DISSECTING THE BIOSYNTHETIC PATHWAY AND BIOLOGICAL SIGNIFICANCE OF UNIQUE LIPOPOLYSACCHARIDE STRUCTURES IN BACTERIAL NITROGEN FIXING ENDOSYMBIONTS

by

DUSTY B. BROWN

(Under the Direction of Russell W. Carlson)

ABSTRACT

Rhizobium are gram negative bacteria that form nitrogen fixing symbiosis (NFS) with compatible leguminous plants. Lipopolysaccharide (LPS) is a major cell surface molecule produced by gram negative bacteria and is crucial for establishing complete NFS in several symbiotic systems. In order to gain a more detailed understanding of LPS and its role in NFS, it is necessary to elucidate the structure, genes, and gene products. The general structure of LPS has been solved for the model organisms *Rhizboium leguminosarum* bv. *viciae* 3841 and *R. etli* bv. *phaseolus* CE3 and most of the biosynthetic genes have been discovered. LPS contains three domains; hydrophobic lipid A, core oligosaccharide, and O-chain polysaccharide. The O-chain is absolutely required to establish NFS, however, the function of the core and lipid A is unknown. The lipid A and core are highly galacturonosylated (GalA) containing one terminal 4` α-GalA on the lipid A and three terminal α-GalAs in the core. Inner membrane *Rhizobium* glycosyl transferases (RgtABCD) transfer GalA from the lipid donor dodecaprenyl-phosphate GalA (Dod-P-GalA) to the lipid A and core. The lipid A portion contains the unique very long chain fatty

acid (VLCFA, 27-hydroxyoctacosanoic) also found in intracellular mammalian pathogens. VLCFA is built off the special acyl carrier protein AcpXL by specific fatty acid biosynthetic elongating enzymes. *R. etli* produces 3-*O*-deacylated lipid A, attributed to outer membrane 3-*O*-deacylase activity also observed in mammalian pathogens due to the action of the outer membrane enzyme PagL. Interestingly, the endosymbiont *Mesorhizobium loti* and plant growth promoting bacteria *Azospirillum lipoferum* produce unusual lipid A α-1,1-GalA. In this work we describe the discovery of the lipid A 4` GalA transferase gene *rgtD*, the Dod-P-GalA biosynthetic gene *rgtE*, and the *M. loti* α-1,1-GalA transferase gene *rgtF* as well as the 3-*O*-deacylase gene *pagL* in *R. etli.* We have prepared single mutants in the *rgtABCDE* genes from *R. leguminosarum* 3841, the *pagL* gene from *R. etli* CE3, and the *acpXL* gene from *R. leguminosarum* bv. *phaseolus* 8002 and describe the effects of these mutations on LPS structure, synthesis, membrane stability, and symbiosis.

INDEX WORDS: lipopolysaccharide, lipid A, core oligosaccharide, long chain fatty acid, galacturonic acid, glycosyl transferase, 3-*O*-deacylase, endosymbiosis, nitrogen fixation, prokaryotic infection

DISSECTING THE BIOSYNTHETIC PATHWAY AND BIOLOGICAL SIGNIFICANCE OF UNIQUE LIPOPOLYSACCHARIDE STRUCTURES IN BACTERIAL NITROGEN FIXING ENDOSYMBIONTS

by

DUSTY B. BROWN

BS, The University of Georgia, 2003

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2011

DISSECTING THE BIOSYNTHETIC PATHWAY AND BIOLOGICAL SIGNIFICANCE OF UNIQUE LIPOPOLYSACCHARIDE STRUCTURES IN BACTERIAL NITROGEN FIXING ENDOSYMBIONTS

by

DUSTY B. BROWN

Major Professor: Russell W. Carlson

Committee: Mark Schell
Elmar Kannenberg

Electronic Version Approved:
Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2011

DEDICATION

I dedicate this work to the late Dr. Chtistian Raetz. As a National Academy of Sciences member, Dr. Raetz was a highly successful researcher at the Duke University Medical Research Center for many years. He was an excellent mentor to many and an inspiration to me.

ACKNOWLEDGEMENTS

I would like to thank Dr. Russell Carlson for his excellent mentorship. I owe much of my professional development to Dr Carlson's guidance, kindness, and unbreakable patience. I would also like to thank Dr. Biswa Choudhury, Dr. Artur Muszynski, Dr. Scott Forsberg, the Carlson lab, and my committee members for their guidance and support in the laboratory. A special thanks to the professors at the Complex Carbohydrate Research Center and Life Sciences (Dr. Lance Wells, Dr. Mike Tiemeyer, Dr. Maor Bar-Peled, Dr. Richard Steet, Dr. Amy Medlock, Dr. Alan Darvill, Dr. Steve Hajduk, and others) for fostering and maintaining a high level, interactive, and enjoyable learning environment. Lastly, thanks to all my friends in the Biochemistry and Molecular Biology graduate program for all the laughs and good times.

TABLE OF CONTENTS

CHAPTER 1

Introduction

Several gram negative bacteria of the *Rhizobiaceae* family are capable of forming nitrogen fixing (NF) endosymbiosis with compatible host plants in the *Leguminosae* family. NF symbiosis is a highly evolved process that involves the genesis of specialized plant root organs called nodules and simultaneous intercellular invasion of the developing nodules by the host specific bacterial symbiont. Eventually, the invading symbionts are endocytosed by specialized nodule tissue within the nodule core. While encased in a host derived membrane (peribacteroid membrane) the symbionts develop into nitrogen fixing bacteroids. The bacteroids persist within the endosomal compartment (symbiosome) resembling an organelle of which the specific function is to provide fixed nitrogen to the host plant. See Oldroyd and Downie (1) for a recent detailed review of the symbiosis process.

Many agriculturally significant legume food and foraging crops are capable of forming NF symbiosis, *e. g.* soy bean, peanut, peas, clover, alfalfa, and multiple varieties of other legumes. In addition, several woody legumes (leguminous trees) form NF symbiosis. Typically, NF legumes act as biofertilizers leaving the soil enriched in nitrogen as well as organic matter and in turn provides sequential crops with nutrients, aids to prevent soil erosion and promotes water retention. In addition, legumes have the propensity for growing in low nutrient soils. For these reasons, NF legumes have been historically important in traditional agricultural practices such as crop rotating. Modern agriculture, however, largely makes use of artificial abiotic fertilizers such as those produced by the Haber-Bosch method. Artificial fertilizers have made it

possible to meet the current global food demands and are the main source of fertilizer used currently worldwide (2,3). However, the over use of artificial fertilizers has led to inefficient nitrogen cycling, potentially dangerous agricultural "run off," soil depletion, and is considered an unsustainable practice (3-8). As the global population increases, the implementation of NF legumes is an attractive alternative for achieving sustainable food agriculture for reasons stated above (9). Furthermore, the cultivation of NF plants is a promising method for reforestation (woody legumes) and soil reconstruction, particularly in tropical areas where nutrients have been lessoned in soils due to unsustainable agricultural practices, industrial mining, and lumber wood production. In order to effectively utilize NF crops, a detailed understanding of the symbiotic process is necessary.

Surface and secreted bacterial polysaccharides are required for complete NF symbiosis in several symbiotic systems (10,11). Lipopolysaccharide (LPS), the focus of this thesis, is a surface polysaccharide located in the outer membrane of gram negative bacteria and is associated with several general biological functions such as, membrane stability, plant and animal pathogenesis, and antibiotic resistance (12-15). Disruption of LPS biosynthesis can lead to disruption of proper symbiosis in many nitrogen fixing symbiotic systems (16,17) as well as disruption of beneficial endophytic relationships (18). Because LPS is a strong initiator of the human innate immune response and is generally essential for the viability and virulence in gram negative pathogens, much attention has been devoted to studying the LPS in pathogens as it is an attractive target for broad based drug therapy and vaccine development for the treatment and prevention of infectious disease in animals. In addition to animal pathology, LPS is emerging as a key factor in plant pathogenesis as well as plant growth promoting symbiosis (15,19,20). However, key fundamental questions need to be answered for a more complete understanding of

LPS biology in plants 1) what is/are the LPS receptor/s in plants? and 2) what are the LPS structural requirements for LPS perception in plants? This thesis focuses on elucidating the biosynthetic components involved in the synthesis of unique LPS structures in the NF symbionts *Rhizobium leguminosarum* and *R. etli* and provides a foundation for answering the second question stated above.

LPS is an amphipathic lipid glycoconjugate located in the outer leaflet of the outer membrane and is typically composed of three domains: 1) Lipid A, the hydrophobic component that anchors the LPS to the outer leaflet of the outer membrane, 2) Core, an oligosaccharide that is attached to the lipid A, 3) O-chain, a highly variable polysaccharide that is attached to the core and generally consists of oligosaccharide repeating units. In order to gain a detailed understanding of LPS as it relates to a particular organism, it is crucial to elucidate the molecular structure and biosynthesis of the LPS in that organism. The combined methods of carbohydrate/lipid structural biology, enzymology, and genetics have greatly contributed to our understanding of *Escherichia coli* LPS biosynthesis and function. As a result, the *E. coli* system has historically served as a general model for LPS biosynthesis and has significantly contributed to our understanding of LPS in many bacterial systems. See Reatz and Whitfield (12) for a detailed review of *E. coli* LPS biosynthesis. In this chapter, I review our current knowledge of the structure, function, and biosynthesis of LPS in *E. coli,* animal pathogens, and the symbiotic model organisms *Rhizobium leguminosarum* bv. *viciae*, *R. leguminosarum* bv. *phaseoli* and *Rhizobium etli* bv. *phaseoli* focusing on the lipid A and core portions of the LPS molecule.

Biosynthesis, Structure, and Function of Lipopolysaccharide (LPS) in *E. coli* and Mammalian Pathogens

The Lipid A biosynthetic Pathway "The Raetz Pathway"

Lipid A is the hydrophobic domain that anchors LPS to the outer leaflet of the outer membrane and is indispensible and ubiquitous among gram negative bacteria. In *E. coli*, the biosynthetic pathway for the synthesis of lipid A and inner core Kdo residues consists of nine genes of which the gene products act in a sequential manner to synthesize the LPS precursor molecule Kdo_2-lipid IV_A. The lipid A biosynthetic pathway is referred to as the "Raetz Pathway" given the large contributions made by the laboratory of Christian Raetz in the elucidation of this pathway. Although a wide variety of lipid As are observed among bacterial species, homologous genes involved in *E. coli* Kdo_2-lipid IV_A biosynthesis are widely observed in sequenced strains suggesting that a general precursor molecule exists for the synthesis of most LPSs in nature. Thus, the constitutive lipid IV_A biosynthetic components are attractive targets for broad based antibiotics and a good starting point for studying novel LPSs. As diagramed in Figure 1.1 and briefly explained here, the lipid A biosynthetic steps begin with the acylation of UDP-GlcNAc by the cytosolic acyl transferase LpxA (21,22). In *E. coli*, LpxA utilizes the thioester linked β-hydroxymyristoyl acyl carrier protein (Acp-OHC14) to transfer OHC14 to UDP-GlcNAc forming an ester linkage at the 3 position (23). LpxA has an intrinsic ability to recognize fatty acid length ("molecular ruler"). For example, *E. coli* LpxA prefers OHC14 while *Pseudomonas aeruginosa* prefers OHC10 (13). This distinction is attributed to a single amino acid residue. Exchanging the *E. coli* "ruler" residue with the *P. aeruginosa* "ruler" residue and vice versa changes the Ec LpxA specificity from Acp-OHC14 to Acp-OHC10 and vice versa (24). The 3-acylated UDP-GlcNAc LpxA product is then deacetylated by the cytosolic deacetylase LpxC to

form 3-acyl UDP-GlcN. The cytosolic acyl transferase LpxD transfers OHC14 from Acp-OHC14 to the 2 position forming an amide linkage to produce 2,3-diacyl-UDP-GlcN. The peripheral membrane protein LpxH is a metallo-pyrophosphorylase that catalyzes the hydrolysis of UDP from diacyl-UDP-GlcN to form diacyl-1-phospho-GlcN and UMP. Interestingly, previous observations revealed that LpxH is missing in roughly one third of gram negative bacteria suggesting that an alternative phosphorylase existed in these bacteria. Metzger and Raetz (25) recently discovered an enzyme (LpxI) in *Caulobacter crescentus* in which its genetic origins localize between *lpxA* and *lpxB*. LpxI catalyzes the hydrolysis of UDP from diacyl-UDP-GlcN and is distinct from LpxH in that hydrolytic attack occurs at the β-phosphate of UDP in contrast to the α-phosphate as is the case with LpxH. Diacyl-UDP-GlcN and diacyl-1-phospho-GlcN are condensed by the peripheral membrane disaccharide synthase LpxB to form tetraacyl 1-phospho-β-1,6 di-GlcN and UMP. The integral membrane protein kinase LpxK utilizes ATP to transfer phosphate to the 4` position of the distal GlcN on the cytoplasmic side of the inner membrane producing 1,4`-*bis*-phosphorylated tetraacyl di-GlcN (lipid IV_A). Core synthesis begins with the sequential addition of two Kdo (3-deoxy-D-*manno*-octulosonic acid) residues to lipid IV_A. The first Kdo (Kdo1) is ketosidically linked at the 6 position of the distal GlcN and the second (Kdo2) is sequentially ketosidically attached to the 4 position of the first Kdo residue to form Kdo_2-lipid IV_A. In *E. coli,* KdtA is a bi-functional integral membrane glycosyl transferase that transfers Kdo residues from CMP-Kdo to lipid IV_A. Finally, secondary acyl chains are sequentially added to 3-hydroxyl groups of the 2` and 3` fatty acids by the integral membrane acyl transferases LpxL and LpxM. LpxL transfers laurate (C12) from Acp-C12 to the 2` fatty acid and LpxM transfers myristate from Acp-C14 to the 3` fatty acid to produce Kdo_2-lipid A. It is worth noting that a wide diversity of lipid A secondary acyl chain patterns exist in nature and

is likely attributed to diverse acyl transferases that are related to LpxM and LpxL, *e. g.* the LpxXL transferase is found in bacteria that have a secondary very long chain fatty acid in their lipid As (26-28).

Lipid A biosynthesis in plants

Interestingly, the model plant *Arabidopsis thaliana* contains nuclear encoded orthologues of the lipid A biosynthetic genes LpxA, LpxC, LpxD, LpxB, LpxK, and KdtA suggesting that plants produce a lipid A precursor molecule and perhaps an LPS like molecule (13,29). Recently, Li and coworkers (29) from the laboratory of Chris Raetz provided the first biochemical evidence that *Arabidopsis* synthesizes a 1,4`-*bis*-phosphorylated lipid IV_A precursor molecule. They created homozygous mutations in the putative canonical lipid A biosynthesis genes in *Arabidopsis* and found that intermediate products of lipid A biosynthesis such as those found in the canonical "Raetz Pathway" accumulated in mutant seedling lipid extracts as compared to wild type. For example, in the *Arabidopsis* $lpxB^{KO/KO}$ the lipid X and UDP-2,3-diacylglucosamine intermediate molecules accumulated in seedling extracts and an accumulation of lipid IV_A was observed in the $kdtA^{KO/KO}$ mutant. Furthermore, GFP fused *Arabidopsis* LpxABCDK and KdtA proteins localized to the mitochondria of *Arabadopsis* root tissues and sub cellular fractionation revealed a 3 to 48 fold increase of lipid X in the mitochondria of the $lpxB^{KO/KO}$ mutant as compared to plastids and whole cell homogenates respectively and no lipid X was found in the cytoplasmic membrane. These findings suggest that a lipid A and/or LPS like molecule exists in plants and is likely synthesized in the mitochondria and possibly plastids through a conserved lipid A biosynthetic pathway. The biological significance of such a molecule in plants is unknown. However, the discovery of lipid A precursors in *Arabidopsis* lays

the foundation for discovering putative enzymes in *Arabidopsis* (likely in the mitochondria) that may modify lipid A precursors to form a larger LPS like glycoconjugate. In addition, knowing the genes involved in plant lipid A biosynthesis allows for more detailed transcription and expression analysis of these genes under various environmental conditions, *e. g.* how are these genes regulated in response to infectious agents such as gram negative pathogens and symbionts?

Core assembly

The presence of the two inner core Kdo residues is well conserved in LPSs with a few exceptions. For example, *Bordetella pertussis* and *Helicobacter pylori* have only one inner core Kdo residue and these organisms represent two mechanisms that lead to the production of one inner core Kdo. It was demonstrated that the KdtA homolog in *Bordetella* is a mono-functional KdtA that can only transfer one Kdo residue (30). Stead and coworkers (31) demonstrated that the *H. pylori* KdtA homolog, like *E. coli* KdtA, was bi-functional and Kdo hydrolase activity in the membranes was responsible for trimming the terminal Kdo2 residue leaving the LPS with one inner core Kdo. Recently, Stead et al (32) discovered the *H. pylori* and *Franacisella novicida* Kdo hydrolases and revealed that two adjacent genes were necessary for activity. One gene encoded a sialidase like glycosyl hydrolase which required the second gene product for activity, a putative transmembrane protein of unknown function. The predicted hydrolase domain faces the periplasmic side of the inner membrane.

Bulkholderia cepacia, Yersinia pestis, and *Acinetobacter heamolyticus* produce a Kdo analogue called KO (D-glycero-D-*talo*-octulosonic acid) whereby the axial hydrogen at the 3 position is replaced with a hydroxyl group. LPSs with KO are more resistant to acid hydrolysis.

Chung and Raetz (33) recently described the KO synthase (KdoO) responsible for KO synthesis. KdoO is a Fe^{2+}, α-ketoglutarate, and O_2 dependent dioxygenase that oxidizes the terminal Kdo2 using Kdo_2-lipid A as a substrate and is only found in gram negative bacteria. KdoO contains a conserved iron binding motif ($HXDX_{n>40}H$) observed in dioxygenases and is distantly related to the *Salmonella* lipid A modifying enzyme LpxO hydroxylase, a dioxygenase that hydroxylates the alpha carbon of the secondary acyl chain C14. Although the biological significance of KO is not understood, it is possible that the presence of KO helps organisms evade host immune detection and may provide stability under hydrolytic conditions such as the phagosome environment. Engineering KdoO mutants in animal pathogens is the next step to understanding the biological/pathological significance of KO in animal pathogens.

The remaining core oligosaccharide moieties are sequentially added to the Kdo1 residue by glycosyl transferases that utilize nucleotide sugar donors and are synthesized on the cytoplasmic side of the inner membrane (12). See Raetz and Whitfield for a detailed review of core structure and synthesis in *E. coli* strains. A wide variety of core oligosaccharide structures are observed in nature. However, some structural trends exist among phylogenetically related organisms. For example, several gamma-proteobacteria such as *Klebsiella, Bordetella, Pseudomonas, Vibrio* produce heptosyl residues in their inner core residues along with common carbohydrates like glucose and galactose in the outer core region while several alpha-proteobacteria such as *Rhizobium, Brucella, Agrobacterium,* and *Sinorhizobium* do not contain core heptosyl residues and may contain uronic acids in their core (20,34). In *E. coli* strain R1, the core synthesis genes (*waa* genes) are localized together in three operons. However, this is not the case for all organisms and elucidating the core biosynthetic genes in these organisms is more difficult. The core region in enteric bacteria typically has an overall negative charge due to the

presence of Kdo and phosphate addition to the heptosyl and Kdo residues by heptosyl and Kdo kinases. In organisms that do not contain heptosyl residues like the aforementioned alpha-2-proteobacteria, the core regions contain negatively charged uronic acids instead of phosphates. The function of negative charge on the core region is discussed below.

Once the lipid A and core portions of LPS are synthesized on the cytoplasmic side of the inner membrane, the molecule is flipped to the periplasmic side of the inner membrane by the ABC transporter MsbA (12) (Figure 1.2). Further processing of the lipid A and core may occur at the periplasmic interface. For example, in *Rhizobium* and *Francisella*, the lipid A phosphates are removed by the 1 and 4` phosphatases LpxE and LpxF respectively (35-38) and the GlcN back bone is glycosylated by a class of enzymes (ArnT-like glycosyl transferases) as discussed below in the lipid A modifications section. Addition of phosphates and phosphoethanolamine to the lipid A and inner core occurs at the periplasmic interface as well (39-43). In *Rhizobium*, the core is glycosylated by ArnT-like *Rhizobium* galacturonosyl transferases (Rgt) (44) demonstrating that ArnT-like enzymes are not limited to lipid A glycosylation. Modifications of the lipid A and core at the periplasmic interface are typically dynamic events in pathogens that occur in response to environmental cues particularly in the host. However, many bacteria have adapted these enzymes for the constitutive *de novo* production of LPS, *e. g. Rhizobium* constitutively express lipid A phosphatases and GalATs and produce stoichiometric amounts of lipid A and core GalAs.

O-chain assembly

O-chain polysaccharide (OPS) in general is a heteropolysaccharide composed of oligosaccharide repeating units and is highly variable from strain to strain. In fact, the O-antigen

is widely used for strain specific serotyping. OPS is attached to the lipid A-core in the inner membrane at the periplasmic interface and can be explained by a three step process (Figure 1.2): i) synthesis and translocation of oligosaccharide diphosphate-bactoprenyl lipid donor that carries the OPS repeating subunit, ii) Polymerization of OPS by the elongation of OPS subunits, and iii) ligation of the polymerized lipid donor linked OPS to the core. The OPS repeating unit is built on to a monophosphate-bactoprenyl lipid at the cytosolic interface of the inner membrane. A series of glycosyl transferases utilize nucleotide sugar donors to form an oligosaccharide pyrophosphate-bactoprenyl conjugate (subunit lipid donor) that is then translocated to the periplasmic interface of the inner membrane by the translocase Wzx (12,13,45). The OPS lipid donor is elongated by the Wzz/Wzy system whereby the polymerase Wzy facilitates the elongation of the subunit lipid donor. Wzy can polymerize lipid linked OPS independently of Wzz, however, genetic studies have suggested that Wzz acts to regulate OPS length distribution (modality) as reflected by the banding patterns of LPS on SDS/DOC-PAGE (12). For example, *E. coli* (46) and *Shigella* (47) strains lacking a functional *wzz* gene are different from wild type in that more LPS with shorter OPS chain is produced and modality is lost.

Recently, Woodward and coworkers (48) developed the first *in vitro* system for studying the *wzz / wzy* interactions in *E. coli* strains. They compared two *E. coli* strains that display different LPS modalities. Strain H2 has a LPS intermediate modal distribution of 9 to 16 OPS subunit repeats and strain B7 has a short LPS modal distribution of 2 to 5. Although the LPS modality is very different in these two strains, their Wzz proteins share 90% amino acid identity. A mutant with a null *wzz* gene was created from strain B7 (B7Δ*wzz*) and the mutant was complemented with a functional *wzz* gene from strain H2 (B7Δ*wzz*/WzzH2). The complemented B7Δ*wzz*/WzzH2 strain now produced LPS with an intermediate modal distribution (9-16 OPS

units) similar to the H2 strain. The role of Wzz in regulating OPS chain length was further verified by an *in vitro* reaction containing an enzymatically produced OPS repeating unit undecacprenyl-pyrophosphate (RU-PP-Und) lipid donor, purified polymerase WzzB7, and purified WzzH2 or WzzB7. As expected, the reaction with only the polymerase Wzy and RU-PP-Und produced OPS-PP-Und species that had a nonmodal distribution. The addition of WzzH2 produced OPS-PP-Und with intermediate modality and the addition of WzzB7 produced OPS-PP-Und with short modality confirming Wzz as an OPS chain length regulator. The development of such an assay will significantly contribute to our understanding of Wzz/Wzy interactions in detail and is a step towards producing specific LPS molecules *in vitro*.

A second pathway for OPS synthesis exists (12) that is independent of the Wzy and has been described in *E. coli* and *Klebsiella pneumoniae*. In this system, OPS is synthesized on to Und-PP at the cytoplasmic interface of the inner membrane. Chain extension does not require in block transfer of OPS subunits from RU-PP-Und by the Wzy polymerase for chain extension. Instead, OPS chain extension is accomplished by processive addition of gylcosyl residues by glycosyl transferases (mono, bifunctional, or both) on the cytoplasmic face of the inner membrane. These OPS structures contain mainly homopolymeric units (i.e. polymannan and galactan) that may contain diverse linkages. After OPS extension is complete, the OPS-PP-Und is transported to the periplasmic face by the Wzm/Wzt ABC transporter.

In each case, Wzy dependent or independent pathways, the OPS is attached to the lipid A-core at the periplasmic interface by the OPS ligase WaaL (12). In *E. coli*, the *waaL* gene is localized on the chromosome in an operon with core biosynthesis genes. WaaL does not likely have a role in OPS chain length due to the fact that OPS-PP-Und species can be produced *in vitro* and contain similar modal distribution patterns as isolated LPS. However, this should be

confirmed in a similar *in vitro* assay as Woodward et al (48) described above. The fully formed LPSs are transported across the periplasm and into the outer membrane. Genetic and biochemical evidence has revealed a number of lipopolysaccharide transport proteins (Lpt) that are required for LPS transport and a model for LPS transport is forming (Figure 1.3). For a review on the emerging topic of LPS transport see Sperandeo et al (49). The Lpt proteins are found in the cytoplasm (LptB), inner membrane (LptF, and LptG), periplasm (LptC [periplasm associated] and LptA), and outer membrane (LptD and LptE). So far, the Lpt proteins have largely been elucidated through bioinformatics and mutational studies that have given rise to a model (Figure 1.3) by which the Lpt proteins may form a protein complex spanning from the cytosol to the outer membrane which utilizes ABC transporter machinery (LptBEF) and LPS chaperones (LptA) to facilitate transport and deposition of nascent LPS from the inner membrane to the outer membrane.

Lipid A: Structure and Function in E. coli and Animal Pathogens

LPS is often referred to as endotoxin due to its toxicity in humans and its ability to elicit the innate immune response by signaling through the tol-like receptor 4 (TLR4) surface receptor molecule. Lipid A is the minimal structure required to elicit the TLR4 mediated immune response in humans (50,51) at low picomolar to nanomolar concentrations (52,53). The general structure of lipid A, particularly observed in enteric bacteria, is a fatty acylated bisphosphorylated β-1,6 glucosamine (GlcN) disaccharide. *E. coli* lipid A has been studied extensively (12-14,54) and much is known about its structure, biosynthesis, and function. *E. coli* lipid A (Figure 1.7) is composed of a β-1,6 di-GlcN disaccharide that is bisphosphorylated at the 1 and 4` positions. Primary β-hydroxymyristate (β-OHC14) fatty acid moieties are amide linked

to the 2 and 2` positions and ester linked at the 3 and 3` positions. Secondary laurate (C12) and myristate (C14) fatty acid moieties are ester linked to the 3` and 2` β-OHC14 fatty acids respectively to form acyl-oxy-acyl moieties on the distal GlcN. As discussed below, the lipid A in *E. coli* and several enteric and pathogenic bacteria functions to stabilize the outer membrane and to elicit and evade host immune defenses.

It has been suggested that LPS molecules are cross linked in the outer membrane via divalent cationic bridging between the negatively charged phosphate moieties and divalent cations such as Mg^{2+} and Ca^{2+} (12,55). The major function of LPS cross bridging is to provide significant stability to the outer membrane. In enteric bacteria like *E. coli,* phosphates are generally attached to the 1 and 4` positions of the lipid A. However, some organisms may only contain monophosphorylated or non-phosphorylated lipid A. For example, *Rhizobium* species contain α-galacturonic acid (GalA) at the 4` position and to some degree contain proximal 2-aminogluconate (GlcNonate) instead of proximal GlcN. Although these organisms do not contain phosphate, negative charge is provided to the lipid A by the GalA and GlcNonate moieties. Interestingly, the plant endosymbionts *Bradyrhizobium elkanii* (56) and *B. japonicum* (unpublished results, personal communication with Artur Muszynski) produce a neutral mannosylated lipid A that apparently lack negatively charged moieties. However, *Bradyrhizobium* lipid A is heavily acylated with secondary long chain fatty acids that may provide additional stability through hydrophobic interactions. Furthermore, negatively charged moieties in the inner core region (in close proximity to lipid A) of LPS may also contribute to membrane stability by cross linking. For example, the inner core heptose region of *Pseudomonas aeroginosa* is heavily phosphorylated and indeed this organism is highly susceptible to lysis by the metal chelator ethylenediaminotetraacetate (EDTA) (57). As demonstrated in chapter 4 of

this work, disruption of core GalA synthesis in *Rhizobium leguminosarum* LPS by genetic engineering decreases overall membrane stability as determined by sensitivities to detergent.

Lipid A elicits the human innate immune response through TLR4/MD2 mediated signaling

Because lipid A is unique, ubiquitous, somewhat structurally conserved, and an indispensable part of the gram negative bacterial envelope, it is not surprising that higher organisms and competing microbes have evolved to recognize and/or target lipid A as a part of their host defense mechanisms. Thus, lipid A is a microbial associated molecular pattern (MAMP). Conversely, bacteria have evolved counter measures by modifying their lipid As in response to host environments which promotes survival. As diagramed in Figure 1.4, LPS binds to the human serum circulating LPS binding protein (LBP), a lipoprotein which facilitates the dissociation and transfer of LPS to the cell surface receptor CD14 on immune tissue such as macrophages and neutrophils (12,13,54). LPS is somehow transferred to TLR4 and the accessory protein MD2. The LPS/TLR4/MD2 complex dimerizes (58) and recruits the cytosolic adaptor protein MyD88. In turn, MyD88 attachment prompts a signal cascade that activates the NF-kB immune response related transcription factor. NF-kB promotes gene expression and the production of cytokines which leads to the recruitment and activation of immune tissues and clearing of potential pathogens. See Akira and Takeda (59) for a detailed review on TLR signaling. Severe bacterial infections (sepsis) in humans can lead to highly fatal gram negative septic shock, an acute acting disease defined by the abnormal hypersensitive immune response elicited by circulating LPS and recognition by TLR4. Septic shock leads to organ failure and eventual death within hours of symptoms and is the leading cause of death in intensive care units (12,60). As described below, pathogenic strains may produce modified lipid As that attenuate the TLR4 signaling pathway. Understanding the structural basis for these modifications and TLR4

binding will be helpful in developing therapeutics for the treatment of septic shock as well as developing whole cell vaccines without the deleterious effects of LPS endotoxicity.

Because several bacteria colonize the guts of animals (61), animals must have a mechanism for wasting constitutively produced LPS and other bacterial particles. Cani and coworkers (62) demonstrated that normally fed mice displayed an uptake and clearing of LPS in a 24hr feeding period by measuring circulating LPS in the blood (measured by the limulus amoebocyte lysate assay). In contrast, mice that were fed a high fat diet developed expected diabetic symptoms and interestingly maintained static amounts of LPS in the circulatory system suggesting a type of LPS intolerance developed in these mice. These and other results from this work suggest that LPS may also play a role in the inflammation associated with type-II diabetes, a human disease defined by insulin resistance that is attributed to inflammation (63) in adipose tissues (*i. e.* adipose associated macrophages). Several studies have shown that excess saturated free fatty acids (SFA) (64-67) can induce TLR4 induced inflammation as well demonstrating a correlation between diet and type-II diabetes. However, Schwartz et al (68) suggest that the metabolic processing of excess SFAs act to amplify the TLR4 response to LPS. It has also been suggested that high fat diets in mice increase the load of circulating LPS (62). Developing Lipid A mimickers, antagonists, and/or LPS inhibitors may prove valuable strategies for treating human metabolic disorders that are associated with the proinflamatory response. As described below, pathogenic strains produce modified lipid As that attenuate the TLR4 signaling pathway and may provide clues to develop non-toxic LPS antagonists.

Lipid A is a target for host produced polycationic antimicrobial peptides

In addition to LPS perception, LPS can be directly targeted by host or competing organisms by antimicrobial peptides (AMP). AMPs are produced in a wide range of organisms from bacteria, insect, mammals, and plants. Polymyxin B (PmxB) is an example of a ten amino acid residue acylated polycationic cyclic AMP that is non-ribosomally produced by *Bacillus polymyxa*. Typical of cationic AMPs, PmxB is an amphipathic molecule that attaches to negatively charged LPSs and phospholipids by ionic (phosphate) and hydrophobic (acyl chains) interactions and eventually forms multimeric complexes that disrupt the cellular envelope likely due to osmotic instability (69). In fact, because PmxB has such a high affinity for LPS, it is commonly used in affinity resins to separate LPS contaminants from samples and to prepare LPS samples to high purity. Vertabrates, invertabrates, and plants produce a class of AMPs that are cysteine rich polycationic peptides (peptides contain multiple di-sulfide linkages and numerous positively charged amino acids; arginines and lysines). Like PmxB, these defensins are short amphipathic peptides that interact with the negatively charged bacterial and fungal cell surface and somehow imbed into and disrupt the cellular envelope. In animals, defensins are found in high concentrations (from 10μg ml^{-1} to >10mg ml^{-1}) within leukocyte and epithelial granules and may be constitutively expressed or up-regulated in response to microbes and LPS (70,71). Infecting bacteria are endocytosed (phagocytosis) by leukocytes and barrier epithelial tissues into endosomal compartments (phagosomes). Granules fuse with and release their contents into phagosomes to kill invading microbes.

Lipid A modifications in human pathogens promote survival and infection

The success of several human pathogenic bacteria has been attributed to the modification to or synthesis of unique lipid A structures in the lipid A carbohydrate backbone and acyl chains (13,54,72). For the definitive reviews on virulence promoting lipid A modifications in gram negative bacteria see reviews by Trent (54) and Raetz (13,14). Lipid A modifications help bacteria evade initial recognition and innate immune defenses such as AMPs, TLR4 associated response, and complement lysis. In *E. coli* and several animal pathogens (*Salmonella, Shigella, Pseudomonas, Yersinia*, *Neisseria, Campylobacter, Helicobacter,* and *Vibrio*) the negatively charged lipid A phosphate moieties can be masked by the positively charged amino containing compounds 4-deoxy-4-aminoarabinose (Ara4N) and phosphoethanolamine (pEtN). These modifications have been shown to be regulated by the two component regulatory systems PmrA/PmrB, when turned on in response to polycationic AMPs promotes resistance to AMPs, and PhoP/PhoQ a major virulence regulator that is activated by the macrophage environment as well as exposure to polycationic AMPs and Mg^{2+} limited media (13,54). Activation of PhoP/PhoQ, PmrAPmrB, and subsequent lipid A modification with Ara4N and pEtN gives rise to polycationic AMP resistance in various pathogens due to the masking of the negatively charged phosphates and charge repulsion by the positively charged amino groups (54,73). In *E. coli,* attachment of the Lipid A Ara4N requires the synthesis of the lipid donor undecaprenyl-phosphate Ara4N (Und-P-Ara4N) by the membrane associated glycosyl transferase ArnC and deformylase ArnD (74,75). Subsequent transfer of Ara4N to the lipid A is facilitated by the integral inner membrane glycosyl transferase ArnT (76). Since the discovery of the ArnC / ArnA system of lipid A glycosylation, several ArnT and ArnC like genes have been discovered and are involved in various glycosylations of the lipid A disaccharide backbone as well as inner core

residues suggesting that conserved classes of enzymes (ArnT and ArnC like proteins) perform these processes (Figures 1.5 and 1.6). For a comprehensive description of the Ara4N – lipid A biosynthetic pathway via the *arn* operon gene cluster in *E. coli*, see work by Yan, Guan, and Raetz (77).

Francisella tularensis, the causative agent of tularemia and a class A bioterrorism agent, and the mouse model pathogen *F. novicida* produce a mixture of lipid A species (78,79) that lacks the 4` phosphate due to the action of the 4` phosphatase, LpxF (36). In addition, the 1-phosphate can be masked by galactosamine (GalN) attachment, and glucose or mannose (Glc or Man) can be attached to the 6` hydroxyl group. The two ArnC-like enzymes (polyprenyl-phosphate glycosyl transferases) responsible for the synthesis of the required lipid donors Und-P-Man or Und-P-Glc, and Und-P-GalN are FlmF1 (79,80) and FlmF2 (79,81) respectively, while the ArnT-like enzyme FlmK (80,81) may be a bifunctional glycosyl transferase that adds Glc or Man to the 6` position, and GalN to the 1 phosphate. Interestingly, in *Francisella*, the amount of free lipid A (lipid A that does not contain core or O-chain carbohydrates) is high (estimated at 90% of total LPS) and indeed lipid A modifications in *Francisella* are important for successful infection (14). The unique modified *Francisella* lipid A structures are poor elicitors of the TLR4/MyD88 signaling pathway and are required for infection of *F. novicida* in a mouse model (82). Inoculation (vaccination) of mice with non-virulent mutant *F. novicida* strains with disruptions in lipid A glycosylation provide protection against subsequent wild type infection (80% - 100% mouse survival) (80) demonstrating the promise of lipid A manipulation and whole cell vaccine development. Marr and coworkers (83) demonstrated that *Bordetella purtussis* (causes whooping cough in humans) and *B. bronchiseptica* produce a lipid A species that is modified with GlcN attached to the 1 and 4` phosphates. A *B. pertussis* strain Tomaha I

transposon mutant was found to be inhibited in the GlcN modification and the transposon insert was mapped to an *arnT* ortholog (accession # BP0398) that was co-transcribed with an adjacent *arnC* ortholog (accession # BP0399). Furthermore, the *arnT / arnC* orthologs were regulated by the BvgAS two component system, a master virulence regulator, suggesting a role for GlcN lipid A modification in virulence.

In addition to lipid A back bone modification, acyl chain modifications may occur in response to host environmental cues. The PhoP/PhoQ regulated *pag* genes (PhoP associated genes) *pagP* and *pagL* express outer membrane proteins responsible for the addition of C16 palmitate (*pagP*) and the removal of 3-*O*-β-hydroxymyristate (*pagL*). In *E. coli* and *Salmonella,* PagP transfers C16 from glycerophospholipids to the hydroxyl group of the 2-primary fatty acid forming an acyl-oxy-acyl linkage (Figure 1.7). Palmitoylation of *E. coli* and *Salmonella* lipid A provides increased resistance to AMPs and attenuation of TLR4 signaling. In *Bordetella bronchiseptica,* PagP adds C16 to the 3` primary fatty acid and is essential for persistent colonization of the respiratory tract in a mouse model and increases resistance to complement mediated lysis (54,84,85). The *Legionella pneumonphila* PagP is important for intracellular infection and virulence as well. *Pseudomonas aeroginosa* adds palmitate to the 3 positioned primary fatty acid under PhoP/Q conditions and is present on *Pseusomonas* lipid As isolated from cystic fibrosis patients. However, the *Pseudomonas* genome does not contain a *pagP* homologue and the function of palmitoylation is not completely understood in this system.

PagL, the topic of Chapter 3, is an outer membrane 3-*O*-deacylase that is wide spread in gram negative bacteria and is PhoP/PhoQ regulated in several enteric pathogens (54). PagL removes the primary hydroxyl fatty acid at the 3 position of the proximal GlcN residue. Because PagL homologues are found in pathogenic and non-pathogenic bacteria, it was suggested that

PagL does not have a committed role in pathogenicity (86). However, upon further investigation, it was demonstrated that purified *E. coli* LPS that was 3-*O*-deacylated by heterologous expression of the *Salmonella pagL* gene attenuated TLR4 signaling compared to wild type LPS (87). In an effort to demonstrate the efficacy of PagL modified lipid A in whole cell vaccines, Geurtsen et al. (88) introduced the *B. bronchiseptica pagL* gene into *B. pertussis* and showed that isolated LPS had lower endotoxicity than wild type LPS *in vitro.* In contrast, a whole cell preparation of the *pagL* complemented *B. pertussis* strain increased the level of endotoxicity *in vitro* as compared to wild type (determined by cytokine production of LPS stimulated macrophages). It was suggested that the paradoxical increased endotoxic effect from the whole cell prep was due to the increased release of LPS from the membrane in this strain as a significant proportion of LPS (cell bound vs. released) was detected in the culture media of the *pagL* complemented strain as compared to the wild type (54% - 66% increase). On the other hand, whole cell preparations of *Neisseria meningitidis* expressing the *B. bronchiseptica pagL* displayed decreased levels of endotoxicity (88) demonstrating that the effects observed in one bacterial system may not reflect that of other systems.

Pseudomonas aeroginosa is a common opportunistic pathogen that infects the respiratory tract in cystic fibrosis patients. Ernst and coworkers (89) showed that *Pseudomonas* isolated from the airways of cystic fibrosis patients had an increased expression of PagL activity as compared to laboratory adapted strains and isolates of acute *Pseudomonas* infection (blood, urinary tract, and eye). Interestingly, a difference in PagL enzyme activity was observed in isolates from mild pulmonary infection versus severe pulmonary infection. Although the apparent full length *pagL* gene was present (determined by PCR) in *Pseudomonas* isolates of 21 patients with severe pulmonary infection, 33% did not have PagL activity suggesting a

mechanism for adaptation/selection which down regulates the activity or expression of PagL in chronic pulmonary infections. As discussed in Chapter 3 of this thesis, 3-*O*-deacylase activity is observed in the endosymbiont *Rhizobium etli* CE3 but not in *R. leguminosarum* 3841 and the related strain *R. leguminosarum* 8002. Upon genome analysis, it was demonstrated that strain CE3 has a full copy of the *pagL* gene, however, strain 3841 has remnants of a severely truncated *pagL* gene. In addition, the closely related strain *R. leguminosarum* WSM1325 has a full copy of the *pagL* gene (lipid A structure has not been determined) further demonstrating the possibility that selective pressure may be necessary for 3-*O*-deacylase stability in some organisms. Like *Pseudomonas, R. etli* bacteroid isolates from infected root nodules display a significant increase in PagL activity compared to laboratory grown free living bacteria suggesting a functional role for PagL in the endosymbiosis process (90).

Core function and modifications

All of the known core structures resolved thus far display anionic characteristics in part due to highly conserved anionic inner core Kdo residues and in some cases uronic acids. As stated above, organisms may phosphorylate core residues to increase the anionic nature of the core, and that these negative charges help to cross bridge LPS molecules through divalent cationic interaction (*e.g.* Ca^{2+} and Mg^{2+}) in the outer membrane to increase membrane rigidity and stability (12). This hypothesis is supported by the fact that core mutants that lack normal negatively charged moieties become more sensitive to lysis by detergents, antimicrobial peptides, and the metal chelator EDTA (12,57, and chapter 4). It is likely that cross bridging of LPS is highly conserved and is a general function necessary for normal membrane stability in gram negative bacteria. Pathogens are able to modify their core structures with the positively

charged amino containing residue pEtN in response to host environments (12,13,54,91). It is thought that these modifications help evade actions of the initial innate immune response because they are regulated by the PmrA/PmrB two component regulatory system. Core modifications do not seem to be as important as lipid A modifications in terms of promoting virulence and resistance to cationic AMPs. For example, *Salmonella* mutants that lack a functional *cptA* (outer heptose pEtN transferase) gene only display a minor increase in sensitivity to the AMP polymyxin B (54) and virulence is not hindered (91). However, a double mutant which lacks a function *cptA* and *pmrC/EptA* (lipid A pEtN transferase) is moderately less virulent in a mouse model than either single mutant and wild type suggesting a synergistic nature of multiple pEtN additions (92).

Antigenic variation in the LOS core region of Neisseria gonorrhoeae and the modulation of the innate immune response through dendritic cell perception

In humans, epithelial derived dendritic cells are major antigen presenting immune tissues that directly regulate primary and adaptive immune responses; *e. g.* promotes T lymphocyte maturation as well as maintaining B cell function (93). The LPS core has been implicated as a ligand for the human dendritic cell (DC) receptor DC-SIGN, a pattern recognition receptor responsible for recognition and phagocytosis of microorganisms and HIV-I (94,95). DC-SIGN is a C-type lectin that is thought to recognize GlcNAc and mannan structures. Klena et al noticed (94) that *E. coli* strain HB101 (does not contain O-chain) was more readily phagocytosed by DCs than pathogenic *E. coli* strains that maintain O-chain. Interestingly, *E. coli* strains that lack O-chain (HB101 and K12) are readily phagocytosed by HeLa cells that express DC-SIGN, but mutants with truncated core, and pathogenic *E. coli* strains that maintain O-chain are not

susceptible to phagocytosis to the same DC-SIGN expressing HeLa cells. These results imply a role for core in the recognition of DC-SIGN and a possible role for O-chain in the evasion of DC-SIGN dependent phagocytosis.

The strictly human sexually transmitted pathogen *Neisseria gonorrhea* does not produce LPS with O-chain but instead produces LOS (lipo-oligosaccharide) composed of lipid A and a core like oligosaccharide portion. The *N. gonorrhoeae* LOS undergoes rapid antigenic variation that is regulated by reversible phase variable gene expression (slip-strand mispairing within polynucleotide tracts that leads to frameshift mutations in the open reading frame) turning "on" or "off" LOS glycosyl transferase genes. Phase variation leads to truncation in the oligosaccharide main α chain (96). *N. gonorrhoeae* produce LOS variants *in vivo* which contain different terminal sugars due to the "off" state of LOS glycosyl transferase genes, *e. g.* 1) terminal GalNAc (wild type), 2) terminal Gal (*lgtD* gene ["off"]), site for sialic acid transfer, 3) terminal Glc (*lgtE*), and 4) terminal GlcNAc (*lgtB*). It was demonstrated by Zhang and coworkers (97) that *N. gonorrhoeae* cells were highly limited in endocytosis by DCs as compared to *E. coli* HB101. HeLa cells expressing DC-SIGN, however, endocytosed *E. coli* HB101 and a *N. gonorrhoeae* mutant (*ΔlgtB*) that produced truncated LOS with terminal GlcNAc. The *ΔlgtB* mutant mimics the above *in vivo* antigenic variant that has terminal GlcNAc. Interestingly, mutants that displayed truncated LOS that mimicked other antigenic variants (*lgtA, B, C* mutants, and wild type) were not subject to endocytosis in a DC-SIGN dependent manner. The *lgtB* mutant produces an LOS that resembles HB101 core in that a terminal GlcNAc is exposed suggesting that terminal GlcNAc is required for DC-SIGN recognition. Zhang et al (95) further demonstrated that the presence of terminal GlcNAc in other bacterial strains was required for DC-SIGN dependent phagocytosis. Vliet et al (96) demonstrated that different *N.*

gonorrhoeae LOS variants modulate DC responses through interaction of different cell receptors. They confirmed that a terminal GlcNAc variant was internalized in a DC-SIGN dependent manner at a significantly higher incident than terminal Gal and GalNAc variants. Interestingly, the terminal GalNAc variant showed significant binding to CHO cells expressing the cell receptor MGL, another C-like receptor expressed on the surface of DCs, while the other variants showed minimal binding. This was further demonstrated by binding studies directly assaying the interaction between purified LOS variants and the MGL receptor. Interestingly, different LOS variants could induce differential expression of DC cytokines and DCs activated by different LOS variants differentially activated T helper cells. The authors proposed a model by which LOS antigenic variation modulates different human innate immune responses depending on the antigen variation in order to promote bacterial survival. Phase variation in *N. gonorrhoeae* is also associated with resistance to serum-antibody dependent complement lysis (98). Mutant strains producing elongated forms of LOS (non-variable) are more susceptible to complement lysis. It is tempting to speculate possible therapeutic uses for *N. gonorrhoeae* LOS like oligosaccharides. For example, it may be possible to facilitate the internalization of a therapeutic agent into dendritic cells and/or macrophages by conjugating to the drug an oligosaccharide containing a terminal GlcNAc similar to *N. gonorrhoeae* LOS. It may also be advantageous to produce a non-endotoxic LOS like molecule that is able to modulate immune tissue for the treatment of auto immune related disorders.

O-chain function

Because this thesis focuses on lipid A and core, I will briefly describe the basic functions of the LPS O-chain polysaccharide (OPS). For a more detailed description of O-chain function

see review by Raetz and Whitfield (12). OPS is highly variable from strain to strain and OPS structure within a strain may also vary in response to the host environment. In general, OPS acts as a barrier to protect against harmful molecules such as host derived antimicrobial agents. In many animal pathogens, disruption in O-chain biosynthesis leads to ineffective host colonization and is therefore a major virulence factor (99,100). Interestingly, the regulation of the modal distribution of OPS chain length has been shown to contribute to virulence and the resistance to serum complement lysis (101-103). The enteric pathogens *E. coli, Shigella flexneri,* and *Salmonella typhimurium* possess two *wzz* genes; one confers shorter modality and the other confers longer modality (the *E.coli fepE* gene and homologues) (103,104). The two *wzz* genes contribute to a bimodal distribution of LPS (shorter and longer) as determined by mutational analysis and SDS-PAGE. The location of *fepE* on a plasmid in *Shigella* and the non syntenic distribution of the *fepE* gene among *E. coli, Shigella,* and *Salmonella* suggests that the *fepE* gene may have been a late acquired LPS biosynthetic gene possibly acquired to promote resistance to complement lysis as high molecular weight LPS (long modality) is associated with complement resistance (105-107).

LPS and host mimicry

Some gram negative pathogens such as *H. pylori, Campylobacter jejuni, and N. meningitidis,* and *N. gonorrhoeae* have acquired the ability to synthesize LPS structures that have host mimicking qualities. For example, *H. pylori* produce Lewis X and Lewis Y carbohydrate structures in the OPS that mimic human blood group antigens. *Neisseria* species are able to add sialic acid to the oligosaccharide portion of LOS and *C. jejuni* produces a human ganglioside like carbohydrate structure on the core (108). In each case, phase variable expression

of genes promotes the production of host mimetic LPS/LOS structures. Host mimicry presumably allows organisms to evade host immune perception as self-antigens should not elicit an immune response. However, persistent infection can lead to the production of autoantibodies and the onset of autoinflammatory disorders as is the case with *C. jejuni* (Guillaine Barré syndrome, neuropathy) and *H. pylori* (gastritis). The sialylation of *N. gonorrhoeae* confers resistance to complement lysis (109,110) however, the role of *N. meningitidis* LOS sialylation is unclear as nonsialydated LOS mutants maintain resistance to complement (111). This could be attributed to the fact that *N. meningitidis* produces a capsule often containing homopolymeric sialic acid which is phase variable and produced by strains in host serum and not normally outside of the host. The sialydated capsule provides resistance to complement in *N. meningitidis* as well as other bacteria that produce salidated capsules and the presence of the capsule may override the effects of LOS sialylation (110). A complement resistant assay of a *N. meningitidis* mutant that lacks both LOS sialylation and capsule compared to a mutant that only lacks capsule and maintains LOS sialylation will help elucidate the role of LOS sialylation in serum resistance. Perhaps the initial phase of infection requires LOS sialylation to promote survival when the organism is uncapsulated and persists long enough to select for the capsulated population.

Structure and Function of Lipopolysaccharide in *Rhizobium* and Contributions of This Work to Existing Knowledge

The general LPS structures from laboratory cultured strains *R. leguminosarum* bv. *viciae* 3841 and *R. etli* bv. *phaseoli* CE3/CFN42 have been solved (Figures 1.8 and 1.9) (20). So far, the observed core and lipid A structures among members of the *Rhizobium* genera are conserved while the O-chain is strain specific. As discussed below, *Rhizobium* LPSs possess unique

structural features when compared to enteric bacteria such as *E. coli* and the elucidation of these structures laid the foundation for discovery of the biosynthetic components, *e. g.* enzymes and genes, especially in the highly unique lipid A portion. Understanding the molecular structure, biochemistry, genetics, and biosynthesis of these LPSs is crucial to understanding their biological functions in *Rhizobium* and like organisms. This section focuses on structure, biosynthesis, and function of the lipid A and core LPS domains from *R. leguminosarum* 3841 and *R. etli* CE3 and the contributions of this thesis to our current knowledge.

Lipid A: structure and biosynthesis

When cultured under laboratory conditions, the general lipid A structure is conserved in bacterial species belonging to the genera *Rhizobium*. *Rhizobium* produce a heterogeneous mixture of lipid As composed of a constitutive trisaccharide backbone that consists of two acylated β-1,6-GlcNs and α-1,4-GalA at the 4` position of the distal GlcN (112,113) (Figure, 1.8, 1.9, and 1.10). The acyl composition consists of a mixture of primary fatty acids (OHC14, OHC15, OHC16, and OHC18) ester linked at the 3 and 3` positions and amide linked at the 2 and 2` positions. A very long chain secondary fatty acid (VLCFA, 27-hydroxy octacosanoic acid) is attached to the hydroxyl group of the 2` primary fatty acid to form an acyl-oxy-acyl moiety. The VLCFA contains either a hydroxyl group or a 3-hydroxy butaroyl group at the C27 position. A partial population of lipid As contain proximal gluconate (GlcNonate) due to oxidation in the outer membrane as described below. *Rhizboium etli* partially produce lipid A species that are 3-*O*-deacylated in the outer membrane (114), *i. e.* the primary fatty acid at the 3 position of the proximal GlcN/GlNonate is removed. However, 3-*O*-deacylated lipid A is not observed in *R. leguminosarum* 3841 or *R. leguminosarum* 8002.

Rhizobium etli synthesizes a precursor molecule similar to the *E. coli* Kdo_2 lipid IV_A (115) (Figure 1.10) and genomic analysis reveals that *Rhizobium* species contain lipid A biosynthetic genes homologous to those found in *E. coli* (20) with the exception of LpxH (UDP pyrophosphorylase). *Rhizobium* species contain a gene homologous to the *Caulobacter crescentus* alternative UDP pyrophosphorylase LpxI (25) described above. *Rhizobium* lipid A biosynthesis diverges from *E. coli* after the synthesis of the Kdo_2 lipid IV_A (Figure 1.10) and downstream biosynthetic steps confer the unique structural features observed in *Rhizobium* lipid A. The secondary VLCFA is synthesized on a special acyl carrier protein (AcpXL) (116) and transferred to the lipid A by the special VLCFA acyl transferase (LpxXL) (26), an enzyme distantly related to the *E. coli* LpxL acyl transferase. In *Rhizobium,* the *acpXL* and *lpxXL* genes fall in a chromosomally located gene cluster which contains homologous genes to the general fatty acid biosynthesis genes (*fab*) (27). Several organisms that produce VLCFA in their lipid As possess the VLCFA biosynthesis gene cluster (27). Genetic studies in various organisms have demonstrated the requirements for each gene within the VLCFA gene cluster, *i.e. acpXL, fabZXL, fabF2XL, fabF1XL, fabIXL/fabGXL,* and *lpxXL* are required for the biosynthesis of lipid A VLCFA (113,117-122) particularly in the model organism *Sinorhizobium meliloti* 1021 where all of the genes within the VLCFA cluster have been individually mutated (120). The origins of the VLCFA 27-hydroxy and 27-β-hydroxybutyryl (BHB) moieties are unknown. It is possible that hydroxylation and subsequent transfer of BHB occurs enzymatically. Hydroxylation may occur early in the biosynthesis as the hydroxyl group is always present on the next to last carbon of elongating intermediates (116) (Figure 1.11, panel A). Another possible origin of the hydroxyl group may be attributed to the premature stoppage of the FabZXL (β-hydroxyacyl-Acp dehdratase) after the condensation of malonyl-AcpXL and acetyl-Acp/CoA (Figure 1.11, panel

B) followed by normal fatty acid elongation and α-keto reduction. A third possible origin of the hydroxyl group may be due to the condensation of malonyl-Acp with a substrate that already contains the hydroxyl group such as β-hydroxybutyryl-CoA/Acp (Figure 1.11, panel C). It is likely that the VLCFA BHB group is added enzymatically to the 27-hydroxyl position as the presence of BHB was not observed in cytosolic fractions of natively purified AcpXL-hydroxyfatty acids (116). An *in vitro* assay containing the VLCFA biosynthesis proteins is crucial in order to resolve the iterative biosynthetic steps for VLCFA biosynthesis and discovering a possible VLCFA BHB transferase and/or hydroxylase.

The lipid A 1 and 4` phosphates are removed on the periplasmic side of the plasma membrane (Figure 1.12) by the integral membrane phosphatases LpxE (38,123) and LpxF (36,123,124) respectively. Dephosphorylation allows for subsequent addition of 4` α-1,4-GalA by the ArnT-like glycosyl transferase RgtD, recently demonstrated in our laboratory and described in Chapter 4 of this thesis. RgtD utilizes the dodecaprenyl-phosphate galactoronosyl (Dod-P-GalA) lipid donor to transfer GalA to the 4` hydroxyl. As described in Chapter 4, Dod-P-GalA is synthesized by the ArnC like glycosyl transferase RgtE. The RgtD/RgtE follows a similar progression as the *E. coli* ArnT/ArnC pathway (Figure 1.5) for lipid A glycosylation. However, lipid A GalA is constitutively produced in *Rhizobium* in contrast to lipid A glycosylation in animal pathogens (13,14,54) which is typically regulated by the host environment. The outer membrane monooxygenase LpxQ (125) converts the proximal GlcN residue to the negatively charged GlcNonate residue. As described in Chapter 3, the *R. etli* outer membrane 3-*O*-deacylase PagL removes primary hydroxyl fatty acid at the 3 position of the proximal GlcN/GlcNonate. The action of PagL and LpxQ produce a mixture of 3-*O*-deacylated and GlcNonate containing LPSs on the outer membrane surface.

Core structure and biosynthesis

The core oligosaccharide structure is conserved between *R. leguminosarum* and *R. etli* and is likely conserved among organisms belonging to the genera *Rhizobium*. *Rhizobium* core is a branched polyanionic octasaccharide that contains the highly conserved lipid A attached inner core Kdo residues (Kdo1 and Kdo2) and an outer distal Kdo (Kdo3) (Figure 1.8 and 9). The core back bone is a tetrasaccharide (α-Kdo3-(2→6)-α-Gal-(1→6)-α-Man-(1→5)-α-Kdo1) that contains three terminal α-GalA residues attached to the 4 and 5 positions of Kdo2 and the 4 position of the Man backbone residue. The inner Kdo residues and core backbone is elongated onto newly synthesized lipid A at the cytoplasmic side of the inner membrane by the membrane bound glycosyl transferases KdtA (inner core Kdo_2), LpcC (mannosyl transferase, ManT), LpcA (GalT), and LpcB (outer KdoT) (20,126). After core backbone synthesis is complete, the lipid A core molecule is transported to the periplasmic side of the inner membrane by MsbA (ABC transporter) where GalA transfer occurs (35). Kanjilal-Kolar et al. (44,127) discovered the three *Rhizobium* GalAT enzymes; RgtA, RgtB, and RgtC; that add GalA to the core. Interestingly, the Rgt proteins are ArnT-like glycosyl transferases that utilize the lipid donor Dod-P-GalA. These results demonstrate that ArnT-like enzymes function not only to glycosylate lipid A but extend to modification of the core oligosaccharide. *In vitro* evidence presented by Kanjilal-Kolar et al. (44) and *in vivo* genetic studies presented in Chapter 4 of this work suggest that GalA transfer follows an order of synthesis first adding GalA to the lipid A by RgtD followed by GalA addition to the 4 position of Kdo2 by RgtA, followed by GalA addition to the 5 position of Kdo2 by RgtB, and finally GalA addition to the 4 position of the Man residue (Figure 1.12).

Biological functions of Rhizobium lipid A

As described above, *Rhizobium* lipid A has several interesting structural features (*i. e.* terminal GalA, GlcNonate, dephosphorylation, VLCFA, and 3-*O*-deacylation). Although a designated role for the unique lipid A structures in symbiosis has not been established, the biosynthetic genes have been discovered. Including this thesis, all of the biosynthetic genes required for these unique lipid A structures have been discovered as described above. Like lipid A found in enteric bacteria such as *E. coli, Rhizobum* lipid A maintains negatively charged residues. However, instead of phosphate, the presence of GalA and Glconate at the 4` and 1 positions respectively contribute to the negatively charged nature of *Rhizobium* lipid A. As discussed in chapter 4 of this thesis, an *rgtD* mutant which lacks lipid A 4` GalA was prepared and displayed increased sensitivity to detergent and polymyxin B (cationic AMP) compared to the parent strain suggesting that the 4` GalA residue functions to stabilize the outer membrane. However, disruption of the *rgtD* gene lead to partial disruption in the production of LPS core GalA and as discussed below, core GalAs also contribute to membrane stability. Lipid A GalA likely functions to cross bridge the LPS through divalent cationic interactions similar to what has been observed in enteric bacteria that produce lipid A phosphates (12,55). The *rgtD* mutant induced pink nodules when inoculated onto the host plant (pea). However, a detailed assessment is required to determine a more accurate symbiotic phenotype of the *rgtD* mutant. Disruption in the *lpxQ* gene will aid in resolving biological functions of the GlcNonate residue. It is tempting to speculate that GlcNonate contributes to ionic cross bridging of LPS as this residue is negatively charged as well. In addition, because LpxQ utilizes oxygen, perhaps LpxQ helps to regulate the oxygen tension at the outer membrane of bacteroids as decreased oxygen tension triggers global transcriptional regulation required for nitrogen fixation, *e. g.* the oxygen sensory

two component regulatory system FixL/FixJ (128) regulates the expression of nitrogen fixing gene products and a low oxygen concentration needs to be maintained for nitrogenase production and activity.

Lipid A VLCFA deficient mutants have been engineered in four endosymbionts: *Sinorhizobium meliloti* 1021 (120,121,129), *R. leguminosarum* bv. *viciae* 3841 (27,117,130), *R. leguminosarum* bv. *phaseoli* 8002 (113), and the broad host range endosymbiont *Sinorhizobium spp.* NGR234 (119). Biological studies concerning VLCFA have not been explored in *R. etli* CE3. Mutations in the *acpXL* gene of the above strains have been created and in each case the mutants produce LPS devoid of VLCFA. However, a small population of lipid A species from these mutants contained palmitate in place of VLCFA likely due to low nonspecific activity of the LpxXL transferase. *lpxXL* mutants in strain 1021 (129) and 3841 (communication with Dianna Bourassa, The University of Georgia) completely lack secondary fatty acids. As expected, the *acpXL* mutants displayed relatively decreased membrane stability as determined by detergent sensitivity which suggests that the VLCFA is important for the overall membrane stability. Interestingly, the *S. meliloti lpxXL* mutant was more susceptible to detergents than the *acpXL* mutant suggesting that the partial addition of palmitate to the lipid A contributes to membrane stability. The *R. leguminosarum* bv *viciae* 3841 (117), *R. leguminosarum* bv *phaseoli* 8002 (113), and *S. meliloti* 1021 (129) *acpXL* mutants shared common physiological phenotypes in that both strains were disrupted in membrane stability and salt tolerance. It is thought that salt tolerance is important for effective symbiosis and, interestingly, isolated *acpXL* mutant nodule bacteria were restored to parent strain levels of salt tolerance further suggesting the necessity for osmotic tolerance during symbiosis. It is not understood what role AcpXL has in salt tolerance. Like *E. coli, Rhziobium* produce osmoregulated periplasmic and excreted β-glucans that have

been implicated in osmo-adaptation (131,132). The general acyl carrier protein (Acp) in *E. coli* is essential for β-glucan synthesis (133-135). Therefore, it is tempting to speculate that AcpXL may be involved in the synthesis of β-glucan in *Rhizobium*. The *R. leguminosarum acpXL* mutants should be assayed for the presence of β-glucan. An *in vitro* assay for the synthesis of β-glucan has been established in *R. leguminosarum* bv. *trifolii* TA1 (136) by assaying for the formation of β-glucans from isotopically labeled UDP-[^{14}C]Glucose and whole cell preparations. A similar assay may be helpful to investigate the role AcpXL may have in β-glucan synthesis.

VLCFA deficient mutants of *R. leguminosarum* bv *viciae* 3841, *R. leguminosarum* bv *phaseoli* 8002, *S. meliloti* 1021, and *Sinorhizobium spp.* NGR234 have been assayed for their ability to form complete nitrogen fixing symbiosis (symbiotic phenotype) with their compatible host legumes (113,119,121,130). The 3841, 8002, and 1021 *acpXL* mutants were able to form nitrogen fixing symbiosis with their respective host plants. However, detailed dissection of the symbiotic process by electron microscopy revealed that the infection process was significantly delayed and aberrant symbiosome formation was observed. The delayed phenotype could be attributed to the disruption in salt tolerance. As discussed in Chapter 2, salt tolerant variants of the *R. leguminosarum* bv *viciae* 3841 *acpXL* mutants are likely selected for by the host plant environment as the nodule isolated bacteria recover salt tolerance to parent strain levels. In fact, when nodule isolated 3841 *acpXL* mutants were inoculated onto the host plant, there was no longer a delay in the infection process (communication with Janine Sherrier, University of Delaware). However, the bacteroids remained aberrantly shaped. The symbiotic phenotype of the NGR234 *acpXL* mutant was more severe. Strain NGR234 forms symbiosis with a broad range of hosts. Interestingly, when inoculated on four host plants, the NGR234 *acpXL* mutant was not able to form symbiosis with three of the four host plants (119). Unlike the other *acpXL* mutants

mentioned, the NGR234 was disrupted in O-chain biosynthesis as determined by LPS PAGE profile but not in the flavanoid induced high molecular weight rhamnose containing LPS (137). The loss of normal O-chain could be the reason for the severe symbiotic phenotype observed for the NGR234 *acpXL* mutant as it has been shown in many systems that O-chain synthesis is required for complete symbiosis (20).

