# IDENTIFICATION AND CHARACTERIZATION OF NUTRIENT TRANSPORTERS IN MYCOBACTERIUM SMEGMATIS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Michiei Sho

May 2011

IDENTIFICATION AND CHARACTERIZATION OF NUTRIENT TRANSPORTERS IN MYCOBACTERIUM SMEGMATIS

Michiei Sho, Ph.D.

Cornell University 2011

Finding cures against tuberculosis (TB) is an on-going battle for humans. TB is one of the world's deadliest diseases and one third of the world's population are infected with TB although only 10 % of the TB infected population developed to TB disease. Decline of TB disease was observed in the United States from over 16,000 to just below 8000 between 1993 and 2008 [13]. Worldwide, an incidence of TB disease increased from 125 cases per 100,000 population in 1990 to 142 cases per 100,000 population in 2004. This increase was primarily due to prevalence of the human immunodeficiency virus (HIV) infection among TB patients and a lack of access to proper treatments among them. In fact, mortality of the TB patients co-infected with HIV was drastically reduced when the patients were treated with antiretroviral therapy (ART) along with TB chemotherapy [13]. Without the concurrent treatment of HIV, up to 50 % of the patients with TB/HIV co-infection will die during the TB treatment.

There are 10 drugs currently approved by the U.S. Food and Drug Administration (FDA) for use in TB chemotherapy. Of those, a cocktail of isoniazid (INH), rifampin (RIF), ethambutal (EMB), and pyrazinamide (PZA) is used as a frontline anti-TB agents, and patients are treated between 4 to 7 months. This long-term anti-TB regimen works efficient as long as the patients religiously take the medication throughout the treatment period. However, the length of this regimen is long and side effects are severe that many patients ended the

regimen without completion after they started to feel better. This inappropriate termination of the regimen lead to the emergence of multi-drug resistant TB (MDR-TB). The outbreak of MDR-TB cases in HIV infected persons has contributed to increase of TB incidents in the U.S.A. from the mid-1980's through early 1990's [12, 81, 1, 91]: TB disease was otherwise believed to be eradicated like smallpox.

All the TB drugs target the genes in metabolic pathways therefore they are more potent against actively growing bacilli. However, slowly growing or metabolizing bacilli (non-replicative bacilli) are not eliminated by the TB drugs thus remain persistent in patients. The non-replicative bacilli can become active by yet unknown environmental factors and re-emerge to develop TB disease long after initial eradication. Also, it is these non-replicative bacilli, which are prone to undergo mutation and emerge as mutlidrug resistant strains [17]. Therefore, it is critical to develop drugs against the non-replicative bacilli. Understanding the biology of the non-replicative bacilli will give us a better strategy to eradicate TB.

The purpose of my Ph.D. dissertation is to contribute to the understanding of how the bacilli acquire nutrients. The non-replicative bacilli are believed to be stable and non-propagative, shown as the loss of acid-fastness. They are also viable but not culturable (VBNC) [15] at this stage, however they still require nutrients to survive. Most transporter studies are focused on drug efflux systems whereas many of the transporters involved in nutrient acquisition remain unknown. Understanding of nutrient uptake systems not only provides us with information as to how the bacilli acquire nutrients but also allows us to design drugs, which can gain access into the cells through essential pathways.

## BIOGRAPHICAL SKETCH

Michiei Sho was born in Osaka, Japan. She became fascinated in science at an early age, when she saw documentaries on how scientific findings unraveled the world's mysteries and improved our daily life. She was particularly drawn into the world of microorganisms. She was fascinated by the unlimited capacities of the tiny life forms, some of which can harm humans by causing infectious diseases and some of them can remediate the environments once contaminated by humans.

She received her undergraduate degree at McGill University in Ste-Anne de Bellevue, Quebec, Canada, where she studied Environmental Microbiology and wrote her senior thesis on development of community DNA extraction techniques from environmental samples. She was supervised by Dr. Charles W. Greer.

In 2000, she pursued M.Sc. with Dr. Chalres W. Greer at Biotechnology Research Institute (BRI) in Montreal and worked on characterization of a bacterium capable of degrading toxic polyaromatic hydrocarbons (PAHs), pyrene and phenanthrene. In this work, she identified the bacterium as *Mycobacterium* sp. Strain S65 and further characterized the genes involved in the degradation using molecular genetics.

In 2004, after working as a research assistant for Drs. Roland Brousseau and Luke Masson at BRI, she enrolled to a Ph.D. program in Microbiology at Cornell University to further pursue her passion in science. She joined the Dr. David G. Russell lab in 2005 and studied nutrient metabolism and acquisition in mycobacteria, mainly in *Mycobacterium smegmatis*.

During her Ph.D. at Cornell, she joined the Big Red Barn, the Graduate and Professional Student Center, originally as a summer staff, later as an assistant

manager for operations. There, she enjoyed meeting students and faculty in different disciplines, and had stimulated discussions with her peers.

After completion of Ph.D. at Cornell, she wishes to pursue her scientific career in environmental microbiology, particularly in the fields where she can study microbial metabolism/physiology in the native environments.

I would like to dedicate this Ph.D. dissertation to my Mother, Naomi Sho. I would not be able to come this far without her support.

**ACKNOWLEDGEMENTS**

First and foremost, I would like to thank my supervisor, Dr. David Russell, for giving opportunities to explore to research fields outside of my speciality. His enthusiasm and curiosity in learning and integrating new technologies strongly stimulated my interest in research and allowed me to explore beyond my research field. I have gained strength and knowledge to become an independent scientist.

I would like to thank Dr. Tadhg P. Begley for providing a lot of resources to help me learn how to make small molecules. Special thanks to his post-doc, Dr. David Hilmey, who taught me how to synthesize molecules with his passion and enthusiasm. I was not interested in organic chemistry until I met him. The fear I had had in organic chemistry disappeared and now I can appreciate how organic chemistry can help understanding biological processes.

I would also like to thank my minor committee members Dr. Joseph E. Peters and Dr. Hening Lin whose contribution provided a broad aspect into my project. This multidisciplinary team made possible to complete my dissertation.

Special thank to Linda Bennett for maintaining the lab so well that we are always pleasant to work. Without her, we would not know where to find reagents and equipments. She is a wonderful lab mother.

During my Ph.D., I had worked at Big Red Barn to have fresh air outside of my research. Working there really helped going through some of the tough time I had in my research. Kris Corda, the Director of BRB, became such a wonderful mentor to me and helped me work out my problems. I would not be here if I were not working there.

I have met many people here at Cornell University but Daniela Bocioaga is a special person who has become my best friend. I have not expected to have

such a amazing friend here. We have a very strong friendship which helped keeping my level of sanity many times during my Ph.D. program.

Another person I met who has a special space in my heart is Miloš Hašan. I do not know how to describe in words how important he is in my life. He supported me with his kindness, patience, and encouragement to complete my dissertation. I am a very lucky person to have him in my life.

Lastly but most importantly, I would like to thank my family for letting me pursue my passion in science. It must have been very difficult for them to have their daughter living so far and not seeing her often. I am very grateful for their understanding of my view for a future career.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

CHAPTER 1

# INTRODUCTION

## 1.1 *Mycobacterium tuberculosis*: the pathogen evolved with humans

*Mycobacterium tuberculosis* is a rod-shape, acid-fast Gram-positive bacterium (Figure 1.1) in the order Actinomycetales. The most prominent feature is its very lipid-rich cell wall, mycolic acids. *Mtb* is almost exclusively associated with its primary host, the human and is not found in a free form. In the host, it typically resides in a specialized cell, macrophage (Figure 1.2); the infected macrophages which later develop into granuloma, a complex structure comprising of many different mononuclear cells.

*Mtb* is the etiological agent of the most deadly infectious disease known to humans, TB, which has caused the most deaths in humans as a single microorganism [23]. It was first suggested by Benjamin Martin, an English physician in his book, "A New Theory of Consumption (1720)" that TB might be infectious

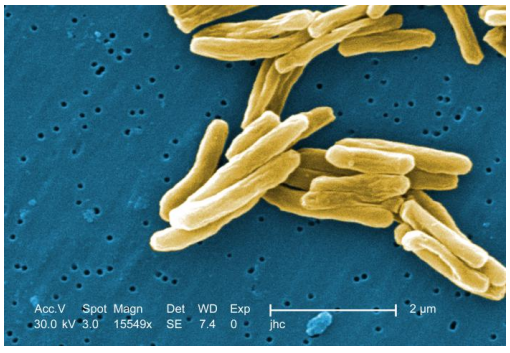



Figure 1.1: Scanning electron micrograph of *Mtb*. source: Centers for Disease Control and Prevention (CDC)

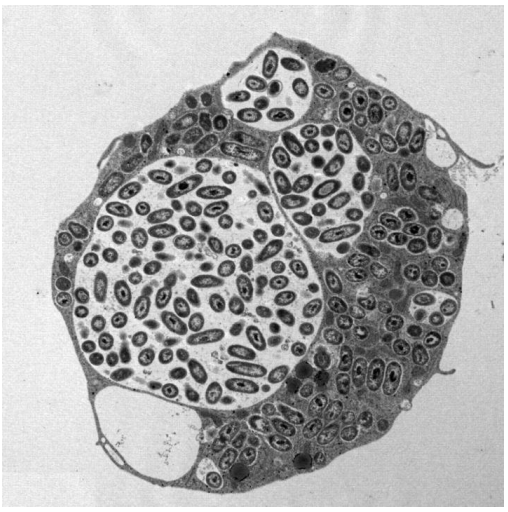


Figure 1.2: Electron micrograph of macrophage loaded with a massive number of *Mtb* bacilli. Picture was extracted from a review by Russell *et al* [75].

and caused by "some certain species of *animalcula*." Since then, TB was called as phthisis, a disease of consumption, or a heritable disease [23]. It was not shown to be caused by infection until 1882, when Robert Koch first identified *Mtb* as the etiological agent of TB and formulated a set of criteria, now known as Koch's postulates, to prove a link between a specific microorganism and a specific disease.

Origin of *Mtb* remains elusive. *Mtb* forms a MTB complex with other mycobacterial pathogens of humans and animals, namely *M. bovis*, *M. bovis* BCG, *M. africanum*, *M. caprae*, *M. microti*, *M. pinnipedii*, and *M. canetti* [86]. All the members of the MTB complex cause TB-like disease in their distinct hosts. They share more than 99.95 % sequence similarity at the nucleotide level. They are distinguished from each other by the distribution pattern of single nucleotide polymorphism (SNPs) and deletions. Most notable marker to differentiate the members of the MTB complex is the deletion of chromosomal region of difference 9 (RD9). RD9 is deleted from all the members which are primarily animal

adapted. There was this persistent idea that *Mtb* was derived from the bovine strain because people believed there might be a possible linkage between the domestication of cattle and the origin of TB in humans with the cattle being less clean than humans. However the smaller genome size of *M. bovis* and the deletion of RD9 in the *M. bovis* genome indicate that *Mtb* did not derive from *M. bovis* [36]. Molecular phylogeny of the MTB complex supports this conclusion and implies there is an ancestral strain for all the members of the MTB complex with a larger chromosome [84]. This ancestral strain, the most recent common ancestor (MRCA), is believed to be the first species which caused TB in humans, and the members of the MTB complex speciated with their adapted animals [84]. The genus *Mycobacterium* includes about 120 or more described species with a growing number of new species isolated from hospital patients [71, 93]. *Mycobacterium spp.* can be divided into 2 different categories: fast-growing and slow-growing. Fast-growers are typically non-pathogenic saprophytic species of *Mycbacterium* and the slow-growers are pathogenic species of *Mycobacterium*. The most well-known species of fast-grower is *M. smegmatis*. The most well-known species of slow-grower is *Mtb*. All the species of the *Mycobacterium* have at least 94.3 % sequence similarity in their 16S rRNA gene but interestingly, the fast- and slow-growing *Mycobacterium spp.* can be easily differentiated in the phylogenetic tree (Fig. 1.3). In addition to the marked phylogenetic pattern of 16S rRNA genes between pathogenic and non-pathogenic species, the genome sizes between pathogenic and non-pathogenic species also differ: non-pathogenic species possess a larger genome size compared to pathogenic species. For example, *M. smegmatis* has a genome size of 7 Mb compared to *Mtb*, whose genome is only 4.4 Mb in size. The etiological agent of leprosy, *M. leprae*, only has 3.31 Mb in its genome [21, 57]. In case of *M. leprae*, which not only has a

small genome but exceptionally large number of pseudogenes [20, 21, 88]. These small genomes in the slow-growing *Mycobacterium spp.* are believed as a result of the extensive reduction and rearrangement of their genomes through co-evolution with their specialized niches (hosts) because the hosts provide many necessary nutrients. This reduction of genome (= reductive evolution) is best reflected in their doubling times: it takes between 15-20 h doubling time for *Mtb* while only 2 h is required for *M. smegmatis* in the 7H9 broth. In case of *M. leprae*, it has a doubling time of 14 days but not in the 7H9 broth. It is unculturable in the synthetic media and is maintained on the feet of SCID mouse or nine banded armadillo [27]. *M. leprae* demonstrates an extensive case of reductive evolution, and tight association with its specialized host in adopting a parasitic lifestyle.

The granuloma in which *Mtb* is constrained does not function as a compartment site in which *Mtb* is killed but instead it prevents the dissemination of *Mtb* to other tissues. From the view of *Mtb*, granuloma provides a containment where the bacilli can hide away from the host's immune system and yet retain access to nutrients. The granuloma is a roughly spherical mass with *Mtb*-infected macrophages as a centered building block, which recruit mononuclear cells and triggers differentiation of macrophages into multinuclear giant cells (MGC) and epitheloid cells. Oxygenated mycolic acids released from *Mtb* triggers the formation of lipid-rich foamy macrophages (FM) [65]. The granuloma is extremely dynamic in structure and forms microenvironments in each layer (Fig. 1.4). Thus, depending on where in the granuloma *Mtb* resides, the available nutrients differ [76], which likely require a shift in bacterial metabolism.

The nutrients available to *Mtb* in the granuloma remain elusive. Yet an increasing body of evidence indicates that fatty acids are the major carbon

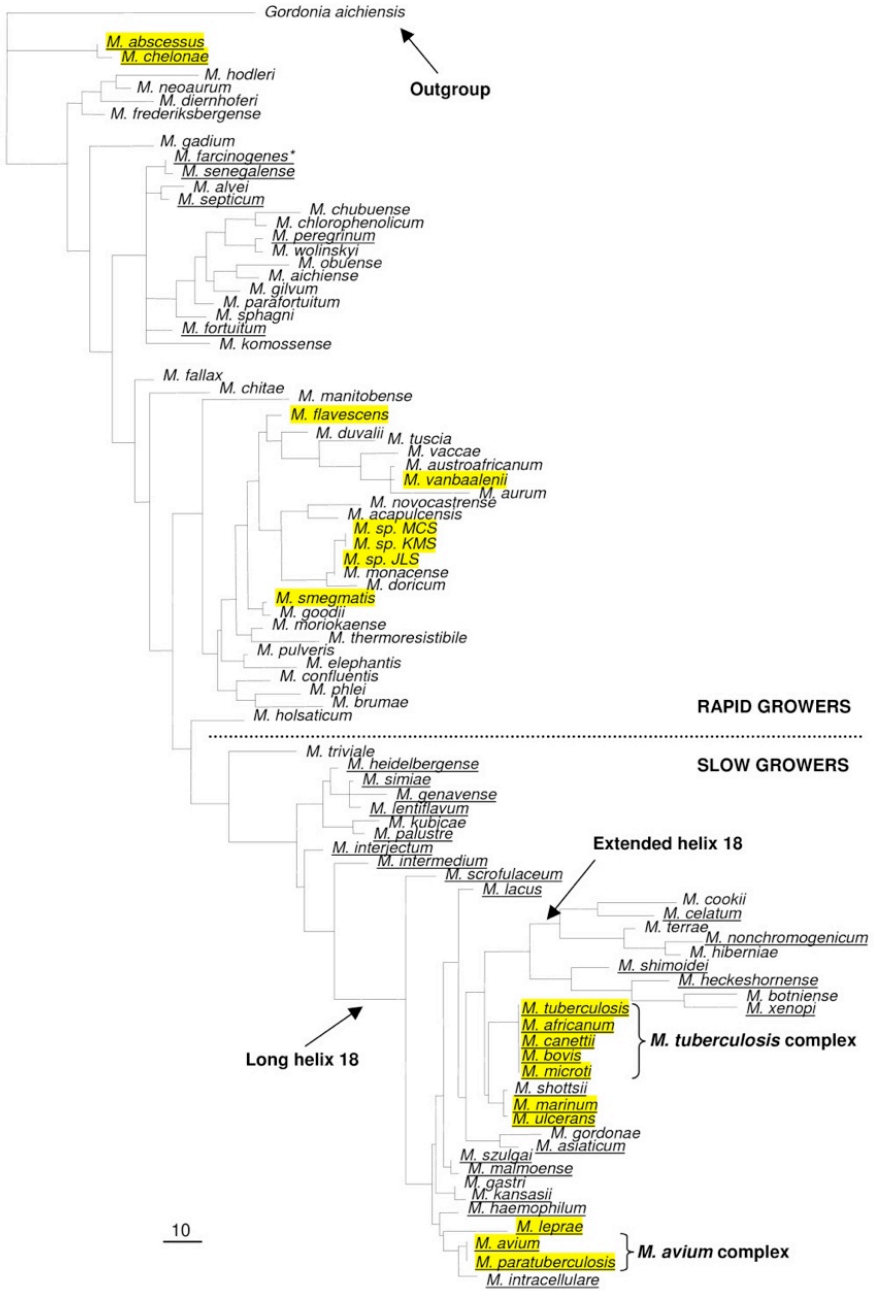



Figure 1.3: A phylogenetic tree of *Mycobacterium* species based on 16S rRNA sequence. Genome sequences of the species highlighted in yellow have been sequenced. The underlined species are considered pathogens. *Gordonia aichiensis* was used as the outgroup.The tree is extracted from Gey van Pittius *et al.*[33].

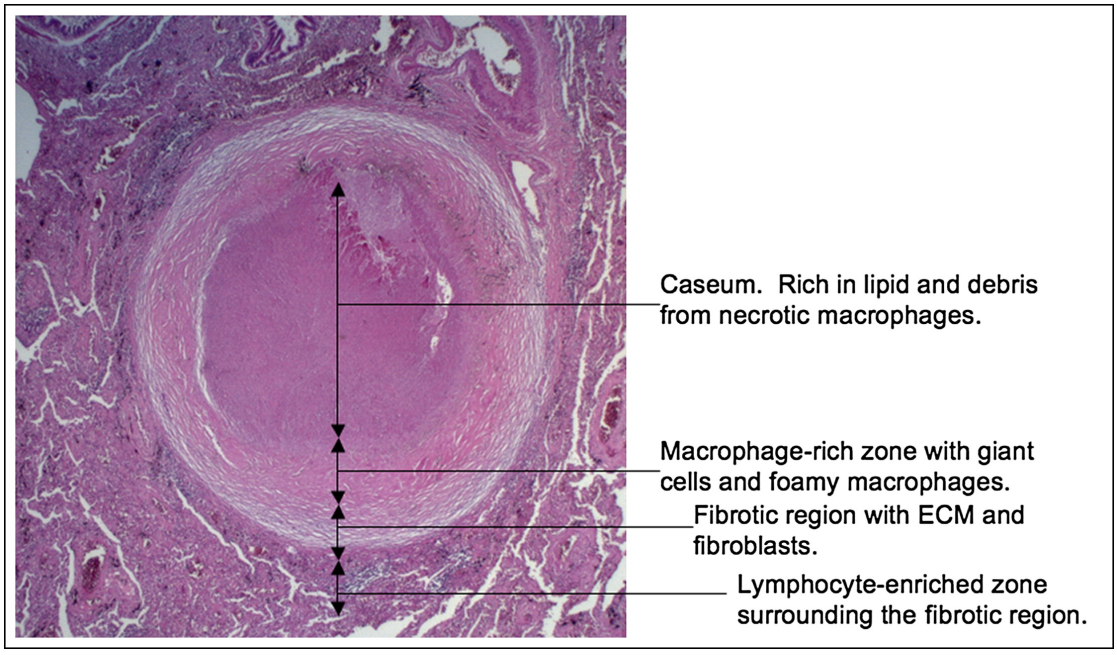



Figure 1.4: Structure of human TB granuloma. The figure is extracted from Russell DG *et al.*[76].

source and energy for *Mtb* in host cells [65]. Analyses of fatty acid composition in the caseum demonstrated that cholesterol, cholesterol esters, triacylglycerols and lactosylceramide are the major species of fatty acid in a caseum center of granuloma. Host genes involved in lipid sequestration and metabolism were highly expressed in the region [48]. Also, transcription profiling of *Mtb* provide a strong evidence that the genes involved in a $C_2$-based metabolism, fatty acid metabolism, gluconeogenesis, and glycerol-2P uptake and metabolism are highly upregulated in bacteria in macrophages and animal models [94, 64, 56, 14, 31]. Together with the loss of several glycolytic enzymes [78], this argues that host lipids are important sources of energy and carbon for *Mtb*.

The numerous studies on transcription profiling of *Mtb* reveal the central

role that carbon metabolism plays in *Mtb* infection. However, the identity of lipid species used by *Mtb* need to be better defined, and how these lipids are acquired by *Mtb* remains a mystery. Analysis of lipid species in granuloma shows the abundance of cholesterol species within granuloma [48] and *Mtb* is able to utilize cholesterol as a source of carbon source. Interestingly, strong attenuation of *Mtb* with a mutation in *mce4*, the gene encoding cholesterol uptake transporter, was observed in murine infection model [64]. Recent studies shed some light in this field, but there are many lipid species other than cholesterol present in granuloma that likely available to *Mtb* as nutrients. What makes it complex is that available lipids would change depending on the layer of granuloma and the stage of infection. Due to extremely hydrophobic characteristics (and some can be very complexed in structure) of lipids, it is reasonable to believe that dedicated transporters are involved in the uptake of different lipids like Mce4 for cholesterol uptake.

A lot of studies emphasize lipids as essential nutrients for *Mtb* during infection, but there are many nutrients other than lipids which are also important for persistence of *Mtb*. Amino acids and micronutrients (trace elements such as iron, magnesium, manganese, etc.) are only a few examples among them. Two ABC transporters, dipeptide (Dpp) and oligopeptide (Opp), were shown to transport peptides into the cell [32]. Commonly peptides less than 2 amino acid residues are transported via Dpp transporter and the peptides longer than 3 amino acid residues are taken up by Opp transporter. Interestingly, the Opp transporter of *Mtb* exhibits functions more than just a peptide transporter: it modulates the host innate immune response by triggering cytokine release and inducing apoptosis of *Mtb* infected macrophages [24]. How does it trigger cytokine release? It turns out *Mtb* impairs the glutathione-redox balance of the

host by importing glutathione via the Opp transporter. Glutathione is important to detoxify methylglyoxal, a toxic by-product from several metabolic pathways in the host, into lactate. By importing glutathione, the concentration of glutathione in the host is depleted, leading to an increase concentration of methylglyoxal resulting in induction of apoptosis of *Mtb* infected macrophages [24].

Another nutrient transporter characterized in *Mtb* is the trehalose transporter, LpqY-SugA-SugB-SugC. It is believed that *Mtb* has limited access to host sugars (based on the transcriptional profiles), yet it encodes for 5 putative sugar transporters [92] including the trehalose transporter. It is unlikely that the trehalose *Mtb* imports is host derived because most sugars would be already broken down into small pieces by the time they are available to *Mtb* [31]. The trehalose which *Mtb* imports therefore must be the trehalose released from the bacterium itself. *Mtb* with mutation in the trehalose transporter genes was attenuated in the mouse infection model, indicating that the recycling mechanism plays an important role in cell wall remodeling. Immunopathological studies of *Mtb* lacking the Opp transporter or the trehalose transporter suggest that *Mtb* exploits the nutrient acquisition systems to maintain virulence in the host [31, 24]. Discrepancy of available host nutrients and the substrate specificities of *Mtb* nutrient transporters might be explained by the additional functions of these nutrient transporters during infection.

The tight link between the roles of some transporters and the pathogenicity of *Mtb* clearly demonstrates the need to understand the transporters required to maintain the infection.

## 1.2 Transporters: The Passage to Essential Nutrients

### 1.2.1 Transport Systems in Bacteria

Bacteria possess two major hurdles to transport molecules into the cells: inner membrane (IM) and outer membrane (OM) in Gram negative bacteria, or thick peptidoglycan cell wall and IM in Gram positive bacteria. Substrates need to gain access to the periplasmic space before being transported into the cytoplasmic space by an ABC importer. The known pathways are via diffusion, a non-specific porin, or specialized porin. In Gram-negative bacteria, the non-specific porins such as OmpF and OmpC allow small substrates (< 650 Da) to cross the outer membrane [61]. Specialized porins such as maltoporin, LamB, and vitamin $B_{12}$ porin, BtuB, initiate uptake of only the specific substrates. In Gram-positive bacteria, it is not known how substrates cross the thick peptidoglycan cell wall. Do Gram positive bacteria possess porins? Gram positive bacteria were long believed to lack porins especially in *Mycobacterium spp.*, which possess highly specialized cell walls for Gram positive bacteria. However, the presence of porins was demonstrated in *M. smegmatis* and it was shown that nutrients such as carbohydrates, amino acids and phosphate pass across the mycolic acid-containing cell wall via the porins [60].

ATP-binding cassette (ABC) transporter system is the largest family of all the transporter families and is largely divided into 3 classes. Mapping of all the known ABC transporters from the 3 domains of life, i.e., in archaea, eukaryotes, and prokaryotes, reveals that all the classes of transporters are found in all 3 domains, indicating that the ABC transporter classes were specialized before life diverged into different domains [25].

**Architecture of ABC transporter**

In all classes of ABC transporters, the common denominators are 2 hydrophobic membrane-spanning (transmembrane) domains (TMDs) and 2 hydrophilic nucleotide-binding domains (NBDs), which are tightly associated with TMDs in the cytosolic side [41]. Most class I ABC transporters have fused TMDs and NBDs (Fig. 1.5A). The common architecture is either the N- or C-terminus of a NBD is fused to TMD "half-size transporter.") Efflux transporters are mostly found in this class. Class II ABC transporters have only NBD domains and are believed to play a role in regulation and/or gene expression (Fig. 1.5B). In class III, all the domains are independently produced and formed a multi-subunit transporter complex (Fig. 1.5C). ABC importers belonging to the class III are exclusively found in archaea and prokaryotes. The Opp and trehalose transporters are ABC importers.

**Architecture of ABC importer**

An ABC importer was the first ABC transporter characterized [41]. Unlike the conventional architecture of ABC transporters, that of an ABC importer is unique because of an additional accessory protein called substrate binding protein (SBP). SBP is a soluble protein freely moving around in periplasmic space in Gram negative bacteria. In Gram positive bacteria, SBP is either anchored to the membrane via lipids or is fused with one of TMDs (Fig. 1.5C). SBP mainly functions to scavenge substrates and is only found in the ABC importer system.

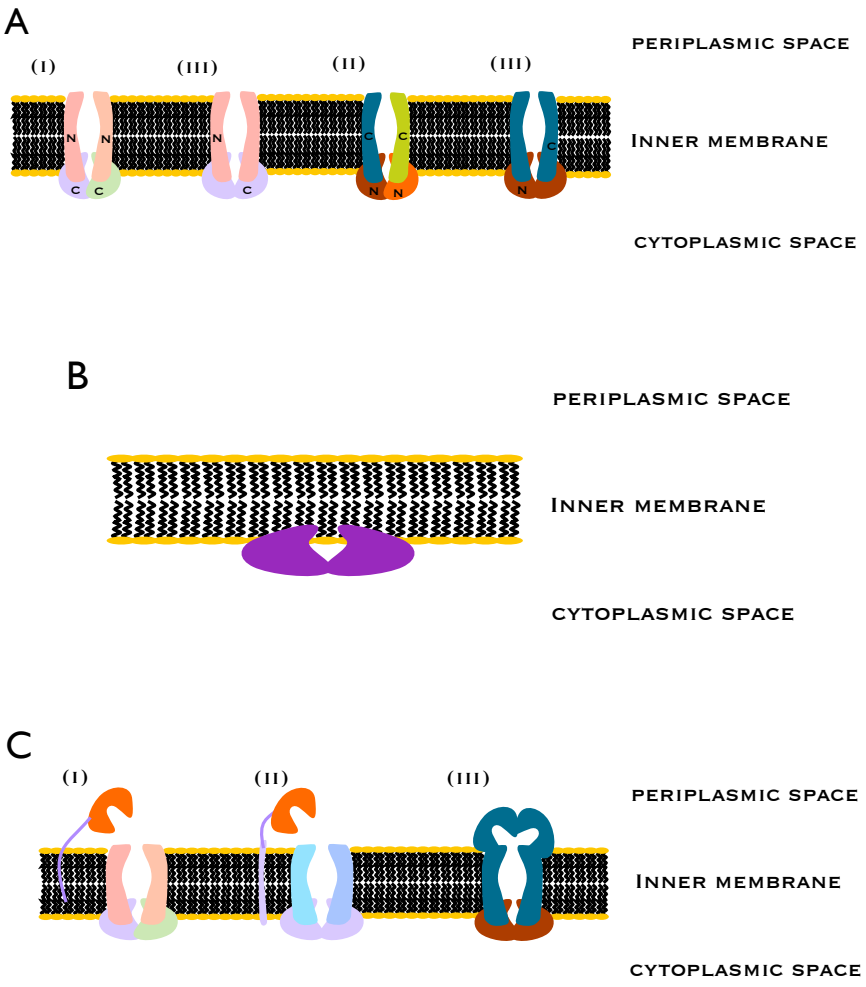



Figure 1.5: Architecture of ABC transporters.(A) Class I ABC transporters have TMD and NBD fused together and this class of transporters is often seen in exporters and is abundant in eukaryotes. (i) "N" in the TMD and "C" in NBD indicate that N-terminal TMD is fused to C-terminal NBD. (ii) Inverted organization was also observed. (iii) Some transporters have two domains of TMD and two domains of NBD fused together. (B) ABC transporters with no associated TMD belong to class II ABC transporter family. The two domains of NBD are fused together. (C) Class III ABC transporter family has SBP in addition. SBP works as a high affinity substrate binding protein. SBP may be simply attached to the membrane via (i) lipid-anchor, (ii) transmembrane-peptide or (iii) as a fused protein with TMDs.

**Determining factor(s) of substrate affinity in ABC importer**

The most important characteristic of SBP is its high affinity to specific substrates. It can transport substrates against concentration gradients greater than $10^5$-fold molarity from the extracellular space to the cytoplasmic space [40]. When organisms are in environments where specific nutrients are very limited, the ABC importer can scavenge as much nutrient as possible, mediated by the high affinity SBP protein. Some experiments however demonstrated that SBP is not necessarily required to transport a substrate into the cells [82]. Maltose binding protein, MalE was deleted from *Escherichia coli* and transport of maltose was tested in the absence of MalE. In the study, the remaining transporter complex, referred to translocator complex, demonstrated the uptake of maltose without MalE, indicating that MalE is not an absolute requirement. However, what this study did not report was kinetic change. It is true that the translocator complex can facilitate uptake of specific substrates without a substrate-binding protein, but at slower rate. After numerous biochemical and structural studies, it has been concluded that both substrate-binding protein and translocator complex are necessary to facilitate high-affinity substrate uptake, with the substrate-binding protein mediating pooling of specific substrate(s) in the proximity to the translocator complex for efficient transportation [25].

Specificity of substrates in a given SBP is predominantly determined by hydrogen bonding [41]. Some SBPs exhibit such high substrate specificity that they interact with only one specific substrate [46], while others like the SBP of the FhuBDC transporter system, can interact with substrates of diverse structures, i.e., coprogen, ferrichrome, and aerobactin [25]. OppA, SBP of oligopeptide ABC transporter, of *Lactococcus lactis* is also known to transport peptides of

lengths between 3 and 35 amino acids [28]. How could $OppA_{L.i.}$ interact with peptides of such variety in lengths? It appears that $OppA_{L.i.}$ is able to accommodate the peptides with various lengths because only the backbone C=O and N-H groups of the peptides interact with $OppA_{L.i.}$ via hydrogen bonding, independent to structures of side chains and sequences of the peptides [29]. The crystal structure study of peptide binding of OppA from *Salmonella typhimurium* demonstrated that only the backbone amino groups of peptide substrates interact with the protein [83], providing an explanation for the diverse substrate specificity of $OppA_{L.i.}$. It appears that there is no direct interaction between side chains and SBP such that the side chains are laid in distinct spacious pockets of the protein.

