

Investigation of the Twin-arginine Translocation Pathway in Mycobacteria

Justin Andrew McDonough

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
2007

Approved by:
Advisor: Miriam Braunstein
Reader: Robert Bourret
Reader: Janne Cannon
Reader: Aravinda de Silva
Reader: Matthew Wolfgang

© 2007
Justin Andrew McDonough
ALL RIGHTS RESERVED

ABSTRACT

JUSTIN ANDREW McDONOUGH: Investigation of the Twin-arginine Translocation Pathway in Mycobacteria
(Under the direction of Miriam Braunstein)

Mycobacterium tuberculosis is an intracellular pathogen that resides primarily in host macrophages. Proteins that reside on the surface of the bacillus or are secreted into environment are ideally positioned to interact with host cell components and are therefore candidate virulence factors and immunogenic antigens. These proteins are actively transported from the cytoplasm to the cell envelope or extracellular space by specific export pathways. In this work, we describe the initial characterization of the *twin-arginine translocation* (Tat) export pathway in mycobacteria. The Tat pathway exports folded proteins, including virulence factors, in a number of bacteria. Proteins targeted for Tat-export carry N-terminal signal sequences that have a conserved twin-arginine motif, referred to as R-R-x- ϕ - ϕ (ϕ = uncharged residue). We provide evidence that the Tat pathway is functional in *Mycobacterium smegmatis* and required for the export of mycobacterial β -lactamases, including BlaC of *M. tuberculosis*. We demonstrate that BlaC can be used as a reporter to exclusively identify Tat-exported fusion proteins that promote resistance to the β -lactam antibiotic carbenicillin. We describe using the BlaC reporter with a *M. tuberculosis* genomic expression library to identify random Tat signal sequence-BlaC fusions that promote Tat export in *M. smegmatis*. Using this approach, we identified eleven *M. tuberculosis* Tat signal sequences shown to direct the export of the BlaC reporter through the Tat pathway, including the known phospholipase C virulence factors. In addition, we

identified a protein lacking the typical twin-arginine motif, suggesting that our current understanding of what defines a Tat signal sequence is limited, and demonstrating the effectiveness of our experimental system. We describe ongoing evaluation to determine the specificity and requirements for export of the authentic full length *M. tuberculosis* proteins. Finally, we demonstrate the novel finding that BlaC can be used as a reporter for the identification of proteins exported by *M. tuberculosis* during growth in macrophages. This may lead to the identification of a subset of proteins exported exclusively within the intracellular environment. This work has important implications in determining the role that protein export plays in *M. tuberculosis*, as well as increasing the understanding of how the Tat pathway functions in bacteria.

TABLE OF CONTENTS

LIST OF TABLES	vi
LIST OF FIGURES	vii
LIST OF ABBREVIATIONS AND SYMBOLS	ix
Chapter	
I. Introduction.....	1
II. The Twin-arginine Translocation (Tat) Pathway of <i>Mycobacterium smegmatis</i> is Functional and Required for the Export of Mycobacterial β -Lactamases	23
III. β -lactamase as a Genetic Reporter Used to Identify Twin-arginine Translocation (Tat) Signal Sequences of <i>Mycobacterium tuberculosis</i>	68
IV. β -lactamase can function as a reporter of bacterial protein export during <i>Mycobacterium tuberculosis</i> infection of host cells	131
V. Discussion.....	162

LIST OF TABLES

Table

2.1	Strains used in this study	28
2.2	Plasmids used in this study	35
2.3	<i>M. tuberculosis</i> signal sequences fused to 'BlaC.....	47
2.4	Export of functional 'BlaC is dependent on a twin-arginine signal sequence and an intact Tat pathway	48
3.1	Strains used in this study	76
3.2	Primers used in this study	78
3.3	Plasmids used in this study	84
3.4	Tat signal sequences identified by 'BlaC fusion.....	89
4.1	Plasmids used in this study	136

LIST OF FIGURES

Figure		
1.1	Sec and Tat export systems of <i>Mycobacterium tuberculosis</i>	8
2.1	Southern analysis of <i>tatA</i> and <i>tatC</i> recombinants in <i>M. smegmatis</i>	39
2.2	The Δ <i>tatA</i> and Δ <i>tatC</i> mutants have growth defects	41
2.3	The Δ <i>tatA</i> and Δ <i>tatC</i> mutants are hypersensitive to SDS	43
2.4	Nitrocefin assays on culture-filtrates and whole-cell extracts of <i>M. smegmatis</i> wild-type and Δ <i>tatA</i> and Δ <i>tatC</i> mutants	45
2.5	Nitrocefin assays on whole-cell extracts of <i>M. smegmatis</i> strains carrying 'BlaC fusion constructs	49
3.1	Construction of multicopy ' <i>blaC</i> expression libraries	77
3.2	Spot test confirmation of carbenicillin-resistant 'BlaC fusions	80
3.3	<i>M. tuberculosis</i> Tat signal sequences are capable of driving the export of 'BlaC	90
3.4	Distribution of predicted and experimentally verified <i>M. tuberculosis</i> twin-arginine signal sequences	93
3.5	Tat-dependent processing of BlaC in <i>M. smegmatis</i>	96
3.6	Assessment of signal sequence processing of <i>M. tuberculosis</i> full length candidate Tat substrates	98
3.7	The PlcB signal sequence is putatively processed in <i>M. tuberculosis</i> , but not in <i>M. smegmatis</i>	101
3.8	BlaC is localized to the detergent phase during Triton X-114 partitioning of <i>M. smegmatis</i>	118
3.9	Rv0315 of <i>M. tuberculosis</i> is homologous to the secreted β -(1,3)- glucanase of the yeast-lytic <i>Cellulosimicrobium cellulans</i>	119
3.10	UgpB and Rv2041c of <i>M. tuberculosis</i> are homologous	120

4.1	Schematic representation of signal sequence-‘BlaTEM-1 fusion constructs	141
4.2	‘BlaTEM-1 does not provide β -lactam-resistance to <i>ΔblaC</i> <i>M. tuberculosis</i>	143
4.3	‘BlaTEM-1 fusion proteins are detected at different amounts in <i>M. tuberculosis</i> whole cell lysates.....	145
4.4	The <i>ΔblaC</i> mutant of <i>M. tuberculosis</i> does not have a growth defect and is sensitive to β -lactam antibiotic in human THP-1 macrophage-like cells	148
4.5	<i>M. tuberculosis</i> signal sequences fused to ‘BlaC and ‘BlaTEM-1 protect intracellular bacilli from β -lactam antibiotics.....	150

LIST OF ABBREVIATIONS AND SYMBOLS

2D	two-dimensional
³² P	phosphorus
A	alanine
A ₄₈₆	absorbance at 486 nanometers
ABC	adenosine triphosphate binding cassette
AIDS	Acquired Immune Deficiency Syndrome
<i>aph</i>	kanamycin resistance gene
ATP	adenosine triphosphate
BCA	bisinchoninic acid
BCG	bacillus Calmette-Guérin
Bla	beta-lactamase
bp	base pair
C	cysteine
C-terminus	carboxy-terminus
carb	carbenicillin
<i>cat</i>	chloramphenicol acetyltransferase gene
CFU	colony-forming unit
D	aspartic acid
dCTP	deoxycytidine triphosphate
DNA	deoxyribonucleic acid
E	glutamic acid
<i>E.</i>	<i>Escherichia</i>

ESAT-6	early secreted antigen target 6 kilodalton
ESX-1	ESAT-6 secretion system
F	phenylalanine
FAD	fatty acid desaturase
FCS	fetal calf serum
Fe/S	iron-sulfur
G	glycine
GC	guanine/cytosine
GFP	green fluorescent protein
H	histidine
HA	hemagglutinin
HIV	Human Immunodeficiency Virus
<i>hyg</i>	hygromycin resistance gene
I	isoleucine
K	lysine
kbp	kilobase pair
kDa	kilodalton
L	leucine
M	methionine
<i>M.</i>	<i>Mycobacterium</i>
Mbp	megabase pair
MDR	multidrug-resistant
mg	milligram

MIC	minimum inhibitory concentration
min	minute
ml	milliliter
MOI	multiplicity of infection
N	asparagine
N-terminus	amino-terminus
ng	nanogram
nm	nanometer
OD ₆₀₀	optical density, 600 nanometers
ORF	open reading frame
ori	origin of replication
P	proline
<i>P.</i>	<i>Pseudomonas</i>
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
Plc	phospholipase C
PMA	phorbol myristate acetate
Q	glutamine
R	arginine
RNA	ribonucleic acid
S	serine
SCID	Severe Combined Immunodeficiency

SDS	sodium dodecyl sulfate
Sec	general secretion pathway
SRP	signal recognition particle
ss	signal sequence
T	threonine
Tat	twin-arginine translocation pathway
TB	tuberculosis
TraSH	transposon site hybridization
V	valine
W	tryptophan
WCL	whole cell lysate
XDR	extensively drug-resistant
Y	tyrosine
β	beta
Δ	deletion
μg	microgram
μl	microliter
μm	micrometer
σ	sigma factor
::	gene insertion

CHAPTER 1

INTRODUCTION

Tuberculosis disease.

Tuberculosis (TB) is an infectious disease caused by bacteria of the *Mycobacterium tuberculosis* complex. Each year, 1.7 million people die from TB, making it the leading cause of death worldwide from a bacterial infectious disease (61). It is estimated that up to one-third of the world's population, around two billion people, are infected with *M. tuberculosis*. The majority of people infected with *M. tuberculosis* are asymptomatic. However, five to ten percent of this population will progress to active, symptomatic disease over the course of a person's lifetime (61). Coinfection with HIV greatly increases the risk of developing TB to ten percent annually. If left untreated, half of the people with active disease will die from TB. Most TB deaths occur in developing countries where poor socio-economic conditions and HIV/AIDS incidence contribute to the transmission and effectiveness of TB disease.

TB can be cured by antibiotic treatment; however, the therapy is long (6-12 months) and involves a combination of antibiotics such as isoniazid and rifampin, to ensure the most effective treatment and to reduce the risk of drug-resistance. However, inadequate treatment, poor quality drugs, and the inability to maintain an appropriate drug regimen, has led to the emergence of drug-resistant strains of *M. tuberculosis* (33). Multidrug-resistant tuberculosis (MDR-TB) refers to strains that are resistant to at least two (isoniazid and rifampin) of the

primary antibiotics used to treat TB. These strains are difficult to treat, but not impossible, with the use of second line drugs. Recently, a second category of extensively drug-resistant strains (XDR-TB) emerged and is named so due to resistance to the primary drugs, isoniazid and rifampin, as well as many second-line drugs (61). Because XDR-TB is resistant to both first and second line drugs, it is virtually untreatable and therefore requires an urgent global response.

TB is primarily a disease of the lungs (pulmonary TB), although other parts of the body can be affected, such as the lymph nodes, kidneys, bones, joints, etc (extrapulmonary TB). The disease is transmitted when a person with active pulmonary TB coughs, sneezes, or talks thereby expelling aerosolized droplet nuclei carrying the bacilli, which can then be inhaled by another person. Upon inhalation, the bacteria are internalized primarily by alveolar macrophages where they survive and replicate. Here the bacteria elicit an immune response that can ultimately lead to containment of the infection or progression to active TB (46). This complex relationship between host and pathogen is not yet understood.

Protein export in *M. tuberculosis*.

The surface and secreted proteins of *M. tuberculosis* are attractive candidates for virulence factors and immunogenic antigens due to their exposure to the host cell environment (26). These proteins utilize specific transport pathways to cross the barriers imposed by the bacterial cytoplasmic membrane or cell wall (collectively termed the cell envelope) to reach their final location, and are therefore referred to as exported proteins. Exported proteins are ideally positioned to sense the environment, interact with the host, and protect the bacillus from host defenses during infection thereby promoting intracellular

survival. In addition, exported proteins can be antigens presented to the immune system during infection that elicit a host immune response. Together, exported proteins have important roles in bacterial physiology and pathogenesis and are of interest as targets for the development of drugs, vaccines, and diagnostics for TB.

Both surface and secreted proteins cross the cytoplasmic membrane and, depending on their final destination, may also navigate the cell wall. The mycobacterial cell wall is a unique, complex, and highly impermeable structure (20). At its core is a network of covalently linked peptidoglycan, arabinogalactan, and long chain (C60-C90) mycolic acids. Beyond the core structure is an outer layer of “free” lipids that is further capped by an exopolysaccharide capsule. This novel cell wall architecture is sure to present special challenges to protein transport and may necessitate mycobacterial-specific pathways.

The protein transport pathways of *M. tuberculosis* are of great interest due to their ability to export these important proteins. Our current understanding of protein export in *M. tuberculosis* remains limited (26, 31). However, the availability of the complete genome sequence of *M. tuberculosis* and the development of genetic tools to work with mycobacteria have stimulated research efforts to understand these complex pathways (6, 13). Two conserved protein export systems have been identified: the general secretion (Sec) pathway and the twin-arginine translocation (Tat) pathway. Together, these two systems probably account for the majority of protein transport across the cytoplasmic membrane. More specialized transport systems that act on select subsets of proteins also exist in *M. tuberculosis*: the SecA2-dependent system and the ESAT-6 secretion system (ESX-1).

In this Chapter, I describe the recent advances made identifying and understanding the protein transport pathways of *M. tuberculosis*, with emphasis on the twin-arginine

translocation (Tat) pathway, which is the focus of the research discussed herein. The Tat pathway is essential for exporting virulence factors in a number of bacterial pathogens (10, 11, 19, 28, 40, 45, 59). Upon initiation of this study, the Tat export pathway had not been characterized in mycobacteria. However, the *M. tuberculosis* genome sequence revealed homologs to components of the Tat translocase and putative Tat-exported proteins (13, 18), including known virulence factors (42). As described in Chapters 2-4, we investigated the Tat pathway of mycobacteria. We showed it to be functional in *M. smegmatis* and we have identified a set of the *M. tuberculosis* exported Tat substrates. In the process, we developed an enzymatic reporter, based on the Tat-exported β -lactamase of *M. tuberculosis* that can be used in a genetic selection to assay for Tat-dependent export in mycobacteria. In addition to being a useful tool for testing the export of predicted Tat substrates, the reporter can also be used to select *M. tuberculosis* sequences that promote Tat-dependent export from a reporter fusion library. A significant feature of this reporter library system is that it can provide information as to the specific sequences required for targeting to the Tat apparatus. This has an advantage over bioinformatic predictions of putative Tat substrates, which rely on information from a relatively small dataset of experimentally proven Tat substrates which do not include mycobacterial examples, and therefore may yield inaccurate predictions.

Much of the work described in the following chapters was carried out in the fast-growing, saprophytic organism *Mycobacterium smegmatis*. Although *M. smegmatis* is not a human pathogen, it is a useful genetic model to study some aspects of mycobacterial physiology, and can be a starting point for subsequent work in the slow-growing, pathogenic *M. tuberculosis* (43). It also provides a convenient model to study mycobacteria without constraints imposed by biosafety level 3 containment.

Sec Pathway

The general secretion (Sec) pathway is a conserved protein export system of bacteria. Years of study, primarily in *Escherichia coli*, provide an advanced understanding of this pathway (for a comprehensive review and detailed references see 16, 35). The Sec pathway transports proteins from the cytosol across the cytoplasmic membrane to the cell envelope. These proteins may be further secreted into the extracellular space through subsequent action of Sec-independent mechanisms.

At the core of the Sec machinery is the Sec translocase composed of the integral membrane proteins SecY, SecE, and SecG, which serve as the channel through which proteins cross the cytoplasmic membrane. Peripherally associated with the Sec translocase is the multifunctional SecA protein. SecA binds to cytoplasmic precursors of exported proteins and to associated chaperones, such as SecB, that maintain these precursors in a translocation-competent unfolded state. SecA binds to membrane phospholipids and to the integral membrane components of the translocase. SecA is also an ATPase that provides energy for translocation. Through cycles of ATP binding and hydrolysis, SecA undergoes conformational changes that drive stepwise export of the precursor through the Sec channel and across the membrane. Because many Sec-exported proteins play essential roles in cell physiology, the central components of the Sec apparatus (SecY, SecE, and SecA) are essential for bacterial viability. Other proteins that are not required for Sec export but increase efficiency of the process (SecG, SecDF, YajC) are not essential.

Sec exported proteins are synthesized as precursors that have N-terminal signal peptides required for transport (41). Typically, the signal peptide is 18-30 amino acids long with a tripartite structure: a positively charged N-region, a hydrophobic H-region, and a C-

region with a cleavage site. During or immediately following translocation of the protein, a signal peptidase cleaves the signal peptide to produce the mature protein. For standard Sec signal peptides, cleavage is performed by signal peptidase I (LepB) and the cleavage site (often denoted A-X-A) has short-chain amino acids at the -1 and -3 positions relative to the first amino acid of the cleaved mature protein. The conserved elements of Sec signal peptides enable bioinformatic prediction of Sec-exported proteins.

The Sec pathway also exports lipoproteins that, by virtue of a covalently attached lipid, remain anchored to the cell envelope. Bacterial lipoproteins are synthesized with N-terminal signal peptides that are similar in overall structure to standard Sec signal peptides, but possess a lipobox motif (L-A/S-G/A-C₊₁) in the C-region (2). The lipobox serves as both the site of lipid modification at the invariant cysteine and the site of signal peptide cleavage, immediately prior to the invariant cysteine. The requirement for a lipobox in relation to the other signal peptide elements enables bioinformatic prediction of lipoproteins. In contrast to standard Sec signal peptides, lipoprotein signal peptides are cleaved by the lipoprotein-specific signal peptidase II, LspA (for a review of lipoprotein processing see 62). In Gram-positive and Gram-negative bacteria, lipoproteins are found in the cytoplasmic membrane. In Gram-negative bacteria, lipoproteins can also be routed to the outer membrane through a process involving the Lol sorting system (reviewed in 56).

In conjunction with additional factors, the Sec pathway also participates in the signal recognition particle (SRP)-mediated process of positioning integral membrane proteins in the cytoplasmic membrane (reviewed in 14). Bacterial SRP, a protein-RNA complex comprised of the Ffh protein and *ffs* (a 4.5S RNA), recognizes transmembrane domains as they emerge from the ribosome and delivers them to the SRP receptor protein FtsY, which in turn directs

the complex to the Sec translocase for export across the membrane. The membrane protein YidC aids in the stable integration of the transmembrane domain into the lipid bilayer of the membrane. This SRP/Sec/YidC-dependent pathway is used by many membrane proteins; however, there are also examples of transmembrane proteins localized in a Sec-independent but YidC-dependent manner (58).

Sec Pathway in *M. tuberculosis*

M. tuberculosis appears to have a functional Sec pathway (Fig. 1.1). The *M. tuberculosis* genome contains *open reading frames* (ORFs) predicted to encode orthologs of the Sec pathway components (7, 13). Genomic analysis also reveals the unusual presence of two SecA homologs: SecA1 and SecA2. The only Sec factor not obvious in the mycobacterial genomes is the SecB chaperone; however, SecB is not present in all bacteria (53). Consistent with the lack of SecB is the observation that the *M. tuberculosis* SecAs lack consensus SecB-binding sites.

The identification of surface and secreted proteins synthesized with recognizable Sec signal peptides provides further support for a functional Sec pathway in *M. tuberculosis* (26). In addition, application of bioinformatic methods to identify Sec signal peptides and transmembrane domains in the ORFs of the *M. tuberculosis* genome predict as much as 23% of the *M. tuberculosis* genome encodes Sec-exported proteins. This prediction is on the order of that proposed for *E. coli* and most other bacteria (23, 48, 52, 55).

So far, the only Sec proteins directly investigated in mycobacteria are SecA1 and SecA2. The presence of two SecAs in *M. tuberculosis* appears to be a property shared with all mycobacterial species including non-pathogenic *M. smegmatis* (8, 13), as well as some

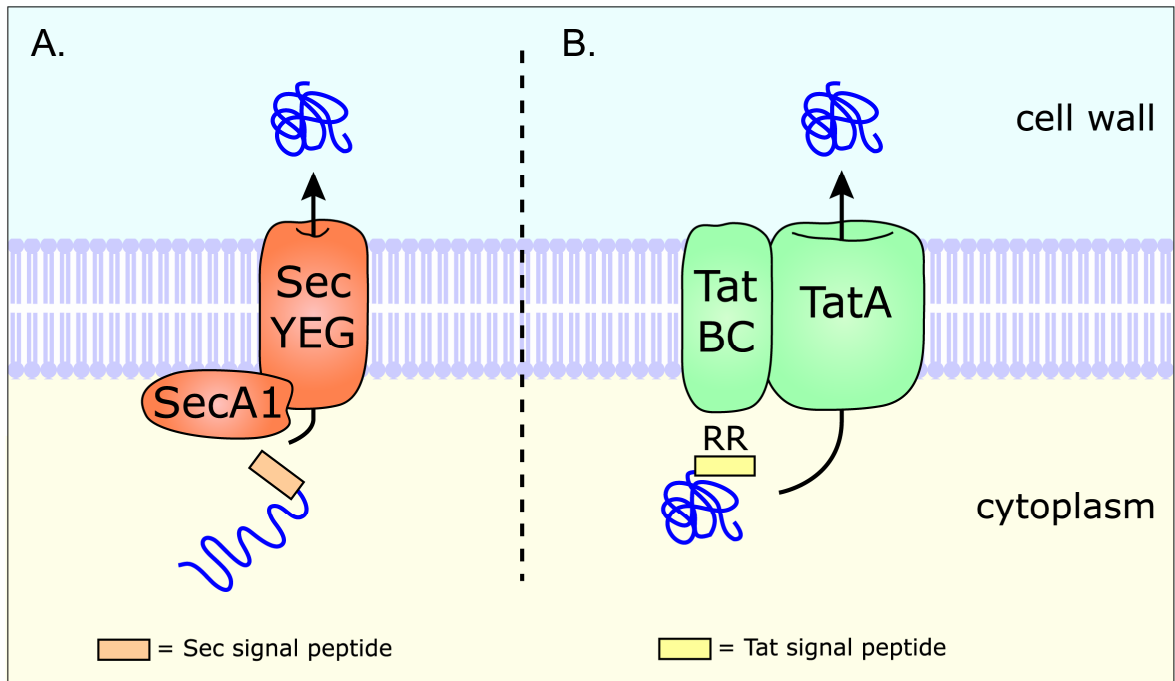


Figure 1.1. Sec and Tat export systems of *Mycobacterium tuberculosis*.

Protein export across the cytoplasmic membrane in *M. tuberculosis*. A) Proposed export through the Sec pathway in *M. tuberculosis*. Unfolded proteins with signal peptides (open box) are recognized by SecA1 and transported through the SecYEG channel. After transport across the membrane, the signal peptide is cleaved and the proteins fold. SecA2 is not shown in this diagram. B) Proposed export through the Tat pathway in *M. tuberculosis*. Folded precursor proteins with Tat signal peptides are delivered to TatBC and transported through the TatA channel.

Gram positive bacteria (5, 12, 30). Both SecA1 and SecA2 are highly homologous to SecA of other bacteria and both include predicted ATP binding sites. However, of the two proteins, SecA1 shares more sequence similarity with the singular and essential SecA of other bacteria. Furthermore, only SecA1 is essential (8, 51). Finally, the N-domain of *M. tuberculosis* SecA1 is functional in *E. coli* as a chimeric protein composed of the N-domain of SecA1 and the C-domain of *E. coli* SecA (36). These findings lead to the conclusion that SecA1 is the “housekeeping” SecA with a major function in transporting proteins beyond the cytoplasm including proteins involved in essential activities. Interestingly, SecA2 cannot fulfill the essential role of SecA1. Even when *secA2* is overexpressed from a plasmid, the chromosomal copy of *secA1* in *M. smegmatis* cannot be deleted (8). This indicates that SecA1 and SecA2 are not functionally equivalent or redundant. SecA2 in mycobacteria is involved in the export of a subset of proteins including lipoproteins with N-terminal signal sequences and secreted proteins that lack recognizable signal sequences (8, 9, 21). SecA2 in *M. tuberculosis* is required for virulence and modulation of the host immune response (9, 27).

Lipoprotein Processing and Export in *M. tuberculosis*

Young and Garbe (1991) were the first to demonstrate the presence of lipoproteins in *M. tuberculosis* (65). More recently, bioinformatic analysis revealed 99 proteins with predicted Sec signal sequences and lipobox motifs in the *M. tuberculosis* genome (55). Demonstration of lipid modification is lacking for the majority of assumed mycobacterial lipoproteins and some may be misidentified. *M. tuberculosis* has a functional lipoprotein signal peptidase LspA as demonstrated by the accumulation of modified but uncleaved

lipoprotein precursors in a *M. tuberculosis* *lspA* mutant (4, 49). Testing of *M. tuberculosis* *lspA* mutants in mice and macrophages further showed that lipoprotein processing is important for *M. tuberculosis* virulence and stimulation of the host Toll-like receptor 2 signaling pathway, an important arm of innate immunity (4, 49).

The final location of lipoproteins in mycobacteria is an unresolved issue. Mycobacterial lipoproteins may remain anchored to the cytoplasmic membrane or reach more external locations such as the outer lipid layer of the cell wall in a manner akin to Lol sorting of lipoproteins in Gram-negative bacteria. There are no obvious orthologs of the Lol system in mycobacteria. However, the crystal structure of *M. tuberculosis* LppX, which functions in the transport of phthiocerol dimycocerosate lipid to the cell surface, is structurally similar to *E. coli* LolA and LolB (54). This raises the possibility that functional homologs of Lol exist in *M. tuberculosis*. Additionally, there are a small number of lipoproteins reported in *M. tuberculosis* culture filtrates (44); however, it remains to be demonstrated that these secreted species are still lipidated and not trimmed free of the lipid moiety.

Tat Pathway

The twin-arginine translocation (Tat) pathway is a more recently identified system for protein export in bacteria (for comprehensive reviews and references see 29, 50). In contrast to the Sec pathway, not all bacteria have a recognizable Tat pathway (18, 64). Like the Sec pathway, the Tat pathway exports proteins across the cytoplasmic membrane. The proteins may be further secreted beyond the cell envelope by a subsequent system (59). Although

there are similarities between the Tat and Sec pathways, Tat-mediated export is Sec-independent and has the distinct mechanism of translocating folded proteins.

The majority of research on the Tat pathway has been performed in *E. coli* (29, 50). In *E. coli*, the Tat translocation apparatus is composed of integral membrane proteins TatA, TatB, and TatC. Some bacteria have an additional TatA homolog with overlapping functions, designated TatE. TatA/E and TatB are similar in structure; both are small with a single N-terminal transmembrane domain and an amphipathic tail. TatC is a larger protein with six transmembrane domains. *In vitro*, TatA forms large homo-oligomeric rings of varying size with a central cavity (22). This led to the hypothesis that TatA complexes represent the membrane-spanning channel through which Tat substrates are translocated. The size variation in TatA complexes could account for the wide variety of folded proteins that transit the Tat channel, which includes multiprotein and cofactor-bound complexes.

The current model for Tat export is as follows (22, 29, 50) (Fig. 1.1B). A folded Tat precursor is targeted to a TatBC complex via the Tat signal peptide (1). After initial recognition, TatA channels are recruited to the complex and the folded precursor is passed along to the channel and transported across the membrane. The proton motive force provides energy for this process. Protein folding is a prerequisite for Tat export (17), and some cofactor-bound Tat precursors have cognate cytoplasmic chaperones that prevent targeting to the Tat pathway until cofactor insertion and/or folding is complete (25, 37, 38). Some bacteria lack a *tatB* ortholog (18). In these cases, TatA appears to fulfill the function of TatB, making the minimum functional Tat components TatA and TatC.

Proteins exported by the Tat pathway are synthesized as precursors with N-terminal signal peptides that have similar tripartite structures to Sec signal peptides. This includes the

presence of a recognition site that is cleaved during export to generate the mature protein. In Tat precursors the standard A-X-A cleavage site is also cleaved by the LepB peptidase (63). There are also Tat signal peptides with lipoboxes and LspA cleavage sites (60). However, there has yet to be direct demonstration of lipidation or LspA cleavage of a Tat exported protein. Additionally, there are examples of Tat substrates that are cytoplasmic membrane proteins. In *E. coli*, there are examples of proteins transported to the cytoplasmic membrane in a Tat-dependent fashion but remain integrated into the membrane via the C-terminus of the protein (24). There are also examples of Fe/S-containing Rieske proteins that are directed to the membrane by N-terminal Tat signal sequences but since the signal sequence is not cleaved they remain integral membrane proteins (3, 15, 34).

The major distinction between Sec and Tat signal peptides is the presence of a consensus twin-arginine motif in the charged N-region, designated R-R-x- ϕ - ϕ (ϕ = uncharged residue) (reviewed in 29, 50). The twin-arginines 'RR' is an important feature of the signal peptide since conservative substitution with twin-lysines 'KK' prevents export. However, the requirement for 'RR' is not absolute since exceptions to this rule exist and there are other features believed to help define a Tat signal peptide. Because Tat substrates must be folded prior to export, the ability of the mature domain of the precursor to fold in the cytoplasm also plays a role in determining whether a protein can be exported by the Tat pathway (17). In connection with this data, it was reported recently that some *E. coli* twin-arginine signal sequences direct the export of an unfolded reporter to the Sec pathway and a folded reporter to the Tat pathway, and therefore represent a category of signal sequences that are promiscuous for export (57).

Tat pathway in *M. tuberculosis*

M. tuberculosis has a functional Tat pathway (Figure 1.1). The genome of *M. tuberculosis* contains orthologs of *tatABC* (13). The *tatA* and *tatC* genes are in an operon and cotranscribed (47). The *tatB* gene is distantly located from the *tatAC* operon. ORFs predicted to encode proteins with Tat signal peptides are also present in the genome (18). The application of bioinformatics-based Tat substrate recognition programs to the *M. tuberculosis* genome is discussed in Chapter 3.

The *M. tuberculosis* *tatABC* genes appear to be essential under standard laboratory conditions. Disruption of *tatAC* or *tatB* in *M. tuberculosis* was unsuccessful unless a second copy of the genes was being expressed from the chromosome (47). These findings are consistent with *transposon site hybridization* (TraSH) analysis of *M. tuberculosis* that reported *tatC* as essential for optimal growth (51). In other bacteria, *tat* mutants often display growth defects; however, *M. tuberculosis* is the first reported example where the Tat pathway is essential for *in vitro* growth. The basis of the essential nature of the *M. tuberculosis* Tat pathway is not clear. Mislocalization of Tat substrates involved in nutrient uptake seems the likely explanation, since several putative *M. tuberculosis* Tat substrates are homologous to periplasmic substrate-binding proteins of ABC transporters (18).

In *M. smegmatis*, *tatABC* genes are also present and have the same genomic arrangement as in *M. tuberculosis*. In *M. smegmatis* it is possible to construct individual deletion mutants of *tatA*, *tatB*, and *tatC*; although, these mutants have *in vitro* growth defects (32, 39). The characterization of the Tat pathway in *M. smegmatis* is discussed in greater detail in Chapter 2 of this work. In addition, Posey *et al.* (39) independently reported results with *M. smegmatis* Δ *tat* mutants that confirmed our results and also identified some

additional properties. *M. smegmatis tat* mutants have several additional phenotypes: hypersensitivity to sodium dodecyl sulfate, changes in colony morphology, inability to use succinate as a sole carbon source, and increased sensitivity to β -lactam antibiotics. The pleiotropic effects of the *tat* mutations suggest that the mycobacterial Tat pathway exports multiple substrates. Comparative 2D-PAGE of *M. smegmatis* culture filtrate proteins revealed five proteins that were *tat*-dependent for secretion (39). Tat substrates are also likely to exist in the cell envelope.

Several lines of evidence demonstrate that export of the *M. smegmatis* β -lactamase (BlaS) is Tat-dependent and that the β -lactam sensitivity phenotype of *M. smegmatis tat* mutants is due to a defect in BlaS export. Because β -lactams target the cell wall, β -lactamases must be exported to protect the cell. Culture filtrates produced from *M. smegmatis tat* mutants contain significantly less β -lactamase, as shown by 2D-PAGE analysis and by *in vitro* enzyme activity assay (32, 39). Further, *M. smegmatis* BlaS as well as the *M. tuberculosis* β -lactamase BlaC have predicted Tat signal peptides. Experiments in which the *M. tuberculosis* BlaC is expressed in a *blaS* β -lactamase mutant of *M. smegmatis* or a *tatA/blaS* double mutant show that BlaC can protect from β -lactams only in the presence of a functional Tat pathway (32). Site-directed mutagenesis experiments further demonstrate that this ability depends on the 'RR' dipeptide in the twin-arginine motif of BlaC. These experiments strongly suggest that *M. tuberculosis* BlaC, like *M. smegmatis* BlaS, is exported by the Tat pathway and that the Tat pathway is required for the β -lactam resistance of *M. tuberculosis*.

One can exploit the Tat-export specific properties of BlaC as a reporter tool for study and identification of *M. tuberculosis* proteins with Tat signal peptides. A truncated 'BlaC

lacking its native Tat signal peptide is not exported and does not protect against β -lactams. However, fusion of a Tat signal peptide to the N-terminus of 'BlaC promotes export as measured by β -lactam resistance, and this export requires a functional Tat pathway since protection from β -lactams is not seen in *M. smegmatis* *tat* mutants (32). The utilization of the 'BlaC reporter to identify *M. tuberculosis* Tat-exported fusion proteins is discussed at greater length in Chapter 3.

The Tat pathway functions in several bacterial pathogens and contributes to virulence (reviewed in 29). *Pseudomonas aeruginosa* uses the Tat pathway for export of the phospholipase C virulence factor PlcH across the cytoplasmic membrane. It is likely that the Tat pathway also contributes to the virulence of *M. tuberculosis*. The cell wall-localized phospholipase C proteins (Plc) of *M. tuberculosis* have predicted Tat signal peptides. These Plc enzymes are virulence factors in *M. tuberculosis* (42). Using the 'BlaC reporter, the PlcB signal peptide was shown to export functional β -lactamase in a Tat and twin 'RR'-dependent manner (32). This indicates that the PlcB signal peptide of *M. tuberculosis* is functional in Tat export. Furthermore, export of a functional PlcB fusion to β -lactamase protected *M. tuberculosis* from β -lactam antibiotics during growth in macrophages, indicating that this reporter can report on proteins exported within host cells. This will be discussed in greater detail in Chapter 4 of this work. Of the other predicted Tat exported proteins in *M. tuberculosis*, the majority are uncharacterized. Interestingly, a *M. tuberculosis* mutant of *Rv2525c*, which encodes a predicted Tat exported protein, exhibits increased virulence in macrophages and in SCID mice (47).

Summary

Over recent years, our understanding of the protein transport systems operating in *M. tuberculosis* has greatly improved. This success was largely driven by the ability to construct mutant strains of *M. tuberculosis* and test individual gene products for roles in protein transport. We anticipate that future research will extend these efforts, combine them with sophisticated biochemical characterization of the pathways, and uncover additional pathways that transport proteins across the novel mycobacterial cell wall.

In the following chapters, we describe the initial characterization of the Tat pathway in *M. smegmatis* and demonstrate that mycobacterial β -lactamases are native Tat-exported proteins (Chapter 2). We describe the development of a β -lactamase reporter for Tat-dependent export in mycobacteria (Chapter 2). We demonstrate the use of this reporter to identify *M. tuberculosis* signal sequences capable of promoting Tat export of the reporter (Chapter 3). We also describe ongoing evaluation of the export of the respective full length *M. tuberculosis* proteins from the signal sequences we identified (Chapter 3). We conclude with the development and application of a β -lactamase reporter system that works to identify Tat proteins exported during growth of *M. tuberculosis* within macrophages (Chapter 4). Together this represents the first steps in determining the role that the Tat pathway plays in processes essential for growth of *M. tuberculosis in vitro* and intracellularly. These studies may pave the way for the development of newer effective anti-TB therapies and better diagnostics to report on infection. Additionally, the ability of the Tat translocase to export folded proteins and complexes may be exploited for the improved secretion of protective antigens and heterologous proteins.

References

1. **Alami, M., I. Luke, S. Deitermann, G. Eisner, H. G. Koch, J. Brunner, and M. Muller.** 2003. Differential interactions between a twin-arginine signal peptide and its translocase in *Escherichia coli*. *Mol Cell* **12**:937-46.
2. **Babu, M. M., M. L. Priya, A. T. Selvan, M. Madera, J. Gough, L. Aravind, and K. Sankaran.** 2006. A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *J Bacteriol* **188**:2761-73.
3. **Bachmann, J., B. Bauer, K. Zwicker, B. Ludwig, and O. Anderka.** 2006. The Rieske protein from *Paracoccus denitrificans* is inserted into the cytoplasmic membrane by the twin-arginine translocase. *FEBS J* **273**:4817-30.
4. **Banaiee, N., E. Z. Kincaid, U. Buchwald, W. R. Jacobs, Jr., and J. D. Ernst.** 2006. Potent inhibition of macrophage responses to IFN-gamma by live virulent *Mycobacterium tuberculosis* is independent of mature mycobacterial lipoproteins but dependent on TLR2. *J Immunol* **176**:3019-27.
5. **Bensing, B. A., and P. M. Sullam.** 2002. An accessory sec locus of *Streptococcus gordonii* is required for export of the surface protein GspB and for normal levels of binding to human platelets. *Mol Microbiol* **44**:1081-94.
6. **Braunstein, M., S. S. Bardarov, and W. R. J. Jacobs.** 2002. Genetic methods for deciphering virulence determinants of *Mycobacterium tuberculosis*, p. 67-99. In V. L. Clark and P. M. Bavoil (ed.), *Methods in Enzymology*, vol. 358. Academic Press, London.
7. **Braunstein, M., and J. T. Belisle.** 2000. Genetics of protein secretion, p. 203-220. In G. F. Hatfull and W. R. J. Jacobs (ed.), *Molecular genetics of mycobacteria*. ASM Press.
8. **Braunstein, M., A. M. Brown, S. Kurtz, and W. R. Jacobs, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. *J Bacteriol* **183**:6979-90.
9. **Braunstein, M., B. J. Espinosa, J. Chan, J. T. Belisle, and W. R. Jacobs, Jr.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* **48**:453-64.
10. **Bronstein, P. A., M. Marrichi, S. Cartinhour, D. J. Schneider, and M. P. DeLisa.** 2005. Identification of a twin-arginine translocation system in *Pseudomonas syringae* pv. tomato DC3000 and its contribution to pathogenicity and fitness. *J Bacteriol* **187**:8450-61.

11. **Caldelari, I., S. Mann, C. Crooks, and T. Palmer.** 2006. The Tat pathway of the plant pathogen *Pseudomonas syringae* is required for optimal virulence. *Mol Plant Microbe Interact* **19**:200-12.
12. **Chen, Q., H. Wu, and P. M. Fives-Taylor.** 2004. Investigating the role of secA2 in secretion and glycosylation of a fimbrial adhesin in *Streptococcus parasanguis* FW213. *Mol Microbiol* **53**:843-56.
13. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, F. Tekaiia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, and B. G. Barrell.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**:537-44.
14. **Dalbey, R. E., and M. Chen.** 2004. Sec-translocase mediated membrane protein biogenesis. *Biochim Biophys Acta* **1694**:37-53.
15. **De Buck, E., L. Vranckx, E. Meyen, L. Maes, L. Vandersmissen, J. Anne, and E. Lammertyn.** 2007. The twin-arginine translocation pathway is necessary for correct membrane insertion of the Rieske Fe/S protein in *Legionella pneumophila*. *FEBS Lett* **581**:259-64.
16. **de Keyzer, J., C. van der Does, and A. J. Driessen.** 2003. The bacterial translocase: a dynamic protein channel complex. *Cell Mol Life Sci* **60**:2034-52.
17. **DeLisa, M. P., D. Tullman, and G. Georgiou.** 2003. Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. *Proc Natl Acad Sci U S A* **100**:6115-20.
18. **Dilks, K., R. W. Rose, E. Hartmann, and M. Pohlschroder.** 2003. Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. *J Bacteriol* **185**:1478-83.
19. **Ding, Z., and P. J. Christie.** 2003. *Agrobacterium tumefaciens* twin-arginine-dependent translocation is important for virulence, flagellation, and chemotaxis but not type IV secretion. *J Bacteriol* **185**:760-71.
20. **Draper, P., and M. Daffe.** 2005. The cell envelope of *Mycobacterium tuberculosis* with special reference to the capsule and outer permeability barrier, p. 261-273. In S. T. Cole, K. D. Eisenach, D. N. McMurray, and W. R. J. Jacobs (ed.), *Tuberculosis and the Tubercle Bacillus*. ASM Press, Washington, DC.
21. **Gibbons, H. S., F. Wolschendorf, M. Abshire, M. Niederweis, and M. Braunstein.** 2007. Identification of two *Mycobacterium smegmatis* lipoproteins exported by a SecA2-dependent pathway. *J Bacteriol.* **189**:5090-100.

22. **Gohlke, U., L. Pullan, C. A. McDevitt, I. Porcelli, E. de Leeuw, T. Palmer, H. R. Saibil, and B. C. Berks.** 2005. The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. *Proc Natl Acad Sci U S A* **102**:10482-6.
23. **Gomez, M., S. Johnson, and M. L. Gennaro.** 2000. Identification of secreted proteins of *Mycobacterium tuberculosis* by a bioinformatic approach. *Infect Immun* **68**:2323-7.
24. **Hatzixanthis, K., T. Palmer, and F. Sargent.** 2003. A subset of bacterial inner membrane proteins integrated by the twin-arginine translocase. *Mol Microbiol* **49**:1377-90.
25. **Hatzixanthis, K., D. J. Richardson, and F. Sargent.** 2005. Chaperones involved in assembly and export of N-oxide reductases. *Biochem Soc Trans* **33**:124-6.
26. **Kurtz, S., and M. Braunstein.** 2005. Protein secretion and export in *Mycobacterium tuberculosis*, p. 71-138. *In* T. Parish (ed.), *Mycobacterium* molecular biology. Horizon bioscience, Norfolk, UK.
27. **Kurtz, S., K. P. McKinnon, M. S. Runge, J. P. Ting, and M. Braunstein.** 2006. The SecA2 secretion factor of *Mycobacterium tuberculosis* promotes growth in macrophages and inhibits the host immune response. *Infect Immun* **74**:6855-64.
28. **Lavander, M., S. K. Ericsson, J. E. Broms, and A. Forsberg.** 2006. The twin arginine translocation system is essential for virulence of *Yersinia pseudotuberculosis*. *Infect Immun* **74**:1768-76.
29. **Lee, P. A., D. Tullman-Ercek, and G. Georgiou.** 2006. The bacterial twin-arginine translocation pathway. *Annu Rev Microbiol* **60**:373-395.
30. **Lenz, L. L., and D. A. Portnoy.** 2002. Identification of a second *Listeria secA* gene associated with protein secretion and the rough phenotype. *Mol Microbiol* **45**:1043-56.
31. **McDonough, J. A., and M. Braunstein.** 2007. Protein transport in *Mycobacterium tuberculosis*. *In* S. H. Kaufmann and E. J. Rubin (ed.), *MTB Handbook Volume 1: Molecular Genetics and Biochemistry*, vol. In press. Wiley-VCH Verlag Press.
32. **McDonough, J. A., K. E. Hacker, A. R. Flores, M. S. Pavelka, Jr., and M. Braunstein.** 2005. The twin-arginine translocation pathway of *Mycobacterium smegmatis* is functional and required for the export of mycobacterial beta-lactamases. *J Bacteriol* **187**:7667-79.

33. **Meya, D. B., and K. P. McAdam.** 2007. The TB pandemic: an old problem seeking new solutions. *J Intern Med* **261**:309-29.
34. **Molik, S., I. Karnauchoy, C. Weidlich, R. G. Herrmann, and R. B. Klosgen.** 2001. The Rieske Fe/S protein of the cytochrome b6/f complex in chloroplasts: missing link in the evolution of protein transport pathways in chloroplasts? *J Biol Chem* **276**:42761-6.
35. **Mori, H., and K. Ito.** 2001. The Sec protein-translocation pathway. *Trends Microbiol* **9**:494-500.
36. **Owens, M. U., W. E. Swords, M. G. Schmidt, C. H. King, and F. D. Quinn.** 2002. Cloning, expression, and functional characterization of the *Mycobacterium tuberculosis secA* gene. *FEMS Microbiol Lett* **211**:133-41.
37. **Palmer, T., F. Sargent, and B. C. Berks.** 2005. Export of complex cofactor-containing proteins by the bacterial Tat pathway. *Trends Microbiol* **13**:175-80.
38. **Perez-Rodriguez, R., A. C. Fisher, J. D. Perlmutter, M. G. Hicks, A. Chanal, C. L. Santini, L. F. Wu, T. Palmer, and M. P. Delisa.** 2007. An essential role for the DnaK molecular chaperone in stabilizing over-expressed substrate proteins of the bacterial twin-arginine translocation pathway. *J Mol Biol.* **367**:715-30.
39. **Posey, J. E., T. M. Shinnick, and F. D. Quinn.** 2006. Characterization of the twin-arginine translocase secretion system of *Mycobacterium smegmatis*. *J Bacteriol* **188**:1332-40.
40. **Pradel, N., C. Ye, V. Livrelli, J. Xu, B. Joly, and L. F. Wu.** 2003. Contribution of the twin arginine translocation system to the virulence of enterohemorrhagic *Escherichia coli* O157:H7. *Infect Immun* **71**:4908-4916.
41. **Pugsley, A. P.** 1993. The complete general secretory pathway in gram-negative bacteria. *Microbiol Rev* **57**:50-108.
42. **Raynaud, C., C. Guilhot, J. Rauzier, Y. Bordat, V. Pelicic, R. Manganelli, I. Smith, B. Gicquel, and M. Jackson.** 2002. Phospholipases C are involved in the virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* **45**:203-17.
43. **Reyrat, J. M., and D. Kahn.** 2001. *Mycobacterium smegmatis*: an absurd model for tuberculosis? *Trends Microbiol* **9**:472-4.
44. **Rosenkrands, I., A. King, K. Weldingh, M. Moniatte, E. Moertz, and P. Andersen.** 2000. Towards the proteome of *Mycobacterium tuberculosis*. *Electrophoresis* **21**:3740-56.