As determined by mass spectrometry, some endosymbionts produce a lipid A species consistent with a 3-*O*-deacylated structure (90,138). It was observed that the outer membranes of some *Rhizboium* strains contained 3-*O*-deacylase activity (114) and was likely due to a PagL like enzyme (3-*O*-deacylase) that is also present in the outer membranes of pathogens (54). PagL modification of lipid A has been implicated in the evasion and resistance of the host immune system in animal pathogens (54). D`Haeze et al. (90) demonstrated that *R. etli* bacteroids exclusively produced 3-*O*-deacylated lipid A while free living *R. etli* produces a mixture of 3-*O*-deacylated lipid A suggesting the importance of a PagL like enzyme during symbiosis with *Phaseolus*. Chapter 3 of this thesis describes the characterization of a gene (*pagL*) required for 3-*O*-deacylation in *R. etli* CE3 and this is the first *pagL* described for a symbiont. Analysis of the *R. etli* CE3 genome sequence allowed identification of a *pagL* homologue in *R. etli* CE3 and in the genomes of symbiotic strains known to produce 3-*O*-deacylated lipid A. A *pagL*-minus mutant was created in *R. etli* CE3 and this mutant was severely disrupted in membrane stability in response to detergent. Interestingly, the *pagL* mutant was more resistant to polymyxin B than parent strain. When inoculated onto the host plant (common bean), the *pagL* mutant induced pink nodules, however, the plants remained yellow at 26 days post inoculation (indicative of nitrogen starvation) suggesting that the PagL protein is important for *R. etli* CE3 - bean symbiosis. The membrane instability of the *pagL* mutant may be the reason for the observed symbiotic

phenotype. However, isolated *pagL* mutant nodule bacteria had comparable membrane stability to parent strain isolates and maintained increased resistance to polymyxin B. When the isolated *pagL* mutant nodule bacteria where re-inoculated onto host plants, a similar plant phenotype was observed (yellow leaves) suggesting that the membrane stability alone was not responsible for the observed phenotype (data not shown). The biological function of PagL in symbionts remains unclear. Because the broad host range symbiont NGR234 has a *pagL* homologue and produces 3-*O*-deacylated lipid A, a NGR234 *pagL* mutant may provide insight into the requirements of PagL in different symbiotic systems.

Biological Functions of Rhizobium LPS galacturonic acid residues

As described above, *Rhizobium* core is heavily galacturonosylated and the genes necessary for the biosynthesis of the four terminal GalA residues have been elucidated. As discussed in Chapters 4, in order to dissect the gene function of *rgtE* and *rgtD* as well as determine the biological function of lipid A and core GalAs, five single gene mutants were prepared in the LPS GalA transferases (*RgtABCDE*) of *R. leguminosarum* 3841. The *rgtE* mutant did not contain the necessary lipid donor (Dod-P-GalA) for LPS GalA transfer and therefore the LPS completely lacked GalA. These results confirmed that Dod-P-GalA is necessary for addition of all GalA residues to the LPS GalA (127), and also showed that *rgtE* encodes the GalAT required for the synthesis of Dod-P-GalA. The *rgtD* mutant did not contain 4` GalA on the lipid A confirming that this gene is the 4` GalA transferase. The *rgtD* gene is present in all of the sequenced genomes of *Rhizobium* species and *rgtD* homologues are found in the

hyperthermophile *Aquifex aeolicus* and the stalk forming *Caulobacter crescentus* (Figure 1.6), two organisms that contain 4` GalA in their lipid A (139,140).

Interestingly, the *rgtA, rgtB,* and *rgtC* mutants (disrupted in core GalA production) displayed relatively increased resistance to the cationic AMP polymyxin B. It is likely that the negatively charged GalA residues aid in the recruitment of polymyxin B and the removal of the core GalAs as in the case of the *rgtABC* mutants lessens the efficiency of polymyxin B binding to core. In contrast, the *rgtE* mutant (completely devoid of GalA) and *rgtD* mutant (devoid of lipid A GalA and partially disrupted in core GalA production) were hypersensitive to polymyxin B suggesting that the presence of 4`GalA is necessary for normal polymyxin B resistance. The *rgt* mutants displayed varying susceptibility to detergent. The *rgtC* was as resistant to detergent as the parent strain while the *rgtA, B, D,* and *E* mutants were more susceptible; the *rgtE* mutant was the most susceptible. These results suggest that the core GalA as well as the lipid A GalA contribute to membrane stability likely through ionic cross bridging of the LPSs in the outer membrane. The GalA residues may also aid in the recruitment of nodule specific cysteine rich peptides (NCRs) that resemble cationic AMPs. NCRs interact with the bacterial cell surface and are the causative agent for the terminal differentiation of *S. meliloti* 1021 bacteroids (141), a general feature of bacteroid formation in indeterminate nodulating systems and is defined by disruption in cell division (cell enlargement) and endoreduplication of replicons (polyploid). This represents one example demonstrating how the symbiotic process likely originated from the manipulation of the host defense system. It is possible that NCRs regulate bacteroid formation in *Rhizobium* species as well and it is likely that NCRs interact with LPS. Although *R. etli* is a determinate system and does not undergo terminal differentiation, the bacteroids become enlarged and it is not known if endoreduplication occurs. In addition, *S. meliloti* contains

terminal GalA and GlcA in the LPS core. Interestingly, there are only two genes in the *S. meliloti* genome that are annotated as ArnT-like bacterial glycosyl transferases and these genes are likely responsible for the attachment of GalA and GlcA to the core of *S. meliloti* 1021. Engineering GalA and GlcA core mutants in *Sinorhizobium* will be helpful in determining the possible role core anionic residues have in NCR perception.

The *rgt* mutants were able to form nitrogen fixing symbiosis in the laboratory indicating that the presence of core GalA is not absolutely required for complete symbiosis. However, plants inoculated with the *rgtE* mutant displayed premature wilting of the early leaves and yellowing at 14 days post inoculation. These effects could be due to early nitrogen starvation or an early plant defense response. A more accurate symbiotic phenotype needs to be defined for the *rgt* mutants and host plant (pea).

An unusual lipid A α-1,1-GalA residue has been reported in several interesting organisms (Figure 1.6); *Azospirillum lipoferum* (142), *Caulobacter crescentus* (140) , *Mesorhizobium huakuii* (143), and *Aquifex aeolicus* (144). *C. crescentus* and *A. aeolicus* lipid As are bis-galacturonylated at the 4` and 1 positions and indeed these organisms contain *rgtD* homologues. Chapter 5 of this thesis describes the elucidation of the α-1,1-GalA transferase gene. from *Mesorhizobium loti* MAFF303099. Genome comparison of the organisms that produce lipid A α-1,1-GalA revealed that a common *arnT*-like gene exists between organisms that contain lipid A α-1,1-GalA and was given the gene name *rgtF*. The *rgtF* function was elucidated by introducing the *rgtF* gene into *R. etli* strains and indeed the *rgtF* complemented strains produced α-1,1-GalA. Interestingly, the *rgtF* complemented *R. etli* CE3 contained a mixture of mono and bisgalacturonylated lipid A and this strain became significantly more resistant to detergent and polymyxin B compared to parent strain indicating an increased stability to the outer membrane.

A. aeoloicus and *C. crescentus* may have evolved to produce bis-galacturonosylated lipid A in order to maintain a more ridged cellular envelope. These organisms contain a diaminoglucosamine (DAG) in their lipid A backbones as well. DAG containing lipid As are more resistant to hydrolysis. *A. aeolicus* is a hyperthermophillic bacterium that optimally grows at 95^{o}C and it may be advantageous to maintain a more ridged heat resistant cellular envelope as these organisms additionally contain a large amount of membrane derived ether linked glycerides (145). Disruption of the *rgtF* genes in these organisms will aid in elucidating the functions of lipid A α-1,1-GalA.

Summary

In order to begin to understand in detail the biological significance of LPS in pathogenic and symbiotic infection, it is crucial to elucidate the LPS structure and the biosynthetic genes and gene products to determine the structural basis for biological activity. As discussed in this chapter, the major pathways for LPS perception are defined in mammalian systems and we are learning how pathogens manipulate these pathways by varying their LPS structure. While LPS is a known virulence and symbiotic factor for plant infection and a known elicitor of the plant defense response, the underlying pathway/s for LPS perception by plants is unknown. It is possible that the unique LPS structural features produced by plant endosymbionts help promote infection and survival in the various stages of plant infection. In this thesis, we describe novel genes involved in the biosynthesis of these unique LPS structures produced by the endosymbionts *Rhizobium leguminosarum, R. etli,* and *Mesorhizobium loti.* We have created several mutants impaired in specific LPS biosynthetic steps and describe the effects of these mutations concerning LPS structure, membrane stability, and symbiosis with host plants. In

addition, these mutants will aid in future studies regarding the structural basis for LPS and the plant defense response.

References

1. Oldroyd, G. E. D., and Downie, J. M. (2008) *Annu. Rev. Plant Biol.* **59**, 519-546
2. Sprent, J. (2005) Nodulated Legume Trees. in *Nitrogen Fixation in Agriculture, Forestry, Ecology, and the Environment* (Werner, D., and Newton, W. eds.), Springer Netherlands. pp 113-141
3. Gruber, N., and Galloway, J. N. (2008) *Nature* **451**, 293-296
4. Tilman, D. (1999) *Proc. Natl. Acad. Sci.* **96**, 5995-6000
5. Michael Beman, J., Arrigo, K. R., and Matson, P. A. (2005) *Nature* **434**, 211-214
6. Liu, J., You, L., Amini, M., Obersteiner, M., Herrero, M., Zehnder, A. J. B., and Yang, H. (2010) *Proc. Natl. Acad. Sci.* **107**, 8035-8040
7. Galloway, J. N., Townsend, A. R., Erisman, J. W., Bekunda, M., Cai, Z., Freney, J. R., Martinelli, L. A., Seitzinger, S. P., and Sutton, M. A. (2008) *Science* **320**, 889-892
8. Guo, J. H., Liu, X. J., Zhang, Y., Shen, J. L., Han, W. X., Zhang, W. F., Christie, P., Goulding, K. W. T., Vitousek, P. M., and Zhang, F. S. (2010) *Science* **327**, 1008-1010
9. Palm, C. A., Smukler, S. M., Sullivan, C. C., Mutuo, P. K., Nyadzi, G. I., and Walsh, M. G. (2010) *Proc. Natl. Acad. Sci.* **107**, 19661-19666
10. Kannenberg, E. L., and Brewin, N. J. (1994) *Trends in Microbiol.* **2**, 277-283
11. Nicolas, F., François, C., and Véréna, P. (2003) *Euro. J. Biochem.* **270**, 1365-1380
12. Raetz, C. R., and Whitfield, C. (2002) *Annu. Rev. Biochem.* **71**, 635-700

13. Raetz, C. R. H., Reynolds, C. M., Trent, M. S., and Bishop, R. E. (2007) *Annu. Rev. Biochem.* **76**, 295-329

14. Raetz, C. R., Guan, Z., Ingram, B. O., Six, D. A., Song, F., Wang, X., and Zhao, J. (2009) *J. Lipid Res.* **50 Suppl**, S103-108

15. Newman, M. A., Dow, J. M., Molinaro, A., and Parrilli, M. (2007) *J. Endotoxin Res.* **13**, 69-84

16. Carlson, R. W., Forsberg, L. S., and Kannenberg, E. L. (2010) Lipopolysaccharides in Rhizobium-Legume Symbioses. in *Subcellular Biochemestry: Endotoxins: Structure, Function and Recognition*, P. J. Quinn and X. Wang Ed., Springer. pp 339-386

17. Kannenberg, E. L., Reuhs, B. L., Forsberg, L. S., and Carlson, R. W. (1998) Lipopolysaccharides and K-Antigens: Their Structures, Biosynthesis, and Functions. in *The Rhizobiaceae: molecular biology of model plant-associated bacteria* (Spaink, H. P., Kondorosi, A., and Hooykaas, P. J. J. eds.), Kluwer Academic Publishers, Dordrecht, The Netherlands. pp 119-154

18. Ormeño-Orrillo, E., Rosenblueth, M., Luyten, E., Vanderleyden, J., and Martínez-Romero, E. (2008) *Environ. Microbiol.* **10**, 1271-1284

19. Erbs, G., and Newman, M. A. (2011) *Mol. Plant Pathol.*

20. Carlson, R. W., Forsberg, L. S., and Kannenberg, E. L. (2010) Lipopolysaccharides in *Rhizobium*-Legume Symbioses in *Endotoxins: Structure, Function and Recognition* (Wang, X., and Quinn, P. J. eds.), Springer Netherlands. pp 339-386

21. Anderson, M. S., and Raetz, C. R. (1987) *J. Biol. Chem.* **262**, 5159-5169

22. Anderson, M. S., Bulawa, C. E., and Raetz, C. R. (1985) *J. Biol. Chem.* **260**, 15536-15541

23. Anderson, M. S., Bull, H. G., Galloway, S. M., Kelly, T. M., Mohan, S., Radika, K., and Raetz, C. R. (1993) *J. Biol. Chem.* **268**, 19858-19865

24. Wyckoff, T. J., Lin, S., Cotter, R. J., Dotson, G. D., and Raetz, C. R. (1998) *J .Biol. Chem.* **273**, 32369-32372

25. Metzger, L. E., and Raetz, C. R. H. (2010) *Biochem.* **49**, 6715-6726

26. Basu, S. S., Karbarz, M. J., and Raetz, C. R. H. (2002) *J .Biol. Chem.* **277**, 28959-28971

27. Vedam, V., Kannenberg, E., Datta, A., Brown, D., Haynes-Gann, J. G., Sherrier, D. J., and Carlson, R. W. (2006) *J. Bacteriol.* **188**, 2126-2133

28. Ramada Bhat, U., Carlson, R. W., Busch, M., and Mayer, H. (1991) *Internatl. J. Syst. Bacteriol.* **41**, 213-217

29. Li, C., Guan, Z., Liu, D., and Raetz, C. R. (2011) *Proc. Natl. Acad. Sci.* **108**, 11387-11392

30. Isobe, T., White, K. A., Allen, A. G., Peacock, M., Raetz, C. R., and Maskell, D. J. (1999) *J. Bacteriol.* **181**, 2648-2651

31. Stead, C., Tran, A., Ferguson, D., Jr., McGrath, S., Cotter, R., and Trent, S. (2005) *J. Bacteriol.* **187**, 3374-3383

32. Stead, C. M., Zhao, J., Raetz, C. R., and Trent, M. S. (2010) *Mol. Microbiol.* **78**, 837-852

33. Chung, H. S., and Raetz, C. R. H. (2011) *Proc. Natl. Acad. Sci.* **108**, 510-515

34. Moreno, E., Pitt, M. W., Jones, L. M., Schurig, G. G., and Berman, D. T. (1979) *J. Bacteriol.* **138**, 361-369

35. Wang, X., Karbarz, M. J., McGrath, S. C., Cotter, R. J., and Raetz, C. R. H. (2004) *J. Biol.Chem.* **279**, 49470-49478

36. Wang, X., McGrath, S. C., Cotter, R. J., and Raetz, C. R. H. (2006) *J .Biol. Chem.* **281**, 9321-9330

37. Karbarz, M. J., Kalb, S. R., Cotter, R. J., and Raetz, C. R. (2003) *J Biol Chem* **278**, 39269-39279

38. Karbarz, M. J., Six, D. A., and Raetz, C. R. H. (2009) *J .Biol. Chem.* **284**, 414-425

39. Reynolds, C. M., Kalb, S. R., Cotter, R. J., and Raetz, C. R. H. (2005) *J .Biol. Chem.* **280**, 21202-21211

40. Touzé, T., Tran, A. X., Hankins, J. V., Mengin-Lecreulx, D., and Trent, M. S. (2008) *Mol. Microbiol.* **67**, 264-277

41. Lee, H., Hsu, F.-F., Turk, J., and Groisman, E. A. (2004) *J. Bacteriol.* **186**, 4124-4133

42. Brabetz, W., Müller-Loennies, S., and Brade, H. (2000) *J. Biol.Chem.* **275**, 34954-34962

43. White, K. A., Lin, S., Cotter, R. J., and Raetz, C. R. H. (1999) *J. Biol. Chem.* **274**, 31391-31400

44. Kanjilal-Kolar, S., Basu, S. S., Kanipes, M. I., Guan, Z., Garrett, T. A., and Raetz, C. R. (2006) *J. Biol. Chem.* **281**, 12865-12878

45. Feldman, M. F., Marolda, C. L., Monteiro, M. A., Perry, M. B., Parodi, A. J., and Valvano, M. A. (1999) *J. Biol. Chem.* **274**, 35129-35138

46. Bastin, D. A., Stevenson, G., Brown, P. K., Haase, A., and Reeves, P. R. (1993) *Mol. Microbiol.* **7**, 725-734

47. Morona, R., van den Bosch, L., and Manning, P. (1995) *J. Bacteriol.* **177**, 1059-1068

48. Woodward, R., Yi, W., Li, L., Zhao, G., Eguchi, H., Sridhar, P. R., Guo, H., Song, J. K., Motari, E., Cai, L., Kelleher, P., Liu, X., Han, W., Zhang, W., Ding, Y., Li, M., and Wang, P. G. (2010) *Nat. Chem. Biol*. **6**, 418-423

49. Sperandeo, P., Deho, G., and Polissi, A. (2009) *Biochim. Biophys. Acta.* **1791**, 594-602

50. Kotani, S., Takada, H., Tsujimoto, M., Ogawa, T., Takahashi, I., Ikeda, T., Otsuka, K., Shimauchi, H., Kasai, N., Mashimo, J., and et al. (1985) *Infect. Immun.* **49**, 225-237

51. Galanos, C., Luderitz, O., Rietschel, E. T., Westphal, O., Brade, H., Brade, L., Freudenberg, M., Schade, U., Imoto, M., Yoshimura, H., and et al. (1985) *Euro. J. of Biochem.* **148**, 1-5

52. Loppnow, H., Brade, H., Durrbaum, I., Dinarello, C., Kusumoto, S., Rietschel, E., and Flad, H. (1989) *J. of Immunol.* **142**, 3229-3238

53. Raetz, C. R. H., Garrett, T. A., Reynolds, C. M., Shaw, W. A., Moore, J. D., Smith, D. C., Ribeiro, A. A., Murphy, R. C., Ulevitch, R. J., Fearns, C., Reichart, D., Glass, C. K., Benner, C., Subramaniam, S., Harkewicz, R., Bowers-Gentry, R. C., Buczynski, M. W., Cooper, J. A., Deems, R. A., and Dennis, E. A. (2006) *J. Lipid Res.* **47**, 1097-1111

54. Trent, M. S., Stead, C. M., Tran, A. X., and Hankins, J. V. (2006) *J. Endotoxin Res.* **12**, 205-223

55. Snyder, S., Kim, D., and McIntosh, T. J. (1999) *Biochem.* **38**, 10758-10767

56. Komaniecka, I., Choma, A., Lindner, B., and Holst, O. (2010) *Chem.* **16**, 2922-2929

57. Sánchez Carballo, P. M., Rietschel, E. T., Kosma, P., and Zähringer, U. (1999) *Euro. J. Biol. Chem.* **261**, 500-508

58. Park, B. S., Song, D. H., Kim, H. M., Choi, B. S., Lee, H., and Lee, J. O. (2009) *Nature* **458**, 1191-U1130

59. Akira, S., and Takeda, K. (2004) *Nat. Rev. Immunol.* **4**, 499-511

60. Nguyen, H. B., Rivers, E. P., Abrahamian, F. M., Moran, G. J., Abraham, E., Trzeciak, S., Huang, D. T., Osborn, T., Stevens, D., and Talan, D. A. (2006) *Annu. Emerg. Med.* **48**, 28-54

61. Bäckhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A., and Gordon, J. I. (2005) *Science* **307**, 1915-1920

62. Cani, P. D., Amar, J., Iglesias, M. A., Poggi, M., Knauf, C., Bastelica, D., Neyrinck, A. M., Fava, F., Tuohy, K. M., Chabo, C., Waget, A., Delmée, E., Cousin, B., Sulpice, T., Chamontin, B., Ferrières, J., Tanti, J.-F., Gibson, G. R., Casteilla, L., Delzenne, N. M., Alessi, M. C., and Burcelin, R. (2007) *Diabetes* **56**, 1761-1772

63. Permana, P. A., Menge, C., and Reaven, P. D. (2006) *Biochem. Biophy. Res. Commun.* **341**, 507-514

64. Ajuwon, K. M., and Spurlock, M. E. (2005) *J.Nutr.* **135**, 1841-1846

65. Han, C. Y., Kargi, A. Y., Omer, M., Chan, C. K., Wabitsch, M., O'Brien, K. D., Wight, T. N., and Chait, A. (2010) *Diabetes* **59**, 386-396

66. Suganami, T., Nishida, J., and Ogawa, Y. (2005) *Arterioscler. Thromb. Vas. Biol.* **25**, 2062-2068

67. Suganami, T., Tanimoto-Koyama, K., Nishida, J., Itoh, M., Yuan, X., Mizuarai, S., Kotani, H., Yamaoka, S., Miyake, K., Aoe, S., Kamei, Y., and Ogawa, Y. (2007) *Arterioscler. Thromb. Vas. Biol.* **27**, 84-91

68. Schwartz, E. A., Zhang, W.-Y., Karnik, S. K., Borwege, S., Anand, V. R., Laine, P. S., Su, Y., and Reaven, P. D. (2010) *Arterioscler. Thromb. Vas. Biol.* **30**, 802-808

69. Clausell, A., Garcia-Subirats, M., Pujol, M., Busquets, M. A., Rabanal, F., and Cajal, Y. (2006) *J. Phy. Chem. B* **111**, 551-563

70. Ganz, T. (2003) *Nat. Rev. Immunol.* **3**, 710-720

71. Oppenheim, J. J., Biragyn, A., Kwak, L. W., and Yang, D. (2003) *Annu. Rheum. Dis.* **62 Suppl 2**, ii17-21

72. Berezow, A. B., Ernst, R. K., Coats, S. R., Braham, P. H., Karimi-Naser, L. M., and Darveau, R. P. (2009) *Microb. Pathog.* **47**, 68-77

73. Lewis, L. A., Choudhury, B., Balthazar, J. T., Martin, L. E., Ram, S., Rice, P. A., Stephens, D. S., Carlson, R., and Shafer, W. M. (2009) *Infect. Immun.* **77**, 1112-1120

74. Breazeale, S. D., Ribeiro, A. A., McClerren, A. L., and Raetz, C. R. H. (2005) *J. Biol. Chem.* **280**, 14154-14167

75. Trent, M. S., Ribeiro, A. A., Doerrler, W. T., Lin, S., Cotter, R. J., and Raetz, C. R. H. (2001) *J. Biol. Chem.* **276**, 43132-43144

76. Trent, M. S., Ribeiro, A. A., Lin, S., Cotter, R. J., and Raetz, C. R. H. (2001) *J. Biol. Chem.* **276**, 43122-43131

77. Yan, A., Guan, Z., and Raetz, C. R. H. (2007) *J. Biol. Chem.* **282**, 36077-36089

78. Wang, X., Ribeiro, A. A., Guan, Z., McGrath, S. C., Cotter, R. J., and Raetz, C. R. H. (2006) *Biochem.* **45**, 14427-14440

79. Song, F., Guan, Z., and Raetz, C. R. (2009) *Biochem.* **48**, 1173-1182

80. Kanistanon, D., Hajjar, A. M., Pelletier, M. R., Gallagher, L. A., Kalhorn, T., Shaffer, S. A., Goodlett, D. R., Rohmer, L., Brittnacher, M. J., Skerrett, S. J., and Ernst, R. K. (2008) *PLoS Pathog* **4**, e24

81. Wang, X., Ribeiro, A. A., Guan, Z., and Raetz, C. R. H. (2009) *Biochem.* **48**, 1162-1172

82. Wang, X., Ribeiro, A. A., Guan, Z., Abraham, S. N., and Raetz, C. R. H. (2007) *Proc. Natl. Acad. Sci.* **104**, 4136-4141

83. Marr, N., Tirsoaga, A., Blanot, D., Fernandez, R., and Caroff, M. (2008) *J. Bacteriol.* **190**, 4281-4290

84. Pilione, M. R., Pishko, E. J., Preston, A., Maskell, D. J., and Harvill, E. T. (2004) *Infect. Immun.* **72**, 2837-2842

85. Preston, A., Maxim, E., Toland, E., Pishko, E. J., Harvill, E. T., Caroff, M., and Maskell, D. J. (2003) *Mol. Microbiol.* **48**, 725-736

86. Geurtsen, J., Steeghs, L., Hove, J. T., van der Ley, P., and Tommassen, J. (2005) *J. Biol. Chem.* **280**, 8248-8259

87. Kawasaki, K., Ernst, R. K., and Miller, S. I. (2004) *J. Biol. Chem.* **279**, 20044-20048

88. Geurtsen, J., Steeghs, L., Hamstra, H.-J., ten Hove, J., de Haan, A., Kuipers, B., Tommassen, J., and van der Ley, P. (2006) *Infect. Immun.* **74**, 5574-5585

89. Ernst, R. K., Adams, K. N., Moskowitz, S. M., Kraig, G. M., Kawasaki, K., Stead, C. M., Trent, M. S., and Miller, S. I. (2006) *J. Bacteriol.* **188**, 191-201

90. D'Haeze, W., Leoff, C., Freshour, G., Noel, K. D., and Carlson, R. W. (2007) *J. Biol. Chem.* **282**, 17101-17113

91. Gunn, J. S. (2008) *Trends Microbiol.* **16**, 284-290

92. Tamayo, R., Choudhury, B., Septer, A., Merighi, M., Carlson, R., and Gunn, J. S. (2005) *J. Bacteriol.* **187**, 3391-3399

93. Auffray, C., Sieweke, M. H., and Geissmann, F. (2009) *Annu. Rev. Immunol.* **27**, 669-692

94. Klena, J., Zhang, P., Schwartz, O., Hull, S., and Chen, T. (2005) *J. Bacteriol.* **187**, 1710-1715

95. Zhang, P., Snyder, S., Feng, P., Azadi, P., Zhang, S., Bulgheresi, S., Sanderson, K. E., He, J., Klena, J., and Chen, T. (2006) *J. Immunol.* **177**, 4002-4011

96. van Vliet, S. J., Steeghs, L., Bruijns, S. C. M., Vaezirad, M. M., Snijders Blok, C., Arenas Busto, J. A., Deken, M., van Putten, J. P. M., and van Kooyk, Y. (2009) *PLoS Pathog.* **5**, e1000625

97. Zhang, P., Schwartz, O., Pantelic, M., Li, G., Knazze, Q., Nobile, C., Radovich, M., He, J., Hong, S.-C., Klena, J., and Chen, T. (2006) *J. Leukoc. Biol.* **79**, 731-738

98. Shafer, W. M., Datta, A., Kumar Kolli, V. S., Mahbubur Rahman, M., Balthazar, J. T., Martin, L. E., Veal, W. L., Stephens, D. S., and Carlson, R. (2002) *J. Endotoxin Res.* **8**, 47-58

99. Burns, S. M., and Hull, S. I. (1998) *Infect. Immunol.* **66**, 4244-4253

100. Cortes, G., Borrell, N., de Astorza, B., Gomez, C., Sauleda, J., and Alberti, S. (2002) *Infect. Immunol.* **70**, 2583-2590

101. Morona, R., Daniels, C., and Van Den Bosch, L. (2003) *Microbiol.* **149**, 925-939

102. Murray, G. L., Attridge, S. R., and Morona, R. (2006) *J. Bacteriol*. **188**, 2735-2739

103. Murray, G. L., Attridge, S. R., and Morona, R. (2003) *Mol. Microbiol.* **47**, 1395-1406

104. Stevenson, G., Kessler, A., and R. Reeves, P. (1995) *FEMS Microbiol. Letters* **125**, 23-30

105. Joiner, K., Grossman, N., Schmetz, M., and Leive, L. (1986) *J. Immunol.* **136**, 710-715

106. Grossman, N., Schmetz, M. A., Foulds, J., Klima, E. N., Jimenez-Lucho, V. E., Leive, L. L., Joiner, K. A., and Jiminez, V. (1987) *J. Bacteriol.* **169**, 856-863

107. Tomás, J. M., Ciurana, B., Benedí, V. J., and Juarez, A. (1988) *J. Gen. Microbiol.* **134**, 1009-1016

108. Moran, A. P., and Prendergast, M. M. (2001) *J. Autoimmun.* **16**, 241-256

109. Nairn, C. A., Cole, J. A., Patel, P. V., Parsons, N. J., Fox, J. E., and Snith, H. (1988) *J. Gen. Microbiol.* **134**, 3295-3306

110. Kugelberg, E., Gollan, B., and Tang, C. M. (2008) *Vaccine* **26**, I34-I39

111. Tsai, C. M., Mocca, L. F., and Frasch, C. E. (1987) *Infect. Immun.* **55**, 1652-1656

112. Bhat, U. R., Forsberg, L. S., and Carlson, R. W. (1994) *J. Biol. Chem.* **269**, 14402-14410

113. Brown, D. B., Huang, Y.-C., Kannenberg, E. L., Sherrier, D. J., and Carlson, R. W. (2011) *J. Bacteriol.*, JB.00392-00311

114. Basu, S. S., White, K. A., Que, N. L. S., and Raetz, C. R. H. (1999) *J. Biol. Chem.* **274**, 11150-11158

115. Price, N. P., Kelly, T. M., Raetz, C. R., and Carlson, R. W. (1994) *J. Bacteriol.* **176**, 4646-4655

116. Brozek, K. A., Carlson, R. W., and Raetz, C. R. H. (1996) *J. Biol. Chem.* **271**, 32126-32136

117. Vedam, V., Kannenberg, E. L., Haynes, J. G., Sherrier, D. J., Datta, A., and Carlson, R. W. (2003) *J. Bacteriol.* **185**, 1841-1850

118. Sharypova, L. A., Niehaus, K., Scheidle, H., Holst, O., and Becker, A. (2003) *J. Biol. Chem.* **278**, 12946-12954

119. Ardissone, S., Kobayashi, H., Kambara, K., Rummel, C., Noel, K. D., Walker, G. C., Broughton, W. J., and Deakin, W. J. (2011) *J. Bacteriol.* **193**, 2218-2228

120. Haag, A. F., Wehmeier, S., Muszynski, A., Kerscher, B., Fletcher, V., Berry, S. H., Hold, G. L., Carlson, R. W., and Ferguson, G. P. (2011) *J. Biol. Chem.* **286**, 17455-17466

121. Haag, A. F., Wehmeier, S., Beck, S., Marlow, V. L., Fletcher, V., James, E. K., and Ferguson, G. P. (2009) *J. Bacteriol.* **191**, 4681-4686

122. Vanderlinde, E. M., Muszynski, A., Harrison, J. J., Koval, S. F., Foreman, D. L., Ceri, H., Kannenberg, E. L., Carlson, R. W., and Yost, C. K. (2009) *Microbiology* **155**, 3055-3069

123. Ingram, B. O., Sohlenkamp, C., Geiger, O., and Raetz, C. R. (2010) *Biochim. Biophys. Acta.* **1801**, 593-604

124. Price, N. P., Jeyaretnam, B., Carlson, R. W., Kadrmas, J. L., Raetz, C. R., and Brozek, K. A. (1995) *Proc.Natl. Acad. Sci.* **92**, 7352-7356

125. Que-Gewirth, N. L., Karbarz, M. J., Kalb, S. R., Cotter, R. J., and Raetz, C. R. (2003) *J. Biol. Chem.* **278**, 12120-12129

126. Kadrmas, J. L., Allaway, D., Studholme, R. E., Sullivan, J. T., Ronson, C. W., Poole, P. S., and Raetz, C. R. (1998) *J. Biol. Chem.* **273**, 26432-26440

127. Kanjilal-Kolar, S., and Raetz, C. R. (2006) *J. Biol. Chem.* **281**, 12879-12887

128. Buelow, D. R., and Raivio, T. L. (2010) *Mol. Microbiol.***75**, 547-566

129. Ferguson, G. P., Datta, A., Carlson, R. W., and Walker, G. C. (2005) *Mol. Microbiol.* **56**, 68-80

130. Vedam, V., Haynes, J. G., Kannenberg, E. L., Carlson, R. W., and Sherrier, D. J. (2004) *Mol. Plant Microbe Interact.* **17**, 283-291

131. Breedveld, M. W., and Miller, K. J. (1994) *Microbiol. Mol. Biol. Rev.* **58**, 145-161

132. Bohin, J.-P. (2000) *FEMS Microbiol. Letters* **186**, 11-19

133. Therisod, H., and Kennedy, E. P. (1987) *Proc. Natl. Acad. Sci.* **84**, 8235-8238

134. Tang, L., Weissborn, A., and Kennedy, E. (1997) *J. Bacteriol.* **179**, 3697-3705

135. Therisod, H., Weissborn, A. C., and Kennedy, E. P. (1986) *Proc.Natl. Acad. Sci.* **83**, 7236-7240

136. Breedveld, M. W., Zevenhuizen, L. P., and Zehnder, A. J. (1992) *J. Bacteriol.* **174**, 6336-6342

137. Reuhs, B. L., Relic, B., Forsberg, L. S., Marie, C., Ojanen-Reuhs, T., Stephens, S. B., Wong, C.-H., Jabbouri, S., and Broughton, W. J. (2005) *J. Bacteriol.* **187**, 6479-6487

138. Gudlavalleti, S. K., and Forsberg, L. S. (2003) *J. Biol. Chem.* **278**, 3957-3968

139. Plotz, B. M., Lindner, B., Stetter, K. O., and Holst, O. (2000) *J. Biol. Chem. Chemistry* **275**, 11222-11228

140. Smit, J., Kaltashov, I. A., Cotter, R. J., Vinogradov, E., Perry, M. B., Haider, H., and Qureshi, N. (2008) *Inn. Immun.* **14**, 25-36

141. Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., Farkas, A., Mikulass, K., Nagy, A., Tiricz, H., Satiat-Jeunemaitre, B., Alunni, B., Bourge, M., Kucho, K., Abe, M., Kereszt, A., Maroti, G., Uchiumi, T., Kondorosi, E., and Mergaert, P. (2010) *Science* **327**, 1122-1126

142. Choma, A., and Komaniecka, I. (2008) *Carbohydr. Res.* **343**, 799-804

143. Choma, A., and Sowinski, P. (2004) *Euro. J.Biochem.* **271**, 1310-1322

144. Plotz, B. M., Lindner, B., Stetter, K. O., and Holst, O. (2000) *J. Biol. Chem.* **275**, 11222-11228

145. Jahnke, L. L., Eder, W., Huber, R., Hope, J. M., Hinrichs, K.-U., Hayes, J. M., Des Marais, D. J., Cady, S. L., and Summons, R. E. (2001) *Appl. Environ. Microbiol.* **67**, 5179-5189

Figure Legends

Figure 1.1. The *E. coli* lipid A biosynthetic pathway as described in the text. Enzymes are in parenthesis and common name intermediate products are in quotations.

Figure 1.2. A general schematic diagram of LPS assembly in gram negative bacteria. 1. Lipid A and core synthesis occurs at the cytosolic/inner membrane interface and is 2. flipped to the periplasmic side by the ABC transporter MsbA. 3. Further core and lipid A assembly or modifications may occur at the periplasmic interface. At this point, lipid A-core may be transported to the outer membrane (5.) or 4. O-chain is attached as described in the text. 5. Finally, the fully formed LPS is transported across the periplasm and inserted into the outer membrane likely due to the combined actions of Lpt proteins. 6. and 7. Some organisms are able to further modify their LPS/LOS in the outer membrane.

Figure 1.3. A representative model for the LPS transport and Lpt protein complex assembly. Adapted from Sperandeo et al (49).

Figure4. The progression of LPS perception and initiation of innate immune tissues (monocytes) as described in the text. Adapted and modified from Raetz and Whitfield (12).

Figure 1.5. A conserved lipid A glycosylation pathway. The formation of an inner membrane lipid sugar donor is synthesized by the ArnC like glycosyltransferase (distantly related to the eukarotic dolichol phosphate mannosyl transferase catalytic domain, DPM1) on the cyplasmic

side of the inner membrane and flipped to the periplasmic side. Integral ArnT-like enzymes (distantly related to the eukaryotic phosphomannosyl transferase, PMT) facilitate transfer of the glycosyl residue from the lipid donor to the lipid A.

Figure 1.6. Lipid A carbohydrate backbone structures and the required ArnT/C enzymes.

Figure 1.7. Lipid A modification in *Salmonella*. Enzymes are in parenthesis and described in the text.

Figure 1.8. The general LPS structure of *Rhizobium leguminosarum* bv. *viciae* 3841.

Figure 1.9. The general LPS structure from *Rhizobium etli* bv. *phaseolus* CE3/CFN42.

Figure 1.10. The point of LPS divergence between *E. coli* and *Rhizobium* from the conserved precursor molecule Kdo_2-lipid IV_A.

Figure 1.11. Proposed origins of the very long chain fatty acid 27-hydroxyl group. Panel A. Observed AcpXL-fatty acid elongation intermediates from cytosolic fractions (116). Panels B and C. Proposed origins of the 27-hydroxyl group as described in the text.

Figure 1.12. LPS assembly in *Rhizobium.* Adapted from Carlson et al (20).

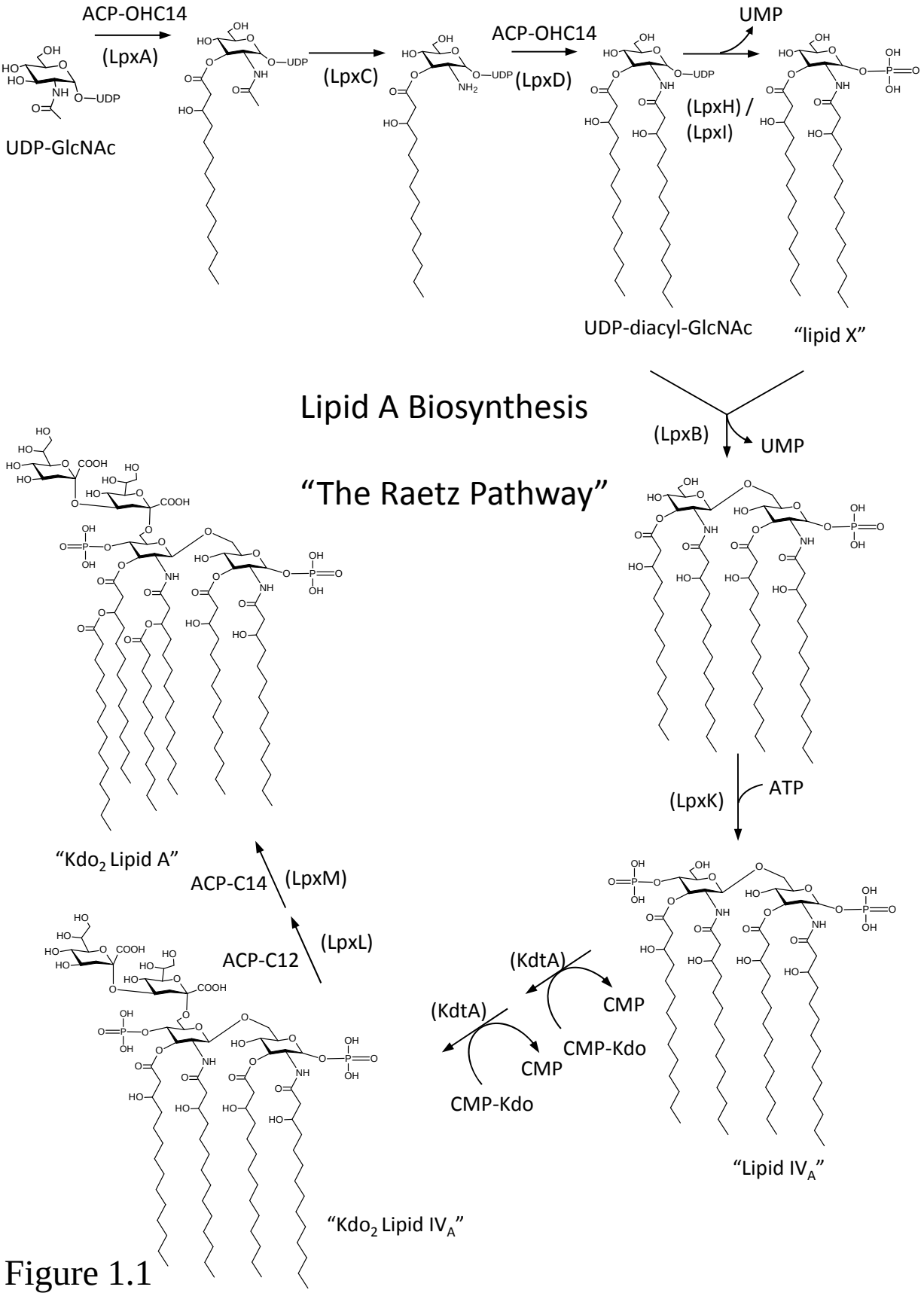

ACP-OHC14
(LpxA)
UDP-GlcNAc
(LpxC)
ACP-OHC14
(LpxD)
UMP
(LpxH) /
(LpxI)
UDP-diacyl-GlcNAc
"lipid X"
Lipid A Biosynthesis
"The Raetz Pathway"
(LpxB)
UMP
"Kdo2 Lipid A"
(LpxK)
ATP
ACP-C14
(LpxM)
ACP-C12
(LpxL)
(KdtA)
CMP
(KdtA)
CMP-Kdo
CMP
CMP-Kdo
"Lipid IVA"
"Kdo2 Lipid IVA"

Figure 1.1

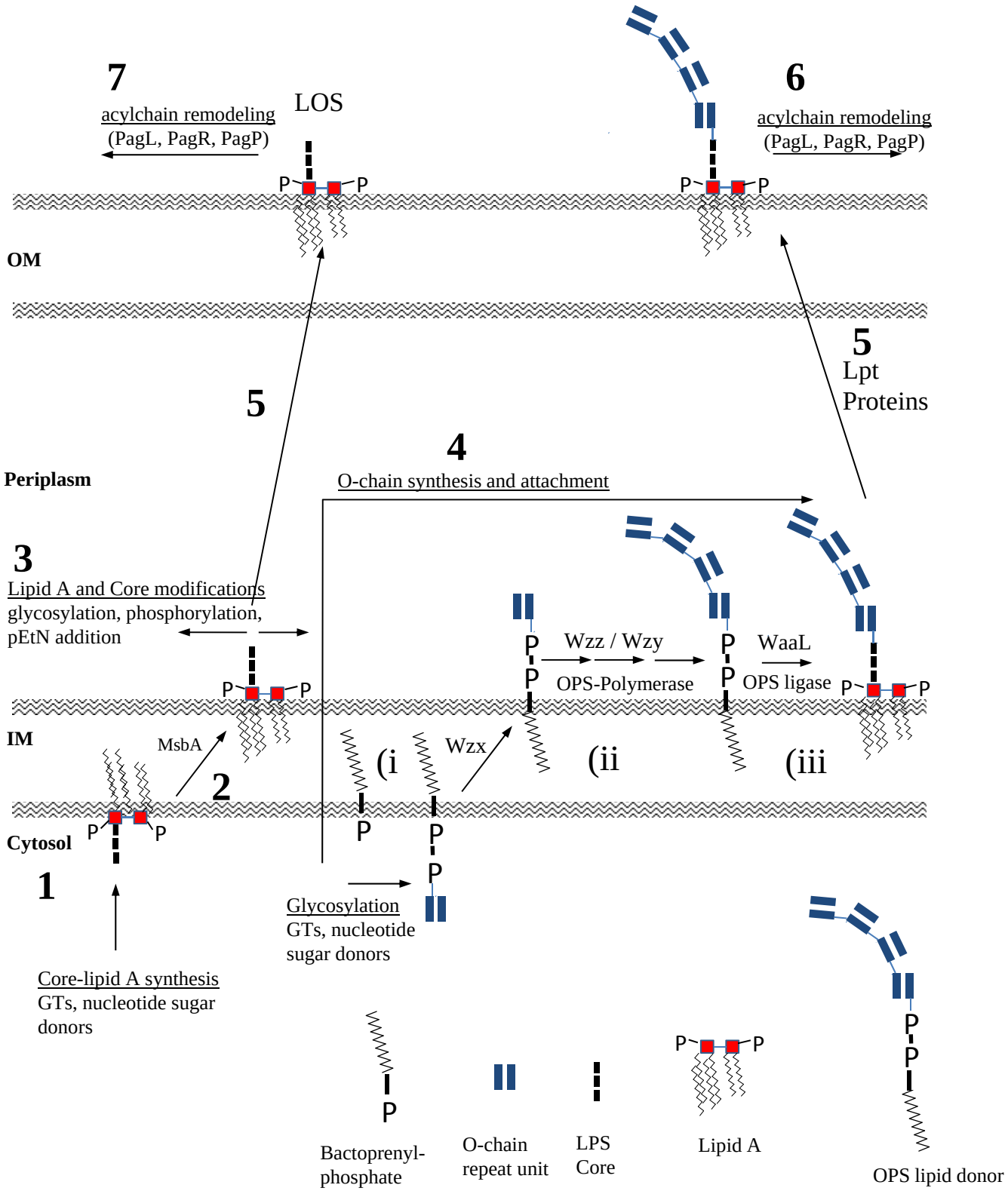

7
acylchain remodeling
(PagL, PagR, PagP)
LOS
6
acylchain remodeling
(PagL, PagR, PagP)
OM
5
Lpt
Proteins
5
4
O-chain synthesis and attachment
Periplasm
3
Lipid A and Core modifications
glycosylation, phosphorylation,
pEtN addition
Wzz / Wzy
OPS-Polymerase
WaaL
OPS ligase
IM
MsbA
2
(i
Wzx
(ii
(iii
Cytosol
1
Glycosylation
GTs, nucleotide
sugar donors
Core-lipid A synthesis
GTs, nucleotide sugar
donors
Bactoprenyl-
phosphate
O-chain
repeat unit
LPS
Core
Lipid A
OPS lipid donor

Figure 1.2

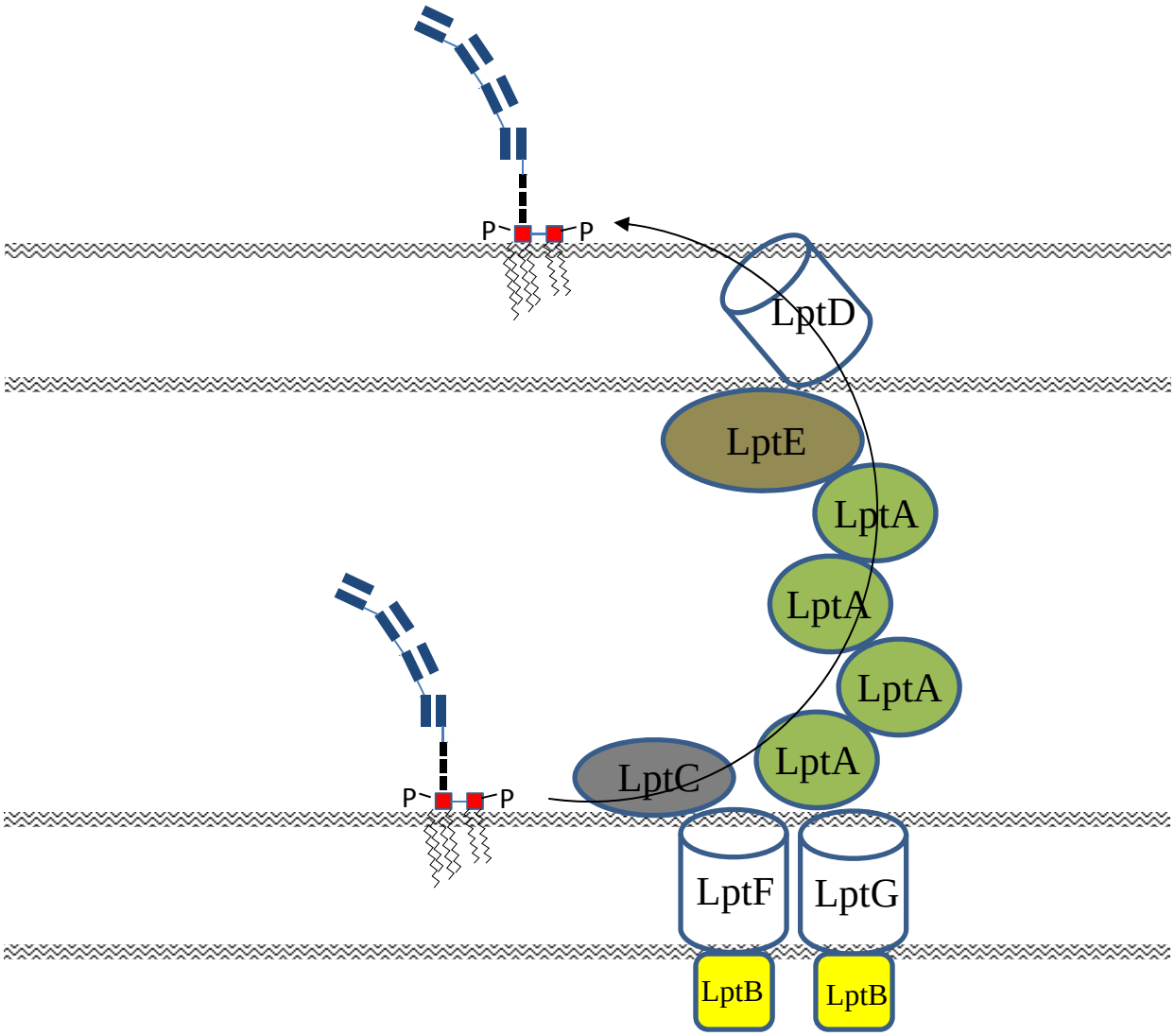


Figure 1.3

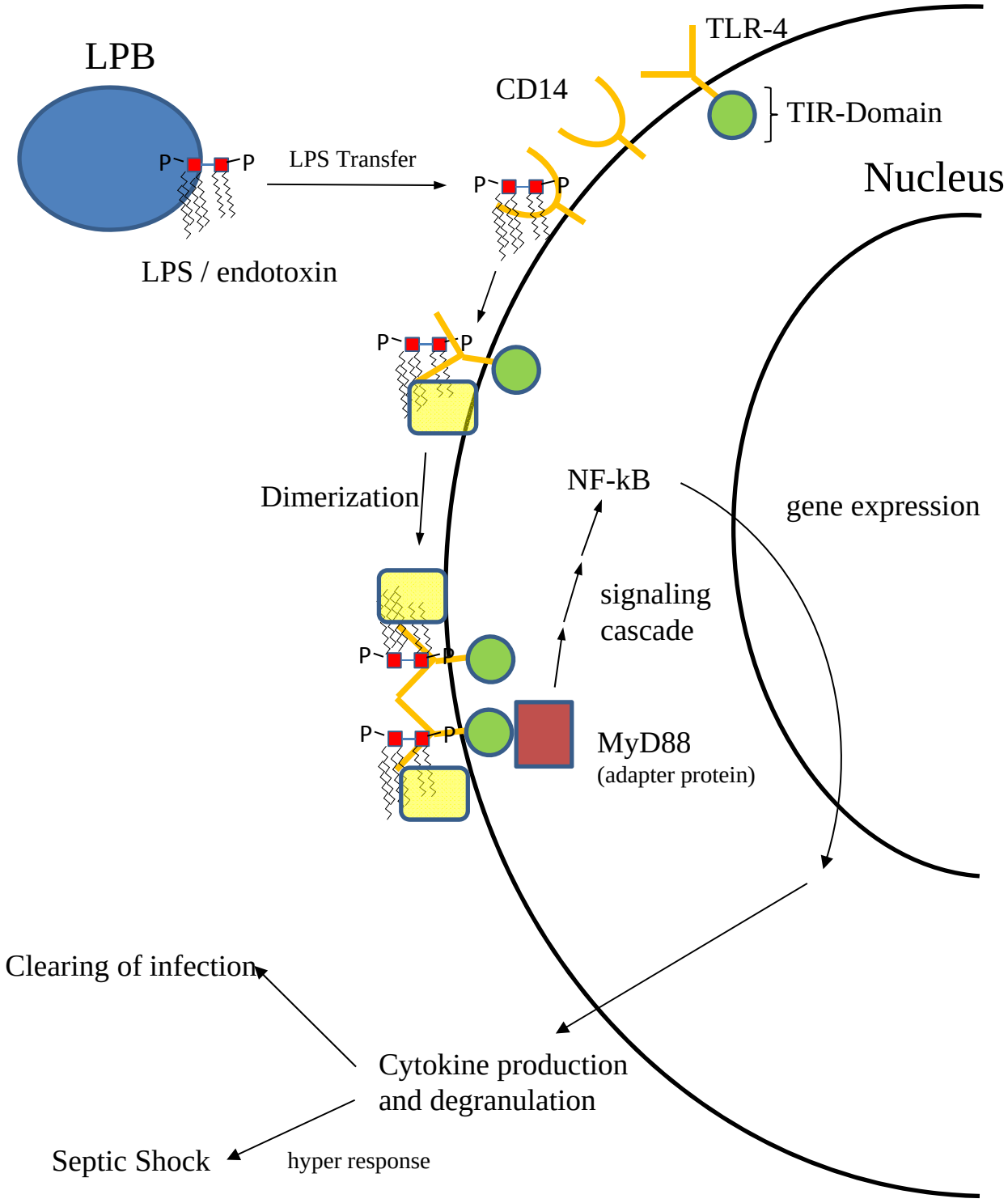

LPB
P
P
LPS Transfer
LPS / endotoxin
TLR-4
CD14
TIR-Domain
Nucleus
P
P
P
P
Dimerization
NF-kB
gene expression
signaling
cascade
P
P
P
P
MyD88
(adapter protein)
Clearing of infection
Cytokine production
and degranulation
Septic Shock
hyper response

Figure 1.4

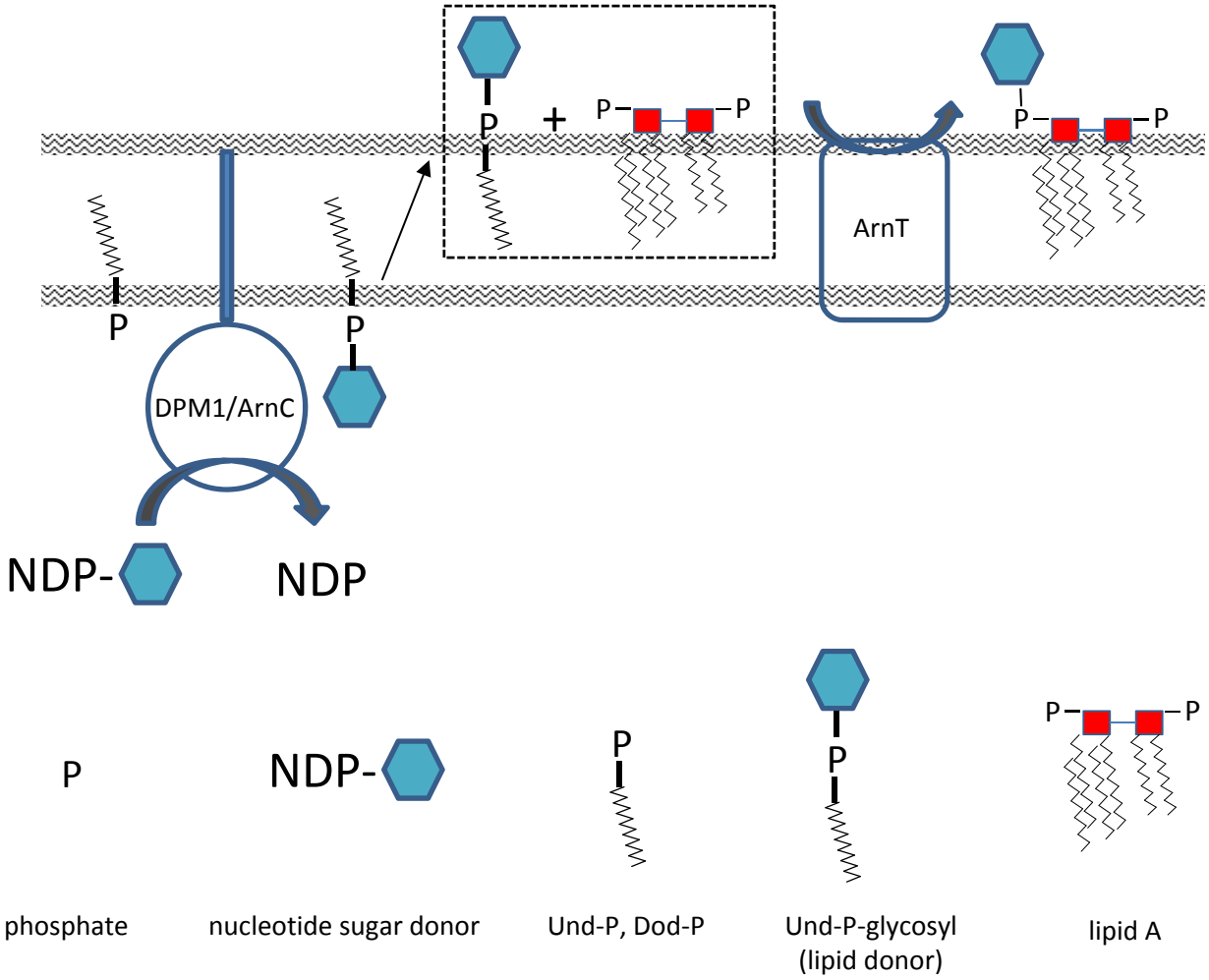
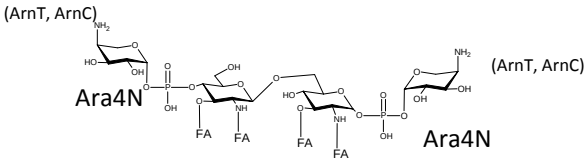
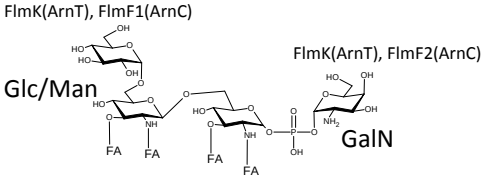

P
+
P
P
P
P
ArnT
P
P
DPM1/ArnC
NDP-
NDP
P
NDP-
P
P
P
P
phosphate
nucleotide sugar donor
Und-P, Dod-P
Und-P-glycosyl
(lipid donor)
lipid A

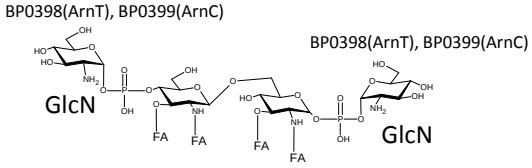
Figure 1.5



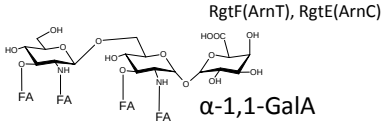
Escherichia, Shigella, Salmonella, Yersinia, Pseudomonas



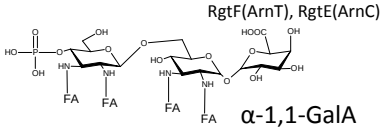
Francisella tulerensis, F. novicida



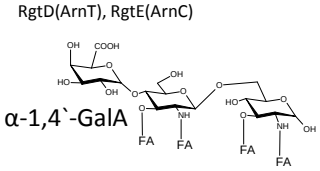
Bordetella pertussis, B. bronchiseptica, B. avium



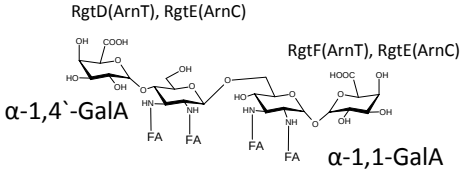
Bordetella pertussis, B. bronchiseptica, B. avium



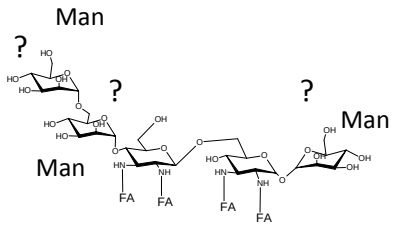
Azospirillum lipoferum



Rhizobium leguminosarum, R. etli



Aquifex aeolicus, Caulobacter crescentus



Bradyrhizobium elkanii, B. japonicum

Figure 1.6

***E. Coli* and *Salmonella* general lipid A structure**

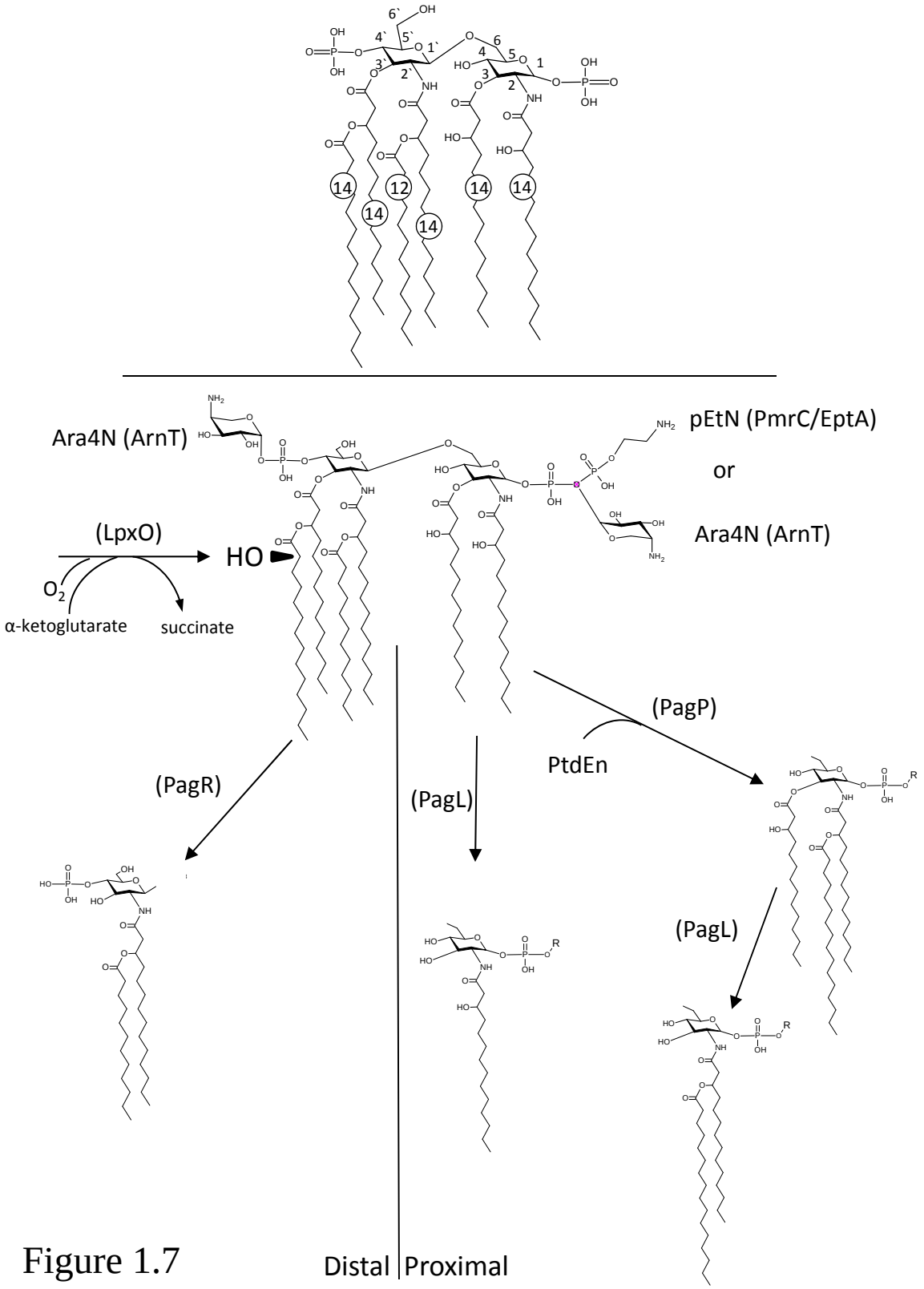


Figure 1.7

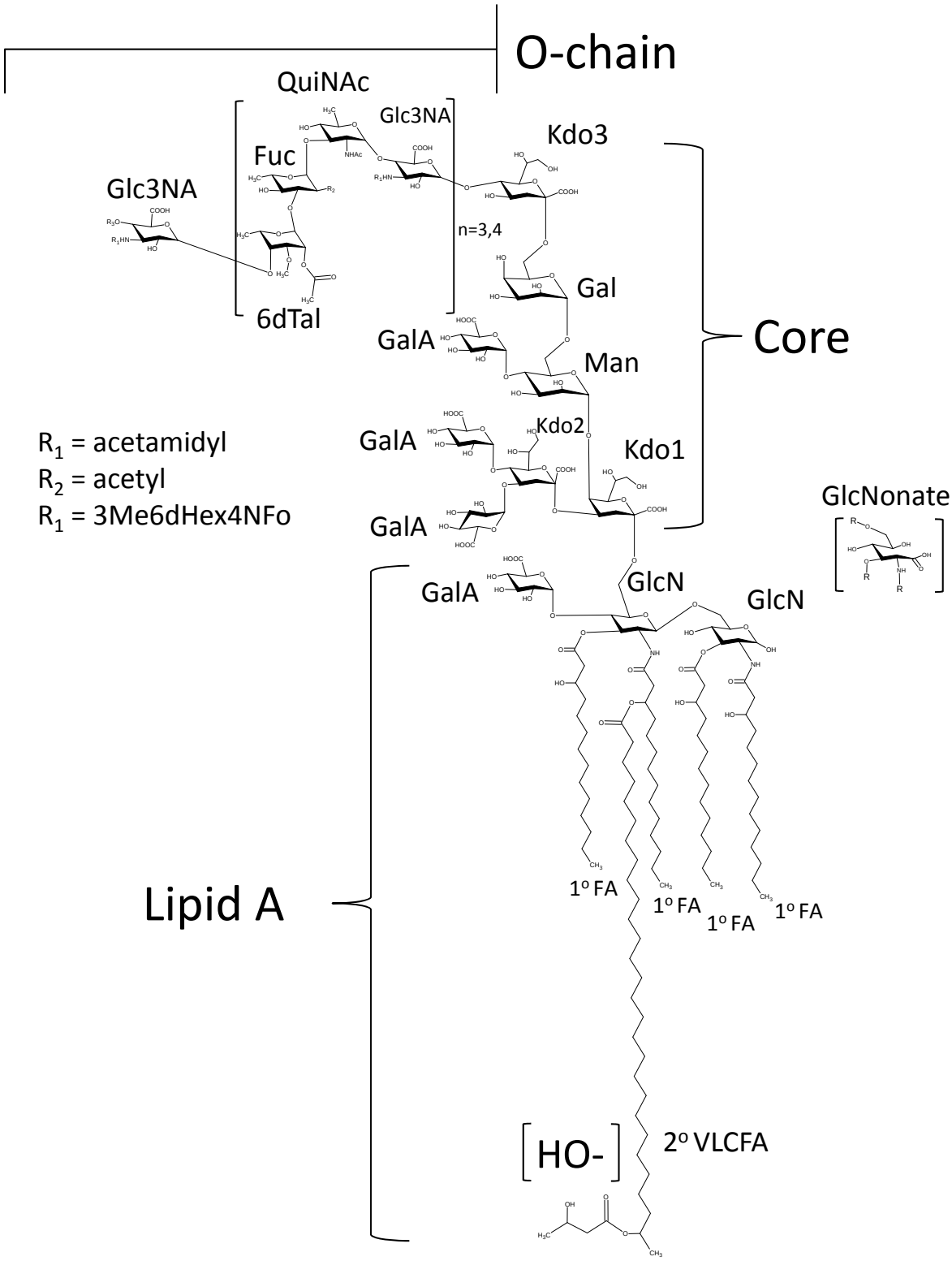

O-chain
QuiNAc
Glc3NA
Fuc
Glc3NA
Kdo3
n=3,4
Gal
6dTal
GalA
Man
Core
R1 = acetamidyl
R2 = acetyl
R1 = 3Me6dHex4NFo
GalA
Kdo2
Kdo1
GlcNonate
GalA
GlcN
GlcN
Lipid A
1º FA
1º FA
1º FA
1º FA
HO-
2º VLCFA

