TOLL-LIKE RECEPTOR 2-MEDIATED RECOGNITION OF MYCOBACTERIAL LIPOPROTEINS AND GLYCOLIPIDS

By

MICHAEL GERALD DRAGE

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Thesis Advisor: Dr. Clifford V. Harding, III
Thesis Co-advisor: Dr. W. Henry Boom

Department of Pathology: Immunology Training Program

CASE WESTERN RESERVE UNIVERSITY

August 2009
We hereby approve the dissertation of Michael Gerald Drage, candidate for the Ph.D. degree*.

(signed)

Eric Pearlman
(Chair of the committee)

Brian Cobb

Clive Hamlin

W. Henry Boom

Clifford V. Harding III

(date) 27 April 2009

*We also certify that written approval has been obtained for any proprietary material contained therein.
Dedication

To my better half and the parasite.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>List of Figures</strong></td>
<td>vi</td>
</tr>
<tr>
<td><strong>Acknowledgements</strong></td>
<td>ix</td>
</tr>
<tr>
<td><strong>List of Abbreviations</strong></td>
<td>x</td>
</tr>
<tr>
<td><strong>Abstract</strong></td>
<td>xi</td>
</tr>
<tr>
<td><strong>Chapter 1: Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td>Tuberculosis</td>
<td>2</td>
</tr>
<tr>
<td>Mycobacteria</td>
<td>2</td>
</tr>
<tr>
<td>TB treatment and prophylaxis</td>
<td>3</td>
</tr>
<tr>
<td>The course of mycobacterial infection</td>
<td>4</td>
</tr>
<tr>
<td>The host-pathogen balance</td>
<td>6</td>
</tr>
<tr>
<td>Survival strategies of mycobacteria</td>
<td>7</td>
</tr>
<tr>
<td>The Mtb cell envelope</td>
<td>9</td>
</tr>
<tr>
<td>PI-based glycolipids: structure and synthesis</td>
<td>12</td>
</tr>
<tr>
<td>Bacterial lipoproteins: structure and synthesis</td>
<td>15</td>
</tr>
<tr>
<td>Pattern recognition receptors</td>
<td>18</td>
</tr>
<tr>
<td>Toll-like receptors: structure and signaling</td>
<td>19</td>
</tr>
<tr>
<td>Long-term TLR signaling leads to down-regulation of specific immune functions</td>
<td>20</td>
</tr>
<tr>
<td>Toll-like receptor 2 is a sensor of lipoproteins and is important for immunity to mycobacteria</td>
<td>21</td>
</tr>
<tr>
<td>Co-receptors and accessory receptors of TLR2</td>
<td>23</td>
</tr>
<tr>
<td>Ligand structural determinants of TLR2-mediated detection: lipopeptides and lipoproteins</td>
<td>24</td>
</tr>
<tr>
<td>Ligand structural determinants of TLR2-mediated detection: glycolipids</td>
<td>25</td>
</tr>
<tr>
<td>Leucine rich repeats</td>
<td>27</td>
</tr>
<tr>
<td>Receptor structural determinants of TLR-mediated detection of PAMPs</td>
<td>27</td>
</tr>
</tbody>
</table>
Chapter 2: TLR2 and its co-receptors determine responses of macrophages and dendritic cells to lipoproteins of Mycobacterium tuberculosis

Abstract

Introduction

Materials and Methods

Results

Discussion

Chapter 3: Mycobacterium tuberculosis lipoprotein LprG (Rv1411c) is a glycolipid carrier and delivers triacylated glycolipids to Toll-like receptor 2

Abstract

Introduction

Materials and Methods

Results

Discussion

Chapter 4: Binding of mycobacterial lipoproteins by TLR2 and TLR1 extracellular domains

Introduction

Methods

Results

Discussion

Chapter 5: Discussion and future directions

Appendix 1: CD14 is important for detection of LprG and NA-LprG

Appendix 2: Proposed model of LprG:glycolipid detection by TLR2

Appendix 3: Correlation of thesis figures with experiments

Appendix 4: Publications

Works Cited
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Potential outcomes of exposure to M. tuberculosis</td>
<td>5</td>
</tr>
<tr>
<td>1.2</td>
<td>The TB tubercle is a granuloma</td>
<td>6</td>
</tr>
<tr>
<td>1.3</td>
<td>The cell envelope of mycobacteria</td>
<td>11</td>
</tr>
<tr>
<td>1.4</td>
<td>Structure and synthesis of PI-based glycolipids</td>
<td>13</td>
</tr>
<tr>
<td>1.5</td>
<td>Bacterial lipoprotein synthesis</td>
<td>17</td>
</tr>
<tr>
<td>1.6</td>
<td>The human Toll-like receptor family and some typical agonists</td>
<td>19</td>
</tr>
<tr>
<td>1.7</td>
<td>Common structural motifs of some TLR2 agonists</td>
<td>22</td>
</tr>
<tr>
<td>1.8</td>
<td>Structures of TLRs bound to their agonists demonstrate different</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>mechanisms of detection</td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>Mycobacterial lipoproteins are TLR2 agonists</td>
<td>47</td>
</tr>
<tr>
<td>2.2</td>
<td>TLR1 is more important than TLR6 for response to mycobacterial</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>lipoproteins</td>
<td></td>
</tr>
<tr>
<td>2.3</td>
<td>CD14 contributes to the detection of LpqH, LprA, and LprG, but</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>not PhoS1 or FSL-1</td>
<td></td>
</tr>
<tr>
<td>2.4</td>
<td>CD36 contributes to the detection of LprA, but not LpqH, LprG, or</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>PhoS1</td>
<td></td>
</tr>
<tr>
<td>2.5</td>
<td>Isolation and characterization of lung APCs by flow cytometry</td>
<td>53</td>
</tr>
<tr>
<td>2.6</td>
<td>Differing levels of TLR2 on lung APC subsets</td>
<td>55</td>
</tr>
<tr>
<td>3.1</td>
<td>NA-LprG carries a mycobacterial TLR2 agonist</td>
<td>75</td>
</tr>
<tr>
<td>3.S-1</td>
<td>TLR2 activity of NA-LprG/NA-LprA chimeric molecules</td>
<td>77</td>
</tr>
<tr>
<td>3.2</td>
<td>Crystal structure of NA-LprG reveals a hydrophobic pocket with the</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>potential to carry a TLR2 agonist</td>
<td></td>
</tr>
<tr>
<td>3.S-3</td>
<td>Comparison of LprG with LprA and LppX</td>
<td>83</td>
</tr>
<tr>
<td>3.3</td>
<td>Single amino acid change alters TLR2 agonist activity</td>
<td>86</td>
</tr>
<tr>
<td>3.4</td>
<td>NA-LprG activity is dependent on TLR1 and TLR2</td>
<td>88</td>
</tr>
</tbody>
</table>
3.5 Mycobacterial glycolipids are associated with NA-LprG
3.6 Triacylated Ac$_1$PIM$_2$ is specifically associated with NA-LprG
3.8 Total spectra of methanol eluates of NA-LprG and PabC
3.8-5 Collision-induced dissociation mass spectrometry of compounds detected in Figure 3.6
3.7 Structure of PIM bound to NA-LprG
3.8 NA-LprG binds purified mycobacterial glycolipids and facilitates their recognition by TLR2
3.9 The glycolipid binding site contributes significantly to the total TLR2 activity of acylated LprG
4.1 Enrichment for highly expressing cells by flow cytometry
4.2 FhT9XaFc is expressed but not secreted efficiently from HEK.293 cells
4.3 Comparison of expression of FhT2XaFc and the leucine zipper construct TLR2-Fos-FLAG (2FF)
4.4 Specificity of Factor Xa-mediated cleavage of FhT2XaFc
4.5 Pull-down of LpqH by FhT2XaFc captured on Protein A beads
4.6 Size exclusion binding assay of FhT2XaFc with LpqH in the absence or presence of 0.9 mM CaCl$_2$
4.7 Schemes of plate-based binding assays
4.8 Validation of plate coating and detection reagents
4.9 FhT2XaFc binds to the plate non-specifically
4.1 No detection of LprG binding to FhT2XaFc immobilized on plate
4.11 Non-specific binding of lipoproteins to the plate.
4.12 Association of protein A/G to CM5 chip at various pH values
4.13 Real-time analysis of LpqH interacting with FhT2XaFc captured on protein A/G immobilized to the chip
5.1 Structural rearrangements induced by V91W point mutation
5.2 Residues in LprA define a narrower pocket opening and conflict with the structure of Ac$_1$PIM$_2$
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3</td>
<td>Proposed binding site of Mannan chain of LM and of LAM to LprG surface</td>
<td>150</td>
</tr>
<tr>
<td>5.4</td>
<td>M. smegmatis expressing LprG-V91W has an altered cell envelope.</td>
<td>154</td>
</tr>
<tr>
<td>5.5</td>
<td>Model of sliding motility</td>
<td>155</td>
</tr>
<tr>
<td>5.6</td>
<td>Residues that define the hydrophobic pocket of TLR2 and TLR1 may be mutated</td>
<td>163</td>
</tr>
<tr>
<td></td>
<td>to prevent ligand binding</td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>CD14 is important for detection of LprG and NA-LprG.</td>
<td>165</td>
</tr>
<tr>
<td>A2</td>
<td>Proposed model of LprG:Glycolipid detection by TLR2.</td>
<td>166</td>
</tr>
<tr>
<td>A3</td>
<td>Correlation of thesis figures with experiments</td>
<td>167</td>
</tr>
<tr>
<td>A4</td>
<td>Publications</td>
<td>169</td>
</tr>
</tbody>
</table>
Acknowledgements

I have benefitted tremendously from working with such excellent and admirable colleagues. The combined group of the Harding and Boom laboratories is a highly interactive and dynamic group of people with numerous ongoing projects relating to basic immunology and infectious disease. It is uncommon to chance upon a group of people who, although working projects with little overlap, are willing and engaged to help each other trouble-shoot and generate ideas to move each project forward. I have much to thank my thesis committee for; their technical advice has come in handy on more than one occasion. More importantly, their urging has helped me to focus my efforts and has lent a pragmatic lens through which I will judge future projects. Of my committee members, Cliff and Henry deserve special gratitude for the incredible amount of time (and grant money) that they have invested in my development, including the design of experiments, analysis and interpretation of data, and the communication of the above to the rest of the world. Furthermore, a benefit of working with Cliff and Henry is that you have the opportunity to interact with people in various stages of career development. This not only benefits one as a trainee, but also as a trainer. While all of the above apply to everyone in the group, there are a few individuals who deserve special recognition. Nancy Nagy is the essential ingredient to the Harding lab, and has been particularly essential to my progress. Over the past two years, she has made hundreds of liters of medium in which I have cultured the M. smegmatis I needed to purify the proteins in the quantities I required. Aside from that and many lab-related things, Nancy has been a steadfast friend for both Nicole and I. Roxana Rojas has also been a tremendous asset for my work, both in pragmatic terms and in conceptual terms. She supported my use of the AKTA and helped me get it running again. Her intellectual contributions to the project, and her expertise regarding glycolipids were critical to the success of the LprG project. Xuedong Ding is an excellent colleague; his extensive experience in biochemistry was critical to the success of the LprG project. In their lab meetings and grant proposals, Roxana Rojas, David Canaday, Scott Fulton, Lakshmi Ramachandra, and Chris Lancioni have all helped me see more clearly the what lies ahead in the future stages of my career. I want to thank David Canaday in particular for his energy, enthusiasm, and his unfailing willingness to engage. I need also to thank Gareth, Yi, Daimon, Supriya, Ying, Ahmad, and Kitty for their input, friendship, and for making the lab a fun place to work.

Finally, special thanks are reserved for my amazing classmate, lab-mate, and life-mate, Nicole Pecora. Aside from her tremendous scientific input into these projects, her tireless support as my friend and wife have inspired me to be a better person.
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AM</td>
<td>Alveolar macrophage</td>
</tr>
<tr>
<td>BCG</td>
<td>Bacille calmette-guerin</td>
</tr>
<tr>
<td>BMM</td>
<td>Bone marrow-derived macrophage</td>
</tr>
<tr>
<td>DC</td>
<td>Dendritic cell</td>
</tr>
<tr>
<td>HEK293</td>
<td>Human embryonic kidney 293 cells</td>
</tr>
<tr>
<td>IFN</td>
<td>Interferon</td>
</tr>
<tr>
<td>LAM</td>
<td>Lipoarabinomman</td>
</tr>
<tr>
<td>LM</td>
<td>Lung macrophage</td>
</tr>
<tr>
<td>LM</td>
<td>Lipomannan</td>
</tr>
<tr>
<td>LPS</td>
<td>Lipopolysachharide</td>
</tr>
<tr>
<td>LRR</td>
<td>Leucine rich repeat</td>
</tr>
<tr>
<td>MHC2</td>
<td>Major histocompatibility complex 2</td>
</tr>
<tr>
<td>Mtb</td>
<td>Mycobacterium tuberculosis</td>
</tr>
<tr>
<td>NA</td>
<td>Non-acylated</td>
</tr>
<tr>
<td>NLRs</td>
<td>Nod-like receptors</td>
</tr>
<tr>
<td>Pam3CSK4</td>
<td>N-Palmitoyl-S-[2,3-bis(palmitoyloxy)-(2RS)-propyl]-[R]-cysteinyl-[S]-seryl-[S]-lysyl-[S]-lysyl-[S]-lysyl-[S]-lysine x 3HCl</td>
</tr>
<tr>
<td>PAMP</td>
<td>Pathogen associated molecular pattern</td>
</tr>
<tr>
<td>PDIMs</td>
<td>Phthiocerol dimycocerosates</td>
</tr>
<tr>
<td>PIMs</td>
<td>Phosphatidyl-my0-inositol mannosides</td>
</tr>
<tr>
<td>TLRs</td>
<td>Toll-like receptors</td>
</tr>
</tbody>
</table>
Toll-like Receptor 2-Mediated Recognition of Mycobacterial Lipoproteins and Glycolipids

by

MICHAEL GERALD DRAGE

Abstract

Mycobacterium tuberculosis has established infection in one third of the world population. Toll-like receptor 2 is important for control of mycobacterial infection, and is important for numerous immunological processes. Toll-like receptor 2 has an unusually diverse ligand repertoire, including lipoproteins, glycolipids, proteins, and glycans. The purpose of the studies described in this thesis was to better characterize the mechanisms of TLR2 activation by mycobacterial lipoproteins. In the first part, we examined co-receptor and accessory receptor contribution to TLR2-mediated detection of the mycobacterial lipoproteins LpqH, LprA, LprG, and PhoS1. We found that while most of the lipoproteins required TLR2, TLR1, and CD14, some did not require CD14, and LprA did not require TLR1. These results suggest that the protein component of bacterial lipoproteins contributes significantly to their ability to activate TLR2. Furthermore, in an effort to determine whether different TLR2 co- and accessory receptor requirements might impact mycobacterial infection, we analyzed murine lung antigen presenting cell subsets and found that expression of TLR2, CD14, and CD36 varied amongst them. This impacted their ability to respond to the mycobacterial lipoprotein LpqH. Lung macrophages expressed less TLR2 than alveolar macrophages or dendritic cells, and were significantly less responsive to LpqH than alveolar macrophages.
A second focus of this thesis was to describe the structural motifs responsible for the ability of LprG to signal independent of its covalently-attached acyl chains. Toward this end, we found that LprG carried glycolipids in a hydrophobic pocket and delivered them to TLR2 for recognition. A point mutation in the hydrophobic pocket abrogated glycolipid binding and TLR2 activation by the non-acylated LprG. We found that while LprA could carry diacylated glycolipids, only LprG carried triacylated glycolipids, and that these glycolipids signaled in a TLR2 and TLR1 dependent manner. Furthermore, we found that LprG increased the potency of glycolipids for TLR2 activation by several orders of magnitude, and that the glycolipids carried in the pocket of acylated LprG explain the difference between LprA and LprG. We propose that LprG functions to carry mycobacterial glycolipids for the purpose of envelope construction. Together, the studies described in this thesis demonstrate that TLR2 detects its agonists via multiple mechanisms, and include the description of a novel recognition mechanism by which TLR2 co-opts a bacterial glycolipid carrier as a ligand delivery system.
Chapter 1: Introduction
**Tuberculosis**

Tuberculosis (TB) has plagued humans for the past 40,000 years [1]. Currently, TB is pandemic; one third of the human population is infected with the etiologic agent, *Mycobacterium tuberculosis* (Mtb). Each year Mtb infects ~9 million people and causes approximately 2 million deaths. Tuberculosis is a major economic burden, causing the loss of 34.7 million disability-adjusted life years [2] [3]), predominantly in impoverished regions. Furthermore, the organism has shown increasing resistance to available antibiotics, necessitating the creation of a new category of drug resistance, termed “extensively drug resistant” (XDR)-TB [4]. There are now major initiatives (i.e. Bill and Melinda Gates Foundations [5], FIND [5], AERAS [5], TB Alliance [5], and the Global Fund for AIDS, Tuberculosis, and Malaria [6]) to improve diagnostics, develop an effective vaccine, and discover new chemotherapeutic agents. However, this problem includes challenges outside the realm of the biomedical sciences. While these efforts are encouraging, it is sobering to know that the advances they provide will likely only be pertinent if the healthcare infrastructure in TB endemic countries is improved.

**Mycobacteria**

Mycobacteria are members of the phylum Actinobacteria, gram-positive bacteria characterized by a high (>55%) genomic G + C content [7, 8]. Within Actinomyces, mycobacteria belong to a branch referred to as the Corynebacterium-Mycobacterium-Nocardia (CMN) branch. There are over eighty species of mycobacteria, most of which are nonpathogenic saprophytes [9]. Several species cause disease in humans.
*Mycobacterium leprae* and *Mycobacterium ulcerans* cause the disease leprosy and Buruli ulcers, respectively. This thesis focuses on *Mycobacterium tuberculosis* (Mtb), the flagship species of the Mtb complex. The Mtb complex includes six closely related species capable of causing TB, including *M. tuberculosis* (the most common agent of human TB), *M. africanum* (an agent of TB in sub-Saharan Africa), *M. bovis* (agent of TB in an array of mammals), *M. canettii* (a rarely encountered agent of human TB), and *M. microti* (the agent of TB in voles) [10]. Laboratory strains of Mtb referenced in this thesis include the virulent strain Mtb H37Rv and its avirulent derivative, Mtb H37Ra. Other mycobacterial species can cause TB-like disease in immunocompromised individuals, such as *M. kansasii* [11] and members of the *M. avium* complex [12].

*M. smegmatis* is an exclusively saprophytic species used extensively in the work described in this thesis. Because of its 3 h doubling time, ability to take up DNA, and its inability to cause disease, *M. smegmatis* (most commonly the strain MC²-155 [13]), is useful for mycobacterial molecular biology [14]. In this thesis, we utilize *M. smegmatis* MC²6/1-2c, a strain that is similarly amenable to manipulation [15].

**TB treatment and prophylaxis**

Mtb is susceptible to relatively few antibiotics, a trait ascribed to the impermeability of the mycobacterial cell envelope (discussed below). First-line drugs include isoniazid, ethambutol, pyrazinamide, and rifampin [11]. While these are cheap, effective drugs, they require months of administration to clear the infection [16]. Interrupted drug administration has led to the development of increasingly drug-resistant
strains of Mtb. While the use of fluoroquinolone adjunctive therapy may improve the efficacy of current drug regimens and shorten the required duration of therapy [17], widespread use of this drug for other infections has already led to increasing fluorquinolone resistance among Mtb isolates [18].

The only currently available vaccine for Mtb infection is the *M. bovis* bacille Calmette-Guerin (BCG). BCG was derived from a virulent *M. bovis* isolate by 230 serial passages in a broth containing glycerol, potato-extract and bile salts [10], leading to attenuation due to loss of genetic material [19]. BCG is used in newborns to prevent disseminated Mtb infection in young children [20]. However, this vaccine is not effective at preventing reactivation disease in adolescents and adults [21], which is the disease manifestation responsible for the vast majority of TB in the world. Because of this lack, there are currently several new vaccines in clinical trials, with vigorous discussion regarding which outcomes represent the best surrogate markers of protection [22, 23].

**The course of mycobacterial infection**

While Mtb typically causes chronic infection, the disease can follow a variety of courses (Fig 1.1). Although difficult to quantify, it is estimated that the majority (~70%) of exposures do not lead to successful infection. Of the people in whom infection is established, about 90% limit the disease
Figure 1.1. Potential outcomes of exposure to *M. tuberculosis*.

to granulomas within the lung and draining lymph nodes, successfully containing the infection and restricting inflammation-induced tissue damage to discrete foci (Fig. 1.2). Tumor necrosis factor alpha [24] and IL-12/Interferon-gamma (IFN-g) [25] signaling axes are important for granuloma formation and maintenance. Although the majority of people initially limit the disease to granulomas, some 10% of apparently immunocompetent individuals progress to primary disseminated disease, a phenomenon not well understood. Of the 90% that successfully limit the infection to granulomas, virtually none achieve sterilizing immunity, and about 5%-10% eventually develop secondary “reactivation” disease, with rupture of the granuloma, aerosolization of bacteria (with the potential to infect others), and often the development of extrapulmonary infection of the pleura, lymphatics, bone, meninges, genito-urinary system, and skin [26]. The phenomenon of reactivation is responsible for the majority of TB disease in the world. While most infected individuals have disease that is described as clinically “latent”, the situation is likely one of dynamic balance between host and pathogen [27].
The Host-pathogen balance

The balance between pathogen and host is established and maintained by interactions between the host innate immune cells (working in concert with other immune system cells) and their intracellular parasites, the Mtb bacilli. Infection typically occurs in the lung, requiring inhalation of only ~10 bacilli [11]. After sensing a bacterium with Pattern Recognition Receptors (PRRs) (discussed below), the alveolar macrophage engulfs the pathogen and translocates to the lung parenchyma, where bacteria and/or components thereof can be taken up by lung macrophages and dendritic cells. Dendritic cells migrate to the draining lymph node, where they produce IL-12, IL-18, and IL-23, increase surface expression of costimulatory molecules and initiate the adaptive immune response by priming naïve T cells. Meanwhile, the host macrophage in the lung produce
proinflammatory cytokines and induce intracellular bactericidal processes, including phagosomal acidification, free radical production, phago-lysosomal fusion, and autophagy [28, 29]. While most bacterial species are destroyed by these processes, some mycobacteria persist and replicate relatively unhindered. After priming, T cells travel to the lung, and upon stimulation by the appropriate antigen in the context of MHC, secrete IFN-gamma, a powerful enhancer of bactericidal functions crucial for controlling Mtb growth [30]. Development of the adaptive immune response correlates with a decreased bacterial load of several orders in magnitude, but does not achieve sterilizing immunity. The importance of T cell function in the control of Mtb infection is evidenced on a global scale by the synergy existing between the Mtb and HIV epidemics [31], and by the finding that deficiencies in IFN-gamma function predispose to infection by intracellular pathogens [25]. The question of how Mtb survives an immune response including vigorous adaptive immunity has been the subject of intensive research.

**Survival strategies of mycobacteria**

Numerous mechanisms have been proposed to explain the ability of mycobacteria to survive the immune response, including direct resistance to microbicidal free radicals, manipulation of phagosome trafficking, inhibition of antigen presentation, and escape from the phagosome. Mtb directly inhibits phagosomal acidification by producing ammonia [32]. To counter free radicals produced by the oxidative burst of phagocytic cells, Mtb has a cell envelope laden with free-radical scavengers [33], and secretes superoxide dismutase [34], both inactivating superoxide, and preventing host macrophage apoptosis and consequent cross-presentation to CD8 T cells [35]. The phosphatidyl-myo-
inositol (PI)-based glycolipids (discussed below) of its cell envelope delay phagosome acidification/maturation and prevent fusion with the lysosome [36, 37], and simultaneously enhance fusion with other early endosomes, presumably to ensure nutrient delivery [38]. Independent of the delay of phagosome maturation provided by glycolipids [39], Mtb lipoproteins LpqH (Rv3763, aka the 19 kD lipoprotein antigen) [40-43], LprG (Rv1411c, aka p27) [44], and LprA (Rv1270c) [45] inhibit antigen presentation. This inhibition is caused by long-term signaling by Toll-like receptor 2 (discussed below) and occurs by transcriptional repression of C2TA, the major regulator of MHC2 expression [46-48]. Recent work has demonstrated that inhibition of antigen presentation occurs in vivo. While uninfected macrophages in the diseased lung have increased MHC2, cells that harbor bacilli are decreased in surface MHC2, despite being in the same inflammatory environment [49, 50]. Pathogenic Mtb also signal through TLR2 in a non-canonical mechanism that broadly inhibits other TLR proinflammatory signaling [51]. Should these mechanisms fail and the mycobacterial phagosome become acidic, the secreted ESAT-6/CFP10 heterodimer can uncouple, allowing ESAT-6 to insert into the phagosomal membrane and promote its rupture [52], a finding that may relate to the contentious report that pathogenic mycobacteria can escape from the phagosome at later time points of infection [53]. By these and other mechanisms, Mtb controls its host cell, creating a niche within which it survives. Additionally, Mtb can also affect other cells of the immune system without being internalized by them. Mtb envelope constituents insert into host membranes [54], are released from the host cell [55] and can directly inhibit the function of other cells including T cells [56-58].

Although it is understood that states of immune compromise such as HIV
infection, corticosteroid use, anti-TNF therapy [59], aging, and drug abuse make reactivation disease more likely [60], the specific factors that determine the balance between these host effector mechanisms and Mtb survival are poorly understood. Since Mtb survives within cells of the innate immune system, it is intuitive that this interaction may be crucial to the establishment and maintenance of the balance between host and pathogen. This interaction is the main focus of the work described in this thesis, which focuses on two classes of molecules present in the mycobacterial envelope; the PI-based glycolipids and the lipoprototeins, and their interaction with TLR2 and its co-receptors and accessory receptors.