45. **Rossier, O., and N. P. Cianciotto.** 2005. The *Legionella pneumophila* *tatB* gene facilitates secretion of phospholipase C, growth under iron-limiting conditions, and intracellular infection. *Infect Immun* **73**:2020-32.
46. **Russell, D. G.** 2007. Who puts the tubercle in tuberculosis? *Nat Rev Microbiol* **5**:39-47.
47. **Saint-Joanis, B., C. Demangel, M. Jackson, P. Brodin, L. Marsollier, H. Boshoff, and S. T. Cole.** 2006. Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of *Mycobacterium tuberculosis*, increases {beta}-lactam susceptibility and virulence. *J Bacteriol* **188**:6669-79.
48. **Saleh, M. T., M. Fillon, P. J. Brennan, and J. T. Belisle.** 2001. Identification of putative exported/secreted proteins in prokaryotic proteomes. *Gene* **269**:195-204.
49. **Sander, P., M. Rezwani, B. Walker, S. K. Rampini, R. M. Kroppenstedt, S. Ehlers, C. Keller, J. R. Keeble, M. Hagemeyer, M. J. Colston, B. Springer, and E. C. Bottger.** 2004. Lipoprotein processing is required for virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* **52**:1543-52.
50. **Sargent, F., B. C. Berks, and T. Palmer.** 2006. Pathfinders and trailblazers: a prokaryotic targeting system for transport of folded proteins. *FEMS Microbiol Lett* **254**:198-207.
51. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48**:77-84.
52. **Schneider, G.** 1999. How many potentially secreted proteins are contained in a bacterial genome? *Gene* **237**:1113-1121.
53. **Scott, J. R., and T. C. Barnett.** 2006. Surface proteins of gram-positive bacteria and how they get there. *Annu Rev Microbiol* **60**:397-423.
54. **Sulzenbacher, G., S. Canaan, Y. Bordat, O. Neyrolles, G. Stadthagen, V. Roig-Zamboni, J. Rauzier, D. Maurin, F. Laval, M. Daffe, C. Cambillau, B. Gicquel, Y. Bourne, and M. Jackson.** 2006. LppX is a lipoprotein required for the translocation of phthiocerol dimycocerosates to the surface of *Mycobacterium tuberculosis*. *EMBO J* **25**:1436-44.
55. **Sutcliffe, I. C., and D. J. Harrington.** 2004. Lipoproteins of *Mycobacterium tuberculosis*: an abundant and functionally diverse class of cell envelope components. *FEMS Microbiol Rev* **28**:645-59.
56. **Tokuda, H., and S. Matsuyama.** 2004. Sorting of lipoproteins to the outer membrane in *E. coli*. *Biochim Biophys Acta* **1693**:5-13.

57. **Tullman-Ercek, D., M. P. Delisa, Y. Kawarasaki, P. Iranpour, B. Ribnicky, T. Palmer, and G. Georgiou.** 2007. Export pathway selectivity of *Escherichia coli* twin-arginine translocation signal peptides. *J Biol Chem.* **282**:8309-16.
58. **van der Laan, M., N. P. Nouwen, and A. J. Driessen.** 2005. YidC--an evolutionary conserved device for the assembly of energy-transducing membrane protein complexes. *Curr Opin Microbiol* **8**:182-7.
59. **Voulhoux, R., G. Ball, B. Ize, M. L. Vasil, A. Lazdunski, L. F. Wu, and A. Filloux.** 2001. Involvement of the twin-arginine translocation system in protein secretion via the type II pathway. *EMBO J* **20**:6735-41.
60. **Widdick, D. A., K. Dilks, G. Chandra, A. Bottrill, M. Naldrett, M. Pohlschroder, and T. Palmer.** 2006. The twin-arginine translocation pathway is a major route of protein export in *Streptomyces coelicolor*. *Proc Natl Acad Sci U S A.* **103**:17927-32.
61. **WorldHealthOrganization** 2007, posting date. WHO Information tuberculosis fact sheet. [Online.]
62. **Wu, C. H.** 1996. Biosynthesis of lipoproteins, p. 1005-1014. *In* F. C. Neidhart (ed.), *Escherichia coli* and *Salmonella typhimurium*. ASM Press, Washington, DC.
63. **Yahr, T. L., and W. T. Wickner.** 2001. Functional reconstitution of bacterial Tat translocation in vitro. *EMBO J* **20**:2472-9.
64. **Yen, M. R., Y. H. Tseng, E. H. Nguyen, L. F. Wu, and M. H. Saier, Jr.** 2002. Sequence and phylogenetic analyses of the twin-arginine targeting (Tat) protein export system. *Arch Microbiol* **177**:441-50.
65. **Young, D. B., and T. R. Garbe.** 1991. Lipoprotein antigens of *Mycobacterium tuberculosis*. *Res Microbiol* **142**:55-65.

CHAPTER 2

The Twin-arginine Translocation (Tat) Pathway of *Mycobacterium smegmatis* is Functional and Required for the Export of Mycobacterial β -Lactamases

Justin A. McDonough¹, Kari E. Hacker¹, Anthony R. Flores², Martin S. Pavelka, Jr.², and Miriam Braunstein¹

¹Department of Microbiology and Immunology,
University of North Carolina School of Medicine, Chapel Hill, North Carolina

²Department of Microbiology and Immunology,
University of Rochester School of Medicine and Dentistry, Rochester, New York

The twin-arginine translocation (Tat) pathway exports folded proteins across the bacterial cytoplasmic membrane and is responsible for the proper extracytoplasmic localization of proteins involved in a variety of cellular functions, including pathogenesis. The *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* genomes contain open reading frames with homology to components of the Tat export system (TatABC) as well as potential Tat-exported proteins possessing N-terminal signal sequences with the characteristic twin-arginine motif. Due to the importance of exported virulence factors in the pathogenesis of *M. tuberculosis* and the limited understanding of mycobacterial protein export systems, we sought to determine the functional nature of the Tat export pathway in mycobacteria. Here we describe phenotypic analyses of $\Delta tatA$ and $\Delta tatC$ deletion mutants of *M. smegmatis* which demonstrated that *tatA* and *tatC* encode components of a functional Tat system capable of exporting characteristic Tat substrates. Both mutants displayed a growth defect on agar medium and hypersensitivity to SDS. The mutants were also defective in the

export of active β -lactamases of *M. smegmatis* (BlaS) and *M. tuberculosis* (BlaC), both of which possess twin-arginine signal sequences. The Tat-dependent nature of BlaC was further revealed by mutation of the twin-arginine motif. Finally, we demonstrated that replacement of the native signal sequence of BlaC with the predicted Tat signal sequences of *M. tuberculosis* phospholipase C proteins (PlcA and PlcB) resulted in the Tat-dependent export of an enzymatically active 'BlaC. Thus, 'BlaC can be used as a genetic reporter for Tat-dependent export in mycobacteria.

Introduction

Tuberculosis (TB), caused by the bacillus *Mycobacterium tuberculosis*, remains a global health problem accounting for over two million deaths annually (23). The ability to survive in host phagocytes is a poorly understood virulence property of *M. tuberculosis*. The *M. tuberculosis* proteins exported to the bacterial cell surface or secreted by the bacillus are ideally positioned to interact with host cell components and include virulence factors that promote intracellular survival.

Our understanding of how exported mycobacterial proteins are transported beyond the cytoplasmic membrane is limited (42). In bacteria, the well-characterized general secretion (Sec) pathway transports unfolded proteins across the cytoplasmic membrane to their final destination (for reviews, see 19, 56). Proteins targeted to this system contain specific N-terminal signal sequences (ss) that consist of a positively-charged region followed by a stretch of hydrophobic residues and end with a short uncharged polar region. Upon export, the signal sequence is cleaved yielding the mature protein. The Sec pathway is also

involved in the insertion of some integral membrane proteins and in these cases the signal sequences are not always cleaved (18). Numerous Sec proteins are required for the operation of this pathway including SecA, which provides the energy to move unfolded proteins through the Sec pore (19, 56). Mycobacteria possess a functional Sec pathway and the unusual property of possessing two SecA proteins (11, 12).

Recent studies reveal that many bacteria also utilize a twin-arginine translocation (Tat) pathway to transport proteins across the cytoplasmic membrane. The Tat pathway operates independently of the Sec pathway and is distinguished by the ability to translocate proteins in a folded state. Although studies in *Escherichia coli* provide the majority of information regarding the Tat pathway (for reviews, see 13, 59), the list of prokaryotic organisms experimentally shown to possess a functional Tat export pathway is expanding and includes Gram-positive bacteria such as *Bacillus subtilis* and *Streptomyces* spp., (40, 66, 67). In addition, the Tat pathway was shown to function in some bacterial pathogens and to contribute to virulence in *Pseudomonas aeruginosa*, *Agrobacterium tumefaciens*, enterohemorrhagic *E. coli* O157:H7, and *Legionella pneumophila* (22, 55, 60, 79).

Substrates of the Tat pathway are synthesized as precursor proteins containing N-terminal signal sequences with the same overall structure as Sec signal sequences. The major distinction between Sec and Tat signal sequences is the presence of a characteristic twin-arginine motif in the charged region of Tat signal sequences, defined as R-R-x- ϕ - ϕ (ϕ = uncharged residue) (4, 6). The importance of the twin arginines 'RR' to Tat export was shown by conservative substitution with twin lysines 'KK' in Tat signal sequences, which eliminates the export of many native Tat substrates and Tat-dependent reporter proteins (17, 30, 73). However, the requirement for twin arginines in a Tat-exported protein is not

absolute since there are reported examples of Tat substrates that lack a characteristic ‘RR’ signature (33, 36). Additional elements such as the degree of hydrophobicity in the region following the R-R-x- ϕ - ϕ motif and the presence of a positively-charged residue near the end of the signal sequence may also play a role in Tat-dependent export (17).

In *E. coli*, the Tat translocation apparatus is composed of the TatA, TatB, and TatC proteins (63). *E. coli* has an additional TatA homologue, designated TatE, which has overlapping functions with TatA. TatA and TatB are similar in structure, with each having a single transmembrane domain and a cytoplasmic tail (64). A number of GC-rich Gram-positive bacteria lack *tatB* (21), and in these bacteria TatA likely fulfills the function of both TatA and TatB. TatC is a more complex protein containing four to six transmembrane domains (9, 28). Current models suggest that TatC and TatB are involved in the initial recognition and delivery of Tat substrates to the protein conducting channel, composed of an oligomer of TatA (5, 27). In all reported cases, deletions of *tatA*, *tatB*, or *tatC* yield viable mutants with defects in the export of Tat signal sequence-containing proteins (9, 22, 40, 63, 66, 79).

We describe here analysis of the Tat pathway in the fast-growing nonpathogen *M. smegmatis*, often used as a genetic model for investigating physiologic properties of mycobacteria. We demonstrate that the Tat pathway is functional in *M. smegmatis*. Furthermore, we show that an intact Tat system is required for optimal growth of *M. smegmatis in vitro* and for the export of active mycobacterial β -lactamases. We also demonstrate the potential for using a truncated β -lactamase (‘BlaC’) lacking its native Tat signal sequence as a reporter for Tat-dependent exported proteins, including virulence factors of *M. tuberculosis*.

Materials and Methods

Bacterial strains and culture methods.

The strains used during this work are listed in Table 2.1. *E. coli* strains DH5 α , TOP10 (Invitrogen), EP1300 (Epicentre), and EZ (Qiagen) were used for DNA cloning procedures. Luria-Bertani (LB) medium (Fisher) was used for culturing of *E. coli* and antibiotics were added at the following concentrations: 150 $\mu\text{g/ml}$, hygromycin B (Roche Applied Science); 100 $\mu\text{g/ml}$, ampicillin; or 40 $\mu\text{g/ml}$, kanamycin (Acros Chemicals). Middlebrook 7H9 or 7H10 medium (Difco; BD Biosciences) was used for the culturing of *M. smegmatis* and *M. tuberculosis*. For growth of *M. smegmatis*, Middlebrook medium was supplemented with 0.5% glycerol and 0.2% dextrose. For growth of *M. tuberculosis*, Middlebrook medium was supplemented with 0.5% glycerol and 1 \times ADS (0.5% bovine serum albumin, fraction V [Roche]; 0.2% dextrose; and 0.85% NaCl). When necessary, media was supplemented with 0.05-0.1% Tween-80 (Fisher). Antibiotics were added to Middlebrook media at the following concentrations: 50 $\mu\text{g/ml}$, hygromycin B; 50 $\mu\text{g/ml}$, carbenicillin; or 20 $\mu\text{g/ml}$, kanamycin. When necessary, sucrose was added to 7H10 plates at a concentration of 4.5%. L-lysine at 40 and 80 $\mu\text{g/ml}$ was added to agar and liquid media, respectively, for growth of *M. smegmatis* strains PM759 and JM578.

Molecular biology procedures.

Standard molecular biology techniques were employed as previously described (61). The Expand high fidelity PCR system (Roche) was used in all PCR reactions from chromosomal DNA and dimethyl sulfoxide at 5.0% was included in select PCR reactions.

TABLE 2.1. Strains used in this study

Strain	Description	Reference
<i>E. coli</i>		
K-12 DH5 α	F- [ϕ 80d Δ lacZM15] Δ (lacZYA-argF) U169 deoR recA1 endA1 hsdR17 glnV44 thi-1 gyrA96 relA1	Gibco-BRL
TOP10	F- mcrA [ϕ 80d Δ lacZM15] Δ (mrr-hsdRMS-mcrBC) Δ lacX74 deoR recA1 endA1 araD139 Δ (ara, leu)7697 galU galK rpsL nupG	Invitrogen
EP1300	F- mcrA [ϕ 80d Δ lacZM15] Δ (mrr-hsdRMS-mcrBC) Δ lacX74 recA1 endA1 araD139 Δ (ara, leu)7697 galU galK λ - rpsL nupG trfA	Epicentre
Qiagen EZ	[F':::Tn10(Tc ^r) proA ⁺ B ⁺ lacI ^q Z Δ M15] recA1 end A1 hsdR17 (Γ_{K12} - m _{K12} ⁺) lac glnV44 thi-1 gyrA96 relA1	Qiagen
<i>M. smegmatis</i>		
mc ² 155	<i>ept-1</i>	(70)
MB679	mc ² 155::pKEH4, <i>tatA</i> single-crossover strain	This work
JM499	mc ² 155::pJM112, <i>tatC</i> single-crossover strain	This work
MB692	mc ² 155, Δ <i>tatA</i>	This work
JM567	mc ² 155, Δ <i>tatC</i>	This work
PM759	mc ² 155, Δ <i>blaS1</i> Δ <i>lysA4</i> <i>rpsL6</i>	(25)
JM575	PM759::pKEH4, <i>tatA</i> single-crossover strain	This work
JM578	PM759, Δ <i>tatA</i>	This work
<i>M. tuberculosis</i>		
H37Rv	Virulent	(2)
PM638	H37Rv, Δ <i>blaC1</i>	(25)

DNA sequencing was performed by the UNC-CH Automated DNA Sequencing Facility.

Construction of Δ *tat* suicide plasmids. (i) Δ *tatA* suicide plasmid pKEH4.

Genomic sequence immediately upstream and downstream of the *tatA* gene, including the 5' and 3' termini of *tatA* were amplified from *M. smegmatis* genomic DNA using the following primers: *tatA*Upstream1 (5'- AGATATCCTCCGCAGGCGTGGATCAT), *tatA*Upstream2 (5'- ACATATGGGTTACCTCCAGAACGTCGGTG), *tatA*Downstream1 (5'- TCATATGCAGTCCACGGACCGGCACAC), and *tatA*Downstream2 (5'- CGACTTCCAGCAGCACCGTCA) (underlined residues denote the *NdeI-XbaI* restriction sites). The PCR products were cloned into the pCR2.1 vector (Invitrogen) to generate pKEH1 and pKEH2. The *NdeI-XbaI* fragment from pKEH2 containing the downstream *tatA*

flank was cloned into *NdeI-XbaI* cut pKEH1. The resulting plasmid, pKEH3, contained a 221-bp unmarked in-frame deletion of *tatA*. The 1.7-kbp *tatA* deletion fragment was then cut out of pKEH3 by *EcoRV* and cloned into the *EcoRV* site on counterselectable suicide vector pYUB657 (51), generating pKEH4. DNA inserts in the above plasmids were sequenced and determined to be error-free.

(ii) Δ *tatC* suicide plasmid pJM112. Genomic sequence immediately upstream and downstream of the *tatC* gene, including the 5' and 3' termini of *tatC* were amplified from *M. smegmatis* genomic DNA using the following primers: SmTatC1 (5'-GCGTGGATGTTCTCGACTACTACC), SmTatC2 (5'-AAGCTTGTTGCGAAGCTCGTGGAGGTG), SmTatC3 (5'-AAGCTTCTCAACGACCGCAGGAAGGC), and SmTatC4 (5'-GCAGGTGCAGGATGACCTCTTC) (underlined residues denote the *HindIII* restriction sites). The PCR products were cloned into the pCC1 vector using the CopyControl PCR cloning kit (Epicentre) to generate pJM104 and pJM105. A 753-bp *HindIII* fragment from pJM104 containing the upstream *tatC* flank was cloned into pJM105 cut with *HindIII*. The resulting plasmid, pJM110, contained a 690-bp unmarked in-frame deletion of the *tatC* gene. The 1.3-kbp *tatC* deletion fragment was then cut out of pJM110 by *BamHI* and cloned into the *BglII* site on counterselectable suicide vector pYUB657 (51), producing pJM112. DNA inserts were sequenced and determined to be error-free.

Two-step allelic exchange to create Δ *tat* mutants of *M. smegmatis*.

To construct the Δ *tatA* mutants, pKEH4 was electroporated into mc²155 as described previously (50, 51) and hygromycin-resistant transformants were selected. Individual

transformants were subjected to Southern analysis and shown to contain the suicide vector integrated into the *tatA* region of the chromosome by means of a single-crossover event. One of these strains, MB679, was employed in the subsequent steps. For the second homologous recombination event, MB679 was grown to saturation in 7H9 medium with hygromycin. This culture was then subcultured at a 1:100 dilution into the same medium lacking hygromycin and incubated overnight. Dilutions were plated onto 7H10 plates containing 4.5% sucrose to select for sucrose-resistant colonies. Sucrose-resistant strains were patched onto 7H10 plates containing hygromycin. Strains that were both sucrose-resistant and hygromycin-sensitive were analyzed by PCR and Southern analysis for a deletion at the *tatA* locus. The same procedure was performed in the $\Delta blaS$ PM759 background to generate a $\Delta blaS \Delta tatA$ deletion strain. To obtain $\Delta tatC$ mutants, pJM112 was electroporated into mc²155 and hygromycin-resistant transformants were confirmed to be the result of a single-crossover event by Southern analysis. One of these clones, JM499, was used for the subsequent work presented in this report. The second homologous recombination event was selected by using the protocol described above for MB679.

Southern analysis.

To analyze *tatA* recombinants, genomic DNA was isolated from *M. smegmatis* strains as previously described (10) and digested with *NdeI-FspI*. The probe used was a 910-bp *SalI* fragment of pJM120 containing *tatC* of *M. smegmatis*. To analyze *tatC* recombinants, genomic DNA was isolated and digested with *FspI*. The probe used was a 753-bp *HindIII* fragment obtained from pJM104 that contains *tatA* of *M. smegmatis*. Southern analysis was

performed as previously described (61), and probes were labeled with [³²P]dCTP using the Ready-to-Go Labeling kit (Amersham).

Antimicrobial susceptibility testing.

The replica-inoculating MIC method used was adapted from a previously described protocol (71). Mid-exponential-phase *M. smegmatis* cultures were diluted ten-fold in 7H9 liquid medium and spotted using a multi-pronged replica-inoculating device onto 7H10 agar plates containing 2-fold dilutions of carbenicillin. Culture dilutions containing an estimated 10² to 10⁴ colony-forming units (CFU) of *M. smegmatis* resulted in confluent growth on agar plates lacking carbenicillin. For plates containing antibiotic, the lowest concentration of carbenicillin that inhibited growth completely after incubation for four days was designated the minimum inhibitory concentration (MIC). Each MIC determination was performed in triplicate. For determining sensitivity to sodium dodecyl sulfate (SDS), mid-exponential-phase *M. smegmatis* cultures were diluted ten-fold in 7H9 liquid medium and spotted using a multi-pronged replica-inoculating device onto 7H10 agar plates containing 0.0025% SDS.

Preparation of culture filtrates and whole-cell extracts from *M. smegmatis* for nitrocefin analysis.

To obtain cell-free culture filtrates, exponential-phase *M. smegmatis* cultures were diluted to an optical density (OD₆₀₀) = 0.1 and allowed to grow at 37°C for two generations. When the culture reached an OD₆₀₀ = 0.4, 500 µl was applied to a 0.2 µm NANOSEP centrifugal filter (Pall Filtron) and the culture filtrate was harvested by centrifugation at 13,000 rpm. Alternatively, culture filtrates were harvested using a 0.22 µm Millex-GV

syringe-driven filter unit (Millipore). Whole cell extracts were prepared from the remaining culture by harvesting 10 ml of culture by centrifugation and the pellets stored at -20°C. At a later time, the pellets were resuspended in 200 µl of 1×PBS with protease inhibitors (10 µg/ml aprotinin, 10 µg/ml E-64, 10 µg/ml leupeptin, 500 µg/ml Pefabloc SC, 10 µg/ml pepstatin A) to which 0.4 g of 106 µm glass beads (Sigma) was added. The cells and beads were then vortexed twice for 5 min at 4°C with a 5 min rest on ice. 200 µl of 1×PBS with protease inhibitors was added, followed by vortexing, and centrifugation for 10 minutes at 4°C. Culture filtrates and cell extracts were assayed for total protein content using the bicinchoninic acid (BCA) protein quantification kit (Pierce).

Nitrocefin assays.

The β-lactamase activity of *M. smegmatis* whole-cell extracts and culture filtrates was determined using the chromogenic β-lactam nitrocefin (46), as previously described (25). Assays were performed on triplicate wells set up from a single culture, and β-lactamase activity, expressed as $A_{486} \text{ min}^{-1} \text{ mg total protein}^{-1}$, was reported as an average with standard deviation. Each experiment was repeated three times and a representative experiment is presented.

Construction of '*blaC* fusion plasmids.

(i) *ss-plcA*. Genomic sequence encoding the predicted signal sequence of *M. tuberculosis plcA* was amplified from *M. tuberculosis* genomic DNA using the following primers: *plcA-1* (5' - TGGCCAGC-CGTCGAGAGTTTTGACAAAG) (underlined residues denote the *MscI* restriction site) and *plcA-3* (5' - GGATC-CGCCGTAGGCCTTTTCAATCA)

(underlined residues denote the *Bam*HI restriction site). The 104-bp PCR product was cloned into the pCR2.1 vector (Invitrogen) to generate pJM128. A 96-bp *Msc*I-*Bam*HI fragment from pJM128 was cloned into pMV261 cut with *Msc*I-*Bam*HI. The resulting plasmid, pJM129, contained the predicted signal sequence of *plcA* cloned downstream of the *hsp60* promoter.

(ii) *ss-plcB*. Genomic sequence encoding the predicted signal sequence of *M. tuberculosis plcB* was amplified from *M. tuberculosis* genomic DNA using the following primers: *plcB-1* (5'- ATGGCC-ACCCGCCGACAATTTTTTGC) (underlined residues denote the *Msc*I restriction site) and *plcB-3* (5'- GGATCC-TCCGTAGGCTTTTTTCGATA) (underlined residues denote the *Bam*HI restriction site). The 104-bp PCR product was cloned into the pCR2.1 vector (Invitrogen) to generate pMB221. A 104-bp *Msc*I-*Bam*HI fragment from pMB221 was cloned into pMV261 cut with *Msc*I-*Bam*HI. The resulting plasmid, pMB222, contained the predicted signal sequence of *plcB* cloned downstream of the *hsp60* promoter.

(iii) *ss-mpt63*. Genomic sequence containing the 5' region of the *mpt63/Rv1926c* gene was amplified from *M. tuberculosis* genomic DNA using the following primers: *Rv1926c-5* (5'- CTGCGCAT-GAAGCTCACCACAATGATCAAG) (underlined residues denote the *Fsp*I restriction site) and *Rv1926c-6* (5'- AGGATCC-AGCCAACGCGACCGGTGCCGCAAAG) (underlined residues denote the *Bam*HI restriction site). The 94-bp PCR product was cloned into the pDRIVE vector (Qiagen) to generate pMB226. A 94-bp *Fsp*I-*Bam*HI fragment from pMB226 was cloned into pMV261 cut with *Msc*I-*Bam*HI. The resulting plasmid, pMB228, contained the predicted signal sequence of *mpt63* cloned downstream of the *hsp60* promoter.

(iv) '*blaC*. Sequence encoding a truncated version of *M. tuberculosis blaC* lacking its predicted signal sequence ('*blaC*) was amplified from pMP159 (25) using the following

primers: *blaCfor1* (5'- AGATCTGCG-CGTCCGGCATCGACAACCTTGC) and *blaCrev* (5'- AGATCT-CCACGAGCCTATGCAAGCACAC) (underlined residues denote the *Bgl*II restriction sites). The 854-bp PCR product was cloned into the pCC1 vector (Epicentre) to generate pJM106. A 848-bp *Bgl*II fragment from pJM106 was end-filled with Klenow and cloned into pMV261 cut with *Msc*I. The resulting plasmid, pJM113, contained the predicted mature sequence of *blaC* cloned downstream of the *hsp60* promoter. All other plasmids used in this work are described in Table 2.2. Translational fusion plasmids were sequenced and shown to be error-free.

Site-directed mutagenesis of '*blaC* fusion plasmids.

Site-directed mutagenesis was employed to replace the twin-arginine residues of the *M. tuberculosis* PlcB signal sequence with twin-lysine residues. pJM111 was used as the template in an inverse PCR using primers plcB-RR2KKfor (5'- AAGCAATTTTTTGCCAAAGCCGCC) and plcB-RR2KKrev (5'- CTTGGTGGCCATTGCGAAGTGATT) (underlined residues denote mutated bases). The resulting PCR product was gel purified, 5'-phosphorylated using the End-It DNA End-Repair Kit (Epicentre), and self-ligated. The resulting plasmid, pJM118, was sequenced to verify the presence of the mutated bases. Likewise, to replace the twin-arginine residues of the BlaC signal sequence with twin-lysine residues, pMP327 was used as the template in an inverse PCR using primers blaC-RR2KKfor (5'- AAGGAACTGCTGGTAGCGATGGCAATG) and blaC-RR2KKrev (5'- CTTACCGAATCCTCTGTTGCGCATGCC). The resulting plasmid, pJM117, was sequenced to verify the presence of the mutated bases.

TABLE 2.2. Plasmids used in this study

Plasmid	Genotype	Description	Source
pCC1	<i>cat oriV ori2</i>	CopyControl (single copy) blunt cloning vector	Epicentre
pCR2.1	<i>bla aph ColE1</i>	TA cloning vector	Invitrogen
pDRIVE	<i>bla aph ColE1</i>	UA cloning vector	Qiagen
pMV261.kan	<i>aph P_{hsp60} oriM ColE1</i>	Multicopy mycobacterial shuttle plasmid	(75)
pJSC77	<i>aph P_{hsp60} -HA oriM ColE1</i>	HA tag cloned into pMV261	(26)
pMB198	<i>hyg int attP ColE1</i>	Single-copy mycobacterial shuttle plasmid	(11)
pYUB657	<i>bla hyg P_{hsp60}-sacB ColE1</i>	Counterselectable suicide plasmid for mycobacteria	(51)
pJM104	<i>cat oriV ori2</i>	<i>M. smegmatis tatC</i> upstream flank cloned into pCC1	This work
pJM105	<i>cat oriV ori2</i>	<i>M. smegmatis tatC</i> downstream flank cloned into pCC1	This work
pJM110	<i>cat ΔtatC (M. smegmatis) oriV ori2</i>	<i>tatC</i> upstream flank from pJM104 cloned adjacent to <i>tatC</i> downstream flank in pJM105	This work
pJM112	<i>bla hyg P_{hsp60}-sacB ΔtatC (M. smegmatis) ColE1</i>	1.3-kbp <i>Bam</i> HI fragment containing <i>ΔtatC</i> from pJM110 cloned into pYUB657	This work
pKEH1	<i>bla aph ColE1</i>	<i>M. smegmatis tata</i> upstream flank cloned into pCR2.1	This work
pKEH2	<i>bla aph ColE1</i>	<i>M. smegmatis tata</i> downstream flank cloned into pCR2.1	This work
pKEH3	<i>bla aph ColE1 Δtata (M. smegmatis)</i>	<i>tata</i> downstream flank from pKEH2 cloned adjacent to <i>tata</i> upstream flank in pKEH1	This work
pKEH4	<i>bla hyg P_{hsp60}-sacB Δtata (M. smegmatis) ColE1</i>	1.7-kbp <i>Eco</i> RV fragment containing <i>Δtata</i> from pKEH3 cloned into pYUB657	This work
pJM119	<i>cat tatC (M. smegmatis) oriV ori2</i>	<i>M. smegmatis tatC</i> ORF cloned into pCC1	This work
pJM120	<i>aph P_{hsp60}-tatC-HA (M. smegmatis) oriM ColE1</i>	<i>M. smegmatis tatC</i> ORF from pJM119 cloned into pJSC77; HA-tagged <i>M. smegmatis tatC</i> under control of <i>hsp60</i> promoter	This work
pMB234	<i>bla aph tata (M. smegmatis) ColE1</i>	<i>M. smegmatis tata</i> ORF cloned into pCR2.1	This work
pMB236	<i>aph P_{hsp60}-tata (M. smegmatis) oriM ColE1</i>	<i>M. smegmatis tata</i> ORF from pMB234 cloned into pMV261; <i>M. smegmatis tata</i> under control of <i>hsp60</i> promoter	This work
pJM124	<i>hyg P_{hsp60}-tata (M. smegmatis) int attP ColE1</i>	<i>P_{hsp60}-tata</i> from pMB236 cloned into pMB198	This work
pMP159	<i>bla blaC (M. tuberculosis) ColE1</i>	Fragment containing <i>blaC</i> from MTCY49 in pKS+	(25)
pMP327	<i>aph P_{blaC}-blaC (M. tuberculosis) oriM ColE1</i>	<i>M. tuberculosis blaC</i> from pMP159 cloned into pMV261; <i>M. tuberculosis blaC</i> under control of endogenous promoter	This work
pJM106	<i>cat 'blaC (M. tuberculosis) oriV ori2</i>	Predicted <i>M. tuberculosis 'blaC</i> mature sequence cloned into pCC1	This work
pMM7-2	<i>aph P_{hsp60}-ssfbpB (M. tuberculosis) oriM ColE1</i>	<i>M. tuberculosis fbpB</i> signal sequence in pMV261 under control of <i>hsp60</i> promoter	(82)
pJM109	<i>aph P_{hsp60}-ssfbpB-'blaC (M. tuberculosis) oriM ColE1</i>	' <i>blaC</i> from pJM106 cloned into pMM7-2	This work
pJM128	<i>bla aph ColE1</i>	Predicted <i>M. tuberculosis plcA</i> signal sequence cloned into pCR2.1	This work
pJM129	<i>aph P_{hsp60}-ssplcA (M. tuberculosis) oriM ColE1</i>	<i>M. tuberculosis plcA</i> signal sequence from pJM128 cloned into <i>MscI</i> - <i>Bam</i> HI cut pMV261	This work
pJM130	<i>aph P_{hsp60}-ssplcA-'blaC (M. tuberculosis) oriM ColE1</i>	' <i>blaC</i> from pJM106 cloned into pJM129	This work
pMB221	<i>bla aph ColE1</i>	Predicted <i>M. tuberculosis plcB</i> signal sequence cloned into pCR2.1	This work
pMB222	<i>aph P_{hsp60}-ssplcB (M. tuberculosis) oriM ColE1</i>	<i>M. tuberculosis plcB</i> signal sequence from pMB221 cloned into <i>MscI</i> - <i>Bam</i> HI cut pMV261	This work

TABLE 2.2. (Continued) Plasmids used in this study

pJM111	<i>aph P_{hsp60}-ssplcB-'blaC (M. tuberculosis) oriM</i> ColE1	' <i>blaC</i> from pJM106 cloned into pMB222	This work
pMB226	<i>bla aph</i> ColE1	<i>M. tuberculosis mpt63</i> signal sequence cloned into pDRIVE	This work
pMB227	<i>aph P_{hsp60}-ssmpt63 (M. tuberculosis) oriM</i> ColE1	<i>M. tuberculosis mpt63</i> signal sequence from pMB226 cloned into <i>MscI-BamHI</i> cut pMV261	This work
pMB228	<i>aph P_{hsp60}-ssmpt63-'blaC (M. tuberculosis) oriM</i> ColE1	' <i>blaC</i> from pJM106 cloned into pMB227	This work
pJM113	<i>aph P_{hsp60}-'blaC (M. tuberculosis) oriM</i> ColE1	' <i>blaC</i> from pJM106 cloned into <i>MscI</i> -linearized pMV261	This work
pJM117	<i>aph P_{blaC}-(KK)blaC (M. tuberculosis) oriM</i> ColE1	<i>M. tuberculosis ssblaC(KK)</i> under control of endogenous promoter	This work
pJM118	<i>aph P_{hsp60}-ssplcB(KK)-'blaC (M. tuberculosis) oriM</i> ColE1	<i>M. tuberculosis ssplcB(KK)-'blaC</i> under control of <i>hsp60</i> promoter	This work

Nucleotide sequence accession numbers.

The DNA sequence of a 2018 bp cloned product containing the *M. smegmatis tata* and *tatC* ORFs was submitted to GenBank and given the accession number AY998985.

Results

Identification of a putative Tat system in mycobacteria.

The genome of *M. tuberculosis* strain H37Rv contains two distinct loci with open reading frames (ORFs) homologous to components of the bacterial Tat system. *M. tuberculosis* TatA (Rv2094c), TatB (Rv1224), and TatC (Rv2093c) are 44%, 38%, and 43% similar at the amino-acid level to TatA, TatB, and TatC of *E. coli* (16). The proximity of *tatA* and *tatC* (17-bp between ORFs) suggests that these genes are located in an operon. The *tatB* gene is located approximately 1-Mbp away from *tatAC*. Using sequence data obtained from The Institute for Genomic Research (<http://www.tigr.org>) we identified homologues of *tatABC* in the genome of *M. smegmatis* strain mc²155. The organization of the *tat* genes is

conserved in all the available mycobacteria genome sequences including *M. tuberculosis*, *M. bovis*, *M. leprae*, *M. avium*, and *M. smegmatis* (<http://www.tigr.org>).

The *M. tuberculosis* genome also contains 31 ORFs with predicted Tat signal sequences, as identified by the TATFIND search algorithm (21). The TATFIND algorithm is based on experimentally demonstrated Tat substrates and sequences of putative Tat substrates in halophilic archaea. The list of TATFIND-predicted *M. tuberculosis* Tat substrates is available at the following URL:

<http://www.sas.upenn.edu/~pohlschr/tatprok.html>. Notable among the predicted *M. tuberculosis* Tat substrates are the virulence factor phospholipase C (PlcB) (58) and the major β -lactamase BlaC (25, 78). Our inspection of the amino-terminus of the major β -lactamase of *M. smegmatis*, BlaS (25, 44), also revealed a potential Tat signal sequence. The identification of putative Tat components and Tat-exported substrates in mycobacteria strongly suggested the existence of a functional Tat pathway.

***M. smegmatis* *tatA* and *tatC* deletion mutants have growth defects.**

To examine the functional nature of the predicted Tat pathway in mycobacteria, in-frame and unmarked deletion alleles of the *tatA* and *tatC* genes were generated in *M. smegmatis* using a two-step allelic exchange strategy (10, 50).

To construct the Δ *tatA* mutant, we electroporated the suicide plasmid pKEH4 containing an in-frame unmarked deletion within *M. smegmatis* *tatA* into wild-type *M. smegmatis* mc²155. This Δ *tatA* deletion allele is predicted to produce a truncated protein of 34 amino acids (in comparison to the wild type protein of 106 amino acids). Integration of pKEH4 at the *tatA* chromosomal locus via a single homologous recombination event yielded

the single-crossover strain MB679. This single-crossover strain contained both a wild-type and a deleted copy of *tatA* along with intervening pKEH4 vector sequence carrying the selectable hygromycin-resistance gene and the counterselectable *sacB* marker. Southern blot analysis confirmed successful construction of the single-crossover strain (Fig. 2.1).

We employed the same strategy to construct a single-crossover strain for generation of a Δ *tatC* mutant using the suicide plasmid pJM112, which contained an in-frame deletion within *M. smegmatis* *tatC*. This *tatC* deletion is predicted to produce a truncated protein of 87 amino acids (in comparison to the wild type protein of 317 amino acids) which lacks the six transmembrane domains. Southern blot analysis confirmed construction of the *tatC* single crossover strain JM499 (Fig. 2.1).

Our initial attempts to isolate secondary recombinants with *tat* deletions from the single-crossover strains yielded only wild-type alleles at both the *tatA* and *tatC* loci. This strongly suggested that under the growth conditions tested, a *tat* deletion in *M. smegmatis* was lethal and that the Tat pathway was essential for growth. However, further incubation for two additional days at 37°C yielded a population of slow-growing secondary recombinants. These small colonies were subsequently proven to be *tatA* or *tatC* deletion mutants by PCR (data not shown) and Southern blot analysis (Fig. 2.1). The Δ *tatA* and Δ *tatC* mutants represented 65% and 32% of the total secondary recombinants, respectively. All of the Δ *tatA* and Δ *tatC* mutants obtained formed smaller single colonies on both minimal 7H10 and rich LB agar media compared to wild-type *M. smegmatis* (Fig. 2.2a). The *M. smegmatis* Δ *tatA* mutant strain MB692 and Δ *tatC* mutant strain JM567 were used for all subsequent analyses (Table 2.1).

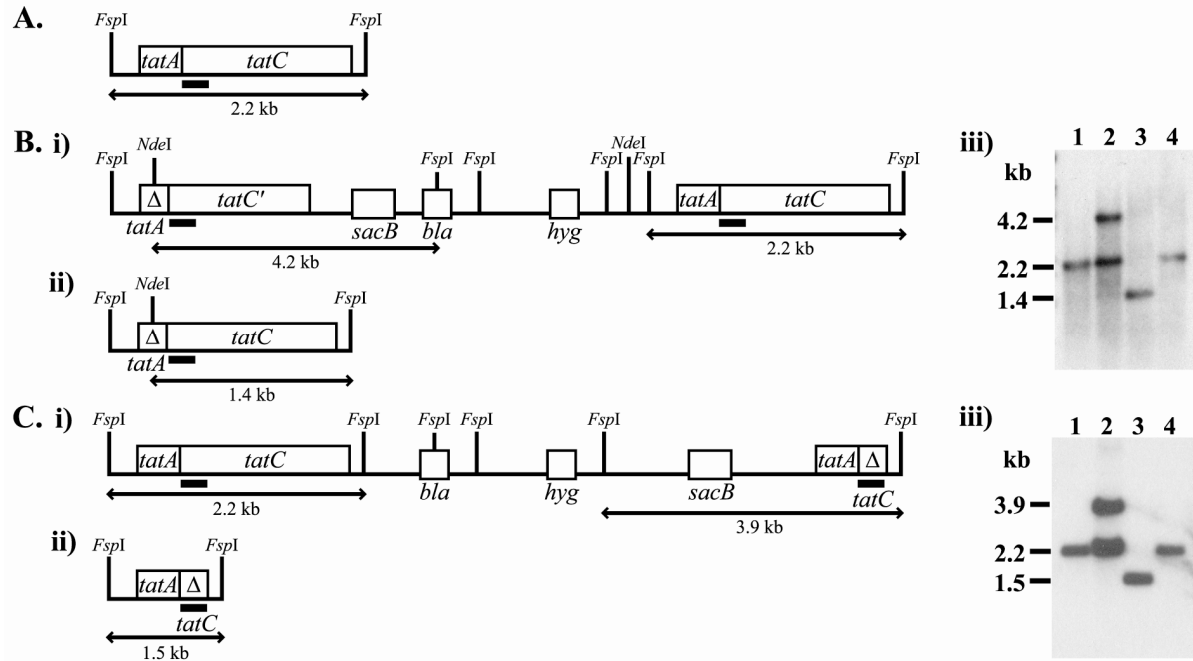


Fig. 2.1. Southern analysis of *tatA* and *tatC* recombinants in *M. smegmatis*. Panels: A, diagram of the mc²155 (wild-type) chromosomal *tatAC* locus; B, Construction of *M. smegmatis* Δ *tatA* allele. i) *tatA* single-crossover strain MB679, ii) Δ *tatA* allele, iii) Southern blot of *FspI/NdeI*-digested genomic DNA. Lanes: 1, mc²155; 2, MB679; 3, MB692 (Δ *tatA*); 4, wild-type recombinant. C, Construction of *M. smegmatis* Δ *tatC* allele. i) *tatC* single-crossover strain JM499, ii) Δ *tatC* allele, iii) Southern blot of *FspI*-digested genomic DNA. Lanes: 1, mc²155; 2, JM499; 3, JM567 (Δ *tatC*); 4, wild-type recombinant. The *FspI/NdeI* restriction endonuclease sites are indicated, as well as the size of the DNA fragments detected by hybridization with the probe used (bold line). Diagrams are not drawn to scale.

We compared growth of the Δtat mutants to that of wild-type *M. smegmatis* in Middlebrook 7H9 liquid medium. Growth of the mutants over time was determined by measuring the optical density (OD₆₀₀) and by plating for colony-forming units (CFU). The $\Delta tatC$ mutant had lower OD₆₀₀ readings than the wild-type strain throughout the growth curve (Fig. 2.2b). However, there was no difference in the number of counted CFU during liquid growth between the $\Delta tatC$ mutant and wild-type *M. smegmatis* strains (Fig. 2.2c). The same phenotypes were observed for the $\Delta tatA$ mutant (data not shown). Thus, the Δtat mutants appeared to replicate at the same rate as wild-type *M. smegmatis* in liquid medium, even though slower growing colonies were observed on agar medium. The basis for the discrepancy between OD₆₀₀ readings and viable CFU in culture is not clear.

Complementation analysis of the $\Delta tatA$ mutant (MB692) was performed with a single-copy plasmid pJM124 that expresses *M. smegmatis* *tatA* from the constitutive *hsp60* promoter. For the $\Delta tatC$ mutant (JM567) the multicopy plasmid pJM120, which constitutively expresses *M. smegmatis* *tatC* from the *hsp60* promoter, was used. Introduction of these *tat* expression constructs into the respective Δtat mutants successfully rescued the mutant phenotypes: small colony growth on agar and reduced OD₆₀₀ readings during liquid culture (Fig. 2.2). The complementation experiments demonstrated that the Δtat mutations are responsible for the observed phenotypes and they indicate that both TatA and TatC of *M. smegmatis* are functional and required for optimal *in vitro* growth.

