UTILIZING FLUORESCENTLY TAGGED UNDECAPRENYL PHOSPHATE AS A MEANS TO STUDY LIPID A MODIFICATION IN *BORDETELLA PERTUSSIS*

by

Theresa Ann Black

A thesis submitted to the faculty
of The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Chemistry

Charlotte

2022

Approved by:

_____
Dr. Jerry Troutman

_____
Dr. Brian Cooper

_____
Dr. Juan Vivero-Escoto

_____
Dr. Irina Nesmelova

ABSTRACT

THERESA ANN BLACK. Utilizing fluorescently tagged undecaprenyl phosphate as a means to study lipid A modification in *Bordetella pertussis.*
(Under the direction of DR. JERRY TROUTMAN)

Bordetella pertussis is a gram-negative bacterium that is the causative agent of whooping cough, a highly contagious disease that can be fatal in very young children. The outer membrane of the bacteria consists mainly of lipolysaccharide (LPS), and when LPS is modified with glucosamine, the bacteria can repel cationic antimicrobial peptides (CAMPs) which are currently the last known defense against multi-drug resistant bacteria. The mechanism by which the bacteria can modify its outer membrane has been predicted based on genetic analysis of *B. pertussis*, but up until now the activity of the enzymes responsible for the pathway has not been individually observed. Using a fluorescently tagged undecaprenyl phosphate analogue (an essential lipid carrier utilized in bacterial glycan synthesis pathways), I was able to monitor each step in the glucosamine modification pathway using fluorescence HPLC. Knowledge of the biochemical steps of this pathway is essential to further research into other ways to treat *B. pertussis* and other multi-drug resistant bacteria.

The glucosamine modification of lipopolysaccharide in *B. pertussis* is a three-step process thought to be performed by three enzymes: LgmA, LgmC, and LgmB (listed in the order in which they are proposed to act). In the first step, LgmA appends N-acetylglucosamine to undecaprenyl phosphate. In the second step, LgmC deacetlyates the product of LgmA to form undecaprenyl phosphate-glucosamine. Finally, LgmB appends the glucosamine to lipid A, a component of LPS. These three steps involve undecaprenyl phosphate, which is why fluorescently tagged bactoprenyl phosphate was chosen to probe this pathway. The difference in retention time of the fluorescent undecaprenyl phosphate analogue and fluorescent undecaprenyl

phosphate-linked sugars was directly observed via HPLC, with supporting data from mass spectrometry, to show the expected products after incubation of the enzymes with their proposed substrates, for the first time producing evidence of the individual enzymatic steps of glucosamine modification.

ACKNOWLEDGMENTS

I would like to thank everyone who has assisted me in this journey of attaining knowledge. First, I would like to thank soon-to-be Dr. Beth Scarbrough for hours spent teaching me how to express proteins and understand HPLC and LC-MS chromatograms, and for providing deletion mutants essential to this work. It was an immeasurable gift to have such a knowledgeable and kind mentor in my first few months in the lab, and a blessing to continue to work with you throughout my two years here. I would also like to thank Dr. Amanda Reid for introducing me to organic synthesis in a way that was extremely useful to my biochemistry research, and for your insatiable curiosity that usually meant if I had a question about something, you had already asked the same question and dove into the literature or took apart the LC-MS until you found the answer. I am grateful to you for sharing many of those answers with me in the short time that we were in lab.

Foremost I am grateful to Dr. Jerry Troutman for fostering a community of scientists that are curious, respectful, helpful, and team oriented. I have very much enjoyed my time in the Chemistry department of UNC Charlotte and in the company of the other Troutman group members. I have always known that I wanted to be a teacher “when I grew up,” but could never figure out what or where I would want to teach. Dr. Troutman was gracious enough to let me out of the lab to teach General Chemistry I for a semester, and I will forever be grateful for that opportunity, because I now know that I will happily be a chemistry instructor forever.

Many thanks to my committee for your helpful questions, comments, and direction. Thanks also to my other lab members for your insights and pleasant company for the last two years. I leave you my plants and hope that they bring you some happiness along with the sunshine that we are now able to see in our new lab!

DEDICATION

To all my former and future students, for the motivation to see this project through.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

LPS	lipolysaccharide
CAMP	cationic antimicrobial peptide
B. pertussis	*Bordetella pertussis*
HPLC	high performance liquid chromatography
LgmA	a glycosyltransferase
LgmC	a deacetylase
LgmB	a glycosyltransferase
Ara4N	4-aminoarabinose
CPS	capsular polysaccharide
MOE	Metabolic Oligosaccharide Engineering
FucNAc	N-acetylfucosamine
DATDG	2,4-diacetamido-2,4,6-trideoxygalactose
GlcNAc	N-acetylglucosamine
UDP-GlcNAc	uridine diphosphate-linked GlcNAc
^{3}H-GalNAc	^{3}H-labeled N-acetylgalactosamine
PglC	a glycosyltransferase
PglA	a glycosyltransferase
Bac	bacillosamine
^{32}P-ATP	^{32}P-labeled adenosine triphosphate
C. jejuni	*Campylobacter jejuni*
E. coli	*Escherichia coli*
ABC	ATP-binding cassette

UppS	undecaprenyl pyrophosphate synthase
IPP	isopentenyl diphosphate
FPP	farnesyl diphosphate
UppP	undecaprenyl pyrophosphate phosphatase
S. mutans	*Streptococcus mutans*
TLC	thin layer chromatography
2NA	*p*-nitro-anilino
ESI-MS	electrospray ionization mass spectrometry
m/z	mass/charge ratio
BP	bactoprenyl phosphate
2AA	2-antranilimide
WcfS	phosphoglycosyltransferase
B. fragilis	*Bacterioides fragilis*
AADGal-1-P	N-acetyl-4-amino-galactosamine-1-phosphate
2CN	2-nitrileaniline
Fl	fluorescently tagged
PGT	phosphoglycosyltransferase
CPS2E	a phosphoglycosyltransferase
WbaP	a phosphoglycosyltransferase
WecA	a phosphoglycosyltransferase
WecP	a phosphoglycosyltransferase
Kdo	Ketodeoxyoctonic acid
TLR4	toll-like receptor 4

LpxA	enzyme in the lipid A synthesis pathway
LpxC	enzyme in the lipid A synthesis pathway
LpxD	enzyme in the lipid A synthesis pathway
LpxH	enzyme in the lipid A synthesis pathway
LpxB	enzyme in the lipid A synthesis pathway
LpxK	enzyme in the lipid A synthesis pathway
KdtA	enzyme in the lipid A synthesis pathway
LpxL	enzyme in the lipid A synthesis pathway
LpxM	enzyme in the lipid A synthesis pathway
ACP	acyl carrier protein
UMP	uridine monophosphate
CMP	cytidine monophosphate
GlcN	glucosamine
Ara4FN	N-formylated Ara4N
ArnC	a glycosyl transferase
ArnD	a deformylase
ArnT	a glycosyl transferase
F. novicida	*Francisella novicida*
FlmF2	a glycosyltransferase
NaxD	a deacetylase
FlmK	a glycosyltransferase
Lgm	lipid A glucosamine modification
LgmD	a putative inner membrane flippase

MALDI	Matrix-Assisted Laser Desorption Ionization
YdjC	a superfamily of deacetylases
C41	*E. coli* expression strain
RP	*E. coli* expression strain
LB	lysogeny broth (Luria Bertani)
OD_{600}	optical density at 600 nm
cef	cell envelope fraction
PCR	polymerase chain reaction
DMSO	dimethyl sulfoxide
MSD	mass selective detection
FLD	fluorescence detector
kD	kiloDalton
Ni-NTA	nickel nitrilotriacetic acid
MntS	small protein in *E. coli*
Tdap	tetanus and diphtheria toxoids with acellular pertussis
SIM	selected ion monitoring

CHAPTER 1: INTRODUCTION

1.1: Basic Background of Bacteria

The amount of bacteria on Earth is astounding. Biomass is defined as the total mass of organisms in a given area or volume, and the total biomass of planet Earth is estimated to be 546 gigatons. Humans weigh about 0.06 gigatons, which accounts for only 0.01% of the total biomass. In contrast, the biomass of bacteria is about 70 gigatons, which means bacteria make up as much as 12.8% of the total biomass (Figure 1). There are about 30,000 named species of bacteria, but this is estimated to be only 0.001% of the total number of bacterial species, which leaves 3×10^9 species yet to be discovered, isolated, and named.[1] All of these numbers indicate that bacteria are extremely common and our knowledge of them is limited to a small subset.

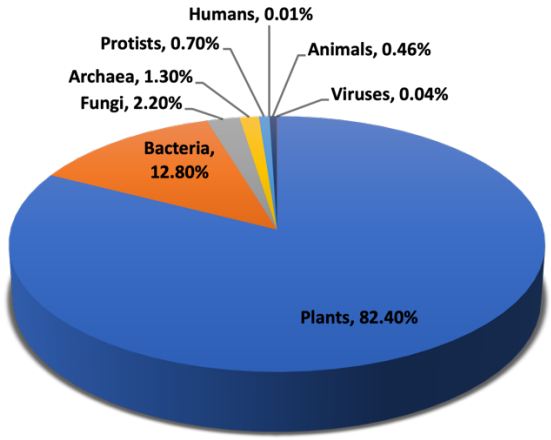

Figure 1: Breakdown of Earth's Biomass.[1]

While the thought of 70 gigatons of bacteria might be conceerning, many of these bacterial species are innocuous, and many bacteria coexist in the human body with no detrimental effects. In fact, several species, such as *Bacterioides*, are responsible for aiding in the digestion of plant materials. About three pounds of bacteria are said to live in the human gut, specifically the large intestine, and are able to digest carbohydrates in plant cell walls that human enzymes cannot digest on their own. Another example of beneficial bacteria is *Lactobacillus*, for which many benefits have been shown, one of which is fighting off *Helicobacter pylori*, the harmful bacteria that causes stomach ulcers.[6]

Current major issues in biomedicine as it relates to bacteria include bacterial infections, multi-drug resistant bacteria, and issues with vaccines losing efficacy over time. There are many common bacterial infections, some of which are fairly low in danger, and others which can be life-threatening. A major issue has been overtreatment of infections with antibiotics that kills beneficial bacteria as well as harmful ones, and repeated treatment with antibiotics that creates multi-drug resistant bacteria especially when full antibiotic regimens are not completed. The third issue is the decline in vaccine effectiveness which can be due to a number of things and can lead to a resurgence in previously eradicated diseases.

1.2: Bacterial Glycans

One of the ways that bacteria infect their hosts and gain drug and vaccine resistance is through glycans. Glycans are saccharides that are attached to the bacterial surface or secreted outside the cell and play a major role in bacteria-host-cell interactions (Figure 2).[7, 8] The modification of the outer membrane of bacteria with glycans causes resistance to cationic antimicrobial peptides (CAMPs), and many glycans are crucial to bacterial survival.[9-12] With the large number of monosaccharides, sugar modifications, and different linkages, glycans have great potential for variety and investigation, and a lack of knowledge of glycan assembly pathways has seriously impeded research into specific inhibitors of these pathways in many priority pathogens.[13]

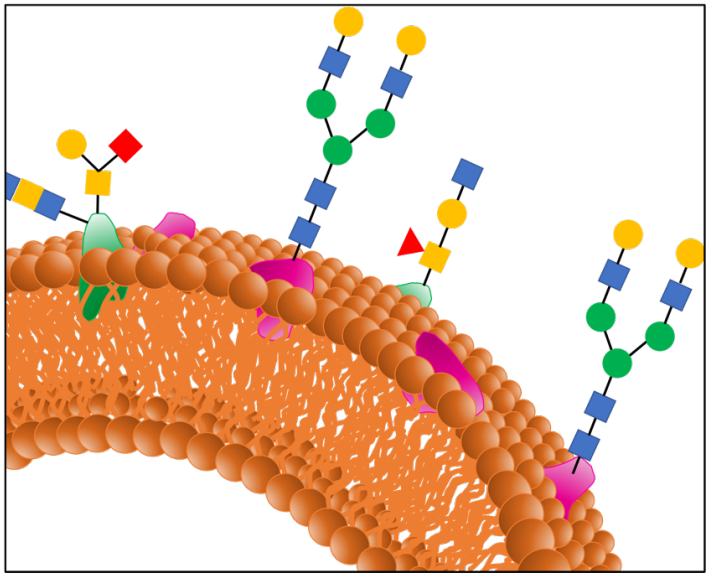

Figure 2: Schematic representation of a cell surface studded with glycans.

1.3: Methods for Investigating Bacterial Glycan Synthesis Pathways

1.3.1: Overview of Glycan Investigation Methods

Although the variety of bacterial glycans is vast, there is a common sequence that is usually followed in the gaining of knowledge about the synthesis pathway of each one.[14] First a phenotypic difference is observed, the presence or absence of a glycan in certain strains or species of bacteria. Genomic studies in the fields of bacterial genetics and bioinformatics may find a single gene or perhaps a locus, a cluster of co-located genes, that seems to be responsible for the phenotypic difference: the presence of the glycan is associated with the presence of certain genes. Then the role of each individual gene that produces an enzyme must be determined. First the roles of enzymes can be proposed based on a number of different analyses and comparisons. Sequence analysis of the enzyme may find homologs with known function; this information can be used to assign an enzyme into a superfamily and propose a broad function.[15] Crystal structure analysis can also be used to find structurally similar enzymes.[16] Sequence analysis of the gene locus may find similarity to other glycan synthesis pathways, this similarity can be used to propose possible functions for each of the enzymes.[2, 17] Sequence analysis of the proteins can also suggest a location for each enzyme, whether membrane bound, membrane associated, or soluble.[10] The vast majority of glycans are assembled in similar pathways, by which one monosaccharide at a time is appended to a lipid anchor in the correct order to form the resulting complex polysaccharide glycan.[14, 18, 19] The final polysaccharide unit provides more insight into which monosaccharides are appended and in which order.[20, 21] After taking all the information together, scientists can propose a comprehensive picture of the biochemical synthesis pathway for a glycan of interest (Figure 3).[10]

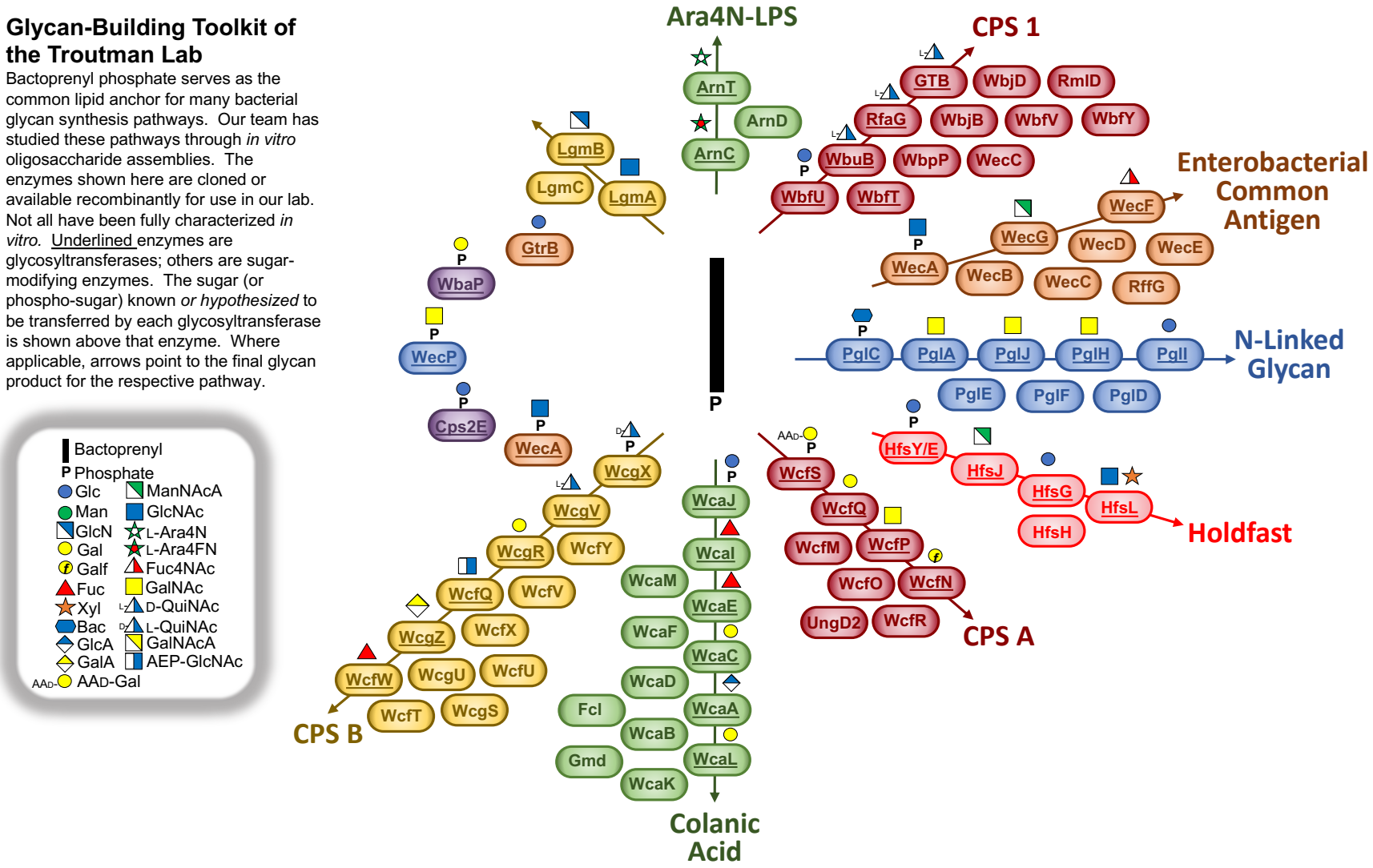

Figure 3: Glycan synthesis pathways investigated in the Troutman group.[3]

Once a scheme is predicted, experiments must be designed to probe each step in the enzymatic pathway and monitor the function of each enzyme. The gap in knowledge is identifying the precise role of each enzyme in the enzymatic pathway. There can be many enzymes in a single pathway, some that add sugars to the glycan and some that modify sugars, and a host of different substrates. One glycan investigation technique is Metabolic Oligosaccharide Engineering (MOE), which involves labeling monosaccharides that are taken up in to the cells then observing the structurally altered glycans that result.[22, 23] Several different labeling techniques have been utilized, some of which are azide labeling and radio labeling.[13, 22, 24, 25] Another technique does not involve any special labeling, but looks at mass spectrometry of cell lysates to observe the presence or absence of certain glycan intermediates in deletion mutants.[26] I will give a brief overview and discuss an example of each of these techniques.

1.3.2: Metabolic Oligosaccharide Engineering – Azide Labeling

The first example of MOE is azide labeling of monosaccharides. These unnatural azide-containing sugars are incorporated into cell media, taken up into the cells, and then metabolically incorporated into bacterial glycans (Figure 4).[23, 27, 28] Azides are useful because of their participation in "click" chemistry, where they can be tagged with alkyne or phosphine reaction partners using highly thermodynamically favorable reactions to enable easier detection, identification, and visualization.[13, 29, 30] Azide labeling (as well as other methods of MOE) can be useful in finding inhibitors for bacterial pathways without investigating all the individual steps,

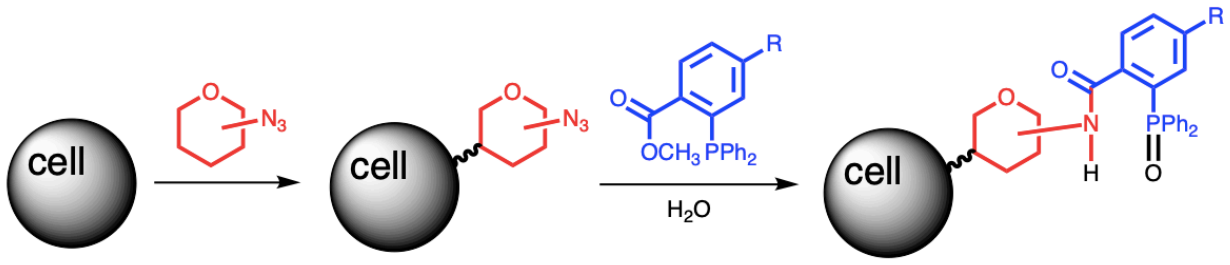

Figure 4: Azide labeling of monosaccharides coupled to the Staudinger ligation.

and it can also can be used to shed light on individual enzymatic steps. Azide labeling is particularly useful in determining which rare or unusual sugars are utilized by particular bacterial species, a process which is commonly a bottleneck in enzymatic assays and can provide insight into narrow-spectrum antibiotics or inhibitors based on the rare sugar usage.

An important use of azide labeling is the Staudinger ligation, which is a highly selective reaction between azide sugars and triaryl phosphines (Figure 4). Neither of these functional groups react with other common biological entities, and both are stable in biological systems. A biological probe can be incorporated with the phosphine that allows for easy visualization of the azide complex after reaction, either on the live cell surface or in cell lysates.[22, 27, 28, 31] The Staudinger ligation was first applied to sialic acid in eukaryotic cells, but has since been shown useful in prokaryotic applications as well.[22] Dube and coworkers investigated the use of azide-containing analogues of the rare bacterial sugars N-acetylfucosamine (FucNAc), bacillosamine, and 2,4-diacetamido-2,4,6-trideoxygalactose (DATDG) in pathogenic bacteria and found that the rare azidosugars were indeed taken up into the cells and eventually displayed on the surface in the final glycan structure.[31] These rare bacterial sugar analogues were more readily displayed than the common monosaccharide precursor N-acetylglucosamine (GlcNAc) and indicated that these analogues were suitable for additional studies that could provide insight into bacterial glycan assembly pathways.[13, 22, 31]

1.3.3: Metabolic Oligosaccharide Engineering – Radiolabeling

Radiolabeling techniques in bacterial cells has mostly been limited to classical ^{3}H labeling of monosaccharides, a process that had been utilized previously in eukaryotic cells specifically in other studies of sialic acid.[28] ^{14}C labeling of monosaccharides has also been employed, but radiolabeled sugars are extremely expensive. One can purchase 25 mg of uridine

diphosphate-linked GlcNAc (UDP-GlcNAc) for about $80, but the smallest available unit of ^{3}H-labled UDP-GlcNAc costs over $2,000. Imperiali and coworkers utilized a similar radiolabeled monosaccharide, ^{3}H-labeled N-acetylgalactosamine (^{3}H-GalNAc), to show the activity of enzymes PglC and PglA in the *Campylobacter jejuni* pathway that synthesizes the N-linked glycan. The product was formed by combining undecaprenyl phosphate with PglC and PglA, the first two glycosyltransferases in the pathway. PglC transfers 1-phospho-2,4-diacetamido bacillosamine (Bac) from UDP-Bac to give undecaprenyl diphosphate-linked Bac. The second enzyme in the pathway is PglA, which transfers GalNAc from UDP-GalNAc to form undecaprenyl diphosphate-linked Bac-GalNAc. Formation of the product was monitored by observing transfer of the radiolabeled ^{3}H-GalNAc from the aqueous phase to the organic phase.[32-34]

Another type of radiolabel, ^{32}P, has been useful in studying a category of enzymes that catalyze the production of precursor molecules essential to glycan synthesis. Undecaprenyl phosphate is the 55-carbon lipid scaffold common to most bacteria on which glycans are sequentially synthesized (Figure 5). Undecaprenyl phosphate can be formed by phosphorylating undecaprenol, a function which was shown to belong to a diacylglycerol kinase in *Streptococcus mutans* by Imperiali and coworkers. The researchers used ^{32}P-ATP as the phosphate donor in these experiments then normal phase HPLC to isolate the

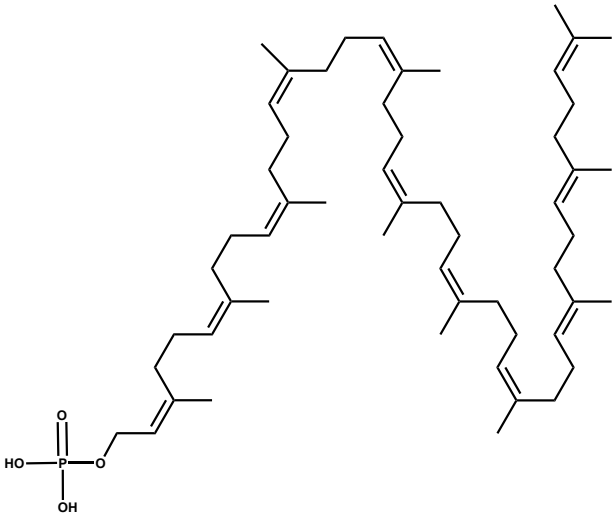

Figure 5: Undecaprenyl phosphate.

product, TLC to confirm the undecaprenyl phosphate product formation, and scintillation counting to monitor the reaction rate of the diacylglycerol kinase.[34]

1.3.4: Mass Spectrometry Analysis of Deletion Mutants

An excellent example of illuminating an entire glycan synthesis scheme via *in vivo* investigation of deletion mutants is the work by Wren and coworkers involving the N-linked glycosylation process in *C. jejuni*.[26] Eleven total putative genes had been assigned functions in the construction of a heptasaccharide glycan product based on genetic analysis and comparison to other enzymes of known function. The total function of the gene locus had been determined by importing the locus into *Escherichia coli* cells and observing N-linked protein glycosylation where none had been previously observed. Wren and coworkers proceeded to delete genes one by one, let the *E. coli* cells grow, harvest the cells, and use mass spectrometry to analyze the lysate. They found evidence of the expected N-linked glycan intermediates that built up because of each subsequent gene deletion using tandem mass spectrometry. One observation they made with two of the N-linked glycosylation enzymes was that even with the deletion of either or both of the genes, the complete heptasaccharide was still formed. They attributed the formation of product to two structurally similar enzymes endogenous to *E. coli*, a putative inner membrane-located ABC transporter and a putative acetyltransferase. This observation is something that occurs in many glycan synthesis pathways from other prokaryotic organisms and is one drawback of recreating synthesis pathways *in vivo*.[35] There are ways to circumvent this occurrence, one of which is also deleting the homologous genes from *E. coli*. However, this only works if the homologous genes have already been identified. There are no doubt many enzymes in *E. coli* whose functions are not yet known and which may interfere with *in vivo* experiments.

