasm. Studies in these organisms show that glycosylation of the preprotein blocks their export by the canonical Sec pathway (7, 16). In addition, at least two targeting elements have been described that route these modified substrates to the accessory SecA2–SecY2 machinery. These features include conserved amino acid residues in the signal peptide and the AST domain (for accessory sec transport) contained

in the N-terminus of the mature domain of the protein, the latter of which is required for binding SecA2 (5, 6, 8). Prior to the studies described in this dissertation, nothing was known about the feature(s) that impose SecA2-dependency on substrates of SecA2-only systems.

In *M. smegmatis*, SecA2 is required for export of Ms1704 and Ms1712 to the cell wall (31). These proteins are predicted solute-binding (SBP) components of two ABC transporters with putative roles in sugar uptake (70). Additionally, both are lipoproteins with seemingly typical N-terminal Sec signal peptides (54, 68, 72). By studying export of several signal peptide chimeric proteins, we demonstrated that although a signal peptide is required for Ms1704 and Ms1712 export, SecA2-specific signal peptides do not appear to exist in mycobacteria (Chapter 2) (26, 31). Hence, one or more features of the mature domains of Ms1704 and Ms1712 impart SecA2-dependency. Unlike the glycoprotein substrates of SecA2-SecY2 systems, neither Ms1704 nor Ms1712 appear to be glycosylated. In Chapter 2, we also determined that the extreme N-terminus of the Ms1704 mature domain is not required for export by the SecA2-dependent pathway, suggesting that mycobacterial SecA2 substrates do not contain a similar AST domain required for SecA2-targeting (26).

A mycobacterial SecA2 substrate can be exported by the Tat pathway

A remarkable finding described in Chapter 2, is that the mature domain of Ms1704 can be efficiently exported by the Tat pathway when fused to a Tat signal peptide (26). With the resulting ssPlcB-Ms1704 chimera, export of the Ms1704 protein occurred in a strictly SecA2-independent but Tat-dependent manner (26). Because the Tat

system only accommodates folded preproteins, this result implied that at least some cytoplasmic Ms1704 preprotein is folded (19). This result led to our hypothesis that a tendency to fold in the cytoplasm is the defining feature of the mature domains of SecA2-dependent preproteins in mycobacteria.

Inspection of the native signal peptide of Ms1704 does not reveal a twin arginine motif characteristic of Tat signal peptides, and the signal peptide is not predicted to be a Tat signal peptide by any of the current Tat signal peptide prediction programs (4, 62). Therefore, we were surprised to discover that Ms1704 containing its native signal peptide is reduced in the cell wall of an *M. smegmatis* Δ*tatC* mutant (Chapter 4). This suggested that some amount of Ms1704 is exported by the Tat system. In the Δ*tatC* mutant the export defect of native Ms1704 is not 100%. Similarly, a partial export defect for native Ms1704 is seen in the Δ*secA2* mutant. However, Ms1704 export is completely absent in a Δ*secA2* Δ*tatC* double mutant. Furthermore, export of Ms1704 could be restored in a strain lacking *secA2* by over-expressing *tatC*. The simplest interpretation of these data is that the Ms1704 preprotein is compatible with two separate export pathways. Ms1704 can either be exported in a SecA2-dependent fashion, presumably being translocated across the membrane via SecYEG in an unfolded state, or it can be exported in a folded conformation by the Tat machinery.

The idea that the signal peptide of Ms1704 can functionally target this preprotein to either the SecA2/Sec or Tat machinery is not out of the realm of possibility. This is because Tat and Sec signal peptides do share the same basic tripartite structure, including a positively charged N-terminus and hydrophobic stretch. Furthermore, there are examples of promiscuous signal peptides that can successfully route certain proteins to

either the Sec or Tat pathway depending on the folded state of the fused mature domain (10, 39, 73). Further, there are examples of functional Tat signal peptides that lack one of the twin arginine residues (35, 36, 46, 49, 66). Some of these Tat motif variants contain a "KR" pair similar to the native signal peptide of Ms1704. Efforts are currently underway to test the ability of the Ms1704 signal peptide to promote Tat-dependent export. If our model holds true, where some proteins like Ms1704 can be exported by two parallel systems (SecA2/Sec or Tat), this would significantly improve our understanding of both the SecA2-dependent and Tat export pathways of mycobacteria. Further, investigation of the signal peptides on such proteins may help to reveal other proteins that are exported by the SecA2 and/or Tat systems of mycobacteria that are not currently recognizable. When considered, this last possibility could aid efforts to identify novel virulence factors exported by the *M. tuberculosis* SecA2 system.

Similar substrates are exported by SecA2 and Tat pathways

The only two known SecA2-dependent exported proteins of *M. smegmatis* are solute-binding proteins (SBPs). SBPs are a diverse family of periplasmic proteins with roles in binding molecules for uptake via ABC-type membrane importers (11, 45). For reasons unknown at this point, SBPs are also a common class of Tat exported protein across a diverse set of organisms, as SBPs with predicted Tat signal peptides are widespread (64). Experimentally demonstrated Tat-dependent SBPs have also been described in *E. coli*, *Agrobacterium tumefaciens*, and *Bacillus subtilis* (21, 38, 57). In Chapter 3, we discovered that almost all the SBPs of *M. tuberculosis* are reduced in the cell wall of the Δ*secA2* mutant, notably four of these SBPs contain predicted Tat signal

peptides. Thus, it appears SBPs have features that frequently demand export by the Tat system and/or SecA2-dependency in *M. smegmatis* and *M. tuberculosis*. These observations are consistent with our model that folded preproteins are likely to demand export by SecA2, and provide further support that the defining feature of mycobacterial SecA2 substrates are cytoplasmic folding tendencies. The relationship of SecA2 to Tat export could also extend to other SecA2-only systems outside mycobacteria. In support of this idea, there are also three SBPs exported by the SecA2-only system of *Listeria monocytogenes* (58). Interestingly, a SBP of *E. coli* (the ribose-binding protein) is exported by the Sec system, but a subset of this SBP is also exported by the Tat system, particularly when Sec export is chemically inhibited (56). While *E. coli* does not possess a SecA2 protein, this report reaffirms a model where SBPs are a family of proteins that can exist in the cytoplasm in folded conformations competent for Tat export.

A model for SecA2 in protein export

The experiments described in this dissertation support the following model for SecA2-only export (Figure 5.1). The majority of Sec exported proteins are unfolded preproteins that are recognized by the canonical SecA1 protein and translocated in an unfolded conformation through SecYEG through cycles of SecA1 ATP hydrolysis. A subset of preproteins can fold in the cytoplasm, a property of the mature domain that demands specialized machinery for translocation across the cytoplasmic membrane. Folded preproteins containing strict Tat-targeting signal peptides are exported in a folded conformation through the Tat machinery. However, the folded state of a cytoplasmic preprotein can range from partially folded to a stable fully folded conformation (28, 41),

the latter of which is only compatible with the Tat system for export. If a partially folded preprotein has a promiscuous signal peptide (that can function in both Sec- and Tat-targeting) then this preprotein can follow one of two independent pathways. The preprotein can adopt a fully folded conformation for export by the Tat system. Alternatively, the partially folded preprotein can be translocated through SecYEG, but this requires the assistance of SecA2. Some SecA2 substrates might be preproteins containing strict Sec signal peptides. Other SecA2 substrates (like Ms1704) may have promiscuous signal peptides and mature domains capable of either Tat-mediated export or Sec-mediated export with SecA2.

The exact function of SecA2 during export of such select substrates (and how that function differs from SecA1) remains elusive. The first possibility is that SecA2 performs an "unfoldase" function, which has been described for other proteins like the bacterial AAA+ proteases (71). In this model, SecA2 would actively unfold the mature domains of preproteins in order to maintain competence for SecYEG-mediated export. However, this possibility seems unlikely because SecA2 does not have any predicted structural similarities to unfoldase proteins. The *M. tuberculosis* SecA1 and SecA2 proteins exhibit 50% similarity at the amino acid level, and most of the functional domains are conserved between these proteins. SecA2 is an ATPase and this enzymatic activity is required for its function, just like the canonical SecA1 protein (61). Given the high similarity to the canonical SecA1 proteins, it is more plausible that SecA2 performs a similar function to SecA1 in export.

Another possible function for SecA2 is to interact with preproteins in the cytoplasm that are overlooked by SecA1 and target these proteins to SecYEG for export.

In mycobacteria, SecA2 exhibits a primarily cytoplasmic localization, while SecA1 is evenly distributed between the cytoplasm and cytoplasmic membrane (61). Because it appears SecA2 is specifically needed for export of proteins more prone to cytoplasmic folding, it is possible there are chaperones that first interact with the preprotein before recognition by SecA2. These chaperones might include those with known roles in Sec export including GroEL or DnaK, which interact with nascent preproteins in order to maintain a conformation suitable for export and prevent inappropriate interactions leading to protein aggregation. (28). SecA2 could then deliver these proteins from the cytoplasmic chaperones to the SecA1/SecYEG machinery. Alternatively, SecA2 alone may be sufficient for recognizing preproteins, maintaining them in an unfolded state, then targeting them to the Sec translocase. This targeting function is analogous to the activity of the SecB chaperone described in *E. coli* canonical Sec export.

Another possibility is that SecA2 is required for energizing transport of certain preproteins through SecYEG. Proteins need to be unfolded for transport through SecYEG, and sometimes cytosolic chaperones aid in this process. However, it has also been shown that the force applied to a preprotein from translocation itself (powered by SecA) can unfold proteins (1, 52). It is possible that SecA2 is able to provide more energy for protein translocation of partially folded proteins while SecA1 cannot. In this case, the forward movement of the preproteins with the assistance of SecA2 ATP hydrolysis could unfold some regions of the preprotein. In fact, it has been demonstrated that loss of a particular structural loop in the motor portion of *E. coli* SecA results in hyperactive ATPase activity and an accelerated rate of ADP release (18). In mycobacteria, this variable subdomain (VAR) is present in SecA1 but absent in SecA2. If

SecA2 had this accelerated ATPase activity, it could power partially folded preproteins through SecYEG more efficiently than SecA1, quickly resulting in SecYEG channel access for translocation of the next preprotein. This proposed role for SecA2 in energizing export is not mutually exclusive from a possible function in Sec translocase-targeting.

Future investigating of the mycobacterial SecA2 system will undoubtedly benefit from a more reductionist approach afforded by in vitro translocation systems, which have been so valuable in generating the large body of literature on canonical Sec export. In vitro translocation is also starting to be applied to studying mechanisms of Tat export (78). In vitro translocation techniques could help elucidate potential energetic advantages of routing preproteins to either the Sec or Tat export pathways. This technique could also help determine if SecA2 has a SecA1/SecYEG-targeting function and exactly what level of preprotein folding SecA2 is capable of handling.

Proposed advantage of SecA2-mediated export in mycobacteria

If our model is correct that SecA2-mediated export is a strategy used by mycobacteria to export folded proteins, then an important question remains: why is SecA2 necessary and Tat export not sufficient for handling export of folded proteins? One plausible explanation for the existence of SecA2 in mycobacteria is that utilization of both Sec and Tat pathways allows a fine-tuning of protein export to prevent overwhelming of either system during times of stress on protein export. SecA2 could play an integral role in connecting Sec- and Tat-mediated export, by adapting over-flow from the Tat machinery for export via SecYEG. For preproteins containing promiscuous signal

peptides, the state of the mature domain will dictate whether preprotein export will occur via Sec or Tat. The proportion of unfolded, partially folded and fully folded preprotein in the cytoplasm could be influenced by environmental stress. Stresses that can contribute to preprotein unfolding are changes in temperature or oxidative stress (2, 29). Other environmental conditions that influence the Sec-Tat balance (presumably because they affect cytoplasmic folding) are salinity and protein over-expression (40, 74, 75). Thus perhaps, SecA2 has a vital role in allowing mycobacterial species to adapt protein export to changing environmental conditions. Certain environmental stresses could lead to an increase in the partially folded preprotein pool that SecA2 assists in exporting. In *M. smegmatis* these stresses might be changes in salinity or temperature that perturb the kinetics of protein folding in the cytoplasm and demand adjustment of SecA2 and Tat usage for protein export. This might explain why the severity of the SecA2-dependent phenotype of substrates, like Ms1704, varies across experiments. For *M. tuberculosis*, the stressful environment inside the macrophage could promote cytoplasmic preprotein unfolding of some Tat substrates, demanding that SecA2 assist in exporting these partially folded substrates through SecYEG. In support of this possibility it is known that oxidative stress, which is experienced by *M. tuberculosis* in the phagosome, can induce protein unfolding and aggregation in the cytoplasm in other bacteria (33, 37). While the effects of the phagosome environment on cytoplasmic protein folding have not been studied in *M. tuberculosis*, if bacterial preprotein unfolding increases inside the macrophage this could help explain why SecA2 is required for bacterial survival and growth during infection (13, 42).

Contribution of SecA2 exported proteins to *M. tuberculosis* virulence

Finally, the contribution of SecA2 to *M. tuberculosis* virulence remains to be reconciled. Our proteomic study of SecA2 export in *M. tuberculosis* revealed many changes in cell wall protein abundance upon deletion of *secA2*, but these changes were seen over a large set of proteins (Chapter 3). Because many of the export defects in the Δ*secA2* mutant were incomplete, it is likely that the attenuation of the *M. tuberculosis* Δ*secA2* mutant during infection is not due to reduced export of a single protein. Of particular interest for validation of SecA2-dependency and future work are the large array of solute-binding proteins and Mce transporters. Collectively, the impact of *secA2* deletion on the export of many SBPs with putative roles in nutrient acquisition could severely reduce the fitness of *M. tuberculosis*. Additionally, the *M. tuberculosis* Δ*secA2* mutant also exhibited reduced levels of some exported proteins with proposed roles in virulence, such as arresting phagosomal maturation. Together, the reduction in export of many proteins with contributions to *M. tuberculosis* survival and growth in the macrophage could explain the attenuation of the Δ*secA2* mutant. Finally, if there is overlap between the substrates exported by SecA2 and Tat pathways, and environmental stress leads to greater reliance on SecA2, then the full impact of SecA2 on protein export may not be apparent until *M. tuberculosis* is exposed to the stresses associated with infection. In line with this last idea, work in other bacterial species demonstrates that whether some exported proteins utilize Sec or Tat can change under certain conditions (40, 74, 75).

Taken together, the results described in this dissertation have advanced our understanding of the mechanisms of SecA2-only export. Importantly, our studies lead to a model where a defining feature of proteins exported by the SecA2 pathway of mycobacteria is a tendency to fold prior to export. In addition, we demonstrated a relationship between the types of preproteins exported by the SecA2 and Tat export systems. The idea that some proteins can use both the canonical Sec and Tat systems for export with the help of promiscuous signal peptides is not new. However, this is the first time that a function for SecA2 in adapting preproteins with Tat-compatible folding properties for export through the Sec machinery has been proposed. Our work suggests that SecA2 is an important addition to the canonical Sec and Tat translocation systems of mycobacteria, allowing efficient protein translocation of folded preproteins across the cytoplasmic membrane.

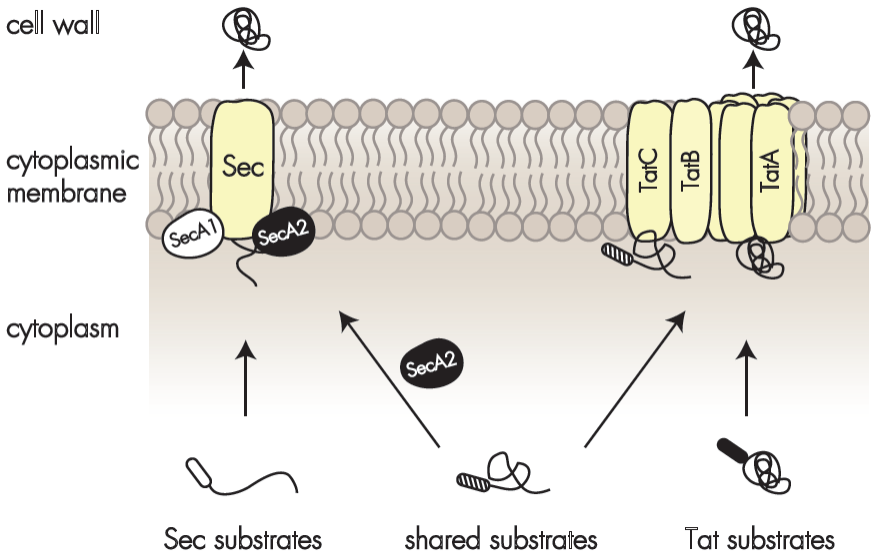

Figure 5.1 Model for mycobacterial SecA2 export
Substrates destined for Sec export are synthesized with Sec signal peptides (white oval). The essential SecA1 protein recognizes these preproteins and using cycles of ATP hydrolysis, drives the unfolded preproteins through the SecYEG membrane channel. Conversely, substrates destined for Tat export are synthesized with Tat signal peptides (black ovals). Tat preproteins fold in the cytoplasm, and are translocated across the cytoplasmic membrane by the TatABC machinery. SecA2 functions with the canonical SecA1/SecYEG machinery to assist in Sec export of some partially folded preproteins. We propose that there are shared substrates, like SBPs, that are partially or fully folded in the cytoplasm. If these shared substrates possess promiscuous signal peptides, they can be exported by Tat in a folded conformation or exported through SecYEG in an unfolded conformation with assistance from SecA2.

References

1. **Arkowitz, R. A., J. C. Joly, and W. Wickner.** 1993. Translocation can drive the unfolding of a preprotein domain. Embo J **12:**243-253.

2. **Baneyx, F., and M. Mujacic.** 2004. Recombinant protein folding and misfolding in *Escherichia coli*. Nat Biotechnol **22:**1399-1408.

3. **Bechtluft, P., N. Nouwen, S. J. Tans, and A. J. Driessen.** 2010. SecB--a chaperone dedicated to protein translocation. Mol Biosyst **6:**620-627.

4. **Bendtsen, J. D., H. Nielsen, D. Widdick, T. Palmer, and S. Brunak.** 2005. Prediction of twin-arginine signal peptides. BMC Bioinformatics **6:**167.

5. **Bensing, B. A., I. R. Siboo, and P. M. Sullam.** 2007. Glycine residues in the hydrophobic core of the GspB signal sequence route export toward the accessory Sec pathway. J Bacteriol **189:**3846-3854.

6. **Bensing, B. A., and P. M. Sullam.** 2010. Transport of preproteins by the accessory Sec system requires a specific domain adjacent to the signal peptide. J Bacteriol **192:**4223-4232.

7. **Bensing, B. A., D. Takamatsu, and P. M. Sullam.** 2005. Determinants of the streptococcal surface glycoprotein GspB that facilitate export by the accessory Sec system. Mol Microbiol **58:**1468-1481.

8. **Bensing, B. A., Y. T. Yen, R. Seepersaud, and P. M. Sullam.** 2012. A Specific interaction between SecA2 and a region of the preprotein adjacent to the signal peptide occurs during transport via the accessory Sec system. J Biol Chem **287:**24438-24447.

9. **Berks, B. C.** 1996. A common export pathway for proteins binding complex redox cofactors? Mol Microbiol **22:**393-404.

10. **Blaudeck, N., P. Kreutzenbeck, R. Freudl, and G. A. Sprenger.** 2003. Genetic analysis of pathway specificity during posttranslational protein translocation across the *Escherichia coli* plasma membrane. J Bacteriol **185:**2811-2819.

11. **Braibant, M., P. Gilot, and J. Content.** 2000. The ATP binding cassette (ABC) transport systems of *Mycobacterium tuberculosis*. FEMS Microbiol Rev **24:**449-467.

12. **Braunstein, M., A. M. Brown, S. Kurtz, and W. R. Jacobs, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. J Bacteriol **183:**6979-6990.

13. **Braunstein, M., B. Espinosa, J. Chan, J. T. Belisle, and W. R. J. Jacobs.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium tuberculosis*. Mol Microbiol **48:**453-464.

14. **Caspers, M., and R. Freudl.** 2008. *Corynebacterium glutamicum* possesses two *secA* homologous genes that are essential for viability. Arch Microbiol.

15. **Chatzi, K. E., M. F. Sardis, S. Karamanou, and A. Economou.** 2013. Breaking on through to the other side: protein export through the bacterial Sec system. Biochem J **449:**25-37.

16. **Chen, Q., B. Sun, H. Wu, Z. Peng, and P. M. Fives-Taylor.** 2007. Differential roles of individual domains in selection of secretion route of a *Streptococcus parasanguinis* serine-rich adhesin, Fap1. J Bacteriol **189:**7610-7617.

17. **Crick, D. E., L. Quadri, and P. J. Brennan.** 2008. Biochemistry of the cell envelope of *Mycobacterium tuberculosis*, p. 1-20. *In* K. S.H.E and E. J. Rubin (ed.), Handbook of Tuberculosis: Molecular Biology and Biochemistry. WILEY-VCH Verlag GmbH & Co.

18. **Das, S., L. M. Grady, J. Michtavy, Y. Zhou, F. M. Cohan, M. M. Hingorani, and D. B. Oliver.** 2012. The variable subdomain of *Escherichia coli* SecA functions to regulate SecA ATPase activity and ADP release. J Bacteriol **194:**2205-2213.

19. **DeLisa, M. P., D. Tullman, and G. Georgiou.** 2003. Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. Proc Natl Acad Sci U S A **100:**6115-6120.

20. **Dilks, K., R. W. Rose, E. Hartmann, and M. Pohlschroder.** 2003. Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. J Bacteriol **185:**1478-1483.

21. **Ding, Z., and P. J. Christie.** 2003. *Agrobacterium tumefaciens* twin-arginine-dependent translocation is important for virulence, flagellation, and chemotaxis but not type IV secretion. J Bacteriol **185:**760-771.

22. **du Plessis, D. J., N. Nouwen, and A. J. Driessen.** 2011. The Sec translocase. Biochim Biophys Acta **1808:**851-865.

23. **Economou, A., and W. Wickner.** 1994. SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. Cell **78:**835-843.

24. **Fagan, R. P., and N. F. Fairweather.** 2011. *Clostridium difficile* has two parallel and essential Sec secretion systems. J Biol Chem **286:**27483-27493.

25. **Feltcher, M. E., and M. Braunstein.** 2012. Emerging themes in SecA2-mediated protein export. Nat Rev Microbiol **10:**779-789.

26. **Feltcher, M. E., H. S. Gibbons, L. S. Ligon, and M. Braunstein.** 2013. Protein export by the mycobacterial SecA2 system Is determined by the preprotein mature domain. J Bacteriol **195:**672-681.

27. **Feltcher, M. E., J. T. Sullivan, and M. Braunstein.** 2010. Protein export systems of *Mycobacterium tuberculosis*: novel targets for drug development? Future Microbiol **5:**1581-1597.

28. **Fisher, A. C., and M. P. DeLisa.** 2004. A little help from my friends: quality control of presecretory proteins in bacteria. J Bacteriol **186:**7467-7473.

29. **Frees, D., K. Savijoki, P. Varmanen, and H. Ingmer.** 2007. Clp ATPases and ClpP proteolytic complexes regulate vital biological processes in low GC, Gram-positive bacteria. Mol Microbiol **63:**1285-1295.

30. **Frobel, J., P. Rose, F. Lausberg, A. S. Blummel, R. Freudl, and M. Muller.** 2012. Transmembrane insertion of twin-arginine signal peptides is driven by TatC and regulated by TatB. Nat Commun **3:**1311.

31. **Gibbons, H. S., F. Wolschendorf, M. Abshire, M. Niederweis, and M. Braunstein.** 2007. Identification of two *Mycobacterium smegmatis* lipoproteins exported by a SecA2-dependent pathway. J Bacteriol **189:**5090-5100.

32. **Gohlke, U., L. Pullan, C. A. McDevitt, I. Porcelli, E. de Leeuw, T. Palmer, H. R. Saibil, and B. C. Berks.** 2005. The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. Proc Natl Acad Sci U S A **102:**10482-10486.

33. **Graf, P. C., and U. Jakob.** 2002. Redox-regulated molecular chaperones. Cell Mol Life Sci **59:**1624-1631.

34. **Guo, X. V., M. Monteleone, M. Klotzsche, A. Kamionka, W. Hillen, M. Braunstein, S. Ehrt, and D. Schnappinger.** 2007. Silencing *Mycobacterium smegmatis* by using tetracycline repressors. J Bacteriol **189:**4614-4623.

35. **Hinsley, A. P., N. R. Stanley, T. Palmer, and B. C. Berks.** 2001. A naturally occurring bacterial Tat signal peptide lacking one of the 'invariant' arginine residues of the consensus targeting motif. FEBS Lett **497:**45-49.

36. **Ignatova, Z., C. Hornle, A. Nurk, and V. Kasche.** 2002. Unusual signal peptide directs penicillin amidase from *Escherichia coli* to the Tat translocation machinery. Biochem Biophys Res Commun **291:**146-149.

37. **Jakob, U., W. Muse, M. Eser, and J. C. Bardwell.** 1999. Chaperone activity with a redox switch. Cell **96:**341-352.

38. **Jongbloed, J. D., H. Antelmann, M. Hecker, R. Nijland, S. Bron, U. Airaksinen, F. Pries, W. J. Quax, J. M. van Dijl, and P. G. Braun.** 2002. Selective contribution of the twin-arginine translocation pathway to protein secretion in *Bacillus subtilis*. J Biol Chem **277:**44068-44078.

39. **Kolkman, M. A., R. van der Ploeg, M. Bertels, M. van Dijk, J. van der Laan, J. M. van Dijl, and E. Ferrari.** 2008. The twin-arginine signal peptide of *Bacillus subtilis* YwbN can direct either Tat- or Sec-dependent secretion of different cargo proteins: secretion of active subtilisin via the *B. subtilis* Tat pathway. Appl Environ Microbiol **74:**7507-7513.

40. **Kouwen, T. R., R. van der Ploeg, H. Antelmann, M. Hecker, G. Homuth, U. Mader, and J. M. van Dijl.** 2009. Overflow of a hyper-produced secretory protein from the *Bacillus* Sec pathway into the Tat pathway for protein secretion as revealed by proteogenomics. Proteomics **9:**1018-1032.

41. **Kulothungan, S. R., M. Das, M. Johnson, C. Ganesh, and R. Varadarajan.** 2009. Effect of crowding agents, signal peptide, and chaperone SecB on the folding and aggregation of *E. coli* maltose binding protein. Langmuir **25:**6637-6648.

42. **Kurtz, S., K. P. McKinnon, M. S. Runge, J. P. Ting, and M. Braunstein.** 2006. The SecA2 secretion factor of *Mycobacterium tuberculosis* promotes growth in macrophages and inhibits the host immune response. Infect Immun **74:**6855-6864.

