The Pennsylvania State University

The Graduate School

Eberly College of Science

SYNTHESIS OF LYSO-FORM LIPOPROTEINS AND THEIR IMPLICATIONS IN COPPER RESISTANCE AND THE HOST IMMUNE RESPONSE

A Dissertation in

Biochemistry, Microbiology, and Molecular Biology

by

Krista M. Armbruster

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2019

The dissertation of Krista M. Armbruster was reviewed and approved* by the following:

Timothy C. Meredith
Assistant Professor of Biochemistry and Molecular Biology
Dissertation Advisor
Chair of Committee

Sarah Ades
Associate Professor of Biochemistry and Molecular Biology
Associate Dean of the Graduate School

Paul Babitzke
Professor of Biochemistry and Molecular Biology

Kathleen Postle
Professor of Biochemistry and Molecular Biology

Margherita T. Cantorna
Distinguished Professor of Molecular Immunology

Wendy Hanna-Rose
Professor of Biochemistry and Molecular Biology
Head of the Department of Biochemistry and Molecular Biology

*Signatures are on file in the Graduate School

ABSTRACT

Ubiquitous in bacteria, lipoproteins are involved in an array of biological processes, including nutrient acquisition, signal transduction, protein folding, and more. While their globular protein domain can vary greatly, all lipoproteins are anchored to the cell membrane by a lipidated N-terminal cysteine residue. How lipoprotein acylation occurs is well-characterized in model Gram-negative bacteria: (i) following insertion into the cytoplasmic membrane by an N-terminal signal peptide, lipoprotein diacylglycerol transferase (Lgt) attaches a diacylglycerol moiety from a neighboring phospholipid to the thiol of the conserved cysteine, (ii) lipoprotein signal peptidase (Lsp) cleaves the signal peptide immediately N-terminal to the cysteine, then (iii) lipoprotein *N*-acyltransferase (Lnt) attaches a third acyl chain to the exposed α-amino group, completing the mature triacylated lipoprotein. Based on the absence of *lnt* sequence orthologs in their genomes, lipoproteins of Gram-positive bacteria were assumed to be diacylated. However, studies have demonstrated lipoprotein triacylation in organisms lacking *lnt*. Additional lipoprotein forms were also discovered in low-GC Firmicutes, named the peptidyl form, the *N*-acetyl form, and the lyso form, all featuring N-terminal modifications. Characterized by an *N*-acyl-*S*-monoacylglyceryl cysteine structure, the lyso form was found in *Enterococcus faecalis*, *Bacillus cereus*, *Lactobacillus delbrueckii*, and *Streptococcus sanguinis*. Following the lyso form's discovery, the enzyme(s) responsible for its synthesis, how it is synthesized, and the overall role of N-terminal lipoprotein modification in Gram-positive bacteria, which appears to differ from that of Gram-negative bacteria, remained unknown. This Dissertation has sought to answer each of these questions.

Using a complementation rescue assay of a conditional-lethal *lnt* mutant in *Escherichia coli*, we identified the enzyme that converts lipoproteins to the lyso form in *E. faecalis* and *B. cereus*, named lipoprotein intramolecular transacylase (Lit). Using a combination of microbial

genetics, traditional molecular biology approaches, and intensive characterization of lipoproteins by matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS), we found that expression of Lit in *E. coli* converts lipoproteins to the lyso form, thereby allowing studies of Lit and the lyso form in a model system. Deletion of Lit from its native organism abolishes maturation to the lyso form, resulting in diacylglycerol-modified lipoproteins.

Discovery of Lit allowed for investigation into its phylogeny, revealing an unexpected distribution of *lit* in a transmissible operon encoding several copper resistance determinants. This operon is found in environmental isolates of *Enterococcus* spp. and *Listeria monocytogenes*, the latter previously only known to elaborate diacylglycerol-modified lipoproteins. Transcriptional and mass spectrometric analyses demonstrated that expression of this Lit ortholog is induced by elevated levels of copper, in turn converting lipoproteins to the lyso form in *L. monocytogenes*. We propose that copper coordinates with the exposed α-amino group of diacylated lipoproteins at the membrane interface, facilitating entry into the cell and causing cellular stress. Lyso-lipoprotein formation may weaken this coordination and subsequent uptake, suggesting a greater physiological role for *N*-terminal lipoprotein modifications. We also demonstrated that conversion to the lyso form markedly decreases recognition by Toll-like receptor 2 (TLR2), occurring predominantly through the TLR2/6 heterodimer, providing the first insight into the innate immune response to the lyso form versus its diacylglycerol-modified precursor.

It was previously theorized that Lit functions by a novel intramolecular transacylation mechanism, by which an acyl chain from the thiol-bound diacylglycerol moiety is transferred to the α-amino group of the cysteine. To test this hypothesis, an elaborate *E. coli* mutant was constructed that is able to incorporate exogenous deuterium-labeled fatty acids into its diacylglycerol-modified lipoproteins. Conditions were identified in which membrane-reconstituted Lit appears to be active *in vitro*. Combining Lit and the labeled diacylated substrate

in a reaction, the resulting lyso-form lipoproteins will be analyzed by mass spectrometry. This may reveal a labeled *N*-acyl chain, which would confirm Lit's proposed mechanism of action.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbrev.	Definition
ABC	ATP-binding cassette
ACN	Acetonitrile
ATP	Adenosine triphosphate
Bam	β-barrel assembly machinery
BLAST	Basic Local Alignment Search Tool
BSA	Bovine serum albumin
CFU	Colony forming unit
CHCA	α-cyano-4-hydroxycinnamic acid
DA-LP	Conventional diacylated lipoprotein
DDM	n-dodecyl-β-D-maltopyranoside
DHB	Dihydroxybenzoic acid
DMEM	Dulbecco's modified eagle medium
EC_{50}	Half-maximal activation
EDTA	ethylenediaminetetraacetic acid
Eep	Enhanced expression of pheromone
FBS	Fetal bovine serum
FSL-1	Pam_2CGDPKHPKSF ligand
HEK	Human embryonic kidney
HRP	Horseradish peroxidase
HTM	Hsaing-Ning Tsai medium
IM	Inner membrane
LES	Lipoprotein export signal
Lgt	Lipoprotein diacylglycerol transferase
Lit	Lipoprotein intramolecular transacylase
Lnt	Lipoprotein N-acyltransferase
Lol	Localization of lipoproteins
Lpp	Braun's lipoprotein
LPS	Lipopolysaccharide
Lpt	Lipopolysaccharide transport
Lsp	Lipoprotein signal peptidase II
Lyso-LP	Lyso-form lipoprotein
m/z	mass-to-charge ratio

Abbrev.	Definition
MALDI	Matrix-assisted laser desorption-ionization
MS	Mass spectrometry
MWCO	Molecular weight cut off
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NJ	Neighbor-joining
NMR	Nuclear magnetic resonance
OD	Optical density
OM	Outer membrane
OMP	Outer membrane protein
PCR	Polymerase chain reaction
PMF	Peptide mass fingerprinting
PMSF	Phenylmethyl sulfonyl fluoride
POTRA	Polypeptide transport-associated
PUL	Polysaccharide utilization loci
PVDF	Polyvinylidene difluoride
ROS	Reactive oxygen species
SCAM	Substituted cysteine accessibility method
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEAP	Secreted alkaline phosphatase
Sec	Secretory
Slam	Surface lipoprotein assembly modulator
Sus	Starch utilization system
TA-LP	Triacylated lipoprotein
TAT	Twin-arginine translocation
TBSE	Tris-buffered saline with EDTA
TFA	Trifluoroacetic acid
TM	Transmembrane
TOF	Time of flight
TLR	Toll-like receptor
TSB	Tryptic soy broth
TT	Transcription terminator
wgs	Whole-genome shotgun

ACKNOWLEDGEMENTS

I'd first like to thank those who have contributed most to my scientific growth.

To my mentor Dr. Tim Meredith, thank you for the time and energy you pour into your students, not just to me but with every student that you work with, including those you teach. Together we accomplished so many things that I am extremely proud of. I am grateful to have been your first graduate student.

To Dr. Gloria Komazin, both you and Tim have a level of patience I could only dream about having myself.

To my first mentor Dr. Andrew Gulick, and to all the Gulick Lab members of that era. You took a chance on me, and I would not be in the spot I am today without your early support and encouragement. You showed me what a fun (but productive) lab environment looks like. And you unexpectedly introduced me to microbiology!

To Dr. Tania Laremore, I couldn't call myself a mass spectrometrist without you.

To my committee, Drs. Sarah Ades, Paul Babitzke, Margherita Cantorna, and Kotty Postle, thank you for your feedback and knowledge.

I've met so many people along this journey who have all impacted me so much.

To Kelvin, I don't think I could've made it through this without you. Who would've thought that we would become such good friends? You are always willing to lend an ear or give advice on an experiment. I will miss our lab banter.

To Michael and Alex, my first undergrads, turned friends.

To Dr. Briana Sprague, you are a vision of intelligence, wit, and taste.

To my fellow BMMB graduate students, the sense of community that we have built has been a source of wisdom, camaraderie, and friendship for me.

There are those who have been there for me from the very beginning. I can't thank you enough.

To Aunt Rose and Uncle Roger, for inspiring me and always reminding me to be proud of myself and my accomplishments.

To my Aunt Sciandra, for always encouraging me to exercise the right side of my brain and for showing me that your story is always being written.

To Colleen, whose example of always following your heart, even when it brings difficult and life-changing decisions, is something I admire.

To Louise, my lifetime penpal.

To those I lost along the way, Grandma and Grandpa, Aunt Jeannette and Uncle Joe, Aunt Anne, Uncle Larry. I carry you in my heart.

Of course to my mom and dad. I don't know what you envisioned for me when I was a tiny Kiddo, a Poopas if you will, but I hope that I have made you proud. You have always supported me, even when you weren't quite sure what I was doing (and maybe neither was I!). Thank you for giving me everything I've ever needed and more. I am who I am today because of you.

Thank you to Jon. We took a chance on each other, and I am so happy that we did. I have never felt alone in this because of you. You have always believed in me, supported and encouraged me, and have helped me to grow in so many ways. I couldn't ask for a better partner and friend.

And thank you to our tiny kitty Albany, who makes me smile every day with her bright green eyes, love of Fortune Cookies and meowing, and her ongoing battle with the feet under the blanket.

This work was funded in part by the National Institutes of Health (R01GM127482). The findings and conclusions herein do not necessarily reflect the view of this funding agency.

Chapter 1

Introduction to Bacterial Lipoproteins

Ubiquitous among bacteria, lipoproteins serve a wide variety of functions central to cell envelope physiology and beyond. Approximately 1 to 5% of all open reading frames are predicted to be lipoproteins depending on the species (1). They are characterized by a post-translationally lipidated N-terminal cysteine that anchors the globular protein domain to the cell's extracytoplasmic surfaces. The lipidation pattern of this cysteine was largely thought to diverge based on the organism's classification as Gram-positive or Gram-negative, however more variations can be found based on species. These variations include a thiol-linked mono- or di-acylglycerol moiety, paired with one of several possible α-amino modifications or an exposed α-amino terminus. A lipoprotein's final N-terminal structure significantly influences how bacteria are sensed by the innate immune system via Toll-like receptors (TLRs).

In this chapter, the roles of select lipoproteins are highlighted, demonstrating their importance in general bacterial physiology. The modification pathway for the conventional triacyl and diacyl lipoproteins is discussed in detail, along with the current knowledge of additional lipoprotein forms found in Gram-positive Firmicutes. A brief description of methods to study lipoproteins is provided. Finally, the connection between lipoproteins and the innate immune response is explored.

1. Physiological Role of Lipoproteins

1.1 In outer membrane biogenesis and integrity

Due to their localization in the cell envelope, many lipoproteins play roles crucial to membrane biogenesis and integrity, particularly in Gram-negative bacteria (Figure 1-1). The most notable of these is the first characterized lipoprotein: Braun's lipoprotein, or Lpp (2). With upwards of one million copies per cell, Lpp is the most abundant protein in *Escherichia coli* and was used as the model protein in the key studies elucidating the lipoprotein modification pathway (3–7). Following lipidation, Lpp is transported across the periplasm to the inner leaflet of the outer membrane, where it covalently crosslinks to peptidoglycan via its C-terminal lysine K58 (8), providing structural integrity to the cell (9). Despite its abundance and role in cell envelope physiology, Lpp can be deleted with no significant growth defects to the cell under normal conditions. However, retention of Lpp in the inner membrane is lethal, as mislocalized Lpp will also crosslink to peptidoglycan, disrupting cell homeostasis resulting in death (10). This phenomenon typically occurs when Lpp is not properly acylated or exported to the outer membrane (sections 3.1 and 3.2) (11, 12). Thus, many studies on lipoprotein modification and related downstream processes involve deleting Lpp or mutating the C-terminal lysine so that it can no longer crosslink to peptidoglycan, allowing for perturbations to these processes without the ensured lethality caused by mislocalized Lpp (11–14).

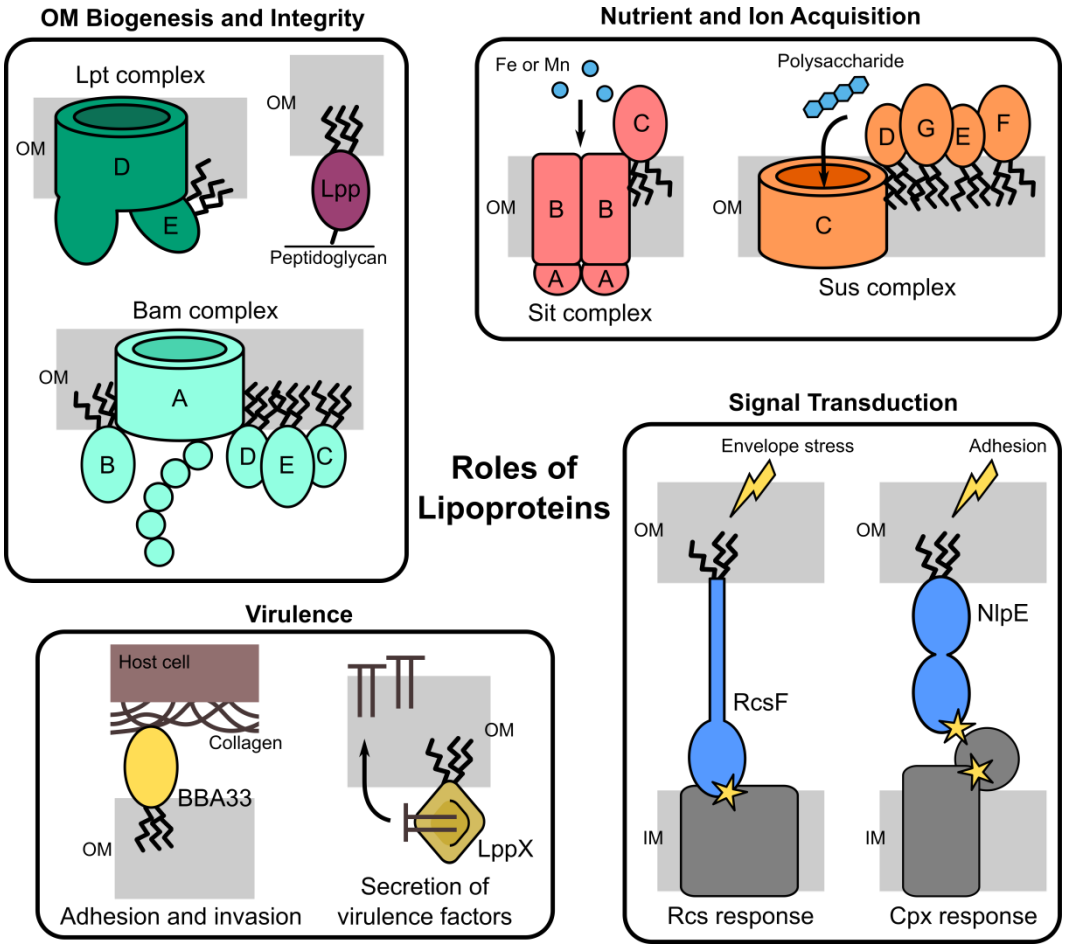


Figure **1-1**: Roles of lipoproteins. Select lipoproteins are highlighted, playing key roles in outer membrane biogenesis and integrity, nutrient and ion acquisition, signal transduction, and virulence. See text for more details.

Similarly localized to the outer membrane and consisting of a β-barrel associated with four lipoproteins, the β-barrel assembly machinery (BAM) integrates fellow outer membrane proteins (OMPs) into their functional conformation within the membrane (Figure 1-1) (15). The central player in this complex, the β-barrel BamA, features five polypeptide transport-associated (POTRA) domains that extend into the periplasm. Its partner lipoproteins, BamB, BamC, BamD, and BamE, associate with BamA and bridge the various POTRA domains to form a ring-like structure beneath the BamA β-barrel (16). Together, they facilitate docking of unfolded OMP substrate-carrying chaperones in the periplasm to BamA, allowing OMP substrate to enter the

barrel and integrate into the membrane. While all five protein components are necessary for maximal BAM activity, only BamA and the lipoprotein BamD are essential for cell viability (17, 18). Deletion or perturbations to BamD disrupts the assembly of OMPs, in turn decreasing the density of the outer membrane (17). Deletions of BamB, BamC, and BamE result in viable cells but also display defects in OMP assembly, partially through destabilization of the BAM complex (19). These mutant strains are more sensitive to rifampin, sodium dodecyl sulfate, and an array of membrane-impermeable agents (20). Thus, the BAM lipoproteins are crucial in maintaining outer membrane homeostasis.

Along with BAM, the lipopolysaccharide (LPS) transport (Lpt) machinery is another important system that contributes to outer membrane biogenesis via transport and incorporation of LPS into the outer membrane (Figure 1-1) (21). After synthesis, LPS is extracted from the inner membrane by the ATP-binding cassette (ABC) transporter LptBFG and transferred to LptC, an integral inner membrane protein (22, 23). LPS is shuttled across the periplasm along a bridge of LptA oligomers until it reaches the outer membrane complex LptDE, which then inserts LPS into the outer leaflet of the membrane (23, 24). LptD and LptE are both essential for cell viability, where LptD is a β-barrel folded by the BAM system and LptE is a lipoprotein. LptE is located almost entirely within the barrel of LptD, an association that not only plugs and stabilizes LptD, but also allows for direct interaction of LptE to LPS as it is inserted into the membrane (24–26). The presence of LptE is essential for the proper folding and assembly of LptD (25). Depletion of either LptD or LptE leads to degradation of LptA, presumably resulting in the collapse of the LptA periplasmic bridge (27). Perturbations at any level in LPS biogenesis, transport, and assembly can result in increased outer membrane permeability, altered or abnormal membranes, and even cell death (25, 28–30). These far-reaching affects highlight the importance of the lipoprotein LptE in bacterial physiology.

1.2 In nutrient and ion acquisition

Lipoproteins are frequently involved in nutrient and ion acquisition because of their localization to the extracytoplasmic membrane and cell surface (sections 3.2 and 3.3) (Figure 1-1). Exemplary in the role of lipoproteins in nutrient acquisition, genomes of the Gram-negative phylum Bacteroidetes include numerous polysaccharide utilization loci (PULs) that allow these gut inhabitants to capture and metabolize various polysaccharides. PULs are largely classified based on their similarity to the canonical starch utilization system (Sus) of *Bacteroides thetaiotaomicron* (31, 32). Sus operons encode seven proteins directly involved in polysaccharide uptake and degradation, SusABCDEFG, plus its divergently-transcribed regulator SusR (33). SusDEFG are all surface-exposed lipoproteins. SusD and the amylase SusG have one starch binding site, while SusE and SusF have multiple binding sites, allowing for binding and breakdown of polysaccharides from the environment (33). These lipoproteins form a complex with SusC, an outer membrane β-barrel that imports the oligosaccharide products into the periplasm in a TonB-dependent manner, where they are further degraded by the water-soluble amylases SusA and SusB (33). Initial studies suggested that SusCD is the main starch-binding subunit along with SusG, without which *B. thetaiotaomicron* cannot grow on starch (34, 35). However, SusEF is highly represented within gut bacterial genomes (36), suggesting the importance of all four lipoproteins of the Sus system. *B. thetaiotaomicron* encodes 88 distinct PULs, each responsible for the uptake of a different substrate, and their prevalence is shared among other Bacteroidetes species (37, 38). These PULs not only allow for effective nutrient acquisition and adaptation to changing environments, but can also greatly influence host health via microbe-assisted catabolism of dietary fibers and production of key metabolites (39, 40). The determinants of polysaccharide recognition and capture, lipoproteins of PULs are crucial in this process.

As seen with PULs, substrate-induced expression of lipoproteins involved in nutrient or ion acquisition is widespread. In the staphylococci, for example, an array of lipoproteins are significantly upregulated during iron starvation that help bacteria scavenge for this essential metal, particularly within a host (Figure 1-1) (41, 42). These lipoproteins have a number of substrates, including siderophores and heme or hemoproteins, while some interact directly with the iron ion itself (41). All are complexed with at least one transmembrane protein, typically an ABC transporter, through which the substrate enters the cell once bound by the lipoprotein (41). The SitABC complex is one such transporter, with the lipoprotein SitC binding free ions from the environment. Originally implicated in iron uptake in *Staphylococcus epidermidis* (43), the *S. aureus* SitC homolog has been shown to bind manganese (44). Regardless of its substrate, SitC is the most abundant lipoprotein in these species and has become the model for lipoprotein studies in *Staphylococcus* spp. (section 4.2) (45–48). Overall, *Staphylococcus* spp. devote approximately 30-40% of their total encoded lipoproteins to nutrient and ion uptake (42, 45). While the percentage may vary across different genomes, the importance of lipoproteins in nutrient or ion acquisition is shared by bacteria (42, 49, 50).

1.3 Other common roles of lipoproteins

A connection to the extracellular environment, many lipoproteins are involved in signal transduction and sensory systems, with roles in maintaining metal homeostasis and responding to various stress responses (Figure 1-1) (51–55). Perhaps the best known examples are the *E. coli* Rcs and Cpx systems, responding to envelope stressors such as perturbations to membrane integrity and misfolded or mislocalized proteins (55). Both systems are activated by an outer membrane lipoprotein, RcsF for Rcs and NlpE for Cpx, in a two-component signaling pathway (55). RcsF associates with outer membrane β-barrels, including the Bam complex (section 1.1), to monitor β-barrel activity and LPS levels (56, 57). Under stress conditions, RcsF interacts with an

inner membrane inhibitor of the Rcs system, in turn prompting a phosphorelay between Rcs proteins, leading to regulation of target genes (56). NlpE responds to its mislocalization, indicative of defects in lipoprotein biogenesis and trafficking (sections 3.1 and 3.2), and to cellular adhesion to a surface (58, 59). When activated, similar to RcsF, NlpE interacts with the sensor kinase CpxA in the periplasm, triggering the phosphorelay and transcriptional response associated with the Cpx system (60). Across the many various two-component and other sensory systems, lipoproteins contribute significantly to many aspects of cell homeostasis.

Lipoproteins are highly implicated in the virulence of pathogens by mediating processes such as cellular adhesion, invasion, evasion, immunomodulation (section 6), and survival in nutrient-limiting conditions (Figure 1-1) (41, 51, 61–63). Numerous individual lipoproteins have been characterized with respect to virulence, such as *Mycobacterium tuberculosis* LppX that transports virulence factors to the outer membrane (64), and BBA33 of the Lyme disease-causing *Borrelia burgdorferi*, a collagen-binding lipoprotein implicated in host infection (65). Reports of lipoproteins encoded on pathogenicity islands further highlight their role in virulence (66–68), with many explored as vaccine targets (51, 63). Studies perturbing the greater lipoprotein modification pathway show the global influence of lipoproteins on virulence, explored in more detail in sections 2.2 and 2.3.

2. Conserved Lipoprotein Modification Pathway

2.1 The preprolipoprotein signal peptide

Lipoproteins are first translated in the cytoplasm as preprolipoprotein precursors that are targeted to the cytoplasmic membrane by a signal peptide. The signal peptide is typically approximately 20 amino acids in length and can be divided into three main regions: (i) the amino-terminal n-region, characterized by the presence of positively-charged amino acids, (ii) the

middle h-region, consisting of hydrophobic amino acids that comprise the transmembrane pass, and (iii) the carboxy-terminal c-region, encoding the conserved four amino acids of what is known as the "lipobox" (Figure 1-2) (69, 70). The lipobox, presently characterized as $[LVI]_{-3}[ASTVI]_{-2}[GAS]_{-1}C_{+1}$, flanks the site at which the signal peptide is cleaved from the lipoprotein precursor, where the invariable +1 cysteine becomes the N-terminal lipidated residue characteristic of mature lipoproteins (1, 51, 69, 71). The signal peptide and lipobox is the basis for several algorithms and searches to identify putative lipoproteins in bacterial genomes, with the initial lipobox sequence widening as more lipoproteins have been identified (1, 72–75). Ultimately, upon recognition the preprolipoprotein crosses the membrane in an unfolded state by the general secretory (Sec) pathway, or less frequently in a folded state by the twin-arginine translocation (TAT) pathway (51, 71, 76–78). As the preprolipoprotein is not yet lipidated at this stage, the transmembrane signal peptide is its sole tether to the cytoplasmic membrane.

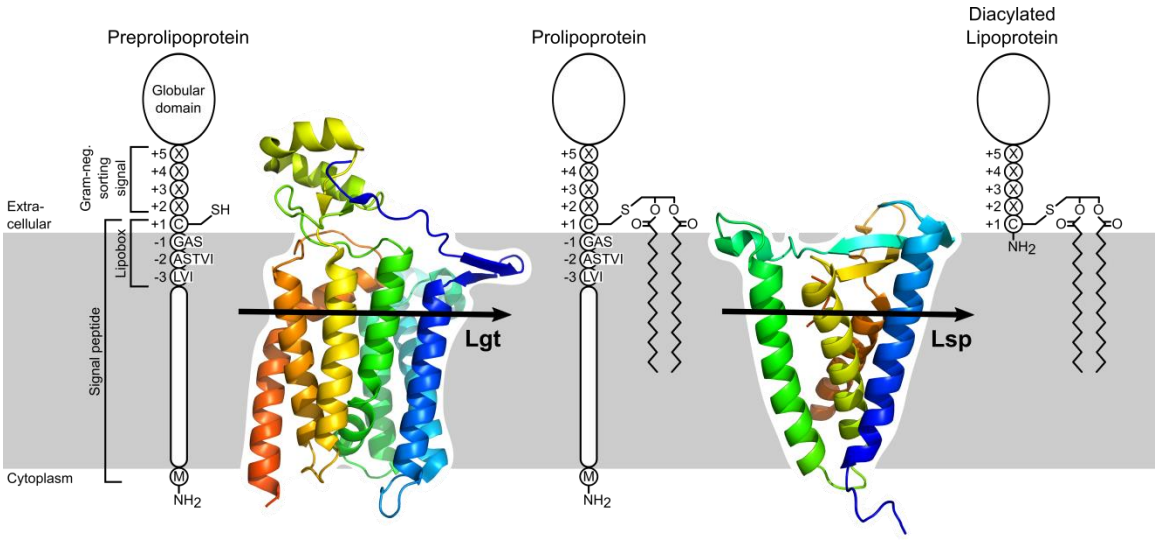


Figure **1-2**: Conserved lipoprotein modification pathway. The preprolipoprotein is translocated across the cytoplasmic membrane by the Sec or TAT machinery (not shown) and is anchored by its transmembrane signal peptide. At the C-terminus of the signal peptide, the lipobox consists of four conserved residues, including the essential +1 cysteine of lipoproteins. Beyond the cysteine, the next four residues comprise a sorting signal in Gram-negative species, dictating lipoprotein retention in the inner membrane or export to the outer membrane primarily via the +2 amino acid

(section 3.2). These residues also dictate surface exposure of outer membrane lipoproteins. Lipoprotein diacylglycerol transferase (Lgt) transfers a diacylglycerol moiety from a membrane phospholipid to the cysteine thiol to create the prolipoprotein. Then, lipoprotein signal peptidase cleaves the signal peptide, liberating the lipoprotein's N-terminus at the acylated cysteine. These steps in lipoprotein modification appear to be conserved across all bacteria, with further triacylation and other N-terminal modifications varying between Gram-positive and Gram-negative species. Structures of Lgt from *E. coli* and Lsp from *P. aeruginosa* are available in the Protein Data Bank (PDB) under accession numbers 5AZB and 5DIR, respectively (79, 80).

2.2 The lipoprotein diacylglycerol transferase (Lgt)

The preprolipoprotein is first lipidated by the lipoprotein diacylglycerol transferase (Lgt), which catalyzes the transfer of a diacylglycerol moiety from a membrane phospholipid to the sulfhydryl of the preprolipoprotein's +1 cysteine to form a thioester bond, resulting in the prolipoprotein (Figure 1-2) (5). It was first definitively characterized from a temperature-sensitive *lgt* mutant of *Salmonella typhimurium* that accumulated unmodified Lpp (81). Lgt is functionally conserved across the domain bacteria, with some species even encoding two copies of the gene (82–85). Indeed, cross-complementation of Lgt between genera is possible, as demonstrated by complementation of a temperature-sensitive *lgt* mutant of *E. coli* by *S. aureus lgt* (86). Its conservation within bacteria suggests that the initial stages of lipoprotein modification are similar even between distantly related bacteria.

Despite Lgt's ubiquity among all bacteria, it is only essential in Gram-negative species, mostly due to the lethality caused by deficient processing of Lpp (section 1.1). Conversely, Lgt has largely been shown as dispensable for cell viability in Gram-positive species, though its loss does result in an array of phenotypes (61, 87). Deletion of *lgt* leads to the release of lipoproteins from the cell surface, a method that has been used to experimentally identify lipoproteins in some strains (50, 83, 85, 88–90). Several *lgt* mutants show reduced growth in minimal media, as many lipoproteins are involved in nutrient and ion acquisition (section 1.2) (45, 50, 91). Sporulation is negatively impacted in *Bacillus subtilis* and *Bacillus anthracis* (92, 93), as is protein secretion

(85, 94), while another common phenotype of Lgt inactivation is increased sensitivity to oxidative stress (88, 95).

Lgt, and by extension lipoproteins, has repeatedly been implicated in the virulence of Gram-positive pathogens, though to varying degrees depending on the organism. Many *lgt*-deficient strains show decreased virulence and impaired TLR2-mediated activation of immune cells (section 6), including *Streptococcus pneumoniae*, *Enterococcus faecalis*, and *Listeria monocytogenes* and *Listeria innocua* (89, 96–98). Due to defects in sporulation, Lgt is necessary for the full virulence of *B. anthracis* (93). A strain of *S. pneumoniae* lacking *lgt* also displayed reduced growth and virulence in mouse models (86), however mutation to Lgt in the related *Streptococcus agalactiae* instead resulted in increased lethality in mice (91). Similarly, hypervirulence was observed from Δ*lgt* mutants of *S. aureus* (99, 100). Together, these studies highlight the impact of lipoproteins while also demonstrating the phenotypic differences that result between bacteria even on a species-specific level.

While studies over several decades have sought to characterize Lgt, the crystal structure of *E. coli* Lgt has offered the greatest insights into its elusive mechanism (79). An inner membrane protein, Lgt was shown to have seven transmembrane domains, validating previous topology mapping studies that used a combination of Lgt fusions to β-galatosidase and alkaline phosphatase, and the substituted cysteine accessibility method (SCAM) (101). These domains comprise the core of the protein and enclose a hydrophobic cavity where several conserved and potentially catalytic residues reside (79). Investigation into these residues, and those identified previously (101), by complementation of a conditional-lethal *lgt*-depletion strain of *E. coli* with Lgt harboring targeted point mutations, revealed many to be essential or important to transacylation activity (79). This implicated several catalytically-essential and hydrogen-bond forming residues as contributing to the observed lipid substrate specificity of Lgt, which has long been known to prefer phosphatidylglycerol and, to a lesser extent, phosphatidic acid (5). The

crystal structure showed two putative binding pockets for phosphatidylglycerol, flanked by clefts at the front and side of the enzyme that could serve as potential entrances for the lipid and preprolipoprotein substrates, as well as an exit for the glycerol-1-phosphate byproduct (79). This led to two theories of the mechanism of Lgt: (i) one where both pockets are occupied with lipid substrate, continually positioned for transfer to the next preprolipoprotein, or (ii) where the lipid substrate enters the first pocket, then slides into the second when preprolipoprotein substrate also binds (79). Better understanding of Lgt and its mechanism may lead to development of inhibitors and novel antibiotics.

2.3 The lipoprotein signal peptidase (Lsp)

The next step in lipoprotein maturation is cleavage of the prolipoprotein's signal peptide at the amino-terminal side of the diacylglycerol-modified +1 cysteine by lipoprotein signal peptidase II (SPase II, or more commonly Lsp) (Figure 1-2) (6). The cleaved signal peptide is rapidly degraded primarily by protease IV, a membrane-bound signal peptide peptidase, along with the cytoplasmic oligopeptidase A (102). Similar to Lgt (section 2.2), Lsp is functionally conserved throughout bacteria, with an additional copy of the gene found in some genomes (61, 82, 103). Different Lsp enzymes can cross-complement between genera, again demonstrated with a conditional-lethal *lsp* mutant of *E. coli* and the *lsp* of *S. aureus*, as well as that of *B. subtilis* (13, 104).

Lsp is essential in Gram-negative bacteria, as retention of the signal peptide prevents further lipoprotein modification and transport to the outer membrane, resulting in death. As seen with Lgt (section 2.2), Lsp is largely nonessential to cell viability in Gram-positive bacteria, though deletion of the gene does lead to several phenotypes, including the variability in virulence observed in *lgt* mutants (61, 87). While a Δ*lsp* strain of *Streptococcus suis* was show to be as virulent as the wildtype strain in a piglet infection model (105), deletion of *lsp* in *S. pneumoniae*

instead decreased virulence in mouse models (106). Studies have also implicated *lsp* in the full virulence of *S. aureus* in mice (100, 107). Similarly, both *L. monocytogenes* and *Mycobacterium tuberculosis* Δ*lsp* mutants displayed attenuated virulence and reduced growth within, and phagosomal escape from, macrophages (103, 108). In fact, expression of *lsp* is strongly induced in *L. monocytogenes* when engulfed by a macrophage (103).

Early studies on Lsp showed that diacylglycerol-modification of the prolipoprotein by Lgt is a prerequisite for Lsp-catalyzed cleavage of the signal peptide in Proteobacteria (109), however there is increasing evidence that this does not hold true for all bacteria. In fact, many of the studies on Lgt that showed shedding of unmodified lipoproteins into the culture media were due to promiscuous Lsp enzymes (section 2.2). Lsp activity has been observed in Δ*lgt* strains of *S. agalactiae*, *L. monocytogenes*, *Streptomyces coelicolor*, and *Streptomyces scabies* (50, 83, 84, 91). Cleavage of the preprolipoprotein signal peptide has also been observed of type I signal peptidases and the enhanced expression of pheromone (Eep) metallopeptidase (90, 91, 110). More often, deletion of Lsp simply yields the expected diacylglycerol-modified lipoprotein with the signal peptide attached. Regardless of the presence or absence of Lgt, a retained signal peptide anchors the lipoprotein to the cell surface, which may mediate phenotypes caused by lipoprotein release into the environment. This is not guaranteed, however, as an *lsp*-deficient strain of *S. pneumoniae* that maintains its lipoproteins at the cell surface still demonstrated defects in lipoprotein function and virulence (106). Taken together, lipoprotein modification by both Lgt and Lsp is an intertwined and complex process that is necessary for normal cell physiology.

The crystal structure of *Pseudomonas aeruginosa* Lsp has been solved in complex with globomycin, an antibiotic produced by certain *Streptomyces strains* (80, 111). Globomycin is a well-established inhibitor of Lsp, enabling numerous studies on Lsp and the greater lipoprotein modification pathway, as well as a method to identify putative lipoproteins (87, 112). In

agreement with previous topology studies on Lsp from *E. coli* (113), *P. aeruginosa* Lsp is an inner membrane protein with four transmembrane helices and two periplasmic loops that sit at the membrane interface, the larger of which forms an amphiphilic β-sheet structure (80). These two loops comprise a cleft that may act to clamp prolipoprotein substrate in the active site as the signal peptide is cleaved (80). Embedded within the core of Lsp and facing the periplasmic side of the membrane, the active site features many of the residues identified as important in *B. subtilis* Lsp activity, supporting their roles in catalysis and substrate recognition (80, 114). The cocrystallized globomycin molecule sits within the active site, stabilized by hydrogen bonding and hydrophobic interactions, sterically blocking entry of prolipoprotein substrate into Lsp for cleavage (80). Studies on Lsp and its inhibition by globomycin continue in the efforts of effective antibiotic development.

3. Lipoprotein *N*-terminal Modification in Gram-negative Bacteria

3.1 The lipoprotein *N*-acyltransferase (Lnt)

While it appears that all bacteria process their lipoproteins by Lgt and Lsp, further lipoprotein modification deviates in what was thought to be based on classification of Gram-positive versus Gram-negative. Emerging studies have challenged this paradigm, explored in detail in section 4. Nevertheless, once the signal peptide is cleaved by Lsp, Gram-negative bacteria attach a third acyl chain taken from a neighboring phospholipid to the exposed α-amino terminus of the +1 cysteine of their lipoproteins, catalyzed by the lipoprotein *N*-acyltransferase (Lnt) (Figure 1-3) (7). Triacylation in Gram-negative bacteria was determined shortly after the discovery of lipoproteins (2), with most early studies on Lnt conducted in *E. coli* using Lpp as a model lipoprotein (section 1.1).

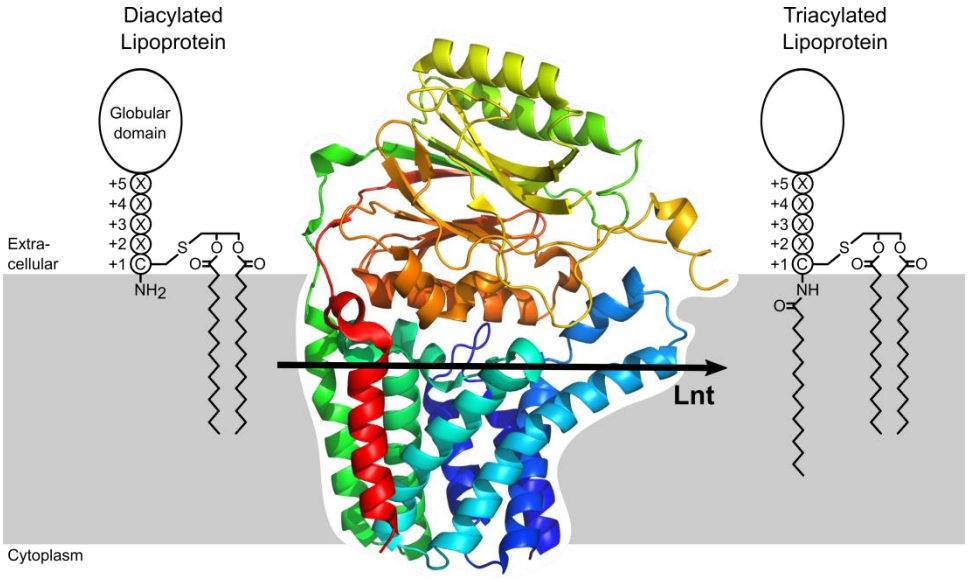


Figure **1-3**: Lipoprotein triacylation in most Gram-negative bacteria. Lipoprotein *N*-acyltransferase (Lnt) catalyzes the addition of a third acyl chain, taken from the *sn*-1 position of a membrane phospholipid, to the α-amino terminus of a diacylated lipoprotein to create the mature triacylated form found in most Gram-negative bacteria. A similar process occurs in Gram-positive Actinobacteria. Structure of Lnt from *E. coli* is available under the PDB accession number 5XHQ (115).

Lipoprotein *N*-acylation by Lnt is crucial for recognition of mature lipoproteins to be exported to the outer membrane, discussed further in section 3.2, where they perform many important functions in outer membrane biogenesis, nutrient uptake, and more (section 1). It is the resulting accumulation of diacylated lipoproteins in the inner membrane that is the main cause of death in *lnt*-deficient cells, largely through the linkage of mislocalized Lpp to peptidoglycan from the inner membrane, and why Lnt is essential in Gram-negative organisms (section 1.1) (10, 11). Disruption of Lnt also results in cell envelope instability and restructuring, increased copper sensitivity, and sensitivity at both high and low temperatures (11, 14, 116, 117). Deletion or mutation of Lpp, as well as overexpression of the machinery that exports lipoproteins to the outer membrane (sections 1.1 and 3.2), can suppress the phenotypes associated with loss of Lnt (11, 14). Further investigation into Lnt's essentiality have suggested that it is dispensable for growth in species that encode an alternate protein in the canonical lipoprotein export machinery (section

3.2) (118–120). Moreover, genomes of the Gram-negative *Wolbachia*, *Buchnera*, and *Fusobacterium* spp. do not encode *lnt*, and presumably produce diacylated lipoproteins (121–123). These examples again demonstrate the complexity of lipoprotein modification and the respective differences that can occur between species.

In 2017, three independent publications simultaneously reported the structure of Lnt: two from *E. coli* (115, 124), and the third from both *E. coli* and *P. aeruginosa* (125). Similar to Lgt (section 2.2), these studies have provided significant insight into Lnt's mechanism of *N*-acylation. In agreement with previous topology mapping studies (11, 126), Lnt is an inner membrane protein with eight transmembrane helices and a large carbon-nitrogen hydrolase domain located in the periplasm. This domain harbors the active site and known catalytic residues in a cavity that is slightly raised above the outer leaflet of the inner membrane (115, 124, 125). All three studies described a hydrophobic groove leading from this cavity towards the membrane, by which lipoprotein and phospholipid substrate might enter and exit the active site (115, 124, 125). Each study also explored the essential residues of Lnt via complementation of Lnt-depleted cells, validating previously-identified residues and implicating others (115, 124–127). Many of these residues are involved in the established mechanism of Lnt as a two-step "ping-pong" reaction, where the *sn*-1-acyl chain from a membrane phospholipid, preferably phosphatidylethanolamine, is first transferred to the sulfhydryl of the catalytic C387 to form a thioester acyl-enzyme intermediate before attachment to the lipoprotein (128, 129). Release of the resulting lyso-phospholipid byproduct seemingly allows for entry of lipoprotein substrate into the cavity, where the acyl chain is transferred from Lnt to the α-amino terminus of the lipoprotein (115, 124, 125, 128–130). Structures of Lnt corroborated this mechanism, though questions remain as to how both lipoprotein and phospholipid substrate are raised above the membrane into the active site.

3.2 The localization of lipoprotein (Lol) machinery

Once *N*-acylated by Lnt, mature triacylated lipoproteins in Gram-negative bacteria are sorted by the localization of lipoprotein (Lol) machinery for either retention in the inner membrane or export to the outer membrane (77). Lipoproteins remaining in the inner membrane encode what is called the "Lol avoidance signal," characterized by an aspartate residue at the +2 position upstream of the lipidated cysteine (Figure 1-2) (131). It is also referred to as the "+2 rule." Five additional amino acid substitutions at this position also result in retention, while further studies have implicated roles of the +3 and +4 positions as well as interactions with phosphatidylethanolamine (132–134). Lipoproteins with these criteria avoid recognition by the Lol machinery (135, 136).

Lipoproteins destined for the outer membrane are exported by the Lol machinery in three phases: (i) sorting and extraction from the inner membrane by LolCDE, (ii) cross-periplasmic shuttling by LolA, and (iii) insertion into the inner leaflet of the outer membrane by LolB (Figure 1-4) (76, 77). The LolCDE complex is a mechanistically unique ABC transporter, in that it does not transport its substrate across the membrane (137). It consists of one copy each of the homologous transmembrane proteins LolC and LolE, and two copies of the cytoplasmic nucleotide binding subunit LolD (137). LolE preferentially interacts with mature *N*-acylated lipoproteins, creating a lipoprotein-ligand bound intermediate (138–140). ATP hydrolysis by LolD weakens the association of LolE to the lipoprotein, which is consequently transferred to the chaperone protein LolA via scaffolding by LolC (139, 141, 142). The hydrophobic cavity of LolA provides a binding site for the lipoprotein's lipid moieties as the LolA-lipoprotein complex traverses the aqueous periplasmic space, apparently diffusing through the peptidoglycan layer (143, 144). At the inner leaflet of the outer membrane, the lipoprotein is handed off to the receptor LolB, itself a lipoprotein bearing strikingly similar structure to LolA, in a mouth-to-mouth manner through connection of their hydrophobic cavities (143, 145). Key differences

between LolA and LolB give LolB higher affinity for the lipoprotein, facilitating the transfer (143, 146). LolB then inserts the lipoprotein into the membrane (147). All five proteins of the Lol system are essential for cell viability due to the need for proper lipoprotein localization in the outer membrane (136, 148, 149).

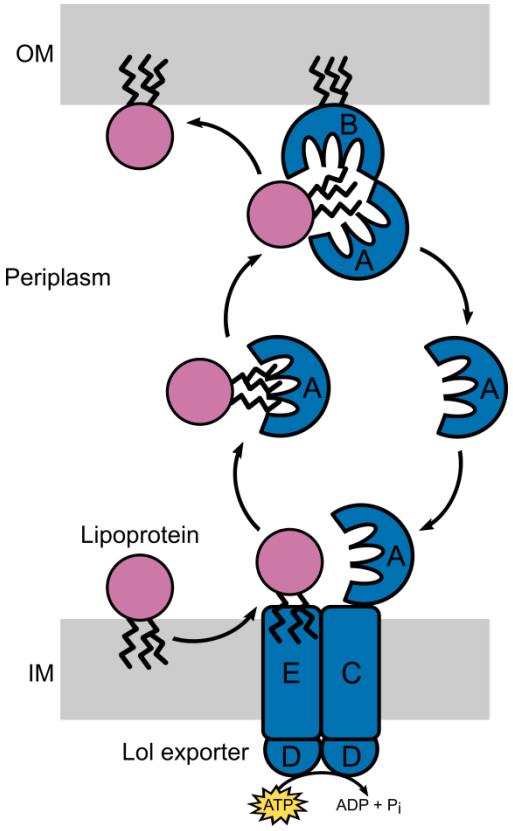


Figure **1-4**: The Lol exporter system. Lipoproteins are extracted from the inner membrane by the ATP-dependent LolCDE complex. The lipoprotein is transferred to the chaperone protein LolA, which shuttles the lipoprotein across the periplasm before handing it off to LolB in a mouth-to-mouth exchange. The lipoprotein is then inserted into the outer membrane by LolB. LolA is free to reassociate with LolCDE to shuttle another lipoprotein.

It was previously determined that *N*-acylation by Lnt is essential for recognition of lipoprotein substrate by the Lol machinery, however it was later shown that LolCDE can recognize the immature diacylated form as well, albeit at lower efficiency, and that overexpression of LolCDE suppresses the lethality caused by the absence of Lnt (section 3.1) (14,

140). Furthermore, some species encode a single transmembrane unit LolF in place of LolC and LolE that appears to readily recognize diacylated lipoproteins, allowing for deletion of Lnt (section 3.1) (118–120). LolF may be the common ancestor to LolC and LolE (139). Recent studies also suggested the presence of an alternate lipoprotein trafficking pathway that can replace the function of LolA and LolB under certain conditions (12). Regardless, the overwhelming majority of Gram-negative bacteria require Lnt and the Lol machinery for export of lipoproteins to the outer membrane to maintain normal cell physiology.

3.3 Surface exposure of lipoproteins

Insertion of lipoproteins into the membrane by LolB inherently places them in the inner leaflet of the outer membrane, where LolB is localized. For some lipoproteins, and more so in some species than others, additional processes function to expose or secrete them to the bacterial surface (76, 150, 151). It has long been known that about two-thirds of Lpp exists in a surface-exposed state, unbound to peptidoglycan, though the physiological role of this form of Lpp and how it crosses the membrane remains unclear (152, 153). Lpp is one of only four known surface-exposed lipoproteins out of the approximately 90 lipoproteins made by *E. coli* (76). The Lyme disease spirochete *Borrelia burgdorferi*, on the other hand, exposes 86 of its 125 predicted lipoproteins to the cell surface (154). It does appear that surface exposure of lipoproteins is widespread in Gram-negative bacteria, if only for just a few lipoproteins per species, and may play a role in virulence, entry of molecules into the cell, and more (151).

Efforts to determine how lipoproteins are surface exposed led to the discovery of the surface lipoprotein assembly modulator (Slam) (155). First identified in *Neisseria meningitidis*, further bioinformatic analyses suggest that Slam-like systems are conserved amongst Gram-negative bacteria (156). An outer membrane protein with a unique soluble β-barrel domain, Slam homologs demonstrate distinct lipoprotein substrate preferences (155, 156). The mechanism of

Slam-dependent surface exposure is yet to be determined. Although many of their lipoproteins are surface-exposed in Sus-like systems (section 1.2), no Slam homologs could be identified in the phylum Bacteroidetes (157). A lipoprotein export signal (LES), however, was found by bioinformatic analyses of known surface-exposed lipoproteins in Bacteroidetes, suggesting that these species share another system for lipoprotein export to the cell surface (157). Still, there are some lipoproteins whose surface exposure is executed by established bacterial systems, including the Type II secretion system in *Klebsiella pneumoniae* and the Bam system in *E. coli* (section 1.1) (57, 158, 159). Altogether, surface exposure is another facet of lipoproteins by which bacteria vary significantly.

4. Lipoprotein *N*-terminal Modification in Gram-positive Bacteria

4.1 The lipoprotein *N*-acyltransferase (Lnt) of Actinobacteria

As Gram-positive bacteria lack the outer membrane to which lipoprotein export, beginning with *N*-acylation, is so crucial (sections 3.1 and 3.2), it was assumed that lipoproteins in Gram-positive organisms were diacylated with an exposed α-amino terminus. Regardless, putative *lnt* orthologs can be found in the genomes of high-GC Actinobacteria by BLASTp searches against *E. coli* Lnt, though early studies could not validate their activity in established *E. coli* systems (127). Lnt activity in Actinobacteria was first definitively shown in the fast-growing *Mycobacterium smegmatis* (160, 161), and later in the slow-growing *Mycobacterium tuberculosis* (162). Various studies have also verified Lnt activity in other Actinobacteria, some of which encode multiple *lnt* genes (83, 87, 163).

Despite their shared homology, *E. coli* Lnt and the Lnt of Actinobacteria differ significantly. Though many of Lnt's essential residues are conserved, a catalytic serine has been described in place of the canonical active site cysteine (162), a mutation which has been shown to

abolish Lnt activity in *E. coli* (127, 129). Several Actinobacterial *lnt* genes could not rescue a conditional-lethal *lnt* mutant of *E. coli* (127, 164), challenging the precedent set by both *lsp* and *lgt* of *S. aureus* and other Gram-positive species to complement that of *E. coli* (sections 2.2 and 2.3) (13, 86), suggesting differences in acyl donor and/or substrate specificity. Indeed, the mycobacterial Lnt transfers either palmitate or tuberculostearic acid from the respective *sn*-2 and *sn*-1 positions of phospholipids (162), in contrast with the palmitate transferred from the *sn*-1 position of phospholipids by *E. coli* Lnt (128, 130). Furthermore, unlike the Lnt of Gram-negative bacteria, most *lnt* genes in Actinobacteria are dispensable for cell viability, similar to *lsp* and *lgt* in Gram-positive species (sections 2.2 and 2.3) (87). Tolerance to loss of Lnt, along with the lack of an outer membrane in Gram-positive bacteria, raises questions towards the role of lipoprotein *N*-acylation in these species.

4.2 Novel *N*-terminal lipoprotein modifications in Firmicutes

No identifiable *lnt* ortholog can be found in low-GC Gram-positive Firmicutes (160), which include the classes *Bacilli* and *Clostridia*, supporting the assumption that lipoproteins in Gram-positive organisms were diacylated. However, this does not preclude lipoprotein *N*-acylation, as the abundant ion-transport lipoprotein SitC (section 1.2) was shown to be triacylated in the Firmicute *S. aureus* (47). Several additional lipoproteins exhibited *N*-acylation across multiple strains of *S. aureus*, as well as in *S. epidermidis* (48). Although the enzyme responsible for lipoprotein *N*-acylation in these species has not yet been identified, further studies in *S. aureus* have suggested that its expression and/or activity is influenced by environmental factors, including pH, bacterial growth phase, salt concentration, and temperature (165). Lipoprotein triacylation was also described in the cell wall-less Mollicutes *Acholeplasma laidlawii* and *Spiroplasma apis*, along with several mycoplasma (161, 166–169). This indicates the existence a

functional ortholog of Lnt in Firmicutes that is structurally distinct from the Gram-negative and Acintobacterial Lnt.

Using intensive mass spectrometric analyses, further investigations across a panel of Firmicutes species led to the discovery of three novel lipoprotein N-terminal structures named the peptidyl, *N*-acetyl, and lyso forms (Figure 1-5) (167). Only found thus far in *Mycoplasma fermentans*, the peptidyl form features two additional amino acids before the diacylglycerol-modified cysteine (167). Interestingly, this strain also produces the conventional diacyl form, with lipoproteins ultimately ending in only one of the two structures (167). Thus, the peptidyl form may be the result of an Lsp with a unique cleavage site and lipoprotein substrate specificity, though this has not been experimentally confirmed (167). Similarly diacylglycerol-modified but instead featuring an amide-linked acetyl group, the *N*-acetyl form of lipoproteins is produced by the soil bacteria *Bacillus licheniformis*, *Bacillus subtilis*, and *Bacillus halodurans*, and the deep-sea bacteria *Oceanobacillus iheyensis* and *Geobacillus kaustophilus* (167). An additional study found the *N*-acetyl form in *Staphylococcus carnosus* (46), but it is not currently known how this form is made in any species.

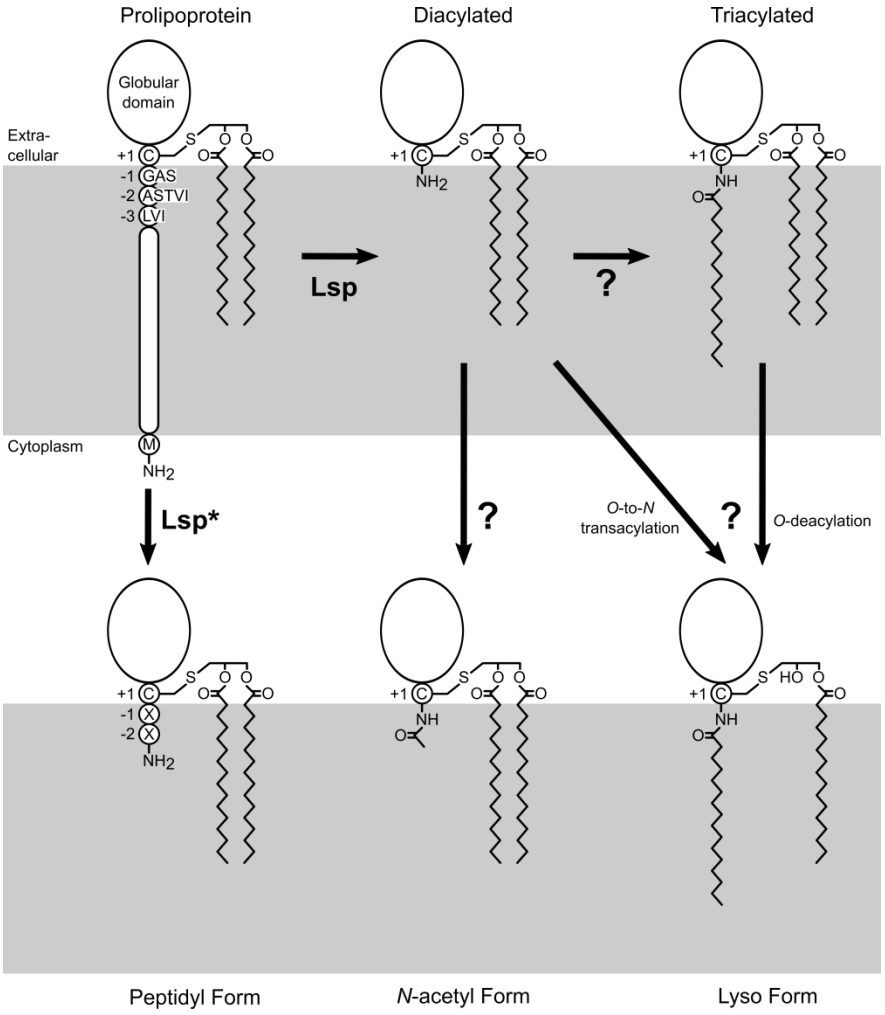


Figure **1-5**: Novel lipoprotein forms in Gram-positive Firmicutes. Starting with a prolipoprotein, it is hypothesized that the peptidyl form, featuring two amino acids before the conserved +1 cysteine, is the product of a unique Lsp enzyme (denoted with an asterisk). Processing of the prolipoprotein by the canonical Lsp results in the diacylated lipoprotein, the presumed precursor of several lipoprotein forms. The *N*-acetyl form is characterized by an amide-linked acetyl group at the *N*-terminus. The triacyl form is identical to that of Gram-negative bacteria, despite the Gram-positive species that produce it lacking an *lnt* ortholog. The lyso form, bearing an *S*-monoacylglycerol moiety and *N*-acyl chain, may be synthesized by: (i) a novel *O*-to-*N* intramolecular transacylation event, or (ii) *O*-deacylation of the triacyl form. The enzymes catalyzing lipoprotein maturation to the *N*-acetyl and triacyl forms are currently unknown, while that of the lyso form was unknown until the studies presented herein.

Lyso-form lipoproteins are made in *Enterococcus faecalis*, *Bacillus cereus*, *Lactobacillus bulgaricus*, and *Streptococcus sanguinus* (167). It is perhaps the most unique lipoprotein structure discovered to date, in the fact that it features a thiol-linked monoacylglycerol moiety, whereas all

other known lipoprotein forms have a diacylglycerol moiety (167). Its α-amino terminus is also modified by an additional acyl chain (167). Therefore, despite having two acyl chains, the lyso form is distinct from the conventional diacyl form in the positions at which the lipids are attached. It was theorized that the diacyl form is the precursor to the lyso form, with an ester-linked acyl chain moved to the *N*-terminus in a novel intramolecular transacylation event (Figure 1-5) (167). Alternately, a second theory proposes that the diacyl form is *N*-acylated in an Lnt-like mechanism concurrent with *O*-deacylation to create the lyso form (167). Until the publication that is the basis for Chapter 2 of this dissertation, it was not known what enzyme catalyzed synthesis of the lyso form.

The discovery of these novel lipoprotein forms has raised questions not just concerning their synthesis, but also the greater physiological role of their shared *N*-terminal modifications, particularly in the absence of an outer membrane. Overall, there appears to be a trend towards lipoprotein *N*-terminal modification, with Firmicutes developing or acquiring multiple systems to modify lipoproteins. Even within the same genera, more than one lipoprotein form has been observed (Figure 1-6). For example, although both *S. aureus* and *S. epidermidis* produce the triacyl form, *S. carnosus* makes the *N*-acetyl form. Similarly, several *Bacillus* spp. produce the *N*-acetyl form, while *B. cereus* makes the lyso form, suggesting that these systems were acquired independently post-speciation. Still, some species, *L. monocytogenes* for one, maintain an exposed α-amino terminus on their lipoproteins to keep the conventional diacyl form (167), adding yet another layer of species-to-species complexity to bacterial lipoproteins.

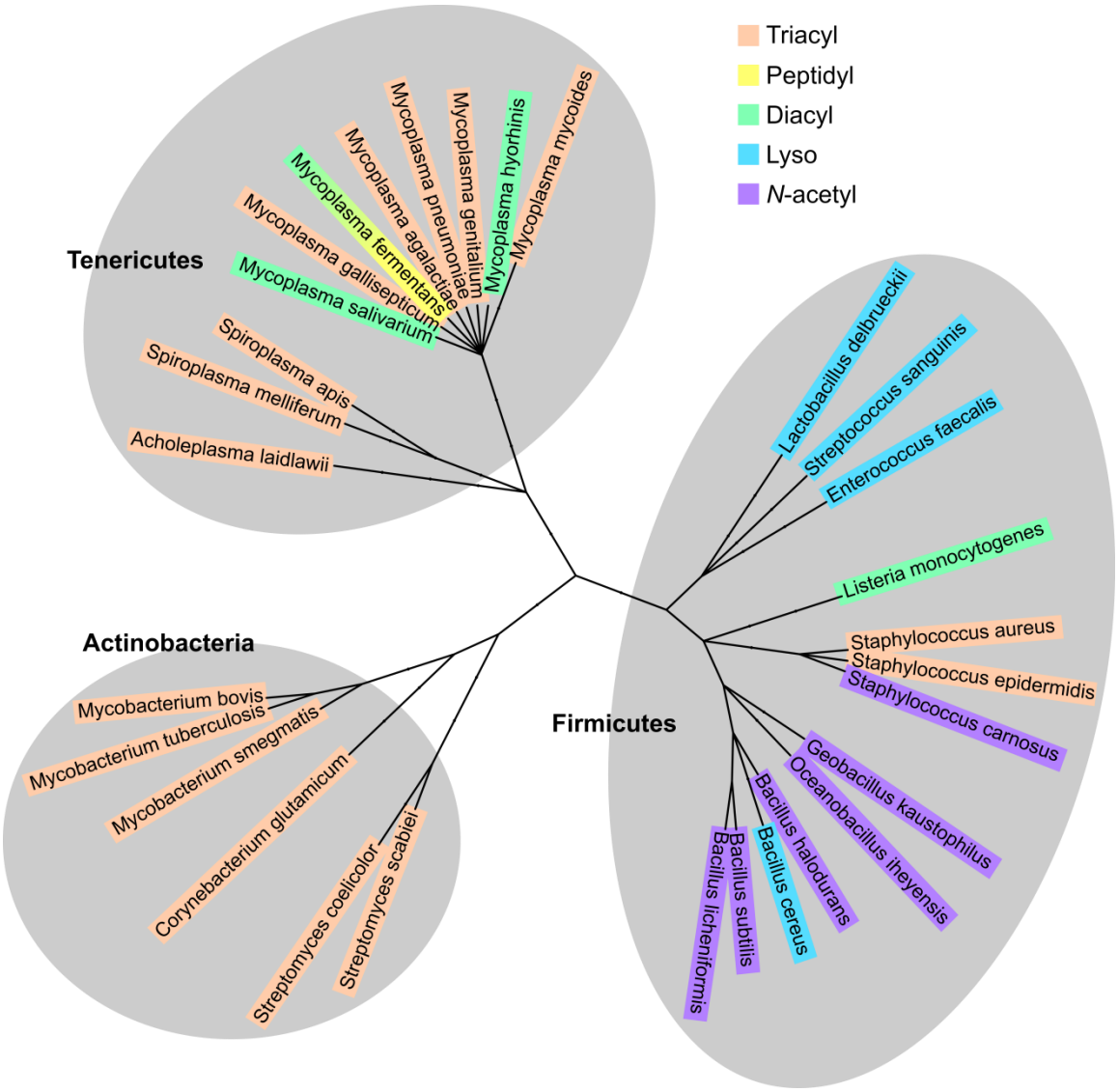


Figure **1-6**: Distribution of lipoprotein structures in Gram-positive species. An unrooted cladogram of Gram-positive Actinobacteria (83, 84, 160, 162, 163), Tenericute (166–174), and Firmicute (46, 167) species whose lipoprotein structures have been experimentally characterized, demonstrates a trend towards *N*-terminal modifications. Triacylation in Actinobacteria occurs by an ortholog of *E. coli* Lnt, while it is unknown how lipoproteins are modified in Tenericutes and Firmicutes. *Mycoplasma fermentans* produces both diacyl and peptidyl lipoproteins; see text for more detail.