**Mechanisms of substrate uptake via ABC importer**

Transport of substrate via ABC importer is a highly concerted mechanism as depicted in the maltose transport. There are 3 main steps involved in the transport: (i) interaction of substrate-bound SBP with TMDs; (ii) hydrolysis of ATP triggered by the interaction; and (iii), release of the substrate into cytoplasmic space by conformational change. Figure 1.6 illustrates the step-wise maltose transport mechanism by *E. coli* maltose transporter. Binding of a substrate causes MalE to close its substrate binding cleft. Only this closed MalE structure is recognized by MalF/G/K complex. The Q loop of MalK, ATPase interacts with the cytoplasmic loop of the MalF/G TMD complex forming a tight transport complex. This form of complex is called pathway (P)-closed because the periplasmic site of the transporter complex is not opened. The binding of a liganded MalE to the MalF/G/K complex causes a conformational change of MalK. The presence or

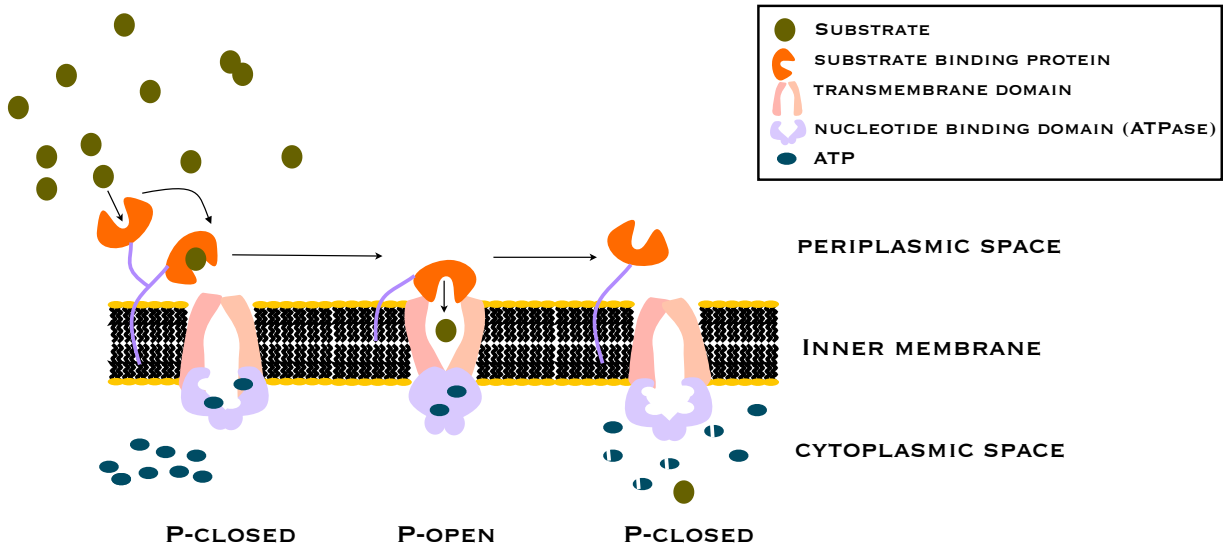



Figure 1.6: A model mechanism of substrate transport via ABC importer. Adapted from [25] though there is no lipid anchor in the Maltose transport example shown in the reference. The anchor was added as we are explaining about ABC importers in *Mycobacterium spp*.

absence of ATP in MalK influences the conformational change as ATP-binding and hydrolysis rearrange the forms in the complex. Once hydrolysis of ATP occurs, MalK is no longer able to hold the closed conformation, going back to its P-closed position, which releases the substrate into the cells. How conformational change occurs in the transmembrane domain, MalF/G needs to be determined, though it is believed that the conformational changes of the peripherally associated proteins, MalE and MalK dimers, stimulates a conformational change of MalF/G, allowing a transfer of substrate from MalE to the substrate binding pocket of MalF/G.

**Known functions of ABC importers**

A diverse array of nutrients are taken up by ABC importers: sugar, amino acids, iron, nickels, phosphates, and much more [25]. Pathogenic bacteria particularly benefit by acquisition of nutrients via the high-affinity substrate uptake system because available nutrients in the host are very limited and competition is high to scavenge necessary nutrients. ABC importers can also be used as a osmosensor [101]. When bacteria face a harmful shift in osmotic potential, i.e., hypertonic condition, compatible solutes such as glycine betaine are take up via ABC importer to counterbalance the hyperosmotic pressure [18, 95, 85]. Also, signaling molecules such as Phr peptides of *Bacillus subtilis* species and pheromones of *Enterococcus faecalis* enter to the cells via ABC importers [52]. Phr peptides regulate sporulation of *B. subtilis* and the *E. faecalis* pheromones regulate mating process with other strains of *E. faecalis*. Both mechanisms are regulated by the changes in environments and are crucial for survival of the bacteria.

### 1.2.2 Ph.D. projects

A considerable amount of efforts has been put into understanding of cellular metabolism and responses to host immune systems in *Mtb*. However, understanding of mechanisms in nutrient acquisition is very limited and is still at its infancy. Recent studies on Opp and trehalose transporters of *Mtb* reveal not only the importance of ABC importers in nutrient acquisition but their importance in the pathogenicity of *Mtb* in the host. There are many proteins annotated as transporter proteins in *Mtb* (TransportDB) but the functions of most of the proteins remain elusive. Understanding substrates of the transporters in *Mtb*

provides us a tool to design antimycobacterial drugs which can be efficiently transported into the bacterial cells via known mechanism. In my Ph.D. dissertation, we present a screen that we developed to identify transporter mutants in *M. smegmatis*, a non-pathogenic relative of *Mtb*, and the characterization of the transporter mutants with respect to their substrate specificity and biological activity.

CHAPTER 2

# DEVELOPING A SCREENING TOOL TO IDENTIFY UPTAKE TRANSPORT DEFICIENT MUTANTS IN *MYCOBACTERIUM SMEGMATIS* $MC^2$ 155

## 2.1 Abstract

Genes responsible for uptake of fluorescein di-$\beta$-D-galactopyranoside (FDG), a synthetic fluorogenic substrate, were screened using a transposon mutagenized library of recombinant *Mycobacterium smegmatis* $MC^2$ 155, *Bluesmeg*. *Bluesmeg* possesses a plasmid, pML803, which encodes *lacZ* as a reporter gene. We exploited this reporter system and various fluorogenic *lacZ* substrates to identify mutants with defective uptake transport mechanism. FDG is a fluorogenic substrate of $\beta$-galactosidase that emits fluorescent signal only when $\beta$-D-galactopyranoside is cleaved from the substrate, fully exposing the fluorescein moiety. Screening of the transposon library in combination with FDG identified four uptake/retention mutants (UMs) designated UM10, UM80, UM83, and UM130. With exception of UM10, all the UM strains belong to oligo/dipeptide or sugar ABC transport family, all of which are unique members of the ABC importer family. By this method, we were able to develop a screening tool which successfully identified transporters for a target substrate.

## 2.2 Introduction

A challenging issue to study mechanisms of nutrient acquisition is to first determine whether there are any specific transporters involved in the acquisition. According to the work performed by Titgemeyer *et al*. [92], *M. smegmatis* and *Mtb* encode carbohydrate transporters predicted by the homology search against other bacteria whose carbohydrate transporters were biochemically characterized. 28 putative carbohydrate transporters are predicted in *M. smegmatis* and 5 similar transporters in *Mtb*. Among the 28 transporters in *M. smegmatis*, 19 transporters belong to ATP-binding cassette (ABC) transporter family. The same study also shows redundancy in substrate specificity in some transporters: xylose is transported via 3 distinctive transporters and ribose is transported via 5 distinctive transporters. TransportDB (`http://www.membranetransport.org/`) is a relational database which provides a list of the predicted cytoplasmic membrane transport proteins, each categorized in a specific family, from organisms whose genome sequences were completed and available. Up to date, the cytoplasmic membrane transport proteins from 365 organisms from eukaryotes to prokaryotes have been compiled in the database. 423 cytoplasmic membrane transport proteins are predicted in *M. smegmatis*, and 148 such proteins in *Mtb*. Interestingly, many of the transport proteins are predicted to play a role in acquisition of nutrients from macromolecules, i.e., sugar and peptides, to small molecules, i.e., metals and ions, and many of them belong to the ABC transport family. Like the transporters predicted to play a role in carbohydrate uptake [92], the functions of most of the transport proteins in the TransportDB database have not been experimentally verified. However, both studies present that active transporters play a role in uptake of nutrients and some transporters have

substrate redundancy.

Recently, cholesterol catabolic locus in *Rhodococcus* sp. strain RHA1 was discovered [94]. *Rhodococcus* sp. strain RHA1 is an environmental species relative of mycobacteria. In the study, the authors showed that both the *mce4ABCDEF* gene cluster and *supAB* genes in RHA1 are important for uptake of cholesterol because mutants lacking either *mce4* or *supAB* were not able to grow on the substrate. This was the first study which identified the core genes involved in cholesterol uptake, and it was a very exciting finding for the field of tuberculosis (TB) because *Mtb* shares synteny with the cholesterol catabolic locus of RHA1. It has become possible to manipulate the genes involved in cholesterol catabolism in *Mtb* to study the importance of cholesterol during infections. Several groups have reported that *Mtb* mutants lacking genes for cholesterol metabolic pathway were highly attenuated during chronic infection in mice and in IFN-$\gamma$-activated macrophages in vitro [64, 56, 14]. This information provides us an insight to physiologically important nutrients for *Mtb* in host cells and within the lipid-rich environment of the granuloma [20, 59]. The host environment is dynamic and there is no guarantee that a carbon source available in one host environment is also available in other host environments. Human lung granulomas in TB patients develops different pathohistological characteristics depending on the stage of progression [48], suggesting that nutrients highly vary during the life cycle.

Transposon site hybridization (TraSH) analysis has indicated that genes annotated as putative disaccharide transporter were essential for *Mtb* to survive during the first week of infection in mice [78]. The authors believed that metabolic shift of main carbon source have occurred from carbohydrates to

lipids and showed the importance of carbon sources other than lipids in the early stage of infection. Mycobacterial metabolism on various nutrient sources has been extensively studied [2, 9, 26, 44, 54]. However, nutrient transporter systems have rarely been documented. This was not due to a lack of interest in nutrient transport mechanisms, but rather a lack of appropriate substrates to study them.

We have constructed a *lacZ* based transposon mutant library in *M. smegmatis*, a non-pathogenic species of mycobacteria. We decided to use *M. smegmatis* instead of *Mtb* because of its faster growth rate and the ease of handling. Fluorescein di-$\beta$-D-galactopyranoside (FDG) is a fluorogenic substrate, in which fluorescein moiety is conjugated with two $\beta$-D-galactopyranosides. The sugar moieties need to be cleaved off by $\beta$-galactosidase to reveal fluorescence. $\beta$-galactosidase is a cytosolic protein, thus the cleavage action should only be seen within the bacteria as fluorescence. In this screen, we were able to identify four transporters which are responsible for FDG uptake in *M. smegmatis*. Interestingly, three of them are ABC importers, a family of ABC transporter which possesses an additional high affinity substrate binding protein. Here, we demonstrated that an active transport system is involved in the FDG uptake.

## 2.3 Materials and Methods

### 2.3.1 Construction of a $\beta$-Galactosidase Reporter Strain of *Mycobacterium smegmatis* MC$^2$ 155

To use $\beta$-galactosidase activity as a bioprobe to detect fluorescein di-$\beta$-D-galactopyranoside (FDG) (Anaspec, San Jose, CA, USA) uptake in *M.smegmatis*, the *lacZ* gene from *E.coli* was stably incorporated into the cell on the replicating plasmid, pML803, a generous gift from Dr. Michael Niederweis at University of Alabama. This plasmid expressed *lacZ* under regulation of psmyc promoter and the transcriptional terminator sequence, $tt_{T4g32}$, at the end of *lacZ* to prevent leaky expression [43]. In brief, pML803 was electroporated into *M.smegmatis* and the transformant was plated on Middlebrook 7H11 agar supplemented with Middlebrook OADC enrichment (Becton Dickinson, NJ) in presence of 50 $\mu$g/ml of hygromycin. Miller assay [55] was performed to confirm $\beta$-galactosidase activity of the recombinant *M.smegmatis* with slight modification. After measuring $OD_{600}$, 50 $\mu$l of the culture was mixed with 950 $\mu$l substrate solution (0.06 M $Na_2HPO_4{\cdot}7H_2O$, 0.04 M $NaH_2PO_4{\cdot}H_2O$, 0.01 M KCl, 0.001 M $MgSO_4$, and 0.05 M $\beta$-mercaptoethanol), then 10 $\mu$l each of chloroform and 2% SDS was added to the mixture to lyse the cell. The lysate was incubated at 37°C for 30 min followed by addition of 50 $\mu$l of ONPG (4 mg/ml) to start the reaction. 50 $\mu$l of 1M $Na_2CO_3$ was added to stop the reaction. The successful recombinant strain of *M. smegmatis* with pML803 was designated as *Bluesmeg* and was maintained on a solid plate containing 50 $\mu$g/ml of hygromycin and 2% X-Gal.

### 2.3.2 Measurement of Fluorescence in *Bluesmeg*

Flow cytometry was employed to detect the level of fluorescence in *Bluesmeg*. A single colony was picked from the plate and inoculated into 2 ml 7H9/OADC/glycerol/0.05% Tween-80 (7H9 media) and incubated at 37°C on a rotary shaker at 100 rpm overnight. 1 ml of the culture was taken and washed once with (vol/vol) PBS/0.05% Tween-80. The culture was resuspended in 10 ml 7H9/OADC/glycerol/0.05% Tween-80 containing 15 $\mu$M fluorogenic $\beta$-galactosidase substrate, either fluorescein di-$\beta$-D-galactopyranoside (FDG) or 5-dodecanoylaminofluorescein di-$\beta$-D-galactopyranoside ($C_{12}$FDG). Aliquot was taken and washed in (vol/vol) PBS/0.05% Tween-80 and resuspended in the same buffer. The suspension was dispersed by passage through a 25G 5/8 needle and filtered through 5 $\mu$M sterile filter (Minisart, Satroius Stedim Biotech, GmbH, Germany) before being applied to flow cytometry (FACSCalibur, BD Biosciences, San Jose, CA, USA). Fluorescence was detected using a FITC (530 ± 15 nm) filter.

### 2.3.3 Determining the Uptake Pathway(s) of FDG

To determine whether active or passive transport is involved in FDG uptake, $NaN_3$ was added to a *Bluesmeg* culture in the presence of 15 $\mu$M FDG. Briefly, *Bluesmeg* was added to the 7H9 medium containing 15 $\mu$M FDG and either 10 $\mu$M or 100 $\mu$M $NaN_3$ and grown at 37°C at 100 rpm. Bacterial cells were harvested at 24 h, washed with (vol/vol) PBS/0.05% Tween-80 and resuspended in the same buffer. Intensity of fluorescence was measured at FL1-H channel in flow cytometry. The fluorescence from the cultures in presence or absence was

compared and used to determine the uptake pathway(s) of FDG in *Bluesmeg*.

To make sure that we were not incorporating surface associated fluorescence, methyl red (MR) was added to quench dyes accessible as extracellular signals. In brief, the culture was grown in the 7H9 media with 15 $\mu$M FDG and aliquot was processed to read in flow cytometry. Just after the first reading, MR was added to the cell suspension and the fluorescence was measured again in flow cytometry. Fluorescent intensity with and without MR was compared and the difference was used to determine the level of extracellular signal.

### 2.3.4 Generation of a Transposon Mutant Library of *Bluesmeg*

A transposon mutant library was generated in the *Bluesmeg* background using the phagemid $\phi$MycoMarT7, which was kindly donated by Dr. Eric Rubin at Harvard Medical School. This phagemid was based on $\phi$AE87, a temperature sensitive mutant of Mycobacteriophage TM4 containing double *cos* sites derived from pYUB328 [4, 6, 50, 77]. $\phi$MycoMarT7 contained a transposon vector MycoMarT7 (GenBank AF411123). The MycoMarT7 vector possessed a kanamycin resistance gene and OriR6K allowing replication in *pir+ E. coli*. This vector also contained a T7 promoter sequence at 5 and 3 ends directed outwards. *Himar1* is a *marinar* family transposase, isolated from *Haematobia irritans*, a horn fly. Transposons in this family insert a gene fragment between inverted terminal repeats (ITRs) into any thymine-adenine (TA) nucleotide site on the chromosome. $\phi$MycoMarT7 contained a hyperactive mutant of the *Himar1* transposase gene, *C9 Himar1*, whose gene product increased transpositional activity and randomizes transposition. Transduction was performed as follows: 50 ml of *Bluesmeg*

culture was grown to saturation in the 7H9 media, washed twice with vol/vol MP buffer (50 mM Tris, pH 7.5/150 mM NaCl/10 mM $MgSO_4$/2 mM $CaCl_2$), and resuspended in 3.15 ml MP buffer. Bacterial cells were passaged through a 25G 5/8 needle and infected with 1.85 ml of $\phi$MycoMarT7 at concentration 5.45 $X10^{10}$. The mixture was incubated at 30°C for 10 min followed by further incubation at 37°C for 4 h. Transduced cells were plated onto 7H10 solid agar containing 50 $\mu$g/ml each of hygromycin and kanamycin, and incubated at 37°C for 3 days. All the colonies were scraped off from plates and resuspended in PBS containing 30% glycerol. The pools of transposon mutant library were kept at -80°C until use.

### 2.3.5 Measuring Loss of Fluorescence from the Bacterial Cells

Prior to sorting, loss of fluorescent from the cells after removal of FDG was measured. This was particularly important because the transposon mutants grown with FDG will be washed and resuspended in a fresh media without FDG prior to the screening process in flow cytometry. Measurement of the natural loss of fluorescent signal from the bacterial cells would define the time allows us to process bacterium with minimal introduction of false-positive transport mutants. Briefly, the *Bluesmeg* culture was grown with FDG until the fluorescent signal reached equilibrium. The bacterial cells were then washed once with (vol/vol) PBS/0.05% Tween-80. The cell was resuspended in the same medium without FDG and aliquot was tested for the retention of fluorescence in the cell.

### 2.3.6 Sorting Transposon Mutants Lacking Fluorescence in Flow Cytometry

Pools of transposon mutant library were grown in the 7H9 media containing 15 $\mu$M FDG at 37°C with agitation at 100 rpm overnight. The pools were washed in PBS/0.05% Tween-80 and were passaged through a 25G 5/8 needle. The cell suspension was filtered through 5 $\mu$M sterile filter (Millipore, MA) to remove clumps of cells. The pool was briefly run through the flow cytometry to determine the distribution of non-fluorescent and fluorescent cells and was immediately subjected to sorting in flow cytometry. A sorting gate was selected in the area where minimal fluorescence was detected from the cells. The sorted cells were collected in PBS and filtered through 0.22 $\mu$m sterile nitrocellulose membrane (Millipore, MA) in a vacuum filter system (Millipore, MA). One ml of PBS/0.05% Tween-80 was added onto the membrane to recover as many cells from the membrane. The cell suspension was plated on 150 x 15 mm 7H11/OADC/glycerol agar plate containing 50 $\mu$g/ml each of hygromycin and kanamycin to check purity of the sorted cells. Colonies were isolated from the plate and subjected to two more rounds of enrichment. In short, they were grown in the 7H9 media containing 15 $\mu$M FDG and the sorting process in flow cytometry was repeated. The sorting process is briefly summarized in Fig. 2.1.

### 2.3.7 Screening of the Mutants from the FACS Selection

Individual colonies from the FACS selection were screened for FDG uptake deficiency in 96-well plates. Colonies were picked and inoculated in 100 $\mu$l the 7H9 media containing either FDG or X-Gal. The plates were incubated at 37°C

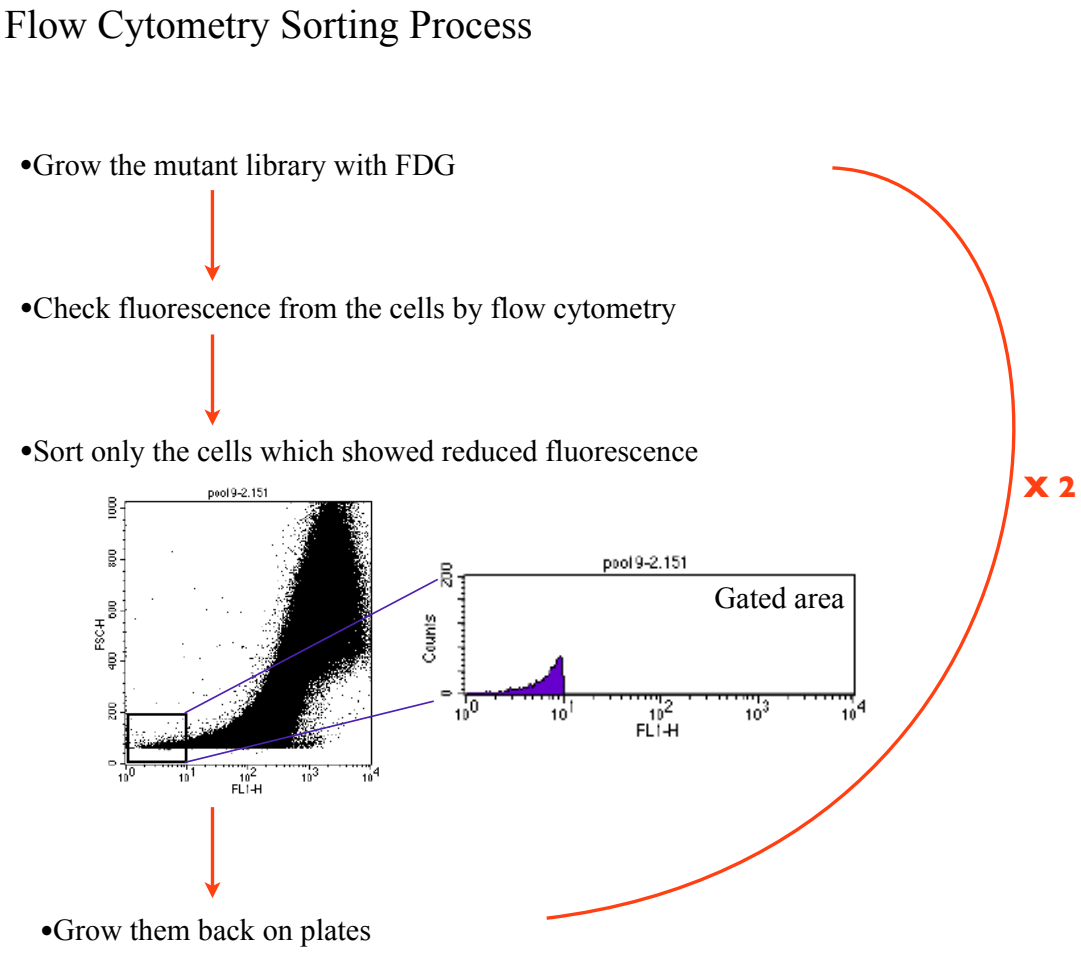


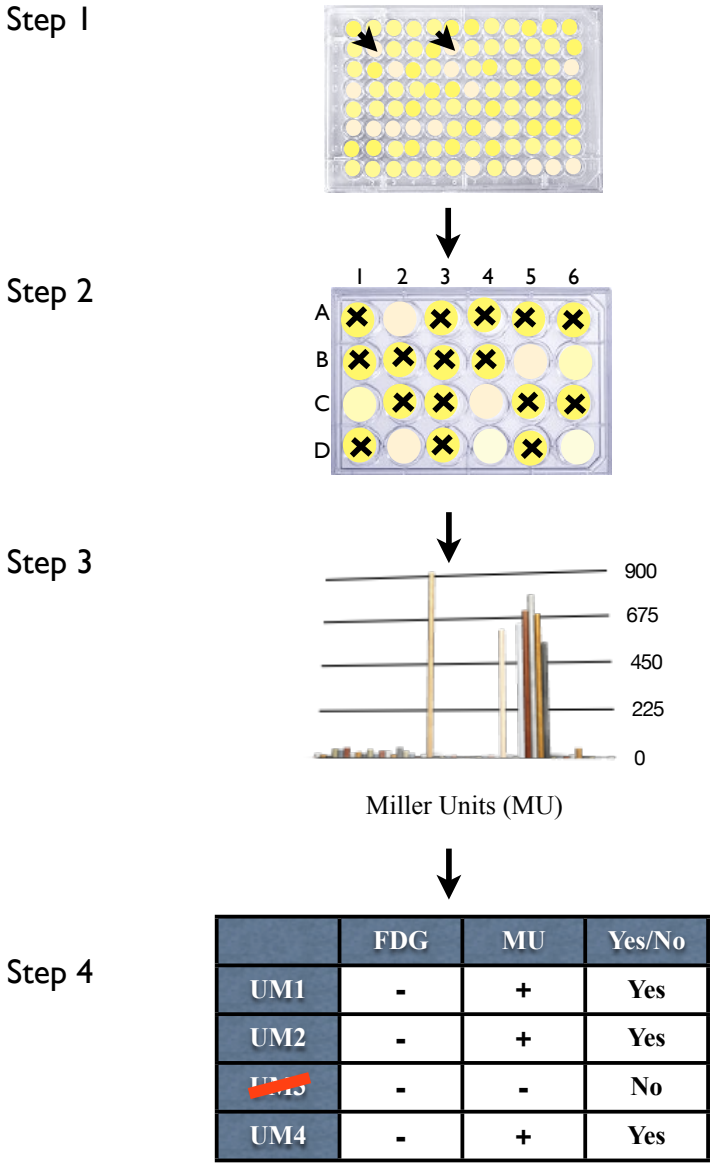

Figure 2.1: Flowchart of the screening process to identify FDG uptake deficient transposon mutants from pools of transposon library of *Bluesmeg* using flow cytometry. Mutants in the gated area were sorted and further identified.

for 24 h. Bacterial cells with weak fluorescence in the FDG plate but strong blue pigmentation in the corresponding X-Gal plate were selected and the same screening process was repeated in 1 ml media in 24-well plates. Bacterial cells that showed minimal fluorescence with strong $\beta$-galactosidase activity by X-Gal cleavage were further tested their $\beta$-galactosidase activity by Miller assay to ensure their FDG uptake deficient phenotype. Uptake mutants were all designated

with UM (uptake mutant) suffix (Fig. 2.2).

### 2.3.8 Identification of Mutants Possibly Possessing FDG Uptake Defective Phenotype

Genomic DNA was extracted from isolated mutants. 600 $\mu$l of culture (in an 1.5 ml microcentrifuge tube) was centrifuged at 3,000 x g at RT for 5 min and washed with 1 ml 10% glycerol. 500 $\mu$l of lysis buffer (3% SDS, 1 mM $CaCl_2$, 10 mM Tris-HCl, pH 8.0, and 100 mM NaCl) was added to the tube and briefly vortexed. The cell suspension was sonicated for 10 sec twice with an 15 sec interval on ice. The lysate was incubated at 56°C for 30 min followed by 1 h incubation at RT with 40 $\mu$g of Proteinase K added. After the incubation, (vol/vol) phenol/chloroform/iso-amyl alcohol (25:24:1) was added and vortexed briefly. The tube was centrifuged at 16,000 x g at RT for 10 min to separate phases. The aqueous phase (upper phase) was transferred to a new microcentrifuge tube and subjected the repeat process. The aqueous phase was transferred to a fresh microcentrifuge tube containing 1/3 vol 3 M NaOAc and 0.8 vol ice cold isopropanol. Genomic DNA was precipitated at -20°C overnight. After centrifugation at 16,000 x g at RT for 10 min, the genomic DNA was washed with 1 ml 70% ice cold ethanol, vacuum dried, and dissolved in sterile deionized water. Ligated mediated PCR (LM-PCR) was performed as described by Prod'hom *et al*. [68] to identify the transposon insertion region. Genomic DNA digested with *Sal*I (New England Biolabs, MA) was ligated with LM-PCR linker mixture (*Sal*gd and *Sal*pt, Table 2.1) using Fast-Link DNA Ligase kit (Epicentre, WI). 5 $\mu$l of *Sal*I digested genomic DNA was added to a mixture containing 3.6 $\mu$M *Sal*I




| | FDG | MU | Yes/No |
|---|---|---|---|
| UM1 | - | + | Yes |
| UM2 | - | + | Yes |
| [redacted] | - | - | No |
| UM4 | - | + | Yes |

Figure 2.2: Flowchart of the identification process to select uptake mutant (UM) after the flow cytometry screening process. In step 1, colonies from the sorting process were transferred to 96-well plates containing 7H9 medium and FDG. Only the mutants with reduced fluorescence were transferred (mutants in the arrowed wells) to 24-well plates. In step 2, the same process described in the step 1 was repeated with a larger volume (1 ml medium instead of 100 $\mu$l. In step 3, the mutants with reduced fluorescence were checked for their $\beta$-galactosidase activity. In step 4, only the mutants with both reduced fluorescence with FDG and and strong Miller Units were selected and designated with uptake mutant (UM) number.

| Primers | Sequences |
|---|---|
| *Sal*gd | 5′-TAG CTT ATT CCT CAA GGC ACG AGC-3′ |
| *Sal*pt | 5′-TCG AGC TCG TGC-3′ |
| Primer 1 | 5′-CTA GAG ACC GGG GAC TTA TCA GCC A-3′ |
| Primer 2 | 5′-TAG CTT ATT CCT CAA GGC ACGG AGC-3′ |
| Primer 3 | 5′-CCC GAA AAG TGC CAC CTA AAT TGT AAG CG-3′ |
| Primer 4 | 5′-CGC TTC CTC GTG CTT TAC GGT ATC G-3′ |

Table 2.1: Sequences of primers used to construct the *Sal*I linker and LM-PCR.

linker, 1.5 $\mu$l each of 10X FastLink ligase buffer and 10 mM ATP, 1 $\mu$l ligase, and water up to 15 $\mu$l. Ligation was performed at 16°C for 2 h followed by inactivation at 65°C for 20 min. An aliquot of the ligated DNA sample was used as a template for PCR. Combination of primers (Table 2.1) was optimized for each sample to get best amplification. PCR was performed using HotStart Plus (Qiagen, GambH) with the following reaction cycles: an initial denaturation step at 95°C for 5 min, followed by 30 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 1.5 min. Final extension was done at 72°C for 10min. PCR amplicons were visualized in agarose gel and distinct bands were excised from the gel and purified using QIAQuick Gel Extraction Kit (Qiagen, GambH) for sequencing.

## 2.4 Results

### 2.4.1 Greater Fluorescence Signal was Detected in the *Bluesmeg* Culture with FDG

The efficiency of substrate processivity of $\beta$-galactosidase expressed in *Bluesmeg* was determined. At first, Miller assay was performed on *Bluesmeg* to measure

the level of $\beta$-galactosidase activity. Figure 2.3A shows a high level of Miller units produced in *Bluesmeg* compared to WT *M. smegmatis*. The $\beta$-galactosidase activity of *Bluesmeg* was easily visualized on X-Gal plate as blue colony (Fig. 2.3B). FDG and $C_{12}$FDG were tested as a $\beta$-galactosidase substrate in *Bluesmeg*. Two different substrates were tested because we wanted to compare which substrate would yield better uptake and retention of fluorescence inside the cell. Flow cytometry data showed that the fluorescent signal was easily detected from FDG after 1 h and increased with time (Fig. 2.3C). With $C_{12}$FDG, hardly any increase of fluorescence was observed in the first 5 h. The shift of fluorescence was eventually observed at 24 h, though it was significantly lower than that from FDG (Fig. 2.3D). WT *M. smegmatis* showed no fluorescence throughout time, indicating that background signal was minimal. This result was consistent with previously reported data [67]. Plovins, A. *et al*. showed that the addition of the $C_{12}$ acyl chain reduced penetration of FDG into viable bacteria as well as yeast due to the increased hydrophobicity. Only non-viable cells showed fluorescence as indicated by co-staining with propidium iodide, which only stained cells that had lost their membrane integrity.

Fluorescence in the cells grown in presence of FDG was measured up to 72 h to determine whether fluorescence increased after 24 h. Longer incubation time did not increase fluorescence. The fluorescence was similar to that at 24 h (data not shown), suggesting that uptake and efflux met equilibrium at 24 h.