43. **Lausberg, F., S. Fleckenstein, P. Kreutzenbeck, J. Frobel, P. Rose, M. Muller, and R. Freudl.** 2012. Genetic evidence for a tight cooperation of TatB and TatC during productive recognition of twin-arginine (Tat) signal peptides in *Escherichia coli*. PLoS One **7:**e39867.

44. **Ligon, L. S., J. D. Hayden, and M. Braunstein.** 2012. The ins and outs of *Mycobacterium tuberculosis* protein export. Tuberculosis (Edinb) **92:**121-132.

45. **Linton, K. J., and C. F. Higgins.** 1998. The *Escherichia coli* ATP-binding cassette (ABC) proteins. Mol Microbiol **28:**5-13.

46. **Marrichi, M., L. Camacho, D. G. Russell, and M. P. DeLisa.** 2008. Genetic toggling of alkaline phosphatase folding reveals signal peptides for all major modes of transport across the inner membrane of bacteria. J Biol Chem **283:**35223-35235.

47. **McDevitt, C. A., G. Buchanan, F. Sargent, T. Palmer, and B. C. Berks.** 2006. Subunit composition and in vivo substrate-binding characteristics of *Escherichia coli* Tat protein complexes expressed at native levels. FEBS J **273:**5656-5668.

48. **McDonough, J. A., K. E. Hacker, A. R. Flores, M. S. Pavelka, Jr., and M. Braunstein.** 2005. The twin-arginine translocation pathway of *Mycobacterium smegmatis* is functional and required for the export of mycobacterial beta-lactamases. J Bacteriol **187:**7667-7679.

49. **McDonough, J. A., J. R. McCann, E. M. Tekippe, J. S. Silverman, N. W. Rigel, and M. Braunstein.** 2008. Identification of functional Tat signal sequences in *Mycobacterium tuberculosis* proteins. J Bacteriol **190:**6428-6438.

50. **McNicholas, P., T. Rajapandi, and D. Oliver.** 1995. SecA proteins of Bacillus subtilis and Escherichia coli possess homologous amino-terminal ATP-binding domains regulating integration into the plasma membrane. J Bacteriol **177:**7231-7237.

51. **Niederweis, M.** 2008. Nutrient acquisition by mycobacteria. Microbiology **154:**679-692.

52. **Nouwen, N., G. Berrelkamp, and A. J. Driessen.** 2007. Bacterial sec-translocase unfolds and translocates a class of folded protein domains. J Mol Biol **372:**422-433.

53. **Palmer, T., and B. C. Berks.** 2012. The twin-arginine translocation (Tat) protein export pathway. Nat Rev Microbiol **10:**483-496.

54. **Petersen, T. N., S. Brunak, G. von Heijne, and H. Nielsen.** 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods **8:**785-786.

55. **Poirier, V., and Y. Av-Gay.** 2012. *Mycobacterium tuberculosis* modulators of the macrophage's cellular events. Microbes Infect **14:**1211-1219.

56. **Pradel, N., C. L. Santini, C. Y. Ye, L. Fevat, F. Gerard, M. Alami, and L. F. Wu.** 2003. Influence of *tat* mutations on the ribose-binding protein translocation in *Escherichia coli*. Biochem Biophys Res Commun **306:**786-791.

57. **Pradel, N., C. Ye, V. Livrelli, J. Xu, B. Joly, and L. F. Wu.** 2003. Contribution of the twin arginine translocation system to the virulence of enterohemorrhagic *Escherichia coli* O157:H7. Infect Immun **71:**4908-4916.

58. **Renier, S., C. Chambon, D. Viala, C. Chagnot, M. Hebraud, and M. Desvaux.** 2013. Exoproteomic analysis of the SecA2-dependent secretion in *Listeria monocytogenes* EGD-e. J Proteomics **80:**183-195.

59. **Richter, S., and T. Bruser.** 2005. Targeting of unfolded PhoA to the TAT translocon of *Escherichia coli*. J Biol Chem **280:**42723-42730.

60. **Rigel, N. W., and M. Braunstein.** 2008. A new twist on an old pathway--accessory Sec systems. Mol Microbiol **69:**291-302.

61. **Rigel, N. W., H. S. Gibbons, J. R. McCann, J. A. McDonough, S. Kurtz, and M. Braunstein.** 2009. The accessory SecA2 system of mycobacteria requires ATP binding and the canonical SecA1. J Biol Chem **284:**9927-9936.

62. **Rose, R. W., T. Bruser, J. C. Kissinger, and M. Pohlschroder.** 2002. Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. Mol Microbiol **45:**943-950.

63. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2001. Comprehensive identification of conditionally essential genes in mycobacteria. Proc Natl Acad Sci U S A **98:**12712-12717.

64. **Shruthi, H., M. M. Babu, and K. Sankaran.** 2010. TAT-pathway-dependent lipoproteins as a niche-based adaptation in prokaryotes. J Mol Evol **70:**359-370.

65. **Smith, T., K. A. Wolff, and L. Nguyen.** 2012. Molecular biology of drug resistance in *Mycobacterium tuberculosis*. Curr Top Microbiol Immunol.

66. **Stanley, N. R., T. Palmer, and B. C. Berks.** 2000. The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. J Biol Chem **275:**11591-11596.

67. **Sullivan, J. T., E. F. Young, J. R. McCann, and M. Braunstein.** 2012. The *Mycobacterium tuberculosis* SecA2 system subverts phagosome maturation to promote growth in macrophages. Infect Immun **80:**996-1006.

68. **Sutcliffe, I. C., and D. J. Harrington.** 2004. Lipoproteins of *Mycobacterium tuberculosis*: an abundant and functionally diverse class of cell envelope components. FEMS Microbiol Rev **28:**645-659.

69. **Tarry, M. J., E. Schafer, S. Chen, G. Buchanan, N. P. Greene, S. M. Lea, T. Palmer, H. R. Saibil, and B. C. Berks.** 2009. Structural analysis of substrate binding by the TatBC component of the twin-arginine protein transport system. Proc Natl Acad Sci U S A **106:**13284-13289.

70. **Titgemeyer, F., J. Amon, S. Parche, M. Mahfoud, J. Bail, M. Schlicht, N. Rehm, D. Hillmann, J. Stephan, B. Walter, A. Burkovski, and M. Niederweis.** 2007. A genomic view of sugar transport in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. J Bacteriol **189:**5903-5915.

71. **Truscott, K. N., A. Bezawork-Geleta, and D. A. Dougan.** 2011. Unfolded protein responses in bacteria and mitochondria: a central role for the ClpXP machine. IUBMB Life **63:**955-963.

72. **Tschumi, A., T. Grau, D. Albrecht, M. Rezwan, H. Antelmann, and P. Sander.** 2012. Functional analyses of mycobacterial lipoprotein diacylglyceryl transferase and comparative secretome analysis of a mycobacterial *lgt* mutant. J Bacteriol **194:**3938-3949.

73. **Tullman-Ercek, D., M. P. DeLisa, Y. Kawarasaki, P. Iranpour, B. Ribnicky, T. Palmer, and G. Georgiou.** 2007. Export pathway selectivity of *Escherichia coli* twin arginine translocation signal peptides. J Biol Chem **282:**8309-8316.

74. **van der Ploeg, R., U. Mader, G. Homuth, M. Schaffer, E. L. Denham, C. G. Monteferrante, M. Miethke, M. A. Marahiel, C. R. Harwood, T. Winter, M. Hecker, H. Antelmann, and J. M. van Dijl.** 2011. Environmental salinity determines the specificity and need for Tat-dependent secretion of the YwbN protein in *Bacillus subtilis*. PLoS One **6:**e18140.

75. **van der Ploeg, R., C. G. Monteferrante, S. Piersma, J. P. Barnett, T. R. Kouwen, C. Robinson, and J. M. van Dijl.** 2012. High-salinity growth conditions promote Tat-independent secretion of Tat substrates in *Bacillus subtilis*. Appl Environ Microbiol **78:**7733-7744.

76. **White, G. F., S. M. Schermann, J. Bradley, A. Roberts, N. P. Greene, B. C. Berks, and A. J. Thomson.** 2010. Subunit organization in the TatA complex of the twin arginine protein translocase: a site-directed EPR spin labeling study. J Biol Chem **285:**2294-2301.

77. **WorldHealthOrganization.** 2012. WHO Information tuberculosis fact sheet.

78. **Yahr, T. L., and W. T. Wickner.** 2001. Functional reconstitution of bacterial Tat translocation in vitro. EMBO J **20:**2472-2479.

Appendix

Label-free quantitation of *M. tuberculosis* cell wall proteins using a spectral counting comparison between wild type (H37Rv) and a Δ*secA2* mutant

Locus	Gene Name	Description/Function	$\log_2$ Δ*A2*/WT	*p*	Avg SpC	Sec[a]	Tat[b]	TM[c]	Lipo[d]
Rv0001	dnaA	chromosomal replication initiation protein	-1.482	0.002	47.4	-	-	-	-
Rv0003	recF	recombination protein F	-0.4678	0.278	11.6	-	-	-	-
Rv0005	gyrB	DNA gyrase subunit B	0.633	0.005	15.5	-	-	-	-
Rv0006	gyrA	DNA gyrase subunit A	0.5363	0.131	48.2	-	-	-	-
Rv0007		hypothetical protein	-0.5502	0.132	17.7	-	-	X	-
Rv0008c		hypothetical protein	0.1811	0.573	6.6	-	-	X	-
Rv0009	ppiA	iron-regulated peptidyl-prolyl cis-trans isomerase	1.107	0.153	18.8	-	-	-	-
Rv0010c		hypothetical protein	-0.667	0.149	3.5	-	-	X	-
Rv0011c		putative septation inhibitor protein	0.5996	0.086	11.7	-	-	X	-
Rv0012		hypothetical protein	0.723	0.034	7.6	-	-	X	-
Rv0013	trpG	para-aminobenzoate synthase component II	1.5612	0.159	4.6	-	-	-	-
Rv0014c	pknB	transmembrane serine/threonine-protein kinase B	-0.7508	0.042	55.7	-	-	X	-
Rv0015c	pknA	transmembrane serine/threonine-protein kinase A	-0.8217	0.126	23.3	-	-	X	-
Rv0016c	pbpA	penicillin-binding protein	-0.9748	0.005	31.5	X	-	X	-
Rv0017c	rodA	cell division protein	-0.1698	0.241	12.5	-	-	X	-
Rv0018c	ppp	serine/threonine phosphatase	-0.3504	0.411	31.2	-	-	X	-
Rv0019c		hypothetical protein	-0.7787	0.057	28.0	-	-	X	-
Rv0020c	TB39.8	hypothetical protein	0.0978	0.743	50.7	-	-	-	-
Rv0021c		hypothetical protein	-0.1757	0.607	4.9	-	-	-	-
Rv0025		hypothetical protein	-0.1354	0.66	2.0	-	-	-	-
Rv0029		hypothetical protein	-0.0684	0.831	9.2	-	-	-	-
Rv0030		hypothetical protein	0.8605	0.076	1.2	-	-	-	-
Rv0034		hypothetical protein	1.8005	0.002	1.7	-	-	-	-
Rv0037c		integral membrane protein	1.1548	0.141	0.9	-	-	X	-
Rv0038		hypothetical protein	-0.0178	0.923	7.0	-	-	-	-
Rv0042c		MarR family transcriptional regulator	-0.2968	0.625	18.5	-	-	-	-
Rv0043c		GntR family transcriptional regulator	1.4092	0.167	3.3	-	-	-	-
Rv0045c		hydrolase	-1.9126	0.061	5.5	-	-	-	-
Rv0046c	ino1	myo-inositol-1-phosphate synthase	-1.884	0.124	4.5	-	-	-	-
Rv0047c		hypothetical protein	-0.654	0.529	6.6	-	-	-	-
Rv0048c		hypothetical protein	0.2827	0.403	21.3	-	-	X	-
Rv0049		hypothetical protein	0.4228	0.42	5.0	-	-	-	-
Rv0050	ponA1	bifunctional penicillin-binding protein 1A/1B	-0.2526	0.456	51.9	-	-	-	-
Rv0051		transmembrane protein	-0.8364	0.017	6.7	-	-	X	-
Rv0052		hypothetical protein	-0.4078	0.628	6.8	-	-	-	-
Rv0053	rpsF	30S ribosomal protein S6	0.186	0.798	9.9	-	-	-	-
Rv0054	ssb	single-stranded DNA-binding protein	1.3518	0.113	4.7	-	-	-	-
Rv0055	rpsR	30S ribosomal protein S18	0.0227	0.898	18.7	-	-	-	-
Rv0056	rplI	50S ribosomal protein L9	0.487	0.33	7.1	-	-	-	-
Rv0060		hypothetical protein	1.1137	0.019	1.1	-	-	-	-
Rv0062	celA1	endo-1,4-beta-glucanase	0.3526	0.09	3.7	-	-	X	-
Rv0066c	icd2	isocitrate dehydrogenase	0.8478	0.133	9.0	-	-	-	-
Rv0068		short chain dehydrogenase	-0.3862	0.179	23.4	-	-	-	-
Rv0070c	glyA	serine hydroxymethyltransferase	-0.848	0.311	15.3	-	-	-	-
Rv0072		glutamine-transport transmembrane protein	-0.4502	0.083	57.8	-	-	X	-
Rv0073		glutamine-transport ATP-binding protein	0.8449	0.086	12.1	-	-	-	-
Rv0077c		oxidoreductase	0.5782	0.092	28.3	-	-	-	-
Rv0078		transcriptional regulator	1.1452	0.017	8.6	-	-	-	-
Rv0079		hypothetical protein	2.113	0.022	12.3	-	-	-	-
Rv0080		hypothetical protein	1.4222	0.002	7.4	-	-	-	-

Rv0088		hypothetical protein	0.0004	0.999	86.5	-	-	-	-
Rv0090		hypothetical protein	-0.3842	0.1	8.6	-	-	X	-
Rv0091	mtn	5'-methylthioadenosine nucleosidase	-0.1379	0.414	1.6	-	-	-	-
Rv0092	ctpA	cation transporter P-type ATPase A	0.0846	0.48	13.6	-	-	X	-
Rv0093c		hypothetical protein	0.1136	0.541	1.4	-	-	X	-
Rv0102		integral membrane protein	1.0688	0.004	1.3	-	-	X	-
Rv0103c	ctpB	cation-transporter P-type ATPase B	0.5272	0.605	2.2	-	-	X	-
Rv0107c	ctpI	cation-transporter ATPase I	0.7322	0.083	25.6	-	-	-	-
Rv0108c		hypothetical protein	0.9965	0.249	3.7	-	-	-	-
Rv0110		integral membrane protein	1.1137	0.019	1.1	-	-	X	-
Rv0112	gca	GDP-mannose 4,6-dehydratase	-0.02	0.943	5.4	-	-	-	-
Rv0113	gmhA	phosphoheptose isomerase	0.912	0.137	3.3	-	-	-	-
Rv0117	oxyS	oxidative stress response regulatory protein	-0.4362	0.129	19.0	-	-	-	-
Rv0118c	oxcA	putative oxalyl-CoA decarboxylase	-0.3648	0.432	15.6	-	-	-	-
Rv0119	fadD7	acyl-CoA synthetase	0.198	0.79	1.8	-	-	-	-
Rv0120c	fusA2	elongation factor G	0.1534	0.355	25.4	-	-	-	-
Rv0123		hypothetical protein	0.512	0.471	1.3	-	-	-	-
Rv0125	pepA	serine protease	0.686	0.332	7.6	X	-	X	-
Rv0127		hypothetical protein	-0.207	0.488	11.5	-	-	-	-
Rv0129c	fbpC	secreted antigen 85-C	0.9697	0.135	26.2	-	X	X	-
Rv0131c	fadE1	acyl-CoA dehydrogenase	-0.3572	0.627	10.9	-	-	-	-
Rv0133		acetyltransferase	0.5786	0.007	11.1	-	-	-	-
Rv0134	ephF	epoxide hydrolase	0.302	0.252	19.9	-	-	-	-
Rv0136	cyp138	cytochrome P450 138	-0.2028	0.719	21.3	-	-	-	-
Rv0139		oxidoreductase	-0.1426	0.685	11.8	-	-	-	-
Rv0144		TetR family transcriptional regulator	1.0018	0.069	9.2	-	-	X	-
Rv0145		hypothetical protein	1.0252	0.104	2.2	-	-	-	-
Rv0146		hypothetical protein	1.6103	0.066	3.6	-	-	-	-
Rv0147		aldehyde dehydrogenase	-0.9752	0.067	99.6	-	-	-	-
Rv0148		short-chain type dehydrogenase/reductase	1.7242	0.067	7.9	-	-	-	-
Rv0149		quinone oxidoreductase	-1.0261	0.622	1.7	-	-	-	-
Rv0153c	ptbB	phosphotyrosine protein phosphatase) (PTPase)	0.336	0.672	3.5	-	-	-	-
Rv0155	pntAa	NAD(P) transhydrogenase subunit alpha	0.5328	0.085	12.7	-	-	-	-
Rv0156	pntAb	NAD(P) transhydrogenase subunit alpha	-2.0296	0.017	8.8	-	-	X	-
Rv0157	pntB	NAD(P) transhydrogenase subunit beta	-0.3026	0.006	119.0	-	-	X	-
Rv0164	TB18.5	hypothetical protein	0.881	0.232	3.0	-	-	-	-
Rv0165c		GntR family transcriptional regulator	-0.0751	0.86	18.1	-	-	-	-
Rv0166	fadD5	acyl-CoA synthetase	0.802	0.01	29.0	-	-	-	-
Rv0167	yrbE1A	MCE1-associated	-0.59	0.179	14.9	-	-	X	-
Rv0168	yrbE1B	MCE1-associated	-1.1818	0.018	4.4	-	-	X	-
Rv0169	mce1A	MCE1-associated	-1.1012	0.018	80.5	-	-	X	-
Rv0170	mce1B	MCE1-associated	-0.7972	0.012	125.1	-	-	X	-
Rv0171	mce1C	MCE1-associated	-0.9534	0.003	144.6	-	-	X	-
Rv0172	mce1D	MCE1-associated	-0.932	0.001	119.8	-	-	X	-
Rv0173	lprK	MCE1-associated lipoprotein	-0.8102	0.006	113.7	-	-	X	X
Rv0174	mce1F	MCE1-associated	-1.1782	0.009	155.4	-	-	X	-
Rv0175		MCE1-associated	0.941	0.028	19.6	-	-	X	-
Rv0176		MCE1-associated	-0.8535	0.005	30.9	-	-	X	-
Rv0177		MCE1-associated	-0.1526	0.608	26.5	-	-	X	-
Rv0178		MCE1-associated	0.0686	0.87	56.4	-	-	X	-
Rv0179c	lprO	lipoprotein	-0.1354	0.718	2.0	X	-	X	X
Rv0180c		transmembrane protein	-0.8432	0.066	24.3	-	-	X	-
Rv0182c	sigG	RNA polymerase factor sigma-70	-0.5794	0.255	8.6	-	-	-	-
Rv0183		lysophospholipase	0.1954	0.201	40.2	-	-	-	-
Rv0184		hypothetical protein	-0.5503	0.092	3.6	-	-	-	-
Rv0185		hypothetical protein	0.2548	0.281	15.8	-	-	-	-

Rv0187		O-methyltransferase	1.3045	0.014	1.3	-	-	-	-
Rv0189c	ilvD	dihydroxy-acid dehydratase	1.7486	0.023	7.8	-	-	-	-
Rv0190		hypothetical protein	-0.3648	0.575	2.6	-	-	-	-
Rv0194		ATP-binding protein ABC transporter	-0.6233	0.057	2.8	-	-	X	-
Rv0196		transcriptional regulator	0.356	0.364	1.0	-	-	-	-
Rv0199		hypothetical protein	0.0042	0.976	14.3	-	-	X	-
Rv0200		transmembrane protein	-0.4634	0.156	12.7	-	-	X	-
Rv0201c		hypothetical protein	-0.5597	0.405	5.5	-	-	-	-
Rv0202c	mmpL11	transmembrane transport protein MmpL11	-0.6872	0.033	21.9	X	-	X	-
Rv0204c		transmembrane protein	-0.3647	0.178	3.8	-	-	X	-
Rv0206c	mmpL3	transmembrane transport protein MmpL3	-0.152	0.198	153.6	-	-	X	-
Rv0207c		hypothetical protein	1.5364	0.107	1.3	-	-	-	-
Rv0215c	fadE3	acyl-CoA dehydrogenase	1.393	0.234	2.6	-	-	-	-
Rv0217c	lipW	esterase	0.0994	0.841	11.2	-	-	-	-
Rv0220	lipC	esterase	-0.1892	0.412	15.1	-	-	-	-
Rv0221		hypothetical protein	-0.0705	0.806	22.9	-	-	-	-
Rv0224c		methyltransferase (methylase)	-0.6335	0.248	1.3	-	-	-	-
Rv0227c		hypothetical protein	-1.2276	0.003	66.9	-	-	X	-
Rv0229c		hypothetical protein	-0.4214	0.464	2.1	-	-	X	-
Rv0233	nrdB	ribonucleotide-diphosphate reductase subunit	-0.6164	0.448	4.4	-	-	-	-
Rv0236c		transmembrane protein	-0.1788	0.849	1.2	X	-	X	-
Rv0237	lpqI	lipoprotein	-0.3204	0.352	22.5	X	-	-	X
Rv0241c		hypothetical protein	0.652	0.102	22.1	-	-	-	-
Rv0242c	fabG	3-ketoacyl-(acyl-carrier-protein) reductase	0.3664	0.03	44.7	-	-	-	-
Rv0243	fadA2	acetyl-CoA acetyltransferase	0.265	0.288	27.2	-	-	-	-
Rv0245		oxidoreductase	0.4127	0.364	2.1	-	-	-	-
Rv0247c		fumarate reductase iron-sulfur subunit	0.144	0.682	60.1	-	-	-	-
Rv0248c	sdhA	succinate dehydrogenase flavoprotein subunit	-0.8976	0.002	152.4	-	-	-	-
Rv0249c		succinate dehydrogenase membrane anchor	0.0596	0.689	0.9	-	-	X	-
Rv0255c	cobQ1	cobyric acid synthase	-0.4952	0.298	12.1	-	-	-	-
Rv0262c	aac	aminoglycoside 2'-N-acetyltransferase (AAC(2')-IC)	-0.308	0.413	5.0	-	-	-	-
Rv0265c		periplasmic IRON-transport lipoprotein	-1.0434	0.014	13.0	-	X	-	X
Rv0270	fadD2	acyl-CoA synthetase	-0.0387	0.818	19.6	-	-	-	-
Rv0271c	fadE6	acyl-CoA dehydrogenase	-0.2178	0.68	11.8	-	-	-	-
Rv0273c		transcriptional regulator	-0.4214	0.261	3.2	-	-	-	-
Rv0274		hypothetical protein	0.2547	0.323	2.5	-	-	-	-
Rv0275c		TetR family transcriptional regulator	0.6602	0.248	6.0	-	-	-	-
Rv0276		hypothetical protein	0.8892	0.031	13.3	-	-	-	-
Rv0281		hypothetical protein	0.9686	0.066	2.1	-	-	-	-
Rv0282		hypothetical protein	1.2378	0.022	4.6	-	-	-	-
Rv0283		hypothetical protein	-0.5958	0.006	42.5	-	-	X	-
Rv0284		hypothetical protein	-0.2322	0.321	317.0	-	-	X	-
Rv0289		hypothetical protein	1.4734	0.102	6.7	-	-	-	-
Rv0290		transmembrane protein	-0.0584	0.763	99.4	-	-	X	-
Rv0291	mycP3	membrane-anchored mycosin	-0.9572	0.035	7.8	X	-	X	-
Rv0292		transmembrane protein	-0.4428	0.037	68.5	-	-	X	-
Rv0293c		hypothetical protein	-0.252	0.883	2.2	-	-	-	-
Rv0299		hypothetical protein	-0.0903	0.524	1.4	-	-	-	-
Rv0303		dehydrogenase/reductase	0.7608	0.22	11.7	-	-	-	-
Rv0307c		hypothetical protein	0.2601	0.467	15.3	-	-	-	-
Rv0308		integral membrane protein	-0.2633	0.207	19.4	-	-	X	-
Rv0312		hypothetical protein	-0.0314	0.929	5.7	-	-	X	-
Rv0313		hypothetical protein	-0.2362	0.236	6.6	-	-	-	-
Rv0314c		hypothetical protein	-0.8364	0.358	8.1	-	-	X	-
Rv0315		beta-1,3-glucanase precursor	0.0665	0.773	12.4	X	X	-	-
Rv0317c	glpQ2	glycerophosphoryl diester phosphodiesterase	1.2016	0.011	29.9	-	-	-	-
Rv0328		TetR/AcrR family transcriptional regulator	1.1355	0.028	3.2	-	-	-	-
Rv0329c		hypothetical protein	1.655	0.056	3.4	-	-	-	-

