ANALYSIS AND DEVELOPMENT OF TARGETING STRATEGIES FOR BACTERIAL SUGARS

by

Katelyn Marie Erickson

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Nanoscale Science

Charlotte

2018

Approved by:

_____
Dr. Jerry Troutman

_____
Dr. Juan Vivero-Escoto

_____
Dr. Craig Ogle

_____
Dr. Kirill Afonin

_____
Dr. Christine Richardson

ABSTRACT

KATELYN MARIE ERICKSON. Analysis and development of detection and targeting strategies for bacterial sugars. (Under the direction of DR. JERRY TROUTMAN)

Bacterial cell surface exopolysaccharides aid in protecting microbes from the defense mechanisms of the host, and differ substantially in structure between species. Exploiting these structural differences would be advantageous to target certain types of bacteria over others for organism selective antimicrobial agents or reagents to detect specific bacteria in the environment. Using the N-linked oligosaccharide from *Campylobacter jejuni* as a model system, I have developed a method for producing the oligosaccharide repeat unit of a bacterial surface polysaccharide. Using this system, I will now develop a new way to immobilize bacterial sugars for the isolation of agents that specifically interact with the cell surface of bacteria. Not only was this completed for the glycan from *C. jejuni,* but also for the newly elucidated colanic acid tetrasaccharide from *E. coli.* Alongside the developed immobilization platform, our lab has acquired mutant strains of *E. coli,* one that contains the N-linked oligosaccharide, and one that does not. Utilizing the difference between the *E. coli* strains, I have selected for agents specific to the *C. jejuni* polymer in a whole cell experiment. The selection of the whole cell and immobilized agents have been completed for the two glycans of interest and for two strains of *E. coli* and visualized by gel electrophoresis, flow cytometry and plate reader analysis. For the immobilization method, synthesis of an analogue with a different chemical handle has been completed, followed by successful building of two different

sugar repeating units onto the new analogue, as seen by HPLC analysis and ESI-MS analysis.

DEDICATION

I would like to dedicate my dissertation to my late grandfather, Dr. Robert Erickson. Without his inspiration I would not have followed in his footsteps in the sciences. I would not have chosen the path towards a graduate degree that I find myself in now. Without his general interest in my studies, along with encouraging magazine clippings about "women in science," I would not have this additional drive to pursue a higher degree of study in the sciences.

ACKNOWLEDGEMENTS

I would like to acknowledge current and past members of the Troutman Lab, specifically Amanda Reid (for all of her amazing synthesis skills with the azide materials and plate reader analysis knowledge) and Beth Scarbrough (with her LIC skills and UppP production) who have joined me in the PhD program and in this lab as the "PhD Trio of mass production" and have always been there for me personally and professionally. I would like to thank Tiffany Williams and Colleen Eade for being the microbiology experts that I needed in the lab (BFP bacteria and P1 phage transduction, and Colleen's amazing editing) - I just wish I had more time to spend with the both of you other than just the end of my PhD adventure. I would also like to thank Phillip Scott for all of his work on the colanic acid pathway and my chance to get into that project to finish what he had begun. I would like to thank Claire Gates for being one of the reasons to smile every day because of her infectious personality. In addition to my lab support I would also like to thank David Lee for putting up with my crazy hours and busy weekends, and emotionally supporting me throughout the years during this difficult task (of obtaining a PhD). And of course I would like to thank Jay Troutman for giving me a chance to be in his lab with little previous biochemistry experience, for dealing with me for so long (Masters and PhD), and allowing me the freedom to follow my desires to hone my characterization and instrumentation skills. Without all of your support I wouldn't have figured out what I wanted to do now or later in science (and in life).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

E. coli	*Escherichia coli*
BPP	bactoprenyl diphosphate
BP	bactoprenyl phosphate
cef	cell envelope fraction
CPS	capsular polysaccharide
ESI-MS	electrospray ionization mass spectrometry
HPLC	high performance liquid chromatography
GlcNAc	*N*-acetylglucosamine
LB	lysogeny broth
NMR	nuclear magnetic resonance spectroscopy
PCR	polymerase chain reaction
P. aeruginosa	*Pseudomonas aeruginosa*
S. aureus	*Staphylococcus aureus*
UDP	uridine diphosphate
CPS2E	a hexose-1-phosphate transferase
WcaJ	a hexose-1-phosphate transferase
WcaI	fucose transferase

WcaF	acetyltransferase
WcaE	fucose transferase
WcaC	galactose transferase
WcaA	glucuronic acid transferase
WcaB	acetyltransferase
WcaL	a proposed galactose transferase
WcaK	a proposed pyruval transferase
2CN-BP	2-cyanoanilino bactoprenyl monophosphate
Az-BP	3-azidoanilino bactoprenyl monophosphate

CHAPTER 1: INTRODUCTION

Short synopsis

The primary goal of my dissertation research was to develop a new method for the detection and potential antibiotic targeting of bacteria by exploiting unique bacterial glycan fingerprints on the surface of these organisms. To do this I have helped develop and apply a new system for the interrogation of bacterial systems involved in the production of complex bacterial surface glycans. Using these chemical tools, unique to the Troutman group, I have prepared new agents that selectively interact with bacterial glycans and demonstrated their ability to interact with whole cells. In this introduction I will provide a brief overview of the specific organisms that my efforts have been focused on. I will also highlight the specific carbohydrate-based features of the cell surface of these organisms, and why they could be effective targets. I will also provide an overview of how bacteria make their cell surface glycans and how that relates to our methods for targeting them. Finally, I will provide a brief overview of current methods for the detection of specific sugars on the surface of bacteria to provide a background on the state of the art of bacterial glycoscience.

Primary organism targets of this dissertation

Escherichia coli

In this dissertation my efforts have focused primarily on two important microorganisms. The first is *Escherichia coli*, a representative member of the *Enterobacteriacae* family, which includes additional dangerous pathogens from the genus *Salmonella* and *Klebsiella*. *E. coli* was chosen because of its well-defined genome and simplicity for genetic manipulation, in addition to its well documented role in food-borne illnesses. Virulent strains of *Escherichia coli* are responsible for as many as 70,000 infections, 2,000 hospitalizations and 60 deaths every year in the United States alone.[1-2] Pathogenic strains of the bacteria can exist harmlessly in the natural microbiome of the bovine digestive system. Therefore, it is not surprising that many virulent *E. coli* outbreaks are directly linked to cattle or indirectly through the crops that come into contact with their feces.[3]

Campylobacter jejuni

The second microorganism of focus is that of *Campylobacter jejuni,* which is known for its sporadic transmission and whose infection is one of the most common causes of gastroenteritis in the world, and is commonly contracted from the handling of poultry.[4-6] Generally, *C. jejuni* coincides with another microbe, *Salmonella enterica,* in providing the source of contamination in the meat industry. *C. jejuni* was chosen in this work because it has a well-defined genome and its lifecycle has been widely studied. Overall, *C. jejuni* is known for having a collection of factors and pathways that are not normally associated together in pathogens, that aids in its establishment of infection in

the host.[7] Because of the varied infection pathways in *C. jejuni*, the variation provides more targets to impact the lifecycle of this bacteria.

Glycans are critical and unique cell surface features in bacteriology.

Bacterial Sugars

Bacterial glycans are central to microbiology and are extraordinarily diverse relative to mammalian and most other eukaryotic systems. Eukaryotic glycans are primarily based on seven sugar residues (Man, Glc, GlcNAc, GalNAc, Gal, NeuAc, and Fuc, **Figure 1.1**).[8] Due to the ability to assemble in branched and linear chains and the availability of multiple hydroxyl groups for conjugation, even this limited number of sugars can be assembled into a wide set of potential structures. The variety of potential linkages holds true for bacterial sugars, but by comparison, bacteria use a wide array of sugars not available in most eukaryotic systems. The unusual sugars include rare carbohydrates like L and D quinovosamine, N-acetyl-rhamnosamine, just to name a few. In addition, bacterial sugars can have various modifications not common in eukaryotes including additional acetyl groups, pyruvate, amines, and other functionalities. Therefore, bacteria not only have the ability to 'mimic' the mammalian sugar code with the use of the seven sugar building blocks mentioned, they also are able to modify those sugars.

Modification of the standard mammalian bacterial sugars can provide additional immune responses or even result in total evasion of immune response altogether using mammalian sugar mimicry.[8] Although it seems counterintuitive to provoke an immune response, research suggests that the benefit of generating a large variety of sugar combinations overcomes the cost of being known to the host immune system.[9] This flexibility is

advantageous to the bacteria because of the significant adaptation to the environment and stress that it has to undergo during its lifetime. Furthermore, these sugar modifications also aid the bacteria in antimicrobial resistance due to the ever-changing bacterial defenses.[10] However, the complexity of the bacterial sugar combinations and modifications only further hinders the ability of researchers to determine how the bacteria thrive and create this library of diverse sugar products. Studying the mammalian sugar code is feasible in the sense of sugar types, but due to the diverse bacterial sugar modifications the elucidation of the bacterial code has been minimal. Understanding the complex bacterial sugar code in identity, function, and impact on health, provides the information needed to combat the bacterial antibiotic resistance problem that occurs today.

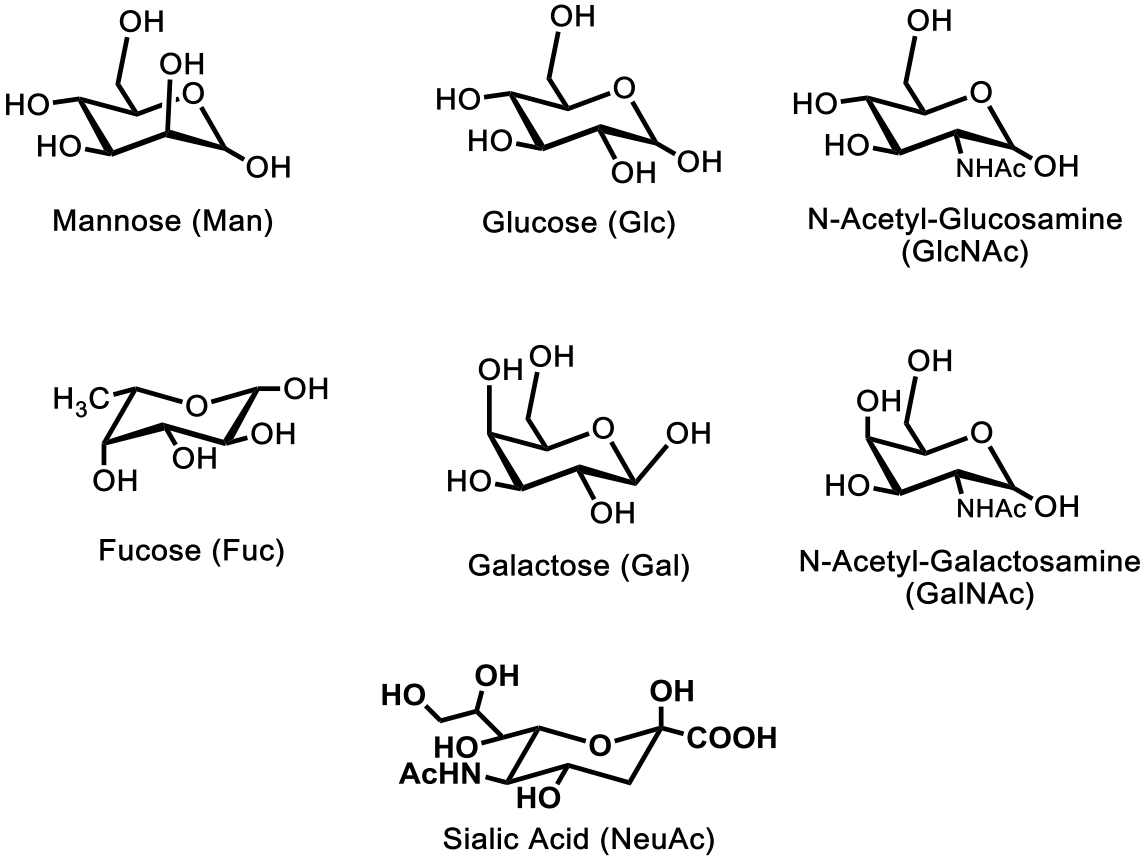


Figure 1.1 The seven mammalian sugar code structures. The seven basic mammalian sugars include mannose (Man), glucose (Glc), N-acetyl-glucosamine (GlcNAc), fucose (Fuc), galactose (Gal), N-acetyl-galactosamine (GalNAc) and sialic acid (NeuAc).

Bacterial polysaccharides

Polysaccharides are complex sugar polymers central to the biology of numerous microorganisms.[11] These materials are broadly defined here as carbohydrate polymers of multi-sugar oligosaccharide repeating units that can be separated into two categories. The first category is the capsular polysaccharides, which are carbohydrate materials held tightly

to the cell surface of an organism and are thought to play important roles in the infection of mammalian hosts.[12] The second type is the secreted exopolysaccharide, which is a material exported out of the cell into the environment, and thought to play an important role in biofilm formation and survival in specific environmental niches.[13] The structures associated with bacterial polysaccharides are often unique to the organism species, and often even the specific strain of that species. The sheer diversity of the sugar structures is evident in the fact that *Escherichia coli* has 80 capsular serotypes.[14] One polysaccharide important for *E. coli* is colanic acid, which is an exopolysaccharide made up of a repeating oligosaccharide consisting of glucose, galactose, pyruvylated galactose, glucuronic acid acetylated fucose, and fucose (**Figure 1.2b).** Colanic acid was first isolated and named by Dr. Goebel in 1963,[15] and is thought to provide a protective barrier for *E. coli* that allows the microbe to withstand low pH.[16-18] Unlike many capsular materials this particular polysaccharide, also known as M-antigen, is found associated with several members of the *Enterobacteriaceae* family, making it an important broad target for detecting the presence of this family of organisms in food or environmental samples. It would also potentially make a good target for selectively targeting antimicrobial agents to pathogenic organisms from the *Enterobacteriaceae* family.

Protein glycosylation

Complex bacterial surface glycans with unique structures are not restricted to polymeric materials. In addition to polysaccharides, cell surface proteins can be modified with a range of different oligosaccharide units. These modifications are typically O-linked where serine (Ser) or threonine (Thr) residues in a protein are linked to a particular oligosaccharide.[19-20]

However, another system of interest is that of the N-linked oligosaccharide, which is a less common modification of protein asparagine (Asn) residues located within a well-defined consensus motif Asn-X-Ser.[21] Up until the early 2000s N-linked protein glycosylation was thought to be unique to eukaryotes. Szymanski and co-workers noted the first bacterial N-linked glycosylation system in *Campylobacter jejuni* in 2002.[22] Since then several other eubacteria have been discovered that form N-linked glycans.[23]

The *C. jejuni* N-linked glycan is made up of a single heptasaccharide that includes the rare bacterial sugar di-N-acetyl-bacillosamine, five N-acetylgalactosamines, and a glucose (**Figure 1.2a**). The N-linked glycan is transferred to some proteins destined for localization to the bacterial outer membrane and is thought to maintain mammalian host cell interactions.[24-25] This maintained interaction leads to the infection of that host, and therefore the N-linked glycosylation of *C. jejuni* bacteria aids in the fitness of that bacteria.[4] Therefore, study of this unique method of bacterial glycosylation aids in the understanding of the bacterial lifecycle.

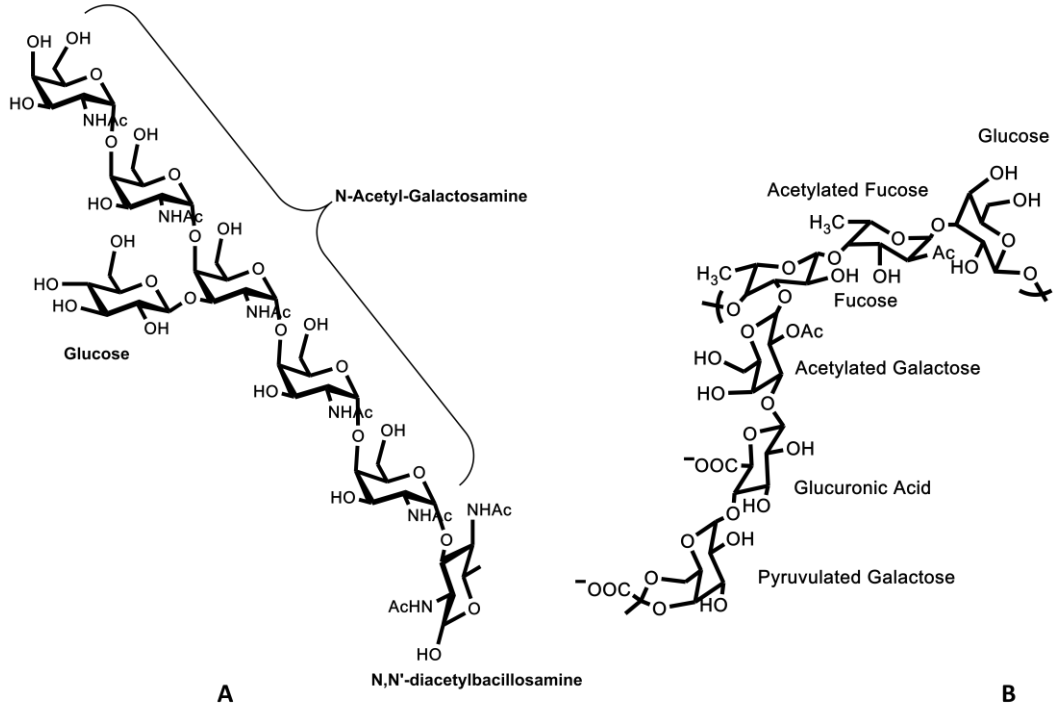


Figure 1.2. N-linked oligosaccharide and the colanic acid repeating unit. The N-linked oligosaccharide (A) is composed of five N-acetyl galactosamines, one diacetylBacillosamine and one glucose. The CA repeating unit (B) consists of six sugars; one glucose, two fucoses (one is acetylated), two galactoses (one is acetylated and one is pyruvylated), and a glucuronic acid.

The production of bacterial sugars and the bactoprenyl that links them.

Bacterial glycan biosynthesis

Even though there is extraordinary diversity in individual sugars, as well as overall bacterial glycan structure, a common pathway is utilized for the assembly of complex glycans in prokaryotes. Interestingly, the bacterial system for assembly of these structures is similar to that found in eukaryotes where each sugar transfer step occurs on a membrane-anchored isoprenoid, and sugars are appended one at a time from activated nucleotide-

linked sugars.[26] The prokaryotic isoprenoid is a C55 material called bactoprenyl diphosphate[27-28] (BPP, **Figure 1.3**) and the eukaryotic membrane anchor is a variable length dolichyl diphosphate.[29]

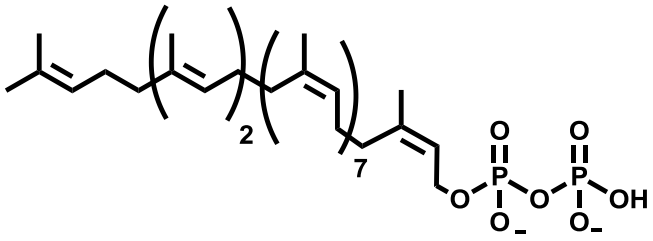


Figure 1.3. Bactoprenyl diphosphate (BPP) structure.

There are three main categories of enzymes in the polysaccharide biosynthesis pathways: the sugar modifying enzymes; a phosphoglycosyltransferase, which initiates assembly on a bactoprenyl phosphate (BP); and glycolsyltransferases, or the enzymes that transfer each of the remaining individual sugars (**Figure 1.4**). The sugar modifying enzymes catalyze changes, such as dehydrogenations, epimerizations, isomerizations and other alterations, to sugars. The first initiating step in assembling the oligosaccharide unit is the transfer of a sugar phosphate from a nucleoside diphosphate linked sugar to the membrane bound BP.[30] After formation of the BPP-monosaccharide additional sugars are added sequentially by highly specific glycosyltransferases.[31] Once the repeating unit of a polysaccharide, or the oligosaccharide to be transferred to protein, is completed it is flipped from facing the interior of the cell to facing the periplasm and is then either polymerized or transferred to

a protein. Polymers and glycosylated proteins then have unique transport systems associated with them for transfer to the cell surface or export into the environment.[32]

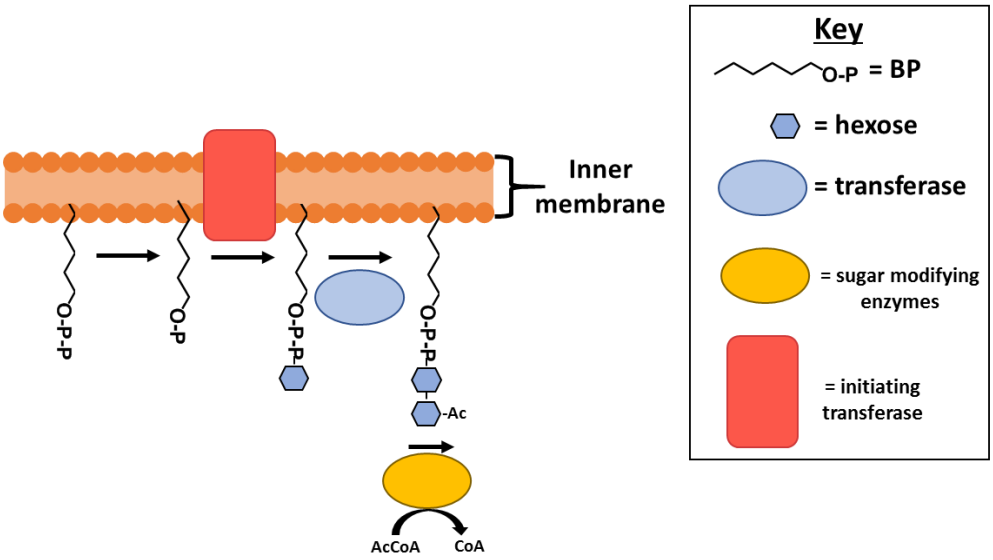


Figure 1.4. General scheme for types of enzymes in biosynthetic pathways. The types of enzymes associated with the biosynthetic pathways are the initiating transferase, otherwise known as the phosphoglycosyltransferase (red square), the glycosyltransferases that transfer the sugars onto the BP (blue oval), and the sugar modifying enzymes that make additions or modifications (e.g. acetylations, aminations, pyruvylations, etc.; yellow oval) to the sugars throughout the pathway.

Commonly in bacteria proteins involved in the assembly of a particular material are expressed from a single genetic locus in the bacterial chromosome. The genes for colanic acid are part of a regulon for responding to alterations in the environment, such as damage to integral parts of the cell, osmotic shock to the system, as well as the bacteria's growth on surfaces.[30] Even though colanic acid has been well studied, not much is known about the exact role of each enzyme in the gene locus. It is known what set of genes are responsible for the production of colanic acid by mutation studies,[33] but the role of each enzyme has not been elucidated. The methods developed for this dissertation to design materials that could selectively detect complex glycans, like colanic acid, were dependent on our ability to first elucidate the role of these key genes in the production of the polysaccharide.

As will be discussed later in this dissertation there was significant risk in developing reagents to target an exopolysaccharide like colanic acid. In addition, colanic acid is present in a wide variety of species and did not provide me with an effective way to determine whether I could design material selective for only one particular species in the environment. As an alternative, I have focused on the N-linked oligosaccharide from *Campylobacter jejuni*. The genes responsible for *C. jejuni* N-linked glycan production have been previously defined, and the production of the glycan has been reconstituted *in vitro*.[34-36] The N-linked glycan is unique to *C. jejuni* and a system to target the organism would have important benefits.

Current Methods of Detection for Bacterial Sugars

While the biological role and therapeutic potential of bacterial sugars has been well established for years, exploitation of this knowledge to develop therapeutics has been limited.[37-38] The major source for capsular sugars has been the isolation of these materials from the surface of cultured bacteria.[39-40] Procedures to isolate them are similar to the methods for specific protein isolation and purification before the establishment of tags and of other molecular biology tools commonly utilized in protein purification, where the methods for isolation of every polysaccharide can differ. Purity of these polysaccharides is also questionable because they are often contaminated with lipopolysaccharides, other surface capsules associated with the bacterium of interest, and because they exhibit impurities in chemical composition potentially induced by the purification procedure itself.[41] Identification of the correct polysaccharide product is also dependent upon having enough material for nuclear magnetic resonance (NMR) analysis and structural characterization. Alternatively, some antibodies and proteins have been developed or isolated to detect particular capsules, but the ability to detect a number of different sugars, along with rare bacterial sugars, is limited.[5-6, 42] With the use of a bactoprenyl anchor analogue that will be described in this work, the development of aptamers for a specific sugar(s) is possible, allowing for more specific detection of bacterial surface components. Along with the ability to detect the sugar components, recognition of bacterial sugars will lead to developments in targeting specific bacterial surfaces for therapeutic applications.

Isolation and Conventional Characterization

In order to study the bacterial sugar coating, the bacteria first have to be grown in abundance and the coating has to be separated from the rest of the bacteria and other polysaccharides present on that bacteria. Generally, the outer bacterial material is centrifuged (fractionated) and extracted in various organic solvents.[40, 43-44] From there, the use of traditional analytical techniques like NMR and mass spectrometry (MS) is possible, if enough material is obtained for such analyses. The aforementioned isolation methodology, however, can modify the original structure of the sugars (e.g. hydrolysis, link breakage, etc.) and not provide a clear depiction of the natural structure. Additionally, this method depends on the natural production of this material from bacteria, which can be an issue for low producing polysaccharides. Furthermore, the purity of the polysaccharide is questionable because of the inherent similarities between the sugar polymers being isolated and the non-specific isolation methods, which generally result in the mixtures of different polysaccharides in the final isolated component.[44] Overall, the use of centrifugation and chemical methods are laborious and may not result in pure material for the study of bacterial polysaccharides.

Glycan-Based Antibodies

Antibodies are blood-based proteins that are produced in response to an "invader" such as bacteria, viruses, or other foreign entities found in blood. Because antibodies are produced via an immune response in the blood of an animal, obtaining the antibodies is time consuming and labor intensive.[45-46] Therefore, not only is a target needed that produces an immune response, there is an additional cost associated with the animals, the

maintenance of the animals, and the final isolation and purification of the antibodies from the animal. These proteins are utilized in a variety of applications, from diagnostics of specific biological markers in medicine to determining if a histidine tag is present on a recombinant protein in the lab. Although antibodies have a plethora of targets, the use of antibodies for carbohydrates are limited.[47] Currently, a glycan-based antibody is used in determining the presence of a neuroblastoma by association with a disialganglioside on the tumor cell.[48-50] As evidenced by this example, the use of antibodies for glycans are directed to mammalian-based diseases and very few are being utilized against bacterial targets, leaving this area of research wide open. One challenge associated with the production of antibodies for bacterial glycans is the ability to produce enough immunogenic material to inject into animals. Common techniques used to produce, identify, and monitor proteins are not applicable to carbohydrates. However, due to the inability to directly detect glycans in complex biological samples, studies tend to focus on the use of antibodies to monitor carbohydrate expression.[47] Although antibodies are widespread and growing in use throughout the scientific and medical community, it has still yet to be explored for the bacterial problem for use in diagnosis of infections or targeting of those infections.

Lectins

Lectins are proteins that bind to carbohydrates. The main function of lectins is to aid in binding of bacteria to host cells, and is a popular target for disruption of the interactions between bacteria and mammalian cells.[51-52] As a result of this functionality, lectins are naturally designed to seek out the more common sugars to mammalian cells (mannose, galactose, etc.), therefore making lectins quite redundant in nature. Although lectins are

plentiful in any given biological setting (i.e. bacterial cells, mammalian cells, etc.), the variety of sugar specificity and generally the binding of bacterial lectins are low (in the micromolar range).[53-55] The most characterized bacterial lectins are the mannose-specific type-1 fimbriae, the galabiose-specific P fimbriae, and the N-acetylglucosamine-binding-F17 fimbriae.[53] Each of these lectins are specific for the common mammalian sugars without the ability to bind to modified and rarer bacterial sugars. In view of this knowledge, the use of lectins to target bacteria are not ideal and would only provide a limited cover of the vast sugar library that bacteria can produce.

Click-Based Methods

Bringing the fields of biology and chemistry together, the use of chemically synthesized tags have become a way to determine the use or production of certain sugar molecules. With chemical modifications, such as with an azide addition, it is feasible to conjugate fluorophores for aid in the detection of these otherwise undetectable sugars. Bertozzi and coworkers have pioneered the use of 'click' based methods for the detection of glycans and other biomolecules.[56-58] In order to utilize the click methodology, however, the click enabled molecules need to be chemically synthesized, which can be long and tedious, and specific to a certain pathway of interest (i.e. at least one of the sugars is known). Alternatively, the use of a click-enabled molecule that is constant throughout a variety of sugar pathways is ideal and is discussed in our work.

Sugar Based Aptamers

On the exterior of many bacteria there are sugar polymers that aid in the virulence of that bacteria that have yet to be utilized as a direct target for bacterial detection. Sialic acid

sugars, with the most abundant being N-acetylneuraminic acid (Neu5Ac), are widely studied due to the overexpression of this sugar on cancer cells, and are a popular target explored within cancer research.[59-60] Development of an aptamer for the Neu5Ac sugar has been done using systematic evolution of ligands by exponential enrichment (SELEX) with RNA, and has a nanomolar binding affinity. [61] Access to the RNA aptamer allows for the targeting of the overexpressed sugar on the outside of cancer cells, as well as the use in biosensors for sialic acid. DNA or RNA aptamers have been in use for years, but the application to bacterial sugars is limited. A study of DNA aptamers for cellulose, which is composed of cellobiose sugar chains, established the foundation of sugar specific DNA aptamers, which are suggested to bind more tightly than that of sugar-specific lectin proteins.[62] Another study found DNA aptamers to sugar-containing antibiotics, such as kanamycin, neomycin, and gentamicin.[63] Although DNA aptamers have been developed for sugar-containing molecules, aptamers to bacteria specific sugars have yet to be studied. Since the bacterial sugar coating accounts for the stability and in some cases, the virulence of many bacteria. Targeting of the sugars specific to bacteria would prove beneficial to destruction of the protective layer that aids in the bacterial infection, as well as detection of bacteria containing those rare sugars. Therefore, this work will discuss the design and production of a platform to obtain sugar specific aptamers for future targeting and detection applications.

Unique chemical probes for bacterial glycan assembly systems

Nearly all bacterial glycans share the bactoprenyl diphosphate (BPP) anchor molecule that links them to the bacterial surface. The BPP anchor is synthesized from eight

condensation reactions of farnesyl diphosphate (FPP) and isopentenyl diphosphate (IPP) catalyzed by the enzyme undecaprenyl pyrophosphate synthase (UPPS).[64] Through an initiating hexose-1-transferase, a linkage is formed between the phosphate of the isoprenoid anchor and the sugar phosphate of nucleotide-linked sugar to form a phosphoanhydride linked B-PP-sugar.[65-67] The glycolsyltransferases then take the BPP-linked sugar and add another sugar, or sugars, to the growing monomer, that is eventually polymerized into the full sugar polymer. Originally, the study of sugar-isoprenoid chemistry was completed with the use of radiolabeled sugars to determine the linkage of a sugar to the bactoprenyl anchor.[34, 68-69] Recently, the development of chemically synthesized fluorescent analogues in the Troutman lab allows for tracking of the sugar attachments to the anchor molecule with high performance liquid chromatography (HPLC) analysis, where a change in retention time of a fluorescent peak correlates to the sugar addition.[70] In order to utilize the fluorescent analogue in determination of the sugar addition, the elongation of the analogue by the previously described UPPS enzyme is needed before sugar attachment can occur. The ability to track the sugar addition is vital in both polysaccharide biosynthesis and the production of the materials described within this proposal. The UPPS enzymatic elongation process will be utilized for a novel FPP analogue that will be further discussed in this research scheme.

Research Approach

To develop selective agents for bacterial sugars, the glycans first need to be synthetically built. The roles of the biosynthetic enzymes for *C. jejuni* have been known for over ten years.[5-6, 34-36, 71] In contrast to the *C. jejuni* pathway, the functions of the *E. coli* colanic acid

biosynthetic pathway enzymes are less clear. In Chapter 2, I describe the functions of the colanic acid biosynthetic enzymes to produce a majority of the sugar repeat unit. In Chapter 3, I describe the use of a new immobilization platform that the glycans of both *C. jejuni* and *E. coli* were built on. An anchor molecule similar to the natural isoprenoid used within capsular biosynthesis was functionalized with an azide moiety. Because of the azide moiety, immobilization of the sugar target is feasible. In Chapter 4, I describe the work of selecting for agents specific to the glycans of both of these organisms. With the combination of capsular biosynthesis sugar-modifying enzymes, sugar isoprenoids, as well as sugar-based target aptamer development; the design of a novel foundation for the procurement of sugar specific aptamers for both the *C. jejuni* N-linked oligosaccharide and the colanic acid exopolysaccharide from *E. coli* will be discussed in this work. The known enzymes for the N-linked oligosaccharide biosynthesis as well as the recently elucidated functions of the biosynthetic enzymes for colanic acid provided the bacterial sugar targets for aptamer development.

