Stony Brook University

Exploration of Lipid Biosynthesis and Transport Mechanisms in Mycobacteria

A Dissertation Presented

by

Megan A. Touchette

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Chemistry

Stony Brook University

August 2015

Stony Brook University
The Graduate School

Megan A. Touchette

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation.

_____
Jessica C. Seeliger – Dissertation Advisor
Assistant Professor of Pharmacological Sciences, Stony Brook University

_____
Isaac Carrico - Chairperson of Defense
Associate Professor of Chemistry, Stony Brook University

_____
Peter Tonge – Committee Member of Defense
Professor of Chemistry, Stony Brook University

_____
David Thanassi – Outside Member of Defense
Professor of Molecular Genetics and Microbiology, Stony Brook University

This dissertation is accepted by the Graduate School

_____
Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Exploration of Lipid Biosynthesis and Transport Mechanisms in Mycobacteria

by

Megan A. Touchette

Doctor of Philosophy

in

Chemistry

Stony Brook University

2015

The inherent drug resistance of Mycobacterium tuberculosis (Mtb) has been attributed in part to the low permeability of its asymmetric outer membrane, which consists of extremely long-chain mycolic acids intercalated with a variety of distinctive surface lipids such as polyacyltrehalose (PAT) and phthiocerol dimycocerosate (PDIM). The biosynthesis of these virulence-associated lipids requires acylation by polyketide-associated protein acyltransferases (PapA). We show that PapA5 is the terminal enzyme in PDIM biosynthesis by demonstrating its dual esterification activity and chain-length preference with synthetic alkyl beta-diol substrate analogues. We also demonstrate that PapA5 is phosphorylated by Mtb kinases PknB and PknE which suggests that post-translational modification may regulate PDIM biosynthesis. Analogously, PapA3 acylates trehalose to form diacyltrehalose (DAT), but the downstream enzyme(s) that generate the final product, PAT, have not been identified. We provide evidence that the serine hydrolase Chp2 is essential for the final steps of PAT biosynthesis and its activity appears to be regulated by the putative lipid transporter MmpL10.

In comparison to biosynthesis, little is known about lipid transport between the inner and outer membranes. In order to identify proteins that may be involved, we initially explored a chemoenzymatic strategy to generate reactive crosslinking lipid analogues. We examined the enzymatic synthesis of various acyl-CoAs by the fatty acyl-CoA ligase FadD6 as in vitro substrates for PapA family enzymes. While FadD6 is capable of generating substrates in situ for PapA3 and PapA5, this is not an ideal method for generating lipid analogues due to low product yield. In other work, we characterized LppX, which has been implicated in transporting PDIM to the outer membrane. We developed a competitive fluorescence reporter assay to characterize LppX-ligand binding and thereby provide insights into PDIM recognition by LppX. LprG, a homolog of LppX, is also required for Mtb virulence and has been implicated in triacylglyceride (TAG) transport and lipoarabinomannan (LAM) display at the cell surface. We hypothesized that LprG interacts with other cell wall proteins in the process of transporting lipids between membranes. Using complementary crosslinking methods, we identified several LprG-interacting proteins that may contribute to lipid trafficking and corroborate LprG involvement in mycobacterial membrane biogenesis.

Dedication

To my loving husband, James, your unwavering love, patience, encouragement and support have driven me through these past five years. I could not have done it without you. I love you forever and for always.

To my parents, Patrick and Ann Himmler, you have loved me unconditionally, taught me to work hard for the things I aspire to achieve and inspired me to follow my heart. I love you and cannot thank you enough for all you have done for me.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACN	acetonitrile
AcOEt	ethyl acetate
AG	arabinogalactan
ANS	8-anilinonaphthalene-1-sulfonate
AP	alkaline phosphatase
Ara	arabinan
ATP	adenosine triphosphate
CD	circular dichroism
CE	cell envelope
CL	cardiolipin
CoA	coenzyme A
DAT	diacyltrehalose
DC-SIGN	dendritic cell-specific intercellular adhesion molecule-3 grabbing non integrin
DCC	N,N'-dicyclohexylcarbodiimide
DCM	methylene chloride
DDM	*N*-dodecyl-β-D-maltoside
DHB	2,5-dihydroxybenzoic acid
DIPEA	*N*,*N*-Diisopropylethylamine
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DSP	dithiobis(succinimidylpropionate)
DTNB	5,5'-dithiobis-(2-nitrobenzoic acid)
DTT	Dithiothreitol

EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
ESI-MS	electrospray ionization-mass spectrometry
FA	formic acid
FABP	fatty acid binding protein
FACL	fatty-acyl CoA synthase
FASII	fatty acid synthase II
Gal	galactose
GFP	green fluorescent protein
HIV	human immunodeficiency virus
ICMT	isotope-coded mass tagging
IM	inner membrane
IPTG	isopropyl β-D-thiogalactopyranoside
kDa	kilo Daltons
LAM	lipoarabinomannan
LM	lipomannan
LPS	lipopolysaccharide
Lpt	LPS transport
MALDI-TOF	matrix assisted laser desorption/ionization-time of flight
ManLAM	mannose-capped lipoarabinomannan
Mas	mycocerosic acid synthase
MBP	maltose-binding protein
MDR-TB	multidrug-resistant tuberculosis
MES	2-(*N*-Morpholino)ethanesulfonic acid
Msm	*Mycobacterium smegmatis*

Mtb	*Mycobacterium tuberculosis*
MW	molecular weight
MWCO	molecular weight cutoff
MyBP	myelin basic protein
NADH	nicotinamide adenine dinucleotide
NBD	nucleotide binding domain
NBD	7-nitrobenzo-2-oxa-1,3-diazole (only referred to in Chapter 5)
Ni-NTA	Nickel-nitrilotriacetic acid
NMR	nuclear magnetic resonance
OCT	1-octanol
OD_{600}	optical density at 600 nm
OM	outer membrane
PA	palmitic acid
PapA	polyketide associated protein acyltransferase
PAT	polyacyltrehalose (also known as pentaacyl or polyphthienoyl trehalose)
pBpa	*para* benzoyl phenylalanine
PBS	phosphate-buffered saline
PCoA	palmitoyl coenzyme A
PCR	polymerase chain reaction
PDB	protein data bank
PDD	(5*R*,7*R*)-pentadecane-5,7-diol
PDIM	phthiocerol dimycocerosate
PE	phosphatidylethanolamine
PG	peptidoglycan
GPL	glycosylphenolphthiocerol dimycocerosates

PGL	phenolic glycolipid
PI	phosphatidylinositol
PIM	phosphatidylinositol mannoside
PIPES	piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)
Pks	polyketide synthase
PMSF	phenylmethanesulfonylfluoride
PSNAC	palmitoyl-*N*-acetylcysteamine
PTG	phosphatidylglycerol
RHD	3*R*,5*R*-heptane-3,5-diol
RNS	reactive nitrogen species
SDC	sodium deoxycholate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SHD	3*S*,5*S*-heptane-3,5-diol
SILAC	stable isotope labeling by amino acids in culture
SL-1	sulfolipid-1
SNAC	*N*-acetylcysteamine
T2P	trehalose-2-palmitate
TAG	triacylglyceride
TAT	triacyltrehalose
TB	tuberculosis
TCEP	tris(2-carboxyethyl)phosphine
TDM	trehalose dimycolate
TetraAT	tetraacyltrehalose
TEV	Tobacco Etch Virus
TFA	trifluoroacetic acid

THF	tetrahydrofuran
THL	tetrahydrolipstatin
TLC	thin layer chromatography
TLR2	Toll-like receptor 2
TM	transmembrane
TMD	transmembrane domain
tRNA	transfer ribonucleic acid
TyrRS	tyrosyl-tRNA synthetase
UAAI	unnatural amino acid incorporation
UDD	(5*R*,7*R*)-undecane-5,7-diol
UPLC	ultra performance liquid chromatography
UV	ultraviolet
WT	wild-type
XDR-TB	extensively drug-resistant tuberculosis
βGal	β-galactosidase

Acknowledgements

I would like to express my sincere appreciation and gratitude for my advisor, Jessica C. Seeliger, for her constant support and guidance. Her enthusiasm and encouragement motivated me and helped to mold me into the independent scientist I am today.

I would also like to thank my committee, Isaac Carrico, Peter Tonge and David Thanassi for all of the helpful discussion and suggestions. It is greatly appreciated.

I would like to express my appreciation for all members of the Seeliger lab both past and present for their helpful discussions but most importantly their day to day support and encouragement.

I would like to thank Cynthia M. Holsclaw, Mary L. Previti, Viven C. Solomon, Julie A. Leary, Carolyn R. Bertozzi, and Jessica C. Seeliger for their work on characterizing Chp2.

I would like to thank Gopal R. Bommineni, Richard J. Delle Bovi, John E. Gadbery, Carrie D. Nicora, Anil K. Shukla, Jennifer E. Kyle, Thomas O. Metz, Dwight W. Martin, Nicole S. Sampson, W. Todd Miller, Peter J. Tonge, and Jessica C. Seeliger for their help in characterizing PapA5.

I would like to thank my mentee, Julia Joseph for her hard work developing the competitive binding fluorescence assay for studying lipid-protein interactions.

I would like to thank Mary L. Previti and Hiren V. Patel for their help with the crosslinking in live cells project as their help with cloning and initial troubleshooting helped propel this project forward in a short time.

I would like to thank Teresa Garrett, Bela Ruzsicska, Markus Seeliger, Christina Baer, Lauren Spagnuolo, and Nadine Henderson for technical assistance, reagents, and helpful discussions.

Publications

- **Touchette, M. H.**, Holsclaw, C. M., Previti, M. L., Solomon, V. C., Leary, J. A., Bertozzi, C. R., Seeliger, J. C. The *rv1184c* Locus Encodes Chp2, an Acyltransferase in *Mycobacterium tuberculosis* Polyacyltrehalose Lipid Biosynthesis. *Journal of Bacteriology* **2015**, 197 (1), 201-210.

- **Touchette, M. H.,** Bommineni, G. R., Delle Bovi, R., Gadbery, J. E., Nicora, C. D., Shukla, A., K., Kyle, J. E., Metz, T. O., Martin, D. W., Sampson, N. S., Miller, W. T., Tonge, P. J., Seeliger, J. C. Diacyltransferase Activity and Chain Length Specificity of *Mycobacterium tuberculosis* PapA5 in the Synthesis of Alkyl Beta-Diol Lipids, *Biochemistry,* **just accepted**.

Chapter 1 Introduction

Tuberculosis is a serious, worldwide public health threat

Mycobacterium tuberculosis (*Mtb*), the bacterium that causes tuberculosis (TB) in humans, is a serious worldwide public health threat. It infects an estimated one third of the world's population and TB is the second leading cause of death worldwide from a single infectious agent. In 2013 there were approximately 9 million new TB cases, 1.1 million of which were co-infections with human immunodeficiency virus (HIV) and approximately 1.5 million deaths from TB (1,2). TB is a treatable disease, but it requires a 6-9 month therapeutic regimen consisting of multiple antibiotics. The majority of those infected live in developing countries and the lengthy chemotherapy results in poor patient compliance (3). In recent years, multidrug-resistant (MDR) and extensively drug-resistant (XDR) *Mtb* strains have emerged and are likely the result of poor patient compliance. In 2013, 480,000 people developed MDR-TB and there were 210,000 deaths from MDR-TB infection (1). The public health crisis developing from these drug-resistant strains necessitates the development of novel treatments (4).

Mtb is spread via inhalation of bacilli-containing aerosols and therefore the primary site of infection is the host lung (5-7). The bacteria are phagocytosed by alveolar macrophages and subsequently inhibit phagosome-lysosome fusion, preventing acidification of the vacuole and allowing for an environment in which the bacteria can replicate (7). The macrophages respond by producing cytokines and chemokines to recruit T-cells and neutrophils. The accumulation of these immune cells initiates the formation of a granuloma (8). Granulomas are a tissue structure made up of different cell types. The core is made of infected macrophages, which are surrounded by uninfected macrophages and monocytes. This is encircled by a layer of extracellular components

and collagen. Finally, the whole structure is surrounded by T and B lymphocytes. TB granulomas tend to be necrotic and are in such cases referred to as caseous due to the accumulation of cellular material in the center of the granuloma (9). Mycobacteria in the granuloma are thought to be protected as they enter into a latent state in which replication slows and basic bacterial metabolism changes. *Mtb* can persist in this state for decades and only 10% of those latently infected will reactivate into active, infectious disease. This conversion can be precipitated by a change in immune status (10,11).

TB remains difficult to treat due to *Mtb*'s inherent drug resistance mechanisms, which can be attributed to drug efflux and low permeability of the cell wall (12-14). The low permeability of the mycobacterial cell wall is thought to be advantageous under stressful conditions (15). The mycobacterial cell wall is essential for *Mtb* growth and survival in host macrophages and the most effective drugs for treating TB inhibit cell wall biosynthesis processes (13). For example, the current front-line drugs isoniazid and ethambutol inhibit cell wall fatty acid and carbohydrate biosynthesis (13,16,17).

Mycobacterial cell envelope

In contrast to other bacteria, mycobacteria are characterized by a dense, multilayer cell wall. They are classified as Gram-positive, but their cell wall resembles that of diderm Gram-negative bacteria. Mycobacteria belong to a group of Gram-positive bacteria with GC-rich DNA and are referred to as actinomycetes. Within this group, mycobacteria fall on the *Corynebacterium-Mycobacterium-Nocardia* branch and this group of bacteria is known to have cell walls with unique composition and structure (13).

The major lipid components of the inner membrane in mycobacteria are the phospholipids cardiolipin (CL), phosphatidylethanolamine (PE), phosphatidylinositol (PI) and glycosylated PIs such as phosphatidylinositol mannoside (PIM), lipomannan (LM) and lipoarabinomannan (LAM). The relative amounts of these lipids varies between species. For instance, it was reported that the *M. smegmatis* (*Msm*) inner membrane (IM) consists of approximately 37% CL, 32% PE and 28% PI/PIMs and *M. phlei* IM consists of about 50% CL, 10% PE and 20% PI/PIMs (18,19). PI and glycosylated PIs are essential for viability in *Mtb* and *Msm* (18,20). Although common in other bacteria, phosphatidylglycerol is a relatively minor IM component in mycobacterial species.

The low permeability of the cell wall contributes to *Mtb* resistance to current therapeutics is due to the presence and variety of lipids in the outer membrane, especially the mycolic acid layer (Figure 1-1). The mycobacterial cell wall is a non-fluid barrier and is responsible for the low permeability (21-23). The mycobacterial cell wall is comprised of three covalently linked layers and a non-covalently associated outer leaflet of the OM (Figure 1-1). The peptidoglycan (PG) polymer that most bacterial species have surrounding their cytosolic inner membrane (IM) is covalently attached to a layer of arabinogalactan (AG) which is esterified at its terminal, non-reducing ends with extremely long $C_{60\text{-}90}$ chain fatty acids known as mycolic acids (Figure 1-1) (24-26). PG, AG and mycolic acid precursors are synthesized in the cytoplasm as glyco- or isoprenyl-phosphate conjugated monomers that are transported across the IM. The final polymerization and crosslinking steps, occur in the periplasm (27).

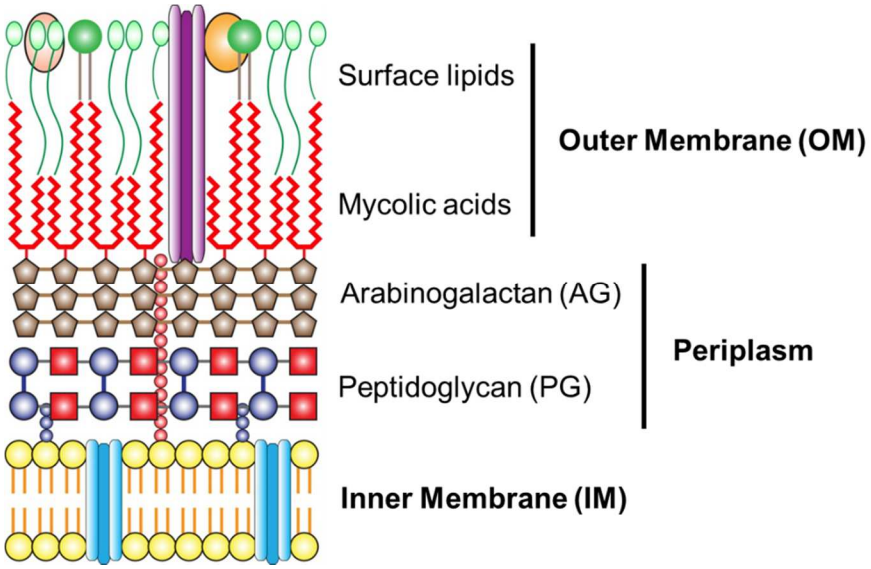


Figure 1-1. Cartoon depicting the mycobacterial cell envelope. The term cell wall refers to the periplasm and OM components of the cell envelope.

PG is composed of repeating *N*-acetyl-β-D-glucos-aminyl-(1→4)-*N*-acyl-muramic acid monomers. A short polypeptide chain consisting of L-alanine, D-glutamic acid, *meso*-diaminopimelic acid and D-alanine is attached to the muramic acid moieties. Mycobacterial peptidoglycan differs from the most common PG form found in bacteria in two ways. First, the muramic acid is *N*-glycosylated whereas in other bacterial species, PG muramic acid is *N*-acetylated. Second, each monomer peptide is crosslinked to an adjacent monomer peptide between two diaminopimelic acids as well as between diaminopimelic acid and D-alanine whereas in other bacterial species, only the latter crosslink is common (28). The AG layer is attached to PG via phosphodiester linkages to approximately 10-12% of the muramic acid residues (24).

AG is not a common periplasmic component in most bacterial species and is unique in its structure and its sugar components. The arabinan (Ara) and galactose (Gal) residues of AG are the furanose form and the Ara units are crosslinked via the C5 and sometimes the C6 of the Gal units.

The galactan region is comprised of crosslinked Gal residues via alternating C5- and C6-linkages. The galactan also crosslinks to some of the muramic acids of PG. The arabinan chain consists of five Ara residues, referred to as the hexaarabinofuranosyl unit. Finally, about 66% of these units are esterified with mycolic acids, which cluster in groups of four (25).

Mycolic acids are α-alkyl-β-hydroxy fatty acids that are primarily covalently attached to the hexaarabinofuranosyl unit of the AG but also act as acyl chains for extractable surface lipids such as trehalose-6,6′-dimycolate (TDM). Mycolic acids are produced from a head-to-head condensation of two fatty acyl parts, a meromycolic acid and the α-chain. The meromycolic acid is 50-60 carbons in length and the α-chain is 24-26 carbons in length and is generally less functionalized that the meromycolic acid. Mycobacterial mycolic acids contain one or two unsaturated C-C bonds or cyclopropane rings on the meromycolate chain. They may also have multiple methyl branches along the meromycolate chain (29-33).

Outer membrane surface lipids and their role in pathogenesis

Another distinct feature of the mycobacterial cell wall is the presence and variety of surface lipids (26). Surface lipids are of special interest because they are at the interface of interaction between the bacterium and the host. They have been previously described as signaling effector molecules for *Mtb* (34). Inhibiting cell wall biosynthesis processes has been an effective therapeutic route in the past. Targeting biosynthesis of surface membrane lipids is hypothesized to be an effective therapeutic because in addition to compromising the integrity of the cell wall, the surface for interaction with the host immune system may be altered. Thought to be intercalated with the covalently attached mycolic acids, several classes of “free lipids” can be extracted from

intact cells and are considered surface lipids. The loss of some classes of these lipids attenuates *Mtb* virulence (35).

LAM and LM are multiglycosylated extensions of PIMs, are considered to be virulence factors of *Mtb*, and are agonists of the innate immune receptor Toll-like receptor 2 (TLR2). Pathogenic mycobacterial LAM is mannose-capped and referred to as ManLAM. (36,37). LAM, LM and PIMs are surface-exposed antigens that bind to the lipoprotein LprG and, upon binding, enhance recognition of triacylated glycolipids by TLR2 (36,38). Another class of surface lipids is glycosylphenolphthiocerol dimycocerosates (PGLs). PGLs comprise a C_{36} anti-diol modified with an oligosaccharide and esterified with two mycocerosic acids. The oligosaccharide consists of one to four deoxy sugars (39). Structurally related to PGLs are a class of wax ester OM surface lipids that are exclusive to slow-growing pathogenic mycobacteria. They generally comprise a long chain β-diol (phthiocerol A/B, phthiodiolone, or phthiotriol) esterified at both hydroxyl moieties with mycocerosic acids. This class of lipids is referred to as DIMs, as they all are dimycocerosates modified diols. Clinical *Mtb* strains that lack DIM and mutants defective in DIM biosynthesis or transport are attenuated for virulence in animal-based infection models (40-43). Another class of OM lipids are the trehalose-based lipids. Trehalose is modified with 2-5 acyl chains with a variety of structures. Mycocerosic acid-modified trehalose (TDM) or cord factor, diacyltrehalose (DAT) and pentaacyltrehalose (PAT) are a few examples. DAT and PAT are modified with saturated fatty acids as well as methylated and unsaturated chains. These lipids have immunosuppressive functions (44-46) and are thought to contribute to phagosome-lysosome fusion arrest (47). Finally, the sulfolipid class of OM lipids also have virulence-associated phenotypes (48). Trehalose-2′-sulfate are esterified with multimethyl-branched hydroxyphthioceranic and phthioceranic fatty

acids and C_{16}-C_{18} saturated straight chain fatty acids to form sulfolipids such as sulfolipid-1 (SL-1).

The work presented here focuses on the mycobacterial outer membrane (OM) lipids PDIM, PAT, SL-1, and LAM. PDIM, PAT and SL-1 are of particular interest because they are mycobacteria-specific, associated with therapeutic resistance and each esterified by different members from the same enzyme family (PapA) that all share acyltransferase activity. These lipids are synthesized largely in the cytosol; these pathways are described in greater detail in Chapters 2 and 3. They are then transported across the IM and through the periplasm and OM where they are assembled. Chapters 5 and 6 detail projects geared towards understanding these transport processes and therefore, broadly, membrane lipid biogenesis. The lipoprotein LppX has been implicated in PDIM translocation to the OM and LprG has been implicated in LAM cell surface display (49,50).

Mycobacteria are phagocytosed via several host cell receptors, including complement receptors, scavenger receptors, C-type lectins and several mycobacterial glycolipids serve as ligands for these receptors (51). Complement receptors recognize mycobacterial glycolipids via their lectin domain (52). The C-type lectins mannose receptor and dendritic cell-specific intercellular adhesion molecule-3 grabbing non integrin (DC-SIGN) both bind the mannose cap of ManLAM, therefore specifically recognizing pathogenic mycobacteria (53-56).

Mycobacterial lipids have also been shown to play crucial roles in arresting phagosome maturation. ManLAM and DIM do so by intercalating into host cell membranes whereas TDM intercalation does not seem to be necessary (57,58). The mechanism by which ManLAM plays a role in arresting phagosome acidification is not fully understood but is associated with recognition by mannose receptors and it is suggested that ManLAM inhibits steps involved in H^+-ATPase

delivery via the phosphatidylinositol-3-phosphate-dependent pathway (59-61). Several gene involved in DIM, acyltrehalose and sulfoglycolipid biosynthesis have been identified as being involved in phagosome maturation arrest but the mechanisms by which this occurs is not fully understood (41,42,57,62,63). Sulfoglycolipids have been implicated in the modulation of host cell responses, including phagosome-lysosome fusion and cytokine production but this is also not well-understood (64-67).

In addition to mediating phagocytosis and being involved in phagosome maturation arrest, mycobacterial lipids affect host cell signaling and secretion of inflammatory cytokines necessary for granuloma formation. TDM is a C-type lectin and scavenger receptor ligand that induces proinflammatory cytokine production leading to granuloma formation (68-70). The *Mtb* glycolipids LAM and LM interact with pattern recognition receptors, such as Toll-like receptor-2 (TLR-2), inducing TNF, an inflammatory cytokine (36,37). Mycobacterial glycolipids interact with multiple host cell receptors and elicit co-signaling, which is critical for granuloma formation (71,72). *Mtb* glycolipids are not only involved in innate immune responses but are also antigens to the adaptive immune system as T cells presentation of diacyl sulfoglycolipid has been observed (73,74). Further work is necessary to identify other specific lipid antigens for the adaptive immune response.

Bacterial elongation and division via polar extension

Due to the complexity of the cell wall, elongation and cell division are a challenge to mycobacteria as they require unique mechanisms for cell wall biogenesis (15).The cell wall is an essential bacterial component as it dictates cell shape and acts as a barrier to exogenous insults and

its synthesis allows cell growth and division. The major cell wall component PG is central to cell wall integrity, is recycled and undergoes multiple rounds of synthesis during each cell cycle. The PG synthase PonA1 is a penicillin binding protein with a transmembrane (TM) helix and two periplasmic catalytic domains with transpeptidation and transglycosylation activities, both of which are critical for normal proliferation (75,76).

Bacterial growth and cell division is determined by the synthesis and crosslinking of its cell wall and this mechanism varies between organisms. In *Escherichia coli* and *Bacillus subtilis*, both rod-shaped bacteria, elongation occurs by deposition of new PG along the cylindrical part of the cell body (77,78). In contrast, mycobacteria, also rod-shaped bacteria, have been shown to extend via polar extension at a significantly more variable rate than in *E. coli* (79,80). Mycobacteria preferentially elongate at the "old pole", the pole that is inherited from the original cell. Therefore one daughter cell inherits the growing pole and the other is left to generate its own. The sister cells elongate at different rates and are different sizes at separation; the sister that inherited the growing pole elongates faster and is longer. These cells are also differentially susceptible to antibiotics. The faster-growing cells are more sensitive to antibiotics that target PG biosynthesis and the slower growing cells are more susceptible to antibiotics that inhibit mycolic acid synthesis (Figure 1-2) (80).

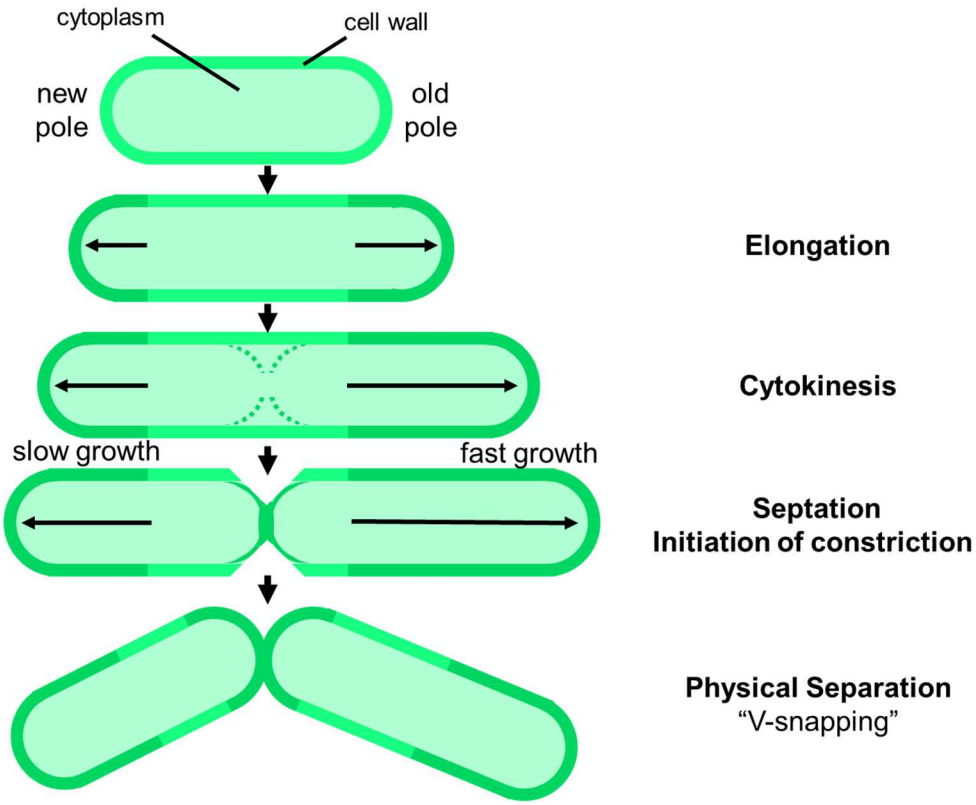


Figure 1-2. Mycobacterial elongation and division results in asymmetric daughter cells. Elongation occurs via cell wall deposition at the bacterial poles. All cells contain an old pole, the pole inherited from its mother cell, and a new pole. Elongation occurs preferentially and more rapidly at the old pole (80). Cell division takes place in several steps. First, IM synthesis occurs during cytokinesis in the center of the cell. PG is synthesized at the septum (a process known as septation). Finally, the cell wall constricts and hydrolysis occurs at the septum for the daughter cells to asymmetrically and physically separate. This figure has been adapted from a recent review on mycobacterial growth and division (75).

This process of cellular elongation in *E. coli* and *B. subtilis* requires the actin-like protein MreB which organizes the cell wall synthesis machinery by directly binding to the membrane via an N-terminal α-helix (81). Mycobacteria, however, do not encode an obvious MreB ortholog, indicating that, like elongation, they organize cell wall synthesis in a different manner. Actinobacteria, in which mycobacteria comprise a genus (82), encode a tropomyosin-like protein, DivIVA, that is required for cell growth and has been shown to localize at the poles and septa

(83,84). Meniche and co-workers showed that DivIVA directs cell wall synthetic complexes to the growing pole in mycobacteria. Negative membrane curvature is sufficient for DivIVA N-terminus-mediated recognition of the pole (85). However, *B. subtilis* DivIVA was shown to interact with liposomes composed of pure phospholipids, indicating that DivIVA can bind phospholipid-containing membranes in the absence of negative curvature. This membrane-binding capability was abrogated in DivIVA constructs that lacked the N-terminus or lacked the hydrophobic residues at the N-terminus. Across species, there is a conserved 20-amino acid amphiphilic helix at the N-terminus, suggesting this may be a conserved mechanism for DivIVA membrane binding (86). These data indicate both membrane binding and recognition of negative curvature as important factors in DivIVA polar localization.

Cell wall synthesis proteins co-localize at the poles and septa

MurG, Glf2 and Pks13, which are PG, AG and mycolic acid biosynthesis proteins, respectively, co-localize asymmetrically at the poles and the septa along with DivIVA. The pole at which they concentrate was shown to be the growing pole. Throughout cell division, they remain localized at the septa and after division, stayed associated with the new pole albeit at a lower concentration at than the growing, old pole. Although co-localized at the pole, MurG, Glf2 and Pks13 are found in a region distinct from where DivIVA binds and has been described as being the subpolar space. FadD32, the acyl-AMP ligase involved in one of the final mycolic acid biosynthetic steps, was shown to interact with Pks13 and therefore also co-localize to the subpolar space. DivIVA was shown to physically interact with AccA3, AccD4 and AccD5 of the acyl-CoA carboxylase (ACC) complex, enzymes which are involved early in fatty acid and mycolic acid biosynthesis. The current model of cell elongation is that new cell wall deposition occurs in the

subpolar region, a space distinct from where DivIVA binds and mycolic acid precursor biosynthesis occurs (Figure 1-3) (85).

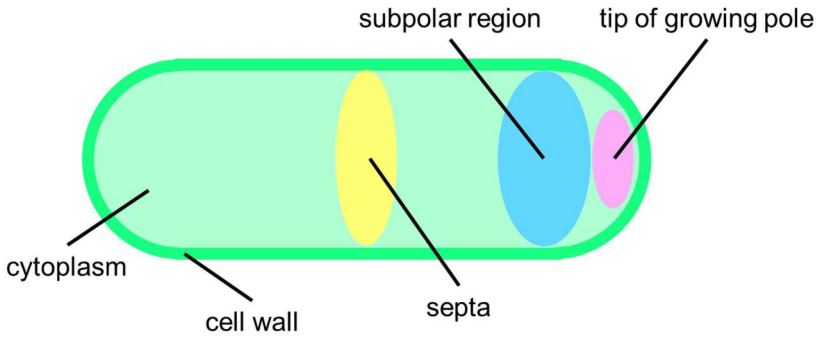


Figure 1-3. Model of mycobacterial cell with subcellular regions relevant to cell growth and division. The tip of the growing pole, subpolar region and septa are designated by the pink, blue and yellow areas respectively. They do not represent membrane-bound organelles, only subcellular location.

Other proteins that have been localized to mycobacterial poles include RipA, RpfB, PonA1, ParA and MmpL3. RipA is an endopeptidase that hydrolyzes PG and forms a complex with RpfB, a lytic transglycosylase. PonA1, a PG synthase, has been identified as a RipA-interactor and is capable of inhibiting the synergistic PG hydrolysis of the RipA-RpfB complex, thereby demonstrating a balance between cell wall degradation and synthesis. This has been shown to occur at the cell poles and septa (87). The *Msm* ATPase ParA mediates chromosome segregation and coordinates this process with cell division and elongation via direct interaction with the *Msm* DivIVA ortholog Wag31 (88). MmpL3, the mycolic acid transmembrane (TM) transporter, has been co-localized with AccA3 and AccD4 to the growing pole of mycobacteria and all three have been identified as DivIVA-interacting proteins (89).

Research Projects Overview

There are several examples in the literature for the colocalization of cell wall synthesis and transport machinery (85,89,90). These functions are essential for bacterial elongation and division and overall survival of the bacterium. It is therefore necessary to understand membrane biogenesis including lipid synthesis and transport.

Herein, I have provided synopses of my work towards gaining a better understanding of the biosynthesis and regulation of the OM lipids PAT and PDIM through the study of their terminal biosynthetic acyltransferases Chp2 and PapA5, respectively. I have also characterized the acyl-CoA ligase FadD6 and was able to show the feasibility of chemoenzymatically generating lipid analogues in conjunction with the acyltransferases PapA3 and PapA5. Regarding lipid transport mechanisms, I developed a competitive fluorescence assay to characterize lipid-protein interactions of the putative lipid transporter LppX. Finally, live-cell crosslinking studies were developed in order to identify the protein-protein interactions for LprG, a putative lipid transporter from the same family as LppX towards developing a model for mycobacterial OM lipid transport.

Chapter 2 The *rv1184c* Locus Encodes Chp2, an Acyltransferase in *Mycobacterium tuberculosis* Polyacyltrehalose Lipid Biosynthesis

This chapter has been published. *Journal of Bacteriology* and all authors have granted permission to use this manuscript in this dissertation:

Touchette, M. H., Holsclaw, C. M., Previti, M. L., Solomon, V. C., Leary, J. A., Bertozzi, C. R., Seeliger, J. C. The *rv1184c* Locus Encodes Chp2, an Acyltransferase in *Mycobacterium tuberculosis* Polyacyltrehalose Lipid Biosynthesis. *Journal of Bacteriology* **2015,** 197 (1), 201-210.

Introduction

In bacteria, glycolipids based on the disaccharide trehalose have been identified in *Mycobacteria*, *Corynebacteria*, *Nocardia*, and *Rhodococcus* (91). These members of the suborder Corynebacterineae have distinctive cell envelopes that are dominated by long branched-chain lipids. Trehalose lipids are found in all these genera, but those with multiply methyl-branched acyl chains are specific to virulent strains of mycobacteria, including *Mycobacterium tuberculosis* (*Mtb*), which causes tuberculosis (92). These lipids include the di-, tri- and pentaacyltrehaloses, which are found in the mycobacterial outer membrane (OM) (93,94). Consistent with their surface location, acyltrehaloses have proposed roles in the attachment of the carbohydrate capsule layer and have been identified as antigens, leading to their exploration as serodiagnostic markers with the potential to distinguish between tuberculous and non-tuberculous mycobacteria (95-98).

The strain specificity of acyltrehaloses implies a role in virulence, and several studies support this idea. Diacyltrehaloses inhibit proliferation of naïve T cells and downregulate cytokine production in activated monocytes, indicative of an immunosuppressive function (44-46). Also, impairment of acyltrehalose production in *Mtb* is associated with acidification of *Mtb*-containing phagosomes (99). Acyltrehaloses may thus contribute to phagosome-lysosome fusion arrest, a key

factor in the survival of *Mtb* in host cells. Indeed, the gene cluster associated with acyltrehalose biosynthesis is upregulated during phagosome acidification (47). Overall, however, the functions of these lipids in *Mtb* physiology and pathogenesis are incompletely understood.

The dominant acyltrehalose in *Mtb* is polyacyltrehalose (PAT; also known as pentaacyl or polyphthienoyl trehalose), which comprises trehalose modified by one straight-chain fatty acid and four mycolipenic (phthienoic) acids (Figure 2-1) (100). The mycolipenic acids are produced by Pks3/4 (also known as Msl3) and disruption of the *pks3/4* gene results in the loss of PAT (94). Due to the cyclical processive mechanism of Pks3/4, the mycolipenic acids vary in the degree of methylation (1-3 methyl branches) as well as the in the length of the saturated alkyl chain (C_{16}-C_{18}) (100,101). This variation gives rise to a diverse range of PAT structures, which is observed by mass spectrometry as a distinctive envelope of ions in the *m/z* 2000-2300 range (102). More recently, *papA3* gene was also shown to be essential for PAT biosynthesis and PapA3 was proposed as an acyltransferase that modifies trehalose with a saturated alkyl chain to form trehalose-2-palmitate (T2P) and then with a mycolipenic acid to form diacyltrehalose (DAT) (102). However, it remains unclear whether PapA3 is in fact a bifunctional enzyme in *Mtb*. One hypothesis is that PapA3 also catalyzes further esterifications to form the final PAT product, but at least *in vitro* this activity was not detected.

As noted in the PapA3 study, the PAT biosynthetic gene cluster closely resembles that of sulfolipid-1 (SL-1), a structurally similar trehalose glycolipid that is also unique to pathogenic mycobacteria. The SL-1 cluster contains two PapA3 homologues, PapA2 and PapA1, that esterify sulfotrehalose with straight-chain and phthioceranoic acids to form diacyl sulfotrehalose (also known as SL-1278) (102,103). Recent work showed that subsequent modifications with phthioceranoic acid are mediated by a third enzyme, Chp1 (104). Chp1 is unrelated to the PapA

acyltransferases, but in predicted structure and catalytic mechanism closely resembles the Antigen85 complex (Ag85A, B and C), which catalyze the transesterification of trehalose monomycolate to form trehalose dimycolate and free trehalose (105).

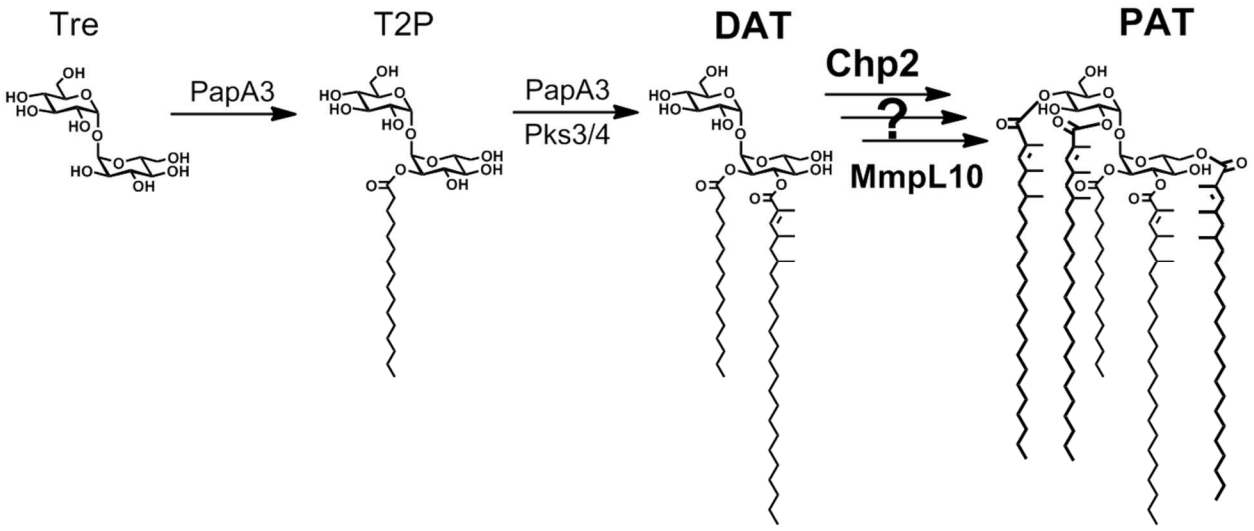


Figure 2-1. The biosynthetic pathway for PAT is incomplete. The acyltransferase PapA3 converts trehalose into T2P and then to DAT using mycolipenic chains produced by Pks3/4. Chp2 and MmpL10 may be involved in the subsequent conversion of DAT into PAT.

The PAT biosynthetic gene cluster includes *rv1184c*, a homologue of *chp1* that we here call *chp2*. The coordinate upregulation of *chp2* and other PAT biosynthetic genes upon phagosome acidification or expression of transcriptional regulators such as PhoP and WhiB3 supports a role for *chp2* in PAT biosynthesis (106,107). Therefore, we hypothesized that Chp2 is an acyltransferase that transforms DAT into PAT. To test this prediction, we characterized the *Mtb Δchp2* knockout strain and showed that it does not produce PAT, but accumulates DAT. We also show that Chp2 is a serine-dependent enzyme that is associated with the cell envelope, but has cytosolic activity that requires the lipid transporter MmpL10. Finally, Chp2 activity is reduced in *Mtb* cells treated with the lipase inhibitor tetrahydrolipstatin (THL).

Experimental Procedures

Bacterial strains, growth media, vector construction and oligonucleotide primers

7H9 and 7H11 growth media and ADC and OADC supplements were obtained from BD Biosciences. Mycobacterial strains used in these studies were *M. tuberculosis* Erdman (ATCC 35801) and *M. smegmatis* (*Msm*) mc^2155 (ATCC 700084). The growth medium was 7H9 (liquid) or 7H11 (solid) with 0.5% glycerol. Additional additives were 0.05% Tween-80 and 10% ADC supplement for *Msm* or 0.025% tyloxapol and 10% OADC for *Mtb*. For selective media, antibiotic concentrations were 100 μg/mL ampicillin, 50 μg/mL kanamycin or 100 μg/mL hygromycin for *E. coli* and 25 μg/mL kanamycin or 50 μg/mL hygromycin for mycobacteria. See Table 2-1 and Table 2-2 for details on vector constructs and oligonucleotide primer sequences.

Table 2-1. Bacterial strains and plasmids used in this study.

Strain	Name	Genotype	Source
M. tuberculosis Erdman			
M. tuberculosis Erdman		Δ*chp2*	This work
M. tuberculosis Erdman		Δ*chp2::chp2*	This work
M. tuberculosis Erdman		Δ*papA3*	Ref (102,108)
M. tuberculosis Erdman		Δ*papA3::papA3*	Ref (102)
M. tuberculosis Erdman	jcm110	Δ*mmpL10*	J. Cox
M. smegmatis mc^2155			

Plasmids

Reference name	Name	Description	Source
	pMV261	Kn^r, pAL5000 origin, ColE1 origin, multiple cloning site, P_{hsp60}	Ref (109)
	pMV306	Kn^r, *int*, attA5, ColE1 origin, multiple cloning site	Ref (110)
	pJSC407	oriE, cosA; HygR flanked by loxP sites	Ref (111)
Δ*chp2* KO construct		pJSC407; HindIII-804bp (first 156bp of *chp1*)-Hyg^r- 813bp (last 157bp of *chp1*)-KpnI	This work
	pRibo-BsaHind	pMV261 derivative; P_{hsp60}-theophylline riboswitch; BsaI for cloning	Ref (112)
Chp2-3xFLAG	pRibo Chp2-3xFLAG	pRibo-BsaHind derivative; contains *chp2* fused to 3xFLAG	This work
AP cytosolic control	pMB111	pMV261 derivative; contains *phoA*	M. Braunstein

AP secreted control	pMB124	pMV261 derivative; contains *fbpB*(aa2-40)-*phoA*	M. Braunstein
βGal cytosolic control	pHsp-*lacZ*	pMWS114 derivative; P_{hsp60}-*lacZ*	Ref (112)
βGal secreted control	sec-lacZ	pMB124 derivative; P_{hsp60}-*fbpB*(aa2-40)-*lacZ*	Ref (112)
full-length Chp2-βGal		pRibo-BsaHind derivative; contains full-length *chp2* fused to *lacZ*	This work
N-term Chp2-βGal		pRibo-BsaHind derivative; contains *chp2* (aa1-39) fused to *lacZ*	This work
cat Chp2-βGal		pRibo-BsaHind derivative; contains *chp2* (aa26-359) fused to *lacZ*	This work

Reference name	**Name**	**Description**	**Source**
full-length Chp2-AP		pRibo-BsaHind derivative; contains full-length *chp2* fused to *phoA*	This work
N-term Chp2-AP		pRibo-BsaHind derivative; contains *chp2*(aa1-39) fused to *phoA*	This work
cat Chp2-AP		pRibo-BsaHind derivative; contains *chp2* (26-359) fused to *phoA*	This work
chp2		pMV306 derivative; contains 1kb upstream of *fadD21*; *chp2*	This work
mmpL10		pMV306 derivative; contains 1kb upstream of *pks3/4*; *mmpL10*	This work
$P_{native2kb}$-*mmpL10*		pMV306 derivative; contains 2kb upstream of *pks3/4*	This work
P_{hsp60}-*mmpL10*		pMV306 derivative; P_{hsp60}-*mmpL10*	This work
P_{gs}-*mmpL10*		pMV306 derivative; P_{gs}-*mmpL10*	This work
	2BT	LIC vector with N-terminal 6xHis, TEV protease site, SspI site	S. Gradia
Chp2-cat		2BT derivative; 6XHis-TEV-*chp2*(aa26-359)	This work

Table 2-2. Primers used in this study.

Name (purpose)	**Sequence** (cloning strategy)
pJSC407 Δ*chp2* (knockout)	(by assembly PCR across HygR cassette)
HindIII 5' flank	ggatccacgaagcttCGAGACACGGCACACCTTTG
5' flank-loxP	ACGCTTCAACCGCCAAAGTCggtggcccggtataacttcg
5' flank-loxP revcomp	cgaagttataccgggccaccGACTTTGGCGGTTGAAGCGT
loxP-3' flank	GAACTGTACCGGATTTCGGACGCTGCAACCACAGATCGACG
loxP-3' flank revcomp	CGTCGATCTGTGGTTGCAGCGTCCGAAATCCGGTACAGTTC
3' flank KpnI revcomp	gactagagggtaccGTTACCCGGTCATGCTGAGG
Δ*chp2* PCR (knockout confirmation)	
chp2 KO confirm REV	CGTCATGTTGTCTTGGTGGCTC
fadD21 FOR	GGCCACATAAAGAGTCGCTTCC
hygR confirm REV	CAGGCTCGCGTAGGAATCATC
Δ*mmpL10* PCR (knockout confirm.)	
papA3 FOR	GCTGAAGTCTGTGTTCCAACGAGTC
mmpL10 internal REV	GTCTTCGTGCGCGCTGACATC
chp2 REV*	GTCGATCGTACGCTAGTTAACCGTGGATCCGGTGCGAG
hygR confirm REV	*same as above*
pMV306 *chp2* (complement)	(in two cloning steps via XbaI-HindIII and HindII-ClaI)
5' fadD21 promoter XbaI	GCTCTAGAgggtgctgcggtgctg
fadD21 promoter 3' HindIII	CTTCATAAGCTTCATtggtgctacattaccgtttc
chp2 5' HindIII for pMV306	ccaATGAAGCTTAAGCGAGTGATTGCGGGAG
chp2 3' ClaI for pMV306	GTCGACATCGATTTAGCCGCCGAAGGCGG
2BT *chp2*-cat (*E. coli* expression)	(by InFusion into SspI-cleaved 2BT)
5’ chp2-cat FOR	GTACTTCCAATCCAATGCAAGCGAACCCGCGTACC
3’ chp2-cat REV	TTATCCACTTCCAATtcattaGCCGCCGAAGGCGG

chp2 S141A QC FOR	GTCGGCCTAgCCCAGGGTTCCCTCGTGCTC
chp2 S141A QC REV	GGAACCCTGGGcTAGGCCGACGGCCGCGG
chp2 SA (Ser141Ala mutagenesis)	(by site-directed mutagenesis)
S141A FOR	GTCGGCCTAgCCCAGGGTTCCCTCGTGCTC
S141A REV	GGAACCCTGGGcTAGGCCGACGGCCGCGG

* Originally designed as a cloning primer. Only underlined portion is complementary to *chp2*.

Sequence homology analysis and structure prediction

The amino acid sequence for Chp2 (Rv1184c) was obtained from Tuberculist (http://genolist.pasteur.fr/TubercuList). Transmembrane helices were predicted by the TMHMM Hidden Markov model (http://www.cbs.dtu.dk/services/TMHMM) (113) and signal peptides by SignalP (http://www.cbs.dtu.dk/services/SignalP/) (114). The Chp2 sequence was also submitted to Phyre for protein fold and structure prediction (http://www.sbg.bio.ic.ac.uk/phyre) (115).

Construction and complementation of gene deletion mutants

The Δ*chp1* mutant strain was created by homologous recombination using specialized phage transduction (116). This deletion replaced 765 bp of *chp2* (corresponding to aa 53-307) with a hygromycin resistance cassette. Flanking sequences for *chp2* were amplified from the H37Rv BAC clone Rv7 (gift of Roland Brosch, Institut Pasteur). Recombinant clones were confirmed by PCR (Table 2-2). The Δ*mmpL10* mutant *Mtb* Erdman strain *jcm110* was a gift of Jeffrey Cox (University of California, San Francisco). The deletion was confirmed by PCR and sequencing and replaced 3000 bp of *mmpL10* (corresponding to aa 3-1002) with a hygromycin resistance cassette. Each deletion strain was complemented with a pMV306 integrating plasmid encoding the

target gene with a native promoter (1 kb upstream of the first gene in the corresponding putative operon) (109). The S141A mutation was introduced into Chp2 by site-directed mutagenesis. For Δ*mmpL10* complementation, additional constructs with an extended native promoter (2 kb upstream of *pks3/4*), the *hsp60* promoter (109), or the glutamine synthase (*gs*) (117) promoter were also tested. The *chp1* integrating plasmid for complementation and the Δ*papA3* mutant and complement strains were previously reported (102,104).

Lipid extraction and mass spectrometry analysis

Mtb strains were grown for 3-5 days to late-log phase. Cultures were diluted in 15 mL detergent-free medium to an OD_{600} of 0.3-0.4 and incubated at 37 °C for two days. Cells were harvested and extracted in 1 mL hexane. The upper organic phase ("surface lipid" fraction) was removed and added to an equal volume of 1:1 chloroform/methanol. All extractions were repeated in at least three independent experiments. Extracts were dried by rotary evaporation and resuspended in 500 µl of 20 mM ammonium acetate in methanol. High-resolution mass spectra were obtained on an LTQ Orbitrap XL mass spectrometer with 100,000 FWHM resolving power (118) and equipped with an electrospray ionization source (ThermoFinnigan, San Jose, CA) operating in the negative ion mode. Ions were introduced into the ion source via direct infusion at a rate of 30 µl/min. The ESI source tuning parameters were: sheath gas at 20 arb, spray voltage at 6 kV, capillary temperature at 275 °C, capillary voltage at -37 V, and tube lens voltage at -230 V. Spectra were acquired from *m/z* 500-3000 using a resolution of 30,000 and a maximum injection time of 500 ms. Spectra are averages of 100 scans acquired using Xcalibur (version 2.0.7 SP1, ThermoFinnigan). The spectra were externally calibrated using the manufacturer's standard calibration mixture and further calibrated internally with the lock mass tool using endogenous

lipids phosphatidyl inositol (*m/z* 851.565473) and methoxymycolic acid (*m/z* 1252.286917); such calibration provides within 2 ppm mass accuracy as reported (118).

^{14}C metabolic labeling and lipid analysis by thin-layer chromatography

Mtb strains were grown to late-log phase. For ^{14}C labeling, 5 μCi ^{14}C-propionic acid was added directly to 10 mL culture at OD_{600} ~1. After overnight incubation, cell pellets were extracted in hexane as described above. Extracts were dried and re-suspended in 1/10 or 1/20 the original extraction volume. Volumes representing equivalent numbers of cells were spotted on silica plates (HPTLC Silica Gel 60, EMD Chemicals) and developed in 92:8 petroleum ether/acetone or 90:10 chloroform/methanol to resolve PAT and DAT, respectively, followed by phosphorimaging (107,119). For THL treatment experiments, THL in DMSO was added to 10-mL cultures at 0, 10, 20, and 40 μg/mL for 6 hours followed by the addition of 5 μCi ^{14}C-propionic acid with further incubation, processing and analysis as above.

Subcellular localization and immunoblot of Chp2

Msm containing pRibo Chp2-3xFLAG (112) was cultured overnight. The cells were pelleted and re-suspended at an OD_{600} of 0.3 in 250 mL Sauton medium with or without 2mM theophylline and grown to an OD_{600} of 1. The cells were harvested by centrifugation at 4,000 x g for 30 minutes. The supernatant was sterilized by passing through a 0.22 μm filter, concentrated ~100-fold by centrifugation to ~2.5 mL through a 10-kDa MWCO membrane, and stored at -20 °C as the culture filtrate fraction (CF). The cell pellet was washed twice with 50 ml PBS. The final pellet was weighed and re-suspended in PBS with 1x protease inhibitor (Roche) at 2 mL/g. The

suspension was lysed by sonication for 2 min 10 s on/off at power level 1-2 with a Model 550 Sonic Dismembrator (Fisher Scientific). Cell debris was removed by centrifugation at 10,000 x g for 10 min. After a reserving sample for immunoblot, the cleared total lysate was subjected to ultracentrifugation at 100,000 x g for 1 h. The supernatant was collected as the cytosol-enriched fraction (cyt) and the pellet as cell envelope-enriched fraction (CE), which includes the cytoplasmic membrane, cell wall, and OM. Protein concentration was determined by the BCA assay (Pierce). For the α-MspA blot, samples were prepared as previously described. Briefly, samples were adjusted to 1 µg/µl in 20 µL of 0.6% octyl-thioglucoside in PBS, heated at 98 °C for 30 min, and centrifuged at 10,000 x g for 5 min to remove protein precipitate. For reducing SDS-PAGE, samples were loaded at 10 µg/well (α-FLAG), 20 µg/well (α-KatG), or 5 µg/well (α-MspA). After transfer and blocking, membranes were probed with α-FLAG (1:1,000; clone M2, Sigma); α-KatG (1:500; Clone IT-57 (CDA4), NR-13793 from BEI Resources, NIAID, NIH); and α-MspA (1:666; gift from Michael Niederweis, University of Alabama Birmingham). Secondary goat anti-mouse IgG-IR800 or goat anti-rabbit IgG-IR700 antibodies (LI-COR Biosciences) were prepared according to manufacturer's recommendation and the signal detected with an Odyssey scanner (LI-COR Biosciences).