**The Mtb Cell Envelope**

The mycobacterial cell envelope (Fig. 1.3) is unusual in structure and remarkably impermeable. With the exception of mycoplasma, the cell envelopes of prokaryotes consist of at least two distinct entities: the inner plasma membrane, responsible for substrate transport, energy production, and the synthesis of macromolecules; and a rigid cell wall, which maintains cell shape and protects the cell against osmotic forces generated by the cytoplasm [61]. Gram-negative bacterial cell walls have an additional outer lipid bilayer. Despite being gram positive, mycobacteria have an outer lipid bilayer, and an electron-dense outer capsule rich in carbohydrate and protein [62]. Although similar in overall structure to gram-negative organisms, the mycobacterial cell envelope has important differences in chemical composition from those of other bacteria. While the lipid content of the cell envelope may represent 5% or 10% of the dry cell mass for gram-positive and gram-negative bacteria, respectively, 40% of the dry cell
mass of mycobacteria is derived from envelope lipids [63]. The lipid content and the cross-linking of the outer membrane to arabinogalactan (see below) are thought to be responsible for the impressive impermeability of the mycobacterial envelope, estimated to be 10 to 100-fold lower than that of the remarkably impermeable *P. aeruginosa* [64]. Structural rigidity is provided by a core of cross-linked peptidoglycan covalently linked to arabinogalactan. The arabinogalactan is linked via ester bonds to a characteristic assortment of long chain, alpha-alkyl, beta-hydroxy fatty acids known as the mycolic acids [65]. While mycolic acids are shared among all Actinomycetes, those of mycobacteria are exceptionally long, ranging from 60-90 carbon units, (compared to 32-50 for other Actinomycetes) [66], and are thought to be responsible for the enhanced acid-fast staining of mycobacteria [67]. Free lipids and glycolipids form an asymmetric bilayer with the mycolic acids. Of these, the phosphatidyl-(myo)-inositol (PI)-based glycolipids (discussed below) are present in tremendous quantities (i.e. PIMs represent 56% of all phospholipids in the cell wall, and 37% of all phospholipids in cytoplasmic membrane [68]). This scheme is supported by electron microscopy studies, which estimate the thickness of the outer and inner membranes to be 9-10 nm, and 4-4.5 nm, respectively [8].
Figure 1.3. The cell envelope of Mycobacteria. The cytoplasmic membrane is the location of glycolipid and lipoprotein synthesis. Lipoproteins and glycolipids are found both in the plasma membrane and the capsule-like layer. Peptidoglycan provides structure, while arabinoglycan is a permeability barrier to hydrophobic molecules. The asymmetric outer membrane represents a formidable permeability barrier to hydrophilic molecules. The outer “capsule-like” layer consists mostly of glycans and protein. The capsule layer may vary the most among mycobacterial species. This general scheme was originally proposed by Minnikin.
PI-based glycolipids: structure and synthesis

Glycolipids represent a major constituent of the mycobacterial envelope, synthesized by lipoproteins and other membrane-associated proteins. While mycobacteria produce several classes of glycolipids (i.e. phenolic glycolipids, sulfolipids, and “cord factors”), this thesis focuses on glycolipids that share the phosphatidyl-myoinositol (PI) anchor, termed PI-based glycolipids. Many of the genes involved in the biosynthesis of PI-based glycolipids are now known (Fig. 1.4). The current paradigm of mycobacterial glycolipid synthesis follows a predominantly linear pathway, PI, PIM, LM, LAM with each progressive step resulting in a more glycosylated species [69]. PI serves as the substrate for the alpha-mannosyltransferase pimA (Rv2610c), which transfers a mannopyranosyl residue from GDP-Mannose to the 2-position of PI to form PIM$_1$ [70]. While there appears to be some redundancy at the next mannosylation [71] pimB (Rv0557) transfers the second mannose residue [72] forming PIM$_2$, either before or after 6-O-acylation of the Manp residue linked to the 2-position of the inositol by the acyltransferase Rv2611c [73], creating Ac$_1$PIM$_2$. PimC (MT1800; identified in CDC1551) is thought to catalyze the addition of a third Manp residue to Ac$_1$PIM$_2$ to form Ac$_1$PIM$_3$, but is lacking in some Mtb isolates, and a M. bovis BCG knockout strain lacks perturbation of PIM, LM or LAM production, suggesting redundancy at this step [74]. Ac$_1$PIM$_3$ is further modified by an unidentified enzyme to make Ac$_1$PIM$_4$. As evidenced by studies utilizing amphomycin, a lipopeptide antibiotic that selectively inhibits polyprenyl-P-requiring synthases, the mannose donor for further mannosylation changes from cytosolic GDP-mannose to periplasmic C$_{35}$/C$_{50}$ polyprenyl monophosphatomanose (C$_{35}$/C$_{50}$-P-Manp) [75],[76]. Therefore, Ac$_1$PIM$_4$ must “flip” to the outer leaflet of the
Figure 1.4. Structure and synthesis of PI-based glycolipids. Structure of PIMs, LMs, and LAMs are illustrated. Enzymes with assigned functions for biosynthesis are labeled in red. Box illustrates structures that are produced using cytosolic GDP-mannose as substrate, occurring at the inner leaflet of plasma membrane. Dashed line demarcates the limit of structures within the PIMs family. PIM$_2$ and PIM$_6$ are by far the most abundant of the PIMs; the others are intermediates. LpqW is thought to be involved in flux at the only branch of the synthetic chain, the creation of the polar PIMs, (PIM$_5$ and PIM$_6$). The mannan chain of LM typically contains 20-26 residues (only 10 are illustrated). It is not known how or where the arabinan is attached to the mannan chain of LM. Adapted from the work of Nigou, Puzo, Chaterjee, Seeberger, Besra, and others.
plasma membrane before further mannosylation. The mechanism for this translocation is as yet unknown. When in the outer leaflet of the plasma membrane, Ac$_1$PIM$_4$ can either be shunted into the production of polar PIM species (PIM$_5$ and PIM$_6$) by the $\alpha(1\rightarrow2)$-mannosyltransferase $pimE$ (Rv1159) [77], or it can be further $\alpha(1\rightarrow6)$-mannosylated to form LMs. The lipoprotein $lpqW$ (Rv1116) appears to regulate the flux of Ac$_1$PIM$_4$ between polar PIMs and LM/LAM synthesis [78]. While the identity of the alpha 1-6 manp-transferase(s) responsible for the synthesis of the core mannan chain of LM is not known, Rv2181 is involved in the addition of alpha 1-2 Manp-linked branches of the mannan backbone in LM and LAM [79]. Mature LM is then glycosylated by the arabinofuranosyl transferase $embC$ (Rv3793) to form LAM [80]. In some mycobacterial species, LAM can be further modified. In fast-growing species such as $M$. smegmatis, LAM can be capped with PI to form PILAM [81]. In slow-growing mycobacteria including $M$. tuberculosis and $M$. leprae, LAM is modified by the addition of 1-3 mannose residues by MT1671 (Rv1635c) to form ManLam [82]. In $M$. cheloniae, LAM is left uncapped, and is called ara-LAM [83]. In $M$. tuberculosis and $M$. kansasii, ManLAM can be terminally modified with 5-deoxy-5-methylthio-xylofuranose (MTX) motif [84]. While knowledge of the biosynthetic enzymes responsible for synthesis of PI-based glycolipids is important, almost nothing is known about the assembly of the cell envelope. In chapters 3 and 5, we hypothesize that the lipoprotein LprG may be involved in transport of PI-based glycolipids for the purposes of cell envelope assembly, and as a consequence of this role, increase the bioavailability of glycolipids for recognition by TLR2 and potentially other receptors.
**Bacterial lipoproteins: structure and synthesis**

The Mtb cell envelope is rich in lipoproteins, which are synthesized as pre-prolipoproteins which undergo a series of post-translational modifications by a series of membrane-integral enzymes (Fig. 1.5) [85]. Prokaryotic lipoproteins are targeted to the cytoplasmic membrane by a type II signal peptide sequence, defined by the PROSITE accession number PS51257, with the additional feature of a lipobox consensus sequence [LVI][ASTVI][GAS][C] [86]. The sequence directs transport of the unfolded preprolipoprotein through the plasma membrane in a SecA-dependent manner [87] where a diacylglyceride unit is added via a thioether linkage to the universally conserved cysteine residue of the lipobox by the enzyme pre-prolipoprotein diglyceryl transferase (Lgt). The signal peptide is then cleaved by prolipoprotein signal peptidase (Lsp), immediately N-terminal to the conserved cysteine of the lipobox. In gram-negative bacteria, a third enzyme, apolipoprotein n-acyl transferase (Lnt) adds an additional lipid to the amino terminus of the N-terminal cysteine. Although it is generally presumed that mycobacterial lipoproteins are triacylated, the presence of the n-linked acyl chain has yet to be biochemically proven [88]. Like secretion, glycosylation of mycobacterial proteins also occurs in a SecA-dependent manner [89]. While lipoproteins are synthesized at the plasma membrane, some, such as LpqH, PhoS1, and LprA are found in the outer capsule-like layer of the envelope and/or in culture filtrate [43, 62, 90], supporting their bioavailability to host cell receptors such as the pattern recognition receptors discussed below. The work in this thesis focuses on LprA (Rv1270c), LprG (Rv1411c), LpqH (Rv3763), and PhoS1 (Rv0934).
While the physiological function for some of these lipoproteins is unknown, their importance for host-pathogen interactions has been the subject of numerous studies. PhoS1 (Rv0934, aka Psts1) is part of a four-gene operon involved in phosphate transport. Whereas most bacteria have a phosphate transporter system, mycobacteria have three such operons [91]. In terms of immunogenic activity, PhoS1 has been characterized as a TLR2 agonist with some cross-reactivity for TLR4 [92]. LpqH is a lipoprotein of unknown physiologic function, but has been the subject of numerous studies because of its immunogenic properties. However, when used in vaccination studies and infection models, LpqH appears to induce an immune response that is detrimental to control of infection [93] [94]. LprA and LprG are members of the LprA/G/F/X family [85], with 34% amino acid identity between them (Fig. 2.1) and 28-30% amino acid identity to the other members of the family, LprF [95] and LppX [96]. While the functions of LprA and LprG are unknown, LprG has been the subject of numerous immunological studies. LprG was initially described as a major antigen in the sera of *M. bovis*-infected cattle [97], suggesting that in the context of infection, it was highly immunogenic and a potential vaccine candidate. Further investigation found LprG to be part of an operon with Rv1410c (aka p55)[98], a putative proton drug antiporter [99]. While knockout of the LprG operon in *Mtb* H37Rv has no effect on growth in culture, there was a marked decrease in splenic CFUs in a murine infection model [100], which suggested that one or both of these genes were important for virulence. Further studies using recombinant purified LprG expressed in *E. coli* showed that LprG induced a strong Th1 type response [101], typically thought to be protective in mycobacterial infections [30]. However, similar to studies using LpqH, priming with LprG reduced BCG-afforded protection.
against subsequent challenge with Mtb [102]. The physiological function of LprG for mycobacteria is of interest, and recent work suggested that the LprG-p55 operon may function in cell envelope synthesis [103]. The work described in chapter 3 of this thesis demonstrates that LprG carries glycolipids in a hydrophobic pocket, and that the carried glycolipids contribute to its activation of Toll-like receptor 2, a pattern recognition receptor.

Figure 1.5. Bacterial lipoprotein synthesis. Each of these enzymatic modifications are thought to occur at the plasma membrane and are likely functionally linked to secretion and glycosylation machinery. Figure c/o Nicole Pecora.
Pathogen Recognition Receptors

The innate immune system serves as an early response system that determines the location of infection, attempts to destroys the pathogen, and facilitates the adaptive immune response [104]. The question of how the immune system initially detects the presence of MTB and other pathogens has in part been answered by the description of pattern recognition receptors (PRRs). PRRs are invariant receptors capable of recognizing structural motifs that are absent or under-represented in the host and relatively conserved among microorganisms. These structural motifs are referred to as Pathogen-Associated Molecular Patterns (PAMPs). Since most, if not all, pathogens express some form of PAMP, PRRs represent an efficient system consisting of a relatively small group of receptors capable of detecting a vast array of pathogens. There are several different families of PRRs, designed to detect pathogens in a wide variety of intra and extracellular compartments: Pentraxins [105], Collectins and Ficollins [106] act in the extracellular milieu; Toll-Like Receptors (TLRs) [107] act at cytoplasmic and vacuolar membranes, and the NOD-like receptors (NLRs, aka NBS-LRR, CATERPILLAR family) [108, 109] [110], RIG-I like helicases (RLRs) [111], DAI [112] and at least one other unidentified receptor [113] function in the cytosol to detect peptidoglycan fragments, RNA, and DNA; respectively. This thesis focuses on the detection of mycobacterial lipoproteins and glycolipids by TLRs, in particular TLR2 and its co-receptors and accessory receptors (discussed below).
**Toll-Like Receptors: Structure and signaling**

Toll-Like Receptors (TLRs) are a family of single-pass type I transmembrane glycoprotein PRRs involved in the initiation of the immune response by the innate immune system [114] and the development of adaptive immunity [115] [116]. TLRs were discovered in mammals through a homology search based on Toll, a drosophila protein important in developmental dorsal-ventral patterning [117], as well as immunity to fungal [118] and gram-positive bacterial [119] pathogens. Homology searches have led thus far to the discovery of six families of homologues, five of which are produced in humans (Fig. 1.6).

![Diagram](image)

**Figure 1.6. The human toll-like receptor family and some typical agonists.** Adapted from [29].
Toll and the TLRs have three domains: an intracellular Toll-Interleukin-1 Resistance (TIR) domain, a transmembrane domain, and an extracellular N-terminus almost entirely composed of Leucine Rich Repeats (LRRs), responsible for pathogen recognition. LRRs are a major focus of the work described in this thesis, and are described in detail below. The TIR domain is thought to mediate interaction of adaptor proteins that initiate a signaling cascade that results in activation of transcription factors including nuclear factor kappa-B (NF-kB), ATF2/c-Jun, and interferon response factors (IRF)-3 and IRF-7 [120]. ATF2/c-Jun is a component of AP-1 and is important in induction of the interferon response along with the IRFs. NF-kB is a master regulator of inflammation and is a strong inducer of cytokine production, such as TNF-a, IL-1, IL-6, IL-8, and others. In chapters 2 and 3, we utilize a bioassay that involves IL-8 production dependent on NF-kB as a measure of TLR2 activation.

**Long term TLR signaling leads to down-regulation of specific immune functions**

Although acute (1-4 hrs) Toll-like receptor signaling tends to induce inflammation and activate antigen presenting cells, several groups have shown that prolonged (24 hr) stimulation of Toll-Like receptor 2 by mycobacterial infection or by incubation with mycobacterial lipoproteins can down-regulate certain immune functions. For instance, LpqH (aka the 19kDa antigen) [40], LprA, [45] and LprG [44] can decrease processing and presentation of antigen by murine bone marrow-derived macrophages and human monocyte-derived macrophages [42]. This inhibition is TLR2-dependent and occurs by transcriptional repression of C2TA, the major regulator of MHC2 expression [47, 48].
Recent work demonstrated that this phenomenon occurs in vivo. While uninfected macrophages in the infected lung have increased MHC2, cells that harbor bacilli are decreased in surface MHC2, despite being in the same inflammatory environment [49][50].

**Toll-Like Receptor 2 is a sensor of lipoproteins and is important for immunity to Mycobacteria**

Toll-Like Receptor 2 has been implicated in the sensing of a broad array of structures from bacteria [121][121-126], fungi [127][128], and protozoa [129]. TLR2 senses at least two classes of molecules produced by *Mycobacteria*: lipoproteins and glycolipids (Fig. 1.7). There are 48 to 100 predicted hypothetical lipoproteins in the MTB genome [88]. The four lipoproteins discussed in this thesis have been characterized as TLR2 agonists: LpqH (the 19 kDa lipoprotein) [40], LprG [44] and LprA [45], and PhoS1 [92]. The glycolipids Lipoarabinomannan (LAM) and its anchor, Phosphatidyl Inositol Mannosides (PIMs) are major structural components of the mycobacterial cell wall, and are also sensed by TLRs: LAM by TLR2, and PIMs by both TLR2 or TLR4 [122]. Since both pro- and anti-inflammatory sequelae of TLR2 signaling have been described, it is important to consider whether the overall effect of TLR2 signaling is beneficial or detrimental to control of infection. Human polymorphism studies suggest that TLR2 plays an important protective role in Mtb infection. Several mutations in the TIR domain of TLR2 are associated with poorly controlled mycobacterial disease; Arg677Trp is associated with the development of lepromatous leprosy [130], and Arg753Gln (shown to reduce response to TLR2/TLR1 ligands [131]) is associated with
increased likelihood of development of active tuberculosis [132]. Interestingly, a TLR1 polymorphism (I602S) is associated with decreased response to TRL2/1 ligands but seems to be protective in the context of *M. leprae* infection [133].

**Figure 1.7.** Common structural motifs of some TLR2 agonists. All three agonists depicted share a diacyl-glycerol motif. (A) The PI-based glycolipid PIM2. One acyl chain is a tuberculostearic acid, and the others are typically palmitate. Red asterisks indicate additional sites of acylation. (B) The synthetic triacylated lipopeptide Pam3CSK4, which consists of three palmitic acids: One n-linked to the free amine terminus of the CSK4 peptide, the other two are part of the diacylglycerol motif thioether-linked to the cysteine side chain. (C) A generic scheme of a mycobacterial lipoprotein, thought to be triacylated similar to Pam3CSK4 (although fatty acids are not defined, and presence of N-linked acyl chain is not biochemically proven). The potential contribution of the protein component to TLR2 activation by full-length lipoproteins is a major focus of this thesis.
Co-receptors and accessory receptors of TLR2

TLR2 cooperates with other surface receptors, including “co-receptors” with which it forms heterodimers and “accessory receptors” that may enhance ligand delivery or recognition. Co-receptors of TLR2 are genetically related to TLR2 and include TLR1 and TLR6. TLR1 contributes to the recognition of triacylated lipopeptides [125] and lipoarabinomannan [134], whereas TLR6 contributes to the detection of diacylated lipopeptides [124] [135]. Although there is now crystallographic data for the structure of a synthetic triacylated lipopeptide (Pam3CSK4) bound to TLR2 and TLR1 [136], a number of questions remain about how TLR2 and its co-receptors recognize the diverse set of physiological ligands with TLR2 agonist activity.

In addition to contributions of the co-receptors that form heterodimers with TLR2, recognition of TLR2 agonists is influenced by several accessory receptors, including CD14 [137], CD36 [138], lipopolysaccharide binding protein [139], CD11b-CD18 integrin [140] and ganglioside GD1a [141]. All of these accessory receptors are hypothesized to deliver agonist to TLR2, and some are also thought to contribute to cytoplasmic signaling. CD14, a GPI-linked protein, directly binds Pam3CSK4 and delivers it to TLR2 [142], increasing the sensitivity of TLR2 responses by approximately 100-fold [143]. The delivery function of CD14 is supported by the finding that antibodies to CD14 block access of Pam3CSK4 to the plasma membrane of human monocytes, whereas although anti-TLR2 antibody blocked responsiveness (TNFa release) of human Peripheral Blood Mononuclear Cells (PBMC), it did not block binding of flag-tagged Pam3CSK4 to the cell surface [144]. CD36 enhances the sensitivity of
macrophages to lipoteichoic acid and MALP-2, but not Pam\textsubscript{3}CSK\textsubscript{4} [138]. Although biochemical evidence of direct binding of CD36 to lipoteichoic acid or MALP-2 is lacking, there are reports that it signals with its C-terminal cytoplasmic tail and that CD36-dependent phagocytosis is important for its ability to augment TLR2 responses [145] [146]. One caveat for some of these studies is their reliance on small synthetic lipopeptide agonists, which may differ from full-length lipoproteins in solubility, mode of interaction with TLR2 and reliance on specific accessory receptors. While there is data that triacylated OspA of Borrelia burgdorferi is sensed in a TLR2 and TLR1 dependent manner [147], there has been little assessment and comparison of accessory receptor usage among different physiological TLR2 ligands, e.g. lipoproteins, so potential heterogeneity in physiological accessory receptor usage remains unexplored. A major aim of this thesis is to describe the co-receptors and accessory receptors important for TLR2-mediated responses to four different Mtb lipoproteins. In chapter 3, we examine the role of co-receptors and the accessory receptor CD14 in response to glycolipids carried by LprG.

**Ligand structural determinants of TLR2-mediated detection: lipopeptides and lipoproteins**

TLR2 agonist activity of many bacterial lipoproteins is dependent upon acylation of the N-terminal cysteine, as demonstrated for OspA [148] [149], LpqH [148] and LprA [45]. However, recent data have suggested that the factors dictating TLR2 agonist activity are more complicated than the initial working model. For instance, it is now known that chirality of the alpha carbon is important for determining sensitivity of
detection; and that cells expressing TLR2 and TLR1 are capable of recognizing monoacylated, diacylated, and triacylated lipopeptides [150, 151]. In addition, there are many reports of TLR2 agonists that do not have any known acylation, including proteins such as Mtb ESAT-6 [51], Porin B from *N. meningitides* [152], fimbriae from both *P. gingivalis* [140] and *S. enterica* [153], Type-II enterotoxins from *E. coli* [154, 155], LcrV from *Y. pestis* [156], and capsular carbohydrates from *B. fragilis* [157]. The specifics of how these non-acylated agonists interact with TLR2 is unknown and likely different from that described for the triacylated agonist Pam₃CSK₄ [136].

A central hypothesis of this thesis is that the peptide component [158] or the glycan component [89] of lipoproteins may influence TLR2 agonist activity. Specifically, in Chapter 2 of this thesis, the lipoprotein LprG is shown to be a more potent TLR2 agonist than LprA, and in chapter 3, the difference in potency is shown to be due to glycolipids that are carried in a hydrophobic pocket of LprG.

**Ligand structural determinants of the TLR2-mediated detection: glycolipids**

Since the PI-based glycolipids contain a diacylglycerol motif (Fig. 1.7), it might be hypothesized that glycolipid TLR2 agonist activity could be explained by the diacylglycerol motif. However, while acylation is universally required, the glycans appear to play an important role in determining TLR2 agonist activity. For instance, PIM₂ and PIM₆ are weak agonists of TLR2 (defined by TNFa and IL-12 production) in murine and human macrophages; but while the TLR2 agonist activity requires acylation, it does not depend on the number of acyl chains (even mono-acylated species were active) [159]. In contrast, LM are strong agonists of TLR2 (requiring TLR1, not
TLR6—see chapter 3), and activity is only found in tri-acylated (Ac₃LM) and tetra-acylated (Ac₄LM) species [160]. Since Ac₃LM has the same core structure and acylation as AcⱼPIM₂, it is logical to conclude that the glycan structure influences TLR2-mediated recognition of these compounds. Indeed, LM purified from *Saccharothrix aerocolonigenes*, characterized by increased dimannoside branches added to the (alpha 1-6) linked mannan chain present in mycobacterial LM, had much greater TLR2 agonist activity than LM from mycobacteria [161]. These studies suggest that greater mannosylation facilitates TLR2-mediated detection of mycobacterial glycolipids. LAM is produced by arabinosylation of LM. Interestingly, ManLAM and AraLAM are both poor TLR2 agonists, a fact explained by arabinose residues masking the LM core of LAM [162], a hypothesis supported by an increase in TLR2 agonist activity after chemical removal of the arabinose chains of LAM to make an LM derivative [163]. As described above, *M. tuberculosis* and *M. leprae* produce ManLAM, and *M. cheloniae* produces uncapped AraLAM. In contrast, *M. smegmatis* produces PILAM, reported to be a strong TLR2 agonist [122] [164]. The strong TLR2 agonist activity of PILAM suggests that the presence of the PI cap overcomes the steric hindrance of the arabinose residues, likely by replacing the hidden LM mannan motif with a directly recognizable diacylglycerol motif.

These glycolipid families are heterogenous molecules with overlapping chemistries, making them challenging to purify. Preparations of LM and LAM typically contain LM, LAM, PIMs, and often a large molecular weight PAS-reactive contaminant that may be arabinan. Difficulties with purification may contribute to the apparent complexity of TLR2-mediated recognition described above. The glycans may affect
TLR2 agonist activity directly by interacting with TLR2, or indirectly, by influencing the behavior of the monomer in aqueous solution (i.e. aggregation [162]). Similarly, glycosylation may directly or indirectly affect accessory receptor requirements for TLR2 activation, or function as an agonist for additional PRR receptors that induce signaling independent of TLR2 (i.e. C-type lectins ). These complexities make the study of glycolipids both interesting and difficult.

**Leucine Rich Repeats**

TLRs are thought to rely on leucine rich repeats (LRR) to bind ligands. LRRs are short sequence motifs present in over 2000 proteins of quite diverse functions, in viruses to eukaryotes, and usually mediate protein-protein interactions [165]. LRR consist of 2-45 repeats of 20-29 amino acids that fold into an arc or horseshoe shape. Each 20-29 amino acid repeat contains a highly conserved sequence LxxLxLxxN/CxL, in which “L” is Val, Leu, or Ile, “N” is Asn, Thr, Ser, or Cys, and “C” is Cys or Ser. Sequence analysis has led to the description of at least seven different independently evolved subfamilies of LRRs defined by differences (external to the LxxLxLxxN/CxL motif) in sequences of the repeated motif.

**Receptor Structural determinants of TLR-mediated detection of PAMPs**

Recent work has yielded a bounty of information regarding how TLR1 [136], TLR2 [136], TLR3 [166] [167], TLR4 [168], TLR5 [169], and TLR9 [170] [171] bind their agonists. While in general terms, TLRs bind their agonists via their LRRs, there is remarkable complexity to the mechanisms of detection utilized by TLRs. This
complexity reflects the challenge of discriminating such diverse ligands. For the purposes of this discussion, it is useful to contrast the detection of a hydrophilic TLR ligand (dsRNA) with that of hydrophobic ligands (LPS and Pam₃CSK₄).

The hydrophilic TLR ligands are sensed by binding to the surface of the LRR of the TLR, and do not require accessory receptors or co-receptors. For instance, TLR3 has two positively charged sites responsible for interaction with the phosphate backbone of dsRNA: one composed of the N-terminal LRR (LRRNT) and LRRs 1-3, and the other composed of LRRs 19-21 [167] (Fig. 1.8). While TLR3 has a small region supportive of homotypic interaction near the LRRCT, binding of ligand drives dimerization of soluble extracellular domain [166] [172].

While TLR3 binds its ligands via ionic interactions mediated by patches of positively charged residues on the surface of its LRRs, TLR4 and TLR2 bind hydrophobic agonists via hydrophobic pockets, and share a requirement for co-receptors and accessory receptors to sense their hydrophobic ligands. The crystal structure of the TLR4/MD-2/LPS complex has been solved [168]. Hexaacylated LPS binds both MD-2 and TLR4, by inserting five of its acyl chains into MD-2 (inducing a conformation change), and the remaining acyl chain inserts into a hydrophobic dimerization interface between the MD-2 and the homodimer partner TLR4. While only one of the six lipid chains of LPS directly contacts TLR4, TLR2 appears to interact directly with 2 lipid chains of its acylated agonists. The TLR2-PAMP interaction that is best defined is the interaction between TLR2 and the synthetic triacylated lipopeptide Pam₃CSK₄. Monomeric TLR2 ECD can bind the diacyl glycerol motif of the synthetic lipopeptide Pam₃CSK₄ [142] via a hydrophobic pocket located at the junction between the central
and C-terminal portions of the ECD of TLR2 (Fig. 1.8) [136]. The amine-linked acyl chain of Pam₃CSK₄ inserts into a narrow groove located between the central and C-terminal domains of TLR1, effectively linking the two TLR ECDs [136].