CHAPTER 2: COLANIC ACID BIOSYNTHESIS

Note: The preliminary data including cloning, initial protein expression, and some HPLC analysis obtained for this chapter was first completed by P. Scott, UNCC Chemistry Master's 2015. The LIC vector was produced by Beth Scarbrough, Nanoscale PhD 2022. The fluorescent analogue precursor was kindly synthesize by Amanda Reid, Nanoscale PhD 2020. This chapter has been submitted for publication in the Journal of Biological Chemistry with P. Scott and K. Erickson as co-first authors. All data shown was acquired and analyzed by K. Erickson unless otherwise noted.

Short Synopsis

The primary goal of Chapter 2 is the elucidation and *in vitro* synthesis of the colanic acid exopolysaccharide repeat unit from *E. coli*. In this chapter, I will provide an overview of the organism *E. coli* and the importance of the colanic acid polymer to this organism. Secondly, I will provide the context for the construction of these complex sugar glycans inside and outside of the bacterial cell. Thirdly, I will explain the use of an anchor molecule utilized in the building of these complex sugars for ease of tracking through the biosynthetic pathway. I will also demonstrate the elucidation of enzyme function in the biosynthetic pathway for colanic acid by way of HPLC and ESI-MS analysis. Finally, I will discuss the impact of building these complex glycans *in vitro* for the study and use of these materials downstream.

Introduction

As explained in the introduction of this dissertation, *E. coli* is an important microbe to study and has a significant impact on human health. More specifically, I have studied the colanic acid exopolysaccharide from *E. coli*. The CA polymer is comprised of hexasaccharide repeat units consisting of glucose, two fucoses, two galactoses, and glucuronic acid (**Figure 2.1**).[30, 72-73] In addition, one fucose and one galactose may be acetylated to varying degrees, and the terminal galactose residue is thought to be pyruvylated. There is considerable discrepancy in the literature on the level of acetylation of CA. Acetyl groups have been reported to be non-stoichiometric with the sugar polymer, and the repeat unit structure has been reported as mono-, di and tri-acetylated.[72] It is not clear when these acetylations occur in the biosynthesis of the polymer.

Interestingly, while there are three possible acetylations of CA known, only two predicted acetyltransferases (*wcaF, wcaB*) are encoded by the CA biosynthesis operon, previously shown to encode all proteins required for production in *E. coli*.[33] Other genes in the locus encode proteins homologous to an initiating glucose-1-phosphate transferase (*wcaJ*), five glycosyltransferases (*wcaI*, *wcaE*, *wcaC*, *wcaL*, *wcaA*), a presumed pyruvyltransferase (*wcaK*), a flippase (*wzx*), and a polymerase (*wzy*) as well as critical transport proteins.[33] Generally, it is thought that CA polymerization is a Wzy-dependent pathway in which each sugar of the polymer repeat unit is added sequentially to the 55-carbon isoprenoid bactoprenyl phosphate (BP). BP anchors the biosynthesis of the glycan repeating unit to the cytoplasmic face of the bacterial inner membrane. Once the repeat unit is complete, the isoprenoid-linked oligosaccharide is translocated across the inner membrane to face the periplasm and is polymerized by Wzy.[30]

Like many polysaccharide biosynthesis pathways, little is known about the precise biochemical roles of the genes associated with CA production. The initiating glucose-1-phosphate transferase (WcaJ) has been biochemically identified and function has been reconstituted *in vitro*.[65, 74] However, which reaction is catalyzed by each of the glycosyltransferases is unknown. There is a severe lack of information on bactoprenyl-dependent biosynthesis pathways, which hampers the ability to predict the functional role of these types of enzymes based on sequence alone.

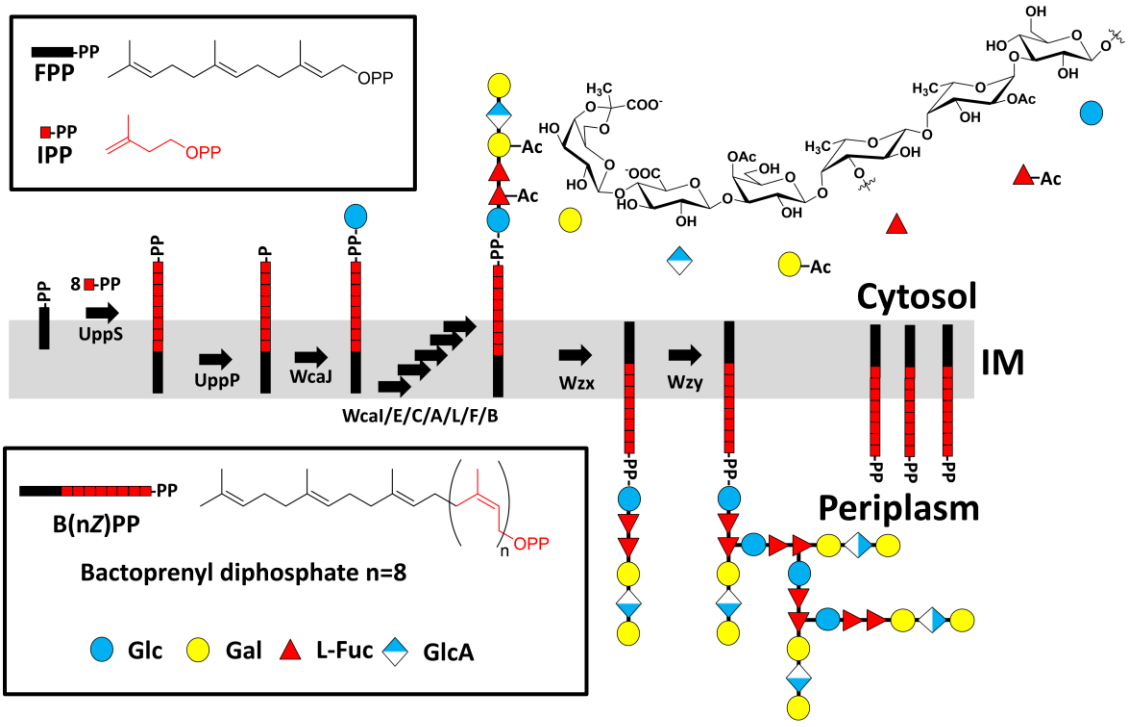


Figure 2.1. CA repeating unit biosynthesis. CA repeating unit consisting of six sugars; one glucose, two fucoses, two galactoses and a glucuronic acid. The repeating unit is assembled on a bactoprenyl phosphate anchor via a wzy-dependent pathway. Once the oligosaccharide repeat unit is formed it is flipped from facing the cytosol to the periplasm where it is polymerized and exported. This repeating unit is polymerized through connections between the glucose and third fucose sugar (wavy line) to form the CA exopolysaccharide.

Recently a new system was developed for characterizing the biosynthesis of complex bacterial polysaccharides that is based on a fluorescent reporter appended to the terminal isoprene of BP (**Figure 2.2**).[70, 75-76] A fluorescent 2-nitrileanilinobactoprenyl phosphate-BP (2CNA-BP) has been used to elucidate each step in the biosynthesis of Capsular Polysaccharide A in *Bacteroides fragilis*.[77] A similar probe was also functional in the early steps of N-linked oligosaccharide biosynthesis in *Campylobacter jejuni*.[78] In this chapter, I have applied this probe to map the early steps of CA production *in vitro*. In this work, the roles of the first four glycosyltransferases associated with the synthesis of the repeat unit have been identified, and function has been reconstituted *in vitro*. These studies highlight several important features of this biosynthesis pathway, including the role of, and centrality of, sugar acetylation in repeat unit production.

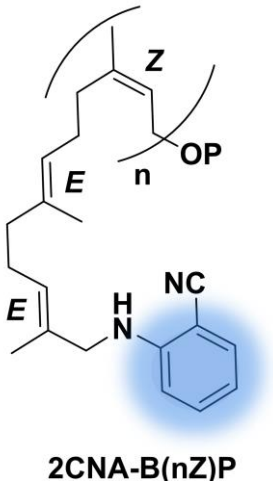


Figure 2.2. The fluorescent reporter 2-nitrileanilinobactoprenyl phosphate (2CNA-B(Z)P) that is appended to the terminal isoprene of BP, has been used to elucidate each step in the biosynthesis of Capsular Polysaccharide A in *Bacteroides fragilis*.

Results

Expression and isolation of CA biosynthesis proteins.

The primary objective was to identify the precise function of each glycosyltransferase (wcaI, wcaE, wcaC, wcaA, wcaL) and acetyltransferase (wcaF, wcaB) gene located in the CA biosynthesis operon. Using ligation independent cloning, each gene was inserted into a modified pET-24a vector encoding a C-terminal hexahistidine tag and overexpressed in C41 *E. coli* cells. Using immobilized metal affinity chromatography, proteins were purified to homogeneity (**Figure 2.3**) and quantified by UV absorbance. Consistent with the lack of predicted transmembrane domains in the proteins based on sequence analysis (**Figure 2.4**), all proteins were soluble and surfactant was not required for isolation. I was unable to obtain large quantities of WcaA and WcaL due to the proteins remaining with lysed cell membranes (**Figure 2.5**). I attempted to extract the proteins from cell envelope fractions (cef) with poor recovery. This data suggested that these proteins either have transmembrane regions that were not predicted, interact with membranes peripherally, or form insoluble inclusion bodies when overexpressed.

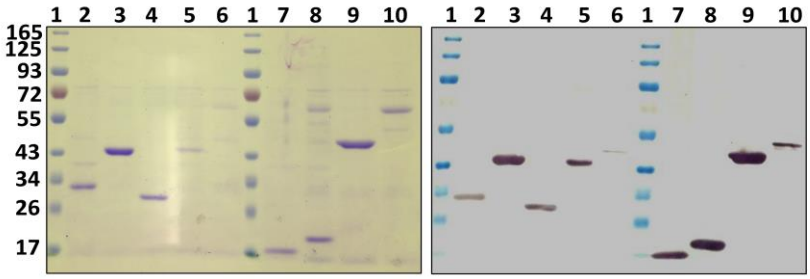


Figure 2.3. Colanic acid protein SDS-PAGE gel displaying the expected molecular weights and purities of all of the CA biosynthesis proteins consisting of the GDP-Fuc synthesis enzymes, the glycolsyltransferases and the sugar modifying enzymes (left panel). Anti-His western of the same enzymes displaying the hexahistidine tag associated with the enzymes from the cloning of the genes for CA biosynthesis (right panel). The lanes are: 1- Ladder (kDa), 2- WcaA (34 kDa). 3- WcaC (46 kDa), 4- WcaE (29 kDa), 5- WcaI (45 kDa), 6- WcaL (46 kDa), 7- WcaB (18 kDa), 8- WcaF (21 kDa), 9- WcaK (48 kDa), 10- WcaM (52 kDa). (Gel and Western analysis by P. Scott).

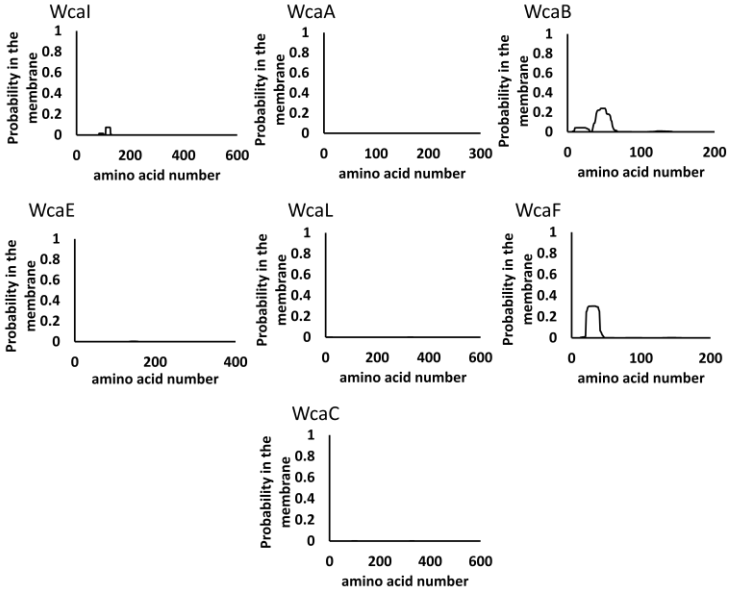


Figure 2.4. Topological prediction of transmembrane domains for the glycosyltransferases from the CA biosynthetic pathway. The only proteins to have any predicted membrane domains are WcaI/B/F.[79]

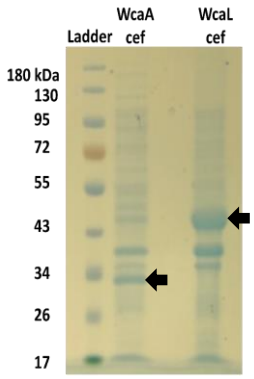


Figure 2.5. SDS-PAGE with Coomassie stain of the cef of cells expressing WcaA (33 kDa) or WcaL (45 kDa).

Assembly of a fluorescent bactoprenyl diphosphate-linked glucose

In previous work, Patel *et al.* demonstrated the *in vitro* activity of the initiating phosphoglycosyltransferase, WcaJ, which catalyzes the addition of glucose phosphate to BP.[74, 80] Another protein, Cps2E from *Streptococcus pneumoniae*, was shown by Cartee and coworkers to catalyze an identical reaction.[81] I chose to use Cps2E in this work to assemble the starting fluorescent 2CNA-BPP-glucose needed to assess the remaining proteins in the pathway. I chose this protein in lieu of WcaJ because it was readily available and was known to accept our 2CNA-BP as a substrate. Cps2E, like WcaJ, is a WbaP family protein with a similar predicted topological structure (**Figure 2.6**).[65, 81] Cps2E was overexpressed in *E. coli* BL-21 cells and membrane fractions were prepared with a vector encoding the protein and with an empty vector control. Using methods previously described, the fluorescent analogue 2-nitrileanilinobactoprenyl phosphate with five *Z*-configuration isoprene units (2CNA-B(5Z)P) was synthesized.[82] The 2CNA-B(5Z)P

configuration analogue was then utilized by Cps2E to provide glucose-linked product. HPLC analysis of the Cps2E reactions indicated the formation of a new product with decreased retention time, consistent with the addition of glucose phosphate to the 2CNA-B(5Z)P analogue. Optimization of the reaction conditions led to a method capable of consuming over 90% of the starting isoprenoid with Cps2E expressing cef, where no turnover was observed with cef from empty vector control cells (**Figure 2.7a**). To ensure that the product formed in the Cps2E reaction was 2CNA-B(5Z)PP-Glc I used electrospray ionization mass spectrometry (ESI -MS) in the negative ion mode to confirm hexose-phosphate addition (**Figure 2.7b**).

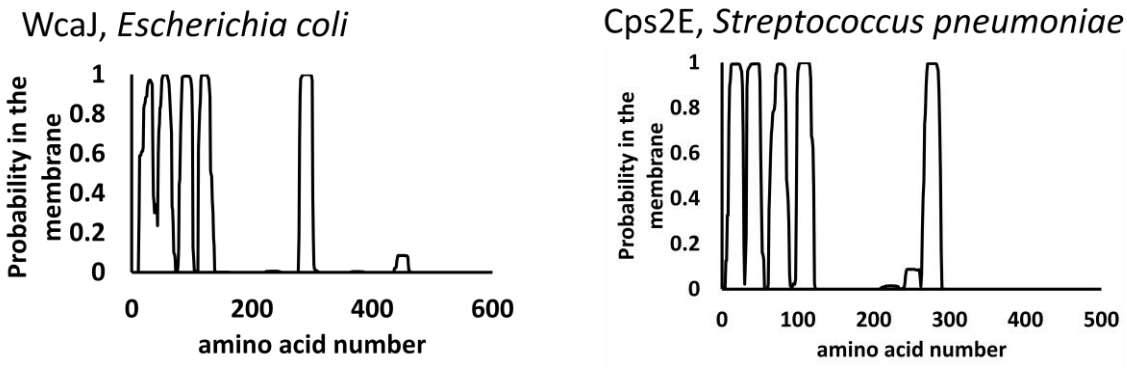


Figure 2.6. TMHMM topological predictions of the the initiating phosphoglycosyltransferases WcaJ compared to Cps2E. [79] Both WcaJ and Cps2E have five predicted transmembrane domains.

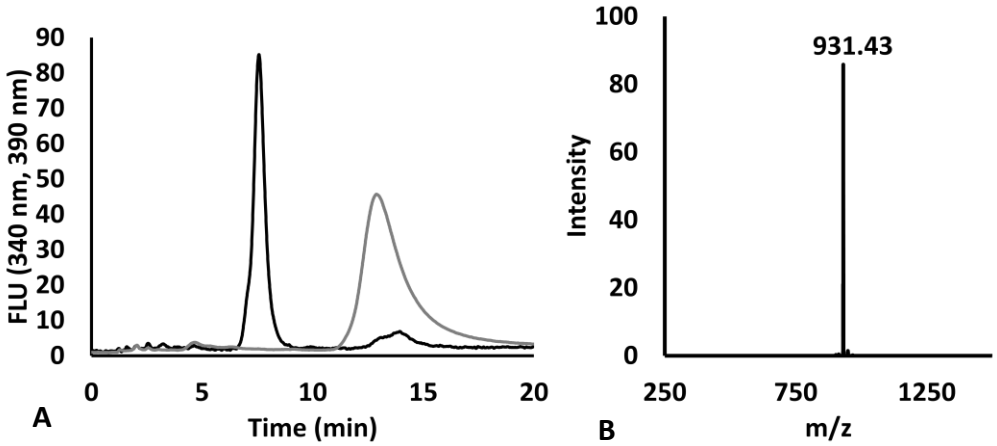


Figure 2.7. Cps2E reaction analyzed by reverse phase HPLC (a) with UDP-Glc and 2CNA-B(5Z)P showing 91.2% turnover (based on the area under the curve of the product compared to the 2CNA-B(5Z)P) with 100X the concentration of UDP-Glc compared to the 2CNA-B(5Z)P. The black line is the 2CNA-B(5Z)PP-Glc product and the gray line is the 2CNA-B(5Z)P starting material. The 2CNA-B(5Z)PP-Glc product was isolated and analyzed by ESI-MS (b) and the mass of 931.43 m/z (expected 931.46 m/z) was obtained. *WcaI transfers unmodified fucose to BPP-Glc*

There were five proposed glycosyltransferases (WcaA, WcaC, WcaE, WcaI, and WcaL) associated with the CA biosynthesis gene cluster that could potentially promote the formation of a disaccharide with BPP-Glc. To test this, I mixed isolated 2CNA-B(5Z)PP-Glc with each glycosyltransferase and GDP-Fuc, then analyzed for product formation by HPLC. Of the five glycosyltransferases, only incubation with WcaI altered the retention time of the substrate. The retention shift from 9.7 to 8.7 min was consistent with the addition of a single sugar moiety and suggested that WcaI was the next enzyme in the pathway (**Figure 2.8a**). The presumed fucosylated product was isolated then confirmed to be the disaccharide 2CNA-B(5Z)PP-Glc-Fuc by ESI-MS (**Figure 2.8b**).

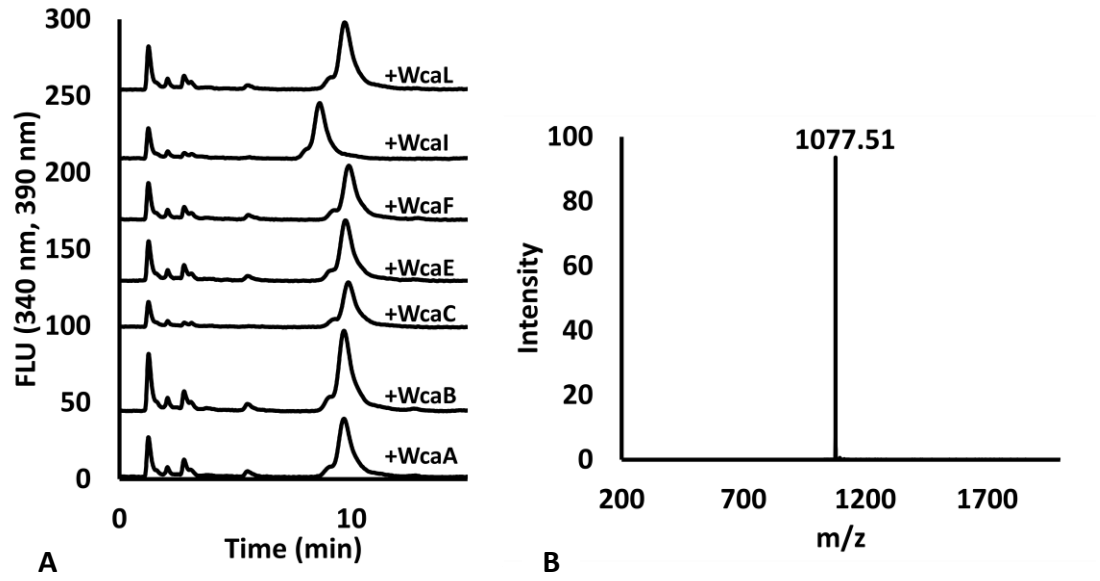


Figure 2.8. Reverse phase HPLC analysis of reactions with 2CNA-B(5Z)PP-Glc and each of the glycolsyltransferases WcaA, WcaC, WcaE, WcaI and WcaL along with the acetyltransferases WcaF and WcaB (a). The enzyme WcaI, in the presence of GDP-Fuc, modifies the 2CNA-B(5Z)PP-Glc . This product was then isolated and analyzed by ESI-MS (b) and found to have an m/z of 1077.51, which is consistent with 2CNA-B(5Z)PP-Glc-Fuc (expected 1077.52 m/z). *WcaF acetylates isoprenoid-linked disaccharide*

There was ambiguity in the next step in the biosynthesis of CA; the fucose residue in 2CNA-B(5Z)PP-Glc-Fuc, could either be acetylated then a second fucose added to form a trisaccharide, or the unmodified fucose could be the acceptor for the addition of a second fucose. To distinguish between these possibilities I prepared B(5Z)PP-Glc-Fuc then added each of the remaining glycosyltransferases and the acetyltransferases in separate reactions with additional GDP-Fuc or acetyl-CoA. I did not observe a shift in retention time with any of the glycosyltransferases under these conditions (**Figure 2.9a**), suggesting that the sugar may need to be acetylated prior to the next step in the pathway. I found that in the

presence of the acetyltransferase WcaF, and not WcaB, that there was a slight retention time change from 8.7 min to 8.8 min, which was consistent with the small increase in hydrophobicity associated with acetylation of the disaccharide (**Figure 2.9a**). A one pot reaction was then prepared with 2CNA-B(5Z)P, Cps2E, UDP-Glc, WcaI, GDP-Fuc, WcaF and acetyl-CoA. WcaF product was isolated and I confirmed the formation of mono-acetylated product by ESI-MS (**Figure 2.9b**). Neither WcaF nor WcaB had any influence on the retention of GDP-fucose alone (data not shown) suggesting that the fucose residue had to be associated with B(5Z)PP-Glc before modification.

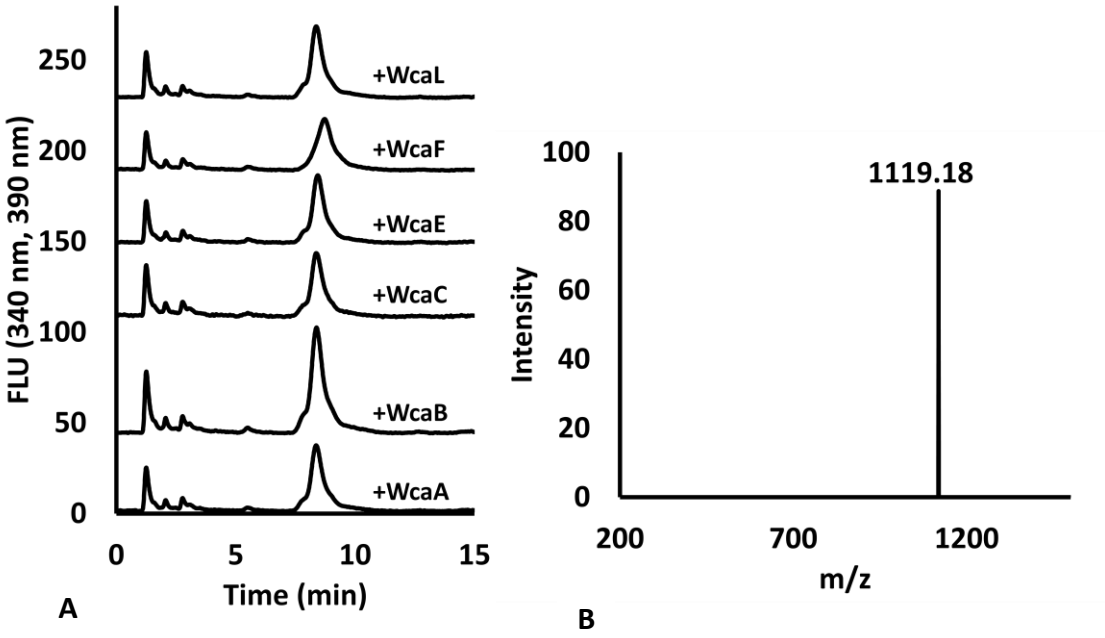


Figure 2.9. Reverse phase HPLC analysis of reactions with 2CNA-B(5Z)PP-Glc-Fuc and each of the glycolsyltransferases WcaA, WcaC, WcaE and WcaL along with the acetyltransferases WcaF and WcaB (a). The enzyme WcaF, in the presence of Acetyl-CoA, modifies 2CNA-B(5Z)PP-Glc-Fuc. This product was then isolated and analyzed by ESI-MS (b) and found to have an m/z of 1119.18 m/z, which is consistent with 2CNA-B(5Z)PP-Glc-AcFuc (expected 1119.53 m/z).

Acetylation is required for WcaE activity

Since WcaF was able to catalyze the transfer of the acetyl group to the B(5Z)PP-Glc-Fuc disaccharide, and none of the glycosyltransferases would utilize the un-acetylated material, it was likely that this modification was required before the formation of the trisaccharide. To test this, reactions were prepared with 2CNA-B(5Z)PP-Glc, WcaI, GDP-Fuc, WcaF, and acetyl-CoA, then were monitored for completion by HPLC. I then aliquoted the reactions into mixtures containing the remaining glycosyltransferase enzymes (WcaA, WcaC, WcaE, and WcaL). Only the addition of WcaE protein led to an 8.8 min to 7.6 min

shift in retention time from the WcaF product B(5Z)PP-Glc-AcFuc (**Figure 2.10a**). MS analysis of the isolated product from a one-pot 2CNA-B(5Z)P to WcaE reaction confirmed the formation of the acetylated trisaccharide 2CNA-B(5Z)PP-Glc-AcFuc-Fuc (**Figure 2.10b**).

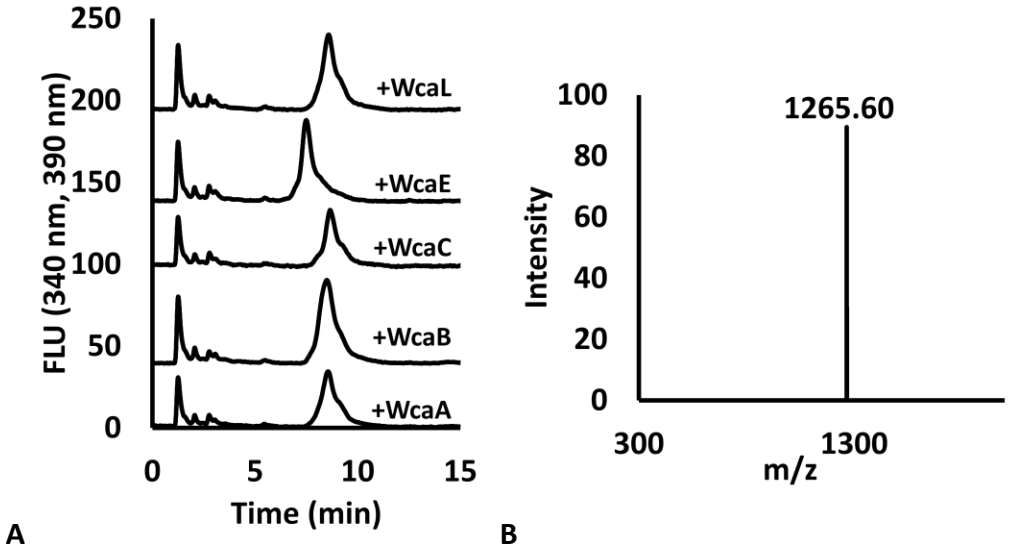


Figure 2.10. Reverse phase HPLC analysis of reactions with 2CNA-B(5Z)PP-Glc-AcFuc and each of the glycolsyltransferases WcaA, WcaC, WcaE and WcaL along with the acetyltransferase WcaB (a). This product was then isolated and analyzed by ESI-MS (b) and found to have an m/z of 1265.60 m/z, which is consistent with 2CNA-B(5Z)PP-Glc-AcFuc (expected 1265.59 m/z).

WcaC catalyzes the formation of an acetylated tetrasaccharide

With the formation of the trisaccharide by WcaE the next step was presumably the introduction of galactose into the repeat unit precursor. I prepared a reaction mixture containing 2CNA-B(5Z)PP-Glc with all components up to WcaE and again tested the

product of WcaE with the remaining glycosyltransferases (WcaA, WcaC, and WcaL) in separate reactions. Again, I found that only one of the remaining glycosyltransferases (WcaC) altered the retention of WcaE product from 7.6 min to 6.6 min (**Figure 2.11a**). Isolated WcaC product was then analyzed by ESI-MS and was found to have the expected mass of the 2CNA-B(5Z)PP-Glc-AcFuc-Fuc-Gal (**Figure 2.11b**).

WcaA adds glucuronic acid to the tetrasaccharide in the presence of the acetyltransferase WcaB

With the formation of the tetrasaccharide 2CNA-B(5Z)PP-Glc-AcFuc-Fuc-Gal by WcaC, the introduction of glucuronic acid into the repeat unit precursor was the expected next addition. I prepared 2CNA-B(5Z)PP-Glc-AcFuc-Fuc-Gal and tested it for turnover with WcaA or WcaL in the presence of UDP-GlcA. With solubilized protein, I observed no product formation. However, when I used the cef of WcaA, but not WcaL, I observed a retention time shift from 6.6 min to 4.3 min consistent with the addition of the more polar glucuronic acid sugar. Limited product formation was observed with WcaA cef alone (**Figure 2.12**) and I were unable to force the reaction to go to completion. Additionally, the use of an empty vector cef control (same plasmid and cells without the gene present) in C41 cells did not modify the 2CNA-B(5Z)PP-Glc-AcFuc-Fuc-Gal product (data not shown). WcaB was then tested for modification of 2CNA-B(5Z)PP-Glc-AcFuc-Fuc-Gal in the presence of acetyl-CoA. A small retention time shift was not readily observed that would be associated with the acetylation of the 2CNA-B(5Z)PP-Glc-AcFuc-Fuc-Gal product. Interestingly, when WcaA and UDP-GlcA were added to the WcaB reaction I observed complete conversion of 2CNA-B(5Z)PP-Glc-AcFuc-Fuc-Gal to form the

presumed di-acetylated pentasaccharide (**Figure 2.13a**). I next isolated WcaA/B product and analyzed by ESI-MS and found the expected mass of the 2CNA-B(5Z)PP-Glc-AcFuc-Fuc-AcGal-GlcA (**Figure 2.13b**). Due to the poor performance of solubilized WcaA there was concern that earlier tests of the protein would be invalid because the functional form of the protein appeared to be associated with the cef. I retested all of the prior reactions with cef preparations of both WcaA and WcaL and still found no activity with the small amounts of solubilized protein or cef preparations (data not shown).