***M. smegmatis* *tat* mutants are sensitive to SDS.**

Tat-deficient mutants of *E. coli* are hypersensitive to the detergent sodium dodecyl sulfate (SDS) (38, 72). To determine if the *M. smegmatis* $\Delta tatA$ and $\Delta tatC$ mutants were

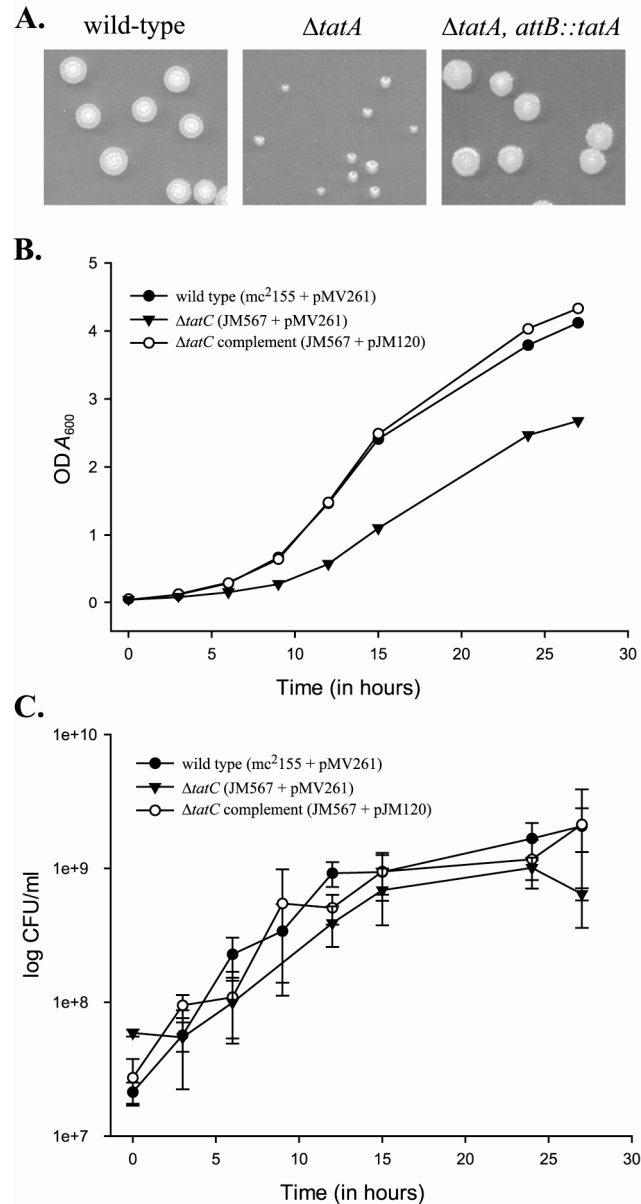


Fig. 2.2. The $\Delta tatA$ and $\Delta tatC$ mutants have growth defects. (A) Single colonies of wild-type (mc²155, pMB198), $\Delta tatA$ mutant (MB692, pMB198), and complemented $\Delta tatA$ *attB::tatA* (MB692, pJM124) strains are shown on 7H10 agar medium containing 0.1% tween-80, grown at 37°C. Not shown are the $\Delta tatC$ mutant and the complemented $\Delta tatC$ strain, which display the same growth phenotypes. (B) Representative growth curves for wild-type (mc²155, pMV261), $\Delta tatC$ mutant (JM567, pMV261), and complemented $\Delta tatC$ (JM567, pJM120) strains in 7H9 liquid medium containing 0.1% tween-80, grown at 37°C, is shown as the optical density at 600 nm. (C) Viable count measurements from the same cultures shown in (B) as the log₁₀ of viable CFU/ml.

similarly hypersensitive to SDS, we compared the ability of the mutant strains to grow on agar media containing SDS using a replica-inoculating protocol (described in Materials and Methods) (71). The $\Delta tatA$ and $\Delta tatC$ mutants failed to grow in the presence of 0.0025% SDS, in contrast to growth observed for the wild-type strain (Fig. 2.3). This was true even after extended incubation (data not shown). Growth of the Δtat mutants on 0.0025% SDS was restored to wild-type levels when complemented with Tat proteins expressed in trans. The failure of the Δtat mutants to grow on medium containing SDS was not a result of the general growth defect on agar medium, as the Δtat mutants grew on agar media lacking SDS, albeit slower (Fig. 2.3).

***M. smegmatis* tat mutants are sensitive to β -lactam antibiotics.**

Tat-deficient mutants of *E. coli* are also hypersusceptible to multiple antibiotics including the β -lactam ampicillin (72). We similarly tested the $\Delta tatA$ and $\Delta tatC$ mutants of *M. smegmatis* for drug susceptibility. β -lactam susceptibility was assayed by determining the minimal inhibitory concentration (MIC) of carbenicillin using a replica-inoculating protocol on 7H10 agar medium (described in Materials and Methods) (71). In contrast to wild-type *M. smegmatis* which is intrinsically resistant to β -lactam antibiotics and had an MIC of >200 $\mu\text{g/ml}$ for carbenicillin, the MIC for the $\Delta tatA$ and $\Delta tatC$ mutants was 5 $\mu\text{g/ml}$. The complemented strains were resistant to >200 $\mu\text{g/ml}$ carbenicillin, indicating that increased susceptibility to carbenicillin of the mutants was due to the Δtat mutations. Similar results were obtained using a disk diffusion assay to test susceptibility to the β -lactams carbenicillin, ampicillin, and amoxicillin (data not shown). In both the MIC and disk diffusion assays, the sensitivity of the Δtat mutants was comparable to the sensitivity of a $\Delta blaS$ mutant lacking

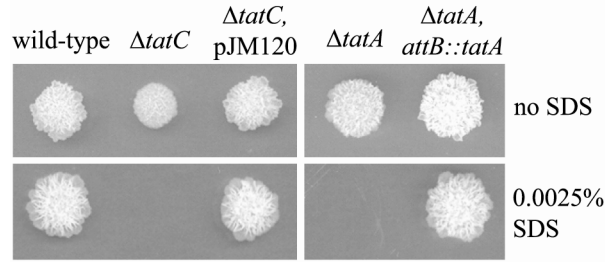


Fig. 2.3. The $\Delta tatA$ and $\Delta tatC$ mutants are hypersensitive to SDS. The following strains were grown in liquid media and approximately 10^3 CFU were spotted onto plates with and without 0.0025% SDS using a multi-pronged replica inoculating device as described in Materials and Methods: wild-type (*mc*²155, pMV261), $\Delta tatC$ mutant (JM567, pMV261), complemented $\Delta tatC$ strain (JM567, pJM120), $\Delta tatA$ mutant (MB692, pMB198), and complemented $\Delta tatA$ *attB::tatA* strain (MB692, pJM124).

the major *M. smegmatis* β -lactamase. In contrast to the results with the β -lactam antibiotics, the $\Delta tatA$ and $\Delta tatC$ mutants exhibited no increased sensitivity to the other antibiotics tested (5 μ g isoniazid, 25 μ g rifampin, 25 μ g ethambutol, 12.5 μ g streptomycin, 20 μ g tetracycline, and 20 μ g vancomycin). Thus, the β -lactam sensitive phenotype of the *M. smegmatis* Δtat mutants is a specific phenomenon and not the result of a general increase in drug susceptibility.

β -lactamase export requires the Tat pathway of *M. smegmatis*.

The major β -lactamases BlaC and BlaS are required for the β -lactam resistance of both *M. tuberculosis* and *M. smegmatis*, respectively (25, 44, 78). Since β -lactam antibiotics target extracytoplasmic cell wall biosynthetic enzymes, a β -lactamase must be exported to the cell envelope or extracellular environment in order to protect the bacterium. Inspection of the N-terminus of BlaC and BlaS revealed a twin-arginine motif, suggesting that both enzymes are exported by the Tat pathway. We hypothesized that the β -lactam sensitivity phenotype of the $\Delta tatA$ and $\Delta tatC$ mutants of *M. smegmatis* was due to a defect in export of the putative Tat substrate BlaS.

It was previously reported that during liquid growth *M. smegmatis* releases β -lactamase into the medium (57). To test whether *M. smegmatis* β -lactamase export depends upon the Tat pathway, we assayed short-term cell-free culture filtrates from the Δtat mutants for β -lactamase activity using the chromogenic β -lactam nitrocefin.

Substantially less β -lactamase activity was observed in culture filtrates produced from the $\Delta tatA$ and $\Delta tatC$ mutants compared to wild-type *M. smegmatis* (Fig. 2.4a). The Δtat

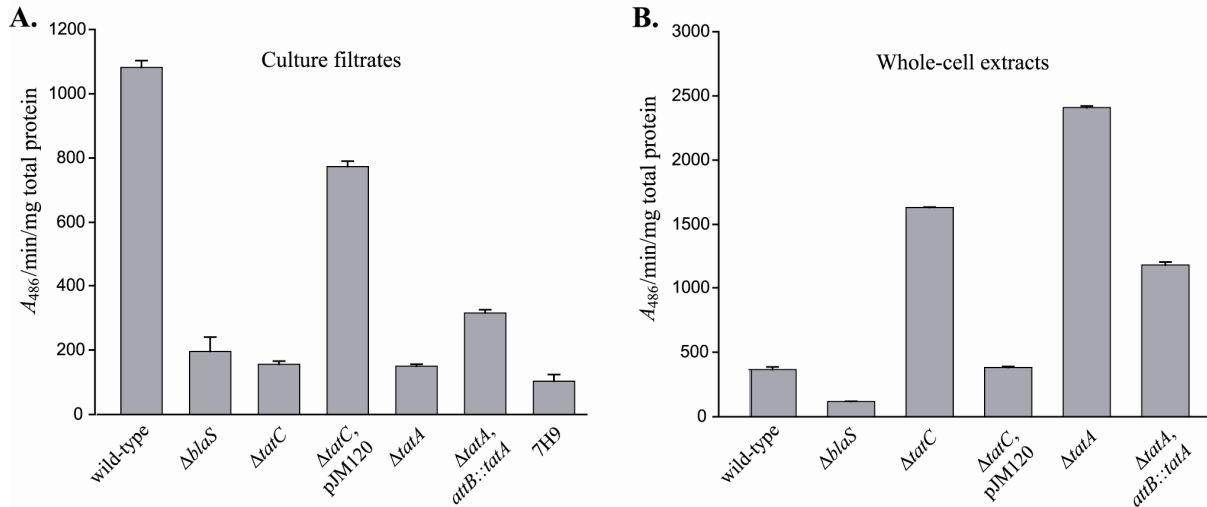


Fig. 2.4. Nitrocefin assays on culture-filtrates and whole-cell extracts of *M. smegmatis* wild-type and $\Delta tatA$ and $\Delta tatC$ mutants. Short-term culture filtrates (A) and whole-cell extracts (B) of log-phase *M. smegmatis* cultures were assayed for the ability to hydrolyze nitrocefin. Strains used: wild-type (mc^2155 , pMV261), $\Delta blaS$ (PM759, pMV261), $\Delta tatC$ mutant (JM567, pMV261), complemented $\Delta tatC$ strain (JM567, pJM120), $\Delta tatA$ mutant (MB692, pMB198), complemented $\Delta tatA$ *attB::tatA* strain (MB692, pJM124), and liquid culture medium (7H9). Nitrocefin hydrolyzing activity was expressed as $A_{486} \text{ min}^{-1} \text{ mg total protein}^{-1}$, and was standardized to PBS. Shown is a representative assay reported as an average \pm standard deviation.

mutant levels were comparable to that observed with the $\Delta blaS$ mutant. The low level of nitrocefin hydrolyzing activity detected in the $\Delta blaS$ and Δtat strains was largely attributable to background from the 7H9 medium (Fig. 2.4a). To rule out the possibility that the reduced β -lactamase activity in culture filtrates of the Δtat mutants was due to reduced expression, we measured the nitrocefin hydrolyzing activity of whole-cell extracts of the strains. This analysis revealed β -lactamase expression with more cell-associated activity for the Δtat mutants in comparison to the $\Delta blaS$ mutant (Fig. 2.4b). Further, a five-fold higher level of β -lactamase activity was associated with whole-cell extracts of the Δtat mutants compared to the wild-type strain. This suggests that in the absence of Tat export β -lactamase accumulates with the cells of the Δtat mutants and that β -lactamase activity is specifically reduced in the culture filtrates of the Δtat mutants. As seen in Fig. 2.4b, whole-cell extracts of the $\Delta blaS$ mutant did not exhibit a complete lack of nitrocefin hydrolyzing activity presumably due to the expression of the BlaE protein. BlaE was recently shown to have a minor contribution to the total β -lactamase activity of *M. smegmatis*, without contributing significantly to resistance (25). Examination of the BlaE protein sequence revealed a Sec-like signal sequence. When complemented, the $\Delta tatC$ mutant showed wild-type β -lactamase activity in the culture filtrate. Complementation of the $\Delta tatA$ mutant also resulted in an increase in culture filtrate β -lactamase activity, although to a reduced degree. These results indicate a defect in β -lactamase export by the Δtat mutants and serve to explain the increased β -lactam sensitivity observed in the absence of TatA and TatC.

Export of active *M. tuberculosis* BlaC depends on the Tat pathway and a twin-arginine signal sequence.

The above results strongly suggest that the *M. smegmatis* β -lactamase is a Tat substrate. Like BlaS, the major β -lactamase of *M. tuberculosis* (BlaC) has a predicted Tat signal sequence (Table 2.3). To determine if BlaC is a Tat substrate, we tested for Tat-dependent export of BlaC expressed in *M. smegmatis* Δ *blaS*, *M. smegmatis* Δ *tatA* Δ *blaS*, and *M. tuberculosis* Δ *blaC* mutants.

TABLE 2.3. *M. tuberculosis* signal sequences fused to 'BlaC

<i>Signal sequence origin</i>	<i>Amino acid sequence</i>
BlaC	MRNRGF GRRELL VAMAMLVSVTGCARHASG <u>ARP</u>
BlaC(KK)	MRNRGF GKKELL VAMAMLVSVTGCARHASG <u>ARP</u>
PlcA	MAS RREFL TKLTGAGAAAF LM DWAAPVIEKAYGGS <u>ARP</u>
PlcB	MAT RRQFF AKAAAAATTAGAFMSLAGPIIEKAYGGS <u>ARP</u>
PlcB(KK)	MAT KKQFF AKAAAAATTAGAFMSLAGPIIEKAYGGS <u>ARP</u>
Mpt63	MGMKLT TM IKTAVAVVAMAAIATFAAPVALAGS <u>ARP</u>
FbpB	MATDVSRKIRAWGRRLMIGTAAAVVLPGLVGLAGGAATAGAGS <u>ARP</u>

Signal sequences of PlcA, and PlcB were predicted by SignalP (3).
 Signal sequences of Mpt63 and FbpB were determined previously (35).
 Underlined residues denote the start of 'BlaC in the fusion constructs,
 residues in italics are derived from the Hsp60 coding region.

Unlike wild-type *M. smegmatis* and *M. tuberculosis*, the corresponding Δ *bla* mutants are sensitive to β -lactams, failing to grow on 7H10 agar medium containing 50 μ g/ml of carbenicillin. When the full-length *M. tuberculosis* BlaC was expressed from plasmid pMP327 in *M. smegmatis* Δ *blaS* and *M. tuberculosis* Δ *blaC* mutants the strains were resistant to carbenicillin (Table 2.4). This β -lactam resistance was dependent on BlaC export since expression of a truncated 'BlaC lacking the native signal sequence, from plasmid pJM113, in either *M. smegmatis* Δ *blaS* or *M. tuberculosis* Δ *blaC* did not allow growth on carbenicillin

TABLE 2.4. Export of functional 'BlaC is dependent on a twin-arginine signal sequence and an intact Tat pathway

Strain	Genotype	Mycobacterial vector*									
		Empty vector	pMP327 (BlaC)	pJM113 ('BlaC)	pJM117 (BlaC-KK)	pJM130 (ssPlcA-'BlaC)	pJM111 (ssPlcB-'BlaC)	pJM118 (ssPlcB-KK-'BlaC)	pMB228 (ssMpt63-'BlaC)	pJM109 (ssFbpB-'BlaC)	
		Carbenicillin Resistance									
<i>M. smegmatis</i>	mc ² 155	WT	+	+	+	+	ND	+	+	+	+
	PM759	$\Delta blaS$	-	+	-	-	+	+	-	-	-
	JM578	$\Delta tatA \Delta blaS$	-	-	-	-	-	-	-	-	-
<i>M. tuberculosis</i>	H37Rv	WT	+	+	+	+	ND	+	+	+	+
	PM638	$\Delta blaC$	-	+	-	-	+	+	-	-	-

* All strains carrying a vector were resistant to 20 μ g/ml kanamycin as a result of the vector resistance marker. Carbenicillin resistance and sensitivity was determined by the presence (+) or absence (-) of single colony growth on 7H10 plates containing 50 μ g/ml carbenicillin and 20 μ g/ml kanamycin. ND, not done.

containing agar medium (Table 2.4). Whole-cell extracts of *M. smegmatis* $\Delta blaS$ carrying pJM113 assayed for nitrocefin activity revealed a 20-fold higher β -lactamase activity compared to that of *M. smegmatis* $\Delta blaS$ confirming that truncated 'BlaC was expressed in this strain (Fig. 2.5a).

Based on the above results, we used β -lactam resistance as an indicator of the export of functional BlaC enzyme in the following experiments. Unlike the results obtained when we expressed full-length BlaC in the $\Delta blaS$ *M. smegmatis* mutant, expression of full-length BlaC in a *M. smegmatis* $\Delta tatA \Delta blaS$ double mutant did not protect against β -lactams. Nitrocefin analysis on whole-cell extracts of *M. smegmatis* $\Delta tatA \Delta blaS$ carrying pMP327 demonstrated that BlaC was expressed in this strain (Fig. 2.5b). Thus, the Tat pathway also appears to function in BlaC export.

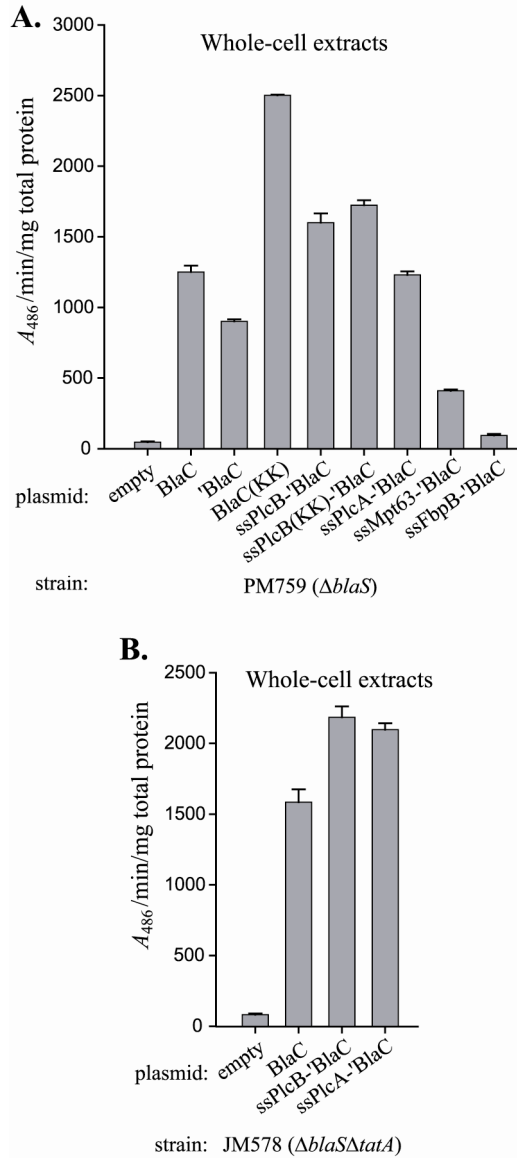


Fig. 2.5. Nitrocefin assays on whole-cell extracts of *M. smegmatis* strains carrying 'BlaC fusion constructs. Whole-cell extracts of *M. smegmatis* cultures were assayed for the ability to hydrolyze nitrocefin. (A) Assays were performed on *M. smegmatis* $\Delta blaS$ (strain PM759) strains carrying the following vectors: pMV261 (empty), pMP327 (BlaC), pJM113 ('BlaC), pJM117 (BlaC(KK)), pJM111 (ssPlcB-'BlaC), pJM118 (ssPlcB(KK)-'BlaC), pJM130 (ssPlcA-'BlaC), pMB228 (ssMpt63-'BlaC), and pJM109 (ssFbpB-'BlaC). (B) Assays were performed on *M. smegmatis* $\Delta tatA\Delta blaS$ (strain JM578) carrying pMV261 (empty), pMP327 (BlaC), pJM111 (ssPlcB-'BlaC), or pJM130 (ssPlcA-'BlaC). Nitrocefin hydrolyzing activity was expressed as $A_{486} \text{ min}^{-1} \text{ mg total protein}^{-1}$, and was standardized to PBS. Shown is a representative assay reported as an average \pm standard deviation.

In addition, we tested the significance of the twin-arginine motif in the BlaC signal sequence to the export of BlaC using mutants in which the twin-arginines 'RR' were substituted with conservative twin-lysine 'KK' residues (Table 2.3). Similar substitutions in Tat substrates from other bacteria have revealed the importance of the 'RR' dipeptide in Tat export (1, 20, 37, 73, 80). Expression of BlaC(KK) from plasmid pJM117 did not protect the Δbla mutants of *M. smegmatis* and *M. tuberculosis* from carbenicillin (Table 2.4). Nitrocefin analysis on *M. smegmatis* whole-cell extracts confirmed that BlaC(KK) was expressed from plasmid pJM117 in the $\Delta blaS$ background (Fig. 2.5a). This result further demonstrates that *M. tuberculosis* BlaC is a Tat substrate.

Truncated 'BlaC can be used as a reporter of Tat export in mycobacteria.

Since BlaC activity was dependent on the Tat pathway and an intact twin-arginine signal sequence, we tested whether truncated 'BlaC could report on the export of other Tat substrates. Plasmids pJM130 and pJM111, that express ssPlcA-'BlaC and ssPlcB-'BlaC fusion proteins in which the putative Tat signal sequences of *M. tuberculosis* PlcA and PlcB were fused in-frame and upstream of truncated 'BlaC (Table 2.3), respectively, were electroporated into $\Delta blaS$ *M. smegmatis* and $\Delta blaC$ *M. tuberculosis*. All the resulting strains were β -lactam resistant (i.e. grew in the presence of carbenicillin) (Table 2.4). This indicates that the PlcA and PlcB signal sequences direct export of functional 'BlaC in *M. smegmatis* and *M. tuberculosis*.

Additional experiments demonstrated that the β -lactam protection observed with the ssPlcA-'BlaC and ssPlcB-'BlaC fusions was dependent on Tat export. First, TatA was required for the β -lactam resistance as shown by the failure of either ssPlcA-'BlaC or ssPlcB-

'BlaC to protect against carbenicillin when expressed in the *M. smegmatis* Δ tat Δ blaS double mutant (Table 2.4). Second, a ssPlcB-'BlaC fusion engineered to have a twin-lysine (KK) substitution for the twin-arginines (RR) in the signal sequence (expressed from plasmid pJM118) did not protect against carbenicillin when expressed in the Δ blaS and Δ blaC strains of *M. smegmatis* and *M. tuberculosis* (Tables 3 and 4). Cell lysates from all these strains possessed β -lactamase activity by nitrocefin analysis (Fig. 2.5). These results reveal a role for the Tat pathway and the 'RR' motif in PlcB export. They further demonstrate that 'BlaC can be used as a reporter of protein export by the Tat pathway.

The 'BlaC reporter does not function with Sec-dependent signal sequences.

Having demonstrated that native BlaC is a Tat substrate that could report on Tat export, we asked whether the 'BlaC reporter also functions when targeted to the Sec pathway. MPT63/Rv1926c and Antigen-85b/FbpB/Rv1886c are well-characterized secreted proteins of *M. tuberculosis* (35) and both of these proteins have putative Sec signal sequences (Table 2.3) (43). Plasmids pMB228 (ssMpt63-'BlaC) and pJM109 (ssFbpB-'BlaC) were tested in *M. smegmatis* Δ blaS and *M. tuberculosis* Δ blaC. Both ssMpt63-'BlaC and ssFbpB-'BlaC failed to protect the *bla* mutants from carbenicillin (Table 2.4). Nitrocefin analysis of whole-cell lysates of Δ blaS *M. smegmatis* carrying pMB228 (ssMpt63-'BlaC) revealed a very low level of β -lactamase activity, while Δ blaS *M. smegmatis* carrying pJM109 (ssFbpB-'BlaC) did not exhibit any detectable activity (Fig. 2.5a). The different outcome of testing these ssSec-'BlaC fusions suggests that, unlike Tat signal sequences, Sec signal sequences are unable to promote export of a functionally active 'BlaC.

Discussion

Here we report the functional characterization of the Tat pathway in mycobacteria. Our study shows that deletion of the *tatA* and *tatC* genes in *M. smegmatis* results in slow growth on agar medium, hypersensitivity to SDS, and sensitivity to β -lactam antibiotics. The diverse phenotypes we observed suggest that the mycobacterial Tat pathway exports multiple substrates.

Our first attempts to construct *tatA* and *tatC* deletion mutants from merodiploid single-crossover strains were problematic and suggested that the Tat pathway was essential in *M. smegmatis*. This was consistent with Transposon Site Hybridization (TraSH) analysis in *M. tuberculosis* which defined *tatC* as being an essential gene for optimal growth (65). However, with extended incubation we recovered deletion mutants of *tatA* and *tatC*. Currently, the basis of the Δ *tat* mutant growth defect is not clear. We hypothesize that mislocalization of a Tat substrate involved in nutrient uptake is responsible, since a number of *M. tuberculosis* Tat substrates, predicted by TATFIND, are homologous to periplasmic substrate-binding proteins of ABC transporters.

The *M. smegmatis* Δ *tatA* and Δ *tatC* mutants exhibited identical phenotypes, consistent with TatA and TatC being components of the same system. Research on other bacteria suggests that a functional Tat pathway requires a minimum of one TatA and one TatC homologue (7, 54). A mycobacterial ORF with similarity to *E. coli* TatB and conserved residues important to TatB function is present in the *M. smegmatis* and *M. tuberculosis* genomes (McDonough, Braunstein unpublished results, TubercuList; Institut Pasteur [<http://genolist.pasteur.fr/TubercuList/>]) (32, 81). However, the role of TatB in mycobacteria

remains to be established. Unlike *tatC*, *tatB* of *M. tuberculosis* was described as non-essential by TraSH analysis (65). Furthermore, the mycobacterial *tatB* is distantly located from *tatA* and *tatC* in the chromosome. Instead *tatB* is located with genes encoding σ^E and potential regulators of σ^E (*sigE*, *rseA*, and *htrA*). This has led to the proposal that TatB functions in σ^E regulation, possibly as a specialized component of the Tat translocation system (45).

Some of the phenotypes reported for *tat* mutants of other bacteria were also exhibited by the *M. smegmatis* Δ *tat* mutants. Interestingly, the basis of these *tat* phenotypes is not necessarily conserved. In *E. coli* Δ *tat* mutants, the SDS hypersensitivity phenotype was linked to impairment of cell division, as indicated by cell filamentation, and attributed to mislocalization of amidases carrying Tat signal sequences (38, 72). Although the *M. smegmatis* Δ *tat* mutants were also hypersensitive to SDS, no filamentation, morphological defect, or size difference was observed by light microscopy (data not shown). Furthermore, analysis of the protein sequences of putative mycobacterial amidases did not reveal recognizable Tat signal sequences. Consequently, we believe the basis of the SDS sensitivity phenotype in *M. smegmatis* Δ *tat* mutants differs from that in *E. coli*. The drug sensitivity phenotype of the *M. smegmatis* Δ *tat* mutants also differs from that in *E. coli* in being specific for β -lactams as opposed to a general increase in drug susceptibility due to a defect in the integrity of the cell envelope (72).

Using multiple approaches, we demonstrated that mycobacterial β -lactamases are native Tat substrates. The Δ *tatA* and Δ *tatC* mutants of *M. smegmatis* failed to grow in the presence of 50 μ g/ml carbenicillin and exhibited a significant reduction in β -lactamase activity in the culture filtrate that was similar to the level observed with the β -lactam

sensitive $\Delta blaS$ β -lactamase mutant. We further showed that expression of *M. tuberculosis* BlaC in the *M. smegmatis* $\Delta blaS$ background protected against β -lactams in a Tat-dependent manner and that protection was dependent on the 'RR' dipeptide' in the twin-arginine motif. The latter experiments were performed in both *M. smegmatis* $\Delta blaS$ and *M. tuberculosis* $\Delta blaC$ strains.

To our knowledge, BlaS and BlaC are the first examples of Tat-exported β -lactamases. Since the Tat pathway exports folded proteins, the Tat-dependent nature of the mycobacterial β -lactamases may reflect rapid intracellular folding of the proteins prior to export. Cofactor-bound proteins are a common category of Tat substrates (48), but BlaC and BlaS are homologous to cofactorless class-A β -lactamases (25, 29, 44). The primary amino acid sequences of BlaC and BlaS are only 37% identical which suggests that the Tat-dependence reflects a common requirement for Tat export as opposed to simple conservation among orthologues. Evaluation of bacterial genome databases reveals that *Mycobacterium leprae* and *Mycobacterium fortuitum* also have β -lactamases with predicted Tat signal sequences, along with a small number of additional non-mycobacterial species including *Burkholderia cepacia* (77), *Burkholderia pseudomallei* (15), *Nocardia asteroides* (53), *Streptomyces clavuligerus* (52), and *Yersinia enterocolitica* (68).

Because β -lactams target proteins in the cell wall, β -lactamases must be exported to protect the cell. This feature of β -lactamase function was previously exploited in developing TEM β -lactamase (TEM Bla) as a genetic reporter of protein export in *E. coli* and other bacteria (14, 49, 69). Despite the appearance of a Sec-like signal sequence on the native TEM Bla, the truncated reporter was shown to work with both Sec and Tat signal sequences (69, 74). In an effort to develop a reporter of protein export that works directly in *M.*

tuberculosis, we tested truncated 'BlaC and protein fusions in which signal sequences were placed in-frame and upstream of 'BlaC in $\Delta blaS$ or $\Delta blaC$ mycobacteria for the ability to protect against β -lactams. Of the four 'BlaC fusions tested, only those with the putative Tat signal sequences (ssPlcA-'BlaC and ssPlcB-'BlaC) promoted growth on carbenicillin plates; the ssSec-'BlaC fusions (Mpt63 and FbpB) did not. These results indicate that Tat signal sequences can direct export of a functional 'BlaC reporter but the Sec signal sequences we tested cannot. This suggests that 'BlaC will be a useful tool for studying the Tat pathway in mycobacteria and for identifying exported Tat substrates in *M. tuberculosis* by selecting active 'BlaC fusions from expression libraries.

The plasmid pMM7-2 we used to express the ssFbpB-'BlaC fusion was successful in expressing an active exported ssFbpB-'PhoA alkaline phosphatase fusion in *M. smegmatis* (data not shown). Thus, it appears that this FbpB signal sequence construct can drive export of functional Sec substrates, such as PhoA, but not Tat substrates like BlaC. Interestingly, the signal sequence of FbpB contains consecutive arginines (R₁₃R₁₄L₁₅M₁₆I₁₇) (Table 2.3). However, it possesses a region of high hydrophobicity following the R-R-x- ϕ - ϕ motif which is not characteristic of Tat substrates, and TATFIND did not predict FbpB to be a Tat substrate (21). Our results with the 'BlaC reporter are consistent with FbpB not being a Tat substrate.

Our inability to detect significant levels of ssSec-'BlaC fusion proteins by nitrocefin analysis of whole-cell extracts may reflect rapid degradation of fusions incompatible with export from the cytoplasm, as this is a common fate of non-exported Tat substrates (8, 17). An alternate possibility is that the ssSec-'BlaC fusions are translocated by the Sec pathway in an unfolded state and that the resulting exported 'BlaC is enzymatically inactive due to

improper folding. In this latter scenario 'BlaC reports on Tat export because it is only functional if exported by the Tat pathway. A precedent for this type of Tat-dependence is the export of green fluorescent protein (GFP) by *E. coli*. When exported by the Tat pathway GFP is active, but when exported by the Sec pathway it is inactive (24, 62, 76).

We used the nitrocefin assay on whole-cell extracts as a way to assess fusion protein expression in carbenicillin sensitive strains (Fig. 2.5). Interestingly, this analysis also revealed substantial cell-associated β -lactamase activity in the $\Delta blaS$ strain expressing full length BlaC. This was surprising since our analysis of native *M. smegmatis* BlaS revealed a relatively low level of β -lactamase activity in the wild-type cell extract (Fig. 2.4b).

Comparison of BlaC and BlaS revealed a potential lipoprotein lipid attachment site at cysteine-24 of the BlaC predicted signal sequence (Table 2.3)(41). Thus, lipid modification of BlaC may tether the protein to the cell wall accounting for the different localization of exported *M. smegmatis* and *M. tuberculosis* β -lactamases.

Examination of the PlcA, PlcB, and PlcC proteins of *M. tuberculosis* reveals the presence of an N-terminal twin-arginine motif for each enzyme, and empirical evidence in *M. tuberculosis* demonstrates that these proteins are cell wall-localized (39, 58). There is a precedent for phospholipase C proteins being Tat substrates in *Pseudomonas aeruginosa* (79). Using ssPlcA-'BlaC and ssPlcB-'BlaC fusions, we demonstrated Tat-dependence for the signal sequences of the *M. tuberculosis* phospholipase C proteins. These 'BlaC fusion proteins were active when expressed in a $\Delta blaS$ background but not in a $\Delta blaS\Delta tata$ background. Furthermore, the 'RR' dipeptide in the PlcB signal sequence was required for the export of an active 'BlaC fusion as shown using the ssPlcB(KK)-'BlaC fusion protein. Interestingly, PlcA was not identified by the TATFIND program, presumably due to the

presence of a negatively charged aspartic acid in the C-terminal region of the predicted signal sequence. In contrast, PlcB lacks this residue and was predicted to be a Tat substrate by TATFIND (21). Our results with the PlcA signal sequence may reflect mycobacterial-specific properties of Tat signal sequences that are not incorporated into the TATFIND algorithm. Because the 'BlaC reporter is a native mycobacterial protein dependent on a Tat signal sequence and Tat pathway for export, it may be able to identify mycobacterial Tat substrates overlooked by a bioinformatic approach.

Using *M. tuberculosis* phospholipase C mutants, Raynaud *et al.* demonstrated the contribution of PlcA, PlcB, and PlcC to the full virulence of *M. tuberculosis* in a mouse model of infection (58). Similarly, a *P. aeruginosa* Δ *tatC* mutant unable to export two phospholipase C proteins, PlcH and PlcN, is attenuated for virulence (47). Additional bacterial pathogens (*Agrobacterium tumefaciens*, enterohemorrhagic *E. coli* O157:H7, and *Legionella pneumophila*) require a functional Tat pathway for full virulence, although the exported factors involved are not clearly defined (22, 55, 60). We expect that the *M. tuberculosis* Tat pathway will also contribute to pathogenesis by exporting virulence factors, including Plc enzymes. *M. tuberculosis* Δ *tat* mutants are required to directly test the role of the Tat pathway in *M. tuberculosis* virulence; however, the potential for a *M. tuberculosis* Δ *tat* mutant to exhibit a growth defect *in vitro* may complicate such analysis.

Our demonstration of the existence of a functional Tat pathway in mycobacteria brings us closer to understanding the protein export systems that function in mycobacterial physiology and pathogenesis. Knowledge of the Tat pathway may also benefit the development of novel tuberculosis vaccines. Recombinant *Mycobacterium bovis* BCG strains with improved secretion capability and/or the ability to export foreign antigens have

been proposed as a vaccine strategy (31, 34). Since the Tat pathway is capable of exporting folded proteins it could be the ideal mechanism for secreting highly expressed proteins and foreign antigens. Finally, the 'BlaC reporter system should serve as a powerful genetic tool for studying protein export directly in *M. tuberculosis* and help us identify the complete set of *M. tuberculosis* Tat-dependent exported proteins.

Acknowledgements

This research was funded in part by grants from the National Institute of Allergy and Infectious Disease of the NIH (AI54540, M.B) (AI47311, M.S.P), a developmental grant from the University of North Carolina at Chapel Hill Center for AIDS Research (NIH #9P30 AI 50410-04), and a Burroughs Wellcome Fund Career Award in the Biomedical Sciences (M.S.P). J.A.M. was also supported by the Training in Sexually Transmitted Diseases and AIDS NIH NIAID training grant 5T32 AI07001-28. A.R.F. was also supported by the Molecular Pathogenesis of Bacteria and Viruses NIH training grant T32 AI07362 and the Medical Scientist Training Program funded by the NIH grant T32 GM07356.

We thank M. Pohlschröder for providing us with the *M. tuberculosis* TATFIND list prior to public release of the program and the members of the Braunstein laboratory for review of the manuscript.

Attributions

The work described here was performed primarily by me with the following exceptions. Kari Hacker and Miriam Braunstein constructed the $\Delta tatA$ deletion mutant. Tony Flores and Marty Pavelka provided the β -lactamase deletion strains of *M. smegmatis* and *M. tuberculosis*, as well as the corresponding complementing vectors. This work has been previously published (*J Bacteriol.* 2005 Nov;187(22):7667-79). Permission to reprint this work has been granted by the publisher.

References

1. **Alami, M., D. Trescher, L. F. Wu, and M. Muller.** 2002. Separate analysis of twin-arginine translocation (Tat)-specific membrane binding and translocation in *Escherichia coli*. *J Biol Chem* **277**:20499-503.
2. **Bange, F. C., F. M. Collins, and W. R. Jacobs, Jr.** 1999. Survival of mice infected with *Mycobacterium smegmatis* containing large DNA fragments from *Mycobacterium tuberculosis*. *Tuber Lung Dis* **79**:171-80.
3. **Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak.** 2004. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* **340**:783-95.
4. **Berks, B. C.** 1996. A common export pathway for proteins binding complex redox cofactors? *Mol Microbiol* **22**:393-404.
5. **Berks, B. C., T. Palmer, and F. Sargent.** 2005. Protein targeting by the bacterial twin-arginine translocation (Tat) pathway. *Curr Opin Microbiol* **8**:174-81.
6. **Berks, B. C., F. Sargent, and T. Palmer.** 2000. The Tat protein export pathway. *Mol Microbiol* **35**:260-74.
7. **Blaudeck, N., P. Kreutzenbeck, M. Muller, G. A. Sprenger, and R. Freudl.** 2005. Isolation and characterization of bifunctional *Escherichia coli* TatA mutant proteins that allow efficient tat-dependent protein translocation in the absence of TatB. *J Biol Chem* **280**:3426-32.
8. **Blaudeck, N., G. A. Sprenger, R. Freudl, and T. Wiegert.** 2001. Specificity of signal peptide recognition in tat-dependent bacterial protein translocation. *J Bacteriol* **183**:604-10.
9. **Bogsch, E. G., F. Sargent, N. R. Stanley, B. C. Berks, C. Robinson, and T. Palmer.** 1998. An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. *J Biol Chem* **273**:18003-6.
10. **Braunstein, M., S. S. Bardarov, and W. R. Jacobs, Jr.** 2002. Genetic methods for deciphering virulence determinants of *Mycobacterium tuberculosis*. *Methods Enzymol* **358**:67-99.
11. **Braunstein, M., A. M. Brown, S. Kurtz, and W. R. Jacobs, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. *J Bacteriol* **183**:6979-90.

12. **Braunstein, M., B. J. Espinosa, J. Chan, J. T. Belisle, and W. R. Jacobs, Jr.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* **48**:453-64.
13. **Bronstein, P., M. Marrichi, and M. P. DeLisa.** 2004. Dissecting the twin-arginine translocation pathway using genome-wide analysis. *Res Microbiol* **155**:803-10.
14. **Broome-Smith, J. K., M. Tadayyon, and Y. Zhang.** 1990. Beta-lactamase as a probe of membrane protein assembly and protein export. *Mol Microbiol* **4**:1637-44.
15. **Cheung, T. K., P. L. Ho, P. C. Woo, K. Y. Yuen, and P. Y. Chau.** 2002. Cloning and expression of class A beta-lactamase gene *blaA*(BPS) in *Burkholderia pseudomallei*. *Antimicrob Agents Chemother* **46**:1132-5.
16. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, F. Tekaiia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, and B. G. Barrell.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**:537-44.
17. **Cristobal, S., J. W. de Gier, H. Nielsen, and G. von Heijne.** 1999. Competition between Sec- and TAT-dependent protein translocation in *Escherichia coli*. *EMBO J* **18**:2982-90.
18. **Dalbey, R. E., and M. Chen.** 2004. Sec-translocase mediated membrane protein biogenesis. *Biochim Biophys Acta* **1694**:37-53.
19. **de Keyzer, J., C. van der Does, and A. J. Driessen.** 2003. The bacterial translocase: a dynamic protein channel complex. *Cell Mol Life Sci* **60**:2034-52.
20. **DeLisa, M. P., P. Samuelson, T. Palmer, and G. Georgiou.** 2002. Genetic analysis of the twin arginine translocator secretion pathway in bacteria. *J Biol Chem* **277**:29825-31.
21. **Dilks, K., R. W. Rose, E. Hartmann, and M. Pohlschroder.** 2003. Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. *J Bacteriol* **185**:1478-83.
22. **Ding, Z., and P. J. Christie.** 2003. *Agrobacterium tumefaciens* twin-arginine-dependent translocation is important for virulence, flagellation, and chemotaxis but not type IV secretion. *J Bacteriol* **185**:760-71.
23. **Dye, C., S. Scheele, P. Dolin, V. Pathania, and M. C. Raviglione.** 1999. Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and

- mortality by country. WHO Global Surveillance and Monitoring Project. JAMA **282**:677-86.
24. **Feilmeier, B. J., G. Iseminger, D. Schroeder, H. Webber, and G. J. Phillips.** 2000. Green fluorescent protein functions as a reporter for protein localization in *Escherichia coli*. J Bacteriol **182**:4068-76.
 25. **Flores, A. R., L. M. Parsons, and M. S. Pavelka, Jr.** 2005. Genetic analysis of the β -lactamases of *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* and susceptibility to β -lactam antibiotics. Microbiology **151**:521-32.
 26. **Glickman, M. S., J. S. Cox, and W. R. Jacobs, Jr.** 2000. A novel mycolic acid cyclopropane synthetase is required for cording, persistence, and virulence of *Mycobacterium tuberculosis*. Mol Cell **5**:717-27.
 27. **Gohlke, U., L. Pullan, C. A. McDevitt, I. Porcelli, E. de Leeuw, T. Palmer, H. R. Saibil, and B. C. Berks.** 2005. The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. Proc Natl Acad Sci U S A **102**:10482-6.
 28. **Gouffi, K., C. L. Santini, and L. F. Wu.** 2002. Topology determination and functional analysis of the *Escherichia coli* TatC protein. FEBS Lett **525**:65-70.
 29. **Hackbarth, C. J., I. Unsal, and H. F. Chambers.** 1997. Cloning and sequence analysis of a class A beta-lactamase from *Mycobacterium tuberculosis* H37Ra. Antimicrob Agents Chemother **41**:1182-5.
 30. **Halbig, D., T. Wiegert, N. Blaudeck, R. Freudl, and G. A. Sprenger.** 1999. The efficient export of NADP-containing glucose-fructose oxidoreductase to the periplasm of *Zymomonas mobilis* depends both on an intact twin-arginine motif in the signal peptide and on the generation of a structural export signal induced by cofactor binding. Eur J Biochem **263**:543-51.
 31. **Hess, J., D. Miko, A. Catic, V. Lehmensiek, D. G. Russell, and S. H. Kaufmann.** 1998. *Mycobacterium bovis* Bacille Calmette-Guerin strains secreting listeriolysin of *Listeria monocytogenes*. Proc Natl Acad Sci U S A **95**:5299-304.
 32. **Hicks, M. G., E. de Leeuw, I. Porcelli, G. Buchanan, B. C. Berks, and T. Palmer.** 2003. The *Escherichia coli* twin-arginine translocase: conserved residues of TatA and TatB family components involved in protein transport. FEBS Lett **539**:61-7.
 33. **Hinsley, A. P., N. R. Stanley, T. Palmer, and B. C. Berks.** 2001. A naturally occurring bacterial Tat signal peptide lacking one of the 'invariant' arginine residues of the consensus targeting motif. FEBS Lett **497**:45-9.

34. **Horwitz, M. A., G. Harth, B. J. Dillon, and S. Maslesa-Galic.** 2000. Recombinant bacillus calmette-guerin (BCG) vaccines expressing the *Mycobacterium tuberculosis* 30-kDa major secretory protein induce greater protective immunity against tuberculosis than conventional BCG vaccines in a highly susceptible animal model. *Proc Natl Acad Sci U S A* **97**:13853-8.
35. **Horwitz, M. A., B. W. Lee, B. J. Dillon, and G. Harth.** 1995. Protective immunity against tuberculosis induced by vaccination with major extracellular proteins of *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* **92**:1530-4.
36. **Ignatova, Z., C. Hornle, A. Nurk, and V. Kasche.** 2002. Unusual signal peptide directs penicillin amidase from *Escherichia coli* to the Tat translocation machinery. *Biochem Biophys Res Commun* **291**:146-9.
37. **Ize, B., F. Gerard, and L. F. Wu.** 2002. In vivo assessment of the Tat signal peptide specificity in *Escherichia coli*. *Arch Microbiol* **178**:548-53.
38. **Ize, B., N. R. Stanley, G. Buchanan, and T. Palmer.** 2003. Role of the *Escherichia coli* Tat pathway in outer membrane integrity. *Mol Microbiol* **48**:1183-93.
39. **Johansen, K. A., R. E. Gill, and M. L. Vasil.** 1996. Biochemical and molecular analysis of phospholipase C and phospholipase D activity in mycobacteria. *Infect Immun* **64**:3259-66.
40. **Jongbloed, J. D., U. Martin, H. Antelmann, M. Hecker, H. Tjalsma, G. Venema, S. Bron, J. M. van Dijk, and J. Muller.** 2000. TatC is a specificity determinant for protein secretion via the twin-arginine translocation pathway. *J Biol Chem* **275**:41350-7.
41. **Juncker, A. S., H. Willenbrock, G. Von Heijne, S. Brunak, H. Nielsen, and A. Krogh.** 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci* **12**:1652-62.
42. **Kurtz, S., and M. Braunstein.** 2005. Protein Secretion and Export in *Mycobacterium tuberculosis*, p. 71-138. In T. Parish (ed.), *Mycobacterium Molecular Microbiology*. Horizon Bioscience, London.
43. **Lee, B. Y., and M. A. Horwitz.** 1999. T-cell epitope mapping of the three most abundant extracellular proteins of *Mycobacterium tuberculosis* in outbred guinea pigs. *Infect Immun* **67**:2665-70.
44. **Li, X. Z., L. Zhang, and H. Nikaido.** 2004. Efflux pump-mediated intrinsic drug resistance in *Mycobacterium smegmatis*. *Antimicrob Agents Chemother* **48**:2415-23.