1.4: Undecaprenyl Phosphate – the Commonality in Bacterial Glycan Synthesis Pathways

1.4.1: Introduction to Undecaprenyl Phosphate

Eukaryotes and prokaryotes alike have wide variation in their glycans, but researchers of eukaryotic protein glycosylation have found many trends and predictable routes that they can exploit when investigating a new glycosylation pathway. In the prokaryotic world, however, no such trends have been yet discovered, and there are many rare and unusual sugars utilized by different species of bacteria. The discovery of a ubiquitous lipid carrier in bacterial glycan synthesis pathways was a major step because at last there was a commonality in bacterial pathways that could be exploited. That lipid carrier is known as undecaprenyl phosphate, a 55-carbon linear polymer of isoprene units that serves as the scaffold or anchor for glycan assembly from individual monosaccharides (Figure 5).[19, 36] Undecaprenyl phosphate anchors the glycan assembly process in the bacterial membrane where glycans are built and modified sequentially, and then it is recycled when the finished glycan is removed.[19, 36] Glycosyltransferases and phosphoglycosyltransferases append sugars or phosphosugars to undecaprenyl phosphate, and sugar-modifying enzymes such as deacetylases or deformylases modify the glycan as well.[19, 36] Usually a flippase is necessary to flip the glycan (still attached to undecaprenyl phosphate) across the membrane, and another enzyme is responsible for removing the glycan from undecaprenyl phosphate and appending it to its final location, whether that is a glycoprotein or glycolipid.[37] It is at this point that undecaprenyl phosphate is recycled and returned to the inner membrane where it will participate in another cycle of the glycan assembly pathway, whether the same one or a different one.[19, 36]

Undecaprenyl phosphate is mainly conserved among almost all species of bacteria, although there are some slight differences in the size in a few species (dodecaprenyl phosphate with 60 carbons has been observed in *Francisella novicida* and *Rhizobium leguminosarum*).[17, 38] The analogue of undecaprenyl phosphate in eukaryotes is identical with the exception of one *cis* bond instead of a *trans* bond and is utilized in plants. Animals utilize dolichyl phosphate which has a similar structure to undecaprenyl phosphate from plants, but the alpha isoprene unit adjacent to the phosphate is saturated.[19, 39] Undecaprenyl diphosphate in bacteria is biosynthesized *de novo* by undecaprenyl pyrophosphate synthase (UppS), a *cis*-prenyltransferase which is highly conserved in bacterial species.[19] UppS catalyzes the condensation of eight isopentenyl diphosphate (IPP) units to a single unit of farnesyl diphosphate (FPP). Undecaprenyl diphosphate is then dephosphorylated to form undecaprenyl phosphate by undecaprenyl pyrophosphate phosphatase (UppP) (Figure 6).[19, 39] Undecaprenyl diphosphate is the form the lipid carrier is in when it is released after the finished glycan is removed, and it can be recycled by UppP to form undecaprenyl phosphate.[19, 39] Undecaprenol (undecaprenyl phosphate with an alcohol functional group instead of a phosphate) can also be phosphorylated by a diacylglycerol kinase to form undecaprenyl phosphate (example from *S. mutans* above).[34]

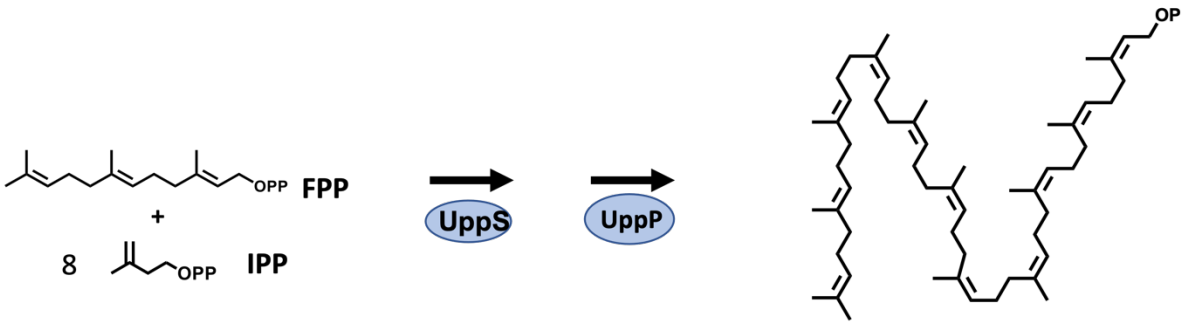

Figure 6: Biosynthesis of undecaprenyl phosphate by UppS and UppP.

As previously mentioned, endogenous enzymes in *E. coli* may interfere with *in vivo* studies, and *in vitro* studies may be preferred in some instances to observe enzyme activity in a more controlled environment. Necessary components for such *in vitro* studies include buffers and salts, purified enzymes, sugar substrates, and undecaprenyl phosphate. Buffers and salts are usually inexpensive and easy to obtain. Some sugar substrates are relatively inexpensive; although as previously mentioned, the rare and unusual sugars are quite expensive or simply not commercially available. Undecaprenyl phosphate, the last component, with its large size and specific bond configurations, would be extremely difficult and time consuming to chemically synthesize. C55 undecaprenol from a plant (*Magnolia kobus*) was sometimes used to form undecaprenyl phosphate, but the plant undecaprenol is a 3-*trans*, 7-*cis* isoprene, while the native form of undecaprenyl phosphate is 2-*trans*, 8-*cis*, so it was not ideal.[36, 40] Chemoenzymatic synthesis using its known enzymes is the preferred way to obtain this substrate and will be explained in further detail in a later section.

When undecaprenyl phosphate is obtained and *in vitro* experiments can be conducted with glycotransferases and sugar-modifying enzymes, the next hurdle is the method of analyzing the results. Previous groups have used radiolabeling and a scintillation counter, TLC, HPLC coupled with various detection methods, and mass spectrometry. Of the four methods listed above, mass spectrometry is thought to give the surest method of determining if an intermediate glycan or final product has indeed been obtained; however, it is also by far the most expensive. HPLC is a cheaper method, but undecaprenyl phosphate and the intermediate coupled glycans do not inherently have any characteristics that make them easy to visualize on HPLC. That is why all the different radiolabeling techniques were initially employed, but radiolabeled sugars are also very expensive. In order to easily visualize glycan intermediates on HPLC (or TLC), the

Troutman group has synthesized several undecaprenyl phosphate analogues with different functional groups. The next sections will summarize the Troutman group's work with undecaprenyl phosphate analogues over the past decade.

1.4.2: Troutman Group – Synthesis and Use of Farnesyl Diphosphate Analogue with UppS

The first step in the long journey to the fluorescently tagged undecaprenyl phosphate analogue (bactoprenyl phosphate) that the Troutman group utilizes currently was the realization that the drawback of the chromophore-lacking ubiquitous undecaprenyl phosphate molecule needed to be turned into an asset. Previous work with *C. jejuni* had indicated that alterations in the polyisoprene structure did not affect the ability of glycosyltransferases to build the N-linked glycan on an isoprenoid, unless the alterations were too close to the monophosphate.[41] Analogues of FPP with additions of various functional groups had been previously synthesized, but few had been tested with UppS, and none had been used as specifically to produce alternative substrates to probe biosynthetic pathways .[42, 43] The decision was made to synthesize analogues of FPP and determine if they could be utilized by UppS.

Instead of radiolabeling (because of its expense and potential safety issues), a moiety with a strong ultraviolet absorbance (E_{395nm}=9500 M^{-1} cm^{-1}) originally designed by the Spielmann group was chosen.[44] The FPP analogue, *p*-nitro-anilinogeranyl diphosphate (NA-GPP) (Figure 7), was incubated with IPP and UppS from *Bacteroides fragilis*

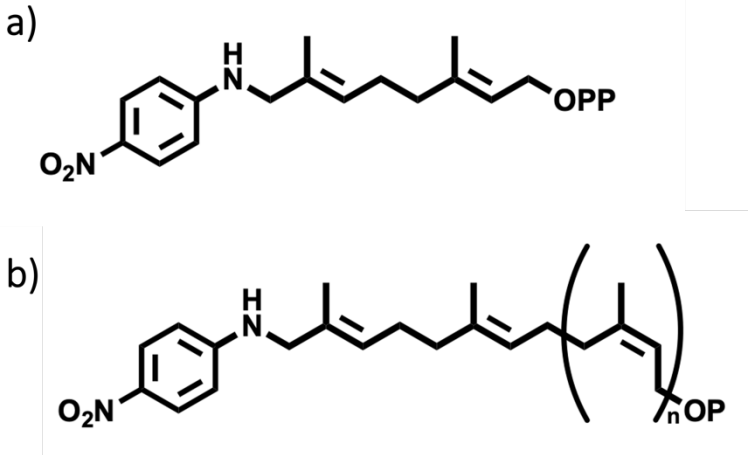

Figure 7: a) FPP analogue 2-anilinogeranyl diphosphate (2NA-GPP) b) P-nitro-anilino bactoprenyl phosphate (2NA-BP).

overnight and the resulting polyisoprenoids were extracted with chloroform and analyzed by TLC. A product was found to have a retention factor nearly identical to the product formed with the native substrate FPP (Figure 7). In order to determine the chain length or number of isoprene additions, reverse phase HPLC on a C18 column and a mobile phase of 1:1 ammonium bicarbonate and *n*-propanol was used. The HPLC had an absorbance detector, and two major peaks were observed. The peaks were isolated and characterized by negative mode electrospray ionization mass spectrometry (ESI-MS), where the *m/z* values were consistent with a product with six and seven isoprene additions to the 2NA-GPP substrate (not shown).[42]

The isoprenoid products obtained were in the diphosphate form, and the group was interested in obtaining the monophosphate for use in glycosyltransferase assays. The previous known method was to use a pyrophosphatase to cleave the diphosphate group completely, and then a kinase would be used to phosphorylate the alcohol. Instead of two steps, the Troutman group used alkaline phosphatase to cleave just one phosphate group to prepare the monophosphate form of bactoprenol. ESI-MS confirmed the *m/z* of the new product consistent with the expected (not shown).[42]

After the isolation and characterization of the *p*-nitro-aniline bactoprenyl phosphate analogue (2NA-BP), it was tested with well characterized glycosyltransferases from the N-linked glycosylation pathway in *C. jejuni* (mentioned in more detail above). The first test was to see if radiolabeled sugars would indeed be incorporated into NA-BP when incubated with PglC and PglA, the first two glycosyltransferases in the *C. jejuni* pathway. After incubation and isolation of the isoprenoids, TLC, scintillation counting, HPLC, and ESI-MS indicated that the sugars were indeed incorporated into the 2NA-BP analogue and that the analogue was an acceptable substrate for all the enzymes tested (not shown).[42]

The 2NA-BP probe was a successful first new probe developed to study oligosaccharide assembly in bacteria. Sufficient quantities of NA-BP-linked disaccharides were obtained to use in downstream assays with other glycosyltransferases, and the strongly absorbent tag was easy to visualize on HPLC. Once the identity of the products were further confirmed by ESI-MS, other assays would be relatively quick to analyze via HPLC without the repeated use of MS. A standard curve of product amounts was even formed that gave insight into enzyme turnover and could be used for kinetic assays. Additional probes with enhanced photophysical properties were being designed by the Troutman lab to increase sensitivity of these assays.[42]

1.4.3: Troutman Group – Fluorescent Undecaprenyl Phosphate Analogue

After the success of a highly absorbent bactoprenyl phosphate (BP) analogue, the Troutman group aimed to design more powerful probes to study complex oligosaccharide biosynthetic pathways, specifically one with a fluorescent tag.[45] Fluorescent tags had been previously utilized in the study of glycosylation pathways, by incorporation of an anthranilamide (AA) into the free oligosaccharide after hydrolysis from the lipid anchor.[33] Instead, the Troutman group was interested to see if the fluorescent AA tag (E_{350nm}=4500 M^{-1} cm^{-1}) could be incorporated directly into BP to form a fluorescent analogue to track glycan synthesis pathways. 2AA-GPP (Figure 8a) was synthesized and incubated with IPP and UppS to form 2AA-BP (Figure 8b) which was extracted with chloroform.[45]

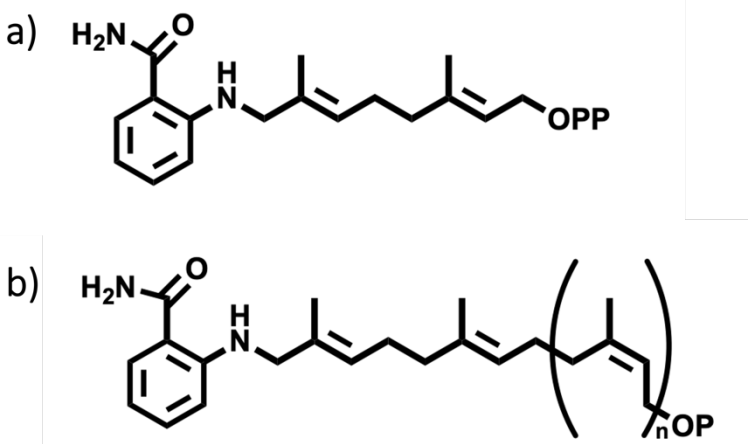

Figure 8: a) 2-Anthranilimide geranyl diphosphate (2AA-GPP) b) 2-Anthranilamide bactoprenyl phosphate (2AA-BP).

Interestingly, during HPLC and ESI-MS analysis of the isoprenoid products, it was discovered that 2AA-BP formed was from six to eight isoprene additions. 2NA-BP (the previous analogue) had only been observed to form from seven additions as its longest isoprenoid. HPLC analysis also indicated that 2AA-BP had a detection sensitivity of more than three orders of magnitude better than 2NA-BP. 2AA-BP was then incubated separately with PglC and its sugar substrate and another phosphoglycosyltransferase, WcfS from *B. fragilis* and its sugar substrate. WcfS catalyzes the transfer of N-acetyl-4-amino-galactosamine-1-phosphate (AADGal-1-P) from UDP-AADGal to undecaprenyl phosphate. HPLC and ESI-MS analysis indicated the production of the expected products (not shown).[45] The main conclusions of these experiments were that the enormous sensitivity of the fluorescent probe was a tool with intriguing potential for studying complex polysaccharide biosynthetic pathways, and that the ability to separate products via HPLC for further analysis with ESI-MS was an important feature. Plans for incorporating these tagged substrates into *in vivo* studies were in progress.[45]

1.4.4: Troutman Group – Synthesis of Another Fluorescently Tagged Isoprenoid

The next improvement to the tagged isoprenoids was the synthesis of a 2-nitrileaniline (2CN, E_{340nm}=2700 M^{-1} cm^{-1}) fluorophore that was incorporated into BP by first synthesizing the 2CN-GPP (Figure 9a) and incubating it with UppS. 2CN-GPP was tested with UppS from three different species and found to be compatible with all three, in contrast to 2AA-GPP which was only compatible with two of the three species. After incubating 2CN-GPP with UppS, the diphosphate was formed. Initial incubation with six equivalents of IPP resulted in isoprenoids with three to six isoprene additions (Figure 9c). Additional equivalents of IPP were incubated with the resultant isoprenoids and were shown to increase the length of the products while maintaining similar relative amounts of products (Figure 9d).[18] Then after incubation of the diphosphate isoprenoids with acid phosphatase, 2CN-bactoprenyl phosphate was formed (Figure 9e). Although the 2-nitrileaniline fluorophore on bactoprenyl phosphate (Figure 9b) was not as potent as the anthranilamide, it showed increased sensitivity to environmental changes and was sufficiently fluorescent to be readily observed via HPLC.[46]

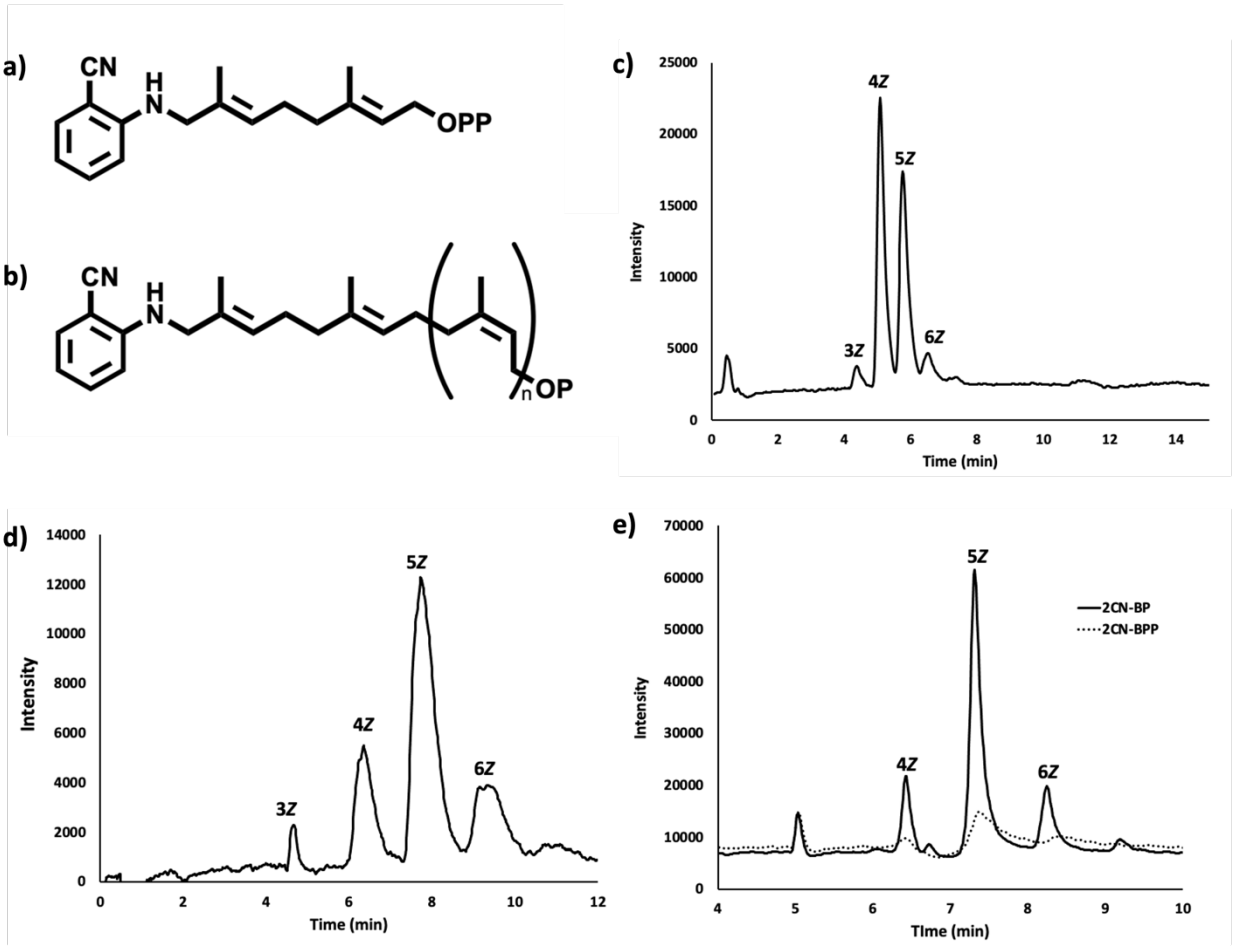

Figure 9: a) 2-Nitrileaniline geranyl diphosphate (2CN-GPP).
b) 2-Nitrileaniline bactoprenyl phosphate (2CN-BP).
c) Range of diphosphate isoprenoids from 3-6 Z additions formed after incubation with UppS and six equivalents of IPP. 2CNc4BPP was the major product.
d) Range of diphosphate isoprenoids still from 3-6 Z additions after four additional equivalents of IPP. 2CNc5BPP was the major product.
e) 2CN-BP is formed after incubation with acid phosphatase.
Note: This work was conducted by the author to replicate previous work done by the Troutman group.

1.4.5: Troutman Group – Variable Length Fluorescent Polyisoprenoids

One challenge with using synthetic isoprenoids in cellular assays is the issue with solubility. The longer chain isoprenoids are increasingly nonpolar and difficult to solubilize in the buffered conditions that allow for efficient enzyme activity. Some enzymes have been shown not to work with extremely short isoprenoids (zero or one *cis* additions) but it was thought that medium-length isoprenoids might have "the best of both worlds" with regards to solubility and

substrate effectiveness. Different surfactants were tested in different concentrations and different length isoprenoids were reliably obtained. This set of experiments allowed for simple smaller isoprenoids to be readily produced for use in glycan assembly assays.[47]

1.4.6: Troutman Group – Polyisoprenoids and Initiating Phosphoglycosyltransferases

The fluorescently tagged bactoprenyl phosphate (Fl-BP) analogues were shown to be successful with a glycosyltransferase and an initiating phosphoglycosyltransferase in previous publications. It was thought that since the glycan gets assembled on the phosphate end of the Fl-BP, as the oligomer grew and got further from the fluorophore, the catalytic site was less likely to be influenced by any tag that was added to the isoprenoid. If this is true, the initial glycosylation step would be the limiting factor for broader application of the Fl-BP probes.[18] Initiating phosphoglycosyltransferases (PGTs) transfer either a hexose-1-phosphate or a N-acetylhexosamine-1-phosphate to BP. This first step is the subject of a series of assays to ascertain whether the Troutman groups Fl-BP analogues may have broader scope of application with the many glycosyl transferases that exist.[18]

Four well-characterized PGTs: CPS2E from *Streptococcus pneumoniae*, WbaP from *Salmonella enterica*, WecA from *E. coli*, and WecP from *Aeromonas hydrophobia* were chosen to test with previously used 2CN-BP and a newly synthesized nitrobenzoxadizol (NBD) tagged BP (Figure 10b). The NBD fluorophore

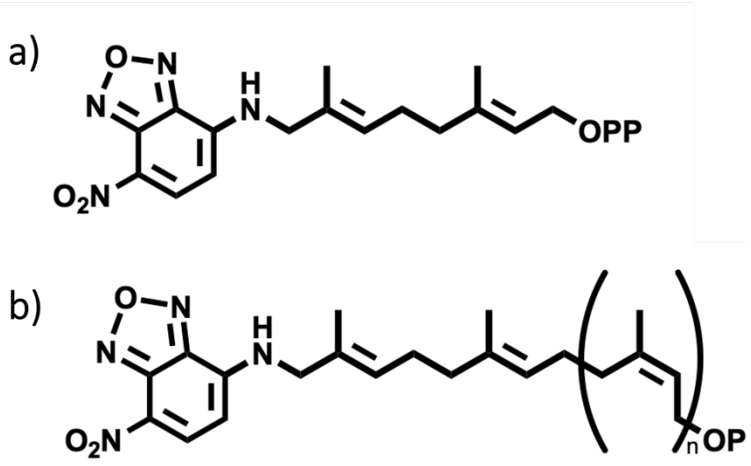

Figure 10: a) Nitrobenzoxadizol geranyl diphosphate (NBD-GPP) b) Nitrobenzoxadizol bactoprenyl phosphate (NBD-BP).

has a higher detection limit and reduced background noise compared to 2CN, and was chosen because the excitation and emission are in the visible range. After incubation of the PGTs with both Fl-BPs and their respective sugar substrates, it was found that both 2CN-BP and NBD-BP were acceptable substrates for all four PGTs.[18]

The conclusions reached by this sequence of publications set up a foundation for a host of bacterial glycan synthesis pathways to be investigated. The Fl-BPs were expected to work with most if not all downstream glycosyltransferases, and there was no longer a need for radiolabeled sugars. It is a slight oversimplification to say that all that is needed to understand a glycan assembly scheme is overexpressing proteins in *E. coli* and obtaining UDP-linked sugars, but that is the starting point. Other pathways that the Troutman group has since published or are currently working on include the production of colanic acid in *E. coli*,[20, 21] the modification of lipid A with 4-amino-arabinose in *E. coli,*[35] the modification of lipid A with glucosamine in *Bordetella pertussis* (unpublished data), and the production of Capsular Polysaccharide A in *B. fragilis*.[48]

1.5: Introduction to Lipid A and Lipid A Modifications – The Raetz Pathway

1.5.1: Lipid A Biosynthesis

Before describing my research into lipid A modification in *B. pertussis* using fluorescently tagged bactoprenyl phosphate, I will summarize the biosynthesis of lipid A and several known modifications of it in different species of bacteria. I will cover the discovery of lipid A, how it is synthesized in bacteria, the differences in lipid A between species and strains of bacteria, including acyl chain number, acyl chain length, and other modifications, the knowledge of which has been discovered largely due to the work of Christian Raetz at Duke University.[5, 49]

Lipopolysaccharide (LPS) is the major component of the outer membrane of most gram-negative bacteria. It is composed of three regions: a hydrophobic moiety named Lipid A, a core

oligosaccharide containing one to four ketodeoxyoctonic acid (Kdo) sugars, and a long chain of repeating oligosaccharide units called O-antigen. Lipid A is the portion of LPS that is recognized by the toll-like receptor 4 (TLR4) receptor in mammals and is integral to host-bacterial interactions. Normal interactions are necessary for host recognition of pathogens, but overactivation by the host immune system can quickly lead to a cascade of inflammation reactions, and Lipid A has been implicated in sepsis.[50] As the major component of the outer membrane, there can be as many as 10^6 lipid A molecules in a single *E. coli* cell.[49] The canonical lipid A from *E. coli* is a glucosamine-based phospholipid with six acyl chains and two phosphate groups, one on the C-1 position of glucosamine and the other on the C-4' position of the other glucosamine.[49] In *E. coli* and most other bacteria, the addition of Kdo sugars to lipid A is necessary for growth and survival, and in virulent strains there are other core sugars and modifications necessary for antibiotic resistance and virulence, but I will initially focus on the synthesis and export of Kdo_2-lipid A.