Figure 1.8

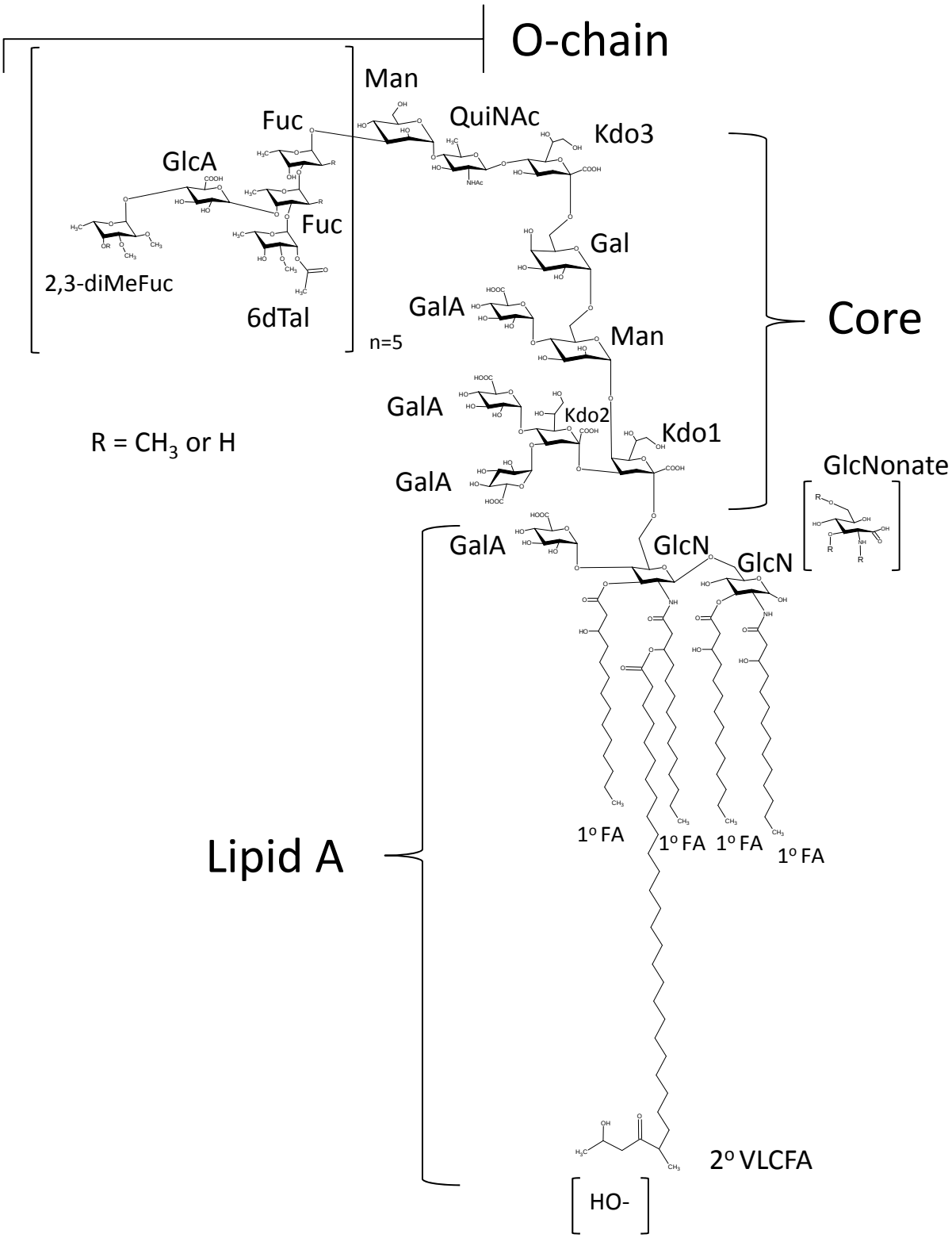

O-chain
Man
Fuc
QuiNAc
Kdo3
GlcA
Fuc
Gal
2,3-diMeFuc
6dTal
GalA
Man
n=5
Core
R = CH3 or H
GalA
Kdo2
Kdo1
GlcNonate
GalA
GlcN
GlcN
Lipid A
1º FA
1º FA
1º FA
1º FA
2º VLCFA
HO-

Figure 1.9

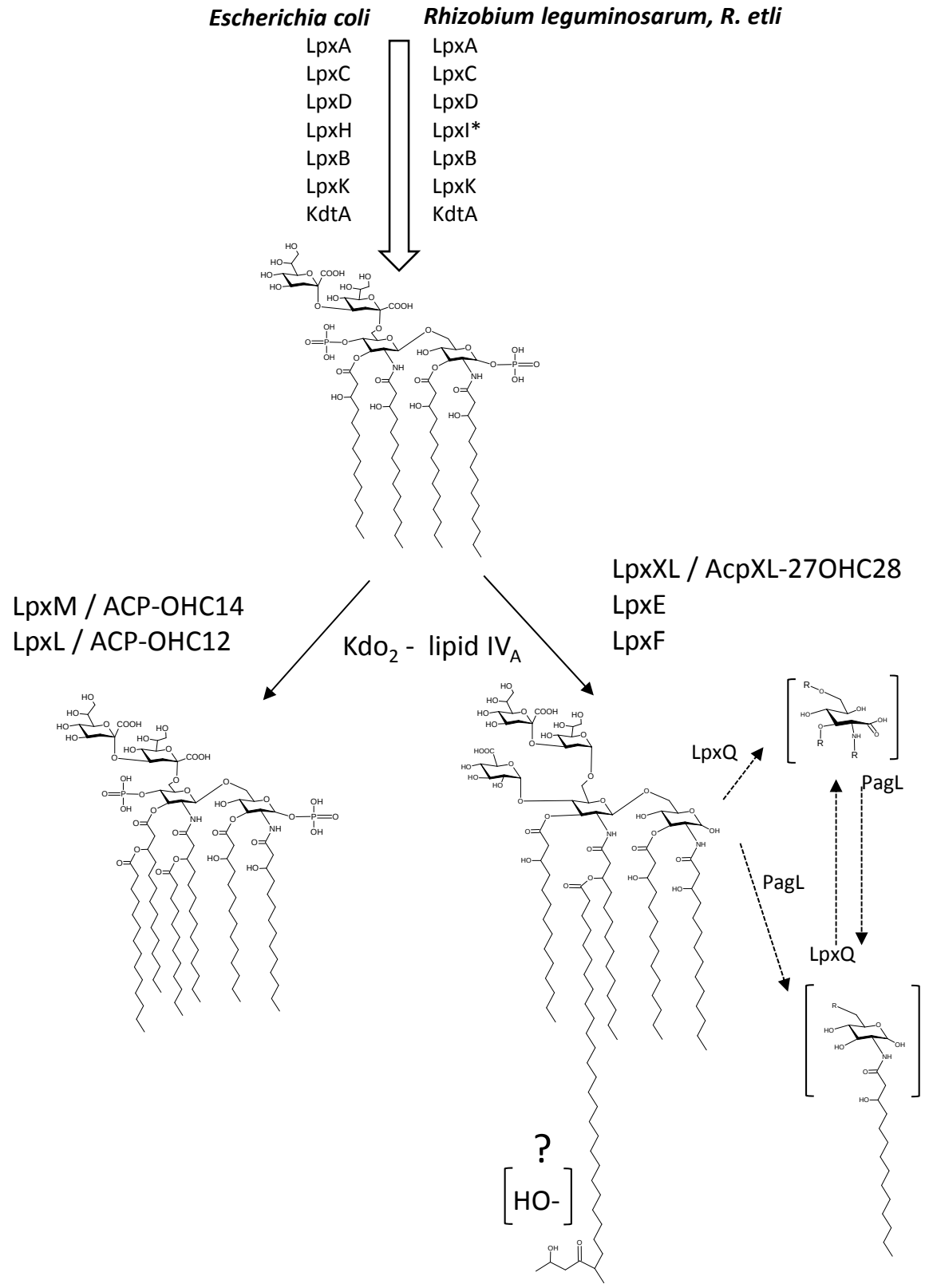

Escherichia coli
LpxA
LpxC
LpxD
LpxH
LpxB
LpxK
KdtA
Rhizobium leguminosarum, R. etli
LpxA
LpxC
LpxD
LpxI*
LpxB
LpxK
KdtA
LpxM / ACP-OHC14
LpxL / ACP-OHC12
Kdo2 - lipid IVA
LpxXL / AcpXL-27OHC28
LpxE
LpxF
LpxQ
PagL
PagL
LpxQ
?
HO-

Figure 1.10

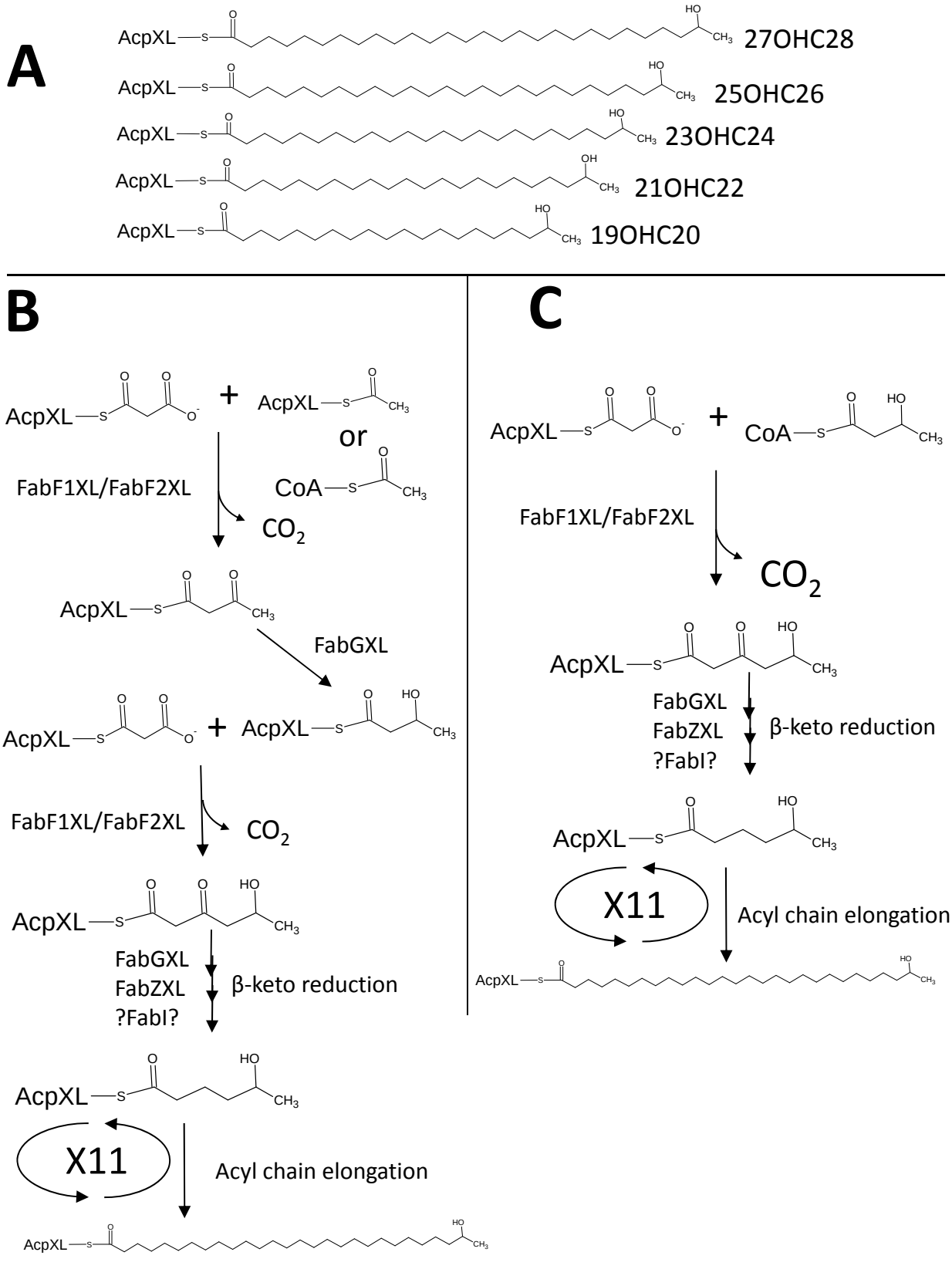

A
AcpXL
27OHC28
25OHC26
23OHC24
21OHC22
19OHC20
B
or
FabF1XL/FabF2XL
CoA
CO2
FabGXL
FabZXL
?FabI?
β-keto reduction
X11
Acyl chain elongation
C

Figure 1.11

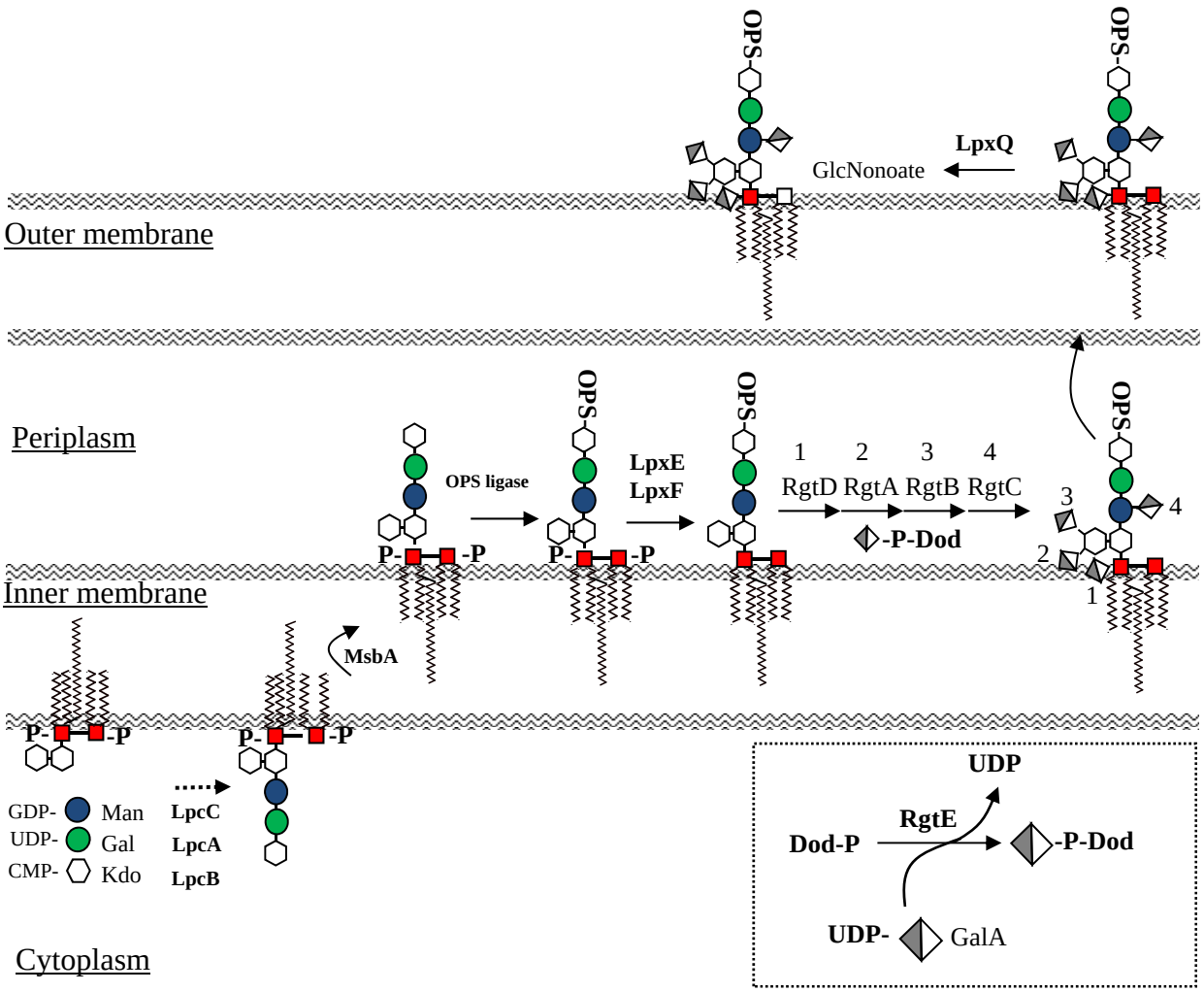

OPS-
LpxQ
GlcNonoate
Outer membrane
Periplasm
OPS ligase
LpxE
LpxF
1 2 3 4
RgtD RgtA RgtB RgtC
-P-Dod
Inner membrane
MsbA
P- -P
GDP- Man
UDP- Gal
CMP- Kdo
LpcC
LpcA
LpcB
Cytoplasm
UDP
RgtE
Dod-P
-P-Dod
UDP- GalA

Figure 1.12

CHAPTER 2

An *acpXL* mutant in *Rhizobium leguminosarum* bv. *phaseolus* lacks 27-hydroxyoctacosanoic acid in its lipid A and is developmentally delayed during symbiotic infection of the determinate nodulating host plant *Phaseolus vulgaris*

Dusty B. Brown, Yu-Chu Huang, L. Elmar Kannenberg, D. Janine Sherrier, and W. Russell Carlson. 2011. *J. Bacteriol.* **193**. 4766-4778.

Reprinted here with the permission of the publisher.

Abstract

Rhizobium leguminosarum is a gram negative bacterium that forms nitrogen fixing symbioses with compatible leguminous plants via intracellular invasion and establishes a persistent infection within host membrane derived subcellular compartments. Notably, an unusual very long chain fatty acid (VLCFA) is found in the lipid A of *R. leguminosarum* as well as in the lipid A of the medically relevant pathogens *Brucella abortus, Brucella melitensis, Bartonella henselae,* and *Legionellae pneumophila* which are also able to persist within intracellular host-derived membranes. These bacterial symbionts and pathogens contain a homologous gene region which is necessary for the synthesis and transfer of the VLCFA to the lipid A. Within this region lies a gene that encodes the specialized acyl carrier protein AcpXL on which the VLCFA is built. This study describes the biochemical and infection phenotypes of an *acpXL* mutant which lacks the VLCFA. The mutation was created in the strain *R. leguminosarum* 8002 bv. *phaseolus* which forms symbiosis with *Phaseolus vulgaris*, a determinate nodulating legume. Structural analysis using gas chromatography and mass spectrometry revealed that the mutant lipid A lacked the VLCFA. When compared to the parent strain, the mutant was more sensitive to the detergents deoxycholate and dodecyl sulfate as well the antimicrobial peptide polymyxin B suggesting a relative compromise to membrane stability. In addition, the mutant was more sensitive to higher salt concentrations. Passage through the plant restored salt tolerance. As observed by electron microscopy, the mutant was developmentally delayed during symbiotic infection of the host plant *Phaseolus vulgaris* and produced abnormal symbiosome structures.

Introduction

Rhizobium leguminosarum is a gram negative Alpha-proteobacterium capable of forming nitrogen fixing symbioses with plants found in the *Leguminosae* family. The symbiotic process involves the formation of root nodules and simultaneous invasion of the legume root by compatible symbiotic bacteria. The invading bacteria are eventually endocytosed by specialized root nodule tissue where they develop into nitrogen fixing bacteroids. Bacteroids persist within plant derived membranes (peribacteroid membrane) which together form organelle-like structures (symbiosomes) of which the primary function is to provide the plant with nitrogen through the conversion of atmospheric nitrogen to ammonia via the bacteroid-derived nitrogenase enzyme. For a recent review of the rhizobial infection process refer to Oldroyd et. al (26).

Depending on the legume host plant, two nodule types, indeterminate and determinate, are known and key fundamental differences exist between the two nodule types. Indeterminate systems are characterized as having an active meristem during the entire life cycle of the nodule. Conversely, meristematic cells from determinate root nodules stop dividing well before nodule senescence. Indeterminate nodules have an oblong shape and harbor terminally differentiated (inhibited in cell division) bacteroids which have characteristic enlarged and branched morphology. Recently discovered (34), the causative agent for the terminal differentiation of the symbiotic bacteria *Sinorhizobium meliloti* has been attributed to plant derived cationic peptides (NCRs, nodule derived cysteine rich peptides). Determinate nodules are typically round shaped and harbor bacteroids with long rod shaped morphology that do not terminally differentiate. In both systems, bacteroids are significantly larger than their non-differentiated free living bacterial counterparts. The fully developed symbiosomes of indeterminate nodules typically contain one

bacteroid per peribacteroid membrane, in contrast to the determinate symbiosomes, that may contain multiple bacteroids per peribacteroid membrane. The type of nodule that forms, indeterminate or determinate, is controlled by the host plant.

Secreted and surface associated bacterial polysaccharides are crucial to the formation of complete symbiosis in both indeterminate and determinate nodule systems (19, 25). One such molecule, lipopolysacharide (LPS) is a major surface glycoconjugate located in the outer membrane and has three major parts; lipid A, core oligosaccharide, and O-chain polysaccharide. For a current and detailed review of *Rhizobium* LPS and its role in *Rhizobium*-legume symbiosis see Carlson et. al (11). *Rhizobium* lipid A, the topic of this study, anchors the LPS molecule to the outer membrane and consists of a di-glucosamine (GlcN) backbone to which primary *O*- and N-β-hydroxy fatty acyl residues are attached (Figure 2.1). The distal GlcN contains a branched α-D-galacturonosyl residue (GalA) at the 4` position while the proximal GlcN may be oxidized to 2-aminogluconate (GlcNonate) in the outer membrane by the monoxygenase LpxQ (28, 29) providing additional negative charge to the molecule. The unique very long chain fatty acid (VLCFA) 27-hydroxy octacosanoic acid is ester linked to the N-β-hydroxy fatty acyl residue on the distal GlcN. Basu et. al have shown that a number of *Rhizobium* species have 3-*O*-deacylase activity in their membranes (2) which is able to cleave the 3-*O* linked β-OH myristate from the proximal GlcN of synthetic lipidIV_A *in vitro*. The acylase activity is likely due to a PagL-like enzyme described in a number of pathogenic strains (30). D'Haeze et al. (12) have presented structural evidence *in vivo* confirming the presence of a 3-*O*-deacylated lipid A species in *Rhizobium etli.*

The lipid As of most symbionts, as well as other medically and agriculturally important intracellular animal and plant pathogens, contain VLCFA and the genes necessary for VLCFA

synthesis (6). Due to these observations, it has been speculated that the VLCFA may be important for pathogenesis and symbiosis by these organisms. The genes which are involved in the synthesis of the *Rhizobium* VLCFA lie in a six gene cluster on the chromosome and are widely conserved among bacteria that synthesize VLCFA (6, 37). The genes within this cluster are *acpXL, fabZXL, fabF2XL, fabF1XL, fabIXL/fabGXL,* and *lpxXL* in order of the transcriptional direction (Figure 2.1). AcpXL is a specialized acyl carrier protein (8, 33, 38) on which the VLCFA is built and the *lpxXL* (1, 14) gene product is required for the transfer of VLCFA from the AcpXL-VLCFA conjugate to the lipid A (Figure 2.1). The genes involved in VLCFA synthesis are specific to VLCFA synthesis. *Rhizobium* strains that contain single mutations in *acpXL*, *fabF1XL, and fabF2XL* genes (33, 35, 38) do not synthesize VLCFA and produce lipid A which lack VLCFA. These results show that the genes required for general fatty acid synthesis (40) do not compensate for the loss of genes required for VLCFA *in vivo*.

Currently, mutagenesis studies concerning the VLCFA in legume symbionts have been devoted to systems which form indeterminate nodules (17, 36-38). In these examples, the loss of VLCFA resulted in the eventual formation of nitrogen fixing symbiosis under laboratory conditions. However, the developmental process was disrupted leading to delayed infection and aberrant symbiosome formation. The requirements for determinate and indeterminate systems can sometimes vary. For example, in the indeterminate systems studied so far, exopolysaccharide production is required for normal symbiosis but not for determinate systems studied (19). Therefore, it is important to investigate each system individually. This study focuses on the *acpXL* gene of *R. leguminosarum* bv. *phaseoli* 8002, a symbiont of *Phaseolus vulgaris* (common bean). *Phaseolus* forms determinate nodules and, as of yet, there has not been much attention devoted to understanding the role of VLCFA in this type of nodule system. In this paper, we

describe the preparation of an *acpXL* mutant and the effect of this mutation on LPS structure, membrane stability, and symbiosis.

Experimental Procedures

Bacterial strains, plasmids, and growth conditions

For a list of bacterial strains and plasmids used in this work, see Table 2.1. *E. coli* strains were grown at 37^{o}C on Luria-Bertani (LB) media. *Rhizobium leguminosarum* strains were grown on tryptone-yeast extract (TY) with 10mM $CaCl_2$ or minimal (Y) medium containing 5% sucrose at 30^{o}C as described previously (5). Where mentioned, media were supplemented with antibiotics in the following concentrations: gentamicin (Gm, 30μg ml^{-1}) and tetracycline (Tc, 10μg ml^{-1} [*E. coli*] and 15μg ml^{-1} [R. *leguminosarum*]).

Construction of the acpXL::aacC1 mutant and complement

Plasmid pVV1, described previously (38), harbors the *acpXL* (327bp) gene from *Rhizobium leguminosarum* within a 3.9kb stretch of cloned DNA. A single *EcoRI* site is located near the center of *acpXL* and was used as a cloning site for insertion mutagenesis. The gentamicin cassette containing the gentamicin resistant gene *aacC1* located in the plasmid pMS255 (3) contains flanking *EcoRI* sites and was inserted into *acpXL* in order to disrupt gene function and positively select for mutants. Enzymes used for restriction digestion and ligation were purchased from Promega, Inc. Mini preps purchased from Qiagen were used for plasmid preparation, enzyme clean up and agarose gel extraction.

Plasmid pMS255 was digested with restriction enzyme *EcoRI* and was subjected to agarose gel electrophoresis. The 1087bp fragment containing the gentamicin cassette was subcloned into plasmid pVV1 at the *EcoRI* to create the plasmid pODB1. Plasmid pODB1 was

transformed into chemically competent *E. coli* DH5α and subsequently isolated from these cells. The region containing the disrupted *acpXL* was subcloned into the suicide vector pEX18-Tc (18) containing the lethal *sacB* gene. For that, the plasmid pODB1 was digested with *XhoI.* A 4.7kb fragment containing the disrupted *acpXL* was ligated into the suicide plasmid pEX18-Tc previously digested with *SalI* in order to create the plasmid pODB3 that was subsequently transformed into *E. coli*.

A tri-parental mating was conducted in order to move pODB3 into *Rhizobium leguminosarum* 8002 via conjugal transfer. The donor strain harboring the plasmid pODB3, the helper strain carrying the plasmid pRK2013 (27) and the recipient strain *R. leguminosarum* 8002 bv. *phaseoli* were cultured overnight in 100ml media. The *E. coli* strains were grown in Luria-Bertani (LB) at 37^{o}C and the *R. leguminosarum* 8002 was grown in Tryptone Yeast (TY) media at 30^{o}C. Cultured donor (2ml) and helper were pelleted and each re-suspended in 4ml TY. Then, 4ml of recipient, 0.5ml donor and 0.5ml helper were mixed and pelleted. The supernatant was discarded until there was just enough left to make a slurry. The entire slurry (approximately 200µl) was spotted on a nitrocellulose membrane (47mm, 0.45µM pore size) laid on a TY agar plate and incubated at 30^{o}C overnight. The resulting paste was washed off the membrane with 4ml sterile de-ionized water. The mating mix was spread plated onto minimal Y agar plates containing 30µg/ml Gm and 5% sucrose. Colonies were picked and streaked on TY/Gm.

Candidate mutants were screened by PCR using primers that recognize either end of the *acpXL* gene (CGATCTTGGCGCAGAGGTTC and CATGAGGTACAGCGCGTGAC). The undisrupted *acpXL* PCR product is about 300bp while the *acpXL::accC1* product is about 1.4kb. An *acpXL* mutant was obtained and stored in 50% glycerol at -80^{o}C.

The mutant was complemented with the *acpXL* gene in order to validate observable differences between the mutant and the parent strain. Plasmid pVV5 (38), described previously, is a pBBRMCS-3 (22) derivative harboring a 2.5kb region that contains the parental *acpXL* gene. Plasmid pVV5 was moved into the mutant via tri-parental mating as described above. The complemented strain was selected for on TY/Tc media plates and confirmed by PCR as described above.

Detergent and sodium chloride susceptibility assays

Bacteria were grown on increasing concentrations of the detergents sodium deoxycholate or sodium dodecyl sulfate. Bacteria were initially grown on TY agar plates. One colony of bacteria from a four day culture on TY media agar plates was streaked on minimal Y agar plates containing 0mM, 1mM and 2mM concentration of sodium deoxycholate. In addition, bacteria were streaked on TY agar plates containing various amounts of sodium dodecyl sulfate (0, 100, and 200μg/ml). In order to assess increasing resistance to high salt concentrations, bacteria were grown on TY agar plates containing 0%, 0.25%, 0.5%, and 0.75% sodium chloride. In all instances bacteria were allowed to grow 4 days at 30^{o}C before visual observation of growth. Each experiment was done in duplicate.

Polymyxin B sensitivity assay

Bacteria were grown in liquid TY media to an O.D. of about 0.8 at A_{600}. Each culture was diluted to an O.D. of about 0.1 using sterile TY liquid media and bacteria were evenly spread onto TY agar plates using sterile swabs and allowed to completely dry. A polymyxin B Etest®

strip (Biomèrieux) was placed in the center of the dry plates and the plates were incubated four days at 30^{o}C before reading. Each experiment was done in duplicate.

Isolation of Nodule Bacteria and Bacteroids

Plants were inoculated with bacteria as described below. At 28 days post inoculation, individual nodules were harvested. Nodules were sterilized in 6% hypochlorite for 30 seconds and diluted 5 times in sterile deionized water for 1 minute. To isolate nodule bacteria, the nodules were squashed to release the inner contents and streaked onto TY media agar plates containing appropriate antibiotics.

Bacteroids were isolated using a modified step wise sucrose gradient method as described previously (12). Briefly, nodules from two batches of 60 plants each, one batch inoculated with wild type and the other with the *R. leguminosarum acpXL* mutant were picked, placed into 10ml grinding buffer (50mMTris, 10mM dithiothreitol, 5% (w/v) polyvinyl polypyrolidone, 5mM *p*-aminobenzamidine) and crushed into a slurry using a mortar and pestle. The slurry was filtered through a plunger syringe fitted with glass wool. The filtrate was centrifuged (4000 x g), the supernatant was discarded and the pellet containing bacteroids was re-suspended in 10ml 50mM Tris/0.5M sucrose/10 mM dithiothreitol buffer (pH7.4). The solution was then passed through a step wise sucrose gradient (from bottom to top: 57%, 52%, 50%, 45% sucrose in 50mM Tris, pH7.4) at 100,000 x g for 1hr as described previously (12). Bacteroids were isolated from band 4 (52% sucrose layer) using a Pasteur pipette and washed twice with 50mM Tris Buffer, pH7.4. The supernatant was discarded and the pellet was stored at -20^{o}C.

Plant Growth and Inoculation for Acetylene Reduction Assay

To test whether or not the *acpXL* mutant bacteroids could fix nitrogen, an acetylene reduction assay was performed on parent and mutant infected bean nodules. Bean seeds, black turtle variety, were surface sterilized by swirling in 70% ethanol for 1 minute followed by 3 washes with sterile water for 1 minute each time. Then, the seeds were swirled in 6% sodium hypochlorite for 4 minutes and extensively washed with sterile water. The seeds were transferred to sterilized Erlenmeyer flasks (1 seed per flask) containing a foam top and hydroponic beads immersed in defined, nitrogen free liquid plant media (Fahreus media (13)). The Erlenmeyer flasks were surrounded with black plastic bags and placed in a growth chamber in order to regulate temperature and light. The temperature was maintained at 24^{o}C with 14hrs of light and 20^{o}C with 10hrs of darkness. When tall enough, the plant stems were carefully pulled between the top of the flask and the foam top exposing the leaves to the outside while leaving the root system in the media. At this point (7 days post germination), the plants were inoculated with 1ml of late-log/early–stationary phase liquid cultures (TY media) of parent and mutant strains. The inoculated plants were allowed to incubate for 30 days.

At 30 days post inoculation, the roots were harvested by cutting just below the seed coat while leaving the root system intact. The roots were placed in a 100ml glass bottle and capped with a septum. Using a plunger syringe, 5ml of acetylene gas was transferred from a liter bottle flushed with acetylene to the bottles containing the plant roots. The roots were allowed to incubate for 30 minutes at room temperature. The evolution of ethylene was measured by extracting 1ml of gas from the incubated bottles using a 1ml plunger syringe and injecting the gas into a GC analyzer (isothermal at 70^{o}C, with a 25m by 0.53mm, 10-μm Al_2O_2-KCL column [Varian, Inc.]).

Plant growth and Inoculation for Transmission Electron Microscopy of Nodules

Beans (*Phaseolus vulgarus* cv. "Tendergreen") were soaked for 3 min in 6% hypochlorite were rinsed 10 times with sterile water, were sterilized 5 min in sterilization solution (30% H_2O_2 in 100% ethanol), and were rinsed 6 times with sterile water. After final wash, the seeds were imbibed in sterile water with agitation for 15 min. The seeds were then transferred into the sterile glass Petri plates and were placed in a dark cabinet at room temperature. After 5 to 7 days, each germinating seed was placed in a 500-mL flask which contained 400-mL volume of hydroton and 300 mL of sterile Fahraeus N-free medium (13). The flasks were capped with sterile foam stoppers and were placed in the dark cabinet again. After 5 days, the hypocotyls were gently pulled out of the mouth of the flask in sterile conditions, and the foam plug was placed back to the mouth of flask to support the shoot. Flasks were wrapped in black plastic and the plants were grown a Conviron growth chamber at 21^{o}C for 2 days before inoculation.

For the inoculation of plants, the wild type (*R. leguminosarum* bv *phaseoli* 8002) and mutant (*R. leguminosarum* bv *phaseoli* 8002 *acpXL*) bacteria were grown at 28 ^{o}C in TY medium. When the optical density of the culture reached to 1.1 at 600 nm ($1x10^8$ cells/mL) the bacterial cells were collected by centrifugation. The cell pellets were resuspended in sterile water. Each plant was inoculated with 1 mL of diluted culture containing approximately 5 $X10^6$ cells.

Transmission Electron Microscopy

Nodules were collected at 14, 21 and 28 dpi and were preserved directly into fixative containing 4% formaldehyde, 1% glutaraldehyde (Electron Microscopy Sciences) in 100 mM

PIPES buffer (Sigma) at pH 7.0. Nodules were cut in thirds longitudinally and the central sections were used for fixation and infiltration. Tissues were fixed overnight at 4^{o}C with rotation, rinsed three times with water and then postfixed with 1% OsO_4 for 2h. Tissues were dehydrated in a graded dilution series of acetone and then infiltrated in a graded dilution series of Epon 812-Araldite resin (mixture of Embed 812, Araldite 502, DDSA and BDMA) (Electron Microscopy Sciences). Tissues were infiltrated in Epon 812-Araldite resin overnight and were heat cured at 65^{o}C for 48 h (24).

For nodule tissues, 0.5 μm sections were collected on coated slides, stained with methylene blue, and documented with Zeiss Axioscope 2 microscope before collecting thin sections. 70-nm-thick sections were then collected onto hexagonal gold grids and were counterstained with 1% uranyl acetate for 5 min and with lead citrate for 5 min. Samples were visualized and imaged on Zeiss Libra 120 TEM (Carl Zeiss, Inc., Germany).

LPS Analysis

Lipopolysaccharide was extracted from cultured bacteria using the hot water/phenol extraction procedure as previously described (39). Briefly, one liter of cultured bacteria (liquid TY, overnight at 30^{o}C) was pelleted and washed 1 time with 0.9% NaCl followed by 3 washes in deionized H_2O. The cells were re-suspended in de-ionized H_2O to a final volume of 20ml. After freeze/thawing 3 times, the sample was subjected to hot water/phenol extraction (39). In order to remove phenol from the sample, the water phase and phenol phase were each transferred to 12,000kDa MWCO dialyses bags and dialyzed against 10L de-ionized H_2O for 5 days, each day exchanging the H_2O. The samples were then freeze dried leaving a white fluffy material enriched in LPS.

LPS from purified bacteroids were extracted using a hot water/phenol micro-extraction. The bacteroid pellet was suspended in 1ml de-ionized H_2O and placed in a 10ml glass culture tube fitted with a Teflon cap. 1ml of 90% liquid phenol was added and the LPS was extracted and dialyzed as described above.

Glycosyl and fatty acid compositions were determined by the preparation and combined gas chromatography-mass spectrometry analysis of trimethylsilyl methylglycosides, with N-acetylation, and fatty acid methyl esters as previously described (7, 41).

Lipid A was partitioned by the standard mild acid hydrolysis procedure (32). Approximately 5mg of extracted LPS was dissolved in 1% glacial acetic acid and incubated at 100^{o}C for 1hr. A white precipitate (lipid A) formed and was separated by centrifugation. The supernatant was washed 3 times with chloroform:methanol:water, 2:2:1.8 (v/v/v) respectively. The chloroform layers were added back to the precipitate and subsequently washed 3 times in de-ionized H_2O. The chloroform layer, containing lipid A, was evaporated under a stream of air.

Mass Spectrometric Analysis

Matrix assisted laser desorption time of flight mass spectrometric (MALDI-TOF/MS) analysis was performed by dissolving lipid A in 50µl chloroform:methanol; 3:1 (v/v). The lipid A solution was then mixed 1:1 (v/v) with 0.5 M 2,4,6-trihydroxyaceto-phenone THAP matrix; 1µl was spotted on a MALDI plate. Spectra were acquired using a MALDI-TOF analyzer (Applied Biosystems, Voyager-DE) in the positive and negative-ion linear mode operating at a 20kV extraction voltage.

DOC-PAGE

Gel electrophoresis in the presence of deoxycholic acid (DOC-PAGE) was carried out as previously described (31). Extracted LPS (2μg) was loaded on a polyacrylamide gel. Gels were stained by the silver stain method previously described (21).

Results

Preparation and isolation of the R. leguminosarum bv. *phaseoli 8002 acpXL::aacC1 mutant*

To understand the biological function of the lipid A derived very long chain fatty acid in a *Rhizobium*-determinant host symbiosis, an *acpXL* mutant was engineered via insertion mutagenesis using the type strain *R. leguminosarum* bv. *phaseoli* 8002. This strain has been used extensively as a model organism in the current field of legume symbiosis research but its genome has not been sequenced. However, it is closely related to the sequenced strain *R. leguminosarum* 3841 bv. *viciae*. Therefore, the plasmid (pVV1) containing the *acpXL* gene from strain 3841 (38) was modified by inserting the gentamicin cassette *aacC1* (3) near the center of the *acpXL* gene in order to disrupt gene function and subsequently introduced into strain 8002 as described in the Experimental Procedures section. Mutants were screened by PCR using primers that recognize either side of the *acpXL* coding region. The wild type PCR product contained the appropriate theoretical size of 300bp. A putative mutant was obtained and contained a PCR product with the appropriate theoretical size of 1.4kb. The mutation was confirmed by restriction digest of the mutant PCR product and the mutant was given the strain number ODB30 and will be referred to as such throughout this study. In order to validate the *acpXL* mutant, ODB30 was

complemented with plasmid pVV5 which harbors the *acpXL* coding region and native promoter. The complemented strain contains two PCR products which represent the inactivated *acpXL* gene from the chromosome (1.4kb) and the normal *acpXL* gene from plasmid pVV5 (300bp).

Structural Analysis of the acpXL Mutant LPS

To determine the affect of *AcpXL* on LPS structure, LPS was extracted from free-living cultured bacteria and nodule derived bacteroids using the hot water / phenol extraction method from parent and mutant bacterial strains (39). Water layer material was collected. In order to assess whether or not mutant ODB30 LPS contained VLCFA, the extracted LPS material was subjected to composition and mass spectroscopy analysis. There are two major LPS species in strain 8002 as observed on DOC-PAGE (Figure 2.2, Lane 1). The high molecular weight bands (LPS-I) represents LPS which contains O-chain polysaccharide (i.e. contains lipid A, Core, and O-chain) while the lower molecular weight band (LPS-II) represents LPS that lacks O-chain (i.e. only contains lipid A and Core). Mutant strain ODB30 also contains two major bands (Figure 2.2, Lane 2). However, a mobility shift is observed between the parent and mutant strains (mutant LPS migrates quicker) and suggests that mutant LPS has a lower molecular weight as compared to the parent strain.

The composition of total LPS and partitioned lipid A from free-living bacteria and bacteroids was determined by gas chromatography (7, 41) in tandem with mass spectrometry (GC/MS). The O-chain structural components of strain 8002 have not been reported. The total LPS composition of strains 8002 and ODB30 revealed the presence of rhamnose (Rha), fucose (Fuc), and mannose (Man) as the major glycosyl component along with the other residues, galacturonic acid (GalA) and Kdo, typically found in the core and lipid A (Table 2.3). Glucose

(Glc) was also present in the LPS preparations, however, this is a likely due to some contaminating EPS or cyclic glucans. Therefore, the O-chain likely contains Rha, Fuc, and Man residues in some sequential combination. A related strain, *R. leguminosarum* 128C53, contains similar amounts of Rha, Fuc, and Man and maintains these residues in the O-chain (10). Other than the level of Glc,

there was no observable difference in the amounts of the predicted O-chain residues between mutant and parent strain suggesting that the disruption to the *acpXL* had no observable affect on the O-chain portion of the LPS. Extensive analysis of O-chain was not conducted as this study focused on the lipid A.

Lipid A was separated from LPS by the standard mild acid hydrolysis (MAH) procedure (32) in order to obtain clear composition and detailed structural information concerning lipid A. As expected, the VLCFA (27OHC28) was present in the lipid A from both free-living and bacteroids of the parent strain (Table 2.3). The VLCFA was not detected in free-living or bacteroid lipid A of mutant ODB30 (Fig 4, panels C and D). The amount of GlcNAc was significantly higher in 8002 bacteroid lipid A when compared to 8002 free-living lipid A (Table 2.3). A similar increase in GlcNAc occurs in mutant ODB30 bacteroid lipid A when compared to free-living lipid A. The 2-aminogluconate (GlcNonate) residue is also found in the carbohydrate backbone of lipid A. The TMS methyl glycoside of this residue has a very low response factor on GC/MS analysis making it difficult to discern quantitative differences between samples. However, the increase in GlcNAc from bacteroid lipid A as compared to free-living lipid A is likely due to a decrease in the presence of GlcNonate in bacteroid lipid A (i.e. there is less GlcN being oxidized to GlcNonate in bacteroids). The amount of palmitic acid (C16:0) was significantly increased in mutant free-living and bacteroid lipid A as compared to the parent free-

living and bacteroid lipid A. The addition of palmitic acid to the mutant lipid A was confirmed by mass spectrometry (below).

Mass Spectroscopic Analysis of Mutant Lipid A

The loss of VLCFA from mutant lipid A was verified by mass spectrometry. The molecular weights of various lipid A from parent and mutant strains were obtained by MALDI-TOF MS analysis. The spectra are shown in Figure 2.3 and a detailed presentation of the observed ions with their proposed compositions is given in Table 2.4. Panels I and III of Figure 2.3 show spectra for the lipid A obtained from wild type strain 8002 free-living bacteria and bacteroids, respectively. Panels II and IV show the spectra of lipid A obtained from mutant strain ODB30 free-living bacteria and bacteroids respectively. There are two major ion groups observed in the spectra obtained from strain 8002 lipid A. The groups containing the higher molecular weight ions are due to structures A and B while the lower molecular weight group contains ions are due to structure C. Each ion group contains a complex mixture of ions due to the heterogeneous nature of *Rhizobium* lipid A such as, variation in the chain length of their β-hydroxy fatty acyl components. The major β-hydroxy fatty acids observed in the composition analysis are 3-OHC14:0, 3-OHC15:0, 3-OHC16:0, and 3-OHC18:0 (Table 2.3). A given lipid A preparation may contain structures with these hydroxy fatty acids present in a number of combinations resulting in an array of ions which have mass differences that vary by 14 (CH_2) and/or 28 (C_2H_4). In addition to differences in fatty acyl chain length, molecules that contain VLCFA may have a β-hydroxybutoxyl (BHB) or a hydroxyl group attached to the C27 position which gives a mass difference of 86. Lipid A molecules that contain GlcNonate are subject to the formation of a lactone, GlcNonolactone, during MAH and subsequent acid catalyzed β-

elimination resulting in the elimination of the 3-*O*-β-hydroxy fatty acid and the formation of 2,3-unsaturated GlcN(acyl)onolactone on the proximal end of the lipid A (Figure 2.3, structures C, F, and I).

The MS spectra obtained from mutant ODB30 contains molecular ions consistent with the loss of the VLCFA (structures D, E, and F). The loss of VLCFA corresponds to a predicted mass differences of 508.8 or 422.8 (-BHB). For example, structure B (2001) – structure D (1493) = 508. Major ions observed in panel II and IV (mutant bacteroid) have corresponding ions in panel I and III (wild type bacteroid) that have mass differences consistent with the loss of VLCFA (i.e. panel II ion 2003.6 – panel IV ion 1493.4 = 510.2). In bacteroid lipid A, there is a significant decrease in the intensity of ions corresponding to the tetraacylated, lipid A having a proximal GlcNono lactone (structures C, F, and I). For example, lipid A from parental bacteroids (panel III) gives a spectrum that contains relatively low intensity structure C ions when compared to the lipid A from a free-living parental culture (panel I). The mutant lipid A spectra contain a group of ions with m/z values consistent with the addition of palmitate to lipid A (structures G, H, and I, panels II and IV, Figure 2.3). The palmitate is predicted to be ester linked to the β-hydroxy fatty acyl group at the 2` position of the distal GlcN in place of the VLCFA.

In order to confirm the position of palmitate in structures G, H, and I (Figure 2.3) found in mutant lipid A, MALDI-TOF spectra were acquired in the positive linear mode to obtain b^+ ions (16). Secondary fragmentation of the lipid A occurs due to cleavage of the glycosidic bond between the proximal and distal GlcN residues and gives rise to b^+ ions that represent the distal end of lipid A. The distal end of the parent strain 8002 lipid A contains GalA, GlcN, β-OH fatty acids, and the VLCFA (OHC28). Panels A and B of Figure 2.4 show the b^+ ions observed for the parent strain 8002 free living and bacteroid lipid A respectively. Ions 1301 and 1215 (panel A

and B, Figure 2.4) differ by a mass of 86 due to the appearance of β-hydroxybutyrate at C27 of the VLCFA. There is also acyl chain variation giving rise to ions that differ by 28 mass units (1243 - 1215 = 28). The b^+ ions acquired from free living and bacteroid mutant lipid A are consistent with lipid A molecules in which the VLCFA has been replaced with palmitate (Panel C and D, Figure 2.4). The ion with mass 1031 (Panel C, Figure 2.4) agree with a lipid A structure composed of GalA, GlcN, two β-hydroxymyristate residues, and a palmitate.

Membrane Stability and Sodium Chloride Sensitivity of the acpXL Mutant

Differences in the membrane stability between mutant ODB30 and wild type 8002 strains were determined by monitoring growth on plates containing increasing concentrations of the detergents sodium deoxycholate (DOC), sodium dodecyl sulfate (SDS) and the antimicrobial peptide polymyxin B (Table 2.2). Compared to wild type, the mutant was more sensitive to DOC, SDS, and polymyxin B. The mutant strain complemented with the normal *acpXL* gene was able to recover DOC and polymyxin B resistance indicating that the disruption of *acpXL* was responsible for the increased sensitivity to DOC. The mutant was also more sensitive to increasing sodium chloride concentrations as compared to wild type (Table 2.2). Vedam et. al. previously described an *acpXL* mutant in *R. leguminosarum* 3841 (37, 38) that was more sensitive to NaCl as compared to the parent strain. Interestingly, those mutant bacteria isolated from pea nodules were restored to wild type NaCl resistance levels. Therefore, nodule bacteria were isolated from bean nodules infected with mutant ODB30 and tested for NaCl sensitivity. As determined by diagnostic PCR (Experimental Procedures), the isolated mutant bacteroids retained the insertion mutation. The ODB30 nodule derived isolates (NB) were restored to wild type NaCl resistance levels (Table 2.2) suggesting that the plant environment may have selected

a variant that compensated for the lack of NaCl tolerance (i.e. a suppressor mutation). Osmotolerance has been shown to be an important factor in the symbiotic process (4). When challenged with DOC detergent, the ODB30 NB resembled the free living ODB30 strain maintaining the same level of sensitivity.

Symbiotic Phenotype

When inoculated on the host plant *Phaseolus vulgaris*, the mutant ODB30 formed nitrogen fixing symbiosis as determined by the acetylene reduction assay. The level of nitrogen fixation by mutant inoculated plants, after 30 days, was similar to that of wild type inoculated plants. A previously described *acpXL* mutant of the pea symbiont *R. leguminosarum* 3841 (36-38) also formed nitrogen fixing symbiosis. However, the infection process was delayed and the bacteroids were enlarged and aberrantly shaped. Therefore, we analyzed the symbiotic development of the mutant via transmission electron microscopy (TEM).

Plants were inoculated and nodules were harvested and sectioned as described in the Experimental Procedures section. At 21 dpi, the wild-type infected cells were enlarged and contained large numbers of bacteroids that contained pronounced refractive poly beta-hydroxybutyrate (PHB) granules (an indication of bacteroid maturity) (Figure 2.5, Panels A and B). Bacteroid density increased further during development (Figure 2.7). Plant mitochondria and plastids accumulated near the cell wall, particularly at the three cell junction where the middle lamella is expanded. This is likely the location of optimal gas (oxygen) diffusion and indicates the hypoxic nature of the plant cytosol. In contrast, at 21dpi the mutant infected plant cells were small, each with a prominent nucleus, a large central vacuole, and organelles dispersed throughout the cytoplasm (Figure 2.5, Panels C and D) and contained few endocytosed rhizobia

(Figure 2.7). At 28dpi, infected cells enlarged to the size of a typical infected cell and plant organelles clustered at the three cell junctions. The mutant bacteroid density also increased to wild type levels (Figure 2.7) and bacteroids produced pronounced PHB granules (Figure 2.6, Panels C and D). However, some structural abnormalities were observed. Many mutant symbiosomes enclosed large numbers of bacteroids and contained excessive matrix material resembling lytic compartments. Mutant symbiosomes had loosely associated peribacteroid membranes which contrasts the tightly associated peribacteroid membrane observed in wild type symbiosomes (Figure 2.6). Taken together, these results suggest that the mutant induces a delayed symbiotic response as compared to wild type and symbiosomes develop abnormally.

Discussion

The acyl portion of *Rhizobium* lipid A contains the unusual very long chain secondary fatty acid (VLCFA) and is the subject of this work. Interestingly, a number of facultative intracellular plant and animal pathogens and plant symbionts also contain VLCFA in their lipid As and also maintain the cluster of genes necessary for VLCFA synthesis (6, 37). Therefore, it has been speculated that the VLCFA may have biological significance concerning rhizobia/legume symbiosis and plant and animal pathogenesis. Currently, the gene responsible for the production of the specialized acyl carrier protein *acpXL* has been mutated in organisms which form symbiosis via indeterminate nodulation and demonstrate the importance of AcpXL for VLCFA synthesis and normal symbiotic development (17, 38). As stated in the introduction, indeterminate systems have key fundamental differences in terms of bacterial infection and nodule development when compared with determinate systems. For this reason, we created an *acpXL* mutant in an organism that forms determinate nodules with its host plant.

In this study we clearly demonstrate the preparation of a mutant in the model organism *R. leguminosarum* bv. *phaseoli* 8002, a symbiont of determinant nodule-forming *Phaseolus vulgaris*, that lacks a functional *acpXL* gene and gave it the strain name ODB30. Our results can be summarized as follows: (i) Structural analysis of the mutant LPS revealed that the lipid A from free living bacteria and nodule derived bacteroids lacked the VLCFA, and contained lipid A structures in which VLCFA was partially replaced with palmitate. Our results indicate a decrease in the conversion of the lipid A proximal glucosamine to 2-aminogluconate in the bacteroid lipid A of both parent and mutant strain when compared with their free living counterparts. Other than GlcN, there was no observable difference between the carbohydrate portion of mutant LPS when compared to parent strain LPS. (ii) Physiological assays revealed that the membrane stability of the mutant was compromised as indicated by increased sensitivity to detergents and NaCl. (iii) Despite these defects, the mutant was able to form a nitrogen fixing symbiosis with the host plant *Phaseolus vulgaris*. However, detailed electron microscopy revealed that the mutant infection process was significantly delayed and produced enlarged, aberrantly shaped symbiosomes which contained loosely attached peribacteroid membranes.

The structural features of the mutant lipid A presented in this report resembled previously characterized *acpXL* mutants in the indeterminate nodulating model organisms *R. leguminosarum* bv. viciae 3841 and *Sinorhizobium meliloti* 1021 (14, 38) in that the lipid A from cultured bacteria lacked VLCFA producing a lipid A species that was tetraacylated. The mutants were able to partially replace the missing VLCFA with palmitate or possibly stearate to produce a pentaacylated lipid A species. In addition, an *lpxXL* mutant was created in *S. meliloti* and only produced tetraacylated lipid A that lacked both VLCFA and palmitate or stearate (14) suggesting that the LpxXL gene product was responsible for the addition of the acyloxy palmitate or

VLCFA which produces pentaacylated lipid A. Therefore, the addition of palmitate to the lipid A of mutant ODB30 is likely due to the undisrupted *lpxXL* gene. The *lpxXL* gene codes for a lipid transferase responsible for the addition of VLCFA to lipid A (8, 9) and seems to be functional in the *acpXL* mutants. It is likely that, in the absence of AcpXL, LpxXL is able to utilize palmitate from the cytosolic pool of acyl carrier proteins carrying palmitate and possibly stearate. Work to create mutations in the *lpxXL* gene, as well as other putative genes involved in VLCFA synthesis, in strain *R. leguminosarum* bv. phaseoli 8002 is currently on going and will further provide information concerning the role of VLCFA and LpxXL for this determinate nodule-forming symbiotic system.

As described previously, an *acpXL* mutant in the model strain *R. leguminosarum* bv. *viciae* 3841, an indeterminate nodulating endosymbiont, was able to restore VLCFA to the lipid A upon infection of the host plant (37). Restoration of VLCFA did not occur in the isolated mutant nodule bacteria when cultured in the free-living state suggesting that restoration of the VLCFA occurred *in planta*. It is unclear how strain 22 was able to restore VLCFA *in planta*. Genome analysis revealed a possible second copy of the *acpXL* and perhaps this gene is expressed *in planta*. Interestingly, a putative dna binding motif (nif box) was observed upstream of this possible *acpXL* suggesting that this gene is positively regulated by the oxygen regulated NifA transcription factor which is expressed in bacteroids. Surprisingly, the *acpXL-minus* mutant strain ODB30 described in this study did not restore VLCFA *in planta*. Whether or not strain ODB30 contains a second *acpXL* is unknown and this information is not readily available as the parent strain's genome is not sequenced. It is worth noting that strain ODB30 infects a determinate host legume while strain 22 infects an indeterminate host legume. The restoration of

VLCFA in strain 22 may reflect a different requirement for symbiotic infection among indeterminate versus determinate systems.

The lipid A from both wild type and mutant bacteroids displayed an increase in the amount of GlcN (Table 2.3) when compared with their free-living counterparts which suggests that the amount of proximal 2-aminogluconate (GlcNonate) decreased in the bacteroid lipid A of both the parent and mutant strain. The conversion of proximal glucosamine to GlcNonate is an oxygen dependent reaction carried out in the outer membrane monoxygenase LpxQ (28, 29). Bacteroid-containing nodule tissue maintains a microaerobic environment and may explain the decreased GlcNonate conversion due to the limiting oxygen available for the LpxQ enzyme. It is also possible that *lpxQ* may be down regulated during bacteroid development.

The VLCFA makes up a considerable amount of the hydrophobic portion of lipid A and its loss likely lessens hydrophobic interactions in the outer membrane that contribute to membrane stability. Therefore, we tested the relative membrane stability of the mutant compared to that of the parent. As expected, the mutant was less resistant to detergents which suggested a compromise to membrane stability. Interestingly, the *lpxXL* mutant previously described in *S. meliloti* (14) was hypersensitive to detergents when compared to the *S. meliloti acpXL* mutant and parent strain suggesting that the presence of palmitoylated lipid A (penta acylated lipid A) added a considerable amount of stability to the outer membrane. The partial replacement of VLCFA by palmitate in mutant ODB30 may also help stabilize the membrane. An *lpxXL* mutant in strain 8002 may provide clues as to whether or not this might be the case. It is important to note that it is unclear if the *acpXL* or *lpxXL* mutant bacteria currently investigated are able to modify the phospholipid portion of the inner and outer membranes or the peptidoglycan in a manner that could possibly have a compensatory affect on the stability of the total bacterial

envelope. Further investigation into the comprehensive make up of the cellular envelope is needed to address these questions.

It has been suggested that salt tolerance is an important symbiotic factor and disruption of mechanisms that regulate osmotolerance in the bacterial symbiont is detrimental to the symbiotic process (4). The mutant ODB30 was less osmotolerant than the parent strain. A previously described *acpXL* mutant in the pea symbiont *R. leguminosarum* bv. viciae 3841 (36-38) also displayed osmosensitivity. However passage of this mutant through the pea nodule resulted in mutant isolates that were no longer osmosenstive. As was the case with the *acpXL* mutant in strain *R. leguminosarum* bv. *viciae* 3841, bean nodule derived ODB30 bacteria were tolerant to increasing salt conditions on the level of wild type (Table 2.2). It is possible that the plant environment selects a salt tolerant bacterial variant which may be required for complete symbiosis. The disruption in salt tolerance may also account for the observed delay in the infection process by the mutant due to the time required to accumulate the variant (Figure 2.5 and 8). When the nodule derived *R. leguminosarum* bv. viciae 3841 *acpXL* mutant salt tolerant isolate was inoculated on pea plants, nodule infection and bacteroid development proceeded on a similar time scale (J. Gann et al., submitted). However, the bacteroids remained enlarged and aberrantly shaped suggesting that the disruption in bacteroid development is not due to the disruption in salt tolerance.

Basu et. al. (2) have previously reported the appearance of 3-*O*-deacylase activity in the membranes of strains *R. leguminosarum* 8401 and *R. etli* CE3 which cleaves the β-hydroxymyristate from the 3 position of the proximal glucosamine residue of lipid A. The observed activity is likely due to a PagL like outer membrane enzyme that has been characterized in the bacterial pathogens *Salmonella enterica, S. typhimurium, Pseudomonas*

aeruginosa and *Bordatella bronchiseptica* (15). The PagL from *S. typhimurium* has been implicated in the modulation of the human innate immune response and adaptation to the host environment (20). Structural evidence provided by D'Haeze et. al. (12) demonstrated that *R. etli* CE3 free living bacteria contain a mixture of lipid A that contains a modified tetraacylated lipid A, likely due to PagL activity, and that bacteroids exclusively produce tetraacylated lipid A suggesting a possible role for PagL in the symbiotic infection of bean by this organism. Evidence provided in this work shows that *R. leguminosarum* bv. phaseoli 8002 does not contain detectable amounts of PagL-modified lipid A in free living bacteria or bacteroids. The tetraacylated lipid A observed in strain 8002 by mass spectrometry (Figure 2.3) is due to beta elimination and these ions are distinguishable from PagL-modified tetraacylated lipid A These results were surprising since 3-*O*-deacylase activity was previously detected in the membranes of the 8002 derived strain 8401 (above). The role of PagL in symbiosis is an unexplored area of research. It is possible that PagL modified lipid A helps regulate the plant defense system and aids in the infection process of symbionts and/or endophytes. The *pagL* gene in *R. etli* CE3 has recently been identified in our laboratory and efforts to examine the function of this enzyme in a variety of symbiotic strains are currently in progress.

Acknowledgments

This work was supported, in whole or in part, by National Institutes of Health Grant GM39583 (to R.W.C.) and by the Department of Energy Grant DE-FG-02-09ER20097 to the CCRC.

Footnotes

Abbreviations used in this study are: LPS, lipopolysaccharide; Kdo, 3-deoxy-*D*-*manno*-2 octulosonic acid; 6dTal, 6-deoxytalose; QuiNAc, quinavosamine; GlcN, glucosamine; Gal, galactose; GalA, galacturonic acid; Man, mannose; Glc, glucose; GlcNonate, 2-aminogluconate; BHB, β-hydroxy butyrate; 3-OHC14:0, β-hydroxy-myristate; OHC16:0, β-hydroxypalmitate; 3-OHC18:0, β-hydroxystearate; dod-P-GalA, dodecaprenyl phosphogalacturonic acid; MALDI-TOF-MS, matrix assisted desorption ionization time of flight mass spectrometry; GC/MS, gas chromatography in line with mass spectrometry; DOC-PAGE, deoxycholic acid polyacrylamide gel electrophoresis; TEM, transmission electron microscopy

References

1. **Basu, S. S., M. J. Karbarz, and C. R. H. Raetz.** 2002. Expression cloning and characterization of the C28 acyltransferase of lipid A biosynthesis in *Rhizobium leguminosarum*. J. Biol. Chem. **277:**28959-28971.
2. **Basu, S. S., K. A. White, N. L. S. Que, and C. R. H. Raetz.** 1999. A deacylase in *Rhizobium leguminosarum* membranes that cleaves the 3-*O*-Linked Beta-hydroxymyristoyl moiety of lipid A precursors. J. Biol. Chem. **274:**11150-11158.
3. **Becker, A., M. Schmidt, W. Jäger, and A. Pühler.** 1995. New gentamicin-resistance and *lacZ* promoter-probe cassettes suitable for insertion mutagenesis and generation of transcriptional fusions. Gene **162:**37-39.
4. **Beedveld, M. W., and Karen J. Miller.** 1998. Cell surface β-glucans. *In* H. P. Spaink, A. Kondorosi, and P. J. J. Hooykaas (ed.), The *Rhizobiaceae*: molecular biology of model plant-associated bacteria. Kluwer Academic Publishers, Dordrecht, The Netherlands.
5. **Beringer, J. E.** 1974. R factor transfer in *Rhizobium leguminosarum*. Microbiol. **84:**188.
6. **Bhat, U. R., R. W. Carlson, M. Busch, and H. Mayer.** 1991. Distribution and phylogenetic significance of 27-hydroxy-octacosanoic acid in lipopolysaccharides from bacteria belonging to the alpha-2 subgroup of Proteobacteria. Int. J. Syst. Bacteriol. **41:**213-217.
7. **Bhat, U. R., L. S. Forsberg, and R. W. Carlson.** 1994. Structure of lipid A component of *Rhizobium leguminosarum* bv. *phaseoli* lipopolysaccharide. Unique nonphosphorylated lipid A containing 2-amino-2-deoxygluconate, galacturonate, and glucosamine. J. Biol. Chem. **269:**14402.

8. **Brozek, K. A., R. W. Carlson, and C. R. H. Raetz.** 1996. A special acyl carrier protein for transferring long hydroxylated fatty acids to lipid A in *Rhizobium*. J. Biol. Chem. **271:**32126-32136.

9. **Brozek, K. A., J. L. Kadrmas, and C. R. H. Raetz.** 1996. Lipopolysaccharide Biosynthesis in *Rhizobium leguminosarum*: Novel enzymes that process precursors containing 3-deoxy-D-manno-octulosonic acid. J. Biol. Chem. **271:**32112-32118.

10. **Carlson, R. W.** 1984. Heterogeneity of *Rhizobium* lipopolysaccharides. J. Bacteriol. **158:**1012-1017.

11. **Carlson, R. W., L. S. Forsberg, and E. L. Kannenberg.** 2010. Lipopolysaccharides in *Rhizobium*-legume symbioses, p. 339-386, Subcellular Biochemestry: Endotoxins: structure, function and recognition, P. J. Quinn and X. Wang ed, vol. 53. Springer.

12. **D'Haeze, W., C. Leoff, G. Freshour, K. D. Noel, and R. W. Carlson.** 2007. *Rhizobium etli* CE3 bacteroid lipopolysaccharides are structurally similar but not identical to those produced by cultured CE3 bacteria. J. Biol. Chem. **282:**17101-17113.

13. **Fahraeus, G.** 1957. The infection of clover root hairs by nodule bacteria studied by a simple glass slide technique. J. Gen. Microbiol. **16:**374-381.

14. **Ferguson, G. P., A. Datta, R. W. Carlson, and G. C. Walker.** 2005. Importance of unusually modified lipid A in Sinorhizobium stress resistance and legume symbiosis. Mol. Microbiol. **56:**68-80.