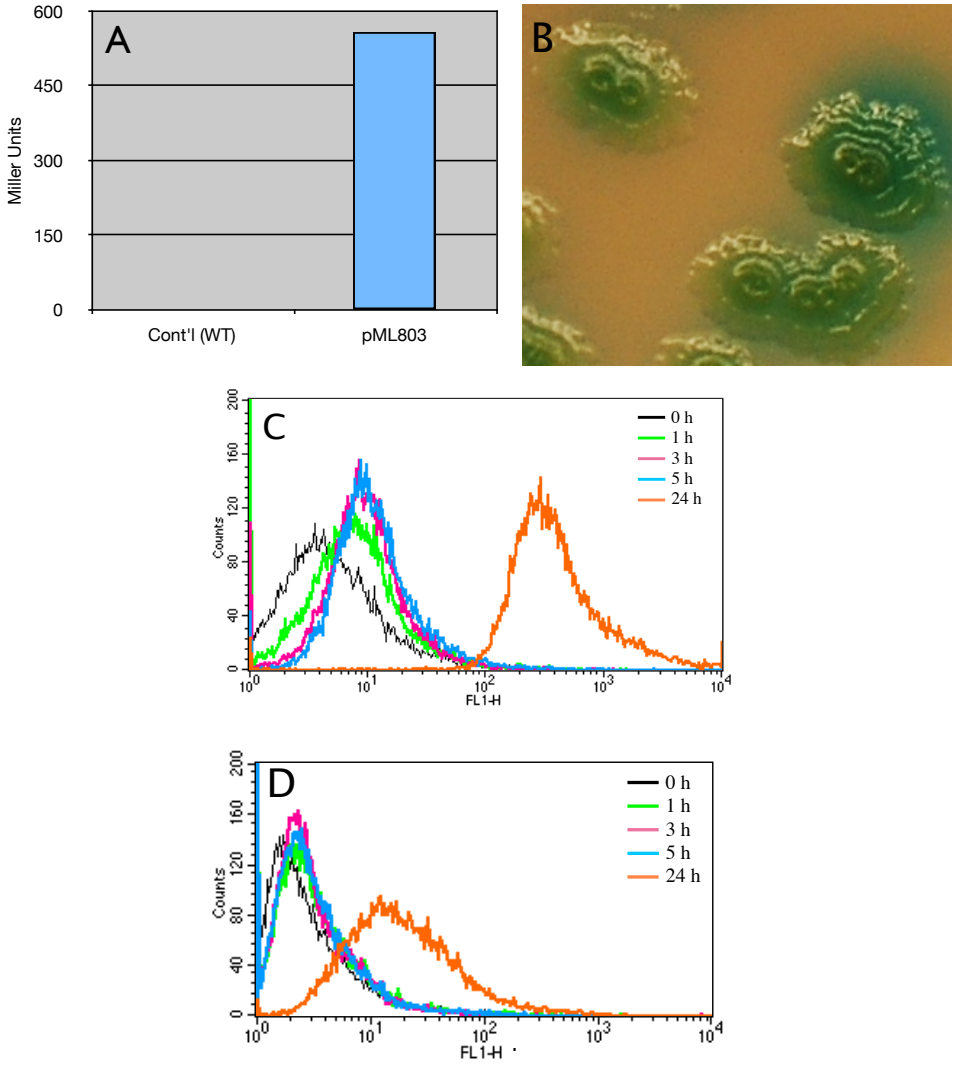



Figure 2.3: (A) Miller assay on WT *M. smegmatis* and *Bluesmeg*. Miller units measured production of *o*-nitrophenol released from ONPG by $\beta$-galactosidase activity at 420 nm. As shown, there was no reaction seen in WT but a strong signal in *Bluesmeg*. (B) Appearance of *Bluesmeg* on agar plate containing X-Gal. Progression of fluorescent signal in *Bluesmeg* was measured over time in presence of FDG (C) or $C_{12}$FDG (D). No signal was detected at 0 h and the signal did not change in the early stage (between 1 - 5 h). At least one log of difference in signal was observed from the cells in presence of FDG and $C_{12}$FDG at 24 h.

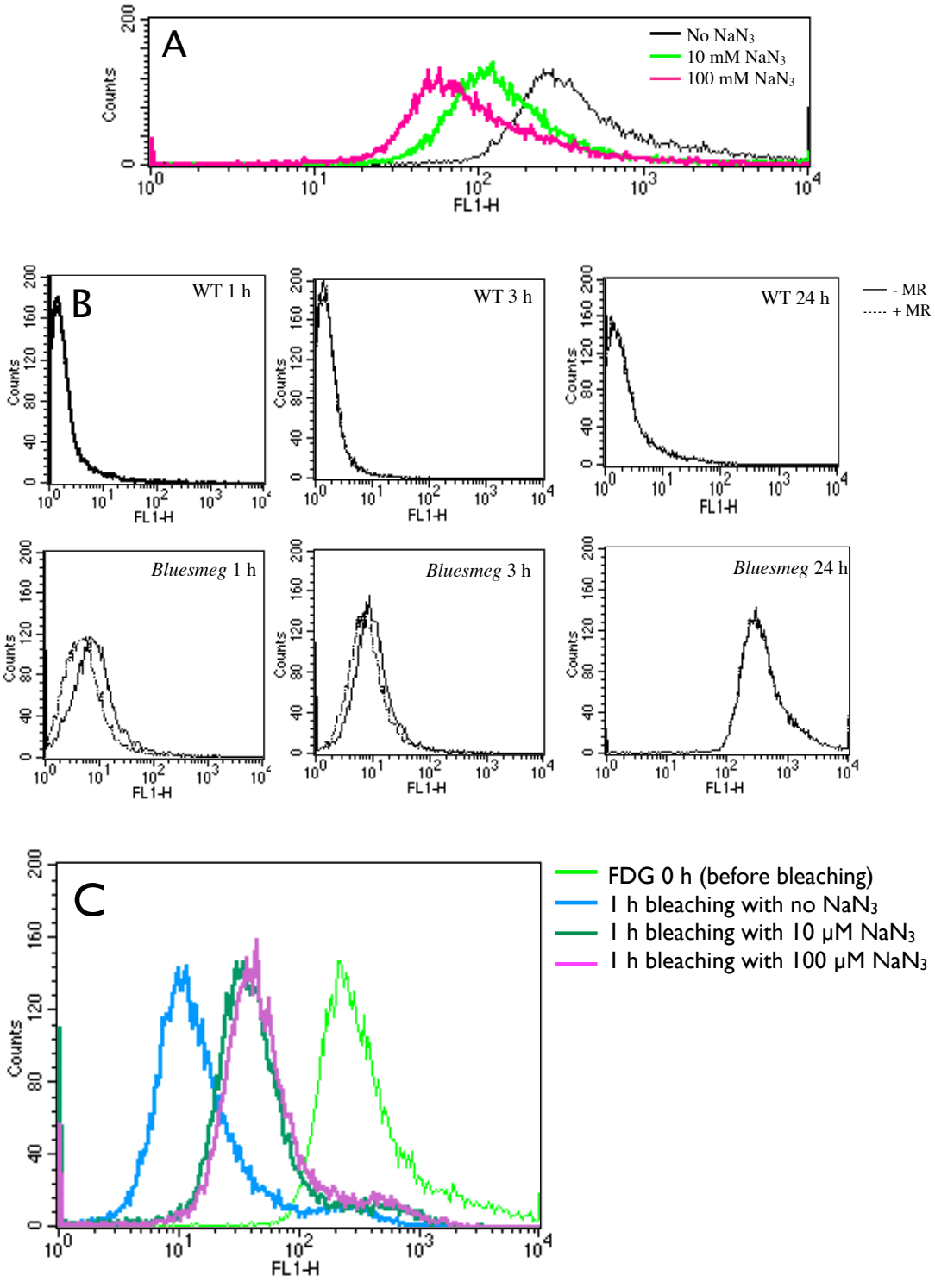

A
No NaN3
10 mM NaN3
100 mM NaN3
Counts
FL1-H
B
WT 1 h
WT 3 h
WT 24 h
- MR
+ MR
Bluesmeg 1 h
Bluesmeg 3 h
Bluesmeg 24 h
C
FDG 0 h (before bleaching)
1 h bleaching with no NaN3
1 h bleaching with 10 µM NaN3
1 h bleaching with 100 µM NaN3

Figure 2.4: (A) $NaN_3$ was added to cultures containing FDG to test whether active or passive transport system was involved in FDG uptake. The result was collected from the same number of cells analyzed in flow cytometry. Less fluorescent signal was detected from the cultures containing 10 $\mu$M $NaN_3$ (green) and 100 $\mu$M $NaN_3$ (pink). (B) Methyl Red (MR) was added to the cell suspension immediately after the first analysis in flow cytometry, and fluorescent signal was measured. In the first 3 h, fluorescent signal was not quite stable. Some signals were derived outside the cells. At 24 h, no difference in signal was detected with or without MR indicating the signal was stabilized and all the signal measured came from the cells. Solid line indicates a cell suspension without MR and dotted line indicates a cell suspension with MR. (C) Bleaching of fluorescent signal was measured after incubating the cells in a medium containing FDG for 24 h. The 0 h time point (green) indicated fluorescent signal before the cells were washed and transferred into a new medium without FDG. After one hour, decrease in fluorescent signal was greatest without $NaN_3$ (blue), but less so with $NaN_3$ (moss green and purple).

### 2.4.2 Both Facilitated and Active Uptake System are Involved in FDG Uptake

To test whether facilitated or passive transport system was involved in uptake of FDG, $NaN_3$ was added to a *Bluesmeg* culture. $NaN_3$ inhibits biological processes which require ATP. By adding $NaN_3$ in a culture containing FDG, we can use the loss of fluorescence as a measure of FDG uptake via active transport. At 24 h, fluorescence from the cell grown in presence of $NaN_3$ was a half to one log less than those in absence of $NaN_3$ (Fig. 2.4A). Bacterial cells grown in 100 $\mu$M $NaN_3$ had a half-log less retention of fluorescence than those in 10 $\mu$M $NaN_3$ indicating that there were still some growth at 10 $\mu$M $NaN_3$. The fluorescence detected at 100 $\mu$M $NaN_3$ indicates the processivity of $\beta$-galactosidase already present at the time of assay. The signal difference between the cells grown at 0 $\mu$M $NaN_3$ and

100 $\mu$M $NaN_3$ might come from FDG entering the cells by facilitated transport and the signal from the cells grown at 100 $\mu$M $NaN_3$ might come solely from FDG entering the cells by passive transport.

Methyl red (MR) was used to quench the extracellular signal so that only internalized substrate was detected in flow cytometry. The fluorescence was less in presence of MR than the total fluorescence at early time point, but the difference gradually diminished over time. At 24h, there was no loss of signal with MR (Fig. 2.4B). This result, combined with the result from the FDG processivity assay, indicates that the fluorescence reached equilibrium at 24 h. This time point was therefore used in all further experiments.

When the bacterial cells, grown in presence of FDG for 24 h, were transferred to a new medium without FDG, rapid loss of fluorescence was detected within the first hour of the transfer (Fig. 2.4C). This rapid loss of signal from the cells indicates that fluorescein was secreted into solution but we did not know whether it was due to active or/and passive efflux transport systems. To determine the secretion system(s), $NaN_3$ was added to the cell suspensions at 0 h, when the bacterial cells were transferred to a fresh medium without FDG to look at the rate of loss by diffusion or active transport: the presence of $NaN_3$ slowed down loss of fluorescence from the cells, but did not completely block the loss of signal. Based on this result, it appears that both active and passive export mechanisms were involved the loss of fluorescence.

### 2.4.3 Identification of Transposon Mutants which Showed Involvement of Oligo-/dipeptide Transporter in FDG Uptake

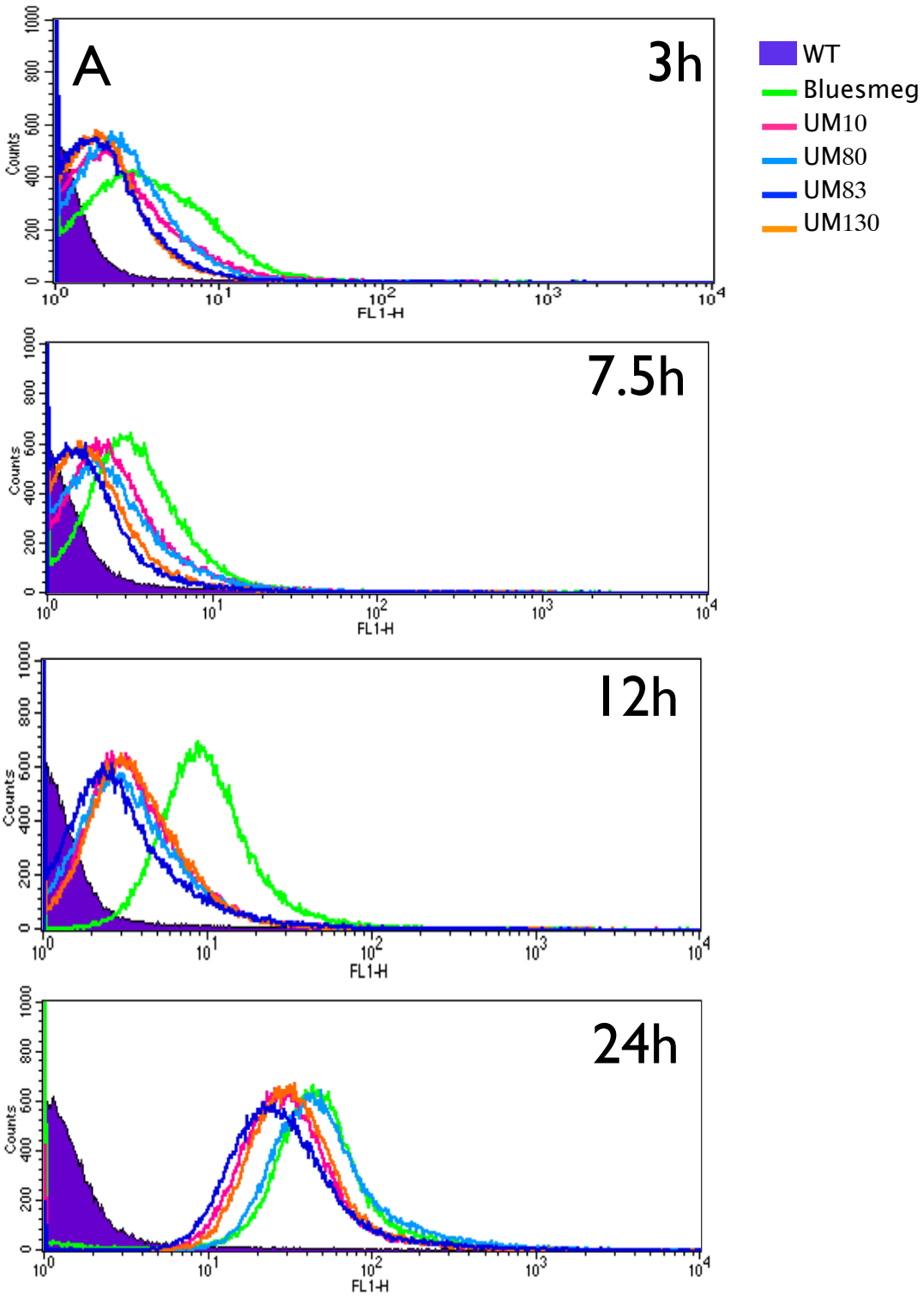
A library of over 30,000 transposon mutants was screened for uptake transport mutants using FDG as a bioprobe in flow cytometry. 8104 transposon mutants were collected from the final enrichment process on solid agar plates and about 2280 mutants were picked and transferred to 96-well plates to test for their FDG uptake deficient phenotype. Those mutants were evaluated for their levels of fluorescence compared to the fluorescence in the *Bluesmeg* control. 521 mutants showed reduced fluorescence and were designated UM for uptake mutant. These UM mutants were transferred to 24-well plates to repeat the screening process. In the first round of the screening process, only lack of fluorescence in the wells was considered. Those UM mutants with lack of fluorescence were again transferred to two 24-well plates for the second round of screening. In the second round, FDG or X-Gal was added as a $\beta$-galactosidase substrate to the plates. Colour development on the corresponding X-Gal plate supports the deficient FDG uptake phenotype of the UM mutants by confirming that $\beta$-galactosidase activity was maintained in these mutants. This screening process was performed to validate the phenotype of the FDG uptake deficient mutants. We speculated that some of the UM mutants showed the FDG uptake deficient phenotype due to insertion of transposon in *lacZ* on pML803 or mutation of the plasmid. The UM mutants lacking both fluorescence from FDG and the strong blue colour from X-Gal were eliminated as false-positive transport deficient mutants. Miller assay on those UM mutants was performed to confirm the FDG trasnport deficient phenotype. 98 UM mutants retained $\beta$-galactosidase activity

| UM # | MSMEG ID | Possible homolog in *Mtb* | Function |
|---|---|---|---|
| 10 | 4363 | Rv1687c,<br>ABC drug resistant family,<br>ATP binding protein | hypothetical protein |
| 80 | 0556 | Rv1238,<br>SugC | ABC transporter<br>ATP binding protein |
| 83 | 4995 | Rv1283c,<br>OppB | oligopeptide transporter system<br>permease, AppB |
| 130 | 4099 | Rv3664c,<br>DppC | ABC transporter<br>permease |

Table 2.2: Identification of the transporter genes disrupted by transposition with its annotated function and possible homolog in *Mtb* are also listed.

as strong as *Bluesmeg* and were characterized further.

Variable template concentration was tested in LM-PCR to optimize the amplification conditions. Out of the 98 UM mutants, 69 UM mutants were identified by sequencing. In those sequenced, 4 UM mutants had a transposon insertion in or near transporter genes (Table 2.2). *MSMEG_0556* (UM80) belongs to a sugar ABC transport superfamily, and *MSMEG_4995* (UM83) and *MSMEG_4099* (UM130) belong to the putative oligo/dipeptide ABC transport superfamily. *MSMEG_4363* (UM10), however, does not belong to any of them and is annotated as a hypothetical protein. Instead, it has the nuclear factor transport 2 (NTF2) superfamily domain. Although *MSMEG_4363* does not belong to the ABC transport family, it was selected to study for further characterization because the NTF2 proteins in eukaryotic cells are known to mitigate a transport of Ran into nucleus [69, 102].


A
3h
7.5h
12h
24h
Counts
FL1-H
WT
Bluesmeg
UM10
UM80
UM83
UM130

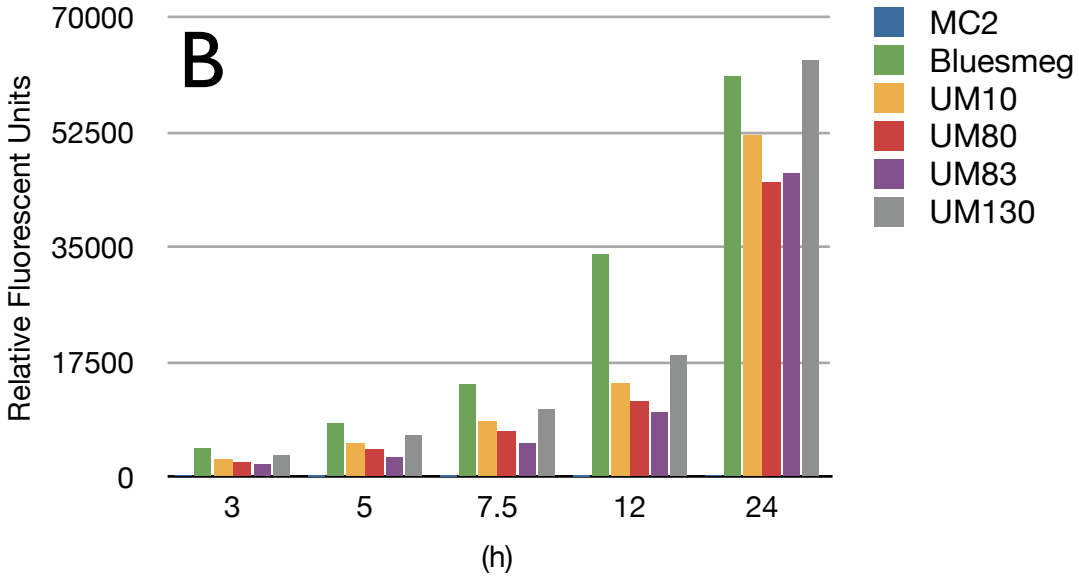



Figure 2.5: (A) Increase of fluorescent signal was measured by flow cytometry over time and compared between *Bluesmeg* and UM mutants 10, 80, 83, and 130. WT *M. smegmatis* was used as a negative control. As seen, no fluorescent signal was detected from WT throughout the assay. At 3 h, fluorescent signal from *Bluesmeg* was already higher than those UM mutants which clustered together between WT and *Bluesmeg*. At 12 h, the difference of fluorescent signal was greatest. The difference was less so at 24 h. (B) Fluorescent signal was also measured using fluorescent plate reader. The same trend was observed confirming the validity of this experiments.

### 2.4.4 UM Mutants Show FDG Uptake Deficient Phenotype

The 4 uptake mutants, UM10, UM80, UM83, and UM130, identified by sequencing were verified for their transport deficient phenotype. Flow cytometry allowed us to compare the change of fluorescence over time between the UM mutants and *Bluesmeg* (Fig. 2.5A). As shown in figure 2.5A, the fluorescence from *Bluesmeg* was always higher than the mutants at 12 h. At 24 h, the difference was less marked and the intensity of fluorescence in UM130 actually reached

the same as *Bluesmeg*. No fluorescence was detected in WT *M. smegmatis* as expected. Aliquots from the bacterial cells harvested for flow cytometry were also analyzed in a fluorescent plate reader to verify the flow cytometry result. The same trend was observed wtih the fluorescent plate reader and the difference between the mutants and *Bluesmeg* was greatest at 12 h and was diminished at 24 h (Fig. 2.5B). Taken together, we hypothesize that FDG is a substrate for several transport systems in *M. smegmatis*.

## 2.5 Discussion

A *lacZ* reporter system has been extensively used as a genetic tool in many organisms due to its versatility [34, 19, 72]. In mycobacteria, there have been several reports in literature which showed successful applications of the system. For instance, Srivastava R. *et al*. [87] developed a tool to rapidly screen efficacy of frontline anti-mycobacterial drugs by measuring $\beta$-galactosidase activity as a sign of cell viability. Raghavan S. *et al*. and Hillmann D. *et al*. [43, 70] constructed a promoter-*lacZ* transcriptional reporter to study gene regulation by adding different lengths of promoter sequence in front of *lacZ*. In our case, we used the *lacZ* reporter system to quantify uptake of $\beta$-galactosidase substrate, FDG, by *M. smegmatis* carrying pML803, a *lacZ* reporter system. We also constructed a recombinant *M. smegmatis* with *lacZ* integrated into the chromosome (data not shown). However, because this recombinant strain only gives one copy of *lacZ* per genome, the expression level of integrated *lacZ* was low and yielded reduced substrate specific signal.

FDG is a fluorogenic substrate in which fluorescein is quenched by di-$\beta$-D-

galactopyranoside. In order to emit fluorescence, di-$\beta$-D-galactopyranoside has to be cleaved off by $\beta$-galactosidase. This substrate was chosen for the following reasons; readily available, inexpensive, detection by FACS, and identification of active transport systems which likely play a role in nutrient transport. Data from the $NaN_3$ experiment (Fig. 2.4A) supported our hypothesis that active transport systems are involved in the uptake of FDG in *M. smegmatis*. The same data also showed that FDG is taken up by diffusion. The latter result was not surprising since FDG is known to be taken up by many organisms [74, 67]. Here, we were able to show that both passive and active transport systems can process uptake of the same substrate.

$C_{12}$FDG was also tested as a candidate substrate because we believed the $C_{12}$ moiety conjugated to fluorescein would allow the dye to retain in the cell much longer, away from worry of rapid efflux of fluorescein dye, which is an issue in FDG. Unfortunately, as reported previously [67, 74], $C_{12}$FDG was cell impermeable in bacteria and yeast and only non-viable cells were permeable to allow $C_{12}$FDG to enter. Detection of fluorescence from $C_{12}$FDG was extremely low and it was very difficult to distinguish from negative control (data not shown).

Sorting the bacterial cells was challenging. The key factor to select the mutants defective in FDG uptake in flow cytometry is a reduced fluorescence compared to the *Bluesmeg* control. However, the reduced fluorescence in the cells could be also seen in the cells losing fluorescence due to different causes. One cause is efflux of fluorescein from the cells by secretion or diffusion. Secretion or diffusion caused reduction of fluorescence in the cells by 2-log within an hour (Fig. 2.4C) after the bacterial cells were transferred to a fresh medium without FDG. To prevent selection of false-positive mutants in the sorting, the sorting

was carried on less than an hour after the harvest. Each pool was harvested while another pool was in the sorting process. This approach minimized the length of time the pool was standing in the wash buffer, in absence of FDG, therefore it minimized selecting the mutants which showed defective FDG uptake phenotype due to enhanced loss of fluorescence. Another possible cause of loss of fluorescence is mutation in *lacZ* or on the plasmid which could cause inactivation of $\beta$-galactosidase due to transposon insertion by *M. smegmatis*. In the latter case, the reason could be that too much of $\beta$-galactosidase is not favourable to the cells. The Miller assay was performed on the initial screen of mutants and showed about 42% of 521 mutants had an inactive *lacZ* gene. It was surprising to see such a large number of mutants with inactive *lacZ*, but given the nature of $\phi$MycoMarT7 random transposition event, this may be predicted. For the next sorting, the factor will be taken into consideration while deciding the numbers of mutants to be sorted in flow cytometry.

Our purpose in this study was to develop a screening tool to identify transporters which are involved in uptake of a target substrate. This aim was accomplished by using a transposon library of a $\beta$-galactosidase reporter strain of *M. smegmatis*, *Bluesmeg*, along with fluorescein di-$\beta$-D-galactopyranoside (FDG) as a target substrate. It was our goal to discover uptake transport genes responsible for the target substrate in the screen, and it was achieved by the identification of the 4 UM mutants. We are hoping to apply the same approach to discover sugar uptake mutants in *Mtb* in the future. Also, it is plausible to expand this approach to discover transporters of different substrates by changing the substrate of interest in screens, and it can be done not only in mycobacteria, but in different organisms. In this screening experiment, we only targeted uptake transport mutants, but it is possible to discover efflux transport mutants by

selecting reverse phenotype in the screen; cells retaining fluorescence.

We used an unbiased screen to identify transporter mutants using FDG as the target substrate. Quite a number of the UM mutants had a transposon insertion on the plasmid and these mutants were eliminated from further identification. Sequences of the remaining UM mutants elucidated that many had a transposon insertion in genes involved in cell membrane/wall syntheses, suggesting "off-target" effects. However, we were excited to identify 3 UM mutants, UM80, UM83, and UM130, belonging to the ABC transporter family. What made us more excited was that all 3 of them share the same transporter architecture with the unique accessory domain, substrate binding domain (SBP), which is exported into extracellular fluid. UM10 was another interesting transposon mutant. It does not belong to the ABC transport family like the other 3 UM mutants, however it possesses the NTF2 domain motif. NTF2 protein in eukaryotic cells plays a crucial role in nuclear import of RanGDP. Also, VirB8, an essential protein in a part of Type IV secretion system in Gram negative pathogens like *Agrobacterium tumefaciens*, possesses the NTF2 domain motif.

Here, we identified 4 UM mutants and the possible transport systems they belong to. Next, we will characterize the 4 UM mutants on their substrate preferences. As these UM mutants were identified via the target substrate, FDG, we are curious to know their true substrates for uptake.

CHAPTER 3

# DISSECTING SUGAR AND PEPTIDE TRANSPORT PHENOTYPES OF THE FDG UPTAKE DEFICIENT MUTANTS

## 3.1 Abstract

Screening for mutants possessing the deficient FDG uptake phenotype identified 4 uptake mutants (UMs), UM10, UM80, UM83, and UM130 in *Mycobacterium smegmatis*. Genetic identification revealed that UM80 (*MSMEG_0556::Tn*), UM83 (*MSMEG_4995::Tn*), and UM130 (*MSMEG_4099::Tn*) all have mutation in loci encoding for the ATP-Binding Cassette (ABC) transporter superfamily. Each locus is predicted to function as sugar, oligopeptide, or dipeptide transporter, respectively. *MSMEG_4363*, the gene disrupted in UM10, encoding a hypothetical protein possessing the nuclear-transporter-factor 2 (NTF2)-like family motif, but with no homology to known transporters. Through extensive bioinformatic search, VirB8, a small protein essential in type IV secretion system (T4SS) also possesses the NTF2-like family motif and its secondary structure shows high similarity to that of MSMEG_4363. In this chapter, the possible roles of these transporters in nutrient acquisition are described.

## 3.2 Introduction

*Mycobacterium smegmatis* is a non-pathogenic microorganism which is widely used as a model system to work on understanding cellular mechanisms of its pathogenic relative, *M. tuberculosis*. Unlike *Mtb*, the maximum biosafety level

required for *M. smegmatis* is the biosafety level 2 (BSL2), which is what most laboratories are equipped. Another benefit of working with *M. smegmatis* is that it grows relatively faster, a colony arises within 3 days, compared to *Mtb*, which takes about a month to show a visible colony on a solid agar plate. Although their genomic size is fairy different (6.9 Mb for *M. smegmatis* and 4.4 Mb for *Mtb*), quite a number of their cellular mechanisms are comparable and considering the benefits described above, it is worth investigating in *M. smegmatis* first. This is the main reason why we looked for potential transporters in *M. smegmatis* for nutrient acquisition.

The ATP-binding cassette (ABC) transporters are transmembrane proteins with representatives in all the domains of life from prokaryotes to archaea and eukaryotes. Among the families of ABC transporters, ABC importers are unique by the additional protein domain, substrate binding protein (SBP), whose main function is to scavenge substrates for transport. SBP is a high-affinity substrate domain which governs substrate specificities of individual transporters. A number of transporters belonging to this family plays a crucial role in importing substrates essential for survival of the organisms. For instance, peptide transporters import peptides which are used as nutrients and function as signals for cellular regulation of gene regulation such as for sporulation, and for developing virulency [42, 52, 32].

A well-studied ABC transporter systems for peptide uptake are the dipeptide (Dpp) and oligopeptide (Opp) transporters in *Lactococcus lactis* [49, 16, 29]. Peptides are generated by degradation of the milk proteins by PrtP, the cell-wall associated proteinase and are the essential source of amino acids for the bacterium to grow on milk. Both Dpp and Opp transporters are required to

transport the peptides into the cells. The Dpp transporter mediates uptake of di- and tripeptides and the Opp transporter mediates uptake of oligopeptides up to 35 amino acids in length [16, 28]. Opp transporters are also found in Archaea. Archaea are known to inhabit in the extreme environments in which most proteins in bacteria are inactive. Palmieri *et al*. identified the first archeal Opp transporter in *Aeropyrum pernix*, which grows optimally at 90°C [63]. The $Opp_{Aper}$ transporter is an ABC transporter and shows a significant similarity to bacterial Opp transporters. An Opp transporter found in another archaeon, *Sulfolobus solfataricus*, is also ATP dependent and shows resemblance to the bacterial Opp transporters [35].

Some carbohydrates are also taken up via the ABC importers. In *Streptococcus mutans*, maltose and maltodextrins are transported via at least two sugar transporters, MsmEFGK and MalXFGK [98]. Many bacteria other than *Streptococcus mutans* also acquire maltose/maltodextrins via the similar transport mechanisms [80]. Detailed biochemical mechanisms of a maltose/maltodextrin transporter are well described in the recent review by Bordignon *et al* [11].

To our knowledge, the first report of peptide acquisition via ABC transporter system in mycobacteria appeared in 2000 when Green *et al*. [38] tested the toxic tripeptides glutathione and *S*-nitrosoglutathione (GSNO) to *M. bovis* BCG and its *opp* mutant strain. In their study, the *opp* mutant strain showed resistance to glutathione concentrations up to 10 mM whereas *M. bovis* BCG was not able to grow in presence of glutathione at 2 mM. Toxicity of GSNO was measured on [$^3$H] uracil incorporation in the cellular mass. When GSNO was present at 2 mM concentration, only 40% of [$^3$H] uracil was incorporated in *M. bovis* BCG whereas 81% of [$^3$H] uracil was incorporated in the mutant strain, strongly

supporting a role of the Opp transporter system in *M. bovis* BCG as a tripeptide transporter. Although interestingly, none of the 25 peptides ranging from two to six amino acids in length supported growth of *M. bovis* BCG as a sole source of nitrogen source.

The Opp transporter in *Mtb* was recently studied by Flores-Valdez *et al*. [32]. In their study, the function of the Opp transporter in modulation of cell wall composition and its involvement to survival of the bacteria in chronic infections in a mouse model was addressed instead of its role in nutrient acquisition. Following the study, another research group published a follow-up paper on the $Opp_{Mtb}$ transporter and described its role as a regulator of cytokine secretion and apoptosis of infected macrophages [24].

In case of carbohydrate transport systems in mycobacteria, a number of putative carbohydrate transport proteins have been predicted in both *Mtb* and *M. smegmatis* [92] but only one carbohydrate transport system has been elucidated experimentally [46]. This carbohydrate transporter system in *Mtb*, LpqY-SugA-SugB-SugC, exclusively transports trehalose which is released from the bacterial cell wall and the transporter functions as a trehalose recycling system. The study showed that the elimination of the transport system strongly perturbed the virulence of *Mtb* in a mouse chronic infection model.

In our previous study (Chapter 2), we identified 4 uptake mutants (UMs), UM10, UM80, UM83, and UM130, by a high-throughput screening using a transposon library of a $\beta$-galactosidase reporter strain of *M. smegmatis*, *Bluesmeg*, along with fluorescein di-$\beta$-D-galactopyranoside (FDG) as a target substrate. The genes disrupted by 3 of the UM mutants, UM80, UM83, and UM130, though they are all annotated to import different substrates, belong to the ABC

importers. In case of UM10, the inactivated gene has no homology to known ABC transporter gene, however its translated sequence shows a NTF2-like domain.

Here, we tested all the UM mutants on both peptides and carbohydrates for potentially preferable substrates without limiting types of substrates. Through our study, we noticed that the predicted substrate specificities of these transporters do not correspond to the substrate specificities observed experimentally.

## 3.3 Materials and Methods

### 3.3.1 Bacterial strains, plasmids, and culture conditions

Bacterial strains and plasmids used in this work are listed in Table 3.1. Plasmids were manipulated and maintained in *Escherichia coli* at 37°C on Luria-Bertani (LB) agar or in LB broth. When necessary, *E. coli* strains were grown in presence of 50 $\mu$g/ml kanamycin or 50 $\mu$g/ml hygromycin. Either Middlebrook 7H9 medium (supplemented with Middlebrook OADC, 0.2% glycerol, and 0.05% Tween-80) or 7H10 agar (supplemented with Middlebrook OADC, 0.2% glycerol) was used for routine maintenance of *M. smegmatis* strains. Either 50 $\mu$g/ml kanamycin or 150 $\mu$g/ml hygromycin was added to the cultures when necessary. Hartmans-de Bont (HB) medium [42] was employed with a slight modification in peptide/sugar uptake assays. Tyloxapol was added at a final concentration of 0.05% (vol/vol) in all the studies instead of 0.05% Tween-80 since Tween-80 can be used as an additional carbon source. Original HB medium contains 0.2% glycerol as a carbon source and 0.02% ammonium sulfate as a

nitrogen source. For peptide uptake assay, the nitrogen source was replaced with 1~10 mM peptide. For carbohydrate uptake assay, the carbon source was replaced with 10 mM carbohydrate.