45. **Manganelli, R., R. Provvedi, S. Rodrigue, J. Beaucher, L. Gaudreau, and I. Smith.** 2004. Sigma factors and global gene regulation in *Mycobacterium tuberculosis*. *J Bacteriol* **186**:895-902.
46. **O'Callaghan, C. H., A. Morris, S. M. Kirby, and A. H. Shingler.** 1972. Novel method for detection of beta-lactamases by using a chromogenic cephalosporin substrate. *Antimicrob Agents Chemother* **1**:283-8.
47. **Ochsner, U. A., A. Snyder, A. I. Vasil, and M. L. Vasil.** 2002. Effects of the twin-arginine translocase on secretion of virulence factors, stress response, and pathogenesis. *Proc Natl Acad Sci U S A* **99**:8312-7.
48. **Palmer, T., F. Sargent, and B. C. Berks.** 2005. Export of complex cofactor-containing proteins by the bacterial Tat pathway. *Trends Microbiol* **13**:175-80.
49. **Palva, I., M. Sarvas, P. Lehtovaara, M. Sibakov, and L. Kaariainen.** 1982. Secretion of *Escherichia coli* beta-lactamase from *Bacillus subtilis* by the aid of alpha-amylase signal sequence. *Proc Natl Acad Sci U S A* **79**:5582-6.
50. **Pavelka, M. S., Jr., and W. R. Jacobs, Jr.** 1996. Biosynthesis of diaminopimelate, the precursor of lysine and a component of peptidoglycan, is an essential function of *Mycobacterium smegmatis*. *J Bacteriol* **178**:6496-507.
51. **Pavelka, M. S., Jr., and W. R. Jacobs, 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**:4780-9.
52. **Perez-Llarena, F., J. F. Martin, M. Galleni, J. J. Coque, J. L. Fuente, J. M. Frere, and P. Liras.** 1997. The *bla* gene of the cephamycin cluster of *Streptomyces clavuligerus* encodes a class A beta-lactamase of low enzymatic activity. *J Bacteriol* **179**:6035-40.
53. **Poirel, L., F. Laurent, T. Naas, R. Labia, P. Boiron, and P. Nordmann.** 2001. Molecular and biochemical analysis of AST-1, a class A beta-lactamase from *Nocardia asteroides* sensu stricto. *Antimicrob Agents Chemother* **45**:878-82.
54. **Pop, O., U. Martin, C. Abel, and J. P. Muller.** 2002. The twin-arginine signal peptide of PhoD and the TatAd/Cd proteins of *Bacillus subtilis* form an autonomous Tat translocation system. *J Biol Chem* **277**:3268-73.
55. **Pradel, N., C. Ye, V. Livrelli, J. Xu, B. Joly, and L. F. Wu.** 2003. Contribution of the twin arginine translocation system to the virulence of enterohemorrhagic *Escherichia coli* O157:H7. *Infect Immun* **71**:4908-16.

56. **Pugsley, A. P.** 1993. The complete general secretory pathway in gram-negative bacteria. *Microbiol Rev* **57**:50-108.
57. **Raynaud, C., G. Etienne, P. Peyron, M. A. Laneelle, and M. Daffe.** 1998. Extracellular enzyme activities potentially involved in the pathogenicity of *Mycobacterium tuberculosis*. *Microbiology* **144 (Pt 2)**:577-87.
58. **Raynaud, C., C. Guilhot, J. Rauzier, Y. Bordat, V. Pelicic, R. Manganelli, I. Smith, B. Gicquel, and M. Jackson.** 2002. Phospholipases C are involved in the virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* **45**:203-17.
59. **Robinson, C., and A. Bolhuis.** 2004. Tat-dependent protein targeting in prokaryotes and chloroplasts. *Biochim Biophys Acta* **1694**:135-47.
60. **Rossier, O., and N. P. Cianciotto.** 2005. The *Legionella pneumophila* *tatB* gene facilitates secretion of phospholipase C, growth under iron-limiting conditions, and intracellular infection. *Infect Immun* **73**:2020-32.
61. **Sambrook, J., and D. W. Russell.** 2001. *Molecular Cloning: A Laboratory Manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
62. **Santini, C. L., A. Bernadac, M. Zhang, A. Chanal, B. Ize, C. Blanco, and L. F. Wu.** 2001. Translocation of jellyfish green fluorescent protein via the Tat system of *Escherichia coli* and change of its periplasmic localization in response to osmotic up-shock. *J Biol Chem* **276**:8159-64.
63. **Sargent, F., E. G. Bogsch, N. R. Stanley, M. Wexler, C. Robinson, B. C. Berks, and T. Palmer.** 1998. Overlapping functions of components of a bacterial Sec-independent protein export pathway. *EMBO J* **17**:3640-50.
64. **Sargent, F., N. R. Stanley, B. C. Berks, and T. Palmer.** 1999. Sec-independent protein translocation in *Escherichia coli*. A distinct and pivotal role for the TatB protein. *J Biol Chem* **274**:36073-82.
65. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48**:77-84.
66. **Schaerlaekens, K., M. Schierova, E. Lammertyn, N. Geukens, J. Anne, and L. Van Mellaert.** 2001. Twin-arginine translocation pathway in *Streptomyces lividans*. *J Bacteriol* **183**:6727-32.
67. **Schaerlaekens, K., L. Van Mellaert, E. Lammertyn, N. Geukens, and J. Anne.** 2004. The importance of the Tat-dependent protein secretion pathway in *Streptomyces* as revealed by phenotypic changes in *tat* deletion mutants and genome analysis. *Microbiology* **150**:21-31.

68. **Seoane, A., and J. M. Garcia Lobo.** 1991. Nucleotide sequence of a new class A beta-lactamase gene from the chromosome of *Yersinia enterocolitica*: implications for the evolution of class A beta-lactamases. *Mol Gen Genet* **228**:215-20.
69. **Smith, H., S. Bron, J. Van Ee, and G. Venema.** 1987. Construction and use of signal sequence selection vectors in *Escherichia coli* and *Bacillus subtilis*. *J Bacteriol* **169**:3321-8.
70. **Snapper, S. B., R. E. Melton, S. Mustafa, T. Kieser, and W. R. Jacobs, Jr.** 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol Microbiol* **4**:1911-9.
71. **Sparling, P. F., F. A. Sarubbi, Jr., and E. Blackman.** 1975. Inheritance of low-level resistance to penicillin, tetracycline, and chloramphenicol in *Neisseria gonorrhoeae*. *J Bacteriol* **124**:740-9.
72. **Stanley, N. R., K. Findlay, B. C. Berks, and T. Palmer.** 2001. *Escherichia coli* strains blocked in Tat-dependent protein export exhibit pleiotropic defects in the cell envelope. *J Bacteriol* **183**:139-44.
73. **Stanley, N. R., T. Palmer, and B. C. Berks.** 2000. The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. *J Biol Chem* **275**:11591-6.
74. **Stanley, N. R., F. Sargent, G. Buchanan, J. Shi, V. Stewart, T. Palmer, and B. C. Berks.** 2002. Behaviour of topological marker proteins targeted to the Tat protein transport pathway. *Mol Microbiol* **43**:1005-21.
75. **Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennett, G. P. Bansal, J. F. Young, M. H. Lee, and G. F. Hatfull.** 1991. New use of BCG for recombinant vaccines. *Nature* **351**:456-60.
76. **Thomas, J. D., R. A. Daniel, J. Errington, and C. Robinson.** 2001. Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in *Escherichia coli*. *Mol Microbiol* **39**:47-53.
77. **Trepanier, S., A. Prince, and A. Huletsky.** 1997. Characterization of the *penA* and *penR* genes of *Burkholderia cepacia* 249 which encode the chromosomal class A penicillinase and its LysR-type transcriptional regulator. *Antimicrob Agents Chemother* **41**:2399-405.
78. **Voladri, R. K., D. L. Lakey, S. H. Hennigan, B. E. Menzies, K. M. Edwards, and D. S. Kernodle.** 1998. Recombinant expression and characterization of the major beta-lactamase of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **42**:1375-81.

79. **Voulhoux, R., G. Ball, B. Ize, M. L. Vasil, A. Lazdunski, L. F. Wu, and A. Filloux.** 2001. Involvement of the twin-arginine translocation system in protein secretion via the type II pathway. *EMBO J* **20**:6735-41.
80. **Yahr, T. L., and W. T. Wickner.** 2001. Functional reconstitution of bacterial Tat translocation in vitro. *EMBO J* **20**:2472-9.
81. **Yen, M. R., Y. H. Tseng, E. H. Nguyen, L. F. Wu, and M. H. Saier, Jr.** 2002. Sequence and phylogenetic analyses of the twin-arginine targeting (Tat) protein export system. *Arch Microbiol* **177**:441-50.
82. **Zeoli, C., M. Maitland, H. Rose, B. Bloom, A. Mittelman, and J. Geliebter.** 2001. Expression of prostate specific molecules in bacille Calmette-Guerin: an immunotherapeutic approach to prostate cancer. *Prostate* **3**:92-97.

CHAPTER 3

β -lactamase as a Genetic Reporter Used to Identify Twin-arginine Translocation (Tat) Signal Sequences of *Mycobacterium tuberculosis*

Justin A. McDonough, Jessica R. McCann, Jason S. Silverman, Erin McElvania TeKippe, and Miriam Braunstein

Department of Microbiology and Immunology,
University of North Carolina School of Medicine, Chapel Hill, North Carolina

The export-dependent activity of β -lactamase makes this enzyme a useful reporter for extracytoplasmic localization of proteins. Chimeric proteins comprised of signal sequences fused to β -lactamase are competent for translocation across the cytoplasmic membrane and delivery to the bacterial cell wall where they can protect bacteria from β -lactam antibiotics. We previously described the use of the *Mycobacterium tuberculosis* β -lactamase 'BlaC as a reporter for protein export in mycobacteria. Unlike β -lactamase reporters described to date, export of functional 'BlaC is dependent on translocation through the Tat pathway. To gain insight into the role the Tat pathway plays in virulence and physiology of *M. tuberculosis*, we utilized the 'BlaC reporter to identify Tat-exported proteins of *M. tuberculosis*. We designed *M. tuberculosis* genomic DNA expression libraries for selection of ORFs fused upstream of 'blaC that result in export of the reporter. We selected for plasmids expressing exported 'BlaC fusion proteins in Δ blaS *M. smegmatis* on media containing the β -lactam antibiotic carbenicillin. Using this approach, we identified 130 carbenicillin-resistant plasmids representing a total of ten unique *M. tuberculosis* ORF-'blaC fusions. All ten have predicted

N-terminal signal sequences that fit the pattern of a Tat signal sequence, with the exception of the predicted oxidoreductase, Rv0063, which lacks a typical twin-arginine motif. When these fusions were tested subsequently in a $\Delta blaS \Delta tatA$ double mutant of *M. smegmatis*, all failed to promote carbenicillin resistance, indicating that export of the fusion protein was Tat-dependent. This work serves to experimentally identify *M. tuberculosis* sequences that promote Tat export and will help better understand the specific features of a Tat signal sequence that are important for export by this pathway. It was then important to confirm that the respective full length proteins were Tat-exported. Analysis of the full length proteins for Tat-dependent processing in *M. smegmatis* indicated that four out of nine proteins showed a Tat-dependent effect. The other five proteins showed no obvious Tat-dependent effect and require further analysis to draw a conclusion as to their export specificity. In these other cases, the mature domain of the protein may play a role in the selectivity of export. We also propose that in some of these cases, the full length proteins may require *M. tuberculosis*-specific factors for Tat export and processing. In conclusion, we identified *M. tuberculosis* signal sequences capable of targeting proteins to the Tat apparatus. The ORFs identified may play a collective role in virulence, antibiotic resistance, copper homeostasis, nutrient import, and cell envelope metabolism. Further elucidation of both the functions of these proteins and their specificity of export is needed to determine the precise role that the Tat pathway plays in these processes.

Introduction

Exported proteins of bacteria function in essential cellular processes such as nutrient acquisition and cell envelope metabolism. These surface-exposed and secreted proteins are also ideally positioned to interact with other cells and environmental factors. This protein subset includes exported proteins that interact with eukaryotic cells, functioning as virulence factors and immunogenic antigens.

In bacteria, protein export across the cytoplasmic membrane represents the first step in the delivery of a large number of proteins to their final location in the cell envelope or extracellular space. The general secretion (Sec) and the twin-arginine translocation (Tat) pathways are the best known examples of systems that transport proteins across the cytoplasmic membrane via protein channels in the membrane (for reviews, see 23, 49, 55, 71). Both systems recognize N-terminal signal sequences of precursor proteins prior to transport. These signal sequences have a charged N-terminal region “N”, a hydrophobic region “H”, and a signal peptidase cleavage region “C”. In most cases, the signal sequence is cleaved during or after translocation, liberating the mature exported substrate. Notable exceptions are the Sec-dependent insertion of transmembrane proteins (21) and the Rieske Fe/S proteins, in which upon export by the Tat pathway in both plant chloroplasts and bacteria the signal sequence is uncleaved and acts as an N-terminal membrane-anchor (2, 22, 54).

The Sec and Tat pathways differ in their ability to translocate proteins. Sec substrates are exported in an unfolded state, while Tat substrates are translocated in a folded state. Additionally, signal sequences targeted to the Tat apparatus (composed of TatA, TatB, and

TatC) carry a consensus motif in the charged N-terminal region called the twin-arginine motif, represented as R-R-x- ϕ - ϕ (where ϕ = an uncharged residue) (7, 8). Although the signal sequence as a whole is required for export, the arginine dipeptide is the major targeting determinant to the Tat apparatus. Conservative substitutions of both arginines for lysines, in a number of substrates studied to date, prevents the Tat-dependent export of these proteins (20, 42, 79). Interestingly, mutants in which only one of the arginines is replaced with a lysine or a glutamine do not always result in a complete block of export (24). Further, a small number of Tat-exported substrates exist in nature that lack the consecutive twin-arginines (44, 46), and this list is likely to grow as the Tat pathway is further investigated in prokaryotes.

Proteins targeted to the Tat apparatus in *E. coli* often contain bound cofactors or subunits that are incorporated in the cytoplasm. This process is dependent on the protein forming and maintaining a folded conformation prior to export, thus necessitating a pathway capable of translocating folded proteins and oligomeric complexes. In fact, in the absence of folding in the cytoplasm, Tat-dependent protein export does not occur, indicating that there is a requirement that the protein be folded prior to Tat export (25). In many cases, Tat signal sequence-containing proteins can only achieve a folded conformation with the aid of dedicated cytoplasmic chaperones, some of which insert cofactors. In *E. coli*, these chaperones can be specific to one or a small number of Tat substrates such as DmsD, which is involved in maturation of redox enzymes, or can be general in facilitating folding and stability of several Tat substrates, such as the DnaK chaperone (43, 56, 60). It is important to mention that not all Tat substrates incorporate complex cofactors. The Tat-exported phospholipase C of *Pseudomonas aeruginosa* PlcH does not bind a cofactor. Interestingly,

its export is still dependent on a chaperone PlcR (19). However, with most cofactorless Tat substrates, it is not yet clear if cytoplasmic chaperones are required.

To date, much of the knowledge of Tat-dependent protein export in bacteria has been the result of studies in model organisms such as *Escherichia coli* and *Bacillus subtilis* (49, 71). However, it is becoming increasingly clear that Tat plays a role in exporting virulence factors in bacterial pathogens such as *Agrobacterium tumefaciens*, enterohemorrhagic *E. coli* O157:H7, *Legionella pneumophila*, *Pseudomonas aeruginosa*, *Pseudomonas syringae*, and *Yersinia pseudotuberculosis* (12, 14, 27, 48, 62, 66, 89). Additionally, the Tat pathway has been speculated to function in the pathogenesis of *Mycobacterium tuberculosis* largely based on the proposal that the phospholipase C enzymes, shown to be important to virulence, are Tat substrates (53, 61, 63, 67).

Infection with *M. tuberculosis* can lead to tuberculosis disease (TB), which causes 1.7 million deaths each year globally (18). The development of genetic tools in mycobacteria, as well as the availability of the *M. tuberculosis* genome sequence, has led to increased research over the past decade to better understand the molecular physiology of this bacterium and its contribution to virulence (10, 16).

We and others previously demonstrated that the Tat secretion apparatus is dispensable in *Mycobacterium smegmatis*, and Δ *tat* mutants have pleiotropic phenotypes including slow growth on plates (53, 61). We speculate that the growth defect in Δ *tat* is due to an inability to localize proteins involved in a variety of cellular processes such as nutrient uptake and cell wall maintenance. Additionally, we demonstrated that the major β -lactamase of *M. smegmatis*, BlaS, is a Tat exported substrate. In a Δ *blaS* mutant, which cannot grow in the presence of β -lactam antibiotics, we were able to functionally complement the Δ *blaS*

mutation with the *M. tuberculosis* ortholog, BlaC. Furthermore, BlaC was shown to be exported in a Tat- and twin-arginine-dependent manner.

Recently, Saint-Joanis *et al.* (67) reported that the *tat* genes of *M. tuberculosis* are essential. A viable deletion mutant of the *M. tuberculosis* *tatA* and *tatC* locus could only be obtained if a second copy of the respective locus was expressed *in trans*. This finding is in agreement with global transposon insertion analysis (TraSH) of *M. tuberculosis*, which failed to identify a viable *tatC* insertion mutant *in vitro* and therefore categorized this as an essential gene (72). Given the slow growth phenotype we observed with the Δ *tat* mutants of *M. smegmatis*, the basis for the essentiality of the Tat pathway in *M. tuberculosis* is not clear. The inability to construct a *tat* mutant of *M. tuberculosis* makes the identification and analysis of putative Tat exported proteins difficult to study directly in the native bacterium. Specifically, comparative proteomic approaches are not an option. Therefore, alternative methods are needed to identify and study Tat exported proteins of *M. tuberculosis*.

To address this, we developed the BlaC β -lactamase as a reporter for Tat-dependent export in mycobacteria. We showed that the putative Tat signal sequences of *M. tuberculosis* phospholipase C enzymes, but not Sec signal sequences, are capable of driving the export of a truncated version of BlaC lacking its signal sequence ('BlaC), in Δ *blaS* *M. smegmatis* and Δ *blaC* *M. tuberculosis* (53). The export of a functionally active 'BlaC in the Δ *bla* background promotes growth of the bacteria in the presence of β -lactam antibiotics which allows us to select for exported fusions. Further, this system can be used to test for Tat-dependent export using the Δ *tat* mutants of *M. smegmatis*.

Here we used 'BlaC to identify Tat signal sequences of *M. tuberculosis* using a genomic fusion library in which we could select for *M. tuberculosis* ORFs fused upstream of

'*blaC* that result in functional export of the reporter. We identified ten unique Tat signal sequences capable of driving the export of 'BlaC in a Tat-dependent manner in *M. smegmatis*. In addition to the virulence factor phospholipase C, we identified proteins with homology to enzymes that function in carbohydrate and lipid metabolism, copper homeostasis, cell envelope maintenance, and nutrient import. Further, this experimental approach led to the identification of a protein that does not fit the typical pattern of what has been reported as a Tat signal sequence. This system proved to be an unbiased approach to identify sequences that promote Tat export and may lead to a better understanding of the elements required for Tat-specific export. Because the mature domain must be folded for Tat export, it is important to verify that the authentic full length proteins were Tat-exported. The basic question of how often it is that a protein with a functional Tat signal sequence, as shown with a reporter, is a true Tat substrate is only just now beginning to be addressed in *E. coli* (87). Analysis of the full length proteins for Tat-dependent processing in *M. smegmatis* indicated that four out of nine proteins showed a Tat-dependent effect. The other five proteins showed no obvious Tat-dependent effect. This latter category requires further analysis to make a definitive conclusion as to the export of these proteins. We additionally hypothesize that additional *M. tuberculosis* factors may be required for export and processing for some of these cases. We conclude that the BlaC genetic fusion library is a novel system to identify Tat signal sequences without the biases imposed by bioinformatics predictions and will be useful for increasing understanding of the elements that define a functional Tat signal sequence. The work so far suggests that four *M. tuberculosis* full length proteins are exported by the Tat pathway in *M. smegmatis* while more work is required to determine the export pathway specificity of the remaining proteins.

Materials and Methods

Bacterial strains and culture methods.

The strains used during this work are listed in Table 3.1. *E. coli* strain DH5 α was used for DNA cloning procedures. Luria-Bertani (LB) medium (Fisher) was used for culturing of *E. coli* and antibiotics were added at the following concentrations: 150 $\mu\text{g/ml}$, hygromycin B (Roche Applied Science); 100 $\mu\text{g/ml}$, carbenicillin; or 40 $\mu\text{g/ml}$, kanamycin (Acros Chemicals). Middlebrook 7H9 or 7H10 medium (Difco; BD Biosciences) was used for the culturing of *M. smegmatis* and *M. tuberculosis*. For growth of *M. smegmatis*, Middlebrook medium was supplemented with 0.5% glycerol and 0.2% dextrose. For growth of *M. tuberculosis*, Middlebrook medium was supplemented with 0.5% glycerol and 1 \times ADS (0.5% bovine serum albumin, fraction V [Roche]; 0.2% dextrose; and 0.85% NaCl). When necessary, media was supplemented with 0.05-0.1% Tween-80 (Fisher). Antibiotics were added to Middlebrook media at the following concentrations: 50 $\mu\text{g/ml}$, hygromycin B; 20 $\mu\text{g/ml}$, kanamycin; or 50 $\mu\text{g/ml}$, carbenicillin. L-lysine at 40 and 80 $\mu\text{g/ml}$ was added to agar and liquid media, respectively, for growth of *M. smegmatis* strains PM759 and JM578.

Molecular biology procedures.

Standard molecular biology techniques were employed as previously described (68). The Expand high fidelity PCR system (Roche) was used in all PCR reactions from chromosomal DNA and dimethyl sulfoxide at 5.0% was included in select PCR reactions. DNA sequencing was performed by either the UNC-CH Automated DNA Sequencing Facility (Chapel Hill, NC) or by Eton Bioscience Inc. (San Diego, CA).

Table 3.1. Strains used in this study

Strain	Description	Reference
<i>E. coli</i>		
K-12 DH5 α	F- [ϕ 80d Δ lacZM15] Δ (lacZYA-argF) U169 deoR recA1 endA1 hsdR17 glnV44 thi-1 gyrA96 relA1	Gibco-BRL
TOP10	F- mcrA [ϕ 80d Δ lacZM15] Δ (mrr-hsdRMS-mcrBC) Δ lacX74 deoR recA1 endA1 araD139 Δ (ara, leu)7697 galU galK rpsL nupG	Invitrogen
XL-1 Blue	[F' proAB, lacI' Δ M15, Tn10(Tet ^r)] recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 λ lac ^c	Stratagene
<i>M. smegmatis</i>		
mc ² 155	ept-1	(76)
MB692	mc ² 155, Δ tatA	(53)
JM567	mc ² 155, Δ tatC	(53)
PM759	mc ² 155, Δ blaS1 Δ lysA4 rpsL6	(33)
JM578	PM759, Δ tatA	(53)
<i>M. tuberculosis</i>		
H37Rv	Virulent	(4)
PM638	H37Rv, Δ blaC1	(33)

Construction of 'BlaC reporter libraries.

Two separate library plasmids were made. For both, a truncated version of *M. tuberculosis blaC* (referred to as '*blaC*') was amplified from pJM106 by PCR using the primers Lib'blaCfor and Lib'blaCrev (Table 3.2).

(i) Library 1. The resulting '*blaC*' amplicon was ligated into the multicopy mycobacterial shuttle vector pMV206.hyg, which had been digested with *Cla*I and *Nco*I. The final plasmid product was named pJES113 (Fig. 3.1). Genomic DNA was isolated from *M. tuberculosis* strain H37Rv as previously described (10), partially digested with *Acc*II and *Hpa*II, and DNA fragments of 0.5 to 5.0-kbp were selected. The genomic digest was cloned into the unique *Cla*I site immediately upstream of '*blaC*' in pJES113. The resulting ligation reaction was transformed into chemically competent *E. coli* XL-1 Blue (Stratagene). Approximately 1×10^6 hygromycin-resistant *E. coli* transformants were pooled for plasmid DNA isolation (Qiagen).

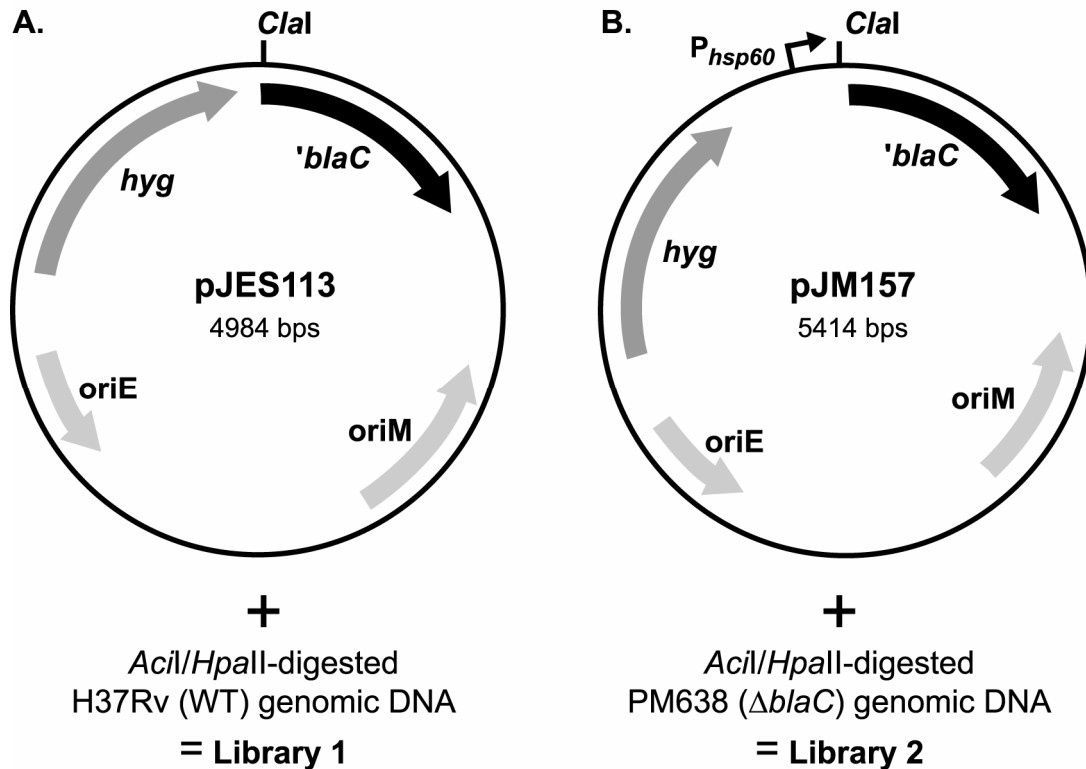


Figure 3.1. Construction of multicopy *'blaC* expression libraries. See Materials and Methods. (A) Library 1 was constructed by ligating ~0.5-5.0 kbp *AciI* or *HpaII* restriction fragments of *M. tuberculosis* H37Rv genomic DNA into the *Clal* site located on the multicopy shuttle plasmid pJES113. The truncated β -lactamase gene, *'blaC*, lacks a promoter, Shine-Dalgarno sequence, start codon, and signal sequence, and therefore depends on transcriptional and translational fusion with N-termini of other proteins for the expression and export of 'BlaC. (B) Library 2 was constructed the same way except *M. tuberculosis* PM638 ($\Delta blaC$) genomic DNA was ligated into the *Clal* site located upstream of *'blaC* on the multicopy shuttle plasmid pJM157. Additionally, pJM157 has the mycobacterial *hsp60* promoter located upstream of the *Clal* insertion site, but lacks both a ribosome binding Shine Dalgarno site and a start codon. Both library plasmids carry a selectable hygromycin-resistance gene (*hyg*).

Table 3.2. Primers used in this study.

Target gene	Oligo name	Sequence (5' → 3')	Restriction enzyme site (5' extension)	Source
Primers used for library construction				
<i>blaC</i>	Lib'blaCfor	ATCGATCCGGCATCGACAACCTTG	<i>ClaI</i>	This work
	Lib'blaCrev	CCATGGCTATGCAAGCACACCGGCAAC	<i>NcoI</i>	This work
<i>hsp60</i>	Hsp60for-BstBI	ATTCGAACGTGGCGAACTCCGTTGTAG	<i>BstBI</i>	This work
	Hsp60rev2-ClaI	TATCGATCGGGGATGAAACGGGGTTC	<i>ClaI</i>	This work
Primers used for plasmid construction (HA-tagged constructs)				
<i>plcB</i>	PlcBMscIF	CGTGGCCATTACCCGCCGACAATTTTTTGC	<i>MscI</i>	This work
	PlcBMscF-2	CGTGGCCATTGGGCTTGCCCCGGACATTTG	<i>MscI</i>	This work
	plcBMscF-3	GTCTAGAGTCCGGTGGCATTTCGTGGTC	<i>XbaI</i>	This work
	plcBXbaF-1	GTCTAGAACCCTTCCCGCTGCGATATCC	<i>XbaI</i>	This work
	PlcBHindR	CGAAGCTTGACAGAGACCCTGGGAATCC	<i>HindIII</i>	This work
<i>Rv0315</i>	315MscF*	CGTGGCCATTCTTATGCCTGAGATGGATCG	<i>MscI</i>	This work
	315MscF-2	GTCTAGAAGAACCAGCACGCCGACAAC	<i>XbaI</i>	This work
	315HindR	CGAAGCTTGAAAGACGCGCACCCAGTCG	<i>HindIII</i>	This work
<i>Rv0519c</i>	519MscF	CGTGGCCATTCTACGGCTGGCTGTGCGGG	<i>MscI</i>	This work
	519HindR	CGAAGCTTGGACGCCATCCAAGCCAGTT	<i>HindIII</i>	This work
<i>lprQ</i>	lprQMscF	CGTGGCCATTAATAACTCCAGCACCCCCAG	<i>MscI</i>	This work
	lprQHindR	CGAAGCTTGCCACCCGGTCCGTTAGTAG	<i>HindIII</i>	This work
<i>ugpB</i>	ugpBXbaF-1	GTCTAGAACTCGCTGCAGGTGTTTCGATGTG	<i>XbaI</i>	This work
	ugpBHindR	CGAAGCTTGGCCATGCCCGCCAGCTTCC	<i>HindIII</i>	This work
<i>Rv0063</i>	63MscF	AGTGGCCATTGCGCGTGAATCTCACGCCAG	<i>MscI</i>	This work
	63XbaF-1	CTCTAGAACGCCGACGCTTACCTGTGG	<i>XbaI</i>	This work
	63HindR	CGAAGCTTGAAGTTCAGCCCGGAGAACA	<i>HindIII</i>	This work
<i>Rv0774c</i>	774MscF*	CGTGGCCATTATGGCCCCGATGCCAGAGTTG	<i>MscI</i>	This work
	774HindR	CGAAGCTTGACCTGTGAGCAGCGCGCGAAC	<i>HindIII</i>	This work
<i>Rv0846c</i>	846XbaF	GTCTAGAGTTTGGGTGAGCCACATGAC	<i>XbaI</i>	This work
	846HindR	CGAAGCTTGAGAATGTAGTCCAGGCGGG	<i>HindIII</i>	This work
<i>Rv2843</i>	2843NruF	ATCGCGATTCTTAGAGCAGCACCAGTCATC	<i>NruI</i>	This work
	2843HindR	CGAAGCTTGTATCGACGGGCTGAAGGCAC	<i>HindIII</i>	This work
Primers used for sequence determination				
<i>blaC</i>	TnBlaCout*	CAGAATGCCAACCCTCATC	n/a	This work
<i>ugpE</i>	ugpE-1	CTTGATGCGCAACCACTTTC	n/a	This work
<i>plcB</i>	plcBmid2	GCGACGGTCGGCGATATGACTTC	n/a	This work
Primers used for identification of high frequency ' <i>blaC</i> fusions				
<i>blaC</i>	pMP327F	GCTTTGGCGCCTGCCATCT	n/a	This work
	<i>blaC</i> for1	AGATCTGCGCGTCCGGCATCGACAACCTTGC	n/a	(53)
	<i>blaC</i> rev	AGATCTCCACGAGCCTATGCAAGCACAC	n/a	(53)

*primers also used for identification of high frequency '*blaC* fusions

(ii) Library 2. We constructed a second library vector that carried a constitutive mycobacterial promoter upstream of the genomic DNA insertion site. A fragment carrying the promoter and the +1 transcriptional start site from *M. tuberculosis hsp60* was amplified by PCR from pMV261 using the primers Hsp60for-BstBI and Hsp60rev2-ClaI (Table 3.2). The resulting amplicon was ligated into the multicopy mycobacterial shuttle vector pJES113, which had been linearized with *ClaI*. The final plasmid product was named pJM157 (Fig.

3.1). In contrast to the protocol described above for Library 1, genomic DNA was isolated from *M. tuberculosis* strain PM638 ($\Delta blaC$) rather than from the H37Rv wild-type strain. As described above for Library 1, 0.5 to 5.0-kbp *AciI* and *HpaII* fragments were cloned into the single *ClaI* site immediately upstream of '*blaC* in pJM157. The resulting ligation was transformed into electrocompetent *E. coli* DH5 α (Invitrogen). Approximately 8×10^5 CFU were pooled and used to isolate plasmid DNA (Qiagen).

'BlaC reporter selection assay.

200 ng of plasmid DNA from either library was electroporated into $\Delta blaS$ *M. smegmatis* strain PM759 (33, 58, 59). The resulting transformation was plated onto 7H10 agar media without tween, containing 50 $\mu\text{g/ml}$ hygromycin and carbenicillin concentrations that ranged from 35-75 $\mu\text{g/ml}$ and incubated at 37°C for a minimum of four days. Transformants were streaked for single colony purification on 7H10 media with hygromycin and further confirmed as being carbenicillin-resistant by spot test analysis. Briefly, single colonies were resuspended in 500 μl aliquots of 7H9 liquid medium with 0.1% tween. 200 μl of the suspensions were transferred to empty wells of a 96-well plate and two-fold serial dilutions were made in 7H9 with 0.1% tween. The dilutions were spotted, using a multipronged replica-inoculating device, onto 7H10 agar plates containing (i) both 50 $\mu\text{g/ml}$ hygromycin and 45 $\mu\text{g/ml}$ carbenicillin, and (ii) 50 $\mu\text{g/ml}$ hygromycin only, and incubated for a minimum of three days at 37°C. Strains were considered for further analysis if they resulted in confluent growth on plates containing both hygromycin and carbenicillin in comparison to a negative control strain carrying a promoterless and signal sequence-less *BlaC*, which fails to grow on media with carbenicillin (53) (Fig. 3.2).

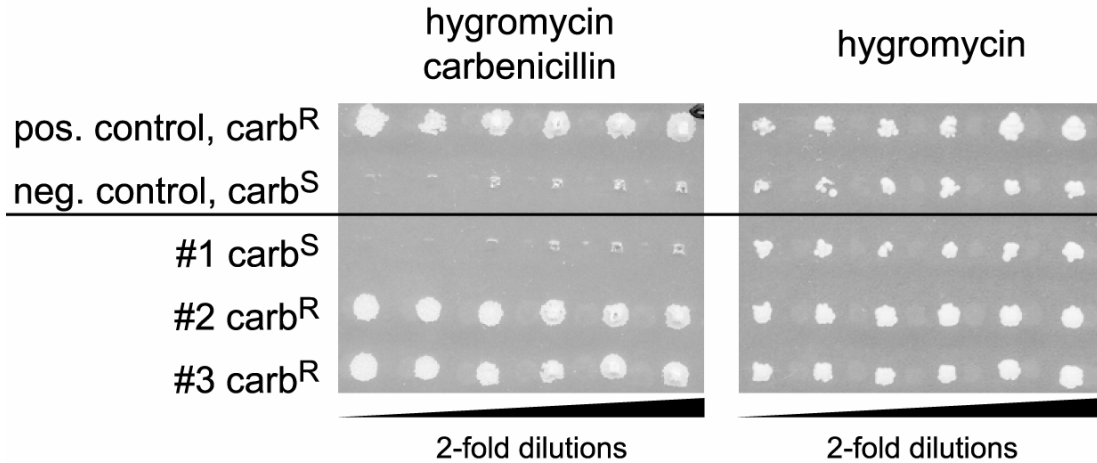


Fig. 3.2. Spot test confirmation of carbenicillin-resistant ‘BlaC fusions. A multipronged replica plater was used to spot 2-fold dilutions of *M. smegmatis* on plates with 50 $\mu\text{g/ml}$ hygromycin and 45 $\mu\text{g/ml}$ carbenicillin (left) and 50 $\mu\text{g/ml}$ hygromycin alone (right), as described in Materials and Methods. *M. smegmatis* ΔblaS expressing full length *M. tuberculosis blaC* was used as a positive control for growth on carbenicillin. *M. smegmatis* ΔblaS carrying truncated *blaC* lacking a promoter and its signal sequence (Δss) was used as a negative control for growth on carbenicillin. Transformants selected on media containing carbenicillin were assayed alongside the controls to confirm or refute the carbenicillin-resistant phenotype. A strain was considered sensitive (carb^{S}) if it failed to grow on hygromycin/carbenicillin but grew on hygromycin alone (transformant #1 above). Resistant strains (carb^{R}) grew on both antibiotics (transformants #2 and #3 above).

Recovery of 'blaC fusion plasmids.

Plasmid DNA was transferred from hygromycin-resistant *M. smegmatis* to *E. coli* DH5 α by electroduction (5). Briefly, a small amount of *M. smegmatis* was transferred from a colony or patch into 20 μ l of ice-cold 10% glycerol. The suspension was mixed by vortexing, incubated on ice for 10 minutes, and added to 40 μ l of electrocompetent *E. coli* DH5 α . The mixture was transferred to a chilled 0.2 cm gap cuvette and pulsed using conditions typical for *E. coli* electroporation (25- μ F, 200 ohms, 2.5 kV). Immediately, 1 ml of LB broth was added to the cuvette and the mixture was transferred to a 1.5 ml centrifuge tube. The tubes were shaken at 37°C for 1 hr. and the mixture plated on LB with 150 μ g/ml hygromycin. Hygromycin-resistant *E. coli* cells were obtained after 16 hours incubation at 37°C, single colony purified, and plasmid DNA was extracted (Qiagen). The *M. tuberculosis* genomic DNA insert immediately upstream of 'blaC was identified by sequencing using the primer TnBlaCout (Table 3.2).

Diagnostic PCR analysis of 'blaC fusions.

To avoid repeat sequencing of 'blaC fusions identified at high frequency, we performed diagnostic colony PCR on carbenicillin-resistant *M. smegmatis*. The primers blaCrev and TnBlaCout are specific to the 3' end of 'blaC on pJES113 and pJM157, and were used as reverse primers in the diagnostic PCRs. Forward primers used were blaCfor1/pMP327F, 315MscF, and 774MscF, and are specific to the 5' end of blaC, Rv0315, or Rv0774c, respectively (Table 3.2).

Anti-BlaC antibody.

A 6× histidine-tagged copy of *M. tuberculosis* BlaC was expressed from Y49 pTrcHisB-BlaC in *E. coli* DH5α (kind gift of Doug Kernodle), and purified by nickel affinity chromatography using HIS-Select Nickel Affinity Gel (Sigma) as described previously (88). Purified BlaC was eluted from the nickel column with 300 mM imidazole at a concentration of 1.3 mg/ml and used to immunize rabbits together with TiterMax Gold adjuvant (Sigma). Rabbit immunizations and polyclonal antisera collection was carried out by BioSource Custom Immunology Department (Hopkinton, MA).

Immunoblot analysis.

Whole cell lysates of *M. smegmatis* strains were prepared as described previously (11). Additionally, whole cell lysates of formalin-killed *M. tuberculosis* strains were prepared as described previously (52). Triton X-114 partitioning of *M. smegmatis* whole cell lysates was prepared as described previously (34). Polyclonal BlaC antiserum was used in immunoblot analysis at a dilution of 1:10,000 and anti-rabbit peroxidase-conjugated antibody was used as secondary antibody (Biorad). Monoclonal HA antiserum (Covance) was used at a dilution of 1:10,000 and anti-mouse peroxidase-conjugated antibody was used as secondary antibody (Biorad).

Expression of full length, HA-tagged *M. tuberculosis* proteins.

To investigate the export of *M. tuberculosis* proteins in *M. smegmatis* and *M. tuberculosis*, we cloned full length *M. tuberculosis* genes in frame with a sequence encoding a hemagglutinin (HA) epitope (YPYDVPDYA), immediately followed by a TGA stop

codon. The resulting C-terminal HA-tagged fusion proteins were expressed from either the constitutive *hsp60* promoter or the native promoter.

(i) *hsp60* promoter-driven HA fusions. Oligonucleotide primers were designed to amplify full length genes from *M. tuberculosis* genomic DNA. Forward and reverse primers were designed each with 5' extension sequences carrying *MscI* and *HindIII* restriction sites, respectively (Table 3.2). The resulting PCR product was first cloned into the pCR2.1 cloning vector (Invitrogen) and sequenced. The cloned gene was then digested with *MscI* and *HindIII* and the appropriate fragment was isolated and ligated into the mycobacterial shuttle vector pJSC77 (35) (digested with *MscI* and *HindIII*) which carries the *hsp60* promoter and multiple cloning site upstream of a C-terminal HA tag (Table 3.3). Due to a *MscI* site within the *Rv2843* gene, the strategy was revised in this instance and a *NruI* site was included instead of *MscI* on the forward primer (Table 3.2).

(ii) native promoter-driven HA fusions. Oligonucleotide primers were designed to amplify a fragment of *M. tuberculosis* genomic DNA which included the full length gene of interest and upstream sequence containing the putative native promoter. Forward primers were designed each with 5' extension sequences carrying *XbaI* and *HindIII* restriction sites, respectively (Table 3.2). The resulting PCR product was first cloned into the pCR2.1 cloning vector (Invitrogen) and sequenced. The cloned gene and promoter was then digested with *XbaI* and *HindIII* and the appropriate fragment was isolated and ligated into the mycobacterial shuttle vector pJSC77 (digested with *XbaI* and *HindIII*) which removes the *hsp60* promoter (Table 3.3).

Table 3.3. Plasmids used in this study.

