Suppressor Analysis to Identify Proteins that Work with the *Mycobacterium smegmatis* SecA2 Protein Export System

Lauren Stone Ligon

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

Thomas Kawula, PhD

Robert Bourret, PhD

Peggy Cotter, PhD

Anthony Richardson, PhD

ABSTRACT

LAUREN STONE LIGON: Suppressor Analysis to Identify Proteins that Work with the *Mycobacterium smegmatis* SecA2 Protein Export System
(Under the direction of Miriam Braunstein)

Mycobacterium tuberculosis, the causative agent of the disease tuberculosis, is a serious threat to human health, responsible for 1.4 million deaths annually. When *M. tuberculosis* is spread from person to person by aerosol and is phagocytosed by alveolar macrophages, the proteins exported by the bacterium are ideally positioned to interact with the host. Therefore, exported proteins and the systems responsible for their export are important for *M. tuberculosis* virulence. All bacteria, including mycobacteria, contain the Sec protein export system, which is responsible for the bulk of protein export and is composed of a motor ATPase protein, SecA, and a membrane-embedded channel complex, SecYEG. Mycobacteria, including both *M. tuberculosis* and the model organism *Mycobacterium smegmatis*, along with some Gram-positive bacteria, are unique in containing two functionally distinct SecA proteins. The SecA protein responsible for housekeeping protein export is termed SecA1 and is essential for cell survival, while the second SecA protein, termed SecA2, is required for the export of a small subset of proteins and is important for *M. tuberculosis* virulence. While the canonical Sec system is well understood, the mechanism of SecA2-dependent export is not. Like canonical SecA proteins, the mycobacterial SecA2 requires ATPase activity to function. Furthermore, an *M. smegmatis secA2* mutant (*secA2 K129R*) that encodes a SecA2 protein defective in ATP binding is dominant negative and appears to be trapped in a complex with interacting proteins

at the membrane. Here, we analyze extragenic suppressors of *secA2 K129R* in an effort to better understand SecA2-dependent export and identify additional components of the SecA2 pathway. Using this approach, we demonstrate a connection between *M. smegmatis* SecA2 and SecY, the major membrane-embedded component of the housekeeping Sec export channel. In addition, we demonstrate a connection between the SecA2 system and Msmeg_1684, a protein of unknown function that is found throughout mycobacteria and may represent a novel component of the SecA2 pathway. Our findings suggest a new model in which the mycobacterial SecA2 export pathway is integrated into the housekeeping Sec pathway and includes an additional SecA2-specific component, Msmeg_1684.

ACKNOWLEDGEMENTS

I would like to thank the following people who contributed to the work described in this dissertation. Dr. Nathan Rigel isolated the suppressor strains discussed in Chapters 2 and 3. Artur Romanchuk and Dr. Corbin Jones performed whole genome alignment and mutation identification for suppressors 4S in Chapter 2 and 7S and 20B in Chapter 3. Brittany Miller performed the SecA1 overexpression experiment in Chapter 2, the quantitative RT-PCR on genes *msmeg_1677-1683* in Chapter 3, and the *msmeg_1684* deletion phenotype testing in Chapter 3.

In addition, I would like to thank several people who cloned or provided plasmids and strains used in this work. Dr. Adrie Steyn provided plasmid pCV125, used in Chapter 2. Dr. William Jacobs provided plasmid pYUB2063, used in Chapter 2. Brandon Anjuwon-Foster created strain BAF1, used in Chapter 3. Chelsea Moriarty created plasmid pCM2 and strains CM11 and CM14, used in Chapter 3. Courtney Sutphen created plasmids pCS1-4, used in Chapter 3. Meagan Martin created plasmids pMEM3 and pMEM4, used in Chapter 3. Dr. Henry Gibbons created plasmid pHSG93, used in Chapter 4.

Finally, I would like to thank Dr. Anthony Richardson for technical assistance with 5'RLM-RACE and quantitative RT-PCR and members of the Braunstein lab for critical reading of Chapters 1 and 2.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure

LIST OF ABBREVIATIONS AND SYMBOLS

Δ	deletion
2D	Two-dimensional
Å	angstrom
β-ME	beta-mercaptoethanol
µg	microgram
µL	microliter
A	adenine
ABC	ATP binding cassette
ADP	adenosine diphosphate
ANOVA	analysis of variance
Asp	aspartate
Atc	anhydrotetracycline
ATP	adenosine triphosphate
ATPase	ATP hydrolase
BLAST	basic local alignment search tool
BLAT	BLAST-like alignment tool
bp	base pair
C	Celsius or cytosine
C.	*Clostridium* or *Corynebacteria*
cDNA	complementary DNA
comp.	complemented
C-terminal	carboxy-terminal
C-terminus	carboxy-terminus

CW	cell wall fraction
Cys	cysteine
DNA	deoxyribonucleic acid
DSP	dithiobis(succinimidyl proprionate)
E	glutamate
E.	*Escherichia*
ECF	enhanced chemifluorescence
ENV	envelope fraction
ESX	ESAT secretion system
F	phenylalanine
g	gravity
G	guanine or glycine
Gly	glycine
HA	hemagglutinin
His	hexahistidine tag
HIV	Human Immunodeficiency Virus
hyg	hygromycin resistance gene
I	isoleucine
IgG	Immunoglobulin G
K	lysine
kan	kanamycin resistance gene
kbp	kilobase pair
kDa	kilodalton
L	leucine

L.	*Listeria*
M	molar or methionine
M.	*Mycobacterium*
Mb	mega base pair
MDR	Multi-drug resistant
MEM	membrane fraction
Met	methionine
mg	milligram
min	minute
mL	milliliter
mM	millimolar
*msmeg_*	*Mycobacterium smegmatis* gene designation
Msmeg_	*Mycobacterium smegmatis* protein designation
n	nucleotide
N	asparagine
N/A	not applicable
NCBI	National Center for Biotechnology Information
Ncgl	*Clostridium glutamicum* gene designation
Ni-NTA	nickel nitrilotriacetic acid
ng	nanogram
nm	nanometer
N-terminal	amino-terminal
N-terminus	amino-terminus
$OD_{600\ nm}$	optical density at 600 nanometers

P	promoter
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
P_i	inorganic phosphate
Pro	proline
Q	glutamine
R	arginine
revTetR	reverse Tet repressor
RLM-RACE	RNA ligase mediated rapid amplification of cDNA ends
RNA	ribonucleic acid
RT	reverse transcriptase
rv	*Mycobacterium tuberculosis* gene designation
Rv	*Mycobacterium tuberculosis* protein designation
S	serine
S.	*Streptococcus*
SDAD	NHS-SS-Diazirine
SDS	sodium dodecyl sulfate
Sec	secretion
SNP	single nucleotide polymorphism
SOAP	short oligonucleotide alignment program
SOD	superoxide dismutase
SOL	soluble fraction
SP	signal peptidase
SRP	signal recognition particle

T	thymine
TAP	tobacco acid pyrophosphatase
Tat	Twin-arginine translocation
TB	tuberculosis
Thr	threonine
Tn-seq	transposon deep sequencing
TraSH	transposon site hybridization
Trp	Tryptophan
UV	ultraviolet
Val	valine
WCL	whole cell lysate
XDR	extensively drug resistant
Y	tyrosine

CHAPTER 1

Introduction[1]

Mycobacterium tuberculosis is the causative agent of the disease tuberculosis, which the World Health Organization estimates kills 1.4 million people every year (1). This pathogen is spread from person to person by aerosols created by sneezing and coughing. When aerosols are inhaled by an uninfected person, *M. tuberculosis* is able to reach the alveolar spaces and is taken up by alveolar macrophages. However, *M. tuberculosis* is not killed by macrophages, but rather can survive and persist, sometimes for decades, before eventually causing active disease. HIV-positive individuals are especially susceptible to developing active tuberculosis disease and, as a result, HIV/tuberculosis co-infection is the primary cause of death worldwide for the HIV-infected population. No effective vaccine is available and although tuberculosis is treatable, it requires a long antibiotic regimen that is expensive and difficult to complete, especially for people in the developing world. Furthermore, multidrug resistant (MDR) and extensively drug resistant (XDR) *M. tuberculosis* strains make treatment even more problematic. Therefore, novel drugs are desperately needed to combat this disease.

Because of their importance to virulence and bacterial viability, the exported proteins of *M. tuberculosis* and their respective protein export systems can be considered potential drug

[1]Adapted for this dissertation from: Ligon LS, Hayden JD, Braunstein M. 2012. The ins and outs of Mycobacterium tuberculosis protein export. Tuberculosis (Edinb) 92(2):121-132.

targets (recently reviewed in (2)). Many exported proteins, which we define as including proteins in the cell envelope of *M. tuberculosis* and proteins secreted by *M. tuberculosis*, interact with the host and are crucial for causing disease. Without its systems for transporting these proteins across the cytoplasmic membrane and to their final destination, *M. tuberculosis* cannot deliver effector proteins that are necessary for virulence. Furthermore, some *M. tuberculosis* protein export systems are essential.

In *M. tuberculosis,* as well as the non-pathogenic model organism *Mycobacterium smegmatis*, there are two highly conserved protein export systems that are responsible for the majority of protein export: the Sec and Tat export pathways. *M. tuberculosis* and *M. smegmatis* also possess specialized protein export systems dedicated to the export of a more limited set of proteins: the accessory SecA2 export pathway and ESX pathways. ESX pathways are also referred to as Type VII secretion systems. Interestingly, both of these specialized protein export systems were first identified in *M. tuberculosis* but later found to also exist in some other bacteria. A better understanding of the mycobacterial protein export systems could lead to new strategies to combat tuberculosis disease. This dissertation describes efforts to understand the mechanism of the SecA2 export pathway, which is important for *M. tuberculosis* virulence.

Housekeeping Sec Export System

The Sec system is highly conserved and present in all bacteria, and it acts as the primary route for exporting proteins to the cytoplasmic membrane and beyond. Because many of the proteins exported by the Sec system perform vital functions that require proper export, the Sec pathway is essential for bacterial viability (Table 1.1). In addition, many bacterial proteins with roles in virulence are exported by the Sec system (3). The current understanding of Sec export

comes from extensive study in *Escherichia coli* (for recent reviews see (4, 5)). While Sec export functions in essentially the same manner in all bacteria in which it has been studied, only a few aspects of this system have been directly investigated in mycobacteria.

Exported proteins and targeting. The Sec system exports proteins across the cytoplasmic membrane post-translationally. Proteins exported by the Sec system, termed preproteins or precursors, are synthesized with an N-terminal signal peptide that is important for targeting the protein for export. Following export, the signal peptide is cleaved to generate the mature protein. Sec signal peptides consist of a positively charged N-terminus, a hydrophobic central domain, and an uncharged polar C-terminus containing the cleavage site (5). Some Sec-exported proteins are lipoproteins, and in these cases the C-terminal end of the signal peptide contains a lipobox motif with a conserved cysteine that is the site of lipid attachment (6). Sec signal peptides and lipoprotein signal peptides can be predicted using the bioinformatic programs PSORTb v3.0 (7), SignalP v4.0 (8), and LipoP v1.0 (9).

Proteins must be in an unfolded state to be exported by the Sec system. In many Gram-negative bacteria there is a Sec export chaperone, SecB, that binds preproteins, maintains them in an unfolded state, and delivers them to SecA – a central component of the Sec system discussed below. However, not all preproteins of Gram-negative bacteria require SecB, and Gram-positive bacteria lack a SecB ortholog (10). In the absence of SecB, it is thought that other proteins fulfill the role of export chaperone. In *M. tuberculosis*, Bordes et al. recently identified Rv1957 as a SecB-like protein (11). Rv1957 prevents proteolysis and aggregation of *M. tuberculosis* HigA, a presumed cytoplasmic protein that is part of the *M. tuberculosis* HigBA toxin-antitoxin system. These activities are consistent with Rv1957 being a chaperone.

However, there is currently no evidence that Rv1957 functions as a chaperone for protein export and it is possible that the function of Rv1957 is restricted to the HigBA system.

System components and mechanism. SecA is a multifunctional component of the Sec export system that works with the heterotrimeric SecYEG membrane channel to export proteins across the cytoplasmic membrane. SecA recognizes and binds to the signal peptide, as well as portions of the mature domain, of preproteins (12, 13). SecA is also an ATPase that harnesses energy from multiple rounds of ATP binding and hydrolysis to "push" preproteins through the SecYEG channel (14, 15). The proton motive force can also contribute to Sec export, but is not absolutely required (15).

Mycobacteria and some Gram-positive bacteria have two SecA proteins, each with distinct functions. The primary housekeeping SecA, which functions like SecA of E. coli, is referred to as SecA1, while the accessory SecA is referred to as SecA2 (see later section for SecA2 discussion). SecA1 is predicted to be essential in *M. tuberculosis* (16, 17) and proven to be essential in the model mycobacterium *M. smegmatis* (18). Because *secA1* deletion mutants are not viable, conditional silencing has been used to study SecA1 function. As would be expected for a housekeeping SecA, depletion of SecA1 in *M. smegmatis* causes both growth inhibition and reduced export of a test protein with a Sec signal peptide (19, 20). *M. tuberculosis* SecA1 is also a demonstrated ATPase (21), and the X-ray crystal structure of *M. tuberculosis* SecA1 is similar to other bacterial SecA structures (22).

The SecYEG complex provides the channel through which proteins travel across the cytoplasmic membrane (23) and together with SecA it forms the "translocase," which is the minimum apparatus needed to reconstitute efficient Sec export *in vitro* (24). SecY is a polytopic membrane protein and the largest component of the SecYEG complex, forming the pore in the

membrane through which proteins pass during export. SecE is suggested to act as a "clamp" to stabilize SecY (25). SecG improves efficiency of protein export (26). Additional Sec components that improve export efficiency are SecD, SecF, and YajC (27).

Once a protein has been exported across the membrane, the signal peptide is removed by one of two signal peptidases – the Type I signal peptidase LepB, or the Type II signal peptidase LspA (28). LspA acts specifically on lipoprotein signal peptides. The *M. tuberculosis* LepB (29) and LspA (30) have been studied and confirmed to function in protein export. Furthermore, *lepB* is essential (29) and an *lspA* mutant of *M. tuberculosis* is attenuated in macrophage and mouse models of infection (30), demonstrating the importance of correctly exported proteins for *M. tuberculosis* survival and virulence.

In addition to the process described above for exporting proteins completely across the cytoplasmic membrane, the SecYEG channel is also used for a significantly different process – co-translational insertion of integral membrane proteins into the cytoplasmic membrane (for a recent review, see (31)). In this process, transmembrane domains of nascent integral membrane proteins are recognized by the signal recognition particle SRP during translation. SRP then delivers the nascent integral membrane protein to the SRP receptor FtsY, which in turn passes the protein to SecYEG for co-translational insertion into the membrane. SecY contains a "lateral gate" which is believed to allow transmembrane domains of integral membrane proteins to pass sideways out of the SecYEG channel and into the membrane, with the help of the YidC protein (32, 33). While not generally required for integral membrane protein insertion, the SecA protein does assist in the case of integral membrane proteins with large hydrophilic domains (34). Integral membrane proteins can be predicted using the bioinformatic program TMHMM v2.0 (35), which searches for transmembrane domains. Specifically in mycobacteria,

co-translational insertion of integral membrane proteins has only barely begun to receive research attention, with the recent confirmation that *M. tuberculosis* SRP displays biochemical properties comparable to SRP of *E. coli* (36).

Contribution to disease and physiology. Because the Sec system performs the bulk of protein export, many Sec-exported proteins are important for both *M. tuberculosis* viability and virulence. Several *M. tuberculosis* proteins proven to function in virulence contain Sec signal peptides (37, 38). Also, as mentioned above, the importance of LspA demonstrates that Sec-exported lipoproteins are important for *M. tuberculosis* virulence (30). In fact, several lipoproteins with Sec signal peptides are known to be important for *M. tuberculosis* virulence, including LpqH (19kD) (39), LppX (40), and LprG (41). While SecYEG is homologous to the Sec61 complex found in eukaryotic cells, there is no eukaryotic homolog of SecA. The lack of SecA in mammalian cells, combined with the critical importance of SecA1 for both viability and virulence in bacteria, makes SecA1 a compelling potential drug target (as reviewed in (2)).

Model, summary, and future questions. At this point, the studies of SecA1, LepB, LspA, and SRP are the only ones to directly investigate housekeeping Sec export in mycobacteria (18-22, 29, 30, 36, 42). However, the data from these studies, combined with the presence of Sec component orthologues (Table 1.1) and exported proteins with Sec signal peptides in both *M. tuberculosis* and *M. smegmatis*, is consistent with the housekeeping Sec system of mycobacteria functioning as described in other bacteria (Fig. 1.1). In this case, preproteins with N-terminal signal peptides are recognized by SecA1, which interacts with the SecYEG channel complex to form the translocase. SecA1 performs repeated cycles of ATP hydrolysis, pushing segments of the preprotein through the SecYEG channel. Signal peptides are removed by a LepB or LspA signal peptidase, and mature exported proteins fold into their final conformations.

In the future it will be important to expand our understanding of the Sec export system in mycobacteria. For example, nothing is currently known about mycobacterial chaperone(s) involved in Sec export. Also, the process of inserting integral membrane proteins into the mycobacterial cytoplasmic membrane is a topic that has so far received little research attention.

Accessory SecA2 Export System

All mycobacteria are unusual in having two SecA proteins. In contrast to SecA1, SecA2 is non-essential and *secA2* deletion mutants have been constructed in several mycobacteria, including *M. tuberculosis* (43), *M. smegmatis* (18), and *Mycobacterium marinum* (44). SecA2 functions in protein export, but its role appears limited to a much smaller subset of proteins than those exported by SecA1. SecA2 is necessary for the full virulence of *M. tuberculosis* (43, 45) and *M. marinum* (44), suggesting that one or more of the proteins it exports are important in pathogenesis. Accessory SecA2 export has been studied directly in mycobacteria; however, there is still much that is unknown.

Exported proteins and targeting. Comparative 2D-PAGE analysis of exported proteins from wild-type and *secA2* deletion mutant strains has been used to identify proteins that require SecA2 for their export (43, 46). Only a small number of proteins were identified in these studies, and they include examples both with and without signal peptides. In *M. smegmatis*, cell wall proteins were analyzed by 2D-PAGE, identifying Msmeg_1704 and Msmeg_1712 as proteins exported by the SecA2 system (46). These proteins share many similarities – both are lipoproteins, contain lipoprotein Sec signal peptides (9), and are predicted sugar-binding components of ABC transporter systems. While direct homologues of Msmeg_1704 and Msmeg_1712 are not found in *M. tuberculosis*, similar sugar-binding lipoproteins are present,

although they have yet to be evaluated for SecA2-dependence. The SecA2-dependent proteins with signal peptides are currently indistinguishable from proteins exported by the housekeeping Sec export system, and experimental evidence shows the signal peptide *is* required for export of Msmeg_1712 (46). However, it is the mature domain of *M. smegmatis* SecA2-dependent proteins, *not* the signal peptide, that imparts the requirement for SecA2 in their export (47).

In *M. tuberculosis*, proteins exported into the culture media (culture filtrates) were analyzed by 2D-PAGE (43). Among the very few proteins identified was the antioxidant enzyme SodA (superoxide dismutase). SodA does not have a signal peptide, but its dependence on SecA2 for export was confirmed by western blot and SodA activity assay (43, 48). Given the role of SodA in detoxifying oxygen radicals, the export of *M. tuberculosis* KatG (catalase), another antioxidant enzyme, was directly evaluated for SecA2-dependence. Like SodA, KatG lacks a signal peptide and is dependent on SecA2 for export (43).

System components and mechanism. Accessory SecA2 proteins are found in all mycobacteria and some Gram-positive bacteria (for an extensive review of SecA2 export, see (49)). Several SecA2 export systems include a SecY2 protein. Referred to as SecA2-SecY2 systems, these systems seem to be dedicated to exporting a single large protein that is heavily glycosylated prior to export (50-52). In addition to the SecA2 and SecY2 proteins, proteins referred to as Asps (or alternately Gaps) also participate in SecA2-SecY2 protein export (52, 53). Asp1, Asp2, and Asp3 are found in all SecA2-SecY2 systems. The *Streptococcus gordonii* system also includes proteins Asp4 and Asp5 (54), which have limited homology to the canonical SecE and SecG proteins, suggesting they may function as components of a SecY2 membrane channel complex. Other proteins associated with SecA2-SecY2 systems are involved in glycosylation of preproteins prior to their export (51, 53, 55-58). In contrast, the mycobacterial SecA2 system

does not contain a SecY2 protein or any other known additional components. The mycobacterial SecA2 system is therefore referred to as a SecA2-only system.

SecA2 has been directly studied in both *M. tuberculosis* and *M. smegmatis*. Importantly, complementation experiments show that SecA2 of *M. tuberculosis* and *M. smegmatis* can substitute for one another to complement *secA2* deletion mutant phenotypes (20). Because these cross-species complementation experiments demonstrate that the SecA2 proteins of *M. tuberculosis* and *M. smegmatis* can carry out the same functions, *M. smegmatis* has been used as a model for studying the mechanism of mycobacterial SecA2-dependent export.

SecA2 exhibits several similarities to the housekeeping SecA1 protein. SecA1 and SecA2 proteins are present in equivalent amounts in *M. tuberculosis*, and both proteins have proven ATPase activity (21). Like housekeeping SecA proteins, ATP hydrolysis is also required for SecA2 function in both *M. tuberculosis* (21) and *M. smegmatis* (20).

There are also several differences between SecA1 and SecA2. SecA2 is smaller than SecA1 because it lacks the C-terminal linker region, which in *E. coli* has been implicated in binding to phospholipids, SecB, and zinc (59-61). SecA2 also differs from SecA1 in how it localizes within the bacterium. Like *E. coli* SecA, SecA1 is evenly distributed between soluble and cell envelope fractions, but SecA2 is found predominantly in the cytosol-containing soluble fraction (20). Most importantly, it is clear that SecA1 and SecA2 have independent functions. Even when SecA2 is overexpressed, SecA1 cannot be deleted, indicating that SecA2 cannot substitute for SecA1. Similarly, overexpression of SecA1 does not rescue the phenotypes of a *secA2* deletion mutant (18).

Given that there is no accessory SecY or obvious alternate export channel to work with SecA2 in mycobacteria, it seems likely that SecA2 functions with the help of some or all of the

housekeeping Sec export system. To address this possibility, export of the SecA2-dependent protein Msmeg_1712 was assessed following depletion of SecA1 in *M. smegmatis*. In the presence of SecA1 depletion, export of Msmeg_1712 was found to be significantly compromised (20). The simplest interpretation of this result is that the SecA2 export system needs the housekeeping SecA1 to function. However, the possibility that SecA1 depletion has an indirect effect on SecA2 export cannot currently be ruled out.

Listeria monocytogenes, *Clostridium difficile*, and *Corynebacterium glutamicum* also have SecA2-only type export systems. A notable similarity between the *L. monocytogenes* and mycobacterial SecA2 systems is that both are reported to export proteins with signal peptides as well as proteins without signal peptides (62, 63). In *L. monocytogenes*, this includes export of MnSOD (manganese superoxide dismutase), which lacks a signal peptide (63). In addition, SecA2 of *L. monocytogenes*, like that of *M. tuberculosis*, is important for virulence but is not essential for growth (62, 64). Interestingly, *L. monocytogenes* SecA2 has been found to be enriched at cell division sites, and the polarly-localized cell division protein DivIVA is required for export of SecA2-dependent proteins p60 and MurA (65). It will be interesting to see if mycobacterial SecA2 proteins also exhibit polar localization.

Unlike in mycobacteria, SecA2 of *C. difficile* and *C. glutamicum* are essential for growth (66, 67), which may reflect a difference in the essentiality of proteins exported by each of these systems. Despite this difference, there are also similarities between the mycobacterial and *C. difficile* SecA2 systems. In both species, SecA2 is found localized predominantly to the cytoplasm (20, 66). An amino acid substitution in the Walker box of *M. smegmatis* SecA2, SecA2 K129R, prevents binding and hydrolysis of ATP and results in a dominant negative phenotype (20), as does the comparable SecA2 K106R substitution in *C. difficile* SecA2 (66). In addition,

these non-functional, dominant negative SecA2 KR variants cause growth inhibition in both species, suggesting they interfere with an essential cell process. As discussed in chapter 2 of this dissertation, the essential process disrupted in *M. smegmatis* is believed to be the canonical Sec export pathway.

Contribution to disease and physiology. The *M. tuberculosis secA2* mutant is attenuated for growth in both macrophages (45) and the mouse model of tuberculosis infection (43). Similarly, the *M. marinum secA2* mutant is attenuated for growth in both zebrafish and mouse models of infection (44). These results indicate that the SecA2 system exports proteins important for *M. tuberculosis* and *M. marinum* virulence. The fact that both *M. tuberculosis* SodA and KatG proteins are dependent on SecA2 for export suggests that the SecA2 system may protect *M. tuberculosis* from the oxidative burst of macrophages. Yet, the *secA2* mutant is still attenuated for growth in *phox*$^{-/-}$ macrophages, which cannot produce an oxidative burst (45). While this result does not exclude a role for SecA2 in resisting oxygen radicals during infection, it does reveal the existence of other roles for the SecA2 system in promoting virulence. During macrophage infection, the *M. tuberculosis secA2* mutant induces increased release of proinflammatory cytokines (45) and increased apoptosis (48), as compared to wild-type *M. tuberculosis*. Thus, the SecA2 system may block innate immune responses to enable *M. tuberculosis* growth in macrophages. In addition, the *secA2* mutant fails to arrest phagosome maturation and acidification, and the growth defect of the *secA2* mutant in macrophages is rescued by inhibitors of phagosome acidification (68). These findings demonstrate the importance of phagosome maturation arrest for *M. tuberculosis* growth in macrophages and suggest that the SecA2 system may export proteins important for blocking this host defense during wild type *M. tuberculosis* infection.

Model, summary, and future questions. More research is needed to elucidate the details of SecA2-dependent export. However the studies conducted prior to this dissertation, particularly those using *M. smegmatis*, help build a preliminary model for this system (Fig. 1.1). SecA2 recognizes a small subset of proteins that are normally not compatible with Sec export, and uses its ATPase activity to assist in their export. It seems most likely that SecA2 works with components of the housekeeping Sec export system, either by delivering proteins to SecA1 or by actively participating in exporting proteins through the SecYEG channel. It remains possible, however, that other unknown components are required in addition to or in lieu of the housekeeping Sec components.

Many important questions remain. For the SecA2-dependent proteins with signal peptides, it is not known what distinguishes them from proteins exported by the housekeeping Sec system, or why these SecA2-dependent proteins cannot be exported by the housekeeping Sec system alone. The fact that some SecA2-dependent proteins do not have signal peptides adds an additional layer of complexity to understanding this system. It is possible that both proteins with signal peptides and proteins without signal peptides are directly exported by the SecA2 export system. Conversely, the SecA2-dependence of proteins without signal peptides may be an indirect effect. For example, it is possible that the SecA2 system exports an unidentified protein containing a signal peptide, which in turn participates in the export of proteins lacking signal peptides via another pathway. Finally, it is likely that more SecA2-dependent proteins exist and remain to be identified, including *M. tuberculosis* proteins containing signal peptides. Important future goals are to identify additional proteins exported by the SecA2 system and understand the role this export system plays in the pathogenesis of *M. tuberculosis*.

Summary

While the canonical Sec export system has been studied extensively over several decades and is now well characterized, the SecA2 export system of mycobacteria is poorly understood. In addition to questions about the identity and targeting of SecA2-dependent proteins and the role of SecA2-dependent export in *M. tuberculosis* virulence, the mechanistic steps of SecA2 export remain unknown. In this dissertation, we describe a search for additional proteins that are either required for or connected to SecA2-dependent protein export. We demonstrate a relationship between SecA2 and the canonical Sec pathway, specifically the SecY channel protein. We also identify a genetic link between SecA2 and the conserved hypothetical protein Msmeg_1684. Finally, we demonstrate that the relationship between SecA2 and SecY does not appear to represent a direct, physical interaction.

In chapter 2, we describe a suppressor analysis carried out using the dominant-negative *M. smegmatis secA2 K129R* allele. Interestingly, two extragenic suppressor strains carry mutations located upstream of the *secY* gene, encoding the main component of the canonical Sec channel complex. We demonstrate that these mutations are located within the *secY* promoter and that they cause increased expression of *secY*. This increase in *secY* expression is sufficient to suppress the *secA2 K129R* allele. Furthermore, we find that SecY protein levels are severely reduced in the presence of the non-functional SecA2 K129R protein. This supports a model in which SecA2 K129R becomes locked in non-functional interactions with the Sec pathway, creating stress on housekeeping Sec export, and resulting in degradation of SecY. Suppressor mutations can alleviate this stress by helping to restore the amount of SecY protein available. Additionally, our findings imply that SecA2 works with the SecYEG channel to export proteins. Finally, we revisit the previously reported requirement for SecA1 in SecA2-dependent

export (20). Depletion of SecA1 reduces export of a SecA2-dependent protein, but does not cause degradation of SecY, confirming the requirement for SecA1 in SecA2-dependent protein export.

In chapter 3, we extend our suppressor analysis of the *M. smegmatis secA2 K129R* allele. Here, we identify several extragenic suppressors affecting the conserved hypothetical protein Msmeg_1684. We further show that an in-frame, unmarked deletion of *msmeg_1684* is also able to suppress the *secA2 K129R* allele. The function of Msmeg_1684 is unknown. However, the homologous Rv3311 protein of *M. tuberculosis*, like SecA2, is predicted by TraSH to be important for growth in macrophages (69). Intriguingly, Msmeg_1684-like proteins and SecA2 proteins frequently co-occur within the order Actinomycetales, of which mycobacteria are members. Our findings suggest that Msmeg_1684 is required for the interaction of SecA2 with the housekeeping Sec pathway and may participate directly in SecA2-dependent protein export.

In chapter 4, we describe the use of *in vivo* protein crosslinking to assess the possibility of a physical interaction between the *M. smegmatis* SecA2 and SecY proteins. The results of chapter 2 were consistent with a model in which these two proteins would interact during protein export, similar to the interaction between SecA and SecY during canonical Sec export. However, we found no evidence of a physical interaction between SecA2 and SecY. While it remains possible that these proteins do interact, and that we were simply unable to detect this interaction with the techniques used, our current results support a role for SecA2 that does not involve a direct physical interaction with SecY. Rather, our findings suggest that SecA2 interacts with the canonical Sec pathway either through an intermediary protein or at a step prior to protein export through SecY.

Our results suggest a model in which SecA2-dependent protein export is integrated into the canonical Sec pathway. SecA1 is required for SecA2-dependent export, and SecA2-dependent proteins most likely cross the cell membrane through the canonical SecYEG channel complex. While the canonical Sec components appear to be shared between Sec and SecA2 export pathways, additional components, such as Msmeg_1684, may be specific to the SecA2 system and assist in the export of SecA2-dependent proteins. SecA2 may act either to deliver a specific subset of proteins to the canonical Sec apparatus, or to participate directly in energizing export of these proteins. These results are significant in improving our understanding of the mechanism of SecA2-dependent protein export.

Table 1.1. *M. smegmatis* and *M. tuberculosis* homologs of conserved export system components.

Export system components	Required for *E. coli* growth[a]	*M. smegmatis* homolog	*M. tuberculosis* homolog	Required for *M. tuberculosis* growth[b]
SecA	yes	Msmeg_1881	Rv3240c (SecA1)	yes
		Msmeg_3654	Rv1821 (SecA2)	no*† (43)
SecY	yes	Msmeg_1483	Rv0732	yes
SecE	yes	Msmeg_1344	Rv0638	yes
SecG	no	Msmeg_3087	Rv1440	no
SecD	no	Msmeg_2961	Rv2587c	yes
SecF	no	Msmeg_2962	Rv2586c	yes
YajC	no	Msmeg_2960	Rv2588c	no
SecB	no		Rv1957‡	no
Ffh (SRP)	yes	Msmeg_2430	Rv2916c	yes
FtsY	yes	Msmeg_2424	Rv2921c	yes
YidC	yes	Msmeg_6942	Rv3921c	yes
LepB	yes	Msmeg_2441	Rv2903c	yes* (29)
LspA	yes	Msmeg_3174	Rv1539	no*† (30)

[a] information compiled from references (70-78)
[b] unless marked by *, requirement for growth is predicted by deep sequencing of transposon libraries (Tn-seq) (17)
* experimentally demonstrated
† experimentally demonstrated required for *M. tuberculosis* virulence
‡ see discussion of Rv1957 in text

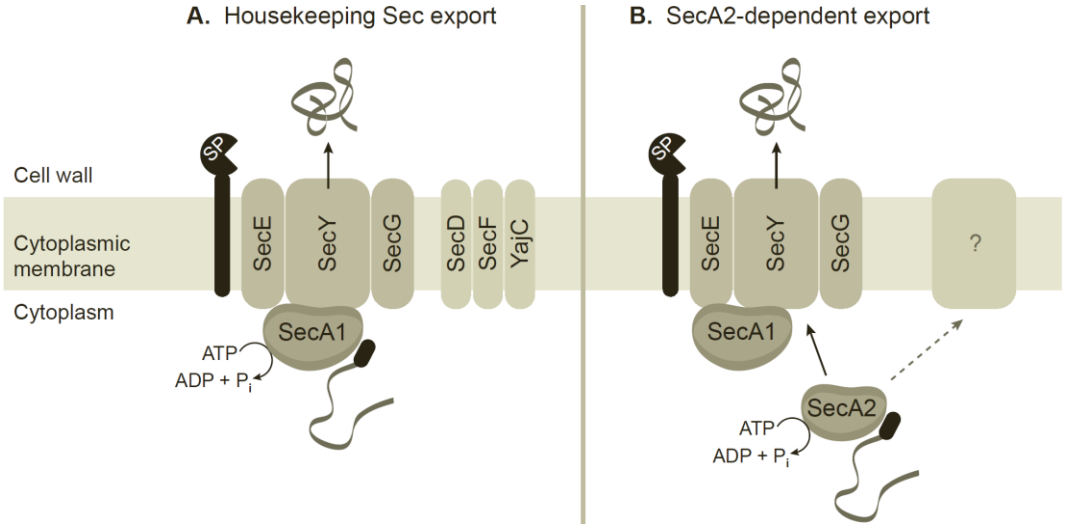

Figure 1.1. A preliminary model of the mycobacterial housekeeping Sec and SecA2-dependent protein export systems. **(A)** Preproteins (grey ribbon) with N-terminal signal peptides (black oval) are recognized by SecA1, which interacts with the SecYEG channel complex to form the translocase. SecA1 performs repeated cycles of ATP hydrolysis, pushing the unfolded preprotein through the SecYEG channel. SecD, SecF, and YajC increase efficiency of protein export. Signal peptides are removed by a LepB or LspA signal peptidase (SP), and mature proteins fold into their final conformations. **(B)** SecA2 recognizes a small subset of proteins and uses its ATPase activity to assist in their export. In the most likely scenario, SecA2 works with components of the housekeeping Sec export system and exports proteins across the cytoplasmic membrane through the SecYEG channel complex. However, it remains possible that other unknown components are required in addition to or in lieu of the housekeeping Sec components. The role of SecA2 in the export of proteins lacking signal peptides (not shown) is currently not understood.

References

1. 2012. Global Tuberculosis Report 2012. World Health Organization.

2. **Feltcher ME, Sullivan JT, Braunstein M.** 2010. Protein export systems of *Mycobacterium tuberculosis*: novel targets for drug development? Future Microbiol **5:**1581-1597.

3. **Finlay BB, Falkow S.** 1997. Common themes in microbial pathogenicity revisited. Microbiol Mol Biol Rev **61:**136-169.

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

5. **Driessen AJ, Nouwen N.** 2008. Protein translocation across the bacterial cytoplasmic membrane. Annu Rev Biochem **77:**643-667.

6. **Braun V, Wu HC.** 1994. Lipoproteins, structure, function, biosynthesis and models for protein export., p. 319-342. *In* Ghuysen JM, Hackenbeck R (ed.), Bacterial cell wall. Elsevier, Amsterdam; New York.

7. **Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ, Brinkman FS.** 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics **26:**1608-1615.

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

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

10. **van der Sluis EO, Driessen AJ.** 2006. Stepwise evolution of the Sec machinery in *Proteobacteria*. Trends Microbiol **14:**105-108.

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

12. **Cunningham K, Wickner W.** 1989. Specific recognition of the leader region of precursor proteins is required for the activation of translocation ATPase of *Escherichia coli*. Proc Natl Acad Sci U S A **86:**8630-8634.

13. **Lill R, Dowhan W, Wickner W.** 1990. The ATPase activity of SecA is regulated by acidic phospholipids, SecY, and the leader and mature domains of precursor proteins. Cell **60:**271-280.

14. **van der Wolk JP, de Wit JG, Driessen AJ.** 1997. The catalytic cycle of the *escherichia coli* SecA ATPase comprises two distinct preprotein translocation events. EMBO J **16:**7297-7304.

15. **Schiebel E, Driessen AJ, Hartl FU, Wickner W.** 1991. Delta mu H+ and ATP function at different steps of the catalytic cycle of preprotein translocase. Cell **64:**927-939.

16. **Sassetti CM, Boyd DH, Rubin EJ.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol **48:**77-84.

17. **Griffin JE, Gawronski JD, Dejesus MA, Ioerger TR, Akerley BJ, Sassetti CM.** 2011. High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. PLoS Pathog **7:**e1002251.

18. **Braunstein M, Brown AM, Kurtz S, Jacobs WR, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. J Bacteriol **183:**6979-6990.

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

20. **Rigel NW, Gibbons HS, McCann JR, McDonough JA, Kurtz S, Braunstein M.** 2009. The Accessory SecA2 System of Mycobacteria Requires ATP Binding and the Canonical SecA1. J Biol Chem **284:**9927-9936.

21. **Hou JM, D'Lima NG, Rigel NW, Gibbons HS, McCann JR, Braunstein M, Teschke CM.** 2008. ATPase activity of *Mycobacterium tuberculosis* SecA1 and SecA2 proteins and its importance for SecA2 function in macrophages. J Bacteriol **190:**4880-4887.

22. **Sharma V, Arockiasamy A, Ronning DR, Savva CG, Holzenburg A, Braunstein M, Jacobs WR, Jr., Sacchettini JC.** 2003. Crystal structure of *Mycobacterium tuberculosis* SecA, a preprotein translocating ATPase. Proc Natl Acad Sci U S A **100:**2243-2248.

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

24. **Hanada M, Nishiyama KI, Mizushima S, Tokuda H.** 1994. Reconstitution of an efficient protein translocation machinery comprising SecA and the three membrane proteins, SecY, SecE, and SecG (p12). J Biol Chem **269:**23625-23631.

25. **Van den Berg B, Clemons WM, Jr., Collinson I, Modis Y, Hartmann E, Harrison SC, Rapoport TA.** 2004. X-ray structure of a protein-conducting channel. Nature **427:**36-44.

26. **Flower AM, Hines LL, Pfennig PL.** 2000. SecG is an auxiliary component of the protein export apparatus of *Escherichia coli*. Mol Gen Genet **263:**131-136.

27. **Duong F, Wickner W.** 1997. Distinct catalytic roles of the SecYE, SecG and SecDFyajC subunits of preprotein translocase holoenzyme. EMBO J **16:**2756-2768.

28. **Paetzel M, Karla A, Strynadka NC, Dalbey RE.** 2002. Signal peptidases. Chem Rev **102:**4549-4580.

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

30. **Sander P, Rezwan M, Walker B, Rampini SK, Kroppenstedt RM, Ehlers S, Keller C, Keeble JR, Hagemeier M, Colston MJ, Springer B, Bottger EC.** 2004. Lipoprotein processing is required for virulence of *Mycobacterium tuberculosis*. Mol Microbiol **52:**1543-1552.