### 3.3.2 Efflux of fluorescein

To confirm the reduction of fluorescence observed from the UM mutants in flow cytometry was solely due to defective uptake systems, we tested the level of fluorescence secretion. Briefly, all the cultures were grown in the 7H9 media in presence of FDG (Invitrogen, Carlsbad, CA, USA). Aliquot was taken at time intervals up to 24 h and was centrifuged at 16, 000 x g at RT for 5 min to separate into cells and supernatant. 200 $\mu$l of the supernatant was transferred to a 96-well plate, with black wall and clear bottom for maximum fluorescent reading (Costar 3603, Corning, NY, USA). Fluorescent in the supernatant was measured at 514 nm with excitation wavelength at 490 nm in Spectra MAX GEMINI EM (Molecular Devices, Sunnyvale, CA, USA).

Additionally, to verify FDG is specifically taken up via active transporters the UM mutants were tested on another fluorogenic substrate, fluorescein diacetate, FDA (Invitrogen), and the fluorescent profiles were compared to those of FDG. FDA is a fluorogenic substrate like FDG, but unlike FDG, it is a cell-permeant substrate. It can be readily taken up into the cell via diffusion. Also unlike FDG, it does not require specific enzyme to be activated. Diacetate moiety of FDA can be cleaved off by any cellular esterase. Briefly, WT, *Bluesmeg*, and the UM mutants were grown to late log phase ($OD_{600}$ 1.2 to 1.6) in the 7H9 medium with appropriate antibiotics at 37°C with agitation at 100 rpm.

Aliquots were used to seed 20 ml cultures in the same medium condition with a starting optical density at 600 nm ($OD_{600}$) of 0.1. The cultures were incubated at 37°C with agitation at 100 rpm. One ml was taken at 3 h, 6 h, and 12 h and was centrifuged at 16 000 x g at RT for 5 min. The supernatant was transferred to a 96-well plate to measure fluorescence in the Spectra MAX GEMINI EM. The cell pellet was washed once with (vol/vol) PBS/0.05% Tween-80 and resuspended in 1 ml of the same buffer for analysis in flow cytometry. The cell was homogenized well using 25G 5/8 needle before applied to FACSCalibour (BD Biosciences, San Jose, CA, USA ). Fluorescence was detected using a FITC (530 ± 15 nm) filter.

### 3.3.3 SDS sensitivity assay

The sensitivity to SDS was used as a measure of cell membrane integrity. The method described by White *et al*. [100] was employed here with modification. Briefly, WT and the UM mutants were grown to a mid-log phase ($OD_{600}$ between 0.4 and 0.6) in the 7H9 medium and treated with 0.1 % SDS. The cultures were incubated at 37°C with agitation at 100 rpm and aliquots were taken at 2 h, 4 h, 8 h, and 12 h post exposure to SDS. The cultures were diluted up to 7-fold and 3 consecutive dilution series were plated onto solid agar medium. Plates were incubated at 37°C for 3-4 days and cell wall integrity was scored based on the CFU counts. For disc diffusion assays, cultures were added to 5 ml 0.6 % 7H9 top agar at final $OD_{600}$ of 0.6 and 1 ml was poured onto each quarter of bottom agar. Sterile filter disc was placed on each quarter and impregnated with either 5 $\mu$l of SDS at increment concentrations of 0, 0.1, 1, and 10 % or 5 $\mu$l of $_D$-cycloserine at increment concentrations of 0, 1, 10, and 100 mg/ml. The

plates were incubated at 37°C and the zones of inhibition were measured after 2 days.

### 3.3.4 Curing pML803 plasmid from UM mutants

To minimize possible effect of pML803 products, such as $\beta$-galactosidase and hygromycin B phosphotransferase, in characterizing the UM mutants, pML803 was cured from the UM mutants. Briefly, the cultures were inoculated in 7H9 medium containing kanamycin, without hygromycin B, and grown overnight at 37°C. Aliquots (1/20 vol of the medium) were transferred to a fresh medium and incubated further at 37°C. This process was repeated 3 times and the cultures from the last passage was plated on 7H10 solid agar containing kanamycin. Individual colonies were picked and replica plated on the 7H10 solid agar containing either kanamycin or hygromycin B. Only UM mutants that grew on the kanamycin plate, but not on the corresponding hygromycin plate, were selected as pML803 cured UM mutants and these mutants were used for further experiments. These UM mutants were designated with ".1" added after their UM name (Table 1).

### 3.3.5 Bialaphos toxicity assay

Bialaphos is a toxic tripeptide compound which is transported via Ddpp/opp transport mechanism(s) in *Sinorhizobium meliloti* and Opp in *Mtb* [62, 32]. The authors claimed that resistance against bialaphos toxicity in those mutants was due to the absence of the Dpp/Opp transport system.

| Strains or plasmids | Descriptions | Sources or references |
|---|---|---|
| Strains | | |
| *Mycobacterium smegmatis* $MC^2$ 155 | Laboratory strain | ATCC 700084 |
| *Bluesmeg* | $MC^2$ 155 carrying pML803 $Hyg^r$ | This study |
| $MC^2$ 155/pJV53 | $MC^2$ 155 carrying pJV53 for recombineering of linear DNA into $MC^2$ 155 $Km^r$ | [96] |
| UM10 | $MC^2$ 155 *MSMEG_4363*::Tn; $Km^r$ pML803 *E.coli*-mycobacterial shuttle vector encoding *lacZ* and $Hyg^r$ | This study |
| UM80 | $MC^2$ 155 *MSMEG_0556*::Tn; $Km^r$ pML803 *E.coli*-mycobacterial shuttle vector encoding *lacZ* and $Hyg^r$ | This study |
| UM83 | $MC^2$ 155 *MSMEG_4995*::Tn; $Km^r$ pML803 *E.coli*-mycobacterial shuttle vector encoding *lacZ* and $Hyg^r$ | This study |
| UM130 | $MC^2$ 155 *MSMEG_4099*::Tn; $Km^r$ pML803 *E.coli*-mycobacterial shuttle vector encoding *lacZ* and $Hyg^r$ | This study |
| UM10.1 | UM10 pML803 cured strain; $Km^r$ $Hyg^s$ | This study |
| UM80.1 | UM80 pML803 cured strain; $Km^r$ $Hyg^s$ | This study |
| UM83.1 | UM83 pML803 cured strain; $Km^r$ $Hyg^s$ | This study |
| UM130.1 | UM130 pML803 cured strain; $Km^r$ $Hyg^s$ | This study |
| Plasmids | | |
| pML803 | a *E.coli*-mycobacterial shuttle vector carrying *lacZ*; $Hyg^r$ | a gift from Michael Niederweis |

Table 3.1: All the strains and plasmids used in this study

As we have two UM mutants each disrupting an ATP transporter permease predicted as a part of either Dpp or Opp ATP transporter, we investigated whether our UM mutants also exhibit resistance against a tripeptide antibiotic, bialaphos (Gold Biotechnology, St. Louis, MO, USA). The UM mutants were grown in the 7H9 media in presence of bialaphos, and growth was used as an indication of resistance. The cultures were grown overnight to a late log phase ($OD_{600}$ = 1.2 to 1.6) and inoculated at starting $OD_{600}$ of 0.01 into a fresh 7H9 medium containing 25 $\mu$g/ml bialaphos. The cultures were grown at 37°C with shaking at 100 rpm and the growth was measured at $OD_{600}$ periodically until

the growths reached stationary phase.

### 3.3.6 Bialaphos/peptides uptake competition assays

To prove peptides and bialaphos share the same transport systems, WT was challenged under bialaphos toxicity in presence of various peptides as competitive substrates for transport. Briefly, the cells were inoculated in the 7H9 media in presence of bialaphos. Various peptides were separately introduced to the media either at the same ratio or up to 100 times the molar concentration of bialaphos. The culture was incubated at 37°C with agitation at 100 rpm. Survival of *M. smegmatis* was determined by measuring growth at $OD_{600}$.

### 3.3.7 Peptide and sugar transport assays

To distinguish substrate specificity of each UM mutant, target substrates were added to a medium as a sole nutrient source. In case of peptides, they were used as a sole nitrogen (N) source in the HB medium at 1 mM concentration. In case of carbohydrates, they were used as a sole carbon (C) source in the HB medium at 10 mM concentration. Briefly, overnight cultures were harvested and aliquot was inoculated at starting $OD_{600}$ of 0.01 into the HB medium containing appropriate source of nutrient in triplicate. The cultures were incubated at 37°C with agitation at 100 rpm and samples were taken periodically to measure the growth at $OD_{600}$.

### 3.3.8 Uptake and utilization of biotinylated peptides

To eliminate the possibility that the results from the peptides/sugar uptake transport assays are due to the inability of the UM mutants to utilize the peptides as nutrient sources, biotin conjugated peptides (Anaspec, Inc) were synthesized and used as substrates. Biotin was chosen as a conjugate probe because there is a specific transporter for biotin in bacteria [39]. The cultures were grown overnight and inoculated in triplicates at starting $OD_{600}$ of 0.01 into a well containing 1 ml of a fresh HB medium without N-source in the 24-well plates. In one plate, the biotin-conjugated form of peptides were added in triplicate at a final concentration of 5 $\mu$M and the non-conjugate form of peptides were added in the same manner in the corresponding wells in another plate. All the plates were incubated at 37°C without shaking and $OD_{600}$ was measured directly in the plates using the KC junior software in PowerWave XS (BioTek, Winooski, VT, USA).

### 3.3.9 BIOLOG phenotypic microarray assays to elucidate N-source preference of each transporter system

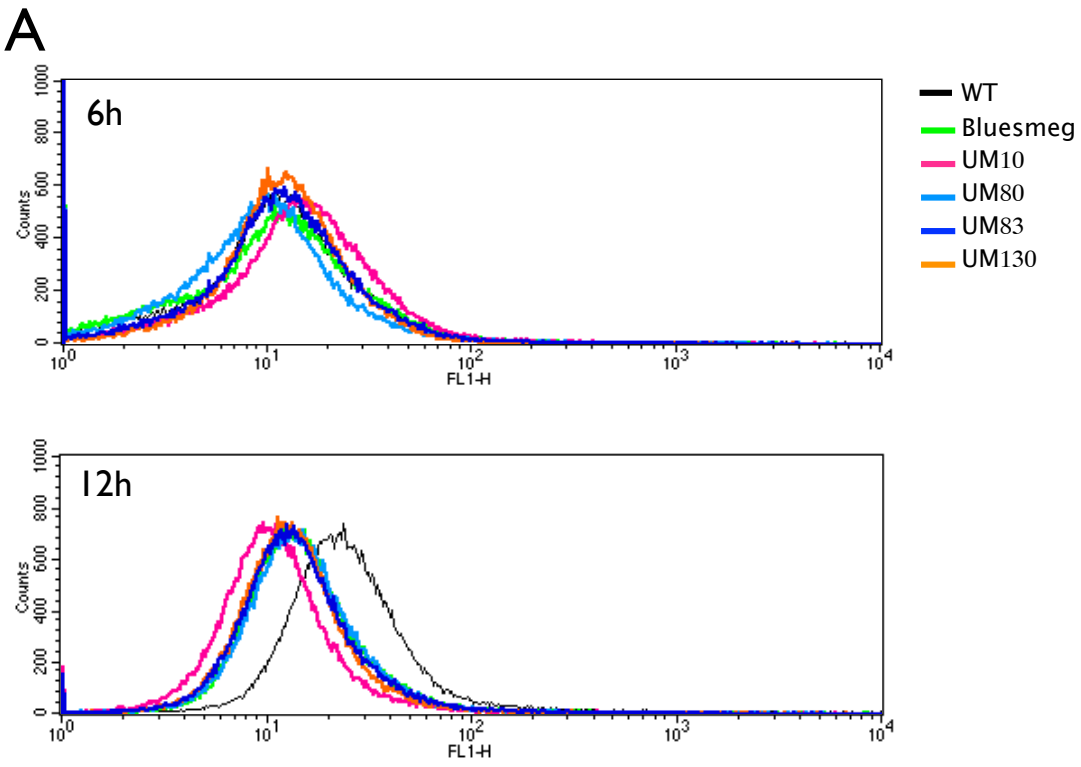
All the UM mutants and WT were tested on BIOLOG phenotypic microarray (PM) plates (BIOLOG, Hayward CA) for their ability to utilize various N-sources. BIOLOG PM 3B, 6, and 7 contain various kinds of N-source and were used in this study. Each well of the PM 3B plate contains amino acids at concentration of 4.5 mM. The wells of the PM 6 and 7 plates contain peptides at concentration of 2 mM except the negative wells contain no amino acid and the positive wells contain $_L$-Glutamine at 4.5 mM. Inoculum for the PM assay

plates were prepared on the 7H10 agar plate. Individual colonies were scraped off from the plate and resuspended into the HB medium which contains no N-source. The bacterial cell suspension was left at RT without shaking for 2 days to deplete intracellular N-source before the assay. The cell suspension was adjusted to achieve $OD_{600}$ of 0.092, equivalent to 81% light transmittance (%T), in inoculating fluid, IF-0a (BIOLOG), as recommended by manufacturer. The complete assay suspension for each PM plate was prepared as follows: 0.88 ml of the inoculum, 1 ml 12X PM additive (2 mM $MgCl_2 \cdot 6H_2O$, 1 mM $CaCl_2 \cdot 2H_2O$, 0.01% Tween-80, and 5 mM $_D$-glucose), 10 ml IF-0a, and 0.12 ml Dye mix H. 100 $\mu$l of the complete assay suspension was added to each well and the plates were sealed with sterile breathable films (Axygen Scientific, Union City, CA, USA) to minimize contamination from neighboring wells. All the plates were incubated at 30°C. Metabolism of individual N-sources was determined by measuring the conversion of tetrazolium-based chemical to formazan at 590 nm using the KC junior software in PowerWave XS. Formazan is soluble purple colour compound which is formed by reduction of tetrazolium salt by dehydrogenase and reductases. Increased concentration of formazan is used as a sign of metabolic activity.

## 3.4 Results

### 3.4.1 Efflux systems are not responsible for the deficient FDG uptake phenotype in the UM mutants

Our screening in the previous chapter identified 4 UM mutants, which showed reduced fluorescent in the cells with FDG, compared to WT, by flow cytometry. We believed this phenotype was solely due to defective uptake transport systems in the mutants. To exclude that the reduction of fluorescence in the bacterial cells was the result of a secondary effect, i.e., accelerated efflux of fluorescence in the UM mutants, we employed another fluorogenic substrate, FDA. FDA is a cell-permeant fluorogenic substrate which emits fluorescent when cellular esterase cleaves off diacetate on the molecule. The purposes of using this substrate were; 1) to measure and compare level of fluorescence secretion and 2), to eliminate efflux as an explanation for differential intra-bacterial signal. Since FDA can be readily taken up by cells via diffusion, there should be no difference of fluorescence in the cells. This was shown in Figure 3.1A. No significant difference of fluorescence level was observed in all UM mutants compared to WT and *Bluesmeg*. When fluorescence in the supernatants was analysed, it was clear that the UM mutants did not secrete more fluorescence (Fig. 3.1B). In fact, all the UM mutants secreted much less fluorescence compared to *Bluesmeg*. When FDA was used as a substrate, level of fluorscein secretion was same in all the strains. Combined these two results, we conclude that the UM mutants have mutation in uptake transport mechanisms.


A
6h
12h
Counts
FL1-H
WT
Bluesmeg
UM10
UM80
UM83
UM130

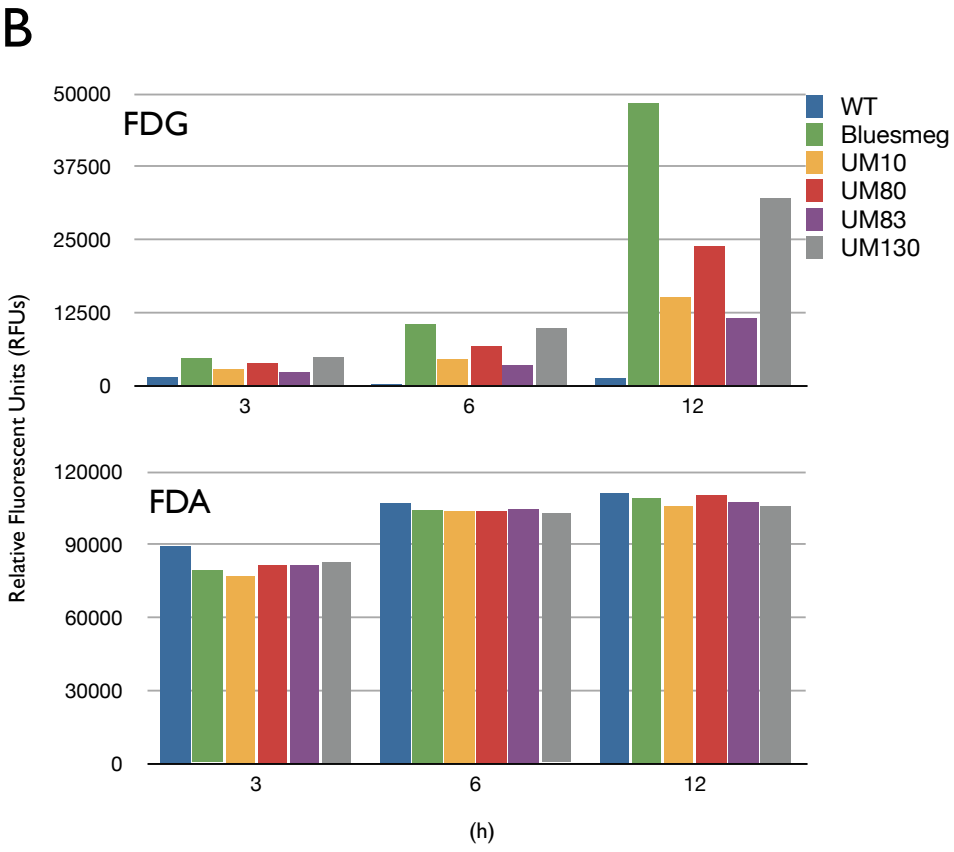

B
FDG
FDA
Relative Fluorescent Units (RFUs)
(h)
WT
Bluesmeg
UM10
UM80
UM83
UM130

Figure 3.1: Fluorescent was measured from the cultures growing in presence of either FDG or FDA. (A) showed comparison of fluorescence secreted when FDG was used as a substrate or FDA as a substrate. It is clear from the FDG graph that less fluorescence in the mutant cells previously observed in flow cytometry was not due to excess secretion of fluorescence. No difference of secreted fluorescence was observed in the UM mutants compared to *Bluesmeg* and WT when FDA was used as a substrate, suggesting that the transposon mutation in these UM mutants did not affect their secretion mechanism. (B) showed no difference in fluorescence in the cells when tested with FDA.

### 3.4.2 Genetic characterization of the UM mutants

Because FDG is not a nutrient, we were curious to see what transporters would be identified from the screening. Figure 3.2 shows the genetic maps of the genes disrupted in the UM mutants and the loci these genes are encoded. Identification of the 4 UM mutants revealed that, except for UM10, they belong to a unique family of ABC transporter, the ABC importer family. Typical ABC transporter family is organized in a characteristic fashion: 2 transmembrane domains and 2 ATP-binding domains. The transmembrane domains play a role in import/export of molecules and the ATP-binding domains hydrolyze ATP to provide energy for the transport. These domains are expressed as separate polypeptides but in some cases, they might be fused into multifunctional polypeptides. ABC importer family has an additional domain: a substrate-binding protein (SBP), which is usually attached to transmembrane with its lipoprotein anchor in Gram positive bacteria. In Gram negative bacteria, SBP can be freely secreted to the periplasmic space.

UM10 has a transposon insertion in *MSMEG_4363* which is an orphan gene and encodes a 208 aa polypeptide, annotated as a hypothetical gene

with unknown function. In its TIGR gene list (http://cmr.jcvi.org/tigr-scripts/CMR/shared/AllGeneList.cgi?sub_org_val=gms&feat_type=ORF), it is annotated to encode a signal peptide. A predicted secondary structure of MSMEG_4363 by PredictProtein [73] possesses its N-terminus located inside the cytoplasm and spaned once in transmembrane and translocated its C-terminus in periplasm (Appendix 1). It is also predicted to show amino acid sequences for protein-protein interaction. By conserved domain search, *MSMEG_4363* encodes a domain for a nuclear transport factor 2 (NTF2-like) superfamily (NCBI cl09109). Proteins in this superfamily are found widespread from prokaryotes to eukaryotes and share many common details in secondary structure. However, their functions are greatly diverse and proteins in this superfamily are further categorized into up to 30 families. For instance, as the name states, NTF2 proteins in mammalian cells and *S. ceravisiae* are important in nuclear localization of Ran, mediating nuclear import of RanGDP, providing a gradient of Ran between cytoplasm and nucleoplasm [69, 47].

For *MSMEG_4363*, it likely plays a role in transporting molecules into the cell, shown by the FDG uptake defect phenotype of its mutant. No homologous gene was found in *Mtb* and other species in the *Mtb* complex, though these species also possess proteins with the NTF2-like motif. Genes from other mycobacterial species which showed high identity were *MSMEG_6770* and *Mvan_5451* of *M. vanbaalenii* PYR-1, an environmental species which is known to utilize polycyclic aromatic hydrocarbons (PAHs). The gene downstream of *MSMEG_4363*, *MSMEG_4362*, is annotated as a universal stress protein, and has UspA of *M. marinum* as the closest homolog. The gene, *MSMEG_4365*, just downstream of *MSMEG_4363* encodes for a hypothetical protein of unknown function. Its predicted secondary structure shows no transmembrane spans

and no regulated secondary structures (NORS), but is predicted to possess a signal sequence at its N-terminus. PSI-BLAST search of the protein against Genbank mostly shows homology to predicted uncharacterized proteins but also shows an interesting hit with proteins predicted as carbohydrate-binding proteins. The closest homolog is Bcav_1631 of *Beutenbergia cavernae*, a Gram-positive bacterium belonging to the suborder Micrococcineae.

UM80 has a mutation in *MSMEG_0556* which possesses a signature motif of an ABC transporter ATPase. It is annotated as a glycerol-3-phosphate transporter ATP-binding subunit, *ugpC*, and shows 54 % identities to *sugC* in *Mtb* by BLASTP search. The individual genes in the locus showed high identities to the corresponding genes in the *Mtb*LpaY-SugA-SugB-SugC locus by the BLAST search, however the *MSMEG_0553-0556* transporter locus does not align with the *Mtb*LpaY-SugA-SugB-SugC locus. This suggests that the *MSMEG_0553-0556* locus is not homologous to the *Mtb* trehalose transporter[46].

The *MSMEG_0553-0556* transporter locus is organized with an ABC transporter substrate binding protein (SBP) gene, *MSMEG_0553*, which possesses TAT signal sequence, two ABC transporter permeases, *MSMEG_0554* and *MSMEG_0555*, and an ABC transporter ATPase, *MSMEG_0556*. Just upstream of the locus, the gene *MSMEG_0552* encodes for an OsmC-like protein. The family of OsmC is involved in response mechanisms to environmental changes such as hyper/hypo-osmotic condition [99] and in defense against organic hydrogen peroxide [22]. Relation between the role of *MSMEG_0552* and the *MSMEG_0553-0556* locus is not known. However, like OpuA in *Lactococcus lacits*, whose locus functions to mediate uptake of organic osmolyte, glycine betaine, under hyper-osmotic conditions and belongs to the ABC transporter superfamily

[95, 101], we cannot eliminate the possibility that the *MSMEG_0553-0556* locus and *MSMEG_0552* also have such relation.

UM83 mutant disrupts the gene, *MSMEG_4995*, which encodes for AppB, oligopeptide transport system permease protein in the *MSMEG_4995-4999* locus. The organization of genes is different from that in the the *MSMEG_0553-0556* locus, that two permeases are encoded in the beginning of the locus followed by an ATPase and SBP. *Mtb* homolog of *MSMEG_4995* is *Rv1283c* encoding for OppB.

UM130 mutant has the gene, *MSMEG_4099*, disrupted. *MSMEG_4099* encodes for an ABC transporter permease like *MSMEG_4995* in UM83. The gene organization of *MSMEG_4101-4098* is somewhat similar to that of *MSMEG_0553-0556*, yet this locus extends to include genes involved in a downstream metabolic pathway. The function of those metabolic genes in conjunction to the ABC transporter genes is yet to be identified. *MSMEG_4099* shares a homology to *Rv3664c* in *Mtb*, which plays a role in Dpp ABC transporter [32, 24]. However like the case in the *MSMEG_0553-0556* transporter locus, the rest of the genes in the *MSMEG_4101-4098* locus do not show homology to the corresponding genes in *Mtb dpp* ATP transporter, suggesting that the MSMEG_4101-4098 locus does not share biological functions to the *Mtb* Dpp ATP transporter.

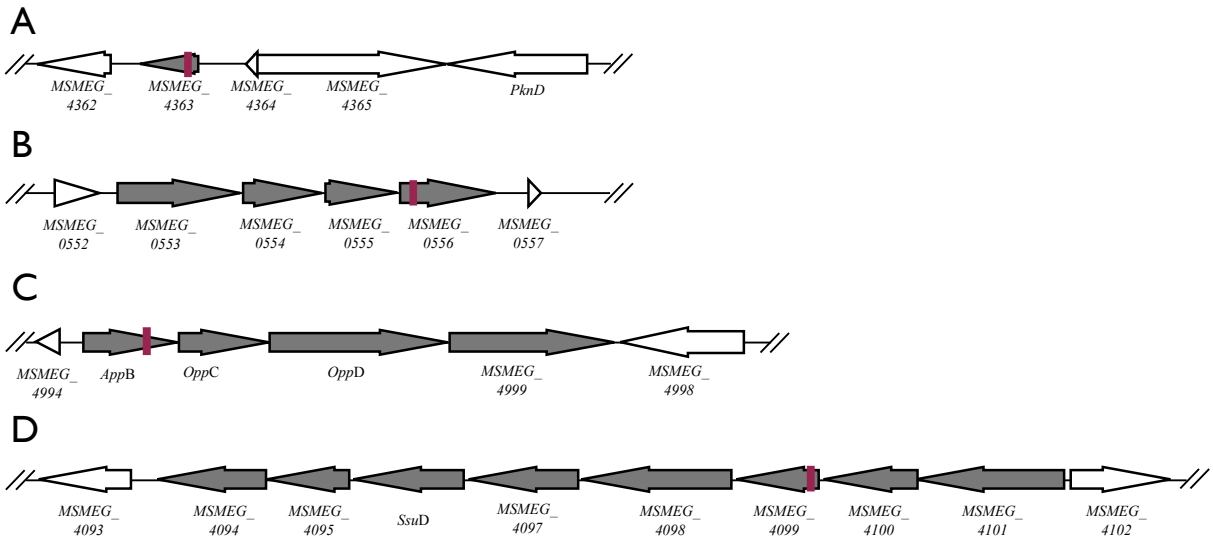



Figure 3.2: Genetic maps of the loci and surrounding genes of the UM mutants. (A) UM10 mutant, (B) UM80 mutant, (C) UM83 mutant and (D) UM130 were shown here. A burgundy bar in the arrow indicates location of transposon insertion site. Arrow indicate transcriptional orientations of each gene. Arrows in gray are in the same locus except *MSMEG_4363*, which is an orphan gene.

### 3.4.3 Resistance to bialaphos was enhanced by the presence of another antibiotic

Bialaphos is a toxic tripeptide which is commonly used as a herbicide and antimicrobial agent [79]. It contains 2 alanine residues and phosphinothricin, an analog of glutamine. It is a prodrug that requires host enzymes to act upon bialaphos. Intracellular peptidase cleaves off the 2 alanine residues and releases phosphinothricin. This compound, phosphinothricin, inhibits the activity of glutamine synthetase, leading to a rapid accumulation of intracellular ammonia levels. Early study by Higgins and Gibson [42] showed that mutation in the *Escherichia coli* Opp permease conferred resistance to bialaphos. Recently, a study

by Flores-Valdez *et al*. also showed that deletion of the Opp transporter locus (*Rv1283c-Rv1280c*), but not the Dpp transporter locus (*Rv3665c-Rv3663c*) in *Mtb* caused *Mtb* to be resistant against bialaphos [32]. Among the UM mutants, the transporter locus of UM83 is the only locus annotated as an Opp ABC transporter. As all the UM mutants exhibit similar phenotype in the FDG uptake, this assay was originally performed to discriminate one mutant from the others. The concentration of bialaphos, 25 $\mu$g/ml, used in this assay was 5 times higher than the concentration used in the Flores-Valdez *et al*. study. When bialaphos was added at the lower concentration, WT was able to grow without any growth defect. This was probably due to 1) *Mtb* has higher sensitivity to bialaphos and 2) liquid culture was used in this study whereas solid agar was used in their study. Therefore, a different concentration was required to perform our assay.

All the UM mutants were capable of growing in the 7H9 media with no significant difference to WT (Fig. 3.3A). This result indicates that the genes disrupted by the mutants are not essential during growth *in vitro*. UM83 showed resistance to bialaphos (Fig. 3.3) as predicted from its homology to the Opp permease in *Mtb*. However, all the other mutants also showed the same degree of resistance against bialaphos to the same extent as UM83. This result was unpredicted because Flores-Valdez *et al*. reported that the *Mtb dpp* knock-out (KO) strain exhibited WT-like susceptibility. It might be difficult to directly compare our result to their result since our methodology was different and we tested on a transposon mutant.

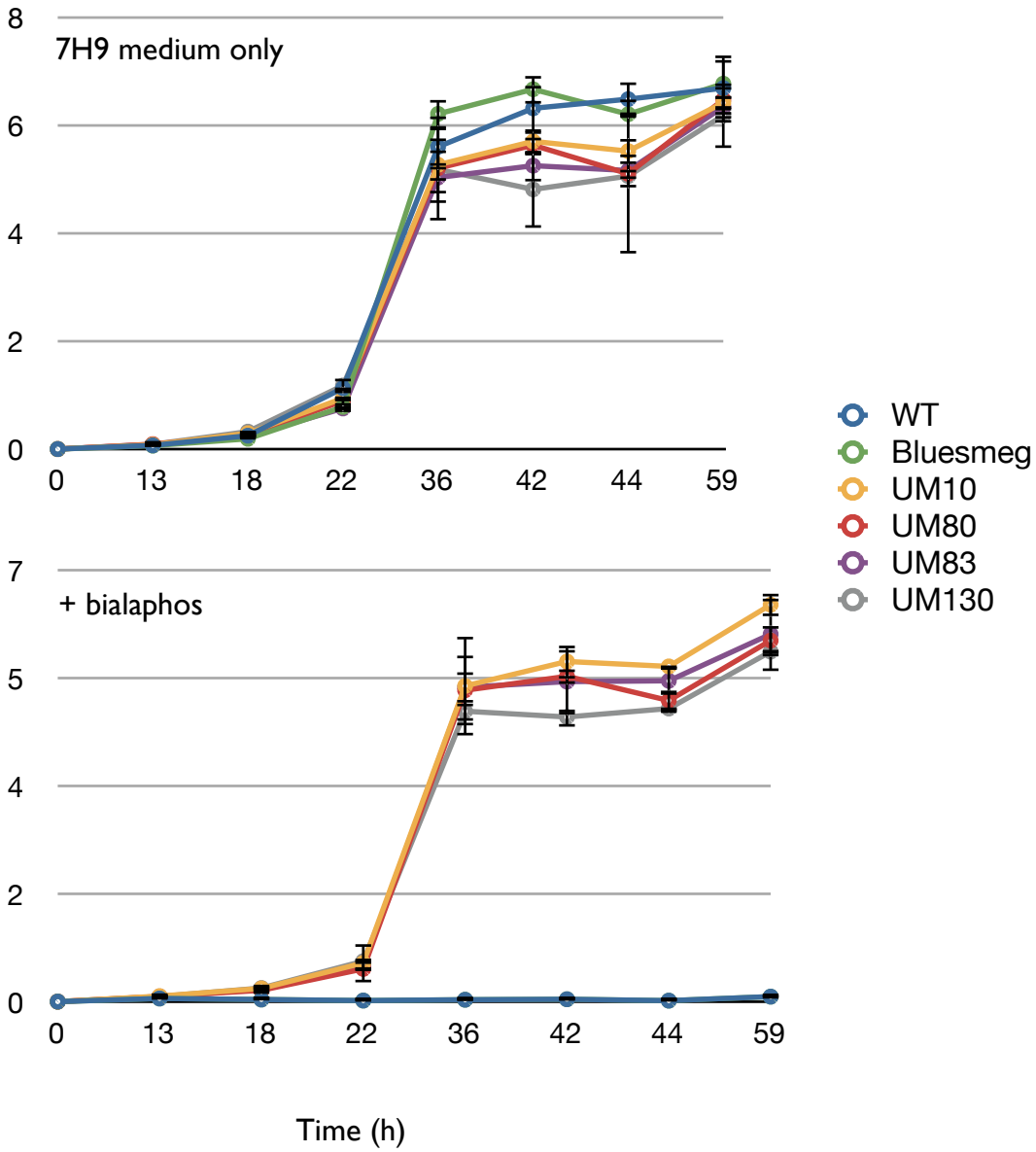
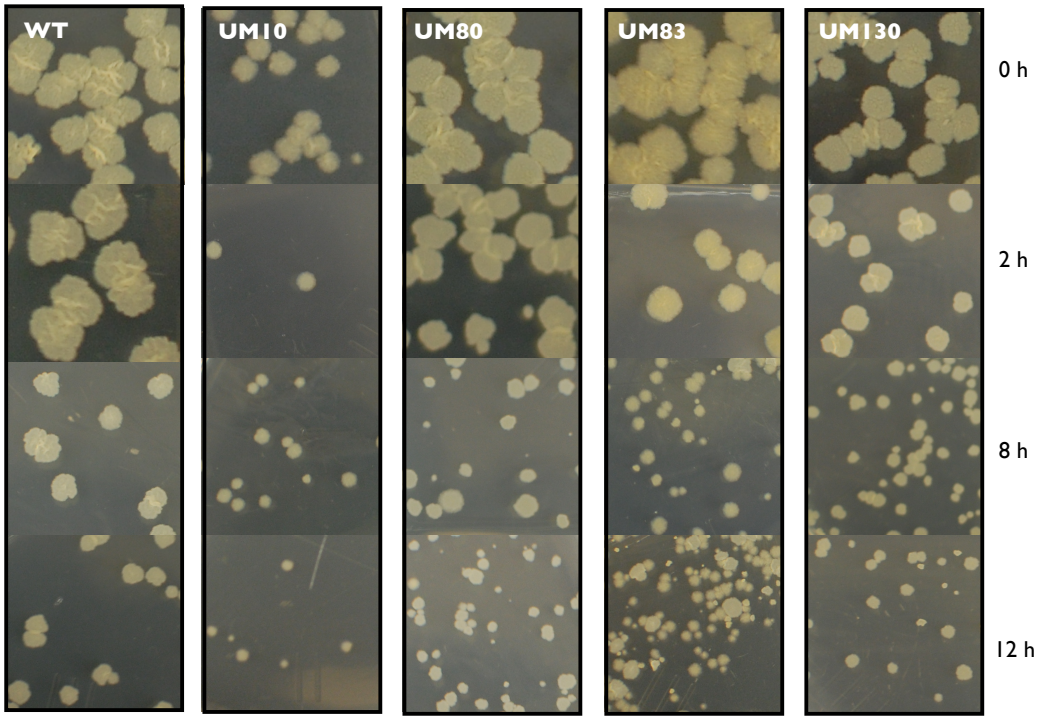



Figure 3.3: (Growth curve ($OD_{600}$) of WT and UM mutants with and without bialaphos.