The Raetz group has described the functions of many if not all of the enzymes involved in the production, modification, and export of lipid A in *E. coli*, and homologs of these enzymes have been found in many other bacteria.[5] The biosynthesis of Kdo_2-lipid A is attributed to nine enzymes (encoded by nine distinct genes) named LpxA, LpxC, LpxD, LpxH, LpxB, LpxK, KdtA, LpxL, and LpxM (Figure 11, largely copied from source). The homologs of these

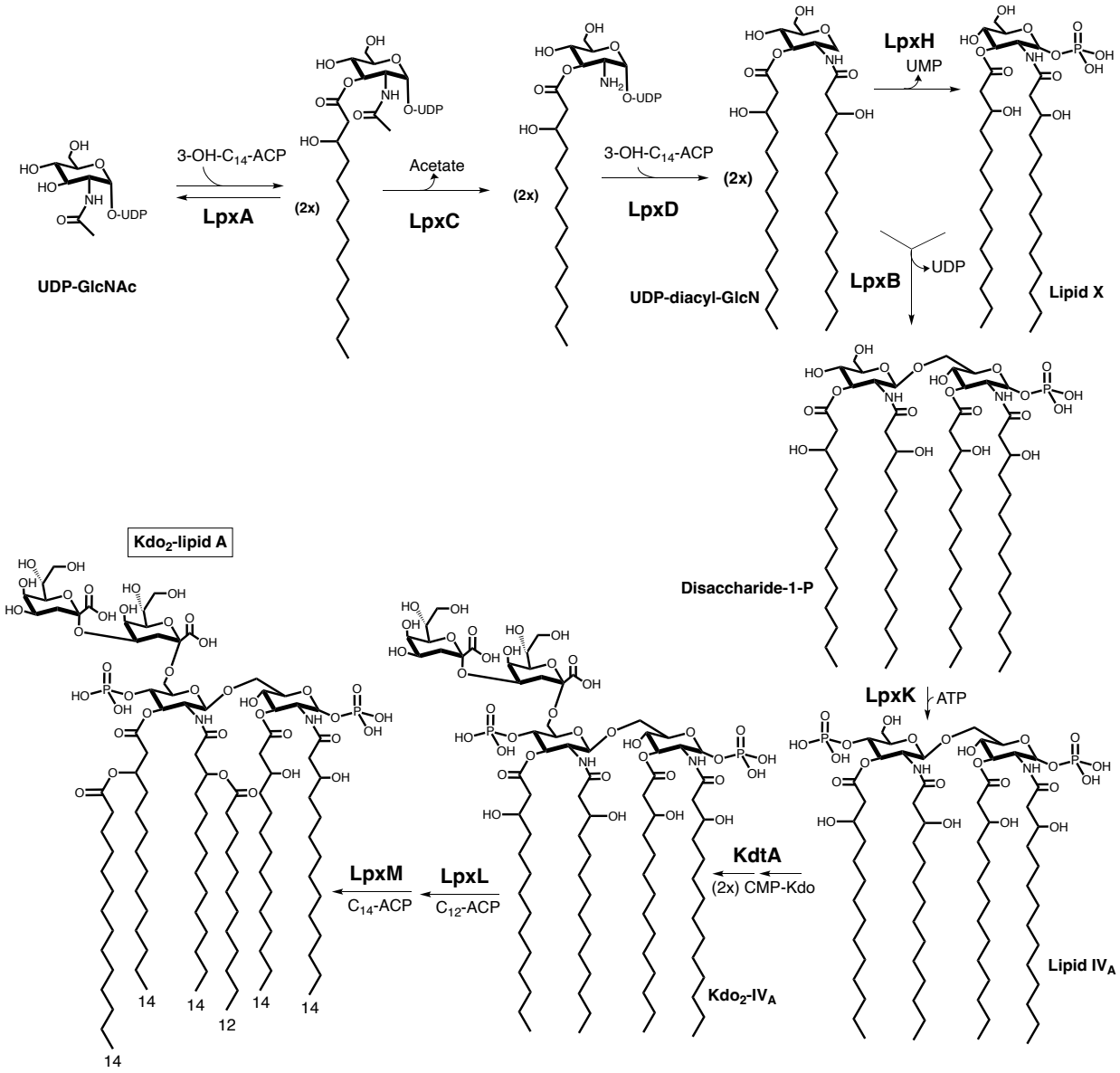

Figure 11: Biosynthesis of E. coli Kdo2-lipid A as described by the Raetz group showing the canonical form of lipid A with six acyl chains, two Kdo sugars, phosphorylated on both sides. ACP=acyl carrier protein. Adapted from Raetz et al.[5]

enzymes are responsible for the slight variations of lipid A found in other species, but the overall structure of lipid A is highly conserved.[49] The discovered variations have been found to play a role in susceptibility to antibiotics as well as immunostimulatory features of gram-negative bacteria.[4] These enzymes are mostly located near the cytoplasmic face or in the inner membrane.[49] There are other enzymes in addition to the nine listed above that further modify lipid A before it is completely assembled and exported to the outer membrane as part of LPS. These enzymes are located facing the periplasm or in the outer membrane.[5] The Raetz group along with others has also elucidated the function of a few of these additional modification enzymes in several different species of bacteria.[5, 49]

The functions of the nine lipid A synthesis enzymes are shown in Figure 11. LpxA is responsible for the incorporation of the initial fatty acyl chains onto UDP-GlcNAc. These chains are fourteen carbons long in *E. coli*, but different bacteria have homologs of LpxA that utilize acyl chains of different lengths.[49] For example, in *B. pertussis* lipid A one of the original fatty acyl chains is only ten carbons long (Figure 15).[51] LpxC deacetylates the GlcNAc moiety to form acetylated glucosamine (GlcN), and LpxD appends an additional fatty acyl chain. LpxH is a kinase that phosphorylates the LpxD product, then LpxB combines the phosphorylated LpxH product with a non-phosphorylated LpxD product to form the singly phosphorylated disaccharide (disaccharide-1-P). LpxK is another kinase that phosphorylates disaccharide-1-P to form lipid IV_A. KdtA appends two Kdo sugars to form Kdo_2-IV_A. Finally, LpxL adds a 12-C acyl chain, and LpxM adds a 14-C acyl chain which results in the final product Kdo_2-lipid A that has six acyl chains. *B. pertussis* lacks a LpxL or LpxM homolog and only adds one additional acyl chain resulting in Kdo_2-lipid A with only five acyl chains.[51]

The general biosynthesis of lipid A described above is considered the conserved pathway, and it occurs on or near the inner membrane.[49] The negative charge on the lipid A phosphate groups was demonstrated to be responsible for the susceptibility of bacteria to cationic antimicrobial peptides, and certain species have only one or zero phosphate groups instead of two.[52, 53] Other modifications beyond the number and length of the acyl chains vary from species to species, and these modifications mainly transpire as lipid A is assembled into LPS and exported to the outer membrane.[14] Most of the enzymes responsible for these further modifications exist in the periplasmic space or embedded in the outer membrane.[49] These modifications are thought to be induced by changes in growth conditions such as pH, concentration of divalent cations, or presence of antimicrobial peptides.[49] The major category of variable lipid A modifications is the addition of polar or charged groups to lipid A, specifically via attachment to the phosphate groups.[14] These modifications are unique from other glycan assembly pathways in bacteria such as capsular and exopolysaccharides which depend on an undecaprenyl diphosphoryl donor: lipid A modification pathways begin with an undecaprenyl monophosphoryl glycan donor.[54] The following sections will describe several known lipid A modifications and the similarities to lipid A modification in *B. pertussis*.

1.5.2: Lipid A modification in *E. coli* – a Parallel Pathway

One of the variable lipid A modifications is that 4-aminoarabinose has been found as a substituent of the phosphate group of lipid A in *E. coli*.[5, 55] The drive for this modification has been attributed to the presence of cationic antimicrobial peptides.[9] The addition of 4-aminoarabinose (Ara4N), which is positively charged at biological pH, increases the overall charge of the membrane and causes resistance to cationic antimicrobial peptides such as colistin

and polymyxin.[5] The functions of the enzymes responsible for this modification were first elucidated in 2002 and will be described below.[56]

As the first step in the lipid A modification pathway, N-formylated Ara4N is appended to the ubiquitous lipid anchor undecaprenyl phosphate by the enzyme ArnC at the cytoplasmic face of the inner membrane (Figure 12).[5] In the second step, ArnD deformylates the ArnC product to form undecaprenyl phosphate-Ara4N. ArnE and ArnF form an inner membrane flippase that flips undecaprenyl phosphate-Ara4N across the inner membrane to face the periplasm where ArnT transfers the Ara4N moiety to lipid A and core sugars. The modified lipid A is then assembled into the complete LPS structure and exported to the outer membrane. Disregarding the

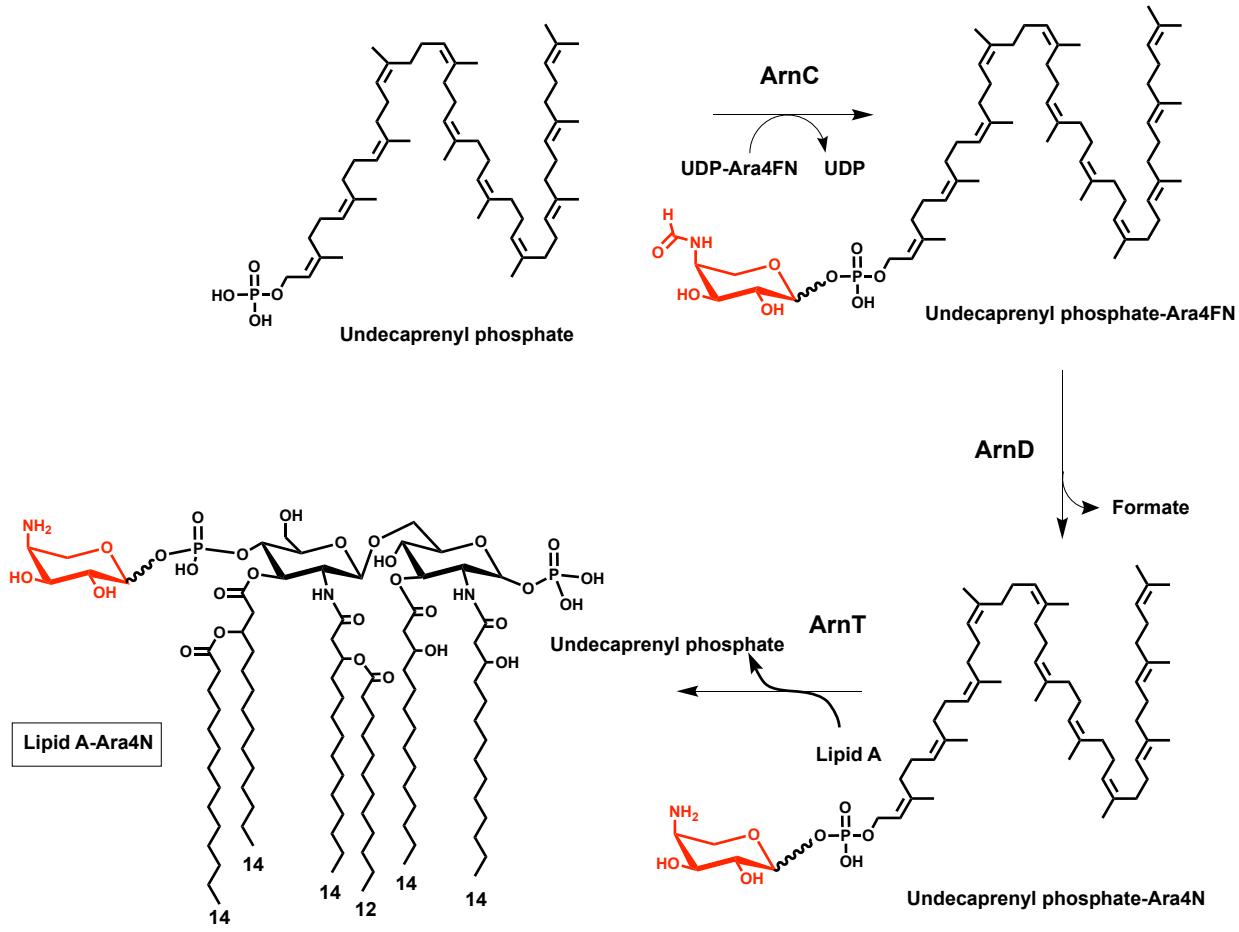

Figure 12: Lipid A 4-aminoarabinose modification in *E. coli*. 4-aminoarabinose and N-formylated 4-aminorarabinose are shown in red. Note: Kdo and other core sugars omitted for simplicity.

flippase for simplicity's sake, this process is considered a three-step pathway where an activated monosaccharide is attached to the lipid anchor, modified, then appended to lipid A. Parallels to this three-step pathway have been observed in other gram-negative bacteria as well.

1.5.3: Lipid A modification in *Francisella novicida* – a Parallel Pathway

Francisella tularensis causes tularemia, a highly contagious pulmonary disease in humans. *Francisella novicida* is the murine equivalent and it does not affect humans, which has made it a model organism. *F. novicida* has been shown to modify lipid A with a galactosamine

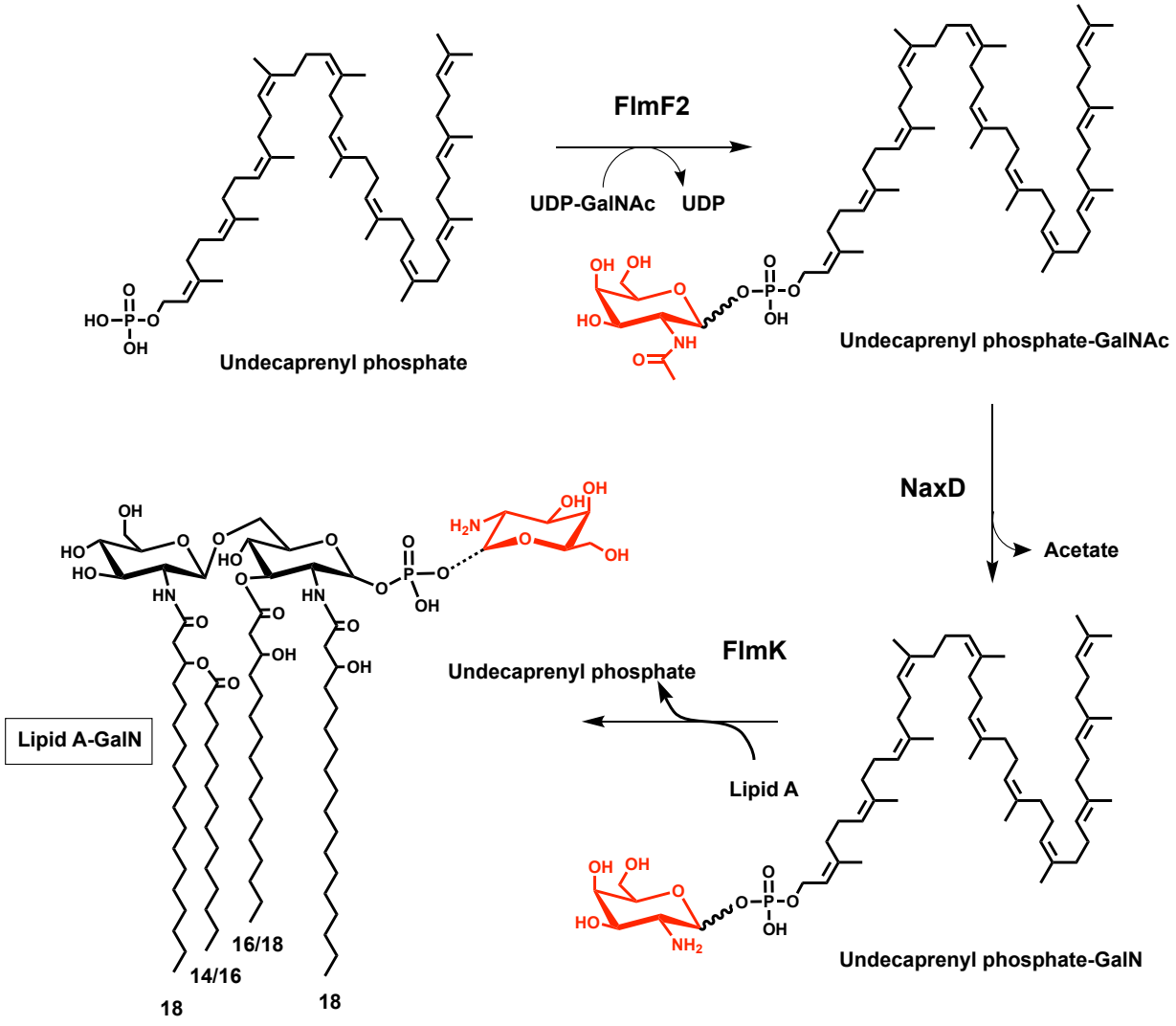

Figure 13: Lipid A is modified with galactosamine in *F. novicida*.[2] Galactosamine and N-Acetylgalactosamine residues are shown in red. Note: Kdo and other core sugars omitted for simplicity.

residue that increases the overall charge of the membrane and cause resistance to cationic antimicrobial peptides.[49] It has also been demonstrated that over 90% of *F. novicida* lipid A exists as "free" lipid A, without Kdo sugars, core sugars, or O-antigen; and that *F. novicida* has an enzyme that removes one of them. From 2009-2012, Raetz and coworkers discovered that the enzymes responsible for galactosamine modification of lipid A in *F. novicida* shared several similarities with the Arn pathway in *E. coli* (Figure 13).[2, 15, 17] The first enzyme in the three-step pathway is FlmF2 that appends N-acetylgalactosamine from UDP-GalNAc to undecaprenyl phosphate.[17] In the second step, NaxD deacetylates undecaprenyl phosphate-GalNAc to form undecaprenyl phosphate-galactosamine (GalN) which is then flipped across the inner membrane.[15] In the third step, FlmK appends GalN to lipid A.[2] When assembled into LPS, the positively charged GalN residue increase the overall charge on the cell surface and confer cationic antimicrobial peptide resistance.[17, 49, 53]

The three genes *flmF2*, *naxD*, and *flmK* were discovered to be co-located in the *Francisella* genome, and they all shared sequence homology with the genes of the *arn* locus. This discovery led to several different assays to confirm that the activity of the enzymes FlmF2, NaxD, and FlmK was homologous to the activity of the Arn enzymes. Raetz and coworkers used a combination of radiolabeled sugars in TLC assays and non-radiolabeled sugars in ESI-LC-MS assays to confirm the function of FlmF2 as a GalNAc transferase. The group grew both wild type *F. novicida* and *flmF2* knockout mutants, harvested the membrane fraction, and used the FlmF2- and non-FlmF2-containing membrane fractions in *in vitro* assays. The assays were conducted with appropriate buffers, undecaprenyl phosphate, and radiolabeled UDP-GalNAc as a substrate. The assays with wild type (FlmF2-containing) membrane fraction showed the appearance of a product assessed to be undecaprenyl phosphate-GalNAc, while the knockout mutant showed no

such product. Additional assays using non-radiolabeled substrates were conducted, and ESI-LC-MS peaks were interpreted as confirming the same data collected with TLC.[2]

The second enzyme in the GalN modification pathway is NaxD (N-acetylhexosamine deacetylase), which was actually the last of the three enzymes to have its function elucidated. Raetz and coworkers discovered by sequence analysis that *FTN_0544*, a gene co-located with *flmF2* and *flmK,* coded for a protein that was a member of the YdjC family of deacetylases that are encoded by multiple other gram-negative bacteria.[15] They generated an *F. novicida naxD* deletion mutant as well as a complemented strain, and noted that *naxD* was necessary for growth in a hostile environment. They further showed that the mutant displayed dose-dependent sensitivity to polymyxin, a cationic antimicrobial peptide. They proposed that its function would be to deacetylate undecaprenyl phosphate-GalNAc to yield undecaprenyl phosphate-GalN. They had previously shown that FlmF2 did not accept UDP-GalN as a substrate, and that the downstream enzyme FlmK did not accept undecaprenyl-GalNAc as a substrate, and so proposed that deacetylation was a necessary step in the pathway.[2, 17]

An interesting experiment described in the NaxD report was the measurement of the zeta electrokinetic potential of different strains of *Francisella* to show how the overall charge of the outer membrane was altered with galactosamine modification. Zeta electrokinetic potential gives an indirect reading of the bacterial surface charge. Galactosamine (positively charged at biological pH) modification would increase the surface charge that would normally be negative due to the phosphate groups on lipid A. Their experiment showed a twofold decrease in surface charge on the *naxD* mutant compared to the wild type. Taken together, the decrease in surface charge and increased susceptibility to polymyxin in the *naxD* mutant indicate that NaxD could be required for galactosamine modification of lipid A in *Francisella*.[15]

ESI-LC-MS was used to analyze the lipid fractions of both wild type and *naxD* mutants, and lipid A peaks were observed in both strains, though a difference in mass-to-charge ratio between them was discovered that corresponded to the addition of galactosamine to lipid A in the wild type. The researchers next examined precursor lipids to see where NaxD might act in the galactosamine modification pathway. They had already determined that GalN was added to lipid A from undecaprenyl phosphate-GalN,[17] and ESI-LC-MS of the precursor lipids indicated the absence of undecaprenyl phosphate-GalN in the mutant, while it was present in the wild type. Instead of undecaprenyl phosphate-GalN, the mutant had a large peak that corresponded to undecaprenyl phosphate-GalNAc. This suggested that NaxD was responsible for the deacetylation of undecaprenyl phosphate-GalNAc to undecaprenyl phosphate-GalN, and to determine this, NaxD was overproduced in *E. coli* (which does not contain a NaxD homolog, does not synthesize undecaprenyl phosphate-GalNAc or undecaprenyl phosphate-GalN, and does not modify lipid A with galactosamine). Assays with the NaxD membrane fraction (as well as an empty vector control) were conducted, and the presence of undecaprenyl phosphate-GalN was detected only in the membrane fractions containing NaxD. Taken all together, the results of these experiments indicated that NaxD was necessary and suggested that it was sufficient for deacetylation of undecaprenyl phosphate-GalNAc.[15]

The third enzyme in the galactosamine modification pathway is FlmK, which was actually the first of the three enzymes to be elucidated. The 4-aminoarabinose donor in *E. coli* is undecaprenyl phosphate-Ara4N, and by analogy, the Raetz group reasoned that the galactosamine donor in *Francisella* might be undecaprenyl phosphate-GalN. They purified and characterized undecaprenyl phosphate-GalN from *F. novicida* and used it in assays with FlmK (an ArnT homolog). The assays showed that FlmK utilized undecaprenyl phosphate-GalN as the

galactosamine donor to modify lipid A.[17] Interestingly, Raetz showed data that indicated that F. novicida will modify free lipid A (lipid A with no Kdo sugars, core sugars, or O-antigen) as well attached lipid A.[2, 53, 57] This information influenced one of my experiments with the FlmK analogue LgmB.

1.5.4: Lipid A Glucosamine Modification in *Bordetella pertussis*

1.5.4.1: Introduction to *Bordetella pertussis*

The genus *Bordetella* includes *pertussis, parapertussis,* and *bronchiseptica. B. pertussis* and *B. parapertussis* are human pathogens of the respiratory tract that cause pertussis or whooping cough. *B. bronchiseptica* is the canine version that causes kennel cough in dogs. *B. pertussis* can cause coughing fits that last for weeks, pneumonia, seizures, brain damage, and even death. Newborns are the most susceptible to *B. pertussis*, and about fifteen infants die each year in the US alone. Globally over 150,000 people die each year due to *B. pertussis*. Although vaccines for *B. pertussis* do exist, the deaths have been attributed to decreasing vaccine efficacy over time as well as the ability of the bacteria to resist antibiotics such as colistin and polymyxin. The antibiotic resistance of *B. pertussis* has been attributed to its ability to modify its outer membrane with glucosamine in a way that is thought to parallel the Ara4N modification of Lipid A in *E. coli* and the GalN modification of lipid A in *F. novicida*.[10]

Lipid A glucosamine modification (Lgm) in *B. pertussis* has been attributed to three genes of the Lgm locus: *lgmA*, *lgmC*, and *lgmB*.[10, 58] (There is an additional gene, *lgmD*, that codes for a putative undecaprenyl phosphate-donor flippase, but it has been shown to not be necessary for lipid A glucosamine modification in *B. pertussis*, and it will not be discussed in detail.[4, 37]) The Lgm pathway was proposed to be a three-step lipid A modification pathway based on sequence homology of the genes to the Arn and Flm pathways and the fact that the final

modification of lipid A is a amino-monosaccharide like Ara4N and GalN. In 2008, the Fernandez group in British Columbia and Martine Caroff in France published the first description of glucosamine modification in *Bordetella*.[4] They found glucosamine substituents on both phosphates of lipid A in both *B. pertussis* and *B. bronchiseptica*, neither of which had been previously known. They also suggested the genetic basis for this modification based on the discovery of a gene encoding a *Bordetella* ArnT ortholog. When this gene was interrupted, the *Bordetella* strain of interest no longer exhibited glucosamine modification.[4]

Previous work by this same group had revealed additional unexplained peaks in *Bordetella* lipid A samples.[59] In order to determine the identity of those peaks, they performed MALDI-MS under mild conditions and were able to assign the substitutions on lipid A as two phosphates and two hexosamine moieties. They noted that a phosphate could not be liberated

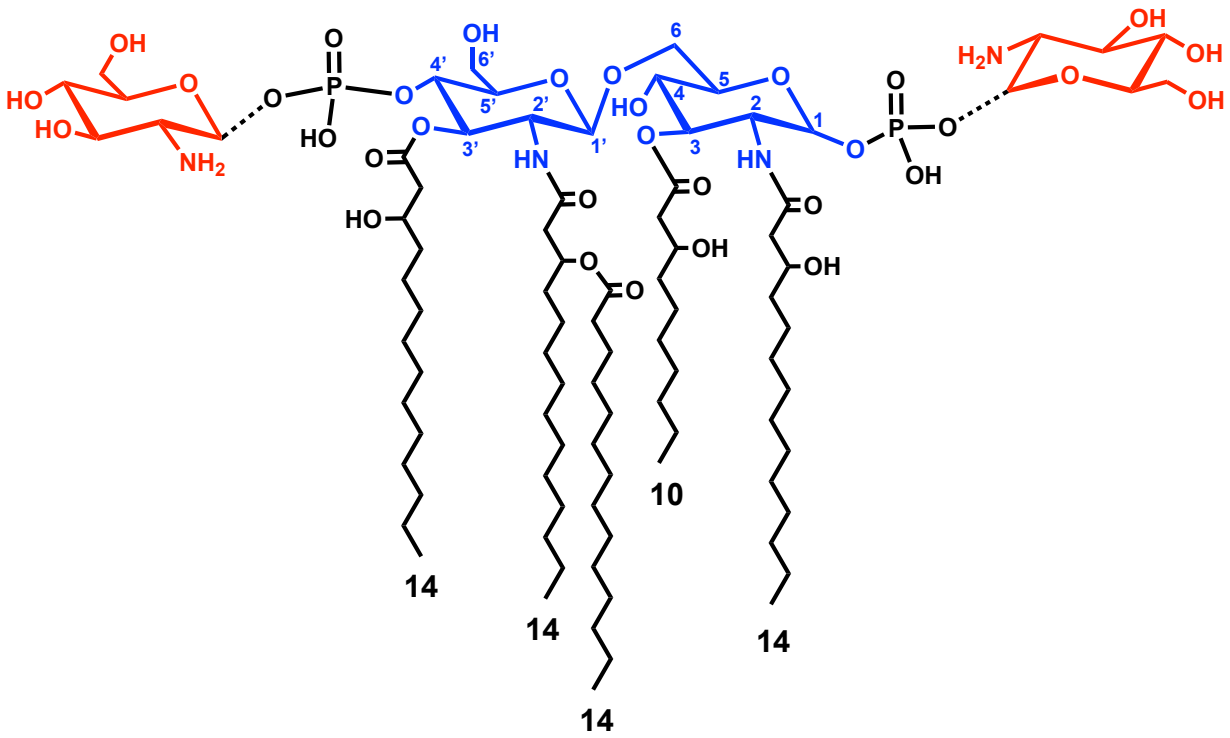

Figure 14: Proposed structure for *B. pertussis* lipid A with two glucosamine substitutions shown in red. Lipid A backbone glucosamine residues shown in blue. Note both 1 and 4' phosphates are present and substituted, but dotted lines show possible incompleteness of substitutions.[4]

without also losing a corresponding hexosamine (indicating that the hexosamine was on the outside of the molecule) and performed a ninhydrin color test to detect the free amino groups in the lipid A structure to confirm the presence of a hexosamine.[4] The hexosamine substituents were further characterized by purifying the lipid A, hydrolyzing with strong acid, removing fatty acids by extraction, peracetylating the reduced sugars, and then testing by GC. They observed a single peak corresponding to peracetylated glucosaminol and concluded that the substituents were glucosamines identical to the glucosamine that constitutes the lipid A backbone (Figure 14).[4]