Enzyme activity of Chp2-βGal and Chp2-AP fusions

Msm strains containing pRibo Chp2-alkaline phosphatase (AP) or Chp2-β-galactosidase (βGal) fusion constructs were grown in 7H9 medium containing 0 or 2 mM theophylline. βGal and AP activity in lysates and whole cells, respectively, were determined essentially as reported using the substrates 2-nitrophenyl β-D-galactopyranoside and 4-nitrophenyl-phosphate (120,121). Activity is reported in Miller units [rxn OD_{420}/(culture OD_{600} x Vs x min)], where Vs = volume of

original culture used in the reaction. Control vectors pMB111 and pMB124 encoding a secretion signal and secreted AP were a kind gift of Miriam Braunstein (University of North Carolina, Chapel Hill). To detect βGal activity on chromogenic medium, cells were streaked on 7H9 agar containing 50 μg/mL X-Gal, 2 mM theophylline, 10% ADC supplement, 0.5% glycerol and 0.05% Tween.

Expression and purification of His-tagged Chp2 catalytic domain in E. coli

A truncated construct of Chp2 lacking the first 25 amino acids (Chp2-cat) was cloned by InFusion (Clontech) into the ligation-independent cloning vector 2BT (gift of Scott Gardia, UC Berkeley) to yield Chp2-cat with an N-terminally fused 6xHis purification tag. Following a 4-hour induction with 0.2 mM isopropyl-β-D-thiogalactopyranoside at 37 °C, cells were lysed in 20 mM Tris, pH 7.4, 200 mM NaCl, 1 mM EDTA, 1 mM DTT. The clarified crude lysate was loaded onto a 5 mL HisTrap FF (GE Biosciences) column and washed with 100 mL 50 mM Tris, pH 7.4, 1 mM DTT, 10% glycerol (buffer A). Bound protein was eluted over 15 column volumes in a gradient of 0-50% 1 M imidazole in buffer A. Fractions containing Chp2-cat were pooled and dialyzed overnight at 4 °C into buffer A. Dialyzed Chp2-cat was loaded onto a 5 mL HiTrap Q column (GE Biosciences), washed with 25 mL buffer A, and eluted in over 15 column volumes in a gradient of 0-100% 1 M NaCl in buffer A. The fractions containing purified Chp2-cat were concentrated and stored at -80 °C. The S141A mutation was introduced by site-directed mutagenesis and the resulting Chp2-cat(SA) protein was purified as above.

Enzymatic reactions

PapA3 was expressed and purified as described (102). For reconstitution of acyltrehalose synthesis, reactions contained 2 μM PapA3, 20 μM palmitoyl-CoA, 1 mM trehalose, and 4 μM Chp2-cat or Chp2-cat(SA) in 25 μL reaction buffer (100 mM sodium phosphate pH 7.2, 1 mM DTT, 10% glycerol). Reactions were incubated at room temperature for 12-16 hours and then quenched with an equal volume of ethanol. Reactions were analyzed by TLC (65:35 chloroform:methanol) and visualized by phosphorimaging. For radiolabeled assays, 1-^{14}C-palmitoyl-CoA was substituted. For biphasic reactions, twice the reaction volume of hexanes was slowly added as a separate phase on top of the aqueous reaction mixture immediately after all components had been combined. Biphasic reactions were not quenched; instead, the two phases were mixed immediately prior to spotting onto the TLC plate.

Reactivity of Chp2-cat with a serine hydrolase-specific activity-based probe

TMR-FP (tetramethylrhodamine linked to fluorophosphonate via a PEG linker; gift of Benjamin Cravatt, Scripps Institute) (122) was incubated at 2 μM for 1 hour at room temperature with 2 μM Chp2-cat or Chp2-cat(SA) in 50 mM Tris pH 7.4, 1 mM DTT, and 10% glycerol either with or without pre-heating at 98 °C for 5 min. For inhibition experiments, THL was added to each sample at 20 μM and incubated for 30 min at room temperature prior to labeling with TMR-FP as above. Samples were separated by SDS-PAGE and scanned for TMR fluorescence with a Typhoon FLA 7000 scanner (GE Healthcare).

Results

The *chp2* gene in *Mtb* Erdman was replaced with a hygromycin resistance to yield the knockout mutant Δ*chp2*. Both this and the Δ*mmpL10* were verified by PCR (Figure 2-2). Surface lipids extracts were analyzed by direct infusion ESI-MS and compared with extracts from wild-type Erdman, a Δ*papA3* mutant, and a strain deficient in *mmpL10*. In the wild type an envelope of ions was detected at *m/z* 1900-2200 that is characteristic of PAT (Figure 2-3A) (102,123). Observed peaks were within 5 ppm of the predicted acetate pseudomolecular ions (data not shown). In addition, a set of peaks centered at approximately *m/z* 1790 is consistent with tetraacyltrehaloses containing one fewer mycolipenoate modification than PAT. Although MSn analysis could not be completed on any of these ions due to low yield, the exact masses match the theoretical masses of predicted tetraacyltrehalose structures (Figure 2-3B). To avoid confusion with the literature acronym TAT for triacyltrehalose, we refer to tetraacyltrehalose as TetraAT. The ions for PAT and TetraAT were not observed in the Δ*papA3*, Δ*chp2*, or Δ*mmpL10* mutant strains.

Lane #	Template DNA	Expected length (bp)
1	wild type	1245
2	wild type	0
3	Δ*chp2*	0
4	Δ*chp2*	1611
5	wild type	3096
6	wild type	0
7	Δ*mmpL10*	0
8	Δ*mmpL10*	944

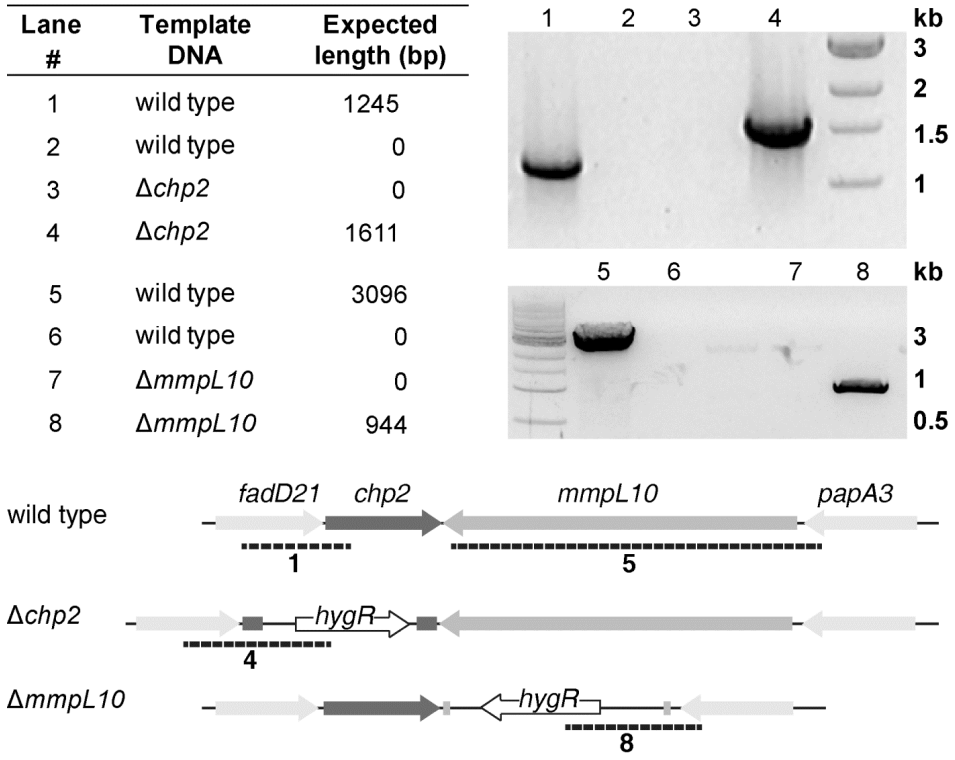
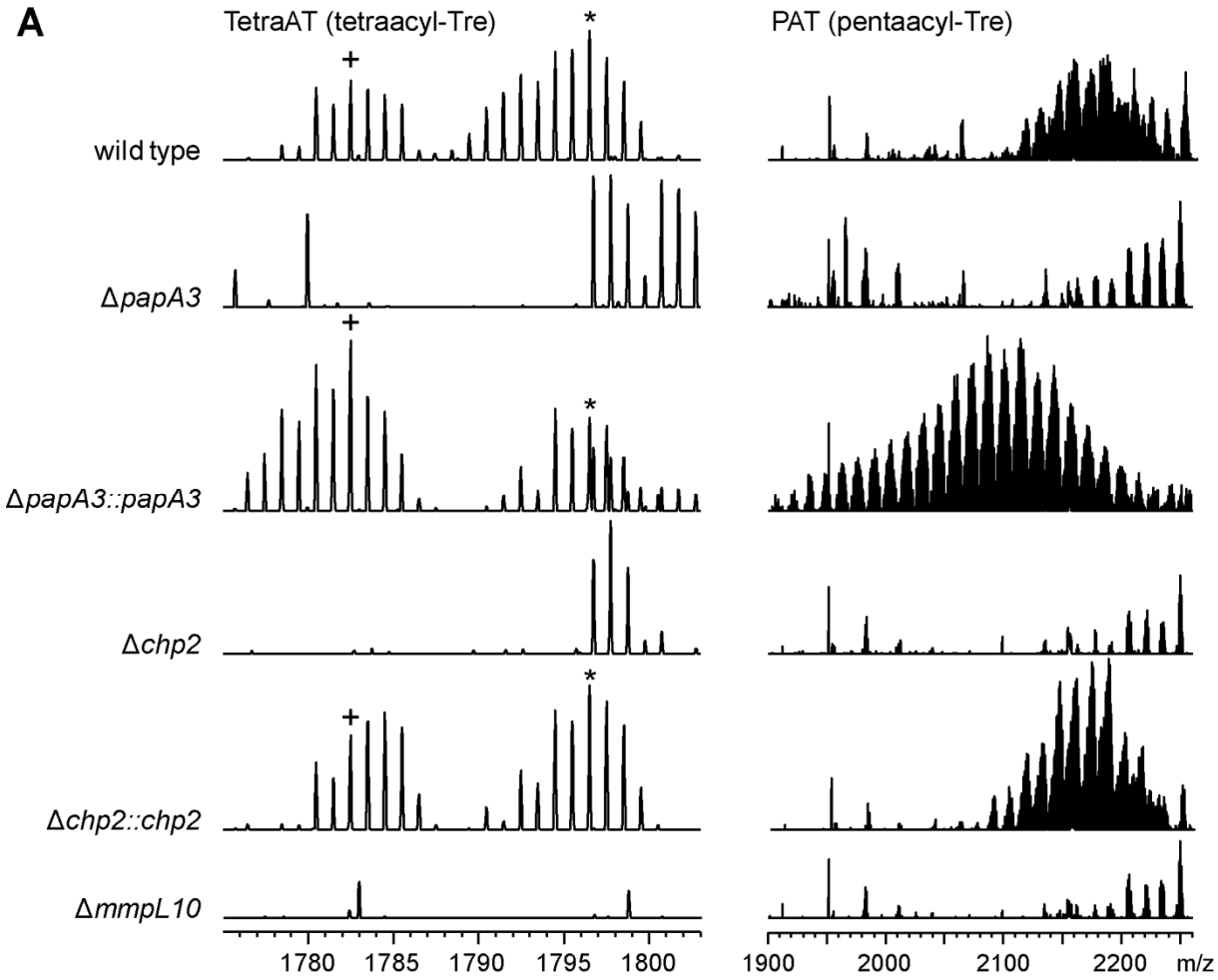
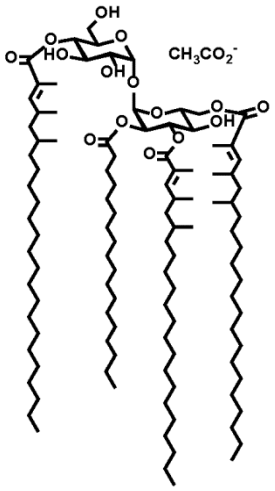


Figure 2-2. Confirmation of Δchp2 and ΔmmpL10 by PCR. For chp2, the 5′ primer annealed ~900 bp upstream of chp2. For lanes 1 and 3, the 3′ primer annealed to a region of chp2 that is absent in the knockout strain. For mmpL10, the 5′ primer annealed ~100 bp upstream of mmpL10. For lanes 5 and 7, the 3′ primer annealed to a region of mmpL10 that is absent in the knockout strain. For lanes 2, 4, 6 and 8, the 3′ primer annealed to the hygR marker that replaces chp2 and mmpL10 in the knockout strains. See supplemental Table 2 for oligonucleotide sequences.



B



TetraAT ions	+	Δppm	*	Δppm
theoretical	1782.487		1796.503	
wild type	1782.491	2.244	1796.499	-2.227
papA3::papA3	1782.485	-1.122	1796.502	-0.557
chp2::chp2	1782.484	-1.683	1796.498	-2.783

Figure 2-3. Chp2 and MmpL10 are required biosynthesis of PAT and the novel acyltrehalose TetraAT. *Mtb* wild type, Δ*papA3*, Δ*chp2*, and Δ*mmpL10* strains were analyzed by ESI-MS. (A) TetraAT and PAT were absent from all three mutants, but were restored in the Δ*papA3* and Δ*chp2* complement strains. PAT appears as a characteristic envelope of peaks centered at approximately *m/z* 2100-2200. A representative segment of the TetraAT spectrum is shown in order highlight two

of the peaks used for assignment. (B) TetraAT was assigned by exact mass and observed ions were within 3 ppm of the predicted *m/z*. As an example, a possible structure for the *m/z* 1782.487 [M+CH_3CO_2-] ion is shown.

Complementary results were obtained by TLC and phosphorimage analysis of lipids from *Mtb* labeled metabolically with ^{14}C-propionate, which is incorporated into methyl-branched lipids (Figure 2-4). While the migration of TAT and TetraAT has not been assigned under the TLC conditions used, PAT and DAT were resolved. All the mutants lack PAT, but while DAT is lost in the Δ*papA3* mutant, both Δ*chp2* and Δ*mmpL10* still produce this biosynthetic intermediate. The wild-type lipid profile is restored in the Δ*papA3::papA3* and Δ*chp2::chp2* complement strains. In addition, *chp1* and *chp2* did not cross-complement the respective Δ*chp2* and Δ*chp1* knockout strains in merodiploid strains containing a second copy of the corresponding gene. Multiple attempts and strategies to complement the Δ*mmpL10* mutant were not successful (data not shown).

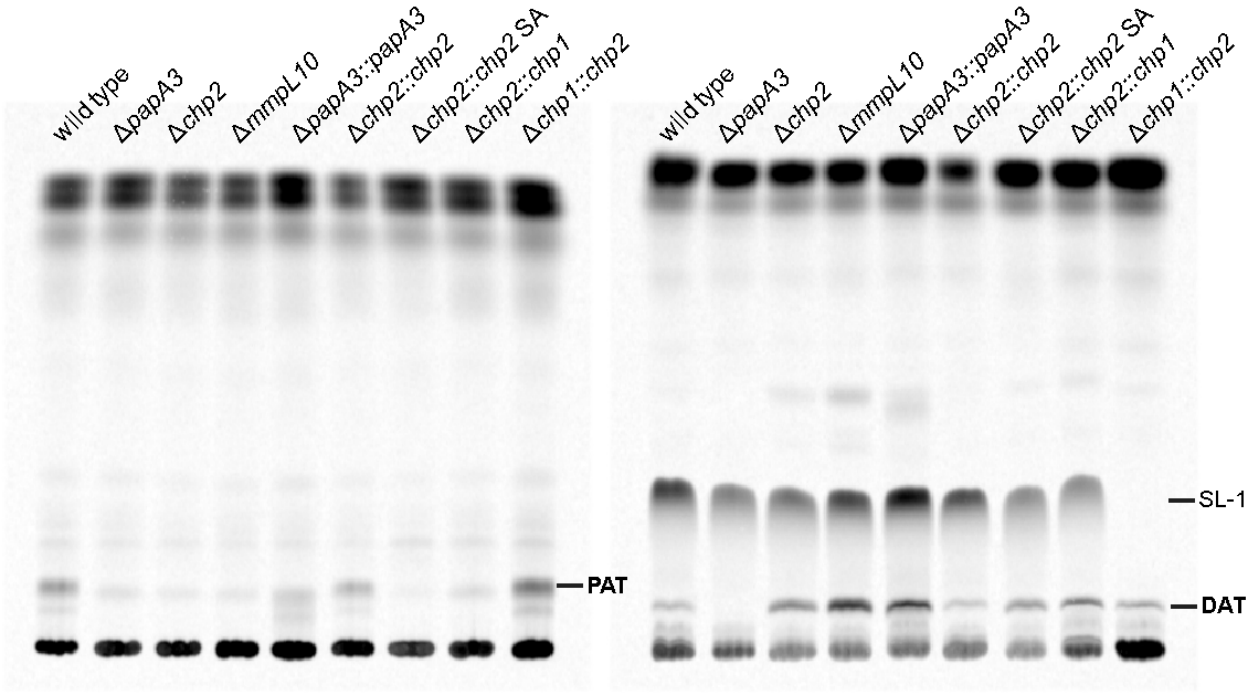


Figure 2-4. Chp2 and MmpL10 function downstream of DAT and Chp2 is specific to the PAT biosynthesis. *Mtb* wild type, ΔpapA3, Δchp2, and ΔmmpL10 strains were analyzed by metabolic labeling with ^{14}C-propionate followed by TLC of lipid extracts. In addition to loss of PAT in all mutants (A), DAT was absent from ΔpapA3 (B). PAT and DAT were restored upon complementation, except in Δchp2::chp2 SA. The chp1 and chp2 homologues did not cross-complement the Δchp2 and Δchp1 mutants. Mobile phases were (A) 92:8 petroleum ether/acetone and (B) 90:10 chloroform/methanol.

Bioinformatic analysis of the Chp2 amino acid sequence revealed several predicted features: (i) a signal sequence (aa 1-26); (ii) a transmembrane helix (aa 5-27); and (iii) a conserved PE-PPE C-terminal domain (aa 79-316) that is a predicted serine hydrolase (Figure 2-5A). In addition, a structural homology model based on the *P. purpurogenum* acetylxylan esterase structure (PDB ID: 1G66) identified the putative catalytic triad (S141-D226-H248; catalytic residue underlined), similar to a previously reported prediction based on the structure of a *Msm* lipase (124). To determine its subcellular location, Chp2 was expressed in *Msm* as a theophylline-inducible full-length C-terminal fusion to a 3xFLAG epitope tag. When cells were grown in theophylline-supplemented medium, FLAG-tagged Chp2 was enriched in the cell envelope

fraction over the cytosolic and culture filtrate fractions (Figure 2-5B). To determine whether the C-terminal catalytic domain is located in the cytosol or the periplasm, Chp2 was also expressed as a theophylline-inducible C-terminal fusion to either alkaline phosphatase (AP) or b-galactosidase (βGal). No AP activity was detected when Chp2 was expressed as the full-length protein, N-terminus (N-term), or catalytic domain (cat) (Figure 2-5C). On the other hand, βGal activity was confirmed in both cell lysates and whole cells (Figure 2-5D). These data suggest that Chp2 has its C-terminus located in the cytoplasm and is associated with the cytoplasmic membrane and possibly via an N-terminal transmembrane helix, similar to Chp1.

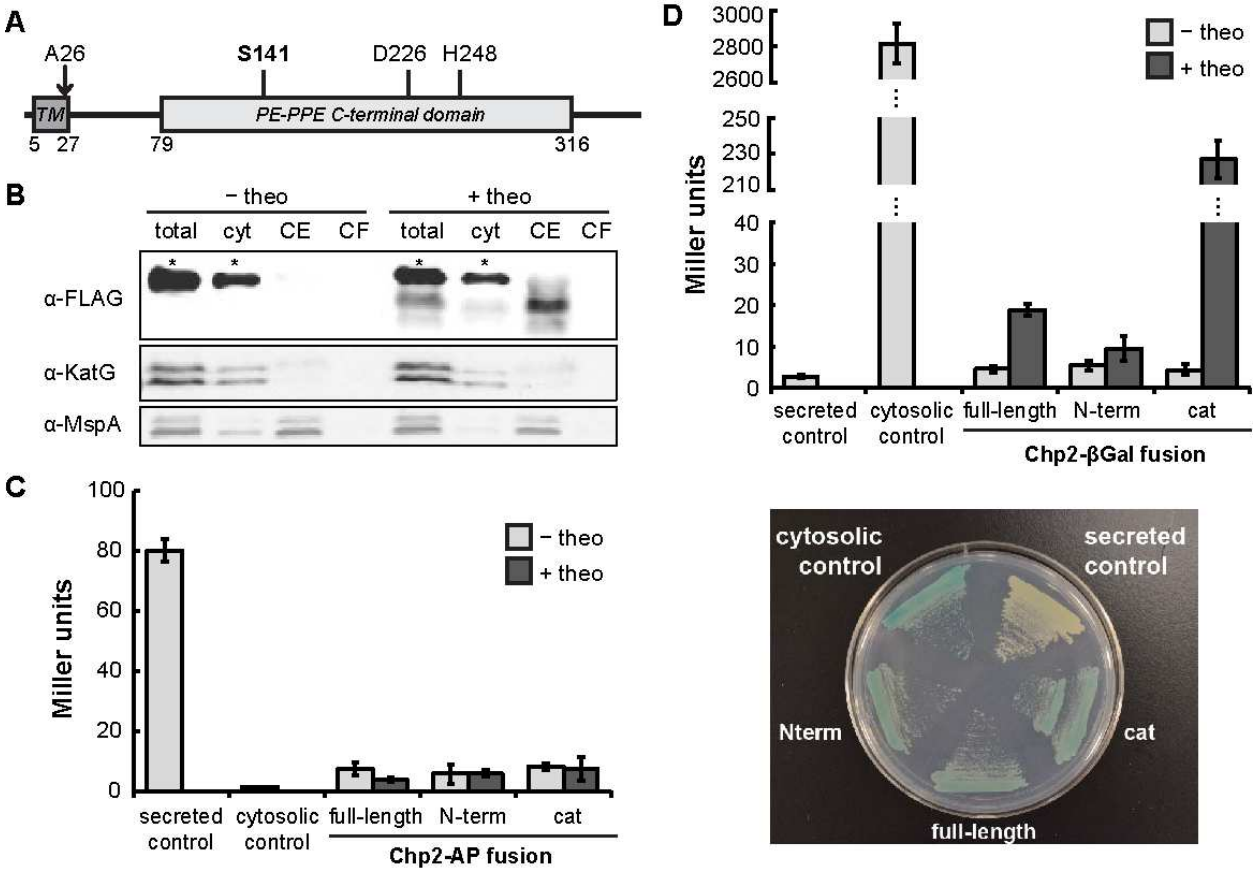


Figure 2-5. Chp2 localizes to the cell envelope with the catalytic domain in the cytosol. (A) Primary structure of Chp2 with the predicted positions of the transmembrane helix (TM), signal peptide cleavage site (A27), conserved domain, and catalytic residues. (B) Immunoblot of

subcellular fractions of *Msm* expressing theophylline-inducible Chp2-3xFLAG showed Chp2 enrichment in the cell envelope (CE) fraction. KatG and MspA are markers for the cytosol and outer membrane, respectively. Asterisk indicates a non-specific α-FLAG-reactive band. In *Msm* expressing theophylline-inducible Chp2 full-length, N-terminal or catalytic domain fused to (C) AP or (D) βGal , enzymatic activity was observed only from βGal fusions, which were further confirmed by growing on chromogenic medium containing X-Gal. Strains grown without theophylline or expressing AP or βGal +/- an N-terminal secretion signal served as negative and positive controls. Turnover of colorimetric substrates is expressed in Miller units.

To explore the proposed enzymatic activity of Chp2, the conserved C-terminal PE-PPE domain (aa 26-359) was expressed and purified as an N-terminally 6xHis-tagged construct (Chp2-cat). Diverse strategies were pursued in an effort to reconstitute and detect trehalose glycolipid acyltransferase activity for Chp2-cat. The reaction of PapA3, trehalose and palmitoyl-CoA yielded the expected mixture of trehalose monopalmitate and trehalose dipalmitate, which is the predicted substrate for Chp2. Addition of Chp2-cat to the PapA3 reaction mixture did not, however, yield any Chp2 activity-dependent changes, as detected by either by TLC or MS analysis (data not shown). To control for possible inhibition of Chp2 by excess trehalose or palmitoyl-CoA from the PapA3 reaction, trehalose mono- and dipalmitate were isolated by preparative TLC and incubated with Chp2-cat (with or without additional palmitoyl-CoA), but again, no additional products were observed. We further reasoned that the predicted products of Chp2 may aggregate or be inaccessible for further reaction in aqueous reaction buffer, since these lipids are neutral and highly hydrophobic. Although mixed organic:aqueous conditions have proved useful for promoting enzymatic wax ester synthesis (Touchette and Seeliger, unpublished results), applying a biphasic system to all of the reaction conditions above did not yield Chp2-dependent products.

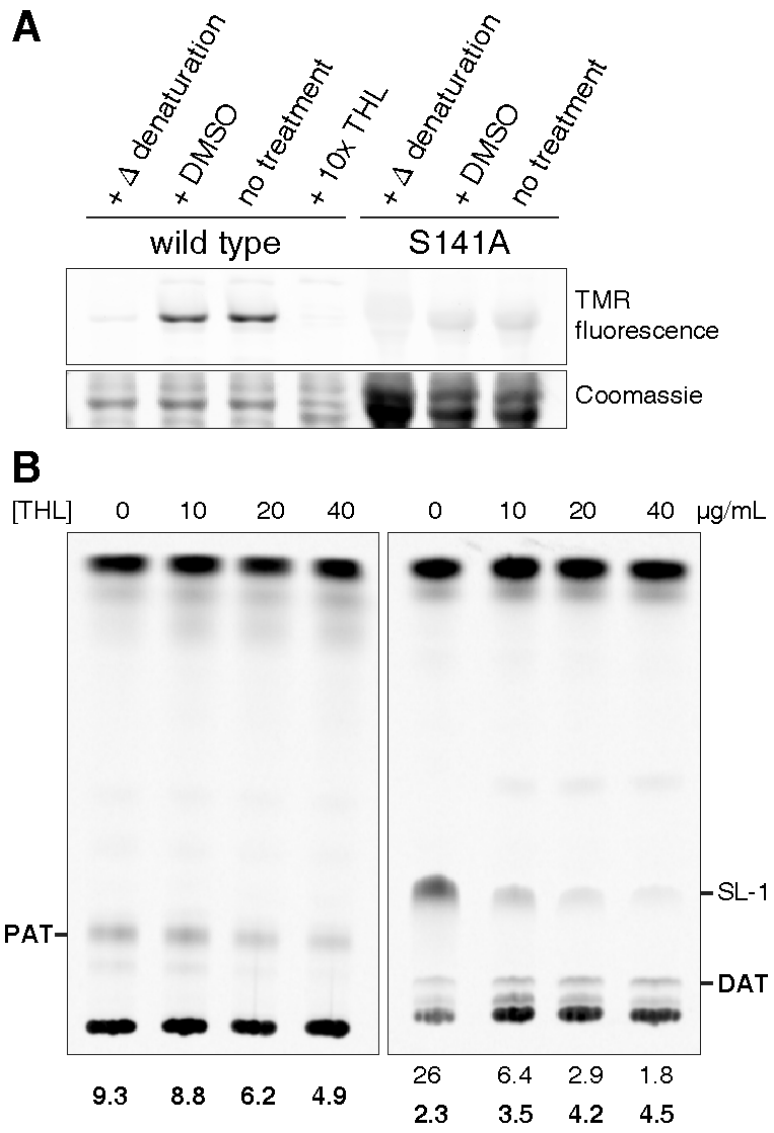


Figure 2-6. Chp2 is a serine hydrolase-type enzyme that is inhibited by THL. (A) Labeling of purified Chp2-cat domain by TAMRA-FP was detected by fluorescence and inhibited by heat denaturation, treatment with THL, or mutation of the catalytic residue (S141A). (B) Lipid extracts from *Mtb* treated with different concentrations of THL for 6 h followed by ^{14}C-propionate labeling revealed the dose-dependent, but incomplete, inhibition of PAT and accumulation of DAT (mobile phases as in Figure 2-4B). SL-1 inhibition by THL has been previously reported. Bold numbers denote the percent of total integrated intensity for PAT and DAT in each lane (analysis for SL-1 shown for comparison (104)).

Chp2 enzymatic activity was instead confirmed using the activity-based probe TMR-FP, which is specific for serine hydrolases (Figure 2-6A). No labeling was observed following heat

denaturation or upon mutation of the putative catalytic residue Ser141 to Ala (SA). Similarly, a mutant *chp2* containing the SA mutation did not complement the Δ*chp2* knockout strain (Figure 2-4). TMR-FP labeling was inhibited by pre-incubating Chp2-cat with the lipophilic lipase inhibitor tetrahydrolipistatin (THL). Also, treatment of a wild-type *Mtb* mid-log culture with THL prior to metabolic incorporation of ^{14}C-propionate revealed a moderate dose-dependent inhibition of PAT biosynthesis and concomitant accumulation of DAT (Figure 2-6B). In further attempts to reconstitute acyltransferase activity, additional constructs of Chp2 (aa 45-359, aa 52-359, and aa 79-359) were expressed as fusions to an N-terminal 6xHis-maltose binding protein linked by a TEV protease site. In all cases, both the fusion and purified cleaved protein were active by TMR-FP labeling, but none yielded products in biochemical reactions with PapA3 (data not shown).

Discussion

In this study we confirmed that Chp2 has a C-terminal serine hydrolase domain that is inhibited by THL. Within the PAT biosynthesis pathway, THL inhibits Chp2 but not PapA3, leading to decreased levels of PAT and the accumulation of the DAT precursor in *Mtb*. Although we were unable to detect Chp activity on trehalose glycolipid analogues, the biochemical data, along with the lipid profiling of the Δ*chp2* mutant, place Chp2 downstream of PapA3 and likely the sole acyltransferase that catalyzes the three esterifications necessary to convert DAT to PAT. Our data also confirm PapA3 as a bifunctional acyltransferase, an assignment that remained in question from an earlier study (102). We further identified a previously uncharacterized acyltrehalose, TetraAT, as a PAT precursor that is absent in the Δ*papA3*, Δ*chp2* and Δ*mmpL10* mutants.

The function of Chp2 is therefore directly analogous to the role of Chp1 in SL-1 sulfoglycolipid biosynthesis. Although Chp1 activity has been successfully reconstituted *in vitro*, it is possible that Chp2 is not active under similar conditions because it is more dependent on interactions with the cytoplasmic membrane or with other proteins such as MmpL10 for folding and/or catalysis. Chp1 is more susceptible to THL inhibition than Chp2 upon treatment of *Mtb* (Figure 2-4B) (104), but in the absence of structural data on these two enzymes, the origins of this specificity are not obvious.

In the SL-1 pathway, the membrane protein MmpL8 is a sulfoglycolipid transporter, but also promotes Chp1 activity (125,126). MmpL10 appears to act similarly within the PAT pathway, as the Δ*mmpL10* mutant is defective in PAT biosynthesis, but still produces DAT. We would further predict that DAT would be absent in the OM of *ΔmmpL10*, but this could not be confirmed. While DAT was detectable by ^{14}C-radioisotope labeling, DAT ions were not observed by mass spectrometry and therefore could not be compared between the surface and cell pellet lipid fractions as done previously (104). Since MS analysis was performed without chromatographic separation, DAT was likely suppressed in our spectra by abundant ions in the same *m/z* range (e.g., theoretical *m/z* 1057.777 for DAT-acetate adduct with palmitate and C_{24} mycolipenoate modifications). Since *mmpL10* is the last gene in the predicted *pks3/4-papA3-mmpL10* operon, its deletion is not expected to affect other loci. Since the Δ*mmpL10* mutant was not complemented despite numerous attempts, it is nonetheless possible that *mmpL10* disruption has a polar effect on the *rv1184c* locus, which has opposite orientation but is separated from *mmpL10* by only three nucleotides.

Chp2 has a high-probability signal sequence (based on empirical algorithms for Gram-negative bacteria (114)) that indicates extracytoplasmic export, although this sequence also has the

properties of a transmembrane helix, as is typical for signal peptides. Chp2 was found in the cell envelope fraction, consistent with either export into the cell wall or membrane localization. The results of the complementary βGal/AP-fusion activity assay indicate that the catalytic domain is in the cytosol and are therefore consistent with membrane localization. Furthermore, the activity of βGal was lowest when fused directly to the Chp2 N-terminal helix domain, consistent with βGal being constrained at the membrane-cytosol interface and therefore hindered in the formation of the active tetramer. This is in contrast to the presumably greater conformational freedom when βGal is fused to either full-length Chp2 or Chp2-cat, which have correspondingly higher observed activities. This observed trend further supports our model for Chp1 topology.

Based on these data, we conclude that PAT biosynthesis in *Mtb* proceeds similarly to sulfoglycolipid biosynthesis, in line with the organizational similarity between the two biosynthetic clusters. PapA3 modifies trehalose to yield DAT, which is then further esterified by Chp2 to form TAT, TetraAT, and PAT (Figure 2-1). Given their sequence and predicted structural similarity, Chp2 likely functions by a similar mechanism to that of Chp1 and Ag85, that is, Chp2 catalyzes transesterification using DAT as the acyl donor and generating T2P as a side product.

As with Chp1 and MmpL8, the activity of Chp2 may be facilitated by the transporter MmpL10 (Figure 2-7). Based on our data, Chp2 at the cytoplasmic leaflet synthesizes PAT and in the process produces T2P that can immediately reenter PAT biosynthesis via modification by PapA3. Because DAT, TAT, TetraAT and PAT are all found at the cell surface, this model implies that MmpL10 transports all of these lipids across the cytoplasmic membrane. This substrate promiscuity would distinguish MmpL10 from previously characterized MmpLs such as MmpL7 and MmpL3, which transport exclusively phthiocerol mycocerosate (PDIM) and trehalose monomycolate (41,127,128). Interestingly, both *mmpL10* and *chp2* have been characterized as

essential *Mtb* genes *in vivo* (129), but an *Mtb* strain disrupted in *pks3/4* (and therefore lacking all PAT-related acyltrehaloses) was not attenuated for growth in a mouse infection (130). These data suggest that the accumulation of DAT compromises *Mtb* survival *in vivo*.

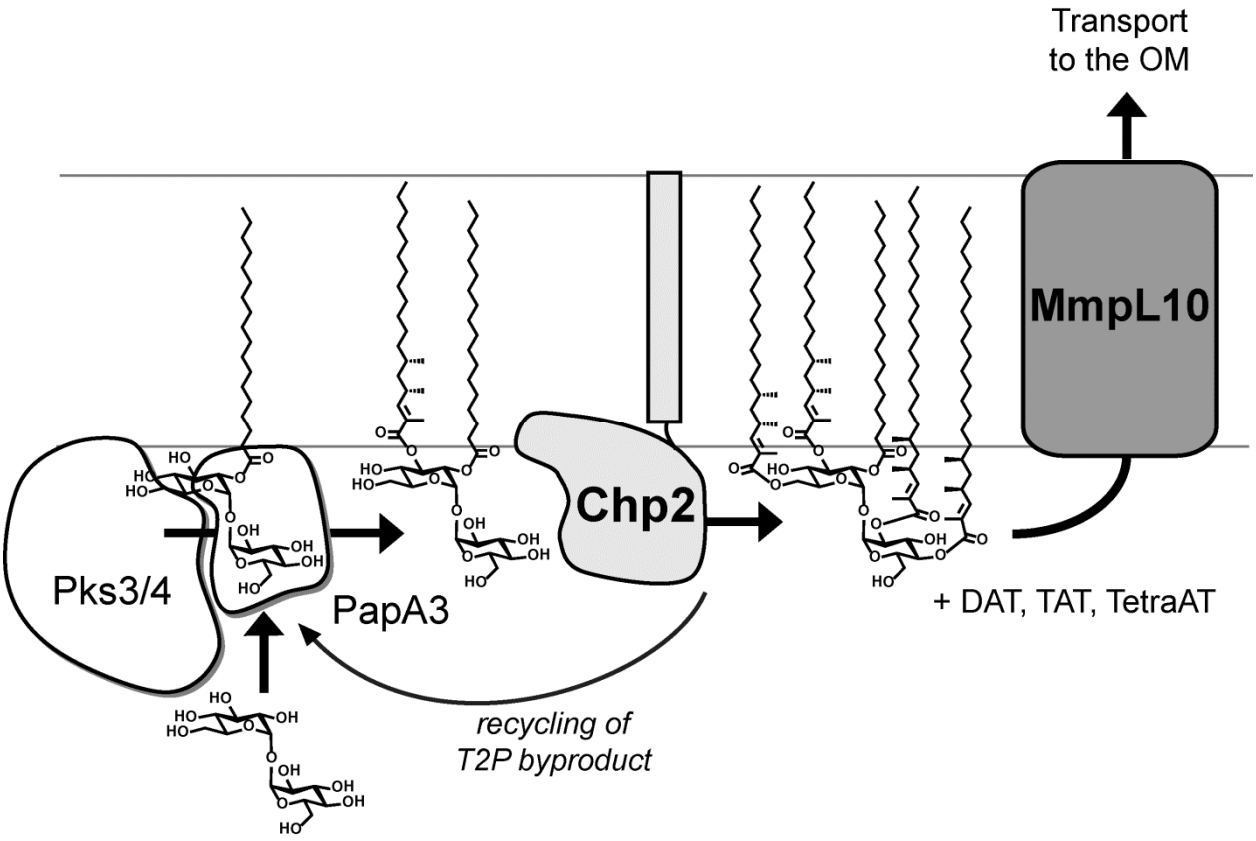


Figure 2-7. Proposed model for acyltrehalose biosynthesis and transmembrane transport. The data are most consistent with a model in which Chp2 completes PAT biosynthesis in the cytosolic leaflet and MmpL10 transports acyltrehaloses across the membrane. The AcTre side products generated by Chp2 can be recycled by PapA3. (Note that for simplicity the mycolipenic groups are truncated in this figure.)

In summary, we have assigned Chp2 as a DAT acyltransferase and thereby completed the biosynthetic pathway for PAT. While the overall synthetic scheme is now defined for PAT, as well as for other virulence-associated OM lipids such as SL-1 and PDIM (131), many questions still remain regarding how MmpLs influence the activity of biosynthetic enzymes and how lipids are

transported and assembled in the OM. Elucidation of these mechanisms will have implications for our understanding of how the structure and composition of the OM are regulated, what structural and functional roles these lipids play in *Mtb* survival and pathogenesis, and the consequences of targeting cell wall-associated processes for tuberculosis drug therapy.

Chapter 3 Diacyltransferase Activity and Chain Length Specificity of *Mycobacterium tuberculosis* PapA5 in the Synthesis of Alkyl Beta-Diol Lipids

This chapter has been published. *Biochemistry* and all authors have granted permission to use this manuscript in this dissertation:

Touchette, M. H., Bommineni, G. R., Delle Bovi, R., Gadbery, J. E., Nicora, C. D., Shukla, A., K., Kyle, J. E., Metz, T. O., Martin, D. W., Sampson, N. S., Miller, W. T., Tonge, P. J., Seeliger, J. C. Diacyltransferase Activity and Chain Length Specificity of *Mycobacterium tuberculosis* PapA5 in the Synthesis of Alkyl Beta-Diol Lipids, *Biochemistry,* **just accepted**.

Introduction

Bacteria that belong to the suborder Corynebacterineae are classified as Gram positive, but their cell wall is distinguished by long-chain fatty acids, known as mycolic acids, that are esterified to an underlying arabinogalactan carbohydrate layer (132). In their covalent attachment to the cell wall, mycolic acids are analogous to the peptidoglycan-anchored wall teichoic acids of other Gram-positive bacteria (133), but constitute a unique, well-ordered layer that is believed to form the inner leaflet of a second lipid bilayer (sometimes referred to as the mycobacterial outer membrane or "mycomembrane") that is distinct from the cytosolic membrane. The double-membrane structure of Corynebacterineae is therefore analogous to that of Gram-negative bacteria, which also possess an asymmetric outer membrane. In addition, just as lipopolysaccharide at the bacterial cell surface restricts cell permeability and plays key roles in host-pathogen interactions of Gram-negative bacteria (134), non-covalently attached lipids in the outer membrane have been linked to the inherent drug resistance and virulence of pathogenic Corynebacterineae (135).

Unsurprisingly, the more prominent examples of such virulence-associated lipids come from the human pathogen *Mycobacterium tuberculosis* (*Mtb*), whose outer membrane contains

glycolipids and ß-diol wax diesters that are unique to pathogenic strains. Of these, the dimycocerosates (DIMs) comprise two families of ß-diol diesters that participate in *Mtb*-host interactions: phthiocerol dimycocerosates (PDIM) and the structurally related phenolic glycolipids (PGL) (136). Clinical *Mtb* strains that lack PDIM and mutants defective in PDIM biosynthesis or transport are attenuated for virulence in animal-based models of infection (40-43). Diverse roles for PDIM in *Mtb*-host biology have been described (137-140) that underscore the importance of these unusual lipids in promoting the ability of *Mtb* to cause disease.

Given the importance of DIMs in *Mtb*-host biology, DIM structure and biosynthesis have been extensively studied and recently summarized in a comprehensive review (131). Briefly, both PDIMs (Figure 3-1) and PGLs comprise aliphatic ß-diols, known as phthiocerols, esterified by long branched-chain fatty acids, known as mycocerosic acids. The PpsABCDE cluster of modular polyketide synthases (PKS) synthesizes phthiocerol ß-diols. Pks1/15 makes the *p*-hydroxyphenylalkanoate precursor that FadD29 then passes to the Pps complex to form the phenolphthiocerol ß-diol of PGL (141). The mycocerosic acids are produced by another specialized PKS, mycocerosic acid synthase (Mas). Additional structural diversity within the DIM families arises from methoxy- and keto- versions of phthiocerol as well as variations in the length and number of branches in the mycocerosic chains.

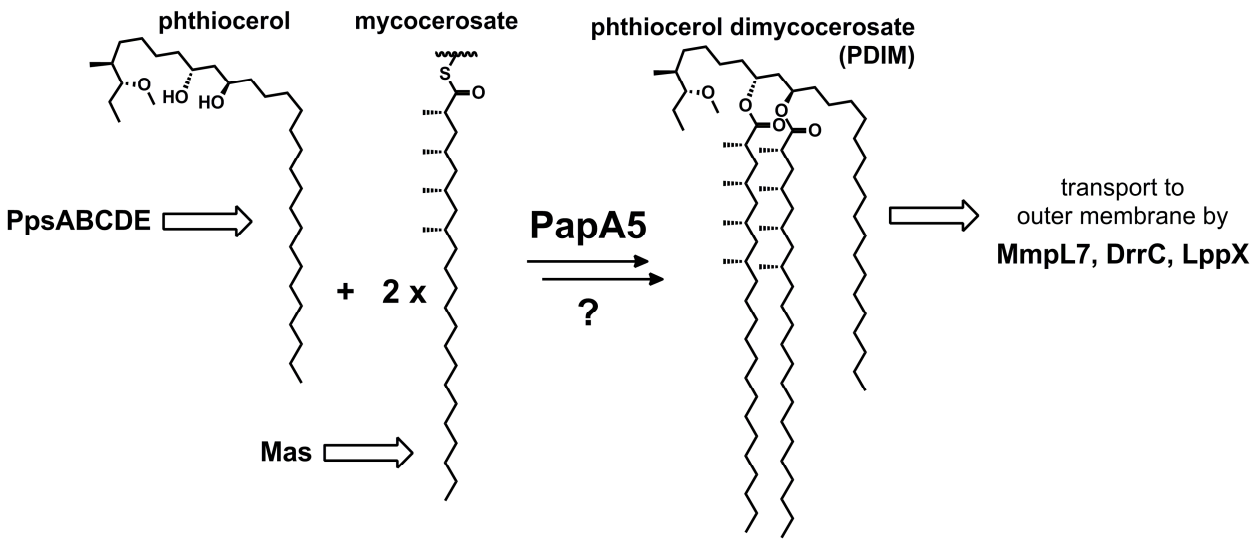


Figure 3-1 Biosynthesis of phthiocerol dimycocerosate (PDIM). The phthiocerol ß-alkyl diol is synthesized by the modular polyketide synthase (PKS) PpsABCDE. The iterative PKS Mas generates mycocerosate, which is esterified directly onto the phthiocerol by PapA5 at least once. PapA5 may also be responsible for a second esterification that generates the final PDIM product, which is then transported to the outer membrane by MmpL7 and DrrC and LppX.

The acyltransferase PapA5 has been implicated in the esterification of ß-diols with mycocerosic acids to generate both PDIM (Figure 3-1) and PGL (142,143). PapA5 has a preference for longer-chain (C_{12}-C_{18}) fatty acid-CoA substrates as acyl donors (143,144) and catalyzes the *in vitro* esterification of a variety of aliphatic alcohols, including vicinal and ß-diols. When PapA5 was incubated with phenolphthiocerol, a putative native substrate isolated from *M. leprae*, only a singly esterified product was detected (142), leaving as an ongoing question whether PapA5 catalyzes successive reactions or, alternatively, a second and as yet uncharacterized acyltransferase follows PapA5 in DIM biosynthesis. Following synthesis in the cytosolic membrane, DIM localization in the outer membrane depends on the integral membrane transporters MmpL7 and DrrC and the cell wall lipoprotein LppX (Figure 3-1) (41,145,146).

There is precedent for dual esterification activity among related acyltransferases. PapA5 belongs to an enzyme family that includes PapA1 and PapA2, which are involved in

sulfoglycolipid biosynthesis, and PapA3, which initiates acyltrehalose biosynthesis (102,103,147). Like DIMs, the products of these acyltransferases are surface-localized outer membrane lipids. PapA3 is unique in that it catalyzes two reactions and esterifies trehalose successively at the 1- and 2-positions with a straight chain fatty acid and a branched chain mycolipenic acid, respectively (102). Based on this example, and the absence of any other predicted acyltransferase near the DIM biosynthetic locus, we hypothesized that PapA5 is indeed the sole acyltransferase in the DIM pathway.

Diverse conditions affect DIM production and PapA5 activity in *Mtb*, including carbon source, reductive stress, and transcriptional regulation in response upon macrophage infection (106,148-151). PDIM biosynthesis may also be modulated by the activity of mycobacterial protein kinases. PDIM levels are attenuated in a PknH knockout strain, and *in vitro*, PknD modifies the PDIM transporter MmpL7 and PknB phosphorylates PapA5 (152-154). Phosphorylated peptides corresponding to Mas and Pks1/15 have also been detected in *Mtb* cell lysates (155,156). We hypothesized that phosphorylation modulates PDIM biosynthesis in part via a direct effect on the catalytic activity of PapA5, as has been shown for other lipid biosynthetic enzymes in *Mtb* (157-159).

Here we present data that establish the dual esterification activity of PapA5 using synthetic ß-diol substrate analogues and show that PapA5 has a preference for longer-chain ß-diols. Using a series of mutations, we define the substrate-binding sites suggested by the previously published crystal structure and propose a model for the interaction of PapA5 with the phthiocerol and acyl donor substrates. We also show that PapA5 is a substrate for the *Mtb* kinases PknB and PknE. Phosphorylation by PknB and PknE overlaps at several Thr within an unresolved segment distal to the active site; modification by PknE was detected at one additional site.

Experimental Procedures

Materials and reagents

E. coli XL-1 Blue (Stratagene) or Stellar (Clontech Laboratories) strains were used for cloning. *E. coli* BL21(DE3) cells were used for protein expression. Cells were grown in LB supplemented with 50 μg/mL kanamycin (for pET28b vectors) or 50 μg/mL streptomycin (for pCDFDuet-1 vectors) unless otherwise indicated. See Table 3-1 for details on vector constructs and oligonucleotide primer sequences. The alkyl alcohols (3*R*,5*R*)-heptane-3,5-diol and (3*S*,5*S*)-heptane-3,5-diol were purchased from Strem Chemicals, Inc. (Germany).

Synthesis of alkyl anti-beta-diol compounds

Overall synthetic scheme for 1,3-*anti*-diols: Compound **4** (Figure 3-2C) was synthesized as reported (160,161). Compound **4** was coupled with hydroxylamine hydrochloridein the presence of EDC/DIPEA to form **5** which, upon reaction with the corresponding Grignard or lithium-derived agent, yielded **6** and **7** in good yield. Stereoselective reduction of compound **6** & **7** with tetramethylammonium triacetoxyborohydride at -20 °C gave compounds **8** & **9**.

(*R*)-3-hydroxy-*N*-methoxy-*N*-methylheptanamide (**5**) (161): *N*,*O*-Dimethylhydroxylamine hydrochloride (521 mg, 5.3 mmol, Acros Organics), EDC (1.73 mg, 9.0 mmol), *N*,*N*'-diisopropylethylamine (0.93 mL, 5.3 mmol, DIPEA, Alfa Aesar) and a catalytic amount of 4-dimethylaminopyridine (~50 mg, 4-DMAP, Acros Organics) were added to a solution of compound **4** (600 mg, 4.1 mmol) in dry DMF. The mixture was stirred for 15 hours at 24 °C and extracted with 10% HCl followed by a wash with saturated sodium bicarbonate. The extract was dried with magnesium sulfate, filtered and evaporated under reduced pressure. The residue was

purified by silica gel chromatography with petroleum ether/AcOEt (7:3, v/v) to obtain compound **5** (498 mg, 83% yield) as a white powder.

$[\alpha]_D^{23}$= - 58 (c 1, CH_2Cl_2). ^{1}H NMR (500 MHz, $CDCl_3$): 3.86 (br s, 1 H), 3.58 (br s, 1 H), 3.55 (s, 3 H), 3.04 (s, 3 H), 2.51 – 2.31 (m, 2 H), 1.38 – 1.16 (m, 6 H), 0.75 (t, *J* = 7.2 Hz, 3 H). ^{13}C NMR (125 MHz, $CDCl_3$): δ 173.4, 67.5, 60.8, 38.0, 36.0, 31.2, 27.3, 22.2, 13.5. MS (ESI^+) m/z (%): 190.1 (100) $[M + H]^+$.

(*R*)-7-hydroxyundecan-5-one (**6**) (162): Compound **5** (203 mg, 1.1 mmol) was dried under argon and then reacted with 2 M *n*-butyl lithium in cyclohexane (1.5 mL, 18.2 mmol, Sigma-Aldrich) in dry THF under nitrogen at -78 °C for 2 hours. The reaction mixture was extracted with saturated sodium bicarbonate, dried with magnesium sulfate, filtered and evaporated under reduced pressure. The resulting residue was purified by silica gel chromatography with petroleum ether/AcOEt (4:1, v/v) to obtain compound **6** (137 mg, 68% yield) as a white powder.

$[\alpha]_D^{23}$= - 42 (c 1, CH_2Cl_2). ^{1}H NMR (500 MHz, $CDCl_3$): 3.95 (br s, 1 H), 3.31 (br s, 1 H), 2.54 – 2.35 (m, 4 H), 1.51-1.46 (m, 2 H), 1.36 – 1.22 (m, 8 H), 0.84 (m, 6 H). ^{13}C NMR (125 MHz, $CDCl_3$): δ 212.2, 67.5, 48.9, 43.2, 36.1, 27.5, 25.5, 22.4, 22.1, 13.8, 13.6. MS (ESI^+) m/z (%): 169.2 (60), 187.2 (40) $[M + H]^+$.

(5*R*,7*R*)-undecane-5,7-diol (**8**; UDD) (163,164): Tetramethylammonium triacetoxyborohydride (775 mg, 2.9 mmol, Sigma-Aldrich) was added to a stirred solution of acetonitrile (3.7 mL, 70.4 mmol, Fisher) and acetic acid (3.7 mL, 130 mmol, Fisher) at 24 °C. After 30 minutes, the reaction mixture was cooled to -78 °C and compound **6** (137 mg, 0.74 mmol) was added The resulting mixture stirred for 40 minutes followed by 36 hours at -20 °C. The reaction mixture was extracted with saturated sodium bicarbonate, dried with magnesium sulfate, filtered and evaporated under reduced pressure. The resulting residue was purified by silica gel

chromatography with petroleum ether/AcOEt (4:1, v/v) to obtain compound **8** (36 mg, 26% yield) as a white powder.

$[\alpha]_D^{23}$= - 6.0 (c 1, CH_2Cl_2). ^{1}H NMR (500 MHz, $CDCl_3$): 3.92 (t, J = 6.5 Hz, 2 H), 2.57 (br s, 1 H), 1.60 – 1.29 (m, 14 H), 0.90 (t, *J* = 6.5 Hz, 6 H). ^{13}C NMR (125 MHz, $CDCl_3$): δ 69.3, 42.3, 37.1, 22.6, 22.5 22.1, 14.0. MS (ESI^+) m/z (%): 189.2 (10), 169.2 (60) $[M + H]^+$.