31. **Facey SJ, Kuhn A.** 2010. Biogenesis of bacterial inner-membrane proteins. Cell Mol Life Sci **67:**2343-2362.

32. **Egea PF, Stroud RM.** 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.

33. **Dalbey RE, Kuhn A.** 2004. YidC family members are involved in the membrane insertion, lateral integration, folding, and assembly of membrane proteins. J Cell Biol **166:**769-774.

34. **Dalbey RE, Chen M.** 2004. Sec-translocase mediated membrane protein biogenesis. Biochim Biophys Acta **1694:**37-53.

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

36. **Palaniyandi K, Veerasamy M, Narayanan S.** 2012. Characterization of Ffh of Mycobacterium tuberculosis and its interaction with 4.5S RNA. Microbiological research **167:**520-525.

37. **McCann JR, Kurtz S, Braunstein M.** 2009. Secreted and Exported Proteins Important to *Mycobacterium tuberculosis* Pathogenesis, p. 265-297. *In* Wooldridge K (ed.), Bacterial secreted proteins : secretory mechanisms and role in pathogenesis. Caister Academic Press, Wymondham.

38. **McCann JR, McDonough JA, Sullivan JT, Feltcher ME, Braunstein M.** 2011. Genome-wide identification of *Mycobacterium tuberculosis* exported proteins with roles in intracellular growth. J Bacteriol **193:**854-861.

39. **Henao-Tamayo M, Junqueira-Kipnis AP, Ordway D, Gonzales-Juarrero M, Stewart GR, Young DB, Wilkinson RJ, Basaraba RJ, Orme IM.** 2007. A mutant of *Mycobacterium tuberculosis* lacking the 19-kDa lipoprotein Rv3763 is highly attenuated in vivo but retains potent vaccinogenic properties. Vaccine **25:**7153-7159.

40. **Sulzenbacher G, Canaan S, Bordat Y, Neyrolles O, Stadthagen G, Roig-Zamboni V, Rauzier J, Maurin D, Laval F, Daffe M, Cambillau C, Gicquel B, Bourne Y, Jackson M.**

2006. LppX is a lipoprotein required for the translocation of phthiocerol dimycocerosates to the surface of *Mycobacterium tuberculosis*. EMBO J **25:**1436-1444.

41. **Bigi F, Gioffre A, Klepp L, Santangelo MP, Alito A, Caimi K, Meikle V, Zumarraga M, Taboga O, Romano MI, Cataldi A.** 2004. The knockout of the *lprG-Rv1410* operon produces strong attenuation of *Mycobacterium tuberculosis*. Microbes Infect **6:**182-187.

42. **Owens MU, Swords WE, Schmidt MG, King CH, Quinn FD.** 2002. Cloning, expression, and functional characterization of the *Mycobacterium tuberculosis secA* gene. FEMS Microbiol Lett **211:**133-141.

43. **Braunstein M, Espinosa BJ, Chan J, Belisle JT, Jacobs WR, Jr.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of Mycobacterium tuberculosis. Mol Microbiol **48:**453-464.

44. **Watkins BY, Joshi SA, Ball DA, Leggett H, Park S, Kim J, Austin CD, Paler-Martinez A, Xu M, Downing KH, Brown EJ.** 2012. Mycobacterium marinum SecA2 promotes stable granulomas and induces tumor necrosis factor alpha in vivo. Infect Immun **80:**3512-3520.

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

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

47. **Feltcher ME, Gibbons HS, Ligon LS, Braunstein M.** 2013. Protein export by the mycobacterial SecA2 system is determined by the preprotein mature domain. J. Bacteriol. **195:**672-681.

48. **Hinchey J, Lee S, Jeon BY, Basaraba RJ, Venkataswamy MM, Chen B, Chan J, Braunstein M, Orme IM, Derrick SC, Morris SL, Jacobs WR, Jr., Porcelli SA.** 2007. Enhanced priming of adaptive immunity by a proapoptotic mutant of *Mycobacterium tuberculosis*. J Clin Invest **117:**2279-2288.

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

50. **Bensing BA, Sullam PM.** 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.

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

52. **Siboo IR, Chaffin DO, Rubens CE, Sullam PM.** 2008. Characterization of the accessory Sec system of *Staphylococcus aureus*. J Bacteriol **190:**6188-6196.

53. **Takamatsu D, Bensing BA, Sullam PM.** 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.

54. **Takamatsu D, Bensing BA, Sullam PM.** 2005. Two additional components of the accessory sec system mediating export of the Streptococcus gordonii platelet-binding protein GspB. J Bacteriol **187:**3878-3883.

55. **Bensing BA, Gibson BW, Sullam PM.** 2004. The Streptococcus gordonii platelet binding protein GspB undergoes glycosylation independently of export. J Bacteriol **186:**638-645.

56. **Mistou MY, Dramsi S, Brega S, Poyart C, Trieu-Cuot P.** 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.

57. **Wu H, Zeng M, Fives-Taylor P.** 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.

58. **Zhou M, Zhu F, Dong S, Pritchard DG, Wu H.** 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.

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

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

61. **Dempsey BR, Wrona M, Moulin JM, Gloor GB, Jalilehvand F, Lajoie G, Shaw GS, Shilton BH.** 2004. Solution NMR structure and X-ray absorption analysis of the C-terminal zinc-binding domain of the SecA ATPase. Biochemistry **43:**9361-9371.

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

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

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

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

66. **Fagan RP, Fairweather NF.** 2011. *Clostridium difficile* has two parallel and essential Sec secretion systems. J. Biol. Chem. **286:**27483-27493.

67. **Caspers M, Freudl R.** 2008. *Corynebacterium glutamicum* possesses two *secA* homologous genes that are essential for viability. Archives of microbiology **189:**605-610.

68. **Sullivan JT, Young EF, McCann JR, Braunstein M.** 2012. The Mycobacterium tuberculosis SecA2 system subverts phagosome maturation to promote growth in macrophages. Infect Immun **80:**996-1006.

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

70. **Murphy CK, Beckwith J.** 1996. Export of Proteins to the Cell Envelope in *Escherichia coli*, p. 967-978. *In* Neidhardt FC (ed.), *Escherichia coli* and *Salmonella* : cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.

71. **Shiba K, Ito K, Yura T, Cerretti DP.** 1984. A defined mutation in the protein export gene within the *spc* ribosomal protein operon of *Escherichia coli*: isolation and characterization of a new temperature-sensitive *secY* mutant. EMBO J **3:**631-635.

72. **Pogliano KJ, Beckwith J.** 1994. Genetic and molecular characterization of the *Escherichia coli secD* operon and its products. J Bacteriol **176:**804-814.

73. **Phillips GJ, Silhavy TJ.** 1992. The *E. coli ffh* gene is necessary for viability and efficient protein export. Nature **359:**744-746.

74. **Gill DR, Salmond GP.** 1990. The identification of the *Escherichia coli ftsY* gene product: an unusual protein. Mol Microbiol **4:**575-583.

75. **van Bloois E, Nagamori S, Koningstein G, Ullers RS, Preuss M, Oudega B, Harms N, Kaback HR, Herrmann JM, Luirink J.** 2005. The Sec-independent function of *Escherichia coli* YidC is evolutionary-conserved and essential. J Biol Chem **280:**12996-13003.

76. **Bogsch EG, Sargent F, Stanley NR, Berks BC, Robinson C, Palmer T.** 1998. An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. J Biol Chem **273:**18003-18006.

77. **Sargent F, Bogsch EG, Stanley NR, Wexler M, Robinson C, Berks BC, Palmer T.** 1998. Overlapping functions of components of a bacterial Sec-independent protein export pathway. EMBO J **17:**3640-3650.

78. **Sargent F, Stanley NR, Berks BC, Palmer T.** 1999. Sec-independent protein translocation in *Escherichia coli*. A distinct and pivotal role for the TatB protein. J Biol Chem **274:**36073-36082.

CHAPTER 2

Suppressor Analysis Reveals a Role for SecY in the SecA2-Dependent Protein Export Pathway of Mycobacteria[1]

All bacteria use the conserved Sec pathway to transport proteins across the cytoplasmic membrane, with the SecA ATPase playing a central role in the process. Mycobacteria are part of a small group of bacteria that have two SecA proteins: the canonical SecA (SecA1) and a second, specialized SecA (SecA2). The SecA2-dependent pathway exports a small subset of proteins and is required for *Mycobacterium tuberculosis* virulence. The mechanism by which SecA2 drives export of proteins across the cytoplasmic membrane remains poorly understood. Here we performed suppressor analysis on a dominant negative *secA2* mutant (*secA2 K129R*) of the model mycobacterium *Mycobacterium smegmatis* to better understand the pathway used by SecA2 to export proteins. Two extragenic suppressor mutations were identified as mapping to the promoter region of *secY*, which encodes the central component of the canonical Sec export channel. These suppressor mutations increased *secY* expression, and this effect was sufficient to alleviate *secA2 K129R* phenotypes. We also discovered that the level of SecY protein was greatly diminished in the *secA2 K129R* mutant, but at least partially restored in the suppressors. Furthermore, the level of SecY in a suppressor strongly correlated with the degree of

[1]Adapted for this dissertation from: Ligon LS, Rigel NW, Romanchuk A, Jones CD, Braunstein M. Suppressor Analysis Reveals a Role for SecY in the SecA2-Dependent Protein Export Pathway of Mycobacteria. J Bacteriol (*in submission*). Additional contributing author: Brittany Miller (Department of Microbiology and Immunology, School of Medicine, University of North Carolina, Chapel Hill, North Carolina, 27599-7290)

suppression. Our findings reveal a detrimental effect of SecA2 K129R on SecY, arguing for an integrated system in which SecA2 works with SecY and the canonical Sec translocase to export proteins.

Introduction

Bacteria use both conserved and specialized protein export systems to deliver proteins to the bacterial cell surface and to the extracellular environment. These exported proteins are important for critical cell functions like nutrient acquisition, cell structure, and in the case of pathogens, virulence. All bacteria possess the conserved Sec protein export system, which performs the bulk of "housekeeping" protein export. More recently, it was discovered that mycobacteria and some Gram-positive bacteria possess a second, specialized Sec export system (1-7). This specialized protein export system is only partially understood, but in pathogens is often required for virulence.

In the housekeeping Sec export system, the motor protein SecA plays a central role in exporting unfolded proteins through a membrane channel composed of the integral membrane proteins SecY, SecE, and SecG, where SecY is the major component of the translocon (8). SecA is a cytoplasmic ATPase that provides energy for protein export through successive rounds of ATP binding and hydrolysis (9, 10). Proteins exported by the Sec system possess an N-terminal signal peptide, which is cleaved from the mature protein following export. Because many important proteins rely on the Sec system for export, many Sec proteins are essential, including SecA and SecY (6, 11). Bacteria containing a second, specialized Sec export system are characterized by the presence of two SecA proteins with distinct functions. The second SecA protein, called SecA2, also possesses ATPase activity (12, 13) and is generally non-essential (14).

In mycobacteria, SecA2 is responsible for the export of a small subset of proteins (15, 16) and is required for virulence of the pathogen *Mycobacterium tuberculosis* (16). The model organism *Mycobacterium smegmatis* also possesses a SecA2 system, which is functionally conserved with that of *M. tuberculosis* (6, 17). Studies with *M. smegmatis* show that SecA2-dependent proteins can contain N-terminal signal peptides indistinguishable from classical Sec signal peptides and that it is the mature domain of the protein that determines the requirement for SecA2 (18). However, the mechanism of SecA2-dependent protein export in mycobacteria is unknown. Some SecA2 export systems, referred to as "SecA2-SecY2" systems, include a second distinct SecY protein (named SecY2) that is required for export of SecA2-dependent proteins. In these systems, SecY2 is presumed to function as the channel for protein translocation across the membrane (5, 7, 14). In mycobacteria, there is no SecY2 protein evident, making it unclear how SecA2-dependent proteins are exported in this system. One possibility is that the mycobacterial SecA2 system exports proteins through the SecYEG channel, sharing this channel with the housekeeping Sec export system; however, this model has not been proven.

To increase our understanding of SecA2 export and identify additional components of the SecA2 system, we carried out a suppressor analysis of a *secA2* mutant of *M. smegmatis*. For this purpose, we used a *secA2 K129R* mutant that encodes for a SecA2 protein with an amino acid substitution (K129R) in the Walker box, which is an amino acid motif important for ATP binding and hydrolysis. We previously demonstrated that this K129R substitution disrupts the ATPase activity of SecA2, rendering it non-functional (12, 17). In addition, SecA2 K129R has a dominant negative effect on wild type SecA2 (17). This suggests that SecA2 K129R is still able to interact with its normal binding partners within the cell, but because it is non-functional, interferes with the function of these partners. Further, SecA2 K129R must disrupt an important process in the cell, because the phenotypes of the *secA2 K129R* allele are worse than those of a

secA2 deletion mutant. These properties make the *secA2 K129R* allele a good starting point for suppressor analysis.

In the present study, we characterized two extragenic suppressor mutations of *secA2 K129R*, both of which are located in the promoter region of the only *secY* gene of mycobacteria. We also discovered that SecY protein levels are drastically reduced in the *secA2 K129R* mutant, and that these suppressor mutations increase *secY* expression to partially restore SecY protein levels and suppress *secA2 K129R* phenotypes. Finally, we found that SecY levels were increased relative to the *secA2 K129R* strain in six additional extragenic suppressors. Taken together, these findings suggest that SecA2 K129R disrupts the housekeeping Sec export system, causing SecY degradation and thus the severe phenotypes of the *secA2 K129R* mutant. The data presented here argues for SecA2 working in concert with SecY and the housekeeping Sec pathway to export SecA2-dependent proteins.

Materials and Methods

Bacterial strains and culture conditions. *M. smegmatis* strains used in this study are described in Table 2.1, and were grown at 37°C or 30°C, using Middlebrook 7H9/7H10 or Mueller-Hinton media. To limit acquisition of suppressors when working with the *secA2 K129R* strain (NR178), starter cultures were generally grown at 30°C, but the actual experiments performed at 37°C. Media were supplemented with 0.5% glycerol plus 0.2% glucose (7H9/7H10 media only) and 0.1% tween-80 (all media). Antibiotics kanamycin (20 µg/mL) and hygromycin B (50 µg/mL) were added as needed. When required, plasmids were introduced into *M. smegmatis* strains by electroporation (19). *Escherichia coli* strains were grown at 37°C in

lysogeny broth or on lysogeny broth agar (otherwise known as Luria-Bertani media). Antibiotics kanamycin (40 μg/mL) and hygromycin B (150 μg/mL) were added as needed.

Suppressor collection. The *secA2 K129R* strain encodes for a SecA2 protein in which the lysine at position 129 is substituted with an arginine (amino acid numbering based on NCBI GenBank accession number AF287049 (6)). Suppressors of *secA2 K129R* were isolated by plating 38 independently grown cultures of the *secA2 K129R* strain onto Mueller-Hinton agar at 37°C. Spontaneous suppressors (i.e. colonies that grew on Mueller-Hinton agar) were obtained, and one small (S) and one large (B) suppressor colony was chosen from each independent culture.

Plasmid construction. Plasmids used in this study are described in Table 2.2 and primers are described in Table 2.3. In all cases, newly constructed plasmids were sequenced. To create *secY'-'lacZ* fusion plasmids, a region upstream of *secY* was PCR amplified from *M. smegmatis* genomic DNA of strains NR178 and NR236 (suppressor 4S) to generate wild type and 4S *secY* promoter sequences, respectively. Each PCR product contained 379 bp upstream of *secY* along with 34 bp of *secY* coding sequence and engineered EcoRI restriction sites, and was cloned into pCR2.1-TOPO (Invitrogen), yielding plasmids pLL5 and pLL6. The EcoRI *secY* promoter fragments were cut from pLL5 and pLL6 and ligated into EcoRI-digested pCV125, yielding plasmids pLL11 and pLL8, which contain *secY'-'lacZ* translational fusions. pLL11 was mutated by site-directed mutagenesis (Stratagene QuikChange II) to re-create the mutation found upstream of *secY* in suppressor NR151 (24S), yielding plasmid pLL15.

To construct integrating *secY* expression plasmids, plasmid pYUB2063 was first digested with PciI and re-ligated to create a smaller plasmid of 5198 bp in size named pLL2. This plasmid integrates at the Tweety mycobacteriophage *attB* site in the *M. smegmatis* genome (20). The entire *secY* gene along with its upstream promoter was PCR amplified from NR178 and NR236

genomic DNA, with engineered NotI and EcoRV restriction sites for cloning purposes. Each PCR product was cloned into pCR2.1-TOPO (Invitrogen), yielding plasmids pLL17 and pLL19. The NotI-EcoRV *secY* fragments were cut from pLL17 and pLL19 and ligated into NotI-EcoRV-digested pLL2, yielding plasmids pLL21 and pLL23.

Azide sensitivity assays. Cultures were plated for azide sensitivity as previously described (17), by mixing 200 µL of a saturated culture with 7H9 top agar and pouring over 7H10 agar plates (with tween supplementation omitted) in three technical replicates. Discs soaked with 10 µL of 0.15 M sodium azide were added to the center of each plate. The diameter of the zone of inhibition was measured after two days and presented as a percentage of the entire plate diameter, yielding percent azide inhibition.

Whole genome and directed sequencing. Genomic DNA was isolated (as described previously (21)) from the 4S suppressor strain (NR236) and submitted for whole genome sequencing at the High-Throughput Sequencing Facility at the University of North Carolina at Chapel Hill. Sequencing was performed using Illumina GA II technology with 36 bp single-end reads. Reads were aligned to the *M. smegmatis* mc^2155 reference genome (NCBI RefSeq accession number NC_008596.1) using SOAP (22) (with default parameters), resulting in an average sequence coverage of 47.5. Single nucleotide polymorphisms, insertions, and deletions were located using SOAP and BLAT (23) (default parameters), resulting in a total of 86 mutations identified. Mutations also identified in other mc^2155-derived strains sequenced in our laboratory were discarded as background mutations (i.e. mutations already present in the parent strain), leaving 15 mutations appearing unique to the 4S strain. Following further confirmatory sequencing of PCR amplified products (Eton Bioscience, Inc. or Genewiz, Inc.), all but one of these mutations were eliminated as either background or false positives.

5' RLM-RACE. RNA ligase-mediated rapid amplification of cDNA ends (RLM-RACE) was performed using the GeneRacer kit (Invitrogen) according to manufacturer instructions. Briefly, total RNA was isolated from strain NR172, treated with TAP to remove 5' triphosphates, and ligated to the GeneRacer RNA oligo. *secY* 5' ends were reverse-transcribed to cDNA using a gene-specific primer located within the *secY* coding sequence (primer R1, see Fig. 2.2B). 5' end cDNA was then amplified using a nested PCR strategy involving a first round of PCR with a second gene-specific primer located within the *secY* coding sequence (primer R2) and the GeneRacer 5' primer (homologous to the GeneRacer RNA oligo), followed by a second round of PCR with a primer located upstream of the *secY* coding sequence (primer R3) and the GeneRacer 5' primer. PCR products were separated by agarose gel electrophoresis, individually gel purified, and cloned into pCR2.1-TOPO (Invitrogen). A minimum of 4 clones originating from each PCR product were submitted to Genewiz, Inc. for sequencing. Transcript 5' ends were identified as the first nucleotide following the sequence of the GeneRacer RNA oligo. Nucleotide positions were numbered relative to the translational start site of *secY* (NCBI GenBank accession number ABK75688).

Quantitative RT-PCR. Strains were grown in 7H9 broth to an $OD_{600\,nm}$ of approximately 1.0, pelleted by centrifugation at 1,600 x g for 10 minutes, and flash-frozen. RNA was isolated using one of two protocols. For both protocols, bacteria were lysed in 1 mL 3:1 chloroform:methanol, and then vortexed with 5 mL Trizol and incubated 10 minutes at room temperature. Phases were separated by centrifugation at 1,600 x g for 15 minutes at 4°C. For the first protocol, the upper phase was mixed with 1x volume of isopropanol to precipitate overnight at 4°C. RNA was pelleted by centrifugation at 20,000 x g for 30 minutes at 4°C, washed twice with cold 70% ethanol, and resuspended in RNAse-free water. For the second protocol, the upper phase was mixed with 0.625x volume of 95% ethanol and column-purified

(Promega SV Total RNA Isolation System). All RNA samples were treated with DNAse (Promega) and then column-purified (Zymo DNA-Free RNA kit). Following RNA isolation, quantitative RT-PCR was performed in triplicate technical replicates using 25 or 50 ng RNA in each reaction. Products were reverse transcribed and amplified from total RNA using the Bioline SensiMix SYBR & Fluorescein One-Step Kit, and amplified from a DNA standard using the Bioline SensiMix SYBR & Fluorescein Kit. For quantitative RT-PCR on *secY*, products were amplified from the 5' ends of *secY* and *rpoB* gene sequences, starting quantity of each transcript was calculated relative to DNA standards, and *secY* transcripts were normalized to *rpoB* transcripts (as a control) in each sample (24).

LacZ (β-galactosidase) activity assays. LacZ activity assays were performed similar to those described previously (25). Each strain was grown in 7H9 broth to late-exponential or saturated phase and 800 µL was pelleted in a microcentrifuge. Cell pellets were resuspended in 800 µL Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 50 mM β-mercaptoethanol), then lysed with 35 µL chloroform and 1 µL 0.1% SDS, and vortexed for 30 seconds. 160 µL ONPG (4 mg/mL in Z buffer) was added to each reaction and incubated for 1 hour at room temperature. Reactions were terminated by addition of 400 µL of 1 M Na_2CO_3. Debris was removed by centrifugation at 16,000 x g for 3 minutes, and the $OD_{420\ nm}$ was read from the supernatant. LacZ activity (Miller units) was calculated by the formula: (1,000 x $OD_{420\ nm}$) / (reaction time (minutes) x culture volume used in reaction (mL) x $OD_{600\ nm}$).

Subcellular fractionation and western blotting. Each strain was grown in Mueller-Hinton broth to an $OD_{600\ nm}$ of approximately 1.0. Subcellular fractions were prepared as previously described (15). Briefly, cell pellets were lysed by passage through a French press. Cell debris was removed by centrifugation at 1,600 x g for 30 minutes and clarified whole cell

lysates (supernatants) were standardized for equal protein content using the Pierce BCA Protein Assay. Clarified whole cell lysates were centrifuged at 100,000 x g for 2 hours to separate cell envelope (pellet) and soluble (supernatant) fractions. Alternately, clarified whole cell lysates were centrifuged at 50,000 x g for 30 minutes to separate cell wall (pellet) and the resulting supernatant was centrifuged at 100,000 x g for 2 hours to separate cell membrane (pellet) and soluble (supernatant) fractions. In some cases, 1% Igepal CA-630 was added during cell envelope resuspension to facilitate solubilization of SecY.

Whole cell lysates (WCL, loaded for equal protein content) or subcellular fractions (loaded for equivalent starting cell material to WCL samples) were boiled, separated by SDS-PAGE, and transferred to nitrocellulose membranes for western blotting. Proteins were detected using a rabbit αSecY polyclonal antiserum at a 1:150 dilution, a rabbit αSecA1 polyclonal antiserum at a 1:50,000 dilution (26), and a mouse αHA monoclonal antibody at a 1:10,000 dilution (Covance). Primary antibodies were detected using alkaline phosphatase-conjugated goat anti-rabbit or goat anti-mouse IgG secondary antibodies, ECF substrate (GE Healthcare), and a Molecular Dynamics Storm 860 phosphorimager or Syngene G:BOX machine. To measure localization of Msmeg_1712-HA in subcellular fractions, fluorescence was quantified using GeneTools software (Syngene) and the percent located in each fraction was calculated relative to the total fluorescence of cell wall + cell membrane + soluble fractions.

SecY antiserum production. Anti-SecY antiserum was produced for this work by Yenzym Antibodies, LLC, and was raised in rabbits to a peptide corresponding to the C-terminal 16 amino acids of *M. smegmatis* SecY (QIESQLMQRNYEGFLK), using TiterMax Gold adjuvant. Affinity purified antibody was used in this study.

Construction of SecA1 depletion strain. The SecA1 depletion strain, JM693, was created as described previously (26). Briefly, wild type *M. smegmatis* mc^2155 was transformed with the pKIsecA1 suicide plasmid (gift of Dr. Sabine Ehrt, Weill Cornell Medical College) and transformants were selected by plating on media containing hygromycin B. Integration of pKIsecA1 at the native *secA1* locus in the chromosome was confirmed by southern blot. Regulation of SecA1 in strain JM693 was achieved by introducing plasmid pNR55 (expressing the revTetR repressor) and growing cultures in the dark with anhydrotetracycline (1200 ng/mL) supplementation. This level of anhydrotetracycline was sufficient to completely deplete SecA1 protein levels (as assessed by western blotting) without impacting cell growth.

Results

Spontaneous suppressor mutations alleviate the severe phenotypes of the *secA2 K129R* allele. Compared to wild type *M. smegmatis*, a *secA2* deletion mutant exhibits a moderate growth defect on Mueller-Hinton agar and increased sensitivity to azide (6). This phenotype is complemented by adding the wild type *secA2* allele, expressed from the constitutive *hsp60* promoter, into the *secA2* deletion mutant (Fig. 2.1A) (17). However, introduction of the *secA2 K129R* allele, also expressed from the *hsp60* promoter, into the *secA2* deletion mutant fails to complement *secA2* deletion phenotypes. SecA2 K129R is non-functional as a result of an amino acid substitution in the ATP binding Walker box. Moreover, the *secA2 K129R* allele exacerbates both the growth defect on Mueller-Hinton agar and sensitivity to azide phenotypes (i.e. *secA2 K129R* phenotypes are more severe than phenotypes of the *secA2* deletion) (Fig. 2.1B) (17). Reported previously, the *secA2 K129R* allele is also dominant negative, exhibiting phenotypes even in the presence of a wild type *secA2* allele (17).

Suppressor mutations spontaneously arise in the *secA2 K129R* background, and are easily identified by their ability to grow on Mueller-Hinton agar. To begin our suppressor analysis, we collected 63 suppressor mutants. All suppressors alleviated the severe phenotypes associated with the *secA2 K129R* allele, improving both growth on Mueller-Hinton agar and resistance to azide (Fig. 2.1B). Throughout this study, we routinely compared suppressor strains to the complemented *secA2* strain (NR172) as a control, instead of the wild type strain. By doing so, we could study the effect of the suppressor mutations across a series of strains that all express *secA2* alleles from the same *hsp60* promoter. In no case did suppressors rescue phenotypes to the level of the complemented *secA2* strain. Instead, suppressor phenotypes more closely resembled the phenotypes of the *secA2* deletion strain, suggesting that suppressor mutations overcome the detrimental effect of SecA2 K129R, but do not necessarily restore SecA2-dependent export. Of the 63 suppressors isolated, 40 did not produce SecA2 K129R protein (by western blot, data not shown), suggesting that these suppressors alleviate *secA2 K129R* phenotypes by preventing SecA2 K129R production. This is a category of suppressors that we expected to observe. The remaining 23 suppressors produced detectable levels of full-length SecA2 K129R protein.

We next sequenced the *secA2 K129R* allele in the 23 suppressors producing SecA2 K129R protein. Of these 23 suppressors, 9 strains contained mutations in the *secA2 K129R* allele (intragenic), while the remaining 14 strains had no mutations in the *secA2 K129R* allele and are therefore extragenic suppressors. Because our goal was to identify proteins that work with SecA2 during export, we focused on characterizing extragenic suppressors.

Suppressors 4S and 24S contain mutations upstream of *secY*. We chose a single extragenic suppressor with normal SecA2 K129R protein levels, identified as 4S (strain NR236),

and performed whole genome sequencing. Alignment of sequence reads to the *M. smegmatis* mc^2155 reference genome resulted in average sequence coverage of 47.5, sufficient to identify mutations in the 4S genome. Each candidate suppressor mutation was PCR amplified and re-sequenced from both 4S and its parent *secA2 K129R* strain. Following this directed sequencing, false positives and background mutations were discarded, leaving only a single mutation uniquely present in suppressor 4S. This unique mutation was a 2 bp insertion located 137 bp upstream of the translational start site for the predicted *M. smegmatis secY* gene, *msmeg_1483* (Fig. 2.2A). The *msmeg_1483* open reading frame is the only *secY* homolog in *M. smegmatis* (43% identical and 63% similar at the amino acid level to the canonical *E. coli* SecY).

To determine whether any other extragenic suppressors in our collection contained mutations upstream of *secY*, we sequenced approximately 1000 bp upstream of *secY* in the 13 additional extragenic suppressor strains. One additional suppressor, identified as 24S (strain NR151), contained a single C→G polymorphism located 134 bp upstream of the translational start site for *secY* (Fig. 2.2A). The close proximity of the 4S and 24S mutations suggests that they act in a similar manner to suppress the *secA2 K129R* allele.

4S and 24S suppressor mutations are located within the *secY* promoter. Because the mutations identified in suppressors 4S and 24S are located upstream of the *secY* coding sequence, we hypothesized that they would map to the *secY* promoter and suppress the *secA2 K129R* phenotypes through an effect on *secY* transcription. Mycobacterial promoters do not always resemble the classical promoter structure defined for *E. coli* and other bacteria, and it can be difficult to predict their location. However, manual inspection of the sequence upstream of *secY* revealed two potential *M. smegmatis* -10 regions, each beginning with the highly conserved nucleotides “TA” (27) and matching at least 3/6 nucleotides of the -10 region

consensus sequence. These two -10 regions, positioned at nucleotides -132 and -140 upstream of the *secY* start codon, are the closest potential -10 regions to *secY*, and are located in the same region as the 4S and 24S mutations (Fig. 2.2A). No obvious -35 region was found, but this is not surprising given that mycobacteria have low homology at this region (27), and may not even require the -35 region for promoter function (28).

The location of the potential promoter -10 regions suggests that the 4S and 24S mutations could directly impact transcription. The 4S mutation is an insertion that improves the -10 region at nucleotide -140 from a 3/6 nucleotide match to the -10 region consensus to a 4/6 nucleotide match. The 24S mutation is a C→G polymorphism upstream of the -10 region at nucleotide -132. Interestingly, this mutation creates a "TGn" motif upstream of the -10 region. Addition of a "TGn" motif to a -10 region can result in a 3-5 fold increase in transcription (29).

To determine the location of the *secY* transcriptional start site in relation to the 4S and 24S mutations, we performed 5' RLM-RACE on *M. smegmatis* RNA (using primer R1, Fig. 2.2B). Following the RACE procedure, we PCR amplified the 5' ends of cDNA created from *secY* transcripts, using a nested PCR strategy. The first round of PCR (using primer R2) amplified out from the *secY* coding sequence and yielded multiple products. A second round of PCR with a primer upstream of *secY* (primer R3) narrowed in on the longest *secY* transcripts. This nested PCR amplification produced two products (Fig. 2.2B) that we cloned and sequenced to determine the 5' end of each species. The longer of the two species represents *secY* transcripts with a 5' end located 127-129 bp upstream of the *secY* translational start site, and the shorter of the two species represents *secY* transcripts with a 5' end located 75 bp upstream of *secY*. The transcripts beginning 127-129 bp upstream of *secY* were the longest species detected, making it likely that they represent true transcriptional start sites. This start site region corresponds

perfectly with the potential -10 regions we identified upstream of *secY*, and supports the conclusion that the 4S and 24S suppressor mutations are located within the *secY* promoter. The transcripts beginning 75 bp upstream of *secY* could result from an alternate transcriptional start site, although there are no obvious promoter sequences located adjacent to this site. Alternatively, these shorter transcripts could be a result of transcript processing/degradation at this location.

4S and 24S suppressor mutations cause increased *secY* transcript levels. Given the location of the 4S and 24S mutations, we next tested whether the mutations affect *secY* expression by measuring *secY* transcript levels in the complemented *secA2*, *secA2* deletion mutant, *secA2 K129R* mutant, and 4S and 24S suppressor strains, along with additional suppressor strains 7S and 20B for comparison. Total RNA was isolated from each strain and was analyzed by quantitative RT-PCR (Fig. 2.3A). There was no significant difference in *secY* transcript levels between the complemented *secA2*, *secA2* deletion mutant, *secA2 K129R* mutant, and 7S and 20B suppressor strains. However, relative to the *secA2 K129R* parent strain, both suppressors 4S and 24S exhibited a reproducible increase in *secY* transcript levels, although this result is statistically significant only for suppressor 24S.

We also tested the effects of the 4S and 24S mutations in the absence of the *secA2 K129R* allele (i.e. in a wild type background) using *secY'-'lacZ* fusions. To do this, we constructed plasmids containing 379 bp upstream of *secY* plus 34 bp of *secY* coding sequence translationally fused to *lacZ*, introduced these plasmids into wild type *M. smegmatis* (mc^2155), and measured LacZ (β-galactosidase) activity as a readout of *secY* expression (Fig. 2.3B). Relative to the wild type *secY* promoter, constructs containing the 4S or 24S *secY* promoters produced significantly

higher levels of LacZ activity. These results reinforce the conclusion that the 4S and 24S mutations increase *secY* expression.

SecY protein is undetectable in the *secA2 K129R* strain, but recovered in 4S and 24S suppressor strains. We next tested for an increase in SecY protein levels in the 4S and 24S suppressors. To measure SecY protein, we generated an antibody against a peptide matching the C-terminus of *M. smegmatis* SecY, and performed αSecY western blots on whole cell lysates (WCL) and on subcellular fractions containing the cell envelope (ENV) or the cytoplasm (SOL). The purity of the subcellular fractions was confirmed by western blot analysis with antibodies to the cell envelope protein MspA and the cytoplasmic protein GroEL1 (data not shown). αSecY western blotting on the WCL of the complemented *secA2* strain detected a prominent protein species at 37 kDa (Fig. 2.4). Although the predicted size of *M. smegmatis* SecY is 48 kDa, full-length SecY in *E. coli* (30) and *Synechococcus PCC7942* (31) is also observed to migrate aberrantly at about 37 kDa. Additionally, lower molecular weight species were sometimes apparent on the αSecY western blots, which are likely degradation products similar to those detected for *E. coli* SecY (8, 32). In support of the 37 kDa species being full-length SecY, strains expressing higher levels of *secY* exhibited an increase in this product (as discussed below). In addition, this full-length *M. smegmatis* SecY species is detected almost exclusively in the cell envelope (ENV) fraction, as expected for an integral membrane protein.

Next, we performed αSecY western blots on complemented *secA2*, *secA2* deletion mutant, *secA2 K129R* mutant, and 4S and 24S suppressor strains (Fig. 2.4). Both complemented *secA2* and *secA2* deletion mutant strains exhibited a strong signal for full-length SecY protein. In stark contrast, we observed a complete loss of detectable full-length SecY protein in the *secA2 K129R* mutant, revealing a link between SecA2 K129R and SecY. As there was no transcriptional

effect on *secY* observed in the *secA2 K129R* strain (Fig. 2.3A), the change in SecY protein levels in the *secA2 K129R* mutant does not appear to be caused by a difference in *secY* expression. An alternate possibility is that stability of the SecY protein is reduced by SecA2 K129R. This explanation is supported by an increase in the SecY degradation products observed in the *secA2 K129R* strain. In addition, experiments in *E. coli* show that jamming export through the SecYEG channel results in degradation of SecY (33). Therefore, the most likely explanation for the lack of detectable SecY protein in the *secA2 K129R* strain is that SecA2 K129R causes stress on the housekeeping SecY export channel, resulting in drastic SecY degradation.

Relative to the *secA2 K129R* strain, suppressors 4S and 24S exhibited increased levels of full-length SecY protein. While the SecY level in 4S and 24S was still lower than that in the complemented *secA2* or *secA2* deletion mutant strains, the increase relative to the *secA2 K129R* strain is consistent with the increased transcription of *secY* (Fig. 2.3A). This increase in *secY* expression would not be expected to prevent the SecY degradation caused by SecA2 K129R. In fact, the presumed SecY degradation products remained high in both suppressor strains. Thus, it appears that the 4S and 24S mutations are able to suppress *secA2 K129R* by altering the balance between *secY* expression and SecY degradation, shifting the steady-state SecY protein level high enough to allow the recovered growth seen in these strains.

Increased *secY* expression is sufficient to suppress *secA2 K129R*. The above results demonstrate that the 4S and 24S mutations cause increased expression of *secY*. However, the question remained whether this relatively small increase in *secY* expression alone was sufficient to suppress the *secA2 K129R* phenotypes. To address this question, we created integrating plasmids carrying the *secY* gene under the control of its native promoter, either with no mutation (wild type) or carrying the 4S mutation upstream of *secY*. These plasmids were then

introduced into the *secA2 K129R* strain, creating merodiploid strains that contain the endogenous *secY* gene at its chromosomal location and a second copy of *secY* integrated in the chromosome at the Tweety mycobacteriophage attachment site. Addition of an extra copy of *secY*, whether driven by the 4S promoter or the native promoter, was indeed sufficient to suppress *secA2 K129R* phenotypes, and both versions suppressed *secA2 K129R* equally as well as the original 4S suppressor (Fig. 2.5A). In support of this result, western blotting revealed that when a second copy of *secY* was introduced into the *secA2 K129R* strain, the level of full-length SecY protein increased to the same level as seen in the 4S suppressor strain (Fig. 2.5B). These results indicate that the increased *secY* expression observed in the 4S and 24S suppressor strains can indeed explain their ability to suppress *secA2 K129R* phenotypes.

Additional extragenic suppressors are also associated with increased SecY levels.

Given the above results, we selected six additional extragenic suppressors and tested for an effect on SecY. These six suppressors expressed normal levels of SecA2 K129R protein and did not carry mutations in the *secY* gene or promoter (data not shown). Relative to the *secA2 K129R* mutant, each of these suppressors exhibited improved azide resistance and growth on Mueller-Hinton agar (Fig. 2.6A), with the degree of suppression varying from strain to strain. Western blotting for SecY revealed that all six additional suppressors have an increased steady-state level of SecY protein relative to the parent *secA2 K129R* strain (Fig. 2.6B). However, unlike the case with the 4S and 24S suppressors, none of these six additional suppressors had altered *secY* transcription, as assessed by quantitative RT-PCR (Fig. 2.6C). Therefore, these suppressors must affect SecY levels in a completely different manner.

The finding that multiple extragenic suppressors affect the SecY protein level, despite differences in their genotypes and phenotypes, reinforces the conclusion that SecA2 K129R has

a detrimental effect on SecY. Furthermore, the level of SecY protein observed correlates with the phenotype of the respective strain. The *secA2 K129R* mutant had no detectable SecY protein, little resistance to azide, and very poor growth on Mueller-Hinton agar. Intermediate suppressors such as 4S, 24S, 1S, and 3S had low but detectable SecY protein levels and moderate resistance to azide and growth on Mueller-Hinton agar. Strong suppressors such as 29S, 33S, 33B, and 10S had near wild type levels of SecY protein and the strongest azide resistance and growth on Mueller-Hinton agar observed, comparable to the phenotypes of the *secA2* deletion mutant. Based on this pattern, we conclude that a critical problem caused by the presence of SecA2 K129R is reduced SecY protein level, and that extragenic suppressor mutations can correct this problem by increasing the amount of SecY protein to a functional level.