The *arnT* homolog locus was identified by genetic analysis, and an ortholog of *arnC* was also discovered nearby. Both were assigned putative functions as glycosyl transferases and appeared to be present in all the strains of *Bordetella* that evidenced glucosamine modification.[4] Although homologs of the enzymes thought to be responsible for this modification had been found in other species, no other glucosamine modification of lipid A had previously been documented.[4]

After the structure of *B. pertussis* modified lipid A was elucidated, researchers in the Caroff and Fernandez groups set out to determine exactly which genes were responsible for this modification. A cluster of four genes was discovered and named the Lgm locus. This locus included *lgmA*, *lgmB*, *lgmC*, and *lgmD*. Bioinformatic analysis showed that LgmA and LgmB were homologs of ArnC and ArnT, respectively.[58] LgmC showed similarity to the YdjC superfamily of deacetylases, and LgmD had indications of being transmembrane bound, but its function was not predicted.[58] The proposed pathway for lipid A modification in *B. pertussis* is shown in Figure 15. The start codons of these genes overlapped, another indication that they all functioned together. The Fernandez group generated deletion mutants of each of these individual

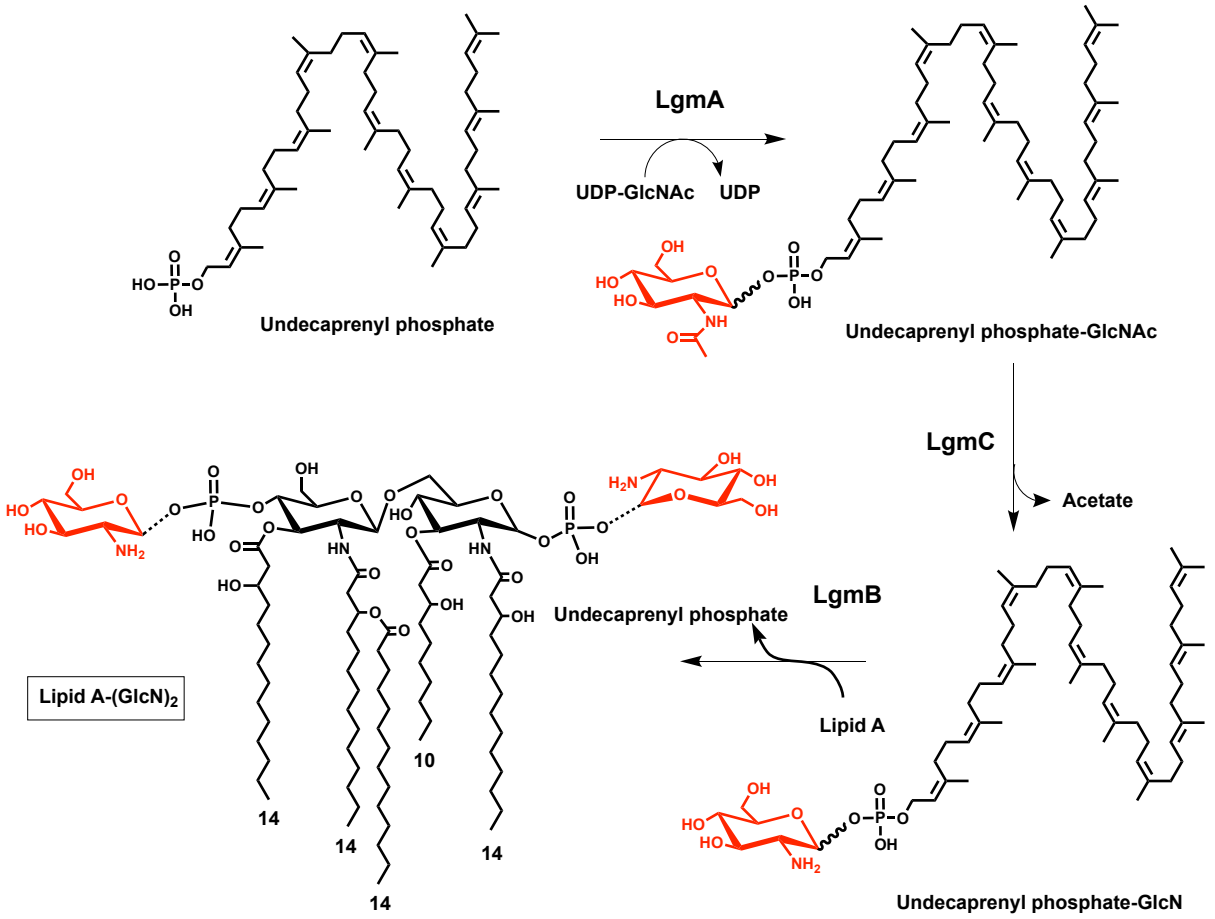

Figure 15: Lipid A is modified with one or two glucosamine residues in *B. pertussis*. Glucosamine and N-acetylglucosamine residues are shown in red. Dotted lines indicate possible incomplete substitution of glucosamine, leading to structures with one or no substituting glucosamine residues.[4] Note: Kdo and other core sugars omitted for simplicity.

genes as well as a deletion of the entire Lgm locus. After growing up the strains, the lipid A from each of the mutants as well as the wild type was analyzed by MALDI-MS. The only strains that still had lipid A modified with glucosamine were the wild type and the *lgmD* deletion, while the other strains did not. This suggested that only *lgmA*, *lgmB*, and *lgmC* were required for glucosamine modification of lipid A in *B. pertussis*.[58] This specific study and one previous focused on the interaction between lipid A and TLR4 and showed that lipid A modification with glucosamine increased TLR4 activation.[58, 60]

Other studies have shown that modification with a positively charged sugar also confers resistance to cationic antimicrobial peptides, and another group in 2014 showed that lipid A

modification specifically with glucosamine was required to protect *B. bronchiseptica* from cationic antimicrobial peptides and was required for transmission of the bacteria.[5, 9] In the same year, the Fernandez lab went on to show that *B. pertussis* modification by glucosamine decreased sensitivity to cationic antimicrobial peptides and increased resistance to outer membrane perturbation.[10] The gap in knowledge remained that the precise function of each of the Lgm enzymes was not yet demonstrated. Armed with the identification of the functions of lipid A-modifying enzymes in other gram-negative bacteria as well as the putative functions of the Lgm enzymes in *B. pertussis,* I proposed to study the function of these enzymes and design assays using the Troutman group's fluorescently tagged undecaprenyl phosphate (bactoprenyl phosphate) analogues in order to assign the precise biochemical roles of LgmA, LgmB, and LgmC.

CHAPTER 2: MATERIALS AND METHODS

2.1: Materials

Lgm plasmids were obtained from the Fernandez lab.[58] Overproduction of proteins was conducted in commercially obtained *E. coli* C41 or RP cells. Lipid A and Kdo_2-lipid A were purchased from Sigma-Aldrich. 2CN-GPP was synthesized by Amanda Reid and Hailey Houde of the Troutman group. IPP was synthesized with the assistance of Amanda Reid. Cell culture media, buffers, and salts purchased from Thermo Scientific. UDP-linked sugars were purchased from Carbosynth. Ni-NTA resin was purchased from Thermo Scientific.

2.2: Protein Overproduction

Lgm plasmids were transformed into *E. coli* expression line C41 or RP cells according to commonly used protocol. Cells were diluted from 5 mL starter cultures and grown in 1L lysogeny broth (LB) at 37 °C for approximately 4 hours until optical density at 600 nm (OD_{600}) between 0.4-0.8 was reached. Cultures were induced with 1 mM IPTG, temperature was decreased to 16 °C overnight. Cells were harvested by centrifugation, resuspended in 50 mM Tris, 200 mM NaCl, 20 mM imidazole buffer and lysed by sonification. Unlysed cells and debris were removed by centrifugation at 5,000 x g, and the supernatant centrifuged under vacuum at 195,000 x g to collect the membrane fraction. Membrane fractions were homogenized in 50 mM Tris, 200 mM NaCl dialysis buffer and concentration was approximated as total protein content by Bradford assay.[61]

2.3: Isoprenoid Synthesis Reactions

IPP was synthesized according to previously published methods.[62] UppS and phosphatase reactions to obtain 2CN-BPs were conducted according to previously published methods.[18] 2CN-BPs were isolated using methods described in the same report.[18] 2CN-c6-BP was chosen as the

undecaprenyl phosphate analogue for all of the reactions in this report due to its solubility and the broad acceptability and strong signal of its fluorescent tag shown by previous group members.[18]

2.4: LgmA Reactions

LgmA activity was tested in 200 mM Bicine buffer at a pH of 8.0, 2.5 mM $MgCl_2$, 15 mM cholate, 2.5 μM 2CNc6BP, 250 μM UDP-GlcNAc, and 0.25 mg/mL LgmA cell envelope fraction (cef). These conditions had been previously optimized by a former laboratory member.[63] Reaction volumes were 40 μL unless otherwise noted. Reaction contents were added to PCR tubes, vortexed briefly, spun down, and incubated at 37 °C for one hour, unless otherwise noted. Reaction mixtures were usually evaluated by HPLC or LC-MS immediately, but occasionally were stored at -20 °C overnight then thawed, sonicated, and evaluated the following day. No difference was observed between samples that were evaluated immediately or frozen.

2.3: Isolating BP-GlcNAc

To isolate BP-GlcNAc, the product of the LgmA reaction, 200 μL of a more conentrated LgmA reaction mixture (20 μM 2CNc6BP and 2 mM UDP-GlcNAc) was added to an Eppendorf tube, vortexted briefly, spun down, and incubated at 37 °C for one hour. An equal amount of *n*-butanol was added to the reaction mixture, vortexed, and spun down. The top organic layer (containing *n*-butanol and the lipid product) was extracted, and the process was repeated once more for a total of ~400 μL of suspended lipid product. The *n*-butanol was dried under a gentle stream of air and the isolated lipids were resuspended in 100 μL of DMSO yielding an approximate 2CNc6BP-GlcNAc concentration of 80 μM (assuming complete turnover) for use in downstream reactions. Samples were either stored at -20 °C for up to several months or used immediately.

2.4: LgmC Reactions

LgmC activity was tested in 100 mM HEPES buffer at a pH of 7.0, 100 mM NaCl, 1 mM $MnCl_2$, ~5 μM 2CNc6BP-GlcNAc (isolated from LgmA reaction, assuming full turnover of 2CNc6BP to 2CNc6BP-GlcNAc), and 0.5 mg/mL LgmC cef. Reaction volumes were 40 μL unless otherwise noted. Reaction contents were added to PCR tubes, vortexed briefly, spun down, and incubated at 37 °C for one hour, unless otherwise noted. Reaction mixtures were usually evaluated by HPLC or LC-MS immediately, but occasionally were stored at -20 °C overnight then thawed, sonicated, and evaluated the following day. No difference was observed between samples that were evaluated immediately or frozen.

2.5: LgmB Reactions

LgmB activity was tested by first incubating LgmC reaction in the conditions described above, then adding 0.5 mg/mL LgmB cef and 0.1 mg/mL free *E. coli* lipid A or *E. coli* Kdo_2-lipid A, vortexing, spinning down, and incubating at 37 °C for an additional hour. An aliquot of the LgmC reaction was saved each time for evaluation by HPLC or LC-MS to ensure 2CNc6BP-GlcN formation.

2.6: HPLC

All HPLC analysis was conducted using an Agilent 1100 HPLC system. Evaluation of Lgm intermediates was performed on a reverse-phase Eclipse XDB-C18, 4.6 x 50 mm, 3.5 μm column with an isocratic mobile phase of 50% n-propanol and 50% 100 mM ammonium bicarbonate (pH 8.0). A fluorescence detector with excitation at 340 nm and emission at 390 nm was attached to monitor the fluorescent isoprenoids.

2.7: LC-MS

All LC-MS analysis was conducted using a Waters Xbridge Peptide BEH C18, 3.5 μm, 4.6 x 50 mm column on an Agilent 1260 Single Quadrupole. To monitor fluorescent isoprenoids, the LC-MS setup included a TEE connector splitting the column eluant between the MSD and FLD (Ex: 340 nm, Em: 390 nm). To monitor non-fluorescent isoprenoids and lipid A, only the MSD was utilized. The solvent conditions utilized to monitor 2CN-BPP and 2CN-BP formation was a gradient from 15% n-propanol to 65% n-propanol, with 0.1% ammonium hydroxide as the other solvent. Conditions used to monitor Lgm intermediates and products was a gradient from 25% n-propanol to 65% n-propanol with 0.1% ammonium hydroxide as the other solvent.

The following mass-to-charge (m/z) ratios were used for analysis on MS:

Table 1: Selected m/z ratios of 2CN-BPPs and 2CN-BPs. The presented values are the $[M-H]^-$ species.

Species	**m/z**
2CN-c3-BPP	633.3
2CN-c4-BPP	701.4
2CN-c5-BPP	769.4
2CN-c6-BPP	837.5
2CN-c3-BP	553.4
2CN-c4-BP	621.4
2CN-c5-BP	689.4
2CN-c6-BP	757.5

Table 2: Selected m/z ratios of 2CN-c6-BP Lgm intermediates. The presented values are the $[M-H]^-$ species.

Species	m/z
2CN-c6-BP	757.5
2CN-c6-BP-GlcNAc	960.6
2CN-c6-BP-GlcN	918.6
2CN-c6-BP-Ara4N	888.6

Table 3: Selected m/z ratios for lipid A and lipid A-linked sugars. Lipid A and Kdo2-lipid A were from *E. coli*. The presented values are the $[M-2H]^{2-}$ species.

Species	m/z
Lipid A	898.1
Lipid A-GlcN	978.6
Lipid A-$(GlcN)_2$	1059.2
Kdo_2-Lipid A	1118.2
Kdo_2-Lipid A-GlcN	1198.7
Kdo_2-Lipid A-$(GlcN)_2$	1279.2

CHAPTER 3: RESULTS AND DISCUSSION

3.1: LgmA Transfers GlcNAc to Undecaprenyl Phosphate

LgmA is a putative glycosyl transferase of approximately 38 kiloDaltons (kD).[58] According to TMHMM 2.0, an online server that analyzes protein sequences, LgmA has two probable transmembrane domains.[64] B. Scarbrough's unpublished data suggested that LgmA is an integral membrane-bound protein and remained with the cell envelope fraction (cef) after expression.[65] LgmA was overproduced in an RP strain of *E. coli* (it did not express well in C41 strain) and the resulting cef was used for the following experiments (Figure 16 and Figure 17).

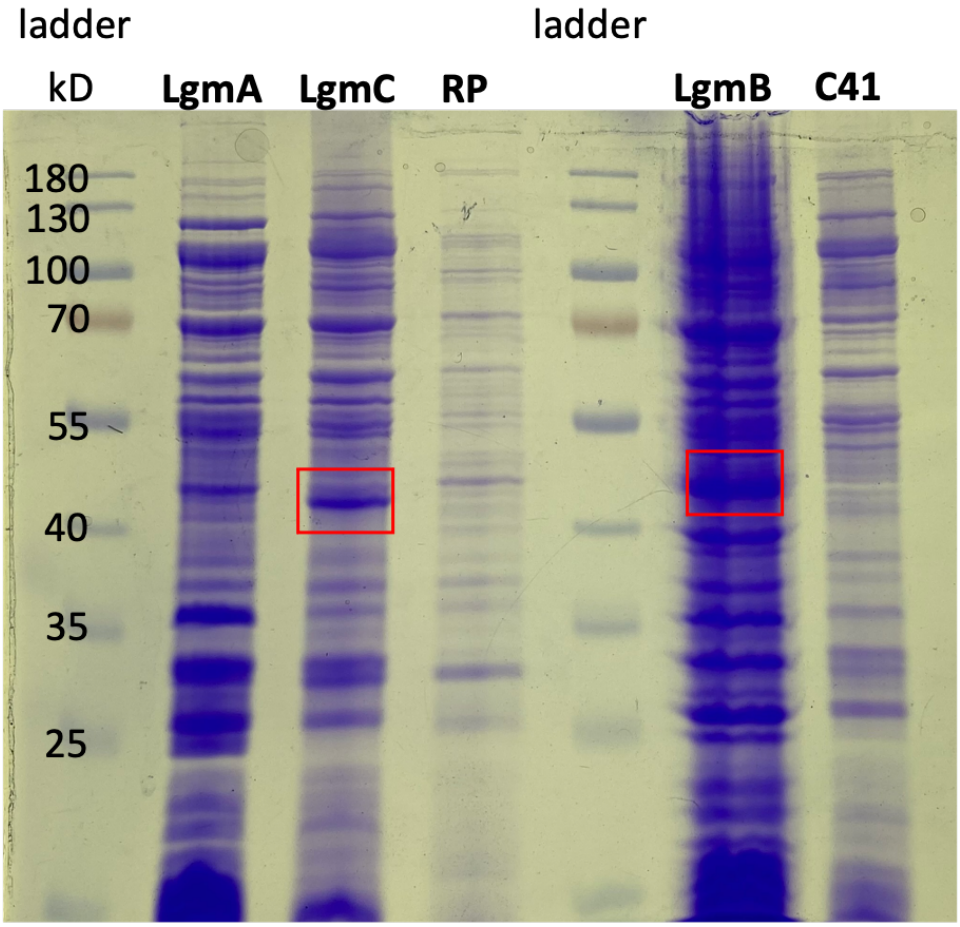

Figure 16: Coomassie stain of Lgm proteins expressed in RP (LgmA and LgmC) and C41 (LgmB). LgmA (~39 kD), LgmB (~61 kD), LgmC (~43 kD). kD, kiloDalton.

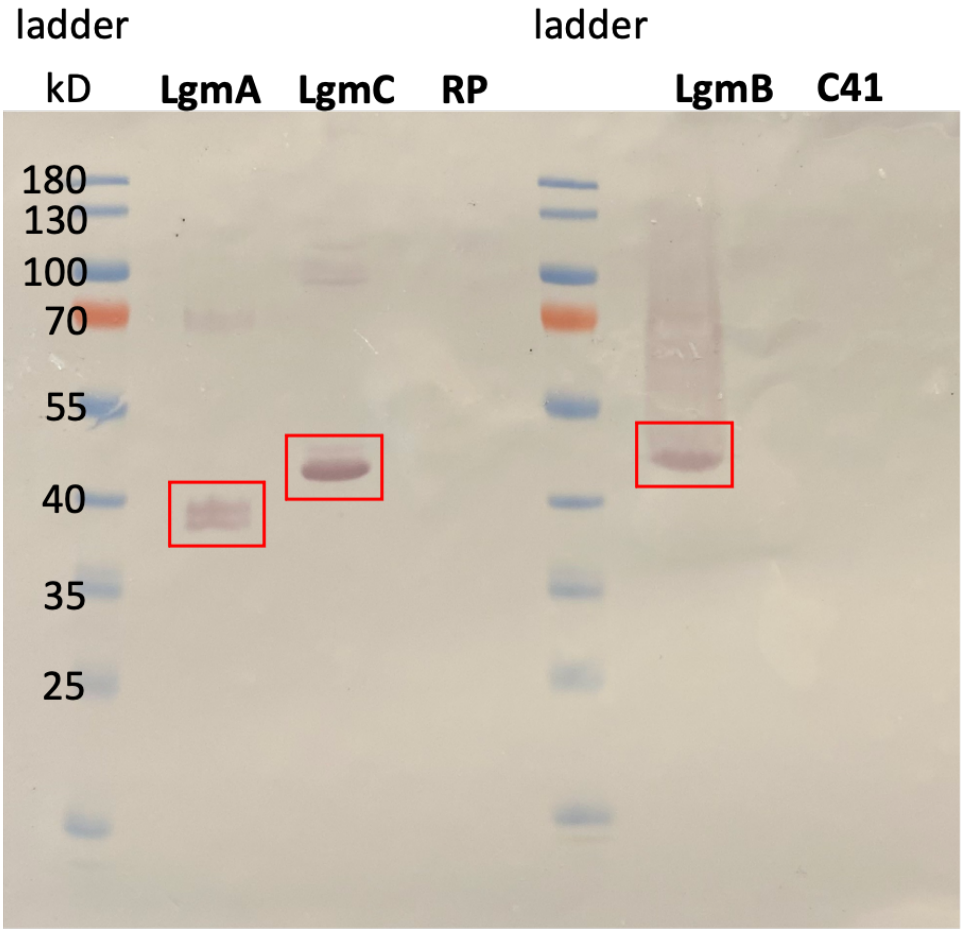

Figure 17: Western blot of Lgm proteins expressed in RP (LgmA and LgmC) and C41 (LgmB). LgmA (~39 kD), LgmB (~61 kD), LgmC (~43 kD). kD, kiloDalton.

The LgmA cef and an empty vector RP control in a buffered solution were both separately incubated with 2CN-c6-BP and UDP-GlcNAc according to E. Okwei's previously optimized conditions.[63] The reactions were evaluated using HPLC and a new fluorescent peak around 4.5 minutes was observed in both the LgmA reaction and the control that had a shift in retention time corresponding to an addition of a hydrophilic moiety to 2CN-c6-BP (Figure 18). The LgmA reaction had an additional fluorescent peak with an even shorter retention time of 4 minutes. The peak that appeared in both the reaction and the control had been previously observed in the Troutman group's experiments with other glycosyltransferases and appeared only in reactions with cef regardless of whether a sugar substrate was added.[35]

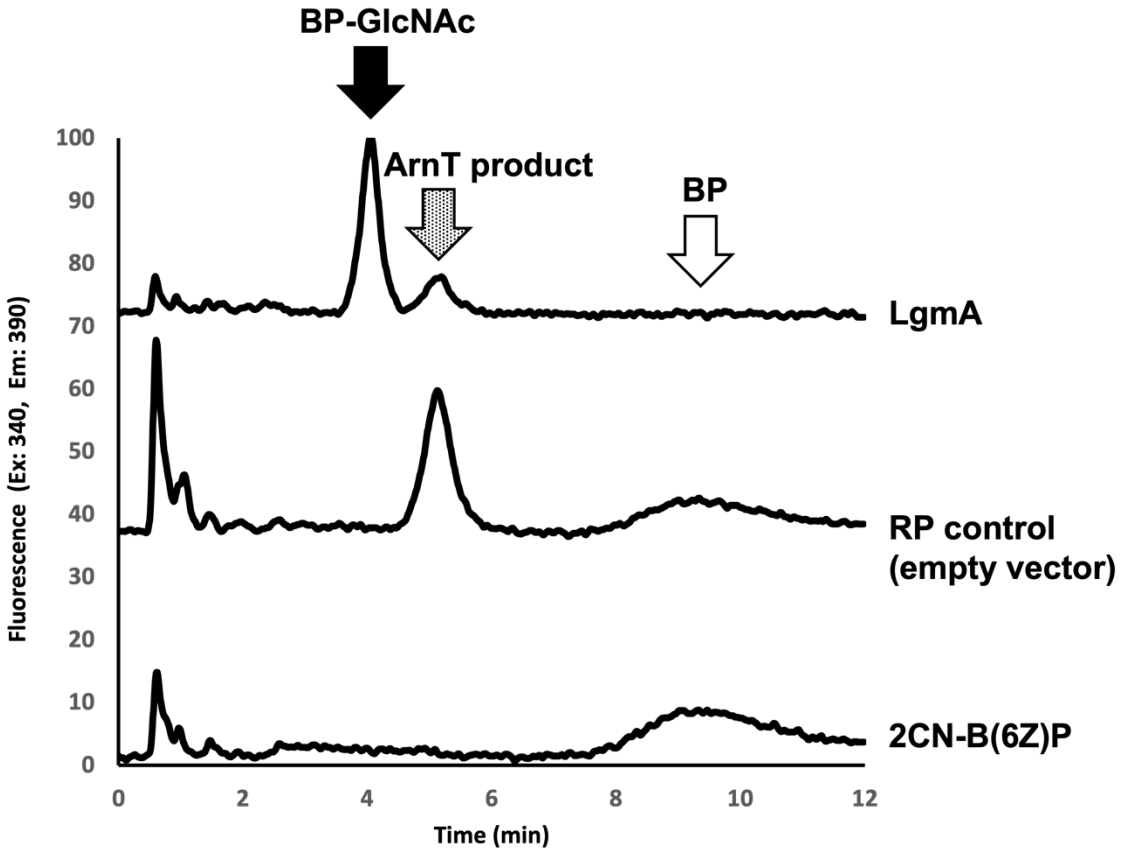

Figure 18: HPLC analysis indicates that LgmA catalyzes the formation of 2CN-c6-BP-GlcNAc from 2CN-c6-BP and UDP-GlcNAc.

The peak was later determined by Scarbrough to be 2CN-c6-BP-Ara4N, a product of the reverse reaction of the endogenous *E. coli* enzyme ArnT (mentioned above in the 4-aminoarabinose modification of lipid A in *E. coli* discussion).[35] Endogenous lipid A-Ara4N was determined to be the 4-aminoarabinose donor for this reaction, which explained how a sugar substrate was available in a membrane fraction.[35] The 2CN-c6-BP peak was also observed to disappear completely in the LgmA reaction but there was still some remaining (indicating incomplete turnover) in the control. ESI-LC-MS was used to confirm that the mass of the new peak corresponded to 2CN-c6-BP-GlcNAc (not shown). The resulting glycan, presumed to be 2CN-c6-BP-GlcNAc, was extracted with *n*-butanol and stored at -20 °C for use with LgmC.