(*R*)-5-hydroxypentadecan-7-one (**7**) (162): Compound **5** (104.4 mg, 0.55 mmol) was dissolved in dry DMF under vacuum. 2 M octylmagnesium chloride in THF (0.6 mL, 3.2 mmol, Sigma-Aldrich) was added dropwise to the reaction mixture and cooled to -50 °C for 2 hours. The reaction mixture was quenched with ammonium chloride, extracted with ethyl acetate twice, dried with magnesium sulfate, filtered and evaporated under reduced pressure. The resulting residue was purified by silica gel chromatography with petroleum ether/AcOEt (4:1, v/v) to obtain compound **7** (70 mg, 67% yield) as a white powder.

$[\alpha]_D^{23}$= - 45 (c 1, CH_2Cl_2). ^{1}H NMR (400 MHz, $CDCl_3$): 4.01 (br s, 1 H), 3.05 (d, *J* = 3.2 Hz, 1 H), 2.58 (dd, *J* = 17.6, 2.8 Hz, 1 H) 2.51 – 2.39 (m, 3 H), 1.60 – 1.26 (m, 18 H), 0.91 – 0.85 (m, 6 H). MS (ESI^+) m/z (%): 189.2 (10), 243.1 (80) $[M + H]^+$.

(5*R*,7*R*)-pentadecane-5,7-diol (**9**; PDD) (163,164): Tetramethylammonium triacetoxyborohydride (304 mg, 1.2 mmol, Alfa Aesar) was added to a stirred solution of acetonitrile (1.45 mL, 27.5 mmol, Fisher) and acetic acid (1.45 mL, 50.7 mmol, Fisher) at 24 °C. After 30 minutes, compound **7** (70 mg, 0.29 mmol) was added and the reaction mixture was cooled to -78 °C for 40 minutes. The reaction stirred at -20 °C for 20 additional hours. The reaction mixture was extracted with saturated sodium bicarbonate, dried with magnesium sulfate, filtered and evaporated under reduced pressure. The resulting residue was purified by silica gel

chromatography with petroleum ether/AcOEt (4:1, v/v) to obtain compound **9** (17 mg, 24% yield) as a white powder.

$[\alpha]_D^{23}$= - 5.0 (c 1, CH_2Cl_2). ^{1}H NMR (400 MHz, $CDCl_3$): 3.94 – 3.90 (m, 2 H), 2.40 (br s, 1 H), 1.60 (t, *J* = 6.2, Hz, 2 H) 1.52 – 1.26 (m, 22 H), 0.92 – 0.85 (m, 6 H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 69.4, 42.2, 37.5, 37.1, 31.8, 29.6, 29.5, 29.2, 27.9, 25.7, 22.69, 22.65, 14.0. MS (ESI^+) m/z (%): 189.2 (10), 243.1 (80) $[M + H]^+$.

*Synthesis of 3R,5R-undecane-3,5-diol dipalmitate (**10**) and 3R,5R-pentadecane-3,5-diol dipalmitate (**11**)*

Compounds **10** & **11** were synthesized from diols **8** & **9** respectively according to the protocol reported for PDIM A synthesis (165,166). Compound **8** (10 mg, 53.11 µmol), palmitic acid (40.9 mg, 159.33 µmol, 3.0 eq, Sigma-Adrich), DCC (43.8 mg, 212.44 µmol, 4.0 eq) and DMAP (26.0 mg, 212.44 µmol, 4.0 eq) were dissolved in dry DCM (0.5 mL) and the resulting mixture was stirred at 25 °C for 48 hours under nitrogen. Compound **9** (10 mg, 41.08 µmol), palmitic acid (31.6 mg, 123.24 µmol, 3.0 eq, Sigma-Aldrich), DCC (33.9 mg, 164.32 µmol, 4.0 eq) and DMAP (20.1 mg, 163.32 µmol, 4.0 eq) were dissolved in dry DCM (0.5 mL) and the resulting mixture was stirred at 25 °C for 48 hours under nitrogen. For both reactions, the solvent was removed under reduced pressure and the product was purified by silica gel chromatography with hexanes/ethyl ether (50:1, v/v) to obtain compound **10** (30 mg, 85.0% yield) or **11** (16.9 mg, 57.1% yield) as a white wax.

3*R*,5*R*-undecane-3,5-diol dipalmitate: ^{1}H NMR (400 MHz, $CDCl_3$): 4.92 (m, 2 H), 2.27 (t, *J* = 8.0 Hz, 4 H), 1.53 – 1.77 (m, 10 H), 1.25 – 1.33 (m, 56 H), 0.88 – 0.90 (m, 12 H). ^{13}C (400

MHz, $CDCl_3$): δ 173.4, 70.1, 38.5, 34.6, 34.5, 31.9, 29.7, 29.6, 29.5, 29.4, 29.3, 29.2, 27.3, 25.0, 22.7, 22.6, 14.1, 14.0.

3*R*,5*R*-pentadecane-3,5-diol dipalmitate: ^1H NMR (400 MHz, $CDCl_3$): 4.92 (m, 2 H), 2.27 (t, *J* = 7.2 Hz, 4 H), 1.49 – 1.60 (m, 10 H), 1.26 – 1.32 (m, 64H), 0.87 – 0.90 (m, 12 H). ^{13}C (400 MHz, $CDCl_3$): δ 173.2, 70.0, 69.9, 38.3, 34.6, 34.4, 34.3, 31.8, 31.7, 29.6, 29.5, 29.4, 29.3, 29.2, 29.1, 29.0, 27.1, 25.0, 24.9, 22.6, 22.5, 22.4, 14.0, 13.8.

Cloning, expression and purification of 6xHis-PapA5

PapA5 (Rv2939) was cloned from H37Rv genomic DNA into pET28b (Novagen) via the NheI and HindIII restriction sites and a TEV protease cleavage site was inserted at the N-terminus before the 6xHis tag (Table 3-1). To investigate the putative substrate binding channels, the following mutations were made in the pET28b 6xHis-TEV-PapA5 vector by site-directed mutagenesis using overlapping primers (Table S1): G129L, S380M, S380F, A382M, Q19R, and Q19K, and V16C/G328C (generated with two successive mutagenesis reactions). The sequence-verified pET28b 6xHis-TEV-PapA5 constructs were transformed into *E. coli* BL21(DE3) cells. N-terminally 6xHis-tagged PapA5 was expressed and purified as described (143). PapA5 aliquots were snap frozen in 50 mM Tris pH 7.4, 100 mM NaCl, 1 mM DTT, 10% glycerol and stored at -80 °C.

Table 3-1. Expression constructs and oligonucleotide primers used in this study.

Name	Description	Mutation	Forward primer[1]	Reverse primer
6xHis-TEV-PapA5	pET28b-*NheI*-TEV-PapA5-*HindIII*		ATGGCTAGC*GAAAACCTGTACTTCCAGG GT*TTTCCCGGATCTGTGATCCG	CGCAAGCTTTCATCATCACTCCATGATCC AGCCATAC
G129L	6xHis-TEV-PapA5 G129L	G129L	CGCTATACCTCCATCACTGCATGGCCGA TCTTCATCACGGGGCCGTTCTCG	TCGTCGACGAGAACGGCCCCGTGATGAA GATCGGCCATGCAG
Q19L	6xHis-TEV-PapA5 Q19L	Q19L	CGCTATACCTCCATCACTGCATGGCCGA TGTTCATCACGGGGCCGTTCTCG	AACACCGCGCAGCTGGATTGTCATGGAA GTAAAAACCTCGTACTTCGCGAAGACTTC CT
S380M	6xHis-TEV-PapA5 S380M	S380M	CGATCAGCGTCCCCCTCGATCTGTACAT GTGTGCCGTTTACGCAGGACA	CGATGATCAGTTGTCCTGCGTAAACGGC ACACATGTACAGATCGAGGGGGACGCT
A382F	6xHis-TEV-PapA5 A382F	A382F	CGTCCCCCTCGATCTGTACTCGTGTTTC GTTTACGCAGGACAACTGATCATCGAGC	CAGTTGTCCTGCGTAAACGAAACACGAG TACAGATCGAGGGGGACGCTGATCG
A382M	6xHis-TEV-PapA5 A382M	A382M	CGTCCCCCTCGATCTGTACTCGTGTATG GTTTACGCAGGACAACTGATCATCGAGC	CAGTTGTCCTGCGTAAACCATACACGAGT ACAGATCGAGGGGGACGCTGATCG
V16C/ G328C	6xHis-TEV-PapA5 V16C/G328C	V16C	CCGGATCTGTGATCCGAAAGCTGTCGC ACAGCGAGGAATGCTTCGCGCAGTAC	AAAACCTCGTACTGCGCGAAGCATTCCT CGCTGTGCGACAGCTTTCGGATCACAG
		G328C	TGGTGTGATTCAGCAGTCGGGGCTCCA CTTCTGCACGGCATTCGAAG	GAGGAGTTCCTTCGAATGCCGTGCAGAA GTGGAGCCCCGACTGCT
PknB 1-330	pCDFDuet-*NdeI*-PknB(aa1-330)-*KpnI*		AGAAGGAGATATACATATGATGACCACC CCTTCCCACC	TTACCAGACTCGAGGGTACCTTATTAGCC CACCGAACCGATGC
PknE 1-290	pCDFDuet-*NdeI*-PknE(aa1-290)-*KpnI*		AGGAGATATACATATGGATGGCACCGC GG	CAGACTCGAGGGTACCTTATTAGCTGCG CCGCAAGATGTC

[1] Restriction sites are underlined. TEV protease recognition site is indicated in italics.

Isotope-coded mass tagging

PapA5 (V16C/G328C) (70 μM) was allowed to thaw on ice. A 50-μL reaction containing 12 μM PapA5 (V16C/G328C), 50 mM HEPES pH 7, and 1% (w/v) sodium deoxycholate (SDC) was prepared at 25 °C before the addition of 5 mM 4-(2,5-dioxo-2H-pyrrol-1(5H)-yl)-N,N,N-trimethyl-d9-butan-1-aminium iodide (167). After incubating for 10 min at 25 °C, excess label was removed by adding 450 μL 50 mM HEPES pH 7 and centrifuging for 6 min at 12,000 x g in a centrifugal filter (Amicon Ultra-0.5mL 10k MWCO, Millipore). The wash step was repeated four more times to reduce the label concentration to 0.5 μM. tris(2-carboxyethyl)phosphine (TCEP, 4.5 mM final concentration) was added to the retentate and the solution was incubated for 2 h at 25 °C. The TCEP concentration was reduced to 5 μM by centrifugal washing and filtration as above. The reduced protein was incubated with 5 mM 4-(2,5-dioxo-2H-pyrrol-1(5H)-yl)-N,N,N-trimethylbutan-1-aminium iodide (167) and 1% SDC for 10 min at 25 °C, then 1 h at 70 °C. The labeled protein was digested with chymotrypsin (10 ng/μL final concentration) and the solution was incubated for 16 h at 25 °C. To quench the digestion and precipitate the SDC, 3.5% (v/v final concentration) trifluoroacetic acid (TFA) was added and the mixture was incubated for 10 min at 25 °C. The precipitate was removed by centrifugation for 10 min at 18,000 x g. The supernatant containing the digested peptides was recovered and an equal volume of supernatant and 2,5-dihydroxybenzoic acid (DHB) matrix was mixed and immediately spotted for peptide fingerprint analysis by MALDI-TOF (Bruker AutoFlexII).

PapA5 enzymatic reactions

Reactions contained 2 μM PapA5, 18 μM [1-^{14}C]palmitoyl-CoA (PerkinElmer), and 180 μM OCT, UDD or PDD or 100 mM heptanediol in 25 μL reaction buffer (MES pH 6.5, 100 mM NaCl). Reactions were initiated by the addition of PapA5, incubated for 12–16 h at 20–25 °C and quenched with an equal volume of ethanol. Specific activity was determined from reactions incubated at 20–25 °C for 45 min. A 2 μL aliquot of each quenched reaction was spotted on silica thin-layer chromatography (TLC) plates (high-performance TLC Silica Gel 60; EMD Chemicals) and developed in 3:1 hexanes:ethyl acetate. For biphasic reactions, immediately after all components were combined, twice the reaction volume of hexanes was added slowly as a separate phase on top of the aqueous reaction mixture. Biphasic reactions were not quenched; instead, 3 μL of the hexanes layer was spotted for TLC. All TLC plates were visualized by phosphorimaging and densitometry analysis was performed using ImageJ software. For isolation of monoester PDD, a 1.25 mL reaction contained 2 μM PapA5, 18 μM [1-^{14}C]palmitoyl-CoA and 420 μM PDD in reaction buffer was incubated for 16 hours at 25 °C. The reaction was spotted on a TLC plate and developed in 85:15 hexanes:ethyl acetate. The silica containing monoester product was scraped off the plate and the monoester product was extracted from the silica with hexanes. The resulting monoester product was incubated with 2 μM PapA5 and 18 μM [1-^{14}C]palmitoyl-CoA for 16 hours at 25 °C. The reactions were spotted and TLC developed in 3:1 hexanes:ethyl acetate.

Mass spectrometry analysis of PapA5 reactions

Biphasic reactions contained 2 μM PapA5, 18 μM palmitoyl-CoA and 180 μM phthiocerol analogue in 500 μL reaction buffer with 1 mL hexanes overlaid on the aqueous reaction mixture. The hexanes layer was removed after 12-16 h incubation at 20–25 °C and allowed to evaporate to

dryness. The resulting film was re-solubilized in 12 µL 2-propanol per sample and analyzed by UPLC-ESI-MS as previously described (168).

Kinase phosphorylation assays

Mtb kinases Pkn B, D, E, H, K and L were purified as His-tagged fusions to maltose-binding protein (MBP-Pkn) and were a kind gift of Dr. Christina Baer (169). Radiolabeled phosphorylation reactions contained 2.8 µM PapA5, 1 µM MBP PknB. Final reaction volumes were 25 µL and contained 20 mM PIPES (pH 7.2), 5 mM $MgCl_2$, 5 mM $MnCl_2$, 400 µM ATP, and 10 µCi [γ-^{32}P]-ATP (PerkinElmer, 10 mCi/mL). Reactions were incubated at 30 °C for 20 min and were quenched by the addition of SDS-PAGE sample buffer with DTT. The reactions were analyzed by SDS-PAGE (12% acrylamide) with Coomassie blue staining. After drying, the gels were analyzed by autoradiography (22 h exposure). For the kinase activity screen with Pkn B, D, E, H, K and L, radiolabeled phosphorylation reactions contained 1 µM MBP-Pkn (all variants) and 1 µM PapA5 or myelin basic protein (MyBP). Final reaction volumes were 50 µL and contained 50 mM Tris (pH 7.4), 5 mM $MgCl_2$, 5 mM $MnCl_2$, 50 µM ATP, 1 µCi [γ-^{32}P]-ATP (PerkinElmer, 10 mCi/mL). Reactions were incubated at 30 °C for 1 h and quenched by the addition of SDS-PAGE sample buffer. Proteins were loaded and separated by SDS PAGE (8% acrylamide for PapA5 gel; 12% acrylamide for MyBP gel), stained, and analyzed as above (24 h exposure).

Co-expression with Pkn and purification of 6xHis-PapA5

The PknB (Rv0014c) catalytic domain (amino acids 1-330) was cloned from *Mtb* H37Rv genomic DNA into pCDFDuet-1 (Novagen) via the NdeI and KpnI restriction sites (Table S1). PknE (Rv1743; aa 1-290) was similarly cloned into pCDFDuet-1 (Table S1). PknB and PknE

expressed from the final vectors are untagged. The sequence-verified constructs pCDFDuet PknB 1-330 and PknE 1-290 were each transformed into *E. coli* BL21(DE3) containing pET28b 6xHis-TEV-PapA5. Double-transformed cells were grown in LB broth (Neogen) with 30 µg/mL streptomycin and 30 µg/mL kanamycin. Cultures were grown at 37 °C with shaking at 250 rpm to an OD_{600} of 1.0, induced with 1 mM isopropyl-β-D-thiogalactopyranoside and incubated for an additional 4 hours. Cells were lysed in 20 mM Tris, pH 7.4, 200 mM NaCl, 15 mM imidazole, 1 mM EDTA, 1 mM DTT (lysis buffer). The clarified crude lysate was incubated with Ni-NTA agarose resin (Qiagen) for 30 min at 4 °C. It was subsequently washed with 10 column volumes of lysis buffer and 15 column volumes of 50 mM Tris, pH 7.4, 100 mM NaCl, 1 mM DTT, 10% glycerol (storage buffer) with 15 mM imidazole. PapA5 was eluted in storage buffer with 250 mM imidazole. Fractions containing purified PapA5 were pooled and dialyzed overnight at 4 °C into storage buffer. PapA5 was co-expressed with PknE under the same conditions as above with PknB. The protocol for purification under denaturing conditions was adapted from the Qiagen manual (The QIAexpressionist 6/2013). Buffers used for the purification were variations on 100 mM NaH_2PO_4, 10 mM Tris, 8 M urea. Buffer D1 was adjusted to pH 8.0, buffer D2 was adjusted to pH 6.3, buffer D3 was adjusted to pH 5.9 and included 300 mM imidazole, and buffer D4 was adjusted to pH 4.5 and included 300 mM imidazole. The cells were resuspended in 5 mL buffer D1 per 1 g cell pellet weight and stirred for 30 minutes at 25 °C. Cells were then lysed by sonication and the lysate was spun for 30 min at 10,000 × g to pellet unbroken cells. Lysates were incubated with Ni-NTA resin equilibrated in buffer D1 (1 mL resin slurry per 4 mL lysate) for 60 min at 25 °C. Resin was pelleted at 1,000 × g for 2 min to removed flow through. The resin was then washed with 1 × 5 mL buffer D1 and 3 × 5 mL buffer D2. PapA5 was eluted from the resin in 3 × 1 mL buffer D3

and 3 × 1 mL buffer D4 and fractions 1-5 were pooled and concentrated using a 30 kDa MWCO concentrator (Millipore).

Mass spectrometry analysis of purified PapA5

PapA5 purified as above was diluted in 100 mM ammonium bicarbonate, reduced with 4 mM DTT and alkylated with 8.4 mM iodoacetamide. The protein was then digested with trypsin (Trypsin Gold, Mass Spectrometry Grade, Promega, USA) at a 25:1 protein:trypsin mass ratio for 16 h at 37 °C. The digests were brought to 2% (v/v) formic acid (FA) and desalted with Supel-Tips C18 Micropipette Tips (Sigma-Aldrich) using FA-containing solutions containing varying proportions of acetonitrile (ACN) according to vendor instructions. Eluted peptides were dissolved in 2% ACN, 0.1% FA (buffer A) for analysis on an LTQ Orbitrap XL ion trap mass spectrometer (Thermo Fisher, San Jose, CA) equipped with a nano-liquid chromatography electrospray ionization source. The peptides were eluted at 300 nL min-1 from a homemade 5 μm ProntoSil 120-5-C18H (Bischoff Chromatography, Leonberg, Germany) capillary column [2% buffer B (98% ACN, 0.1% FA) to 40% buffer B over 115 min; 40% to 80% buffer B over 3 min and held for 3 min; 80% to 2% buffer B over 0.1 min and held for 29 min]. Full mass spectra (MS) were recorded over m/z 400 to 2000, 60,000 resolution, followed by top-five MS/MS scans in the ion trap. Peptides with a charge /state of +2 or higher were analyzed. MS/MS spectra were extracted from the RAW file with ReAdW (http://sourceforge.net/projects/sashimi). The resulting mzXML data files were searched with Inspect against a custom database composed of the Uniprot EColi_K12 proteome with added sequences for the expressed PapA5 proteins and common contaminants. The data were also analyzed using the GPM X! Tandem and MaxQuant Andromeda search engines.

Results

PapA5 activity with alkyl-1,3-diol substrates

The substrate specificity of PapA5 has been probed previously with an array of commercially available alkyl alcohols that included primary aliphatic alcohols (C_2-C_{20}), tertiary alcohols (2-alkyl alcohols, C_6 and C_8) and both vicinal and ß-diols (C_6-C_{10}) (143,144). Of these, the highest activity was observed with the primary aliphatic alcohol 1-octanol (OCT). In comparison, all diols showed at least 5-fold less product formation under the same conditions, although even with these low levels of activity a proportion of diester was observed (2-4% compared to OCT product formation) (143). Given the strong preference for primary alkyl alcohols over diol substrates, the question of whether dual esterification is a physiological function for PapA5 still remained.

We hypothesized that PapA5 discriminates both the relative position of the two hydroxyl groups and the chain length on either side of the ß-diol. We therefore tested PapA5 activity with the commercially available 3*S*,5*S*- and 3*R*,5*R*-heptanediols (SHD and RHD, respectively) since they are longer than previously tested 2,4-pentanediols, which were not esterified (143). PapA5 was expressed and purified as previously described, and purity was confirmed by Coomassie stain (Figure 3-8A) and tryptic digest LC-MS/MS (>90% spectral counts assigned to PapA5). When incubated with PapA5 and ^{14}C-palmitoyl-coenzyme A (PCoA) as the acyl donor, products consistent with the monoester and diester were detected by thin-layer chromatography (TLC) only at substrate concentrations 200 times higher than the K_m of 0.5 mM for OCT (Figure 3-2A) (143). Therefore, the data from these substrates did not support diesterification as a native function of PapA5. Interestingly, RHD, which has the same stereochemistry as phthiocerol, gave a higher

yield of monoester (6.7 ± 1.2% of input for RHD vs. 1.9 ± 0.24% for SHD), suggesting that PapA5 can discriminate substrate stereochemistry at the first acylation step.

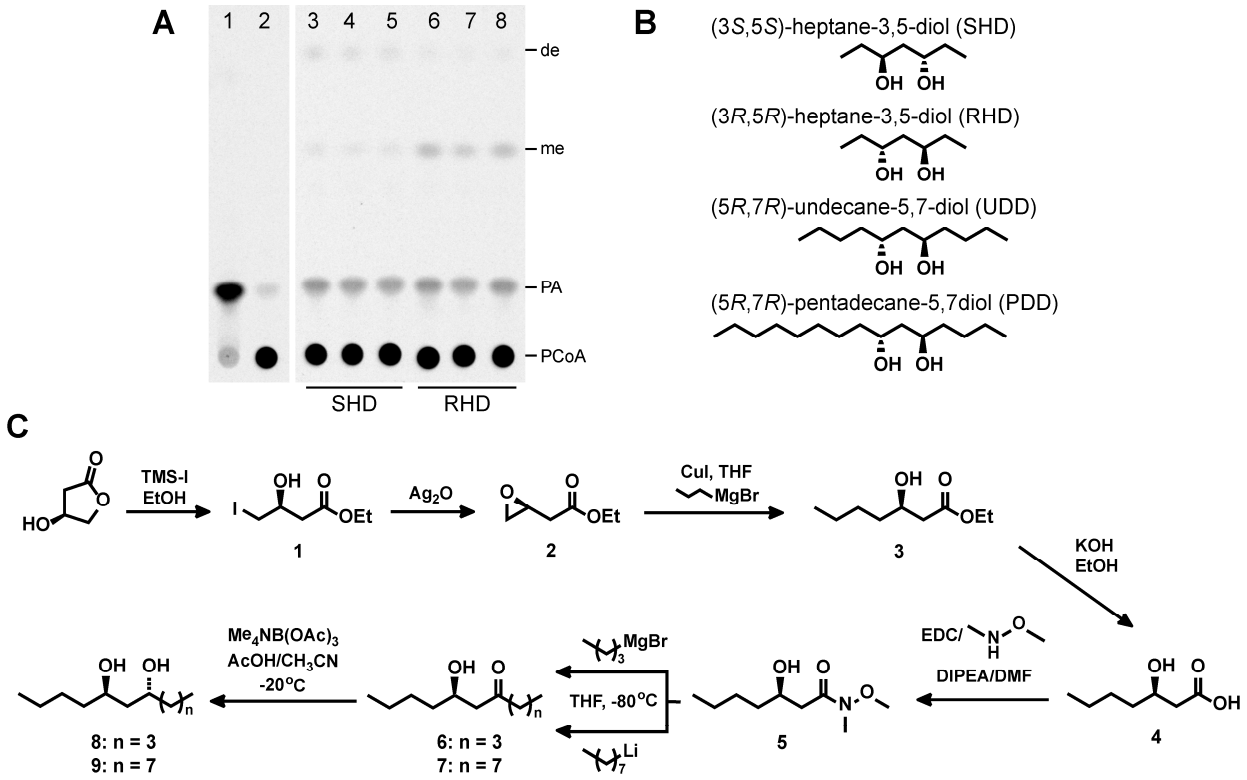


Figure 3-2. Substrate analogues recapitulate phthiocerol diol regio- and stereoselectivity. (A) PapA5 (2 μM) was incubated overnight with 18 μM [^{14}C-1]-palmitoyl-CoA (PCoA) and 100 mM 3*S*,5*S*-heptane-3,5-diol (SHD; lanes 3-5) or 3*R*,5*R*-heptane-3,5-diol (RHD; lanes 6-8). The migration of the diesters (de) and monoesters (me) are indicated. Triplicate reactions for each substrate are shown. Products were analyzed by TLC (3:1 hexanes:ethyl acetate) and analyzed by phosphorimaging visualization and densitometry. PCoA and [1-^{14}C]-palmitic acid (PA; product of competing PCoA hydrolysis) were included as migration standards (lanes 1-2). (B) Structures of SHD, RHD and the longer-chain analogues 3*R*,5*R*-undecane-5,7-diol (UDD) and 3*R*,5*R*-pentadecane-5,7-diol (PDD). (C) Synthetic scheme for UDD (**8**) and PDD (**9**).

We then synthesized the analogues 5*R*,7*R*-undecane-5,7-diol (UDD) and 5*R*,7*R*-pentadecane-5,7-diol (PDD) to recapitulate the stereochemistry of the native phthiocerol and to extend the chain length (Figure 3-2B,C). The desired anti stereochemistry was obtained using well

documented methods for the reduction of the ß-*R*-hydroxy ketone precursor with high diastereoselectivity (~95:5 anti:syn) (Figure 3-2C) (170).

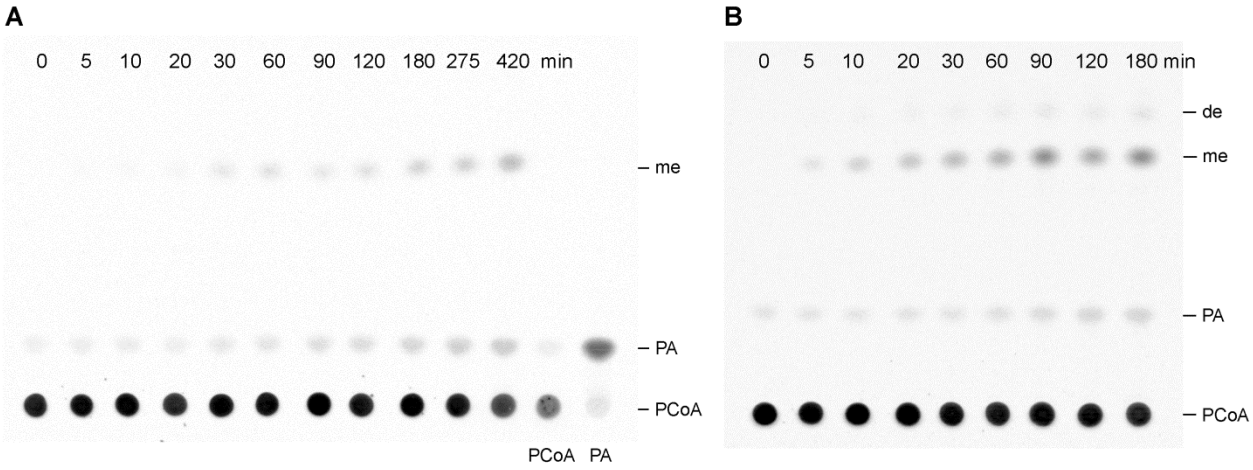


Figure 3-3. PDD diester forms under aqueous conditions at high diol concentration. PapA5 was incubated with 18 μM [1-^{14}C]-PCoA and (A) 360 μM UDD or (B) 360 μM PDD in buffer. Reactions were quenched at the indicated timepoints and analyzed by phosphorimaging and densitometry analysis. The diester formed in (B) indicates that the monoester product competes effectively with PDD for binding to PapA5 under these conditions and therefore the Michaelis-Menten steady-state assumption for monoester formation no longer holds. PA and PCoA (shown in A) served as migration standards. (me = monoester, de = diester)

Under previously reported aqueous buffer conditions (143), both UDD and PDD yielded a single product spot by TLC (Figure 3-4A, lanes 4-6). However, mass spectrometry confirmed that these products were the monoester forms (data not shown). Nevertheless, these conditions revealed that PapA5 has a higher specific activity with UDD and PDD than OCT and also prefers the longer-chain PDD over UDD for formation of the monoester (Figure 3-4B). Also, the apparent K_m for UDD (~150 μM) was 3-fold lower than for OCT (500 μM) and the k_{cat} for UDD was four orders of magnitude faster (2.84 μM^{-1} min^{-1} vs. 4.4×10^{-5} μM^{-1} min^{-1}) (Figure 3-4C). In contrast, at PDD concentrations ≥ 360 μM, the diester product was detected, indicating that under these conditions monoester competes with PDD for binding to PapA5 (Figure 3-3). Thus, the Michaelis-Menten

steady-state assumption no longer holds and the rate of monoester formation could not be directly determined. Nevertheless, the results with UDD still underscore that, unlike previously tested substrates, PapA5 is significantly more active with UDD than OCT.

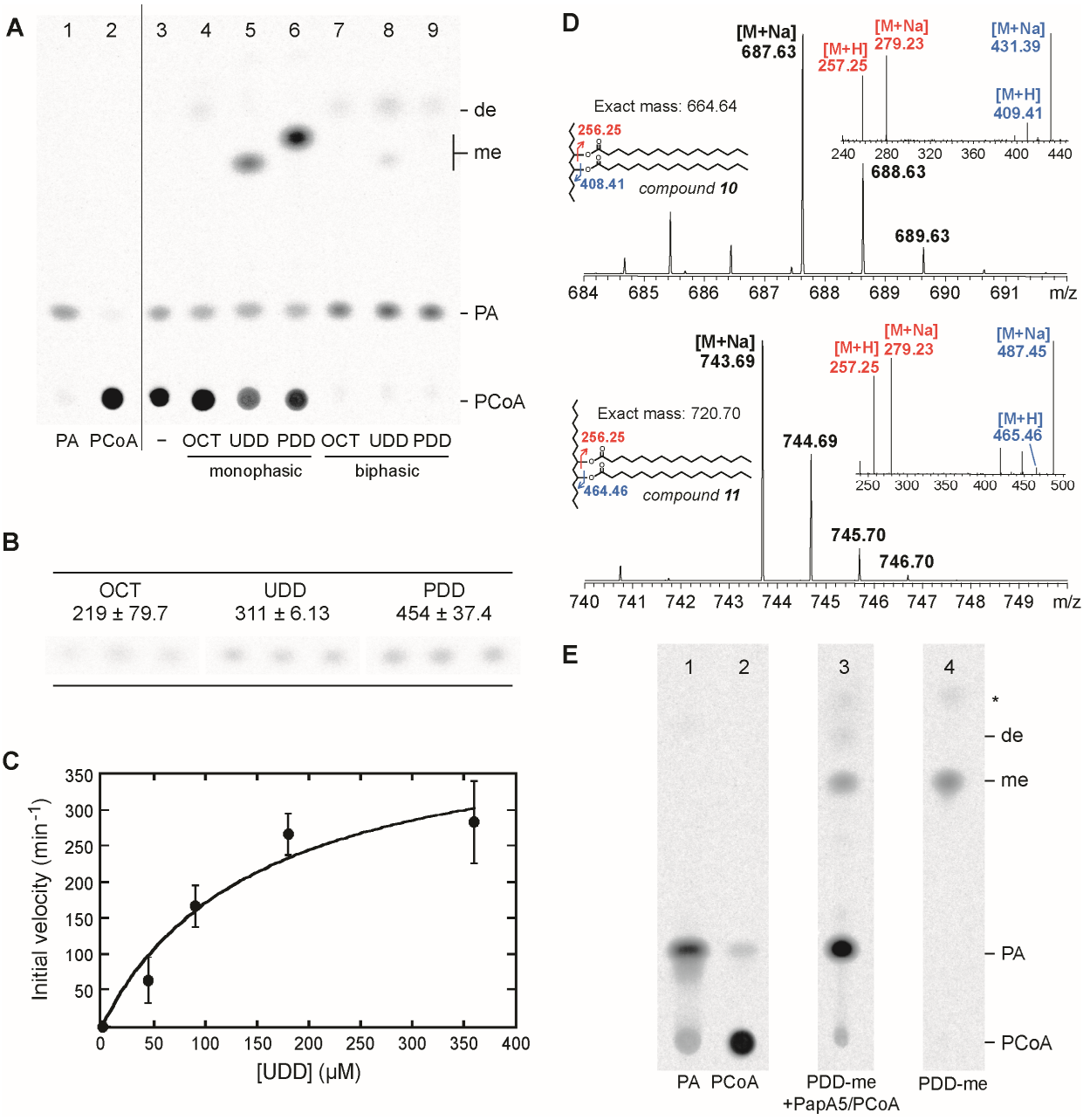


Figure 3-4. PapA5 diacylates ß-diol substrates and prefers longer-chain substrates. (A) PapA5 was incubated overnight with 18 μM [1-^{14}C]-PCoA and DMSO vehicle (lane 3) or 180 μM of the indicated substrate analogues under monophasic aqueous conditions (lanes 3-6) or biphasic

buffer:hexanes conditions (lanes 7-9; only organic phase spotted). Products were analyzed as in Figure 3-2A. Palmitoyl-OCT and the diesters (de) of UDD and PDD have similar R_f values in this mobile phase. Monoesters (me) of UDD and PDD are also indicated. OCT served as a positive control for PapA5 activity. Data are representative of three replicate experiments. (B) Specific activity for monoester formation in μM min^{-1} mg^{-1} PapA5 at 18 μM [1-^{14}C]-PCoA and 180 μM of the indicated substrate under monophasic conditions (22 °C, 45 min.). Activities are the average of three replicates (shown) ± S.D. (C) Initial velocity of palmitoyl-UDD formation as a function of UDD concentration (K_m = ~150 μM, k_{cat} = 2.84 μM^{-1} min^{-1}). (D) The diacyl products from biphasic reactions with UDD (left) and PDD (right) were confirmed by UPLC/ESI-tandem mass spectrometry in positive mode. The sodium adduct ion was observed at the predicted m/z for both substrates (top row). MS2 analysis (bottom row) of the m/z 687.63 and m/z 743.69 parent ions confirmed fragmentation at the ester bond (R = -$C_{15}H_{31}$). (E) Reactions were performed as in (A) except PapA5 was incubated with PDD-[1-^{14}C]-monopalmitate (PDD-me) and PCoA (lane 3; only organic phase spotted). PA and PCoA (lanes 1 and 2 in both A and B) and TLC-purified PDD-me (lane 3 in B) served as migration standards. The asterisk indicates an impurity carried over from purified PDD-me.

We hypothesized that the hydrophobicity of the monoester product leads to aggregation that hinders access by PapA5 and inhibits diester formation. This could also explain why only the monoester product was observed when PapA5 was incubated with phenolphthiocerol (142). We therefore screened PapA5 activity with PDD in a series of phase-separated aqueous-organic systems (1:2 buffer:solvent where the solvent was diethyl ether, chloroform, ethyl acetate, heptanes or hexanes) and found that the highest, most reproducible yield of PDD diester was obtained with hexanes (data not shown). We further verified that PapA5 activity with OCT was similar in aqueous and biphasic buffer:hexanes conditions and that PapA5 is therefore not adversely affected by the inclusion of the organic layer (Figure 3-5). Buffer:hexanes (2:1) was therefore used for all subsequent experiments.. Under these conditions, additional products with higher R_f for both UDD and PDD were observed by TLC (Figure 3-4A, lanes 8-9) and confirmed as the predicted dipalmitate products using tandem mass spectrometry (Figure 3-4D) and comparison to synthetic UDD-dipalmitate (**10**) and PDD-dipalmitate (**11**) standards (data not shown). Finally, PDD-1-^{14}C-monopalmitate was isolated by preparative TLC and shown to be a

substrate for PapA5, thereby confirming that PapA5 successively esterifies the alkyl diol to form the diester (Figure 3-4E).

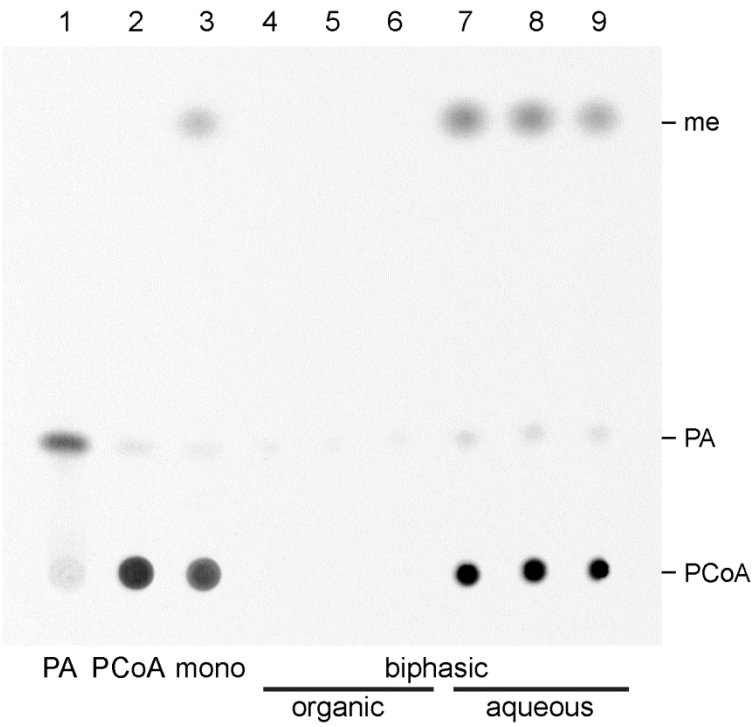


Figure 3-5. Biphasic reaction conditions do not affect PapA5 activity with OCT. PapA5 was incubated overnight with 18 μM [1-^{14}C]-PCoA and 180 μM OCT in reaction buffer (monophasic; lane 3) or in 1:2 buffer:hexanes (biphasic; lanes 4-9). Reactions were performed in triplicate and the organic (lanes 4-6) and aqueous phases (7-9) were spotted separately. The reaction volume spotted in lanes 7-9 was twice that loaded in other lanes. PA and PCoA (lanes 1 and 2) served as migration standards. (me = monoester)

Analysis of mutations that affect substrate binding to PapA5

Based on the PapA5 crystal structure and its similarity to other acyl-CoA-dependent acyltransferases, Buglino *et al.* described two putative substrate binding channels and hypothesized that the channels accommodate the two arms of phthiocerol with the hydroxyl nucleophiles positioned near the active site (Figure 3-6A) (144). We therefore predicted that bulky mutations at different locations along channel 1 would differentially affect the turnover of OCT,

UDD and PDD based on their chain lengths, whereas a bulky mutation in the predicted pantetheine-binding channel should have the same effect for all substrates (Figure 3-6A). Buglino *et al.* also noted that an additional helix, helix H, shortens channel 2 in PapA5 relative to that of related acyltransferases and proposed that this helix must move to allow phthiocerol bind. Thus, mutations designed to constrain helix H by forming a salt bridge or disulfide bond with the adjacent helix A would be predicted to decrease turnover of UDD and PDD, but not OCT since this primary alcohol would not require access to channel 2 (Figure 3-6A).

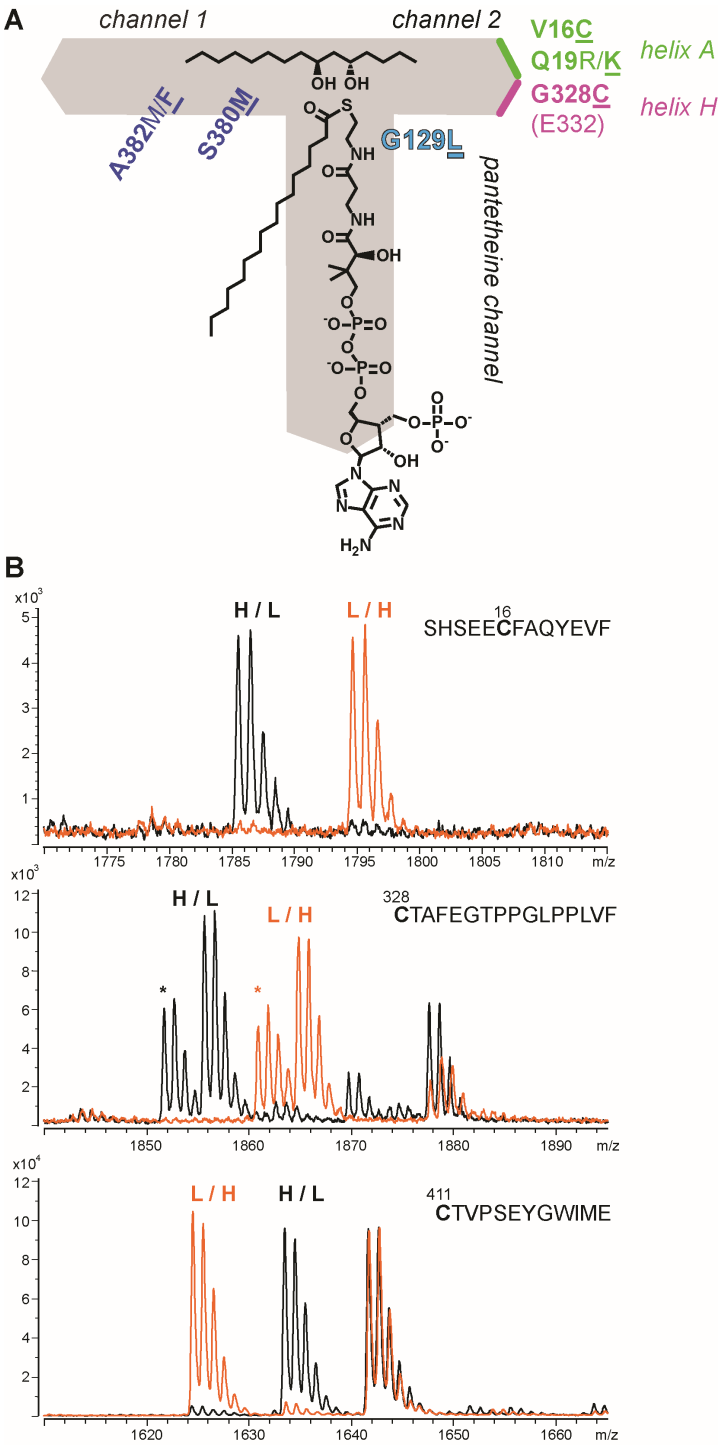


Figure 3-6. Substrate-blocking mutations modulate PapA5 activity. (A) Schematic of the predicted phthiocerol and acyl-pantetheine binding channels based on previous analysis by Onwueme *et al*. of the PapA5 crystal structure (PDB ID: 1Q9J). Bulky mutations were introduced as indicated to disrupt substrate binding. Mutations were also made to introduce a salt bridge (Q19R, Q19K) to E332 or a disulfide bond (V16C/G328C) between helices A and H. Mutants that displayed less than 20% wild-type enzyme activity for all substrates are indicated in bold underline

(see Table 2). (B) Isotope-coded mass tagging of the V16C/G328C mutant. PapA5 was treated with isotope-tagged labels before and after reduction (orange = light/heavy, L/H; black = heavy/light, H/L). Peptides corresponding to C16 and C328 were labeled only after treatment with reducing agent, confirming that the cysteines are oxidized. The native cysteine C411, which was not expected to be oxidized, served as a control. Peaks marked with an asterisk were also observed in an identically treated wild-type PapA5 control sample (data not shown).

To test these predictions we created a series of mutations (Figure 3-6A). The purified mutant proteins were all isolated solubly and in good yield, and the mutations did not have a significant effect on the overall PapA5 protein fold as indicated by circular dichroism spectra (Figure 3-7). As predicted, G129L, which introduces steric hindrance in the pantetheine-binding channel and would therefore affect PCoA binding, reduced PapA5 activity for the three substrates (Table 3-2). In channel 1, S380M completely abrogated activity for all substrates, consistent with the proximity of this position to the active site (~12 Å from H124, which is required for activity and thought to serve as a catalytic base (143)). However, contrary to predictions, the effects of all other mutations also did not vary with substrate. For example, A382 is ~15 Å from H124 and a bulky sidechain would be expected to hinder PDD (~12 Å for the longer arm) and OCT (~11 Å), but not UDD (~7.5 Å). While mutation of A382 to Phe had a more severe effect on activity than mutation to Met, the level of catalytic impairment of each mutant was similar with all three substrates. For mutations designed to constrain helix H, mutation of Q19 to Lys reduced activity more than mutation to Arg, suggesting that Lys is more effective at forming a salt bridge with E332 and thereby constraining helix H than Arg. The double cysteine mutant V16C/G328C forms a disulfide bond between these two residues on helix A and H, as verified by the dependence of isotope-labeled maleimide modification upon treatment with reducing agent (Figure 3-6B) (167). The effect of the disulfide bond on PapA5 activity was more severe than for the salt bridge Q19K mutation, but again, the reduction in catalytic activity was similar across all substrates. Since all

mutants had similar circular dichroism spectra to the wild type (Figure 3-7), the observed changes in activity are likely due to the specific mutation rather than unfolded or unstable protein.

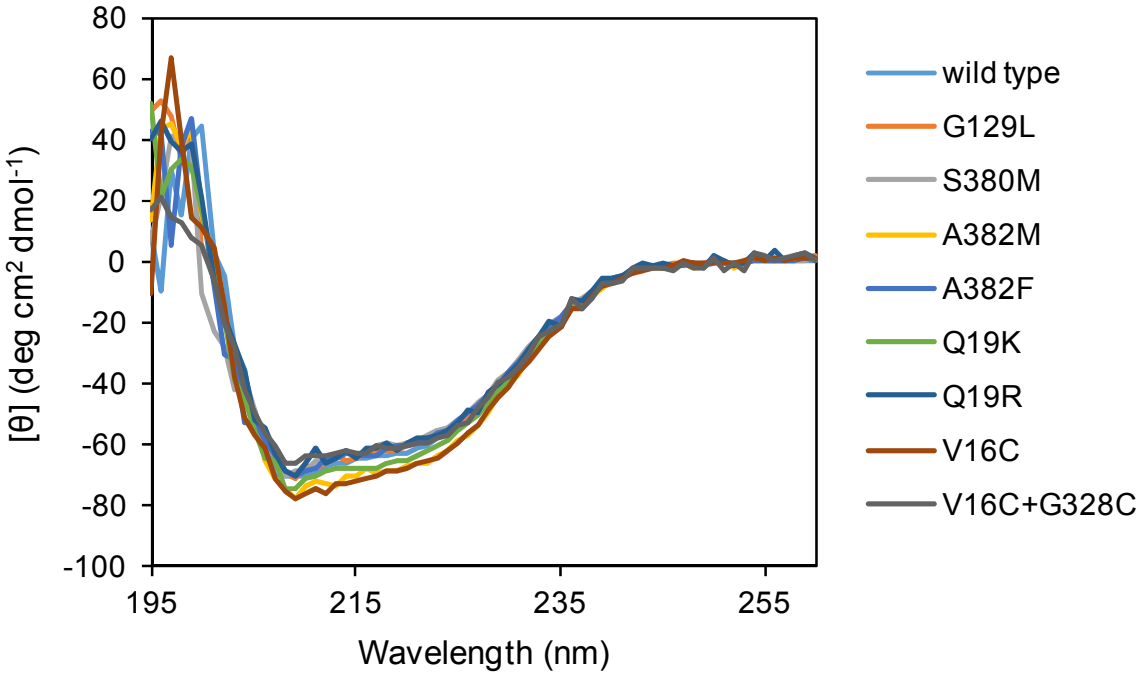


Figure 3-7. Circular dichroism of purified wild-type and mutant PapA5. Wild-type, PapA5, as expressed and purified from *E. coli*, display similar circular dichroism spectra. These data indicate that the mutants have similar secondary structure to the wild-type and therefore the mutations do not have a significant effect on the protein fold.

Table 3-2. Specific activities of PapA5 mutants for formation of the monoester.

	Specific Activity ($\mu M\ min^{-1}\ mg^{-1}$)					
Mutation(s)	**OCT**		**UDD**		**PDD**	
wild type[1]	219 ± 79.7		311 ± 6.13		454 ± 37.4	
G129L	28.2 ± 7.2	**(13%)**[2]	11.4 ± 3.3	**(4%)**	58.8 ± 17.5	**(13%)**
S380M	< 5.0		< 5.0		< 5.0	
A382M	76.0 ± 7.7	(35%)	121 ± 12.1	(39%)	220 ± 16.1	(48%)
A382F	< 5.0		< 5.0		18.7 ± 14.9	**(4%)**
Q19R	150 ± 14.6	(68%)	150 ± 10.4	(48%)	249 ± 36.7	(55%)
Q19K	35.8 ± 6.5	**(16%)**	< 5.0		50.4 ± 10.5	**(11%)**
V16C/G328C	< 5.0		26.1 ± 9.5	**(8%)**	13.6 ± 7.0	**(3%)**

[1] Raw data is shown in Figure 3-4B.

[2] Percent wild-type activity. Mutants that displayed less than 20% wild-type activity for all substrates are indicated in bold.

Phosphorylation of PapA5 by Mtb Ser/Thr kinases

Recently, biosynthetic enzymes involved in mycolic acid biosynthesis have been shown to be phosphorylated by Ser/Thr kinases with consequent downregulation of catalytic activity. For example, phosphorylation of the fatty acid synthase II (FASII) enoyl-ACP reductase InhA reduces activity by lowering the affinity for the NADH cofactor (157). Inhibition of β-ketoacyl acyl carrier protein synthase KasB upon phosphorylation is attributed to the proximity of the modified Thr residues to the catalytic triad (159). Since PDIM levels are modulated by growth conditions, infection or the presence of PknH, we hypothesized that *Mtb* Ser/Thr kinases may also regulate PDIM production by modifying PapA5. We first confirmed that PapA5 is phosphorylated only in the presence of kinase and [γ-^{32}P]-ATP and thereby confirmed that PapA5 is a substrate for PknB, as previously observed (Figure 3-8A) (154). We then tested Pkn D, E, H, K and L for their ability to phosphorylate PapA5 as well. All kinases were active, as demonstrated by autophosphorylation

and phosphorylation of myelin basic protein (MyBP), a model substrate that is phosphorylated by many *Mtb* Ser/Thr kinases (171,172), but only PknB and PknE phosphorylated PapA5 (Figure 3-8B).

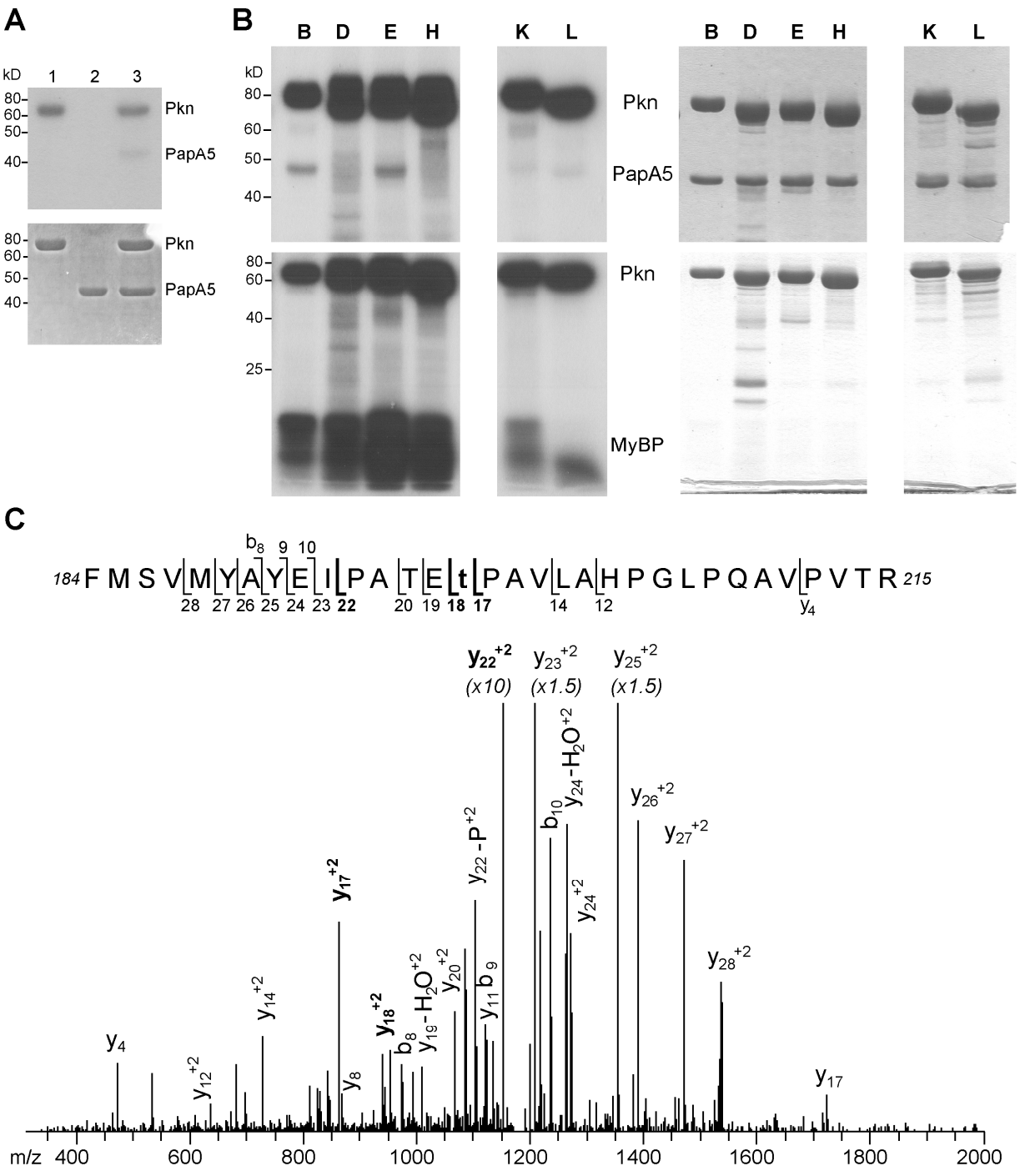


Figure 3-8. PknB and PknE phosphorylate PapA5. *Mtb* Pkn kinase domains were purified as fusions to MBP and incubated with γ-^{32}P-ATP. (A) PknB autophosphorylates (lane 1) and PapA5 is phosphorylated only in the presence of PknB (lane 3). Autoradiogram (24-hour exposure; top) and Coomassie-stained SDS-PAGE gel (bottom) (B) PknB, D, E, H, K and L were each incubated with PapA5 (top) or myelin basic protein as a positive control (MyBP; bottom). Autoradiogram (24-hour exposure; left) and Coomassie-stained SDS-PAGE (right). Approximate M.W.: MBP-Pkn ~80 kDa; PapA5 49 kDa; MyBP 18 kDa. (C) An example MS2 spectrum for the aa184-215

peptide of PapA5 co-expressed with PknB shows phosphorylation of T198 (indicated by **t** in the peptide sequence). The triply charged ion has predicted m/z 1179.58 (monoisotopic) and m/z 1180.36 (avg); the parent ion was detected at m/z 1180.25. Phosphorylation at T198 is supported by the peak for **y18** and the strong proline peaks for **y17** and **y22**.