### 3.4.4 Increased SDS sensitivity was shown in all UM mutants

The objective of this study was to investigate the effect of mutation on integrity of cell membrane. Because transposon gene was inserted into the genes which encode for transmembrane proteins and ATPases, we hypothesized that a formation of a transporter complex with the cryptic domain might disrupt a proper structure of cell membrane leading to a permeability defect. We used SDS, an anionic detergent, to test this hypothesis as SDS has been used to study the cell wall/membrane integrity of membrane protein mutants [5, 100]. Cultures were grown to a mid-log phase and SDS at final concentration of 0.1% was added to the cultures. Cells were plated on the 7H10 agar immediately after SDS was added. This point was considered as 0 h and the CFU/ml at this time point was used as 100% survival. The 7H9 medium was used as an inoculum solution to minimize unnecessary loss of viability during plating process since it took up to 1.5 h to complete plating at each time point. Drastic response was observed from UM10 (Fig. 3.4A). At 0 h, morphology of the mutant did not change significantly and possessed wrinkly appearance, the signature colony morphology of mycobacteria. However, the size of individual colonies was much reduced, at least by 50%. This was very surprising since the cells of the UM10 culture were plated within 30 min after sampled and the number of CFU/ml was comparable to that of WT and the other mutants. At 2 h, the colonies showed much severe morphological defects. These colonies were needle-point size with fuzzy or protruding edge, sheet-like (flat-center), and loss of pigmentation. With other mutants, these morphological defects were not apparent at the early incubation periods and only started to be observed at 8 h. With WT, these defects were also seen at 8 h but with less severity.

A




B

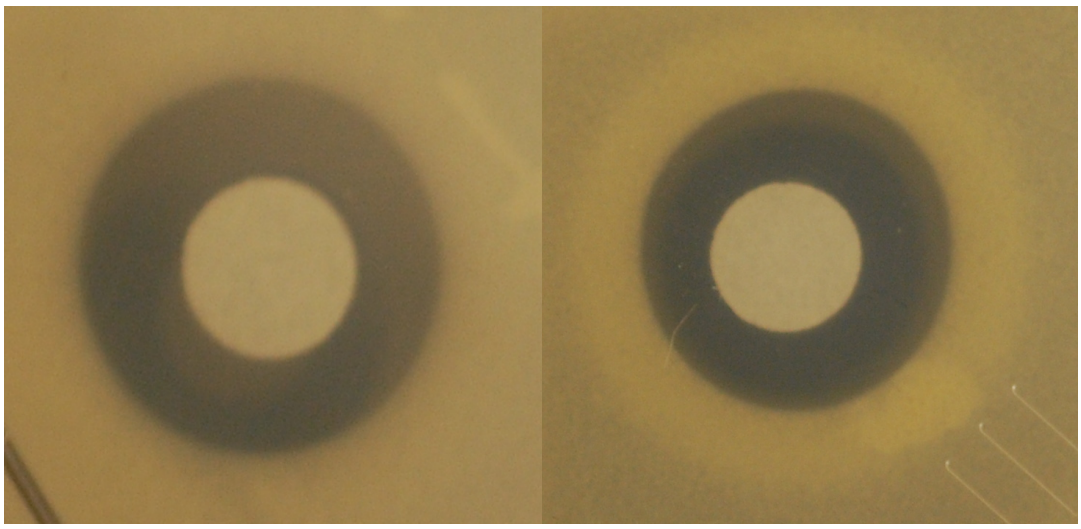


Figure 3.4: Effects of SDS appear as colony morphology. (A) Pictures of colonies of each strain were taken at time points 0, 2, 8, and 12 h and compared side by side. As seen at 0 h, a strong effect of SDS on was already observed on UM10 appeared as reduction of colony size. (B) Zone of inhibition of UM10 (right) with 5 $\mu$l of 10 % SDS concentration was shown. A formation of thicker culture around the edge of the zone was observed. Typical appearance of zone of inhibition is shown on the left.

Severe response to SDS stress was seen from all UM mutants in the SDS sensitivity assay (Table 3.2 and Fig. 3.5). At 2 h post treatment, the survival of UM mutants was drastically reduced to below 3 %. At the same time point, 50 % of WT still survived. Much severe reduction of population was seen in UM10 and UM80 mutants that the percentage survival dropped to 0.015 % and 0.08 %, respectively. At 4 h post treatment, all the mutants already showed below 1 % survival whereas 35 % of the WT population survived at the time point. The percentage survival of WT was comparable to to that described in White *et al*. [100]. Interestingly, no significant changes in the measure of $OD_{600}$ in all the cultures during the SDS sensitivity assay (data not shown). This indicates that $OD_{600}$ does not change with relative to a live population in the culture within the short period of time, and it cannot be used as a measurement of survival in broth.

$_D$-cycloserine is inhibitors of core enzymes for synthesis of peptidoglycan in cell wall, alanine racemase and $_D$-alanine-$_D$-alanine ligase. No zone of inhibition was observed until the concentration reached 10 mg/ml. Diameters of zone of inhibition from the mutants showed no significant difference compared to that of WT (data not shown). The result from the $_D$-cycloserine disc assay suggests that the mutants maintain the same integrity of cell wall structures as WT. When sensitivity to SDS was tested on plates, higher concentration than the broth as-

| Time (hr) | Strains (% survival) | | | | |
|---|---|---|---|---|---|
| | WT | UM10 | UM80 | UM83 | UM130 |
| 0 | 100 | 100 | 100 | 100 | 100 |
| 2 | 50 | .015 | .08 | 2.6 | 1.4 |
| 4 | 35 | .002 | .0175 | .55 | .45 |
| 8 | 16 | .0018 | .0036 | .072 | .0072 |
| 12 | 4.3 | .00008 | .001 | .028 | .003 |

Table 3.2: % survival of cells in effects of 0.1% SDS compared to WT.

say was required to observe the effect. Zone of inhibition was slightly observed at 1 % SDS, and it was clearly shown at 10 %. Interestingly, no significant difference in the diameters of zone of inhibition in the mutants compared to that of WT, however, UM10 mutant shows interesting appearance: accumulation of growth around the edge of the zone of inhibition, observed as a thicker lawn of bacteria, making a ring appearance (Fig. 3.6B). This apparent phenotype is evident when compared to the zones of inhibition from other UM mutants and WT.

Taken together, these results showed that disruption of the genes involved in an ABC uptake transporter complex cause adverse effects on the integrity of cell membrane in the UM mutants. Flores-Valdez *et al*. showed that phthiocerol dimycocerosate (PDIM) and Myc ratio was deregulated in their *opp* mutant [32]. It is possible that the UM mutants also have some defects in cell wall.

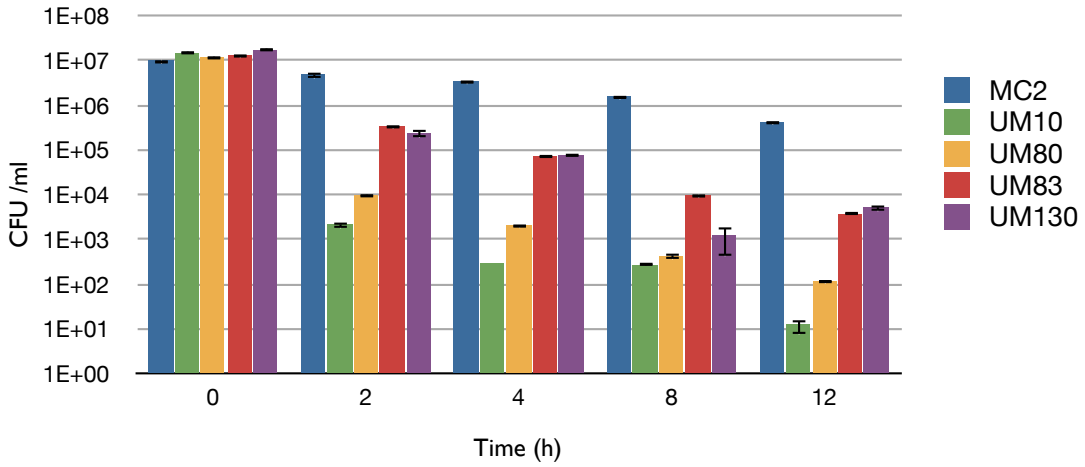



Figure 3.5: % survival of cells in effects of 0.1% SDS shown in histogram.

### 3.4.5 Uptake of biotinylated peptides was via different transporter(s)

Relative growths of the UM mutants on peptides and the corresponding biotinylated peptides were measured and compared (Fig. 3.6). In this experiment, we modified the scale and the growth condition. We used 1 ml medium in 24-well plates instead of a medium of 20 ml, the volume which is routinely used, because we only had a small quantity of the biotinylated peptides synthesized. Also, the plates were incubated without agitation to minimize contamination from neighboring wells. $OD_{600}$ was measured directly in the plates using BioTeck.

The rates of relative growth were measured by comparing the $OD_{600}$ of the cultures in the peptides (or biotinylated peptides) to those in the negative well. As shown in figure 3.6, all UM mutants grew much better when the biotinylated

peptides were used as a N-source. This increased in growth with the corresponding biotinylated peptides indicate that all UM mutants are capable of taking up the biotinylated peptides and utilizing them as N-source. When glycine and phenylalanine were conjugated with biotin the relative growths increased. We speculate that the peptides GF and FGG are better assimilated in the cells, however this needs to be experimentally confirmed.

In this experiment, we were able to show that all UM mutants were able to grow on the peptides as long as they were transported into the cells. Additionally, this indicates that the biotinylated peptides were taken up via the transport systems other than the ones disrupted in the mutants.

### 3.4.6 Peptides are competitive inhibitors of bialaphos

We were curious to see whether bialaphos toxicity could be mitigated when added together with other ABC transporter substrates such as peptides since bialaphos is known to enter the cells via Dpp and/or Opp transporters in other bacteria [62]. We tested several peptides ranging from di- to tripeptides with increased molar ratio to bialaphos on WT and measured growth as an indicator of how effective the peptide competed against bialaphos for the same ligand binding site. Figure 3.8 shows that diminished bialaphos toxicity in presence of peptide. Surprisingly, not only in the media containing tripeptide, Phe-Gly-Gly, but dipeptide, Gly-Phe or Ala-Ala, WT growth was rescued almost to the level in 7H9 media only. Higher molar ratio of peptide to bialaphos improved growth. At 1:1 ratio of peptide and bialaphos, the tripeptide Phe-Gly-Gly rescued WT growth better than the dipeptides but these dipeptides were able to rescue WT

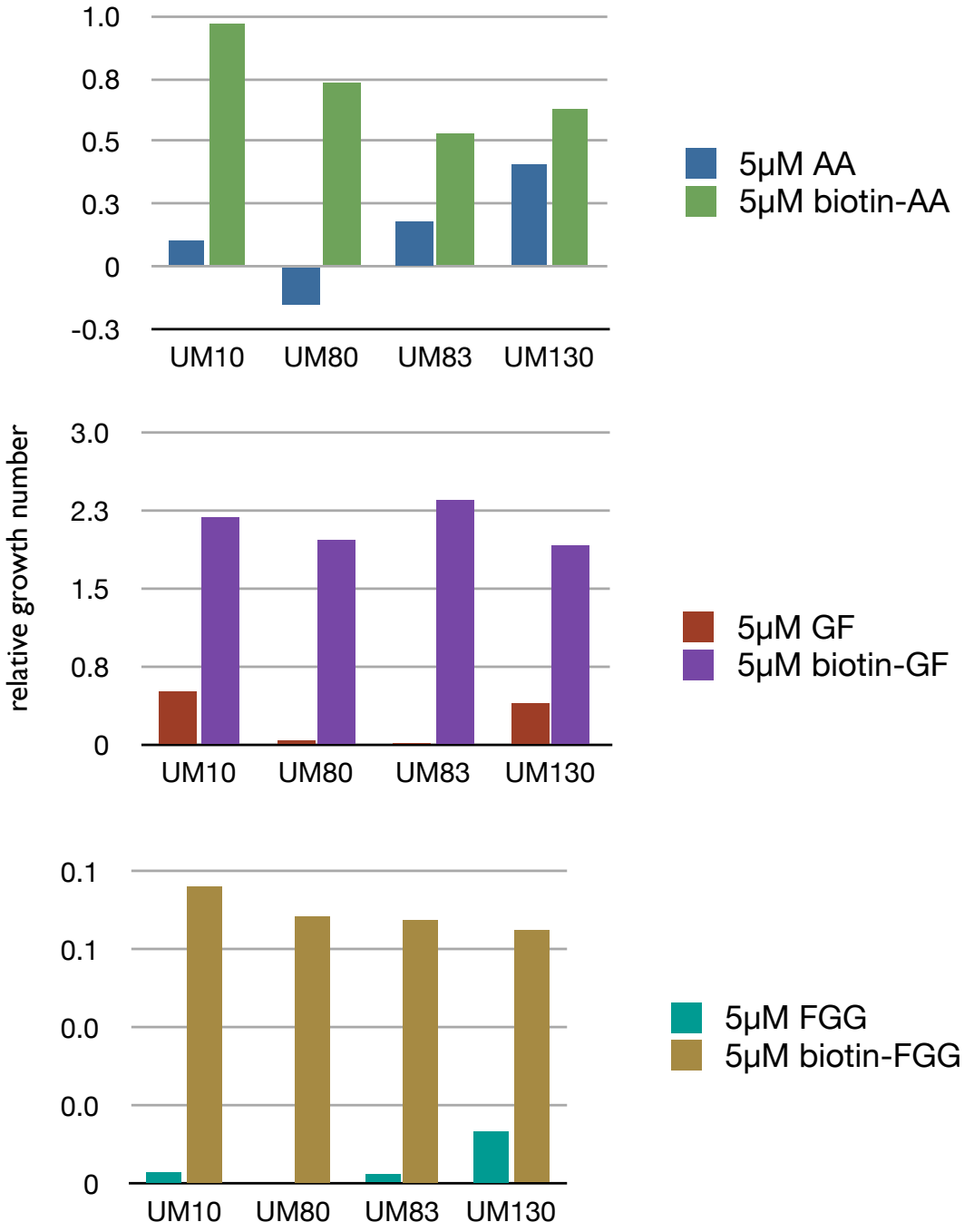



Figure 3.6: Relative growth of UM mutants with biotinylated peptides as a sole N-source in HB media. The number was determined relative to the corresponding negative control, which is only the HB without a N-source.

growth with increased molar ratio. This result suggests that bialaphos competes for the same transporter pathways as the dipeptides and oligopeptides tested. This indicates that unlike *Mtb*, both Dpp and Opp ATP transporters in *M. smegmatis* are involved in bialaphos uptake. Given the fact that all the UM mutants show resistance to bialaphos, it is likely that multiple transporters can mediate uptake of bialaphos in *M. smegmatis*.

When FDG was added as a competitor, the growth was rescued to the same extent as Gly-Phe at 1:1 molar ratio, but it did not improve even FDG was added 10 times more. This result suggests that FDG might be transported via same transport mechanism as bialaphos.

### 3.4.7 BIOLOG PM plates

BIOLOG PM plates were employed to obtain an array of nitrogen substrates, mainly peptides, taken up via the transporter systems disrupted by the UM mutants. Metabolism of the target substrates was extrapolated by measuring the reduction of tetrazolium salts to formazan at 590 nm. In BIOLOG PM plates, tetrazolium salts were added to wells to measure the rates of metabolism on individual nutrient sources. The tetrazolium salts were cleaved to soluble formazan by bacterial reductase or dehydrogenase. This by-product, formazan, emits a purple colour and can be measured at 590 nm. Thus, if the bacterial cells can utilize the substrate, the bacterial population increases which lead to an increase amount of reductase/dehydrogenase released into the culture supernatant to reduce the tetrazolium salts to formazan. This sequence of event causes more production of formazan, leading to a stronger purple colour on the

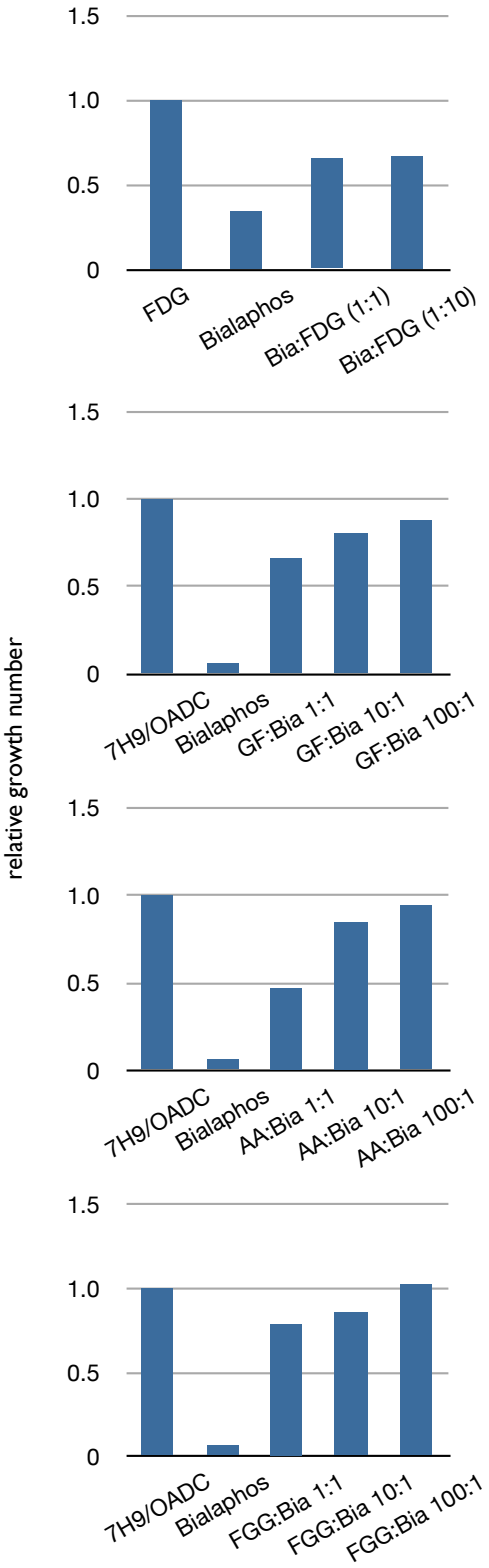

relative growth number
1.5
1.0
0.5
0
FDG
Bialaphos
Bia:FDG (1:1)
Bia:FDG (1:10)
7H9/OADC
Bialaphos
GF:Bia 1:1
GF:Bia 10:1
GF:Bia 100:1
AA:Bia 1:1
AA:Bia 10:1
AA:Bia 100:1
FGG:Bia 1:1
FGG:Bia 10:1
FGG:Bia 100:1

Figure 3.7: Competitive uptake of bialaphos and peptides was measured. The growth of WT in 7H9 media was used as 1. Affinity of peptides over bialaphos was determined by the growth over that of bialaphos alone.

wells.

For our purpose, we looked for peptides which were metabolized the least. Selection of the peptides was performed as follows: the metabolism of each substrate was normalized against the negative control in whose well had no source of nitrogen but the same amount of the bacterial cell was inoculated as the rest of the wells in the plate. The values of these normalized substrate wells of the UM mutants were subtracted from the corresponding wells of WT. Table 3.3 shows the 30 N-sources, arranged in the order from the least value, from each UM mutant. Comparing the first 5 N-sources from the UM mutants, we noticed that peptides containing amino acids His, Phe, or Trp, were the least metabolized, suggesting that these peptides are preferentially taken up by the transporters disrupted in the UM mutants. The similarity between these peptides are that all possess amino acids with a bulky side chain. Tyrosine was not included in the list because WT was also unable to metabolize peptides containing tyrosine. Growth of the UM mutants on some of the peptides listed in the first 5 in the Table 3 were tested in a large scale and the result was described in detail in the next section.

Preferential C-source of the UM mutants were not tested using the BIOLOG PM assay because the PM plates do not provide a diverse sample of di- and oligosaccharides.

| No. | UM10 | | UM80 | | UM83 | | UM130 | |
|---|---|---|---|---|---|---|---|---|
| | Peptides | Relative Value | Peptides | Relative Value | Peptides | Relative Value | Peptides | Relative Value |
| 1 | Ala-His | -0.44 | Phe-Phe | -0.256 | Leu-His | -0.398 | Gly-His | -0.383 |
| 2 | Leu-His | -0.403 | Trp-Tyr | -0.226 | Gly-His | -0.368 | Leu-His | -0.311 |
| 3 | Pro-Phe | -0.348 | Ala-His | -0.211 | Phe-Phe | -0.301 | Ala-His | -0.242 |
| 4 | Gly-His | -0.33 | Trp-Glu | -0.207 | Trp-Phe | -0.3 | Glu-Trp | -0.235 |
| 5 | Phe-Phe | -0.297 | Trp-Gly | -0.204 | Arg-Ser | -0.284 | Trp-Arg | -0.22 |
| 6 | Trp-Phe | -0.282 | Trp-Phe | -0.195 | Pro-Trp | -0.276 | Ser-Phe | -0.209 |
| 7 | Ala-Thr | -0.224 | Pro-Trp | -0.147 | Arg-Trp | -0.248 | Trp-Lys | -0.194 |
| 8 | Gly-Asn | -0.219 | Trp-Asp | -0.134 | Glu-Asp | -0.236 | Glu-Asp | -0.179 |
| 9 | Trp-Gly | -0.219 | Trp-Ser | -0.123 | Glu-Trp | -0.213 | Pro-Pro | -0.161 |
| 10 | Pro-Trp | -0.219 | Leu-His | -0.12 | Leu-Leu-Leu | -0.205 | Gly-D-Val | -0.155 |
| 11 | Glu-Asp | -0.198 | Phe-Trp | -0.111 | Trp-Val | -0.178 | Arg-Ser | -0.138 |
| 12 | Arg-Ser | -0.171 | Gly-D-Val | -0.1 | Trp-Tyr | -0.173 | Leu-D-Leu | -0.135 |
| 13 | Trp-Glu | -0.167 | Pro-Phe | -0.088 | Trp-Leu | -0.17 | Asp-Phe | -0.128 |
| 14 | Glu-Trp | -0.164 | Leu-Leu-Leu | -0.087 | Pro-Phe | -0.169 | D-Ala-D-Ala | -0.119 |
| 15 | Trp-Leu | -0.164 | Pro-Hyp | -0.079 | Trp-Gly | -0.158 | Pro-Trp | -0.115 |
| 16 | Leu-Leu-Leu | -0.164 | Glu-Trp | -0.077 | Trp-Glu | -0.155 | Leu-Leu-Leu | -0.11 |
| 17 | Leu-D-Leu | -0.149 | Trp-Leu | -0.076 | Gly-D-Ala | -0.153 | D-Leu-D-Leu | -0.106 |
| 18 | D-Ala-Gly | -0.148 | Glu-Asp | -0.071 | D-Ala-Gly | -0.151 | D-Leu-Tyr | -0.106 |
| 19 | Trp-Ala | -0.144 | Gly-D-Thr | -0.071 | Leu-D-Leu | -0.145 | Gly-D-Asp | -0.104 |
| 20 | Gly-D-Val | -0.138 | Trp-Ala | -0.061 | Gly-D-Val | -0.142 | Asp-Glu | -0.1 |
| 21 | Phe-Trp | -0.132 | Gly-D-Ala | -0.045 | gamma-Glu-Gly | -0.14 | gamma-Glu-Gly | -0.094 |
| 22 | Nitrate | -0.13 | D-Ala-D-Ala | -0.043 | Phe-Trp | -0.133 | Gly-D-Thr | -0.08 |
| 23 | Trp-Asp | -0.122 | Trp-Val | -0.041 | His-Asp | -0.131 | Trp-Val | -0.075 |
| 24 | Arg-Trp | -0.115 | Gly-D-Asp | -0.036 | D-Ala-D-Ala | -0.131 | D-Ala-Gly | -0.075 |
| 25 | Tyr-Ile | -0.113 | Gly-His | -0.032 | Gly-D-Thr | -0.13 | Gly-D-ser | -0.07 |
| 26 | Trp-Tyr | -0.108 | Leu-D-Leu | -0.03 | Tyr-Ile | -0.114 | Asp-Leu | -0.055 |
| 27 | D-Ala-D-Ala | -0.103 | Trp-Lys | -0.028 | Asp-Leu | -0.112 | Phe-Ser | -0.05 |
| 28 | Gly-D-Thr | -0.099 | D-Leu-D-Leu | -0.027 | D-Leu-D-Leu | -0.102 | Tyr-Ile | -0.046 |
| 29 | Pro-Hyp | -0.096 | Gly-D-ser | -0.018 | Asp-Phe | -0.101 | Gly-D-Ala | -0.045 |
| 30 | Trp-Arg | -0.086 | D-Leu-Tyr | -0.013 | Trp-Asp | -0.092 | Asp-Asp | -0.045 |

Table 3.3: The growth of UM mutants was measured and determined as described in text. The metabolic value of individual UM mutants was normalized against the negative control and subtracted from that of WT. The 30 peptides in which UM mutants showed the least growth compared to WT are listed.

### 3.4.8 The mutants show similar phenotype in uptake defects

Usually substrate uptake study uses radiolabeled substrate and purified SBP and measures the kinetics of substrate affinity [28, 24, 35]. Alternatively, incorporation of radiolabeled substrate into biomass could be used as a measure of uptake transport [24, 46]. In this case, uptake of the radiolabeled substrate was determined by measuring radioactivity in the biomass and calibrated with cell numbers at $OD_{600}$. In our case, we simply add substrates as either a N-source, in case of peptides, or a C-source, in case of carbohydrates, in the HB medium and used the growth rate at $OD_{600}$ as an indicator of uptake. This strategy works because the only defect in the UM mutants is in the transporter system and no inherent defect in growth was detected from any of the UM mutants growing in a rich medium such as the 7H9 medium. Reduced growth was used as an indicator of the substrate preference by a specific UM mutant. We decided to use this approach because not all the peptides and carbohydrates were available in radiolabeled form and we wanted to standardize the assay so all the data were comparable.

#### Peptide transport

Peptide uptake assay was performed by measuring growth in a medium containing a target peptide as a sole N-source (Fig. 3.8). Ala-Ala, Ala-Glu, Ala-His, Phe-Phe, and Phe-Val peptides were chosen because it was described in the previous study [38] that *Mycobacterium bovis* BCG was not able to utilize them as a sole N-source. We were curious to see whether this would hold true in *M. smegmatis*. Other peptides chosen for this experiment were either a preferred substrate in the BIOLOG PM assay or to expand the list of test peptides.

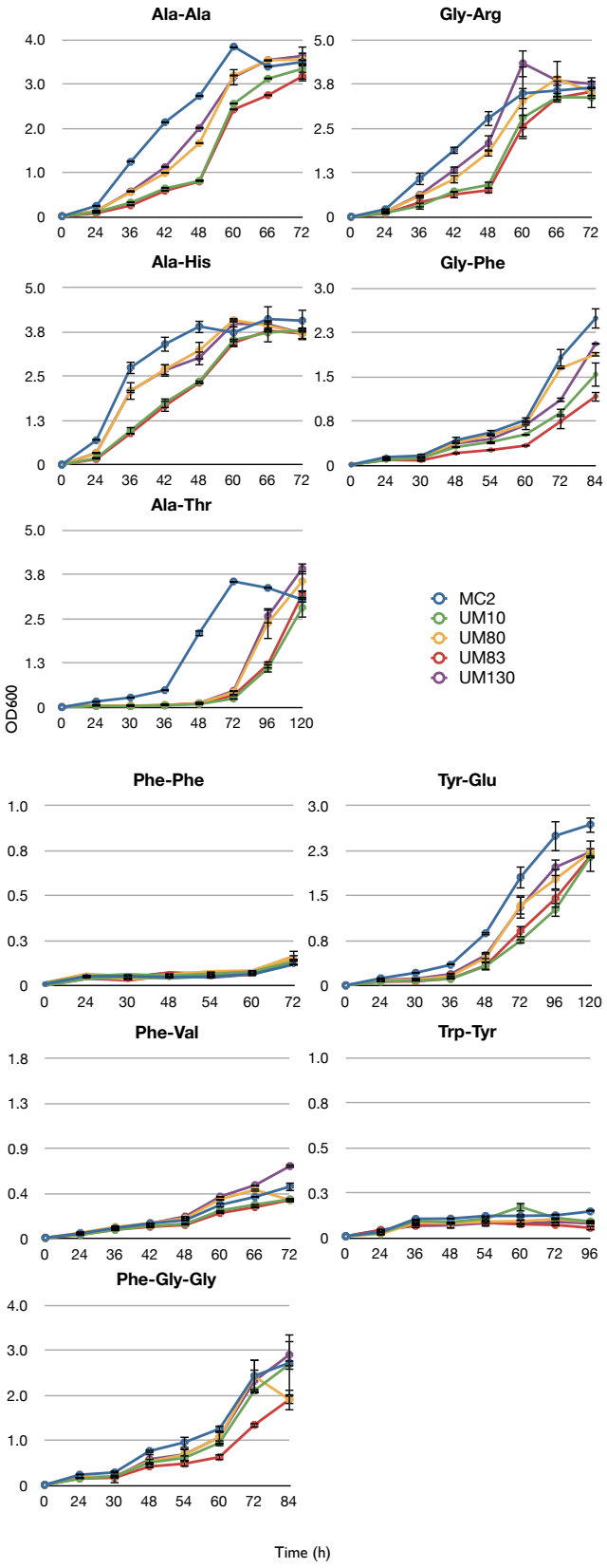

Ala-Ala
Gly-Arg
Ala-His
Gly-Phe
Ala-Thr
MC2
UM10
UM80
UM83
UM130
OD600
Phe-Phe
Tyr-Glu
Phe-Val
Trp-Tyr
Phe-Gly-Gly
Time (h)

Figure 3.8: Substrate specificity of mutants measured by a means of growth at $OD_{600}$. Growth of the UM mutants and WT was compared with peptides as a N-source.

In the hydrophobic and bulky dipeptides (Phe-Phe, Phe-Val, Trp-Tyr) media, WT as well as the UM mutants did not grow. This result demonstrates that *M. smegmatis* is not able to transport these peptides. An alternative explanation is that *M. smegmatis* lacks extracellular peptidases to break the dipeptides into amino acids to transport. In the BIOLOG PM assay (Table 3.3), Phe-Phe and Trp-Tyr peptides were the substrates UM mutants showed less growth compared to WT. However, when carefully looked at the growth, WT also showed a negative growth, therefore, therefore subtraction of the negative values of the UM mutants from the negative value of WT led to a greater difference in the data (data not shown).