To investigate the specificity of LgmA, the cef was incubated with 2CN-c6-BP and four UDP-linked sugars: UDP-Glc, UDP-GlcNAc, UDP-Gal, and UDP-GalNAc. The only reaction that showed a shift in retention time was the one that contained UDP-GlcNAc, indicating that LgmA exhibits a strong preference for UDP-GlcNAc (Figure 19). All of the data collected for LgmA was consistent with previous data collected by Okwei.[63]

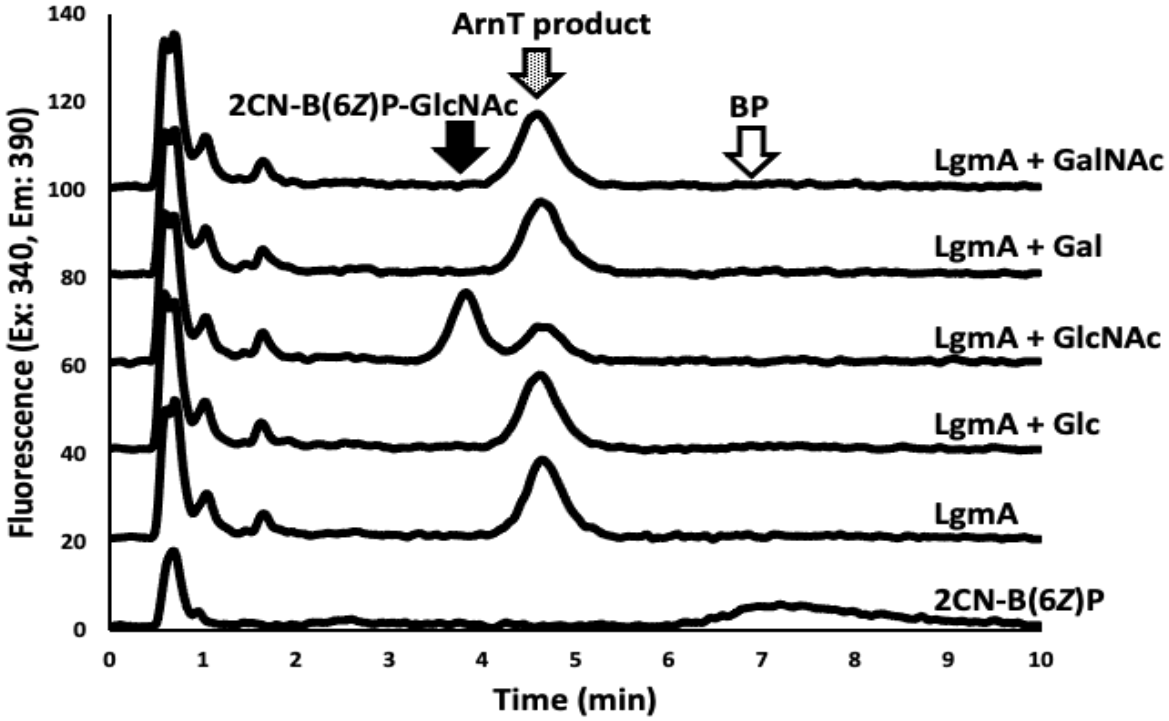

Figure 19: HPLC analysis indicates that LgmA exhibits strong specificity for UDP-GlcNAc over other UDP-linked sugars.

3.2: LgmC Deacetylates Undecaprenyl Phosphate-GlcNAc

LgmC is a putative deacetylase of approximately 36 kD.[58] According to TMHMM 2.0, it has zero transmembrane domains, but is thought to be associated with the membrane.[63, 64] It is a homologue of N-acetylhexosamine deacetylase (NaxD) in *F. novicida* and is expected to deacetylate undecaprenyl-GlcNAc to form undecaprenyl phosphate-GlcN.[15, 58] I first produced LgmC without attempting to solubilize it from the membrane, and it appeared to remain in the cef. Since the Troutman laboratory had not observed LgmC deacetylase activity yet, I first attempted to optimize the reaction conditions by testing the LgmC cef activity with two different salts (NaCl and KCl) and three different divalent cations (Mg^{2+}, Mn^{2+}, and Zn^{2+}).[63, 66]

The cef was incubated in buffered solutions with 2CN-c6-BP-GlcNAc that had been isolated (along with 2CN-c6-BP-Ara4N) from the LgmA reaction, and the divalent cations

caused very different activity (Figure 20). No difference in activity was observed between NaCl and KCl (not shown). Using HPLC for analysis, with Mg^{2+} a double peak was observed. One peak was consistent with the LgmA product, and the other eluted approximately thirty seconds later. With Mn^{2+} only one peak was observed, and it eluted approximately thirty seconds after the LgmA product. With Zn^{2+} only one peak was observed, and it was consistent with the LgmA product. The slight difference in the elution times of the peaks is in accordance with the expected difference between 2CN-c6-BP-GlcNAc and 2CN-c6-BP-GlcN. These results suggest that Mg^{2+} allows partial turnover, Mn^{2+} allows full turnover, and Zn^{2+} allows no turnover.

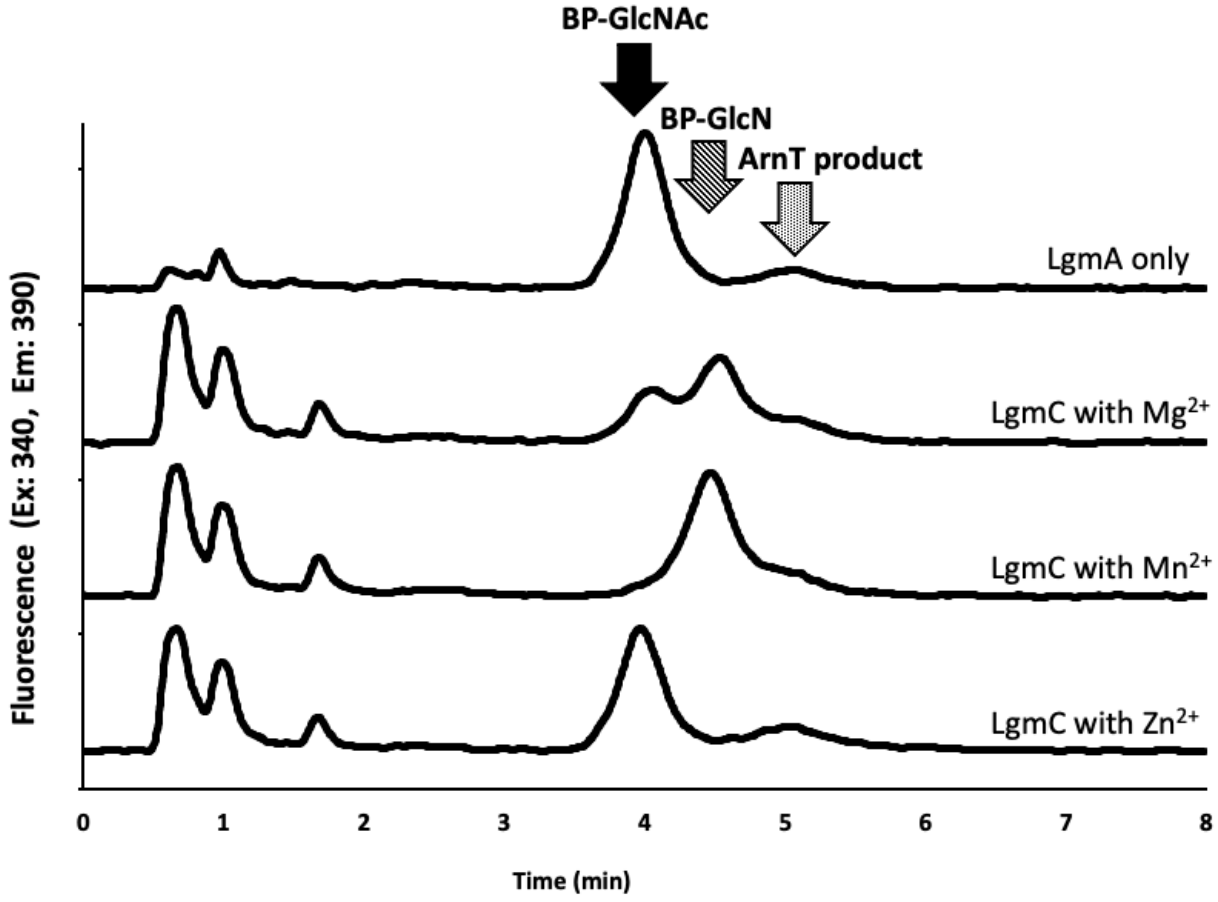

Figure 20: HPLC analysis indicates that LgmC catalyzes the deacetylation of 2CN-c6-BP-GlcNAc to 2CN-c6-BP-GlcN in the presence of manganese (II). Note the presence of 2CN-c6-BP-Ara4N in each of these reactions caused by carryover from the LgmA reaction.

To further confirm these results, the same reactions were analyzed using ESI-LC-MS (Figure 21). In the LgmC reaction with Mg^{2+} there were two separate mass peaks, one that corresponded to the mass of 2CN-c6-BP-GlcNAc and one that corresponded to 2CN-c6-BP-GlcN. In the Mn^{2+} reaction there was only one mass peak that corresponded to the mass of 2CN-c6-BP-GlcN. In the Zn^{2+} reaction there was again only one mass peak, and it corresponded to the mass of 2CN-c6-BP-GlcNAc. These results support the conclusion that LgmC deacetylates undecaprenyl phosphate-GlcNAc to form undecaprenyl phosphate-GlcN in conditions of 100 mM sodium chloride and 1 mM Mn^{2+}, and these conditions were selected for future reactions.

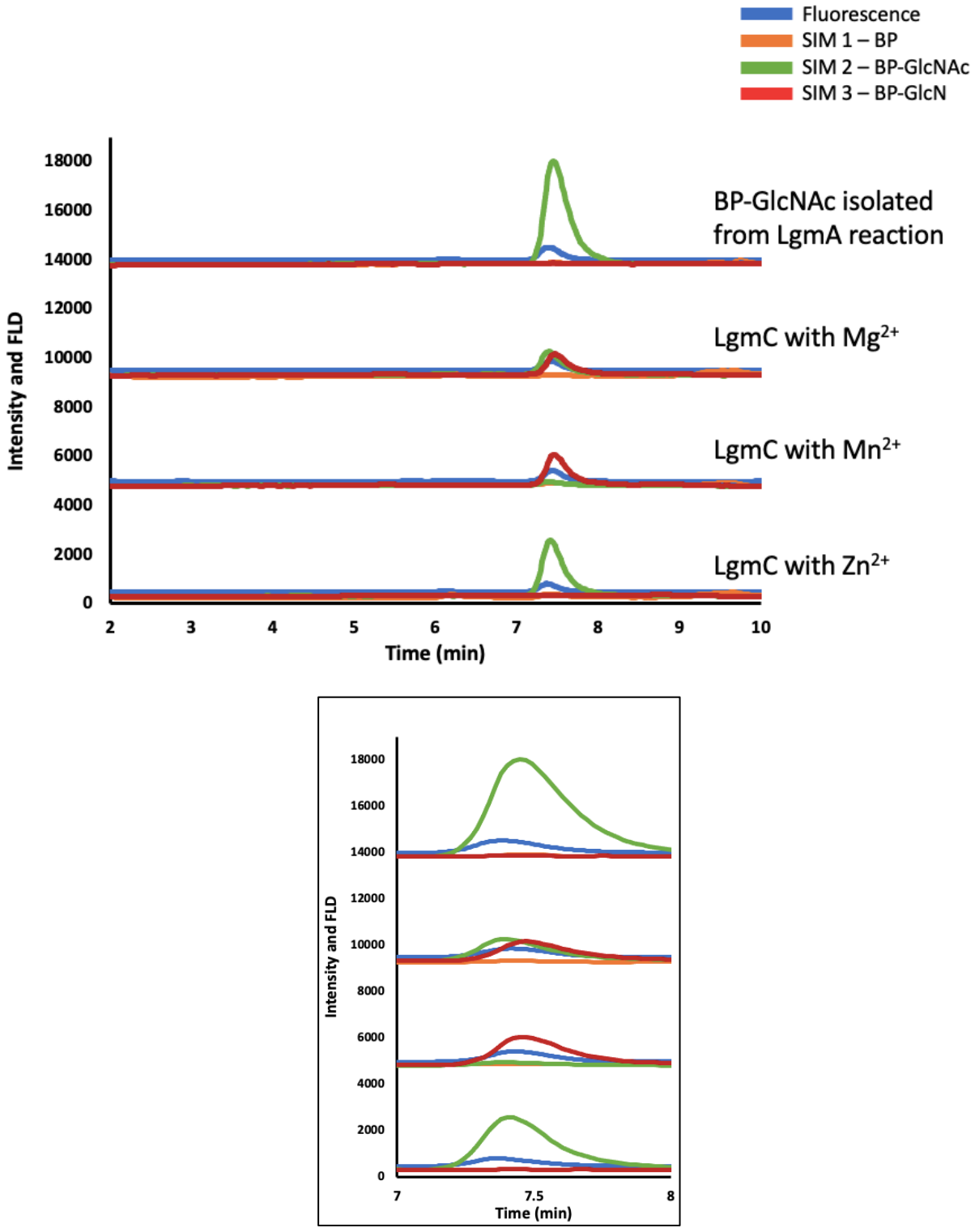

Figure 21: ESI-LC-MS data corroborates peak identity of LgmA and LgmC products. Inset shows the peaks of interest.

Because LgmC has zero transmembrane domains, it was thought that it could be easily purified. After re-expressing LgmC and purifying it by histidine association on a nickel-nitrilotriacetic acid (Ni-NTA) column with the surfactant 0.1% Triton X-100 in the elution buffer, evidence of LgmC as well as trace amounts of other proteins were observed in the soluble fraction by Coomassie stain (Figure 22).

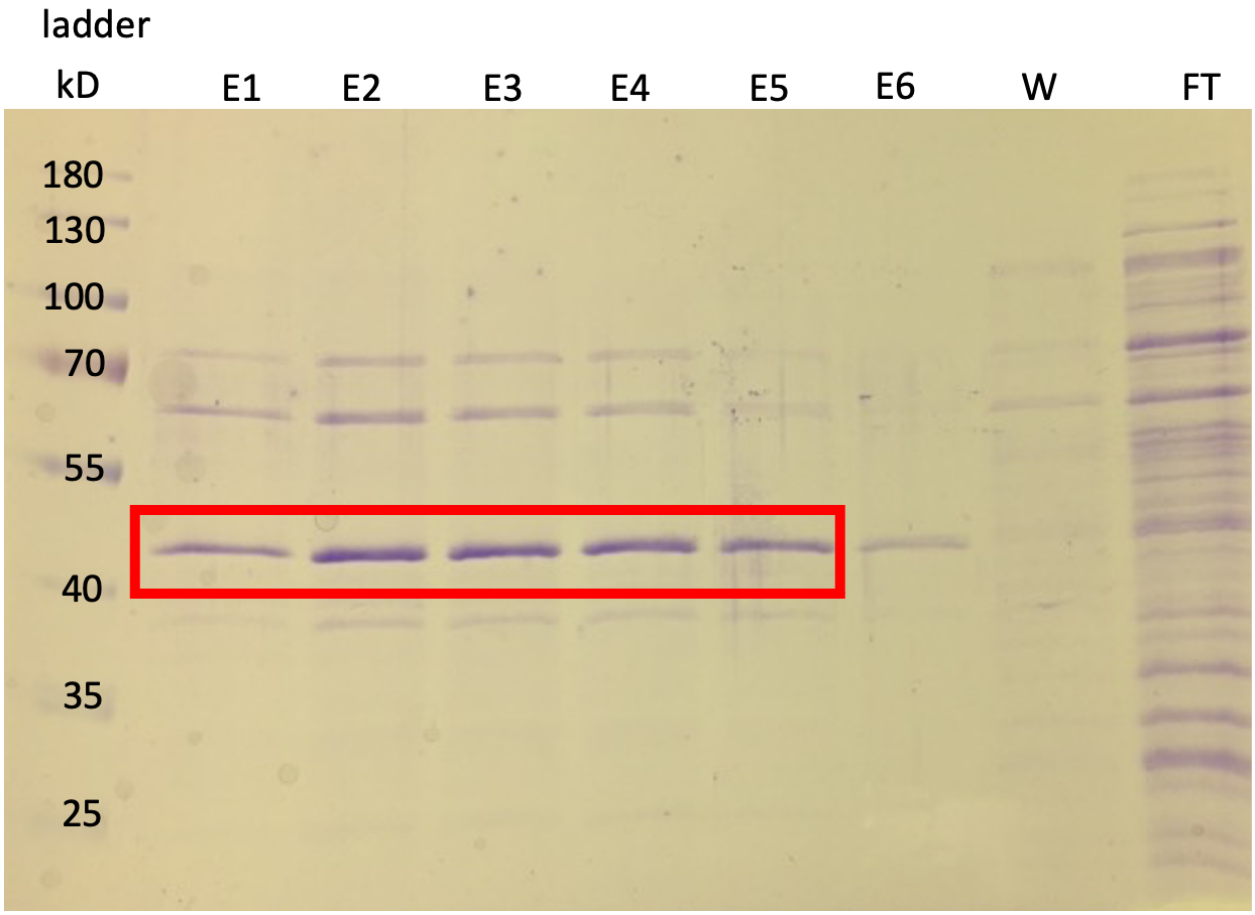

Figure 22: LgmC (~43 kD) purified protein Coomassie stain after Ni-NTA column purification. Evidence of LgmC in elutions 1-5, E1-5 were chosen for dialysis. kD, kiloDalton; E, elution; W, wash; FT, flow through.

The soluble fraction was dialyzed and stored at -20 °C. The purified LgmC was tested using the previously optimized conditions, but no turnover from 2CN-c6-BP-GlcNAc to 2CN-c6-BP-GlcN was observed (Figure 23). Because 0.1% Triton X-100 was used in the elution buffer, and because Triton X-100 micelles are too large to be dialyzed away, there was an estimated 0.01% Triton X-100 present in the reaction (1 µL of purified LgmC in 10 µL reaction).

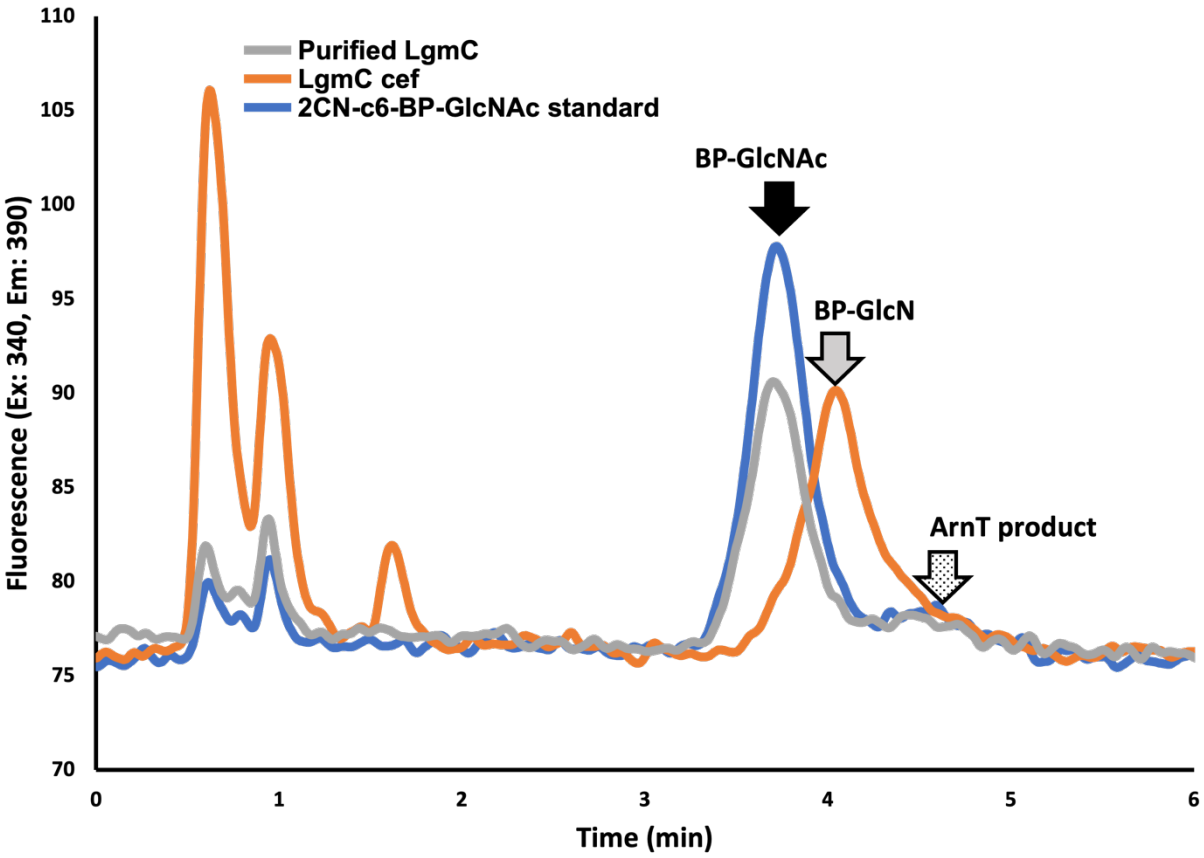

Figure 23: Purified LgmC does not deacetylate 2CN-c6-BP-GlcNAc.

In order to investigate the cause of the purified LgmC not performing, I tested LgmC cef (previously shown to be fully functional) with the same concentration of Triton X-100 that the purified LgmC contained (0.01%) and found the LgmC cef did not function with the addition of surfactant (Figure 24). I also tested different concentrations of Triton-X as well as varying concentrations of cholate and n-dodecyl-β-D-maltoside (DDM) from approximately the critical micelle concentration (CMC) to several times the CMC. I also tested lower concentrations of purified LgmC in the reaction in order to have lower concentrations of Triton X-100. The purified LgmC never showed evidence of turnover (not shown), but LgmC cef was able to turn over 2CN-BP-GlcNAc to 2CN-BP-GlcN with the lower end of the tested concentrations of the three different detergents (Figure 25).

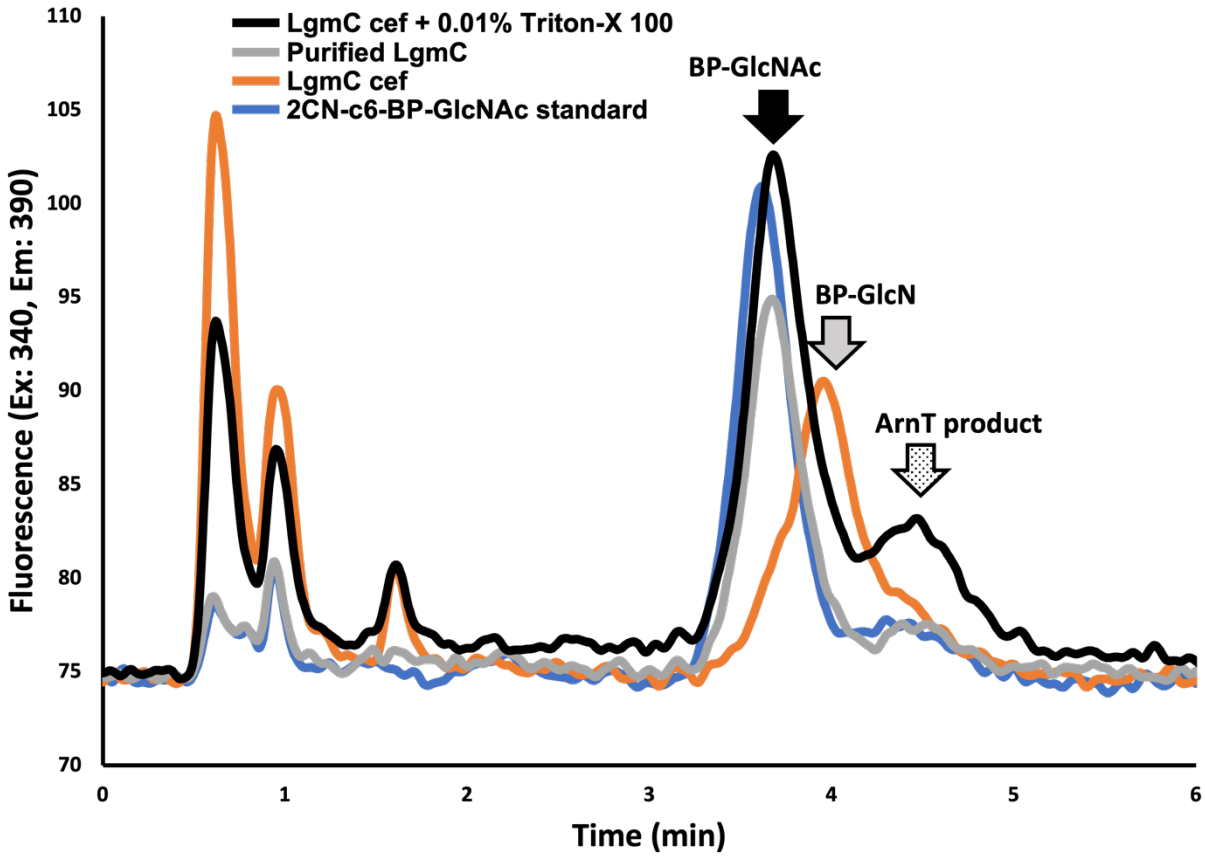

Figure 24: LgmC cef with the addition of 0.01% Triton X-100 did not deactylate 2CN-c6-BP-GlcNAc.

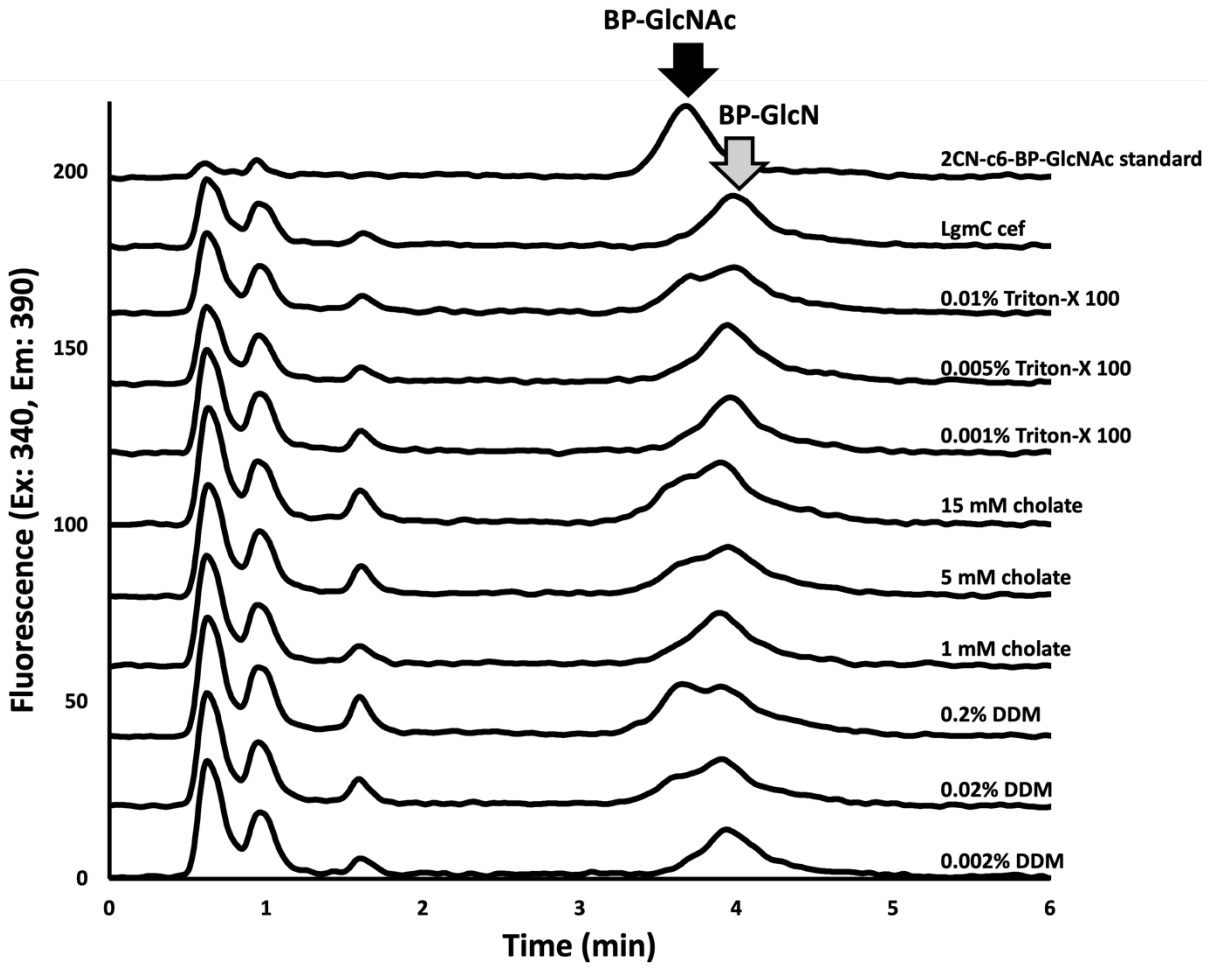

Figure 25: LgmC cef activity tested with different concentrations of Triton X-100, cholate, and DDM. Note: It is difficult to distinguish the ArnT product peak in this figure, but a distinction can be seen between 2CN-c6-BP-GlcNAc and 2CN-c6-BP-GlcN.