15. **Geurtsen, J., L. Steeghs, J. t. Hove, P. van der Ley, and J. Tommassen.** 2005. Dissemination of lipid A deacylases (PagL) among gram-negative bacteria. J. Biol. Chem. **280:**8248-8259.

16. **Gudlavalleti, S. K., and L. S. Forsberg.** 2003. Structural characterization of the lipid A component of *Sinorhizobium sp*. NGR234 rough and smooth form lipopolysaccharide. J. Biol. Chem. **278:**3957-3968.

17. **Haag, A. F., S. Wehmeier, S. Beck, V. L. Marlow, V. Fletcher, E. K. James, and G. P. Ferguson.** 2009. The *Sinorhizobium meliloti* LpxXL and AcpXL proteins play important roles in bacteroid development within alfalfa. J. Bacteriol. **191:**4681-4686.

18. **Hoang, T. T., R. A. R. Karkhoff-Schweizer, A. J. Kutchma, and H. P. Schweizer.** 1998. A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants. Gene **212:**77-86.

19. **Kannenberg, E. L., and N. J. Brewin.** 1994. Host-plant invasion by *Rhizobium*: the role of cell-surface components. Trends Microbiol. **2:**277-283.

20. **Kawasaki, K., R. K. Ernst, and S. I. Miller.** 2004. 3-*O*-Deacylation of lipid A by PagL, a PhoP/PhoQ-regulated deacylase of *Salmonella typhimurium*, modulates signaling through Toll-like Receptor 4. J. Biol. Chem. **279:**20044-20048.

21. **Kittelberger, R., and F. Hilbink.** 1993. Sensitive silver-staining detection of bacterial lipopolysaccharides in polyacrylamide gels. J. Biochem. Biophys. Meth. **26:**81-86.

22. **Kovach, M. E., P. H. Elzer, D. Steven Hill, G. T. Robertson, M. A. Farris, R. M. Roop, and K. M. Peterson.** 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene **166:**175-176.

23. **Lamb, J. W., G. Hombrecher, and A. W. B. Johnston.** 1982. Plasmid-determined nodulation and nitrogen-fixation abilities in *Rhizobium phaseoli.* Mol. Gen. Genet. **186:**449-452.

24. **Mollenhauer, H. H.** 1964. Plastic embedding mixture for use in electron microscope. Stain Technol. **39:**111-114.

25. **Nicolas, F., C. François, and P. Véréna.** 2003. Surface polysaccharide involvement in establishing the *Rhizobium* / legume symbiosis. Eur. J. Biochem. **270:**1365-1380.

26. **Oldroyd, G. E. D., and J. A. Downie.** 2008. Coordinating nodule morphogenesis with rhizobial Infection in legumes. Annu. Rev. Plant Biol. **59:**519-546.

27. **Phadnis, S. H., and H. K. Das.** 1987. Use of the plasmid pRK 2013 as a vehicle for transposition in *Azotobacter vinelandii*. J. Biosciences **12:**131-135.

28. **Que-Gewirth, N. L. S., M. J. Karbarz, S. R. Kalb, R. J. Cotter, and C. R. H. Raetz.** 2003. Origin of the 2-amino-2-deoxy-gluconate unit in *Rhizobium leguminosarum* Lipid A. J. Biol. Chem. **278:**12120-12129.

29. **Que-Gewirth, N. L. S., S. Lin, R. J. Cotter, and C. R. H. Raetz.** 2003. An outer membrane enzyme that generates the 2-amino-2- deoxy-gluconate moiety of *Rhizobium leguminosarum* Lipid A. J. Biol. Chem. **278:**12109-12119.

30. **Raetz, C. R. H., C. M. Reynolds, M. S. Trent, and R. E. Bishop.** 2007. Lipid A modification systems in gram-negative bacteria. Annu. Rev. Biochem. **76:**295-329.

31. **Reuhs, B. L., R. W. Carlson, and J. S. Kim.** 1993. *Rhizobium fredii* and *Rhizobium meliloti* produce 3-deoxy-D-manno-2-octulosonic acid-containing polysaccharides that are structurally analogous to group II K-antigens (capsular polysaccharides) found in *Escherichia coli*. J. Bacteriol. **175:**3570-3580.

32. **Ryan, J. M., Conrad, H. E.** 1974. Structural heterogeneity in the lipopolysaccharide of *Salmonella newington*. Arch. Biochem. Biophys. **162:**530-536.

33. **Sharypova, L. A., K. Niehaus, H. Scheidle, O. Holst, and A. Becker.** 2003. *Sinorhizobium meliloti acpXL* mutant lacks the C28 hydroxylated fatty acid moiety of lipid A and does not express a slow migrating form of lipopolysaccharide. J. Biol. Chem. **278:**12946-12954.

34. **Van de Velde, W., G. Zehirov, A. Szatmari, M. Debreczeny, H. Ishihara, Z. Kevei, A. Farkas, K. Mikulass, A. Nagy, H. Tiricz, B. Satiat-Jeunemaitre, B. Alunni, M. Bourge, K. Kucho, M. Abe, A. Kereszt, G. Maroti, T. Uchiumi, E. Kondorosi, and P. Mergaert.** 2010. Plant peptides govern terminal differentiation of bacteria in symbiosis. Science **327:**1122-1126.

35. **Vanderlinde, E. M., A. Muszynski, J. J. Harrison, S. F. Koval, D. L. Foreman, H. Ceri, E. L. Kannenberg, R. W. Carlson, and C. K. Yost.** 2009. *Rhizobium leguminosarum* biovar *viciae* 3841, deficient in 27-hydroxyoctacosanoate-modified lipopolysaccharide, is impaired in desiccation tolerance, biofilm formation and motility. Microbiol. **155:**3055-3069.

36. **Vedam, V., J. G. Haynes, E. L. Kannenberg, R. W. Carlson, and D. J. Sherrier.** 2004. A *Rhizobium leguminosarum* lipopolysaccharide lipid-A mutant induces nitrogen-fixing nodules with delayed and defective bacteroid formation. Mol. Plant-Microbe Interact. **17:**283-291.

37. **Vedam, V., E. Kannenberg, A. Datta, D. Brown, J. G. Haynes-Gann, D. J. Sherrier, and R. W. Carlson.** 2006. The pea nodule environment restores the ability of a *Rhizobium leguminosarum* lipopolysaccharide *acpXL* mutant to add 27-hydroxyoctacosanoic acid to Its lipid A. J. Bacteriol. **188:**2126-2133.

38. **Vedam, V., E. L. Kannenberg, J. G. Haynes, D. J. Sherrier, A. Datta, and R. W. Carlson.** 2003. A *Rhizobium leguminosarum acpXL* mutant produces lipopolysaccharide lacking 27-hydroxyoctacosanoic acid. J. Bacteriol. **185:**1841-1850.

39. **Westphal, O., and K. Jann.** 1965. Endotoxins: Extraction with phenol-water and further applications of the procedure. Methods Carbohydr. Chem. **5:**83–91.

40. **White, S. W., J. Zheng, Y.-M. Zhang, and C. O. Rock.** 2005. The structural biology of type II fatty acid biosynthesis. Annu. Rev. Biochem. **74:**791-831.

41. **York, W. S., Darvill, A. G., McNeil, M., Stevenson, T. T., and Albersheim, P.** 1986. Isolation and characterization of plant-cell walls and cell-wall components Methods Enzymol. **118:**3-40.

Figure Legends

Figure 2.1. A model for the synthesis of very long chain fatty acid (VLCFA) in *Rhizobium.* The six genes predicted to be responsible for synthesis of the VLCFA in *Rhizobium* are localized on the chromosome in a gene cluster (arrows indicate the direction of transcription). The VLCFA is built on the specialized acyl carrier protein AcpXL and the *fab* gene products are predicted to be involved in the fatty acid elongation process. The acyl transferase LpxXL transfers the VLCFA to the N-linked 3-hydroxymyristate at the 2` position of the distal glucosamine of Lipid A. GalA = galacturonic acid, GlcN 2–aminoglucose, GlcNonate = 2-aminogluconate, 3-OHC14:0 = hydroxymyristate, Core = core oligosaccharide, O-Chain = O-Chain polysaccharide

Figure 2.2. DOC-PAGE / Silver Stained mutant LPS has a different migration pattern as compared to wild type. 2µg of extracted LPS was loaded in each well. Lane 1) Wild Type 8002. Lane 2) Mutant OLDB30. LPS-II = molecular species composed of Lipid A and Core (lacks O-Chain), LPS-I = molecular species composed of Lipid A, Core, and O-Chain.

Figure 2.3. MALDI-TOF / MS analysis of lipid A from strains *R. leguminosarum* 8002 and OLDB30. Lipid A from free-living bacteria and bacteroids was separated from LPS and analyzed by MALDI-TOF / MS. Data was acquired in the negative linear mode. Panel I) Free-living 8002. Panel II) Free-living OLDB30. Panel III) Bacteroid 8002. Panel IV) Bacteroid OLDB30. The predicted minimal lipid A structures and their calculated molecular weights are displayed and are labeled A through I. Structures that contain the VLCFA may contain β-hydroxybutoxy (BHB) or a hydroxy group at position 27 of the VLCFA. For a comprehensive list of ions see Table 2.4. B = Structures A containing proximal GlcNonate, C = Structures A containing proximal GlcNonolactone minus 3-*O*-hydroxy acyl chain, E = Structures D containing GlcNonate, F = Structures D containing proximal 2,3-unsaturated GlcNonolactone

minus 3-*O*-hydroxy acyl chain, H = Structures G containing proximal 2,3-unsaturated GlcNonate, I = Structures G containing proximal GlcNonolactone minus 3-*O*-hydroxy acyl chain.

Figure 2.4. b^+ ions of lipid A acquired by MALDI-TOF/MS. Data was acquired in the positive linear mode. The spectra show the region where b^+ ions occur. Panel A) Free living wild type 8002 lipid A. Panel B) Bacteroid 8002 lipid A. Panel C) Free living mutant OLDB30 lipid A. Panel D) Bacteroid OLDB30 lipid A. BHB = β-hydroxybutyrate, F = the $[M+Na]^+$ ions of the structure explained in the legend of Figure 2.3. B =

Figure 2.5. Transmission electron microscopy of wild type and mutant-infected nodules at 21dpi. A. Low magnification view of *R. leguminosarum* bv. *phaseoli* 8002 wild type infected cells in the fixation zone, showing wild type infected cells contained a great number of bacteroids and large vacuoles. **B.** Higher-magnification view of wild type infected cells **C.** Low magnification image of mutant infected cells in the fixation zone, illustrating the smaller size of infected cells and the low density of bacteroids compared to WT infected cells. **D.** Higher-magnification of mutant infected cells**.**

Figure 2.6. Transmission electron microscope of wild and mutant-infected nodules at 28dpi. A. Low magnification view of *Rhizobium leguminosarum* bv. *phaseoli* 8002 wild type bacteroids in fixation zone, showing each infected cell contained a great level of mature bacteroids. **B.** Higher-magnification view of panel A, indicating mostly one bacteroid in one intact symbiotic compartment. Mature bacteroids were dividing and mitochondria were lined along the cell wall. **C.** The low magnification view of mutant infected cells, demonstrating no significant differences in cell size and number of bacteroids as compared to wild type infected cells. **D.** shows abnormalities in mutant infected cells. Multiple bacteroids located in one irregular-shaped and

lytic symbiotic compartment and the infected cell has lost a great level of cellular content. The arrows in panel **C** indicate multiple bacteroids in a single symbiotic compartment. Scale bars = 1 μm.

Figure 2.7. Bacteroid measurement under TEM. Bacteroids were counted in a whole area of one infected cell using Image J and re-calculated as bacteroid numbers per 100 μm^2. Bars represent the mean value of bacteroid numbers per 100 μm^2. **A,** Mean values of bacteroid numbers per 100 μm^2 of wild type (n = 6) and mutant (n = 16), showing statistically few bacteroids in mutant-inoculated nodules at 21 dpi. **B,** Mean values of bacteroid numbers per 100 (μm)2 of wild type (n = 5) and mutant (n = 5), demonstrating statistically no significant differences in bacteroid numbers between wild type and mutant-inoculated nodules at 28dpi. Bacteroids in infected threads were excluded in the present analysis. Comparison between the wild type and mutant bacteroids were made using student -test. *** Indicates the significant difference at $p<0.001$. N.S. indicates no significant difference at $p = 0.1377$.

Table 2.1. Bacterial strains and plasmids used in this report

Strains or plasmids	Characteristics	Source
E. coli		
DH5α	Φ80d*lacZ _M15 recA1 endA1 gyrA96 thi-1 hsdR17* (rK- mK-) *supE44 relA1 deoR* Δ(*lacZYA argF) U169*	Stratagene
R. lequminosarum* bv. *phaseoli		
8002	Isolated from Bean, Fix^+	(23)
ODB30	Strain 8002 *acpXL::accC1*, Gm^r Fix^+	This Study
ODB30/pVV5	Strain ODB30 complemented with plasmid pVV5, Gm^r Tc^r	This Study
Plasmids		
pMS255	Source of *aacC1 Gm* cassette, Gm^r	(3)
pRK2013	Mobilizing plasmid for pEX-Tc, Col E1 replicon, Kan^r	(27)
pEX18-Tc	Suicide vector, allows positive selection for integration, Tc^r	(18)
pVV1	pBS SK derivative with 3.9-kb insert containing *acpXL,* Amp^r	(38)
pODB1	pVV1 derivative containing *Gm* cassette from pMS255	This Study
pODB3	pEX-Tc containing a 4.7-kb insert from pODB1	This Study
pVV5	pBBR1MCS-3 containing a 2.5-kb region from pVV1 that harbors *acpXL,* used for complementation of OLDB30*,* Tc^r	(38)

Table 2.2. *Sensitivity to Detergents and Sodium Chloride

DOC (mM)

Strain	1	2	3
8002	+	+	-
ODB30	+	-	-
Complement	+	+	-
ODB30 NB	+	-	-

SDS (µg/ml)

Strain	0	100	200
8002	+	+	+
ODB30	+	-	-

NaCL (%)

0	0.5	0.75	1.0
8002	+	+	+	-
ODB30	+	-	-	-
Complement	+	+	+	-
ODB30 NB	+	+	+	-

Strain

Strain	Polymyxin B MIC (µg/ml)
8002	16
ODB30	8
Complement	16
ODB30 NB	8

*Strains were grown on Y minimal media containing the indicated concentrations of sodium desoxycholate (DOC), TY media containing sodium dodecyl sulfate (SDS), sodium chloride (NaCl) or polymyxin B. (+) indicates that growth was observed visually. (-) indicates no growth was observed. MIC = minimal inhibitory concentration. 8002 = parent strain, ODB30 = *acpXL* mutant, Complement = mutant complemented with functional *acpXL* gene, ODB30 NB = nodule derived mutant bacteria

Table 2.3. Relative percent of peak ratios obtained from GC/MS†

Free-living Bacteria

Total LPS	Rha	Fuc	Man	Gal	Kdo	GlcNAc	GalA	Glc	OHC14:0	OHC15:0	OHC16:0	OHC18:0	OHC28:0	C16:0
8002	22.9	18.1	26.4	1.9	3.8	Tr	4.4	16.5	4.4	Tr	0.6	0.6	-*	-
ODB30	30.4	23.9	17.4	2.7	6.7	Tr	6.5	0.8	6.5	0.4	1.3	4.3	-	-
LipidA	Rha	Fuc	Man	Gal	Kdo	GlcNAc	GalA	Glc	OHC14:0	OHC15:0	OHC16:0	OHC18:0	OHC28:0	C16:0
8002	-	-	-	-	-	4.0	7.0	-	29.6	6.0	8.2	12.0	32.1	0.7
ODB30	-	-	-	-	-	5.2	10.9	-	39.0	7.8	11.7	19.2	-	6.0

Bacteroids

Total LPS	Rha	Fuc	Man	Gal	Kdo	GlcNAc	GalA	Glc	OHC14:0	OHC15:0	OHC16:0	OHC18:0	OHC28:0	C16:0
8002	17.3	15.6	15.6	1.7	3.4	0.5	5.2	31.3	5.2	1.0	0.3	0.6	3.4	-
ODB30	24.6	21.0	29.5	2.2	3.1	0.4	3.1	8.9	3.1	0.8	Tr	2.6	-	-
LipidA	Rha	Fuc	Man	Gal	Kdo	GlcNAc	GalA	Glc	OHC14:0	OHC15:0	OHC16:0	OHC18:0	OHC28:0	C16:0
8002	-	-	-	-	-	14.5	3.6	-	32.7	10.5	6.7	7.9	22.5	1.2
ODB30	-	-	-	-	-	8.2	10.0	-	38.5	16.5	8.3	9.6	-	8.0

†Composition of total LPS and Lipid A from Free-Living and Bacteroid strains was obtained by GC/MS analysis of trimethyl silylated methyl ester and methyl glycoside derivitives. Displayed is the relative percentage of integrated peaks within each chromatogram obtained. Rhamnose (Rha), Fucose (Fuc), Mannose (Man), Galactose (Gal), 3-deoxy-*D*-*manno*-2-octulosonic acid (Kdo), N-acetyl glucosamine (GlcNAc), Galacturonic acid (GalA), Glucose (Glc), Hydroxymyristate (3-OHC14:0), Hydroxypentadecanoate (3-OHC15:0), hydroxypalmitate (3-OHC16:0), Hydroxystearate, (3-OHC18:0), 27-hydroxyoctacosanoate (27-OHC28:0). Tr, trace amounts. * It is not clear why 27-OHC28:0 was not detected in the total LPS of the parent strain. However, this issue is resolved by composition (above) and mass spectroscopic (Figure 2.3) analysis of partitioned Lipid A molecules which clearly demonstrates the presence of 27-OHC28:0 in the parent strain lipid A and the absence of 27-OHC28:0 in the mutant lipid A.

Table 2.4. T\he major ions observed from MALDI/TOF MS spectra of free-living *R. leguminosarum* 8002 and the mutant *R. leguminosarum* ODB30 lipid

Strain	Observed [M-H]-	Calculated [M-H]-	Proposed Composition
8002			
Stucture A	1931.2	1930.6	GalA, †$GlcN_2$, 3-OHC14:0_4, 27-OHC28:0, 3-OHC4:0
	1945.0	1944.7	GalA, $GlcN_2$, 3-OHC14:0_3, 3-OHC15:0, 27-OHC28:0, 3-OHC4:0
	1959.7	1958.7	GalA, $GlcN_2$, 3-OHC14:0_3, 3-OHC16:0, 27-OHC28:0, 3-OHC4:0
Stucture A (-BHB)	1889.4	1900.7	GalA, $GlcN_2$, 3-OHC14:0_3, 3-OHC18:0, 27-OHC28:0
Stucture B1	1959.7	1960.7	GalA, GlcN, GlcNonate, 3-OHC14:0_3, 3-OHC15:0, 27-OHC28:0, 3-OHC4:0
	1974.6	1974.7	GalA, GlcN, GlcNonate, 3-OHC14:0_3, 3-OHC16:0, 27-OHC28:0, 3-OHC4:0
	1988.4	1988.7	GalA, GlcN, GlcNonate, 3-OHC14:0_2, 3-OHC15:0, 3-OHC16:0, 27-OHC28:0, 3-OHC4:0
	2003.4	2002.7	GalA, GlcN, GlcNonate, 3-OHC14:0_3, 3-OHC18:0, 27-OHC28:0, 3-OHC4:0
	2017.2	2016.8	GalA, GlcN, GlcNonate, 3-OHC14:0_2, 3-OHC15:0, 3-OHC18:0, 27-OHC28:0, 3-OHC4:0
	2032.1	2030.8	GalA, GlcN, GlcNonate, 3-OHC14:0_2, 3-OHC16:0, 3-OHC18:0, 27-OHC28:0, 3-OHC4:0
	2044.5	2044.8	GalA, GlcN, GlcNonate, 3-OHC14:0_2, 3-OHC15:0, 3-OHC16:0, 3-OHC18:0, 27-OHC28:0, 3-OHC4:0
	2058.7	2058.9	GalA, GlcN, GlcNonate, 3-OHC14:0_2, 3-OHC18:0_2, 27-OHC28:0, 3-OHC4:0
Stucture B (-BHB)	1874.4	1874.6	GalA, GlcN, GlcNonate 3-OHC14:0_3, 3-OHC15:0, 27-OHC28:0
	1889.4	1888.6	GalA, GlcN, GlcNonate 3-OHC14:0_3, 3-OHC16:0, 27-OHC28:0
	1902.5	1902.6	GalA, GlcN, GlcNonate 3-OHC14:0_2, 3-OHC15:0, 3-OHC16:0, 27-OHC28:0
	1917.3	1916.7	GalA, GlcN, GlcNonate 3-OHC14:0_3, 3-OHC18:0, 27-OHC28:0
	1931.2	1930.7	GalA, GlcN, GlcNonate 3-OHC14:0_2, 3-OHC15:0, 3-OHC18:0, 27-OHC28:0
	1945.0	1944.7	GalA, GlcN, GlcNonate 3-OHC14:0_2, 3-OHC16:0, 3-OHC18:0, 27-OHC28:0
	1974.6	1972.8	GalA, GlcN, GlcNonate 3-OHC14:0_2, 3-OHC18:02, 27-OHC28:0
Structure C	1685.0	1684.3	GalA, GlcN, *GlcNonolactone, 3-OHC14:0_3, 27-OHC28:0, 3-OHC4:0
	1714.2	1712.3	GalA, GlcN, *GlcNonolactone, 3-OHC14:0_2, 3-OHC16:0, 27-OHC28:0, 3-OHC4:0
	1742.6	1740.4	GalA, GlcN, *GlcNonolactone, 3-OHC14:0_2, 3-OHC18:0, 27-OHC28:0, 3-OHC4:0
	1770.6	1768.4	GalA, GlcN, *GlcNonolactone, 3-OHC14:0, 3-OHC16:0, 3-OHC18:0, 27-OHC28:0, 3-OHC4:0
Structure C (-BHB)	1628.4	1626.2	GalA, GlcN, *GlcNonolactone, 3-OHC14:0_2, 3-OHC16:0, 27-OHC28:0
	1656.4	1654.3	GalA, GlcN, *GlcNonolactone, 3-OHC14:0_2, 3-OHC18:0, 27-OHC28:0
ODB30			

Structure D	1449.9	1449.9	GalA, $GlcN_2$, 3-OHC14:0_3, 3-OHC16:0
	1478.4	1477.9	GalA, $GlcN_2$, 3-OHC14:0_3, 3-OHC18:0
Structure E	1478.4	1479.9	GalA, GlcN, GlcNonate, 3-OHC14:0_2, 3-OHC15:0, 3-OHC16:0
	1494.9	1493.9	GalA, GlcN, GlcNonate, 3-OHC14:0_2, 3-OHC16:0_2
Structure F	1177.3	1175..4	GalA, GlcN, *GlcNonolactone, 3-OHC14:0_3
	1205.2	1203.5	GalA, GlcN, *GlcNonolactone, 3-OHC14:0_2, 3-OHC16:0
	1233.4	1231.5	GalA, GlcN, *GlcNonolactone, 3-OHC14:0_2, 3-OHC18:0
	1261.5	1259.6	GalA, GlcN, *GlcNonolactone, 3-OHC14:0_2, 3-OHC16:0, 3-OHC18:0
	1289.7	1287.6	GalA, GlcN, *GlcNonolactone, 3-OHC14:0, 3-OHC18:0_2
Structure G	1717.5	1716.3	GalA, GlcN2, 3-OHC14:0_3, 3-OHC18:0, C16:0
	1732.8	1730.4	GalA, GlcN2, 3-OHC14:0_2, 3-OHC15:0, 3-OHC18:0, C16:0
	1760.5	1758.4	GalA, GlcN2, 3-OHC14:0, 3-OHC15:0, 3-OHC16:0, 3-OHC18:0, C16:0
	1774.8	1772.4	GalA, GlcN2, 3-OHC14:0_2, 3-OHC18:02, C16:0
Structure H	1717.5	1718.3	GalA, GlcN, GlcNonate, 3-OHC14:0_2, 3-OHC15:0, 3-OHC16:0, C16:0
	1732.8	1732.8	GalA, GlcN, GlcNonate, 3-OHC14:0_3, 3-OHC18:0, C16:0
	1760.5	1760.4	GalA, GlcN, GlcNonate, 3-OHC14:0_2, 3-OHC16:0, 3-OHC18:0, C16:0
	1774.8	1774.4	GalA, GlcN, GlcNonate, 3-OHC14:0, 3-OHC15:0, 3-OHC16:0, 3-OHC18:0, C16:0
Structure I	1444.0	1441.9	GalA, GlcN, *GlcNonolactone, 3-OHC14:0_2, 3-OHC16:0, C16:0
	1471.9	1470.0	GalA, GlcN, *GlcNonolactone, 3-OHC14:0_2, 3-OHC18:0, C16:0
	1499.9	1498.0	GalA, GlcN, *GlcNonolactone, 3-OHC14:0, 3-OHC16:0, 3-OHC18:0, C16:0
	1528.7	1526.1	GalA, GlcN, *GlcNonolactone, 3-OHC14:0, 3-OHC18:0_2

* 2, 3-anhydro-2-aminogluconic acid, † the subscripted 2 refers to the number of glucosamine (GlcN) residues

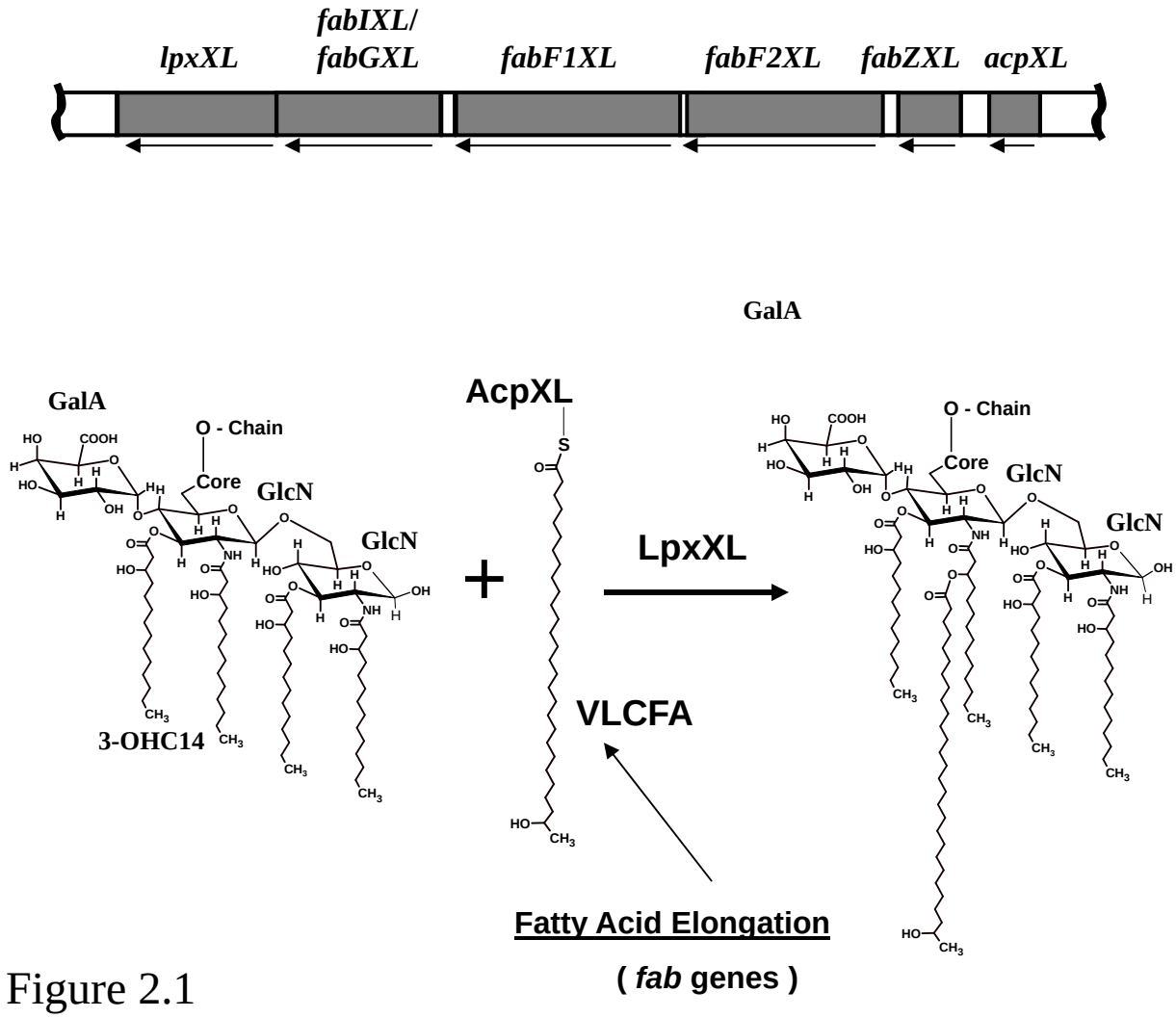

fabIXL/
lpxXL
fabGXL
fabF1XL
fabF2XL
fabZXL
acpXL
GalA
GalA
O - Chain
Core
GlcN
GlcN
3-OHC14
AcpXL
LpxXL
VLCFA
Fatty Acid Elongation
(fab genes)
O - Chain
Core
GlcN
GlcN

Figure 2.1

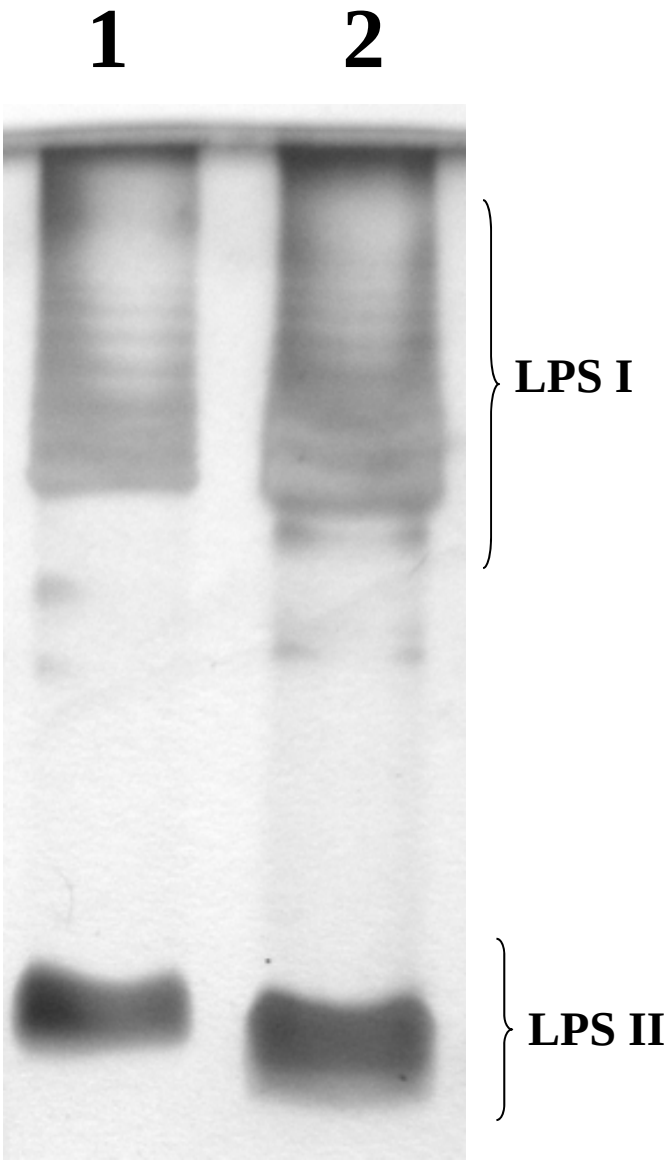


Figure 2.2

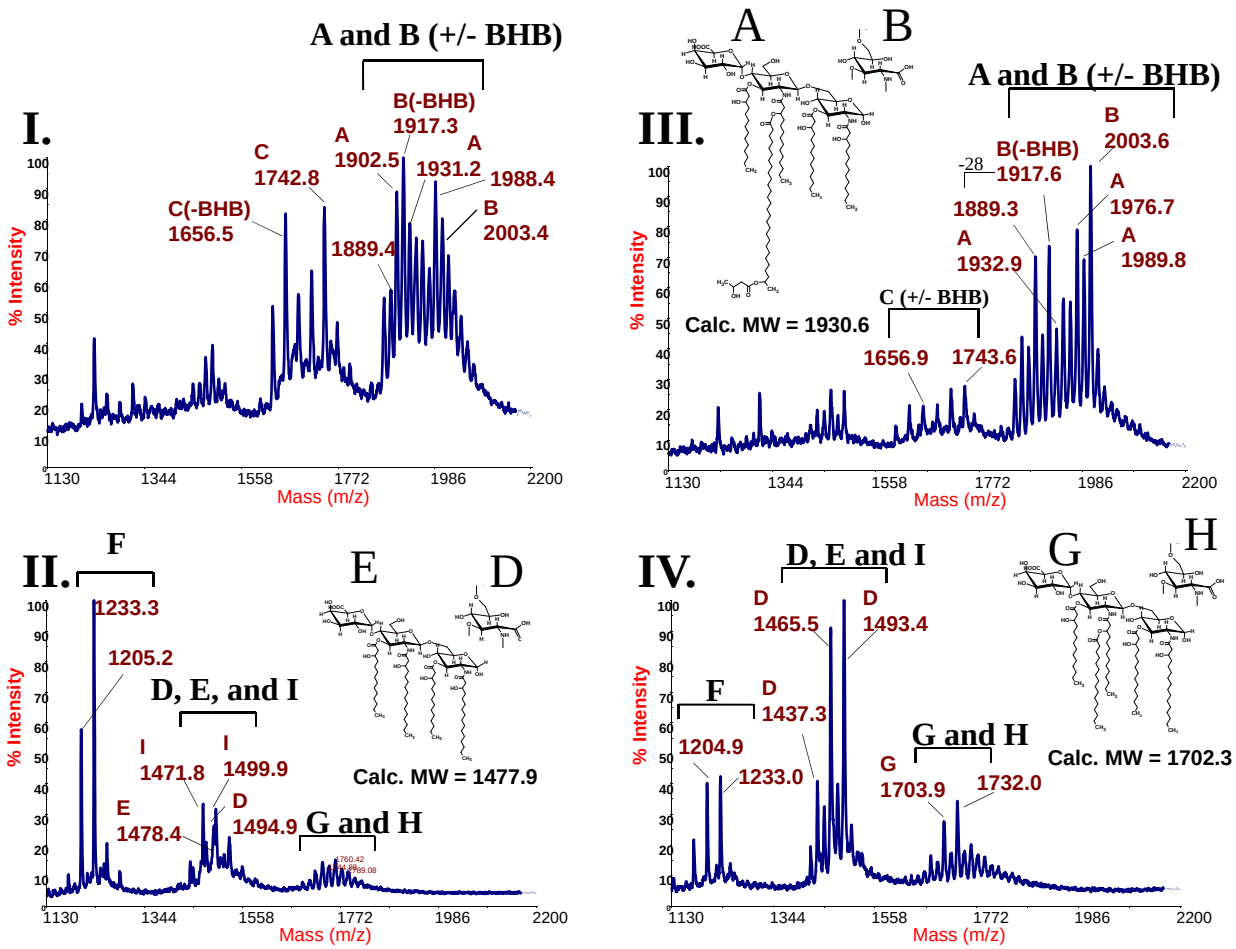

I.
A and B (+/- BHB)
B(-BHB)
1917.3
A
1902.5
C
1742.8
A
1931.2
1988.4
C(-BHB)
1656.5
B
2003.4
1889.4
% Intensity
Mass (m/z)
III.
A
B
A and B (+/- BHB)
B
2003.6
B(-BHB)
1917.6
-28
A
1976.7
1889.3
A
1989.8
A
1932.9
C (+/- BHB)
Calc. MW = 1930.6
1656.9
1743.6
II.
F
1233.3
E
D
1205.2
D, E, and I
I
1471.8
I
1499.9
Calc. MW = 1477.9
E
1478.4
D
1494.9
G and H
IV.
D, E and I
G
H
D
1465.5
D
1493.4
F
D
1437.3
1204.9
1233.0
G and H
Calc. MW = 1702.3
G
1703.9
1732.0

Figure 2.3

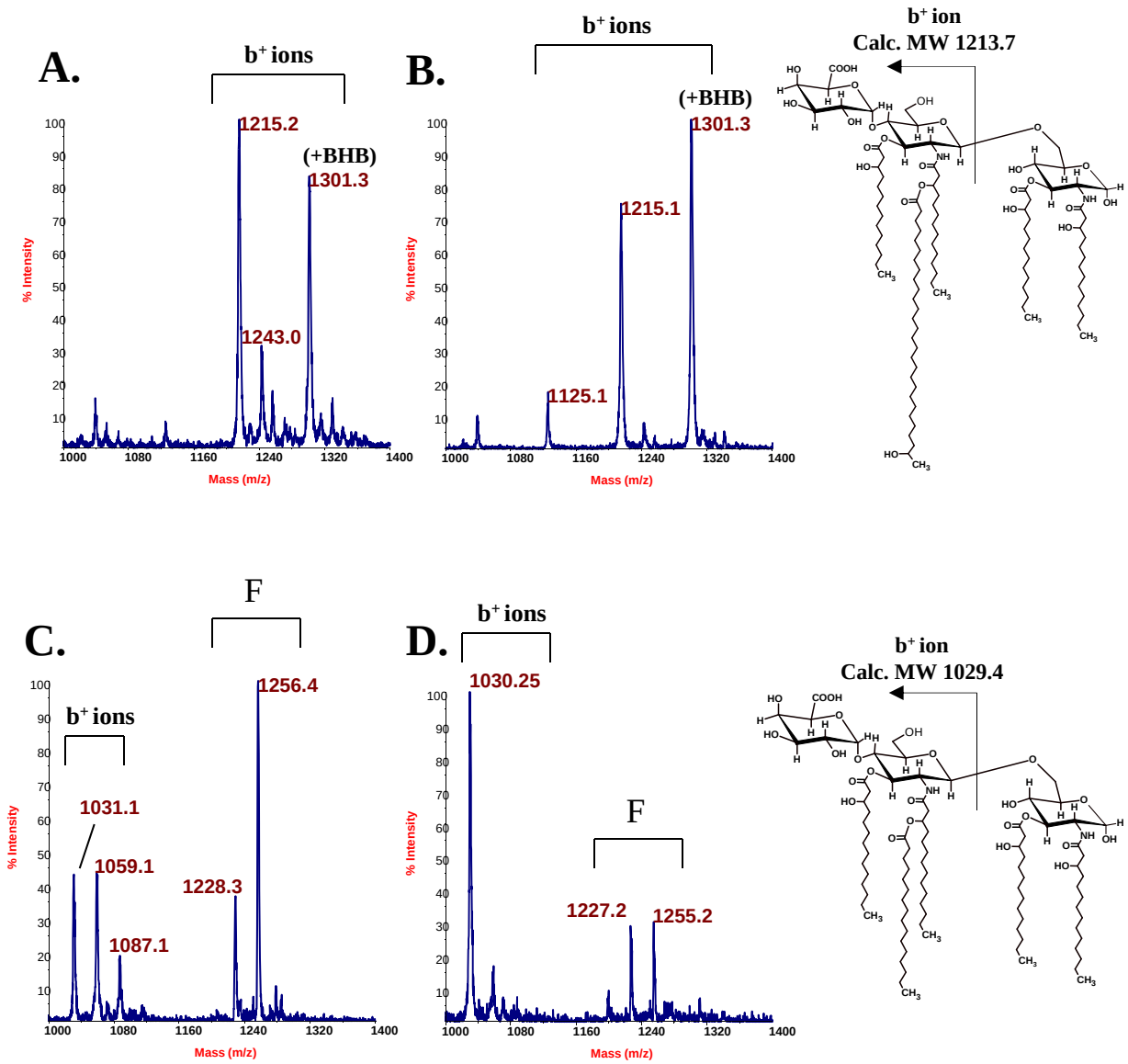

A.
b+ ions
1215.2
(+BHB)
1301.3
1243.0
% Intensity
Mass (m/z)
B.
b+ ions
(+BHB)
1301.3
1215.1
1125.1
% Intensity
Mass (m/z)
b+ ion
Calc. MW 1213.7
C.
F
b+ ions
1256.4
1031.1
1059.1
1228.3
1087.1
% Intensity
Mass (m/z)
D.
b+ ions
1030.25
F
1227.2
1255.2
% Intensity
Mass (m/z)
b+ ion
Calc. MW 1029.4

Figure 2.4

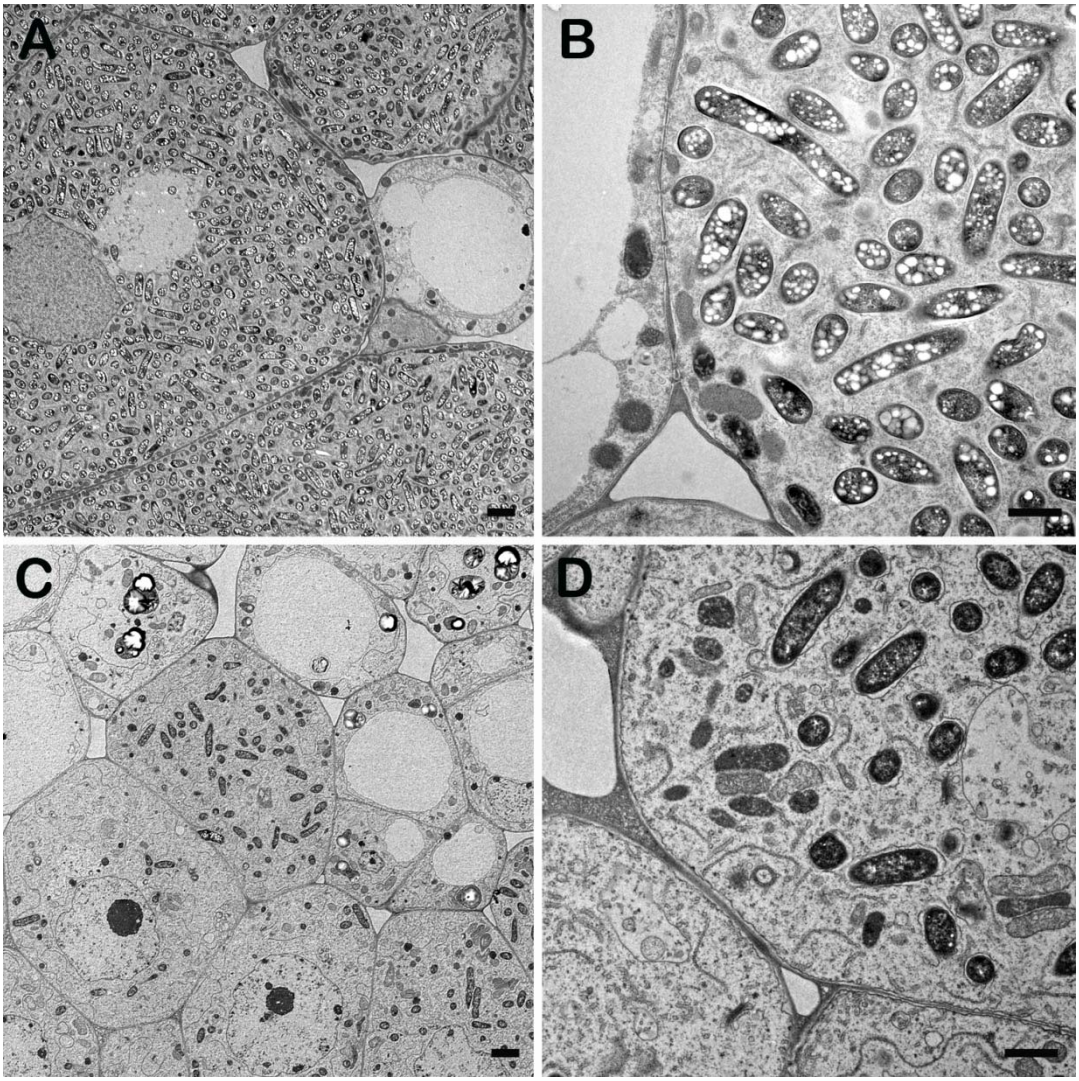


Figure 2.5

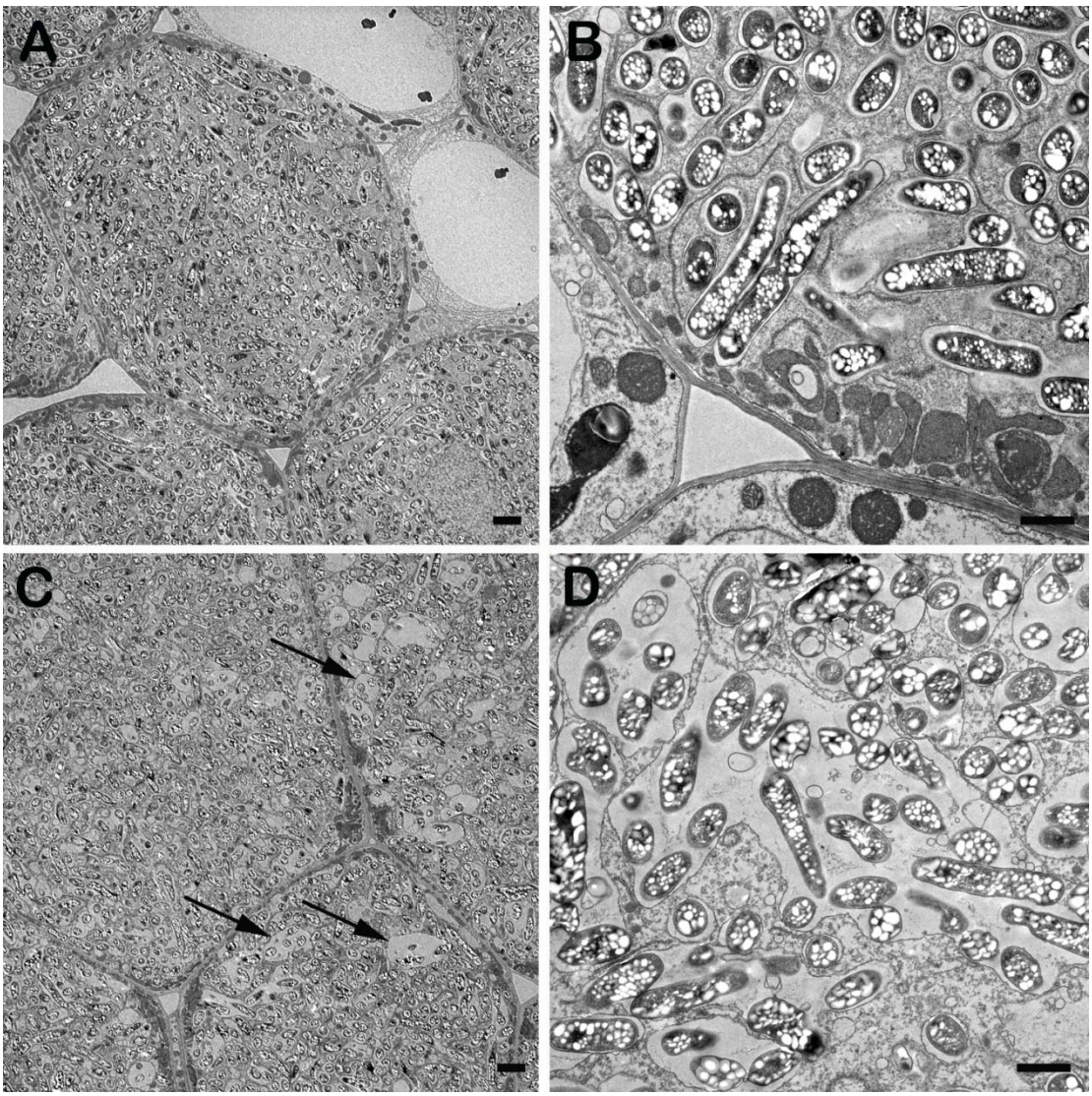


Figure 2.6

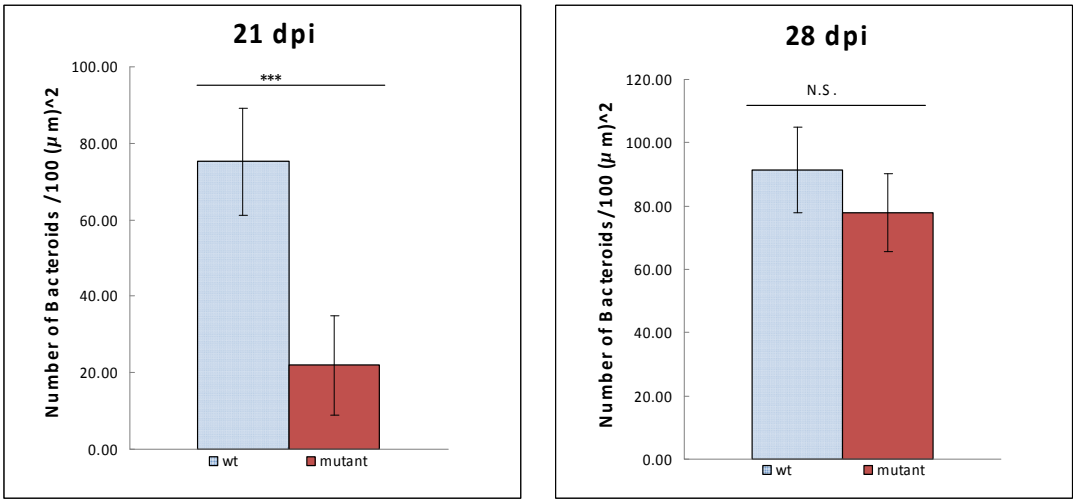

21 dpi
Number of Bacteroids /100 (μm)^2
100.00
80.00
60.00
40.00
20.00
0.00
***
wt
mutant
28 dpi
Number of Bacteroids/100 (μm)^2
120.00
100.00
80.00
60.00
40.00
20.00
0.00
N.S.
wt
mutant

Figure 2.7

CHAPTER 3

Characterization of the 3-*O*-deacylase gene, *pagL*, required for the removal of primary β-hydroxy fatty acid from the lipid A in the nitrogen-fixing endosymbiont *Rhizobium etli* CE3

Dusty B. Brown, Omar Salas, Kacie Speed, and Russell W. Carlson

Submitted to *The Journal of Biological Chemistry,* 10/14/2011.

Abstract

Until now, the gene responsible for the 3-*O*-deacylation of lipid A among nitrogen-fixing endosymbionts has not been characterized. Several gram negative animal pathogens such as *Salmonella enterica,* and *Pseudomonas aeroginosa* contain an outermembrane 3-*O*-deacylase (PagL) which has been implicated in host immune evasion. The role of 3-*O*-deacylated lipid A among nitrogen-fixing endosymbionts, plant endophytes and plant pathogens has not been studied. However, D'Haeze *et al.* (D'Haeze *et al.*. 2007. *J. Biol. Chem. 282*, 17101-17113) reported that the lipopolysaccharide isolated from the bean symbiont, *R. etli* CE3, contained exclusively tetraacylated lipid A that lacked a lipid A β-hydroxymyristyl residue; an observation that is consistent with the possibility of PagL activity being important in symbiosis. A putative *pagL* gene was identified in the *R. etli* genome sequence. With this information, we created a *pagL*$^-$ mutant strain derived from *R. etli* CE3. Using mass spectrometry, we demonstrated that the mutant lacks 3-*O*-deacylated lipid A. As determined by gas chromatography / mass spectrometry and gel electrophoresis of lipopolysaccharide (LPS), the mutant contained similar carbohydrate composition as parent strain. However, the mutant LPS contained larger amounts of β-hydroxypentadecanoic acid. Furthermore, relative to the parent strain, the mutant had increased sensitivity to detergent, displayed an increased resistance to the antimicrobial peptide polymyxin B, and was negatively affected in establishing symbiosis with its host, *Phaseolus vulgaris*.

Introduction

The gram negative endosymbiont *Rhizobium etli* forms nitrogen-fixing symbiosis with the legume host *Phaseolus vulgaris* (common black bean) and is a major endophyte of *Zea mays* (corn) in the agricultural soils of Mexico where the co-culture of bean and corn has been a common practice for thousands of years (1). In each case, the symbiotic/endophytic relationships promote plant growth. Nitrogen-fixing symbiosis is initiated through chemical signaling between the symbiont (bacteria) and legume host which leads to organogenesis of root nodules and the internalization of the bacteria. The invading symbionts travel through infection threads to the root nodule where they are endocytosed. After endocytosis, the bacteria develop into bacteroids within plant derived peribacteroid membranes and form nitrogen-fixing symbiosomes. For a current review on legume symbiosis, refer to Oldroyd *et al.* (2). The molecular basis for the endophytic relationship between *R. etli* and corn is less understood. *R. etli* isolates have been found to colonize the roots of corn in Mesoamerican areas and in many cases have been estimated to be the major microbial occupant (1). Colonization of corn roots by *R. etli* promotes plant growth independently of nitrogen fixation (1).

Surface and secreted bacterial polysaccharides are necessary for complete symbiosis (3,4). Lipopolysaccharide (LPS), a surface polysaccharide located in the outer membrane, is the focus of this study. Disruption of LPS biosynthesis leads to disruption of proper symbiosis in many nitrogen-fixing symbiotic systems (5,6) as well as disruption of endophytic relationships (7). For a detailed review on the role of LPS in symbiosis see Carlson *et al.* (5). The general structure of *R. etli* CE3 LPS has been extensively studied (5,8-10). D`Haeze *et al.* (8) determined that subtle structural changes occurred to the O-chain polysaccharide (OPS) of *R. etli* CE3 bacteroid LPS when compared to that of free-living CE3 whereby a single methyl group was

added to the 2-position of a fucosyl residue. In addition, the lipid A of CE3 bacteroid LPS was exclusively tetraacylated as compared to the free-living lipid A which contained a mixture of penta- and tetraacylated lipid A. It was speculated that tetraacylated lipid A resulted from the removal of a primary β-hydroxymyristoyl residue located at the 3-position of the proximal glucosamine (or 2-aminogluconoate) residue by a 3-*O*-deacylase. Indeed, Basu *et al.* (11) reported 3-*O*-deacylase activity in the membranes of a number of symbionts including *R. etli* CE3. The outer membrane 3-*O*-deacylase enzyme PagL has been reported in a number of gram negative pathogens including *Salmonella typhimurium, S. enterica, Pseudomonas aeruginosa*, and *Bordetella bronchiseptica* (12) and has been implicated in the evasion of the host immune response (13,14). It is possible that plant symbionts, endophytes, and pathogens modify their lipid A in a similar way to promote infection.

To understand the role of PagL concerning symbiosis, pathogenesis, and endophytic association with agriculturally and economically significant host plants, it is important to locate and characterize the gene that encodes PagL in these organisms. Although the membranes from *R. etli* CE3 demonstrated 3-*O*-deacylase activity *in vitro* (11), the isolated enzyme has not been reported and consequently the amino acid sequence and the gene encoding PagL is not known. Furthermore, *R. etli* PagL could not be identified by alignment with known PagL amino acid sequence due to low sequence identity. The crystal structure of the *P. aeruginosa* PagL was solved (15) and a conserved domain was elucidated which directly enabled us to identify a putative *pagL* gene in the *R. etli* CE3 genome (16). Therefore, we hypothesized that this gene is responsible for 3-*O*-deacylation of *R. etli* CE3 lipid A. In addition, we have identified *R. etli* CE3 PagL homologues in the genomes of a number of endosymbionts including, *Sinorhizobium meliloti, S. medicae, S. fredii, R. leguminosarum, R. etli, Methylobacterium nodulans,* as well as

the plant pathogens of the genus *Agrobacterium* and the poplar tree endophyte *Methylobacterium populi* (17) (S1).

In order to study the function of the putative *pagL* gene in the model organism *R. etli* CE3, we created a *pagL*$^-$ mutant via insertion mutagenesis. In this study, we characterized the *R. etli pagL* gene and determined its functional involvement in the 3-*O*-deacylation of lipid A. We further characterized the effect of the mutation on LPS synthesis, membrane stability, and symbiosis with the host plant *Phaseolus vulgaris.*

Experimental Procedures

Bacterial Strain, Plasmids and Growth Conditions—Bacterial strains and plasmids are described in Table 3.1 *Escherichia coli* strains were grown on Luria-Burtani (LB) media at 37 ^oC. *Rhizobium* strains were grown on tryptone yeast media containing 10 mM $CaCl_2$ or Y minimal media containing 2 % sucrose at 30 ^oC (18). Antibiotics were added in the following concentrations where indicated: *E. coli* (gentamicin [15 μg/ml], tetracycline [10 μg/ml], ampicillin [50 μg/ml]) and *R. etli* (gentamicin [30 μg/ml], tetracycline [20 μg/ml]).

Mutagenesis of the 3-O-deacylase gene—The putative 3-*O*-deacylase (*pagL*) gene from *Rhizobium etli* bv. *phaseoli* CE3 was mutated using insertion mutagenesis. PCR cloning was used to create a DNA fragment with the disrupted *pagL* gene and homolgous flanking regions to be cloned into a suicide plasmid. Conjugal transfer of the constructed plasmid into *R. etli* CE3 was performed and double recombinants were selected. Plasmid extractions, gel extractions and PCR/enzyme cleanup was performed using Quiagen mini prep kits. Restriction enzymes were

purchased from Promega, Inc. PCR reagent (iProof™ High-Fidelity polymerase) was purchased from BioRad.

A primer pair was engineered with *XbaI* (underlined) and *BamHI* (underlined) restriction sites (AGTCCTCAGACCGTTATTGCCGCTGT and AGTCGGATCCAGCTGTTGCTGCCAGC) in order to construct an approximately 1.1 kb PCR product containing 225 bp of the *pagL* gene (5` region) and 873 bp DNA upstream of the *pagL* gene. The PCR product was cloned into pUC18 to create plasmid pPagL-Up. A second PCR product was constructed using a primer pair engineered with *BamHI* (underlined) and *XmaI* (underlined) cloning sites to create an approximately 1.5 kb product containing 239 bp of the *pagL* gene (3` region) and 916 bp of downstream DNA. The product was cloned into pUC18 to create the plasmid pPagL-Dwn. The two inserts were tethered together by subcloning the insert from pPagL-Up into pPagL-Dwn (double digest with *XbaI* and *BamHI*) to create pPagL-UpDwn which contained an insert region with upstream DNA, *pagL* coding region (approximately 100 bp deletion near the center of the *pagL* gene), and downstream DNA. The center of the *pagL* region was engineered with a *BamHI* cloning site. The gentamicin resistant cassette *aacC1* from plasmid pMS255 (19) was cloned into the *pagL* region at the *BamHI* cloning site to create plasmid pPagL-UpGmDwn. The insert from pPagL-UpGmDwn was subcloned into the suicide vector pEX18-Tc (20) which contains the *sacB* lethal gene (double digest with *XbaI* and *XmaI*) to create pPagL-KO. The plasmid pPagL-KO was moved into *R. etli* CE3 via tri-parental mating as described previously (21). A *pagL*$^{-}$ mutant was obtained and was given the strain number *R. etli* ODB31. A construct containing the *pagL* gene was created in order to complement the mutant strain. The *pagL* gene was PCR cloned into plasmid pRK404E1(22) to create plasmid pRePagL using a forward primer engineered with a *BamHI* cloning site (underlined) and

homologous dna starting 14 bp upstream of the *pagL* gene (ATCGATGGATCCGCTTGGTGCGAATCAAAG) The reverse primer was engineered with an *EcoRI* cloning site (underlined) that recognized homologous DNA 30 bp downstream of the *pagL* coding region (ATCGATGAATTCGCAGGCTTTTCTCGTAAAG). The insert was oriented in parallel with the lac promoter of pRK404E1 to promote transcription. Plasmid pRePagL was moved into strain ODB31 via tri-parental mating as described above. All strains were verified by PCR with the above primers that were used to create pRePagL (Figure 3.1).

Lipopolysaccharide isolation and analysis—Briefly, lipopolysaccharide (LPS) was extracted by the hot water phenol extraction method (23) and dialyzed against deionized water using 12-14,000 MWCO dialyses bags and subsequently lyophilized. Composition of LPS was determined by gas chromatography in tandem with mass spectrometry (GC/MS) of trimethylsilyl (TMS) derived methyl glycosides and fatty acid methyl esters (FAME) as described previously (9,24) using a Hewlett Packard 5890 Series II GC/MS analyzer fitted with a 30 m DB-1 column (J & W Scientific).

Matrix assisted desorption ionization of lipid A—Lipid A was isolated by mild acid hydrolysis as described previously (25). Briefly, approximately 5 mg of extracted LPS was treated with 2% acetic acid (100 ^{o}C, 1 hr) to partition the lipid A. Partitioned Lipid A was separated by three times chloroform:methanol:water (2: 2: 1.8, v: v: v) extraction. The lipid A containing chloroform layers were pooled and dried under a stream of air. Isolated lipid A was dissolved in a 3: 1 (v: v) chloroform:methanol mixture, mixed 1: 1 (v: v) with 2,4,6-trihydroxyaceto-phenone THAP matrix and spotted on a MALDI plate. Spectra were acquired by a MALDI-TOF analyzer (Applied Biosystems, 5800) in the positive reflectron mode operating at a 20 kV extraction voltage.

DOC-PAGE and silver stain of lipopolysaccharide—Polyacrylamide gel electrophoresis in the presence of deoxycholate (DOC) of was adapted from Reuhs *et al.* (26). A loop full of bacteria grown on TY agar media was dissolved in 100 µl lysis buffer (1 % SDS, 50 mM Tris, and 10 mM EDTA) and incubated at 100 ^{o}C for 10 minutes to release LPS from the cell envelope. The sample was micro-centrifuged (14,000 rpm) to pellet cell debris. The supernatant (containing LPS) was mixed 1 to 1 with sample loading buffer and 2 µl was loaded for DOC-PAGE. The gel was silver stained as described previously (27).

Deoxycholic acid and Polymyxin B resistance assay—Polymyxin B gradient plates were made by pouring at an angle 25 ml TY agar media containing 50 µg/ml in a square 100 mm x 15 mm petri dish. The agar slants were placed level and 40 ml TY agar media was poured on top and allowed to dry. The Polymyxin B was allowed to diffuse throughout the plate overnight. A loop full of bacteria grown for four days on TY agar media was dissolved in 1 ml sterile phosphate buffered saline (pH 7.4). Then, 10 µl of bacterial suspension was spotted on the gradient plate and streaked from highest to lowest Polymyxin B concentration. The plates were incubated for four days at 30 ^{o}C. For bacteria isolated from host root nodules, 10 µl from a stock suspension was diluted in 90 µl of TY liquid and 10 µl of this dilution was streaked onto gradient plates. The stock suspension was prepared as follows: Nodules (28 days post inoculated, dpi, as described below) were surface sterilized in 2 % sodium hypochlorite and washed in excess sterile deionized water. The contents of the nodules were released into 150 µl TY liquid media by thorough crushing with forceps. 150 µl glycerol was added and the nodulate was stored at -80 ^{o}C.

Deoxycholic acid (DOC) gradient plates were made as described above using TY containing 2 mM DOC. Bacteria were prepared and streaked onto the gradient plates and incubated as described above. Each experiment was repeated more than three times.

Plant growth and inoculation—Organic black turtle beans (Lot # HP-BTB-04, Wheatgrasskits.com, Springville, Utah) were surfaced sterilized in 95% ethanol for 1 minute followed by 3 rinses in sterile deionized water. The beans were further sterilized in 5 % hypochlorite (2 minutes) followed by 10 rinses in sterile deionized water. Then, the beans were transferred to 0.8 % tap water agar plates and germinated in the dark for 3 days at room temperature. Sprouts were transferred to sterile 500 ml Erlenmeyer flasks containing a foam top, hydroponic beads immersed in defined, nitrogen free liquid plant media (Fahreus media (28)), and surrounded with a brown paper bag. Plants were kept in a Conviron growth chamber and were maintained at 24 ^{o}C with 14 h of light and 20 ^{o}C with 10 h of darkness. Bacteria used as inoculants were grown to mid/late log phase in liquid TY media containing appropriate antibiotics. The bacteria were pelleted, washed in sterile phosphate buffered saline (pH 7.4) and resuspended to an approximate O.D. of 0.15 (600 nm wave length). Then, 1 ml inoculate was added to the plants 11 days post germination. Plant phenotypes were observed and scored. The experiment was repeated 3 times. 30 plants per inoculate strain was used in one experiment and 15 plants per inoculate strain was used in subsequent experiments.

Results

Creation of the pagL^{-} *mutant in* Rhizobium etli *CE3*—In order to identify the gene responsible for 3-*O*-deacylase activity in *Rhizobium etli* and study its function, the putative *pagL* gene was mutated. The putative *pagL* gene (accession number RHE_CH02285) in *R. etli* CE3 (isogenic to *R. etli* CFN 42) is located on the chromosome and contains the conserved PagL catalytic domain (15,16) and predicted outer membrane protein structure. The *pagL*$^{-}$ mutant was created via

insertion mutagenesis as described in the experimental procedures section and was given the strain name *Rhizobium etli* ODB31. In order to validate phenotypic differences observed between the mutant and parent strain, the mutant ODB31 was complemented with the native *pagL* gene from parent strain CE3 (Figure 3.1).

Lipopolysaccharide composition of the pagL$^-$ *mutant*—In order to study differences between parent and mutant strain lipopolysaccharide (LPS), the composition of each LPS was determined. The LPS of *R. etli* CE3 has been studied extensively and its structure is known (5,8,29). When compared to parent strain CE3, the mutant ODB31 contained the known glycosyl residues found in the OPS (GlcA, QuiNAc, 3Me6dTal, Man, and Fuc), the core (Man, Gal, GalA, Kdo) and lipid A (GlcN, GalA). In addition, the DOC-PAGE profiles of parent and mutant LPS displayed similar migration patterns (Figure 3.2). These results suggest that the mutant maintained the general carbohydrate structure as the parent strain LPS. However, the primary hydroxy fatty acid portion of the LPS was variable between parent and mutant strains. The mutant LPS contained a significantly larger amount of β-hydroxypentadecanoate (β-OHC15:0). When the mutant was complemented with the native *pagL* gene, the amount of β-OHC15:0 decreased to parent strain levels. These results suggest that β-OHC15:0 is located on the lipid A at the site of PagL activity. There was no observable difference in the amount 27-OHC28:0 between parent and mutant LPS.