Growths of UM mutants and WT on the Gly-Phe and Phe-Gly-Gly peptides suggest that the structure of phenylalanine conjugate is critical for a successful transport. All the cultures were able to grow in the media containing these peptides despite of an extensive lag period. This is also applied for the Tyr-Glu peptides.

In all the peptides tested, UM mutants and WT grew at relatively faster rates with the peptides conjugated with alanine (Fig. 3.8) except the Ala-Thr peptides. Lag periods of all the cultures were the longest (36 h for WT and 72 h for UM mutants) in all the peptides tested. It is very interesting to compare the lag periods of the cultures in the Ala-Thr peptides to that in other alanine conjugated peptides. It is very evident that addition of threonine to alanine drastically slowed the growth levels.

We only tested a few peptides from the list of the BIOLOG PM assay in this study, however we noticed that the substrate specificities observed in the BIOLOG PM assay did not correlate to what we observed in this study. One possible explanation could be that we were much prone to introduce errors, i.e., false positive or negative results, in the BIOLOG PM assay when we only used the reduction of tetrazolium salts to formazan as an indicator of growth. The manufacturer recommends us to use the OmniLog software which can trace growths on individual substrate (well) every hour and provide a metabolic curve on the substrate.

In the peptide growth assay, all UM mutants grew in media containing the peptides which WT also grew, but with extensive lag periods. The lag periods might suggests that the growth rates of the UM mutants are strongly affected. Because mutation occurred in the transporter genes, it is very likely that transport of the peptides was affected in the UM mutants, observed as the extensive lag periods. Recovery of the growth of the UM mutants at later time point indicates that the peptides were transported by alternative mechanisms yet to be determined.

**Carbohydrate transport**

All the mutants were able to grow as well as WT in media containing fructose or glucose as a sole C-source (Fig. 3.9), suggesting that the transporters encoded in the UM mutants are not involved in uptake of these carbohydrates. All the cultures showed very smooth homogenized cultures throughout their growth. None of the UM mutants and WT was able to grow on maltodextrin. Other disaccharides, maltose and melibiose, were also tested in this study but none

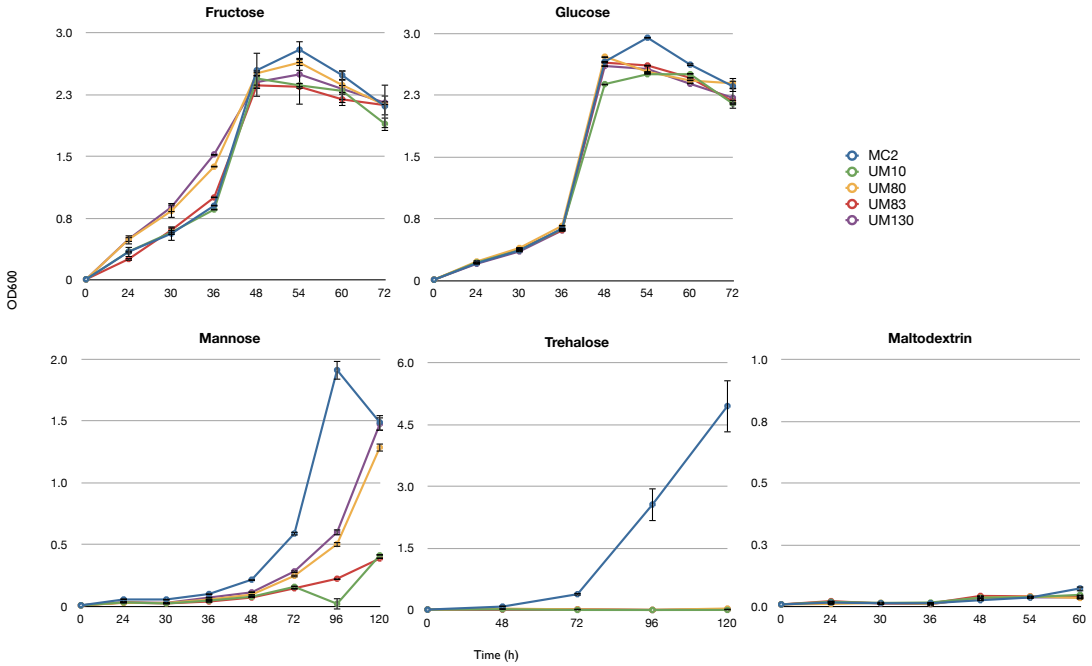



Figure 3.9: Substrate specificity of mutants measured by a means of growth at $OD_{600}$. Comparison of growth of UM mutants and WT with various carbohydrates a C-source.

including WT showed growth (data not shown).

Growth defect was observed in the UM mutants when mannose or trehalose was added as a sole C-source with much severe growth defect observed with trehalose (Fig. 3.10). Both WT and the UM mutants showed extensive lag periods in the trehalose medium but the lag period was less extent in WT compared to that in the UM mutants. We almost missed such the extensive delay of growth from the UM mutants because when WT reached the stationary phase, only a little growth was observed in the UM cultures.

We noticed that the growth pattern of the UM mutants with trehalose, after the extensive lag period, was similar to WT and we wondered whether their growth was adaptive. To test this hypothesis, we grew the cultures to a mid-log

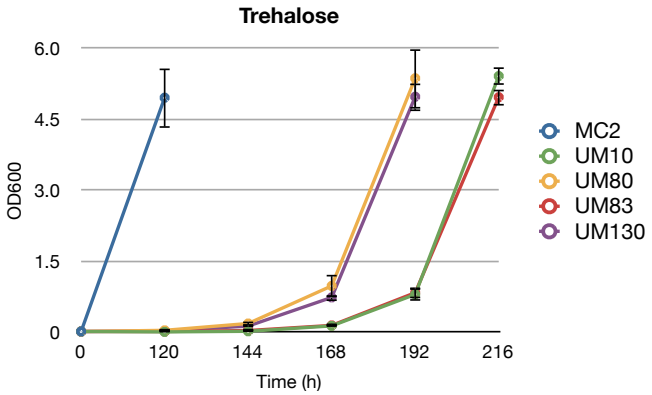



Figure 3.10: Substrate specificity of mutants measured by a means of growth at $OD_{600}$. Measurement of growth of UM mutants grown on 10 mM trehalose. No UM mutants were able to grow at the same growth rate as that of WT. Growth was observed at an extended period of time after the growth of WT reached the stationary phase.

phase in a trehalose medium, transferred to a fresh trehalose medium and monitored their growth, particularly length of the lag period. The growth profile of WT did not change in the second passage. It still showed the same length of lag period as observed in the first passage (Fig. 3.11). Contrary to WT, shorter lag period was observed from UM80, 83, and 130 in the second passage. Particularly with UM80 and 130, no difference was observed in their growth compared to WT. The length of lag period in UM83 was also shortened by 3 days. The shorter lag period seen in these mutants in the fresh trehalose medium after they were pre-grown in a medium containing trehalose as a sole C-source indicates the presence of alternative transport mechanism for uptake of trehalose. No change in the length of lag period was observed from the UM10 culture in the second passage.

In both peptide and carbohydrate growth assays, we eliminated the possibility that the mutants have adverse defects in metabolic pathways due to the

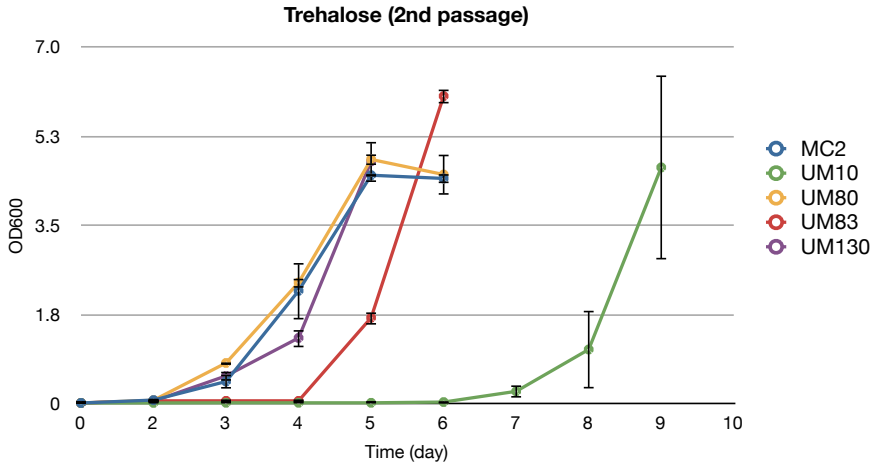



Figure 3.11: Determine adaptation of the UM mutants on trehalose. Cultures grown to a mid-log phase in the trehalose medium were transferred to a fresh trehalose medium and monitored for the length of their lag period.

transposon insertion since the UM mutants are able to grow in a complex rich medium like 7H9 media comparable to that of WT. Instead, the phenotypes are likely due to replacement of regulatory pathways and the length of delays implies that each UM mutant is under different regulation and/or differences in their substrate specificities.

Another notable phenotype we observed was the paring of two UM mutants in their growth trend. Pairs of UM10 and UM83, and UM80 and UM130 almost always followed the same kinetics of growth on all the substrates tested in this study. We first thought that the function of genes mutated by transposon insertion might be the cause of this phenotype. However, this was not the case since the mutants, UM83 and UM130 in which a permease gene was affected (Chapter 2), did not exhibit resembling growth trends.

## 3.5 Discussion

In this study, we have characterized 4 new transporters in *M. smegmatis* for their possible biological roles in nutrient acquisition. 3 transporters, namely MSMEG_0553-0556 (UM80), MSMEG_4101-4098 (UM130), and MSMEG_4995-4999 (UM83), were previously annotated as sugar, dipeptide, and oligopeptide transporters, respectively. However, contradictory to their annotations, our results (Figs. 3.8 to 3.10) showed that these transporter mutants showed growth defects in both peptides and carbohydrates. The growth defects across different families of substrates suggest that these transporters are capable of transporting a broad range of substrates. It has been previously reported that a single transporter is able to transport multiple substrates but only within the same category. For example, AgaRZSXPAEFGK-MspB in *M. smegmatis* is annotated to transport both $\alpha$-galactosides and melibiose [60]. Also, some peptide transporters have shown to transport peptides with variable lengths [29].

Instead, the broad substrate preference observed in the UM mutants, resembled some drug efflux transporters. The multidrug transporter complex AcrAB-TolC, MdfA, and EmrE of *E.coli* allows the bacterium to extrude a wide range of drugs, which share no resemblance in structure, size or ionic charges [89]. The same phenomenon has also been observed in mammalian cells during treatments of cancer that multidrug resistance has been a major impediment in curing cancer [37]. We hypothesized that the growth defects observed from the UM mutants are a result of decreased or elimination of the substrate uptake due to mutation in the transporter genes.

Among the substrates tested, all UM mutants exhibited the most severe

growth defect with trehalose. Trehalose is a disaccharide formed from two glucose units joined by a 1-1 alpha bond. It has been documented that *M. smegmatis* and *Mtb* can utilize trehalose as a sole C-source [92, 46]. Trehalose is an important sugar comprising a part of mycobacterial cell wall and is recognized as an inflammatory molecule by host cells. The study by Kalscheuer *et al*. showed a recycling mechanism of released trehalose into *Mtb* via trehalose transporter, whose mechanism contributes to the survival of *Mtb* during infection. In our study, we only characterized trehalose uptake efficiency of these transporters by measuring the growth using the UM mutants. Further characterization of substrate specificity and affinity of these transporters to trehalose may help elucidate how trehalose transport works in *Mtb*. Furthermore, characterizing mechanisms of these transporters would help design drug compounds which can be transported via the *Mtb* trehalose transporter.

Earlier, we hypothesized that the growth defects of the UM mutants are due to reduced uptake systems by mutations in the transporter genes. However, this phenotype does not explain the extended lag periods. RT-PCR on adjacent genes in the transporter loci, affected by transposon insertion, showed that expression of the adjacent genes was not affected (data not shown).

Adaptation to trehalose by the UM mutants (Fig. 3.11, particularly with the ABC importer mutants (UM80, 83 and 130), suggests that the length of lag period can be explained by the time it takes for the cells to induce transcription of alternative transporter. To prove this hypothesis, changes in the transcriptional profile, particularly upregulation of transporter genes, during growth from the lag phase to the log phase need to be elucidated.

The Opp transporters in *B. subtilis* and *Enterococcus faecalis* have a role to im-

port signaling peptides. The signaling peptides, Phr, in *B. subtilis* is released to extracellular medium and enters into the cell via an Opp transporter to signal sporulation. In *E. faecalis*, the mating pheromones enter into the cell via an Opp transporter and start the signaling pathway to mediate conjugation with another cell to share DNA, mainly virulence factors [52]. *M. smegmatis* is not known to sporulate nor conjugate, however, we can hypothesize that the transporters characterized in this chapter may play a role in uptake of signaling molecules to regulate expression of nutrient transports and to shift metabolic pathways. This idea came from a study by Douka and his colleagues. [30]. In their study, the authors isolated a mutant of *Zymomonas mobilis* which showed a prolonged lag phase for growth in medium at high glucose concentration. However, when the mutant was transferred to a fresh medium which had been conditioned by WT grown to early log phase and removed, the lag phase disappeared. The authors claimed that the growth stimulating molecules were released and present in the preconditioned medium, thus the mutant was able to grow like WT. Based on these studies, it is possible to explain the extended lag phases in the mutants by lack of uptake mechanisms of signaling peptides to stimulate upregulation of the downstream metabolic cascade. UM10 (*MSMEG_4363* mutant) is the only mutant which is not an ABC importer in this study. It is an orphan gene, whose mutation shows the worst growth defects or longest extensive lag periods(Figs. 3.8 to 3.10). Its deduced amino acid sequence possesses the NTF2-like motif. The NTF2 protein was originally isolated in *Xenopus* ovary cells as a soluble protein essential for translocation of the Ran-GDP complex into nucleus [58] and its roles as a nuclear importer of the Ran-GDP complex have been extensively documented [53, 47, 102]. In prokaryotes, its function is barely characterized. Only protein with the NTF2-like mo-

tif characterized in prokaryotes is VirB8 proteins of *Agrobacterium tumerfaciens* and *Brucella suis*, and VirB8 like protein, ComB8, in *Helicobacter pylori* [90, 3]. VirB8 is an essential assembly component of type IV secretion systems (T4SSs). T4SSs are exclusively found in Gram negative bacteria and play a crucial role in translocation of substrates across the cell envelope. Importantly, this system allows pathogens such as *A. tumerfaciens* to directly insert virulence factors into a host cell. VirB8 is a core component of the complex: it initiates an assembly of other components of T4SS and interactions with VirB8 directs the other components to a proper assembly site [8]. No such function of proteins possessing NTF2-like motif has been demonstrated in Gram-positive bacteria, but it is exciting to speculate that, based on the similarity in secondary structures of Vir8 of *A. tumerfaciens* and *B.suis* to MSMEG_4363, and the transport defect phenotype of MSMEG_4363, MSMEG_4363 could also be a part of a transport system yet to be discovered.

In this study, all UM mutants show the least growth with trehalose but also showed some degrees of growth defects with other nutrient sources tested. The outcome suggests that our screening method with FDG might identify transporters which can transport a broad range of substrates. None of these transporters was previously characterized to play a role in nutrient acquisition. It is exciting to unravel possible pathways of nutrients in *M. smegmatis* because this data can be applied to study those in *Mtb*. To further understand substrate specificities of these transporters, a clean KO of the *sbp* genes need to be constructed and tested on the affect on the ability to transfer the substrates tested here.

CHAPTER 4

# DETERMINING UPTAKE OF BIALAPHOS VIA MSMEG0553, THE SUBSTRATE BINDING PROTEIN OF THE UM80 LOCUS

## 4.1 Abstract

Substrate binding protein (SBP) is an extracellular protein, which exhibits high affinity to target substrates and modulates substrate specificities of a transporter complex. We characterized 4 uptake mutant (UM) strains in the previous chapter. Out of which, 3 UM strains have a transposon insertion in the ABC transporter superfamily but there was no transposon insertion in any of the *sbp* genes, making difficult to delineate the role(s) in those transporters. Here, we constructed a *sbp* knockout (KO) strain, *MSMEG_0553* KO in *Mycobacterium smegmatis* $MC^2$ 155. With the KO strain, we showed that MSMEG_0553 functions as an entry site for a tripeptide toxic compound, bialaphos. WT *M. smegmatis* was not able to grow in the 7H9 medium in presence of bialaphos, but the *MSMEG_0553* KO strain grew well without any effect of bialaphos toxicity.

## 4.2 Introduction

Transport of nutrients across the cell membranes is crucial for survival in all organisms. The binding protein-dependent ABC transporters are the unique members of the ABC transporter family and have been only found in bacteria and archaea up to date. Transporters in this superfamily function merely to import solutes into the cells and are distinguished from other ABC transporters

by the additional extracellular receptor domain, substrate binding protein (SBP) [41, 25]. This accessory protein functions as a high affinity receptor in the transporter complex, therefore, the binding protein-dependent ABC transporter has a different transport mechanism: (i) a target substrate first binds to the SBP; (ii) the liganded SBP docks onto its associated TMDs and this signals the NBDs to recruit ATP; (iii) binding of ATP allows a conformational change of TMDs, transferring the substrate from the SBP to the binding site in the TMDs; and (iv) hydrolysis of ATP causes dissociation of the NBD dimers, leading to the conformational change from an outward- to an inward-facing forms and release of the substrate into the cell [10].

In this study, we tried to answer a question, whether SBP of the transporter mutants characterized in the previous chapter modulates substrate specificities. Because SBP is located in the periplasmic space, it is the least perturbed subunit in the ABC importer complex for cell membrane. We chose *MSMEG_0553*, a *sbp* gene in a sugar ABC transport complex, to study its role on substrate specificity in the transport complex. If we observe a same transport phenotype from the SBP deletion mutant as we observed from its corresponding transposon mutant, it suggests that the phenotype observed in the transposon mutant is not due to perturbation of cell membrane.

## 4.3 Materials and Methods

### 4.3.1 Bacterial strains and plasmids

*M. smegmatis* $MC^2$ 155 was used to construct a *sbp* KO strain. To manipulate all the plasmic constructs, *Escherichia coli* TOP10 (Invitrogen, Carlsbad, CA, USA) was used. Zero Blunt® TOPO® PCR cloning kit (Invitrogen) was employed in the subcloning process. An *E. coli-mycobacteria* shuttle vector, pYUB854, was used to construct a flanking region to delete *MSMEG_0553* gene. All the strains and plasmids used in this study are listed in Table 4.1.

### 4.3.2 Growth conditions

Cultures of $MC^2$ 155 and its KO strain were maintained in the Middlebrook 7H9 medium (BD, Franklin Lakes, NJ, USA) supplemented with Middlebrook OADC (BD), 0.2% glycerol, and 0.05% Tween-80) or 7H10 agar (BD) (supplemented with Middlebrook OADC, 0.2% glycerol) at 37°C. When broth was used, the cultures were shaken at 100 rpm. All the *E. coli* strains used for preparations of the cloning fragments were maintained in Luria-Bertani (LB) broth (BD) or on LB agar (BD) plate at 37°C. When grown in broth, the *E. coli* strains were shaken at 200 or 250 rpm. When necessary, the cultures were supplemented with the appropriate antibiotics as listed in Table 4.1. Concentrations of hygromycin B (Roche, Indianapolis, IN, USA) and kanamycin (EMD Biosciences, Gibbstown, NJ, USA) used in $MC^2$ 155 were 150 $\mu$g/ml and 20 $\mu$g/ml, respectively. 50 $\mu$g/ml kanamycin or hygromycin B was used in all the *E. coli* strains throughout the study.

| Strains or plasmids | Descriptions | Sources or references |
|---|---|---|
| Strains | | |
| *Mycobacterium smegmatis* $MC^2$ 155 | Laboratory strain | ATCC 700084 |
| *Bluesmeg* | $MC^2$ 155 carrying pML803 $Hyg^r$ | This study |
| $MC^2$ 155/pJV53 | $MC^2$ 155 carrying pJV53 for recombineering of linear DNA into $MC^2$ 155 $Km^r$ | [96] |
| $MC^2$ 155 △ *MSMEG_0553* | *MSMEG_0553* knock-out $Hyg^r$ | This study |
| Plasmids | | |
| pML803 | a *E.coli*-mycobacterial shuttle vector carrying *lacZ*; $Hyg^r$ | a gift from Michael Niederweis |
| pCR4-TOPO | Cloning vector for PCR products; $Km^r$ | Invitrogen |
| pCR4-0553up | pCR4-TOPO containing PCR product of *MSMEG_0553* upstream flanking fragment to insert into pYUB854 | This study |
| pCR4-0553down | pCR4-TOPO containing PCR product of *MSMEG_0553* downstream flanking fragment to insert into pYUB854 | This study |
| pYUB854 | Mycobacterial cosmid possessing *res-hyg-res* gene cassette for generating gene deletions in mycobacteria; $Hyg^r$ | [7] |
| p0553 | pYUB854 containing the upper and downstream fragments of *MSMEG_0553* from pCR4-0553up and pCR4-0553down flanking *hyg* gene | This study |
| pVV16 | *E. coli* mycobacterial shuttle vector with expression under control of the *hsp60* promoter and encoding a His_*x*6 C-terminal tag, MCSs between the promoter site and His_*x*6 site $Km^r$ $Hyg^r$ | [45] |
| pSM0553 | pVV16 containing *MSMEG_0553* | This study |
| pJV53 | Recombineering vector, derivative of pLAM12 with Che9c 60 and 61 under control of the acetamidase promoter | [96] |

Table 4.1: All the strains and plasmids used in this study

### 4.3.3 Construction of a *M. smegmatis* KO strain

To determine whether a substrate binding protein (SBP) plays a role in substrate specificity, a *sbp* KO strain was constructed in $MC^2$ 155. The *sbp* gene in the UM80 mutant locus, *MSMEG_0553*, was deleted using a homologous recombination by the recombineering method described in van Kessel and Hatfull

(2008) [97]. Briefly, the upstream fragment of *MSMEG_0553*, including a part of 5′ end of *MSMEG_0553*, and the downstream fragment of *MSMEG_0553* were amplified using primer sets #300 (5- TCT AGA GAG GAC CCG CAG GCC GCC AGG GCC GG -3) and #301 (5- CTT AAG GTG ATC GCC GGC AGC TCG GAT CCG GCG -3), and #302 (5′- CCA TGG GCA GGC CGG TGC GCC GAA GCT GTT GC -3′) and #303 (5′- GGA TCC CTG TCG AAG ACC CCT GCC AGA CCT C -3′). The underlined sequences represent external restriction enzyme (RE) sites added to the primers. The PCR amplicon was gel purified using QIAquick Gel Extraction Kit (QIAGEN GmbH, Germany) and was subcloned into TOP10 using Zero Blunt® TOPO® PCR cloning kit by manufacturer's recommendation. Briefly, the ligated plasmid was mixed with TOP10 competent cells and incubated on ice for 45 min, followed by heat shock at 42°C for 30 sec and on ice for 60 sec. SOC medium (provided with the TOP10 competent cells kit by Invitrogen) was added to the vial, incubated at 37°C with agitation at 250 rpm for 1 h. 50 to 250 $\mu$l was spread on LB plates containing kanamycin and the LB plates were incubated at 37°C overnight. The plasmid isolated from the transformants were sent for sequencing to verify successful cloning. Once verified, the fragment was digested with appropriate restriction enzymes (RE) (New England Biolabs, Ipswich, MA, USA) and gel purified. Both upstream and downstream fragments were ligated into pYUB854 one at a time, in such a manner that the $hyg^r$ cassette was flanked between the fragments. The flanking region was amplified by PCR with primers #370 (5′- GGA TCG TGG TGA CGG CGT ACC GCG ACC G -3′) and #371 (5′- CGC ACC AGA GAC GGC GTG AAC GCG GGC G -3′) and the purified PCR amplicon was transformed into electrocompetent *M. smegmatis* cells carrying pJV53, a gift from Dr. Graham Hatfull at Pittsburgh Bacteriophage Institute and Department of Biological

Sciences, University of Pittsburgh, Pittsburgh, PA, USA. The plasmid pJV53 carries Che9c gp60 and gp61. Che9c gp60 and gp61 are homologs of RecE and RecT and encode for exonuclease and DNA-binding activities, respectively [96]. 0.2 % acetamide was added to induce expression of Che9c gp60 and gp61 whose proteins mediate a homolog recombination of the PCR amplicon into the designated region of the genome. The transformants were plated on 7H10 solid agar containing hygromycin B and kanamycin and incubated at 37°C for 4 days until colonies were visible. Genomic DNA from the colonies was isolated to verify the deletion of the *MSMEG_0553* gene. The method to extract genomic DNA was described in Chapter 2. Deletion of the gene in the KO strain was verified by PCR with the primers #219 (5- TAA CAT ATG GCC GAC GCG CCC GTA TCG GGC CGA TTC G -3) and #220 (5- ATT TAA GCT TTC ATG ACG AGG CGC TGT CCC AGC TCG CC -3) which amplify the inner sequence of *MSMEG_0553*.

### 4.3.4 Bialaphos toxicity test

Growth of the *MSMEG_0553* KO was tested in the 7H9 medium containing hygromycin B and kanamycin to determine whether the deletion of the gene causes *in vitro* growth defect. The *MSMEG_0553* KO and $MC^2$ 155 (WT) were grown overnight to a late log phase ($OD_{600}$ = 1.2 1.6) and inoculated at starting $OD_{600}$ of 0.01 into the 7H9 medium with appropriate antibiotics if necessary. The cultures were grown at 37°C and aliquot was periodically removed to measure their growths by $OD_{600}$. Bialaphos toxicity was performed on the *MSMEG_0553* KO as described in Chapter 3. Briefly, the culture was grown overnight in the 7H9 medium containing hygromycin B and kanamycin at 37°C overnight. Aliquot was taken into a fresh medium containing all the antibiotics

and 25 $\mu$g/ml bialaphos. The culture was grown at 37°C and the growth was measured at $OD_{600}$.

## 4.4 Results

### 4.4.1 Deletion of the *MSMEG_0553* gene

To study the function of MSMEG_0553 in the transport of bialaphos, we constructed a deletion strain of the *MSMEG_0553* gene. Deletion of the *MSMEG_0553* gene was verified by PCR amplification. PCR product was obtained from the parent strain, *M. smegmatis*, using the *MSMEG_0553* specific primer set. However, no PCR product was observed with the *MSMEG_0553* KO strain (Fig. 4.1), confirming absence of the *MSMEG_0553* gene in the KO strain. This result shows a successful deletion of the gene using the recombineering technique.

### 4.4.2 No growth defect was observed from the *MSMEG_0553* KO strain

Deletion of the *MSMEG_0553* gene in $MC^2$ 155 did not alter its growth in the 7H9 medium and the $\triangle$ *MSMEG_0553* strain was able to grow as good as WT 4.2. This result indicates that *MSMEG_0553* is not an essential gene for *in vitro* growth in 7H9 medium.

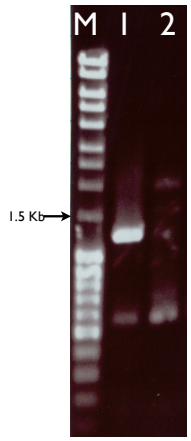



Figure 4.1: Gel electrophoresis of PCR products obtained from amplification of the *MSMEG_0553* gene. Lane M is a molecular weight (MW) marker. The arrow indicates MW of 1.5 Kb. Lane 1 shows the positive PCR amplification of the *MSMEG_0553* gene from the parent strain *M. smegmatis*. No PCR product was obtained from the genomic DNA of the *MSMEG_0553* KO strain (lane 2), indicating a successful deletion of the *MSMEG_0553* gene in the KO strain.

### 4.4.3 The *MSMEG_0553* KO strain is resistant to bialaphos toxicity

To confirm the function of MSMEG_0553 as SBP, toxicity of bialaphos to the KO strain was tested. If MSMEG_0553 does not play a role as an entry site of the substrate, the KO strain should be killed like WT. Figure 4.3 shows that the KO strain grew steadily in presence of bialaphos, whereas WT failed to grow. This result clearly illustrates that MSMEG_0553 functions as an entry site for bialaphos and deletion of *MSMEG_0553* results in resistant to bialaphos toxicity.

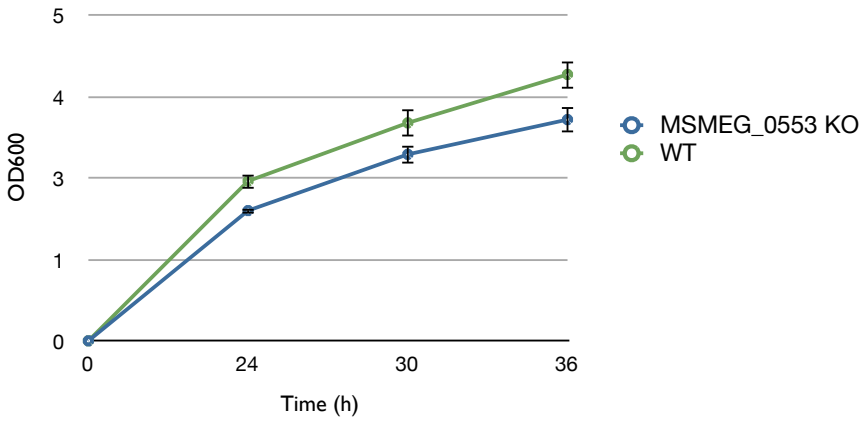



Figure 4.2: Growth of *MSMEG_0553* KO strain in 7H9 medium was compared to that of WT. The *MSMEG_0553* KO strain was able to grow comparable to WT up to 36 h, however no growth was observed after that time point.

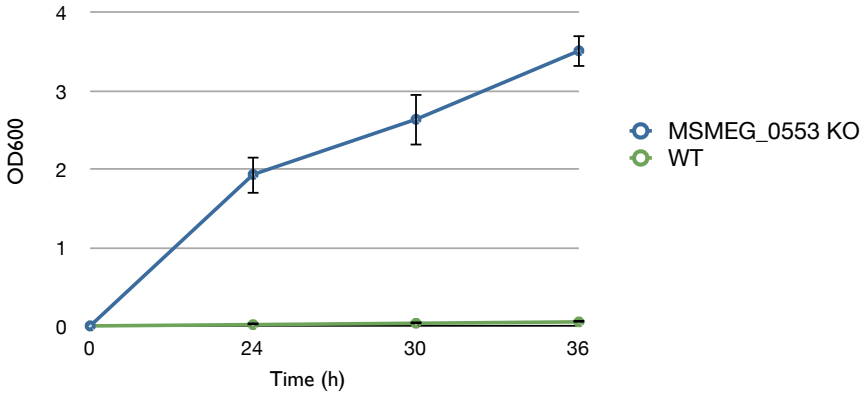



Figure 4.3: Bialaphos resistance of *MSMEG_0553* KO strain was determined as a measure of growth. There is no difference in growth with and without bialaphos in the *MSMEG_0553* KO cultures, providing an evidence that *MSMEG_0553* is used as an entry for bialaphos into the cell.

## 4.5 Discussion

In this study, we tried to understand whether substrate specificity is determined by the SBP or the TMD/NBD transport complex alone or both because all our transporter mutants had a transposon insertion in the genes encoding for permease or ATPase and none had a mutation in the *sbp* genes. Literatures show that both the SBP and the TMD/NBD transport complex are needed for the transport of a substrate [82, 40, 51]. We also wanted to elucidate whether the transport phenotype observed in the transporter mutants is due to defects in the transport process itself or perturbation of the cell membrane with the transposon mutated permease or ATPase.

*Lactococcus lactis* is an important bacterium in dairy industry. It has been extensively used in cheese and buttermilk production. Opp transporter in the bacterium plays the key mechanism for transporting essential nutrients such as carbohydrates and nitrogen in a form of peptides from milk. When OppA, the SBP protein of the Opp transporter is deleted, the mutant is completely abolished in the uptake of oligopeptides [66]. This study suggests that OppA dictates the uptake of substrate in the Opp transporter. However, Charbonnel *et al*. demonstrated that OppA from 6 different lactococcal strains can function exchangeably with the native Opp transporter complex in the $oppA_{L.l}$ deleted *L. lactis* strain [16]. The authors also showed that despite some amino acid sequence differences in the OppA from different lactococcal strains, all showed the identical substrate specificity in the $oppA_{L.l}$ deleted *L. lactis* strain. These results led the authors to present the two-filter theory that specificities to substrate are first filtered via binding affinity to the SBP then successful binding lets the substrate to be transferred to the transporter complex. In order to understand

whether the SBP in our transporter mutants plays a role in determining substrate specificity, we deleted one of the *sbp* genes, *MSMEG_0553*, in *M. smegmatis* and characterized it for bialaphos toxicity.

The *MSMEG_0553* KO strain showed a resistance against bialaphos demonstrated by the growth in the 7H9 medium containing bialpahos. This result suggests that loss of the *MSMEG_0553* gene allowed the mutant to grow under such the condition. It also supports the idea that binding to MSMEG_0553 is crucial for bialaphos to initiate its transport into the cell.