Plasmid	Genotype	Description	Source
pCR2.1	<i>bla aph</i> ColE1	TA cloning vector	Invitrogen
pMV206. <i>hyg</i>	<i>hyg oriM</i> ColE1	Multicopy mycobacterial shuttle plasmid	
pMV261. <i>kan</i>	<i>aph P_{hsp60} oriM</i> ColE1	Multicopy mycobacterial shuttle plasmid	(81)
pMV306. <i>kan</i>	<i>aph int attP</i> ColE1	Single-copy mycobacterial shuttle plasmid	
pJSC77	<i>aph P_{hsp60} -HA oriM</i> ColE1	HA tag cloned into pMV261	(35)
pJM106	<i>cat 'blaC (M. tuberculosis) oriV ori2</i>	Predicted <i>M. tuberculosis 'blaC</i> mature sequence cloned into pCC1 (Epicentre)	(53)
pJES112	<i>bla aph</i> ColE1 ' <i>blaC (M. tuberculosis)</i>	<i>M. tuberculosis 'blaC</i> mature sequence cloned into pCR2.1	This work
pJES113	<i>hyg oriM</i> ColE1 ' <i>blaC (M. tuberculosis)</i>	<i>M. tuberculosis 'blaC</i> mature sequence cloned into pMV206. <i>hyg</i>	This work
pJM153	<i>bla aph</i> ColE1 P _{<i>hsp60</i>} (<i>M. tuberculosis</i>)	<i>M. tuberculosis hsp60</i> promoter cloned into pCR2.1	This work
pJM157	<i>hyg oriM</i> ColE1 P _{<i>hsp60</i>} , ' <i>blaC (M. tuberculosis)</i>	<i>M. tuberculosis hsp60</i> promoter cloned into pJES113	This work
Y49	<i>bla P_{trc}-blaC (M. tuberculosis)</i> ColE1	<i>E. coli</i> expression vector for <i>M. tuberculosis blaC</i>	(88)
pTrcHisB-BlaC			
pJES139	<i>bla aph</i> ColE1 <i>plcB (M. tuberculosis)</i>	<i>M. tuberculosis</i> full length <i>plcB</i> cloned into pCR2.1	This work
pJSS52	<i>aph oriM</i> ColE1 P _{<i>hsp60</i>} <i>plcB-HA</i>	1.5-kbp <i>MscI/HindIII</i> fragment containing full length <i>plcB</i> from pES139 cloned into pJSC77	This work
pJM172	<i>aph int attP</i> ColE1 P _{<i>hsp60</i>} <i>plcB-HA</i>	2.0-kbp <i>XbaI/SalI</i> fragment containing P _{<i>hsp60</i>} <i>plcB-HA</i> from pJSS52 cloned into pMV306. <i>kan</i>	This work
pJSS51	<i>bla aph</i> ColE1 ' <i>plcB (M. tuberculosis)</i>	<i>M. tuberculosis plcB</i> (no signal sequence) cloned into pCR2.1	This work
pJM171	<i>aph oriM</i> ColE1 P _{<i>hsp60</i>} ' <i>plcB-HA</i>	1.4-kbp <i>MscI/HindIII</i> fragment containing <i>plcB</i> (no signal sequence) from pSS51 cloned into pJSC77	This work
pJM168	<i>bla aph</i> ColE1 Rv0315 (<i>M. tuberculosis</i>)	<i>M. tuberculosis</i> full length Rv0315 cloned into pCR2.1	This work
pJM169	<i>aph oriM</i> ColE1 P _{<i>hsp60</i>} Rv0315-HA	890-bp <i>MscI/HindIII</i> fragment containing full length Rv0315 from pJM168 cloned into pJSC77	This work
pJM170	<i>bla aph</i> ColE1 Rv0519c (<i>M. tuberculosis</i>)	<i>M. tuberculosis</i> full length Rv0519c cloned into pCR2.1	This work
pJM173	<i>aph oriM</i> ColE1 P _{<i>hsp60</i>} Rv0519c-HA	904-bp <i>MscI/HindIII</i> fragment containing full length Rv0519c from pJM170 cloned into pJSC77	This work
pJM167	<i>bla aph</i> ColE1 <i>lprQ (M. tuberculosis)</i>	<i>M. tuberculosis</i> full length <i>lprQ</i> cloned into pCR2.1	This work
pJM175	<i>aph oriM</i> ColE1 P _{<i>hsp60</i>} <i>lprQ-HA</i>	1.3-kbp <i>MscI/HindIII</i> fragment containing full length <i>lprQ</i> from pJM167 cloned into pJSC77	This work
pJM210	<i>bla aph</i> ColE1 Rv0774c (<i>M. tuberculosis</i>)	<i>M. tuberculosis</i> full length Rv0774c cloned into pCR2.1	This work
pJM211	<i>aph oriM</i> ColE1 P _{<i>hsp60</i>} Rv0774c-HA	913-bp <i>MscI/HindIII</i> fragment containing full length Rv0774c from pJM210 cloned into pJSC77	This work
pJM212	<i>bla aph</i> ColE1 Rv2843 (<i>M. tuberculosis</i>)	<i>M. tuberculosis</i> full length Rv2843 cloned into pCR2.1	This work
pJM214	<i>aph oriM</i> ColE1 P _{<i>hsp60</i>} Rv2843-HA	547-bp <i>NruI/HindIII</i> fragment containing full length Rv2843 from pJM212 cloned into pJSC77	This work
pJM196	<i>bla aph</i> ColE1 P _{<i>native</i>} <i>plcB (M. tuberculosis)</i>	<i>M. tuberculosis</i> full length <i>plcB</i> and promoter cloned into pCR2.1	This work
pJM199	<i>aph oriM</i> ColE1 P _{<i>native</i>} <i>plcB-HA</i>	2.2-kbp <i>XbaI/HindIII</i> fragment containing full length <i>plcB</i> and promoter from pJM196 cloned into pJSC77	This work
pJM202	<i>bla aph</i> ColE1 P _{<i>native</i>} <i>plcA, plcB (M. tuberculosis)</i>	<i>M. tuberculosis</i> full length <i>plcA, plcB</i> , and promoter cloned into pCR2.1	This work
pJM203	<i>aph oriM</i> ColE1 P _{<i>native</i>} <i>plcA, plcB-HA</i>	3.5-kbp <i>XbaI/HindIII</i> fragment containing full length <i>plcA, plcB</i> , and promoter from pJM202 cloned into pJSC77	This work

Table 3.3. (Continued) Plasmids used in this study.

pJM209	<i>aph oriM</i> ColE1 P _{native} <i>plcA</i> , <i>plcB-FLAG</i>	3.5-kbp <i>XbaI/HindIII</i> fragment containing full length <i>plcA</i> , <i>plcB</i> , and promoter from pJM202 cloned into pJM208	This work
pJM197	<i>bla aph</i> ColE1 P _{native} <i>Rv0315</i> (<i>M. tuberculosis</i>)	<i>M. tuberculosis</i> full length <i>Rv0315</i> and promoter cloned into pCR2.1	This work
pJM206	<i>aph oriM</i> ColE1 P _{native} <i>Rv0315-HA</i>	1.2-kbp <i>XbaI/HindIII</i> fragment containing full length <i>Rv0315</i> and promoter from pJM197 cloned into pJSC77	This work
pJM198	<i>bla aph</i> ColE1 P _{native} <i>ugpE</i> , <i>ugpB</i> (<i>M. tuberculosis</i>)	<i>M. tuberculosis</i> full length <i>ugpE</i> , <i>ugpB</i> , and promoter cloned into pCR2.1	This work
pJM201	<i>aph oriM</i> ColE1 P _{native} <i>ugpE</i> , <i>ugpB-HA</i>	2.4-kbp <i>XbaI/HindIII</i> fragment containing full length <i>ugpE</i> , <i>ugpB</i> , and promoter from pJM198 cloned into pJSC77	This work
pJM204	<i>bla aph</i> ColE1 P _{native} <i>Rv0063</i> (<i>M. tuberculosis</i>)	<i>M. tuberculosis</i> full length <i>Rv0063</i> and promoter cloned into pCR2.1	This work
pJM205	<i>aph oriM</i> ColE1 P _{native} <i>Rv0063-HA</i>	1.8-kbp <i>XbaI/HindIII</i> fragment containing full length <i>Rv0063</i> and promoter from pJM204 cloned into pJSC77	This work
pJM213	<i>bla aph</i> ColE1 P _{native} <i>Rv0846c</i> (<i>M. tuberculosis</i>)	<i>M. tuberculosis</i> full length <i>Rv0846c</i> and promoter cloned into pCR2.1	This work
pJM215	<i>aph oriM</i> ColE1 P _{native} <i>Rv0846c-HA</i>	1.8-kbp <i>XbaI/HindIII</i> fragment containing full length <i>Rv0846c</i> and promoter from pJM213 cloned into pJSC77	This work
pMP327	<i>aph</i> P _{blaC-blaC} (<i>M. tuberculosis</i>) <i>oriM</i> ColE1	<i>M. tuberculosis blaC</i> cloned into pMV261	(53)
pJM113	<i>aph</i> P _{hsp60} - <i>blaC</i> (<i>M. tuberculosis</i>) <i>oriM</i> ColE1	<i>blaC</i> from pJM106 cloned into <i>MscI</i> -linearized pMV261	(53)
pJM117	<i>aph</i> P _{blaC-(KK)} <i>blaC</i> (<i>M. tuberculosis</i>) <i>oriM</i> ColE1	<i>M. tuberculosis ssblaC(KK)</i> cloned into pMV261	(53)

Bioinformatic identification of putative twin-arginine signal sequences.

In silico analysis of the *M. tuberculosis* genome was performed to identify putative Tat-exported proteins based on the presence of a predicted twin-arginine signal sequence at the N-terminus. Protein sequences corresponding to the 4056 predicted open reading frames (ORFs) of *M. tuberculosis* H37Rv were obtained from TubercuList (Institut Pasteur [<http://genolist.pasteur.fr/TubercuList/>]) and were entered as query sequences in the TATFIND v.1.4 (65) (<http://signalfind.org/tatfind.html>) and TatP v.1.0 (6) (<http://www.cbs.dtu.dk/services/TatP/>) search algorithms. The output for TATFIND v.1.4 is either “True” or “False” as to the whether a given peptide sequence has a predicted Tat signal sequence. TatP v.1.0 has multiple outputs including (i) potential cleavage site, (ii) potential Tat signal peptide but no Tat motif, and (iii) potential Tat signal peptide with Tat motif. For

our analysis, we chose to use the default search criteria RRx[FGAVML][LITMVF] and include only those proteins that had both a predicted Tat motif and the predicted components of a Tat signal sequence, as described for known bacterial Tat substrates. Finally, in our *in silico* analysis, we included the predicted Tat exported proteins, as defined by the TigrFAM motif (TIGR01409) (41, 74). This list was obtained directly from the TIGR website (<http://cmr.tigr.org/tigr-scripts/CMR/GenomePage.cgi?database=ntmt02>).

Results

Selection of ‘*blaC* fusions that promote resistance to carbenicillin in *M. smegmatis*.

We previously demonstrated that the predicted twin-arginine signal sequences from the cell wall-associated PlcA and PlcB phospholipase C enzymes of *M. tuberculosis* are capable of directing the export of a truncated *M. tuberculosis* β -lactamase, ‘BlaC lacking its native signal sequence (53). Upon introduction of the PlcA or PlcB signal sequence-‘BlaC chimeras into the β -lactam-sensitive strains of *M. smegmatis* (Δ *blaS*) and *M. tuberculosis* (Δ *blaC*), the strains acquired the ability to grow in the presence of the β -lactam antibiotic carbenicillin. We further demonstrated that the requirements for functional ‘BlaC export in *M. smegmatis* are the presence of (i) a functioning Tat apparatus, and (ii) a Tat-dependent signal sequence fused in-frame with ‘BlaC. This latter requirement was established based on the inability of either Tat signal sequences that had a twin-lysine substitution of the twin-arginine motif or *M. tuberculosis* Sec signal sequences to promote protection against β -lactam antibiotics. Together, this data demonstrates that ‘BlaC can function as a reporter for Tat-dependent export in *M. smegmatis*.

We used this reporter in a genetic selection for *M. tuberculosis* sequences that promote Tat export. We designed reporter libraries to identify *M. tuberculosis* open reading frames (ORFs) that, when fused in frame to truncated 'blaC, generate fusion proteins that are exported by the Tat pathway. We constructed two multicopy expression vectors, pJES113 and pJM157, that carry blaC lacking its signal sequence (referred to as 'blaC), downstream of a unique ClaI restriction site (Fig. 3.1). pJES113 lacked a promoter on the vector and pJM157 carried the constitutively active hsp60 promoter on the vector upstream of the ClaI site. We then partially digested *M. tuberculosis* genomic DNA with the ClaI-compatible endonucleases AclI or HpaII, which cut frequently in the chromosome. Genomic fragments between 0.5 and 5.0 kbp were ligated into ClaI-linearized pJES113 and pJM157 to create genomic DNA Library 1 and Library 2, respectively. Libraries 1 and 2 contained plasmids from 1×10^6 and 8×10^5 pooled *E. coli* transformants, respectively.

The reason for using two libraries is that there may be classes of ORFs only compatible with one of the vectors. We hypothesized that gene fusions not expressed in *M. smegmatis*, or missed because the gene was located downstream in a large operon and therefore unlikely to contain the native promoter, might be expressed and identified with the hsp60 promoter-containing Library 2. Conversely, gene fusions that were toxic because of expression from the hsp60 promoter might be identified by Library 1. Since the Library 2 vector lacks the hsp60 Shine Dalgarno sequence, translation of the fusion protein can only occur if the native Shine Dalgarno site of an ORF is provided on the genomic insert.

Both libraries were electroporated into $\Delta blaS$ *M. smegmatis* and transformants were plated on 7H10 agar medium containing 45 $\mu\text{g/ml}$ of carbenicillin. Viable transformants were further confirmed as being carbenicillin resistant on 45 $\mu\text{g/ml}$, as described in the

Materials and Methods section and shown in Figure 3.2. In an attempt to achieve near saturation of possible targets, a number of factors were considered in determining the total number of CFUs to be plated for each library, as described in greater detail in the Discussion. For selection using Library 1, a total of 1×10^6 CFU were plated, yielding a total of 101 carbenicillin-resistant colonies. For selection using Library 2, a total of 8.7×10^5 CFU were plated, yielding a total of 29 carbenicillin-resistant colonies. Plasmid DNA was isolated from carbenicillin-resistant strains and sequenced to determine the identity of the genomic DNA insert. In later rounds of selection, we prescreened carbenicillin-resistant strains by colony PCR for specific fusions we picked up at high frequency, to eliminate sequencing repeat fusions. From Library 1, which was constructed with wild-type H37Rv DNA, BlaC was identified in 50% of the clones (Table 3.4). Library 2 was subsequently constructed with genomic DNA from the $\Delta blaC$ strain of *M. tuberculosis*, thereby eliminating repeated identification of BlaC.

Identification of *M. tuberculosis* signal sequences that drive the Tat-dependent export of ‘BlaC.

We identified ten unique ORF-*blaC* fusions from the two libraries combined (Table 3.4). All of the ten *M. tuberculosis* fusion proteins had predicted N-terminal signal sequences that fit the tripartite structure of having a charged “N” region, an uncharged or hydrophobic “H” region, and a signal peptidase cleavage “C” region. An amino acid alignment of the identified signal sequences revealed that nine out of ten “N” regions contained Tat motifs that followed the R-R-x- ϕ - ϕ pattern (Fig. 3.3). The exception was the Rv0063 signal sequence, which lacks the arginine dipeptide, and instead has a glutamine in

Table 3.4. Tat signal sequences identified by 'BlaC fusion.

Tat signal sequence	Description	No. Times Identified	TatP	TAT FIND	Tigr FAM
Rv0063	Possible oxidoreductase	1	×
Rv0315	Possible beta-1,3-glucanase precursor	41	×	×	×
Rv0483	Probable conserved lipoprotein (LprQ)	2
Rv0519c	Possible lipase	5	×	...	×
Rv0774c	Possible lipase; short-chain alcohol dehydrogenase family	12	×	×	×
Rv0846c	Probable multicopper oxidase	1	×	×	×
Rv2068c	Class A beta-lactamase (BlaC)	63	×	×	×
Rv2350c	Phospolipase C (PlcB)	1	×	×	×
Rv2351c	Phospolipase C (PlcA)	0*	×	...	×
Rv2833c	Probable glycerol-3-phosphate-binding lipoprotein (UgpB)	3	×	×	×
Rv2843	Probable conserved transmembrane Ala-rich protein	1	...	×	×

*Demonstrated by direct testing only

the position where the second arginine would be (R-Q-T-F-L). Despite this disparity, the Rv0063 sequence otherwise had features typical of a Tat signal sequence (8) (Fig. 3.3a).

With the aid of signal sequence prediction programs, we identified putative signal peptidase cleavage sites in the ten signal sequences (Fig. 3.3a). Of the ten, five possess a predicted cysteine-containing lipobox motif (L-A/S-G/A-C₊₁) in the C-region (1, 82). Lipid attachment occurs at the conserved cysteine, which is also the first residue of the mature protein after signal sequence cleavage by the lipoprotein-specific signal peptidase, LspA (3, 69). In *E. coli*, lipidation is a prerequisite for cleavage of the signal sequence (86). It is not clear whether the above five proteins are true lipoproteins. In addition, some of the signal sequences with lipoboxes also have predicted signal peptidase I cleavage sites near the lipobox (Fig. 3.3a).

Interestingly, we observed that the site of fusion to 'BlaC always occurred very close to the predicted signal peptidase cleavage site of the identified fusions (Fig. 3.3a). Of the fusions we identified, the longest length between a predicted cleavage site and the 'BlaC

fusion junction was 34 amino acids. Never did we identify active fusions of 'blaC to the middle or 3' end of a *M. tuberculosis* ORF.

Previous data indicated that the PlcA and PlcB signal sequences were capable of exporting a functional 'BlaC in the Δ blaS background of *M. smegmatis* but not in a Δ blaS Δ tatA background (53). To test whether the fusions identified by the 'BlaC libraries were similarly exported in a Tat-dependent manner, we introduced all the fusion plasmids into Δ blaS Δ tatA *M. smegmatis*. All ten fusions failed to confer carbenicillin resistance to the Δ blaS Δ tatA strain, even after extended incubation (data not shown). This indicated that all the fusion proteins identified require the Tat pathway for export of functional BlaC.

Bioinformatic identification of putative Tat signal sequences.

Currently, there are three available bioinformatic algorithms used to predict the presence of twin-arginine signal sequences within a given protein sequence. We applied three of these programs, TatP v1.0 (6), TATFIND v1.4 (65), and TigrFAM (TIGR01409) (41, 74), to the analysis of the *M. tuberculosis* H37Rv genome sequence and compared the output to the experimental data we obtained with our 'BlaC library approach (see Materials and Methods). In addition to the ten proteins we identified from the library analysis, we also included PlcA in the *in silico* analysis since we previously demonstrated that a PlcA signal sequence-'BlaC fusion is Tat-exported (53). Of the 4056 ORFs in the *M. tuberculosis* H37Rv genome (15), 95 are predicted to encode proteins with Tat signal sequences by at least one of the three prediction programs. However, the prediction criteria for the programs differs to the extent that there is only limited overlap between the Tat signal sequences predicted by all three of the programs. Of the 95 predicted Tat proteins, only eleven are

predicted by all three programs and ten are predicted by two out of three programs (Fig. 3.4a and Table 3.4). Of the signal sequences identified experimentally to direct export of 'BlaC, six were predicted by all three programs (Rv0315, Rv0774c, Rv0846c, BlaC, PlcB, and UgpB) and three were predicted by 2/3 programs (Rv0519c, PlcA, and Rv2843) (Fig. 3.4b). The Rv0063 signal sequence, which aligns with traditional Tat signal sequences but lacks the twin-arginine dipeptide, was only predicted by TigrFAM. Not surprisingly, when we manually changed the Rv0063 input sequence from its native SRQTFLR to SRRTFLR, the protein was then predicted to be a Tat substrate by both TatP and TATFIND. Finally, we identified the LprQ (Rv0483) signal sequence in our library, which was not predicted by any of the three programs. This is most likely due to an extended "N" region upstream of the twin-arginine motif, which may preclude prediction by the existing programs. It is likely that *lprQ* is annotated incorrectly. In the pJM157 library, we obtained a clone in which the annotated start codon of *lprQ* (TubercuList, Institut Pasteur [<http://genolist.pasteur.fr/TubercuList/>]) was absent and the +1 transcriptional start of the *hsp60* promoter was located 18 bp upstream from a GTG codon which we predict is the actual start codon. The library insert also needs to include the native Shine-Dalgarno site, which supports this conclusion. The corrected LprQ protein, which lacks the N-terminal 14 amino acids in the annotated sequence, was applied to the TatP and TATFIND servers and was then shown to have a predicted Tat signal sequence by TatP but not by TATFIND.

Assessment of signal sequence processing of *M. tuberculosis* full length proteins.

The above selection identified *M. tuberculosis* sequences capable of directing the export of 'BlaC through the Tat pathway in *M. smegmatis*. As there is a strict requirement

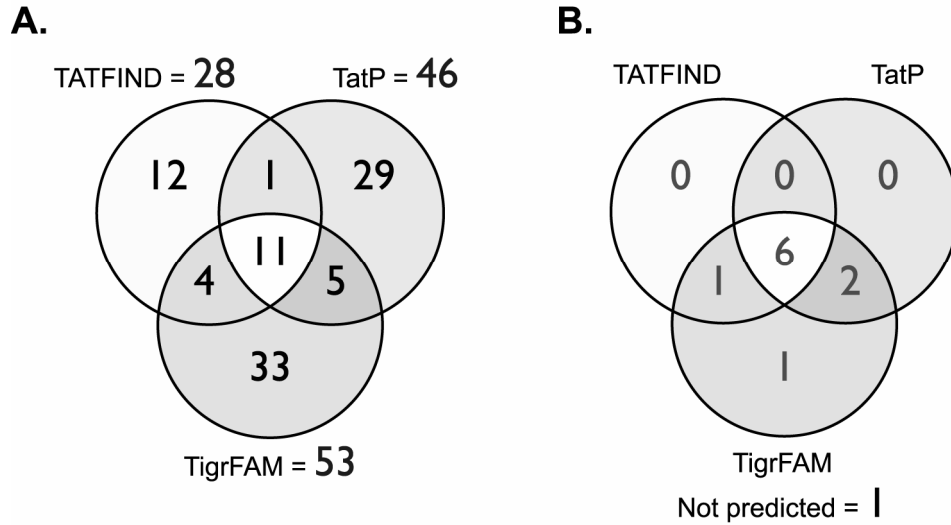


Figure 3.4. Distribution of predicted and experimentally verified *M. tuberculosis* twin-arginine signal sequences. (A) Venn diagram indicating the distribution of *M. tuberculosis* proteins predicted to have N-terminal twin-arginine signal sequences, as gathered by bioinformatic analysis using the TATFIND v.1.4, TatP v.1.0, and TigrFAM (TIGR01409) algorithms. The number of proteins predicted by TatP was 46 (29 unique), TATFIND 28 (12 unique), and TigrFAM 53 (33 unique). The number of proteins predicted by 2/3 programs was 10. The number of proteins predicted by 3/3 programs was 11. (B) Venn diagram indicating the distribution of *M. tuberculosis* proteins identified by signal sequence fusion to 'BlaC among the three prediction programs indicated. The number of demonstrated proteins predicted by 2/3 programs was 3. The number of demonstrated proteins predicted by 3/3 programs was 6. Rv0063 was predicted by TigrFAM only. LprQ was not predicted by any of the programs.

for a Tat substrate to fold prior to export (25), we reasoned that these signal sequences were responsible for targeting folded 'BlaC to the Tat translocase. However, a recent study in *E. coli* indicates that some twin-arginine signal sequences are also able to direct export of unfolded reporter proteins via the Sec pathway (87). This suggests that, in addition to the targeting determinants of the signal sequence, the folded state of the mature domain is important in determining if a protein will be exported by the pathway it is targeted to. We therefore sought to determine if the authentic full length proteins identified by our selection strategy were exported through the Tat pathway.

We chose to evaluate export of the full length *M. tuberculosis* proteins by assaying for signal sequence processing using immunoblot analysis. Here, we can look for the presence or absence of a smaller molecular weight processed form of the protein (approximate 3-4 kDa smaller) under conditions in which Tat export is functioning or not. The analysis was carried out in *M. smegmatis*, where we had the advantage of being able to test for Tat-dependent processing using the Δtat mutants. We first chose to examine export of full length BlaC, since we had previously shown enzymatically that BlaC is exported in a $\Delta blaS$ background, but not in a $\Delta blaS \Delta tata$ background, or when a conservative twin-lysine substitution was made in BlaC (BlaC[KK]) (53).

Polyclonal antibodies were generated against truncated BlaC as described in Materials and Methods and used to detect BlaC expressed in *M. smegmatis* in whole cell lysates. The anti-BlaC antibody does not cross react with BlaS of wild-type *M. smegmatis* (data not shown), or with any additional *M. smegmatis* proteins, as indicated by the absence of detectable bands in a $\Delta blaS$ strain carrying an empty vector (Figure 3.5, lane 1). In the $\Delta blaS$ background, which has an intact *tat* locus, BlaC is present as a predominant band

which runs at about 30 kDa. A minor band of about 33 kDa in size, which is the predicted size of full length BlaC is also present. In the same $\Delta blaS$ background, BlaC[KK] is only present as the larger of the two bands (Fig. 3.5, lane 4). Similarly, when BlaC is expressed in a $\Delta blaS \Delta tatA$ background, there is a notable decrease in the detectable amount of the smaller band and an increase in the larger band (Fig. 3.5, lane 5). Together this data suggests that in the absence of Tat export, there is a defect in BlaC processing, as observed by the absence or reduction of the smaller molecular weight, presumably mature species. Interestingly, in contrast to the complete lack of a processed band for BlaC[KK], we still observed the smaller band of BlaC in the $\Delta blaS \Delta tatA$ strain, albeit at a reduced level. The reason for this is unknown, but may indicate that a small amount of nonfunctional BlaC is being exported by a Tat-independent pathway. It is worthwhile to point out that the $\Delta blaS \Delta tatA$ strain grows slower than the $\Delta blaS$ strain due to the loss of *tatA* (53), and this slower growth may allow BlaC to overcome its strict Tat-dependence and interact with other export pathways.

Finally, it is worth mentioning that a truncated BlaC lacking its signal sequence (Δss) runs as a single band of similar, but not equivalent molecular weight as the processed form seen in the adjacent lane (Figure 3.5, lanes 1 and 2). BlaC is a predicted lipoprotein (82), and therefore lipidation of the exported form of BlaC may explain the slower mobility compared to the nonlipidated Δss form, which lacks the conserved lipobox motif.

Since we were able to detect Tat-dependent signal sequence processing of BlaC in *M. smegmatis*, we went on to test the remaining nine ORFs identified in the ‘BlaC selection for export as evidenced by signal sequence cleavage. We designed expression constructs that carried the full length *M. tuberculosis* gene fused at the C-terminus to a HA epitope tag. These constructs were expressed from either the constitutive mycobacterial *hsp60* promoter

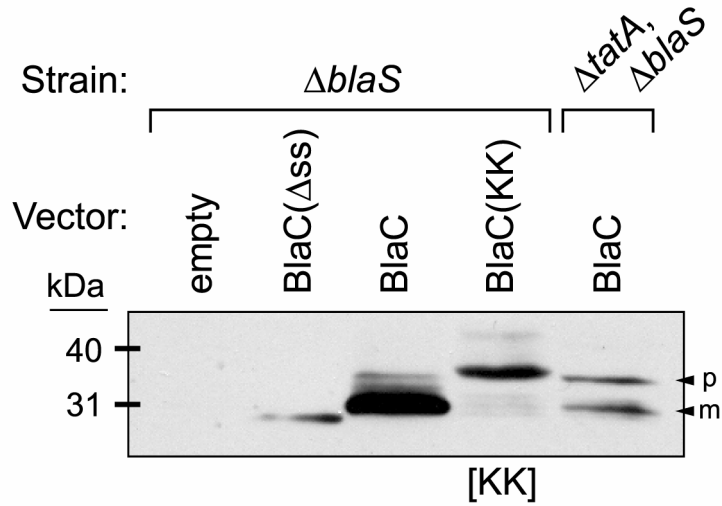


Figure 3.5. Tat-dependent processing of BlaC in *M. smegmatis*. BlaC was assessed for signal sequence processing by immunoblot analysis. Full length *blaC* or *blaC*[KK] was expressed from its native promoter in a *tat*⁺ strain ($\Delta blaS$) or a *tat*⁻ deletion strain ($\Delta blaS \Delta tatA$) of *M. smegmatis*. *blaC* lacking its signal sequence (Δss) was expressed from the mycobacterial *hsp60* promoter. Bands were detected by immunoblot analysis using anti-BlaC antibodies. The presumed precursor (p) and mature (m) forms of select proteins are indicated by arrowheads.

or their native promoter. When native promoters were used, they were chosen on the basis of information from upstream sequences present in 'blaC library fusions. The constructs were introduced into wild-type and Δtat *M. smegmatis*, and whole cell lysates were separated by SDS-PAGE and examined by immunoblot analysis with anti-HA antibodies as described above.

Eight out of nine proteins were detected by immunoblot analysis (Fig. 3.6). The exception was native promoter-driven UgpB, which we failed to detect by immunoblot in both *M. smegmatis* and *M. tuberculosis*.

Of the eight proteins tested, Rv0315, Rv0774c, and Rv2843 showed evidence of a Tat-mediated processing event. This was demonstrated by the presence of a smaller molecular weight form in wild-type and reduction of this species in the $\Delta tatC$ mutant (Fig. 3.6). In a wild-type background, Rv0315 showed three bands equally spaced with about 3-4 kDa size difference between bands. The highest molecular weight band is similar in size (~34 kDa) to that of the predicted Rv0315 precursor form and is present in both wild-type and $\Delta tatC$. Interestingly, the faster of the two smaller bands is completely absent in $\Delta tatC$, whereas the middle band is only reduced in the $\Delta tatC$ mutant, suggesting that both smaller bands are derived from the precursor form in a Tat-dependent manner. This may indicate that a second processing event is occurring during or after export. In contrast, Rv0774c and Rv2843 are represented as two bands in both wild-type and $\Delta tatC$ *M. smegmatis*; however, the smaller band is reduced in quantity in the $\Delta tatC$ mutant. Interestingly, it also appears that the overall protein content for these proteins is decreased in the $\Delta tatC$ mutant, despite each lane being loaded for equal total protein. We conclude that there is a Tat-dependent effect

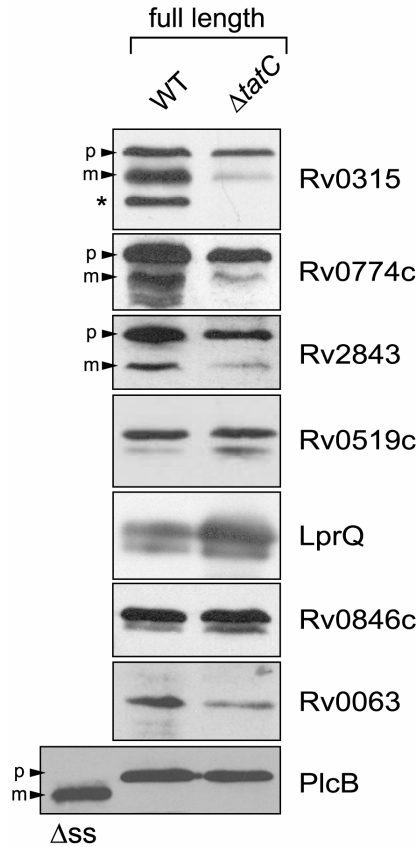


Figure 3.6. Assessment of signal sequence processing of *M. tuberculosis* full length candidate Tat substrates. The proteins indicated were assessed for signal sequence processing by immunoblot analysis. Full length genes were expressed from their native promoter (*plcB*, *Rv0063*, *Rv0315*, *Rv0846c*) or the *hsp60* promoter (Δ *ss-plcB*, *Rv0519c*, *lprQ*, *Rv0774c*, *Rv2843*) in a *tat*⁺ wild-type strain (*mc*²155) or a *tat*⁻ deletion strain (JM567, Δ *tatC*) of *M. smegmatis*. Strains were harvested at OD₆₀₀ = 1.0 for whole cell lysate preparation and equal volumes of the lysates were run on an SDS-PAGE gel. To maximize separation of bands, lysates were run on either 8% (*Rv0846c*, *Rv0063*, and *PlcB*), 12% (*Rv0315*, *Rv0774c*, *Rv0519c*, and *LprQ*), or 15% (*Rv2843*) acrylamide gels based on the size of the targeted protein. For *Rv0774c* and *Rv2843*, protein quantitation was performed to ensure that equivalent amounts of total protein were loaded between wild-type and Δ *tatC* strains. Bands were detected by immunoblot analysis using anti-HA antibodies. The presumed precursor (p) and mature (m) forms of select proteins are indicated by arrowheads. The asterisk indicates a putative proteolytic cleavage product for *Rv0315*. Native promoter-driven *UgpB* was not detected by anti-HA immunoblot.

for these three proteins, although further analysis should be conducted to more definitively establish the identity of the Tat-dependent smaller product.

Surprisingly, five out of the eight proteins displayed a profile by immunoblot analysis in which no differences in size were observed between expression in wild-type or $\Delta tatC$ *M. smegmatis*, (ie. no obvious Tat-dependence in *M. smegmatis*). Rv0483 (LprQ), Rv0519c, and Rv0846c each show two bands in both the wild-type and Δtat *M. smegmatis* backgrounds (Fig. 3.6). For LprQ and Rv0519c, the two bands are about three kDa apart and run at the expected size of a precursor/mature doublet. For Rv0846c the bands are close in size and until further analysis is performed, we cannot determine whether the smaller of the two bands is a potential mature species or a proteolytic degradation product. Rv0063 and PlcB each show only one band in both *M. smegmatis* backgrounds (Fig. 3.6). As described in more detail in the Discussion, the lack of Δtat dependence for these five proteins could indicate that (i) the protein is exported by both Sec and Tat, (ii) the protein is exported by Sec only, (iii) precursor and mature species cannot be resolved on the gel system used, and therefore the use of signal sequence cleavage as a readout for export is not adequate, or (iv) the protein is not exported by *M. smegmatis*. Further experiments are required to determine which of these possibilities is occurring.

Analysis of PlcB export in mycobacteria.

We were interested in establishing that the phospholipase C enzymes are true Tat substrates since they are proven virulence factors of *M. tuberculosis* (63). The *M. tuberculosis* Plc enzymes are good candidates for being true Tat substrates since phospholipase C enzymes are exported by the Tat pathway in a number of bacteria (26, 89).

Therefore, we focused subsequent analysis on PlcB to determine its mode of export in mycobacteria. As shown in Figure 3.6, PlcB was present as a single band in *M. smegmatis* whole cell lysates at the predicted precursor size of 57 kDa, regardless of whether the Tat pathway was present. For size comparison, we expressed an HA epitope-tagged form of PlcB lacking the predicted N-terminal signal sequence (Δ ss). As expected, the Δ ssPlcB migrates faster on the gel compared to the band seen for the full length protein (Figs. 3.6 and 3.7). This suggested that the full length protein is most likely unprocessed in both wild-type and Δ tat *M. smegmatis* backgrounds.

Because *M. smegmatis* lacks phospholipase C homologs we considered the possibility that this nonpathogenic mycobacteria may also lack accessory factors required for PlcB export in *M. tuberculosis*. To test this possibility, we expressed full length PlcB in *M. tuberculosis* H37Rv and assayed whole cell lysates by immunoblot analysis. Unlike the profile seen in *M. smegmatis*, PlcB, when expressed in *M. tuberculosis*, showed two clear bands, the larger of the two migrating at the same location as the band seen when full length PlcB was expressed in *M. smegmatis*, and the faster of the two migrating at the same location as the Δ ssPlcB band (Fig. 3.7). This suggests that the PlcB signal sequence is cleaved in *M. tuberculosis*, but not in *M. smegmatis*. Interestingly, immunoblot analysis of PlcB expressed in slow-growing *M. bovis* bacillus Calmette-Guérin (BCG) showed the same two bands as seen with PlcB in *M. tuberculosis* (data not shown), despite the lack of Plc homologs in BCG. Experiments are planned to further characterize these findings.

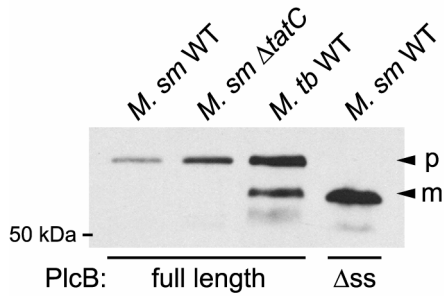


Figure 3.7. The PlcB signal sequence is putatively processed in *M. tuberculosis*, but not in *M. smegmatis*. Full length *plcB* was expressed from its native promoter and *plcB* lacking its signal sequence (Δ ss) was expressed from the mycobacterial *hsp60* promoter in a *tat*⁺ wild-type strain (mc²155) of *M. smegmatis* (*M. sm*), a *tat*⁻ deletion strain (Δ *tatC*, JM567) of *M. smegmatis*, or a *tat*⁺ wild-type strain (H37Rv) of *M. tuberculosis* (*M. tb*). Bands were detected by immunoblot analysis using anti-HA antibodies. The presumed precursor (p) and mature (m) forms of PlcB are indicated by arrowheads.

Discussion

BlaC reporter libraries.

The goal of this work was to identify the substrates of the *M. tuberculosis* Tat pathway in order to determine the role it plays in virulence. Because the Tat pathway in *M. tuberculosis* is essential and mutants cannot be made (67), the ways this can be approached are limited, preventing the possibility of a direct comparative approach in *M. tuberculosis*.

Using reporter fusions to identify exported proteins is a well established strategy. This type of selection has been previously employed using the *E. coli* TEM-1 β -lactamase to identify proteins exported by the Sec and Tat pathway in bacteria (13, 57, 75), and has recently been used in *M. tuberculosis* (52).

We previously reported that BlaC was a Tat-specific reporter (53). Here, we exploited this property in a strategy to identify Tat exported proteins of *M. tuberculosis*. To date, this represents the first unbiased use of a genetic reporter to identify Tat-exported proteins in bacteria. So far, reporters of Tat export have only been used to confirm Tat predictions (70, 84, 87). Moreover, BlaC has the capacity of being both specific for Tat export, and employable in a genetic selection.

Because of the novelty of the BlaC reporter system, this work can also address two basic aspects of Tat biology. First, this system can identify Tat signal sequences in an unbiased way without input from prediction programs. Second, by subsequent evaluation of the full length protein, it can allow us to determine how often it is that a demonstrated functional Tat signal sequence is present on a bona fide Tat substrate, or instead is present on a protein that is exported by multiple or other pathways. Both of these issues have only

begun to be investigated in *E. coli* (87), but are likely to be an important issue in understanding Tat export of proteins.

Our data shows that the BlaC reporter is a powerful system to select for Tat-exported fusion proteins. Our proof of concept experiments had been done previously (53). Here, our data with the library reinforces the power of the system: all '*blaC* fusions obtained were in-frame with an ORF, all unique 'BlaC fusions identified from the library were exported in a Tat-dependent manner, and all but one identified ORF had a twin-arginine motif.

Each library was pooled from approximately 1×10^6 plasmid-containing *E. coli*. Given (i) the 4.4 Mbp size of the *M. tuberculosis* genome, (ii) a library insert size of 500 bp, and (iii) a 1/6 chance of obtaining an in-frame translational fusion between '*blaC* and an ORF, we calculated that selecting from a total of 2.4×10^5 clones was necessary to achieve a 99% probability of obtaining in-frame '*blaC* fusions with every *M. tuberculosis* ORF (68). Therefore, we constructed the libraries to be sufficiently complex in an effort to achieve near saturation of possible targets, and assessed a total of 1.9×10^6 CFU of *M. smegmatis* for carbenicillin-resistance from both libraries combined. Analysis of the identified 'BlaC fusions revealed that only the very N-terminus of a protein was fused to 'BlaC. On average, the fusion junction was 15 amino acids beyond the predicted signal sequence cleavage site. Because of the requirement for proper folding prior to export, the most likely explanation for this is that extended mature sequence disrupts 'BlaC folding and therefore prevents export. The above calculation did not take into account this unanticipated requirement for a fusion to occur near the signal sequence cleavage site. Therefore, our calculations likely underestimate the number of library clones needed to identify all possible targets that can be identified using this system.

Further, given this observation, we conclude that only genes that have recognition sites for the *AciI* or *HpaII* endonucleases that are just beyond the signal sequence cleavage site and in-frame with that gene will be identified using the '*blaC* libraries described here. Therefore, we cannot state that the current analysis using the 'BlaC reporter is comprehensive. There are proteins predicted to be Tat substrates by Tat signal sequence prediction programs that we did not identify by random fusion to 'BlaC which may be a result of the lack of *AciI/HpaII* restriction sites within 30 amino acids of their predicted cleavage site. *Rv2041c* is an example of a gene that falls within this category. Because we now realize it would be missed by the library, a direct demonstration of potential Tat substrates like *Rv2041c* is required to determine if they are Tat-exported proteins. Alternatively, modification of the 'BlaC fusion library to circumvent the requirement for specific restriction sites near the cleavage site may result in the identification of additional Tat-exported proteins.

In total, we selected 130 carbenicillin-resistant *M. smegmatis* colonies carrying active '*blaC* fusions to ten unique genes of *M. tuberculosis*. Of the ten, all have predicted signal sequences. Five of these genes encode proteins (*Rv0063*, *Rv0315*, *Rv0774c*, *BlaC*, and *PlcB*) that have been reported to be secreted or cell wall-associated by proteomics (36, 39, 50, 63, 73). Only *BlaC* and *PlcB* have demonstrated functions (63, 88).

Of the ten unique ORFs identified from both libraries, a small number were picked up at high frequency. Using Library 1, which was constructed with wild-type *M. tuberculosis* genomic DNA, we identified *BlaC* repeatedly. We suspect many of these plasmids carried the *blaC* gene intact. This led us to prescreen carbenicillin-resistant isolates by PCR for *blaC* prior to sequencing. In addition, we designed Library 2 using Δ *blaC* genomic DNA to avoid

identifying BlaC. We also identified fusions to the Rv0315 signal sequence a total of 41 times from both libraries, and the length of the N-terminal fusions showed the greatest variation ranging from 25 to 53 amino acids. These were also later prescreened by PCR.

Five of the ten signal sequences identified from the library are predicted to be lipoproteins. This is not unusual, as several demonstrated Tat-exported proteins in bacteria carry putative lipoboxes (92); however, there is a lack of evidence demonstrating that these proteins are in fact lipidated and anchored to the membrane. A link between Tat export and lipoprotein processing has yet to be established (71). BlaC is a good candidate for this analysis since it has lipobox and we have shown that it is Tat-exported. Additionally, we have observed that BlaC partitions to the lipoprotein-rich detergent phase when treated with Triton X-114, both in *M. tuberculosis* and *M. smegmatis* (H.S. Gibbons, unpublished data and Supplemental Figure 3.8), and this localization is dependent on its native signal sequence in *M. smegmatis*, suggesting it may be lipidated. Further analysis is required to determine if BlaC is a true lipoprotein.

Analysis of Tat-dependent processing of *M. tuberculosis* full length proteins.

We embarked on examination of the export of each of the ten authentic full length proteins in *M. smegmatis* wild-type and Δtat . We undertook this analysis since several putative *E. coli* Tat signal sequences were capable of directing export of a reporter to either the Sec or Tat pathway, depending on the unfolded or folded state of the reporter (87). However, analysis of several of the authentic full length proteins showed them to be exported exclusively by Tat. This suggests that, in *E. coli*, the export of native proteins that carry twin-arginine signal sequences may not reflect what is observed when the respective signal

sequence is fused to a folded or unfolded reporter enzyme. For our analysis, full length proteins were expressed with a C-terminal HA tag in *M. smegmatis* and export was assessed by looking for evidence of Tat-dependent signal peptide cleavage by immunoblot analysis. Four proteins (BlaC, Rv0315, Rv0774c, and Rv2843) showed a Tat-dependent effect, with disappearance of smaller bands in Δtat consistent with export and signal sequence cleavage being Tat dependent. Surprisingly, the remaining five proteins showed no obvious Tat effect, with identical immunoblot profiles in both wild-type and Δtat *M. smegmatis* backgrounds.

There are several possible explanations for why we see no Tat-dependent difference for these five proteins: (i) The protein may be exported by both Sec and Tat, in which case blockage of Tat-export in Δtat does not affect Sec export of the protein. (ii) The protein may be exported by Sec only. (iii) There may be technical problems with the assay and we may not be able to distinguish between the precursor or mature form of the protein by immunoblot. In this case, expression of a truncated control lacking a signal sequence should help reveal this to be the problem. Alternatively, another assay can be used. (iv) The full length protein is not exported by *M. smegmatis*. This is a possibility since we are expressing a *M. tuberculosis* protein in a heterologous system. At this point, we cannot draw any definitive conclusions about these five proteins and further analysis is required to determine if the full length proteins identified from our library are exported by Tat, Sec, both, or none in *M. smegmatis* and *M. tuberculosis*.

Although there are examples of cross-species export by the Tat pathway (80), a number of reports suggest that recognition of a signal sequence by a heterologous host does not always occur (9, 77). In support of this conclusion, we further analyzed PlcB, which lacks a *M. smegmatis* ortholog. We showed that when PlcB was expressed in its native host,

M. tuberculosis, it yielded two species by immunoblot analysis, the faster of the two migrating at the same molecular weight as the expected processed PlcB lacking its signal sequence. To prove this is the processed form of PlcB, we would have to demonstrate the absence or reduction of the smaller band in a strain defective for export of PlcB, ideally in a *tat* mutant; however, this is not an option. It is likely that if PlcB is a Tat substrate in *M. tuberculosis*, a KK substitution of the twin-arginine consensus motif would prevent it from being exported in *M. tuberculosis*. This experiment is underway. We hypothesize that a specific factor is necessary for the export of PlcB in *M. tuberculosis*, which is absent in *M. smegmatis*.

Although additional experiments are necessary, these preliminary results suggest the attractive possibility that PlcB is a *M. tuberculosis*-specific Tat substrate. There are a small number of examples of Tat substrates shown to require dedicated chaperones for export (43, 56). Although the exact function of these chaperones remains to be discerned, proposals for how they may work include (i) protecting the Tat substrate from degradation or premature delivery to the translocon before folding is complete, (ii) delivering the folded substrate to the translocase, or (iii) promoting proper folding (56). Since some of these functions relate to the mature domain of the protein, this may account for why the full-length PlcB and not the ssPlcB-‘BlaC failed to be exported by *M. smegmatis*. Notably, PlcH of *P. aeruginosa* requires PlcR chaperones for its export (19). However, there are no obvious PlcR homologues in *M. tuberculosis*.

Similar to PlcB, we speculated that some of the other proteins we analyzed utilize *M. tuberculosis*-specific chaperones specific to the mature domains of the proteins. Although most previous reports describe dedicated chaperones as mentioned above, there is evidence

that DnaK plays a role as a general chaperone in folding of Tat substrates in *E. coli* (38, 43, 56, 60). Like PlcB, we fail to detect a Tat-dependent effect in *M. smegmatis* for the Rv0063 and Rv0846 proteins, which lack respective orthologs in *M. smegmatis*, and are therefore also candidates for having a *M. tuberculosis*-specific chaperone.

Putative functions of the genes identified to date from the BlaC fusion library.

Here, we provide a more detailed description of the ten genes identified by the selection described in this work, specifically in regard to their demonstrated or putative functions based on sequence homology and their potential justification for export by the Tat pathway.

BlaC (Rv2068c). Mycobacteria and a small number of nonmycobacterial species have β -lactamases with predicted Tat signal sequences. To date, the only β -lactamases demonstrated to be native Tat substrates are BlaS of *M. smegmatis* and BlaC of *M. tuberculosis* (53). In this work, we have further shown that BlaC is processed in a Tat-dependent manner. Due to the intrinsic properties of Tat export, it is reasonable to suggest that BlaS and BlaC simply fold too quickly in the cytoplasm and are therefore incompatible with export through the Sec apparatus (25). Alternatively, these proteins may bind cofactors or associate with additional proteins in the cytoplasm prior to export. The crystal structure of BlaC was recently determined to high resolution and its structure confirmed homology-based hypotheses that BlaC is a class A β -lactamase (40, 90). Although some β -lactamases bind cofactors, Class A β -lactamases do not.