5. Methods for Characterizing Lipoprotein N-terminal Structure

Following the key publications characterizing lipoprotein structure and synthesis, the majority of studies on lipoproteins focused on the activity of a single lipoprotein or lipoprotein-containing complex with little attention to their lipids or N-terminal structure. Lipoproteins are naturally difficult to work with, as most are expressed in low abundance under standard laboratory conditions, especially in Gram-positive organisms (63, 175). Their inherent hydrophobicity also presents challenges, however most of their protein domains are water-soluble and thus more easily manipulated. Therefore, many studies on lipoproteins involve cleaving the lipids or mutating the N-terminal cysteine so that no lipids can be attached to the protein, allowing for crystallographic, nuclear magnetic resonance (NMR), and other studies on lipoproteins, most of which retain their activity in the absence of their lipid domain (176–179).

For those interested in lipoprotein N-terminal structure, different techniques have traditionally been used to assess lipoprotein form. Historically, Edman degradation has been used to determine the presence or absence of an amide-linked modification. Originally developed for peptide sequencing, Edman degradation relies on sequential phenyl isothiocyanate interaction with the amino group of the N-terminal amino acid of a peptide (180). If the α-amino group of a lipoprotein is modified by an acyl chain, for example, Edman degradation will fail. This failure was used as evidence of lipoprotein *N*-acylation, and was acceptable to determine whether lipoproteins were *N*-acylated (6). However, in light of novel lipoprotein forms (section 4.2), Edman degradation is not sufficient to accurately determine lipoprotein structure, though it can still be used as a line of evidence supporting α-amino modifications (167).

Another common method of assessing lipoprotein modifications is SDS-PAGE. Though the difference in mass between the diacylated precursor and mature triacylated lipoprotein is typically around 250 Da, this difference is sufficient for separation of the two forms if the

lipoprotein itself is small, such as Lpp, and a high percentage gel is used. A larger difference in mass between the forms, perhaps a prolipoprotein with the signal peptide still attached, is more easily discernible. Indeed, SDS-PAGE has been used to assess activity of Lsp and Lnt (6, 7, 80, 115). Similar to Edman degradation, SDS-PAGE is not sufficient to distinguish between all lipoprotein forms. Edman degradation could be used to compare the triacyl versus *N*-acetyl or lyso forms, for example, but not diacyl to the lyso form, as their masses are identical. SDS-PAGE also cannot determine the positions of the modifications on a lipoprotein.

In efforts to utilize techniques that accurately determine a lipoprotein's N-terminal structure, the field has recently shifted towards mass spectrometric (MS) analysis of lipoproteins. To prepare lipoproteins for MS, a single lipoprotein is overexpressed and purified, or total lipoproteins enriched from cells using an established Triton X-114-mediated extraction method (181, 182). Lipoproteins are typically then separated by SDS-PAGE, with protocols varying beyond this step. Some protocols trypsinize the globular domain of the lipoprotein within the excised acrylamide gel plug in a buffer containing detergent before extraction of the N-terminal lipopeptide with organic solvents (48, 167, 183). Alternately, lipoproteins can be transferred to a nitrocellulose membrane, which acts as a scaffold for trypsinization of the lipoprotein (168). The resulting N-terminal tryptic lipopeptides are then eluted from the membrane with organic solvents. Both protocols serve to isolate and concentrate the N-terminal lipopeptides for analysis by MS. Some lipopeptides require only single MS analysis to determine their N-terminal structure, as is the case for triacylated lipoproteins (47, 160, 165, 168). For assignment of novel lipopeptide structures, particularly to differentiate the conventional diacyl form from the lyso form, tandem MS/MS is required (167). Lipase treatment followed by MS or MS/MS can also aid in structural determination (48, 167). Overall, MS has become instrumental in studies seeking to accurately characterize lipoprotein N-terminal structure.

6. Lipoproteins and the Host Immune Response

Lipoproteins are potent ligands of Toll-like receptors (TLRs), and as such, are highly implicated in the innate immune response. Innate immunity refers to the body's early nonspecific defense mechanisms against a wide array of antigens, including physical barriers, such as skin, and phagocytotic and immune cells, such as macrophages and monocytes (184). TLRs, of which there are ten identified in humans and thirteen in mice, are typically expressed on the surface of these immune cells (185). Each with a distinct substrate, including LPS for TLR4 and double-stranded RNA for TLR3, TLRs enable the recognition and subsequent response to both bacteria and viruses (185, 186). In general, all TLRs are composed of: (i) an extracellular or ectodomain that includes a leucine-rich repeat (LRR), (ii) a transmembrane region, and (iii) an intracellular domain (185). Binding of ligand at the ectodomain induces hetero- or homo-dimerization of the engaged TLRs, allowing the intracellular domains to interact, triggering a signaling cascade that varies depending on which TLR is stimulated (185, 186). This cascade leads to the activation of various transcription factors, one of which is NF-κB, ultimately resulting in the production of cytokines, chemokines, and interferons (185, 186). These molecules contribute to the inflammatory response, recruiting additional immune cells to the infection site and resulting in more adaptive immune responses (184).

The homologous TLRs 2, 1, and 6 respond to bacterial lipoproteins, whose N-terminal structure can significantly impact how they are sensed. The TLR2/1 heterodimer forms when engaged by triacylated lipoproteins, while the TLR2/6 heterodimer responds to diacylated lipoproteins (185–187). Crystallographic studies of these heterodimers bound by their respective ligands have significantly contributed to understanding of TLR-lipoprotein substrate specificity (Figure 1-7) (188–190). In structures of both TLR2/1 and TLR2/6, the shared diacylglycerol moiety of triacylated and diacylated lipoproteins localizes within the hydrophobic lipid-binding

pocket of TLR2 (188, 189). The *N*-acyl chain of triacylated lipoproteins inserts into a narrow hydrophobic channel of TLR1, thus requiring both TLR2 and TLR1 for proper substrate recognition (188). TLR6 shares this binding pocket, however it is blocked by two phenylalanine residues F343 and F365, precluding acyl chain insertion, giving preference of the TLR2/TLR6 complex to diacylated substrates (189). Beyond these binding pockets, interactions within the final TLR2/1 and TLR2/6 complexes also differs. Following engagement of TLR2/1, the triacylated ligand appears to bridge the two proteins, facilitating their heterodimerization (188, 190). Conversely, the TLR2/6 complex is driven by an H-bond network formed between TLR6 and the peptide of the ligand, as well as extensive hydrophobic and hydrophilic interactions at the TLR2/6 interface (189, 190). Together, these factors contribute to the high degree of substrate specificity of TLR2/1 for triacyl and TLR2/6 for diacyl lipoproteins.

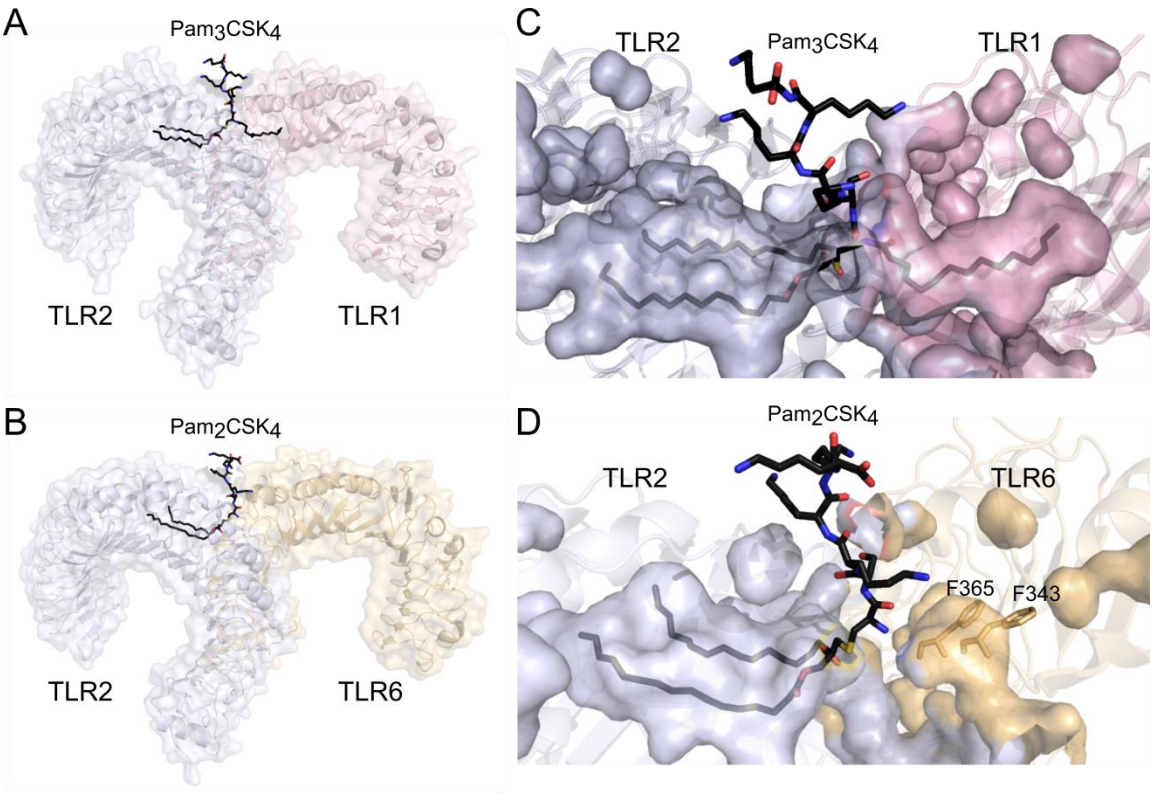


Figure **1-7**: Structures of TLR2/1 and TLR2/6 heterodimers. (A) Crystal structures of TLR2 (light blue) and TLR1 (light pink) in complex with the synthetic triacylated lipopeptide Pam_3CSK_4 (black). PDB Accession Number: 2Z7X (188). (B) Crystal structures of TLR2 (light blue) and TLR6 (light orange) in complex with the synthetic diacylated lipopeptide Pam_2CSK_4 (black). PDB Accession Number: 3A79 (189). (C, D) The cavity-filled models of TLRs 2, 1, and 6 show the binding pockets of TLR2 and TLR1 for the lipopeptide's diacylglycerol moiety and *N*-acyl chain, respectively. The binding pocket of TLR6 is blocked by two phenylalanine residues F343 and F365.

Recognition and activation of TLR2/1 and TLR2/6 to their respective ligands is well-established, however less is known about response to the novel peptidyl, *N*-acetyl, and lyso lipoprotein forms (section 4.2). Although triacylated lipoproteins of *S. aureus* are synthesized in an unknown manner distinct from that of Gram-negative bacteria, as the final lipoprotein structure does not differ, they are also ligands of the TLR2/1 heterodimer (167). The response to the peptidyl form has not yet been characterized, while both the *N*-acetyl and lyso forms signal in a TLR2-dependent manner (46, 167). Initial studies indicated that the TLR2/6 heterodimer is necessary for recognition of the *N*-acetyl form (166). Response to individual lyso-form lipoproteins occurred differentially in TLR1-/- and TLR6-/- macrophages, suggesting the lyso form engages with both TLR2/1 and TLR2/6 heterodimers, which was partially contributed to interactions with the lipoprotein peptide sequence (167). Further studies are necessary to fully characterize TLR2 response to novel lipoprotein forms, including which heterodimers are required for signaling and possible differences in activation potency.

7. Specific Aims

The unexpected discovery of triacylated and novel lipoprotein forms in Gram-positive Firmicutes challenged accepted paradigms of the lipoprotein field, namely the respective bifurcation of triacylated and diacylated lipoproteins between Gram-negative and Gram-positive bacteria. Previously, it was held that lipoprotein triacylation was necessary for proper localization

to the outer membrane. This is likely not its only purpose, however, as triacylated lipoproteins are found in several Firmicute species lacking an outer membrane. Similarly, the novel lipoprotein forms found in Firmicutes also feature various *N*-terminal lipoprotein modifications, suggesting an overall trend towards, and in turn greater physiological role, of a modified α-amino terminus. How these forms are made in Firmicutes remains unknown. Thus, in light of the information presented in this Chapter, the work described in this Dissertation aims to:

(i) Identify the enzyme responsible for lyso-form lipoprotein production in Gram-positive Firmicutes

(ii) Establish a reproducible method to characterize lipoprotein N-terminal structure

(iii) Investigate the greater physiological role of lipoprotein *N*-terminal modifications in the context of the lyso form

a. within bacteria

b. within the host via the innate immune response

(iv) Establish an assay to characterize the mechanism by which lyso-form lipoproteins are made.

References

1. **Babu MM**, **Priya ML**, **Selvan AT**, **Madera M**, **Gough J**, **Aravind L**, **Sankaran K**. 2006. A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. J Bacteriol **188**:2761–2773.
2. **Hantke K**, **Braun V**. 1973. Covalent Binding of Lipid to Protein: Diglyceride and Amide-Linked Fatty Acid at the N-Terminal End of the Murein-Lipoprotein of the *Escherichia coli* Outer Membrane. Eur J Biochem **34**:284–296.

3. **Guo MS**, **Updegrove TB**, **Gogol EB**, **Shabalina SA**, **Gross CA**, **Storz G**. 2014. MicL, a new σE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. Genes Dev **28**:1620–1634.
4. **Li GW**, **Burkhardt D**, **Gross C**, **Weissman JS**. 2014. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. Cell **157**:624–635.
5. **Sankaran K**, **Wu HC**. 1994. Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. J Biol Chem **269**:19701–19706.
6. **Hussain M**, **Ichihara S**, **Mizushima S**. 1982. Mechanism of signal peptide cleavage in the biosynthesis of the major lipoprotein of the *Escherichia coli* outer membrane. J Biol Chem **257**:5177–5182.
7. **Gupta SD**, **Wu HC**. 1991. Identification and subcellular localization of apolipoprotein N-acyltransferase in *Escherichia coli*. FEMS Microbiol Lett **62**:37–41.
8. **Braun V**, **Rehn K**, **Wolff H**. 1970. Supramolecular Structure of the Rigid Layer of the Cell Wall of *Salmonella*, *Serratia*, *Proteus*, and *Pseudomonas fluorescens*. Number of Lipoprotein Molecules in a Membrane Layer. Biochemistry **9**:5041–5049.
9. **Braun V**, **Rehn K**. 1969. Chemical Characterization, Spatial Distribution and Function of a Lipoprotein (Murein-Lipoprotein) of the *E. coli* Cell Wall: The Specific Effect of Trypsin on the Membrane Structure. Eur J Biochem **10**:426–438.
10. **Yakushi T**, **Tajima T**, **Matsuyama S**, **Tokuda H**. 1997. Lethality of the Covalent Linkage between Mislocalized Major Outer Membrane Lipoprotein and the Peptidoglycan of *Escherichia coli*. J Bacteriol **179**:2857–2862.
11. **Robichon C**, **Vidal-Ingigliardi D**, **Pugsley AP**. 2005. Depletion of apolipoprotein *N*-acyltransferase causes mislocalization of outer membrane lipoproteins in *Escherichia coli*. J Biol Chem **280**:974–983.

12. **Grabowicz M**, **Silhavy TJ**. 2017. Redefining the essential trafficking pathway for outer membrane lipoproteins. Proc Natl Acad Sci U S A **114**:4769–4774.
13. **Zhao XJ**, **Wu HC**. 1992. Nucleotide sequence of the *Staphylococcus aureus* signal peptidase II (lsp) gene. FEBS Lett **299**:80–84.
14. **Narita S ichiro**, **Tokuda H**. 2011. Overexpression of LolCDE allows deletion of the *Escherichia coli* gene encoding apolipoprotein *N*-acyltransferase. J Bacteriol **193**:4832–4840.
15. **Leyton DL**, **Belousoff MJ**, **Lithgow T**. 2015. The β-barrel assembly machinery complex. Meth Mol Biol **1329**:1-16.
16. **Han L**, **Zheng J**, **Wang Y**, **Yang X**, **Liu Y**, **Sun C**, **Cao B**, **Zhou H**, **Ni D**, **Lou J**, **Zhao Y**, **Huang Y**. 2016. Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins. Nat Struct Mol Biol **23**:192–196.
17. **Malinverni JC**, **Werner J**, **Kim S**, **Sklar JG**, **Kahne D**, **Misra R**, **Silhavy TJ**. 2006. YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. Mol Microbiol **61**:151–164.
18. **Kahne D**, **Wu T**, **Kim S**, **Malinverni J**, **Silhavy TJ**, **Ruiz N**. 2005. Identification of a Multicomponent Complex Required for Outer Membrane Biogenesis in *Escherichia coli*. Cell **121**:235–245.
19. **Gunasinghe SD**, **Shiota T**, **Stubenrauch CJ**, **Schulze KE**, **Webb CT**, **Fulcher AJ**, **Dunstan RA**, **Hay ID**, **Naderer T**, **Whelan DR**, **Bell TDM**, **Elgass KD**, **Strugnell RA**, **Lithgow T**. 2018. The WD40 Protein BamB Mediates Coupling of BAM Complexes into Assembly Precincts in the Bacterial Outer Membrane. Cell Rep **23**:2782–2794.
20. **Onufryk C**, **Crouch ML**, **Fang FC**, **Gross CA**. 2005. Characterization of six lipoproteins in the σ^e regulon. J Bacteriol **187**:4552–4561.

21. **Sperandeo P**, **Martorana AM**, **Polissi A**. 2017. The lipopolysaccharide transport (Lpt) machinery: A nonconventional transporter for lipopolysaccharide assembly at the outer membrane of Gram-negative bacteria. J Biol Chem **292**:17981-17990.
22. **Narita S ichiro**, **Tokuda H**. 2009. Biochemical characterization of an ABC transporter LptBFGC complex required for the outer membrane sorting of lipopolysaccharides. FEBS Lett **583**:2160–2164.
23. **Okuda S**, **Freinkman E**, **Kahne D**. 2012. Cytoplasmic ATP hydrolysis powers transport of lipopolysaccharide across the periplasm in *E. coli*. Science **338**:1214–1217.
24. **Chimalakonda G**, **Chng S-S**, **Silhavy TJ**, **Kahne D**, **Ruiz N**. 2010. Characterization of the two-protein complex in *Escherichia coli* responsible for lipopolysaccharide assembly at the outer membrane. Proc Natl Acad Sci **107**:5363–5368.
25. **Chimalakonda G**, **Ruiz N**, **Chng S-S**, **Garner RA**, **Kahne D**, **Silhavy TJ**. 2011. Lipoprotein LptE is required for the assembly of LptD by the -barrel assembly machine in the outer membrane of *Escherichia coli*. Proc Natl Acad Sci **108**:2492–2497.
26. **Freinkman E**, **Chng S-S**, **Kahne D**. 2011. The complex that inserts lipopolysaccharide into the bacterial outer membrane forms a two-protein plug-and-barrel. Proc Natl Acad Sci **108**:2486–2491.
27. **Samalikova M**, **Martorana AM**, **Deho G**, **Villa R**, **Polissi A**, **Grandori R**, **Sperandeo P**. 2010. New Insights into the Lpt Machinery for Lipopolysaccharide Transport to the Cell Surface: LptA-LptC Interaction and LptA Stability as Sensors of a Properly Assembled Transenvelope Complex. J Bacteriol **193**:1042–1053.
28. **Nikaido, H; Varra M**. 1985. Molecular basis of bacterial outer membrane permeability. Microbiol Mol Biol Rev **67**:593–656.
29. **Garozzo D**, **Peano C**, **Falchi FA**, **Dehò G**, **Maccagni EA**, **Puccio S**, **Palmigiano A**, **Sperandeo P**, **De Castro C**, **Polissi A**, **Martorana AM**. 2017. Mutation and Suppressor

Analysis of the Essential Lipopolysaccharide Transport Protein LptA Reveals Strategies To Overcome Severe Outer Membrane Permeability Defects in *Escherichia coli*. J Bacteriol **200**: e00487-17.

30. **Ruiz N**, **Gronenberg LS**, **Kahne D**, **Silhavy TJ**. 2008. Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of *Escherichia coli*. Proc Natl Acad Sci **105**:5537–5542.
31. **Anderson KL**, **Salyers AA**. 1989. Biochemical evidence that starch breakdown by *Bacteroides thetaiotaomicron* involves outer membrane starch-binding sites and periplasmic starch-degrading enzymes. J Bacteriol **171**:3192–3198.
32. **Anderson KL**, **Salyers AA**. 1989. Genetic evidence that outer membrane binding of starch is required for starch utilization by *Bacteroides thetaiotaomicron*. J Bacteriol **171**:3199–3204.
33. **Foley MH**, **Cockburn DW**, **Koropatkin NM**. 2016. The Sus operon: a model system for starch uptake by the human gut Bacteroidetes. Cell Mol Life Sci **73**:2603-2617.
34. **Shipman JA**, **Cho KH**, **Siegel HA**, **Salyers AA**. 1999. Physiological characterization of SusG, an outer membrane protein essential for starch utilization by *Bacteroides thetaiotaomicron*. J Bacteriol **181**:7206–7211.
35. **Shipman JA**, **Berleman JE**, **Salyers AA**. 2000. Characterization of four outer membrane proteins involved in binding starch to the cell surface of *Bacteroides thetaiotaomicron*. J Bacteriol **182**:5365–5372.
36. **Ellrott K**, **Jaroszewski L**, **Li W**, **Wooley JC**, **Godzik A**. 2010. Expansion of the protein repertoire in newly explored environments: Human gut microbiome specific protein families. PLoS Comput Biol **6**:1–11.
37. **Gordon JI**, **Himrod J**, **Hooper L V**, **Bjursell MK**, **Xu J**, **Lynn K**, **Deng S**, **Chiang HC**. 2003. A Genomic View of the Human–Bacteroides. Science **299**:2074–2076.

38. **Xu J**, **Mahowald MA**, **Ley RE**, **Lozupone CA**, **Hamady M**, **Martens EC**, **Henrissat B**, **Coutinho PM**, **Minx P**, **Latreille P**, **Cordum H**, **Van Brunt A**, **Kim K**, **Fulton RS**, **Fulton LA**, **Clifton SW**, **Wilson RK**, **Knight RD**, **Gordon JI**. 2007. Evolution of Symbiotic Bacteria in the Distal Human Intestine. PLoS Biol **5**:1574–1586.

39. **Shreiner AB**, **Kao JY**, **Young VB**. 2015. The gut microbiome in health and in disease. Curr Opin Gastroenterol **31**:69–75.

40. **Natarajan N**, **Pluznick JL**. 2014. From microbe to man: the role of microbial short chain fatty acid metabolites in host cell biology. Am J Physiol Cell Physiol **307**:C979-985.

41. **Sheldon JR**, **Heinrichs DE**. 2012. The iron-regulated staphylococcal lipoproteins. Front Cell Infect Microbiol **2**:41.

42. **Shahmirzadi S V**, **Nguyen MT**, **Götz F**. 2016. Evaluation of *Staphylococcus aureus* lipoproteins: Role in nutritional acquisition and pathogenicity. Front Microbiol **7**:1404.

43. **Cockayne A**, **Hill PJ**, **Powell NB**, **Bishop K**, **Sims C**, **Williams P**. 1998. Molecular cloning of a 32-kilodalton lipoprotein component of a novel iron-regulated *Staphylococcus epidermidis* ABC transporter. Infect Immun **66**:3767–3774.

44. **Horsburgh MJ**, **Wharton SJ**, **Cox AG**, **Ingham E**, **Peacock S**, **Foster SJ**. 2002. MntR modulates expression of the PerR regulon and superoxide resistance in *Staphylococcus aureus* through control of manganese uptake. Mol Microbiol **44**:1269–1286.

45. **Stoll H**, **Dengjel J**, **Nerz C**, **Götz F**. 2005. *Staphylococcus aureus* deficient in lipidation of prelipoproteins is attenuated in growth and immune activation. Infect Immun **73**:2411–2423.

46. **Nguyen M-T**, **Uebele J**, **Kumari N**, **Nakayama H**, **Peter L**, **Ticha O**, **Woischnig A-K**, **Schmaler M**, **Khanna N**, **Dohmae N**, **Lee BL**, **Bekeredjian-Ding I**, **Götz F**. 2017. Lipid moieties on lipoproteins of commensal and non-commensal staphylococci induce differential immune responses. Nat Commun **8**:2246.

47. **Kurokawa K**, **Lee H**, **Roh K-B**, **Asanuma M**, **Kim YS**, **Nakayama H**, **Shiratsuchi A**, **Choi Y**, **Takeuchi O**, **Kang HJ**, **Dohmae N**, **Nakanishi Y**, **Akira S**, **Sekimizu K**, **Lee BL**. 2009. The Triacylated ATP Binding Cluster Transporter Substrate-binding Lipoprotein of *Staphylococcus aureus* Functions as a Native Ligand for Toll-like Receptor 2. J Biol Chem **284**:8406–8411.

48. **Asanuma M**, **Kurokawa K**, **Ichikawa R**, **Ryu KH**, **Chae JH**, **Dohmae N**, **Lee BL**, **Nakayama H**. 2011. Structural evidence of α-aminoacylated lipoproteins of *Staphylococcus aureus*. FEBS J **278**:716–728.

49. **Sutcliffe IC**, **Russell RRB**. 1995. Lipoproteins of gram-positive bacteria. J Bacteriol **177**:1123-1128.

50. **Baumgärtner M**, **Kärst U**, **Gerstel B**, **Loessner M**, **Wehland J**, **Jänsch L**. 2007. Inactivation of Lgt allows systematic characterization of lipoproteins from *Listeria monocytogenes*. J Bacteriol **189**:313–324.

51. **Kovacs-Simon A**, **Titball RW**, **Michell SL**. 2011. Lipoproteins of bacterial pathogens. Infect Immun **79**:548–561.

52. **Janulczyk R**, **Ricci S**, **Björck L**. 2003. MtsABC is important for manganese and iron transport, oxidative stress resistance, and virulence of *Streptococcus pyogenes*. Infect Immun **71**:2656–2664.

53. **Steyn AJC**, **Joseph J**, **Bloom BR**. 2003. Interaction of the sensor module of *Mycobacterium tuberculosis* H37Rv KdpD with members of the Lpr family. Mol Microbiol **47**:1075–1089.

54. **Saleh M**, **Bartual SG**, **Abdullah MR**, **Jensch I**, **Asmat TM**, **Petruschka L**, **Pribyl T**, **Gellert M**, **Lillig CH**, **Antelmann H**, **Hermoso JA**, **Hammerschmidt S**. 2013. Molecular architecture of streptococcus pneumoniae surface thioredoxin-fold lipoproteins crucial for extracellular oxidative stress resistance and maintenance of virulence. EMBO Mol Med **5**:1852–1870.

55. **Laloux G**, **Collet J-F**. 2017. Major Tom to Ground Control: How Lipoproteins Communicate Extracytoplasmic Stress to the Decision Center of the Cell. J Bacteriol **199**:e00216-17.

56. **Cho SH**, **Szewczyk J**, **Pesavento C**, **Zietek M**, **Banzhaf M**, **Roszczenko P**, **Asmar A**, **Laloux G**, **Hov AK**, **Leverrier P**, **Van Der Henst C**, **Vertommen D**, **Typas A**, **Collet JF**. 2014. Detecting envelope stress by monitoring β-barrel assembly. Cell **159**:1652–1664.

57. **Konovalova A**, **Perlman DH**, **Cowles CE**, **Silhavy TJ**. 2014. Transmembrane domain of surface-exposed outer membrane lipoprotein RcsF is threaded through the lumen of -barrel proteins. Proc Natl Acad Sci **111**:E4350–E4358.

58. **Shimizu T**, **Ichimura K**, **Noda M**. 2016. The surface sensor NlpE of enterohemorrhagic *Escherichia coli* contributes to regulation of the type III secretion system and flagella by the Cpx response to adhesion. Infect Immun **84**:537–549.

59. **Otto K**, **Silhavy TJ**. 2002. Surface sensing and adhesion of *Escherichia coli* controlled by the Cpx-signaling pathway. Proc Natl Acad Sci **99**:2287–2292.

60. **Delhaye A, Laloux G, Collet J-F**. 2019. The Lipoprotein NlpE Is a Cpx Sensor That Serves as a Sentinel for Protein Sorting and Folding Defects in the *Escherichia coli* Envelope. J Bacteriol **201**:e00611-18.

61. **Hutchings MI**, **Palmer T**, **Harrington DJ**, **Sutcliffe IC**. 2009. Lipoprotein biogenesis in Gram-positive bacteria: knowing when to hold ’em, knowing when to fold ’em. Trends Microbiol **17**:13-21.

62. **Becker K**, **Sander P**. 2016. *Mycobacterium tuberculosis* lipoproteins in virulence and immunity – fighting with a double-edged sword. FEBS Lett **590**:3800-3819.

63. **Nguyen MT**, **Götz F**. 2016. Lipoproteins of Gram-Positive Bacteria: Key Players in the Immune Response and Virulence. Microbiol Mol Biol Rev **80**:891–903.

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

65. **Zhi H**, **Weening EH**, **Barbu EM**, **Hyde JA**, **Höök M**, **Skare JT**. 2015. The BBA33 lipoprotein binds collagen and impacts *Borrelia burgdorferi* pathogenesis. Mol Microbiol **96**:68–83.

66. **Robertson GT**, **Child R**, **Ingle C**, **Celli J**, **Norgard M V.** 2013. IglE Is an Outer Membrane-Associated Lipoprotein Essential for Intracellular Survival and Murine Virulence of Type A *Francisella tularensis*. Infect Immun **81**:4026–4040.

67. **Nguyen MT**, **Kraft B**, **Yu W**, **Demicrioglu DD**, **Hertlein T**, **Burian M**, **Schmaler M**, **Boller K**, **Bekeredjian-Ding I**, **Ohlsen K**, **Schittek B**, **Götz F**. 2015. The νSaα Specific Lipoprotein Like Cluster (lpl) of *S. aureus* USA300 Contributes to Immune Stimulation and Invasion in Human Cells. PLoS Pathog **11**:e1004984.

68. **Whalan RH**, **Funnell SGP**, **Bowler LD**, **Hudson MJ**, **Robinson A**, **Dowson CG**. 2006. Distribution and genetic diversity of the ABC transporter lipoproteins PiuA and PiaA within *Streptococcus pneumoniae* and related streptococci. J Bacteriol **188**:1031–1038.

69. **Von Heijne G**. 1989. The structure of signal peptides from bacterial lipoproteins. Protein Eng **2**:531–534.

70. **Klein P**, **Somorjai RL**, **Lau PC**. 1988. Distinctive properties of signal sequences from bacterial lipoproteins. Protein Eng **2**:15–20.

71. **Hayashi S**, **Wu HC**. 1990. Lipoproteins in bacteria. J Bioenerg Biomembr **22**:451–471.

72. **Harrington DJ**, **Sutcliffe IC**. 2015. Pattern searches for the identification of putative lipoprotein genes in Gram-positive bacterial genomes. Microbiology **148**:2065–2077.

73. **Juncker AS**, **Willenbrock H**, **Von Heijne G**, **Brunak S**, **Nielsen H**, **Krogh A**. 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci **12**:1652–62.

74. **Bagos PG**, **Tsirigos KD**, **Liakopoulos TD**, **Hamodrakas SJ**. 2008. Prediction of lipoprotein signal peptides in Gram-positive bacteria with a Hidden Markov Model. J Proteome Res **7**:5082–5093.

75. **Setubal JC**, **Reis M**, **Matsunaga J**, **Haake DA**. 2006. Lipoprotein computational prediction in spirochaetal genomes. Microbiology **152**:113–121.

76. **Zückert WR**. 2014. Secretion of Bacterial Lipoproteins: Through the Cytoplasmic Membrane, the Periplasm and Beyond. Biochim Biophys Acta **1843**:1509-1516.

77. **Narita S**, **Tokuda H**. 2016. Bacterial lipoproteins; biogenesis, sorting and quality control. Biochim Biophys Acta **1862**:1414-1423.

78. **Tokuda H**. 2009. Biogenesis of Outer Membranes in Gram-Negative Bacteria. Biosci Biotechnol Biochem **73**:465–473.

79. **Zhao Y**, **Mao G**, **Zhang Y**, **Sankaran K**, **Zhang XC**, **Li Z**, **Kang X**, **Wang X**, **Sun F**. 2016. Crystal structure of *E. coli* lipoprotein diacylglyceryl transferase. Nat Commun **7**:10198.

80. **El Arnaout T**, **Vogeley L**, **Bailey J**, **Boland C**, **Caffrey M**, **Stansfeld PJ**. 2016. Structural basis of lipoprotein signal peptidase II action and inhibition by the antibiotic globomycin. Science **351**:876–880.

81. **Gan K**, **Gupta SD**, **Sankaran K**, **Schmid MB**, **Wu HC**. 1993. Isolation and characterization of a temperature-sensitive mutant of *Salmonella typhimurium* defective in prolipoprotein modification. J Biol Chem **268**:16544–16550.

82. **Sutcliffe IC**, **Harrington DJ**, **Hutchings MI**. 2012. A phylum level analysis reveals lipoprotein biosynthesis to be a fundamental property of bacteria. Protein Cell **3**:163–170.

83. **Widdick DA**, **Hicks MG**, **Thompson BJ**, **Tschumi A**, **Chandra G**, **Sutcliffe IC**, **Brülle JK**, **Sander P**, **Palmer T**, **Hutchings MI**. 2011. Dissecting the complete lipoprotein biogenesis pathway in *Streptomyces scabies*. Mol Microbiol **80**:1395–1412.

84. **Widdick DA**, **Hicks MG**, **Chandra G**, **Thompson BJ**, **Palmer T**, **Sutcliffe IC**, **Hutchings MI**. 2010. Investigating lipoprotein biogenesis and function in the model Gram-positive bacterium *Streptomyces coelicolor*. Mol Microbiol **77**:943–957.

85. **Grau T**, **Rezwan M**, **Antelmann H**, **Tschumi A**, **Sander P**, **Albrecht D**. 2012. Functional Analyses of Mycobacterial Lipoprotein Diacylglyceryl Transferase and Comparative Secretome Analysis of a Mycobacterial lgt Mutant. J Bacteriol **194**:3938–3949.

86. **Qi H-Y**, **Sankaran K**, **Gan K**, **Wu HC**. 1995. Structure-Function Relationship of Bacterial Prolipoprotein Diacylglyceryl Transferase: Functionally Significant Conserved Regions. J Bacteriol **177**:6820–6824.

87. **Buddelmeijer N**. 2015. The molecular mechanism of bacterial lipoprotein modification-How, when and why? FEMS Microbiol Rev.

88. **MacDonald N**, **Chimalapati S**, **Brown JS**, **Durmort C**, **Hermans PWM**, **Mitchell T**, **Camberlein E**, **Cohen JM**, **Vernet T**. 2012. Effects of Deletion of the *Streptococcus pneumoniae* Lipoprotein Diacylglyceryl Transferase Gene lgt on ABC Transporter Function and on Growth In Vivo. PLoS One **7**:e41393.

89. **Reffuveille F**, **Serror P**, **Chevalier S**, **Budin-Verneuil A**, **Ladjouzi R**, **Bernay B**, **Auffray Y**, **Rincé A**. 2012. The prolipoprotein diacylglyceryl transferase (Lgt) of *Enterococcus faecalis* contributes to virulence. Microbiology **158**:816–825.

90. **Denham EL**, **Ward PN**, **Leigh JA**. 2009. In the absence of Lgt, lipoproteins are shed from *Streptococcus uberis* independently of Lsp. Microbiology **155**:134–141.

91. **Mancuso G**, **Theilacker C**, **Hübner J**, **Santos-Sierra S**, **Dramsi S**, **Poyart C**, **Chraibi K**, **Pellegrini E**, **Henneke P**, **Trieu-Cuot P**, **Teti G**, **Golenbock DT**. 2014. Lipoproteins Are

Critical TLR2 Activating Toxins in Group B Streptococcal Sepsis. J Immunol **180**:6149–6158.

92. **Igarashi T**, **Setlow B**, **Paidhungat M**, **Setlow P**. 2004. Effects of a gerF (lgt) Mutation on the Germination of Spores of *Bacillus subtilis*. J Bacteriol **186**:2984–2991.
93. **Okugawa S**, **Moayeri M**, **Pomerantsev AP**, **Sastalla I**, **Crown D**, **Gupta PK**, **Leppla SH**. 2012. Lipoprotein biosynthesis by prolipoprotein diacylglyceryl transferase is required for efficient spore germination and full virulence of *Bacillus anthracis*. Mol Microbiol **83**:96–109.
94. **Leskelä S**, **Wahlström E**, **Kontinen VP**, **Sarvas M**. 1999. Lipid modification of prelipoproteins is dispensable for growth but essential for efficient protein secretion in *Bacillus subtilis*: Characterization of the Igt gene. Mol Microbiol **31**:1075–1085.
95. **Bray BA**, **Sutcliffe IC**, **Harrington DJ**. 2009. Impact of lgt mutation on lipoprotein biosynthesis and in vitro phenotypes of *Streptococcus agalactiae*. Microbiology **155**:1451–1458.
96. **Ocvirk S**, **Sava IG**, **Lengfelder I**, **Lagkouvardos I**, **Steck N**, **Roh JH**, **Tchaptchet S**, **Bao Y**, **Hansen JJ**, **Huebner J**, **Carroll IM**, **Murray BE**, **Sartor RB**, **Haller D**. 2015. Surface-Associated Lipoproteins Link *Enterococcus faecalis* Virulence to Colitogenic Activity in IL-10-Deficient Mice Independent of Their Expression Levels. PLoS Pathog **11**:e1004911.
97. **Machata S**, **Tchatalbachev S**, **Jansch L**, **Hain T**, **Chakraborty T**, **Mohamed W**. 2014. Lipoproteins of *Listeria monocytogenes* Are Critical for Virulence and TLR2-Mediated Immune Activation. J Immunol **181**:2028–2035.
98. **Pollard T**, **Stafford S**, **Chimalapati S**, **Vollmer W**, **Noursadeghi M**, **Tomlinson G**, **Periselneris J**, **Cohen J**, **Lapp T**, **Aldridge C**, **Brown J**, **Camberlein E**, **Picard C**, **Casanova J-L**. 2014. TLR-Mediated Inflammatory Responses to *Streptococcus pneumoniae*

Are Highly Dependent on Surface Expression of Bacterial Lipoproteins. J Immunol **193**:3736–3745.

99. **Ferracin F**, **Biswas L**, **Landolt LZ**, **Schmaler M**, **Gotz F**, **Jann NJ**, **Landmann R**. 2009. Lipoproteins in *Staphylococcus aureus* Mediate Inflammation by TLR2 and Iron-Dependent Growth *In Vivo*. J Immunol **182**:7110–7118.

100. **Bubeck Wardenburg J**, **Williams WA**, **Missiakas D**. 2006. Host defenses against *Staphylococcus aureus* infection require recognition of bacterial lipoproteins. Proc Natl Acad Sci **103**:13831–13836.

101. **Pailler J**, **Aucher W**, **Pires M**, **Buddelmeijer N**. 2012. Phosphatidylglycerol: Prolipoprotein diacylglyceryl transferase (Lgt) of *Escherichia coli* has seven transmembrane segments, and its essential residues are embedded in the membrane. J Bacteriol **194**:2142–2151.

102. **Novak P**, **Dev IK**. 1988. Degradation of a signal peptide by protease IV and oligopeptidase A. J Bacteriol **170**:5067–5075.

103. **Réglier-Poupet H**, **Frehel C**, **Dubail I**, **Beretti JL**, **Berche P**, **Charbit A**, **Raynaud C**. 2003. Maturation of lipoproteins by type II signal peptidase is required for phagosomal escape of *Listeria monocytogenes*. J Biol Chem **278**:49469–49477.

104. **Bron S**, **van Dijl JM**, **Bolhuis A**, **Pragai Z**, **Venema G**, **Tjalsma H**. 2009. The signal peptidase II (Isp) gene of *Bacillus subtilis*. Microbiology **143**:1327–1333.

105. **de Greeff A**, **Hamilton A**, **Sutcliffe IC**, **Buys H**, **van Alphen L**, **Smith HE**. 2003. Lipoprotein signal peptidase of *Streptococcus suis* serotype 2. Microbiology **149**:1399–1407.

106. **Khandavilli S**, **Homer KA**, **Yuste J**, **Basavanna S**, **Mitchell T**, **Brown JS**. 2008. Maturation of *Streptococcus pneumoniae* lipoproteins by a type II signal peptidase is required for ABC transporter function and full virulence. Mol Microbiol **67**:541–557.

107. **Mei JM**, **Nourbakhsh F**, **Ford CW**, **Holden DW**. 1997. Identification of *Staphylococcus aureus* virulence genes in a murine model of bacteraemia using signature-tagged mutagenesis. Mol Microbiol **26**:399–407.

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

109. **Tokunaga M**, **Tokunaga H**, **Wu HC**. 1982. Post-translational modification and processing of Escherichia coli prolipoprotein in vitro. Proc Natl Acad Sci **79**:2255–2259.

110. **Denham EL**, **Ward PN**, **Leigh JA**. 2008. Lipoprotein signal peptides are processed by Lsp and Eep of *Streptococcus uberis*. J Bacteriol **190**:4641–4647.

111. **Inukai M**, **Enokita R**, **Torikata A**, **Nakahara M**, **Iwado S**, **Arai M**. 1978. Globomycin, a new peptide antibiotic with spheroplast-forming activity. I. Taxonomy of producing organisms and fermentation. J Antibiot (Tokyo) **31**:410–420.

112. **Dev IK**, **Harvey RJ**, **Ray PH**. 1985. Inhibition of prolipoprotein signal peptidase by globomycin. J Biol Chem **260**:5891–5894.

113. **Munoa FJ**, **Miller KW**, **Beers R**, **Graham M**, **Wu HC**. 1991. Membrane topology of *Escherichia coli* prolipoprotein signal peptidase (signal peptidase II). J Biol Chem **266**:17667–17672.

114. **Tjalsma H**, **Zanen G**, **Venema G**, **Bron S**, **Van Dijl JM**. 1999. The potential active site of the lipoprotein-specific (type II) signal peptidase of *Bacillus subtilis*. J Biol Chem **274**:28191–28197.

115. **Zhai Y**, **Cui L**, **Lou J**, **Li H**, **Zhang XC**, **Sun F**, **Zhang K**, **Xiong Y**, **Lu G**, **Wang X**, **Xu Y**. 2017. Crystal structure of *E. coli* apolipoprotein N-acyl transferase. Nat Commun **8**:15948.

116. **Gan K**, **Gupta SD**, **Sankaran K**, **Schmid MB**, **Wu HC**. 1993. Isolation and characterization of a temperature-sensitive mutant of *Salmonella typhimurium* defective in prolipoprotein modification. J Biol Chem **268**:16544–16550.

117. **Rogers SD**, **Bhave MR**, **Mercer JF**, **Camakaris J**, **Lee BT**. 1991. Cloning and characterization of cutE, a gene involved in copper transport in *Escherichia coli*. J Bacteriol **173**:6742–6748.

118. **Gwin CM**, **Prakash N**, **Christian Belisario J**, **Haider L**, **Rosen ML**, **Martinez LR**, **Rigel NW**. 2018. The apolipoprotein N-acyl transferase Lnt is dispensable for growth in Acinetobacter species. Microbiology **164**:1547–1556.

119. **LoVullo ED**, **Wright LF**, **Isabella V**, **Huntley JF**, **Pavelka MS**. 2015. Revisiting the gram-negative lipoprotein paradigm. J Bacteriol **197**:1705–1715.

120. **Grabowicz M**. 2018. Lipoprotein Transport: Greasing the Machines of Outer Membrane Biogenesis: Re-Examining Lipoprotein Transport Mechanisms Among Diverse Gram-Negative Bacteria While Exploring New Discoveries and Questions. BioEssays **40**:1–11.

121. **Turner JD**, **Langley RS**, **Johnston KL**, **Gentil K**, **Ford L**, **Wu B**, **Graham M**, **Sharpley F**, **Slatko B**, **Pearlman E**, **Taylor MJ**. 2009. Wolbachia Lipoprotein Stimulates Innate and Adaptive Immunity through Toll-like Receptors 2 and 6 to Induce Disease Manifestations of Filariasis. J Biol Chem **284**:22364–22378.

122. **Shigenobu S**, **Sakaki Y**, **Ishikawa H**, **Hattori M**, **Watanabe H**. 2000. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. Nature **407**:81–86.

123. **Bhattacharyya S**, **Ghosh SK**, **Shokeen B**, **Eapan B**, **Lux R**, **Kiselar J**, **Nithianantham S**, **Young A**, **Pandiyan P**, **McCormick TS**, **Weinberg A**. 2016. FAD-I, a *Fusobacterium nucleatum* Cell Wall-Associated Diacylated Lipoprotein That Mediates Human Beta Defensin 2 Induction through Toll-Like Receptor-1/2 (TLR-1/2) and TLR-2/6. Infect Immun **84**:1446–1456.

124. **Xu M**, **Kattke MD**, **Reichelt M**, **Kapadia SB**, **Gloor SL**, **Yan D**, **Zilberleyb I**, **Kang J**, **Katakam AK**, **Murray JM**, **Pantua H**, **Noland CL**, **Diao J**. 2017. Structural insights into lipoprotein *N*-acylation by *Escherichia coli* apolipoprotein *N*-acyltransferase. Proc Natl Acad Sci **114**:E6044–E6053.

125. **Wiktor M**, **Weichert D**, **Howe N**, **Huang CY**, **Olieric V**, **Boland C**, **Bailey J**, **Vogeley L**, **Stansfeld PJ**, **Buddelmeijer N**, **Wang M**, **Caffrey M**. 2017. Structural insights into the mechanism of the membrane integral *N*-acyltransferase step in bacterial lipoprotein synthesis. Nat Commun **8**:15952.

126. **Gélis-Jeanvoine S**, **Lory S**, **Oberto J**, **Buddelmeijer N**. 2015. Residues located on membrane-embedded flexible loops are essential for the second step of the apolipoprotein *N*-acyltransferase reaction. Mol Microbiol **95**:692–705.

127. **Vidal-Ingigliardi D**, **Lewenza S**, **Buddelmeijer N**. 2007. Identification of essential residues in apolipoprotein *N*-acyl transferase, a member of the CN hydrolase family. J Bacteriol **189**:4456–4464.

128. **Hillmann F**, **Argentini M**, **Buddelmeijer N**. 2011. Kinetics and phospholipid specificity of apolipoprotein *N*-acyltransferase. J Biol Chem **286**:27936–27946.

129. **Buddelmeijer N**, **Young R**. 2010. The essential *Escherichia coli* apolipoprotein *N*-acyltransferase (Lnt) exists as an extracytoplasmic thioester acyl-enzyme intermediate. Biochemistry **49**:341–346.

130. **Jackowski S**, **Rock CO**. 1986. Transfer of fatty acids from the 1-position of phosphatidylethanolamine to the major outer membrane lipoprotein of *Escherichia coli*. J Biol Chem **261**:11328–11333.

131. **Yamaguchi K**, **Yu F**, **Inouye M**. 1988. A single amino acid determinant of the membrane localization of lipoproteins in *E. coli*. Cell **53**:423–432.

132. **Seydel A**, **Gounon P**, **Pugsley AP**. 1999. Testing the "+2 rule" for lipoprotein sorting in the *Escherichia coli* cell envelope with a new genetic selection. Mol Microbiol **34**:810–821.

133. **Terada M**, **Kuroda T**, **Matsuyama SI**, **Tokuda H**. 2001. Lipoprotein Sorting Signals Evaluated as the LolA-dependent Release of Lipoproteins from the Cytoplasmic Membrane of *Escherichia coli*. J Biol Chem **276**:47690–47694.

134. **Hara T**, **Matsuyama SI**, **Tokuda H**. 2003. Mechanism underlying the inner membrane retention of *Escherichia coli* lipoproteins caused by Lol avoidance signals. J Biol Chem **278**:40408–40414.

135. **Masuda K**, **Matsuyama S -i.**, **Tokuda H**. 2002. Elucidation of the function of lipoprotein-sorting signals that determine membrane localization. Proc Natl Acad Sci **99**:7390–7395.

136. **Tokuda H**, **Matsuyama SI**. 2004. Sorting of lipoproteins to the outer membrane in *E. coli*. Biochim Biophys Acta **1693**:5-13.

137. **Yakushi T**, **Masuda K**, **Narita SI**, **Matsuyama SI**, **Tokuda H**. 2000. A new ABC transporter mediating the detachment of lipid-modified proteins from membranes. Nat Cell Biol **2**:212–218.

138. **Sakamoto C**, **Satou R**, **Tokuda H**, **Narita S ichiro**. 2010. Novel mutations of the LolCDE complex causing outer membrane localization of lipoproteins despite their inner membrane-retention signals. Biochem Biophys Res Commun **401**:586–591.

139. **Mizutani M**, **Mukaiyama K**, **Xiao J**, **Mori M**, **Satou R**, **Narita SI**, **Okuda S**, **Tokuda H**. 2013. Functional differentiation of structurally similar membrane subunits of the ABC transporter LolCDE complex. FEBS Lett **587**:23–29.

140. **Fukuda A**, **Matsuyama S-I**, **Hara T**, **Nakayama J**, **Nagasawa H**, **Tokuda H**. 2002. Aminoacylation of the N-terminal cysteine is essential for Lol-dependent release of

lipoproteins from membranes but does not depend on lipoprotein sorting signals. J Biol Chem **277**:43512–43518.

141. **Ito Y**, **Kanamaru K**, **Taniguchi N**, **Miyamoto S**, **Tokuda H**. 2006. A novel ligand bound ABC transporter, LolCDE, provides insights into the molecular mechanisms underlying membrane detachment of bacterial lipoproteins. Mol Microbiol **62**:1064–1075.

142. **Kaplan E**, **Greene NP**, **Crow A**, **Koronakis V**. 2018. Insights into bacterial lipoprotein trafficking from a structure of LolA bound to the LolC periplasmic domain. Proc Natl Acad Sci **115**:E7389–E7397.

143. **Takeda K**, **Miyatake H**, **Yokota N**, **Matsuyama SI**, **Tokuda H**, **Miki K**. 2003. Crystal structures of bacterial lipoprotein localization factors, LolA and LolB. EMBO J **22**:3199–3209.

144. **Vollmer W**, **Bertsche U**. 2008. Murein (peptidoglycan) structure, architecture and biosynthesis in Escherichia coli. Biochim Biophys Acta **1778**:1714–1734.

145. **Okuda S**, **Tokuda H**. 2009. Model of mouth-to-mouth transfer of bacterial lipoproteins through inner membrane LolC, periplasmic LolA, and outer membrane LolB. Proc Natl Acad Sci **106**:5877–5882.

146. **Taniguchi N**, **Matsuyama SI**, **Tokuda H**. 2005. Mechanisms underlying energy-independent transfer of lipoproteins from LolA to LolB, which have similar unclosed β-barrel structures. J Biol Chem **280**:34481–34488.

147. **Tsukahara J**, **Mukaiyama K**, **Okuda S**, **Narita SI**, **Tokuda H**. 2009. Dissection of LolB function - Lipoprotein binding, membrane targeting and incorporation of lipoproteins into lipid bilayers. FEBS J **276**:4496–4504.

148. **Narita S**, **Tanaka K**, **Matsuyama S**, **Tokuda H**. 2002. Disruption of lolCDE, encoding an ATP-binding cassette transporter, is lethal for *Escherichia coli* and prevents release of lipoproteins from the inner membrane. J Bacteriol **184**:1417–22.

149. **Tanaka K**, **Matsuyama SI**, **Tokuda H**. 2001. Deletion of lolB, encoding an outer membrane lipoprotein, is lethal for *Escherichia coli* and causes accumulation of lipoprotein localization intermediates in the periplasm. J Bacteriol **183**:6538–42.

150. **Wilson MM**, **Bernstein HD**. 2016. Surface-Exposed Lipoproteins: An Emerging Secretion Phenomenon in Gram-Negative Bacteria. Trends Microbiol **24**:198-208.

151. **Hooda Y**, **Moraes TF**. 2018. Translocation of lipoproteins to the surface of gram negative bacteria. Curr Opin Struct Biol **51**:73-79.

152. **Asmar AT**, **Collet JF**. 2018. Lpp, the Braun lipoprotein, turns 50—major achievements and remaining issues. FEMS Microbiol Lett **365**.

153. **Inouye M**, **Shaw J**, **Shen C**. 1972. The assembly of a structural lipoprotein in the envelope of *Escherichia coli*. J Biol Chem **247**:8154–8159.

154. **Dowdell AS**, **Murphy MD**, **Azodi C**, **Swanson SK**, **Florens L**, **Chen S**, **Zückert WR**. 2017. Comprehensive spatial analysis of the Borrelia burgdorferi lipoproteome reveals a compartmentalization bias toward the bacterial surface. J Bacteriol **199**:e00658-16.

155. **Judd A**, **Lai CC-L**, **Gray-Owen SD**, **Shin HE**, **Hooda Y**, **Moraes TF**, **Buckwalter CM**. 2016. Slam is an outer membrane protein that is required for the surface display of lipidated virulence factors in *Neisseria*. Nat Microbiol **1**:16009.

156. **Hooda Y**, **Lai CCL**, **Moraes TF**. 2017. Identification of a Large Family of Slam-Dependent Surface Lipoproteins in Gram-Negative Bacteria. Front Cell Infect Microbiol **7**:207.

157. **Lauber F**, **Cornelis GR**, **Renzi F**. 2016. Identification of a New Lipoprotein Export Signal in Gram-Negative Bacteria. MBio **7**:e01232-16.

158. **D'Enfert C**, **Chapon C**, **Pugsley AP**. 1987. Export and secretion of the lipoprotein Pullulanase by *Klebsiella pneumoniae*. Mol Microbiol **1**:107–116.

159. **Cho SH**, **Szewczyk J**, **Pesavento C**, **Zietek M**, **Banzhaf M**, **Roszczenko P**, **Asmar A**, **Laloux G**, **Hov AK**, **Leverrier P**, **Van Der Henst C**, **Vertommen D**, **Typas A**, **Collet JF**. 2014. Detecting envelope stress by monitoring β-barrel assembly. Cell **159**:1652–1664.

160. **Tschumi A**, **Nai C**, **Auchli Y**, **Hunziker P**, **Gehrig P**, **Keller P**, **Grau T**, **Sander P**. 2009. Identification of apolipoprotein *N*-acyltransferase (Lnt) in mycobacteria. J Biol Chem **284**:27146–27156.

161. **Nakayama H**, **Kurokawa K**, **Lee BL**. 2012. Lipoproteins in bacteria: structures and biosynthetic pathways. FEBS J **279**:4247–4268.

162. **Brülle JK**, **Tschumi A**, **Sander P**. 2013. Lipoproteins of slow-growing Mycobacteria carry three fatty acids and are N-acylated by apolipoprotein *N*-acyltransferase BCG_2070c. BMC Microbiol **13**.

163. **Mohiman N**, **Argentini M**, **Batt SM**, **Cornu D**, **Masi M**, **Eggeling L**, **Besra G**, **Bayan N**. 2012. The ppm Operon Is Essential for Acylation and Glycosylation of Lipoproteins in *Corynebacterium glutamicum*. PLoS One **7**:e46225.

164. **Baulard AR**, **Gurcha SS**, **Engohang-Ndong J**, **Gouffi K**, **Locht C**, **Besra GS**. 2003. In vivo interaction between the Polyprenol phosphate mannose synthase Ppm1 and the integral membrane protein Ppm2 from *Mycobacterium smegmatis* revealed by a bacterial two-hybrid system. J Biol Chem **278**:2242–2248.

165. **Kurokawa K**, **Kim M-S**, **Ichikawa R**, **Ryu K-H**, **Dohmae N**, **Nakayama H**, **Lee BL**. 2012. Environment-mediated accumulation of diacyl lipoproteins over their triacyl counterparts in *Staphylococcus aureus*. J Bacteriol **194**:3299–3306.

166. **Jan G**, **Fontenelle C**, **Le Hénaff M**, **Wróblewski H**. 1995. Acylation and immunological properties of *Mycoplasma gallisepticum* membrane proteins. Res Microbiol **146**:739–750.

167. **Kurokawa K**, **Ryu K-H**, **Ichikawa R**, **Masuda A**, **Kim M-S**, **Lee H**, **Chae J-H**, **Shimizu T**, **Saitoh T**, **Kuwano K**, **Akira S**, **Dohmae N**, **Nakayama H**, **Lee BL**. 2012. Novel bacterial lipoprotein structures conserved in low-GC content Gram-positive bacteria are recognized by Toll-like receptor 2. J Biol Chem **287**:13170–13181.

168. **Serebryakova M V**, **Demina IA**, **Galyamina MA**, **Kondratov IG**, **Ladygina VG**, **Govorun VM**. 2011. The acylation state of surface lipoproteins of Mollicute *Acholeplasma laidlawii*. J Biol Chem **286**:22769–22776.

169. **Le Hénaff M**, **Crémet JY**, **Fontenelle C**. 2002. Purification and characterization of the major lipoprotein (P28) of *Spiroplasma apis*. Protein Expr Purif **24**:489–496.

170. **Le Hénaff M**, **Fontenelle C**. 2000. Chemical analysis of processing of spiralin, the major lipoprotein of *Spiroplasma melliferum*. Arch Microbiol **173**:339–345.

171. **Shibata K -i.**, **Hasebe A**, **Into T**, **Yamada M**, **Watanabe T**. 2014. The N-Terminal Lipopeptide of a 44-kDa Membrane-Bound Lipoprotein of *Mycoplasma salivarium* Is Responsible for the Expression of Intercellular Adhesion Molecule-1 on the Cell Surface of Normal Human Gingival Fibroblasts. J Immunol **165**:6538–6544.

172. **Le Hénaff M**, **Guéguen MM**, **Fontenelle C**. 2000. Selective acylation of plasma membrane proteins of *Mycoplasma agalactiae*: The causal agent of agalactia. Curr Microbiol **40**:23–28.

173. **Mühlradt PF**, **Kiess M**, **Meyer H**, **Süssmuth R**, **Jung G**. 1998. Structure and specific activity of macrophage-stimulating lipopeptides from *Mycoplasma hyorhinis*. Infect Immun **66**:4804–4810.

174. **Jan G**, **Fontenelle C**, **Verrier F**, **Le Henaff M**, **Wroblewski H**. 1996. Selective acylation of plasma membrane proteins of *Mycoplasma mycoides* subsp. mycoides SC, the contagious bovine pleuropneumonia agent. Curr Microbiol **32**:38–42.

175. **Brokx SJ**, **Ellison M**, **Locke T**, **Bottorff D**, **Frost L**, **Weiner JH**. 2004. Genome-Wide Analysis of Lipoprotein Expression in Escherichia coli MG1655. J Bacteriol **186**:3254–3258.

176. **Glenwright AJ**, **Pothula KR**, **Bhamidimarri SP**, **Chorev DS**, **Baslé A**, **Firbank SJ**, **Zheng H**, **Robinson C V.**, **Winterhalter M**, **Kleinekathöfer U**, **Bolam DN**, **Van Den Berg B**. 2017. Structural basis for nutrient acquisition by dominant members of the human gut microbiota. Nature **541**:407–411.

177. **Shimada I**, **Okuda S**, **Sakakura M**, **Tokuda H**, **Nakada S**, **Takahashi H**. 2009. Structural Investigation of the Interaction between LolA and LolB Using NMR. J Biol Chem **284**:24634–24643.

178. **Matsuyama SI**, **Yokota N**, **Tokuda H**. 1997. A novel outer membrane lipoprotein, LolB (HemM), involved in the LolA (p20)-dependent localization of lipoproteins to the outer membrane of *Escherichia coli*. EMBO J **16**:6947–6955.

179. **Shu W**, **Liu J**, **Ji H**, **Lu M**. 2000. Core structure of the outer membrane lipoprotein from *Escherichia coli* at 1.9 Å resolution. J Mol Biol **299**:1101–1112.

180. **Edman P**. 1950. Method for Determination of the Amino Acid Sequence in Peptides. Acta Chem Scand **4**:283–293.

181. **Feng SH**, **Lo SC**. 1999. Lipid extract of *Mycoplasma penetrans* proteinase K-digested lipid-associated membrane proteins rapidly activates NF-kappaB and activator protein 1. Infect Immun **67**:2951–6.

182. **Feng SH**, **Lo SC**. 1994. Induced mouse spleen B-cell proliferation and secretion of immunoglobulin by lipid-associated membrane proteins of *Mycoplasma fermentans* incognitus and *Mycoplasma penetrans*. Infect Immun **62**:3916–3921.

183. **Ujihara T**, **Sakurai I**, **Mizusawa N**, **Wada H**. 2008. A method for analyzing lipid-modified proteins with mass spectrometry. Anal Biochem **374**:429–431.