SecA1 is required for SecA2-dependent protein export. We have previously shown that depletion of SecA1 in *M. smegmatis* leads to reduced SecA2-dependent export, suggesting a requirement for SecA1 in the SecA2 export pathway (17). However, in light of our findings that SecA2 K129R causes degradation of SecY, we wondered if a reduction in SecA1 might similarly cause a stress on the housekeeping Sec export system resulting in SecY degradation. In that case, the apparent requirement for SecA1 we reported previously might actually have been a reflection of a role for SecY in the SecA2 export pathway. To address this question, we expressed HA-tagged Msmeg_1712 (a known SecA2-dependent exported protein (15)) in a strain (JM693) in which SecA1 depletion can be induced by addition of anhydrotetracycline (Atc) in the presence of the revTetR repressor (26). We then monitored both SecY levels and SecA2-dependent export of Msmeg_1712-HA in the presence and absence of SecA1. Interestingly, the addition of Atc alone resulted in reduced levels of SecY protein (Fig. 2.7A). This occurred even in the absence of the revTetR repressor and was therefore *not* a result of SecA1 depletion. On the

contrary, when compared to a +Atc/-revTetR sample, the SecA1-depleted sample (+Atc/+revTetR) showed no additional reduction in SecY protein levels, indicating that SecA1 depletion does not cause degradation of SecY. Monitoring the subcellular localization of Msmeg_1712-HA (Fig. 2.7), we observed a decrease in SecA2-dependent export when SecA1 was depleted, confirming the previously observed requirement for SecA1 in SecA2-dependent protein export. SecA1 depletion had no effect on protein levels or localization of either SecA2 or a cytoplasmic control protein, SigA (data not shown).

SecA1 overexpression does not cause SecY degradation. We have previously shown that overexpression of SecA1 in a *secA2* deletion mutant background causes a growth defect on Mueller-Hinton agar (6). The reason for this result is not fully understood, but hints at the connection between SecA1 and SecA2-dependent pathways, as discussed above. In light of our findings that SecA2 K129R causes degradation of SecY, we revisited this experiment. If overexpression of SecA1, in the absence of SecA2, exerts a stress on the Sec pathway, it also could have resulted in SecY degradation, causing the observed phenotypes. To evaluate this possibility, we compared strains with and without a multi-copy *M. smegmatis* SecA1 overexpression plasmid, in both wild type and *secA2* deletion mutant backgrounds. As seen previously, SecA1 overexpression in the *secA2* deletion mutant background resulted in a growth defect on Mueller-Hinton agar (Fig. 2.8A). Additionally, SecA1 overexpression resulted in increased sensitivity to azide in both wild type and *secA2* deletion mutant backgrounds. However, SecA1 overexpression did not significantly affect SecY protein levels (Fig. 2.8B), demonstrating that the detrimental effect of high SecA1 levels did not occur through SecY degradation. SecA1 overexpression had no effect on protein levels of a cytoplasmic control protein, SigA (data not shown).

Discussion

While the conserved Sec export system of bacteria is well understood, even the most basic steps of SecA2-dependent export remain to be characterized. For the mycobacterial SecA2-dependent export system, there is no obvious SecY2 available to serve as a dedicated export channel. This type of SecA2 pathway is therefore termed a "SecA2-only" system (14). One of the many basic questions to be answered about the SecA2 pathway of mycobacteria is whether it utilizes the SecYEG translocase or an unknown apparatus to export its specific set of proteins. Here we showed that a SecA2 K129R dominant negative protein leads to reduced levels of the sole SecY of mycobacteria and that increased SecY levels can suppress the severe phenotypes of a *secA2 K129R* mutant. These data are significant in revealing a link between SecA2 and the housekeeping SecY, which provides strong support for the idea that SecA2 works with the canonical SecYEG translocase to export proteins.

The effect of SecA2 K129R on SecY levels indicates a relationship between SecA2-dependent export and the canonical Sec pathway. However, this result also raises the question of why SecA2 K129R leads to lower SecY levels. There is precedent for SecY degradation occurring in response to stress at the Sec translocase. In *E. coli*, when the SecYEG channel complex is artificially "jammed" by attempted export of a folded protein, the SecY protein is degraded by the FtsH protease, which serves to remove the non-functional "jammed" translocon (33). Our results are consistent with there being a similar stress on the SecYEG channel in the presence of the ATP binding-defective SecA2 K129R in *M. smegmatis*. There is also a homolog of *ftsH* in *M. smegmatis*, which could potentially be responsible for SecY degradation as in *E. coli*.

The housekeeping SecA protein drives export of an individual protein in a stepwise fashion through successive cycles of ATP binding and hydrolysis, during which SecA repeatedly releases and re-associates with the translocon. Furthermore, ATP hydrolysis is specifically necessary for SecA to dissociate from the translocase during this process (10). Therefore, a SecA protein that cannot bind and/or hydrolyze ATP (such as a Walker box KR variant) will become trapped at the membrane translocon and will fail to complete protein export. A Walker box substitution in *E. coli* SecA (SecA K108R) that is defective in ATP binding and hydrolysis (34) shifts SecA localization from the soluble fraction towards the cell envelope fraction (35), as does the corresponding SecA2 K129R variant of *M. smegmatis* (17). Thus, the most likely explanation for the effect of SecA2 K129R on SecY is that SecA2 K129R is locked in non-functional complexes with the housekeeping Sec export pathway, which results in SecY degradation in an effort by the cell to eliminate SecA2 K129R-jammed translocons. By analogy to the housekeeping Sec pathway (36), SecA2 could physically "dock" with SecYEG during protein export. However, we were unable to detect any physical interaction between SecA2 and SecY by chemical crosslinking in *M. smegmatis* (see Chapter 4). Therefore, the link we detect between SecA2 export and the SecY channel may reflect either a transient physical interaction or an indirect interaction involving another component of the export machinery.

Because SecY is critical for the export of many essential proteins, the reduced SecY levels observed in the presence of SecA2 K129R would be extremely detrimental to the cell. This helps explain the severe growth inhibition caused by SecA2 K129R. However, it is unlikely that there is no functional SecY protein in the *secA2 K129R* mutant, even though there was no full-length SecY detectable by western blotting in this strain (Fig. 2.4). SecY is thought to be essential for growth in all bacteria, including mycobacteria (37); therefore, a complete loss of

SecY would be lethal. Most likely, the amount of SecY in the *secA2 K129R* strain is below the level of detection with our αSecY antibody.

Given the reduced SecY levels in the *secA2 K129R* strain, it is logical that extragenic suppressor mutations might act to restore SecY abundance to a functional level. In suppressors 4S and 24S, this is accomplished by increasing transcription of *secY*, which must allow sufficient SecY production to replace the SecY channels lost to degradation. Interestingly, while suppressor 4S consistently displayed a less robust effect on *secY* expression than suppressor 24S (Fig. 2.3), these two suppressors are phenotypically identical (Fig. 2.1B). This indicates that the relatively weak effects of the 4S mutation are sufficient to reach a threshold SecY level necessary to restore functional export through the SecYEG translocon.

Six additional extragenic suppressors also restored SecY protein to varying degrees, though not by affecting *secY* transcription. These additional suppressor mutations could act in an alternate way to increase SecY production, for example by increasing the efficiency of *secY* translation. Another possibility is that these additional suppressor mutations serve to avoid the detrimental interaction between SecA2 K129R and the Sec pathway, thereby preventing SecY degradation in the first place. This latter possibility is especially interesting, as it may point to additional components of the machinery required for SecA2-dependent export and help elucidate the interaction points between the SecA2 pathway and housekeeping Sec export. In any case, the correlation between the suppressor phenotype and restoration of detectable SecY protein levels makes a strong case for reduced SecY levels being responsible for the *secA2 K129R* mutant phenotype.

The results of the current study confirm and build upon our previous result that the canonical SecA of mycobacteria (SecA1) is important for SecA2-dependent protein export, and

suggests that the SecA2 export pathway is fully integrated into housekeeping Sec export. It remains possible, however, that there exist additional specialty components that are important for SecA2-dependent protein export. We recently showed that features of a protein's mature domain (i.e. not the signal peptide) determine the requirement for SecA2. One interesting possibility is that the defining feature of the mature domain of SecA2-exported proteins is a propensity to fold in the cytoplasm prior to export (18). As an integrated component of the Sec pathway, SecA2 could assist the canonical Sec pathway to greater or lesser degrees with the export of proteins that are difficult to export due to folding or other features.

A protein export pathway in which SecA2 works with SecYEG may be a common feature of "SecA2-only" type systems. In fact, recent data from the *Clostridium difficile* "SecA2-only" system (4) shows that the corresponding Walker box substitution in the SecA2 of this system (SecA2 K106R) is dominant negative and causes severe growth inhibition when overexpressed, suggesting that it similarly interferes with an essential pathway. This and other data is consistent with the idea that SecA1 and SecA2 may also share use of the SecYEG channel in *C. difficile*. It would be interesting to see whether the detrimental effects of *C. difficile* SecA2 K106R also involve SecY degradation.

In conclusion, we uncovered a connection between SecA2 and the housekeeping SecY of mycobacteria using a classical genetic approach. Our results indicate that the "SecA2-only" system of mycobacteria utilizes the housekeeping SecYEG channel to export its select subset of proteins, addressing a key question about the mechanism of SecA2-dependent export. This pathway for SecA2-dependent export is distinct from that employed by "SecA2-SecY2" systems. Because our study indicates that the mycobacterial SecA2 export system is actually an

adaptation of the housekeeping Sec system, continued study of this system will not only increase our understanding of SecA2-dependent export, but may also shed light on the conserved Sec pathway utilized by all bacteria.

Table 2.1. *Mycobacterium smegmatis* strains used in this study.

***M. smegmatis* strain**	**Description**	**Source**
mc^2155	wild type	(38)
NR116	Δ*secA2*	(17)
NR158	mc^2155+pMV306.kan, wild type+empty plasmid	(17)
NR160	NR116+pMV306.kan, Δ*secA2*+empty plasmid	(17)
NR172	NR116+pYA810, Δ*secA2*+*secA2* complementing plasmid	(17)
NR178	NR116+pNR25, Δ*secA2*+*secA2 K129R* plasmid	(17)
NR236	NR178, with extragenic suppressor mutation 4S	This work
NR151	NR178, with extragenic suppressor mutation 24S	This work
NR154	NR178, with extragenic suppressor mutation 29S	This work
NR155	NR178, with extragenic suppressor mutation 33S	This work
NR156	NR178, with extragenic suppressor mutation 33B	This work
NR230	NR178, with extragenic suppressor mutation 1S	This work
NR234	NR178, with extragenic suppressor mutation 3S	This work
NR248	NR178, with extragenic suppressor mutation 10S	This work
JM693	*secA1* under control of P_{myc1}-*tetO*, for SecA1 depletion	Dr. Justin McDonough

Table 2.2. Plasmids used in this study.

Plasmid	Genotype	Description	Source
pMV306.kan	*aph int attP$_{L5}$* ColE1	Single-copy mycobacterial shuttle vector, integrates in mycobacteriophage L5 *attB* site	(39)
pMV361.kan	*aph* P$_{hsp60}$ *int attP$_{L5}$* ColE1	Single-copy mycobacterial shuttle vector with *hsp60* promoter, integrates in mycobacteriophage L5 *attB* site	(39)
pYA810	*aph* P$_{hsp60}$-*secA2 int attP$_{L5}$* ColE1	*M. smegmatis secA2* in pMV361.kan	(15)
pNR25	*aph* P$_{hsp60}$-*secA2K129R int attP$_{L5}$* ColE1	*M. smegmatis secA2 K129R* in pMV361.kan	(17)
pCR2.1-TOPO	*aph bla* ColE1	TOPO TA cloning plasmid	Invitrogen
pLL5	*aph bla* P$_{secY}$-*secY'* ColE1	*M. smegmatis secY* promoter and 34 bp of *secY* gene in pCR2.1-TOPO	This work
pLL6	*aph bla* P$_{secY(4S)}$-*secY'* ColE1	NR236 (4S) *secY* promoter and 34 bp of *secY* gene in pCR2.1-TOPO	This work
pCV125	*hyg 'lacZ oriM* ColE1	Multi-copy mycobacterial shuttle vector with promoter-less *lacZ* gene	MedImmune
pLL11	*hyg* P$_{secY}$-*secY'-'lacZ oriM* ColE1	*M. smegmatis secY* translationally fused to *lacZ* in pCV125	This work
pLL8	*hyg* P$_{secY(4S)}$-*secY'-'lacZ oriM* ColE1	NR236 (4S) *secY* translationally fused to *lacZ* in pCV125	This work
pLL15	*hyg* P$_{secY(24S)}$-*secY'-'lacZ oriM* ColE1	NR151 (24S) *secY* translationally fused to *lacZ* in pCV125	This work
pLL17	*aph bla secY* ColE1	*M. smegmatis secY* in pCR2.1-TOPO	This work
pLL19	*aph bla secY(4S)* ColE1	NR236 (4S) *secY* in pCR2.1-TOPO	This work
pYUB2063	*hyg bla int attP$_{Tweety}$* ColE1 *cos*$^{\lambda}$	Single-copy mycobacterial shuttle vector, integrates in mycobacteriophage Tweety *attB* site	Dr. WR Jacobs, Albert Einstein Coll. of Med.
pLL2	*hyg bla int attP$_{Tweety}$* ColE1	pYUB2063 with 3299 bp PciI fragment removed, single-copy mycobacterial shuttle vector, integrates in mycobacteriophage Tweety *attB* site	This work
pLL21	*hyg bla secY int attP$_{Tweety}$* ColE1	*M. smegmatis secY* in pLL2	This work
pLL23	*hyg bla secY(4S) int attP$_{Tweety}$* ColE1	NR236 (4S) *secY* in pLL2	This work
pKIsecA1	*hyg P$_{myc1}$-tetO-secA1'* ColE1	Suicide vector containing the first 641 bp of *M. smegmatis secA1* fused to *P$_{myc1}$-tetO*	(26)
pHSG85	*aph msmeg_1712-HA oriM* ColE1	*M. smegmatis msmeg_1712* with HA tag in pMV261.kan	(17)
pNR55	*aph msmeg_1712-HA oriM* ColE1 P$_{smyc}$-*revTetR*	*revTetR* and *msmeg_1712-HA* in pMV261.kan	(17)
pMV261.kan	*aph* P$_{hsp60}$ *oriM* ColE1	Multi-copy mycobacterial shuttle vector with *hsp60* promoter	(39)
pYUB544	*aph* P$_{hsp60}$-*secA1 oriM* ColE1	*M. smegmatis secA1* in pMV261.kan	(6)

Table 2.3. Primers used in this study.

Primer	Sequence and Description
1483promF	5'-GAGAATTCGTCCCACGTGTCGTTCT-3' binds 379 bp upstream of *secY* (*msmeg_1483*), contains engineered EcoRI site, used to amplify *secY* promoter region
1483promR	5'-GAGAATTCCCGTTCGCAGCGACGAGATG-3' binds at bp 34 of *secY* (*msmeg_1483*) coding sequence, contains engineered EcoRI site, used to amplify *secY* promoter
secYFNotI	5'-GAGCGGCCGCGTCCCACGTGTCGTTCTC-3' binds 379 bp upstream of *secY* (*msmeg_1483*), contains an engineered NotI site, used to amplify *secY* with its native promoter
secYREcoRV	5'-GAGATATCAGGCGGTCCGAGCAGAAC-3' binds 23 bp downstream of *secY* (*msmeg_1483*), contains an engineered EcoRV site, used to amplify *secY* with its native promoter
secY5'RACE-R1	5'-GACCTGCGCGATGCATTGCTGAAC-3' binds at bp 153 of *secY* (*msmeg_1483*) coding sequence, used to reverse transcribe *secY* transcript 5' ends to cDNA for 5' RLM-RACE
secY5'RACE-R2	5'-GGATTACCAGGCCCAAGGTGAACAG-3' binds at bp 79 of *secY* (*msmeg_1483*) coding sequence, used to amplify *secY* transcript 5' ends for 5' RLM-RACE
secY5'RACE-R3	5'-GTGGGTCCAGCGTCAATGGTCAAG-3' binds 51 bp upstream of *secY* (*msmeg_1483*), used to amplify *secY* transcript 5' ends for 5' RLM-RACE
secYF1	5'-GTGCTTTCGGCTTTCATCTC-3' binds at bp 1 of *secY* (*msmeg_1483*) coding sequence, used for quantitative RT-PCR
secYR1	5'-CGCGATGCATTGCTGAAC-3' binds at bp 147 of *secY* (*msmeg_1483*) coding sequence, used for quantitative RT-PCR
rpoBF	5'-GTCTCTAGCCAGAGCAAGTC-3' binds at bp 25 of *rpoB* (*msmeg_1367*) coding sequence, used for quantitative RT-PCR
rpoBR	5'-TCGAAGGAATCCGTCTGAAC-3' binds at bp 158 of *rpoB* (*msmeg_1367*) coding sequence, used for quantitative RT-PCR

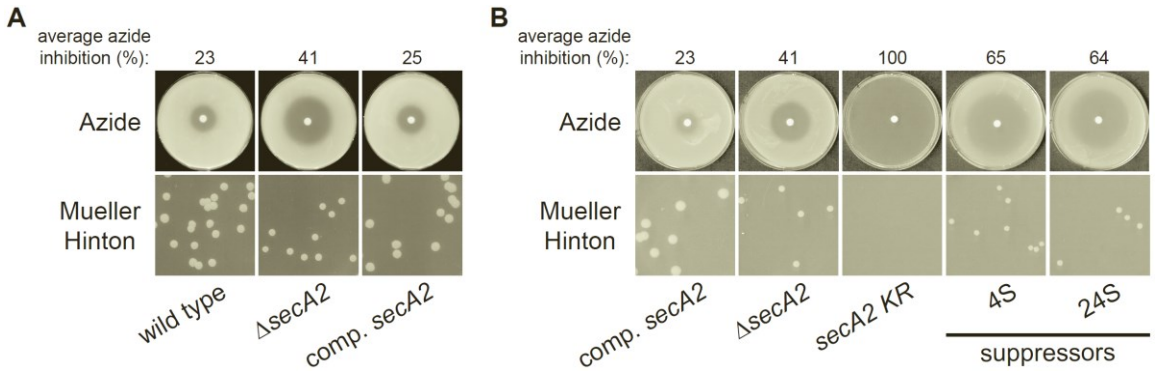

Figure 2.1. Azide sensitivity and Mueller-Hinton agar growth phenotypes. The indicated strains were plated for sensitivity to 10 µL of 0.15 M sodium azide for two days at 37°C and growth on Mueller-Hinton agar plates for six days at 37°C. Average azide inhibition is calculated by measuring the diameter of the zone of azide inhibition as a percentage of the plate diameter and is the mean of three technical replicates. Results shown are representative of at least three independent experiments. Strains tested were: **(A)** wild type+empty plasmid (NR158), Δ*secA2*+empty plasmid (NR160), and complemented *secA2* (NR172); **(B)** complemented *secA2* (NR172), Δ*secA2*+empty plasmid (NR160), *secA2 K129R* (NR178), 4S (NR236), and 24S (NR151).

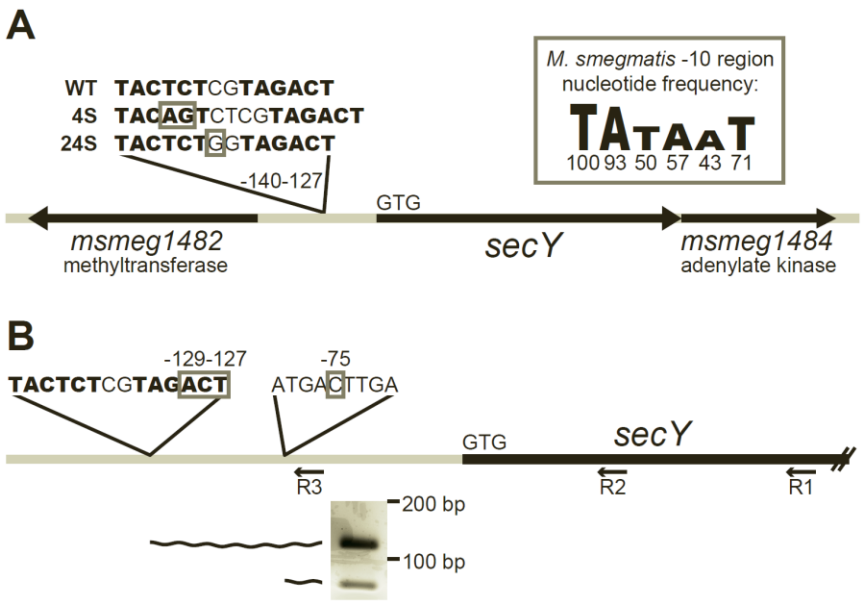

Figure 2.2. 4S and 24S mutations are located within the *secY* promoter. **(A)** Relative to the predicted translational start site of *secY* (indicated by start codon GTG), suppressor 4S (NR236) contains a 2 bp insertion between nucleotides -137 and -138. Suppressor 24S (NR151) contains a single nucleotide polymorphism at nucleotide -134. Suppressor mutations are boxed. Potential promoter -10 regions are shown in bold letters. The *M. smegmatis* -10 region consensus sequence is shown as a sequence logo representing the frequency at which each nucleotide occurs (percentages indicated beneath each nucleotide, adapted from (27)). Drawing not to scale. **(B)** Transcript 5' ends identified upstream of *secY*. Using primer R3, two species of *secY* transcript 5' ends were amplified (see agarose gel inset) from RNA from the complemented *secA2* strain (NR172). Relative to the predicted translational start site of *secY* (GTG), the longer species has a 5' end at nucleotides -129-127, and the shorter species has a 5' end at nucleotide -75. 5' end nucleotides are boxed. Potential promoter -10 regions are shown in bold letters. Amplified 5' end products are represented by wavy lines.

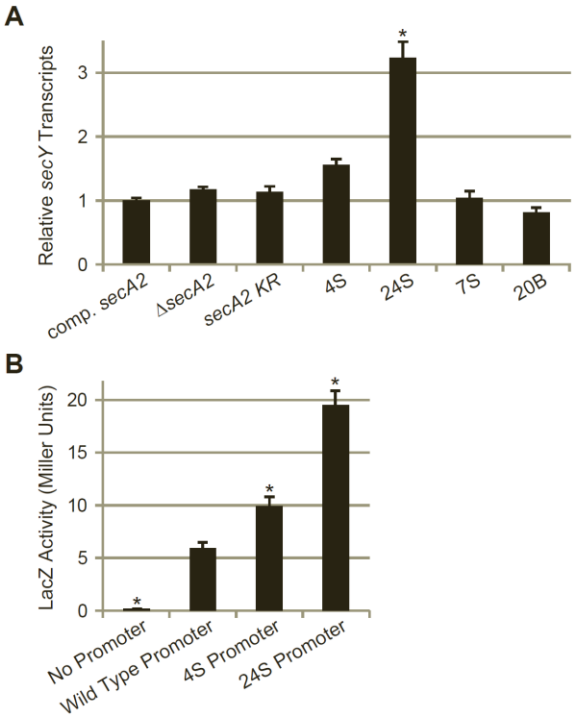

Figure 2.3. Suppressor mutations cause increased *secY* expression. **(A)** *secY* transcript levels measured by quantitative RT-PCR, relative to *rpoB* transcript levels. *secY* transcript level in the complemented *secA2* strain is set to 1. Data represents the mean of three biological replicates, and error bars represent standard error. *Statistically different from the *secA2 K129R* strain ($p<0.001$) by one-way ANOVA with Student-Newman-Keuls test. Strains tested were: complemented *secA2* (NR172), Δ*secA2*+empty plasmid (NR160), *secA2 K129R* (NR178), 4S (NR236), 24S (NR151), 7S (NR242), and 20B (NR260). **(B)** LacZ fusion activity assays. *secY* promoter regions plus 34 bp of *secY* coding sequence from wild type and suppressor strains were translationally fused to *lacZ*, and tested in wild type *M. smegmatis* mc^2155. Data represents the mean of six biological replicates, and error bars represent standard error. *Statistically different from "wild type promoter" strain ($p<0.05$) by one-way ANOVA on ranks with Student-Newman-Keuls test. LacZ fusions tested were: no promoter (pCV125), wild type *secY* promoter (pLL11), 4S *secY* promoter (pLL8), and 24S *secY* promoter (pLL15).

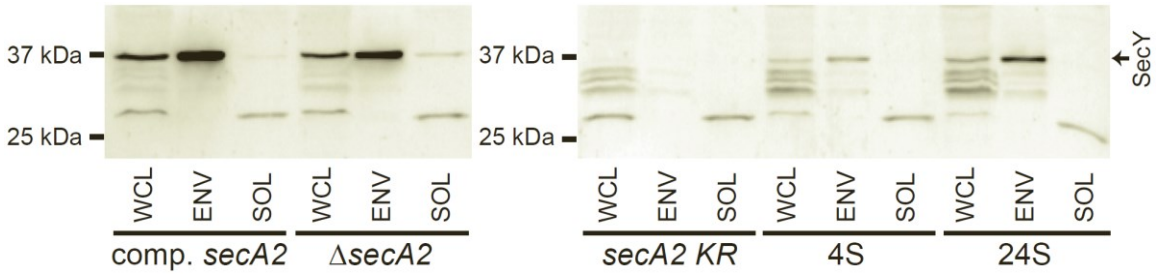

Figure 2.4. Suppressor mutations increase SecY protein levels relative to the *secA2 K129R* strain. Whole cell lysates and subcellular fractions were separated by SDS-PAGE and SecY protein detected by western blotting. All samples were equally loaded. WCL – whole cell lysate, ENV – pellet fraction containing cell envelope, SOL – soluble fraction containing cytoplasm. Strains tested were: complemented *secA2* (NR172), Δ*secA2*+empty plasmid (NR160), *secA2 K129R* (NR178), 4S (NR236), and 24S (NR151).

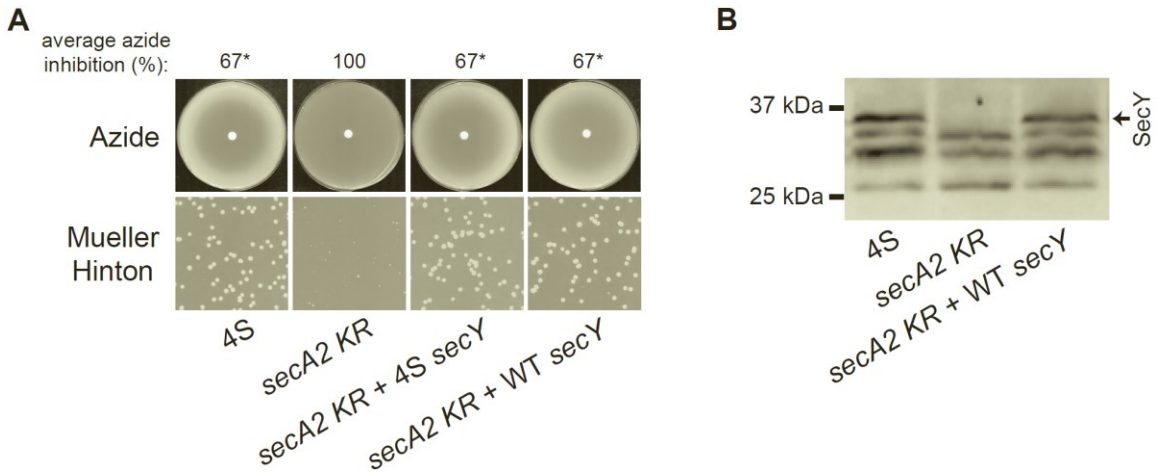

Figure 2.5. Increased SecY levels are sufficient to suppress *secA2 K129R*. Integrating plasmids containing *secY* under the control of the wild type promoter or the 4S promoter were added to the *secA2 K129R* strain and tested for azide sensitivity, colony size on Mueller-Hinton agar, and SecY protein level. Strains tested were: 4S suppressor+empty plasmid (NR236+pLL2), *secA2 K129R*+empty plasmid (NR178+pLL2), *secA2 K129R*+4S *secY* (NR178+pLL23), *secA2 K129R*+WT *secY* (NR178+pLL21). **(A)** Azide sensitivity and Mueller-Hinton agar growth phenotypes. Average azide inhibition is calculated by measuring the diameter of the zone of azide inhibition as a percentage of the plate diameter and is the mean of three biological replicates. *Statistically different from *secA2 K129R* strain ($p<0.001$) by one-way ANOVA with Student-Newman-Keuls test. **(B)** Whole cell lysates separated by SDS-PAGE and SecY protein detected by western blotting. All samples were equally loaded.

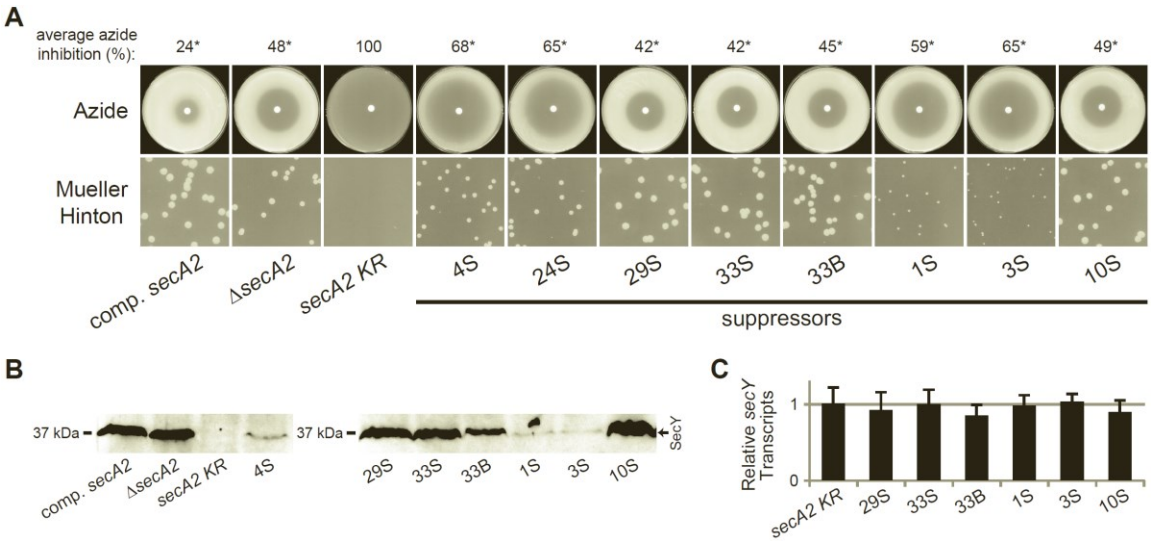

Figure 2.6. Additional extragenic suppressors increase SecY protein levels, but not *secY* transcript levels. Strains tested were: complemented *secA2* (NR172), Δ*secA2*+empty plasmid (NR160), *secA2 K129R* (NR178), 4S (NR236), 24S (NR151), 29S (NR154), 33S (NR155), 33B (NR156), 1S (NR230), 3S (NR234), and 10S (NR248). **(A)** Azide sensitivity and Mueller-Hinton agar growth phenotypes. Average azide inhibition is calculated by measuring the diameter of the zone of azide inhibition as a percentage of the plate diameter and is the mean of two to four biological replicates. *Statistically different from *secA2 K129R* strain ($p<0.001$) by one-way ANOVA with Student-Newman-Keuls test. **(B)** Cell envelope fractions separated by SDS-PAGE and SecY protein detected by western blotting. All samples were equally loaded. **(C)** *secY* transcript levels measured by quantitative RT-PCR, relative to *rpoB* transcript levels. The *secY* transcript level for the *secA2 K129R* strain is set to 1. Data represents the mean of six biological replicates, and error bars represent standard error. No statistical differences by one-way ANOVA.

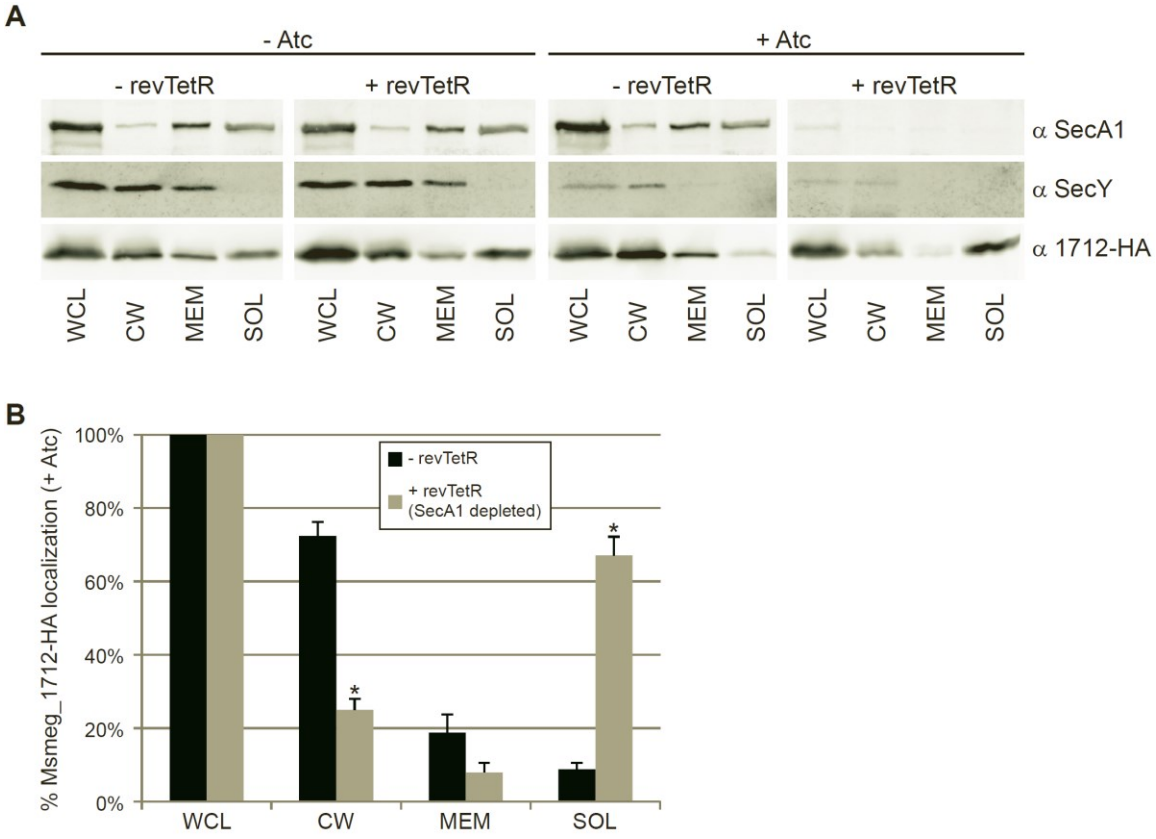

Figure 2.7. SecA1 depletion does not affect SecY and is required for SecA2-dependent protein export. **(A)** Whole cell lysates and subcellular fractions were separated by SDS-PAGE and SecA1, SecY, and Msmeg_1712-HA proteins detected by western blotting. All samples were equally loaded. WCL – whole cell lysate, CW – pellet fraction containing cell wall, MEM – pellet fraction containing cell membrane, SOL – soluble fraction containing cytoplasm. Strains tested were: "- revTetR" (JM693+pHSG85) and "+ revTetR" (JM693+pNR55). Each strain was tested in the presence (+ Atc) or absence (- Atc) of 1200 ng/mL anhydrotetracycline. SecA1 depletion occurs when both revTetR and Atc are present. **(B)** Msmeg_1712-HA localization was quantitated from western blots as in (A). WCL was calculated by adding CW+MEM+SOL and setting to 100%. Subcellular fractions (CW, MEM, and SOL) are presented as a percentage of the total WCL value. Data represents the mean of three biological replicates, and error bars represent standard error. *Statistically different from the "-revTetR" strain ($p<0.001$) by t-test. Strains tested were: "- revTetR" (JM693+pHSG85) and "+ revTetR" (JM693+pNR55). Each strain was tested in the presence (+ Atc) of 1200 ng/mL anhydrotetracycline.

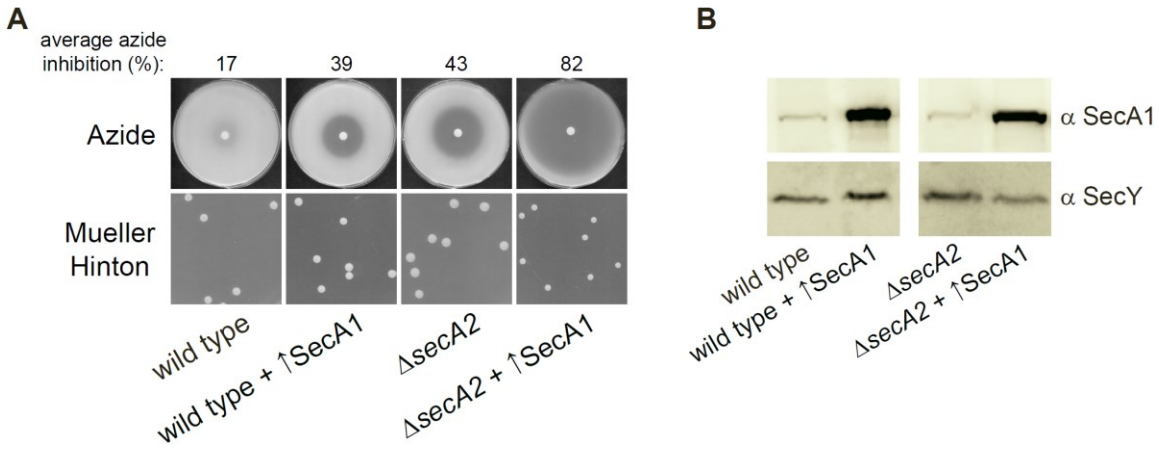

Figure 2.8. Overexpression of *M. smegmatis* SecA1 does not cause SecY degradation. Strains tested were: wild type+empty plasmid (mc^2155+pMV261.kan), wild type+SecA1 overexpression (mc^2155+pYUB544), Δ*secA2*+empty plasmid (NR116+pMV261.kan), and Δ*secA2*+SecA1 overexpression (NR116+pYUB544). **(A)** Azide sensitivity and Mueller-Hinton agar growth phenotypes. Average azide inhibition is calculated by measuring the diameter of the zone of azide inhibition as a percentage of the plate diameter and is the mean of one to three biological replicates. **(B)** Whole cell lysates separated by SDS-PAGE and SecA1 and SecY proteins detected by western blotting. All samples were equally loaded. When quantitated from three biological replicates of each strain and normalized to the cytoplasmic control protein SigA, no significant differences in SecY protein level were detected (data not shown).

References

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

2. **Nguyen-Mau SM, Oh SY, Kern VJ, Missiakas DM, Schneewind O.** 2012. Secretion genes as determinants of *Bacillus anthracis* chain length. J. Bacteriol. **194:**3841-3850.

3. **Caspers M, Freudl R.** 2008. *Corynebacterium glutamicum* possesses two *secA* homologous genes that are essential for viability. Archives of microbiology **189:**605-610.

4. **Fagan RP, Fairweather NF.** 2011. *Clostridium difficile* has two parallel and essential Sec secretion systems. J. Biol. Chem. **286:**27483-27493.

5. **Bensing BA, Sullam PM.** 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.

6. **Braunstein M, Brown AM, Kurtz S, Jacobs WR, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. J Bacteriol **183:**6979-6990.

7. **Siboo IR, Chaffin DO, Rubens CE, Sullam PM.** 2008. Characterization of the accessory Sec system of *Staphylococcus aureus*. J Bacteriol **190:**6188-6196.

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

9. **Schiebel E, Driessen AJ, Hartl FU, Wickner W.** 1991. Delta mu H+ and ATP function at different steps of the catalytic cycle of preprotein translocase. Cell **64:**927-939.

10. **van der Wolk JP, de Wit JG, Driessen AJ.** 1997. The catalytic cycle of the *escherichia coli* SecA ATPase comprises two distinct preprotein translocation events. EMBO J **16:**7297-7304.

11. **Shiba K, Ito K, Yura T, Cerretti DP.** 1984. A defined mutation in the protein export gene within the *spc* ribosomal protein operon of *Escherichia coli*: isolation and characterization of a new temperature-sensitive *secY* mutant. EMBO J **3:**631-635.