PlcA (Rv2351c) and PlcB (Rv2350c). We previously demonstrated that the signal sequences from the phospholipase C enzymes, PlcA and PlcB, drove the export of the BlaC

reporter in a Tat- and twin-arginine-dependent manner (53). As expected, we selected a random PlcB fusion to 'BlaC from our fusion library that promoted growth on carbenicillin. We also predicted we would identify a PlcA fusion; however, we did not. It is possible that, as described above, a random PlcA-'BlaC fusion was not compatible with our selection due to the lack of an appropriately positioned *Acil/HpaII* site within the *plcA* gene. Alternatively, it is possible that the libraries were not complex enough. To date, we have not directly tested a 'blaC fusion to *plcC*, which lies downstream of *plcB*, although it is a close homolog to the other two genes and has a predicted Tat signal sequence.

PlcA and PlcB, have been found in subcellular cell wall fractions of *M. tuberculosis* and have been shown to contribute to virulence in mice (63). Phospholipase C homologs are recognized as having potential Tat signal sequences in many bacteria, and may therefore represent a category of cofactorless proteins that, for unknown reasons, has adapted to export by the Tat pathway (26). Further, a number of phospholipase C enzymes, including those with homology to PlcA and PlcB of *M. tuberculosis*, are demonstrated Tat substrates (12, 66, 89).

The *plcABC* locus is absent in *M. smegmatis* and no orthologs to the *plc* genes exist elsewhere on the *M. smegmatis* chromosome (<http://www.tigr.org/>). As mentioned above, preliminary evidence suggests that PlcB is processed in *M. tuberculosis*, but not in *M. smegmatis* and we propose that a chaperone required for export and/or processing of PlcB is present in *M. tuberculosis*, but absent in *M. smegmatis*.

Rv0519c and Rv0774c. Rv0519c and Rv0774c are homologous proteins, sharing 54% amino acid sequence identity throughout the length of both proteins, suggesting that the two proteins have similar cellular functions. The 'BlaC fusions we obtained to Rv0519c were the

longest in length, having between 74 to 82 amino acids of the N-terminus of Rv0519c fused to 'BlaC. Both proteins each have a predicted transmembrane domain at the N-terminus (78) that overlaps with the hydrophobic domain of the signal sequence. Since the signal sequences have predicted cleavage sites, it is not clear if these proteins are true membrane proteins. However, Rv0774c is detected in the culture filtrate of wild-type *M. tuberculosis* and is therefore presumably secreted and not an integral membrane protein (50). Further, Rv0774c appears to show a Tat-dependent effect in *M. smegmatis*, unlike Rv0519c, which may further indicate that the proteins are localized differently (Fig. 3.6). Rv0519c and Rv0774c are 69 and 80% identical at the amino acid level to a single ortholog in *M. smegmatis* (Msmeg5851), which has a predicted twin-arginine signal sequence.

Both *Rv0519c* and *Rv0774c* are reported as nonessential genes by TraSH analysis (72). Sequence analysis for both proteins (31, 45) reveals putative domains belonging to esterase D and lipase families. In addition, Rv0519c and Rv0774c show some homology to the Ag85 family of mycolyltransferases (FbpA, FbpB, and FbpC) including having conserved Ser, Asp/Glu, and His active site residues. Together, this information has led Takayama *et al.* (83) to propose that these proteins may function as mycolyltransferase enzymes, involved in the transfer of mycolic acid from Myc-PL to trehalose 6-phosphate on or close to the inner membrane surface.

Rv0315. As mentioned above, the Rv0315 signal sequence was identified more times than any fusion, except BlaC. Rv0315 has been identified in two independent analyses of *M. tuberculosis* culture filtrates (50, 73). Additionally, Dubnau *et al.* (28) identified *Rv0315* using a promoter trap for genes preferentially expressed by *M. tuberculosis* in the mouse lung, which suggests that its export is important during infection. Rv0315 shows Tat-

dependent processing in *M. smegmatis*. Rv0315 is 45% and 53% identical at the amino acid level to Msmeg5345 and Msmeg0645, respectively. The latter of the two *M. smegmatis* orthologs has a predicted Tat signal sequence.

Rv0315 is reported as nonessential by TraSH analysis (72). The function of Rv0315 is unknown, although it is a glycosyl hydrolase family 16 member and has homology to secreted β -1,3-glucanases of some yeast-lytic actinomycetes, which use these enzymes to lyse yeast cells by degrading the β -1,3-glucan components of the yeast cell wall. Outside of the mycobacteria, Rv0315 has 44% amino acid similarity to the demonstrated β -1,3-glucanase (β gIII_A) of the yeast-lytic actinomycete *Cellulosimicrobium cellulans* (29, 30) (Supplemental Fig. 3.9). β gIII_A has consecutive arginines at its N-terminus that align with the twin-arginine motif of Rv0315; however, β gIII_A is not predicted to be a Tat substrate by TATFIND or TatP, and instead has a predicted transmembrane domain where the signal sequence cleavage site would be. Previous analysis of a secreted β -1,3-glucanase of *Bacillus circulans* demonstrated that the enzyme is proteolytically modified post-export, resulting in sequential loss of the N-terminus (91). A similar event in the processing of Rv0315 may help to explain the smallest band on the immunoblot seen when Rv0315 is expressed in wild-type *M. smegmatis* (Fig. 3.6).

Rv0846c. Rv0846c is a predicted lipoprotein with homology to bacterial multicopper oxidases and has a putative copper-binding site. *Rv0846c* was not included in the comprehensive TraSH analysis of *M. tuberculosis in vitro* (72), and therefore data is not available as to whether *Rv0846c* is essential for *in vitro* or *in vivo* growth. Rv0846c showed no obvious Tat-dependence in *M. smegmatis* and, like PlcB, lacks an ortholog in *M.*

smegmatis, leaving open the possibility that a *M. tuberculosis*-specific chaperone may be involved in Rv0846c Tat export.

Multicopper oxidases play a critical role in copper homeostasis and iron metabolism in bacteria (47). These are periplasmic enzymes that bind copper in the cytoplasm prior to export of the protein, and are therefore often Tat exported. Its homolog in *E. coli*, *cueO* (formerly *yacK*), is a demonstrated Tat-exported multicopper oxidase and is part of a copper-regulatory operon that has been shown to oxidize a wide variety of substrates including 2,6-dimethoxyphenol, enterobactin, and ferrous iron (37, 47, 79). Due to the presence of an N-terminal Tat signal sequence and homology to other multicopper oxidases, Rv0846c is a good candidate to fulfill this role in *M. tuberculosis*. It is also possible that, like other multicopper oxidases, Rv0846c may have an additional role in iron import.

UgpB (Rv2833c). UgpB of *M. tuberculosis* is a predicted lipoprotein homologous to periplasmic substrate-binding proteins of other bacteria, particularly those involved in glycerol-3-phosphate acquisition and uptake via ATP-binding cassette (ABC) transporters. It is categorized as an essential gene *in vitro* by TraSH analysis (72), which makes it a candidate for explaining why the Tat pathway is essential in *M. tuberculosis*. UgpB has been identified in membrane fractions of wild-type *M. tuberculosis* by two-dimensional electrophoresis (51), which is consistent with it being a membrane-bound lipoprotein. We were not able to express full length UgpB. Until we overcome this problem, we cannot address the issue of full length export. UgpB is 40% similar to the product of the *Rv2041c* gene, which is also a predicted substrate binding protein of a putative sugar import system (Supplemental Fig. 3.10). We did not identify a Rv2041c-‘BlaC fusion in our library. As mentioned above, we predict that our library does not contain a Rv2041c-‘BlaC fusion

protein that is compatible for Tat-export since *Rv2041c* lacks *Acil/HpaII* restriction sites within 30 bp of the predicted signal peptide cleavage site. UgpB has low homology (~40% similar) to two *M. smegmatis* proteins, Msmeg5574 and Msmeg0515 that both lack a twin-arginine signal sequence.

Recently, Titgemeyer *et al.* (85) described the Ugp system as one of four putative ABC transport systems likely involved in sugar import in *M. tuberculosis*. Interestingly, protein motif analysis of *M. tuberculosis* UgpB additionally classifies the protein as a Fe/S Rieske subunit of ubiquinol-cytochrome c reductase. Members of this family of proteins in bacteria often have Tat signal sequences for export with the Fe/S group already bound, and may additionally be anchored to the membrane by an uncleavable signal sequence (2, 22, 54).

LprQ (Rv0483). *Rv0483/lprQ* encodes a putative lipoprotein with no predicted function. LprQ is 81% similar to Msmeg0929, which has a predicted Tat signal sequence.

LprQ belongs to the ErfK/YbiS/YcfS/YnhG family of bacterial proteins whose function is largely unknown but has members that contain putative peptidoglycan binding domains. Interestingly, immediately upstream of *lprQ* is *Rv0482/murB*, which is predicted to encode a potential fatty acid desaturase (FAD)-binding enzyme involved in peptidoglycan biosynthesis. The *murB* and *lprQ* genes are in the same orientation and separated by only 62 bp. It is attractive to speculate that MurB and LprQ interact, perhaps to transfer peptidoglycan from MurB to LprQ for export to the cell wall. Another possibility is that MurB acts as a chaperone for LprQ, functioning in the quality control mechanism of cofactor insertion and folding of the substrate protein prior to Tat export. Alternatively, MurB may

form a complex with LprQ and use a “piggy-back” mechanism for export through the Tat translocase.

Rv2843. Like LprQ, Rv2843 encodes a putative lipoprotein with no predicted function. Of the ten genes identified from the ‘*blaC*’ library, only *Rv2843* and *lprQ* have orthologs in the *Mycobacterium leprae* genome that are not pseudogenes. Since *M. leprae* has undergone reductive evolution and consequently has a large number of pseudogenes, the subset of true genes are thought to be important for growth *in vivo* and pathogenesis (17). Our analysis (Fig. 3.6) revealed that Rv2843 shows a Tat-dependent effect in *M. smegmatis*. Rv2843 is 40% similar to Msmeg2623 in *M. smegmatis*, which has a predicted Tat signal sequence.

Rv2843 is predicted to encode a 181 amino acid alanine-rich protein with no homology outside of mycobacteria. The protein has two predicted transmembrane domains at residues 20-42 and 62-84 (78). ‘*BlaC*’ fusion to Rv2843 occurred at residue 66, although the predicted lipoprotein cleavage site is between residues 36 and 37. At this point it is unclear if Rv2843 is an integral membrane protein. *Rv2843* overlaps by 4 bp with *Rv2844*, which is predicted to encode a 162 amino acid alanine-rich protein of unknown function. A transposon insertion in the downstream gene, *Rv2844*, in a Δ *blaC* background, yields a strain that is hypersusceptible to the β -lactam antibiotic ceftriaxone (32). *Rv2844* does not have a predicted N-terminal signal sequence. Together, this data suggests that the two proteins could perhaps interact and function in the export of a cell envelope component that is required for β -lactam antibiotic resistance (32). In this scenario, *Rv2844* could be passively co-exported, or ‘piggy-backed’, as a complex with Rv2843 through the Tat pathway – a mechanism that has been described for some Tat substrates (64).

Rv0063. The identification of an active 'BlaC fusion to Rv0063 was intriguing since it was the only fusion protein that lacked consecutive arginines in the "N" region of the signal sequence. Despite this anomaly, the Rv0063 signal sequence fit the tripartite pattern of a typical signal sequence. Since Rv0063-'BlaC was exported in a Tat-dependent manner, we conclude that the signal sequence of Rv0063 is competent for Tat export. It joins a small number of examples of bacterial Tat signal sequences that do not fit the twin-arginine pattern. This category includes the TtrB subunit of tetrathionate reductase from *Salmonella enterica*, which has a lysine-arginine dipeptide (44) and the Tat-exported prepropenicillin amidase from *E. coli* has an asparagine residue between the two arginines in the Tat motif (46). These examples of non-RR-dependent export by the Tat pathway suggest that additional features of the signal sequence are important in targeting. They may also demonstrate the existence of individual Tat substrate characteristics specific to a particular organism. Further identification of the exceptions to this rule may help in identifying these additional features.

In addition to these naturally occurring proteins, Tat-exported proteins with conservative substitutions of the RR dipeptide to KK are not exported by the Tat pathway. However, a single substitution of one of the arginines is sometimes recognized by the Tat apparatus. Finally, the presence of a glutamine residue in the second position of the Rv0063 Tat motif is not unprecedented. In fact, a substitution of the second arginine for glutamine in the twin-arginine motif of the *E. coli* TorA signal sequence did not prevent export of the GFP reporter to which it was fused (24).

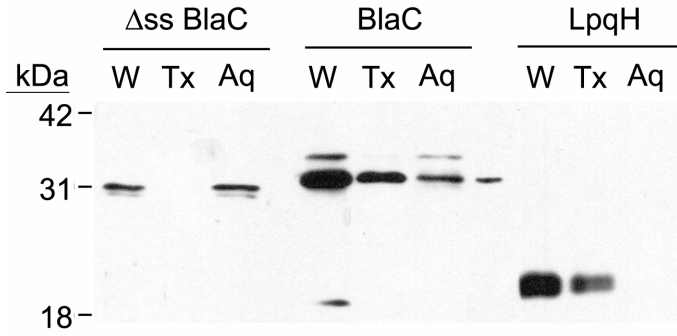
Rv0063 was identified by Malen *et al.* (50) in the culture filtrate of *M. tuberculosis* using two-dimensional gel electrophoresis. Rv0063 is 40% similar to MsmeG1746, which is about the same size, but lacks a signal sequence.

The function of Rv0063 is unknown, although it has been annotated as a putative oxidoreductase based on domain recognition signatures present in the primary amino acid sequence, including that of a potential FAD-binding protein. In this instance, insertion of the FAD cofactor may precede targeting and export of Rv0063, and therefore serve as a requirement for Tat-dependent export of the folded protein. Interestingly, MurB, an enzyme potentially involved in peptidoglycan synthesis, is in the same family of FAD-linked oxidoreductases as Rv0063. As mentioned above, *murB* is located immediately upstream of the *lprQ* gene and could act as a chaperone for Rv0063 and/or LprQ prior to Tat export.

Summary.

The use of a Tat-specific reporter enzyme in conjunction with a genetic fusion library has led to the identification of *M. tuberculosis* proteins having signal sequences capable of Tat-export. In addition to β -lactamase and phospholipase C, we identified eight previously uncharacterized proteins with potential functions that include nutrient import, copper homeostasis, and cell envelope metabolism. Further work is required to determine if these proteins are exclusively exported by the Tat pathway. Demonstration of the true functions of these proteins and their specificity of export with regard to the Sec and Tat pathways can help uncover the overall role that the Tat pathway plays in *M. tuberculosis*, including its potential importance in virulence. Although we believe the selection strategy described here was performed to near saturation, we uncovered unexpected restrictions that may have

resulted in the exclusion of some fusion proteins. A modified library approach or additional methods can be applied, thereby allowing the comprehensive identification of Tat-specific signal sequences of *M. tuberculosis*. Together, this approach can contribute significantly to an increased understanding of how the Tat pathway functions in bacteria.



Supplemental Figure 3.8. BlaC is localized to the detergent phase during Triton X-114 partitioning of *M. smegmatis*. Full length *blaC* was expressed from its native promoter and *blaC* lacking its signal sequence (Δ ss) was expressed from the mycobacterial *hsp60* promoter in the Δ *blaS* strain of *M. smegmatis*. As a control, *M. tuberculosis* LpqH (19 kDa lipoprotein) of was expressed in wild-type *M. smegmatis*. French-pressed whole cell lysates (W) were partitioned using Triton X-114 into detergent (Tx) and aqueous (Aq) phases as described previously (34). Bands were detected by immunoblot analysis using anti-BlaC or anti-LpqH antibodies.

```

Rv0315    1  ---MLMPEDRRR-----MMMMAGFGALAAALPAPTAWADPSRPAAPAG---PTPAPAAP
BglIIIA   1  MSLPHEPSSPSRRTLTLILAAAAGLALVAAWIVIATRSSPPTSPTTEGGQVTTAPANDP

Rv0315    50 AAATGG-LLFHDEFDGPAGSVDPDSKWQVSNHRTPIKNPVGFDRPQFFGQYRDSRQNVFL
BglIIIA   61 TAVTPESLAWSDEFDGAAGSAPNPDVWNHETG-----AGGWGNAELQNYTTSRVNSAL

Rv0315   109 DGNSNLVLRATRE--GNRYFGGLVHGLWRGGIGTTWEARIKFNCLAPGMWPAWWLS----
BglIIIA  114 DGQGNLVITALQESDGSYTSARLTTQGNVQPFQGRIEARIQIPRGO-GIWSAFWMVGANL

Rv0315   163 -NDDPGRSGEIDLIEWYGNGTWPSGTTVHANPDGTAFETCPIGVDGGWHNWRVTWNPSGM
BglIIIA  173 PDTPWPTSGEIDIMENVGNAPHEVHGTVHG--PGYSGDNGIMGTYQHPQGW SFADDFHTF

Rv0315   222 YFWLDYADGIEPYFSVPATGIEDLNEPIREWPFNDPGYKVPVNLNLA VGGSGGDP-ATG
BglIIIA  231 GIDWTPGEITWLV DQGEYHRVTTADV GANQWVFDQP---FFLILNVAIGQWP GNP DATT

Rv0315   281 SYPQEMLVDWVRVF-----
BglIIIA  288 PFPQQMKVDYVRVYDNATQ

```

Supplemental Figure 3.9. Rv0315 of *M. tuberculosis* is homologous to the secreted β -(1,3)-glucanase of the yeast-lytic *Cellulosimicrobium cellulans*. An alignment of the amino acid sequence of Rv0315 and β glIIA of *Cellulosimicrobium cellulans* using ClustalW reveals 44% amino acid similarity over the length of each protein.

UgpB 1 --MDPLNRRQFLALAAAAAGVTAGCAGMGGGGSVKSGSGPIDFWSSHPGQSSAAERELIG
Rv2041c 1 MVNKPFFERRSLLRGAGALT--AASLAPWAAGCAADDDALTFFFAANPDELPRMRVNVN-

UgpB 59 RFQDRFPTLSVKLIDAGKDYDEVAQKFNAALIGTDVPDVVLLDDRWWFHFALSGVLTALD
Rv2041c 58 EFQRRYPDIKVRALLSGPG---VMQQLATFCAGGKCPDVLMAWELTYAELADRGVLLDLN

UgpB 119 DLFG----QVGVDTTDYVDSLLADYEFNGRHAYVPYARSTPLFYYNKAAWQQAGLPDRGP
Rv2041c 115 TLLARDQAFAAELKSDSIGALYETFTFNGGQYAFPEQWSGNFLFYNKQLFDDAGVPPP-P

UgpB 175 QSWS---EFDEWGPQLQRVVVGAGRSAHGWANADLISWTFQGPNWAFGGAYSDKWTLTLTE
Rv2041c 174 GSWERPWSFAEFLDAAQALTKQGRSGRDRQWGFVNAWVSFYAAGLFAMNNGVPWSVPRMN

UgpB 232 P-----ATIAAGNFYRNSIHGKGYAAVAN-----DIANEFATGILASAVASTGSLAG
Rv2041c 234 PTHLNFDHDGFLEAVQFYADLTNKHKVPASAAEQSMSTADLFSVGKAGIALAGHWRYQT

UgpB 279 ITASARFDFGAAPLPTGPDAAAPACPTGGAGLAIPAKLSEERKVNALKFIAFVTNPTNTAY
Rv2041c 294 FDRADGLDFDVAPLPIGPRGRAACSDIGVTGLAIAATSR-RKDQAWFVKFATGPVQAL

UgpB 339 FSQQTGYLPVRKSAVDDASERHYLADNPRARVALDQLPHTRTQDYARVFLPGGDRIISAG
Rv2041c 353 IGESRLFVPVLRSAIN----SHGFANARRVGNLAVLSEGPAYSEGLPVTPAWEKIAALM

UgpB 399 LESIG--LRGADVTKTFTNIQKRLQVILDRQIMRKLAGHG
Rv2041c 409 DRYFGPVLRGSRPATSLTGLSQAVDEVLRNP-----

Supplemental Figure 3.10. UgpB and Rv2041c of *M. tuberculosis* are homologous. An alignment of the potential sugar transport binding proteins UgpB and Rv2041c using ClustalW reveals 40% amino acid similarity over the length of each protein.

Acknowledgements

This research was funded in part by grants from the National Institute of Allergy and Infectious Disease of the NIH (AI54540 to M.B) and a developmental grant from the University of North Carolina at Chapel Hill Center for AIDS Research (NIH #9P30 AI50410-04). J.A.M. was also supported by the Training in Sexually Transmitted Diseases and AIDS NIH NIAID training grant 5T32 AI07001-28 and a Society of Fellows dissertation completion fellowship from the Graduate School at the University of North Carolina at Chapel Hill. J.R.M was also supported by an NIH Cell and Molecular Biology training grant (NIH 5-T32-GM008581). We would like to thank Doug Kernodle for providing us with the HIS-tagged BlaC expression vector and Ruth Silversmith for advice on purification of HIS-tagged BlaC.

Attributions

The work described here was a collaborative effort. I developed the system and was involved in all aspects, and was further responsible for the analysis of the full length proteins exclusively. Significant contributions were made by Jessica McCann, Erin McElvania TeKippe, and Jason Silverman.

References

1. **Babu, M. M., M. L. Priya, A. T. Selvan, M. Madera, J. Gough, L. Aravind, and K. Sankaran.** 2006. A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *J Bacteriol* **188**:2761-73.
2. **Bachmann, J., B. Bauer, K. Zwicker, B. Ludwig, and O. Anderka.** 2006. The Rieske protein from *Paracoccus denitrificans* is inserted into the cytoplasmic membrane by the twin-arginine translocase. *FEBS J* **273**:4817-30.
3. **Banaiee, N., E. Z. Kincaid, U. Buchwald, W. R. Jacobs, Jr., and J. D. Ernst.** 2006. Potent inhibition of macrophage responses to IFN-gamma by live virulent *Mycobacterium tuberculosis* is independent of mature mycobacterial lipoproteins but dependent on TLR2. *J Immunol* **176**:3019-27.
4. **Bange, F. C., F. M. Collins, and W. R. Jacobs, Jr.** 1999. Survival of mice infected with *Mycobacterium smegmatis* containing large DNA fragments from *Mycobacterium tuberculosis*. *Tuber Lung Dis* **79**:171-80.
5. **Baulard, A., C. Jourdan, A. Mercenier, and C. Locht.** 1992. Rapid mycobacterial plasmid analysis by electrofusion between *Mycobacterium* spp. and *Escherichia coli*. *Nucleic Acids Res* **20**:4105.
6. **Bendtsen, J. D., H. Nielsen, D. Widdick, T. Palmer, and S. Brunak.** 2005. Prediction of twin-arginine signal peptides. *BMC Bioinformatics* **6**:167.
7. **Berks, B. C.** 1996. A common export pathway for proteins binding complex redox cofactors? *Mol Microbiol* **22**:393-404.
8. **Berks, B. C., F. Sargent, and T. Palmer.** 2000. The Tat protein export pathway. *Mol Microbiol* **35**:260-74.
9. **Blaudeck, N., G. A. Sprenger, R. Freudl, and T. Wiegert.** 2001. Specificity of signal peptide recognition in tat-dependent bacterial protein translocation. *J Bacteriol* **183**:604-10.
10. **Braunstein, M., S. S. Bardarov, and W. R. J. Jacobs.** 2002. Genetic methods for deciphering virulence determinants of *Mycobacterium tuberculosis*, p. 67-99. In V. L. Clark and P. M. Bavoil (ed.), *Methods in Enzymology*, vol. 358. Academic Press, London.
11. **Braunstein, M., A. M. Brown, S. Kurtz, and W. R. Jacobs, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. *J Bacteriol* **183**:6979-90.

12. **Bronstein, P. A., M. Marrichi, S. Cartinhour, D. J. Schneider, and M. P. DeLisa.** 2005. Identification of a twin-arginine translocation system in *Pseudomonas syringae* pv. tomato DC3000 and its contribution to pathogenicity and fitness. *J Bacteriol* **187**:8450-61.
13. **Broome-Smith, J. K., M. Tadayyon, and Y. Zhang.** 1990. Beta-lactamase as a probe of membrane protein assembly and protein export. *Mol Microbiol* **4**:1637-44.
14. **Caldelari, I., S. Mann, C. Crooks, and T. Palmer.** 2006. The Tat pathway of the plant pathogen *Pseudomonas syringae* is required for optimal virulence. *Mol Plant Microbe Interact* **19**:200-12.
15. **Camus, J. C., M. J. Pryor, C. Medigue, and S. T. Cole.** 2002. Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology* **148**:2967-73.
16. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, F. Tekaiia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, and B. G. Barrell.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**:537-44.
17. **Cole, S. T., K. Eiglmeier, J. Parkhill, K. D. James, N. R. Thomson, P. R. Wheeler, N. Honore, T. Garnier, C. Churcher, D. Harris, K. Mungall, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. M. Davies, K. Devlin, S. Duthoy, T. Feltwell, A. Fraser, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, C. Lacroix, J. Maclean, S. Moule, L. Murphy, K. Oliver, M. A. Quail, M. A. Rajandream, K. M. Rutherford, S. Rutter, K. Seeger, S. Simon, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, K. Taylor, S. Whitehead, J. R. Woodward, and B. G. Barrell.** 2001. Massive gene decay in the leprosy bacillus. *Nature* **409**:1007-11.
18. **Corbett, E. L., C. J. Watt, N. Walker, D. Maher, B. G. Williams, M. C. Raviglione, and C. Dye.** 2003. The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. *Arch Intern Med* **163**:1009-21.
19. **Cota-Gomez, A., A. I. Vasil, J. Kadurugamuwa, T. J. Beveridge, H. P. Schweizer, and M. L. Vasil.** 1997. PlcR1 and PlcR2 are putative calcium-binding proteins required for secretion of the hemolytic phospholipase C of *Pseudomonas aeruginosa*. *Infect Immun* **65**:2904-13.
20. **Cristobal, S., J. W. de Gier, H. Nielsen, and G. von Heijne.** 1999. Competition between Sec- and TAT-dependent protein translocation in *Escherichia coli*. *EMBO J* **18**:2982-90.

21. **Dalbey, R. E., and M. Chen.** 2004. Sec-translocase mediated membrane protein biogenesis. *Biochim Biophys Acta* **1694**:37-53.
22. **De Buck, E., L. Vranckx, E. Meyen, L. Maes, L. Vandersmissen, J. Anne, and E. Lammertyn.** 2007. The twin-arginine translocation pathway is necessary for correct membrane insertion of the Rieske Fe/S protein in *Legionella pneumophila*. *FEBS Lett* **581**:259-64.
23. **de Keyzer, J., C. van der Does, and A. J. Driessen.** 2003. The bacterial translocase: a dynamic protein channel complex. *Cell Mol Life Sci* **60**:2034-52.
24. **DeLisa, M. P., P. Samuelson, T. Palmer, and G. Georgiou.** 2002. Genetic analysis of the twin arginine translocator secretion pathway in bacteria. *J Biol Chem* **277**:29825-31.
25. **DeLisa, M. P., D. Tullman, and G. Georgiou.** 2003. Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. *Proc Natl Acad Sci U S A* **100**:6115-20.
26. **Dilks, K., R. W. Rose, E. Hartmann, and M. Pohlschroder.** 2003. Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. *J Bacteriol* **185**:1478-83.
27. **Ding, Z., and P. J. Christie.** 2003. *Agrobacterium tumefaciens* twin-arginine-dependent translocation is important for virulence, flagellation, and chemotaxis but not type IV secretion. *J Bacteriol* **185**:760-71.
28. **Dubnau, E., P. Fontan, R. Manganelli, S. Soares-Appel, and I. Smith.** 2002. *Mycobacterium tuberculosis* genes induced during infection of human macrophages. *Infect Immun* **70**:2787-95.
29. **Ferrer, P.** 2006. Revisiting the *Cellulosimicrobium cellulans* yeast-lytic beta-1,3-glucanases toolbox: a review. *Microb Cell Fact* **5**:10.
30. **Ferrer, P., T. Halkier, L. Hedegaard, D. Savva, I. Diers, and J. A. Asenjo.** 1996. Nucleotide sequence of a beta-1,3-glucanase isoenzyme IIA gene of *Oerskovia xanthineolytica* LL G109 (*Cellulomonas cellulans*) and initial characterization of the recombinant enzyme expressed in *Bacillus subtilis*. *J Bacteriol* **178**:4751-7.
31. **Finn, R. D., J. Mistry, B. Schuster-Bockler, S. Griffiths-Jones, V. Hollich, T. Lassmann, S. Moxon, M. Marshall, A. Khanna, R. Durbin, S. R. Eddy, E. L. Sonnhammer, and A. Bateman.** 2006. Pfam: clans, web tools and services. *Nucleic Acids Res* **34**:D247-51.

32. **Flores, A. R., L. M. Parsons, and M. S. Pavelka, Jr.** 2005. Characterization of novel *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* mutants hypersusceptible to beta-lactam antibiotics. *J Bacteriol* **187**:1892-900.
33. **Flores, A. R., L. M. Parsons, and M. S. Pavelka, Jr.** 2005. Genetic analysis of the beta-lactamases of *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* and susceptibility to beta-lactam antibiotics. *Microbiology* **151**:521-32.
34. **Gibbons, H. S., F. Wolschendorf, M. Abshire, M. Niederweis, and M. Braunstein.** 2007. Identification of two *Mycobacterium smegmatis* lipoproteins exported by a SecA2-dependent pathway. *J Bacteriol*.
35. **Glickman, M. S., J. S. Cox, and W. R. Jacobs, Jr.** 2000. A novel mycolic acid cyclopropane synthetase is required for cording, persistence, and virulence of *Mycobacterium tuberculosis*. *Mol Cell* **5**:717-27.
36. **Gomez, M., S. Johnson, and M. L. Gennaro.** 2000. Identification of secreted proteins of *Mycobacterium tuberculosis* by a bioinformatic approach. *Infect Immun* **68**:2323-7.
37. **Grass, G., and C. Rensing.** 2001. CueO is a multi-copper oxidase that confers copper tolerance in *Escherichia coli*. *Biochem Biophys Res Commun* **286**:902-8.
38. **Graubner, W., A. Schierhorn, and T. Bruser.** 2007. DnaK plays a pivotal role in Tat targeting of CueO and functions beside SlyD as a general Tat signal binding chaperone. *J Biol Chem* **282**:7116-24.
39. **Gu, S., J. Chen, K. M. Dobos, E. M. Bradbury, J. T. Belisle, and X. Chen.** 2003. Comprehensive proteomic profiling of the membrane constituents of a *Mycobacterium tuberculosis* strain. *Mol Cell Proteomics* **2**:1284-1296.
40. **Hackbarth, C. J., I. Unsal, and H. F. Chambers.** 1997. Cloning and sequence analysis of a class A beta-lactamase from *Mycobacterium tuberculosis* H37Ra. *Antimicrob Agents Chemother* **41**:1182-5.
41. **Haft, D. H., I. T. Paulsen, N. Ward, and J. D. Selengut.** 2006. Exopolysaccharide-associated protein sorting in environmental organisms: the PEP-CTERM/EpsH system. Application of a novel phylogenetic profiling heuristic. *BMC Biol* **4**:29.
42. **Halbig, D., T. Wiegert, N. Blaudeck, R. Freudl, and G. A. Sprenger.** 1999. The efficient export of NADP-containing glucose-fructose oxidoreductase to the periplasm of *Zymomonas mobilis* depends both on an intact twin-arginine motif in the signal peptide and on the generation of a structural export signal induced by cofactor binding. *Eur J Biochem* **263**:543-51.

43. **Hatzixanthis, K., D. J. Richardson, and F. Sargent.** 2005. Chaperones involved in assembly and export of N-oxide reductases. *Biochem Soc Trans* **33**:124-6.
44. **Hinsley, A. P., N. R. Stanley, T. Palmer, and B. C. Berks.** 2001. A naturally occurring bacterial Tat signal peptide lacking one of the 'invariant' arginine residues of the consensus targeting motif. *FEBS Lett* **497**:45-9.
45. **Hulo, N., A. Bairoch, V. Bulliard, L. Cerutti, E. De Castro, P. S. Langendijk-Genevaux, M. Pagni, and C. J. Sigrist.** 2006. The PROSITE database. *Nucleic Acids Res* **34**:D227-30.
46. **Ignatova, Z., C. Hornle, A. Nurk, and V. Kasche.** 2002. Unusual signal peptide directs penicillin amidase from *Escherichia coli* to the Tat translocation machinery. *Biochem Biophys Res Commun* **291**:146-9.
47. **Kim, C., W. W. Lorenz, J. T. Hoopes, and J. F. Dean.** 2001. Oxidation of phenolate siderophores by the multicopper oxidase encoded by the *Escherichia coli yacK* gene. *J Bacteriol* **183**:4866-75.
48. **Lavander, M., S. K. Ericsson, J. E. Broms, and A. Forsberg.** 2006. The twin arginine translocation system is essential for virulence of *Yersinia pseudotuberculosis*. *Infect Immun* **74**:1768-76.
49. **Lee, P. A., D. Tullman-Ercek, and G. Georgiou.** 2006. The bacterial twin-arginine translocation pathway. *Annu Rev Microbiol* **60**:373-395.
50. **Malen, H., F. S. Berven, K. E. Fladmark, and H. G. Wiker.** 2007. Comprehensive analysis of exported proteins from *Mycobacterium tuberculosis* H37Rv. *Proteomics* **7**:1702-18.
51. **Mawuenyega, K. G., C. V. Forst, K. M. Dobos, J. T. Belisle, J. Chen, E. M. Bradbury, A. R. Bradbury, and X. Chen.** 2005. *Mycobacterium tuberculosis* functional network analysis by global subcellular protein profiling. *Mol Biol Cell* **16**:396-404.
52. **McCann, J. R., J. A. McDonough, M. S. Pavelka, Jr., and M. Braunstein.** 2007. Beta-lactamase can function as a reporter of bacterial protein export during *Mycobacterium tuberculosis* infection of host cells. *Microbiology* **In press**.
53. **McDonough, J. A., K. E. Hacker, A. R. Flores, M. S. Pavelka, Jr., and M. Braunstein.** 2005. The twin-arginine translocation pathway of *Mycobacterium smegmatis* is functional and required for the export of mycobacterial beta-lactamases. *J Bacteriol* **187**:7667-79.
54. **Molik, S., I. Karnauchov, C. Weidlich, R. G. Herrmann, and R. B. Klossgen.** 2001. The Rieske Fe/S protein of the cytochrome b6/f complex in chloroplasts:

- missing link in the evolution of protein transport pathways in chloroplasts? J Biol Chem **276**:42761-6.
55. **Mori, H., and K. Ito.** 2001. The Sec protein-translocation pathway. Trends Microbiol **9**:494-500.
 56. **Palmer, T., F. Sargent, and B. C. Berks.** 2005. Export of complex cofactor-containing proteins by the bacterial Tat pathway. Trends Microbiol **13**:175-80.
 57. **Palva, I., M. Sarvas, P. Lehtovaara, M. Sibakov, and L. Kaariainen.** 1982. Secretion of *Escherichia coli* beta-lactamase from *Bacillus subtilis* by the aid of alpha-amylase signal sequence. Proc Natl Acad Sci U S A **79**:5582-6.
 58. **Pavelka, M. S., Jr., and W. R. Jacobs, Jr.** 1996. Biosynthesis of diaminopimelate, the precursor of lysine and a component of peptidoglycan, is an essential function of *Mycobacterium smegmatis*. J Bacteriol **178**:6496-507.
 59. **Pavelka, M. S., Jr., and W. R. Jacobs, 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**:4780-9.
 60. **Perez-Rodriguez, R., A. C. Fisher, J. D. Perlmutter, M. G. Hicks, A. Chanal, C. L. Santini, L. F. Wu, T. Palmer, and M. P. Delisa.** 2007. An essential role for the DnaK molecular chaperone in stabilizing over-expressed substrate proteins of the bacterial twin-arginine translocation pathway. J Mol Biol. **367**:715-30.
 61. **Posey, J. E., T. M. Shinnick, and F. D. Quinn.** 2006. Characterization of the twin-arginine translocase secretion system of *Mycobacterium smegmatis*. J Bacteriol **188**:1332-40.
 62. **Pradel, N., C. Ye, V. Livrelli, J. Xu, B. Joly, and L. F. Wu.** 2003. Contribution of the twin arginine translocation system to the virulence of enterohemorrhagic *Escherichia coli* O157:H7. Infect Immun **71**:4908-4916.
 63. **Raynaud, C., C. Guilhot, J. Rauzier, Y. Bordat, V. Pelicic, R. Manganelli, I. Smith, B. Gicquel, and M. Jackson.** 2002. Phospholipases C are involved in the virulence of *Mycobacterium tuberculosis*. Mol Microbiol **45**:203-17.
 64. **Rodrigue, A., A. Chanal, K. Beck, M. Muller, and L. F. Wu.** 1999. Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway. J Biol Chem **274**:13223-8.
 65. **Rose, R. W., T. Bruser, J. C. Kissinger, and M. Pohlschroder.** 2002. Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. Mol Microbiol **45**:943-50.

66. **Rossier, O., and N. P. Cianciotto.** 2005. The *Legionella pneumophila* *tatB* gene facilitates secretion of phospholipase C, growth under iron-limiting conditions, and intracellular infection. *Infect Immun* **73**:2020-32.
67. **Saint-Joanis, B., C. Demangel, M. Jackson, P. Brodin, L. Marsollier, H. Boshoff, and S. T. Cole.** 2006. Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) System of *Mycobacterium tuberculosis*, increases β -lactam susceptibility and virulence. *J Bacteriol* **188**:6669-79.
68. **Sambrook, J., and D. W. Russell.** 2001. *Molecular Cloning: A Laboratory Manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
69. **Sander, P., M. Rezwan, B. Walker, S. K. Rampini, R. M. Kroppenstedt, S. Ehlers, C. Keller, J. R. Keeble, M. Hagemeier, M. J. Colston, B. Springer, and E. C. Bottger.** 2004. Lipoprotein processing is required for virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* **52**:1543-52.
70. **Santini, C. L., A. Bernadac, M. Zhang, A. Chanal, B. Ize, C. Blanco, and L. F. Wu.** 2001. Translocation of jellyfish green fluorescent protein via the Tat system of *Escherichia coli* and change of its periplasmic localization in response to osmotic up-shock. *J Biol Chem* **276**:8159-64.
71. **Sargent, F., B. C. Berks, and T. Palmer.** 2006. Pathfinders and trailblazers: a prokaryotic targeting system for transport of folded proteins. *FEMS Microbiol Lett* **254**:198-207.
72. **Sasseti, C. M., D. H. Boyd, and E. J. Rubin.** 2001. Comprehensive identification of conditionally essential genes in mycobacteria. *Proc Natl Acad Sci U S A* **98**:12712-7.
73. **Schmidt, F., S. Donahoe, K. Hagens, J. Mattow, U. E. Schaible, S. H. Kaufmann, R. Aebersold, and P. R. Jungblut.** 2004. Complementary analysis of the *Mycobacterium tuberculosis* proteome by two-dimensional electrophoresis and isotope-coded affinity tag technology. *Mol Cell Proteomics* **3**:24-42.
74. **Selengut, J. D., D. H. Haft, T. Davidsen, A. Ganapathy, M. Gwinn-Giglio, W. C. Nelson, A. R. Richter, and O. White.** 2007. TIGRFAMs and Genome Properties: tools for the assignment of molecular function and biological process in prokaryotic genomes. *Nucleic Acids Res* **35**:D260-4.
75. **Smith, H., S. Bron, J. Van Ee, and G. Venema.** 1987. Construction and use of signal sequence selection vectors in *Escherichia coli* and *Bacillus subtilis*. *J Bacteriol* **169**:3321-8.

76. **Snapper, S. B., R. E. Melton, S. Mustafa, T. Kieser, and W. R. Jacobs, Jr.** 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol Microbiol* **4**:1911-9.
77. **Snyder, A., A. I. Vasil, S. L. Zajdowicz, Z. R. Wilson, and M. L. Vasil.** 2006. Role of the *Pseudomonas aeruginosa* PlcH Tat signal peptide in protein secretion, transcription, and cross-species Tat secretion system compatibility. *J Bacteriol* **188**:1762-74.
78. **Sonnhammer, E. L., G. von Heijne, and A. Krogh.** 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* **6**:175-82.
79. **Stanley, N. R., T. Palmer, and B. C. Berks.** 2000. The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. *J Biol Chem* **275**:11591-6.
80. **Stanley, N. R., F. Sargent, G. Buchanan, J. Shi, V. Stewart, T. Palmer, and B. C. Berks.** 2002. Behaviour of topological marker proteins targeted to the Tat protein transport pathway. *Mol Microbiol* **43**:1005-21.
81. **Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennett, G. P. Bansal, J. F. Young, M. H. Lee, and G. F. Hatfull.** 1991. New use of BCG for recombinant vaccines. *Nature* **351**:456-60.
82. **Sutcliffe, I. C., and D. J. Harrington.** 2004. Lipoproteins of *Mycobacterium tuberculosis*: an abundant and functionally diverse class of cell envelope components. *FEMS Microbiol Rev* **28**:645-59.
83. **Takayama, K., C. Wang, and G. S. Besra.** 2005. Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*. *Clin Microbiol Rev* **18**:81-101.
84. **Thomas, J. D., R. A. Daniel, J. Errington, and C. Robinson.** 2001. Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in *Escherichia coli*. *Mol Microbiol* **39**:47-53.
85. **Titgemeyer, F., J. Amon, S. Parche, M. Mahfoud, J. Bail, M. Schlicht, N. Rehm, D. Hillmann, J. Stephan, B. Walter, A. Burkovski, and M. Niederweis.** 2007. A genomic view of sugar transport in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. *J Bacteriol.* **189**:5903-15.
86. **Tokunaga, M., H. Tokunaga, and H. C. Wu.** 1982. Post-translational modification and processing of *Escherichia coli* prolipoprotein *in vitro*. *Proc Natl Acad Sci U S A* **79**:2255-9.

87. **Tullman-Ercek, D., M. P. Delisa, Y. Kawarasaki, P. Iranpour, B. Ribnicky, T. Palmer, and G. Georgiou.** 2007. Export pathway selectivity of *Escherichia coli* twin-arginine translocation signal peptides. *J Biol Chem.* **282**:8309-16.
88. **Voladri, R. K., D. L. Lakey, S. H. Hennigan, B. E. Menzies, K. M. Edwards, and D. S. Kernodle.** 1998. Recombinant expression and characterization of the major beta-lactamase of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **42**:1375-81.
89. **Voulhoux, R., G. Ball, B. Ize, M. L. Vasil, A. Lazdunski, L. F. Wu, and A. Filloux.** 2001. Involvement of the twin-arginine translocation system in protein secretion via the type II pathway. *Embo J* **20**:6735-41.
90. **Wang, F., C. Cassidy, and J. C. Sacchettini.** 2006. Crystal structure and activity studies of the *Mycobacterium tuberculosis* beta-lactamase reveal its critical role in resistance to beta-lactam antibiotics. *Antimicrob Agents Chemother* **50**:2762-71.
91. **Watanabe, T., N. Kasahara, K. Aida, and H. Tanaka.** 1992. Three N-terminal domains of beta-1,3-glucanase A1 are involved in binding to insoluble beta-1,3-glucan. *J Bacteriol* **174**:186-90.
92. **Widdick, D. A., K. Dilks, G. Chandra, A. Bottrill, M. Naldrett, M. Pohlschroder, and T. Palmer.** 2006. The twin-arginine translocation pathway is a major route of protein export in *Streptomyces coelicolor*. *Proc Natl Acad Sci U S A.* **103**:17927-32.

CHAPTER 4

β -lactamase can function as a reporter of bacterial protein export during *Mycobacterium tuberculosis* infection of host cells

Jessica R. McCann¹, Justin A. McDonough¹, Martin S. Pavelka, Jr.², and Miriam Braunstein¹

¹Department of Microbiology and Immunology,
University of North Carolina School of Medicine, Chapel Hill, North Carolina

²Department of Microbiology and Immunology,
University of Rochester School of Medicine and Dentistry, Rochester, New York

Mycobacterium tuberculosis is an intracellular pathogen that is able to avoid destruction by host immune defenses. Exported proteins of *M. tuberculosis*, which include proteins localized to the bacterial surface or secreted into the extracellular environment, are ideally situated to interact with host factors. As a result, these proteins are attractive candidates for virulence factors, drug targets, and vaccine components. Here we describe a new β -lactamase reporter system capable of identifying exported proteins of *M. tuberculosis* during growth in host cells. Because β -lactams target bacterial cell wall synthesis, β -lactamases must be exported beyond the cytoplasm to protect against these drugs. When used in protein fusions, β -lactamase can report on the subcellular location of another protein as measured by protection from β -lactam antibiotics. Here we demonstrate that a truncated TEM-1 β -lactamase lacking a signal sequence for export ('BlaTEM-1) can be used in this manner directly in a mutant strain of *M. tuberculosis* lacking the major β -lactamase, BlaC. The 'BlaTEM-1 reporter conferred β -lactam resistance when fused to both Sec and Tat

export signal sequences. We further demonstrate that β -lactamase fusion proteins report on protein export while *M. tuberculosis* is growing in THP-1 macrophage-like cells. This genetic system should facilitate the study of proteins exclusively exported in the host environment by intracellular *M. tuberculosis*.