Figure 1.8. Structures of TLRs bound to their agonists demonstrate different mechanisms of detection. (A) TLR3 binds the hydrophilic agonists dsRNA via charged patches on its surface [167]. (B) TLR2 and TLR1 detect hydrophobic agonists by binding the acyl chains directly via hydrophobic pockets [136]. Images reproduced with permission.
While the above studies demonstrate that TLR2 binds diacylglycerol motifs and TLR1 binds amine-linked fatty acids, the contribution of the peptide [158] or glycan [89] component of full-length, physiologically important lipoproteins remains unknown, and the reports of non-acylated TLR2 agonists suggest that other mechanisms of TLR2 binding exist. Studies of full-length lipoprotein agonists of TLR2 may shed light on other mechanisms of interaction with TLR2. A major aim of this thesis is to define the interactions between TLR2 and the mycobacterial lipoproteins LpqH, LprA, LprG, and the non-acylated TLR2 agonist, NA-LprG and its carried glycolipids. Work toward this goal is described in chapters 4 and 5.

In summary, the work described in this thesis was directed towards three goals. The first goal was to describe the contribution of co-receptors and accessory receptors to TLR2-mediated recognition of four mycobacterial lipoproteins. The second goal was to investigate the structural determinants responsible for TLR2 agonist activity of LprG. The third goal was to characterize in detail the binding of mycobacterial lipoproteins by TLR2 and TLR1. Regarding the first goal, the co-receptor and accessory receptor requirements of LprA, LprG, LpqH, and PhoS1 were found to differ among the agonists. While all agonists share a requirement for TLR2, and no dependence on TLR6, they showed varying dependence on TLR1, and LprA showed the least amount of dependence. LprA, LprG, and LpqH were detected in a CD14-dependent manner, while PhoS1 was not. CD36 only contributed to the detection of LprA. Lung APC subsets were shown to differ in their expression of TLR2, CD14, and CD36, and expression of TLR2 correlated with responsiveness to the mycobacterial lipoprotein LpqH. The second goal of this
thesis is an extension of work by Nicole Pecora, who found that LprG was a stronger TLR2 agonist than LprA, and that LprG retained TLR2 agonist activity in the absence of its covalent lipid modification. For the second goal of this thesis, it was found that the increased TLR2 agonist activity of LprG is due to glycolipids carried in a hydrophobic pocket and delivered to TLR2. We propose that glycolipid transport for the purpose of cell envelope construction is the physiological function of LprG, and that the increased bioavailability of glycolipid monomers carried by LprG enhances the sensitivity of glycolipid detection by TLR2. This represents the first report of a TLR co-opting a pathogen-derived protein to enhance detection of hydrophobic PAMPs. For the third aim of this thesis, only preliminary data has been accumulated. The extracellular LRRs of TLR2 is shown to be capable of binding LpqH in a pull-down assay. Other studies toward this aim are described in chapter 4, and current directions are described in chapter 5.
Chapter 2: TLR2 and its co-receptors determine responses of macrophages and dendritic cells to lipoproteins of *Mycobacterium tuberculosis*
Abstract

*Mycobacterium tuberculosis* (Mtb) signals through Toll-like receptor 2 (TLR2) to regulate antigen presenting cells (APCs). Mtb lipoproteins, including LpqH, LprA, LprG and PhoS1, are TLR2 agonists, but their co-receptor requirements are unknown. We studied Mtb lipoprotein-induced responses in TLR2−/−, TLR1−/−, TLR6−/−, CD14−/− and CD36−/− macrophages. Responses to LprA, LprG, LpqH and PhoS1 were completely dependent on TLR2. LprG, LpqH, and PhoS1 were dependent on TLR1, but LprA did not require TLR1. None of the lipoproteins required TLR6, although a redundant contribution by TLR6 cannot be excluded. CD14 contributed to detection of LprA, LprG and LpqH, whereas CD36 contributed only to detection of LprA. Studies of lung APC subsets revealed lower TLR2 expression by CD11b^high^/CD11c^low^ lung macrophages than CD11b^low^/CD11c^high^ alveolar macrophages, which correlated with hyporesponsiveness of lung macrophages to LpqH. Thus, lung APC subsets differ in TLR expression, which may determine differences in responses to Mtb. Findings in this chapter are published in Cellular Immunology [173] and are reproduced with permission.
Introduction

Toll-like receptor 2 (TLR2) is important for control of mycobacterial infections, including infections by Mycobacterium tuberculosis (Mtb) [174-177] and M. leprae [130]. Mycobacterial agonists of TLR2 include lipoproteins and glycolipids. The genome of Mtb is predicted to encode 48-100 lipoproteins of largely unknown function, nearly half of which share no conserved domains with lipoproteins of species outside the genus Mycobacteria [88]. LprA [45], LprG [44], LpqH (19-kDa lipoprotein) [40, 41, 46] and PhoS1 (38-kDa lipoprotein) [92] are Mtb lipoproteins with TLR2 agonist activity that modulate antigen presenting cell (APC) functions, which we define to include cytokine production and other innate immune responses of APCs as well as antigen presentation, recapitulating many effects of APC infection with Mtb. Despite the importance of understanding how TLR2 recognizes its ligands, many questions remain about the mechanisms whereby Mtb lipoproteins are recognized. In this study we used four well-characterized mycobacterial lipoproteins to dissect the contributions of TLR2, its co-receptors and accessory receptors to recognition of lipoproteins.

TLR2 cooperates with other surface receptors, including “co-receptors” with which it forms heterodimers and “accessory receptors” that may enhance ligand delivery or recognition. Co-receptors of TLR2 are genetically related to TLR2 and include TLR1 and TLR6. TLR1 contributes to the recognition of triacylated lipopeptides [125] and lipoarabinomannan [134], whereas TLR6 contributes to the detection of diacylated lipopeptides [124] [135]. Although there is now
crystallographic data for the structure of a synthetic triacylated lipopeptide (Pam₃CSK₄) bound to TLR2 and TLR1 [136], a number of questions remain about how TLR2 and its co-receptors recognize the diverse set of physiological ligands with TLR2 agonist activity. For example, the mechanisms by which non-acylated ligands, such as neisserial porins [152] and capsular carbohydrates [157], are recognized by TLR2 and its co-receptors remain unclear. For full-length, natural lipoproteins, it remains unknown how the peptide component [158] or glycan component [53] may influence signaling by TLR2, TLR1 and TLR6. Our studies used co-receptor knock-out cells to assess TLR2 co-receptor requirements for responses to four different Mtb lipoproteins, revealing variable dependence on TLR1 and a lack of dependence on TLR6.

In addition to contributions of the co-receptors that form heterodimers with TLR2, recognition of TLR2 agonists is influenced by several accessory receptors, including CD14 [137], CD36 [138], lipopolysaccharide binding protein [139], CD11b-CD18 integrin [140] and ganglioside GD1a [141]. All of these accessory receptors are hypothesized to deliver agonist to TLR2, and some are also thought to contribute to cytoplasmic signaling. CD14, a GPI-linked protein, directly binds Pam₃CSK₄ and delivers it to TLR2 [142], increasing the sensitivity of TLR2 responses by approximately 100-fold [143]. CD36 enhances the sensitivity of macrophages to lipoteichoic acid and MALP-2, but not Pam₃CSK₄ [138]. Although biochemical evidence of direct binding of CD36 to lipoteichoic acid or MALP-2 is lacking, there are reports that it signals with its C-terminal cytoplasmic tail and that CD36-dependent phagocytosis is important for its ability to augment TLR2
responses [145, 146]. One caveat for some of these studies is their reliance on small synthetic lipopeptide agonists, which may differ from full-length lipoproteins in solubility, mode of interaction with TLR2 and reliance on specific accessory receptors. While there is data that triacylated OspA of Borrelia burgdorferi is sensed in a TLR2 and TLR1 dependent manner [147], there has been little assessment and comparison of accessory receptor usage among different physiological TLR2 ligands, e.g. lipoproteins, so potential heterogeneity in physiological accessory receptor usage remains unexplored. Our studies assessed accessory receptor requirements for responses to four different Mtb lipoproteins, revealing both overlaps and differences in accessory receptors used to recognize these lipoproteins.

Mtb is an intracellular pathogen that infects APCs, and the ability of different APC types to respond to Mtb infection may be governed by their relative expression of TLR2 and its co-receptors and accessory receptors. The primary site of infection by Mtb is the lung. Lung APC populations include lung macrophages and lung dendritic cells (DCs) in the lung parenchyma, as well as alveolar macrophages [178] [179]. Mtb is harbored within alveolar macrophages early in infection and within lung macrophages and lung DCs later in infection [50, 180, 181]. Potential differences between these APC subsets in responsiveness to Mtb or its lipoproteins are largely unexplored and could have significant implications for host responses to Mtb infection. For example, the relative expression of TLR2 has not been assessed on lung APC subsets. We demonstrate that purified APC subsets from murine lungs
differ in expression of TLR2, and the subset with lowest TLR2 expression, lung macrophages, is hyporesponsive to Mtb lipoprotein.
Materials and Methods

Cloning and expression of Histidine-tagged LprG (Rv1411c) and LprA (Rv1270c)

LprA was cloned previously [45]. LprG was amplified from Mtb H37Rv genomic DNA by PCR using the following primers:

5’GCATATCCATATGCGGACCCCCAGACGCCACTG
3’GTACAAGCTTGCTCACCGGGGGCTTCG

The 5’ primers included an Ndel site and the 3’ primer included a HindIII site. The PCR product was digested with Ndel and HindIII (NEB, Ipswitch, MA) and ligated into the shuttle vector pVV16 (provided by J. Belisle, Colorado State University, Fort Collins, CO) behind the mycobacterial hsp60 promoter and in-frame with a C-terminal 6x Histidine (His) tag. All constructs were verified by sequencing and analyzed using Clone Manager (SciEd software, Cary, NC). Chemically competent *E. coli* (Invitrogen, Carlsbad, CA) was transformed according to the manufacturer’s protocol. *M. smegmatis* was transformed by electroporation with a Gene Pulser (Bio-Rad, Hercules, CA) set at 2.5 kV, 25 µF, and 800 Ohms. LprA and LprG were expressed in *M. smegmatis* MC2 1-2C (R. Wilkinson, Imperial College, London, U.K.) cultivated in Middlebrook 7H9 broth (Difco, Lawrence, KS) supplemented with 1% casamino acids (Fisher, Pittsburgh, PA, BP1424), 0.2% glycerol (Fisher G33-1), 0.2% glucose, and 0.05% Tween 80. Kanamycin was used at 30 µg/ml for selection of both *E. coli* and *M. smegmatis*. Bacteria were isolated by centrifugation at 6000 x g for 20 min at 4°C.

*Mtb* strain H37Ra (ATCC 25177) was cultured with shaking at 37°C to late log phase growth (2.5 weeks) in *Mtb* 7H9 broth (4.7 g/l 7H9 (Difco 271310), 5 ml/l
glycerol, 0.5 ml/l Tween-80 (Sigma, St. Louis, MO, P4780) supplemented with 10% albumin/dextrose/catalase (BD, Franklin Lakes, NJ 212352). Bacilli were harvested by centrifugation at 5,000 x $g$ for 20 min at 4°C.

**Lysis and Purification of His$_6$-tagged proteins**

Purification of LprA and LprG was accomplished as reported previously [45]. Cells were resuspended in lysis buffer (10 ml/liter of bacterial culture) consisting of 50 mM NaH$_2$PO$_4$, 300 mM NaCl, 20 mM imidazole, pH 8.0, 2.5% protease inhibitor cocktail (Sigma P8849), 75 U/ml benzonase (Novagen, Madison, WI, 70664-3), and 2.5 mg lysozyme (Sigma, L-3790) and incubated for 15 min at 37°C. Bacteria were disrupted mechanically by 4 passages through a French press (2000 psi). Insoluble material was removed from the lysate by ultracentrifugation at 100,000 x $g$ for 1 h at 4°C, and supernatant was incubated directly with Ni-NTA beads (Qiagen, Valencia, CA, 1018244) for 2-4 h at 4°C. Ni-NTA beads were transferred to polypropylene columns, washed with 75 volumes of wash buffer (50 mM NaH$_2$PO$_4$, 1 M NaCl, 20 mM imidazole, 10% glycerol, pH 8.0), and bound protein was dissociated with elution buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl, 450 mM imidazole, pH 8.0). To prepare for anion exchange chromatography, sample was desalted into 20 mM Tris, pH 8.0 using PD-10 columns (GE Healthcare, Uppsala, Sweden 17-085-01). Samples were subjected to anion exchange chromatography using quaternary ammonium columns (GE Healthcare, 17-5053-01), and eluted with the addition of NaCl in the following steps: 50, 100, 200, 400, 1000 mM. Presence and purity of desired protein
was verified by SDS-PAGE and visualized as single bands by both silver stain and anti-His Western blot; yields were estimated by BCA protein assay (Pierce, Rockford, IL, 23225). Material eluted by 100 mM NaCl was used for all experiments.

*Lysis of Mtb H37Ra and purification of LpqH*

Bacilli from 4 liters of Mtb H37Ra culture were collected by centrifugation as above, resuspended in 40 ml of water supplemented with protease inhibitor cocktail (Sigma P8849), 75 U/ml benzonase (Novagen 70664-3), and 2.5 mg of lysozyme, incubated for 15 min at 37°C, and mechanically disrupted via 4 passages through a French Press (2000 psi). To extract hydrophobic molecules, lysate was combined with 6% Triton X-114 (Sigma, X-114) in Tris buffered saline (TBS; 50 mM Tris, 150 mM NaCl, pH 7.5) to achieve a concentration of 4% Triton X-114. After overnight incubation at 4°C, insoluble material was removed by centrifugation at 100,000 x g for 1 h at 4°C, and supernatant was stored at 4°C overnight. Insoluble material was resuspended in 4% Triton X-114 in TBS at 4°C and incubated overnight at 4°C. Insoluble material was again removed by centrifugation at 100,000 x g for 1 h at 4°C. Soluble material from the two sequential extractions was used for further purification. To remove hydrophilic contaminants, soluble material from Triton X-114 extractions was back-extracted three times by addition of a two-fold excess of ice-cold TBS, incubation on ice for 15 min, incubation at 37°C for 15 min to allow separation of detergent and aqueous phases, centrifugation (3,000 x g for 15 min at 37°C) and removal of the upper aqueous phase by aspiration. To remove TBS, the
detergent phase was similarly back-extracted once with ice-cold water. To remove detergent, protein was precipitated by the addition of 5 volumes of acetone at -20°C and stored overnight at -20°C. The acetone precipitate was collected by centrifugation, washed 3 times with -20°C acetone, and air-dried to yield approximately 100 mg dry weight of a white powder per liter of culture. To exclude glycolipids, the acetone precipitate was solubilized in PBS-saturated phenol and subjected to six sequential extractions by the addition of PBS (1:1 v/v) at room temperature for 4-15 h (rotating in Teflon-coated tubes with sealing caps, Fisher scientific 05-562-16B, 05-563-1c), centrifugation at 15,000 x g for 30 min at room temperature, and removal of the upper aqueous phase. The final phenol phase was dialyzed against distilled water using a 3.5 kD molecular weight cut-off membrane (Corning, Lowell, MA, 132720). The precipitate was removed from dialysis tubing, washed extensively with distilled water at room temperature, and lyophilized. The yield from one liter of culture was approximately 25 mg of precipitate. Eight mg of lyophilized phenol precipitate was subjected to SDS-PAGE through a preparative 20 x 20 cm, 1.5 mm thick 13% polyacrylamide Tris-Cl gel using Protean IIxi electrophoresis apparatus (BioRad, Hercules, CA). The sample was eluted in fractions using a Whole-Gel Eluter (BioRad) into imidazole-HEPES buffer (43 mM imidazole, 35 mM HEPES, pH 7.4). To identify fractions containing LpqH, a portion of each fraction was analyzed by SDS-PAGE with silver staining or Western blotting using polyclonal anti-BCG Ab (DAKO, B0124, 1:30,000) or the LpqH-specific monoclonal antibody IT-19 (1:3000) [182]. Fractions containing LpqH and no other bands by silver gel were pooled and concentrated using 10,000 molecular weight
cut-off tubes (Millipore, Billerica, MA, UFC801096). Yields were determined by BCA protein assay.

Purified PhoS1 was provided by John T. Belisle under NIAID contract HHSN266200400091c “TB Vaccine Testing and Research Materials”; it was purified from Mtb culture supernatant by ammonium sulfate precipitation, lectin chromatography, and hydrophobic interaction chromatography.

Mammalian cell culture

Unless otherwise specified, incubations with eukaryotic cells were performed at 37°C in 5% CO₂ atmosphere. Standard medium was DMEM (Hyclone, Logan, UT, ASK30773) supplemented with 10% heat-inactivated FCS, 50 μM 2-ME, 2 mM L-glutamine, 1 mM sodium pyruvate, 10 mM HEPES, pH 7.4, and penicillin/streptomycin (Hyclone). Female C57BL/6J mice (8-16 weeks old) were obtained from the Jackson Laboratory, housed under specific pathogen-free conditions and used to produce macrophages. TLR1+/−, TLR2+/− and TLR6+/− mice were generously provided by Shizuo Akira (Research Institute for Microbial Disease, Osaka University, Osaka, Japan) and were back-crossed to C57BL/6J mice a minimum of eight times. CD14−/− mice (B6.129S-Cd14tm1Frm/J) were obtained from the Jackson Laboratory, maintained under specific pathogen-free conditions and used to produce macrophages. CD14−/− mice were compared to C57BL/6J mice and F2 hybrids of C57BL/6J and 129sv. CD36−/− mice were obtained and backcrossed onto C57BL/6 background a minimum of eight times as described [183]. Bone
marrow cells were cultured for 7-12 d in standard medium supplemented with 25% LADMAC cell-conditioned medium [184]. HEK293 cells stably expressing TLR2-YFP (HEK293.TLR2) were produced previously [185] [186]. HEK293 cells (ATCC CRL-1573) were stably transfected with the empty vector to produce a control HEK293.pcDNA3 cell line. Transfected HEK293 cell lines were maintained in HEK medium (DMEM supplemented with 10% heat-inactivated FCS (HyClone), ciprofloxacin (10 µg/ml) and geneticin (500 µg/ml)).

Cytokine ELISAs

HEK293.TLR2 or HEK293.pcDNA3 cells were incubated in 96-well plates (20,000 cells/well) for 8 h in 90 µl of HEK medium and then for an additional 16 h with or without TLR2 agonist. Supernatant IL-8 concentration was quantified by ELISA (R&D, Minneapolis, MN, DY208). Bone marrow-derived macrophages were incubated overnight at 100,000 cells/well in standard medium and then with or without lipoprotein for 12 h. Supernatants were collected and stored at -80°C. TNF-alpha in the supernatant was quantified by ELISA (BD Biosciences #558874, R&D DY410). The following synthetic TLR agonists were also used: Ultrapure E. coli LPS (Invivogen, San Diego, CA, tlr1-pelps), FSL-1 (Invivogen, tlr1-fsl) and Pam3CSK4 (Invivogen, tlr1-pms).
Lung APC preparation and flow cytometry

Mice were anesthetized with 1.25% 2,2,2 tribromoethanol (Sigma, T4840-2) solution dissolved in 40% 2-methyl-2-butanol (tert-amyl alcohol; Sigma, A-1685). The pulmonary circulation was perfused with ~10 ml PBS with 2 mM EDTA. Lungs were removed, minced, and incubated in standard medium supplemented with 2.5 mg/ml collagenase D (Roche, Indianapolis, IN) and 75 U/ml benzonase for 45 min at 37°C. Tissue was then passed through a 70-µm filter and rinsed extensively with standard medium. Cells were resuspended in ACK lysis buffer for 5 min at room temperature. Dead cells and tissue debris were removed with Dead Cell Removal kit (Miltenyi Biotec, Auburn, CA, 130-090-101). Crude lung APCs were prepared using a 1:1 mixture of magnetic beads selective for CD11b (Miltenyi biotec, 130-049-601) and CD11c (Miltenyi Biotec, 130-052-001) (100 µl of each slurry per 10⁷ viable cells). For flow cytometry, crude lung APCs were stained using the following antibodies: APC-AF750-conjugated anti-CD11b (eBioscience, San Diego, CA, M1/70, #27-0112, used at 0.4 µg/1x10⁶ cells), PB-conjugated anti-CD11c (eBioscience, N418, #57-0114, used at 0.4 µg/1x10⁶ cells) or APC-conjugated anti-CD11c (eBioscience, N418, #17-0114-82, used at 0.4 µg/1x10⁶ cells), PECy7-conjugated anti-TLR2 (eBioscience, T2.5, #25-9024, used at 0.4 µg/1x10⁶ cells), FITC-conjugated anti-CD14 (eBioscience, Sa 2-8, #11-141-82, used at 1.0 µg/1x10⁶ cells), PE-conjugated anti-CD36 (eBioscience, No.72-1, #12-0261, used at 0.4 µg/1x10⁶ cells) and biotin-conjugated anti-GR1 (eBioscience, RB6-8C5, #13-5931, used at 1.0 µg/1x10⁶ cells). Flow analysis and sorting were performed on a BD FACSaria
cytometer. Specific mean fluorescence intensity (MFI) was calculated as MFI with specific Ab minus MFI with isotype control Ab. Samples to be sorted were stained only with antibodies specific for CD11b, CD11c, and GR-1. Data was analyzed using Winlist 5.0 software (Verity, Topsham, ME). Events representing single cells were gated by forward and side scatter. Of these, only GR1-negative events were included in the analysis and sort. GR1-negative events were sorted by FACS to isolate lung APC subsets based on surface expression of CD11b and CD11c as previously described [49, 178]. Lung APC populations included alveolar macrophages (AM, CD11b\text{low}/CD11c\text{high}), lung macrophages (LM, CD11b\text{high}/CD11c\text{low}), and lung DCs (CD11b\text{high}/CD11c\text{high}). To assess relative responsiveness to a TLR2 agonist, FACS-purified lung macrophages and alveolar macrophages were incubated in 96-well plates (30,000 cells/well) in standard medium for 3 h and then in serum-free medium (Macrophage SFM, Invitrogen, 12065074) with or without LpqH for 16 h. Supernatants were collected, stored at -80°C and assessed for TNF-alpha by ELISA.

Statistical analysis

Two-tailed Student’s T-test was used for comparisons of responses of two cell types (e.g. HEK293.TLR2 vs. HEK293.pcDNA3, wild-type vs. CD14-/- macrophages, or lung macrophages vs. alveolar macrophages). Two-way analysis of variance (ANOVA) with Bonferroni post-test was used comparisons of multiple cell types with different TLR deficiencies.
Results

These studies were designed to enhance our understanding of TLR recognition of Mtb lipoproteins, including the potential contributions of coreceptors or accessory receptors. To assess potential variation in recognition mechanisms and determine general principles, we studied multiple Mtb lipoproteins, including LprA, LprG, and LpqH; a fourth Mtb lipoprotein, PhoS1, was included in some studies. To compare their TLR2 agonist function, Mtb lipoproteins and a TLR2 model agonist (Pam$_3$CSK$_4$) were incubated with HEK293.TLR2 cells (transfected with human TLR2) or HEK293.pcDNA3 cells (transfected with control vector), and induction of IL-8 was assessed by ELISA. Mtb lipoproteins induced expression of IL-8 by HEK293.TLR2 cells but not HEK293.pcDNA3 cells (Fig. 2.1). A concentration of approximately 3-10 nM was sufficient to induce responses by LprA, LprG, LpqH and Pam$_3$CSK$_4$ (Fig. 2.1). These data confirmed that LprA, LprG and LpqH are all agonists of human TLR2, consistent with other data from studies with murine TLR2$^{-/-}$ macrophages (below) that indicate that these lipoproteins, as well as PhoS1, are agonists of TLR2.
Figure 2.1. **Mycobacterial lipoproteins are TLR2 agonists.** HEK293.TLR2 cells (filled symbols) or control HEK293.pcDNA3 cells (empty symbols) were exposed to mycobacterial lipoproteins for 15 h, and the concentration of IL-8 in culture supernatant was assayed by ELISA. Data represent mean +/- SD for triplicate samples, except where error is too small to be visualized. For concentrations above 10 nM, LprG, LpqH, LprA and Pam3CSK4 produced responses in HEK293.TLR2 cells that were significantly different from responses in HEK293.pcDNA3 cells (p < 0.01 for all such comparisons). Reproduced with permission from Cellular Immunology [173].

Since TLR2 is thought to signal as a heterodimer with either TLR1 or TLR6, we investigated contributions of these co-receptors to recognition of Mtb lipoproteins. TLR dependence of TNF-alpha induction by Mtb lipoproteins was studied with bone marrow-derived macrophages from wild-type, TLR2-/-, TLR1-/- or TLR6-/- mice (**Fig. 2.2**). TNF-alpha induction was completely dependent on TLR2 expression for LprA, LprG, LpqH, and PhoS1. With the exception of LprA, TNF-alpha induction by these lipoproteins was dependent on TLR1 expression (responses were significantly reduced with TLR1-/- macrophages, although to a lesser degree than with TLR2-/- macrophages). Lipoprotein-induced responses were not reduced in TLR6-/- cells, although a redundant contribution of TLR6 cannot be excluded. As predicted, TNF-alpha induction by the synthetic diacylated lipopeptide FSL-1 was
dependent on expression of TLR2 and TLR6 but not TLR1, and the response to Pam₃CSK₄ lipopeptide was dependent on expression of TLR2 and TLR1 but not TLR6 (data not shown). Induction of TNF-alpha by LprA was not reduced by single knockout of either TLR1 or TLR6; possible explanations include redundant contributions of TLR1 and TLR6 for LprA. These data suggest that Mtb lipoproteins are generally more dependent on TLR1 than TLR6, although TLR2 co-receptor dependence varies to some degree with different lipoproteins.