12. **Hou JM, D'Lima NG, Rigel NW, Gibbons HS, McCann JR, Braunstein M, Teschke CM.** 2008. ATPase activity of *Mycobacterium tuberculosis* SecA1 and SecA2 proteins and its importance for SecA2 function in macrophages. J Bacteriol **190:**4880-4887.

13. **Bensing BA, Sullam PM.** 2009. Characterization of Streptococcus gordonii SecA2 as a paralogue of SecA. J Bacteriol **191:**3482-3491.

14. **Rigel NW, Braunstein M.** 2008. A new twist on an old pathway--accessory Sec [corrected] systems. Mol Microbiol **69:**291-302.

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

16. **Braunstein M, Espinosa BJ, Chan J, Belisle JT, Jacobs WR, Jr.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of Mycobacterium tuberculosis. Mol Microbiol **48:**453-464.

17. **Rigel NW, Gibbons HS, McCann JR, McDonough JA, Kurtz S, Braunstein M.** 2009. The Accessory SecA2 System of Mycobacteria Requires ATP Binding and the Canonical SecA1. J Biol Chem **284:**9927-9936.

18. **Feltcher ME, Gibbons HS, Ligon LS, Braunstein M.** 2013. Protein export by the mycobacterial SecA2 system is determined by the preprotein mature domain. J. Bacteriol. **195:**672-681.

19. **Snapper SB, Lugosi L, Jekkel A, Melton RE, Kieser T, Bloom BR, Jacobs WR, Jr.** 1988. Lysogeny and transformation in mycobacteria: stable expression of foreign genes. Proc Natl Acad Sci U S A **85:**6987-6991.

20. **Pham TT, Jacobs-Sera D, Pedulla ML, Hendrix RW, Hatfull GF.** 2007. Comparative genomic analysis of mycobacteriophage Tweety: evolutionary insights and construction of compatible site-specific integration vectors for mycobacteria. Microbiology **153:**2711-2723.

21. **Larsen MH, Biermann K, Tandberg S, Hsu T, Jacobs WR, Jr.** 2007. Genetic Manipulation of *Mycobacterium tuberculosis*. Curr. Protoc. Microbiol. **Chapter 10:**Unit 10A 12.

22. **Li R, Li Y, Kristiansen K, Wang J.** 2008. SOAP: short oligonucleotide alignment program. Bioinformatics **24:**713-714.

23. **Kent WJ.** 2002. BLAT--the BLAST-like alignment tool. Genome Res **12:**656-664.

24. **Badejo AC, Badejo AO, Shin KH, Chai YG.** 2013. A gene expression study of the activities of aromatic ring-cleavage dioxygenases in Mycobacterium gilvum PYR-GCK to changes in salinity and pH during pyrene degradation. PLoS One **8:**e58066.

25. **Alland D, Steyn AJ, Weisbrod T, Aldrich K, Jacobs WR, Jr.** 2000. Characterization of the Mycobacterium tuberculosis iniBAC promoter, a promoter that responds to cell wall biosynthesis inhibition. J Bacteriol **182:**1802-1811.

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

27. **Newton-Foot M, Gey van Pittius NC.** 2013. The complex architecture of mycobacterial promoters. Tuberculosis (Edinb) **93:**60-74.

28. **Kenney TJ, Churchward G.** 1996. Genetic analysis of the *Mycobacterium smegmatis rpsL* promoter. J. Bacteriol. **178:**3564-3571.

29. **Bashyam MD, Tyagi AK.** 1998. Identification and analysis of "extended -10" promoters from mycobacteria. J. Bacteriol. **180:**2568-2573.

30. **Ito K.** 1984. Identification of the *secY* (*prlA*) gene product involved in protein export in *Escherichia coli*. Mol. Gen. Genet. **197:**204-208.

31. **Nakai M, Sugita D, Omata T, Endo T.** 1993. Sec-Y protein is localized in both the cytoplasmic and thylakoid membranes in the cyanobacterium *Synechococcus PCC7942*. Biochem. Biophys. Res. Commun. **193:**228-234.

32. **Akiyama Y, Ito K.** 1990. SecY protein, a membrane-embedded secretion factor of *E. coli*, is cleaved by the *ompT* protease in vitro. Biochem. Biophys. Res. Commun. **167:**711-715.

33. **van Stelten J, Silva F, Belin D, Silhavy TJ.** 2009. Effects of antibiotics and a proto-oncogene homolog on destruction of protein translocator SecY. Science **325:**753-756.

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

35. **Jilaveanu LB, Zito CR, Oliver D.** 2005. Dimeric SecA is essential for protein translocation. Proc Natl Acad Sci U S A **102:**7511-7516.

36. **Zimmer J, Nam Y, Rapoport TA.** 2008. Structure of a complex of the ATPase SecA and the protein-translocation channel. Nature **455:**936-943.

37. **Sassetti CM, Boyd DH, Rubin EJ.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol **48:**77-84.

38. **Snapper SB, Melton RE, Mustafa S, Kieser T, Jacobs WR, Jr.** 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. Mol. Microbiol. **4:**1911-1919.

39. **Stover CK, de la Cruz VF, Fuerst TR, Burlein JE, Benson LA, Bennett LT, Bansal GP, Young JF, Lee MH, Hatfull GF, Snapper SB, Barletta RG, Jacobs WR, Jr., Bloom BR.** 1991. New use of BCG for recombinant vaccines. Nature **351:**456-460.

CHAPTER 3

Identification of *Mycobacterium smegmatis secA2 K129R* Suppressor Mutations in Gene *msmeg_1684*[1]

All bacteria contain the conserved Sec pathway, which is responsible for transporting proteins across the cytoplasmic membrane. The SecA motor protein is an ATPase that plays a central role in this process. In mycobacteria, there are two SecA proteins. The canonical SecA responsible for the majority of protein export is referred to as SecA1 and a second SecA protein, referred to as SecA2, is responsible for exporting a small subset of proteins. SecA2-dependent export is also required for virulence of the pathogen *Mycobacterium tuberculosis*. Our previous work (Chapter 2) suggests that SecA2-dependent proteins of the model organism *Mycobacterium smegmatis* are exported through the canonical SecYEG channel with the assistance of SecA1. Despite this advance in our understanding of the mycobacterial SecA2 pathway, the mechanistic details of how SecA2 drives protein export across the cytoplasmic membrane remain poorly understood. Here, we analyzed suppressors of a dominant negative *M. smegmatis secA2* mutant (*secA2 K129R*) in order to improve our understanding of the SecA2 pathway. Eight extragenic suppressor mutations were identified as mapping to conserved hypothetical gene *msmeg_1684*. While the function of Msmeg_1684 is unknown, our results

[1]Contributing authors: Artur Romanchuk and Corbin Jones (Department of Biology, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3280); Nathan Rigel, Brittany Miller, Brandon Anjuwon-Foster, Chelsea Moriarty, Courtney Sutphen, Meagan Martin, and Miriam Braunstein (Department of Microbiology and Immunology, School of Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7290)

suggest that it participates directly in the mycobacterial SecA2-dependent protein export pathway.

Introduction

Mycobacteria use a variety of systems to export proteins to the bacterial cell surface and extracellular environment, including the widely conserved Sec and Tat systems as well as the specialized SecA2 and ESX systems (1). Many of the proteins exported by these systems are important for either basic cellular processes or virulence mechanisms, and are therefore critical to the ability of *Mycobacterium tuberculosis* to cause disease.

The canonical Sec pathway is conserved in all bacteria and is responsible for the majority of housekeeping protein export. In this system, SecY, SecE, and SecG form a membrane-embedded channel complex through which unfolded proteins are able to cross the membrane (2, 3). The cytoplasmic motor protein SecA pushes exported proteins through the SecYEG channel using the energy provided by repeated rounds of ATP binding and hydrolysis (4, 5). However, mycobacteria and also some Gram-positive bacteria are unique in that they contain two functionally distinct SecA proteins (6-12). In these bacteria, the housekeeping SecA protein (SecA1) is essential and performs the bulk of protein export, whereas the accessory SecA protein (SecA2) is generally non-essential (13) and is required for the export of a specific subset of proteins.

In *M. tuberculosis*, the SecA2 protein export system is important for virulence (14). *M. tuberculosis* is a pathogen that survives and grows in macrophages, and the SecA2-dependent protein export system is necessary for *M. tuberculosis* to block phagosome acidification and

maturation, which is important for intracellular growth (15). However, the mechanism of SecA2-dependent protein export is poorly understood. Some SecA2 systems include a second copy of the major Sec channel component SecY, called SecY2, which is believed to provide the channel for SecA2-dependent protein export (11-13). However, no SecY2 protein is present in mycobacteria. As presented in Chapter 2, we demonstrated a genetic relationship between the SecA2 protein of the model mycobacterium *Mycobacterium smegmatis* and the canonical SecY protein. Our data support a model in which the *M. smegmatis* SecA2 system works with the housekeeping Sec pathway to export its select set of proteins through the SecYEG translocon. However, it remains unknown whether additional SecA2-specific components participate in SecA2-dependent export. In fact, some SecA2 export systems of other bacteria are known to include additional protein components (i.e. components other than SecA2 and SecY2 proteins) important for tailoring the pathway to export a specific subset of proteins. For example, "SecA2-SecY2" type systems typically export glycosylated proteins. These systems contain proteins called Asps that assist in protein export, as well as GtfA and GtfB proteins that are important for glycosylation of SecA2-dependent proteins prior to export (16). Glycosylation does not seem to be the defining feature of proteins exported by the "SecA2-only" type mycobacterial systems, but other specific features of the mycobacterial SecA2-dependent proteins may require additional export machinery components.

To improve our understanding of the mycobacterial SecA2 export system, we performed a suppressor analysis using a dominant negative allele of *M. smegmatis secA2*, referred to as *secA2 K129R*. This allele encodes a variant of SecA2 containing a substitution in the Walker box motif, which is important for ATP binding and hydrolysis. The resulting SecA2 K129R protein is unable to hydrolyze ATP, and therefore is non-functional for SecA2-dependent protein export (17, 18). In addition, SecA2 K129R exhibits severe azide sensitivity and growth on Mueller-

Hinton agar defects, even in the presence of wild type SecA2 protein (18). This dominant negative phenotype, along with the fact that the phenotypes of the *secA2 K129R* strain are more severe than those of a *secA2* deletion mutant (18), suggests that SecA2 K129R, though non-functional, is still able to interact with its normal binding partners and interferes with their function. SecA2 K129R also undergoes a shift in localization towards the membrane-containing cell envelope fraction, whereas wild type SecA2 is primarily located in the cytoplasm (18). Therefore, SecA2 K129R would appear to be locked in an interaction at the cell membrane. These features make the *secA2 K129R* allele useful as a starting point for suppressor analysis. The utility of this suppressor approach was demonstrated in Chapter 2 by identifying two suppressor mutations in the *secY* promoter. Examination of these two suppressors showed that the presence of SecA2 K129R results in degradation of SecY and that increased *secY* expression is able to suppress the *secA2 K129R* phenotypes. These findings support a model in which SecA2 K129R interacts with the housekeeping Sec pathway and causes jamming of the SecY channel, leading to the drastic phenotypes of the *secA2 K129R* strain.

In Chapter 2 we characterized two extragenic suppressors; here, we characterized eight additional extragenic suppressors of *secA2 K129R*. Whole genome sequencing of these suppressor mutants demonstrated that all eight had mutations affecting the same gene, *msmeg_1684*. Two of these suppressors contain very large chromosomal deletions including *msmeg_1684*, while the other six suppressors contain various mutations directly in or adjacent to *msmeg_1684*. The function of Msmeg_1684 is unknown, but *msmeg_1684*-like genes co-occur with *secA2* genes in actinomycetes and, similar to SecA2, the Msmeg_1684-like protein Rv3311 is predicted to be important for growth of *M. tuberculosis* in macrophages (19). The data presented here argues for a model in which Msmeg_1684 is directly involved in the SecA2-dependent protein export pathway.

Materials and Methods

Azide sensitivity assays. Percent azide inhibition was measured as described previously (Chapter 2).

Subcellular fractionation and western blotting. Whole cell lysates and subcellular fractions were prepared and analyzed by western blotting as described previously (Chapter 2). Proteins were detected using rabbit polyclonal antisera against SecA2 (20), SecA1 (20), and SigA (gift of Dr. Murty Madiraju), at dilutions of 1:25,000, 1:50,000, and 1:20,000 respectively.

Whole genome and directed sequencing. Genomic DNA was isolated (as described previously (21)) from the 7S (NR242), 20B (NR260), 29S (NR154), 33S (NR155), 33B (NR156), 1S (NR230), 3S (NR234), and 10S (NR248) suppressor strains and submitted for whole genome sequencing at the High-Throughput Sequencing Facility at the University of North Carolina at Chapel Hill.

For strains 7S and 20B, sequencing was performed and analyzed as described previously (Chapter 2). Briefly, sequencing was performed at the High-Throughput Sequencing Facility at the University of North Carolina at Chapel Hill using Illumina GA II technology with 36 bp single-end reads. When reads were aligned to the *M. smegmatis* mc^2155 reference genome (mc^2155 is the parent strain to all strains described here) the resulting average read coverage was 16-20. Relative to the reference genome, a total of 53 mutations were identified in suppressor 7S and 63 mutations were identified in suppressor 20B. Many of these mutations were also identified in other mc^2155-derived strains sequenced in our laboratory and were discarded as background mutations (i.e. mutations already present in our *M. smegmatis* parent strain). The remaining mutations include 5 mutations appearing unique to the 7S strain and 8 mutations appearing unique to the 20B strain.

For strains 29S, 33S, 33B, 1S, 3S, and 10S, multiplex sequencing was performed using Illumina HiSeq technology with 50 bp single-end reads. Reads were aligned to the *M. smegmatis* mc^2155 reference genome (NCBI RefSeq accession number NC_008596.1) using Geneious version 5.5.6 (Biomatters, available from http://www.geneious.com/), resulting in an average coverage ranging between 171 and 251 across the various strains. Single nucleotide polymorphisms, insertions, and deletions were located using Geneious, resulting in a list of the total mutations identified in each strain: 162 mutations in suppressor 29S, 168 in 33S, 161 in 33B, 165 in 1S, 161 in 3S, and 160 in 10S. The vast majority of these mutations are present in all 6 sequenced strains, and are very likely to be background mutations. Therefore, those mutations not identified in all 6 strains were examined further. Four mutations were identified in either 3 or 4 strains each. In previous sequencing projects, identical mutations identified in most but not all strains were checked by direct sequencing and were always found to be present in all strains; therefore, these 4 mutations are also likely to be background mutations. The remaining mutations include 2 mutations appearing common to the 1S and 3S strains, 6 mutations appearing unique to the 29S strain, 9 in 33S, 3 in 33B, 8 in 1S, 5 in 3S, and 5 in 10S.

To assess specific genes or regions of interest, PCR amplified products were directly sequenced by Eton Bioscience, Inc. or Genewiz, Inc.

***M. smegmatis* colony PCR.** *M. smegmatis* cell extracts were prepared by suspending a loop-full of the desired *M. smegmatis* strain (taken from an agar plate) in 100 µL of 0.1% Triton X-100 and boiling for 10 minutes. Extracts were then cooled on ice and cell debris was pelleted for 10 minutes in a microcentrifuge. PCR was performed using 5 µL of the resulting supernatant as the template for a 50 µL PCR reaction.

Plasmid construction. Plasmids used in this study are described in Table 3.1 and primers are described in Table 3.2. In all cases, newly constructed plasmids were sequenced.

All plasmids used for unmarked gene deletion by two-step allelic exchange were built into the pMP62 suicide plasmid and were created by the same general strategy. Regions of approximately 500-1000 bp flanking the left and right ends of each desired deletion were amplified by PCR, fused together, and then sub-cloned into plasmid pMP62.

The suicide plasmid used to delete genes *msmeg_1704-1712*, pMEM4, was created by PCR amplifying the regions immediately flanking the left and right ends of the desired deletion. Splice overlap extension PCR (22) was then used to fuse the two flanking regions together into one PCR product. This product was cloned into pCR2.1-TOPO (Invitrogen), creating plasmid pMEM3. The combined flanking regions were then cut from pMEM3 using EcoRV, and ligated into EcoRV-digested pMP62, creating pMEM4.

The suicide plasmid used to delete genes *msmeg_1700-1726*, pCS4, was created by PCR amplifying the regions immediately flanking the left and right ends of the desired deletion. Each product was cloned into pCR2.1-TOPO (Invitrogen), creating plasmids pCS1 and pCS2, respectively. The left flanking region was cut from pCS1 using HindIII and NheI, and ligated into HindIII/NheI-digested pCS2, creating pCS3 and placing the left and right flanking regions adjacent to one another. The combined flanking regions were then cut from pCS3 using EcoRV, and ligated into EcoRV-digested pMP62, creating pCS4.

The suicide plasmid used to delete genes *msmeg_1692-1726*, pLL13, was created by PCR amplifying the region immediately flanking the left end of the desired deletion. This product was cloned into pCR2.1-TOPO (Invitrogen), creating plasmid pLL4. The left flanking region was cut from pLL4 using HindIII and NheI, and ligated into HindIII/NheI-digested pCS3. HindIII/NheI

digestion of pCS3 releases the left flank previously contained in this plasmid, but maintains the right flanking region. This created pLL12 and placed the left and right flanking regions adjacent to one another. The combined flanking regions were then cut from pLL12 using EcoRV, and ligated into EcoRV-digested pMP62, creating pLL13.

The suicide plasmid used to delete genes *msmeg_1677-1726*, pLL14, was created by PCR amplifying the region immediately flanking the left end of the desired deletion. This product was cloned into pCR2.1-TOPO (Invitrogen), creating plasmid pLL3. The left flanking region was cut from pLL3 using HindIII and NheI, and ligated into HindIII/NheI-digested pCS3. HindIII/NheI digestion of pCS3 releases the left flank previously contained in this plasmid, but maintains the right flanking region. This created pLL7 and placed the left and right flanking regions adjacent to one another. The combined flanking regions were then cut from pLL7 using EcoRV, and ligated into EcoRV-digested pMP62, creating pLL14.

The suicide plasmid used to delete genes *msmeg_1677-1692*, pCM2, was created by PCR amplifying the regions immediately flanking the left and right ends of the desired deletion. Each product was cloned into pCR2.1-TOPO (Invitrogen), creating plasmids pLL27 and pLL29, respectively. The left flanking region was cut from pLL27 using HindIII and NheI, and ligated into HindIII/NheI-digested pLL29, creating pLL30 and placing the left and right flanking regions adjacent to one another. The combined flanking regions were then cut from pLL30 using NdeI, and ligated into NdeI-digested pMP62, creating pCM2.

The plasmid used to express *secA2 K129R* from its native promoter, pLL36, was created by cutting the *secA2 K129R* allele along with its native promoter from pNR21 (23) using NotI and EcoRI. This fragment was then ligated into NotI/EcoRI-digested pMV306.kan, creating pLL36.

The suicide plasmid used to delete gene *msmeg_1684*, pLL50, was created by PCR amplifying the regions immediately flanking the left and right ends of the desired deletion. Each product was cloned into pCR2.1-TOPO (Invitrogen), creating plasmids pLL43 and pLL44, respectively. The left flanking region was cut from pLL43 using HindIII and NheI, and ligated into HindIII/NheI-digested pLL44, creating pLL49 and placing the left and right flanking regions adjacent to one another. The combined flanking regions were then cut from pLL49 using NdeI, and ligated into NdeI-digested pMP62, creating pLL50.

Bacterial strains and culture conditions. *M. smegmatis* strains used in this study are described in Table 3.3, and were grown as described previously (Chapter 2). *Escherichia coli* strains were grown at 37°C using lysogeny broth media (otherwise known as Luria-Bertani media). Antibiotics kanamycin (40 μg/mL) and hygromycin B (150 μg/mL), and carbenicillin (for amp^R plasmids, 100 μg/mL) were added as needed. When hygromycin B and carbenicillin were used together (for pMP62-based plasmids), their concentrations were adjusted to 50 μg/mL each.

Gene deletion by two-step allelic exchange. Unmarked deletion strains were created by two-step allelic exchange, as described previously (10, 24, 25). Briefly, suicide plasmids containing a hygromycin-resistance selectable marker, a *sacB* counter-selectable marker, and flanking regions for the genes to be deleted were transformed into *M. smegmatis* by electroporation. Transformants in which the suicide plasmid recombined into the genome (by homologous recombination into one of the flanking regions) were selected by plating on media containing hygromycin B. Hygromycin-resistant transformants were grown to saturation, diluted 1:100 in media lacking hygromycin B, and then grown overnight at 37°C. Bacteria in which a second recombination event occurred, resulting in loss of the *sacB* marker, were

selected by plating on 7H10 agar supplemented with 0.2% glucose and 4.5% sucrose. Sucrose-resistant colonies were patched on 7H10 agar +/- hygromycin B to verify loss of the hygromycin resistance gene. The resulting strains were assessed for the desired chromosomal deletion first by PCR and subsequently by southern blot. Because some of the desired deletions were very large and technically challenging to create, they were built by sequentially deleting smaller regions of 8-17 kbp at a time. In general, deletion strains were created in the wild type (mc^2155) background to assess the phenotypes of each deletion in isolation. In addition, deletion strains were created in combination with a *secA2* deletion. The *secA2 K129R* allele (pNR25) was then introduced into these double deletion strains to test the ability of each strain to suppress *secA2 K129R* phenotypes.

The *msmeg_1704-1712* deletion strain was created by allelic exchange using suicide plasmid pMEM4 in the wild type *M. smegmatis* strain mc^2155, resulting in strain LL036. The *secA2* deletion was then introduced into this strain (as achieved previously in the *secA2* mutant NR116) using suicide plasmid pNR6, resulting in strain LL062.

The *msmeg_1700-1726* deletion strain was created by allelic exchange using plasmid pCS4 in strain LL036, resulting in strain LL051. The *secA2* deletion was then re-created in this strain using suicide plasmid pNR6, resulting in strain LL066.

The *msmeg_1692-1726* deletion strain was created by allelic exchange using plasmid pLL13 in strain LL051, resulting in strain LL082. The *secA2* deletion was then re-created in this strain using suicide plasmid pNR6, resulting in strain LL089.

The *msmeg_1677-1726* deletion strain was created by allelic exchange using plasmid pLL14 in strain LL082, resulting in strain LL129, and in strain LL089, resulting in strain LL132.

The *msmeg_1677-1692* deletion strain was created by allelic exchange using plasmid pCM2 in strain mc^{2}155, resulting in strain CM14, and in the *secA2* mutant strain NR116, resulting in strain CM11.

The *msmeg_1684* deletion strain was created by allelic exchange using plasmid pLL50 in the *secA2* mutant strain NR116, resulting in strain BAF1.

Quantitative RT-PCR. RNA isolation and quantitative RT-PCR were performed as described previously (Chapter 2). For quantitative RT-PCR on *secA2*, *msmeg_1677, msmeg_1682, and msmeg_1683,* products were amplified from the 5' ends of each gene sequence, as well as from the *rpoB* gene sequence. Starting quantity of each transcript was calculated relative to DNA standards, and transcripts of each gene were normalized to *rpoB* transcripts (as a control) in each sample (26).

Results

Spontaneous suppressor mutations alleviate the severe phenotypes of the *secA2 K129R* allele. The *M. smegmatis secA2* deletion mutant exhibits increased sensitivity to azide and a moderate growth defect on Mueller-Hinton agar, relative to the wild type mc^{2}155 strain (10). These phenotypes can be complemented by adding an integrating plasmid containing the wild type *secA2* allele expressed from the *hsp60* promoter (Fig. 2.1A, Chapter 2) (18). However, addition of a comparable plasmid containing the *secA2 K129R* mutant allele expressed from the *hsp60* promoter cannot complement the *secA2* deletion phenotypes. Instead, the non-functional SecA2 K129R protein, which contains a substitution in the Walker box preventing ATP binding and hydrolysis, causes more severe phenotypes than observed for the *secA2* deletion

strain (Fig. 3.1) (18). The *secA2 K129R* allele is also dominant negative, exhibiting phenotypes in the presence of a wild type *secA2* allele (18).

As described in Chapter 2, suppressor mutations spontaneously arise in the *secA2 K129R* strain and were collected for further study. Throughout this study, we routinely compared these suppressor strains to the complemented strain expressing *secA2* from the *hsp60* promoter, in order to compare strains in which all *secA2* alleles are expressed from the same promoter. We chose two extragenic suppressor strains, 7S (NR242) and 20B (NR260), for identification of suppressor mutations. Each of these suppressors exhibits azide resistance and growth on Mueller-Hinton agar similar to the *secA2* deletion strain (Fig. 3.1), and are therefore referred to as "strong" suppressors. In contrast, some other suppressors, such as the 4S and 24S suppressors studied in Chapter 2, exhibit phenotypes that are "intermediate" between the phenotypes of the *secA2* deletion strain and the *secA2 K129R* mutant strain. Suppressors 7S and 20B also exhibit normal SecA2 K129R protein levels (i.e. unchanged relative to the parent *secA2 K129R* strain) and normal *secY* transcript levels (Fig. 2.3, Chapter 2).

Suppressors 7S and 20B affect subcellular localization of SecA1 and SecA2. In Chapter 2, we described two suppressors (4S and 24S) with intermediate suppressor phenotypes that partially alleviate the effects of the *secA2 K129R* allele by increasing expression of *secY*. Because suppressors 7S and 20B exhibit stronger suppressor phenotypes than 4S and 24S, we hypothesized that they might completely prevent the detrimental effects of the *secA2 K129R* allele. One effect of this allele is that the localization of the SecA2 K129R protein shifts towards the membrane-containing cell envelope fraction, relative to wild type SecA2, which is found in the cytoplasm (18). Therefore, we evaluated the subcellular localization of SecA2, and also SecA1, in these suppressor strains (Fig. 3.2).

As seen previously (18), in the complemented *secA2* strain (NR172) SecA1 is found equally distributed between the cytoplasm-containing soluble fraction and the cell envelope-containing pellet fraction. This is consistent with the localization of the canonical SecA protein of *E. coli* (27). In contrast, SecA2 is found predominantly in the soluble fraction of NR172 (18). In the presence of SecA2 K129R (strain NR178), the percentage of both SecA1 and SecA2 in the cell envelope-containing pellet fraction increased, indicating increased interaction of both proteins with the membrane (or membrane-bound proteins) in the *secA2 K129R* strain. Unlike the SecA1 and SecA2 localization patterns seen in the *secA2 K129R* strain, suppressors 7S and 20B exhibited localization patterns very similar to what is seen in the complemented *secA2* strain. In other words, the 7S and 20B mutations that suppress the *secA2 K129R* plate phenotypes also restore normal localization of both SecA1 and SecA2 proteins. For comparison, suppressor 4S (NR236) was also tested. In the suppressor 4S strain, the localization patterns of SecA1 and SecA2 was also altered in comparison to what was seen in the *secA2 K129R* strain. However, in this case, the percent of membrane-bound SecA1 and SecA2 was not completely reduced to the level seen in the complemented *secA2* strain. As discussed in Chapter 2, suppressor 4S is an intermediate suppressor that acts by increasing *secY* expression. This mode of suppression would not be expected to prevent the interaction of SecA2 K129R with the housekeeping Sec system, but it would compensate for the detrimental effects of that interaction by increasing the available pool of SecY channels. Thus, it is not surprising that 4S would exhibit only a partial effect on localization of SecA1 and SecA2. In contrast, the complete reversal of SecA1/SecA2 localization exhibited in suppressors 7S and 20B suggests that these strong suppressors may act to completely *prevent* the detrimental effects of SecA2 K129R, thereby preventing the shift in localization seen in the *secA2 K129R* strain.

Suppressors 7S and 20B contain large overlapping chromosomal deletions. In preparation for whole genome sequencing to identify suppressor mutations, we first verified that suppressors 7S and 20B contained intact genes for each of the housekeeping Sec system components. It was previously determined that each of these strains contains normal *secA2 K129R*, *secA1*, and *secY* gene sequences (23). In order to rule out the possibility of suppressor mutations occurring in either *secE* or *secG*, we sequenced each of these genes from suppressor strains 7S and 20B as well as the parent strain, the *secA2 K129R* mutant. In all cases, the *sec* gene sequences matched those of the published *M. smegmatis* mc^2155 genome.

After confirming the absence of any *sec* gene mutations in strains 7S and 20B, we performed whole genome sequencing on both strains. Alignment of sequence reads to the *M. smegmatis* mc^2155 reference genome resulted in sequence coverage sufficient to identify mutations in both genomes. To reduce the identification of background mutations, we compared the consensus sequences of 7S and 20B to those of other mc^2155-derived strains sequenced in our laboratory (data not shown). This allowed us to generate a list of 5 unique mutations in the 7S suppressor genome, and 8 unique mutations in the 20B suppressor genome. Interestingly, one of the unique mutations from each strain was a very large deletion, encompassing about 51 kbp in suppressor 7S (genes *msmeg_1678-1726*, bp 1769908-1820883) and about 43 kbp in suppressor 20B (genes *msmeg_1684-1726,* bp 1777694-1820871). These deletions were characterized by extremely low to zero sequence coverage throughout the genomic region that was deleted (Fig. 3.3). The deletions in 7S and 20B are largely overlapping; however, the two deletions arose independently of one another, as these two suppressor strains originated from separate cultures of the *secA2 K129R* strain and differ at the left end of the deletions.

To verify the presence of these large deletions in the 7S and 20B suppressors, we attempted to PCR amplify 3 different regions of about 1000 bp each within the large deletions – one encompassing the right end of the large deletions, one in *msmeg_1712*, and one in *msmeg_1704* (Fig. 3.4). Each of these PCR products was successfully amplified from genomic DNA purified from both the wild type strain and the *secA2 K129R* strain (NR178, parent strain to 7S and 20B), but could not be amplified from suppressors 7S and 20B. As a control, a region of about 700 bp in *secE* and a region of about 500 bp in *secG* could be PCR amplified from all four strains. These results confirm that large deletions are indeed present in suppressors 7S and 20B, while the left end of these deletions was confirmed later (below).

Suppressor NR123 also contains a large deletion overlapping those in suppressors 7S and 20B. Upon identification of the overlapping deletions in suppressors 7S and 20B, we wanted to determine whether similar mutations occurred in other extragenic suppressors. To accomplish this, we attempted to colony PCR amplify the same 3 regions discussed above from other suppressors. One additional suppressor with an overlapping deletion, strain NR123, was identified among a group of 7 suppressors arising from an *M. smegmatis* strain expressing the *M. tuberculosis secA2 K115R* allele (comparable to the *M. smegmatis secA2 K129R* allele).

As discussed above, one PCR test region was located to encompass the right end of the large deletions, which is in approximately the same location in both 7S and 20B suppressors. This test PCR product could not be amplified from suppressor strain NR123, therefore the right end of the deletion in this strain extends at least as far as in 7S and 20B. To assess the left end of the deletion in NR123, we attempted to colony PCR amplify a region of about 850 bp in *msmeg_1684*, the left-most gene deleted in suppressor 20B. While this region could be successfully amplified from both the wild type and *secA2 K129R* strains, it could not be amplified

from suppressor NR123 (or from suppressors 7S and 20B, as expected). Therefore, the left ends of all three deletions were confirmed to extend at least as far as *msmeg_1684* (Fig. 3.4).

Fifty genes are deleted in suppressors 7S and 20B. Between the two deletions found in strains 7S and 20B, fifty genes are deleted (Fig. 3.4 and Table 3.4). Forty-three of these genes are deleted in both strains. Of the remaining seven genes deleted only in suppressor 7S, six are transcribed in a downstream orientation relative to the left end of the smaller deletion in suppressor 20B and therefore may be subject to polar effects on their expression, resulting from the nearby deletion site.

Several genes worth noting are located within these large deletions. Foremost, genes *msmeg_1704* and *msmeg_1712* encode the only two known *M. smegmatis* SecA2-dependent exported proteins (28). In addition, four of the affected genes are predicted to encode amidases (*msmeg_1679*, *msmeg_1686*, *msmeg_1702*, and *msmeg_1703*). Some SecA2-dependent exported proteins of *Listeria monocytogenes* are amidases (29, 30). However, these *L. monocytogenes* amidases contain signal peptides, while none of the *M. smegmatis* amidase genes contain predicted signal peptides or transmembrane domains, and are therefore not predicted to be exported (Table 3.4). None of the other genes affected in the 7S and 20B deletions have any known connection to protein export pathways, though some are predicted to be exported proteins themselves. However, it is also interesting to note the presence of several genes at the right end of the 7S and 20B deletions that are annotated to encode transposase enzymes (*msmeg_1716-1721* and *msmeg_1725-1726*). These genes could conceivably have played a role in initiating the deletion of this large region of the genome.

Deletion of *msmeg_1704* and *msmeg_1712* is not responsible for suppression. Because the large deletions occurred in three independent suppressors, we hypothesized that

they represented the mutations responsible for suppression of *secA2 K129R* phenotypes in these strains. Therefore, we next began to narrow down the deletion regions in an effort to determine the minimum deletion required to recreate the suppressor phenotype (Fig. 3.4). The fact that *msmeg_1704* and *msmeg_1712*, encoding the only two known SecA2-dependent exported proteins in *M. smegmatis* (28), are located within the large deletions was particularly striking to us. This finding suggested a model in which elimination of the substrates of the SecA2 export system could suppress the severe phenotypes caused by SecA2 K129R. If SecA2 K129R must first engage a protein destined for export *before* it moves to and interferes with the membrane-associated export machinery, then elimination of SecA2-dependent exported proteins might prevent this detrimental interaction from occurring. To test this hypothesis, we first created a strain in which genes *msmeg_1704-1712* were deleted. In a wild type background (LL036), this deletion has no effect on azide resistance or growth on Mueller-Hinton agar (data not shown). Next, we deleted genes *msmeg_1704-1712* in a *secA2* deletion background (LL062) and then introduced the *secA2 K129R* allele (plasmid pNR25). By introducing the *secA2 K129R* allele at the last step, we avoided the accumulation of new suppressor mutations prior to completion of the final strain. Finally, we assessed the ability of this deletion to suppress *secA2 K129R* phenotypes (Fig. 3.5). Deletion of genes *msmeg_1704-1712* did *not* suppress *secA2 K129R* phenotypes, suggesting our initial model was incorrect.

Large deletions can be narrowed down to a 15-gene deletion capable of suppression.

Because deletion of *msmeg_1704-1712* did not suppress *secA2 K129R*, we proceeded to create successively larger deletions, eliminating genes *msmeg_1700-1726* and then *msmeg_1692-1726* (Fig. 3.4). Because of their large size, these deletions were technically challenging to construct. Therefore, each deletion was constructed from the previous smaller deletion strain, effectively deleting 8-17 kbp at a time. In a wild type background (strains LL051 and LL082 respectively),

these deletions have no effect on azide resistance or growth on Mueller-Hinton agar (data not shown). Next we created strains carrying these same deletions along with a *secA2* deletion (strains LL066 and LL089, respectively) and subsequently introduced the *secA2 K129R* allele (plasmid pNR25). Both the *msmeg_1700-1726* and *msmeg_1692-1726* deletions did not suppress *secA2 K129R* phenotypes.

Having failed to observe suppression with these deletions, we constructed an even larger deletion spanning *msmeg_1677-1726* (strain LL129 in wild type background and strain LL132 in *secA2* deletion background), which essentially recreates the deletion found in suppressor 7S. Fortunately, this largest deletion suppressed the phenotypes of the *secA2 K129R* allele (Fig. 3.5), which demonstrates that the large deletions found in suppressors 7S and 20B can indeed account for the observed suppression of *secA2 K129R*.

Because the *msmeg_1677-1726* deletion suppressed *secA2 K129R* while the *msmeg_1692-1726* deletion could not, we hypothesized that the deleted gene responsible for the suppressor phenotype resides somewhere in the *msmeg_1677-1691* region. To test this hypothesis, we constructed a strain with a deletion of *msmeg_1677-1692* (Fig. 3.5) in a *secA2* deletion background (CM11), and then introduced the *secA2 K129R* allele (plasmid pNR25) into this strain. In fact, the *msmeg_1677-1692* deletion *was* capable of suppressing *secA2 K129R* phenotypes (Fig. 3.5). This result indicated that the deleted gene(s) responsible for the suppressor phenotype is located within genes *msmeg_1677-1692*, a region of 15 genes.

Suppressor 20B deletion does not prevent expression of downstream genes, narrowing down the suppressing region to a suite of 9 genes. The finding that the 15-gene deletion (CM11, Δ*msmeg_1677-1692*) was capable of suppressing *secA2 K129R* drastically reduced the number of genes we needed to consider as possibly being responsible for the

suppressor phenotype (Fig. 3.6). However, of the remaining 15 genes, none stood out as having the potential to be involved in protein export. In an effort to narrow down this list of genes further, we next examined a group of six genes that are deleted in suppressor 7S but not in suppressor 20B (*msmeg_1677-1683*). Because these six genes are transcribed in a downstream orientation relative to the left end of the 20B deletion, it was possible that, despite being present in suppressor 20B, they were not expressed due to polar effects. Therefore, the 15-gene deletion was engineered to include these six genes. However, if expression of these genes is *not* prevented by the 20B deletion, then we could remove these genes from the list of genes under consideration as potentially responsible for suppression.

Examination of genes *msmeg_1677-1683* using a promoter prediction program (SoftBerry BPROM) revealed a likely promoter upstream of *msmeg_1682* (Fig. 3.6). In addition, genes *msmeg_1677-1682* are closely spaced together on the chromosome (with the last codon of each gene overlapping the start codon of the following gene), suggesting they are likely transcribed as a single operon. Therefore, we decided to assess expression of genes *msmeg_1677* and *msmeg_1682* as representative genes at the beginning and the end of this 5-gene predicted operon. The last of the six genes we wanted to examine, *msmeg_1683*, is not part of the predicted *msmeg_1677-1682* operon and is also directly adjacent to the deletion in suppressor 20B. Therefore, *msmeg_1683* has a strong potential to be affected by the 20B deletion and we decided to separately evaluate its expression.

Expression of each of these three genes (*msmeg_1677, 1682,* and *1683*) was evaluated by quantitative RT-PCR, comparing suppressor strains 7S and 20B to the parent *secA2 K129R* strain (Fig. 3.7). For each gene, transcripts were standardized to *rpoB* transcripts as a control, and the transcript level in the *secA2 K129R* strain was set to 1. In all three cases, the gene

transcripts found in suppressor 7S were much lower than those found in the *secA2 K129R* strain. This was expected, as all three of these genes are deleted in suppressor 7S. While a low amount of transcript was still detected in suppressor 7S for gene *msmeg_1677*, this was found to be caused by a non-specific transcript that was also amplified at a low level in these reactions. In suppressor 20B, where all three genes are still present, the transcript levels of each gene are *not* reduced, but rather increased. It is interesting that the presence of the deletion in suppressor 20B increased expression of the neighboring genes (*msmeg_1677-1683*); however, because expression of these genes is increased rather than decreased, it is unlikely that any of these genes are involved in suppression of *secA2 K129R*. Based on these results, we reasoned that the gene(s) responsible for the 7S and 20B suppression phenotype must be among nine genes: *msmeg_1684-1692*.

Six extragenic suppressors contain mutations affecting *msmeg_1684*. Six additional extragenic suppressors of *M. smegmatis secA2 K129R* were briefly examined in Chapter 2. All of these suppressors exhibited normal SecA2 K129R protein levels and contained normal *secA2 K129R*, *secA1*, and *secY* gene sequences (23), as well as normal *secY* promoter sequences (Chapter 2). Interestingly, while these six suppressors exhibited increased SecY protein levels relative to the parent *secA2 K129R* strain, none exhibited changes in *secY* transcription like those seen in suppressors 4S and 24S (Fig. 2.6, Chapter 2). These six suppressors do not contain large deletions like those found in suppressors 7S and 20B. To further characterize these suppressors, we performed western blots to detect SecA1, SecA2, and SigA proteins (Fig. 3.8). As expected, all strains exhibit equal levels of the cytoplasmic control protein SigA. In addition, all of the extragenic suppressors expressed SecA2 K129R to the same level seen in the parent *secA2 K129R* strain (as seen previously (23)) and all strains exhibited equal levels of SecA1 protein.