Mass spectroscopic analysis of the pagL$^-$ *mutant lipid A*—Mass spectroscopic analysis of mutant ODB31 lipid A was performed in order to characterize the loss of 3-*O*-deacylase activity on its structure. The spectra are shown in Figure 3.3 and proposed compositions for the various ions are listed in Table 3.2. The spectrum obtained from parent strain lipid A (Figure 3.3, panel I) contains two major ion groups each displaying a mixture of microheterogenous ions. Microheterogeneity occurs due to variation in acyl chain length (+/- 28 mass units [C_2H_4] or +/-

14 mass units [CH_2]), the presence of β-hydroxybutyryl group (+/- 86) at the C27 position of 27-hydroxyoctacosanoic acid, and the oxidation of the proximal glucosamine to 2-aminogluconoate (+/- 16). The higher molecular weight ion group, *m/z* 1914 - 2022, represents penta-acylated lipid A while the lower molecular weight ion group, *m/z* 1624 - 1774, represents tetra-acylated lipid A. Two forms of tetraacylated lipid A are present. One form is an artifact of the isolation process due to the acid catalyzed β-elimination of the 3-*O* linked primary fatty acid during the mild acid hydrolysis procedure. This occurs on lipid A species in which the proximal GlcNonate residue forms a lactone during mild acid hydrolysis of the LPS. The resulting structure(s) on MS analysis, structure(s) C, show ions of m/z 1738.7, 1710.6, 1652.5, and 1624.4 (Figure 3.3). The second tetraacylated lipid A is the product of PagL activity that results from the hydrolysis of the 3-*O* linked hydroxyl fatty acid by 3-*O*-deacylase (PagL) leaving a hydroxyl group at the C3 position of the proximal glucosamine/GlcNonate (Figure 3.3, structure(s) D and E). The resulting ions for D and E structures are m/z 1644.5, 1660.6, 1672.5, 1688.7, 1728.7, 1746.7, 1758.7, 1768.7, 1786.7 and 1774.7, 1796.7.

The parent strain CE3 produced a mixture of ions consistent with A, B, C, and D structures while ions due to E structures were low in intensity (Figure 3.3, panel I and Table 3.2.). The lipid A from the mutant ODB31 did not contain detectable ions consistent with structures D or E (Figure 3.3, panel II and Table 3.2). However, structures A, B, and C were present. The lipid A from the complemented mutant *pagL*$^-$ recovered the presence of structures D and also produced ions with m/z values consistent with E structures (Figure 3.3, panel III and Table 3.2). The intensities of ions due to A structures were lower for the lipid A from the complemented *pagL*$^-$ mutant strain compared to the other lipid A preparations. These results

strongly suggest that the *R. etli* CE3 *pagL* gene is responsible for the removal of a primary β-hydroxy fatty acid from the lipid A.

Escherichia coli *LPS is deacylated by the* Rhizobium etli *PagL enzyme*—The *Rhizobium* 3-*O*-deacylase, PagL, has less than 20% amino acid sequence similarity when compared to the PagL of enteric bacteria *Salmonella spp.* and *Pseudomonas aeruginosa*. Enteric bacteria share a similar general lipid A structure defined by a *bis*-phosphorylated diglucosamine backbone containing four β-hydroxymyristic primary fatty acids at the 2`, 3` and 2, 3 positions of the distal and proximal GlcN residues, respectively. The 2` and 3` primary fatty acids contain ester linked myristoyl and lauroyl subsitutents, respectively, at their β-OH groups creating the acyloxyacyl lipid A moieties (Figure 3.4, Structure A). A 3-*O*-deacylase has not been detected in *E. coli* and this is consistent with the fact that the *E. coli* genome does not contain a PagL homologue. Therefore, the *R. etli* PagL was introduced into the *E. coli* Top 10 cloning strain (*E. coli* Top 10/pRePagL) in order to determine if *R. etli* PagL can de-O-acylated *E. coli* lipid A.

Isolated *E. coli* LPS was subjected to the mild acid hydrolysis procedure as described in the experimental procedures section and subjected to MALDI-TOF MS analysis. The mass spectrum of the lipid A from the parent *E. coli* strain (Figure 3.4, panel I) contains three major ion groups. The *E. coli* lipid A structure and the source of the observed ions are as indicated and are consistent with the normal hexaacylated lipid A observed in *E. coli* (30). The lower molecular weight ions are a result of dephosphorylation (-80 mass units) and deacylation (-210 [C14] and -226 [- β-OHC14]); likely aritifacts that occur during the mild acid hydrolysis procedure used to isolate the lipid A. When the *pagL* gene is introduced into *E. coli,* a clear shift of -226 (OHC14:0) mass units occurs for all three groups of ions; e.g. 1796-226 = 1570, 1586-226 = 1360, 1360-226 = 1134, consistent with lipid A which lacks a β-OHC14:0 residue; Figure

3.4, panel II . Low intensity ions due to residual amounts of hexaacylated lipid A were also observed in the *E. coli* Top 10/pRePagL lipid A (1795 mass units). Taken together with the lipid A analysis of the mutant strain ODB31, these results suggest that the *R. etli* PagL has a rather loose substrate specificity in that it can act on these structurally diverse *E. coli* and *R. etli* lipid A species.

Sensitivity of the pagL$^-$ *mutant to deoxycholic acid (DOC) and polymyxin B (pmxB)*—The membrane stability of the *pagL*$^-$ mutant was assessed by determining its relative susceptibility to DOC and PmxB compared to the parent CE3 strain (Figure 3.5). Mutant ODB31 was more susceptible to increasing concentrations of DOC than parent strain (Figure 3.5A) suggesting some compromise to the outer membrane integrity. However, when subjected to increasing concentrations of PmxB (an outer membrane attacking antimicrobial peptide), mutant ODB31 displayed an increased resistance compared to the parent strain (Figure 3.5B). In addition, the relative resistance to DOC and PmxB reverted back to parent strain levels in the complemented strain ODB31/pRePagL (Figure 3.5A & 3.5B) suggesting that the phenotypic differences observed between mutant and parent strains was a result of the non-functional *pagL* gene.

The parent and *pagL*$^-$ mutant bacteria were isolated from root nodules and also assayed for their resistance to DOC and PmxB (Figure 3.5C & 3.5D). Mutant nodule bacteria recovered parental resistance to DOC compared to parent nodule bacteria (Figure 3.5C). However, the mutant nodule bacteria maintained a relatively increased resistance to PmxB when compared to parent nodule bacteria (Figure 3.5D).

The pagL$^-$ *mutant displays an aberrant symbiotic phenotype with the host plant* Phaseolus vulgarus—The ability of mutant ODB31 to form proper symbiosis with the host plant *Phaseolus vulgarus* was assayed by inoculating plants with bacteria as described in the experimental

procedures section. Between 10 and 14 dpi, even though nodules were present, the mutant inoculated plants showed yellowing leaves indicative of nitrogen starvation, while plants inoculated with the parent strain maintained green leaves as well as having root nodules. At 21 dpi, the leaves of the mutant inoculated plants maintained their yellow color (Figure 3.6) and while pink in color, the root nodules were considerably smaller compared to those found on plants inoculated with the parent strain (Figure 3.6).

Discussion

The lipid A 3-*O*-deacylase, PagL, has been characterized in a number of gram negative animal pathogens (12) and has been implicated in the evasion of the host immune response and resistance to antimicrobial peptides (13,14,31). It was determined *in vitro* that the membranes of several nitrogen-fixing endosymbionts contained 3-*O*-deacylase activity able to remove the 3-*O* linked β-hydroxymyristoyl moiety from the lipid IV_A precursor molecule (11). In addition, mass spectrometric evidence suggests that the endosymbionts *R. etli* CE3 (8) and *S. fredii* NGR 234 (32), when grown under laboratory conditions, produce 3-*O*-deacylated lipid A molecules. Furthermore, *R. etli* CE3 bacteroids exclusively produce tetraacylated lipid A, putatively due to 3-*O*-deacylation (8). Therefore, we hypothesized that *R. etli* CE3 produces a PagL enzyme which may play a role in symbiotic infection.

In this study we describe the preparation of a mutant that lacks a functional putative *pagL* gene in *R. etli* bv. *phaseoli* CE3. The *pagL*$^-$ mutant does not contain the tetraacylated lipid A, indicating that enzymatic de-*O*-acylation of lipid A has been disrupted. In addition, the mutant lipid A contains a larger amount of β-OHC15:0 than that of the parent strain. When the native

pagL gene is introduced into the mutant, the enzymatically tetraacylated lipid A structure was recovered and the amount of β-OHC15:0 decreased to parent strain levels. Based on these results, we conclude that the putative *pagL* gene described in this study is indeed responsible for the deacylation of lipid A and that when β-OHC15:0 is present it is located almost exclusively at the 3-position of the lipid A, *i.e.* the target of PagL activity.

The *Pseudomonas aeruginosa* PagL crystal structure has been solved (33) and a catalytic domain was proposed to contain a $HXSX_{(n+18)}E$ motif by which the histidine (H), serine (S), and glutamate (E) residues form a classical serine esterase catalytic triad. The *Rhizobium etli* PagL contains the HXS motif but instead of glutamic acid (E), a threonine (T) was observed (S1). The threonine residue is conserved in PagL homologues from organisms found in the *Rhizobiales* order (S1) and can theoretically perform a similar function as glutamate (*i. e.* stabilization of the deprotonated catalytic histidine by hydrogen bonding). In addition, the *R. etli* PagL contained the conserved phenylalanine (F) hydrophobic residue (S1) previously implicated in positioning the lipid A molecule for hydrolysis by the *P. aeroginosa* PagL (12). However, a lipid A binding motif has not been fully characterized for PagL enzymes. The substrate specificity for the *R. etli* PagL enzyme extends to *E. coli* lipid A (Figure 3.4). *E. coli* and *R. etli* lipid A structures are quite different and our results suggest that the *R. etli* PagL has a broad substrate specificity that may extend to a variety of gram negative lipid A structures.

The *pagL*$^{-}$ mutant grown under laboratory conditions was more susceptible to the detergent deoxycholic acid (DOC) than the parent strain indicating a compromise to membrane stability (Figure 3.5). These results were somewhat surprising since one would speculate that removal of fatty acid from lipid A, as in the case of the parent and complemented mutant strains, would render the lipid A less hydrophobic thereby decreasing hydrophobic interactions that

contribute to overall membrane stability. However, the parent and complemented mutant strains (CE3 and ODB31/pRePagL), which produce increased amounts of tetraacylated lipid A, were more resistant to DOC. Perhaps disruption of *pagL* expression affects, as yet, uncharacterized factors that cause the relative decrease in membrane stability resulting in the observed increased DOC sensitivity. Interestingly, passage of the mutant through the plant restored parental levels of resistance to DOC (Figure 3.5). Membrane stability toward possible host plant factors or conditions is, likely, an important factor in this endosymbiotic relationship and the plant has apparently selected a variant of the pagL$^-$ mutant with increased membrane stability.

The mutant grown under laboratory conditions displayed relatively increased resistance to the polycationic antimicrobial peptide (AMP) polymyxin B (PmxB). Upon passage through the plant, the mutant derived nodule bacteria maintained increased resistance to PmxB (Figure 3.5). D'Haeze *et al.* (8) reported that isolated *R. etli* CE3 bacteroids, which exclusively contain tetraacylated lipid A, were very sensitive to AMPs; a result which is consistent with the increased PmxB resistance of the *pagL*$^-$ mutant that does not form tetraacylated lipid A. In contrast to this result, it has been shown that 3-*O*-deacylated lipid A leads to increased PmxB resistance in *Salmonella enterica* (14). Thus, our results for the *R. etli* CE3 *pagL*$^-$ mutant are counterintuitive, particularly given that the interaction between PmxB and LPS has a hydrophobic component and one would speculate that the removal of fatty acid by PagL may lessen the interaction between PmxB and lipid A as suggested for the *pagL*$^-$ mutant.

In mammalian systems, LPS is recognized by the toll like receptor 4 (TLR-4) which in turn induces a proinflammatory response (34,35). The TLR-4 ectodomain contains leucine rich repeats (LRR), a common feature of the TLR receptor family responsible for the perception of a wide variety of conserved microbial associated molecular patterns (MAMPs). The modification

or synthesis of unique lipid A structures has been implicated in the evasion of the animal LPS recognition system (36-39). Although it is unknown how LPS is recognized in plants, there are numerous transmembrane pattern recognition proteins (PRRs) containing an LRR ectodomain and an intracellular receptor like kinase domain (LRR-RLK proteins) that are involved in plant innate immunity and cell signaling (40). Some of these have been shown to recognize two highly conserved MAMPs; flagellin (flg22) and elongation factor (EF-Tu) (41). Furthermore, LRR-RLKs are crucial for bacterial infection and nodule organogenesis in legume symbiosis (42-44). It is possible that plant-associated bacteria modulate their lipid A similar to mammalian-associated bacteria in order to modulate plant signaling through LPS – LRR-RLK interactions. Elucidating the LPS receptor/s in plants as well as possible LPS associated AMPs/NCRs will aid in our higher understanding of the role LPS has in symbiotic, pathogenic, and endophytic relationships.

It is important to note that the *pagL* gene is not present or expressed in all symbionts and is not absolutely required for complete symbiosis in these systems. Recent work by in our laboratory (21) showed that the bean nodulating strain *R. leguminosarum* bv. *phaseoli* 8002 does not produce a deacylated lipid A and likely does not contain the *pagL* gene. The genome of the closely related organism *R. leguminosarum* bv. *viciae* 3841 has been published (the genome of strain 8002 is not published) and does not contain the *pagL* gene (S2). However, strain 3841 contains remnants of the *pagL* gene suggesting that these organisms may have lost the *pagL* gene possibly through the loss of selective pressure in their natural environment or in the laboratory. Another closely related clover-nodulating strain, *R. leguminosarum* bv. *trifolii* WSM 1325 does contain the *pagL* gene. It is not known whether or not strain WSM 1325 produces 3-*O*-deacylated lipid A. As stated above, *S. meliloti* contains both deacylated lipid A and the *pagL*

gene. The absence of the *pagL* gene from some endosymbionts suggests that *pagL* is not absolutely required for complete symbiosis. However, the results presented in this study supports the conclusion that *R. etli* CE3 requires *pagL* expression for normal symbiosis (Figure 3.6). As mentioned in the introduction, in addition to nitrogen-fixing symbiosis with bean, *R. etli* strains are also recognized for their endophytic relationship with corn. It is tempting to speculate that expression of *pagL* may also play a role in this endophytic relationship. Information provided in this manuscript provides a means in which to study the role of PagL in a variety of symbiotic, endophytic, and pathogenic relationships.

Acknowledgements

**We acknowledge Stefan Eberhard of the CCRC for preparing the plant photographs displayed in this work.*

References

1. Gutiérrez-Zamora, M. L., and Martínez-Romero, E. (2001) *J. Biotechnol.* **91**, 117-126
2. Oldroyd, G. E. D., and Downie, J. A. (2008) *Annu. Rev. Plant Biol.* **59**, 519-546
3. Kannenberg, E. L., and Brewin, N. J. (1994) *Trends Microbiol.* **2**, 277-283
4. Nicolas, F., François, C., and Véréna, P. (2003) *Euro. J. Biochem.* **270**, 1365-1380
5. Carlson, R. W., Forsberg, L. S., and Kannenberg, E. L. (2010) Lipopolysaccharides in Rhizobium-Legume Symbioses. in *Subcellular Biochemestry: Endotoxins: Structure, Function and Recognition*, P. J. Quinn and X. Wang Ed., Springer. pp 339-386
6. Kannenberg, E. L., Reuhs, B. L., Forsberg, L. S., and Carlson, R. W. (1998) Lipopolysaccharides and K-Antigens: Their Structures, Biosynthesis, and Functions. in *The Rhizobiaceae: molecular biology of model plant-associated bacteria* (Spaink, H. P., Kondorosi, A., and Hooykaas, P. J. J. eds.), Kluwer Academic Publishers, Dordrecht, The Netherlands. pp 119-154
7. Ormeño-Orrillo, E., Rosenblueth, M., Luyten, E., Vanderleyden, J., and Martínez-Romero, E. (2008) *Environ. Microbiol.* **10**, 1271-1284
8. D'Haeze, W., Leoff, C., Freshour, G., Noel, K. D., and Carlson, R. W. (2007) *J. Biol. Chem.* **282**, 17101-17113
9. Bhat, U. R., Forsberg, L. S., and Carlson, R. W. (1994) *Journal of Biological Chemistry* **269**, 14402-14410
10. Forsberg, L. S., and Carlson, R. W. (1998) *J. Biol. Chem.* **273**, 2747-2757
11. Basu, S. S., White, K. A., Que, N. L. S., and Raetz, C. R. H. (1999) *J. Biol. Chem.* **274**, 11150-11158

12. Trent, M. S., Stead, C. M., Tran, A. X., and Hankins, J. V. (2006) *J. Endotoxin Res.* **12**, 205-223

13. Kawasaki, K., Ernst, R. K., and Miller, S. I. (2004) *J. Biol. Chem.* **279**, 20044-20048

14. Kawasaki, K., China, K., and Nishijima, M. (2007) *J. Bacteriol.* **189**, 4911-4919

15. Rutten, L., Geurtsen, J., Lambert, W., Smolenaers, J. J. M., Bonvin, A. M., de Haan, A., van der Ley, P., Egmond, M. R., Gros, P., and Tommassen, J. (2006) *Proc. Natl. Acad. Sci.* **103**, 7071-7076

16. Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., Fong, J. H., Geer, L. Y., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Jackson, J. D., Ke, Z., Lanczycki, C. J., Lu, F., Marchler, G. H., Mullokandov, M., Omelchenko, M. V., Robertson, C. L., Song, J. S., Thanki, N., Yamashita, R. A., Zhang, D., Zhang, N., Zheng, C., and Bryant, S. H. *Nucl. Acids Res.*

17. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucl. Acids Res.* **25**, 3389-3402

18. Beringer, J. E. (1974) *Microbiol.* **84**, 188

19. Becker, A., Schmidt, M., Jager, W., and Puhler, A. (1995) *Gene* **162**, 37-39

20. Hoang, T. T., Karkhoff-Schweizer, R. A. R., Kutchma, A. J., and Schweizer, H. P. (1998) *Gene* **212**, 77-86

21. Brown, D. B., Huang, Y.-C., Kannenberg, E. L., Sherrier, D. J., and Carlson, R. W. (2011) *J. Bacteriol.*, JB.00392-00311

22. Lehman, L. J., Fitzmaurice, W. P., and Roberts, G. P. (1990) *Gene* **95**, 143-147

23. Westphal, O., and Jann, K. (1965) *Meth. Carbohydr. Chem.* **5**, 83–91

24. York, W. S., Darvill, A. G., McNeil, M., Stevenson, T. T., and Albersheim, P. (1986) *Meth. Enymol.* **118**, 3-40

25. Ryan, J. M., Conrad, H. E. (1974) *Arch. Biochem. Biophys.* **162**, 530-536

26. Reuhs, B. L., Carlson, R. W., and Kim, J. S. (1993) *J. Bacteriol.* **175**, 3570-3580

27. Kittelberger, R., and Hilbink, F. (1993) *J. Biol. Chem. Biophys. Meth.* **26**, 81-86

28. Fahraeus, G. (1957) *J Gen Microbiol* **16**, 374-381

29. Forsberg, L. S., Bhat, U. R., and Carlson, R. W. (2000) *J. Biol. Chem.* **275**, 18851-18863

30. Raetz, C. R. H., Reynolds, C. M., Trent, M. S., and Bishop, R. E. (2007) *Ann. Rev. Biochem.* **76**, 295-329

31. Guo, L., Lim, K. B., Gunn, J. S., Bainbridge, B., Darveau, R. P., Hackett, M., and I., M. S. (1997) *Science* **276**, 250-253

32. Gudlavalleti, S. K., and Forsberg, L. S. (2003) *J. Biol. Chem.* **278**, 3957-3968

33. Rutten, L., Geurtsen, J., Lambert, W., Smolenaers, J. J. M., Bonvin, A. M., De Haan, A., Van der Ley, P., Egmond, M. R., Gros, P., and Tommassen, J. (2006) *Proc. Natl. Acad. Sci.* **103**, 7071-7076

34. Park, B. S., Song, D. H., Kim, H. M., Choi, B.-S., Lee, H., and Lee, J.-O. (2009) *Nature* **458**, 1191-1195

35. Akira, S., and Takeda, K. (2004) *Nat. Rev. Immunol.* **4**, 499-511

36. Trent, M. S., Stead, C. M., Tran, A. X., and Hankins, J. V. (2006) *J. Endotoxin Res.* **12**, 205-223

37. Berezow, A. B., Ernst, R. K., Coats, S. R., Braham, P. H., Karimi-Naser, L. M., and Darveau, R. P. (2009) *Microb. Pathog.* **47**, 68-77

38. Coats, S. R., Berezow, A. B., To, T. T., Jain, S., Bainbridge, B. W., Banani, K. P., and Darveau, R. P. (2011) *Infect. Immun.* **79**, 203-210

39. Kanistanon, D., Hajjar, A. M., Pelletier, M. R., Gallagher, L. A., Kalhorn, T., Shaffer, S. A., Goodlett, D. R., Rohmer, L., Brittnacher, M. J., Skerrett, S. J., and Ernst, R. K. (2008) *PLoS Pathog.* **4**, e24

40. Erbs, G., Molinaro, A., Dow, J. M., and Newman, M. A. (2010) Lipopolysaccharides and Plant Innate ImmunityEndotoxins: Structure, Function and Recognition. (Wang, X., and Quinn, P. J. eds.), Springer Netherlands. pp 387-403

41. Boller, T., and Felix, G. (2009) *Annu. Rev. Plant Biol.* **60**, 379-406

42. Spaink, H. P. (2004) *Trends Microbiol.* **12**, 201-204

43. Limpens, E., Mirabella, R., Fedorova, E., Franken, C., Franssen, H., Bisseling, T., and Geurts, R. (2005) *Proc. Natl. Acad. Sci.* **102**, 10375-10380

44. Lefebvre, B., Timmers, T., Mbengue, M., Moreau, S., Herve, C., Toth, K., Bittencourt-Silvestre, J., Klaus, D., Deslandes, L., Godiard, L., Murray, J. D., Udvardi, M. K., Raffaele, S., Mongrand, S., Cullimore, J., Gamas, P., Niebel, A., and Ott, T. (2010) *Proc. Natl. Acad. Sci.* **107**, 2343-2348

45. Noel, K. D., Sanchez, A., Fernandez, L., Leemans, J., and Cevallos, M. A. (1984) *J. Bacteriol.* **158**, 148-155

46. Phadnis, S. H., and Das, H. K. (1987) *J. Biosci.* **12**, 131-135

Footnotes

This work was supported, in whole or in part, by National Institutes of Health Grant GM39583 (to R.W.C.) and by the Department of Energy Grant DE-FG-02-09ER20097 to the CCRC. This work is in partial fulfillment for the PhD degree of D.B.B.

Abbreviations

Abbreviations used in this study are: LPS, lipopolysaccharide; Kdo, 3-deoxy-*D*-*manno*-2 octulosonic acid; 3Me6dTal, 3-O-methyl-6-deoxytalose; QuiNAc, N-acetylquinovosamine; GlcN, glucosamine; Gal, galactose; Fuc, fucose; GalA, galacturonic acid; Man, mannose; Glc, glucose; GlcNonate, 2-aminogluconate; BHB, β-hydroxybutyrate; β-OHC14:0, β-hydroxymyristate; β-OHC16:0, β-hydroxypalmitate; β-OHC18:0, β-hydroxystearate; dod-P-GalA, dodecaprenyl phosphogalacturonic acid; MALDI-TOF-MS, matrix assisted laser desorption ionization time of flight mass spectrometry; GC/MS, gas chromatography in line with mass spectrometry; DOC-PAGE, deoxycholic acid polyacrylamide gel electrophoresis; PmxB, polymyxin B: AMP, polycationic antimicrobial peptide

Figure Legends

Figure 3.1. PCR analysis of the *pagL*$^-$ mutant and complemented strains. Lane 1) DNA ladder. Lane 2) Parent strain CE3. Lane 3) Mutant strain ODB31. Lane 4) Complemented mutant strain. The predicted size of the native PCR product (Lane 2) is 635 base pairs (bp). The predicted size of the mutant PCR product (Lane 3) containing the *aacC1* insert is 1531 bp. The mutant complemented strain (Lane 4) contains both PCR native and mutated PCR product.

Figure 3.2. DOC-PAGE analysis of mutant LPS. Lane 1) Parent strain CE3. Lane 2) Mutant strain ODB31. Lane 3) Complemented mutant strain. LPS-I, high molecular weight LPS species containg lipid A, Core, and O-chain. LPS-II, low molecular weight LPS species containg lipid A and Core.

Figure 3.3. MALDI-TOF mass spectrometric analysis of mutant lipid A. Spectra were obtained in the negative reflectron mode. I. Parent strain CE3. II. Mutant strain ODB31. III. Complemented mutant strain. Peaks are labeled with the predicted structures A-D. The general structures of A-D vary due to presence or absence of BHB (+/- 86 mass units) and variation in acyl chain length (+/- 28 or 14 mass units).

Figure 3.4. MALDI-TOF MS analysis of *E. coli* / pRePagL lipid A. Spectra were obtained in the negative reflectron mode. I.) Parent strain *E. coli* Top 10. II.) *E. coli* Top 10 containing plasmid pRePagL. The structures are labeled A or B (de-O-acylated). Microheterogeneity

occurs as a result of deacylation and dephosphorylation of *E. coli* lipid A during the mild acid hydrolysis procedure and variation in acyl chain length (+/- 28 mass units). * These structures are present due to the chemical deacylation of lipid A during mild acid hydrolysis of the LPS. *E. coli* does not normally contain 3-*O*-deacylase activity.

Figure 3.5. The susceptibility of the *pagL*⁻ mutant to DOCand PmxB. **Panel A** – Growth of bacteria on DOC gradient plates. **Panel B** – Growth of bacteria on PmxB gradient plates **Panel C –** Growth of nodule bacteria on DOC gradient plates. **Panel D –** Growth of nodule bacteria on PmxB gradient plates. Lanes 1 and 1a) Parent strain CE3. Lanes 2 and 2a) Mutant strain ODB31. Lanes 3 and 31) Complemented mutant strain

Figure 3.6. Plant phenotype of mutant inoculated *Phaseolus vulgarus* (black bean). Pictures of representative plants were taken 21 dpi. **Panel A**. 1) Top view of parent strain CE3 inoculated plants. 2) Top view of mutant ODB31 inoculated plants. **Panel B**. 1) A representative picture of root nodules from parent strain CE3 inoculated plant. 2) A representative picture of root nodules from mutant strain ODB31 inoculated plant.

Table 3.1 Strains and Plasmids

Strains or Plasmid	Characteristics	Source
E. coli		
Top 10	F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*araleu*) 7697 *gal*U *gal*K *rps*L (Strr) *end*A1 *nup*G	Invitrogen
Top 10/pRePagL	*E. coli* Top 10 containing plasmid pRePagL, Tc^r	This Study
R. etli* bv. *phaseoli		
CE3	Fix^+, Nod^+, Str^r	(45)
ODB31	*pagL::aacC1,* Fix^+, Nod^+, Str^r, Gm^r	This Study
ODB31/pRePagL	strain ODB31 harboring plasmid pRePagL, Fix^+, Nod^+, Str^r, Gm^r, Tc^r	This Study
Plasmids		
pUC18	Cloning vector, Am^r	Fermentas
pMS255	Source of *Gm* cassette, Gm^r	(19)
pRK2013	Mobilizing plasmid for pEX18-Tc, Col E1 replicon, Kan^r	(46)
pEX18-Tc	Suicide vector, allows positive selection for integration, Tc^r	(20)
pRK404E1	Broad host range shuttle vector for, Tc^r	(22)
pPagL-Up	pUC18 containing the upstream region and 5` end of *pagL*, Am^r	This Study
pPagL-Dwn	pUC18 containing the downstream region and 3` end of *pagL*, Am^r	This Study
pPagL-UpDwn	pUC18 containing the inserts from pPagL-Up and pPagL Dwn, (*pagLΔ100bp*), Am^r	This Study
pPagL-UpGmDwn	*Gm* cassette (*aacC1*) from pMS255 cloned into pPagL-UpDwn, Am^r, Gm^r	This Study
pRePagL-KO	Insert from pPagL-UpGmDwn cloned into pEX18-Tc, Tc^r, Gm^r	This Study
pRePagL	*pagL* region from strain CE3 PCR cloned into pRK404E1, Tc^r	This Study

Table 3.2. Lipid A ions ($[M-H]^-$)† observed in the MALDI-TOF-MS spectra and their predicted compositions

Obs.	Calc.	Type*	Proposed Compositions	CE3	OBD31
1984.8	1984.4	A*	$GalA_1GlcN_2$βOHC14:0_3βOHC18:0_1BHBC28:0_1	+	+
2000.8	2000.4	B*	$GalA_1GlcN_1GlcNonate_1$βOHC14:0_3βOHC18:0_1BHBC28:0_1	+	+
1738.7	1738.2	C*	$GalA_1GlcN_1$2,3-unsaturated GlcN(acyl)onolactone $_1$βOHC14:0_2βOHC18:0_1BHBC28:0_1	+	+
1758.7	1758.2	D*	$GalA_1GlcN_2$βOHC14:0_2βOHC18:0_1BHBC28:0_1	+	-
1774.7	1774.2	E*	$GalA_1GlcN_1GlcNonate_1$βOHC14:0_2βOHC18:0_1BHBC28:0_1	+	-
1956.7	1956.4	A	A - 28 (C_2H_4)	+	+
1936.9	1936.4	B	B - 86(BHB) + 22 (Na); i.e. an $[MNa-H]^-$ ion.	+	+
1926.7	1926.8	A	A - 86(BHB) + 28 (C_2H_2)	+	+
1914.9	19.14.4	B	B - 86 (BHB)	+	+
1886.8	1886.3	B	B - 86 (BHB) - 28 (C_2H_4)	+	+
1802.7	1802.2	E	E + 28 (C_2H_4)	+	-
1796.7	1797.2	E	E +22 (Na)	+	-
1786.7	1786.3	D	D +28 (C_2H_4)	+	-
1768.6	1769.2	E	E - 28 (C_2H_4) + 22 (Na)	+	-
1746.7	1746.2	E	E - 28 (C_2H_4)	+	-
1730.5	1730.3	D	D - 28 (C_2H_4)	+	-
1728.7	1728.2	D	D - 86 (BHB) + (2) 28 (C_2H_4)	+	-
1710.6	1710.2	C	C - 28 (C_2H_4)	+	+
1700.5	1700.2	D	D - 86 (BHB) + 28 (C_2H_4)	+	-
1672.5	1672.2	D	D - 86 (BHB)	+	-
1688.7	1688.2	E	E - 86 (BHB)	+	-
1660.6	1660.2	E	E - 86 (BHB) - 28 (C_2H_4)	+	-
1652.5	1652.4	C	C - 86 (BHB)	+	+
1644.5	1644.2	D	D - 86 (BHB) – 28 (C_2H_4)	+	-
1624.4	1624.1	C	C - 86 (BHB) - (3) 28 (C_2H_4)	+	+

†Calibration was done using exact mass values, and the calculated values are exact mass values.

*****Structure types A, B, C, D, and E are shown in Figure 3.3. Structure types D and E are those that result from PagL activity. Structure type C results from acid catalyzed β-elimination of the 3-O-acyl residue from the proximal GlcNonolactone during mild acid hydrolysis of the LPS which converts that residue into 2,3-unsaturated GlcN(acyl)onolactone. +/- Ions observed (+) or absent (-) from the spectrum of the indicated strain.

Supplementary Figure Legends.

S1. Amino acid sequence alignment of the *R. etli* PagL polypeptide and homologues in various endosymbionts and endophytes. The *Pseudomonas aeruginosa* PagL was added as a reference point for the catalytic domain. The strain numbers are given. Re, *Rhiobium etli*; Rl, *Rhizobium leguminosarum*; Smel, *Sinorhizobium meliloti*; Smed, *Sinorhizobium medicae*; Sfred, *Sinorhizobium fredii*; At, *Agrobacterium tumefaciens*; Mp, *Methylobacterium populi*; Mn, *Methylobacterium nodulans*. Shaded amino acids are identical. Boxed regions are predicted catalytic amino acids. Boxed areas are putative catalytic residues. * Indicates a conserved phenylalanine (F) residue implicated in positioning the lipid A for hydrolysis (15).

S2. DNA sequence alignment of the *R. etli* CE3, *R. leguminosarum* WSM1025, and *R. leguminosarum* 3841 *pagL* gene region. * Indicates the start and stop codons in the open reading frame of the template strand.

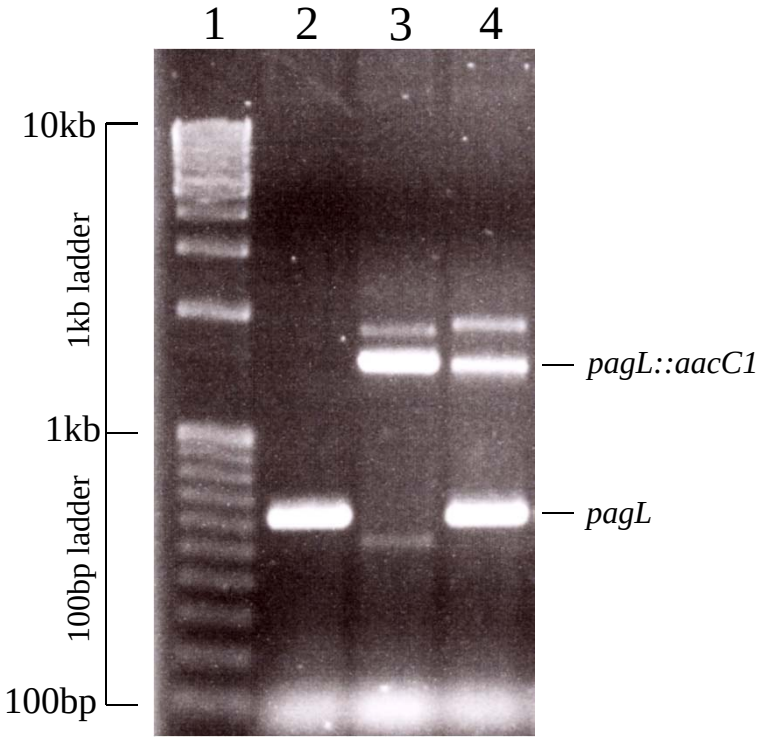


Figure 3.1

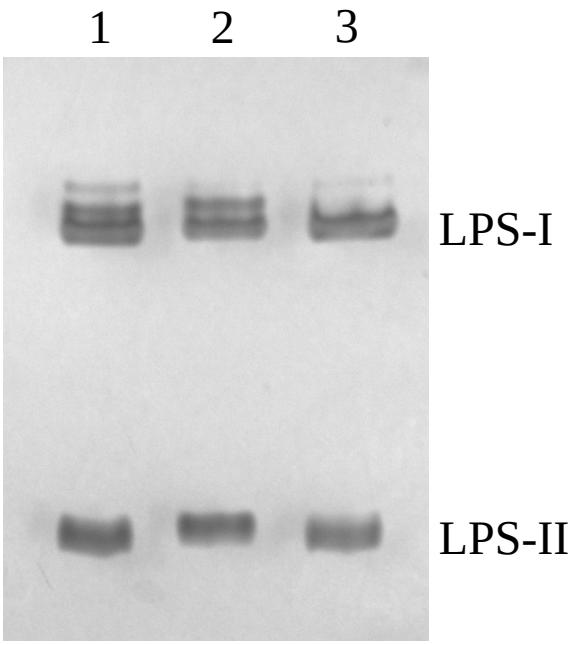


Figure 3.2

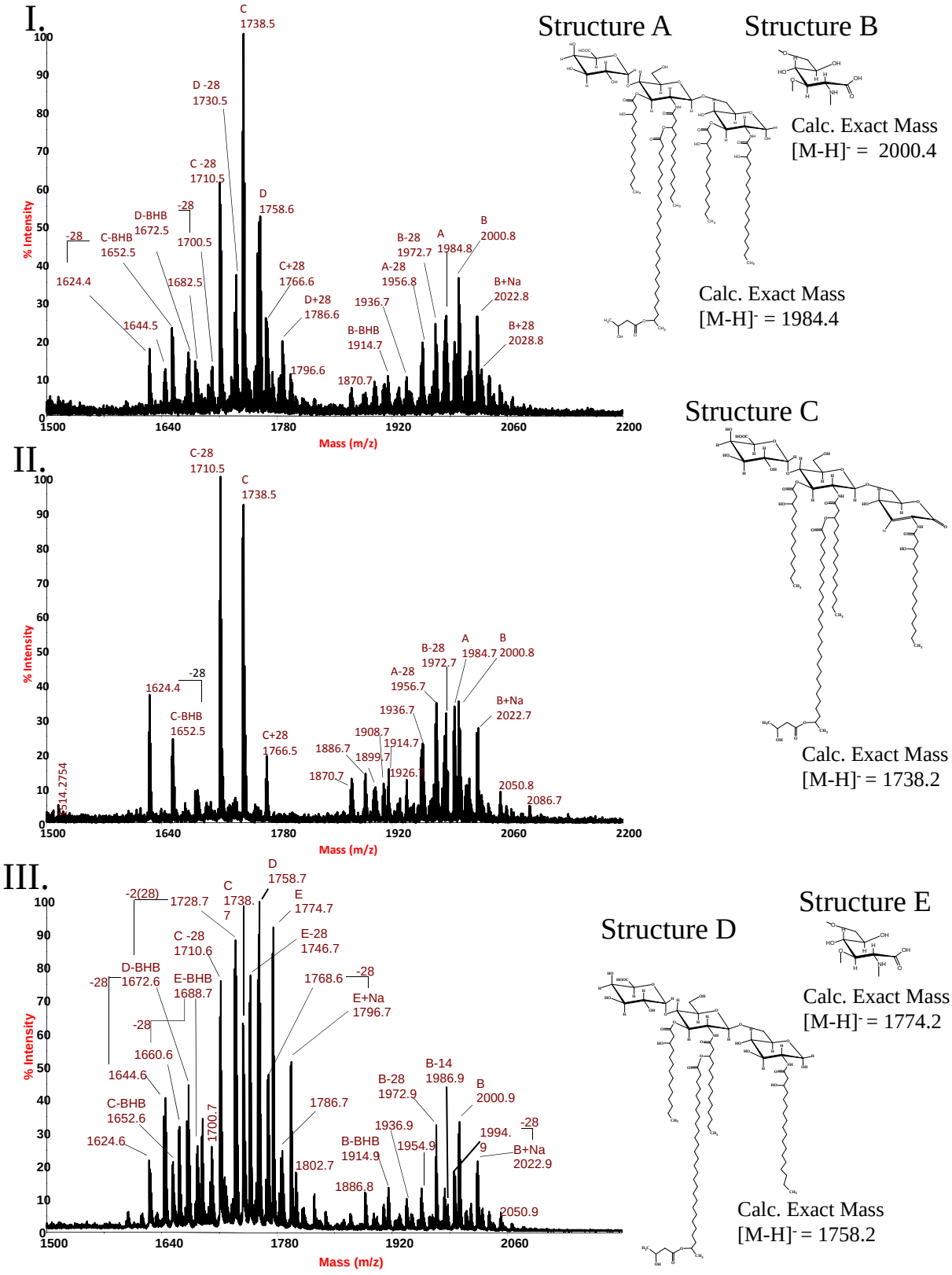
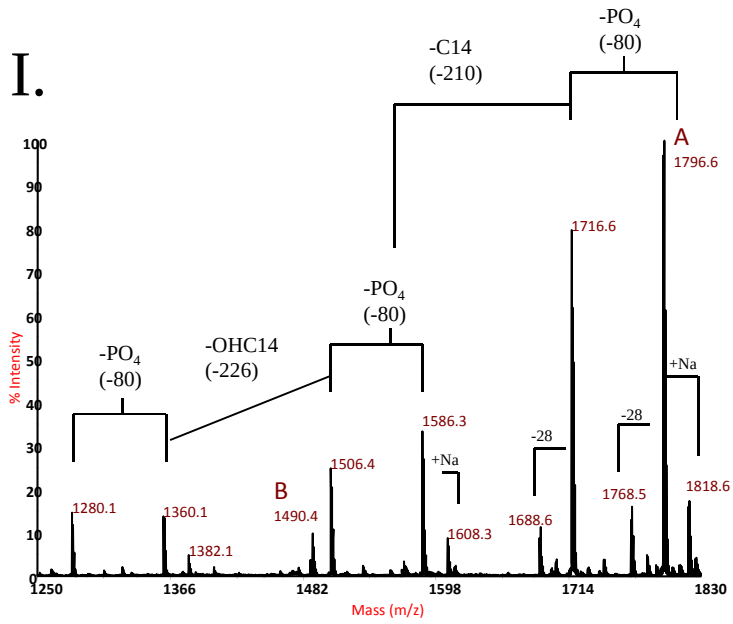
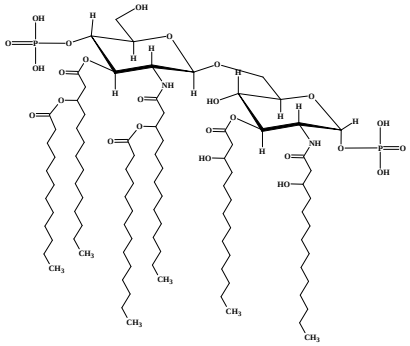

I.
% Intensity
Mass (m/z)
C 1738.5
D -28 1730.5
C -28 1710.5
D 1758.6
-28
D-BHB 1672.5
C-BHB 1652.5
1700.5
1624.4
1682.5
1644.5
C+28 1766.6
D+28 1786.6
1796.6
B-28 1972.7
A 1984.8
B 2000.8
A-28 1956.8
B+Na 2022.8
1936.7
B-BHB 1914.7
B+28 2028.8
1870.7
Structure A
Structure B
Calc. Exact Mass
[M-H]- = 2000.4
Calc. Exact Mass
[M-H]- = 1984.4
II.
% Intensity
Mass (m/z)
C-28 1710.5
C 1738.5
-28
1624.4
C-BHB 1652.5
C+28 1766.5
B-28 1972.7
A 1984.7
B 2000.8
A-28 1956.7
B+Na 2022.7
1936.7
1908.7
1914.7
1886.7
1899.7
1870.7
1926.7
2050.8
2086.7
514.2754
Structure C
Calc. Exact Mass
[M-H]- = 1738.2
III.
% Intensity
Mass (m/z)
D 1758.7
C 1738.7
E 1774.7
-2(28)
1728.7
C -28 1710.6
E-28 1746.7
D-BHB 1672.6
E-BHB 1688.7
-28
1768.6
E+Na 1796.7
1660.6
1644.6
C-BHB 1652.6
1624.6
1700.7
1786.7
B-14 1986.9
B-28 1972.9
B 2000.9
1936.9
1994.9
-28
B-BHB 1914.9
1954.9
B+Na 2022.9
1802.7
1886.8
2050.9
Structure D
Structure E
Calc. Exact Mass
[M-H]- = 1774.2
Calc. Exact Mass
[M-H]- = 1758.2

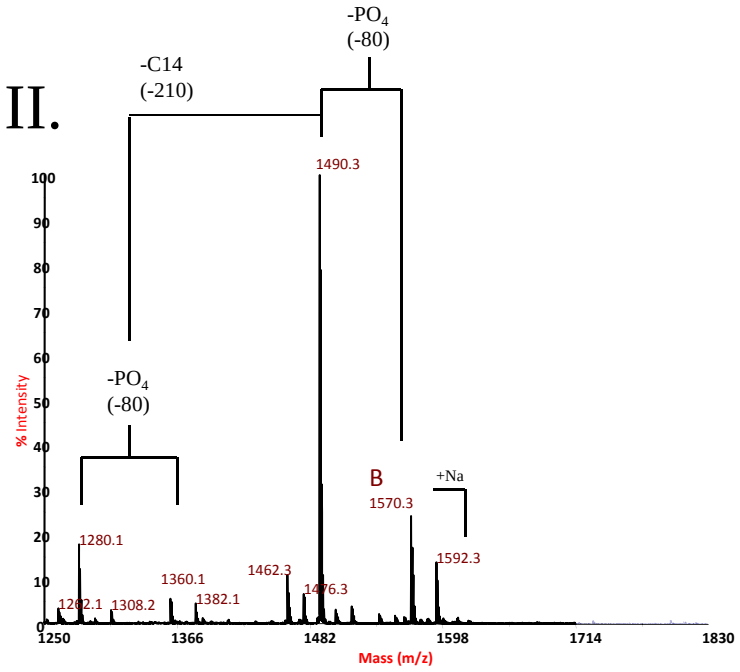
Figure 3.3



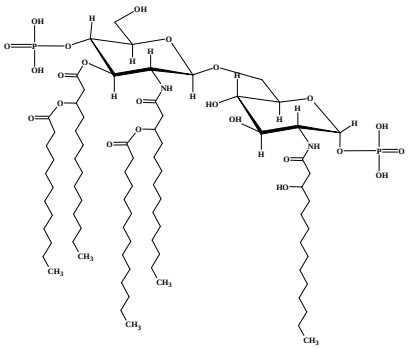
Structure A



Calc. Exact Mass
M-H = 1796.2



Structure B



Calc. Exact Mass
M-H = 1570.0

Figure 3.4

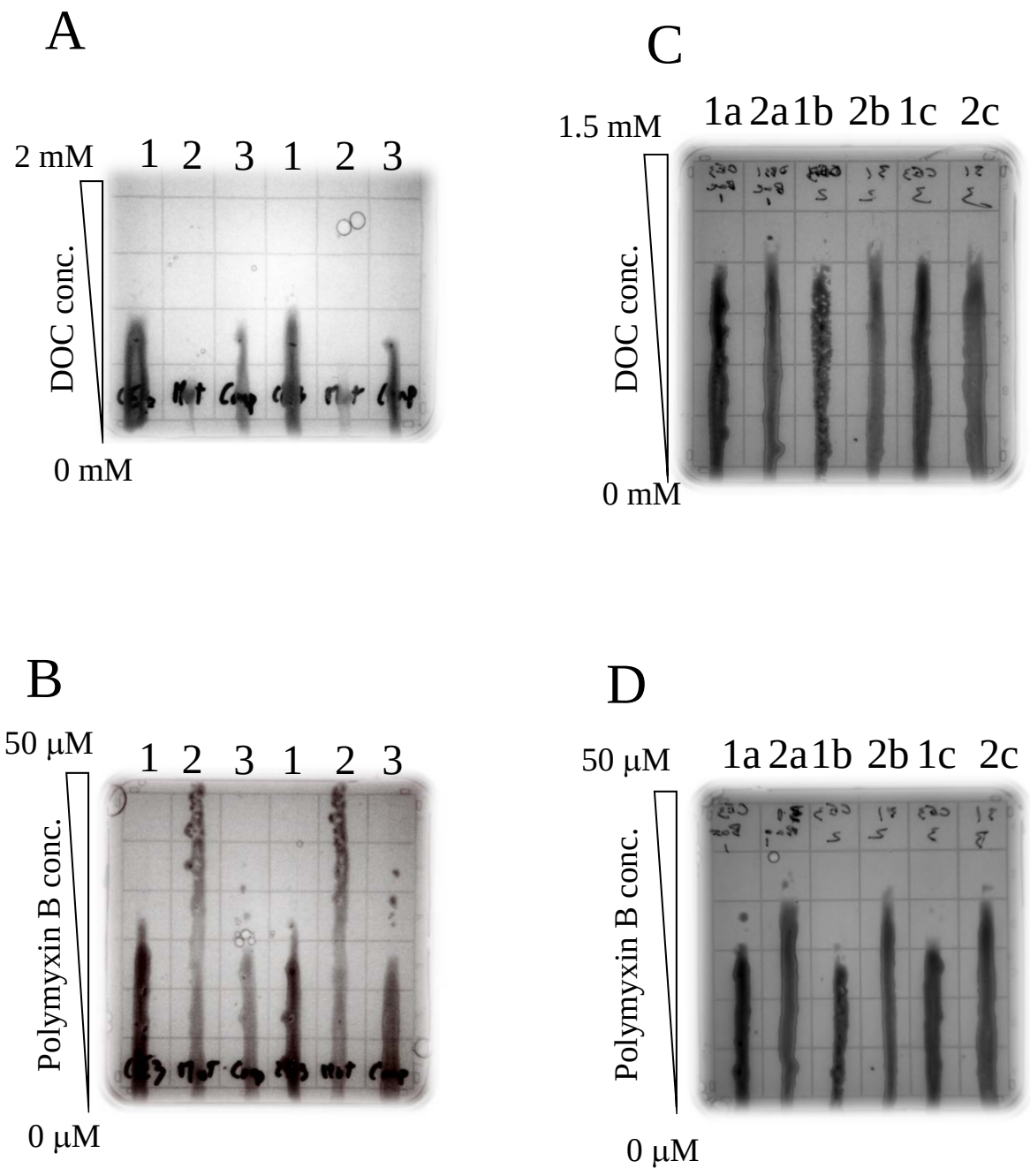

A
2 mM
1 2 3 1 2 3
DOC conc.
0 mM
C
1.5 mM
1a 2a 1b 2b 1c 2c
DOC conc.
0 mM
B
50 µM
1 2 3 1 2 3
Polymyxin B conc.
0 µM
D
50 µM
1a 2a 1b 2b 1c 2c
Polymyxin B conc.
0 µM

Figure 3.5

A

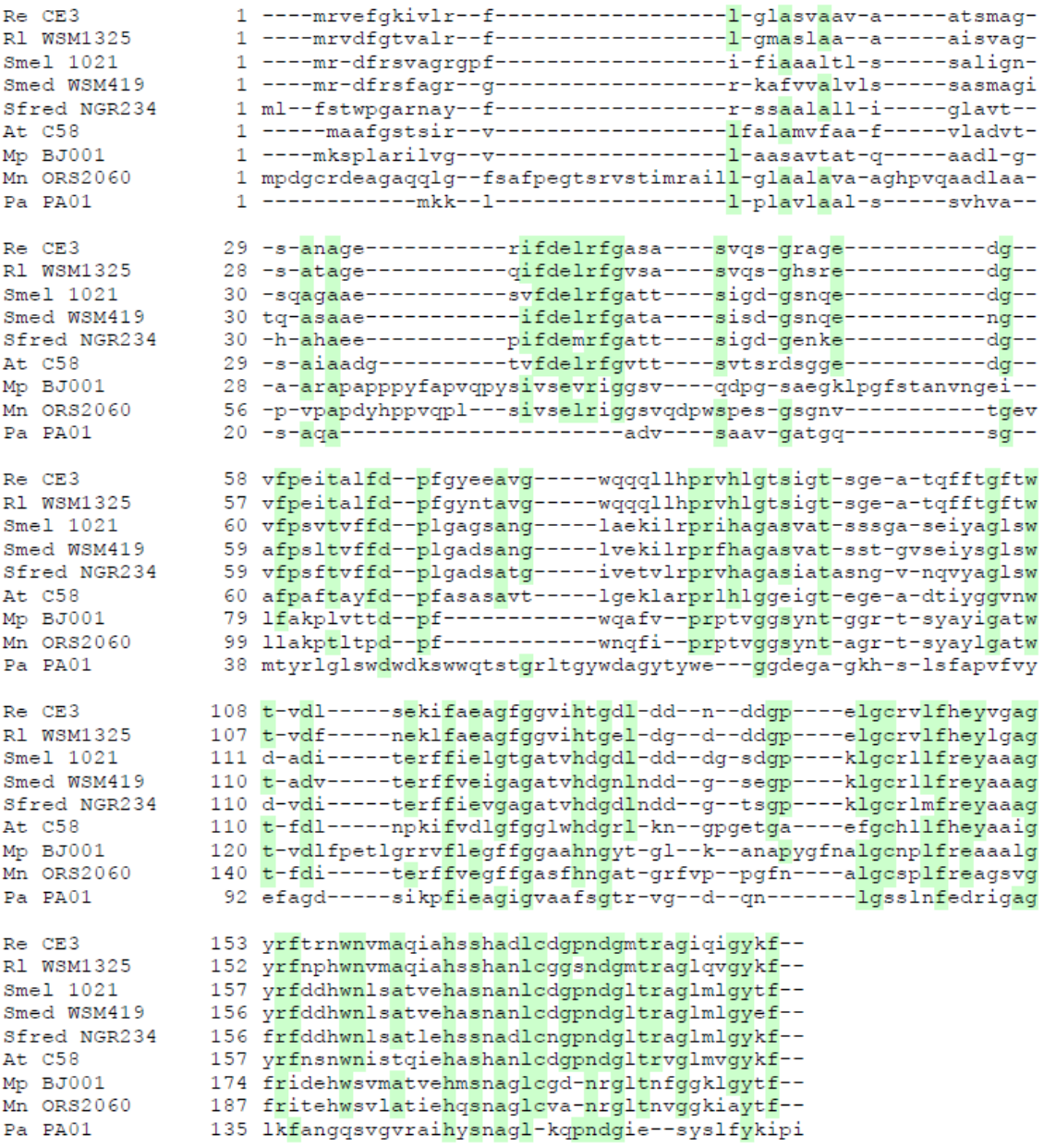
1

2

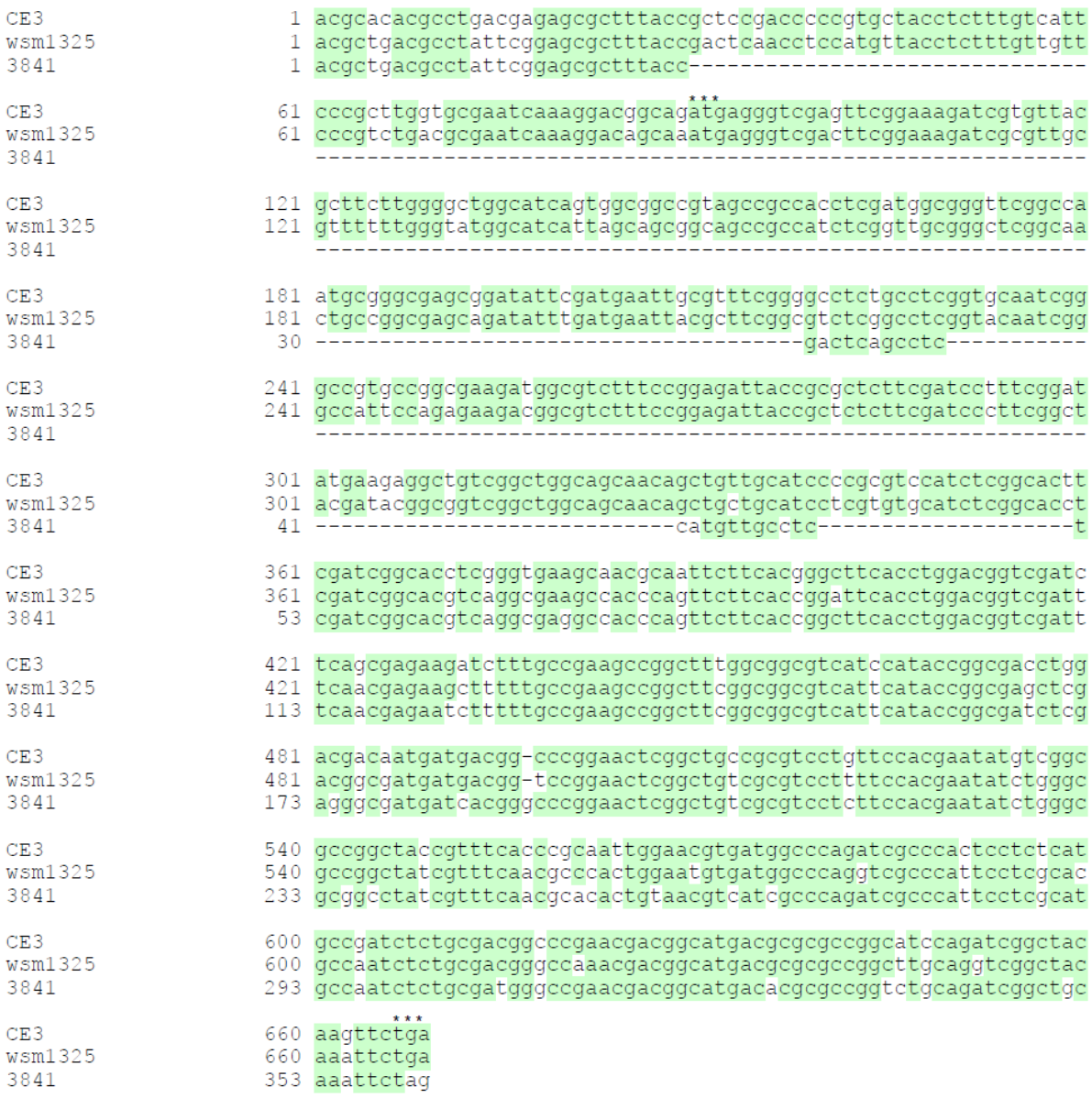
B

1 2

Figure 3.6



Supplemental 3.1



Supplemental 3.2

CHAPTER 4

Characterization of the galacturonosyl transferase genes *rgtA, B, C, D,* and *E* responsible for lipopolysaccharide synthesis in the nitrogen-fixing endosymbiont *Rhizobium leguminosarum*: Lipopolysaccharide core and lipid A galacturonosyl residues confer membrane stability

Dusty B. Brown, L. Scott Forsberg, Elmar L. Kannenberg and Russell W. Carlson

Accepted by *The Journal of Biological Chemistry.* Reprinted here with permission of publisher.

Abstract

Rhizobium lipopolysaccharide (LPS) contains four terminally linked galacturonic acid (GalA) residues: one attached to the lipid A and three attached to the core oligosaccharide moiety. Attachment of the GalA residues requires the lipid donor dodecaprenyl-phosphate GalA (Dod-P-GalA) which is synthesized by the GalA transferase (GalAT) RgtE reported here. The galacturonosyl transferases (GalATs) RgtA, B, and C utilize Dod-P-GalA to attach GalAs on the LPS core region and RgtD attaches GalA to the lipid A 4`-position. As reported here, the functions of the *rgtD* and *rgtE* genes were determined via insertion mutagenesis and structural characterization of the mutant lipid A. The *rgtE*$^{-}$ mutant lacked Dod-P-GalA as determined by mass spectrometry of total lipid extracts and the inability of the *rgtE*$^{-}$ mutant membranes to provide the substrate for heterologously expressed RgtA activity. In addition, we created single mutations in each of the *rgtA, B, C, D,* and *E* genes in order to study the biological function of the GalA residues. The structures of the core oligosaccharide region from each of the *rgt* mutants were elucidated by glycosyl linkage analysis. Each mutant was assayed for its sensitivity to sodium deoxycholate (DOC) and to the antimicrobial cationic peptide, polymyxin B (PmxB). The *rgt* mutants were more sensitive than the parent strain to DOC by varying degrees. However, the *rgtA, B,* and *C* mutants were more resistant to PmxB while the *rgtD* and *E* mutants were less resistant in comparison to the parent strain.

Introduction

The gram negative, nitrogen-fixing, endosymbiotic bacterium *Rhizobium leguminosarum* contains four terminally linked galacturonic acid (GalA) residues on its lipopolysaccharide (LPS) (1,2). Three of the GalA residues are attached to the core region while one GalA is attached to the lipid A (Figure 4.1). The lipid donor dodecaprenyl-phosphate GalA (Dod-P-GalA) is required for the attachment of all four GalA residues to the LPS as reported by Kanjilal-Kolar et al (3) and reported here. Three galacturonic acid transferase (GalAT) enzymes RgtA, B, and C were previously shown, in an *in vitro* assay, to transfer GalA from Dod-P-GalA to the synthetic substrate Man-Kdo_2-[4`-^{32}P]-Lipid IV_A (3,4). These results demonstrated that the GalATs RgtA and RgtB attach GalA to the branching Kdo likely at the 4 and 5 positions. However, it was unknown to which position each enzyme attaches the GalA residue. The RgtC enzyme attaches GalA to the Man residue at the 4 position. In addition, sequential addition of GalA was observed *in vitro* where by the RgtB and RgtC enzymes were active only in the presence of RgtA, and the RgtC enzyme was only active in the presence of both RgtA and RgtB (3,4).

Rhizobia, like several gram negative bacterial species, synthesize the lipid A precursor molecule Kdo_2-lipid IV_A. However, the 1 and 4` phosphates are subsequently removed respectively by phosphatases LpxE and LpxF (5,6) and, in turn, GalA is added to the 4` position. The proximal lipid A GlcN can be oxidized by the outer membrane monoxygenase LpxQ to form a negatively charged 2-aminogluconate (GlcNonate) residue (7). Prior to this report, the putative lipid A-4`-GalAT, RgtD, and the Dod-P-GalAT, RgtE, were unknown. In this report, we describe the function of the *rgtD* and *rgtE* genes by single gene mutagenesis. We also describe the additional preparation of single gene knock outs in *rgtA, B,* and *C* of *R. leguminosarum* bv.

viciae 3841, a nitrogen-fixing endosymbiont of pea. In this study, and the effect of each mutation on membrane stability, LPS structure, and LPS synthesis is descdribed.

Experimental Procedures

Bacterial strains, plasmids, and growth conditions—For a list of bacterial strains and plasmids used in this work, see Table 4.1. *E. coli* strains were grown at 37 ^{o}C on Luria-Burtani (LB) media. *Rhizobium leguminosarum* strains were grown on tryptone-yeast extract (TY) with 10 mM $CaCl_2$ or minimal (Y) medium at 30 ^{o}C as described previously containing 5 % sucrose where needed (8). Antibiotics were used in the following concentrations; tetracyclin, Tc (15 μg/ml), kanamycin, Km (50 μg/ml), gentamicin, Gm (*R. leguminosarum* [30 μg/ml], *E. coli* [(15 μg/ml]).

Insertion mutagenesis of galacturonosyl transferase genes—The GalAT genes, *rgtA, B, C, D,* and *E,* were individually mutated by inserting the gentamicin resistant gene *aacC1* (9) near the center of each gene thereby disrupting gene function. Plasmid extractions, gel extractions and PCR/enzyme clean-up was performed using Quiagen mini prep kits. Restriction enzymes were purchased from Promega, Inc. (Madison, WI). For each mutagenesis procedure, approximately 1 kb fragments containing either half of the open reading frame and upstream DNA or half of the open reading frame and downstream DNA were amplified (iProof^{TM} High-Fidelity polymerase, Biorad, Hercules, CA) from the *R. leguminosarum* bv. *viciae* 3841 genome. The primers were constructed containing restriction enzyme sites for cloning of each fragment (Table 4.2). Each PCR fragment was cloned into plasmid pUC18 by digestion of pUC18 and PCR products with respective restriction enzymes that recognize the engineered primers followed by ligation (Table

4.2). The resulting plasmids created, pRgtAUp, pRgtADwn, pRgtBUp, pRgtBDwn, pRgtCUp, pRgtCDwn, pRgtDUp, pRgtDDwn, pRgtEUp, and pRgtEDwn were transformed into *E. coli*. The cloned upstream fragments were released by restriction digest using the same enzymes that were used to clone them and separated by gel electrophoresis followed by gel extraction. The isolated upstream fragments were subcloned into the plasmids containing the downstream fragments by double restriction digestion and ligation using the following enymes; pRgtADwn (*BamHI, ,EcoRI*), pRgtBDwn (*BamHI, XbaI*), pRgtCDwn (*SmaI, EcoRI*), pRgtDDwn (*BamHI, XbaI*), and pRgtEDwn (*BamHI, EcoRI*). The resulting plasmids contained upstream DNA, the open reading frames with an engineered single cloning site near the center, and downstream DNA (pRgtAUpDwn, pRgtBUpDwn, pRgtCUpDwn, pRgtDUpDwn, and pRgtEUpDwn) and were transformed into *E. coli*. The gentamicin resistant gene cassette *aacC1* (1 kb) was released from plasmid pM255 (9) by restriction digest and isolated by gel electrophoresis followed by gel extraction. The digested *aacC1* fragments were inserted into plasmids containing up and downstream DNA by single restriction digestion and ligation using the following enzymes; pRgtAUpDwn (*BamHI*), pRgtBUpDwn (*BamHI*), pRgtCUpDwn (*SmaI*), pRgtDUpDwn (*BamHI*), and pRgtEUpDwn (*BamHI*). The resulting plasmids (pRgtA-KO, pRgtB-KO, pRgtC-KO, pRgtD-KO, pRgtE-KO) were transformed into *E. coli*. The inserts from each plasmid were released by restriction digestion followed by gel electrophoresis and gel extraction using the following enzymes; pRgtA-KO (*XbaI, EcoRI*), pRgtB-KO (*XbaI, EcoRI*), pRgtC-KO (*EcoRI, HindIII*), pRgtD-KO (*SacI, XbaI*), and pRgtE-KO (*XbaI, EcoRI*). The fragments were cloned into the suicide vector pEX18-Tc (10) which harbors the *sacB* lethal gene to create plasmids pEX/RgtA-KO, pEX/RgtB-KO, pEX/RgtC-KO, pEX/RgtD-KO, and pEX/RgtE-KO which were transformed into *E. coli.*

A tri-parental mating was conducted in order to move each pEX/E-KO into 3841 via conjugal transfer. The donor strain *E. coli* Top 10 harboring pEX/E-KO, the helper strain *E. coli* carrying the helper plasmid pRK2013 (20), and the recipient strain 3841 were cultured overnight in 100ml media. The *E. coli* strains were grown in LB (37 ^{o}C) broth and the *R. leguminosarum* strain 3841 was grown in TY broth (30 ^{o}C). 2 ml of cultured donor and helper were centrifuged at 4000 x g for 10 minutes. The cell pellets were each re-suspended in 4ml TY medium. In a 4000 x g for 10 min. The supernatant was discarded until there was just enough left to make a resuspension. The entire suspension (approximately 500 μl) was spotted on a nitrocellulose membrane previously laid on a TY agar plate and kept at 30 ^{o}C overnight. The resulting paste was washed off the membrane with 4 ml sterile de-ionized water. In order to select for double recombinants, various amounts of the mating mix were plated onto minimal Y agar plates containing 30 μg/ml Gm and 5 % sucrose. Colonies were picked and streaked on TY/Gm plates. Mutant colonies were isolated using a PCR screen and given the strain names EL202 (*rgtE*$^{-}$), EL203 (*rgtA*$^{-}$), EL204 (*rgtB*$^{-}$), EL205 (*rgtC*$^{-}$), and EL206 (*rgtD*$^{-}$). The EL202 (*rgtE*$^{-}$), EL203 (*rgtA*$^{-}$), 204 (*rgtB*$^{-}$), and 205 (*rgtC*$^{-}$) mutants were respectively complemented with the previously described plasmids pMKGE, pRK-RgtA, pRK-RgtB, and pRK-RgtC by tri-parental (3,4).