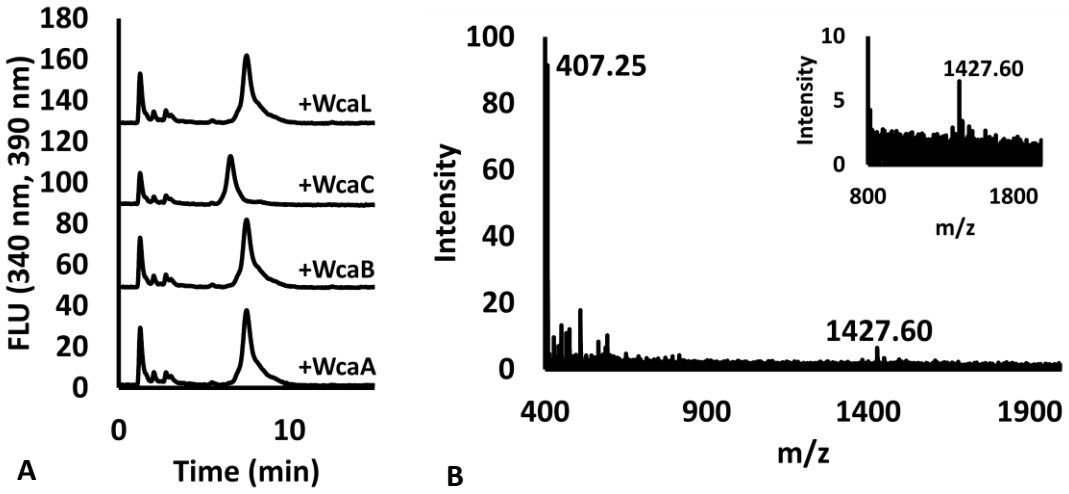


Figure 2.11. Reverse phase HPLC analysis with 2CNA-B(5Z)PP-Glc-AcFuc-Fuc and each of the glycolsyltransferases WcaA, WcaC and WcaL along with the acetyltransferase WcaB(a). The enzyme WcaC, in the presence of UDP-Gal, modifies 2CNA-B(5Z)PP-Glc-AcFuc-Fuc. This product was then isolated and analyzed by ESI-MS (b) and found to have an m/z of 1427.60 m/z, which is consistent with 2CNA-B(5Z)PP-Glc-AcFuc (expected 1427.64 m/z), the 407.25 m/z corresponds to sodium cholate, which is the detergent used in the reaction.

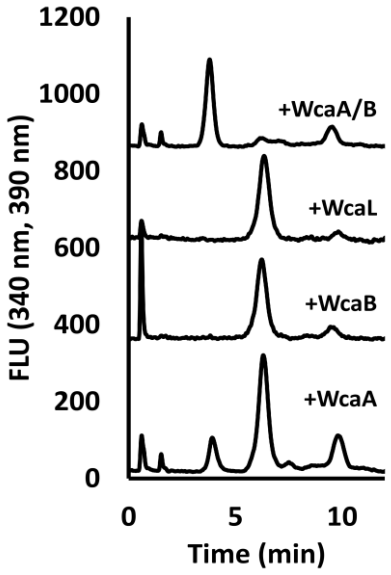


Figure 2.12. HPLC analysis of 2CNA5(Z)PP-Glc-AcFuc-Fuc-Gal, or WcaC product, with WcaA, WcaB, WcaL and the combination of WcaA and WcaB. Only 18% of the WcaC product can be converted with WcaA alone, but over 90% is converted with both WcaA and WcaB present. The peak around 10 minutes if leftover WcaI product from incomplete WcaF conversion.

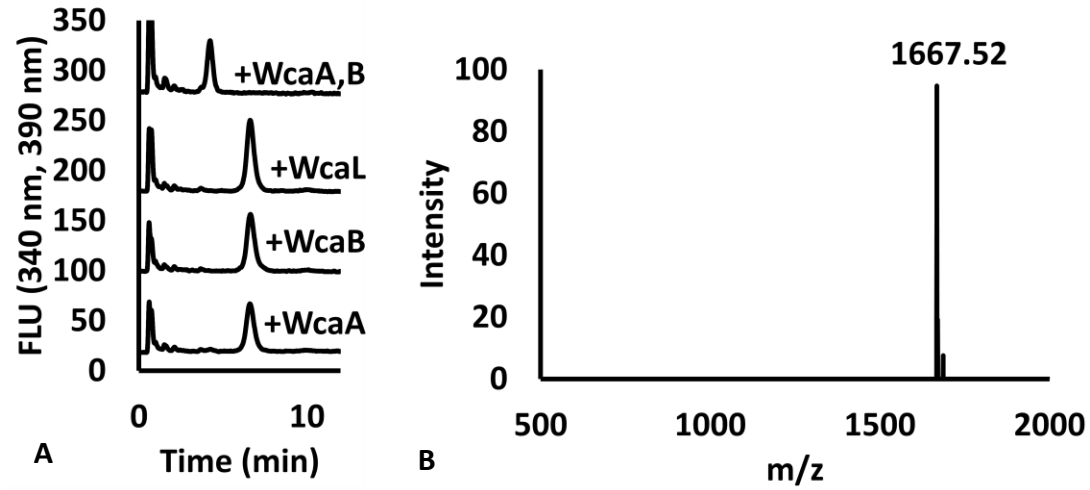


Figure 2.13. Reverse phase HPLC analysis of 2CNA-B(5Z)PP-Glc-AcFuc-Fuc-Gal and each of the glycolsyltransferases WcaA and WcaL along with the acetyltransferase WcaB (a). The enzyme WcaA and WcaB, in the presence of UDP-Glucuronic acid and Acetyl-CoA, modifies 2CNA-B(5Z)PP-Glc-AcFuc-Fuc-Gal. This product was then isolated and analyzed by ESI-MS (b) and found to have an m/z of 1667.52 m/z, which is consistent with 2CNA-B(5Z)PP-Glc-AcFuc (expected 1667.67 m/z +Na).

Discussion

The identification of the functional roles of proteins involved in complex polysaccharide biosynthesis in bacteria remains a major challenge. Radiolabel based methods are most often used, but when the target protein is involved in the incorporation of a second or third sugar, or a sugar modification such as acetylation, these methods become more complex as more of the radiolabel is incorporated. This report highlights the use of fluorescent polyisoprenoids as replacements for traditional radiolabel techniques. Importantly, this work has shown that these probes can be used for assessing the acetylation of sugars, in addition to glycosylation reactions.

Previous work has suggested that there are five glycosyltransferases associated with the production of CA and one initiating glucose-1-phosphate transferase (WcaJ).[65, 74] WcaI was predicted to catalyze the transfer of a fucose based on its proximity to the GDP-fucose biosynthesis genes in the gene cluster. [33] Our data clearly demonstrates that WcaI catalyzes the addition of the second sugar in CA biosynthesis, and no other protein in the CA biosynthesis gene cluster is capable of doing this.

It has been reported that CA has non-stoichiometric acetylations at the 2 and 3 positions of the disaccharide fucose residue.[72] Since this was non-stoichiometric, I initially thought that these modifications would not be required for the formation of the polymer repeat unit, and may occur after repeat unit formation or upon polymer formation. However, our data demonstrate that the activity of the acetyltransferase WcaF on the polyprenyl-linked disaccharide is critical to the formation of the repeat unit. Since WcaF was critical to the formation of the oligosaccharide repeat unit, it is potentially selective for either the 2 or the 3 position. It may be possible that incorporation of the acetyl group in either position could lead to formation of a WcaE competent substrate. Therefore, the non-stoichiometric nature of this modification may be due to the acetyl group being split between these two positions. Our ESI-MS data suggests only one modification per fucose, and I have no evidence of multiple acetylations at this stage of the pathway. However, it is possible that a second acetylation could take place after formation of the oligosaccharide or polymer. The importance of WcaF is consistent with studies showing that this protein is critical to the formation of CA *in vivo* and more recently, the importance of that gene for biofilm formation.[83-84] I have shown that WcaB, the other encoded acetyltransferase, had no influence on the disaccharide, but increases the ability of WcaA to add the fifth sugar onto

the CA repeating unit, and also appears to be a critical protein in the assembly of the oligosaccharide repeat unit.

Inconsistencies in the literature with respect to the structure of CA are problematic in assigning functional roles of proteins in these pathways.[15, 72-73, 85] Our analysis demonstrates that each of the proteins I have described catalyzes a transfer of a specific sugar, but I have not identified the precise linkages of the materials formed. Upon completion of the *in vitro* biosynthesis of the CA repeat unit it would be ideal to also confirm the structure of the product formed to help identify which structure these specific enzymes form. This could shed important light on the precise structure of this complex molecule in nature. After this work, only two steps are required for the formation of the CA repeat unit.[86] It is likely that WcaL transfers a galactose, and there is a pyruvylation by WcaK. However, I have thus far been unable to reconstitute the function of these last two proteins.

Experimental Procedures

Ligation-Independent Cloning (LIC)

A pET-24a vector (Novagen) was altered in two rounds of site-directed mutagenesis using PCR primers (**Table 2.1**) designed to sequentially replace a 5' NdeI with a ScaI restriction site and a XhoI with MscI site. Mutagenesis reactions mixtures contained 0.5 μM each of forward and reverse primer, 4.0 ng of vector DNA, 0.2 mM dNTP mix, 1X Pfu Ultra DNA polymerase buffer (Agilent), and 1 unit of Pfu Ultra DNA Polymerase (Agilent). Mutagenesis was achieved with 25 cycles of 98°C for 10 s, 55°C for 30 s, and 72°C for 30 s. The PCR product was treated with 10 units of DpnI for 3 hrs at 37°C and

transformed into chemically competent DH5α cells. Plasmid was isolated using Fermentas GeneJET Plasmid Miniprep Kit and the process was repeated for the second mutation. ScaI and MscI restriction sites were confirmed with sequencing by Eurofins-Operon.

Linear pET24a-LIC vector was generated by digestion with ScaI and MscI (New England BioLabs) in reaction mixtures containing vector DNA, 1X BSA and 20 units of ScaI and MscI and were incubated for 3 hours at 37°C. Linear plasmid DNA was isolated by gel electrophoresis. Gene inserts were amplified from *E. coli* K12 genomic DNA in 50 µL PCR reactions containing 1X HF buffer (Thermo Fisher Scientific), 0.2 mM dNTPs, 0.2 µM of each primer (**Table 2.1**), and 1 unit of Phusion High Fidelity DNA polymerase (Thermo Fisher Scientific). Amplification was performed with 30 cycles of 98°C for 10 s, 64°C for 30 s, and 72°C for 30 s. Isolated insert and vector were treated with 1 unit of T4 DNA polymerase (Thermo Fisher Scientific) and 2.5 mM dCTP for vector DNA or 2.5 mM dGTP for the gene insert DNA.to produce 20-22 base single-stranded 5' overhangs. All PCR products and T4 digested DNA were purified using the Promega Wizard SV Gel and PCR Clean Up Kit. T4 DNA polymerase treated insert and vector DNA were incubated together in 1:1 or 1:3 vector to insert ratios at 25°C for 30 min and subsequently transformed into competent DH5α cells. After picking a single colony and isolation of vector, clones were confirmed by sequencing (Eurofins-Operon).

Table 2.1. Colanic Acid Primer Sequences

Gene	Forward Primer	Reverse Primer
Nde-ScaI	CTTTAAGAAGGAGATATAAGTACTGCTAGCATGACTGGTGG	CCACCAGTCATGCTAGCAGTACTTATATCTCCTTCTTAAAG
XhoI-MscI	GCTTGCGGCCGCATGGCCACACCACCACCACCACCACTG	CAGTGGTGGTGGTGGTGGTGTGGCCATGCGGCCGCAAGC
WcaA	ATCAGAGGATCCATGAAAAACAATCCGCTGATCTCAATC	ATCAGAGAGCTCGCGCCCCCGAATACCATCAGCCAG
WcaB	ATCAGAGGATCCATGCTGGAAGATCTGCGCGCCAACAGC	ATCAGAGAGCTCTTTAATTACCTTCACTCGCGCTTTTTC
WcaC	ATCAGA GGATCC ATGAATATTTTGCAATTTAATGTGCGA	ATCAGAGAGCTCCAGATTCTGATAGAAGTTGACATAC
WcaE	ATCAGA GGATCC ATGTTGCTTAGCATAATCACTGTCGCG	ATCAGAGAGCTCGACTTTGTTATATAAGGCTTTCGTCT
WcaF	ATCAGAGGATCCATGCAAGATTTAAGCGGTTTCTCGGTGC	ATCAGAGAGCTCTTCAGTTTCAACGCGTTCGCGTATCAC
WcaI	ATCAGAGCTAGCATGAAAATACTGGTCTACGGCATTAAC	ATCAGAGAGCTCTCCCCGAATATCATTTATAAATTGAC
WcaK	ATCAGAGGATCCATGAAATTACTTATTCTGGGCAACCACAC	ATCAGAGAGCTCTTTCACCTCCCCGATGCGCTCAAGCACA
WcaL	ATCAGAGGATCCATGAAGGTCGGCTTCTTTTTACTGAAAT	ATCAGAGAGCTCTAAAGCCTGCAGCAAGCTGGCGAGTT

Expression of the CA Biosynthesis Proteins

All pET-24a vectors containing each of the CA biosynthesis genes were transformed into chemically competent C41 expression cells. Protein was typically expressed in 500 mL cultures. Cultures were incubated after inoculation in a shaker at 37°C until an optical density (O.D.) of at least 0.4. Once an appropriate O.D. was reached, 0.5 mM IPTG was added to each culture flask and the cultures were incubated at 30°C shaking. After 3 hours,

the cultures were pelleted for 15 minutes at 5,000 relative centrifugal force (RCF). The supernatant was then discarded and the pelleted cells were re-suspended in 0.9 % NaCl. The re-suspended cells were transferred to a 50 mL conical tube and spun for 15 minutes at 5,000 RCF. The supernatant was again discarded and the pelleted cells were stored at -80°C.

Induction and Overexpression of Cps2E

A pET-20b vector with *cps2E* gene inserted was previously transformed into BL21-A1 cells for protein expression by the Yother group.[81] A 500 mL culture was prepared from this cell stock with shaking at 37 °C until an O.D. of at least 0.4 was reached. Arabinose was added to induce protein expression at a final concentration of 100 μg/mL, and the temperature was lowered to 16°C and allowed to shake overnight. Cells were pelleted for 15 minutes at 5,000 relative centrifugal force (RCF), re-suspended in 0.9 % NaCl and pelleted again for storage at -80 °C. Cell envelope fractions with Cps2E protein were prepared as described below.

Expression of the CA Biosynthesis Proteins

Each pET-24a-LIC vector containing the CA biosynthesis genes was transformed into chemically competent C41 expression cells. A single colony was picked grown overnight and mixed with 1:1 with glycerol then was stored at -80 °C. Overnight cultures of cells transformed with each vector were then used to inoculate 500 mL cultures, which were then incubated at 37°C with shaking until reaching an optical density (O.D.) of at least 0.4. Once an appropriate O.D. was reached, 0.5 mM IPTG was added to each culture flask and the samples were incubated at 30 °C with shaking. After 3 hours, cells were harvested by

15 minutes of centrifugation at 5,000 RCF. The supernatant was discarded and the cells were re-suspended in 0.9 % NaCl. The re-suspended cells were transferred to a 50 mL conical tube and spun for 15 minutes at 5,000 RCF. The supernatant was discarded and the pelleted cells were stored at -80°C.

Isolation and Purification of Soluble CA Biosynthesis Proteins

Pelleted cells were re-suspended in lysis buffer (50 mM Tris (pH 8), 200 mM NaCl, 20 mM imidazole). The re-suspended cells were lysed by sonication for four minutes (pulse: one second on, one second off). The lysed contents were spun for one hour at 90,140 RCF. Supernatant was pored over a Ni^{2+}-NTA resin, volume 1 mL. After application of the supernatant, 12 mL of wash buffer (50 mM Tris (pH 8), 50 mM imidazole, 200 mM NaCl) was poured through the column and collected. Lastly, six 0.5 mL aliquots of elution buffer (50 mM Tris (pH 8), 500 mM imidazole, 200 mM NaCl) were added and fractions were collected. All fractions containing proteins, based on SDS-PAGE analysis, were combined then dialyzed (3x) at 4°C in 1000 mL of dialysis buffer (50 mM Tris, 300 mM NaCl). Purified protein was stored at -80° C. The concentration of the final purified protein was measured using the calculated extinction coefficients for each protein and UV absorbance (**Table 2.2**).

Table 2.2. CA Soluble Protein Extinction Coefficients [87]

Protein	**Extinction Coefficient (M^{-1} cm^{-1}) at 280 nm**
WcaB	17,085
WcaC	40,308
WcaE	43,890
WcaF	46,075
WcaI	51,113

Preparation of Membrane-Bound Cps2E, WcaA and WcaL

Cell envelope fractions containing Cps2E, WcaA and WcaL were prepared by first lysing overexpressing cells as described above followed by centrifugation for 30 minutes at 2,500 RCF. The pellet was discarded, and the remaining supernatant was centrifuged at 90,140 RCF for one hour. The pellet containing Cps2E, WcaA or WcaL was then homogenized in 1mL of 50 mM Tris (pH 8.0), 300 mM NaCl. Total protein was determined using a Bradford assay with BSA as the standard.

Cps2E Activity Assay

Reaction mixtures were prepared containing 100 mM Bicine (pH 8), 2.5 mM $MgCl_2$, 7.5 mM sodium cholate, 25 μM 2CNA-B(5Z)P, 25 mM UDP-Glc, and 53 μg total protein Cps2E cef in a total volume of 1 mL. The reaction mixture was analyzed by HPLC on a C_{18} (Zorbax XDB-C18 3.5 μm. 4.6 mm x 50 mm) column with an isocratic mobile phase consisting of 42% *n*-propanol and 58% 100 mM ammonium bicarbonate. The 2CNA-B(5Z)PP-Glc was then isolated on a semi preparative C_{18} HPLC column (Eclipse XDB-C18 5 μm, 9.4 mm x 250 mm) with an isocratic method of 50:50 *n*-propanol, 100 mM ammonium bicarbonate at 2 mL/min. The isolated product was used for the remaining reactions.

General Glycosyltransferase Activity Assays

All reaction mixtures were prepared using a master mix of 10.5 μM 2CNA-B(5Z)P P-Glc, 100 mM Bicine (pH 8), 2.5 mM $MgCl_2$, 7.5 mM sodium cholate, 37.5 μM GDP-Fuc, 3.8 mM Acetyl-CoA, 1 mM UDP-Gal, and 1 mM GlcA.. Each CA pathway enzyme (WcaA/B/C/E/F/I/L) was added individually to the reactions at a final concentration of 4 μM, except for WcaA, which was added at 5 μg total protein as a cef. The total volume for each analytical reaction was 20 μL and reactions were carried out at room temperature. Next, products were analyzed stepwise where complete conversion was monitored by HPLC with the preceding enzyme followed by addition of the remaining enzymes at either 4 μM (purified protein) or 5 ug total protein (WcaA or WcaL cef). For example, reactions were prepared to form disaccharide with WcaI, then analyzed by HPLC to ensure turnover before aliquoting the WcaI product into mixtures containing the above proteins. All

glycosyltransferase and acetyltransferase reactions were analyzed by HPLC using a C_{18} (Zorbax XDB-C18 3.5 µm. 4.6 mm x 50 mm) column and an isocratic mobile phase consisting of 35% n-propanol and 65% 100 mM NH_4HCO_3.

ESI-Mass Spectrometry

For ESI-MS analysis the oligosaccharide assembly reactions were scaled up to 20 µM of 2CNA-B(5Z)-linked glycan product in a total volume of 1 mL. Products were isolated on a semi preparative C_{18} HPLC column (Eclipse XDB-C18, 5 µm, 9.4 mm x 250 mm) with an isocratic method of 61% *n*-propanol and 39% 100 mM ammonium bicarbonate at 2 mL/min. The 2CNA-B(5Z)PP-Glc-AcFuc-Fuc-AcGal-GlcA (WcaA/B product) was isolated as previously stated but with an isocratic method of 50% *n*-propanol and 50% 100 mM ammonium bicarbonate. The isolated products were used for ESI-MS analysis performed on the fractions collected from the HPLC. MS analysis was done on a Thermo VELOS Pro Dual-Pressure Linear Ion Trap instrument using ESI introduced by direct flow of the dilute sample into the ESI apparatus at 15 µL/min. Negative ion mode was used with a capillary temperature of 240°C and a spray voltage of 4.0 kV. Once a signal was obtained, MS^n with collision energy of 35 kV was used to select for and fragment the m/z of interest.

CHAPTER 3: CLICK-ENABLED DETECTION OF POLYISOPRENOIDS FOR COMPLEX GLYCAN BIOSYNTHESIS AND IMMOBILIZATION

Note: Contributions to this work include the synthesis of the azide analogue by Amanda Reid (Nanoscale PhD, 2020), and the cloning and expression of the UppP enzyme by Beth Scarbrough (Nanoscale PhD 2022). All data is from K. Erickson unless otherwise noted.

Short Synopsis

The goal of Chapter 3 is to develop an immobilization platform for the production of agents that display selectivity for bacterial sugars. This chapter will describe the synthesis of an analogue with a chemical handle that provides a method of conjugation or immobilization. Secondly, this chapter will display the ability to utilize the chemical handle for conjugation to brighter fluorophores. Thirdly, evidence is provided to show that the analogue is utilized by specific enzymes for biosynthetic pathways, and that the pathways also accept this material. Finally, two different glycans were built on the analogue linked to both a conjugation molecule (biotin) and a fluorophore that was added prior to and after the glycan addition.

Introduction

Major advances in chemical tools in recent years have enabled azide and alkyne-linked sugars to be used to metabolically label glycoproteins and other glycoconjugates in the

study of animal development, glycome analysis, and a wide range of other important areas of biology.[88-91] One problematic area for the use of such tools has been their application in bacterial glycobiology due, in part, to the massive sugar diversity associated with these organisms. Recent advances in this field have begun to close this gap with new methodologies for assembling nucleotide-linked sugar analogues as enzymatic substrates for bacterial systems and the development of fluorescent D-amino acids for installation into peptidoglycan.[92-93]

While advances in new sugar based derivatives have had extraordinary impact on glycobiology and microbiology as a whole, there is an inherent limitation because of the sheer number of sugar structures and specificity of the enzymes involved in bacterial glycan assembly.[9] Our group has recently focused on a common pathway associated with the production of complex bacterial glycans ranging from teichoic acids to capsules and secreted polysaccharides. These pathways depend on the C55 isoprenoid bactoprenyl phosphate (BP), which serves as a lipid carrier for the assembly of the majority of bacterial glycans.[64] Individual sugars are appended sequentially to this lipid carrier (**Figure 2.1**) until oligosaccharide completion, then the pathway diverges to protein glycosylation, glycan polymerization, and other polysaccharide production pathways.[94]

The Poulter group developed a series of azide and alkynyl analogues of FPP for the study of eukaryotic protein prenylation systems.[95] Their group and others have used the analogues in Staudinger ligation protocols to examine the eukaryotic prenylome, and as a tagging strategy for cellular proteins.[96] In this chapter, I test a benzyl azide containing FPP analogue for its application to bacterial glycan assembly systems, and identify it as a

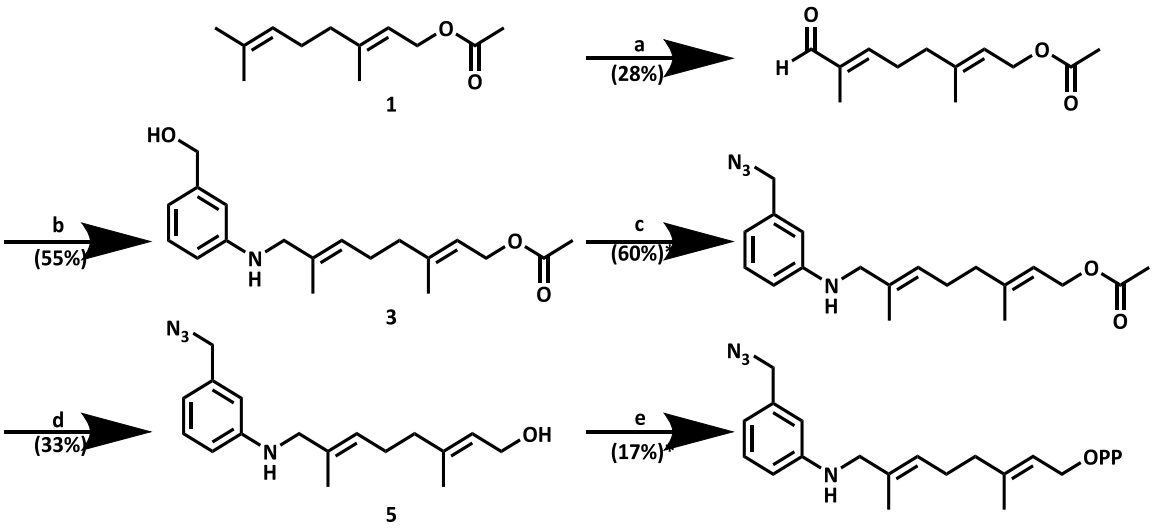
potentially powerful tool for applications in bacterial glycobiology. The azide FPP analogue allows for the assembly of the glycan onto a platform that is easily immobilized to a surface, or resin. This immobilization platform provides a way to select for agents specific to the bacterial glycans of interest, as an extension beyond fluorescent tagging.

Results

UppS forms a "click" ready product with a 3-azido-benzylanilinogeranyl diphosphate

In order to determine whether an azide tag could be incorporated into BPP utilizing UppS and an azide FPP analogue, I synthesized the 3-azido-benzylanilinogeranyl diphosphate (AzBn-GPP, **Figure 3.1**) first reported by the Poulter group (**Scheme 3.1**).[95] To test activity with UppS, I mixed the AzBn-GPP with *Staphylococcus aureus* UppS, isolated as previously described for *Bacterioides fragilis* UppS, and analyzed product formation by HPLC.[82, 97] As previously observed with other anilinogeranyl diphosphate analogues, the azide analogue was a substrate for UppS giving products with varying total isoprenoid length (**Figure 3.2a**). To test whether the elongated products could be labeled, I utilized a dibenzocyclooctyne (DBCO)-linked TAMRA fluorophore (**Figure 3.3**), and analyzed again by HPLC (**Figure 3.2d**). The labeling reaction resulted in a large retention time shift consistent with the addition of the relatively hydrophilic fluorophore. To confirm the formation of the azide UppS products, I scaled the reactions up for isolation and characterization by electrospray ionization mass spectrometry (ESI-MS). Using the MS

data I were able to confirm the addition of 4-7 isoprenes by UppS to the AzBn-GPP (**Figure 3.2b**).



Scheme 3.1. Synthesis of the FPP Analogue a) 70% t-BuOOH, SeO_2, Salicylic Acid, CH_2Cl_2. b) 3-aminobenzyl-OH, AcOH, $NaBH(OAc)_3$, CH_2Cl_2. c) DPPA, DBU, toluene. d) K_2CO_3, MeOH. e) 1) PBr_3, CH_2Cl_2. 2) tris (tetra-N-butyl ammonium) diphosphate, MeCN. The percentages represent the yield for each reaction step. (Synthesis originally performed by P. Scott, continued by K Erickson and modified by A. Reid) * corresponds to percent yields for each reaction step.

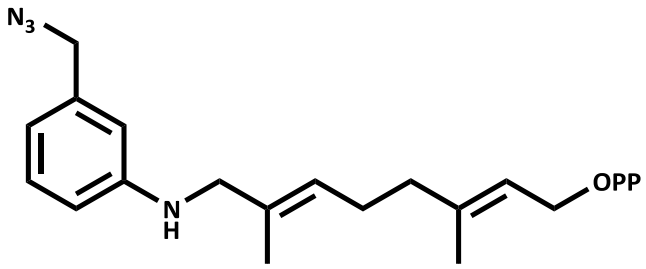


Figure 3.1. The chemically synthesized AzBn-GPP structure.

AzBn-BPP is a substrate for bacterial UppP removing a single phosphate

In order for a bactoprenyl analogue to be useful for glycan biosynthesis a single phosphate must be removed to provide a BP substrate for initiating phosphoglycosyltransferases. UppP is a well characterized phosphatase responsible for this activity in cells.[98-99] C41 cells overexpressing UppP were lysed then membrane components were isolated as a part of the cell envelope fraction (cef). The cef was then utilized with AzBn-BPP to test for the ability of endogenous UppP to cleave the phosphate group. Aliquots of the reaction were removed then labeled with TAMRA for analysis by HPLC (**Figure 3.2d**). There was a clear indication of phosphate removal with a slight increase in retention time of one minute from the starting material. Product was isolated then subjected to ESI-MS analysis where the m/z confirmed the formation of monophosphate product (**Figure 3.2c**). I next tested whether TAMRA labeled BPP could serve as a substrate for the phosphatase. To do this, I mixed crude UppS product labeled with the TAMRA fluorophore with UppP cef. Surprisingly, there appeared to be little influence of the large fluorophore tag on formation of the monophosphate product (**Figure 3.4**).

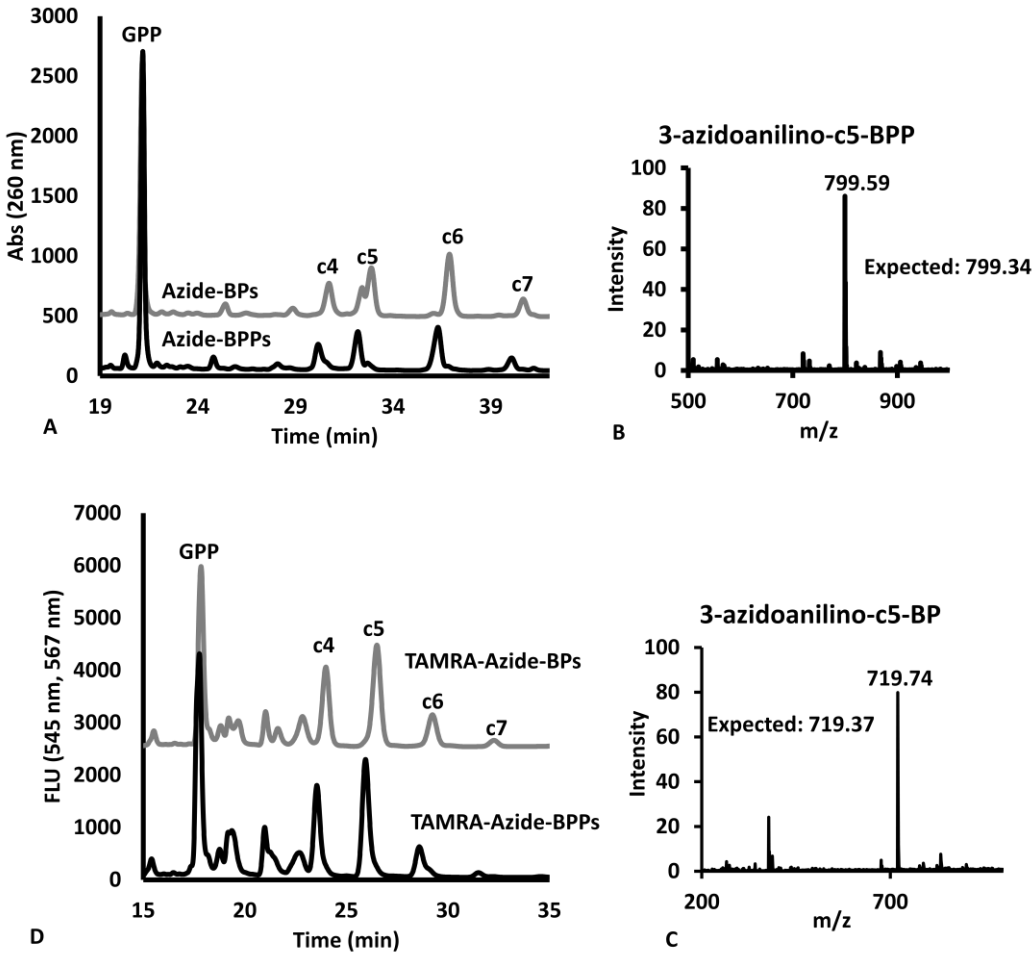


Figure 3.2. HPLC Analysis of AzBn-GPP with *UPPS* enzyme (a, black line) and that reactions with the addition of the BacA enzyme (a, grey line). ESI-MS analysis of the AzBn-BP(P) products from the UPPS (b) and BacA (c) reactions show the expected m/z for the BPP and BP products. The Az-GPP modified by the UPPS (black line) and BacA (grey line) enzymes were labeled with TAMRA-DBCO (d). The HPLC reactions were analyzed with a semi preparative C_{18} HPLC column (Eclipse XDB-C18 5 µm, 9.4 mm x 250 mm) with a gradient method of 2.2% *n*-propanol per minute with the aqueous component of 100 mM ammonium bicarbonate at 1 mL/min.