**Figure 2.2. TLR1 is more important than TLR6 for response to mycobacterial lipoproteins.** Bone marrow-derived macrophages from wild-type, TLR1⁻/⁻, TLR2⁻/⁻, or TLR6⁻/⁻ mice were stimulated for 12 h with mycobacterial lipoproteins in standard medium. Culture supernatants were assayed for TNF-alpha by ELISA. Data are representative of three independent experiments (two for PhoS1) that used independent preparations of lipoproteins and different preparations of bone marrow derived macrophages. Data are shown as the mean +/- SD of triplicate assays (*, p < 0.05; **, p < 0.01; and ***, p < 0.001 for comparison of wild-type vs. TLR1⁻/⁻ or TLR2⁻/⁻ cells). Reproduced with permission from Cellular Immunology [173].
In addition to the co-receptors that form heterodimers with TLR2, lipoprotein recognition may be influenced by accessory receptors that assist in recognition or delivery of ligands. CD14 and CD36 have been reported to contribute to recognition of other ligands by TLR2 [137, 138]. To assess CD14 contribution to TLR2-mediated recognition of Mtb lipoproteins, we tested lipoprotein induction of TNF-alpha expression by bone marrow-derived macrophages from CD14−/− and wild-type mice (Fig. 2.3). CD14−/− macrophages exhibited impaired cytokine responses relative to wild-type macrophages, although the magnitude of the deficit in CD14−/− cells varied with different TLR2 agonists. Among the Mtb lipoproteins, responses to LprA showed the greatest dependence on CD14; responses to LprG and LpqH showed intermediate dependence on CD14, and the response to PhoS1 showed no discernable dependence on CD14. LPS, a TLR4 agonist, showed a predicted dependence on CD14, while FSL-1 did not. Similar CD14 dependence was seen when bone marrow-derived macrophages from CD14−/− mice were compared with B6 x 129sv F2 hybrid mice (results not shown). To investigate the importance of CD36 to recognition of Mtb lipoproteins by TLR2, we compared Mtb lipoprotein induction of TNF-alpha by wild-type and CD36−/− mice (Fig. 2.4). While LprA responses were dependent on CD36, responses to PhoS1, LprG and LpqH were independent of CD36 expression. In agreement with previous literature [138], response to LPS was not dependent on CD36 expression. These results indicate that responses to different Mtb lipoproteins vary in their dependence on accessory receptors, with dependence on CD14 more common than dependence on CD36 in the set of studied lipoproteins.
Figure 2.3. CD14 contributes to the detection of LpqH, LprA and LprG, but not PhoS1 or FSL-1. Bone marrow-derived macrophages derived from wild-type or CD14−/− mice were stimulated with a range of concentrations of agonist in standard medium (LPS) or SFM (all other agonists). TNF-alpha in culture supernatant was measured by ELISA. Data are representative of three independent experiments (two for PhoS1). Data represent mean +/- SD for triplicate samples (*, p < 0.05; **, p < 0.01 for comparison of CD14−/− vs. wild-type cells). Figure reproduced with permission from Cellular Immunology [173].
Figure 2.4. **CD36 contributes to the detection LprA, but not LpqH, LprG, or PhoS1.** Bone marrow-derived macrophages derived from wild-type or CD36 knockout mice were stimulated with a range of concentrations of agonist for 12 h in standard medium (LPS) or SFM (all other ligands). Culture supernatant was assayed for TNF-alpha by ELISA. Data are representative of three independent experiments (two for PhoS1). Data represent mean +/- SD for triplicate samples (*, p < 0.05 for comparison of CD36-/- vs. wild-type cells). Figure reproduced with permission from Cellular Immunology [173].
Since Mtb infection primarily occurs in the lung, differential expression of TLR2 or its co-receptors on lung APC subsets may influence the contributions of these cells to host defense in Mtb infection. Therefore, we assessed receptor expression in several APC populations prepared directly from lung tissue. Crude lung APCs were prepared from lung homogenates with a mixture of CD11b and CD11c affinity magnetic beads, and specific lung APC subpopulations were then purified by FACS to exclude Gr-1+ events (neutrophils) and prepare lung APC subsets, including CD11b\textsubscript{low}/CD11c\textsubscript{high} alveolar macrophages, CD11b\textsuperscript{high}/CD11c\textsuperscript{low} lung macrophages and CD11b\textsuperscript{high}/CD11c\textsuperscript{high} lung DCs. These APC subset definitions are supported by other published studies [178]; furthermore, bronchoalveolar lavage cells are predominantly CD11b\textsubscript{low}/CD11c\textsubscript{high}, supporting the definition of alveolar macrophages by this marker phenotype, and CD11b\textsuperscript{high}/CD11c\textsuperscript{high} lung DCs have high expression of MHC-II and C-type lectin receptors DC-SIGN and Dec-205 [187], consistent with this cell type assignment. By this approach and selective gating (Fig. 2.5), crude lung APCs were approximately 23% alveolar macrophages, 13% lung macrophages and 7% lung DCs, with other cells consisting of CD11b\textsubscript{low}/CD11c\textsubscript{low} cells or cells with intermediate phenotypes lying between the gates for the defined APC subsets. Interestingly, surface expression of TLR2, CD14 and CD36 varied among the APC subsets (Fig. 2.6). For TLR2, CD14 and CD36, respectively, alveolar macrophages had specific MFI of 422, 287 and 1175, lung DCs had specific MFI of 673, 917 and 744, and lung macrophages had specific MFI of 113, 247 and 1007. Thus, expression of TLR2, CD14 and CD36 varies on lung APC subsets, and lung macrophages have particularly low TLR2 surface expression.
Figure 2.5. *Isolation and characterization of lung APCs by flow cytometry.*

Crude lung APCs were gated by scatter properties (A) and GR1-negativity (shaded peak is isotype, dotted line is stain) (B). APC subsets were characterized by expression of CD11b and CD11c (C and D). Alveolar macrophages (AM) are defined as CD11b<sub>low</sub>/CD11c<sub>high</sub>. Lung macrophages (LM) are defined as CD11b<sub>high</sub>/CD11c<sub>low</sub> and consisted of 2 morphologies. Lung DCs are defined as CD11b<sub>high</sub>/CD11c<sub>high</sub>. GR1<sup>high</sup> events were excluded from analysis. PMN, morphology of GR1-positive events. Reproduced with permission from Cellular Immunology [173].

We investigated whether low TLR2 expression by lung macrophages correlated with a lack of responsiveness to the well-characterized mycobacterial
TLR2 ligand, LpqH (the 19-kD lipoprotein). Lung APC populations were purified by FACS with populations defined by expression of CD11b and CD11c (see Methods). Since cell yields were too low to obtain adequate numbers of all three lung APC types by three-way sorting, our analysis was focused on comparison of alveolar macrophages and lung macrophages. Macrophages were incubated with LpqH for 16 h, and TNF-alpha production was then assessed by ELISA. While alveolar macrophages produced TNF-alpha in response to LpqH, lung macrophages were minimally responsive (Fig. 2.6D). Thus, the low level of TLR2 expression on lung macrophages was associated with a decreased ability of these cells to respond to LpqH by induction of TNF-alpha. These data suggest that differences in TLR2 expression have functional consequences for lung APCs.
Figure 2.6. Differing levels of TLR2 on lung APC subsets. Surface staining for TLR2 (A), CD14 (B) and CD36 (C) is shown with dotted lines, and isotype control staining is shown in gray. Specific MFI values (see text) were calculated as MFI with specific Ab minus MFI with isotype control Ab. In panel D, cells were stimulated with LpqH for 16 h, and TNF-alpha in culture supernatant was measured by ELISA. Results are representative of 3 independent experiments. Data points represent the mean +/- SD of results from at least triplicate cell cultures. Due to limited cell numbers, fewer LpqH concentrations were assessed with lung macrophages. At the LpqH concentration points that were common to both cell types (10 and 100 nM), the responses were significantly different (**, p < 0.01; ***, p < 0.001). Reproduced with permission from Cellular Immunology [173].
Discussion

In this work, we assessed the contributions of co-receptors and accessory receptors to the recognition of several Mtb lipoproteins by TLR2. Mtb lipoproteins were generally more dependent on TLR1 than TLR6 as a TLR2 co-receptor, and more dependent on CD14 than CD36 as an accessory receptor. Furthermore, expression of TLR2 and its co-receptors differed between lung APC subsets, and the very low expression of TLR2 by lung macrophages was associated with low responsiveness of these cells to LpqH. These data suggest that variation in expression of TLR2 may influence the ability of different APC types to respond to infection with Mtb.

While TLR1 and TLR6 are both characterized as co-receptors of TLR2, the degree to which each contributes and the mechanisms by which each contributes remain unclear for physiological TLR2 ligands, e.g. Mtb lipoproteins. TLR1 and TLR6 are thought to participate directly in both ligand recognition and adaptor recruitment. Investigations with synthetic lipopeptides suggested that TLR2/TLR1 heterodimers sense triacylated lipopeptides [125] whereas TLR2/TLR6 heterodimers sense diacylated lipopeptides [124]. However, the ability of TLR2 to detect nonacylated ligands [152] and capsular carbohydrates [157] indicates that other structural determinants affect TLR2 co-receptor usage. We compared the responses of TLR2−/−, TLR1−/− or TLR6−/− macrophages to several lipoproteins (Fig. 2.2). These studies revealed three major characteristics of responses to Mtb lipoproteins. First, TLR2 was absolutely required for response to all of the Mtb lipoproteins, consistent with previous studies with LpqH [125] [148]. Second,
lipoproteins differ significantly in their efficacy for cytokine induction in both the HEK293.TLR2 cell line and murine bone marrow-derived macrophages. Third, TLR6 is not important for response to these lipoproteins. Finally, TLR1 contributes significantly to recognition of most of the studied Mtb lipoproteins, although low responses to high concentrations of the lipoproteins may still be achieved in the absence of TLR1; similar results were observed for Pam3CSK4 lipopeptide in our studies and others [125] [151]. Interestingly, LprA showed markedly less dependence on TLR1 than the other lipoproteins. It is possible that some TLR2 agonists have cross-reactivity for TLR1 and another TLR2 co-receptor, e.g. TLR6, TLR10 (which is closely related to TLR1 and TLR6) or a non-TLR receptor. Since Mtb lipoproteins show variability in their co-receptor usage, they may be useful tools for further studies of the structural determinants of TLR2 co-receptor dependence.

The mechanisms by which CD14 contributes as an accessory receptor to promote TLR2 signaling are incompletely understood; prevalent hypotheses include a role in trafficking and/or increasing bioavailability of TLR ligands. Like the TLRs, CD14 is a leucine rich repeat protein that is GPI-linked to the cell surface and cannot signal directly. CD14 can also be released from the cell surface and is present in soluble forms in mammalian serum. The presence of either membrane-associated CD14 or recombinant soluble CD14 enhances Pam3CSK4 binding to TLR2 [142] and resultant signaling [143]. Our data show that not all lipoproteins require CD14 (Fig. 2.3). Among the Mtb lipoproteins, responses to LprA showed the greatest dependence on CD14; responses to LprG and LpqH showed intermediate
dependence on CD14, and the response to PhoS1 showed no discernable dependence on CD14. PhoS1 is particularly striking in its ability to signal in CD14−/− macrophages. Further studies of these lipoproteins may yield insight into the basis of CD14 dependent recognition of bacterial lipoproteins by TLR2.

The role for CD36 in TLR2 signaling is less clear. As a prototypic class B scavenger receptor, CD36 performs homeostatic functions, including recognition and phagocytosis of apoptotic cells, senescent cells, cellular debris, and oxidized LDL [183]. It is involved in the detection and phagocytosis of pathogens, including *Staphylococcus aureus* and, to a lesser extent, *E. coli* [145]. More specifically, CD36 is thought to associate with the TLR2 receptor complex and contribute selectively to the detection of some diacylglycerol motifs by TLR2 [138] [188]. Of the Mtb lipoproteins examined in this work, only LprA induced a response that was dependent on CD36 expression. Therefore, LprA may be useful in future studies examining structural components responsible for CD36 accessory function.

It is interesting that LprA and LprG differ in their receptor usage despite being closely related (~34% identical amino acid sequence). LprA was partially dependent on both CD14 and CD36, whereas LprG showed partial dependence on CD14 but not CD36. LprA, unlike the other three Mtb lipoproteins in this study showed no dependence on TLR1. The structural characteristics of LprA that determine its unusual receptor usage are as yet undefined.

Lung APCs that are infected during the course of aerosol infection with mycobacteria include alveolar macrophages, lung macrophages and lung dendritic
cells. We investigated the expression of TLR2 and two potential accessory receptors (CD14 and CD36) by these lung APC subsets. Lung macrophages expressed less TLR2 than the other lung APC subsets (Fig. 2.6A-C). Consistent with their low level of TLR2 expression, lung macrophages were hyporesponsive to LpqH. Since lung macrophages may be infected by mycobacteria, the low level of TLR2 expression on these cells may limit TLR2-dependent responses to mycobacteria that these cells harbor. Thus, low levels of receptor expression on certain APC types may become a limiting factor for responses to pathogens. Due to the likely contribution of multiple receptors in host cell interactions with a pathogen, reduced expression of a single receptor may not ablate host cell responses to the pathogen, but limitation in expression of innate immune receptors by subsets of APCs may alter the quality or magnitude of responses, potentially affecting pathogenesis of infection.
Chapter 3: *Mycobacterium tuberculosis* lipoprotein LprG (Rv1411c) is a glycolipid carrier and delivers triacylated glycolipids to Toll-like receptor 2
Abstract

*Mycobacterium tuberculosis* (Mtb) expresses lipoproteins (e.g. LprG), whose TLR2 activity is considered to be dependent on their N-terminal triacylation. Surprisingly, non-acylated (NA)-LprG retained TLR2 agonist activity, demonstrating a previously unknown mechanism for TLR2 agonism. Western analysis and mass spectrometry demonstrated association of LprG with lipoarabinomannan, lipomannan and phosphatidylinositol mannoside (PIM), all glycolipid TLR2 agonists with shared core structures. Triacylated glycolipids were associated with LprG but not a homologous lipoprotein, LprA. TLR2 potency of glycolipids was increased by association with NA-LprG. X-ray crystallography of LprG revealed a hydrophobic pocket of 1535 cubic angstroms, sufficient to accommodate three alkyl chains of triacylated glycolipids. Point mutations in this putative binding pocket reduced association of glycolipid TLR2 agonists with LprG. X-ray crystallography found that PIM binds inside the pocket and that binding is in direct conflict with the point mutation that decreases TLR2 activity of LprG. Thus, LprG is a glycolipid carrier with specificity that allows selection of triacylated glycolipids, which may facilitate their delivery for cell wall biogenesis. Furthermore, host cells may co-opt LprG glycolipid carrier function as a novel mechanism to enhance TLR2 recognition of triacylated Mtb glycolipids. This chapter is adapted from work published in Nature Structural and Molecular Biology [189] and is used with permission.
Introduction

*Mycobacterium tuberculosis* (Mtb) infects one third of the world’s population and continues to be a leading cause of death. Additional challenges include the emergence of antibiotic-resistant Mtb strains, the substantially increased incidence of tuberculosis in HIV-infected individuals and the lack of an effective vaccine.

Although the vigorous immune response generated by immunocompetent individuals can eradicate most bacilli and contain infection, sterilizing immunity is not usually achieved [60]. Most infected individuals harbor latent disease with the potential for future reactivation, clinical disease, infectious spread and mortality.

The hydrophobic cell envelope of Mtb may contribute to many aspects of tuberculosis pathogenesis, including long-term survival in the host. Mtb cell wall components stimulate host responses and contribute to the activity of Freund’s adjuvant [190]. The mycobacterial cell wall contains glycolipids, which contribute to resistance to bactericidal free-radicals generated by host cells [33] and modulate immune functions, including phagosome maturation [37] [191] and cytokine production. The cell wall also contains lipoproteins that are triacylated at their N-terminal cysteine residues; the genome of Mtb is predicted to encode 65 or more such cell wall lipoproteins [192] [88].

Toll-like receptor 2 (TLR2), which forms heterodimers with TLR1 or TLR6, is an important contributor to innate immune recognition of Mtb [40, 41, 49, 125, 134, 148, 159, 193-196]. TLR2/TLR1 heterodimers bind lipopeptides that are triacylated on an N-terminal cysteine. Crystal structures show that the thioether-
linked diacylglycerol binds a hydrophobic pocket in TLR2, and the amide-linked third acyl chain binds TLR1 [136]. Alternatively, TLR2 can pair with TLR6, which lacks a hydrophobic groove for binding lipids. The TLR2/TLR6 heterodimer is activated by agonists with two alkyl chains, likely with both alkyl chains bound within TLR2 and hydrophilic elements of the ligand contacting TLR6 [136]. For example, mycoplasma lipopeptides lack the amide-linked acyl chain so that they are diacylated and signal through TLR2/TLR6 heterodimers [124, 197].

TLR2 agonist activity has been demonstrated for the following Mtb lipoproteins (Mtb H37Rv gene nomenclature and protein name synonyms in parentheses): LpqH (Rv3763, 19-kD lipoprotein) [40, 41, 125, 148], LprA (Rv1270c) [45], LprG (Rv1411c, p27) [44], and Psts1 (Rv0934; PhoS1, p38) [92]. Mycobacterial glycolipid agonists of TLR2 include phosphatidyl-(myo)-inositolmannosides (PIMs) lipomannans (LMs), lipoarabinomannans (LAMs), and inositol phosphate-capped LAMs (PI-LAMs) [159, 195, 196]. In this study we focus on the role of LprG in TLR2 activation with comparison to LprA, which is homologous to LprG; the function of LprG is unclear, but it may contribute to the function of the p55 small molecules transporter [103] and to virulence of Mtb [100].

An important determinant of TLR2-mediated recognition of lipoproteins is thought to be TLR2 binding the N-terminal diacylglycerol, in some cases accompanied by binding of an additional acyl chain by TLR1. Surprisingly, our preliminary studies showed that nonacylated (NA)-LprG retained TLR2 stimulatory capacity. Here we report detailed studies of a new mechanism by which LprG serves as a carrier of mycobacterial glycolipids, resulting in substantial enhancement of
glycolipid recognition by TLR2. We report the crystal structure of NA-LprG, which reveals a putative glycolipid binding pocket lined with hydrophobic residues; the dimensions of this pocket allow binding of triacylated glycolipids, and introduction of a single point mutation in this pocket blocks the glycolipid carrier function of LprG. These results present two novel interpretations. First, we propose that LprG functions in mycobacteria as a glycolipid carrier to chaperone glycolipids during their trafficking and delivery to the mycobacterial cell wall, contributing to virulence [100] [103] and providing potential opportunities for targeting in drug design. Second, we propose that the glycolipid carrier function of LprG facilitates recognition of triacylated glycolipids by TLR2, presenting a novel paradigm for recognition of hydrophobic TLR2 agonists that are not sufficiently soluble in an aqueous environment and must be chaperoned for efficient delivery to TLR2. Thus, host cells may co-opt the chaperone function of microbial carriers, providing a mechanism to enhance innate immune recognition of Mtb or other bacteria that express hydrophobic TLR2 agonists.
Materials and Methods

Cloning and expression of His$_6$-tagged proteins

LprA and NA-LprA were cloned previously [45]. LprG was amplified from Mtb H37Rv genomic DNA by PCR using the following oligodeoxynucleotide primers (sequences written 5’ to 3’; underlined portions are Ndel and HindIII restriction enzyme recognition sites): the 5’ primer GCATATCCATATGCGGACCCCCAGACGC ACTG and the 3’ primer GTACAAGCTTGCTCACGGGGGCTTCG. A non-acylated (NA) variant of LprG was cloned by using a 5’ primer that excluded the signal sequence and changed the acylated cysteine to a methionine. NA-LprG was cloned with the 5’ primer GCAATTCCATATGTCGTCCGCTTC and the 3’ primer GTACAAGCTTGCTACCGGGGCTTCG. Fusions of NA-LprA and NA-LprG were produced by digestion of the NA-Lpr constructs with Ndel, mscl, and HindIII, and ligating the 5’ fragment of NA-LprA with the 3’ fragment of NA-LprG to make the A:G fusion protein. For the G:A fusion, the same digest was performed, and 5’ NA-LprG was fused with the 3’ NA-LprA fragment. Site-directed mutagenesis of NA-LprG was performed using the Quikchange site directed mutagenesis kit (#200519; Stratagene, La Jolla, CA) with the 5’ primer GCCGCGACGGGAACCTGGAAGCTCACGCTGGGT and the 3’ primer ACCCAGCGTGAGCTTCCAGTTTCCCGTCGCGGC. For expression in M. smegmatis, constructs were digested with Ndel and HindIII (NEB, Ipswich, MA) and ligated into the shuttle vector pVV16 (provided by J. Belisle, Colorado State University, Fort Collins, CO) behind the constitutively active hsp60 promoter and in-frame with a C-
terminal His$_6$ tag. For expression in *E. coli* Rosetta cells (#71405, EMD, Gibbstown, NJ), constructs were digested with *NdeI* and *HindIII* and ligated with the expression plasmid pET-22b(+) (Novagen) (removing the pelB leader sequence), placing the coding sequence behind the IPTG-inducible T7 promoter and in frame with a C-terminal His$_6$ tag. All constructs were verified by sequencing and analyzed using Clone Manager (SciEd software, Cary, NC). For expression in *M. smegmatis*, strain MC$^2$ 1-2C (R. Wilkinson, Imperial College, London, U.K.) was transformed by electroporation with a Gene Pulser (Bio-Rad, Hercules, CA) set at 2.5 kV, 25 µF, and 800 Ohms. *M. smegmatis* was cultivated in Middlebrook 7H9 broth (Difco, Lawrence, KS) supplemented with 1% casamino acids (Fisher, Pittsburgh, PA, BP1424), 0.2% glycerol, 0.2% glucose, and 0.05% Tween 80; selection was with kanamycin at 30 µg/ml. For expression in *E. coli*, chemically competent *E. coli* Rosetta was transformed and cultured in Luria-Bertani broth (LB); selection was with ampicillin (50 µg/ml) and chloramphenicol (34 µg/ml). When culture OD$_{600}$ was approximately 1.0, gene expression was induced for 2-3 h by addition of 500 nM IPTG (Invitrogen, 15529-019). Bacteria were isolated by centrifugation at 6000 x $g$ for 20 min at 4°C.

**Purification of His$_6$-tagged proteins**

NA-LprG, NA-LprA and NA-LprG-V91W were stably expressed in both *E. coli* and *M. smegmatis*, and were purified as described [45]. Cells were resuspended in lysis buffer (2.5 ml/liter of bacterial culture) consisting of 50 mM NaH$_2$PO$_4$, 300 mM
NaCl, 20 mM imidazole, pH 8.0, 2.5% protease inhibitor cocktail (Sigma P8849), 75 U/ml benzonase (Novagen, Madison, WI, 70664-3), and 2.5 mg lysozyme (Sigma L-3790) and incubated for 15 min at 37°C. Bacteria were disrupted mechanically by 4 passages through a French press (2000 psi). Insoluble material was removed from the lysate by ultracentrifugation at 100,000 x g for 1 h at 4°C, and supernatant was incubated directly with Ni beads (Qiagen, Valencia, CA, 1018244) for 2-4 h at 4°C. Ni beads were transferred to polypropylene columns, washed 3x with 25 volumes of wash buffer (50 mM NaH$_2$PO$_4$, 1 M NaCl, 20 mM imidazole, 10% glycerol, pH 8.0), and bound protein was dissociated with elution buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl, 450 mM imidazole, pH 8.0). Samples were desalted into 20 mM Tris, pH 8.0 using PD-10 columns (GE Healthcare, Uppsala, Sweden 17-085-01), subjected to anion-exchange chromatography using quaternary ammonium columns (GE Healthcare, 17-5053-01), eluted by stepwise addition of 50, 150, 200 and 1000 mM NaCl, and concentrated using 10-kDa cutoff Centricon units (Amicon, UFC801008). Protein purity was verified by SDS-PAGE with silver stain and anti-His$_6$ Western blot; yields were estimated by BCA protein assay (Pierce, Rockford, IL, 23225). Material eluted at 50-200 mM NaCl was used for all experiments.

**Mycobacterial lysates and charging of E. coli-derived proteins**

Mtb strain H37Ra (ATCC 25177) was cultured with shaking at 37°C to late log phase growth (2.5 weeks) in Mtb 7H9 broth (4.7 g/l 7H9 (Difco 271310), 5 ml/l glycerol, 0.5 ml/l Tween-80 (Sigma, St. Louis, MO, P4780) supplemented with 10% albumin/dextrose/catalase (BD, Franklin Lakes, NJ 212352). Bacilli from 100 ml late log phase culture were harvested by centrifugation at 5,000 x g for 20 min at
4°C, resuspended in 5 ml culture supernatant and stored at -80°C. Frozen stocks of
*M. smegmatis* or Mtb H37Ra were thawed and used with or without resuspension in
5 ml HBSS to prepare lysates by sonication with a Misonix Sonicator 3000 (Misonix,
Farmingdale, NY) at amplitude 1 for four 15-min bursts in ice water with a
temperature cut-off of 40°C. Mtb H37Rv lysates were obtained from Karen Dobos-
Elder and John Belisle, Colorado State University, under NIH contract
HHSN266200400091C, N01-AI-40091. Since Mtb H37Rv lysates contained
EDTA, they were passed over a Ni spin column (Qiagen cat#31014) and eluted to
remove EDTA; MgCl₂ and CaCl₂ were added to 1.0 mM and 1.3 mM, respectively.

For charging of *E. coli*-expressed proteins, 300-500 µg of protein (purified by
Ni-affinity and anion-exchange chromatography) was incubated with mycobacterial
sonicate for 3 h at 37°C with rocking. Insoluble material was removed by
centrifugation at 100,000 x g, and charged proteins were repurified from the
supernatant by Ni-affinity and anion-exchange chromatography. For charging with
purified glycolipids, proteins similarly purified from *E. coli* (100 µg) were incubated
for 3 h at 37°C with 0.5-50 µg of purified glycolipid (Invivogen, San Diego, CA; LAM-
MS, LM-Ms) or lysis buffer in a total volume of 100 µl and repurified as above.

**SDS-PAGE and visualization of purified proteins and glycolipids**

Gels (13% acrylamide) were cast and run using a Tris-HCl buffer system.
Proteins were visualized with Silver Stain Plus (BioRad, 161-0449EDU) and
Western blot with mouse anti-His₆ monoclonal Ab (Santa Cruz, Santa Cruz, CA sc-
8036; 1:1000). Carbohydrates (including glycolipids) were visualized with Pro-Q
Emerald 300 (Molecular Probes, Eugene, OR P20495) following periodate oxidation. Mycobacterial proteins and glycolipids were also visualized by Western analysis with rabbit polyclonal anti-BCG (DAKO, Glostrup, Denmark; 1:30,000). For Western analysis, material was transferred to PVDF membranes, which were blocked with 5% milk in PBS supplemented with 0.1% Tween-20 (PBST) for 1 h at 22°C and incubated overnight at 4°C with antibody. Blots were washed three times in PBST and incubated for 2 h at room temperature with secondary antibodies conjugated to horseradish peroxidase. Membranes were washed three times in PBST, and reactive bands were visualized with enhanced chemiluminescence (GE Healthcare; RPN2106)).