To get a clearer picture of the changes being made by the Lgm enzymes without the Arn product visible in all the fluorescent assays, I decided to re-express the Lgm proteins in a cell line that had an ArnC deletion. (ArnC catalyzes the formation of undecaprenyl phosphate-Ara4FN in *E. coli*, and cells with no ArnC should have no undecaprenyl phosphate-Ara4FN, no undecaprenyl phosphate-Ara4N, and no lipid A-Ara4N.) I chose the strain C41 *ΔarnC*:frt:pRP in the hope that the RP plasmid would allow LgmA to be overeexpressed, and because the C41 strain worked well for both LgmC and LgmB expression. The *Lgm* plasmids were transformed

into the C41 *ΔarnC*:frt:pRP expression cells, proteins were overproduced as described previously, and cef was harvested for LgmA, LgmB, and LgmC (Figure 26 and Figure 27).

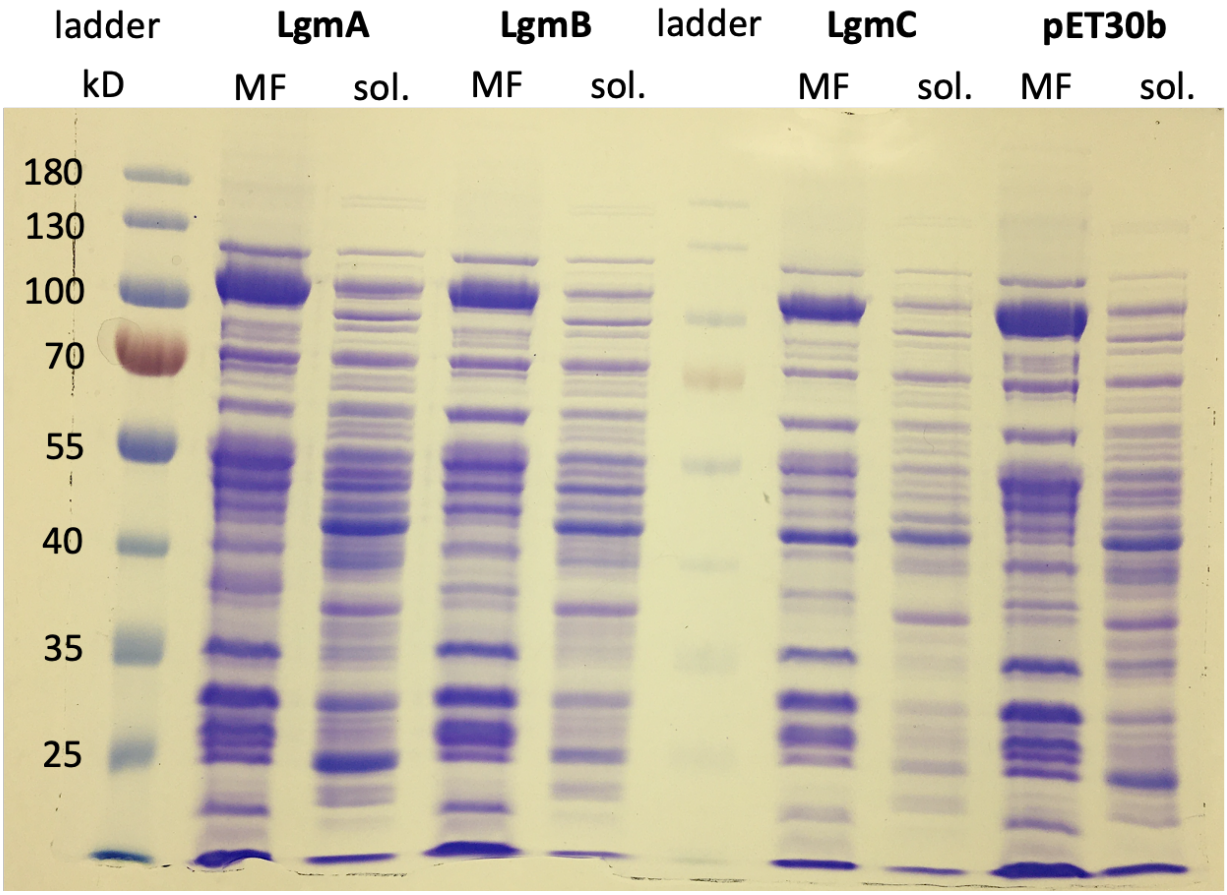

Figure 26: Coomassie stain of Lgm proteins expressed in C41 Δ*arnC*:frt:pRP. LgmA (~39 kD), LgmB (~61 kD), LgmC (~43 kD). kD, kiloDalton; MF, membrane fraction; sol., soluble fraction.

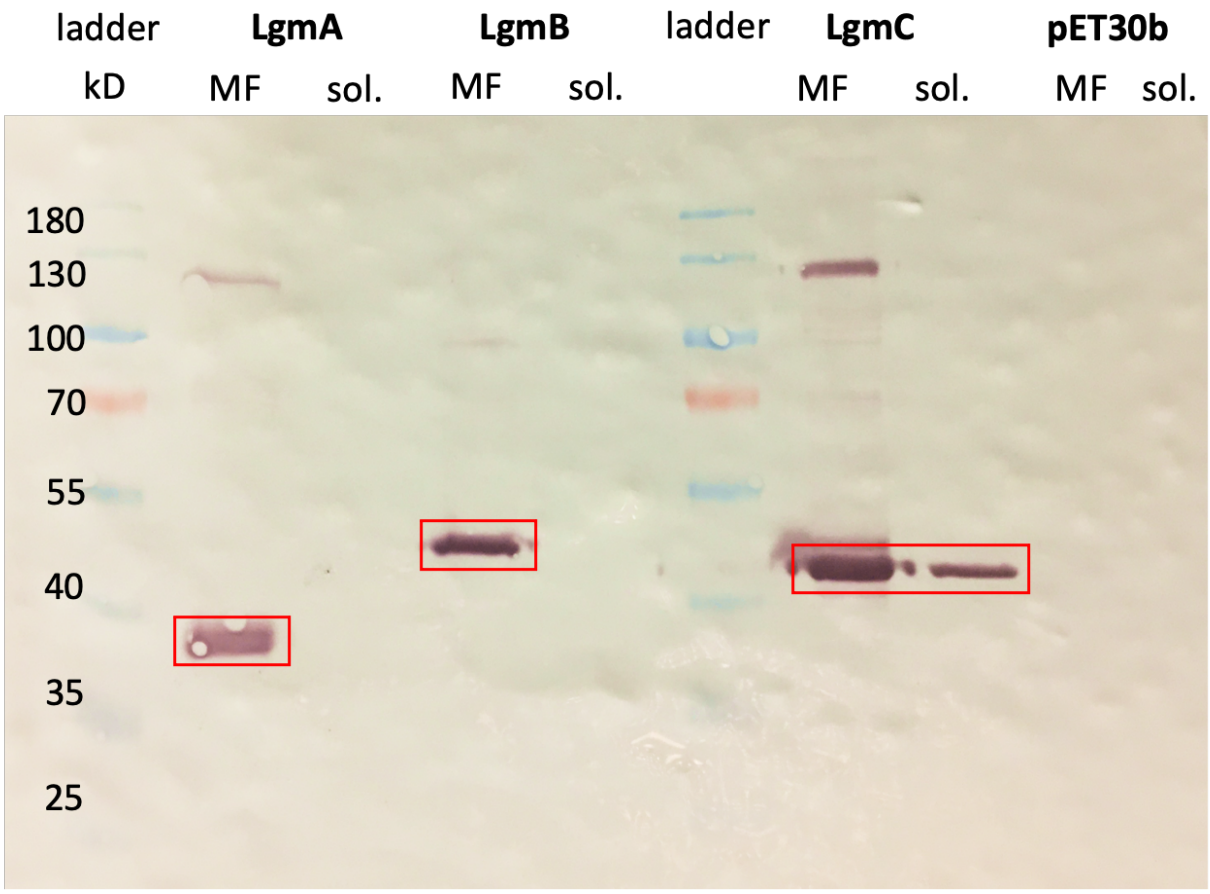

Figure 27: Western blot of Lgm proteins expressed in C41 Δ*arnC*:frt:pRP indicates proteins were present in the cef. LgmC evident in both the membrane fraction and the soluble fraction. LgmA (~39 kD), LgmB (~61 kD), LgmC (~43 kD). kD, kiloDalton; MF, membrane fraction; sol., soluble fraction.

After obtaining the cef containing LgmA, it was incubated with 2CN-c6-BP and UDP-GlcNAc in the same conditions used previously. It was thought that the ArnC deletion would eliminate the undesired 2CN-c6-BP-Ara4N present in previous reactions. HPLC analysis of the reaction contents indicated that indeed 2CN-c6-BP-Ara4N had been eliminated, but full turnover to 2CN-c6-BP-GlcNAc was not observed as desired (not shown). Instead, a relatively small amount of 2CN-c6-BP-GlcNAc was observed, a relatively large amount of 2CN-c6-BP remained, and a new product was observed. The new product was assessed to be 2CN-c6-BPP-GlcNAc, the known product of the endogenous *E. coli* enzyme WecA, based on its earlier elution time than the LgmA product 2CN-c6-BP-GlcNAc and previous characterization by our group.[18]

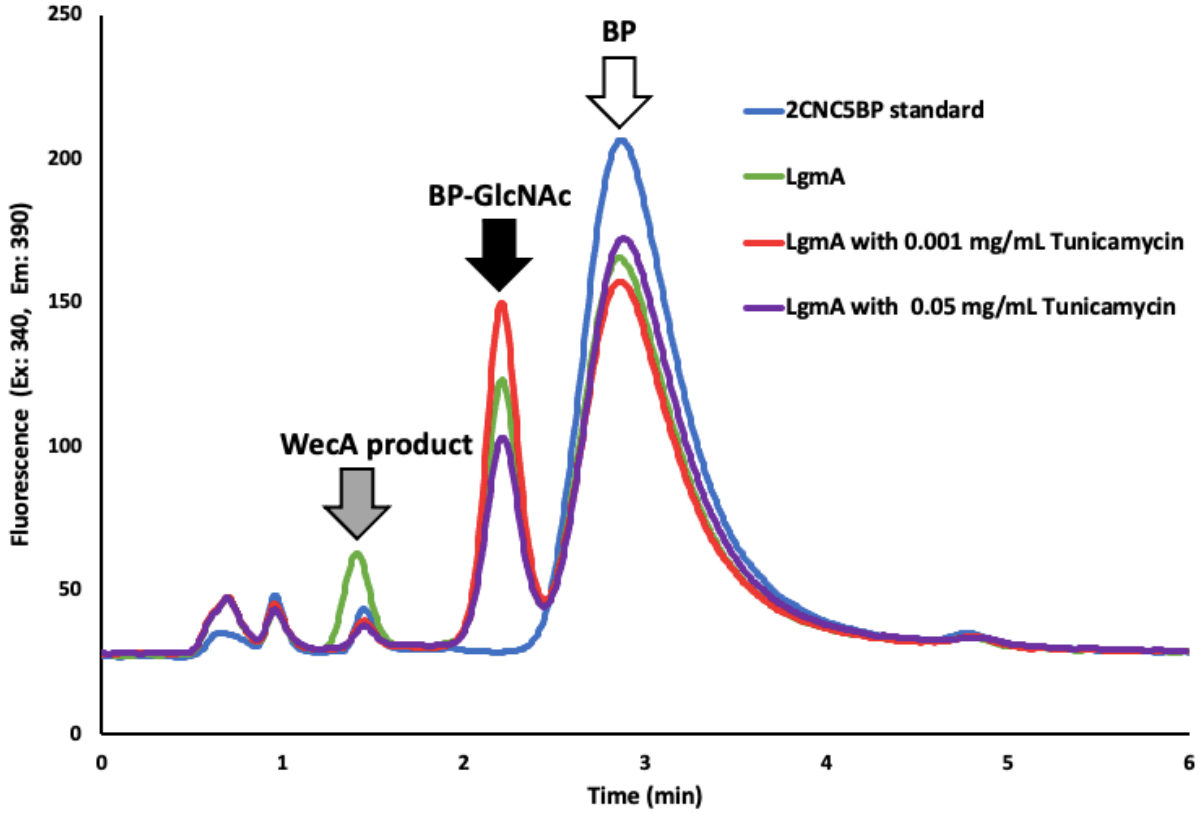

Figure 28: Optimal concentration of 0.001 mg/mL tunicamycin inhibits WecA completely and leads to greatest turnover by LgmA. Complete turnover by LgmA not observed at any concentration of tunicamycin.

Okwei's previous work involving LgmA had shown that tunicamycin was an effective inhibitor of WecA but not an effective inhibitor of LgmA.[63] I incubated LgmA cef with its substrates and various concentrations of tunicamycin in attempt to eliminate the WecA product and increase LgmA production as much as possible. A concentration of 0.001 mg/mL tunicamycin was observed to effectively inhibit WecA without inhibiting LgmA, but LgmA turnover was only marginally improved (Figure 28). There was still a relatively large amount of 2CN-c6-BP remaining, even after lengthening the incubation time to two hours and overnight (not shown). Since only a small amount of 2CN-c6-BP-GlcNAc was obtained with the deletion mutant, I used the LgmA produced in the original strain for moving on to research into the function of LgmB.

3.3: LgmB Appends Glucosamine to Lipid A

LgmB is an undecaprenyl phosphate glycosyltransferase of approximately 61 kD.[58] According to TMHMM 2.0, it has nine transmembrane domains.[64] LgmB is an ArnT homologue and is proposed to append GlcN from BP to lipid A.[58] This protein has been expressed in C41 and is thought to remain in the cef. There were two potential issues that I considered before studying LgmB. The first is that our method of visualizing the glycan on the undecaprenyl phosphate analogue carrier molecule is useful up until LgmB, but after LgmB, the glycan is no longer attached to the fluorescent molecule. That issue could be circumvented by utilizing the mass function of the LC-MS. The other issue is the lack of easily obtainable *B. pertussis* lipid A. As discussed earlier, the main differences between *E. coli* lipid A and *B. pertussis* lipid A are the number and length of acyl chains. *E. coli* lipid A is characterized by six acyl chains that range from twelve to fourteen carbons in length, while *B. pertussis* lipid A has five acyl chains that range from ten to fourteen carbons in length. I only had access to *E. coli* lipid A, and was interested to see if LgmB would utilize *E. coli* lipid A as an acceptable alternative substrate to native *B. pertussis* lipid A.

As a first test of LgmB activity, I incubated LgmB cef and 0.1 mg/mL free lipid A with the contents of an incubated LgmC reaction that would contain 2CN-c6-BP-GlcN. I expected to see 2CN-c6-BP-GlcN and lipid A disappear and 2CN-c6-BP and lipid A-GlcN (or lipid A-$GlcN_2$) appear. After incubating the LgmB-containing reaction contents for one hour, I evaluated the contents via ESI-LC-MS using the $[M-2H]^{2-}$ m/z ratios listed in Table 1 and Table 2. In the reserved aliquot of the LgmC reaction, I observed a SIM peak with a m/z that corresponded to 2CN-c6-BP-GlcN as expected. In the LgmB reaction with free lipid A, I saw a peak for free lipid A and 2CN-c6-BP-GlcN. I saw no new peaks for lipid A-GlcN or 2CN-c6-BP (Figure 29).

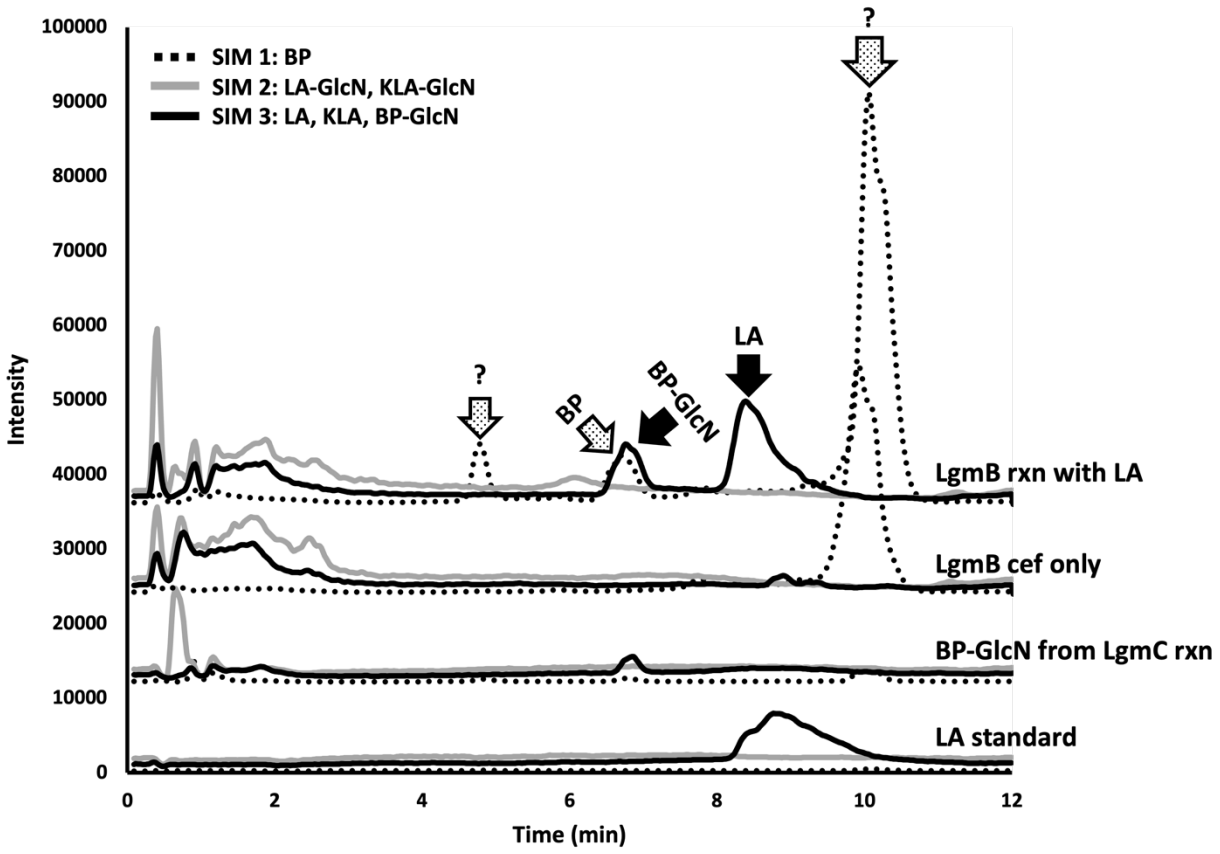

Figure 29: LgmB reaction with free lipid A evaluated by LC-MS shows no indication of LgmB activity due to presence of lipid A and BP-GlcN after the reaction is allowed to proceed. Peak of unknown identity near 10 minutes (m/z ratio same as 2CN-c5-BP) visible in LgmB cef and LgmB reaction. Additional peak with sam m/z ratio also visible in LgmB reaction near 5 minutes. Reactions are offset by 12000 units. Rxn, reaction; LA, lipid A; cef, cell envelope fraction.

Raetz and coworkes summarized the many modifications of lipid A that occur after the incorporation of Kdo and other core sugars.[5] It is possible that LgmB modifies Kdo_2-lipid A and not free lipid A, so I repeated the previous reactions with commercially obtained *E. coli* Kdo_2-lipid A. (I was not able to obtain Kdo_2-lipid A with other core sugars, but the core polysaccharides in *E. coli* lipid A are very different than the core oligosaccharides in *B. pertussis*.[5]) After incubating the LgmB and Kdo_2-lipid A-containing reaction contents for one hour, I again evaluated the contents via ESI-LC-MS using the $[M-2H]^{2-}$ m/z ratios listed in Table 2 and Table 3. In this reaction I observed some of the expected results of an LgmB reaction:

Kdo_2-lipid A disappeared, and 2CN-c6-BP appeared. I did not observe any indication of Kdo_2-lipid A-GlcN or Kdo_2-lipid A-$(GlcN)_2$ in the reaction after LgmB was allowed to react, and the SIM peak for Kdo_2-lipid A overlapped slightly with 2CN-c6-BP-GlcN (they both eluted around 7 minutes), so it was difficult to determine if either of them completely disappeared from the reaction. (Figure 30).

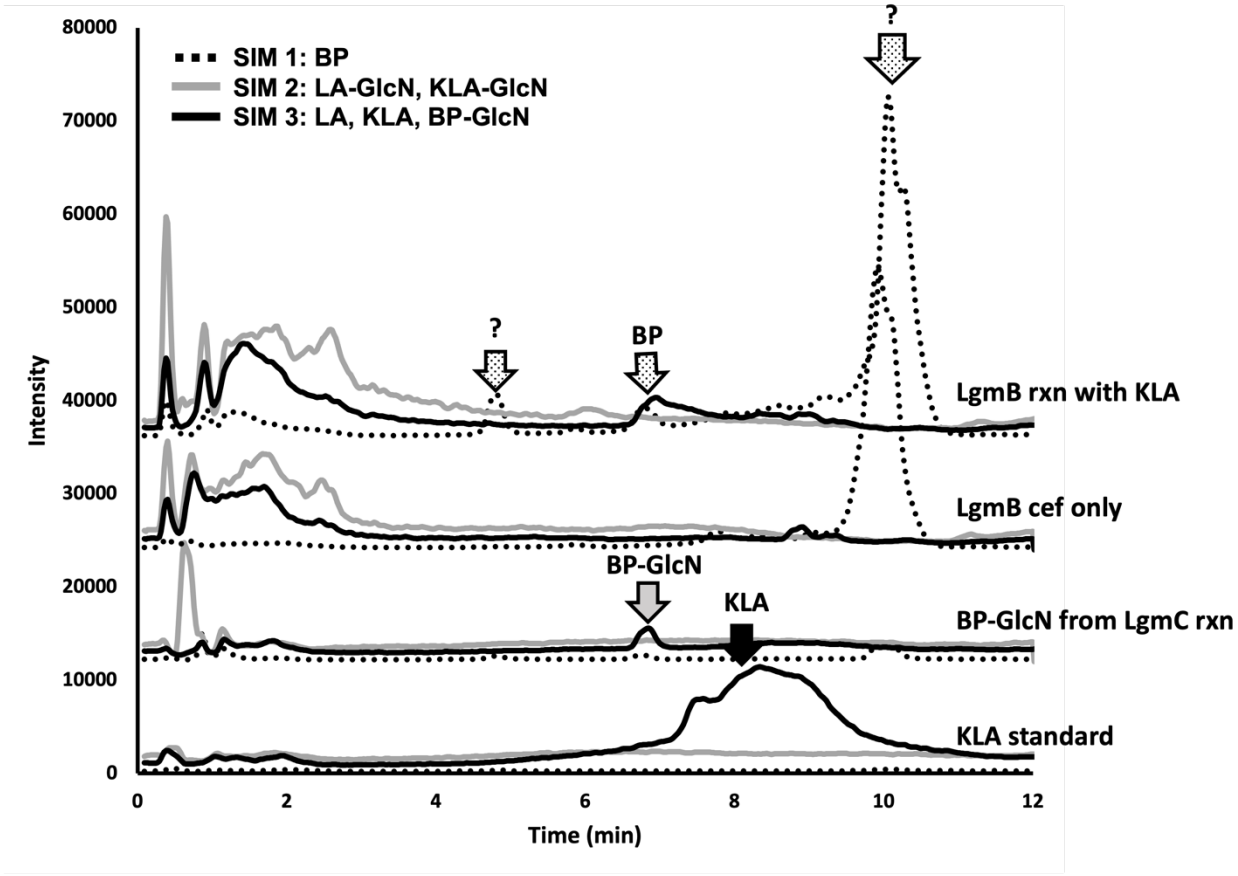

Figure 30: LgmB reaction with Kdo_2-lipid A evaluated by LC-MS indicates possible LgmB activity due to disappearance of Kdo_2-lipid A after the reaction is allowed to proceed. Peak of unknown identity near 10 minutes (m/z ratio same as 2CN-c5-BP) visible in LgmB cef and LgmB reaction. Additional peak with sam m/z ratio also visible in LgmB reaction near 5 minutes. Reactions are offset by 12000 units. Rxn, reaction; KLA, Kdo_2-lipid A; cef, cell envelope fraction.

3.4: LgmD is a Proposed Undecaprenyl Phosphate-Donor Flippase

LgmD is a proposed BP-donor flippase, but its exact role is not known. It has four transmembrane domains and is in the gtrA-like Dfam family of proposed BP-donor flippases.[67] It is not necessary for the modification of lipid A with GlcN and will not be studied in this report.[58]

3.5: Lgm Pathway Summary

I have provided biochemical evidence using fluorescence-linked HPLC and ESI-LC-MS for the putative functions of LgmA, LgmC, and LgmB. Figure 31 illustrates the entirety of the lipid A glucosamine modification pathway in *B. pertussis* as known thus far. LgmA appends GlcNAc to undecaprenyl phosphate, LgmC deacetylates it to form undecaprenyl phosphate-GlcN, and LgmB appends glucosamine onto lipid A.