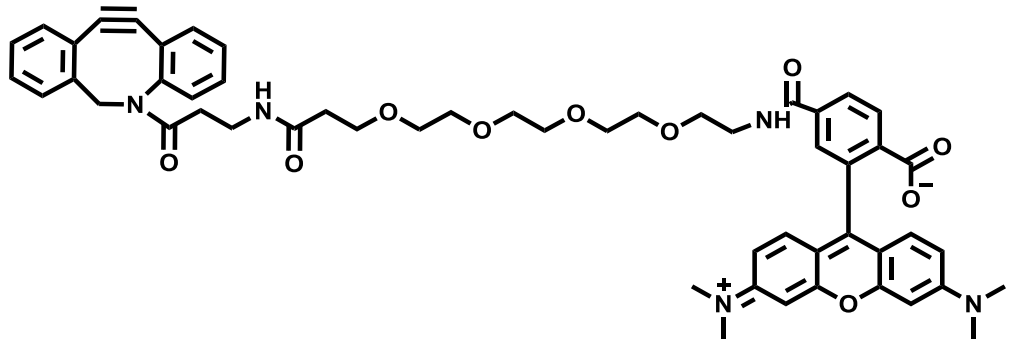


Figure 3.3. The structure of the Dibenzocyclooctyne-PEG_4-TAMRA (TAMRA-DBCO) used in the "click" reactions of the AzBn-GPP/BPP/BPs.

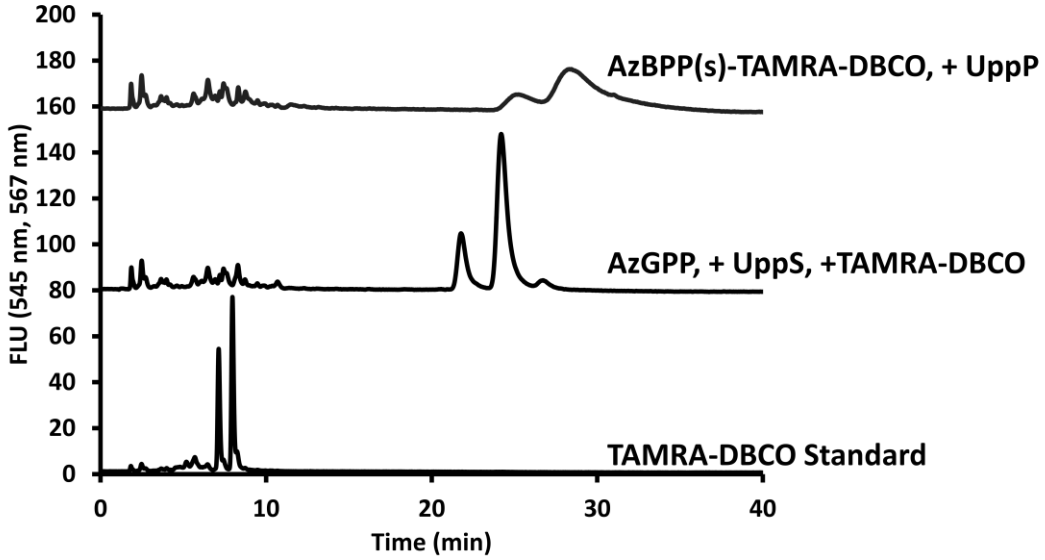


Figure 3.4. HPLC Analysis of the TAMRA-labeled AzBn-BPP products being accepted by UppP and forming TAMRA-labeled monophosphate products. The AzGPP was utilized by the UppS enzyme and then the DBCO-TAMRA was added after the BPP products were formed. The HPLC analysis was done with C_{18} HPLC column (Zorbax XDB-C18 5 µm, 4.6 mm x 50 mm) with a gradient method of 2.2% *n*-propanol per minute with the aqueous component of 100 mM ammonium bicarbonate at 1 mL/min. (Data for this figure was obtained by Amanda Reid).

Campylobacter jejuni N-glycan assembly on a Fluorescent BP

The pathway associated with N-linked glycan assembly in the organism *Campylobacter jejuni* has been well studied and serves as an excellent model system for the application of the azide-linked BP. In order to test the activity of the pathway enzymes, I first screened activity with our more widely used fluorescent 2-nitrileanilinobactoprenyl phosphate (2CNA-BP, **Figure 2.2**). UDP-diNAcBac was assembled using procedures similar to published methods, except an alternative to PglF, WbjB from *Vibrio vulnificus*, was used for formation of a 4-keto-sugar substrate for PglE (**Figure 3.5**).[100] The *pglE, D,* C, *A, J, H* and *I* were all overexpressed and I isolated protein product using published methods (**Figure 3.6**).[35, 101-104] I mixed the 2CNA-BP with PglC and UDP-diNAcBac and I observed a retention time change of 7 to 4 minutes by HPLC, which was consistent with the addition of diNAcBac-phosphate to 2CNA-BP (**Figure 3.7a**). Once I confirmed the formation of the 2CNA-BPP-diNAcBac I then tested the remaining proteins PglA, PglJ, PglH, and PglI in sequential reactions analyzed by HPLC. I observed retention time shifts of about 0.5 minutes, consistent with the addition of GalNAc to the mono and di-saccharides with PglA and PglJ respectively (**Figure 3.7a**). I also observed a larger retention time change of one minute with PglH, consistent with the addition of three GalNAc residues, as well as a smaller retention time change with PglI and UDP-glucose of about 0.5 minutes. Formation of the PglC product was difficult to push to completion, but all other reactions appeared to completely consume the fluorescent substrate. This is a common occurrence with these kinds of transferases and has specifically been noted for the PglC and PglA combination.[71]

The 2CNA-BPP-oligosaccharide was next isolated and I subjected it to ESI-MS analysis confirming the formation of the 2CNA-BPP-heptasaccharide (**Figure 3.7b**).

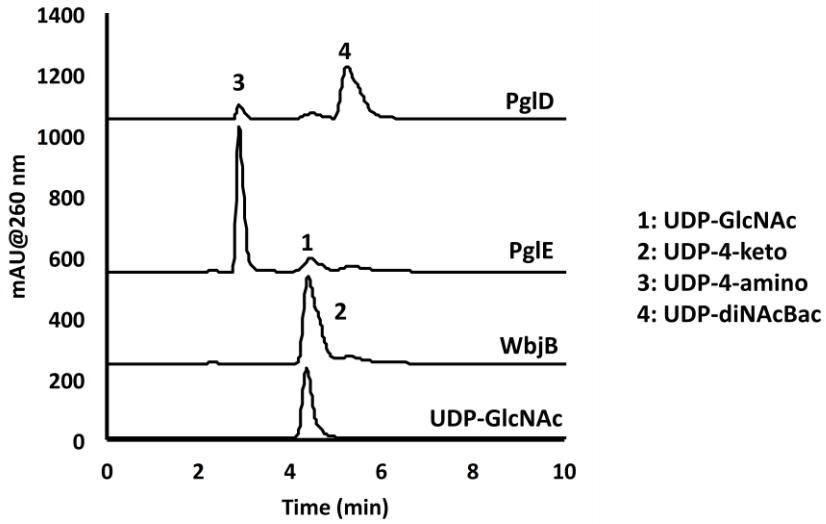


Figure 3.5. HPLC analysis of the sugar modifying enzymes WbjB, PglE and PglD with UDP-GlcNAc sugar on a Zorbax-NH_2 (4.6 x 150 mm, 5 µm) with 250 mM ammonium acetate (pH 4.5).

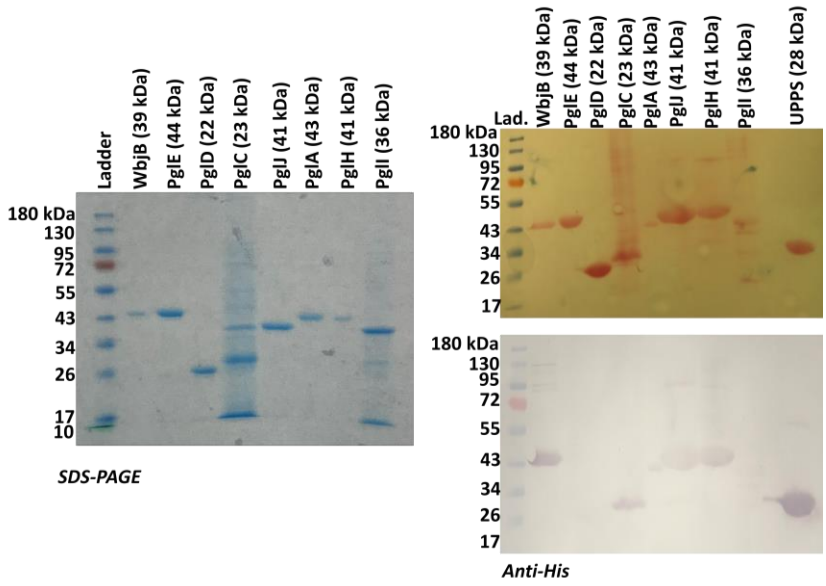


Figure 3.6. SDS-PAGE analysis of the proteins used in the production of the heptasaccharide of *C. jejuni* (left panel). The right panel shows the Ponceau stain, staining all possible protein on the blot, and below is the anti-histidine tag Western blot, specific to the His-containing recombinant proteins.

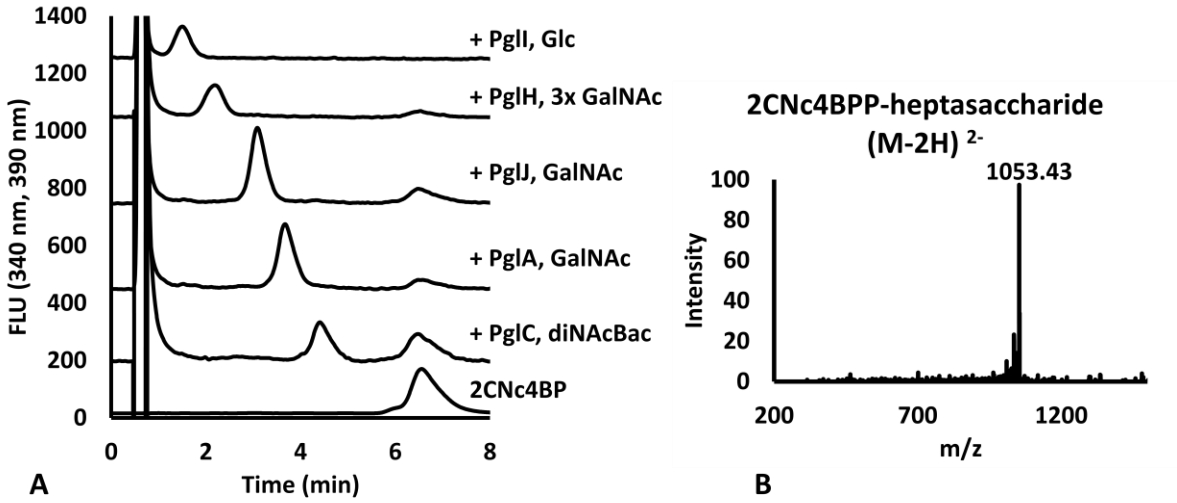


Figure 3.7. HPLC Analysis of 2CNc4BP with the pgl proteins adding each additional sugar onto the bactoprenyl. PglC adds the diNAcBac sugar, PglA adds a GalNAc, PglJ adds another GalNAc, PglH adds three GalNAc sugars and finally PglI adds the Glc sugar to form the heptasaccharide. The HPLC analysis was done with a C_{18} HPLC column (Zorbax XDB-C18 5 µm, 4.6 mm x 50 mm) with a gradient method of

2.2% *n*-propanol per minute with the aqueous component of 100 mM ammonium bicarbonate at 1 mL/min. The ESI-MS of the 2CNc4BPP-heptasaccharide (with the addition of 5% ammonium hydroxide before analysis to obtain the negatively charged species) produced an m/z of 1053.43 m/z (expected 1053.45 m/z).

TAMRA labeled BP serves as a substrate for N-linked glycan biosynthesis proteins

Since the fluorescent 2CNA-BP was a substrate for *C. jejuni* N-linked glycosylation proteins, I were next interested in whether the TAMRA-AzBn-BPP would also serve as an effective substrate. To test this, I first screened for activity of all Pgl proteins with the AzBn-BP as a substrate. Based on HPLC, I did observe product formation with PglC, A, J, H, and I (data not shown). Next, I tested purified TAMRA labeled BP and found that even with the large fluorophore replacing the terminal isoprene unit I observed sugar additions for each enzyme in the pathway (**Figure 3.8**).

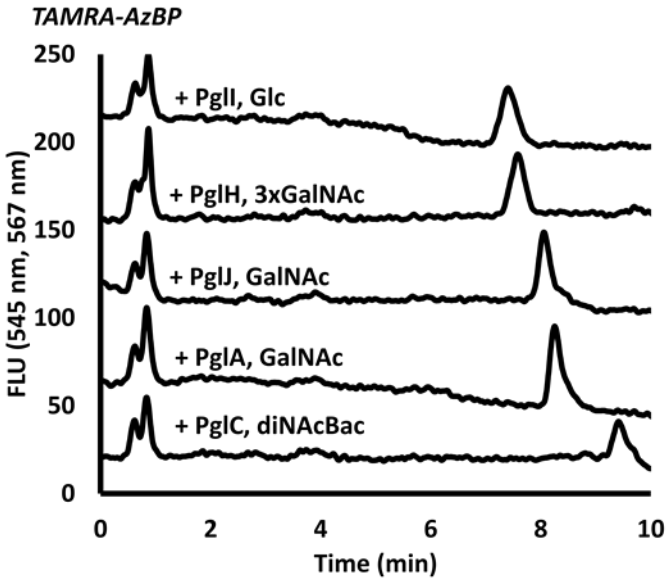


Figure 3.8. HPLC Analysis of the TAMRA-AzBnBP with the pgl proteins adding each additional sugar onto the bactoprenyl. PglC adds the diNAcBac sugar to produce the 9 minute peak, PglA adds a GalNAc to form the product around 8 minutes, PglJ adds another GalNAc with the product at around 7.9 minutes,

PglH adds three GalNAc sugars to form the product at around 7.5 minutes, and finally PglI adds the Glc sugar to form the heptasaccharide at around 7 minutes. The HPLC analysis was done with a C_{18} HPLC column (Zorbax XDB-C18 5 µm, 4.6 mm x 50 mm) with a gradient method of 2.2% *n*-propanol per minute with the aqueous component of 100 mM ammonium bicarbonate at 1 mL/min.

TAMRA labeled BP serves as a substrate for colanic acid biosynthesis proteins

Since 2CNA-BP and the TAMRA-AzBn-BP were both substrates for the N-linked oligosaccharide pathway, and the 2CNA-BP was also a substrate for the colanic acid pathway, I were interested in seeing if the use of the TAMRA-AzBn-BP also was utilized by the colanic acid enzymes. To test this, I used isolated TAMRA-AzBn-BP with the colanic acid proteins (**Figure 3.9**). Although the TAMRA-AzBn-BP is much larger than the 2CNA-BP it was still able to be utilized by the colanic acid enzymes. This is significant because it shows the versatility of this large analogue in multiple pathways.

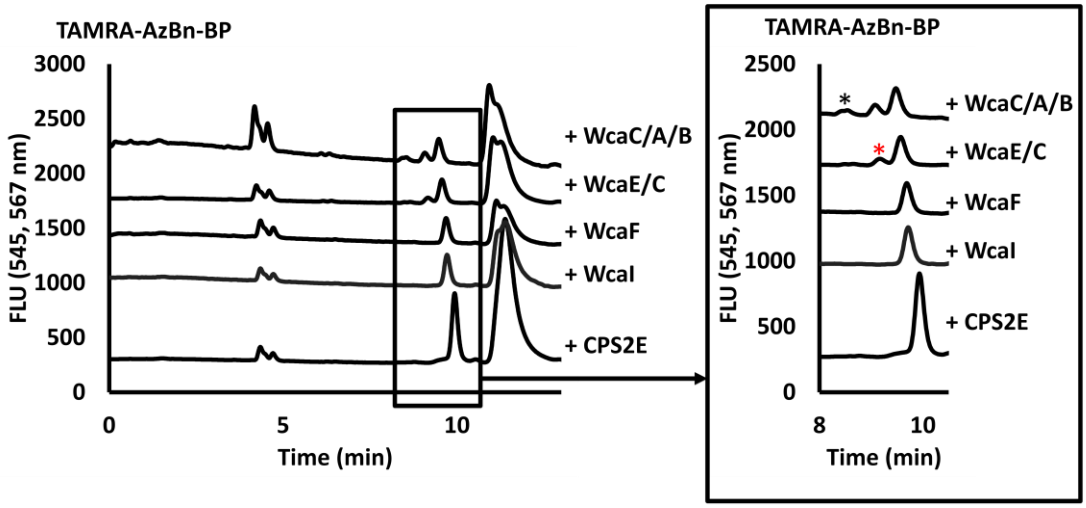


Figure 3.9. HPLC Analysis of the TAMRA-AzBnBP with the colanic acid proteins adding each additional sugar onto the bactoprenyl. These reactions were conducted with a mix of all of the possible sugars and cofactors (UDP-glucose, UDP-galactose, GDP-fucose, Acetyl-CoA, and UDP-glucuronic acid). CPS2E is the initiating hexose-1-phosphate transferase that adds glucose, WcaI adds fucose, WcaF acetylates the fucose, WcaE add another fucose, WcaC adds a galactose, and WcaA adds a glucuronic acid (in the presence of WcaB, which acetylates). Each addition shows a retention time shift of around 0.2 minutes due to the increased hydrophilicity of the sugar addition; except for the acetyl group addition, for which the shift is not detectable with this method and column. The * represents the WcaC product (complete turnover was not achieved). The * represents the WcaA/B product (again, complete turnover was not achieved). There is significant leftover TAMRA-BP (~11 minutes). The HPLC analysis was done with a C_{18} HPLC column (Zorbax XDB-C18 5 µm, 4.6 mm x 50 mm) with a gradient method of 2.2% *n*-propanol per minute with the aqueous component of 100 mM ammonium bicarbonate at 1 mL/min.

Immobilization preparation of the glycans onto the biotinylated anchor

Once I determined that the biosynthetic enzymes utilized the TAMRA-AzBn-BPs, I wanted to see if this could be expanded to a biotinylated octyne. The AzBn-BPPs were biotinylated and cleaved with UppP (**Figure 3.10**). They were then isolated and utilized in the biosynthetic pathways to produce the N-linked oligosaccharide (**Figure 3.11**) and the colanic acid tetrasaccharide (**Figure 3.12**). These products were utilized in studies discussed in Chapter 4.

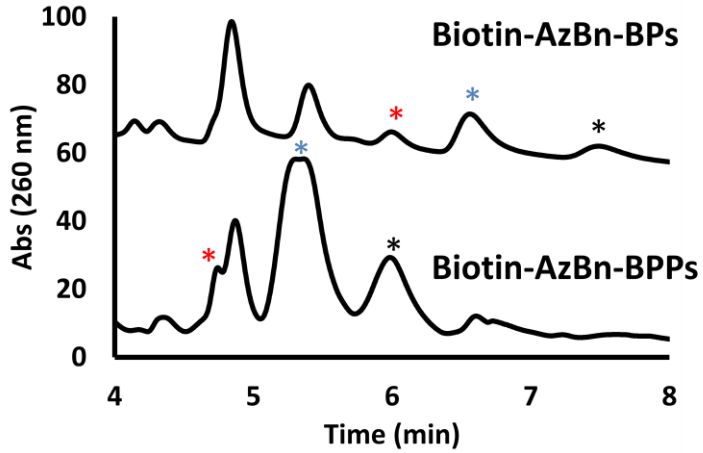


Figure 3.10. HPLC Analysis of the Biotin-AzBnBPPs being cleaved to the Biotin-AzBnBPs where the AzBn-GPP was reacted with UppS, biotinylated with 2 x mole excess of the Biotin-PEG_4-DBCO (Sigma) and a phosphate was cleaved to form the Biotinylated-AzBnBPs, which is displayed with a shift in retention time corresponding to the loss of a phosphate. The HPLC analysis was done with a C_{18} HPLC column (Zorbax XDB-C18 5 µm, 4.6 mm x 50 mm) with a gradient method of 2.2% *n*-propanol per minute with the aqueous component of 100 mM ammonium bicarbonate at 1 mL/min. Each of the matching BPP/BP product peaks is denoted with a similar colored "*" for clarity.

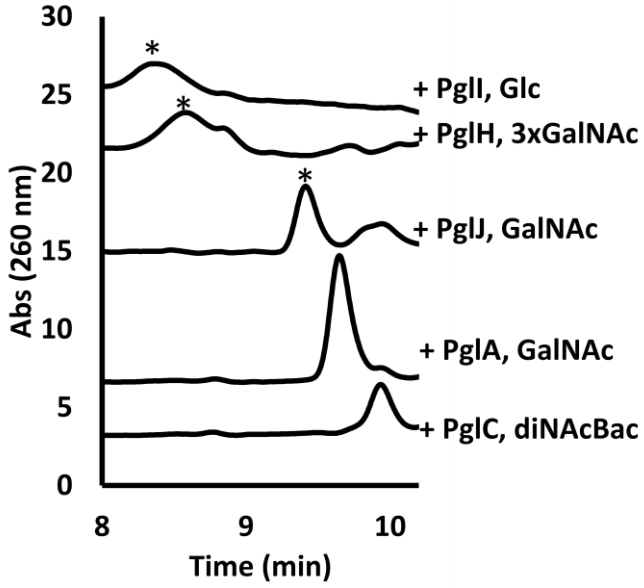


Figure 3.11. HPLC Analysis of the Biotin-AzBnBP with the N-linked oligosaccharide proteins adding each additional sugar onto the bactoprenyl. PglC is the initiating hexose-1-phosphate transferase that adds diNAcBac, PglA, PglJ and PglH add five UDP-GalNAcs and PglI adds a UDP-Glc. The HPLC analysis was done with a C_{18} HPLC column (Zorbax XDB-C18 5 µm, 4.6 mm x 50 mm) with a gradient method of 2.2% *n*-propanol per minute with the aqueous component of 100 mM ammonium bicarbonate at 1 mL/min. The product peak is denoted with a "*" for the PglJ, PglH and PglI reactions.

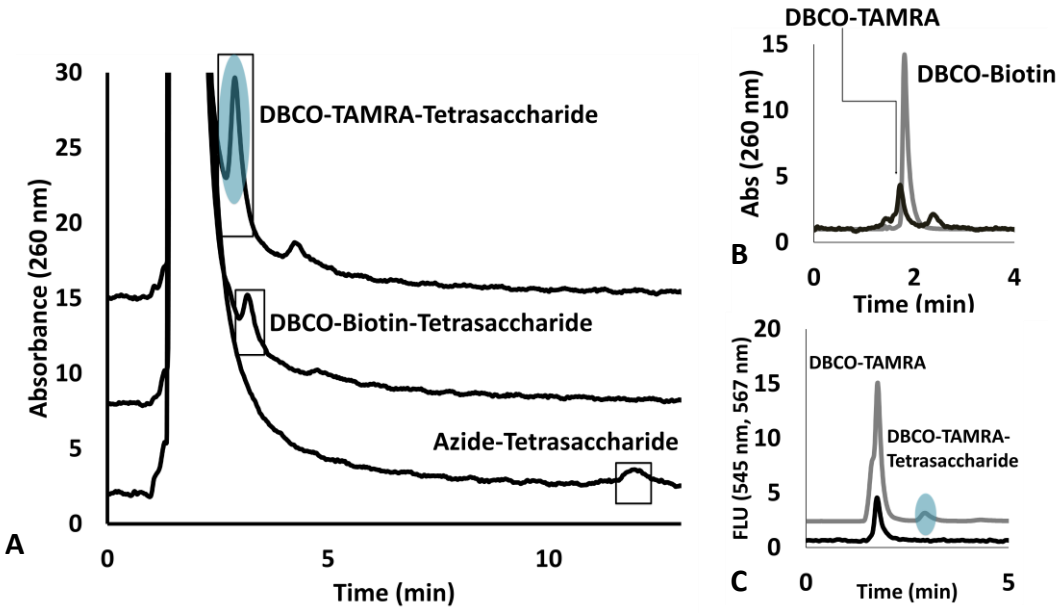


Figure 3.12. HPLC Analysis of the Biotin-AzBnBP with the colanic acid tetrasaccharide. First, the tetrasaccharide was built onto the azide-BP (a, bottom), then biotinylated (a, middle) or clicked to TAMRA (a, top). The difference between the TAMRA-DBCO and the Biotin-DBCO is shown in (b) where the TAMRA-DBCO elutes before the Biotin-DBCO. In (c) I can see that only the TAMRA-tetrasaccharide product is fluorescent (grey line). The HPLC analysis was done with a C_{18} HPLC column (Zorbax XDB-C18 5 µm, 4.6 mm x 50 mm) with a gradient method of 2.4% *n*-propanol per minute with the aqueous component of 100 mM ammonium bicarbonate at 1 mL/min.

Azide labeling enhances detection of UppS products to 40 fmoles

The use of azide-linked polyisoprenoids provides the ability to label the UppS product with a variety of fluorophores to increase detection limits by orders of magnitude over previously reported materials. The HPLC detection limit for 2CNA-BP (**Figure 3.13)** is

in the 1-5 pmole range (**Figure 3.14**). The AzBn-BP detection limit was in the 5-10 nmole range. However, treatment of the AzBn-BP with DBCO-TAMRA (**Figure 3.13)** enhanced detection limit to 40 fmoles (**Figure 3.14**). Labeling of the azide therefore enhanced detection limits by 5 orders of magnitude above the AzBP and 2 orders of magnitude over the nitrileaniline.

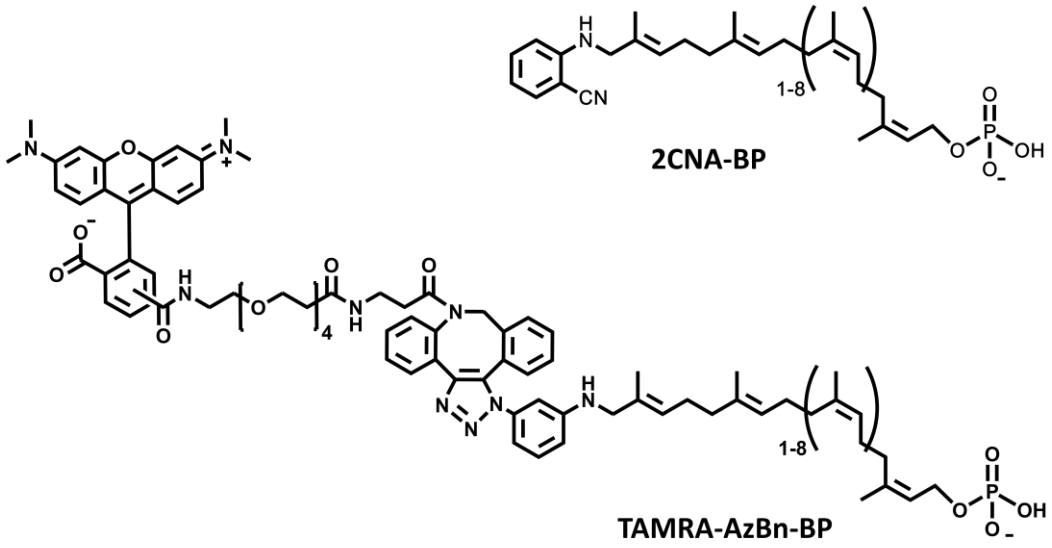
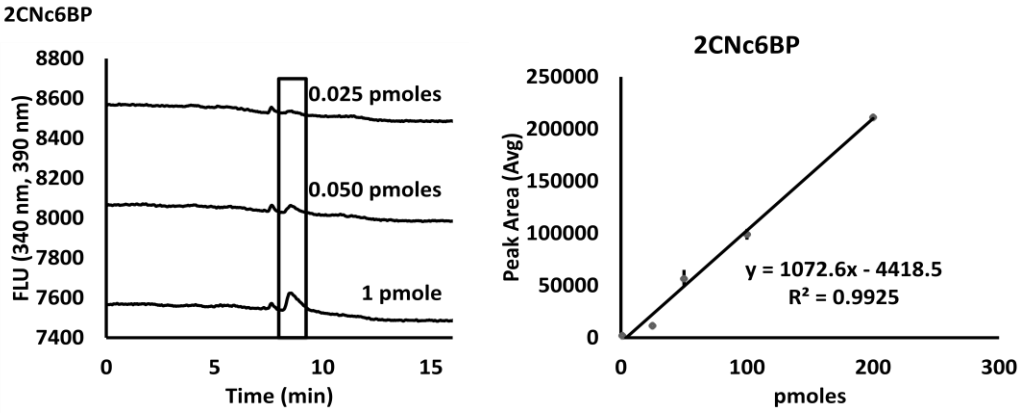
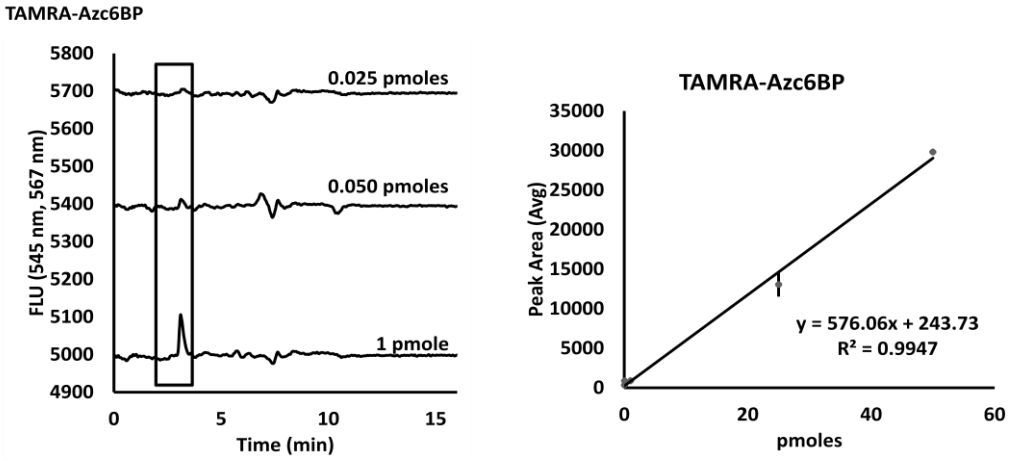


Figure 3.13. Structure of the 2CNA-BP compared to the TAMRA-AzBn-BP analogue. The size of the TAMRA-linked analogue is quite large compared to that of the 2CNA-BP analogue and both can be utilized in biosynthetic pathways.



Sample	DL (pmoles)	QL (pmoles)
2CNc6BP	1.6	5.2



Sample	DL (pmoles)	QL (pmoles)
TAMRA-Azc6BP	0.4	1.3

Figure 3.14. HPLC Analysis of the detection limit of the 2CNc6BP and the TAMRA-Azide-c6BP, the left panels display the minimum amount of the BP that was injected and could be seen by HPLC for comparison to each other; the right panels show the pmoles injected into the HPLC and the average of the peak areas of three injections. From the slope of the peak areas and the pmoles injected the detection limit "DL" and the quantitation limit "QL" were calculated and tabulated. The DL of 2CNc6BP is 1.6 pmoles

compared to 0.4 pmoles for TAMRA-Azide-c6BP and the QL is 5.2 pmoles for 2CNc6BP compared to 1.3 pmoles for the TAMRA-Azide-c6BP.

Discussion

Traditional sugar labeling implies the use of radioactive labels which typically require extractions to follow incorporation of the sugars and thin layer chromatography (TLC) to separate products.[103-105] This methodology requires expensive equipment and upkeep to maintain the use of radiolabels.[68] In lieu of radiolabeled sugars, the sugars are enzymatically added to the bactoprenyl-based click handle that provides the conjugation ability. With this methodology the label is associated with the bactoprenyl anchor and not with one individual sugar, or sugars, in addition to being easily analyzed by HPLC methods rather than radiography.

Here I show that an azide-linked bactoprenyl anchor can be enzymatically produced from a simple chemically produced precursor. This chemoenzymatic strategy is particularly important as elongation with appropriate configuration through chemical methods can be long and tedious, and may not be possible with the azide functionality present.[106] In this work I show that the azide analogue is utilized by the UppS enzyme to form the AzBn-BPPs. Furthermore, the phosphatase UppP utilized the AzBn-BPPs, the TAMRA-AzBn-BPPs, and the Biotin-AzBn-BPPs, which shows the diversity of this enzyme in product utilization. Both of the AzBn-BP(P)s have been characterized by HPLC and ESI-MS.