Rv0333		hypothetical protein	0.5476	0.159	7.1	-	-	-	-
Rv0337c	aspC	aminotransferase	0.8553	0.152	3.1	-	-	-	-
Rv0338c		iron-sulfur-binding reductase	-0.318	0.078	99.7	-	-	X	-
Rv0339c		transcriptional regulator	-2.1756	0.024	8.1	-	-	-	-
Rv0343	iniC	isoniazid inductible gene protein	1.9708	0.063	8.9	-	-	-	-
Rv0344c	lpqJ	lipoprotein	-0.58		0.8	X	-	-	X
Rv0346c	ansP2	L-asparagine ABC transporter permease	0.0248	0.891	25.2	-	-	X	-
Rv0350	dnaK	molecular chaperone DnaK	0.2112	0.069	102.1	-	-	-	-
Rv0352	dnaJ1	chaperone protein	-0.201	0.613	40.0	-	-	-	-
Rv0353	hspR	HEAT shock protein transcriptional repressor	-0.9585	0.319	8.9	-	-	-	-
Rv0356c		hypothetical protein	0.4981	0.117	2.1	-	-	-	-
Rv0358		hypothetical protein	-0.0789	0.919	1.9	-	-	-	-
Rv0360c		hypothetical protein	0.862	0.077	0.7	-	-	-	-
Rv0361		hypothetical protein	-0.1054	0.748	21.8	-	-	X	-
Rv0364		transmembrane protein	-0.0004	0.999	8.3	-	-	X	-
Rv0365c		hypothetical protein	-0.3126	0.691	7.1	-	-	-	-
Rv0369c		membrane oxidoreductase	0.5898		0.8	-	-	X	-
Rv0377		LysR family transcriptional regulator	-1.3956	0.015	9.7	-	-	-	-
Rv0379	secE2	protein transport protein	-1.2116	0.011	69.5	-	-	-	-
Rv0381c		hypothetical protein	-0.952	0.128	2.3	-	-	-	X
Rv0382c	pyrE	orotate phosphoribosyltransferase	1.0048	-	0.5	-	-	-	-
Rv0383c		hypothetical protein	-0.3904	0.225	48.7	-	-	X	-
Rv0384c	clpB	endopeptidase ATP binding protein	1.1132	0.279	8.5	-	-	-	-
Rv0385		hypothetical protein	-0.1697	0.705	6.7	-	-	-	-
Rv0386		LuxR/UHPA family transcriptional regulator	-1.2058	0.126	8.3	-	-	-	-
Rv0392c	ndhA	membrane NADH dehydrogenase	-0.6528	0.024	33.2	-	-	X	-
Rv0394c		hypothetical protein	0.6548	0.476	1.4	-	-	X	-
Rv0399c	lpqK	lipoprotein	-2.4903	0.112	3.2	X	-	-	X
Rv0400c	fadE7	acyl-CoA dehydrogenase	1.0048	-	0.5	-	-	-	-
Rv0402c	mmpL1	transmembrane transport protein MmpL1	0.2858	0.827	0.9	-	-	X	-
Rv0406c		beta lactamase like protein	0.3387	0.366	5.8	-	-	-	-
Rv0408	pta	phosphate acetyltransferase	0.918	0.188	4.5	-	-	-	-
Rv0409	ackA	acetate kinase	1.0596	0.004	1.3	-	-	-	-
Rv0410c	pknG	serine/threonine-protein kinase	-0.1895	0.103	39.4	-	-	-	-
Rv0411c	glnH	glutamine-binding lipoprotein	-1.6566	0.05	6.5	X	-	-	X
Rv0412c		hypothetical protein	-0.257	0.156	26.9	-	-	X	-
Rv0413	mutT3	7,8-dihydro-8-oxoguanine-triphosphatase	0.723	0.255	9.0	-	-	-	-
Rv0415	thiO	thiamine biosynthesis oxidoreductase ThiO	0.1733	0.487	7.6	-	-	-	-
Rv0417	thiG	thiazole synthase	-0.1379	0.414	0.5	-	-	-	-
Rv0418	lpqL	lipoprotein aminopeptidase	-0.0234	0.9	33.9	X	-	X	X
Rv0419	lpqM	lipoprotein peptidase	-2.0968	0.019	11.5	X	-	X	X
Rv0421c		hypothetical protein	1.1548	0.141	0.9	-	-	-	-
Rv0423c	thiC	thiamine biosynthesis protein ThiC	1.8295	0.097	2.3	-	-	-	-
Rv0424c		hypothetical protein	-0.6032	0.17	1.7	-	-	-	-
Rv0425c	ctpH	metal cation transporting P-type ATPase	2.3846	0.005	9.3	X	-	X	-
Rv0426c		hypothetical protein	0.3278	0.024	36.3	X	-	X	-
Rv0431		putative tuberculin related peptide	0.0584	0.869	26.9	-	-	X	-
Rv0432	sodC	periplasmic superoxide dismutase [Cu-Zn]	-0.2983	0.774	26.3	X	-	-	X
Rv0433		carboxylate-amine ligase	0.5188	0.559	2.9	-	-	-	-
Rv0434		hypothetical protein	1.017	0.02	5.4	-	-	-	-
Rv0436c	pssA	CDP-diacylglycerol--serine O-phosphatidyltransferase	1.1698	0.33	4.6	-	-	X	-
Rv0437c	psd	phosphatidylserine decarboxylase	-0.09	0.941	9.7	-	-	-	-
Rv0438c	moeA2	molybdopterin biosynthesis protein	-0.0095	0.942	3.2	-	-	-	-
Rv0439c		short chain dehydrogenase	0.105	0.768	24.9	-	-	-	-
Rv0440	groEL2	chaperonin GroEL	0.6047	0.15	228.2	-	-	-	-
Rv0441c		hypothetical protein	-0.0178	0.965	7.8	-	-	-	-
Rv0444c		hypothetical protein	0.7058	0.102	36.0	-	-	-	-
Rv0445c	sigK	RNA polymerase sigma factor	0.6898	0.031	23.1	-	-	-	-
Rv0450c	mmpL4	transmembrane transport protein MmpL4	0.1022	0.518	49.6	-	-	X	-
Rv0452		transcriptional regulator	0.4493	0.134	2.3	-	-	-	-

Rv0455c		hypothetical protein	0.2278	0.574	9.6	X	-	X	-
Rv0461		hypothetical protein	0.0428	0.939	7.6	-	-	X	-
Rv0462	lpd	dihydrolipoamide dehydrogenase	-0.4934	0.178	21.9	-	-	-	-
Rv0464c		hypothetical protein	0.139	0.828	12.6	-	-	-	-
Rv0465c		transcriptional regulator	0.4168	0.365	5.2	-	-	-	-
Rv0467	icl	isocitrate lyase	0.0652	0.933	5.2	-	-	-	-
Rv0468	fadB2	3-hydroxybutyryl-CoA dehydrogenase	0.3548	0.05	9.4	-	-	-	-
Rv0469	umaA	mycolic acid synthase	0.5474	0.287	11.5	-	-	-	-
Rv0470c	pcaA	mycolic acid synthase	0.9359	0.208	3.5	-	-	-	-
Rv0472c		TetR family transcriptional regulator	0.1568	0.676	2.5	-	-	-	-
Rv0473		transmembrane protein	0.255	0.047	8.9	-	-	X	-
Rv0474		transcriptional regulator	0.3836	0.051	11.1	-	-	-	-
Rv0475	hbhA	iron-regulated heparin binding hemagglutinin	1.1105	0.101	6.1	-	-	-	-
Rv0476		transmembrane protein	-0.977	0.277	6.3	X	-	X	-
Rv0479c		hypothetical protein	-0.5838	0.092	28.8	-	-	X	-
Rv0484c		short-chain type oxidoreductase	1.9622	0.084	2.9	-	-	-	-
Rv0487		hypothetical protein	0.497	0.048	2.6	-	-	-	-
Rv0490	senX3	putative two component sensor histidine kinase	-0.5793	0.099	12.9	-	-	X	-
Rv0492c		oxidoreductase GMC-type	0.1548	0.651	9.5	-	-	-	-
Rv0493c		hypothetical protein	-0.3436	0.243	7.6	-	-	-	-
Rv0495c		hypothetical protein	-0.4955	0.493	4.4	-	-	-	-
Rv0496		hypothetical protein	-0.18	0.297	43.3	-	-	-	-
Rv0497		transmembrane protein	-0.3686	0.162	27.7	-	-	X	-
Rv0498		hypothetical protein	-0.7245	0.543	1.8	-	-	-	-
Rv0500	proC	pyrroline-5-carboxylate reductase	-0.8091	0.186	3.9	-	-	-	-
Rv0500A		hypothetical protein	-1.0903	0.012	1.2	-	-	-	-
Rv0501	galE2	UDP-glucose 4-epimerase	-0.4332	0.208	29.9	-	-	-	-
Rv0502		hypothetical protein	-0.4432	0.07	13.9	-	-	-	-
Rv0503c	cmaA2	cyclopropane-fatty-acyl-phospholipid synthase 2 (	0.8614	0.03	8.1	-	-	-	-
Rv0505c	serB1	phosphoserine phosphatase	-0.8673	0.021	16.5	-	-	X	-
Rv0506	mmpS2	membrane protein	0.107	0.78	2.6	-	-	X	-
Rv0507	mmpL2	transmembrane transport protein MmpL2	-0.1704		0.7	-	-	X	-
Rv0509	hemA	glutamyl-tRNA reductase	0.1848	0.53	14.5	-	-	-	-
Rv0511	hemD	uroporphyrin-III C-methyltransferase	1.265	0.25	7.7	-	-	-	-
Rv0513		transmembrane protein	-0.1076	0.756	18.7	-	-	X	-
Rv0516c		hypothetical protein	1.9686	0.037	6.3	-	-	-	-
Rv0517		membrane acyltransferase	0.1766	0.884	3.1	-	-	X	-
Rv0523c		hypothetical protein	-0.291	0.338	29.0	-	-	-	-
Rv0524	hemL	glutamate-1-semialdehyde aminotransferase	-0.6829	0.191	17.9	-	-	-	-
Rv0525		hypothetical protein	0.4678	0.239	16.7	-	-	-	-
Rv0526		thioredoxin protein	-0.643	0.126	26.2	X	X	-	X
Rv0528		transmembrane protein	-1.0916	0.103	12.1	-	-	X	-
Rv0529	ccsA	cytochrome C-type biogenesis protein	-0.5908	0.373	5.2	-	-	X	-
Rv0530		hypothetical protein	0.0868	0.555	46.1	-	-	-	-
Rv0531		hypothetical protein	0.2513	0.711	5.1	-	-	X	-
Rv0533c	fabH	3-oxoacyl-(acyl carrier protein) synthase III	-1.9896	0.048	6.3	-	-	-	-
Rv0534c	menA	1,4-dihydroxy-2-naphthoate octaprenyltransferase	-0.698	0.085	2.8	-	-	X	-
Rv0537c		integral membrane protein	-0.6042	0.159	15.3	-	-	X	-
Rv0538		hypothetical protein	-1.5286	0.093	9.4	-	-	X	-
Rv0539		dolichyl-phosphate sugar synthase	-0.1379	0.414	1.6	-	-	-	-
Rv0540		hypothetical protein	2.352	0	2.4	-	-	-	-
Rv0543c		hypothetical protein	0.9097	0.016	1.7	-	-	-	-
Rv0544c		transmembrane protein	-0.3647	0.178	3.8	-	-	X	-
Rv0545c	pitA	inorganic phosphate transporter	-0.283	0.344	3.0	X	-	X	-
Rv0547c		short chain dehydrogenase	-0.2094	0.355	26.6	-	-	-	-
Rv0548c	menB	naphthoate synthase	-1.034	0.018	3.6	-	-	-	-
Rv0556		transmembrane protein	0.4589	0.073	12.7	-	-	X	-
Rv0558	ubiE	ubiquinone/menaquinone biosynthesis methyltransferase	0.5589	0.107	15.7	-	-	-	-
Rv0559c		hypothetical protein	-0.3185	0.56	2.7	X	-	-	-

Rv0561c		oxidoreductase	0.302	0.572	3.3	-	-	-	-
Rv0563	htpX	heat shock protein	-0.2586	0.704	10.0	-	-	X	-
Rv0565c		monooxygenase	-0.2526	0.517	7.4	-	-	X	-
Rv0566c		putative nucleotide-binding protein	-0.1698	0.746	4.3	-	-	-	-
Rv0567		methyltransferase/methylase	0.0328	0.913	21.2	-	-	-	-
Rv0568	cyp135B1	cytochrome P450 135B1	0.3232	0.382	17.1	-	-	-	-
Rv0570	nrdZ	ribonucleoside-diphosphate reductase	3.1218	0.001	6.6	-	-	-	-
Rv0580c		hypothetical protein	-0.6192	0.134	102.5	-	-	-	-
Rv0583c	lpqN	lipoprotein	-0.4422	0.071	27.5	X	-	-	X
Rv0598c		hypothetical protein	0.3482	0.376	2.0	X	-	-	-
Rv0604	lpqO	lipoprotein	-0.9362	0.005	14.0	X	-	-	X
Rv0605		resolvase	0.1136	0.541	2.1	-	-	-	-
Rv0606		hypothetical protein	-1.1984	-	0.6	-	-	-	-
Rv0610c		hypothetical protein	-0.7948	0.456	2.8	-	-	-	-
Rv0613c		hypothetical protein	-0.0548	0.779	49.3	-	-	-	-
Rv0617		hypothetical protein	0.0048		0.3	-	-	-	-
Rv0618	galTa	galactose-1-phosphate uridylyltransferase galTa	0.1138	0.919	1.4	-	-	-	-
Rv0625c		transmembrane protein	-1.2215	0.103	6.8	-	-	X	-
Rv0629c	recD	exonuclease V alpha chain	-0.4213	0.478	11.6	-	-	-	-
Rv0630c	recB	exonuclease V beta chain	0.8558	0.218	21.9	-	-	-	-
Rv0631c	recC	exonuclease V gamma chain	1.5026	0.102	11.1	-	-	-	-
Rv0632c	echA3	enoyl-CoA hydratase	0.7328	0.065	38.1	-	-	-	-
Rv0633c		hypothetical protein	-0.6976	0.046	28.1	-	-	X	-
Rv0634c		glyoxalase II	0.0532	0.797	1.5	-	-	-	-
Rv0635	HadA	(3R)-hydroxyacyl-ACP dehydratase subunit	-1.098	0.116	12.6	-	-	-	-
Rv0636	HadB	(3R)-hydroxyacyl-ACP dehydratase subunit	0.3427	0.373	9.5	-	-	-	-
Rv0637	HadC	(3R)-hydroxyacyl-ACP dehydratase subunit	-0.4446	0.482	19.5	-	-	-	-
Rv0638	secE	preprotein translocase subunit SecE	-0.2768	0.234	26.8	-	-	X	-
Rv0639	nusG	transcription antitermination protein NusG	0.2742	0.333	4.2	-	-	-	-
Rv0640	rplK	50S ribosomal protein L11	-0.1131	0.731	6.7	-	-	-	-
Rv0641	rplA	50S ribosomal protein L1	0.048	0.962	1.4	-	-	-	-
Rv0642c	mmaA4	methoxy mycolic acid synthase	0.6825	0.199	11.2	-	-	-	-
Rv0643c	mmaA3	methoxy mycolic acid synthase	-0.31	0.751	3.6	-	-	-	-
Rv0644c	mmaA2	methoxy mycolic acid synthase	0.3143	0.365	10.7	-	-	-	-
Rv0646c	lipG	lipase/esterase	-0.3372	0.265	14.0	-	-	-	-
Rv0647c		hypothetical protein	0.2773	0.263	20.6	-	-	-	-
Rv0651	rplJ	50S ribosomal protein L10	0.2278	0.468	14.1	-	-	-	-
Rv0652	rplL	50S ribosomal protein L7/L12	0.0328	0.934	11.2	-	-	-	-
Rv0654		dioxygenase	0.2584	0.845	1.2	-	-	-	-
Rv0655	mkl	ribonucleotide ABC transporter ATP-binding	-0.8418	0.02	108.0	-	-	-	-
Rv0663	atsD	arylsulfatase	-0.1522	0.858	3.1	-	-	-	-
Rv0667	rpoB	DNA-directed RNA polymerase subunit beta	-0.0008	0.994	59.7	-	-	-	-
Rv0668	rpoC	DNA-directed RNA polymerase subunit beta'	-0.1108	0.508	316.1	-	-	-	-
Rv0669c		hydrolase	-0.5417	0.146	15.9	-	-	-	-
Rv0671	lpqP	lipoprotein	1.4621	0.182	2.9	X	-	X	X
Rv0676c	mmpL5	transmembrane transport protein MmpL5	0.3824	0.021	58.4	-	-	X	-
Rv0677c	mmpS5	hypothetical protein	0.3756	0.201	22.1	-	-	X	-
Rv0678		hypothetical protein	-0.1024	0.77	17.9	-	-	-	-
Rv0679c		putative threonine rich protein	-0.667	0.224	4.1	X	-	-	X
Rv0681		TetR family transcriptional regulator	0.8045	0.339	1.1	-	-	-	-
Rv0682	rpsL	30S ribosomal protein S12	0.3562	0.309	37.6	-	-	-	-
Rv0683	rpsG	30S ribosomal protein S7	0.0546	0.726	26.2	-	-	-	-
Rv0684	fusA1	elongation factor G	0.279	0.301	22.2	-	-	-	-
Rv0685	tuf	elongation factor Tu	-0.2138	0.386	224.8	-	-	-	-
Rv0686		hypothetical protein	-0.0946	0.526	99.6	-	-	X	-
Rv0687	fabG	3-ketoacyl-(acyl-carrier-protein) reductase	-0.1395	0.911	1.6	-	-	-	-
Rv0688		putative ferredoxin reductase	1.1712	0.433	1.5	-	-	-	-

Rv0690c		hypothetical protein	0.2884	0.407	6.9	-	-	-	-
Rv0691c		transcriptional regulator	1.1137	0.019	1.1	-	-	-	-
Rv0694	lldD1	L-lactate dehydrogenase (cytochrome)	-0.5048	0.336	15.2	-	-	-	-
Rv0695		hypothetical protein	1.8843	0.006	2.3	-	-	-	-
Rv0696		membrane sugar transferase	-0.2662	0.487	20.7	-	-	X	-
Rv0700	rpsJ	30S ribosomal protein S10	0.7416	0.048	58.2	-	-	-	-
Rv0701	rplC	50S ribosomal protein L3	1.07	0.042	10.7	-	-	-	-
Rv0702	rplD	50S ribosomal protein L4	0.7682	0.07	6.0	-	-	-	-
Rv0703	rplW	50S ribosomal protein L23	0.5907	0.17	7.9	-	-	-	-
Rv0704	rplB	50S ribosomal protein L2	0.5464	0.025	52.9	-	-	-	-
Rv0705	rpsS	30S ribosomal protein S19	0.9102	0	13.3	-	-	-	-
Rv0706	rplV	50S ribosomal protein L22	0.5564	0.252	5.3	-	-	-	-
Rv0707	rpsC	30S ribosomal protein S3	-0.1758	0.206	110.8	-	-	-	-
Rv0708	rplP	50S ribosomal protein L16	1.1266	0.02	12.7	-	-	-	-
Rv0709	rpmC	50S ribosomal protein L29	-0.4214	0.532	2.1	-	-	-	-
Rv0710	rpsQ	30S ribosomal protein S17	-0.2414	0.411	11.6	X	-	-	-
Rv0713		transmembrane protein	-0.5918	0.015	29.4	-	-	X	-
Rv0714	rplN	50S ribosomal protein L14	0.991	0.028	9.7	-	-	-	-
Rv0715	rplX	50S ribosomal protein L24	1.4092	0.049	7.1	-	-	-	-
Rv0716	rplE	50S ribosomal protein L5	0.026	0.919	17.1	-	-	-	-
Rv0717	rpsN	30S ribosomal protein S14	0.2513	0.661	4.3	-	-	-	-
Rv0718	rpsH	30S ribosomal protein S8	0.5934	0.25	10.0	-	-	-	-
Rv0719	rplF	50S ribosomal protein L6	0.6411	0.142	30.8	-	-	-	-
Rv0720	rplR	50S ribosomal protein L18	0.0227	0.946	12.4	-	-	-	-
Rv0721	rpsE	30S ribosomal protein S5	-1.5274	0.214	33.3	-	-	-	-
Rv0722	rpmD	50S ribosomal protein L30	-0.088	0.862	4.1	-	-	-	-
Rv0723	rplO	50S ribosomal protein L15	-0.3006	0.409	5.5	-	-	-	-
Rv0724	sppA	protease IV	0.2174	0.664	26.2	-	-	-	-
Rv0728c	serA2	D-3-phosphoglycerate dehydrogenase	-1.997	0.002	5.2	-	-	-	-
Rv0729	xylB	D-xylulose kinase	0.302	0.317	2.4	-	-	-	-
Rv0730		hypothetical protein	0.7263	0.252	2.9	-	-	-	-
Rv0732	secY	preprotein translocase subunit SecY	-0.0571	0.76	40.9	-	-	X	-
Rv0733	adk	adenylate kinase	1.1163	0.239	2.0	-	-	-	-
Rv0734	mapA	methionine aminopeptidase	-0.4624	0.572	5.5	-	-	-	-
Rv0735	sigL	RNA polymerase sigma factor	-0.7722	0.289	6.3	-	-	-	-
Rv0737		transcriptional regulator	0.0738	0.861	6.4	-	-	-	-
Rv0744c		transcriptional regulator	0.0328	0.95	9.1	-	-	-	-
Rv0751c	mmsB	3-hydroxyisobutyrate dehydrogenase	0.1136	0.541	1.4	-	-	-	-
Rv0756c		hypothetical protein	0.0204	0.961	13.0	-	-	-	-
Rv0757	phoP	two component system response transcriptional positive regulator	1.0316	0.058	13.4	-	-	-	-
Rv0758	phoR	two component system response sensor kinase membrane associated	-0.3492	0.044	25.8	-	-	X	-
Rv0759c		hypothetical protein	-0.2574	0.394	13.3	-	-	-	-
Rv0760c		hypothetical protein	0.1636	0.594	3.7	-	-	-	-
Rv0761c	adhB	zinc-containing alcohol dehydrogenase	-0.4388	0.582	6.2	-	-	-	-
Rv0769		short chain dehydrogenase	0.438	0.741	1.8	-	-	-	-
Rv0771		4-carboxymuconolactone decarboxylase	-0.4214	0.453	4.1	-	-	-	-
Rv0774c		hypothetical protein	1.5764	0.112	1.4	X	X	X	-
Rv0775		hypothetical protein	-0.0745	0.929	1.3	-	-	-	-
Rv0777	purB	adenylosuccinate lyase	0.113	0.835	4.0	-	-	-	-
Rv0778	cyp126	cytochrome P450 126	-1.802	0.001	11.6	-	-	-	-
Rv0779c		transmembrane protein	0.302	0.122	12.0	-	-	X	-
Rv0781	ptrBa	oligopeptidase B	-0.3844	0.285	4.1	-	-	-	-
Rv0785		putative FAD-binding dehydrogenase	-0.2226	0.34	10.0	-	-	-	-
Rv0788	purQ	phosphoribosylformylglycinamidine synthase I	2.0994	0.054	5.3	-	-	-	-
Rv0790c		hypothetical protein	0.2544	0.503	1.0	-	-	-	-
Rv0792c		GntR family transcriptional regulator	0.0232	0.959	8.6	-	-	-	-
Rv0794c		oxidoreductase	1.5313	0.138	2.4	-	-	-	-
Rv0798c	cfp29	29 kDa antigen	0.6982	0.057	27.4	-	-	-	-
Rv0799c		hypothetical protein	-0.6773	0.015	3.8	-	-	-	-
Rv0801		hypothetical protein	0.7626	0.037	4.2	-	-	-	-

Rv0802c		hypothetical protein	1.862	0.036	1.0	-	-	-	-
Rv0808	purF	amidophosphoribosyltransferase	-0.9404	0.006	1.4	-	-	-	-
Rv0809	purM	phosphoribosylaminoimidazole synthetase	0.039	0.89	6.6	-	-	-	-
Rv0817c		hypothetical protein	0.836	0.187	10.6	X	-	X	-
Rv0818		transcriptional regulator	-0.1054	0.638	38.2	-	-	-	-
Rv0820	phoT	phosphate ABC transporter ATP-binding protein	0.1068	0.837	30.6	-	-	-	-
Rv0821c	phoY2	phosphate transport regulator	0.5935	0.182	6.4	-	-	-	-
Rv0822c		hypothetical protein	0.8646	0.085	2.8	-	-	-	-
Rv0824c	desA1	acyl-[acyl-carrier protein] desaturase	0.5886	0.215	10.1	-	-	-	-
Rv0825c		hypothetical protein	0.1782	0.132	28.8	-	-	-	-
Rv0831c		hypothetical protein	-0.1622	0.83	12.7	-	-	-	-
Rv0835	lpqQ	lipoprotein	-0.607	0.207	1.3	X	-	-	X
Rv0845		two component sensor kinase	-1.8484	0.004	6.2	-	-	X	-
Rv0846c		oxidase	-0.4214	0.014	8.1	-	X	-	X
Rv0851c		short chain dehydrogenase	-0.1698	0.803	2.2	-	-	-	-
Rv0852	fadD16	fatty-acid-CoA ligase	0.0562	0.937	3.3	-	-	-	-
Rv0854		hypothetical protein	-0.4955	0.054	5.5	-	-	-	-
Rv0855	far	fatty-acid-CoA racemase	-0.3345	0.632	2.0	-	-	-	-
Rv0859	fadA	acetyl-CoA acetyltransferase	0.0524	0.869	8.8	-	-	-	-
Rv0860	fadB	fatty oxidation protein	3.0943	0.004	17.6	-	-	-	-
Rv0861c	ercC3	DNA helicase	1.0775	0.026	10.9	-	-	-	-
Rv0862c		hypothetical protein	0.514	0.179	47.0	-	-	-	-
Rv0864	moaC	molybdenum cofactor biosynthesis protein C	-0.4123	0.731	1.9	-	-	-	-
Rv0866	moaE2	molybdenum cofactor biosynthesis protein E2	0.302	0.317	2.4	-	-	-	-
Rv0869c	moaA	molybdenum cofactor biosynthesis protein A	0.302	0.317	1.2	-	-	-	-
Rv0870c		hypothetical protein	0.4404	0.365	6.2	-	-	X	-
Rv0871	cspB	cold shock-like protein B	1.4533	0.202	1.3	-	-	-	-
Rv0873	fadE10	acyl-CoA dehydrogenase	-0.3248	0.101	57.4	-	-	-	-
Rv0874c		hypothetical protein	0.5364	0.287	1.5	-	-	-	-
Rv0875c		hypothetical protein	-0.8598	0.174	4.8	-	-	X	-
Rv0879c		transmembrane protein	-0.4214	0.091	2.5	-	-	X	-
Rv0880		MarR family transcriptional regulator	1.012	0.025	1.1	-	-	-	-
Rv0881		rRNA methyltransferase	-0.4478	0.264	1.2	-	-	-	-
Rv0885		hypothetical protein	-0.3648	0.538	9.3	-	-	-	-
Rv0886	fprB	NADPH:adrenodoxin oxidoreductase FprB	0.0738	0.74	7.5	-	-	-	-
Rv0888		hypothetical protein	-0.4018	0.324	9.9	X	-	X	-
Rv0890c		LuxR family transcriptional regulator	0.5126	0.176	15.1	-	X	-	-
Rv0892		monooxygenase	0.0048		0.7	X	-	X	-
Rv0896	gltA	type II citrate synthase	0.2454	0.719	7.8	-	-	-	-
Rv0899	ompA	outer membrane protein A	-0.2746	0.203	23.1	-	-	X	-
Rv0900		hypothetical protein	-0.4236	0.538	1.7	-	-	X	-
Rv0901		hypothetical protein	-0.661	0.104	5.2	-	-	X	-
Rv0902c	prrB	two component sensor histidine kinase PRRB	-0.678	0.021	30.2	-	-	X	-
Rv0903c	prrA	two component response transcriptional regulatory protein PRRA	1.7033	0.219	1.7	-	-	-	-
Rv0904c	accD3	acetyl-coenzyme A carboxylase	0.9097	0.016	0.9	-	-	-	-
Rv0905	echA6	enoyl-CoA hydratase	0.1898	0.703	6.1	-	-	-	-
Rv0906		hypothetical protein	-0.7756	0.064	15.1	X	-	-	-
Rv0907		hypothetical protein	0.3471	0.777	4.0	-	-	-	-
Rv0908	ctpE	metal cation transporter ATPase P-type	-0.6292	0.318	11.2	-	-	X	-
Rv0910		hypothetical protein	1.6445	0.001	3.3	-	-	-	-
Rv0911		hypothetical protein	-0.076	0.911	2.9	-	-	-	-
Rv0914c		acetyl-CoA acetyltransferase	-0.0498	0.832	19.3	-	-	-	-
Rv0921		resolvase	-0.667	0.265	4.6	-	-	-	-
Rv0923c		hypothetical protein	-0.1182	0.786	13.7	-	-	-	-
Rv0925c		hypothetical protein	0.7827	0.203	1.4	-	-	-	-
Rv0927c		short chain dehydrogenase	-1.073	0.094	3.3	-	-	-	-
Rv0928	pstS3	periplasmic phosphate-binding lipoprotein	0.4424	0.008	9.7	X	-	-	X
Rv0931c	pknD	transmembrane serine/threonine-protein kinase	-0.6422	0.045	105.8	-	-	X	-
Rv0932c	pstS2	periplasmic phosphate-binding lipoprotein	-0.7332	0.224	26.8	X	-	-	X
Rv0933	pstB	phosphate ABC transporter ATP-binding	0.1436	0.521	52.3	-	-	-	X