In order to identify the mutations present in the six additional extragenic suppressor strains, we performed whole genome sequencing on all strains. Alignment of sequence reads to the *M. smegmatis* mc^{2}155 reference genome resulted in sequence coverage for each strain ranging from 171 to 251, sufficient to identify potential mutations in the suppressor genomes. The vast majority of mutations were identified in all 6 sequenced strains, and are very likely to be background mutations. Therefore, we began by examining only those candidate mutations appearing unique to a single strain. Interestingly, four of the suppressor strains were predicted to contain independent mutations in or near conserved hypothetical gene *msmeg_1684*. Further manual inspection of this region revealed that, in fact, all six sequenced suppressors contain candidate mutations in or near *msmeg_1684*, though the mutations in two strains were not identified by automated sequence analysis. Importantly, no mutations are alike, indicating that each *msmeg_1684* mutation arose independently of the others.

To verify the presence of these mutations, we PCR amplified and sequenced *msmeg_1684* from each strain. All six suppressor strains contained the *msmeg_1684* mutations as predicted by whole genome sequencing. In contrast, the parent *secA2 K129R* strain contained an *msmeg_1684* sequence identical to the sequence reported in the published mc^{2}155 reference genome (NCBI Gene accession number YP_886060.1). As shown in Fig. 3.9 and Table 3.5, each *msmeg_1684* mutation is unique. Two mutations create frameshifts in the open reading frame at residues 25 and 261 and one mutation creates a premature stop codon truncating the protein at residue 123 (Msmeg_1684 is 431 amino acids in length). Two other mutations create smaller changes, substituting a single amino acid at residue 134 and deleting two amino acids at residues 269-270. The final mutation is a single nucleotide polymorphism located 6 bp upstream of *msmeg_1684*. While it is not immediately clear if this mutation affects the *msmeg_1684* promoter or shine-delgarno sequences, it seems likely that it affects

expression of *msmeg_1684* in some way. To assess whether any other suppressors contain *msmeg_1684* mutations, *msmeg_1684* was PCR amplified and sequenced from the other 8 (of the original 14) extragenic suppressor strains. No additional *msmeg_1684* suppressor mutations were found in any strain, including the 4S and 24S suppressor strains with mutations in the *secY* promoter examined previously (Chapter 2).

Most interestingly, *msmeg_1684* is located in the large deletions of both suppressors 7S and 20B. Furthermore, *msmeg_1684* is one of the remaining nine genes that we believed could explain the suppression observed in 7S and 20B. Taken together, this work strongly suggests that deletion of the *msmeg_1684* gene can suppress *secA2 K129R* phenotypes and can account for the 7S and 20B suppressor phenotypes.

An *msmeg_1684* deletion can suppress *secA2 K129R* phenotypes. To directly assess whether deletion of *msmeg_1684* can result in suppression of *secA2 K129R* phenotypes, we created an unmarked, in-frame deletion of *msmeg_1684* (leaving behind only 10 amino acids) in a *secA2* deletion mutant background (strain BAF1). We then introduced the *secA2 K129R* allele into this strain and assessed the resulting phenotypes. Deletion of *msmeg_1684* is indeed able to suppress *secA2 K129R* phenotypes. The *secA2 K129R*, deletion *msmeg_1684* double mutant exhibited strong resistance to azide and growth on Mueller-Hinton agar (Fig. 3.10), similar to the phenotypes of suppressor 33S (NR155), which has a frameshift mutation at nucleotide 25 of *msmeg_1684*. We conclude that loss of function mutations affecting *msmeg_1684* can suppress *secA2 K129R* phenotypes. This finding demonstrates that there is a genetic relationship between *msmeg_1684* and *secA2*, and therefore indicates that Msmeg_1684 must perform a function related to the SecA2-dependent protein export pathway.

Suppressors affecting *msmeg_1684* and *secY* do not act through the *hsp60* promoter.

Because the suppressor analysis described in Chapters 2 and 3 was carried out using a strain expressing the *secA2 K129R* allele from the *hsp60* promoter, it was possible that some suppressor mutations would act by preventing or reducing expression from the *hsp60* promoter, resulting in less SecA2 K129R protein and, consequently, a less severe phenotype. The fact that SecA2 K129R is expressed at roughly equal levels, as assessed by western blot analysis, in all extragenic suppressors we have studied (Fig. 3.8) suggests that these suppressor mutations act specifically to suppress the effects of the SecA2 K129R protein, not to alter expression from the *hsp60* promoter. Nevertheless, this was a question that we wanted to investigate further.

We began by measuring expression of *secA2* alleles in various strains to determine whether subtle changes in expression might occur at the level of transcription from the *hsp60* promoter (Fig. 3.11). Relative to the expression of wild type *secA2* from the *hsp60* promoter (in strain NR172), the *secA2* deletion strain with an empty vector (NR160) exhibits no expression of *secA2*, as expected. However, expression of the *secA2 K129R* allele from the *hsp60* promoter (in strain NR178) is about 2.5 fold higher than expression of wild type *secA2* from the same promoter. Plasmids pYA810 (wild type *secA2* on *hsp60* promoter) and pNR25 (*secA2 K129R* on *hsp60* promoter) are identical with the exception of the K129R mutation in *secA2*, therefore the presence of SecA2 K129R appears to affect the cell in such a way as to increase expression from the *hsp60* promoter. While the *hsp60* promoter has previously been shown to respond to stresses such as heat, acid, and peroxide (31), it has not been shown to respond to export-related stress, such as that seemingly caused by the presence of SecA2 K129R. As the *hsp60* promoter is very commonly used to express genes in mycobacteria, this will be important to consider in the design of future experiments.

In suppressor 7S, a strong suppressor affecting *msmeg_1684*, expression of the *secA2 K129R* allele is reduced to the level seen for the wild type *secA2* allele (Fig. 3.11). In contrast, expression of the *secA2 K129R* allele is only partially reduced in suppressor 4S, an intermediate suppressor affecting expression of *secY* (Chapter 2). While on the surface this result seems to support the concern that suppressors might act to affect expression from the *hsp60* promoter, it is important to consider that suppressor strains also prevent or alleviate the export-related stress caused by SecA2 K129R. Because expression from the *hsp60* promoter is sensitive to export-related stress, as demonstrated by the *secA2 K129R* strain (Fig. 3.11), the observed decrease in expression in the suppressor strains may be indicative only of reduced export stress, though we cannot draw a conclusion at this point.

In order to determine conclusively whether our suppressors act to suppress the effects of SecA2 K129R or to reduce expression from the *hsp60* promoter, we created an integrating plasmid carrying the *secA2 K129R* allele expressed from its native promoter (pLL36). We then introduced this plasmid into various strains and assessed the resulting phenotypes. By testing whether suppressor mutations can suppress the native promoter-driven *secA2 K129R*, we can remove any role for the *hsp60* promoter from the experiment. As for P_{hsp60}-*secA2 K129R*, introduction of P_{native}-*secA2 K129R* into the *secA2* deletion strain caused increased sensitivity to azide and decreased growth on Mueller-Hinton agar compared to the *secA2* deletion mutant (Fig. 3.12). However, the effects of the P_{native}-*secA2 K129R* allele are not as severe as those of the P_{hsp60}-*secA2 K129R* allele, due to lower expression from the native promoter (SecA2 K129R is expressed at 36% of the level supported by the *hsp60* promoter, by western blotting, data not shown). Next, the P_{native}-*secA2 K129R* allele was introduced into strains containing a *secA2* deletion in combination with engineered deletions removing genes *msmeg_1677-1726* (strain LL132) or genes *msmeg_1677-1692* (strain CM11). As discussed above, both of these deletions

remove gene *msmeg_1684* and both strains are capable of suppressing the P_{hsp60}-*secA2 K129R* allele. In the presence of the P_{native}-*secA2 K129R* allele, each of these strains again exhibited suppressor phenotypes (Fig. 3.12). We also tested a strain containing an extra copy of *secY* from suppressor 4S (NR116+pLL23). As discussed in Chapter 2 (Fig. 2.5), this strain is capable of suppressing the P_{hsp60}-*secA2 K129R* allele. Again, this engineered suppressor strain is also capable of suppressing the P_{native}-*secA2 K129R* allele (Fig. 3.12). These findings demonstrate conclusively that both *msmeg_1684* and *secY* type suppressor strains act specifically to suppress the effects of SecA2 K129R and not to reduce expression from the *hsp60* promoter.

Discussion

As discovered in Chapter 2, the SecA2-dependent export pathway appears to be integrated with the housekeeping Sec pathway, with SecA2 and SecA1 likely sharing use of the SecYEG channel for protein export. However, it remains unclear whether additional SecA2-specific components exist and participate in SecA2-dependent protein export. For example, in SecA2-SecY2 type systems, additional components called Asps also participate in SecA2-dependent protein export (16).

While the 4S and 24S suppressors (discussed in Chapter 2) are able to partially suppress *secA2 K129R* phenotypes by increasing *secY* expression, they do not completely prevent the detrimental effects of the SecA2 K129R protein, which appears to "jam" the housekeeping Sec export pathway, resulting in degradation of SecY. Rather, these *secY*-type suppressors partially compensate for SecY degradation by increasing the available pool of SecY protein. In contrast, several other extragenic suppressors discussed in Chapter 2 were able to suppress *secA2 K129R*

phenotypes more fully and with normal SecY protein levels, but did so without affecting *secY* expression (Fig. 2.6, Chapter 2).

Here, we examined eight additional extragenic suppressors of the *secA2 K129R* allele, including the six suppressors mentioned above and assessed briefly in Chapter 2. Of these eight suppressors, six exhibited strong suppressor phenotypes, with azide resistance and growth on Mueller-Hinton agar similar to that of the *secA2* deletion mutant. In addition, we demonstrated that two of these strong suppressors, 7S and 20B, completely reverse the SecA2 and SecA1 membrane localization changes seen in the *secA2 K129R* mutant. These findings suggest that, rather than partially compensating for the effects of the SecA2 K129R protein, as seen in the 4S and 24S *secY*-related suppressors (Chapter 2), these other suppressors may completely prevent the detrimental interactions of SecA2 K129R with the housekeeping Sec system from occurring in the first place. These types of suppressors may point to additional components of the SecA2 export system that are necessary for the interaction between SecA2 and the housekeeping Sec system.

To identify the suppressor mutations present in these strains, all eight suppressors were subjected to whole genome sequencing. In two suppressors, 7S and 20B, very large chromosomal deletions were identified including gene *msmeg_1684*. The remaining six suppressors contained either point mutations or small insertions or deletions in or near gene *msmeg_1684*. To confirm that mutations affecting this gene could indeed cause the suppressor phenotype, we engineered an unmarked, in-frame deletion of *msmeg_1684* in a *secA2* deletion background, which we found is also capable of suppressing *secA2 K129R* phenotypes. The fact that deletion of *msmeg_1684* results in a strong suppressor phenotype suggests that Msmeg_1684 must be present for the detrimental effects of SecA2 K129R to occur. This is

consistent with the idea that Msmeg_1684 directly participates in the SecA2-dependent export pathway. In the future, it will be interesting to explore this idea by assessing whether deletion of *msmeg_1684* in a wild type background eliminates export of SecA2-dependent proteins.

A variety of different *msmeg_1684* mutations were identified as suppressor mutations. Five of the eight suppressor mutations either removed the *msmeg_1684* gene (7S and 20B) or caused frameshifts or truncation of *msmeg_1684* at residues 25, 123, and 261 of the 431 amino acid protein (strains 33S, and 33B, and 29S, respectively). Interestingly, these five suppressor strains with drastic *msmeg_1684* mutations also exhibited strong suppressor phenotypes, consistent with the conclusion that complete removal of *msmeg_1684* prevents the detrimental effects of the *secA2 K129R* allele. Because the mutations causing frameshifts or truncation have suppressor phenotypes as strong as the complete *msmeg_1684* deletions, they most likely result in the production of unstable proteins that are quickly degraded. One additional suppressor, 10S, also exhibits strong suppressor phenotypes. However, this suppressor is unique in that it does not directly affect the *msmeg_1684* coding sequence, but rather is a T→C polymorphism located 6 bp upstream of *msmeg_1684*. Given the strength of the 10S suppressor phenotype, we hypothesize that the mutation in this strain must severely compromise expression of *msmeg_1684*; however, the location of the promoter and shine-delgarno sequences upstream of *msmeg_1684* are not immediately evident. Therefore, it will be interesting to discover if the mutation in strain 10S does, in fact, affect expression of *msmeg_1684*, and in what way.

The remaining two suppressors, 3S and 1S, affect *msmeg_1684* in more subtle ways. The mutation in strain 3S causes a single amino acid substitution, changing glycine 134 to an aspartate. Similarly, the suppressor mutation in strain 1S results in an in-frame deletion of two

amino acids, glycine 269 and threonine 270. Both 3S and 1S suppressors exhibit an intermediate level of *secA2 K129R* suppression, suggesting that their respective mutations adversely affect, but do not completely destroy, Msmeg_1684 function. These mutations in particular may prove useful for future study of Msmeg_1684 function. If any of the suppressors are found to produce stable, but abnormal Msmeg_1684 protein, they may pinpoint specific portions of the amino acid sequence required for Msmeg_1684 to function or interact with other proteins.

Msmeg_1684 is annotated as a conserved hypothetical protein and displays no informative homology to known proteins using BLAST. Therefore, its function is completely unknown. However there are a number of interesting features to note. Msmeg_1684-like proteins are found in all mycobacteria and Msmeg_1684 has a strong homolog in *M. tuberculosis*, named Rv3311, which is 65% identical and 79% similar to Msmeg_1684 at the amino acid level (Fig. 3.13). The function of Rv3311 is also unknown, but is predicted by TraSH to be required for growth of *M. tuberculosis* in macrophages (19). Rv3311 was not tested in the TraSH screen performed in mice (32), therefore no prediction exists for this phenotype. In addition, *rv3311* is located downstream of, and predicted to be co-transcribed with, *sapM* (*rv3310*), encoding an exported protein which, like SecA2, is thought to be important for blocking phagosome maturation (15, 33). It will be interesting to determine whether an *M. tuberculosis rv3311* mutant exhibits virulence defects comparable to those of the *secA2* mutant, which would be consistent with these two genes functioning in the same pathway.

Intriguingly, there may be a trend for genes encoding exported proteins to cluster near *msmeg_1684*-like genes. Both *msmeg_1684* and *rv3311* are located in close proximity to genes encoding amidases (genes *msmeg_1679*, *1686*, *1702*, and *1703* in *M. smegmatis* and genes *rv3305c* and *3306c* in *M. tuberculosis*). Amidase genes *msmeg_1702* and *1703* are homologous

to genes *rv3306c* and *3305c*, respectively. The presence of these amidase genes is interesting to note, as amidases are found among the SecA2-dependent proteins identified in *L. monocytogenes* (29, 30). The *msmeg_1684* gene is also located not far from genes *msmeg_1704* and *msmeg_1712*, encoding two solute-binding components of ABC transporter systems, which are the only known SecA2-dependent exported proteins in *M. smegmatis* (28). Similarly, the *msmeg_1684* homolog in *Corynebacterium glutamicum*, *NCgl0651*, is also located near two amidase genes (*NCgl0652* and *0657*), as well as genes encoding two ABC transporter systems (*NCgl0636-0646*). While this idea is purely speculative at this point, the proximity of amidase and ABC transporter-encoding genes to *msmeg_1684*-like genes implies that genes encoding SecA2-dependent exported proteins may cluster near *msmeg_1684*. However, there is currently no evidence that any of the amidases are exported by the SecA2 system.

It is also interesting to examine the occurrence of Msmeg_1684-like proteins in other bacteria. All proteins homologous to Msmeg_1684 are found within the order Actinomycetales, of which mycobacteria are members. Among the currently sequenced genomes of bacteria in the order Actinomycetales (about 370 genomes), approximately half contain homologs of Msmeg_1684, including the mycobacteria, corynebacteria, rhodococci, and gordonia species, as well as several others (Fig. 3.14). SecA2 proteins are also commonly found within the order Actinomycetales and are also present in approximately half of the currently sequenced Actinomycetales genomes. Intriguingly, the species within this order that contain SecA2 proteins almost completely overlap with those species that contain Msmeg_1684-like proteins (Fig. 3.14). We found only 16 species that contain Msmeg_1684-like proteins without a SecA2 protein (notably including the rhodococci) and only 6 species that contain SecA2 proteins without an Msmeg_1684-like protein. In the species containing Msmeg_1684-like proteins only, the Msmeg_1684-like protein is also more distantly related to those of mycobacteria. This

frequent co-occurrence of Msmeg_1684-like and SecA2 proteins is consistent with the idea that they function as part of the same pathway. The presence of an Msmeg-1684-like protein is a feature associated specifically with SecA2 systems closely related to those of mycobacteria. This includes the SecA2 system studied in *C. glutamicum* (8). SecA2 systems found outside of the order Actinomycetales, including those of both the "SecA2-only" and "SecA2-SecY2" varieties, do *not* contain Msmeg_1684-like proteins. Therefore, it appears that Msmeg_1684-like proteins may represent a specific adaptation of the mycobacterial and related Actinomycetales SecA2 systems.

Another interesting feature is that both Msmeg_1684 and Rv3311 contain an unusually high number of aspartate and glutamate residues (Fig. 3.13), making them very acidic proteins (with predicted isoelectric points of 3.83 and 3.90 respectively). Interestingly, protein export chaperones are frequently acidic proteins; for example, SecB, an export-dedicated chaperone of *E. coli*, has a predicted isoelectric point of 4.05. Thus, we can hypothesize that Msmeg_1684 might play a chaperone-like role in the SecA2-dependent export pathway, perhaps maintaining SecA2-dependent proteins in an export-competent state prior to export, similar to the role of SecB in *E. coli* Sec export. Consistent with this idea, it was recently shown that features of the mature domain (i.e. not the signal peptide) of *M. smegmatis* SecA2-dependent proteins determine the requirement for SecA2 in their export (34). Specifically, a tendency towards folding in the cytoplasm prior to export was suggested to play a role in the requirement for SecA2. With this possibility in mind, we can develop a model to account for suppression of *secA2 K129R* phenotypes in the absence of Msmeg_1684 function. If Msmeg_1684 is an export chaperone that serves to keep SecA2-dependent proteins unfolded prior to export, then the absence of Msmeg_1684 in a suppressor strain might prevent SecA2 K129R from interacting with proteins destined for export. Because substrate protein binding by SecA2 could be a

requisite step for interacting with the SecYEG channel, the absence of Msmeg_1684 could then serve to avoid the detrimental interactions of SecA2 K129R with the housekeeping Sec pathway, resulting in a suppressor phenotype.

The path to discovery of Msmeg_1684 was indirect; however, we collected additional pieces of information and useful tools along the way. Notably, while deletion of *msmeg_1704* and *msmeg_1712* was not important for suppressing *secA2 K129R*, as originally hypothesized, we created an *msmeg_1704-1712* deletion strain, which may be useful in future study of the SecA2-dependent exported proteins encoded by these genes. We also demonstrated that our suppressors are capable of suppressing *secA2 K129R* phenotypes, regardless of whether the allele is expressed from its native promoter or the *hsp60* promoter, creating in the process a new version of the *secA2 K129R* allele that is expressed from the native promoter and may also be useful for future studies. In addition, we discovered that the *hsp60* promoter is subject to regulation in response to the export-related stress caused by the *secA2 K129R* allele. This promoter is often considered to be constitutive and is widely used for gene expression in mycobacteria (31). The changes in expression we observed from the *hsp60* promoter under conditions of export-related stress could have unintended consequences, which should be considered in future experiments using this promoter. A final interesting point to discuss is the presence of such large deletions (43-51 kbp) in the 7S and 20B suppressors when smaller mutations in *msmeg_1684* are equally effective at suppressing *secA2 K129R*. Interestingly, there are several genes annotated as transposases found at the right end of both the 7S and 20B deletions. In addition, the right end of each deletion is in the same location, while the left ends of the deletions differ. This suggests that the transposase genes are able to initiate deletion of nearby genes. In the case of the *secA2 K129R* strain, any deletion beginning at this point and stretching far enough to affect *msmeg_1684* would behave as a suppressor strain. In fact, the

deletion in suppressor 20B stretches just far enough to affect *msmeg_1684* but does not affect any additional genes located to the left of this point.

In conclusion, we used a classical genetic approach to discover a novel aspect of mycobacterial protein export – the presence of an Msmeg_1684-like protein. Msmeg_1684 appears likely to participate directly in the SecA2 export pathway. In the future, it will be interesting to characterize this relationship and determine the specific role of Msmeg_1684 in SecA2-dependent export. Because the *M. tuberculosis* homolog of Msmeg_1684, Rv3311, is predicted to be important for virulence, continued study of these proteins will not only increase our understanding of SecA2-dependent export, but may also shed light on virulence mechanisms of *M. tuberculosis*.

Table 3.1. Plasmids used in this study.

Plasmid	Genotype	Description	Source
pMV306.kan	*aph int attP$_{L5}$* ColE1	Single-copy mycobacterial shuttle vector, integrates in mycobacteriophage L5 *attB* site	(31)
pMV361.kan	*aph* P$_{hsp60}$ *int attP$_{L5}$* ColE1	Single-copy mycobacterial shuttle vector with *hsp60* promoter, integrates in mycobacteriophage L5 *attB* site	(31)
pYA810	*aph* P$_{hsp60}$-*secA2 int attP$_{L5}$* ColE1	*M. smegmatis secA2* in pMV361.kan	(28)
pNR25	*aph* P$_{hsp60}$-*secA2K129R int attP$_{L5}$* ColE1	*M. smegmatis secA2 K129R* in pMV361.kan	(18)
pNR7	*aph* P$_{hsp60}$-*secA2K115R int attP$_{L5}$* ColE1	*M. tuberculosis secA2 K115R* in pMV361.kan	(18)
pCR2.1-TOPO	*aph bla* ColE1	TOPO TA cloning plasmid	Invitrogen
pMP62	*bla hyg* P$_{hsp60}$-*sacB* ColE1	empty suicide vector for creating deletions in mycobacteria, also called pYUB657	(24)
pNR6	*bla hyg* P$_{hsp60}$-*sacB* ColE1 Δ*secA2*	suicide vector containing flanking regions to delete *M. smegmatis secA2*	(18)
pMEM3	*aph bla* Δ*msmeg_1704-1712* ColE1	left and right flanking regions for *msmeg_1704-1712* in pCR2.1-TOPO	This work
pMEM4	*bla hyg* P$_{hsp60}$-*sacB* ColE1 Δ*msmeg_1704-1712*	suicide vector containing flanking regions to delete *msmeg_1704-1712*	This work
pCS1	*aph bla msmeg_'1699* ColE1	left flanking region for *msmeg_1700* in pCR2.1-TOPO	This work
pCS2	*aph bla msmeg_1728* ColE1	right flanking region for *msmeg_1726* in pCR2.1-TOPO	This work
pCS3	*aph bla* Δ*msmeg_1700-1726* ColE1	left and right flanking regions for *msmeg_1700-1726* in pCR2.1-TOPO	This work
pCS4	*bla hyg* P$_{hsp60}$-*sacB* ColE1 Δ*msmeg_1700-1726*	suicide vector containing flanking regions to delete *msmeg_1700-1726*	This work
pLL4	*aph bla msmeg_'1691-1692'* ColE1	left flanking region for *msmeg_1692* in pCR2.1-TOPO	This work
pLL12	*aph bla* Δ*msmeg_1692-1726* ColE1	left and right flanking regions for *msmeg_1692-1726* in pCR2.1-TOPO	This work
pLL13	*bla hyg* P$_{hsp60}$-*sacB* ColE1 Δ*msmeg_1692-1726*	suicide vector containing flanking regions to delete *msmeg_1692-1726*	This work
pLL3	*aph bla msmeg_'1678* ColE1	left flanking region for *msmeg_1677* in pCR2.1-TOPO	This work
pLL7	*aph bla* Δ*msmeg_1677-1726* ColE1	left and right flanking regions for *msmeg_1677-1726* in pCR2.1-TOPO	This work
pLL14	*bla hyg* P$_{hsp60}$-*sacB* ColE1 Δ*msmeg_1677-1726*	suicide vector containing flanking regions to delete *msmeg_1677-1726*	This work
pLL27	*aph bla msmeg_'1678* ColE1	left flanking region for *msmeg_1677* in pCR2.1-TOPO	This work
pLL29	*aph bla msmeg_'1692-1693'* ColE1	right flanking region for *msmeg_1692* in pCR2.1-TOPO	This work
pLL30	*aph bla* Δ*msmeg_1677-1692* ColE1	left and right flanking regions for *msmeg_1677-1692* in pCR2.1-TOPO	This work
pCM2	*bla hyg* P$_{hsp60}$-*sacB* ColE1 Δ*msmeg_1677-1692*	suicide vector containing flanking regions to delete *msmeg_1677-1692*	This work
pNR21	*bla* ColE1 *oriF1 lacZ'-msmeg_3652-secA2-msmeg_3654-'lacZ*	*M. smegmatis secA2 K129R* with native promoter in pSKII+	(23)
pLL36	*aph secA2K129R int attP$_{L5}$* ColE1	*M. smegmatis secA2 K129R* with native promoter in pMV306.kan	This work
pLL2	*hyg bla int attP$_{Tweety}$* ColE1	pYUB2063 with 3299 bp PciI fragment removed, single-copy mycobacterial shuttle vector, integrates in mycobacteriophage Tweety *attB* site	This work (Chapter 2)
pLL23	*hyg bla secY(4S) int attP$_{Tweety}$* ColE1	NR236 (4S) *secY* in pLL2	This work (Chapter 2)
pLL43	*aph bla msmeg_'1686-1684'* ColE1	left flanking region for *msmeg_1684* in pCR2.1-TOPO	This work
pLL44	*aph bla msmeg_'1684-1683'* ColE1	right flanking region for *msmeg_1684* in pCR2.1-TOPO	This work
pLL49	*aph bla* Δ*msmeg_1684* ColE1	left and right flanking regions for *msmeg_1684* in pCR2.1-TOPO	This work
pLL50	*bla hyg* P$_{hsp60}$-*sacB* ColE1 Δ*msmeg_1684*	suicide vector containing flanking regions to delete *msmeg_1684*	This work

Table 3.2. Primers used in this study.

Primer	Sequence and Description
1704-1712LF1-2	5'-GATATCATCCCGCAGATCGGCACCC-3' binds at bp 751 of *msmeg_1703*, contains engineered EcoRV site , used to amplify left flank for *Δmsmeg_1704-1712* plasmid
1704-1712LF2	5'-GGAGTCGGCTTGTGAGGAAGTCATAAACGTCGGCGGCTAC-3' binds 278 bp downstream of *msmeg_1703* , used to amplify left flank for *Δmsmeg_1704-1712* plasmid
1704-1712RF1	5'-CTTCCTCACAAGCCGACTCC-3' binds 408 bp upstream of *msmeg_1713*, used to amplify right flank for *Δmsmeg_1704-1712* plasmid
1704-1712RF2	5'-ATGGTGCACGCCGTGAAATC-3' binds at bp 308 of *msmeg_1713*, used to amplify right flank for *Δmsmeg_1704-1712* plasmid
1700-1726LF1-3	5'-GAGATATCGAACGATTCGGCATATGTG-3' binds at bp 540 of *msmeg_1699*, contains engineered EcoRV site, used to amplify left flank for *Δmsmeg_1700-1726* plasmid
1700-1726LF2-2	5'-GAGCTAGCAACAAGTGTAGGCCAACAAG-3' binds 94 bp upstream of *msmeg_1699*, contains engineered NheI site, used to amplify left flank for *Δmsmeg_1700-1726* plasmid
1700-1726RF1-2	5'-GAGCTAGCCAAGGGCAGCATCTAGTGTC-3' binds 303 bp upstream of *msmeg_1728*, contains engineered NheI site, used to amplify right flank for *Δmsmeg_1700-1726* plasmid
1700-1726RF2	5'-TCGATGTAGTCGACGATCAC-3' binds 16 bp downstream of *msmeg_1728*, used to amplify right flank for *Δmsmeg_1700-1726* plasmid
1692-1726LF1	5'-GAGATATCCGGCCTCGCAGTCGTTG-3' binds at bp 595 of *msmeg_1691*, contains engineered EcoRV site, used to amplify left flank for *Δmsmeg_1692-1726* plasmid
1692-1726LF2	5'-GAGCTAGCGGCCTGAAGTCAAACAC-3' binds 114 bp upstream of *msmeg_1691*, contains engineered NheI site, used to amplify left flank for *Δmsmeg_1692-1726* plasmid
1677-1726LF1	5'-GAGATATCGTGCTTCGTGACTCTG-3' binds at bp 309 of *msmeg_1678*, contains engineered EcoRV site, used to amplify left flank for *Δmsmeg_1677-1726* plasmid
1677-1726LF2	5'-GAGCTAGCGTCCGAGAGGGCTAGG-3' binds 11 bp downstream of *msmeg_1678*, contains engineered NheI site, used to amplify left flanks for *Δmsmeg_1677-1726* and *Δmsmeg_1677-1692* plasmids
1677LF1NdeI	5'-GACATATGCGTGCTTCGTGACTCTG-3' binds at bp 308 of *msmeg_1678*, contains engineered NdeI site, used to amplify left flank for *Δmsmeg_1677-1692* plasmid
1692RF1	5'-GAGCTAGCGGGACGCTGATCACTC-3' binds 10 bp downstream of *msmeg_1693*, contains engineered NheI site, used to amplify right flank for *Δmsmeg_1677-1692* plasmid
1692RF2	5'-GACATATGGCAGTTGCTCGACATCAC-3' binds at bp 1,008 of *msmeg_1693*, contains engineered NdeI site, used to amplify right flank for *Δmsmeg_1677-1692* plasmid
1684LF1	5'-AACATATGCGCAACTGGGTGTGCCGTATCACTG-3' binds at bp 758 of *msmeg_1686*, contains engineered NdeI site, used to amplify left flank for *Δmsmeg_1684* plasmid
1684LF2	5'-AAGCTAGCAGCAGCCATGCGGCACAGCCTAAC-3' binds at bp 9 of *msmeg_1684*, contains engineered NheI site, used to amplify left flank for *Δmsmeg_1684* plasmid
1684RF1	5'-ATGCTAGCTCCCGGCTCCGTCAGGAGTAGCG-3' binds at bp 21 of *msmeg_1684*, contains engineered NheI site, used to amplify right flank for *Δmsmeg_1684* plasmid
1684RF2	5'-AACATATGAGCCACCCGGCGAAATTGAAGCCAC-3' binds at bp 614 of *msmeg_1683*, contains engineered NdeI site, used to amplify right flank for *Δmsmeg_1684* plasmid
secA2RTF2	5'-GCTGTCGGAGGTCAAAGG-3' binds at bp 93 of *secA2*, used for quantitative RT-PCR
secA2RTR2	5'-GCGAGAAACTGCGTGATG-3' binds at bp 227 of *secA2*, used for quantitative RT-PCR
1677F1	5'-CGAAGCGAATCGTCAACTC-3' binds at bp 183 of *msmeg_1677*, used for quantitative RT-PCR
1677R1	5'-TCGTTGGCGTTCATGTTGG-3' binds at bp 338 of *msmeg_1677*, used for quantitative RT-PCR
1682F1	5'-GTGGCGGATTATCTGGTCTC-3'

	binds at bp 238 of *msmeg_1682*, used for quantitative RT-PCR
1682R1	5’-TCACGGGCTTCTGGAATG-3’
	binds at bp 415 of *msmeg_1682*, used for quantitative RT-PCR
1683-1F	5’-CCGATGCTGGAACAACAC-3’
	binds at bp 28 of *msmeg_1683*, used for quantitative RT-PCR
1683-1R	5’-CAGGCTCCACCAGAGATTC-3’
	binds at bp 192 of *msmeg_1683*, used for quantitative RT-PCR
rpoBF	5’-GTCTCTAGCCAGAGCAAGTC-3’
	binds at bp 25 of *rpoB* (*msmeg_1367*) coding sequence, used for quantitative RT-PCR
rpoBR	5’-TCGAAGGAATCCGTCTGAAC-3’
	binds at bp 158 of *rpoB* (*msmeg_1367*) coding sequence, used for quantitative RT-PCR

Table 3.3. *Mycobacterium smegmatis* strains used in this study.

***M. smegmatis* strain**	**Description**	**Source**
mc^2155	wild type	(35)
NR116	Δ*secA2*	(18)
NR160	NR116+pMV306.kan, Δ*secA2*+empty plasmid	(18)
NR172	NR116+pYA810, Δ*secA2*+*secA2* complementing plasmid	(18)
NR178	NR116+pNR25, Δ*secA2*+*secA2 K129R* plasmid	(18)
NR242	NR178, with extragenic suppressor mutation 7S	This work (Chapter 2)
NR260	NR178, with extragenic suppressor mutation 20B	This work (Chapter 2)
NR236	NR178, with extragenic suppressor mutation 4S	This work (Chapter 2)
NR151	NR178, with extragenic suppressor mutation 24S	This work (Chapter 2)
NR123	NR116+pNR7, with extragenic suppressor mutation	(23)
NR154	NR178, with extragenic suppressor mutation 29S	This work (Chapter 2)
NR155	NR178, with extragenic suppressor mutation 33S	This work (Chapter 2)
NR156	NR178, with extragenic suppressor mutation 33B	This work (Chapter 2)
NR230	NR178, with extragenic suppressor mutation 1S	This work (Chapter 2)
NR234	NR178, with extragenic suppressor mutation 3S	This work (Chapter 2)
NR248	NR178, with extragenic suppressor mutation 10S	This work (Chapter 2)
LL036	Δ*msmeg_1704-1712*	This work
LL062	Δ*msmeg_1704-1712,* Δ*secA2*	This work
LL051	Δ*msmeg_1700-1726*	This work
LL066	Δ*msmeg_1700-1726,* Δ*secA2*	This work
LL082	Δ*msmeg_1692-1726*	This work
LL089	Δ*msmeg_1692-1726,* Δ*secA2*	This work
LL129	Δ*msmeg_1677-1726*	This work
LL132	Δ*msmeg_1677-1726,* Δ*secA2*	This work
CM14	Δ*msmeg_1677-1692*	This work
CM11	Δ*msmeg_1677-1692,* Δ*secA2*	This work
BAF1	Δ*msmeg_1684,* Δ*secA2*	This work

Table 3.4. Genes found in suppressor 7S and 20B deletions[a].

Msmeg_ Gene #	Coding Strand	Annotation	Exported Protein[c]
1678	+	transcriptional regulator, LysR family protein	No
1677	-	aspartate ammonia-lyase	No
1679	-	AmiB	No
1680	-	conserved hypothetical protein	No
1681	-	endoribonuclease L-PSP superfamily protein	No
1682	-	flavin-containing monooxygenase FMO	No
1683	-	cytosine/purine/uracil/thiamine/allantoin permease family protein	Yes
1684	-	conserved hypothetical protein	No
1685	-	hypothetical protein	No
1686	-	NPL/P60-family secreted protein	No
1687	-	hypothetical protein	No
1688	-	cupin domain protein	No
1689	-	3-oxoacyl-[acyl-carrier-protein] reductase	No
1690	-	putative ECF sigma factor RpoE1	No
1691	-	transcriptional regulatory protein	No
1692	+	ECF-family RNA polymerase sigma factor	No
1693	-	succinate dehydrogenase [ubiquinone] flavoprotein subunit	No
1694	+	uracil phosphoribosyltransferase	No
1695	-	phosphoglucomutase/phosphomannomutase	No
1696	-	regulatory protein, MarR	No
1698	+	putative ammonia monooxygenase superfamily protein	Yes
1697	-	hypothetical protein	No
1699	-	pentachlorophenol 4-monooxygenase	Yes
1700	+	TetR-family protein transcriptional regulator	No
1701	-	purine nucleoside phosphorylase	No
1702	+	Amidohydrolase	No
1703	+	Amidohydrolase	No
1704	+	ABC transporter	Yes
1705	+	D-xylose transport ATP-binding protein XylG	No
1706	+	xylose transport sytem permease protein XylH	Yes
1707	-	phosphatase YfbT	No
1708	-	ribose operon repressor, putative	No
1709	-	inner membrane ABC transporter permease protein YjfF	Yes
1710	-	ribose transport system permease protein RbsC	Yes
1711	-	ATP binding protein of ABC transporter	No
1712	-	ABC transporter periplasmic-binding protein YtfQ	Yes
1713	+	L-ribulokinase	No
1714	+	L-ribulose-5-phosphate 4-epimerase UlaF	No
1715	+	L-arabinose isomerase	No
1716	+	IS3 family element, transposase orfA	No
1718[b]	+	IS3 family element, transposase orfB, interruption-N	N/A
1717	-	ISMsm8, transposase	No
1719[b]	+	IS3 family element, transposase orfB, interruption-C	N/A
1720	+	IS1137, transposase orfA	No
1721	+	IS1137, transposase orfB	No
1722	-	hypothetical protein	Yes
1723	+	hypothetical protein	Yes
1724	-	conserved hypothetical protein	No
1725	+	putative transposase	No
1726	+	Transposase	No

[a] Genes located above red line are deleted in suppressor 7S, but not in suppressor 20B.

[b] Previously annotated as a gene on SmegmaList, but annotation has since been removed.

[c] Proteins predicted to be exported based on presence of a signal peptide (by SignalP v4.1 (36)) or transmembrane domain (by TMHMM v2.0 (37)).

Table 3.5. Suppressors affecting *msmeg_1684.*

Suppressor Strain	Suppressor Strength	Mutation	Effect on Msmeg_1684 Protein [a]
10S (NR248)	Strong	T→C SNP 6 bp upstream of *msmeg_1684*	Possible effect on promoter or ribosome binding site
33S (NR155)	Strong	CCGCGCT duplication beginning at bp 67 of *msmeg_1684*	Frameshift beginning at Trp25
33B (NR156)	Strong	C→A SNP at bp 369 of *msmeg_1684*	Cys123→STOP substitution, causes truncation
3S (NR234)	Intermediate	G→A SNP at bp 401 of *msmeg_1684*	Gly134→Asp substitution
29S (NR154)	Strong	C insertion at bp 782 of *msmeg_1684*	Frameshift beginning at Pro261
1S (NR230)	Intermediate	ACCGGC deletion beginning at bp 805 of *msmeg_1684*	Deletion of Gly269 & Thr270 [b]
20B (NR260)	Strong	Deletion of *msmeg_1684-1726* (removes bp 1-1180 of *msmeg_1684*)	Deletion of Met1 to Val394, should completely prevent Msmeg_1684 production
7S (NR242)	Strong	Deletion of *msmeg_1678-1726*	Complete deletion

[a] Amino acid numbering based on NCBI GenBank accession number ABK72696. The *msmeg_1684* gene is 1296 bp in length and the Msmeg_1684 protein is 431 amino acids in length.

[b] Note that amino acids Gly269 and Thr270 are part of a short repetitive region, "TGTGT," therefore the precise amino acids (or bp) deleted cannot be determined.