*Mass spectrometry and identification of ligands of NA-LprG*

For nanospray mass spectrometry, 12 µg of purified protein was denatured in 500 µL methanol more than 5 min, followed by vortexing. Ten µL of methanol eluate was loaded onto an in-house made glass nanospray tip for negative-mode electrospray ionization mass spectrometry (LCQ Advantage, Thermo Finnigan, Ringoes, NJ). For LC-MS analysis, 4.3 picomoles of each protein was extracted in 500 µL methanol, dried under nitrogen, redissolved in 200 µL of HPLC mobile phase solution (95%A:5% B, see below), and loaded on to a monochrome diol column (46 mm x 250 mm, 3 µm; Varian Inc., Palo Alto, CA) and coupled on-line to a LXQ 2 dimensional ion-trap mass spectrometer (Thermo Finnigan) equipped with an electrospray ionization (ESI) source. The solvent system for mobile phase and gradient method for separation were slightly modified from a previous method
Briefly, mobile phase A was hexane:isopropanol (60:40, vol/vol) containing 0.1% (vol/vol) formic acid, 0.05% (vol/vol) ammonium hydroxide, and 0.05% (vol/vol) triethylamine. Mobile phase B was methanol containing 0.1% (vol/vol) formic acid, 0.05% (vol/vol) ammonium hydroxide, and 0.05% (vol/vol) triethylamine. At a flow rate of 0.7 ml/min, the binary gradient started at 5% mobile phase B, linearly increasing to 15% B in 6 min and then held at 15% B for 10 min, followed by linearly increasing to 95% B in 8 min, held at 95% mobile phase B for 6 min and finally adjusted back to 5% B in 2 min.

*Crystallization and determination of NA-LprG structure*

Purified NA-LprG formed crystals. The structure was solved by multiple-wavelength anomalous dispersion (MAD) and refined using native LprG to a resolution of 2.2 angstroms in the C2 space group. The final model had an $R_{\text{work}}$ of 22% and an $R_{\text{free}}$ of 26% with no outliers in the Ramachandran plot (data not shown). The final structure model was checked with Procheck, which verified that all residues were located in regions of allowed geometry.

Protein structure graphics were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (NIH P41 RR-01081; [199]). Modeling of the LprG-V91W mutant was done with Deepview 4.0 [200]. Predicted structures of LprA and the fusions between LprA and LprG were generated using CPH model v2.0 ([http://www.cbs.dtu.dk/services/CPHmodels/](http://www.cbs.dtu.dk/services/CPHmodels/)).
**Mammalian cell culture**

Unless otherwise specified, incubations with eukaryotic cells were performed at 37°C in 5% CO₂ atmosphere. Standard medium was DMEM (Hyclone, Logan, UT, ASK30773) supplemented with 10% heat-inactivated FCS, 50 μM 2-ME, 2 mM L-glutamine, 1 mM sodium pyruvate, 10 mM HEPES, pH 7.4, and penicillin/streptomycin (Hyclone). Stimulation medium was standard medium with serum concentration reduced to 0.2% FBS. Female C57BL/6J mice (8-16 weeks old) were obtained from the Jackson Laboratory, housed under specific pathogen-free conditions and used to produce macrophages. Bone marrow was obtained from TLR1−/−, TLR2−/− and TLR6−/− mice courtesy of Amy Hise, Case Western Reserve University; these mice were generously provided by Shizuo Akira (Research Institute for Microbial Disease, Osaka University, Osaka, Japan) and were back-crossed to C57BL/6J mice a minimum of eight times. CD14 knockout mice (B6.129S-Cd14tm1Frm/J) were obtained from the Jackson Laboratory and were compared to both C57BL/6J mice and F2 hybrids of C57BL/6J and 129sv. Bone marrow cells were cultured for 7-12 d in standard medium supplemented with 25% LADMAC cell-conditioned medium [184]. HEK293 cells stably expressing TLR2-YFP (HEK293.TLR2) were produced previously [185, 186]. HEK293 cells (ATCC CRL-1573) were stably transfected with the empty vector to produce a control HEK293.pcDNA3 cell line. Transfected HEK293 cell lines were maintained in HEK medium (DMEM supplemented with 10% heat-inactivated FCS (HyClone)) supplemented with ciprofloxacin (10 μg/ml) and geneticin (500 μg/ml). HEK293.TLR2-CD14 cells were purchased from Invivogen (293-htrlr2cd14) and
cultured in HEK medium supplemented with 100 µg/ml Normocin (Invivogen, 100 µg/ml Hygromycin B and 10 µg/ml Blasticidin.

**Cytokine ELISAs**

HEK293 cells were incubated in 96-well plates (20,000 cells/well) in 90 µl of stimulation medium for 5-8 h and then for an additional 16 h with or without TLR2 agonist. Supernatant IL-8 concentration was quantified by ELISA (R&D, Minneapolis, MN, DY208). Bone marrow-derived macrophages were incubated overnight at 100,000 cells/well in standard medium and then for 12 h in stimulation medium with or without agonist. Supernatants were collected and stored at -80°C. TNF-alpha in the supernatant was quantified by ELISA (BD Biosciences #558874, R&D DY410). The following synthetic TLR agonists were also used: FSL-1 (Invivogen, tlr1-fsl), Pam3CSK4 (Invivogen, tlr1-pms), and CpG-B ODN 1826 (5’-tcc atg acg ttc acg tt-3’ lot C44-05225-q1a) provided by Coley Pharmaceutical Group (Wellesly, MA).
**Results**

*LprG has TLR2 agonist activity independent of its N-terminal acylation*

To explore structure-function relationships, we compared TLR2 agonist activity of two homologous Mtb lipoproteins, LprG and LprA. LprG stimulated TLR2-dependent IL-8 secretion more potently than LprA in a bioassay with TLR2-transfected HEK293.TLR2 cells (Fig. 3.1A). Because acylation of these lipoproteins is mediated by a common enzymatic pathway [85], suggesting that they are similarly acylated, we investigated whether the differing protein structures of these lipoproteins could affect TLR2 agonist activity. We designed non-acylated (NA) forms with the leader peptide removed and the N-terminal cysteine replaced with methionine. Recombinant His$_6$-tagged acylated and non-acylated versions of LprA and LprG were expressed and assessed for TLR2 agonist activity. Consistent with prior data [45], NA-LprA did not induce TLR2-dependent responses, indicating an essential role for acylation of LprA in its TLR2 agonist activity (Fig. 3.1A upper panel). In contrast, NA-LprG retained significant TLR2 agonist activity that was less potent than acylated LprG but similar to acylated LprA (Fig. 3.1A, upper panel). This surprising result suggested that LprG possessed a novel determinant of TLR2 agonism, independent of acylation of its N-terminal cysteine.

Interestingly, overexpression of CD14 in addition to TLR2 (in HEK293.TLR2-CD14 cells) enhanced the apparent potency of NA-LprG but not acylated LprG or LprA (Fig. 3.1A lower panel). This observation is consistent with other evidence that the mechanism for recognition of the NA-LprG-associated TLR2 agonist activity
is different than for acylated LprG and LprA, and foreshadows a parallel with recognition of glycolipids (below).
Figure 3.1. **NA-LprG carries a mycobacterial TLR2 agonist.** (A) HEK293.TLR2 cells show a dose-dependent IL-8 response to LprA, LprG, and NA-LprG, but no response to NA-LprA. Top panel, bioassay with HEK293.TLR2 cells. Bottom panel, bioassay with HEK293.TLR2-CD14 cells. Control HEK293 cells lacking TLR2 and CD14 failed to respond to all four proteins, similar to the response seen to NA-LprA in these panels (data not shown). (B) Alignment of primary amino acid sequences of LprG (upper) and LprA (lower). Shared residues are shown in bold font. The cysteine that is acylated in mature LprG or LprA is shown in green. Residues that contribute to the ligand binding pocket defined in later studies are shown in red. Residues below arrows contribute to secondary structure indicated alpha=helix, beta=strand. (C, D) NA-LprG can acquire TLR2 agonist activity from mycobacterial lysates. NA-LprG and NA-LprA were expressed in *E. coli*, purified by Ni-affinity and anion-exchange chromatography, incubated with control buffer or a lysate of *M. smegmatis* (E), *Mtb* H37Ra (F) or *Mtb* H37Rv (F), repurified by Ni-affinity (E) or Ni-affinity and ion exchange chromatography (F), and incubated with HEK293.TLR2 cells for 12 h. For all data panels, IL-8 production was quantified by ELISA, and data are reported as the mean +/- SD of triplicate HEK293.TLR2 assays. Results are representative of at least 3 independent experiments. Figure produced with permission using data published in *Nature Structural and Molecular Biology* [189].
To test whether specific peptide sequences or domains were required for TLR2 agonist activity of NA-LprG, we produced N-terminal truncation mutants of NA-LprG. Truncations within the first 14 amino acids did not affect TLR2 agonist activity, and further truncation involving residues that form the first alpha helix produced instability and poor expression (data not shown). Since LprG and LprA are homologous (Fig. 3.1B), and LprA lacks acyl-independent TLR2 activity, an alternative strategy was to produce full-length chimeric molecules containing different portions of LprG and LprA to determine which portions of LprG conferred TLR2 activity. Chimeric molecules were produced containing the N-terminal half of NA-LprA and the C-terminal half of NA-LprG, or vice versa, and tested for TLR2 agonist activity with HEK293.TLR2 cells (Fig. 3.S-1). Both chimeric molecules were active, suggesting that both halves of the molecule contributed to agonism, perhaps through a conformational determinant.
Figure 3.S-1. TLR2 activity of NA-LprG/NA-LprA chimeric molecules. (A) Structure model of LprA:LprG and LprG: LprA. (B,C) HEK293.TLR2 cells were incubated for 12 h with NA-LprG, NA-LprA or chimeric proteins combining the N-terminal half (NTD) and C-terminal half (CTD) of NA-LprG (green) and NA-LprA (red).

LprG binds a mycobacterial TLR2 agonist and delivers it for recognition by TLR2

LprG belongs to the same lipoprotein family as LppX, which binds and transports hydrophobic phthiocerol dimycerosates to culture supernatant [96]. A hydrophobic cavity was predicted in LprG when its structure was predicted using LppX as a template (data not shown). We hypothesized that LprG might activate TLR2 by non-covalently binding a lipid agonist via a conformation-dependent
binding site. This hypothesis was supported by the observation that NA-LprG had reduced activity when expressed in *E. coli* instead of *M. smegmatis* (Fig. 3.1C), suggesting that TLR2 agonist(s) carried by LprG may be present in mycobacteria but absent in *E. coli*. Further suggesting non-covalent association of a mycobacterial product, the activity of NA-LprG purified from *E. coli* was significantly increased following incubation with lysates of either *M. smegmatis* (Fig. 3.1C), Mtb H37Ra (Fig. 3.1D) or Mtb H37Rv (Fig. 3.1D). These results suggest that NA-LprG binds a mycobacterial TLR2 agonist and delivers it for recognition by TLR2.

*The crystal structure of LprG reveals a hydrophobic binding pocket that serves as the binding site for TLR2 agonists*

To gain insight into the structural basis of TLR2 signaling by LprG, the crystal structure of non-acylated LprG (NA-LprG) was determined. The refined apo structure contained 194 residues of the total 210 residues and water molecules. Though the crystallographic unit contained two molecules of LprG (termed molecule A and molecule B), size exclusion chromatography demonstrated that LprG exists as a monomer in aqueous solution (data not shown). The buried surface between these two molecules is considered physiologically insignificant.
Figure 3.2. Crystal structure of NA-LprG reveals a hydrophobic pocket with the potential to carry a TLR2 agonist. (A) LprG, front view. (B) LprA model based on LprG structure, front view. (C) LppX, front view [96]. (D) LprG front view, surface hydrophobicity (orange=hydrophobic, blue=hydrophilic). (E) LprA front view, surface hydrophobicity (modeled on LprG). (F) LppX surface hydrophobicity. (G) LprG pocket view, surface hydrophobicity. (H) LprA pocket view, surface hydrophobicity (modeled on LprG). Figure produced with permission using data published in Nature Structural and Molecular Biology [189]
LprG is a single-domain protein with an alpha/beta fold consisting of a 10-strand anti-parallel beta sheet (beta 1-10) and six alpha-helices (alpha 1-6) attached to the concave face of the beta sheet, similar to LppX (Fig. 3.2A,C) [96]. Like LppX, LprG possesses a large hydrophobic cavity (Fig. 3.2G) with approximate dimensions of 9 x 14 x25 angstroms, and a volume of ~1535 angstroms$^3$, approximately 1300 angstroms$^3$ smaller in volume than that of LppX (Figs 3.2G and 3.2I). The cavity maintained a diameter of 9 angstroms, which narrowed to two small sub-pockets about 14 angstroms from the entrance (Fig. 3.2G, 3.S-2D). The N-terminus of LprG defines the back wall of the cavity, while the C-terminus defines the front surface of the cavity (Fig. 3.S-2F). While the beta sheet of the upper part of the protein is concave to accommodate the cavity, the lower part of the beta-sheet is flat and defines the terminus of the binding cavity. Most residues inside the large binding pocket are hydrophobic, implying that the ligand is also hydrophobic. The cavity is only accessible from one direction, close to the junction between beta strands 3 and 4. The entrance of the cavity is approximately 9 x 20 angstroms, and is surrounded by hydrophobic residues and three polar residues; Tyr130, Asp128, and Ser72.
While LprG and LppX share 28% identity scattered throughout the amino acid sequence, the sequence conservation is not equally distributed throughout the tertiary structure (Fig 3.S-3A). The conserved residues between LppX and LprG are located distant from the LprG pocket entrance. There are also very few conserved residues in helices alpha 2, alpha 3, and alpha 4 of LprG, which define the front
surface of the pocket (*Fig. 3.5-3A*). In LppX, the residues corresponding to alpha 3 and alpha 4 of LprG form a flexible loop (*Fig. 3.2C,F,I*) that is shifted outward from the cavity and creates the opening to a large binding pocket. When LppX was superimposed on LprG, helix 2 of LprG conflicted with the C17 fatty acid located in LppX (*Fig. 3.5-3C*), and helices 3 and 4 of LprG defined a narrower pocket and occupied the location of the entrance of LppX. Thus the conformation of this region affected both the location of the pocket entrance and the cavity size. Similar to LprG, the pocket opening of LppX is surrounded by hydrophobic residues with only a few polar residues including Lys36, Ser107 and Ser110.
Figure 3.S-3. Comparison of LprG with LprA and LppX. (A) Ribbon of LprG with position of residues (orange) conserved between LprG and LppX. (B) Position of residues conserved between LprG and LprA. (C) Ribbon overlay of LprG (navy) and LppX (light blue) with lipids from LppX structure shown in yellow. (D) Ribbon overlay of LprG (blue) and model of LprA (copper). Images reproduced with permission from Nature Structural and Molecular Biology [189].
LprG and LprA share 34% identity in amino acid sequence, more than that of LprG and LppX, and the conserved residues are distributed more evenly throughout the tertiary structure of the protein (Fig 3.5-3B). Using MODELLER, LprA was predicted to adopt an alpha/beta fold similar to LprG and LppX (Fig. 3.2B). Compared to LppX, LprA has more residues conserved with LprG located in the pocket and at the entrance, suggesting that the cavity of LprA may be more similar in shape and location to LprG than that of LppX (Fig 3.2H, 3.5-2). The cavity of the LprA model has a volume of 1700 \( \text{Å}^3 \), and while large enough to accommodate multiple alkyl chains, has a narrower opening (Fig. 5.2)

To test the hypothesis that the hydrophobic pocket of LprG serves as a binding site of TLR2 agonists, we performed site-directed mutagenesis to create a NA-LprG-V91W mutant, replacing a valine at the entrance to the hydrophobic pocket with a bulky tryptophan. X-ray crystallography of the mutant demonstrated that this single amino acid change decreased the pore diameter by approximately 3.5 angstroms (Fig 3.3A-E). NA-LprG-V91W was stably expressed in *M. smegmatis*, and its ability to activate TLR2 was significantly reduced relative to NA-LprG expressed in *M. smegmatis* (Fig. 3.3F). Mutations at two other sites in the hydrophobic pocket (V194R and V217F) also decreased TLR2 agonist activity of NA-LprG (data not shown). Furthermore, when NA-LprG-V91W was expressed in *E. coli*, purified and then incubated with a sonicate of *M. smegmatis* or Mtb, the V91W mutant lacked the ability to acquire TLR2 agonist activity from mycobacteria (Fig. 3.3G). These results provide strong support for the hypothesis that the
hydrophobic cavity serves as a binding site for TLR2 agonist(s), which are then delivered by LprG for recognition by TLR2.
Figure 3.3. Single amino acid change alters TLR2 agonist activity. (A) Overlay of LprG (brown) and LprG-V91W (tan) ribbon structures with side chain of 91W shown in orange. (B). LprG surface hydrophobicity, pore view with V91 in yellow. (C). LprG-V91W with W91 in yellow. (D) LprG pocket entrance with V91 in orange. (E) LprG pocket entrance with mutated W91 in red. (F) TLR2 Activity of NA-LprG and NA-LprG-V91W expressed in M. smegmatis and tested on HEK293.TLR2 cells (as in Fig. 1E). (G) TLR2 activity of NA-LprG and NA-LprG-V91W expressed in E. coli, purified, incubated with Mtb H37Ra lysate and repurified (as in Fig. 1F; data for wild-type E. coli-derived NA-LprG are from Fig. 1F). Data are reported as mean +/- SD of triplicate HEK293.TLR2 assays. Results are representative of at least 3 independent experiments. Figure produced with permission using data published in Nature Structural and Molecular Biology [189]
**TLR2 agonists associated with NA-LprG have recognition requirements similar to triacylated mycobacterial glycolipids**

The dependence of TLR2 agonists on TLR1 vs. TLR6 is determined by agonist structures [197]. Therefore, we investigated properties of TLR2 agonist(s) putatively associated with NA-LprG by testing NA-LprG signaling in macrophages from mice genetically deficient in TLR1, TLR2 or TLR6. All macrophage types responded to CpG ODN 1826, a TLR9 agonist, by secreting TNF-alpha, confirming responsiveness of these cells to other TLR stimuli (data not shown). Recognition of NA-LprG or associated molecules was deficient in TLR2−/− and TLR1−/− macrophages but not TLR6−/− macrophages (Fig. 3.4A). Similar results were observed with NA-LprG that was expressed in *E. coli* and then incubated with Mtb lysate to allow loading of NA-LprG with Mtb-derived TLR2 agonist(s) (Fig. 3.4B). These results suggest that the activity associated with NA-LprG signals through TLR2/TLR1 heterodimers, a pattern that has been observed with mycobacterial glycolipids including LAM (Fig. 3.4C) [134] and LM (Fig. 3.4D) [194], as well as triacylated lipopeptide (Fig. 3.4F). In contrast, cytokine production by the diacylated lipopeptide FSL-1 was dependent on TLR2 and TLR6, but not TLR1 (Fig. 3.4E). Other studies with macrophages from CD14−/− mice showed that NA-LprG activity was dependent on CD14 (Appendix 1), consistent with observations that CD14 contributes to TLR2 recognition of triacylated lipopeptides [201] and glycolipids [202]. The dependence of NA-LprG and Mtb glycolipid signaling on TLR2/TLR1 and CD14 suggests that putative TLR2 agonist(s) associated with NA-LprG have recognition requirements similar to triacylated mycobacterial glycolipids.
Figure 3.4. NA-LprG activity is dependent on TLR1 and TLR2. (A, B) Macrophage responses to NA-LprG are dependent on TLR2 and TLR1 but not TLR6. Bone marrow-derived macrophages from TLR2−/−, TLR1−/−, TLR6−/− or wild-type mice were incubated for 12-16 h in stimulation medium with NA-LprG, glycolipid or lipopeptide, and TNF-alpha production was determined by ELISA. (A) NA-LprG was purified by Ni-affinity and anion-exchange chromatography from M. smegmatis. (B-D) NA-LprG was purified from E. coli by Ni-affinity and anion-exchange chromatography, incubated with Mtb H37Ra lysate (B), LAM from M. smegmatis (C) or LM from M. smegmatis (D), and then repurified by Ni-affinity and ion exchange chromatography. (E) Response to FSL-1 lipopeptide (model TLR2/TLR6 agonist). (F) Response to Pam3CSK4 lipopeptide (model TLR2/TLR1 agonist). Response to a TLR9 agonist (CpG ODN 1826 TLR9 agonist) was intact in all knockout cells (data not shown). Data are reported as mean +/- SD of triplicate macrophage assays. Results are representative of at least 3 independent experiments (one for panels C and D). Figure reproduced with permission from Nature Structural and Molecular Biology [189].
Mycobacterial glycolipids are associated with LprG

To directly identify molecules that non-covalently associate with NA-LprG, proteins were purified from *M. smegmatis* by Ni-affinity followed by anion-exchange chromatography and subjected to SDS-PAGE with silver stain, Pro-Q stain for carbohydrates following periodate oxidation, or Western blot using a monoclonal anti-His\(_6\) antibody, monoclonal anti-LAM (CS-35) [203], or polyclonal anti-*M. bovis* BCG antibody that recognizes many components of BCG and Mtb (Fig. 3.5). Silver stain and anti-His\(_6\) Western blot of protein preparations showed an isolated band at approximately 24 kDa derived from NA-LprG, NA-LprG-V91W or NA-LprA (Fig. 3.5A, D). Staining of carbohydrates after periodate oxidation showed bands with apparent molecular weights of 25-30 kDa, 14-18 kDa and <10 kDa (Fig. 3.5C), which correspond to the apparent molecular weights by SDS-PAGE of Mtb glycolipids LAM, LM and PIM, respectively (confirmed by running glycolipid standards). These compounds are likely glycolipids because they were seen after periodate oxidation but not on conventional silver stain, and glycolipids in the PIM-LAM series resolve as broad bands based on heterogeneity of the glycan component of each molecular species. Western blot with polyclonal anti-BCG antibody detected NA-LprG and bands with apparent molecular weights consistent with the LAM and LM (Fig. 3.5C). Importantly, the putative glycolipid bands corresponding to LAM and LM were associated with NA-LprG but not NA-LprG-V91W or NA-LprA, indicating that these glycolipid agonists of TLR2 are associated preferentially with NA-LprG. The identity of the putative LAM band was verified with CS-35 (Fig. 3.5D), a monoclonal antibody specific for the tetra-arabinose motif present in LAM [203]. The
carbohydrate stain also revealed clear association of PIM with NA-LprG, lesser association of PIM with NA-LprA, and only minimal association of PIM with NA-LprG-V91W. Western blot revealed PIM staining of NA-LprG but not NA-LprA or NA-LprG-V91W (although the anti-BCG staining may detect only a subset of PIM species). These results provided evidence for preferential association of Mtb glycolipids with NA-LprG.

**Figure 3.5. Mycobacterial glycolipids are associated with NA-LprG.** SDS-PAGE analysis of proteins purified from *M. smegmatis*. Samples were visualized by silver stain (A), Pro-Q stain for carbohydrates (B), polyclonal anti-BCG Western blot (C), and a combination of anti-His (red) and anti-LAM (green) monoclonal antibodies. Arrows on right that are labeled LAM, LM and PIM represent the positions observed with glycolipid standards run in this gel system. Figure reproduced with permission from Nature Structural and Molecular Biology [189].

LAM, LM, and PIM share a common structural core, suggesting that all of these mycobacterial glycolipids might associate with NA-LprG via a shared
structural motif, which may also contribute to TLR2/TLR1 agonist activity. Therefore, to directly determine the molecular structures of small molecules associated with LprG and related proteins, we treated proteins with methanol to denature and solubilize small molecules non-covalently associated with the protein. We then analyzed methanol eluates with nanoelectrospray ionization mass spectrometry to detect compounds with a mass to charge (m/z) ratio up to 2000. Methanol alone (Fig. 3.6A) or methanol elutes of an unrelated protein Pab C (Rv0812) (Fig. 3.5-S-4) did not give detectable ions above the level of randomly detected signal. In contrast, NA-LprG, NA-LprG-V91W, and NA-LprA (Fig. 3.6) expressed in M. smegmatis yielded ions corresponding to mycobacterial phospholipids. Ions detected at m/z 851.4, 1013.5, 1175.5 and 1413.7 corresponded in mass to [M-H]- of diacyl phosphatidylinositol, diacyl phosphatidylinositol monomannoside (PIM₁), diacyl phosphatidylinositol dimannoside (PIM₂) and triacyl PIM₂ (Ac₁PIM₂), respectively. Collision-induced dissociation mass spectrometry (CID-MS) analysis of these compounds yielded product ions expected from these initial structures, confirming the tentative assignments based on mass (Fig. 3.6A inset and Fig. 3.5-S-5). For example, triacyl PIM₂ (Ac₁PIM₂) (m/z 1413.7) yielded products corresponding to the loss of mannose (m/z 1251), loss of acyl mannose (m/z 1013), loss of C16:0 acyl (m/z 1157), loss of C19:0 fatty acyl (m/z 1115) and acyl phosphoinositol dimannoside (m/z 803). Thus, analysis of compounds of molecular weight under 2000 amu showed that LprG binds at least four structurally related molecules that all contain
phosphatidylinositol as the core structure, but differ in the number of mannose units and fatty acyl chains.
Figure 3.6. Triacylated Ac₁PIM₂ is specifically associated with NA-LprG. (A) Negative mode electrospray ionization mass spectrometry of methanol treated NA-LprG yielded ions corresponding in mass to mycobacterial phospholipids. Ions of m/z 851.5, 1013.5 and 1413.7 were subjected to negative mode collision induced dissociation mass spectrometry, yielding ions that correspond to the masses and fragments that are detailed in Fig 3.S-5 and shown in summary in the insets. Ions detected near m/z 823.5 correspond to an alternately acylated form of phosphatidylinositol, and ions detected near m/z 949.9 corresponds to an H₃PO₄ adduction of the phosphatidylinositol. Ions detected near 1175.5 correspond to diacylated PIM2. (B-E) LC-MS analysis was carried out on 4.3 picomoles of NA-LprG (red), NA-LprG-V91W (blue), NA-LprA (green) and solvent blank (black). The total ion current trace shows that signals from non-lipidic components of protein preparations were detected at similar levels, serving as a control for equivalent loading of proteins onto the column (B). Mass chromatograms measured in narrow mass ranges corresponding to the masses of phosphatidylinositol (m/z 851.5) (C) diacyl PIM₁ (m/z 1013.6) (D), and triacyl Ac₁PIM₂ (m/z 1413.76) (E) are shown. Figure produced with permission using data published in Nature Structural and Molecular Biology [189].
In comparative liquid chromatography-mass spectrometry (LC-MS) analysis, it was possible to semi-quantitatively measure mass spectral signals corresponding to individual molecular species within the eluents from the same molar quantity of NA-LprG, NA-LprG-V91W, and NA-LprA (Fig. 3.6B-E). This was accomplished by using signals from protein preparations as a loading control (Fig. 3.6B) and simultaneously monitoring signals in narrow ranges (+/- 0.5 amu) of the expected masses of diacyl PI (Fig. 3.6C), diacyl PIM₁ (Fig. 3.6D) and triacyl PIM₂ (Ac₁PIM₂) (Fig. 3.6E). Diacyl PI and diacyl PIMs, were detected in association with NA-LprG and NA-LprA, but only in lower amounts with NA-LprG-V91W. In contrast, the triacylated molecule Ac₁PIM₂ was only detected in NA-LprG (Fig. 3.6E). We conclude that LprG has binding specificity for lipids in the PI-PIM series and that these lipids bind at or near the tryptophan residue in the binding pocket. Further, whereas LprA binds only diacylated PIMs, LprG can bind both diacylated and triacylated versions of this molecule. Other studies show that a large proportion of naturally occurring LAM is triacylated [204]. Therefore, the ability of LprG to bind triacylated lipids can also explain the association of glycolipids with LprG seen by PAGE analysis (Fig. 3.5). Specifically, this result may explain why LAM is associated with LprG but not LprA, and why the signal for PIM is greater in association with LprG than LprA (Fig. 3.5B-C).
Figure 3.S-4. Total spectra of methanol eluates of NA-LprG and PabC.