		protein							
Rv0934	pstS1	periplasmic phosphate-binding lipoprotein	#NUM!	0.081	137.4	X	-	-	X
Rv0935	pstC1	phosphate ABC transporter transmembrane	-0.6573	0.027	5.9	-	-	X	-
Rv0936	pstA2	phosphate ABC transporter transmembrane	0.4538	0.115	8.0	-	-	X	-
Rv0938		ATP-dependent DNA ligase	-0.2226	0.758	5.4	-	-	-	-
Rv0939		2-hydroxyhepta-2,4-diene-1,7-dioate isomerase	-0.1656	0.78	12.2	-	-	-	-
Rv0949	uvrD1	ATP-dependent DNA helicase II	0.5646	0.186	13.7	-	-	-	-
Rv0951	sucC	succinyl-CoA synthetase subunit beta	1.4092	0.062	3.7	-	-	-	-
Rv0952	sucD	succinyl-CoA synthetase subunit alpha	0.5636	0.376	13.1	-	-	-	-
Rv0954		transmembrane protein	-0.0543	0.558	27.6	-	-	X	-
Rv0955		integral membrane protein	-0.0534	0.977	1.1	-	-	X	-
Rv0957	purH	phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase	-0.412	0.592	3.7	-	-	-	-
Rv0958		magnesium chelatase	-0.4372	0.068	29.0	-	-	-	-
Rv0959		hypothetical protein	0.8958	0.003	31.2	-	-	-	-
Rv0966c		hypothetical protein	-0.088	0.715	32.0	-	-	-	-
Rv0968		hypothetical protein	-0.4094	0.396	3.2	-	-	-	-
Rv0969	CtpV	metal cation transporter P-type ATPase	-0.472	0.055	16.6	-	-	X	-
Rv0972c	fadE12	acyl-CoA dehydrogenase	1.497	0.001	2.1	-	-	-	-
Rv0979A	rpmF	50S ribosomal protein L32	0.5562	0.225	12.7	-	-	-	-
Rv0981	mprA	two component response transcriptional regulatory protein	-0.1838	0.506	13.9	-	-	-	-
Rv0982	mprB	two component sensor kinase	-0.945	0.034	15.9	-	-	X	-
Rv0983	pepD	serine protease	-0.2987	0.431	16.3	-	-	X	-
Rv0984	moaB2	pterin-4-alpha-carbinolamine dehydratase	0.0428	0.913	8.5	-	-	-	-
Rv0985c	mscL	large-conductance mechanosensitive channel	0.2464	0.099	30.5	-	-	X	-
Rv0986		adhesion component transport ATP-binding	-0.6646	0.199	9.5	-	-	-	-
Rv0987		adhesion component transport transmembrane	-1.667	0.085	3.9	-	-	X	-
Rv0988		hypothetical protein	-0.2067	0.47	10.6	-	-	X	-
Rv0992c		hypothetical protein	0.8407	0.054	1.3	-	-	-	-
Rv0993	galU	UTP--glucose-1-phosphate uridylyltransferase	-0.5994	0.004	18.2	-	-	-	-
Rv0994	moeA1	molybdopterin biosynthesis protein	0.1232	0.542	19.5	-	-	-	-
Rv0995	rimJ	ribosomal-protein-alanine acetyltransferase	0.9044	0.068	4.1	-	-	-	-
Rv0999		hypothetical protein	1.2238	0.004	14.9	-	-	-	-
Rv1001	arcA	arginine deiminase	-0.58		0.8	-	-	-	-
Rv1002c		hypothetical protein	0.3697	0.135	5.6	-	-	X	-
Rv1003		hypothetical protein	0.4505	0.266	0.9	-	-	-	-
Rv1006		hypothetical protein	-0.4498	0.084	81.9	X	-	-	-
Rv1007c	metG	methionyl-tRNA synthetase	-0.9404	0.006	2.9	-	-	-	-
Rv1013	pks16	acyl-CoA synthetase	-0.563	0.106	23.1	-	-	-	-
Rv1015c	rplY	50S ribosomal protein L25/general stress protein	0.8531	0.537	1.9	-	-	-	-
Rv1016c	lpqT	lipoprotein	-0.875	0.126	6.5	X	-	-	X
Rv1017c	prsA	ribose-phosphate pyrophosphokinase	0.5428	0.062	35.6	-	-	-	-
Rv1019		TetR family transcriptional regulator	2.0367	0.021	2.9	-	-	-	-
Rv1020	mfd	transcription-repair coupling factor Mfd (TRCF)	-0.2224	0.756	18.3	-	-	-	-
Rv1021		nucleoside triphosphate pyrophosphohydrolase	1.3586	0.056	5.4	-	-	-	-
Rv1022	lpqU	lipoprotein	-0.4954	0.464	5.8	X	-	X	X
Rv1023	eno	phosphopyruvate hydratase	-0.311	0.4	24.4	-	-	-	-
Rv1024		hypothetical protein	0.7736	0.148	6.6	-	X	-	-
Rv1028c	kdpD	sensor protein KDPD	0.0992	0.794	23.9	-	-	X	-
Rv1030	kdpB	potassium-transporting ATPase subunit B	-0.7689	0.027	6.6	-	-	X	-
Rv1031	kdpC	potassium-transporting ATPase subunit C	-0.2521	0.402	3.8	-	-	X	-
Rv1033c	trcR	two component transcriptional regulator	1.9494	0.052	3.7	-	-	-	-
Rv1045		hypothetical protein	0.6354	0.24	2.3	-	-	-	-
Rv1050		oxidoreductase	0.1832	0.514	13.8	-	-	-	-
Rv1056		hypothetical protein	0.4092	0.625	3.5	-	-	-	-
Rv1060		hypothetical protein	0.6784	0.043	15.5	-	-	-	-

Rv1063c		hypothetical protein	-0.698	0.057	1.8	-	-	-	-
Rv1069c		hypothetical protein	-0.7932	0.487	8.8	-	-	X	-
Rv1070c	echA8	enoyl-CoA hydratase	-1.1847	0.287	2.0	-	-	-	-
Rv1072		transmembrane protein	-0.5104	0.078	22.4	-	-	X	-
Rv1077	cbs	cystathionine beta-synthase	-1.1379	0.059	0.9	-	-	-	-
Rv1078	pra	proline-rich antigen	0.0964	0.502	66.4	-	-	X	-
Rv1081c		hypothetical protein	-0.9404	0.006	1.4	-	-	X	-
Rv1085c		hemolysin-like protein	0.6136	0.609	1.9	-	-	X	-
Rv1086		short (C15) chain Z-isoprenyl diphosphate synthase	0.6406	0.091	15.8	-	-	-	-
Rv1092c	coaA	pantothenate kinase	-1.0337	0.342	3.8	-	-	-	-
Rv1093	glyA	serine hydroxymethyltransferase	-0.5924	0.053	26.9	-	-	-	-
Rv1095	phoH2	PhoH-like protein	-0.2142	0.652	8.1	-	-	-	-
Rv1096		glycosyl hydrolase	0.338	0.35	25.2	-	-	X	-
Rv1097c		hypothetical protein	-0.7366	0.1	12.4	-	-	X	-
Rv1098c	fumC	fumarate hydratase	-0.472	0.689	4.3	-	-	-	-
Rv1100		hypothetical protein	-0.1314	0.394	14.7	-	-	X	-
Rv1101c		hypothetical protein	-0.6984	0.035	29.5	-	-	X	-
Rv1106c		cholesterol dehydrogenase	-0.5408	0.365	14.8	-	-	-	-
Rv1109c		hypothetical protein	-0.2032	0.223	20.7	-	-	-	-
Rv1111c		hypothetical protein	-0.7722	0.051	4.2	-	-	X	-
Rv1118c		hypothetical protein	0.8307	0.158	6.3	-	-	-	-
Rv1121	zwf1	glucose-6-phosphate 1-dehydrogenase	-0.0314	0.939	1.4	-	-	-	-
Rv1124	ephC	epoxide hydrolase	0.5832	0.457	11.9	-	-	-	-
Rv1125		hypothetical protein	0.862	0.077	0.7	-	-	-	-
Rv1129c		transcriptional regulator	0.1924	0.487	10.8	-	-	-	-
Rv1130		hypothetical protein	-0.3626	0.22	59.8	-	-	-	-
Rv1131	gltA1	citrate synthase	1.0252	0.104	6.5	-	-	-	-
Rv1132		hypothetical protein	-1.4204	0.14	20.4	-	-	X	-
Rv1133c	metE	5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase	1.3008	0	51.3	-	-	-	-
Rv1138c		oxidoreductase	-0.088	0.861	4.1	-	-	-	-
Rv1140		integral membrane protein	1.1096	0.078	12.1	-	-	X	-
Rv1141c	echA11	enoyl-CoA hydratase	0.386	0.704	9.5	-	-	-	-
Rv1144		short-chain type dehydrogenase/reductase	-0.7976	0.496	11.8	-	-	-	-
Rv1152		transcriptional regulator	0.2628	0.342	9.4	-	-	-	-
Rv1153c	omt	O-methyltransferase	-0.6465	0.074	6.5	-	-	-	-
Rv1154c		hypothetical protein	0.8519	0.28	2.5	-	-	-	-
Rv1160	mutT2	7,8-dihydro-8-oxoguanine-triphosphatase	1.0781	0.163	4.6	-	-	-	-
Rv1161	narG	respiratory nitrate reductase subunit alpha	0.3273	0.413	121.8	-	-	-	-
Rv1162	narH	respiratory nitrate reductase subunit beta	-0.0199	0.944	25.5	-	-	-	-
Rv1163	narJ	respiratory nitrate reductase subunit delta	0.9427	0.008	11.0	-	-	-	-
Rv1164	narI	respiratory nitrate reductase subunit gamma	0.828	0.028	18.8	-	-	X	-
Rv1165	typA	GTP-binding translation elongation factor	-1.5832	0.058	4.7	-	-	-	-
Rv1166	lpqW	lipoprotein	0.051	0.884	21.7	X	-	-	X
Rv1167c		transcriptional regulator	1.148	0.001	8.6	-	-	-	-
Rv1168c	PPE17	PPE family protein	-0.6104	0.323	2.6	-	-	-	-
Rv1169c	PE11	PE family protein	1.0048	-	0.5	-	-	-	-
Rv1172c	PE12	PE family protein	0.088	0.88	2.4	-	-	-	-
Rv1173	fbiC	FO synthase	-0.163	0.625	9.6	-	-	-	-
Rv1175c	fadH	NADPH dependent 2,4-dienoyl-CoA reductase	-0.0464	0.721	69.5	-	-	-	-
Rv1176c		hypothetical protein	0.8634	0.132	8.7	-	-	-	-
Rv1177	fdxC	ferredoxin FdxC	1.3011	0.269	1.7	-	-	-	-
Rv1178		N-succinyldiaminopimelate aminotransferase	-0.7553	0.207	4.1	-	-	-	-
Rv1179c		hypothetical protein	0.3184	0.253	47.0	-	-	-	-
Rv1180	pks3	polyketide beta-ketoacyl synthase	0.5753	0.001	100.1	-	-	-	-
Rv1181	pks4	polyketide beta-ketoacyl synthase	0.3144	0.014	409.4	-	-	-	-
Rv1182	papA3	polyketide synthase associated protein	0.139	0.186	16.8	-	-	-	-
Rv1183	mmpL10	transmembrane transport protein MmpL10	0.1072	0.818	37.2	X	-	X	-

Rv1184c		hypothetical protein	-0.3102	0.662	11.1	X	-	X	-
Rv1185c	fadD21	acyl-CoA synthetase	0.706	0.142	9.4	-	-	-	-
Rv1186c		hypothetical protein	-0.148	0.782	14.5	-	-	-	-
Rv1188		proline dehydrogenase	-1.1354	0.027	3.2	-	-	-	-
Rv1190		hypothetical protein	-0.4252	0.519	10.1	-	-	-	-
Rv1191		hypothetical protein	-0.0404	0.921	31.2	-	-	-	-
Rv1192		hypothetical protein	-0.2126	0.474	17.4	-	-	-	-
Rv1193	fadD36	acyl-CoA synthetase	0.5536	0.225	3.4	-	-	-	-
Rv1194c		hypothetical protein	-0.4054	0.285	13.2	-	-	-	-
Rv1196	PPE18	PPE family protein	-0.9946	0.125	7.6	-	-	-	-
Rv1201c		transferase	-0.2036	0.497	15.9	-	-	-	-
Rv1204c		hypothetical protein	-0.4814	0.446	20.2	-	-	-	-
Rv1206	fadD6	acyl-CoA synthetase	-0.0358	0.893	40.1	-	-	-	-
Rv1207	folP2	dihydropteroate synthase 2	-0.449	0.275	5.8	-	-	-	-
Rv1209		hypothetical protein	0.692	0.144	11.0	-	-	X	-
Rv1211		hypothetical protein	1.5898	-	0.7	-	-	-	-
Rv1215c		hypothetical protein	0.5906	0.465	11.1	-	-	-	-
Rv1216c		integral membrane protein	0.3396	0.14	26.8	-	-	X	-
Rv1217c		tetronasin-transport integral membrane protein	0.738	0.053	33.7	-	-	X	-
Rv1218c		tetronasin-transport ATP-binding protein ABC	0.7154	0.014	57.6	-	-	-	-
Rv1223	htrA	serine protease	-0.2098	0.218	153.4	-	-	X	-
Rv1226c		hypothetical protein	0.0048		0.7	-	-	X	-
Rv1227c		hypothetical protein	-0.1104	0.737	11.5	-	-	X	-
Rv1228	lpqX	lipoprotein	-1.2752	0.241	3.1	-	-	-	X
Rv1229c	mrp	MRP family ATP-binding protein	0.4613	0.332	14.5	-	-	-	-
Rv1231c		hypothetical protein	-0.479	0.166	26.7	-	-	X	-
Rv1232c		hypothetical protein	-1.445	0.047	8.1	-	-	-	-
Rv1234		hypothetical protein	0.6238	0.165	25.1	-	-	X	-
Rv1235	lpqY	sugar-binding lipoprotein	-0.2372	0.27	25.7	X	-	-	X
Rv1237	sugB	sugar-transport integral membrane protein ABC	-0.0214	0.94	7.1	-	-	X	-
Rv1238	sugC	sugar-transport ATP-binding protein ABC	-0.116	0.726	22.0	-	-	-	-
Rv1239c	corA	magnesium/cobalt transporter	0.0142	0.96	24.9	-	-	X	-
Rv1244	lpqZ	lipoprotein	-0.314	0.69	5.2	X	-	-	X
Rv1245c		short-chain type dehydrogenase/reductase	-0.5614	0.145	43.3	-	-	-	-
Rv1247c		hypothetical protein	0.0366	0.916	5.6	-	-	-	-
Rv1248c	kgd	alpha-ketoglutarate decarboxylase	0.3582	0.355	27.8	-	-	-	-
Rv1249c		hypothetical protein	1.1986	0.131	3.3	-	-	X	-
Rv1252c	lprE	lipoprotein	-0.503	0.566	2.7	X	-	-	X
Rv1253	deaD	cold-shock DEAD-box protein A	0.2622	0.611	32.9	-	-	-	-
Rv1254		acyltransferase	-0.1762	0.669	2.5	-	-	X	-
Rv1255c		transcriptional regulator	-0.1698	0.675	4.2	-	-	-	-
Rv1259		hypothetical protein	-1.0362	0.044	0.9	-	-	-	-
Rv1261c		hypothetical protein	0.6536	0.074	34.9	-	-	-	-
Rv1262c		HIT-like protein	1.107	0.202	5.4	-	-	-	-
Rv1263	amiB2	amidase	-0.2635	0.597	13.3	-	-	-	-
Rv1264		adenylyl cyclase	-0.5794	0.163	3.9	-	-	-	-
Rv1265		hypothetical protein	0.4168	0.255	10.2	-	-	-	-
Rv1266c	pknH	transmembrane serine/threonine-protein kinase H	-0.4281	0.079	41.2	-	-	X	-
Rv1269c		hypothetical protein	-1.0238	0.115	6.6	X	X	-	-
Rv1270c	lprA	lipoprotein	-0.309	0.148	121.3	X	-	X	X
Rv1272c		drugs-transport transmembrane ATP-binding protein ABC transporter	-0.667	0.164	8.1	-	-	X	-
Rv1274	lprB	lipoprotein	1.1946	0.024	6.4	X	-	-	X
Rv1275	lprC	lipoprotein	0.2462	0.197	21.1	X	-	-	X
Rv1278		hypothetical protein	1.0868	0.03	14.6	-	-	-	-
Rv1279		dehydrogenase FAD flavoprotein	-0.2874	0.327	25.2	-	-	-	-
Rv1280c	oppA	periplasmic oligopeptide-binding lipoprotein	-1.0478	0.001	22.5	-	X	X	X
Rv1281c	oppD	oligopeptide-transport ATP-binding protein	-0.0404	0.892	52.5	-	-	-	-

		ABC							
Rv1283c	oppB	oligopeptide-transport integral membrane protein	-0.0645	0.934	5.2	-	-	X	-
Rv1284		hypothetical protein	1.4947	0.006	1.0	-	-	-	-
Rv1285	cysD	sulfate adenylyltransferase subunit 2	0.1811	0.859	3.0	-	-	-	-
Rv1286	cysN	bifunctional sulfate adenylyltransferase subunit 1	-0.4376	0.13	15.8	-	-	-	-
Rv1288		hypothetical protein	-0.5469	0.045	28.5	-	-	-	-
Rv1290c		hypothetical protein	1.0531	0.028	1.1	-	X	X	-
Rv1294	thrA	homoserine dehydrogenase	0.2318	0.554	17.7	-	-	-	-
Rv1297	rho	transcription termination factor Rho	-0.0612	0.851	49.9	-	-	-	-
Rv1304	atpB	F0F1 ATP synthase subunit A	0.1508	0.538	26.8	-	-	X	-
Rv1306	atpF	F0F1 ATP synthase subunit B	0.0746	0.643	60.1	-	-	X	-
Rv1307	atpH	F0F1 ATP synthase subunit delta	-0.5629	0.368	131.7	-	-	X	-
Rv1308	atpA	F0F1 ATP synthase subunit alpha	-0.2608	0.471	333.1	-	-	-	-
Rv1309	atpG	F0F1 ATP synthase subunit gamma	0.123	0.578	150.0	-	-	-	-
Rv1310	atpD	F0F1 ATP synthase subunit beta	-0.556	0.113	204.3	-	-	-	-
Rv1311	atpC	F0F1 ATP synthase subunit epsilon	0.7152	0.036	22.9	-	-	-	-
Rv1312		hypothetical protein	0.1226	0.4	15.7	-	-	X	-
Rv1314c		hypothetical protein	0.9515	0.17	0.9	-	-	-	-
Rv1316c	ogt	-cysteine methyltransferase	0.6354	0.478	5.0	-	-	-	-
Rv1319c		adenylate cyclase	-0.607	0.207	2.5	-	-	X	-
Rv1320c		adenylate cyclase	-1.5118	0.004	3.9	-	-	X	-
Rv1321		hypothetical protein	-0.0413	0.917	8.0	-	-	-	-
Rv1327c	glgE	glucanase	-0.698	0.105	14.3	-	-	-	-
Rv1328	glgP	glycogen phosphorylase GlgP	0.7288	0.345	15.0	-	-	-	-
Rv1331	clpS	ATP-dependent Clp protease adaptor protein	0.0742	0.877	3.2	-	-	-	-
Rv1332		transcriptional regulator	0.107	0.738	2.6	-	-	-	-
Rv1333		hydrolase	-0.0543	0.879	2.0	-	-	-	-
Rv1335	CFP10A	9.5 kDa culture filtrate antigen	0.0048		0.7	-	-	-	-
Rv1337		integral membrane protein	0.1136	0.541	1.4	-	-	X	-
Rv1338	murI	glutamate racemase	1.5825	0.013	5.0	-	-	-	-
Rv1348		drugs-transport transmembrane ATP-binding protein ABC transporter	-0.7045	0.131	1.4	-	-	X	-
Rv1362c		hypothetical protein	-0.0978	0.735	14.5	-	-	X	-
Rv1363c		hypothetical protein	-0.9352	0.001	19.5	-	-	X	-
Rv1368	lprF	lipoprotein	0.5293	0.037	55.3	X	-	X	X
Rv1373		glycolipid sulfotransferase	0.7736	0.399	3.8	-	-	-	-
Rv1376		hypothetical protein	-0.0486	0.882	1.1	-	-	-	-
Rv1379	pyrR	uracil phosphoribosyltransferase	0.7116	0.027	5.7	-	-	-	-
Rv1380	pyrB	aspartate carbamoyltransferase catalytic subunit	-1.1472	0.006	15.6	-	-	-	-
Rv1382		export or membrane protein	-0.2244	0.099	11.3	-	-	X	-
Rv1383	carA	carbamoyl phosphate synthase small subunit	1.1832	0.011	4.0	-	-	-	-
Rv1384	carB	carbamoyl phosphate synthase large subunit	1.1218	0.065	14.9	-	-	-	-
Rv1386	PE15	PE family protein	-0.8136	0.053	2.1	-	-	-	-
Rv1387	PPE20	PPE family protein	-0.5368	0.024	21.7	-	-	-	-
Rv1388	mihF	putative integration host factor	-0.1307	0.817	15.3	-	-	-	-
Rv1391	dfp	phosphopantothenoylcysteine decarboxylase	-1.497	0.011	5.5	-	-	-	-
Rv1392	metK	S-adenosylmethionine synthetase	0.3613	0.419	11.7	-	-	-	-
Rv1393c		monoxygenase	-0.2614	0.221	14.4	-	-	-	-
Rv1394c	cyp132	cytochrome P450 132	-0.64	0.118	4.8	-	-	-	-
Rv1397c		hypothetical protein	1.107	0.285	3.0	-	-	-	-
Rv1398c		hypothetical protein	0.1326	0.51	4.5	-	-	-	-
Rv1399c	lipH	lipase	0.2544	0.503	2.1	-	-	-	-
Rv1400c	lipI	lipase	0.1045	0.727	14.1	-	-	-	-
Rv1401		hypothetical protein	1.1137	0.019	1.1	-	-	X	-
Rv1404		transcriptional regulator	0.1036	0.694	12.9	-	-	-	-
Rv1407	fmu	Fmu protein (SUN protein)	-1.147	0.351	2.8	-	-	-	-

Rv1409	ribG	Probable bifunctional riboflavin biosynthesis	-0.277	0.647	7.1	-	-	-	-
Rv1410c		aminoglycosides/tetracycline-transport integral	-0.6584	0.602	14.4	-	-	X	-
Rv1411c	lprG	lipoprotein	-0.0143	0.91	60.3	X	-	-	X
Rv1412	ribC	riboflavin synthase subunit alpha	1.0662	0.058	6.2	-	-	-	-
Rv1417		hypothetical protein	0.144	0.661	5.3	-	-	X	-
Rv1418	lprH	lipoprotein	-0.7945	0.087	5.5	X	-	X	X
Rv1421		hypothetical protein	0.3736	0.318	15.4	-	-	-	-
Rv1422		hypothetical protein	1.0596	0.004	1.3	-	-	-	-
Rv1423	whiA	transcriptional regulator	1.1254	0.061	15.4	-	-	-	-
Rv1424c		hypothetical protein	0.1636	0.584	7.6	-	-	-	-
Rv1425		hypothetical protein	-0.1127	0.587	21.4	-	-	-	-
Rv1426c	lipO	esterase	-1.1127	0.006	2.5	X	-	X	-
Rv1427c	fadD12	acyl-CoA synthetase	-0.2941	0.425	27.0	-	-	-	-
Rv1428c		hypothetical protein	-0.7018	0.156	9.0	-	-	-	-
Rv1429		hypothetical protein	-0.1584	0.829	5.7	-	-	-	-
Rv1431		hypothetical protein	-0.9437	0.176	3.0	-	-	X	-
Rv1436	gap	glyceraldehyde-3-phosphate dehydrogenase	-0.1016	0.178	8.9	-	-	-	-
Rv1442	bisC	biotin sulfoxide reductase	0.6656	0.022	16.6	-	-	-	-
Rv1443c		hypothetical protein	-0.2274	0.134	21.8	-	-	-	-
Rv1444c		hypothetical protein	0.3896	0.395	7.5	-	-	-	-
Rv1446c	opcA	putative OXPP cycle protein	-0.0952	0.806	25.2	-	-	-	-
Rv1447c	zwf2	glucose-6-phosphate 1-dehydrogenase	0.5152	0.465	12.2	-	-	-	-
Rv1449c	tkt	transketolase	0.6488	0.123	3.8	-	-	-	-
Rv1454c	qor	quinone reductase	0.107	0.849	1.7	-	-	-	-
Rv1455		hypothetical protein	0.0231	0.965	8.4	-	-	-	-
Rv1457c		unidentified antibiotic-transport integral	0.1136	0.541	0.7	-	-	X	-
Rv1458c		unidentified antibiotic-transport ATP-binding	-0.3761	0.218	25.4	-	-	-	-
Rv1459c		integral membrane protein	-0.698	0.461	4.5	-	-	X	-
Rv1461		hypothetical protein	0.588	0.456	4.8	-	-	-	-
Rv1462		hypothetical protein	-1.802	0.001	2.3	-	-	-	-
Rv1463		ABC transporter ATP-binding protein	1.9777	0.003	2.6	-	-	-	-
Rv1464	csd	cysteine desulfurase	-1.5526	0.178	5.3	-	-	-	-
Rv1470	trxA	thioredoxin TRXA	1.2431	0.246	1.6	-	-	-	-
Rv1472	echA12	enoyl-CoA hydratase	0.5395	0.382	16.9	-	-	-	-
Rv1473		macrolide ABC transporter ATP-binding protein	0.3928	0.189	3.3	-	-	-	-
Rv1474c		transcriptional regulator	-0.1076	0.676	12.7	-	-	-	-
Rv1476		hypothetical protein	0.1913	0.068	24.7	-	-	X	-
Rv1479	moxR1	transcriptional regulator	-0.2554	0.467	77.8	-	-	-	-
Rv1480		hypothetical protein	0.9896	0.104	10.8	-	-	-	-
Rv1481		hypothetical protein	-0.7714	0.069	23.1	-	-	X	-
Rv1483	fabG1	3-oxoacyl-[acyl-carrier protein] reductase	0.4404	0.522	4.2	-	-	-	-
Rv1484	inhA	enoyl-(acyl carrier protein) reductase	0.3246	0.694	3.8	-	-	-	-
Rv1486c		hypothetical protein	-0.2824	0.271	11.2	-	-	-	-
Rv1487		hypothetical protein	-0.034	0.957	2.7	-	-	X	-
Rv1488		hypothetical protein	-0.3374	0.228	180.8	-	-	X	-
Rv1489		hypothetical protein	1.93	0	10.3	-	-	X	-
Rv1498A		hypothetical protein	1.1357	0.517	1.9	-	-	-	-
Rv1500		glycosyltransferase	-0.7782	0.264	6.6	-	-	X	-
Rv1501		hypothetical protein	1.8408	0.025	1.9	-	-	-	-
Rv1502		hypothetical protein	0.4403	0.647	3.5	-	-	-	-
Rv1508c		hypothetical protein	0.243	0.426	50.4	-	-	X	-
Rv1509		hypothetical protein	-0.6436	0.247	5.4	-	-	-	-
Rv1513		hypothetical protein	1.1124	0.344	7.2	-	-	-	-
Rv1514c		hypothetical protein	1.3586	0.102	5.7	-	-	-	-
Rv1515c		hypothetical protein	0.3128	0.487	12.1	-	-	-	-
Rv1516c		sugar transferase	0.0906	0.798	17.5	-	-	-	-
Rv1521	fadD25	acyl-CoA synthetase	0.2142	0.782	4.4	-	-	X	-
Rv1523		methyltransferase	-0.4008	0.258	17.9	-	-	-	-
Rv1524		glycosyltransferase	2.5467	0.268	2.3	-	-	-	-
Rv1526c		glycosyltransferase	-0.2737	0.755	2.3	-	-	-	-