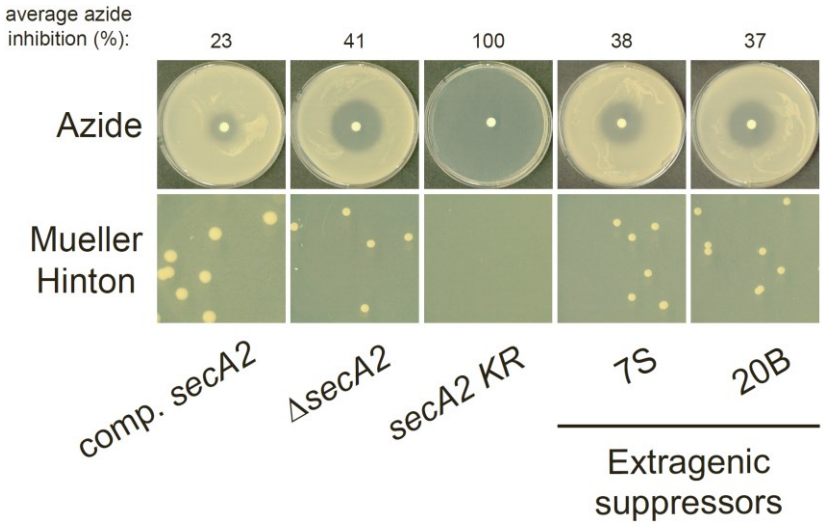

Figure 3.1. Azide sensitivity and Mueller-Hinton agar growth phenotypes. The indicated strains were plated for sensitivity to 10 µL of 0.15 M sodium azide for two days at 37°C and growth on Mueller-Hinton agar plates for six days at 37°C. Average azide inhibition is calculated by measuring the diameter of the zone of azide inhibition as a percentage of the plate diameter and is the mean of three technical replicates. Results shown are representative of at least three independent experiments. Strains tested were: complemented *secA2* (NR172), Δ*secA2*+empty plasmid (NR160), *secA2 K129R* (NR178), 7S (NR242), and 20B (NR260).

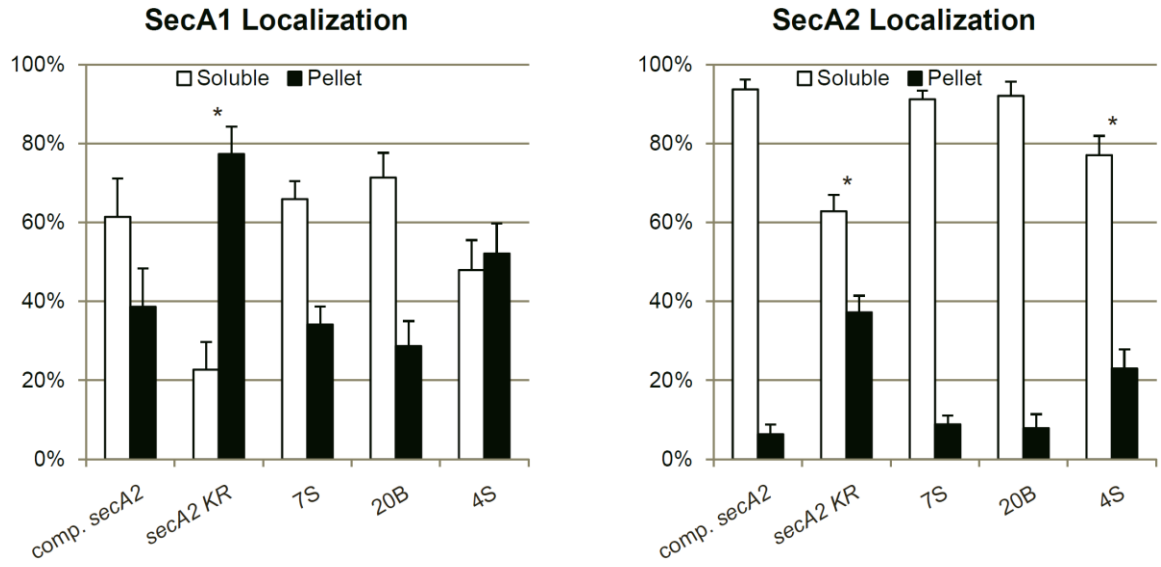

Figure 3.2. Suppressors affect subcellular localization of SecA1 and SecA2. The percentage of SecA1 or SecA2 detected by western blot in soluble (cytoplasmic) or pellet (cell envelope) fractions is plotted as a percentage of the total protein detected. Data represents the mean of six biological replicates, and error bars represent standard error. For SecA1 localization, the *secA2 K129R* strain is statistically different from all other strains (*p≤0.018) by one-way ANOVA with Student-Newman-Keuls test. For SecA2 localization, both the *secA2 K129R* and 4S strains are statistically different from all other strains (*p≤0.021). Strains tested were: complemented *secA2* (NR172), *secA2 K129R* (NR178), 7S (NR242), 20B (NR260), and 4S (NR236).

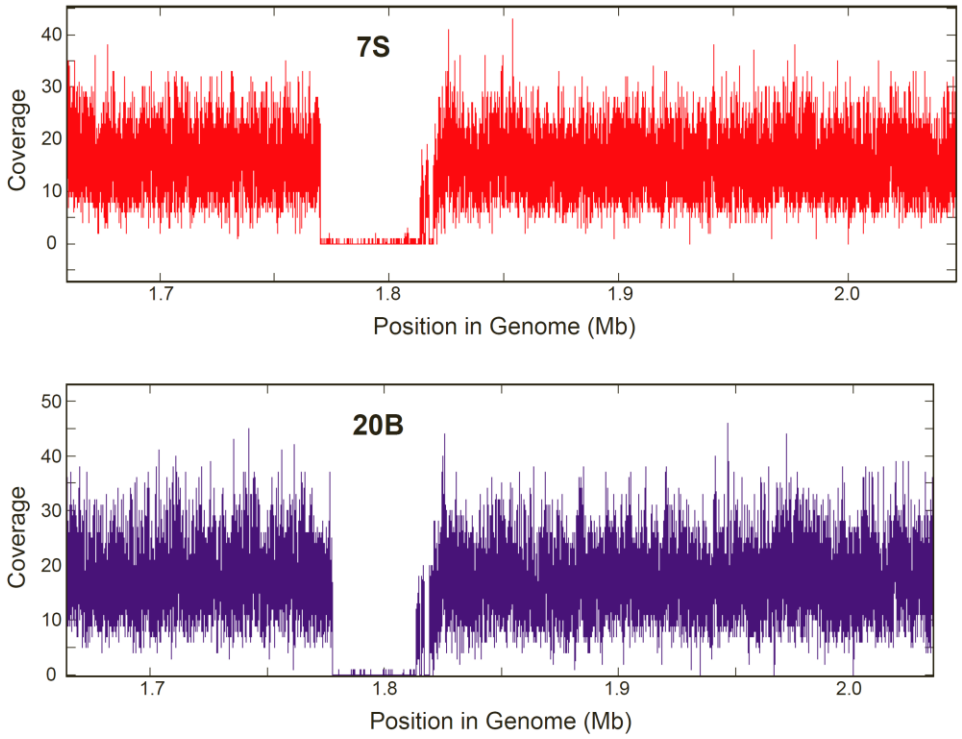

Figure 3.3. Suppressors 7S and 20B contain large deletions. Sequencing read coverage (number of times a given location was sequenced) plotted by position in the *M. smegmatis* mc^2155 reference genome. Both suppressors contain large gaps in sequencing coverage, indicating that the unsequenced region has been deleted from the suppressor genome.

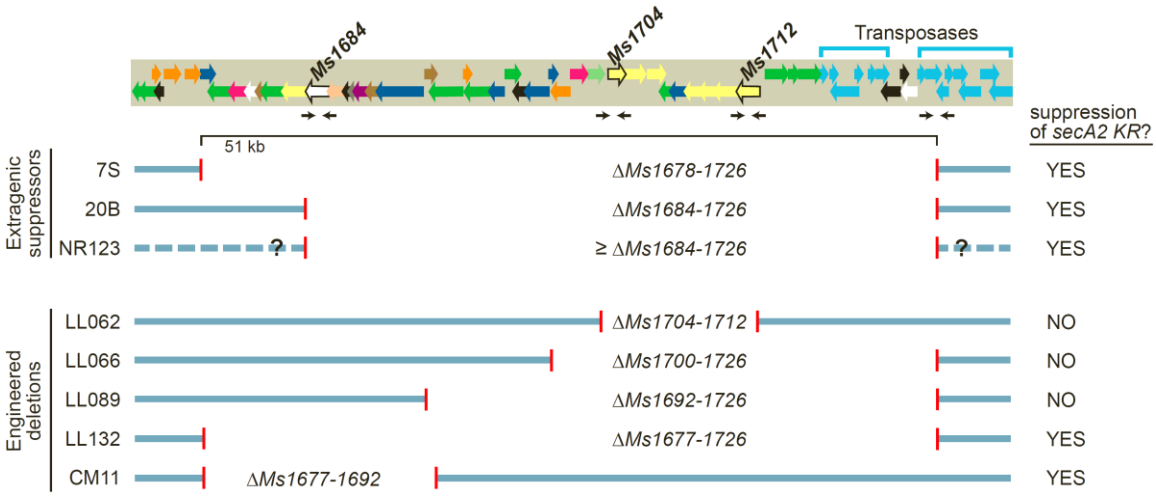

Figure 3.4. Genomic region affected by 7S and 20B suppressors. Genes of interest are indicated at the top. The first 3 lines represent naturally occurring extragenic suppressors of *secA2 KR* alleles. Small black arrows indicate the location of PCR primer pairs used to confirm deletions. The region deleted in suppressor NR123 is at least as large as that in suppressor 20B, but the exact boundaries of the NR123 deletion are unknown. The last 5 lines represent engineered deletions. Of these, only Δ*msmeg_1677-1726* and Δ*msmeg_1677-1692* are able to suppress *secA2 K129R* phenotypes.

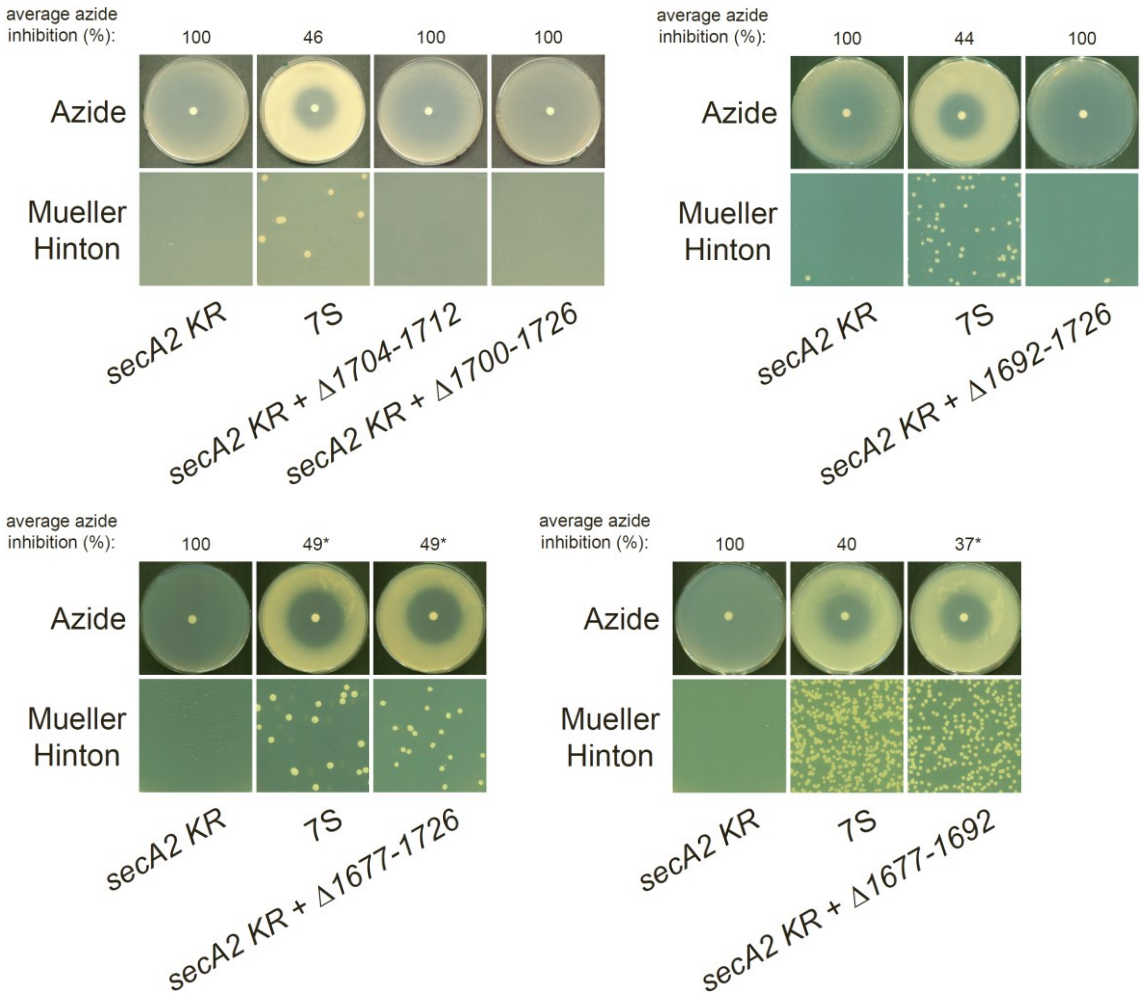

Figure 3.5. Strains Δ*msmeg_1677-1726* and Δ*msmeg_1677-1692* are suppressors of *secA2 K129R* phenotypes. Azide sensitivity and Mueller-Hinton agar growth phenotypes. Average azide inhibition is calculated by measuring the diameter of the zone of azide inhibition as a percentage of the plate diameter and is the mean of one to three biological replicates (with three technical replicates for each biological replicate). Each panel represents a separate experiment. *Statistically significant suppression ($p \leq 0.002$) by one-sample t-test (calculated versus 100% azide inhibition). (Note that when suppressor 7S is *not* marked as significant, it is because only a single biological replicate of 7S was tested in that experiment.) Strains tested were: *secA2 K129R* (NR178), 7S (NR242), *secA2K 129R* + Δ*msmeg_1704-1712* (LL062+pNR25), *secA2 K129R* + Δ*msmeg_1700-1726* (LL066+pNR25), *secA2 K129R* + Δ*msmeg_1692-1726* (LL089+pNR25), *secA2 K129R* + Δ*msmeg_1677-1726* (LL132+pNR25), and *secA2 K129R* + Δ*msmeg_1677-1692* (CM11+pNR25).

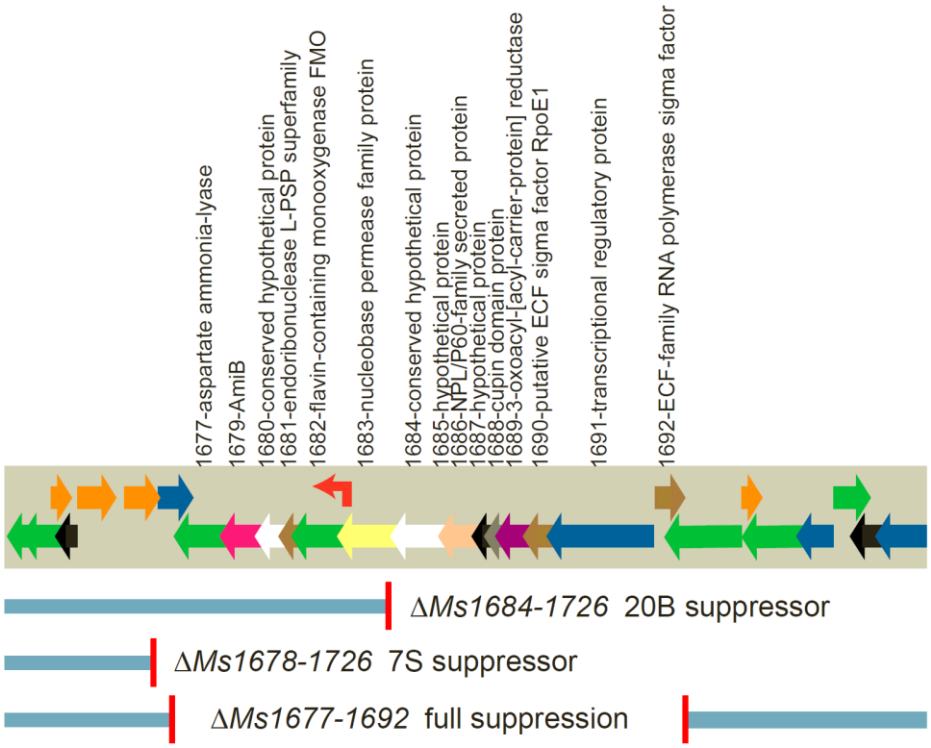

Figure 3.6. A 15-gene region is responsible for suppression. Protein annotations are indicated at the top. Naturally occurring suppressors 20B and 7S contain deletions beginning at *msmeg_1684* and *msmeg_1678* respectively. An artificially constructed deletion of genes *msmeg_1677-1692* also exhibits full suppression of *secA2 K129R*. (Note that *msmeg_1678* is located to the left of *msmeg_1677*.) Red bent arrow represents a putative promoter upstream of *msmeg_1682*.

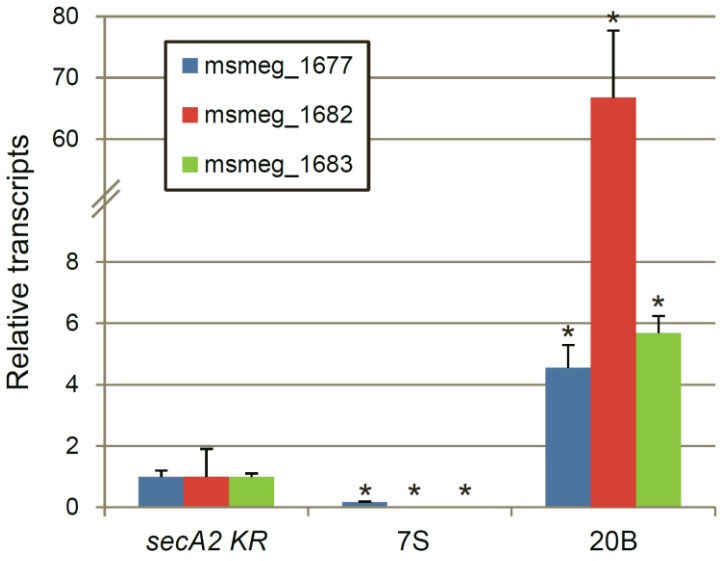

Figure 3.7. Genes *msmeg_1677-1683* are expressed in suppressor 20B. *msmeg_1677, msmeg_1682, and msmeg_1683* transcript levels measured by quantitative RT-PCR, relative to *rpoB* transcript levels. Relative transcript levels in the *secA2 K129R* strain are set to 1. Data represents the mean of three biological replicates, and error bars represent standard error. *Statistically different from the *secA2 K129R* strain ($p<0.05$) by one-way ANOVA on ranks with Student-Newman-Keuls test. Strains tested were: *secA2 K129R* (NR178), 7S (NR242), and 20B (NR260).

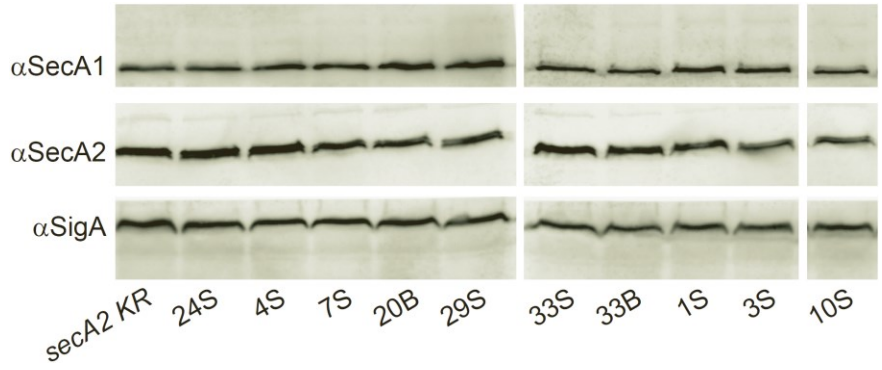

Figure 3.8. Suppressor mutations do not affect SecA1 and SecA2 protein levels. Whole cell lysates were separated by SDS-PAGE and SecA1, SecA2, and SigA (as a loading control) proteins detected by western blotting. All samples were equally loaded. Strains tested were: complemented *secA2* (NR172), Δ*secA2*+empty plasmid (NR160), *secA2 K129R* (NR178), 24S (NR151), 4S (NR236), 7S (NR242), 20B (NR260), 29S (NR154), 33S (NR155), 33B (NR156), 1S (NR230), 3S (NR234), and 10S (NR248).

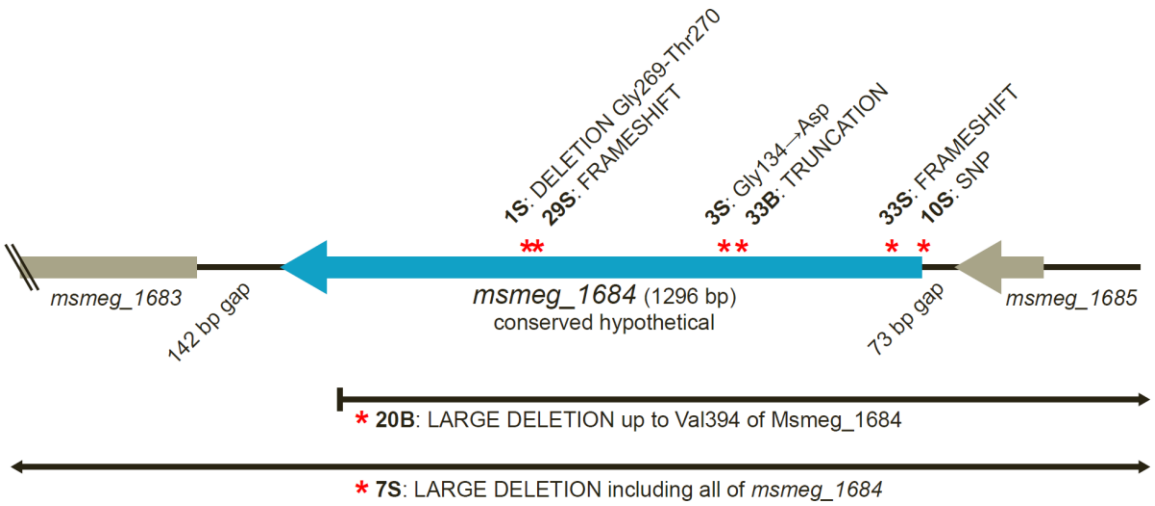

Figure 3.9. Eight suppressor mutations affect *msmeg_1684*. Red stars indicate approximate locations of suppressor mutations. Suppressors 20B and 7S are indicated by black bars and contain large deletions encompassing *msmeg_1684-1726* and *msmeg_1678-1726* respectively. Drawing not to scale.

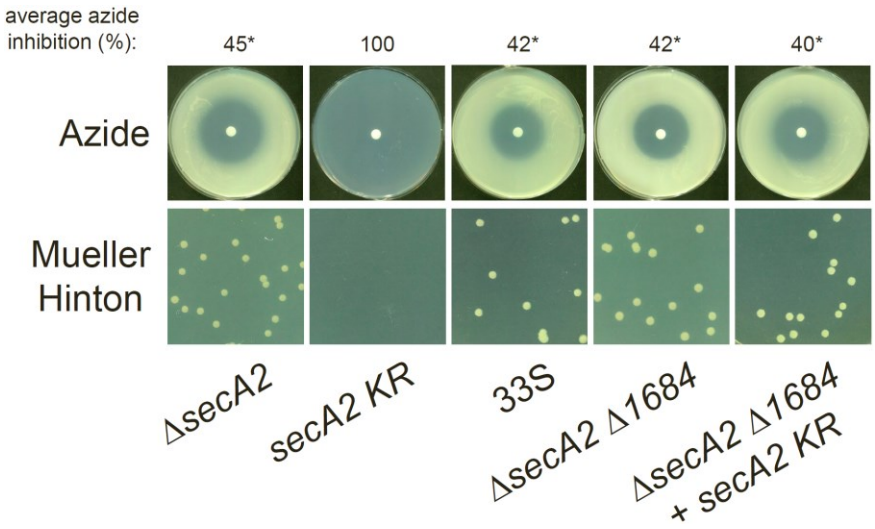

Figure 3.10. Δ*msmeg_1684* can suppress *secA2 K129R* phenotypes. Azide sensitivity and Mueller-Hinton agar growth phenotypes. Average azide inhibition is calculated by measuring the diameter of the zone of azide inhibition as a percentage of the plate diameter and is the mean of three to four biological replicates (with three technical replicates for each biological replicate). *Statistically different from *secA2 K129R* strain ($p<0.001$) by one-way ANOVA with Student-Newman-Keuls test. Strains tested were (from left to right): Δ*secA2*+empty plasmid (NR160), *secA2 K129R* (NR178), 33S (NR155), Δ*secA2* Δ*msmeg_1684*+empty plasmid (BAF1+pMV306.kan), and Δ*secA2* Δ*msmeg_1684*+*secA2 K129R* (BAF1+pNR25).

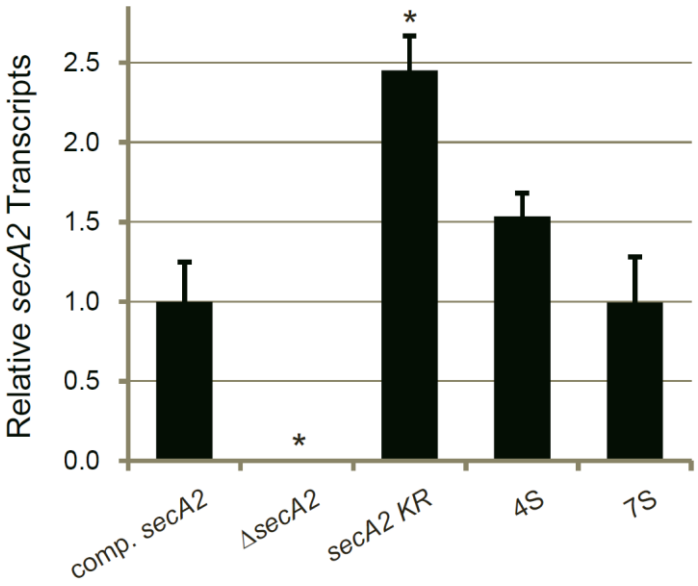

Figure 3.11. Export-related stress causes increased *secA2* expression from the *hsp60* promoter. *secA2* transcript levels measured by quantitative RT-PCR, relative to *rpoB* transcript levels. *secA2* transcript level in the complemented *secA2* strain is set to 1. Data represents the mean of three biological replicates, and error bars represent standard error. Both the Δ*secA2* and *secA2 K129R* strains are statistically different from all other strains (*p≤0.017) by one-way ANOVA with Student-Newman-Keuls test. Strains tested were: complemented *secA2* (NR172), Δ*secA2*+empty plasmid (NR160), *secA2 K129R* (NR178), 4S (NR236), and 7S (NR242).

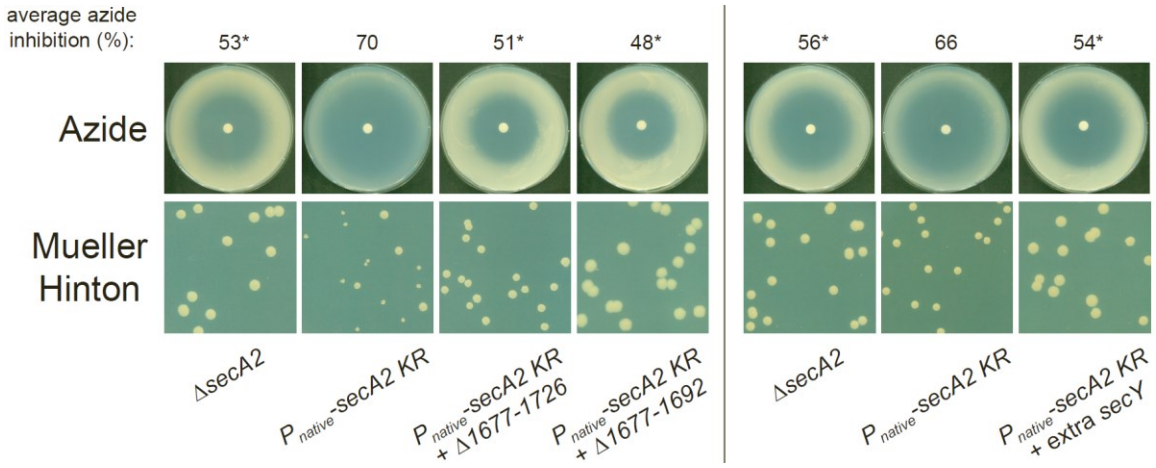

Figure 3.12. Suppressor strains can suppress *secA2 K129R* expressed from its native promoter. Azide sensitivity and Mueller-Hinton agar growth phenotypes. Average azide inhibition is calculated by measuring the diameter of the zone of azide inhibition as a percentage of the plate diameter and is the mean of three to four biological replicates (with three technical replicates for each biological replicate). *Statistically different from *secA2 K129R* strain ($p<0.001$) by one-way ANOVA with Student-Newman-Keuls test. Strains tested were (from left to right): Δ*secA2*+empty plasmid (NR160), P_{native}-*secA2 K129R* (NR116+pLL36), P_{native}-*secA2 K129R* + Δ*msmeg_1677-1726* (LL132+pLL36), P_{native}-*secA2 K129R* + Δ*msmeg_1677-1692* (CM11+pLL36), Δ*secA2*+empty plasmid (NR160+pLL2), P_{native}-*secA2 K129R* (NR116+pLL2+pLL36), and P_{native}-*secA2 K129R* + extra *secY* (NR116+pLL23+pLL36).

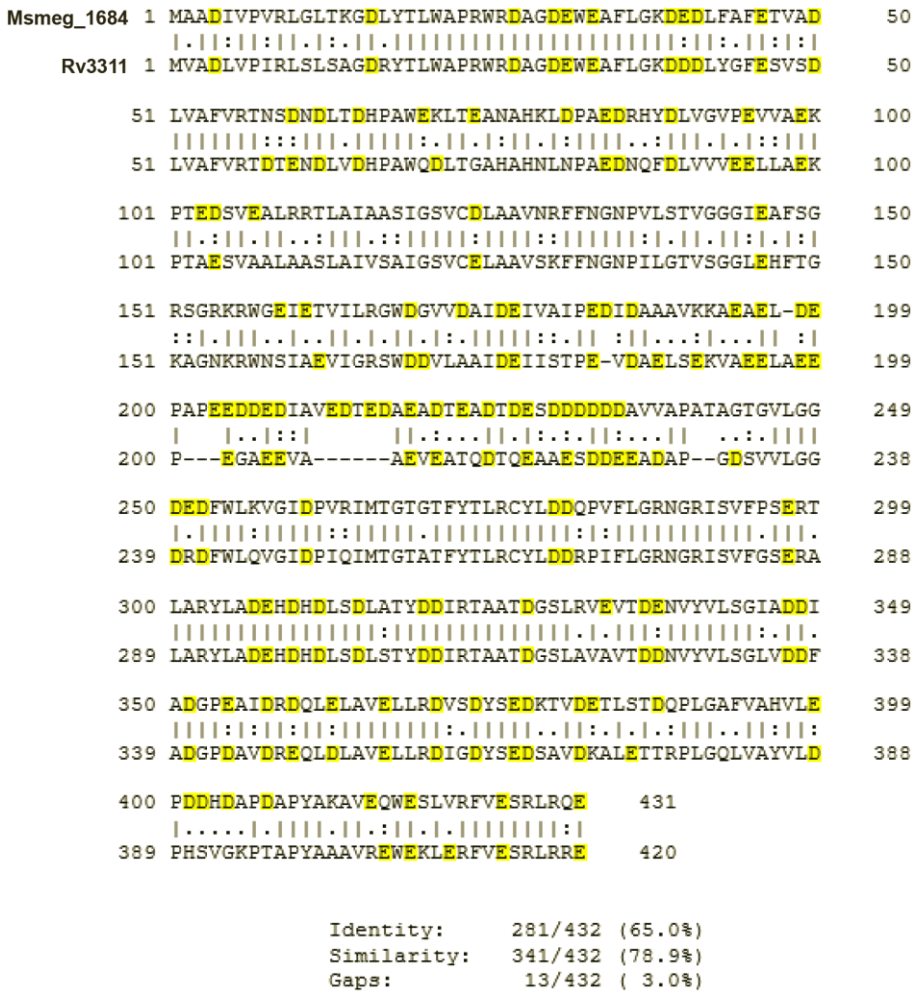

Figure 3.13. Alignment of Msmeg_1684 and Rv3311 proteins. Msmeg_1684 protein sequence was aligned to the homologous Rv3311 protein sequence from *M. tuberculosis* using the EBLOSUM62 algorithm available from EMBOSS Needle. Acidic residues are highlighted in yellow.

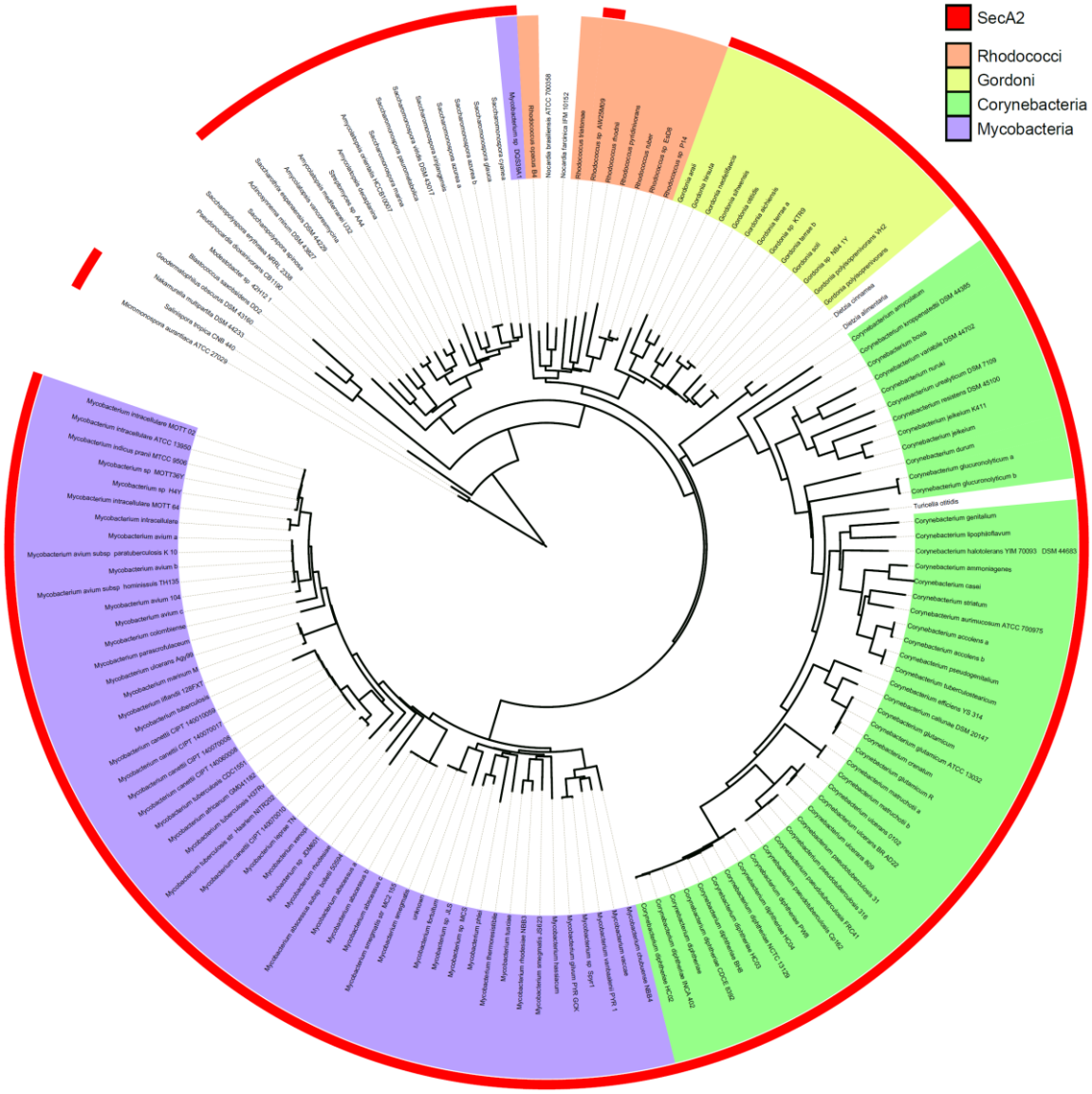

Figure 3.14. Phylogenetic tree of Msmeg_1684-like proteins showing co-occurrence of SecA2 proteins. All proteins homologous to Msmeg_1684 were identified by NCBI Protein BLAST and exported as a phylogenetic tree, which was then visualized using iTOL. Color blocks indicate groups of related organisms and the presence of a SecA2 protein in each species is indicated by a red bar.

References

1. **Ligon LS, Hayden JD, Braunstein M.** 2012. The ins and outs of *Mycobacterium tuberculosis* protein export. Tuberculosis (Edinb) **92:**121-132.

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

3. **Flower AM, Hines LL, Pfennig PL.** 2000. SecG is an auxiliary component of the protein export apparatus of *Escherichia coli*. Mol Gen Genet **263:**131-136.

4. **Schiebel E, Driessen AJ, Hartl FU, Wickner W.** 1991. Delta mu H+ and ATP function at different steps of the catalytic cycle of preprotein translocase. Cell **64:**927-939.

5. **van der Wolk JP, de Wit JG, Driessen AJ.** 1997. The catalytic cycle of the *escherichia coli* SecA ATPase comprises two distinct preprotein translocation events. EMBO J **16:**7297-7304.

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

7. **Nguyen-Mau SM, Oh SY, Kern VJ, Missiakas DM, Schneewind O.** 2012. Secretion genes as determinants of *Bacillus anthracis* chain length. J. Bacteriol. **194:**3841-3850.

8. **Caspers M, Freudl R.** 2008. *Corynebacterium glutamicum* possesses two *secA* homologous genes that are essential for viability. Archives of microbiology **189:**605-610.

9. **Fagan RP, Fairweather NF.** 2011. *Clostridium difficile* has two parallel and essential Sec secretion systems. J. Biol. Chem. **286:**27483-27493.

10. **Braunstein M, Brown AM, Kurtz S, Jacobs WR, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. J Bacteriol **183:**6979-6990.

11. **Bensing BA, Sullam PM.** 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.

12. **Siboo IR, Chaffin DO, Rubens CE, Sullam PM.** 2008. Characterization of the accessory Sec system of *Staphylococcus aureus*. J Bacteriol **190:**6188-6196.

13. **Rigel NW, Braunstein M.** 2008. A new twist on an old pathway--accessory Sec [corrected] systems. Mol Microbiol **69:**291-302.

14. **Braunstein M, Espinosa BJ, Chan J, Belisle JT, Jacobs WR, Jr.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of Mycobacterium tuberculosis. Mol Microbiol **48:**453-464.

15. **Sullivan JT, Young EF, McCann JR, Braunstein M.** 2012. The Mycobacterium tuberculosis SecA2 system subverts phagosome maturation to promote growth in macrophages. Infect Immun **80:**996-1006.

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

17. **Hou JM, D'Lima NG, Rigel NW, Gibbons HS, McCann JR, Braunstein M, Teschke CM.** 2008. ATPase activity of *Mycobacterium tuberculosis* SecA1 and SecA2 proteins and its importance for SecA2 function in macrophages. J Bacteriol **190:**4880-4887.

18. **Rigel NW, Gibbons HS, McCann JR, McDonough JA, Kurtz S, Braunstein M.** 2009. The Accessory SecA2 System of Mycobacteria Requires ATP Binding and the Canonical SecA1. J Biol Chem **284:**9927-9936.