Isolation of lipopolysaccharide—Briefly, bacteria were grown to mid-late log phase, pelleted and washed with 0.9 % sodium chloride, followed by two washes with deionized water. The pelleted bacteria were subjected to hot phenol-water extraction as described previously (11). Water layer material was dialyzed in 12-14 kDa MWCO dialysis bags against deionized 10 L deionized water for five days, each day exchanging the water. The sample was lyophilized and subjected to

DNase, RNase, and protease treatment as described previously (12) and dialyzed in 12-14 kDa MWCO bags for three days and lyophilized.

DOC-PAGE analysis—Gel electrophoresis in the presence of DOC was performed as previously described (13). Briefly, 1 μg of extracted LPS material was loaded onto an 18 % acrylimide gel and current was applied (30 mA [constant], 400 V) in the presence of DOC buffer. The gel was subjected to Alcian blue/silver staining as previously described (14).

Glycosyl and fatty acid composition analysis of lipopolysaccharide—The composition of LPS was determined by gas chromatography/mass spectrometry (GC/MS) of trimethylsilyl (TMS) methyl glycosides and methyl ester fatty acids as described previously (15,16). Isolated LPS, 300 μg per sample, was subjected to methanolysis (80 ^{o}C, 16-18 hrs) followed by *N*-acetylation (100 ^{o}C, 1 hr) and trimethylsylation (80 ^{o}C, 30 min). The TMS derivatives were dissolved in hexane and injected into a Hewlett Packard 5890 Series II GC/MS analyzer fitted with a 30m EC-1 column (Alltech).

Lipid A isolation and analysis—Lipid A was isolated by the standard mild acid hydrolysis procedure (17). Briefly, approximately 5 mg of isolated LPS was treated with 1 % acetic acid (100 ^{o}C, 1 hr) to cleave the labile Kdo ketosidic linkage between the core moiety and the lipid A. The lipid A/core-polysaccharide mixture was partitioned three times by chloroform: methanol: water (2: 2: 1.8, v: v: v) extraction. The lipid A-containing chloroform layers were pooled and dried under a stream of air.

The lipid A preparations were analyzed by matrix assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF-MS). Isolated lipid A was dissolved in a 3: 1 (v: v) chloroform: methanol mixture, mixed 1: 1 (v: v) with 2,4,6-trihydroxyacetophenone (THAP) matrix and 0.5 μl was spotted on a MALDI plate. Spectra were acquired by a MALDI-TOF

analyzer (Applied Biosystems 5800) in the positive reflectron mode operating at a 20 kV extraction voltage scanning from 800-3,000 mass units.

Linkage Analysis of the Kdo and Uronosyl Residues in LPS—The Kdo and uronic acid linkages in the intact LPS samples were identified by preparing the partially methylated carboxyl-reduced alditol acetate (PMAA) derivatives, following the procedures described previously for *R. etli* LPS (2,18) with additional modifications. The LPS samples were first subjected to permethylation using the sodium hydroxide procedure (19), modified to promote recovery of carboxymethylated Kdo residues as described for sialic acid containing glycoconjugates (20).

The purpose of these modified procedures is to preserve the Kdo carboxymethyl esters during the aqueous work-up following methylation, thus minimizing base catalyzed elimination from the Kdo and subsequent loss of linkage. This modified work-up consists of: 1) cooling the methylation reaction mixture and adding excess chloroform (2 ml), 2) low speed centrifugation to remove excess precipitated NaOH and NaI, and 3) acidification of the chloroform supernatant by washing with dilute 5 % aqueous acetic acid (3 ml). Following low speed centrifugation, the chloroform layer, containing the permethylated LPS, is washed 9 more times with water and then dried under nitrogen stream.

The resulting permethylated LPS was then subjected to reduction of the Kdo and uronosyl carboxymethyl groups with lithium triethylborodeuteride (Superdeuteride, Aldrich) in THF (2 h at room temperature); mild acid hydrolysis (0.1 M trifluoroacetic acid, 100 °C, 30 min) to cleave the Kdo ketosidic linkages; reduction of Kdo residues at C-2 carbonyl group (using $NaBD_4$ in water/ethanol); normal hydrolysis (2 M trifluoroacetic acid, 121 °C, 2 h) to cleave the remaining uronosyl and neutral sugar linkages; reduction of the newly formed aldehydo sugars (with $NaBD_4$ in 50 mM NH_4OH); and acetylation of the resulting partially methylated alditols to

yield the PMAA derivatives. Acetylation was performed in acetonitrile-pyridine-acetic anhydride containing 4-*N,N′*-dimethylaminopyridine as catalyst for 4 h at room temperature as described (21). The PMAA derivatives were analyzed by GC/MS (electron

impact) using a 30-m SP-2330 capillary column (Supelco) programmed from 80 to 250 °C.

Preparation of E. coli *membranes*—Cultures, 1 L, of *E. coli* harboring plasmids pET23a or pRgtA were grown to mid-logarithmic phase and harvested by centrifugation at 4000 x g and 4 ^{o}C for 20 min. The cell pellets were re-suspended in pre-chilled 50 mM HEPES, pH 7.5. Membranes were prepared as previously described (4).

Extraction and purification of total lipids from the rgtE^{-} mutant—Cultures, 1 L each, of the 3841 parent, *rgtE^{-}* mutant and complemented strains were grown to mid/late logarithmic phase and harvested by centrifugation at 4000 x g for 15min and washed 1 time with phosphate buffered saline. Lipids were extracted from the pellets as described previously (3).

Analysis of membrane lipids by combined liquid chromatography-electrospray ionization mass spectrometry (LC-ESI/MS)—Normal phase LC-ESI/MS of lipids was performed using an Agilent 1200 Quaternary LC system coupled to a QSTAR XL quadrupole time-of-flight tandem mass spectrometer (Applied Biosystems, Foster City, CA). An Ascentis® Si HPLC column (5 μm, 25 cm × 2.1 mm) was used. Mobile phase A consisted of chloroform/methanol/aqueous ammonium hydroxide (800: 195: 5, v/v/v). Mobile phase B consisted of chloroform/methanol/water/ aqueous ammonium hydroxide (600: 340: 50: 5, v/v/v/v). Mobile phase C consisted of chloroform/methanol/water/aqueous ammonium hydroxide (450: 450: 95: 5, v/v/v/v). The elution program consisted of the following: 100 % mobile phase A was held isocratically for 2 min and then linearly increased to 100 % mobile phase B over 14 min and held at 100 % B for 11 min. The LC gradient was then changed to 100 % mobile phase C over 3 min

and held at 100 % C for 3 min, and finally returned to 100 % A over 0.5 min and held at 100 % A for 5 min. The total LC flow rate was 300 µl/min. The post-column splitter diverted ~10 % of the LC flow to the ESI source of the Q-Star XL mass spectrometer, with MS settings as follows: IS = -4500 V, CUR = 20 psi, GS1 = 20 psi, DP = -55 V, and FP = -150 V. Nitrogen was used as the collision gas for MS/MS experiments. Data acquisition and analysis were performed using the instrument's Analyst QS software.

In Vitro *reconstitution of GalAT activity*—The activity assay methods were adapted from Kanjilal et al (3,4). Briefly, the standard reaction mixture contained 50 mM MES, pH 6.5, 0.05 % Nonidet P-40, 2 mM $MgCl_2$ and 2.5 µM Kdo_2-1-dephospho-[4`-^{32}P]lipid IV_A (10^6 cpm/nmol). When mentioned, total extracted lipid was added to the reaction at a final concentration of 0.25 mg/ml and purified donor was added at a final concentration of 5 µM. Washed membranes were used as the enzyme source at a final concentration of 0.25 mg/ml. The reaction mixtures were kept at 30 ^{o}C for 10 or 30 min. Reactions were stopped by spotting 4 µl on a Silica Gel 60 plate. Plates were developed in a solvent system containing chloroform, pyridine, 88 % formic acid and water (30: 70: 16: 10, v/v/v/v), dried and exposed overnight on a Phosphorimager screen. Product was detected using PhosphorImager (Storm 840, Amersham Biosciences).

Polymyxin B and DOC resistance assay—Polymyxin B gradient plates were made by pouring at an angle 25 ml TY agar media containing 50 µg/ml of PmxB in a square 100 mm x 15 mm petri dish. After the agar solidified, they were placed level and 40 ml TY agar media was poured on top and the plates were allowed to cool. The PmxB was allowed to diffuse throughout the plates overnight. A loop full of bacteria grown for four days on TY agar media was dissolved in 1 ml sterile phosphate buffered saline (pH 7.4). Then, 10 µl of bacterial suspension was spotted on the

gradient plate and streaked from highest to lowest PmxB concentration. The plates were incubated for four days at 30 ^{o}C.

Sodium deoxycholate gradient plates were made as described above using TY containing 2 mM DOC. Bacteria were prepared and streaked onto the gradient plates and incubated as described above.

Plant growth and inoculation—Surface sterilized pea seeds, variety Early Alaska, (Bunton Sead Company, Inc.) were germinated in the dark at room temperature for three days on 0.8 % water agar plates and transferred to sterile Erlenmeyer flasks (1 seed per flask) containing a foam top and 300 ml nitrogen free 1.5 % agar plant media (Fahraeus media). The Erlenmeyer flasks were surrounded with brown paper bags and placed in a Conviron growth chamber. Approximately 10 days post germination, plants were inoculated with bacteria (1 ml of late-log/early–stationary phase liquid TY cultures). The plants were maintained at 19 ^{o}C with 14 h of light and 15 ^{o}C with 10 h of dark.

Results

Discovery of the rgtD *and* rgtE *genes*—The predicted peptide sequence from the previously described LPS core GalATs, RgtA, B, and C (4), were aligned with the *R. leguminosarum* bv. *viciae* 3841 genome using the basic local alignment search tool (22). Results revealed a chromosomally localized single cistronic open reading frame (accession number RL0684) with appreciable similarities to RgtA, B, and C (S4.1). In addition, RL0684 is predicted to be an ArnT-like protein, a class of proteins responsible for the transfer of monosaccharides from bactoprenyl-phosphate monosaccharide lipid donors to the lipid A back bone (23,24).

Furthermore, the *rgtD* gene is found in the genomes of a number of *R. etli* and *R. leguminosarum* strains including *R. etli* CE3/CFN42, *R. etli* CIAT652, *R. leguminosarum* bv. *viciae* 3841, and *R. leguminosarum* bv. *trifolii* WSM1325. The LPS lipid A and core oligosaccharide structures are conserved in these *Rhizobium* species and, thus, it is expected that genes responsible for the synthesis of LPS lipid A and core oligosaccharide are also highly conserved.

The *rgtE* gene, previously referred to as *orf3,* was predicted to encode the GalAT the synthesizes Dod-P-GalA by Kanjilal-Kolar et al (3,4), however, the gene function was not determined. The predicted amino acid sequence of RgtE is similar to a class of glycosyl transferases that have been demonstrated experimentally to synthesize polyprenyl-phosphate glycosyl lipid donors necessary for the glycosylation of lipid A (S4.2) (25-27). The peptides are predicted bacteria like dolichol phosphate mannosyl (dpm1) type two glycosyl transferases and are of similar sizes containing two predicted C-terminal transmembrane domains that likely tether the proteins to the inner membrane while the catalytic portion resides on the inner membrane / cytoplasm interface.

A*nalysis of total lipids from the* rgtE$^-$ *mutant by LC-ESI-MS*—Total lipids were collected from cell pellets of the parent *R. leguminosarum* 3841, mutant EL202, and the complemented mutant strain using the Bligh and Dyer extraction method as described in experimental procedures. Kanjilal-Kolar *et al.* (3) previously reported the lipid donor Dod-P-GalA to be involved in the transfer of GalA to *R. leguminosarum* LPS by the *Rhizobium* GalATs; RgtA, B, and C. Therefore, if RgtE is required for the synthesis of Dod-P-GalA, then mutant EL202 lipids would not contain Dod-P-GalA. Extracted lipids were subjected to LC-ESI-MS analysis for the presence of Dod-P-GalA. As expected, *R. leguminosarum* 3841 lipids contained the appropriate ions observed previously for Dod-P-GalA (Figure 4.2, panel A). At 20 minutes into the LC run a

peak was observed that contained Dod-P-GalA singly and doubly charged signature ions of *m/z* ratios of 1089 and 544, respectively. Ion *m/z* 913 agrees with the molecular weight of Dod-P. The signature ions for Dod-P-GalA were not detected in extracted lipids from mutant EL202 (Figure 4.2, panel B). However, strain EL202 lipids did contain ion *m/z* 913 due to Dod-P. Complementing the mutant with the *rgtE* gene recovered the appearance of Dod-P-GalA in the lipid composition (Figure 4.2, Panel C).

The inability of the rgtE$^-$ *mutant lipid extracts to act as a GalA donor in an enzyme activity assay with the* R. leguminosarum *bv.* viciae 3841 *GalAT, RgtA*—As previously described (3), extracted lipids from *R. leguminosarum* bv. *viciae* 3841, due to the presence of Dod-P-GalA, are able to act as a donor in the conversion of the radio-labeled substrate Kdo_2-1-dephospho-[4`-$^{32}$P]lipid IV$_A$ to the GalA-$Kdo_2$-1-dephospho-[4`-^{32}P]lipid IV$_A$ product in the presence of membranes from *E. coli* over expressing the *R. leguminosarum* bv. *viciae* 3841 RgtA. Total lipids extracted from strains 3841, EL202, and complemented EL202 mutant were assayed for their ability to serve as a donor for this reaction. Standard reaction mixtures were prepared with equal amounts of lipid extracts, *E. coli* membranes containing RgtA, and radio-labeled Kdo_2-1-dephospho-[4`-^{32}P]lipid IV$_A$ substrate. Reactions were performed for 10 or 30 minutes and stopped by spotting 4 µl of the reaction mixture on a Silica Gel 60 plate. The Silica Gel plate containing each reaction was subsequently developed using the solvent system described (see Experimental Procedures), and the products were observed on a PhosphoImager (Amersham Biosciences) (Figure 4.3). Reaction mixtures containing parent strain 3841 lipid extract (lane 5) were able to convert the substrate into the product. In contrast, no product was observed from reactions containing mutant EL202 lipid extract (lane 6). Lipid extract from mutant EL202 complemented with the *rgtE* gene was able to carry out conversion of substrate to product (lane 7). These results

are consistent with the requirement of RgtE for the production of Dod-P-GalA which is necessary for the RgtA catalyzed transfer of GalA to Kdo_2-1-dephospho-[4`-^{32}P]lipid IV_A.

Composition and gel electrophoresis of lipopolysaccharides from the rgt *mutants*—The presence of LPS derived GalA in each strain was determined by GC/MS of trimethylsilyl (TMS) derived methyl glycosides and fatty acid methyl esters. As expected, strain EL202 (*rgtE*$^-$) did not contain detectable amounts of GalA. Because there is a single galactosyl residue on each LPS molecule, the Gal: GalA ratio was used to determine differences in the amounts of GalA. Strains EL203 (*rgtA*$^-$) (Gal : GalA = 1.0: 2.0), EL205 (*rgtC*$^-$) (Gal : GalA = 1.0: 2.9) and EL206 (*rgtD*$^-$) (Gal : GalA = 1.0: 2.9) LPS had less GalA than that of the parent strain 3841 (Gal : GalA = 1.0: 3.9). There was not an observable decrease in the amount of GalA in the LPS of strain EL204 (*rgtB*$^-$) (Gal : GalA = 1.0: 4.0). However, as discussed further below, linkage analysis showed that strain EL204 (*rgtB*$^-$) was disrupted in its ability to add one of the two GalA residues to the branching Kdo residue; i.e. Kdo2 in Figure 4.1. Together, these results suggest that all of the *rgt* mutants were disrupted to varying degrees in GalA addition to the LPS.

Gel electrophoresis (DOC-PAGE) of the *rgt* mutant LPSs was performed and compared with parent strain (Figure 4.4). Surprisingly, there was no observable difference in the migration of the *rgt* mutants LPSs when compared to parent strain 3841. This was unexpected, particularly in the case of strain EL202 (*rgtE*$^-$) as this strain lacks four GalA residues when compared to strain 3841 which calculates as a mass difference of approximately 704 mass units. It may be that any change in migration due to the reduction in size is offset by the loss of charge due to the absence of GalA residues. The banding pattern of the *rgt* mutants suggests that there was no major disruption to OPS biosynthesis. In agreement, composition analysis of the LPS from each *rgt* mutant shows that they all contain the normal OPS glycosyl residues present in the parent

strain LPS; i.e. *N*-acetyl quinovosamine (QuiNAc), fucose (Fuc), and 3-O-methyl-6-deoxytalose (3Me6dTal). The LPSs from strains EL202 (*rgtE*$^-$) and EL206 (*rgtD*$^-$) displayed a slightly different banding pattern than those of the parent strain and the other *rgt* mutants (Figure 4.4, Lanes 5 and 6) in that the intensity of one band was greatly reduced. The molecular basis for this difference is not known.

MALDI-TOF/MS of the lipid A from the rgt *mutants*—The *rgt* mutant lipid A preparations were studied in detail by MALDI-TOF/MS. The lipid A observed in *Rhizobium* species is microheterogeneous in nature. The microheterogeneity is due to several factors including the presence or absence of β-hydroxybutyryl (BHB) linked at the C27 position of the secondary long chain fatty acid 27-hydroxyoctacosanoic acid (mass difference of 86), the non-stoichiometric oxidation of the proximal GlcN to GlcNonate (a mass difference of 16), and variation in acyl chain length (mass differences in series of 28 or 14 mass units). In addition, chemically modified lipid A species occur due to the mild acid hydrolysis procedure whereby acid leads conversion of the proximal GlcNonate to GlcNonolactone and resulting acid catalyzed β-elimination of β-hydroxymyristate from the 3-position, which results in the formation of 2,3-dideoxy-2-amino-2-ene-gluconolactone as the proximal lipid A glycosyl residue (Figure 4.5, structures A and D).

MALDI-TOF MS analysis of the *rgtA*$^-$*, B*$^-$*,* and *C*$^-$ mutant lipid A preparations (data not shown) produced ions identical to those observed for the parental lipid A (Figure 4.5, structures A, B, and C) showing that each of these mutants produce a normal lipid A structure with GalA at the 4'-position. The EL202 (*rgtE*$^-$) and EL206 (*rgtD*$^-$) mutant lipid A lacked GalA. (Figure 4.5, panel III). The composition of the EL206 *(rgtD*$^-$*)* lipid A was compared with that of the parent lipid A (Figure 4.6) and confirmed that the EL206 (*rgtD*$^-$) mutant produces lipid A that lacks any detectable GalA. This was also the case for the EL202 (*rgtE*$^-$) mutant (data not shown).

Linkage analysis of rgt *mutants*—In order to determine the structures of the *rgtA*$^-$, *B*$^-$, *C*$^-$, and *D*$^-$ mutant LPS core oligosaccharides, glycosyl linkage analysis was performed on their LPSs as well as the LPS from the parent 3841 strain, which contains all LPS GalA residues, and the LPS from EL202 (*rgtE*$^-$) mutant, which lacks all the LPS GalA residues. The results of this analysis are shown in Table 4.3.

The core oligosaccharide structure of the parent *R. leguminosarum* bv. *viciae* 3841 LPS has been determined (1,2) and shown to contain three Kdo residues as shown in Figure 4.1. The proximal Kdo residue (Kdo1) is ketosidically linked the to the distal GlcN residue of the lipid A backbone, the branching Kdo residue (Kdo2) is linked to the 4-position of the proximal Kdo1 residue, and the distal Kdo residue (Kdo3) is linked to the 6-position of the core oligosaccharide Gal residue. Two of the three core GalA residues are linked to the 4- and 5-positions of the branching Kdo2 residue; i.e. Kdo2 is 4,5-linked. The core Man residue is linked to the 5-position of the proximal Kdo1 residue; i.e. Kdo1 is also 4,5-linked. The OPS, which is present in non-stoichiometric amounts, is linked to the 4-position of the distal Kdo3 residue; i.e. Kdo3 is 4-linked when OPS is present and terminally linked on LPS molecules without OPS. Thus, if one considers the parent LPS structure (Figure 4.4) the ratio of 4,5-linked Kdo (Kdo1+Kdo2): Kdo3 (4-linked Kdo3 + terminally linked Kdo3) would be 2: 1, and the ratio of terminally linked Kdo: 4-linked Kdo, i.e. Kdo3 without and with OPS attached, would indicate the ratio of LPS structures that lack OPS. Table 4.3 shows that the ratio of 4,5-linked Kdo: 4-linked + terminally linked Kdo is 1.7 which is consistent with the expected 2: 1 ratio of (Kdo1 + Kdo2): Kdo3. The ratio of terminally linked to 4-linked Kdo3 from Table 4.3 is 0.35 which reflects the ratio of LPS molecules without OPS to those with OPS. In addition to the linkage of the Kdo residues, the expected ratio of 6-linked Gal + terminally linked GalA (which would also appear as 6-linked

residue in the method used) to 4,6-linked Man would be 4: 1 and Table 4.3 shows that it is 3.3 which is reasonably consistent with the expected value since the recovery of terminally linked GalA is always low with the method used. Thus, the linkage analysis results of the core Kdo, Gal, GalA, and Man residues are consistent with the known core oligosaccharide structure (see Figure 4.4, and structure A in Figure 4.7) for this LPS.

The *rgtE*$^{-}$ mutant produces an LPS that is devoid of GalA. Thus, if one considers parental LPS core structure A shown in Figure 4.7, the absence of all GalA residues would result in one 4,5-linked Kdo (the proximal Kdo1 residue), one terminally linked Kdo (the branching Kdo2 residue), and one 4-linked Kdo (the distal Kdo3 residue); i.e. a 1: 1: 1 ratio of 4,5-: 4-: terminally linked Kdo residues. Since the LPS preparation consists of LPS with and without OPS, a certain amount of terminally linked Kdo3 would be present due to structures in which there is no OPS attached to the 4-position of the distal Kdo3 residue. The actual ratio of 4,5-: 4-: terminally linked Kdo is 1.0: 0.41: 1.3. Again the lower than expected value of 4-linked Kdo and the higher than expected value of terminally linked Kdo are due to the fact that there is, as with the parent LPS, a portion of LPS that lacks OPS. The reduction in the relative amount of 4-linked Kdo3 compared to the parent LPS suggests that the *rgtE*$^{-}$ mutant may be somewhat reduced in the proportion of LPS molecules containing OPS in comparison to parental LPS; however, as described above this possible reduction is apparently not sufficient to detect with regard to the LPS composition and DOC-PAGE, described above. The observed linkages of the Gal, GalA, and Man residues are also consistent with the absence of terminal GalA. The mass spectrum of the partially methylated alditol acetate (PMAA) derived from the 6-linked Gal residue shows that it does not contain deuterium at C6 confirming that terminal-GalA is not present (the PMAA derivatives and their diagnostic primary ion fragments are listed in Table 4.3). The fact that there

is no detectable 4,6-linked Man and only 6-linked Man confirms that there is no GalA attached to the 4-position of the Man residue, and the 1: 1 ratio of 6-linked Gal: 6-Man is completely consistent with a core oligosaccharide that is devoid of GalA residues. Thus, the *rgtE*$^-$ mutant has the core structure F shown in Figure 4.7.

It is known that RgtA is the GalAT that adds one of the two GalA residues to either the 4-, or 5-position of the branching Kdo2 residue in an *in vitro* assay using Man-Kdo_2-lipid-IV$_A$ as the substrate (4). It was not known which of these two GalA residues is added by RgtA, or how the lack of either one of these GalA residues might affect the addition of the remaining GalA residues to the LPS in *R. leguminosarum* bv. *viciae* 3841. The results for the Kdo linkages (see Table 4.3) show that the 4,5-: 4-: terminal- ratio of Kdo is 1: 0.77: 1.5 which is quite similar to the results for the Kdo linkages of the *rgtE*$^-$ mutant which lacks both of the GalA residues attached to the branching Kdo2 residue. These results support the conclusion that the *rgtA*$^-$ is unable to add either GalA residue to the branching Kdo2 residue. Analysis, GC/MS, of the PMAAs for the Gal, GalA, and Man residues show that a portion of the PMAA of the 6-linked Gal is due to terminal GalA and that the presence of 4,6-linked Man indicates that some GalA is added to the 4-position of the Man residue. The percentage of Man residues containing this GalA residue is 59 % while 41 % lack the GalA residue. Thus, inactivation of *rgtA* prevents GalA addition to both the 4- and 5-positions of the branching Kdo2 residue, i.e. the absence of RgtA activity also prevents RgtB activity, and also reduces the level of GalA added to the Man residue, i.e. negatively affects the activity of RgtC. Thus, the *rgtA*$^-$ mutant contains a mixture of the core structures B and F shown in Figure 4.7.

As with RgtA, it is known that RgtB adds one of the two GalA residues to the 4- or 5-position of the branching Kdo2 residue in an *in vitro* assay using Man-Kdo_2-lipid-IVA as the substrate, but the identity of the position was unknown (4). Table 4.3 shows that the ratio of 4,5-: 4-: terminal Kdo in the *rgtB*$^-$ mutant is 1.0: 2.0: 0.20. This result supports that RgtA activity in the absence of RgtB adds GalA to the 4-position of the branching Kdo2 residue. Therefore, the functions of RgtB and RgtA are to add GalA to the 5- and 4-positions, respectively, of the branching Kdo2 residue. Further, given the fact that inactivation of RgtA entirely prevents addition of either of these GalA residues, these results strongly support the conclusion that RgtA addition of GalA to the 4-position of the branching Kdo2 residue is required for RgtB activity. The linkages of the Gal, GalA, and Man residues of the *rgtB*$^-$ mutant, Table 4.3, show a 4,6-Man: 6-Man: 6-Gal + terminal-GalA ratio of 1.0: 0.09: 2.1. This result supports a core structure in which the Man residue is almost completely substituted with GalA at the 4-position. Thus, in the *rgtB*$^-$ mutant, RgtA activity is able to add GalA to the 4-position of the branching Kdo2 residue and, in contrast to the *rgtA*$^-$ mutant, almost complete GalA substitution of the Man residue at O-4 occurs. Therefore, the *rgtB*$^-$ mutant largely contains structure C shown in Figure 4.7.

Previous work showed that RgtC added GalA to the 4-position of the Man residue in an *in vitro* assay using Man-Kdo_2-lipid-IVA as the substrate (4). However, it was not known if the absence of this activity would affect the addition of any of the other core GalA residues during LPS synthesis in *R. leguminosarum* bv. *viciae* 3841. Table 4.3 shows that the ratio of 4,5-: 4-: terminally-linked Kdo residues is similar to that observed for the 3841 parent strain with the exception that the relative level of 4-linked Kdo is somewhat reduced (22 rather that 28 %) while the level of terminal Kdo is somewhat increased (16 rather than 9.8 %). These results support

the conclusion that the *rgtC*- branching Kdo2 residue is fully substituted by GalA while the level of OPS substitution at O-4 of the distal Kdo3 may be slightly reduced compared to the parent LPS resulting in the slight decrease in 4-linked Kdo and slight increase of terminally linked Kdo. The linkages of Gal, GalA, and Man residues, Table 4.3, clearly indicate a large decrease in 4,6-linked Man and an increase in 6-linked Man; i.e. the ratio of these residues is 0.24: 1.0 indicating that the *rgtC*- mutant is essentially unable to add GalA to the 6-position of the core Man residue. This result is consistent with those previously published regarding RgtC activity in an *in vitro* assay using Man-Kdo_2-lipid-IV_A as the substrate (4). The small amounts of 4,6-linked Man indicates that the *rgtC*- mutant may be slightly leaky and, therefore, there is a low level of RgtC activity. In summary, these results support the conclusion that the *rgtC*- mutant mainly produces structure D shown in Figure 4.7 with a small amount of the parent structure A.

We describe above that *rgtD* encodes the GalAT which adds GalA to the 4'-position of the lipid A. In order to determine if the *rgtD*- mutant was altered in ability to add GalA to the core oligosaccharide, glycosyl linkages were determined and the results are shown in Table 4.3. The level of 4,5-linked Kdo is decreased in comparison to the parent LPS to levels similar to those observed in the *rgtA*-, *rgtB*- and *rgtE*- mutants. This result indicates that the branching Kdo2 residue may not be fully substituted with GalA at either the 4- or both the 4- and 5-positions. However, the increased level of 4-linked Kdo relative to that observed in the *rgtA*- and the *rgtE*- mutants (43 % compared to 15 % and 24 %, respectively) indicates that the *rgtD*- mutant is able to add some GalA to the 4-position of the branching Kdo2 residue. The level of terminally linked Kdo in the *rgtD*- mutant is also increased compared to that observed in the parent and *rgtB*- mutant, but less than that observed for either the *rgtE*- or *rgtA*- mutants. Since the high levels of terminally linked Kdo in the latter mutants are due to a complete inability to

add GalA to either the 4- or 5-positions of the branching Kdo2, the relatively reduced level of terminally linked Kdo in the *rgtD*$^{-}$ mutant is likely due to its ability to partially substitute the branching Kdo2 residue at the 4- or both the 4- and 5-positions. The linkage positions of Gal, GalA, and Man are also given in Table 4.3. Gal is 6-linked and GalA is terminally linked as would be expected. The relative level of 6-linked Man (18 %) is very similar to the level of 4,6-linked Man (15 %) indicating that a little less than 50 % of the Man residues is substituted with GalA at the 4-position. Thus, these data indicate that inactivating *rgtD*, in addition causing the loss of GalA on the 4'-position of the lipid A, results in a product that is affected in the addition of GalA to all positions in the core oligosaccharide, possibly by acting as a less than an optimal substrate for RgtA, B, and C. The linkage data indicate that the *rgtD*$^{-}$ mutant is able to produce a mixture of core structures A – F shown in Figure 4.7.

Susceptibility of the rgt *mutants to deoxycholic acid and PmxB*-Because the *rgt* mutants were disrupted in their ability to add GalA to the LPS and it has been previously suggested that negatively charged moieties provide stability to the outer-membrane of gram negative bacteria (28,29), we hypothesized that the membranes may be less stable due to the loss of charge provided by GalA residues. Therefore, we assayed their relative susceptibility to DOC using agar plates prepared with a DOC gradient (Figure 4.8). With the exception of strain EL205 (*rgtC*$^{-}$), the *rgt* mutants were somewhat more susceptible to DOC when compared with the parent strain (Figure 4.8, panel A). Of the susceptible *rgt* mutants, strain EL202 (*rgtE*$^{-}$), which lacks all GalA residues, showed the highest degree of susceptibility while EL203 (*rgtA*$^{-}$), EL204 (*rgtB*$^{-}$), and EL206 (*rgtD*$^{-}$) showed comparable susceptibility to each other and slightly more than the parent. These results suggest that, of the four GalA residues, those on the branched Kdo2 residue and

lipid A contribute to the greatest extent to the stability of the outer-membrane in response to DOC.

In addition, we tested the relative susceptibilities of the *rgt* mutants to the antimicrobial compound PmxB. PmxB is a polycationic antimicrobial peptide that reacts with bacterial LPS due to interactions with the negative charged groups, *e.g.* phosphates, on the LPS as well as hydrophobic interactions. It is known that reduction of the negative charges on the LPS (*e.g.* from *Salmonella*) via introduction of positively charged groups such as 4-aminoarabinose or phosphoethanol amine increase resistance toward PmxB (28,30). Strains EL203 (*rgtA*$^-$), EL204 (*rgtB*$^-$), and EL205 (*rgtC*$^-$) were more resistant to PmxB when compared to parent strain (Figure 4.8, panel B). As described above these strains produce LPS with less GalA on the core region as compared to parent strain (Table 4.3) while maintaining normal amounts of GalA on the lipid A. Strains EL202 (*rgtE*$^-$) and EL206 (*rgtD*$^-$), both of which produce an LPS that lacks GalA on the 4'-position of the lipid A, were more susceptible to PmxB than the parent and the other *rgt* mutants (with EL202 (*rgtE*$^-$) being slightly more susceptible than EL206 (*rgtD*$^-$)). Since removal of GalA reduced the negative charge of the LPS, it was expected that these mutants would have increased resistance to PmxB, and this, indeed appears to be the case for mutants EL203 (*rgtA*$^-$), EL204 (*rgtB*$^-$), and EL205 (*rgtC*$^-$). However, unexpectedly mutants EL202 (*rgtE*$^-$) and EL206 (*rgtD*$^-$) which, respectively, produce LPS that lack the GalA attached to the 4'-position of the lipid A, as well as all or a portion of the core GalA residues, were significantly more sensitive to PmxB than the parent and other *rgt* mutants. Thus the GalA residue attached to the 4'-position of the lipid A appears to be important in conferring resistance, rather than susceptibility, to PmxB.

Symbiotic phenotype between the rgt *mutants and the host plant (Pea)*—Pea plants were inoculated with the *rgtA*$^-$*, B*$^-$*, C*$^-$ and *D*$^-$ mutants in order to determine if disruption of these genes resulted in any obvious symbiotic defects. At 7 days post inoculation (dpi), nodules were formed in plants inoculated with the either the parent strain or the mutant strains. At 21 dpi the parent and mutant strains contained pink nodules indicative of infected and nitrogen-fixing nodules. Thus, it appears that nitrogen-fixing symbiosis occurred in each of these *rgt* mutants.

Interestingly, in the case with the *rgtE*$^-$ mutant, even though at 28 dpi, plants inoculated with parent strain *R. leguminosarum* bv. *viciae* 3841 or *rgtE*$^-$ mutant both displayed pink nodules, developed comparable numbers of nodules, and had equal levels of acetylene reduction indicating that nitrogenase activity in the *rgtE*$^-$ mutant was not impaired, a phenotype was observed. All *rgtE*$^-$ mutant inoculated plants showed discoloration and wilting of the early leaves between 10 and 14 dpi and by 28 dpi all mutant-inoculated plants (40 of 40) showed this wilting response (Figure 4.9). None of the parent strain-inoculated plants had wilted leaves. Detailed microscopy studies on the development of nodules and bacteroids for each of these mutants is required to fully understand any possible symbiotic function of these GalA residues. These more detailed studies are needed particularly in view of the already published data showing that disruption of *acpXL*, the gene that encodes the acyl carrier protein required for the synthesis of the very long chain fatty acid (VLCFA) that is present on the lipid A, also results in nitrogen-fixing nodules, however, microscopic examination showed that bacteroid formation was severely affected (31-33). This work is currently in progress for each of the *rgt* mutants and will be described in an additional manuscript.

Discussion

A number of gram negative bacteria contain GalA as part of their LPS core and/or lipid A. In order to study the functions of these GalA residues with regard to LPS synthesis, symbiosis, pathogenesis, and bacterial physiology, it was necessary to identify and disrupt each of the genes that encode for the various GalATs that add these GalA residues to the LPS. In this study, we describe the identification and function of the *rgtD* and *rgtE* genes from the nitrogen-fixing endosymbiont *Rhizboium leguminosarum* bv. *viciae* 3841. We demonstrate that the *rgtD* gene product is responsible for the addition of GalA to the 4'-position of the lipid A distal GlcN and that the *rgtE* product is responsible for the synthesis of the Dod-P-GalA lipid donor. We also additionally prepared single gene mutations in each of the *rgtA, B,* and *C* genes. Our results, together with the known structure of the *R leguminosarum* bv *viciae* 3841 LPS, allowed us to determine the structures for each of the *rgt* mutant LPSs, demonstrate the order of GalA substitution during LPS synthesis, and determine the effects of each mutation on membrane stability and resistance to the antimicrobial peptide, PmxB.

Analysis of the lipid A from mutant EL206 (*rgtD*$^{-}$) demonstrated the absence of GalA from the 4'-position and, therefore, showed that *rgtD* encodes the GalAT that transfers GalA to this position of the lipid A. Bacteria that contain 4` galacturonosylated lipid A include *R. leguminosarum* and *R. etli* strains, which are members of the *Rhizobiales* (1), the stalk forming *Caulobacter crescentus* (34), and the hyperthermophile *Aquifex aeolicus* (35). Indeed, strains *C. crescentus* CB15 and *A. aeolicus* VF5 contain RgtD homologues (S4.3) and these proteins are likely involved in the synthesis of 4` galacturonosylated lipid A in these bacteria as well. The

information provided in this study concerning the *rgtD* will aid in future studies regarding 4` galacturonylated lipid A in a variety of organisms.

Total lipid extracts from *rgtE*⁻ mutant did not contain detectable amounts of the lipid donor Dod-P-GalA (Figure 4.2) suggesting that the *rgtE* gene is necessary for the synthesis of Dod-P-GalA. In addition, a membrane preparation from the *rgtE*⁻ mutant could not serve as a substrate to support RgtA catalyzed addition of GalA to Kdo_2lipid-IV$_A$, and it produces an LPS devoid of GalA. These results, along with amino acid comparisons with other proteins that synthesize lipid-linked sugars (S4.2), support the conclusion that RgtE is the GalAT responsible for the synthesis of the Dod-P-GalA lipid donor required for the addition of GalA residues to the core region by RgtA, B, C and to the lipid A by RgtD.

The determination of each of the *rgt* mutant core structures allowed us to predict the order of the transfer of the various GalA residues during LPS biosynthesis in *R. leguminosarum* bv. *viciae* 3841; shown in Figure 4.4. Disruption of RgtD, which adds the GalA to the 4'-position of the lipid A perturbed addition of all of the remaining GalA residues, resulting in a mixture of core oligosaccharides containing various structures with and without GalA addition to the branching Kdo2 residue as well as to the Man residue. Since the parent LPS contains stoichiometric levels of all four GalA residues, it is likely that the RgtD product is required as the optimal substrate for the subsequent GalAT and, therefore, that RgtD adds the first of the LPS GalA residues. In addition, for all of the remaining *rgtABC*⁻ mutants the 4'-lipid A GalA residue is present showing that RgtD activity is not affected in any of these mutants. Disruption of RgtA results in the complete absence of GalA on both the 4- and 5- positions of the branching Kdo2 residue, and incomplete addition of GalA to the Man residue, while disruption of RgtB results in addition of GalA to the 4-position of Kdo2 and nearly complete addition of GalA to the Man

residue. This implies that the RgtA activity is needed prior to RgtB and, further, that RgtA adds GalA to the 4-position of Kdo2 and, therefore, RgtB adds GalA to the 5-position of Kdo2. In addition, since a small percentage of the Man residue in the *rgtB* mutant LPS lacks GalA, the likely order is RgtA, followed by RgtB, and then RgtC. This is further supported by the observation that, while disruption of RgtC eliminated addition of GalA to the Man residue, it did not affect addition of GalA to the 4- and 5-positions of Kdo2. Thus, in *R. leguminosarum* bv. viciae 3841, the predicted order for the transfer of the GalA residues, as shown in Figure 4.4, is as follows: Transfer to the 4'-position of the lipid A by RgtD, followed by transfer to the 4-position of the branching Kdo2 by RgtA, then to the 5-position of the Kdo2 residue by RgtB, and lastly, to the 4-position of the Man residue by RgtC.

With regard to membrane stability, all of the *rgt* mutants were more sensitive to DOC than the parent strain (Figure 4.8, panel B) except for EL205 (*rgtC*$^-$). Strain EL202 (*rgtE*$^-$) was the most sensitive while EL203 (*rgtA*$^-$), EL204 (*rgtB*$^-$), and EL206 (*rgtD*$^-$) were comparable to each other in their DOC susceptibility, but less sensitive than EL202 (*rgtE*$^-$). The complemented strains recovered parent strain levels of DOC resistance. These results suggest that the GalA residues present on the branching Kdo2 residue and on the 4'-position of the lipid A distal GlcN confer membrane stability towards disruption by DOC. It is possible, that the presence of GalA contributes to the cross linking of LPS molecules in the outer membrane through ionic bridging between the negatively charged carboxyl moiety and divalent cations such as Ca^{2+} and Mg^{2+}. Typically, as is the case with enteric bacteria, lipid A is *bis*-phosphorylated at the 1 and 4` positions and in some cases on inner core residues. These negatively charged phosphates contribute to membrane stability through LPS / divalent cation bridging (29,36). *Rhizobium leguminosarum* and *R. etli* lipid A does not contain the 1 and 4` phosphates but, rather, contains

4`-GalA and, to some degree, proximal 2-aminogluconic acid (Figure 4.4), both negatively charged moieties. The GalA residues likely add negative charge *in lieu* of phosphate and results presented in this study suggest that the LPS GalA residues function similarly to phosphate groups in stabilizing the outer membrane.

In addition to affecting sensitivity to DOC, the *rgt* mutants varied in their susceptibility to PmxB). PmxB is an acylated polycationic cyclic peptide that interacts with the negatively charged moieties and hydrophobic acyl portion of LPS. Mutant strains EL203 (*rgtA*$^-$), EL204 (*rgtB*$^-$), and EL205 (*rgtC*$^-$) were somewhat more resistant to PmxB compared to parent strain. In the case of these mutants, reduction of the anionic GalA residues would be expected to reduce interaction with the cationic PmxB and, therefore, it is not surprising that increased PmxB resistance was observed. However, strains EL202 (*rgtE*$^-$) and EL206 (*rgtD*$^-$) were significantly less resistant to PmxB compared to parent strain. These strains differ from the other *rgt* mutant strains in that the LPSs from both completely lack the lipid A 4`-GalA residue while the other *rgt* mutants and parent strain maintain this residue. It is possible that the lipid A 4'-GalA residue is a major contributor for conferring membrane stability against disruption by PmxB, and, therefore, even though the absence of GalA may reduce PmxB binding, the greater reduction in membrane stability results in increased disruption of the outer membrane and, therefore, increased sensitivity to PmxB.

It is unclear why *R. leguminosarum* and *R. etli* strains modify their LPS with GalA. Perhaps bacteria that attach GalA instead of phosphate have a competitive advantage in phosphate limiting environments. In addition, GalA is a common component of the plant cell wall (e.g. such as galacturonan polysaccharides) and the presence of GalA on the LPS may provide host mimicry allowing it to escape or manipulate the host's defense response. Because

bacteroids maintain a tight association with the plant derived peribacteroid membrane, it is tempting to speculate that LPS GalA residues may interact in an ionic manner with the peribacteroid membrane during endosymbiosis to help maintain this tight association. Furthermore, the GalA residues present on the LPS may aid in the recruitment of plant defensin-like molecules. The causative agent for terminal bacteroid differentiation of *Sinorhizobum meliloti* was recently reported (37) to be host-derived defensin like peptides, NCRs (nodule-specific cysteine rich peptides). NCRs act at the cell surface and cause elongation, branching, and endoreduplication of free-living *S. meliloti*. A similar process likely occurs during symbiosis between *R. leguminosarum* bv. *viciae* 3841 and its host, *Pisum sativum* (pea), and it is possible that LPS GalA residues contribute to the interaction between pea NCRs and the bacterial cell surface. Further work is in progress to determine the role of the LPS GalA residues with regard to its interaction with pea defensins.

Acknowledgements

*****_The authors acknowledge Hak Suk Chung, Ziqiang Guan, and Chris Raetz for their help in preparing and performing LC-ESI-MS analysis of the total membrane extracts and in performing the RgtA assays._

References

1. Carlson, R. W., Forsberg, L. S., and Kannenberg, E. L. (2010) Lipopolysaccharides in *Rhizobium*-Legume Symbioses. in *Subcellular Biochemestry: Endotoxins: Structure, Function and Recognition*, P. J. Quinn and X. Wang Ed., Springer. pp 339-386
2. Forsberg, L. S., and Carlson, R. W. (1998) *J. Biol. Chem.* **273**, 2747-2757
3. Kanjilal-Kolar, S., and Raetz, C. R. (2006) *J. Biol. Chem.* **281**, 12879-12887
4. Kanjilal-Kolar, S., Basu, S. S., Kanipes, M. I., Guan, Z., Garrett, T. A., and Raetz, C. R. (2006) *J. Biol. Chem.* **281**, 12865-12878
5. Karbarz, M. J., Six, D. A., and Raetz, C. R. H. (2009) *J. Biol. Chem.* **284**, 414-425
6. Price, N. P., Jeyaretnam, B., Carlson, R. W., Kadrmas, J. L., Raetz, C. R., and Brozek, K. A. (1995) *Proc. Natl. Acad. Sci.* **92**, 7352-7356
7. Que-Gewirth, N. L. S., Karbarz, M. J., Kalb, S. R., Cotter, R. J., and Raetz, C. R. H. (2003) *J. Biol. Chem.* **278**, 12120-12129
8. Beringer, J. E. (1974) *Microbiol.* **84**, 188
9. Becker, A., Schmidt, M., Jager, W., and Puhler, A. (1995) *Gene* **162**, 37-39
10. Hoang, T. T., Karkhoff-Schweizer, R. A. R., Kutchma, A. J., and Schweizer, H. P. (1998) *Gene* **212**, 77-86
11. Westphal, O., and Jann, K. (1965) *Meth. Carbohydr. Chem.* **5**, 83–91
12. Fraysse, N., Jabbouri, S., Treilhou, M., Couderc, F., and Poinsot, V. (2002) *Glycobiol.* **12**, 741-748
13. Reuhs, B. L., Carlson, R. W., and Kim, J. S. (1993) *J. Bacteriol.* **175**, 3570-3580
14. Kittelberger, R., and Hilbink, F. (1993) *J. Biochem. Biophys. Meth.* **26**, 81-86

15. York, W. S., Darvill, A. G., McNeil, M., Stevenson, T. T., and Albersheim, P. (1986) *Meth. Enzymol.* **118**, 3-40

16. Bhat, U. R., Forsberg, L. S., and Carlson, R. W. (1994) *J. Biol. Chem.* **269**, 14402

17. Ryan, J. M., Conrad, H. E. (1974) *Arch. Biochem. Biophys.* **162**, 530-536

18. Forsberg, L. S., Noel, K. D., Box, J., and Carlson, R. W. (2003) *J. Biol. Chem.* **278**, 51347-51359

19. Ciucanu, I., and Kerek, F. (1984) *Carbohydr. Res.* **131**, 209-217

20. Cointe, D., Leroy, Y., and Chirat, F. (1998) *Carbohydr. Res.* **311**, 51-59

21. Anumula, K. R., and Taylor, P. B. (1992) *Anal. Biochem.***203**, 101-108

22. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucl. Acids Res.***25**, 3389-3402

23. Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., Fong, J. H., Geer, L. Y., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Jackson, J. D., Ke, Z., Lanczycki, C. J., Lu, F., Marchler, G. H., Mullokandov, M., Omelchenko, M. V., Robertson, C. L., Song, J. S., Thanki, N., Yamashita, R. A., Zhang, D., Zhang, N., Zheng, C., and Bryant, S. H. *Nucl. Acids Res.* **39**, D225-D229.

24. Trent, M. S., Ribeiro, A. A., Lin, S., Cotter, R. J., and Raetz, C. R. H. (2001) *J. Biol. Chem.* **276**, 43122-43131

25. Trent, M. S., Ribeiro, A. A., Lin, S., Cotter, R. J., and Raetz, C. R. H. (2001) *J. Biol. Chem.* **276**, 43122-43131

26. Song, F., Guan, Z., and Raetz, C. R. (2009) *Biochem.* **48**, 1173-1182

27. Wang, X., Ribeiro, A. A., Guan, Z., and Raetz, C. R. H. (2009) *Biochem.* **48**, 1162-1172

28. Raetz, C. R. H., Reynolds, C. M., Trent, M. S., and Bishop, R. E. (2007) *Annu. Rev. Biochem.* **76**, 295-329

29. Snyder, S., Kim, D., and McIntosh, T. J. (1999) *Biochem.* **38**, 10758-10767

30. Trent, M. S., Stead, C. M., Tran, A. X., and Hankins, J. V. (2006) *J. Endotoxin Res.* **12**, 205-223

31. Vedam, V., Haynes, J. G., Kannenberg, E. L., Carlson, R. W., and Sherrier, D. J. (2004) *Mol. Plant-Microbe Interact.* **17**, 283-291

32. Brown, D. B., Huang, Y.-C., Kannenberg, E. L., Sherrier, D. J., and Carlson, R. W. (2011) *J. Bacteriol.* **193**, 4766-4778.

33. Haag, A. F., Wehmeier, S., Beck, S., Marlow, V. L., Fletcher, V., James, E. K., and Ferguson, G. P. (2009) *J. Bacteriol.* **191**, 4681-4686

34. Smit, J., Kaltashov, I. A., Cotter, R. J., Vinogradov, E., Perry, M. B., Haider, H., and Qureshi, N. (2008) *Inn. Immun.***14**, 25-36

35. Plotz, B. M., Lindner, B., Stetter, K. O., and Holst, O. (2000) *J. Biol. Chem.* **275**, 11222-11228

36. Raetz, C. R., and Whitfield, C. (2002) *Annu. Rev. Biochem.* **71**, 635-700

37. Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., Farkas, A., Mikulass, K., Nagy, A., Tiricz, H., Satiat-Jeunemaitre, B., Alunni, B., Bourge, M., Kucho, K., Abe, M., Kereszt, A., Maroti, G., Uchiumi, T., Kondorosi, E., and Mergaert, P. (2010) *Science* **327**, 1122-1126

38. Wang, T. L., Wood, E. A., and Brewin, N. J. (1982) *Planta* **155**, 345-349

39. Phadnis, S. H., and Das, H. K. (1987) *J. Biosci.* **12**, 131-135

Footnotes

*This work was supported, in whole or in part, by National Institutes of Health Grant GM39583 (to R.W.C.) and by the Department of Energy Grant DE-FG-02-93ER20097 to the CCRC. This work is also in partial fulfillment for the PhD degree of D.B.B.

Abbreviations

LPS, lipopolysaccharide; Kdo, 3-deoxy-D-manno-2 octulosonic acid;3Me 6dTal, 3-O-methyl-6-deoxytalose; QuiNAc, quinavosamine; GlcN, glucosamine; Gal, galactose; GalA, galacturonic acid; Man, mannose; Glc, glucose; GlcNonate, 2-aminogluconate; BHB, β-hydroxybutyrate; β-OHC14:0, β-hydroxy-myristate; β-OHC16:0, β-hydroxypalmitate; β-OHC18:0, β-hydroxystearate; Dod-P-GalA, dodecaprenyl phosphate galacturonic acid; LC-ESI-MS/MS, liquid chromatography electrospray ionization tandem mass spectrometry; MALDI-TOF-MS, matrix assisted desorption ionization time of flight mass spectrometry; GC/MS, combined gas chromatography\mass spectrometry; DOC-PAGE, deoxycholic acid polyacrylamide gel electrophoresis; PmxB, polymyxin B.

Figure Legends

Figure 4.1. The general structure of the *R. leguminosarum* bv *viciae* 3841 LPS core and lipid A regions. GlcN = 2-aminoglucosamine, GlcNonate = 2-aminogluconic acid, Kdo = 3-deoxy-d-*manno*-2-octulosonic acid, GalA = galacturonic acid, Man = mannose, Gal = galactose, 1^{o} OH-FA = the primary β -hydroxy fatty acids are attached to the GlcN backbone of the lipid A, 2^{o} FA = the secondary very long chain fatty acid (VLCFA), 27OHC28, is attached to the 3` 1^{o} OH-FA to form an acyloxyacyl moiety, BHB = β-hydroxybutyryl group. The proximal GlcN can be oxidized to GlcNonate by the outermembrane oxidase LpxQ. Position 27 of the VLCFA can contain a hydroxyl or BHB group. As described in this study, RgtE is required for the synthesis of the lipid donor dod-P-GalA which can then be utilized by the galacturonosyl transferases RgtA, B, C, and D in the transfer of GalA to lipid A and Core where indicated.

Figure 4.2. LC-ESI-MS of membrane lipid components. Total lipids were examined for the appearance of Dod-P-GalA by LC-ESI-MS. An ion scan for 1089 was performed on data acquired from the total lipid extracts of 3841 (C), EL202 (B), and complemented strain (C). The ion scan revealed the retention time of Dod-P-GalA to be 20 minutes where a sharp peak was observed on the LC chromatogram (not shown). Dod-P-GalA ions were not detectable in the total lipids from EL202. MS/MS data confirmed the identification of the labeled ions (not shown).

Figure 4.3. RgtA activity assay with total lipid extracts. Total lipids were tested for their ability to serve as a GalA donor in an activity assay with RgtA over expressed in *E. coli*

membranes and the radio labeled acceptor substrate 1-dephospho-[4`-^{32}P]lipid IVa (Structure I). Standard reaction mixtures are described in the experimental procedures section. Products were separated by thin layer chromatography and viewed by a radio imager. Lanes 1 and 2 lack enzyme. Lanes 3-7contain, as an enzyme source, membranes from *E. coli* over expressing RgtA. Lipid substrate was added where indicated by (+). The product in Lane 3 was identified by Kanjilal-Kolar et. al (3,4) to be GalA modified 1-dephospho-[4`-^{32}P]lipid IVa (Structure II). Conversion of substrate was not observed from reactions containing EL202 lipids (Lane 6) indicating the lack of the necessary lipid donor (dod-P-GalA) required for activity. Lipids from EL202 complemented with the *rgtE* gene (EL202/pMKGE) were able to recover RgtA activity (Lane 7).

Figure 4.4. DOC-PAGE LPS profiles. Extracted LPS (1μg) was loaded into each and stained as described in the Experimental Methods section. Lane 1) Parent strain 3841. Lane 2) Strain EL203 (*rgtA*$^-$). Lane 3) Strain EL204 (*rgtB*$^-$). Lane 4) Strain EL205 (*rgtC*$^-$). Lane 5) Strain EL206 (*rgtD*$^-$). Lane 6) Strain EL202 (*rgtE*$^-$). Arrowheads indicate the reduction of an LPS band from strains EL202 (*rgtE*$^-$) and EL206 (*rgtD*$^-$) that is present in the other LPS DOC-PAGE profiles.

Figure 4.5. MALDI-TOF MS analysis of the *rgtD*$^-$ and *rgtE*$^-$ lipid A. Spectra were acquired in the positive reflectron mode. I.) Parent Strain 3841. II.) Mutant strain EL202 (*rgtE*$^-$). III.) Mutant strain EL206 (*rgtD*$^-$). Structures A-F indicated ions observed in the spectra and the calculated exact masses (Calc. EM) are labeled. Variations occur to structures A-F by the presence or absence of BHB (-86) and variation in the chain length of the hydroxy

fatty acids (+/- 28 and/or 14). Panels II and III contain ions with masses consistent with the absence of GalA (-176 mass units).

Figure 4.6. GC/MS of the EL206 (*rgtD*$^{-}$) mutant lipid A. A) Parent strain 3841. B) Mutant strain EL206 (*rgtD*$^{-}$). Trimethylsilyl methylglycosides and fatty acid methyl esters were derived from partitioned lipid A and analyzed by GC/MS. The lipid A from the *rgtA*$^{-}$, *B*$^{-}$, and *C*$^{-}$ mutants gave profiles similar to that of the parent strain, while the lipid A from the rgtE^{-} mutant gave a GC/MS profile similar to that of the *rgtD*$^{-}$ mutant; i.e. it was devoid of GalA residues. GlcNAc is derived from the lipid A GlcN residues.

Figure 4.7. The proposed core structures of the LPS from *rgt* mutants based on glycosyl linkage analysis, as described in the text. Dashed lines represent non-stoichiometric amounts of the attached residue.

Figure 4.8. Susceptibility of the *rgt* mutants to deoxycholic acid (DOC) and PmxB (PmxB). Panel A – Growth of bacteria on DOC gradient plates. 1) Parent strain 3841. 2) Mutant strain EL203 (*rgtA*$^{-}$). 3) Mutant strain EL204 (*rgtB*$^{-}$). 4) Mutant strain EL205 (*rgtC*$^{-}$) 5) Mutant strain EL206 (*rgtD*$^{-}$). 6) Mutant strain EL202. Panel B – Growth of bacteria on PmxB gradient plates. 1) Parent strain 3841. 2) Mutant strain EL203 (*rgtA*$^{-}$). 3) Mutant strain EL204 (*rgtB*$^{-}$). 4) Mutant strain EL205 (*rgtC*$^{-}$) 5) Mutant strain EL206 (*rgtD*$^{-}$). 6) Mutant strain EL202 (*rgtE*$^{-}$).

Figure 4.9. The symbiotic phenotype of plants inoculated with the *rgtE*⁻ mutant inoculated. A) A representative plant, of 40 plants, at 21 dpi inoculated with the parent strain 3841 . B) A representative plant, of 40 plants, at 21d.i inoculated with mutant EL202. The early leaves of all 40 mutant inoculated plants displayed necrosis between 7 and 14 dpi and were completely wilted by 28 dpi. All of the parent strain 3841inoculated plants maintained green non-wilted leaves even at 28 dpi.

Figure 4.10. Schematic diagram showing the biosynthetic steps for *R. leguminosarum* bv. *viciae* LPS and indicating the order by which the GalA residues are added to the lipid A and core oligosaccharide. The LPS core backbone is built off of the lipid A on the cytosolic side of the innermembrane and is flipped to the periplasmic side of the innermembrane where the transmembrane Rgt proteins transfer GalA from the lipid donor Dod-P-GalA to the LPS (1). The LPS is transported to the outer membrane where the oxidase LpxQ oxidizes the proximal GlcN to GlcNonate. The Rgt proteins are labeled 1-4 to signify the predicted order of GalA addition. LpcA, B, and C = glycosyl transferases, RgtA, B, C, D, and E = GalA transferases. The boxed region shows the synthesis of Dod-P-GalA from UDP-GalA by RgtE.

Table 4.1. Bacterial Strains and Plasmids

Strains and Plasmid	Characteristics	Source
E. coli		
Top10	F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *araD*139 Δ(*araleu*) 7697 *gal*U *gal*K *rps*L (Str^r) *end*A1 *nup*G	Invitrogen
R. lequminosarum* bv. *viciae		
3841	Strain 300 Str^r , Nod^+, Fix^+	(38)
EL203 (*rgtA*$^-$)	Strain 3841 *rgtA::accC1*, Gm^r Nod^+	This Study
EL204 (*rgtB*$^-$)	Strain 3841 *rgtB::accC1*, Gm^r Nod^+	This Study
EL205 (*rgtC*$^-$)	Strain 3841 *rgtC::accC1*, Gm^r Nod^+	This Study
EL206 (*rgtD*$^-$)	Strain 3841 *rgtD::accC1*, Gm^r Nod^+	This Study
EL202 (*rgtE*$^-$)	Strain 3841 *rgtE::accC1*, Gm^r, Nod^+, Fix^+	This Study
EL203 (*rgtA*$^-$)/pRK-RgtA	Strain EL203 (*rgtA*$^-$) complemented with plasmid pRK-RgtA (4), Gm^rTc^r	This Study
EL204 (*rgtB*$^-$)/pRK-RgtB	Strain EL204 (*rgtB*$^-$) complemented with plasmid pRK-RgtB (4), Gm^r Tc^r	This Study
EL205 (*rgtC*$^-$)/pRK-RgtC	Strain EL205 (*rgtC*$^-$) complemented with plasmid pRK-RgtC (4), Gm^r Tc^r	This Study
EL202 (*rgtE*$^-$)/pMKGE	Strain EL202 (*rgtE*$^-$) complemented with plasmid pMKGE (4), Gm^r Tc^r	This Study
Plasmids		
pUC18	Cloning vector, Am^r	Fermentas
pMS255	Source of *aacC1 Gm* cassette, Gm^r	(9)
pRK2013	Mobilizing plasmid for pEX-Tc, Col E1 replicon, Kan^r	(39)
pEX18-Tc	Suicide vector, allows positive selection for integration, Tc^r	(10)
pRgtA-KO	pEX18 containing *rgtA::aacC1* insert used to engineer EL203 (*rgtA*$^-$), Gm^r, Tc^r	This Study
pRgtB-KO	pEX18 containing *rgtB::aacC1* insert used to engineer EL204 (*rgtB*$^-$), Gm^r, Tc^r	This Study
pRgtC-KO	pEX18 containing *rgtC::aacC1* insert used to engineer EL205 (*rgtC*$^-$), Gm^r, Tc^r	Study
pRgtD-KO	pEX18 containing *rgtD::aacC1* insert used to engineer EL206 (*rgtD*$^-$), Gm^r, Tc^r	This Study
pRgtE-KO	pEX18 containing *rgtE::aacC1* insert used to engineer EL202 (*rgtE*$^-$), Gm^r, Tc^r	This Study

Table 4.2. Primer List

Gene	Primer*
RgtAUp	Fwd ACGTGGATCC*GCAGGTGAAGCTGATGG*
RgtAUp	Rev ACGTGAATTC*GCAGCACCAGCAGGAAG*
RgtADwn	Fwd ACGTTCTAGA*CTGGGTGCCCGACAATC*
RgtADwn	Rev ACGTGGATCC*GGATCTTGCGCAGGAAC*
RgtBUp	Fwd ACGTTCTAGATTTCGCGCTGATGCCTG
RgtBUp	Rev ACGTGGATCC*CTTCTCCAGTGTGCGTG*
RgtBDwn	Fwd ACGTGGATCC*GCTCTCGCCATCATCA*
RgtBDwn	Rev ACGTGAATTC*TGGCCCGCGTCTATTG*
RgtCUp	Fwd ACGTGAATTC*GCATGCATCGGCAGCA*
RgtCUp	Rev ACGTCCCGGG*TCGGCAGTCATTCGC*
RgtCDwn	Fwd ACGTCCCGGG*TTTTTCGCGCGCTCT*
RgtCDwn	Rev ACGTAAGCTT*GTGCCGATGTCGTTGC*
RgtDUp	Fwd ACGTTCTAGA*GAGATCCCTGAGGCCTG*
RgtDUp	Rev ACGTGGATCC*GCCGCACCGTCGTATTG*
RgtDDwn	Fwd ACGTGGATCC*TCAGAGGCAGGAGCGAC*
RgtDDwn	Rev ACGTGAGCTC*AGCGCCTTCAGCAGCAC*
RgtEUp	Fwd ACGTTCTAGA*CTGGCTGCCCGACAATC*
RgtEUp	Rev ACGTGGATCC*GGATCTTGCGCAGGAAC*
RgtEDwn	Fwd ACGTGGATCC*GCAGGTGAAGCTGATGG*
RgtEDwn	Rev ACGTGAATTC*GCAGCACCAGCAGGAAG*

*Underlined regions indicate the engineered endonuclease restriction sites. Italic letters indicate regions homologous to *R. leguminosarum* bv. *viciae* 3841 genomic dna.

Table 4.3. Linkage analysis of the *rgt* mutant LPSs

	89 *m/z* ion extraction				118 *m/z* ion extraction		
Strains	T-Kdo	4-Kdo	5-Kdo	4,5-Kdo	6-Gal (T-GalA)[a]	6-Man	4,6-Man
3841 (Parent Strain)	9.8	28	0.0	62	77	0.0	23
EL202 (*rgtE*$^-$)	49	15	0.0	36	51	49	0.0
EL203 (*rgtA*$^-$)	45	24	0.0	31	66	14	20
EL204 (*rgtB*$^-$)	6.2	63	0.0	31	66	2.9	31
EL205 (*rgtC*$^-$)	16	22	0.0	62	77	18	4.4
EL206 (*rgtD*$^-$)	20	43	0.0	36	67	18	15

Permethylated alditol acetates (PMAA) were prepared from isolated LPS and analyzed by GC/MS in order to determine carbohydrate linkages. The relative abundance of Kdo residues having different linkages was estimated by selective ion monitoring for the diagnostic *m/z* 89 primary fragment ion, derived from carbons 7 and 8 of each Kdo derivative. The relative abundance of Gal/GalA and Man residues was determined from the intensity of the shared *m/z* 118 primary ion. [a] 6 linked Gal and terminal GalA can be distinguished by the relative abundance of ions *m/z* 189/191, and *m/z* 233/235. T = terminally linked residue. Numbers before each residue (i.e. 4-Kdo) represents the linked position. Derivatives and primary ion fragments: 6-linked Man: 1,5,6-tri-*O*-acetyl-2,3,4-tri-*O*-methyl-mannitol (*m/z* 118, 162, 189, 233); 4,6-linked Man: 1,4,5,6-tetra-*O*-acetyl-2,3-di-*O*-methyl-mannitol (*m/z* 118, 261); 6-linked Gal: 1,5,6-tri-*O*-acetyl-2,3,4-tri-*O*-methyl-galactitol (*m/z* 118, 162, 189, 233); T-GalA: 1,5,6-tri-*O*-acetyl-2,3,4-tri-*O*-methyl-6,6-^{2}H-galactitol (*m/z* 118, 162, 191, 235); T-Kdo (1,2,6-tri-*O*-acetyl-3-deoxy-4,5,7,8-tetra-*O*-methyl-1,1-^{2}H-octitol (*m/z* 89, 146, 205, 206, 250, 366); 4,5-linked Kdo (1,2,4,5,6-penta-*O*-acetyl-3-deoxy-7,8-di-*O*-methyl-1,1-^{2}H-octitol (*m/z* 89, 186, 228, 348, 422); 4-linked Kdo (1,2,4,6-tetra-*O*-acetyl-3-deoxy-5,7,8-tri-*O*-methyl-1,1-^{2}H-octitol (*m/z* 89, 205, 278, 320, 394).