Rv1531		hypothetical protein	-0.9126	0.155	3.6	-	-	-	-
Rv1532c		hypothetical protein	0.5898		0.8	-	-	-	-
Rv1534		transcriptional regulator	1.302	0.193	3.4	-	-	-	-
Rv1535		hypothetical protein	0.0759	0.87	8.0	-	-	-	-
Rv1538c	ansA	L-aparaginase ansA	-1.7834	-	0.7	-	-	-	-
Rv1539	lspA	lipoprotein signal peptidase	0.3646	0.664	2.5	-	-	X	-
Rv1543		fatty acyl-CoA reductase	-0.443	0.139	80.6	-	-	-	-
Rv1544		ketoacyl reductase	0.8782	0	34.9	-	-	-	-
Rv1546		hypothetical protein	1.2343	0.032	3.6	-	-	-	-
Rv1547	dnaE	DNA polymerase III subunit alpha	2.3813	0.132	2.2	-	-	-	-
Rv1549	fadD11.1	fatty-acid-CoA ligase	-0.0003	0.999	8.2	-	-	-	-
Rv1552	frdA	fumarate reductase flavoprotein subunit	0.5764	0.422	1.0	-	-	-	-
Rv1553	frdB	fumarate reductase iron-sulfur subunit	-0.5876	0.17	3.7	-	-	-	-
Rv1556		regulatory protein	-0.8864	0.029	1.1	-	-	-	-
Rv1563c	treY	maltooligosyltrehalose synthase	0.5933	0.446	0.7	-	-	-	-
Rv1569	bioF1	8-amino-7-oxononanoate synthase	0.302	0.381	1.9	-	-	-	-
Rv1578c		phiRv1 phage protein	0.0169	0.898	3.3	-	-	-	-
Rv1586c		phiRv1 integrase	1.333	0.147	5.0	-	-	-	-
Rv1589	bioB	biotin synthase	1.3586	0.047	3.2	-	-	-	-
Rv1591		hypothetical protein	-0.3784	0.3	10.3	-	-	X	-
Rv1592c		hypothetical protein	0.9686	0.066	2.1	-	-	-	-
Rv1593c		hypothetical protein	1.012	0.025	3.3	-	-	-	-
Rv1594	nadA	quinolinate synthetase	-0.1572	0.731	13.0	-	-	-	-
Rv1595	nadB	L-aspartate oxidase	0.6353	0.339	1.5	-	-	-	-
Rv1596	nadC	nicotinate-nucleotide pyrophosphorylase	-1.5556	0.028	12.1	-	-	-	-
Rv1598c		hypothetical protein	0.9708	0.003	29.6	-	-	-	-
Rv1602	hisH	imidazole glycerol phosphate synthase subunit	-0.8364	0.08	3.9	-	-	-	-
Rv1605	hisF	imidazole glycerol phosphate synthase subunit	0.5768	0.288	13.7	-	-	-	-
Rv1608c	bcpB	peroxidoxin	-0.3844	0.353	10.3	-	-	-	-
Rv1610		hypothetical protein	-0.1078	0.944	1.9	-	-	X	-
Rv1611	trpC	indole-3-glycerol-phosphate synthase	1.5326	0.057	16.9	-	-	-	-
Rv1612	trpB	tryptophan synthase subunit beta	0.2453	0.803	4.5	-	-	-	-
Rv1613	trpA	tryptophan synthase subunit alpha	0.012	0.948	1.5	-	-	-	-
Rv1614	lgt	prolipoprotein diacylglyceryl transferase	-0.3585	0.02	45.9	-	-	X	-
Rv1615		hypothetical protein	0.6137	0.506	1.6	-	-	X	-
Rv1617	pykA	pyruvate kinase	0.3586	0.242	7.2	-	-	-	-
Rv1618	tesB1	acyl-CoA thioesterase II	0.7955	0.172	1.9	-	-	-	-
Rv1620c	cydC	cytochrome' transport ABC transporter	0.2454	0.719	7.8	-	-	X	-
Rv1621c	cydD	cytochrome' transport t ABC transporter	0.1203	0.797	2.3	-	-	X	-
Rv1622c	cydB	cytochrome D ubiquinol oxidase (subunit II)	1.1946	0.111	5.9	-	-	X	-
Rv1623c	cydA	cytochrome D ubiquinol bd-I oxidase subunit I)	-0.4472	0.024	28.4	-	-	X	-
Rv1625c	cya	adenylyl cyclase	0.2504	0.341	15.9	-	-	X	-
Rv1627c		lipid-transfer protein	0.3075	0.102	4.0	-	-	-	-
Rv1628c		hypothetical protein	0.1136	0.541	1.4	-	-	-	-
Rv1629	polA	DNA polymerase I	0.5526	0.175	18.7	-	-	-	-
Rv1630	rpsA	30S ribosomal protein S1	1.3664	0.025	18.1	-	-	-	-
Rv1633	urvB	excinuclease ABC subunit B	0.0981	0.69	11.9	-	-	-	-
Rv1636	TB15.3	hypothetical protein	0.4164	0.274	20.7	-	-	-	-
Rv1638	uvrA	excinuclease ABC subunit A	-0.1486	0.487	57.5	-	-	-	-
Rv1640c	lysS	lysyl-tRNA synthetase	0.4374	0.031	59.3	-	-	X	-
Rv1641	infC	translation initiation factor IF-3	-1.988	0.007	1.8	-	-	-	-
Rv1643	rplT	50S ribosomal protein L20	0.3813	0.731	2.4	-	-	-	-
Rv1647		hypothetical protein	0.0464	0.842	8.7	-	-	-	-
Rv1649	pheS	phenylalanyl-tRNA synthetase subunit alpha	0.5917	0.359	1.9	-	-	-	-
Rv1650	pheT	phenylalanyl-tRNA synthetase subunit beta	-0.997	0.277	1.6	-	-	-	-
Rv1652	argC	N-acetyl-gamma-glutamyl-phosphate reductase	-1.2029	0.157	8.1	-	-	-	-
Rv1658	argG	argininosuccinate synthase	-1.324	0.23	5.8	-	-	-	-
Rv1660	pks10	chalcone synthase	-0.377	0.04	4.9	-	-	-	-

Rv1661	pks7	polyketide synthase	-0.7654	0.349	6.1	-	-	-	-
Rv1662	pks8	polyketide synthase	-0.3324	0.377	9.3	-	-	-	-
Rv1664	pks9	polyketide synthase	-0.0828	0.682	0.5	-	-	-	-
Rv1665	pks11	chalcone synthase	-0.3648	0.333	6.4	-	-	-	-
Rv1666c	cyp139	cytochrome P450 139	0.302	0.381	3.9	-	-	-	-
Rv1667c		macrolide-transport ATP-binding protein ABC	1.3888	0.233	1.8	-	-	-	-
Rv1674c		transcriptional regulator	1.012	0.025	1.1	-	-	-	-
Rv1676		hypothetical protein	-0.1698	0.675	2.1	-	-	-	-
Rv1678		integral membrane protein	0.302	0.444	3.8	-	-	X	-
Rv1679	fadE16	acyl-CoA dehydrogenase	-0.6144	0.012	8.8	-	-	-	-
Rv1680		hypothetical protein	0.6588	0.474	9.0	-	-	-	-
Rv1683		acyl-CoA synthetase	0.132	0.556	52.1	-	-	-	-
Rv1684		hypothetical protein	-0.0903	0.524	2.1	-	-	-	-
Rv1685c		hypothetical protein	0.2544	0.503	1.0	-	-	-	-
Rv1687c		ABC transporter ATP-binding protein	0.8921	0.098	4.6	-	-	-	-
Rv1689	tyrS	tyrosyl-tRNA synthetase	1.0252	0.069	2.1	-	-	-	-
Rv1697		hypothetical protein	1.076	0.22	6.7	-	-	X	-
Rv1698		hypothetical protein	0.0717	0.885	20.9	-	-	X	-
Rv1699	pyrG	CTP synthetase	0.7736	0.079	12.8	-	-	-	-
Rv1701	xerD	site-specific tyrosine recombinase	0.396	0.21	25.5	-	-	-	-
Rv1703c		catechol-o-methyltransferase	-0.3472	0.174	18.6	-	-	-	-
Rv1707		transmembrane protein	-0.1774	0.401	36.8	-	-	X	-
Rv1708		putative initiation inhibitor protein	0.192	0.584	41.6	-	-	-	-
Rv1711		hypothetical protein	0.2364	0.696	17.7	-	-	-	-
Rv1731	gabD2	succinic semialdehyde dehydrogenase	-0.2186	0.169	40.8	-	-	-	-
Rv1738		hypothetical protein	1.704	0.018	6.2	-	-	-	-
Rv1742		hypothetical protein	-0.186	0.888	3.8	-	-	-	-
Rv1743	pknE	transmembrane serine/threonine-protein kinase E	0.0498	0.917	24.1	-	-	X	-
Rv1746	pknF	membrane serine/threonine-protein kinase	-0.0206	0.928	26.3	-	-	X	-
Rv1747		ATP-binding protein ABC transporter	-0.0584	0.53	78.3	-	-	X	-
Rv1748		hypothetical protein	0.7884	0.058	10.9	-	-	X	-
Rv1749c		integral membrane protein	0.886	0.186	16.8	-	-	X	-
Rv1751		hypothetical protein	-0.0767	0.841	28.1	-	-	-	-
Rv1754c		hypothetical protein	-2.0167	0.005	2.7	-	-	X	-
Rv1760		hypothetical protein	0.5336	0.375	5.5	-	-	-	-
Rv1762c		hypothetical protein	0.4736	0.359	11.7	-	-	-	-
Rv1766		hypothetical protein	0.862	0.077	0.7	-	-	-	-
Rv1767		hypothetical protein	1.356	0.027	1.7	-	-	-	-
Rv1769		hypothetical protein	1.0504	0.242	8.8	-	-	-	-
Rv1770		hypothetical protein	0.0926	0.859	26.1	-	-	-	-
Rv1771		oxidoreductase	-0.211	0.097	66.4	-	-	-	-
Rv1772		hypothetical protein	-0.1955	0.791	1.4	-	-	-	-
Rv1775		hypothetical protein	0.0874	0.929	3.9	-	-	-	-
Rv1778c		hypothetical protein	-0.802	0.015	2.7	-	-	-	-
Rv1780		hypothetical protein	0.393	0.376	1.6	-	-	-	-
Rv1782		hypothetical protein	-0.2788	0.469	25.9	-	-	X	-
Rv1783		hypothetical protein	0.1368	0.553	60.1	-	-	X	-
Rv1784		hypothetical protein	-0.0566	0.707	156.9	-	-	-	-
Rv1794		hypothetical protein	-0.4	0.027	25.8	-	-	-	-
Rv1795		hypothetical protein	0.1228	0.191	80.7	-	-	X	-
Rv1796	mycP5	proline rich membrane-anchored mycosin	-0.8455	0.135	14.4	X	-	X	-
Rv1797		hypothetical protein	-0.324	0.305	24.3	-	-	X	-
Rv1798		hypothetical protein	-0.0486	0.882	0.6	-	-	-	-
Rv1808	PPE32	PPE family protein	0.6588	0.03	6.4	-	-	-	-
Rv1812c		dehydrogenase	0.1784	0.425	9.2	-	-	-	-
Rv1819c		drugs-transport ABC transporter	-0.0061	0.983	50.5	-	-	X	-
Rv1820	ilvG	hypothetical protein	-0.1411	0.449	14.4	-	-	-	-
Rv1822	pgsA2	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase	-0.4848	0.017	20.0	-	-	X	-
Rv1823		hypothetical protein	-0.2029	0.665	31.4	-	-	-	-
Rv1824		hypothetical protein	0.7867	0.022	1.7	-	-	X	-

Rv1825		hypothetical protein	-0.7684	0.126	22.4	-	-	X	-
Rv1827	cfp17	hypothetical protein	0.8304	0.16	4.9	-	-	-	-
Rv1828		hypothetical protein	0.1043	0.47	17.7	-	-	-	-
Rv1829		hypothetical protein	0.4627	0.269	11.3	-	-	-	-
Rv1830		hypothetical protein	1.0048	-	1.0	-	-	-	-
Rv1833c		haloalkane dehalogenase	-0.2483	0.415	44.9	-	-	-	-
Rv1834		hydrolase	-1.277	0.182	7.3	-	-	-	-
Rv1835c		hypothetical protein	0.2137	0.839	4.4	-	-	-	-
Rv1836c		hypothetical protein	0.1826	0.076	93.9	-	-	X	-
Rv1837c	glcB	malate synthase G	1.935	0.005	7.0	-	-	-	-
Rv1841c		hypothetical protein	-0.9672	0.06	15.6	-	-	X	-
Rv1842c		hypothetical protein	-0.2008	0.581	8.2	-	-	X	-
Rv1843c	guaB1	inosine 5-monophosphate dehydrogenase	-0.2127	0.396	9.1	-	-	-	-
Rv1844c	gnd1	6-phosphogluconate dehydrogenase	-1.8696	0.006	10.9	-	-	-	-
Rv1847		hypothetical protein	0.1121	0.897	3.4	-	-	-	-
Rv1852	ureG	urease accessory protein	-0.1576	0.589	9.3	-	-	-	-
Rv1853	ureD	urease accessory protein	-0.0486	0.882	1.1	-	-	-	-
Rv1854c	ndh	NADH dehydrogenase	0.096	0.51	11.9	-	-	X	-
Rv1855c		oxidoreductase	0.9686	0.027	3.7	-	-	-	-
Rv1860	apa	hypothetical protein	0.619	0.411	3.1	X	-	X	-
Rv1865c		short chain dehydrogenase	0.3451	0.344	18.6	-	-	-	-
Rv1866		hypothetical protein	1.0048	-	0.5	-	-	-	-
Rv1868		hypothetical protein	0.1857	0.583	12.6	-	-	-	-
Rv1870c		hypothetical protein	0.1712	0.917	1.0	-	-	-	-
Rv1871c		hypothetical protein	0.1066	0.595	64.0	-	-	-	-
Rv1872c	lldD2	L-lactate dehydrogenase (cytochrome)	-0.3222	0.21	641.4	-	-	-	-
Rv1874		hypothetical protein	0.0072	0.985	2.0	-	-	-	-
Rv1875		hypothetical protein	0.302	0.533	4.6	-	-	-	-
Rv1876	bfrA	bacterioferritin	-0.7856	0.159	4.0	-	-	-	-
Rv1880c	cyp140	cytochrome p450 140	-0.5624	0.028	40.2	-	-	-	-
Rv1882c		short chain dehydrogenase	-0.0388	0.888	16.9	-	-	-	-
Rv1883c		hypothetical protein	0.2142	0.764	4.7	-	-	-	-
Rv1885c		chorismate mutase	-0.6305	0.161	9.1	X	-	X	-
Rv1886c	fbpB	secreted antigen 85-B fbpB	0.005	0.985	73.7	-	X	X	-
Rv1890c		hypothetical protein	0.4403	0.597	2.1	-	-	-	-
Rv1893		hypothetical protein	-1.1252	0.272	5.0	-	-	-	-
Rv1899c	lppD	lipoprotein	0.0602	0.871	7.7	X	-	-	X
Rv1900c	lipJ	lignin peroxidase	0.0252	0.899	4.7	-	-	-	-
Rv1901	cinA	competence damage-inducible protein A	-0.4803	0.43	1.8	-	-	-	-
Rv1904		hypothetical protein	1.3782	0.041	5.1	-	-	-	-
Rv1907c		hypothetical protein	0.393	0.376	1.6	-	-	-	-
Rv1908c	katG	catalase-peroxidase-peroxynitritase T	1.2972	0	8.6	-	-	-	-
Rv1912c	fadB5	oxidoreductase	0.0364	0.866	30.2	-	-	-	-
Rv1914c		hypothetical protein	0.862	0.077	1.4	-	-	-	-
Rv1915	aceAa	isocitrate lyase	0.0206	0.909	25.8	-	-	-	-
Rv1916	aceAb	isocitrate lyase	0.0596	0.689	0.9	-	-	-	-
Rv1919c		hypothetical protein	-0.5452	0.071	45.2	-	-	-	-
Rv1920		hypothetical protein	0.3529	0.137	31.3	-	-	-	-
Rv1922		lipoprotein	-1.441	0.001	3.9	X	-	-	X
Rv1923	lipD	lipase	-0.3646	0.473	1.6	-	-	-	-
Rv1924c		hypothetical protein	-0.5449	0.044	13.3	-	-	X	-
Rv1925	fadD31	acyl-CoA synthetase	-0.3254	0.173	78.0	-	-	-	-
Rv1926c	mpt63	immunogenic protein	-0.8622	0	72.2	X	-	X	-
Rv1927		hypothetical protein	-0.088	0.921	6.5	-	-	-	-
Rv1929c		hypothetical protein	1.3896	0.103	5.8	-	-	-	-
Rv1930c		hypothetical protein	0.686	0.032	4.7	-	-	-	-
Rv1937		oxygenase	0.1163	0.777	2.7	-	-	-	-
Rv1940	ribA1	riboflavin biosynthesis protein	-0.0903	0.524	0.7	-	-	-	-
Rv1942c		hypothetical protein	1.6893	0.057	2.1	-	-	-	-
Rv1956		transcriptional regulator	0.754	0.006	6.7	-	-	-	-
Rv1957		hypothetical protein	1.2544	0.033	3.5	-	-	-	-
Rv1976c		hypothetical protein	-0.0486	0.882	1.1	-	-	-	-

Rv1977		hypothetical protein	0.3526	0.642	3.7	-	-	X	-
Rv1978		hypothetical protein	0.3174	0.365	58.1	-	-	-	-
Rv1980c	mpt64	immunogenic protein (antigen MPT64/MPB64)	0.5872	0.025	18.6	X	-	X	-
Rv1984c	cfp21	cutinase precursor CFP21	-0.1976	0.606	3.0	X	-	-	-
Rv1992c	ctpG	metal cation transporter P-type ATPase G	0.5144	0.141	6.3	-	-	X	-
Rv1993c		hypothetical protein	-0.3612	0.435	10.2	-	-	-	-
Rv1996		hypothetical protein	1.1724	0	10.4	-	-	-	-
Rv1998c		hypothetical protein	0.223	0.4	12.8	-	-	-	-
Rv2000		hypothetical protein	-1.1469	0.425	1.4	-	-	-	-
Rv2002	fabG3	20-beta-hydroxysteroid dehydrogenase	1.9686	0.001	8.0	-	-	-	-
Rv2005c		hypothetical protein	0.9686	0.015	21.7	-	-	-	-
Rv2010		hypothetical protein	-0.4214	0.2	3.1	-	-	-	-
Rv2011c		hypothetical protein	0.012	0.948	1.5	-	-	-	-
Rv2012		hypothetical protein	-0.0903	0.524	1.4	-	-	-	-
Rv2016		hypothetical protein	0.9994	0.11	9.7	-	-	-	-
Rv2017		transcriptional regulator	1.3927	0.009	5.5	-	-	-	-
Rv2018		hypothetical protein	1.4471	0.046	0.9	-	-	-	-
Rv2021c		transcriptional regulator	0.302	0.317	2.4	-	-	-	-
Rv2026c		hypothetical protein	1.377	0.007	12.8	-	-	-	-
Rv2027c		histidine kinase response regulator	0.6792	0.174	24.7	-	-	-	-
Rv2030c		hypothetical protein	2.6715	0.013	12.4	-	-	-	-
Rv2031c	hspX	heat shock protein	2.4246	0.036	36.2	-	-	-	-
Rv2033c		hypothetical protein	0.7932	0.264	6.6	-	-	-	-
Rv2035		hypothetical protein	-0.3647	0.178	1.9	-	-	-	-
Rv2036		hypothetical protein	0.4947	0.053	1.6	-	-	-	-
Rv2037c		transmembrane protein	-1.7546	0.083	6.7	-	-	-	-
Rv2041c		sugar-binding lipoprotein	-1.704	0.084	5.3	X	X	-	X
Rv2042c		hypothetical protein	1.7263	0.12	6.9	-	-	-	-
Rv2045c	lipT	carboxylesterase	-0.0248	0.928	17.6	-	-	-	-
Rv2047c		hypothetical protein	0.1262	0.759	18.6	-	-	-	-
Rv2048c	pks12	polyketide synthase	0.1356	0.025	80.3	-	-	-	-
Rv2051c	ppm1	polyprenol-monophosphomannose synthase Ppm1	-0.4334	0.345	36.8	-	-	X	-
Rv2054		hypothetical protein	0.1084	0.577	40.4	-	-	-	-
Rv2061c		hypothetical protein	0.2498	0.364	20.3	-	-	-	-
Rv2062c	cobN	cobaltochelatase subunit CobN	0.4878	0.066	35.7	-	-	-	-
Rv2066	cobI	S-adenosyl-L-methionine-precorrin-2 methyl transferase/precorrin-3 methylase	0.4228	0.665	4.6	-	-	-	-
Rv2068c	blaC	class A BETA-lactamase	0.0746	0.811	35.0	X	X	-	X
Rv2070c	cobK	cobalt-precorrin-6x reductase	2.3339	0.005	2.2	-	-	-	-
Rv2071c	cobM	precorrin-4 C11-methyltransferase	1.302	0.022	2.4	-	-	-	-
Rv2072c	cobL	precorrin-6y methyltransferase	-1.2264	0.168	4.4	-	-	-	-
Rv2073c		shortchain dehydrogenase	-0.7055	0.57	12.9	-	-	-	-
Rv2074		hypothetical protein	0.4551	0.017	10.6	-	-	-	-
Rv2080	lppJ	lipoprotein	-1.044	0.122	2.2	X	-	X	X
Rv2082		hypothetical protein	-0.2737	0.531	9.7	-	-	-	-
Rv2084		hypothetical protein	0.0048		0.7	-	-	-	-
Rv2088	pknJ	transmembrane serine/threonine-protein kinase J	0.6042	0.299	14.3	-	-	-	-
Rv2091c		hypothetical protein	-0.6344	0.026	69.6	-	-	X	-
Rv2093c	tatC	Sec-independent protein translocase	0.1136	0.541	1.4	-	-	X	-
Rv2094c	tatA	twin arginine translocase protein A	0.3761	0.443	7.7	-	-	X	-
Rv2097c		hypothetical protein	-0.4078	0.19	13.9	-	-	-	-
Rv2101	helZ	helicase	0.2432	0.525	36.4	-	-	-	-
Rv2102		hypothetical protein	0.0594	0.921	6.0	-	-	-	-
Rv2110c	prcB	proteasome (beta subunit)	0.0522	0.932	1.0	-	-	-	-
Rv2112c		hypothetical protein	-0.513	0.187	6.8	-	-	-	-
Rv2113		integral membrane protein	-0.7273	0.101	23.1	-	-	X	-
Rv2114		hypothetical protein	-0.0314	0.946	1.8	-	-	-	-
Rv2115c		ATPase	1.302	0.085	3.2	-	-	-	-
Rv2116	lppK	lipoprotein	0.034	0.877	10.9	-	-	-	X
Rv2118c		RNA methyltransferase	1.3526	0.024	5.2	-	-	-	-