Introduction

Tuberculosis is responsible for nearly two million deaths each year (47). *Mycobacterium tuberculosis*, the causative agent of this disease, is an intracellular pathogen and the ability of this bacterium to survive and grow in macrophages is essential to its virulence. Multiple processes are likely employed by *M. tuberculosis* to avoid destruction in macrophages. These include residing in a phagosome that fails to mature into an acidified phagolysosome and resisting reactive radicals (as reviewed in 32, 48). As in other bacterial pathogens, *M. tuberculosis* proteins exported beyond the cytoplasm to the bacterial cell envelope (comprised of the cytoplasmic membrane and cell wall) or secreted into the environment are ideally positioned to interact with host cell components and promote survival in macrophages. Consequently, exported and secreted proteins make good candidates for virulence factors, drug targets for disease intervention, and vaccine antigens.

Mycobacteria possess two conserved pathways for exporting proteins: the general secretion (Sec) pathway and the twin-arginine translocation (Tat) pathway (2, 21, 25, 28, 29, 35). These systems recognize precursor proteins synthesized with amino-terminal signal sequences and transport them across the cytoplasmic membrane (10, 26). The proteins exported by these pathways can remain associated with the cell envelope or be further

secreted by the bacterium. The signal sequences of Sec and Tat substrates share a similar domain structure; however, Tat substrates are distinguished by the presence of the twin-arginine motif, R-R-x- ϕ - ϕ (ϕ = uncharged residue). The two pathways also differ in their mode of transport. Sec substrates are translocated across the cytoplasmic membrane in an unfolded state, whereas Tat substrates are translocated in a folded conformation. *M. tuberculosis* also has at least two specialized protein export pathways: the SecA2-dependent system and the ESX-1 (ESAT-6) system (3, 15, 18, 30, 40). Interestingly, both pathways appear capable of secreting specific subsets of proteins that lack conventional Sec or Tat signal sequences.

In *M. tuberculosis*, proteomic and genetic methods have been used to experimentally identify proteins exported beyond the cytoplasm (reviewed in (21)). The genetic methods rely on reporter enzymes that are fused to *M. tuberculosis* protein sequences and report on the subcellular location of the fusion proteins (4, 8, 12, 23, 46). Surrogate hosts such as non-pathogenic *Mycobacterium smegmatis* or *Escherichia coli* have been used in most of these studies, often because endogenous enzyme activities in *M. tuberculosis* precluded their use directly in the pathogen. The use of surrogate hosts is a problem for identifying proteins that are only exported by pathogenic *M. tuberculosis*.

β -lactamase is an export reporter that was not initially employed directly in *M. tuberculosis* because of endogenous β -lactam resistance. β -lactamase catalyzes the hydrolysis of β -lactams, a class of antibiotic that targets cell wall biosynthetic enzymes located outside of the cytoplasmic membrane. Therefore, β -lactamase must be exported beyond the cytoplasm to protect the bacterium from the drug. For this reason, when fused to another protein, it can be used as an export reporter with β -lactam resistance as a powerful

indicator of export. We recently reported that a *ΔblaC* mutant of *M. tuberculosis*, lacking the chromosomally encoded β-lactamase BlaC, is β-lactam sensitive (14). Further, we showed that BlaC is a native Tat substrate and that a truncated 'BlaC lacking a signal sequence can function as a reporter of Tat-dependent export directly in a *ΔblaC* mutant of *M. tuberculosis* (25). This was shown by fusing a Tat signal sequence to 'BlaC and demonstrating that the resulting hybrid protein confers resistance to the β-lactam antibiotic carbenicillin in the *ΔblaC* background. Interestingly, the 'BlaC reporter works with Tat but not Sec exported proteins. Here we expanded the β-lactamase tools that can be used directly in *M. tuberculosis* by demonstrating that the TEM-1 β-lactamase (BlaTEM-1), originally identified in a clinical isolate of *E. coli* (9), functions as an export reporter in the *ΔblaC* mutant of *M. tuberculosis*. The 'BlaTEM-1 reporter has the significant advantage of being compatible with both Sec and Tat signal sequences.

The proteomic and genetic approaches used in previous work for identifying exported proteins of *M. tuberculosis* are limited by their reliance on *in vitro* grown bacteria. Consequently, a potentially interesting collection of proteins only exported or secreted while *M. tuberculosis* are inside host cells are missed. In this report, we demonstrate that β-lactamase reporters have the novel capability of identifying *M. tuberculosis* proteins that are exported during intracellular growth in β-lactam treated THP-1 macrophage-like cells. The system we describe will be of significant value for identifying the most interesting category of exported *M. tuberculosis* proteins – those exported during growth in the host environment.

Materials and Methods

Bacterial strains, media and growth conditions.

Escherichia coli DH5 α was grown in Luria-Bertani medium (Fisher) supplemented with the following concentrations of antibiotics as required: carbenicillin, 100 μ g/ml; kanamycin, 40 μ g/ml. *M. tuberculosis* strains H37Rv (WT), PM638 (Δ blaC, H37Rv) (14) and all derivative strains were cultured in Middlebrook 7H9 medium or on Middlebrook 7H10 agar medium (Difco; BD Biosciences) supplemented with 10% ADS (0.5% BSA, fraction V [Roche]; 0.2% dextrose; and 0.85% NaCl), 0.5% glycerol, and 0.05% Tween 80 (Fisher). Antibiotics for mycobacteria were used at the following concentrations: carbenicillin, 50 μ g/ml; kanamycin, 20 μ g/ml. 7H10 plates supplemented with carbenicillin lacked tween, as the combination of tween and carbenicillin appeared detrimental to growth of fusion-expressing strains.

Construction of 'blaTEM-1 fusion plasmids.

Plasmids used in this study are listed in Table 4.1. All subcloned PCR products were sequenced and determined to be error free. Sequence encoding the mature domain (lacking the N-terminal signal sequence) of *E. coli* BlaTEM-1 was amplified from pUC19 plasmid DNA (Invitrogen) using the following primers: TEMbla1 (5'-AGATCTCACCCAGAAACGCTGGTGAAAG) and TEMbla2 (5'-GTTACCAATGCTTAATCAGTGAGGCACC). The resulting PCR product was cloned into the pCC1 vector (Epicentre) to generate pJM114. The 'blaTEM-1 reporter was subcloned as a BglII-BamHI fragment into each of the multi-copy vectors described below. (i) Δ ss,

Table 4.1. Plasmids used in this study

Plasmid	Genotype	Description	Source
pCC1	<i>cat oriV ori2</i>	CopyControl (single copy) blunt cloning vector	Epicentre
pCR2.1	<i>bla aph ColE1</i>	TA cloning vector	Invitrogen
pMV261. <i>kan</i>	<i>aph P_{hsp60} oriM ColE1</i>	Multicopy mycobacterial shuttle plasmid	41
pMB219	<i>aph oriM ColE1</i>	Multicopy mycobacterial shuttle plasmid	This work
pMB222	<i>aph P_{hsp60}-ssplcB (M. tuberculosis) oriM ColE1</i>	<i>M. tuberculosis plcB</i> signal sequence in pMV261 under control of <i>hsp60</i> promoter	25
pMB227	<i>aph P_{hsp60}-ssmpt63 (M. tuberculosis) oriM ColE1</i>	<i>M. tuberculosis mpt63</i> signal sequence in pMV261 under control of <i>hsp60</i> promoter	25
pMB228	<i>aph P_{hsp60}-ssmpt63-'blaC (M. tuberculosis) oriM ColE1</i>	<i>M. tuberculosis ssmpt63-'blaC</i> in pMV261 under control of <i>hsp60</i> promoter	25
pJM109	<i>aph P_{hsp60}-ssfbpB-'blaC (M. tuberculosis) oriM ColE1</i>	<i>M. tuberculosis ssfbpB-'blaC</i> in pMV261 under control of <i>hsp60</i> promoter	25
pJM111	<i>aph P_{hsp60}-ssplcB-'blaC (M. tuberculosis) oriM ColE1</i>	<i>M. tuberculosis ssplcB-'blaC</i> in pMV261 under control of <i>hsp60</i> promoter	25
pJM113	<i>aph P_{hsp60}-'blaC (M. tuberculosis) oriM ColE1</i>	<i>M. tuberculosis 'blaC</i> (no signal sequence) in pMV261 under control of <i>hsp60</i> promoter	25
pJM114	<i>cat oriV ori2</i>	<i>E. coli 'blaTEM-1</i> cloned into pCC1	This work
pJES101	<i>aph P_{hsp60}-ssplcB-'blaTEM-1 (E. coli) oriM ColE1</i>	' <i>blaTEM-1</i> from pJM114 cloned into pMB222	This work
pJES102	<i>aph P_{hsp60}-'blaTEM-1 (E. coli) oriM ColE1</i>	' <i>blaTEM-1</i> from pJM114 cloned into pMV261	This work
pJES103	<i>aph P_{hsp60}-ssmpt63-'blaTEM-1 (E. coli) oriM ColE1</i>	' <i>blaTEM-1</i> from pJM114 cloned into pMB227	This work
pJES125	<i>bla aph ColE1</i>	<i>M. tuberculosis mpt83</i> signal sequence and upstream sequence cloned into pCR2.1	This work
pJES128	<i>aph oriM ColE1</i>	' <i>blaTEM-1</i> from pJM114 cloned into <i>Bam</i> HI-linearized pMB219	This work
pJES129	<i>aph P_{mpt83}-ssmpt83-'blaTEM-1 (E. coli) oriM ColE1</i>	<i>M. tuberculosis mpt83</i> signal sequence and upstream sequence from pJES125 cloned into pJES128	This work

'*blaTEM-1*. '*blaTEM-1* was digested from pJM114, end-filled with Klenow and cloned into

*Msc*I cut pMV261. The resulting plasmid, pJES102, contains the '*blaTEM-1* reporter

without a fused signal sequence cloned downstream of the *hsp60* promoter. (ii) **ssplcB-**

'*blaTEM-1*. The '*blaTEM-1* fragment was subcloned into *Bam*HI cut pMB222. The

resulting plasmid, pJES101, contains an in-frame fusion of DNA encoding the signal

sequence of *PlcB/Rv2350c* (*ssplcB*) to '*blaTEM-1* under the control of the *hsp60* promoter.

(iii) **ssmpt63-'blaTEM-1**. The '*blaTEM-1* fragment was subcloned into *Bam*HI cut pMB227.

The resulting plasmid, pJES103, contains an in-frame fusion of *ssmpt63* (*Rv1926c*) to

'*blaTEM-1* under the control of the *hsp60* promoter. (iv) **ssmpt83-'blaTEM-1**. DNA

encoding the signal sequence and the first 31 amino acids of the mature *M. tuberculosis*

Mpt83 (Rv2873) protein along with the native *mpt83* promoter (20) was amplified from *M. tuberculosis* genomic DNA using the following primers: mpt83HindIIIF (5'-CAAGCTTCGTCGGATCCGTGGTAGGGGATGTC) and mpt83HindIIIR (5'-CAAGCTTCGGGGTCAGCCATTGCCGCCGTGG) and cloned into the pCR2.1 vector (Invitrogen) to generate pJES125. A *HindIII* fragment from pJES125, carrying *ssmpt83* and upstream genomic sequence, was cloned into *HindIII* cut pJES128 (Table 4.1). The resulting plasmid, pJES129, contains an in-frame fusion of *ssmpt83* to '*blaTEM-1* under the control of the native *mpt83* promoter (P_{mpt83}).

Protein quantification by immunoblot.

Whole cell lysates of *M. tuberculosis* strains were prepared as described previously (2) with the following modifications. *M. tuberculosis* cultures were grown in 5 ml volumes to mid-exponential phase. The cultures were washed twice and resuspended in PBS 0.02% Tween 80. An equal volume of 10% formalin was added to the washed cultures, which were then incubated at room temperature for 1 hour with frequent mixing by inversion. The formalin fixation step was necessary to kill *M. tuberculosis* before further processing. Bacteria were then harvested by centrifugation at 3000 RPM, washed once in PBS 0.02% tween to remove residual formalin, and bead-beaten lysates were then obtained from each sample. Protein concentration for each lysate was measured using a bicinchoninic acid protein quantification kit (Pierce). Lysates were boiled for 10 minutes, subjected to SDS-PAGE and immunoblots were performed using standard conditions. Primary antibodies specific for BlaTEM-1 were used at a concentration of 1:5000 (QED Biosciences), and horseradish peroxidase-conjugated anti-mouse secondary antibodies were used at a

concentration of 1:20,000. Bands were visualized using Western Lightning Chemiluminescent Reagent Plus (PerkinElmer) and quantified using ImageJ Image Processing and Analysis software (<http://rsb.info.nih.gov/ij/>). Whole cells lysates with the highest level of expression were diluted to enable direct comparison of all hybridization signals on a single blot. The comparative quantification was determined by measuring pixel density of an equal area for each blotted lysate in duplicate. Signal intensity per μg of whole cell lysate loaded was determined and is reported as the amount relative to protein detected in the 'BlaTEM-1 expressing strain.

Macrophage infections.

THP-1 cells were maintained in RPMI (Gibco)/10% heat inactivated fetal calf serum (FCS) at 37 °C and 5% CO₂. To prepare THP-1 monolayers for infection, cells were centrifuged at 300 g, washed once in RPMI, then resuspended in RPMI /10% FCS at a concentration of 1×10^6 cells/ml. Cells were seeded into 8-well tissue culture slides at 2×10^5 cells/well and treated with phorbol myristate acetate (PMA) at a final concentration of 50 ng/ml for 48 hours.

M. tuberculosis was grown to mid-exponential phase OD₆₀₀ of 0.5-1.0. Immediately prior to infection, the bacterial culture was pelleted, washed once in PBS containing 0.05% Tween 80 (PBS-Tw), and resuspended in an equal volume of PBS-Tw. The culture was then briefly sonicated to break up clumps of bacteria, diluted in RPMI/10%FCS medium and added to the THP-1 monolayer at MOI = 0.1.

THP-1 monolayers were infected with *M. tuberculosis* strains for 4 hours at 37 °C and 5% CO₂. Overlaying medium was then removed, the monolayers were washed 3 times

with RPMI to remove non-cell associated bacteria, and triplicate wells were lysed and plated to determine uptake (day 0 time-point). The infected monolayers were then overlaid with RPMI/10% FCS, or RPMI/10% FCS supplemented with carbenicillin and maintained at 37°C and 5% CO₂. At 3 days post infection, the overlying medium was replenished with RPMI/10% FCS media or media supplemented with carbenicillin, as appropriate. On days 1, 3 and 5 post-infection, triplicate wells for each infection were washed to remove antibiotic and lysed with 0.05% SDS. The resulting lysates were diluted and plated on 7H10 agar to enumerate intracellular bacteria during the course of infection. On day 0 and day 5 of the infection, cell lysates were also plated on 7H10 agar supplemented with 50 µg/ml carbenicillin. This demonstrated that selection of spontaneous β-lactam resistant mutants did not occur during the course of infection. To determine the appropriate carbenicillin concentration necessary to kill intracellular bacteria, THP-1 infection experiments were performed with a range of antibiotic concentrations (Fig. 4.4b). Carbenicillin at 1 mg/ml was determined to be the lowest concentration of antibiotic that caused optimal killing of sensitive intracellular *M. tuberculosis* and was used in subsequent experiments.

Results

'BlaTEM-1 is exported by Sec and Tat signal sequences in *M. tuberculosis*.

β-lactamase is an ideal reporter for protein export because it must be localized beyond the bacterial cytoplasmic membrane to effectively protect the bacterium from β-lactam antibiotics. Therefore, it can be used in protein fusions to identify proteins that are extracytoplasmic. An attractive feature of a β-lactamase reporter is that a selection for β-

lactam resistant colonies can be performed, as opposed to a more labor-intensive screen. In the past, we showed that the endogenous β -lactamase of *M. tuberculosis* BlaC can function as a reporter of export exclusively by the Tat pathway when expressed in the β -lactam sensitive Δ *blaC* mutant of *M. tuberculosis* or Δ *blaS* mutant of *M. smegmatis* (25). Since the Sec pathway is likely responsible for the majority of protein export in *M. tuberculosis*, we were interested in utilizing a β -lactamase reporter that additionally works with Sec exported proteins. For this reason, we tested the *E. coli* TEM-1 β -lactamase (BlaTEM-1) which has been used in other bacteria to report on proteins exported by Sec, Tat, Type II and Type III secretion systems (5, 7, 37, 39).

A series of multi-copy kanamycin-marked '*blaTEM-1* plasmids were constructed and electroporated into the Δ *blaC* mutant of *M. tuberculosis* (Fig. 4.1). The resulting kanamycin resistant strains were tested for the ability to grow in the presence of 50 μ g/ml of the β -lactam carbenicillin. When the truncated '*blaTEM-1* reporter without a signal sequence was expressed in the Δ *blaC* mutant of *M. tuberculosis*, the strain remained carbenicillin-sensitive. In fact, no colonies of the strain expressing the truncated 'BlaTEM-1 grew on agar containing carbenicillin even after extended incubation (Figs. 4.1 and 4.2). However, expression of a hybrid protein comprised of a Sec signal sequence from Mpt63, a well-established secreted protein of *M. tuberculosis* (17, 24), fused to 'BlaTEM-1 (ssMpt63-'BlaTEM-1) protected the Δ *blaC* mutant from carbenicillin, as was evident by the ability of this strain to grow on carbenicillin agar plates (Figs. 4.1 and 4.2). We similarly tested a fusion protein in which the Sec signal sequence of a proven cell wall-associated lipoprotein, Mpt83 (16), was fused to 'BlaTEM-1. This construct also conferred β -lactam resistance to Δ *blaC* *M. tuberculosis* (Fig. 4.1). Of note, the ssMpt83-'BlaTEM-1 fusion protein also

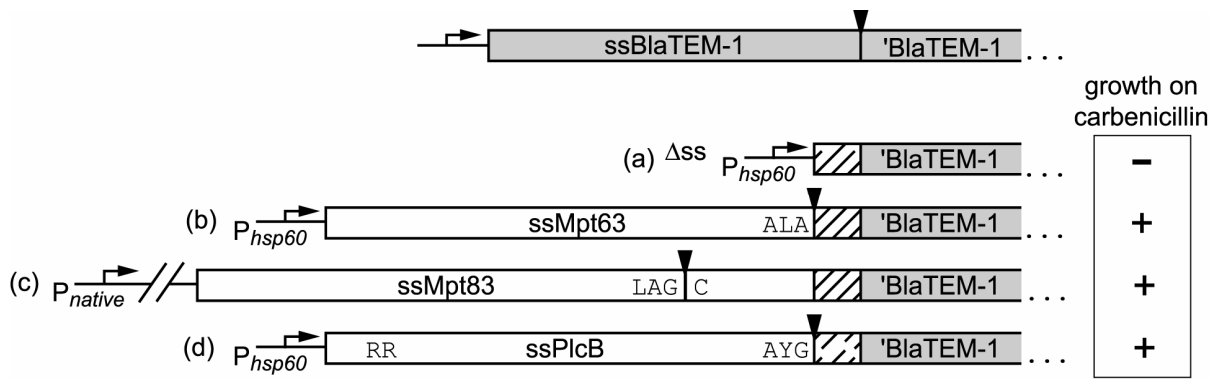


Fig. 4.1. Schematic representation of signal sequence-'BlaTEM-1 fusion constructs.

Mycobacterial shuttle plasmids were designed to encode fusion proteins of *M. tuberculosis* peptide sequence (open boxes) with a truncated 'BlaTEM-1 protein (gray boxes) lacking its native signal sequence. The hatched boxes indicate plasmid-derived peptide sequence that is present as a result of the cloning process. The constructs were driven off the constitutive *M. tuberculosis hsp60* promoter for (a) pJES102/'*blaTEM-1*, (b) pJES103/*ssmpt63*-'*blaTEM-1*, and (d) pJES101/*ssplcB*-'*blaTEM-1*. The native *M. tuberculosis* promoter located upstream of the *mpt83* operon was used to drive expression of (c) pJES129/ *ssmpt83*-'*blaTEM-1* (promoters indicated by arrows). Signal peptidase cleavage sites are indicated by arrowheads and by the AxA/G recognition motif for PlcB and Mpt63, and the LAGC lipobox recognition motif for Mpt83. Diagram not to scale; ss, signal sequence.

included the first 31 amino acids of the mature Mpt83 protein as well as the native *mpt83* promoter which is reported to be active at very low levels *in vitro* (16, 34).

Finally, we tested the signal sequence of PlcB, a proven cell wall-associated phospholipase C, for the ability to promote export of enzymatically active 'BlaTEM-1 (19, 31). PlcB has a predicted Tat signal sequence, and the ssPlcB-'BlaTEM-1 fusion also allowed *ΔblaC M. tuberculosis* to grow in the presence of carbenicillin (Fig. 4.1).

To determine whether the ssPlcB-'BlaTEM-1 fusion was exported by the Tat pathway, it was tested in *ΔblaS M. smegmatis* and in a *ΔtatA ΔblaS M. smegmatis* double mutant (25) in two independent experiments. When the ssPlcB-'BlaTEM-1 fusion protein was expressed in *ΔblaS M. smegmatis*, 92% of colonies were carbenicillin resistant. However, when the same construct was expressed in the *ΔtatA ΔblaS* mutant only an average 7% of colonies were carbenicillin resistant indicating that the Tat pathway functions in the export of this fusion protein. To show that a functional Tat pathway was not required for export of the Sec signal sequence-'BlaTEM-1 fusion, we similarly evaluated export of ssMpt63-'BlaTEM-1. When expressed in *ΔblaS* and the *ΔtatA ΔblaS* mutants, ssMpt63-'BlaTEM-1 conferred carbenicillin resistance to 90% and 95% of colonies, respectively. This indicated, as expected, no role for the Tat pathway in exporting a Sec signal sequence-'BlaTEM-1 fusion.

In each example where a *M. tuberculosis* signal sequence (Sec or Tat) was fused to 'BlaTEM-1, *ΔblaC M. tuberculosis* was protected from β-lactam attack. To demonstrate that the inability of the 'BlaTEM-1 reporter lacking a signal sequence to protect against carbenicillin was due to lack of export, as opposed to lack of expression, whole cell extracts of 'BlaTEM-1 expression strains were prepared and assayed for cell-associated β-lactamase.

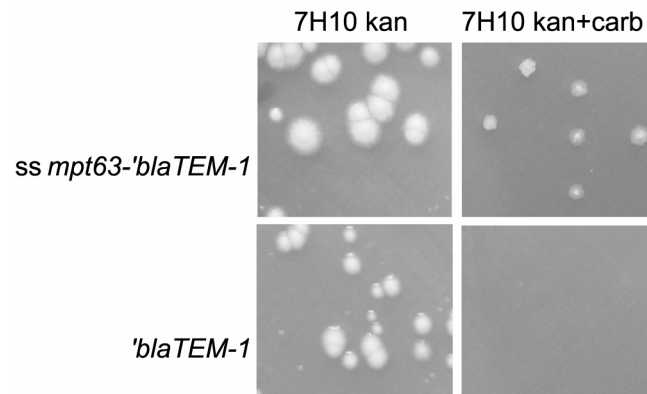


Fig. 4.2. 'BlaTEM-1 does not provide β -lactam-resistance to Δ blaC *M. tuberculosis*. Plasmids encoding the indicated 'blaTEM-1 fusions were electroporated into *M. tuberculosis* Δ blaC. The resulting strains were then plated on 7H10 plates supplemented with either kanamycin and 0.05% tween or kanamycin and carbenicillin without tween. Plates were inspected for growth following 21-25 days of incubation. Not shown are colonies expressing ssMpt83-'BlaTEM-1 and ssPlcB-'BlaTEM-1; growth on plates containing carbenicillin for these strains was similar to that conferred by ssMpt63-'BlaTEM-1.

To test for enzyme activity, we used the chromogenic β -lactam nitrocefin, which turns red following cleavage by β -lactamase (27). During a 15 minute incubation the nitrocefin was hydrolyzed by all strains expressing 'BlaTEM-1 constructs, while $\Delta blaC M. tuberculosis$ demonstrated no activity, similar to PBS alone (data not shown). Importantly, β -lactamase activity was detected with the truncated 'BlaTEM-1 reporter lacking a signal sequence. In fact, the lysate from the 'BlaTEM-1 strain converted nitrocefin to the red product almost instantaneously and faster than any other strain tested. We similarly detected β -lactamase activity in whole cell lysates of $\Delta blaC M. tuberculosis$ expressing the 'BlaC reporter lacking its native signal sequence.

We also compared the level of each 'BlaTEM-1 fusion protein present in whole cell lysates from the respective *M. tuberculosis* strains by immunoblots with antibodies specific for BlaTEM-1. This revealed a wide variation in the amount of 'BlaTEM-1 protein produced by the different strains (Fig. 4.3). The non-exported 'BlaTEM-1 expressed off the *hsp60* promoter (P_{hsp60}) was the most abundant protein detected. P_{hsp60} is considered a relatively strong promoter and is, therefore, present on many mycobacterial shuttle vectors (42). In comparison, the P_{hsp60} driven ssPlcB-'BlaTEM-1 and ssMpt63-'BlaTEM-1 were expressed at lower levels (59% and 0.9% of the level of the non-exported 'BlaTEM-1 construct, respectively). Since *mpt83* is expressed at relatively low levels *in vitro* we expected the ssMpt83-'BlaTEM-1 fusion to be weakly expressed (16, 34, 38). In fact, it was nearly undetectable by immunoblot, present at only 0.4% of the amount of non-exported 'BlaTEM-1 construct. The bands detected on the immunoblot are in general agreement with the predicted molecular weight of the expressed proteins. 'BlaTEM-1, lacking a signal sequence, has a predicted size of 28 kDa. Since whole cell lysates were analyzed in these

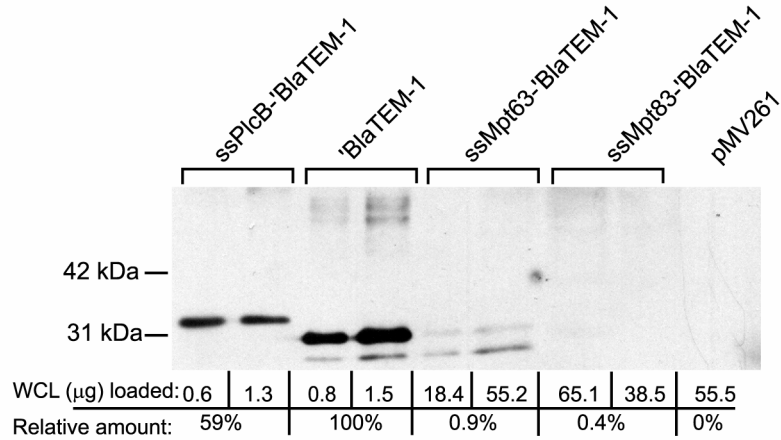


Fig. 4.3. 'BlaTEM-1 fusion proteins are detected at different amounts in *M. tuberculosis* whole cell lysates. Protein present in whole cell lysates (WCL) from each of the indicated Δ *blaC* strains were separated by SDS-PAGE and immunoblotted using primary antibody specific for BlaTEM-1. Comparative signal was quantified by measuring pixel density of an equal area for each blotted lysate in duplicate. Average signal intensity per μ g of WCL is reported as the amount relative to protein detected in the 'BlaTEM-1 expressing strain. Due to the different amounts of protein in each strain, it was necessary to load dilutions of the 'BlaTEM-1 and ssPlcB-'BlaTEM-1 expressing lysates so that signal from less abundant protein fusions could be simultaneously detected. There was no detectable signal with the WCL from the Δ *blaC* mutant carrying empty pMV261 plasmid.

experiments it is possible to see processed protein and/or uncleaved cytosolic precursor, which may explain the larger sized ssPlcB-‘BlaTEM-1 product. The signal sequences of PlcB and Mpt63 would add approximately 3 and 4 kDa, while the Mpt83 signal sequence and fused portion of the mature protein would add approximately 11 kDa, if left intact.

These observations suggested that even though ‘BlaTEM-1 does not promote growth in the presence of carbenicillin, a significant amount of β -lactamase was produced and accumulated within the bacterium. Together, our results indicated that in *ΔblaC M. tuberculosis* ‘BlaTEM-1 must be exported to confer protection against β -lactam antibiotics, that β -lactam resistance can be used to report on export, and that this reporter can be exported by Sec or Tat signal sequences and is compatible with different levels of expression.

The *ΔblaC* mutant of *M. tuberculosis* is sensitive to β -lactams during intracellular growth in human THP-1 cells.

β -lactam antibiotics can be used for clinical treatment of intracellular pathogens such as *Listeria monocytogenes* (33), and have been shown to reduce the population of phagocytosed *Staphylococcus aureus* (1). This indicates that β -lactams can enter macrophages and inhibit intracellular growth of some bacteria. The *ΔblaC* mutant of *M. tuberculosis* is sensitive to β -lactams *in vitro*, and we set out to test if this mutation also makes *M. tuberculosis* susceptible to β -lactams during growth in host cells.

Intracellular growth of the *ΔblaC* mutant was not previously evaluated; therefore, we first tested the ability of this mutant to grow within human monocytic THP-1 cells. THP-1 cells were infected at a MOI of 0.1 with either the *ΔblaC* mutant or the virulent parental

H37Rv strain. After a four hour period of infection, the THP-1 monolayer was washed to remove non-cell associated bacilli and fresh media was added back. Growth over a five day period was assessed by plating of infected host cell lysates for viable bacilli. The *ΔblaC* mutant showed no difference in intracellular growth when compared to H37Rv (Fig. 4.4a). Of note, we confirmed that *M. tuberculosis* does not grow in the THP-1 culture medium as previously reported (49).

To determine if the *ΔblaC* mutant was sensitive to β -lactams during intracellular growth, THP-1 cells were infected with *ΔblaC M. tuberculosis* and, following the washes to remove extracellular bacilli, media containing different concentrations of carbenicillin was added to the infected monolayers. After five days incubation, the infected monolayers were washed to remove carbenicillin and lysed to plate for viable bacilli. In the absence of carbenicillin, the *ΔblaC* mutant grew in THP-1 cells as previously seen. However, as the concentration of carbenicillin during the intracellular growth period increased, growth of the mutant diminished. At carbenicillin concentrations of ≥ 0.8 mg/ml substantial killing of the mutant was observed (Fig. 4.4b). These results indicated that the *ΔblaC* mutant is sensitive to β -lactam antibiotics during intracellular growth, and it suggested that the β -lactamase reporters could be used to study protein export during intracellular growth. Additional experiments showed that a concentration of 1 mg/ml carbenicillin was sufficient to achieve significant killing of the *ΔblaC* mutant of *M. tuberculosis* in THP-1 cells, and this concentration was used in all subsequent experiments.

Figure 4.

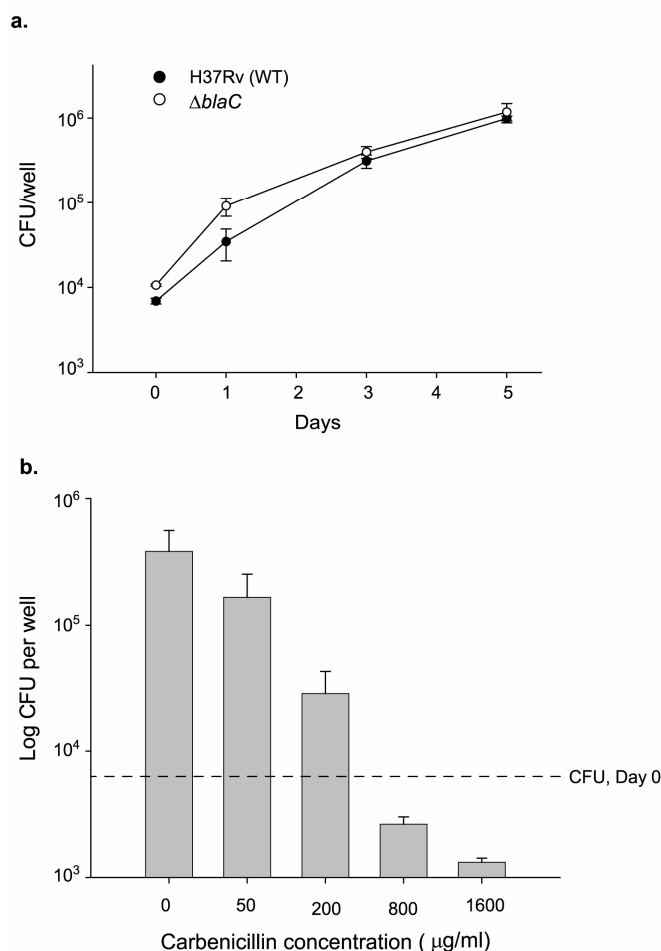


Fig. 4.4. The $\Delta blaC$ mutant of *M. tuberculosis* does not have a growth defect and is sensitive to β -lactam antibiotic in human THP-1 macrophage-like cells. (a) THP-1 cells were seeded into 8 well chamber slides, and triplicate wells were infected with either WT H37Rv or $\Delta blaC$ *M. tuberculosis* at a MOI of 0.1 bacilli per macrophage. At 4 hours (Day 0), 1, 3 and 5 days post infection, infected wells were washed, lysed, and plated for intracellular bacteria. Error bars represent standard error of the mean of c.f.u. in triplicate wells. (b) THP-1 cells were infected with $\Delta blaC$ *M. tuberculosis* as in (a). Following a 4-hour uptake period, wells were washed and indicated concentrations of carbenicillin were added to infected wells. Infected cells were lysed and plated at 4 hours (Day 0) and 5 days post infection to enumerate intracellular bacteria. Dashed line represents average intracellular CFU at 4 hours post infection. Error bars represent standard error of the mean of quadruplicate wells combined from two replicates.

Export of β -lactamase protects intracellular $\Delta blaC$ *M. tuberculosis* from β -lactam antibiotics.

A reporter system that works with intracellularly growing *M. tuberculosis* would be of great value for identifying exported proteins that are expressed and exported only during infection. Having shown that the $\Delta blaC$ mutant was sensitive to β -lactams during intracellular growth, we tested if β -lactamase could be used to report on protein export by *M. tuberculosis* growing in host cells. We tested fusion proteins expressing the 'BlaC and 'BlaTEM-1 reporters for the ability to protect the $\Delta blaC$ mutant in β -lactam treated THP-1 cells. In each experiment we compared an exported fusion protein to the truncated reporter alone. To test the 'BlaC reporter, which works with Tat exported proteins only, THP-1 cells were infected with the *M. tuberculosis* $\Delta blaC$ mutant expressing ssPlcB-'BlaC or 'BlaC only. Media with or without 1 mg/ml carbenicillin was added and the course of infection was monitored over a five day period. In the absence of carbenicillin, both strains grew in THP-1 cells during the course of the experiment. However, in the presence of carbenicillin, the strain expressing the truncated reporter alone did not grow and was reduced 10-fold over five days while the strain expressing the exported ssPlcB-'BlaC fusion protein was protected from carbenicillin and grew normally (Fig. 4.5a).

The 'BlaTEM-1 fusions were similarly tested. When THP-1 cells were infected with $\Delta blaC$ *M. tuberculosis* expressing either the exported ssMpt63-'BlaTEM-1 or the 'BlaTEM-1 reporter alone, only the strain expressing ssMpt63-'BlaTEM-1 fusion grew in THP-1 cells in the presence of carbenicillin. The non-exported 'BlaTEM-1 strain was sensitive to the β -lactam and was reduced in number 10-fold (Fig. 4.5b). Similarly, $\Delta blaC$ *M. tuberculosis* exporting ssMpt83-'BlaTEM-1 fusion was able to grow in carbenicillin treated THP-1 cells,

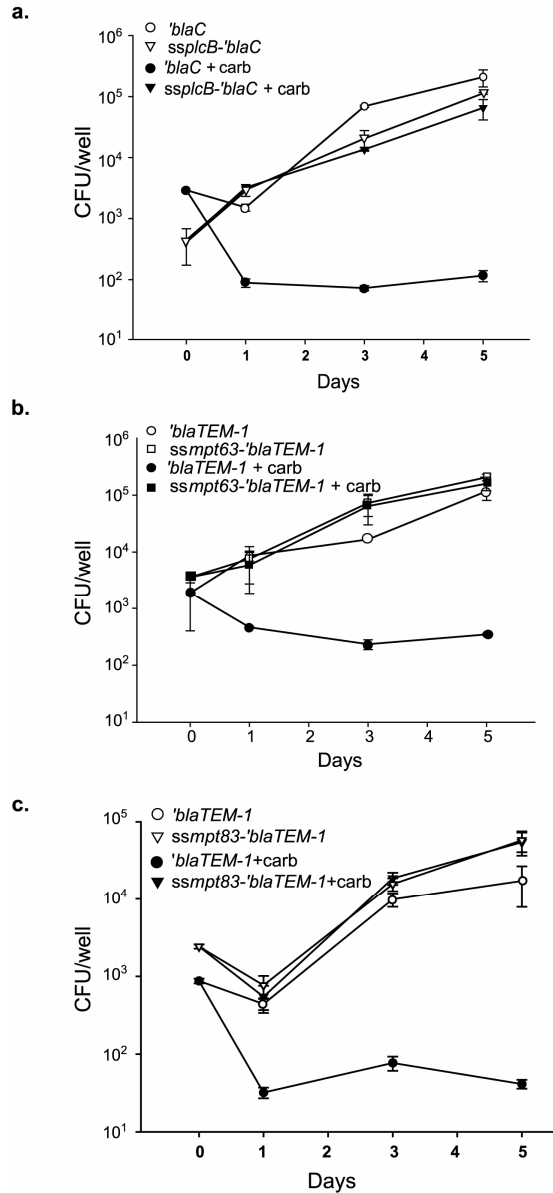


Fig. 4.5. *M. tuberculosis* signal sequences fused to 'BlaC and 'BlaTEM-1 protect intracellular bacilli from β -lactam antibiotics. THP-1 macrophage like cells were infected in triplicate wells with (a) *M. tuberculosis* Δ *blaC* expressing either ssPlcB-'BlaC or 'BlaC, (b) *M. tuberculosis* Δ *blaC* expressing either 'BlaTEM-1 or ssMpt63-'BlaTEM-1 or (c) *M. tuberculosis* Δ *blaC* expressing 'BlaTEM-1 or ssMpt83-'BlaTEM-1. Infected cells were then left untreated or treated with 1 mg/ml carbenicillin (carb). Wells were washed, lysed and plated 4 hours (d 0), 1, 3 and 5 days post infection. Each experiment was replicated 3 times in the case of (a) and (b), and 4 times in (c), with similar results.

while the non-exported 'BlaTEM-1 construct did not confer resistance to the *AblaC* mutant (Fig. 4.5c).

These experiments demonstrated that both the Tat specific 'BlaC reporter and the more permissive 'BlaTEM-1 reporter can report on protein export while *M. tuberculosis* is growing in β -lactam treated host cells. The use of β -lactamase reporters with intracellular *M. tuberculosis* represents a powerful tool for the study and identification of proteins exported during growth in host cells.

Discussion

The exported proteins of *M. tuberculosis* have been the subject of research attention for some time. This stems from the well-established fact that the majority of bacterial virulence factors and antigens are proteins exported out of the cytoplasm to the bacterial cell envelope or secreted out from the bacterium (13). In fact, there is a growing list of *M. tuberculosis* exported and secreted proteins shown to contribute to virulence or to development of a host immune response (21). Genetic reporters have proven to be powerful tools for identifying these extracytoplasmic proteins. The construction of a β -lactam sensitive *AblaC* mutant of *M. tuberculosis* opened the door for using β -lactamases as reporters of protein export directly in *M. tuberculosis*. The 'BlaC reporter can be used as a Tat specific reporter while the 'BlaTEM-1 reporter, shown here, can work with Sec or Tat signal sequences. An advantage of β -lactamase reporters is that they can be used to select for exported fusion proteins, as opposed to more labor intensive screening. In addition, we showed here for the first time that resistance to β -lactam antibiotics can be used to report on

protein export during intracellular growth of bacteria. Even in more genetically tractable bacterial pathogens, the identification of proteins exported or secreted from within host cells is a challenge.

Because β -lactams target cell wall modifying enzymes, β -lactamases must be exported in order to protect against these drugs. This export requirement was previously exploited with fusion proteins expressed in *E. coli* and other bacteria grown *in vitro* (5, 22). Here we showed that BlaTEM-1 can also report on protein export directly in *AblaC M. tuberculosis*. The three *M. tuberculosis* signal sequences tested in our study are from well-established secreted or cell-wall associated proteins. Mpt63 (Rv1926c, 16kDa protein) has a predicted Sec signal sequence and is one of the four most abundant *M. tuberculosis* proteins secreted into culture media during *in vitro* growth (17). Mpt83 (Rv2873) is a glycosylated lipoprotein (16, 43) that is exported to the cell wall of *M. tuberculosis*. Mpt83 has a predicted Sec signal sequence with a lipoprotein signal peptidase (LspA) cleavage site and the requisite conserved cysteine for lipid modification. PlcB (Rv2350c, phospholipase C) is a cell wall associated protein of *M. tuberculosis* shown to function in virulence (19, 31). Unlike Mpt63 and Mpt83, PlcB has a predicted Tat signal sequence including a twin-arginine motif (11). Signal sequences from all three of these proteins were able to promote export of a fused 'BlaTEM-1 reporter on the basis of production of β -lactam resistance. Notably, the ssMpt83-'BlaTEM-1 fusion protein was expressed from the native *mpt83* promoter and the fusion protein included the predicted signal sequence plus 31 amino acids of the mature Mpt83 protein. This demonstrated the ability of the reporter to work with different strength promoters and extended protein sequences. It is important to note that even though variable levels of fusion protein were detected in *M. tuberculosis* whole cell lysates as determined by

immunoblot, each exported fusion provided sufficient protection against 50 µg/ml carbenicillin while the most abundant 'BlaTEM-1 without an export signal did not confer β-lactam resistance.

Previously, we showed that the PlcB signal sequence is able to drive export of functional 'BlaC in a Tat- and RR-dependent manner (25). In *E. coli* the 'BlaTEM-1 reporter works with both Sec and Tat signal sequences (5, 39). The Sec and Tat pathways appear essential in *M. tuberculosis* (2, 35, 36). Therefore, to investigate the mode of export of the ssPlcB-'BlaTEM-1 fusion protein it was tested in *M. smegmatis* Δ *blaS* and in a *M. smegmatis* Δ *tatA* Δ *blaS* double mutant. A 93% reduction in β-lactam resistant colonies was observed in the *M. smegmatis* Δ *tatA* Δ *blaS* double mutant. Thus, the Tat pathway is involved in the export of ssPlcB-'BlaTEM-1, although other export pathways participate as well. The signal sequence of PlcB may be promiscuous in targeting the Tat or Sec pathway for export depending on the folded or unfolded nature of a fused reporter element. Similar results were recently shown for some predicted Tat signal sequences in *E. coli* (45).

In addition to working with the Sec and Tat pathways, the 'BlaTEM-1 reporter has been used with type II and type III secretion systems of Gram-negative bacteria (7, 37). Since substrates of the type III secretion system lack conventional N-terminal signal sequences, it remains possible that the 'BlaTEM-1 reporter will also work with non-conventional exported proteins of *M. tuberculosis*.

An interesting category of exported proteins that has been largely overlooked are those proteins only expressed and/or exported during the course of infection. We hypothesize that these are proteins exclusively exported in the host environment including virulence factors and protective antigens. Further, only a small number of the exported *M.*

tuberculosis proteins identified *in vitro* have ever been directly investigated during intracellular growth in host cells (21). For most of these studies, immunomicroscopy was used to localize the proteins in *M. tuberculosis* infected macrophages, which required development of suitable antibodies. We reasoned that if β -lactam antibiotics can reach intracellular Δ *blaC M. tuberculosis*, β -lactamase reporters should additionally work during intracellular growth. β -lactam antibiotics do not normally accumulate in eukaryotic cells; however, antibiotics of this class freely diffuse in and out of host cells (44), and β -lactam antibiotics are used to treat some intracellular bacterial infections (33). More specifically, β -lactams reach intracellular *Staphylococcus aureus* and *Listeria monocytogenes* and prevent growth of these organisms in THP-1 cells (1, 6). Here we showed that Δ *blaC M. tuberculosis* in THP-1 cells was also susceptible to carbenicillin. Thus, BlaC is responsible for *M. tuberculosis* resistance to β -lactam antibiotics during intracellular growth, indicating that the chromosomal *blaC* is a key factor preventing the use of β -lactams to treat *M. tuberculosis* infection.

When the set of exported β -lactamase fusion proteins was tested for the ability to protect Δ *blaC M. tuberculosis* from β -lactam treatment during intracellular growth, all exported fusions conferred resistance. In contrast, the truncated non-exported β -lactamase reporters were not protective. These experiments demonstrated the effectiveness of both 'BlaC and 'BlaTEM-1 reporters to identify *M. tuberculosis* sequences that drive export of each reporter during growth within host cells. Because the ssMpt83-'BlaTEM-1 fusion was expressed from the native promoter, our results indicate that Mpt83, a protein of unknown function, is expressed and exported during intracellular infection. This result is consistent with the reported induction of *mpt83* in macrophages (38) .