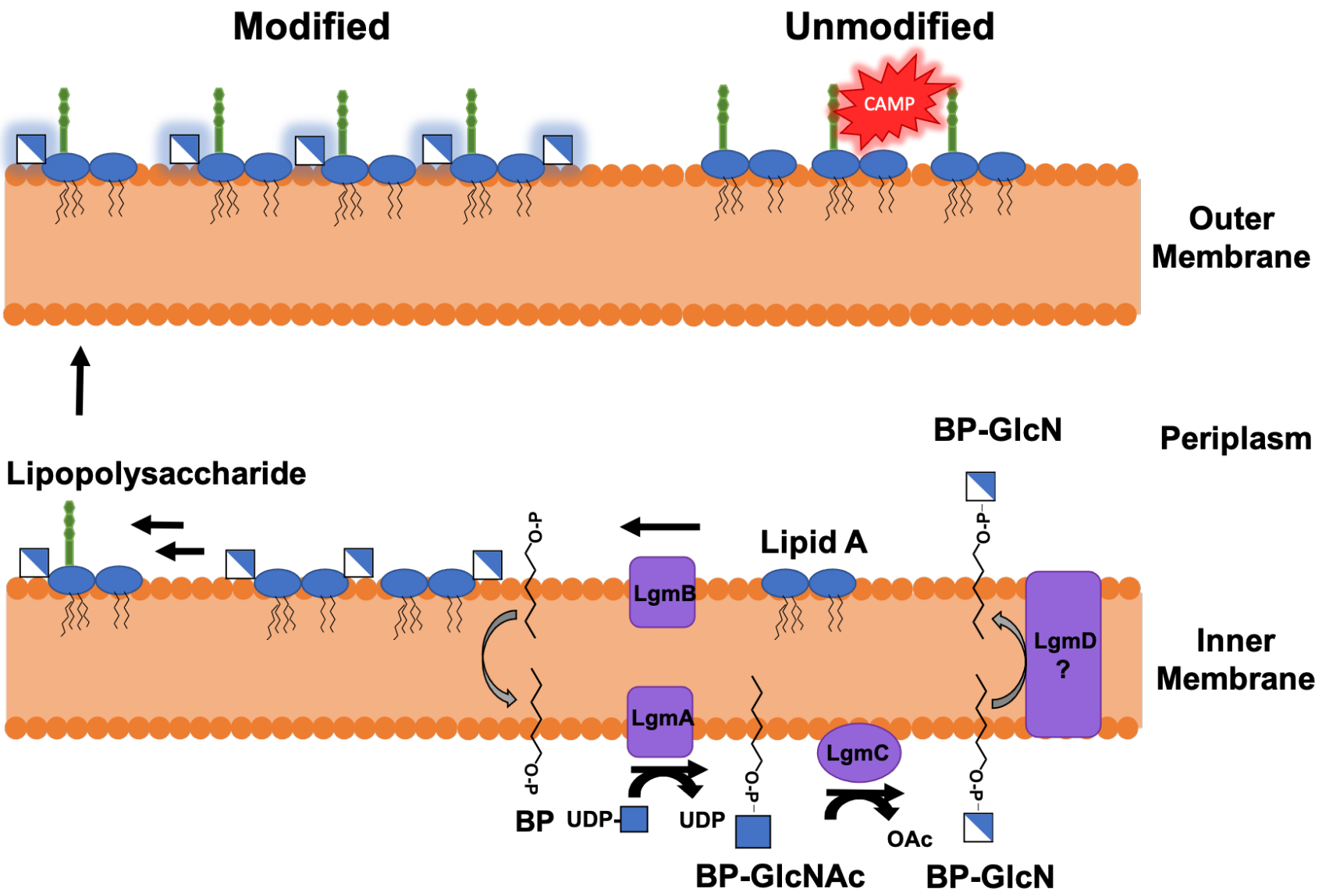

Figure 31: Schematic representation of the lipid A glucosamine modification pathway in *B. pertussis* as confirmed by this work.

CHAPTER 4: CONCLUSIONS AND FUTURE WORK

4.1: Concluding the Lgm Pathway Research

Infections of *B. pertussis* have not been eradicated despite decades of high vaccine coverage. Both whole cell and acellular vaccines, many different versions of each, have been in use in different countries with differing efficacy.[68] One of the things blocking production of a more effective vaccine is the lack of knowledge about the varied causes for pathogenesis and host colonization by *B. pertussis*, but one of the major contributors to pathogenesis has been shown to be *B. pertussis* LPS because of its role in TLR4-signaling modulation.[60, 68] In the year 2000 the structure of *B. pertussis* LPS was elucidated, with the exception of the glucosamine modification of lipid A.[51] As early as 2008 it was proposed that the lipid A modification in *B. pertussis* might be similar to the lipid A modification with 4-aminoarabinose in *E. coli,* and the modification was shown to be one or two glucosamine residues.[4] However, it took five years until genetic evidence suggested potential genes responsible for this modification and their possible functions in 2013.[58] Now nine years after the discovery of the genes of interest, we have provided biochemical evidence for the functions of the gene products attributed to glucosamine modification.

4.2: Presence of Endogenous Enzymes

One major difficulty in conducting this research is caused by the presence of endogenous enzymes in the *E. coli* cef used for *in vitro* assays of membrane bound proteins. Because of the diverse nature of bacterial glycans, there are many glycosyl transfer enzymes in bacteria, and many of them are membrane bound. Figure 3 depicts fourteen full and partial glycan assembly pathways that the Troutman group has studied, and that is only a tiny fraction of the glycan pathways that exist. There are at least two ways that endogenous enzymes can interfere with *in*

vitro assays that I have observed. The first involves a glycosyltransferase that utilizes the same activated sugar substrate as the enzyme of interest, and the second involves a lipid-bound glycan that remains with the cef. Both of these interferences were observed during the study of Lgm.

4.2.1: Arn Product in LgmA Reactions

The first way that an endogenous enzyme interfered with the Lgm pathway was the observed presence of an unknown peak in all LgmA reactions. This fluorescent peak was observed via HPLC in reactions with LgmA with and without the substrate UDP-GlcNAc, and with the empty vector control cef with and without the substrate UDP-GlcNAc (Appendix A). The two conditions necessary to produce the peak were cef and the fluorescent undecaprenyl phosphate analogue. Scarbrough determined that this peak was undecaprenyl phosphate analogue-Ara4N.[35] This may sound improbable, because no 4-aminoarabinose was added to the reactions, and all the soluble sugars should have theoretically been removed when the soluble fraction was discarded. However, it was determined that lipid A-Ara4N remained in the cef, and the enzyme ArnT was able to perform the reverse reaction and append 4-aminoarabinose to undecaprenyl phosphate, causing a fluorescent peak to appear.[35] Once the identity of the additional fluorescent peak was determined, we were able to reintroduce the *lgm* genes to a cell line with an *arnC* deletion. We proposed that a deletion of *arnC* meant that no lipid A-Ara4N would be produced, and therefore no undecaprenyl phosphate-Ara4N could be produced later when the fluorescent analogue was added.

After proteins were produced in the deletion mutants, I attempted to reproduce the functional assays, hoping that there would be no undecaprenyl phosphate-Ara4N present, and indeed there was not. However, I was never able to observe full turnover using LgmA in the

mutant. After getting rid of the Arn product, I saw a new undesired product that I determined to be resulting from endogenous WecA, which leads to discussion of the other type of issue.

4.2.2: WecA Product in LgmA Reactions

WecA is a phosphoglycosyltransferase in *E. coli* that transfers a phosphosugar from UDP-GlcNAc to undecaprenyl phosphate to form undecaprenyl diphosphate-GlcNAc. After the Arn product was eliminated, I observed a significant peak that I determined to be undecaprenyl diphosphate-GlcNAc from endogenous WecA activity. The relative position of the well-characterized WecA product had previously been used to help assign the identity of the LgmA product peak, and I was convinced of its identity.[63]

WecA was shown to be effectively inhibited by tunicamycin and LgmA was not. Tunicamycin was utilized to eliminate the WecA product, but I still did not see full turnover into the LgmA product. Trying to get rid of all of the possible glycosyltransferase pathways seemed like an insurmountable task. There are currently no known additional endogenous glycan assembly pathways in *E. coli* that utilize UDP-GlcNAc as a starting material, and no other reversible pathways thought to be able to utilize undecaprenyl phosphate, but that does not mean no other pathways exist. It is possible that a currently unknown pathway utilizes UDP-GlcNAc and is diverting it away from the LgmA pathway. It is also possible that undecaprenyl phosphate is being diverted to another pathway that does not produce a distinguishable peak on HPLC, but the large amount of remaining 2CN-c6-BP observed on the HPLC chromatogram causes me to think that this is not a significant reason for the lack of LgmA turnover.

4.3: Determining LgmC Reaction Conditions

When LgmC activity was first being probed, the conditions for the enzymatic assay needed to be set. Different divalent cations, salts, surfactants, buffers, and pH levels were

considered. Each of these considerations will be discussed in more detail here, although some conditions took more consideration than others. Many of the conditions were based on the recent research into the LgmC homolog NaxD which used an enzymatic assay of a recombinant protein, with considerations taken from common conditions in other biochemical research.

The first consideration was the divalent cation. Many enzymes, not just redox enzymes, depend on ferrous iron (Fe^{2+}) as a cofactor. However, atmospheric iron is very low, and many bacteria have evolved methods to function in low-iron environments. Magnesium(II) and zinc(II) were early replacements for iron(II) in *in vitro* assays, but in 1999, manganese(II) was named as a substitute to magnesium and zinc in many enzymatic assays.[66] Studies have suggested that a recently discovered *E. coli* small protein MntS somehow promotes Mn^{2+} insertion into non-redox enzymes to allow the enzymes to function without iron.[69, 70] In a low iron environment such as an *in vitro* experiment, it is unclear whether endogenous MntS would still function (or be required to function), or whether the presence of excess Mn^{2+} in the buffered conditions would be sufficient to allow it to function as the required cofactor.

Many bacteria, including *B. pertussis*, have been known to have replaced the iron(II) cofactor of many enzymes with manganese(II), which is more readily available.[71-73] It is not yet clear whether an MntS homolog is present in *B. pertussis*, but it does contain other manganese transport proteins.[74] Whether the *B. pertussis* enzyme LgmC natively functions with manganese(II) as a cofactor or not cannot be stated definitively from this work; however the greatest turnover that I observed with LgmC was in conditions with divalent manganese as the cofactor. I cannot determine whether manganese(II) functioned as the cofactor alone, or with the help of the small protein MntS, but it would be interesting to determine if the *E. coli* small protein was present in the cef or not. The forty-seven amino acid protein is predicted to be

soluble but has been shown to be associated with the inner membrane of *E. coli*.[64, 75] Perhaps it even remains in cef when *E. coli* is used for recombinant protein expression, as in the case with LgmC. It would also be interesting to determine if *B. pertussis* possesses an MntS homolog, and if so, how similar the proteins were.

Regarding the salt chosen for LgmC enzymatic assay, I tested two common salts: NaCl and KCl, and observed equal activity in the presence of both. I continued with NaCl as the salt for further experiments because it seemed like the most commonplace.

Surfactants are often necessary in assays containing lipids like undecaprenyl phosphate in order to help solubilize the lipids. The Raetz group used 0.1% Triton in their NaxD assays, but I did not find any surfactant necessary to observe LgmC activity. I attempted to purify LgmC and used Triton-X to accomplish this. After using 0.1% Triton-X in the elution buffer during Ni-NTA purification, there was an approximate concentration of 0.01% Triton-X in the purified LgmC reaction mixtures, and I did not see any activity. I was unsure if the lack of activity was due to the surfactant in the reaction mixture or was simply due to the fact that LgmC needed to be stabilized by the membrane and was not functional apart from the membrane. To test whether Triton-X eliminated LgmC activity, I incubated LgmC cef (previously shown to work) with different concentrations of Trition-X. At a Triton-X concentration of ##, LgmC cef exhibited deacetylase activity. I then decreased the amount of purified LgmC in the reaction mixture to decrease the concentration of Triton-X. At this same concentration of Triton-X that LgmC cef was shown to be functional, and even lower concentrations (down to 0.001 % Triton-X), purified LgmC did not exhibit any deactylase activity (not shown). This suggested that the lack of activity exhibited by the purified enzyme was due to malformation or instability apart from the membrane, and not due to the presence of Triton-X alone. I also tested other surfactants (cholate

and DDM) and observed similar patterns: that LgmC cef was functional in low concentrations of other surfactants. I resorted to using LgmC cef for all further enzymatic reactions.

There are many commercially available buffer systems; the Raetz group used 50 mM HEPES buffer (pH 7.5), and I used similar conditions with 100 mM HEPES buffer (pH 8.0) which was chosen partially for similarity to the Raetz conditions and partially for convenience. The initial test conditions varied the salts and divalent cations, and I was able to observe full turnover by LgmC with both salts and one of the divalent cations (Mn^{2+}) in the other described conditions, so I did not attempt to alter the buffer or pH.

4.4: Future Work in Assessing LgmC Specificity

Because of the extensive work by the Troutman laboratory with various glycan assembly pathways, our laboratory has access to many fluorescently tagged undecaprenyl phosphate analogue glycan intermediates. I have evidence that LgmC is a deacetylase that efficiently deacetylates undecaprenyl phosphate-GlcNAc, and further experiments could be designed to test its specificity with other modified sugars such as *N*-formylated 4-amino-4-deoxy-L-arabinose, an intermediate in the Arn pathway of *E. coli*, to see if it can deformylate the intermediate monophosphoryl acetylated glycan. LgmC activity could also be tested on two diphosphoryl acetylated glycans (the products of WecA and WecP: BPP-GlcNAc and BPP-GalNAc, respectively) to evaluate the difference that an additional phosphate group has on how the substrate interacts with the active site of LgmC.

4.4: LC-MS Concerns

One difficulty with using LC-MS to evaluate membrane bound enzymatic assays is sifting through all the debris that comes along with the membrane fraction. It can be difficult to find the m/z ratio of interest amidst the many peaks observed. One way to mitigate this difficulty

is to isolate or at least partially isolate the compound of interest after the enzymatic assay is complete. One way I have done this is with *n*-butanol extraction of 2CN-c6-BP-GlcNAc (after LgmA reaction) and a two-phase Bligh-Dyer system to isolate 2CN-c6-BP-GlcNAc (after LgmC reaction) and lipid A (after LgmB reaction). These methods are not perfect isolation techniques however, because many lipids besides the desired lipids are also included.

When a sample is evaluated via LC-MS, I have the opportunity to set up three selected ion monitoring (SIM) channels. I do my best to assign different m/z ratios to the SIMs so that I can quickly and easily see whether or not a selected ion is present. For the LgmC reactions, I was able to couple fluorescence detection with mass detection. This coupling allowed me to assign peak identities based on elution time (fluorescence) as well as m/z ratio (mass detection). When the expected elution times matched the m/z ratios, it was much greater support for my assignments than just one or the other.

For the LgmB reactions, I wanted to monitor 2CN-c6-BP-GlcN, 2CN-c6-BP, free lipid A, Kdo_2-lipid A, free lipid A-GlcN, Kdo_2-lipid A-GlcN, free lipid A-$(GlcN)_2$, and Kdo_2-lipid A-$(GlcN)_2$. I distributed these eight ions to the SIMS in the way that made the most sense to me in looking for the appearance and disappearance of them in the LgmB reaction, but it was not possible to only have one ion per SIM. I used the LC-MS in negative ion mode with a ammonium hydroxide buffer with pH 11. I had to take into account the charges that the ions of interest would possess in such buffer conditions in order to determine the m/z ratio to monitor. Knowledge of the approximate pK_a values of the phosphate and hydroxyl functional groups of the large molecules was essential to determine the proposed charge, but experimental data was necessary as well because the observed m/z ratio of complex phospholipids in our laboratory has been different than what we predicted (not shown). For example, undecaprenyl diphosphate has

two phosphate groups. The approximate pK_as of phosphate are 2.2, 7.2, and 12.4. At a pH of 11, at least two of the hydroxyl groups on one of the phosphates are predicted to be deprotonated, leading to a charge of -2. The m/z ratio would then be $[M-2H]^{2-}$; however, what has actually been observed is the $[M-H]^-$ peak (not shown).

Another predicament comes when the analyte of interest is uncharged in the selected buffer. The buffer used in our LC-MS is 0.1% ammonium hydroxide with a pH of approximately 11. At this pH, it is likely that at least one hydroxyl group of an unmodified phosphate on lipid A would be deprotonated to give a charge of minus one, but it is difficult to say what the pKa of a phosphate linked to a glucosamine residue would be. The amine on glucosamine is likely protonated at biological pH, but it is unlikely to be protonated at pH 11; it is likely neutral. If one of the phosphates on lipid A is deprotonated giving the entire structure a minus one charge, the m/z ratio would be over 2000, which is the limit of our detector. If both of the phosphates are deprotonated, the charge would be minus two, and the m/z ratio of the $[M-2H]^{2-}$ species would be in the range of the detector (Table 3). Even though the predicted species has a m/z value within the range of the detector, it was not observed. At this point in time, it is not possible to say with certainty whether the species was not observed because it was not present, or because it was not able to be detected for some reason, perhaps because the charge (or lack of charge) made it difficult to detect.

The reactions conducted with LgmB suggested but in no way proved that LgmB is necessary or sufficient to transfer glucosamine from undecaprenyl phosphate to lipid A. More reactions with LgmB are needed to repeat the data that I already collected, and some controls are necessary to isolate LgmB as the source of the transfer, and not something already present in the cef. These controls include but are not limited to 1) incubating lipid A with LgmB cef without

BP-GlcN, 2) incubating lipid A with an empty vector control cef with BP-GlcN, 3) incubating lipid A with an empty vector control cef without BP-GlcN, 4) incubating lipid A with LgmC and BP-GlcN, and 5) incubating lipid A with LgmA and BP-GlcN. It would be helpful to pinpoint the correct conditions that would cause lipid A-GlcN to be charged and show up in the mass spectrum instead of depending on visualizing the disappearance of the other reactants. It would also be beneficial to obtain *B. pertussis* lipid A (with and without the core oligosaccharides) in case LgmB is specific to *B. pertussis* lipid A.

4.5: Significance of This Research

Many bacterial infections and diseases that were once almost eradicated have become a threat once again. Multi-drug resistant bacteria and the decrease in efficacy of some vaccines has led to the resurgence of bacterial infections. Understanding the biochemical role of enzymes in lipid A modification pathways and their impact on pathogenesis is crucial to the potential for novel treatment options. Pinpointing the role of each enzyme in a pathway based on biochemical assays is one step further than assigning the putative functions based on sequence analysis, and the next step is gaining more information into the mechanism by which the enzymes catalyze the formation of toxic or immunosuppressant substances. Understanding the biochemistry of lipid A modifications and their impact on pathogenesis could lead to novel treatment options for these diseases. Such information could lead to improvement upon the traditional gram-negative bacterial live vaccines as well as heat-killed vaccines like Tdap (tetanus and diphtheria toxoids with acellular pertussis).

REFERENCES

1. Bar-On, Y. M.; Phillips, R.; Milo, R., The biomass distribution on Earth. *Proceedings of the National Academy of Sciences* **2018,** *115* (25), 6506-6511.

2. Song, F.; Guan, Z.; Raetz, C. R., Biosynthesis of undecaprenyl phosphate-galactosamine and undecaprenyl phosphate-glucose in Francisella novicida. *Biochemistry* **2009,** *48* (6), 1173-82.

3. Eade, C., Glycan-Building Toolkit of the Troutman Lab. UNC Charlotte: 2019.

4. Marr, N., Tirsoaga, A., Blanot, D., Fernandez, R., and Caroff, M., Glucosamine found as a substituent of both phosphate groups in Bordetella lipid A backbones: role of a BvgAS-activated ArnT ortholog. *J Bacteriol* **2008,** *190*, 4281–4290.

5. Christian R.H. Raetz, C. M. R., M. Stephen Trent, and Russell E. Bishop, Lipid A Modification Systems in Gram-Negative Bacteria. *Annu. Rev. Biochem.* **2007,** *76*, 295–329.

6. Westerik, N.; Reid, G.; Sybesma, W.; Kort, R., The Probiotic Lactobacillus rhamnosus for Alleviation of Helicobacter pylori-Associated Gastric Pathology in East Africa. *Frontiers in Microbiology* **2018,** *9*.

7. Lee, S.; Inzerillo, S.; Lee, G. Y.; Bosire, E. M.; Mahato, S. K.; Song, J., Glycan-mediated molecular interactions in bacterial pathogenesis. *Trends in Microbiology* **2022,** *30* (3), 254-267.

8. Comstock, L. E. a. K., Dennis L., Bacterial Glycans: Key Mediators of Diverse Host Immune Responses. *Minireview* **2006,** *126* (5), 847-850.

9. Rolin, O.; Muse, S. J.; Safi, C.; Elahi, S.; Gerdts, V.; Hittle, L. E.; Ernst, R. K.; Harvill, E. T.; Preston, A., Enzymatic modification of lipid A by ArnT protects Bordetella bronchiseptica against cationic peptides and is required for transmission. *Infect Immun* **2014,** *82* (2), 491-9.

10. Shah, N. R.; Hancock, R. E.; Fernandez, R. C., Bordetella pertussis lipid A glucosamine modification confers resistance to cationic antimicrobial peptides and increases resistance to outer membrane perturbation. *Antimicrob Agents Chemother* **2014,** *58* (8), 4931-4.

11. Mao, Y.; Doyle, M. P.; Chen, J., Insertion mutagenesis of wca reduces acid and heat tolerance of enterohemorrhagic Escherichia coli O157:H7. *J Bacteriol* **2001,** *183* (12), 3811-5.

12. Morrison, D. C.; Jacobs, D. M., Binding of polymyxin B to the lipid A portion of bacterial lipopolysaccharides. *Immunochemistry* **1976,** *13* (10), 813-8.

13. Williams, D. A.; Pradhan, K.; Paul, A.; Olin, I. R.; Tuck, O. T.; Moulton, K. D.; Kulkarni, S. S.; Dube, D. H., Metabolic inhibitors of bacterial glycan biosynthesis. *Chemical Science* **2020,** *11* (7), 1761-1774.

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

15. Llewellyn, A. C., et al., NaxD is a deacetylase required for lipid A modification and Francisella pathogenesis. *Mol Microbiol* **2012,** *86*, 611-627.

16. Takahito Imagawa, H. I., Mayumi Kanagawa, Akio Ebihara, Seiki Kuramitsu, Hideaki Tsuge, Crystal structure of the YdjC-family protein TTHB029 from Thermus thermophilus HB8: Structural relationship with peptidoglycan N-acetylglucosamine deacetylase. *Biochemical and Biophysical Research Communications* **2008,** *367* (3), 535-541.

17. Wang, X.; Ribeiro, A. A.; Guan, Z.; Raetz, C. R., Identification of undecaprenyl phosphate-beta-D-galactosamine in Francisella novicida and its function in lipid A modification. *Biochemistry* **2009,** *48* (6), 1162-72.

18. Reid, A. J.; Scarbrough, B. A.; Williams, T. C.; Gates, C. E.; Eade, C. R.; Troutman, J. M., General Utilization of Fluorescent Polyisoprenoids with Sugar Selective Phosphoglycosyltransferases. *Biochemistry* **2020,** *59* (4), 615-626.

19. Manat, G.; Roure, S.; Auger, R.; Bouhss, A.; Barreteau, H.; Mengin-Lecreulx, D.; Touzé, T., Deciphering the metabolism of undecaprenyl-phosphate: the bacterial cell-wall unit carrier at the membrane frontier. *Microb Drug Resist* **2014,** *20* (3), 199-214.

20. Reid, A. J.; Eade, C. R.; Jones, K. J.; Jorgenson, M. A.; Troutman, J. M., Tracking Colanic Acid Repeat Unit Formation from Stepwise Biosynthesis Inactivation in Escherichia coli. *Biochemistry* **2021**.

21. Scott, P. M.; Erickson, K. M.; Troutman, J. M., Identification of the Functional Roles of Six Key Proteins in the Biosynthesis of Enterobacteriaceae Colanic Acid. *Biochemistry* **2019,** *58* (13), 1818-1830.

22. Dube, D. H.; Bertozzi, C. R., Metabolic oligosaccharide engineering as a tool for glycobiology. *Current Opinion in Chemical Biology* **2003,** *7* (5), 616-625.

23. Laughlin, S., Bertozzi, C., Metabolic labeling of glycans with azido sugars and subsequent glycan-profiling and visualization via Staudinger ligation. *Nature Protocols* **2007,** *2*, 2930–2944.

24. Mulloy, B. H., G.W.; and Stanley, P. , Structural Analysis of Glycans. In *Essentials of Glycobiology*, 2nd edition ed.; Varki, A. C., R.D.; Esko, J.D.; et al., Ed. Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2009.

25. Marando, V. M.; Kim, D. E.; Calabretta, P. J.; Kraft, M. B.; Bryson, B. D.; Kiessling, L. L., Biosynthetic Glycan Labeling. *Journal of the American Chemical Society* **2021,** *143* (40), 16337-16342.

26. Linton, D.; Dorrell, N.; Hitchen, P. G.; Amber, S.; Karlyshev, A. V.; Morris, H. R.; Dell, A.; Valvano, M. A.; Aebi, M.; Wren, B. W., Functional analysis of the Campylobacter jejuni N-linked protein glycosylation pathway. *Mol Microbiol* **2005,** *55* (6), 1695-703.

27. Saxon, E.; Bertozzi, C. R., Cell surface engineering by a modified Staudinger reaction. *Science* **2000,** *287* (5460), 2007-10.

28. Saxon, E.; Luchansky, S. J.; Hang, H. C.; Yu, C.; Lee, S. C.; Bertozzi, C. R., Investigating Cellular Metabolism of Synthetic Azidosugars with the Staudinger Ligation. *Journal of the American Chemical Society* **2002,** *124* (50), 14893-14902.

29. Evans, R. A., The Rise of Azide–Alkyne 1,3-Dipolar 'Click' Cycloaddition and its Application to Polymer Science and Surface Modification. *Australian Journal of Chemistry* **2007,** *60* (6), 384-395.

30. Reactions, C. C. D. C. F. f. a. F. G., Hartmuth C. Kolb Dr., M. G. Finn Prof., K. Barry Sharpless Prof. *Angew. Chem. Int. Ed.* **2001,** *40* (11), 2004-2021.

31. Clark, E. L.; Emmadi, M.; Krupp, K. L.; Podilapu, A. R.; Helble, J. D.; Kulkarni, S. S.; Dube, D. H., Development of Rare Bacterial Monosaccharide Analogs for Metabolic Glycan Labeling in Pathogenic Bacteria. *ACS Chem Biol* **2016,** *11* (12), 3365-3373.

32. Glover, K. J.; Weerapana, E.; Chen, M. M.; Imperiali, B., Direct biochemical evidence for the utilization of UDP-bacillosamine by PglC, an essential glycosyl-1-phosphate transferase in the Campylobacter jejuni N-linked glycosylation pathway. *Biochemistry* **2006,** *45* (16), 5343-50.

33. Glover, K. J.; Weerapana, E.; Imperiali, B., In vitro assembly of the undecaprenylpyrophosphate-linked heptasaccharide for prokaryotic N-linked glycosylation. *Proc Natl Acad Sci U S A* **2005,** *102* (40), 14255-14259.