PknB phosphorylation of threonines on PapA5 has been reported, but the location and functional consequences of the modification have not been examined (154). To determine the effect of phosphorylation on catalytic activity, PapA5 was isolated from *E. coli* cells co-expressing the PknB kinase domain, but no change in activity was detected relative to unmodified PapA5 using OCT, UDD or PDD as substrates (data not shown). To verify phosphorylation by PknB and determine the site(s) of modification, PapA5 was analyzed by tandem mass spectrometry. PapA5 expressed by itself showed no evidence for phosphorylation as determined by three independent search programs (MaxQuant, InsPecT and GPM X!Tandem). In contrast, phosphorylation was detected in PapA5 isolated from PknB co-expressing cells. Specifically, both MaxQuant and InsPecT detected phosphorylation on the peptide aa184-215. Although quantification of the degree of phosphorylation was not possible, manual inspection of the MS2 spectra confirmed peptides with modification at T196, T198 or T214. A representative spectrum for phosphorylation at T198 is shown in Figure 3-8.

PknE phosphorylation of PapA5 has not been previously reported. Upon co-expression with PknE, PapA5 was insoluble and was purified under denaturing conditions. PapA5 expressed by itself and co-expressed with PknE were isolated in parallel under denaturing conditions and subjected to tandem mass spectrometric analysis as above. Both MaxQuant and GPM X!Tandem detected phosphorylation at T196, T198 or T214, suggesting overlap with PknB phosphorylation sites. One additional phosphorylation site at T144 was detected by both algorithms (Figure 3-9C). T144 is located at the end of helix C, the N-terminus of which forms part of the active site. The

T144 sidechain hydroxyl is appropriately positioned to form a hydrogen bond with the backbone amide proton of T355 (T144 hydroxyl O to T355 amide N atom-to-atom distance, 3.3 Å). PknE phosphorylation of T144 may disrupt this interaction and destabilize PapA5, leading to the insolubility and aggregation observed upon co-expression in *E. coli*.

Discussion

Our observation that PapA5 catalyzes dual esterification with a preference for longer-chain ß-diols provides strong support for the assignment of PapA5 as a diacyltransferase and the final enzyme in DIM biosynthesis (Figure 3-1). Furthermore, the results of our mutational analysis are not consistent with the previously reported model for substrate binding to PapA5 and suggest that a revised model is required. Indeed, after the structure of PapA5 was reported, structures of mammalian carnitine acyltransferases in complex with CoA, carnitine or substrate-competitive inhibitors revealed that channel 1 accommodates the alkyl chain of the acyl-CoA substrate with the electrophilic center at the active site, while channel 2 binds to the carnitine nucleophile (173-176). In line with these structures and our activity data, we propose that in PapA5 mycocerosic acid binds to channel 1 and phthiocerol binds to channel 2 (Figure 3-9A). None of the mutants in channel 1 (S380M, A382M/F) showed significant discrimination between OCT, UDD and PDD substrates despite the varying chain lengths, suggesting that channel 1 does not bind to the diol substrate, but rather accommodates the alkyl chain of the acyl donor. In this model, constraining helix H affects substrates of all chain lengths equally, as observed, since helix H is only ~7 Å from the active site H124. Major structural rearrangements proximal to helix H would therefore be required to position the hydroxyl nucleophile(s) at the active site while accommodating the two arms of phthiocerol. Flexibility in this region of PapA5 is suggested by regions around helix H

that are unresolved in the crystal structure (aa 176-180 and aa 192-204) (144). For the second esterification, channel 2 may accommodate the mycocerosate chain as well as the phthiocerol of the monoester substrate. Given the size of phthiocerol monomycocerosate, channel 1 may also be required to bind to the monoester, although this possibility could not be determined from the existing mutants and activity data.

Our revised model for substrate binding, in which phthiocerol and mycocerosic acid occupy distinct channels that orient the nucleophile and electrophile in the active site at the channel junction, implies that phthiocerol is only partially bound by PapA5. Even if helix H moves away, channel 2 is too shallow (~11 Å from H124 to the surface) to accommodate the longer arm (C_{21}-C_{24}) of phthiocerol. Moreover, for the second esterification reaction PapA5 must access the unmodified hydroxyl on the highly hydrophobic phthiocerol monomycocerosate. Our results also show that for efficient diester synthesis PapA5 requires a hydrophilic-hydrophobic interface, which in *Mtb* could be provided by the cytosol-membrane interface. Although PapA5 is soluble, it may therefore be proximal to the membrane in *Mtb*, consistent with other soluble biosynthetic enzymes that act on lipid intermediates in cell wall biosynthesis and have been localized to polar sites of cell wall growth (177,178).

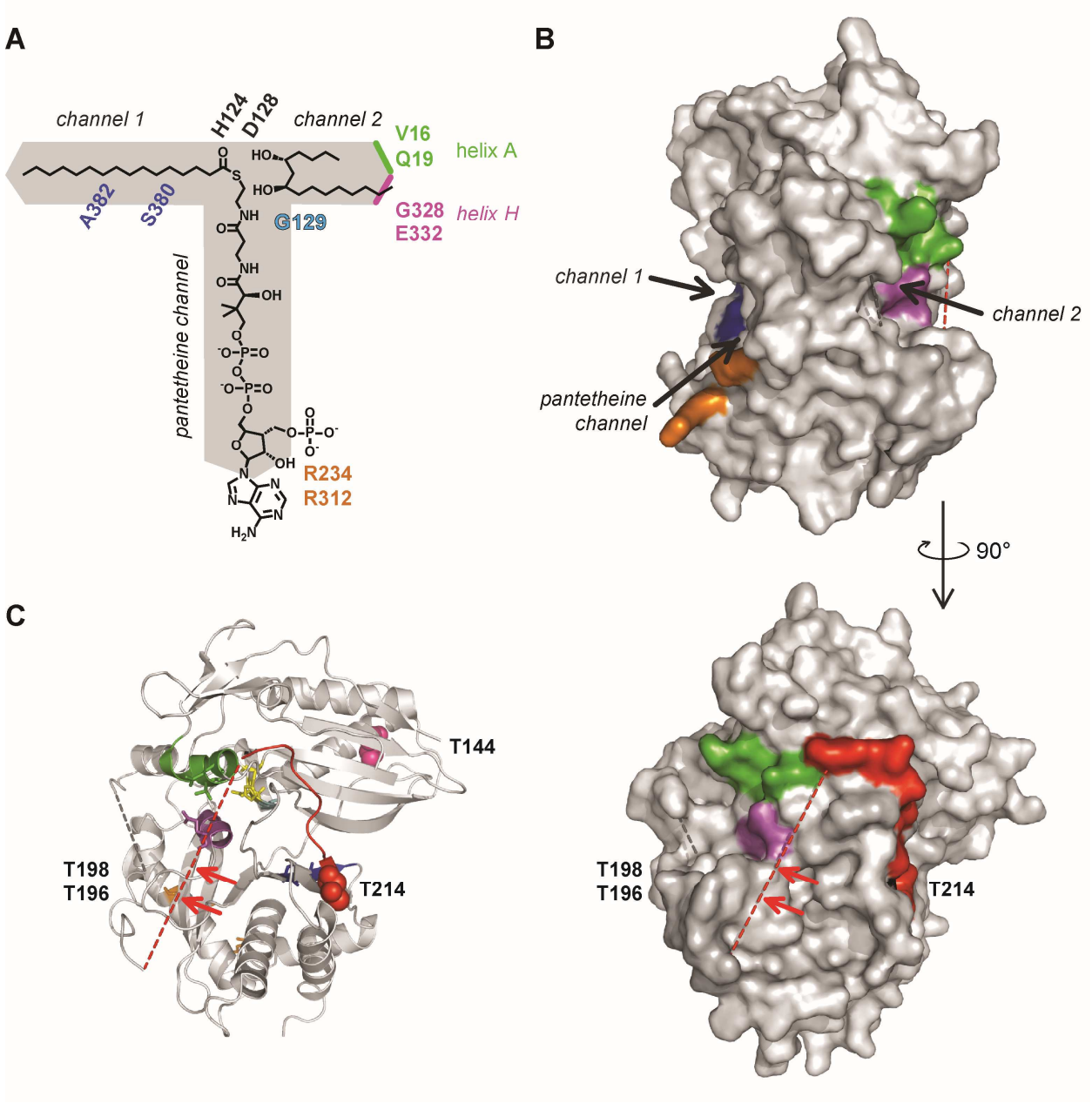


Figure 3-9. Model for PapA5 substrate binding and protein interactions. (A) Modified schematic of the substrate channels showing the proposed binding modes for the acyl acceptor (here, palmitate) in channel 1 and the alkyl diol (PDD) in channel 2. (B) PapA5 crystal structure (PDB ID: 1Q9J) showing the apposition of the Mas-ACP-interacting residues R234 and R312 (orange) and helices H and A. Dashed lines designate unresolved regions. (C) The region highlighted in red encompasses the residues T196, T198 and T214 that are phosphorylated by both PknB and PknE. Residue T144 (pink) is also phosphorylated by PknE. Arrows indicate the approximate positions of T196 and T198. Catalytic residue H124 and active-site residue D128 are shown in yellow.

PknB and PknE phosphorylate PapA5 at overlapping but distinct sites

Unexpectedly, PknB phosphorylation did not affect PapA5 activity *in vitro*. Although phosphorylated PapA5 was generated in a recombinant overexpression system, modification by PknB in this context appears to be specific, as phosphorylation was detected at only 3 of 28 Thr residues in PapA5, ~20 of which are at the protein surface. Interestingly, the modified residues T196 and T198 are located in the middle of an unresolved region (aa192-204) at the surface of PapA5 (Figure 3-9C). Along with T214, these residues encompass a region proximal to helix H and the entrance to the putative phthiocerol binding channel and may therefore modulate the interaction of PapA5 with other proteins (Figure 3-9C). The phosphorylation of T196 and T198 in particular supports this idea, as unstructured segments are common motifs in protein-protein interactions (179). Enzymes that catalyze successive steps can form multiprotein complexes to promote efficient turnover and access to substrates, and interactions have been reported between the DIM transporter MmpL7 and PpsE; between PpsE and the PDIM-associated thioesterase TesA; and between the acyl carrier domain of mycocerosic acid synthase (Mas-ACP) and PapA5 (180-182). Residues implicated in the interaction with Mas-ACP (R234 and R312) are located on a facet orthogonal to T186, T198 and T214 (Figure 3-9B,C), suggesting that post-translational modification is unlikely to affect Mas binding (182). Rather, phosphorylation of these residues may modulate PapA5 interactions with the negatively charged plasma membrane (132,183) or with other phthiocerol biosynthetic enzymes, particularly the terminal phthiocerol enzyme PpsE, with consequences for DIM levels in the cell. Interestingly, co-expression with PknE resulted in insoluble PapA5, perhaps due to the disruption of a hydrogen bond by phosphorylation at T144. While PknB and PknE modifications overlap at T196, T198 and T214, PknE phosphorylation may affect PapA5 stability by modifying T144 and thus regulate PapA5 in a manner distinct from PknB

modification. These results, along with previous reports of MmpL7, Mas, and Pks1/15 phosphorylation, warrant further investigation into the potentially diverse roles of phosphorylation in modulating DIM biosynthesis (153,155,156).

Chapter 4 The Fatty Acyl-CoA Ligase FadD6 is Capable of Generating Substrates for Polyketide-Associated Protein Acyltransferases *in situ*

Introduction

While not much is known about lipid transport in *Mtb*, there are some hints from a few specific examples. The lipoprotein LppX has been implicated in transport of the outer membrane (OM) lipid phthiocerol dimycocerosate (PDIM). PDIM is a wax ester, outer membrane, surface lipid found only in pathogenic mycobacteria. The *Mtb* gene knockout strain, Δ*lppX*, has a decreased virulence phenotype in the murine model. PDIM biosynthesis, however, is not affected as PDIM was found fully intact. PDIM accumulated in the IM and was not found to be released into the culture medium suggesting a role for LppX in transporting it across the periplasm to the OM (49). The interaction between PDIM and LppX has never been shown biochemically due to the extreme hydrophobic nature of the wax ester, which is incompatible with typical aqueous biochemical assays.

The structure of LppX comprises an α/β-fold with 11 antiparallel β-sheets that form a U-shaped, half β-barrel and 3 opposing α-helices. This creates a large hydrophobic cavity (25 Å × 12 Å × 12 Å ; ~2835 $Å^3$) capable of binding lipids, as suggested by the three fatty acids that co-purified and co-crystallized in this pocket (49). Generating lipid analogues would facilitate studying the putative LppX-PDIM relationship. We proposed to synthesize a bifunctional lipid analogue such that (i) it is less hydrophobic than the natural lipid and may be compatible with traditional biochemical assays, (ii) it contains a photoreactive moiety to covalently crosslink the lipid to its binding protein and (iii) it contains a functional group that allows for enrichment of the

lipid-protein complex. Figure 4-1 outlines this proposal, in which a PDIM analogue would be generated using the truncated 1,3-β-alkyl diol phthiocerol analogue synthesized and characterized in Chapter 2, 5*R*, 7*R*-undecane-5, 7-diol (UDD). UDD would be acylated with bifunctional fatty acids that each contain the relatively small photoreactive crosslinker diazarine and an alkyne that can be modified via copper-catalyzed cycloaddition with an azide-modified affinity tag, such as a 6xHis-tag.

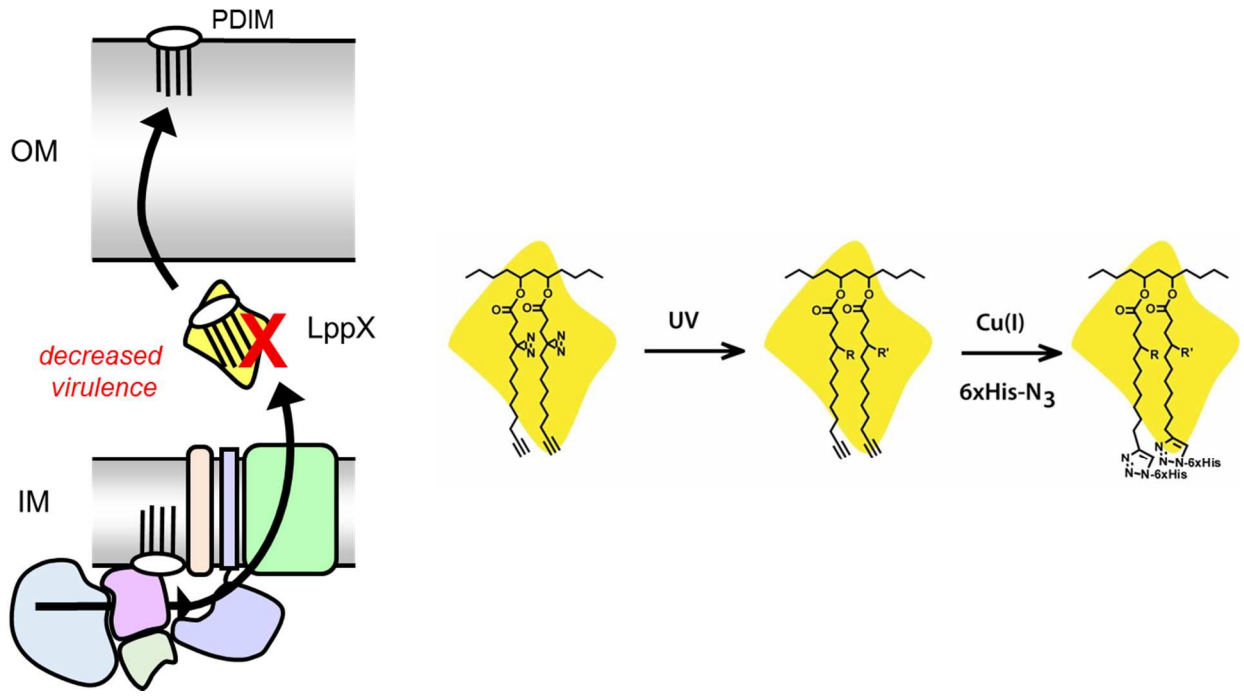


Figure 4-1. Generation of photoreactive lipid analogues to identify lipid-protein interactions. As a proof-of-principle experiment, the interaction between LppX and a bifunctional PDIM analogue will allow us to develop the method to identify other lipid-protein interactions. (*left*) In a strain lacking *lppX*, PDIM accumulates in the inner membrane, implicating LppX in transport of PDIM to the outer membrane. (*right*) Using bifunctional lipid analogues will allow us to covalently link to carrier proteins and subsequently purify the lipid-protein adduct.

The synthesis of a bifunctional, diazarine-alkyne fatty acid has already been reported (Figure 4-2A). We suggest using this methodology to synthesize a longer fatty acid shown since the chain-length specificity of FadD6 requires long chain fatty acids C_{12}-C_{18} (Figure 4-2B) (184). These fatty acids would then serve as substrates for synthesizing the lipid analogues chemoenzymatically by taking advantage of *Mtb* lipid biosynthetic enzymes. Previously reported chemoenzymatic synthetic approaches have used glycosyltransferases (185-187) to generate various glycosidic linkages in a highly stereo- and regioselective manner (185). Currently, there are no examples in the literature of lipid chemoenzymatic synthesis but there is an example system where α-2,6-linked sialosides were synthesized in a single-pot reaction containing three functionally different enzymes (186). The major advantages of chemoenzymatic synthesis include ease, simplicity of synthesis, short length of synthesis, regioselectivity, stereoselectivity as well as cost (185).

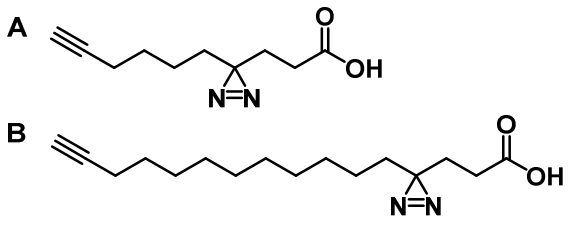


Figure 4-2. Bifunctional fatty acids proposed for use in chemoenzymatic synthesis of lipid analogues. (A) The synthesis of this fatty acid has been previously reported (188). (B) The structure of the proposed fatty acid to be used in generating *Mtb* lipid analogues.

We therefore propose to take advantage of the known chemistries of the polyketide-associated protein acyltransferases (PapA), within this mycobacterial protein family, PapA3 and PapA5 were described in Chapters 1 and 2 esterify hydroxyls on a sugar or alkyl diol with fatty

acids from an acyl-CoA donor. To do this, we require a fatty acid-CoA ligase to couple the bifunctional fatty acid with coenzyme-A (CoA) to generate the acyl-CoA substrate for the PapA enzymes. Fatty acyl-CoA ligases (FACL), as their name suggests, ligate fatty acids with coenzyme A (CoA) (Figure 4-3) (184). The ligase FadD6 is a FACL and has been characterized *in vitro* to ligate fatty acids with CoA as well as the inexpensive, truncated CoA analogue *N*-acetylcysteamine (SNAC). FadD6 has previously been described as promiscuous with respect to fatty acid chain-length specificity (184).

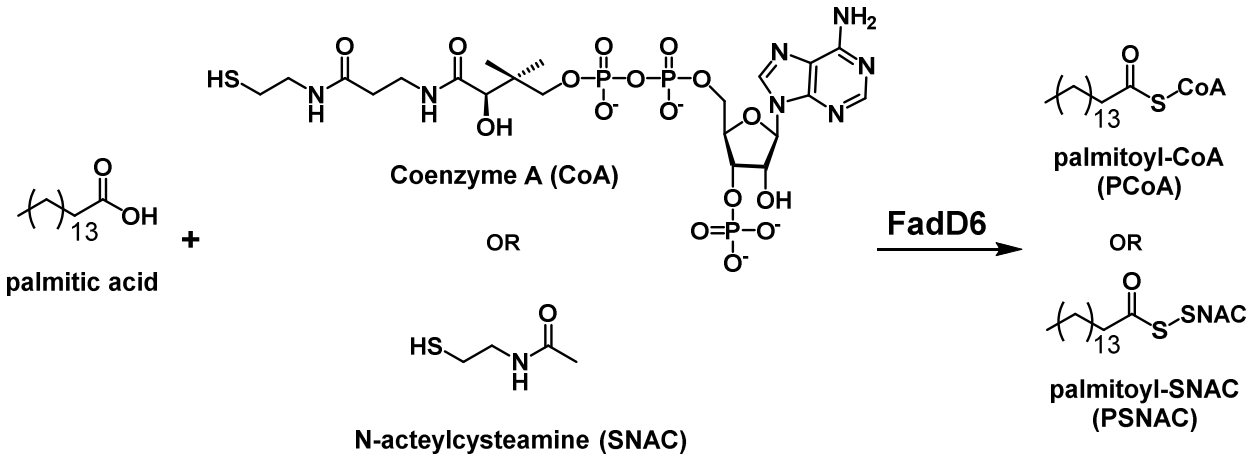


Figure 4-3. FadD6 ligates carboxylic acids with CoA to produce thioester adducts. FadD6 has been shown to have preference for long chain fatty acids (184). This schematic describes the ability of FadD6 to ligate the C_{16} fatty acid, palmitic acid, with CoA or SNAC to yield the palmitoyl-thioester product.

The ultimate goal of the project presented here was to chemoenzymatically synthesize a chemical probe to validate LppX as a PDIM-binding protein. Our strategy was to use PapA5 to generate a bifunctional PDIM analogue. PapA5 has been previously implicated in the final PDIM

biosynthesis step and acylates phthiocerol with mycocerosic acids (142,189). Our long-term intent would be to apply this technique to elucidate binding proteins for other virulence-associated, surface lipids polyacyltrehalose (PAT) and sulfolipid-1 (SL-1). PAT and SL-1 are also acylated by PapA family enzymes. PapA3 has been characterized as an acyltransferase in PAT biosynthesis and PapA1 and PapA2 have been characterized as acyltransferases in SL-1 biosynthesis. There are currently no clues for what proteins are involved in their transport from the IM to the surface of the OM. In this work, we characterize the ability of FadD6 to generate substrates *in situ* for PapA3 and PapA5 and therefore validate chemoenzymatic synthesis as a route to generate lipid analogues to interrogate interactions between surface lipids and their potential binding partners.

Experimental Procedures

Optimization of FadD6 overexpression conditions in E. coli.

A 4-mL culture of *E. coli* BL21(DE3) cells containing pET28b-Nterm6xHis-FadD6 was inoculated from a frozen glycerol cell stock and grown overnight at 37 °C at 250 rpm in Terrific Broth (BD Difco) supplemented with 50 μg/mL kanamycin. The starter culture was used to inoculate 100 mL Terrific Broth supplemented with 50 μg/mL kanamycin. At OD_{600} 0.5 and 1.0, 3 × 10 mL cultures were induced with 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG, Alfa Aesar) and were incubated further at 18, 25 or 37 °C. At time points 0, 2, 4, 6, 8, 10 hours post induction, 1 mL samples were removed and pelleted. Cell pellets were re-suspended in lysis buffer (20 mM Tris, pH 7.4, 200 mM NaCl, 1 mM EDTA, 1 mM DTT) to OD_{600} 1.0 and lysed by sonication 5 sec on/off for 2 min total sonication. The debris was cleared by centrifugation at 5,000 x *g* for 10 min. The cleared lysates were analyzed by SDS-PAGE and Coomassie Blue staining.

Expression and purification of His-tagged FadD6 in E. coli

A starter culture of *E. coli* BL21(DE3) cells containing pET28b-6xHis-FadD6 was inoculated from a frozen glycerol cell stock and cultured in Terrific broth (BD Difco) supplemented with 50 μg/mL kanamycin overnight at 37 °C with 250 rpm shaking. The starter culture was used to inoculate 1 L Terrific Broth supplemented with 50 μg/mL kanamycin. At OD_{600} 1.0, the cells were induced with 0.5 mM IPTG, cultured at 25 °C for 6 hours and harvested at 5,000 x *g* for 10 min. Cell pellets were re-suspended in 15 mL lysis buffer and lysed by sonication for 30 sec on/off for 10 min total sonication. The debris was cleared by centrifugation at 12,500 x *g* for 10 min. The clarified crude lysate was loaded onto a 5-mL HisTrap FF (GE Biosciences) column and washed with buffer A (100 mL 50 mM Tris, pH 7.4, 1 mM DTT, 10% glycerol). Bound protein was eluted over 20 column volumes in a gradient of 0-50% 1 M imidazole in buffer A. Fractions containing FadD6 were pooled and dialyzed overnight at 4°C into buffer A.

Enzymatic reactions

For reconstitution of FadD6 acyl-CoA ligase activity *in vitro*, enzymatic assays were conducted as previously described (184). Briefly, reactions contained 1 μM FadD6, 120 μM ^{14}C-palmitic acid, 8 mM $MgCl_2$, 2 mM ATP, 1 mM CoA or SNAC in 25 μL reaction buffer (50 mM Tris, pH 7.5). Reactions were initiated by the addition of ^{14}C-PA, incubated at 30 °C for 15 minutes, and quenched with an equal volume of 10% acetic acid. Reactions were analyzed by TLC (80/40/25, v/v/v, *n*-butanol:water:acetic acid; HPTLC glass, Silica gel 60, EMD) and visualized by phosphorimaging. For coupled enzymatic reactions with PapA enzymes, substrate

concentrations, buffers and TLC conditions were previously described in Chapter 1 and 2 for PapA3 and PapA5, respectively.

Results

Initial FadD6 expression and purification attempts yielded pure but inactive protein (data not shown). We therefore first validated the quality of all reagents used in the enzymatic assay. We tested a new ATP stock, used Ellman's reagent [5,5'-dithiobis-(2-nitrobenzoic acid) or DTNB] to confirm the concentration of the CoA stock and measured the circular dichroism (CD) spectra of purified FadD6. The Ellman's reagent assay confirmed CoA free thiol concentration (data not shown) and the CD spectra revealed a primarily alpha-helical fold that was partially denatured with heating at 50 °C (Figure 4-4). The structure of FadD6 has not been solved but the structure of *Mtb* FadD28, a protein from the same enzyme family, shows primarily alpha-helical fold (PDB: 3T5A), and therefore the data suggest that purified FadD6 is folded.

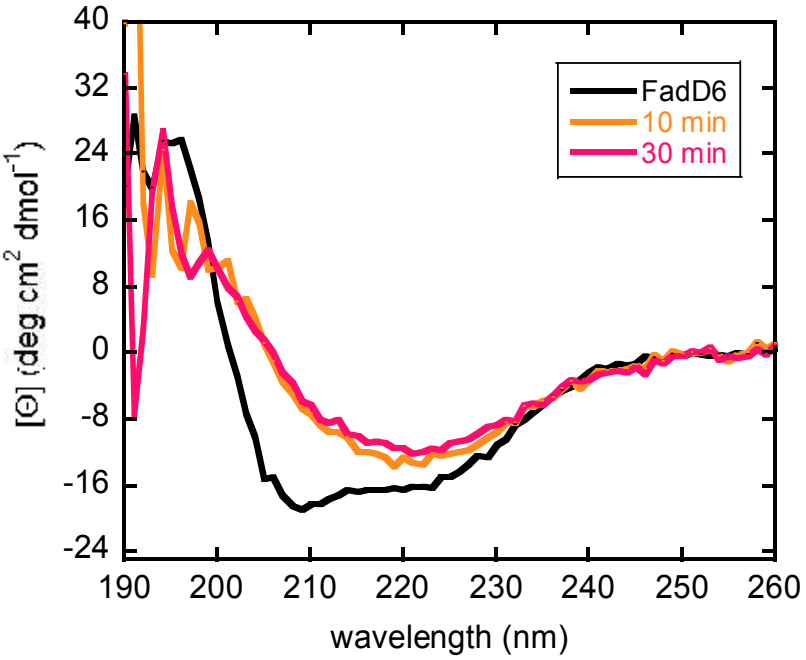


Figure 4-4. Circular dichroism spectra of purified FadD6. The black line shows the spectrum for 9 µM purified FadD6 and the orange and pink lines are the spectra for heated purified FadD6 at 50 °C for 10 and 30 minutes respectively.

We determined that induction at OD_{600} 1.0 for 6 hours at 25 °C yielded the greatest amount of soluble FadD6. FadD6 purified from these optimized expression conditions was then tested for *in vitro* fatty acyl CoA-ligase activity. TLC and Phosphorimage analysis revealed that FadD6 ligates palmitic acid to CoA and SNAC (Figure 4-5). A time course was conducted in order to determine conditions that would yield the maximal product formation (Figure 4-6). This revealed that incubating the reaction longer than 15 minutes did not increase the amount of product formed in the CoA and SNAC reactions. Approximately 15% of palmitic acid was converted to PCoA and approximately 38% of palmitic acid was converted to PSNAC.

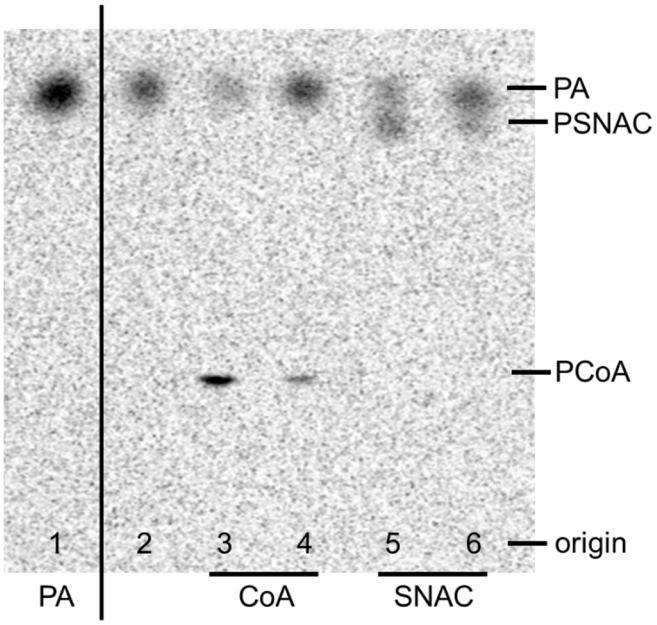


Figure 4-5. Phosphorimage of TLC shows FadD6 *in vitro* fatty acyl CoA-ligase activity. FadD6 was incubated with 120 µM ^{14}C-palmitic acid (PA), 1 mM CoA or SNAC and the cofactors ATP and Mg^{2+}. Lane 1 is the migration control for PA. Reactions 2-6 contained PA, FadD6, $MgCl_2$ and ATP. Reactions 3-4 differ in the ATP stock used, but both used the thiol substrate CoA. Reactions 5-6 also differed in the ATP stock used, but both used the thiol substrate SNAC.

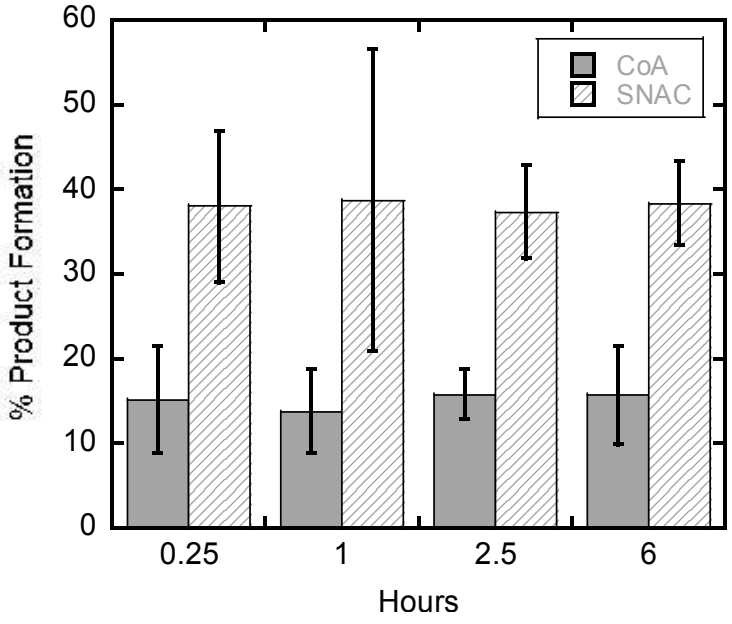


Figure 4-6. FadD6 reaction time course reveals maximum product formation occurs within 15 minutes and twice as much acyl-SNAC is formed compared to acyl-CoA. Reactions were initiated simultaneously and quenched 0.25, 1, 2.5 and 6 hours post PA addition. Percent product formation reported is the average of three replicates and the error bars represent ± standard deviation.

Figure 4-7 shows coupled enzymatic reactions containing both FadD6 and either PapA3 or PapA5 with all of their cofactors and substrates. Under these conditions, FadD6 is still capable of generating PCoA or PSNAC, but PapA3 and PapA5 only use PCoA as acyl substrates. In the FadD6-SNAC reactions, no PapA products were observed by TLC despite PSNAC formation. In the FadD6-CoA-PapA3 reactions, both trehalose-2-palmitate (TreP) and trehalose-2,3-palmitate ($TreP_2$) were detected as migrating alongside migration standards. However, in the FadD6-CoA-PapA5 reactions, only UDD monopalmitate was observed and not UDD dipalmitate.

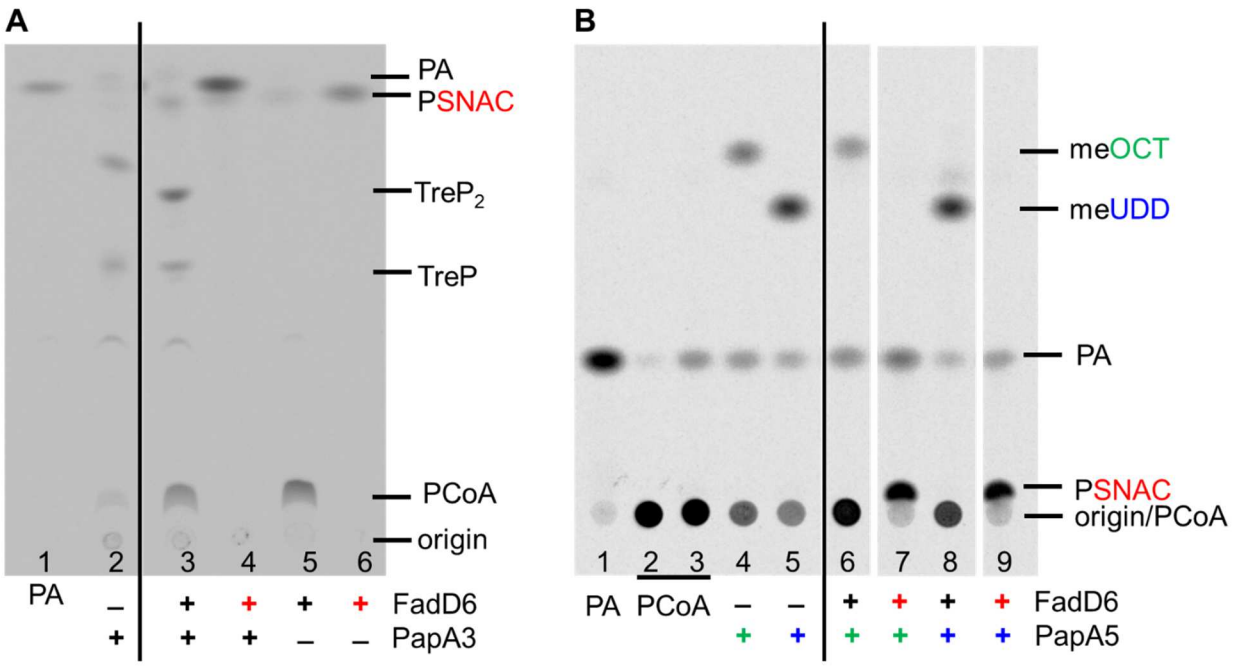


Figure 4-7. FadD6 produces PCoA substrates *in situ* for PapA3 and PapA5 use in lipid acylation. In both panels, reactions to the left of the vertical lines are migration controls. FadD6 reactions contained 120 µM ^{14}C-PA, 1 mM CoA or SNAC and the cofactors ATP and Mg^{2+}. In coupled reactions containing FadD6 and PapA, FadD6 generates PCoA in situ for PapA-catalyzed acylation of their respective head groups. PSNAC produced by FadD6, however, did not make a good substrate for the PapA enzymes. (A) PapA3 reactions contained 1 mM trehalose and lane 2 in panel A is a migration control for PapA3 products and contained 20 µM PCoA. (B) PapA5 reactions contained 180 µM 1-octanol (OCT) or 5*R*, 7*R*-undecane-5,7-diol (UDD) and lanes 4-5 in panel B are migration controls for PapA5 products and contained 20 µM PCoA. Lane 2-3 are both PCoA migration standards, but lane 3 was incubated at room temperature alongside enzymatic reactions and shows non-enzymatic hydrolysis of PCoA. PA, palmitic acid; PCoA, palmitoyl-CoA; PSNAC, palmitoyl-SNAC; TreP, trehalose-2-palmitate; $TreP_2$, trehalose-2,3-dipalmitate; meUDD, monoester UDD; meOCT, monoester OCT.

Discussion

We have shown that FadD6 has *in vitro* fatty acyl-CoA ligase activity (Figure 4-5) and that it is capable of using both CoA and SNAC as thiol substrates. FadD6 makes 2-3 times more acyl-SNAC than it does acyl-CoA and under the conditions tested, product formation is complete within 15 minutes. Although acyl-CoA is the physiological product, the lower yield of this product

compared to acyl-SNAC may be explained by PCoA inhibition of ligase activity since PCoA substrate inhibition has been demonstrated for PapA1, PapA2, PapA3 and PapA5 (108,189,190). Coupled FadD6 reactions with PapA3 and PapA5 demonstrated the feasibility of chemoenzymatic synthesis of lipid analogues (Figure 4-7). Both acyltransferases preferred PCoA over PSNAC as an acyl chain donor, suggesting that the PapAs require recognition of CoA beyond the pantetheine arm for activity. The crystal structure of PapA5 (see Chapter 2) supports this idea. PapA5 has a channel for acyl chain binding which is adjacent to a CoA-binding channel that suggests specific interactions of the pantetheine arm in this CoA-binding channel. It is also important to note that UDD diester formation this these reactions was minimal. This is likely due to the hydrophobicity of the diester product and is a function of the conditions in which the reactions were carried out.

We have shown that chemoenzymatic synthesis of lipids is possible using FadD6 and PapA enzymes PapA3 or PapA5. In each case, the PapA enzyme transferred the acyl chain from the acyl-CoA adducts produced by FadD6 to either trehalose or UDD. These data support a chemoenzymatic strategy for the production of lipid analogue containing the bifunctional fatty acids shown in Figure 4-2, since FadD6, PapA3 and PapA5 have been shown to have promiscuous recognition of acyl chains (108,184,189,191).

One drawback to this approach for identifying novel lipid-binding proteins is that these experiments would be most easily performed with cell lysates. Probing cellular lysates has the disadvantage that the setting is non-physiological for lipid-protein interactions and eliminating non-specific lipid-protein interactions due to general hydrophobic interactions would be difficult. Given that we propose to use OM lipid analogue probes and assuming the probes would incorporate into the OM, whole cells could be treated with the probe, thus allowing for the identification of OM proteins as a similar method previously described (192).

We have shown the feasibility of lipid analogue chemoenzymatic synthesis by demonstrating that FadD6-PapA coupled reactions generate lipid analogues from fatty acids. Based on this work, this method could be used to incorporate bifunctional fatty acids into lipid analogues. Before moving forward with the synthesis of the diazarine-alkyne fatty acid, optimization of the coupled reactions should be completed. For example, another aliquot of FadD6 could be added after 15 min to increase yield in the coupled reaction assuming that denaturation/deactivation, rather than product inhibition or substrate consumption, is the reason for low yield. The FadD6-PapA5 reactions could be optimized in biphasic conditions, as described in Chapter 2, to maximize diester product formation. These experiments are important next steps to assess the likelihood of generating an experimentally useful quantity of bifunctional, lipid analogue for subsequent biochemical crosslinking studies.

Chapter 5 Quantification of Mycobacterium tuberculosis Lipid-Protein Interactions Using a Competitive Binding Fluorescence Assay

Attribution: The cloning, expression, purification and de-lipidation of LppX used in this chapter were done by me. The characterization of LppX by the competitive binding fluorescence assay was done by my mentee, Julia K. Joseph, who has given permission to include this work in this dissertation.

Introduction

The success of *Mycobacterium tuberculosis* (*Mtb*) as a pathogen is attributed, in part, to its unusually hydrophobic cell envelope, which includes unique surface lipids. Research suggests that these lipids are important in host immune responses and pathogenesis (193). Despite the demonstrated importance of these surface lipids, relatively little is known about the mechanism by which they contribute to *Mtb* pathogenicity. A better understanding of the role of these unique surface lipids is vital for creating innovative therapeutic strategies.

Recent studies have concentrated on the surface lipid, phthiocerol dimycocerosate (PDIM), which has been implicated in *Mtb* virulence (194,195). PDIM plays a key role in bacterial engulfment by the macrophage and PDIM biosynthesis is essential for phagocytosis. It is hypothesized to interfere directly with host cell membrane functions, such as phagosome maturation. Specifically, PDIM contributes to the prevention of lysosomal acidification of *Mtb*-containing phagosomes, which allows the bacteria to survive and persist inside the macrophage (57).

PDIM is synthesized in the cytosol and is then transported to the mycobacterial outer membrane (OM), where it plays a role in virulence. Though the molecular mechanism by which this transport occurs is unknown, research suggests that the lipoprotein LppX plays a key role in

the process. The *Mtb* gene knockout mutant *ΔlppX* accumulates PDIM in the inner membrane (IM) but not in the OM (49). Significantly, this mutant demonstrated attenuated virulence in the murine model of infection. The crystal structure of LppX shows has a large hydrophobic cavity, which is hypothesized to be large enough to encapsulate PDIM. These data support the hypothesis that LppX plays an important role in the translocation of PDIM from the IM to the OM. Its mechanism of transport may be similar to LolA, a structural homologue of LppX, which binds lipoproteins and transports them through the periplasm of *E. coli* from the IM to the OM (196). A scheme representing a model for PDIM transport is shown in Figure 5-1 and this process is explained in detail in Chapter 6.

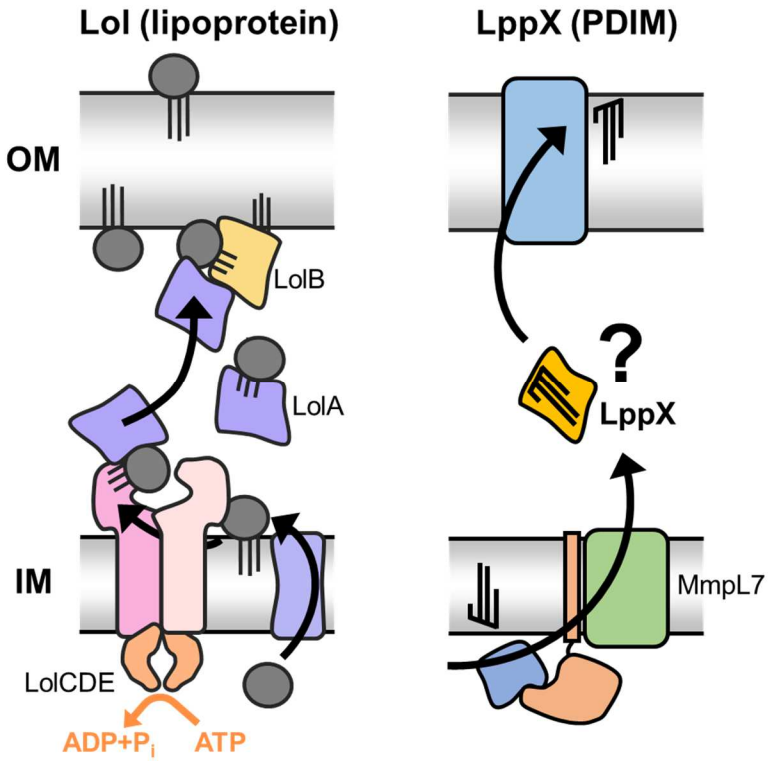


Figure 5-1. Proposed mechanism of PDIM transport by LppX may be analogous to lipoprotein transport in *E. coli*. (*left*) Schematic of Lol-mediated lipoprotein transport in *E. coli*. LolA is a chaperone that transports lipoproteins from the IM to the OM protein LolB. LolA and LolB are structurally homologous to LppX and have a large hydrophobic pocket to bind the acyl chains of lipoproteins. (*right*) LppX is hypothesized to bind PDIM and transport it from the inner

membrane (IM) to the outer membrane (OM). The *Mtb* gene knockout mutant Δ*lppX* yielded an accumulation of PDIM in the IM.

This work aims to elucidate the substrate specificity of LppX for PDIM by taking advantage of a competitive fluorescence assay, which has previously been used to interrogate ligand binding of sterol carrier-2 and fatty acid binding proteins (FABPs) (197). The assay first characterizes the binding of a fluorescent reporter to a protein of interest, which in this work, is LppX. The reporter, in this case, either NBD-stearate or bis-ANS, is excited at a specific wavelength. The resulting emission spectrum is used to characterize the environment of the fluorescent reporter. For instance, the hydrophobic fluorescent reporter used in this work was hypothesized to emit higher signal at a shorter wavelength upon binding the hydrophobic pocket of LppX, as compared to when the emission spectrum was detected without the presence of LppX. This could be due to its interactions within the hydrophobic binding pocket. The binding of the reporter to LppX would then be used to characterize the K_d of the reporter and the protein.

Once this interaction was characterized, a competing ligand was then added to displace the fluorescent reporter from the binding pocket. The resulting emission spectrum was hypothesized to decrease due to the displacement of the reporter from the hydrophobic pocket. The decrease in signal with the addition of competing ligands would then be used to determine the K_d of the ligand for the protein. Therefore, this assay indirectly characterizes ligand binding to protein through competition with the fluorescent reporter (198).

By validating a competitive fluorescence assay for LppX, important insights into the molecular determinants of PDIM recognition by LppX may be uncovered. In addition, the results of this study may corroborate the assay's potential as a general method to quantify interactions with lipids and their respective lipid-binding proteins in *Mtb*.

Experimental Procedures

Cloning, expression and purification of LppX from E. coli

LppX (Rv2945c) was cloned from H37Rv genomic DNA into pET28b (Novagen) via the NdeI and HindIII restriction sites and a TEV protease cleavage site was inserted at the N-terminus before the 6xHis tag. The sequence-verified pET28b-6xHis-TEV-LppX constructs were transformed into *E. coli* BL21(DE3) cells. N-terminally 6xHis-tagged PapA5 was expressed and purified as described (49). PapA5 aliquots were snap frozen in buffer A (50 mM Tris pH 7.4, 100 mM NaCl, 1 mM DTT, 10% glycerol) and stored at -80 °C.

To investigate the putative ligand binding pocket (Figure 5-2), the following mutations were made in the pET28b-6xHis-TEV-LppX vector by site-directed mutagenesis using overlapping primers (

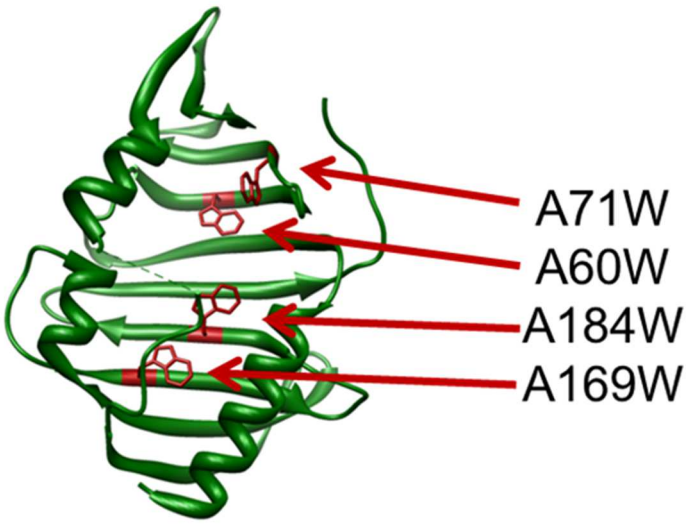


Figure 5-2. Location of Ala to Trp mutations in apo-LppX. Four alanines within the hydrophobic cavity of LppX were mutated to tryptophan to sterically hinder different parts of the binding pocket. A71W and A60W mutations were made to disrupt the hypothesized entrance of the binding pocket, while A184W and A169W were made to disrupt the inner part of the pocket.

Table 5-1): A60W, A71W, A169W, A184W, A60W/A184W (generated with two successive mutagenesis reactions). The sequence-verified mutants were purified as described above.

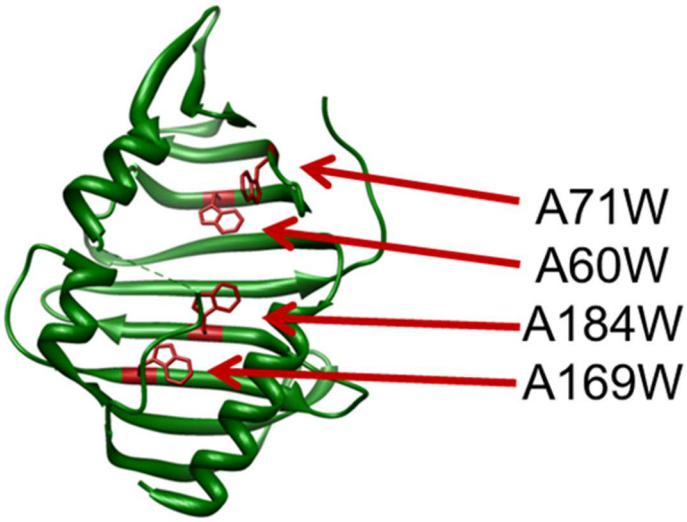


Figure 5-2. Location of Ala to Trp mutations in apo-LppX. Four alanines within the hydrophobic cavity of LppX were mutated to tryptophan to sterically hinder different parts of the binding pocket. A71W and A60W mutations were made to disrupt the hypothesized entrance of the binding pocket, while A184W and A169W were made to disrupt the inner part of the pocket.

Table 5-1. Expression constructs and oligonucleotide primers used in this study.

Name	Description	Mutation	Forward Primer	Reverse Primer
6xHis-TEV-LppX	pET28b-6xHis-TEV-LppX		CGCGGCAGCCATATGGAAAACCTGTACT	CGAGTGCGGCCGCAAGCTTCTACTACTAGTCGA
A60W	pET28b-6xHis-TEV-LppX-A60W	A60W	CGACAGCTTGCTGGGTATTACCAGTTGGGATGTCGACGTCCGGGCCAATCCGCTCGCG	CAACCGGGAAAGTCGACAGCTTGCTGGGTATTACCAGTTGGGATGTCGACGTCCGGGCC
A71W	pET28b-6xHis-TEV-LppX-A71W	A71W	GCCAATCCGCTCGCGTGGAAGGGCGTATGCACCTACAACGACGAGCAGG	GACGTCCGGGCCAATCCGCTCGCGTGGAAGGGCGTATGCACC
A169W	pET28b-6xHis-TEV-LppX-A169W	A169W	AGTGCAAGGCCGTGGACCGTGTGGATTGCCCAGGACGGCTCGCAC	CCTGGCGCCAAGAGTGCAAGGCCGTGGACCGTGTGGATTGCC
A184W	pET28b-6xHis-TEV-LppX-A184W	A184W	CACCACCTCGTCCGATGGAGCATCGACCTCGGATCCGGGTCGATTCAGCTCACG	CAGGACGGCTCGCACCACCTCGTCCGATGGAGCATCGACCTCGG

Removal of fatty acids bound to ligand-binding pocket of LppX

Fatty acids bound in the ligand binding pocket of LppX were removed using hydroxyalkoxypropyl dextran type VI (Sigma-Aldrich) resin, which has a high affinity for hydrophobic molecules, such as fatty acids. The resin was loaded in a gravity column and equilibrated in buffer A at 25 °C. Purified LppX was added to the column contain resin and incubated at 37 °C for 30 min with stirring every 5 min (binding capacity for 1 g resin is 10 mg protein). LppX was eluted from the resin at 25 °C in 3 x volume of LppX added. De-lipidated LppX is referred to in this work as apo-LppX.