184. **Alberts B**, **Johnson A**, **Lewis J**, **Raff M**, **Roberts K**, **Walter P**. 2002. Innate Immunity. Mol Biol Cell **4th Edition.**

185. **Jiménez-Dalmaroni MJ**, **Gerswhin ME**, **Adamopoulos IE**. 2016. The critical role of toll-like receptors - From microbial recognition to autoimmunity: A comprehensive review. Autoimmun Rev **15**:1-8.

186. **Beutler BA**. 2009. TLRs and innate immunity. Blood **113**:1399-1407.

187. **Oliveira-Nascimento L**, **Massari P**, **Wetzler LM**. 2012. The role of TLR2 ininfection and immunity. Front Immunol **3**:79.

188. **Jin MS**, **Kim SE**, **Heo JY**, **Lee ME**, **Kim HM**, **Paik S-G**, **Lee H**, **Lee J-O**. 2007. Crystal Structure of the TLR1-TLR2 Heterodimer Induced by Binding of a Tri-Acylated Lipopeptide. Cell **130**:1071–1082.

189. **Kang JY**, **Nan X**, **Jin MS**, **Youn S-J**, **Ryu YH**, **Mah S**, **Han SH**, **Lee H**, **Paik S-G**, **Lee J-O**. 2009. Recognition of Lipopeptide Patterns by Toll-like Receptor 2-Toll-like Receptor 6 Heterodimer. Immunity **31**:873–884.

190. **Manavalan B**, **Basith S**, **Choi S**. 2011. Similar structures but different roles-an updated perspective on TLR structures. Front Physiol **2**:1–13.

Chapter 2

Identification of the Lyso-Form *N*-Acyl Intramolecular Transferase in Low-GC Firmicutes

Adapted from:

Armbruster KM, Meredith TC. 2017. Identification of the Lyso-Form *N*-Acyl Intramolecular Transferase in Low-GC Firmicutes. *J. Bacteriol.* 199(11): e00099-17.

Abstract

Bacterial lipoproteins are embedded in the cell membrane of both Gram-positive and Gram-negative bacteria where they serve numerous functions central to cell envelope physiology. Lipoproteins are tethered to the membrane by an *N*-acyl-*S*-(mono/di)-acyl-glyceryl-cysteine anchor that is variably acylated depending on genera. In several low-GC Gram-positive Firmicutes, a monoacyl-glyceryl-cysteine with an *N*-terminal fatty acid (known as the lyso form) has been reported, though how it is formed is unknown. Here, through an intergenic complementation rescue assay in *Escherichia coli*, we report the identification of a common orthologous transmembrane protein in both *Enterococcus faecalis* and *Bacillus cereus* capable of forming lyso-form lipoproteins. When deleted from the native host, lipoproteins remain diacylated with a free *N*-terminus, as maturation to the *N*-acylated lyso-form is abolished. Evidence is presented suggesting that the previously unknown gene product functions through a novel intramolecular transacylation mechanism, transferring a fatty acid from the diacylglycerol moiety to the α-amino group of the lipidated cysteine. As such, the discovered gene has been named lipoprotein intramolecular transacylase (*lit*) to differentiate from the intermolecular *N*-acyltransferase (*lnt*) involved in triacyl lipoprotein biosynthesis in Gram-negative organisms.

Importance: This study identifies a new enzyme, conserved among low-GC, Gram positive bacteria, that is involved in bacterial lipoprotein biosynthesis and synthesizes lyso-form lipoproteins. Its discovery is an essential first step in determining the physiological role of N-terminal lipoprotein acylation in Gram-positive bacteria and how these modifications impact bacterial cell envelope function.

Introduction

Lipoproteins are bacterial cell membrane components common throughout Gram-negative and Gram-positive bacteria, constituting 2-5% of all cellular proteins (1–5). Lipoproteins invariably contain a signature acylated *N*-terminal cysteine residue that tethers an otherwise soluble globular protein domain to the membrane surface. Positioned at the membrane-environment interface, lipoproteins play critical roles in ion and nutrient capture, solute transport, cell adhesion, assembly of protein complexes, and as protein-folding chaperones.

All lipoproteins are initially translated as precursors containing an *N*-terminal signal peptide harboring a conserved amino acid lipobox motif sequence preceding an invariant cysteine residue (6, 7). Once transported to the outer face of the membrane, prolipoprotein diacylglycerol transferase (Lgt) attaches a diacylglycerol residue from a phospholipid donor through a thioether bond (8). The signal peptidase II (Lsp) then cleaves the leader peptide immediately upstream of the cysteine to liberate the α-amino group (9). In *Escherichia coli* and other Gram-negative bacteria, biosynthesis is completed by apolipoprotein *N*-acyltransferase (Lnt) which transfers a fatty acid from a membrane phospholipid to form an amide linkage (10). While both Lgt and Lsp are highly conserved across genera, *lnt* sequence orthologs are confined to Gram-negative bacteria and mycobacteria, which also make triacylated lipoproteins (5, 11). However, lipoprotein *N*-terminal modification in other Gram-positive bacteria still occurs, and in low-GC Firmicutes in particular, is more complex and varied than previously appreciated (2). In addition to the conventional unmodified *N*-terminal and *N*-acylated forms, Kurokawa et al. discovered three novel lipoprotein forms: the peptidyl form, the *N*-acetyl form, and the lyso form (12). Lyso-form lipoproteins contain an *N*-acyl chain but differ in having a monoacyl-glyceryl group (akin to lysophospholipids) attached to cysteine. Lyso-form lipoproteins are found in several Gram-positive Firmicutes, including *Enterococcus faecalis* and *Bacillus cereus* (12). It is currently

unclear how the mature lyso-form lipoprotein is made, but likely occurs in *E. faecalis* after canonical processing by Lgt and Lsp (13).

To make the final *N*-acyl-*S*-monoacyl-glyceryl-cysteine structure characteristic of lyso-form lipoproteins, two theories have been proposed: *O*-deacylation of the conventional triacyl form in tandem with *N*-terminal acylation via an Lnt-like reaction, or intramolecular transacylation from the *S*-diacylglyceryl group to the α-amino group of cysteine (12). However, there are no apparent *lnt* sequence orthologs in Firmicute genomes, nor are there any known examples of enzyme-catalyzed lipoprotein intramolecular transacylation (5). The general function of lipoprotein aminoacylation among Firmicutes also remains to be seen. In Gram-negative organisms, Lnt-catalyzed lipoprotein *N*-acylation is necessary for efficient localization of outer membrane (OM) lipoproteins and ensuing cell division (14–16). This cellular role for *N*-acylation is not applicable to lyso-form lipoproteins, which are found in cell envelopes lacking an OM. Therefore, identification of the enzyme that *N*-terminally acylates lyso-form lipoproteins is a necessary first step not just in determining how these lipoproteins are made, but in elucidating the physiological role of *N*-acylation in Gram-positive bacteria as well. Using cross-complementation of an Lnt depletion strain in *E. coli* as a selection strategy, coupled with mass spectrometry (MS) analyses of lipoproteins, we herein report the identification of a single gene capable of forming lyso-form lipoproteins in two low-GC Firmicute hosts: *E. faecalis* and *B. cereus*.

Results

Intergenic lipoprotein *N*-acylation cross-complementation screen design

In contrast to the essential role of Lnt in *E. coli*, lipoproteins are non-essential in many Firmicutes (3, 5) and there are no known phenotypes attributable to the *N*-terminal acylation state

that could be used for direct selection in native hosts. However, we hypothesized that lyso-form lipoproteins would be compatible substrates for the *E. coli* localization of lipoproteins (Lol) export machinery since triacyl and lyso-form lipoproteins both share an α-amino linked acyl chain (4). Fukuda et al. have extensively characterized the substrate specificity of Lol transport and have shown *N*-acylation to be a critical component in substrate recognition (16). Following acylation, the LolCDE complex extracts lipoproteins from the inner membrane (IM) and transfers them to LolA (Figure 2-1A), which shuttles lipoproteins across the periplasm (30). Lipoproteins are then transferred to LolB, itself a lipoprotein, which subsequently inserts the acylated *N*-terminal anchor of OM-destined lipoproteins into the membrane bilayer (31). This data, when coupled with previous studies showing the cross recognition of *E. coli* lipoprotein substrates by Firmicute lipoprotein biosynthetic proteins when heterologously expressed (32, 33), suggested an intergenic complementation screen to find the gene(s) responsible for lyso-form lipoprotein *N*-acylation would be feasible. To address potentially incompatible acyl chain substrate specificities between the putative Firmicute lyso-form *N*-acyl lipoprotein transferase and the acyl chain composition available in *E. coli*, we built two separate genomic DNA libraries of small inserts (3-5 kb) from both *B. cereus* and *E. faecalis*. While *B. cereus* elaborates mainly saturated branched-chain fatty acids, the complement of fatty acids in *E. faecalis* more closely matches *E. coli* in being even-numbered, linear acyl chains and containing unsaturation units (34). High quality libraries were built in the multi-copy number plasmid pUC19 using genomic DNA randomly sheered by NEBNext dsDNA Fragmentase, allowing for at least 3-fold genomic coverage per transformation reaction, with more than 90% of library plasmids harboring Firmicute genomic DNA insert.

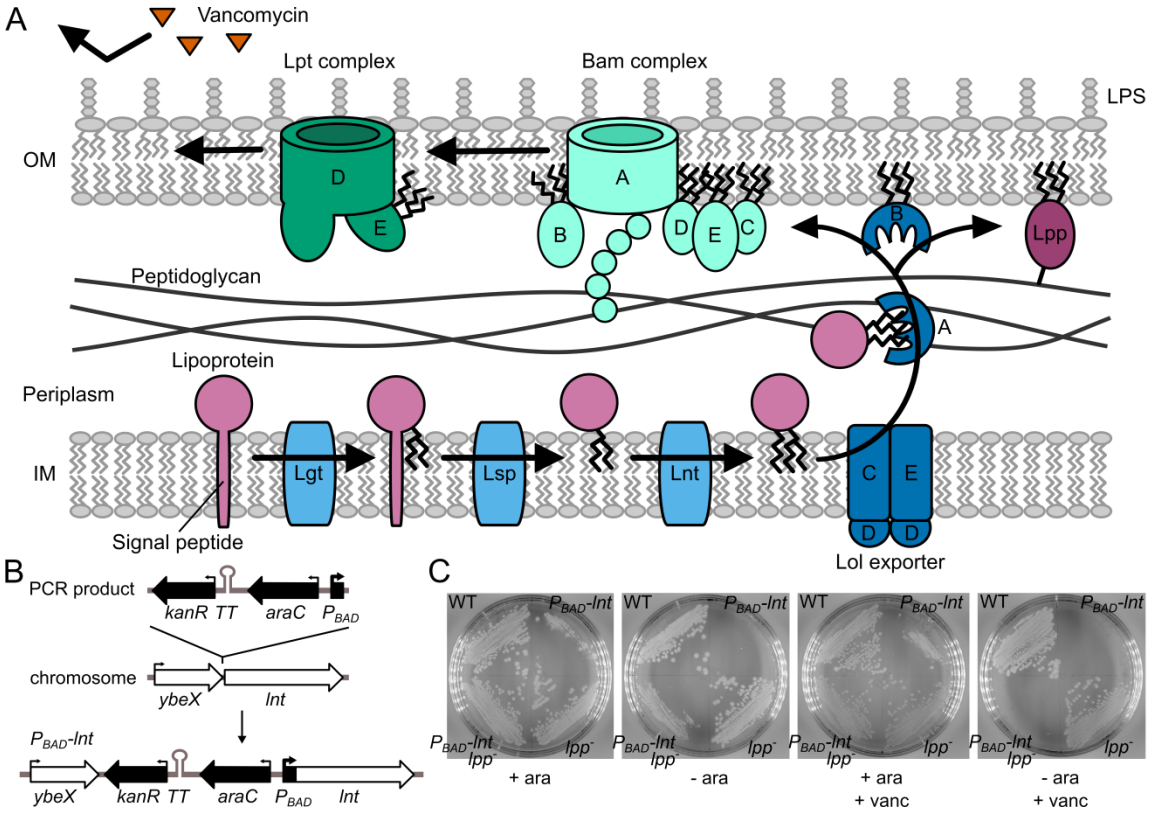


Figure **2-1**: Lipoprotein maturation in *E. coli* and complementation strategy. (A) Lipoproteins are sequentially modified by Lgt, Lsp, and Lnt. Once triacylated, lipoproteins are shuttled across the periplasm where they are inserted into the OM by the Lol export machinery. The Bam complex, catalyzing β-barrel protein assembly and insertion into the membrane, folds LptD in association with LptE, in turn ensuring proper LPS assembly at the outer leaflet of the membrane. The LPS leaflet excludes large hydrophillic molecules, such as vancomycin, from entering the cell. The highly abundant lipoprotein Lpp forms covalent linkages to peptidoglycan through residue K58. (B) The *kanR-TT-araC-*P_{BAD} cassette was constructed and inserted directly upstream of *lnt* by Red-mediated recombination to generate L-arabinose-dependent expression in *E. coli*. The cassette contained a transcriptional terminator to prevent read-through transcription. (C) Wildtype *E. coli* BW25113 (WT), P_{BAD}-*lnt* (KA325), *lpp*$^-$ (TXM327), and P_{BAD}-*lnt lpp*$^-$ (KA349) strains were streaked onto LB agar plates with or without L-arabinose and vancomycin. Plates lacking arabinose contained glucose to suppress basal P_{BAD} expression, while all plates were supplemented with palmitate as a potential source for fatty acids.

In order to identify Firmicute candidate proteins able to *N*-terminally acylate lipoproteins and thus restore OM transport in *E. coli*, a conditional-lethal mutant of *lnt* was constructed in *E. coli*. The *kanR-TT-araC-*P_{BAD} cassette was integrated directly upstream of the chromosomal *lnt* gene (Figure 2-1B), and the resulting strain KA325 was indeed dependent on L-arabinose for

growth (Figure 2-1C). Phenotypic rescue by growth in the absence of *lnt* inducer L-arabinose would thus nominate candidate Firmicute lipoprotein *N*-acyltransferase genes.

Vancomycin reduces background growth in *lpp* null strains

Despite the high library coverage (see Materials and Methods), initial attempts to identify rescuing DNA fragments from either *B. cereus* or *E. faecalis* were unsuccessful (data not shown). Presumably, this was due to incomplete rescue through low expression, protein misfolding, and/or low affinity for lipoprotein substrates at either the *N*-acylation or transport stage. Transport is critical to cell viability, as the abundant OM lipoprotein Lpp (Braun's lipoprotein) will continue to be covalently crosslinked to peptidoglycan from the IM, ultimately leading to cell lysis (35). With upwards of one million Lpp molecules per cell (36), robust transit to the OM must be maintained so as to prevent Lpp accumulation in the IM and the ensuing Lpp-peptidoglycan adducts from forming. To reduce the bar for rescue, we deleted Lpp, which is not essential (37), from our assay strain background in order to facilitate phenotypic rescue. However, growth of this strain (KA349) was observed even in the absence of *lnt* inducer L-arabinose (Figure 2-1C). This growth may be due to residual Lnt activity or slow but sufficient transit of diacylated lipoproteins to the OM by the Lol export machinery. Regardless, it was necessary to reduce this background growth in order to perform the screen. We reasoned other lipoprotein-dependent cell processes may be compromised in an *lnt*-depleted state, such as the OM β-barrel assembly machinery (Bam) and lipopolysaccharide transport (Lpt) complexes. The Bam complex, composed of an OM β-barrel and several lipoproteins, folds OM proteins (OMPs) and integrates them into the membrane (4). LptD, an OM β-barrel folded by the Bam complex, can only be folded when associated with its lipoprotein partner LptE (38). Together, these proteins assemble the OM lipopolysaccharide (LPS) layer in the outer leaflet. Since LPS forms a highly selective

permeability barrier that protects the cell from harmful agents (39), a decrease in OM-located Bam, Lpt, and LPS due to retention of lipoprotein components in the IM would enhance susceptibility to normally-excluded molecules, such as the large glycopeptide antibiotic vancomycin (Figure 2-1A). Indeed, low levels of vancomycin completely abrogated growth of the *lpp*$^-$ strain KA349 in the absence of L-arabinose-induced *lnt* (Figure 2-1C).

E. faecalis WMC_RS08810 and *B. cereus* BC1526 complement Lnt-depleted *E. coli*

With selection assay conditions established, *E. faecalis* and *B. cereus* genomic DNA was introduced into KA349 by electroporation and plated on LB agar containing: carbenicillin, to select for successful plasmid transformants; glucose, to repress gratuitous expression from the P_{BAD}-*lnt* construct; vancomycin, to reduce background growth as described above; and palmitate, to provide a potential acyl chain substrate. Plasmids from colonies showing enhanced growth on selection plates were sequenced. Of sixteen chosen colonies transformed with *E. faecalis* genomic DNA, twelve contained overlapping inserts that shared a single intact, unannotated open reading frame (WMC_RS08810; Figure 2-2A). Similarly, seven of eight chosen colonies transformed with *B. cereus* genomic DNA contained an overlapping insert encoding the open reading frame BC1526 (Figure 2-2B). The Basic Local Alignment Search Tool (BLAST) search algorithm identified the two genes as orthologs sharing 34% amino acid identity. Both genes are predicted to be membrane proteins containing a domain of unknown function (DUF1461). One plasmid encoding the candidate *lnt* rescue gene of interest from each strain, designated pEF4 and pBC7, was chosen for further analysis.

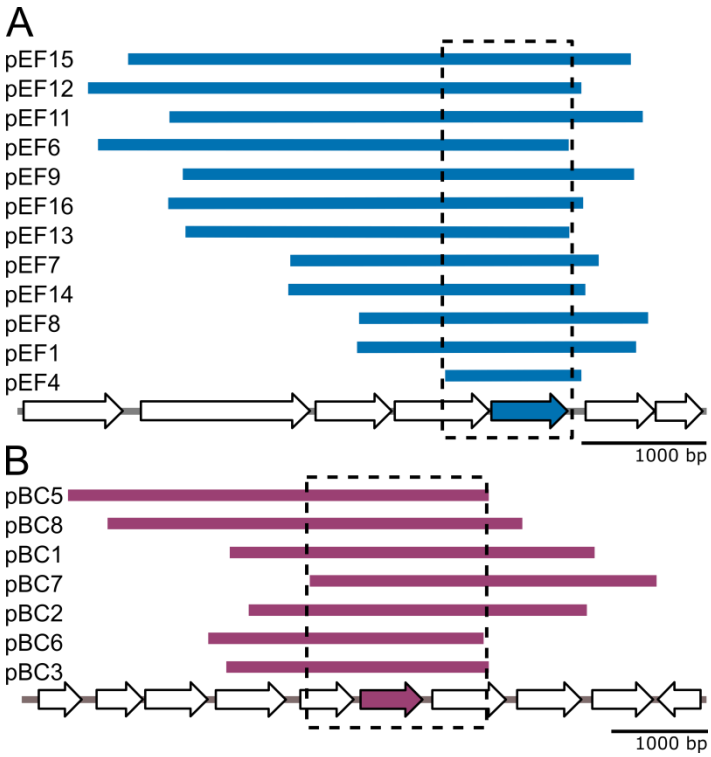


Figure **2-2**: Genetic maps of DNA fragments conveying a viable phenotype in Lnt-depleted *E. coli* cells. DNA fragments that rescued growth in the Lnt-depleted *E. coli* mutant strain KA349 (P_{BAD}-*lnt*, *lpp*$^-$) were mapped to the *E. faecalis* genome (A) and the *B. cereus* genome (B). Fragments are arranged in size top-down from largest to smallest. The overlapping inserts are boxed in each panel, encoding a single unannotated open reading frame indicated by the filled-in arrow. These genes are *E. faecalis* WMC_RS08810 and *B. cereus* BC1526. Plasmids pEF4 and pBC7 were used in subsequent studies.

E. faecalis WMC_RS08810 and *B. cereus* BC1526 can fully substitute for *lnt* in *E. coli* only when *lpp* is deleted

To confirm rescue, growth of P_{BAD}-*lnt* conditional *lpp*$^-$ strains harboring plasmids (pUC19, pEF4, and pBC7) was monitored with and without L-arabinose (Figure 2-3A). While the control strain was dependent on L-arabinose induction of *lnt*, growth was restored in strains harboring either pEF4 or pBC7 in the absence of L-arabinose, suggesting that pEF4 and pBC7 rescue the lethal phenotype caused by Lnt depletion. To determine if *E. faecalis* WMC_RS08810 and *B. cereus* BC1526 could fully substitute for *lnt* in *E. coli*, we performed co-transduction

linkage analysis (Figure 2-3B). The donor strain TXM541 was constructed with two selectable alleles, *chiQ::*Apr^r and *lnt::*Spt^r, placed approximately ~19-kb apart on the chromosome, along with a functional copy of *lnt* reinserted in a distant chromosomal position. A P1*vir* lysate was made using TXM541 and transduced into a panel of recipient strains with or without an intact *lpp* gene and harboring potential rescue plasmids (pUC19, pUC19-*lnt*, pEF4, and pBC7). Transductants were first plated on apramycin to select for colonies that successfully received the donor DNA fragment of interest, then subsequently patched onto spectinomycin to assess for the *lnt::*Spt^r allele. Colonies with co-resistance to both apramycin and spectinomycin would indicate that *lnt* can be functionally replaced in the recipient strain as it is no longer essential. The observed positive control (pUC19-*lnt*) co-transduction efficiency using P1*vir* agreed well with the predicted value of 50% (40), while the *lnt::*Spt^r allele could not be established in the pUC19 empty vector control (Figure 2-3B). Co-transduction comparable to pUC19-*lnt* was observed in recipient strains harboring pEF4 and pBC7 and lacking a functional *lpp* gene. However, co-transduction was not observed in the same recipient strains with an intact *lpp* gene. This indicates that the Firmicute candidate genes cannot rescue cell viability in an *lnt* null background when *lpp* is present.

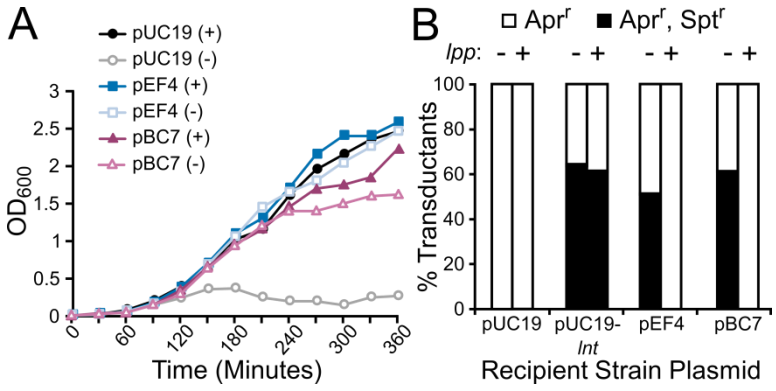


Figure **2-3**: Complementation of *lnt* depletion by pEF4 and pBC7 in *E. coli*.(A) The growth of *lnt*-inducible strains harboring pUC19 (●), pEF4 (■), or pBC7 (▲) in an *lpp* null background grown with (closed symbols) or without (open symbols) L-arabinose was measured in LB broth.

(B) The *lnt::*Spt^r and *chiQ::*Apr^r linked markers were packaged into P1*vir* and transduced into recipient strains harboring pUC19, pUC19-*lnt*, pEF4, or pBC7 in an *lpp* null or wildtype *lpp* background. The percentage of transductant colonies resistant to apramycin only (Apr^r; open bars) or co-resistant to both apramycin and spectinomycin (Apr^r, Spt^r; shaded bars) are shown (*n* = 21 to 60 colonies).

OM lipoprotein localization is stimulated by *E. faecalis* WMC_RS08810

We hypothesized that the observed phenotypic rescue was due to restoration of Lol-mediated lipoprotein transport to the OM through α-aminoacylation. To determine the distribution of lipoproteins in the cell envelope, the OM and IM fractions were separated by centrifugation through a discontinuous sucrose gradient. Two distinct samples of strain KA528, featuring the P_{BAD}-*lnt* construct, were harvested: one grown with L-arabinose and one grown without, the latter to serve as the Lnt-depleted sample. The third sample was harvested from KA532 (Δ*lnt* + pEF4), previously constructed by P1*vir* transduction of the *lnt::*Spt^r allele, so as to monitor WMC_RS08810 activity in the complete absence of Lnt. Chromosomal *lpp* was deleted in both strains and back-complemented with a plasmid encoding a *C*-terminally Strep-tagged LppK58A allele. Mutation of K58 prevents covalent crosslinking of Lpp to peptidoglycan (35, 41).

Following sample centrifugation and fractionation, the protein abundance in each fraction was measured by the Bradford reagent (Figure 2-4A). Fractions containing protein corresponding to each membrane were pooled and their volumes normalized. The protein profile of the Lnt-depleted sample notably differed from the profile of the sample from KA528 grown with L-arabinose, in that the density of the IM increased while that of the OM decreased. The profile of the *E. faecalis*-complemented strain KA532 mirrored the Lnt-induced sample. Coomassie blue staining of the PAGE-separated pooled membrane fractions showed enrichment of porins in the OM fraction for each of the three samples (Figure 2-4B). To further assess membrane separation,

the activity of the IM-associated enzyme NADH oxidase was measured (Figure 2-4C). While the IM from KA532 and KA528 grown with L-arabinose contained the majority of the total NADH oxidase activity, near-equivalent NADH oxidase activity was observed between the OM and IM from the Lnt-depleted sample (Figure 2-4C). The blending of OM and IM markers has previously been observed upon Lnt depletion (14, 35). Collectively, the data indicates WMC_RS08810 can at least partially restore OM biogenesis in the absence of the *lnt N*-acyltransferase.

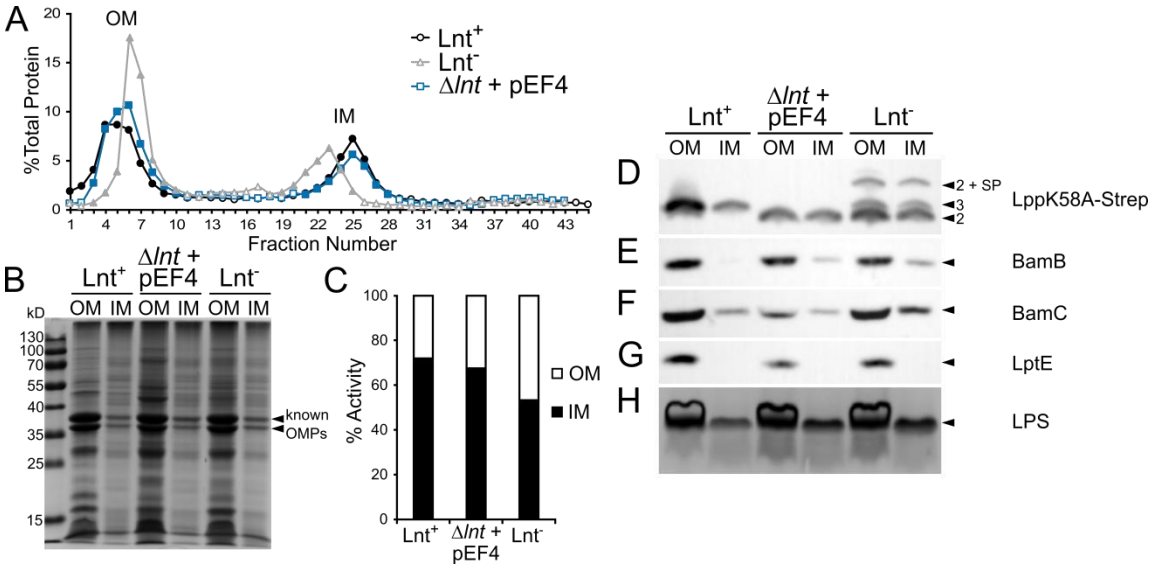


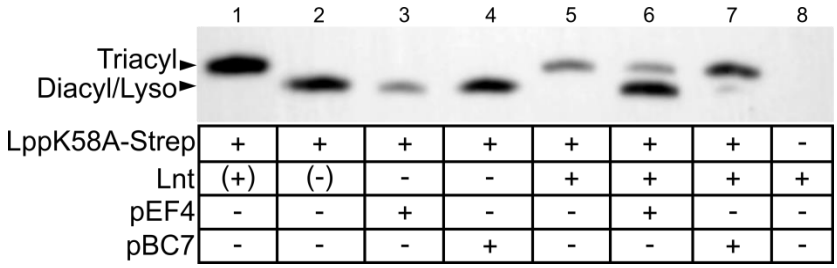
Figure **2-4**: Membrane localization of lipoproteins in KA528 and KA532. The total membrane fraction from KA528 (P_{BAD}-*lnt*) grown in LB with arabinose (Lnt$^+$, ●), or without (Lnt$^-$, ▲), and from KA532 (Δ*lnt* + pEF4, ■) were separated by discontinuous sucrose gradient ultracentrifugation. (A) The protein content of the collected fractions was determined by Bradford assay. OM and IM fractions containing protein were pooled (closed symbols) and their final volumes normalized. (B) Protein profiles of the OM and IM samples were analyzed by SDS-PAGE and Coomassie blue staining. (C) The NADH oxidase activity of the OM and IM samples were measured by following consumption of NADH at 340 nm. Samples were immunoblotted with antibodies against LppK58A-Strep (D), BamB (E), BamC (F), and LptE (G) and detected by HRP chemiluminescence. Lipoproteins having three acyl chains (3), two acyl chains (2), and the signal peptide (SP) are indicated. (H) The LPS profiles of whole cell samples were detected by SDS-PAGE separation and visualized by silver staining.

To further assess the extent of complementation, fractions were analyzed by immunoblotting for several key lipoproteins involved in OM biogenesis (see Figure 2-1A). Immunoblots against LppK58A, BamB, BamC, and LptE revealed lipoprotein enrichment in the OM for all three samples (Figures 2-4D-G). The most telling results were seen for Strep-tagged LppK58A (Figure 2-4D), for which distinct bands corresponding to lipoproteins with two and three acyl chains could be observed. The Lnt-depleted sample contained a mixture of triacylated and diacylated LppK58A in both the OM and IM fractions. Unproteolyzed Lpp with an intact signal peptide was also detected, including in the OM fraction, which is in agreement with the previously observed blending of OM and IM (14, 35). Lpp precursor accumulation in the IM due to the absence of Lnt-catalyzed *N*-acylation is consistent with slowed Lol transport of diacylated substrates. Unprocessed signal peptide, however, was not observed in the WMC_RS08810-complemented strain even though all LppK58A signal ran with an electrophoretic mobility consistent with diacylation. Localization of BamB, BamC, and LptE was not appreciably perturbed in any of the samples (Figures 2-4E-G). This may reflect the ability of Lol transport to keep pace with *de novo* diacylglycerol lipoprotein intermediates of low abundance or inherent substrate preference. Likewise, the amount and distribution of LPS remained comparable across all samples (Figure 2-4H).

E. coli **Lnt and the Firmicute candidate proteins compete for a common lipoprotein substrate**

In order to investigate the structure of the lower molecular weight bands observed in the WMC_RS08810 membrane-separated sample (Figure 2-4D), we designed a substrate competition experiment between Lnt and candidate Firmicute *N*-acyltransferases (Figure 2-5). If WMC_RS08810 synthesizes a lyso-form lipoprotein from a diacylglycerol, then Lnt should not

be able to convert this pool into mature triacylated lipoprotein since the α-amino terminus is already modified. Conversely, if WMC_RS08810 induces phenotypic rescue through a mechanism other than lipoprotein aminoacylation, then co-expression of Lnt should convert the entire lipoprotein population into the mature triacylated form with a higher molecular weight band. Similar to WMC_RS08810, LppK58A from a strain expressing BC1526 (Δ*lnt* + pBC7) runs at the same position as diacylglycerol-modified LppK58A from Lnt-depleted samples (Figure 2-5, lanes 3 and 4 versus 2). Cells expressing both Lnt and either Firmicute candidate gene show two distinct bands on the gel in comparison to Lnt alone (lanes 6 and 7 versus 5), indicative of two non-interconvertible populations of LppK58A. Differences in band intensity between these two strains (lane 6 versus 7) are likely due to enzyme-substrate compatibility, particularly acyl chain substrate. As the straight-chained, unsaturated fatty acid composition of *E. faecalis* more closely matches that of *E. coli*, WMC_RS08810 may be more efficient in *E. coli* than BC1526, which normally utilizes branched-chained fatty acids. The observed outcome supports a model whereby the two lipoprotein *N*-acylating enzymes compete for diacylglycerol-modified LppK58A substrate having a free α-amino group. This results in two structures of LppK58A product depending on which enzyme modifies the lipoprotein: Lnt forming triacylated lipoprotein and the Firmicute gene product presumably forming a diacylated lipoprotein, one of which is an *N*-acyl chain since this structure cannot be further processed by Lnt. As the total number of acyl chains does not change, *N*-acylation would have to accompany *O*-deacylation in order to complete the lyso-form. Taken together, these results indicate that lyso-form lipoproteins are synthesized by an intramolecular transacylation mechanism, in which an acyl chain from the diacylglycerol moiety is transferred to the α-amino group of cysteine.



LppK58A-Strep	+	+	+	+	+	+	+	-
Lnt	(+)	(-)	-	-	+	+	+	+
pEF4	-	-	+	-	-	+	-	-
pBC7	-	-	-	+	-	-	+	-

Figure **2-5**: Lipoprotein substrate competition assay. Extracts from *lpp* null strains harboring various combinations of lipoprotein *N*-acylating enzymes (+ or -) and LppK58A-Strep tagged substrate were separated by SDS-PAGE and immunoblotted with anti-Strep HRP conjugate. Strain KA528 was grown with (lane 1) and without (lane 2) L-arabinose, while KA532 (lane 3), KA583 (lane 4), KA548 (lane 5), KA599 (lane 6), KA600 (lane 7), and TXM327 (lane 8) were grown in LB media. Parentheses indicate *lnt* expression was controlled by adding L-arabinose to growth media.

LppK58A is the lyso form when processed by *E. faecalis* WMC_RS08810

To confirm the lyso-form structure, we purified LppK58A-Strep from *E. coli* strains KA548 (WT *lnt*) and KA532 (Δ*lnt* + pEF4) and analyzed their *N*-terminal tryptic peptides by MALDI-TOF MS. The MS spectrum of Lpp purified from strain KA548 showed a prominent ion at *m/z* 1396 (Figure 2-6A), corresponding to the mass of conventionally triacylated *N*-terminal peptide possessing acyl chains totaling 48:1 (with 48 and 1 referring to the number of total carbons and a double bond, respectively). This peak was notably missing from the MS spectrum of LppK58A purified from strain KA532 (Figure 2-6D). Rather, a peak was detected at *m/z* 1157, equivalent to an *N*-terminal lipopeptide containing two acyl chains totaling 32:1. To determine the position of the fatty acids on the lipopeptides, MS/MS analyses were conducted on the *m/z* 1396 and 1157 parent ions. Both MS/MS spectra share the *y*-series ions corresponding to the SSNAK peptide of Lpp (Figure 2-6B and E). In the MS/MS spectrum of *m/z* 1396, an ion at *m/z* 1141 corresponds to the product ion that has lost a $C_{16:1}$ fatty acid ($C_{15}H_{29}COOH$), and an ion at *m/z* 1157 corresponds to the loss of a $C_{16:0}$ fatty acid ($C_{15}H_{30}COOH$) from either the diacylglycerol or the *N*-terminus (Figure 2-6B). Two additional characteristic fragment ions at

m/z 813 and 845 correspond to the *N*-acyl($C_{16:0}$)-dehydroalanyl peptide, generated by the neutral loss of the diacylthioglycerol, and the thiolated *N*-acyl($C_{16:0}$)-peptide, respectively. Similar *N*-acylated peptide fragment ions are also evident in the MS/MS spectrum of the *m/z* 1157 ion from the WMC_RS08810-expressing strain; however these ions appear at *m/z* 811 and 843, indicative of a monounsaturated *N*-terminal fatty acid ($C_{16:1}$) (Figure 2-6E). Additional ions at *m/z* 901 and 921 correspond to product ions that have lost a $C_{16:0}$ fatty acid ($C_{15}H_{30}COOH$) and the *N*-terminal $C_{16:1}$ fatty acid ($C_{15}H_{29}COOH$). Further MS/MS fragment analysis of the *m/z* 1185 parent ion from the WMC_RS08810-expressing strain, which is 28 Da heavier than the *m/z* 1157 ion due to two additional methylene (-CH_2-) groups, also contained the signature *m/z* 811 ion supporting assignment of an *N*-acylated parent ion structure (Figure 2-7).

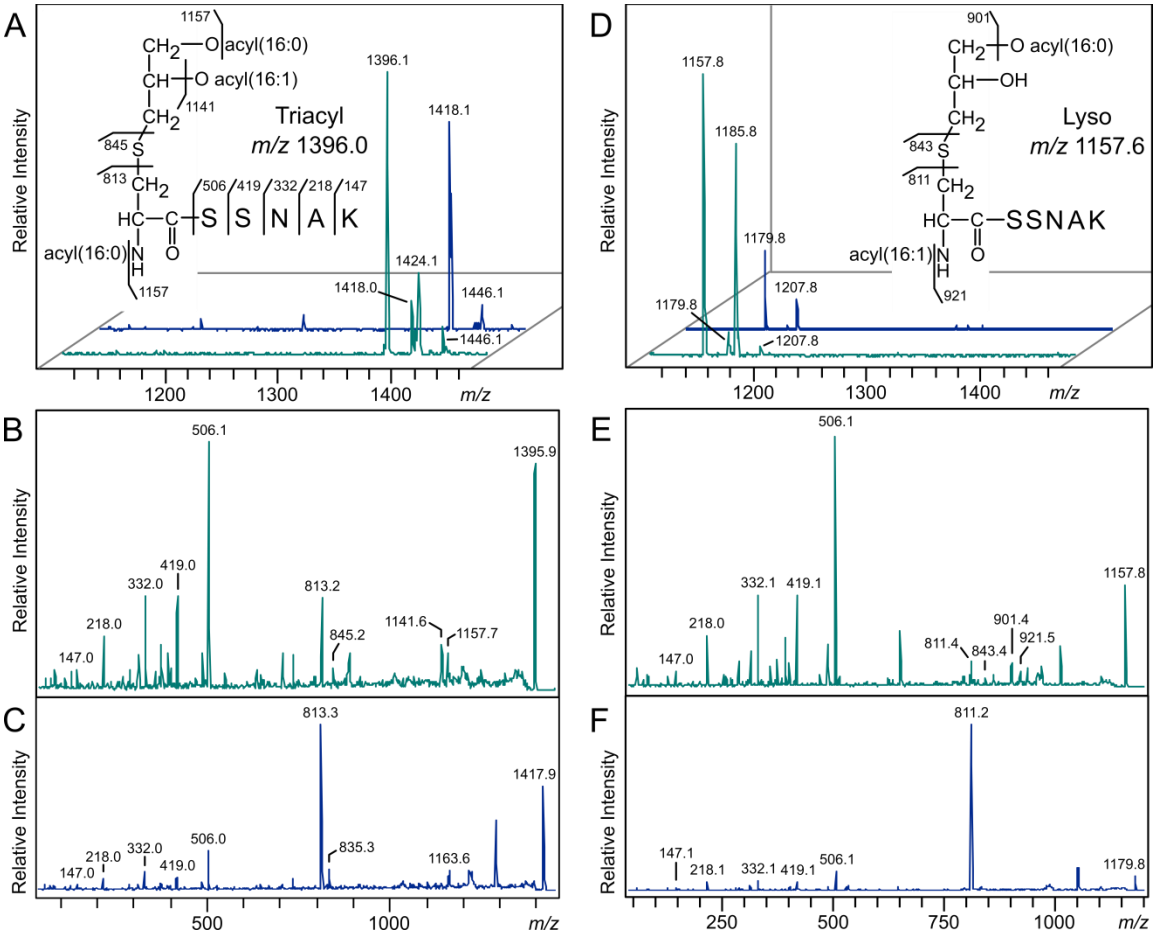


Figure **2-6**: MALDI-TOF of *E. coli* LppK58A Processed by Lnt and *E. faecalis* WMC_RS08810. Trypsinized Lpp lipopetides purified from the wildtype *lnt* strain KA548 (A) and the *lnt* null, WMC_RS08810-expressing strain KA532 (D) were eluted from nitrocellulose bands and analyzed by MALDI-TOF MS (turquoise trace). Sodium was added to a subset of samples to promote sodiated adduct formation (blue trace). The MS/MS spectra of the protonated *m/z* 1396 (B) and sodiated *m/z* 1418 (C) parent ions were used to elucidate the triacylated *N*-terminal structure of Lpp from KA548 (inset of A). Likewise, the MS/MS spectra of the ions *m/z* 1157 (E) and 1179 (F) were used to assign the lyso-form *N*-terminal peptide structure of Lpp from KA532 (inset of D). The position of the *O*-acyl chains could not be determined on either the triacyl or lyso-form Lpp molecules.

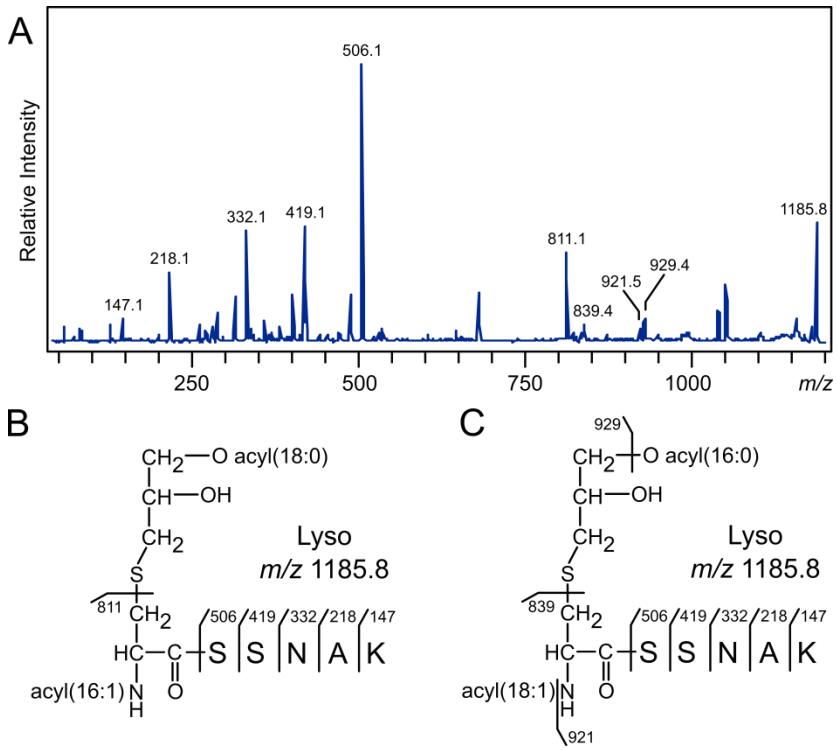


Figure **2-7**: MALDI-TOF MS/MS Spectrum of the *E.coli* Lpp ion *m/z* 1185. MS/MS spectrum of the *m/z* 1185 ion of Lpp purified from KA532 (A) and the two elucidated structures (B and C) of lyso-form Lpp featuring lipidated cysteine with two fatty acids totaling 34:1.

Both MS/MS spectra of the parent ions *m/z* 1396 and 1157 showed strong preferential fragmentation to the *m/z* 506 SSNAK peptide ion, diminishing the diagnostic *N*-acylated peptide signal intensity. For certain samples, this proved to be problematic and hampered confident structure assignments. Since sodium adducts can improve signal from triacylglycerides and

phospholipids (42), we promoted sodium adduct formation by spiking samples with sodium bicarbonate in order to coerce potentially more informative fragmentation (Figure 2-6). While overall signal intensity declined, sodium adducts were readily formed as determined by the 22 Da shift in the *m/z* 1396 and 1157 ions to *m/z* 1418 and 1179, respectively. Subsequent MS/MS fragmentation of these sodium adducts generated strong *N*-acylated dehydroalanyl peptide ion signals at *m/z* 811 and 813 (Figure 2-6C and F). Taken together, the absence of the *m/z* 1396 triacylated peptide peak in the MS spectrum and the prominent *N*-acylated peptide peaks in the MS/MS spectrum of LppK58A purified from KA532 confirm that LppK58A is converted into the lyso-form when processed by *E. faecalis* WMC_RS08810.

Deletion of WMC_RS08810 and BC1526 prevents lyso-form lipoprotein formation in *E. faecalis* and *B. cereus*

To confirm the role of *E. faecalis* WMC_RS08810 and *B. cereus* BC1526 in lyso-lipoprotein formation, the genes were deleted in their native organisms. Native lipoproteins from the wildtype, deletion, and complemented strains were enriched using the Triton X-114 phase partitioning method and separated by SDS-PAGE. The lipoprotein profile and relative abundance were unchanged when either gene was deleted and then back complemented with a plasmid-encoded copy (Figure 2-8). This was consistent with the inability of SDS-PAGE to discriminate between diacylglycerol and lyso-forms of lipoproteins (see Figure 2-5). The most abundant lipoproteins from *E. faecalis* (PnrA) and *B. cereus* (PrsA), previously characterized by Kurokawa et al. (12), were chosen for further analysis by MALDI-TOF MS/MS.

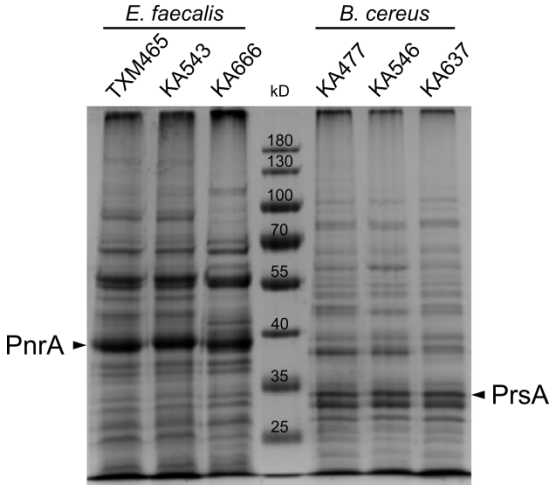


Figure **2-8**: Lipoproteins from *E. faecalis* and *B. cereus*. Using the Triton X-114 phase partitioning method, lipoproteins were enriched from *E. faecalis* strains TXM465 (WT), KA543 (ΔWMC_RS08810), and KA666 (ΔWMC_RS08810 + pMSP3535-WMC_RS08810), along with *B. cereus* strains KA477 (WT), KA546 (ΔBC1526), and KA637 (ΔBC1526 + pSP-BC1526-hp). The resulting Triton X-114 phases were analyzed by SDS-PAGE and visualized by Coomassie blue staining.

Based on the proposed intramolecular transacylation mechanism for lipoprotein modification when expressed heterologously in *E. coli*, deletion of the candidate genes should result in conventional diacylglycerol modified lipoproteins of identical mass to lyso-form. Indeed, peaks corresponding to the *N*-terminal peptide (*m/z* 997 for *E. faecalis* PnrA and *m/z* 1335 for *B. cereus* PrsA) were seen in parent spectra from both wildtype and deletion strains (Figure 2-9). MS/MS spectra contained *y*-series ions corresponding to the *N*-terminal amino acid sequence GGGK of PnrA (Figure 2-10A and B) and GTSSSDK of PrsA (Figure 2-10F and G), confirming lipoprotein identity. However, the parent ions fragmented differently depending on whether the WMC_RS08810 or BC1526 genes were present. Differences in the lipoproteins purified from the wildtype and deletion strains arise from the acylated or deacylated product ions. In the MS/MS spectrum of *E. faecalis* PnrA from wildtype cells, ions at *m/z* 715, 733, and 759 correspond to parent ions having lost an $C_{18:1}$ fatty acid ($C_{17}H_{33}COOH$), $C_{18:1}$ ketene ($C_{16}H_{31}CH{=}C{=}O$), and

$C_{16:0}$ ketene ($C_{14}H_{29}CH=C=O$), respectively. Two additional characteristic fragment ions observed at *m/z* 625 and 657 correspond to the *N*-acyl($C_{16:0}$)-dehydroalanyl peptide, resulting from the neutral loss of the monoacyl($C_{18:1}$)-thioglycerol, and the thiolated *N*-acyl($C_{16:0}$)-peptide, respectively (Figure 2-10A). These peaks are absent from the spectrum of PnrA purified from WMC_RS08810 deletion mutants (Figure 2-10B). Rather, this spectrum features ions at *m/z* 741 and 759, corresponding to the loss of an $C_{16:0}$ fatty acid ($C_{15}H_{30}COOH$) and $C_{16:0}$ ketene ($C_{14}H_{29}CH=C=O$), respectively, and an additional ion at *m/z* 459 whose mass is consistent with the loss of both fatty acids (34:1). Two characteristic peaks at *m/z* 387 and 419, corresponding to dehydroalanyl and thiolated CGGGK peptides with free α-amino groups, respectively, both support the diacylglycerol lipoprotein structural assignment.

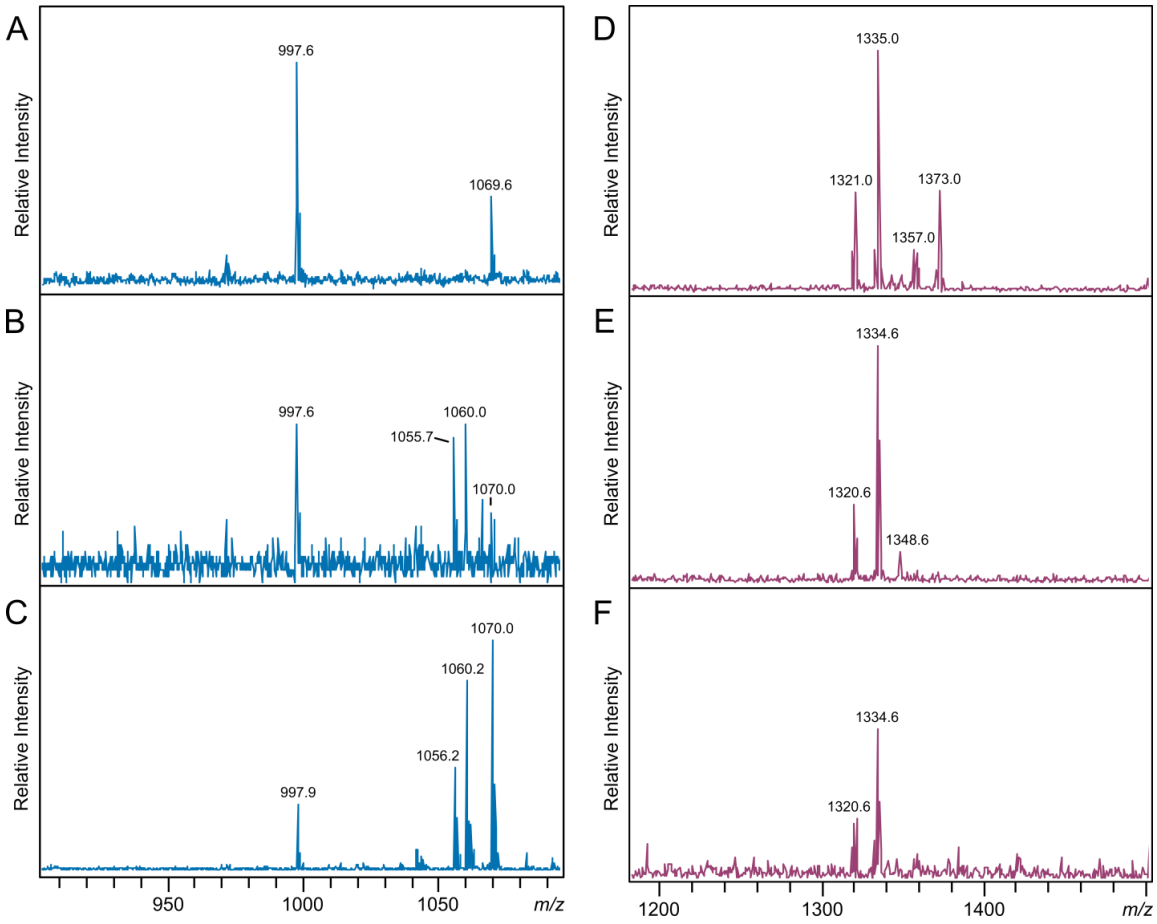


Figure **2-9**: MALDI-TOF MS of *E. faecalis* PnrA and *B. cereus* PrsA. Parent MS spectra of the *m/z* 997 ion region corresponding to the *N*-terminal lipopeptide of *E. faecalis* PnrA purified from the wildtype strain (A), the deletion strain (B), and the back-complemented strain (C). MS spectra of the *m/z* 1334 parent ion region corresponding to the *N*-terminal lipopeptide of *B. cereus* PrsA purified from the wildtype strain (D), the deletion strain (E), and the back-complemented strain (F).

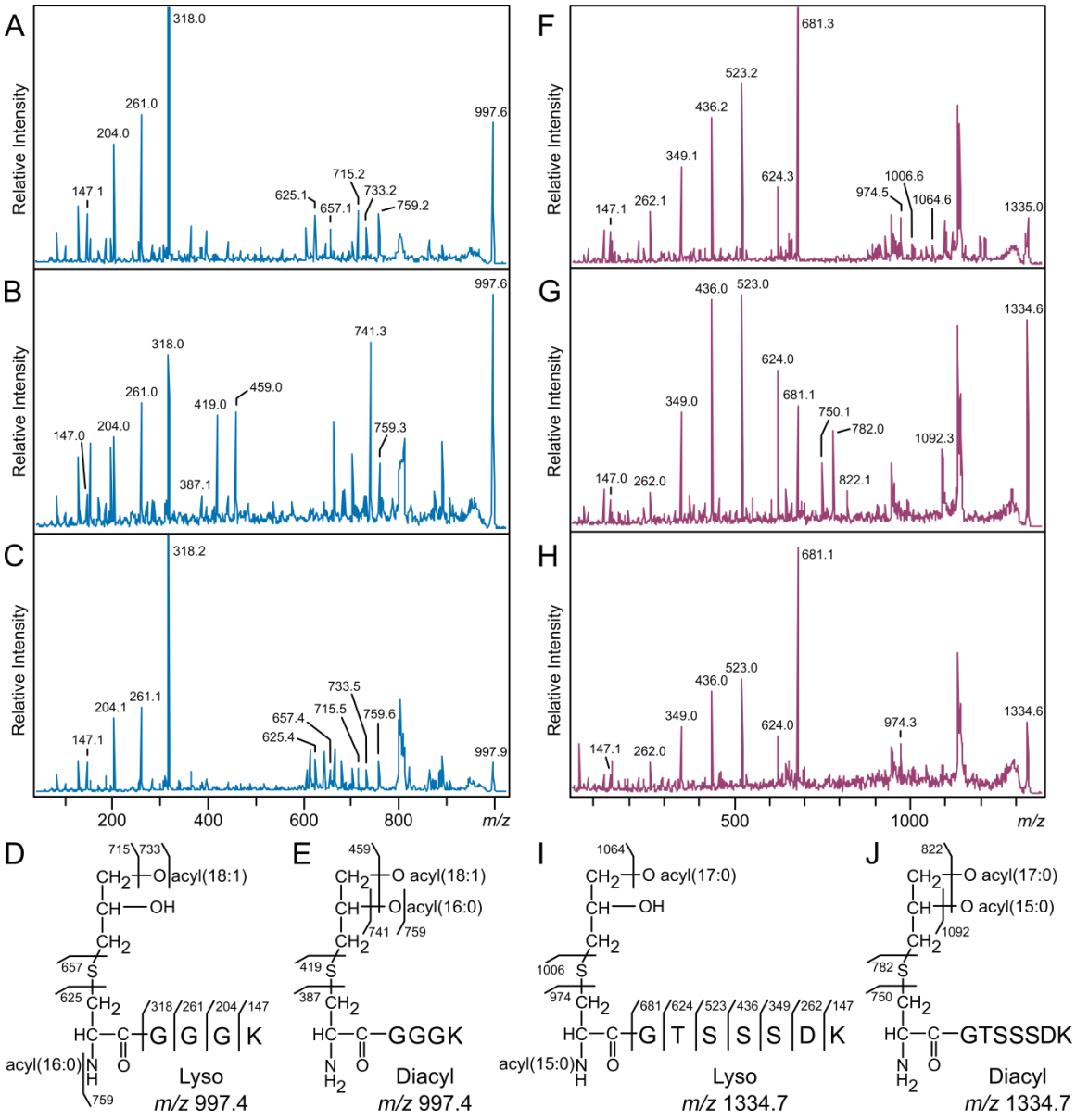


Figure **2-10**: MALDI-TOF MS/MS of *E. faecalis* PnrA and *B. cereus* PrsA. MS/MS spectra of the *m/z* 997 parent ion corresponding to the *N*-terminal lipopeptide of *E. faecalis* PnrA purified from the wildtype strain (A), the deletion strain (B), and the plasmid back-complemented strain (C). MS/MS spectra of the *m/z* 1334 parent ion corresponding to the *N*-terminal lipopeptide of *B. cereus* PrsA purified from the wildtype strain (F), the deletion strain (G), and the back-complemented strain (H). The elucidated lyso and diacylglycerol structures are shown for PnrA (D and E) and PrsA (I and J).

Similar differences were observed in the MS/MS spectra of *B. cereus* PrsA purified from wildtype and BC1526 deletion strains (Figure 2-10F-H). The MS/MS spectrum of PrsA from wildtype features the *N*-acyl($C_{16:0}$)-dehydroalanyl peptide and thiolated *N*-acyl($C_{16:0}$)-peptide peaks at *m/z* 974 and 1006, respectively, and is consistent with the lyso-form (Figure 2-10F). These peaks are absent in the MS/MS spectrum of PrsA purified from the deletion strain (Figure 2-10G). Instead, two fragment ions at *m/z* 750 and 782 appeared that correspond to the nonacylated dehydroalanyl and thiolated CGTSSSDK peptides, respectively. Additional ions at *m/z* 1092 and 822 corresponding to the product ions that have lost a $C_{15:0}$ fatty acid ($C_{14}H_{28}COOH$) or both fatty acids (32:0) were also detected. While the exact positions of the *O*-acyl chains could not be determined for either lyso form of PnrA or PrsA, back complementation in both *E. faecalis* and *B. cereus* deletion strains with plasmids encoding the candidate genes restored lyso-form lipoprotein maturation (Figure 2-10C and H).

Discussion

Lipoproteins are highly conserved components of bacterial cell envelopes from diverse genera that thus far have been thought to share a common biosynthetic pathway (5). The sequential action of diacylglcyerol transferase (Lgt) and the signal peptide II protease (Lsp) is invariant, while those organisms with an OM further process diacylglyceryl lipoproteins by *N*-acylation (Lnt) (Figure 2-1). This modification is important as the Lol transport machinery preferentially recognizes triacylated lipoprotein substrates and is therefore required for robust translocation rates (14, 16). Since Firmicutes lack an OM, it came as no surprise when genomic sequencing failed to identify *lnt* orthologs. No other enzymes have been implicated in lipoprotein biosynthesis, though *lnt* has been characterized outside of OM-containing bacteria in mycobacteria and high-GC Gram-positive organisms (11, 43–45). However, when Kurokawa et

al. surveyed the *N*-terminal structure among a panel of Firmicutes, they uncovered novel lipoprotein structural diversity that included *N*-acylation (12). The discovery of lyso-form lipoproteins, and other *N*-terminal modified forms in low-GC Gram-positive Firmicutes lacking apparent *lnt* sequence orthologs, suggests unique biosynthesis pathways that have been tailored for host organism-specific physiological roles. Herein, we have identified one such protein responsible for lyso-form lipoprotein biosynthesis in *E. faecalis* and *B. cereus*.

To identify putative genes involved in lyso-lipoprotein formation, we engineered a selection strain for intergenic complementation (Figure 2-1). The essential role of Lnt in *N*-acylating substrate for Lol-mediated transport in *E. coli*, in combination with the reduced burden of lipoprotein processing in an Lpp null background, allowed us to identify two previously unknown candidate genes: *E. faecalis* WMC_RS08810 and *B. cereus* BC1526. These genes supported growth in Lnt-deficient *E. coli* cells, prevented mislocalization of LppK58A, and restored general cell envelope integrity as judged by phenotypic resistance to the cell wall biosynthesis inhibitor vancomycin. Vancomycin is normally unable to cross the OM and inhibit peptidoglycan synthesis, unless there are permeability defects. As such, vancomycin is a useful probe of membrane integrity (46, 47), and was critical in reducing background growth in the Lpp null selection strain background. Suppression of Lnt essentiality through loss of Lpp has previously been observed (14, 48). Curiously, localization of lipoproteins directly involved in OM porin assembly (BamB and BamC), LPS transport (LptE), and LPS levels did not reveal gross defects. This suggests loss of OM integrity due to Lnt depletion in an Lpp null background may be more complex than originally envisioned. Regardless, substrate competition experiments, in tandem with MS analyses, confirmed that lyso-form lipoproteins improve cell fitness through restoration of cell envelope integrity. Once more, deletion of either of these genes alone in *E. faecalis* or *B. cereus* produced conventional diacylglycerol-modified lipoproteins with free α-amino termini. As lipoprotein *N*-acylation occurs concomitant with *O*-deacylation, we have

named this protein lipoprotein intramolecular transacylase (Lit) to differentiate from the canonical *N*-acyl transferase Lnt (Figure 2-11).