Rv2120c		integral membrane protein	0.5148	0.562	1.5	-	-	X	-
Rv2124c	metH	5-methyltetrahydrofolate-homocystein methyltransferase	0.5996	0.236	18.8	-	-	-	-
Rv2127	ansP1	L-asparagine permease	-0.006	0.991	10.7	-	-	X	-
Rv2129c		short chain dehydrogenase	0.5487	0.016	70.9	-	-	-	-
Rv2131c	cysQ	monophosphatase	0.7772	0.565	9.0	-	-	-	-
Rv2137c		hypothetical protein	0.476	0.382	15.2	-	-	-	-
Rv2138	lppL	lipoprotein	-0.3479	0.202	6.5	X	-	X	X
Rv2139	pyrD	dihydroorotate dehydrogenase 2	-0.1678	0.226	12.5	-	-	-	-
Rv2145c	wag31	hypothetical protein	-0.0072	0.956	205.3	-	-	-	-
Rv2146c		transmembrane protein	1.0252	0.105	5.5	-	-	X	-
Rv2151c	ftsQ	cell division protein	-0.2456	0.547	25.0	-	-	X	-
Rv2153c	murG	undecaprenyldiphospho-muramoylpentapeptide beta-N-acetylglucosaminyltransferase	0.038	0.924	15.8	-	-	-	-
Rv2155c	murD	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase	-0.1386	0.88	6.9	-	-	-	-
Rv2156c	mraY	phospho-N-acetylmuramoyl-pentapeptide-transferase	0.8407	0.054	0.7	-	-	X	-
Rv2159c		hypothetical protein	-0.5178	0.221	351.7	-	-	-	-
Rv2161c		hypothetical protein	0.264	0.376	27.4	-	-	-	-
Rv2171	lppM	lipoprotein	-0.9928	0.329	9.1	X	-	X	X
Rv2176	pknL	transmembrane serine/threonine-protein kinase L	-0.3958	0.511	5.8	-	-	X	-
Rv2179c		hypothetical protein	-0.1697	0.804	3.5	-	-	-	-
Rv2182c		1-acylglycerol-3-phosphate O-acyltransferase	0.1576	0.587	11.7	-	-	-	-
Rv2183c		hypothetical protein	-0.802	0.015	2.7	-	-	-	-
Rv2184c		hypothetical protein	0.571	0.256	8.7	-	-	-	-
Rv2185c	TB16.3	hypothetical protein	-0.2403	0.72	17.8	-	-	-	-
Rv2186c		hypothetical protein	-0.0004	0.999	4.2	-	-	-	-
Rv2187	fadD15	long-chain-fatty-acid-CoA ligase	0.0846	0.694	173.9	-	-	-	-
Rv2188c		hypothetical protein	0.0428	0.933	4.2	-	-	-	-
Rv2190c		hypothetical protein	0.7872	0.008	9.2	X	-	X	-
Rv2192c	trpD	anthranilate phosphoribosyltransferase	0.7	0.386	3.3	-	-	-	-
Rv2193	ctaE	cytochrome C oxidase subunit III	0.107	0.765	3.3	-	-	X	-
Rv2194	qcrC	ubiquinol-cytochrome C reductase QcrC(cytochrome C subunit)	0.4444	0.018	47.7	X	-	X	-
Rv2195	qcrA	Rieske iron-sulfur protein	-0.4236	0.116	166.9	-	-	X	-
Rv2196	qcrB	ubiquinol-cytochrome C reductase QcrB (cytochrome B subunit)	0.0176	0.871	133.3	-	-	X	-
Rv2197c		transmembrane protein	-0.7501	0.128	7.5	-	-	X	-
Rv2198c	mmpS3	membrane protein	-0.3684	0.067	22.5	-	-	X	-
Rv2200c	ctaC	transmembrane cytochrome C oxidase subunit II	-0.6208	0.009	123.6	-	-	X	-
Rv2201	asnB	asparagine synthetase	-0.6016	0.057	21.1	-	-	-	-
Rv2202c	cbhK	carbohydrate kinase	0.959	0.252	8.3	-	-	-	-
Rv2203		hypothetical protein	0.0444	0.925	28.4	-	-	X	-
Rv2204c		hypothetical protein	0.4156	0.029	19.0	-	-	-	-
Rv2206		transmembrane protein	-0.6378	0.339	16.7	-	-	X	-
Rv2210c	ilvE	branched-chain amino acid aminotransferase	0.302	0.381	1.9	-	-	-	-
Rv2212		hypothetical protein	1.4533	0.202	1.3	-	-	-	-
Rv2213	pepB	leucyl aminopeptidase	-0.364	0.321	244.8	-	-	-	-
Rv2214c	ephD	short chain dehydrogenase	-0.3788	0.074	36.9	-	-	X	-
Rv2215	dlaT	dihydrolipoamide acetyltransferase	0.8504	0.003	76.3	-	-	-	-
Rv2216		hypothetical protein	0.0337	0.883	22.1	-	-	-	-
Rv2219		transmembrane protein	-0.6598	0.502	26.6	-	-	X	-
Rv2219A		hypothetical protein	0.2378	0.385	6.2	-	-	X	-
Rv2220	glnA1	glutamine synthetase	-0.0314	0.913	17.4	-	-	-	-

Rv2221c	glnE	glutamate-ammonia-ligase adenylyltransferase	-0.6164	0.537	3.2	-	-	-	-
Rv2222c	glnA2	glutamine synthetase	0.1636	0.649	2.9	-	-	-	-
Rv2224c	caeA	exported protease	0.0632	0.804	36.1	-	-	X	X
Rv2226		hypothetical protein	0.4635	0.078	11.6	-	-	-	-
Rv2232		hypothetical protein	-0.0221	0.968	2.9	-	-	-	-
Rv2234	ptpA	phosphotyrosine protein phosphatase (protein-tyrosine-phosphatase) (PTPase) (LMW phosphatase)	0.6353	0.309	2.6	-	-	-	-
Rv2235		transmembrane protein	-0.1474	0.774	8.8	-	-	X	-
Rv2237		hypothetical protein	0.4328	0.697	11.5	-	-	-	-
Rv2240c		hypothetical protein	-0.9955	0.015	6.8	-	-	-	-
Rv2241	aceE	pyruvate dehydrogenase subunit E1	1.0948	0.189	10.3	-	-	-	-
Rv2242		hypothetical protein	-0.006	0.989	23.6	-	-	-	-
Rv2243	fabD	acyl-carrier-protein S-malonyltransferase	-1.8061	0.072	1.4	-	-	-	-
Rv2244	acpP	acyl carrier protein	-0.2532	0.37	172.8	-	-	-	-
Rv2245	kasA	3-oxoacyl-(acyl carrier protein) synthase II	-0.3158	0.65	21.6	-	-	-	-
Rv2246	kasB	3-oxoacyl-(acyl carrier protein) synthase II	-0.1726	0.729	21.3	-	-	-	-
Rv2247	accD6	acetyl/propionyl-CoA carboxylase beta subunit	-0.8242	0.126	34.1	-	-	-	-
Rv2249c	glpD1	glycerol-3-phosphate dehydrogenase	-1.5188	0.102	9.9	-	-	-	-
Rv2250c		transcriptional regulator	0.5786	0.004	7.5	-	-	-	-
Rv2251		flavoprotein	1.5877	0.011	2.1	-	-	-	X
Rv2252		diacylglycerol kinase, Involved in synthesis of phosphatidylinositol mannosides	0.2914	0.567	19.5	-	-	-	-
Rv2258c		transcriptional regulator	-0.3186	0.266	17.0	-	-	-	-
Rv2259	adhE2	zinc-dependent alcohol dehydrogenase	0.2278	0.617	4.8	-	-	-	-
Rv2263		short chain dehydrogenase	-1.0726	0.139	17.1	-	-	-	-
Rv2272		transmembrane protein	1.0662	0.058	6.2	-	-	X	-
Rv2280		dehydrogenase	-1.0548	0.07	5.1	-	-	-	-
Rv2282c		LysR family transcriptional regulator	0.0836	0.88	11.8	-	-	-	-
Rv2284	lipM	esterase	-0.3014	0.549	10.7	-	-	X	-
Rv2285		hypothetical protein	-0.8544	0.03	24.2	-	-	-	-
Rv2286c		hypothetical protein	1.9097	0.004	1.2	-	-	-	-
Rv2287	yjcE	integral membrane transport protein	-0.7743	0.074	3.7	-	-	X	-
Rv2289	cdh	CDP-diacylglycerol pyrophosphatase	-0.567	0.035	55.7	X	-	X	-
Rv2290	lppO	lipoprotein	-0.8726	0.138	9.4	X	-	-	X
Rv2291	sseB	thiosulfate sulfurtransferase SseB	-0.8023	0.195	2.1	-	-	-	-
Rv2294		aminotransferase	-1.1163	0.108	10.2	-	-	-	-
Rv2296		haloalkane dehalogenase	0.5966	0.094	103.4	-	-	-	-
Rv2297		hypothetical protein	0.1313	0.51	14.6	-	-	-	-
Rv2298		hypothetical protein	-0.1716	0.384	55.1	-	-	-	-
Rv2299c	htpG	heat shock protein 90	0.8264	0.035	10.5	-	-	-	-
Rv2301	cut2	cutinase CUT2	0.0874	0.73	4.7	X	X	X	-
Rv2302		hypothetical protein	0.4434	0.728	1.3	-	-	-	-
Rv2303c		antibiotic-resistance protein	-0.1357	0.829	2.8	-	-	-	-
Rv2305		hypothetical protein	-0.1435	0.67	13.6	-	-	-	-
Rv2307c		hypothetical protein	0.8983	0.259	2.2	-	-	X	-
Rv2324		AsnC family transcriptional regulator	0.7486	0.436	5.4	-	-	-	-
Rv2325c		hypothetical protein	0.0148	0.982	2.4	-	-	X	-
Rv2326c		transmembrane ATP-binding protein ABC transorter	0.2938	0.454	44.6	-	-	X	-
Rv2328	PE23	PE family protein	-1.1644	0.006	3.9	-	-	-	-
Rv2330c	lppP	lipoprotein	0.4009	0.424	2.8	X	-	X	X
Rv2332	mez	malate dehydrogenase	0.5046	0.152	10.0	-	-	-	-
Rv2334	cysK1	cysteine synthase A	-0.8864	0.029	1.1	-	-	-	-
Rv2336		hypothetical protein	-0.8693	0.23	7.9	-	-	-	-
Rv2337c		hypothetical protein	1.107	0.001	5.1	-	-	-	-
Rv2338c	moeW	hypothetical protein	-0.2856	0.303	17.6	-	-	-	-
Rv2339	mmpL9	transmembrane transport protein MmpL9	0.6054	0.08	8.1	-	-	X	-
Rv2342		hypothetical protein	1.302	0.101	4.1	-	-	-	-

Rv2343c	dnaG	DNA primase	-0.1076	0.665	4.4	-	-	-	-
Rv2345		transmembrane protein	-0.2659	0.136	130.7	X	-	X	-
Rv2346c	esxO	putative ESAT-6 like protein	-0.9952	-	0.5	-	-	-	-
Rv2350c	plcB	membrane-associated phospholipase C	0.3586	0.568	7.9	X	X	-	-
Rv2351c	plcA	membrane-associated phospholipase C	-2.1356	0.008	3.2	-	X	-	-
Rv2357c	glyS	glycyl-tRNA synthetase	-0.4664	0.084	17.6	-	-	-	-
Rv2360c		hypothetical protein	-0.526	0.417	2.6	-	-	-	-
Rv2363	amiA2	amidase	-0.8004	0.01	24.7	-	-	-	-
Rv2364c	era	GTP-binding protein	-0.0495	0.945	1.1	-	-	-	-
Rv2366c		transmembrane protein	-0.1387	0.781	7.3	-	-	X	-
Rv2370c		hypothetical protein	-0.283	0.579	7.0	-	-	-	-
Rv2373c	dnaJ2	chaperone protein	0.1089	0.578	20.4	-	-	-	-
Rv2374c	hrcA	heat-inducible transcription repressor	-0.0609	0.458	74.0	-	-	-	-
Rv2375		hypothetical protein	0.887	0.028	3.1	-	-	-	-
Rv2376c	cfp2	low molecular weight antigen 2	0.274	0.786	2.5	X	-	X	-
Rv2383c	mbtB	phenyloxazoline synthase MBTB (phenyloxazoline synthetase)	1.0184	0.157	11.5	-	-	-	-
Rv2385	mbtJ	putative acetyl hydrolase	0.0633	0.885	2.4	-	-	-	-
Rv2387		hypothetical protein	-0.4672	0.205	7.9	-	-	X	-
Rv2388c	hemN	coproporphyrinogen III oxidase	-0.1379	0.414	0.5	-	-	-	-
Rv2392	cysH	phosphoadenosine phosphosulfate reductase	0.6884	0.23	2.4	-	-	-	-
Rv2394	ggtB	gamma-glutamyltranspeptidase precursor	-0.4633	0.074	25.0	X	X	-	X
Rv2397c	cysA1	sulfate-transport ATP-binding protein ABC transporter	1.0596	0.004	3.8	-	-	-	-
Rv2399c	cysT	sulfate-transport integral membrane protein ABC transporter	0.2996	0.427	1.7	-	-	X	-
Rv2400c	subI	sulfate-binding lipoprotein	-2.2542	0.048	2.6	-	-	-	X
Rv2402		hypothetical protein	0.6329	0.532	1.6	-	-	-	-
Rv2403c	lppR	lipoprotein	0.5377	0.681	2.1	X	-	X	X
Rv2404c	lepA	GTP-binding protein	1.218	0.084	3.5	-	-	-	-
Rv2405		hypothetical protein	0.8242	0.085	15.2	-	-	-	-
Rv2406c		hypothetical protein	1.0252	0.123	5.8	-	-	-	-
Rv2409c		hypothetical protein	0.8674	0.036	9.3	-	-	-	-
Rv2410c		hypothetical protein	1.6886	0.119	7.3	-	-	-	-
Rv2411c		hypothetical protein	1.9436	0.055	6.8	-	-	-	-
Rv2412	rpsT	30S ribosomal protein S20	0.3526	0.521	5.4	-	-	-	-
Rv2413c		hypothetical protein	-1.2806	0.027	15.1	-	-	-	-
Rv2418c		hypothetical protein	-0.7056	0.063	7.0	-	-	-	-
Rv2420c		hypothetical protein	-0.1379	0.414	0.5	-	-	-	-
Rv2425c		hypothetical protein	-1.1622	0.315	5.9	-	-	-	-
Rv2426c		hypothetical protein	-0.1634	0.597	15.8	-	-	-	-
Rv2427c	proA	gamma-glutamyl phosphate reductase	-0.9436	0.39	4.9	-	-	-	-
Rv2429	ahpD	alkyl hydroperoxide reductase subunit D	-0.8054	0.146	8.1	-	-	-	-
Rv2441c	rpmA	50S ribosomal protein L27	1.0252	0.058	9.8	-	-	-	-
Rv2442c	rplU	50S ribosomal protein L21	0.5388	0.382	8.5	-	-	-	-
Rv2444c	rne	ribonuclease E	1.2378	0.018	6.1	-	-	-	-
Rv2445c	ndk	nucleoside diphosphate kinase	1.8303	0.049	4.1	-	-	-	-
Rv2447c	folC	folylpolyglutamate synthase protein	-1.0903	0.012	1.2	-	-	-	-
Rv2449c		hypothetical protein	-0.3092	0.28	34.2	-	-	-	-
Rv2453c	mobA	molybdopterin-guanine dinucleotide biosynthesis protein A	0.7618	0.054	10.8	-	-	-	-
Rv2456c		integral membrane transport protein	-0.4045	0.244	4.5	-	-	X	-
Rv2457c	clpX	ATP-dependent protease ATP-binding subunit	-0.0946	0.901	16.2	-	-	-	-
Rv2460c	clpP2	ATP-dependent Clp protease proteolytic subunit	0.4684	0.05	28.7	-	-	-	-
Rv2461c	clpP	ATP-dependent Clp protease proteolytic subunit	0.692	0.009	7.9	-	-	-	-
Rv2462c	tig	trigger factor	1.0596	0.004	1.3	-	-	-	-
Rv2463	lipP	esterase/lipase	0.097	0.607	8.2	-	-	-	-
Rv2464c		DNA glycosylase	1.012	0.025	1.1	-	-	-	-
Rv2467	pepN	aminopeptidase N	-1.0868	0.018	4.2	-	-	-	-
Rv2468c		hypothetical protein	-0.7128	0.117	21.4	-	-	-	-

Rv2470	glbO	globin	0.862	0.077	2.1	-	-	-	-
Rv2473		alanine and proline rich membrane protein	-0.9497	0.039	8.2	X	-	X	-
Rv2475c		hypothetical protein	0.3212	0.716	1.4	-	-	-	-
Rv2476c	gdh	NAD-dependent glutamate dehydrogenase	0.806	0.021	80.4	-	-	-	-
Rv2477c		putative ABC transporter ATP-binding protein	0.0814	0.827	32.1	-	-	-	-
Rv2482c	plsB2	glycerol-3-phosphate acyltransferase	-0.3362	0.353	33.6	-	-	-	-
Rv2483c	plsC	bifunctionnal putative L-3-phosphoserine phosphatase/1-acyl-SN-glycerol-3-phosphate acyltransferase	0.7386	0.026	42.2	-	-	-	-
Rv2484c		hypothetical protein	0.2688	0.383	10.3	-	-	-	-
Rv2485c	lipQ	carboxylesterase	-0.6806	0.39	5.8	-	-	-	-
Rv2488c		LuxR family transcriptional regulator	0.571	0.591	5.4	-	-	-	-
Rv2495c	pdhC	branched-chain alpha-keto acid dehydrogenase subunit E2	1.8238	0.033	2.4	-	-	-	-
Rv2496c	pdhB	pyruvate dehydrogenase E1 component beta subunit	-0.0486	0.882	1.7	-	-	-	-
Rv2497c	pdhA	pyruvate dehydrogenase E1 component alpha subunit	0.6786		0.5	-	-	-	-
Rv2498c	citE	citrate (Pro-3S)-lyase beta subunit	0.0964	0.299	4.8	-	-	-	-
Rv2500c	fadE19	acyl-CoA dehydrogenase	-0.6335	0.248	1.3	-	-	-	-
Rv2509		short-chain type dehydrogenase/reductase	0.4886	0.053	61.1	-	-	-	-
Rv2516c		hypothetical protein	-1.0658	0.009	11.6	-	-	-	-
Rv2518c	lppS	lipoprotein	-1.5564	0.094	8.6	-	-	X	X
Rv2520c		hypothetical protein	-0.2326	0.403	24.0	-	-	X	-
Rv2521	bcp	bacterioferritin comigratory protein	1.1234	-	1.2	-	-	-	-
Rv2524c	fas	fatty acid synthase	0.3092	0.013	384.6	-	-	-	-
Rv2526		hypothetical protein	-1.9952	-	0.8	-	-	-	-
Rv2527		hypothetical protein	0.5302	0.432	6.7	-	-	-	-
Rv2529		hypothetical protein	0.5272	0.695	1.2	X	-	-	-
Rv2531c		amino acid decarboxylase	1.1306	0.042	10.5	-	-	-	-
Rv2532c		hypothetical protein	-1.1932	0.161	6.4	X	-	X	-
Rv2533c	nusB	transcription antitermination protein NusB	0.393	0.376	1.6	-	-	-	-
Rv2534c	efp	elongation factor P	2.2588	0.014	2.6	-	-	-	-
Rv2536		transmembrane protein	-0.4196	0.182	135.9	-	-	X	-
Rv2553c		hypothetical protein	-0.5794	0.105	12.3	-	-	X	-
Rv2554c		Holliday junction resolvase-like protein	0.862	0.077	0.7	-	-	-	-
Rv2555c	alaS	alanyl-tRNA synthetase	0.6818	0.039	18.5	-	-	-	-
Rv2556c		hypothetical protein	-0.3647	0.567	2.1	-	-	-	-
Rv2559c	rarA	recombination factor protein	-0.6953	0.03	12.8	-	-	-	-
Rv2560		hypothetical protein	-0.2835	0.658	2.7	-	-	X	-
Rv2563		glutamine-transport transmembrane protein ABC transporter	-0.3675	0.009	92.0	-	-	X	-
Rv2564	glnQ	glutamine-transport ATP-binding protein ABC transporter	0.0617	0.724	22.3	-	-	-	-
Rv2565		hypothetical protein	-0.4116	0.069	40.7	-	-	-	-
Rv2566		hypothetical protein	0.4162	0.266	30.5	-	-	-	-
Rv2567		hypothetical protein	0.6068	0.513	5.0	-	-	-	-
Rv2571c		transmembrane alanine and valine and leucine rich protein	-0.1718	0.73	7.1	-	-	X	-
Rv2574		hypothetical protein	-0.7068	0.152	8.5	-	-	-	-
Rv2575		hypothetical protein	1.302	0.071	3.5	-	X	X	-
Rv2581c		glyoxalase II	-0.5656	0.444	11.9	-	-	-	-
Rv2582	ppiB	peptidyl-prolyl cis-trans isomerase B	-1.5034	0.041	26.1	-	-	X	-
Rv2583c	relA	GTP pyrophosphokinase	-0.232	0.552	15.1	-	-	-	-
Rv2584c	apt	adenine phosphoribosyltransferase	0.2931	0.064	69.5	-	-	-	-
Rv2585c		lipoprotein	-0.8512	0.148	13.5	X	-	-	X
Rv2586c	secF	preprotein translocase subunit SecF	-0.602	0.002	55.6	-	-	X	-
Rv2587c	secD	preprotein translocase subunit SecD	0.0349	0.857	31.1	-	-	X	-
Rv2588c	yajC	preprotein translocase subunit YajC	0.1811	0.677	21.4	-	-	X	-
Rv2594c	ruvC	Holliday junction resolvase	0.7661	0.145	5.4	-	-	-	-
Rv2597		hypothetical protein	1.04	0.294	8.4	-	-	X	-

Rv2599		hypothetical protein	-2.3303	0.002	3.2	X	-	X	-
Rv2604c		glutamine amidotransferase subunit PdxT	1.0252	0.005	3.3	-	-	-	-
Rv2606c		pyridoxal biosynthesis lyase PdxS	-0.0102	0.975	6.8	-	-	-	-
Rv2609c		hypothetical protein	-0.56	0.091	21.0	-	-	X	-
Rv2610c	pimA	alpha-mannosyltransferase	0.5898		0.8	-	-	-	-
Rv2611c		lipid A biosynthesis lauroyl acyltransferase	-1.3215	0.127	8.0	-	-	-	-
Rv2612c	pgsA1	CDP-diacylglycerol--inositol 3-phosphatidyltransferase	0.2696	0.493	10.3	-	-	X	-
Rv2616		hypothetical protein	0.2378	0.49	8.7	-	-	-	-
Rv2617c		hypothetical protein	-1.6763	0.059	3.9	-	-	X	-
Rv2619c		hypothetical protein	0.5272	0.695	1.2	-	-	-	-
Rv2620c		transmembrane protein	0.5919	0.474	3.0	-	-	X	-
Rv2622		methyltransferase (methylase)	0.8686	0.024	27.9	-	-	-	-
Rv2623	TB31.7	hypothetical protein	3.7	0	7.6	-	-	-	-
Rv2625c		hypothetical protein	3.202	0.019	9.8	-	-	X	-
Rv2626c		hypothetical protein	1.6694	0.04	1.9	-	-	-	-
Rv2627c		hypothetical protein	4.0448	0.001	12.0	-	-	-	-
Rv2632c		hypothetical protein	0.3112	0.587	2.6	-	-	-	-
Rv2633c		hypothetical protein	0.148	0.337	28.2	-	-	-	-
Rv2637	dedA	transmembrane protein DedA	-0.43	0.329	9.7	-	-	X	-
Rv2638		hypothetical protein	0.198	0.746	1.8	-	-	-	-
Rv2643	arsC	arsenic-transport integral membrane protein	1.3238	0.124	2.2	-	-	X	-
Rv2652c		phiRv2 prophage protein	0.9638	0.047	0.7	-	-	-	-
Rv2658c		prophage protein	0.2709	0.664	3.2	-	-	-	-
Rv2672		secreted protease	-0.2942	0.406	25.2	X	-	X	X
Rv2673		integral membrane protein	-0.4608	0.035	18.5	-	-	X	-
Rv2676c		hypothetical protein	1.1266	0.004	5.7	-	-	-	-
Rv2682c	dxs1	1-deoxy-D-xylulose-5-phosphate synthase	-0.0254	0.947	25.6	-	-	-	-
Rv2683		hypothetical protein	1.978	0.044	3.2	-	-	-	-
Rv2688c		antibiotic ABC transporter ATP-binding protein	-2.6338	0.001	4.8	-	-	-	-
Rv2690c		hypothetical protein	-0.2414	0.579	12.3	-	-	X	-
Rv2691	ceoB	TRK system potassium uptake protein	-0.1696	0.782	5.0	-	-	-	-
Rv2692	ceoC	TRK system potassium uptake protein	-1.7229	0.039	1.3	-	-	-	-
Rv2693c		integral membrane alanine and leucine rich protein	1.1137	0.019	1.1	-	-	X	-
Rv2694c		hypothetical protein	0.6353	0.012	7.7	-	-	-	-
Rv2696c		hypothetical protein	0.0252	0.899	4.7	-	-	-	-
Rv2697c	dut	deoxyuridine 5'-triphosphate nucleotidohydrolase	-0.8587	0.114	2.9	-	-	-	-
Rv2698		alanine rich transmembrane protein	-0.1152	0.748	4.9	-	-	X	-
Rv2700		secreted alanine rich protein	-1.3106	0.032	10.6	-	-	X	-
Rv2702	ppgK	polyphosphate glucokinase PPGK (polyphosphate-glucose phosphotransferase)	0.2544	0.503	4.2	-	-	-	-
Rv2703	sigA	RNA polymerase sigma factor	1.4092	0.123	6.9	-	-	-	-
Rv2713	sthA	soluble pyridine nucleotide transhydrogenase	0.1914	0.483	30.9	-	-	-	-
Rv2715		hydrolase	-0.2794	0.086	29.8	-	-	-	-
Rv2718c	nrdR	transcriptional regulator	0.6588	0.115	21.3	-	-	-	-
Rv2721c		hypothetical protein	-1.2432	0.009	106.7	X	-	X	-
Rv2724c	fadE20	acyl-CoA dehydrogenase	-0.6414	0.506	4.2	-	-	-	-
Rv2725c	hflX	GTP-binding protein	0.1285	0.902	1.9	-	-	-	-
Rv2726c	dapF	diaminopimelate epimerase	-0.5507	0.087	3.0	-	-	-	-
Rv2728c		hypothetical protein	0.2453	0.494	3.1	X	-	-	-
Rv2732c		transmembrane protein	0.0002	1	13.7	-	-	X	-
Rv2734		hypothetical protein	0.6352	0.148	1.8	-	-	-	-
Rv2736c	recX	recombination regulator	1.0596	0.004	1.3	-	-	-	-
Rv2737c	recA	DNA recombination protein	0.5886	0.334	5.4	-	-	-	-
Rv2739c		alanine rich transferase	-0.8152	0.261	5.6	-	-	-	-
Rv2740		hypothetical protein	0.6461	0	44.8	-	-	-	-
Rv2743c		hypothetical protein	-0.2412	0.756	8.1	-	-	X	-
Rv2744c	35kd_ag	hypothetical protein	0.4848	0.162	117.3	-	-	-	-