Previously, we showed that UM80, a transposon mutant of *MSMEG_0556* was resistant to bioalaphos toxicity (Chapter 3). MSMEG_0556 is an ABC transporter nucleotide-binding protein and functions to provide energy for a substrate transfer by hydrolysis of ATP inside the cell. Therefore, it does not directly involve in binding to the substrate. UM80 mutant, however, exhibited a resistance against bialaphos, demonstrating that MSMEG_0556 functions very tightly and specific to the rest of the transporter complex. The result from Chapter 3 and the result here combined together suggest that the uptake of substrates is mediated in a two-step system and supports the study of Charbonnel and her colleagues. Inactivation or deletion of one of them is, therefore, enough to lose its transport ability.

Here, we only tested the binding affinity of *MSMEG_0553* on bialaphos. OppA from different species have shown to bind an array of substrates of different size, shape, and charge [83, 66, 28, 16]. In fact, our uptake assay with the UM mutants in Chapter 3 demonstrate that these transporter mutants exhibit specificities to a broad range of substrates (Chapter 3). Unfortunately, all our UM mutants do not have mutation in *sbp*. Our future work will focus on making *M.*

*smegmatis* mutants with all the 3 *sbp* genes deleted one at a time and characterizing the roles in substrate specificity. Also, all the UM mutants are transposon mutants so the gene products might be produced but cryptic. Mutants with clean deletion of these transposon disrupted genes should be constructed and tested in the uptake assay whether we are able to observe the same results. At last, to study the role(s) of the transporter complexes, a uptake assay needs to be performed on the mutant with the entire locus deleted.

CHAPTER 5

# CONCLUDING REMARKS

The main goals of my Ph.D. dissertation are (i) to develop a screening method to identify transporter genes responsible for uptake of the target substrates and (ii), to characterize the phenotypes of the mutants selected in the screen.

## 5.1 Development of the screen method

The fluorogenic substrate, FDG, was used to identify active transporters. FDG was a very useful target substrate as it possesses two sugar moieties, di-$\beta$-$_D$-galactopyranoside, which are a substrate for $\beta$-galactosidase. By constructing a library of transposon mutants possessing a *lacZ* reporter plasmid, uptake of FDG was easily detected by fluorescence from the cells in flow cytometry. This screen is a negative genetic selection, in which mutants with reduced fluorescent were selected and sorted in flow cytometry. The mutants with the defective *lacZ* gene or with off-target effects were eliminated from the final selection of transporter mutants. 4 transportation mutants were identified that exhibited defect in the FDG uptake and all except one (UM10), belong to the ABC importer family (UM80, UM83, and UM130).

## 5.2 Characterization of the UM mutants as uptake deficient mutants

To prove that the phenotype observed from the mutants in the screen was due to the reduction of uptake function (not due to the increase in efflux function), retention of fluorescence from the cell-permeant fluorogenic substrate, FDA, was examined. Because FDA is a cell-permeant substrate, it can be readily transported into the cells via diffusion and there should be no difference of fluorescence in the cells of UM mutants and WT. If an increase of fluorescence was observed from the UM mutants with FDA, it would indicate that the efflux systems in the mutants were defective. Or, if an decrease of fluorescence was observed from the UM mutants with FDA, it would indicate that the efflux systems in the mutants were overly active compared to WT. The result from FDA uptake assay showed no significant difference in the level of fluorescence in all the strains tested. This result is consistent with our hypothesis that the deficient FDG uptake phenotype was due to reduced substrate uptake in the UM mutants.

Results generated in the bialaphos toxicity assay provided additional support that reduced FDG uptake was due to defect in internalization. Bialaphos is a prodrug which needs to be activated by cellular enzyme to exert its cytotoxicity. Previous studies demonstrated that Opp/Dpp transporters are the specific entry site for bialaphos and deletion of genes in the operons is associated with bialaphos resistance [62, 32]. We have no concrete evidence that this holds true in *M. smegmatis*, however we were able to demonstrate that bialaphos also enters via peptide transporters in *M. smegmatis*: when *M. smegmatis* was grown in the 7H9 medium containing bialaphos alone, no growth was observed in the

culture. However, presence of the peptides in the bialaphos media competitively inhibited toxicity of bialaphos, allowing *M. smegmatis* to grow. This result strongly suggests that bialaphos and peptides enter via the same route(s). Contrary to the bialaphos toxicity to *M. smegmatis*, all UM mutants demonstrated resistance to bialaphos. No growth defect was observed in the UM mutant cultures in presence of bialaphos. Because bialaphos is known to be taken up via peptide transporter(s), it is very likely that the transporters in the UM mutants are also related to peptide transporters. In fact, UM83 and UM130 are annotated as an Opp and Dpp transporter, respectively.

### 5.2.1 Growth defects of UM mutants in peptides and carbohydrates

Three out of four mutants have transposon insertion in loci encoding ABC transporters but they are annotated to transport different substrates, from sugar to di- and oligopeptides. To distinguish the mutants on preferential substrates, their growth on peptides and carbohydrates was tested, with growth defects as an indicator of deficient uptake function. In this study, all mutants showed a prolonged lag period compared to WT with most peptides and carbohydrates tested. In case of trehalose, the lag period was so extensive that visible growth in the mutants was only observed after WT reached the stationary phase. We believed that the length of lag periods indicates the time it takes for the mutants to adapt. However, despite the prolonged lag periods, their growth recovered and they grew to the same extent as that of WT. This growth recovery suggests that cellular mechanisms were not affected once the substrate was internalized.

This was also shown by their comparable growth in the 7H9 medium.

Interestingly, the mutants appeared to segregate in pairs with UM10/UM83, and UM80/UM130 during the growth assay. This phenotype could not be explained by the genes disrupted in the two mutants in the pairs because the subunits encoded by the genes in the pair are different. The gene organizations of UM80 and UM130 are similar but it is unlikely that the gene organization is the cause since that of UM10 and UM83 does not share similarity. More experiments are needed to fully elucidate this phenotype.

### 5.2.2 Cell wall defect is not associated with transport deficiency

It was previously studied that mutation in the Opp transporter genes of *Mtb* modulats cell wall lipid compositions[32]. We were also curious to see whether the mutations in the UM mutants altered cell wall lipid compositions and its possible effect on transport. We used sensitivity to SDS as an indicator for alteration of cell wall lipid compositions as SDS is a well known ionic detergent to disrupt lipid-lipid, lipid-protein, and protein-protein interactions. Sensitivity of SDS to the mutants was tested by 2 methods; a CFU-base survival assay, and zone of inhibition assay. When SDS toxicity was tested in the form of zone of inhibition (ZOI), no significant difference in size of ZOI was observed from all UM mutants and WT, except that a denser growth around the edge of ZOI was observed from UM10. Additionally, no difference was observed with $_D$-cycloserine. However, when the toxicity was measured in broth (the CFU-base survival assay), a drastic reduction of cell numbers was detected from all UM mutants, particularly with UM10. At 2 h in the 7H9 medium containing 0.1%

SDS, the total CFU was decreased to 0.015% of its starting CFU and the size of UM10 colonies was greatly reduced and lost its pigment. At 8 h incubation with 0.1% SDS, not only UM10 but also other mutants showed the morphological defects. The CFU-base survival assay in broth showed that cell wall of the mutants was much sensitive to SDS toxicity. This result indicates that cell wall lipid compositions in all the mutants were altered which caused the rapid killing and morphological alterations in the environment containing SDS. However, this phenotype is not related to the transport defect phenotype in the UM mutants. If cell wall defect was the cause of the transport defect in the mutants, we should have observed an increase of fluorescence from the UM mutants in presence of FDG because FDG influx should have increased. Also, killing of the UM mutants should have been observed in presence of bialaphos due to weakening of the cell walls, allowing unspecific entry of bialaphos. However, these results were not what we observed. Instead, we observed reduction of fluorescence with FDG and bialaphos resistance in the UM mutants as explained earlier in this chapter. Therefore, we concluded that the phenotypes from the SDS sensitivity and the transport assays showed effects of the mutation independently.

## 5.3 All UM mutants have overlapping phenotypes

All the four independent transport mutants were identified based on their FDG uptake deficient phenotypes. We observed the uptake deficient phenotype from all the four independent mutants. At the beginning, we suspected that the FDG uptake deficient phenotype might have derived from the slower growth in the UM mutants. However, all mutants grew in the 7H9 medium at the same

growth rate as of WT, suggesting that there was no adverse defect in their essential metabolic pathways. Furthermore, to measure fluorescence by flow cytometry, the same number of cells were used that the only difference was the level of fluorescence detected per cell. Therefore, we concluded that their growth did not affect the FDG uptake deficient phenotype and the overlapping phenotype could not be explained by it. Next, we hypothesized that the overlapping phenotype might be due to an increased secretion of fluorescence by some of the mutants. However, supernatants from all the mutants demonstrated the comparable level of fluorescence with the cell permeable fluorogenic substrate, FDA, to that of WT. Furthermore, when we measured the level of fluorescence in the cells, it was also same as that of WT. Both results indicate that the secretion system in the mutants and WT is comparable and is not the cause of the overlapping phenotype. We also looked at the cell wall integrity of the mutants whether the alteration of cell wall lipid compositions caused an off-target FDG uptake deficient phenotype. We challenged the mutants to grow in presence of SDS. With SDS, the survival rates in the mutants were severely reduced and their colonies lost usual mycobacterial colony morphology, i.e., orange pigmentation, waxy appearance, and wrinkling of the colony surface. Would this defect be the cause of the overlapping FDG uptake deficient phenotype? As explained earlier in this chapter, this phenotype only explains the relationship between the mutation in the transporter genes and weakening of cell wall integrity, but it is difficult to explain the overlapping phenotypes by the weakening of cell wall integrity.

We also observed an overlapping phenotype on bialaphos from the UM mutants. When bialaphos was added to 7H9 medium, no growth was observed in theWT culture but all UM mutants grew well. No effect of bialaphos was ob-

served in their growth and their growth patterns with and without bialaphos werecomparable. To pursue the basis of this phenotype, it was suggested that we test bialaphos toxicity on the UM mutants in absence of kanamycin. We always addedkanamycin in the UM mutant cultures because the transposon marker possesses a kanamycin resistant gene and the cultures can be kept free of contamination by inclusion of the antibiotic.However, it is not required as a selectable marker because the transposon with the $Kan^r$ cassette is integrated into the genome. It was suggested that addition of other antibiotics like kanamycin could inducebialaphos resistance, therefore we tested bialaphos toxicity on the UM mutants in absence of the antibiotic. Preliminary result shows that when kanamycin waseliminated from the 7H9 medium containing bialaphos, growth of the UM mutants was greatly reduced and minimal growth could be detected in the UM mutantsafter 24 h of incubation, at which time point, we typically observe strong growth in these cultures. This result indicates that the presence of kanamycin protected the UMmutants from bialaphos toxicity and induced the overlapping phenotype, but more thorough experiments are needed to elucidate the effect ofkanamycin on the induction of bialaphos resistance in the UM mutants. The effect could be due to reduced uptake of bialaphos, altered bacterial metabolism, or enhanced efflux. These require additional experiments to resolve.

## 5.4 Future work

In this study, we partially characterized 4 individual transporters for their possible roles in nutrient acquisition. The most puzzling phenotypes observed were the overlapping phenotypes of the UM mutants in the FDG defective uptake.

Elucidating how, why, and what caused the overlapping phenotypes may help us to fully understand their roles in nutrient acquisition and their other possible roles, as well as their substrate specificity.

First approach we could take is to determine any alteration in the cellular mechanisms, i.e., regulation and signaling pathways, due to the transposon mutations. Comparative analysis of transcription profiles of the UM mutants growing in substrates such as peptides or carbohydrates might unravel not only shifts of the cellular mechanisms but also the identification of alternative transport pathway(s).

Along with the overlapping phenotypes, we also need to elucidate substrate specificities of individual transporters as the emerging evidence provides us that *Mtb* and other pathogens exploit the nutrient transport systems for their survival during the infection. The biggest hurdle is to find (or design) a bioprobe which is not structurally perturbing and is comparable to the target substrate. In the beginning of my Ph.D. program, we tried to employ click chemistry to identify transporter proteins. In this technique, non-functional end of a target substrate could be modified to either an alkyne or azide form. The transporter protein and the target substrate complex could be detected using a counterpart probe which was conjugated to the modified end of the target substrate. Because the alkyne and azide are very small molecules, they are good candidates as a probe. Caveat in our experiments was that we chose biotin to conjugate to the counterpart probe without realizing that biotinylated proteins are abundant in *M. smegmatis*. We also used a fluorescent molecule to conjugate to the counterpart probe to detect the transporter protein-target molecule complex. However, the background fluorescent was too high to identify the protein.

Lastly, it will be very interesting to study the transport mechanism involved with *MSMEG_4363*. UM10, *MSMEG_4363* mutant, showed the same substrate uptake deficient phenotype and bialaphos resistant to the other UM mutants, however as mentioned before, *MSMEG_4363* does not belong to any ABC transporter family. Instead, it possesses a NTF2-like motif. Vir8 of *A. tumerfaciens* and *B.suis*, a part of Type IV secretion system (T4SS), also possess the NTF2-like motif. Gram positive bacteria including *M. smegmatis* do not have T4SS. We cannot help speculate that MSMEG_4363 is a part of a transport systems yet to be uncovered. To test, immunoprecipitation using an epitope-tag MSEMG_4363 is one method to identify proteins interacting with MSMEG_4363. This might uncover a new transport system in mycobacteria.

Transporters are the first proteins which encounter external molecules. Substrate specificities of the transporters determine what molecules to be transported into the cells. Substrates can range from nutrients, signaling peptides, antigens to drugs and multiple molecules. Some of them can be transported via the same transporter(s) or via multiple transporters. Functions and mechanisms of transporters in *Mtb* mostly remain elusive. Exploitation of trehalose and Opp transport systems by *Mtb* to gain virulency in the host cells tells us that elucidating nutrient transporters are not only to understand what *Mtb* eats in the host cells but also how *Mtb* maintains its pathogenicity within the host cells. It is urgently needed to understand the transport mechanisms to fully appreciate the true survival mechanisms of *Mtb* during infection.

APPENDIX A

## SECONDARY STRUCTURAL PREDICTION OF MSMEG_4363

### Input

```
>query
MCHTPTCRLE LSSPGVVMRV PGLQRFSRGN PSSRDARRIN AGLVRATAGR
VPPAVAVLAA GALALSGCAN ETPPEQPGAH TRAVLEDVFA TTSQTGFGET
FLDRLSDDVT FTATGTSPVA GQYHGKTEYR EKVLSRLHDH LATPMRPQLD
QMIVDGDWAA VRFHAEGVHG TNGSDASMQY CWVMRVAGDQ IVDVIGYYDT
AKMAGLFV
```

## Secondary Structure

**Prediction (brief)** (Show Landscape View)

```
          ....,....1....,....2....,....3....,....4....,....5....,....6
AA        MCHTPTCRLELSSPGVVMRVPGLQRFSRGNPSSRDARRINAGLVRATAGRVPPAVAVLAA
OBS_sec
PROF_sec                    EEEE             HHHHHHHHHHHHHHHH    HHHHHHHH
Rel_sec   954664324455563244135400000257402443321012233210235226888888
SUB_sec     .  .....    ...... .......  ..................... ..

O_3_acc   bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P_3_acc   e  e ebebeee ebbbb b  b ebeee eebeebee beebbe b eeb  bbbbbbb
Rel_acc   311100322110010123012101111321223222230011122120110012522262
SUB_acc   ..................................................... ... .


          ....,....7....,....8....,....9....,....10...,....11...,....12
AA        GALALSGCANETPPEQPGAHTRAVLEDVFATTSQTGFGETFLDRLSDDVTFTATGTSPVA
OBS_sec
PROF_sec  HHHHH              HHHHHHHHHHHHHHHH   HHHHHHHH    EEE
Rel_sec   654211124567755430345788998887530366278875421340322146666 55
SUB_sec     .......      .....            ...  .     ...........

O_3_acc   bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P_3_acc   bbbb     eeeeee eee b ebb  bbe b eee be bbe bbeebeb b ee e
Rel_acc   311111100120022011413333933442231202003280332022120100100110
SUB_acc   .................. ..... ..  ........... ...................


          ....,....13...,....14...,....15...,....16...,....17...,....18
AA        GQYHGKTEYREKVLSRLHDHLATPMRPQLDQMIVDGDWAAVRFHAEGVHGTNGSDASMQY
OBS_sec
PROF_sec       HHHHHHHHHHHHHHH      EEEEEEEEE   EEEEEEEEEEEE       EEEEE
Rel_sec   432343788898888665311454201234455316436888863000124564000014
SUB_sec   ......            .... ........  .. ..      .......  .......

O_3_acc   bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P_3_acc       b e bbe bbe bbe beeebeb b  bbbeee bbb b beb   eeeeebeb b
Rel_acc   001113135124513350320111230212121110216362230101101300312021
SUB_acc   ........ ..  ... ..................... . ...................


          ....,....19...,....20...,...
AA        CWVMRVAGDQIVDVIGYYDTAKMAGLFV
OBS_sec
PROF_sec  EEEEEE   EEEEEEEEE  HHHHHH
Rel_sec   57888517606665321040233220 28
SUB_sec         .  .    ............

O_3_acc   bbbbbbbbbbbbbbbbbbbbbbbbbbbb
P_3_acc   bbbb b e  bb b   b  e bbe be
Rel_acc   233623120231211100121140300 3
SUB_acc   ... .................. .....
```

## Transmembrane

**PHDhtm summary**

- **NHTM=1**
  PHDhtm detected **1** membrane helix for the **best** model.The second best model contained 0 helix.
- **TOP=in**
  PHDhtm predicted the topology in, i.e. the first loop region is in (Note: this prediction may be problematic when the sequence you sent starts or ends with a region predic
  membrane helix!)
- Reliability of best model=2 (0 is low, 9 is high)
- Zscore for best model=0.000
- Difference of positive charges (K+R) inside - outside=-2.900 (the higher the value, the more reliable)
- Reliability of topology prediction =2 (0 is low, 9 is high)
- Details of the strength of each predicted membrane helix:
  (sorted by strength, strongest first)
- N HTM
- Total score
- Best HTM
- c-N
- 1
- 0.9567
- 0.6444
- 48 - 65
- Overview over transmembrane segments:

- Positions
- Segments
- Explain
- 1- 47
- i1
- inside region 1
- 48- 65
- T1
- ?=error 1
- 66- 208
- o1
- outside region 1

**PHD results (normal)** (Show Landscape View)

```
          ....,....1....,....2....,....3....,....4....,....5....,....6....,....7....,....8....,....9....,....1
AA        MCHTPTCRLELSSPGVVMRVPGLQRFSRGNPSSRDARRINAGLVRATAGRVPPAVAVLAAGALALSGCANETPPEQPGAHTRAVLEDVFATTSQTGFGET
PHD_htm                                                  MMMMMMMMMMMM
Rel_htm   99999999999999999999999999999999999999999999998753311356777655540357889999999999999999999999999999
SUB_htm                                                 ........   ........

PHDrhtm                                                 MMMMMMMMMMMMMMMM
PiMohtm   iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiTTTTTTTTTTTTTTTTTooooooooooooooooooooooooooooooooooo

          ....,....11...,....12...,....13...,....14...,....15...,....16...,....17...,....18...,....19...,....20
AA        FLDRLSDDVTFTATGTSPVAGQYHGKTEYREKVLSRLHDHLATPMRPQLDQMIVDGDWAAVRFHAEGVHGTNGSDASMQYCWVMRVAGDQIVDVIGYYDT
PHD_htm
Rel_htm   99999999999999999999999999999999999999999999999999999999999999999999999999999999999999999999999999
SUB_htm

PHDrhtm
PiMohtm   oooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooo

          ....,....2
AA        AKMAGLFV
PHD_htm
Rel_htm   99999999
SUB_htm

PHDrhtm
PiMohtm   oooooooo
```

## Protein Localization Prediction

**Localization prediction in prokaryotes**

```
== Summary of subcellular localization prediction using LOCtree ==
Protein Identifier  Predicted Localization   Reliability Index   Intermediate localization prediction    Reliability index of intermediate
localization predictions
query     Cytoplasmic    2    Cytoplasmic    2
```

## Protein-Protein binding

```
>query: strech=5 crowd_predictions=7 gap=20 itr=0
MCHTPTCRLELSSPGVVMRVPGLQRFSRGNPSSRDARRIN
PP-PPP-P-PP---------------PPP-------PP--

AGLVRATAGRVPPAVAVLAAGALALSGCANETPPEQPGAH
------------------------------P--PPP----

TRAVLEDVFATTSQTGFGETFLDRLSDDVTFTATGTSPVA
-----------------------------------PP---

GQYHGKTEYREKVLSRLHDHLATPMRPQLDQMIVDGDWAA
-----------P----------P-----------------

VRFHAEGVHGTNGSDASMQYCWVMRVAGDQIVDVIGYYDT
----------PP---------------------------

AKMAGLFV
--------
```

## Prosite

```
----------------------------------------------------------------
Pattern-ID: ASN_GLYCOSYLATION PS00001 PDOC00001
Pattern-DE: N-glycosylation site
Pattern:    N[^P][ST][^P]
   172      NGSD

Pattern-ID: PKC_PHOSPHO_SITE PS00005 PDOC00005
Pattern-DE: Protein kinase C phosphorylation site
Pattern:    [ST].[RK]
   6        TCR
   32       SSR
   200      TAK

Pattern-ID: CK2_PHOSPHO_SITE PS00006 PDOC00006
Pattern-DE: Casein kinase II phosphorylation site
Pattern:    [ST].{2}[DE]
   32       SSRD
   72       TPPE
   100      TFLD

Pattern-ID: MYRISTYL PS00008 PDOC00008
Pattern-DE: N-myristoylation site
Pattern:    G[^EDRKHPFYW].{2}[STAGCN][^P]
   29       GNPSSR
   42       GLVRAT
   121      GQYHGK
   167      GVHGTN
   173      GSDASM

Pattern-ID: ATP_GTP_A PS00017 PDOC00017
Pattern-DE: ATP/GTP-binding site motif A (P-loop)
Pattern:    [AG].{4}GK[ST]
   120      AGQYHGKT
```

APPENDIX B

# SECONDARY STRUCTURAL PREDICTION OF VIRB8 OF *AGROBACTERIUM TUMEFACIENS*

## Input

```
>query
MKGSEYALLV ARETLAEHYK EVEAFQTARA KSARRLSKVI AAVATIAVLG
NVAQAFTIAT MVPLIRLVPV YLWIRPDGTV DSEVSVSRLP ATQEEAVVNA
SLWEYVRLRE SYDADTAQYA YDLVSNFSAP MVRQNYQQFF NYPNPTSPQV
ILGKHGRLEV EHIASNDVTP GVQQIRYKRT LIVDGKMPMA STWTATVRYE
KVTSLPGRLR LTNPGGLVVT SYQTSEDTVS NAGHSEP
```

## Secondary Structure

**Prediction (brief)** (Show Landscape View)

```
          ....,....1....,....2....,....3....,....4....,....5....,....6
AA        MKGSEYALLVARETLAEHYKEVEAFQTARAKSARRLSKVIAAVATIAVLGNVAQAFTIAT
OBS_sec
PROF_sec                 HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHE
Rel_sec   976412334553110145666654334334567777777888888888888876542 01
SUB_sec      ......  ......      .......                              ....

O_3_acc   bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P_3_acc   eeeeeeeee eeee eebbeebeeb ee be beeb  bbbbbbbbbbbbbbbbbbbbbb
Rel_acc   330221200111320250162333011120202334121513896868986975657751
SUB_acc   ..............  ..  ..............  ...  ..                  .


          ....,....7....,....8....,....9....,....10...,....11...,....12
AA        MVPLIRLVPVYLWIRPDGTVDSEVSVSRLPATQEEAVVNASLWEYVRLRESYDADTAQYA
OBS_sec
PROF_sec  E       EEEEEEE    EEEEEE        HHHHHHHHHHHHHHHHH   HHHHHHH
Rel_sec   202443112688871577423543102244561046778989998875420361245544
SUB_sec   .........     .   ... ........  ...             .... ...  ..

O_3_acc   bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P_3_acc   bbe eebbbbbb b ee eb bbb bee e e eebbb  bbb bbeb eebe e b eb
Rel_acc   301112002246152120202211201202011010141209233721121232213222
SUB_acc   ..........  . ......................... ... ... ..............


          ....,....13...,....14...,....15...,....16...,....17...,....18
AA        YDLVSNFSAPMVRQNYQQFFNYPNPTSPQVILGKHGRLEVEHIASNDVTPGVQQIRYKRT
OBS_sec
PROF_sec  HHHHHH   HHHHHHHHHHH         EEE    EEEEEEEEEEE    EEEEEEEEE
Rel_sec   442200144136777777631457655210102430477888875322674257887643
SUB_sec   ...........        ...     ..........        ...  ..      ..

O_3_acc   bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P_3_acc   be b  bb eeb e b ebbeeeeee     beee eb b bbbb  beee b b be
Rel_acc   032721113315232533004011131011100201233656033120022232734222
SUB_acc   ... ....... ... .... ..................   ............ . ...


          ....,....19...,....20...,....21...,....22...,....23...,..
AA        LIVDGKMPMASTWTATVRYEKVTSLPGRLRLTNPGGLVVTSYQTSEDTVSNAGHSEP
OBS_sec
PROF_sec  EEE       EEEEEEEEEEE               EEEEEEE
Rel_sec   330465542102578888753256540000126640477531013111356765778
SUB_sec   ....   .....        ..   .......  ...   .........

O_3_acc   bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P_3_acc      eee ee e b bbb b   ee eee   ebbbbbbbbbb   ee beeeeeeee
Rel_acc   221222020112138172331113012211110123617200120021011023316
SUB_acc   .............. . .................. . .................
```

### Transmembrane

**PHDhtm summary**

- **NHTM=1**
  PHDhtm detected **1** membrane helix for the **best** model.The second best model contained 0 helix.
- **TOP=in**
  PHDhtm predicted the topology in, i.e. the first loop region is in (Note: this prediction may be problematic when the sequence you sent starts or ends with a region predicted in a membrane helix!)
- Reliability of best model=8 (0 is low, 9 is high)
- Zscore for best model=0.000
- Difference of positive charges (K+R) inside - outside=-13.001 (the higher the value, the more reliable)
- Reliability of topology prediction =9 (0 is low, 9 is high)
- Details of the strength of each predicted membrane helix:
  (sorted by strength, strongest first)
- N HTM
- Total score
- Best HTM
- c-N
- 1
- 0.9463
- 0.8925
- 42 - 66
- Overview over transmembrane segments:

- Positions
- Segments
- Explain
- 1- 41
- i1
- inside region 1
- 42- 66
- T1
- ?=error 1
- 67- 237
- o1
- outside region 1

**PHD results (normal)** (Show Landscape View)

```
          ....,....1....,....2....,....3....,....4....,....5....,....6....,....7....,....8....,....9....,....1
AA        MKGSEYALLVARETLAEHYKEVEAFQTARAKSARRLSKVIAAVATIAVLGNVAQAFTIATMVPLIRLVPVYLWIRPDGTVDSEVSVSRLPATQEEAVVNA
PHD_htm                                        MMMMMMMMMMMMMMMMMMMMMMMMMMMMMM
Rel_htm   99999999999999999999999999999999998875035677888877777777777777778776542035788999999999999999999998
SUB_htm                                       .....                         .......

PHDrhtm                                          MMMMMMMMMMMMMMMMMMMMMM
PiMohtm   iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiTTTTTTTTTTTTTTTTTTTTTTTTooooooooooooooooooooooooooooooooooo

          ....,....11...,....12...,....13...,....14...,....15...,....16...,....17...,....18...,....19...,....20
AA        SLWEYVRLRESYDADTAQYAYDLVSNFSAPMVRQNYQQFFNYPNPTSPQVILGKHGRLEVEHIASNDVTPGVQQIRYKRTLIVDGKMPMASTWTATVRYE
PHD_htm
Rel_htm   899999999999999999999999999999999999999999999999999999999999999999999999999999999999999999999999999999999999999999999999999999999988899999
SUB_htm

PHDrhtm
PiMohtm   oooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooo

          ....,....21...,....22...,....23...,....24
AA        KVTSLPGRLRLTNPGGLVVTSYQTSEDTVSNAGHSEP
PHD_htm
Rel_htm   99999999999999999999999999999999999
SUB_htm

PHDrhtm
PiMohtm   ooooooooooooooooooooooooooooooooooo
```

## Protein Localization Prediction

### Localization prediction in prokaryotes

```
== Summary of subcellular localization prediction using LOCtree ==
Protein Identifier  Predicted Localization   Reliability Index   Intermediate localization prediction    Reliability index of intermediate
localization predictions
query     Cytoplasmic    2    Cytoplasmic    2
```

## Protein-Protein binding

```
>query: strech=5 crowd_predictions=7 gap=20 itr=0
MKGSEYALLVARETLAEHYKEVEAFQTARAKSARRLSKVI
--------------------P------------PP-----

AAVATIAVLGNVAQAFTIATMVPLIRLVPVYLWIRPDGTV
----------------------------------------

DSEVSVSRLPATQEEAVVNASLWEYVRLRESYDADTAQYA
----------------------------------------

YDLVSNFSAPMVRQNYQQFFNYPNPTSPQVILGKHGRLEV
------------P---------------------------

EHIASNDVTPGVQQIRYKRTLIVDGKMPMASTWTATVRYE
----------------------------------------

KVTSLPGRLRLTNPGGLVVTSYQTSEDTVSNAGHSEP
P-----------------------------------
```

**Prosite**

```
-----------------------------------------------------------------
Pattern-ID: ASN_GLYCOSYLATION PS00001 PDOC00001
Pattern-DE: N-glycosylation site
Pattern:    N[^P][ST][^P]
   99       NASL
   126      NFSA

Pattern-ID: CAMP_PHOSPHO_SITE PS00004 PDOC00004
Pattern-DE: cAMP- and cGMP-dependent protein kinase phosphorylation site
Pattern:    [RK]{2}.[ST]
   34       RRLS

Pattern-ID: PKC_PHOSPHO_SITE PS00005 PDOC00005
Pattern-DE: Protein kinase C phosphorylation site
Pattern:    [ST].[RK]
   27       TAR
   32       SAR
   196      TVR

Pattern-ID: CK2_PHOSPHO_SITE PS00006 PDOC00006
Pattern-DE: Casein kinase II phosphorylation site
Pattern:    [ST].{2}[DE]
   14       TLAE
   92       TQEE
   101      SLWE
   224      TSED

Pattern-ID: MYRISTYL PS00008 PDOC00008
Pattern-DE: N-myristoylation site
Pattern:    G[^EDRKHPFYW].{2}[STAGCN][^P]
   3        GSEYAL
   78       GTVDSE
   216      GLVVTS
```

APPENDIX C

## SECONDARY STRUCTURAL PREDICTION OF VIRB8 OF *BRUCELLA SUIS*

## Input

```
>query
MFGRKQSPQK SVKNGQGNAP SVYDEALNWE AAHVRLVEKS ERRAWKIAGA
FGTITVLLGI GIAGMLPLKQ HVPYLVRVNA QTGAPDILTS LDEKSVSYDT
VMDKYWLSQY VIARETYDWY TLQKDYETVG MLSSPSEGQS YASQFQGDKA
LDKQYGSNVR TSVTIVSIVP NGKGIGTVRF AKTTKRTNET GDGETTHWIA
TIGYQYVNPS LMSESARLTN PLGFNVTSYR VDPEMGVVQ
```

## Secondary Structure

**Prediction (brief)** (Show Landscape View)

```
          ....,....1....,....2....,....3....,....4....,....5....,....6
AA        MFGRKQSPQKSVKNGQGNAPSVYDEALNWEAAHVRLVEKSERRAWKIAGAFGTITVLLGI
OBS_sec
PROF_sec                          HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
Rel_sec   955435643335666655420456654433444433456565567677777766666553
SUB_sec      ..  ....       ....    ...........                        .

O_3_acc   bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P_3_acc   e eeeeeeeee eeeeeeeee beeb eb  e be beebb  bbbbbbbbbbbbbbbb
Rel_acc   312002111211112231103212252312111222323002242289688789887766
SUB_acc   ......................... ................. ..


          ....,....7....,....8....,....9....,....10...,....11...,....12
AA        GIAGMLPLKQHVPYLVRVNAQTGAPDILTSLDEKSVSYDTVMDKYWLSQYVIARETYDWY
OBS_sec
PROF_sec  HEEEEE      EEEEEEE     EEEEEE        HHHHHHHHHHHHHHHHH   HH
Rel_sec   101331145531168887256761567751024555500467788898876533102601
SUB_sec   ........  ...     .    .     ....    ...            ..... ..

O_3_acc   bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P_3_acc   bbbbbbe ee bbbbb bbeee eb bb  beeee e eebb  bbbb bb b ee e e
Rel_acc   775131012100327726001112112202101321122224021272236221011311
SUB_acc      ...........  . ......................... .... ... .........


          ....,....13...,....14...,....15...,....16...,....17...,....18
AA        TLQKDYETVGMLSSPSEGQSYASQFQGDKALDKQYGSNVRTSVTIVSIVPNGKGIGTVRF
OBS_sec
PROF_sec  HHHHHHHEEEEE   HHHHHHHHHH             EEEEEEEEEEE    EEEEEE
Rel_sec   343221022320261256777665304566310013420367788877423675257888
SUB_sec   ............. ..        ...   ..........        ...   .

O_3_acc   bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P_3_acc    b eebe b bbb eebbe b e b eee  ee  eee  b b bbbbb  eeebb b b
Rel_acc   131321223210120221334132012320012000221223647054311221032734
SUB_acc   ..................... ......................   .  ......... .


          ....,....19...,....20...,....21...,....22...,....23...,....24
AA        AKTTKRTNETGDGETTHWIATIGYQYVNPSLMSESARLTNPLGFNVTSYRVDPEMGVVQ
OBS_sec
PROF_sec  EEEEEE         EEEEEEEEEEE                EEEEEEEE
Rel_sec   754310246655421036888888740565565100012653036764220110010018
SUB_sec     ......    .....        ..      ......  ...   ...........

O_3_acc   bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P_3_acc       e  eeeeeee   bbbbbbb b  eeeeeee   ebbbbbbb   e eee eeee
Rel_acc   222121112111101222194935112012011211110123606110301020101 6
SUB_acc   ...................   . ................... . ............
```