Due to the diversity of the substrates that UppS and UppP utilize, this clickable bactoprenyl platform has potential for expansion to other biosynthetic pathways and the

use of more powerful fluorophores and versatile conjugation methods. For utility in biosynthetic pathways I have shown the use of the TAMRA and the Biotin-AzBn-BP in the *C. jejuni* N-linked oligosaccharide and *E. coli* colanic acid pathways by HPLC and ESI-MS analysis. The unexpected ability of the biosynthetic enzymes to utilize such a large analogue (the Biotin or TAMRA-AzBn-BP) displays the possibility of unlimited modifications to the non-phosphorylated end of the farnesyl analogues. This is a significant find because not only does it show the versatility of these analogue modifications, but that it is possible in more than one biosynthetic pathway.

For fluorescent applications, our data shows that between the established 2CNA-B(Z)P analogue and the TAMRA-AzBn-BP analogue that the detection limit of the TAMRA fluorophore is 2 orders of magnitude higher. This is just one of the possibilities of fluorophores and conjugations available that are 'click' ready. In addition to more powerful fluorophores, the use of method specific chromophores that are needed in the visible range for confocal microscopy, for example, are accessible as long as it contains a 'clickable' moiety. Furthermore, the click platform provides a way to attach the sugar units to a surface, or a resin for downstream analysis, providing the needed platform to select for agents specific to bacterial glycans.

Experimental Procedures

Synthesis of the Azide-Benzyl Geranyl Diphosphate Analogue

The synthesis protocol for the FPP analogue has been adapted and modified from Chehade *et. al* and Labadie *et. Al*. [76, 95] Note that precursors to the final diphosphate are

characterized only to the extent needed to compare with literature values (the aldehyde precursor has been previously characterized by our group).

Synthesis of (E,E)-3,7-Dimethyl-1-acetoxy-2,6-octadien-8-al

Excess CH_2Cl_2, 0.283 g SeO_2 (0.283 g, 2.55 mmols), 0.354 g salicylic acid (0.354 g, 2.56 mmols) and 70% tert-butyl hydroperoxide (13.1 mL, 135 mmols) were added to a round bottom flask and placed in an ice bath. Once the mixture was homogeneous, geranyl acetate (5 g, 25.47 mmols) was added and the reaction was left stirring overnight. The reaction was diluted with ether and extracted in a separatory funnel with the following: 5% $NaHCO_3$, saturated $CuSO_4$, saturated $Na_2S_2O_3$ and 0.9% NaCl. Each washing step was repeated twice. The organic layer was dried with anhydrous $MgSO_4$ and the solvent was removed by reverse pressure. The remaining oil was purified using silica flash chromatography using a 5% (v/v) EtOAc/Hexanes solution.

Synthesis of Acetoxygeranyl Benzyl Alcohol

A round bottom flask was flame-dried under argon gas to remove any moisture. Once cooled, excess CH_2Cl_2, 3-aminobenzyl alcohol (0.604 g, 4.90 mmols) and of geranyl aldehyde (0.936 g, 4.45 mmols) were added to the flask followed by glacial acetic acid (307 µL, 5.54 mmols) and 1.3 g of $Na(OAc)_3BH$ (1.3 g, 6.13 mmols). The reaction was generally left overnight and extracted with chloroform the following day. The product was purified first using 10% EtOAc/Hexanes to remove any leftover geranyl aldehyde followed by 30% EtOAc/Hexanes to isolate the desired product. (^{1}H, $CDCl_3$, δ): 7.19 (t, 1H), 6.73-6.59 (m, 3H), 5.44 (m, 2H), 4.65 (s, 2H), 4.63 (d, 2H), 3.70 (s, 2H), 2.22-2.11 (m, 7H), 1.78 (s, 3H), 1.75 (s, 3H). Yield = 55%.

Synthesis of 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadien-1-ol

The benzylic alcohol was then converted to an azide. Excess toluene was added to a round bottom flask. Next, DPPA (0.831 g, 3.02 mmols), 442.6 mg of DBU (0.443 g, 2.91 mmols) and 775.1 mg of benzylic alcohol geranyl acetate (0.775 g, 2.44 mmols) were added and the mixture stirred on ice for 2 hours and at room temperature overnight. The reaction was purified using 10% EtOAc/Hexanes to remove the product along with a second spot with a very similar Rf. Converting to the alcohol generated a much larger difference between R_f values and allowed for easier purification of the benzyl azido geranyl alcohol. The 8-N-m-benzyl alcohol-amino-3,7-dimethyl-2,6 octadien-1-ol product was formed by reacting acetoxy 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphate (0.5 g, 1.46 mmols) with excess MeOH and K_2CO_3 (0.608 g, 4.40 mmols). The reaction was left overnight and purified using 10% EtOAc/Hexanes followed with 50%EtOAc/Hexanes. (1H, $CDCl_3$, δ): 7.19 (m, 1H), 6.61 (m, 3H), 5.41 (m, 2H), 4.23-4.12 (m, 4H), 3.70 (s, 2H), 2.20-2.04 (m, 4H), 1.75 (m, 6H). Rf = 0.3. Yield = 33%.

Synthesis of 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphate

Next, the benzyl azido geranyl alcohol was brominated followed by subsequent diphosphorylation. PBr_3 (7.44 mg, 27.49 µmols) was added dropwise to a mixture containing 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadien-1-ol (10 mg, 33.3 µmols) and excess CH_2Cl_2. The reaction occurred almost instantaneously. Tetra butyl ammonium diphosphate (210 mg, 233 µmols) was added dropwise to this mixture and left stirring overnight at room temperature. The following day, the solvent is removed and the remaining material is dissolved in 25 mM ammonium bicarbonate in 20% propanol and

run through an ion-exchange column to replace the tetrabutyl ammonium with ammonium. The purified analogue is then frozen and lyophilized prior to purification by HPLC and characterization by ^{1}H-NMR, ^{13}C-NMR, ^{31}P-NMR and mass spectrometry. (^{1}H, D_2O, δ): 7.13 (t, 1H), 6.67 (d, 3H), 5.25 (q, 2H), 4.30 (t, 2H), 4.17 (s, 2H), 3.54 (s, 2H), 2.07-1.94 (m, 4H), 1.52 (s, 3H), 1.47 (s, 3H). (^{13}C, D_2O, δ): 148.2, 142.4, 136.4, 132.0, 129.5, 126.2,119.4, 118.2, 114.7, 114.2, 62.5, 54.0, 50.7, 38.2, 25.0, 15.2, 13.3. (^{31}P, D_2O, δ): -8.91 (1P), -10.14 (1P). Expected m/z 459.12, obtained 459.27 m/z; Extinction coefficient ε= 4345 M^{-1} cm^{-1} at 260 nm; Yield =16%

Expression and Isolation of the C. jejuni enzymes

The enzymes from the *C. jejuni* pathway were expressed as previously published with a few modifications.[35, 102-104] PglC and PglI were expressed in cell envelope fractions (cef) and induced by IPTG, whereas PglA, PglH and PglJ were auto induced[107], soluble proteins, with no triton used in the sonication or the isolation. PglD and PglE were soluble and isolated in the presence of 1% Triton. Concentrations for the proteins were determined by Bradford for the cefs (PglC and PglI) and the triton containing proteins (PglD and PglE). Using the extinction coefficients for PglA, PglH and PglJ, (**Table 3.1)** the concentrations of the protein were determined with a NanoDrop Spectrophotometer (Thermo).

Table 3.1. N-linked oligosaccharide *C. jejuni* protein extinction coefficients.

Protein	**Extinction Coefficient (M^{-1} cm^{-1})**
PglA	44,725
PglJ	18,380
PglH	29,380

Formation of diNAcBac with WbjB, PglE and PglD

To form the UDP-diNAcBac sugar product, 5 mM UDP-GlcNAc was utilized by 5 µM WbjB from *V. vulnificus* in 50 mM Tris-Acetate (pH 7.5) and 50 mM NaCl. Then, 1 mM of this product was used with 13 mM L-glutamate, 1.5 mM pyridoxal 5′-phosphate (PLP) and 2 µg of PglE enzyme to form the amino sugar. Finally, 1 mM of the WbjB product that was turned over by the PglE enzyme was added to 1 mM acetyl-CoA and 2 µg of PglD. Both PglE and PglD proteins were isolated with triton and concentrations were found using a Bradford assay. The reactions were monitored at the absorbance of 260 nm (for UDP) with an HPLC on a Zorbax-NH_2 (4.6 x 150 mm, 5 µm) with 250 mM ammonium acetate (pH 4.5).

Formation of AzBn-BPPs with UPPS

UppS enzyme from *S. aureus* (50 µM) was utilized in a reaction with AzBn-GPP (10 mM), IPP (6 mM), 25 mM Bicine (pH 7), 0.5 mM $MgCl_2$, 5 mM KCl, 2% OTG in a total volume of 200 µL. The reaction was analyzed by reverse phase HPLC with an Eclipse XDB-C18 (9.4 x 250 mm, 5 µm) column and a gradient of 2.2% n-propanol per minute for 36.9 minutes with 25 mM ammonium bicarbonate as the aqueous component.

Cleavage of the AzBn-BPPs with UppP

Both of the UPPS reactions from above (0.5 mM), unlabeled and TAMRA-labeled with were utilized by UppP (1 µg, cef) in a solution of 50 mM HEPES (pH 7), 150 mM NaCl, 200 µM $CaCl_2$, and 10 mM sodium cholate. The reaction was analyzed by reverse phase HPLC with Zorbax C18 (4.6 x 50 mm, 3.5 µm) column and a gradient of 2.2% n-propanol

per minute for 36.9 minutes with 25 mM ammonium bicarbonate as the aqueous component.

Huisgen-Cycloaddition of the AzBn-BPPs with TAMRA- or Biotin-DBCO

The UppS/UppP reaction from "Cleavage of AzBn-BPPs with UppP" was taken and labeled with 2x excess (by moles) of the Biotin-PEG_4-DBCO (Sigma) or TAMRA-PEG_4-DBCO (Sigma). The TAMRA-labeled BPs and BPPs and were isolated on an Eclipse XDB-C18 (9.4 x 250 mm, 5 µm) column with a gradient. The gradient began at 20% n-propanol, 80% 25 mM ammonium bicarbonate (pH 8.4) and increased by 5% n-propanol every 5 minutes for 70 minutes (modified to start at 30% n-propanol for the biotin-DBCO labeled BPs to decrease the isolation time to 60 minutes).

N-linked oligosaccharide enzymes with 2CNA-BPs

2CNA-BPs were produced as previously described.[82] For the N-linked oligosaccharide biosynthesis pathway enzymes with 2CNA-BPs, the reactions in a total volume of 1 mL with 0.1% Triton, 10 mM $MgCl_2$, 100 mM Bicine (pH 8), 20 µM 2CNc4BP, 3 mM UDP-diNAcBac, 3 mM UDP-GalNAc, 2 mM UDP-Glc, and each enzyme added separately to this reaction mixture at 10 µg PglC (cef), 5 µM PglA, PglJ and PglH, and 30 µg of PglI (cef). The reactions were analyzed with a C_{18} HPLC column (Zorbax XDB-C18 5 µm, 4.6 mm x 50 mm) with a gradient method of 2.2% *n*-propanol per minute with the aqueous component of 100 mM ammonium bicarbonate at 1 mL/min.

N-linked oligosaccharide enzymes with TAMRA-AzBn-BPs

The reactions were set up in a total volume of 100 µL with 0.1% Triton, 10 mM $MgCl_2$, 100 mM Bicine (pH 8), 1 µM TAMRA-DBCO-AzBn-BP, 2 mM UDP-diNAcBac, 4 mM

UDP-GalNAc, 1 mM UDP-Glc, and each enzyme added separately to this reaction mixture at 5 µg PglC (cef), 2 µM PglA, PglJ or PglH, and 30 µg of PglI (cef). The reactions were analyzed with a C_{18} HPLC column (Zorbax XDB-C18 5 µm, 4.6 mm x 50 mm) with a gradient method of 2.2% *n*-propanol per minute with the aqueous component of 100 mM ammonium bicarbonate at 1 mL/min.

N-linked oligosaccharide enzymes with Biotin-AzBn-BPs

The reactions were set up in a total volume of 10 µL with 0.1% Triton, 10 mM $MgCl_2$, 100 mM Bicine (pH 8), 10 µM Biotin-DBCO-AzBn-BP, 2 mM UDP-diNAcBac, 4 mM UDP-GalNAc, 1 mM UDP-Glc, and each enzyme added separately to this reaction mixture at 5 µg PglC (cef), 2 µM PglA, PglJ or PglH, and 30 µg of PglI (cef). The reactions were analyzed with a C_{18} HPLC column (Zorbax XDB-C18 5 µm, 4.6 mm x 50 mm) with a gradient method of 2.2% *n*-propanol per minute with the aqueous component of 100 mM ammonium bicarbonate at 1 mL/min.

Colanic acid enzymes with TAMRA-AzBn-BPs

The reactions were set up in a total volume of 100 µL with 10 mM sodium cholate, 10 mM $MgCl_2$, 100 mM Bicine (pH 8), 1 µM TAMRA-DBCO-AzBn-BP, 4 mM UDP-Glc, 80 µM GDP-Fucose, 4 mM acetyl-CoA, 1 mM UDP-Gal, 2 mM UDP-GlcA, and each enzyme added separately to this reaction mixture at 10 µg CPS2E (cef), 5 µM WcaI, WcaF, WcaE, WcaC, or WcaB, and 10 µg of WcaA (cef). The reactions were analyzed with a C_{18} HPLC column (Zorbax XDB-C18 5 µm, 4.6 mm x 50 mm) with a gradient method of 2.2% *n*-propanol per minute with the aqueous component of 100 mM ammonium bicarbonate at 1 mL/min.

Colanic acid enzymes with Biotin-AzBn-BPs

The reactions were set up in a total volume of 10 µL with 10 mM sodium cholate, 10 mM $MgCl_2$, 100 mM Bicine (pH 8), 10 µM Biotin-DBCO-AzBn-BP, 4 mM UDP-Glc, 80 µM GDP-Fucose, 4 mM acetyl-CoA, 1 mM UDP-Gal, and each enzyme added separately to this reaction mixture at 10 µg CPS2E (cef), 5 µM WcaI, WcaF, WcaE, and WcaC. The reactions were analyzed with a C_{18} HPLC column (Zorbax XDB-C18 5 µm, 4.6 mm x 50 mm) with a gradient method of 2.2% *n*-propanol per minute with the aqueous component of 100 mM ammonium bicarbonate at 1 mL/min.

Isolation of the 2CNA-BP C. jejuni N-linked oligosaccharide Pathway Products

The final 2CNc4BPP-heptasaccharide (PglI product) from the reactions prepared above were isolated on an Eclipse XDB-C18 (9.4 x 250 mm, 5 µm) column with an isocratic 20% n-propanol, 90% 25 mM ammonium bicarbonate (pH 8.4). The product was collected in glass tubes, dried down and taken up in 200 µL of 20% n-propanol, 80% 25 mM ammonium bicarbonate (pH 8.4). Then 5% of NH_4OH (v/v) was added and it was analyzed by ESI-MS.

ESI-MS Analysis of the 2CNA-BP C. jejuni N-linked oligosaccharide Pathway Product

ESI-MS analysis was done on a Thermo VELOS Pro Dual-Pressure Linear Ion Trap instrument using ESI introduced by infused flow with 50:50 acetonitrile : water along with the sample into the ESI apparatus at 10 µL/min. Negative ion mode was used with a capillary temperature of 150°C and a spray voltage of 3.9 kV. Once a signal was obtained, MS^n with collision energy of 35 kV was used to select for and fragment the m/z

of interest. The expected mass was 1053.45 m/z (for the $[M-2H]^{2-}$) and 1053.43 m/z was obtained.

CHAPTER 4: THE SELECTION AND ANALYSIS OF BACTERIAL SUGAR SPECIFIC APTAMERS FOR CAMPYLOBACTER JEJUNI N-LINKED OLIGOSACCHARIDE AND COLANIC ACID FROM ESCHERICHIA COLI

Short Synopsis

In this Chapter the primary goal was to select for agents with specificity to specific bacterial glycans. I chose to use DNA aptamers as the agent and the glycans of interest are colanic acid from *E. coli* and the N-linked oligosaccharide from *C. jejuni.* In this chapter I will describe the process of selecting for DNA aptamers for the two glycans. In this work, two different selection methods were utilized, one where the glycan is produced *in vitro* and immobilized, and the other is a whole cell selection experiment where the glycan of interest is either present or knocked out. Secondly, I will discuss the preparation of the selected DNA pools for analysis with both flow cytometry and fluorescent plate reader assays to probe the specificity of the aptamer pools to the two glycans mentioned. Additionally, I will discuss the sequence identities of candidates from the aptamer pools and their interaction with the glycans of interest. Finally, I will describe the need for glycan specific detection in biosynthetic pathway intermediates and how this work can begin to fulfill that need.

Introduction

Bacterial surfaces are a target rich environment composed of glycan molecules such as capsular polysaccharides, exopolysaccharides, and oligosaccharides, which vary significantly between species. These diverse surface glycans are important because they aid in the protection of the microbes from their environment, as well as enhance the virulence of some bacteria.[8, 105, 108-109] The structural difference of the glycans provide a distinction between bacterial surfaces in the environment and can provide a handle for targeting a specific bacterium. Therefore, the development of agents specific to these surface glycans would provide the material needed for detection and targeting of specific types of bacteria.

The targeting agent of interest in this study is that of a DNA aptamer. Aptamers are made up of single-stranded DNA or RNA, which proves to be a versatile tool and provides multiple advantages to that of antibodies.[110-111] Although antibodies have been at the forefront of the biomedical field, one disadvantage is the need of the immune system of a living animal for the production of the antibody. The use of a living animal drives the expense associated with antibody production, which includes animal upkeep. This expense is much more than that of aptamer production, specifically with DNA-based aptamers, and there is no *in vitro* alternative to animals for new antibody production.[112] Furthermore, the versatility of an aptamer stems from the fact that one can be made for a variety of targets, without an immune system stimulating agent, such as with antibodies. The timescale for the production of aptamers is also much shorter where they can be produced in a week (once methods are well established) compared to months for antibodies.

Single stranded DNA libraries can be selected using the SELEX method to isolate sequences specific for a target of interest. SELEX is a method of selecting for an agent specific for a target through a series of steps.[110] These steps include first introducing your agent to your target, generally with the target immobilized for ease of use. Then the agents that are not involved in binding initially are 'washed' away or otherwise separated from the material that is interacting with the target. From there, the partitioning of the bound agents from the target is needed in order to obtain the selective material, or 'pool'. After that the bound material is amplified to produce more of the same agents that were bound to the target. Finally, the process repeats enriching the 'pool' until there are agents that show specificity to the given target. In addition to selections with the target, there are generally ways to prevent non-specific binding and binding to similar materials to the target by introducing other proteins and similar materials to the target.

Typically when developing aptamers for cell surface materials this methodology focuses on selecting DNA or RNA molecules that interact with cells that produce the material, and negatively selected against cells that do not.[113-116] However, this approach is limiting in that one must be able to readily culture the particular cell of interest, and the organism must constitutively produce the target material, or the production of that material must be induced in the targeted cells. Cell surface materials that are only transiently produced cannot be targeted using whole cell selection very effectively, and therefore it is even difficult to know when and how readily any such materials are produced at any given time.

Although more challenging than cell-based targets, aptamers have been selected against small molecules, including some simple sugars and sugar-containing molecules.[63, 111] Since the bacterial sugar coating accounts for the stability and protection of many bacteria,

targeting of the sugars specific to bacteria would prove beneficial to destruction of the protective layer that aids in bacterial infection[9, 11], as well as detection of bacteria containing sugars specific to that particular microbe.

While this dissertation does describe whole cell selection of aptamers to specific bacterial surface glycans, one of the major advancements in this work is the ability to build cell surface glycans *in vitro* and to immobilize those materials, which can then be used to select for aptamers that allow the detection of the glycans in or on living cells. The methodology could be applied to any bacterial glycan in which the genes responsible for the glycan are known, and function has been mapped (as in chapters 2 and 3). The tools generated here would allow for targeting specific glycans as an antimicrobial strategy, detection of these organisms in the environment, and could be used as research tools for detecting transiently produced materials.

In this dissertation I have focused my efforts on two distinct glycans produced by bacteria that highlight the two major potential uses of these materials for either targeting or detection of transiently produced molecular species. First, I have focused on the *Campylobacter jejuni* N-linked oligosaccharide described in chapter 1 and 3. *Campylobacter jejuni* infections are the most common causes of gastroenteritis in the world, and is commonly contracted from the handling of poultry. [4-6] Therefore, targeting or detecting these pathogens in the environment would be beneficial. Importantly, I focus on the development of aptamers to this oligosaccharide through the constitutive production of this glycan in a genetically modified *E. coli* strain. Using this system, I can test the ability to produce aptamers to specific bacterial glycans using both whole cell and *in vitro*

produced glycan strategies. After completion of this dissertation, plans are in place for the application of these aptamers to the detection of the glycan in *C. jejuni*.

The second glycan that I have targeted in this dissertation is the secreted polysaccharide colanic acid produced by *E. coli*. This glycan is produced only under specific growth conditions. However, the methodology used to detect its presence is not selective, and is ambiguous. A new reagent is desperately needed. In addition, recent work in the laboratory of one of our key collaborators, Dr. Kevin Young, University of Arkansas Medical Sciences, has led us to consider another important application for our aptamers. The Young group has recently proposed that many of the phenotypic properties of bacterial glycans may in fact be incorrectly assigned to the polysaccharide itself, and instead may be associated with a much more difficult to detect and measure species.[83] For example, his group has shown that deletion of genes in colanic acid biosynthesis leads to major changes in bacterial cell shape, which was ascribed to the build-up of colanic acid intermediates. [83] According to this work there is a finite amount of BP material, and when it is tied up in the production of the pathway intermediates, it distorts the cell overall. This distortion has been attributed to the sequestration of limiting cellular BP from cell wall biosynthesis including the essential peptidoglycan. By being able to produce specific reagents that could detect the build-up of such intermediate glycan materials I could finally quantify this contribution, along with ongoing efforts unrelated to my project to measure cellular BP levels.

Results

Acquisition of aptamer targets

For this dissertation I needed a set of targets for whole cell and immobilized glycan selected aptamer production. The production of biotinylated *E. coli* colanic acid tetrasaccharide and *C. jejuni* heptasaccharide were described in chapter 3, and were immobilized by taking the biotinylated glycans and introducing the glycans to streptavidin beads. The interaction between biotin and streptavidin is well known and used for a large number of bioconjugation reactions.[117-120] For whole cell selections against *C. jejuni* heptasaccharide I acquired *E. coli* strains from the laboratory of Dr. Christine Szymanski of the University of Georgia Complex Carbohydrate Research Center (Table 4.1). The strains were produced from *E. coli* K12/MG1655 where the N-linked oligosaccharide gene locus was inserted into the chromosome in place of a gene called *rfc* (O-antigen polymerase gene). The strain control for that experiment was an *E. coli* K12/MG1655 strain with a kanamycin cassette in place of the *rfc* gene. For attempts to produce colanic acid aptamers (not described in this dissertation) and as targets for colanic acid tetrasaccharide selected aptamers I obtained single gene deletion strains of MG1655 *E. coli* from the laboratory of Dr. Kevin Young (Table 4.1). The strains were produced with lamda red recombineering that replaces the gene of interest with an antibiotic resistance cassette.[121] These deletions represented various points in the biosynthesis pathway for colanic acid.

Table 4.1. Strains for N-linked oligosaccharide whole cell aptamer experiments.

Strain	Description
K12 +pgl	*E. coli* with N-linked oligosaccharide loci in the genome (of *rfc* gene)
K12 *rfc::kan*	*E. coli* with Kan resistance in the *rfc* gene; parent strain of +pgl

Table 4.2. Strains used in the colanic acid aptamer experiments.

Strain	Description
MG1655	Common laboratory strain of *E. coli* (same K12 but with no Kan resistance) F- λ – *ilvG rfb- 50 rph-1* (No O-antigen, or bacteriophage F or λ, decreased *pyrE* expression)
DR38	MG1655 Δ[*wza – wcaM]::frt* (missing colanic acid gene loci)
ΔWcaA	MG1655 Δ*wcaA::frt*
ΔWcaC	MG1655 Δ*wcaC::frt*

Aptamer library design

The ssDNA aptamer library was prepared with a 20 nucleotide randomized region with two flanking 20 nucleotide known DNA sequences. Primers were prepared for these 20 nucleotide known regions in order to amplify the ssDNA that was associated with the targets (**Table 4.3**). The melting temperatures (T_m) of the primers were kept at similar

values (58 ºC) and checked for interactions with each other. There were no interactions found with the Multiple Primer Analyzer (Thermo Fisher Scientific) [122] between the two primers used, therefore significant binding between just the primers is unlikely. There is one predicted self-dimer with the forward primer, but this was minimized from three interactions to one interaction by modifying the G/C content but keeping the T_m the same. Other primers that were utilized in this work include the FAM-labeled forward primer and the phosphorylated reverse primer. Both of the sequences for the forward and reverse primer were the same with an exception of the 5' of the forward was modified with a 6-carboxyfluorescein (FAM). The reverse primer was modified at the 5' with a phosphate for use with lambda exonuclease, which favors the phosphorylated 5' DNA strand in a digestion reaction.

Table 4.3. ssDNA Library construction and primers. The ssDNA library was constructed to have 20 randomized nucleotides between two known regions that were matched with primers for amplification of the unknown aptamer sequences.

Library
5'-GGATACGCAGTCAGTGGCTA –N_{20}- TCGTCGTGTCTCGACTCTAC -3'
Forward Primer
5'-GGATACGCAGTCAGTGGCTA-3'
Reverse Primer
5'-GTAGAGTCGAGACACGACGA- 3'

Optimization of Library and Aptamer Pool Polymerase Chain Reactions

To obtain aptamer pools, first the synthesized ssDNA library (IDT Technologies) was incubated with the target of interest (the whole cells or the immobilized glycans from chapter 3). Whole cells and immobilized glycans were washed using 'wash buffer' (1X PBS, 1.4 mM $MgCl_2$) to remove weakly binding DNA. Next, to amplify the bound DNA by the polymerase chain reaction (PCR), either whole cells with DNA bound were used as substrate or DNA that was ethanol precipitated off of the immobilized glycan was used for amplification. This reaction amplifies the DNA pool to use for the next selection.

Various sources describe the production of non-specific amplicons when undergoing PCR with a large aptamer library (~ 10^{14} different ssDNAs) and to keep the cycles under

12 for the least non-specific amplicons.[116] Therefore, the PCR conditions for the purchased ssDNA library and the following selection pools were optimized for both reaction conditions and PCR cycles. In separate experiments with all other variables held constant, the concentration of the primers were varied from 0.5 to 2 µM, the dNTPS were varied from 0.1 to 0.5 µM, the template concentration was varied from 0.1 to 1 µg of DNA (according to the Nanodrop spectrophotometer absorbance at 260 nm, Thermo), and the addition and exclusion of DMSO in the reaction mixture was analyzed (data not shown). The optimal production of amplified pool DNA was determined using a band at 60 base pairs with native polyacrylamide gel electrophoresis (PAGE), which is the expected size of the PCR product based on the ssDNA library size, including the primer regions. Once the optimized reaction conditions were obtained, the PCR cycles were varied from 7 to 20 cycles (**Figure 4.1**). A second band begins to be formed around 11-13 cycles so 9 cycles was selected for the remainder of the PCR reactions. In order to label the PCR pools with a fluorescent moiety to make it detectable without staining, a Fluorescein (FAM) primer was used in place of the forward primer. Optimization with the FAM primer was also carried out the same as above, and the use of 20 cycles was needed for the addition of the FAM primer (**Figure 4.1**).

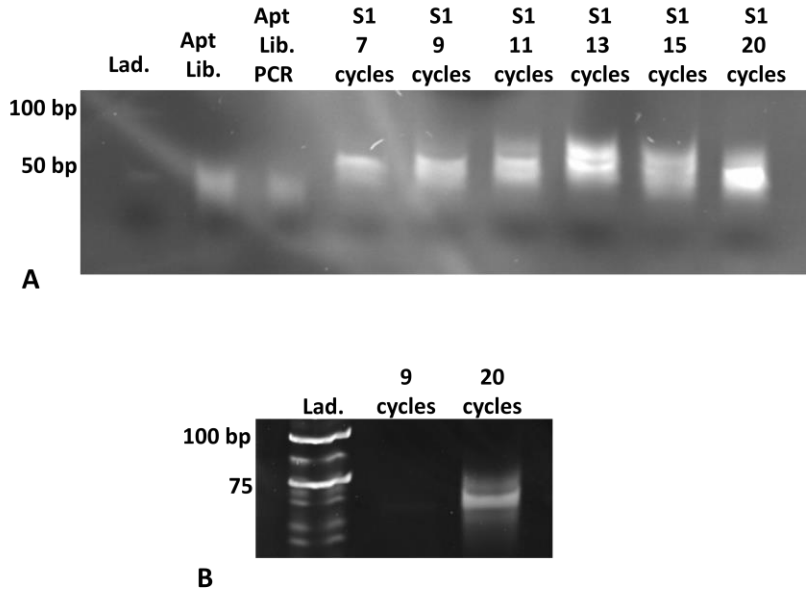


Figure 4.1. Native-PAGE gel of the DNA aptamer PCR optimization of the number of cycles. The number of cycles utilized for the unlabeled PCR experiments were 9 cycle based on the least number of bands present combined with the least amount of time for the experiment (a). For the FAM and phosphorylated primers, the cycles were extended to 20 (b).

Production of ssDNA from selection PCR products

In order for DNA to be considered an effective aptamer and bind to a biomolecule rather than to another DNA strand, it needs to be single-stranded. There is a significant variability in the literature as to how to produce ssDNA after the PCR amplification of the selected pool. Production of ssDNA either by heat denaturation, chemical denaturation, the use of streptavidin bead removal via a biotinylated second strand, or the use of a phosphorylated second strand and the use of lambda exonuclease to degrade one strand of the dsDNA have been proposed.[115, 123-124] For our purposes I have utilized the lambda exonuclease (New England Biolabs) method to render the PCR products single stranded and confirmed by

Native-PAGE (**Figure 4.2**). This is made possible with the use of a phosphorylated reverse primer during PCR, which is the strand that the exonuclease preferentially digests. With exonuclease treatment, a second lower molecular weight band is also present on the Native-PAGE gel slightly under the PCR product, which is consistent with literature.[125] The presence of this band only after the addition of the lamda exonuclease indicates that ssDNA is present from the PCR product that had been produced from the template DNA.

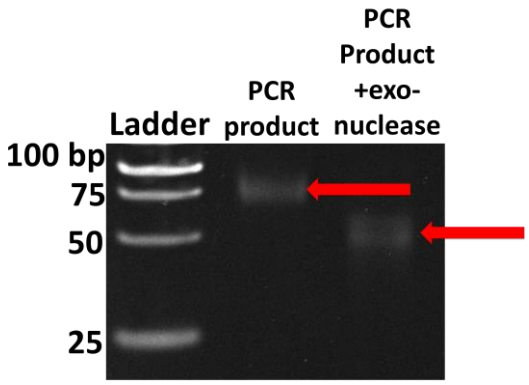


Figure 4.2. Native-PAGE gel of the DNA aptamer library PCR product with lamda exonuclease to form ssDNA. Subjecting the PCR product to lambda exonuclease displays the presence of a band at a lower molecular weight. The red arrows indicate where each band is located for the PCR product and the ssDNA product.

The expected PCR product size is associated with the immobilized glycan(s) and whole cells throughout the selection processes

The ssDNA library was prepared with a 20 nucleotide randomized region with two flanking 20 nucleotide known DNA sequences. Primers were prepared for these 20 nucleotide known regions in order to amplify the ssDNA that was associated with the sugar

target(s). Throughout the selection process the obtained PCR product from the target(s) was around 60 base pairs when compared to a low molecular weight DNA ladder (Thermo Fisher Scientific). This product was only formed when both primers were present and coincides with the size of the introduced aptamer pool. This product size was seen throughout all of the selection processes described here, the immobilized colanic acid tetrasaccharide and heptasaccharide selections, and the whole cell selections with an *E. coli* strain containing the N-linked oligosaccharide. Throughout the cell-based and immobilized glycan-based selection processes the native-PAGE gel analysis has shown the preservation of a 60 base pairs PCR product (**Figure 4.3**). This indicates that there is DNA associated with the cell fraction or the glycan-linked beads after stringent washing and multiple selections.