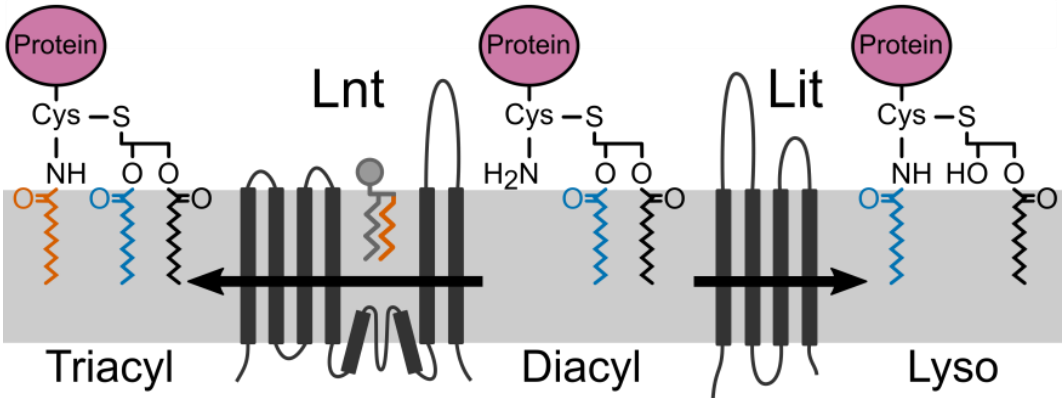


Figure **2-11**: Bacterial lipoprotein *N*-acylation. In Gram-negative bacteria, Lnt transfers the *sn*-1 acyl chain (orange) of a membrane phospholipid to form triacylated lipoprotein. In low-GC Gram-positive Firmicutes that make lyso-form lipoproteins, it is proposed that Lit internally transfers the *sn*-2 acyl chain (blue) from the diacylglycerol moiety to the α-amino group of cysteine. When heterologously expressed in *E. coli* (see Figure 2-6), the *N*-terminal dehydroalanyl fragment ion of Lpp shifts from an Lnt-catalyzed intermolecular *sn*-1 $C_{16:0}$ acyl chain transfer (*m/z* 813, orange acyl chain) to a WMC_RS08810-catalyzed intramolecular *sn*-2 $C_{16:1}$ acyl chain transfer (*m/z* 811, blue acyl chain).

The Lit family of enzymes have no detectable amino acid sequence similarity or predicted membrane topology to that of Lnt (14, 49), as would be predicted for the proposed intramolecular *O*- to *N*-transacylation involving the *sn*-2 fatty acid. At present, a bifunctional enzyme that *O*-deacylates, and adds an *N*-acyl chain from a neighboring phospholipid cannot be definitively ruled out. This would seem unlikely, however, given the small size (~220 amino acids with four predicted transmembrane domains (50); Figure 2-11). While we could not directly determine the exact position of the remaining *O*-acyl chain on the lyso-form lipoproteins by MS/MS analyses, fatty acid composition of the membrane phospholipid pool supports a *sn*-2 origin for the *N*-terminal acyl chain. In *E. coli*, *de novo* phospholipids are synthesized by PlsB and PlsC through sequential addition to the glycerol phosphate backbone of the *sn*-1 and *sn*-2

fatty acids, respectively (51, 52). Both enzymes have well-characterized and distinct preference for acyl chain substrates. PlsB prefers $C_{16:0}$ and $C_{18:1}$ fatty acids and typically does not accept $C_{16:1}$ acyl chains as substrates (53), while PlsC lacks strict substrate specificity and can incorporate saturated or unsaturated fatty acids (51). Interestingly, there was a two-mass-unit difference in the *N*-acylated lipopeptide fragment ions of LppK58A when processed by *E. faecalis* WMC_RS08810 versus *E. coli* (Figure 2-6C and F). Previous studies on acyl chain specificity show that Lnt transfers the *sn*-1 acyl chain of phospholipid donors, preferring $C_{16:0}$ saturated acyl chains over $C_{16:1}$ (54, 55). Indeed, when processed by Lnt, the dominant Lpp fragment ion at *m/z* 813 contained a $C_{16:0}$ *N*-acyl chain (Figure 2-6C). As WMC_RS08810-processed Lpp contained a $C_{16:1}$ *N*-acyl chain (*m/z* 811; Figure 2-6F), this suggests that this fatty acid was added to the diacylglycerol donor by PlsC and thus, *E. faecalis* WMC_RS08810 likely catalyzes the transfer of the *sn*-2 fatty acid to the free α-amino group of the *N*-terminal cysteine. Mechanistically, two distinct scenarios can be proposed. In the first, *O*-deacylation through side chain nucleophilic attack of an ester-bound acyl chain forms a covalent enzyme-fatty acid intermediate, which then reacylates to form an amide at the α-amino group of cysteine. Alternatively, the enzyme simply positions the α-amino terminus for direct nucleophilic attack on the ester-linked fatty acid. In either case, formation of an amide from an ester would be thermodynamically favored and result in net conversion of diacylglycerol lipoprotein precursors into lyso-form lipoproteins. Reconstitution of Lit and *in vitro* characterization with a defined phospholipid pool representative of the Firmicute native host membrane environment will be necessary to confirm acyl chain specificity.

The physiological role of lyso-form lipoproteins and the greater role of *N*-terminal lipoprotein modifications in Gram-positive bacteria in general remains unknown. In contrast to Gram-negative bacteria, enzymes involved in lipoprotein biosynthesis are not essential in Gram-positive organisms (3, 5). Indeed, deletions of WMC_RS08810 and BC1526 were readily

constructed in *E. faecalis* and *B. cereus* without any apparent phenotypes. Lit orthologs are present in *Lactobacillus bulgaricus* and *Streptococcus sanguinis*, both also known to make lyso-form lipoproteins (12), suggesting Lit is conserved amongst organisms that make the lyso-form. Conversely, Lit is absent from Firmicute species that are known to not make the lyso form, including *Staphylococcus aureus* RN4220 which makes triacylated lipoproteins, *Listeria monocytogenes* strain ATCC 15313 which has diacylglycerol-modified lipoproteins, and *Bacillus subtilis* strain 168 which elaborates *N*-acetyled lipoproteins (12, 56, 57). Lit is not present in any other phylum of bacteria, suggesting that the gene has been independently acquired in Firmicutes at the species level as a possible "post-edit" to diacylglycerol modified lipoproteins. Consistent with a relatively recent acquisition, there is no genomic synteny between WMC_RS08810 and BC1526. While the selective pressure operative for *N*-acylation is currently unknown, the identification of the Lit enzyme family provides opportunities for further study.

Materials and Methods

Bacterial strains and growth conditions

E. coli strains are derivatives of the reference strain BW25113 and were grown in LB-Miller medium at 37°C with agitation. Cultures were supplemented with 0.2% L-arabinose when necessary. Antibiotic markers were selected with carbenicillin (50 or 100 μg/mL), kanamycin (30 μg/mL), spectinomycin (25 or 50 μg/mL), apramycin (100 μg/mL), and chloramphenicol (25 μg/mL) unless otherwise noted. *E. faecalis* and *B. cereus* strains are derivatives of reference strains ATCC 19433 and ATCC 14579, respectively. All strains were grown in tryptic soy broth (TSB) at 37°C, and selected with erythromycin (10 μg/mL) or tetracycline (5 μg/mL) when

appropriate. *E. faecalis* was grown with nisin (100 ng/mL) when appropriate. Strains used in this study are listed in Table 2-1.

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

Strain/Plasmid	Relevant Genotype/Phenotype[a]	Source or Ref.
E. coli Strains[b]		
BW25113	*E. coli* K-12 wildtype [Δ*(araDaraB) 567* Δ*lacZ4787 (::rrnB-3)* λ⁻ *rph-1* Δ*(rhaD-rhaB)568 hsdR514*]	CGSC7636[c]
TXM315	*gut::*Kan^r *-rrnB TT-araC-*P_{BAD}*-lnt*	This study
TXM327	*lpp::*Chl^r	This study
KA325	*ybeX-(kanR-rrnB TT-araC-*P_{BAD}*)-lnt*	This study
KA349	TXM327 *ybeX-(kanR-rrnB TT-araC-*P_{BAD}*)-lnt*	This study
KA472	TXM315 *lnt::*Spt^r	This study
KA486	TXM327 + pEF4	This study
KA489	TXM327 + pBC7	This study
KA494	TXM327 + pUC19	This study
KA495	TXM327 + pUC19-*lnt*	This study
KA503	pUC19	This study
KA504	pUC19-*lnt*	This study
KA505	pEF4	This study
KA528	KA349 + pKA524	This study
KA532	KA486 *lnt::*Spt^r by transduction + pKA522	This study
TXM537	KA325 + pEF4	This study
TXM538	KA325 + pUC19	This study
TXM539	KA349 + pEF4	This study
TXM540	KA349 + pUC19	This study
TXM541	KA472 *chiQ::* Apr^r	This study
KA548	TXM327 + pKA522	This study
KA555	pBC7	This study
KA583	KA489 *lnt::* Spt^r by transduction + pKA522	This study
KA597	KA325 + pBC7	This study
KA598	KA349 + pBC7	This study
KA599	TXM327 + pEF4 + pKA522	This study
KA600	TXM327 + pBC7 + pKA522	This study
Firmicutes Strains		
TXM465	*E. faecalis* ATCC 19433	Sigma
KA477	*B. cereus* ATCC 14579	American Type Culture Collection
KA543	TXM465 ΔWMC_RS08810	This study
KA546	KA477 ΔBC1526	This study
KA637	KA546 + pKA637	This study
KA666	KA543 + pKA635	This study
Plasmids		
pUC19	General cloning vector; Car^r	
pCL25		
pKFC		
pEF4	pUC19-EF genomic DNA; Car^r	This study
pBC7	pUC19-BC genomic DNA; Car^r	This study
pKA495	pUC19-*lnt*; Car^r	This study
pKA522	*lppK58A-Strep*; Kan^r	This study
pKA524	*lppK58A-Strep*; Spt^r	This study
pKA635	pMSP3535-WMC_RS08810; Ery^r	This study
pKA637	pSP-BC1526-hp; Tet^r	This study

[a] Resistance: Car^r, carbenicillin; Kan^r, kanamycin; Spt^r, spectinomycin; Chl^r, chloramphenicol; Ery^r, erythromycin; Tet^r, tetracycline; Apr^r, apramycin.
[b] All *E. coli* strains are derivatives of BW25113.
[c] Strain CGSC7636 at the Coli Genetic Stock Center (CGSC).

Construction of deletion strains, the P_{BAD}-*lnt* cassette, and plasmids

All gene deletions in *E. coli* were constructed by the Red recombinase system and transduced into recipient strains by P1*vir* when appropriate (17). Primers used in this study are listed in Table 2-2.

Table **2-2**: Primers used in this study.

Description	**Sequence (5' to 3')**
5'- Kan^r	CAGACCGCTTCTGCGCAGTGGGCTTACATG
Kan^r-3'	GACGGCCAGTGAATTGAAATCTCGTGATGG
5'-*rrnB* terminator	CGCAGAAGCGGTCTGATAAAACAG
rrnB terminator-3'	GGAGAGCGTTCACCGACAAACAAC
5'-*araC* P_{BAD}	TGATTACGCCAAGCTTGTTAGCCCAAAAAACGG
araC P_{BAD}-3'	CGGTGAACGCTCTCCCAATTGTCTGATTCG
5'-*ybeX* P_{BAD}-*lnt* integration	GTCAAAATCCCGGATGACTCACCCCAGCCGAAGCTGGATGAATAAGAAATCTCGTGATGGCAGGTTGGGC
ybeX P_{BAD}-*lnt* integration-3'	TTCAATTAATGAGGCAAAAGCCATGTAGTTATCTATCCAGTTTCGGGTTAGCCCAAAAAACGGGTATGGAG
5'-*lpp::*Chl^r	CTGGGCGCGGTAATCCTGGGTTCTACTCTGCTGGCAGGTTGCTATGAATATCCTCCTTAG
*lpp::*Chl^r-3'	GGCGCACAATGTGCGCCATTTTTCACTTCACAGGTACTATTAGTAGGCTGGAGCTGCTTC
5'-*gutA::*P_{BAD}-*lnt*	CATGGTGCAGAGTGGTTTATCGGGCTGTTCCAAAAGGGCGGAGAGGTGTTTAGACCATGATTACGCCAAGC
gutA::P_{BAD}-*lnt*-3'	CGATCCCGGCGCGATATCGTCGGGGACAGGGCCTGCCACATGGACAGTGCCCGTAAAACGACGGCCAGTG
5'-*chiQ::*Apr^r	CTCATCGCCATAATGGCATCGGGGCTGGTAGCTTGTGCGCAATCACCTAGATCCTTTTGG
*chiQ::*Apr^r-3'	CTGCGTATTTCCTGCCAGACTTTATCACAATCGGCTTTCACCCGTTCTCCGCTCATGAGC
5'-*lnt::*Spt^r	GCCTCATTAATTGAACGCCAGCGCATTCGCCTGCTGCTGGCGATTCCCCTGCTCGCGCAG

*lnt::*Spt^r-3'	CACAGCAGCAAAACCAAACAATGCCGTCAGCACCCACAGCG GGCTTGAACGAATTGTTAG
5'-*lnt* pUC19	GATTACGCCAAGCTTAGCCGAAGCTGGATG
lnt pUC19-3'	GACGGCCAGTGAATTCATGTATTCCGGCACG
5'-*lppK58A*	GCTTTTGTGAATTAATTTGTATATCG
lppK58A-3'	CACTTCACAGGTACTATTACGCGCGGTATTTAGTAG
5'-Kan^r	CCTCGACTTCGCTGCAAGATCCCCTCACGCTG
Kan^r-3'	AGTACCTGTGAAGTGAAGCTTTTGGTCGGTCATTTCG
5'-*repA*	GCAGCGAAGTCGAGGCATTTCTGTC
repA-3'	TTAATTCACAAAAGCGAATTCCCGACAGTAAG
5'-Spt^r	CAGGGGATCAAGATCTATTCCCCTGCTCGCGCAG
Spt^r-3'	CTGTGAAGTGAAGCTTGCTTGAACGAATTGTTAG
5'-*lppK58A-Strep*	CATCCTCAATTTGAAAAATAATAGTACCTGTGAAGTG
lppK58A-Strep-3'	ACTCCACGCTGAACCGGATCCCGCGCGGTATTTAGTAG
5'-pKFC WMC_RS08810 upstream	GACGGCCAGTGAATTCAAGCCAGTGATTCG
pKFC WMC_RS08810 upstream-3'	GAGTCCAAGTGTTTCCCTTAAACGCAATTTCG
5'-WMC_RS08810 downstream	GAAACACTTGGACTCTTATTAGAAGGCCTG
WMC_RS08810 downstream-3'	TGATTACGCCAAGCTCATCTAGTCCTTCGG
5'-pKFC BC1526 upstream	GACGGCCAGTGAATTTCCTGTGGCAACTGG
pKFC BC1526 upstream-3'	CCATTGTTGCGAGTGCAAAAATAGAAAATGC
5'-pKFC BC1526 downstream	GCACTCGCAACAATGGTCTCTGCTTAAAGAG
pKFC BC1526 downstream-3'	TGATTACGCCAAGCTAGGCACTACGTACTC

5'-WMC_RS08810 pMSP3535	CCACTAGTCCCGGGCTGCAGCCTATGTTGTTGATTCC
WMC_RS08810 pMSP3535-3'	GAGACCGGCCTCGAGGCGAGCAATTTTGGAATTG
5'-BC1526 pSP-hp	CAAGGAGGTGAATGTACATTGGATCGATTGATTAC
BC1526 pSP-hp-3'	CGGCCGGTACCGGATCCTAGAACAGAACTAGACCG

The P_{BAD}-*lnt* cassette was constructed similarly to previously described, with some modifications (14). DNA fragments encoding the kanamycin resistance gene (Kan^r), the *rrnB* transcription terminator (*TT*), the *araC* gene, and the L-arabinose-inducible P_{BAD} promoter were PCR-amplified and assembled into pUC19 vector using the In-Fusion HD Cloning Kit (Clontech). The cassette was then PCR-amplified from the resulting plasmid with ~45-bp homology regions (primers 5'-*ybeX* P_{BAD}-*lnt* integration and *ybeX* P_{BAD}-*lnt* integration-3') and integrated directly upstream of *lnt* by Red-mediated recombination. Integrants were selected on L-arabinose and kanamycin and verified by PCR. The arabinose-dependent growth phenotype was confirmed in the resulting KA325 strain, into which the *lpp::*Chl^r allele was transduced from TXM327 using P1*vir* to create KA349.

To create the plasmid encoding a C-terminal Strep-tagged LppK58A allele, DNA fragments encoding the Kan or Spt resistance gene, the *repA* gene from pCL25, and *lppK58A* were first PCR-amplified and assembled using the In-Fusion HD Cloning Kit. With the resulting plasmid as template, the Strep-tag was added by inverse PCR (primers 5'-*lppK58A-Strep* and *lppK58A-Strep*-3').

Unmarked, internal gene deletions of *E. faecalis* WMC_RS08810 and *B. cereus* BC1526 were generated using the pKFC plasmid (18). The *E. faecalis* deletion strain was back-complemented by cloning WMC_RS08810 into the shuttle expression vector pMSP3535, a gift from Gary Dunny (Addgene plasmid #46886), under the control of a nisin-inducible promoter

(19). The *B. cereus* deletion strain was back complemented by cloning BC1526 into the shuttle expression vector pSPNprM-hp, a gift from Dieter Jahn (Addgene plasmid #48120), under the control of a xylose-inducible promoter (20). All plasmids used in this study are listed in Table 2-1.

Construction of *E. faecalis* and *B. cereus* genomic fragment libraries and complementation of the *lnt*-inducible *E. coli* strain KA349

E. faecalis and *B. cereus* genomic DNA was digested with NEBNext dsDNA Fragmentase (NEB) per manufacturer's instructions for 0.5-2 minutes at 37°C. After size fractionation by agarose gel electrophoresis, DNA fragments ranging from 3- to 5-kb were isolated. Fragments were end-repaired with components from the pWEB Cosmid Cloning Kit (Epicentre) and ligated using T4 ligase into pUC19 that had been *Sma*I-linearized and dephosphorylated with calf intestinal phosphatase (NEB). The resulting libraries were transformed into the conditional-lethal arabinose-regulated *lnt* mutant strain KA349 by electroporation. Transformants were selected on LB agar containing 100 μg/mL carbenicillin, 10 μg/mL vancomycin, 50 μg/mL palmitate, and 0.2% (w/v) glucose at 37°C. As a positive control, a portion of the electroporated cells were also plated on the same media with L-arabinose in place of glucose. Based on the number of colonies observed on the positive control plates, the average insert size of 4-kb, and the known size of the *E. faecalis* and *B. cereus* genomes (2.9 and 5.4 Mb, respectively), approximately 61-fold total coverage of the *E. faecalis* genome and 51-fold total coverage of the *B. cereus* genome were screened. Plasmids were isolated from colonies growing in the absence of L-arabinose, screened by restriction digest to confirm insert size, and the junction sites sequenced to identify the DNA insert.

Growth analysis

Starter cultures were diluted into fresh LB containing 50 μg/mL carbenicillin with or without L-arabinose to an optical density at 600 nm (OD_{600}) of 0.001. When the prediluted cultures reached an OD_{600} of 0.5, cultures were rediluted into fresh media to an OD_{600} of 0.01. The OD_{600} was recorded every 30 minutes.

Genetic linkage analysis

To construct a passively marked linked antibiotic resistance marker for co-transduction analysis, the full *kanR-TT-araC-*P_{BAD}*-lnt* cassette from KA325 was PCR-amplified (5'-*gutA::*P_{BAD}*lnt* and *gutA::*P_{BAD}*-lnt*-3') and integrated into the glucitol locus by the Red recombinase system. Then, the chitin metabolism gene *chiQ* and wildtype *lnt* genes were replaced with apramycin and spectinomycin resistance genes, respectively, while maintained on L-arabinose to create strain TXM541. Donor lysates of TXM541 were prepared using P1*vir* and transduced into experimental recipient strains using standard protocols (21). Candidate transductants were first selected on LB agar with apramycin then patched onto LB agar with spectinomycin to score for co-transduction of *lnt::*Spt^r.

Separation of inner and outer membranes by discontinuous sucrose gradient centrifugation

Starter cultures of KA528 and KA532 were inoculated 1:1000 into 1L LB with L-arabinose and 25 μg/mL spectinomycin or 50 μg/mL each of carbenicillin and kanamycin, respectively. An additional starter culture of KA528 was washed three times, resuspended in fresh LB, and inoculated 1:2000 into 1L LB with 25 μg/mL spectinomycin in order to deplete Lnt protein. Cultures were grown to an OD_{600} of 1.0-1.2, harvested by centrifugation, washed with 25

mL PBS, and stored frozen. Cell pellets were processed as previously described (22, 23) and were resuspended in 25 mL of 50 mM Tris-HCl pH 7.8 containing 1 mM EDTA and 0.1 mg/mL each of DNAse and RNAse. Cell suspensions were passed through a French pressure cell four times at 14,000 psi followed by the addition of PMSF protease inhibitor to 1 mM and $MgCl_2$ to 2 mM. Unbroken cells were removed by centrifugation (3,200xg, 10 minutes, 4°C) before collecting membranes by ultracentrifugation (110,000xg, 90 minutes, 4°C). The supernatant was discarded, and the membrane pellet was resuspended in 10 mM HEPES pH 7.4 (adjusted with NaOH) and collected by centrifugation as before. The resulting total membrane fractions were homogenized in 3 mL of HEPES buffer with the aid of a small-bore 26 G needle and layered onto a discontinuous sucrose gradient containing 5 mL of 2.02 M, 18 mL of 1.44 M, and 12 mL of 0.77 M sucrose in HEPES buffer. Tubes were centrifuged at 80,000xg in a swinging bucket rotor for 19 hours at 4°C. One-mL fractions were collected by piercing the bottom of the centrifuge tubes with a syringe needle. The amount of protein in each fraction was measured by Bradford reagent. Fractions containing protein corresponding to the OM and IM were pooled and their total volumes normalized with HEPES buffer. Membrane fractions were stored frozen at -20°C.

NADH oxidase assay

Incubation mixtures contained 50 mM Tris-HCl pH 7.5, 0.12 mM NADH, 0.2 mM dithiothreitol, and 3 μL of the pooled and normalized membrane fraction in a volume of 1.0 mL. The rate of decrease in absorbance at 340 nm was assayed at room temperature as has been described (24).

Immunoblotting for LppK58A-Strep, BamB, BamC, and LptE from membrane fractions

Normalized membrane fractions were separated by SDS-PAGE using a 16.5% Tris-tricine gel (25), transferred to a PVDF membrane (0.2 μM), incubated with a 1:5000 dilution of Precision Protein StrepTactin-HRP Conjugate (Bio-Rad) per manufacturer's instructions, and LppK58A-Strep levels detected by enhanced chemiluminescence (SuperSignal West Pico). To detect BamB and BamC, membrane fractions were separated by SDS-PAGE using a 10% Tris-glycine gel and a 12% Tris-glycine gel for LptE. After transfer to PVDF membranes, samples were incubated with a 1:5000 dilution of polyclonal antibodies against BamB, BamC, and LptE overnight at 4°C. Bands were visualized using HRP-conjugated secondary antibody (1:5000) and detected by chemiluminescence.

Lipoprotein *N*-acylation whole cell competition assay

Strains expressing *lnt* and/or the Firmicute candidate *N*-acyl transferase genes, along with a selection-compatible LppK58A-Strep reporter low-copy number plasmid construct (pKA522 or pKA524), were grown to mid-log (OD_{600} = 0.5-0.6), pelleted by centrifugation (5,000xg, 2 minutes, room temperature), and washed once with PBS. The supernatant was discarded and the cell pellet stored frozen. To obtain Lnt-depleted samples, KA528 was inoculated 1:6000 into LB with 0.2% glucose and grown until cell division stopped as monitored by OD_{600}. Aliquots of OD_{600}-normalized whole cells were boiled in SDS-PAGE sample loading buffer and separated by SDS-PAGE using a 16.5% Tris-tricine gel. Immunoblots for LppK58A-Strep were performed as described above.

LPS gel and silver stain

The LPS profiles of the normalized membrane fractions were analyzed by SDS-PAGE (16.5% Tris-tricine gel) and visualized by silver staining (26).

Purification of LppK58A-Strep from *E. coli*

One-liter cultures of *E. coli* expressing Lnt only (KA548) or the *E. faecalis* candidate *N*-acyltransferase WMC_RS08810 (KA532) along with LppK58A-Strep were grown to late exponential phase (OD_{600} = 1.5) and harvested by centrifugation (3,200xg, 10 minutes, 4°C). Cells were resuspended in 25 mL of 50 mM Tris-HCl pH 7.8 containing 1 mM EDTA and 0.1 mg/mL each of DNAse and RNAse, and broken by French press as described above. Membranes were collected by ultracentrifugation (110,000xg, 90 minutes, 4°C), and stored at -20°C. Membranes were resuspended in 10 mM HEPES pH 7.4 with 1 mM EDTA to 1 mg/mL of protein. Membrane-bound proteins were solubilized by supplementing to 2% Triton X-100 (v/v) and incubating samples at room temperature for 30 minutes with gentle shaking (27, 28). At the end of the incubation period, samples were clarified by centrifugation (20,000xg, 20 minutes, rt). Supernatants were diluted with 3 volumes of HEPES buffer containing 0.5 mM *n*-dodecyl-β-D-maltopyranoside (DDM) and passed over StrepTactin Sepharose resin (IBA Biosciences). Protein was eluted as directed by manufacturer, except all wash and elution solutions were supplemented with 0.5 mM DDM. Proteins were concentrated to 1 mg/mL using a Pierce Protein Concentrator (9 K MWCO), aliquoted, and frozen at -80°C.

Purification of native lipoproteins by Triton X-114 phase partitioning

E. faecalis and *B. cereus* cultures were grown to late exponential phase (OD_{600} = 1.5) in 15 mL of TSB, washed once with TBSE solution (20 mM Tris-HCl pH 8.0, 130 mM NaCl and 5 mM EDTA), and frozen until use. Native bacterial lipoproteins were enriched using the Triton X-114 phase partitioning method with some modifications (12). Briefly, the harvested bacterial cells were resuspended in 800 µL TSBE with 1 mM PMSF and 0.5 mg/mL lysozyme at 37 °C for 30 minutes. Cells were disrupted by bead beating using a MagNA Lyser (7000 setting, equal volume of 0.1 mm zirconia silica beads, 5 by 20-second cycles with 2 minute chilling period on ice in between cycles). Unbroken cells and beads were removed by low speed centrifugation (3,000xg, 5 minutes, 4°C). Beads were washed with a second aliquot of TBSE, the supernatants pooled, and then supplemented with Triton X-114 to a final concentration of 2% (v/v). After incubation at 4 °C for 1 hour, the mixture was incubated at 37°C for 10 minutes to induce phase separation. Samples were centrifuged (10,000xg, 10 minutes, rt), the upper aqueous phase was removed and replaced with the same volume of TBSE solution. This procedure was repeated twice more, with subsequent incubations lessened to 10 minutes. The final Triton X-114 phase was chilled on ice, bought to 500 µL with ice cold TSBE, and the single phase solution immediately centrifuged (16,000xg, 2minutes, 4°C) to remove insoluble integral membrane proteins. The lipoprotein-enriched fraction was obtained from the supernatant by precipitation with 3 volumes of acetone and overnight incubation at -20 °C. Pellets were washed twice with acetone, air dried, and resuspended in 10 mM Tris-HCl pH 8.0 or directly in SDS-PAGE sample buffer.

Electroblotting, tryptic digestion, and lipopeptide extraction from nitrocellulose membranes

To prepare lipopeptides for MS analysis, the strategy of Serebryakova et al. was utilized with minor modifications (29). After SDS-PAGE separation (described above for immunoblotting), purified LppK58A-Strep and enriched lipoproteins from *E. faecalis* and *B. cereus* were transferred onto BioTract NT nitrocellulose membrane (PALL Life Sciences) and stained with Ponceau S (0.2% in 5% acetic acid). The stained protein bands were excised, thoroughly destained with water, and diced. Nitrocellulose pieces were washed twice with 50 mM NH_4HCO_3 pH 7.8, and digested overnight at 37°C in 20 µl of a 20 µg/mL solution of Trypsin Gold (Promega) in 50 mM NH_4HCO_3 pH 7.8. Hydrophilic peptides were sequentially extracted stepwise with 50 µl of aqueous 0.5% trifluoroacetic acid, 10% acetonitrile, and 20% acetonitrile, using a 10-minute room temperature incubation for each solution. A final 10-minute extraction with 15 µl of a 10 mg/mL α-cyano-4-hydroxycinnamic acid (CHCA, Sigma) matrix solution dissolved in chloroform/methanol (2:1, v/v) was performed to elute tightly bound hydrophobic peptides, including the bulk fraction of acylated *N*-terminal lipoprotein-derived tryptic peptides.

MALDI-TOF MS and MS/MS

One µL of the extracted lipopeptide in CHCA was deposited onto a steel target plate. An optional second µL was layered onto the same spot when needed. Where indicated, aqueous $NaHCO_3$ was added to the tryptic peptide extract (final concentration 1 mM) to enrich sodiated adduct formation. MALDI-TOF MS was conducted using an Ultraflextreme (Bruker Daltonics) MALDI-TOF mass spectrometer in positive reflectron mode. The MS/MS spectra were acquired using a MALDI-TOF/TOF instrument (Ultraflextreme; Bruker Daltonics) in LIFT mode.

References

1. **Babu MM**, **Priya ML**, **Selvan AT**, **Madera M**, **Gough J**, **Aravind L**, **Sankaran K**. 2006. A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. J Bacteriol **188**:2761–2773.
2. **Nakayama H**, **Kurokawa K**, **Lee BL**. 2012. Lipoproteins in bacteria: structures and biosynthetic pathways. FEBS J **279**:4247–4268.
3. **Buddelmeijer N**. 2015. The molecular mechanism of bacterial lipoprotein modification-How, when and why? FEMS Microbiol Rev **39**:246–261.
4. **Narita S**, **Tokuda H**. 2016. Bacterial lipoproteins; biogenesis, sorting and quality control. Biochim Biophys Acta **1862**(11):1414–1423.
5. **Nguyen MT**, **Götz F**. 2016. Lipoproteins of Gram-Positive Bacteria: Key Players in the Immune Response and Virulence. Microbiol Mol Biol Rev **80**:891–903.
6. **Juncker AS**, **Willenbrock H**, **Von Heijne G**, **Brunak S**, **Nielsen H**, **Krogh A**. 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci **12**:1652–1662.
7. **Bagos PG**, **Tsirigos KD**, **Liakopoulos TD**, **Hamodrakas SJ**. 2008. Prediction of lipoprotein signal peptides in Gram-positive bacteria with a Hidden Markov Model. J Proteome Res **7**:5082–5093.
8. **Sankaran K**, **Wu HC**. 1994. Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. J Biol Chem **269**:19701–19706.
9. **Hussain M**, **Ichihara S**, **Mizushima S**. 1982. Mechanism of signal peptide cleavage in the biosynthesis of the major lipoprotein of the *Escherichia coli* outer membrane. J Biol Chem **257**:5177–5182.

10. **Gupta SD**, **Wu HC**. 1991. Identification and subcellular localization of apolipoprotein N-acyltransferase in *Escherichia coli*. FEMS Microbiol Lett **62**:37–41.
11. **Tschumi A**, **Nai C**, **Auchli Y**, **Hunziker P**, **Gehrig P**, **Keller P**, **Grau T**, **Sander P**. 2009. Identification of apolipoprotein N-acyltransferase (Lnt) in mycobacteria. J Biol Chem **284**:27146–27156.
12. **Kurokawa K**, **Ryu K-H**, **Ichikawa R**, **Masuda A**, **Kim M-S**, **Lee H**, **Chae J-H**, **Shimizu T**, **Saitoh T**, **Kuwano K**, **Akira S**, **Dohmae N**, **Nakayama H**, **Lee BL**. 2012. Novel bacterial lipoprotein structures conserved in low-GC content Gram-positive bacteria are recognized by Toll-like receptor 2. J Biol Chem **287**:13170–13181.
13. **Reffuveille F**, **Serror P**, **Chevalier S**, **Budin-Verneuil A**, **Ladjouzi R**, **Bernay B**, **Auffray Y**, **Rincé A**. 2012. The prolipoprotein diacylglyceryl transferase (Lgt) of Enterococcus faecalis contributes to virulence. Microbiology **158**:816–825.
14. **Robichon C**, **Vidal-Ingigliardi D**, **Pugsley AP**. 2005. Depletion of apolipoprotein N-acyltransferase causes mislocalization of outer membrane lipoproteins in *Escherichia coli*. J Biol Chem **280**:974–983.
15. **Vidal-Ingigliardi D**, **Lewenza S**, **Buddelmeijer N**. 2007. Identification of essential residues in apolipoprotein N-acyl transferase, a member of the CN hydrolase family. J Bacteriol **189**:4456–4464.
16. **Fukuda A**, **Matsuyama S-I**, **Hara T**, **Nakayama J**, **Nagasawa H**, **Tokuda H**. 2002. Aminoacylation of the N-terminal cysteine is essential for Lol-dependent release of lipoproteins from membranes but does not depend on lipoprotein sorting signals. J Biol Chem **277**:43512–43518.

17. **Yakushi T**, **Masuda K**, **Narita S**, **Matsuyama S**, **Tokuda H**. 2000. A new ABC transporter mediating the detachment of lipid-modified proteins from membranes. Nat Cell Biol **2**:212–218.

18. **Matsuyama SI**, **Yokota N**, **Tokuda H**. 1997. A novel outer membrane lipoprotein, LolB (HemM), involved in the LolA (p20)-dependent localization of lipoproteins to the outer membrane of *Escherichia coli*. EMBO J **16**:6947–6955.

19. **Zhao XJ**, **Wu HC**. 1992. Nucleotide sequence of the *Staphylococcus aureus* signal peptidase II (lsp) gene. FEBS Lett **299**:80–84.

20. **Qi H-Y**, **Sankaran K**, **Gan K**, **Wu HC**. 1995. Structure-Function Relationship of Bacterial Prolipoprotein Diacylglyceryl Transferase: Functionally Significant Conserved Regions. J Bacteriol **177**:6820–6824.

21. **Saito HE**, **Harp JR**, **Fozo EM**. 2014. Incorporation of exogenous fatty acids protects *Enterococcus faecalis* from membrane-damaging agents. Appl Environ Microbiol **80**:6527–6538.

22. **Yakushi T**, **Tajima T**, **Matsuyama S**, **Tokuda H**. 1997. Lethality of the Covalent Linkage between Mislocalized Major Outer Membrane Lipoprotein and the Peptidoglycan of *Escherichia coli*. J Bacteriol **179**:2857–2862.

23. **Guo MS**, **Updegrove TB**, **Gogol EB**, **Shabalina SA**, **Gross CA**, **Storz G**. 2014. MicL, a new σE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. Genes Dev **28**:1620–1634.

24. **Hirota Y**, **Suzuki H**, **Nishimura Y**, **Yasuda S**. 1977. On the process of cellular division in *Escherichia coli*: a mutant of *E. coli* lacking a murein-lipoprotein. Proc Natl Acad Sci **74**:1417–1420.

25. **Chimalakonda G**, **Ruiz N**, **Chng S-S**, **Garner RA**, **Kahne D**, **Silhavy TJ**. 2011. Lipoprotein LptE is required for the assembly of LptD by the -barrel assembly machine in the outer membrane of *Escherichia coli*. Proc Natl Acad Sci **108**:2492–2497.

26. **Nikaido, H; Varra M**. 1985. Molecular basis of bacterial outer membrane permeability. Microbiol Mol Biol Rev **67**:593–656.

27. **Wu TT**. 1966. A model for three-point analysis of random general transduction. Genetics **54**:405–410.

28. **Braun V**, **Rotering H**, **Ohms J-P**, **Haganmaier H**. 1976. Conformational Studies on Murein-Lipoprotein from the Outer Membrane of *Escherichia coli*. Eur J Biochem **70**:601–610.

29. **Al-Saad KA**, **Zabrouskov V**, **Siems WF**, **Knowles NR**, **Hannan RM**, **Hill HH**. 2003. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of lipids: ionization and prompt fragmentation patterns. Rapid Commun Mass Spectrom **17**:87–96.

30. **Brülle JK**, **Tschumi A**, **Sander P**. 2013. Lipoproteins of slow-growing Mycobacteria carry three fatty acids and are N-acylated by apolipoprotein N-acyltransferase BCG_2070c. BMC Microbiol **13**:223.

31. **Mohiman N**, **Argentini M**, **Batt SM**, **Cornu D**, **Masi M**, **Eggeling L**, **Besra G**, **Bayan N**. 2012. The ppm Operon Is Essential for Acylation and Glycosylation of Lipoproteins in *Corynebacterium glutamicum*. PLoS One **7**:e46225.

32. **Widdick DA**, **Hicks MG**, **Thompson BJ**, **Tschumi A**, **Chandra G**, **Sutcliffe IC**, **Brülle JK**, **Sander P**, **Palmer T**, **Hutchings MI**. 2011. Dissecting the complete lipoprotein biogenesis pathway in *Streptomyces scabies*. Mol Microbiol **80**:1395–1412.

33. **Ruiz N**, **Falcone B**, **Kahne D**, **Silhavy TJ**. 2005. Chemical conditionality: A genetic strategy to probe organelle assembly. Cell **121**:307–317.

34. **Ruiz N**, **Wu T**, **Kahne D**, **Silhavy TJ**. 2006. Probing the barrier function of the outer membrane with chemical conditionality. ACS Chem Biol **1**:385–395.

35. **Narita S ichiro**, **Tokuda H**. 2011. Overexpression of LolCDE allows deletion of the *Escherichia coli* gene encoding apolipoprotein N-acyltransferase. J Bacteriol **193**:4832–4840.

36. **Gélis-Jeanvoine S**, **Lory S**, **Oberto J**, **Buddelmeijer N**. 2015. Residues located on membrane-embedded flexible loops are essential for the second step of the apolipoprotein N-acyltransferase reaction. Mol Microbiol **95**:692–705.

37. **Omasits U**, **Ahrens CH**, **Müller S**, **Wollscheid B**. 2014. Protter: Interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics **30**:884–886.

38. **Zhang Y-M**, **Rock CO**. 2008. Thematic review series: Glycerolipids. Acyltransferases in bacterial glycerophospholipid synthesis. J Lipid Res **49**:1867–1874.

39. **Yao J**, **Rock CO**. 2013. Phosphatidic acid synthesis in bacteria. Biochim Biophys Acta **1831**:495–502.

40. **Rock C**, **Goelzg SE**, **Grant A**. 1981. Phospholipid synthesis in *Escherichia coli*. Characteristics of fatty acid transfer from acyl-acyl carrier protein to sn-glycerol-3-phosphate. J Biol Chem **258**:462–464.

41. **Jackowski S**, **Rock CO**. 1986. Transfer of fatty acids from the 1-position of phosphatidylethanolamine to the major outer membrane lipoprotein of *Escherichia coli*. J Biol Chem **261**:11328–11333.

42. **Hillmann F**, **Argentini M**, **Buddelmeijer N**. 2011. Kinetics and phospholipid specificity of apolipoprotein N-acyltransferase. J Biol Chem **286**:27936–27946.

43. **Kurokawa K**, **Lee H**, **Roh K-B**, **Asanuma M**, **Kim YS**, **Nakayama H**, **Shiratsuchi A**, **Choi Y**, **Takeuchi O**, **Kang HJ**, **Dohmae N**, **Nakanishi Y**, **Akira S**, **Sekimizu K**, **Lee BL**. 2009. The Triacylated ATP Binding Cluster Transporter Substrate-binding Lipoprotein of

Staphylococcus aureus Functions as a Native Ligand for Toll-like Receptor 2. J Biol Chem **284**:8406–8411.

44. **Asanuma M**, **Kurokawa K**, **Ichikawa R**, **Ryu KH**, **Chae JH**, **Dohmae N**, **Lee BL**, **Nakayama H**. 2011. Structural evidence of α-aminoacylated lipoproteins of *Staphylococcus aureus*. FEBS J **278**:716–728.
45. **Datsenko KA**, **Wanner BL**. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci **97**:6640–6645.
46. **Kato F**, **Sugai M**. 2011. A simple method of markerless gene deletion in *Staphylococcus aureus*. J Microbiol Methods **87**:76–81.
47. **Bryan EM**, **Bae T**, **Kleerebezem M**, **Dunny GM**. 2000. Improved Vectors for Nisin-Controlled Expression in Gram-Positive Bacteria. Plasmid **44**:183-190.
48. **Stammen S**, **Müller BK**, **Korneli C**, **Biedendieck R**, **Gamer M**, **Franco-Lara E**, **Jahn D**. 2010. High-yield intra- and extracellular protein production using *Bacillus megaterium*. Appl Environ Microbiol **76**:4037–4046.
49. **Thomason LC**, **Costantino N**, **Court DL**, **Thomason LC**, **Costantino N**, **Court DL**. 2007. *E. coli* Genome Manipulation by P1 Transduction. Curr Protoc Mol Biol **79**:1.17.1–1.17.8.
50. **Koplow J**, **Goldfine H**. 1974. Alterations in the Outer Membrane of the Cell Envelope of Heptose-Deficient Mutants of *Escherichia coli*. J Bacteriol **117**:527–543.
51. **Meredith TC**, **Aggarwal P**, **Mamat U**, **Lindner B**, **Woodard RW**. 2006. Redefining the Requisite Lipopolysaccharide Structure in *Escherichia coli*. ACS Chem Biol **1**:33–42.
52. **Osborn MJ**, **Gander J. E**, **Parisi E**, **Carson J**. 1972. Mechanism of Assembly of the Outer Membrane of *Salmonella typhimurium*. J Biol Chem **247**:3962–3972.
53. **Schägger H**. 2006. Tricine-SDS-PAGE. Nat Protoc **1**:16–22.

54. **Hitchcock PJ**, **Brown TM**. 1983. Morphological Heterogeneity Among *Salmonella* Lipopolysaccharide Chemotypes in Silver-Stained Polyacrylamide Gels. J Bacteriol **154**:269–277.

55. **Schnaitman CA**. 1971. Solubilization of the cytoplasmic membrane of *Escherichia coli* by Triton X-100. J Bacteriol **108**:545–552.

56. **Filip C**, **Fletcher G**, **Wulff JL**, **Earhart CF**. 1973. Solubilization of the cytoplasmic membrane of *Escherichia coli* by the ionic detergent sodium-lauryl sarcosinate. J Bacteriol **115**:717–722.

57. **Serebryakova M V**, **Demina IA**, **Galyamina MA**, **Kondratov IG**, **Ladygina VG**, **Govorun VM**. 2011. The acylation state of surface lipoproteins of Mollicute *Acholeplasma laidlawii*. J Biol Chem **286**:22769–22776.

Chapter 3

Enrichment of bacterial lipoproteins and preparation of N-terminal lipopeptides for structural determination by mass spectrometry

Adapted from:

Armbruster, KM, Meredith, TC. 2018. Enrichment of bacterial lipoproteins and preparation of N-terminal lipopeptides for structural determination by mass spectrometry. J Vis Exp. 135:e56842.

Abstract

Lipoproteins are important constituents of the bacterial cell envelope and potent activators of the mammalian innate immune response. Despite their significance to both cell physiology and immunology, much remains to be discovered about novel lipoprotein forms, how they are synthesized, and the effect of the various forms on host immunity. To enable thorough studies on lipoproteins, this protocol describes a method for bacterial lipoprotein enrichment and preparation of N-terminal tryptic lipopeptides for structural determination by matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS). Expanding on an established Triton X-114 phase partitioning method for lipoprotein extraction and enrichment from the bacterial cell membrane, the protocol includes additional steps to remove non-lipoprotein contaminants, increasing lipoprotein yield and purity. Since lipoproteins are commonly used in Toll-like receptor (TLR) assays, it is critical to first characterize the N-terminal structure by MALDI-TOF MS. Herein, a method is presented to isolate concentrated hydrophobic peptides enriched in N-terminal lipopeptides suitable for direct analysis by MALDI-TOF MS/MS. Lipoproteins that have been separated by Sodium Dodecyl Sulfate Poly-Acrylamide Gel Electrophoresis (SDS-PAGE) are transferred to a nitrocellulose membrane, digested *in situ* with trypsin, sequentially washed to remove polar tryptic peptides, and finally eluted with chloroform-methanol. When coupled with MS of the more polar trypsinized peptides from wash solutions, this method provides the ability to both identify the lipoprotein and characterize its N-terminus in a single experiment. Intentional sodium adduct formation can also be employed as a tool to promote more structurally informative fragmentation spectra. Ultimately, enrichment of lipoproteins and determination of their N-terminal structures will permit more extensive studies on this ubiquitous class of bacterial proteins.

Introduction

Bacterial lipoproteins are characterized by a conserved N-terminal lipid-modified cysteine that anchors the globular protein domain to the cell membrane surface. They are universally distributed in bacteria, constituting 2 to 5% of all cellular genes within a typical genome[1]. Lipoproteins play critical roles in a wide variety of cellular processes, including nutrient uptake, signal transduction, assembly of protein complexes, and in maintaining cell envelope structural integrity[2]. In pathogenic bacteria, lipoproteins serve as virulence factors[3,4]. During an infection, recognition of N-terminal lipopeptides by Toll-like receptor (TLR) 2 incites an innate immune response to remove invading pathogens. Depending on the N-terminal acylation state, lipoproteins are generally recognized by alternate TLR2 heterodimeric complexes. TLR2-TLR1 recognizes N-acylated lipopeptides, while TLR2-TLR6 binds free lipopeptide α-amino termini. Once bound, the signaling pathways converge to induce secretion of proinflammatory cytokines [3,4].

It was previously thought that lipoproteins from Gram-positive bacteria were diacylated and those from Gram-negative bacteria were triacylated, differing in the absence or presence of an amide-linked fatty acid on the conserved N-terminal cysteine residue. This assumption was supported by the lack of sequence orthologs in Gram-positive genomes to Lnt, the Gram-negative N-acyl transferase that forms triacylated lipoproteins[5]. However, recent studies have revealed lipoprotein triacylation in Gram-positive Firmicutes that lack *lnt*, as well as three novel N-terminal lipoprotein structures, termed the peptidyl, lyso, and *N*-acetyl forms[6–8]. These findings raise questions about possible yet-to-be-discovered lipoprotein forms, alongside fundamental questions about how these novel lipoproteins are made and what physiological purpose or advantage the various forms impart. Furthermore, they clearly demonstrate the current inability of genomics to predict lipoprotein structure. Indeed, we recently identified a novel class of

lipoprotein N-acyl transferases, called Lit, from *Enterococcus faecalis* and *Bacillus cereus* that makes lyso-form lipoproteins[9]. This indicates the need to experimentally verify lipoprotein structure, which can be challenging due to their extremely hydrophobic nature and limited methods available to characterize their molecular structure.

To facilitate studies on lipoprotein induction of the host immune response, as well as N-terminal structural determination, we have adapted several previously-described protocols in order to purify bacterial lipoproteins and prepare the N-terminal tryptic lipopeptides for analysis by MALDI-TOF MS[6,10–12]. Lipoproteins are enriched using an established Triton X-114 phase partitioning method, with optimization to remove contaminating non-lipoproteins and increase lipoprotein yield. These lipoproteins are suitable for direct use in TLR assays or for further purification by SDS-PAGE. For MALDI-TOF MS, transfer of the lipoproteins to nitrocellulose membrane provides a scaffold for efficient *in situ* trypsin digestion, washing, and subsequent elution from the membrane surface, resulting in highly purified N-terminal lipopeptides. Nitrocellulose has been shown to facilitate sample handling and improve sequence coverage for highly hydrophobic peptides from integral membrane proteins[13,14], as well as lipoproteins[9,10]. The method has the additional advantage of fractionating peptides based on polarity, so that intermediate wash solutions can be analyzed for high confidence protein identification simultaneously with N-terminal structural determination in a single experiment. This protocol uniquely features intentional sodium adduct formation to promote parent ion fragmentation towards dehydroalanyl ions during MS/MS, aiding in structural assignment of the *N*-acylation state. The N-terminus is both the most variable and key feature related to TLR recognition of lipoproteins. Taken together, this protocol has allowed intensive and reproducible studies on lipoproteins, with the individual stages of purification and structural determination by MALDI-TOF MS easily adapted depending on the overall goal of the experiment.

Protocol

1. Cell Growth and Lysis

1.1) Grow bacteria in 15 mL tryptic soy broth (TSB) or similar rich media to late exponential phase (OD600 of 1.0-1.5). Harvest cells by centrifugation, wash once with Tris-buffered saline/EDTA (TBSE: 20 mM Tris-hydrochloride (HCl), pH 8.0, 130 mM sodium chloride (NaCl), and 5 mM ethylenediaminetetraacetic acid (EDTA)), and continue with protocol or freeze until use. Note: Lipoprotein expression and modification can be influenced by growth conditions (e.g. acidity, salinity, growth media) and growth phase[15]. Cell growth can be scaled as desired, but 15-mL of cells is recommended for a single preparation of lipoproteins as excess biomass can decrease lipoprotein yield. One 15-mL preparation generally yields enough sample for optimal loading of ~2-4 lanes on a standard SDS-PAGE mini gel.

1.2) Resuspend cells in 800 µL TBSE with 1 mM phenylmethyl sulfonyl fluoride (PMSF) and 0.5 mg/mL lysozyme. Transfer solution to a 2.0-mL threaded microcentrifuge tube with screw cap and O-ring and incubate for 20 min at 37 °C. Note: Certain species are not susceptible to lysis by lysozyme. An appropriate lytic enzyme should be substituted to aid in lysis.

1.3) Add ~800 µL 0.1 mm zirconia/silica beads to the tube and disrupt cells by shaking at maximum speed (7,000 rpm) on a homogenizer for 5 cycles of 30 s each, with 2 min rest on ice between each cycle.

1.4) Centrifuge sample at 3000 x g for 5 min at 4 °C to pellet beads and unbroken cells. Transfer the supernatant to a new 2.0-mL microcentrifuge tube and keep on ice.

1.5) Add 200 μL TBSE to the remaining pellet and return to the homogenizer for an additional cycle. Centrifuge as above and combine supernatant with the previous supernatant (should be ~800-1000 μL total volume).

2. Enrichment of Lipoproteins by Triton X-114 Phase Partitioning

2.1) Supplement the supernatant with Triton X-114 to a final concentration of 2% (vol/vol) by adding an equal volume of 4% (vol/vol) Triton X-114 in ice-cold TBSE and incubate on ice for 1 h, mixing by inversion every ~15 min. When chilled, the supernatant and Triton X-114 will be miscible.

2.2) Transfer the tube to a 37 °C water bath and incubate for 10 min to induce phase separation. Centrifuge the sample at 10,000 x g for 10 min at room temperature to maintain bi-phasic separation.

2.3) Gently pipette off the upper aqueous phase and discard. Add ice-cold TBSE to the lower Triton X-114 phase to refill the tube to its original volume and invert to mix. Incubate on ice for 10 min.

2.4) Transfer the tube to a 37 °C water bath and incubate for 10 min to induce phase separation, then centrifuge at 10,000 x g for 10 min at room temperature.

2.5) Repeat steps 2.3-2.4 once more for a total of 3 separations. Remove the upper aqueous phase and discard.

2.6) A pellet of precipitated proteins that formed during the course of extractions may be visible at the bottom of the tube. To remove, add 1 volume of ice-cold TBSE to the Triton X-114 phase. Centrifuge at 4 °C at 16,000 x g for 2 min to pellet insoluble protein. Note: Sample should remain a single phase. If phase separation occurs, re-chill sample and centrifuge again. The pellet is generally composed of non-lipoprotein contaminants, but can be retained for further analysis.

2.7) Immediately transfer the supernatant to a fresh 2.0-mL microcentrifuge tube containing 1250 µL of 100% acetone. Pipette up and down thoroughly to wash the sample from the tip as it will be viscous. Mix by inversion and incubate overnight at -20 °C to precipitate protein.

2.8) Centrifuge the sample at 16,000 x g for 20 min at room temperature to pellet lipoproteins with attention to the orientation of the tube. Lipoproteins will form a thin, white film along the outside wall of the tube.

2.9) Wash pellet twice with 100% acetone. Decant acetone and allow sample to air dry.

2.10) Add 20-40 µL of 10 mM Tris-HCl, pH 8.0 or a standard Laemmli SDS-PAGE sample buffer[16] and thoroughly resuspend by pipetting up and down against the wall with the precipitated lipoproteins. Scrape the side of the tube with the pipette tip to dislodge lipoproteins. Lipoproteins will not dissolve in Tris buffer, but rather form a suspension. Store lipoproteins at -20 °C until use.

3. SDS-PAGE, Electroblotting, and Staining with Ponceau S

3.1) Separate lipoproteins by SDS-PAGE using standard methods[16]. The appropriate gel acrylamide percentage will vary depending on its intended use and size of lipoproteins. Here, *E. faecalis* lipoproteins were separated over a 10% Tris-glycine gel, while a 16.5% Tris-tricine gel was used for the smaller *E. coli* lipoprotein Lpp[17].

3.2) Transfer the lipoproteins to a nitrocellulose transfer membrane using a standard electroblotting procedure. A semi-dry transfer using Bjerrum Schafer-Nielsen buffer plus 0.1% SDS was performed at 25 V, 1.3 mA for 15 min[18]. Alternately, polyvinylidene difluoride (PVDF) membrane could be used, however as it is more hydrophobic than nitrocellulose, it may reduce efficient elution of hydrophobic lipopeptides from the membrane at later steps.

3.3) Transfer the nitrocellulose membrane to a container and cover with Ponceau S solution (0.2% (w/v) Ponceau S in 5% acetic acid). Rock gently for 5 min or until red-pink bands are visible.

3.4) Pour off Ponceau S solution and carefully rinse the nitrocellulose membrane with dH_2O to remove excess stain. **Note:** Ponceau-stained bands will destain rapidly. If bands disappear, repeat staining process.

3.5) With a clean razor blade, excise the desired band and transfer to a microcentrifuge tube. Wash three times with 1 mL of dH_2O to completely destain the band. **Note:** The protocol can be paused here. Store the excised section covered with water at -20 °C until use.

3.6) Transfer the section to a clean surface and with a clean razor blade, dice the nitrocellulose strip into small pieces of approximately 1 mm x 1 mm. Collect the pieces into a 0.5-mL low protein binding microcentrifuge tube.

3.7) Wash the pieces twice with 1 mL of freshly-prepared 50 mM ammonium bicarbonate (NH_4HCO_3), pH 7.8 in HPLC grade water.

4. Tryptic Digestion, Lipopeptide Extraction from Nitrocellulose Membrane, Deposition onto MALDI Target, and Data Acquisition

4.1) Resuspend nitrocellulose pieces in 20 μL of a 20 μg/mL solution of trypsin in 50 mM NH_4HCO_3, pH 7.8 in HPLC grade water. Vortex to mix, then spin briefly to ensure that all pieces are fully covered by the trypsin solution. Parafilm the tube lid to prevent evaporation and incubate digest overnight at 37 °C.

4.2) Spin the sample at 16,000 x g for 30 s and remove the liquid by pipetting. Note: This removes trypsinized hydrophilic peptides that can be examined later by MALDI-TOF MS for protein identification.

4.3) Add 50 μL of 0.5% trifluoroacetic acid (TFA) in HPLC grade water. Vortex to mix, then spin the sample briefly to ensure all pieces are covered by the solution. Incubate for 10 min at room temperature. Remove the liquid by pipetting. CAUTION: TFA is harmful if inhaled. Handle with caution and use in the chemical fume hood. Avoid contact with skin.

4.4) Repeat step 4.3 with 50 µL of 10% acetonitrile in HPLC grade water. CAUTION: Acetonitrile is harmful if inhaled. Handle with caution and use in the chemical fume hood. Avoid contact with skin.

4.5) Repeat step 4.3 with 50 µL of 20% acetonitrile in HPLC grade water. This removes any moderately hydrophobic peptides loosely bound to the nitrocellulose.

4.6) To elute the tightly-bound lipopeptides, add 15 µL of freshly-made 10 mg/mL α-cyano-4-hydroxycinnamic acid (CHCA) matrix dissolved in chloroform-methanol (2:1, v/v) and incubate for 10 min at room temperature with intermittent vortexing. Alternately, add 15 µL chloroform-methanol to elute lipopeptides and mix with matrix at a later time. Other matrices, such as 2,5-dihydroxybenzoic acid (DHB), should be tested as alternatives to CHCA if unsatisfactory results are obtained for a particular lipopeptide composition. CAUTION: Both chloroform and methanol are harmful if inhaled. Handle with caution and use in the chemical fume hood. Avoid contact with skin.

4.6.1) Optional: To promote sodium adduct formation, supplement the solution of CHCA in chloroform-methanol with aqueous sodium bicarbonate ($NaHCO_3$) to a final concentration of 1 mM.

4.7) Spin the sample briefly. With a pipette, carefully transfer the liquid to a new low protein binding microcentrifuge tube. This solution will contain the bulk of the N-terminal lipopeptides. Note: The nitrocellulose may partially or fully dissolve in the chloroform-methanol solution. This will not negatively affect MS results, and may be used as an alternate method to increase lipopeptide return. PVDF membrane will not dissolve in the same manner.

4.8) Deposit 1 μL of the eluted lipopeptides with CHCA onto a polished steel MALDI target. The chloroform-methanol will evaporate quickly, leaving behind crystallized CHCA and lipopeptides. An optional second 1 μL aliquot can be deposited onto the same spot to increase sample concentration. Chloroform-methanol may spread significantly after deposition onto the target and as such, care should be taken to avoid sample mixing on the target. It is recommended to deposit and shoot multiple spots of each sample.

4.9) Immediately proceed to mass spectrometry. Scan all areas of the spot for lipopeptide signal, especially if the spot contains two layers of sample, as the second layer may push the lipopeptides to the spot's outer rim. Recommended starting laser intensity is 25%, though it may be necessary to increase intensity to acquire signal and sum multiple spectra (20 or more scans) to achieve adequate signal-to-noise ratio. Here, spectra were acquired on a MALDI-TOF-TOF instrument (see Materials) using a factory-configured instrument method for the reflector positive-ion detection over the 700-3500 m/z range. The instrument was calibrated with a bovine serum albumin (BSA) tryptic peptide mixture.

Representative Results

A schematic of the protocol is provided in **Figure 3-1**. The lipoprotein-enriched fraction extracted from *Enterococcus faecalis* ATCC 19433 by Triton X-114 is shown in **Figure 3-2**. For comparison, the banding pattern of the precipitated protein fraction from step 2.6 is also shown. Proteins from this fraction were confirmed by MALDI-MS to be highly abundant contaminating proteins other than lipoproteins (**Table 3-1**). The mass spectra in **Figure 3-3** demonstrate the tryptic peptide ion profile of the *E. faecalis* lipoprotein PnrA that occurs with subsequent washes

of nitrocellulose-bound PnrA with solvents of increasing polarity (peak assignments listed in **Table 3-2**). **Figure 3-4** shows the *N*-terminal structural characterization of PnrA as determined by MALDI-TOF MS/MS, revealing diagnostic *N*-acylated dehydroalanyl peaks consistent with the lyso-lipoprotein form. **Figure 3-5** illustrates the effect of sodium adduct formation on fragmentation, with the sodiated parent ion preferentially fragmenting in favor of the *N*-acylated dehydroalanyl ion.

Figures and Tables

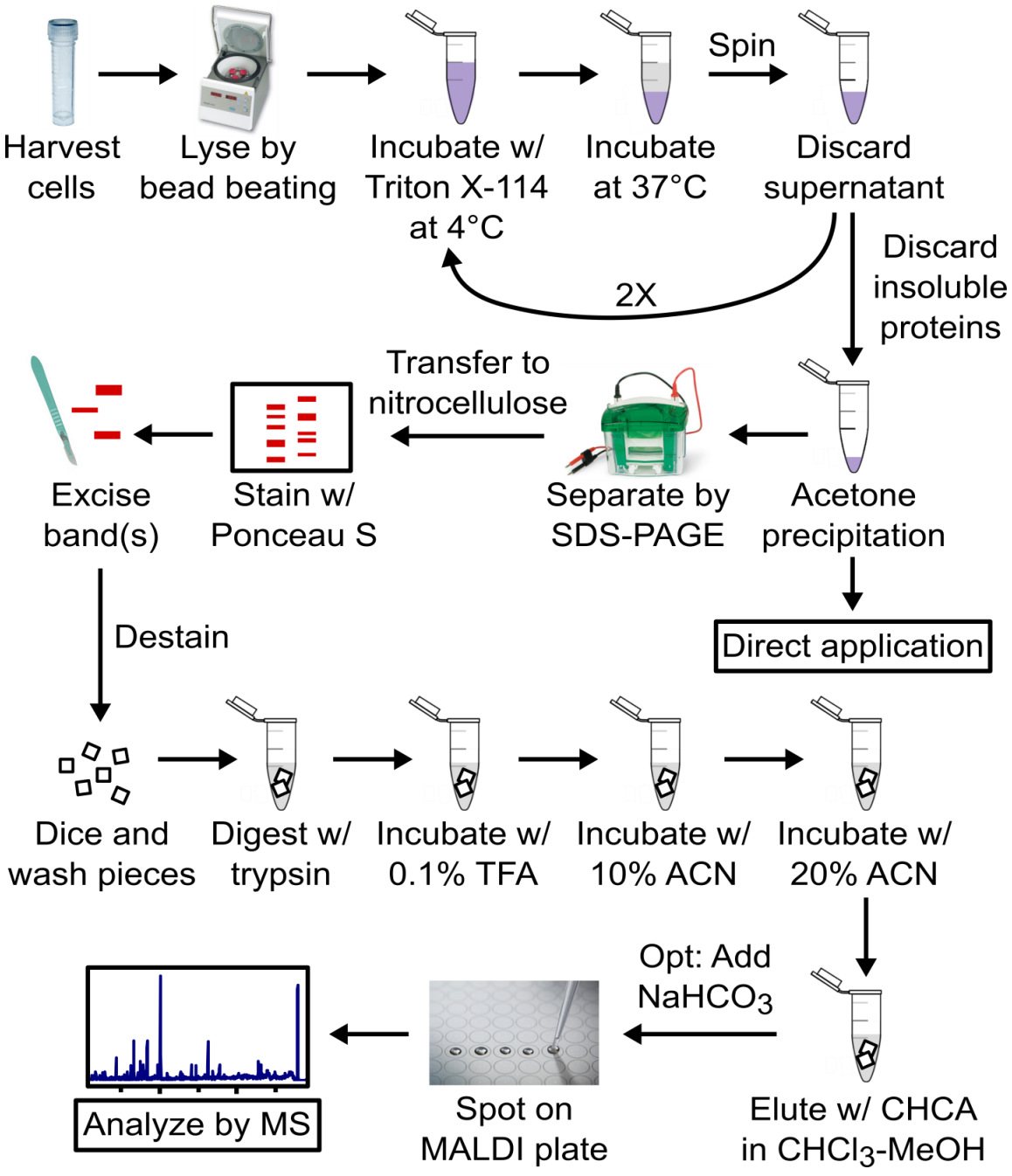


Figure **3-1**: Schematic of the protocol. Lipoproteins are enriched by Triton X-114 phase partitioning and can be used directly, most commonly in TLR assays, or further purified by SDS-PAGE for structural determination. Lipoproteins are transferred to nitrocellulose, digested with trypsin, washed stepwise, and the resulting lipopeptides eluted with chloroform-methanol for structural analysis by MALDI-MS. The trypsin and nitrocellulose wash solutions may be saved for protein identification and MS analysis. w/: with; TFA: trifluoroacetic acid; ACN: acetonitrile;

CHCA: α-cyano-4-hydroxycinnamic acid; opt: optional; MALDI: matrix-assisted laser desorption ionization; MS: mass spectrometry.