Figure reproduced with permission from Nature Structural and Molecular Biology [189].
Figure 3.S-5. Collision-induced dissociation mass spectrometry analysis of compounds detected in Fig 3.6. Breakdown products of PI (A), PIM<sub>1</sub> (B) and Ac<sub>1</sub>PIM<sub>2</sub> (C) verify the tentative identities assigned by mass. Figure reproduced with permission from Nature Structural and Molecular Biology [189].
Mycobacterial glycolipids bind inside the hydrophobic pocket of LprG.

To understand the mechanism of glycolipid binding to LprG, co-crystalization studies of glycolipid and NA-LprG were performed. A mixture of phosphatidylinositol mannoside and phosphatididylinositol was incubated with NA-LprG at a molar ratio of 1:1, and crystals were obtained with the same hit condition used to solve the structure of NA-LprG. The resulting structure was solved to 1.8 angstroms in the C2 space group, with an $R_{\text{work}}$ and $R_{\text{free}}$ of 22% and 26% respectively. The final refined structure including all residues was within allowed geometry, with the general structure remaining the same as the NA-LprG structure. Similar to previous results, NA-LprG formed an asymmetric crystallographic unit with two molecules of NA-LprG. While molecule B of the structure was empty, molecule A of the structure bound triacylated PIM$_2$ (Ac$_1$PIM$_2$) via insertion of three acyl chains into the hydrophobic pocket, with the glycans exposed at the surface (Fig. 3.7).
Figure 3.7. Structure of PIM bound to NA-LprG  (A) Surface hydrophobicity of NA-LprG showing surface of PIM in pocket with sugars exposed.  (B) Pocket view showing molecular structure of Ac$_1$PIM$_2$ in the pocket.  (C) View of Ac$1$PIM$2$ in pocket of NA-LprG showing that acyl chains of Ac$_1$PIM$_2$ fill the pocket opening, and glycans associate with polar residues near the pocket entrance.  (D) V91W point mutation directly conflicts with PIM in pocket.  Surface of V91 shown in yellow, surface of W91 shown in red.  Figure produced with permission using data published in Nature Structural and Molecular Biology [189].
The crystal structure of NA-LprG-V91W demonstrates conflict between the point mutation and the acyl chain of Ac_1PIM_2.

To investigate the reason why NA-LprG-V91W has lower TLR2 activity, the NA-LprG-V91W structure was solved to 2.3 angstroms. The mutation of valine to tryptophan caused the pocket wall to shift by 3.5 angstroms at Val91 (Fig. 3.3A). Because a small residue was mutated into a large bulky residue, the pocket in the Leu73 to Leu76 region bulged into the cavity, narrowing it. The cavity size of NA-LprG-V91W calculated by castP is ~1190 angstroms³, slightly smaller than that of NA-LprG (1535 angstroms³). There are conformational changes in the two loop regions near the entrance. The loop between beta 3 and beta 4 was pushed away from the entrance. The loop between beta 5 and alpha 2 was also moved away from entrance to form an extra beta strand in the mutant NA-LprG-V91W structure. Some residues contributing to the cavity entrance in wild-type NA-LprG were shifted in the point mutant (Fig. 5.1). Because of the movement of the two loops and the tryptophan mutation, the size of the pocket entrance became 8 Å x 13 Å, smaller than the wild-type NA-LprG. At the location of the point mutation, the channel diameter was decreased by about 3.5 angstroms.

To determine whether the V91W caused structural rearrangements that would conflict with the binding of PIM, the structures of NA-LprG-PIM and NA-LprG-V91W were superimposed. This analysis revealed that the bulky tryptophan side-chain directly conflicted with the acyl chain of Ac_1PIM_2 (Fig. 3.7D). This structural
data is consistent with the MS and SDS-PAGE data that suggest the single amino acid mutation abrogates the ability of LprG to bind PI-based glycolipids.

*LprG enhances TLR2 recognition of mycobacterial glycolipids*

To directly test the ability of NA-LprG to bind specific candidate glycolipids and deliver them for recognition by TLR2, we purified NA-LprG from *E. coli* (with little or no TLR2 activity), incubated it with LM or LAM, repurified it by Ni-affinity and anion-exchange chromatography, and tested it for TLR2 agonist activity. NA-LprG was able to bind LM and LAM and deliver them for recognition by TLR2 (Fig. 3.8). In contrast, NA-LprG-V91W (Fig. 3.8) and NA-LprA (data not shown) lacked the ability to bind and deliver LM and LAM for TLR2 recognition, consistent with prior evidence for their specific association of with NA-LprG (Fig. 3.5). Importantly, association with NA-LprG enhanced the apparent potency of LAM and LM by at least 1.5 log orders of magnitude (Fig. 3.8). Of note, this comparison is based on the molar concentrations of NA-LprG shown in Fig. 3.8, assuming that 100% of the NA-LprG molecules were loaded with LM or LAM; in the likelihood that glycolipid loading was less than 100% efficient, the factor by which association with NA-LprG enhanced glycolipid potency would be greater than estimated. We conclude that NA-LprG binds mycobacterial glycolipids and enhances their recognition by TLR2.
Our studies indicate the existence of two determinants of the TLR2 agonist activity of LprG, including N-terminal acylation and chaperoned glycolipids. The studies indicated above were performed largely with NA-LprG in order to reveal the chaperoned TLR2 agonist activity. One question is what proportion of the total TLR2 agonist activity of LprG is provided by chaperoned glycolipids. Acylated LprG-V91W was found to be significantly less potent than wild-type acylated LprG (Fig. 3.8).
dose response studies showed that the V91W mutation reduced TLR2 potency of acylated LprG by a log or more. Acylated LprG-V91W had potency similar to acylated LprA (Fig. 3.9), indicating that the TLR2 activity remaining when glycolipids were not associated with LprG was similar to that of LprA, perhaps reflecting similar N-terminal acylation of these molecules. These results indicate that the higher potency of LprG relative to LprA results from the TLR2 agonist activity of LprG-associated glycolipids. Thus, TLR2 agonist activity of LprG reflects significant contributions by glycolipid TLR2 agonists chaperoned by LprG as well as acylation of LprG itself.

Figure 3.9. The glycolipid binding site contributes significantly to the total TLR2 agonist activity of acylated LprG. HEK293.TLR2-CD14 cells were used to assess TLR2 agonist activity of acylated LprG, LprA, and LprG-V91W as in Fig. 1. Data are reported as mean +/- SD of triplicate HEK293.TLR2-CD14 assays. Similar results were obtained in 4 independent experiments. Figure reproduced with permission from Nature Structural and Molecular Biology [189].


**Discussion**

Previous studies have demonstrated that TLR2 agonist activity of bacterial lipoproteins is dependent upon acylation of the N-terminal cysteine, as demonstrated for OspA [148, 149], LpqH [148] and LprA [45]. In contrast, we report that LprG retains TLR2 agonist activity in the absence of acylation (Fig. 1A). While NA-LprG expressed in *E. coli* had minimal TLR2 activity, it gained TLR2 activity after exposure to mycobacterial lysate (Fig. 1E, F) or mycobacterial glycolipids (Fig. 8). Moreover, Western blot analysis revealed association of mycobacterial glycolipids with LprG. Mass spectrometry revealed that diacylated glycolipids were associated with both LprA and LprG, whereas triacylated glycolipids were associated preferentially with LprG. These studies reveal that LprG binds mycobacterial glycolipid TLR2 agonists and delivers them for recognition by TLR2.

X-ray crystallography studies revealed the structure of LprG, yielding information about potential conformations, ligands, and the basis of TLR2 activity. The crystallographic unit was an asymmetric dimer, and the differences of the two molecules within the unit may reflect physiologically important conformation states of LprG. Molecule A had a larger pocket and was characterized by more helical content. Specifically, Helix α3 in molecule A was a loop in molecule B, and the respective amino acids (Phe123 to Gln134) were shifted toward the cavity, reducing the cavity volume of molecule B by 400 angstroms³. These differences may have precluded PIM binding to molecule B. These differences are reflected in the position of Tyr130. In molecule A, Tyr130 is out of the pocket, while in molecule B, Try130 is...
in the pocket, narrowing the channel. The loop between beta 3 and beta 4 was moved outward from the protein, widening the entrance to the cavity in molecule B. Molecule A was used to represent the structure of LprG in this thesis.

The pocket was large enough to fit three alkyl chains of the PI-based glycolipids and the entrance was able to hold the remainder of the Ac₁PIM₂. The size of the binding pocket was reduced upon mutation of Val91 in the pocket to a bulky tryptophan. The V91W point mutation blocked one of the alkyl chain binding sites (Fig. 3.7). In the LC-MS study, NA-LprG was associated with several cell wall lipid precursors but mutant NA-LprG-V91W only lost its affinity for triacylated glycolipid. Interestingly, this point mutation in the pocket also disrupted TLR2 agonist activity. Overall, crystallographic structures support the hypothesis that the pocket in LprG is a binding site for triacylated glycolipid and establishes LprG as a glycolipid carrier that delivers triacylated glycolipid for recognition by TLR2.

Helices alpha 2, alpha 3, and alpha 4 in LprG are critical distinguish the cavity size and shape from LppX. In the sequence alignment of LppX with LprG, most conserved residues were located distant from the pocket of LprG. However, LprA had more conserved residues around the pocket. Helix alpha 4 in LprG is well conserved in LprA, which might help to maintain the pocket entrance in the same position as that of LprG. The binding pocket of model LprA obtained from MODELLER is lined with hydrophobic residues of the same size as the residues in the LprG binding pocket. The pocket size of the LprA model can theoretically accommodate a triacylated structure such as Ac₁PIM₂, as determined by
superimposing the LprA model and the LprG-PIM structure. Thus, while LprA is potentially another lipid carrier for TLR2 recognition in mycobacteria, it does not carry triacylated.

X-ray crystallography of the NA-LprG-V91W point mutant demonstrated that the tryptophan side directly conflicts with binding of Ac_1PI M_2. This point mutant is also markedly decreased in TLR2 activity, suggesting that triacylated glycolipids bound in the pocket of LprG are responsible for a portion of its TLR2 activity, and entirely responsible for the TLR2 activity present in NA-LprG. Overall, these results establish that LprG is a glycolipid carrier that can deliver its cargo for recognition by TLR2.

While these studies focused primarily on the role of LprG as a carrier of agonists for TLR2 activation with implications for host defense, our findings also have implications for the function of LprG in bacterial physiology. LprG is widely present in mycobacteria with 100% sequence identity among Mtb complex species and significant sequence conservation in other mycobacterial species. LprG (Rv1411c) forms an operon with the transmembrane protein p55 (Rv1410c) [98], a putative drug antiporter [99]. The operon is important for virulence [100], and the proteins may function together in construction of the mycobacterial envelope [103]. We propose that LprG serves as a carrier to facilitate trafficking of glycolipids from their site of synthesis at the plasma membrane to their location in the outer membrane of the bacterium [205, 206]. This mechanism is similar to that proposed for LppX in transport of pDIMs to the outer membrane of mycobacteria [96] and the
role of lipoproteins in *E. coli* in transport of membrane components, including other lipoproteins [207] and LPS [208, 209]. Since glycolipids have important roles in bacterial physiology and host-pathogen interactions, a potential future goal is inhibition of these processes by targeting the pocket of LprG with chemotherapeutics.

In addition to its function in bacterial physiology, our results suggest a role for LprG in facilitating TLR2 recognition of triacylated Mtb glycolipids. LprG was unique among the studied proteins in ability to bind triacylated PIM$_2$ (Ac$_1$PIM$_2$) as seen in LC-MS studies (Fig. 3.6), and this may correspond to the specificity of LprG for the larger and more heterogeneous LM and LAM molecules (Fig. 5) that may contain a triacylated PIM core structure. LAM is a complex polymer with a high molecular weight and cannot be analyzed by the MS systems reported here, but other studies show that a large proportion of naturally occurring LAM is triacylated [204]. Therefore, the preferential association of LAM with LprG as compared with LprA (Fig. 3.5) may be explained by acyl chain interactions. These findings are consistent with the observed receptor dependence of Mtb glycolipid signaling [134, 159, 193, 194] and the concept that TLR2 activity of mycobacterial glycolipids requires their triacylation or tetraacylation [160, 210]. The capacity of the LprG binding pocket to accommodate lipids with up to three alkyl chains (Fig. 3.7) may confer its ability to chaperone TLR2/TLR1 agonist activity of triacylated mycobacterial glycolipids. Although LprG was distinct from LprA in the ability to chaperone TLR2 agonists, other mycobacterial lipoproteins may be carriers of hydrophobic ligands, some of which may be TLR2 agonists. Thus, the potential
contribution of chaperoned molecules to TLR2 activity of other lipoproteins should be considered.

Glycolipids chaperoned by NA-LprG have potency similar to that of lipoprotein TLR2 agonists, e.g. the potency of NA-LprG is similar to the potency of acylated LprA (responses detected at <30 nM NA-LprG). In contrast, unchaperoned “naked” glycolipids were far less potent (responses detected at 20 µg/ml, or approximately 1 µM) and efficacious (i.e. produced a lower maximum cytokine response). Thus, association of glycolipids with NA-LprG increased their apparent potency by a factor of 1.5-2 log; moreover, this calculation assumes that all molecules of NA-LprG were loaded with glycolipid and likely is a significant underestimate of the increase in potency afforded by association of glycolipid with NA-LprG. Thus, the carrier function of NA-LprG dramatically increases recognition of glycolipids by TLR2.

Currently, it is not known how insoluble lipid agonists of TLR2 leave their phospholipid bilayers for transport through aqueous environments to reach and bind to TLR2. For LPS and TLR4, lipid transfer to the TLR is highly dependent on a host-derived carrier, LPS binding protein [211], and CD14 [212]. Our results suggest that the carrier function of LprG may solve a similar transport problem for mycobacterial glycolipids. To our knowledge this is the first report of a pathogen-derived protein serving as a carrier for delivery of hydrophobic agonists for innate immune recognition. We propose that other pathogen-derived carriers may similarly enhance the potency and efficacy of hydrophobic PAMPs for induction of host responses due to the common need of the pathogen and the host to chaperone
these molecules during transit through aqueous environments and deliver them at their site of function.

Our data indicate that the ability of LprG to facilitate TLR2 activation is influenced by its ligand specificity, which includes triacylated glycolipids. This property may allow LprG to deliver glycolipids that are enriched in TLR2 agonists, essentially editing the types of lipids delivered for potential TLR2 recognition. Among the entire spectrum of phospholipids, sphingolipids or other lipids produced by the host or pathogen, the large majority are diacylated. We hypothesize that LprG has an editing function that preferentially selects triacylated lipids from among the much larger pool of naturally occurring diacylated lipids, leading to selective activation of the TLR2-TLR1 heterodimer.

These findings indicate a novel strategy for the host to co-opt function of microbial carriers to facilitate delivery of hydrophobic agonists for recognition by host receptors. Contributions of this mechanism to recognition of Mtb glycolipids may enhance host immunity to Mtb, but we cannot exclude the possibility that the pathogen benefits from this mechanism, perhaps via TLR2-dependent induction of down-regulatory mechanisms that may contribute to immune evasion by Mtb [40, 41, 49, 193]. Similar mechanisms may contribute to innate immune recognition of other bacterial pathogens.
Chapter 4: Binding of mycobacterial lipoproteins by TLR2 and TLR1 extracellular domains.
Introduction

Toll-like receptors are single-pass type I transmembrane receptors with an extracellular domain composed almost entirely of leucine rich repeats, and an intracellular TIR domain which mediates signal transduction [29]. A major goal of this thesis is to describe the physical interactions of mycobacterial lipoproteins and glycolipids with the extracellular LRRs of TLR2.

Recent work has yielded a bounty of information regarding how TLR1 [136], TLR2 [136], TLR3 [166, 167], TLR4 [168], TLR5 [169], and TLR9 [170, 171] bind their agonists. While in general terms, TLRs bind their agonists via their LRRs, there is remarkable complexity to the mechanisms of detection utilized by TLRs. This complexity is likely due to specific challenges imposed by the ligands. For the purposes of this discussion, it is useful to contrast the detection of a hydrophilic ligand with that of hydrophobic ligands.

The hydrophilic TLR ligands are sensed by binding to the surface of the LRR of the TLR, and do not require accessory receptors or co-receptors. For instance, TLR3 has two positively charged sites responsible for interaction with the phosphate backbone of dsRNA: one composed of the N-terminal LRR (LRRNT) and LRRs 1-3, and the other composed of LRRs 19-21 [167]. While TLR3 has a small region supportive of homotypic interaction near the LRRCT, binding of ligand drives dimerization of soluble extracellular domain [166, 172].

While TLR3 binds its ligands via ionic interactions mediated by patches of positively charged residues on the surface of its LRRs, TLR4 and TLR2 bind
hydrophobic agonists via hydrophobic pockets, and share a requirement for co-
receptors and accessory receptors to sense their hydrophobic ligands. The crystal
structure of the TLR4/MD-2/LPS complex has been solved [168]. Hexaacylated LPS
binds both MD-2 and TLR4, by inserting five of its acyl chains into MD-2 (inducing a
conformation change), and the remaining acyl chain forms a hydrophobic interface
between MD-2 and TLR4. While hexa-acylated LPS contacts the surface of TLR4
with only one acyl chain, TLR2 appears to interact heavily with its acylated agonists.
The TLR2-PAMP interaction that is best defined is the interaction between TLR2
and the synthetic triacylated lipopeptide Pam3CSK4. Monomeric TLR2 ECD can bind
the diacyl glycerol motif of the synthetic lipopeptide Pam3CSK4 [142] via a
hydrophobic pocket located at the junction between the central and C-terminal
portions of the ECD of TLR2 (Fig. 1.8) [136]. Residues of LRRs 9-12 contribute to
the surface of the hydrophobic pocket. The amine-linked acyl chain of Pam3CSK4
inserts into a narrow groove located between the central and C-terminal domains of
TLR1, effectively linking the two TLR ECDs [136].

The above studies demonstrate that TLR2 binds diacylglycerol motifs and
TLR1 binds amine-linked fatty acids. However, the contribution of the peptide [158]
or glycan [89] component of full-length, physiologically important lipoproteins
remains unknown, and the reports of non-acylated TLR2 agonists suggest that other
mechanisms of TLR2 binding exist. Studies of full-length lipoprotein agonists of
TLR2 may shed light on other mechanisms of interaction with TLR2. We therefore
set out to design experiments to define binding characteristics of TLR2 with the
mycobacterial lipoproteins LpqH, LprA, LprG, and the non-acylated TLR2 agonist,
NA-LprG. This chapter of this thesis describes studies designed to answer the following questions: 1. Do mycobacterial lipoproteins interact directly with TLR2 ECD? 2. Does TLR1 also bind mycobacterial lipoproteins? 3. What are the affinities of interaction between TLR2 and/or TLR1 and each mycobacterial lipoproteins? 4. Does NA-LprG and its carried glycolipid form a ternary complex with TLR2, or does NA-LprG dissociate after delivery glycolipid to TLR2? 5. Does soluble CD14 contribute to mycobacterial lipoprotein or glycolipid binding to TLR2? 6. Will full-length lipoproteins and glycolipids induce dimerization of soluble extracellular domains of TLR2 and TLR1? 6. Does NA-LprA, inactive as a TLR2 PAMP, behave similarly in these experimental systems to TLR2 agonists?

To address these questions, purified lipoproteins were used in binding assays with purified TLRs. While preliminary data of direct interaction between TLR2 and mycobacterial lipoproteins have been obtained, a detailed pharmacological description of the interaction between TLR2 and mycobacterial lipoproteins remains to be accomplished. This chapter details the approaches that have been taken to produce these molecules and the experimental designs used to describe binding between TLR2 and mycobacterial lipoproteins.
Methods

Mammalian expression systems

Creation and expression of TLR2-Fc constructs

Since Toll-like receptor 2 is known to be N-glycosylated at four residues in the extracellular domain [213], the first expression and purification approach utilized a human cell line known to be capable of mammalian N-linked glycosylation, the HEK293 cells (ATCC, CRL-1573). These were either transduced with retrovirus to produce human TLR2 ECD fused with mouse IgG2a Fc [152], (kindly provided by Alberto Visintin while in the laboratory of Douglas Golenbock), or transfected with the pcDNA3-based construct “pILN-FLAG-TLR2XaFc”, encoding a pre-pro-trypsin leader sequence, a FLAG epitope tag, human TLR2 ECD, a Factor Xa cleavage recognition sequence, and mouse IgG2a Fc, a protein referred to as “FhT2XaFc”. This construct was received from Brian Monks, also working in Douglas Golenbock’s laboratory, and has not yet been used in any publications. A similar construct, pMGD01, was created with encoding pre-pro-trypsin leader sequence, a FLAG epitope tag, human TLR9 ECD, a Factor Xa cleavage recognition sequence, and mouse IgG2a Fc, encoding a protein referred to as “FhT9XaFc”.

113
Figure 4.1. Enrichment for highly expressing cells by flow cytometry.
HEK.293 cells transfected with pILN-FhT2XaFc (A) or pILN-FhT9XaFc (B) were selected for surface staining of an IRES-linked protein (LNGFR). Consecutive sorting shows enrichment for cells with greater levels of surface marker.
Expression of TLR2-Fc constructs in HEK293 cells

For growth in adherent conditions, all transfected HEK lines were maintained in HEK medium (see chapter 2). For suspension culture, several different serum-free mediums were tested (CD293, 293SFM, Freestyle SFM). Freestyle SFM was chosen as it permitted the greatest cell density.

HEK293 cells were stably transfected with plasmids pILN-FLAG-TLR2X-Xa-Fc and pMGD01 using standard lipofection. After 3 weeks of culture under G418 selection (50 ug/ml), cells were gradually adapted to growth in serum-free medium (Invitrogen, 10486) and converted to growth in suspension. Cells were considered fully adapted when the density of viable cells approximately doubled every 24 h and consistently maintained viability above 85% (assessed by trypan blue exclusion). Cells were then sorted for surface expression of LNGFR by staining with FITC-anti-LNGFR antibody (1:11; Miltenyi, 130-097-917) and sorted on an ELITE II flow cytometer such that cells with the top 5% of LNGFR-FITC MFI were selected for passage (MGD117, MGD264), (Fig. 4.1). These cells were frozen soon after the second sort (further sorts did not increase purity of population) and are the laboratory stock of HEK293-FhT2XaFc and HEK293-FhT9XaFc cells. While the FhT2XaFc construct was found in the culture filtrate of transfected cells with yields of approximately 100-200 ug/L, FhT9XaFc was barely detectable in culture filtrate even after protein A affinity enrichment (Fig. 4.2), and purification from cell lysate was not pursued.
Figure 4.2. FhT9XaFc is expressed but not secreted efficiently from HEK.293 cells. SDS-PAGE analysis with visualization by silver stain (A) or anti-FLAG (B) and anti-mouse IgG (C) western blot. lane 2 is cell lysate (non-denaturing lysis). Lanes 3-7 are flow through and washes, lanes 8-9 are elution by pH 3.5 and 3.33, respectively.

Leucine Zipper constructs

While the previously discussed constructs allowed the production of soluble TLR2 ECD, we wished to study mycobacterial lipoproteins known to signal via both TLR2 and TLR1 (see chapter 2). Since TLR2 and TLR1 are thought to cooperate in
direct binding of their ligands, we designed soluble TLR2/TLR1 and TLR2/TLR6 and TLR1/TLR6 complexes that would make ligand-independent heterodimers using leucine zippers cloned from c-Fos and c-Jun. Constructs containing Fos contained FLAG epitope, and constructs containing Jun contained streptavidin binding peptide (SBP) motif. This design was intended to allow purification of heterodimers by tandem anti-FLAG and streptavidin affinity purifications. These constructs are based on pcDNA3, and use the endogenous TLR leader sequence as the only secretion signal. Constructs are named pMGD02 (human TLR2-Fos-FLAG), pMGD03 (human TLR6-Fos-FLAG) pMGD04 (human TLR1-Jun-SBP) and pMGD05 (human TLR6-Jun-SBP).