Characterization of reporter-protein binding interactions by fluorescence spectroscopy

12-*N*-methyl-(7-nitrobenz-2-oxa-1,3-diazo)aminostearic acid (NBD-stearate), an 18-carbon fatty acid modified with an extrinsic fluorescent group, was purchased from Avanti Polar Lipids, Inc. This fluorescent reporter has been used previously to characterize substrates for lipid-binding proteins (197). As a hydrophobic fatty acid, NBD-stearate was predicted to also bind to the hydrophobic cavity of LppX. NBD-stearate was dissolved in dimethyl sulfoxide (DMSO) and kept at a stock concentration of 50 mM, while stored at -20 °C and protected from light. The structure of NBD-stearate can be found in Figure 5-3.

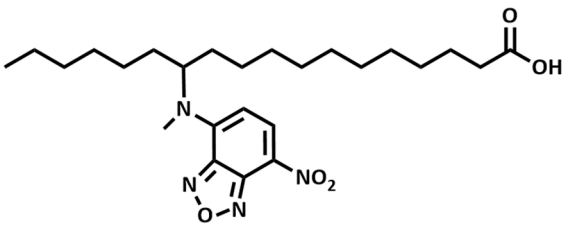


Figure 5-3. Structure of NBD-stearate. The reporter used in this work is an 18-carbon fatty acid modified with an extrinsic fluorescent group, NBD.

Bis-ANS, a hydrophobic fluorescent reporter, has been previously used to characterize T cell recognition of microbial lipids and lipoglycans (199,200). Bis-ANS was dissolved in DMSO and kept at a stock concentration of 40 mM. The structure of bis-ANS is shown in Figure 5-4.

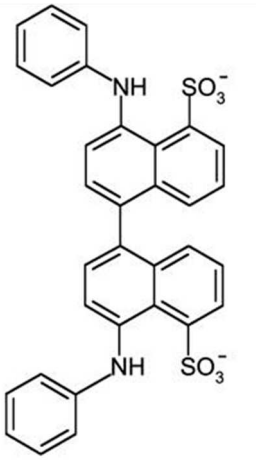


Figure 5-4 Structure of bis-ANS. This hydrophobic fluorescent reporter was previously used to characterize T cell recognition of microbial lipids and lipoglycans (199). For our purposes, it was used for competitive binding studies once ligands were unsuccessful in displacing NBD-stearate from apo-LppX.

The controls of the competitive assay were first established using FluoroMax®-4 Spectrofluorimeter. This included the background fluorescence emission from the protein, fluorescent reporter, and Tris buffer. Once this was established, SpectraMax® M2 microplate reader was then used to measure the fluorescent signal from the protein-reporter interactions. After several optimization experiments to determine the ideal conditions for detecting fluorescence, the excitation wavelength for NBD-stearate was set at 466 nm, while emission was measured over a range of 500 – 600 nm. The cutoff value, which eliminated any interference from the excitation spectra, was set at 515 nm. For plate reader experiments using bis-ANS as the reporter, bis-ANS was excited at 380 nm, and emission was detected from 420 – 700 nm. 455 nm was used as the cutoff value. In addition, 200 nM NBD-stearate and 5 μM bis-ANS was determined to be the optimal concentration for the LppX binding studies, as these concentrations produced significant fluorescence signal while also minimizing the formation of micelles.

All protein samples were prepared in 50 mM Tris pH 7.5, 1 mM DTT, 100 mM NaCl and were read by the plate reader using a 96-well Costar black plate. To first characterize the binding of NBD-stearate to LppX, the fluorescence intensity of concentrations up to 38 μM of apo-LppX binding to 200 nM NBD-stearate were measured. The binding curve of this interaction, subsequently referred to in this work as "K_{d1}," was then constructed by using fluorescence values at the maximum emission intensity, 530 nm. The dissociation constant was then calculated from the binding curve by the software program, Kaleidagraph.

Characterization of ligand-protein binding interactions by fluorescence spectroscopy

15 μM of apo-LppX was used for the competition assay to measure the competitive binding of three ligands. These ligands were MAC13243, a known inhibitor of LolA, a structural homologue of LppX found in *E. coli*; docosanoic acid, a twenty-carbon fatty acid co-crystallized in the hydrophobic cavity of LppX; and a PDIM analogue synthesized (synthesis detailed in Chapter 3). The structures of these ligands can be found in Figure 5-5.

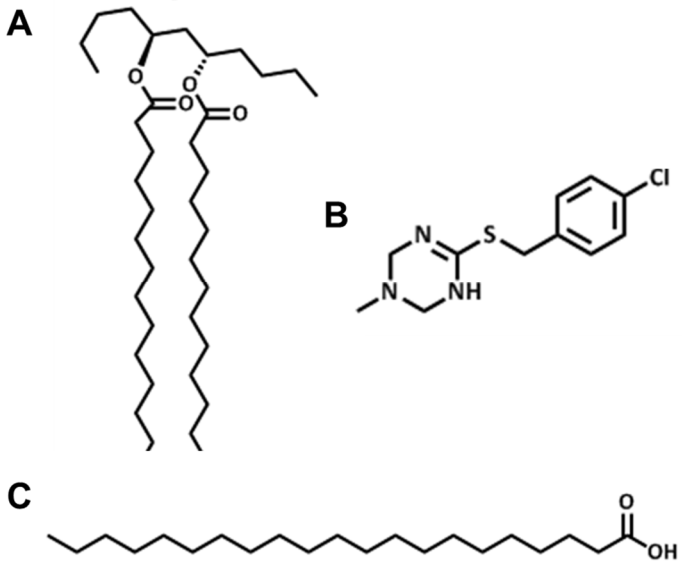


Figure 5-5. Structures of Competing Ligands. The following structures were used in the fluorescence assay to compete with NBD-stearate for the LppX binding pocket. (A) (5*R*,7*R*)-dipalmitoyl-undecane-5, 7-diol, a PDIM analogue, was synthesized. PDIM could not be used, as its extreme hydrophobicity renders it incompatible with biochemical assays. (B) MAC13243 (Maybridge, Inc.) is an inhibitor of LolA. LolA is a structural homologue of LppX. (C) Docosanoic acid is a 20-carbon saturated fatty acid that co-crystallized within the binding pocket of LppX.

Results and Discussion

Characterizing NBD-stearate and LppX interactions

The controls of this experiment were first established using FluoroMax®-4 Spectrofluorimeter. The maximum fluorescence emission from NBD-stearate was detected at 550 nm. Significant fluorescence was not observed from protein or buffer (Figure 5-6).

Upon the addition of protein to NBD-stearate, fluorescence emission increased 50-fold for LppX, and 140-fold for apo-LppX. A 30 nm blue shift in the emission maximum was also observed. This result was expected, as the single-chain fatty acid fluorescent reporter had increased opportunities to interact favorably with the hydrophobic pocket of the protein and escape the hydrophilic buffer environment. This allowed for an increase in detected fluorescence signal, as well as the corresponding blue shift in the emission maximum (Figure 5-6).

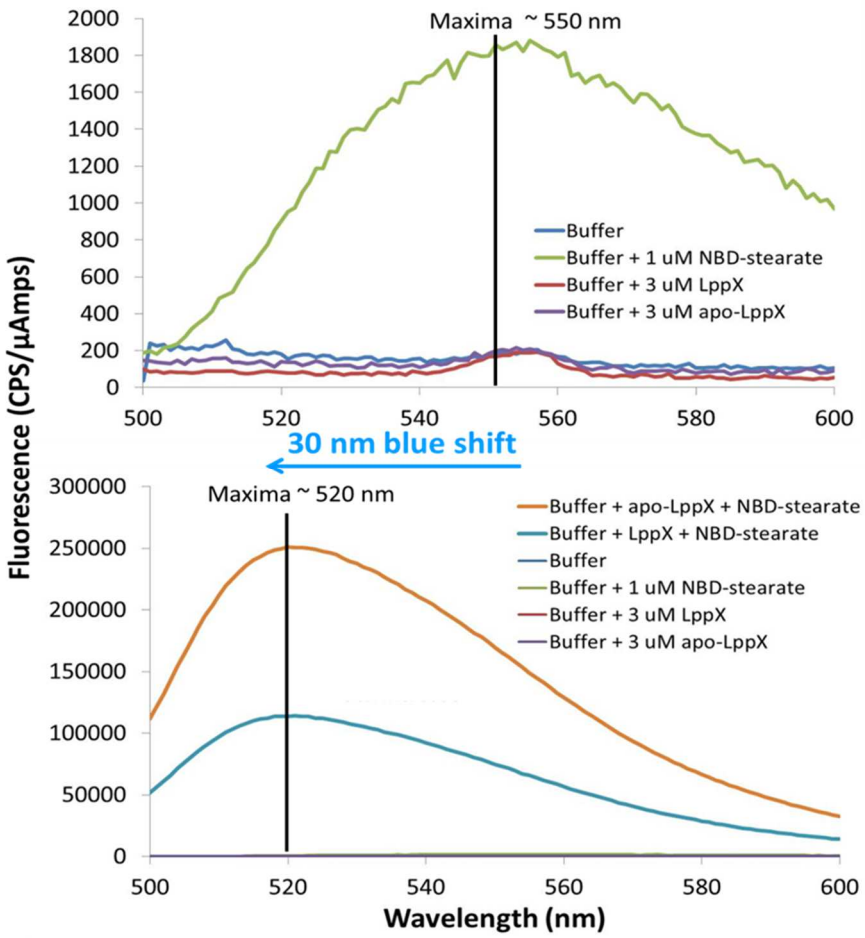


Figure 5-6. Controls for fluorescence spectroscopy measurements. (Top) The emission maximum from NBD-stearate was observed at ~550 nm. Significant background fluorescence was not observed from protein or buffer alone. (Bottom) Upon addition of apo-LppX and LppX, there was a 140-fold increase and 50-fold increase in fluorescence emission, respectively. In addition, a 30 nm blue shift in the emission maximum was observed. Signal from buffer, protein, and NBD-stearate alone were too low to be detected and thus, were determined not to interfere with fluorescence emission from protein-NBD-stearate interactions.

To construct the K_{d1} curve, fluorescence emission from NBD-stearate with increasing concentrations of LppX was detected using the SpectraMax® M2 microplate reader. A positive correlation was observed: as the concentration of LppX increased, the fluorescence emission also

increased. The maximum emission for all measured interactions occurred at 530 nm. This was seen for apo-LppX, as well as the mutant protein, apo-LppX A71W, as shown in Figure 5-7.

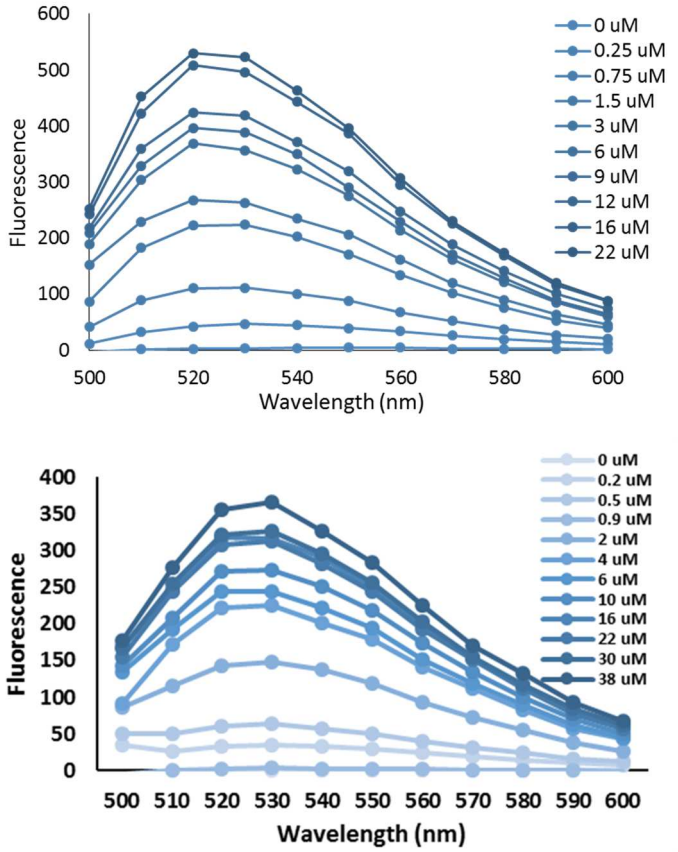


Figure 5-7. Emission spectra from protein-NBD-stearate interactions. Increasing concentrations of apo-LppX (Top) and apo-LppX A71W (Bottom) binding to 200 nM NBD-stearate correlated with increasing fluorescence signal. The maximum emission was observed at 530 nm.

The detected fluorescence signal at 530 nm for the varying concentrations of LppX were then used to construct the K_{d1} curve of LppX for NBD-stearate using Kaleidagraph. The K_{d1} curve for apo-LppX and the apo-LppX mutant A71W, are shown in Figure 5-8 as examples. The dissociation constant for this binding interaction was calculated to be 3.2 µM and 3.6 µM, respectively.

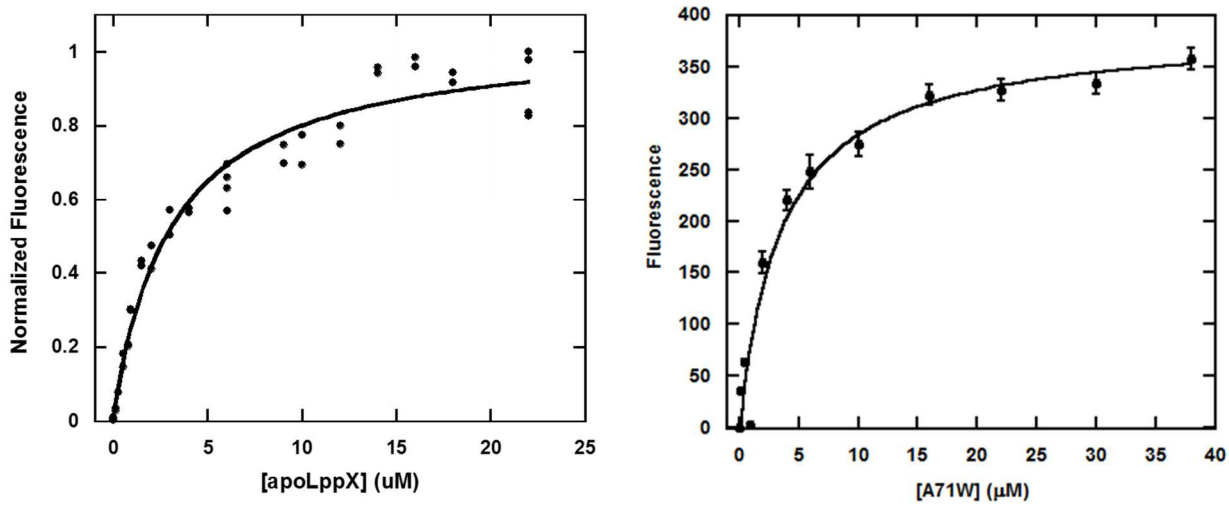


Figure 5-8. K_{d1} binding curve of apo-LppX and A71W to NBD-stearate. Using the maximum fluorescence emission spectrum values at 530 nm, the binding curve was plotted against the concentration of protein. The dissociation constant was calculated by Kaleidagraph, and was determined to be 3.2 μM for apo-LppX (left) and 3.6 μM for A71W (right). The K_{d1} was determined as the result of three separate trials.

The K_{d1} values were calculated for all LppX mutants in a similar manner. A full compendium of these measured binding interactions can be found in Table 5-2.

Table 5-2. K_{d1} Values of LppX, apo-LppX and apo-LppX mutants. K_{d1} for NBD-stearate and LppX proteins were calculated using Kaleidagraph and the fluorescence emission values at 530 nm. Each K_{d1} was calculated by three separate trials.

Protein	**Kd (μM)**
apo-LppX	3.2 +/- 0.4
LppX	2.3 +/- 0.2
A169W	1.4 +/- 0.2
A184W	2.9 +/- 0.9
A60W	1.8 +/- 0.6
A71W	3.6 +/- 0.4
A184W+A60W	0.8 +/- 0.1

The dissociation constant for apo-LppX and NBD-stearate was 3.2 μM, while the K_{d1} for LppX was calculated to be lower, at 2.3 μM. This may indicate the presence of lipids still bound to the hydrophobic cavity of LppX, which may favorably interact with the hydrophobic fluorescent reporter and demonstrate tighter binding. The "de-lipidation" protocol was thus necessary to effectively remove any bound fatty acids after purification of LppX from *E. coli*, as the presence of fatty acids still bound to the hydrophobic pocket were observed to distort the measured K_{d1} values.

The K_{d1} values of the LppX mutants did not demonstrate a clear trend of disrupted binding. In fact, apo-LppX mutants A169W and A60W had lower dissociation constants than both LppX and apo-LppX, while A71W and A184W had comparable K_{d1} values to apo-LppX (Table 5-2). Interestingly, the double apo-LppX mutant, which had both A60W and A184W mutated, had the lowest measured K_{d1} when compared to all tested mutants. These results may indicate that the mutations introduced were not enough to significantly block the binding of NBD-stearate. In addition, as a hydrophobic amino acid, tryptophan may have contributed to increasing the hydrophobic character of the binding pocket. This may have resulted in tighter binding, as observed with the measured K_{d1} of the double apo-Lppx mutant.

Competition with MAC13243, docosanoic acid, and a PDIM analogue

After characterizing the interaction between the NBD-stearate and apo-LppX, competition with three ligands was performed. The competition assay was first simulated on Microsoft Excel, where the K_{d2} values ranged from 0.001 μM to 100 μM. The results of this simulation can be found in Figure 5-9.

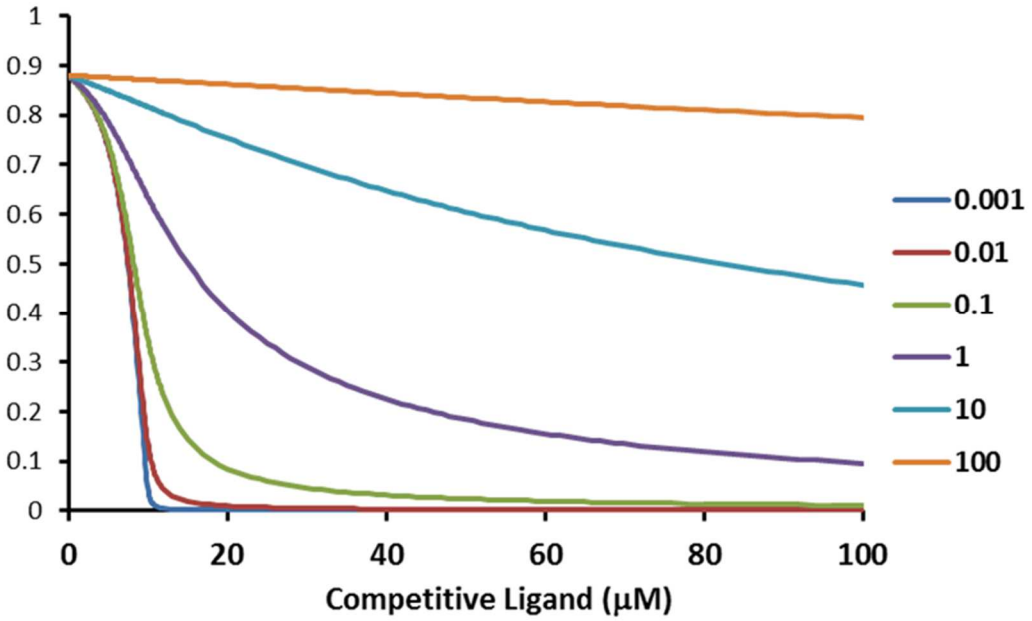


Figure 5-9. Simulation of Competition and K_{d2} binding curves. Using Microsft Excel, K_{d2} values ranging from .001 to 100 µM were modeled with 200 nM NBD-stearate and 15 µM apo-LppX set as the condition. The sharpest decrease in relative fluorescence signal was observed with the smallest K_{d2} value, 0.001 µM, while a very slight decrease in signal was observed with the largest K_{d2} value, 100 µM. The simulation was based on the exact analytical expression for complete competitive binding (201).

The results of this simulation were used to estimate the K_{d2} values of the competing ligands. Docosanoic acid is a 20-carbon fatty acid and was one of the fatty acids co-crystallized with LppX. It is thus known to bind to the hydrophobic pocket (49). In addition, its chain length is comparable to the 22-carbon mycocerosic acids that modify the phthiocerol moiety in PDIM. Therefore, docosanoic acid was used as the first ligand to compete with NBD-stearate. However, even with concentrations as high as 500 µM, no significant decrease in fluorescent signal was observed. A comparable result was observed with MAC13243, an inhibitor of LolA, a structural homolog of LppX from *E. coli* (Figure 5-10).

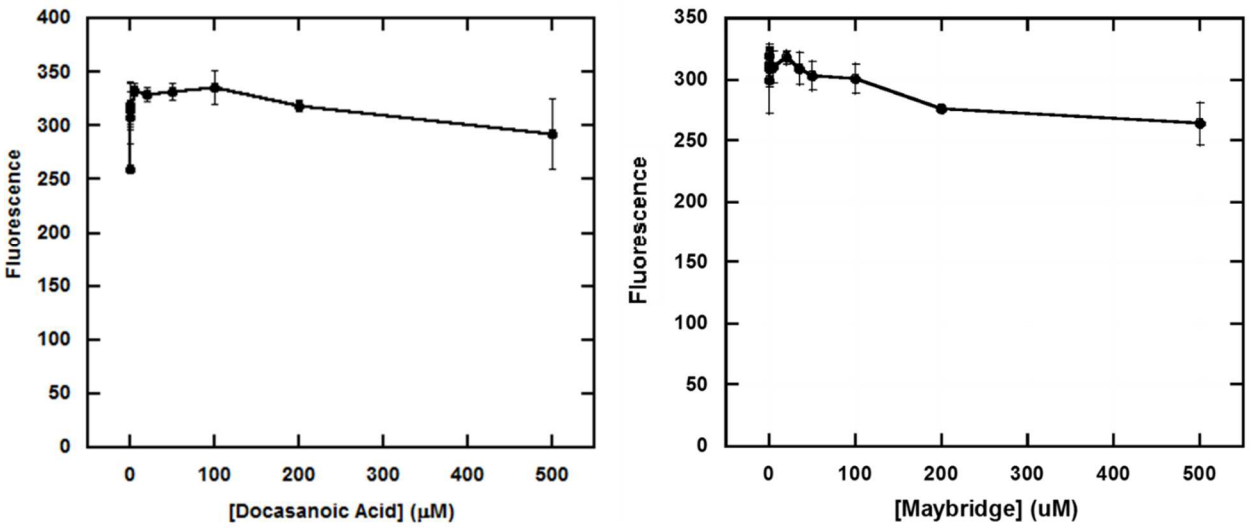


Figure 5-10. Competition with Docosanoic Acid and MAC13243. (Left) Increasing concentrations of docosanoic acid competed with 200 nM NBD-stearate. Even at 500 µM, no significant decrease in fluorescent signal was observed. The graph was constructed as a result of three separate trials. (Right) Increasing concentrations of MAC13243 (denoted here as "Maybridge") competed with 200 nM NBD-stearate. Even at 500 µM, no significant decrease in fluorescent signal was observed. The graph was constructed as a result of three separate trials.

Although a decrease in fluorescence was expected, an increase in fluorescence emission was observed while attempting competition with the synthesized PDIM analogue (Figure 5-11).

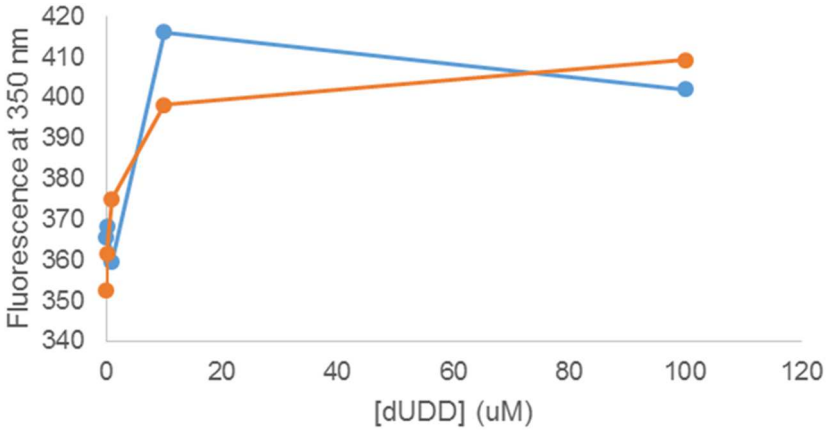


Figure 5-11. Competition with the synthesized PDIM analogue. Increasing concentrations of the synthesized PDIM analogue (denoted here as "dUDD" for (5*R*,7*R*)-dipalmitoyl-undecane-5, 7-diol) competed with 200 nM NBD-stearate. At concentrations up to 100 µM, a steady increase in fluorescent emission was observed. Two separate trials of this experiment are shown.

The results of these trials indicate that the competition assay with these ligands was not successful. The fluorescent reporter as a single fatty acid chain may not have been large enough to quantify the binding interaction with the large binding pocket of LppX. More than one reporter may have been binding to the same pocket. LppX was also observed to have a higher affinity for NBD-stearate than the competing ligands. Therefore, instead of using NBD-stearate, the fluorescent reporter bis-ANS was used.

Characterization of bis-ANS – apo-LppX interactions

In order to determine whether the competing ligands could displace a different fluorescent reporter, bis-ANS was used instead of NBD-stearate. When bis-ANS was excited in the absence of apo-LppX, the maximum emission was observed at 540 nm. Upon the addition of apo-LppX, a blue shift of the emission was observed, as the maximum emission was now at 470 nm. In addition,

a 180-fold increase in signal was observed (Figure 5-12). This result attested to the binding of the hydrophobic reporter to the hydrophobic pocket of apo-LppX.

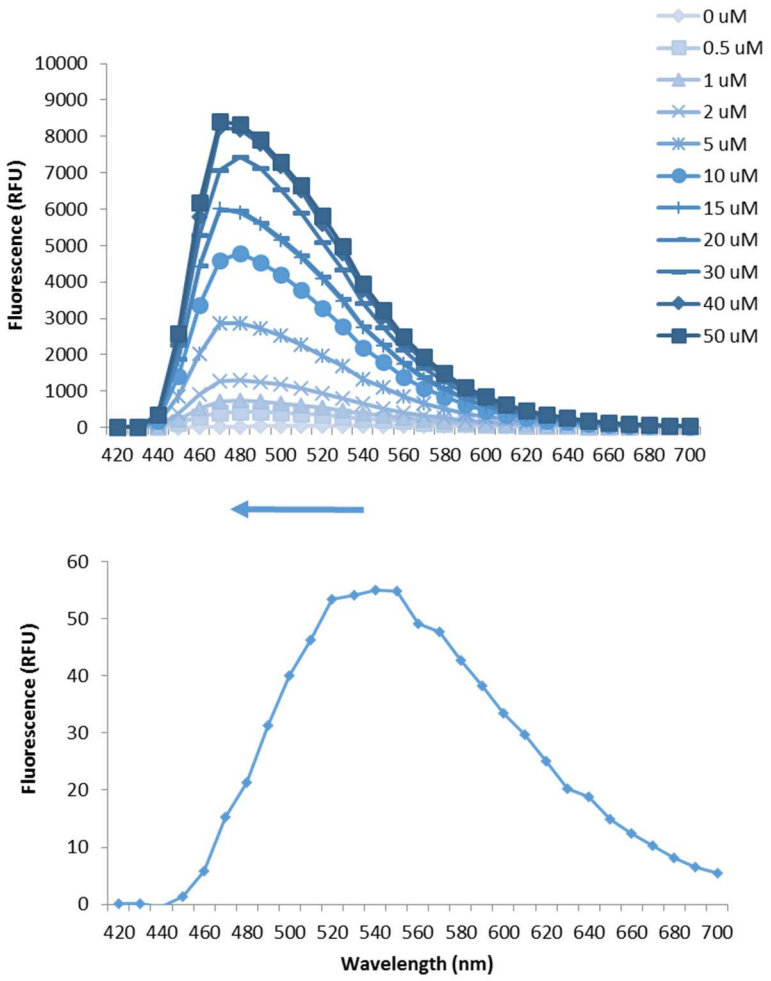


Figure 5-12. Controls for spectroscopy measurements with bis-ANS. (Bottom) The emission maximum from bis-ANS was observed at ~540 nm. Significant background fluorescence was not observed from the reporter alone. (Top) Upon addition of apo-LppX, there was a 180-fold increase in fluorescence emission. In addition, a 70 nm blue shift in the emission maximum was observed. This attests to the binding of bis-ANS to the binding pocket of LppX.

Using the maximum fluorescence values at 470 nm, the K_{d1} curve of apo-LppX for NBD-stearate using Kaleidagraph was constructed (Figure 5-13). The dissociation constant for this binding interaction was calculated to be 13.6 μM, which was higher than that of NBD-stearate to

apo-LppX. The higher dissociation constant denotes the weaker binding of this reporter for the binding pocket. This may be due to the bulkier size of this reporter as compared to the NBD-stearate.

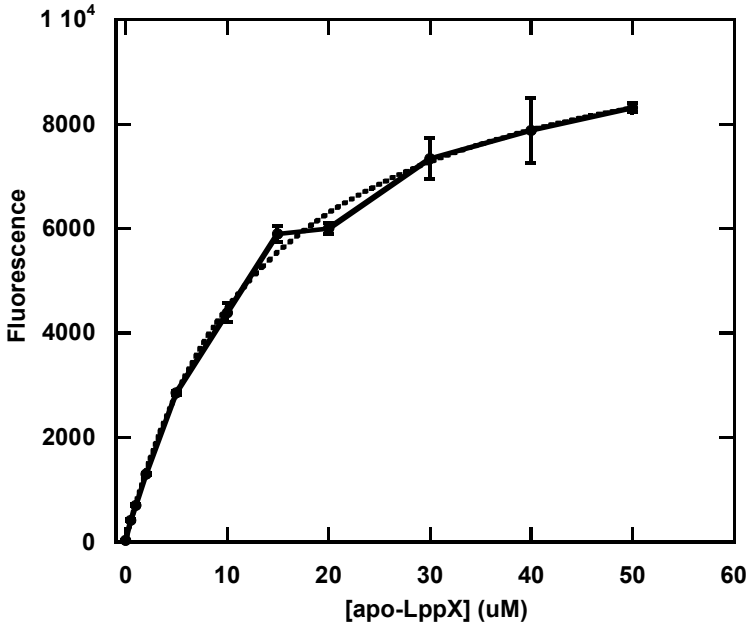


Figure 5-13. K_{d1} binding curve of bis-ANS to apo-LppX. Using the maximum fluorescence emission spectrum values at 470 nm, the binding curve was plotted against the concentration of protein. The dissociation constant was calculated by Kaleidagraph, and was determined to be 13.6 μM. This dissociation constant was higher than the measured K_{d1} of apo-LppX and NBD-stearate, perhaps indicating the bulkier size of the reporter. The K_{d1} was determined as the result of three separate trials.

Competition with docosanoic acid and MAC13243

After characterizing the binding of apo-LppX to bis-ANS, docosanoic acid was then used to attempt to compete out the reporter. A significant decrease in signal was not observed until concentrations reach 100 μM (right panel, Figure 5-14).

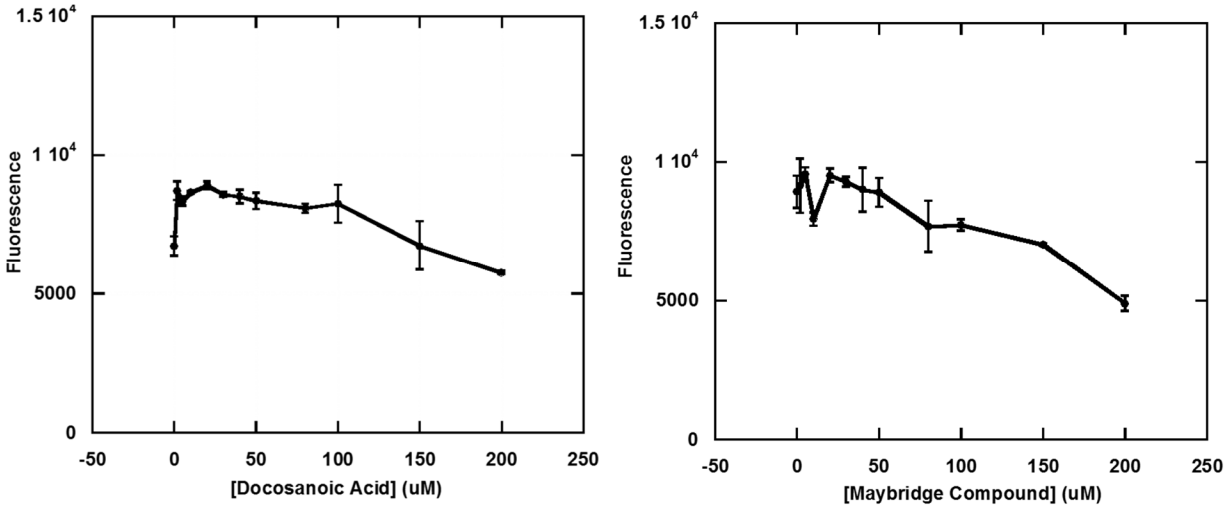


Figure 5-14. Competition with docosanoic acid and MAC13243. (*left*) At concentrations up to 100 μM, a significant and specific decrease in fluorescence due to the addition of docosanoic acid was not observed. (*right*) At concentrations up to 50 μM, a significant and specific decrease in fluorescence due to the addition of MAC13243 was not observed. Though a decrease in signal was observed at concentrations higher than 100 μM, this was later concluded to be the result of increasing amounts of the solvent, DMSO, in which the compound was stored, being added to the protein and possibly denaturing it.

A similar result was also observed with competition with MAC13243 (left panel, Figure 5-14). Though a decrease in signal at concentrations higher than 50 μM was observed, this was later concluded to be the result of increasing amounts of DMSO, the solvent in which the compound was stored, being added to the protein and possibly denaturing it. A successful competitive assay with bis-ANS as the reporter was thus not observed. Due to these results, the competitive assay with the synthesized PDIM analogue was not attempted.

Conclusions

The competitive assay described in this work was not successful in characterizing the interaction between LppX and PDIM. The hydrophobic nature of the tested ligands as well as NBD-stearate and bis-ANS may have led to the observed phenomenon of the fluorescent reporter interacting with micelles formed by the hydrophobic ligands, rather than successful binding to the LppX pocket. An increase in fluorescence emission from competition with the synthesized PDIM analogue appears to support this conclusion. In addition, as the protein is not in its native environment, the binding of the ligands may not have been accurately assessed by this assay. Quantification of the interaction of LppX and the tested ligands may be possible by other methods, such as surface plasmon resonance or isothermal titration calorimetry.

Although the competition assay was unsuccessful for characterizing the binding interactions of LppX, the assay may still be utilized for investigating other proteins with smaller binding pockets. Currently, this assay is being used in the lab to quantify the ligand binding for other lipid-binding proteins. For example, the periplasmic Mtb protein LprG is being studied for its interaction with its putative substrate, triacylglycerol (TAG). LprG is a structural homologue of LppX and has also been implicated in transporting other surface lipids from the inner to outer membrane in mycobacteria. We also have data supporting the ability of LprG to bind triacylated lipids, such as triacylglycerol, acylated phosphoinnositolmannoside, and lipoarabinomannan (50). The assay described in this work may thus be useful as a tool for characterizations of other putative lipid-binding proteins in mycobacteria.

Chapter 6 Understanding Lipid Transport Mechanisms in Mycobacteria through Identification of Protein-Protein Interactions

Attribution: LprG TAG mutants discussed in this chapter were generated by Mary L. Previti (Seeliger Lab, Stony Brook University); proteomic analyses were completed by Dwight W. Martin (Proteomics Facility, Stony Brook University); analysis of isotopically labeled MSn data was completed by Armand B. Cognetta III (Cravatt Lab, Scripps Research Institute, La Jolla, California).

Introduction

Review of LprG literature

While the biosynthesis of mycobacterial outer membrane (OM) lipids has been relatively well studied, comparatively little is known about their transport from the inner membrane (IM) to the OM. One clue about lipid transport mechanisms comes from the implication of MmpL7 and LppX in the proper localization of phthiocerol dimycocerosate (PDIM), a virulence-associated surface lipid (49,202). MmpL7 is essential for *Mtb* virulence and is an integral IM transporter proposed to transport PDIM across the IM (41,194,202). LppX is a cell wall-associated protein and was extensively described in Chapter 5. Briefly, an *Mtb lppX* transposon mutant is attenuated for virulence *in vivo* and lacks PDIM at the cell surface, suggesting a role for LppX in transporting PDIM to the OM (49). LppX is part of a structurally homologous family of lipoproteins that includes LprA, LprF and LprG. LprG is the only family member that is conserved across mycobacterial species including *Mycobacterium smegmatis* (*Msm*), a fast-growing, non-pathogenic strain, suggesting that LprG plays a fundamental role in mycobacterial physiology in addition to being required for *Mtb* virulence. LprG encoded by *rv1411c* in *Mtb str.* H37Rv and *MSMEG_3070* in *Mms* str. mc^2155 and is co-transcribed with the major facilitator superfamily

transporter encoded by *rv1410c* (*MSMEG_3069*) (203). We therefore propose a role for Rv1410c and LprG in lipid transport to the OM in mycobacteria that is analogous to that of MmpL7 and LppX in PDIM transport.

LprG has been described to bind triacylated glycolipids and to be necessary for proper display of lipoarabinomannan (LAM) at the *Mtb* surface (50,204). Carbohydrate staining of an SDS-PAGE gel and immunodetection of a Western blot containing purified *Msm* LprG showed co-purification of LAM, lipomannan (LM) and phosphatidylinositol mannosides (PIM). Mass spectrometric analysis of organic extracts of *Msm* LprG purified from *M. smegmatis* yielded masses corresponding to diacyl phosphatidylinositol, diacyl phosphatidylinositol mannoside (PIM_1), diacyl phosphatidylinositol dimannoside (PIM_2), and triacyl PIM_2 (Ac_1PIM_2) (38). LM and LAM share the Ac_1PIM_2 core, but have more extensive glycosylation. All of these lipids contain the phosphatidylinositol core, but differ in the number of mannose residues and acyl chains. Together these data indicated that *Msm* LprG can bind diacyl and triacyl glycolipids (38). The ability of *Mtb* LprG to bind LM and LAM has also been confirmed. Immunoblotting of protein purified from *Mtb* showed co-purification of PIM, LM and LAM. Surface plasmon resonance of several structural variants of PIM, LM and LAM revealed that these lipids bind *Mtb* LprG *in vitro* via the lipid acyl chains binding the hydrophobic pocket (50).

Multiple X-ray crystal structures of *Mtb* LprG have been solved, each one of a truncated form lacking the first 35 residues which contain the Sec secretion signal and the lipobox (signal for lipoprotein acylation of the conserved Cys). The solved structures of wild-type LprG; LprG bound to PIM; and LprG bound to triacylglyceride (TAG) show that LprG is a half beta-barrel with a hydrophobic ligand-binding pocket enclosed by α-helices that binds acyl chains of PIM and TAG (38,205). The structure of LprG(V91W), a ligand-binding pocket mutant, was also solved;

the mutation was shown to decrease the size of the entrance to the pocket and the overall volume of the ligand-binding pocket; and provided steric hindrance to limit ligand binding (38).

In unpublished work from our and our collaborators' labs, we showed that the an *Mtb* strain lacking the operon *rv1410c-lprG*, TAG accumulates inside the cell and that purified, truncated soluble *Msm* and *Mtb* LprG are capable of transporting TAG between lipid bilayer membrane vesicles (205). Because of the ability of LprG to bind PIM, LM and LAM, and its ability to bind and transfer TAG, we hypothesize LprG plays a role in lipid transport in mycobacteria. With no other clues for the mechanism by which this occurs in mycobacteria, we look to characterized transport systems in Gram-negative bacteria.

Review of Gram-negative cell wall transport systems

The cell envelope of Gram-negative bacteria consists of two membranes flanking a peptidoglycan-containing compartment called the periplasm. Both the IM and OM are lipid bilayers: the IM is symmetrical and made of glycerophospholipids while the OM is asymmetrical with the inner leaflet made of glycerophospholipids and the outer leaflet made of lipopolysaccharides (LPS). The OM functions as a protective barrier for the bacterium and the virulence-associated lipid LPS contributes to the impermeable nature of the OM (206). Both membranes also have proteins integral to and associated with them. Proteins integral to the IM span the membrane with hydrophobic α-helices and OM integral proteins have amphipathic β-strands that form a β-barrel structure (207). Lipoproteins are lipoylated, membrane-associated proteins that are found in either the OM or the periplasmic leaflet of the IM. Both protein and lipid components of the OM are synthesized in the cytosol or at the cytoplasmic leaflet of the IM, are

transported across the IM and periplasm and then assembled in the OM. The transport systems for both lipoproteins and LPS have been studied extensively in *E. coli* and recently reviewed (208,209). Although *E. coli* is only distantly related to Gram-positive mycobacteria, the problem of transport across the periplasm to the OM is structurally analogous. We therefore review below the literature on LPS and lipoprotein transport systems as potential models for OM lipid transport in mycobacteria.

LPS is composed of two to three components: lipid A, which anchors the lipid in the membrane; an oligosaccharide, which is referred to as the core; and a polysaccharide called O-antigen. O-antigen is not a component of LPS in all bacterial species and in such cases, it is then referred to as lipooligosaccharide (LOS). Lipid A and the core oligosaccharide are synthesized and coupled in the cytoplasm and subsequently flipped across the IM by the ABC transporter MsbA in the periplasm where the lipid A/core moiety is then coupled with O-antigen in the periplasm by the enzyme WaaL (210). After LPS is assembled and flipped to the periplasmic side of the IM, it is then shuttled across the periplasm and assembled in the OM by the LPS transport (Lpt) machinery (211). LptB, LptF, and LptG form a hetero-oligomeric complex in a 2:1:1 stoichiometry with ABC transporter activity in the IM. LptF and LptG form the transmembrane domain (TMD) and LptB the cytoplasmic nucleotide binding domain (NBD) (212-215). LptC is anchored to the IM via its N-terminal transmembrane (TM) helix and has a large, soluble, periplasmic domain that is capable of binding LPS (216,217). LptC has also been shown to form a complex with LptBFG (215). LptA is a periplasmic protein that also has the ability to bind the lipid A portion of LPS (212,218). LptD and LptE form a hetero-oligomeric complex in the OM where LptD is a β-barrel TM protein and LptE is a lipoprotein with its soluble domain facing the periplasm (219,220). LptE

has been shown to directly interact with LPS, suggesting a direct role in its transport to the outer leaflet of the OM (Figure 6-1, left) (221).

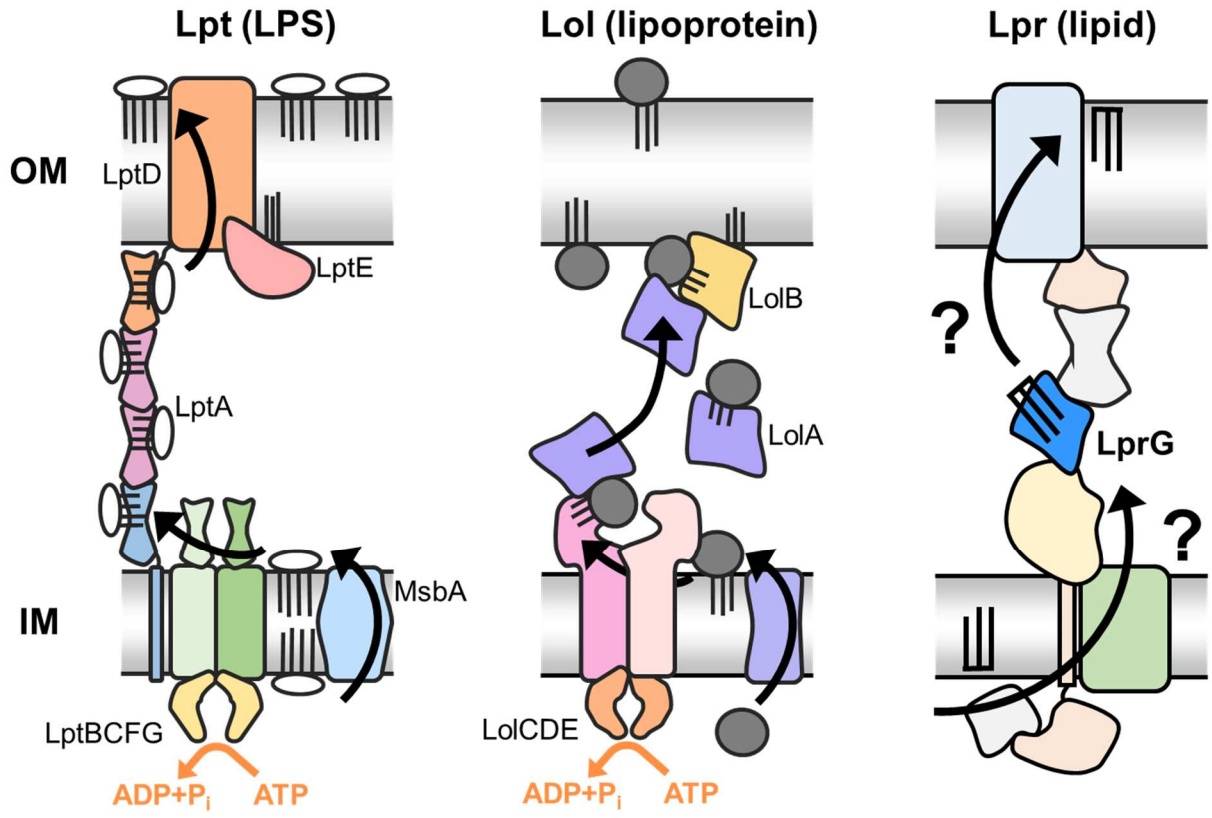


Figure 6-1. LprG-mediated lipid transport may have similarities to *E. coli* transport systems. Mycobacterial outer membrane (OM) lipid transport may have similarities to Gram-negative Lpt-mediated lipopolysaccharide (LPS) transport and/or Lol-mediated lipoprotein transport. (*left*) The LptBCFG inner membrane (IM) complex extracts LPS from the periplasmic side of the IM, transfers it to LptC and then LptA. LptA forms a channel from the IM complex to the OM complex, which is formed by LptDE. (*middle*) Lipoproteins are extracted from the IM by the LolCDE complex and transported to the OM by the chaperone LolA to the OM anchored LolB. (*right*) Putative LprG-mediated lipid transport system in mycobacteria may be similar to LPS and/or lipoprotein transport mechanisms in *E. coli*.

The structure of LtpA revealed a novel protein fold: a twisted β-jellyroll comprising of 16 consecutive anti-parallel β-strands (222). Although there is no sequence similarity to LptA, LptC was found to have a similar structure (216). The interior of the cavity formed by the β-jellyroll in both LptA and LptC contains hydrophobic sidechains and is predicted to be deep enough to harbor the acyl chains of lipid A (223). The periplasmic domains of LptD, LptF and LptG were predicted to have a similar β-jellyroll fold to LptA by sequence homology and the crystal structure of LptD confirmed this hypothesis for LptD (224,225). This conserved β-jellyroll fold in LptACDFG further suggested that these proteins work in the same pathway. Moreover, the transfer of LPS from the IM to LptC and the subsequent transfer from LptC to LptA require ATP hydrolysis, suggesting that ATP powers LPS transport through the entire transmembrane complex (Figure 6-1, left) (223).

Another Gram-negative cell wall transport system that has been well-characterized is lipoprotein transport. Before discussing lipoprotein transport, it is important to understand lipoprotein processing and the parallels between *E. coli* and mycobacteria. Proteins destined for the cell wall are synthesized as nascent peptide chains in the cytosol and encode an N-terminal secretion signal that determines whether they will be translocated across the IM via the Sec system or Tat pathway (226-229). Lipoproteins are a class of functionally heterogeneous secreted proteins that contain an *N*-acyl-diacylglyceryl lipid moiety on the N-terminal Cys residue. Bacterial lipoproteins play a variety of roles in the cell wall: they are key in membrane biogenesis, act as substrate transporters and are involved in drug efflux (230-232). They are synthesized in the cytosol as nascent polypeptides and contain a signal sequence that targets them for translocation across the IM, usually via the Sec system (233,234). Lipoproteins also have a consensus sequence known as a lipobox that is a signal for lipidation at a conserved Cys residue (235). In mycobacteria,

the lipobox sequence is (Leu/Val/Ile)-(Ala/Ser/Thr/Val/Ile)-(Gly/Ala/Ser)-C and differs from the lipobox sequence in *E. coli* Leu-(Ala/Ser)-(Gly/Ala)-Cys (233,234,236). Once the nascent polypeptide is translocated across the IM, the pre-prolipoprotein diacyl glyceryl transferase or Lgt adds a diacylglycerol moiety onto the thiol of the conserved Cys residue. The Lgt-modified prolipoprotein is now subject to the signal peptidase LspA, which cleaves the secretion signal and lipobox from the modified polypeptide, leaving the Cys of the lipobox as the new N-terminus. Finally, the apolipoprotein is modified by *N*-acyltransferase Lnt, which adds a third acyl chain to the N-terminal amino group of the modified Cys (235). Proper lipoprotein processing has been shown to be essential for *Mtb* virulence, highlighting the important of this modification in lipoprotein function and of lipoproteins in general during infection (237).

In *E. coli*, mature lipoproteins either remain at the periplasmic side of the IM or are transported to the OM by the localization of lipoproteins (Lol) system. Approximately 90% of lipoproteins are localized to the OM either at the periplasmic side or at the cell surface (238). Lipoprotein localization is determined by the residue adjacent to the N-terminal cysteine, often referred to as the +2 rule. If the +2 residue is Ser, the lipoprotein is destined for the OM, whereas if the +2 residue is Asp, the lipoprotein remains in the IM (239). The +3 residue can also play a role in membrane localization as when Asp is followed by His or Lys, the lipoprotein may end up in the OM (240). The Lol system comprises of five proteins, LolABCDE. LolC, LolD and LolE form the IM ABC transporter in a 1:2:1 stoichiometry. LolC and LolE are the IM membrane components each with four TM helices and a large periplasmic domain each. LolD is the NBD component of the ABC transporter (241,242). LolA is a soluble periplasmic chaperone that is capable of binding the lipid moiety of lipoproteins and has been implicated in transporting them from the IM to the OM (243,244). LolB is the OM component of the Lol transport system, is itself

a lipoprotein and is capable of binding the lipid moiety of lipoproteins passed from LolA (Figure 6-1, middle) (245).

Structural analysis of LolA and LolB revealed that they are similar although their amino acid sequences are not. They have 11 antiparallel β-strands with 3 α-helices that enclose a hydrophobic cavity (196). Intriguingly, the *Mtb* lipoproteins LppX and LprG are structural homologs of LolA and LolB, indicating that the hydrophobic cavity enables transport of acyl chains whether they are part of a lipid or lipoprotein (38,49). This similarity of functions between the *E.coli* Lol and Lpt transport systems suggests that the mechanism of transport in mycobacteria may similarly require multiple proteins to transport cargo from the inner to outer membrane. We thus hypothesize that LprG functions in collaboration with other, as yet unidentified, proteins to traffic lipids in mycobacteria.

LprG protein interactions identified by complementary crosslinking methods and tandem mass spectrometry

To interrogate the hypothesized mechanism for LprG-mediated lipid transport in mycobacteria, we identified and applied two complementary crosslinking methods in live cells as the approach best suited to identify physiological, possibly transient, interactions between LprG and other proteins that therefore may function in lipid transport. The first crosslinking method uses the small amine-reactive molecule, dithiobis(succinimidylpropionate) or DSP (Figure 6-2). DSP can covalently crosslink proximal primary amines via two NHS-ester reactive ends joined by an eight-carbon spacer arm with a disulfide bond. The NHS-esters react with primary amines at pH 7-9, releasing the N-hydroxy-succinimide leaving groups and forming covalent amide bonds with

proteins, which generally have multiple primary amines located in the amino acid side chain of lysine residues and at the N-terminus. DSP is membrane permeable and thus allows for covalent crosslinking of intracellular protein interactions. DSP has been used to identify protein interactions within the cell wall as well as in the cytosol, underscoring its utility in multiple cellular contexts (246-248).

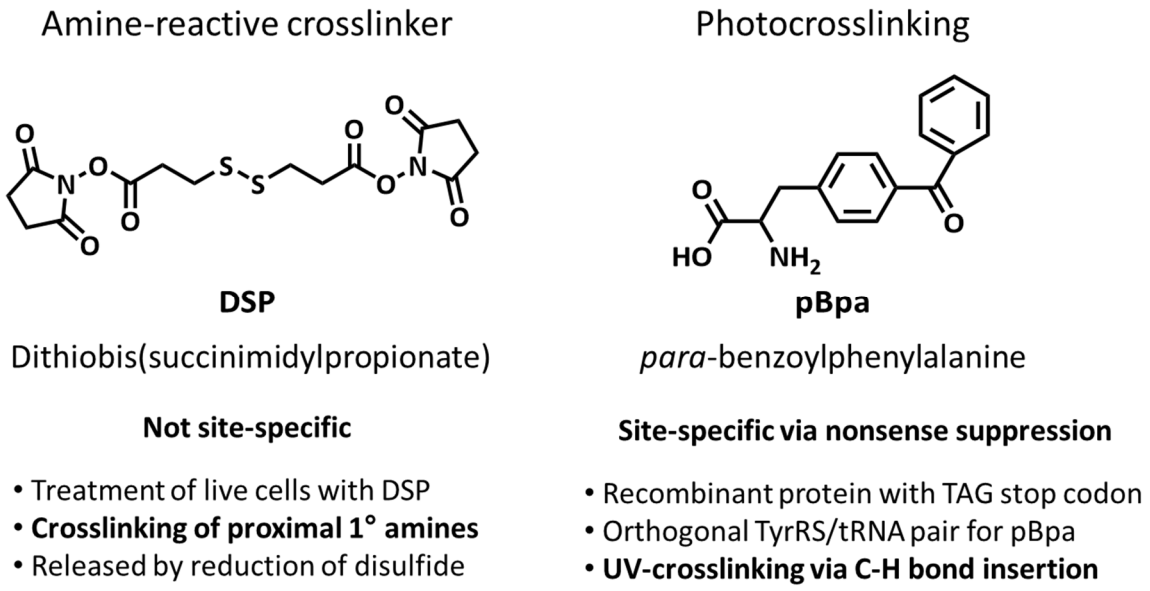


Figure 6-2. Complementary crosslinking methods to identify LprG interactors. Both methods can crosslink proteins in live cells and therefore capable of identifying physiological protein-protein interactions.