Supplementary Figure Legends

Supplemental 4.1. Amino acid sequence alignment of the RgtA, B, C, and D polypeptides. The alignmentwas performed using the multiple sequence alignment scoring matrix BLOSUM 62 application in the Clone Manager program. Shaded letters represent sequence identity.

Supplemental 4.2. Relationship of RgtE to ArnC, FlmF1 and FlmF2. The RgtE amino acid sequence shares similar qualities with known bactoprenyl glycosyl transferases that function in the synthesis of monosaccharide lipid donors which are utilized in the biosynthesis of LPS. Like the known glycosyl transferases ArnC, FlmF1,and FlmF2, the RgtE amino acid sequence contains a glycosyl transferase 2 (GT2) domain which contains the signature DXD catalytic motif followed by two signature putative C-terminaltransmembrane (TM) domains. The enzymes are likely anchored to the inner membrane by the transmembrane domains leaving the catalytic domain (GT2) exposed to the cytoplasmic/inner membrane interface.

Supplemental 4.3. Amino acid sequence alignment of the *R. leguminosarum* bv. *viciae* RgtD and putative RgtD proteins in *Caulobacter crescentus* CB15 and *Aquifex aeolicus* VF5. Alignments were made in the Clone Manager program using the BLOSUM 62 matrix scoring system. Shaded letters are identical amino acids. The putative RgtD peptides share over 50% similarity (positives) with the *Rhizobium* RgtD and are share highly conserved regions towards the N-terminus.

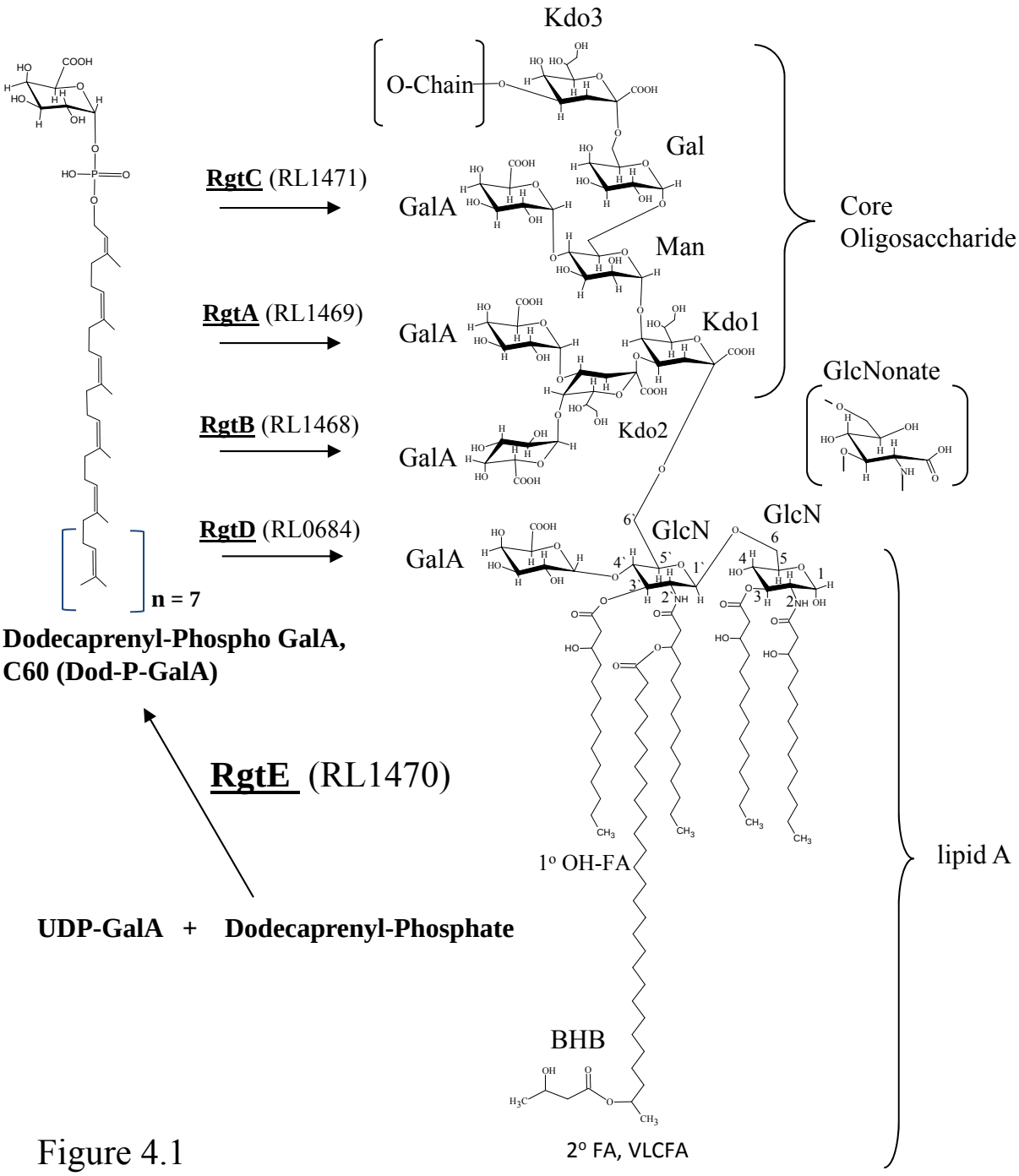

Kdo3
O-Chain
Gal
RgtC (RL1471)
GalA
Man
Core
Oligosaccharide
RgtA (RL1469)
GalA
Kdo1
GlcNonate
RgtB (RL1468)
GalA
Kdo2
RgtD (RL0684)
GalA
GlcN
GlcN
n = 7
Dodecaprenyl-Phospho GalA,
C60 (Dod-P-GalA)
RgtE (RL1470)
UDP-GalA + Dodecaprenyl-Phosphate
1º OH-FA
lipid A
BHB
2º FA, VLCFA

Figure 4.1

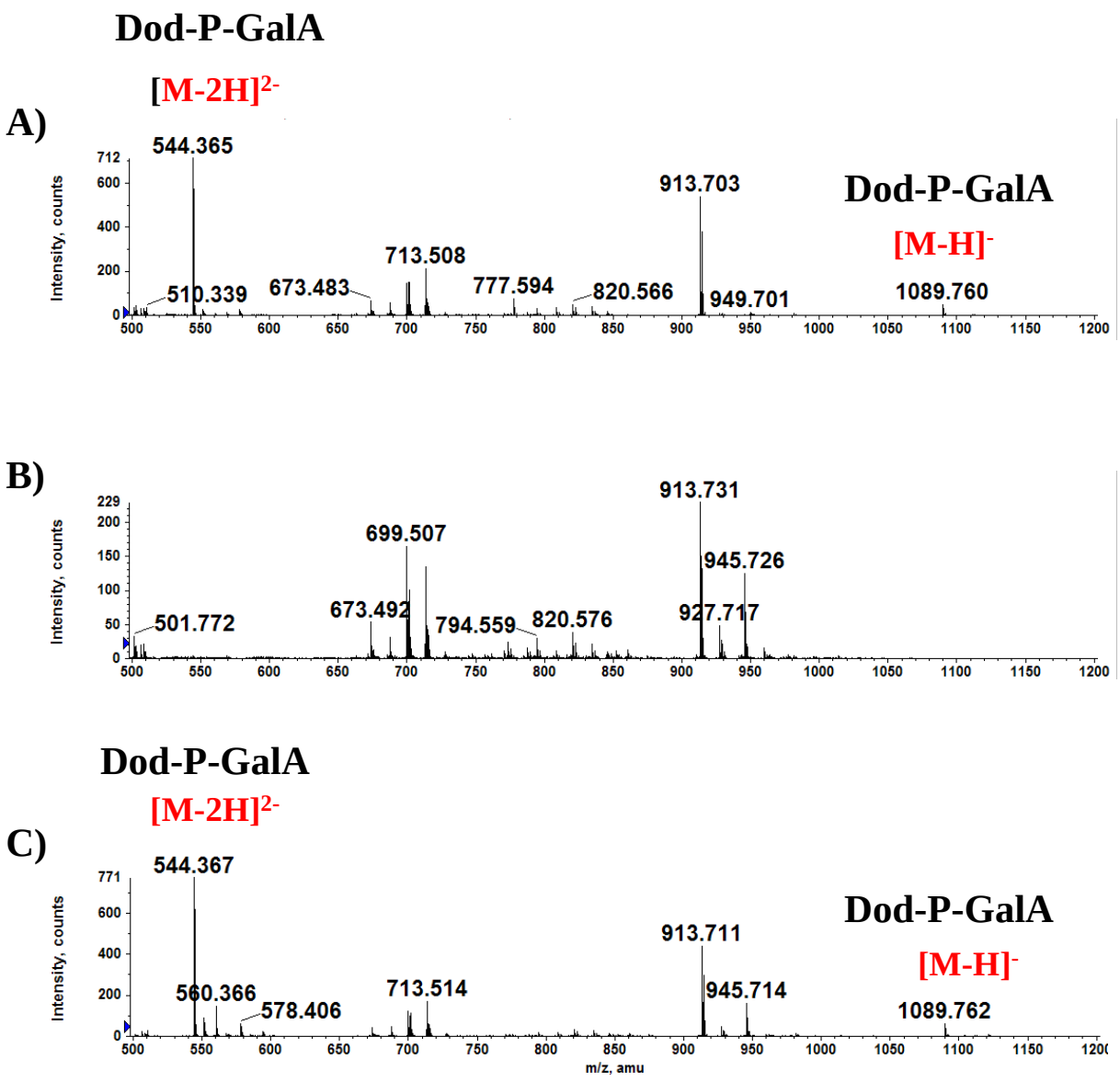
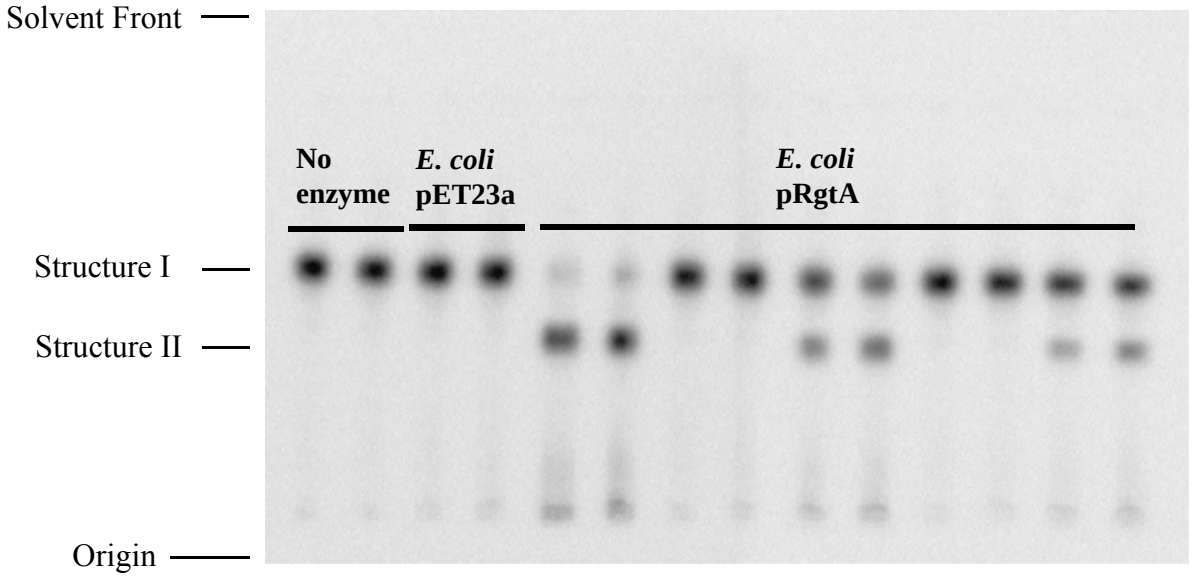

Dod-P-GalA
[M-2H]2-
A)
544.365
913.703
Dod-P-GalA
[M-H]-
713.508
510.339
673.483
777.594
820.566
949.701
1089.760
Intensity, counts
B)
913.731
699.507
945.726
673.492
501.772
794.559
820.576
927.717
Dod-P-GalA
[M-2H]2-
C)
544.367
Dod-P-GalA
[M-H]-
913.711
560.366
713.514
578.406
945.714
1089.762
m/z, amu

Figure 4.2



Time (min)	10 30	10 30	10 30	10 30	10 30	10 30	10 30
	1	2	3	4	5	6	7
Purified donor	+	+	+	-	-	-	-
Totol lipid extracts from							
3841	-	-	-	-	+	-	-
EL202	-	-	-	-	-	+	-
EL202/pMKGE	-	-	-	-	-	-	+

Figure 4.3

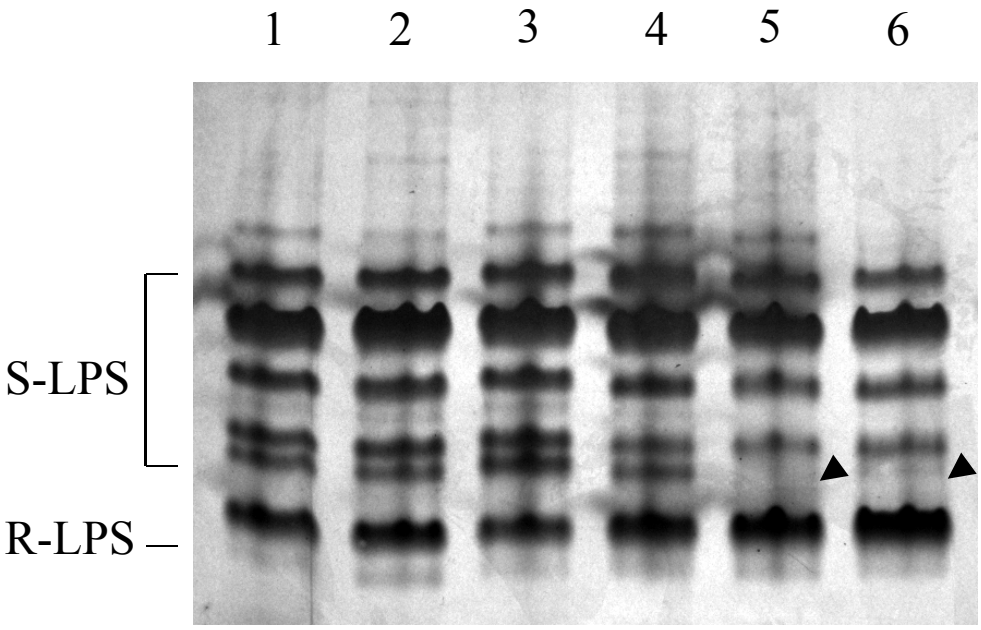


Figure 4.4

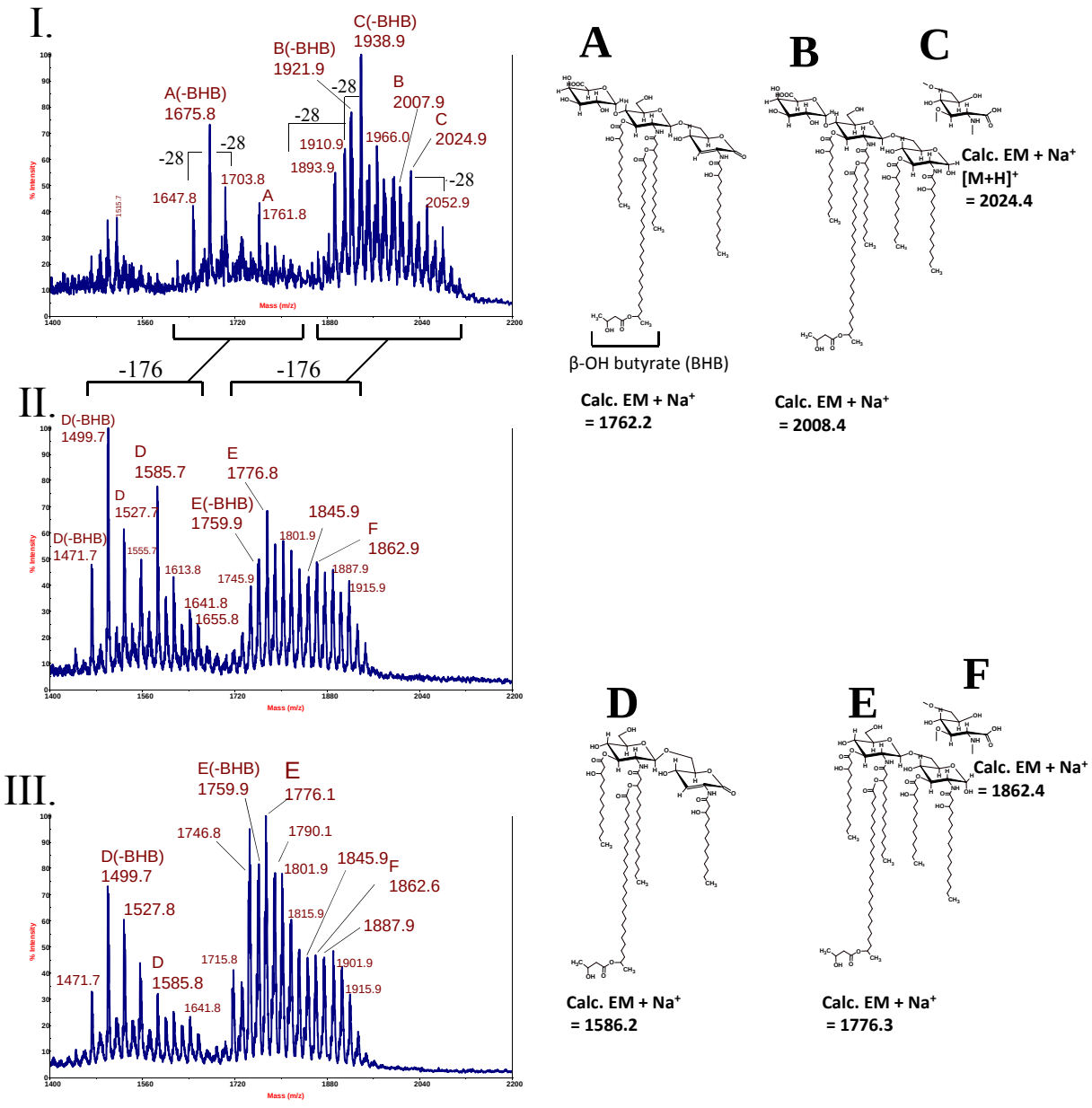
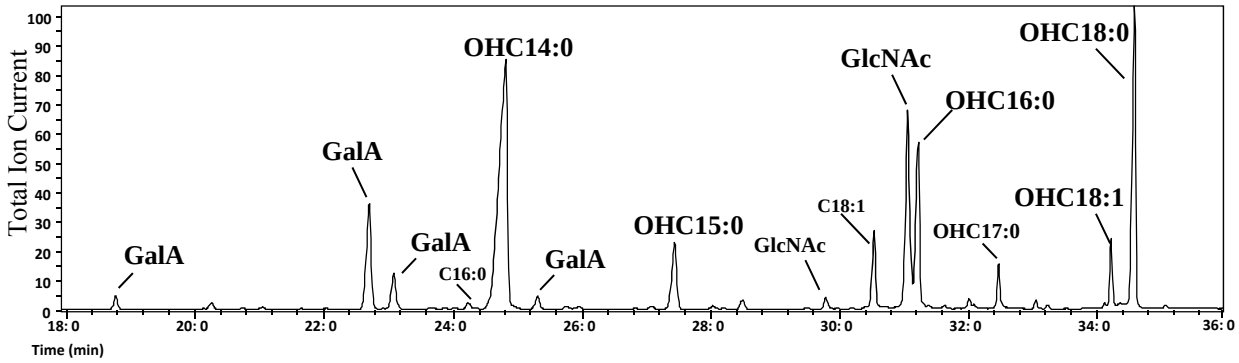

I.
C(-BHB)
1938.9
B(-BHB)
1921.9
A(-BHB)
1675.8
B
2007.9
C
2024.9
-28
1910.9
1966.0
1893.9
1703.8
1647.8
A
1761.8
2052.9
-176
-176
II.
D(-BHB)
1499.7
D
1585.7
E
1776.8
D
1527.7
E(-BHB)
1759.9
1845.9
F
1862.9
D(-BHB)
1471.7
1801.9
1887.9
1915.9
1613.8
1745.9
1641.8
1655.8
III.
E(-BHB)
1759.9
E
1776.1
1746.8
1790.1
D(-BHB)
1499.7
1801.9
1845.9
F
1862.6
1527.8
1815.9
1887.9
D
1585.8
1715.8
1901.9
1471.7
1915.9
1641.8
A
B
C
β-OH butyrate (BHB)
Calc. EM + Na+
= 1762.2
Calc. EM + Na+
= 2008.4
Calc. EM + Na+
[M+H]+
= 2024.4
D
E
F
Calc. EM + Na+
= 1586.2
Calc. EM + Na+
= 1776.3
Calc. EM + Na+
= 1862.4

Figure 4.5

A


Total Ion Current
Time (min)
GalA
GalA
GalA
C16:0
OHC14:0
OHC15:0
GlcNAc
C18:1
GlcNAc
OHC16:0
OHC17:0
OHC18:1
OHC18:0

B

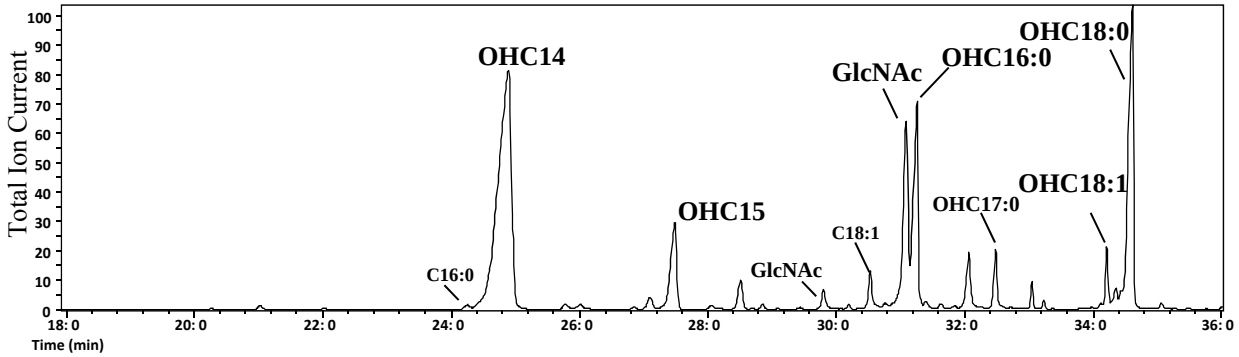

Total Ion Current
Time (min)
C16:0
OHC14
OHC15
GlcNAc
C18:1
GlcNAc
OHC16:0
OHC17:0
OHC18:1
OHC18:0

Figure 4.6

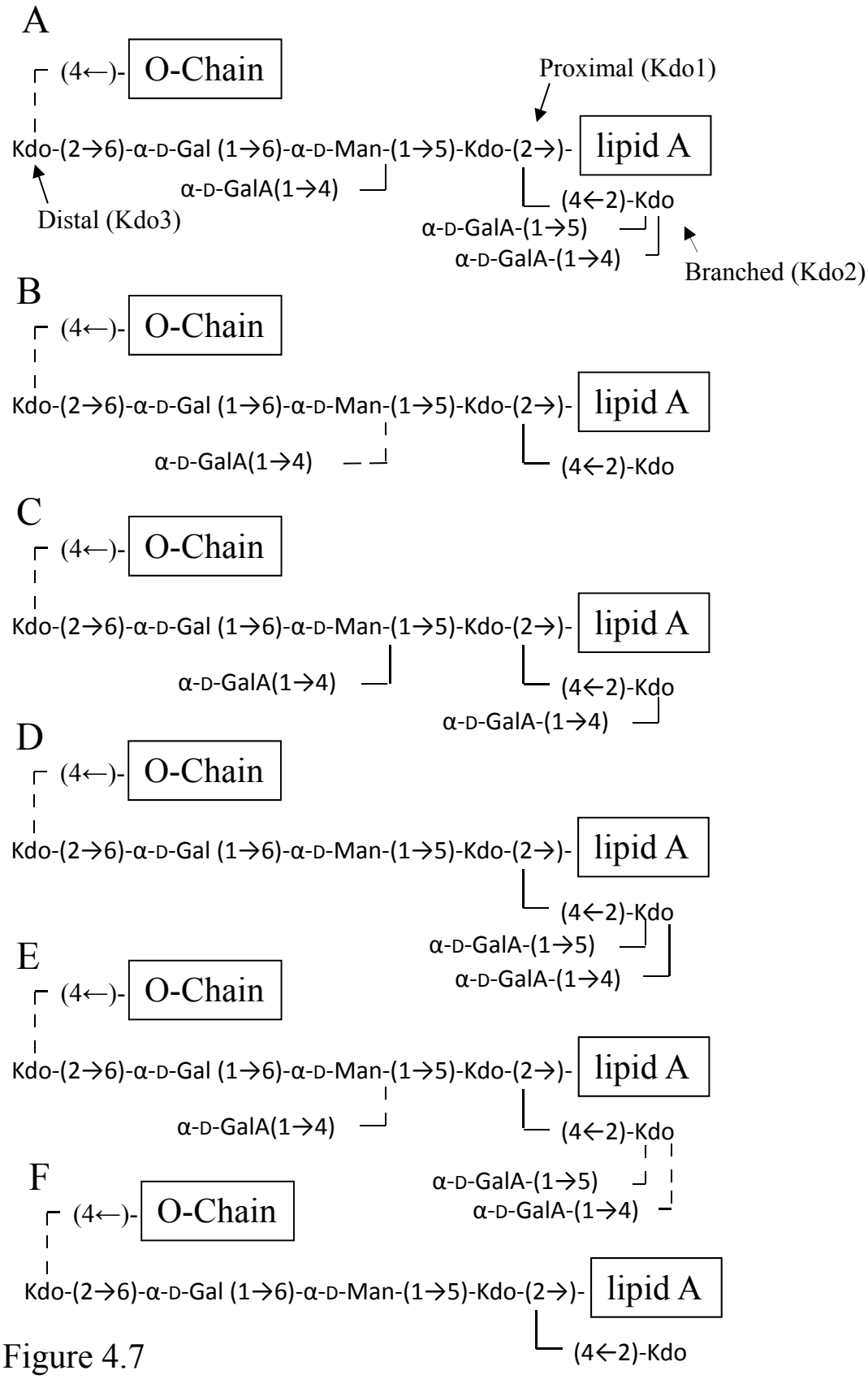

A
(4←)- O-Chain
Proximal (Kdo1)
Kdo-(2→6)-α-D-Gal (1→6)-α-D-Man-(1→5)-Kdo-(2→)- lipid A
α-D-GalA(1→4)
(4←2)-Kdo
α-D-GalA-(1→5)
α-D-GalA-(1→4)
Distal (Kdo3)
Branched (Kdo2)
B
(4←)- O-Chain
Kdo-(2→6)-α-D-Gal (1→6)-α-D-Man-(1→5)-Kdo-(2→)- lipid A
α-D-GalA(1→4)
(4←2)-Kdo
C
(4←)- O-Chain
Kdo-(2→6)-α-D-Gal (1→6)-α-D-Man-(1→5)-Kdo-(2→)- lipid A
α-D-GalA(1→4)
(4←2)-Kdo
α-D-GalA-(1→4)
D
(4←)- O-Chain
Kdo-(2→6)-α-D-Gal (1→6)-α-D-Man-(1→5)-Kdo-(2→)- lipid A
(4←2)-Kdo
α-D-GalA-(1→5)
α-D-GalA-(1→4)
E
(4←)- O-Chain
Kdo-(2→6)-α-D-Gal (1→6)-α-D-Man-(1→5)-Kdo-(2→)- lipid A
α-D-GalA(1→4)
(4←2)-Kdo
α-D-GalA-(1→5)
α-D-GalA-(1→4)
F
(4←)- O-Chain
Kdo-(2→6)-α-D-Gal (1→6)-α-D-Man-(1→5)-Kdo-(2→)- lipid A
(4←2)-Kdo

Figure 4.7

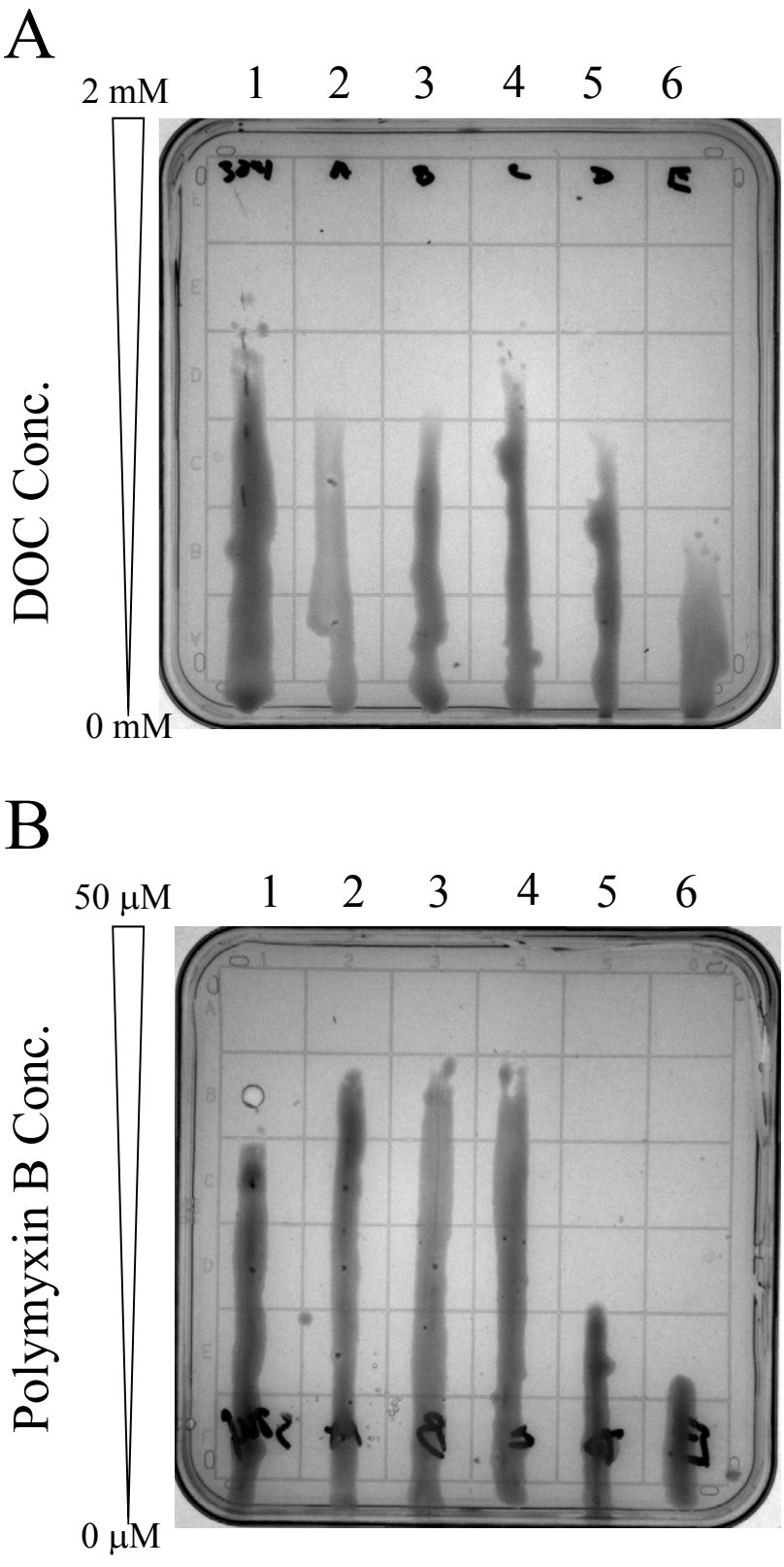


Figure 4.8

Figure 4.9

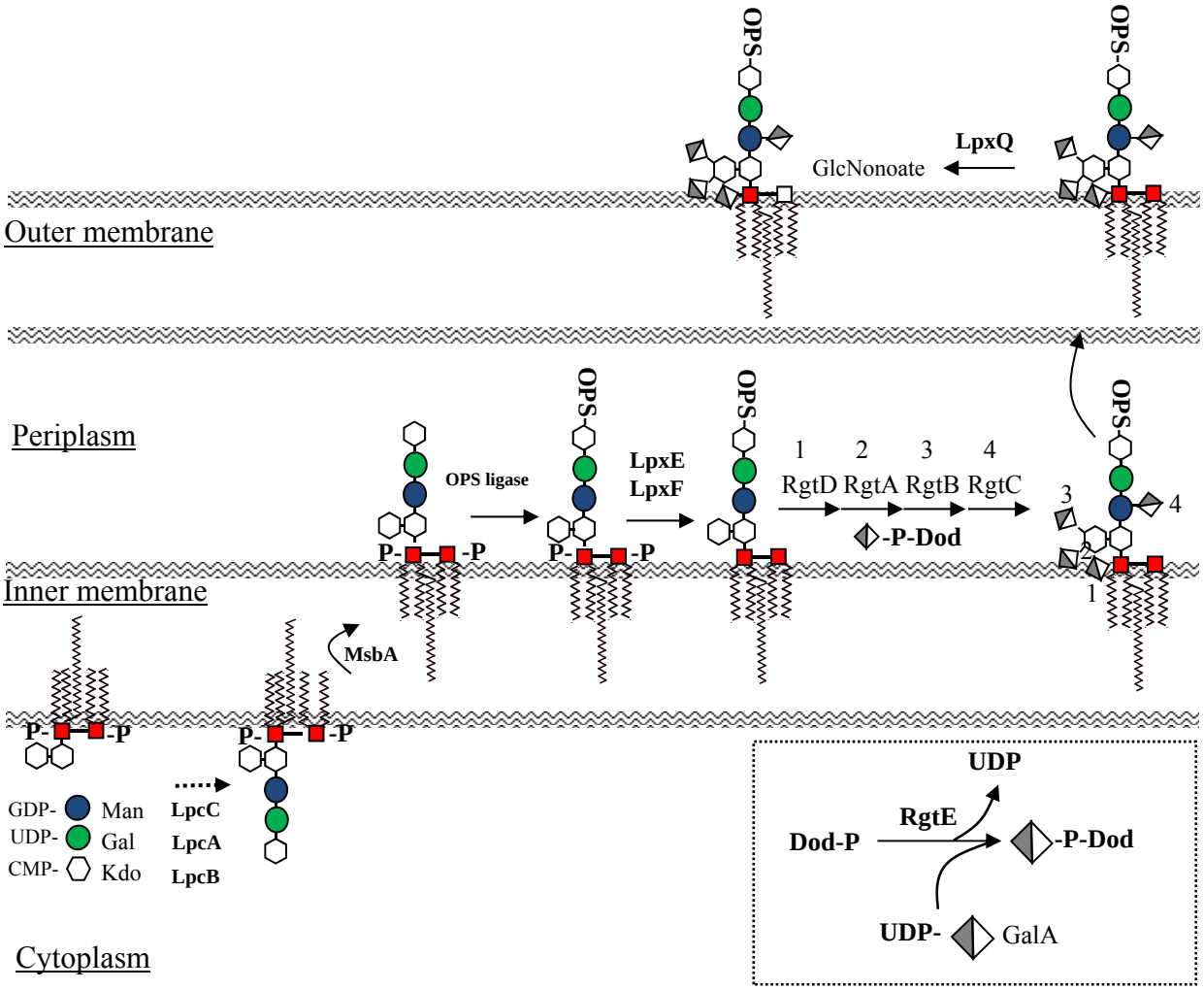
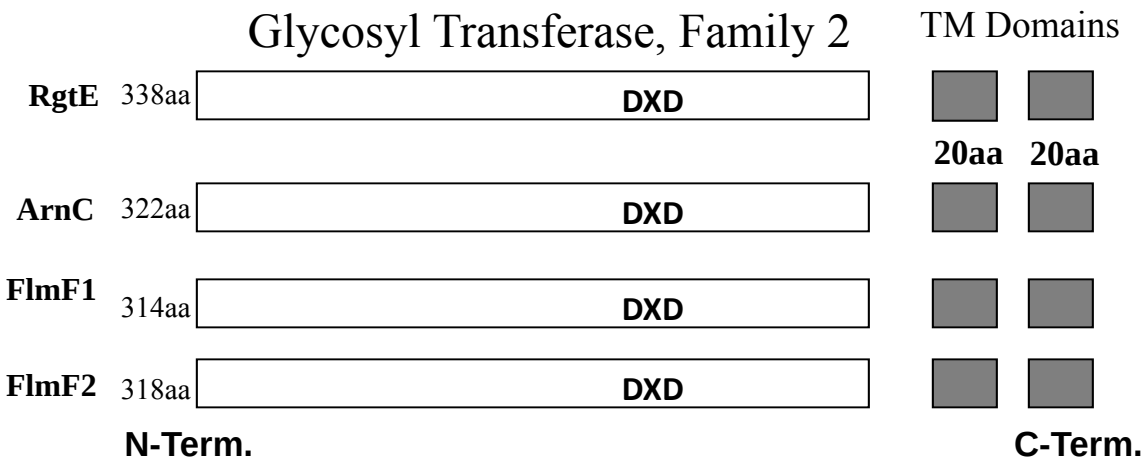

OPS-
LpxQ
GlcNonoate
Outer membrane
Periplasm
OPS ligase
LpxE
LpxF
1 2 3 4
RgtD RgtA RgtB RgtC
-P-Dod
Inner membrane
MsbA
P- -P
GDP- Man
UDP- Gal
CMP- Kdo
LpcC
LpcA
LpcB
Cytoplasm
UDP
RgtE
Dod-P
-P-Dod
UDP- GalA

Figure 4.10

```
Summary of Percent Matches:
   Ref:  RgtA(RL1469)          1 to    499   (   499 aa)     --
     2:  RgtB(RL1468)          1 to    494   (   494 aa)     45%
     3:  RgtC(RL1471)          1 to    501   (   501 aa)     59%
     4:  RgtD(RL0684)          1 to    473   (   473 aa)     21%


RgtA(RL1469)          1 MLERATRTIKTAGLLLAAYFVL-NIVLRIVLPHSLELDEAEQSFFSQYLLAGYGPQPPFY
RgtB(RL1468)          1 MTESNRRDISWIFALLAAYFVL-QVGVRLATSHSLDLDEAEQAFRSQWLAAGYGPQPPFY
RgtC(RL1471)          1 MLERITRSITSASIFLAAYFLL-NIALRIALPHTLDLDEAEQSFYSQYLLAGYGPQPPFY
RgtD(RL0684)          1 MSPR-------SGLLIVLGFTLWRVVMLNFDATDFFVDEAQYWFWSQNLDLGYYSKPPMI

RgtA(RL1469)         60 NWMQYAVVSVTGIS-------IGALIVPKNILLFLSYLFYGLAGRRVLKDEALAAVGMLA
RgtB(RL1468)         60 NWLQYTVFQFAGVS-------LTALSIVKNLLLFISYLLYGLTARLVLRDKALVAIATLG
RgtC(RL1471)         60 NWIQYAIVSVTGIS-------MWVLSVPKNIILFGCYLFYGLAAREVLKSRSLAALAMLS
RgtD(RL0684)         54 AWVIRAMTELSGSNAIYWIRLLGPLIHMAAALVLMK------TAKRFVGPEIEGWTG-AT

RgtA(RL1469)        113 LITLPQVSYMAQQDLTHTTALLFASSLFLYGFFRTLDRPDMASYLLLGLATGIGLISKYN
RgtB(RL1468)        113 LLTIPQMAFEMQRDLTHTVAVFFSASIFFYGFIRSLKQPSLASYLIAGIGIGFGLLAKYN
RgtC(RL1471)        113 LITLPQVGLMAQRELTHTVALLFATSLFLFGFFRTLRQPTIGSYLLIGIATGIGLISKYN
RgtD(RL0684)        107 YITLPGVA-LSSVFFSTDVILLFFIAIALLAYFGLTQRRSVGLALVMGLGVGLAFLTKYA

RgtA(RL1469)        173 FALMPVVALIAILPDAEWRRRALDWRMLAAITVALVIVLPHAVWLQGNLAFASSDTLVK-
RgtB(RL1468)        173 FAILPAAALIAALSDARLRPRIFDWRLGLTAAVALVITLPHLFWLKDNLDFATARTLEK-
RgtC(RL1471)        173 FAILPFAALIAVLPEREWRSRLIDWRLLPAAVLAILIVLPHALWLPDNLASASAPTLER-
RgtD(RL0684)        166 VLFVVPGGAIALLLIPAARIAVRD--VIIAVAVAAVVALPNLWW---NL--QHDNTTVRH

RgtA(RL1469)        232 ---MAAGS---EPAGAVRIGKGLLAFLVAIIAFAALPVVIFA---ATFRRD--FVRALSA
RgtB(RL1468)        232 ---MTASG---DASYLTQVAMGVSSLALAIISFAALTVAVFA---IVFGKS--LRPALGS
RgtC(RL1471)        232 ---MTADPGHLAPAGLPRIGQGLLSLVIAVLGFVALPIVLIA---AAFRRDSRVFRALSS
RgtD(RL0684)        219 TQDIAHWS---ELGINLRRGLEFFAAQFGVVG----PIIFFAMLWAVYR----MIRGRSD

RgtA(RL1469)        281 GNRWTGMMERMML-ASLAGIALIVLFTGSTTVRERWLDPFLL---VLPIYFLAKMQAAGL
RgtB(RL1468)        281 GSEWTRLLERMML-VFLAGILLLIVFGGAAGIKDRWLVPMLF---ILPLYFCLKIEAAGV
RgtC(RL1471)        286 SSPMIRVIERMMV-ISLLAFVGVVLFAGASDIHERWLDPCLL---VLLIYLFLKLETADI
RgtD(RL0684)        268 DR------EKMLVWLSMPVVLLITLQATVAKAYANWAVTAYVAGTILAVWLLYLKWPKGL

RgtA(RL1469)        337 DLSAGLRRFRPVLPVLMACVLIALGFRVVGAGLIGTY-SRPNVPMAGFAREMTRQAEPAL
RgtB(RL1468)        337 ETGKALRRFIPVVAVIMIGVPAALYGSVAAARFTGHY-ERLNRPYAGMLEILRKQAEPAA
RgtC(RL1471)        342 DLSAGLARFRPVVPVFMVVILSILLFRIVGIQYIGTY-TRTNVPFSGYVAELTATRKPVL
RgtD(RL0684)        322 RLSLTINGIASLL-FPLATIFPHQLLLPNGDALMKRYLGRAEVSRE--AAALATQAGTDI

RgtA(RL1469)        396 VIASDTYIGGNMRLQFPDVPVVIPDFPAPGIP-AY--------AEAKGPVLIVWRGKKTA
RgtB(RL1468)        396 ILAGDSLLAGNLRQDIPGVPILSADY--PGFN-PD--------LTSRRPLLLVWLLPK--
RgtC(RL1471)        401 IVAGTKFIAGNMRLQFPDVPVVIPFFPGPGVP-EY--------ADAKGPVLVIWRGE---
RgtD(RL0684)        379 IVTDNRDMVADLFYTLRDASYRIYARAPAGLPESYYEQEFALPADITGKVLFLTDGAFTC

RgtA(RL1469)        447 TAADAVMPERFSSALTAAGIALQEIGSLSLPYYFGRQGDNFALGYAWVRPETR
RgtB(RL1468)        443 GGSEALPPDMAEWLQANLGTSAPEASVIDVPYFYGRGDDRYRFGYAWVNQPG-
RgtC(RL1471)        449 TADDPTISPGFANDLVKSGIHLPELKTLTLPYLFGDGKRSFSIGYSWVEGGAK
RgtD(RL0684)        439 ATE---TPEVLKNWQPTEG------------NYKGKTLSIYKVSATCLAP---
```

Supplement 4.1

Strains	Related Enzymes	Reaction
Rhizobium leguminosarum	RgtE, 100% i.d.	UDP-D-GalA → Dod-P-GalA
Escherichia coli	ArnC, 32% i.d.	UDP-L-AraNFormyl → Und-P-L-AraNFormyl
Francisella novicida	FlmF1, 24% i.d. FlmF2, 23% i.d.	UDP-D-GalNAc → Und-P-D-GalNAc UDP-D-Glc → Und-P-D-Glc



Supplement 4.2

```
Summary of Percent Matches:
   Ref:  Rl RgtD(RL0684)        1 to    473   (   473 aa)     --
     2:  Aq RgtD(aq_765)        1 to    480   (   480 aa)     27%
     3:  Cc RgtD(CC_0209)       1 to    524   (   524 aa)     29%


Rl RgtD(RL0684)       1 -------MSP-------RSGLLIVLG--FTLWRVVMLNFDATDFFVDEAQYWFWSQNLDL
Aq RgtD(aq_765)       1 -----------------MFGLALLINTFFLVFRVLYVLFYPVDLSPEEAQYWDWSRHLDL
Cc RgtD(CC_0209)      1 MQTAPDASSPPIENRAWRLTLLMIGG--LTIVRLAALFLTPLELYPDEAQYWLWSRELAF

Rl RgtD(RL0684)      45 GYYSKPPMIAWVIRAMTELSGSNAIYWIRLLGPLIHMAAALVLMKTAKRFVGPEIEGWTG
Aq RgtD(aq_765)      44 SYYSKPPMVAYMNFLSTHVFG-NTELGVRINAILLSFLLSLITYFFAKKLFSEKVAFVAS
Cc RgtD(CC_0209)     59 GYFSKPPMIAWLIWATTQIGDTEA--WVRLSAPFLHGATALVIHRIARRLYG----GWAG

Rl RgtD(RL0684)     105 ----ATYITLPGVALSSVFFSTDVILLFFIAIALLAYFGL---TQRRSVGLALVMGLGVG
Aq RgtD(aq_765)     103 ----VVPNLFTGFSINSVLFTTDSPLIFFWALSVISFYFA---IEKNTLSLWILTGVFSG
Cc RgtD(CC_0209)    113 LAAAAIYSLMPGVVLSSGLIATDAPLLFFLSLTVWAYVSLPDASARRRYALAAGMGAALG

Rl RgtD(RL0684)     158 LAFLTKYAVLFVVPGGAIALLLIPAARIAVRDVIIAVAVAA--VVALPNLWWNLQHDNTT
Aq RgtD(aq_765)     156 LAFLSKYPAVFLLPLGILYLYLTKKELLKDLKIFSSVLVAF--LIALPVLIWNAKHDFIS
Cc RgtD(CC_0209)    173 LAFLSKYAAVYALGSVALHFAISSEARRRWSPALVGLFIVAFALVLAPNLLWNAANQFST

Rl RgtD(RL0684)     216 VRHTQDIAHWSELGI-NLRRGLEFFAAQFGVVGPIIFFAMLW-AVYRMIRGRSDDREKML
Aq RgtD(aq_765)     214 FKHVSNLAQKHAHFP-NFSTFFEYLGGQVLLLSVIPFFFVLY-GWVRTFK----ERNKRL
Cc RgtD(CC_0209)    233 VKHTAANANWNAHQLFNVRELIEFVGSQFGVFGPVPFAVLIGGAIWLGVKRKLQSPDLLL

Rl RgtD(RL0684)     274 VWL---SMPVVLLITLQATVAKAYANWAVTAYVAGTIL-AVWLLYLKWPKGLR-LSLTIN
Aq RgtD(aq_765)     268 IFLTTFSLPVFLFFAFLSLKKRVYANWAGFGYYTASIL-FAY-YFLKSPKSLKFLTLILS
Cc RgtD(CC_0209)    293 LCF---ALPPLIVVAGEAFVSRANANWAGAAFVSGSVIVAGWLLRWNARRWLI-GGLVLQ

Rl RgtD(RL0684)     329 GIASLLFPLATIFPH----QLLLPNGDALMKRYLG-----RAEVSREAAALATQAGTDII
Aq RgtD(aq_765)     326 AFLTLLLHFTPLFDYLGLRNLLPPKRDPA-KLLVG-----WEDLGKEVGRFYT--GKELI
Cc RgtD(CC_0209)    349 AAFAAFFVACMVNPKVAD-AAGLSNG---FKRVRGWDQTVEAVIARVREEQALRGPLSAV

Rl RgtD(RL0684)     380 VTDNRDMVADL-FYTLRDASYRIYARAPA-------GLPESYYEQEFALPADITGK-VLF
Aq RgtD(aq_765)     378 FSTAYQISAELAFYV--PGNPRTYVFHVN-------RYTQYYLWREGL--KNFKGKDAVF
Cc RgtD(CC_0209)    405 AMDDRFVYNAA-AYYGRD--YFGQPGAPPLRMWVHEAYPQNQAETETPLDADYGRR-ALI

Rl RgtD(RL0684)     431 LT-DGAFTCAT----ETPEVLKNWQPTEG---------NYKGKTLSIYKVS----A----
Aq RgtD(aq_765)     427 VS-YGGVPKEVMRSFEGKEFLKEVRVVWR---------NQVVRKFYIYKLKNFKGE----
Cc RgtD(CC_0209)    461 VSLEGGY---------RPEIEQDFKAVSGLQIARVRLDKTRSRRVDLFIAE----GFAPL

Rl RgtD(RL0684)     469 -----TCLAP-------
Aq RgtD(aq_765)     473 -----FYENPKGY----
Cc RgtD(CC_0209)    508 PRDPITGLPPKPKTAAR
```

Supplement 4.3

CHAPTER 5

Characterization of a novel lipid A α-(1,1)-galacturonosyl transferase (RgtF) from the nitrogen fixing endosymbiont *Mesorhizobium loti* : heterologous expression of the *rgtF* gene in *Rhizobium etli* increases membrane stability.

Dusty B. Brown, Artur Muszynski, and Russell W. Carlson.

To be submitted to *The Journal of Biological Chemistry*.

Abstract

An unusual α-(1,1)-galacturonosyl (GalA) lipid A modification has been reported in the lipopolysaccharide (LPS) of a number of interesting gram negative bacteria including the nitrogen fixing bacteria *Azospirillum lipoferum, Mesorhizobium huakui* and *M. loti* (our laboratory, Dr. Artur Muszinski), the stalk forming bacterium *Caulobacter crescentus*, and the hyperthermophillic bacterium *Aquifex aeolicus*. However, the α-(1,1)-GalA transferase (GalAT), which we have named RgtF, was unknown. Species of the *Rhizobium* genre produce lipid A with α-(1,4`)-GalA but not α-(1,1)-GalA. The *Rhizobium* GalAT, RgtD, is the lipid A α-(1-4`)-GalAT which utilizes the lipid donor dodecaprenyl-phosphate galacturonic acid (Dod-P-GalA) for GalA transfer. An additional *Rhizobium* GalAT, RgtE, is required for the biosynthesis of Dod-P-GalA. We predicted candidate *rgtF* genes in bacterial species known to produce lipid A with α-(1,1)-GalA. In order to determine the predicted *rgtF* gene function, we cloned the *M. loti rgtF* gene into an expression plasmid and introduced that plasmid into *Rhizobium etli* strains that do not contain the *rgtF* gene or produce lipid A α-(1,1)-GalA. MALDI-TOF-MS/MS analysis revealed that the lipid As from these *rgtF* complemented strains were modified with an additional GalA on the proximal glucosamine. Expression of *rgtF* in the nitrogen fixing endosymbiotic model organism *R. etli* conferred increased membrane stability.

Introduction

Lipopolysaccharide (LPS) is a major cell surface glycoconjugate found in the outermembranes of gram negative bacteria and contains a carbohydrate and lipid portion. The lipid A, the lipid portion of LPS, anchors the LPS molecule to the outer leaflet of the outermembrane lipid bilayer and is the main focus of this study. Furthermore, lipid A (endotoxin) in general is an elicitor of the animal defense response and is crucial for outer membrane stability and viability in gram negative bacteria (1,2). However, mutants that produce LPS with truncated carbohydrate domains (rough and deep rough LPS) are typically viable under laboratory conditions. These observations emphasize the important role that lipid A has in maintaining the integrity of the gram negative bacterial cellular envelope and as a major recognition molecule by mammalian host defense response systems.

The model organism *Mesorhizobium loti* MAFF303099 forms nitrogen fixing endosymbiosis with the host legume plant *Lotus japonicum* (3). There are many examples that demonstrate the importance of lipopolysaccharide (LPS) in the formation of complete nitrogen fixing symbiosis between bacterial symbionts and host plants (4). The lipid A structure of *M loti* has not been published. However, the lipid A structure of the closely related strain *M huakuii* has been determined (5) and presents a unique lipid A structure that contains a non-reducing trisaccharide carbohydrate backbone. The backbone comprises a β-(1,6)-2,3-diamino-2,3-dideoxyglucosamine (DAG) disaccharide that is partially substituted with phosphate at the 4` position and contains an unusual α-(1,1) linked galacturonosyl residue attached to the proximal DAG in stoichiometric amounts. Organisms that produce lipid A with α-(1,1) GalA include *Azospiriullum lipoferum, Caulobacter crescentus,* and *Aquifex aeolicus. C. crescentus* and *A. aeolicus* produce *bis*-galacturonosylated lipid A with GalA attached to the 4` (distal) and 1

(proximal) positions. The nitrogen fixing endosymbionts *Rhizobium etli* and *R. leguminosarum* produce α-(1,4`) mono-galacturonosylated lipid A as well as three terminally linked GalAs on the LPS core. The biosynthetic genes responsible for GalA attachment to the lipid A and core have been determined in *R. leguminosarum* 3841 (Chapter4 and (6)). The *Rhizobium* GalATs (Rgt) responsible for core GalA addition, RgtA, B, and C as well as the lipid A (1,4`)-GalAT, RgtD, are ArnT (4-deoxy-4-aminoarabanosyl transferase) like (7) integral innermembrane proteins of the glycosyl transferase family 39 (GT-39) and require the lipid donor dodecaprenyl-phosphate galacturonic acid (Dod-P-GalA) for GalA transfer to the LPS. An additional GalAT, RgtE, is a bacterial like dpm-1 (dolichol-phosphate mannosyl transferase) innermembrane associated protein of the glycosyl tranferase family 2 (GT-2) and is responsible for the biosynthesis of Dod-P-GalA. This type of biosynthetic pathway where by an ArnT like protein and polyprenyl-phosphate (bactoprenyl) glycosyl lipid donor is likely a general pathway for the observed glycosylation of the lipid A disaccharide backbone and LPS core among gram negative bacteria as this phenomenon is generally observed (Chapter 4 and (6,8-10). Therefore, we hypothesize that the GalA transferase responsible for the biosynthesis of α-(1,1)-GalA in the aforementioned organisms is an ArnT like protein and that a bactoprenyl-phosphate GalA lipid donor will be required.

Our laboratory has previously characterized the general lipid A structure of the sequenced strain *M. loti* MAFF303099 (data not shown) and determined that the lipid A backbone composition was similar to the fully characterized lipid A of *M. huakuii* (5). We were able to predicte the α-(1-1)-GalAT in *M. loti* MAFF303099, which we have named RgtF. Here, we describe the function of the *M. loti* MAFF303099 *rgtF* to be an ArnT-like lipid A α-(1,1)-GalAT. In addition, we introduced the *rgtF* gene into the agriculturally significant nitrogen fixing bean

symbiont *R. etli* bv. *phaseoli* and describe the effects on lipid A biosynthesis and membrane stability.

Experimental Procedures

Bacterial strains, plasmids, and growth conditions-For a list of strains and plasmids used in this work see Table 5.1. *E. coli* strains were grown at 37 ^{o}C on Luria-Burtani (LB) media and *Rhiobium etli* and *Mesorhizoboium loti* strains were grown at 30 ^{o}C on tryptone-yeast extract media (TY) (11) containing 10mM $CaCl_2$. Antibiotics were used at the following concentrations where indicated: Streptomycin (50 μg/ml), Spectinomycin (50 μg/ml), and Tetracycline (15 μg/ml).

Creation of the rgtF *expression vector and conjugal transfer into* Rhizobium etli *strains*-The putative *M. loti* MAFF303099 *rgtF* gene (gene i.d., mlr0011) was PCR cloned from the MAFF303099 genomic DNA into plasmid pRK404E1 (12) using the following primers engineered with *BamHI* (underlined) and *EcoRI* (underlined) respectively (ATCGATGGATCC*GTCACTTCCGAGAAACTG* and ATCGATGAATTC*CCAAGGAGCCGACTGG*). The gene was oriented parallel to the lac operon promoter and the resulting plasmid was named pMl_RgtF. Plasmid pMl_RgtF was transformed into *E. coli* Top10 (Invitrogen Inc.) cloning strain. In order to introduce plasmid pMl_RgtF into *Rhizobium* strains, a tri-parental mating was performed as previously described (13) using *E. coli/*pMl_RgtF as a donor, *E. coli/*pRK2013 (14) as a helper, and *Rhizobium* strains as acceptors. The following diagnostic primers were created to recognize regions outside of the pRK404E1 multiple cloning site (GTGGCGAAACCCGACAG and GACTGGAAAGCGGGCAG).

Isolation of LPS and lipid A-Bacteria were grown under vigorous shaking at 30 ºC in TY media for 48 hrs. Finally, cells were harvested by centrifugation at 3500 × g and then washed with sterile PBS followed by washing with distilled water. LPSs were isolated from bacteria using the hot phenol water extraction procedure (15). Crude phenol and water phases were dialyzed against distilled water using 12-14,000 MWCO dialysis bag, freeze-dried and was further purified by ultracentrifugation at 100,000 × *g* at 4 ºC for 18h. The LPS was mainly recovered from water phase and resulting LPS pellets were suspended in a small volume of deionized water, lyophilized, and characterized as described below. LPS was washed with 95% ethanol (v/v) at 4 °C to avoid any possible contamination with phospholipids, re-suspended in water and freeze-dried. LipidA was released by mild hydrolysis. Briefly, samples (5 mg/ml, 1% HOAc) were hydrolyzed at ~100 °C for 70 min, with constant stirring, freeze-dried and extracted with Bligh-Dyer solvent (2 $CHCl_3$/2 MeOH/1.8 H2O; v/v/v). Organic phase was collected and extraction was repeated with chloroform only. Combined organic phases were extracted with H_2O and dried with stream of N_2.

MALDI MS analysis of lipid A-Lipid A samples were dissolved in CH_3Cl/IPA/H_20 mixture (5/3/0.25; v/v/v) and then mixed 1:1 (v/v) with 0.5 M 2.4.6- trihydroxyacetophenone monohydrate (THP) matrix in methanol and 0.5ul (1µg) were spotted onto a stainless-steel target plate. Finally Spectra were recorded in the negative or positive reflectron ion mode using an ABSCIEX TOF/TOF **5800** System, Aplied Biosystems, USA. The MS/MS spectra were acquired in the positive mode. All acquisitions were calibrated with a standard of CE3 lipid A and verified with peptide standards and observed structures represent a monoisotopic most abundant $[M-H]^-$ or $[M+Na]^+$ ions.

Glycosyl and fatty acid composition analysis of lipid A—The composition of lipid A was determined by gas chromatography/mass spectrometry (GC/MS) of trimethylsilyl (TMS) methyl glycosides and fatty acid methyl esters (FAME) as described previously (16,17). For FAME analysis, lipid A along with an inositol internal standard was dissolved in 2 M methanolic HCl and incubated at 80 ºC for 18 h and analyzed by combined gas chromatography-mass spectrometry (GC-MS). TMS methyl glycosides were derived by methanolysis of lipid A. FAMES were removed from the reaction by hexane extraction. The remaining sugar residues were treated for 1 h with 0.5 M HCl-MeOH, *N*-acetylated, converted to TMS methyl glycosides, and analyzed by GC-MS.

DOC, SDS, PmxB, and NaCl resistance assay- Gradient plates were made by pouring at an angle 25 ml TY agar media containing either sodium deoxycholate (DOC), sodium dodecyl sulfate (SDS), polymyxin B (PmxB), or sodium chloride (NaCl) at indicated concentrations in a square 100 mm x 15 mm petri dish. After the agar solidified, they were placed level and 40 ml of TY agar media was poured evenly on top and the top layer was allowed to solidify. The plates were allowed to diffuse overnight. A bacterial stock solution was prepared by dissolving a loop full of bacteria grown for four days on TY agar media into 1 ml sterile phosphate buffered saline (pH 7.4). Then, 10 µl of bacterial suspension was spotted on the gradient plate and streaked from highest to lowest concentration of DOC, SDS, PmxB, or NaCl. The plates were incubated for four days at 30 ^{o}C. Serial dilutions of the bacterial stocks were prepared and enumerated to insure comparable concentrations. Each experiment was repeated 3 times.

Results

Discovery of the putative lipid A α-(1,1)-GalA transferase-The presence of lipid A α-(1,1)-GalA has been reported in *Caulobacter crescentus* (18), *Mesorhizobium huakuii* (5), *Azospirillum lipoferum* (19), and *Aquifex aeolicus* (20). So far, the glycosylatioin of the lipid A backbone in gram negative bacteria has been attributed to ArnT (4-deoxy-4-amino-arabinosyl transferase) like glycosyl transferases that are distantly related to the eukaryotic PMT (phosphate mannosyl transferase) family of glycosyl transferases. ArnT is an integral inner membrane protein that acts on the periplasmic face of the inner membrane and facilitates the transfer of 4-deoxy-4-amino-arabinose (Ara4N) from the lipid donor undecaprenyl-phosphate-Ara4N (Und-P-Ara4N) to the lipid A 4` or 1 phosphate in *E. coli* and other enteric pathogens (1,2). Before deformylation, the Und-P-*N*-formyl arabinose precursor lipid donor (Und-P-Ara4FN) is synthesized at the inner membrane cytosolic interface by ArnC, a dpm1-like glycosyl transferase (bacterial like dolichol-phosphate mannosyl transferse) that utilizes UDP-Ara4FN as a substrate (1,2). ArnT and ArnC like enzymes have since been discovered in *Francisella* (9,10,21), *Bordetella* (22), and *Rhizobium* (Chapter 4) and were shown to be involved in the synthesis of lipid donor substrates and subsequent transfer of glycosyl residues to the lipid A disaccharide backbone. Therefore, we hypothesized that the lipid A α-(1,1)-GalAT is an ArnT like protein. Indeed, *C. crescentus, A. aeolicus, A. lipoferum,* and *Mesorhizobium spp.* contain a similar predicted *arnT* like gene with high peptide similarities which we named *rgtF* (23) (Table 5.2.). In addition, *C. crescentus* and *A. aeolicus* contain homologues to the recently reported lipid A (1,4`)-GalAT, RgtD, which is consistent with the presence of 1 and 4` GalA in these organisms' lipid A (Table 5.2). Also, these organisms contain homologues of RgtE (Table 5.2), an ArnC like enzyme responsible for the synthesis of the lipid donor dodecaprenyl-phosphate GalA which is required for the

galacturonosylation of LPS in *R. leguminosarum* (Chapter 4). In *M. loti* MAFF303099, the *rgtF* gene is in a four gene cluster that contains the predicted lipid A biosynthetic genes *lpxE* and *rgtE* and a gene of unknown function (Figure 5.1 Table 5.2). The lipid A 1-phosphatase *lpxE* (mlr0012) which is required for the removal of 1-phosphate in *R. leguminosarum* (24) and *R. etli* (25), is more than likely required for the 1-dephosphorylation and subsequent addition of the α-(1,1)-GalA to the 1 position of the proximal diaminoglucosamine (DAG) (Figure 5.1). The predicted *M. loti rgtE* gene is likely required for the synthesis of a bactoprenyl-P-GalA lipid donor (Chapter 4 and (8)) that can be utilized by RgtF (Figure 5.1).

Analysis of lipid A isolated from Rhizobium etli *strains CE3, CS506, ODB32 and ODB33* –In order to determine the presence and location of GalA on the lipid A and confirm the function of the *rgtF* gene, MALDI TOF MS/MS analysis was performed on isolated lipid A. MALDI TOF analysis of lipid A from the CE3 parent strain indicated the presence of two major clusters of ions due to pentaacyl and tetraacyl (lacking 3OH-C14:0 acyl chain) lipid A species in which the proximal glycosyl residue can be either 2-amino-2-deoxy-gluconate (GlcNonate) or GlcN. Each cluster consists of multiple ions due to heterogeneity in fatty acyl chain length (±14 amu, ±28 amu) and LCFA (27-OHC28:0) with or without 3-hydroxybutyrate (3-OHC4:0, ±86 amu) substitution. For details refer to Figure 5.2, panel I; Figure 5.5, structure I; and Table 5.3, panel I. These ions and proposed compositions are in agreement with published structures (17,25).

In the Δ*lpxF* 4`-phosphatase mutant strain CS506, the ions were consistent with pentaacyl (in the region between m/z 1774 and 1971) and tetraacyl- lipidA species phosphorylated at the distal GlcN and devoid of the 4` GalA. The observed clusters of ions are shifted down by ~96 amu with respect to what is observed in CE3 lipid A (-GalA, +P). Similar to CE3 lipid A, the long chain fatty acid residue was substituted with a hydroxyl group (27OHC28:0) or 3-

hydroxybutyrate (27(3-OHC4:0)C28:0) at the C27 position. These structures were consistent with previously reported data (25). Interestingly, we observed an increase in the number of structures with proximal GlcNonate in the *lpxE*$^-$ mutant (Figure 5.2, panel II; Figure 5.5, structure II; and Table 5.3, panel II).