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

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

21. **Larsen MH, Biermann K, Tandberg S, Hsu T, Jacobs WR, Jr.** 2007. Genetic Manipulation of *Mycobacterium tuberculosis*. Curr. Protoc. Microbiol. **Chapter 10:**Unit 10A 12.

22. **Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR.** 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene **77:**61-68.

23. **Rigel NW.** 2008. Characterization of the SecA2 Protein Export Pathway of Mycobacteria. University of North Carolina at Chapel Hill, Chapel Hill, NC.

24. **Pavelka MS, Jr., Jacobs WR, Jr.** 1999. Comparison of the construction of unmarked deletion mutations in Mycobacterium smegmatis, Mycobacterium bovis bacillus Calmette-Guerin, and Mycobacterium tuberculosis H37Rv by allelic exchange. J Bacteriol **181:**4780-4789.

25. **Braunstein M, Bardarov SS, Jacobs WR, Jr.** 2002. Genetic methods for deciphering virulence determinants of Mycobacterium tuberculosis. Methods Enzymol **358:**67-99.

26. **Badejo AC, Badejo AO, Shin KH, Chai YG.** 2013. A gene expression study of the activities of aromatic ring-cleavage dioxygenases in Mycobacterium gilvum PYR-GCK to changes in salinity and pH during pyrene degradation. PLoS One **8:**e58066.

27. **Cabelli RJ, Dolan KM, Qian LP, Oliver DB.** 1991. Characterization of membrane-associated and soluble states of SecA protein from wild-type and SecA51(TS) mutant strains of Escherichia coli. J Biol Chem **266:**24420-24427.

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

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

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

31. **Stover CK, de la Cruz VF, Fuerst TR, Burlein JE, Benson LA, Bennett LT, Bansal GP, Young JF, Lee MH, Hatfull GF, Snapper SB, Barletta RG, Jacobs WR, Jr., Bloom BR.** 1991. New use of BCG for recombinant vaccines. Nature **351:**456-460.

32. **Sassetti CM, Rubin EJ.** 2003. Genetic requirements for mycobacterial survival during infection. Proc Natl Acad Sci U S A **100:**12989-12994.

33. **Vergne I, Chua J, Lee HH, Lucas M, Belisle J, Deretic V.** 2005. Mechanism of phagolysosome biogenesis block by viable Mycobacterium tuberculosis. Proc Natl Acad Sci U S A **102:**4033-4038.

34. **Feltcher ME, Gibbons HS, Ligon LS, Braunstein M.** 2013. Protein export by the mycobacterial SecA2 system is determined by the preprotein mature domain. J. Bacteriol. **195:**672-681.

35. **Snapper SB, Melton RE, Mustafa S, Kieser T, Jacobs WR, Jr.** 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. Mol. Microbiol. **4:**1911-1919.

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

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

CHAPTER 4

Searching for Evidence of a Physical Interaction Between *Mycobacterium smegmatis* SecA2 and SecY[1]

Bacteria use the conserved Sec protein export system to transport proteins across the cell membrane. In this system, the motor protein SecA interacts with the membrane channel complex SecYEG and harnesses energy from ATP hydrolysis to push proteins through the SecYEG channel and across the membrane. In mycobacteria, a second SecA protein, called SecA2, is responsible for a subset of protein export. Unlike the canonical Sec system, the mechanism of protein export in the mycobacterial SecA2 system is not well understood. However, SecA2-dependent exported proteins are thought to be exported through the canonical SecYEG membrane channel complex with the assistance of SecA1 (Chapter 2). This suggests a model for SecA2-dependent export in which SecA2 interacts physically with SecY during export, analogous to the physical interaction of *Escherichia coli* SecA with SecY during canonical Sec export. Here, we evaluate this model by performing *in vivo* protein crosslinking of *M. smegmatis* SecA2 and a dominant negative variant of SecA2 (SecA2 K129R). Crosslinked complexes containing SecA2 proteins were purified and tested for co-purification of SecY. In no case were we able to detect SecY crosslinked to SecA2 proteins. Our findings suggest an alternate model for SecA2-

[1]Contributing authors: Henry Gibbons and Miriam Braunstein (Department of Microbiology and Immunology, School of Medicine, University of North Carolina, Chapel Hill, North Carolina, 27599-7290)

dependent protein export, in which SecA2 and SecY proteins both participate in the SecA2 system, but do not physically interact or form stable complexes during protein export.

Introduction

Bacteria must place proteins on their cell surface and in their extracellular environment in order to survive. Many of these exported proteins are required for essential cell functions such as nutrient acquisition, while others are important for pathogenesis. All bacteria contain the conserved Sec protein export system, which is responsible for the bulk of protein export. This system transports unfolded proteins across the cell membrane through a membrane-embedded channel complex, composed of SecY, SecE, and SecG (1, 2). Protein export is driven by the motor protein, SecA, which is an ATPase that pushes proteins through the SecYEG translocon using energy derived from successive rounds of ATP binding and hydrolysis (3, 4). In *Escherichia coli* and other bacteria, it has been shown that SecA engages in a direct physical interaction with SecY during protein export (5-8). This interaction is integral to the Sec export mechanism, inducing activation of SecA ATPase activity (9, 10) and possibly opening of the SecYEG channel (6).

In addition to the canonical Sec system, Mycobacteria and some Gram-positive bacteria also contain a specialized export system characterized by the presence of a second, functionally distinct SecA protein (11-17). In these cases, the "housekeeping" SecA protein is referred to as SecA1, while the second SecA protein is referred to as SecA2 and is responsible for a subset of protein export. SecA2 systems are typically non-essential, but in the case of pathogens such as *Mycobacterium tuberculosis*, are often important for virulence (18, 19). Like SecA1, SecA2 is an

ATPase, and this activity is required for SecA2 to function normally (20, 21). However, the mechanism of mycobacterial SecA2-dependent protein export is poorly understood.

In the model mycobacterium *Mycobacterium smegmatis*, we have analyzed suppressor mutants of a dominant negative *secA2* allele in order to learn more about the mycobacterial SecA2 export pathway. This mutant *secA2* allele, referred to as *secA2 K129R*, encodes a SecA2 protein with an amino acid substitution in the Walker box, a region which is important for ATP binding and hydrolysis. The resulting SecA2 K129R protein is non-functional (20, 22) and appears to become "locked" in detrimental interactions with its protein binding partners. In addition, while wild type SecA2 is found primarily in the cytoplasm, SecA2 K129R exhibits a shift in localization towards the cell envelope, suggesting it is locked in an interaction with membrane-localized proteins. Mutations that suppress the phenotypes of the *secA2 K129R* allele act to avoid or compensate for these detrimental interactions and can therefore point to proteins that are connected in some way to the SecA2-dependent export pathway. Previously (Chapter 2), this approach identified a genetic relationship between SecA2 and SecY, the main component of the channel complex used for housekeeping Sec export. In the presence of SecA2 K129R, SecY protein levels dramatically decrease. This reduction in SecY levels is most likely a result of SecY degradation that is caused by "jamming" or stress on the SecYEG translocon in the presence of the non-functional SecA2 K129R protein. Mutations increasing *secY* expression are able to compensate somewhat for the ongoing degradation of SecY, and thereby suppress the phenotypes of the *secA2 K129R* allele.

These findings suggest a model in which SecA2-dependent proteins are exported through the canonical SecYEG channel. Thus, we hypothesized that SecA2 interacts directly with SecY during protein export, in the same way that SecA of other bacteria interacts with SecY to

export proteins (5, 6). Here, we tested for a physical SecA2-SecY interaction in *M. smegmatis* using *in vivo* protein crosslinking followed by purification of SecA2-containing protein complexes. Our experiments failed to demonstrate a physical interaction between SecA2 and SecY. This result suggests that the genetic relationship between SecA2 and SecY does *not* represent a direct physical interaction, but rather a transient or indirect interaction. Alternate models for the interaction of SecA2 with SecY are discussed.

Materials and Methods

Bacterial strains and culture conditions. *M. smegmatis* strains used in this study are described in Table 4.1, and were grown at 37°C, using Mueller-Hinton media supplemented with 0.1% tween-80 and kanamycin (20 µg/mL). Hygromycin B (50 µg/mL) was added as needed. To induce SecA2 K129R expression in strain LL159, cultures at an $OD_{600\,nm}$ of 0.3 were supplemented with 400 ng/mL anhydrotetracycline (Atc) and grown in the dark for 4 hours. When required, plasmids were introduced into *M. smegmatis* strains by electroporation (23). *Escherichia coli* strains were grown at 37°C in lysogeny broth or on lysogeny broth agar (otherwise known as Luria-Bertani media). Antibiotics kanamycin (40 µg/mL) and hygromycin B (50 µg/mL) were added as needed.

Plasmid construction. Plasmids used in this study are described in Table 4.2 and primers are described in Table 4.3. In all cases, newly constructed plasmids were sequenced. To construct an integrating plasmid expressing *secY* from the 24S (NR151) mutant promoter, we began with plasmid pLL17 (Chapter 2), which contains *M. smegmatis secY* on its native promoter. Site-directed mutagenesis (Stratagene QuikChange II) was performed on pLL17 to introduce the 24S *secY* promoter mutation (C→G polymorphism 134 bp upstream of the *secY*

coding sequence), resulting in plasmid pLL41. A NotI-EcoRV fragment (containing *secY* and its promoter) was then cut from pLL41 and ligated into NotI-EcoRV-digested pLL2, yielding plasmid pLL42.

Intracellular crosslinking and SecA2-His purification. Each strain was grown in Mueller-Hinton broth to an $OD_{600\ nm}$ of approximately 1.0. For each condition to be tested, 200 mL culture was pelleted at room temperature, washed, and resuspended in 2 mL PBS + 5x protease inhibitor cocktail (500x stock contains: 1 mg/ml each of aprotinin, E-64, leupeptin, and pepstatin A and 50 mg/ml Pefabloc SC in DMSO). Fresh crosslinkers were added (2.5% formaldehyde, 10 mM DSP in DMSO, or 20 mM SDAD in DMSO) and incubated for 30 minutes at room temperature, vortexing periodically to mix. Tris (1.5 M, pH 8.8) was added to a final concentration of 100 mM to quench crosslinkers and was incubated for 15 minutes at room temperature, vortexing periodically to mix. Cells were then pelleted, washed, and resuspended in 3 mL PBS + 5x protease inhibitor cocktail. Specifically for SDAD-treated samples, cells were moved to a glass flask and exposed to UV light (320-370 nm) for 5 x 3 minutes to complete crosslinking, cooling with ice during exposure. Cells were lysed by passage through a French press, and cell debris was removed by centrifugation at 1,600 x g for 30 minutes to generate clarified whole cell lysates (WCL, supernatants). A small amount of WCL was set aside for SDS-PAGE and 2.2 mL WCL was centrifuged at 100,000 x g for 2 hours to separate cell envelope (pellet) and soluble (supernatant) fractions. Cell envelope pellets (ENV) were resuspended in 2.2 mL of Qiagen lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, adjusted to pH 8.0 using NaOH) + 1% Igepal CA-630 to facilitate solubilization of membrane proteins. A small amount of ENV was set aside for SDS-PAGE.

To purify His-tagged SecA2 proteins, 2 mL of each ENV sample was mixed with 0.5 mL Qiagen Ni-NTA agarose and incubated for 1 hour at 4°C on a rocking platform. Slurry was then loaded into a column created by packing a 1 inch square of glass wool into the tip of a 3 mL plastic syringe. The slurry was allowed to settle, collecting the flow-through for future analysis. Columns were washed twice with 2 mL of Qiagen wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, adjusted to pH 8.0 using NaOH) + 1% Igepal CA-630, collecting each wash. Next, SecA2-His proteins were eluted twice with 1 mL of Qiagen elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, adjusted to pH 8.0 using NaOH) + 1% Igepal CA-630, collecting each eluate.

SDS-PAGE and western blotting. WCL, ENV, and samples from SecA2-His purification were either boiled 15 minutes in SDS-PAGE sample buffer + 5% β-mercaptoethanol (β-ME) to reverse protein crosslinks, or mixed with SDS-PAGE sample buffer lacking β-ME at room temperature to maintain intact crosslinks. Samples were then loaded for equivalent starting material and separated by SDS-PAGE. Gels were either stained with coomassie or transferred to nitrocellulose membranes for western blotting. Membranes were temporarily stained with Ponceau S to detect all protein prior to western blotting, and individual proteins were detected using a rabbit αSecY polyclonal antiserum at a 1:150 dilution (Chapter 2) and a rabbit αSecA2 polyclonal antiserum at a 1:25,000 dilution (24). Primary antibodies were detected using alkaline phosphatase-conjugated goat anti-rabbit IgG secondary antibodies, ECF substrate (GE Healthcare), and a Syngene G:BOX machine.

Results

SecY does not crosslink to SecA2-His. We previously found that SecA2 interacts genetically with SecY (Chapter 2). By analogy to the housekeeping Sec pathway (6), SecA2 could physically "dock" with SecYEG during protein export. We therefore performed *in vivo* protein crosslinking (Fig. 4.1) and evaluated whether or not crosslinking could be detected between a C-terminally His-tagged SecA2 protein and SecY. His-tagged SecA2 (expressed from the *hsp60* promoter, plasmid pHSG93) has previously been shown to complement a *secA2* deletion mutant (25), demonstrating that it is capable of normal function despite the presence of the His tag.

Membrane-permeable protein crosslinkers formaldehyde, DSP, and SDAD were chosen and were added to intact *M. smegmatis* cells. This whole cell *in vivo* crosslinking method was chosen in an effort to avoid disruption of native protein interactions, which can occur when cells are lysed prior to crosslinking. Formaldehyde is a widely used crosslinker that reacts with primary amine groups and has been used successfully for whole cell *in vivo* crosslinking of SecA to SecY in *E. coli* (8). DSP also reacts with primary amines but has a longer spacer arm (12.0 Å) than formaldehyde (2.0 Å), enabling it to react with more distant amine groups. SDAD is a photoactivatable crosslinker that also has a long spacer arm (13.5 Å). This crosslinker reacts first at one end with primary amines and then, upon UV exposure, reacts non-specifically at the opposite end, allowing it to crosslink proteins even in the absence of adjacent amine groups.

Following *in vivo* crosslinking, complexes containing SecA2-His were purified and analyzed for the presence of SecY. While SecA2 is normally found primarily in the cytoplasm (22), SecY is a membrane protein and is found in the cell envelope fraction (Fig. 2.4, Chapter 2). Therefore, any SecA2 protein that has been crosslinked to SecY will become trapped at the membrane. In order to enrich for crosslinked SecA2 complexes that might contain SecY, we first

isolated the cell envelope fraction from crosslinker-treated cells and solubilized this fraction using 1% Igepal CA-630, a detergent we found effective for solubilizing SecY (data not shown). Finally, crosslinked complexes containing SecA2-His were purified from the solubilized cell envelope fraction by allowing the His-tag to bind to a nickel-affinity column, washing away unbound proteins, and then eluting SecA2-His by addition of high levels of imidazole.

SDS-PAGE of crosslinked samples demonstrated that all crosslinkers used were active in the cell envelope. This was evidenced by changes in the banding pattern observed when crosslinked cell envelope samples are compared to untreated samples (Fig. 4.2A). In addition, all crosslinkers induced the appearance of higher molecular weight SecA2-containing species (Fig. 4.2B), indicating they successfully reacted with SecA2-His. Despite its typical localization in the cytoplasm, SecA2-His was successfully purified from the cell envelope and was present in column eluates (Fig. 4.2C). Finally, western blotting was performed to detect any SecY protein crosslinked to SecA2-His (Fig. 4.2D). The SecY monomer band was weak to undetectable in crosslinked samples, consistent with crosslinking of SecY to other proteins (though no discrete higher molecular weight bands were visible). Upon reversal of crosslinks, the SecY monomer was again detectable. However, in no case was SecY detected in the purified SecA2-His column eluate samples.

SecY does not crosslink to SecA2 K129R-His. Because the genetic relationship between SecA2 and SecY was originally detected using the *secA2 K129R* allele, we also tested whether SecY could be chemically crosslinked to SecA2 K129R-His. SecA2 K129R also exhibits a shift in localization towards the cell envelope relative to wild type SecA2 (22), consistent with increased interaction with SecY or other membrane proteins. In setting up this experiment we took into consideration the fact that the presence of SecA2 K129R is very detrimental to the cell and

results in SecY degradation (Chapter 2). In order to successfully express both SecA2 K129R-His and SecY at the same time, we used a strain carrying an anhydrotetracycline (Atc)-inducible *secA2 K129R-His* allele (strain LL159). Atc was added during the last 4 hours of culture growth to induce expression of the *secA2 K129R-His* allele. By inducing *secA2 K129R-His* expression only briefly before crosslinking, we could limit the length of time during which SecY would be subject to degradation induced by the SecA2 K129R protein. In addition, this strain is also a *secY* merodiploid, containing both the chromosomal copy of *secY* and an extra copy of *secY* driven from the overexpressing 24S mutant *secY* promoter and integrated into the chromosome at the Tweety mycobacteriophage attachment site (26). Crosslinking and purification of SecA2 K129R-His was performed as described above. Again, all crosslinkers were active in the cell envelope (Fig. 4.3A) and all crosslinkers induced the appearance of higher molecular weight SecA2-containing species (Fig. 4.3B), indicating they successfully reacted with SecA2 K129R-His. In addition, SecA2 K129R-His was successfully purified from the cell envelope fraction (Fig. 4.3C). However, as before, no SecY was detected in purified SecA2 K129R-His samples (Fig. 4.3D). Therefore, even using the SecA2 K129R protein, which we believe is trapped in non-functional complexes with the Sec pathway, no interaction between SecA2 and SecY could be detected.

Discussion

As shown previously (Chapter 2), SecA2-dependent protein export appears to be integrated into the housekeeping Sec pathway and SecA1 and SecA2 appear to share use of the SecYEG translocon to export proteins across the cell membrane. The most obvious model for this relationship is one in which SecA2 directly interacts with SecY during protein export, in the same way observed for the canonical Sec pathway. The interaction between SecY and the

canonical SecA protein has been demonstrated in *E. coli* using numerous methods (including ligand affinity/far western blotting (5), surface plasmon resonance (7), and *in vivo* crosslinking (8)) and a co-crystal structure of the *Thermotoga maritima* SecA-SecY complex has been solved (6). Here, we evaluated this model by attempting to chemically crosslink SecY to SecA2-His or SecA2 K129R-His. Under the conditions tested, no crosslinking was observed. These results suggest that the interaction of SecA2 with the housekeeping SecY protein may not be as direct as we initially hypothesized.

The primary amine-reactive crosslinker formaldehyde has been successfully used in *E. coli* to crosslink SecA to SecY *in vivo* (8) and was thus our first-line choice for crosslinking. In addition, we also used DSP, a second primary amine-reactive crosslinker with a longer spacer arm. Neither of these chemicals were successful at crosslinking SecY to either SecA2-His or SecA2 K129R-His. To assess the likelihood of obtaining amine-amine crosslinking between SecA2 and SecY, we modeled hypothetical SecA2-SecY and SecA1-SecY protein complexes onto the *T. maritima* SecA-SecY co-crystal structure (6). Next we identified the location of all primary amines (i.e. all lysine residues and the protein N-terminus) in both the SecA1-SecY and SecA2-SecY modeled structures. For the modeled SecA1-SecY structure, there were SecA1 amine groups in close proximity to SecY amine groups, suggesting an amine-amine crosslinker would successfully crosslink these two proteins. However, for the modeled SecA2-SecY structure, there were no closely spaced pairings of SecA2 amine groups with SecY amine groups observed. This suggests that an amine-amine crosslinker, such as Formaldehye or DSP, might fail to crosslink these two proteins. To account for this potential disadvantage, we additionally used a photo-activatable crosslinker, SDAD. While this crosslinker reacts at one end with a primary amine group (which is readily available in either SecA2 or SecY), the second end of the SDAD crosslinker reacts non-specifically upon exposure to UV light, allowing it to react with any

protein in close proximity at the time of UV exposure. This feature should overcome the lack of closely spaced amine group pairs in the hypothetical SecA2-SecY complex.

Despite the use of a variety of carefully chosen chemical crosslinkers, no crosslinking was observed between SecY and either SecA2-His or SecA2 K129R-His. This suggests that SecA2 does not participate in a direct physical interaction with SecY. However, there are many possible explanations for this finding. One possibility is that SecA2 and SecY *do* interact physically, but that this interaction could not be detected under the present experimental conditions. There are many variables that can affect the success of a crosslinking experiment, such as type of crosslinker, crosslinker concentration, exposure time and temperature, and protein context at the time of crosslinking (i.e. *in vivo* vs. *in vitro* crosslinking). While it would be technically prohibitive to try all possible crosslinking conditions, the conditions used here were carefully chosen to maximize our chances of finding a SecA2-SecY interaction, yet they failed to identify an interaction.

The lack of a detectable interaction between SecA2 and SecY may indicate that any interactions between these two proteins are weak or transient in nature. The fact that SecA2 is found predominantly in the cytoplasm (Fig. 3.2, Chapter 3) is consistent with this conclusion. In contrast, SecA1, which is believed to interact directly with SecY, is found evenly distributed between the cell envelope and the cytoplasm. It would be interesting to test whether a physical interaction can be detected between SecA1 and SecY in *M. smegmatis*. Perhaps SecA2 interacts with SecY only briefly to deliver SecA2-dependent proteins, allowing SecA1 to take over and drive export of these proteins through the SecY channel and across the membrane. In fact, SecA1 appears to be required for SecA2-dependent protein export ((25) and Fig. 2.7, Chapter 2).

Relative to wild type SecA2, the SecA2 K129R protein variant is located to a larger extent in the cell envelope, suggesting it is trapped in an interaction with a membrane protein(s) such as SecY. Therefore, it was possible that using SecA2 K129R in our crosslinking experiments would improve our ability to detect a SecY interaction. However, we previously found that the presence of SecA2 K129R results in degradation of SecY (Chapter 2). To work around this potential problem, we expressed SecA2 K129R only briefly from an inducible promoter and also used a *secY* merodiploid strain (LL159) that overexpresses *secY*. While this strain exhibited normal levels of SecY protein, no interaction was detected between SecA2 K129R and SecY. However, we do not know how quickly SecY is degraded following interaction with SecA2 K129R. It is possible that SecA2 K129R-SecY complexes do form, but are immediately degraded, and that any remaining SecY protein has not yet engaged in an interaction with SecA2 K129R. This could be another reason we were unable to detect crosslinking of SecY to SecA2 K129R-His.

Another possible explanation for the lack of a detectable interaction between SecA2 and SecY is that, while SecA2 and SecY participate in the same SecA2-dependent export pathway, these two proteins may not directly interact with one another. Rather, other components of the canonical Sec pathway, such as SecA1, may mediate the genetic relationship between SecA2 and SecY, interacting with both proteins during different steps of protein export. This type of intermediary function could also be performed by other unknown SecA2-specific components of the protein export machinery. Through the course of our suppressor analysis we also identified Msmeg_1684, a protein that appears to play a role in the process of SecA2-dependent protein export (Chapter 3). The function of Msmeg_1684 is currently unknown. It will be interesting in the future to assess whether this protein interacts physically with either SecA2 or SecY.

In conclusion, we used *in vivo* protein crosslinking to assess whether SecY interacts with SecA2-His or SecA2 K129R-His. No physical interaction was detected between these proteins. However, this does not contradict our findings that SecA2 and SecY interact genetically (Chapter 2). Rather, the nature of the interactions between SecA2 and the housekeeping Sec pathway, including SecY, are not immediately evident and will require additional study to understand fully.

Table 4.1. *Mycobacterium smegmatis* strains used in this study.

***M. smegmatis* strain**	**Description**	**Source**
NR116	Δ*secA2*	(22)
LL133	NR116+pHSG93, Δ*secA2*+*secA2-6xHis* plasmid	This work
LL143	NR116+pNR54, Δ*secA2*+Atc-inducible *secA2 K129R-6xHis* plasmid	This work
LL159	LL143+pLL42, Δ*secA2*+Atc-inducible *secA2 K129R-6xHis* plasmid+24S *secY* plasmid	This work

Table 4.2. Plasmids used in this study.

Plasmid	Genotype	Description	Source
pMV306.kan	*aph int attP$_{L5}$* ColE1	Single-copy mycobacterial shuttle vector, integrates in mycobacteriophage L5 *attB* site	(27)
pMV361.kan	*aph* P$_{hsp60}$ *int attP$_{L5}$* ColE1	Single-copy mycobacterial shuttle vector with *hsp60* promoter, integrates in mycobacteriophage L5 *attB* site	(27)
pHSG93	*aph* P$_{hsp60}$*-secA2-6xHis int attP$_{L5}$* ColE1	*M. smegmatis secA2-6xHis* in pMV361.kan	(25)
pNR54	*aph oriM* ColE1 P$_{myc1}$*-tetO-secA2 K129R-6xHis* P$_{smyc}$*-tetR*	Atc-inducible *secA2 K129R-6xHis* plasmid	(22)
pCR2.1-TOPO	*aph bla* ColE1	TOPO TA cloning plasmid	Invitrogen
pLL17	*aph bla secY* ColE1	*M. smegmatis secY* in pCR2.1-TOPO	This work (Chapter 2)
pLL41	*aph bla secY(24S)* ColE1	NR151 (24S) *secY* in pCR2.1-TOPO	This work
pLL2	*hyg bla int attP$_{Tweety}$* ColE1	pYUB2063 with 3299 bp PciI fragment removed, single-copy mycobacterial shuttle vector, integrates in mycobacteriophage Tweety *attB* site	This work (Chapter 2)
pLL42	*hyg bla secY(24S) int attP$_{Tweety}$* ColE1	NR151 (24S) *secY* in pLL2	This work

Table 4.3. Primers used in this study.

Primer	Sequence and Description
NR151sdmF	5'- GGCCGACGCAAGGGCTCCTGTTACTCTGGTAGACTGTTTGAC -3' binds 161 bp upstream of *secY* (*msmeg_1483*), used for site-directed mutagenesis to recreate 24S (NR151) *secY* promoter mutation
NR151sdmR	5'- CTCAGCCGTCAAACAGTCTACCAGAGTAACAGGAGCCCTTGC -3' binds 113 bp upstream of *secY* (*msmeg_1483*), used for site-directed mutagenesis to recreate 24S (NR151) *secY* promoter mutation

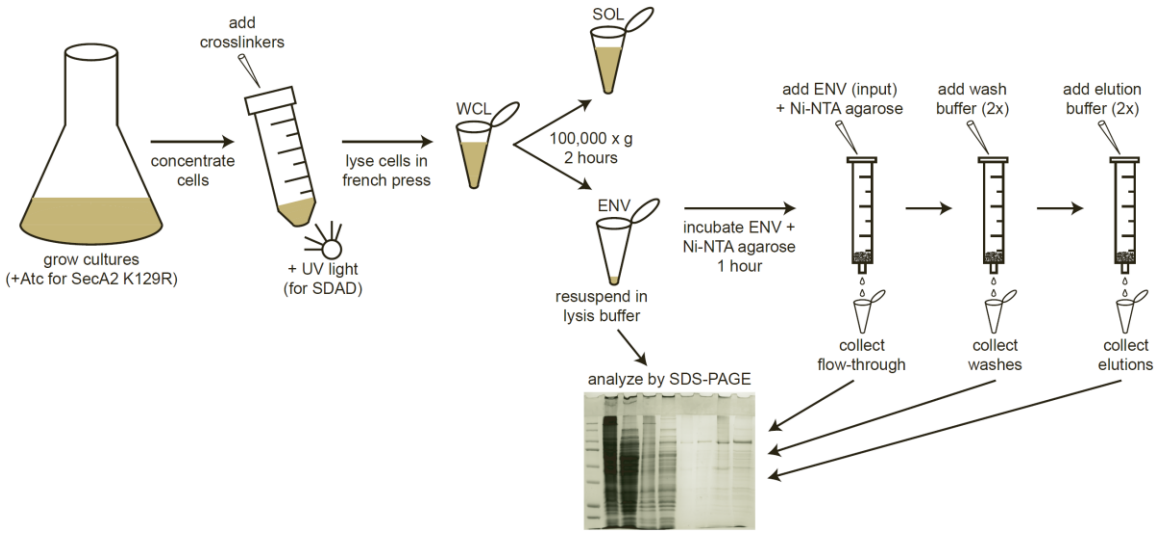

Figure 4.1. Schematic representation of experimental design for protein crosslinking and purification. Briefly, cultures were grown containing either SecA2-His (strain LL133) or Atc-induced SecA2 K129R-His (strain LL159). Cells were concentrated and crosslinked at room temperature. For samples containing the crosslinker SDAD, cells were also exposed to UV light to complete crosslinking. Cells were lysed by passage through a French press, creating whole cell lysates (WCL), which were then fractionated to cell envelope-containing (ENV, pellet) and cytoplasm-containing (SOL, supernatant) fractions by ultracentrifugation. The ENV fraction was resuspended in Qiagen lysis buffer and incubated with Ni-NTA agarose to allow binding of His-tagged proteins. ENV/agarose slurry was loaded into a column and the flow-through was collected. Columns were washed twice with Qiagen wash buffer and His-tagged proteins were eluted twice with Qiagen elution buffer. All samples were collected during column purification steps and were subsequently analyzed by SDS-PAGE and Western blotting. See Materials and Methods section for additional details.

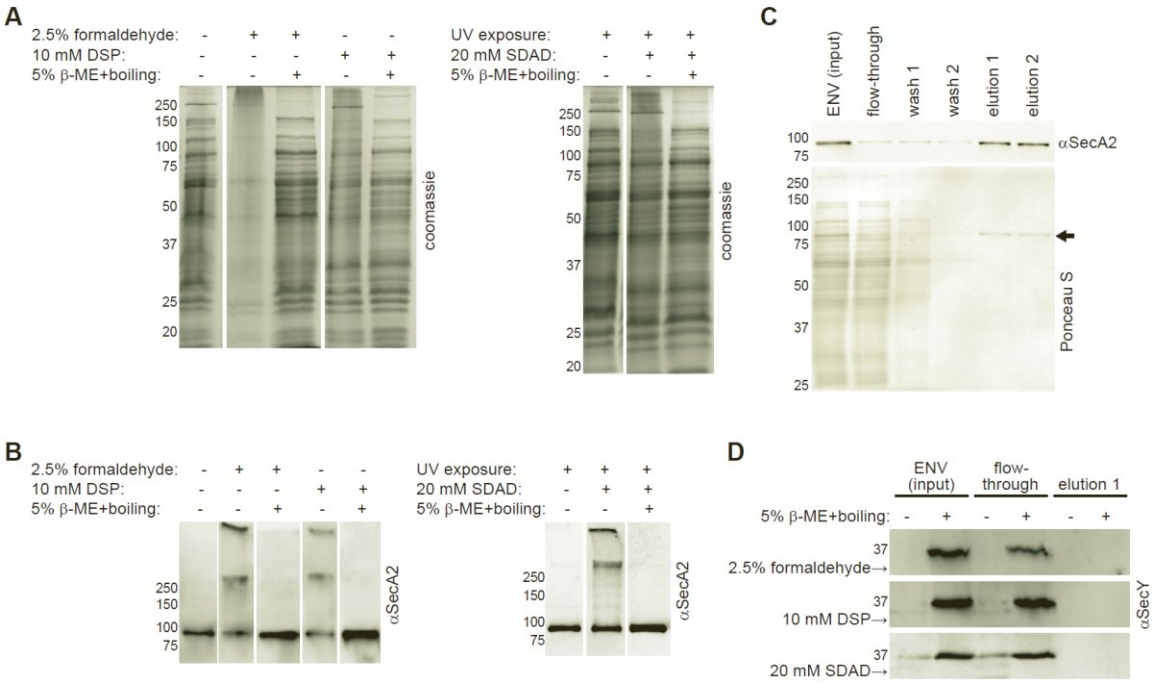

Figure 4.2. SecY does not crosslink to SecA2-His. Whole cells of strain Δ*secA2*+*secA2-His* (LL133) were crosslinked with the indicated chemicals and SecA2-His was purified from cell envelope (ENV) fractions. Samples were either mixed with SDS-PAGE sample buffer lacking β-ME at room temperature (-) to maintain protein crosslinks, or mixed with SDS-PAGE sample buffer containing 5% β-ME and boiled for 15 min. (+) to reverse protein crosslinks. Approximate molecular weights (in kDa) are indicated. All images shown are representative of at least 2 experiments. **(A)** ENV fractions separated by SDS-PAGE and all proteins detected with coomassie stain. **(B)** ENV fractions separated by SDS-PAGE and SecA2 detected by Western blotting. **(C)** SecA2-His purification samples separated by SDS-PAGE and transferred to a membrane, all proteins detected with Ponceau S stain (bottom), then SecA2 detected by Western blotting (top). Samples shown were not crosslinked and were treated with 5% β-ME + 15 min. boiling. Arrow indicates location of SecA2 band on Ponceau S-stained membrane. **(D)** Selected SecA2-His purification samples separated by SDS-PAGE and SecY detected by Western blotting.

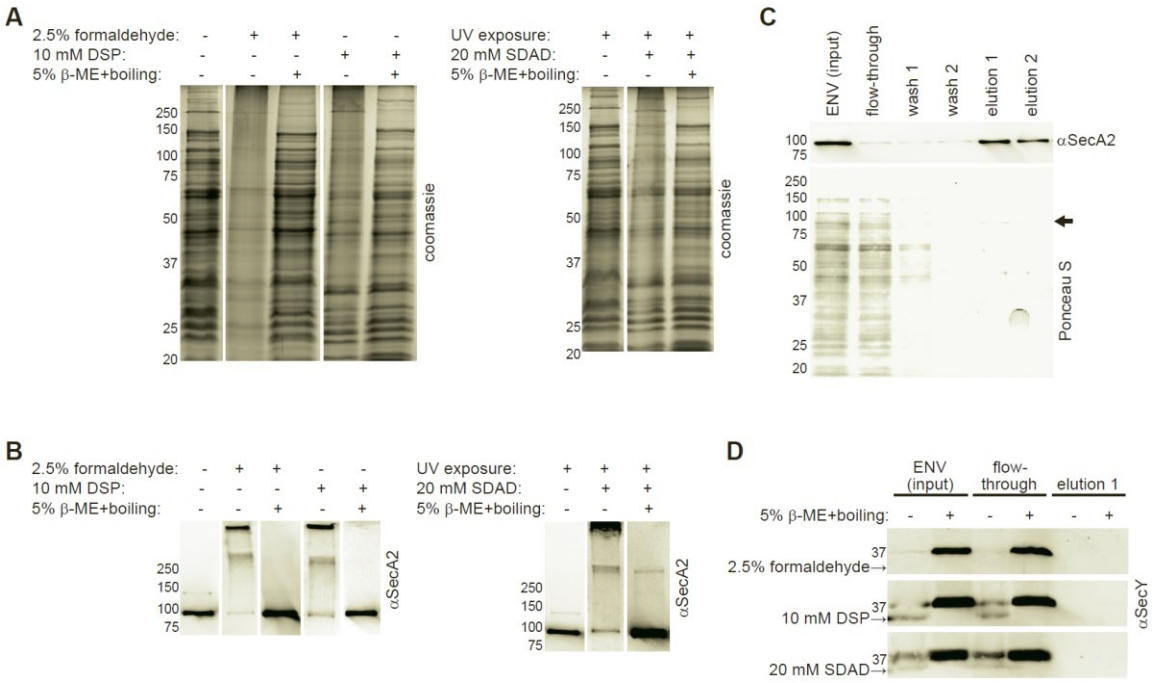

Figure 4.3. SecY does not crosslink to SecA2 K129R-His. Strain Δ*secA2*+Atc-inducible *secA2 K129R-His*+24S *secY* (LL159) was treated with 400 ng/mL Atc for four hours to induce production of SecA2 K129R-His. Whole cells were then crosslinked with the indicated chemicals and SecA2 K129R-His was purified from cell envelope (ENV) fractions. Samples were either mixed with SDS-PAGE sample buffer lacking β-ME at room temperature (-) to maintain protein crosslinks, or mixed with SDS-PAGE sample buffer containing 5% β-ME and boiled for 15 min. (+) to reverse protein crosslinks. Approximate molecular weights (in kDa) are indicated. All images shown are representative of at least 2 experiments. **(A)** ENV fractions separated by SDS-PAGE and all proteins detected with coomassie stain. **(B)** ENV fractions separated by SDS-PAGE and SecA2 detected by Western blotting. **(C)** SecA2-His purification samples separated by SDS-PAGE and transferred to a membrane, all proteins detected with Ponceau S stain (bottom), then SecA2 detected by Western blotting (top). Samples shown were not crosslinked and were treated with 5% β-ME + 15 min. boiling. Arrow indicates location of SecA2 band on Ponceau S-stained membrane. **(D)** Selected SecA2 K129R-His purification samples separated by SDS-PAGE and SecY detected by Western blotting.

References

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

2. **Flower AM, Hines LL, Pfennig PL.** 2000. SecG is an auxiliary component of the protein export apparatus of *Escherichia coli*. Mol Gen Genet **263:**131-136.

3. **Schiebel E, Driessen AJ, Hartl FU, Wickner W.** 1991. Delta mu H+ and ATP function at different steps of the catalytic cycle of preprotein translocase. Cell **64:**927-939.

4. **van der Wolk JP, de Wit JG, Driessen AJ.** 1997. The catalytic cycle of the *escherichia coli* SecA ATPase comprises two distinct preprotein translocation events. EMBO J **16:**7297-7304.

5. **Snyders S, Ramamurthy V, Oliver D.** 1997. Identification of a region of interaction between Escherichia coli SecA and SecY proteins. J Biol Chem **272:**11302-11306.

6. **Zimmer J, Nam Y, Rapoport TA.** 2008. Structure of a complex of the ATPase SecA and the protein-translocation channel. Nature **455:**936-943.

7. **de Keyzer J, van der Does C, Kloosterman TG, Driessen AJ.** 2003. Direct demonstration of ATP-dependent release of SecA from a translocating preprotein by surface plasmon resonance. J Biol Chem **278:**29581-29586.

8. **Manting EH, van der Does C, Driessen AJ.** 1997. In vivo cross-linking of the SecA and SecY subunits of the Escherichia coli preprotein translocase. J Bacteriol **179:**5699-5704.

9. **Duong F.** 2003. Binding, activation and dissociation of the dimeric SecA ATPase at the dimeric SecYEG translocase. EMBO J **22:**4375-4384.

10. **Natale P, Swaving J, van der Does C, de Keyzer J, Driessen AJ.** 2004. Binding of SecA to the SecYEG complex accelerates the rate of nucleotide exchange on SecA. J Biol Chem **279:**13769-13777.

11. **Bensing BA, Sullam PM.** 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.

12. **Braunstein M, Brown AM, Kurtz S, Jacobs WR, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. J Bacteriol **183:**6979-6990.

13. **Caspers M, Freudl R.** 2008. *Corynebacterium glutamicum* possesses two *secA* homologous genes that are essential for viability. Archives of microbiology **189:**605-610.

14. **Fagan RP, Fairweather NF.** 2011. *Clostridium difficile* has two parallel and essential Sec secretion systems. J. Biol. Chem. **286:**27483-27493.

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

16. **Nguyen-Mau SM, Oh SY, Kern VJ, Missiakas DM, Schneewind O.** 2012. Secretion genes as determinants of *Bacillus anthracis* chain length. J. Bacteriol. **194:**3841-3850.

17. **Siboo IR, Chaffin DO, Rubens CE, Sullam PM.** 2008. Characterization of the accessory Sec system of *Staphylococcus aureus*. J Bacteriol **190:**6188-6196.

18. **Rigel NW, Braunstein M.** 2008. A new twist on an old pathway--accessory Sec [corrected] systems. Mol Microbiol **69:**291-302.