The second crosslinking method used in this study was site-specific incorporation of the unnatural photoreactive amino acid *para*-benzoylphenylalanine (pBpa) (Figure 6-2). Upon UV irradiation at 350-365 nm, pBpa forms a diradical and undergoes C-H bond insertion within nearby peptide backbones or amino acid side chains (249). Schultz and co-workers developed an

orthogonal tRNA/aminoacyl-tRNA synthetase (aaRS) pair incorporates unnatural amino acids site-specifically in proteins using nonsense or frameshift codons. They engineered the *Methanococcus jannaschii* (*Mj*) tyrosyl-RS to incorporate pBpa at the amber UAG nonsense codon in *E. coli* (250). The engineered *Mj*tRNA^{Tyr}/*Mj*TyrRS pair for use in *E. coli* was orthogonal in mycobacteria due to similar identity elements of the endogenous Tyr-tRNA^{Tyr} in eubacteria which are distinct from *Mj*Tyr-tRNA^{Tyr}, as previously described (251). This method can be used to generate unnatural amino acid-modified proteins directly in living cells, including mycobacteria.

Unnatural amino acid incorporation (UAAI) was used to elucidate the mechanism of lipopolysaccharide (LPS) transport from the inner to outer membrane in *E. coli*. Freinkman *et al.* generated multiple LptA mutants with pBpa incorporated throughout and were able to show that the N-terminus of LptA interacts with LptC at the inner membrane, the C-terminus of LptA interacts with LptD in the outer membrane of *E. coli* and that LptA is capable of forming homodimers *in vivo* in a head-to-tail fashion. These data imply that LptABCDEFG interact to form a transenvelope protein complex (Figure 6-1, left) (252). The UAAI method was also used to generate a model of lipoprotein transfer between Lol transporters. Okuda and Tokuda created multiple mutants of LolA and LolB such that pBpa would be incorporated throughout their three-dimensional structures. They were able to show that LolA and LolB interact with each other at the entrance to their hydrophobic ligand-binding cavities, or in a mouth-to-mouth manner. Using UAAI, they were also able to show that the mouth of LolA also interacts with LolC but not LolD or LolE (Figure 6-1, middle) (253). The success of UAAI in identifying protein-protein interactions associated with transport from the inner to outer membranes in *E. coli* encouraged us to try this tool in mycobacteria.

Experimental Procedures

Bacterial strains, growth media, vector construction and oligonucleotide primers

M. smegmatis str. mc^2155 (ATCC 700084) was grown in 7H9 (liquid), 7H11 (solid), Sauton defined medium (254) and Roisin minimal medium (255). 7H9 and 7H11 were supplemented with 0.2% casamino acids (BD), 0.5% glycerol, 0.05% glucose and 0.05% Tween80. For selective media, antibiotic concentrations were as follows: 25 µg/mL hygromycin and 12.5 µg/mL kanamycin for strains containing pMV361 and pSMT plasmids or 50 µg/mL hygromycin for strains only containing pSMT. Note: cells expressing both plasmids have a 6 hour doubling time. See Table 6-1 for vector constructs, oligonucleotide primer sequences and cloning strategies.

Attenuated Mtb strains mc^26220 and mc^26230 were grown in 7H9 (liquid), 7H11 (solid) supplemented with 10% Middlebrook OADC supplement (BD), 0.5% glycerol (Fisher), 0.2% casaminoacids, 24 µg/mL pantothenate (Sigma), 80 µg/mL lysine and 0.05% tyloxapol (liquid, Sigma) (256,257). For selective media, antibiotic concentrations used as denoted above.

Table 6-1. Table of vector constructs and cloning primers for LprG site-directed mutagenesis. Site-directed mutagenesis was done in either pET24b or pMV261 plasmid, then the mutated *lprG* was amplified using the forward primer 5'-GAATCACTTCGCACATATGCAGACGCGCCCAC-3' and reverse primer 5'-CCTTGTAGTCGCTAGCGGCCGCGGGCTTG-3' and subcloned into pSMT using Infusion (Clontech). Restriction sites used were NdeI and NheI.

Mutation	Description	Forward primer (5' to 3')	Reverse primer (5' to 3')	Template
Y132TAG	pSMT-pHsp60-MsmLprG-Y132TAG-3xFLAG-MjtRNA	GGCCAACATCTAGGACGTGTCGGCGATCCTG	CCGACACGTCCTAGATGTTGGCCGCGGCACC	pET24b-MsmLprG-TEV-6xHis
L70 TAG	pSMT-pHsp60-MsmLprG-L70TAG-3xFLAG-MjtRNA	GGTCGACGGGTAGCCCGTGGAGAAGCTCGAC	CTCCACGGGCTACCCGTCGACCTTGCCCTG	pET24b-MsmLprG-TEV-6xHis
F96TAG	pSMT-pHsp60-MsmLprG-F96TAG-3xFLAG-MjtRNA	CCTGATCGCGTAGGGCCAGAAGATCGCCGAC	CTTCTGGCCCTACGCGATCAGGTCGGCGGTG	pMV261-pHsp60-MsmLprG-TEV-6xHis-MscI+EcoRI
Q98TAG	pSMT-pHsp60-MsmLprG-Q98TAG-3xFLAG-MjtRNA	CGCGTTCGGCTAGAAGATCGCCGACGCGAAG	CGGCGATCTTCTAGCCGAACGCGATCAGGTC	pMV261-pHsp60-MsmLprG-TEV-6xHis-MscI+EcoRI
D102TAG	pSMT-pHsp60-MsmLprG-D102TAG-3xFLAG-MjtRNA	GAAGATCGCCTAGGCGAAGTTCGTGATCGCCG	CGAACTTCGCCTAGGCGATCTTCTGGCCGAAC	pMV261-pHsp60-MsmLprG-TEV-6xHis-MscI+EcoRI
L37TAG	pSMT-pHsp60-MsmLprG-L37TAG-3xFLAG-MjtRNA	CCGACGCACCGTAGCCCGACGGTGCCGCGCTG	CACCGTCGGGCTACGGTGCGTCGGAGGTCTCCG	Assembly
N130TAG	pSMT-pHsp60-MsmLprG-N130TAG-3xFLAG-MjtRNA	GTGCCGCGGCCTAGATCTACGACGTGTCGGCG	CGTCGTAGATCTAGGCCGCGGCACCGTAGTTC	pMV261-pHsp60-MsmLprG-TEV-6xHis-MscI+EcoRI
K187TAG	pSMT-pHsp60-MsmLprG-K187TAG-3xFLAG-MjtRNA	CACCGGCGCTCTAGGCCGACGGCCCGGTGCCC	GGCCGTCGGCCTAGAGCGCCGGTGCGATCTTG	Assembly

K181TAG	pSMT-pHsp60-MsmLprG-K181TAG-3xFLAG-MjtRNA	CGCGGTCAACTAGATCGCACCGGCGCTCAAG	CCGGTGCGATCTAGTTGACCGCGTCGGCGCTG	pMV261-pHsp60-MsmLprG-TEV-6xHis-MscI+EcoRI
A149TAG	pSMT-pHsp60-MsmLprG-A149TAG-3xFLAG-MjtRNA	CAACGTGCTGTAGAACTTCTCCGACGCCACG	CGGAGAAGTTCTACAGCACGTTGGCCAGGC	pET24b-MsmLprG-TEV-6xHis
A136TAG	pSMT-pHsp60-MsmLprG-A136TAG-3xFLAG-MjtRNA	CGACGTGTCGTAGATCCTGAACCCCGACACG	GGTTCAGGATCTACGACACGTCGTAGATGTTG	pMV261-pHsp60-MsmLprG-TEV-6xHis-MscI+EcoRI
E160TAG	pSMT-pHsp60-MsmLprG-E160TAG-3xFLAG-MjtRNA	CGACGGCCGCTAGTCGATCAACGGCACCGAG	CGTTGATCGACTAGCGGCCGTCGGCCGTGGC	pET24b-MsmLprG-TEV-6xHis
D109TAG	pSMT-pHsp60-MsmLprG-D109TAG-3xFLAG-MjtRNA	CGTGATCGCCTAGGGCAATCTCTACGCGGC	GAGATTGCCCTAGGCGATCACGAACTTCGC	pMV261-pHsp60-MsmLprG-TEV-6xHis-MscI+EcoRI
L60TAG	pSMT-pHsp60-MsmLprG-L60TAG-3xFLAG-MjtRNA	CGTGCACCTGTAGCTGACGGTGCAGGGCAAG	GCACCGTCAGCTACAGGTGCACGCTCTGCTGCG	pET24b-MsmLprG-TEV-6xHis
L93TAG	pSMT-pHsp60-MsmLprG-L93TAG-3xFLAG-MjtRNA	CACCGCCGACTAGATCGCGTTCGGCCAGAAG	CGAACGCGATCTAGTCGGCGGTGCCCTCCGC	pET24b-MsmLprG-TEV-6xHis
L75TAG	pSMT-pHsp60-MsmLprG-L75TAG-3xFLAG-MjtRNA	CGTGGAGAAGTAGGACGGCGACCTGACCAAC	GGTCGCCGTCCTACTTCTCCACGGGCAGCCCG	Traditional ligation
V218TAG	pSMT-pHsp60-MsmLprG-V218TAG-3xFLAG-MjtRNA	GGGCAACAGCTAGACGATGACGCTCTCGGAC	GCGTCATCGTCTAGCTGTTGCCCGGCGTGGGC	pET24b-MsmLprG-TEV-6xHis
K227TAG	pSMT-pHsp60-MsmLprG-K227TAG-3xFLAG-MjtRNA	GACTGGGGTTAGCAGGTCAACGTCACCAAG	CGTTGACCTGCTAACCCCAGTCCGAGAGCG	pET24b-MsmLprG-TEV-6xHis

T232TAG	pSMT-pHsp60-MsmLprG-T232TAG-3xFLAG-MjtRNA	GAATCACTTCGCACATATGCAGACGCGCCCAC	CCTTGTAGTCGCTAGCGGCCGCGGGCTTC	pET24b-MsmLprG-TEV-6xHis
A236TAG	pSMT-pHsp60-MsmLprG-A236TAG-3xFLAG-MjtRNA	GAATCACTTCGCACATATGCAGACGCGCCCAC	CCTTGTAGTCGCTAGCCTACGCGGGCTTGGTG	pET24b-MsmLprG-TEV-6xHis
Y113H	pSMT-pHsp60-MsmLprG-Y115H-3xFLAG-MjtRNA	CGGCAATCTCCACGCGGCGCTGACGCCCGGCG	CAGCGCCGCGTGGAGATTGCCGTCGGCGATCAC	Assembly
Y125H	pSMT-pHsp60-MsmLprG-Y127H-3xFLAG-MjtRNA	GCTGTCGAACCACGGTGCCGCGGCCAACATC	GCCGCGGCACCGTGGTTCGACAGCGGATCGCCG	Assembly
Y132H	pSMT-pHsp60-MsmLprG-Y132H-3xFLAG-MjtRNA	GCCAACATCCACGACGTGTCGGCGATCCTG	GACACGTCGTGGATGTTGGCCGCGGCACC	Assembly

Protein expression and DSP-crosslinking for proteomics

A 100-mL culture of *Msm* containing pMV261-pHsp60-*Msm*LprG-6xHis or pSMT-pHsp60-*Msm*Lprg-FLAG was grown for 9 hours at 37 °C with shaking at 225 rpm to OD_{600} ~1.0, pelleted, resuspended and used to inoculate 1 L of 7H9 at an OD_{600} ~0.05. This culture was incubated 15 hours at 37 °C at 225 rpm to an OD_{600} ~1.5. Cells were harvested at 5,000 x *g* for 25 min at 4 °C and split into two pellets. Cell pellets were washed 3 x 100 mL with PBS with 0.05% Tween80. Washed pellets were resuspended in 50 mL PBS with 0.05% Tween80 (1/10 of original culture volume). Dithiobis(succinimidyl propionate) or DSP (Thermo Scientific Pierce) was prepared fresh by solubilizing solid in DMSO dried over molecular sieves to make a 100 mM stock solution. Control cells were treated with an equivalent volume of dry DMSO. Cells were incubated with 2 mM DSP with gentle agitation for 30 min at 25 °C. To quench DSP crosslinking, 50 mM Tris pH 7.5 final concentration (from a 1 M stock) was added to the cell suspension and incubated another 15 min at 25 °C with gentle agitation. Cells were pelleted at 4,000 x *g* for 10 min and stored at -20 °C until further processing.

GFP fluorescence measurements

A 10-mL culture of *Msm* transformed with pSMT-pHsp60-MjtRNA-3xFLAG and pMV361-pHsp60-MjTyrRS was grown in 7H9 for 9 hours at 37 °C at 225 rpm to an OD_{600} ~0.2. A stock solution of pBpa was made by solubilizing the solid in 1 N NaOH at 500 mM. Once dissolved, four parts MQ water was added to yield 100 mM pBpa. Filter sterilized (0.2 μm pore size, Millipore) 100 mM pBpa was added to the cells to a final concentration of 2 mM pBpa and incubated at 37 °C with shaking at 225 rpm in the dark for another 15 hours. Control cells were treated with an equivalent volume of 1:4 1 N NaOH:MQ water. Cells were harvested at 4,000 rpm

and resuspended in PBS with 0.05% Tween80 to OD_{600} 1.0 and 150 µL were aliquoted in a 96-well plate (Corning). Fluorescence was measured using an emission filter at 485 nm and an excitation filter for 538 nm.

Protein expression and photocrosslinking

A 10-mL culture of *Msm* containing pSMT-pHsp60-*Mj*tRNA-FLAG and pMV361-pHsp60-*Mj*TyrRS was cultured in selective Sauton medium for 15 hours overnight at 37 °C at 225 rpm to an OD_{600} ~1.5. Cells were pelleted and resuspended in Sauton medium supplemented with 25 µg/mL hygromycin and used to inoculate 30 mL at an OD_{600} ~0.05. Cultures were incubated 9 hours at 37 °C with shaking at 225 rpm. Filter sterilized 100 mM pBpa was added to the cells at OD_{600} 0.2 to a final concentration of 2 mM pBpa and incubated at 37 °C with shaking at 225 rpm in the dark for another 15 hours. Control cells were treated with an equivalent volume of 1:5 1 N NaOH:MQ water. At OD_{600} 1.0 – 1.5, cells were harvested at 5,000 x *g* for 25 min and resuspended in 1:10 of the original culture volume in PBS. Cells for crosslinking were irradiated with 365 nm UV light for 30 min on ice with stirring every 5 min. Control cells were shielded from ambient light, not exposed to UV irradiation and kept on ice. Cell suspensions were pelleted at 4,000 x *g* for 10 min and stored at 20 °C until further processing.

Protein expression and photocrosslinking for proteomics

The protocol remained the same as described above in "Protein expression and photocrosslinking" section except culture volumes were 100 mL for the initial culture and 2 L for the subculture. Sauton medium was used for unlabeled LprG(Q98pBpa) expression and Roisin

medium was used for stable isotope labeling by amino acids in culture (SILAC) samples. Medium for SILAC contained either $^{14}NH_4Cl$ (used for UV-treated sample) or $^{15}NH_4Cl$ (used for control sample) as the sole nitrogen source.

Cell lysis, lysate fractionation and immunoprecipitation

Cell pellets were thawed on ice and resuspended 1:40 of the original culture volume in PBS. Protease inhibitor was added from a 100x stock [17 g/L phenylmethanesulfonylfluoride (PMSF, Sigma-Aldrich), 33 g/L benzamidine hydrochloride (Sigma-Aldrich), 0.137 g/L pepstatin A (MP Biomedicals), 0.03 g/L leupeptin (Roche), 0.2 g/L chymostatin (Sigma-Aldrich) in ethanol]. Cells were lysed in screw cap tubes containing 0.1 mm silica beads (Spectrum Chemical Manufacturing) at 5.0 m/s for 1 min on/off for 8 min total processing time using an MP Biomedical FastPrep-24 beat beater at 4 °C. Cell debris was pelleted by centrifuging at 5,000 x *g* for 10 min and the cleared lysate was transferred to 4 mL ultracentrifuge tubes (Fisher). The cytosol-enriched fraction was separated from the cell envelope (CE)-enriched fraction by ultracentrifugation at 100,000 x *g* for 1 hour. The CE-enriched fraction was solubilized in 1% N-dodecyl-β-D-maltoside (DDM, Anatrace) in PBS at 4 °C for 15 hours overnight with gentle agitation. The resuspended CE was subjected to ultracentrifugation at 100,000 x *g* for 1 hour to remove insoluble debris. The final pellet was resolubilized in 8 M urea (Sigma-Aldrich) to be analyzed by SDS-PAGE. The resulting supernatant or solubilized CE was transferred to a 15-mL conical tube containing 0.5 mL equilibrated in buffer A (300 mM NaCl, 1x protease inhibitor, 0.02% NaN_3, 1% DDM in PBS, pH 7.4) anti-FLAG M2 Affinity Gel (Sigma-Aldrich) resin and was incubated on a rotator at 25 °C for 2 hours. The resin was pelleted at 1,000 x *g* for 5 min and the flow-through was removed. Using multiple fractions of buffer A, the resin was washed in 60 column volumes (6 x 5 mL

washes) by pelleting the resin at 1,000 x *g* for 5 min between wash steps. The resin was resuspended in a final volume of 1 mL buffer A and transferred to a 2 mL tube. The 15-mL conical tube was washed to remove residual resin (3 x 0.5 mL) and the washes were added to the 2 mL tube containing the resin. The resin was pelleted at 1,000 x *g* for 5 min, supernatant removed, and 500 µL of 300 µg/mL FLAG peptide (Sigma-Aldrich) in buffer A was added to elute the enriched FLAG species from the resin. The elution slurry was incubated for 1 hour at 25 °C on a rotator before the resin was pelleted at 1,000 x *g* for 5 min and the eluant removed. A second 500 µL elution fraction was collected in the same manner, but without the 1 hour incubation.

Proteomic sample prep and analysis

All samples were analyzed by the Stony Brook University Proteomics Facility. For isotopically labeled samples, the control and photocrosslinked samples were mixed at this point. The protein was digested with trypsin as follows. The sample was reduced with 4 mM DTT for 30 min at 25 °C and alkylated in 8.4 mM iodoacetamide for 30 min at 25 °C in the dark. The solution was diluted to 1.7 M urea with 50 mM $(NH_4)HCO_3$ and 1 µg trypsin (in 25 mM $(NH_4)HCO_3$) (Trypsin Gold, Mass Spectrometry Grade, Promega, USA) was added. The digestion mixture was incubated at 37 °C for approximately 18 hours.

The digests were brought to 2% formic acid (FA), and the volume reduced by vacuum centrifugation to a final concentration of about 4 M urea. This solution was desalted with Supel-Tips C18 Micropipette Tips (Sigma-Aldrich) using FA-containing solutions with varied acetonitrile (ACN), essentially as described in vendor's bulletin. The eluted peptides were dried by vacuum centrifugation and dissolved in 2% ACN, 0.1% FA (buffer MS1) for analysis by

automated microcapillary liquid chromatography-tandem mass spectrometry. Fused-silica capillaries (100 μm inner diameter) were pulled using a P-2000 CO_2 laser puller (Sutter Instruments, Novato, CA) to a 5 μm inner diameter tip and packed with 10 cm of 5 μm ProntoSil 120-5-C18H (Bischoff Chromatography, Leonberg, Germany) using a pressure bomb.

This column was then placed in-line with a Dionex Ultimate 3000 equipped with an autosampler. The column was equilibrated in buffer MS1, and the peptide mixture was loaded onto the column using the autosampler at a flow rate of 2 μl/min. The HPLC pump flowed at 100 μL/min and the flow rate to the electrospray tip was reduced to ~ 200-300 nL/min by a split. The HPLC separation was provided by a gradient between Buffer MS1, Buffer MS2 (98% ACN, 0.1% FA) and Buffer MS3 (50% ACN, 0.1% FA). The HPLC gradient was held constant at 100% buffer MS1 for 5 min after peptide loading followed by an 80 min gradient from 100% buffer MS1 to 100% buffer MS3. Then, there was a 1 min transition as the gradient was switched to 51% buffer MS2 and this was followed by a 34 min gradient to 80% buffer MS2 and then held constant for 3 min. Finally, the gradient was changed from 80% buffer MS2 to 100% buffer MS1 over 2 min and then held constant at 100% buffer MS1 for 40 more min. The application of a 1.8 kV distal voltage electrosprayed the eluted peptides directly into a Thermo Fisher Scientific LTQ XL ion trap mass spectrometer equipped with a nanoLC electrospray ionization source (ThermoFinningan, San Jose, CA). Full mass spectra (MS) were recorded on the peptides over a 400-2000 m/z range, followed by five tandem mass (MS/MS) events sequentially generated in a data-dependent manner on the first, second, third, fourth and fifth most intense ions selected from the full MS spectrum (at 35% collision energy). Mass spectrometer scan functions and HPLC solvent gradients were controlled by the Xcalibur data system (ThermoFinnigan, San Jose, CA).

MS/MS spectra were extracted from the RAW file with ReAdW.exe (http://sourceforge.net/projects/sashimi). The resulting mzXML data files were searched using The GPM X!Tandem search engine (http://www.thegpm.org/tandem/) against the Uniprot *Mycobacterium smegmatis* (strain ATCC 700084 / mc^2155) proteome (dated 12/08/2014) with added sequences for common contaminants.

Immunoblot analysis

For reducing SDS-PAGE, 30 µg/well were loaded for whole cell lysate samples and 0.5 µg/well for enriched samples. Protein from the SDS_PAGE gel was transferred to nitrocellulose membranes (0.45 µm, Bio-Rad), blocked and probed with mouse α-FLAG (1:1,000; clone M2, Sigma Aldrich); rabbit α-6xHis HRP (1:10,000; Abcam); mouse α-GroEL (1:5,000; Abcam); and rabbit α-MspA (1:1,000; gift of Michael Niederweis, University of Alabama Birmingham) for 1 hour with agitation at room temperature. After washing, membranes were probed with secondary antibodies goat α-mouse IgG HRP (1:10,000; Abcam) or goat α-rabbit IgG HRP (1:10,000; Abcam). Immunoblots were visualized by chemiluminescence by probing washed membranes with Immobilon Western Chemiluminescent HRP Substrate (Millipore).

Results

Primary amine crosslinking by DSP to identify LprG-interacting proteins

To identify the LprG interactome, *Msm* containing pMV261-pHsp60-*Msm*LprG-6xHis was cultured and split into two cultures, one for DSP treatment and the other as the DMSO vehicle-treated control. Higher molecular-weight adducts observed by an anti-His tag immunoblot were

reduced following incubation with the reducing agent DTT, indicating that adducts are specific to the disulfide-containing DSP crosslinking agent. The truncated version of LprG lacks both the secretion signal and lipobox and is therefore not be secreted to the cell wall or modified with acyl chains. We hypothesized that unlike a truncated, cytosolic form, native full-length, cell wall-localized LprG, will have access to its physiological binding partners and form adducts. Adduct formation was indeed observed with full-length LprG, but not the truncated form, suggesting that the higher molecular-weight adducts observed for full-length LprG are physiological and not an artifact of overexpressing LprG in *Msm* (Figure 6-3).

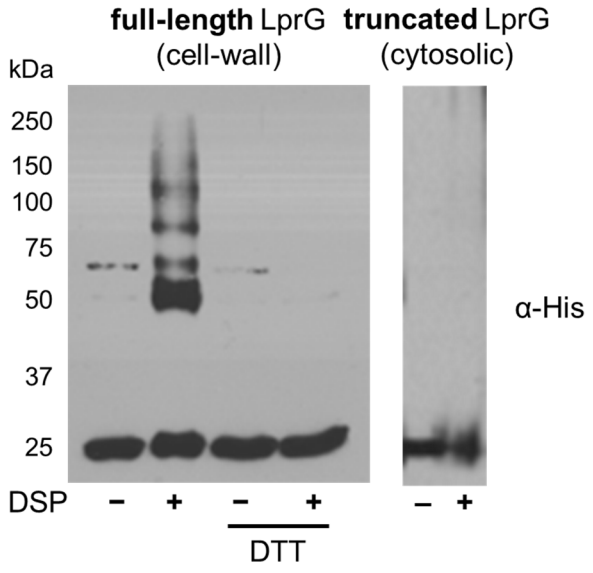


Figure 6-3. Amine-reactive crosslinker, DSP, generates LprG-protein adducts *in vivo*. *Msm* containing pMV261-pHsp60-*Msm*LprG-6xHis were incubated with DSP and the total cell lysates analyzed by immunoblot. Only full-length LprG yields higher molecular-weight adducts following DSP treatment. LprG-protein adducts can be reduced by the addition of DTT. Truncated LprG lacks a secretion signal, is not exported to the cell wall and is therefore be unable to access its native interactome.

In order to reliably identify the proteins crosslinked to LprG by DSP, an enrichment step was necessary. We initially pursued affinity purification with Ni-NTA resin and were able to detect enrichment by anti-His tag immunoblot (Figure 6-4) but were unable to obtain LprG and LprG adducts in sufficient yield and purity for downstream proteomic identification, as crosslinked and control samples were indistinguishable by total protein staining (data not shown). In contrast, anti-FLAG immunoprecipitation from cells expressing *Msm*LprG-FLAG yielded a higher degree of enrichment. Both the anti-FLAG immunoblot (Figure 6-4) and total protein stain (data not shown) indicated higher molecular weight adducts in the crosslinked sample and not in the control.

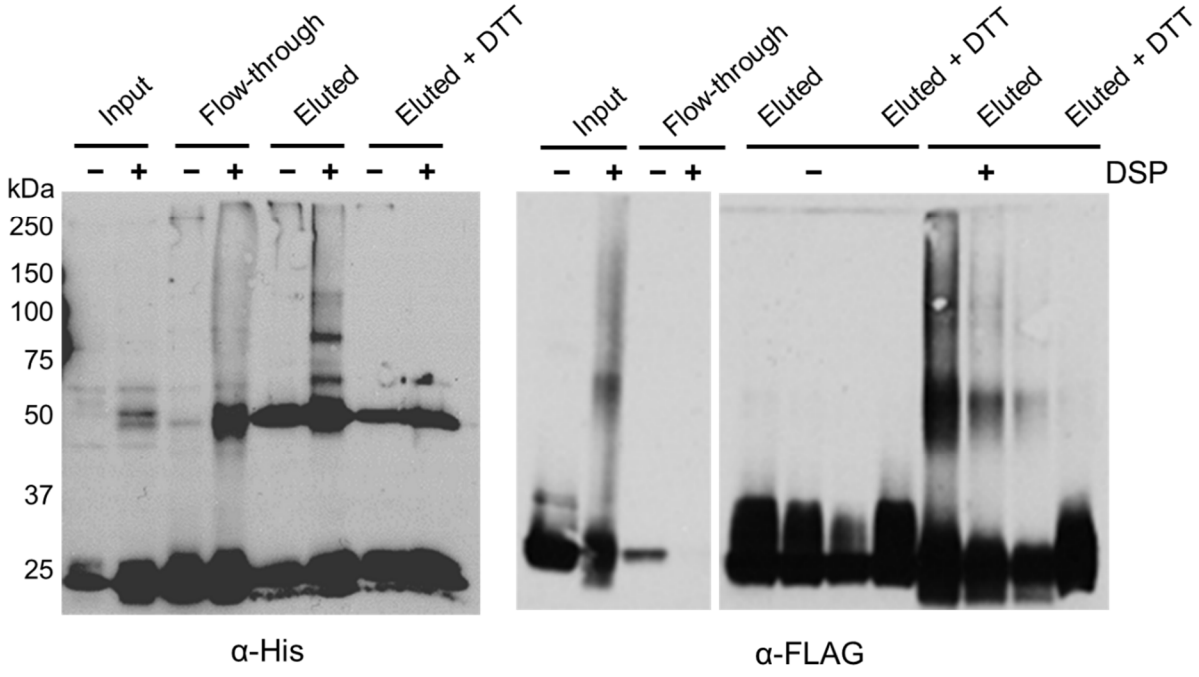


Figure 6-4. Affinity chromatography enriches for LprG-protein adducts. LprG and protein adducts formed after DSP treatment were enriched with either Ni-NTA resin (left panel) or anti-FLAG M2 conjugated beads (right two panels). In both cases, the enriched adducts were unconjugated by the addition of the reducing agent DTT.

To identify proteins crosslinked to LprG by DSP, we submitted two immunoprecipitated, trypsin-digested samples for LC-MS/MS analysis. One sample was the DMSO vehicle control and the second was DSP-treated. Hits from the GPM X!Tandem search engine output are listed in Table 6-2. Hits were defined by the following criteria: (1) detected in both technical replicates of the DSP-crosslinked samples, (2) not detected in either of the DMSO vehicle control technical replicates, (3) at least two quantified peptides detected per protein and (4) at least two spectral counts detected per protein. Spectral counts represent the number of spectra identified for a given protein and is used as a semiquantitative method for measuring protein abundance (258). Further examination revealed that the majority of hits have cell wall-associated functions. Several of the hits, including AccA3 and AccD4, are soluble proteins that are involved in fatty acid synthesis. Of the 19 hits with at least two detected peptides, 7 are confirmed or predicted transmembrane proteins and 15 are predicted to be localized to the cell wall. This includes secreted, IM-associated, periplasmic and OM-associated proteins. The predominance of cell wall- and membrane-associated proteins suggests that our crosslinking protocol and proteomics analysis reveal true LprG interactors, given that LprG is localized exclusively to the cell wall compartment (Bai and Seeliger, unpublished data).

Table 6-2. Table of top MS hits from DSP crosslinking of LprG. Proteins listed below were identified in both of two technical replicate LC-MS/MS runs of a FLAG IP of *Msm*LprG-FLAG. Only hits that (1) were detected only in crosslinked samples; (2) had > 2 peptides identified per protein; and (3) had > 2 spectral counts per protein are listed below. Hits are ranked from highest to lowest average spectral counts. Data for replicates 1 and 2 indicate the number of **peptides identified (spectral counts)**; NH, no homolog; blank spaces indicate lack of information.

Gene	Name	H37Rv Homolog	Description	Replicate 1	Replicate 2
MSMEG_0250	*mmpL3*	Rv0206c	Possible conserved transmembrane transport protein implicated in mycolic acid transport	**7 (14)**	**6 (13)**
MSMEG_1807	*accA3*	Rv3285	Probable bifunctional protein acetyl-/propionyl-coenzyme A carboxylase (alpha chain): biotin carboxylase + biotin carboxyl carrier protein	**7 (15)**	**5 (11)**
MSMEG_2410		Rv2969c	Possible conserved membrane or secreted protein	**6 (17)**	**5 (8)**
MSMEG_0317		Rv0227c	Probable conserved membrane protein	**4 (10)**	**3 (11)**
MSMEG_0835	*sodC*	Rv0432	Periplasmic superoxide dismutase [Cu-Zn]	**3 (11)**	**2 (9)**
MSMEG_4533	*subI*	Rv2400c	Probable sulfate-binding lipoprotein	**5 (11)**	**3 (4)**
MSMEG_3618	*apa*	Rv1860	Alanine and proline rich secreted protein (fibronectin attachment protein) (antigen MPT-32)	**4 (8)**	**3 (7)**
MSMEG_5070	*htrA*	Rv1223	Probable serine protease (DEGP protein)	**4 (6)**	**4 (8)**
MSMEG_4296	*caeA*	Rv2224c	Probable carboxylesterase, lipoprotein	**5 (10)**	**2 (3)**
MSMEG_4268	*ctaC*	Rv2200c	Probable transmembrane cytochrome C oxidase (subunit II)	**4 (8)**	**2 (4)**
MSMEG_0690		Rv0338c	Probable iron-sulfur-binding reductase	**4 (5)**	**3 (7)**
MSMEG_4263	*qcrB*	Rv2196	Probable ubiquinol-cytochrome C reductase (cytochrome B subunit)	**3 (5)**	**3 (7)**
MSMEG_6524		Rv0338c	ABC Polyamine/Opine/Phosphonate transporter, periplasmic ligand binding protein	**2 (6)**	**2 (6)**
MSMEG_6391	*accD4*	Rv2196	Probable propionyl-CoA carboxylase beta chain 4 (pccase) (propanoyl-CoA: carbon dioxide ligase)	**2 (5)**	**3 (6)**
MSMEG_5368		NH	Ectoine/hydroxyectoine ABC transporter solute-binding protein	**2 (4)**	**4 (6)**

MSMEG_2351	*fixA*	Rv3799c	Probable electron transfer flavoprotein (beta-subunit) (electron transfer flavoprotein small subunit)	**2 (5)**	**3 (5)**
MSMEG_2695		NH	Conserved 35 kDa alanine rich protein	**3 (5)**	**2 (4)**
MSMEG_6900	*ponA1*	Rv3029c	Probable bifunctional penicillin-binding protein 1A/1B (murein polymerase): penicillin-insensitive transglycosylase + penicillin-sensitive transpeptidase (DD-transpeptidase)	**3 (6)**	**2 (2)**
MSMEG_1642		Rv2744c	Probable conserved transmembrane ATP-binding protein ABC transporter	**2 (3)**	**2 (2)**

Site-specific incorporation of the photocrosslinking unnatural amino acid pBpa via nonsense suppression

Unnatural amino acid incorporation (UAAI) can be used to incorporate unnatural crosslinker amino acids site-specifically and therefore interrogate protein interactions at a specific location within a protein. To use this system in mycobacteria to identify LprG protein-protein interactions, we first verified its utility in mycobacteria as previously shown (251). The plasmid pSMT-*Mj*tRNA-pHsp60-GFP47TAG, (gift of Peter Schultz, Scripps Research Institute, La Jolla, California) encodes green fluorescent protein (GFP) mutant in which codon 47 was mutated to TAG. Only upon addition of the unnatural amino acid pBpa and expression of the engineered *Mj*TyrRS and *Mj*tRNA^{Tyr} was the observed fluorescence comparable to wild-type levels (Figure 6-5).

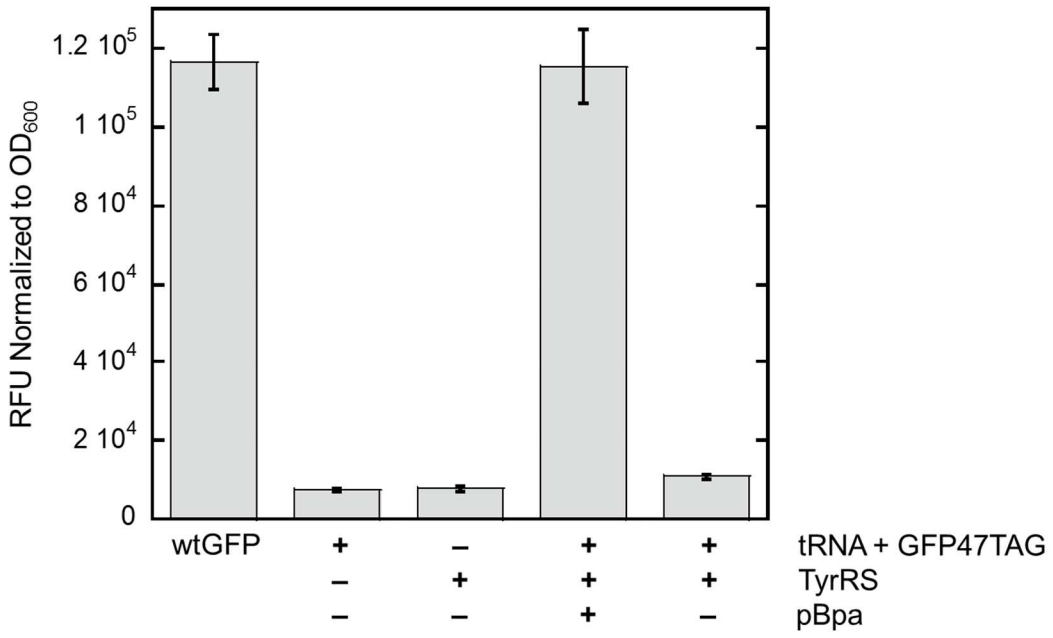


Figure 6-5. pBpa is incorporated at the TAG codon in *Msm* expressing GFP47TAG. Nonsense suppression was initiated in cells at OD_{600} 0.2 with the addition of 2 mM pBpa and incubated overnight. Fluorescence was normalized to OD_{600} for each culture after incubation. Only cells containing the synthetase, $tRNA^{Tyr}$ and GFP47TAG allele fluoresced above background levels upon the addition of pBpa. Fluorescence from pBpa-treated *Msm* containing *Mj*TyrRS and *Mj*$tRNA^{Tyr}$ + GFP47TAG is equivalent to *Msm* constitutively expressing WT-GFP. Data are the average of three biological replicates. Error bars represent ± S.D.

Because we anticipate future studies using UAAI in *Mtb*, we also attempted to validate the method in an attenuated *Mtb* strain mc^26220 but were unsuccessful. We did not detect fluorescence from bacteria treated with pBpa irrespective of treatment time, which ranged from a few hours to several days. In case the expression level of GFP was below the limit of fluorescence detection, we also performed an anti-GFP immunoblot, which, however, also indicated that full-length GFP was not expressed. We then attempted to transform the plasmids into a different attenuated *Mtb*

strain, mc^{2}6230, but were still unable to recapitulate the fluorescence observed in the *Msm* UAAI system (data not shown).

Nonetheless, since LprG is conserved across mycobacteria, we pursed further studies using UAAI on LprG in *Msm*. Sites for pBpa incorporation were chosen based on a predicted homology structure of *Msm*LprG generated by Protein Homology/Analogy Recognition Engine (Phyre) ver. 2.0 and based on the solved structure of *Mtb* LprG (PDB: 3MHA) (38,259). We chose residues all across the protein structure, including positions at the entrance to and within the ligand-binding pocket; on the α-helical surface (Figure 6-6, left); and on the β-sheet surface (Figure 6-6, right) to provide a survey of the entire protein surface.

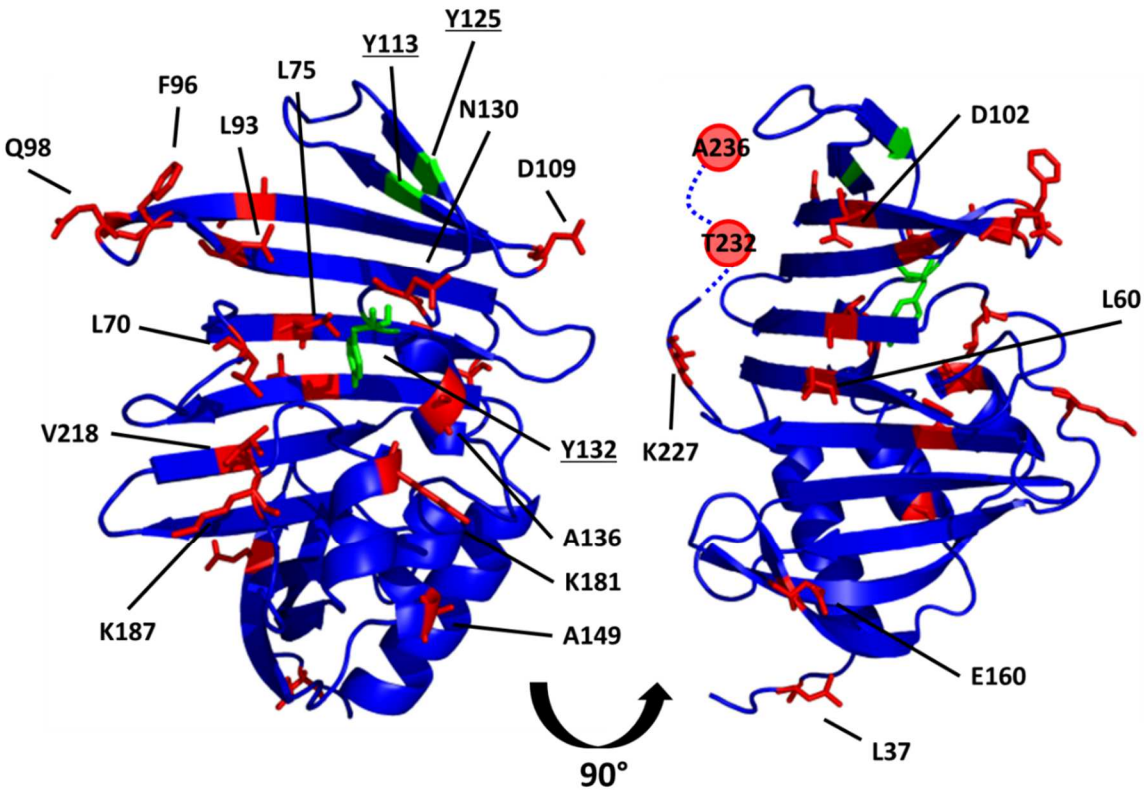


Figure 6-6. Homology model structure of *Msm*LprG idicating sites of photocrosslinker incorporation. A model of MsmLprG was generated by the Phyre2 server (259), using the MtbLprG structure (PDB: 3MHA) (38). All labeled, red rsidues were mutated to TAG for pBpa incorporation. The underlined residues were mutated to His (Figure 6-8). Both TAG and His mutations were made at Y132. Because the peptide encompassing T232 and A236 is unstructured, their locations are approximated by the red circles. (*left*) The view of the α-helical surface and (*right*) β-sheet surface of *Msm*LprG.

To validate UAAI for LprG, we verified full-length LprG expression via immunoblot against the C-terminal 3xFLAG tag. Only in the presence of pBpa was full-length LprG detected. Treatment with pBpa did not alter the expression level of WT-LprG. Sample loads were normalized to A_{280} of the whole cell lysate and showed that the D109, N130 and A136 LprG mutants were expressed at lower levels than WT-LprG (Figure 6-7).

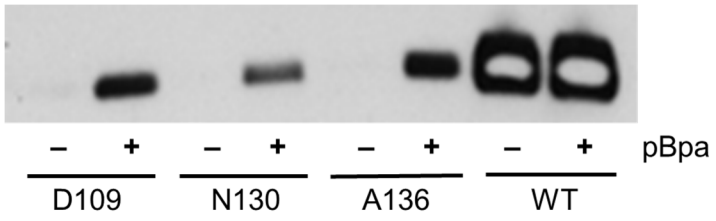


Figure 6-7. Full-length LprG is expressed only in the presence of pBpa. *Msm* containing MjTyrRS, MjtRNA^{Tyr} + *Msm*LprG-FLAG were treated at OD_{600} 0.2 with 2 mM pBpa and incubated for 16 hours. An equivalent amount of lysate as determined by A_{280} was analyzed by SDS-PAGE and anti-FLAG immunoblot.

Initial experiments with WT-LprG, pBpa treatment and crosslinking indicated that WT-LprG shows photo-induced higher molecular-weight adducts indicative of crosslinking (Figure 6-8, left). We hypothesized that at this level of LprG expression, sufficient background incorporation of pBpa by the *Mj*tRNA^{Tyr}/*Mj*TyrRS pair occurs at positions encoding Tyr to yield detectable photocrosslinking to WT-LprG. To test this hypothesis, we mutated the three endogenous Tyr to His, a conservative mutation that maintains the polar, aromatic nature of Tyr. All three endogenous Tyr are located at the entrance of the ligand-binding pocket and correspondingly (and in support of our hypothesis), the crosslinking pattern resembles that of TAG mutants at the entrance to the ligand-binding pocket (e.g., Q98; see Figure 6-10). Upon mutation of Tyr to His, single Tyr to His mutants yielded similar crosslinking patterns as those for WT-LprG, indicating that background crosslinking is likely attributed pBpa incorporation at more than one position. The double mutant Y113/125H (Y2xH) showed significantly less crosslinking, and the triple mutant Y113/125/132H (Y3xH) yielded no crosslinking (Figure 6-8).

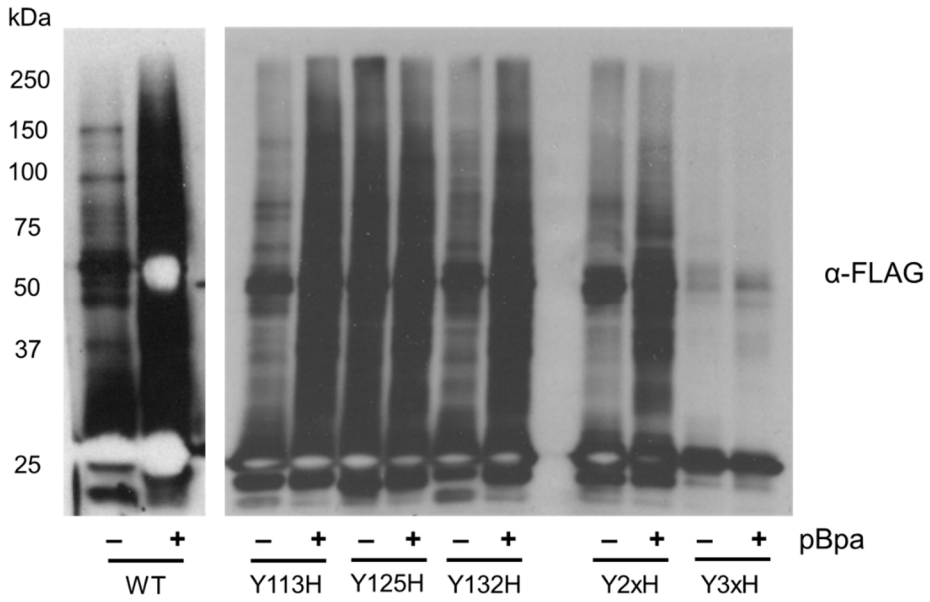


Figure 6-8. WT-LprG shows significant crosslinking that may result from background incorporation of pBpa. Samples were UV-irradiated for 10 min and total lysates analyzed by SDS-PAGE and anti-FLAG immunoblot. Sample loads were normalized to A_{280}. A series of Tyr to His mutants were generated as follows: Y2xH, Y113/125H; Y3xH, Y113,125,132H. (*left*) WT-LprG shows significant crosslinking. (*right*) All three native Tyr are located near the mouth of the ligand-binding pocket (Figure 6-6) and their crosslinking patterns resemble those of mutants in the same vicinity (Figure 6-10). LprG is a 24 kDa protein. LprG and adducts were detected by anti-FLAG immunoblot

In order to optimize UV irradiation time for maximal crosslinking yield, cells expressing *Msm*LprG(Y3xH) and *Msm*LprG(Q98pBpa) were subjected to 0, 5, 10, 20 and 30 min of UV-irradiation at 4 °C. For time points beyond 10 min, detected bands intensify, but additional adducts are not detected, suggesting that UV exposure up to 30 min does not introduce non-specific crosslinks (Figure 6-9).

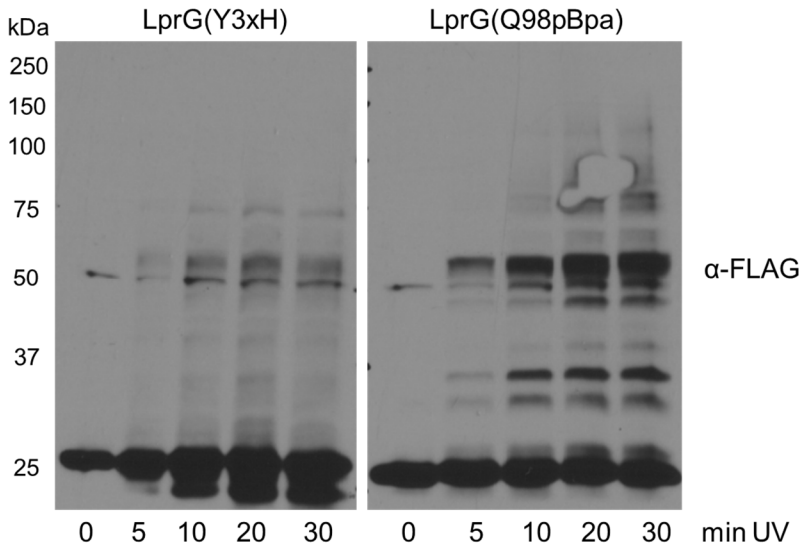


Figure 6-9. UV time course to determine optimal crosslinking time for UAAI studies. LprG(Y3xH) was used as a control (Figure 6-8). LprG(Q98pBpa) was used as a representative TAG mutant as it generates a high yield of crosslinked adducts (Figure 6-10). Beyond 10 min, no additional bands are detected, only an intensifying of bands, indicating that non-specific crosslinking does not occur within 30 min of UV irradiation. LprG is 24 kDa. LprG and adducts were detected by anti-FLAG immunoblot

Using the Y3xH WT control and optimized crosslinking protocol, all of the TAG mutants were analyzed qualitatively by comparing UV-treated samples with and without pBpa treatment (Figure 6-10). Mutations in the ligand-binding pocket did not yield significant protein crosslinking as indicated by the absence of higher molecular-weight (MW) bands. Mutations at the entrance to the ligand-binding pocket, however, yielded many higher MW bands. Mutations on the α-helical and β-sheet surfaces of LprG did not yield as many LprG adducts. Of these nine mutants, only L37, K181 and K187 have higher MW adducts with only a few distinct bands each. Mutants on the unstructured C-terminal tail of LprG have several distinct higher MW bands.

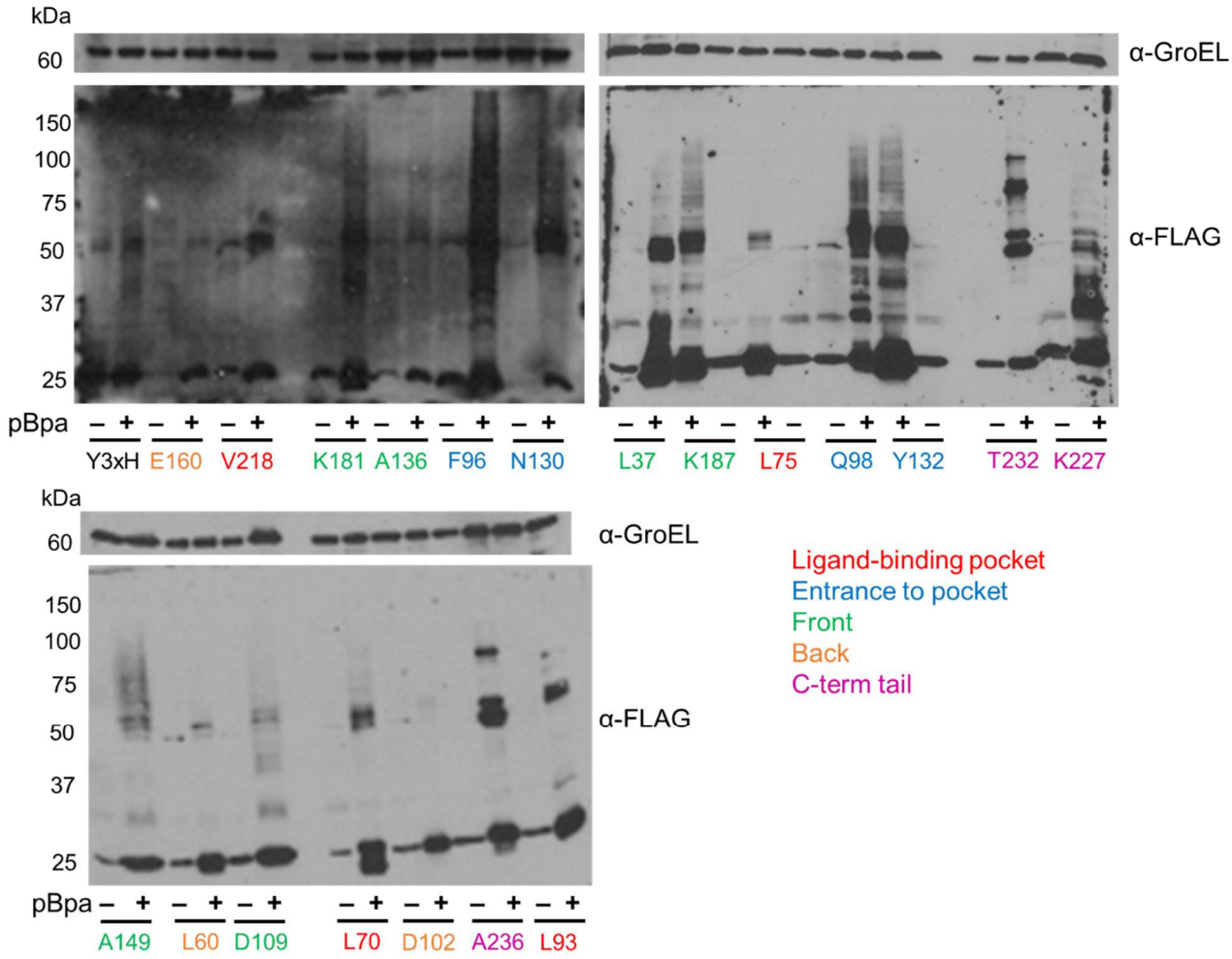


Figure 6-10. Live-cell photocrosslinking to detect LprG protein-protein interactions. Distinct banding patterns arise from mutations in various locations throughout LprG. Mutants are categorized by location (shown in different colors) throughout the structure of LprG. Using these data, interaction interfaces can be mapped to the entrance of the ligand-binding pocket and the C-terminal tail. LprG and adducts were detected by anti-FLAG immunoblot; GroEL was probed as a protein loading control.