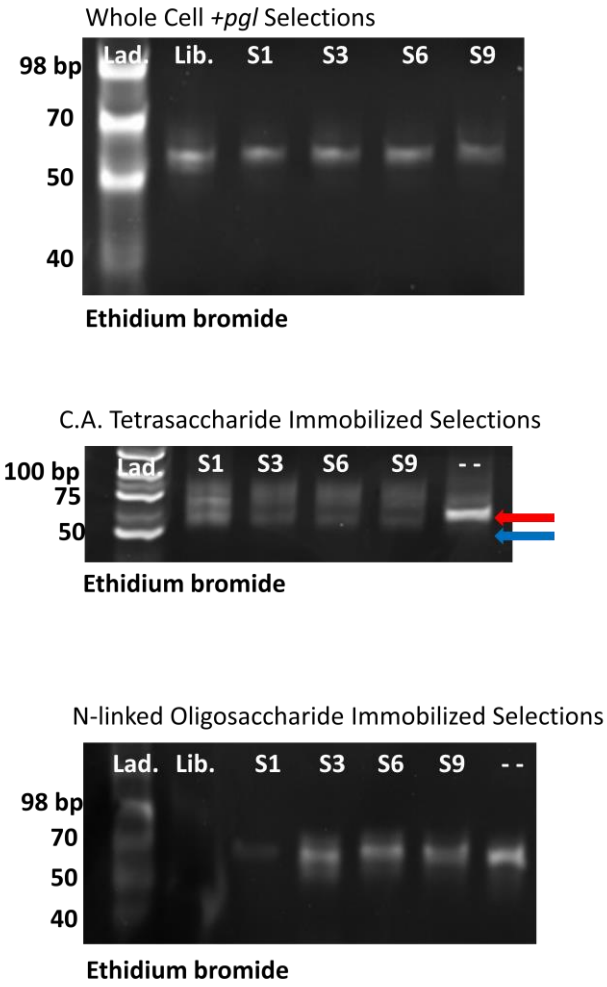


Figure 4.3. Native-PAGE gel of the DNA aptamer pools over the course of the selections. For each of the whole cell selections for the N-linked oligosaccharide containing *E. coli*, and the immobilized colanic acid tetrasaccharide and N-linked oligosaccharide selections, a band is clearly visible that corresponds to the expected PCR product of 60 base pairs designed from a ssDNA library through selections 1, 3, 6, 9 and the final pool after two negative selections (or the BP loaded beads for the immobilization method, - -). The red arrow represents a FAM-labeled band with a slightly higher molecular weight than the unlabeled (blue arrow).

Flow Cytometry of Bacteria

To test for interactions between our aptamers and targeted bacteria I chose to utilize flow cytometry, which allows us to rapidly distinguish between cells bound and unbound to a fluorescent aptamer. However, there was one major challenge associated with this that relates to the relative size of prokaryotes. Flow cytometry instrumentation at UNC-Charlotte is primarily designated for use with mammalian cells which are 1-2 orders of magnitude larger than bacteria. It is in fact difficult on this instrumentation to distinguish between background particulates and our bacterial cells. I therefore needed to find a way to distinguish the bacteria from particulates in flow cytometry. This was done with the use of a blue fluorescent protein (BFP) gene controlled by a constitutive promoter originally placed into *Salmonella enterica* and transferred to our *E. coli* strains for this project.[126-127] BFP expression was transduced into the *E. coli* strains of bacteria utilized in the selection process containing the N-linked oligosaccharide using a P1 bacteriophage. The expression of BFP allowed for ready differentiation of intact cells from particulate background, which enabled us to easily gate on the bacteria and omit any background particulates.

The N-linked oligosaccharide aptamer pools display enrichment as the selections proceed

Beginning with the ssDNA library I expected to see an enrichment of the aptamer pools to the target cells as the selections progress. The aptamer pools obtained throughout the selection process were amplified with the FAM primer in PCR and equal concentrations were incubated with the N-linked oligosaccharide containing cells and *E. coli* cells. These samples were then analyzed by flow cytometry for the mean fluorescence of the FAM

label (e.g. the FAM-labeled aptamer) over the background cells (*E. coli*). As the selections progressed, the fluorescent signal by flow cytometry of the selection pools increased with increasing selections. Specifically, a significant difference was observed between the beginning library to the fourth, eighth, and final selection pools, respectively (**Figure 4.4**). Negative selections were included at the end of the tenth selection to try to increase the enrichment further (e.g. final). Additionally, the immobilized aptamer pool also showed an increase in the fluorescence as the selections progressed with the whole cells (**Figure 4.5**). Specifically, there is a significant difference between the beginning library and the first, third, sixth, ninth and final pools for the immobilized glycan method for the N-linked oligosaccharide containing cells. The flow cytometry data for both the whole cell and immobilized glycan methods for the N-linked oligosaccharide provide evidence of enrichment of the aptamer pools to the glycan target as the selections proceed.

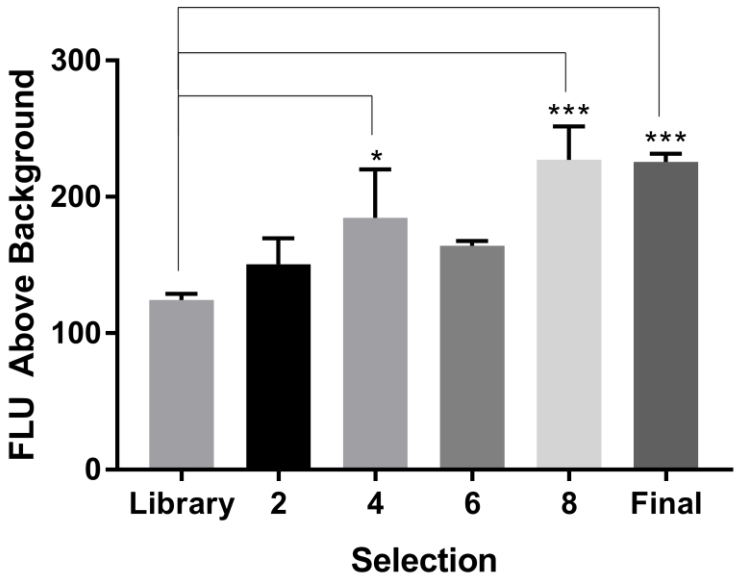


Figure 4.4. Flow Cytometry of the selection process for the N-linked oligosaccharide whole cells. The fluorescently labeled (FAM) selection pools were incubated with BFP-labeled N-linked oligosaccharide containing cells. The enrichment increases as the selections progress based on the FAM fluorescence mean from the aptamer pool. The background is the signal associated with the aptamer pool with the N-linked oligosaccharide negative bacteria. The error bars represent the standard deviation of three replicates. One-way ANOVA; * $P \leq 0.01$; *** $P \leq 0.0001$.

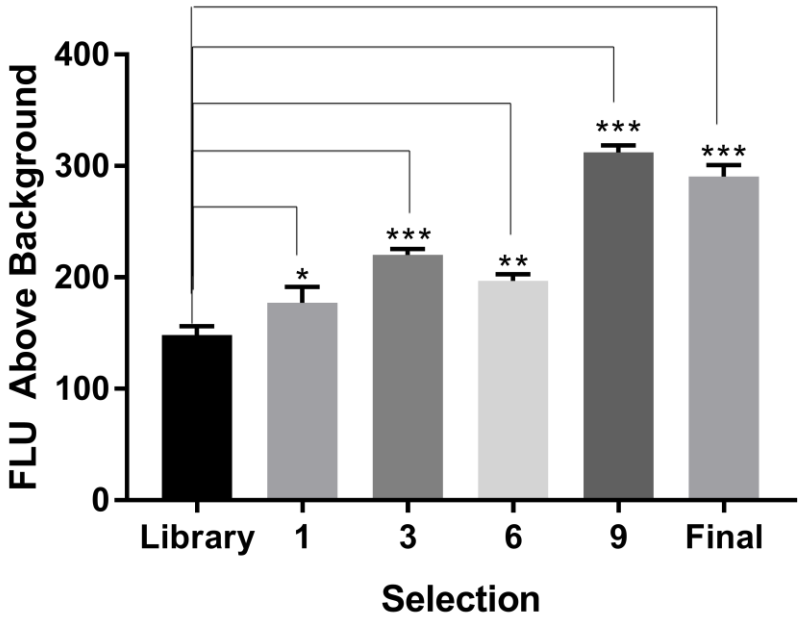


Figure 4.5. Flow Cytometry of the selection process for the immobilized N-linked oligosaccharide. The fluorescently labeled (FAM) selection pools were incubated with BFP-labeled N-linked oligosaccharide containing cells. The enrichment increases as the selections progress based on the FAM fluorescence mean from the aptamer pool. The background is the signal associated with the aptamer pool with the N-linked oligosaccharide negative bacteria. One-way ANOVA; * $P \leq 0.05$; ** $P \leq 0.001$; *** $P \leq 0.0005$

Whole cell aptamer pool and heptasaccharide aptamer pool bind 2-fold higher to N-linked oligosaccharide containing cells

Once I obtained a potential aptamer pool for further analysis, I utilized flow cytometry to determine the selectivity of these pools between the N-linked oligosaccharide containing cells and the background *E. coli.* For this I incubated the fluorescent aptamer pools with the heptasaccharide negative and heptasaccharide positive cells. What I found was that the aptamer pool FAM signal associated with the N-linked oligosaccharide containing bacteria

was consistently two-fold higher than the *E. coli* K12 control strain (**Figure 4.6**) based on the FAM mean fluorescence. This two-fold increase was derived from the mean total fluorescence of the FAM signal that was around 200 mean fluorescence units for the K12 *E. coli* N-linked oligosaccharide negative cells compared to 400 mean fluorescence units for the N-linked oligosaccharide positive cells (**Figure 4.7**).

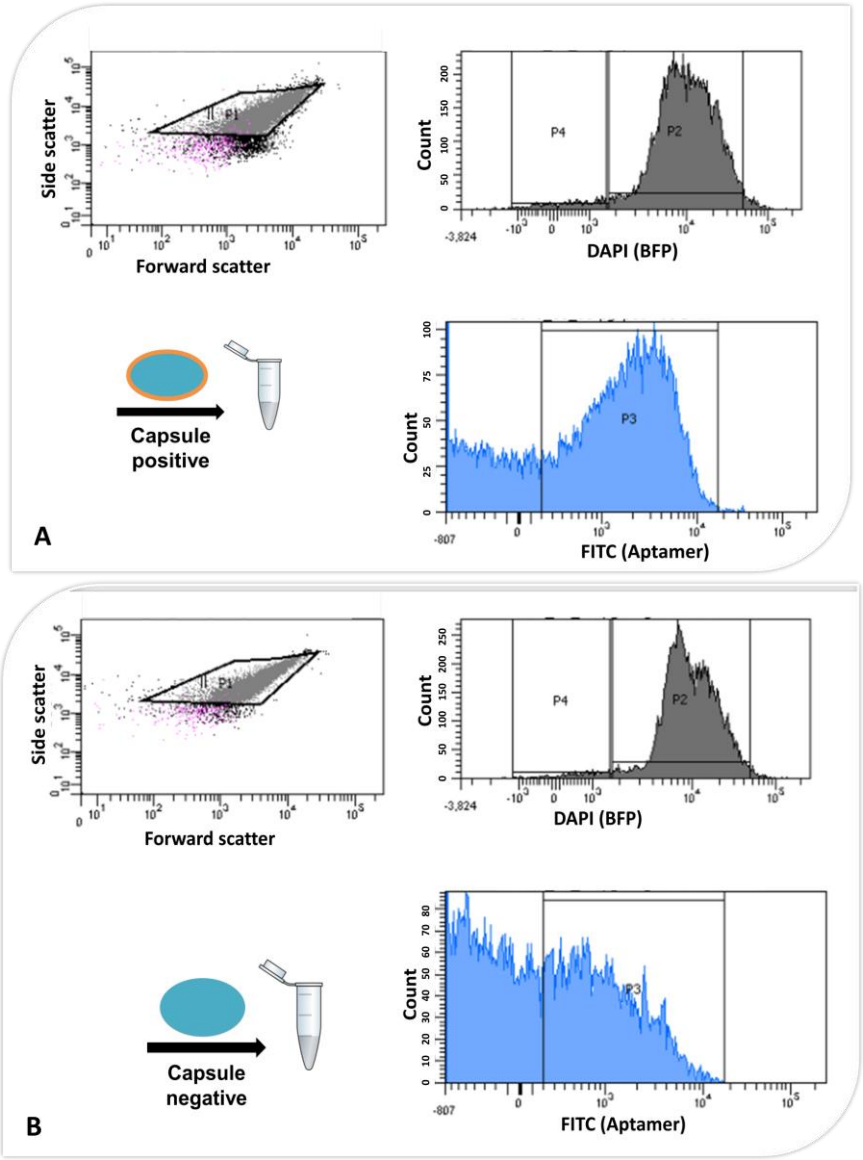


Figure 4.6. Flow Cytometry of the whole cell N-linked oligosaccharide. The fluorescently labeled (FAM) selection pools were incubated with BFP-labeled N-linked oligosaccharide containing cells (positive) or

BFP labeled (DAPI signal) *E. coli* cells (negative). The fluorescent signal from the FAM-labeled library is two-fold higher for the positive (~400 fluorescent units) (a) than the negative cells (~200 fluorescence units) (b) based on the FAM fluorescence mean associated with the labeled aptamer.

Bead-based N-linked oligosaccharide aptamers have similar selectivity to the cells as the whole cell aptamers

I was interested to see how the immobilized N-linked oligosaccharide aptamer pool compared to the cell-based aptamer pool to see if an aptamer can be selected for bacterial cells through the enzymatic formation of a glycan from that bacteria instead of utilizing the whole bacteria. This is an ideal methodology because cells would not be necessary and *in vitro* selected aptamers would be just as effective for whole cells as they would be for the individual glycan of interest. When the cell-based and immobilized based aptamer pools were compared using flow cytometry, I found that the bead-based aptamer did indeed have a similar selectivity (no statistical difference between) for the N-linked oligosaccharide containing cells compared to that of the whole cell aptamer (**Figure 4.7**). There is an increase in selectivity of the aptamer pools for both the immobilized and whole cell experiments between the *E. coli* containing the N-linked oligosaccharide (positive) and the background *E. coli* strain (negative) ($P \leq 0.05$). The FAM mean fluorescence is around 200 units for the negative cells and around 400 for the positive cells. There is not a statistical significance between the whole cells aptamer pool and the bead-based aptamer pool for the positive cells ($P = 0.5$), with an average of 373 and 488 fluorescence units, respectively.

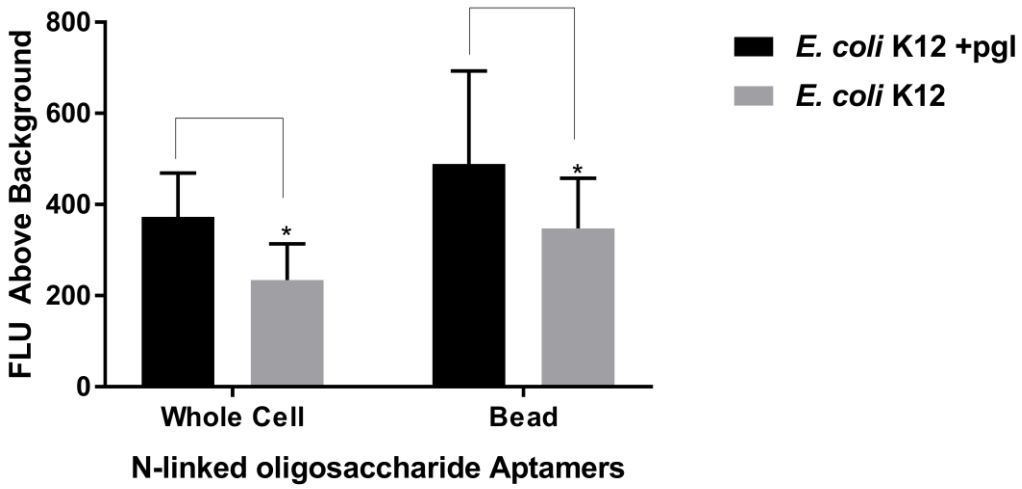


Figure 4.7. Flow cytometry data of N-linked oligosaccharide whole cell and immobilized aptamer pools. The N-linked oligosaccharide whole cell and immobilized bead-based aptamer pools were tested with N-linked oligosaccharide (positive) and *E. coli* (negative) cells that were BFP-labeled to determine the binding of the aptamer pool (FAM) to the N-linked oligosaccharide containing *E. coli* "+pgl" compared to K12 *E. coli*. The background pertains to the BFP-labeled cells without the aptamer present for each of the +pgl and *E. coli* K12. The error bars represent the standard deviation of three replicates. T-test; * $P \leq 0.05$

A Comparison of Specific aptamer sequences from the aptamer pools

Once the aptamer selections were carried out the final aptamer pools were cloned using TOPO cloning (Thermo Fisher Scientific) to determine some sequence identity of the aptamer pool. TOPO cloning is a technique in which DNA fragments are cloned into vectors without the use of DNA ligase. The Taq DNA polymerase adds a single deoxyadenosine (A) to the 3'-end of the PCR product.[128] Then, the inherent activity of the enzyme topoisomerase binds to a vector that encodes a specific binding sequence that ligates the DNA into the vector. The PCR product for each selection process, the whole

cell and immobilized N-linked oligosaccharide and the immobilized colanic acid tetrasaccharide, were amplified by PCR with Taq polymerase, incubated with the TOPO cloning vector, and transformed into TOP10 *E. coli* cells. After plating, colonies were picked to isolate the sequences that were inserted into the vector. Ten colonies were picked and the vector was isolated by Sanger sequencing (Eurofins). Out of ten picked colonies the cloning resulted in four sequences for the immobilized N-linked oligosaccharide, three for the colanic acid tetrasaccharide, and five sequences for the N-linked oligosaccharide containing whole cells (**Figure 4.8**).

Sequence analysis of the aptamer candidates indicated a few important properties. There was significant G and C content (50% or greater of the randomized region) throughout the aptamer sequences (orange and blue respectively, **Figure 4.8**), presumably resulting in more stable structures due to the larger amount of hydrogen bonding (three for a G/C compared to two for an A/T pair). Additionally, there is a conserved residue throughout the N-linked oligosaccharide aptamer clones at position 36/60 (**Figure 4.8**). Although only one residue is conserved, out of the four clones obtained for the immobilized experiment with this similarity, the two (c4 and c8) both have a higher selectivity for the N-linked oligosaccharide than the others (**Figure 4.10**). Based on the predicted secondary structures of these clones (NUPACK) [129], the clone 4 and 8 have that cytosine involved in the interactive strands of DNA compared to clone 7 and 8 (predicted to be in the loop region, **Figure 4.9**).

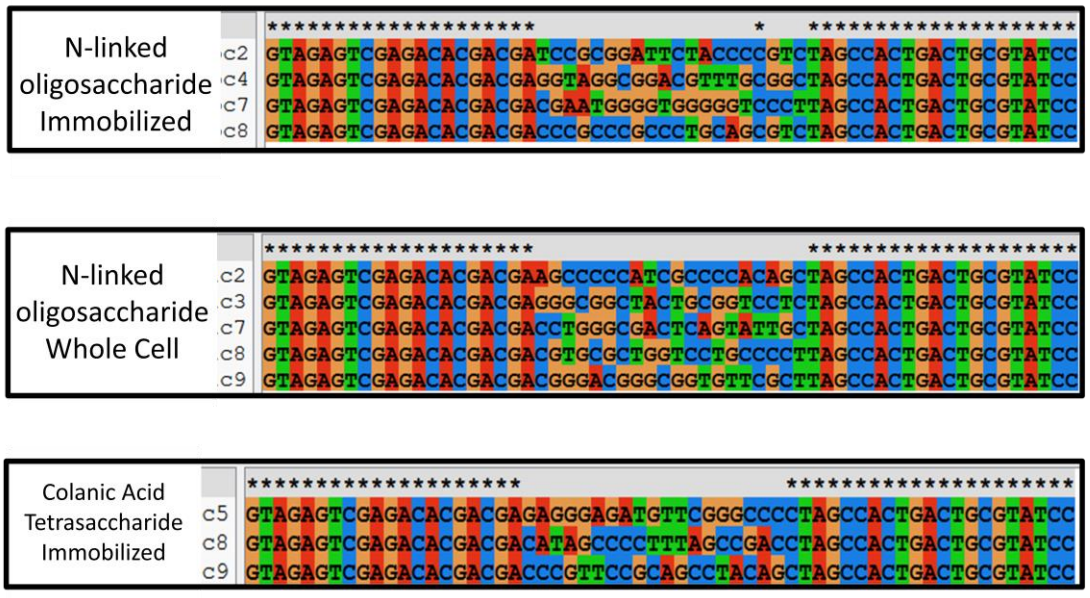


Figure 4.8. TOPO sequencing data of N-linked oligosaccharide whole cell and immobilized aptamer pools. The N-linked oligosaccharide whole cell and immobilized bead-based aptamer pools were cloned through TOPO sequencing and four sequences were obtained for the N-linked oligosaccharide immobilization selections, five for the N-linked oligosaccharide whole cell selections and three for the colanic acid tetrasaccharide selections.

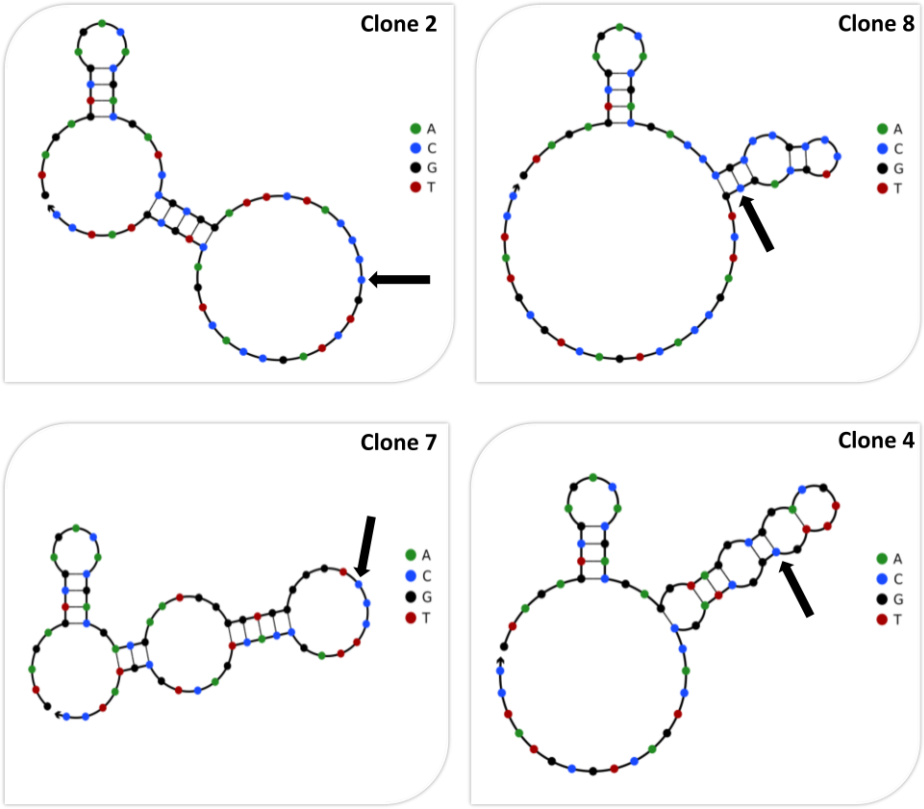


Figure 4.9. Predicted secondary structure of the immobilized N-linked oligosaccharide aptamer clones. The N-linked oligosaccharide immobilized clone sequences shared a cytosine residue indicated in all four clones with a black arrow. Clone 4 and clone 8 both have the cytosine residue predicted to be utilized in the secondary structure as an interacting base, whereas the other two clones have the cytosine predicted to be in a loop domain and not significantly contributing to the secondary structure.

To test the effectiveness of the aptamer sequences, aptamer DNA from the TOPO cloning was amplified with a fluorescent FAM primer and a phosphorylated reverse primer for digestion of this strand by lambda exonuclease to form ssDNA. These amplifications were then used to analyze binding to the bacterial strains containing N-linked oligosaccharide (above the background of the FAM fluorescent mean of the clones binding to the *E. coli* background cells). For the clones from the aptamer pools, candidates were desired that had a fluorescence mean above the library and more closely associated with the pool they were produced from. The higher performing aptamer candidates, based on a higher fluorescence associated with the N-linked oligosaccharide positive cells in the flow cytometry data (**Figure 4.10**) of the library (with an average of 120 fluorescence units), were the sequences from clone 4 and clone 8 from the bead-based N-linked oligosaccharide selections (with an average of 250 and 200 fluorescence units respectively) as well as clone 7 of the whole cell selection (an average of 200 fluorescence units). These specific sequences contain slightly more A/T (6/20 and 7/20 in the randomized region, respectively) compared to over 50% G/C content in the remaining aptamer candidates (**Figure 4.8**).

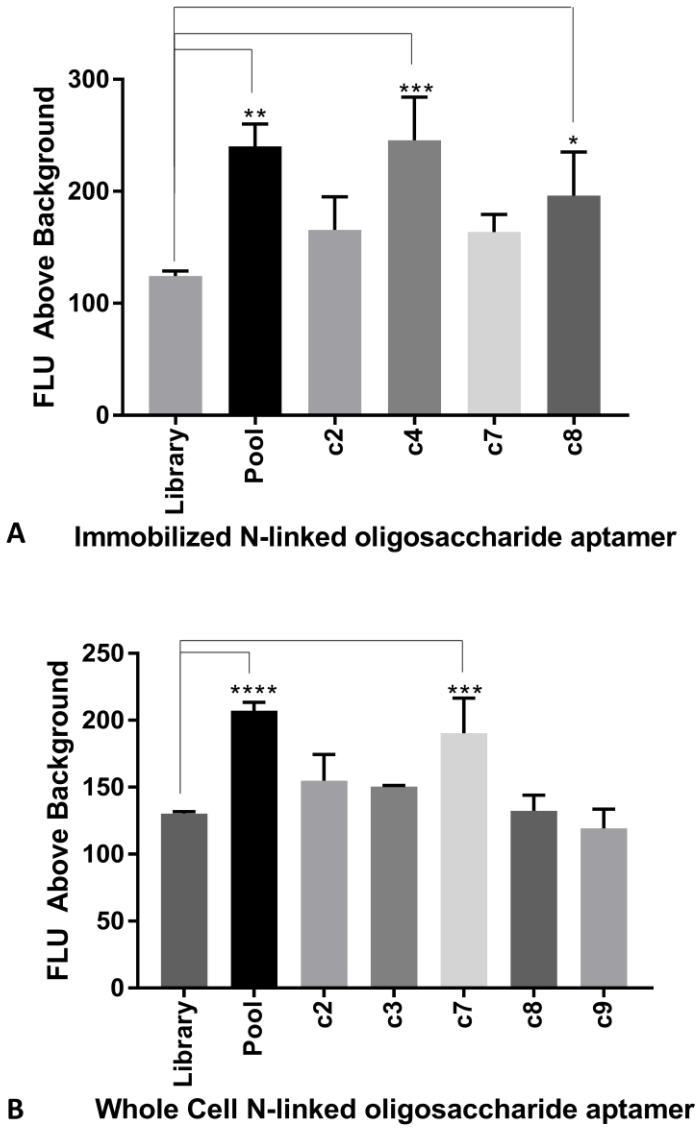


Figure 4.10. Flow cytometry data of immobilized N-linked oligosaccharide and whole cell aptamer candidates from sequecing. The N-linked oligosaccharide immobilized bead-based (a) and whole cell (b) aptamer pools were analyzed by flow cytometry and the mean fluorescence from the FAM-labeled aptamer clones was compared to the starting library, the final aptamer pool (after selections) and the individual clones from TOPO sequencing. The error bars represent three replicates. One-way ANOVA; * $P \leq 0.05$, ** $P \leq 0.001$, *** $P \leq 0.0001$ and **** $P < 0.0001$

Imaging N-linked oligosaccharide bound to a selected aptamer

I was next interested in whether aptamers could be imaged bound to the oligosaccharide producing cells. To do this I used the N-linked oligosaccharide whole cell clone 7 aptamer and the Amnis ImageStream®^{X} Mark II Imaging Flow Cytometer available for demonstration purposes in the UNC-Charlotte Biological Sciences Department. Based solely on the flow analysis, the labeling of the fluorescent aptamer clone 7 was only around 0.72% and 0.70% of the total cells for both the N-linked oligosaccharide containing cells and *E. coli* background cells respectively. As this analysis was during a demo and preceded the use of the BFP-labeled bacteria, previous gating on the bacterial population may have included other debris. From the imaging analysis, the 0.7% of that population was visually inspected and designated as single or multiple bacteria (aggregates). I also observed that the whole cell clone 7 aptamer candidate demonstrated binding to the cells containing the N-linked oligosaccharide and not to the K12 *E. coli* (**Figure 4.11**). Out of thousands of images of the FAM-labeled bacterial cells, there were only six positive FAM images from the N-linked oligosaccharide negative cells, with the remaining being on the N-linked oligosaccharide positive cells. Out of the six FAM-labeled N-linked oligosaccharide negative cells one of those six was an artifact, where the fluorescence was not associated with a bacterial cell in the microscopy image, and not truly a labeled cell (data not shown). From the 'aggregate' bacteria, minimal fluorescence was seen for the *E. coli* background cells (**Figure 4.11a**). More interestingly, the majority of the aptamer signal in the analysis was bound to populations of N-linked oligosaccharide cells, rather than to single bacteria. This potentially signifies that the N-

linked oligosaccharide helps the bacteria aggregate and the aptamer pool is able to concentrate on a population of the sugar coating via the bacterial aggregation.

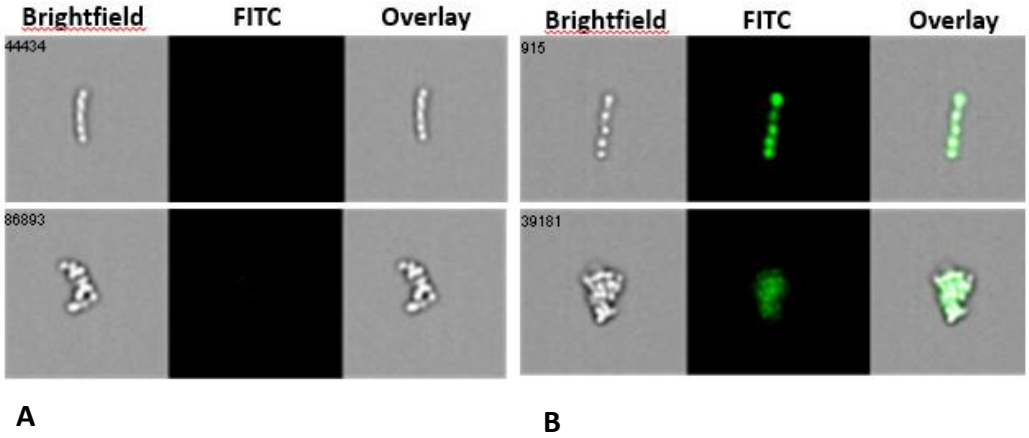


Figure 4.11. Flow cytometry imaging data. The N-linked oligosaccharide whole cell clone 7 was analyzed with the Amnis ImageStream®^{X} Mark II Imaging Flow Cytometer with *E. coli* cells (a) and N-linked oligosaccharide containing *E. coli* cells (b) and where the cells could be individually imaged and analyzed for fluorescence of the FAM-labeled aptamers.

Fluorescently Labeled Tetrasaccharide aptamer pools are able to detect colanic acid biosynthetic intermediates

As previously mentioned, the ability to quantify biosynthetic pathway intermediates is essential to understand how these materials are utilized in the cell, and what effects they have on the bacterial lifecycle. For this purpose I have selected for a DNA aptamer pool for the tetrasaccharide portion of the colanic acid exopolysaccharide repeat unit (See chapter 2 and chapter 3). From the immobilized tetrasaccharide selection process I produced an aptamer pool that was useful in the detection of the build-up of biosynthetic pathway intermediates that was monitored with a plate reader assay. As before, I labeled

the PCR product from the selection process with FAM during PCR, as well as the phosphorylated complement for digestion with lambda exonuclease for ssDNA. Then, I took different cells lines with mutations in the colanic acid biosynthetic pathway (Table 4.2) where mutations were located before and after the genes that produce the tetrasaccharide unit. For example, the Δ*wcaC* would be expected to provide a build-up of the trisaccharide intermediate and the Δ*wcaA* would be expected to provide a build-up of the tetrasaccharide unit.