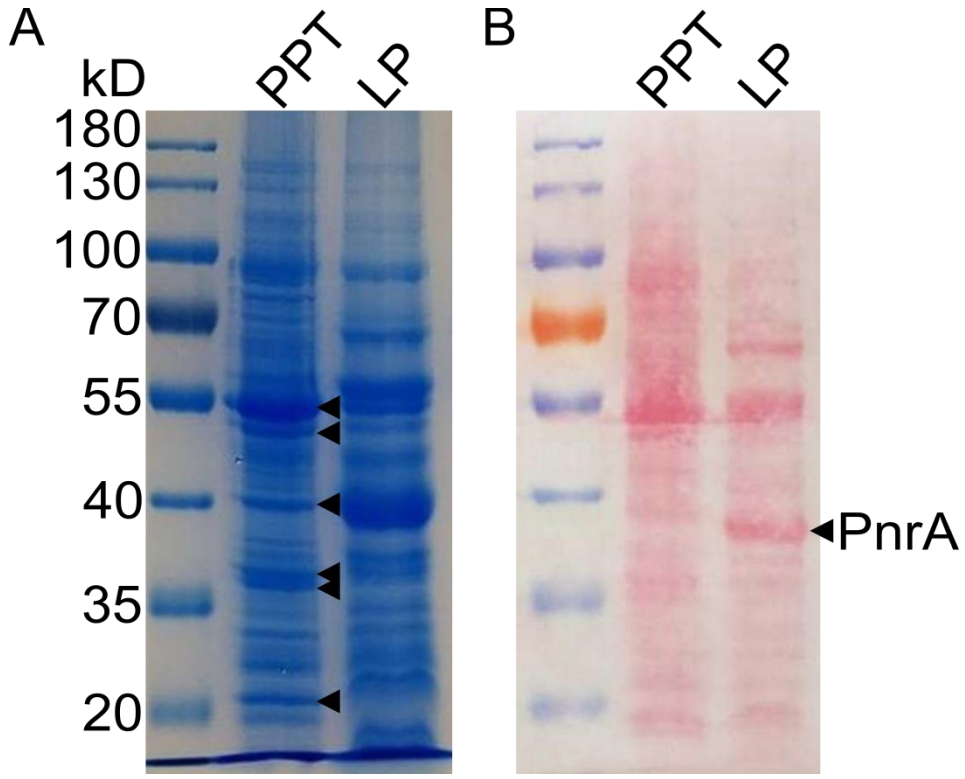


Figure **3-2**: Profile of Triton X-114 enriched proteins from *E. faecalis*. A 10% Tris-glycine SDS-PAGE gel stained with Coomassie blue (A) and the corresponding nitrocellulose membrane stained with Ponceau S (B) reveal a different banding pattern between the precipitated protein fraction ("PPT") and the purified lipoproteins ("LP"). The indicated Coomassie-stained bands were excised and identified as non-lipoproteins by MALDI-MS of tryptic peptides (see Table 3-1). PnrA was identified and analyzed by the protocol described herein.

Table **3-1**: Precipitated proteins are non-lipoprotein contaminants. Proteins were identified by tryptic digest and MALDI-MS using standard in gel digestion protocols[19] and are listed from top to bottom in the same order they appear on the Coomassie gel in Figure 3-2. Each protein was identified with a Confidence Interval (C.I.) greater than 95%.

Band	Protein ID	Accession No.	Est. Molecular Weight	Peptide Count
1	Translation elongation factor Tu	gb|EOL37301.1|	43,387	15
2	NADH peroxidase	gb|EOL34572.1|	49,520	9
3	pyruvate dehydrogenase E1 component, alpha subunit	gb|EOL34709.1|	41,358	12
4	pyruvate dehydrogenase E1 component, beta subunit	gb|EOL34710.1|	35,373	19
5	30S ribosomal protein S2	gb|EOL33066.1|	29,444	16
6	30S ribosomal protein S3	gb|EOL37312.1|	24,355	14

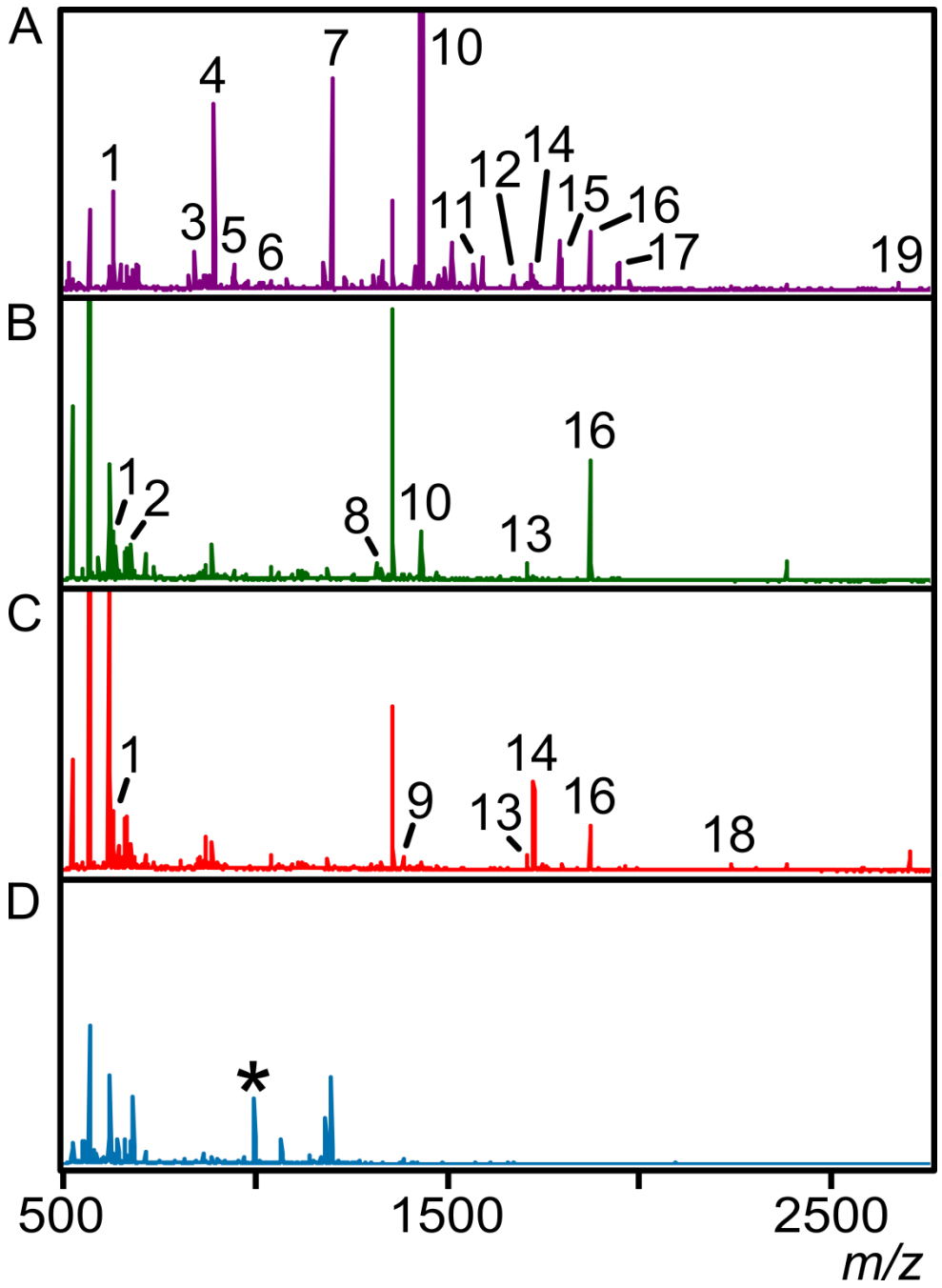


Figure **3-3**: Ion profile and abundance changes with polarity of nitrocellulose wash solutions. (A) The trypsin fraction shows several internal peptide fragments corresponding to the *E. faecalis* lipoprotein PnrA indicated in Figure 3-2. The 10% (B) and 20% (C) acetonitrile wash fractions show specific fractionation of the peptide sample. (D) The final elution fraction is highly enriched with the N-terminal lipopeptide at *m/*z 997, indicated by an asterisk (*). Intensity of each spectra is normalized to (A) for comparison of signal intensity. Peaks masses and assigned sequences are listed in Table 3-2.

Table **3-2**: Observed masses and the corresponding tryptic peptides. *In silico* trypsin digestion of *E. faecalis* PnrA (GenBank: EOL37280.1) was performed using PeptideMass[20,21], allowing up to two missed cut sites. The theoretical masses of the peaks observed on the spectra in Figure 3-3 are listed, along with the corresponding peptide sequences.

Peak Number	Theoretical Mass	Protein Position	No. of Missed Cut Sites	Peptide Sequence
1	631.3409	170-176	0	GVADAAK
2	675.4035	316-321	1	TKEAVK
3	826.4093	163-169	0	FQAGFEK
4	893.4839	121-128	0	NVVSATFR
5	943.5359	240-247	0	VWVIGVDR
6	1040.5047	188-197	0	YAASFADPAK
7	1200.6331	273-284	0	GVGTAVQDIANR
8	1316.6997	290-301	0	FPGGEHLVYGLK
9	1385.7827	83-94	0	FNTIFGIGYLLK
10	1432.743	149-162	0	VGFVGGEEGVVIDR
11	1568.7802	259-272	1	DGKEDNFTLTSTLK
12	1672.8289	240-254	1	VWVIGVDRDQDADGK
13	1707.8184	129-145	0	DNEAAYLAGVAAANETK
14	1724.8027	35-49	0	SFNQSSWEGLQAWGK
15	1797.9228	257-272	2	TKDGKEDNFTLTSTLK
16	1872.9854	285-301	1	ALEDKFPGGEHLVYGLK
17	1945.9613	230-247	1	DLNESGSGDKVWVIGVDR
18	2240.1345	149-169	1	VGFVGGEEGVVIDRFQAGFEK
19	2675.2543	230-254	2	DLNESGSGDKVWVIGVDRDQDADGK

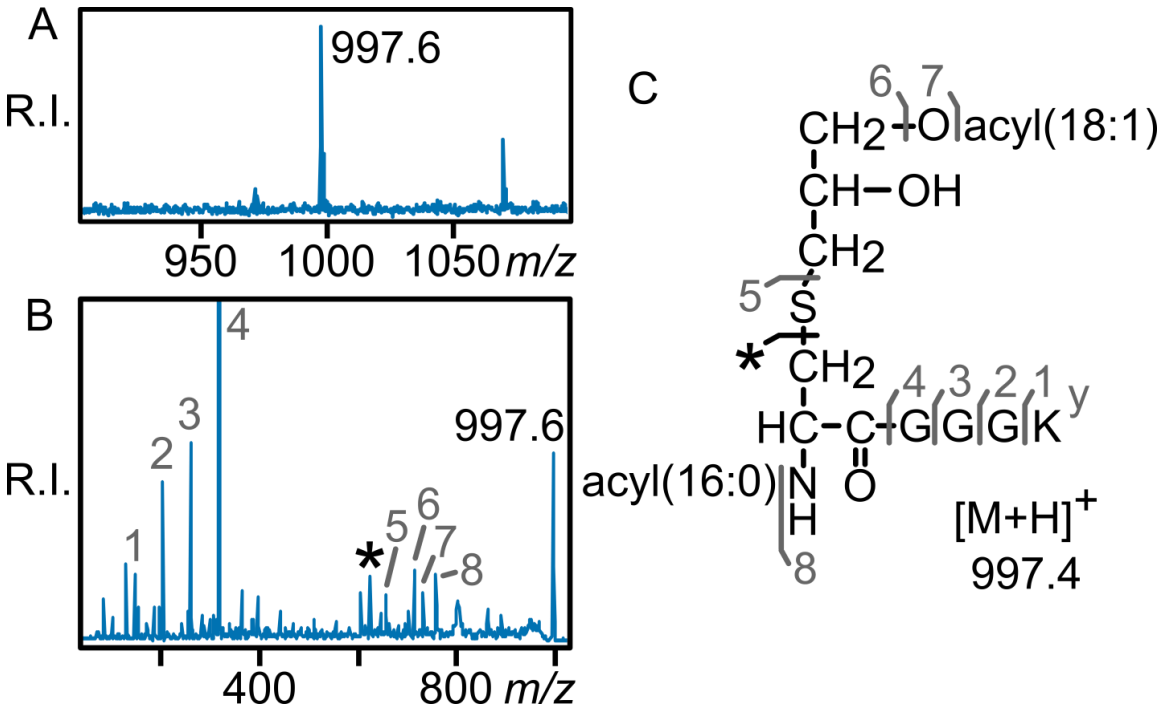


Figure **3-4**: MALDI-TOF MS of the *E. faecalis* lipoprotein PnrA. (A) Parent MS spectrum of the *m/z* 997 region corresponding to the N-terminal lipopeptide of *E. faecalis* PnrA. (B) MS/MS of the lipopeptide peak reveals that it is the lyso form, with the diagnostic *N*-acylated dehydroalanyl peptide fragment ion peak indicated by an asterisk (*). The elucidated structure is shown (C). This figure has been modified from a previous publication[9]. R.I.: relative intensity

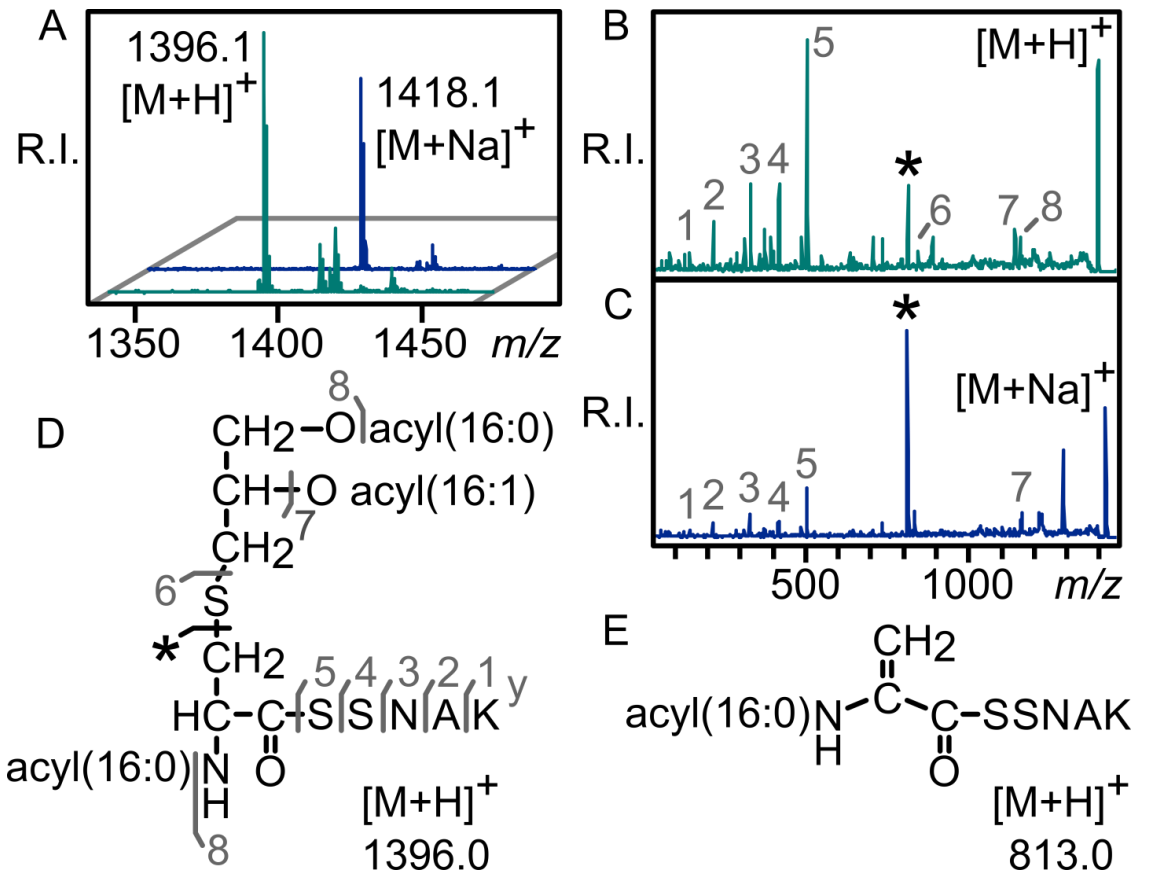


Figure **3-5**: Sodium adduct formation promotes parent ion fragmentation towards dehydroalanyl lipopeptide ions. (A) MALDI-TOF spectra of the *E. coli* lipoprotein Lpp N-terminal peptide without (turquoise traces) and with (blue traces) the addition of sodium bicarbonate. Addition of sodium bicarbonate to the final eluted fraction results in a 22 Da increase from the calculated mass of the parent lipopeptide. When compared with the MS/MS spectrum of the protonated ion (B), the MS/MS spectrum of the corresponding sodiated ion (C) shows significant preferential fragmentation toward the *N*-acyl-dehydroalanyl ion through neutral elimination of the diacylthioglyceryl moiety, indicated by an asterisk (*). The structure of the parent triacylated Lpp N-terminal tryptic peptide (D) and the *N*-acyl-dehydroalanyl peptide fragment ion (E) are depicted. This figure has been modified from a previous publication[9]. R.I.: relative intensity

Discussion

The protocol herein describes two distinct stages of lipoprotein characterization: enrichment by Triton X-114 phase partitioning and structural determination by MALDI-TOF MS. During Triton X-114 extraction, additional centrifugation removes contaminating proteins that precipitate during this process, followed by acetone precipitation to yield highly enriched lipoproteins. By limiting the scale of each preparation to 15-mL worth of cells, several samples can be easily processed in parallel, and if desired, pooled at the end of the protocol.

For structural analysis by MS, transfer of the lipoproteins to nitrocellulose facilitates trypsin digestion, washing, and subsequent elution from the membrane. In our hands, this approach has proven to be preferable to detergent-mediated extraction from traditional polyacrylamide gel plugs, as the lipoproteins are associated with the surface of the nitrocellulose membrane and not embedded in polyacrylamide. Lipoprotein association with the nitrocellulose surface also largely prevents the common problem of nonspecific adsorption to container surfaces by hydrophobic peptides. Stepwise elution of the nitrocellulose pieces removes more hydrophilic, internal tryptic peptides which can be analyzed by MALDI-TOF MS to achieve high confidence protein identification assignments. The final elution step reproducibly yields concentrated lipopeptides in a small volume that is largely free of interfering salts and ion suppressing contaminants.

Successful MS analysis of a lipoprotein is subject to its abundance, and as such, the darkest bands from the Ponceau S-stained membrane are likely to give the best results. However, it is possible that several lipoproteins may migrate in a single band during PAGE separation, complicating the resulting spectra. Therefore, it is highly recommended to first identify the protein by peptide mass fingerprinting (PMF), which can be accomplished by collecting MS spectra on the total trypsin fraction or either acetonitrile wash fraction and inputting the resulting

peaks into a software such as Mascot[22]. Once the protein sequence has been determined, calculating the mass of the tryptic N-terminal lipopeptide significantly aids in choosing which ion to fragment by MS/MS.

Additional sample heterogeneity can result from varying acyl chain lengths on lipoproteins within a population and their N-terminal modifications, both depending on the source bacteria[6]. While variation in chain length makes for distinctive peak clusters on parent spectra, characterized by 14 Da increments corresponding to methylene groups (-CH_2-), it can split the overall lipopeptide signal and decrease sensitivity. It is also likely that the different lipoprotein forms influence enrichment and fragmentation, though we have successfully identified triacylated, diacylated, and lyso-form lipoproteins using the described protocol. Similarly, the varying peptide moieties of lipoproteins add another level of complexity to analyses, regardless of their N-terminal structure, as a very hydrophilic amino acid composition may prevent lipopeptide partitioning into the organic chloroform phase using other extraction methods[8]. Transfer of the lipoprotein to nitrocellulose would assist in studying such lipopeptides, as all elution fractions can be analyzed in this method.

Drawing from previous literature involving triacylglycerides and phospholipids[23], intentional sodium adduct formation can be used to promote more informative fragmentation via formation of the (*N*-acyl)-dehydroalanyl peptide ion. These characteristic ions are key to assigning structure, since the acylation state of the N-terminus is isolated from the acyl substitutions on the glyceryl moiety. Sodium adduct formation provides a convenient alternate option to sulfur oxidation with hydrogen peroxide, which has been shown to also promote *N*-acyl-dehydroalanyl peptide fragmentation[6,8]. Although sodium adducts may show an overall suppression of the parent ion signal, the specific increase in fragmentation towards dehydroalanyl ions compensates for diminished parent ion intensity. It may be worth exploring other matrices and fragmentation of additional adducts to elicit the most informative spectra.

This protocol improves on the traditional Triton X-114 phase partitioning method for lipoprotein enrichment, while transfer of PAGE-separated lipoproteins to nitrocellulose membrane facilitates trypsin digestion and elution of lipopeptides. MALDI-TOF MS analysis on the total tryptic or wash fractions allows for protein identification in tandem with N-terminal structural determination by MS/MS in a single experiment. Optional sodium adduct formation highlights differences at the lipopeptide N-terminus by promoting more informative ion fragmentation. The ability to routinely purify lipoproteins and characterize their N-termini will enable extensive studies on how novel lipoprotein forms are made, their physiological role within the bacterial cell envelope, and how they are detected by the mammalian immune system.

References

1. Babu, M. M. *et al.* A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *J Bacteriol*. **188**(8), 2761–2773 (2006).
2. Narita, S., Tokuda, H. Bacterial lipoproteins; biogenesis, sorting and quality control. *Biochimica et Biophysica Acta*. **1862**(11), 1414-1423 (2016).
3. Kovacs-Simon, A., Titball, R. W., Michell, S. L. Lipoproteins of bacterial pathogens. *Infect Immun.* **79**(2), 548–61 (2011).
4. Nguyen, M. T., Götz, F. Lipoproteins of Gram-Positive Bacteria: Key Players in the Immune Response and Virulence. *Microbiol Mol Biol Rev*. **80**(3), 891–903 (2016).
5. Nakayama, H., Kurokawa, K., Lee, B. L. Lipoproteins in bacteria: structures and biosynthetic pathways. *FEBS J.* **279**(23), 4247–4268 (2012).

6. Kurokawa, K. *et al.* Novel bacterial lipoprotein structures conserved in low-GC content Gram-positive bacteria are recognized by Toll-like receptor 2. *J Biol Chem.* **287**(16), 13170–13181 (2012).
7. Kurokawa, K. *et al.* The Triacylated ATP Binding Cluster Transporter Substrate-binding Lipoprotein of *Staphylococcus aureus* Functions as a Native Ligand for Toll-like Receptor 2. *J Biol Chem.* **284**(13), 8406–8411 (2009).
8. Asanuma, M. *et al.* Structural evidence of α-aminoacylated lipoproteins of *Staphylococcus aureus*. *FEBS J.* **278**(5), 716–728 (2011).
9. Armbruster, K. M., Meredith, T. C. Identification of the Lyso-Form N-Acyl Intramolecular Transferase in Low-GC Firmicutes. *J Bacteriol* **199**(11), e00099-17 (2017).
10. Serebryakova, M. V, Demina, I. A., Galyamina, M. A., Kondratov, I. G., Ladygina, V. G., Govorun, V. M. The acylation state of surface lipoproteins of Mollicute *Acholeplasma laidlawii*. *J Biol Chem.* **286**(26), 22769–22776 (2011).
11. Feng, S. H., Lo, S. C. Induced mouse spleen B-cell proliferation and secretion of immunoglobulin by lipid-associated membrane proteins of *Mycoplasma fermentans* incognitus and *Mycoplasma penetrans. Infect Immun.* **62**(9), 3916–3921 (1994).
12. Feng, S. H., Lo, S. C. Lipid extract of *Mycoplasma penetrans* proteinase K-digested lipid-associated membrane proteins rapidly activates NF-kappaB and activator protein 1. *Infect Immun.* **67**(6), 2951–2956 (1999).
13. Luque-Garcia, J. L., Zhou, G., Spellman, D. S., Sun, T. T., Neubert, T. A. Analysis of electroblotted proteins by mass spectrometry: protein identification after Western blotting. *Mol Cell Proteomics.* **7**(2), 308-314 (2008).
14. Luque-Garcia, J. L., Zhou, G., Sun, T.T., Neubert, T. A. Use of Nitrocellulose Membranes for Protein Characterization by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *Anal Chem.* **78**(14), 5102–5108 (2006).

15. Kurokawa, K. *et al.* Environment-mediated accumulation of diacyl lipoproteins over their triacyl counterparts in *Staphylococcus aureus*. *J Bacteriol* **194**(13), 3299–3306 (2012).
16. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* **227**(5259), 680–685 (1970).
17. Schägger, H. Tricine-SDS-PAGE. *Nat Protoc.* **1**, 16–22 (2006).
18. Gravel, P. Protein Blotting by the Semidry Method. *The Protein Protocols Handbook.* 321–334 Springer Protocols. (2002).
19. Webster, J., Oxley, D. Protein Identification by Peptide Mass Fingerprinting using MALDI-TOF Mass Spectrometry. *The Protein Protocols Handbook*. 1117–1129 Springer Protocols. (2009).
20. Wilkins, M. R. *et al.* Detailed peptide characterization using PEPTIDEMASS - a World-Wide-Web-accessible tool. *Electrophoresis* **18**(3–4), 403–408 (1997).
21. Gasteiger, E. *et al.* Protein Identification and Analysis Tools on the ExPASy Server. *The Proteomics Protocols Handbook* , 571–607 Springer Protocols (2005).
22. Perkins, D. N., Pappin, D. J. C., Creasy, D. M., Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**(18), 3551–3567 (1999).
23. Al-Saad, K. A., Zabrouskov, V., Siems, W. F., Knowles, N. R., Hannan, R. M., Hill, H. H. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of lipids: ionization and prompt fragmentation patterns. *Rapid Commun Mass Spectrom.* **17**(1), 87–96 (2003).

Chapter 4

Copper-induced expression of a transmissible lipoprotein intramolecular transacylase alters lipoprotein acylation and the Toll-like receptor 2 response to *Listeria monocytogenes*

Adapted from:

Armbruster KM, Komazin G, Meredith TC. 2019. Copper-induced expression of a transmissible lipoprotein intramolecular transacylase alters lipoprotein acylation and the Toll-like receptor 2 response to *Listeria monocytogenes*. J Bacteriol. 201:e00195-19.

Journal of Bacteriology Spotlight: Copper induced remodeling of lipoproteins alters the Toll-like receptor 2 response

Abstract

Bacterial lipoproteins are globular proteins anchored to the extracytoplasmic surfaces of cell membranes through lipidation at a conserved *N*-terminal cysteine. Lipoproteins contribute to an array of important cellular functions for bacteria, as well as being a focal point for innate immune system recognition through binding to Toll-like receptor 2 (TLR2) heterodimer complexes. Although lipoproteins are conserved among nearly all classes of bacteria, the presence and type of α-amino-linked acyl chain is highly variable and even strain specific within a given bacterial species. In *Enterococcus faecalis*, lipoproteins are anchored by an *N*-acyl-*S*-monoacyl-glyceryl cysteine (lyso-form) moiety installed by a chromosomally-encoded lipoprotein intramolecular transacylase (Lit). The reason for lyso-lipoprotein formation and *N*-acylation variability in general is presently not fully understood. Herein, we describe a mobile genetic element common to environmental isolates of *Listeria monocytogenes* and *Enterococcus* spp. encoding a functional Lit ortholog (Lit2) that is cotranscribed with several well established copper resistance determinants. Expression of Lit2 is tightly regulated, and induction by copper converts lipoproteins from the diacylglycerol-modified form characteristic of *L. monocytogenes* type strains to the α-amino-modified lyso form observed in *E. faecalis*. Conversion to the lyso form through either copper addition to media or constitutive expression of *lit2* decreases TLR2 recognition when using an activated NF-κB secreted embryonic alkaline phosphatase reporter assay. While lyso formation significantly diminishes TLR2 recognition, lyso-modified lipoprotein is still predominantly recognized by the TLR2/6 heterodimer.

Importance: The induction of lipoprotein *N*-terminal remodeling in response to environmental copper in Gram-positive bacteria suggests a more general role in bacterial cell envelope physiology. *N*-terminal modification by lyso formation in particular simultaneously modulates the TLR2 response in direct comparison to their diacylglycerol-modified precursors.

Thus, use of copper as a frontline antimicrobial control agent and ensuing selection raises the potential of diminished innate immune sensing and enhanced bacterial virulence.

Introduction

Lipoproteins are highly conserved bacterial cell surface-bound proteins, and consequently are a focal recognition point for innate immune pathways involved in microbial pathogen recognition (1–4). Lipoproteins are structurally unique in having post-translational acylation at a conserved *N*-terminal cysteine that anchors variable globular protein domains to the cell membrane surface where they perform a myriad of functions (2, 3, 5–7). Preprolipoproteins are first exported across the cytoplasmic membrane, after which they become modified by lipoprotein diacylglyceryl transferase (Lgt). Lgt recognizes a short, conserved amino acid sequence called a lipobox containing an invariant cysteine residue that is modified by transferring a diacylglycerol (DA) moiety from a membrane phospholipid to the cysteine thiol to form a thioether (8, 9). Then, lipoprotein signal peptidase II (Lsp) cleaves the signal peptide, exposing the cysteine α-amino group (10). Beyond this step, further lipoprotein modification varies between species and in a way that does not simply correspond to Gram-negative versus Gram-positive bacteria as had previously been accepted (6, 11–14).

In most Gram-negative bacteria, acylation of the α-amino cysteine terminus is catalyzed by the essential lipoprotein *N*-acyl transferase (Lnt) using a phospholipid acyl donor to create the mature triacylated lipoprotein (TA-LP) (15). This *N*-acyl chain is critical for lipoprotein recognition and transport to the outer membrane (16, 17). However, an *lnt* ortholog and thus TA-LP is also present in many high-GC Gram-positive Actinobacteria (6, 13, 18). More surprisingly, TA-LP are found in some low-GC Gram-positive Firmicutes (*Staphylococcus aureus*, *Staphylococcus epidermidis*) having no identifiable *lnt* sequence ortholog (11, 19, 20). Further

studies by Kurokawa et al. among a panel of Firmicutes revealed several novel lipoprotein forms, all featuring novel *N*-terminal modifications: the *N*-acetyl form, having an amide-linked acetyl group (*Bacillus subtilis*, among others); the lyso form (lyso-LP), with an *N*-acyl *S*-monoacyl-glyceryl (*Enterococcus faecalis, Bacillus cereus*, and others); and a peptidyl form, containing two additional amino acids before the lipid-modified cysteine (*Mycoplasma fermentans*) (11). More recently, Nguyen et al. identified the *N*-acetyl form in *Staphylococcus carnosus* as well (21). The widespread phylogenetic conservation of *N*-terminally modified lipoprotein forms, when considered with the seemingly random *intra*-species lipoprotein type, suggests a strong selective pressure favoring *N*-terminal modifications that appeared post-speciation. The underlying physiological purpose of *N*-modification though remains unclear, especially as some Gram-positive species continue to elaborate unmodified DA-LP (diacylglycerol-modified lipoprotein), as in *Listeria monocytogenes* (11). In either case, however, lipoprotein *N*-terminal structural diversity is accommodated by the innate immune response through cognate Toll-like receptor 2 (TLR2) heterodimerization (TLR2/1 or TLR2/6 for TA-LP and DA-LP ligands, respectively) (4, 21–23).

Capitalizing on the shared *N*-acyl chain of TA-LP and lyso-LP as means for rescue of the otherwise lethal *lnt*-null phenotype in *Escherichia coli*, we identified the previously uncharacterized gene lipoprotein intramolecular transacylase (*lit*) as responsible for the synthesis of lyso-LP in Firmicutes (12). Homology searches among type strains revealed a relatively narrow phylogenomic distribution in comparison to the highly conserved Lgt and Lsp proteins. Herein, however, we report the surprising distribution of a distinct Lit-type protein (named Lit2) encoded on a mobile genetic element embedded within a copper resistance operon. We investigate the effect of Lit2 on lipoprotein formation and subsequent TLR2 signaling using two environmental isolates, *Enterococcus faecalis* TX1342 and *Listeria monocytogenes* CFSAN023459. We show that Lit2 expression is induced by copper, in turn converting the

lipoprotein profile from DA-LP to lyso form in *L. monocytogenes*. We also compare the TLR2 response of the lyso-LP form, using both synthetic lipopeptide standards and heat-inactivated whole bacterial cells, and demonstrate that lyso-LP signals preferentially through the TLR2/6 heterodimer. Overall, however, the lyso form is shown to be a markedly weaker ligand than either conventional DA-LP or TA-LP forms. Taken together, the connection between copper resistance, lyso-form lipoproteins, and altered TLR2 innate immune recognition suggests a greater role of *N*-terminal lipoprotein modification in Gram-positive bacteria in copper resistance and potentially virulence through muted TLR2 recognition.

Results

Phylogenetic analyses reveal a *lit* sequence ortholog co-localized within a putative copper resistance operon

When the *E. faecalis* ATCC 19433 Lit protein sequence is queried using a BLASTp search, Lit orthologs bifurcate into two distinct clades (Figure 4-1A). In the first clade are species known to produce lyso-LP (11), as well as clinically relevant organisms such as *E. faecium*, *B. anthracis*, and *S. pneumoniae*. All genes are chromosomally located, though there is no shared cross-genera genomic synteny. This is in stark contrast to the second clade of Lit orthologs, where all are flanked by highly similar genes within a larger putative mobile element, on either a freely-replicating plasmid (*L. monocytogenes*) or integrated into the chromosome (*Enterococcus* spp.). In *Enterococcus* spp. isolates, these strains therefore contain two copies of *lit*, while isolates of *L. monocytogenes* harboring this plasmid contain a single copy. To differentiate between the two, we will hereby refer to the chromosomal copy as "*lit1*" and the accessory genome copy as "*lit2*".

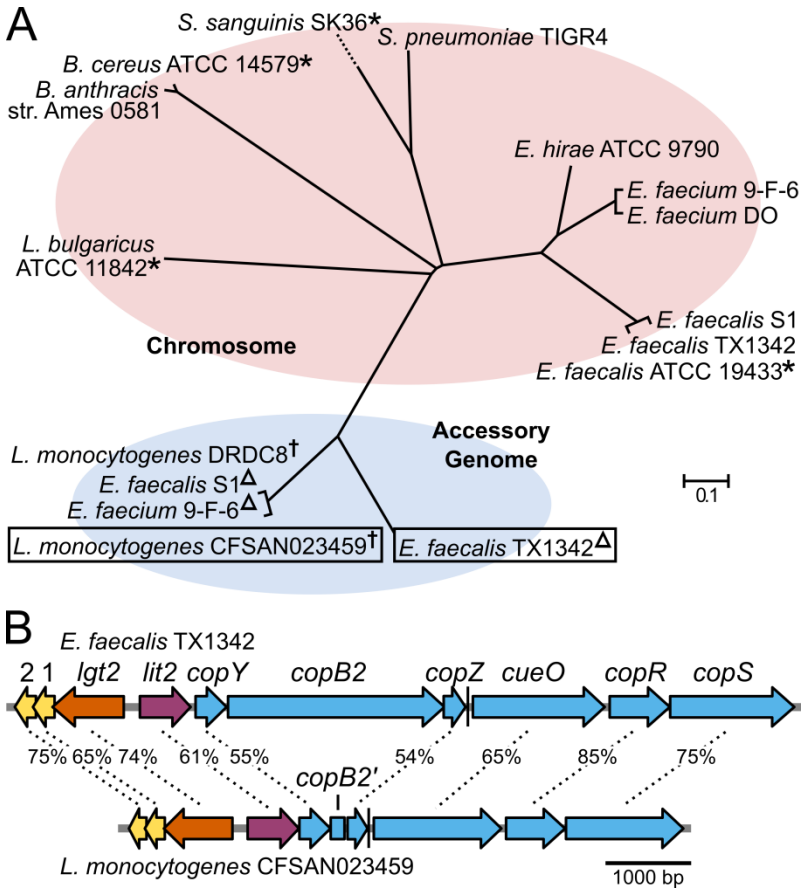


Figure **4-1**: Phylogenomic distribution of Lit. (A) A non-exhaustive phylogenetic tree was generated using the amino acid sequence of Lit from select strains. Sequences of Lit1 and Lit2 separate into two clades based on their location in the chromosome or the accessory genome, respectively. Strains previously demonstrated to make lyso-form lipoproteins are noted with an asterisk (11, 12). Other strains are assumed to make the lyso form based on the presence of *lit1*. The strains *E. faecalis* TX1342 and *L. monocytogenes* CFSAN023459 characterized in this study are boxed. Plasmid-borne copies of *lit2* are denoted with a dagger (†) and chromosomally-integrated copies with a triangle. (B) Schematic representations of the *E. faecalis* TX1342 and *L. monocytogenes* CFSAN023459 *lit2*-copper resistance operons are shown, with ortholog sequence identity indicated. The proteins of unknown function are designated "1" and "2" for clarity, and the incomplete CopB2 fragment as *copB2'*. A predicted intrinsic terminator following *copZ* is indicated by a vertical line.

The overall genetic architecture and gene similarity between *E. faecalis* TX1342 (*Ef*TX1342) and *L. monocytogenes* CFSAN023459 (*Lm*CFSAN) indicates a common origin and transmission potential between *Enterococcus* and *Listeria* sp (Figure 4-1B). Intriguingly, *lit2* appears to be the first gene in a polycistronic operon with well characterized copper resistance

determinants (24, 25). CopB is a P-type ATPase that exports copper across the cell membrane, thereby reducing intracellular copper concentrations (26). While *Lm*CFSAN does not encode an intact *copB*, a *copB'* fragment is present, suggesting a gene deletion event. The metallochaperone CopZ binds cytoplasmic copper and shuttles it to CopB to be extruded from the cell, or to CopY, a copper-responsive transcriptional regulator that derepresses expression of target copper resistance-related genes (27–29). These genes are followed by CueO, a multi-copper oxidase that helps detoxify copper by oxidizing Cu^{+} to Cu^{2+} (30), and a two-component signal transduction system CopRS that senses extracellular copper and induces expression of copper resistance genes (31–33). Together, these proteins maintain copper homeostasis and provide resistance to elevated levels of copper. Divergently transcribed from the *lit2*-copper resistance operon is a second copy of another lipoprotein biosynthetic gene *lgt*, hereby known as "*lgt2*", again to differentiate from the chromosomal copy *"lgt1*".

Lit2 is a lipoprotein intramolecular transacylase

While the co-localization of *lgt2* with *lit2* strongly suggests Lit2 is indeed a lipoprotein intramolecular transacylase, the Lit2 orthologs are more similar to each other than to the experimentally characterized *E. faecalis* ATCC 19433 Lit1 (Figure 4-1A). To confirm, a P1*vir* cotransduction linkage analysis assay was performed to assess whether *lit2* could functionally replace the lipoprotein *N*-acyltransferase *lnt* in *E. coli*. If functional, lyso-LP formation will allow proper lipoprotein trafficking to the outer membrane (12). Donor lysate of the *E. coli* strain TXM541 (*lnt*::Spt^{r} *chiQ*::Apr^{r}) was transduced into recipient strains harboring experimental plasmids pUC19, p*lnt*, p*Eflit2*, or p*Lmlit2*. Apramycin-resistant (Apr^{r}) colonies mark successful transduction events, while co-transduction of spectinomycin (Spt^{r}) demonstrates that *lnt* can be functionally replaced (Figure 4-2A). While *lnt* cannot be deleted from the control pUC19 strain,

the *lnt*::Spt^r allele can be established at frequencies comparable to the *lnt*-expressing positive control when either *Ef*TX1342 *lit2* or *Lm*CFSAN *lit2* is present (12). This confirms Lit2 can functionally substitute for *lnt*.

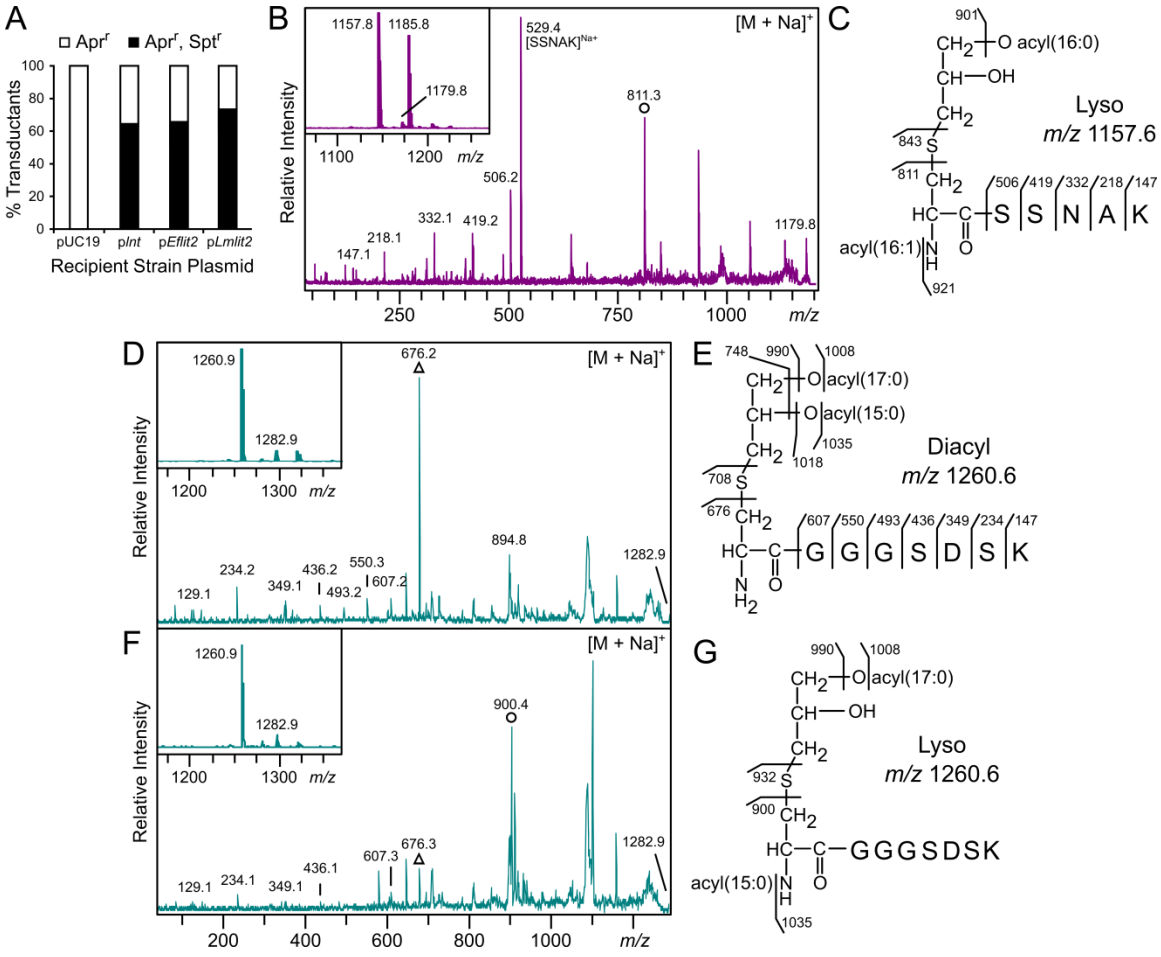


Figure **4-2**: Lit2 is a functional lipoprotein intramolecular transacylase. (A) The linked *lnt*::Spt^r and *chiQ*::Apr^r markers were co-transduced by P1*vir* into recipient strains carrying pUC19, p*lnt*, p*Eflit2*, or p*Lmlit2*. The percentage of transductants resistant to apramycin alone versus both apramycin and spectinomycin are shown (n = 24 – 40) and compared to negative (pUC19) and positive (p*lnt*) controls (12). (B) Trypsinized *N*-terminal lipopeptides of Lpp from the *lnt*-null strain KA811 expressing *Eflit2* were analyzed by MALDI-TOF MS. The protonated *m/z* 1157.8 and sodiated *m/z* 1179.8 parent ions are shown (B; inset), with the latter further fragmented by tandem MS/MS (B). This spectrum was used to assign Lpp structure from KA811 as the lyso form (C). Note the *m/z* 1185.8 peak in the parent spectrum is consistent with a $C_{16:0}$, $C_{18:1}$ acyl chain combination. (D, F) Trypsinized *N*-terminal lipopeptides of the lipoprotein KO07_11695 from wildtype *Listeria monocytogenes* ATCC 19115 type (D) and the corresponding *Lmlit2* expressing strain KA849 (F) were analyzed by MALDI-TOF MS. The parent spectra (insets of D

and F) display the protonated *m/z* 1260.9 and sodiated *m/z* 1282.9 parent ions. The sodiated ions were fragmented by tandem MS/MS (D and F), revealing production of DA-LP in KA847 (E) and the lyso-LP when *Lmlit2* is expressed (G). The structurally diagnostic dehydroalanyl ions for the diacylglycerol-modified (triangle) and the lyso (circle) lipopeptide are indicated.

To determine lipoprotein form, Lpp(K58A) was analyzed from the *E. coli* strain KA811 (*lnt*::Spt^r + p*Eflit2*) by MALDI-TOF MS. Similar to when Lpp is modified by Lit1 (12), the parent spectrum revealed ions at *m/z* 1157.8 and 1179.8, consistent with the predicted mass of the *N*-terminal CSSNAK tryptic lipopeptide possessing two acyl chains ($C_{16:0}$ and $C_{16:1}$) and ionizing as protonated and sodium adducts, respectively (Figure 4-2B, inset). Fragmentation of the sodium adduct by tandem MS/MS (Figure 4-2B), which fragments preferentially toward the dehydroalanyl peptide (Figure 4-3A) (34), exhibited Lpp's y-series ions and a prominent ion at *m/z* 811.3. This ion is diagnostic of a monounsaturated $C_{16:1}$ *N*-acyl chain on the lipopeptide (Figure 4-2C), demonstrating that Lpp is converted to the lyso form when *Ef*TX1342 *lit2* is expressed in *E. coli*.

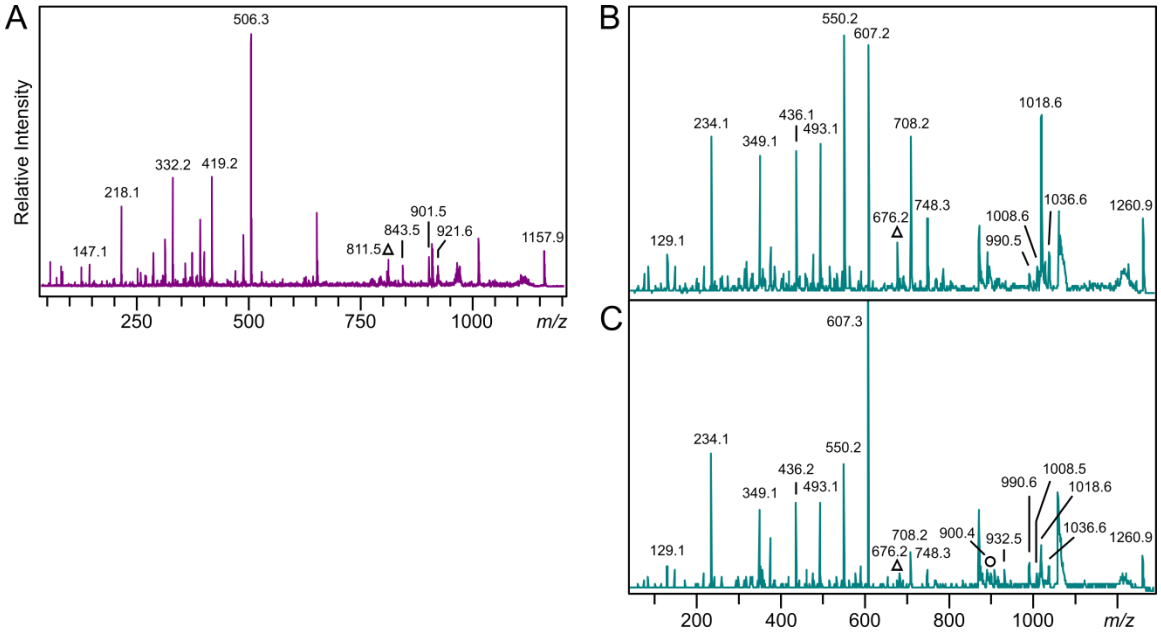


Figure **4-3**: MALDI-TOF MS/MS spectrum of the protonated parent ions *m/z* 1157.9 and 1260.9. (A) MS/MS spectrum of the protonated parent ion *m/z* 1157.9 of Lpp purified from strain KA811 expressing *E. faecalis* TX1342 *lit2*. MS/MS spectra of the protonated parent ion *m/z* 1260.9 of the *L. monocytogenes* lipoprotein KO07_11695 from strain KA847 (ATCC 19115) (B), and KA849 (ATCC 19115 + pP_{xyl}*Lmlit2*) plus 2% xylose (C). The diagnostic dehydroalanyl ions for the diacylglycerol-modified (triangle) and the lyso (circle) lipopeptide are indicated.

To show that *Lm*CFSAN Lit2 similarly functions as an intramolecular transacylase, *Lmlit2* was cloned under a xylose-inducible promoter into *Listeria monocytogenes* ATCC 19115, a type strain of *L. monocytogenes* that lacks any *lit* sequence orthologs. Lipoproteins were extracted from cultures grown with 2% xylose and wildtype ATCC 19115 cells, then separated by SDS-PAGE (Figure 4-4). A predicted peptide ABC transporter substrate-binding lipoprotein, KO07_11695, previously studied by Kurokawa et al. (11), was chosen for further analysis by MALDI-TOF MS. The parent spectra of both strains, regardless of *Lmlit2* expression, contained peaks at *m/z* 1260.9 and 1282.9 corresponding to the protonated and sodiated ions of the diacylated *N*-terminal CGGGSDSK tryptic lipopeptide (Figure 4-2D and E, insets). To differentiate between the DA- and lyso-lipopeptide, which have identical masses, the sodiated ion *m/z* 1282 from each sample was further fragmented. While the MS/MS spectra of both strains displayed the same CGGGSDSK y-series peptide ions, the dehydroalanyl *m/z* ion differed depending on *lit2* expression. The type strain spectrum had a prevalent ion at *m/z* 676, corresponding to the dehydroalanyl CGGGSDSK peptide with a free α-amino terminus (Figure 4-2D and E), while introduction of *lit2* resulted in a shift to *m/z* 900.4. This mass is consistent with an added $C_{15:0}$ *N*-acyl chain (Figure 4-2F and G), and is further supported by MS/MS spectra of the protonated parent *m/z* 1260 ion (Figure 4-3B and C). Thus, lipoproteins in the *L. monocytogenes* type strain ATCC 19115 are DA-LP, and expression of *Lm*CFSAN*lit2* alone is sufficient for lipoprotein conversion.

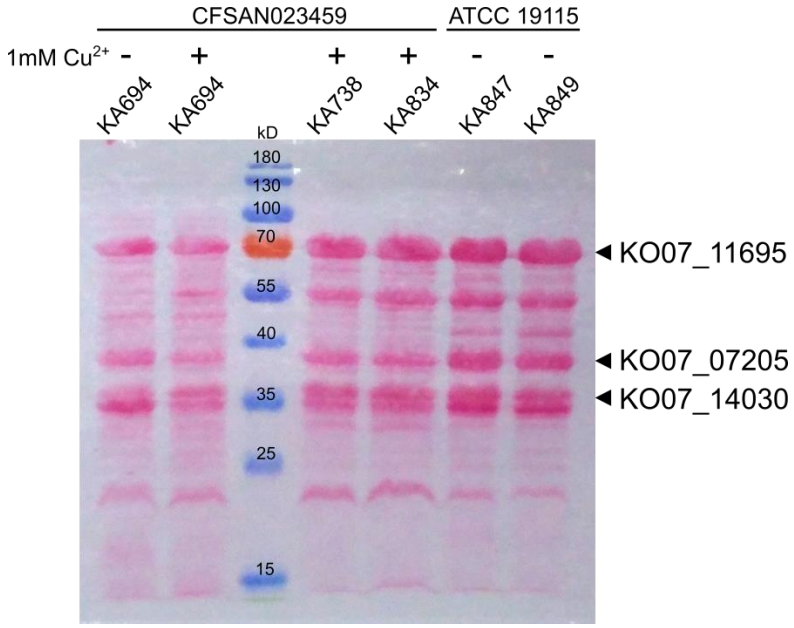


Figure **4-4**: Lipoprotein profiles from *Listeria monocytogenes* CFSAN023459 and ATCC 19115. Using the Triton X-114 phase partitioning method, lipoproteins were enriched from the following *L. monocytogenes* strains grown in HTM+ with the indicated amount of $CuCl_2$: KA694 (CFSAN023459) without and with 1 mM $CuCl_2$, KA738 (CFSAN023459 Δ*lit2*) with 1 mM $CuCl_2$, KA834 (CFSAN023459 Δ*lit2* + p*lit2*) with 1 mM $CuCl_2$ plus 2% xylose, KA847 (ATCC 19115) in HTM+only, and KA849 (ATCC 19115 + p*xyllit2*) in HTM+ plus 2% xylose. The Triton X-114 phases were separated by SDS-PAGE, transferred to nitrocellulose membrane, and visualized by Ponceau S staining. The lipoproteins chosen for analysis by MALDI-TOF are indicated.

Copper induces expression of *lit2*

As the *lit2*-copper resistance operon (Figure 4-1B) contains orthologs of CopY and CopR, two known copper-responsive transcriptional regulators (27, 31), we hypothesized that *lit2* expression is induced by excess copper. To test this, we performed a series of Northern blots in *Ef*TX1342 (Figure 4-5A) and *Lm*CFSAN (Figure 4-5B). Copper concentrations were chosen so that there was no effect on growth rates (data not shown). In both strains, *lit2* transcripts were only detected when cells were grown with 1 mM copper, indicating tight basal regulation. Observed *lit2* transcript sizes in both strains are consistent with a polycistronic operon beginning with *lit2* and terminating after *copZ* at the predicted intrinsic terminator (Figure 4-1B). Copper

induced expression of select downstream genes (*copB2* for *Ef*TX1342 and *copY* for *Lm*CFSAN) produced identical transcript lengths, confirming co-expression on a single transcript. We also probed for expression of the divergently oriented lipoprotein-related gene *lgt2* in *Lm*CFSAN and likewise observed copper-dependent expression. To check whether copper induced expression of all lipoprotein-related genes is a general response, we probed expression of either chromosomal *lit1* (in *Lm*CFSAN) or *lgt1* (in *Ef*TX1342). Transcripts were constitutively expressed at similar levels regardless of the presence of copper in both cases. Induction was specific to copper, as neither strain upregulated *lit2* when grown with alternative metals (silver, cadmium, cobalt, magnesium, nickel, or zinc) (Figure 4-5C). Quantification of the copper-induced transcriptional response by RT-qPCR was consistent with Northern blotting (Figure 4-5D and E). Collectively, the data indicates that *lit2* and *lgt2* are an integral part of the copper-responsive regulon.

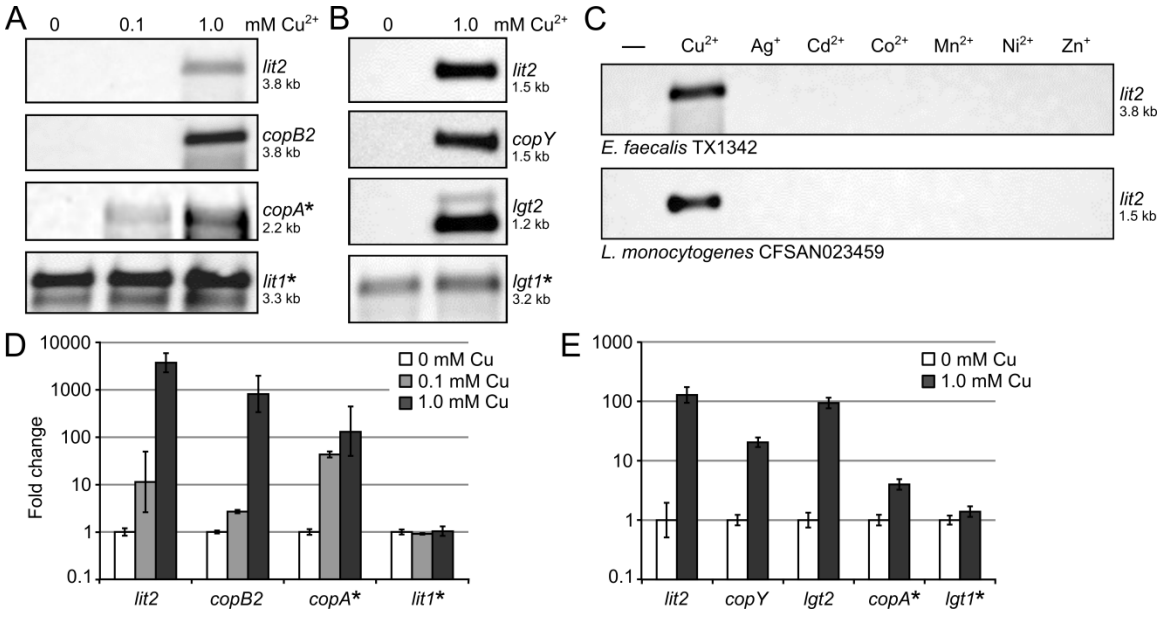


Figure **4-5**: Copper induces expression of the *lit2*-copper resistance operon. (A, B) RNA was extracted from *E. faecalis* TX1342 cells grown with 0, 0.1, or 1.0 mM $CuCl_2$ (A) and *L. monocytogenes* CFSAN023459 cells grown with 0 or 1.0 mM $CuCl_2$ (B) and probed for expression of the indicated genes. Chromosomal genes are noted with an asterisk. (C) RNA was extracted from *E. faecalis* TX1342 (upper) and *L. monocytogenes* CFSAN023459 (lower) cells induced with various metal ions and probed for *lit2* expression. Estimated transcript lengths are

indicated. (D, E) Expression of the indicated target genes was measured by RT-qPCR from *E. faecalis* TX1342 (D) and *L. monocytogenes* CFSAN023459 (E) grown with 0, 0.1, or 1.0 mM copper. Expression was normalized in each strain against the internal control gene *gyrA*. The data are shown as the mean ± the standard deviation of three replicates.

Copper induces conversion of lipoproteins from diacylglycerol-modified to lyso form in *Lm*CFSAN

While *Ef*TX1342 already encodes a constitutively-expressed chromosomal *lit* gene and thus addition of a second copy in *lit2* would not be expected to alter lipoprotein form, *L. monocytogenes* normally makes DA-LP lacking *N*-terminal modifications (Figure 4-2D). Therefore, environmental isolates of *L. monocytogenes* like *Lm*CFSAN that have acquired the *lit2*-copper resistance operon may elaborate lyso-LP, specifically when exposed to elevated copper.

Lipoproteins were extracted following growth with or without 1 mM copper, which did not alter the overall lipoprotein profile or relative abundance according to SDS-PAGE (Figure 4-4). The lipoprotein KO07_11695 was again chosen for analysis in *Lm*CFSAN, and the two samples had identical parent spectra by MALDI-TOF MS analysis (Figure 4-6). Prevalent ions at *m/z* 1260 and *m/z* 1282, along with the same corresponding y-series ions, could likewise be assigned to the diacylated *N*-terminal CGGGSDSK peptide whether copper was added or not. However, tandem MS/MS analysis revealed pronounced differences in acylation of the product ions (Figure 4-7; see Figure 4-2E and 2G for KO07_11695's *N*-terminal lipopeptide structures). When grown without copper, fragmentation of the protonated parent *m/z* 1260 yielded DA-LP related ions at *m/z* 676, 708, 748, and 1018, corresponding to dehydroalanyl CGGGSDSK, thiolated peptide, and the peptide having lost either both fatty acids (32:0) or a single $C_{15:0}$ fatty acid ($C_{14}H_{28}COOH$) (Figure 4-7A). Additional ions at *m/z* 900 and 932 were diagnostic of the lyso form, and correspond to the *N*-acyl($C_{15:0}$)-dehydroalanyl peptide and the thiolated *N*-

acyl($C_{15:0}$)-peptide. Additional ions at *m/z* 990, 1008, and 1035 can be assigned to either DA-LP or lyso-LP forms, and correspond to the parent ions having lost a $C_{17:0}$ fatty acid ($C_{16}H_{32}COOH$), $C_{17:0}$ ketene ($C_{15}H_{29}CH=C=O$), and $C_{15:0}$ ketene ($C_{13}H_{27}CH=C=O$), respectively. Fragmentation of the sodiated parent *m/z* 1282 also results in abundant peaks at *m/z* 676 and 900, evidence of both non-acylated and *N*-acylated dehydroalanyl forms (Figure 4-7B; see Figure 4-2G for lipopeptide structure). From these results, we conclude that a mixture of DA-LP and lyso-LP exists within a cell population even without added copper, suggesting some basal *lit2* expression. We could not reliably determine the ratio of DA-LP to lyso-LP due to inherent differences in ionization efficiency and multiple ion adduct formation.