### Transmembrane

**PHDhtm summary**

- **NHTM=1**
  PHDhtm detected **1** membrane helix for the **best** model.The second best model contained 0 helix.
- **TOP=in**
  PHDhtm predicted the topology in, i.e. the first loop region is in (Note: this prediction may be problematic when the sequence you sent starts or ends with a re membrane helix!)
- Reliability of best model=6 (0 is low, 9 is high)
- Zscore for best model=0.000
- Difference of positive charges (K+R) inside - outside=-9.537 (the higher the value, the more reliable)
- Reliability of topology prediction =9 (0 is low, 9 is high)
- Details of the strength of each predicted membrane helix:
  (sorted by strength, strongest first)
- N HTM
- Total score
- Best HTM
- c-N
- 1
- 0.9601
- 0.8825
- 49 - 67
- Overview over transmembrane segments:

- Positions
- Segments
- Explain
- 1- 48
- i1
- inside region 1
- 49- 67
- T1
- ?=error 1
- 68- 239
- o1
- outside region 1

**PHD results (normal)** (Show Landscape View)

```
          ....,....1....,....2....,....3....,....4....,....5....,....6....,....7....,....8....,....9....,....1
AA        MFGRKQSPQKSVKNGQGNAPSVYDEALNWEAAHVRLVEKSERRAWKIAGAFGTITVLLGIGIAGMLPLKQHVPYLVRVNAQTGAPDILTSLDEKSVSYDT
PHD_htm                                                MMMMMMMMMMMMMMMMMMMMM
Rel_htm   99999999999999999999999999999999999999999998763135677887777777877776410356778999999999999999999999999
SUB_htm                                               ......               .......

PHDrhtm                                                  MMMMMMMMMMMMMMMMMMM
PiMohtm   iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiTTTTTTTTTTTTTTTTTTTooooooooooooooooooooooooooooooooooo

          ....,....11...,....12...,....13...,....14...,....15...,....16...,....17...,....18...,....19...,....20
AA        VMDKYWLSQYVIARETYDWYTLQKDYETVGMLSSPSEGQSYASQFQGDKALDKQYGSNVRTSVTIVSIVPNGKGIGTVRFAKTTKRTNETGDGETTHWIA
PHD_htm
Rel_htm   99999999999999999999999999999999999999999999999999999999999999999999999999888999999999999999999999999998
SUB_htm

PHDrhtm
PiMohtm   ooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooooo

          ....,....21...,....22...,....23...,....24
AA        TIGYQYVNPSLMSESARLTNPLGFNVTSYRVDPEMGVVQ
PHD_htm
Rel_htm   888899999999999999999999999999999999
SUB_htm

PHDrhtm
PiMohtm   oooooooooooooooooooooooooooooooooooo
```

## Protein Localization Prediction

**Localization prediction in prokaryotes**

```
== Summary of subcellular localization prediction using LOCtree ==
Protein Identifier  Predicted Localization   Reliability Index   Intermediate localization prediction    Reliability index of intermediate
localization predictions
query     Cytoplasmic    5    Cytoplasmic    5
```

## Protein-Protein binding

```
>query: strech=5 crowd_predictions=7 gap=20 itr=0
MFGRKQSPQKSVKNGQGNAPSVYDEALNWEAAHVRLVEKS
P-----------------------------------PP-

ERRAWKIAGAFGTITVLLGIGIAGMLPLKQHVPYLVRVNA
--P-------------------------------------

QTGAPDILTSLDEKSVSYDTVMDKYWLSQYVIARETYDWY
------------PP------------------------PP

TLQKDYETVGMLSSPSEGQSYASQFQGDKALDKQYGSNVR
-------P--------------------------------

TSVTIVSIVPNGKGIGTVRFAKTTKRTNETGDGETTHWIA
---------------------------P-----------

TIGYQYVNPSLMSESARLTNPLGFNVTSYRVDPEMGVVQ
-------------------------P------------
```

## Prosite

```
-----------------------------------------------------------
Pattern-ID: ASN_GLYCOSYLATION PS00001 PDOC00001
Pattern-DE: N-glycosylation site
Pattern:    N[^P][ST][^P]
   188      NETG
   225      NVTS

Pattern-ID: CAMP_PHOSPHO_SITE PS00004 PDOC00004
Pattern-DE: cAMP- and cGMP-dependent protein kinase phosphorylation site
Pattern:    [RK]{2}.[ST]
   4        RKQS

Pattern-ID: PKC_PHOSPHO_SITE PS00005 PDOC00005
Pattern-DE: Protein kinase C phosphorylation site
Pattern:    [ST].[RK]
   11       SVK
   40       SER
   177      TVR
   183      TTK
   215      SAR
   228      SYR

Pattern-ID: CK2_PHOSPHO_SITE PS00006 PDOC00006
Pattern-DE: Casein kinase II phosphorylation site
Pattern:    [ST].{2}[DE]
   21       SVYD
   89       TSLD
   100      TVMD
   134      SPSE

Pattern-ID: TYR_PHOSPHO_SITE PS00007 PDOC00007
Pattern-DE: Tyrosine kinase phosphorylation site
Pattern:    [RK].{2,3}[DE].{2,3}Y
   149      KALDKQY

Pattern-ID: MYRISTYL PS00008 PDOC00008
Pattern-DE: N-myristoylation site
Pattern:    G[^EDRKHPFYW].{2}[STAGCN][^P]
   17       GNAPSV
   49       GAFGTI
   59       GIGIAG
   138      GQSYAS

Pattern-ID: AMIDATION PS00009 PDOC00009
Pattern-DE: Amidation site
Pattern:    .G[RK][RK]
   2        FGRK
```

## BIBLIOGRAPHY


[1] Tuberculosis crisis. *World Health Forum*, 15(3):292–3, 1994.

[2] Johannes Amon, Fritz Titgemeyer, and Andreas Burkovski. A genomic view on nitrogen metabolism and nitrogen control in mycobacteria. *J Mol Microbiol Biotechnol*, 17(1):20–9, 2009.

[3] Susan Bailey, Doyle Ward, Rebecca Middleton, J Gunter Grossmann, and Patricia C Zambryski. Agrobacterium tumefaciens virb8 structure reveals potential protein-protein interaction sites. *Proc Natl Acad Sci U S A*, 103(8):2582–7, Feb 2006.

[4] V Balasubramanian, M S Pavelka, Jr, S S Bardarov, J Martin, T R Weisbrod, R A McAdam, B R Bloom, and W R Jacobs, Jr. Allelic exchange in mycobacterium tuberculosis with long linear recombination substrates. *J Bacteriol*, 178(1):273–9, Jan 1996.

[5] Niaz Banaei, Eleanor Z Kincaid, S-Y Grace Lin, Edward Desmond, William R Jacobs, Jr, and Joel D Ernst. Lipoprotein processing is essential for resistance of mycobacterium tuberculosis to malachite green. *Antimicrob Agents Chemother*, 53(9):3799–802, Sep 2009.

[6] S Bardarov, J Kriakov, C Carriere, S Yu, C Vaamonde, R A McAdam, B R Bloom, G F Hatfull, and W R Jacobs, Jr. Conditionally replicating mycobacteriophages: a system for transposon delivery to mycobacterium tuberculosis. *Proc Natl Acad Sci U S A*, 94(20):10961–6, Sep 1997.

[7] Stoyan Bardarov, Svetoslav Bardarov Jr, Jr, Martin S Pavelka Jr, Jr, Vasan Sambandamurthy, Michelle Larsen, JoAnn Tufariello, John Chan, Graham Hatfull, and William R Jacobs Jr, Jr. Specialized transduction: an efficient method for generating marked and unmarked targeted gene disruptions in mycobacterium tuberculosis, m. bovis bcg and m. smegmatis. *Microbiology*, 148(Pt 10):3007–17, Oct 2002.

[8] Christian Baron. Virb8: a conserved type iv secretion system assembly factor and drug target. *Biochem Cell Biol*, 84(6):890–9, Dec 2006.

[9] Michael Berney and Gregory M Cook. Unique flexibility in energy metabolism allows mycobacteria to combat starvation and hypoxia. *PLoS One*, 5(1):e8614, 2010.

[10] Esther Biemans-Oldehinkel, Mark K Doeven, and Bert Poolman. Abc transporter architecture and regulatory roles of accessory domains. *FEBS Lett*, 580(4):1023–35, Feb 2006.

[11] Enrica Bordignon, Mathias Grote, and Erwin Schneider. The maltose atp-binding cassette transporter in the 21st century–towards a structural dynamic perspective on its mode of action. *Mol Microbiol*, 77(6):1354–66, Sep 2010.

[12] M F Cantwell, D E Snider, Jr, G M Cauthen, and I M Onorato. Epidemiology of tuberculosis in the united states, 1985 through 1992. *JAMA*, 272(7):535–9, Aug 1994.

[13] Centers for Disease Control and Prevention (CDC). Mortality among patients with tuberculosis and associations with hiv status — united states, 1993-2008. *MMWR Morb Mortal Wkly Rep*, 59(46):1509–13, Nov 2010.

[14] Jennifer C Chang, Maurine D Miner, Amit K Pandey, Wendy P Gill, Nada S Harik, Christopher M Sassetti, and David R Sherman. igr genes and mycobacterium tuberculosis cholesterol metabolism. *J Bacteriol*, 191(16):5232–9, Aug 2009.

[15] Michael C Chao and Eric J Rubin. Letting sleeping dos lie: does dormancy play a role in tuberculosis? *Annu Rev Microbiol*, 64:293–311, Oct 2010.

[16] Pascale Charbonnel, Mauld Lamarque, Jean-Christophe Piard, Christophe Gilbert, Vincent Juillard, and Danièle Atlan. Diversity of oligopeptide transport specificity in lactococcus lactis species. a tool to unravel the role of oppa in uptake specificity. *J Biol Chem*, 278(17):14832–40, Apr 2003.

[17] Chen-Yuan Chiang, Rosella Centis, and Giovanni Battista Migliori. Drug-resistant tuberculosis: past, present, future. *Respirology*, 15(3):413–32, Apr 2010.

[18] Gwénaëlle Choquet, Nathalie Jehan, Christine Pissavin, Carlos Blanco, and Mohamed Jebbar. Ousb, a broad-specificity abc-type transporter from erwinia chrysanthemi, mediates uptake of glycine betaine and choline with a high affinity. *Appl Environ Microbiol*, 71(7):3389–98, Jul 2005.

[19] Weihua Chu, Dhiraj A Vattem, Vatsala Maitin, Mary B Barnes, and Robert J C McLean. Bioassays of quorum sensing compounds using agrobac-

terium tumefaciens and chromobacterium violaceum. *Methods Mol Biol*, 692:3–19, 2011.

[20] S T Cole, R Brosch, J Parkhill, T Garnier, C Churcher, D Harris, S V Gordon, K Eiglmeier, S Gas, C E Barry, 3rd, F Tekaia, K Badcock, D Basham, D Brown, T Chillingworth, R Connor, R Davies, K Devlin, T Feltwell, S Gentles, N Hamlin, S Holroyd, T Hornsby, K Jagels, A Krogh, J McLean, S Moule, L Murphy, K Oliver, J Osborne, M A Quail, M A Rajandream, J Rogers, S Rutter, K Seeger, J Skelton, R Squares, S Squares, J E Sulston, K Taylor, S Whitehead, and B G Barrell. Deciphering the biology of mycobacterium tuberculosis from the complete genome sequence. *Nature*, 393(6685):537–44, Jun 1998.

[21] S T Cole, K Eiglmeier, J Parkhill, K D James, N R Thomson, P R Wheeler, N Honoré, 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. Massive gene decay in the leprosy bacillus. *Nature*, 409(6823):1007–11, Feb 2001.

[22] José R R Cussiol, Thiago G P Alegria, Luke I Szweda, and Luis E S Netto. Ohr (organic hydroperoxide resistance protein) possesses a previously undescribed activity, lipoyl-dependent peroxidase. *J Biol Chem*, 285(29):21943–50, Jul 2010.

[23] Thomas M Daniel. The history of tuberculosis. *Respir Med*, 100(11):1862–70, Nov 2006.

[24] Arunava Dasgupta, Kamakshi Sureka, Devrani Mitra, Baisakhee Saha, Sourav Sanyal, Amit K Das, Parul Chakrabarti, Mary Jackson, Brigitte Gicquel, Manikuntala Kundu, and Joyoti Basu. An oligopeptide transporter of mycobacterium tuberculosis regulates cytokine release and apoptosis of infected macrophages. *PLoS One*, 5(8):e12225, 2010.

[25] Amy L Davidson, Elie Dassa, Cedric Orelle, and Jue Chen. Structure, function, and evolution of bacterial atp-binding cassette systems. *Microbiol Mol Biol Rev*, 72(2):317–64, table of contents, Jun 2008.

[26] Luiz Pedro S de Carvalho, Steven M Fischer, Joeli Marrero, Carl Nathan, Sabine Ehrt, and Kyu Y Rhee. Metabolomics of mycobacterium tubercu-

losis reveals compartmentalized co-catabolism of carbon substrates. *Chem Biol*, 17(10):1122–31, Oct 2010.

[27] A M Dhople. Factors influencing the in vitro growth of mycobacterium leprae: effect of inoculum. *Microbios*, 94(378):103–12, 1998.

[28] Mark K Doeven, Rupert Abele, Robert Tampé, and Bert Poolman. The binding specificity of oppa determines the selectivity of the oligopeptide atp-binding cassette transporter. *J Biol Chem*, 279(31):32301–7, Jul 2004.

[29] Mark K Doeven, Jan Kok, and Bert Poolman. Specificity and selectivity determinants of peptide transport in lactococcus lactis and other microorganisms. *Mol Microbiol*, 57(3):640–9, Aug 2005.

[30] E Douka, A I Koukkou, G Vartholomatos, S Frillingos, E M Papamichael, and C Drainas. A zymomonas mobilis mutant with delayed growth on high glucose concentrations. *J Bacteriol*, 181(15):4598–604, Aug 1999.

[31] Wolfgang Eisenreich, Thomas Dandekar, Jürgen Heesemann, and Werner Goebel. Carbon metabolism of intracellular bacterial pathogens and possible links to virulence. *Nat Rev Microbiol*, 8(6):401–12, Jun 2010.

[32] Mario Alberto Flores-Valdez, Rowan P Morris, Françoise Laval, Mamadou Daffé, and Gary K Schoolnik. Mycobacterium tuberculosis modulates its cell surface via an oligopeptide permease (opp) transport system. *FASEB J*, 23(12):4091–104, Dec 2009.

[33] Nicolaas C Gey van Pittius, Samantha L Sampson, Hyeyoung Lee, Yeun Kim, Paul D van Helden, and Robin M Warren. Evolution and expansion of the mycobacterium tuberculosis pe and ppe multigene families and their association with the duplication of the esat-6 (esx) gene cluster regions. *BMC Evol Biol*, 6:95, 2006.

[34] Cheol-Min Ghim, Sung Kuk Lee, Shuichi Takayama, and Robert J Mitchell. The art of reporter proteins in science: past, present and future applications. *BMB Rep*, 43(7):451–60, Jul 2010.

[35] M Gogliettino, M Balestrieri, G Pocsfalvi, I Fiume, L Natale, M Rossi, and G Palmieri. A highly selective oligopeptide binding protein from the archaeon sulfolobus solfataricus. *J Bacteriol*, 192(12):3123–31, Jun 2010.

[36] S V Gordon, R Brosch, A Billault, T Garnier, K Eiglmeier, and S T Cole.

Identification of variable regions in the genomes of tubercle bacilli using bacterial artificial chromosome arrays. *Mol Microbiol*, 32(3):643–55, May 1999.

[37] Michael M Gottesman and Victor Ling. The molecular basis of multidrug resistance in cancer: the early years of p-glycoprotein research. *FEBS Lett*, 580(4):998–1009, Feb 2006.

[38] R M Green, A Seth, and N D Connell. A peptide permease mutant of mycobacterium bovis bcg resistant to the toxic peptides glutathione and s-nitrosoglutathione. *Infect Immun*, 68(2):429–36, Feb 2000.

[39] Peter Hebbeln, Dmitry A Rodionov, Anja Alfandega, and Thomas Eitinger. Biotin uptake in prokaryotes by solute transporters with an optional atp-binding cassette-containing module. *Proc Natl Acad Sci U S A*, 104(8):2909–14, Feb 2007.

[40] R Hengge and W Boos. Maltose and lactose transport in escherichia coli. examples of two different types of concentrative transport systems. *Biochim Biophys Acta*, 737(3-4):443–78, Aug 1983.

[41] C F Higgins. Abc transporters: from microorganisms to man. *Annu Rev Cell Biol*, 8:67–113, 1992.

[42] C F Higgins and M M Gibson. Peptide transport in bacteria. *Methods Enzymol*, 125:365–77, 1986.

[43] Dietmar Hillmann, Iris Eschenbacher, Anja Thiel, and Michael Niederweis. Expression of the major porin gene mspa is regulated in mycobacterium smegmatis. *J Bacteriol*, 189(3):958–67, Feb 2007.

[44] K Höner zu Bentrup and D G Russell. Mycobacterial persistence: adaptation to a changing environment. *Trends Microbiol*, 9(12):597–605, Dec 2001.

[45] M Jackson, D C Crick, and P J Brennan. Phosphatidylinositol is an essential phospholipid of mycobacteria. *J Biol Chem*, 275(39):30092–9, Sep 2000.

[46] Rainer Kalscheuer, Brian Weinrick, Usha Veeraraghavan, Gurdyal S Besra, and William R Jacobs, Jr. Trehalose-recycling abc transporter lpqy-sugasugb-sugc is essential for virulence of mycobacterium tuberculosis. *Proc Natl Acad Sci U S A*, Nov 2010.

[47] Joshua B Kelley and Bryce M Paschal. Hyperosmotic stress signaling to the nucleus disrupts the ran gradient and the production of rangtp. *Mol Biol Cell*, 18(11):4365–76, Nov 2007.

[48] Mi-Jeong Kim, Helen C Wainwright, Michael Locketz, Linda-Gail Bekker, Gabriele B Walther, Corneli Dittrich, Annalie Visser, Wei Wang, Fong-Fu Hsu, Ursula Wiehart, Liana Tsenova, Gilla Kaplan, and David G Russell. Caseation of human tuberculosis granulomas correlates with elevated host lipid metabolism. *EMBO Mol Med*, 2(7):258–74, Jul 2010.

[49] E R Kunji, E J Smid, R Plapp, B Poolman, and W N Konings. Di-tripeptides and oligopeptides are taken up via distinct transport mechanisms in lactococcus lactis. *J Bacteriol*, 175(7):2052–9, Apr 1993.

[50] D J Lampe, B J Akerley, E J Rubin, J J Mekalanos, and H M Robertson. Hyperactive transposase mutants of the himar1 mariner transposon. *Proc Natl Acad Sci U S A*, 96(20):11428–33, Sep 1999.

[51] F C Lanfermeijer, A Picon, W N Konings, and B Poolman. Kinetics and consequences of binding of nona- and dodecapeptides to the oligopeptide binding protein (oppa) of lactococcus lactis. *Biochemistry*, 38(44):14440–50, Nov 1999.

[52] B A Lazazzera. The intracellular function of extracellular signaling peptides. *Peptides*, 22(10):1519–27, Oct 2001.

[53] Daniel L Levy and Rebecca Heald. Nuclear size is regulated by importin and ntf2 in xenopus. *Cell*, 143(2):288–98, Oct 2010.

[54] Joeli Marrero, Kyu Y Rhee, Dirk Schnappinger, Kevin Pethe, and Sabine Ehrt. Gluconeogenic carbon flow of tricarboxylic acid cycle intermediates is critical for mycobacterium tuberculosis to establish and maintain infection. *Proc Natl Acad Sci U S A*, 107(21):9819–24, May 2010.

[55] J.H. Miller. *Experiments in Molecular Genetics*. Cold Spring Harbor, CSH Laboratory Press, , NY., 1972.

[56] Maurine D Miner, Jennifer C Chang, Amit K Pandey, Christopher M Sassetti, and David R Sherman. Role of cholesterol in mycobacterium tuberculosis infection. *Indian J Exp Biol*, 47(6):407–11, Jun 2009.

[57] Marc Monot, Nadine Honoré, Thierry Garnier, Nora Zidane, Diana Sher-

afi, Alberto Paniz-Mondolfi, Masanori Matsuoka, G Michael Taylor, Helen D Donoghue, Abi Bouwman, Simon Mays, Claire Watson, Diana Lockwood, Ali Khamesipour, Ali Khamispour, Yahya Dowlati, Shen Jianping, Thomas H Rea, Lucio Vera-Cabrera, Mariane M Stefani, Sayera Banu, Murdo Macdonald, Bishwa Raj Sapkota, John S Spencer, Jérôme Thomas, Keith Harshman, Pushpendra Singh, Philippe Busso, Alexandre Gattiker, Jacques Rougemont, Patrick J Brennan, and Stewart T Cole. Comparative genomic and phylogeographic analysis of mycobacterium leprae. *Nat Genet*, 41(12):1282–9, Dec 2009.

[58] M S Moore and G Blobel. Purification of a ran-interacting protein that is required for protein import into the nucleus. *Proc Natl Acad Sci U S A*, 91(21):10212–6, Oct 1994.

[59] Ernesto J Muñoz-Elías and John D McKinney. Carbon metabolism of intracellular bacteria. *Cell Microbiol*, 8(1):10–22, Jan 2006.

[60] Michael Niederweis. Nutrient acquisition by mycobacteria. *Microbiology*, 154(Pt 3):679–92, Mar 2008.

[61] Hiroshi Nikaido. Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev*, 67(4):593–656, Dec 2003.

[62] Joaquina Nogales, Socorro Muñoz, José Olivares, and Juan Sanjuán. Genetic characterization of oligopeptide uptake systems in sinorhizobium meliloti. *FEMS Microbiol Lett*, 293(2):177–87, Apr 2009.

[63] Gianna Palmieri, Annarita Casbarra, Immacolata Fiume, Giuliana Catara, Antonio Capasso, Gennaro Marino, Silvia Onesti, and Mosé Rossi. Identification of the first archaeal oligopeptide-binding protein from the hyperthermophile aeropyrum pernix. *Extremophiles*, 10(5):393–402, Oct 2006.

[64] Amit K Pandey and Christopher M Sassetti. Mycobacterial persistence requires the utilization of host cholesterol. *Proc Natl Acad Sci U S A*, 105(11):4376–80, Mar 2008.

[65] Pascale Peyron, Julien Vaubourgeix, Yannick Poquet, Florence Levillain, Catherine Botanch, Fabienne Bardou, Mamadou Daffé, Jean-François Emile, Bruno Marchou, Pere-Joan Cardona, Chantal de Chastellier, and Frédéric Altare. Foamy macrophages from tuberculous patients' granulomas constitute a nutrient-rich reservoir for m. tuberculosis persistence. *PLoS Pathog*, 4(11):e1000204, Nov 2008.

[66] A Picon, E R Kunji, F C Lanfermeijer, W N Konings, and B Poolman. Specificity mutants of the binding protein of the oligopeptide transport system of lactococcus lactis. *J Bacteriol*, 182(6):1600–8, Mar 2000.

[67] A Plovins, A M Alvarez, M Ibañez, M Molina, and C Nombela. Use of fluorescein-di-beta-d-galactopyranoside (fdg) and c12-fdg as substrates for beta-galactosidase detection by flow cytometry in animal, bacterial, and yeast cells. *Appl Environ Microbiol*, 60(12):4638–41, Dec 1994.

[68] G Prod'hom, B Lagier, V Pelicic, A J Hance, B Gicquel, and C Guilhot. A reliable amplification technique for the characterization of genomic dna sequences flanking insertion sequences. *FEMS Microbiol Lett*, 158(1):75–81, Jan 1998.

[69] B B Quimby, T Lamitina, S W L'Hernault, and A H Corbett. The mechanism of ran import into the nucleus by nuclear transport factor 2. *J Biol Chem*, 275(37):28575–82, Sep 2000.

[70] Sridharan Raghavan, Paolo Manzanillo, Kaman Chan, Cole Dovey, and Jeffery S Cox. Secreted transcription factor controls mycobacterium tuberculosis virulence. *Nature*, 454(7205):717–21, Aug 2008.

[71] N Rastogi, E Legrand, and C Sola. The mycobacteria: an introduction to nomenclature and pathogenesis. *Rev Sci Tech*, 20(1):21–54, Apr 2001.

[72] Celia Rodriguez, Paula Tejera, Braulio Medina, Rosa Guillen, Angel Dominguez, Jose Ramos, and Jose M Siverio. Ure2 is involved in nitrogen catabolite repression and salt tolerance via ca2+ homeostasis and calcineurin activation in the yeast hansenula polymorpha. *J Biol Chem*, Sep 2010.

[73] Burkhard Rost, Guy Yachdav, and Jinfeng Liu. The predictprotein server. *Nucleic Acids Res*, 32(Web Server issue):W321–6, Jul 2004.

[74] B Rowland, A Purkayastha, C Monserrat, Y Casart, H Takiff, and K A McDonough. Fluorescence-based detection of lacz reporter gene expression in intact and viable bacteria including mycobacterium species. *FEMS Microbiol Lett*, 179(2):317–25, Oct 1999.

[75] David G Russell, Henry C Mwandumba, and Elizabeth E Rhoades. Mycobacterium and the coat of many lipids. *J Cell Biol*, 158(3):421–6, Aug 2002.

[76] David G Russell, Brian C VanderVen, Wonsik Lee, Robert B Abramovitch, Mi-jeong Kim, Susanne Homolka, Stefan Niemann, and Kyle H Rohde. Mycobacterium tuberculosis wears what it eats. *Cell Host Microbe*, 8(1):68–76, Jul 2010.

[77] C M Sassetti, D H Boyd, and E J Rubin. Comprehensive identification of conditionally essential genes in mycobacteria. *Proc Natl Acad Sci U S A*, 98(22):12712–7, Oct 2001.

[78] Christopher M Sassetti and Eric J Rubin. Genetic requirements for mycobacterial survival during infection. *Proc Natl Acad Sci U S A*, 100(22):12989–94, Oct 2003.

[79] Eva Schinko, Klaus Schad, Sema Eys, Ullrich Keller, and Wolfgang Wohlleben. Phosphinothricin-tripeptide biosynthesis: an original version of bacterial secondary metabolism? *Phytochemistry*, 70(15-16):1787–800, 2009.

[80] E Schneider. Abc transporters catalyzing carbohydrate uptake. *Res Microbiol*, 152(3-4):303–10, 2001.

[81] B G Shearer. Mdr-tb. another challenge from the microbial world. *J Am Dent Assoc*, 125(1):42–9, Jan 1994.

[82] H A Shuman. Active transport of maltose in escherichia coli k12. role of the periplasmic maltose-binding protein and evidence for a substrate recognition site in the cytoplasmic membrane. *J Biol Chem*, 257(10):5455–61, May 1982.

[83] S H Sleigh, J R Tame, E J Dodson, and A J Wilkinson. Peptide binding in oppa, the crystal structures of the periplasmic oligopeptide binding protein in the unliganded form and in complex with lysyllysine. *Biochemistry*, 36(32):9747–58, Aug 1997.

[84] Noel H Smith, R Glyn Hewinson, Kristin Kremer, Roland Brosch, and Stephen V Gordon. Myths and misconceptions: the origin and evolution of mycobacterium tuberculosis. *Nat Rev Microbiol*, 7(7):537–44, Jul 2009.

[85] Sander H J Smits, Marina Höing, Justin Lecher, Mohamed Jebbar, Lutz Schmitt, and Erhard Bremer. The compatible-solute-binding protein opuac from bacillus subtilis: ligand binding, site-directed mutagenesis, and crystallographic studies. *J Bacteriol*, 190(16):5663–71, Aug 2008.

[86] Akos Somoskovi, Jillian Dormandy, Linda M Parsons, Michel Kaswa, Khye Seng Goh, Nalin Rastogi, and Max Salfinger. Sequencing of the pnca gene in members of the mycobacterium tuberculosis complex has important diagnostic applications: Identification of a species-specific pnca mutation in "mycobacterium canettii" and the reliable and rapid predictor of pyrazinamide resistance. *J Clin Microbiol*, 45(2):595–9, Feb 2007.

[87] R Srivastava, D Kumar, P Subramaniam, and B S Srivastava. beta-galactosidase reporter system in mycobacteria and its application in rapid antimycobacterial drug screening. *Biochem Biophys Res Commun*, 235(3):602–5, Jun 1997.

[88] Timothy P Stinear, Torsten Seemann, Sacha Pidot, Wafa Frigui, Gilles Reysset, Thierry Garnier, Guillaume Meurice, David Simon, Christiane Bouchier, Laurence Ma, Magali Tichit, Jessica L Porter, Janine Ryan, Paul D R Johnson, John K Davies, Grant A Jenkin, Pamela L C Small, Louis M Jones, Fredj Tekaia, Françoise Laval, Mamadou Daffé, Julian Parkhill, and Stewart T Cole. Reductive evolution and niche adaptation inferred from the genome of mycobacterium ulcerans, the causative agent of buruli ulcer. *Genome Res*, 17(2):192–200, Feb 2007.

[89] Nir Tal and Shimon Schuldiner. A coordinated network of transporters with overlapping specificities provides a robust survival strategy. *Proc Natl Acad Sci U S A*, 106(22):9051–6, Jun 2009.

[90] Laurent Terradot, Richard Bayliss, Clasien Oomen, Gordon A Leonard, Christian Baron, and Gabriel Waksman. Structures of two core subunits of the bacterial type iv secretion system, virb8 from brucella suis and comb10 from helicobacter pylori. *Proc Natl Acad Sci U S A*, 102(12):4596–601, Mar 2005.

[91] Parvathi Tiruviluamala and Lee B Reichman. Tuberculosis. *Annu Rev Public Health*, 23:403–26, 2002.

[92] Fritz Titgemeyer, Johannes Amon, Stephan Parche, Maysa Mahfoud, Johannes Bail, Maximilian Schlicht, Nadine Rehm, Dietmar Hillmann, Joachim Stephan, Britta Walter, Andreas Burkovski, and Michael Niederweis. A genomic view of sugar transport in mycobacterium smegmatis and mycobacterium tuberculosis. *J Bacteriol*, 189(16):5903–15, Aug 2007.

[93] Enrico Tortoli. The new mycobacteria: an update. *FEMS Immunol Med Microbiol*, 48(2):159–78, Nov 2006.

[94] Robert Van der Geize, Katherine Yam, Thomas Heuser, Maarten H Wilbrink, Hirofumi Hara, Matthew C Anderton, Edith Sim, Lubbert Dijkhuizen, Julian E Davies, William W Mohn, and Lindsay D Eltis. A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into mycobacterium tuberculosis survival in macrophages. *Proc Natl Acad Sci U S A*, 104(6):1947–52, Feb 2007.

[95] T van der Heide, M C Stuart, and B Poolman. On the osmotic signal and osmosensing mechanism of an abc transport system for glycine betaine. *EMBO J*, 20(24):7022–32, Dec 2001.

[96] Julia C van Kessel and Graham F Hatfull. Recombineering in mycobacterium tuberculosis. *Nat Methods*, 4(2):147–52, Feb 2007.

[97] Julia C van Kessel and Graham F Hatfull. Mycobacterial recombineering. *Methods Mol Biol*, 435:203–15, 2008.

[98] Alexander J Webb, Karen A Homer, and Arthur H F Hosie. Two closely related abc transporters in streptococcus mutans are involved in disaccharide and/or oligosaccharide uptake. *J Bacteriol*, 190(1):168–78, Jan 2008.

[99] Arnim Weber, Stephanie A Kögl, and Kirsten Jung. Time-dependent proteome alterations under osmotic stress during aerobic and anaerobic growth in escherichia coli. *J Bacteriol*, 188(20):7165–75, Oct 2006.

[100] Mark J White, Hongjun He, Renee M Penoske, Sally S Twining, and Thomas C Zahrt. Pepd participates in the mycobacterial stress response mediated through mprab and sige. *J Bacteriol*, 192(6):1498–510, Mar 2010.

[101] Janet M Wood. Osmosensing by bacteria. *Sci STKE*, 2006(357):pe43, Oct 2006.

[102] Qiao Zhao, Sara Leung, Anita H Corbett, and Iris Meier. Identification and characterization of the arabidopsis orthologs of nuclear transport factor 2, the nuclear import factor of ran. *Plant Physiol*, 140(3):869–78, Mar 2006.