The MG1655 strain that contains the full colanic acid exopolysaccharide, the DR38 strain that is a knock out mutant of colanic acid entirely, and the two Δ*wcaA* and Δ*wcaC* strains were grown and lysed. From here, the cell lysates were incubated with the same known concentration of the fluorescent aptamer pool obtained from the immobilized tetrasaccharide selection experiment. The cells were pelleted and the fluorescence of the supernatant was compared to the total DNA input. A decrease in the fluorescent signal indicates binding of the DNA aptamer pool to the cell lysates. When compared to MG1655 (contains the full colanic acid exopolysaccharide) and DR38 (a knock out mutant of colanic acid entirely) *E. coli* strains, the fluorescently labeled aptamer pool was able to reproducibly detect the intermediates in the colanic acid mutants of Δ*wcaA* and Δ*wcaC* respectively (**Figure 4.12**). Based off of the plate reader results, the colanic acid tetrasaccharide aptamer pool has equal selectivity for the tri- or tetra-saccharide unit, some selectivity for the full colanic acid polymer (MG1655), and little interaction with the strain with no colanic acid polymer (DR38).

The aptamer candidates from the aptamer pool for the colanic acid tetrasaccharide (candidates 5, 8 and 9) were also analyzed by plate reader for the selectivity of these candidates to biosynthesis pathway intermediates (**Figure 4.13**). Compared to the DNA control, a significant decrease is shown for the Δ*wcaC* mutant for both candidates 5 and 6 but not candidate 8. There were only three candidates obtained from the TOPO cloning of this pool, and two of the three show some selectivity for one of the mutants, but not for the Δ*wcaA* mutant expected. This may be due to the small number of candidates obtained and it is possible I have obtained aptamers more specific for the trisaccharide (Δ*wcaC*) over the tetrasaccharide (Δ*wcaA*) but there exists at least some aptamers in the pool that have some selectivity for both of the mutants.

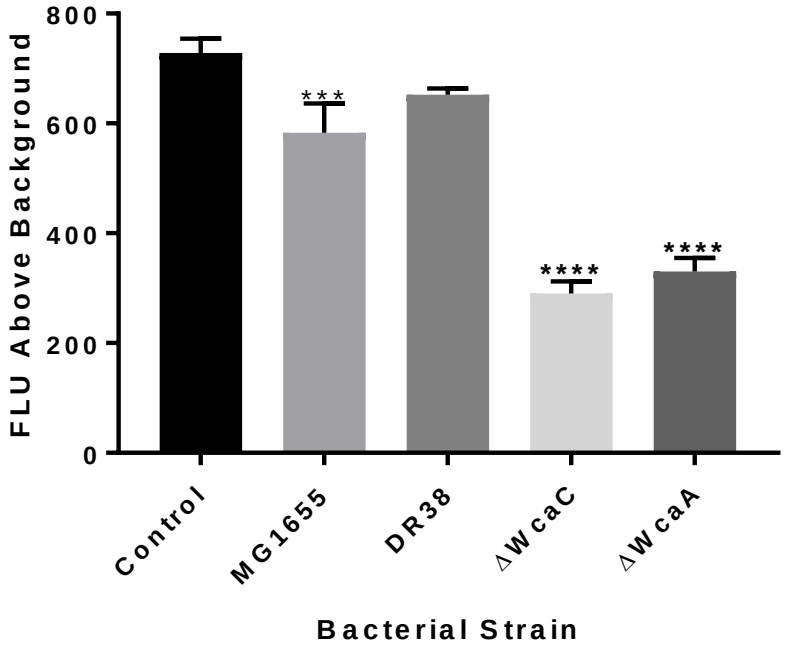


Figure 4.12. Plate Reader analysis of the fluorescently labeled aptamer pool for the colanic acid tetrasaccharide with different lysates representing variables of the colanic acid pathway. The MG1655

contains the whole colanic acid polymer, the DR38 contains no colanic acid polymer, the Δ*wcaC* and Δ*wcaA* are mutants of the colanic acid pathway that would produce more colanic acid tri- and tetrasaccharide respectively. The control is the DNA aptamer pool. Compared to the control, the measurements are of the supernatant after incubation with the lysates, the decrease in signal corresponds to the amount of aptamer associated with the cells. The error bars represent the standard deviation of three replicates. One-way ANOVA; *** $P \leq 0.001$ and **** $P \leq 0.0001$

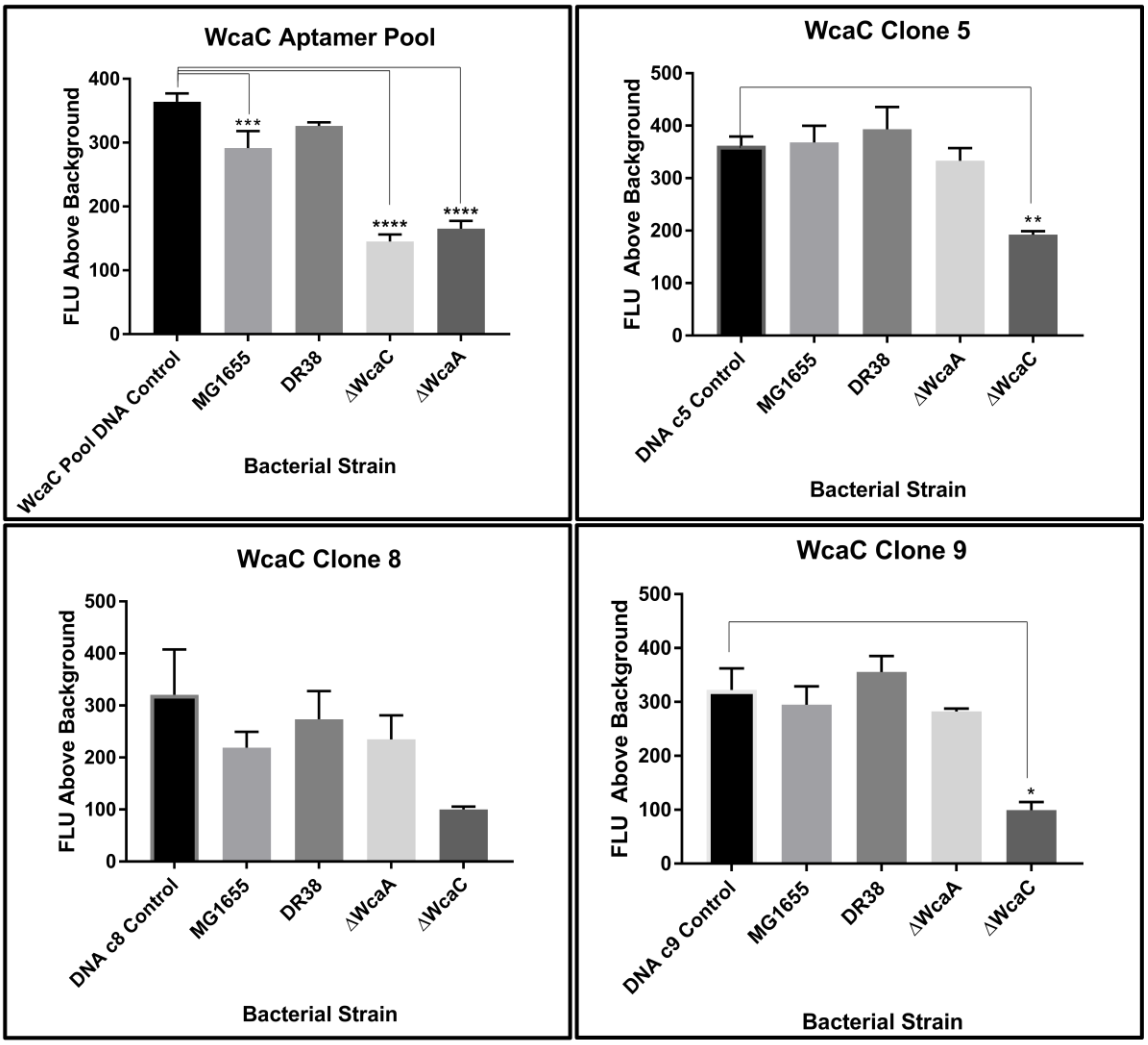


Figure 4.13. Plate Reader analysis of the fluorescently labeled aptamer pool and candidates for the colanic acid tetrasaccharide with different lysates representing variables of the colanic acid pathway. The MG1655

contains the whole colanic acid polymer, the DR38 contains no colanic acid polymer, the Δ*wcaC* and Δ*wcaA* are mutants of the colanic acid pathway that would produce more colanic acid tri- and tetrasaccharide respectively. The control is the DNA aptamer pool or candidate. Compared to the control, the measurements are of the supernatant after incubation with the lysates, the decrease in signal corresponds to the amount of aptamer associated with the cells. The error bars represent the standard deviation of three replicates. One-way ANOVA; * $P \leq 0.01$ and ** $P \leq 0.001$

Discussion

The goal this work was to find N-linked oligosaccharide and colanic acid specific DNA aptamers for use in detection of bacteria via the sugar coatings. The 'positive' bacterial strain is *E. coli* that was modified to contain the N-linked oligosaccharide on the surface of the bacteria.[5] *E. coli* does not naturally contain this oligosaccharide, therefore, this strain allows us to study and target the *C. jejuni* oligosaccharide while using a safe and easy to grow (aerobic, facultative anaerobe) *E. coli* strain in the laboratory. With the use of this strain compared to the K12 *E. coli* background strain, the ssDNA pool of 10^{14} molecules was selected for the oligosaccharide material in both a cell-based aptamer selection and an immobilized glycan selection. Additionally, the ability to construct the glycans *in vitro* onto an azide-bactoprenyl allowed for the selection of aptamers that would bind to the lipid-linked sugar unit without the cell background and without disturbing the enzymatic conformation of the sugar molecules (i.e. no chemical modification or harsh extractions). Importantly, this provides access to otherwise difficult to work with bacteria because you can produce the glycan for any type of bacteria as long as the enzymes are known or predicted for that biosynthetic pathway.

Based on flow cytometry, the aptamer pool was two-fold more selective for the N-linked oligosaccharide containing bacteria than for *E. coli* K12. Additionally, the immobilized glycan method competed well with the whole cell based aptamers and also provided a potential method for quantification of pathway intermediates in the case of colanic acid. Within that aptamer pool nine candidates were found through TOPO sequencing and out of those candidates, three showed similar selectivity to the N-linked oligosaccharide containing bacteria as the aptamer pool as a whole.

Currently, there is a need for a cell-based fluorescent probe for biosynthetic pathway intermediates. Generally only morphological effects of mutations in biological pathways, or the *in vitro* fluorescent assays with built intermediates of the pathways are utilized, and the ability to capture these intermediates in cells as they are occurring is useful to predict how the cells respond to outside stimuli and efflux of starting material for those pathways. Additionally, the ability to detect these intermediates leads to the possible quantification of these pathway materials to determine if there is a definitive number of starting material, such as the bactoprenyl diphosphate to the lipid-linked sugar molecules in the biosynthetic pathways of polysaccharides in bacteria. The colanic acid tetrasaccharide aptamer pool was able to be utilized in a plate reader assay to determine a build-up of the tetrasaccharide unit in cell lysates. Furthermore, candidates from the colanic acid tetrasaccharide showed selectivity for the trisaccharide build-up than for both the tri- and tetrasaccharide intermediates as the whole pool suggests. This may be due to the small number of candidates obtained from the aptamer pool and further work would need to be completed to extract those sequences that preferentially bind to either the tri- or tetrasaccharide.

Experimental Procedures

Selection of Escherichia coli Whole Cell targeting agents: N-linked Oligosaccharide

Whole cell bacterial cell systematic evolution of ligands by exponential enrichment (SELEX) was carried out with a randomized DNA library and two strains of *Escherichia coli.* The *E. coli* constructs were provided by Dr. Christine Szymanski (University of Georgia). The *pgl* locus was incorporated into the *E. coli K12* strain to replace the O-antigen polymerase gene (*rfc)* and provide the oligosaccharide at the surface of the bacteria. The *pgl* positive strain was used for the positive selection and the K12 *rfc::kan* strain was used for the negative selection. A total of nine positive selections with the *pgl* positive *E. coli* strain and four negative selections against *rfc::kan E. coli* were completed and PCR product was collected for each. The negative selections were done after selection 3, 6 and two of them after selection 9 to increase enrichment of the selected pool. The aptamer library was purchased from Integrated DNA Technologies (IDT) with 20 randomized nucleotides between two known regions, as follows:

5’- GGATACGCAGTCAGTGGCTA –N_{20}- TCGTCGTGTCTCGACTCTAC -3’

The two known regions are used for amplification of the DNA obtained from the selections with the polymerase chain reaction (PCR) with the following primers:

Forward: 5’- GGATACGCAGTCAGTGGCTA-3’

T_m= 58.3, CG= 55%, 20 nucleotides

Reverse: 5’- GTAGAGTCGAGACACGACGA-3’

Starter cultures of *E. coli* were grown overnight at 37 °C in Luria-Bertoni (LB) broth (50 μg/mL kanamycin). Then the subculture was diluted 1:100 in fresh LB/kanamycin and grown at 37 °C for 2-3 hours until an OD of 0.6 (which corresponds to $3x10^8$ colony forming units (CFU) per mL by plate count for both strains). The *E. coli* strain containing the *pgl* loci grew slower than the K12 *E. coli* strain and therefore had to grow an hour longer (3 hours) than the parent strain (2 hours) to obtain the same CFU/mL. The cultures were pelleted at 6,000 x g for 8 minutes and washed with 1X PBS for use in a selection. Meanwhile, 1-2 nmoles of the DNA library was diluted in wash buffer (1X PBS with 1.4 mM $MgCl_2$). The cells were incubated with the DNA library, or 1.5 μg (~300 nM) of DNA (PCR product from selections), at 25 °C with constant agitation (220 rpm) for 45 minutes in a total volume of 1 mL of binding buffer (1X PBS, 1.4 mM $MgCl_2$, 1% BSA). After incubation, the cells were pelleted at 8,000 x g for 6 minutes and washed once with 1 mL of washing buffer [1X PBS with 1.4 mM $MgCl_2$]. The washes were increased to two after five selections to increase the stringency of the selections. The pools were then amplified with the polymerase chain reaction (PCR) for further selections or for analysis.

Aptamer Pool Polymerase Chain Reaction (PCR)

The cell fraction containing the bound DNA pool was re-suspended in 150 μL 1X PCR buffer, and PCR was completed on this fraction. The reaction conditions for a 50 μL PCR reaction are: 4 mM $MgCl_2$, 0.5 μM of each primer, 0.4 mM dNTPs, 1.75 U Taq polymerase, and 10 μL of the collected fraction (~200 ng/μL on the Nanodrop at 260 nm). The thermocycler (Eppendorf MasterCycler Pro) conditions began with 94°C for 3 minutes of initial denaturing, and 9 cycles of denaturation at 94°C for 30 seconds, annealing at 51°C

for 30 seconds, extension at 72 °C for 20 seconds with a final extension of 72 °C for 5 minutes, and then it is held at 4 °C. The cycles were optimized to prevent unspecific amplicons from forming, which were seen after 13 PCR cycles with the unlabeled primers. With the incorporation of the FAM forward primer, the use of 20 cycles was needed. The PCR product was analyzed with a 8% Native-PAGE gel alongside a GeneRuler low molecular weight DNA ladder (50-1000 bp, Thermo Fisher Scientific) dyed with 6X loading buffer containing xylene cyanol and bromophenol blue dyes at 120 V for 25 minutes and stained with ethidium bromide before imaging (BioRad ChemiDoc MP Imaging System).

TOPO Cloning

TOPO (topoisomerase) cloning (Invitrogen, Thermo Fisher Scientific) was completed after the negative selections after selection 9 that produced several aptamer candidates. TOPO cloning was completed with a topoisomerase I- activated pCR4-TOPO TA Vector and TOP10 *E. coli* competent cells. The PCR product was produced with Taq polymerase, as specified previously, that provides a 3'A-overhang that fits into the TOPO vector, that has T overhangs, and with the associated topoisomerase in the vector, the PCR product is ligated into the plasmid and ready for isolation and Sanger sequencing. First, the TOPO vector reaction that consisted of a total volume of 6 µL with 4 µL of the PCR product (at 5 ng/µL), 1 µL of salt solution (1.2 M NaCl, 0.06 M $MgCl_2$), 1 µL of the pCR4-TOPO vector was incubated at room temperature for 30 minutes and then placed on ice. From there, the vector was transformed into TOP10 *E. coli* competent cells by taking 2 µL of the TOPO vector reaction into 1 vial (50 µL) of TOP10 chemically competent cells, incubating

for 30 minutes on ice, then heat shocking for 30 seconds at 42ºC, and cooling on ice for 2 minutes. The cells were recovered in 250 µL of Super Optimal Broth with catabolic suppression (SOC) media for 1 hour at 37ºC with constant agitation at 220 rpm. The recovered cells (10-50 µL) were plated in the presence of 50 µg/µL of carbenicillin for selection of the TOPO vector containing cells. A control was alongside with just the TOP10 cells (that has no antibiotic resistance) and with a pUC19 control vector that is the positive control with the TOP10 cells.

Selection of Immobilized targeting agents

The aptamer library utilized for the whole cell selection was also used for the azide analogue sugars. The enzymes from the known *C. jejuni* N-linked oligosaccharide biosynthesis pathway were used to build the heptasaccharide on the azide analogue or the enzymes from the colanic acid biosynthesis described previously (CPS2E, WcaI/F/E/C) were used to build the tetrasaccharide on the azide analogue. This azide-BP-tetrasaccharide or *C. jejuni* heptasaccharide was then attached to a streptavidin bead with a Huisgen-cycloaddition reaction of the azide analogue to a biotinylated octyne (ring-strained alkyne). A 1:1 mole ratio of the octyne to azide was used, with the detection limit of the azide analogue by UV-Vis (260 nm) being a minimum of 1 nmole. The beads (50 µL, 0.5 mg) were loaded with 10 µg of the biotinylated Azide-c6-BPP-Glc-AcFuc-Fuc-Gal, Az-heptasaccharide or the control Azide-c6-BP overnight with rotation at 4 °C in 300 µL total volume of bead storage buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Tween-20). The loaded beads were washed twice with bead storage buffer and ready for use.

The biotinylated Azide-c6-BPP-Glc-AcFuc-Fuc-Gal, Azide-heptasaccharide, or Azide-c6BPP loaded streptavidin magnetic beads (Thermo Fisher Scientific) were incubated with 0.02% glucose and 1 nmole of ssDNA aptamer library (or 300 µL PCR product, ≥ 400 ng/µL) in 300 µL total volume (or 400 µL for aptamer pool) of bead storage buffer for 45 minutes at 25 °C with constant rotation (200 rpm). The beads were then washed twice with bead storage buffer and the supernatant was saved for PCR analysis. The beads were then subjected to ethanol precipitation for obtaining the bound DNA off of the beads. To the bead fraction in 300 µL storage buffer, 30 µL (1/10 volume) of sodium acetate (pH 5.2) and 990 µL (3X volume) of 95% ethanol was added. The solution was incubated overnight at 4 °C. The next day, the solution was centrifuged at 14,100 x g for 30 minutes at room temperature. The supernatant was kept for PCR analysis alongside the DNA-containing fraction. The pellet was washed with 70% ethanol and centrifuged at 14,100 x g for 15 minutes at room temperature. The pellet was taken up in 100 µL of Tris-EDTA (TE) buffer. Using a magnetic tube stand, the remaining beads were removed from the fraction, re-suspended in 100 µL of bead storage buffer, and the supernatant was kept as the eluted DNA fraction. PCR was performed as above with the supernatant, the bead fraction, and the eluted DNA fraction.

Preparation of samples for and analysis by flow cytometry

PCR products were prepared like above for the aptamer pools and the candidates from TOPO cloning with a FAM-labeled forward primer. The fluorescent aptamers that were tested were as follows: the "final" aptamer pool that was amplified from the end of the eleven selections total, selections 2, 4, 6, 8 to check for enrichment, and the candidates

that were amplified from the TOPO cloning (five candidates obtained for the *pgl* whole cell experiment – candidates 2, 3, 7, 8, and 9; four for the immobilized heptasaccharide – candidates 2, 4, 7, and 8, and three for the immobilized colanic acid tetrasaccharide- candidate 5, 8 and 9 (that will be discussed further later). These candidates were analyzed by flow cytometry for binding to *E. coli* with and without the *pgl* loci in the genome (negative and positive) by looking at the mean total fluorescence of the FAM. In order to differentiate the bacteria from the debris, the bacterial strains were made to constitutively express blue fluorescent protein (BFP) via P1 phage transduction (described in detail in "Preparation of BFP producing E. coli for analysis with flow cytometry").

Glycerol stock starter cultures of the K12 *E. coli* strains (*pgl* positive and K12 control) were grown overnight at 37 °C in Luria-Bertoni (LB). The next day, the cultures were centrifuged at 8,000 x g for 8 minutes, washed with 1X PBS once and diluted to an OD of 1.0 in 1X PBS. 10-100 nM of FAM-labeled PCR products obtained from TOPO cloning were rendered single-stranded by lamda exonuclease before incubation in a 300 μL total volume with the diluted cells for 45 minutes at 25 °C at 220 rpm. The reaction was centrifuged at 8,000 x g for 8 minutes, washed with 1X PBS, taken up in 200 μL total volume and were directly analyzed by flow cytometry. Flow cytometry analysis was performed with a BD Biosciences LSRFortessa instrument analyzing side scatter and forward side scatter where 10,000 events were collected with gating on the population of BFP bacteria (PBS blank excluded). The blue signal from the BFP (380 nm excitation, 450 nm emission) containing bacteria was excited with a 405 nm laser and monitored by

the DAPI filter (450/50 nm) on the flow cytometer. The subpopulation of blue positive bacteria were analyzed for the FITC positive signal, or the signal from the aptamers.

Preparation of BFP producing E. coli for analysis with flow cytometry

The preparation of the BFP bacterial strain used in the P1 phage transduction was completed by Dr. C. Eade.

A previous construct of Blue Fluorescent Protein (BFP, Evrogen) served as the template for our BFP+ mutant. The previous construct was codon-optimized for expression in Enterobacteria and inserted into the *Salmonella enterica* subsp *enterica* serotype Typhimurium chromosome, preceded by a constitutive promoter and followed by an FRT-flanked *cat* gene conferring resistance to chloramphenicol (REFs). This strain served as template, and colony PCR was performed to amplify the construct (promoter, BFP, and *cat* gene) with primers that appended 40 bp homologous the *E. coli* genome directly flanking the *lacZ* gene (BFPinLacZ_Fwd and BFPinLacZ_Rev, sequences below). Using the method of Datsenko and Wanner, this amplicon was electroporated into *E. coli* MG1655 that had been transformed with pKD46 and induced with arabinose. Electroporated cells were selected on chloramphenicol, and colonies that grew were purified, and the insertion was verified by PCR using primers that spanned the insertion site (Blue-ck-2 and LacY_Rev, sequences below).

BFPinlacZ_Fwd:

ggaattgtgagcggataacaatttcacacaggaaacagctGAGCTGTTGACAATTAATCATCGGC

BFPinlacZ_Rev:

ttacgcgaaatacgggcagacatggcctgcccggttattaGATTGTGTAGGCTGGAGCTGCTTCG

Blue-ck-2: CTGGGCTGGGAAGCATTC

LacY_Rev: GGAAAAACGGGAAGTAGGCTC

Preparation of BFP producing E. coli by P1 Phage Transduction

The preparation of the BFP containing strains was completed by Dr. T. Williams.

The P1 Bacteriophage for *E. coli* was purchased as a frozen liquid stock (ATCC® 25404-B1™). First, the host strain (MG1655 *E. coli*) was grown to early log phase, meanwhile, 0.5 mL of LB was added to the liquid phage stock. Soft agar plates were prepared containing the MG1655 *E. coli* strain on LB. Then, the phage was serial diluted and 100 µL of the solution was plated. Lysis was visible within 24 hours at 37ºC. After 24 hours the soft agar is scraped off of the surface of the agar plates, centrifuged at 1000 x g for 25 minutes and the supernatant is passed through a 0.22 µm filter. This is stored at 4ºC for future use. To use the phage, a donor and accepter strain is needed. The donor in our case was the MG1655 that contains the BFP provided by Dr. Eade. This strain was grown in 5 mL of LB broth supplemented with 5 mM $CaCl_2$ overnight. Then, 100 µL of the culture was inoculated with 1 µL of phage stock at 37ºC shaking at 200 rpm for 20 minutes and added to 4 mL LB top agar and poured over LB agar. The plate was incubated at 37ºC overnight. The plate appears lysed if phage is working. The next day the plates were scraped into 5 mL of LB then lysed with 100 µL of chloroform was added and vortexed

for 1 minutes to ensure lysis. The culture was centrifuged at 5,000 x g for 10 minutes and then 13,000 x g for 10 minutes. The supernatant was collected and stored over chloroform at 4ºC. To make the acceptor strain, the acceptor strain was grown to log phase (pgl containing *E. coli* and K12 *E. coli)* at 37ºC shaking at 200 rpm with 5 mM $CaCl_2$. Then 100 µL of the acceptor was combined with 1 µL of the donor lysate at 37ºC shaking at 200 rpm for 20 minutes. Then 1 mL of LB with 5 mM sodium citrate was added and grown for 1 hour at the same conditions. The culture was centrifuged at 10,000 x g for 5 minutes and the pellet was plated on LB with 5 mM sodium citrate with selections for 50 µg/mL of kanamycin for the pgl modification and 20 µg/mL for the BFP. The colonies were re-streaked on plates of the same composition twice and the construct was confirmed by PCR (data not shown).

Lysate Plate Reader Assays

E. coli strains were grown overnight in LB with 20 µg/mL of chloramphenicol for the colanic acid mutants Δ*wcaA* and Δ*wcaC* and with 50 µg/mL kanamycin for the MG1655 (with CA operon; pET24a plasmid) and DR38 (no CA operon). The next day the cells were pelleted at 5,000 x g for 10 minutes, washed twice with 1X PBS, and taken up in a total of 1 mL of 1X PBS. The ODs were matched to 1.0 in a total volume of 1 mL, corresponding to 10^8 cfu/mL. The cells were lysed by sonication for three minutes (pulse one second on, one second off) at 25% amplitude. In a total volume of 100 µL, 10 µL of the 10^8 cfu/mL were incubated with 500 nM FAM-labeled aptamer PCR products of the pool or the candidates were rendered single stranded via lamda exonuclease (PCR with FAM forward primer and phosphorylated reverse primer). The cells were than pelleted at 8,000 x g for 6

minutes and the supernatant was checked for the remaining fluorescence signal compared to the DNA control. The fluorescent signal of DNA attached to the cells was calculated based on the DNA put into the experiment and the DNA that was left in the supernatant, giving the total fluorescent signal of DNA that was associated with the bacterial cell lysates. These experiments were repeated three times.

CHAPTER 5: CONCLUSIONS AND FUTURE DIRECTIONS

Colanic Acid Biosynthesis

E. coli is a known food-borne pathogen that affects many all over the world. Colanic acid is an exopolysaccharide that protects *E. coli* from low pH environments. Understanding how the bacteria produce this protective coating is beneficial in prevention and treatment of bacterial infections. Prior to this work the gene locus was known that is responsible for the production of colanic acid [33] and our lab has been interested in elucidating the functions of these genes in the production of the sugar repeat unit. To date, I have determined the roles of four glycosyltransferases and two acetyl transferases that are responsible for producing five of the six sugars of the colanic acid repeat unit. Currently there are only two proteins left in the pathway that are in need of characterization, one being WcaL, a predicted glycosyltransferase that is expected to transfer the last sugar from UDP-Gal. The other enzyme is WcaK, a predicted pyruval transferase that would modify the last added Gal sugar.

Our lab has already characterized one of the first pyruval transferases that is responsible for the pyruvulation of a BPP-linked disaccharide substrate in the biosynthesis of capsular polysaccharide A (CPSA) of *Bacteroides fragilis*.[77] Although the amino acid sequence similarity is small between these two enzymes (24%), they are both predicted to

have a similar function. Due to the specificity of WcfO for the BPP-linked disaccharide and not the lone sugar or the trisaccharide, it is possible that WcaK has yet to be functional because the preceding sugar addition has yet to be successful (the addition of Gal by WcaL). The characterization of these proteins are included in the future goals of the colanic acid biosynthetic project to complete the six sugar repeat unit for use in future studies.

Click-Enabled Detection and Immobilization of Glycans

The ability to detect complex glycan molecules is an area of research that is in dire need of expansion. The use of radiolabels are common practice in glycan research, but are difficult to utilize in everyday experiments. Presently, the use of chemically synthesized azide probes are popular and provide a handle for further conjugation of useful fluorophores or other chemical handles. Traditionally, the azide has been located on the sugars of interest and integrated into a pathway that is known to take that modified sugar. However, it is unlikely that these modified sugars can be universally accepted by highly specific glycosyltransferases and phosphoglycosyltransferases.

Within our lab I have focused on the anchor molecule that the sugars become attached to in the biosynthetic process. This anchor molecule, the bactoprenyl diphosphate, has been modified with an azide, allowing for the clickable functionality to be utilized but not interfere with the biosynthetic process. I have shown that I can successfully make the azide analogue, utilize the analogue to prepare it for the biosynthetic pathways with two enzymes, UppS and UppP, and click the azide analogue throughout the process. Additionally, I have shown that I can utilize the plain and unclicked azide analogue in the

known N-linked oligosaccharide biosynthetic pathway from *C. jejuni.* Furthermore, I have utilized the N-linked oligosaccharide and the colanic acid tetrasaccharide built onto the azide analogue to select for DNA that specifically binds to those sugar units. This was done through a clickable biotin functionality to the azide analogue after the sugars were added. This provides a very useful tool that allows glycans to be easily built and attached to surfaces for future analysis, such as in a resin or array to determine what proteins or other biomolecules bind to these bacterial sugars. One of the prospects of attaching glycans to a surface are in those of microarrays. There is a growing need for a diverse set of glycans in microarrays to test against unknown proteins, for example. This platform provides a way of procuring and immobilizing a large set of diverse bacterial glycans.

Another potential application of the 'clickable' bactoprenyl is that of a non-HPLC FRET analysis. FRET is the Forster Resonance Energy Transfer where when two fluorophores that have overlapping excitation and emission properties (the emission of one overlaps with the excitation of the other) come within a certain distance of each other, there is a measurable energy transfer. One fluorophore can be clicked onto the bactoprenyl and the other onto a clickable sugar where an assessment of the incorporation of that sugar onto that bactoprenyl can occur with fluorescence detection on a plate reader, without the use of HPLC.

DNA Aptamers for Bacterial Sugars

Although there are ways to detect sugars, from lectins to antibodies, there is a significant lack of variety when it comes to looking at bacterial sugars. The diversity and variety of modifications bacterial sugars undergo in comparison to the smaller sugar code

of mammalian cells, makes it difficult to design methods to detect these sugars. Utilizing our knowledge of biosynthetic pathways of bacterial sugars I have focused on the use of DNA aptamers to specifically target these sugars.

The focus of this work was the N-linked oligosaccharide from *C. jejuni* and colanic acid from *E. coli*. One of the two methodologies that was utilized, was a whole cell-based approach where one cell line was 'positive' and contained the N-linked oligosaccharide, and the other was the background control cell line that was 'negative'. The other utilized the ability to make the sugar units in the lab, were I produced the full heptasaccharide of the N-linked oligosaccharide and the colanic acid tetrasaccharide and linked those glycans to beads for the use in an immobilized selection process. With these two methodologies, I have provided evidence that the whole cell and immobilized N-linked oligosaccharide aptamer pools and candidates have specificity for the positive cells over the background control cells. Interestingly, the immobilized selection method was able to produce DNA aptamer candidates that were just as selective as the whole-cell based candidates. This has great potential in that the isolation of glycan-specific agents for any bacterial glycan can be done without the use of whole cells. This provides a cleaner method as well as functions as a platform for other analyses.

Furthermore, the immobilized colanic acid aptamer pool was able to detect the build-up of the tetrasaccharide intermediates within the *E. coli* lysates. This provides a way to detect specific glycans within cell experiments to give insight into the workings of the biosynthetic pathways in whole cells.