34. Hartley, M. D.; Larkin, A.; Imperiali, B., Chemoenzymatic synthesis of polyprenyl phosphates. *Bioorg Med Chem* **2008,** *16* (9), 5149-56.

35. Scarbrough, B. A.; Eade, C. R.; Reid, A. J.; Williams, T. C.; Troutman, J. M., Lipopolysaccharide Is a 4-Aminoarabinose Donor to Exogenous Polyisoprenyl Phosphates through the Reverse Reaction of the Enzyme ArnT. *ACS Omega* **2021,** *6* (39), 25729-25741.

36. Hartley, M. D.; Imperiali, B., At the membrane frontier: a prospectus on the remarkable evolutionary conservation of polyprenols and polyprenyl-phosphates. *Arch Biochem Biophys* **2012,** *517* (2), 83-97.

37. Yan, A.; Guan, Z.; Raetz, C. R., An undecaprenyl phosphate-aminoarabinose flippase required for polymyxin resistance in Escherichia coli. *J Biol Chem* **2007,** *282* (49), 36077-89.

38. Kanjilal-Kolar, S.; Raetz, C. R. H., Dodecaprenyl phosphate-galacturonic acid as a donor substrate for lipopolysaccharide core glycosylation in Rhizobium leguminosarum. *The Journal of biological chemistry* **2006,** *281* (18), 12879-12887.

39. TouzÉ, T.; Mengin-Lecreulx, D.; Slauch, J. M., Undecaprenyl Phosphate Synthesis. *EcoSal Plus* **2008,** *3* (1).

40. Troutman, J. M.; Imperiali, B., Campylobacter jejuni PglH is a single active site processive polymerase that utilizes product inhibition to limit sequential glycosyl transfer reactions. *Biochemistry* **2009,** *48* (12), 2807-16.

41. Chen, M. M.; Weerapana, E.; Ciepichal, E.; Stupak, J.; Reid, C. W.; Swiezewska, E.; Imperiali, B., Polyisoprenol specificity in the Campylobacter jejuni N-linked glycosylation pathway. *Biochemistry* **2007,** *46* (50), 14342-8.

42. Lujan, D. K.; Stanziale, J. A.; Mostafavi, A. Z.; Sharma, S.; Troutman, J. M., Chemoenzymatic synthesis of an isoprenoid phosphate tool for the analysis of complex bacterial oligosaccharide biosynthesis. *Carbohydr. Res.* **2012,** *359*, 44-53.

43. Kuo, C.-J.; Guo, R.-T.; Lu, I. L.; Liu, H.-G.; Wu, S.-Y.; Ko, T.-P.; Wang, A. H. J.; Liang, P.-H., Structure-Based Inhibitors Exhibit Differential Activities against Helicobacter pylori and Escherichia coli Undecaprenyl Pyrophosphate Synthases. *Journal of Biomedicine and Biotechnology* **2008,** *2008*, 841312.

44. Chehade, K. A. H.; Kiegiel, K.; Isaacs, R. J.; Pickett, J. S.; Bowers, K. E.; Fierke, C. A.; Andres, D. A.; Spielmann, H. P., Photoaffinity Analogues of Farnesyl Pyrophosphate Transferable by Protein Farnesyl Transferase. *Journal of the American Chemical Society* **2002,** *124* (28), 8206-8219.

45. Mostafavi, A. Z.; Lujan, D. K.; Erickson, K. M.; Martinez, C. D.; Troutman, J. M., Fluorescent probes for investigation of isoprenoid configuration and size discrimination by bactoprenol-utilizing enzymes. *Bioorg. Med. Chem.* **2013,** *21* (17), 5428-35.

46. Dodbele, S.; Martinez, C. D.; Troutman, J. M., Species differences in alternative substrate utilization by the antibacterial target undecaprenyl pyrophosphate synthase. *Biochemistry* **2014,** *53* (30), 5042-50.

47. Troutman, J. M.; Erickson, K. M.; Scott, P. M.; Hazel, J. M.; Martinez, C. D.; Dodbele, S., Tuning the production of variable length, fluorescent polyisoprenoids using surfactant-controlled enzymatic synthesis. *Biochemistry* **2015,** *54* (18), 2817-27.

48. Sharma, S. E., K. M.; and Troutman, J. M., Complete Tetrasaccharide Repeat Unit Biosynthesis of the Immunomodulatory Bacteroides fragilis Capsular Polysaccharide A. *ACS Chem Biol.* **2017,** *12* (1), 92-101.

49. Christian R. H. Raetz, Z. G., Brian O. Ingram, David A. Six, Feng Song, Xiaoyuan Wang, and Jinshi Zhao, Discovery of new biosynthetic pathways: the lipid A story. *J. Lipid Res.* **2009,** *50*, S103-S108.

50. Raetz, C. R.; Whitfield, C., Lipopolysaccharide endotoxins. *Annu Rev Biochem* **2002,** *71*, 635-700.

51. Caroff, M.; Brisson, J.; Martin, A.; Karibian, D., Structure of the Bordetella pertussis 1414 endotoxin. *FEBS Lett* **2000,** *477* (1-2), 8-14.

52. Phillips, N. J.; Schilling, B.; McLendon, M. K.; Apicella, M. A.; Gibson, B. W., Novel modification of lipid A of Francisella tularensis. *Infect Immun* **2004,** *72* (9), 5340-8.

53. Wang, X.; Ribeiro, A. A.; Guan, Z.; Abraham, S. N.; Raetz, C. R., Attenuated virulence of a Francisella mutant lacking the lipid A 4'-phosphatase. *Proc Natl Acad Sci U S A* **2007,** *104* (10), 4136-41.

54. Whitfield, C.; Trent, M. S., Biosynthesis and export of bacterial lipopolysaccharides. *Annu Rev Biochem* **2014,** *83*, 99-128.

55. Trent, M. S.; Ribeiro, A. A.; Lin, S.; Cotter, R. J.; Raetz, C. R., An inner membrane enzyme in Salmonella and Escherichia coli that transfers 4-amino-4-deoxy-L-arabinose to lipid A: induction on polymyxin-resistant mutants and role of a novel lipid-linked donor. *J Biol Chem* **2001,** *276* (46), 43122-31.

56. Breazeale, S. D.; Ribeiro, A. A.; Raetz, C. R., Oxidative decarboxylation of UDP-glucuronic acid in extracts of polymyxin-resistant Escherichia coli. Origin of lipid a species modified with 4-amino-4-deoxy-L-arabinose. *J Biol Chem* **2002,** *277* (4), 2886-96.

57. Wang, X.; Ribeiro, A. A.; Guan, Z.; McGrath, S. C.; Cotter, R. J.; Raetz, C. R., Structure and biosynthesis of free lipid A molecules that replace lipopolysaccharide in Francisella tularensis subsp. novicida. *Biochemistry* **2006,** *45* (48), 14427-40.

58. Shah, N. R.; Albitar-Nehme, S.; Kim, E.; Marr, N.; Novikov, A.; Caroff, M.; Fernandez, R. C., Minor modifications to the phosphate groups and the C3' acyl chain length of lipid A in two Bordetella pertussis strains, BP338 and 18-323, independently affect Toll-like receptor 4 protein activation. *J Biol Chem* **2013,** *288* (17), 11751-60.

59. Tirsoaga, A.; El Hamidi, A.; Perry, M. B.; Caroff, M.; Novikov, A., A rapid, small-scale procedure for the structural characterization of lipid A applied to Citrobacter and Bordetella strains: discovery of a new structural element. *J Lipid Res* **2007,** *48* (11), 2419-27.

60. Marr, N.; Hajjar, A. M.; Shah, N. R.; Novikov, A.; Yam, C. S.; Caroff, M.; Fernandez, R. C., Substitution of the Bordetella pertussis lipid A phosphate groups with glucosamine is required for robust NF-kappaB activation and release of proinflammatory cytokines in cells expressing human but not murine Toll-like receptor 4-MD-2-CD14. *Infect Immun* **2010,** *78* (5), 2060-9.

61. Bradford, M. M., A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **1976,** *72*, 248-54.

62. Davisson, V. J. W., A. B.; Neal, T. R.; Stremler, K. E.; Muehlbacher, M. and Poulter, C. D., Phosphorylation of Isoprenoid Alcohols. *J. Org. Chem.* **1986,** *51*, 4768-4779.

63. Okwei, E. N. N. Elucidating the Molecular Basis of Inhibition of Enzymes Essential for Baterialvirulence: The Activity of Three Glycosyltransferases and a Potential Drug Target, Undecaprenyl Pyrophosphate Synthase. UNCC, 2016.

64. TMHMM Server v. 2.0. http://www.cbs.dtu.dk/services/TMHMM/ (accessed March 12).

65. Scarbrough, B. A., Lgm pathway. Unpublished work. 2021.

66. Bock, C. W.; Katz, A. K.; Markham, G. D.; Glusker, J. P., Manganese as a Replacement for Magnesium and Zinc: Functional Comparison of the Divalent Ions. *Journal of the American Chemical Society* **1999,** *121* (32), 7360-7372.

67. Family: GtrA (PF04138). https://pfam.xfam.org/family/PF04138 (accessed March 12).

68. Crowcroft, N. S.; Pebody, R. G., Recent developments in pertussis. *The Lancet* **2006,** *367* (9526), 1926-1936.

69. Correction: The Escherichia coli Small Protein MntS and Exporter MntP Optimize the Intracellular Concentration of Manganese. *PLoS Genet* **2015,** *11* (6), e1005322.

70. Martin, J. E.; Waters, L. S.; Storz, G.; Imlay, J. A., The Escherichia coli small protein MntS and exporter MntP optimize the intracellular concentration of manganese. *PLoS Genet* **2015,** *11* (3), e1004977.

71. Zaharik, M. L.; Finlay, B. B., Mn2+ and bacterial pathogenesis. *Front Biosci* **2004,** *9*, 1035-42.

72. Kehres, D. G.; Maguire, M. E., Emerging themes in manganese transport, biochemistry and pathogenesis in bacteria. *FEMS Microbiol Rev* **2003,** *27* (2-3), 263-90.

73. Bosma, E. F.; Rau, M. H.; van Gijtenbeek, L. A.; Siedler, S., Regulation and distinct physiological roles of manganese in bacteria. *FEMS Microbiol Rev* **2021,** *45* (6).

74. Brickman, T. J.; Armstrong, S. K., Bordetella pertussis fur gene restores iron repressibility of siderophore and protein expression to deregulated Bordetella bronchiseptica mutants. *Journal of Bacteriology* **1995,** *177* (1), 268-270.

75. Fontaine, F.; Fuchs, R. T.; Storz, G., Membrane localization of small proteins in Escherichia coli. *J Biol Chem* **2011,** *286* (37), 32464-74.

APPENDIX A: LgmA cef control

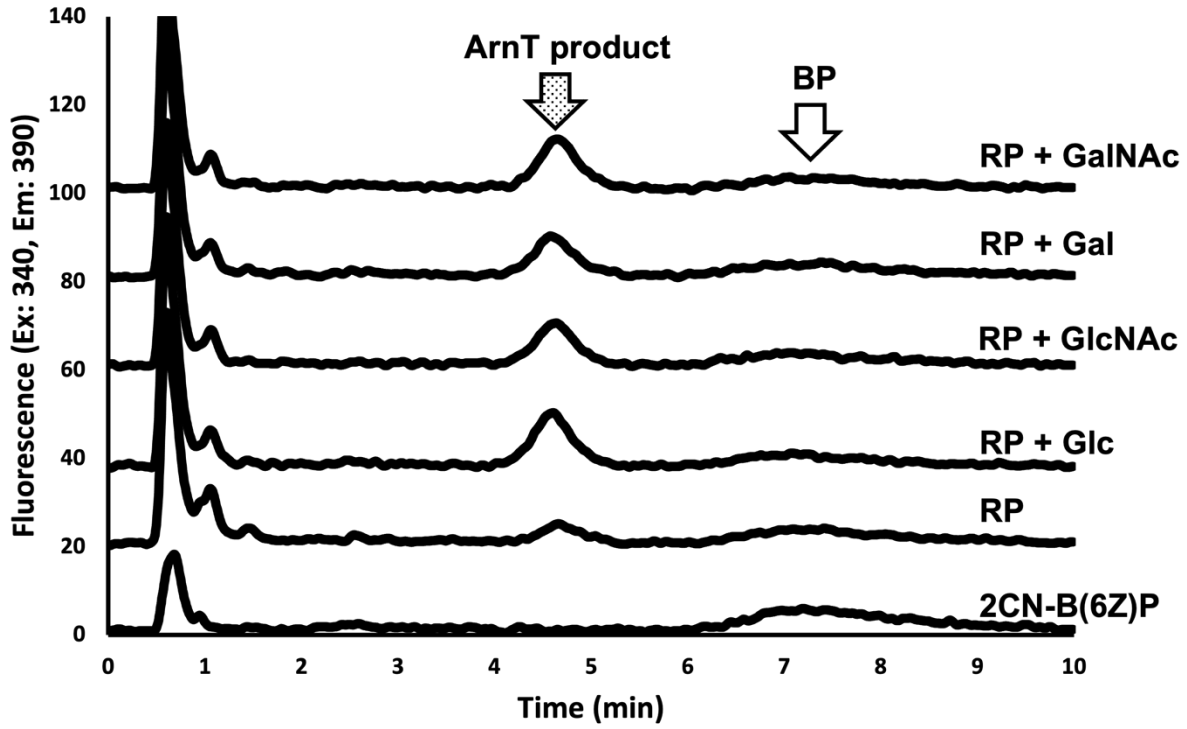

Figure 32: RP empty vector control for LgmA specificity assay. The ArnT product is formed in the presence of all four sugars and even in the absence of added sugar. No LgmA product formation is observed with the cef control.

APPENDIX B: *E.coli* lipid A

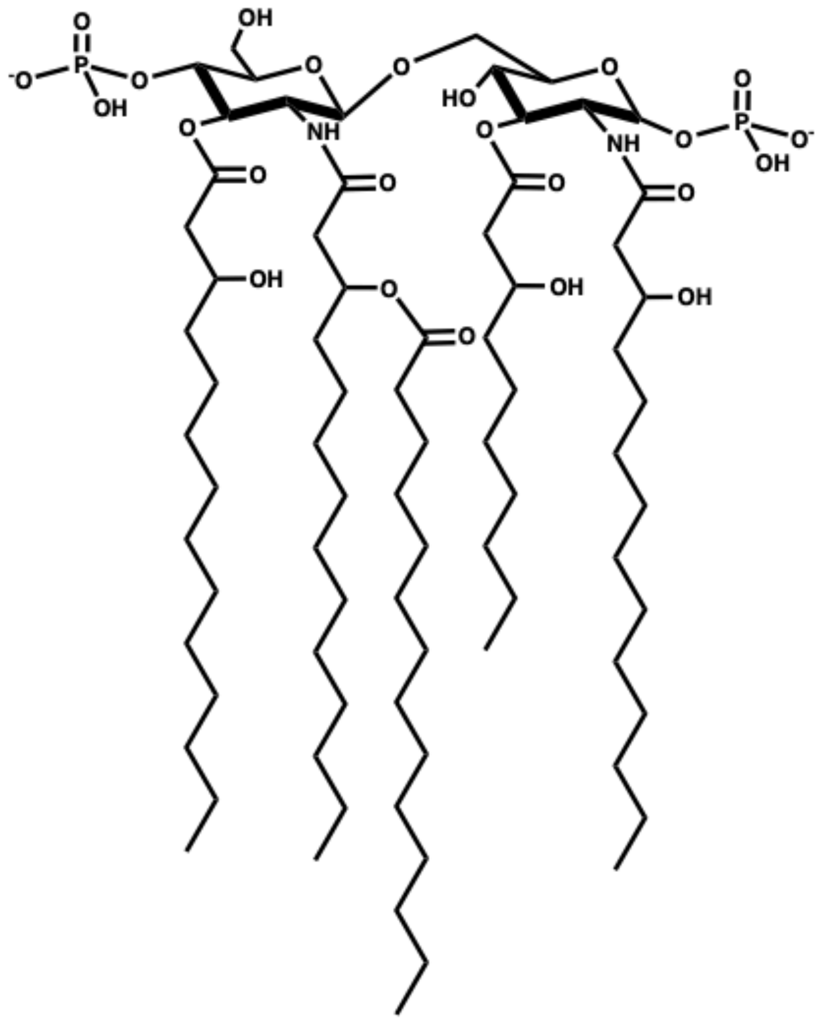

Figure 33: *E.coli* free lipid A.

APPENDIX C: *E. coli* Kdo_2-lipid A

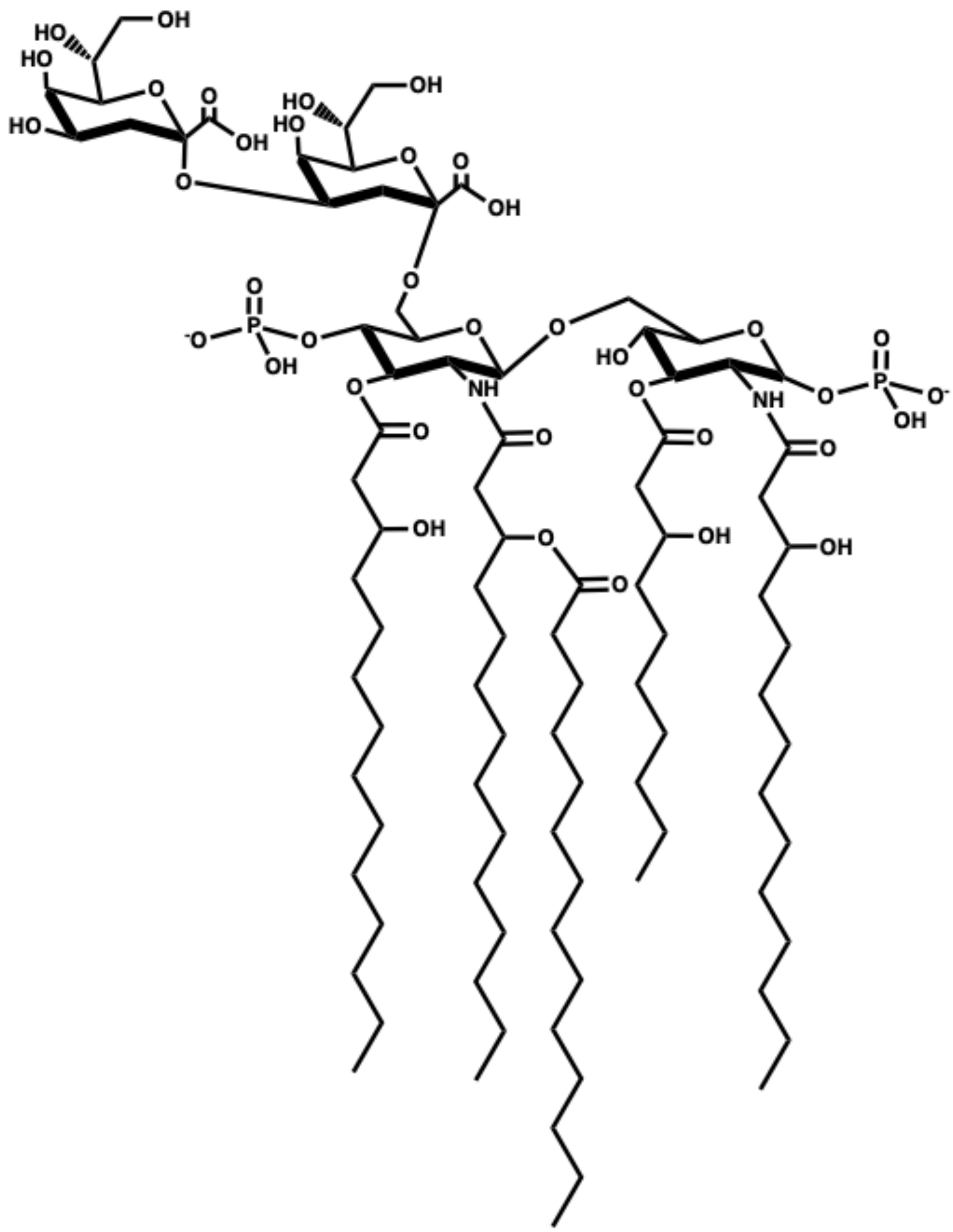

Figure 34: *E. coli* Kdo_2-lipid A.

APPENDIX D: *B. pertussis* lipid A

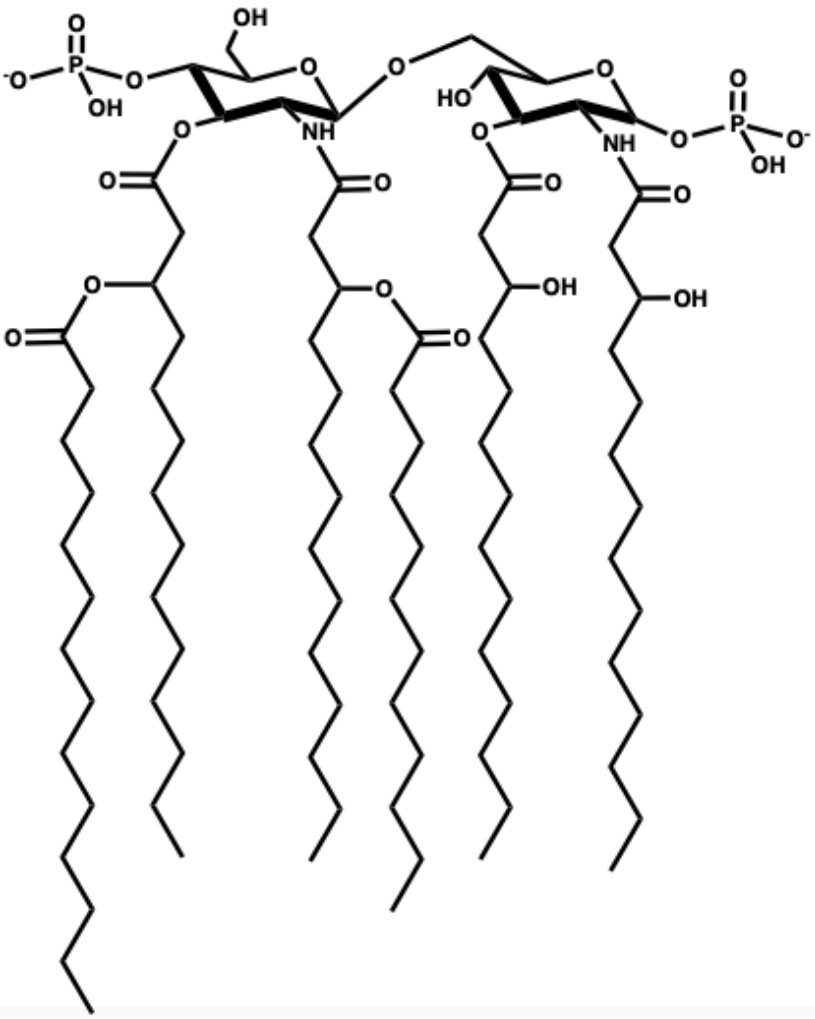

Figure 35: *B. pertussis* free lipid A.

APPENDIX E: *B. pertussis* Kdo2-lipid A

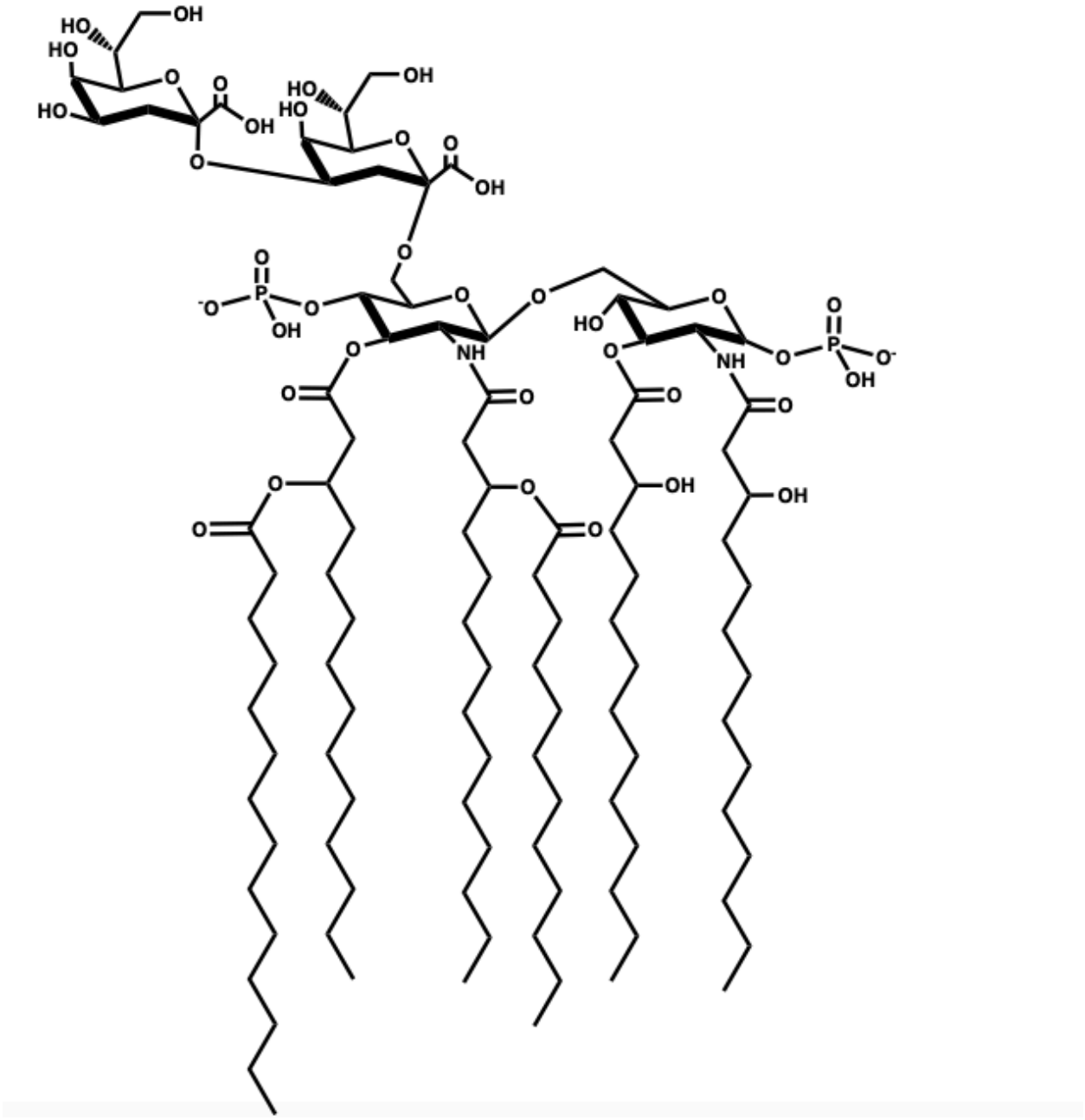

Figure 36: *B. pertussis* Kdo_2-lipid A.