*Msm*LprG(Q98pBpa) was chosen for further proteomic analysis of the crosslinked proteins because it is located at the entrance to the ligand-binding pocket and yielded many higher MW adducts. We hypothesized that this mutant would crosslink with proteins involved in lipid transport as lipids bind LprG at this interface. To identify proteins interacting with LprG, we submitted two

immunoprecipitated, trypsin-digested samples for LC-MS/MS analysis. Both samples contained pBpa-incorporated LprG but only one was UV-irradiated for 30 min (the other sample, the non UV-treated served as the control). Top hits from GPM X!Tandem search engine output are listed in Table 6-3, which summarizes the proteomic analysis with proteins listed from highest to lowest spectral counts. Hits were defined by the following criteria: (1) detected in only the photocrosslinked sample, (2) not detected in the non UV-treated control sample, (3) at least two quantified peptides per protein, and (4) at least two spectral counts per protein. Of the 23 proteins that made this cutoff, 7 are confirmed or putative integral membrane proteins and 16 have predicted cell wall-associated functions.

Table 6-3. Top MS hits from FLAG IP of LprG(Q98pBpa)-FLAG UV-crosslinked adducts. Proteins listed below were identified in one LC-MS/MS run of a *Msm*LprG(Q98pBpa)-FLAG FLAG IP. Only proteins that (1) were detected in the UV-treated, crosslinked sample; (2) were not detected in the non UV-treated control sample; (3) had > 2 peptides identified per protein; and (4) had > 2 spectral counts per protein are listed below. Hits are ranked by highest to lowest spectral counts. NH, no homolog; blank spaces indicate lack of information; *these proteins are highly similar with respect to amino acid sequence and may be not distinguishable by peptides identified (see Discussion section for further clarification).

Gene	Name	H37Rv Homolog	Description	# of peptides (spectral counts)
MSMEG_6398	*ag85A*	Rv3804c	Antigen 85-A (Diacylglycerol acyltransferase/mycolyltransferase) (Putative esterase)	**5 (16)**
MSMEG_1469	*rpsH*	Rv0718	30S ribosomal protein S8	**3 (12)**
MSMEG_4561		NH	Putative ABC Fe^{3+}-siderophores transporter, periplasmic binding protein (ABC transporter substrate-binding protein) (Putative solute-binding protein)	**3 (10)**
MSMEG_0020		NH	ABC transporter substrate-binding protein (Periplasmic binding protein)	**3 (8)**
MSMEG_2961	*secD*	Rv2587c	Protein translocase subunit	**3 (8)**
*MSMEG_0520 *MSMEG_0965 *MSMEG_5483	*mspB* *mspA* *mspC*	NH	Porin	**3 (7)**
MSMEG_2628	*infB*	Rv2839c	Translation initiation factor IF-2	**3 (6)**
MSMEG_4271		Rv2203	Uncharacterized protein, possible conserved membrane protein	**2 (6)**
MSMEG_0919	*hbhA*	Rv0475	HBHA-like protein (Heparin-binding hemagglutinin) (Uncharacterized protein)	**2 (6)**
MSMEG_4999	*oppA*	Rv1280c	Bacterial extracellular solute-binding protein, family protein 5 (Uncharacterized protein)	**4 (5)**
MSMEG_6900	*ponA1*	Rv0050	Penicillin-binding protein 1A	**4 (5)**
MSMEG_3027		Rv2553c	Aminodeoxychorismate lyase (Secreted solute-binding protein,aminodeoxychorismate lyase-like protein) (Uncharacterized protein)	**3 (4)**

MSMEG_5017		NH	Putative lipoprotein	**2 (4)**
MSMEG_3851	*lppI*	Rv2046	Putative lipoprotein	**2 (4)**
MSMEG_6524		NH	ABC Polyamine/Opine/Phosphonate transporter, periplasmic ligand binding protein (ABC transporter substrate-binding protein) (Putative solute binding protein)	**2 (4)**
MSMEG_4328	*kasB*	Rv2246	3-oxoacyl-(Acyl-carrier-protein) synthase 1	**2 (4)**
MSMEG_4263	*qcrB*	Rv2196	Menaquinol-cytochrome C reductase (Ubiquinol-cytochrome C reductase) (b subunit)	**2 (4)**
MSMEG_4957	*thrA*	Rv1294	Homoserine dehydrogenase	**2 (4)**
MSMEG_6941		Rv3920c	R3H domain-containing protein (Single-stranded DNA-binding protein) (Single-stranded nucleic acid binding R3H)	**2 (3)**
MSMEG_6220		Rv3693	Lipoprotein (Putative conserved membrane protein)	**2 (2)**
MSMEG_3796	*lysX*	Rv1640c	Lysylphosphatidylglycerol biosynthesis bifunctional protein [Includes: Lysine--tRNA ligase (Lysyl-tRNA synthetase); Phosphatidylglycerol lysyltransferase (Lysylphosphatidylglycerol synthetase)	**2 (2)**
MSMEG_0736		Rv0383c	Putative conserved secreted protein (Uncharacterized protein)	**2 (2)**
MSMEG_6393	*fadD32*	Rv3801c	Acyl-CoA synthase (Fatty acid-CoA ligase)	**2 (2)**

SILAC combined with proteomics analysis affords greater accuracy in identifying hits by enabling comparison of experiment vs. control within a single LC-MS/MS run and analysis. This approach increases both accuracy and sensitivity by removing instrument run-to-run variation and allowing quantitative analysis of spectral counts. Table 6-4 summarizes the isotopically labeled proteomics analysis. To identify proteins that interact with LprG, we analyzed an

immunoprecipitated, trypsin-digested sample by LC-MS/MS. The sample was a 1:1 mixture of ^{14}N-labeled, pBpa-incorporated LprG that was UV-irradiated for 30 min and ^{15}N-labeled, pBpa-incorporated LprG (which served as the control). The proteins listed in Table 6-4 were enriched in the ^{14}N-photocrosslinked sample relative to the ^{15}N-control by a ratio of at least two and had at least two peptides identified per protein. Data are shown for both technical replicates. At first glance there is little overlap between hits identified in the unlabeled and the isotopically labeled experiments, as there are only 3 proteins that appear on both lists. All but 3 of the top 23 hits from the isotopically-labeled samples appear in the full dataset for unlabeled photocrosslinking (proteins identified in this dataset but did not make the cutoff for the top hits listed in Table 6-3), 14 of which appear in the non UV-treated control for the unlabeled UV-crosslinked *Msm*LprG(Q98pBpa). Fifteen of the top hits have been or are thought to be cell wall-associated and include integral membrane, periplasmic or secreted proteins.

Table 6-4. Top MS hits from FLAG IP of isotopically-labeled LprG(Q98pBpa)-FLAG photocrosslinked adducts. Proteins listed were identified in either one or both of two replicate LC-MS/MS runs of a FLAG IP of *Msm*LprG(Q98pBpa)-FLAG. Only hits that had ≥ 2 peptides identified and a $^{14}N/^{15}N$ ratio > 2 are listed below. NH, no homolog; *also top hit in unlabeled LprG(Q98pBpa) UV-crosslinking IP MS; **also top hit in DSP-crosslinking IP MS; §proteins that were in both unlabeled Q98pBpa no UV and UV-crosslinked datasets and were discarded as hits in Table 6-3; blank spaces indicate lack of information.

Protein information				Replicate 1			Replicate 2		
Gene	**Name**	**H37Rv Homolog**	**Description**	**Ratio**	**Std Dev**	**# Quant Peptides**	**Ratio**	**Std Dev**	**# Quant Peptides**
MSMEG_4217§	*wag31*	Rv2145c	DivIVA family protein	2.98	0.26	12	3.77	0.58	13
MSMEG_0913§	*umaA*	Rv0469	Possible mycolic acid synthase	3.42	0.37	6	3.58	0.45	9
MSMEG_5415	*eno*	Rv1023	Probable membrane-associated enolase	1.93	0.16	6	1.89	0.25	7
MSMEG_4692		Rv2468c	Conserved protein	1.96	0.36	4	2.06	0.17	4
MSMEG_0835**	*sodC*	Rv0432	Periplasmic superoxide dismutase [Cu-Zn]	2.69	0.01	2	3.66	0.42	3
MSMEG_0643§		NH	Putative extracellular solute-binding protein, family protein 5	2.29	0.22	6	2.8	0.18	3
MSMEG_4262§	*qcrA*	Rv2195	Probable rieske iron-sulfur protein	1.97	0.17	5	2.25	0.24	3
MSMEG_3904§	*lppK*	Rv2116	Conserved lipoprotein	12.04	11.26	2	10.8	13	2
MSMEG_0520* MSMEG_0965* MSMEG_5483*	*mspB* *mspA* *mspC*	NH NH NH	Porin	4.07	0.45	2	4.97	0.11	2
MSMEG_2727§		NH	Putative glutamate binding transport protein	2.09	0.34	2	2.25	0.71	2

MSMEG_6759§	*glpK*	NH	Glycerol kinase, ATP binding, glycerol-3-phosphate process				2.1	0.24	20
MSMEG_1441§	*sucB*	Rv0706	50S ribosomal protein L22 RplV				2.59	0.45	5
MSMEG_1060§		NH	Putative Lsr2 protein, regulatory protein involved in multiple cellular processes including cell wall biosynthesis & antibiotic resistance				1.86	0.15	4
MSMEG_0317**§		Rv0227c	Probable conserved membrane protein				3.56	0.77	3
MSMEG_3058§		NH	Lipoprotein, nlpa family protein				3.3	2.14	3
MSMEG_2410**§		Rv2969c	Possible conserved membrane or secreted protein				1.79	0.14	3
MSMEG_1642**§		Rv1747	Probable conserved transmembrane ATP-binding protein ABC transporter				1.74	0.24	3
MSMEG_3903§	*cfp2*	Rv2376c	Low molecular weight antigen (low molecular weight protein antigen 2)				20	0	2
MSMEG_4561*		NH	ABC Fe3+-siderophores transporter, periplasmic binding protein				20	0	2
MSMEG_6391**	*accD4*	Rv3799c	Probable propionyl-CoA carboxylase beta chain 4 (pccase) (propanoyl-CoA:carbon dioxide ligase)				10.64	13.23	2
MSMEG_4658		NH	Sugar ABC transporter substrate-binding protein				2.64	0.65	2
MSMEG_3851*	*lppI*	Rv2046	Probable lipoprotein	20	0	2			
MSMEG_4773		NH	Amidohydrolase-family protein	20	0	2			

As a proof-of-principle confirmation, the interaction of *Msm*LprG and MspA was validated by anti-MspA of whole-cell lysate and FLAG IP samples. Immunoblot of whole-cell lysates from *Msm* expressing WT-LprG or *Msm*LprG(Q98pBpa) indicated the presence of the intact MspA porin octamer, which is detergent resistant and therefore runs as an intact oligomer by SDS-PAGE (Figure 6-11, left). In the UV-treated, crosslinked *Msm*LprG(Q98pBpa) FLAG IP sample, a band is detected at ~46 kDa, which is consistent with MspA monomer (22 kDa) crosslinked to LprG (24 kDa) (Figure 6-11, right).

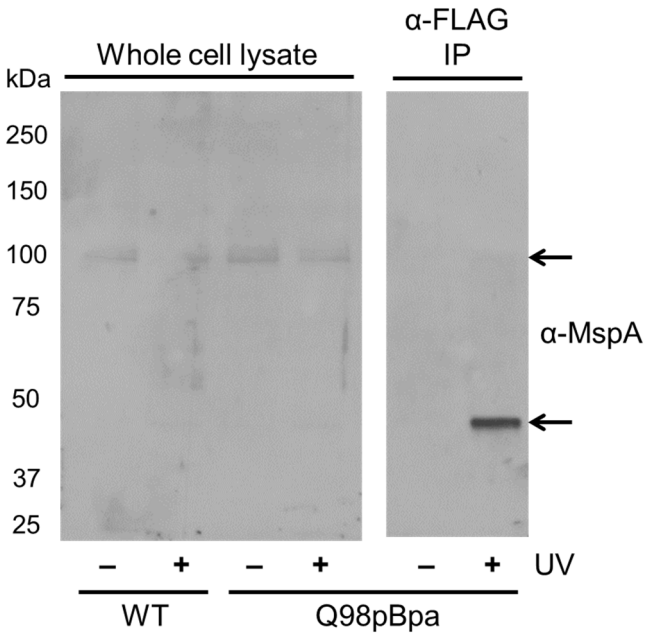


Figure 6-11. Anti-MspA immunoblot supports MS identification of LprG-interaction. (*left*) immunoblot against *Msm* outer membrane porin MspA reveals the expected SDS-resistant octamer at ~100 kDa (indicated by top arrow) in whole cell lysates. (*right*) Upon immunoprecipitation of LprG(Q98pBpa)-FLAG, a band at ~46 kDa (indicated by bottom arrow) is detected only in the UV-treated sample and is consistent with LprG (24 kDa) crosslinked to MspA monomer (22 kDa).

Discussion

Primary amine crosslinking by DSP to identify LprG-interacting proteins

We have shown that interactions captured by DSP are likely physiological since only full-length LprG, which is located in the cell wall, generates higher molecular-weight adducts indicative of covalent LprG-protein interactions. Immunoprecipitation of FLAG-tagged LprG and LprG-adducts was shown to be a preferred method for enrichment for downstream proteomic applications due to a higher yield and purity of tagged proteins. This result is not surprising as the FLAG antibody is highly specific for the FLAG epitope whereas Ni-NTA resin has the greater potential for nonspecific binding (260).

Proteins identified by LC-MS/MS in DSP-crosslinked LprG samples revealed that more than 75% of the hits have predicted cell wall-associated functions (Table 6-2). The remaining 4 proteins are AccA3, AccD4, MSMEG_2695 and FixA. AccA3 and AccD4 are soluble enzymes involved in fatty acid biosynthesis. Very little is known about MSMEG_2695: it is a conserved hypothetical protein and its *Mtb* homolog has been identified in the membrane fraction and in the cytosolic fraction, but not the culture filtrate (261). FixA is a subunit of the electron transport chain. Detection of AccA3, AccD4 and FixA as top LprG-interacting hits indicates a potential pitfall of DSP crosslinking as these are soluble proteins lacking known secretion signals. DSP is a nonspecific amine-reactive crosslinker and may therefore crosslink entire multiprotein complexes. This would lead to the identification of proteins that do not necessarily interact directly with LprG and highlights a disadvantage of this crosslinking technique. For example, AccA3 and AccD4 function in fatty acid biosynthesis and the fatty acids products are subsequently incorporated into lipids. After synthesis in the cytosol, OM-bound lipids must be transported across the IM, through the periplasm and assembled in the OM. Previous studies indicate that lipid biosynthetic enzymes

and transporters may associate to promote efficient substrate processing. Based on this model and given that LprG has been implicated in the lipid transport to the OM, AccA3 and AccD4 may appear as hits as a result of their interactions with IM lipid transporters rather than direct association with LprG.

Given their demonstrated genetic and functional association, we predicted that LprG would crosslink with the co-transcribed MSMEG_3069 (*Mtb* ortholog Rv1410c) transporter (203). We were able to identify two peptides from MSMEG_3069 in one technical replicate of the DSP-crosslinked sample and therefore did not report this potential interaction in Table 6-2. Upon further inspection, the bioinformatically predicted orientation of MSMEG_3069 in the IM indicated only one Lys residue in a periplasmic loop and therefore only one possible location for DSP-crosslinking. This result highlighted another major disadvantage of the amine-crosslinker: its dependence on primary amines of protein interaction partners being in close enough proximity for covalent modification.

Site-specific incorporation of the photocrosslinking unnatural amino acid pBpa via nonsense suppression

We demonstrated the incorporation of pBpa site-specifically into proteins in mycobacteria, as previously demonstrated with GFP. We confirmed that full-length GFP is expressed in the presence of the engineered *Mj*TyrRS, *Mj*tRNA^{Tyr} and pBpa (Figure 6-5). We also showed for multiple TAG mutants of *Msm*LprG that full-length protein is only expressed in the presence of pBpa (Figure 6-7). Regardless of expression level, the truncated, cytosolic control for DSP-crosslinking (Figure 6-3) indicates that only properly processed LprG will form covalent adducts

with physiological interaction partners. In other work, we have shown that overexpressing truncated LprG in *Msm* results in similar protein levels as overexpressing full-length protein suggesting that this result is not an artifact of differential protein expression levels (data not shown).

Our data suggests that in the presence of pBpa, *Mj*TyrRS is capable of incorporating pBpa at endogenous Tyr codons of WT-*Msm*LprG (Figure 6-8). This could be a result of an altered stoichiometry of synthetase to tRNA as the specific recognition of tRNA by its orthogonal synthetase requires the correct balance (i.e., in the absence of sufficient *Mj*tRNA, the synthetase may use *Msm*tRNA^{Tyr}) (262). Furthermore, the *Mj*TyrRS/*Mj*tRNA^{Tyr} pair used in this study was optimized for use in *E. coli*, not mycobacteria (250). Although *Msm*tRNA^{Tyr} and *Ecoli*tRNATyr are similar and therefore expected both to be orthogonal to *Mj*tRNA^{Tyr}, there may be key differences that allow for a low level of pBpa incorporation into Tyr codons in mycobacteria.

Crosslinking by photoreactive pBpa occurs via C-H bond insertion and may therefore covalently modify proximal proteins or even ligand; it is not specifically a protein-protein covalent modifier. LprG-ligand adducts are therefore a possibility, but will not necessarily result in a shift in MW detectable by SDS-PAGE, and that proteomics analysis methods were not optimized to account for such adducts. However, if LprG were bound the acyl chains of LAM, LM or PIM and crosslinked to these ligands, there would be a detectable shift in MW as these lipids can be separated by SDS-PAGE and are observed at 25, 15 and 10 kDa, respectively. The anti-*Mtb*LAM antibody has been shown to cross-react with *Msm*LAM; however immunoblot did not reveal a shift in LAM that would indicate crosslinking to any of the mutants assayed (data not shown). However, bulky mutations in the binding pocket of LprG can significantly reduce lipid binding, as in the well-characterized pocket mutant V91W. Insertion of pBpa in this pocket could have a

similar effect, which would reduce the yield of any protein-ligand adducts, possibly below the limit of detection.

Introduction of pBpa at the entrance to the ligand-binding pocket, or mouth, yielded many higher MW adducts (Figure 6-10, Figure 6-6). This result is consistent with our hypothesis that the mouth of LprG interacts with other proteins involved in lipid trafficking, based on previous studies on the structural homolog and lipoprotein transport protein LolA. LolA acquires lipoproteins from LolC at the IM and passes this cargo to LolB at the OM via mouth-to-mouth transfer (253). Therefore, the Q98pBpa mutant was chosen for MS analysis to identify interacting proteins that may be related to the function of LprG in lipid transport.

Introduction of pBpa within the C-terminal tail of *Msm*LprG yielded several distinct higher MW bands, indicating specific interactions within this domain (Figure 6-10, Figure 6-6). This novel finding is supported by the mechanism by which *E.coli* LptA transfers LPS. LptA forms oligomers with itself and interacts with IM protein LptC and OM protein LptD in a head-to-tail manner with the "head" referring to the N-terminus and "tail" referring to the C-terminus of the proteins (252). *In vitro*, the N- and C- termini of LprG are dispensable for ligand binding and TAG transport but these unstructured regions may be critical for *in vivo* protein-protein interactions and therefore lipid transport (Bai and Seeliger, unpublished data).

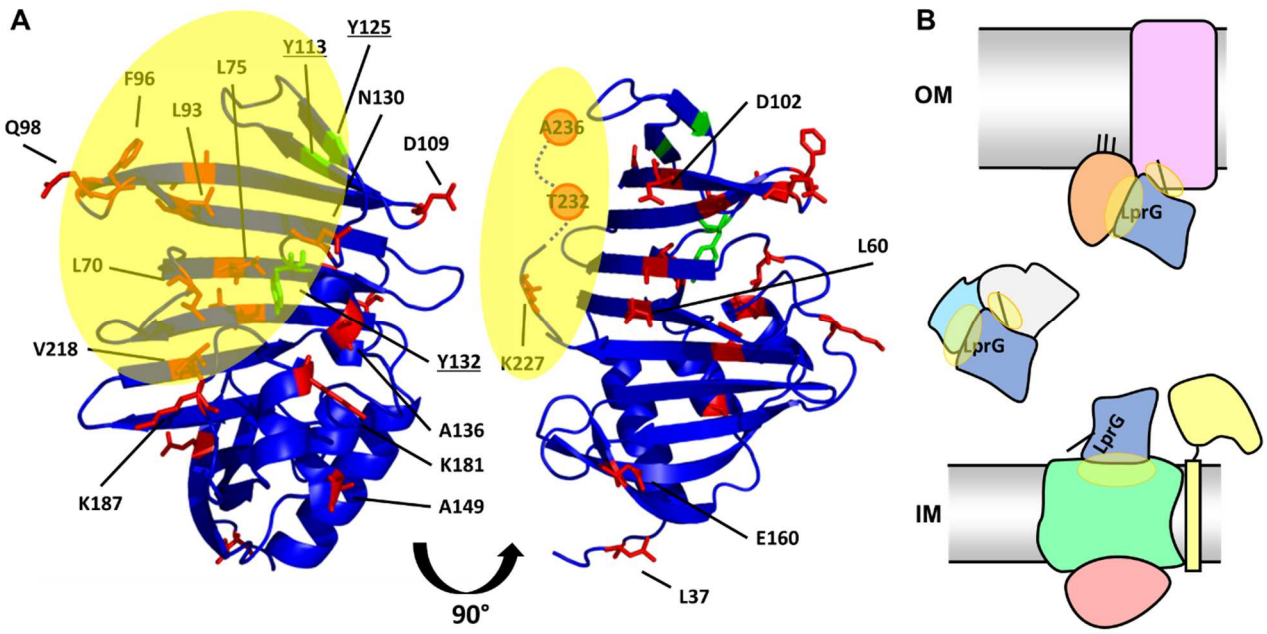


Figure 6-12. *Msm* LprG interaction interfaces identified by photocrosslinking of pBpa mutants. (A) The protein-protein interaction interfaces of LprG are indicated by the yellow shading. Residues at the entrance of the ligand binding pocket and at the C-terminal tail of LprG have distinct crosslinking banding patterns. (B) Schematic outlining a general interaction mechanism for LprG with proteins in the IM, periplasmic space and OM.

LprG interaction partners identified by proteomics

Although expected, an interaction with MSMEG_3069 was not reliably detected in any crosslinking experiment. As noted, this could be due to the low probability of crosslinking amines in the conjugate proteins by DSP. Also, absence of crosslinking to MSMEG at Q98 may indicate that MSEMG_3069 does not interact with LprG at the mouth interface, but may interact at another interface. It is also possible that they do not physically interact. It is possible that more than one transporter is involved in lipid transport through and extraction from the IM in mycobacteria. For example, in *E. coli,* the transporter MsbA flips LPS from the cytosolic side of the IM to the periplasmic side and the ABC transporter LptBCFG extracts LPS from the membrane such that it can be passed in a head-to-tail fashion via LptA to the OM (214,215). Of the three datasets reported

here, six ABC transporters were identified as interactors of LprG. Only one of those has a homolog in *Mtb,* MSMEG_1642 (Rv1747). We are interested in elucidating a transport mechanism conserved across mycobacterial species and therefore will focus in the future on hits with *Mtb* homologs. In order to characterize MSMEG_1642, which is an uncharacterized protein, gene knockout strains will be tested for a sliding motility phenotype as a readout for LprG function and profiled for changes in OM lipid composition since LprG affects lipid localization to the OM (50,204,205,263).

Only a few proteins overlap between any two datasets reported in Table 6-2, Table 6-3 and Table 6-4 (Figure 6-13). Further examination revealed that more than half (14 of 23) of the isotopically-labeled UV-crosslinked samples were also detected in the unlabeled photocrosslinked dataset as well as in the no UV control sample. These proteins did not meet our cutoff criteria for the analysis of the unlabeled proteomics experiment because they appeared in both the control and crosslinked samples. Relative spectral counts between the two samples could not be quantified because of high run-to-run variability in the LC-MS/MS. This highlights the potential gain in accuracy afforded by the isotopically labeled proteomics, which provides quantitative measures of enrichment.

Our crosslinking data suggests that LprG interacts with MspA, MspB and/or MspC at the entrance to the LprG hydrophobic ligand-binding pocket. Immunoblot analysis corroborated this proteomics finding (Table 6-3, Table 6-4, and Figure 6-11). MspA is the major OM porin in *Msm,* but there are three additional *Msm* porins, MspB, MspC and MspD which are identical in sequence to MspA with the exception of 2, 4 and 18 residues, respectively (264). Because of high sequence identity between the three porins, distinguishing between the three in this experiment is highly unlikely unless the peptide containing the distinguishing residues is identified. Also, because of

the high sequence identity, the anti-MspA antiserum likely cross-reacts with MspB or MspC. Regardless, this is the first experimental evidence supporting LprG localization to the OM in mycobacteria and, more broadly, the first evidence that any mycobacterial lipoprotein reaches the OM. However, more data is needed to reveal whether or not there is a functional relationship between LprG and MspA/B/C in lipid transport and/or assembly at the bacterial surface. This may represent a non-canonical feature of mycobacterial OM porins as OM porins are generally thought to transport small hydrophilic molecules across the OM (265).

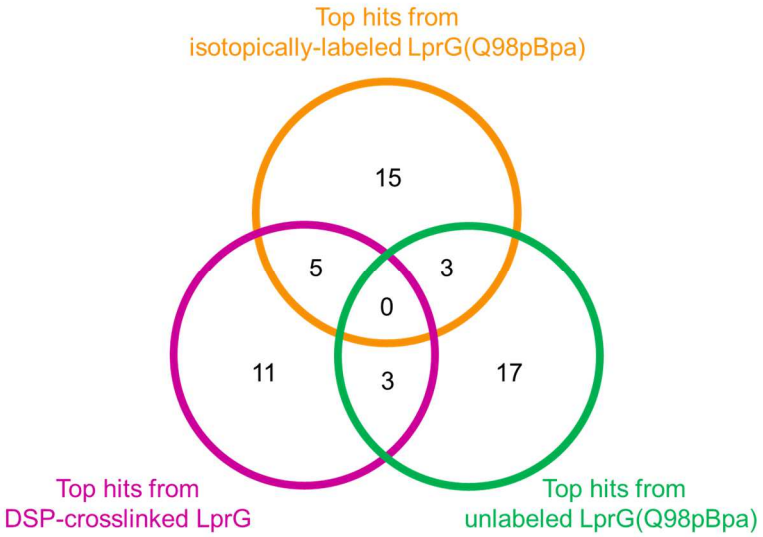


Figure 6-13. Summary of MS hits among proteomic datasets.

Interacting proteins may be functionally associated with either cell wall biogenesis or lipid transport

Interestingly, mycolic acid biosynthesis proteins were detected in every proteomic dataset. In the DSP crosslinked dataset, MmpL3, AccA3 and AccD4 were identified as top LprG-interacting proteins. MmpL3 is an integral transmembrane transporter involved in translocating mycolic acids across the IM (100,128). AccA3 and AccD4 are part of the acyl-CoA carboxylase (ACC) complex and are enzymes involved in the early steps of fatty acid and mycolic acid biosynthesis (266,267). In the unlabeled photocrosslinked dataset, antigen-85A (Ag85A) was identified as a top LprG-interacting protein. Ag85A is a cell wall-associated mycolyltransferase that is essential for the synthesis of the virulence-associated lipid trehalose dimycolate (268,269). Another top hit was FadD32, which adenylates and ligates the meromycolic chain onto Pks13, the terminal condensation protein in mycolic acid synthesis. Both Pks13 and FadD32 have been co-localized to the subpolar region of the growing pole of mycobacteria (85). Finally, in the isotopically-labeled photocrosslinked proteomic dataset, UmaA and AccD4, both enzymes involved in mycolic acid biosynthesis, were identified (270).

DivIVA is required for cell growth and is responsible for localization of cell wall synthesis protein complexes at the growing pole (85). The *B. subtilis* DivIVA ortholog has been shown to bind membranes via an amphiphilic N-terminus that is conserved across species, including mycobacteria (86). Wag31, the *Msm* DivIVA ortholog and a putative LprG-interacting protein (Table 4), may directly interact with and localize LprG to the growing pole of *M. smegmatis*. DivIVA has previously been shown to form an SDS-resistant complex with AccA3, AccD4 and AccD5, and this strong interaction could explain the detection of these three cytosolic proteins in each MS dataset (85). The identification of proteins lacking a predicted secretion signal or

transmembrane domain in the cell wall fraction is unprecedented (271-275). It is possible that these proteins are exported by a secretion system other than Sec or Tat, or be recognized by a secretion system by another mean than a canonical signal peptide. MmpL3, AccA3 and AccD4 have also been co-localized to the growing pole of mycobacteria and have been identified as DivIVA-interacting proteins (89).

Another protein identified as an LprG-interaction partner was PonA1 (Table 6-2 and Table 6-3). PonA1 is a major peptidoglycan (PG) synthase capable of regulating mycobacterial polar growth. It is a bifunctional enzyme that has a TM-spanning helix and two catalytic periplasmic domains that catalyze transglycosylation and transpeptidation of PG (90). PonA1 has also been localized to the cell poles and septa (87).

The identification of proteins involved in cell wall synthesis as LprG interactors is reasonable as we hypothesis LprG has a role in membrane biogenesis via OM lipid transport. We focused on Q98TAG, a mutation at the mouth of the ligand-binding pocket, to identify proteins that may collaborate with LprG in lipid trafficking. We hypothesized that lipid transport may occur via mouth-to-mouth transfer, as occurs for the structurally homologous lipoprotein transporters LolA and LolB. Our data suggest that some LprG interactions may not be directly related to lipid trafficking, but instead to coordination of protein localization for polar growth. The putative interactions with mycolic acid biosynthesis may be related to the co-localization of these proteins rather than a role for LprG in the transport of cell-wall mycolic acid precursors such as trehalose dimycolate, since the ligand-binding pocket of LprG is not large enough to accommodate the C_{60} mycolic acids in *Msm* or $C_{60\text{-}90}$ mycolic acids in *Mtb* (276).

General discussion of MS results and implications for future experiments

Our finding that LprG localizes to the OM supports the existence of a lipoprotein transport system in mycobacteria; Gram positive bacteria that as monoderms do not require such a system, and no such system has been identified in mycobacteria or related diderms within the Corynebacterineae. Unfortunately, we cannot depend on sequence homology to help identify a possible Lol-like system as *E. coli* LolA and LolB have very low sequence identity (11%) but are structurally homologous and the same is true for LptA and LptC of the *E. coli* LPS transport system. Since sequence homology does not provide clues as to which genes could be responsible, as homologues of the Lol system are present in mycobacteria, other methods are necessary to make this determination. The *E. coli* LPS and lipoprotein transport systems provide functional models, but each is a distinct mechanism for a distinct purpose. At this point, we are unable to make any conclusions as to which, if either system, the mycobacterial lipid transport mechanism may more closely resemble. Using the methodology described here, the incorporation of pBpa at the N-terminus of LprG may lead not only to additional proteins within the LprG-mediated lipid transport pathway, but also proteins that transport LprG itself, as well as other lipoproteins. In order to confirm the roles of the putative lipoprotein transport system, functional cross-species complementation (i.e., complementing MSMEG knock-out strains with E. coli Lol components) may be used.

The studies proposed above may also lead to the identification of a consensus signal for the localization of lipoproteins in the IM or OM. For example, the +2 rule, where the residue C-terminal to the conserved Cys dictates localization in *E. coli*. Also in *E. coli*, a negative charge or amide group in the +3 position acts as a retention signal. This is consistent with our results localizing *Msm*LprG to the OM as the residues following Cys in *Msm*LprG are Ser-Ser-Ser and in

*Mtb*LprG are Ser-Ser-Gly. It is also worth noting that the signal for OM localization by the Lol system in *E. coli* is a Ser in the +2 position and *Msm*LprG and *Mtb*LprG both have Ser in this position (239).

The N- and C- termini of LprG are hypothesized to play important roles in mycobacterial lipid trafficking. Studies to elucidate the significance of the C-terminus in LprG function are underway in the Seeliger lab using the sliding motility assay as a readout for LprG activity. Sliding motility is thought to reflect changes in membrane composition at the cell surface. Compared to WT, strains lacking *lprG*, have decreased sliding motility. This phenotype can be complemented with WT-*lprG* and partially complemented with the pocket mutant *lprG(V91W)*. We are interested in testing the sliding motility of mutants lacking the C-terminus to determine whether or not this portion of the protein mediates protein interactions that are required for lipid transport function.

Creating TAG mutations at the N-terminus for UAAI studies will allow us to define interactions at this terminus of the protein. These studies are expected to identify the lipoprotein-processing proteins Lgt, LspA and Lnt. Finally, we propose to study the subcellular localization of LprG and its interaction partners' based on our hypothesis that these protein complexes promote coordinated cell wall biogenesis and therefore localize at the cellular poles and septa.

References

1. WorldHealthOrganization. WHO Global Tuberculosis Report 2014.

2. Corbett, E. L., Watt, C. J., Walker, N., Maher, D., Williams, B. G., Raviglione, M. C., and Dye, C. (2003) The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. *Archives of Internal Medicine* **163**, 1009-1021

3. Zhang, Y. (2005) The magic bullets and tuberculosis drug targets. *Annual Review of Pharmacology and Toxicology* **45**, 529-564

4. Fischbach, M. A., and Walsh, C. T. (2009) Antibiotics for emerging pathogens. *Science* **325**, 1089-1093

5. Flynn, J. L., and Chan, J. (2001) Tuberculosis: latency and reactivation. *Infection and Immunity* **69**, 4195-4201

6. Flynn, J. L., and Chan, J. (2001) Immunology of tuberculosis. *Annual Review of Immunology* **19**, 93-129

7. Houben, E. N., Nguyen, L., and Pieters, J. (2006) Interaction of pathogenic mycobacteria with the host immune system. *Current Opinion in Microbiology* **9**, 76-85

8. Russell, D. G., Cardona, P. J., Kim, M. J., Allain, S., and Altare, F. (2009) Foamy macrophages and the progression of the human tuberculosis granuloma. *Nature Immunology* **10**, 943-948

9. Mukhopadhyay, S., Farver, C. F., Vaszar, L. T., Dempsey, O. J., Popper, H. H., Mani, H., Capelozzi, V. L., Fukuoka, J., Kerr, K. M., Zeren, E. H., Iyer, V. K., Tanaka, T., Narde, I., Nomikos, A., Gumurdulu, D., Arava, S., Zander, D. S., and Tazelaar, H. D. (2012) Causes of pulmonary granulomas: a retrospective study of 500 cases from seven countries. *Journal of Clinical Pathology* **65**, 51-57

10. Caceres, N., Tapia, G., Ojanguren, I., Altare, F., Gil, O., Pinto, S., Vilaplana, C., and Cardona, P. J. (2009) Evolution of foamy macrophages in the pulmonary granulomas of experimental tuberculosis models. *Tuberculosis* **89**, 175-182

11. Peyron, P., Vaubourgeix, J., Poquet, Y., Levillain, F., Botanch, C., Bardou, F., Daffe, M., Emile, J. F., Marchou, B., Cardona, P. J., de Chastellier, C., and Altare, F. (2008) Foamy macrophages from tuberculous patients' granulomas constitute a nutrient-rich reservoir for M. tuberculosis persistence. *PLoS Pathogens* **4**, e1000204

12. De Rossi, E., Ainsa, J. A., and Riccardi, G. (2006) Role of mycobacterial efflux transporters in drug resistance: an unresolved question. *FEMS Microbiology Reviews* **30**, 36-52

13. Brennan, P. J., and Nikaido, H. (1995) The Envelope of Mycobacteria. *Annual Review of Biochemistry* **64**, 29-63

14. Jarlier, V., and Nikaido, H. (1994) Mycobacterial cell wall: structure and role in natural resistance to antibiotics. *FEMS Microbiology Letters* **123**, 11-18

15. Hett, E. C., and Rubin, E. J. (2008) Bacterial Growth and Cell Division: a Mycobacterial Perspective. *Microbiology and Molecular Biology Reviews* **72**, 126-156

16. Takayama, K., and Kilburn, J. O. (1989) Inhibition of synthesis of arabinogalactan by ethambutol in Mycobacterium smegmatis. *Antimicrobial Agents and Chemotherapy* **33**, 1493-1499

17. Zhang, Y., Heym, B., Allen, B., Young, D., and Cole, S. (1992) The catalase[mdash]peroxidase gene and isoniazid resistance of Mycobacterium tuberculosis. *Nature* **358**, 591-593

18. Jackson, M., Crick, D. C., and Brennan, P. J. (2000) Phosphatidylinositol is an essential phospholipid of mycobacteria. *Journal of Biological Chemistry* **275**, 30092-30099

19. Akamatsu, Y., and Nojima, S. (1965) Separation and analyses of the individual phospholipids of mycobacteria. *Journal of Biochemistry* **57**, 430-439

20. Sassetti, C. M., Boyd, D. H., and Rubin, E. J. (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular Microbiology* **48**, 77-84

21. Liu, J., Barry, C. E., 3rd, Besra, G. S., and Nikaido, H. (1996) Mycolic acid structure determines the fluidity of the mycobacterial cell wall. *Journal of Biological Chemistry* **271**, 29545-29551

22. Liu, J., Rosenberg, E. Y., and Nikaido, H. (1995) Fluidity of the lipid domain of cell wall from Mycobacterium chelonae. *Procedings of the National Academy of Science U S A* **92**, 11254-11258

23. Hoffmann, C., Leis, A., Niederweis, M., Plitzko, J. M., and Engelhardt, H. (2008) Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. *Procedings of the National Academy of Science* **105**, 3963-3967

24. Daffe, M., Brennan, P. J., and McNeil, M. (1990) Predominant structural features of the cell wall arabinogalactan of Mycobacterium tuberculosis as revealed through characterization of oligoglycosyl alditol fragments by gas chromatography/mass spectrometry and by 1H and 13C NMR analyses. *Journal of Biological Chemistry* **265**, 6734-6743

25. McNeil, M., Daffe, M., and Brennan, P. J. (1991) Location of the mycolyl ester substituents in the cell walls of mycobacteria. *Journal of Biological Chemistry* **266**, 13217-13223

26. Chatterjee, D. (1997) The mycobacterial cell wall: structure, biosynthesis and sites of drug action. *Current Opinion in Chemical Biology* **1**, 579-588

27. Daffe, M., and Draper, P. (1998) The envelope layers of mycobacteria with reference to their pathogenicity. *Advances in microbial physiology* **39**, 131-203

28. Phiet, P. H., Wietzerbin, J., Zissman, E., Petit, J. F., and Lederer, E. (1976) Analysis of the cell wall of five strains of Myocbacterium tuberculosis BCG and of an attenuated human strain, W 115. *Infection and Immunity* **13**, 677-681

29. Collins, M. D., Goodfellow, M., and Minnikin, D. E. (1982) A survey of the structures of mycolic acids in Corynebacterium and related taxa. *Journal of General Microbiology* **128**, 129-149

30. Minnikin, D. E., Minnikin, S. M., Goodfellow, M., and Stanford, J. L. (1982) The mycolic acids of Mycobacterium chelonei. *Journal of General Microbiology* **128**, 817-822

31. Davidson, L. A., Draper, P., and Minnikin, D. E. (1982) Studies on the mycolic acids from the walls of Mycobacterium microti. *Journal of General Microbiology* **128**, 823-828

32. Draper, P., Dobson, G., Minnikin, D. E., and Minnikin, S. M. (1982) The mycolic acids of Mycobacterium leprae harvested from experimentally infected nine-banded armadillos. *Annales de Microbiologie* **133**, 39-47

33. Collins, M. D., Goodfellow, M., and Minnikin, D. E. (1982) Fatty acid composition of some mycolic acid-containing coryneform bacteria. *Journal of General Microbiology* **128**, 2503-2509

34. Brennan, P. J. (2003) Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis. *Tuberculosis* **83**, 91-97

35. Neyrolles, O., and Guilhot, C. (2011) Recent advances in deciphering the contribution of Mycobacterium tuberculosis lipids to pathogenesis. *Tuberculosis* **91**, 187-195

36. Quesniaux, V. J., Nicolle, D. M., Torres, D., Kremer, L., Guerardel, Y., Nigou, J., Puzo, G., Erard, F., and Ryffel, B. (2004) Toll-like receptor 2 (TLR2)-dependent-positive and TLR2-independent-negative regulation of proinflammatory cytokines by mycobacterial lipomannans. *Journal of Immunology* **172**, 4425-4434

37. Riviere, M., Moisand, A., Lopez, A., and Puzo, G. (2004) Highly ordered supra-molecular organization of the mycobacterial lipoarabinomannans in solution. Evidence of a relationship between supra-molecular organization and biological activity. *Journal of Molecular Biology* **344**, 907-918

38. Drage, M. G., Tsai, H.-C., Pecora, N. D., Cheng, T.-Y., Arida, A. R., Shukla, S., Rojas, R. E., Seshadri, C., Moody, D. B., Boom, W. H., Sacchettini, J. C., and Harding, C. V. (2010) Mycobacterium tuberculosis lipoprotein LprG (Rv1411c) binds triacylated glycolipid agonists of Toll-like receptor 2. *Nature Structural and Molecular Biology* **17**, 1088-1095

39. Minnikin, D. E., Dobson, G., Parlett, J. H., Goodfellow, M., and Magnusson, M. (1987) Analysis of dimycocerosates of glycosylphenolphthiocerols in the identification of some clinically significant mycobacteria. *European Journal of Clinical Microbiology* **6**, 703-707

40. Goren, M. B., Brokl, O., and Schaefer, W. B. (1974) Lipids of putative relevance to virulence in Mycobacterium tuberculosis: phthiocerol dimycocerosate and the attenuation indicator lipid. *Infection and Immunity* **9**, 150-158

41. Cox, J. S., Chen, B., McNeil, M., and Jacobs, W. R. (1999) Complex lipid determines tissue-specific replication of Mycobacterium tuberculosis in mice. *Nature* **402**, 79-83

42. Camacho, L. R., Ensergueix, D., Perez, E., Gicquel, B., and Guilhot, C. (1999) Identification of a virulence gene cluster of Mycobacterium tuberculosis by signature-tagged transposon mutagenesis. *Molecular Microbiology* **34**, 257-267

43. Yu, J., Tran, V., Li, M., Huang, X., Niu, C., Wang, D., Zhu, J., Wang, J., Gao, Q., and Liu, J. (2012) Both phthiocerol dimycocerosates and phenolic glycolipids are required for virulence of Mycobacterium marinum. *Infection and Immunity* **80**, 1381-1389

44. Saavedra, R., Segura, E., Leyva, R., Esparza, L. A., and López-Marín, L. M. (2001) Mycobacterial Di-O-Acyl-Trehalose Inhibits Mitogen- and Antigen-Induced Proliferation of Murine T Cells In Vitro. *Clinical and Diagnostic Laboratory Immunology* **8**, 1081-1088

45. Saavedra, R., Segura, E., Tenorio, E. P., and Lopez-Marin, L. M. (2006) Mycobacterial trehalose-containing glycolipid with immunomodulatory activity on human CD4+ and CD8+ T-cells. *Microbes and Infection / Institut Pasteur* **8**, 533-540

46. Lee, K.-S., Dubey, V. S., Kolattukudy, P. E., Song, C.-H., Shin, A. R., Jung, S.-B., Yang, C.-S., Kim, S.-Y., Jo, E.-K., Park, J.-K., and Kim, H.-J. (2007) Diacyltrehalose of Mycobacterium tuberculosis inhibits lipopolysaccharide- and mycobacteria-induced proinflammatory cytokine production in human monocytic cells. *FEMS Microbiology Letters* **267**, 121-128

47. Rohde, K. H., Abramovitch, R. B., and Russell, D. G. (2007) Mycobacterium tuberculosis Invasion of Macrophages: Linking Bacterial Gene Expression to Environmental Cues. *Cell Host Microbe* **2**, 352-364

48. Guiard, J., Collmann, A., Garcia-Alles, L. F., Mourey, L., Brando, T., Mori, L., Gilleron, M., Prandi, J., De Libero, G., and Puzo, G. (2009) Fatty acyl structures of mycobacterium tuberculosis sulfoglycolipid govern T cell response. *Journal of Immunology* **182**, 7030-7037

49. Sulzenbacher, G., Canaan, S., Bordat, Y., Neyrolles, O., Stadthagen, G., Roig-Zamboni, V., Rauzier, J., Maurin, D., Laval, F., Daffe, M., Cambillau, C., Gicquel, B., Bourne, Y., and Jackson, M. (2006) LppX is a lipoprotein required for the translocation of phthiocerol dimycocerosates to the surface of Mycobacterium tuberculosis. *EMBO J* **25**, 1436-1444

50. Shukla, S., Richardson, E. T., Athman, J. J., Shi, L., Wearsch, P. A., McDonald, D., Banaei, N., Boom, W. H., Jackson, M., and Harding, C. V. (2014) Mycobacterium tuberculosis

lipoprotein LprG binds lipoarabinomannan and determines its cell envelope localization to control phagolysosomal fusion. *PLoS Pathogens* **10**, e1004471

51. Tailleux, L., Maeda, N., Nigou, J., Gicquel, B., and Neyrolles, O. (2003) How is the phagocyte lectin keyboard played? Master class lesson by Mycobacterium tuberculosis. *Trends in Microbiology* **11**, 259-263

52. Villeneuve, C., Gilleron, M., Maridonneau-Parini, I., Daffe, M., Astarie-Dequeker, C., and Etienne, G. (2005) Mycobacteria use their surface-exposed glycolipids to infect human macrophages through a receptor-dependent process. *Journal of Lipid Research* **46**, 475-483

53. Schlesinger, L. S. (1993) Macrophage phagocytosis of virulent but not attenuated strains of Mycobacterium tuberculosis is mediated by mannose receptors in addition to complement receptors. *Journal of Immunology* **150**, 2920-2930

54. Venisse, A., Fournie, J. J., and Puzo, G. (1995) Mannosylated lipoarabinomannan interacts with phagocytes. *European Journal of Biochemistry* **231**, 440-447

55. Tailleux, L., Schwartz, O., Herrmann, J. L., Pivert, E., Jackson, M., Amara, A., Legres, L., Dreher, D., Nicod, L. P., Gluckman, J. C., Lagrange, P. H., Gicquel, B., and Neyrolles, O. (2003) DC-SIGN is the major *Mycobacterium tuberculosis* receptor on human dendritic cells. *Journal of Experimental Medicine* **197**, 121-127

56. Maeda, N., Nigou, J., Herrmann, J. L., Jackson, M., Amara, A., Lagrange, P. H., Puzo, G., Gicquel, B., and Neyrolles, O. (2003) The cell surface receptor DC-SIGN discriminates between Mycobacterium species through selective recognition of the mannose caps on lipoarabinomannan. *Journal of Biological Chemistry* **278**, 5513-5516

57. Astarie-Dequeker, C., Le Guyader, L., Malaga, W., Seaphanh, F. K., Chalut, C., Lopez, A., and Guilhot, C. (2009) Phthiocerol dimycocerosates of M. tuberculosis participate in macrophage invasion by inducing changes in the organization of plasma membrane lipids. *PLoS Pathogens* **5**, e1000289

58. Welin, A., Winberg, M. E., Abdalla, H., Sarndahl, E., Rasmusson, B., Stendahl, O., and Lerm, M. (2008) Incorporation of Mycobacterium tuberculosis lipoarabinomannan into macrophage membrane rafts is a prerequisite for the phagosomal maturation block. *Infection and Immunity* **76**, 2882-2887

59. Fratti, R. A., Backer, J. M., Gruenberg, J., Corvera, S., and Deretic, V. (2001) Role of phosphatidylinositol 3-kinase and Rab5 effectors in phagosomal biogenesis and mycobacterial phagosome maturation arrest. *The Journal of Cell Biology* **154**, 631-644

60. Kang, P. B., Azad, A. K., Torrelles, J. B., Kaufman, T. M., Beharka, A., Tibesar, E., DesJardin, L. E., and Schlesinger, L. S. (2005) The human macrophage mannose receptor

directs Mycobacterium tuberculosis lipoarabinomannan-mediated phagosome biogenesis. *Journal of Experimental Medicine* **202**, 987-999

61. Vergne, I., Chua, J., and Deretic, V. (2003) Tuberculosis toxin blocking phagosome maturation inhibits a novel Ca2+/calmodulin-PI3K hVPS34 cascade. *Journal of Experimental Medicine* **198**, 653-659

62. Rousseau, C., Sirakova, T. D., Dubey, V. S., Bordat, Y., Kolattukudy, P. E., Gicquel, B., and Jackson, M. (2003) Virulence attenuation of two Mas-like polyketide synthase mutants of *Mycobacterium tuberculosis*. *Microbiology* **149**, 1837-1847

63. Rousseau, C., Winter, N., Pivert, E., Bordat, Y., Neyrolles, O., Ave, P., Huerre, M., Gicquel, B., and Jackson, M. (2004) Production of phthiocerol dimycocerosates protects Mycobacterium tuberculosis from the cidal activity of reactive nitrogen intermediates produced by macrophages and modulates the early immune response to infection. *Cell Microbiology* **6**, 277-287

64. Goren, M. B., D'Arcy Hart, P., Young, M. R., and Armstrong, J. A. (1976) Prevention of phagosome-lysosome fusion in cultured macrophages by sulfatides of Mycobacterium tuberculosis. *Procedings of the National Academy of Science U S A* **73**, 2510-2514

65. Pabst, M. J., Gross, J. M., Brozna, J. P., and Goren, M. B. (1988) Inhibition of macrophage priming by sulfatide from Mycobacterium tuberculosis. *Journal of Immunology* **140**, 634-640

66. Zhang, L., Goren, M. B., Holzer, T. J., and Andersen, B. R. (1988) Effect of Mycobacterium tuberculosis-derived sulfolipid I on human phagocytic cells. *Infection and Immunity* **56**, 2876-2883

67. Brozna, J. P., Horan, M., Rademacher, J. M., Pabst, K. M., and Pabst, M. J. (1991) Monocyte responses to sulfatide from Mycobacterium tuberculosis: inhibition of priming for enhanced release of superoxide, associated with increased secretion of interleukin-1 and tumor necrosis factor alpha, and altered protein phosphorylation. *Infection and Immunity* **59**, 2542-2548

68. Bowdish, D. M., Sakamoto, K., Kim, M. J., Kroos, M., Mukhopadhyay, S., Leifer, C. A., Tryggvason, K., Gordon, S., and Russell, D. G. (2009) MARCO, TLR2, and CD14 are required for macrophage cytokine responses to mycobacterial trehalose dimycolate and Mycobacterium tuberculosis. *PLoS Pathogens* **5**, e1000474

69. Schoenen, H., Bodendorfer, B., Hitchens, K., Manzanero, S., Werninghaus, K., Nimmerjahn, F., Agger, E. M., Stenger, S., Andersen, P., Ruland, J., Brown, G. D., Wells, C., and Lang, R. (2010) Cutting edge: Mincle is essential for recognition and adjuvanticity of the mycobacterial cord factor and its synthetic analog trehalose-dibehenate. *Journal of immunology* **184**, 2756-2760

70. Ishikawa, E., Ishikawa, T., Morita, Y. S., Toyonaga, K., Yamada, H., Takeuchi, O., Kinoshita, T., Akira, S., Yoshikai, Y., and Yamasaki, S. (2009) Direct recognition of the mycobacterial glycolipid, trehalose dimycolate, by C-type lectin Mincle. *Journal of Experimental Medicine* **206**, 2879-2888

71. Reed, M. B., Domenech, P., Manca, C., Su, H., Barczak, A. K., Kreiswirth, B. N., Kaplan, G., and Barry, C. E., 3rd. (2004) A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response. *Nature* **431**, 84-87

72. Tanne, A., Ma, B., Boudou, F., Tailleux, L., Botella, H., Badell, E., Levillain, F., Taylor, M. E., Drickamer, K., Nigou, J., Dobos, K. M., Puzo, G., Vestweber, D., Wild, M. K., Marcinko, M., Sobieszczuk, P., Stewart, L., Lebus, D., Gicquel, B., and Neyrolles, O. (2009) A murine DC-SIGN homologue contributes to early host defense against Mycobacterium tuberculosis. *Journal of Experimental Medicine* **206**, 2205-2220

73. Gilleron, M., Stenger, S., Mazorra, Z., Wittke, F., Mariotti, S., Bohmer, G., Prandi, J., Mori, L., Puzo, G., and De Libero, G. (2004) Diacylated sulfoglycolipids are novel mycobacterial antigens stimulating CD1-restricted T cells during infection with Mycobacterium tuberculosis. *Journal of Experimental Medicine* **199**, 649-659

74. Cohen, N. R., Garg, S., and Brenner, M. B. (2009) Antigen Presentation by CD1 Lipids, T Cells, and NKT Cells in Microbial Immunity. *Advances in Immunology* **102**, 1-94

75. Kieser, K. J., and Rubin, E. J. (2014) How sisters grow apart: mycobacterial growth and division. *Nature Reviews Microbiology* **12**, 550-562

76. Typas, A., Banzhaf, M., Gross, C. A., and Vollmer, W. (2012) From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nature Reviews Microbiology* **10**, 123-136

77. de Pedro, M. A., Quintela, J. C., Holtje, J. V., and Schwarz, H. (1997) Murein segregation in Escherichia coli. *Journal of Bacteriology* **179**, 2823-2834

78. Daniel, R. A., and Errington, J. (2003) Control of cell morphogenesis in bacteria: two distinct ways to make a rod-shaped cell. *Cell* **113**, 767-776