Glycosyl composition analysis of isolated lipid A revealed that GalA is not present in the lipid A of strain CS506. However, GalA is present in the lipid A of the *rgtF* complemented strain ODB33 demonstrating that the *rgtF* gene was responsible for adding GalA to the lipid A (Figure 5.7, panels C and D). Furthermore, in the ODB33 strain (CS506/pMl_RgtF) MALDI TOF MS spectrum, the higher mass cluster was represented by phosphorylated pentaacyl- and tetraacyl-LipidA species with or without 3-OHC4:0 on the LCFA with an the addition of a GalA residue likely at the reducing end. A thorough analysis of the region between m/z 1950.40 and 2120.52 (Figure 5.2, panel III and Table 5.3, panel III) has revealed that the structures with proximal GlcNonate could not be formed. This clearly suggests that proximal GlcN is substituted with a GalA residue due to RgtF activity (Figure 5.5, structure III) and blocks oxidation of the proximal GlcN by the outer membrane monooxygenase LpxQ. The only structures found in ODB33 to produce GlcNonate are lipid A species without additional GalA (Compare Table 5.3, panels II and III with Figure 5.2, panels II and III). Taken together, these results suggest that only partial substitution of lipid A with GalA occurred in strain ODB33. To further support the hypothesis that the RgtF protein is transferring a GalA residue to the proximal GlcN residue, we performed MALDI TOF MS/MS analysis on lipid A in the positive ionization mode. The major ions of the pentaacyl phosphorylated and galacturonosylated ODB33 lipid A were sodiated forms $[M+Na]^+$ at m/z 2088.4 (P-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_3$,3-OHC18:0,27-OHC28:0, 3-OHC4:0) and m/z 2060.4 (P-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_3$,3-OHC16:0,27-OHC28:0, 3-OHC4:0) and these

were selected for MS/MS cleavage (Figure 5.3, panel III'). MS/MS analysis of the m/z 2088.4 ion generated characteristic fragment ions (Figure 5.3, panel III'-a). We observed a fragment ion at m/z 1912.4 due to cleavage of the GalA residue from the main structure (P-GlcN-GlcN, (3-OHC14:0)$_3$, 3-OHC18:0, 27-OHC28:0, 3-OHC4:0). Identical $[M+Na]^+$ ion was observed in the MALDI spectrum of intact ODB33 lipid A (Figure 5.3, panel III') further supporting that GalA is an integral part of the main structure. Moreover, the observed *Y* fragmented ion at m/z 886.5 was consistent with the proximal GlcN substituted with GalA at the reducing end and acylated with 3OH-C18:0 and 3-OHC14:0. The observed B+ fragment ion m/z 1224.8 is consistent with the distal side of the molecule containing 4' phosphorylated GlcN that is acylated with two 3-OHC14:0 and 27-OH28:0 and the ion at m/z 698.4 is consistent with the distal phosphorylated GlcN substituted with only two 3-OHC14:0 acyl groups. The fragmented ions at m/z 2008.5 and 1990.4 were due to the cleavage of the distally located 4` phosphate group.

MS/MS analysis of the precursor ion at m/z 2060.4 generated fragments identical to those observed in the MS/MS spectrum of the parent ion with m/z 2088.4 as well as new fragments 28 amu less likely due to substitution of the proximal GlcN by 3-OHC16:0 instead of 3-OHC18:0 (compare Figure 5.3, panels III'-a and b). We observed a fragment ion at m/z 1884.3 due to cleavage of the GalA residue (-176 amu) from the main structure (P-GlcN-GlcN, (3-OHC14:0)$_3$, 3-OHC16:0, 27-OHC28:0, 3-OHC4:0). This ion was identical to the parent $[M+Na]^+$ ion 1884.3 observed in the MALDI MS spectrum of intact lipid A (Figure 5.3, panel III'). The *Y* ion with m/z 858.5 is consistent with the proximal GlcN being substituted at the reducing end with GalA and acylated with 3-OHC14:0 and 3-OHC16:0. The B^+ fragment ion with m/z 1224.8 was similar to the B+ ions from the fragmentation of parent ion 2088.4 obtained from strain ODB33 as well as B+ ions obtained from strain CS506 further supports that the GalA is added to the

proximal GlcN. The loss of phosphate from the distal GlcN generated m/z 1980.4 and 1962.4 ions that are 28 amu lower than the counterpart ions (m/z 2008.5 and 1990.4) observed in the MS/MS spectrum of the m/z 2088.4 parent ion.

In the MALDI-MS analysis of strain ODB32 (CE3/pMl_RgtF) lipid A, the highest observed cluster of ions was in the region between m/z 2074.38 and 2230.47 and were +176 amu higher than ions observed in the CE3 parent strain spectrum consistent with the addition of GalA (Figure 5.4, panels I and II). These ions represented pentaacyl lipid A species with a tetrasacharide sugar backbone containing two GalA residues (Figure 5.5, structure IV). This was supported by a lack of structures with GlcNonate in this ion range, as was observed for the high molecular weight ions in the ODB33 MALDI TOF MS spectrum (Figure 5.4, panel II and Table 5.3, panel IV). In this pentaacyl lipid A, the LCFA was substituted with a hydroxyl group or β-hydroxybutyrate at the C27 position. The intensity of MALDI signals for lipid A substituted with 2 GalA residues was very low suggesting only partial substitution of proximal GlcN with GalA. The majority of signals in the ODB32 spectrum was due to structures also found in the parent strain (compare Figure 5.3, panels I and II and Table 5.3, panels I and IV).

Expression of the rgtF *gene* in Rhizobium etli *CE3 confers resistance to detergents and polymyxin B*-In order to determine membrane stability of the *rgtF* complemented strains ODB32 and ODB33, strains were tested for their relative susceptibility to the detergents sodium dodecyl sulfate (SDS), deoxycholic acid (DOC), and the outer membrane attacking antimicrobial peptide polymyxin B (PmxB). Interestingly, strain ODB32 (CE3/pMl_RgtF) was more resistant to SDS (Figure 5.6, panel C), DOC (Figure 5.6, panel D), and PmxB (Figure 5.6, panel B) when compared with parent strain indicating that the ODB32 strain maintained a more stable membrane. It was previously shown that the CS506 mutant was more susceptible to PmxB

compared to parent strain CE3 (25) and we have reproduced those results (Figure 5.6, panel A). Here, we demonstrate that mutant CS506 is also more susceptible to SDS but not DOC (Figure 5.6, panels C and D). Interestingly, strain ODB33 (CS506/pMl_RgtF) did not demonstrate an increased resistance to detergent or PmxB (Figure 5.6, panel A) suggesting that the expression of *rgtF* and lipid A galacturonosylation had little effect on membrane stability in this strain. Because membrane stability may potentially influence osmotolerance, we tested the *rgtF* complemented strains for the ability to grow in the presence of high sodium chloride concentrations (Figure 5.6, Panel B). There was no observable difference between the strains and their ability to grow at high salt concentrations suggesting that the expression of *rgtF* or the loss of *lpxF* does not significantly affect osmotolerance.

Discussion

In this study, we describe for the first time the unique lipid A α-(1,1)-GalA transferase gene *rgtF* from *M. loti* MAFF303099. The MAFF303099 putative *rgtF* gene was PCR cloned into a plasmid vector and introduced into *R. etli* strains that do not produce α-(1,1)-GalA. Mass spectrometric and carbohydrate composition analysis of the lipid A from the *rgtF* complemented strains unequivocally demonstrated that expression of *rgtF* leads to the addition of GalA to the proximal GlcN. In addition, organisms that have been shown to produce lipid A α-(1,1)-GalA contain *rgtF* homologues in their genomes along with the 1-phosphatase *lpxE* gene and the bactoprenyl-phosphate galacturonosyl transferase *rgtE* gene (Table 5.2 and Figure 5.1). It was previously shown that the lipid donor dodecaprenyl-phosphate GalA is required for GalA attachment to the LPS in *R. leguminosarum* bv. *viciae* (Chapter 3 and (8)) and that LpxE removes the lipid A 1-phosphate (24,25). Interestingly, the *M. loti* MAFF303099 genes *rgtF, rgtE,* and *lpxF* are found together in a gene cluster on the chromosome (Figure 5.1). In addition,

the *rgtF* and *lpxE* genes are localized together on the chromosomes of *C. crecsentus, A. lipoferum,* and *Mesorhizobium spp.* (Table 5.2) suggesting a functional relationship between the 1-phosphatase and the RgtF GalAT. Therefore, we suggest a model for the biosynthesis of the lipid A α-(1,1)-GalA (Figure 5.1) whereby the 1-phosphate is removed by LpxE following transfer of GalA from bactoprenyl-P-GalA to the 1 position of the lipid A by RgtF.

It is not known why some organisms have evolved to produce lipid A with GalA. *E. coli* and many other well studied enteric pathogens contain lipid A with 1 and 4` phosphates and it has been suggested that the phosphate groups facilitate the cross bridging of LPS molecules in the outer membrane through divalent cationic interactions with metal ions, *e. g.* Ca^{2+} and Mg^{2+} (26,27). Because the biosynthesis of the precursor molecule lipid IV_A is a well conserved pathway (2) that follows a sequential order and the lipid IV_A biosynthetic genes are wide spread throughout the sequenced genomes of gram negative bacteria, it is thought that most gram negative organisms initially produce 1 and 4` phosphates (2). Organisms that do not produce 1 and/or 4` phosphate express inner membrane phosphatases that remove them on the periplasmic side of the inner membrane. The 1 and/or 4` phosphatases have been discovered in *R. leguminosarum, R. etli, Francisella spp.,* and *Porphyromonas gingivalis.* The periodontal pathogen *P. gingivalis* may produce 4` monphosphorylated lipid A and nonphosphorylated lipid A (28). Interestingly, the lipid A 1-dephosphorylation activity is regulated by host derived haemin, a common nutrient in the oral cavity. The removal of the lipid A phosphates in *P. gingivalis* is crucial for resistance to cationic antimicrobial peptides and modulation and evasion of the tol like receptor 4 (TLR4) immune response (29,30). The 4` phosphatase *lpxF* of *Francisella* has been shown to be crucial for establishing normal virulence (31). In addition to phosphate removal, some mammalian pathogens demonstrate the ability to dynamically modify

their lipid A backbone with glycosyl residues in response to the host environment. For example, several pathogens such as *E. coli, Salmonella spp., Yersinia pestis, Pseudomonas aeroginosa,* and *Bordetella spp.* can mask their phosphates with positively charged glycosyl residues (Ara4N, GalN, and GlcN) (1,22). These modifications promote survival within the host and result in the organisms being less immunogenic and more resistant to cationic antimicrobial peptides due to charge repulsion. *F. novicida* is able to glycosylate its nonphosphorylated distal glucosamine with a neutral sugar (Glc/Man) and mask the 1-phosphate with galactosamine (32). These glycosylations are important for establishing a normal infection and mutant strains that lack proper lipid A glycosylation are promising candidates for whole cellular vaccine development (10).

Unlike the lipid A glycosylation in mammalian pathogens, the lipid A GalA glycosyl residues are constitutively produced in the nitrogen fixing endosymbionts *Rhizobium spp*. and *Mesorhizobium spp.* We have demonstrated in our laboratory that *R. leguminosarum* mutants that are disrupted in the production of GalA in their lipid A backbone or LPS core are compromised in membrane stability (Chapter 4). An *rgtE-minus* mutant of *R. leguminosarum* produces LPS completely devoid of GalA and is severely compromised in membrane stability (Chapter 4). Therefore, the presence of GalA on the LPS functions to stabilize the outer membrane likely through divalent cationic bridging similar to mammalian pathogens that produce phosphate residues in the lipid A and LPS core (27). It was also demonstrated that removal of only the core GalA residues from the LPS of *R. leguminosarum* resulted in a slight increase in the resistance to the cationic antimicrobial peptide polymyxin B (PmxB) likely due to the removal of the negative charge produced by GalA that may aid in the recruitment of the positively charged PmxB. However, the removal of the 4` GalA and total removal of LPS GalA residues resulted in the

organism being significantly more susceptible to PmxB demonstrating that the 4` GalA was crucial for membrane stability and overrides the effect of PmxB resistance due to the loss of negative charge. In this study, we have demonstrated that introducing the *M. loti rgtF* gene into *R. etli* CE3 results in the organism producing a mixture of mono and *bis*-galacturonosylated lipid A (Figure 5.4). Interestingly, strain ODB32 (CE3/pMl_RgtF) was more resistant to detergents and PmxB suggesting that the addition of GalA to the lipid A backbone caused the membrane to be more stable. These results support the hypothesis that the presence of GalA on the LPS promotes membrane stability. These findings have potential implications for the engineering of agriculturally significant plant growth promoting organisms such as *Rhizobium, Azospirillum, Bradyrhizobium,* and others to promote survival in the soil and rhizosphere. Further studies involving the competitiveness, symbiotic/endophytic effectiveness, and survival of *rgtF* complemented strains under various stress conditions such as temperature variability, desiccation, plant derived antimicrobial resistance, etc., are ongoing.

References

1. Trent, M. S., Stead, C. M., Tran, A. X., and Hankins, J. V. (2006) *J. Endotoxin. Res.* **12**, 205-223
2. Raetz, C. R. H., Reynolds, C. M., Trent, M. S., and Bishop, R. E. (2007) *Annu. Rev. Biochem.* **76**, 295-329
3. Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S., Watanabe, A., Idesawa, K., Ishikawa, A., Kawashima, K., Kimura, T., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Mochizuki, Y., Nakayama, S., Nakazaki, N., Shimpo, S., Sugimoto, M., Takeuchi, C., Yamada, M., and Tabata, S. (2000) *DNA Res.* **7**, 331-338
4. Carlson, R. W., Forsberg, L. S., and Kannenberg, E. L. (2010) Lipopolysaccharides in *Rhizobium*-Legume Symbioses in *Endotoxins: Structure, Function and Recognition* (Wang, X., and Quinn, P. J. eds.), Springer Netherlands. pp 339-386
5. Choma, A., and Sowinski, P. (2004) *Euro. J. Biochem.* **271**, 1310-1322
6. Kanjilal-Kolar, S., Basu, S. S., Kanipes, M. I., Guan, Z., Garrett, T. A., and Raetz, C. R. (2006) *J. Biol. Chem*. **281**, 12865-12878
7. Trent, M. S., Ribeiro, A. A., Lin, S., Cotter, R. J., and Raetz, C. R. (2001) *J. Biol. Chem.* **276**, 43122-43131
8. Kanjilal-Kolar, S., and Raetz, C. R. (2006) *J.Biol. Chem.* **281**, 12879-12887
9. Song, F., Guan, Z., and Raetz, C. R. (2009) *Biochemistry* **48**, 1173-1182
10. Kanistanon, D., Hajjar, A. M., Pelletier, M. R., Gallagher, L. A., Kalhorn, T., Shaffer, S. A., Goodlett, D. R., Rohmer, L., Brittnacher, M. J., Skerrett, S. J., and Ernst, R. K. (2008) *PLoS Pathog.* **4**, e24

11. Beringer, J. E. (1974) *Microbiol.* **84**, 188

12. Lehman, L. J., Fitzmaurice, W. P., and Roberts, G. P. (1990) *Gene* **95**, 143-147

13. Brown, D. B., Huang, Y.-C., Kannenberg, E. L., Sherrier, D. J., and Carlson, R. W. (2011) *J. Bacteriol.*, JB.00392-00311

14. Phadnis, S. H., and Das, H. K. (1987) *J.Biosci.s* **12**, 131-135

15. Westphal, O., and Jann, K. (1965) *Meth.Carbohydr.Chem.* **5**, 83-91

16. York, W. S., Darvill, A. G., McNeil, M., Stevenson, T. T., and Albersheim, P. (1986) *Meth. Enzymol.* **118**, 3-40

17. Bhat, U. R., Forsberg, L. S., and Carlson, R. W. (1994) *J.Biol.Chem.***269**, 14402

18. Smit, J., Kaltashov, I. A., Cotter, R. J., Vinogradov, E., Perry, M. B., Haider, H., and Qureshi, N. (2008) *Inn. Immun.* **14**, 25-36

19. Choma, A., and Komaniecka, I. (2008) *Carbohydr.Res.* **343**, 799-804

20. Plotz, B. M., Lindner, B., Stetter, K. O., and Holst, O. (2000) *J.Biol.Chem.***275**, 11222-11228

21. Wang, X., Ribeiro, A. A., Guan, Z., and Raetz, C. R. H. (2009) *Biochem.* **48**, 1162-1172

22. Marr, N., Tirsoaga, A., Blanot, D., Fernandez, R., and Caroff, M. (2008) *J.bacteriol.* **190**, 4281-4290

23. Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., Fong, J. H., Geer, L. Y., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Jackson, J. D., Ke, Z., Lanczycki, C. J., Lu, F., Marchler, G. H., Mullokandov, M., Omelchenko, M. V., Robertson, C. L., Song, J. S., Thanki, N., Yamashita, R. A., Zhang, D., Zhang, N., Zheng, C., and Bryant, S. H. *Nucl. Acids Res.*

24. Karbarz, M. J., Kalb, S. R., Cotter, R. J., and Raetz, C. R. (2003) *J. Biol. Chem.* **278**, 39269-39279

25. Ingram, B. O., Sohlenkamp, C., Geiger, O., and Raetz, C. R. (2010) *Biochim.biophys. acta.* **1801**, 593-604

26. Snyder, S., Kim, D., and McIntosh, T. J. (1999) *Biochem.* **38**, 10758-10767

27. Raetz, C. R., and Whitfield, C. (2002) *Annu.Rev. Biochem.* **71**, 635-700

28. Coats, S. R., Jones, J. W., Do, C. T., Braham, P. H., Bainbridge, B. W., To, T. T., Goodlett, D. R., Ernst, R. K., and Darveau, R. P. (2009) *Cellular Microbiol.* **11**, 1587-1599

29. Coats, S. R., To, T. T., Jain, S., Braham, P. H., and Darveau, R. P. (2009) *Internatl. J. oral sci.* **1**, 126-135

30. Coats, S. R., Berezow, A. B., To, T. T., Jain, S., Bainbridge, B. W., Banani, K. P., and Darveau, R. P. (2011) *Infecti.immun.* **79**, 203-210

31. Wang, X., Ribeiro, A. A., Guan, Z., Abraham, S. N., and Raetz, C. R. H. (2007) *Proc.Natl. Acad. Sci.* **104**, 4136-4141

32. Raetz, C. R., Guan, Z., Ingram, B. O., Six, D. A., Song, F., Wang, X., and Zhao, J. (2009) *J. Lipid Res.* **50 Suppl**, S103-108

33. Noel, K. D., Sanchez, A., Fernandez, L., Leemans, J., and Cevallos, M. A. (1984) *J.bacteriol.* **158**, 148-155

Figure Legends

Figure 5.1. Chromosomal location of the *rgtF* gene and proposed biosynthesis of the lipid A α-(1,1)-GalA in *M. loti* MAFF303099. The predicted *rgtE* and *lpxE* genes are located proximal to the *rgtF* gene. LpxE is a predicted lipid A 1-phosphatase that likely removes the 1-phosphate and RgtE is a dpm1-like glycosyl transferase that likely synthesizes a bactoprenyl-P-GalA lipid donor for the transfer of GalA to the 1 position of the lipid A by the ArnT like glycosyl transferase RgtF.

Figure 5.2. MALDI TOF MS spectra of isolated lipid A. **I)** parent strain CE3. **II)** strain CS506, the Δ*lpxE* mutant of CE3. **III)** strain ODB33 (CS506/pMl_RgtF).

Figure 5.3. MALDI TOF MS/MS spectra of strain ODB33 (CS506/pMl_RgtF). **III`)** single MS displaying parent ions from the lipid A of ODB33. **a)** MS/MS spectrum obtained from the parent ion 2088.4. **b)** MS/MS spectrum obtained from the parent ion 2060.4.

Figure 5.4. MALDI TOF MS spectra of isolated lipid A. **I)** parent strain CE3. **II)** strain ODB32 (CE3/pMl_RgtF).

Figure 5.5. The predicted lipid A structures determined by MALDI TOF MS (Figures 2 and 4). **I)** A general lipid A structure observed in strain CE3. **II)** A General lipid A structure observed in strain CS506. **III)** A general lipid A structure observed in strain ODB33. **IV)** A general lipid A structure observed in the ODB32.

Figure 5.6. DOC, SDS, PmxB, and NaCl susceptibility assay. **Panel A)** PmxB gradient plate. **Panel B)** NaCl gradient plate. **Panel C)** SDS gradient plate. **Panel D)** DOC gradient plate. **1)** Strain CE3. **2)** Strain ODB32. **3)** Strain CS506**. 4)** Strain ODB33.

Figure 5.7. The composition of tri-methyl silyl derived methyl glycosidic GalA residue from isolated lipid A. **Panel A)** strain CE3. **Panel B)** strain ODB32. **Panel C)** Strain CS506. **Panel D)** Strain ODB33. The GalA residues were determined by retention time and extracting the diagnostic m/z 234 ion. GlcNAc residues were determined by retention time and extracting the diagnostic ion m/z 173. Strain CS506 did not contain measurable amounts of GalA. Introducing the *rgtF* gene into CS506 (strain ODB33) results in the production of GalA on the lipid A.

Table 5.1. Bacterial Strains and Plasmids

Strains and Plasmids	Characterization	Source
E. coli		
Top10	F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*araleu*) 7697 *gal*U *gal*K *rps*L (Str^r) *end*A1 *nup*G	Invitrogen
Rhizobium etli		
CE3	Parent strain, Nod^+, Fix^+, Str^r	(33)
CS506	Δ*lpxF* mutant of strain CE3, Nod^+, Fix^+, Str^r, $Spec^r$	(25)
ODB32	Strain CE3 containing plasmid pMl_RgtF, Str^r, Tet^r	This Study
ODB33	Strain CS506 containing plasmid pMl_RgtF, Str^r, $Spec^r$, Tet^r	This Study
Mesorhizobium loti		
MAFF303099	Nod^+, Fix^+	MAFF GenBank
Plasmids		
pRK404E1	Broad host range shuttle vector for, Tc^r	(12)
pRK2013	Mobilizing plasmid for pMl_RgtF, Col E1 replicon, Kan^r	(14)
pMl_RgtF	Plasmid pRK404E1 containing the *rgtF* gene from strain MAFF303099	This Study

Table 5.2. Predicted genes involved in the biosynthesis of galacturonosylated lipid A in a variety of bacterial species

	LpxE	LpxF	RgtD	RgtF	RgtE
Rhizobium leguminosarum 3841	RL4708 Referance	RL1570 Referance	RL0684 Referance	ND	RL1470 Reference
Mesorhizobium loti MAFF303099	mlr0012 Expect $7e^{-41}$	ND	ND	mlr0011 Referance	Mlr0009 Expect $2e^{-33}$
M. ciceri WSM1271	Mesci_4393 Expect $1e^{-38}$	ND	ND	Mesci_4394 Expect 0.0	Mesci_4733 Expect $1e^{-135}$
M. opportunistum WSM2075	Mesop_4837 Expect $8e^{-37}$	ND	ND	Mesop_4838 Expect 0.0	Mesop_4840 Expect $3e^{-32}$
Caulobacter crescentus CB15	cc_3019 Expect $7e^{-09}$	ND	cc_0209 Expect $7e^{-51}$	cc_0468 Expect $2e^{-81}$	Cc_0469 Expect $6e^{-31}$
Aquifex aeolicus VF5	ND	ND	aq_765 Expect $1e^{-45}$	aq_1695a Expect $7e^{-16}$	Aq_1899 Expect $2e^{-94}$
Azospirillum lipoferum B510	AZL_d01960 Expect $2e^{-09}$	ND	ND	AZL_e02810 Expect $2e^{-84}$	AZL_e02830 Expect$4e^{-36}$

ND, not detected; LpxE, 1-phosphatase; LpxF, 4`-phosphatase; RgtD, 4`-galacturonosyl transferase; RgtF, α-(1,1)-galacturonosyl transferase; RgtF, undecaprenyl-phosphate galacturonosyl transferase

Table 5.3. Proposed structures for the highest cluster of ions observed in MALDI-TOF analysis (Figures 2 and 4) of Lipid A isolated from *Rhizobium* CE3, its *lpxE*$^-$ mutant (SC506), the *rgtF* complemented CS506 strain ODB33 and the *rgtF* complemented CE3 strain ODB32 . The recorded m/z values represent the monoisotopic $[M-H]^-$ ions observed in negative reflectron mode.

Strain	Observed * $[M-H]^-$	Calc. * $[M-H]^-$	Molecular formula	Proposed Composition (pentaacyl Lipid A)
CE 3				
	1914.43	1914.41	$C_{106}H_{198}N_2$	GalA-GlcN-GlcNonate, $(3\text{-OHC14:0})_3$, 3-OHC18:0, 27-
	1886.41	1886.39	$C_{104}H_{194}N_2$	GalA-GlcN-GlcNonate, $(3\text{-OHC14:0})_3$, 3-OHC16:0, 27-
	1942.45	1942.45	$C_{108}H_{202}N_2$	GalA-GlcN-GlcNonate, $(3\text{-OHC14:0})_2$, 3-OHC18:0, 3-
	1972.44	1972.42	$C_{108}H_{200}N_2$	GalA-GlcN-GlcNonate, $(3\text{-OHC14:0})_3$, 3-OHC16:0, 27-
I-A	2000.47	2000.45	$C_{110}H_{204}N_2$	GalA-GlcN-GlcNonate, $(3\text{-OHC14:0})_3$, 3-OHC18:0, 27-
	2028.50	2028.45	$C_{112}H_{208}N_2$	GalA-GlcN-GlcNonate, $(3\text{-OHC14:0})_2$, 3-OHC18:0, 3-
	1870.39	1870.39	$C_{104}H_{194}N_2$	GalA-GlcN-GlcN, $(3\text{-OHC14:0})_3$, 3-OHC16:0, 27-OHC28:0
	1898.42	1898.42	$C_{106}H_{198}N_2$	GalA-GlcN-GlcN, $(3\text{-OHC14:0})_3$, 3-OHC18:0, 27-OHC28:0
	1926.43	1926.45	$C_{108}H_{202}N_2$	GalA-GlcN-GlcN, $(3\text{-OHC14:0})_2$, 3-OHC18:0, 3-OHC16:0, 27-
	1942.45	1942.41	$C_{107}H_{198}N_2$	GalA-GlcN-GlcN, $(3\text{-OHC14:0})_3$, 3-OHC15:0, 27-OHC28:0, 3-
	1954.42	1954.48	$C_{110}H_{206}N_2$	GalA-GlcN-GlcN, $(3\text{-OHC14:0})_2$, $(3\text{-OHC18:0})_2$, 27-OHC28:0
I-B	1956.44	1956.43	$C_{108}H_{200}N_2$	GalA-GlcN-GlcN, $(3\text{-OHC14:0})_3$,3-OHC16:0, 27-OHC28:0, 3-
	1970.45	1970.44	$C_{109}H_{202}N_2$	GalA-GlcN-GlcN, $(3\text{-OHC14:0})_2$, 3-OHC15:0, 3-OHC16:0, 27-
	1984.46	1984.46	$C_{110}H_{204}N_2$	GalA-GlcN-GlcN, $(3\text{-OHC14:0})_3$, 3-OHC18:0, 27-OHC28:0, 3-
	1998.47	1998.47	$C_{111}H_{206}N_2$	GalA-GlcN-GlcN, $(3\text{-OHC14:0})_2$, 3-OHC18:0, 3-OHC15:0, 27-
	2012.48	2012.49	$C_{112}H_{208}N_2$	GalA-GlcN-GlcN, $(3\text{-OHC14:0})_2$, 3-OHC18:0, 3-OHC16:0, 27-
	2040.48	2040.52	$C_{114}H_{212}N_2$	GalA-GlcN-GlcN, $(3\text{-OHC14:0})_2$, $(3\text{-OHC18:0})_2$, 27-OHC28:0,
	2054.46	2054.53	$C_{115}H_{214}N_2$	GalA-GlcN-GlcN, $(3\text{-OHC14:0})_2$, 3-OHC15:0, $(3\text{-OHC18:0})_2$,

Strain	Observed * $[M-H]^-$	Calc. *$[M-H]^-$	Molecular formula	Proposed Composition (pentaacyl Lipid A)
CS506				
	1790.27	1790.32	$C_{98}H_{187}N_2O_2$	P-GlcN-GlcNonate, $(3\text{-OHC14:0})_3$, 3-OHC16:0, 27-OHC28:0
	1804.31	1804.33	$C_{99}H_{189}N_2O_2$	P-GlcN-GlcNonate, $(3\text{-OHC14:0})_2$, 3-OHC15:0, 3-OHC16:0,
	1818.32	1818.35	$C_{100}H_{191}N_2O$	P-GlcN-GlcNonate, $(3\text{-OHC14:0})_3$, 3-OHC18:0, 27-OHC28:0
	1832.32	1832.37	$C_{101}H_{193}N_2O$	P-GlcN-GlcNonate, $(3\text{-OHC14:0})_2$, 3-OHC15:0, 3-OHC18:0,
	1846.34	1846.38	$C_{102}H_{195}N_2O$	P-GlcN-GlcNonate, $(3\text{-OHC14:0})_2$, 3-OHC18:0, 3-OHC16:0,
	1876.32	1876.36	$C_{102}H_{193}N_2O$	P-GlcN-GlcNonate, $(3\text{-OHC14:0})_3$, 3-OHC16:0, 27-OHC28:0,
	1890.35	1890.37	$C_{103}H_{195}N_2O$	P-GlcN-GlcNonate, $(3\text{-OHC14:0})_2$, 3-OHC15:0, 3-OHC16:0,

II-A	1904.35	1904.39	$C_{104}H_{197}N_2O$	P-GlcN-GlcNonate, $(3\text{-OHC14:0})_3$, 3-OHC18:0, 27-OHC28:0,
	1918.37	1918.40	$C_{105}H_{199}N_2O$	P-GlcN-GlcNonate, $(3\text{-OHC14:0})_2$, 3-OHC15:0, 3-OHC18:0,
	1932.38	1932.42	$C_{106}H_{201}N_2O$	P-GlcN-GlcNonate, $(3\text{-OHC14:0})_2$, 3-OHC18:0,3-OHC16:0,
	1946.38	1946.43	$C_{107}H_{203}N_2O$	P-GlcN-GlcNonate, 3-OHC14:0,3-OHC15:0, 3-OHC18:0, 3-
	1960.38	1960.45	$C_{108}H_{205}N_2O$	P-GlcN-GlcNonate, 3-OHC14:0, $(3\text{-OHC16:0})_2$, 3-OHC18:0,
	1774.29	1774.32	$C_{98}H_{187}N_2O_2$	P-GlcN-GlcN, $(3\text{-OHC14:0})_3$, 3-OHC16:0, 27-OHC28:0
	1802.32	1802.36	$C_{100}H_{191}N_2O$	P-GlcN-GlcN, $(3\text{-OHC14:0})_3$, 3-OHC18:0, 27-OHC28:0
	1816.33	1816.37	$C_{101}H_{193}N_2O$	P-GlcN-GlcN, $(3\text{-OHC14:0})_2$, 3-OHC15:0, 3-OHC18:0, 27-
	1830.32	1830.39	$C_{102}H_{195}N_2O$	P-GlcN-GlcN, $(3\text{-OHC14:0})_2$, 3-OHC18:0, 3-OHC16:0, 27-
	1858.30	1858.42	$C_{104}H_{199}N_2O$	P-GlcN-GlcN, $(3\text{-OHC14:0})_2$, $(3\text{-OHC18:0})_2$, 27-OHC28:0
II-B	1860.32	1860.36	$C_{102}H_{193}N_2O$	P-GlcN-GlcN, $(3\text{-OHC14:0})_3$, 3-OHC16:0, 27-OHC28:0, 3-
	1874.34	1874.38	$C_{103}H_{193}N_2O$	P-GlcN-GlcN, $(3\text{-OHC14:0})_2$, 3-OHC15:0, 3-OHC16:0, 27-
	1888.35	1888.39	$C_{104}H_{197}N_2O$	P-GlcN-GlcN, $(3\text{-OHC14:0})_3$, 3-OHC18:0, 27-OHC28:0
	1902.36	1901.41	$C_{105}H_{199}N_2O$	P-GlcN-GlcN, $(3\text{-OHC14:0})_2$, 3-OHC15:0, 3-OHC18:0, 27-
	1916.37	1916.42	$C_{106}H_{201}N_2O$	P-GlcN-GlcN, $(3\text{-OHC14:0})_2$, 3-OHC18:0, 3-OHC16:0, 27-
	1944.37	1944.45	$C_{108}H_{205}N_2O$	P-GlcN-GlcN, $(3\text{-OHC14:0})_2$, $(3\text{-OHC18:0})_2$, 27-OHC28:0, 3-

ODB33	**Observed * $[M-H]^-$**	**Calc. *$[M-H]^-$**	**Molecular formula**	**Proposed Composition (pentaacyl Lipid A)**
III-A	nd	nd	-	-
	1950.40	1950.36	$C_{104}H_{195}N_2O$	P-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_3$, 3-OHC16:0, 27-OHC28:0
	1964.41	1964.37	$C_{105}H_{197}N_2O$	P-GlcN-GlcN-GalA, 3-OHC14:0_2, 3-OHC15:0, 3-OHC16:0,
	1978.43	1978.39	$C_{106}H_{199}N_2O$	P-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_3$, 3-OHC18:0, 27-OHC28:0,
	1992.43	1992.40	$C_{107}H_{201}N_2O$	P-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_2$, 3-OHC15:0, 3-OHC18:0,
	2006.45	2006.42	$C_{108}H_{203}N_2O$	P-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_2$, 3-OHC18:0, 3-OHC16:0,
	2034.47	2034.45	$C_{110}H_{217}N_2O$	P-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_2$, $(3\text{-OHC18:0})_2$, 27-
III-B	2036.44	2036.39	$C_{108}H_{201}N_2O$	P-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_3$, 3-OHC16:0, 27-OHC28:0,
	2050.45	2050.41	$C_{109}H_{203}N_2O$	P-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_2$, 3-OHC15:0, 3-OHC16:0,
	2064.46	2064.42	$C_{110}H_{205}N_2O$	P-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_3$, 3-OHC18:0, 27-OHC28:0,
	2078.48	2078.43	$C_{111}H_{207}N_2O$	P-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_2$, 3-OHC15:0, 3-OHC18:0,
	2092.49	2092.45	$C_{112}H_{209}N_2O$	P-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_2$, 3-OHC18:0, 3-OHC16:0,
	2106.51	2106.47	$C_{113}H_{211}N_2O$	P-GlcN-GlcN-GalA, 3-OHC14:0, 3-OHC15:0, 3-OHC16:0, 3-
	2120.52	2120.49	$C_{114}H_{213}N_2O$	P-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_2$, $(3\text{-OHC18:0})_2$, 27-

ODB32	**Observed * $[M-H]^-$**	**Calc. *$[M-H]^-$**	**Molecular formula**	**Proposed Composition (pentaacyl Lipid A)**
IV-A	nd	nd	nd	-
	2074.38	2074.45	$C_{112}H_{206}N_2O$	GalA-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_3$, 3-OHC18:0, 27-
	2088.39	2088.49	$C_{113}H_{208}N_2O$	GalA-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_2$, 3-OHC18:0, 3-

	2102.40	2102.48	$C_{114}H_{210}N_2O$	GalA-GlcN-GlcN, $(3\text{-OHC14:0})_2$, 3-OHC18:0, 3-OHC16:0,
	2116.42	2116.50	$C_{115}H_{212}N_2O$	GalA-GlcN-GlcN-GalA, 3-OHC14:0, 3-OHC15:0, 3-OHC16:0,
	2132.40	2130.46	$C_{114}H_{208}N_2O$	GalA-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_3$, 3-OHC16:0, 27-
IV-B	2146.41	2146.47	$C_{115}H_{210}N_2O$	GalA-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_2$, 3-OHC15:0, 3-
	2160.42	2160.49	$C_{116}H_{212}N_2O$	GalA-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_3$, 3-OHC18:0, 27-
	2174.43	2174.51	$C_{117}H_{214}N_2O$	GalA-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_2$, 3-OHC18:0, 3-
	2188.44	2188.52	$C_{118}H_{216}N_2O$	GalA-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_2$, 3-OHC18:0, 3-
	2202.44	2202.53	$C_{119}H_{218}N_2O$	GalA-GlcN-GlcN-GalA, 3-OHC14:0, 3-OHC15:0, 3-OHC16:0,
	2216.46	2216.55	$C_{120}H_{220}N_2O$	GalA-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_2$, $(3\text{-OHC18:0})_2$, 27-
	2230.47	2230.57	$C_{121}H_{222}N_2O$	GalA-GlcN-GlcN-GalA, $(3\text{-OHC14:0})_2$, 3-OHC15:0, (3-

Legend: "A"- Highest observed LipidA structures with proximal GlcNonate; "B"- Highest observed LipidA structures with proximal GlcN; nd- not detected

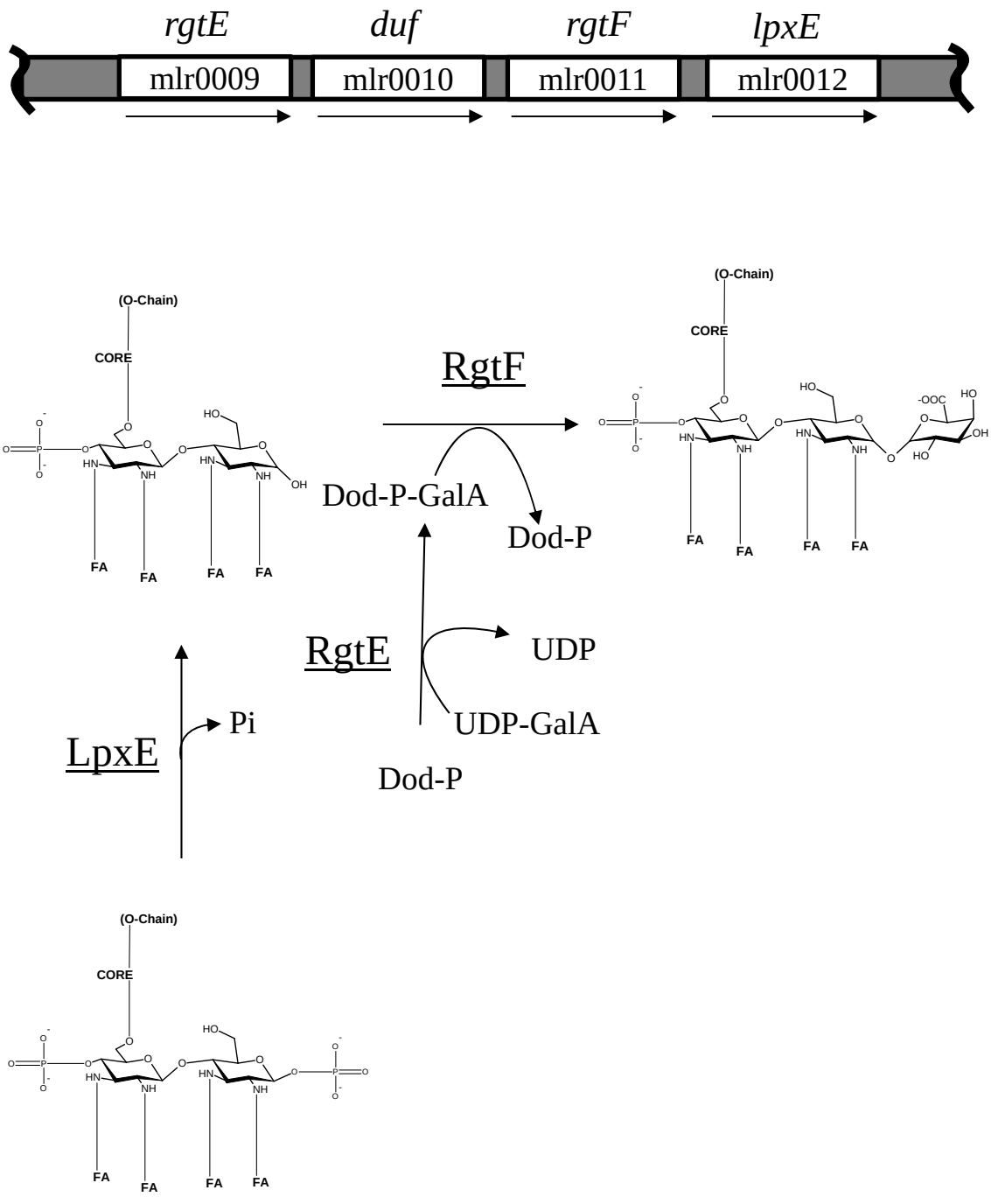

rgtE
duf
rgtF
lpxE
mlr0009
mlr0010
mlr0011
mlr0012
(O-Chain)
CORE
FA
RgtF
Dod-P-GalA
Dod-P
RgtE
UDP
UDP-GalA
Dod-P
LpxE
Pi

Figure 5.1

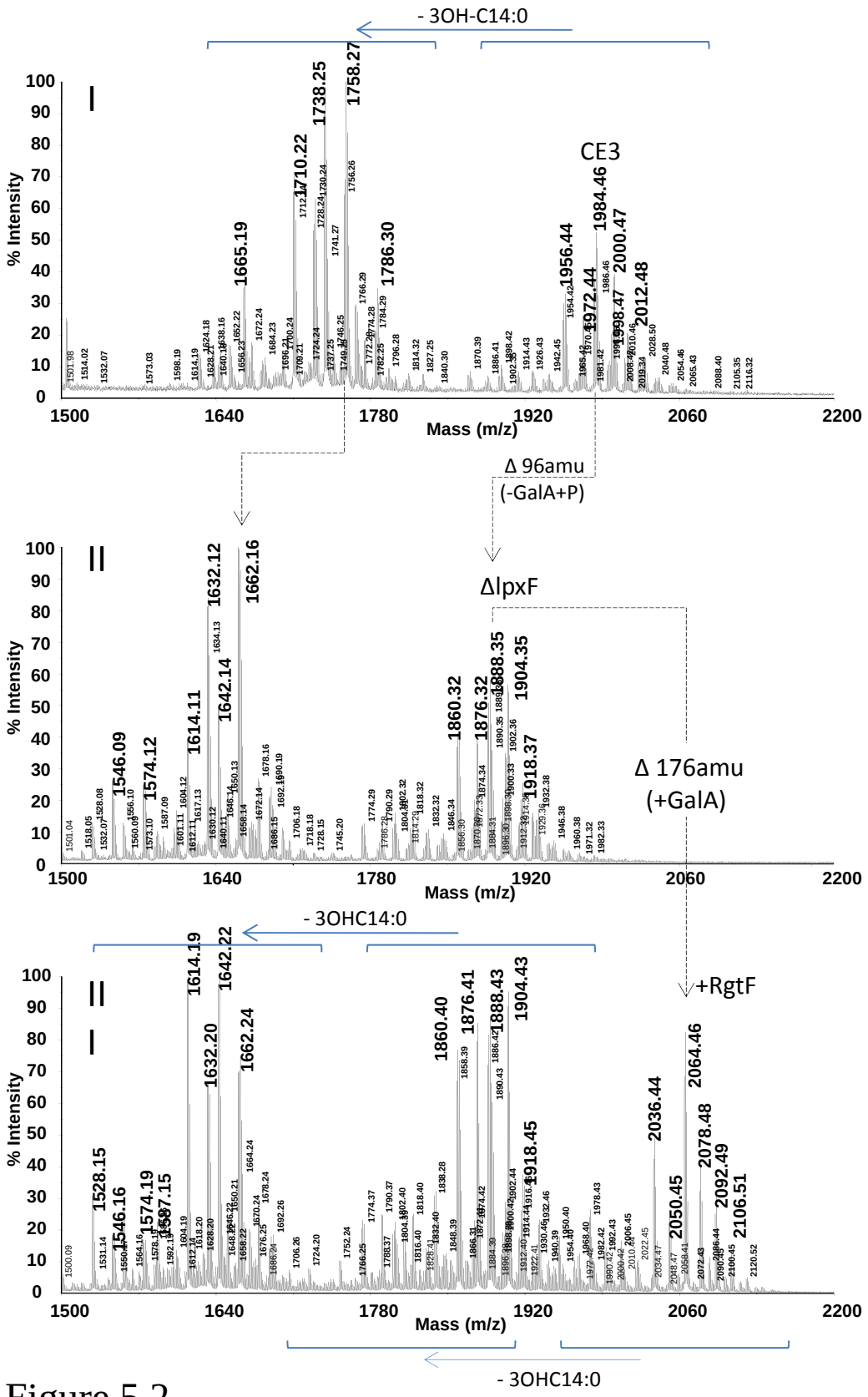

- 3OH-C14:0
I
CE3
1758.27
1738.25
1710.22
1665.19
1786.30
1956.44
1984.46
1972.44
2000.47
2012.48
% Intensity
Mass (m/z)
Δ 96amu
(-GalA+P)
II
ΔlpxF
1662.16
1632.12
1614.11
1642.14
1546.09
1574.12
1860.32
1876.32
1888.35
1904.35
1918.37
Δ 176amu
(+GalA)
- 3OHC14:0
+RgtF
II
I
1614.19
1642.22
1632.20
1662.24
1860.40
1876.41
1888.43
1904.43
1528.15
1546.16
1574.19
1597.15
1918.45
2036.44
2064.46
2050.45
2078.48
2092.49
2106.51
- 3OHC14:0

Figure 5.2

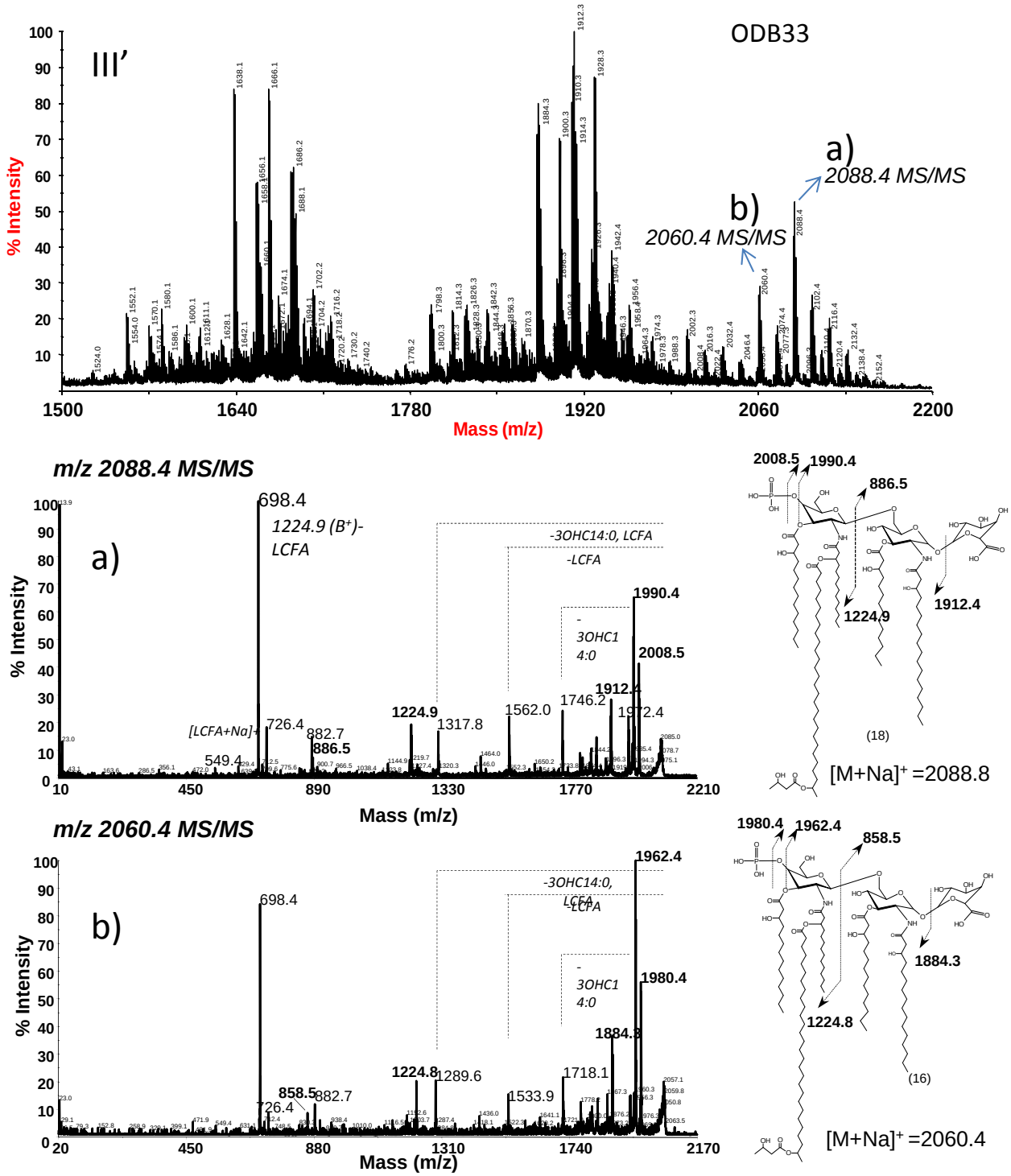

III'
ODB33
% Intensity
Mass (m/z)
a)
2088.4 MS/MS
b)
2060.4 MS/MS
m/z 2088.4 MS/MS
a)
698.4
1224.9 (B+)-
LCFA
-3OHC14:0, LCFA
-LCFA
1990.4
3OHC1
4:0
2008.5
1912.4
1224.9
1562.0
1746.2
1972.4
1317.8
[LCFA+Na]+
726.4
882.7
886.5
549.4
2008.5
1990.4
886.5
1912.4
1224.9
(18)
[M+Na]+ =2088.8
m/z 2060.4 MS/MS
b)
698.4
1962.4
-3OHC14:0,
LCFA
-LCFA
1980.4
3OHC1
4:0
1884.3
1224.8
1289.6
1718.1
858.5
882.7
726.4
1533.9
1980.4
1962.4
858.5
1884.3
1224.8
(16)
[M+Na]+ =2060.4

Figure 5.3

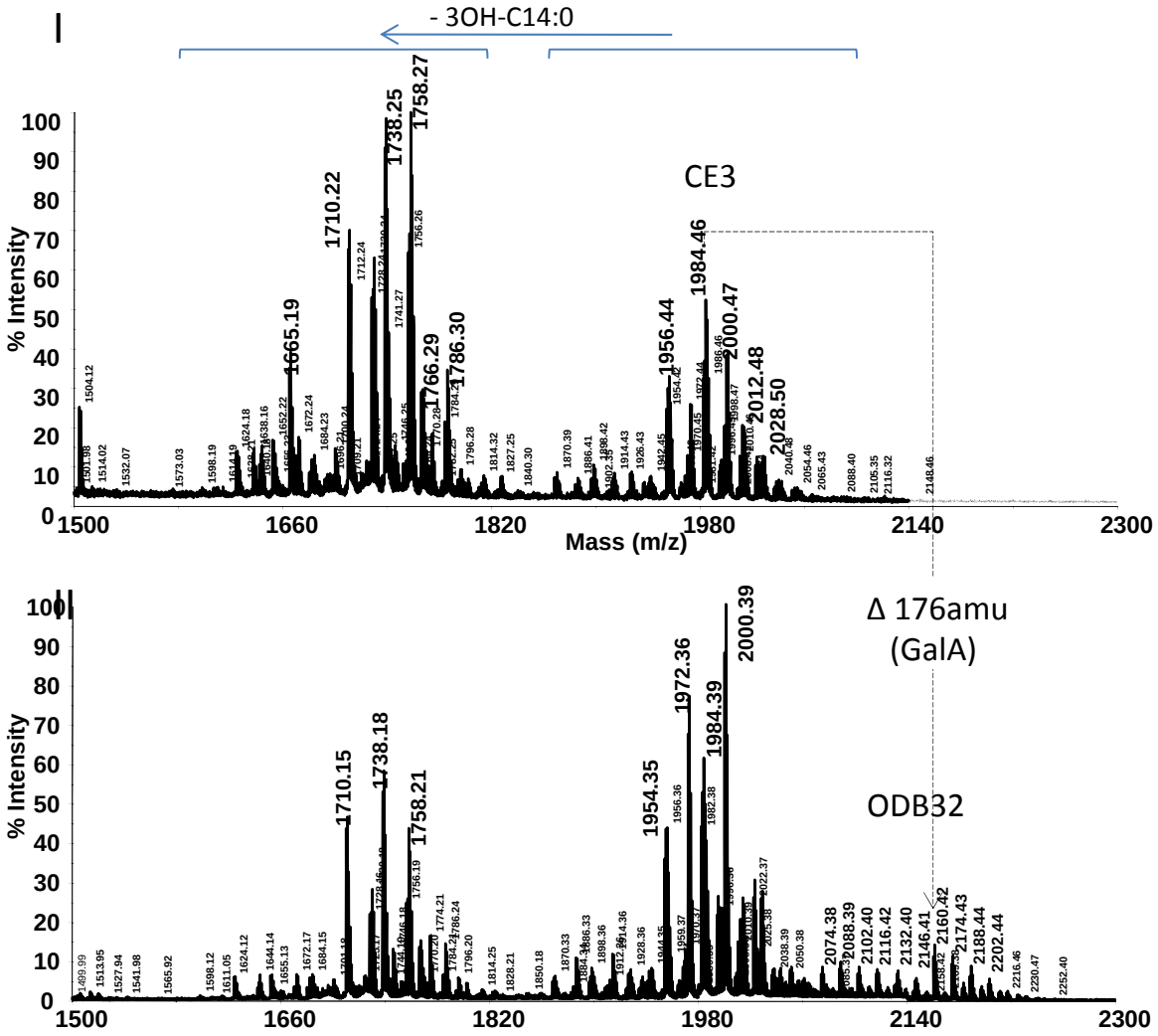

I
- 3OH-C14:0
CE3
% Intensity
Mass (m/z)
1758.27
1738.25
1710.22
1665.19
1766.29
1786.30
1956.44
1984.46
2000.47
2012.48
2028.50
Δ 176amu
(GalA)
ODB32
II
1710.15
1738.18
1758.21
1954.35
1972.36
1984.39
2000.39
2088.39
2102.40
2116.42
2132.40
2146.41
2160.42
2174.43
2188.44
2202.44

Figure 5.4

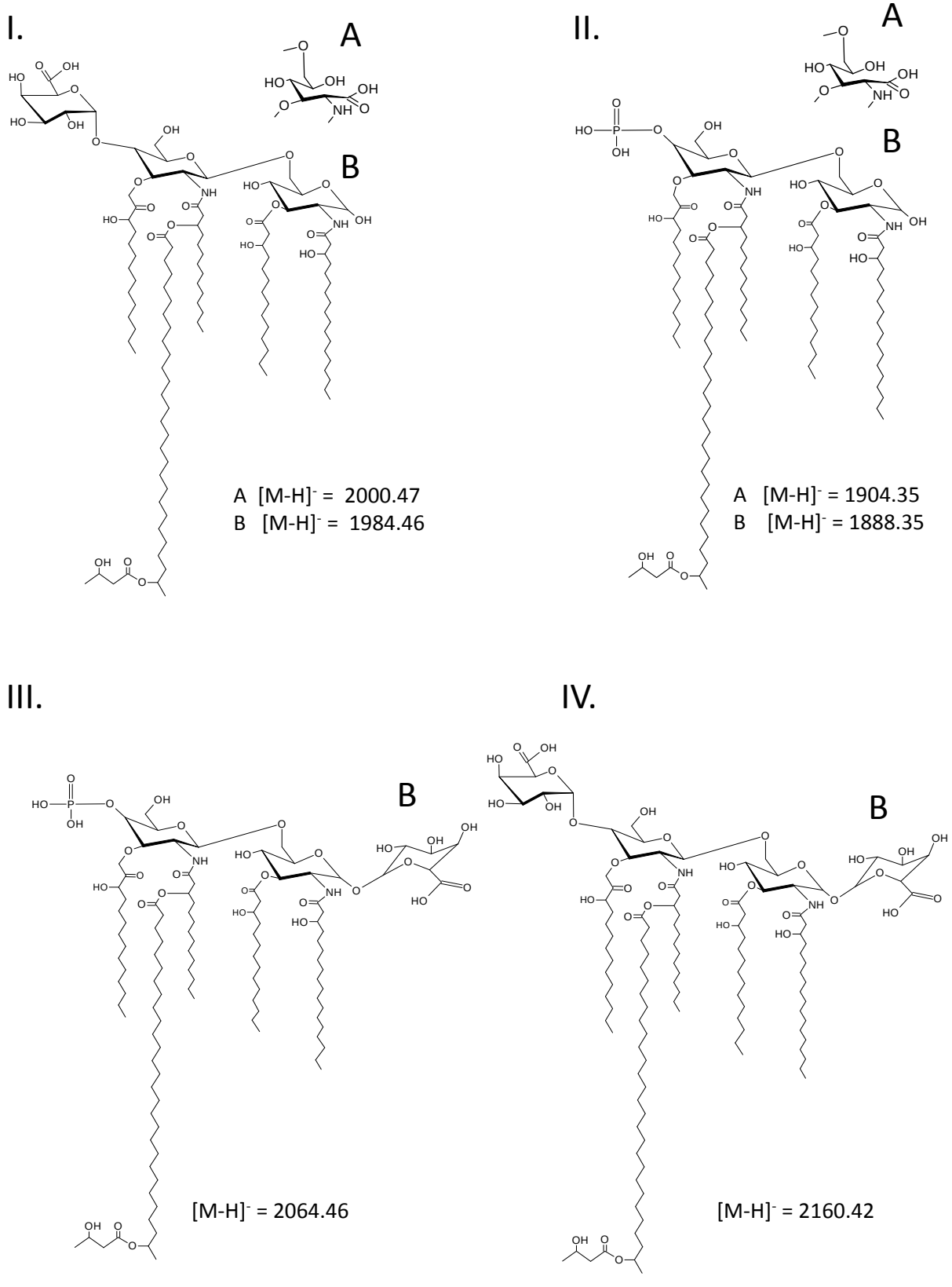

I.
A
B
A [M-H]- = 2000.47
B [M-H]- = 1984.46
II.
A
B
A [M-H]- = 1904.35
B [M-H]- = 1888.35
III.
B
[M-H]- = 2064.46
IV.
B
[M-H]- = 2160.42

Figure 5.5

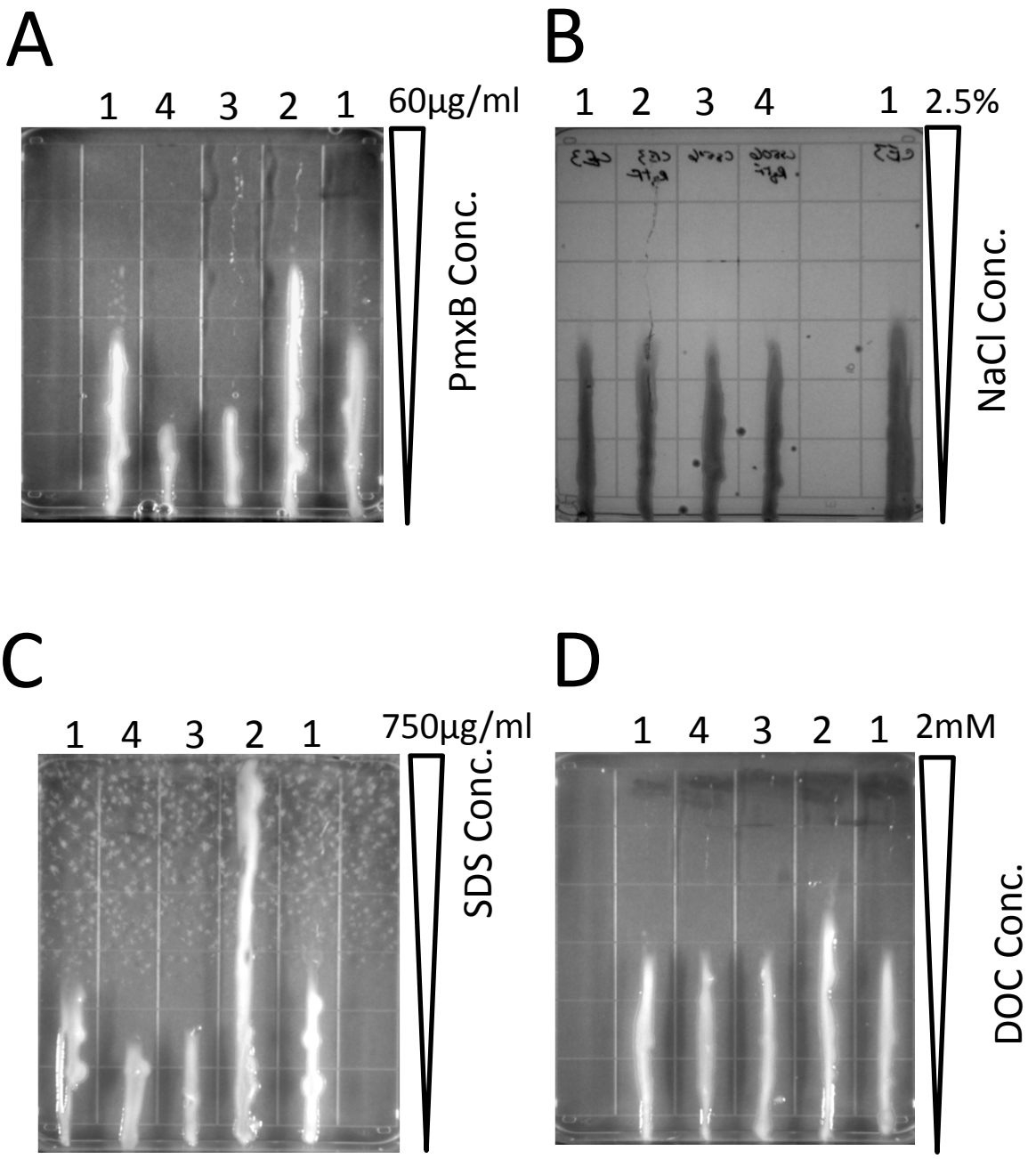

A
1 4 3 2 1
60µg/ml
PmxB Conc.
B
1 2 3 4
1 2.5%
NaCl Conc.
C
1 4 3 2 1
750µg/ml
SDS Conc.
D
1 4 3 2 1
2mM
DOC Conc.

Figure 5.6

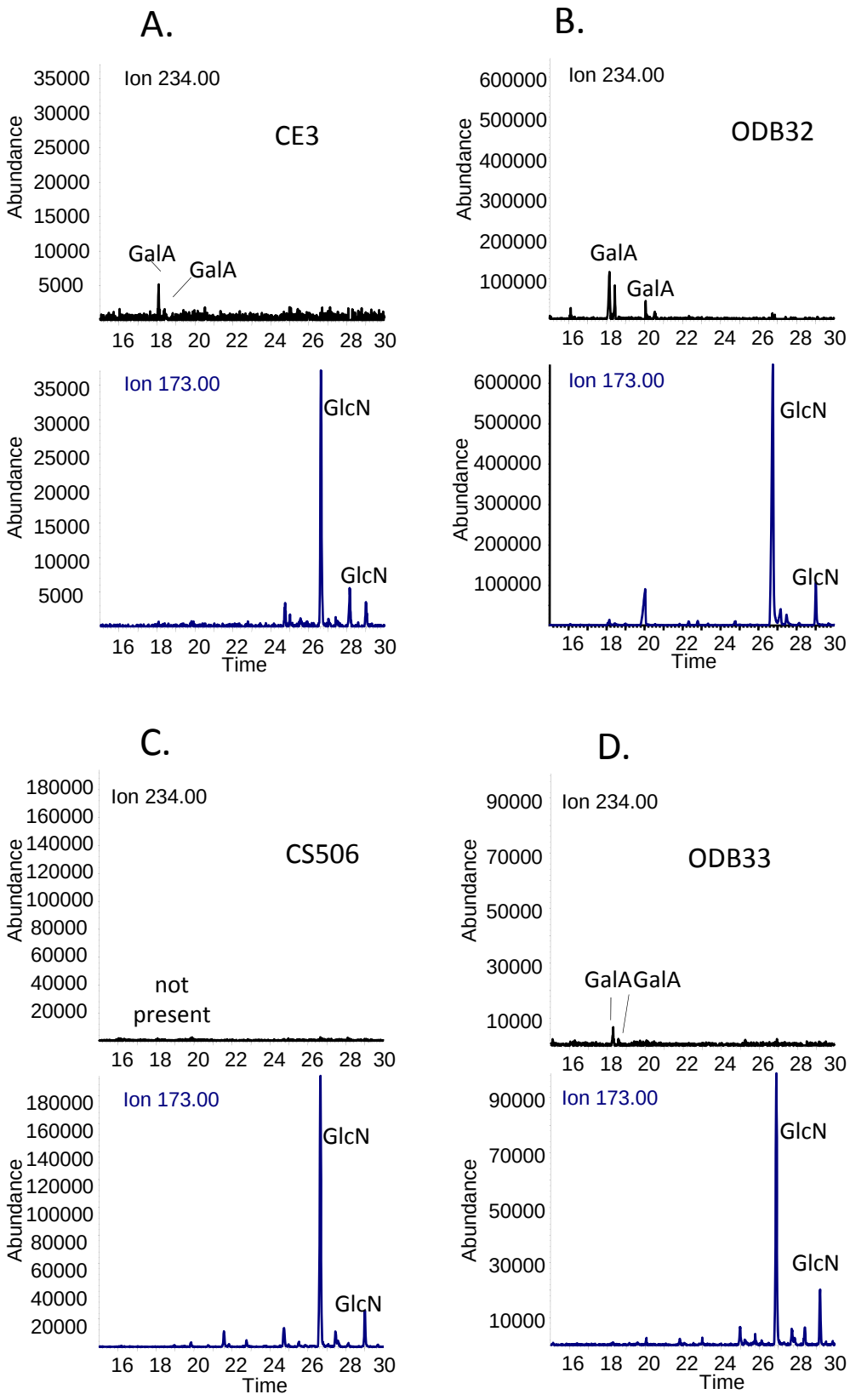

A.
Ion 234.00
CE3
GalA
GalA
Ion 173.00
GlcN
GlcN
Abundance
Time
B.
Ion 234.00
ODB32
GalA
GalA
Ion 173.00
GlcN
GlcN
Abundance
Time
C.
Ion 234.00
CS506
not present
Ion 173.00
GlcN
GlcN
Abundance
Time
D.
Ion 234.00
ODB33
GalA GalA
Ion 173.00
GlcN
GlcN
Abundance
Time

Figure 5.7