Rv2745c		transcriptional regulator	0.1136	0.541	0.7	-	-	-	-
Rv2746c	pgsA3	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase	-0.613	0.019	11.5	-	-	X	-
Rv2748c	ftsK	cell division transmembrane protein	3.08	0.007	7.5	-	-	X	-
Rv2749		hypothetical protein	0.2431	0.781	1.2	-	-	-	-
Rv2751		hypothetical protein	0.2326	0.451	27.5	-	-	-	-
Rv2756c	hsdM	type I restriction/modification system DNA methylase	-0.1152	0.923	7.6	-	-	-	-
Rv2766c	fabG	3-ketoacyl-(acyl-carrier-protein) reductase	1.5313	0.039	4.6	-	-	-	-
Rv2772c		transmembrane protein	0.1345	0.46	17.6	-	-	X	-
Rv2773c	dapB	dihydrodipicolinate reductase	1.1548	0.141	1.8	-	-	-	-
Rv2777c		hypothetical protein	1.7263	0.03	2.2	-	-	-	-
Rv2778c		hypothetical protein	0.5432	0.158	26.9	-	-	-	-
Rv2782c	pepR	zinc protease	-0.217	0.711	3.2	-	-	-	-
Rv2783c	gpsI	polynucleotide phosphorylase/polyadenylase	0.016	0.904	25.5	-	-	-	-
Rv2784c	lppU	lipoprotein	-1.0253	0.033	4.3	X	-	-	X
Rv2785c	rpsO	30S ribosomal protein S15	0.6114	0.117	23.1	-	-	-	-
Rv2794c		hypothetical protein	-1.1379	0.059	0.9	-	-	-	-
Rv2795c		hypothetical protein	0.2039	0.774	2.8	-	-	-	-
Rv2796c	lppV	lipoprotein	1.0252	0.06	5.3	-	-	-	X
Rv2799		hypothetical protein	-1.232	0.019	14.1	-	-	X	-
Rv2800		hydrolase	-0.607	0.207	2.5	-	-	-	-
Rv2803		hypothetical protein	-0.1379	0.414	0.5	-	-	-	-
Rv2809		hypothetical protein	0.4848	0.574	3.3	-	-	-	-
Rv2821c		hypothetical protein	1.3387	0.177	4.4	-	-	-	-
Rv2822c		hypothetical protein	-1.0646	0.034	9.2	-	-	-	-
Rv2823c		hypothetical protein	2.5487	0.018	1.3	-	-	-	-
Rv2824c		hypothetical protein	0.5412	0.358	5.5	-	-	-	-
Rv2837c		hypothetical protein	-0.9693	0.455	6.5	-	-	-	-
Rv2839c	infB	translation initiation factor IF-2	0.4748	0.296	19.1	-	-	-	-
Rv2842c		hypothetical protein	0.3216	0.178	7.3	-	-	-	-
Rv2843		hypothetical protein	-0.4572	0.05	17.3	X	-	X	X
Rv2844		hypothetical protein	-0.711	0.194	15.4	-	-	-	-
Rv2846c	efpA	integral membrane efflux protein EfpA	0.1136	0.541	0.7	-	-	X	-
Rv2848c	cobB	cobyrinic acid a,c-diamide synthase	0.4496	0.28	2.4	-	-	-	-
Rv2849c	cobO	cob(I)yrinic acid a,c-diamide adenosyltransferase	-0.5794	0.163	3.9	-	-	-	-
Rv2850c		magnesium chelatase	-0.0042	0.995	12.2	-	-	-	-
Rv2852c	mqo	malate:quinone oxidoreductase	-0.2526	0.517	7.4	-	-	-	-
Rv2854		hypothetical protein	0.3287	0.354	5.8	-	-	-	-
Rv2855	mtr	mycothione reductase	0.054	0.952	2.1	-	-	-	-
Rv2860c	glnA4	glutamine synthetase	-0.4236	0.538	0.9	-	-	-	-
Rv2861c	mapB	methionine aminopeptidase	-0.1354	0.718	2.0	-	-	-	-
Rv2862c		hypothetical protein	-1.4634	0.152	5.4	-	-	-	-
Rv2868c	ispG	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	0.9969	0.193	8.1	-	-	-	-
Rv2869c		transmembrane protein	-0.5489	0.121	8.2	-	-	X	-
Rv2873	mpt83	cell surface lipoprotein (lipoprotein P23)	-0.1796	0.556	8.1	X	-	-	X
Rv2883c	pyrH	uridylate kinase	1.6382	0.012	3.0	-	-	-	-
Rv2884		transcriptional regulator	0.6994	0.481	10.5	-	-	-	-
Rv2886c		resolvase	0.6352	0.431	3.7	-	-	-	-
Rv2888c	amiC	amidase	-0.2714	0.235	38.0	-	-	-	-
Rv2889c	tsf	elongation factor Ts	0.512	0.471	1.3	-	-	-	-
Rv2890c	rpsB	30S ribosomal protein S2	-0.3647	0.356	41.6	-	-	-	-
Rv2894c	xerC	site-specific tyrosine recombinase	1.0563	0.125	4.9	-	-	-	-
Rv2900c	fdhF	formate dehydrogenase H	0.2082	0.773	7.3	-	-	-	-
Rv2901c		hypothetical protein	0.4812	0.027	13.2	-	-	-	-
Rv2902c	rnhB	ribonuclease HII	0.862	0.077	0.7	-	-	-	-
Rv2903c	lepB	signal peptidase I	-0.499	0.554	11.2	-	-	X	-
Rv2904c	rplS	50S ribosomal protein L19	0.91	0.016	17.0	-	-	-	-
Rv2907c	rimM	16S rRNA-processing protein	0.6353	0.047	1.4	-	-	-	-
Rv2909c	rpsP	30S ribosomal protein S16	0.6042	0.045	12.6	-	-	-	-
Rv2910c		hypothetical protein	0.302	0.609	3.2	-	-	-	-

Rv2911	dacB2	D-alanyl-D-alanine carboxypeptidase	1.6452	0.01	6.6	X	-	-	-
Rv2913c		D-amino acid aminohydrolase	0.7244	0.168	12.9	-	-	-	-
Rv2914c	pknI	transmembrane serine/threonine-protein kinase I	0.6408	0.13	21.3	-	-	X	-
Rv2915c		hypothetical protein	1.4504	0.039	1.4	-	-	-	-
Rv2918c	glnD	PII uridylyl-transferase	0.3526	0.411	3.3	-	-	-	-
Rv2919c	glnB	nitrogen regulatory protein P-II	1.1168	0.006	5.6	-	-	-	-
Rv2921c	ftsY	cell division protein	-0.3002	0.152	41.9	-	-	X	-
Rv2922c	smc	chromosome partition protein Smc	0.0607	0.724	54.8	-	-	-	-
Rv2924c	fpg	formamidopyrimidine-DNA glycosylase	-1.1955	0.205	2.2	-	-	-	-
Rv2927c		hypothetical protein	2.7992	0.001	4.4	-	-	-	-
Rv2928	tesA	thioesterase TESA	1.1163	0.039	3.8	-	-	-	-
Rv2931	ppsA	phenolpthiocerol synthesis type-I polyketide synthase	0.683		0.7	-	-	-	-
Rv2932	ppsB	phenolpthiocerol synthesis type-I polyketide synthase PPSB	-2.0039	0.006	2.6	-	-	-	-
Rv2933	ppsC	phenolpthiocerol synthesis type-I polyketide synthase PPSC	-0.472	0.426	2.8	-	-	-	-
Rv2934	ppsD	phenolpthiocerol synthesis type-I polyketide synthase PPSD	0.3204	0.317	15.4	-	-	-	-
Rv2935	ppsE	phenolpthiocerol synthesis type-I polyketide synthase PPSE	-0.4842	0.021	14.8	-	-	-	-
Rv2936	drrA	daunorubicin-DIM-transport ATP-binding protein ABC transporter	-0.971	0.003	15.5	-	-	-	-
Rv2937	drrB	daunorubicin-DIM-transport integral membrane protein ABC transporter	-0.1379	0.414	1.1	-	-	X	-
Rv2938	drrC	daunorubicin-DIM-transport integral membrane protein ABC transporter	-0.4687	0.581	1.3	-	-	X	-
Rv2940c	mas	multifunctional mycocerosic acid synthase membrane-associated MAS	1.0815	0.011	70.4	-	-	-	-
Rv2941	fadD28	acyl-CoA synthetase	-0.32	0.752	22.1	-	-	-	-
Rv2942	mmpL7	transmembrane transport protein MmpL7	-0.7492	0.014	28.8	-	-	X	-
Rv2945c	lppX	lipoprotein	0.0556	0.648	38.9	X	-	X	X
Rv2946c	pks1	polyketide synthase	0.0684	0.871	11.6	-	-	-	-
Rv2947c	pks15	polyketide synthase	-0.7941	0.272	17.5	-	-	-	-
Rv2949c		hypothetical protein	0.7326	0.209	7.8	-	-	-	-
Rv2950c	fadD29	acyl-CoA synthetase	-0.4713	0.094	1.8	-	-	-	-
Rv2951c		oxidoreductase	0.1736	0.236	46.2	-	-	-	-
Rv2952		methyltransferase (methylase)	0.6225	0.048	34.4	-	-	-	-
Rv2953		hypothetical protein	-0.1027	0.779	37.8	-	-	-	-
Rv2954c		hypothetical protein	0.2006	0.745	9.6	-	-	-	-
Rv2955c		hypothetical protein	-0.1244	0.785	15.8	-	-	-	-
Rv2959c		methyltransferase (methylase)	0.727	0.026	13.3	-	-	-	-
Rv2962c		glycosyl transferase	-0.2452	0.213	12.2	-	-	-	-
Rv2967c	pca	pyruvate carboxylase	0.585	0.006	65.5	-	-	-	-
Rv2969c		hypothetical protein	-1.246	0.436	28.0	-	-	X	-
Rv2970c	lipN	lipase/esterase	-0.3442	0.463	23.3	-	-	-	-
Rv2971		oxidoreductase	1.271	0	3.7	-	-	-	-
Rv2975c		hypothetical protein	-1.0903	0.012	1.2	-	-	-	-
Rv2976c	ung	uracil-DNA glycosylase	1.302	0.107	3.3	-	-	-	-
Rv2977c	thiL	thiamine monophosphate kinase	0.0048		0.7	-	-	-	-
Rv2980		hypothetical protein	0.302	0.573	3.4	-	-	X	-
Rv2984	ppk	polyphosphate kinase	0.8605	0.076	1.2	-	-	-	-
Rv2986c	hupB	DNA-binding protein HU	0.626	0.056	335.2	-	-	-	-
Rv2987c	leuD	isopropylmalate isomerase small subunit	-0.088	0.812	14.8	-	-	-	-
Rv2988c	leuC	isopropylmalate isomerase large subunit	1.5764	0.112	1.4	-	-	-	-
Rv2989		transcriptional regulator	1.3268	0.002	9.5	-	-	-	-
Rv2994		integral membrane protein	-0.045	0.927	9.2	-	-	X	-
Rv2995c	leuB	3-isopropylmalate dehydrogenase	1.9376	0.029	2.9	-	-	-	-
Rv2996c	serA1	D-3-phosphoglycerate dehydrogenase	-0.246	0.614	6.3	-	-	-	-
Rv2997		alanine rich dehydrogenase	-0.5309	0.57	1.9	-	-	-	-
Rv2999	lppY	lipoprotein	-0.3189	0.154	12.7	X	-	-	X

Rv3001c	ilvC	ketol-acid reductoisomerase	-0.476	0.219	22.3	-	-	-	-
Rv3002c	ilvH	acetolactate synthase 3 regulatory subunit	-0.0144	0.96	11.7	-	-	-	-
Rv3003c	ilvB1	acetolactate synthase 1 catalytic subunit	0.8128	0.066	15.1	-	-	-	-
Rv3004	cfp6	low molecular weight protein antigen 6	0.2607	0.78	4.4	-	-	X	-
Rv3005c		hypothetical protein	-0.1814	0.595	15.2	-	-	X	-
Rv3006	lppZ	lipoprotein	-0.4463	0.395	34.0	X	-	-	X
Rv3007c		oxidoreductase	0.42		1.2	-	-	-	-
Rv3011c	gatA	aspartyl/glutamyl-tRNA amidotransferase subunit	-0.4999	0.042	3.5	-	-	-	-
Rv3016	lpqA	lipoprotein	0.4404	0.309	4.7	X	-	-	X
Rv3028c	fixB	electron transfer flavoprotein subunit alpha	0.9548	0.024	20.5	-	-	-	-
Rv3029c	fixA	electron transfer flavoprotein subunit beta	1.1501	0.121	7.5	-	-	-	-
Rv3033		hypothetical protein	0.1723	0.214	11.4	X	-	-	-
Rv3034c		transferase	0.2857	0.622	3.4	X	-	-	-
Rv3035		hypothetical protein	-1.1469	0.31	2.8	-	-	-	-
Rv3036c	TB22.2	hypothetical protein	1.0596	0.004	1.3	X	-	-	-
Rv3038c		hypothetical protein	-0.6704	0.184	17.0	-	-	-	-
Rv3040c		hypothetical protein	-0.1294	0.801	3.1	-	-	-	-
Rv3041c		ABC transporter ATP-binding protein	-0.9004	0.075	17.5	-	-	-	-
Rv3043c	ctaD	cytochrome C oxidase polypeptide I	0.061	0.754	33.6	-	-	X	-
Rv3044	fecB	FEIII-dicitrate-binding periplasmic lipoprotein	-0.8278	0.136	22.2	X	-	-	X
Rv3046c		hypothetical protein	-0.7569	0.427	1.1	-	-	-	-
Rv3049c		monooxygenase	-0.0026	0.991	27.3	-	-	-	-
Rv3050c		AsnC family transcriptional regulator	0.0532	0.797	0.7	-	-	X	-
Rv3051c	nrdE	ribonucleotide-diphosphate reductase subunit	1.4303	0.008	19.4	-	-	-	-
Rv3052c	nrdI	ribonucleotide reductase stimulatory protein	1.3316	0.002	18.2	-	-	-	-
Rv3053c	nrdH	glutaredoxin electron transport protein NrdH	-0.1379	0.414	0.5	-	-	-	-
Rv3057c		short chain dehydrogenase	-0.0914	0.865	17.2	-	-	-	-
Rv3058c		TetR family transcriptional regulator	0.3072	0.145	15.6	-	-	-	-
Rv3059	cyp136	cytochrome P450 136	-1.4278	0.018	26.0	-	-	-	-
Rv3062	ligB	ATP-dependent DNA ligase	-1.0887	0.003	4.7	-	-	-	-
Rv3066		DeoR family transcriptional regulator	-0.277	0.551	4.6	-	-	-	-
Rv3069	ccrB	camphor resistance protein CrcB	0.302	0.513	4.1	-	X	X	-
Rv3071		hypothetical protein	-0.0994	0.773	7.0	-	-	-	-
Rv3075c		hypothetical protein	0.9255	0.339	1.2	-	-	-	-
Rv3076		hypothetical protein	-0.2082	0.187	13.2	-	-	-	-
Rv3077		hydrolase	-1.6661	0.045	9.5	-	-	-	-
Rv3080c	pknK	serine/threonine- kinase transcriptional regulator	0.5997	0.685	4.2	-	-	-	-
Rv3081		hypothetical protein	0.2589	0.591	8.4	-	-	-	-
Rv3083		monooxygenase	-1.4806	0.001	12.1	-	-	-	-
Rv3084	lipR	acetyl-hydrolase/esterase	-0.9057	0.019	11.7	-	-	-	-
Rv3085		short-chain type dehydrogenase/reductase	0.6372	0.031	21.4	-	-	-	-
Rv3087		hypothetical protein	0.4924	0.153	14.2	-	-	-	-
Rv3088		hypothetical protein	-0.4674	0.052	34.7	-	-	-	-
Rv3090		hypothetical protein	-0.3662	0.363	49.5	-	-	X	-
Rv3091		hypothetical protein	0.0474	0.736	37.0	-	-	-	-
Rv3092c		integral membrane protein	0.0596	0.931	1.9	-	-	X	-
Rv3093c		oxidoreductase	-0.9952	-	0.5	-	-	-	-
Rv3099c		hypothetical protein	2.7391	0.024	1.6	-	-	-	-
Rv3101c	ftsX	putative cell division (septation integral membrane protein ABC transporter)	-0.5414	0.017	37.1	-	-	X	-
Rv3102c	ftsE	putative cell division septation component-transport ATP-binding protein ABC transporter)	-0.3242	0.698	33.2	-	-	-	-
Rv3104c		transmembrane protein	-0.1472	0.653	14.4	-	-	X	-
Rv3116	moeB2	molybdenum cofactor biosynthesis protein	-0.3922	0.658	7.4	-	-	-	-
Rv3119	moaE1	molybdenum cofactor biosynthesis protein E	1.4228	0.076	7.2	-	-	-	-
Rv3120		hypothetical protein	0.0192	0.938	11.5	-	-	-	-
Rv3130c	tgs1	triacylglycerol synthase	1.2116	0.033	19.9	-	-	-	-
Rv3132c	devS	two component sensor histidine kinase	0.7346	0.274	27.8	-	-	-	-
Rv3133c	devR	two component transcriptional regulatory protein	1.692	0.001	4.6	-	-	-	-

Rv3134c		hypothetical protein	1.8568	0.032	7.5	-	-	-	-
Rv3136	PPE51	PPE family protein	0.0564	0.926	8.3	-	-	-	-
Rv3139	fadE24	acyl-CoA dehydrogenase	0.1516	0.77	6.7	-	-	-	-
Rv3140	fadE23	acyl-CoA dehydrogenase	0.0058	0.981	14.9	-	-	-	-
Rv3141	fadB4	NADPH quinone oxidoreductase	0.9213	0.444	4.5	-	-	-	-
Rv3143		response regulator	0.2823	0.553	5.2	-	-	-	-
Rv3144c	PPE52	PPE family protein	-0.5309	0.57	1.9	-	-	-	-
Rv3145	nuoA	NADH dehydrogenase subunit A	0.4282	0.273	13.6	X	-	X	-
Rv3146	nuoB	NADH dehydrogenase subunit B	-0.0777	0.622	18.8	-	-	-	-
Rv3147	nuoC	NADH dehydrogenase subunit C	1.2836	0.031	15.3	-	-	-	-
Rv3148	nuoD	NADH dehydrogenase subunit D	0.3928	0.053	20.7	-	-	-	-
Rv3149	nuoE	NADH dehydrogenase subunit E	0.0564	0.859	14.1	-	-	-	-
Rv3150	nuoF	NADH dehydrogenase I chain F	-0.4154	0.304	13.7	-	-	-	-
Rv3151	nuoG	NADH dehydrogenase subunit G	0.0504	0.79	53.0	-	-	-	-
Rv3152	nuoH	NADH dehydrogenase subunit H	-0.5226	0.249	11.3	-	-	X	-
Rv3153	nuoI	NADH dehydrogenase subunit I	0.4202	0.442	9.2	-	-	-	-
Rv3154	nuoJ	NADH dehydrogenase subunit J	0.466	0.006	5.8	-	-	X	-
Rv3155	nuoK	NADH dehydrogenase subunit K	-0.547	0.079	12.5	-	-	X	-
Rv3156	nuoL	NADH dehydrogenase subunit L	-0.3279	0.368	27.2	-	-	X	-
Rv3157	nuoM	NADH dehydrogenase subunit M	0.245	0.314	49.9	-	-	X	-
Rv3158	nuoN	NADH dehydrogenase subunit N	-0.3958	0.072	40.3	-	-	X	-
Rv3161c		dioxygenase	1.7992	0.034	4.1	-	-	-	-
Rv3165c		hypothetical protein	0.302	0.456	7.4	-	-	X	-
Rv3166c		hypothetical protein	-0.3648	0.422	2.5	-	-	X	-
Rv3168		hypothetical protein	0.8303	0.229	6.6	-	-	-	-
Rv3170	aofH	flavin-containing monoamine oxidase	-0.0314	0.939	2.8	-	-	-	-
Rv3173c		TetR/ACRR family transcriptional regulator	0.1412	0.731	8.2	-	-	-	-
Rv3188		hypothetical protein	1.1137	0.019	1.1	-	-	-	-
Rv3190c		hypothetical protein	-1.082	0.257	9.7	-	-	-	-
Rv3193c		hypothetical protein	-0.3314	0.013	158.5	-	-	X	-
Rv3194c		hypothetical protein	-1.009	0.018	28.0	X	-	X	-
Rv3196A		hypothetical protein	-0.0903	0.524	1.4	-	-	-	-
Rv3197		ABC transporter ATP-binding protein	-0.2111	0.069	13.0	-	-	-	-
Rv3198A		glutaredoxin protein	-0.503	0.345	2.4	-	-	-	-
Rv3198c	uvrD2	ATP-dependent DNA helicase II	0.2998	0.226	11.6	-	-	-	-
Rv3199c	nudC	NADH pyrophosphatase	0.302	0.764	3.0	-	-	-	-
Rv3200c		transmembrane cation transporter	0.3548	0.007	38.5	-	-	X	-
Rv3202c		ATP-dependent DNA helicase	-0.58		0.8	-	-	-	-
Rv3204		DNA-methyltransferase (modification methylase)	0.4471	0.147	1.2	-	-	-	-
Rv3205c		hypothetical protein	0.7827	0.001	5.3	-	-	-	-
Rv3206c	moeB1	molybdopterin biosynthesis-like protein	-0.7085	0.468	3.5	-	-	X	-
Rv3207c		hypothetical protein	-0.3036	0.475	12.9	X	-	X	-
Rv3208A	TB9.4	hypothetical protein	-1.4236	0.132	1.6	-	-	-	-
Rv3211	rhlE	ATP-dependent RNA helicase RhlE	0.1266	0.804	8.1	-	-	-	-
Rv3212		hypothetical protein	-0.0894	0.721	11.5	X	-	X	-
Rv3213c		SOJ/PARA-like protein	-0.7544	0.392	22.5	-	-	-	-
Rv3215	entC	isochorismate synthase	0.0638	0.864	10.3	-	-	-	-
Rv3218		hypothetical protein	-0.2646	0.35	14.4	-	-	-	-
Rv3220c		two component sensor kinase	0.7116	0.293	7.2	-	-	-	-
Rv3224		short chain dehydrogenase	2.1636	0.022	4.4	-	-	-	-
Rv3225c		transferase	-0.1777	0.729	7.6	-	-	-	-
Rv3229c		linoleoyl-CoA desaturase	-0.186	0.676	3.1	-	-	-	-
Rv3230c		oxidoreductase	-0.3277	0.568	6.9	-	-	-	-
Rv3232c	pvdS	transcriptional regulator	0.0946	0.582	49.2	-	-	-	-
Rv3233c		hypothetical protein	-0.1697	0.603	5.4	-	-	-	-
Rv3234c		hypothetical protein	0.0192	0.958	7.9	-	-	-	-
Rv3236c		integral membrane transport protein	1.2164	0.005	3.6	-	-	X	-
Rv3237c		hypothetical protein	0.302	0.4	22.0	-	-	-	-
Rv3238c		integral membrane protein	0.7426	0.064	9.5	-	-	X	-
Rv3240c	secA1	preprotein translocase subunit SecA	-0.0558	0.709	56.3	-	-	-	-
Rv3241c		hypothetical protein	-0.5072	0.114	45.9	-	-	-	-

Rv3243c		hypothetical protein	-0.3539	0.023	11.1	-	-	-	-
Rv3244c	lpqB	lipoprotein	-0.3566	0.218	32.3	X	-	-	X
Rv3245c	mtrB	two component sensory histidine kinase	-0.4842	0.203	31.6	-	-	X	-
Rv3246c	mtrA	two component sensory transduction transcriptional regulatory protein	1.6682	0.2	10.0	-	-	-	-
Rv3248c	sahH	S-adenosyl-L-homocysteine hydrolase	0.2557	0.381	47.1	-	-	-	-
Rv3249c		TetR family transcriptional regulator	2.3168	0.016	6.1	-	-	-	-
Rv3253c		cationic amino acid transport protein	0.239	0.846	0.9	-	-	X	-
Rv3254		hypothetical protein	-0.5187	0.089	12.8	-	-	-	-
Rv3256c		hypothetical protein	0.8354	0.297	5.0	-	-	-	-
Rv3259		hypothetical protein	-0.3646	0.445	3.1	-	-	-	-
Rv3264c	manB	D-alpha-D-mannose-1-phosphate guanylyltransferase	0.2502	0.286	14.1	-	-	-	-
Rv3265c	wbbL1	dTDP-RHA:A-D-GlcNAc-diphosphoryl polyprenol	-0.0488	0.824	14.8	-	-	-	-
Rv3266c	rmlD	dTDP-6-deoxy-L-lyxo-4-hexulose reductase	-1.223	0.053	10.2	-	-	-	-
Rv3268		hypothetical protein	-0.7412	0.417	3.3	-	-	-	-
Rv3269		hypothetical protein	0.8702	0.028	42.3	-	-	-	-
Rv3270	ctpC	metal cation-transporting P-type ATPase C	0.3544	0.016	118.2	-	-	-	-
Rv3271c		integral membrane protein	1.6276	0.01	5.7	-	-	X	-
Rv3272		hypothetical protein	-0.1672	0.759	13.5	-	-	-	-
Rv3273		transmembrane carbonic anhydrase	-0.1814	0.01	101.0	-	-	X	-
Rv3274c	fadE25	acyl-CoA dehydrogenase	-0.123	0.819	7.9	-	-	-	-
Rv3275c	purE	phosphoribosylaminoimidazole carboxylase catalytic subunit	0.3112	0.587	2.6	-	-	-	-
Rv3276c	purK	phosphoribosylaminoimidazole carboxylase ATPase subunit	0.0048		0.7	X	-	-	-
Rv3277		transmembrane protein	-0.5794	0.316	4.0	-	-	X	-
Rv3278c		transmembrane protein	0.3008	0.049	20.2	-	-	X	-
Rv3280	accD5	propionyl-CoA carboxylase beta chain	2.0638	0.175	2.5	-	-	-	-
Rv3282	maf	Maf-like protein	0.0252	0.963	3.6	-	-	-	-
Rv3285	accA3	bifunctional acetyl-/propionyl-coenzyme A carboxylase subunit alpha	0.8874	0.225	18.0	-	-	-	-
Rv3288c	usfY	hypothetical protein	0.9515	0.17	0.9	-	-	X	-
Rv3289c		transmembrane protein	0.1136	0.541	0.7	-	-	X	-
Rv3291c		AsnC family transcriptional regulator	-0.7484	0.429	3.2	-	-	-	-
Rv3295		TetR family transcriptional regulator	2.571	0	3.8	-	-	-	-
Rv3296	lhr	ATP-dependent helicase	0.3618	0.48	13.3	-	-	-	-
Rv3297	nei	endonuclease VIII	1.302	0.008	1.9	-	-	-	-
Rv3298c	lpqC	esterase	-1.4236	0.132	3.1	-	-	-	X
Rv3299c	atsB	arylsulfatase	0.7129	0.045	18.9	-	-	-	-
Rv3301c	phoY1	phosphate transporter PhoU	-0.0949	0.816	25.6	-	-	-	-
Rv3302c	glpD2	glycerol-3-phosphate dehydrogenase	0.0514	0.844	33.9	-	-	-	-
Rv3303c	lpdA	flavoprotein disulfide reductase	-0.9174	0.023	29.4	-	-	-	-
Rv3311		hypothetical protein	0.7786	0.089	27.6	-	-	-	-
Rv3316	sdhC	succinate dehydrogenase cytochrome B-556 subunit	-1.044	0.168	2.3	-	-	X	-
Rv3317	sdhD	succinate dehydrogenase hydrophobic membrane anchor subunit	0.1231	0.496	13.4	-	-	X	-
Rv3318	sdhA	succinate dehydrogenase flavoprotein subunit	-0.224	0.266	47.2	-	-	-	-
Rv3319	sdhB	succinate dehydrogenase iron-sulfur subunit	-0.096	0.803	26.1	-	-	-	-
Rv3329		hypothetical protein	-0.1016	0.541	9.0	-	-	-	-
Rv3330	dacB1	penicillin-binding protein	-0.4404	0.204	1.6	X	-	X	-
Rv3331	sugI	sugar-transport integral membrane protein	0.3896	0.347	4.3	-	-	X	-
Rv3334		MerR family transcriptional regulator	-0.1338	0.574	8.7	-	-	-	-
Rv3335c		integral membrane protein	0.2453	0.348	3.9	-	-	X	-
Rv3338		hypothetical protein	0.85	0.058	4.9	-	-	-	-
Rv3359		oxidoreductase	-0.0534	0.977	2.5	-	-	-	-
Rv3362c		ATP/GTP-binding protein	-0.2178	0.191	17.4	-	-	-	-
Rv3363c		hypothetical protein	0.107	0.797	3.3	-	-	-	-
Rv3365c		hypothetical protein	-0.2214	0.047	127.	-	-	X	-