19. **Braunstein M, Espinosa BJ, Chan J, Belisle JT, Jacobs WR, Jr.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of Mycobacterium tuberculosis. Mol Microbiol **48:**453-464.

20. **Hou JM, D'Lima NG, Rigel NW, Gibbons HS, McCann JR, Braunstein M, Teschke CM.** 2008. ATPase activity of *Mycobacterium tuberculosis* SecA1 and SecA2 proteins and its importance for SecA2 function in macrophages. J Bacteriol **190:**4880-4887.

21. **Bensing BA, Sullam PM.** 2009. Characterization of Streptococcus gordonii SecA2 as a paralogue of SecA. J Bacteriol **191:**3482-3491.

22. **Rigel NW, Gibbons HS, McCann JR, McDonough JA, Kurtz S, Braunstein M.** 2009. The Accessory SecA2 System of Mycobacteria Requires ATP Binding and the Canonical SecA1. J Biol Chem **284:**9927-9936.

23. **Snapper SB, Lugosi L, Jekkel A, Melton RE, Kieser T, Bloom BR, Jacobs WR, Jr.** 1988. Lysogeny and transformation in mycobacteria: stable expression of foreign genes. Proc Natl Acad Sci U S A **85:**6987-6991.

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

25. **Rigel NW.** 2008. Characterization of the SecA2 Protein Export Pathway of Mycobacteria. University of North Carolina at Chapel Hill, Chapel Hill, NC.

26. **Pham TT, Jacobs-Sera D, Pedulla ML, Hendrix RW, Hatfull GF.** 2007. Comparative genomic analysis of mycobacteriophage Tweety: evolutionary insights and construction of compatible site-specific integration vectors for mycobacteria. Microbiology **153:**2711-2723.

27. **Stover CK, de la Cruz VF, Fuerst TR, Burlein JE, Benson LA, Bennett LT, Bansal GP, Young JF, Lee MH, Hatfull GF, Snapper SB, Barletta RG, Jacobs WR, Jr., Bloom BR.** 1991. New use of BCG for recombinant vaccines. Nature **351:**456-460.

CHAPTER 5

Discussion

As the causative agent of the disease tuberculosis, the bacterium *Mycobacterium tuberculosis* is responsible for 1.4 million deaths annually and is the primary cause of death for HIV-positive individuals (1). In addition, multiply and extensively drug-resistant strains of *M. tuberculosis* (MDR and XDR-TB) pose a serious challenge for future tuberculosis treatment. When *M. tuberculosis* is spread from person to person by aerosol and is phagocytosed by alveolar macrophages, the proteins exported by the bacterium are ideally positioned to interact with the host. The pool of *M. tuberculosis* exported proteins, in addition to performing functions critical for *M. tuberculosis* growth and survival, are able to modulate host immune responses, allowing the bacterium to survive within the host and cause disease.

M. tuberculosis and other mycobacteria, including the model organism *Mycobacterium smegmatis*, have several systems dedicated to the export of proteins to the bacterial cell surface and the extracellular environment (2). Like the proteins they export, these systems are often required for the survival and/or virulence of the bacterium. The conserved Sec export system is responsible for the bulk of protein export and is essential for cell survival (3). In this system, unfolded proteins with N-terminal signal peptides are post-translationally exported through a membrane-embedded channel composed of SecY, SecE, and SecG (4, 5). The SecA motor protein recognizes the proteins destined for Sec export (6, 7) and interacts with the SecYEG

translocon (8, 9), harvesting the energy from repeated cycles of ATP binding and hydrolysis to push proteins through the SecYEG channel and across the membrane (10, 11).

Mycobacteria, along with some Gram-positive bacteria, are unique in containing two functionally distinct copies of the SecA motor protein. The SecA protein responsible for housekeeping Sec export is referred to as SecA1, while the second SecA protein is referred to as SecA2, is generally non-essential, and is responsible for a subset of protein export (12). When found in pathogens such as *M. tuberculosis*, the SecA2 export system is often important for virulence. In some bacteria, the SecA2 protein is accompanied by a SecY2 protein that is also required for SecA2-dependent protein export (12). These systems are referred to as "SecA2-SecY2" systems. Mycobacteria do not contain a SecY2 protein or any other obvious SecA2 system components and the mycobacterial SecA2 export pathway is therefore referred to as a "SecA2-only" system. Unlike the well-studied, conserved Sec export pathway, the mechanism of mycobacterial SecA2-dependent export is poorly understood.

This dissertation describes experiments that advance our understanding of the *M. smegmatis* SecA2 export pathway. Prior to this dissertation work, we knew that SecA2 had ATPase activity and that this activity was important for its function (13, 14). We also knew that SecA2-dependent export could not proceed in the absence of the housekeeping SecA1 protein (14). However, we did not know whether SecA2 worked with other housekeeping Sec components or if any additional factors were involved in the SecA2 export pathway. Here, we demonstrated for the first time a connection between *M. smegmatis* SecA2 and SecY, the major component of the housekeeping Sec export channel. In addition, we demonstrated a connection between the SecA2 export pathway and Msmeg_1684, a protein of unknown function that is conserved throughout mycobacteria and may represent a novel component of

the SecA2 export system. The findings described in this dissertation suggest a new model in which the mycobacterial SecA2 export pathway is integrated into the housekeeping Sec pathway and includes an additional SecA2-specific component, Msmeg_1684.

SecA2 K129R Causes SecY Degradation

In Chapter 2, we identified a relationship between SecA2 and SecY by studying *secA2 K129R*, an allele encoding a substitution in the Walker box of SecA2. This Walker box substitution prevents binding and hydrolysis of ATP, rendering SecA2 K129R non-functional (13, 14). In addition, the Walker box substitution produces a dominant negative SecA2 protein (14). Dominant negative proteins exhibit phenotypes even in the presence of a wild type copy of the protein, and are often still able to interact with their normal binding partners. Consistent with the idea that it is interacting with other proteins, SecA2 K129R exhibits a shift in localization from the cytoplasm towards the cell envelope (14), implying that it is trapped in a complex with a membrane protein.

During protein export, the canonical SecA protein repeatedly associates with and releases from the SecYEG translocon, pushing exported proteins through the SecYEG channel in a stepwise fashion (10, 11). ATP hydrolysis is necessary for SecA to release from the SecYEG translocon (10). An *E. coli* SecA variant containing a Walker box substitution (15) comparable to that in *M. smegmatis* SecA2 K129R is therefore unable to release from and becomes trapped at the cell membrane (16). If SecA2 interacts with the canonical Sec pathway in a similar way as *E. coli* SecA does, we would expect SecA2 K129R to become trapped at the translocon while attempting to export a protein. As demonstrated in Chapter 2, the presence of SecA2 K129R results in a drastic reduction in SecY protein levels. This effect on SecY makes sense in light of

previous studies performed in *Escherichia coli*. "Jamming" of Sec protein export in *E. coli* by attempted export of a folded protein results in degradation of SecY by the protease FtsH (17). Therefore, SecY degradation in the *M. smegmatis secA2 K129R* mutant might occur in an effort by the cell to remove SecY channels rendered non-functional by trapped SecA2 K129R proteins.

In addition to being a dominant negative protein, SecA2 K129R causes severe growth inhibition in *M. smegmatis*. The observed connection between SecA2 K129R and SecY explains this growth defect. Because SecY is an essential protein critical for cell growth and viability (18, 19), the severe decrease in SecY protein levels in the presence of SecA2 K129R would be detrimental to the cell. SecA1 depletion could also be stressful on the Sec export system and could potentially result in SecY degradation, as seen for SecA2 K129R. Therefore, we also re-evaluated the previously observed (14) requirement for SecA1 in SecA2-dependent export. Our results showed that SecA1 depletion does *not* affect SecY protein levels, and therefore that SecY degradation does not account for the apparent role of SecA1 in the SecA2 export system.

Suppressor Mutations Increase SecY Protein Levels

Given the degradation of SecY observed in the presence of SecA2 K129R, it is logical that suppressor mutations might alleviate the phenotypes of the *secA2 K129R* strain by either avoiding SecY degradation or restoring SecY protein levels. Consistent with this hypothesis, we found two extragenic suppressors containing mutations affecting *secY* expression, strains 4S and 24S. Each of these suppressors contains a mutation upstream of *secY*. In Chapter 2, we demonstrated that these mutations are located within the *secY* promoter and cause an increase in *secY* transcript levels. At the protein level, suppressors 4S and 24S exhibit increased SecY levels relative to the *secA2 K129R* strain, but do not fully restore SecY to the level seen in the

presence of the wild type *secA2* allele. This increase in *secY* expression is sufficient to suppress the *secA2 K129R* allele. These findings demonstrate that suppressors 4S and 24S do not prevent SecY degradation, but rather act to partially restore SecY protein levels, balancing SecY degradation with increased *secY* expression.

In addition to the two suppressors that directly affect *secY* expression, we examined six additional extragenic suppressors. All six suppressors exhibited increased levels of SecY protein relative to the *secA2 K129R* strain. However, these six suppressors do not contain mutations in or near *secY* and do not affect *secY* transcript levels. The suppressor mutations in these strains might affect SecY production in some other way; alternately, these mutations might act to avoid SecY degradation in the first place. For instance, these suppressor mutations may act by preventing the non-functional interaction of SecA2 K129R with components of the housekeeping Sec export pathway, preventing the stress on Sec export that results in SecY degradation. This possibility is especially interesting, as suppressor mutations that act in this manner may point to additional proteins required for SecA2-dependent protein export. It is also interesting to note that when the eight total extragenic suppressor strains examined in Chapter 2 are taken as a group, the strength of the suppressor phenotype correlates with the level of SecY protein in each strain. This provides support for the conclusion that degradation of SecY is responsible for the severe phenotypes of the *secA2 K129R* mutant strain and that the level of suppression reflects the degree of restoration of SecY levels.

Msmeg_1684 is Genetically Connected to SecA2-Dependent Protein Export

In Chapter 3, we examined additional extragenic suppressors of the dominant negative *secA2 K129R* allele. Two suppressors, strains 7S and 20B, contain very large deletions (43-51

kbp) in the same region of the genome. Among the many genes located within these deletion regions, we ruled out all but nine, leaving genes *msmeg_1684-1692* as candidates for explaining the suppression phenotype. Whole genome sequencing of six additional suppressors revealed mutations affecting gene *msmeg_1684* in all six strains. Because *msmeg_1684* is also deleted in suppressors 7S and 20B, we actually collected a total of 8 suppressor strains that appear to act through *msmeg_1684* mutations. To confirm the role of *msmeg_1684* in the suppressor phenotype, we evaluated a strain containing an unmarked, in-frame deletion of *msmeg_1684*. Like the eight suppressor strains, the *msmeg_1684* deletion strain is also able to suppress the *secA2 K129R* allele.

Our findings in Chapter 3 demonstrate that the presence of Msmeg_1684 contributes to the SecA2 K129R phenotype and that the loss of Msmeg_1684 creates a suppressor phenotype. This implies that Msmeg_1684 is connected in some way to the SecA2 protein export pathway. The role of Msmeg_1684 in SecA2-dependent export is yet to be determined; however, there are several interesting clues to its function that are worthy of discussion. Foremost, Msmeg_1684 appears to be required for the interaction of SecA2 K129R with the housekeeping Sec pathway. In support of this conclusion, the shift in localization of SecA2 K129R to the cell envelope is completely reversed in suppressors 7S and 20B. These two suppressors, which contain large deletions removing gene *msmeg_1684*, exhibit SecA2 K129R localization to the cytoplasm, resembling that of wild type SecA2. Consistent with the localization change, these suppressors are also able to completely reverse the SecA2 K129R azide sensitivity and growth on Mueller-Hinton phenotypes, exhibiting phenotypes similar to the *secA2* deletion strain. In contrast, suppressors 4S and 24S, which act by increasing *secY* expression, are only able to partially reverse the effects of the *secA2 K129R* allele. The complete reversal of SecA2 K129R localization and phenotypes in suppressors 7S and 20B demonstrates that Msmeg_1684 must be

required for SecA2 K129R to become trapped in complexes at the membrane. In suppressor strains lacking Msmeg_1684, the detrimental interaction of SecA2 K129R with the Sec pathway is completely prevented, avoiding the resulting stress on export and degradation of SecY entirely. In the future, the requirement for Msmeg_1684 in the interaction of SecA2 K129R with the cell membrane should be confirmed in the in-frame *msmeg_1684* deletion mutant.

An additional point supporting the observed connection of Msmeg_1684 to the SecA2 export pathway is found in the pattern of occurrence of Msmeg_1684-like proteins in other species. Proteins with strong homology to Msmeg_1684, as well as SecA2 proteins, are found throughout mycobacteria. In addition, Msmeg_1684-like proteins are found in about half of the approximately 370 currently sequenced species within the order Actinomycetales, of which mycobacteria are members. Similarly, SecA2 proteins are also found in about half of the sequenced Actinomycetales species. Strikingly, the species containing Msmeg_1684 proteins and the species containing SecA2 proteins are largely overlapping. We identified only 16 species containing Msmeg_1684-like proteins without SecA2 proteins, and only 6 species containing SecA2 proteins without Msmeg_1684-like proteins. However, bacteria outside of the Actinomycetales that contain SecA2 systems, including those of both the “SecA2-only” and “SecA2-SecY2” types, are not associated with Msmeg_1684-like proteins. The frequent co-occurrence of SecA2 and Msmeg_1684-like proteins among the Actinomycetales supports the conclusion that these two proteins operate as part of the same pathway and suggests that the presence of an Msmeg_1684-like protein is a specific adaptation of the mycobacterial and related SecA2 systems.

Msmeg_1684 is 431 amino acids in size and contains no predicted signal peptide or transmembrane domains, supporting a predicted cytoplasmic location. Despite the fact that

Msmeg_1684 is conserved with similar proteins found throughout the order Actinomycetales, it has no useful similarity to proteins with known functions. Therefore, the function of Msmeg_1684 cannot be predicted. However, it is interesting to note that Msmeg_1684 contains many aspartate and glutamate residues and is therefore a highly acidic protein (isoelectric point is 3.83). Protein export chaperones are often acidic proteins, such as the SecB chaperone involved in housekeeping Sec export in Gram-negative bacteria (20) (isoelectric point is 4.05). In addition, numerous acidic chaperones are involved in type three secretion (21). Thus, one possibility to consider is that Msmeg_1684 might perform a chaperone role in SecA2-dependent protein export. Export chaperones can have functions both in maintaining proteins in an unfolded state and in delivering them to the export apparatus, thus it will be interesting to evaluate Msmeg_1684 for these roles in the future.

The *M. tuberculosis* homolog of Msmeg_1684, Rv3311, is predicted by TraSH to be important for growth in macrophages (22), indicating that Rv3311 likely plays a role in *M. tuberculosis* virulence. Rv3311 was not tested in the TraSH screen performed in mice (23), therefore no prediction exists for this phenotype. Similarly, the *M. tuberculosis* SecA2 protein is also predicted by TraSH to be important for growth in macrophages (22), as well as in mice (23). These SecA2 phenotypes have also been demonstrated directly (24, 25). If Rv3311 and SecA2 participate in the same pathway, then mutant strains of each would be expected to display similar phenotypes. Therefore, it will be interesting to assess an *rv3311* deletion strain for phenotypes exhibited by the *secA2* deletion strain, such as failure to grow in macrophages(25), reduced early growth in mice (24), smooth colony morphology on agar plates (14), and failure to block phagosome maturation (26). The *rv3311* gene is also predicted to be co-transcribed with *sapM* (*rv3310*). Like SecA2, SapM is thought to be important for the ability of *M. tuberculosis* to

resist phagosome maturation (26, 27). This close association of *rv3311* with *sapM* provides yet another indirect clue connecting Msmeg_1684 to the SecA2 export system.

A Physical Interaction Between SecA2 and SecY Cannot be Detected

In Chapter 4, we assessed the possibility of a physical interaction between the *M. smegmatis* SecA2 and SecY proteins. The results presented in Chapter 2 demonstrated a connection between SecA2 and SecY and suggested a model in which SecA2 would interact physically with SecY during protein export, pushing exported proteins through the canonical SecYEG channel. To assess this model, we performed *in vivo* protein crosslinking using a variety of chemical crosslinkers. One of the crosslinkers used, formaldehyde, has been successfully used to crosslink SecA to SecY *in vivo* in *E. coli* (28). Following chemical crosslinking, we purified His-tagged SecA2 and SecA2 K129R proteins from the cell envelope using a nickel-affinity column and evaluated the resulting eluates for co-purification of SecY. In no case did we identify SecY in association with either SecA2 or SecA2 K129R. This finding suggests that the observed genetic relationship between *secA2* and *secY* does not represent a direct physical interaction between the SecA2 and SecY proteins.

There are several possible explanations for this finding. First, while the crosslinking conditions used were carefully chosen to maximize our chances of success, it is possible that the crosslinking conditions used for this experiment were simply not capable of trapping a SecA2-SecY interaction. Second, it is possible that a physical interaction between SecA2 and SecY does occur, but is so weak or transient that it could not be detected. This possibility is supported by the fact that SecA2 is normally found primarily in the cytoplasm (14), suggesting that it spends

less time interacting with membrane proteins than SecA1, which is found in both the cytoplasm and the cell envelope.

It is also possible that our failure to detect an interaction between SecA2 and SecY is indicative of the true situation within the cell. Perhaps, unlike the physical interaction between the canonical SecA and SecY proteins, SecA2 never engages in a direct physical interaction with SecY. Rather, the genetic relationship between *secA2* and *secY* may represent an indirect interaction mediated through other proteins involved in SecA2-dependent export. This is an interesting possibility as it would point to a function for SecA2 that is very different than the role of SecA1 as a motor protein. Instead of physically pushing proteins through the SecYEG channel, perhaps SecA2 delivers proteins to SecA1, which then performs the final export steps. As shown previously (14) and confirmed here, SecA1 does play a role in SecA2-dependent protein export. Alternately, the connection between SecA2 and SecY could be mediated by some other protein involved in SecA2-dependent export, such as Msmeg_1684, the protein of unknown function we identified in Chapter 3. The exact relationships between SecA2, SecY, SecA1, and Msmeg_1684 will require additional study in the future to fully understand how these proteins interact during SecA2-dependent protein export.

Suppressor Analysis Through Genome Sequencing is a Successful Approach for Exploring SecA2-Dependent Export

Extragenic suppressor analysis is a classical genetic method that can be used to identify unknown genes involved in various bacterial pathways, and has been used extensively in studying the canonical Sec system in *E. coli*. However, this approach has only rarely been used

in mycobacteria (29, 30). Here, we described the analysis of 10 extragenic suppressors of the *M. smegmatis secA2 K129R* allele. This approach was successful in identifying two genes, *secY* and *msmeg_1684*, that are genetically connected to SecA2-dependent export.

Our original suppressor collection included 63 independent suppressor strains capable of improved growth on Mueller-Hinton agar relative to the parent *secA2 K129R* mutant strain (as described in Chapter 2 and (31)). Of these 63 suppressors, 23 expressed SecA2 K129R protein. The remaining 40 suppressors lost expression of the *secA2 K129R* allele, suggesting they act either through intragenic mutations within the *secA2 K129R* allele, through mutations affecting the *hsp60* promoter from which *secA2 K129R* was expressed, or through mutations affecting genes that regulate P_{hsp60}-*secA2 K129R* expression or SecA2 K129R protein levels. While these 40 suppressors could potentially inform on SecA2 regulation, they are unlikely to provide mechanistic information on SecA2-dependent protein export and were not studied further. Of the 23 suppressors still expressing SecA2 K129R protein, 9 were found to have intragenic mutations within the *secA2 K129R* allele (31). These intragenic mutations could provide information on specific regions of the SecA2 protein that are important for its function and would be interesting to study in the future, particularly in the context of exploring physical interactions between SecA2 and other proteins involved in SecA2-dependent export. The remaining 14 suppressors expressing SecA2 K129R protein do not have mutations in the *secA2 K129R* allele and are therefore extragenic suppressor strains (see Appendix for detailed information on extragenic suppressors). This category of suppressor mutants is where we focused our study for this dissertation, in an effort to identify additional components of the SecA2-dependent protein export system.

Among the 14 extragenic suppressors of *secA2 K129R*, we found two strains to have mutations upstream of *secY*, as discussed in Chapter 2. The remaining 12 extragenic suppressors do not contain mutations affecting *secY* or its promoter. We found 8 of the remaining 12 extragenic suppressors of *secA2 K129R* to contain mutations affecting gene *msmeg_1684*, as discussed in Chapter 3. Mutations in *msmeg_1684* are not found in the other 6 of 14 extragenic suppressors of *secA2 K129R* (Appendix).

Therefore, of the 14 extragenic suppressors of *secA2 K129R* the mutations responsible for suppression in 4 strains have yet to be identified. These strains do not carry mutations in either the *secY* promoter or *msmeg_1684* (Appendix). Therefore, future identification of the suppressor mutations in these strains may provide yet additional information about the SecA2 pathway. Notably, two of these suppressors, strains 5S and 13B, do express SecA2 K129R protein, but do so at a reduced level relative to the parent *secA2 K129R* strain. It is possible that this reduction in SecA2 K129R levels is responsible for the suppressor phenotype of these two strains.

In addition to the suppressors of *M. smegmatis secA2 K129R* examined here, a separate group of 7 extragenic suppressors was isolated from an *M. smegmatis* strain expressing the *M. tuberculosis secA2 K115R* allele (comparable to *M. smegmatis K129R*) (31). Though these 7 suppressor strains were not studied in depth in this dissertation work, they were screened for mutations affecting the *secY* promoter. No mutations affecting *secY* were found among this group of extragenic suppressors (Appendix). However, one suppressor from this group has a large chromosomal deletion affecting *msmeg_1684*. The remaining 6 suppressors of *M. tuberculosis secA2 K115R* have not been evaluated for *msmeg_1684* mutations, an experiment which should be performed in the future.

The recent advent of whole genome sequencing technology has significantly contributed to the success of the suppressor analysis strategy described in this dissertation. Previous methods used to locate mutations responsible for suppression, such as the screening of genomic DNA librairies, had significant drawbacks. For example, a dominant mutation, such as those affecting *secY*, could not be identified in the same way as a recessive mutation, such as those affecting *msmeg_1684*. Therefore, if the wrong approach was used, the suppressor mutation could not be identified. In comparison, we are now able to quickly sequence and analyze whole genomes, pinpointing hundreds of mutations of all types in a single genome in a matter of weeks. However, this approach comes with its own challenges. Initially, genome analysis and identification of mutations was a laborious process requiring us to recruit the assistance of collaborators capable of constructing custom computer algorithms tailored to the project at hand. In addition, the first genomes we sequenced required an entire lane of Illumina sequencing for each genome and provided average sequence coverage as low as 16. Even after mutations are successfully identified, the identification of multiple mutations in a single strain presents the challenge of determining which mutation(s) is responsible for the suppressor phenotype. During the course of this dissertation work, genome sequencing and analysis has improved drastically, to the point where we can now sequence 6 genomes (or more) in a single lane of Illumina sequencing, resulting in average coverage of around 200 for each genome. In addition, user-friendly computer software, such as the Geneious program used here, now allows quick automated data analysis and mutation identification without the assistance of computer programmers.

Previously, the idea of performing a suppressor analysis in *M. tuberculosis* would have been daunting. The laborious process required to locate a suppressor mutation combined with the slow growth and difficulty of working with *M. tuberculosis* would likely have prohibited such

an effort. However, the use of whole genome sequencing now makes the suppressor analysis strategy described here feasible in *M. tuberculosis*. In the future, it would be interesting to assess whether suppressors of *M. tuberculosis secA2* mutants occur in genes similar to those identified in *M. smegmatis* suppressor mutants. In addition, this strategy could also be used to explore novel molecular pathways in *M. tuberculosis*, providing an unbiased method of identifying unknown genes related to a gene of interest.

A Model of the SecA2-Dependent Protein Export Pathway

Combining the findings of this dissertation with our previous knowledge of the mycobacterial SecA2 system, we can now propose an improved model for SecA2-dependent protein export (Fig. 5.1). We knew previously that depletion of SecA1 reduced export of SecA2-dependent proteins, reflecting a possible requirement for SecA1 in the SecA2 pathway (14). However, we did not know if the SecYEG channel was also used for SecA2-dependent export. We find here that SecY *is* also connected to the SecA2 system. In addition, we find that the observed requirement for SecA1 is not a consequence of SecY degradation induced upon SecA1 depletion. This result does not prove a direct role for SecA1 in exporting SecA2-dependent proteins, as it is still possible that SecA1 is required for export of another protein that is in turn involved in SecA2-dependent export. However, the connection of SecA2 to **both** the SecA1 and SecY components of the housekeeping Sec pathway strongly supports a model in which SecA2-dependent export is fully integrated into the Sec export system. Therefore, we propose that SecA2-dependent proteins are exported through the SecYEG channel with the direct assistance of SecA1.

In general, SecA proteins are known to physically engage the SecYEG translocon while pushing exported proteins across the membrane (8, 9, 28). However, we failed to detect a physical interaction between SecA2 and SecY. Therefore, we propose that SecA1 performs the canonical SecA function as the motor protein and interacts with SecYEG to drive export of SecA2-dependent proteins across the membrane. Meanwhile, rather than simply substituting for SecA1 in certain cases, SecA2 must perform a completely different mechanistic function from that of canonical SecA proteins. One possibility is that SecA2 is responsible for a targeting step, recognizing a specific subset of proteins that are not recognized by the housekeeping system and then delivering them to the SecA1-SecYEG translocase for subsequent export. While it is not mutually exclusive with the previous possibility, another possibility is that SecA2 plays a chaperone-like role, maintaining a subset of proteins in an export-competent state prior to or during export by the housekeeping Sec system. In fact, this potential chaperone function is consistent with another recent study from our laboratory (32). In this study, it was shown that the N-terminal signal peptides found on *M. smegmatis* SecA2-dependent proteins *are* required for their export, but do not determine the requirement for SecA2. Rather, it is the mature domain of these SecA2-dependent proteins that carries the SecA2-specificity feature. This feature that determines SecA2-dependence is proposed to be a propensity to fold in the cytoplasm prior to export. Perhaps SecA2 is required either to prevent folding or to provide additional energy necessary to unfold proteins during export. Also, while folding or some other specific feature of the mature domain targets these proteins for SecA2-dependent export, the simultaneous requirement for the N-terminal signal peptide in their export is consistent with our model. After being delivered and/or unfolded by SecA2, the N-terminal signal peptides of SecA2-dependent proteins would be important for normal interaction with both SecA1 and SecY, as in canonical Sec export.

Further, we propose that Msmeg_1684 also participates in SecA2-dependent protein export. While the connection we observed between Msmeg_1684 and the SecA2 system has not yet been proven to represent a direct role in export, our findings support this model. The 7S and 20B suppressors, in which *msmeg_1684* is deleted, completely reverse the association of SecA2 K129R with the membrane, returning it to a cytoplasmic localization resembling wild type SecA2. In addition, *msmeg_1684* suppressor mutations are able to completely reverse the phenotypes of the parent *secA2 K129R* strain. These findings suggest that *msmeg_1684*-type suppressor strains prevent the detrimental interaction of SecA2 K129R with the Sec pathway, preventing degradation of SecY from ever occurring. In the future, pulse chase experiments could be used to confirm the apparent effects of various suppressor mutations on SecY degradation rates.

Based on the above arguments, we propose that Msmeg_1684 must be present in order for SecA2 to interact with the housekeeping Sec system. This requirement for Msmeg_1684 might occur at a variety of points in the SecA2 export pathway. For example, one possibility is that Msmeg_1684 acts as an adaptor protein and is necessary to mediate the interaction between SecA2 and other Sec components such as SecA1 or SecY. In the case of a suppressor strain lacking Msmeg_1684, no interaction would occur between SecA2 and the housekeeping Sec pathway, thereby preventing any detrimental effects caused by the SecA2 K129R variant. Alternately, Msmeg_1684 might act at an earlier point in SecA2-dependent export. For example, Msmeg_1684 might act as a chaperone that maintains SecA2-dependent proteins in an export-competent state prior to their interaction with SecA2. In fact, Msmeg_1684 is a very acidic protein, similar to other known protein export chaperones. In this case, Msmeg_1684 would be required for SecA2 to interact with proteins destined for SecA2-dependent export. Therefore, in a suppressor strain lacking Msmeg_1684, SecA2 K129R would not be able to

interact with a SecA2-dependent exported protein and, as a result, might never proceed to interact with the housekeeping Sec pathway, preventing the detrimental effects of SecA2 K129R.

In conclusion, our findings as presented in this dissertation have significantly improved our understanding of the mechanism of SecA2-dependent protein export. Because the mycobacterial SecA2 system appears to be a specialized adaptation of the housekeeping Sec pathway, an improved understanding of SecA2-dependent export will also inform on currently unknown limitations of the Sec pathway that necessitate this adaptation. Therefore, continued study of this system will shed light not only on SecA2-dependent export, a pathway important for *M. tuberculosis* virulence, but also on the canonical Sec pathway utilized by all bacteria and essential for bacterial survival.

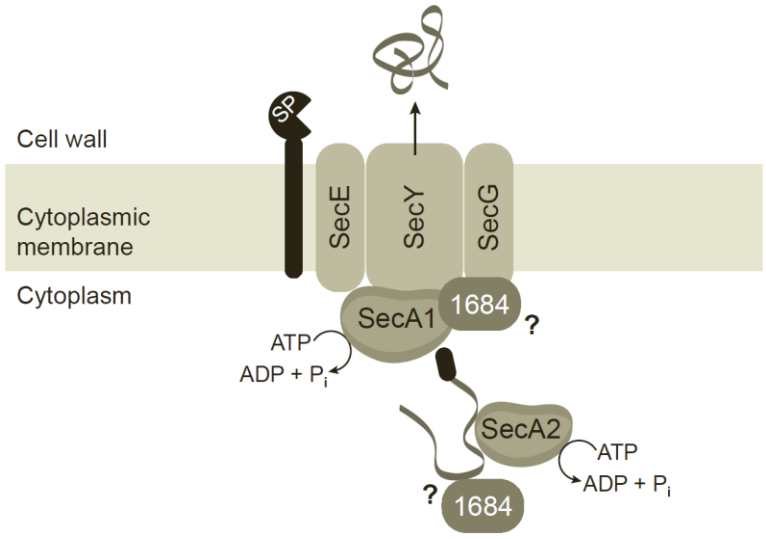

Figure 5.1. An improved model of the SecA2-dependent protein export system. SecA2 recognizes the mature domain of a small subset of proteins and uses its ATPase activity to assist in their export. Msmeg_1684 influences the SecA2 export pathway at an unknown step. SecA1 is also required for SecA2-dependent export and performs repeated cycles of ATP hydrolysis, pushing the unfolded preprotein (grey ribbon) through the SecYEG channel. Signal peptides (black oval) are removed by a LepB or LspA signal peptidase (SP), and mature proteins fold into their final conformations.

References

1. 2012. Global Tuberculosis Report 2012. World Health Organization.

2. **Ligon LS, Hayden JD, Braunstein M.** 2012. The ins and outs of *Mycobacterium tuberculosis* protein export. Tuberculosis (Edinb) **92:**121-132.

3. **Murphy CK, Beckwith J.** 1996. Export of Proteins to the Cell Envelope in *Escherichia coli*, p. 967-978. *In* Neidhardt FC (ed.), *Escherichia coli* and *Salmonella* : cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.

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

5. **Driessen AJ, Nouwen N.** 2008. Protein translocation across the bacterial cytoplasmic membrane. Annu Rev Biochem **77:**643-667.

6. **Cunningham K, Wickner W.** 1989. Specific recognition of the leader region of precursor proteins is required for the activation of translocation ATPase of *Escherichia coli*. Proc Natl Acad Sci U S A **86:**8630-8634.

7. **Lill R, Dowhan W, Wickner W.** 1990. The ATPase activity of SecA is regulated by acidic phospholipids, SecY, and the leader and mature domains of precursor proteins. Cell **60:**271-280.

8. **Snyders S, Ramamurthy V, Oliver D.** 1997. Identification of a region of interaction between Escherichia coli SecA and SecY proteins. J Biol Chem **272:**11302-11306.

9. **Zimmer J, Nam Y, Rapoport TA.** 2008. Structure of a complex of the ATPase SecA and the protein-translocation channel. Nature **455:**936-943.

10. **van der Wolk JP, de Wit JG, Driessen AJ.** 1997. The catalytic cycle of the *escherichia coli* SecA ATPase comprises two distinct preprotein translocation events. EMBO J **16:**7297-7304.

11. **Schiebel E, Driessen AJ, Hartl FU, Wickner W.** 1991. Delta mu H+ and ATP function at different steps of the catalytic cycle of preprotein translocase. Cell **64:**927-939.

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

13. **Hou JM, D'Lima NG, Rigel NW, Gibbons HS, McCann JR, Braunstein M, Teschke CM.** 2008. ATPase activity of *Mycobacterium tuberculosis* SecA1 and SecA2 proteins and its importance for SecA2 function in macrophages. J Bacteriol **190:**4880-4887.

14. **Rigel NW, Gibbons HS, McCann JR, McDonough JA, Kurtz S, Braunstein M.** 2009. The Accessory SecA2 System of Mycobacteria Requires ATP Binding and the Canonical SecA1. J Biol Chem **284:**9927-9936.

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

16. **Jilaveanu LB, Zito CR, Oliver D.** 2005. Dimeric SecA is essential for protein translocation. Proc Natl Acad Sci U S A **102:**7511-7516.

17. **van Stelten J, Silva F, Belin D, Silhavy TJ.** 2009. Effects of antibiotics and a proto-oncogene homolog on destruction of protein translocator SecY. Science **325:**753-756.

18. **Sassetti CM, Boyd DH, Rubin EJ.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol **48:**77-84.

19. **Shiba K, Ito K, Yura T, Cerretti DP.** 1984. A defined mutation in the protein export gene within the *spc* ribosomal protein operon of *Escherichia coli*: isolation and characterization of a new temperature-sensitive *secY* mutant. EMBO J **3:**631-635.

20. **Weiss JB, Ray PH, Bassford PJ, 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.

21. **Ghosh P.** 2004. Process of protein transport by the type III secretion system. Microbiol Mol Biol Rev **68:**771-795.

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

23. **Sassetti CM, Rubin EJ.** 2003. Genetic requirements for mycobacterial survival during infection. Proc Natl Acad Sci U S A **100:**12989-12994.

24. **Braunstein M, Espinosa BJ, Chan J, Belisle JT, Jacobs WR, Jr.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of Mycobacterium tuberculosis. Mol Microbiol **48:**453-464.

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

26. **Sullivan JT, Young EF, McCann JR, Braunstein M.** 2012. The Mycobacterium tuberculosis SecA2 system subverts phagosome maturation to promote growth in macrophages. Infect Immun **80:**996-1006.

27. **Vergne I, Chua J, Lee HH, Lucas M, Belisle J, Deretic V.** 2005. Mechanism of phagolysosome biogenesis block by viable Mycobacterium tuberculosis. Proc Natl Acad Sci U S A **102:**4033-4038.

28. **Manting EH, van der Does C, Driessen AJ.** 1997. In vivo cross-linking of the SecA and SecY subunits of the Escherichia coli preprotein translocase. J Bacteriol **179:**5699-5704.

29. **Consaul SA, Jacobs WR, Jr., Pavelka MS, Jr.** 2003. Extragenic suppression of the requirement for diaminopimelate in diaminopimelate auxotrophs of Mycobacterium smegmatis. FEMS Microbiol Lett **225:**131-135.

30. **Joshi SM, Pandey AK, Capite N, Fortune SM, Rubin EJ, Sassetti CM.** 2006. Characterization of mycobacterial virulence genes through genetic interaction mapping. Proc Natl Acad Sci U S A **103:**11760-11765.

31. **Rigel NW.** 2008. Characterization of the SecA2 Protein Export Pathway of Mycobacteria. University of North Carolina at Chapel Hill, Chapel Hill, NC.

32. **Feltcher ME, Gibbons HS, Ligon LS, Braunstein M.** 2013. Protein export by the mycobacterial SecA2 system is determined by the preprotein mature domain. J. Bacteriol. **195:**672-681.

APPENDIX

Extragenic Suppressors of Walker Box Mutant *secA2* Alleles

Official strain name	Strain isolation number	Genome sequenced	*secY* promoter mutation	*msmeg_1684* mutation	Azide resistance	Mueller-Hinton colony size	*secY* transcript level	SecY protein level	SecA2 protein level	SecA1 protein level
NR151	24S		Yes	No	Intermediate	Normal	High	Low	Normal	Normal
NR154	29S	Yes	No	Yes	Strong	Normal	Normal	Normal	Normal	Normal
NR155	33S	Yes	No	Yes	Strong	Normal	Normal	Normal	Normal	Normal
NR156	33B	Yes	No	Yes	Strong	Normal	Normal	Normal	Normal	Normal
NR201	39S		No	No						
NR230	1S	Yes	No	Yes	Intermediate	Small	Normal	Low	Normal	Normal
NR234	3S	Yes	No	Yes	Intermediate	Small	Normal	Low	Normal	Normal
NR236	4S	Yes	Yes	No	Intermediate	Normal	High	Low	Normal	Normal
NR238	5S		No	No	Intermediate	Normal		Normal	Low	Normal
NR242	7S	Yes	No	Yes[a]	Strong	Normal	Normal	None[b]	Normal	Normal
NR243	7B		No	No					None[c]	
NR248	10S	Yes	No	Yes	Strong	Normal	Normal	Normal	Normal	Normal
NR254	13B		No	No	Intermediate	Normal		Normal	Low	Normal
NR260	20B	Yes	No	Yes[a]	Strong	Normal	Normal	None[b]	Normal	Normal
NR122[a]	SSW1		No							
NR123[a]	SSW2		No	Yes[a]						
NR128[a]	SSW7		No							
NR129[a]	SSW8		No							
NR140[a]	SSW9		No							
NR141[a]	SSW10		No							
NR142[a]	SSW11		No							

Suppressors in top portion of table were isolated from an *M. smegmatis* strain expressing the *M. smegmatis secA2 K129R* allele (as described in Chapter 2 and (1)). This group of suppressors was pre-screened for the presence of a normal *secA2 K129R* allele (i.e. with no intragenic mutations) and for expression of SecA2 K129R protein. Suppressors in bottom portion of table were isolated from a *Mycobacterium smegmatis* strain expressing the *Mycobacterium tuberculosis secA2 K115R* allele (as described in (1)), which is comparable to the *M. smegmatis secA2 K129R* allele. This group of suppressors was screened for normal *secA2 K115R* alleles and for expression of SecA2 K115R protein upon initial isolation by Nathan Rigel, but have not been re-tested.

Empty cells indicate experiment has not been done. Mutations in *secY* promoter were assessed by sequencing approximately 1 kbp upstream of *secY*. "Strong" azide resistance and "normal" Mueller-Hinton colony size refers to phenotypes similar to the *secA2* deletion mutant. "Intermediate" azide resistance and "small" Mueller-Hinton colony size refers to phenotypes intermediate between those of the *secA2* deletion and *secA2 K129R* mutants. "Normal" transcript or protein levels refers to levels similar to the *secA2 K129R* mutant.

[a] In these suppressors, *msmeg_1684* is removed by a large chromosomal deletion.

[b] These large deletion type suppressors do not produce detectable levels of SecY protein, unlike other suppressors affecting *msmeg_1684*.

[c] This suppressor exhibited a normal SecA2 protein level when originally tested by Nathan Rigel, but no SecA2 was detected upon re-testing. This phenotype should be re-evaluated.

References

1. **Rigel NW.** 2008. Characterization of the SecA2 Protein Export Pathway of Mycobacteria. University of North Carolina at Chapel Hill, Chapel Hill, NC.