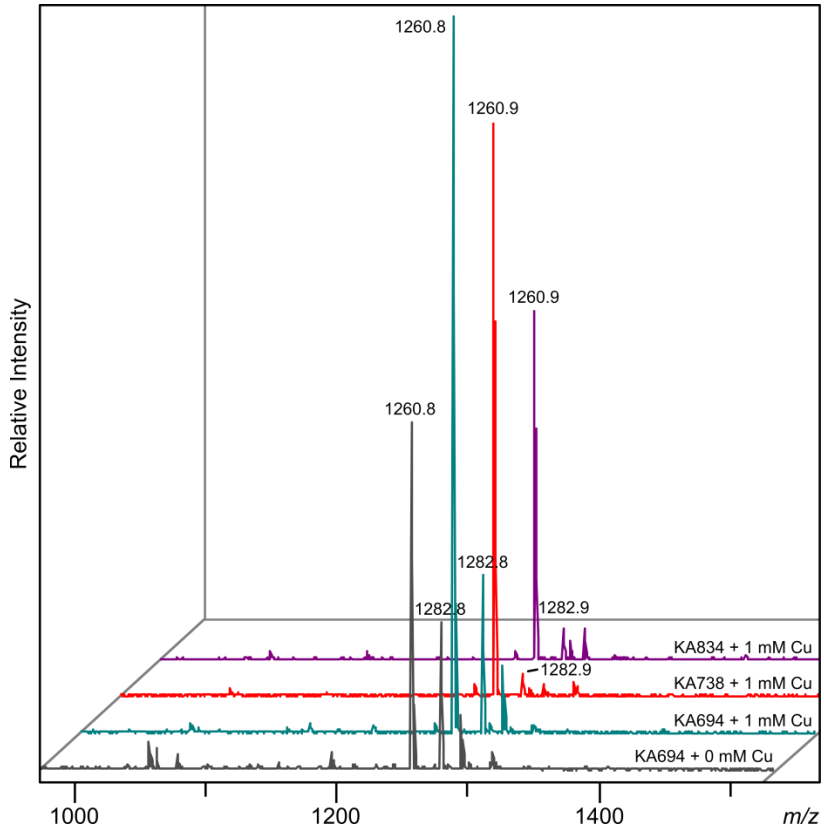


Figure **4-6**: MALDI-TOF MS of KO07_11695, a predicted peptide ABC transporter substrate-binding lipoprotein. The stacked parent MS spectra of the *m/z* 1260 ion region corresponding to the *N*-terminal lipopeptide of KO07_11695 from *L. monocytogenes* CFSAN023459 are shown. The natural abundance of the sodium adduct *m/z* 1282 is also indicated. The lipoprotein was

purified from KA694 (CFSAN023459) grown without and with 1 mM $CuCl_2$, KA738 (CFSAN023459 Δ*lit2*) grown with 1 mM $CuCl_2$, and KA834 (CFSAN023459 Δ*lit2* + *pxyllit2*) grown with 1 mM $CuCl_2$ plus 2% xylose.

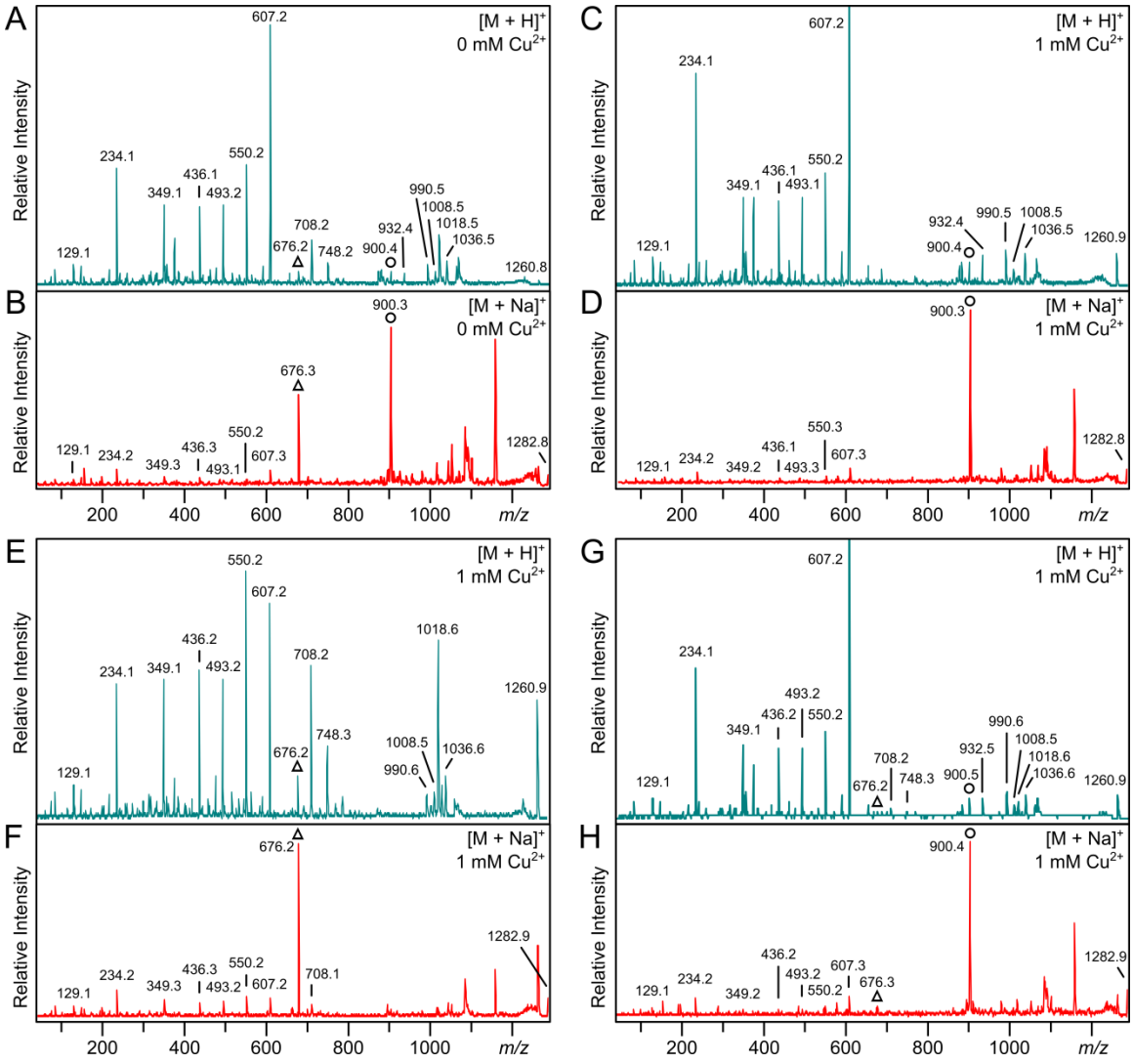


Figure **4-7**: Copper induces conversion of lipoproteins from DA-LP to lyso-LP. (A, B) The MS/MS spectra of the protonated *m/z* 1260 (A) and sodiated *m/z* 1282 (B) parent ions of lipoprotein KO07_11695 from *L. monocytogenes* CFSAN023459 grown in the absence of copper. (C, D) The MS/MS spectra of the protonated *m/z* 1260 (C) and sodiated *m/z* 1282 (D) parent ions of the *Lm*CFSAN strain grown with 1 mM copper. (E, F) The MS/MS spectra of the protonated *m/z* 1260 (E) and sodiated *m/z* 1282 (F) parent ions of the Δ*lit2* strain grown with 1 mM $CuCl_2$. (G, H) The MS/MS spectra of the protonated *m/z* 1260 (G) and sodiated *m/z* 1282 (H) parent ions of the Δ*lit2* strain back-complemented with pP_{xyl}*Lmlit2* grown with 1 mM $CuCl_2$ and 2% xylose. The diagnostic dehydroalanyl ions for the diacylglycerol-modified (triangle) and the lyso (circle) lipopeptide are indicated. Parent spectra are shown in Figure 4-6.

In contrast to the mixed lipoprotein profile from extracts grown without copper, only the lyso-LP specific ions at *m/z* 900 and 932 were observed when *Lm*CFSAN was cultured with copper (Figure 4-7C from protonated and Figure 4-7D from sodiated parent ions). There was also a marked absence of DA-LP related ions. To see whether copper induced lipoprotein conversion is global, two additional lipoproteins were analyzed (Figure 4-8 and 4-9). A clear enrichment in the lyso-LP population as seen with KO07_11695 was likewise observed. To demonstrate that copper itself is not responsible for conversion of lipoproteins to the lyso form, we deleted *lit2* from *Lm*CFSAN and grew the resulting strain with 1 mM copper. Even when grown with copper, only DA-LP could be detected (Figure 4-7E and 4-7F). Back-complementation with a plasmid-borne *lit2* copy of the gene completely restored lyso-LP production (Figure 4-7G and 4-7H), consistent with global conversion of lipoproteins to the lyso form upon induction of the *lit2*-copper resistance operon cassette.

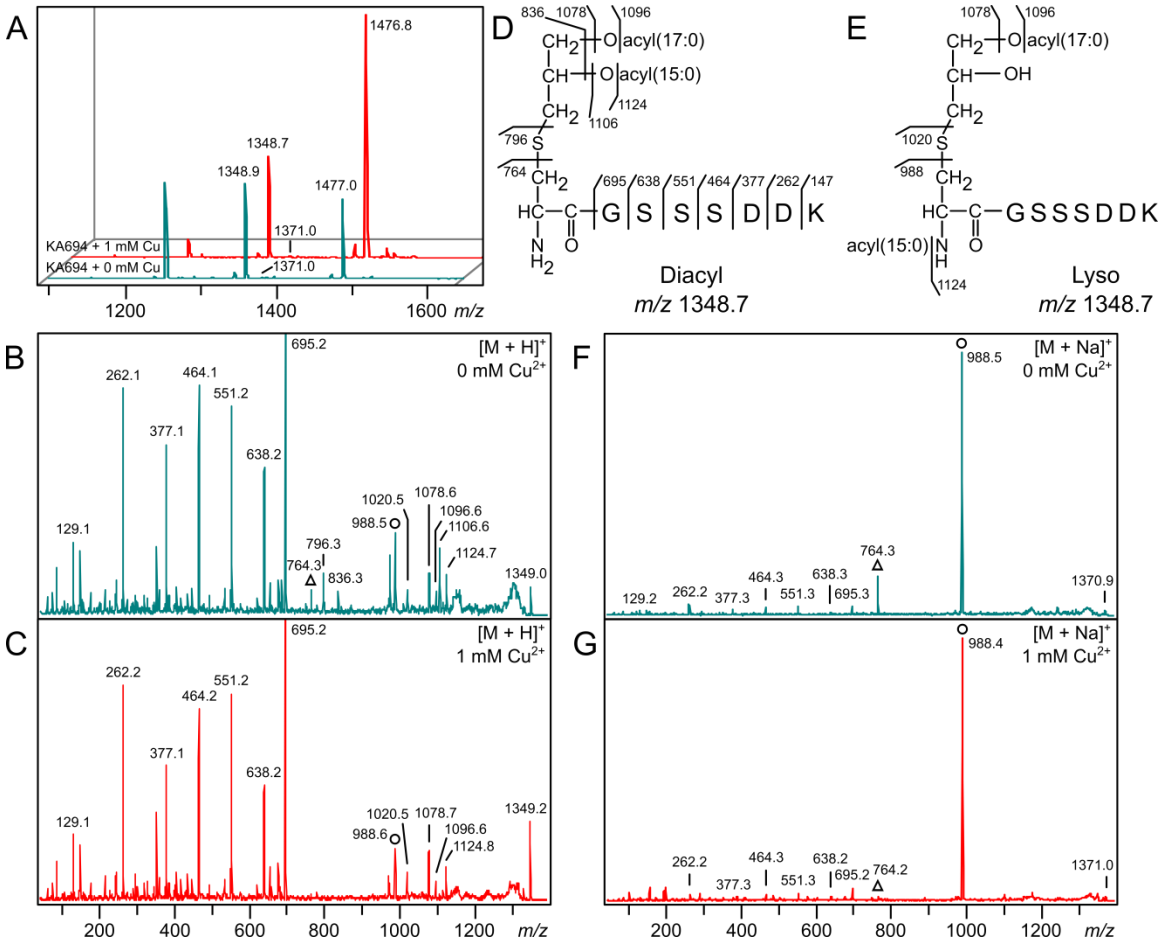


Figure **4-8**: MALDI-TOF MS and MS/MS of KO07_07205, a predicted PnrA-like lipoprotein. (A) The stacked parent MS spectra of the *m/z* 1348 ion region corresponding to the *N*-terminal lipopeptide of KO07_07205 from *L. monocytogenes* CFSAN023459 (KA694) grown with (red traces) and without (turquoise traces) 1 mM $CuCl_2$ are shown. The natural abundance of the sodium adduct *m/z* 1370 is also indicated. (B, F) The MS/MS spectra of the protonated *m/z* 1349.0 (B) and the sodiated *m/z* 1370.9 (F) parent ions reveal both diacylglycerol-modified and lyso-form lipopeptides when KA694 is grown without $CuCl_2$. (C, G) The MS/MS spectra of the protonated *m/z* 1349.2 (C) and the sodiated *m/z* 1371.0 (G) parent ions reveal near-conversion to the lyso form when exogenous $CuCl_2$ is added to the growth media. (D, E) The elucidated structures of the diacylglycerol-modified (D) and the lyso-form (E) N-terminal lipopeptides are shown. The diagnostic dehydroalanyl ions for the diacylglycerol-modified (triangle) and the lyso (circle) lipopeptide are indicated.

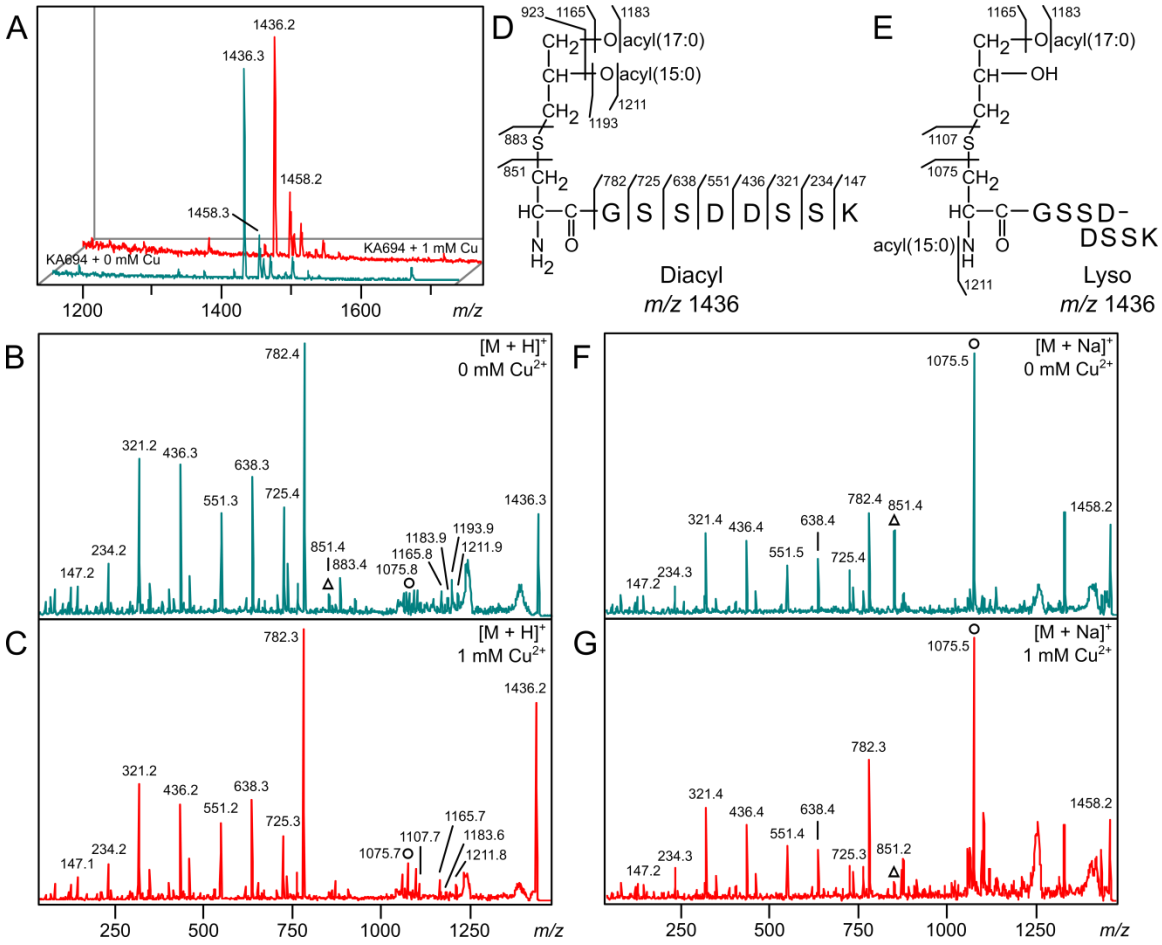


Figure **4-9**: MALDI-TOF MS and MS/MS of KO07_14030, a predicted FMN-binding lipoprotein. (A) The stacked parent MS spectra of the *m/z* 1436 ion region corresponding to the *N*-terminal lipopeptide of KO07_14030 from *L. monocytogenes* CFSAN023459 (KA694) grown with and without 1 mM $CuCl_2$ are shown. The natural abundance of the sodium adduct *m/z* 1458 is also indicated. (B, F) The MS/MS spectra of the protonated m/z 1436.3 (B) and the sodiated m/z 1458.2 (F) parent ions reveal both diacylglycerol-modified and lyso-form lipopeptides when KA694 is grown without copper. (C, G) The MS/MS spectra of the protonated m/z 1436.2 (C) and the sodiated m/z 1458.2 (G) parent ions reveal near-conversion to the lyso form when exogenous $CuCl_2$ is added to the growth media. (D, E) The elucidated structures of the diacylglycerol-modified (D) and the lyso-form (E) N-terminal lipopeptides are shown. The diagnostic dehydroalanyl ions for the diacylglycerol-modified (triangle) and the lyso (circle) lipopeptide are indicated.

TLR2 stimulation by synthetic lyso-form lipopeptides

As lipoproteins are converted from DA-LP to the lyso form in *Lm*CFSAN, we questioned if this alteration affects recognition by TLR2. While it has been established that DA-LP are sensed by the TLR2/6 heterodimer and TA-LP are sensed by TLR2/1 (4, 6, 22, 23, 35), it is not clear which heterodimer senses the lyso form and with what sensitivity. Since little is known about the specific TLR2 response to lyso-LP (11), we first measured TLR2 signaling activity using a set of defined synthetic lipopeptides and a HEK-Blue-TLR/2/1/6 reporter cell line that secretes alkaline phosphatase (SEAP) upon activation of the TLR2 responsive NF-κB transcription factor. TLR2 activation was measured using Pam_3CSK_4, Pam_2CSK_4, and $PamC(Pam)SK_4$, representing the TA-LP, DA-LP, and lyso-LP forms, respectively, as well as the *Mycoplasma salivarium*-derived DA-LP FSL-1 ligand ($Pam_2CGDPKHPKSF$). The DA-LP Pam_2CSK_4 and FSL-1 ligands elicited signal at the lowest concentration (0.01 ng/well), with an approximately 10-fold higher concentration (0.1 ng/well) needed for half-maximal activation (EC_{50}) (Figure 4-10A). Both the detection limit (0.1 ng/well) and the EC_{50} (1 ng/well) was shifted to slightly less than 10-fold higher for the TA-LP Pam_3CSK_4 in comparison to the DA-LP ligand set. The lyso-LP $PamC(Pam)SK_4$ was by far, however, the weakest TLR2/1/6 ligand, with a shift higher in detection limit and EC_{50} well over 2-logs in comparison to the DA-LP standards (Figure 4-10A).

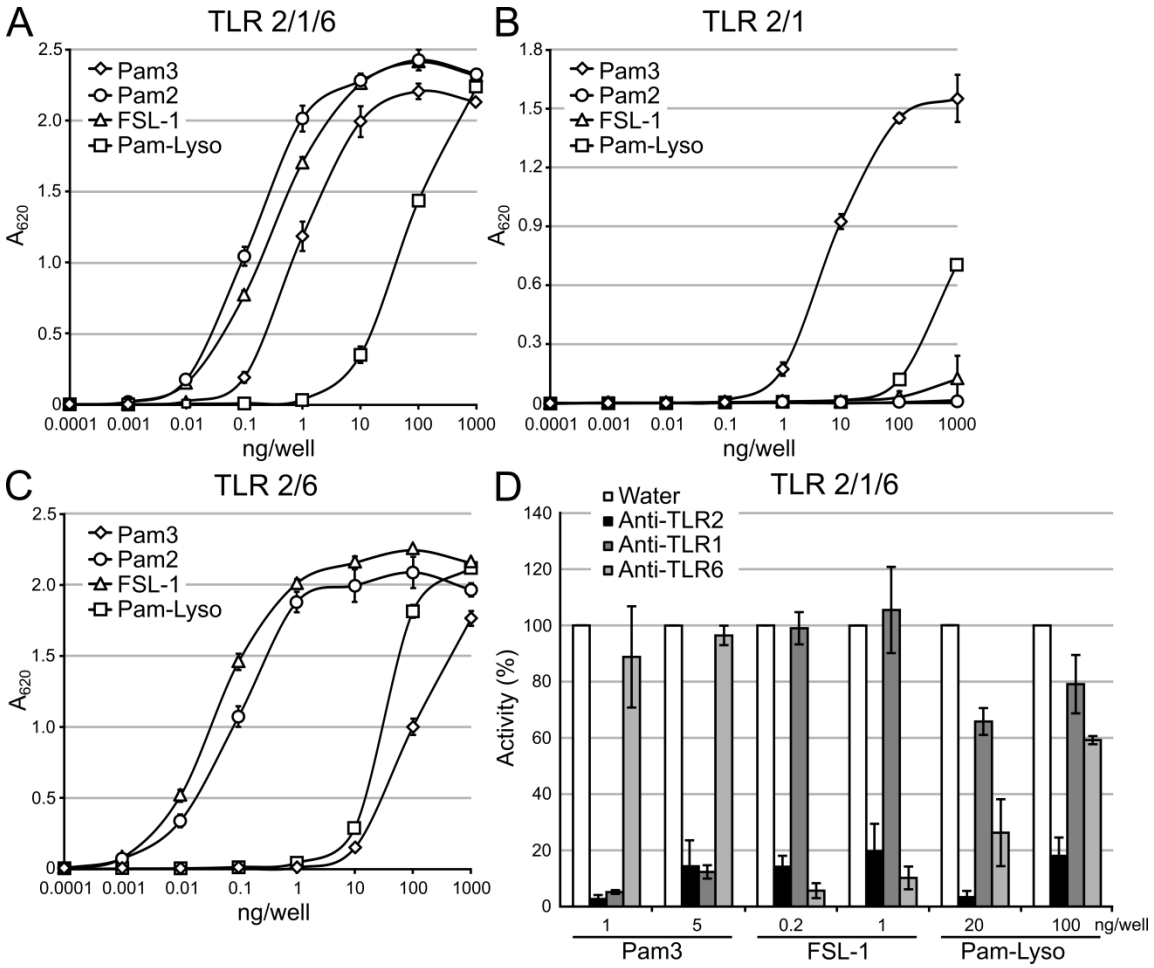


Figure **4-10**: TLR2 response to synthetic lipopeptides. (A) HEK-Blue-TLR2/1/6 cells were exposed to 10-fold dilutions of the synthetic lipopeptides Pam_3CSK_4 (Pam3) and $PamC(Pam)SK_4$ (Pam-Lyso), respectively representing the TA-LP and lyso-LP, as well as Pam_2CSK_4 (Pam2) and FSL-1, both DA-LP ligands. The same lipopeptides were exposed to HEK-Blue cells expressing only TLR2/1 (B) or TLR2/6 (C). (D) HEK-Blue-TLR2/1/6 cells pretreated with 10 ug/mL of TLR-neutralizing antibodies were then exposed to two optimized concentrations of Pam_3CSK_4, $PamC(Pam)SK_4$, and FSL-1. The percent activity when normalized to the water control wells is indicated. The data shown are the mean ± the standard deviation of three biological replicates.

Structural and biochemical studies of TLRs 2/1 and 2/6 have shown that conventional DA-LP lipopeptides signal through TLR2/6 while TA-LP lipopeptides signal through TLR2/1 (22, 23). Since TLR2 has a larger, diffuse binding pocket for the thioether-linked diacylglycerol moiety, ligand specificity is predominantly imparted by TLR1 accommodating the extra *N*-acyl chain of TA-LP in a hydrophobic binding pocket that is inaccessible in TLR6 (23, 35). At least

two distinct binding modes could be envisioned for lyso-LP: one where the monoacyl-glycerol acyl chain remains in TLR2 while the *N*-acyl is bound within TLR1 in a TLR2/1 heterocomplex, and a second whereby both acyl chains bind within TLR2 in a TLR2/6 complex. Therefore, we used TLR2/1 and TLR2/6 specific reporter cell lines to probe specificity. Using HEK-Blue-TLR2/1 cells, Pam_3CSK_4 was the highest affinity ligand, while the DA-LPs Pam_2CSK_4 and FSL-1 elicited little signal even at the highest concentration tested (1000 ng/well), as expected (Figure 4-105B). The lyso-form PamC(Pam)SK_4 ligand was 2-logs weaker ligand than TA-LP with respect to detection limit and EC_{50}, albeit still a slightly better ligand than the DA-LPs at the highest concentration tested (Figure 4-10B). In HEK-Blue-TLR2/6 cells, the PamC(Pam)SK_4 detection limit and EC_{50} was also over 2-logs weaker than the canonical DA-LP ligand (Figure 4-10C). Collectively, the data is consistent with the lyso-form PamC(Pam)SK_4 lipopeptide being predominantly sensed by the TLR2/6 heterodimer, with limited TLR2/1 activation when present at high concentrations. Engagement of TLR2/6, however, remained much weaker than the cognate DA-LP standards.

To support this finding, we treated HEK-Blue-TLR2/1/6 cells with neutralizing antibodies specific for each TLR component, then challenged with an optimized concentration of lipopeptide standard. Two separate ligand concentrations, each chosen to achieve robust knockdown of TLR2 signaling (greater than 80%) and within the linear range of the dose response curve (Figure 4-10A), were utilized with equal amounts of neutralizing antibody (Figure 4-10D). A more pronounced signal knockdown of the lyso-form PamC(Pam)SK_4 standard was observed when TLR6 was neutralized in comparison to TLR1 at both ligand concentrations tested. Taken together, this suggests that while the lyso form can indeed be sensed by TLR2/1, the bulk of signaling activity observed in HEK-Blue-TLR2/1/6 cells is due to TLR2/6.

TLR2 stimulation by whole bacteria

As copper-induced expression of *lit2* converts lipoproteins to the lyso form in *Lm*CFSAN (Figure 4-7) and lyso-LP is a weaker TLR2 agonist than the cognate DA-LP (Figure 4-10), we hypothesized that copper-rich growth environments could modulate TLR2 detection as well. Thus we repeated the previous TLR2 assays using heat-inactivated, whole *Lm*CFSAN cells grown with or without 1 mM copper (Figure 4-11A). Using HEK-Blue-TLR2/1/6 reporter cells, *Lm*CFSAN cells elaborating mostly DA-LP (wildtype grown without copper, and Δ*lit2* grown with or without copper) are detected at approximately 5-fold lower CFU/mL with respect to detection limit and EC_{50} than *lit2*-induced *Lm*CFSAN cells (wildtype grown with copper). When the contribution of TLR2/1 and TLR2/6 was isolated using TLR2 heterodimer-specific reporter assays, there was a similar ligand potency trend using TLR2/6 cells (Figure 4-12A). Interestingly, none of the samples including those with mostly lyso-LP compositions were able to measurably induce NF-κB in the TLR2/1 reporter cell line (Figure 4-12B). We next used neutralizing antibodies to probe specificity in TLR2/1/6 when using heat-inactivated bacterial preparations (Figure 4-11B). There was little to no reduction in signal when TLR1 was neutralized, while TLR6 neutralization reduced TLR2 stimulation by more than 80%. This implicates the TLR2/6 heterodimer as being the critical TLR2 heterodimer for sensing lyso-LP.

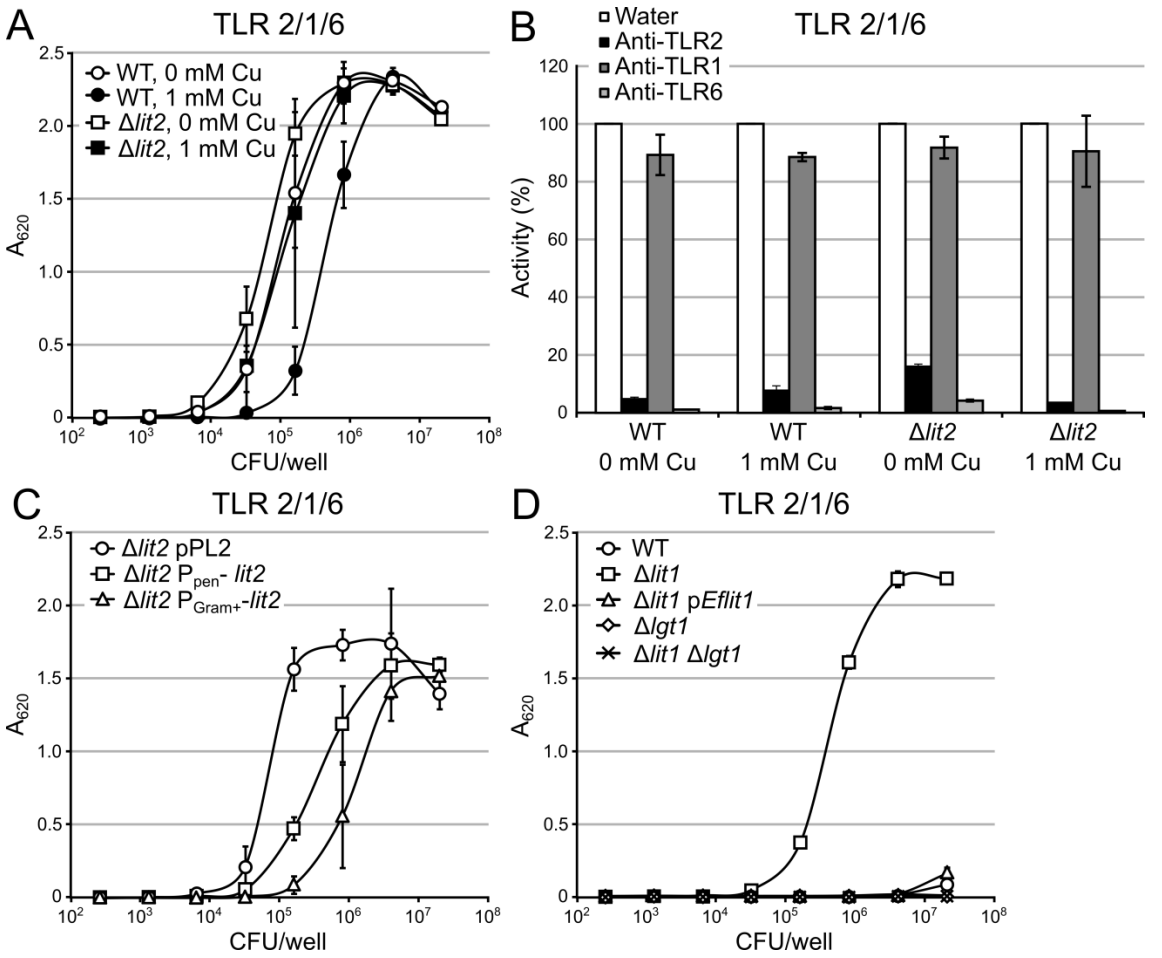


Figure **4-11**: TLR2 response to whole bacteria. (A) HEK-Blue-TLR2/1/6 cells were exposed to 5-fold dilutions of heat-inactivated, whole bacteria cells of *L. monocytogenes* CFSAN023459 (WT) grown with or without 1 mM copper, as well as the derivative Δ*lit2* cells grown with or without 1 mM $CuCl_2$. (B) HEK-Blue-TLR2/1/6 cells were pretreated with 10 ug/mL of TLR-neutralizing antibodies then exposed to the same bacterial cell preparations as in panel A. The "WT + 1 mM Cu" sample was added to 8.0 x 10^5 CFU/mL, while the others were added to 3.2 x 10^4 CFU/mL. The percent activity normalized to the water control is indicated. (C, D) HEK-Blue-TLR2/1/6 cells were exposed to 5-fold dilutions of heat-inactivated, whole bacteria cells of the indicated *Lm*CFSAN (C) and *E. faecalis* ATCC 19433 strains (D). The data are shown as the mean ± the standard deviation of three biological replicates.

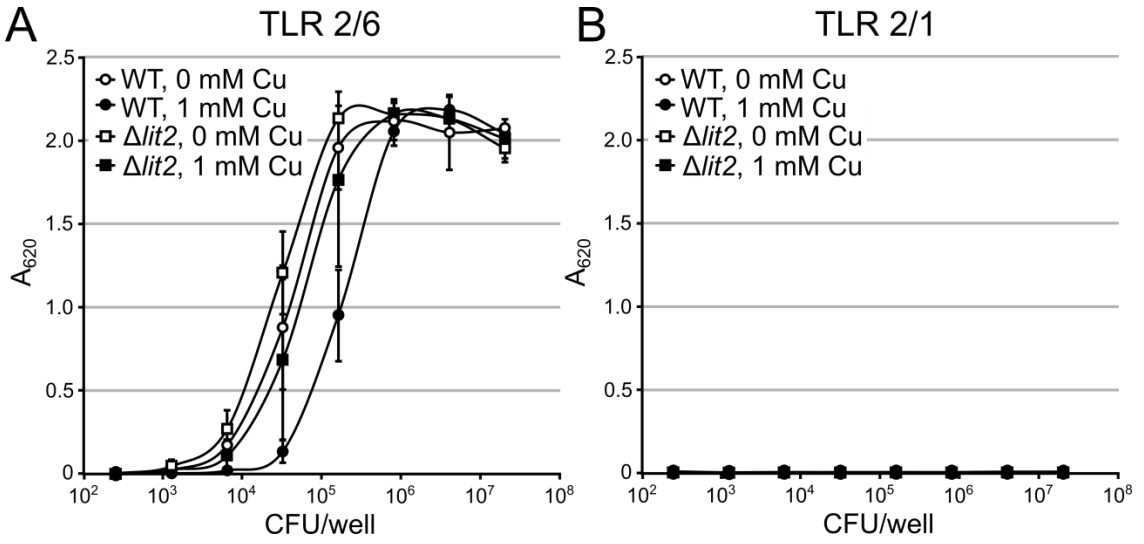


Figure **4-12**: TLR2 response to whole bacteria. (A) HEK-Blue-TLR2/6 cells were exposed to 5-fold dilutions of heat-inactivated, whole bacteria cells of *L. monocytogenes* CFSAN023459 (WT) grown with or without 1 mM $CuCl_2$, as well as the derivative Δ*lit2* cells grown with or without 1 mM $CuCl_2$. (B) HEK-Blue-TLR2/1 cells were exposed to the same bacterial cell preparations as in panel A. The data are shown as the mean ± the standard deviation of three biological replicates.

Lipoprotein conversion in the wildtype *Lm*CFSAN strain even when grown with copper under our conditions is not complete (Figures 4-7, 4-8 and 4-9). Considering the difference in potency between the DA-LP and lyso-LP, even a small residual population of DA-LP would be expected to make an outsized contribution to total signal in TLR2/1/6 reporter cells. To measure the effect of *lit2* expression independent of endogenous copper induction, *Lm*CFSAN *lit2* was placed under the control of the constitutive promoters P_{pen} and P_{Gram+}, an optimized Gram-positive promoter (36). According to the MS/MS spectra for these strains, P_{Gram+}-*lit2* (strain KA1179) realizes a more complete conversion to lyso-LP than P_{Pen}-*lit2* (KA1178) (Figure 4-13). The lyso-form producing strains P_{pen}-*lit2* and P_{Gram+}-*lit2* elicited signal at approximately 10- and 25-fold higher CFU/mL loads than the parent Δ*lit2* strain (Figure 4-11C). The TLR2 response thus correlates with the relative extent of lipoprotein population conversion from DA-LP to lyso-LP.

This trend was consistent when measured with TLR2/6 reporter cells (Figure 4-14A), and once again little to no signal was detected with TLR2/1 cells (Figure 4-14B).

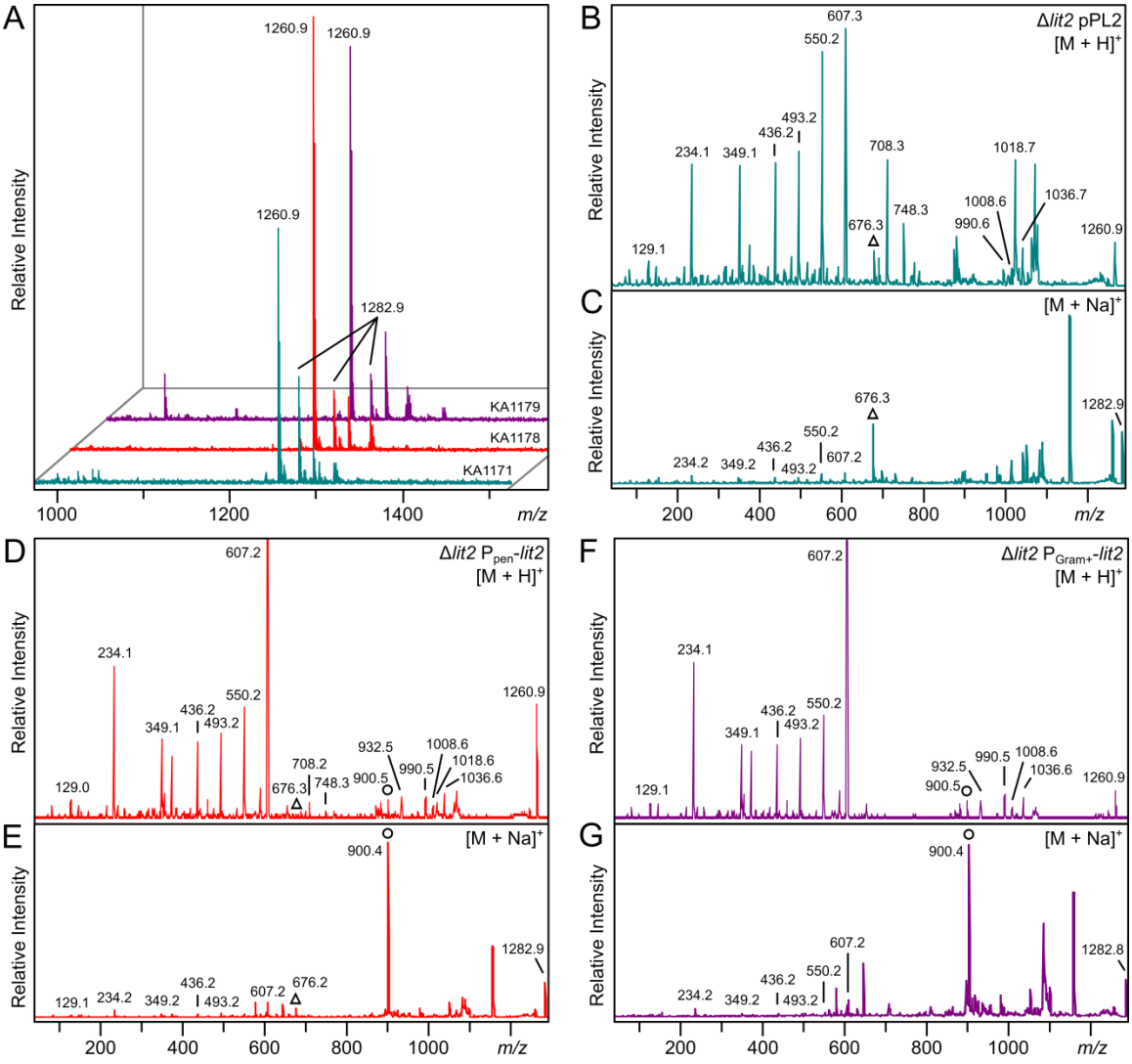


Figure **4-13**: MALDI-TOF MS and MS/MS of KO07_11695 from strains KA1171, KA1178, and KA1179. (A) The stacked parent MS spectra of the *m/z* 1260 ion region corresponding to the *N*-terminal lipopeptide of KO07_11695 from *L. monocytogenes* CFSAN023459 are shown. The natural abundance of the sodium adduct *m/z* 1282 is also indicated. (B, C) The MS/MS spectra of the protonated *m/z* 1260 (B) and sodiated *m/z* 1282 (C) parent ions of strain KA1171 (Δ*lit2 attB*::pPL2). (D, E) The MS/MS spectra of the protonated *m/z* 1260 (D) and sodiated *m/z* 1282 (E) parent ions of strain KA1178 (Δ*lit2 attB*::pPL2-P_{pen}-*lit2*). (F, G) The MS/MS spectra of the protonated *m/z* 1260 (F) and sodiated *m/z* 1282 (G) parent ions of strain KA1179 (Δ*lit2 attB*::pPL2-P_{Gram+}-*lit2*). The elucidated structures of the diacylglycerol-modified and the lyso-form *N*-terminal peptides are shown in Figure 4-2 (E and G, respectively). The diagnostic dehydroalanyl ions for the diacylglycerol-modified (triangle) and the lyso (circle) lipopeptide are indicated.

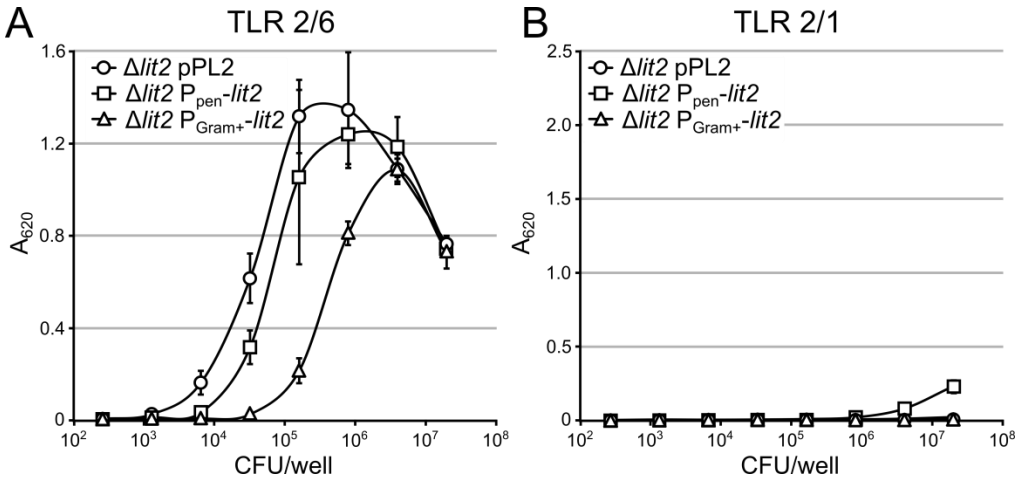


Figure **4-14**: TLR2 response to whole bacteria. (A) HEK-Blue-TLR2/1 cells were exposed to 5-fold dilutions of heat-inactivated, whole bacteria cells of the *L. monocytogenes* CFSAN023459 Δ*lit2* derivative strains with the pPL2 empty vector, or *lit2* under the control of the P_{pen} or P_{Gram+} promoters, integrated into the chromosome. (B) HEK-Blue-TLR2/6 cells were exposed to the same bacterial cell preparations as in panel A. The data are shown as the mean ± the standard deviation of three biological replicates.

As lyso-LP formation in *Lm*CFSAN reduced detection by TLR2/6, we sought to test whether this is true in other lyso-LP producing bacteria where the *lit* gene is naturally present and chromosomally integrated. We thus measured TLR2 response to heat-inactivated *E. faecalis* ATCC 19433 cells and compared to the isogenic Δ*lit1* derivative strain KA543 (12) using HEK-Blue-TLR2/1/6 reporter cells. This resulted in the most remarkable shift, with enhanced TLR2 mediated detection of the DA-LP Δ*lit1* strain at CFU/mL inputs 2- to 3-logs lower than the lyso-LP producing wildtype (Figure 4-11D). Complementation of *lit1* on a plasmid restored signal to that of wildtype cells. Deletion of *lgt*, thus abrogating production of lipoproteins altogether, resulted in no detectable signal and confirms the bulk of observed signal in fact originates from differences in lipoprotein acylation. These results were echoed in TLR2/6 cells, while no signal was detected with TLR2/1 cells for either DA-LP or lyso-LP (Figure 4-15).

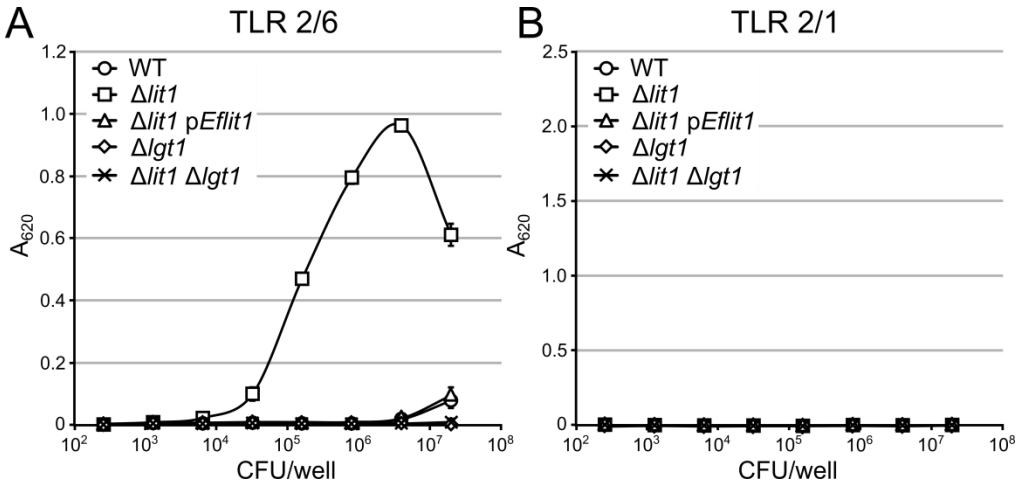


Figure **4-15**: TLR2 response to whole bacteria. (A) HEK-Blue-TLR2/6 cells were exposed to 5-fold dilutions of heat-inactivated, whole bacteria cells of *Enterococcus faecalis* ATCC 19433 and the indicated derivatives. (B) HEK-Blue-TLR2/1 cells were exposed to the same bacterial cell preparations as in panel A. The data are shown as the mean ± the standard deviation of three biological replicates.

Discussion

We originally identified Lit1 in *E. faecalis* and noted the narrow distribution of similar sequences in the NCBI database to only those bacteria expressing lyso-LP (12). Further phylogenomic analysis herein has revealed two distinct clades: *i.*) *lit1* located on the chromosome without synteny; and *ii.*) *lit2* located on a mobile genetic element within a copper resistance operon (Figure 4-1). We confirmed Lit2 is indeed a functional lipoprotein transacylase (Figure 4-2), that expression of *lit2* is induced specifically by copper (Figure 4-5), and that *lit2* alone is sufficient to convert lipoproteins from DA-LP to lyso-LP in *L. monocytogenes* CFSAN023459 (Figure 4-7). Finally, we demonstrate that the shift in lipoprotein *N*-terminal structure to lyso tangibly impacts TLR2 detection (Figure 4-10 and 4-11).

Although copper is an essential nutrient, it is highly toxic in excess and is an often used antimicrobial in healthcare and agriculture settings (37–39). Bacteria have in turn evolved

mechanisms to combat copper toxicity, many of which are subject to horizontal gene transfer (40, 41). Indeed, the presence of a *lit2*-copper resistance operon has been noted before in the *L. monocytogenes* dairy isolate DRDC8 where it is part of a freely-replicating plasmid containing multiple heavy metal resistance determinants (42), similar to the *Lm*CFSAN plasmid studied herein, and again in a comparative genomics study of copper resistance determinants in *E. faecalis* strains isolated from farms using copper-supplemented pig feed (43). In the latter, the *lit2*-copper resistance element is chromosomally-integrated in a pathogenicity/fitness island likely through transposition, as in *Ef*TX1342. Copper resistance orthologs of CopY, a copper-sensing transcriptional regulator (27, 44); CopB, a copper efflux pump (26); CopZ, a metallochaperone (28); CueO, a copper oxidase (30); and CopRS, a two-component response regulator (33) are all associated with *lit2* (Figure 4-1B). Thus, selection for copper resistance may also lead to lipoprotein conversion and in turn the unanticipated modulation of TLR2 signaling (Figure 4-16).

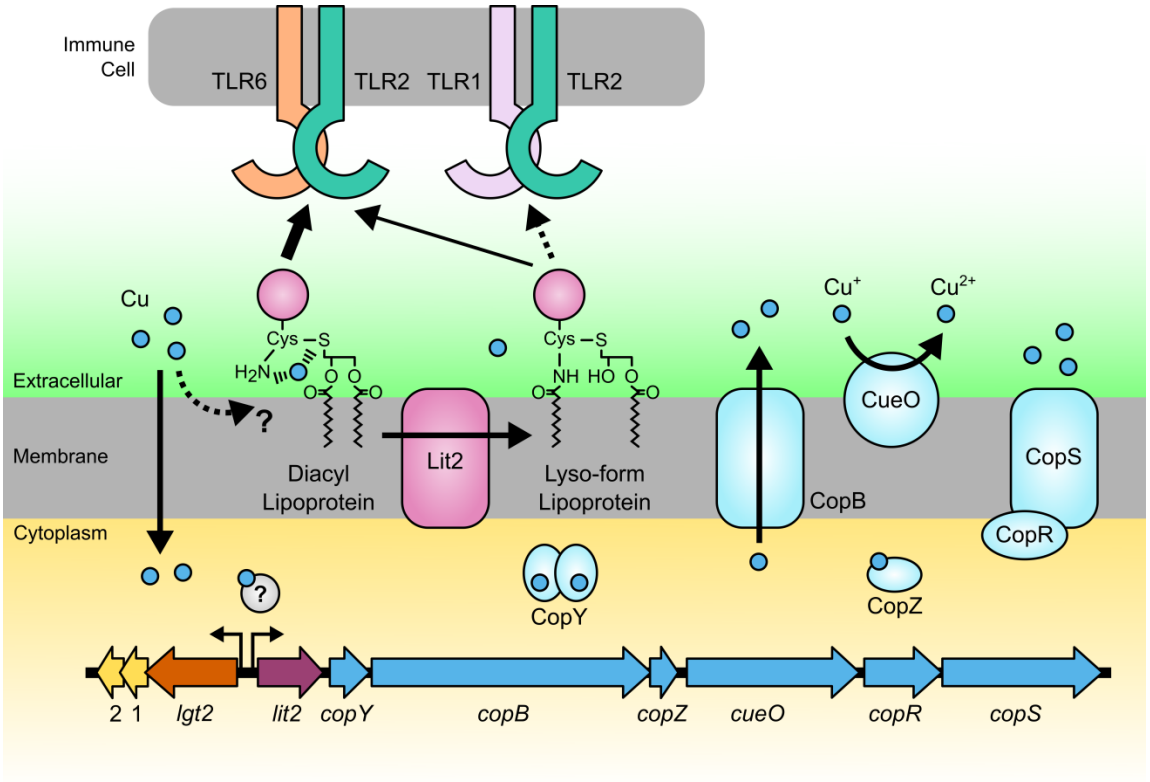


Figure **4-16**: Model of transposon-mediated response to copper. Copper enters the cell nonspecifically or via transporters. The proposed high affinity association of copper at the *N*-terminus of DA-LP is shown. Elevated copper levels induce transcription through an undetermined copper-dependent regulator. Expression of genes downstream of *lit2* contributes to copper resistance by various mechanisms, including copper efflux, copper oxidation, and regulation of additional genes. Co-induction of Lit2 simultaneously converts lipoproteins from the DA-LP form to the lyso-LP, which is proposed to reduce copper coordination at the lipoprotein *N*-terminus. While both DA-LP and lyso-LP are sensed by the TLR2/6 heterodimer, the lyso form is overall a less potent ligand at TLR2/6 than DA-LP and a poor TLR2/1 ligand.

There is a growing appreciation of the structural diversity found amongst the *N*-termini of bacterial lipoproteins (6, 11, 19, 21). While understanding the physiological function and consequences of *N*-acylation has lagged, it is clear that *N*-acylation is not simply required for lipoprotein trafficking from the inner to outer membrane of Gram-negative bacteria by the Lol-transport system (16, 17, 45, 46). Rather, the Lol transporter substrate selectivity may be more of a quality control check to ensure that only mature *N*-acylated TA-LP are transported to the outer membrane. Why lipoprotein *N*-acylation is advantageous in bacterial membranes is an outstanding question. The widespread occurrence of *N*-terminal modification though does suggest a more universal selective pressure in nature beyond Gram-negative-specific transport (14, 47). The genetic co-regulation between *lit2* and copper tolerance determinants in Firmicutes reported here, in tandem with the peculiar distribution of *N*-terminal modification machinery, may provide a clue. Bioavailable copper in the more soluble Cu^{2+} is thought to have accumulated largely after environmental oxygenation (48). Thus, a pre-existing Lgt-Lsp lipoprotein pathway would need to be edited post-speciation on a strain-by-strain basis. Acute copper challenge typical in farms and healthcare settings may likewise be contributing to acquisition of *N*-acylation genes within DA-LP-producing bacteria. When lipoprotein *N*-terminal modification is viewed as a post-pathway edit, the seemingly random species-level *N*-terminal structural variation becomes less puzzling.

This is not the first experimental connection between lipoproteins and copper, as the lipoprotein *N*-acylating *E. coli lnt* gene was originally named *cutE* due to accumulation of

intracellular copper and enhanced susceptibility in *cutE/lnt* defective mutants (49). We have however yet to observe comparable changes in copper sensitivity in Δ*lit1/lit2* mutants in *Ef*TX1342 or *Lm*CFSAN under standard growth conditions (unpublished data). Nevertheless, TA-LPs, including Braun's lipoprotein (Lpp) (50), are among the most abundant protein classes in Gram-negative bacteria (51). Thus, lipoprotein *N*-acylation appears to be a general defense strategy against copper. DA-LPs have a thioether sulfur atom as well as a free α-amino group on the *N*-terminal cysteine residue, which can conceivably coordinate copper in a high affinity complex (Figure 4-16). Any structural modifications that weaken copper cation coordination, such as a decrease in the nucleophilicity of the amino terminus, would be expected to limit lipoprotein-copper interactions. Indeed, the common denominator among all lipoprotein *N*-terminal modifications discovered to date (TA-LP with *N*-acyl/*N*-acetyl, *N*-peptidyl and lyso-LP) is α-amide formation, which delocalizes electron density on the α-amino nitrogen and the corresponding copper coordination potential. This would be advantageous, as copper binding at the cell membrane interface can increase copper uptake and disrupt metal homeostasis (52), or directly impact membrane integrity through promoting reactive oxygen species formation (53). Lipoprotein function itself could also be impacted, through enhancing cysteine thioether oxidation. Further studies will be necessary to quantify copper binding to lipoproteins of various *N*-terminal structures.

What is clear is that induction by copper of the *lit2*-copper resistance operon leads to lipoprotein conversion from DA-LP to the lyso form, and this alters recognition by TLR2 in HEK-TLR2 reporter cell assays (Figure 4-10 and 4-11). DA-LP are sensed by the TLR2/6 heterodimer, while TA-LP are sensed by TLR2/1 (4). Limited studies measuring TLR2 response to the lyso form using single, purified lipoproteins have thus far conducted have revealed complex, mixed TLR2 heterodimer activation (11). Our results show that while the lyso form can be sensed by both the TLR2/1 and TLR2/6 heterodimers, lyso-LP is a less potent TLR2 agonist

than either DA-LP or TA-LP. There was a clear inverse correlation between lyso-LP content and NF-κB reporter activation, whether lipopeptide standards or heat-inactivated bacteria were used. Once more, there was a strong preference for lyso-LP engaging TLR2/6 over TLR2/1, which is somewhat surprising given crystallographic structures clearly showing the TA-LP *N*-acyl chain binding to the hydrophobic channel of TLR1 (22). That lyso-LP, which also contains an *N*-acyl chain, interacts preferentially with the TLR2/6 heterodimer suggests complete TLR2 ligand engagement with both acyl chains is likely the most critical determinant driving receptor heterodimerization and that lyso-LP acyl chain distribution is non-optimal for this binding mode. This in turn may dampen the host innate immune response, particularly during the onset of infection when TLR2 response is most crucial (4). It will be interesting to determine whether the results observed here in non-immune HEK reporter cells directly translates to primary immune cells and ultimately infection models. Roles for other chaperones and adapter proteins in recognizing non-canonical or weak TLR2 ligands, that may not be expressed in HEK cells, have been proposed (4). Regardless, the potential of copper use in healthcare and agriculture to not only spread copper resistance-conferring genes (43, 54), but also alter the bacterial lipoprotein structural landscape and hence the TLR2 response, needs to be considered.

Materials and Methods

Phylogenetic analysis

Strains of interest were identified in a BLAST search using the protein sequence of *Enterococcus faecalis* ATCC 19433 Lit as query. To identify isolates from partially-assembled genomes, a tblastn search was performed against the database of whole-genome shotgun (wgs)

contigs. Sequences were aligned using Muscle. A phylogenetic tree was constructed using the neighbor-joining (NJ) method in MEGA7 (55).

Bacterial strains and growth conditions

Strains and plasmids used in this study are listed in Table 4-1 and primers in Table 4-2. *Enterococcus faecalis* TX1342 was grown in tryptic soy broth (TSB) at 37°C with agitation. Strain KA666 was grown with chloramphenicol (5 µg/mL) and nisin (100 ng/mL) when appropriate. *Listeria monocytogenes* CFSAN023459 contains two plasmids, CFSAN023459_01 (12,949 bp; Accession: NZ_CP014253.1) and CFSAN023459_02 (52,687 bp; Accession: NZ_CP014254.1); the latter containing the *lit2*-copper resistance operon. *Listeria* strains were grown in modified Hsiang-Ning Tsai medium (HTM) at 37°C with agitation (56). To create HTM+, HTM was supplemented with 0.1 mg/mL each of alanine, arginine, asparagine, aspartic acid, glutamine, glutamate, glycine, histidine, isoleucine, leucine, lysine, phenylalanine, proline, serine, threonine, and valine. Antibiotic markers were selected with chloramphenicol (2.5 µg/mL). Cells were induced with a final concentration of 1 mM copper (II) chloride or 2% (w/v) xylose, each added when cultures had reached an OD_{600} of 0.1 unless noted.

Table **4-1**: Bacterial strains and plasmids used in this study.

Strain	Relevant Genotype[1]	Reference
Escherichia coli S17-1	recA pro hsdR RP42Tc::Mu-Km::Tn7 integrated into the chromosome	
Escherichia coli BW25113	*E. coli* K-12 wildtype [Δ*(araD-araB)567* Δ*lacZ4787(::rrnB-3)* λ⁻ *rph-1* Δ*(rhaD-rhaB)568 hsdR514*]	
TXM327	*lpp*::Cm^r	(12)
TXM541	*gut*::Kan^r-*rrnB* TT-*araC-P$_{BAD}$-lnt*, *lnt*::Spt^r, *chiQ*::Apr^r	(12)
KA707	TXM327 + p*Lmlit2*	This study
KA708	TXM327 + p*Eflit2*	This study

KA808	KA708 *chiQ*::Apr^r, *lnt*::Spt^r	This study
KA811	KA808 + pKA810	This study
Listeria monocytogenes		
KA694	*Listeria monocytogenes* CFSAN023459 with plasmid CFSAN023459_02 (courtesy of Dwayne Roberson)	
KA738	KA694 Δ*lit2*	This study
KA834	KA834 + pP_{xyl}*Lmlit2*	This study
KA847	*Listeria monocytogenes* L2 ATCC 19115	
KA849	KA847 + pP_{xyl}*Lmlit2*	This study
KA1171	KA738 *attB*::pPL2	This study
KA1178	KA738 *attB*::P_{pen}*Lm*CFSAN*lit2*	This study
KA1179	KA738 *attB*::P_{Gram+}*Lm*CFSAN*lit2*	This study
Enterococcus faecalis		
TXM465	*Enterococcus faecalis* ATCC 19433	
KA543	TXM465 Δ*lit1*	(12)
KA666	KA543 + pKA635	(12)
KA693	*Enterococcus faecalis* TX1342 (courtesy of Barbara E. Murray)	
GKM744	KA543 Δ*lgt*	This study
GKM760	TXM465 Δ*lgt*	This study
Plasmid Name		
pKFC	Temperature-sensitive shuttle vector; Cm^r	
pKFC (ts fix)	Cm^r	This study
pP_{xyl}*Lmlit2*	pKFC (ts fixed)-P_{xyl}*Lm*CFSANl*it2*; Cm^r	This study
pKA635	pMS3535-*lit1*; Ery^r	(12)
pKA810	pCL25-*E. coli lppK58A*-Strep-tag; Trim^r	This study
p*lnt*	pUC19-*E. coli lnt*; Car^r	(12)
p*Eflit2*	pUC19-*E. faecalis* TX1342 *lit2*; Car^r	This study
p*LmLit2*	pUC19-*L. monocytogenes* CFSAN023459 *lit2*; Car^r	This study
pPL2	integration into *Lm* tRNA^{Arg} site; Cm^r (courtesy of Richard Calendar)	(60)
pTXM1170	pPL2-P_{pen}*Lm*CFSAN*lit2*; Cm^r	This study
PTXM1169	pPL2-P_{Gram+}*Lm*CFSAN*lit2*; Cm^r	This study

[1]Kn, kanamycin; Kan^r, kanamycin resistance; Chl^r, chloramphenicol resistance; Spt^r, spectinomycin resistance; Apr^r, apramycin resistance; Ery^r, erythromycin resistance; Trim^r, trimethoprim resistance; Car^r, carbenicillin resistance.

Table **4-2**: Primers used in this study.