					1				
Rv3373	echA18	enoyl-CoA hydratase	-0.103	0.795	11.2	-	-	-	-
Rv3376		hypothetical protein	-0.429	0.201	8.7	-	-	-	-
Rv3377c		cyclase	0.2431	0.781	2.5	-	-	-	-
Rv3378c		hypothetical protein	0.8538	0.357	4.9	-	-	-	-
Rv3384c		hypothetical protein	-0.3648	0.561	2.6	-	-	-	-
Rv3389c	htdY	dehydrogenase	-0.0486	0.882	1.1	-	-	-	-
Rv3390	lpqD	lipoprotein	0.0188	0.913	82.4	X	-	X	X
Rv3391	acrA1	short chain dehydrogenase	0.31	0.637	10.2	-	-	-	-
Rv3392c	cmaA1	cyclopropane-fatty-acyl-phospholipid synthase 1	0		0.0	-	-	-	-
Rv3398c	idsA1	multifunctional dimethylallyltransferase/farnesyl diphosphate synthetase	-0.5053	0.051	2.6	-	-	-	-
Rv3401		hypothetical protein	1.6537	0.008	15.2	-	-	-	-
Rv3404c		hypothetical protein	1.477	0.242	4.4	-	-	-	-
Rv3405c		transcriptional regulator	-0.6164	0.242	6.4	-	-	-	-
Rv3406		dioxygenase	-1.261	0.053	1.8	-	-	-	-
Rv3407		hypothetical protein	0.4402	0.136	4.2	-	-	-	-
Rv3408		hypothetical protein	0.2181	0.78	8.9	-	-	-	-
Rv3409c	choD	cholesterol oxidase precursor	0.063	0.784	23.8	-	-	-	-
Rv3410c	guaB3	inosine 5-monophosphate dehydrogenase	0.1136	0.541	0.7	-	-	-	-
Rv3411c	guaB2	inosine 5'-monophosphate dehydrogenase	1.5698	0.099	13.9	-	-	-	-
Rv3417c	groEL1	chaperonin GroEL	-0.0179	0.939	115.8	-	-	-	-
Rv3418c	groES	co-chaperonin GroES	0.9476	0.002	25.1	-	-	-	-
Rv3419c	gcp	DNA-binding/iron metalloprotein/AP endonuclease	0.0252	0.98	6.8	-	-	-	-
Rv3422c		hypothetical protein	0.3586	0.438	4.0	-	X	-	-
Rv3423c	alr	alanine racemase	-0.172	0.843	2.3	-	-	-	-
Rv3435c		transmembrane protein	0.698	0.388	2.2	-	-	X	-
Rv3436c	glmS	glucosamine-fructose-6-phosphate aminotransferase	-0.2594	0.069	43.7	-	-	-	-
Rv3437		transmembrane protein	0.6353	0.339	1.5	-	-	X	-
Rv3438		hypothetical protein	0.4586	0.157	22.8	-	-	-	-
Rv3441c	glmM	phospho-sugar mutase / MRSA protein	0.3662	0.563	7.4	-	X	-	-
Rv3442c	rpsI	30S ribosomal protein S9	0.9984	0.116	11.7	-	-	-	-
Rv3443c	rplM	50S ribosomal protein L13	-0.2102	0.411	16.0	-	-	-	-
Rv3455c	truA	tRNA pseudouridine synthase A	-0.1379	0.414	1.1	-	-	-	-
Rv3456c	rplQ	50S ribosomal protein L17	-0.3646	0.715	8.8	-	-	-	-
Rv3457c	rpoA	DNA-directed RNA polymerase subunit alpha	-0.4164	0.594	6.1	-	-	-	-
Rv3458c	rpsD	30S ribosomal protein S4	0.8707	0.034	81.3	-	-	-	-
Rv3459c	rpsK	30S ribosomal protein S11	-0.1174	0.761	42.8	-	-	-	-
Rv3460c	rpsM	30S ribosomal protein S13	0.1398	0.581	18.2	-	-	-	-
Rv3462c	infA	translation initiation factor IF-1	-0.3844	0.381	8.2	-	-	-	-
Rv3464	rmlB	dTDP-glucose 4,6-dehydratase	0.1939	0.743	1.7	-	-	-	-
Rv3477	PE31	PE family protein	-0.0238	0.973	4.7	X	-	-	-
Rv3478	PPE60	PE family protein	-0.9364	0.003	32.3	-	-	-	-
Rv3479		hypothetical protein	0.1012	0.511	93.7	-	-	X	-
Rv3480c		hypothetical protein	-0.4562	0.097	26.1	-	-	-	-
Rv3481c		integral membrane protein	0.0522	0.932	1.0	-	-	X	-
Rv3482c		hypothetical protein	0.4152	0.411	9.5	-	-	X	-
Rv3483c		hypothetical protein	0.2054	0.496	10.7	-	-	X	-
Rv3485c		short chain dehydrogenase	0.048	0.979	1.7	-	-	-	-
Rv3486		hypothetical protein	0.6352	0.367	1.9	-	-	-	-
Rv3487c	lipF	esterase/lipase	-0.1527	0.667	14.2	-	-	-	-
Rv3490	otsA	alpha,alpha-trehalose-phosphate synthase	-0.2456	0.208	20.6	-	-	-	-
Rv3491		hypothetical protein	-0.051	0.819	4.9	X	-	X	-
Rv3492c		MCE4-associated	-0.6102	0.007	18.9	-	-	X	-
Rv3493c		MCE4-associated	-0.0276	0.941	21.9	-	-	X	-
Rv3494c	mce4F	MCE4-associated	-1.685	0.026	16.5	-	-	X	-
Rv3495c	lprN	MCE4-associated lipoprotein	-1.43440	0.002	28.3	X	-	-	X
Rv3496c	mce4D	MCE4-associated	-1.9556	0.023	19.3	-	-	X	-

Rv3497c	mce4C	MCE4-associated	-1.4807	0.002	43.5	-	-	X	-
Rv3498c	mce4B	MCE4-associated	-1.237	0.006	23.0	-	-	X	-
Rv3499c	mce4A	MCE4-associated	-1.0598	0.001	19.8	X	-	X	-
Rv3501c	yrbE4A	MCE4-associated	-1.3764	0.003	11.5	-	-	X	-
Rv3504	fadE26	acyl-CoA dehydrogenase	1.8532	0.093	2.8	-	-	-	-
Rv3509c	ilvX	hypothetical protein	-0.4154	0.462	8.8	-	-	-	-
Rv3516	echA19	enoyl-CoA hydratase	-0.1379	0.414	0.5	-	-	-	-
Rv3519		hypothetical protein	0.4348	0.479	24.2	-	-	-	-
Rv3523	ltp3	acetyl-CoA acetyltransferase	-0.7177	0.471	3.0	-	-	-	-
Rv3525c		siderophore-binding protein	-0.3646	0.659	1.8	-	-	-	-
Rv3528c		hypothetical protein	-0.3535	0.44	54.9	-	-	-	-
Rv3537		3-ketosteroid-delta-1-dehydrogenase	-0.0684	0.836	8.4	-	-	-	-
Rv3545c	cyp125	cytochrome P450 125	0.6445	0.018	2.2	-	-	-	-
Rv3547		hypothetical protein	-0.1242	0.638	43.5	-	-	-	-
Rv3553		oxidoreductase	-0.8919	0.102	2.0	-	-	-	-
Rv3554	fdxB	electron transfer protein FdxB	-0.5258	0.19	14.4	-	-	X	-
Rv3568c	bphC	biphenyl-2,3-diol 1,2-dioxygenase	-0.698	0.057	1.8	-	-	-	-
Rv3572		hypothetical protein	-1.6637	0.002	2.1	X	-	-	-
Rv3574		transcriptional regulatory protein TetR-family	-0.9672	0.196	5.1	-	-	-	-
Rv3575c		transcriptional regulatory protein LacI-family	1.2831	0.379	5.9	-	-	-	-
Rv3576	lppH	lipoprotein	0.0532	0.797	1.5	X	-	-	X
Rv3581c	ispF	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	-0.2048	0.849	1.7	-	-	-	-
Rv3582c	ispD	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase	-0.2304	0.852	11.4	-	-	-	-
Rv3584	lpqE	lipoprotein	0.4314	0.13	80.1	X	-	X	X
Rv3586	disA	DNA integrity scanning protein	0.1636	0.811	2.5	-	-	-	-
Rv3587c		hypothetical protein	0.3586	0.673	5.6	-	-	X	-
Rv3588c	canB	carbonic anhydrase,	-0.5598	0.427	2.2	-	-	-	-
Rv3589	mutY	adenine glycosylase	-0.503	0.024	12.7	-	-	-	-
Rv3592	TB11.2	hypothetical protein	1.8302	0.006	2.5	-	-	-	-
Rv3594		hypothetical protein	-0.5596	0.404	4.1	-	-	-	-
Rv3596c	clpC1	ATP-dependent protease ATP-binding subunit	0.2184	0.486	73.0	-	-	-	-
Rv3597c	lsr2	iron-regulated LSR2 protein precursor	-0.1848	0.716	10.7	-	-	-	-
Rv3598c	lysS	lysyl-tRNA synthetase	1.3864	-	1.4	-	-	-	-
Rv3600c		pantothenate kinase	1.0596	0.004	1.3	-	-	-	-
Rv3601c	panD	aspartate alpha-decarboxylase	0.7263	0.252	1.5	-	-	-	-
Rv3602c	panC	pantoate--beta-alanine ligase	1.7828	0.034	2.4	-	-	-	-
Rv3604c		transmembrane protein	-0.1854	0.309	38.9	-	-	X	-
Rv3606c	folK	2-amino-4-hydroxy-6-hydroxymethyldihydropteridi ne pyrophosphokinase	-1.0579	0.459	1.6	-	-	-	-
Rv3607c	folB	dihydroneopterin aldolase	0.6353	0.047	2.8	-	-	-	-
Rv3609c	folE	GTP cyclohydrolase I	-1.7412	0.036	4.8	-	-	-	-
Rv3610c	ftsH	membrane-bound protease (cell division protein)	0.6192	0.01	63.5	-	-	X	-
Rv3614c		hypothetical protein	0.0795	0.871	32.4	-	-	-	-
Rv3615c		hypothetical protein	0.6645	0.04	15.0	-	-	-	-
Rv3616c		hypothetical protein	-0.406	0.484	13.5	-	-	-	-
Rv3617	ephA	epoxide hydrolase	2.4916	0.015	2.4	-	-	-	-
Rv3623	lpqG	lipoprotein	-1.3507	0.216	25.0	X	-	-	X
Rv3626c		hypothetical protein	-0.3686	0.35	10.3	-	-	-	-
Rv3627c		hypothetical protein	-1.1495	0.072	20.3	X	-	X	-
Rv3631		transferase	1.0053	0.156	4.7	-	-	-	-
Rv3632		hypothetical protein	0.0596	0.689	1.8	-	-	X	-
Rv3633		hypothetical protein	0.3713	0.09	11.4	-	-	-	-
Rv3644c		DNA polymerase III subunit delta'	0.0266	0.943	11.4	-	-	-	-
Rv3645		transmembrane protein	-0.2994	0.359	11.4	-	-	X	-
Rv3646c	topA	DNA topoisomerase I	-0.2202	0.623	50.6	-	X	-	-
Rv3661		hypothetical protein	0.738	0.017	13.1	-	-	X	-
Rv3662c		hypothetical protein	-0.698	0.104	3.5	-	-	-	-
Rv3663c	dppD	peptide ABC transporter ATP-binding protein	-1.387	0.052	7.7	-	-	-	-
Rv3666c	dppA	periplasmic dipeptide-binding lipoprotein	-1.7659	0.045	5.7	X	-	X	X

Rv3667	acs	acetyl-CoA synthetase	-0.398	0.41	6.9	-	-	-	-
Rv3669		transmembrane protein	-0.1136	0.654	56.4	-	-	X	-
Rv3670	ephE	epoxide hydrolase	0.0759	0.883	6.1	-	-	-	-
Rv3671c		membrane-associated serine protease	-0.783	0.123	19.1	-	-	X	-
Rv3672c		hypothetical protein	0.3586	0.359	3.2	-	-	-	-
Rv3673c		membrane-anchored thioredoxin-like protein	-0.1054	0.415	7.2	-	-	X	-
Rv3674c	nth	endonuclease III	-0.1354	0.66	2.0	-	-	-	-
Rv3675		hypothetical protein	0.6353	0.339	3.0	-	-	X	-
Rv3676	crp	CRP/FNR family transcriptional regulator	2.2418	0.001	15.3	-	-	-	-
Rv3678c		hypothetical protein	-0.893	0.081	2.8	-	-	-	-
Rv3679		anion transporter ATPase	-0.0596	0.829	10.0	-	-	-	-
Rv3680		anion transporter ATPase	-0.3952	0.257	45.7	-	-	-	-
Rv3682	ponA2	penicillin-binding protein 1A/1B	-0.1182	0.633	56.7	X	-	X	-
Rv3683		hypothetical protein	0.5526	0.2	18.2	-	-	X	-
Rv3684		lyase	0.0104	0.976	17.9	-	-	-	-
Rv3685c	cyp137	cytochrome P450 137	-0.7612	0.065	20.9	-	-	-	-
Rv3688c		hypothetical protein	0.961	0.13	5.3	-	-	-	-
Rv3689		transmembrane protein	-0.0418	0.928	2.4	-	-	X	-
Rv3690		hypothetical protein	0.7304	0.079	11.7	-	-	X	-
Rv3691		hypothetical protein	-0.8094	0.014	8.7	-	-	-	-
Rv3692	moxR2	methanol dehydrogenase transcriptional regulator	1.1832	0.111	8.8	-	-	-	-
Rv3693		hypothetical protein	-0.6176	0.037	26.1	-	-	X	-
Rv3694c		transmembrane protein	-0.5802	0.283	22.8	-	-	X	-
Rv3695		hypothetical protein	-0.5462	0.212	7.3	-	-	X	-
Rv3701c		hypothetical protein	-0.019	0.958	7.6	-	-	-	-
Rv3703c		hypothetical protein	0.0596	0.689	1.8	-	-	-	-
Rv3704c	gshA	glutamate--cysteine ligase	-1.0903	0.012	1.2	-	-	-	-
Rv3705c		hypothetical protein	-0.4214	0.201	2.5	X	-	-	-
Rv3708c	asd	aspartate-semialdehyde dehydrogenase	0.1164	0.504	17.5	-	-	-	-
Rv3709c	ask	aspartate kinase	-0.3052	0.4	29.5	-	-	-	-
Rv3712		ligase	-0.5696	0.461	4.6	-	X	-	-
Rv3713	cobQ2	cobyric acid synthase	-0.8724	0.034	11.1	-	-	-	-
Rv3714c		hypothetical protein	0.9797	0.029	4.3	-	-	-	-
Rv3715c	recR	recombination protein	1.7992	0.006	6.2	-	-	-	-
Rv3716c		hypothetical protein	-0.1697	0.81	4.1	-	-	-	-
Rv3718c		hypothetical protein	1.0138	0.075	21.4	-	-	-	-
Rv3719		hypothetical protein	0.285	0.079	29.9	-	-	-	-
Rv3720		fatty acid synthase	0.463	0.199	35.9	-	-	-	-
Rv3721c	dnaZX	DNA polymerase III subunits gamma and tau	0.0628	0.75	24.7	-	-	-	-
Rv3723		transmembrane protein	0.1058	0.47	22.2	-	-	X	-
Rv3726		dehydrogenase	0.3592	0.199	23.9	-	-	-	-
Rv3729		transferase	0.8644	0.006	31.3	-	-	-	-
Rv3731	ligC	ATP-dependent DNA ligase	0.3112	0.587	2.6	-	-	-	-
Rv3732		hypothetical protein	-0.0851	0.758	26.6	X	-	X	-
Rv3734c		hypothetical protein	-0.0652	0.852	56.1	-	-	-	-
Rv3747		hypothetical protein	0.277	0.23	1.8	-	-	-	-
Rv3753c		hypothetical protein	-0.1379	0.414	0.5	-	-	-	-
Rv3758c	proV	(glycine betaine/carnitine/choline/L-proline) ATP-binding protein ABC transporter	-0.1698	0.503	6.8	-	-	-	-
Rv3759c	proX	(glycine betaine/carnitine/choline/L-proline) binding lipoprotein	-0.4214	0.407	7.4	X	-	-	X
Rv3760		hypothetical protein	0.0252	0.964	6.2	-	-	X	-
Rv3761c	fadE36	acyl-CoA dehydrogenase	0.113	0.803	8.3	-	-	-	-
Rv3762c		hydrolase	-0.3796	0.453	15.3	-	-	-	-
Rv3763	lpqH	19 kDa lipoprotein antigen precursor	-0.0884	0.364	72.2	X	-	-	X
Rv3764c		two component sensor kinase	-1.0903	0.012	1.2	-	-	X	-
Rv3765c		two component transcriptional regulatory protein	0.2648	0.838	7.4	-	-	-	-
Rv3766		hypothetical protein	1.2316	0.009	1.9	-	-	-	-
Rv3767c		hypothetical protein	0.6354	0.453	2.5	-	-	-	-
Rv3774	echA21	enoyl-CoA hydratase	0.8392	0.045	8.0	-	-	-	-
Rv3775	lipE	lipase	-0.4557	0.184	30.5	-	-	-	-

Rv3778c		aminotransferase	-1.8567	0.098	3.3	-	-	-	-
Rv3779		transmembrane protein alanine and leucine rich	-0.6338	0.64	3.8	-	-	X	-
Rv3780		hypothetical protein	1.9032	0.059	1.0	-	-	-	-
Rv3781	rfbE	O-antigen/lipopolysaccharide ABC transporter	0.4018	0.241	19.7	-	-	-	-
Rv3782		L-rhamnosyltransferase	0.1796	0.542	17.5	-	-	-	-
Rv3783	rfbD	O-antigen/lipopolysaccharide ABC transporter	-1.0733	0.33	2.3	-	-	X	-
Rv3789		integral membrane protein	0.393	0.376	1.6	-	-	X	-
Rv3790		oxidoreductase	-1.3336	0.127	4.5	-	-	-	-
Rv3791		short chain dehydrogenase	-0.4329	0.036	18.0	-	-	-	-
Rv3792		transmembrane protein	-0.0744	0.831	5.0	-	-	X	-
Rv3793	embC	indolylacetylinositol arabinosyltransferase (arabinosylindolylacetylinositol synthase)	-0.421	0.016	52.2	-	-	X	-
Rv3794	embA	indolylacetylinositol arabinosyltransferase/ arabinosylindolylacetylinositol synthase	-0.3445	0.054	74.8	-	-	X	-
Rv3795	embB	indolylacetylinositol arabinosyltransferase (arabinosylindolylacetylinositol synthase)	-0.2764	0.068	58.2	-	-	X	-
Rv3796		hypothetical protein	-0.2244	0.748	8.5	-	X	-	-
Rv3799c	accD4	propionyl-CoA carboxylase beta chain	1.2202	0.008	6.6	-	-	-	-
Rv3800c	pks13	polyketide synthase	0.6632	0.002	221.7	-	-	-	-
Rv3801c	fadD32	acyl-CoA synthetase	-0.365	0.152	163.1	-	-	-	-
Rv3802c		hypothetical protein	-0.698	0.12	13.6	-	-	X	-
Rv3803c	fbpD	secreted MPT51/MPB51 antigen protein FBPD (	-0.52	0.428	12.4	X	-	X	-
Rv3804c	fbpA	secreted antigen 85-A (	1.1412	0.017	50.0	-	X	X	-
Rv3805c		transmembrane protein	-0.4632	0.352	17.2	-	-	X	-
Rv3806c		phosphoribose diphosphate:decaprenyl-phosphate phosphoribosyltransferase	-0.0646	0.963	11.9	-	-	X	-
Rv3807c		transmembrane protein	-1.2853	0.124	1.4	-	-	X	-
Rv3808c	glfT	UDP-galactofuranosyl transferase GLFT	-0.3156	0.204	24.3	-	-	-	-
Rv3813c		hypothetical protein	1.9376	0.008	4.2	-	-	-	-
Rv3814c		acyltransferase	1.1851	0.168	20.6	-	-	-	-
Rv3815c		acyltransferase	1.0363	0.083	4.6	-	-	-	-
Rv3816c		acyltransferase	0.7852	0.07	27.6	-	-	-	-
Rv3818		hypothetical protein	0.5364	0.287	1.5	-	-	-	-
Rv3819		hypothetical protein	1.2606	0.015	17.3	-	-	-	-
Rv3824c	papA1	polyketide synthase associated protein	-0.0498	0.807	9.8	-	-	-	-
Rv3825c	pks2	polyketide synthase	0.2949	0.012	300.0	-	-	-	-
Rv3826	fadD23	acyl-CoA synthetase	0.4595	0.694	2.4	-	-	X	-
Rv3834c	serS	seryl-tRNA synthetase	-0.6753	0.029	1.9	-	-	-	-
Rv3835		hypothetical protein	0.9963	0.022	10.5	-	-	X	-
Rv3841	bfrB	bacterioferritin BfrB	0.5666	0.038	21.7	-	-	-	-
Rv3842c	glpQ1	glycerophosphoryl diester phosphodiesterase	-0.4524	0.099	11.2	-	-	-	-
Rv3843c		transmembrane protein	0.179	0.17	22.2	-	-	X	-
Rv3846	sodA	superoxide dismutase [Fe] SODA	0.1136	0.541	0.7	-	-	-	-
Rv3849		hypothetical protein	0.5447	0.288	7.7	-	-	-	-
Rv3850		hypothetical protein	0.7854	0.036	42.0	-	X	-	-
Rv3852	hns	histone-like protein	-0.0552	0.751	16.5	-	-	X	-
Rv3854c	ethA	monooxygenase	0.5332	0.072	20.4	-	-	-	-
Rv3864		hypothetical protein	0.0609	0.756	11.5	-	-	X	-
Rv3865		hypothetical protein	1.1636	0.031	4.1	-	-	-	-
Rv3866		hypothetical protein	-0.2498	0.624	6.6	-	-	-	-
Rv3867		hypothetical protein	-0.3746	0.629	5.6	-	-	-	-
Rv3868		hypothetical protein	2.0177	0.012	8.9	-	-	-	-
Rv3869		hypothetical protein	0.0702	0.81	21.9	-	-	X	-
Rv3870		transmembrane protein	0.0664	0.792	58.8	-	-	X	-
Rv3871		hypothetical protein	0.2591	0.503	55.1	-	-	-	-
Rv3874	esxB	10 kDa culture filtrate antigen	-0.836	0.387	3.3	-	-	-	-
Rv3876		hypothetical protein	0.5104	0.176	35.0	-	-	-	-
Rv3877		transmembrane protein	0.1586	0.745	15.0	-	-	X	-

Rv3879c		hypothetical protein	0.6919	0.372	2.5	-	-	-	-
Rv3880c		hypothetical protein	1.1092	0.002	11.3	-	-	-	-
Rv3882c		hypothetical protein	-0.3346	0.155	39.2	-	-	X	-
Rv3883c	mycP1	membrane-anchored mycosin	-0.2362	0.595	6.0	X	-	X	-
Rv3884c	eccA2	ESX-2 type VII secretion system	1.0068	0.105	5.4	-	-	-	-
Rv3885c		hypothetical protein	-0.028	0.898	31.6	-	-	X	-
Rv3886c	mycP2	alanine and proline rich mycosin	-0.802	0.019	11.0	X	-	X	-
Rv3887c		transmembrane protein	-0.7782	0.213	10.2	-	-	X	-
Rv3888c		hypothetical protein	-0.1161	0.476	47.4	-	-	-	-
Rv3894c		hypothetical protein	0.001	0.995	68.9	-	-	X	-
Rv3895c		hypothetical protein	-0.3499	0.185	20.3	-	-	X	-
Rv3902c		hypothetical protein	1.0048	-	0.5	-	-	-	-
Rv3908		hypothetical protein	-0.6252	0.018	11.7	-	-	-	-
Rv3909		hypothetical protein	1.0258	0.003	52.5	X	-	-	-
Rv3910		transmembrane protein	-0.3912	0.008	116.5	-	-	X	-
Rv3914	trxC	thioredoxin trxC (TRX) (MPT46)	-0.959	0.538	5.1	-	-	-	-
Rv3915		hydrolase	0.0428	0.864	9.5	-	-	-	-
Rv3917c	parB	chromosome partitioning protein	0.3125	0.032	22.0	-	-	-	-
Rv3918c	parA	chromosome partitioning protein	-1.1878	0.016	12.4	-	-	-	-
Rv3921c	yidC	putative inner membrane protein translocase component	-0.7448	0.471	19.5	-	-	X	-
Rv3923c	rnpA	ribonuclease P	-0.7412	0.196	5.8	-	-	-	-

a – An "X" indicates the presence of an N-terminal signal peptide predicted using SignalP v. 4.0 (2)

b – An "X" indicates the presence of an N-terminal Tat signal peptides as predicted using TatP and TatFind

c – An "X" indicates one or more putative transmembrane domains predicted using using TMHMM v. 2.0 (3)

d – An "X" indicates a potential lipoprotein containing a putative lipobox-type signal peptide as predicted by the LipoP v. 1.0 program (1) or using a lipoprotein pattern search against the *M. tuberculosis* H37Rv genome (4)

References

1. **Juncker, A. S., H. Willenbrock, G. Von Heijne, S. Brunak, H. Nielsen, and A. Krogh.** 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci **12:**1652-1662.

2. **Petersen, T. N., S. Brunak, G. von Heijne, and H. Nielsen.** 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods **8:**785-786.

3. **Sonnhammer, E. L., G. von Heijne, and A. Krogh.** 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc Int Conf Intell Syst Mol Biol **6:**175-182.

4. **Sutcliffe, I. C., and D. J. Harrington.** 2004. Lipoproteins of *Mycobacterium tuberculosis*: an abundant and functionally diverse class of cell envelope components. FEMS Microbiol Rev **28:**645-659.