Expression of leucine zipper constructs

HEK293 and E310-CHO cell lines were stably transfected with either one or two plasmids in the following combinations (number=TLR, FF= Fos-FLAG, JS= Jun-SBP): 2FF, 1JS, 6JS, 6FF, 2FF/1JS, 2FF/6JS, 2FF/2JS, 6FF/1JS. After maintenance under G418 selection for 3 weeks, cells were stained for their IRES-linked surface markers and sorted via flow cytometry (Fig. 4.3). Unfortunately, TLR2-fos-FLAG and TLR1-jun-SBP were produced and secreted at barely detectable levels in both cell systems; and TLR6 was produced weakly and not secreted (data not shown, MGD247). Because their expression was so poor, purification of these constructs was not pursued.
Figure 4.3. Comparison of expression of FhT2XaFc and the leucine zipper construct TLR2-Fos-FLAG (2FF). LNGFR surface staining of HEK.293 cells transfected with FhT2XaFc (A) or TLR2-Fos-FLAG (B). Anti-FLAG western analysis of cell lysates (C) or culture filtrate (D) of HEK.293 cells transiently transfected with FhT2XaFc or TLR2-Fos-FLAG (2FF). Numbers above gels reflect time (days) post-transfection.
TLR-VLR fusion proteins

Because many groups have reported difficulty in expression of soluble TLR ECDs, Jie-Oh Lee took the ingenious strategy of expressing parts of TLRs fused with another leucine rich repeat protein, the variable lymphocyte receptor (VLR), which he had previously crystallized [214]. Constructs encoding TLR-VLR fusion proteins were received from Jie-Oh Lee. These are named as follows: 2932 (human TLR1), 2394 (human TLR2), 2812 (mouse TLR2), TV8-Fc (human TLR4), mTLR4 (full-length mouse TLR4 ECD, no VLR fusion). FLAG-tagged 2394 and Myc-tagged 2932 constructs were produced in the pAcGP67 plasmid.

Purification of FhT2XaFc or FhT2x from HEK293-FhT2XaFc culture filtrate

After culture filtrate was collected, sodium azide (0.5%) was added, and the filtrate was subjected to protein A affinity chromatography by gravity flow through a manually-poured column of protein A-agarose beads (250 µl of protein A beads per L filtrate). As an alternate approach, anti-FLAG affinity chromatography was also utilized, but was found to be less efficient than protein A affinity at capture of the construct (data not shown). For this reason, protein A affinity was chosen as the standard approach. The protein A beads were washed with 500 ml (per ml of beads) PBS pH 7.4, and eluted in one of two methods: 5x1 ml acidic buffer (100 mM Glycine pH 3.5 or 100 mM sodium citrate, pH 3.5) or by cleavage by Factor Xa (EMD, 69036) under the following reaction conditions: 10 U/ml in 50 mM Tris-HCl, 100 mM NaCl, 5 mM CaCl₂, pH 8.0, four h incubation at room temperature with gentle
agitation. The acid and enzymatic elution methods allowed elution of full-length construct (FhT2XaFc) or the FLAG-tagged TLR2 ECD (FhT2x), respectively. Cleavage specificity was demonstrated by SDS-PAGE analysis of products (Fig. 4.4).
Figure 4.4. Specificity of Factor Xa-mediated cleavage of FhT2XaFc. Purified FhT2XaFc was incubated with a range of concentrations of Factor Xa at room temp for 2, 4, 8, or 17 h. Products were analyzed by SDS-PAGE and visualized with silver stain (A), or western blots specific for FLAG epitope (B), or mouse IgG Fc (C). Predicted sizes of products are indicated in panel D.
A recent modification to the protocol for purifying TLR2 ECD from HEK293-FhT2XaFc culture filtrate was to concentrate the culture filtrate using an Amicon stirred cell (model 8400) with a 50,000 MWCO membrane and affinity-purify the protein using a pre-packed protein A-sepharose column (HiTrap rProtA, GE Healthcare, 17-5079-01) using an AKTA Purifier 10 FPLC system. The column was washed with 5-10 ml of 20 mM sodium phosphate, pH 7.0 and eluted with 100 mM sodium citrate, pH 3.5 into neutralization buffer (1M Tris-HCl, pH 9.0; 200 ul per ml of eluate). While no direct comparison was made, this approach seemed to garner greater yields, most likely because of reduced volume of wash buffer.

For details of expression and purification of mycobacterial lipoproteins, please refer to chapter 2.

**Results**

**Pull-down assay**

To assess whether lipoproteins interacted directly with TLR2 ECD, protein A beads were coated with FhT2XaFc or mouse IgG specificity control or PBS. The coated beads were then incubated in the presence or absence of LpqH, washed extensively, and material complexed to the beads was eluted with SDS-PAGE buffer, resolved by SDS-PAGE, and visualized by silver stain and anti-LpqH (IT-19, 1:3000) western blot (Fig. 4.5). LpqH was pulled down selectively by FhT2XaFc coated to
the bead, consistent with the hypothesis that LpqH binds directly to TLR2 ECD.

Interestingly, the presence of FBS inhibited binding to FhT2XaFc, possibly due to competition for LpqH binding by PRRs present in serum.

Figure 4.5. Pull-down of LpqH by FhT2XaFc captured on a protein A bead. Lane 2 is LpqH, lane 5-7 protein A beads coated with isotype control antibody and incubated with PBS (5), PBS + LpqH (6), or 10%FBS+LpqH (7). Protein A beads coated with FhT2XaFc and incubated with PBS (8), PBS + LpqH (9), or 10% FBS + LpqH (10). Proteins visualized with silver gel (upper panel) or anti-LpqH (IT-19) (lower panel).
Size exclusion binding assay

To assess whether lipoproteins interact with TLR2 ECD in solution, purified LpqH was incubated with purified FhT2XaFc in PBS for 3h at 37°C. This mixture was then subjected to gel filtration chromatography using a Superose 12 10/300 HR column (GE Healthcare, c/o Brian Cobb lab) and elution profiles were compared to profiles of unmixed proteins (**Fig. 4.6A**). This experiment yielded no evidence of Ve shift of the FhT2XaFc peak, implying that no change of stokes readius (due to binding) of the protein is detected. A second experiment was performed with two critical differences in design: 1. Since interactions between biological macromolecules often occur in a calcium-dependent manner, this experiment was performed in PBS supplemented with 1 mM CaCl\(_2\) (DPBS). 2. Since elution with acid may have denatured TLR2 ECD, In this experiment, FhT2x was used instead of full-length FhT2XaFc. The peak of FhT2x shifted approximately 600 ul, equivalent to an apparent gain in MW of 8 kD (**Fig. 4.6B**). While this experiment needs to be repeated, if there is real binding, it may either be attributable to a requirement for CaCl\(_2\) or the more conformationally active TLR2, due to enzymatic elution of FhT2x compared to acid-elution of FhT2XaFc.
**Figure 4.6. Size exclusion binding assay of FhT2XaFc with LpqH.** Experiment was performed in the absence (A) or presence (B) of 0.9 mM calcium chloride.

**Plate-based binding assays**

As described above, a plate-based binding assay can be used to demonstrate specific and saturable binding and as a means of generating an estimate of binding affinity. Four plate-based binding assays were utilized (Fig. 4.7). While all setups yielded uninterpretable results, there may be some remaining options worth exploring, explained below.
Establishment of coating conditions and detection reagents

To establish that each protein can effectively adhere to the ELISA plate, plates were coated with a range of concentrations of protein, washed, and probed directly with various detection reagents. These experiments were intended to determine appropriate amounts of material for coating, as well as demonstrate effective detection reagents appropriate for this system (Fig. 4.8).
Figure 4.8. Validation of plate coating and detection reagents. (A) Detection of His$_6$-tagged Lprs coated to plate by Abcam anti-His antibody. (B) Detection of His$_6$-tagged LprG by Ni-HRP. (C) Detection of LpqH-biotin coated plate by streptavidin-HRP. (D) Detection of FhT2XaFc coated plate by HRP-conjugated anti-FLAG and anti-mouse IgG-Fc antibodies.

Plate-based assay #1: Plate coated with lipoprotein, incubated with range of concentrations of full-length FhT2XaFc.

Plates can be coated successfully with lipoprotein (Fig. 4.8A). However, full-length FhT2XaFc exhibited high non-specific binding to the plate with an apparent affinity of approximately 1.5 nM (Fig. 4.9). While BSA reduced binding of FhT2XaFc
to the plate to some degree, it still could not prevent non-specific binding above 1 nM.

Figure 4.9. FhT2XaFc binds to plate non-specifically.

Plate-based assay #2: Plate coated with TLR, incubated with His$_6$-tagged Lprs.

Plates can be coated successfully with FhT2XaFc (Fig. 4.8D), and His$_6$-tagged Lprs can be detected by anti-His antibody (Fig. 4.8A), but no interaction of His$_6$-LprG with immobilized FhT2XaFc could be detected. The presence of CaCl$_2$ and/or CD14 did not increase detection (Fig. 4.10). This experiment was blocked with superblock for 1h at room temperature, and did not suffer from background signal. However, unlike previous experiments, the binding step and washes were performed with PBS supplemented with 0.05% Tween-20. The presence of detergent may have been responsible for the lack of background, but also may abrogate acyl chain binding into the hydrophobic pocket of TLR2.
Figure 4.10. No detection of LprG binding to FhT2XaFc coated to the plate.

Plate-based assay #3: Plate coated with protein A/G, TLR captured on Protein A/G, and incubated with His$_6$-tagged Lprs or LpqH-biotin.

Coating of the ELISA plate with Protein A/G was predicted to offer two advantages: 1. Allow for dense coating of plate to reduce potential for background binding to plate, and 2. Bind FhT2XaFc in a more uniform conformation predicted to increase availability of LRR for binding Lprs. The Protein A/G coat precluded antibody-based detection. Thus, Ni-HRP was used to detect the His$_6$-tagged Lprs (validated in Fig. 4.8B), and Sav-HRP was used to detect LpqH-biotin (validated in Fig. 4.8C). LpqH-biotin bound to the plate non-specifically above 1 nM, and His$_6$-tagged Lprs bound non-specifically to the plate at concentrations above 100 nM (Fig. 4.11).
Figure 4.11. Non-specific binding of lipoproteins to the plate. LpqH (A) and the His$_6$-tagged lipoproteins in the absence (B) or presence (C) of calcium. (D) Positive control for detection.

Surface plasmon resonance (SPR)

SPR was measured by a BIACore 3000 system (c/o Satya Yadav, PhD, CCF) using a CM5 sensor chip with a carboxymethyl-dextran matrix. This is a versatile chip; capable of covalent attachment of a protein “ligand” via free amine (-NH$_2$), or –SH, -CHO, -OH, or –COOH chemistry. Binding of another protein “analyte” is assessed by flowing a solution of the analyte over the ligand-coated surface and
measuring in real time the change in refractive index of the solution within 300 nm of the surface of the chip. The first approach involved directly conjugating FhT2XaFc as the ligand. To accomplish this, solutions (85 µl) of FhT2XaFc were made in 10 mM Acetate pH 5.0, 4.5, and 4.0. Each solution was pumped over the surface at a rate of 20 µl/min, and the change in refractive index (relative units) was measured over time. The predicted isoelectric point of FhT2XaFc is 6.0; we found that while the protein did not interact with the carboxy-dextran matrix at pH 5.0, it did interact when in pH 4.5 and 4.0 (Fig. 4.12). Therefore pH 4.5 was used for immobilization of FhT2XaFc. One drawback to this approach is the relatively harsh conditions required for immobilization of the TLR2 ECD to the chip. Binding analysis was first attempted using Pam3CSK4 and FSL-1 as analytes. These analytes bound with very high affinity to both the experimental (with ligand) and reference (without ligand) surfaces. Harsh conditions (0.1% SDS and 5 mM NaOH) were required to strip the lipopeptides from the surfaces. Subsequent to this harsh treatment, no binding was detected when LpqH was used as an analyte with these surfaces (see MGD267). Given the preliminary evidence for direct interaction between TLR2 ECD and LpqH (Fig. 4.5), the lack of binding may be due to inappropriate buffer conditions for binding or denaturing of the FhT2XaFc during purification and/or immobilization.

To avoid exposure of FhT2XaFc to the harsh conditions used for the immobilization of the ligand, Protein A/G was immobilized to the surface using free amine chemistry in 10mM acetate pH 4.5 (Fig. 4.12). This allowed for capture of the FhT2XaFc under neutral conditions (10 mM HEPES, 150 mM NaCl, MGD271)
(Fig. 4.13). To analyze whether LpqH could bind FhT2XaFc captured by immobilized Protein A/G, we injected LpqH in the presence and absence of Calcium. We found that signal intensity increased only in the presence of Calcium (Fig. 4.13). This experiment offers preliminary hint of an interaction, but must be interpreted cautiously as it lacks two important controls: 1. Injection of PBS/CaCl$_2$ solution lacking LpqH (to control for possible charge-charge interactions of Ca$^{2+}$ with the carboxymethylidextran matrix) and 2. Injection of LpqH in PBS/CaCl$_2$ solution over a chip coated with isotype control antibody or a TLR LRR not predicted to bind LpqH (i.e. TLR4). These controls are discussed in more detail in chapter 5.

![Graph showing association of protein A/G to CM5 chip at various pH values.](image)

**Figure 4.12.** Association of protein A/G to CM5 chip at various pH values.
Figure 4.13. Realtime analysis of LpqH interacting with FhT2XaFc captured on protein A/G immobilized to the chip. Capture of FhT2XaFc onto protein A/G (injections 1 and 2). Gradual loss of FhT2XaFc from dissociation seen in slope before injection 3. Injection of LpqH in PBS (3) or LpqH in PBS + 1 mM CaCl$_2$ (4).
Discussion

The studies described in this chapter were designed to demonstrate direct physical interaction between human TLR2 ECD and mycobacterial lipoproteins (when the studies were designed and performed, it was thought that NA-LprG was signaling directly via its peptide component or a covalent post-translational modification). Current goals include analysis of “delivery” of glycolipid to TLR2, discussed below. Another goal of the experiments was to estimate affinity of the interaction between each of the lipoproteins and TLR2, to determine if the differences in TLR2 response of LprA and LprG (Fig. 3.1) and LpqH (Fig. 2.1) correlated with binding affinity. Since many of the experiments did not achieve these goals, it is worth discussing a few critical technical issues that may be at play.

Since TLR2 is expected to bind its hydrophobic agonists via hydrophobic interactions, detergents were avoided in most experiments. It is interesting that including 0.05% Tween-20 in a plate-based binding assay removed all background signal, as well as any potential binding signal. This suggests that if binding is occurring between TLR2 and lipoproteins, detergent is effective at disrupting it. It may be worth testing some very minute concentrations of detergent, with the goal of finding a dose of detergent that will remove background but leave specific binding intact.

Evidence from the literature suggests that the interaction between TLR2 and a diacylglycerol motif is very stable [136]. This study allowed Pam₃CSK₄ to bind to soluble TLR2 ECD, then added TLR1 to the mixture, and re-purified the ternary
complex from any free TLR ECD or Pam₃CSK₄ monomers by size exclusion chromatography. The complex was consistent enough in conformation to generate a crystal structure with resolution of all three molecules in their entirety. The ability of the complex to remain uniformly intact throughout a size exclusion run suggests that the complex is very stable. Presumably, mycobacterial lipoproteins would be predicted to bind in a similar manner, and with similar stability. For this reason, a size-exclusion readout may be a more optimal approach to discriminating whether the proteins interact. While the size exclusion experiments that have been performed thus far are flawed by the use of LpqH, use of the His6-tagged lipoproteins should be less ambiguous.

Another approach to improving the binding assays is to use delivery molecules. While binding occurs between purified TLR2 and purified Pam₃CSK₄ in the absence of any other co-receptors, cellular responses are dramatically enhanced if the Pam₃CSK₄ is first bound to soluble CD14 before use in bioassays [143]. Importantly, although soluble CD14 enhances cell sensitivity to Pam₃CSK₄, it does not form a ternary complex with TLR2, suggesting that its role is not to stabilize interactions between TLR2 and Pam₃CSK₄, but rather to deliver Pam₃CSK₄ to TLR2 [142]. This proposed delivery function [144] [201] is also proposed for CD14 in LPS signaling. Since the immunogenic activity of hydrophobic molecules such as LPS [215] [216] [217], Lipid A [218] and lipopeptides [219] is influenced by their aggregation state, it can be inferred that bioavailability may be a limiting factor in the binding and detection of hydrophobic agonists. Thus, including protein carrier molecules such as CD14 or NA-LprG in plate-based binding assays may
simultaneously increase the apparent affinity of binding and decrease the troublesome background noted in the absence of detergent. Although it may be less important for the size exclusion assay, CD14 may also be useful in this setting.

Future studies using constructs based on the TLR-VLR hybrid proteins designed by the laboratory group of Jie-Oh Lee will have several advantages. First, we anticipate greater yields of TLRs. Second, this system will allow purification of TLR1 and TLR2, TLR4, and possibly TLR6 (construct in progress). The constructs are similar in design to the FhT2XaFc construct that has already been used, and thus should transition into the assays already in use.
Chapter 5: Discussion and Future Directions
Recognition of microbes and their products by PRRs is one of the earliest events in host-pathogen interaction. Specifically, TLRs are critical for initiation of innate immune responses such as the induction of pro-inflammatory cytokines, activation of microbicidal functions, and the initiation of the adaptive immune response [29]. While there is a tremendous body of knowledge regarding the functions of TLRs and the sequelae of their activation, many questions remain regarding their mechanism of activation. This thesis focuses on the mechanisms of TLR2 activation by mycobacterial lipoproteins.

Among the TLRs expressed in humans, TLR2 is the most confusing. It is reported to sense the broadest array of molecules, including lipopeptides/lipoproteins [135, 148], glycolipids [134, 196], proteins [152], peptides [51], and capsular carbohydrates [157]. No common theme, structural motif, or single mechanism of binding can easily be imagined to accommodate the detection of such disparate agonists by TLR2.

How, then, does TLR2 recognize such a wide variety of natural compounds (glycans, lipids, amino acids)? Several facts may help provide an explanation. TLR2 is the only known TLR to form heterodimers with other TLRs; TLR1 and TLR6, which are closely related to TLR2. Since TLR1 and TLR6 are thought to participate in binding of agonists as well as signaling, we refer to them as co-receptors. While there is data to support ligand specificity for these heterodimers, there is no data to support differential transcriptional sequelae of signaling by them. This suggests that TLR1 and TLR6 are maintained in the genome to expand the ligand repertoire
of TLR2. However, there is a fourth TLR in this clade, the oft-forgotten TLR10, which is more closely related to TLR2 than TLR1 or TLR6 [29]. Thus far, neither agonist nor function has been assigned to TLR10; it remains an orphan receptor with very little known about it.

TLR2 also utilizes more accessory receptors than any other TLR, and how they contribute remains largely ambiguous. The best-characterized accessory receptor is CD14, which also plays a critical role in LPS detection by TLR4 [212]. Other co-receptors of TLR2 include CD36, [138] lipopolysaccharide binding protein [139], CD11b-CD18 integrin [220], ganglioside GD1a [141], and RP105 [221]. Many of these accessory receptors have been characterized only for a single ligand or a subset of ligands. While some of these accessory receptors are hypothesized to deliver ligands to TLR2, CD36 [146] and the CD11b-CD18 integrins [220] are reported to directly contribute to signaling other than that typically induced by TLR2.

It is not known if the other accessory receptors can contribute to TLR2 signaling in ways other than ligand delivery. Furthermore, these accessory receptors seem to have some ligand specificity of their own, and may contribute to the ligand specificity observed for TLR1 and TLR6. Many of these important findings, in particular regarding ligand specificity, were obtained by studies using a small set of synthetic TLR2 agonists. These studies have led to some “rules of thumb” of TLR2 activation: For instance, triacylated lipoproteins are recognized by TLR2/TLR1 heterodimers, and diacylated ligands are recognized by TLR2/TLR6
heterodimers. There are fewer proposed ligand structural determinants for the specificity of the accessory receptors, except perhaps that CD36 is important for recognition of diacylated structures such as R-MALP-2 and Lipoteichoic acid (but not all diacylated structures; i.e. S-MALP-2, Pam\(_2\)CSK\(_4\)) [138]. Since lipoproteins have a much larger protein component than these synthetic molecules, we hypothesized that the protein component may influence the co-receptor and accessory receptor requirements of TLR2 activation.

Chapter 2 of this thesis describes work that addresses the question of whether the protein component of full-length lipoproteins contributes to TLR2 signaling and/or co-receptor and accessory receptor requirements. Since mycobacterial lipoproteins are predicted to be triacylated, the prevailing hypothesis of the field would predict their recognition by TLR2/TLR1 heterodimers. We tested LpqH, PhoS1, LprA, and LprG for their TLR2 co-receptor requirements. While most of the proteins showed dependence on TLR1, the dependence varied, and LprA did not require TLR1. We propose that LprA may exhibit some cross-reactivity for TLR1, TLR6, TLR10, or another PRR. Since these proteins are all predicted to be similarly acylated, we hypothesized that the protein component influences the receptor requirements for TLR2 agonist activity. However, in addition to the possibility of direct recruitment of co-receptors by the protein component, the protein component may be playing an indirect role in recruitment of TLR1 or TLR6 co-receptors (i.e. by carrying another TLR2 agonist, as described for LprG in chapter 3).
The mechanisms by which CD14 contributes as an accessory receptor to promote TLR2 signaling are incompletely understood; prevalent hypotheses include a role in trafficking and/or increased bioavailability of TLR ligands. CD14 has LRRs similar to TLRs, but is GPI-linked to the cell surface and cannot directly signal. CD14 is also released from the cell surface; the presence of either membrane-associated or soluble CD14 enhances Pam3CSK4 binding to TLR2 [142] and resultant signaling [143]. Membrane CD14 may contribute to TLR2 signaling via additional mechanisms than strict delivery of agonists by soluble CD14. The role of CD14 in LPS signaling offers some clues as to other ways CD14 may contribute. CD14 is required for internalization of TLR4 after LPS stimulation [222] and induction of IFN-beta [212], consistent with the reported requirement of internalization for TRAM recruitment and IFN-beta induction [223]. TLR2 and TLR4 are the only TLRs that utilize CD14, and are also the only TLRs that require TIRAP for MyD88 recruitment [224] to membrane microdomains [225]. Thus, recruitment of TLR2 [226] and TLR4 [227] to membrane microdomains by agonist-loaded CD14 may contribute indirectly to intracellular signaling events.

Work described in this thesis demonstrates that not all lipoproteins require CD14. Among the Mtb lipoproteins, response to LprA showed the greatest dependence on CD14; responses to LprG and LpqH showed intermediate dependence, while responses to PhoS1 showed no detectable dependence on CD14 expression. Contribution by CD36 to detection of lipoproteins also varied. Only responses to LprA showed a requirement for CD36. These proteins, in particular
LprA and LprG, may be useful candidates to study the structural determinants important for accessory receptors of TLR2.

It is interesting that LprA and LprG differ in their receptor usage despite being closely related (~34% identical amino acid sequence). LprA was partially dependent on both CD14 and CD36, whereas LprG showed partial dependence on CD14 but not CD36. LprA, unlike the other three Mtb lipoproteins in this study showed no dependence on TLR1. The structural characteristics of LprA that determine its unusual receptor usage are as yet undefined.

A recent study has reported a novel contribution to TLR2-mediated detection of LpqH by the protein RP105 [221]. RP105 is a LRR protein related to TLR4. TLR4 interacts with MD-2 [228], and RP105 interacts with MD-1 [229]. RP105 has a transmembrane domain but only a small (11 amino acid) cytoplasmic domain that may associate with other signaling molecules to mediate signaling. The study by Blumenthal et al. is particularly interesting because RP105 is thought to decrease macrophage TNF response to LPS [230], while this study demonstrates that RP105 enhances macrophage TNF response to LpqH. Strangely, the cell type that expresses RP105 also determines RP105 function in sensing LPS. While macrophage cytokine response to LPS was decreased with expression of RP105, B cell proliferation in response to LPS was enhanced by expression of RP105. These ligand-specific and cell type-specific effects on TLR responsiveness reveal novel complexity to TLR signaling. Future studies of the involvement of RP105 in
macrophage responses to mycobacterial lipoproteins may shed light on agonist structural requirements of signaling via RP105.

The above studies were performed using bone marrow-derived macrophages. The use of bone marrow-derived macrophages of knockout mice addressed the importance of receptors for TLR2-mediated response to mycobacterial lipoproteins. However, receptor expression levels may vary within wild-type animals dependent on cell type. As mycobacterial infection occurs primarily in the lung, we wished to define receptor expression profiles of the antigen presenting cells present in the lung: the alveolar macrophages, lung macrophages, and lung dendritic cells [179]. These subsets have been previously defined by flow cytometry [178]. We found that lung macrophages had lower levels of TLR2 surface expression than either alveolar macrophages or dendritic cells. By sorting these populations and testing their responsiveness to LpqH, we found that lung macrophages were less responsive than alveolar macrophages. These data suggest that receptor expression level may be a limiting factor for host responses to pathogens.

While the finding that alveolar macrophages were the most responsive to Mtb lipoproteins was unexpected, recent advances in the understanding of lung mucosal immunity may offer an explanation. Alveolar macrophages have historically been described as a relatively quiescent, anti-inflammatory APC subtype important for maintaining immune homeostasis within the lung. For instance, in vivo elimination of alveolar macrophages using clodronate-filled liposomes leads to
overt inflammatory reactions to otherwise harmless particulate and soluble antigens [231]. It has recently become clear that the anti-inflammatory state of the AM is not an intrinsic trait, but rather dependent on the lung microenvironment. For instance, while the major lung collectin surfactant protein A (SP-A) increases phagocytic uptake of \textit{P. aeruginosa} [232], it decreases TNFα production in response to LPS [233] [234]. In addition, the resting AM maintains close association to type I alveolar epithelial cells (AEC) [235], which express, bind and activate TGFβ [236]. This is thought to maintain the AM in a quiescent state in the face of the many non-self (but non-injurious) stimuli that might otherwise excite them [237]. However, TLR stimulation of alveolar macrophages induces a rapid loss of contact with type I AECs, allowing production of the proinflammatory cytokines TNFα and IL-6 [235]. The relative responsiveness of alveolar macrophages in this work may in part be explained by their removal from the lung microenvironment before stimulation.

In addition to characterizing receptor requirements for TLR2 mediated responses to mycobacterial lipoproteins, work in this thesis has also been directed towards characterizing the agonist structural determinants of TLR2 signaling, in particular for LprG. Previous studies have demonstrated that TLR2 agonist activity of bacterial lipoproteins is dependent upon acylation of the N-terminal cysteine, as demonstrated for OspA [148] [149], LpqH [148] and LprA [45]. In contrast, work by Nicole Pecora demonstrated that LprG retains TLR2 agonist activity in the absence of acylation. Several lines of evidence led to the conclusion that the TLR2 activity of NA-LprG is due to a carried molecule. First, while NA-LprG expressed in \textit{E. coli} had
minimal TLR2 activity, it gained TLR2 activity after exposure to mycobacterial lysate or mycobacterial glycolipids. Second, N- and C-terminal fusion proteins of NA-LprA and NA-LprG did not localize the activity to a single portion of NA-LprG. Third, the crystal structure revealed a hydrophobic pocket in LprG defined by residues in both the N- and C-terminus of the protein. The pocket was found to be responsible for TLR2 activity, as a point mutation predicted to provide steric hindrance abrogated the intrinsic activity of the protein purified from *M. smegmatis*, and the increase in activity due to charging of *E. coli* protein with mycobacterial lysate. Further experiments demonstrated that NA-LprG carried glycolipids in the hydrophobic pocket. First, western blot analysis revealed association of mycobacterial glycolipids with LprG. Mass spectrometry revealed that diacylated glycolipids were associated with both LprA and LprG, whereas triacylated glycolipids were associated preferentially with LprG. These studies reveal that LprG binds mycobacterial glycolipid TLR2 agonists and delivers them for recognition by TLR2.