Primer Name	Description	Primer Sequence

KA1061	5'-pUC19 *Lm*CFSAN*lit2* for	CGGTACCCGGGGATCCCTCCTTAGGCTATTC
KA1062	pUC19 *Lm*CFSAN*lit2* rev-3'	CCATGATTACGCCAAGCTTCATTATATTAAAAGG
KA1063	5'-pUC19 *Ef*TX1342*lit2* for	CGGTACCCGGGGATCCCGCTAGTGGGAACATC
KA1064	pUC19 *Ef*TX1342*lit2* rev-3'	CCATGATTACGCCAAGCTTCTCTCAAAGTCATTCC
KA1065	5'-*Lm*CFSAN*lit2* upstream	GACGGCCAGTGAATTCATGTCCGTGCCCTTTATG
KA1066	*Lm*CFSAN*lit2* upstream-3'	TAATAGAAATAGCATGC
KA1067	5'-*Lm*CFSAN*lit2* downstream	GCATGCTATTTCTATTAACTTTGACAACAATTCATC AT
KA1068	*Lm*CFSAN*lit2* downstream-3'	TGATTACGCCAAGCTGGACATTCTACTTGAGA
KA1155	5'-P_{xyl}	CAAGGAGGTGAATGTACAATGAAAAAAAGCATGC
KA1156	P_{xyl}-3'	CGGCCGGTACCGGATCCCTCCTTAGGCTATTC
TM1172	5'-pKFC ori QC for	GTGTATTAGCACCGTTATTATATCATG
TM1173	pKFC ori QC rev-3'	CTTTTTTCATCCTACCTTCTGTATCAG
TM1174	5'-pKFC plas for	GTAGGATGAAAAAAGAGCATTATCATATTC
TM1175	pKFC plas rev-3'	ACGGTGCTAATACACTTAACAAAATTTAG
KA1511	5'-*Lm*CFSAN*lit2* qPCR	GTCTCAGATCGCGGTATAAA
KA1512	*Lm*CFSAN*lit2* qPCR-3'	GCTACTCGTAGAAAGAAGCA
KA1515	5'-*Lm*CFSAN*copA* qPCR	TCTTTGCCGAACAACTACC
KA1516	*Lm*CFSAN*copA* qPCR-3'	TGCCATCACCAACGAATG
TM1554	*Ef*TX1342*lit2* T7-3'	GTTTATAATACGACTCACTATAGGGAGACAATCGT TGTTAGTG
TM1555	5'-*Ef*TX1342*lit2* T7	GTCTTTGCTTTGTCGCTTGC
TM1556	*Ef*TX1342*copB2* T7-3'	GTTTATAATACGACTCACTATAGGGAGAAGCATGT GCGCACCAG
TM1557	5'-*Ef*TX1342*copB2* T7	CCAATGTCAATTTGGCTACAG
TXM1560	*Lm*CFSAN*lit2* T7-3'	GTTTATAATACGACTCACTATAGGGAGAGTCAAA GTAGAAACAC
TXM1561	5'-*Lm*CFSAN*lit2* T7	CAAAACTGGAGTAGATCTCAAG
TM1564	*Ef*TX1342*lit1* T7-3'	GTTTATAATACGACTCACTATAGGGAGAAAATAA AGGACACTAGG
TM1565	5'-*Ef*TX1342*lit1* T7	GCGTTTAAGGGAAACACTTGGAC
KA1566	*Lm*CFSAN*lgt2* T7-3'	GTTTATAATACGACTCACTATAGGGAGACAATTT CTCCATAGG
KA1567	5'-*Lm*CFSAN*lgt2* T7	GCTAATAGAAGAGCAAAGCGTG

TM1568	*Ef*TX1342*copA* T7-3'	GTTTATAATACGACTCACTATAGGGAGAGTCGCATGTGTCATTGG
TM1569	5'-*Ef*TX1342*copA* T7	CTGTGAACCAATTATCTGGTGTTC
KA1576	5'-*Ef*TX1342gyrA qPCR	TTTCCAACAGGCGGTTTAG
KA1577	*Ef*TX1342gyrA qPCR-3'	TCCATTCGGCATTTCAGTC
TM1594	5'-*Ef*TX1342*lit2* qPCR	GTCGCTTGCAATTTCAGTTAC
TM1595	*Ef*TX1342*lit2* qPCR-3'	GGATTGTTCAAGTAGTTCATC
TM1596	5'-*Ef*TX1342*lit1* qPCR	GGCTTTAAGTATCACGATCAC
TM1597	*Ef*TX1342*lit1* qPCR-3'	AATCAAGGTCACACGATCAAC
TM1598	5'-*Ef*TX1342*copA* qPCR	ACGATTGAAAAAGCTGTGAAC
TM1599	*Ef*TX1342*copA* qPCR-3'	ATAACCTGCATCCGTAACTG
KA1604	*Lm*CFSAN*copY* T7-3'	GTTTATAATACGACTCACTATAGGGAGACATTGAATTCCTCCG
KA1605	5'-*Lm*CFSAN*copY* T7	CAGGTATCGAATTCAGAG
KA1628	5'-*Lm*CFSAN*lgt1* T7	GCGCTAATTGGTGCAG
KA1629	*Lm*CFSAN*lgt1* T7-3'	GTTTATAATACGACTCACTATAGGGAGACTTCCAAATGAATACC
TM1786	5'-*Lm*CFSAN*lit2* P_{Gram+}	TAGAACTAGTGGATCAACGACGGCCAGTGAATTGACAAAAATTGGTATATATGATATAATATAATCAAGGAGTGATCTA
TM1787	*Lm*CFSAN*lit2* P_{Gram+}-3'	AGGGAACAAAAGCTGACCATCCCTCCTTAGG
TM1800	5'-*Lm*CFSAN*lit2* Ppen	AACAGCGCGTGTATTAGGAGTGATCTAATAT
TM1801	*Lm*CFSAN*lit2* Ppen-3'	ACGGCCAGTGAATTAAGCTAATTCCGGTGG
TM1841	5'-*Lm*CFSAN*copY* qPCR	CAGGTATCGAATTCAGAGTTAGATG
TM1842	*Lm*CFSAN*copY* qPCR-3'	CCAACTATTACGCTCTTGCATTC

Construction of deletion strains and *lit2* complementation plasmid

An unmarked, internal deletion of *lit2* from plasmid CFSAN02359_02 was generated using the temperature-sensitive pKFC plasmid (57) and verified by PCR. To complement this deletion and for *lit2* expression in *L. monocytogenes* ATCC 19115, the xylose-inducible promoter from pSPNprM-hp (a gift from Dieter Jahn; Addgene plasmid number 48120) and *lit2* from CFSAN023459_02 were cloned into pKFC-ts fix that had been repaired for stable replication.

Transformation of *L. monocytogenes*

Either electroporation or conjugation from *E. coli* S17-1 was employed to transform strains of *L. monocytogenes*. Electroporation was performed following the protocol described by Monk et al. (58). Target genes were cloned into the phage attachment site integrating vector pPL2 under the control of the indicated promoters (a gift from Dr. Richard Calendar (59)) by a 3-piece DNA fragment assembly (InFusion, Clontech). Constructs were introduced into recipient *L. monocytogenes* through bi-parental conjugation.

Total RNA isolation

RNA was extracted using the RNeasy Mini Kit (Qiagen) with the following modifications. One-mL of cells at an OD_{600} of 1.0 were pelleted by centrifugation, treated with 1 mL of RNAlater for 20 min, and stored at -20°C until use. Cells were washed with 1 mL of TE buffer (10 mM Tris, pH 8.0, 1mM EDTA), resuspended in 900 µL QIAzol Lysis Reagent, and combined with an equal volume of 0.1-mm zirconium beads. Cells were disrupted using a MagNA Lyser at 7000 speed for two 20-sec cycles with 2-min rest on ice in between. Beads and unbroken cells were removed by centrifugation (3,000 x g for 1 min), washed with an additional

300 µL of QIAzol Lysis Reagent, and the supernatants pooled. After chloroform-induced phase separation, 500 µL of the aqueous phase was collected for subsequent steps. The optional centrifugation at max speed for 1 min to dry the column was performed. RNA was eluted with two 40-µL volumes of RNase-free water for a total of 80 µL. RNA was quantified by absorbance at 260 nm and the rRNA integrity assessed by MOPS-formaldehyde agarose gel.

Total RNA was prepared as described above for metal induction experiments. The minimum inhibitory concentration for each metal in TSB was determined for *Ef*TX1342 (data not shown). Overnight cultures were then diluted into metal supplemented TSB (1:100 v/v, 1 mM $CuCl_2$, $NiCl_2$, $MnCl_2$, or $ZnSO_4$, 0.015 mM $AgNO_3$ or $CdCl_2$, 0.25 mM $CoCl_2$) and grown at 37°C until OD_{600} reached 1 before harvesting cells for RNA isolation. For *Lm*CFSAN, cells were grown to an OD_{600} of 0.3, induced with 1 mM of each metal for 1 hr before harvesting.

Northern blotting

Northern blots were performed using the NorthernMax Kit (Ambion) per manufacturer's instructions. Briefly, 500 ng of total RNA from *L. monocytogenes* and 2 µg from *E. faecalis* were separated by MOPS-formaldehyde agarose gel and transferred to BrightStar-Plus positively charged nylon membrane (Invitrogen) using the Whatman Nytran SuPerCharge TurboBlotter Kit (GE Healthcare Life Sciences) for 3.5 hr. Membranes were crosslinked by baking at 80°C for 20 min. Biotin-labeled RNA probes were synthesized from DNA using the MAXIscript T7 Transcription Kit (Thermo Fisher), including the optional DNase digest and clean-up with NucAway Spin Columns (Invitrogen), along with gene-T7 specific primer sets (Table 4-2). Probes were added to 10 ng/mL in ULTRAhyb Ultrasensitive hybridization buffer (Invitrogen) and incubated at 72°C for 20 hr. Membranes were washed as directed by the NorthernMax Kit,

with the two high stringency washes performed at 68°C. RNA was visualized by the Chemiluminescent Nucleic Acid Detection kit according to instructions (Thermo Fisher).

Reverse-Transcription Quantitative PCR (RT-qPCR)

Maxima First Strand cDNA Synthesis Kit for RT-qPCR, with dsDNase treatment (Thermo Fisher) was used to synthesize cDNA template from 150 ng of total RNA isolated as described above. PowerUp SYBR Green Master Mix (Applied Biosystems) was used for RT-qPCR reactions, with gene-qPCR specific primer sets (Table 4-2) and measured using an Applied Biosystems 7300 Real-Time PCR machine. All data were measured in triplicate, normalized to internal *gyrA* control, and relative expression levels calculated using the $2^{-\Delta\Delta CT}$ method (60).

Purification of Lpp(K58A)-Strep-tag and lipoprotein extraction

Lpp(K58A)-Strep-tag was affinity column-purified from the *E. coli* strain KA811 and separated by SDS-PAGE using a 16.5% Tris-tricine gel as described (12). Unlabeled lipoproteins were extracted using the Triton X-114 phase partitioning method and separated with a 12% Tris-glycine SDS-PAGE gel (12, 34).

Preparation of N-terminal tryptic lipopeptides for MALDI-TOF MS and MS/MS analysis

Following transfer to nitrocellulose membrane, bands were visualized by Ponceau S staining. The bands selected for mass spectrometric analysis were trypsinized overnight and eluted from the membrane as described (12, 34). Single or multiple 0.5-µL layers of lipopeptide in 10 mg/mL CHCA were deposited onto a steel target plate. Mass spectra were collected using

an Ultraflextreme (Bruker Daltonics) MALDI-TOF mass spectrometer in positive reflectron mode. Tandem MS/MS spectra were acquired on the same instrument in Lift mode.

TLR2 assays

HEK293 NF-κB/SEAP-reporter cells (HEK-Blue) expressing human TLRs 2/1/6, 2/1, and 2/6 (InvivoGen, San Diego, CA) were cultivated in 75-cm^2 culture flasks in 15 mL Dulbecco's modified eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine, 50 U/mL penicillin, 50 mg/mL streptomycin, 100 mg/mL Normocin (InvivoGen), and 4 μL/mL HEK-Blue Selection antibiotics at 37°C in 5% CO_2. The triacylated Pam_3CSK_4 and diacylglycerol-modified Pam_2CSK_4 and FSL-1 (Pam_2CGDPKHPKSF) synthetic lipopeptides were purchased from InvivoGen, and the lyso-form PamC(Pam)SK_4 from EMC Microcollections (Tubingen, Germany). Cells were seeded in a 96-well plate to 25,000 cells per well in 180 μL of growth media and stimulated for 20 hr at 37°C in 5% CO_2 with 20 μL of synthetic lipopeptide or heat-inactivated whole bacteria cells diluted in endotoxin-free water. Bacteria were grown to an OD_{600} of 0.5 in HTM+ for *L. monocytogenes* or TSB for *E. faecalis*, washed once with water, heat-inactivated at 58°C for 1 hr, and stored at -20°C. Heat inactivation of cells was verified by plating on solid media. Where indicated, anti-hTLR2-IgA, anti-TLR1-IgG, and anti-hTLR6-IgG neutralizing antibodies (InvivoGen) were added to a final concentration of 10 μg/mL and incubated for 1 hr before challenge. NK-κB-dependent secreted embryonic alkaline phosphatase (SEAP) activity was measured using QUANTI-Blue Solution (InvivoGen) per manufacturer's instructions.

References

1. **Kovacs-Simon A**, **Titball RW**, **Michell SL**. 2011. Lipoproteins of bacterial pathogens. Infect Immun **79**:548–561.
2. **Buddelmeijer N**. 2015. The molecular mechanism of bacterial lipoprotein modification-How, when and why? FEMS Microbiol Rev. **39**:246-261.
3. **Nguyen MT**, **Götz F**. 2016. Lipoproteins of Gram-Positive Bacteria: Key Players in the Immune Response and Virulence. Microbiol Mol Biol Rev **80**:891–903.
4. **Oliveira-Nascimento L**, **Massari P**, **Wetzler LM**. 2012. The role of TLR2 ininfection and immunity. Front Immunol. **3**:79.
5. **Szewczyk J**, **Collet JF**. 2016. The Journey of Lipoproteins Through the Cell: One Birthplace, Multiple Destinations. Adv Microb Physiol **69**: 1–50.
6. **Nakayama H**, **Kurokawa K**, **Lee BL**. 2012. Lipoproteins in bacteria: structures and biosynthetic pathways. FEBS J **279**:4247–4268.
7. **Sutcliffe IC**, **Russell RRB**. 1995..Lipoproteins of Gram-Positive Bacteria. J Bacteriol **177**:1123-1128.
8. **Sankaran K**, **Wu HC**. 1994. Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. J Biol Chem **269**:19701–19706.
9. **Babu MM**, **Priya ML**, **Selvan AT**, **Madera M**, **Gough J**, **Aravind L**, **Sankaran K**. 2006. A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. J Bacteriol **188**:2761–2773.
10. **Hussain M**, **Ichihara S**, **Mizushima S**. 1982. Mechanism of signal peptide cleavage in the biosynthesis of the major lipoprotein of the *Escherichia coli* outer membrane. J Biol Chem **257**:5177–5182.

11. **Kurokawa K**, **Ryu K-H**, **Ichikawa R**, **Masuda A**, **Kim M-S**, **Lee H**, **Chae J-H**, **Shimizu T**, **Saitoh T**, **Kuwano K**, **Akira S**, **Dohmae N**, **Nakayama H**, **Lee BL**. 2012. Novel bacterial lipoprotein structures conserved in low-GC content Gram-positive bacteria are recognized by Toll-like receptor 2. J Biol Chem **287**:13170–13181.
12. **Armbruster KM**, **Meredith TC**. 2017. Identification of the Lyso-Form N-Acyl Intramolecular Transferase in Low-GC Firmicutes. J Bacteriol **199**:e00099-17.
13. **Tschumi A**, **Nai C**, **Auchli Y**, **Hunziker P**, **Gehrig P**, **Keller P**, **Grau T**, **Sander P**. 2009. Identification of apolipoprotein N-acyltransferase (Lnt) in mycobacteria. J Biol Chem **284**:27146–27156.
14. **Sutcliffe IC**, **Harrington DJ**, **Hutchings MI**. 2012. A phylum level analysis reveals lipoprotein biosynthesis to be a fundamental property of bacteria. Protein Cell **3**:163–170.
15. **Gupta SD**, **Wu HC**. 1991. Identification and subcellular localization of apolipoprotein N-acyltransferase in *Escherichia coli*. FEMS Microbiol Lett **62**:37–41.
16. **Fukuda A**, **Matsuyama S-I**, **Hara T**, **Nakayama J**, **Nagasawa H**, **Tokuda H**. 2002. Aminoacylation of the N-terminal cysteine is essential for Lol-dependent release of lipoproteins from membranes but does not depend on lipoprotein sorting signals. J Biol Chem **277**:43512–43518.
17. **Robichon C**, **Vidal-Ingigliardi D**, **Pugsley AP**. 2005. Depletion of apolipoprotein N-acyltransferase causes mislocalization of outer membrane lipoproteins in *Escherichia coli*. J Biol Chem **280**:974–983.
18. **Widdick DA**, **Hicks MG**, **Thompson BJ**, **Tschumi A**, **Chandra G**, **Sutcliffe IC**, **Brülle JK**, **Sander P**, **Palmer T**, **Hutchings MI**. 2011. Dissecting the complete lipoprotein biogenesis pathway in *Streptomyces scabies*. Mol Microbiol **80**:1395–1412.

19. **Asanuma M**, **Kurokawa K**, **Ichikawa R**, **Ryu KH**, **Chae JH**, **Dohmae N**, **Lee BL**, **Nakayama H**. 2011. Structural evidence of α-aminoacylated lipoproteins of *Staphylococcus aureus*. FEBS J **278**:716–728.

20. **Kurokawa K**, **Lee H**, **Roh K-B**, **Asanuma M**, **Kim YS**, **Nakayama H**, **Shiratsuchi A**, **Choi Y**, **Takeuchi O**, **Kang HJ**, **Dohmae N**, **Nakanishi Y**, **Akira S**, **Sekimizu K**, **Lee BL**. 2009. The Triacylated ATP Binding Cluster Transporter Substrate-binding Lipoprotein of *Staphylococcus aureus* Functions as a Native Ligand for Toll-like Receptor 2. J Biol Chem **284**:8406–8411.

21. **Nguyen M-T**, **Uebele J**, **Kumari N**, **Nakayama H**, **Peter L**, **Ticha O**, **Woischnig A-K**, **Schmaler M**, **Khanna N**, **Dohmae N**, **Lee BL**, **Bekeredjian-Ding I**, **Götz F**. 2017. Lipid moieties on lipoproteins of commensal and non-commensal staphylococci induce differential immune responses. Nat Commun **8**:2246.

22. **Jin MS**, **Kim SE**, **Heo JY**, **Lee ME**, **Kim HM**, **Paik S-G**, **Lee H**, **Lee J-O**. 2007. Crystal Structure of the TLR1-TLR2 Heterodimer Induced by Binding of a Tri-Acylated Lipopeptide. Cell **130**:1071–1082.

23. **Kang JY**, **Nan X**, **Jin MS**, **Youn S-J**, **Ryu YH**, **Mah S**, **Han SH**, **Lee H**, **Paik S-G**, **Lee J-O**. 2009. Recognition of Lipopeptide Patterns by Toll-like Receptor 2-Toll-like Receptor 6 Heterodimer. Immunity **31**:873–884.

24. **Solioz M**, **Stoyanov J V.** 2003. Copper homeostasis in *Enterococcus hirae*. FEMS Microbiol Rev **27:**183-195.

25. **Solioz M**, **Abicht HK**, **Mermod M**, **Mancini S**. 2010. Response of Gram-positive bacteria to copper stress. J Biol Inorg Chem. **15:**3-14.

26. **Odermatt A**, **Krapf R**, **Solioz M**. 1994. Induction of the putative copper ATPases, CopA and CopB, of *Enterococcus hirae* by Ag^{+} and Cu^{2+}, and Ag^{+} extrusion by CopB. Biochem Biophys Res Commun **202**:44–48.

27. **Strausak D**, **Solioz M**. 1997. CopY is a copper-inducible repressor of the *Enterococcus hirae* copper ATPases. J Biol Chem **272**:8932–8936.

28. **Wickramasinghe WA**, **Dameron CT**, **Weber T**, **Solioz M**, **Cobine P**, **Harrison MD**. 2002. The *Enterococcus hirae* copper chaperone CopZ delivers copper(I) to the CopY repressor. FEBS Lett **445**:27–30.

29. **Multhaup G**, **Strausak D**, **Bissig KD**, **Solioz M**. 2001. Interaction of the CopZ copper chaperone with the CopA copper ATPase of *Enterococcus hirae* assessed by surface plasmon resonance. Biochem Biophys Res Commun **288**:172–177.

30. **Grass G**, **Rensing C**. 2001. CueO is a multi-copper oxidase that confers copper tolerance in *Escherichia coli*. Biochem Biophys Res Commun **286**:902–908.

31. **Mills SD**, **Lim CK**, **Cooksey DA**. 1994. Purification and characterization of CopR, a transcriptional activator protein that binds to a conserved domain (cop box) in copper-inducible promoters of *Pseudomonas syringae*. Mol Gen Genet **244**:341–351.

32. **Schelder S**, **Zaade D**, **Litsanov B**, **Bott M**, **Brocker M**. 2011. The two-component signal transduction system CopRS of *Corynebacterium glutamicum* is required for adaptation to copper-excess stress. PLoS One **6**:e22143.

33. **Quintana J**, **Novoa-Aponte L**, **Argüello JM**. 2017. Copper homeostasis networks in the bacterium *Pseudomonas aeruginosa*. J Biol Chem **292**:15691–15704.

34. **Armbruster KM**, **Meredith TC**. 2018. Enrichment of Bacterial Lipoproteins and Preparation of N-terminal Lipopeptides for Structural Determination by Mass Spectrometry. J Vis Exp 136:e56842–e56842.

35. **Manavalan B**, **Basith S**, **Choi S**. 2011. Similar structures but different roles-an updated perspective on TLR structures. Front Physiol **2**:41.

36. **Bouloc P**, **Lartigue M-F**, **Glaser P**, **Trieu-Cuot P**, **Villain A**, **Sismeiro O**, **Dillies M-A**, **Sauvage E**, **Da Cunha V**, **Rosinski-Chupin I**, **Caliot M-E**. 2015. Single nucleotide

resolution RNA-seq uncovers new regulatory mechanisms in the opportunistic pathogen *Streptococcus agalactiae*. BMC Genomics **16**:419.

37. **Hodgkinson V**, **Petris MJ**. 2012. Copper homeostasis at the host-pathogen interface. J Biol Chem **287**:13549–13555.
38. **Djoko KY**, **Ong C-LY**, **Walker MJ**, **McEwan AG**. 2015. The Role of Copper and Zinc Toxicity in Innate Immune Defense against Bacterial Pathogens. J Biol Chem **290**:18954–18961.
39. **Samanovic MI**, **Ding C**, **Thiele DJ**, **Darwin KH**. 2012. Copper in microbial pathogenesis: Meddling with the metal. Cell Host Microbe. **11:**106-115.
40. **Hobman JL**, **Crossman LC**. 2015. Bacterial antimicrobial metal ion resistance. J Med Microbiol. **64**:471-497.
41. **Silveira E**, **Freitas AR**, **Antunes P**, **Barros M**, **Campos J**, **Coque TM**, **Peixe L**, **Novais C**. 2014. Co-transfer of resistance to high concentrations of copper and first-line antibiotics among *Enterococcus* from different origins (humans, animals, the environment and foods) and clonal lineages. J Antimicrob Chemother **69**:899–906.
42. **Bell FY**, **Adelaide BBH**. 2010. Copper Tolerance of Listeria monocytogenes strain DRDC8. PhD Thesis. The University of Adelaide, Adelaide, AU.
43. **Zhang S**, **Wang D**, **Wang Y**, **Hasman H**, **Aarestrup FM**, **Alwathnani HA**, **Zhu YG**, **Rensing C**. 2015. Genome sequences of copper resistant and sensitive *Enterococcus faecalis* strains isolated from copper-fed pigs in Denmark. Stand Genomic Sci **10**:35.
44. **Portmann R**, **Magnani D**, **Stoyanov J V**, **Schmechel A**, **Multhaup G**, **Solioz M**. 2004. Interaction kinetics of the copper-responsive CopY repressor with the cop promoter of Enterococcus hirae. J Biol Inorg Chem **9**:396–402.
45. **Grabowicz M**, **Silhavy TJ**. 2017. Redefining the essential trafficking pathway for outer membrane lipoproteins. Proc Natl Acad Sci **114**:4769–4774.

46. **Narita S ichiro**, **Tokuda H**. 2011. Overexpression of LolCDE allows deletion of the *Escherichia coli* gene encoding apolipoprotein N-acyltransferase. J Bacteriol **193**:4832–4840.

47. **Grabowicz M**. 2018. Lipoprotein Transport: Greasing the Machines of Outer Membrane Biogenesis: Re-Examining Lipoprotein Transport Mechanisms Among Diverse Gram-Negative Bacteria While Exploring New Discoveries and Questions. BioEssays **40**:e1700187.

48. **Rodushkin I**, **El Albani A**, **Chi Fru E**, **Lalonde S V**, **Konhauser KO**, **Andersson P**, **Partin CA**, **Rodríguez NP**, **Weiss DJ**. 2016. Cu isotopes in marine black shales record the Great Oxidation Event. Proc Natl Acad Sci **113**:4941–4946.

49. **Rogers SD**, **Bhave MR**, **Mercer JF**, **Camakaris J**, **Lee BT**. 1991. Cloning and characterization of cutE, a gene involved in copper transport in *Escherichia coli*. J Bacteriol **173**:6742–6748.

50. **Asmar AT**, **Collet JF**. 2018. Lpp, the Braun lipoprotein, turns 50—major achievements and remaining issues. FEMS Microbiol Lett **365**:fny199.

51. **Guo MS**, **Updegrove TB**, **Gogol EB**, **Shabalina SA**, **Gross CA**, **Storz G**. 2014. MicL, a new σE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. Genes Dev **28**:1620–1634.

52. **Macomber L**, **Imlay JA**. 2009. The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity. Proc Natl Acad Sci **106**:8344–8349.

53. **Grass G**, **Rensing C**, **Solioz M**. 2011. Metallic copper as an antimicrobial surface. Appl Environ Microbiol **77**:1541–1547.

54. **Yazdankhah S**, **Rudi K**, **Bernhoft A**. 2014. Zinc and copper in animal feed – development of resistance and co-resistance to antimicrobial agents in bacteria of animal origin. Microb Ecol Heal Dis **25**:25862.

55. **Kumar S**, **Stecher G**, **Tamura K**. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol **33**:1870–1874.

56. **Tsai H-N**, **Hodgson DA**. 2003. Development of a synthetic minimal medium for *Listeria monocytogenes*. Appl Environ Microbiol **69**:6943–6945.

57. **Kato F**, **Sugai M**. 2011. A simple method of markerless gene deletion in *Staphylococcus aureus*. J Microbiol Methods **87**:76–81.

58. **Monk IR**, **Gahan CGM**, **Hill C**. 2008. Tools for Functional Postgenomic Analysis of *Listeria monocytogenes*. Appl Environ Microbiol **74**:3921–3934.

59. **Lauer P**, **Chow MYN**, **Loessner MJ**, **Portnoy DA**, **Calendar R**. 2002. Construction, characterization, and use of two *Listeria monocytogenes* site-specific phage integration vectors. J Bacteriol **184**:4177–4186.

60. **Livak KJ**, **Schmittgen TD**. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. Methods **25**:402–408.

Chapter 5

Mechanistic characterization of the lipoprotein intramolecular transacylase Lit*

*Dr. Timothy C. Meredith conceived and constructed strain TXM1111. Dr. Gloria Komazin purified both the unlabeled and deuterium-labeled Lpp(K58A)-Streptag proteins. Sarah Staskiewicz assisted with cloning and construction of the Lit-LacZ/PhoA reporter strains.

Abstract

While their globular protein domains differ, all bacterial lipoproteins share an N-terminal cysteine residue that is variably acylated depending on the species. Lipoproteins in Gram-negative bacteria feature a thiol-linked diacylglycerol moiety and *N*-acyl chain, with the latter transferred from a membrane phospholipid to the α-amino terminus by the lipoprotein *N*-acyltransferase (Lnt). Though they do not encode Lnt in their genomes, some Gram-positive Firmicutes also *N*-acylate their lipoproteins. Characterized by an *S*-monoacylglycerol moiety and *N*-acyl chain, lyso-form lipoproteins are produced in *Enterococcus faecalis*, among others. *N*-acylation of the lyso form is catalyzed by the lipoprotein intramolecular transacylase (Lit), predicted to transfer an ester-linked lipid from the diacylglycerol moiety of a lipoprotein precursor to the α-amino position. To determine the mechanism of Lit, we aimed to combine reconstituted Lit *in vitro* with acyl-labeled, diacylated lipoprotein substrate and unlabeled phospholipids. Mass spectrometric analyses of the resulting lyso-form lipoproteins will identify the source of the *N*-acyl chain as either an unlabeled membrane phospholipid, similar to Lnt, or a labeled lipid from the diacylglycerol moiety. Laying the foundation for such an experiment, a complex strain of *Escherichia coli* was constructed that is able to incorporate exogenous, deuterium-labeled lipids into diacylated lipoproteins. Conditions were then identified in which membrane-reconstituted Lit appears to be active. If Lit functions as predicted, this will be the first demonstration of intramolecular transacylation activity. Characterization of the membrane topology of Lit is also presented.

Introduction

Ubiquitous among bacteria, lipoproteins contribute to a wide variety of functions in the cell. They are characterized by a conserved N-terminal cysteine residue that is post-translationally modified with lipids, which serve to anchor the lipoprotein to the membrane. Two covalently-attached lipid chains are first added to the cysteine thiol by lipoprotein diacylglycerol transferase (Lgt), which catalyzes the transfer of a diacylglycerol moiety (DA) from a neighboring phospholipid to the lipoprotein (1). Then, lipoprotein signal peptidase (Lsp) cleaves the signal peptide of the lipoprotein, liberating the cysteine's *N*-terminus for optional further modifications depending on the species (2). In Gram-negative bacteria and some high-GC Gram-positive species, lipoprotein *N*-acyltransferase (Lnt) attaches a single acyl chain, again from a neighboring phospholipid, to the exposed α-amino group of the modified cysteine, creating the mature triacylated lipoprotein in these organisms (3). However, despite lacking an *lnt* sequence ortholog, several low-GC Gram-positive Firmicutes also elaborate *N*-acylated lipoproteins (4). *Staphylococcus aureus* and *S. epidermidis* make triacylated lipoproteins, while *Enterococcus faecalis*, *Bacillus cereus*, *Streptococcus sanguinis*, and *Lactobacillus delbrueckii* produce the recently discovered lyso form, characterized by an *N*-acyl-*S*-monoacyl cysteine structure, that is distinct from the triacyl form (Figure 5-1) (5–7).

Initially, there were two theories as to how the lyso form is made: (i) transacylation from the DA moiety to the *N*-terminus within an individual lipoprotein, or (ii) *N*-acylation by an Lnt-like mechanism coupled with *O*-deacylation (Figure 5-1) (5). Previously, no precedence of intramolecular transacylation had existed in literature, and its proposal inferred an enzyme mechanism distinct from that of Lnt. Efforts to determine how the lyso form is synthesized led to the discovery of an uncharacterized integral membrane protein in Firmicutes that converts diacylglycerol-modified lipoproteins (DA-LP) to the lyso form (8). Studies on this protein

revealed evidence in favor of a transacylation mechanism, and the protein was consequently named lipoprotein intramolecular transferase (Lit). This evidence primarily consisted of mass spectrometric (MS) analyses of the lipoprotein Lpp from *E. coli* strains expressing either Lnt or Lit, revealing differences in the mature lipoprotein's *N*-acyl chain (saturated for triacyl versus monounsaturated for lyso), ultimately suggesting that Lit's lipid donor comes from within the existing DA moiety of the lipoprotein (8). However, further studies were necessary to confirm that Lit functions in this proposed manner.

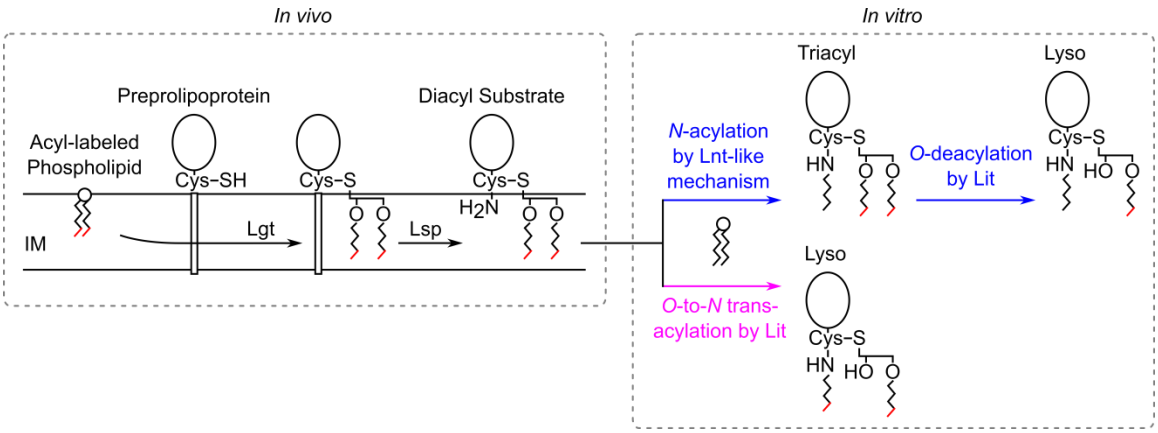


Figure **5-1**: Experimental design. Strain TXM1111 was generated (see Figure 5-2) to produce acyl-labeled membrane phospholipids (labels colored red) that are incorporated into diacylated lipoproteins *in vivo*. *In vitro*, the labeled, diacylated substrate is reconstituted with inner membranes containing unlabeled phospholipids and *E. faecalis* Lit, proposed to function by one of two mechanisms: (i) *N*-acylation concurrent with *O*-deacylation, with Lit acting as a bifunctional enzyme (blue arrows), or (ii) *O*-to-*N* transacylation (pink arrows). The first scenario results in lyso-form lipoproteins with a labeled *O*-acyl chain and unlabeled *N*-acyl chain, while both lipids are labeled in the second scenario.

We previously showed that Lnt can be deleted in *E. coli* when Lit is expressed and the abundant lipoprotein Lpp is mutated to prevent crosslinking to peptidoglycan (8). The resulting production of lyso-form lipoproteins has no deleterious effects on growth, allowing for in-depth studies on the Firmicute enzyme Lit in the model organism *E. coli*. To determine if *E. faecalis* Lit functions as an intramolecular transacylase, we aimed to combine acyl-labeled, DA-LP substrate with Lit-containing membranes and analyze the resulting lyso-form lipoproteins by MS (Figure 5-

1). A strain of *E. coli* that produces the DA-LP substrate was constructed, with deuterium labels incorporated into both lipids of the DA moiety. Reaction conditions were identified in which membrane-reconstituted Lit appeared to be active, however deuterium labeling on the final lyso-form lipoproteins has not yet been determined. Once completed, this would be the first demonstration of lipoprotein intramolecular transacylation. Investigation into the membrane topology of Lit is also presented.

Results

Construction of strain TXM1111 as source of labeled, DA-LP substrate

To test whether Lit functions as proposed, reconstituted Lit will be given acyl-labeled, DA-LP substrate in a reaction mixture (Figure 5-1). Analyzing the resulting lyso-form lipoproteins by MS, if Lit indeed functions intramolecularly, we would observe a peak corresponding to the N-terminal lipopeptide with a deuterium-labeled acyl chain. If Lit functions intermolecularly, using a phospholipid substrate provided by the membrane similar to that of Lnt, we would observe an unlabeled acyl chain on the N-terminal lipopeptide. Thus, we first needed a strain that would produce labeled, DA-LP to be used in such a reaction (Figure 5-2A).

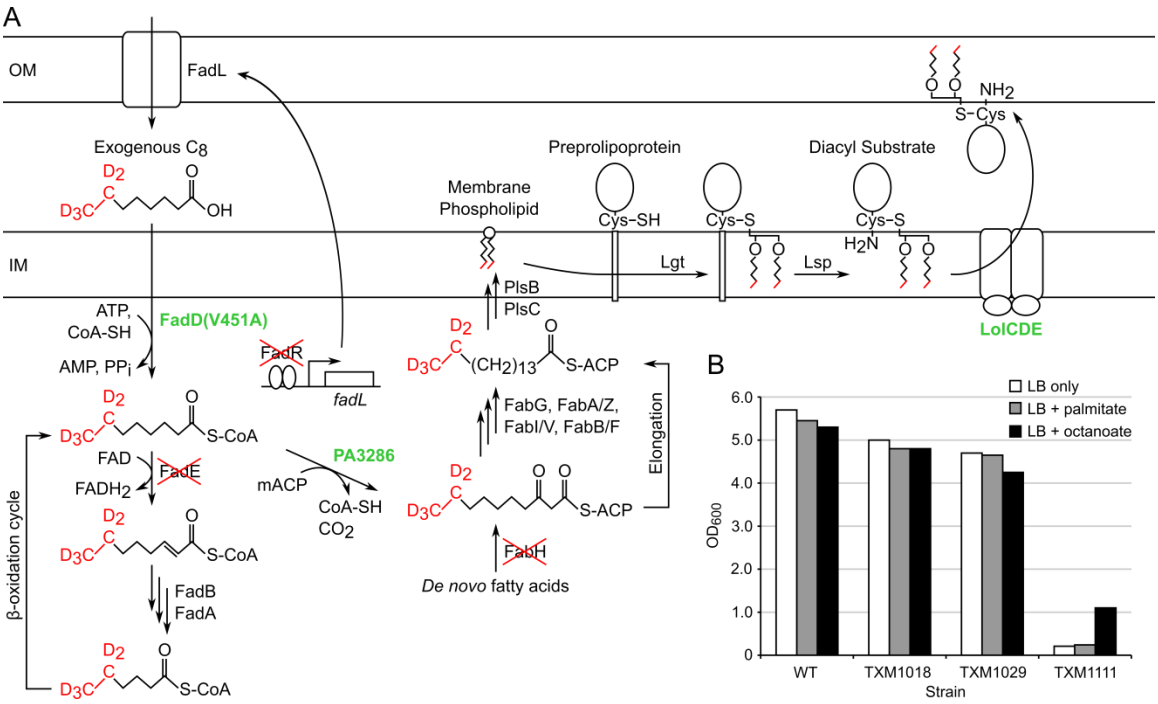


Figure **5-2**: Production of acyl-labeled DA-LP by strain TXM1111. (A) Exogenous, terminally-deuterated octanoate (labels shown in red) is transported across the OM by FadL following deletion of FadR. The octanoate is brought into the cytoplasm and esterified by the optimized FadD(V451A) (overexpressed genes shown in green). Entry of deuterated-C_8 substrate into the *E. coli* β-oxidation pathway is prevented by the deletion of FadE. Instead, it is shunted into the fatty acid biosynthesis pathway by the plasmid-encoded *P. aeruginosa* protein PA3286, while deletion of FabH precludes *de novo* fatty acids from doing the same. Following cycles of elongation and desaturation by multiple Fab proteins, deuterated acyl chains are incorporated into membrane phospholipids by PlsB and PlsC. Lipoprotein biosynthesis proceeds as normal, with the deuterium-labeled DA moiety transferred by Lgt from a phospholipid donor to the cysteine thiol of the lipoprotein. Lsp cleaves the signal peptide of the lipoprotein, exposing the α-amino group of the lipidated cysteine, to yield labeled DA-LP. The overexpressed ABC transporter LolCDE, along with LolAB (not shown), exports the non-preferred DA-LP substrate to the OM. (B) BW25113 wildtype (WT) cells and its derivatives TXM1018 (Δ*lpp*, Δ*fadE*), TXM1029 (Δ*lpp*, Δ*fadE* + *PA3286* + *lolCDE*), and TXM1111 [Δ*lpp*, Δ*fadE*, Δ*fabH*, Δ*lnt* + *PA3286* + *lolCDE* + *fadD(V451A)*] were grown in LB only, or LB supplemented with 100 μg/mL palmitate or octanoate. The final OD_{600} was measured after 19 hr.

As protein substrate, we chose a variant of the abundant lipoprotein Lpp: Lpp(K58A)-Strep-tag, which we have previously purified and analyzed by MALDI-TOF MS (8). The K58A mutation abolishes the point of covalent crosslinking of Lpp to peptidoglycan, traditionally allowing for perturbations to the lipoprotein maturation system (9–12). The chromosomal *lpp* was deleted and the aforementioned variant expressed on plasmid pTM1100, driven by the native

promoter P_{lpp}. We also expressed LolCDE, the ABC transporter essential in trafficking lipoproteins from the inner to outer membrane in *E. coli*, on plasmid pTXM1026. Driven by the constitutive P_{Kan} promoter, overexpression of LolCDE, along with the K58A mutation in Lpp, allowed us to delete Lnt (10), yielding a strain that produces only DA-LP to be used as substrate in a reaction with Lit.

Next, we chose to label lipoproteins with octanoic-7, 7, 8, 8, 8,-d_5 acid, a fatty acid featuring five terminal deuterium atoms, to differentiate labeled Lpp from unlabeled by MS (see Figure 5-3 and Table 5-1). However, as *E. coli* will not readily take up exogenous short chain fatty acids, we deleted the fatty acid metabolism regulatory protein FadR (13). This results in higher expression of the outer membrane fatty acid transporter FadL, allowing for increased uptake of the labeled octanoate from growth media (13, 14). Once inside the cell, esterification of the labeled octanoate with CoA was promoted by overexpression of the increased-activity variant FadD(V451A) on pTM1100 (15, 16). To prevent entry of the labeled octanoate into the β-oxidation cycle, which would lead to catabolism of the octanoate and random distribution of the deuterium atoms as they are recycled by the cell, the first enzyme in this pathway FadE was deleted (17). Instead, the *Pseudomonas aeruginosa* protein PA3286, encoded on pTXM1026, directly shunts the intact labeled octanoate into the fatty acid elongation cycle (18). Additionally, FabH was deleted to inhibit incorporation of *de novo* fatty acids into the elongation pathway, leaving the labeled octanoate shunted by PA3286 as the main source of fatty acids for this strain (19). Once elongated, the labeled fatty acids are incorporated into membrane phospholipids by PlsB and PlsC, then transferred to lipoproteins by Lgt (1, 20).

Growth of the final strain TXM1111 is over 20-fold less robust when compared to that of wildtype BW25113 cells and intermediate strains TXM1018 and TXM1029 due to the number and nature of its mutations (Figure 5-2B). However, addition of octanoate, but not palmitate, to both liquid (Figure 5-2B) and solid media (not shown) moderately rescued growth, attesting to

the substrate specificity of PA3286 and its crucial role in the viability of TXM1111. Ultimately, this complex strain results in acyl-labeled, DA-LP, of which we focused only on Lpp(K58A)-Strep-tag for this study, referred to as DA*-Lpp when labeled and DA-Lpp when unlabeled.

DA*-Lpp from strain TXM1111 is mostly dually-labeled

MS analyses were performed on DA*-Lpp purified from TXM1111 to determine the level of labeling (ie. one labeled acyl chain vs. two labeled chains). Incorporation of only one labeled lipid into DA*-Lpp would complicate interpretation of downstream results, as the initial position of the acyl chain (*sn*-1 or *sn*-2) would be unclear. It would also be difficult to determine if any unlabeled acyl chains on the N-terminus of the product lyso-Lpp came from the DA*-Lpp precursor or from unlabeled phospholipids provided in the *in vitro* reaction. Therefore, DA*-Lpp with both acyl chains deuterium-labeled is preferred for this experiment.

MADLI-TOF MS of trypsinized DA*-Lpp revealed a series of peaks corresponding to different levels of labeling, varying acyl chain lengths, and adducts, shown in Figure 5-3 and further detailed in Table 5-1. The peak of most interest appears at *m/z* 1167.8, which can be assigned to the dually-labeled N-terminal lipopeptide of Lpp with one monounsaturated and one saturated C_{16} acyl chains (Figure 5-3A). The partner sodium adduct of this peak is also present at *m/z* 1189.7. A low abundance peak corresponding to the singly-labeled lipopeptide can be observed at *m/z* 1162.6. Notably missing from this spectrum are peaks of the unlabeled lipopeptide, which appears at *m/z* 1157.6 in the spectrum of DA-Lpp isolated from strain KA775, an *lnt*-null suppressor mutant that spontaneously arose in early attempts to construct a labeling strain (Figure 5-3D) (8). Together, this indicates that the majority of DA*-Lpp from strain TXM1111 is dually-labeled.

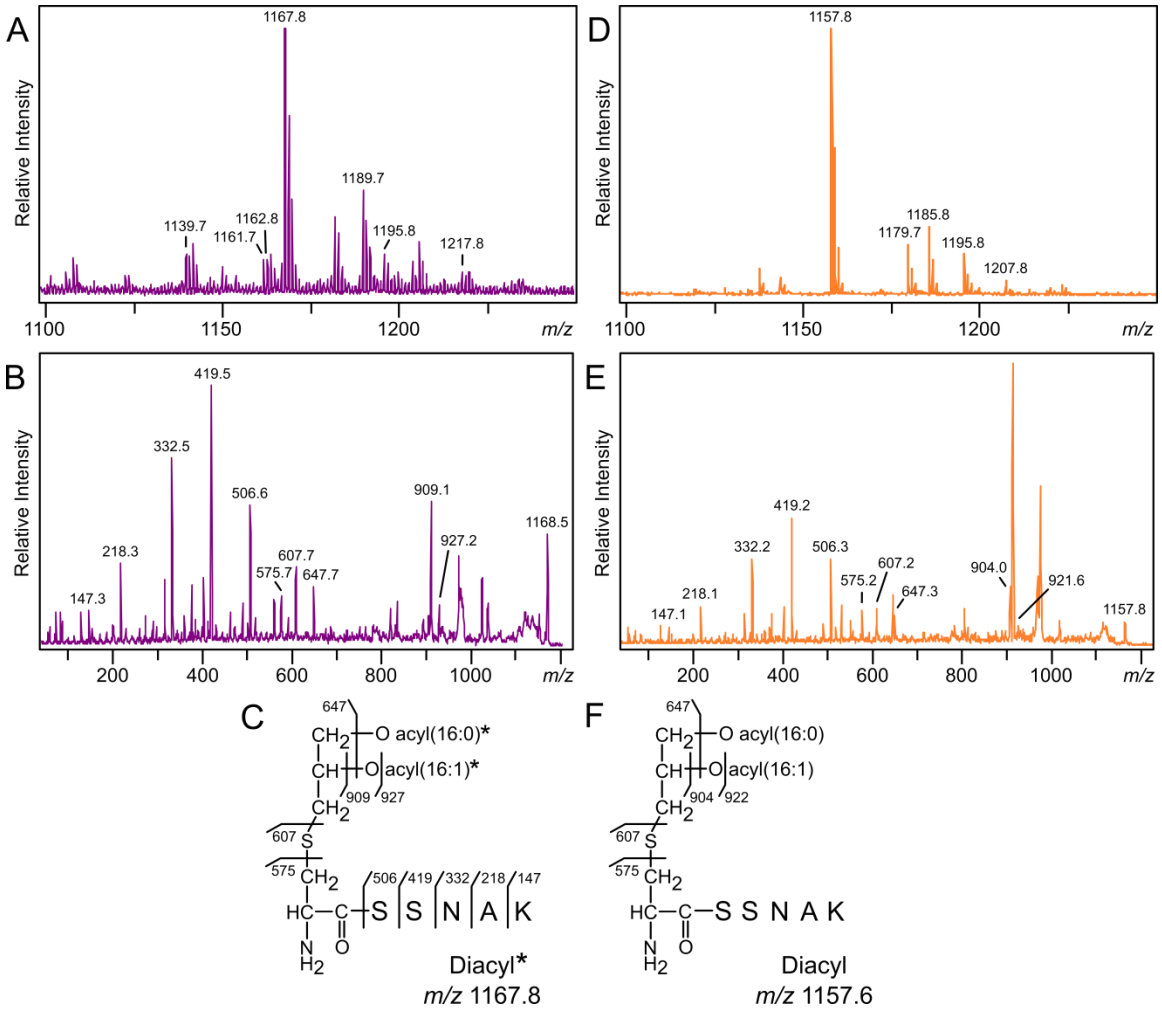


Figure **5-3**: MALDI-TOF MS of DA*-Lpp purified from strains TXM1111 and KA775. (A, D) Trypsinized Lpp lipopeptides purified from strain TXM1111 grown with terminally-deuterated octanoate (A) and the *lnt*-null suppressor mutant KA775 (D) were analyzed by MALDI-TOF MS. The major peaks are detailed in Table 3. (B, C) MS/MS spectra of the *m/z* 1167.8 parent ion (B) was used to elucidate the N-terminal structure of Lpp as DA-modified with two labeled acyl chains (C). (E, F) MS/MS spectra of the *m/z* 1157.8 parent ion (E) elucidates the N-terminal structure of unlabeled DA-Lpp (F).

Table **5-1**: Assignment of Prominent Ions from MS Analyses of DA(*)-Lpp

Theoretical *m/z*	**Corresponding Sodium Adduct**	**Acyl Chain Length and Saturation[3]**	**Total Acyl Chain Length**	**No. of Labeled Acyl Chains**
1139.6	1161.6	14:0, 16:1; or 14:1, 16:0	30:1	2
1157.6	1179.6	16:0, 16:1	32:1	0
1162.6	1184.6[1]	16:0, 16:1	32:1	1

1167.6	1189.6	16:0, 16:1	32:1	2
1185.6	1207.6	16:1, 18:0; or 16:0, 18:1	34:1	0
1190.6	1212.6[1]	16:1, 18:0; or 16:0, 18:1	34:1	1
1195.7[2]	1217.7	16:1, 18:0; or 16:0, 18:1	34:1	2

[1] Peaks were not observed in the spectra in Figure 5-3.
[2] May be a potassium adduct of *m/z* 1157.6 in the spectrum of KA775 (Figure 5-3D).
[3] The *sn*-1 or *sn*-2 positions of the chains were not determined.

For further confirmation of both lipoprotein structure and labeling, the *m/z* 1167.8 parent ion was fragmented by tandem MS/MS (Figure 5-3B). The *y*-series ions corresponding to the SSNAK peptide of Lpp were observed as expected, while ions 575.6, 607.7, and 647.7 are indicative of the lipopeptide's nonacylated α-amino terminus (Figure 5-3C). Ions *m/z* 575.6 and 607.7 respectively represent the dehydroalanyl peptide and the thiolated peptide, generated by the neutral loss of the diacylthioglycerol and DA moieties. Ion *m/z* 647.7 is consistent with the loss of both labeled fatty acids, a characteristic fragment peak that has traditionally only been observed of DA-LP (5, 8). The MS/MS spectrum of the *m/z* 1157.8 parent ion of DA-Lpp from KA775 shared all these peaks, confirming its peptide identity and N-terminal structure (Figure 5-3E and F). Distinct between the two spectra, additional ions at *m/z* 909.1 and 927.2 of DA*-Lpp from TXM1111 correspond to the parent ion having lost a labeled $C_{16:1}$ fatty acid ($C_{15}H_{29}COOH$) and a labeled $C_{16:1}$ ketene ($C_{14}H_{27}CH=C=O$), respectively (Figure 5-3B). While DA-Lpp from KA775 displays the same fragmentation products, the ions instead appear at *m/z* 904.0 and 921.6, reflecting the lack of deuterium-labeling (Figure 5-3E and F). Taken together, these peaks confirm the incorporation of two deuterium-labeled fatty acids into Lpp.

Lit is active *in vitro*

With DA*-Lpp in hand, we next needed reconstituted Lit enzyme and reaction conditions. We chose published reactions conditions previously used for studies on the activity of

Lnt: 50 mM Tris-HCl pH 7.2, 150 mM NaCl, and 0.1% Triton X-100 (21). To avoid purifying recombinant Lit or Lnt, we followed a procedure where the source of enzyme is inner membranes collected from *E. coli* cells overexpressing the enzyme of interest (22). No further conditions were explored for optimization.

To our knowledge, only MS can definitively confirm conversion of DA-LP to the lyso form, but this can be laborious and require large amounts of protein to yield confident results. To avoid MS analyses while testing reaction conditions on a small scale, we first analyzed reactions with unlabeled DA-Lpp(K58A)-Strep-tag isolated from KA775 by Western blot. DA-modified and lyso-form lipoproteins are indistinguishable by SDS-PAGE, however, since they are the same mass (8). Thus, we relied on the persistence of a DA-LP band when both Lnt and Lit are present as indirect evidence of Lit activity. In other words, in a reaction with just Lnt, most DA-Lpp substrate is converted to the triacyl form, resulting in a visible band shift to a higher molecular weight. If Lit is also present in the reaction, a band persists at the "diacylated" weight, presumably because Lit converts DA-Lpp to the lyso form, which cannot then be converted to the triacyl form by Lnt. Indeed, there is a significantly darker "diacylated" Lpp band in the Lnt/Lit reaction at the 24 hr time point when contrasted with the Lnt-only reaction. We assume that this band is the lyso form, suggesting activity of Lit in these conditions (Figure 5-4).

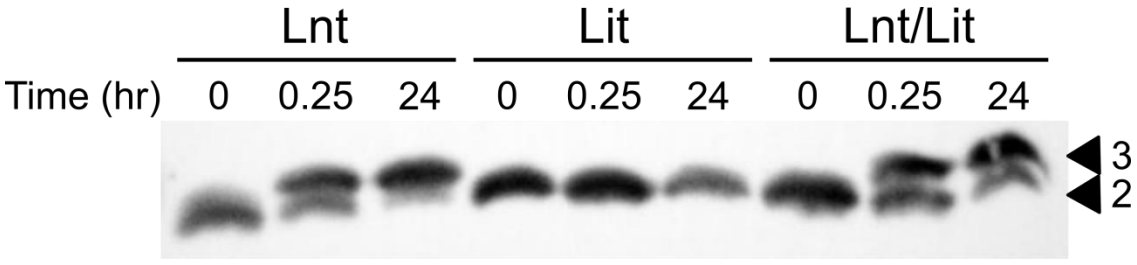


Figure **5-4**: *In vitro* activity of Lit. DA-Lpp was combined with inner membranes containing Lnt, Lit, or both Lnt/Lit for 24 hr, with aliquots removed at 0 and 0.25 hr. Reactions containing Lnt show conversion to the triacyl form ("3"), while the reaction containing both Lnt and Lit shows persistence of Lpp with two acyl chains ("2"), presumably the lyso form. The reaction with just Lit maintains a "diacyl" band throughout, but is also presumed to be the lyso form by the 24 hr time point.

Together with the acyl-labeled DA*-Lpp substrate, we have both components for the final experiment in which DA*-Lpp will be combined with membrane-reconstituted Lit. The resulting lyso-form Lpp will be analyzed by MS to determine the position of the labeled acyl chains, namely the *N*-acyl chain. Preliminary attempts have revealed technical difficulties with recovery of Lpp from the reaction mixture. Efforts are ongoing to troubleshoot this problem.

Partial membrane topology of Lit

As Lit belongs to a novel class of enzymes with no similarity to characterized acyltransferases, we sought to map its membrane topology in *E. coli*. According to several membrane topology mapping algorithms, Lit is predicted to have four transmembrane (TM) passes and two periplasmic loops, with both the N- and C-termini located in the cytoplasm (Figure 5-5A). To experimentally verify this prediction, we cloned a series of fusions of Lit to either β-galactosidase (LacZ) or alkaline phosphatase (PhoA) after each predicted TM pass, with the fusions made from the +9 residue of LacZ and +13 residue of PhoA so as to not encode their native signal peptides (11). The activity of the resulting strains was tested using a colorimetric assay on solid media, whereby colonies appear blue when LacZ is localized to the cytoplasm and PhoA in the periplasm (23). In line with topology predictions, LacZ activity was exhibited by the constructs with fusions following TM pass 2 (LitR129-LacZ) and pass 4 (LitK212-LacZ) (Figure 5-5B). PhoA activity was exhibited by the fusion following TM pass 1 (LitD52-PhoA) as predicted, however not after TM pass 3 (LitD173-PhoA) (Figure 5-5C). Additional fusion sites after residues L159, A165, L179, and M185 were tested, but not did show PhoA activity (data not shown). Similarly, a construct encoding the 18-residue linker VPDSYTQVASWTEPFPFC between LitD173 and PhoA also did not exhibit PhoA activity (data not shown) (24). Therefore,

as neither LacZ nor PhoA activity was observed at fusions following TM pass 3, we cannot fully verify the topology of Lit using this approach.

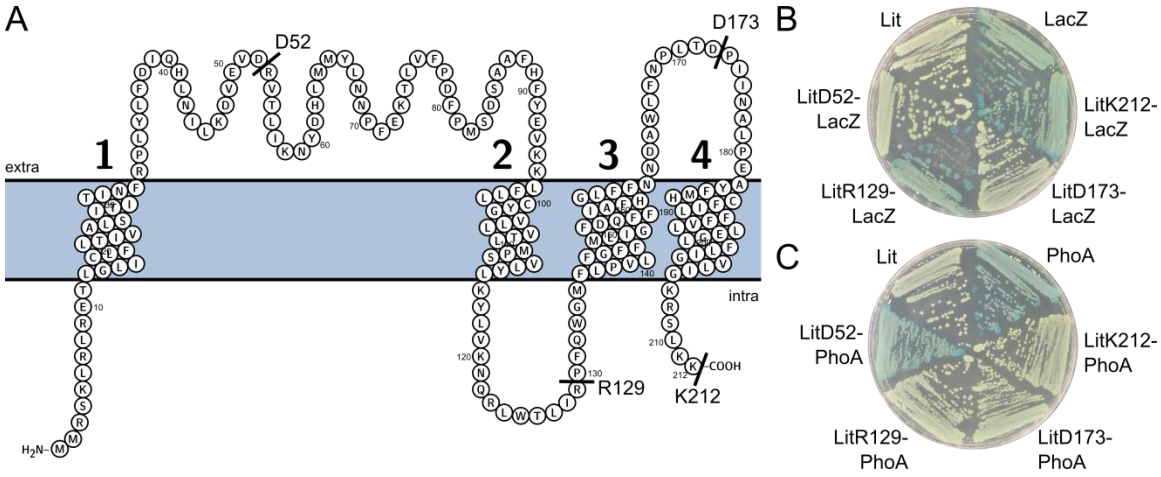


Figure **5-5**: Proposed membrane topology of Lit. (A) The membrane topology of Lit as predicted by Protter (25). The amino acids after which the LacZ and PhoA fusions to Lit were made are indicated. (B) Strains expressing full-length Lit (SS851), full-length LacZ (SS904), and the fusions LitD52-LacZ (SS951), LitR129-LacZ (SS935), LitD173-LacZ (SS932), and LitK212-LacZ (SS915) were struck on solid media containing X-Gal. Blue colonies, exhibited by cells expressing the full-length LacZ, LitR129-LacZ, and LitK212-LacZ, indicate active LacZ localized in the cytoplasm, in line with the predicted topology of Lit. (C) Strains expressing full-length Lit (SS851), full-length PhoA (SS964), and the fusions LitD52-PhoA (SS950), LitR129-PhoA (SS936), LitD173-PhoA (SS940), and LitK212-PhoA (SS916) were struck on solid media containing BCIP. Blue colonies, exhibited by cells expressing the full-length PhoA and LitD52-PhoA, indicate active PhoA localized in the periplasm, as expected. In contrast to the predicted topology of Lit, colonies with LitD173-PhoA did not show PhoA activity.

Discussion

Discovery of lyso-form lipoproteins and the enzyme that catalyzes their synthesis raised questions about the mechanism of Lit. While preliminary evidence supports intramolecular transacylation activity (8), further studies were necessary to demonstrate transfer of the ester-linked lipid from the DA moiety of the lipoprotein to the *N*-acyl position. Herein, we constructed a complex strain of *E. coli* that incorporates exogenously-added, deuterium-labeled lipids into its DA-LP. MALDI-TOF of DA*-Lpp purified from this strain indicated labeling on both acyl chains of the DA moiety, an important factor for downstream experiments. By SDS-PAGE,

membrane-reconstituted Lit appears to be active *in vitro*. These two components have laid the foundation for an experiment in which lyso*-Lpp from a reaction of Lit and DA*-Lpp will be analyzed by MS to determine the final positions of the labeled acyl chains. Ultimately, this may confirm the mechanism of Lit as a lipoprotein intramolecular transacylase.

Lit, and thus the lyso form, is only found in certain low-GC Firmicutes (5). Various Firmicutes produce additional *N*-terminally modified lipoproteins, including the conventional triacyl form, but also the peptidyl and *N*-acetyl forms (5). Overall, there is a greater trend towards *N*-terminal lipoprotein modification that likely arose as a post-speciation edit to the established Lgt-Lsp modification pathway, presenting an opportunity for several different enzymes and/or mechanisms to evolve. Assuming that Lit functions by intramolecular transacylation, what is its advantage over other known mechanisms of *N*-acylation, such as that of *E. coli* Lnt? Perhaps the answer lies in the reaction byproducts. Lnt, which transfers an acyl chain from a membrane phospholipid to the lipoprotein α-amino terminus (3), results in not just the triacylated lipoprotein, but a lysophospholipid as well. Physically similar to detergents, lysophospholipids can be cytolytic and perturb membrane integrity (26). In *E. coli*, lysophospholipids are reacylated by an acyltransferase enzyme (Aas), a system that is absent in Gram-positive bacteria (27). In light of this, *N*-acylation by intramolecular transacylation may be advantageous, as it avoids formation of lysophospholipids.

Lit represents a novel class of proteins that, despite also catalyzing lipoprotein *N*-acylation, shares no homology with *E. coli* Lnt. Therefore, characterization of the membrane topology of Lit may provide further insight into its mechanism. Predicted to have four TM passes with both N- and C-termini in the cytoplasm, we experimentally verified localization of the loop between TM passes 1 and 2 as periplasmic, and the C-terminus and loop between TM passes 2 and 3 as cytoplasmic (Figure 5-5). We were unable to verify localization of the predicted periplasmic loop between TM passes 3 and 4 using the colorimetric reporter approach taken

herein. It is possible that this loop is embedded within the membrane in a way that is not conducive to PhoA fusion activity. An embedded intracellular loop was described of *E. coli* Lnt (11, 28, 29) and was recently confirmed by crystallography (30–32). It is also possible that fusions within this loop leave Lit susceptible to proteolysis, explaining the lack of colorimetric signal. Confirmation of protein expression by Western blot would alleviate this concern. Alternate methods of determining membrane topology, such as the substituted cysteine accessibility method (SCAM) (33), may successfully verify orientation of this loop.