79. Thanky, N. R., Young, D. B., and Robertson, B. D. (2007) Unusual features of the cell cycle in mycobacteria: polar-restricted growth and the snapping-model of cell division. *Tuberculosis* **87**, 231-236

80. Aldridge, B. B., Fernandez-Suarez, M., Heller, D., Ambravaneswaran, V., Irimia, D., Toner, M., and Fortune, S. M. (2012) Asymmetry and aging of mycobacterial cells lead to variable growth and antibiotic susceptibility. *Science* **335**, 100-104

81. Salje, J., van den Ent, F., de Boer, P., and Lowe, J. (2011) Direct membrane binding by bacterial actin MreB. *Molecular Cell* **43**, 478-487

82. Flardh, K. (2003) Growth polarity and cell division in Streptomyces. *Current Opinion in Microbiology* **6**, 564-571

83. Kang, C. M., Nyayapathy, S., Lee, J. Y., Suh, J. W., and Husson, R. N. (2008) Wag31, a homologue of the cell division protein DivIVA, regulates growth, morphology and polar cell wall synthesis in mycobacteria. *Microbiology* **154**, 725-735

84. Letek, M., Ordonez, E., Vaquera, J., Margolin, W., Flardh, K., Mateos, L. M., and Gil, J. A. (2008) DivIVA is required for polar growth in the MreB-lacking rod-shaped actinomycete Corynebacterium glutamicum. *Journal of Bacteriology* **190**, 3283-3292

85. Meniche, X., Otten, R., Siegrist, M. S., Baer, C. E., Murphy, K. C., Bertozzi, C. R., and Sassetti, C. M. (2014) Subpolar addition of new cell wall is directed by DivIVA in mycobacteria. *Proceedings of the National Academy of Sciences* **111**, E3243-E3251

86. Lenarcic, R., Halbedel, S., Visser, L., Shaw, M., Wu, L. J., Errington, J., Marenduzzo, D., and Hamoen, L. W. (2009) Localisation of DivIVA by targeting to negatively curved membranes. *EMBO Journal* **28**, 2272-2282

87. Hett, E. C., Chao, M. C., and Rubin, E. J. (2010) Interaction and Modulation of Two Antagonistic Cell Wall Enzymes of Mycobacteria. *PLoS Pathogens* **6**, e1001020

88. Ginda, K., Bezulska, M., Ziolkiewicz, M., Dziadek, J., Zakrzewska-Czerwinska, J., and Jakimowicz, D. (2013) ParA of Mycobacterium smegmatis co-ordinates chromosome segregation with the cell cycle and interacts with the polar growth determinant DivIVA. *Molecular Microbiology* **87**, 998-1012

89. Carel, C., Nukdee, K., Cantaloube, S., Bonne, M., Diagne, C. T., Laval, F., Daffé, M., and Zerbib, D. (2014) Mycobacterium tuberculosis Proteins Involved in Mycolic Acid Synthesis and Transport Localize Dynamically to the Old Growing Pole and Septum. *PLoS ONE* **9**, e97148

90. Kieser, K. J., Boutte, C. C., Kester, J. C., Baer, C. E., Barczak, A. K., Meniche, X., Chao, M. C., Rego, E. H., Sassetti, C. M., Fortune, S. M., and Rubin, E. J. (2015) Phosphorylation of the Peptidoglycan Synthase PonA1 Governs the Rate of Polar Elongation in Mycobacteria. *PLoS Pathogens* **11**, e1005010

91. Khan, A. A., Stocker, B. L., and Timmer, M. S. (2012) Trehalose glycolipids--synthesis and biological activities. *Carbohydrate Research* **356**, 25-36

92. Daffe, M., Lacave, C., Laneelle, M. A., Gillois, M., and Laneelle, G. (1988) Polyphthienoyl Trehalose, Glycolipids Specific for Virulent-Strains of the Tubercle Bacillus. *European Journal Biochemistry* **172**, 579-584

93. Ortalo-Magné, A., Lemassu, A., Lanéelle, M. A., Bardou, F., Silve, G., Gounon, P., Marchal, G., and Daffé, M. (1996) Identification of the surface-exposed lipids on the cell envelopes of Mycobacterium tuberculosis and other mycobacterial species. *Journal of Bacteriology* **178**, 456-461

94. Dubey, V. S., Sirakova, T. D., and Kolattukudy, P. E. (2002) Disruption of msl3 abolishes the synthesis of mycolipanoic and mycolipenic acids required for polyacyltrehalose synthesis in Mycobacterium tuberculosis H37Rv and causes cell aggregation. *Molecular Microbiology* **45**, 1451-1459

95. Rousseau, C., Neyrolles, O., Bordat, Y., Giroux, S., Sirakova, T. D., Prevost, M.-C., Kolattukudy, P. E., Gicquel, B., and Jackson, M. (2003) Deficiency in mycolipenate- and mycosanoate-derived acyltrehaloses enhances early interactions of Mycobacterium tuberculosis with host cells. *Cellular Microbiology* **5**, 405-415

96. Tortola, M. T., Laneelle, M. A., and Martin-Casabona, N. (1996) Comparison of two 2,3-diacyl trehalose antigens from Mycobacterium tuberculosis and Mycobacterium fortuitum for serology in tuberculosis patients. *Clinical and Diagnostic Laboratory Immunology* **3**, 563-566

97. Muñoz, M., Lanéelle, M.-A., Luquin, M., Torrelles, J., Julián, E., Ausina, V., and Daffé, M. (1997) Occurrence of an antigenic triacyl trehalose in clinical isolates and reference strains of Mycobacterium tuberculosis. *FEMS Microbiology Letters* **157**, 251-259

98. Julian, E., Matas, L., Alcaide, J., and Luquin, M. (2004) Comparison of antibody responses to a potential combination of specific glycolipids and proteins for test sensitivity improvement in tuberculosis serodiagnosis. *Clinical and Diagnostic Laboratory Immunology* **11**, 70-76

99. Neyrolles, O., and Guilhot, C. (2011) Recent advances in deciphering the contribution of Mycobacterium tuberculosis lipids to pathogenesis. *Tuberculosis* **91**, 187-195

100. Besra, G. S., Bolton, R. C., McNeil, M. R., Ridell, M., Simpson, K. E., Glushka, J., van Halbeek, H., Brennan, P. J., and Minnikin, D. E. (1992) Structural elucidation of a novel family of acyltrehaloses from Mycobacterium tuberculosis. *Biochemistry* **31**, 9832-9837

101. MINNIKIN, D. E., DOBSON, G., SESARDIC, D., and RIDELL, M. (1985) Mycolipenates and Mycolipanolates of Trehalose from Mycobacterium tuberculosis. *Journal of General Microbiology* **131**, 1369-1374

102. Hatzios, S. K., Schelle, M. W., Holsclaw, C. M., Behrens, C. R., Botyanszki, Z., Lin, F. L., Carlson, B. L., Kumar, P., Leary, J. A., and Bertozzi, C. R. (2009) PapA3 is an acyltransferase required for Polyacyltrehalose biosynthesis in Mycobacterium tuberculosis. *Journal of Biological Chemistry* **284**, 12745-12751

103. Kumar, P., Schelle, M. W., Jain, M., Lin, F. L., Petzold, C. J., Leavell, M. D., Leary, J. A., Cox, J. S., and Bertozzi, C. R. (2007) PapA1 and PapA2 are acyltransferases essential for the biosynthesis of *Mycobacterium tuberculosis* virulence factor Sulfolipid-1. *Procedings of the National Academy of Science U. S. A.* **104**, 11221-11226

104. Seeliger, J. C., Holsclaw, C. M., Schelle, M. W., Botyanszki, Z., Gilmore, S. A., Tully, S. E., Niederweis, M., Cravatt, B. F., Leary, J. A., and Bertozzi, C. R. (2012) Elucidation and Chemical Modulation of Sulfolipid-1 Biosynthesis in Mycobacterium tuberculosis. *Journal of Biological Chemistry* **287**, 7990-8000

105. Belisle, J. T., Vissa, V. D., Sievert, T., Takayama, K., Brennan, P. J., and Besra, G. S. (1997) Role of the major antigen of Mycobacterium tuberculosis in cell wall biogenesis. *Science* **276**, 1420-1422

106. Singh, A., Crossman, D. K., Mai, D., Guidry, L., Voskuil, M. I., Renfrow, M. B., and Steyn, A. J. C. (2009) *Mycobacterium tuberculosis* WhiB3 Maintains Redox Homeostasis by Regulating Virulence Lipid Anabolism to Modulate Macrophage Response. *PLoS Pathogens* **5**, e1000545

107. Walters, S. B., Dubnau, E., Kolesnikova, I., Laval, F., Daffe, M., and Smith, I. (2006) The Mycobacterium tuberculosis PhoPR two-component system regulates genes essential for virulence and complex lipid biosynthesis. *Molecular Microbiology* **60**, 312-330

108. Hatzios, S. K., Schelle, M. W., Holsclaw, C. M., Behrens, C. R., Botyanszki, Z., Lin, F. L., Carlson, B. L., Kumar, P., Leary, J. A., and Bertozzi, C. R. (2009) PapA3 Is an Acyltransferase Required for Polyacyltrehalose Biosynthesis in Mycobacterium tuberculosis. *Journal of Biological Chemistry* **284**, 12745-12751

109. Stover, C. K., de la Cruz, V. F., Fuerst, T. R., Burlein, J. E., Benson, L. A., Bennett, L. T., Bansal, G. P., Young, J. F., Lee, M. H., Hatfull, G. F., Snapper, S. B., Barletta, R. G., Jacobs, W. R., and Bloom, B. R. (1991) New use of BCG for recombinant vaccines. *Nature* **351**, 456-460

110. Stover, C. K., Bansal, G. P., Hanson, M. S., Burlein, J. E., Palaszynski, S. R., Young, J. F., Koenig, S., Young, D. B., Sadziene, A., and Barbour, A. G. (1993) Protective immunity elicited by recombinant bacille Calmette-Guerin (BCG) expressing outer surface protein A (OspA) lipoprotein: a candidate Lyme disease vaccine. Journal of Experimental Medicine **178**, 197-209

111. Stanley, S. A., Raghavan, S., Hwang, W. W., and Cox, J. S. (2003) Acute infection and macrophage subversion by Mycobacterium tuberculosis require a specialized secretion system. *Procedings of the National Academy of Science U. S. A.* **100**, 13001-13006

112. Seeliger, J. C., Topp, S., Sogi, K. M., Previti, M. L., Gallivan, J. P., and Bertozzi, C. R. (2012) A Riboswitch-Based Inducible Gene Expression System for Mycobacteria. *PLoS ONE* **7**, e29266

113. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. L. (2001) Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *Journal of General Microbiology* **305**, 567-580

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

115. Kelley, L. A., and Sternberg, M. J. E. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols* **4**, 363-371

116. Bardarov, S., Bardarov, S., Jr., Pavelka, M. S., Jr., Sambandamurthy, V., Larsen, M., Tufariello, J., Chan, J., Hatfull, G., and Jacobs, W. R., Jr. (2002) Specialized transduction: an efficient method for generating marked and unmarked targeted gene disruptions in *Mycobacterium tuberculosis*, *M. bovis* BCG and *M. smegmatis*. *Microbiology* **148**, 3007-3017

117. Harth, G., and Horwitz, M. A. (1997) Expression and Efficient Export of Enzymatically Active Mycobacterium tuberculosis Glutamine Synthetase in Mycobacterium smegmatis and Evidence That the Information for Export is Contained within the Protein. *Journal of Biological Chemistry* **272**, 22728-22735

118. Makarov, A., Denisov, E., Kholomeev, A., Balschun, W., Lange, O., Strupat, K., and Horning, S. (2006) Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Analytical Chemistry* **78**, 2113-2120

119. Asensio, J. G., Maia, C., Ferrer, N. L., Barilone, N., Laval, F., Soto, C. Y., Winter, N., Daffe, M., Gicquel, B., Martin, C., and Jackson, M. (2006) The Virulence-associated Two-component PhoP-PhoR System Controls the Biosynthesis of Polyketide-derived Lipids in Mycobacterium tuberculosis. *Journal of Biological Chemistry* **281**, 1313-1316

120. Hillmann, D., Eschenbacher, I., Thiel, A., and Niederweis, M. (2007) Expression of the Major Porin Gene mspA Is Regulated in Mycobacterium smegmatis. *Journal Bacteriology* **189**, 958-967

121. Griffin, T. J. t., Parsons, L., Leschziner, A. E., DeVost, J., Derbyshire, K. M., and Grindley, N. D. (1999) In vitro transposition of Tn552: a tool for DNA sequencing and mutagenesis. *Nucleic Acids Research* **27**, 3859-3865

122. Patricelli, M. P., Giang, D. K., Stamp, L. M., and Burbaum, J. J. (2001) Direct visualization of serine hydrolase activities in complex proteomes using fluorescent active site-directed probes. *Proteomics* **1**, 1067-1071

123. Layre, E., Sweet, L., Hong, S., Madigan, Cressida A., Desjardins, D., Young, David C., Cheng, T.-Y., Annand, John W., Kim, K., Shamputa, Isdore C., McConnell, Matthew J., Debono, C. A., Behar, S. M., Minnaard, Adriaan J., Murray, M., Barry III, Clifton E.,

Matsunaga, I., and Moody, D. B. (2011) A Comparative Lipidomics Platform for Chemotaxonomic Analysis of Mycobacterium tuberculosis. *Chemical Biology* **18**, 1537-1549

124. Sultana, R., Tanneeru, K., and Guruprasad, L. (2011) The PE-PPE Domain in Mycobacterium Reveals a Serine α/β Hydrolase Fold and Function: An *In Silico* Analysis. *PLoS ONE* **6**, e16745

125. Domenech, P., Reed, M. B., Dowd, C. S., Manca, C., Kaplan, G., and Barry, C. E., III. (2004) The Role of MmpL8 in Sulfatide Biogenesis and Virulence of Mycobacterium tuberculosis. *Journal of Biological Chemistry* **279**, 21257-21265

126. Converse, S. E., Mougous, J. D., Leavell, M. D., Leary, J. A., Bertozzi, C. R., and Cox, J. S. (2003) MmpL8 is required for sulfolipid-1 biosynthesis and Mycobacterium tuberculosis virulence. *Procedings of the National Academy of Science* **100**, 6121-6126

127. Tahlan, K., Wilson, R., Kastrinsky, D. B., Arora, K., Nair, V., Fischer, E., Barnes, S. W., Walker, J. R., Alland, D., Barry, C. E., and Boshoff, H. I. (2012) SQ109 Targets MmpL3, a Membrane Transporter of Trehalose Monomycolate Involved in Mycolic Acid Donation to the Cell Wall Core of Mycobacterium tuberculosis. *Antimicrobial Agents and Chemotherapy* **56**, 1797-1809

128. Grzegorzewicz, A. E., Pham, H., Gundi, V. A., Scherman, M. S., North, E. J., Hess, T., Jones, V., Gruppo, V., Born, S. E., Kordulakova, J., Chavadi, S. S., Morisseau, C., Lenaerts, A. J., Lee, R. E., McNeil, M. R., and Jackson, M. (2012) Inhibition of mycolic acid transport across the Mycobacterium tuberculosis plasma membrane. *Nature Chemical Biology* **8**, 334-341

129. Sassetti, C. M., and Rubin, E. J. (2003) Genetic requirements for mycobacterial survival during infection. *Antimicrobial Agents and Chemotherapy* **100**, 12989-12994

130. Passemar, C., Arbues, A., Malaga, W., Mercier, I., Moreau, F., Lepourry, L., Neyrolles, O., Guilhot, C., and Astarie-Dequeker, C. (2013) Multiple deletions in the polyketide synthase gene repertoire of Mycobacterium tuberculosis reveal functional overlap of cell envelope lipids in host-pathogen interactions. *Cell Microbiol*ogy

131. Quadri, L. E. (2014) Biosynthesis of mycobacterial lipids by polyketide synthases and beyond. *Critical Reviews in Biochemistry and Molecular Biology*

132. Crellin, P. K., Luo, C.-Y., and Morita, Y. S. (2013) Metabolism of Plasma Membrane Lipids in Mycobacteria and Corynebacteria. in *Lipid Metabolism* (Baez, R. V. ed.), InTech, Rijeka, Croatia. pp 119-148

133. Weidenmaier, C., and Peschel, A. (2008) Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nature Review Microbiology* **6**, 276-287

134. Silhavy, T. J., Kahne, D., and Walker, S. (2010) The Bacterial Cell Envelope. *Cold Spring Harbor Perspectives in Biology* **2**

135. Brennan, P. J., and Nikaido, H. (1995) The envelope of mycobacteria. *Annual Review Biochemistry* **64**, 29 - 63

136. Daffe, M., and Laneelle, M. A. (1988) Distribution of phthiocerol diester, phenolic mycosides and related compounds in mycobacteria. *Journal General Microbiology* **134**, 2049-2055

137. Rousseau, C., Winter, N., Pivert, E., Bordat, Y., Neyrolles, O., Avé, P., Huerre, M., Gicquel, B., and Jackson, M. (2004) Production of phthiocerol dimycocerosates protects Mycobacterium tuberculosis from the cidal activity of reactive nitrogen intermediates produced by macrophages and modulates the early immune response to infection. *Cell Microbiology* **6**, 277-287

138. Murry, J. P., Pandey, A. K., Sassetti, C. M., and Rubin, E. J. (2009) Phthiocerol Dimycocerosate Transport Is Required for Resisting Interferon-γ–Independent Immunity *Journal Infectious Disease* **200**, 774-782

139. Kirksey, M. A., Tischler, A. D., Siméone, R., Hisert, K. B., Uplekar, S., Guilhot, C., and McKinney, J. D. (2011) Spontaneous Phthiocerol Dimycocerosate-Deficient Variants of Mycobacterium tuberculosis Are Susceptible to Gamma Interferon-Mediated Immunity. *Infection and Immunity* **79**, 2829-2838

140. Cambier, C. J., Takaki, K. K., Larson, R. P., Hernandez, R. E., Tobin, D. M., Urdahl, K. B., Cosma, C. L., and Ramakrishnan, L. (2014) Mycobacteria manipulate macrophage recruitment through coordinated use of membrane lipids. *Nature* **505**, 218-222

141. Vergnolle, O., Chavadi, S. S., Edupuganti, U. R., Mohandas, P., Chan, C., Zeng, J., Kopylov, M., Angelo, N. G., Warren, J. D., Soll, C. E., and Quadri, L. E. (2015) Biosynthesis of mycobacterial cell-envelope-associated phenolic glycolipids in Mycobacterium marinum. *Journal Bacteriology*

142. Chavadi, S. S., Onwueme, K. C., Edupuganti, U. R., Jerome, J., Chatterjee, D., Soll, C. E., and Quadri, L. E. N. (2012) The mycobacterial acyltransferase PapA5 is required for biosynthesis of cell wall-associated phenolic glycolipids. *Microbiology* **158**, 1379-1387

143. Onwueme, K. C., Ferreras, J. A., Buglino, J., Lima, C. D., and Quadri, L. E. N. (2004) Mycobacterial polyketide-associated proteins are acyltransferases: Proof of principle with Mycobacterium tuberculosis PapA5. *Procedings of the National Academy of Science U. S. A.* **101**, 4608-4613

144. Buglino, J., Onwueme, K. C., Ferreras, J. A., Quadri, L. E. N., and Lima, C. D. (2004) Crystal Structure of PapA5, a Phthiocerol Dimycocerosyl Transferase from Mycobacterium tuberculosis. *Journal of Biological Chemistry* **279**, 30634-30642

145. Camacho, L. R., Constant, P., Raynaud, C., Laneelle, M. A., Triccas, J. A., Gicquel, B., Daffe, M., and Guilhot, C. (2001) Analysis of the phthiocerol dimycocerosate locus of Mycobacterium tuberculosis. Evidence that this lipid is involved in the cell wall permeability barrier. *Journal of Biological Chemistry* **276**, 19845-19854

146. Sulzenbacher, G., Canaan, S., Bordat, Y., Neyrolles, O., Stadthagen, G., Roig-Zamboni, V., Rauzier, J., Maurin, D., Laval, F., Daffé, M., Cambillau, C., Gicquel, B., Bourne, Y., and Jackson, M. (2006) LppX is a lipoprotein required for the translocation of phthiocerol dimycocerosates to the surface of *Mycobacterium tuberculosis*. *EMBO Journal* **25**, 1436-1444

147. Bhatt, K., Gurcha, S. S., Bhatt, A., Besra, G. S., and Jacobs, W. R. (2007) Two polyketide-synthase-associated acyltransferases are required for sulfolipid biosynthesis in Mycobacterium tuberculosis. *Microbiology* **153**, 513-520

148. Jain, M., Petzold, C. J., Schelle, M. W., Leavell, M. D., Mougous, J. D., Bertozzi, C. R., Leary, J. A., and Cox, J. S. (2007) Lipidomics reveals control of Mycobacterium tuberculosis virulence lipids via metabolic coupling. *Procedings of the National Academy of Science U. S. A.* **104**, 5133-5138

149. Yang, X., Nesbitt, N. M., Dubnau, E., Smith, I., and Sampson, N. S. (2009) Cholesterol Metabolism Increases the Metabolic Pool of Propionate in Mycobacterium tuberculosis. *Biochemistry* **48**, 3819-3821

150. Shi, L., Sohaskey, C. D., Kana, B. D., Dawes, S., North, R. J., Mizrahi, V., and Gennaro, M. L. (2005) Changes in energy metabolism of Mycobacterium tuberculosis in mouse lung and under in vitro conditions affecting aerobic respiration. *Procedings of the National Academy of Science U. S. A.* **102**, 15629-15634

151. Rohde, K. H., Veiga, D. F., Caldwell, S., Balazsi, G., and Russell, D. G. (2012) Linking the transcriptional profiles and the physiological states of Mycobacterium tuberculosis during an extended intracellular infection. *PLoS Pathogens* **8**, e1002769

152. Gomez-Velasco, A., Bach, H., Rana, A. K., Cox, L. R., Bhatt, A., Besra, G. S., and Av-Gay, Y. (2013) Disruption of the serine/threonine protein kinase H affects phthiocerol dimycocerosates synthesis in Mycobacterium tuberculosis. *Microbiology* **159**, 726-736

153. Perez, J., Garcia, R., Bach, H., de Waard, J. H., Jacobs, W. R., Jr., Av-Gay, Y., Bubis, J., and Takiff, H. E. (2006) Mycobacterium tuberculosis transporter MmpL7 is a potential substrate for kinase PknD. *Biochemical and Biophysical Research Communications* **348**, 6-12

154. Gupta, M., Sajid, A., Arora, G., Tandon, V., and Singh, Y. (2009) Forkhead-associated domain-containing protein Rv0019c and polyketide-associated protein PapA5, from substrates of serine/threonine protein kinase PknB to interacting proteins of Mycobacterium tuberculosis. *Journal of Biological Chemistry***284**, 34723-34734

155. Prisic, S., Dankwa, S., Schwartz, D., Chou, M. F., Locasale, J. W., Kang, C. M., Bemis, G., Church, G. M., Steen, H., and Husson, R. N. (2010) Extensive phosphorylation with overlapping specificity by Mycobacterium tuberculosis serine/threonine protein kinases. *Procedings of the National Academy of Science* **107**, 7521-7526

156. Fortuin, S., Tomazella, G. G., Nagaraj, N., Sampson, S. L., Gey van Pittius, N. C., Soares, N. C., Wiker, H. G., de Souza, G. A., and Warren, R. M. (2015) Phosphoproteomics analysis of a clinical Mycobacterium tuberculosis Beijing isolate: expanding the mycobacterial phosphoproteome catalog. *Frontiers in Microbiology* **6**, 6

157. Molle, V., Gulten, G., Vilcheze, C., Veyron-Churlet, R., Zanella-Cleon, I., Sacchettini, J. C., Jacobs, W. R., Jr., and Kremer, L. (2010) Phosphorylation of InhA inhibits mycolic acid biosynthesis and growth of Mycobacterium tuberculosis. *Molecular Microbiology* **78**, 1591-1605

158. Corrales, R. M., Molle, V., Leiba, J., Mourey, L., de Chastellier, C., and Kremer, L. (2012) Phosphorylation of mycobacterial PcaA inhibits mycolic acid cyclopropanation: consequences for intracellular survival and for phagosome maturation block. *Journal of Biological Chemistry* **287**, 26187-26199

159. Vilcheze, C., Molle, V., Carrere-Kremer, S., Leiba, J., Mourey, L., Shenai, S., Baronian, G., Tufariello, J., Hartman, T., Veyron-Churlet, R., Trivelli, X., Tiwari, S., Weinrick, B., Alland, D., Guerardel, Y., Jacobs, W. R., Jr., and Kremer, L. (2014) Phosphorylation of KasB Regulates Virulence and Acid-Fastness in Mycobacterium tuberculosis. *PLoS Pathogens* **10**, e1004115

160. Huang, G. F., and Hollingsworth, R. I. (1998) An efficient synthesis of (R)-3-hydroxytetradecanoic acid. *Tetrahedron Asymmetry* **9**, 4113-4115

161. Larcheveque, M., and Henrot, S. (1990) Enantiomerically Pure Beta,Gamma-Epoxyesters from Beta-Hydroxylactones - Synthesis of Beta-Hydroxyesters and (-)-GABOB. *Tetrahedron* **46**, 4277-4282

162. Roche, C., Labeeuw, O., Haddad, M., Ayad, T., Genet, J. P., Ratovelomanana-Vidal, V., and Phansavath, P. (2009) Synthesis of anti-1,3-Diols through RuCl3/PPh3-Mediated Hydrogenation of beta-Hydroxy Ketones: An Alternative to Organoboron Reagents. *European Journal of Organic Chemistry*, 3977-3986

163. Zorb, A., and Bruckner, R. (2010) Conversion of Conjugated Enones into Enantiomerically Pure beta-Hydroxy Ketones or 1,3-Diols - Samarium(II) Bromide Reductions of Protected alpha,beta-Dihydroxy Ketones. *European Journal of Organic Chemistry*, 4785-4801

164. Evans, D. A., and Chapman, K. T. (1986) The Directed Reduction of Beta-Hydroxy Ketones Employing Me4NHB(OAc)3. *Tetrahedron Letters* **27**, 5939-5942

165. Casas-Arce, E., ter Horst, B., Feringa, B., and Minnaard, A. (2008) Asymmetric Total Synthesis of PDIM A: A Virulence Factor of Mycobacterium tuberculosis. *Chem.istry* **14**, 4157-4159

166. ter Horst, B., Feringa, B. L., and Minnaard, A. J. (2007) Catalytic Asymmetric Synthesis of Phthioceranic Acid, a Heptamethyl-Branched Acid from Mycobacterium tuberculosis. *Organic Letters* **9**, 3013-3015

167. Su, C. Y., London, E., and Sampson, N. S. (2013) Mapping peptide thiol accessibility in membranes using a quaternary ammonium isotope-coded mass tag (ICMT). *Bioconjugate Chemistry* **24**, 1235-1247

168. Gao, X., Zhang, Q., Meng, D., Isaac, G., Zhao, R., Fillmore, T., Chu, R., Zhou, J., Tang, K., Hu, Z., Moore, R., Smith, R., Katze, M., and Metz, T. (2012) A reversed-phase capillary ultra-performance liquid chromatography–mass spectrometry (UPLC-MS) method for comprehensive top-down/bottom-up lipid profiling. *Analytical and Bioanalytical Chemistry* **402**, 2923-2933

169. Baer, C. E., Iavarone, A. T., Alber, T., and Sassetti, C. M. (2014) Biochemical and spatial coincidence in the provisional Ser/Thr protein kinase interaction network of Mycobacterium tuberculosis. *Journal of Biological Chemistry***289**, 20422-20433

170. Evans, D. A., Chapman, K. T., and Carreira, E. M. (1988) Directed Reduction of Beta-Hydroxy Ketones Employing Tetramethylammonium Triacetoxyborohydride. *Journal of the American Chemical Society* **110**, 3560-3578

171. Grundner, C., Gay, L. M., and Alber, T. (2005) Mycobacterium tuberculosis serine/threonine kinases PknB, PknD, PknE, and PknF phosphorylate multiple FHA domains. *Protein Science* **14**, 1918-1921

172. Sharma, K., Chandra, H., Gupta, P. K., Pathak, M., Narayan, A., Meena, L. S., D'Souza, R. C., Chopra, P., Ramachandran, S., and Singh, Y. (2004) PknH, a transmembrane Hank's type serine/threonine kinase from Mycobacterium tuberculosis is differentially expressed under stress conditions. *FEMS Microbiology Letters* **233**, 107-113

173. Hsiao, Y. S., Jogl, G., and Tong, L. (2004) Structural and biochemical studies of the substrate selectivity of carnitine acetyltransferase. *Journal of Biological Chemistry* **279**, 31584-31589

174. Mattevi, A. (2006) A monotopic membrane protein goes solo. *Structure* **14**, 628-629

175. Rufer, A. C., Lomize, A., Benz, J., Chomienne, O., Thoma, R., and Hennig, M. (2007) Carnitine palmitoyltransferase 2: analysis of membrane association and complex structure with a substrate analog. *FEBS Letters* **581**, 3247-3252

176. Perspicace, S., Rufer, A. C., Thoma, R., Mueller, F., Hennig, M., Ceccarelli, S., Schulz-Gasch, T., and Seelig, J. (2013) Isothermal titration calorimetry with micelles: Thermodynamics of inhibitor binding to carnitine palmitoyltransferase 2 membrane protein. *FEBS Open Bio* **3**, 204-211

177. Carel, C., Nukdee, K., Cantaloube, S., Bonne, M., Diagne, C. T., Laval, F., Daffe, M., and Zerbib, D. (2014) Mycobacterium tuberculosis proteins involved in mycolic acid synthesis and transport localize dynamically to the old growing pole and septum. *PLoS One* **9**, e97148

178. Meniche, X., Otten, R., Siegrist, M. S., Baer, C. E., Murphy, K. C., Bertozzi, C. R., and Sassetti, C. M. (2014) Subpolar addition of new cell wall is directed by DivIVA in mycobacteria. *Procedings of the National Academy of Science* **111**, E3243-3251

179. Neduva, V., and Russell, R. B. (2005) Linear motifs: evolutionary interaction switches. *FEBS Letters* **579**, 3342-3345

180. Jain, M., and Cox, J. S. (2005) Interaction between polyketide synthase and transporter suggests coupled synthesis and export of virulence lipid in M. tuberculosis. *PLoS Pathogens* **1**, e2

181. Rao, A., and Ranganathan, A. (2004) Interaction studies on proteins encoded by the phthiocerol dimycocerosate locus of Mycobacterium tuberculosis. *Molecular Genetics and Genomics* **272**, 571-579

182. Trivedi, O. A., Arora, P., Vats, A., Ansari, M. Z., Tickoo, R., Sridharan, V., Mohanty, D., and Gokhale, R. S. (2005) Dissecting the Mechanism and Assembly of a Complex Virulence Mycobacterial Lipid. *Molecular Cell* **17**, 631-643

183. Thelen, M., Rosen, A., Nairn, A. C., and Aderem, A. (1991) Regulation by phosphorylation of reversible association of a myristoylated protein kinase C substrate with the plasma membrane. *Nature* **351**, 320-322

184. Arora, P., Vats, A., Saxena, P., Mohanty, D., and Gokhale, R. S. (2005) Promiscuous Fatty Acyl CoA Ligases Produce Acyl-CoA and Acyl-SNAC Precursors for Polyketide Biosynthesis. *Journal of the American Chemical Society* **127**, 9388-9389

185. Flitsch, S. L., Goodridge, D. M., Guilbert, B., Revers, L., Webberley, M. C., and Wilson, I. B. H. (1994) The chemoenzymatic synthesis of neoglycolipids and lipid-linked oligosaccharides using glycosyltransferases. *Bioorganic and Medicinal Chemistry* **2**, 1243-1250

186. Yu, H., Huang, S., Chokhawala, H., Sun, M., Zheng, H., and Chen, X. (2006) Highly Efficient Chemoenzymatic Synthesis of Naturally Occurring and Non-Natural α-2,6-Linked Sialosides: A P. damsela α-2,6-Sialyltransferase with Extremely Flexible Donor–Substrate Specificity. *Angewandte Chemie International Edition* **45**, 3938-3944

187. Zhang, J., Kowal, P., Fang, J., Andreana, P., and Wang, P. G. (2002) Efficient chemoenzymatic synthesis of globotriose and its derivatives with a recombinant α-(1→4)-galactosyltransferase. *Carbohydrate Research* **337**, 969-976

188. Dubinsky, L., Jarosz, L. M., Amara, N., Krief, P., Kravchenko, V. V., Krom, B. P., and Meijler, M. M. (2009) Synthesis and validation of a probe to identify quorum sensing receptors. *Chemical Communications*, 7378-7380

189. Onwueme, K. C., Ferreras, J. A., Buglino, J., Lima, C. D., and Quadri, L. E. N. (2004) Mycobacterial polyketide-associated proteins are acyltransferases: Proof of principle withMycobacterium tuberculosisPapA5. *Proceedings of the National Academy of Sciences* **101**, 4608-4613

190. Kumar, P., Schelle, M. W., Jain, M., Lin, F. L., Petzold, C. J., Leavell, M. D., Leary, J. A., Cox, J. S., and Bertozzi, C. R. (2007) PapA1 and PapA2 are acyltransferases essential for the biosynthesis of the Mycobacterium tuberculosis virulence factor Sulfolipid-1. *Proceedings of the National Academy of Sciences* **104**, 11221-11226

191. Touchette, M. H., Holsclaw, C. M., Previti, M. L., Solomon, V. C., Leary, J. A., Bertozzi, C. R., and Seeliger, J. C. (2015) The rv1184c Locus Encodes Chp2, an Acyltransferase in Mycobacterium tuberculosis Polyacyltrehalose Lipid Biosynthesis. *Journal of Bacteriology* **197**, 201-210

192. Niphakis, M. J., Lum, K. M., Cognetta, A. B., 3rd, Correia, B. E., Ichu, T. A., Olucha, J., Brown, S. J., Kundu, S., Piscitelli, F., Rosen, H., and Cravatt, B. F. (2015) A Global Map of Lipid-Binding Proteins and Their Ligandability in Cells. *Cell* **161**, 1668-1680

193. Rao, V., Fujiwara, N., Porcelli, S. A., and Glickman, M. S. (2005) Mycobacterium tuberculosis controls host innate immune activation through cyclopropane modification of a glycolipid effector molecule. *Journal Experimental Medicine* **201**, 535-543

194. Camacho, L. R., Constant, P., Raynaud, C., Laneelle, M. A., Triccas, J. A., Gicquel, B., Daffe, M., and Guilhot, C. (2001) Analysis of the phthiocerol dimycocerosate locus of Mycobacterium tuberculosis. Evidence that this lipid is involved in the cell wall permeability barrier. *Journal Biological Chemistry* **276**, 19845-19854

195. Goren, M. B., Brokl, O., and Schaefer, W. B. (1974) Lipids of putative relevance to virulence in Mycobacterium tuberculosis: phthiocerol dimycocerosate and the attenuation indicator lipid. *Infection and Immunity* **9**, 150-158

196. Takeda, K., Miyatake, H., Yokota, N., Matsuyama, S., Tokuda, H., and Miki, K. (2003) Crystal structures of bacterial lipoprotein localization factors, LolA and LolB. *EMBO Journal* **22**, 3199-3209

197. Schroeder, F., Myers-Payne, S. C., Billheimer, J. T., and Wood, W. G. (1995) Probing the ligand binding sites of fatty acid and sterol carrier proteins: effects of ethanol. *Biochemistry* **34**, 11919-11927

198. Eftink, M. R. (1997) Fluorescence methods for studying equilibrium macromolecule-ligand interactions. *Methods in Enzymology* **278**, 221-257

199. Hawe, A., Sutter, M., and Jiskoot, W. (2008) Extrinsic fluorescent dyes as tools for protein characterization. *Pharmaceutical Research* **25**, 1487-1499

200. Mancek-Keber, M., and Jerala, R. (2006) Structural similarity between the hydrophobic fluorescent probe and lipid A as a ligand of MD-2. *FASEB journal* **20**, 1836-1842

201. Roehrl, M. H., Wang, J. Y., and Wagner, G. (2004) A general framework for development and data analysis of competitive high-throughput screens for small-molecule inhibitors of protein-protein interactions by fluorescence polarization. *Biochemistry* **43**, 16056-16066

202. Domenech, P., Reed, M. B., and Barry, C. E., 3rd. (2005) Contribution of the Mycobacterium tuberculosis MmpL protein family to virulence and drug resistance. *Infection and Immunity* **73**, 3492-3501

203. Bigi, F., Alito, A., Romano, M. I., Zumarraga, M., Caimi, K., and Cataldi, A. (2000) The gene encoding P27 lipoprotein and a putative antibiotic-resistance gene form an operon in Mycobacterium tuberculosis and Mycobacterium bovis. *Microbiology* **146**, 1011-1018

204. Gaur, R. L., Ren, K., Blumenthal, A., Bhamidi, S., Gonzalez-Nilo, F. D., Jackson, M., Zare, R. N., Ehrt, S., Ernst, J. D., and Banaei, N. (2014) LprG-mediated surface expression of lipoarabinomannan is essential for virulence of Mycobacterium tuberculosis. *PLoS Pathogens* **10**, e1004376

205. Martinot, A. J., Farrow, M., Bai, L., Layre, E., Cheng, T.Y., Tsai, J.H.C., Iqbal, J., Annand, J., Sullivan, Z., Hussain, M., Sacchettini, J., Moody, D.B., Seeliger, J., Rubin, E.J. (2015) Mycobacterial Metabolic Syndrome: LprG and Rv1410 Regulate Triacylglyceride Levels, Growth Rate and Virulence in Mycobacterium tuberculosis. *in review*

206. Nikaido, H. (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiology and Molecular Biology Reviews* **67**, 593-656

207. Fairman, J. W., Noinaj, N., and Buchanan, S. K. (2011) The structural biology of beta-barrel membrane proteins: a summary of recent reports. *Current Opinions in Structural Biology* **21**, 523-531

208. Putker, F., Bos, M. P., and Tommassen, J. (2015) Transport of lipopolysaccharide to the Gram-negative bacterial cell surface,

209. Okuda, S., and Tokuda, H. (2011) Lipoprotein Sorting in Bacteria. *Annual Review Microbiology* **65**, 239-259

210. Raetz, C. R., and Whitfield, C. (2002) Lipopolysaccharide endotoxins. *Annual Review Biochemistry* **71**, 635-700

211. Kalynych, S., Morona, R., and Cygler, M. (2014) Progress in understanding the assembly process of bacterial O-antigen. *FEMS Microbiology Reviews* **38**, 1048-1065

212. Sperandeo, P., Cescutti, R., Villa, R., Di Benedetto, C., Candia, D., Deho, G., and Polissi, A. (2007) Characterization of lptA and lptB, two essential genes implicated in lipopolysaccharide transport to the outer membrane of Escherichia coli. *Journal of Bacteriology* **189**, 244-253

213. Stenberg, F., Chovanec, P., Maslen, S. L., Robinson, C. V., Ilag, L. L., von Heijne, G., and Daley, D. O. (2005) Protein complexes of the Escherichia coli cell envelope. *Journal of Biological Chemistry* **280**, 34409-34419

214. Ruiz, N., Gronenberg, L. S., Kahne, D., and Silhavy, T. J. (2008) Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of Escherichia coli. *Procedings of the National Academy of Science* **105**, 5537-5542

215. Narita, S., and Tokuda, H. (2009) Biochemical characterization of an ABC transporter LptBFGC complex required for the outer membrane sorting of lipopolysaccharides. *FEBS Letters* **583**, 2160-2164

216. Tran, A. X., Dong, C., and Whitfield, C. (2010) Structure and functional analysis of LptC, a conserved membrane protein involved in the lipopolysaccharide export pathway in Escherichia coli. *Journal of Biological Chemistry* **285**, 33529-33539

217. Sestito, S. E., Sperandeo, P., Santambrogio, C., Ciaramelli, C., Calabrese, V., Rovati, G. E., Zambelloni, L., Grandori, R., Polissi, A., and Peri, F. (2014) Functional characterization of E. coli LptC: interaction with LPS and a synthetic ligand. *ChemBioChem* **15**, 734-742

218. Tran, A. X., Trent, M. S., and Whitfield, C. (2008) The LptA protein of Escherichia coli is a periplasmic lipid A-binding protein involved in the lipopolysaccharide export pathway. *Journal of Biological Chemistry* **283**, 20342-20349

219. Bos, M. P., Tefsen, B., Geurtsen, J., and Tommassen, J. (2004) Identification of an outer membrane protein required for the transport of lipopolysaccharide to the bacterial cell surface. *Procedings of the National Academy of Science* **101**, 9417-9422

220. Wu, T., McCandlish, A. C., Gronenberg, L. S., Chng, S. S., Silhavy, T. J., and Kahne, D. (2006) Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of Escherichia coli. *Procedings of the National Academy of Science* **103**, 11754-11759

221. Chng, S. S., Ruiz, N., Chimalakonda, G., Silhavy, T. J., and Kahne, D. (2010) Characterization of the two-protein complex in Escherichia coli responsible for lipopolysaccharide assembly at the outer membrane. *Procedings of the National Academy of Science* **107**, 5363-5368

222. Suits, M. D., Sperandeo, P., Deho, G., Polissi, A., and Jia, Z. (2008) Novel structure of the conserved gram-negative lipopolysaccharide transport protein A and mutagenesis analysis. *Journal of Molecular Biology* **380**, 476-488

223. Okuda, S., Freinkman, E., and Kahne, D. (2012) Cytoplasmic ATP hydrolysis powers transport of lipopolysaccharide across the periplasm in E. coli. *Science* **338**, 1214-1217

224. Qiao, S., Luo, Q., Zhao, Y., Zhang, X. C., and Huang, Y. (2014) Structural basis for lipopolysaccharide insertion in the bacterial outer membrane. *Nature* **511**, 108-111

225. Villa, R., Martorana, A. M., Okuda, S., Gourlay, L. J., Nardini, M., Sperandeo, P., Deho, G., Bolognesi, M., Kahne, D., and Polissi, A. (2013) The Escherichia coli Lpt transenvelope protein complex for lipopolysaccharide export is assembled via conserved structurally homologous domains. *Journal of Bacteriology* **195**, 1100-1108

226. Economou, A. (2002) Bacterial secretome: the assembly manual and operating instructions (Review). *Molecular Membrane Biology* **19**, 159-169

227. Driessen, A. J., and Nouwen, N. (2008) Protein translocation across the bacterial cytoplasmic membrane. *Annual Review Biochemistry* **77**, 643-667

228. Chatzi, K. E., Sardis, M. F., Karamanou, S., and Economou, A. (2013) Breaking on through to the other side: protein export through the bacterial Sec system. *The Biochemical Journal* **449**, 25-37

229. Palmer, T., and Berks, B. C. (2012) The twin-arginine translocation (Tat) protein export pathway. *Nature Reviews Microbiology* **10**, 483-496

230. Clavel, T., Germon, P., Vianney, A., Portalier, R., and Lazzaroni, J. C. (1998) TolB protein of Escherichia coli K-12 interacts with the outer membrane peptidoglycan-associated proteins Pal, Lpp and OmpA. *Molecular Microbiology* **29**, 359-367

231. Ehrmann, M., Ehrle, R., Hofmann, E., Boos, W., and Schlosser, A. (1998) The ABC maltose transporter. *Molecuar Microbiology* **29**, 685-694

232. Nikaido, H. (1998) Multiple antibiotic resistance and efflux. *Current Opinions in Microbiology* **1**, 516-523

233. Hayashi, S., and Wu, H. C. (1990) Lipoproteins in bacteria. *Journal of Bioenergetics and Biomembranes* **22**, 451-471

234. Tokuda, H., and Matsuyama, S. (2004) Sorting of lipoproteins to the outer membrane in E. coli. *Biochimica et Biophysica Acta* **1694**, IN1-9

235. Sankaran, K., and Wu, H. C. (1994) Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. *Journal of Biological Chemistry* **269**, 19701-19706

236. Babu, M. M., Priya, M. L., Selvan, A. T., Madera, M., Gough, J., Aravind, L., and Sankaran, K. (2006) A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *Journal of Bacteriology* **188**, 2761-2773

237. Sander, P., Rezwan, M., Walker, B., Rampini, S. K., Kroppenstedt, R. M., Ehlers, S., Keller, C., Keeble, J. R., Hagemeier, M., Colston, M. J., Springer, B., and Bottger, E. C. (2004) Lipoprotein processing is required for virulence of Mycobacterium tuberculosis. *Mol Microbiol* **52**, 1543-1552

238. Sutcliffe, I. C., and Russell, R. R. (1995) Lipoproteins of gram-positive bacteria. *Journal of Bacteriology* **177**, 1123-1128

239. Yamaguchi, K., Yu, F., and Inouye, M. (1988) A single amino acid determinant of the membrane localization of lipoproteins in E. coli. *Cell* **53**, 423-432

240. Gennity, J. M., and Inouye, M. (1991) The protein sequence responsible for lipoprotein membrane localization in Escherichia coli exhibits remarkable specificity. *Journal of Biological Chemistry* **266**, 16458-16464

241. Yakushi, T., Masuda, K., Narita, S., Matsuyama, S., and Tokuda, H. (2000) A new ABC transporter mediating the detachment of lipid-modified proteins from membranes. *Nature Cell Biology* **2**, 212-218

242. Yasuda, M., Iguchi-Yokoyama, A., Matsuyama, S., Tokuda, H., and Narita, S. (2009) Membrane topology and functional importance of the periplasmic region of ABC transporter LolCDE. *Bioscience, Biotechnology, and Biochemistry* **73**, 2310-2316

243. Matsuyama, S., Tajima, T., and Tokuda, H. (1995) A novel periplasmic carrier protein involved in the sorting and transport of Escherichia coli lipoproteins destined for the outer membrane. *EMBO Journal* **14**, 3365-3372

244. Tajima, T., Yokota, N., Matsuyama, S., and Tokuda, H. (1998) Genetic analyses of the in vivo function of LolA, a periplasmic chaperone involved in the outer membrane localization of Escherichia coli lipoproteins. *FEBS Letters* **439**, 51-54

245. Matsuyama, S., Yokota, N., and Tokuda, H. (1997) A novel outer membrane lipoprotein, LolB (HemM), involved in the LolA (p20)-dependent localization of lipoproteins to the outer membrane of Escherichia coli. *EMBO Journal* **16**, 6947-6955

246. Huang, B. X., and Kim, H. Y. (2013) Effective identification of Akt interacting proteins by two-step chemical crosslinking, co-immunoprecipitation and mass spectrometry. *PLoS ONE* **8**, e61430

247. Maguire, P. B., Briggs, F. N., Lennon, N. J., and Ohlendieck, K. (1997) Oligomerization Is an Intrinsic Property of Calsequestrin in Normal and Transformed Skeletal Muscle. *Biochemical and Biophysical Research Communications* **240**, 721-727

248. de Gunzburg, J., Riehl, R., and Weinberg, R. A. (1989) Identification of a protein associated with p21ras by chemical crosslinking. *Procedings of the National Academy of Science* **86**, 4007-4011

249. Tanaka, Y., Bond, M. R., and Kohler, J. J. (2008) Photocrosslinkers illuminate interactions in living cells. *Molecular BioSystems* **4**, 473-480

250. Young, T. S., Ahmad, I., Yin, J. A., and Schultz, P. G. (2010) An Enhanced System for Unnatural Amino Acid Mutagenesis in E. coli. *Journal of Molecular Biology* **395**, 361-374

251. Wang, F., Robbins, S., Guo, J., Shen, W., and Schultz, P. G. (2010) Genetic Incorporation of Unnatural Amino Acids into Proteins in <italic>Mycobacterium tuberculosis</italic>. *PLoS ONE* **5**, e9354

252. Freinkman, E., Okuda, S., Ruiz, N., and Kahne, D. (2012) Regulated Assembly of the Transenvelope Protein Complex Required for Lipopolysaccharide Export. *Biochemistry* **51**, 4800-4806

253. Okuda, S., and Tokuda, H. (2009) Model of mouth-to-mouth transfer of bacterial lipoproteins through inner membrane LolC, periplasmic LolA, and outer membrane LolB. *Proceedings of the National Academy of Sciences* **106**, 5877-5882

254. Allen, B. W. (1998) Mycobacteria. pp 15-30

255. Beste, D. J. V., Peters, J., Hooper, T., Avignone-Rossa, C., Bushell, M. E., and McFadden, J. (2005) Compiling a Molecular Inventory for Mycobacterium bovis BCG at Two Growth Rates: Evidence for Growth Rate-Mediated Regulation of Ribosome Biosynthesis and Lipid Metabolism. *Journal of Bacteriology* **187**, 1677-1684

256. Sambandamurthy, V. K., Derrick, S. C., Jalapathy, K. V., Chen, B., Russell, R. G., Morris, S. L., and Jacobs, W. R., Jr. (2005) Long-term protection against tuberculosis following vaccination with a severely attenuated double lysine and pantothenate auxotroph of Mycobacterium tuberculosis. *Infection and Immunity* **73**, 1196-1203

257. Sambandamurthy, V. K., Derrick, S. C., Hsu, T., Chen, B., Larsen, M. H., Jalapathy, K. V., Chen, M., Kim, J., Porcelli, S. A., Chan, J., Morris, S. L., and Jacobs, W. R., Jr. (2006) Mycobacterium tuberculosis DeltaRD1 DeltapanCD: a safe and limited replicating mutant strain that protects immunocompetent and immunocompromised mice against experimental tuberculosis. *Vaccine* **24**, 6309-6320

258. Lundgren, D. H., Hwang, S. I., Wu, L., and Han, D. K. (2010) Role of spectral counting in quantitative proteomics. *Expert Review of Proteomics* **7**, 39-53

259. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10**, 845-858

260. Constans, A. (2002) Protein Purification II: Affinity Tags. *The Scientist* **16**

261. de Souza, G. A., Arntzen, M. O., Fortuin, S., Schurch, A. C., Malen, H., McEvoy, C. R., van Soolingen, D., Thiede, B., Warren, R. M., and Wiker, H. G. (2011) Proteogenomic analysis of polymorphisms and gene annotation divergences in prokaryotes using a clustered mass spectrometry-friendly database. *Molecular and Cellular Proteomics* **10**, M110 002527

262. Bedouelle, H. (2000) Tyrosyl-tRNA Synthetases. *Madame Curie Bioscience Database*

263. Farrow, M. F., and Rubin, E. J. (2008) Function of a mycobacterial major facilitator superfamily pump requires a membrane-associated lipoprotein. *Journal of Bacteriology* **190**, 1783-1791

264. Stahl, C., Kubetzko, S., Kaps, I., Seeber, S., Engelhardt, H., and Niederweis, M. (2001) MspA provides the main hydrophilic pathway through the cell wall of Mycobacterium smegmatis. *Molecular Microbiology* **40**, 451-464

265. Faller, M., Niederweis, M., and Schulz, G. E. (2004) The structure of a mycobacterial outer-membrane channel. *Science* **303**, 1189-1192

266. Gago, G., Kurth, D., Diacovich, L., Tsai, S. C., and Gramajo, H. (2006) Biochemical and structural characterization of an essential acyl coenzyme A carboxylase from Mycobacterium tuberculosis. *Journal of Bacteriology* **188**, 477-486

267. Portevin, D., de Sousa-D'Auria, C., Montrozier, H., Houssin, C., Stella, A., Laneelle, M. A., Bardou, F., Guilhot, C., and Daffe, M. (2005) The acyl-AMP ligase FadD32 and AccD4-containing acyl-CoA carboxylase are required for the synthesis of mycolic acids

and essential for mycobacterial growth: identification of the carboxylation product and determination of the acyl-CoA carboxylase components. *Journal of Biological Chemistry* **280**, 8862-8874

268. Kremer, L., Maughan, W. N., Wilson, R. A., Dover, L. G., and Besra, G. S. (2002) The M. tuberculosis antigen 85 complex and mycolyltransferase activity. *Letters in Applied Microbiology* **34**, 233-237

269. Ronning, D. R., Vissa, V., Besra, G. S., Belisle, J. T., and Sacchettini, J. C. (2004) Mycobacterium tuberculosis antigen 85A and 85C structures confirm binding orientation and conserved substrate specificity. *Journal of Biological Chemistry* **279**, 36771-36777

270. Yuan, Y., Lee, R. E., Besra, G. S., Belisle, J. T., and Barry, C. E., 3rd. (1995) Identification of a gene involved in the biosynthesis of cyclopropanated mycolic acids in Mycobacterium tuberculosis. *Procedings of the National Academy of Science* **92**, 6630-6634

271. Gunawardena, H. P., Feltcher, M. E., Wrobel, J. A., Gu, S., Braunstein, M., and Chen, X. (2013) Comparison of the membrane proteome of virulent Mycobacterium tuberculosis and the attenuated Mycobacterium bovis BCG vaccine strain by label-free quantitative proteomics. *Journal of Proteome Research* **12**, 5463-5474

272. Mawuenyega, K. G., Forst, C. V., Dobos, K. M., Belisle, J. T., Chen, J., Bradbury, E. M., Bradbury, A. R., and Chen, X. (2005) Mycobacterium tuberculosis functional network analysis by global subcellular protein profiling. *Molecular Biology of the Cell* **16**, 396-404

273. Wolfe, L. M., Mahaffey, S. B., Kruh, N. A., and Dobos, K. M. (2010) Proteomic definition of the cell wall of Mycobacterium tuberculosis. *Journal of Proteome Research* **9**, 5816-5826

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

275. He, Z., and De Buck, J. (2010) Cell wall proteome analysis of Mycobacterium smegmatis strain MC2 155. *BMC Microbiology* **10**, 121

276. Liu, J., and Nikaido, H. (1999) A mutant of Mycobacterium smegmatis defective in the biosynthesis of mycolic acids accumulates meromycolates. *Procedings of the National Academy of Science* **96**, 4011-4016