The crystal structure of LprG revealed a hydrophobic pocket with a volume of 1535 cubic angstroms. Mutation of a single residue 91V to 91W abrogated intrinsic TLR2 activity of the protein purified from *M. smegmatis*, and the extrinsic activity gained after exposure to mycobacterial lysate. The crystal structure of the point mutant demonstrates that in addition to the steric hindrance directly caused by the presence of the aromatic side chain (decreasing the diameter of the pocket from 9 angstroms to 4.5 angstroms), the point mutation also caused shifts in several other side chains, including an approximately 1 angstrom shift in the loop between
β3 and β4, which opened the pore of the pocket (Fig. 5.1). These changes led to a narrower pocket with a smaller total volume (1195 cubic angstroms). Furthermore, the crystal structure of PIM bound to wild-type NA-LprG was solved, and demonstrates that an acyl chain of PIM$_2$ would directly conflict with the bulky side chain of the tryptophan mutant.

![Figure 5.1](image_url)

**Figure 5.1. Structural rearrangements induced by V91W point mutation.** Ribbon structures of LprG (tan) and LprG-V91W (brown) overlaid show near-perfect overlap except for a 1.3 angstrom shift of the loop between beta3 and beta 4 (red dashed circle) resulting in a contribution to pocket opening (pink). LprG surface is shown in blue. 91W further induced a shift of residues L73-L76 into the pocket (surfaces of 91W and L73-L76 shown in red). Figure produced with permission using data published in Nature Structural and Molecular Biology [189].

One important question that remains unanswered is why NA-LprA does not bind triacylated glycolipids. The 34% sequence conservation and the relatively uniform three-dimentional distribution of the conservation suggest that LprA has a
pocket similar to that of LprG. When the structure of LprA is modeled using LprG as a template, LprA is predicted to have a larger (1700 angstroms$^3$) volume pocket. Why then, does NA-LprA only associate with diacylated glycolipids, and why does it lack TLR2 activity? There are two subtle structural differences that may explain in part why NA-LprA lacks TLR2 activity. The first difference involves the unstructured loop between beta 1 and beta 2 in LprG and LprA (LprG residues 66-72), which contributes to the pocket opening. While the backbone of the loop in LprG and LprA model overlap quite well, they differ in amino acid sequence. In LprG, the residues G70 and S72 have small side chains and thus define a large opening. In LprA, the corresponding residues are asparagine and arginine, which narrow the potential pocket opening considerably. When the LprG-PIM structure is overlayed with that of the LprA model, the side chains of the LprA sequence conflict with the acyl chains of Ac$_1$PIM$_2$ (Fig. 5.2). This is a tentative hypothesis and will be verified by X-ray crystallography of NA-LprA.
Figure 5.2. Residues in LprA define a narrower pocket opening and conflict with the structure of Ac₁PIM₂. (A) Ribbon structure of LprA model (green) and LprG (blue) overlapped with the structure of Ac₁PIM₂ in the pocket. Red circle designates residues examined in panel B. (B) Surface of residues 70-72 of LprG shown in gray, surface of corresponding residues of LprA model shown in red. Surface of Ac₁PIM₂ in pocket of LprG shown in blue. Figure produced with permission using data published in Nature Structural and Molecular Biology [189].

The second reason why LprA may not bind LM and LAM is due to a surface feature that appears to differ between LprA and LprG. The glycans of the glycolipids may contribute to binding of glycolipids to their protein carriers. In the crystal structure of PIM bound to LprG, the glycans are positioned closest to the few polar residues near the pocket opening. Specifically, the mannose added by pimB is over a groove that wraps around the front surface of LprG, ending near the bottom of the protein (Fig. 5.3). Since that mannose is the site upon which further mannosylation
occurs (Fig. 1.4), it may be that the hydrophilic groove serves as a binding site for the mannan chain of LM and LAM. This groove may be important for the ability of LprG to transport LM and LAM for its physiologic function. The model of LprA lacks such a groove (Fig 3.2), suggesting that the groove may be an important determinant of the ability to bind and carry LM and LAM. Point mutations in the groove of NA-LprG may test this hypothesis. If binding of LM and LAM can be prevented by point mutations in the groove, then such a mutant could be used to distinguish the importance of LM and LAM transport vs that of PIMs in experimental systems discussed below for the LprG-V91W point mutant.
Figure 5.3. Proposed binding site of mannann chain of LM and LAM to LprG surface. (A) While acyl chains are embedded in hydrophobic channel of LprG, the PI moiety and the mannose residues associate with hydrophilic residues at the pore pening. Specifically, pimB-added mannose (dashed circle) is positioned over a groove which exits the pore and wraps around LprG, as shown in panel B (red dashed line). Groove continues around LprG to reach the bottom of the protein (data not shown). Figure produced with permission using data published in Nature Structural and Molecular Biology [189].
While this thesis focused primarily on the role of LprG as a carrier of agonists for TLR2 activation with implications for host defense, the findings also have implications for the function of LprG in bacterial physiology. LprG is widely present in mycobacteria with 100% sequence identity among Mtb complex species and significant sequence conservation in other mycobacterial species. LprG (Rv1411c) forms an operon with the transmembrane protein p55 (Rv1410c) [98], a putative drug antiporter [99]. The operon is important for virulence [100], and the proteins may function together in construction of the mycobacterial envelope [103]. The data presented here is consistent with the hypothesis that LprG serves as a carrier to facilitate synthesis and/or trafficking of glycolipids from their site of synthesis at the plasma membrane to their location in the outer membrane of the bacterium [205] [206]. We investigated this hypothesis with a preliminary experiment, utilizing the sliding motility assay [238]. Sliding motility is a passive spreading process driven by the growth of the bacterial colony. In this assay, bacteria are spotted onto a charged agarose surface and observed for the ability to spread outward as a single unit. If the surface of the bacterium is hydrophobic, it will not interact with the charged surface of the agarose, and will therefore be able to slide. If the surface is more hydrophilic, the bacterium will be unable to spread outward. Previous studies have demonstrated that *M. smegmatis* mc²155 undergoes sliding motility in this assay [238] [103]. In this strain, mutants deficient in glycopeptidolipids are unable to undergo sliding motility [239]. Similar findings were found in *M. smegmatis* mc²155 deficient in LprG [103]. In this thesis, we use a different strain of *M. smegmatis*, mc²6/1-2c. A preliminary experiment suggests that
our strain lacks the ability to undergo sliding motility (Fig. 5.4). However, when LprG-V91W is overexpressed, the strain acquires the ability to undergo sliding motility, suggesting that the presence of LprG-V91W is associated with an increase in surface hydrophobicity (Fig. 5.5). As this strain also expresses wild-type LprG, expression of LprG-V91W may have interfered with the ability of wild-type LprG to participate in the synthesis and/or transport of glycolipids to the cell surface, possibly by competing with wild-type LprG for interaction with other proteins associated with the membrane (i.e. p55). The expression of NA-LprG-V91 (which is retained in the cytoplasm) had no effect on colony morphology, suggesting that the mutant has to localize to the cytoplasmic membrane to exert its effect. We conclude that expression of the tryptophan mutant has a dominant-negative effect on LprG function dependent on targeting of the mutant to the cell membrane.

The question as to why *M. smegmatis* mc²/1-2c does not undergo sliding motility has not been investigated. Since LprG is expressed in slow-growing mycobacteria, where glycopeptidolipids are absent, we do not think that GPLs are an important substrate for LprG. Furthermore, Recht et al. proposed that glycopeptidolipids are present on the outer surface of the bacterium with their acyl chains sticking out towards the medium, making the surface of the bacterium hydrophobic. While the structure of the surface is not at all certain, this model seems to present an energetically unfavorable situation. It is commonly proposed that Mycobacteria typically maintain a hydrophilic surface with the acyl chains of extractable lipids forming the opposing leaflet of an asymmetric bilayer with the mycolic acids (Fig. 1.3). Since there are other freely extractable lipids in the cell
surface, we propose that there may be a balance between hydrophobicity and hydrophilicity determined by the relative prevalence of each class of molecule. Our preliminary data suggest that the PI-based glycolipids carried by LprG may be an important constituent of the hydrophilic surface (Fig. 5.5)

The function we propose for LprG is similar to that proposed for LppX in the transport of pDIMs to the outer membrane of mycobacteria [96] and the role of lipoproteins in *E. coli* in transport of membrane components, including other lipoproteins [207] and LPS [208] [209]. Since glycolipids have important roles in bacterial physiology and host-pathogen interactions that benefit the pathogen, a potential future goal is inhibition of these processes by targeting the pocket of LprG with chemotherapeutics.
Figure 5.4. *M. smegmatis* expressing LprG-V91W has an altered cell envelope. Colony morphology under carbon replete (high glucose (HG), 2% glucose) or carbon-limiting (low glucose (LG), 0.02% glucose) conditions suggests a change in surface properties of the *M. smegmatis* strain over-expressing the LprG-V91W protein. *M. smegmatis* strains overexpressing LprA (C) and LprG (B) behave the same as vector-transfected strain (A). LprG-V91W acquires the ability to undergo sliding motility in glucose-limiting conditions, suggesting increased hydrophobic character of the bacterium. Interestingly, NA-LprG-V91W did not change colony morphology, suggesting that the point mutant has to traffic to the cytoplasmic membrane to exert its dominant negative phenotype. This data is preliminary and needs to be repeated.
Figure 5.5. Model of sliding motility. (A) In vector control strain of *M. smegmatis*, the presence of glycolipids at the cell surface may contribute to charge/charge interactions between the surface of the bacterium and the charged surface of the agarose polymer, preventing sliding motility. (B) In *M. smegmatis* transformed with LprG-V91W, the mutant LprG may compete with the wild-type LprG for interaction with proteins that contribute to the function of LprG (i.e. p55, loading proteins or transport proteins), thus preventing synthesis and/or transport of glycolipids to the surface of the bacterium. The surface in the absence of the glycolipids is hydrophobic, a trait permissive to sliding motility.
In addition to its function in bacterial physiology, our results suggest a role for LprG in facilitating TLR2 recognition of triacylated Mtb glycolipids. LprG was unique among the studied proteins in ability to bind triacylated PIM$_2$ (Ac$_1$PIM$_2$) as seen in LC-MS studies, and this may correspond to the specificity of LprG for the larger and more heterogeneous LM and LAM molecules that may contain a triacylated PIM core structure. LAM is a complex polymer with a high molecular weight and cannot be analyzed by the MS systems reported here, but other studies show that a large proportion of naturally occurring LAM is triacylated [204]. Therefore, the preferential association of LAM with LprG as compared with LprA (Fig. 5) may be explained by acyl chain interactions. These findings are consistent with the observed receptor dependence of Mtb glycolipid signaling [193] [194] [159] [134] and the concept that TLR2 activity of mycobacterial glycolipids requires their triacylation or tetraacylation [210] [160]. The capacity of the LprG binding pocket to accommodate lipids with up to three alkyl chains may confer its ability to chaperone TLR2/TLR1 agonist activity of triacylated mycobacterial glycolipids. Although LprG was distinct from LprA in the ability to chaperone TLR2 agonists, other mycobacterial lipoproteins may be carriers of hydrophobic ligands, some of which may be TLR2 agonists. Thus, the potential contribution of chaperoned molecules to TLR2 activity of other lipoproteins should be considered.

Glycolipids chaperoned by NA-LprG have potency similar to that of lipoprotein TLR2 agonists, e.g. the potency of NA-LprG is similar to the potency of acylated LprA (responses detected at <30 nM NA-LprG). In contrast, unchaperoned “naked” glycolipids were far less potent (responses detected at 20 µg/ml, or
approximately 1 μM). Thus, association of glycolipids with NA-LprG increased their apparent potency by a factor of 1.5-2 log; moreover, this calculation assumes that all molecules of NA-LprG were loaded with glycolipid and likely is a significant underestimate of the increase in potency afforded by association of glycolipid with NA-LprG. Thus, the carrier function of NA-LprG dramatically increases recognition of glycolipids by TLR2.

Currently, it is not known how insoluble lipid agonists of TLR2 leave their phospholipid bilayers for transport through aqueous environments to reach and bind to TLR2. For LPS and TLR4, lipid transfer to the TLR is highly dependent on a host-derived carrier, LPS binding protein [211], and CD14 [212]. Our results suggest that the carrier function of LprG may solve a similar transport problem for mycobacterial glycolipids. To our knowledge this is the first report of a pathogen-derived protein serving as a carrier for delivery of amphipathic agonists for innate immune recognition. We propose that other pathogen-derived carriers may similarly enhance the potency and efficacy of amphipathic PAMPs for induction of host responses due to the common need of the pathogen and the host to chaperone these molecules during transit through aqueous environments and deliver them at their site of function.

These findings indicate a novel strategy for the host to co-opt the function of microbial carriers to facilitate delivery of hydrophobic agonists for recognition by host receptors. Contributions of this mechanism to recognition of Mtb glycolipids may enhance host immunity to Mtb, but we cannot exclude the possibility that the pathogen benefits from this mechanism, perhaps via TLR2-dependent induction of
down-regulatory mechanisms that may contribute to immune evasion by Mtb [193] [40] [41] [49]. Similar mechanisms may contribute to innate immune recognition of other bacterial pathogens.

While we propose that LprG carries glycolipids for the purpose of envelope construction, many important questions remain about the function of LprG. To pursue this question appropriately, we have received Mtb H37Rv LprG KO from Naiz Banaiee (Stanford). Since knockout of LprG also dramatically reduces the function of p55 (and possibly the expression of p55), the knockout should be reconstituted with either LprG or LprG-V91W. These three strains and wild-type Mtb H37Rv should be used for further comparisons. This will allow for direct query of the function of the LprG pocket in a virulent strain of Mtb. Further studies could examine whether envelope construction was perturbed using assays relevant to some important concepts in the field, including drug resistance, phagosome maturation and survival within the macrophage, and virulence.

As discussed in chapter 1, the intrinsic resistance of mycobacteria to many antibiotics is ascribed to the remarkable impermeability of the cell envelope. The outer membrane of mycobacteria is relatively impermeable to hydrophilic compounds due to lipid bilayers and relatively low expression of porins [240]. The outer membrane of mycobacteria is also impermeable to hydrophobic molecules due to two traits. First, low fluidity due to long saturated hydrocarbons (i.e. the mycolic acids), and second, the presence of tri- and tetra-acylated molecules with a single head-group (the abundant PI-based glycolipids located in the outer leaflet of the outer membrane [8]. In this thesis, we propose that LprG is an important
carrier of triacylated glycolipids to the outer membrane of the mycobacterium.
Assuming the H37Rv reconstituted with LprG-V91W is an otherwise “healthy” strain, a defect in transport of triacylated molecules should selectively affect permeability to hydrophobic molecules, leaving permeability to hydrophilic molecules unchanged. Therefore, one might predict that compared to the Mtb LprG knockout reconstituted with wild-type LprG, the Mtb H37Rv LprG knockout reconstituted with LprG-V91W would be more susceptible to hydrophobic drugs, such as rifamycins, tetracyclins, macrolides, and some fluorquinolones, while resistance to hydrophilic drugs would remain unchanged. If a difference in susceptibility to hydrophobic drugs is found, it can be verified to be occurring predominantly by passage directly through the membrane (i.e not by an aqueous route) by demonstrating that permeability increases after increasing the temperature by ~10°C, as has been demonstrated for the hydrophobic fluoroquinolone norfloxacin [8]. One method to measure permeability of the outer membrane to hydrophilic solutes has been applied to the fast-grower *M. cheloniae* [64]. Jarlier *et al.* exposed mycobacteria to cephalosporins and measured the permeability of the outer membrane by the amount of cleavage product caused by beta-lactamases present in the periplasm. This has the benefit of being a more direct measure of permeability of the outer membrane than cell viability.

There are several potential flaws to the use of antibiotics to study cell wall permeability as a consequence of loss of LprG carrier function. One is the possibility that mycolic acids may be responsible for a large portion of the permeability barrier to hydrophobic antibiotics. This is suggested by numerous studies using unbiased
approaches (such as transposon mutagenesis) to describe drug resistance genes. Many of the genes involved in mycolic acid synthesis have been discovered in this manner. To my knowledge, none of the genes involved in glycolipid synthesis have been discovered in this manner. Therefore, loss of triacylated glycolipids may not change the overall permeability of the outer membrane to hydrophobic drugs.

Secondly, testing viability of the mycobacterium to certain drugs that need to enter the cytosol to exert their effect (such as the fluoroquinolones) may add two factors that complicate interpretation. First, an antibiotic that needs to enter the cytoplasm to exert its effect can be pumped out by any of of the multi-drug resistance pumps (one of which is p55). Secondly, if said drug is a hydrophobic drug, it would also require passage through the arabinogalactan layer, another barrier to hydrophobic molecules expected to be constructed independent LprG function. Still, past studies (reviewed in [8]) offer hope for this approach, and a preliminary experiment (Fig. 5.4) suggests there is some difference in the cell envelope when LprG-V91W is expressed. Testing outer membrane permeability with the use of antibiotics would demonstrate the importance of the carrier function of LprG by demonstrating that LprG has a role in envelope construction, and that said role is relevant to drug resistance. This would suggest that targeting LprG may provide a means to potentiate the activity of common, already available drugs that are currently therapeutically ineffective for mycobacterial infection because of inability to penetrate the unusual cell envelope.

Due to the caveats of testing for drug susceptibility in using these strains, a lack of change in drug susceptibility would not necessarily indicate that the cell
envelope remained unchanged due to the mutation in LprG. Thus, it would be worthwhile to take a more direct approach to examining glycolipid and lipid transport to the cell surface. Since alteration of a subset of molecules may have effects on other extractable lipids, it may be beneficial to screen broadly for typical free, extractable lipids using approaches similar to those described to elucidate the function of LppX [96].

Mycobacterial glycolipids are released from infected cells in exosomes and microvesicles [55]. Since LprG seems able to carry glycolipids in aqueous solutions, it may be worth exploring if the profile of released glycolipids is different from antigen presenting cells infected with LprG KO strain or the LprG KO strain reconstituted with LprG-V91W.

Since glycolipids are capable of delaying phagosome maturation [241], it may be worthwhile to test if glycolipids carried by LprG can also perform this function. One possible approach to test this is to load bone marrow macrophages with Ni beads coated with NA-LprG or NA-LprG-V91W purified from M. smegmatis, and quantify phagosome maturation. Readouts based on confocal microscopy developed by David Russell and his group [242] may be used to assess phagosome maturation under these conditions.

Reconstitution of the Mtb H37Rv LprG knockout with the LprG-V91W mutant would also allow testing of the importance of the carrier function of LprG for virulence. One approach to measuring virulence is to compare colony forming units (CFUs) from lungs, LN, and spleens of mice at various time points after aerosol
infection with Mtb H37Rv, Mtb H37Rv LprG kO, the Mtb H37Rv LprG KO reconstituted with LprG, and the Mtb H37Rv LprG knockout reconstituted with LprG-V91W. Readouts of growth rate, plateau CFUs, and animal survival should provide a reasonable functional readout of virulence in the mouse model.

Another goal of this work is to analyze more thoroughly the physical interaction between TLR2 extracellular domain and mycobacterial lipoproteins. The acylated lipoproteins may be anticipated to bind similar to Pam₃CSK₄ by insertion of the diacyl glycerol into the hydrophobic pocket of TLR2. The importance of the pocket for the detection of mycobacterial TLR2 agonists can be tested by mutation of the pocket of TLR2 and TLR1 (Fig. 5.6). Given the difficulties with binding studies performed thus far, it may be worthwhile to use a bioassay readout (i.e. by expressing TLR2 mutants in HEK.293 cells compared to cells expressing wild-type TLR2; controlled for TLR2 surface expression). If a decrease in responsiveness were observed, it would be interesting to produce these point mutants in our TLR-VLR constructs for use in binding assays.
Figure 5.6. Residues that define the hydrophobic pocket of TLR2 and TLR1 may be mutated to prevent ligand binding. (A) Ribbon model of TLR2 (beta strands blue, alpha helices yellow) showing five residues that contribute to the hydrophobic pocket. Because these residues are part of a beta strand, rotation around the alpha carbon is constrained, making the effects of mutations easier to predict. Shown are I341 (black) V343 (red), L312 (green), I314 (blue), and L289 (cyan, in foreground). Mutation of V343 to a bulky side chain such as tryptophan may have the best chance of success. (B) Ribbon model of TLR1 colored as in panel A, showing a candidate residue (V333) that contributes to the binding pocket and can be mutated to provide steric hindrance.

A more immediate goal is to determine the nature of glycolipid “delivery” to TLR2 by NA-LprG. One question is whether NA-LprG-glycolipid “hands-off” glycolipid to TLR2 for detection, or whether TLR2 forms a ternary complex with NA-LprG-glycolipid. This will first be tested by size-exclusion analysis NA-LprG interaction with purified TLR2 extracellular domain. There are reasons to propose
either mechanism. In this thesis, diacyl-glycerol motif shared among lipopeptides and glycolipids has been discussed as a major structural determinant of TLR2 agonists. If glycolipids are detected by binding of acyl chains into the hydrophobic pocket of TLR2 as we suggest, then one would predict a hand-off mechanism, with only a transient ternary complex, not detected in a size exclusion assay. However, non-acylated glycans can also be sensed by TLR2 through unknown mechanisms [157]. This raises the possibility that TLR2 interacts directly with carbohydrates, likely independent of its hydrophobic pocket. Thus, if TLR2 is detecting glycolipids by binding the glycans, a stable ternary complex may form. This hypothesis can be tested by creating mutants of the TLR2 pocket that interfere with detection of Pam3CSK4 that should not interfere with detection of glycolipids.
Appendix 1. CD14 is important for the detection of LprG and NA-LprG.

Figure reproduced with permission from Nature Structural and Molecular Biology [189].
Appendix 2. Proposed model of LprG:Glycolipid detection by TLR2.
Appendix 3: Correlation of thesis figures with experiment numbers.

<table>
<thead>
<tr>
<th>Figure</th>
<th>Experiment</th>
<th>Book</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.3</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.4</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.6</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.8</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>MGD258</td>
<td>4</td>
<td>42-46</td>
</tr>
<tr>
<td>2.2</td>
<td>MGD433</td>
<td>6</td>
<td>43-67</td>
</tr>
<tr>
<td>2.3</td>
<td>MGD429,430</td>
<td>6</td>
<td>34-37</td>
</tr>
<tr>
<td>2.4</td>
<td>MGD429,430</td>
<td>6</td>
<td>34-37</td>
</tr>
<tr>
<td>2.5</td>
<td>MGD425</td>
<td>6</td>
<td>10 thru 23</td>
</tr>
<tr>
<td>2.6</td>
<td>MGD420,421, 425, 428</td>
<td>5;6</td>
<td>139-145; 10-33</td>
</tr>
<tr>
<td>3.1</td>
<td>MGD539, NDP267, AA35</td>
<td>7</td>
<td>135-139</td>
</tr>
<tr>
<td>3.S-1</td>
<td>NDP257</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.2</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.S-2</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.S-3</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.3</td>
<td>AA8, AA22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.4</td>
<td>MGD542, MGD543</td>
<td>7</td>
<td>143-146</td>
</tr>
<tr>
<td>3.5</td>
<td>MGD545-548, 568</td>
<td>8</td>
<td>19-28</td>
</tr>
<tr>
<td></td>
<td>N/A, (protein from MGD514, 515, 519,KD16, KD18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.6</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.S-4</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.S-5</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>3.7</td>
<td>N/A (protein from MGD560)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.8</td>
<td>MGD526, 550-552, MGD557</td>
<td>8</td>
<td>30-53</td>
</tr>
<tr>
<td>3.9</td>
<td>MGD559</td>
<td>8</td>
<td>65-67</td>
</tr>
<tr>
<td>4.1</td>
<td>MGD171, MGD293</td>
<td>3;4</td>
<td>22,23; 91</td>
</tr>
<tr>
<td>4.2</td>
<td>MGD257</td>
<td>4</td>
<td>40-41</td>
</tr>
<tr>
<td>4.3</td>
<td>MGD247</td>
<td>4</td>
<td>30</td>
</tr>
<tr>
<td>4.4</td>
<td>MGD322</td>
<td>4</td>
<td>137-139</td>
</tr>
<tr>
<td>4.5</td>
<td>MGD337</td>
<td>5</td>
<td>6,7</td>
</tr>
<tr>
<td>4.6</td>
<td>MGD360, MGD377</td>
<td>5</td>
<td>32-33, 54-59</td>
</tr>
<tr>
<td>4.7</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.8</td>
<td>MGD259, 320, 339, 275</td>
<td>4;5</td>
<td>47,68,128-136;9</td>
</tr>
<tr>
<td>4.9</td>
<td>MGD205</td>
<td>3</td>
<td>102-104</td>
</tr>
<tr>
<td>4.10</td>
<td>MGD284</td>
<td>4</td>
<td>81-82</td>
</tr>
<tr>
<td>4.11</td>
<td>MGD311; 340-342,</td>
<td>4;5</td>
<td>110-111;10-15,</td>
</tr>
<tr>
<td>4.12</td>
<td>MGD269, 270</td>
<td>4</td>
<td>58-59</td>
</tr>
<tr>
<td>4.13</td>
<td>MGD272, 296</td>
<td>4</td>
<td>61, 92-100</td>
</tr>
<tr>
<td>5.1</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.2</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.3</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.4</td>
<td>MGD563</td>
<td>8</td>
<td>70,71</td>
</tr>
<tr>
<td>5.5</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.6</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 4: List of Publications


Abstracts, posters, presentations


Pecora ND, **Drage MG,** Boom WH, Harding CV. Distinct structural determinants of TLR2 signaling by mycobacterial lipoproteins LprA and LprG. Journal of Immunology 176:S95-S95 Suppl. S April 1 2006. Distinct structural determinants of TLR2 signaling by mycobacterial lipoproteins LprA and LprG. Abstract and oral presentation]
Works Cited


80. Zhang, N., et al., *The Emb proteins of mycobacteria direct arabinosylation of lipoarabinomannan and arabinogalactan via an N-terminal