Though not explored in this current study, identification of the essential amino acids of Lit may prove intriguing. Studies on the mechanism of *E. coli* Lnt show that Lnt functions by forming an acyl-enzyme intermediate at the catalytic cysteine C387 that is located on a large extracellular loop (11, 21, 29, 34). *E. faecalis* Lit contains three cysteine residues of its 212 total amino acids, C16, C100, and C187, predicted to be membrane-embedded in TM passes 1, 2, and 4, respectively (see Figure 5-5). In alignments of the primary sequence of the Lit1 and Lit2 proteins used to generate the phylogenetic tree in Figure 4-1, only C187 appears to be the most conserved, along with several residues of TM pass 3 and the loop between TM passes 3 and 4 (Figure 5-6). This may be indicative of the active site of Lit, whose localization within the membrane or near the membrane interface would theoretically place C187 in close proximity to the lipoprotein N-terminus, where an acyl-enzyme intermediate could form as the ester-linked lipid of the DA moiety is transferred to the α-amino position.

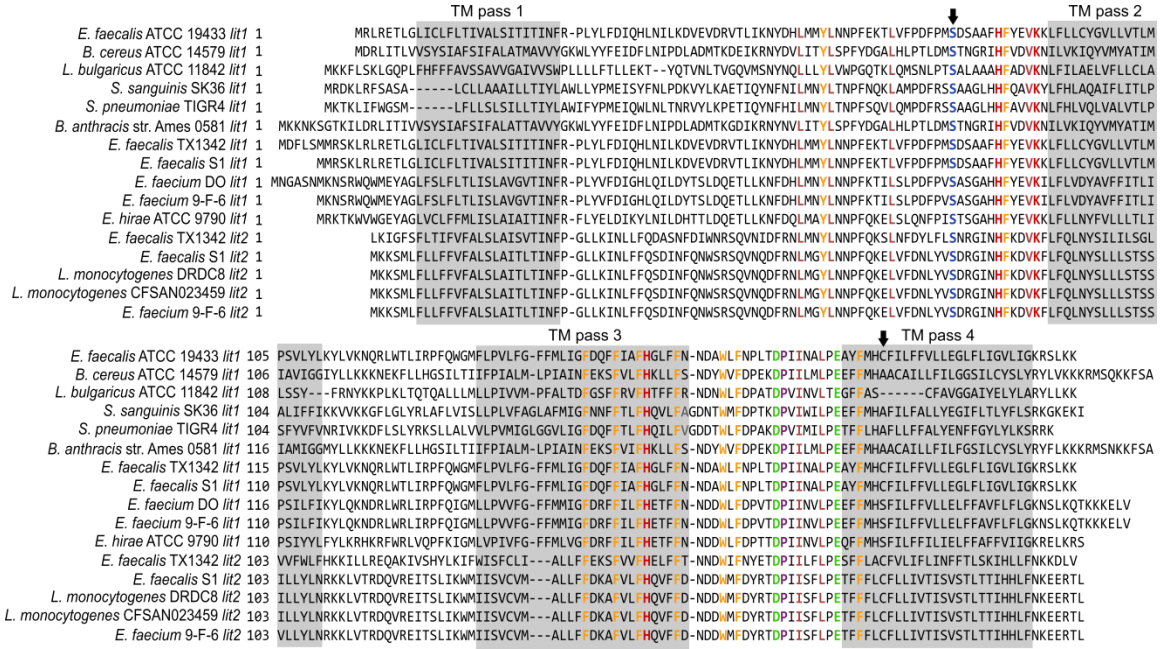


Figure **5-6**: Primary sequence alignment of Lit1 and Lit2 proteins. The primary sequence alignment of the chromosomal Lit1 and accessory Lit2 of the indicated strains is shown, as generated by STRAP (35). The predicted transmembrane (TM) passes are boxed in gray. Amino acids that are 100% conserved between all sequences are colored. Residues S83 and C187 are indicated with an arrow.

Similarly conserved across several Lit proteins is serine S83 of the extracellular loop between TM passes 1 and 2 (Figure 5-6). Interestingly, a catalytic serine has been described for Lnt of the Gram-positive *Mycobacterium bovis* (36). However, the C387S mutation in *E. coli* Lnt yields an inactive protein that is able to form the acyl-enzyme intermediate, but unable to transfer the lipid to a lipoprotein (29, 34). Studies have shown that Firmicutes favor cysteine-exclusion from their exported proteins, as they lack machinery to reduce wayward disulfide bond formation in extracellular proteins, namely the disulfide oxidoreductases of the Dsb family, and exposed cysteine residues are more susceptible to deleterious modifications (37). In light of this, a catalytic serine may be advantageous over a catalytic cysteine for Lit activity in *E. faecalis* and other lyso-form producing bacteria.

Until the position of the labeled acyl chains on lyso*-Lpp are determined, we cannot definitively conclude that Lit functions by intramolecular transacylation. It remains possible that Lit is a bifunctional enzyme that concurrently catalyzes lipoprotein *O*-deacylation and *N*-acylation to create the lyso form. We find this unlikely, however, due to the size of Lit. Regardless of which mechanism it uses, considering what is currently known about lipoprotein *N*-acylation, Lit is a unique enzyme worthy of further study. As more is learned about the various lipoprotein forms in Firmicutes and how they are made, we may see additional enzymes and mechanisms arise in the field.

Materials and Methods

Bacterial strains and growth conditions

E. coli strains used in this study, listed in Table 5-2, are derivatives of either BW25113 or BL21(DE3). All strains were grown in LB-Miller medium at 37°C with agitation. Antibiotic markers were selected with carbenicillin (100 μg/mL), apramycin (100 μg/mL), spectinomycin (50 μg/mL), tetracycline (5 μg/mL), kanamycin (15 or 30 μg/mL), and trimethoprim (50 μg/mL). Where appropriate, fatty acids were supplemented to 100 μg/mL.

Table **5-2**: Bacterial strains and plasmids used in this study.

Strain or Plasmid[1]	Relevant Genotype/Phenotype[2]	Reference
Escherichia coli BL21(DE3) strains		
KA729	pET22b(+)-*Eflit* (from *E. faecalis* ATCC 19433)	This study
KA736	*lpp*::Cm^r + pET22b(+)-*Eflit*	This study
KA801	*lpp*::Cm^r *lnt*::Spt^r *chiQ*::Apr^r + pET22b(+)-*Eflit*	This study
KA818	*lpp*::Cm^r *lnt*::Spt^r *chiQ*::Apr^r + pET22b(+)-*Eflit* + pKA810	This study

Escherichia coli BW25113 strains		
TXM327	*lpp*::Cm^r	(8)
TXM541	*gut*::Kan^r-*rrnB* TT-*araC*-P_{BAD}-*lnt*, *lnt*::Spt^r, *chiQ*::Apr^r	(8)
KA775	TXM327 *lnt*::Spt^r + pKA763 (suppressor mutant)	This study
KA827	*phoA*::Kan^r	Keio Collection
KA845	*fadR*::Tmp^r	This study
SS851	KA827 + pET22b(+)-*Eflit*; Car^r	This study
SS904	KA827 + pET22b(+)-*lacZ*; Car^r	This study
SS915	KA827 + pET22b(+)-*Eflit*K212-(+9)*lacZ* fusion; Car^r	This study
SS916	KA827 + pET22b(+)-*Eflit*K212-(+13)*phoA* fusion; Car^r	This study
SS932	KA827 + pET22b(+)-*Eflit*D173-(+9)*lacZ* fusion; Car^r	This study
SS935	KA827 + pET22b(+)-*Eflit*R129-(+9)*lacZ* fusion; Car^r	This study
SS936	KA827 + pET22b(+)-*Eflit*R129-(+13)*phoA* fusion; Car^r	This study
SS940	KA827 + pET22b(+)-*Eflit*D173-(+13)*phoA* fusion; Car^r	This study
SS950	KA827 + pET22b(+)-*Eflit*D52-(+13)*phoA* fusion; Car^r	This study
SS951	KA827 + pET22b(+)-*Eflit*D52-(+9)*lacZ* fusion; Car^r	This study
SS964	KA827 + pET22b(+)-*phoA*; Car^r	This study
TXM1014	KA845 *lpp*::Kan^r flanked by FRT sites	This study
TXM1015	TXM1014 *lpp*::FRT	This study
TXM1018	TXM1015 *fadE*::Tet^r	This study
TXM1029	TXM1015 + pTXM1026	This study
TXM1032	TXM1029 *fabH*::Cm^r	This study
TXM1037	TXM1032 *lnt*::Spt^r	This study
TXM1058	TXM1037 with *fabH*::Cm^r flanked by FRT sites	This study
TXM1067	TXM1058 *fabH*::FRT	This study
TXM1111	TXM1067 + pTM1100	This study
Plasmids		
pCL25, pBBR1	General cloning vectors	(38, 39)
pKA522	pCL25 with *lpp*(K58A)-Strep-tag; Kan^r	(8)
pKA763	pKA522 with *PA3286* (from *P. aeruginosa*)	This study
pKA810	pCL25 with *lpp*(K58A)-Strep-tag; Tmp^r	This study
pTXM1026	pBBR1 with P_{Kan}-*lolCDE-PA3286*; Kan^r	This study
pTM1100	pCL25 with P_{lpp}-*lpp*(K58A)-Strep-tag-*fadD*(V451A); Cm^r	This study

[1] Letters designate the primary constructor of the respective strain and/or plasmid: KA, Krista Armbruster; TXM or TM, Timothy Meredith; SS, Sarah Staskiewicz.

[2] Resistance phenotypes: Apr^r, apramycin; Car^r, carbenicillin; Cm^r, chloramphenicol; Kan^r, kanamycin; Spt^r, spectinomycin; Tet^r, tetracycline; Tmp^r, trimethoprim.

Construction of deletion strains and plasmids

Gene deletions in *E. coli* were constructed using the Red-recombinase method and transduced into recipient strains by P1*vir* when appropriate (40). Plasmids were assembled using the In-Fusion HD cloning kit (Clontech). Primers used in this study are listed in Table 5-3.

Table **5-3**: Primers used in this study.

Primer Name	**Description**	**Primer Sequence**
TM692	5'-*lnt*::Spt^r	GCCTCATTAATTGAACGCCAGCGCATTCGCCTG CTGCTGGCGATTCCCCTGCTCGCGCAG
TM693	*lnt*::Spt^r-3'	CACAGCAGCAAAACCAAACAATGCCGTCAGCA CCCACAGCGGGCTTGAACGAATTGTTAG
KA1053	5'-pET22b(+) *Eflit*	CTTTAAGAAGGAGATATACATATGATGAGGTCG AAATTGCG
KA1054	pET22b(+) *Eflit*-3'	GTGGTGGTGGTGGTGGTGGCCGCCGCCTTTTTTT AAAGAACGTTTTCCG
TM1121	5'-*fadE*::Tet^r	GATTTTGAGTATTCTCGCTACGGTTGTCCTGCTC GGCGCGTTCAAATGTAGCACCTGAAG
TM1122	*fadE::*Tet^r-3'	GTCGCCAGCTCTTCCGGATCAAAGTCATCAACG TTAATACTGGAATCCGTTAGCGAGGTG
TM1123	5'-*fadE* upcheck	CCATATCATCACAAGTGGTCAGACC
TM1124	*fadE* downcheck-3'	CAACTTTCCGCACTTTCTCCGGCAAC
TM1130	5'-pKA522 PA3286	CTGTGAAGTGAAGCTGAGAGAAGAGGAACAAC TATGCATAAAGCC
TM1131	pKA522 PA3286-3'	GGACTCTGGGGTTCGGCCACGATCAGTGTTTAC
KA1245	5'-*fadR::*Tmp^r	GTCATTAAGGCGCAAAGCCCGGCGGGTTTCGC GGAAGAGTACGGATAGACGGCATGCACG
KA1246	*fadR::*Tmp^r-3'	CCAAATCTCGCCACTCTCATGCCCATAGCGAC GCACTGTGAATCCGTTGCTGCCAC
KA1247	5'-*fadR* upcheck	GGTATGATGAGTCCAAC
KA1248	*fadR* downcheck-3'	CATCGAGTTGCTGGAACG
TM1476	5'-*lolCDE*	GACCGACCAAAAGCTACAGCAACCAGACGG
TM1477	*lolCDE*-3'	TTACTGTCGGGAATTCTTGTTTTAATGTACTGCC
TM1614	5'-*fabH* upcheck	GGACACCTCTGACGAGTGGATTG
TM1615	*fabH* downcheck-3'	CACGATCGGTTGGATCGCAGGTG

TM1689	5'-*fabH::*Cm^{r}FRT	CGCCACATTGCCGCGCCAAACGAAACCGTTTC AACCATGGGCTATGAATATCCTCCTTAG
TM1690	*fabH::*Cm^{r}FRT-3'	CCGCCCCAGATTTCACGTATTGATCGGCTACGC TTAATGCATGTAGGCTGGAGCTGCTTC
TM1722	5'-*fadD(*V451A)	GCTAATGCGCAGGAATCCTTCTTCATCC
TM1723	*fadD*(V451A)-3'	TTCCTGCGCATTAGCGATCGTAAAAAAGAC
SS1314	pET22b(+) *Eflit*-3'	GGTGGTGGTGCTCGAGTTCACTTTTCTCTC
SS1562	5'-pET22b(+) *phoA*	GAAGGAGATATACATATGAAACAAAGCACTA TTGC
SS1316	pET22b(+) *phoA*-3'	GGTGGTGGTGCTCGACGCGGTTTTATTTCA
SS1317	5'-pET22b(+) *lacZ*	CTTTAAGAAGGAGATATACAAACAGCTATGAC CATG
SS1318	pET22b(+) *lacZ*-3'	GGTGGTGGTGCTCGAGTTCAAGCAAGCTTTT
SS1502	5'-(+9) *lacZ*	GCCGTCGTTTTACAACG
SS1503	5'-(+13) *phoA*	TTACTGTTTACCCCTGTG
SS1275	*Eflit*K212 fusion-3'	TTTTTTTAAAGAACGTTTTCCG
SS1276	5'-*Eflit*K212 *lacZ* fusion	CGTTCTTTAAAAAAAGCCGTCGTTTTACAAC
SS1504	5'-*Eflit*K212 *phoA* fusion	CACAGGGGTAAACAGTAATTTTTTTAAAGAACG
SS1505	*Eflit*D52 *lacZ* fusion-3'	CGTTGTAAAACGACGGCATCAACCTCTACGTC
SS1506	*Eflit*D52 *phoA* fusion-3'	CACAGGGGTAAACAGTAAATCAACCTCTACGTC
SS1507	*Eflit*R129 *lacZ* fusion-3'	CGTTGTAAAACGACGGCCCGTATCAATGTCCA
SS1508	*Eflit*R129 *phoA* fusion-3'	CACAGGGGTAAACAGTAACCGTATCAATGTCCA
SS1509	*Eflit*D173 *lacZ* fusion-3'	CGTTGTAAAACGACGGCATCAGTCAGCGGATT
SS1510	*Eflit*D173 *phoA* fusion-3'	CACAGGGGTAAACAGTAAATCAGTCAGCGGATT

Labeling and column purification of Strep-tagged Lpp(K58A)

Strain TX1111 constructed to produce labeled Lpp is described in the Results and illustrated in Figure 5-2. For labeling, octanoic-7, 7, 8, 8, 8,-d_5 acid (CDN Isotopes) was added to solid media at 100 μg/mL. After incubation of strain TXM1111, colonies were scraped and collected from the agar surface. Lpp(K58A)-Strep-tag was purified from the resulting biomass using the method described previously (8). Briefly, cells were passed through a French pressure cell four times at 14,000 lb/in^2. Membranes were collected from the resulting lysate by ultracentrifugation (110,000 x g, 90 minutes, 4°C) and resuspended to 1 mg/mL of protein in 10 mM HEPES pH 7.4 with 1 mM EDTA. The resuspended membranes were supplemented with 2%

Triton X-100 (v/v), incubated at room temperature for 30 min with gentle shaking, then clarified by centrifugation (20,000 x g, 20 minutes, room temperature). The supernatant was diluted with HEPES buffer with 0.5 mM *n*-dodecyl-β-D-maltopyranoside (DDM) before being passed over StrepTactin Sepharose resin (IBA Biosciences). Protein was eluted following manufacturer's instructions, with all wash and elution buffers containing 0.5 mM DDM. Protein was aliquoted and stored at -80°C.

Growth assay with fatty acids

Single colonies of strains BW25113, TXM1018, TXM1029, and TXM1111 were chosen into LB media. For TXM1111, media contained octanoate. Once cultures had become cloudy, cells were washed once with LB, inoculated into fresh LB to an OD_{600} of 0.0001, and supplemented with 100 μg/mL palmitic or octanoic acid. Cultures were grown for 19 hr until the final OD_{600} was measured.

Reconstitution of inner membranes for activity assays

As a source of Lnt, Lit, and Lnt/Lit, inner membranes of KA893, KA801, and KA736 respectively were harvested as previously described, with some modifications (8). One liter of cells were grown to an OD_{600} of 1.0, harvested by centrifugation, washed once with phosphate-buffered saline (PBS), and stored frozen until use. Cell pellets were resuspended in ~35 mL of 50 mM Tris, pH 7.8 (adjusted with HCl) containing 1 mM EDTA and disrupted by passage through a French pressure cell three times at 14,000 lb/in^2. Phenylmethylsulfonyl fluoride (PMSF) was then added to 1 mM and the lysates centrifuged at 3,200 x g for 10 min at 4°C to pellet unbroken cells. Total membranes were collected by ultracentrifugation at 110,000 x g for 90 min at 4°C,

then homogenized in 3 mL of 10 mM Hepes, pH 7.4 (adjusted with NaOH) with a 26-gauge needle. The membrane suspension was layered on top of a discontinuous sucrose gradient as described previously and centrifuged at 80,000 x g for 19 hr at 4°C. The top volume of sucrose was discarded and the visible inner membrane layer (~4-6 mL) transferred to a new ultracentrifuge tube before the volume was adjusted to ~40 mL with Hepes buffer. The solution was centrifuged again at 110,000 x g for 90 min at 4°C. The final inner membrane pellet was resuspended in Hepes buffer supplemented with 10% glycerol and normalized across the three samples to the same protein concentration as determined by Bradford regent. Aliquots were stored frozen at -80°C.

Activity assays

The reaction buffer, composed of 50 mM Tris, pH 7.2, 150 mM NaCl, and 0.1% Triton X-100, was optimized previously by Hillmann et al. in activity assays with purified Lnt enzyme (21). To the reaction, inner membrane fractions containing Lnt, Lit, or Lnt/Lit were added to a final concentration of 0.75 mg/mL and sonicated for 1 min in a sonicator bath. Purified DA-Lpp(K58A) was then added to 0.004 mg/mL and sonicated again before incubation at 37°C. At the indicated time intervals, 10-µL samples were removed and the reaction halted by the addition of loading buffer and flash freezing. Samples were separated by SDS-PAGE over a 16.5% Tris-tricine gel and immunblotted for Lpp(K58A)-Strep-tag using Precision Protein StrepTactin-horseradish peroxidase conjugate (Bio-Rad) as described previously (8).

MALDI-TOF

For MS analysis of DA-Lpp substrate, purified Lpp was transferred from a 16.5% Tris-tricine SDS-PAGE gel to nitrocellulose membrane. The band corresponding to Lpp was excised, digested in a trypsin solution overnight, then sequentially washed with 0.1% trifluoroacetic acid (TFA), 10% and 20% acetonitrile before eluting the N-terminal lipopeptides in 10 mg/mL α-cyano-4-hydroxycinnamic acid (CHCA) in 2:1 chloroform-methanol, as described previously (8, 41). Multiple layers of the lipopeptide in CHCA were deposited onto a steel target plate. MALDI-TOF MS was conducted in positive reflectron mode on an Ultraflextreme (Bruker Daltonics) mass spectrometer. Tandem MS/MS spectra were acquired on the same instrument in LIFT mode. A protocol for MS analysis of Lpp recovered from *in vitro* reactions is currently being developed.

LacZ and PhoA activity assays

Listed in Table 5-2, a series of constructs were made fusing *lit* to either β-galactosidase (*lacZ*) or alkaline phosphatase (*phoA*) lacking their N-terminal signal peptides (from amino acid +13 for PhoA and +9 for LacZ) (11). Fusions were made after each predicted transmembrane pass of Lit. Constructs encoding the full-length *lacZ* and *phoA* genes were also generated. For colorimetric activity assays on solid media, the indicated strains were struck on LB agar containing 100 μg/mL carbenicillin and 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). To assay for LacZ activity, media also contained 40 μg/mL X-Gal, and for PhoA activity, 40 μg/mL 5-bromo-4-chloro-3'-indoly phosphate (BCIP). Plates were imaged after ~20 hr incubation at 37°C.

References

1 **Sankaran K**, **Wu HC**. 1994. Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. J Biol Chem **269**:19701–19706.

2. **Hussain M**, **Ichihara S**, **Mizushima S**. 1982. Mechanism of signal peptide cleavage in the biosynthesis of the major lipoprotein of the *Escherichia coli* outer membrane. J Biol Chem **257**:5177–5182.

3. **Gupta SD**, **Wu HC**. 1991. Identification and subcellular localization of apolipoprotein N-acyltransferase in *Escherichia coli*. FEMS Microbiol Lett **62**:37–41.

4. **Nakayama H**, **Kurokawa K**, **Lee BL**. 2012. Lipoproteins in bacteria: structures and biosynthetic pathways. FEBS J **279**:4247–4268.

5. **Kurokawa K**, **Ryu K-H**, **Ichikawa R**, **Masuda A**, **Kim M-S**, **Lee H**, **Chae J-H**, **Shimizu T**, **Saitoh T**, **Kuwano K**, **Akira S**, **Dohmae N**, **Nakayama H**, **Lee BL**. 2012. Novel bacterial lipoprotein structures conserved in low-GC content Gram-positive bacteria are recognized by Toll-like receptor 2. J Biol Chem **287**:13170–13181.

6. **Kurokawa K**, **Lee H**, **Roh K-B**, **Asanuma M**, **Kim YS**, **Nakayama H**, **Shiratsuchi A**, **Choi Y**, **Takeuchi O**, **Kang HJ**, **Dohmae N**, **Nakanishi Y**, **Akira S**, **Sekimizu K**, **Lee BL**. 2009. The Triacylated ATP Binding Cluster Transporter Substrate-binding Lipoprotein of *Staphylococcus aureus* Functions as a Native Ligand for Toll-like Receptor 2. J Biol Chem **284**:8406–8411.

7. **Asanuma M**, **Kurokawa K**, **Ichikawa R**, **Ryu KH**, **Chae JH**, **Dohmae N**, **Lee BL**, **Nakayama H**. 2011. Structural evidence of α-aminoacylated lipoproteins of *Staphylococcus aureus*. FEBS J **278**:716–728.

8. **Armbruster KM**, **Meredith TC**. 2017. Identification of the Lyso-Form N-Acyl Intramolecular Transferase in Low-GC Firmicutes. J Bacteriol **199**:e00099-17.

9. **Yakushi T**, **Tajima T**, **Matsuyama S**, **Tokuda H**. 1997. Lethality of the Covalent Linkage between Mislocalized Major Outer Membrane Lipoprotein and the Peptidoglycan of *Escherichia coli*. J Bacteriol **179**:2857–2862.
10. **Narita S ichiro**, **Tokuda H**. 2011. Overexpression of LolCDE allows deletion of the *Escherichia coli* gene encoding apolipoprotein N-acyltransferase. J Bacteriol **193**:4832–4840.
11. **Robichon C**, **Vidal-Ingigliardi D**, **Pugsley AP**. 2005. Depletion of apolipoprotein N-acyltransferase causes mislocalization of outer membrane lipoproteins in *Escherichia coli*. J Biol Chem **280**:974–983.
12. **Grabowicz M**, **Silhavy TJ**. 2017. Redefining the essential trafficking pathway for outer membrane lipoproteins. Proc Natl Acad Sci U S A **114**:4769–4774.
13. **DiRusso CO**, **Metzger AK**, **Heimert TL**. 1993. Regulation of transcription of genes required for fatty acid transport and unsaturated fatty acid biosynthesis in *Escherichia coli* by FadR. Mol Microbiol **7**:311–322.
14. **Black PN**. 1991. Primary sequence of the *Escherichia coli* fadL gene encoding an outer membrane protein required for long-chain fatty acid transport. J Bacteriol **173**:435–442.
15. **Ford TJ**, **Way JC**. 2015. Enhancement of *E. coli* acyl-CoA synthetase FadD activity on medium chain fatty acids. PeerJ **3**:e1040.
16. **Black PN**, **DiRusso CC**, **Metzger AK**, **Heimert TL**. 1992. Cloning, sequencing, and expression of the fadD gene of *Escherichia coli* encoding acyl coenzyme A synthetase. J Biol Chem **267**:25513–25520.
17. **Campbell JW**, **Cronan JE**. 2002. The enigmatic *Escherichia coli* fadE gene is yafH. J Bacteriol **184**:3759–3764.
18. **Yuan Y**, **Leeds JA**, **Meredith TC**. 2012. *Pseudomonas aeruginosa* directly shunts β-oxidation degradation intermediates into de novo fatty acid biosynthesis. J Bacteriol **194**:5185–5196.

19. **Lai CY**, **Cronan JE**. 2003. β-Ketoacyl-Acyl Carrier Protein Synthase III (FabH) Is Essential for Bacterial Fatty Acid Synthesis. J Biol Chem **278**:51494–51503.
20. **Zhang Y-M**, **Rock CO**. 2008. Thematic review series: Glycerolipids. Acyltransferases in bacterial glycerophospholipid synthesis. J Lipid Res **49**:1867–1874.
21. **Hillmann F**, **Argentini M**, **Buddelmeijer N**. 2011. Kinetics and phospholipid specificity of apolipoprotein N-acyltransferase. J Biol Chem **286**:27936–27946.
22. **Trent MS**, **Ribeiro AA**, **Lin S**, **Cotter RJ**, **Raetz CRH**. 2002. An Inner Membrane Enzyme in *Salmonella* and *Escherichia coli* That Transfers 4-Amino-4-deoxy-l-arabinose to Lipid A. J Biol Chem **276**:43122–43131.
23. **Manoil C**. 1991. Analysis of membrane protein topology using alkaline phosphatase and beta-galactosidase gene fusions. Methods Cell Biol **34**:61–75.
24. **Drew D**, **Sjöstrand D**, **Nilsson J**, **Urbig T**, **Chin C-N**, **de Gier J-W**, **von Heijne G**. 2002. Rapid topology mapping of *Escherichia coli* inner-membrane proteins by prediction and PhoA/GFP fusion analysis. Proc Natl Acad Sci U S A **99**:2690–2695.
25. **Omasits U**, **Ahrens CH**, **Müller S**, **Wollscheid B**. 2014. Protter: Interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics **30**:884–886.
26. **Weltzien HU**. 1979. Cytolytic and membrane-perturbing properties of lysophosphatidylcholine. Biochim Biophys Acta **559**:259–287.
27. **Lin Y**, **Bogdanov M**, **Lu S**, **Guan Z**, **Margolin W**, **Weiss J**, **Zheng L**. 2018. The phospholipid-repair system LplT/Aas in Gram-negative bacteria protects the bacterial membrane envelope from host phospholipase A 2 attack. J Biol Chem **293**:3386–3398.
28. **Gélis-Jeanvoine S**, **Lory S**, **Oberto J**, **Buddelmeijer N**. 2015. Residues located on membrane-embedded flexible loops are essential for the second step of the apolipoprotein N-acyltransferase reaction. Mol Microbiol **95**:692–705.

29. **Buddelmeijer N**, **Young R**. 2010. The essential *Escherichia coli* apolipoprotein N-acyltransferase (Lnt) exists as an extracytoplasmic thioester acyl-enzyme intermediate. Biochemistry **49**:341–346.

30. **Zhai Y**, **Cui L**, **Lou J**, **Li H**, **Zhang XC**, **Sun F**, **Zhang K**, **Xiong Y**, **Lu G**, **Wang X**, **Xu Y**. 2017. Crystal structure of *E. coli* apolipoprotein N-acyl transferase. Nat Commun **8**:15948.

31. **Xu M**, **Kattke MD**, **Reichelt M**, **Kapadia SB**, **Gloor SL**, **Yan D**, **Zilberleyb I**, **Kang J**, **Katakam AK**, **Murray JM**, **Pantua H**, **Noland CL**, **Diao J**. 2017. Structural insights into lipoprotein N-acylation by *Escherichia coli* apolipoprotein N-acyltransferase. Proc Natl Acad Sci **114**:E6044–E6053.

32. **Wiktor M**, **Weichert D**, **Howe N**, **Huang CY**, **Olieric V**, **Boland C**, **Bailey J**, **Vogeley L**, **Stansfeld PJ**, **Buddelmeijer N**, **Wang M**, **Caffrey M**. 2017. Structural insights into the mechanism of the membrane integral N-acyltransferase step in bacterial lipoprotein synthesis. Nat Commun **8**:15952.

33. **Zhu Q**, **Casey JR**. 2007. Topology of transmembrane proteins by scanning cysteine accessibility mutagenesis methodology. Methods **41**:439–450.

34. **Vidal-Ingigliardi D**, **Lewenza S**, **Buddelmeijer N**. 2007. Identification of essential residues in apolipoprotein N-acyl transferase, a member of the CN hydrolase family. J Bacteriol **189**:4456–4464.

35. **Gille C**, **Fähling M**, **Weyand B**, **Wieland T**, **Gille A**. 2014. Alignment-Annotator web server: Rendering and annotating sequence alignments. Nucleic Acids Res **42**:W3-6.

36. **Brülle JK**, **Tschumi A**, **Sander P**. 2013. Lipoproteins of slow-growing *Mycobacteria* carry three fatty acids and are N-acylated by apolipoprotein N-acyltransferase BCG_2070c. BMC Microbiol **13**.

37. **Daniels R**, **Mellroth P**, **Bernsel A**, **Neiers F**, **Normark S**, **Von Heijne G**, **Henriques-Normark B**. 2010. Disulfide bond formation and cysteine exclusion in gram-positive bacteria. J Biol Chem **285**:3300–3309.

38. **Luong TT**, **Lee CY**. 2007. Improved single-copy integration vectors for *Staphylococcus aureus*. J Microbiol Methods **70**:186–190.

39. **Antoine R**, **Locht C**. 1992. Isolation and molecular characterization of a novel broad-host-range plasmid from *Bordetella bronchiseptica* with sequence similarities to plasmids from Gram-positive organisms. Mol Microbiol **6**:1785–1799.

40. **Datsenko KA**, **Wanner BL**. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A **97**:6640–6645.

41. **Armbruster KM**, **Meredith TC**. 2018. Enrichment of Bacterial Lipoproteins and Preparation of N-terminal Lipopeptides for Structural Determination by Mass Spectrometry. J Vis Exp e56842–e56842.

Chapter 6

Conclusions

Introduction

Recent studies have challenged the accepted paradigms of the lipoprotein field. It had been assumed that Gram-positive bacteria produce only conventionally diacylated lipoproteins, however orthologs of *Escherichia coli lnt,* and thus triacylated lipoproteins, were unexpectedly found in Gram-positive Actinobacteria. Triacylated lipoproteins are also synthesized in several Tenericutes and Firmicutes spp. in a manner distinct from that of Lnt. Furthermore, additional lipoprotein forms featuring N-terminal modifications, the peptidyl, *N*-acetyl, and lyso forms, can be found in various Firmicutes. Once highly implicated in lipoprotein localization to the outer membrane, the occurrence of N-terminal modifications in bacteria lacking an outer membrane suggests an alternate role. In this Dissertation, I have identified the enzyme Lit that converts diacylated lipoproteins to the lyso form in its native organisms *Enterococcus faecalis* and *Bacillus cereus*, and demonstrated that this enzyme also functions in *E. coli*. Further investigations into Lit revealed its colocalization on a transmissible copper resistance operon, suggesting a role of N-terminal lipoprotein modifications in resistance to copper. Additionally I have shown that conversion of lipoproteins to the lyso form significantly decreases recognition by Toll-like receptor (TLR) 2 with implications for the host immune response. Finally, preliminary experiments have laid the groundwork to determine if Lit functions through a novel intramolecular transacylation mechanism, whereby an ester-linked acyl chain from the diacylglycerol moiety is transferred to the α-amino terminus. Over the course of these studies, I have developed and optimized a method to determine lipoprotein N-terminal structure by mass

spectrometry. Still, many outstanding questions remain, including how lipoprotein forms other than the lyso form are made in Firmicutes. The work presented in this Dissertation has also opened new avenues of study, namely the possible contributions of lipoprotein structure to copper resistance, and the impact on the host immune response. These questions and more will be discussed in this chapter.

1. Synthesis of Additional Lipoprotein Forms in Firmicutes

1.1 The triacyl form

Staphylococcus aureus is not only a model Gram-positive organism, but can also be a formidable hospital- and community-acquired pathogen (1, 2). Therefore, the discovery of triacylated lipoproteins, identical in N-terminal structure to those of Gram-negative bacteria, in *S. aureus* and *S. epidermidis* remains of particular interest. In both Tenericutes and Firmicutes spp. that elaborate triacylated lipoproteins, the enzyme responsible for addition of the *N*-acyl chain is unknown. In efforts to identify the enzyme, we constructed a library of *S. aureus* genomic DNA in the hopes of rescuing the conditional lethal *E. coli* mutant of *lnt*. This experiment was identical to that presented in Chapter 2, which successfully identified Lit in both *E. faecalis* and *B. cereus*. Inserts from thirteen rescued colonies mapped to the region of the *S. aureus* chromosome encoding CymR (SAOUHSC_001732), a well-characterized transcription factor involved in cysteine metabolism (Figure 6-1A) (3). However, mass spectrometric analyses of the major lipoprotein SitC from the *S. aureus* type strain RN4220 and its cognate Δ*cymR* strain revealed no change in lipoprotein structure (Figure 6-1B and C). Indeed, the prominent peaks present in both spectra at *m/*z 1365, 1379, 1394, 1407, 1421, 1435, and 1449 are indicative of the triacylated N-terminal tryptic lipopeptide CGTGGK with varying acyl chain lengths, as has been observed previously (4, 5). Expression of *S. aureus* CymR in *E. coli* allowed for deletion of *lnt* but resulted

in diacylated lipoproteins, as shown by Western blot against the Strep-tagged Lpp(K58A) variant (Figure 6-1D). Therefore, the phenotypic rescue initially implicating CymR is by a mechanism other than lipoprotein triacylation, perhaps by mediating responses to stress caused by lipoprotein mislocalization, as CymR is a known factor in the *S. aureus* stress response (6, 7).

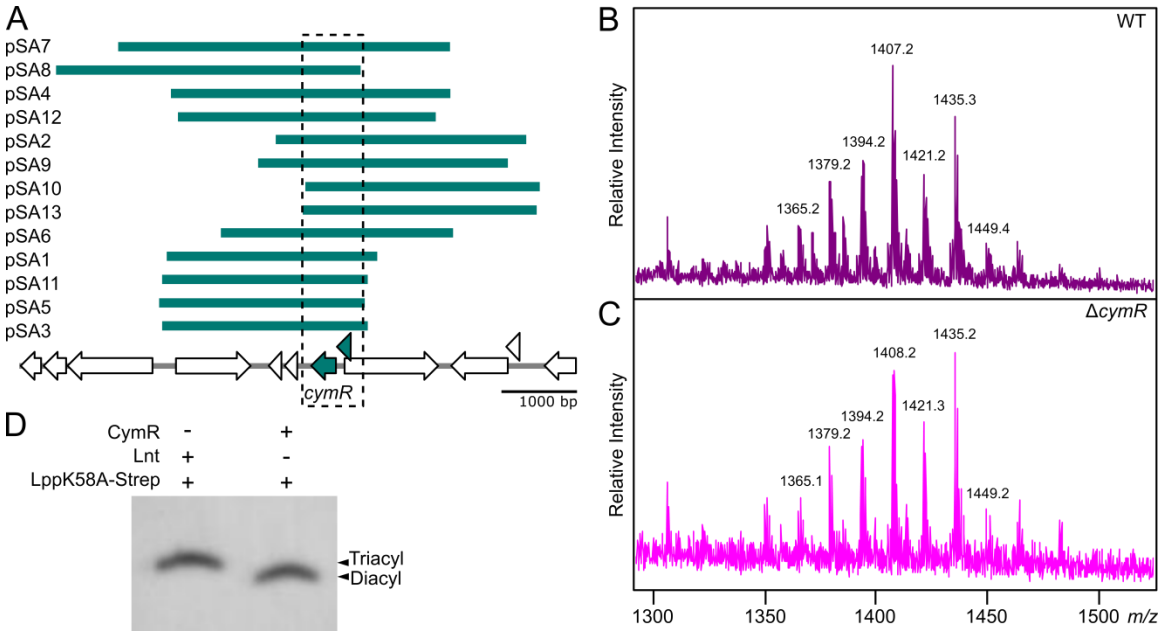


Figure **6-1**: *S. aureus* CymR conveys a viable phenotype in Lnt-depleted *E. coli* cells but does not *N*-acylate lipoproteins. (A) Arranged top-down by size from largest to smallest, DNA fragments that rescued growth in Lnt-depleted *E.coli* cells were mapped to the *S. aureus* RN4220 genome, implicating CymR in the observed phenotypic rescue. (B, C) MALDI-TOF of the N-terminal tryptic lipopeptide of SitC from wildtype RN4220 cells (B) and Δ*cymR* cells (C) each show an identical series of peaks corresponding to the triacylated lipopeptide. (D) A Western blot against the Strep-tag of Lpp(K58A) reveals bands corresponding to triacylated Lpp in wildtype *E. coli* cells (Lnt^+) and diacylated Lpp in *lnt*-null *E. coli* cells expressing *S. aureus* CymR.

Despite both *S. aureus lgt* and *lsp* complementing their corresponding deletions in *E. coli* (8, 9), our attempts to complement *lnt* were not successful, thus alternate approaches must be considered to identify the *N*-acylating protein in *S. aureus*. Similar to the *in vitro* assay presented in Chapters 2 and 5, *E. coli* Lpp could be used as a reporter lipoprotein whose conversion to the triacyl form by the *S. aureus* candidate protein would be easily visualized by SDS-PAGE. Once activity is observed from whole *S. aureus* cells, the cells can be systematically fractionated and used in reactions until the purest fraction still exhibiting *N*-acylation activity is found. This

fraction can then be analyzed by mass spectrometry to identify the proteins contained within and the candidate(s) further characterized in *S. aureus* to determine activity on lipoproteins in the native host. Though simple in theory, we hypothesize that the *S. aureus N*-acylating enzyme may use a unique acyl chain substrate, therefore it may prove challenging to reconstitute the system after each round of fractionation. However, if successful, identification of the enzyme would be a significant contribution to the lipoprotein field, answering long-standing questions about staphylococci physiology that may be extended into other Gram-positive phyla.

1.2 The *N*-acetyl form

N-acetyl form lipoproteins, characterized by an amide-linked acetyl group on the diacylglycerol-modified cysteine, are produced by the model Gram-positive organism *Bacillus subtilis*, along with several related strains, *Bacillus licheniformis*, *Bacillus halodurans*, *Oceanobacillus iheyensis* and *Geobacillus kaustophilus* (10). The *N*-acetyl form is also found in *Staphylococcus carnosus* (11). As the enzyme that catalyzes synthesis of the *N*-acetyl form is unknown, we again attempted to rescue the conditional lethal *E. coli* mutant of *lnt* with a library of *B. subtilis* genomic DNA, however no candidate could be identified (Armbruster and Meredith, unpublished data). As this experiment primarily relies on proper lipoprotein transport to the outer membrane as the means of rescue, it is possible that the *N*-acetyl form is not a substrate for the LolCDE export system. It is also possible that the *N*-acetyl form could not successfully be produced in *E. coli*, perhaps due to lipoprotein and/or acetyl substrate incompatibilities, or from deficient expression or proteolysis of the candidate enzyme. Repeating the rescue experiment using libraries of genomic DNA from other *N*-acetyl form-producing organisms may yield different results, though it is not guaranteed. That several species produce the *N*-acetyl form, while close relatives thereof do not (*B. cereus* and *S. aureus/S. epiderimidis*), also presents opportunities for using comparative genomics to identify potential candidates.

1.3 The peptidyl form

The peptidyl form, featuring two additional amino acids before the diacylglycerol-modified cysteine, has thus far only been found in *Mycoplasma fermentans* and is proposed to be the result of Lsp with a unique cleavage site (10). Interestingly, of the four lipoproteins characterized from *M. fermentans*, two were the peptidyl form, while the other two were the conventional diacyl form, suggesting not only a unique cleavage site of Lsp but also distinct lipoprotein substrate specificity (10). A thorough exploration of lipoproteins in this strain will be necessary to determine the ratio of peptidyl to diacylated lipoproteins. Knowing the identities of individual peptidyl lipoproteins may reveal trends in their function or a novel cleavage site of Lsp. Expressing the *M. fermentans* Lsp in a closely-related strain not known to make the peptidyl form would confirm the role of Lsp in the peptidyl form's synthesis, and also may provide insight into the specificity of Lsp.

In light of the proposed role of N-terminal modifications of lipoproteins in resistance to copper (Chapter 4), Lsp may have evolved to leave additional residues before the cysteine. It is possible that variations to the peptidyl form also exist, having only one, or more than two amino acids before the canonical cysteine. These residues may serve to weaken copper coordination at the lipoprotein N-terminus, as we have suggested of the *N*-acyl chain of the lyso form. If this is indeed the case, we may see Lsp enzymes of other strains evolving in this manner as well, especially among environmental isolates.

2. Lipoproteins and Copper

2.1 *In vivo* copper phenotype

While copper is an essential nutrient to life, excess copper can be detrimental to bacteria. Copper can generate reactive oxygen species (ROS) by Fenton-like chemistry that can damage lipids, proteins, and nucleic acids (12). Increased intracellular copper can also displace the iron-sulfur clusters that are required for the function of a wide array of proteins (13). Thus, bacteria have evolved a number of copper resistance mechanisms, including efflux pumps and copper binding proteins (14–16).

In Chapter 4, we presented genetic evidence of the co-selection of lipoprotein-modifying and copper resistance genes whose expression are copper-induced. Along with previous studies linking lipoprotein *N*-acylation to copper resistance (17), a greater connection between lipoproteins and copper is apparent even in the absence of a copper sensitivity phenotype *in vivo*. The chromosomal *lit* of *E. faecalis* and *B. cereus*, as well as the transmissible *lit2* of *Listeria monocytogenes* CFSAN023459, were all readily deleted with no obvious phenotypes (Chapters 2 and 4). Aerobic growth of the wildtype *E. faecalis* or *L. monocytogenes* versus their Δ*lit* derivatives with various concentrations of copper revealed no significant differences (Armbruster and Meredith, unpublished data). This is not altogether surprising, as Gram-positive bacteria are tolerant to loss of both Lgt and Lsp under normal laboratory conditions (18, 19). The phenotypes resulting from perturbing the Gram-positive lipoprotein modification pathway are typically evident only under certain conditions, such as nutrient starvation, oxidative stress, and in virulence models (18, 19). Furthermore, bacteria have several robust methods of combating copper toxicity, therefore lipoprotein N-terminal modification is likely not the first line of defense against excess copper. Aerobic growth conditions will only assay the effect of copper(II) ions, not copper(I), which is generally thought to be more toxic (13). In short, our preliminary attempts to observe an *in vivo* copper phenotype in relation to lipoprotein modification are thus far non-exhaustive. Several conditions may need to be tested before a more subtle phenotype is found.

Once observed, an *in vivo* lipoprotein-copper phenotype may significantly contribute to studies on lipoproteins in Gram-positive organisms, particularly in identifying the enzymes that produce the *N*-acetyl and triacyl forms in Firmicutes. Challenging a comprehensive single-gene deletion library of the Firmicutes strain of interest with copper under the correct conditions may reveal novel factors in copper resistance, potentially implicating candidate genes in lipoprotein modification. Analysis by mass spectrometry would confirm differences in lipoprotein acylation by the candidate protein. Following identification, a lipoprotein-copper phenotype would also facilitate further characterization of the enzyme, such as determining its essential amino acids.

2.2 *In vitro* copper binding to lipoproteins

Though the greater lipoprotein-copper phenotype and its effect on cell homeostasis remains unknown, we have proposed an interaction between lipoproteins and copper at the molecular level (Chapter 4). Diacylated lipoproteins with an exposed α-amino terminus, along with the thioether sulfur atom of the diacylglycerol moiety, theoretically provide an attractive site for copper to coordinate with a lipoprotein. This would increase copper localization at the membrane interface, potentially facilitating ROS formation within the membrane or entry into the cell where it can exert its toxic effects. Modification to the lipoprotein N-terminus that weakens cation coordination by delocalizing electron density at the α-amino nitrogen would decrease copper affinity to a lipoprotein. Such modifications may include lipoprotein *N*-acylation, *N*-acetylation, or retention of amino acids before the diacylglycerol-modified cysteine observed of the respective triacyl and lyso, *N*-acetyl, and peptidyl forms.

Direct *in vitro* assays for copper-lipoprotein binding are needed to complement copper phenotype studies. To demonstrate differential copper binding affinity to lipoproteins with various N-terminal structures, a copper-fluorophore quenching assay can be used (20). With this assay, triacylated, diacylated, and lyso-form *E. coli* Lpp and/or synthetic lipopeptides

(Pam_3CSK_4, Pam_2CSK_4, and PamC(Pam)SK_4, respectively) can be reconstituted into supported lipid bilayers with membrane-embedded fluorophores and a defined phospholipid composition. When copper is introduced to the system, the fluorophore is quenched only when in close proximity to copper, such as when associated with a phospholipid head group, or in this case a lipoprotein. Under our theory, we would expect to observe more fluorophore quenching with diacylated substrates (higher copper binding with exposed α-amino terminus) than triacylated or lyso-form substrates (lower copper binding to *N*-acylated lipoprotein). Though study of lipoproteins by this assay has not yet been reported, due to the nature of the experiment, if successful it will provide stronger evidence for copper association specifically with the N-terminus of lipoproteins than traditional methods of characterizing binding affinity.

3. Lyso-form Lipoproteins and Host Immunity

3.1 Recognition by Toll-like receptor 2 complexes

Our evidence indicates that lyso-form lipoproteins can signal through both the TLR2/1 and TLR2/6 heterodimers, as has previously been suggested (10). We have shown a significant preference for the TLR2/6 complex, though signaling via the lyso form is much weaker in comparison to their diacylglycerol-modified counterparts (Chapter 4). This suggests two lyso-form ligand-binding theories whereby (i) the lyso-form's monoacylglycerol moiety inserts into TLR2 and the *N*-acyl chain into TLR1, bridging the two proteins, as occurs with the triacyl ligand (21), or (ii) the lyso-form ligand contorts itself in such a way that both acyl chains insert into TLR2, similar to the diacyl ligand (Figure 6-2) (22). Altogether, lyso-form lipoproteins are poorer ligands than conventional diacyl or triacyl lipoproteins for either complex, which can perhaps be explained by the proposed binding theories. In the first theory, as only one lipid engages TLR2 instead of the traditional two, this may weaken engagement or association with TLR2, in turn

decreasing ligand recognition and subsequent signaling. In the second theory, depending on the ligand's conformation as it inserts both acyl chains into TLR2, steric effects may impact formation of the TLR2/6 complex, which is known to rely on interactions at the protein interface for heterodimerization (22). Decreased or weakened interaction between TLR2 and TLR6 would result in reduced signal due to unproductive complex formation. For either case, structural studies may be the only method to definitively determine lyso-form lipoprotein binding to the TLR2/1 and TLR2/6 heterodimers.

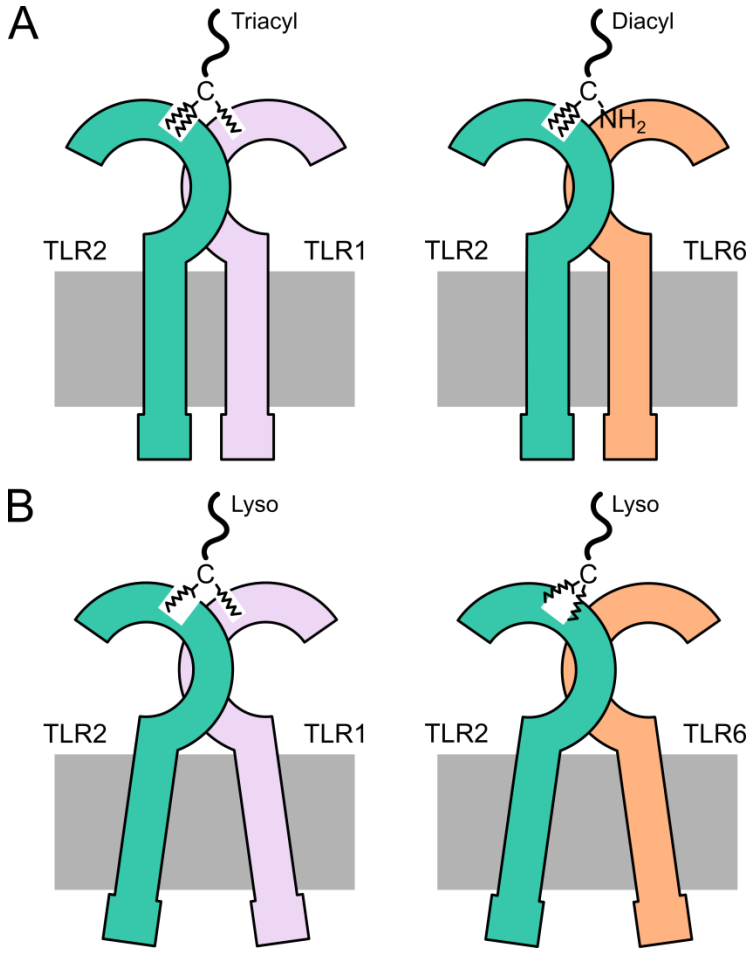


Figure **6-2**: Model of lyso-form lipoprotein binding to TLR2/1 and TLR2/6 heterodimers. (A) The shared diacylglycerol moiety of conventional triacylated and diacylated lipoproteins inserts into the larger binding pocket of TLR2, while the *N*-acyl chain of triacyl lipoproteins inserts into the smaller binding tunnel of TLR1 to trigger a signal cascade. (B) The *N*-acyl chain of lyso-form lipoproteins inserts into TLR1 with only a monoacylglycerol engaging with TLR2, making the lyso form a poor ligand for the TLR2/1 complex. In the TLR2/6 heterodimer, both the

monoacylglycerol and *N*-acyl lipids insert into TLR2 in a way that inhibits productive complex formation.

3.2 Macrophages

At the cross-section of copper resistance and host immunity are macrophages, specialized immune cells that detect, phagocytose, and destroy bacteria and other foreign bodies within the host. Macrophages utilize a multitude of methods to kill bacteria, including phagosome acidification, bacterial nutrient deprivation and starvation, generation of reactive oxygen and nitrogen species, and production of antimicrobial peptides (23, 24). Of particular relevance to this Dissertation, macrophages also use zinc and copper metal ions to disrupt bacterial cells (25–27). In response to proinflammatory agents, macrophages upregulate a copper importer, increasing copper uptake and capacity for killing copper-sensitive bacteria (28). Many pathogens have evolved mechanisms to combat these antimicrobial systems, including copper resistance, allowing for intracellular growth and phagosomal escape from macrophages (29–31). In light of the connection between N-terminal lipoprotein structure and copper resistance, it is possible that *N*-terminal modification to lipoproteins contribute to evasion of macrophage-mediated immune responses.

Again, although deletion of the lipoprotein modification pathway in Gram-positive bacteria largely results in no significant phenotypes in rich media, this may not hold true within the hostile environment of a macrophage. Lgt and Lsp have shown to be important in nutrient-limiting conditions and during oxidative stress (18, 19), common in macrophages. The added assault by copper from macrophages may further elicit a lipoprotein-related phenotype. To explore this possibility, the intracellular survival of diacyl- versus lyso-form-lipoprotein producing bacteria within macrophages can be measured. In this experiment, bacteria are phagocytosed by macrophages and incubated, then the macrophages lysed. Lysates are plated on solid media, resulting in viable colony counts. With this method, bacteria can also be harvested

for gene expression studies (32, 33). If lyso-form production does contribute to intracellular survival in macrophages, we may observe upregulation of *lit* and increased bacterial survival of lyso-form- over their diacyl-producing counterparts. This data would serve to further highlight the connection between lipoprotein form, copper resistance, and host immunity.

References

1. **Turner NA**, **Sharma-Kuinkel BK**, **Maskarinec SA**, **Eichenberger EM**, **Shah PP**, **Carugati M**, **Holland TL**, **Fowler VG**. 2019. Methicillin-resistant *Staphylococcus aureus*: an overview of basic and clinical research. Nat Rev Microbiol **17**:203–218. 2.
2. **Gnanamani A**, **Hariharan P**, **Paul-Satyaseela M**. 2017. *Staphylococcus aureus*: Overview of Bacteriology, Clinical Diseases, Epidemiology, Antibiotic Resistance and Therapeutic Approach. Frontiers in *Staphylococcus aureus*. InTech.
3. **Soutourina O**, **Poupel O**, **Coppée JY**, **Danchin A**, **Msadek T**, **Martin-Verstraete I**. 2009. CymR, the master regulator of cysteine metabolism in *Staphylococcus aureus*, controls host sulphur source utilization and plays a role in biofilm formation. Mol Microbiol **73**:194–211.
4. **Kurokawa K**, **Lee H**, **Roh K-B**, **Asanuma M**, **Kim YS**, **Nakayama H**, **Shiratsuchi A**, **Choi Y**, **Takeuchi O**, **Kang HJ**, **Dohmae N**, **Nakanishi Y**, **Akira S**, **Sekimizu K**, **Lee BL**. 2009. The Triacylated ATP Binding Cluster Transporter Substrate-binding Lipoprotein of *Staphylococcus aureus* Functions as a Native Ligand for Toll-like Receptor 2. J Biol Chem **284**:8406–8411.
5. **Asanuma M**, **Kurokawa K**, **Ichikawa R**, **Ryu KH**, **Chae JH**, **Dohmae N**, **Lee BL**, **Nakayama H**. 2011. Structural evidence of α-aminoacylated lipoproteins of *Staphylococcus aureus*. FEBS J **278**:716–728.

6. **Soutourina O**, **Dubrac S**, **Poupel O**, **Msadek T**, **Martin-Verstraete I**. 2010. The pleiotropic CymR regulator of *Staphylococcus aureus* plays an important role in virulence and stress response. PLoS Pathog **6**:1–13.
7. **Ji Q**, **Zhang L**, **Sun F**, **Deng X**, **Liang H**, **Bae T**, **He C**. 2012. *Staphylococcus aureus* CymR is a new thiol-based oxidation-sensing regulator of stress resistance and oxidative response. J Biol Chem **287**:21102–21109.
8. **Qi H-Y**, **Sankaran K**, **Gan K**, **Wu HC**. 1995. Structure-Function Relationship of Bacterial Prolipoprotein Diacylglyceryl Transferase: Functionally Significant Conserved Regions. J Bacteriol **177**:6820–6824.
9. **Zhao XJ**, **Wu HC**. 1992. Nucleotide sequence of the *Staphylococcus aureus* signal peptidase II (lsp) gene. FEBS Lett **299**:80–84.
10. **Kurokawa K**, **Ryu K-H**, **Ichikawa R**, **Masuda A**, **Kim M-S**, **Lee H**, **Chae J-H**, **Shimizu T**, **Saitoh T**, **Kuwano K**, **Akira S**, **Dohmae N**, **Nakayama H**, **Lee BL**. 2012. Novel bacterial lipoprotein structures conserved in low-GC content Gram-positive bacteria are recognized by Toll-like receptor 2. J Biol Chem **287**:13170–13181.
11. **Nguyen MT**, **Götz F**. 2016. Lipoproteins of Gram-Positive Bacteria: Key Players in the Immune Response and Virulence. Microbiol Mol Biol Rev **80**:891–903.
12. **Grass G**, **Rensing C**, **Solioz M**. 2011. Metallic copper as an antimicrobial surface. Appl Environ Microbiol **77**:1541–1547.
13. **Macomber L**, **Imlay JA**. 2009. The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity. Proc Natl Acad Sci **106**:8344–8349.
14. **Solioz M**, **Abicht HK**, **Mermod M**, **Mancini S**. 2010. Response of Gram-positive bacteria to copper stress. J Biol Inorg Chem **15**:3-14.
15. **Solioz M**, **Stoyanov J V.** 2003. Copper homeostasis in *Enterococcus hirae*. FEMS Microbiol Rev **27**:183-195.

16. **Chaturvedi KS**, **Henderson JP**. 2014. Pathogenic adaptations to host-derived antibacterial copper. Front Cell Infect Microbiol **4**:3.
17. **Rogers SD**, **Bhave MR**, **Mercer JF**, **Camakaris J**, **Lee BT**. 1991. Cloning and characterization of cutE, a gene involved in copper transport in *Escherichia coli*. J Bacteriol **173**:6742–6748.
18. **Buddelmeijer N**. 2015. The molecular mechanism of bacterial lipoprotein modification-How, when and why? FEMS Microbiol Rev **39**:246-261.
19. **Hutchings MI**, **Palmer T**, **Harrington DJ**, **Sutcliffe IC**. 2009. Lipoprotein biogenesis in Gram-positive bacteria: knowing when to hold ’em, knowing when to fold ’em. Trends Microbiol **17**:13-21.
20. **Monson CF**, **Cong X**, **Robison AD**, **Pace HP**, **Liu C**, **Poyton MF**, **Cremer PS**. 2012. Phosphatidylserine reversibly binds Cu 2+ with extremely high affinity. J Am Chem Soc **134**:7773–7779.
21. **Jin MS**, **Kim SE**, **Heo JY**, **Lee ME**, **Kim HM**, **Paik S-G**, **Lee H**, **Lee J-O**. 2007. Crystal Structure of the TLR1-TLR2 Heterodimer Induced by Binding of a Tri-Acylated Lipopeptide. Cell **130**:1071–1082.
22. **Kang JY**, **Nan X**, **Jin MS**, **Youn S-J**, **Ryu YH**, **Mah S**, **Han SH**, **Lee H**, **Paik S-G**, **Lee J-O**. 2009. Recognition of Lipopeptide Patterns by Toll-like Receptor 2-Toll-like Receptor 6 Heterodimer. Immunity **31**:873–884.
23. **Flannagan R**, **Heit B**, **Heinrichs D**. 2015. Antimicrobial Mechanisms of Macrophages and the Immune Evasion Strategies of *Staphylococcus aureus*. Pathogens **4**:826–868.
24. **Weiss G**, **Schaible UE**. 2015. Macrophage defense mechanisms against intracellular bacteria. Immunol Rev **264**:182–203.

25. **Djoko KY**, **Ong C-LY**, **Walker MJ**, **McEwan AG**. 2015. The Role of Copper and Zinc Toxicity in Innate Immune Defense against Bacterial Pathogens. J Biol Chem **290**:18954–18961.

26. **Hodgkinson V**, **Petris MJ**. 2012. Copper homeostasis at the host-pathogen interface. J Biol Chem **287**:13549–13555.

27. **Stafford SL**, **Bokil NJ**, **Achard MES**, **Kapetanovic R**, **Schembri MA**, **McEwan AG**, **Sweet MJ**. 2013. Metal ions in macrophage antimicrobial pathways: emerging roles for zinc and copper. Biosci Rep **33**:541–554.

28. **White C**, **Lee J**, **Kambe T**, **Fritsche K**, **Petris MJ**. 2009. A role for the ATP7A copper-transporting ATPase in macrophage bactericidal activity. J Biol Chem **284**:33949–33956.

29. **Portnoy DA**, **Chen C**, **Mitchell G**. 2016. Strategies Used by Bacteria to Grow in Macrophages. Microbiol Spectr **4**.

30. **Johnson MDL**, **Kehl-Fie TE**, **Klein R**, **Kelly J**, **Burnham C**, **Mann B**, **Rosch JW**. 2015. Role of Copper Efflux in Pneumococcal Pathogenesis and Resistance to Macrophage-Mediated Immune Clearance. Infect Immun **83**:1684–1694.

31. **Ladomersky E**, **Khan A**, **Shanbhag V**, **Cavet JS**, **Chan J**, **Weisman GA**, **Petris MJ**. 2017. Host and Pathogen Copper-Transporting P-Type ATPases Function Antagonistically during Salmonella Infection. Infect Immun **85**:e00351-17.

32. **Chatterjee SS**, **Hossain H**, **Otten S**, **Kuenne C**, **Kuchmina K**, **Machata S**, **Domann E**, **Chakraborty T**, **Hain T**. 2006. Intracellular gene expression profile of Listeria monocytogenes. Infect Immun **74**:1323–1338.

33. **Eriksson S**, **Lucchini S**, **Thompson A**, **Rhen M**, **Hinton JCD**. 2003. Unravelling the biology of macrophage infection by gene expression profiling of intracellular *Salmonella enterica*. Mol Microbiol **47**:103–118.

VITA

Krista M. Armbruster

Krista Armbruster was born in Buffalo, New York where she attended the State University of New York at Buffalo (UB). There, she received her Bachelor of Science (BS) in Biological Sciences, with a Concentration in Neuroscience, and a Bachelor of Arts (BA) in Spanish in 2014. Continuing her undergraduate research under Dr. Andrew Gulick in the Department of Structural Biology at UB, she worked as a Research Associate before starting her doctoral studies at The Pennsylvania State University at University Park, Pennsylvania. Advised by Dr. Timothy Meredith in the Biochemistry, Microbiology, and Molecular Biology (BMMB) Ph.D. Program, Krista published several articles focusing on the N-terminal structure of bacterial lipoproteins:

Armbruster KM, Meredith TC. 2017. Identification of the Lyso-Form *N*-Acyl Intramolecular Transferase in Low-GC Firmicutes. J Bacteriol. 199(11): e00099-17.

Armbruster KM, Meredith TC. 2018. Enrichment of bacterial lipoproteins and preparation of N-terminal lipopeptides for structural determination by mass spectrometry. J Vis Exp. 135:e56842.

Armbruster KM, Komazin G, Meredith TC. 2019. Copper-induced expression of a transmissible lipoprotein intramolecular transacylase alters lipoprotein acylation and the Toll-like receptor 2 response to *Listeria monocytogenes*. J Bacteriol. 201:e00195-19.

Spotlight: Copper induced remodeling of lipoproteins alters the Toll-like receptor 2 response

Her Doctor of Philosophy was conferred in August 2019.