Several approaches have described proteins exported by *M. tuberculosis in vitro*, but a different suite of proteins may be exported during infection of the host. The intracellular β -lactamase reporter system we describe represents a new genetic tool for studying protein export in *M. tuberculosis*. It can be used to directly test the intracellular export of a protein of interest. We also hope to use it in combination with multiple rounds of infection and selection of β -lactam resistant clones from a *M. tuberculosis* fusion library. This should serve to identify the most interesting category of proteins; namely, those that are exported during intracellular growth and missed by alternative methods.

Acknowledgements

This research was funded in part by grants from the National Institute of Allergy and Infectious Disease of the NIH (AI54540 and AI070928 to M.B., and AI47311 to M.S.P.) and a developmental grant from the University of North Carolina at Chapel Hill Center for AIDS Research (NIH #9P30 AI 50410-04). J.R.M was supported by an NIH Cell and Molecular Biology training grant (NIH 5-T32- GM008581). J.A.M. was also supported by the Training in Sexually Transmitted Diseases and AIDS NIH NIAID training grant 5T32 AI07001-28 and a Society of Fellows dissertation completion fellowship from the Graduate School at the University of North Carolina at Chapel Hill. We thank Nathan Rigel for assistance with quantitative immunoblots and the members of the Braunstein laboratory for review of the manuscript.

Attributions

The work described here was a collaborative effort between Jessica McCann and myself. I developed the system and was involved in all aspects, and Jessica was further responsible for the subsequent analysis of the TEM-1 β -lactamase data. This work has been previously published (*Microbiology*. 2007 Oct;153(10):3350-9). Permission to reprint this work has been granted by the publisher.

References

1. **Barcia-Macay, M., C. Seral, M. P. Mingeot-Leclercq, P. M. Tulkens, and F. Van Bambeke.** 2006. Pharmacodynamic evaluation of the intracellular activities of antibiotics against *Staphylococcus aureus* in a model of THP-1 macrophages. *Antimicrob Agents Chemother* **50**:841-51.
2. **Braunstein, M., A. M. Brown, S. Kurtz, and W. R. Jacobs, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. *J Bacteriol* **183**:6979-90.
3. **Braunstein, M., B. J. Espinosa, J. Chan, J. T. Belisle, and W. R. Jacobs, Jr.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* **48**:453-64.
4. **Braunstein, M., T. I. Griffin, J. I. Kriakov, S. T. Friedman, N. D. Grindley, and W. R. Jacobs, Jr.** 2000. Identification of genes encoding exported *Mycobacterium tuberculosis* proteins using a Tn552 phoA in vitro transposition system. *J Bacteriol* **182**:2732-40.
5. **Broome-Smith, J. K., M. Tadayyon, and Y. Zhang.** 1990. Beta-lactamase as a probe of membrane protein assembly and protein export. *Mol Microbiol* **4**:1637-44.
6. **Carryn, S., F. Van Bambeke, M. P. Mingeot-Leclercq, and P. M. Tulkens.** 2003. Activity of beta-lactams (ampicillin, meropenem), gentamicin, azithromycin and moxifloxacin against intracellular *Listeria monocytogenes* in a 24 h THP-1 human macrophage model. *J Antimicrob Chemother* **51**:1051-2.
7. **Charpentier, X., and E. Oswald.** 2004. Identification of the secretion and translocation domain of the enteropathogenic and enterohemorrhagic *Escherichia coli* effector Cif, using TEM-1 beta-lactamase as a new fluorescence-based reporter. *J Bacteriol* **186**:5486-95.
8. **Chubb, A. J., Z. L. Woodman, F. M. da Silva Tatley, H. J. Hoffmann, R. R. Scholle, and M. R. Ehlers.** 1998. Identification of *Mycobacterium tuberculosis* signal sequences that direct the export of a leaderless beta-lactamase gene product in *Escherichia coli*. *Microbiology* **144**:1619-29.
9. **Datta, N., and P. Kontomichalou.** 1965. Penicillinase synthesis controlled by infectious R factors in Enterobacteriaceae. *Nature* **208**:239-41.
10. **DeLisa, M. P., D. Tullman, and G. Georgiou.** 2003. Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. *Proc Natl Acad Sci U S A* **100**:6115-20.

11. **Dilks, K., R. W. Rose, E. Hartmann, and M. Pohlschroder.** 2003. Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. *J Bacteriol* **185**:1478-83.
12. **Downing, K. J., R. A. McAdam, and V. Mizrahi.** 1999. *Staphylococcus aureus* nuclease is a useful secretion reporter for mycobacteria. *Gene* **239**:293-299.
13. **Finlay, B. B., and S. Falkow.** 1997. Common themes in microbial pathogenicity revisited. *Microbiol Mol Biol Rev* **61**:136-69.
14. **Flores, A. R., L. M. Parsons, and M. S. Pavelka, Jr.** 2005. Genetic analysis of the beta-lactamases of *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* and susceptibility to beta-lactam antibiotics. *Microbiology* **151**:521-32.
15. **Guinn, K. M., M. J. Hickey, S. K. Mathur, K. L. Zakel, J. E. Grotzke, D. M. Lewinsohn, S. Smith, and D. R. Sherman.** 2004. Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* **51**:359-70.
16. **Hewinson, R. G., S. L. Michell, W. P. Russell, R. A. McAdam, and W. J. Jacobs.** 1996. Molecular characterization of MPT83: a seroreactive antigen of *Mycobacterium tuberculosis* with homology to MPT70. *Scand J Immunol* **43**:490-9.
17. **Horwitz, M. A., B. W. Lee, B. J. Dillon, and G. Harth.** 1995. Protective immunity against tuberculosis induced by vaccination with major extracellular proteins of *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* **92**:1530-4.
18. **Hsu, T., S. M. Hingley-Wilson, B. Chen, M. Chen, A. Z. Dai, P. M. Morin, C. B. Marks, J. Padiyar, C. Goulding, M. Gingery, D. Eisenberg, R. G. Russell, S. C. Derrick, F. M. Collins, S. L. Morris, C. H. King, and W. R. Jacobs, Jr.** 2003. The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue. *Proc Natl Acad Sci U S A* **100**:12420-5.
19. **Johansen, K. A., R. E. Gill, and M. L. Vasil.** 1996. Biochemical and molecular analysis of phospholipase C and phospholipase D activity in mycobacteria. *Infect Immun* **64**:3259-66.
20. **Juarez, M. D., A. Torres, and C. Espitia.** 2001. Characterization of the *Mycobacterium tuberculosis* region containing the *mpt83* and *mpt70* genes. *FEMS Microbiol Lett* **203**:95-102.
21. **Kurtz, S., and M. Braunstein.** 2005. Protein secretion and export in *Mycobacterium tuberculosis*, p. 71-138. In T. Parish (ed.), *Mycobacterium molecular biology*. Horizon bioscience, Norfolk, UK.

22. **Lee, H. J., and K. T. Hughes.** 2006. Posttranscriptional control of the *Salmonella enterica* flagellar hook protein FlgE. *J Bacteriol* **188**:3308-16.
23. **Lim, E. M., J. Rauzier, J. Timm, G. Torrea, A. Murray, B. Gicquel, and D. Portnoi.** 1995. Identification of *Mycobacterium tuberculosis* DNA sequences encoding exported proteins by using *phoA* gene fusions. *J Bacteriol* **177**:59-65.
24. **Manca, C., K. Lyashchenko, H. G. Wiker, D. Usai, R. Colangeli, and M. L. Gennaro.** 1997. Molecular cloning, purification, and serological characterization of MPT63, a novel antigen secreted by *Mycobacterium tuberculosis*. *Infect Immun* **65**:16-23.
25. **McDonough, J. A., K. E. Hacker, A. R. Flores, M. S. Pavelka, Jr., and M. Braunstein.** 2005. The twin-arginine translocation pathway of *Mycobacterium smegmatis* is functional and required for the export of mycobacterial beta-lactamases. *J Bacteriol* **187**:7667-79.
26. **Mori, H., and K. Ito.** 2001. The Sec protein-translocation pathway. *Trends Microbiol* **9**:494-500.
27. **O'Callaghan, C. H., A. Morris, S. M. Kirby, and A. H. Shingler.** 1972. Novel method for detection of beta-lactamases by using a chromogenic cephalosporin substrate. *Antimicrob Agents Chemother* **1**:283-8.
28. **Owens, M. U., W. E. Swords, M. G. Schmidt, C. H. King, and F. D. Quinn.** 2002. Cloning, expression, and functional characterization of the *Mycobacterium tuberculosis secA* gene. *FEMS Microbiol Lett* **211**:133-41.
29. **Posey, J. E., T. M. Shinnick, and F. D. Quinn.** 2006. Characterization of the twin-arginine translocase secretion system of *Mycobacterium smegmatis*. *J Bacteriol* **188**:1332-40.
30. **Pym, A. S., P. Brodin, L. Majlessi, R. Brosch, C. Demangel, A. Williams, K. E. Griffiths, G. Marchal, C. Leclerc, and S. T. Cole.** 2003. Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis. *Nat Med* **9**:533-9.
31. **Raynaud, C., C. Guilhot, J. Rauzier, Y. Bordat, V. Pelicic, R. Manganelli, I. Smith, B. Gicquel, and M. Jackson.** 2002. Phospholipases C are involved in the virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* **45**:203-17.
32. **Russell, D. G.** 2007. Who puts the tubercle in tuberculosis? *Nat Rev Microbiol* **5**:39-47.

33. **Safdar, A., and D. Armstrong.** 2003. Antimicrobial activities against 84 *Listeria monocytogenes* isolates from patients with systemic listeriosis at a comprehensive cancer center (1955-1997). *J Clin Microbiol* **41**:483-5.
34. **Said-Salim, B., S. Mostowy, A. S. Kristof, and M. A. Behr.** 2006. Mutations in *Mycobacterium tuberculosis* Rv0444c, the gene encoding anti-SigK, explain high level expression of MPB70 and MPB83 in *Mycobacterium bovis*. *Mol Microbiol* **62**:1251-63.
35. **Saint-Joanis, B., C. Demangel, M. Jackson, P. Brodin, L. Marsollier, H. Boshoff, and S. T. Cole.** 2006. Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of *Mycobacterium tuberculosis*, increases {beta}-lactam susceptibility and virulence. *J Bacteriol* **188**:6669-79.
36. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48**:77-84.
37. **Sauvonnnet, N., and A. P. Pugsley.** 1996. Identification of two regions of *Klebsiella oxytoca pullulanase* that together are capable of promoting beta-lactamase secretion by the general secretory pathway. *Mol Microbiol* **22**:1-7.
38. **Schnappinger, D., S. Ehrt, M. I. Voskuil, Y. Liu, J. A. Mangan, I. M. Monahan, G. Dolganov, B. Efron, P. D. Butcher, C. Nathan, and G. K. Schoolnik.** 2003. Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: insights into the phagosomal environment. *J Exp Med* **198**:693-704.
39. **Stanley, N. R., F. Sargent, G. Buchanan, J. Shi, V. Stewart, T. Palmer, and B. C. Berks.** 2002. Behaviour of topological marker proteins targeted to the Tat protein transport pathway. *Mol Microbiol* **43**:1005-21.
40. **Stanley, S. A., S. Raghavan, W. W. Hwang, and J. S. Cox.** 2003. Acute infection and macrophage subversion by *Mycobacterium tuberculosis* require a specialized secretion system. *Proc Natl Acad Sci U S A* **100**:13001-6.
41. **Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennett, G. P. Bansal, J. F. Young, M. H. Lee, and G. F. Hatfull.** 1991. New use of BCG for recombinant vaccines. *Nature* **351**:456-60.
42. **Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennett, G. P. Bansal, J. F. Young, M. H. Lee, G. F. Hatfull, and et al.** 1991. New use of BCG for recombinant vaccines. *Nature* **351**:456-60.
43. **Sutcliffe, I. C., and D. J. Harrington.** 2004. Lipoproteins of *Mycobacterium tuberculosis*: an abundant and functionally diverse class of cell envelope components. *FEMS Microbiol Rev* **28**:645-59.

44. **Tulkens, P. M.** 1991. Intracellular distribution and activity of antibiotics. *Eur J Clin Microbiol Infect Dis* **10**:100-6.
45. **Tullman-Ercek, D., M. P. Delisa, Y. Kawarasaki, P. Iranpour, B. Ribnicky, T. Palmer, and G. Georgiou.** 2007. Export pathway selectivity of *Escherichia coli* twin-arginine translocation signal peptides. *J Biol Chem.* **282**:8309-16.
46. **Wiker, H. G., M. A. Wilson, and G. K. Schoolnik.** 2000. Extracytoplasmic proteins of *Mycobacterium tuberculosis* - mature secreted proteins often start with aspartic acid and proline. *Microbiology* **146 (Pt 7)**:1525-33.
47. **World Health Organization** 2007, posting date. WHO Information tuberculosis fact sheet. [Online.]
48. **Zahrt, T. C., and V. Deretic.** 2002. Reactive nitrogen and oxygen intermediates and bacterial defenses: unusual adaptations in *Mycobacterium tuberculosis*. *Antioxid Redox Signal* **4**:141-59.
49. **Zhang, M., J. Gong, Y. Lin, and P. F. Barnes.** 1998. Growth of virulent and avirulent *Mycobacterium tuberculosis* strains in human macrophages. *Infect Immun* **66**:794-9.

CHAPTER 5

DISCUSSION

Tuberculosis (TB) is responsible for 1.7 million deaths each year and remains a global health problem despite the existence of anti-mycobacterial therapies (53). Several major factors contribute to the continued existence of this disease such as the lack of an effective vaccine, the increased development of active TB in individuals with AIDS, and the emergence of drug resistant strains of *Mycobacterium tuberculosis* (27). These problems can be addressed by achieving a greater understanding of the mechanisms that *M. tuberculosis* uses to cause disease. Specifically, further insight into the means by which *M. tuberculosis* interacts with the host may lead to the development of better vaccines, diagnostics, and anti-mycobacterial therapies with the ultimate goal of combating TB.

We are interested in how protein export pathways of *M. tuberculosis* contribute to virulence. In bacterial pathogens, protein export pathways play important roles in transporting virulence factors and immunogenic antigens to their proper external location to enable interaction with the host (11). The protein export pathways of *M. tuberculosis* have only begun to be studied (21, 25). Knowledge of these pathways and their exported substrates is important in determining their contribution to virulence and could lead to new disease intervention and prevention strategies.

The Tat pathway was first described in the thylakoid membranes of plant chloroplasts and subsequently in bacteria as a novel means of transporting folded proteins across

membranes (2, 38, 40, 44, 51). Since then, a large amount of work, conducted primarily with *E. coli*, has contributed to a basic understanding of the unique properties of Tat export (23, 39). More recently, the Tat pathway has begun to be investigated in other bacterial species (23). From these analyses, it is clear that specific properties of Tat export are conserved between unrelated bacteria. However, bacteria-specific differences also exist. For example, some Gram-positive bacteria with a functioning Tat pathway lack a *tatB* ortholog (9), which is an essential component of the Tat translocase in Gram-negative bacteria (41). Finally, the studies also reveal that the Tat pathway is involved in virulence of some bacterial pathogens and may represent a unique means of localizing important virulence factors (3, 5, 10, 22, 32, 35, 50). Investigation of Tat-export in other bacteria and identification of the cognate exported substrates may help to uncover novel properties of Tat-export and help pinpoint elements that are conserved or unique to specific groups or families of bacteria.

The work presented in this dissertation dealt with the initial characterization of the Tat pathway in mycobacteria. We hypothesized that like other bacterial pathogens, *M. tuberculosis* utilized the Tat pathway for the export of a subset of folded proteins including virulence factors. Additionally, we sought to use the information obtained from our analyses to better understand the basic properties of Tat export in bacteria. Here, we discuss the contribution of this work to (i) determining a role for the Tat pathway in *M. tuberculosis* physiology and pathogenesis, (ii) better understanding the properties of Tat export in bacteria, and (iii) exploiting the ability of the mycobacterial Tat pathway to export folded proteins of interest.

Determining a role for the Tat pathway in *M. tuberculosis* pathogenesis and physiology.

The Tat pathway is important for the virulence of several bacterial pathogens and in some cases has specifically been shown to export virulence factors (3, 5, 10, 22, 32, 35, 50). It has become clear that the phospholipase C enzymes, which are virulence factors in many bacteria, are commonly Tat-exported proteins (9). This is true in *Pseudomonas aeruginosa* where Δtat mutants fail to export two phospholipase C enzymes critical for virulence (50). The phospholipase C enzymes of *M. tuberculosis* are cell-wall associated virulence factors with predicted twin-arginine signal sequences (33). We hypothesized that the Tat pathway was involved in the export of the phospholipase C virulence factors and therefore is likely to be important to virulence. In addition, there may be other proteins exported by the Tat pathway that contribute to the virulence of *M. tuberculosis*.

In Chapter 2, we presented data that demonstrated the functionality of the Tat pathway, by making *tat* mutants, in the genetic model system *Mycobacterium smegmatis*. We showed that deletion of the predicted components of the Tat translocase resulted in slow growth and hypersensitivity to SDS (26). This latter phenotype is suggestive of a loss in function due to the mislocalization of one or more proteins involved in cell envelope metabolism. We further showed that the chromosomally-encoded β -lactamases of *M. smegmatis* and *M. tuberculosis* were Tat-exported proteins. Notably these had predicted Tat signal sequences.

We initially planned to investigate the Tat pathway in *M. tuberculosis* using Δtat mutants. However, the *tat* genes proved to be essential in *M. tuberculosis* (36, 42). It is probable that one or more proteins exported by the *M. tuberculosis* Tat system are required for *in vitro* growth. Several predicted Tat substrates were previously identified as essential

for *in vitro* growth by TraSH analysis (42) and are therefore likely candidates for contributing to the essential nature of a *tat* mutant. Also, since *M. tuberculosis* is a slow grower, it is conceivable to think that a growth defect similar to that seen in Δ *tat* *M. smegmatis* (26), would result in a *M. tuberculosis* strain that takes significantly longer to grow on plates than the wild-type. The inability to obtain *tat* mutants in *M. tuberculosis* makes comparative studies between mutant and wild-type unfeasible. Therefore, alternative methods were needed to investigate the *M. tuberculosis* Tat pathway.

To gain insight into the contribution that Tat export plays in *M. tuberculosis* physiology and pathogenesis, we designed a strategy to identify the Tat-exported substrates of *M. tuberculosis* in a *M. smegmatis* background. To do this, we developed the native β -lactamase of *M. tuberculosis*, BlaC into a reporter for export by the Tat pathway. *M. tuberculosis* BlaC represents the first example of a genetic reporter specific for Tat export that can be used directly in a selection on agar plates. This can be a powerful method to identify Tat exported proteins in other bacterial systems. This idea is discussed in the next section.

We demonstrated that the twin-arginine signal sequences of *M. tuberculosis* phospholipase C (Plc) enzymes were capable of directing export via the Tat pathway (26). This data demonstrated these Plcs have functional Tat signal sequences and is consistent with them being exported by the Tat pathway. In turn, this would mean that Tat export contributes to virulence. However, ultimately we must demonstrate that the full length authentic Plc enzymes are true Tat-exported proteins. This analysis is underway and is discussed further in the following section.

Other than phospholipase C, we identified additional *M. tuberculosis* ORFs with signal sequences capable of directing the export of the BlaC reporter. These proteins have homology to lipases, substrate binding proteins, and redox proteins (see Chapter 3 Discussion); however, they have yet to be characterized in *M. tuberculosis*. Until their function is determined, it is unclear if they play a role in virulence. In addition, there may be other proteins exported by the Tat pathway that we have not identified due to limitations of our reporter system; therefore, alternative methods may be required to identify these proteins, as described in the next section.

A limitation of the library approach we employed is that it relied on *in vitro* growth conditions. We hypothesize that there is a subset of proteins that are exclusively exported by *M. tuberculosis* while the bacteria are inside macrophages and are therefore likely to be important for pathogenesis. This subset of proteins would not have been identified in the library selection described in Chapter 3. In Chapter 4, we describe the first-ever use of a β -lactamase reporter of export with intracellularly growing bacteria (24). We showed that the Δ *blaC* mutant of *M. tuberculosis* was susceptible to carbenicillin while growing inside macrophages. Further, we showed that a PlcB signal sequence fusion to BlaC was exported and protected Δ *blaC* *M. tuberculosis* from carbenicillin inside macrophages. Thus, there is potential to adapt the 'BlaC reporter and BlaC expression library, described in Chapter 3, in *M. tuberculosis* for the selection of strains that export BlaC fusions during bacterial growth in macrophages. This would be a novel means of identifying proteins exported exclusively in infection, and would be an important step in determining the role of the Tat pathway in the virulence of *M. tuberculosis*.

The Tat pathway is present in many bacteria, but it is not in mammalian cells. Thus, the Tat apparatus is a potential target for novel drug development. Our finding that the Tat pathway may be involved in the export of the Plc virulence factors suggests that inhibition of Tat export in *M. tuberculosis* would prevent localization of phospholipase C enzymes and result in an attenuation of virulence similar to that described with a $\Delta plcABC$ mutant. Furthermore, since the Tat pathway is essential for *in vitro* growth, we speculate that inhibition of Tat export by a specific drug would result in a bactericidal effect *in vivo*.

Better understanding of the properties of Tat export in bacteria

We undertook an unbiased experimental approach to identify *M. tuberculosis* proteins with functional Tat signal sequences. Using a random BlaC fusion library, we identified *M. tuberculosis* Tat signal sequences that drove export of the reporter by the Tat pathway. Most analyses of Tat export are based on initial bioinformatic prediction of Tat signal sequences (46, 48). Since the dataset of Tat-exported proteins from which these programs are created is biased towards a small number of prokaryotes (1, 34, 43), the programs may be limited in their ability to predict true substrates in mycobacteria. In fact, we showed in Chapter 3 (Fig. 3.4) that there was very little overlap between the outputs of three Tat signal sequence prediction programs when applied to the *M. tuberculosis* genome, thus strengthening this claim. The data obtained from our analysis was not dependent on bioinformatic analysis, and therefore represents an unbiased approach with potential to provide new information that could be used to better understand Tat-export in bacteria. Here, we present an overview of this information and provide speculation of the roles that both the signal sequences and mature domains of the identified proteins might play in Tat export. We propose that specific

chaperones or additional factors may be required for export of Tat-substrates in *M. tuberculosis*. Finally, we conclude with the idea that BlaC can be used as an easy-to-assay reporter for Tat-dependent export in other bacteria.

Using the BlaC reporter library, we identified ten signal sequences that drove the Tat export of a functional BlaC. These signal sequences fit the pattern of typical Tat signal sequences described to date, specifically with respect to the consensus twin-arginine motif. Interestingly, most of the proteins we identified were predicted by at least two of the bioinformatic programs described in Chapter 3. This suggests that the sequence elements predicted by all three programs may be the best indicators of Tat export and could be used to either improve the current programs or to design programs better suited for mycobacteria and related bacteria. The Rv0063 signal sequence was an exception to this rule since it lacked a twin-arginine dipeptide, making it at least the third reported example of a non-RR signal sequence in bacteria (17, 19). At this point in time, we do not know if the atypical motif of Rv0063 is specifically required for mycobacterial Tat export, or instead is a general, but rare property of Tat export in bacteria. Our analysis shows that an unbiased experimental method can provide new information on functional signal sequences. Below, we discuss ways in which we might identify additional Tat signal sequences using the BlaC reporter, which could yield more examples of Tat signal sequences with atypical Tat motifs.

Although we were successful in identifying Tat-exported fusion proteins using the BlaC reporter, we encountered an unexpected limitation to our strategy that may have prevented us from selecting additional fusions. As described in Chapter 3, we only obtained fusions of BlaC immediately downstream of the predicted signal sequence cleavage site of an ORF. This suggests that inclusion of the mature domain of a protein when fused to the 'BlaC

reporter prevents export. A likely explanation for this is that the foreign sequence disrupts the folding of BlaC and prevents its functional export by the Tat pathway. It also suggests that some ORFs might not be identified using our current library approach. The overlooked ORFs would be those that lack restriction sites positioned to generate in-frame '*blaC* fusions near the cleavage site. One way to address this problem would be to directly construct individual BlaC fusions to predicted Tat signal sequences and demonstrate export using the selection strategy we described. However, as we demonstrated in Chapter 3, a more comprehensive and unbiased approach might result in the identification of Tat signal sequences that do not fit the pattern of a typical Tat signal sequence. Since this is a category of proteins we are interested in identifying, an unbiased approach is preferable. An alternate approach would be to design a new library in which obtaining an in-frame fusion is not dependent on the presence of appropriately positioned restriction sites within an ORF. This could be achieved with a library constructed from random fragments of sonicated genomic DNA upstream of the '*blaC* reporter to be used in the same selection described in Chapter 3. This can potentially result in a more random and complex fusion library. With this approach, we would expect to select for fusions to the ORFs described in Chapter 3, in addition to new ORF fusions, thus expanding our list. Furthermore, we could perform the selection directly in *M. tuberculosis* to potentially identify fusion proteins that are not compatible with the *M. smegmatis* Tat system.

In a recent study in *E. coli*, a subset of Tat signal sequences show specificity to the Tat translocase, while others are promiscuous and can direct export of a reporter through either the Sec or Tat translocase (48). Because export is dependent on the folded state of a protein (8), it is not surprising that the mature domain is a factor in determining whether a

given protein can be exported by the Tat apparatus. This study also brings up the question: how often is it that a protein with a functional signal sequence is a true Tat substrate? To address this issue in mycobacteria, we showed preliminary analysis of export of the authentic full length proteins encoded from nine of the ten ORFs identified from the library. Using signal sequence processing as a readout for export, we detected a Tat-dependent effect for full length BlaC, Rv0315, Rv0774c, and Rv2843 by immunoblot. The remaining five proteins showed no obvious Tat-dependent difference. As described in Chapter 3, the lack of *Δtat* dependence for these five proteins could indicate that (i) the protein is exported by both Sec and Tat, (ii) the protein is exported by Sec only, (iii) precursor and mature species cannot be resolved on the gel system used, and therefore the use of signal sequence cleavage as a readout for export is not adequate, or (iv) the protein is not exported by *M. smegmatis*. For this last possibility, we cannot exclude the possibility that the C-terminal HA tag is preventing the proper folding and/or targeting of the full length protein to the export apparatus. Specific analyses to address these possibilities are ongoing and more work needs to be done before we can make a conclusion.

The possibility of the authentic protein being a Sec substrate could be explored by assessing precursor processing when Sec export is inhibited. Guo *et al.* (14) developed a method in which expression of a target gene is reduced in response to increased levels of a transcriptional repressor specific to that target gene. This method has been applied in *M. smegmatis* to demonstrate that expression of the essential *secA1* gene is downregulated in response to addition of the repressor (14). This strain can be used as a background for the introduction of our existing full length constructs. Here, we can utilize this method to potentially demonstrate Sec-dependent export of the full length proteins that show no Tat-

dependent effect. Specifically, we would look for evidence of a processing defect (accumulation of precursor) in response to decreased *secA1* expression. Finally, we can also deplete SecA1 in a Δ *tat* background to better demonstrate if a protein is exported in a truly Sec-dependent manner. The demonstration that a protein with a functional Tat signal sequence is exported by the Sec pathway suggests that the full length *M. tuberculosis* protein is unfolded in *M. smegmatis* and further work will be necessary to determine if this is indicative of the authentic protein in its native host.

Another factor that may help explain why some full length *M. tuberculosis* proteins fail to show Tat-dependence is the lack of an accessory factor in the *M. smegmatis* model required for export. Chaperones have been described in a small number of bacteria and in these cases, are required for folding of a precursor protein prior to export by the Tat pathway (15, 30). Of these few reported examples, most function in insertion of cofactors into the precursor protein (28). In contrast, the PlcR protein of *P. aeruginosa* is an example of a chaperone that is not involved in cofactor insertion but is still required for efficient folding of the PlcH phospholipase C prior to export (6). We demonstrated that full length PlcB may be processed in slow-growing *M. tuberculosis* and *M. bovis* BCG, but not in fast-growing *M. smegmatis* suggesting that *M. tuberculosis*-specific factors may be required for the export of full length PlcB. We are currently testing a full length PlcB mutant in which the twin-arginine dipeptide was substituted with twin-lysines. If PlcB is exported by the Tat pathway in *M. tuberculosis*, as our data suggests, we hypothesize that the PlcB[KK] protein would be defective for export and would therefore only be present as an uncleaved precursor form. The PlcB data suggests that additional *M. tuberculosis* factors may be required for the export of several other full length proteins that failed to show processing in *M. smegmatis*. Future

work may focus on the identification and characterization of specific accessory factors or chaperones that contribute to Tat export in *M. tuberculosis*. To date, no examples of Sec or Tat chaperones have been identified in *M. tuberculosis*, and therefore, this finding might be significant in better understanding the requirements for export in mycobacteria.

It is well established that only folded proteins are compatible for export by the Tat pathway and unfolded proteins are incompatible for export (8). To date, a number of Tat-specific reporters that fold prior to export have been described, including green fluorescent protein (GFP), which can be detected by fluorescence analysis in the periplasm (37, 47). In addition, Posey *et al.* (31) have shown that an *E. coli* Tat signal sequence fusion to GFP is exported in a Tat-dependent manner in *M. smegmatis* by screening. Recently, the use of GFP as a reporter for Tat export was enhanced by the addition of a small C-terminal tag (SsrA) that targets the GFP fusion proteins for degradation by the cytoplasmic ClpXP machinery (7). Using this approach, only exported fusion proteins escape degradation and accumulate in the periplasm where they can be detected by fluorescence analysis. The addition of the SsrA tag is a marked improvement of this system; however, flow cytometric screening is required for the identification of exported GFP-SsrA fusion proteins.

Our approach was to utilize the *M. tuberculosis* BlaC as a reporter for Tat-export. The use of a β -lactamase selection assay eliminates labor intensive screening and allows us to perform a simple assay for the presence or absence of viable colonies on agar media with carbenicillin. This is a classic approach to identifying exported proteins in bacteria (4, 29, 45, 46), although not in a Tat-dependent manner. However, until recently, it had not been performed directly in mycobacteria due to the presence of endogenous β -lactamase activity

(12, 49). We were able to circumvent this problem by using Δbla mutants of *M. smegmatis* and *M. tuberculosis*, which are sensitive to β -lactam antibiotics (12).

Many of the existing Tat-dependent reporters have been used successfully in a variety of bacterial hosts (20, 31, 52). We speculate that BlaC can be utilized as a reporter for Tat export in other bacteria, and given the above-mentioned advantages it has over existing reporters, would be an effective means for identifying Tat-exported proteins. This would be especially useful in *E. coli*, where using an easy-to-assay unbiased approach to identify Tat-exported proteins has not been performed and therefore might result in the identification of atypical Tat signal sequences. As mentioned earlier, this information may uncover novel properties of Tat export, such as a relaxed requirement for the RR dipeptide, therefore contributing to understanding of Tat biology overall. Finally, the use of this reporter in other pathogens might result in the identification of Tat-exported virulence factors. Since it appears that phospholipase C enzymes are a category of Tat-exported virulence factors (9, 35, 50), it is conceivable that other subsets of virulence factors may rely exclusively on the Tat pathway for export.

Exploiting the ability of the mycobacterial Tat pathway to export folded proteins of interest

The export of heterologous proteins in bacteria has been exploited for a variety of purposes, and until recently, was accomplished mainly by export through the Sec translocase (13). However, many proteins proved to be incompatible for export by the Sec pathway due to strict folding requirements. Because of its ability to export folded proteins, the Tat

pathway may be a useful alternative to this problem, especially when exporting foreign (non-native proteins) and therefore has important implications in biotechnology.

Here, we present two ways in which the mycobacterial Tat pathway might be utilized in the export of folded proteins. First, large-scale efforts to construct better TB vaccines include development of recombinant strains of the existing vaccine (rBCG) in an attempt to produce a lasting immune response (16). Export of large amounts of immunogenic antigens by the Tat pathway may be an effective means of localizing the proteins where they can interact efficiently with the host. By this rationale, the use of polyvalent rBCG strains has been proposed as a method to provide protection against other diseases (18). In this instance, the ability to efficiently export foreign proteins, which may fold in the bacterial cytoplasm, may rely on the Tat pathway.

Second, the overexpression of proteins in bacteria for purification and other analyses is often hampered by problems such as protein insolubility. These problems can sometimes be overcome by directing the secretion of the protein, and may be further improved by export through the Tat apparatus (13). This might be an important option in the purification of mycobacterial proteins, especially if expression directly in mycobacteria is desired. These two methods represent novel means of utilizing the Tat pathway in mycobacteria for the improved export of proteins and may be the foundation for many innovative studies to come.

References

1. **Bendtsen, J. D., H. Nielsen, D. Widdick, T. Palmer, and S. Brunak.** 2005. Prediction of twin-arginine signal peptides. *BMC Bioinformatics* **6**:167.
2. **Berks, B. C.** 1996. A common export pathway for proteins binding complex redox cofactors? *Mol Microbiol* **22**:393-404.
3. **Bronstein, P. A., M. Marrichi, S. Cartinhour, D. J. Schneider, and M. P. DeLisa.** 2005. Identification of a twin-arginine translocation system in *Pseudomonas syringae* pv. tomato DC3000 and its contribution to pathogenicity and fitness. *J Bacteriol* **187**:8450-61.
4. **Broome-Smith, J. K., M. Tadayyon, and Y. Zhang.** 1990. Beta-lactamase as a probe of membrane protein assembly and protein export. *Mol Microbiol* **4**:1637-44.
5. **Caldelari, I., S. Mann, C. Crooks, and T. Palmer.** 2006. The Tat pathway of the plant pathogen *Pseudomonas syringae* is required for optimal virulence. *Mol Plant Microbe Interact* **19**:200-12.
6. **Cota-Gomez, A., A. I. Vasil, J. Kadurugamuwa, T. J. Beveridge, H. P. Schweizer, and M. L. Vasil.** 1997. PlcR1 and PlcR2 are putative calcium-binding proteins required for secretion of the hemolytic phospholipase C of *Pseudomonas aeruginosa*. *Infect Immun* **65**:2904-13.
7. **DeLisa, M. P., P. Samuelson, T. Palmer, and G. Georgiou.** 2002. Genetic analysis of the twin arginine translocator secretion pathway in bacteria. *J Biol Chem* **277**:29825-31.
8. **DeLisa, M. P., D. Tullman, and G. Georgiou.** 2003. Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. *Proc Natl Acad Sci U S A* **100**:6115-20.
9. **Dilks, K., R. W. Rose, E. Hartmann, and M. Pohlschroder.** 2003. Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. *J Bacteriol* **185**:1478-83.
10. **Ding, Z., and P. J. Christie.** 2003. *Agrobacterium tumefaciens* twin-arginine-dependent translocation is important for virulence, flagellation, and chemotaxis but not type IV secretion. *J Bacteriol* **185**:760-71.
11. **Finlay, B. B., and S. Falkow.** 1997. Common themes in microbial pathogenicity revisited. *Microbiol Mol Biol Rev* **61**:136-69.

12. **Flores, A. R., L. M. Parsons, and M. S. Pavelka, Jr.** 2005. Genetic analysis of the beta-lactamases of *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* and susceptibility to beta-lactam antibiotics. *Microbiology* **151**:521-32.
13. **Georgiou, G., and L. Segatori.** 2005. Preparative expression of secreted proteins in bacteria: status report and future prospects. *Curr Opin Biotechnol* **16**:538-45.
14. **Guo, X. V., M. Monteleone, M. Klotzsche, A. Kamionka, W. Hillen, M. Braunstein, S. Ehrt, and D. Schnappinger.** 2007. Silencing essential protein secretion in *Mycobacterium smegmatis* using tetracycline repressors. *J Bacteriol.* **189**:4614-23.
15. **Hatzixanthis, K., T. Palmer, and F. Sargent.** 2003. A subset of bacterial inner membrane proteins integrated by the twin-arginine translocase. *Mol Microbiol* **49**:1377-90.
16. **Hernandez-Pando, R., M. Castanon, C. Espitia, and Y. Lopez-Vidal.** 2007. Recombinant BCG vaccine candidates. *Curr Mol Med* **7**:365-72.
17. **Hinsley, A. P., N. R. Stanley, T. Palmer, and B. C. Berks.** 2001. A naturally occurring bacterial Tat signal peptide lacking one of the 'invariant' arginine residues of the consensus targeting motif. *FEBS Lett* **497**:45-9.
18. **Honda, M., K. Matsuo, T. Nakasone, Y. Okamoto, H. Yoshizaki, K. Kitamura, W. Sugiura, K. Watanabe, Y. Fukushima, S. Haga, Y. Katsura, H. Tasaka, K. Komuro, T. Yamada, T. Asano, A. Yamazaki, and S. Yamazaki.** 1995. Protective immune responses induced by secretion of a chimeric soluble protein from a recombinant *Mycobacterium bovis* bacillus Calmette-Guerin vector candidate vaccine for human immunodeficiency virus type 1 in small animals. *Proc Natl Acad Sci U S A* **92**:10693-7.
19. **Ignatova, Z., C. Hornle, A. Nurk, and V. Kasche.** 2002. Unusual signal peptide directs penicillin amidase from *Escherichia coli* to the Tat translocation machinery. *Biochem Biophys Res Commun* **291**:146-9.
20. **Kikuchi, Y., M. Date, H. Itaya, K. Matsui, and L. F. Wu.** 2006. Functional analysis of the twin-arginine translocation pathway in *Corynebacterium glutamicum* ATCC 13869. *Appl Environ Microbiol* **72**:7183-92.
21. **Kurtz, S., and M. Braunstein.** 2005. Protein secretion and export in *Mycobacterium tuberculosis*, p. 71-138. *In* T. Parish (ed.), *Mycobacterium* molecular biology. Horizon bioscience, Norfolk, UK.
22. **Lavander, M., S. K. Ericsson, J. E. Broms, and A. Forsberg.** 2006. The twin arginine translocation system is essential for virulence of *Yersinia pseudotuberculosis*. *Infect Immun* **74**:1768-76.

23. **Lee, P. A., D. Tullman-Ercek, and G. Georgiou.** 2006. The bacterial twin-arginine translocation pathway. *Annu Rev Microbiol* **60**:373-395.
24. **McCann, J. R., J. A. McDonough, M. S. Pavelka, Jr., and M. Braunstein.** 2007. Beta-lactamase can function as a reporter of bacterial protein export during *Mycobacterium tuberculosis* infection of host cells. *Microbiology*, **153**:3350-9.
25. **McDonough, J. A., and M. Braunstein.** 2007. Protein Transport in *Mycobacterium tuberculosis*. In S. H. Kaufmann and E. J. Rubin (ed.), *MTB Handbook Volume 1: Molecular Genetics and Biochemistry*, vol. In press. Wiley-VCH Verlag Press.
26. **McDonough, J. A., K. E. Hacker, A. R. Flores, M. S. Pavelka, Jr., and M. Braunstein.** 2005. The twin-arginine translocation pathway of *Mycobacterium smegmatis* is functional and required for the export of mycobacterial beta-lactamases. *J Bacteriol* **187**:7667-79.
27. **Meya, D. B., and K. P. McAdam.** 2007. The TB pandemic: an old problem seeking new solutions. *J Intern Med* **261**:309-29.
28. **Palmer, T., F. Sargent, and B. C. Berks.** 2005. Export of complex cofactor-containing proteins by the bacterial Tat pathway. *Trends Microbiol* **13**:175-80.
29. **Palva, I., M. Sarvas, P. Lehtovaara, M. Sibakov, and L. Kaariainen.** 1982. Secretion of *Escherichia coli* beta-lactamase from *Bacillus subtilis* by the aid of alpha-amylase signal sequence. *Proc Natl Acad Sci U S A* **79**:5582-6.
30. **Perez-Rodriguez, R., A. C. Fisher, J. D. Perlmutter, M. G. Hicks, A. Chanal, C. L. Santini, L. F. Wu, T. Palmer, and M. P. Delisa.** 2007. An essential role for the DnaK molecular chaperone in stabilizing over-expressed substrate proteins of the bacterial twin-arginine translocation pathway. *J Mol Biol.* **367**:715-30.
31. **Posey, J. E., T. M. Shinnick, and F. D. Quinn.** 2006. Characterization of the twin-arginine translocase secretion system of *Mycobacterium smegmatis*. *J Bacteriol* **188**:1332-40.
32. **Pradel, N., C. Ye, V. Livrelli, J. Xu, B. Joly, and L. F. Wu.** 2003. Contribution of the twin arginine translocation system to the virulence of enterohemorrhagic *Escherichia coli* O157:H7. *Infect Immun* **71**:4908-4916.
33. **Raynaud, C., C. Guilhot, J. Rauzier, Y. Bordat, V. Pelicic, R. Manganelli, I. Smith, B. Gicquel, and M. Jackson.** 2002. Phospholipases C are involved in the virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* **45**:203-17.

34. **Rose, R. W., T. Bruser, J. C. Kissinger, and M. Pohlschroder.** 2002. Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. *Mol Microbiol* **45**:943-50.
35. **Rossier, O., and N. P. Cianciotto.** 2005. The *Legionella pneumophila* *tatB* gene facilitates secretion of phospholipase C, growth under iron-limiting conditions, and intracellular infection. *Infect Immun* **73**:2020-32.
36. **Saint-Joanis, B., C. Demangel, M. Jackson, P. Brodin, L. Marsollier, H. Boshoff, and S. T. Cole.** 2006. Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of *Mycobacterium tuberculosis*, increases {beta}-lactam susceptibility and virulence. *J Bacteriol* **188**:6669-79.
37. **Santini, C. L., A. Bernadac, M. Zhang, A. Chanal, B. Ize, C. Blanco, and L. F. Wu.** 2001. Translocation of jellyfish green fluorescent protein via the Tat system of *Escherichia coli* and change of its periplasmic localization in response to osmotic up-shock. *J Biol Chem* **276**:8159-64.
38. **Santini, C. L., B. Ize, A. Chanal, M. Muller, G. Giordano, and L. F. Wu.** 1998. A novel sec-independent periplasmic protein translocation pathway in *Escherichia coli*. *EMBO J* **17**:101-12.
39. **Sargent, F., B. C. Berks, and T. Palmer.** 2006. Pathfinders and trailblazers: a prokaryotic targeting system for transport of folded proteins. *FEMS Microbiol Lett* **254**:198-207.
40. **Sargent, F., E. G. Bogsch, N. R. Stanley, M. Wexler, C. Robinson, B. C. Berks, and T. Palmer.** 1998. Overlapping functions of components of a bacterial Sec-independent protein export pathway. *EMBO J* **17**:3640-50.
41. **Sargent, F., N. R. Stanley, B. C. Berks, and T. Palmer.** 1999. Sec-independent protein translocation in *Escherichia coli*. A distinct and pivotal role for the TatB protein. *J Biol Chem* **274**:36073-82.
42. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48**:77-84.
43. **Selengut, J. D., D. H. Haft, T. Davidsen, A. Ganapathy, M. Gwinn-Giglio, W. C. Nelson, A. R. Richter, and O. White.** 2007. TIGRFAMs and Genome Properties: tools for the assignment of molecular function and biological process in prokaryotic genomes. *Nucleic Acids Res* **35**:D260-4.
44. **Settles, A. M., A. Yonetani, A. Baron, D. R. Bush, K. Cline, and R. Martienssen.** 1997. Sec-independent protein translocation by the maize Hcf106 protein. *Science* **278**:1467-70.

45. **Smith, H., S. Bron, J. Van Ee, and G. Venema.** 1987. Construction and use of signal sequence selection vectors in *Escherichia coli* and *Bacillus subtilis*. *J Bacteriol* **169**:3321-8.
46. **Stanley, N. R., F. Sargent, G. Buchanan, J. Shi, V. Stewart, T. Palmer, and B. C. Berks.** 2002. Behaviour of topological marker proteins targeted to the Tat protein transport pathway. *Mol Microbiol* **43**:1005-21.
47. **Thomas, J. D., R. A. Daniel, J. Errington, and C. Robinson.** 2001. Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in *Escherichia coli*. *Mol Microbiol* **39**:47-53.
48. **Tullman-Ercek, D., M. P. Delisa, Y. Kawarasaki, P. Iranpour, B. Ribnicky, T. Palmer, and G. Georgiou.** 2007. Export pathway selectivity of *Escherichia coli* twin-arginine translocation signal peptides. *J Biol Chem.* **282**:8309-16.
49. **Voladri, R. K., D. L. Lakey, S. H. Hennigan, B. E. Menzies, K. M. Edwards, and D. S. Kernodle.** 1998. Recombinant expression and characterization of the major beta-lactamase of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **42**:1375-81.
50. **Voulhoux, R., G. Ball, B. Ize, M. L. Vasil, A. Lazdunski, L. F. Wu, and A. Filloux.** 2001. Involvement of the twin-arginine translocation system in protein secretion via the type II pathway. *EMBO J* **20**:6735-41.
51. **Weiner, J. H., P. T. Bilous, G. M. Shaw, S. P. Lubitz, L. Frost, G. H. Thomas, J. A. Cole, and R. J. Turner.** 1998. A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell* **93**:93-101.
52. **Widdick, D. A., K. Dilks, G. Chandra, A. Bottrill, M. Naldrett, M. Pohlschroder, and T. Palmer.** 2006. The twin-arginine translocation pathway is a major route of protein export in *Streptomyces coelicolor*. *Proc Natl Acad Sci U S A.* **103**:17927-32.
53. **World Health Organization** 2007, posting date. WHO Information tuberculosis fact sheet. [Online.]