Future work includes the whole cell and immobilized selection of aptamers for the full colanic acid polymer and repeating unit, respectively. Potential aptamer candidates have been identified by flow cytometry that function similarly to the aptamer pool that show selectivity for the N-linked oligosaccharide containing cells and further analysis with these specific sequences may provide a better tool for specific bacterial detection. For the colanic acid tetrasaccharide aptamer pool, candidates from the immobilized tetrasaccharide have shown promise and further analysis with these specific sequences may provide a better tool for pathway intermediates. From there this library of glycan aptamers can be expanded from just the tetrasaccharide to the other glycan intermediates of the pathway to detect differences in intermediate build-up when bacteria is subjected to different environments or introduced to outside materials.

Further analysis is needed on the obtained candidates that includes full library sequencing on all of the aptamer pools to determine the number of aptamer candidates left at the end of the selection process for each target. Additionally, the affinity of each of the aptamer candidates is crucial in comparing the binding of these potential aptamer sequences to the target glycans. Furthermore, the preparation and purification of large amounts of the FAM-labeled products for analysis is time consuming and because some potential candidates are known, or from the pools being fully sequenced, it is possible to purchase the DNA already fluorescently labeled for more analysis, specifically with plate reader (where you need a larger amount of material for detection than flow cytometry).

There are plans in place to expand beyond colanic acid exopolysaccharide and the N-linked oligosaccharide to a capsular polysaccharide. This capsular polysaccharide is

called Enterobacterial Common Antigen (ECA), and as a capsule, has an interaction with the bacterial surface, keeping it in place compared to the excreted colanic acid exopolysaccharide. With the addition of ECA, the DNA aptamers could expand to three different glycan structures found on bacteria. With this is mind, it is also possible to expand beyond these three glycans to other pathways of interest, such as the previously elucidated CPSA pathway from our lab.

DNA Aptamers for Detection of the Colanic acid Exopolysaccharide

I have completed selections of whole cells containing the N-linked oligosaccharide, and have omitted the trials of colanic acid containing cells. In previous attempts at isolating aptamers for the colanic acid exopolysaccharide I have not been successful at producing selective aptamer pools. Although I have maintained a PCR product throughout the whole cell selection process, the verification of the production of colanic acid in enough quantity to visualize on the cells has not been accomplished with published staining methods. [130] This may be due to the fact that an exopolysaccharide is 'loosely' associated with the bacterial cell and is therefore found in the media more than with the cells. Therefore, I are not positive that the aptamer candidates obtained are specific for colanic acid and may just be selective for *E. coli.* In addition to this drawback, it has been discovered that colanic acid is produced on the cell in times of stress, and not widely produced when the *E. coli* bacteria is grown in normal laboratory strains at 37ºC.[131] Although attempts were made to grow this strain at lower temperatures to counteract that fact, I still have not confirmed the production of colanic acid on the current strains in our laboratory.

Generally, the detection of colanic acid from cells is accomplished with colorimetric assays, but are not very sugar specific and are prone to background interference from other sugar-containing glycans that may be isolated alongside the colanic acid.[132] Currently, I have made efforts to overproduce colanic acid with the use of a promotor of the colanic acid operon for use in selecting for whole cell agents for colanic acid. With confirmed overproduction of colanic acid on the strains used in the selection of DNA aptamers specific for the colanic acid exopolysaccharide, I would provide more specific detection of colanic acid over the fucose colorimetric assay.

REFERENCES

1. Rangel, J. M.; Sparling, P. H.; Crowe, C.; Griffin, P. M.; Swerdlow, D. L., Epidemiology of Escherichia coli O157 : H7 outbreaks, United States, 1982-2002. *Emerg Infect Dis* **2005,** *11* (4), 603-609.

2. Raum, E.; Lietzau, S.; von Baum, H.; Marre, R.; Brenner, H., Changes in Escherichia coli resistance patterns during and after antibiotic therapy: a longitudinal study among outpatients in Germany. *Clin Microbiol Infec* **2008,** *14* (1), 41-48.

3. Nguyen, Y.; Sperandio, V., Enterohemorrhagic E. coli (EHEC) pathogenesis. *Front Cell Infect Mi* **2012,** *2*.

4. Alemka, A.; Nothaft, H.; Zheng, J.; Szymanski, C. M., N-Glycosylation of Campylobacter jejuni Surface Proteins Promotes Bacterial Fitness. *Infect Immun* **2013,** *81* (5), 1674-1682.

5. Nothaft, H.; Davis, B.; Lock, Y. Y.; Perez-Munoz, M. E.; Vinogradov, E.; Walter, J.; Coros, C.; Szymanski, C. M., Engineering the Campylobacter jejuni N-glycan to create an effective chicken vaccine. *Sci Rep-Uk* **2016,** *6*.

6. Nothaft, H.; Perez-Munoz, M. E.; Gouveia, G. J.; Duar, R. M.; Wanford, J. J.; Lango-Scholey, L.; Panagos, C. G.; Srithayakumar, V.; Plastow, G. S.; Coros, C.; Bayliss, C. D.; Edison, A. S.; Walter, J.; Szymanski, C. M., Coadministration of the Campylobacter jejuni N-Glycan-Based Vaccine with Probiotics Improves Vaccine Performance in Broiler Chickens. *Appl Environ Microb* **2017,** *83* (23).

7. Burnham, P. M.; Hendrixson, D. R., Campylobacter jejuni: collective components promoting a successful enteric lifestyle. *Nat Rev Microbiol* **2018,** *16* (9), 551-565.

8. Brockhausen, I., Crossroads between bacterial and mammalian glycosyltransferases. *Front Immunol* **2014,** *5*, 1-21.

9. Mostowy, R. J.; Holt, K. E., Diversity-Generating Machines: Genetics of Bacterial Sugar-Coating. *Trends Microbiol* **2018**.

10. Zaman, S. B.; Hussain, M. A.; Nye, R.; Mehta, V.; Mamun, K. T.; Hossain, N., A Review on Antibiotic Resistance: Alarm Bells are Ringing. *Cureus* **2017,** *9* (6), e1403.

11. Tan, F. Y. Y.; Tang, C. M.; Exley, R. M., Sugar coating: bacterial protein glycosylation and host-microbe interactions. *Trends Biochem Sci* **2015,** *40* (7), 342-350.

12. Corbett, D.; Roberts, I. S., The role of microbial polysaccharides in host-pathogen interaction. *F1000 Biol Rep* **2009,** *1*, 30.

13. Whitfield, C.; Valvano, M. A., Biosynthesis and expression of cell-surface polysaccharides in gram-negative bacteria. *Adv Microb Physiol* **1993,** *35*, 135-246.

14. Whitfield, C.; Paiment, A., Biosynthesis and assembly of Group 1 capsular polysaccharides in Escherichia coli and related extracellular polysaccharides in other bacteria. *Carbohyd Res* **2003,** *338* (23), 2491-2502.

15. Goebel, W. F., Colanic Acid. *P Natl Acad Sci USA* **1963,** *49* (4), 464-&.

16. Mao, Y.; Doyle, M. P.; Chen, J., Role of colanic acid exopolysaccharide in the survival of enterohaemorrhagic Escherichia coli O157 : H7 in simulated gastrointestinal fluids. *Lett Appl Microbiol* **2006,** *42* (6), 642-647.

17. Mao, Y.; Doyle, M. P.; Chen, J. R., Insertion mutagenesis of wca reduces acid and heat tolerance of enterohemorrhagic Escherichia coli O157 : H7. *J Bacteriol* **2001,** *183* (12), 3811-3815.

18. Stincone, A.; Daudi, N.; Rahman, A. S.; Antczak, P.; Henderson, I.; Cole, J.; Johnson, M. D.; Lund, P.; Falciani, F., A systems biology approach sheds new light on Escherichia coli acid resistance. *Nucleic Acids Res* **2011,** *39* (17), 7512-7528.

19. Valguarnera, E.; Kinsella, R. L.; Feldman, M. F., Sugar and Spice Make Bacteria Not Nice: Protein Glycosylation and Its Influence in Pathogenesis. *J Mol Biol* **2016,** *428* (16), 3206-3220.

20. Nothaft, H.; Szymanski, C. M., Protein glycosylation in bacteria: sweeter than ever. *Nat Rev Microbiol* **2010,** *8* (11), 765-78.

21. Rangarajan, E. S.; Bhatia, S.; Watson, D. C.; Munger, C.; Cygler, M.; Matte, A.; Young, N. M., Structural context for protein N-glycosylation in bacteria: The structure of PEB3, an adhesin from Campylobacter jejuni. *Protein Sci* **2007,** *16* (5), 990-5.

22. Young, N. M.; Brisson, J. R.; Kelly, J.; Watson, D. C.; Tessier, L.; Lanthier, P. H.; Jarrell, H. C.; Cadotte, N.; St Michael, F.; Aberg, E.; Szymanski, C. M., Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, Campylobacter jejuni. *J Biol Chem* **2002,** *277* (45), 42530-9.

23. Nothaft, H.; Szymanski, C. M., Bacterial protein N-glycosylation: new perspectives and applications. *J Biol Chem* **2013,** *288* (10), 6912-20.

24. Day, C. J.; Semchenko, E. A.; Korolik, V., Glycoconjugates play a key role in Campylobacter jejuni infection: interactions between host and pathogen. *Front Cell Infect Mi* **2012,** *2*.

25. Szymanski, C. M.; Logan, S. M.; Linton, D.; Wren, B. W., Campylobacter - a tale of two protein glycosylation systems. *Trends Microbiol* **2003,** *11* (5), 233-238.

26. Lange, B. M.; Rujan, T.; Martin, W.; Croteau, R., Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes. *P Natl Acad Sci USA* **2000,** *97* (24), 13172-13177.

27. Rohmer, M.; Knani, M.; Simonin, P.; Sutter, B.; Sahm, H., Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate. *Biochem J* **1993,** *295 (Pt 2)*, 517-24.

28. Horbach, S.; Sahm, H.; Welle, R., Isoprenoid biosynthesis in bacteria: two different pathways? *Fems Microbiol Lett* **1993,** *111* (2-3), 135-40.

29. Schenk, B.; Fernandez, F.; Waechter, C. J., The ins(ide) and out(side) of dolichyl phosphate biosynthesis and recycling in the endoplasmic reticulum. *Glycobiology* **2001,** *11* (5), 61R-70R.

30. Whitfield, C., Biosynthesis and assembly of capsular polysaccharides in Escherichia coli. *Annu Rev Biochem* **2006,** *75*, 39-68.

31. Miyake, K.; Iijima, S., Bacterial capsular polysaccharide and sugar transferases. *Adv Biochem Eng Biotechnol* **2004,** *90*, 89-111.

32. Tytgat, H. L. P.; Lebeer, S., The Sweet Tooth of Bacteria: Common Themes in Bacterial Glycoconjugates. *Microbiol Mol Biol R* **2014,** *78* (3), 372-417.

33. Stevenson, G.; Andrianopoulos, K.; Hobbs, M.; Reeves, P. R., Organization of the Escherichia coli K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. *J Bacteriol* **1996,** *178* (16), 4885-4893.

34. 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-Us* **2006,** *45* (16), 5343-5350.

35. Olivier, N. B.; Chen, M. M.; Behr, J. R.; Imperiali, B., In vitro biosynthesis of UDP-N,N '-diacetylbacillosamine by enzymes of the Campylobacter jejuni general protein glycosylation system. *Biochemistry-Us* **2006,** *45* (45), 13659-13669.

36. 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-Us* **2009,** *48* (12), 2807-2816.

37. Park, I. H.; Lin, J. S.; Choi, J. E.; Shin, J. S., Characterization of Escherichia coli K1 colominic acid-specific murine antibodies that are cross-protective against Neisseria meningitidis groups B, C, and Y. *Mol Immunol* **2014,** *59* (2), 142-153.

38. Lattar, S. M.; Llana, M. N.; Denoel, P.; Germain, S.; Buzzola, F. R.; Lee, J. C.; Sordelli, D. O., Protein Antigens Increase the Protective Efficacy of a Capsule-Based Vaccine against Staphylococcus aureus in a Rat Model of Osteomyelitis. *Infect Immun* **2014,** *82* (1), 83-91.

39. Goncalves, V. M. M.; Takagi, M.; Lima, R. B.; Massaldi, H.; Giordano, R. C.; Tanizaki, M. M., Purification of capsular polysaccharide from Streptococcus pneumoniae serotype 23F by a procedure suitable for scale-up. *Biotechnol Appl Bioc* **2003,** *37*, 283-287.

40. Guexholzer, S.; Tomcsik, J., The Isolation and Chemical Nature of Capsular and Cell-Wall Haptens in a Bacillus Species. *J Gen Microbiol* **1956,** *14* (1), 14-+.

41. Hashimoto, M.; Kirikae, F.; Dohi, T.; Kusumoto, S.; Suda, Y.; Kirikae, T., Structural elucidation of a capsular polysaccharide from a clinical isolate of Bacteroides vulgatus from a patient with Crohn's disease. *Eur J Biochem* **2001,** *268* (11), 3139-3144.

42. Torensma, R.; Vanwijk, A.; Visser, M. J. C.; Bouter, A.; Rozenbergarska, M.; Verhoef, J., Monoclonal-Antibodies Specific for the Phase-Variant O-Acetylated K1 Capsule of Escherichia-Coli. *J Clin Microbiol* **1991,** *29* (7), 1356-1358.

43. Staats, N.; De Winder, B.; Stal, L. J.; Mur, L. R., Isolation and characterization of extracellular polysaccharides from the epipelic diatoms Cylindrotheca closterium and Navicula salinarum. *Eur J Phycol* **1999,** *34* (2), 161-169.

44. Whitfield, C.; Perry, M. B., Isolation and purification of cell surface polysaccharides from Gram-negative bacteria. *Method Microbiol* **1998,** *27*, 249-258.

45. Singer, S. J.; Fothergill, J. E.; Shainoff, J. R., A General Method for the Isolation of Antibodies. *J Am Chem Soc* **1960,** *82* (3), 565-571.

46. Singer, S. J.; Fothergill, J. E.; Shainoff, J. R., A New and General Method for the Isolation of Anti-Protein Antibodies. *J Am Chem Soc* **1959,** *81* (9), 2277-2278.

47. Sterner, E.; Flanagan, N.; Gildersleeve, J. C., Perspectives on Anti-Glycan Antibodies Gleaned from Development of a Community Resource Database. *Acs Chem Biol* **2016,** *11* (7), 1773-1783.

48. Yu, A. L.; Gilman, A. L.; Ozkaynak, M. F.; London, W. B.; Kreissman, S. G.; Chen, H. X.; Smith, M.; Anderson, B.; Villablanca, J. G.; Matthay, K. K.; Shimada, H.; Grupp, S. A.; Seeger, R.; Reynolds, C. P.; Buxton, A.; Reisfeld, R. A.; Gillies, S. D.; Cohn, S. L.; Maris, J. M.; Sondel, P. M.; Children's Oncology, G., Anti-GD2 antibody with GM-CSF, interleukin-2, and isotretinoin for neuroblastoma. *N Engl J Med* **2010,** *363* (14), 1324-34.

49. Cheung, I. Y.; Kushner, B. H.; Modak, S.; Basu, E. M.; Roberts, S. S.; Cheung, N. V., Phase I trial of anti-GD2 monoclonal antibody hu3F8 plus GM-CSF: Impact of body weight, immunogenicity and anti-GD2 response on pharmacokinetics and survival. *Oncoimmunology* **2017,** *6* (11), e1358331.

50. Cheung, N. K.; Cheung, I. Y.; Kramer, K.; Modak, S.; Kuk, D.; Pandit-Taskar, N.; Chamberlain, E.; Ostrovnaya, I.; Kushner, B. H., Key role for myeloid cells: phase II results of anti-G(D2) antibody 3F8 plus granulocyte-macrophage colony-stimulating factor for chemoresistant osteomedullary neuroblastoma. *Int J Cancer* **2014,** *135* (9), 2199-205.

51. Ferreira, F., The use of recombinant lectins for the bioanalysis of cell surface glycosylation. *New Biotechnol* **2018,** *44*, S86-S86.

52. Imberty, A., Recombinant and engineered lectins for research and diagnostics. *Febs Open Bio* **2018,** *8*, 54-54.

53. Nizet, V.; Varki, A.; Aebi, M., Microbial Lectins: Hemagglutinins, Adhesins, and Toxins. In *Essentials of Glycobiology*, rd; Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.;

Prestegard, J. H.; Schnaar, R. L.; Seeberger, P. H., Eds. Cold Spring Harbor (NY), 2015; pp 481-491.

54. Sharon, N., Bacterial lectins, cell-cell recognition and infectious disease. *Febs Lett* **1987,** *217* (2), 145-57.

55. Weimar, T.; Haase, B.; Kohli, T., Low affinity carbohydrate lectin interactions examined with surface plasmon resonance. *J Carbohyd Chem* **2000,** *19* (8), 1083-1089.

56. Woo, C. M.; Lund, P. J.; Huang, A. C.; Davis, M. M.; Bertozzi, C. R.; Pitteri, S. J., Mapping and Quantification of Over 2000 O-linked Glycopeptides in Activated Human T Cells with Isotope-Targeted Glycoproteomics (Isotag). *Mol Cell Proteomics* **2018,** *17* (4), 764-775.

57. Kamariza, M.; Shieh, P.; Ealand, C. S.; Peters, J. S.; Chu, B.; Rodriguez-Rivera, F. P.; Sait, M. R. B.; Treuren, W. V.; Martinson, N.; Kalscheuer, R.; Kana, B. D.; Bertozzi, C. R., F Rapid detection of Mycobacterium tuberculosis in sputum with a solvatochromic trehalose probe. *Sci Transl Med* **2018,** *10* (430).

58. Gordon, C. G.; Mackey, J. L.; Jewett, J. C.; Sletten, E. M.; Houk, K. N.; Bertozzi, C. R., Reactivity of Biarylazacyclooctynones in Copper-Free Click Chemistry. *J Am Chem Soc* **2012,** *134* (22), 9199-9208.

59. Bull, C.; Boltje, T. J.; van Dinther, E. A. W.; Peters, T.; de Graaf, A. M. A.; Leusen, J. H. W.; Kreutz, M.; Figdor, C. G.; den Brok, M. H.; Adema, G. J., Targeted Delivery of a Sialic Acid-Blocking Glycomimetic to Cancer Cells Inhibits Metastatic Spread. *Acs Nano* **2015,** *9* (1), 733-745.

60. Schultz, M. J.; Swindall, A. F.; Bellis, S. L., Regulation of the metastatic cell phenotype by sialylated glycans. *Cancer Metast Rev* **2012,** *31* (3-4), 501-518.

61. Cho, S.; Lee, B. R.; Cho, B. K.; Kim, J. H.; Kim, B. G., In vitro selection of sialic acid specific RNA aptamer and its application to the rapid sensing of sialic acid modified sugars. *Biotechnol Bioeng* **2013,** *110* (3), 905-913.

62. Yang, Q.; Goldstein, I. J.; Mei, H. Y.; Engelke, D. R., DNA ligands that bind tightly and selectively to cellobiose. *P Natl Acad Sci USA* **1998,** *95* (10), 5462-5467.

63. Nikolaus, N.; Strehlitz, B., DNA-Aptamers Binding Aminoglycoside Antibiotics. *Sensors-Basel* **2014,** *14* (2), 3737-3755.

64. Teng, K. H.; Liang, P. H., Undecaprenyl diphosphate synthase, a cis-prenyltransferase synthesizing lipid carrier for bacterial cell wall biosynthesis. *Mol Membr Biol* **2012,** *29* (7), 267-273.

65. Furlong, S. E.; Ford, A.; Albarnez-Rodriguez, L.; Valvano, M. A., Topological analysis of the Escherichia coli WcaJ protein reveals a new conserved configuration for the polyisoprenyl-phosphate hexose-1-phosphate transferase family. *Sci Rep-Uk* **2015,** *5*.

66. 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. *Bioorgan Med Chem* **2013,** *21* (17), 5428-5435.

67. Mostafavi, A. Z.; Troutman, J. M., Biosynthetic assembly of the Bacteroides fragilis capsular polysaccharide A precursor bactoprenyl diphosphate-linked acetamido-4-amino-6-deoxygalactopyranose. *Biochemistry-Us* **2013,** *52* (11), 1939-49.

68. Tao, J. Y.; Diaz, R. K.; Teixeira, C. R. V.; Hackmann, T. J., Transport of a Fluorescent Analogue of Glucose (2-NBDG) versus Radiolabeled Sugars by Rumen Bacteria and Escherichia coli. *Biochemistry-Us* **2016,** *55* (18), 2578-2589.

69. Troutman, J. M.; Sharma, S.; Erickson, K. M.; Martinez, C. D., Functional identification of a galactosyltransferase critical to Bacteroides fragilis Capsular Polysaccharide A biosynthesis. *Carbohyd Res* **2014,** *395*, 19-28.

70. 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. *Carbohyd Res* **2012,** *359*, 44-53.

71. Lukose, V.; Whitworth, G.; Guan, Z.; Imperiali, B., Chemoenzymatic Assembly of Bacterial Glycoconjugates for Site-Specific Orthogonal Labeling. *J Am Chem Soc* **2015,** *137* (39), 12446-9.

72. Meredith, T. C.; Mamat, U.; Kaczynski, Z.; Lindner, B.; Holst, O.; Woodard, R. W., Modification of lipopolysaccharide with colanic acid (M-antigen) repeats in Escherichia coli. *J Biol Chem* **2007,** *282* (11), 7790-7798.

73. Sutherland, I. W., Structural Studies on Colanic Acid, Common Exopolysaccharide Found in Enterobacteriaceae, by Partial Acid Hydrolysis - Oligosaccharides from Colanic Acid. *Biochem J* **1969,** *115* (5), 935-+.

74. Patel, K. B.; Toh, E.; Fernandez, X. B.; Hanuszkiewicz, A.; Hardy, G. G.; Brun, Y. V.; Bernards, M. A.; Valvano, M. A., Functional Characterization of UDP-Glucose:Undecaprenyl-Phosphate Glucose-1-Phosphate Transferases of Escherichia coli and Caulobacter crescentus. *J Bacteriol* **2012,** *194* (10), 2646-2657.

75. Aragao-Leoneti, V.; Campo, V. L.; Gomes, A. S.; Field, R. A.; Carvalho, I., Application of copper(I)-catalysed azide/alkyne cycloaddition (CuAAC) 'click chemistry' in carbohydrate drug and neoglycopolymer synthesis. *Tetrahedron* **2010,** *66* (49), 9475-9492.

76. Chehade, K. A. H.; Andres, D. A.; Morimoto, H.; Spielmann, H. P., Design and synthesis of a transferable farnesyl pyrophosphate analogue to Ras by protein farnesyltransferase. *J Org Chem* **2000,** *65* (10), 3027-3033.

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

78. 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), 14255-9.

79. Sonnhammer, E. L.; von Heijne, G.; Krogh, A., A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* **1998,** *6*, 175-82.

80. Saldias, M. S.; Patel, K.; Marolda, C. L.; Bittner, M.; Contreras, I.; Valvano, M. A., Distinct functional domains of the Salmonella enterica WbaP transferase that is involved in the initiation reaction for synthesis of the O antigen subunit. *Microbiol-Sgm* **2008,** *154*, 440-453.

81. Cartee, R. T.; Forsee, W. T.; Bender, M. H.; Ambrose, K. D.; Yother, J., CpsE from type 2 Streptococcus pneumoniae catalyzes the reversible addition of glucose-1-phosphate to a polyprenyl phosphate acceptor, initiating type 2 capsule repeat unit formation. *J Bacteriol* **2005,** *187* (21), 7425-7433.

82. 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.

83. Ranjit, D. K.; Young, K. D., Colanic Acid Intermediates Prevent De Novo Shape Recovery of Escherichia coli Spheroplasts, Calling into Question Biological Roles Previously Attributed to Colanic Acid. *J Bacteriol* **2016,** *198* (8), 1230-1240.

84. Zhang, J. Y.; Poh, C. L., Regulating exopolysaccharide gene wcaF allows control of Escherichia coli biofilm formation. *Sci Rep-Uk* **2018,** *8*.

85. Ratto, M.; Verhoef, R.; Suihko, M. L.; Blanco, A.; Schols, H. A.; Voragen, A. G. J.; Wilting, R.; Siika-aho, R.; Buchert, J., Colanic acid is an exopolysaccharide common to many enterobacteria isolated from paper-machine slimes. *J Ind Microbiol Biot* **2006,** *33* (5), 359-367.

86. 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. In *J Biol Chem*, 2018.

87. Wilkins, M. R.; Gasteiger, E.; Bairoch, A.; Sanchez, J. C.; Williams, K. L.; Appel, R. D.; Hochstrasser, D. F., Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* **1999,** *112*, 531-52.

88. Xiao, H.; Tang, G. X.; Wu, R., Site-Specific Quantification of Surface N-Glycoproteins in Statin-Treated Liver Cells. *Anal Chem* **2016,** *88* (6), 3324-32.

89. Andersen, K. A.; Aronoff, M. R.; McGrath, N. A.; Raines, R. T., Diazo groups endure metabolism and enable chemoselectivity in cellulo. *J Am Chem Soc* **2015,** *137* (7), 2412-5.

90. Beahm, B. J.; Dehnert, K. W.; Derr, N. L.; Kuhn, J.; Eberhart, J. K.; Spillmann, D.; Amacher, S. L.; Bertozzi, C. R., A visualizable chain-terminating inhibitor of glycosaminoglycan biosynthesis in developing zebrafish. *Angew Chem Int Ed Engl* **2014,** *53* (13), 3347-52.

91. Luchansky, S. J.; Hang, H. C.; Saxon, E.; Grunwell, J. R.; Yu, C.; Dube, D. H.; Bertozzi, C. R., Constructing azide-labeled cell surfaces using polysaccharide biosynthetic pathways. *Methods Enzymol* **2003,** *362*, 249-72.

92. Shieh, P.; Siegrist, M. S.; Cullen, A. J.; Bertozzi, C. R., Imaging bacterial peptidoglycan with near-infrared fluorogenic azide probes. *Proc Natl Acad Sci U S A* **2014,** *111* (15), 5456-61.

93. Siegrist, M. S.; Whiteside, S.; Jewett, J. C.; Aditham, A.; Cava, F.; Bertozzi, C. R., (D)-Amino acid chemical reporters reveal peptidoglycan dynamics of an intracellular pathogen. *ACS Chem Biol* **2013,** *8* (3), 500-5.

94. Coker, O. O.; Palittapongarnpim, P., Current understanding of de novo synthesis of bacterial lipid carrier (undecaprenyl phosphate): More enzymes to be discovered. *Afr J Microbiol Res* **2011,** *5* (18), 2555-2565.

95. Labadie, G. R.; Viswanathan, R.; Poulter, C. D., Farnesyl diphosphate analogues with omega-bioorthogonal azide and alkyne functional groups for protein farnesyl transferase-catalyzed Ligation reactions. *J Org Chem* **2007,** *72* (24), 9291-9297.

96. Viswanathan, R.; Labadie, G. R.; Poulter, C. D., Regioselective Covalent Immobilization of Catalytically Active Glutathione S-Transferase on Glass Slides. *Bioconjugate Chem* **2013,** *24* (4), 571-577.

97. 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.

98. El Ghachi, M.; Howe, N.; Huang, C. Y.; Olieric, V.; Warshamanage, R.; Touze, T.; Weichert, D.; Stansfeld, P. J.; Wang, M. T.; Kerff, F.; Caffrey, M., Crystal structure of undecaprenyl-pyrophosphate phosphatase and its role in peptidoglycan biosynthesis. *Nat Commun* **2018,** *9*.

99. El Ghachi, M.; Bouhss, A.; Blanot, D.; Mengin-Lecreulx, D., The bacA gene of Escherichia coli encodes an undecaprenyl pyrophosphate phosphatase activity. *J Biol Chem* **2004,** *279* (29), 30106-30113.

100. Erickson, K. Characterization of Sugar-Modifying Enzymes in the Production of the Capsular Polysaccharides of *B. fragilis* and *V. vulnificus* University of North Carolina at Charlotte, 2014.

101. De Schutter, J. W.; Morrison, J. P.; Morrison, M. J.; Ciulli, A.; Imperiali, B., Targeting Bacillosamine Biosynthesis in Bacterial Pathogens: Development of Inhibitors to a Bacterial Amino-Sugar Acetyltransferase from Campylobacter jejuni. *J Med Chem* **2017,** *60* (5), 2099-2118.

102. 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.

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

104. 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.

105. James, D. B. A.; Gupta, K.; Hauser, J. R.; Yother, J., Biochemical Activities of Streptococcus pneumoniae Serotype 2 Capsular Glycosyltransferases and Significance of Suppressor Mutations Affecting the Initiating Glycosyltransferase Cps2E. *J Bacteriol* **2013,** *195* (24), 5469-5478.

106. Bai, B.; Chu, C. J.; Lowary, T. L., Lipooligosaccharides from Mycobacteria: Structure, Function, and Synthesis. *Isr J Chem* **2015,** *55* (3-4), 360-372.

107. Studier, F. W., Protein production by auto-induction in high-density shaking cultures. *Protein Expres Purif* **2005,** *41* (1), 207-234.

108. Chatzidaki-Livanis, M.; Jones, M. K.; Wright, A. C., Genetic variation in the Vibrio vulnificus group 1 capsular polysaccharide operon. *J Bacteriol* **2006,** *188* (5), 1987-1998.

109. Pacheco, A. R.; Sperandio, V., Shiga toxin in enterohemorrhagic E.coli: regulation and novel anti-virulence strategies. *Front Cell Infect Mi* **2012,** *2*.

110. Blind, M.; Blank, M., Aptamer Selection Technology and Recent Advances. *Mol Ther-Nucl Acids* **2015,** *4*.

111. Ruscito, A.; DeRosa, M. C., Small-Molecule Binding Aptamers: Selection Strategies, Characterization, and Applications. *Front Chem* **2016,** *4*.

112. Klutz, S.; Holtmann, L.; Lobedann, M.; Schembecker, G., Cost evaluation of antibody production processes in different operation modes. *Chem Eng Sci* **2016,** *141*, 63-74.

113. Chen, F.; Zhou, J.; Luo, F. L.; Mohammed, A. B.; Zhang, X. L., Aptamer from whole-bacterium SELEX as new therapeutic reagent against virulent Mycobacterium tuberculosis. *Biochem Bioph Res Co* **2007,** *357* (3), 743-748.

114. Dua, P.; Ren, S.; Lee, S. W.; Kim, J. K.; Shin, H. S.; Jeong, O. C.; Kim, S.; Lee, D. K., Cell-SELEX Based Identification of an RNA Aptamer for Escherichia coli and Its Use in Various Detection Formats. *Mol Cells* **2016,** *39* (11), 807-813.

115. Moon, J.; Kim, G.; Park, S. B.; Lim, J.; Mo, C., Comparison of Whole-Cell SELEX Methods for the Identification of Staphylococcus Aureus-Specific DNA Aptamers. *Sensors-Basel* **2015,** *15* (4), 8884-8897.

116. Sefah, K.; Shangguan, D.; Xiong, X. L.; O'Donoghue, M. B.; Tan, W. H., Development of DNA aptamers using Cell-SELEX. *Nat Protoc* **2010,** *5* (6), 1169-1185.

117. Marttila, A. T.; Laitinen, O. H.; Airenne, K. J.; Kulik, T.; Bayer, E. A.; Wilchek, M.; Kulomaa, M. S., Recombinant NeutraLite avidin: a non-glycosylated, acidic mutant of chicken avidin that exhibits high affinity for biotin and low non-specific binding properties. *Febs Lett* **2000,** *467* (1), 31-6.

118. Argarana, C. E.; Kuntz, I. D.; Birken, S.; Axel, R.; Cantor, C. R., Molecular cloning and nucleotide sequence of the streptavidin gene. *Nucleic Acids Res* **1986,** *14* (4), 1871-82.

119. Livnah, O.; Bayer, E. A.; Wilchek, M.; Sussman, J. L., Three-dimensional structures of avidin and the avidin-biotin complex. *Proc Natl Acad Sci U S A* **1993,** *90* (11), 5076-80.

120. Green, N. M., Avidin. 3. The Nature of the Biotin-Binding Site. *Biochem J* **1963,** *89*, 599-609.

121. Yu, D.; Ellis, H. M.; Lee, E. C.; Jenkins, N. A.; Copeland, N. G.; Court, D. L., An efficient recombination system for chromosome engineering in Escherichia coli. *Proc Natl Acad Sci U S A* **2000,** *97* (11), 5978-83.

122. Breslauer, K. J.; Frank, R.; Blocker, H.; Marky, L. A., Predicting DNA duplex stability from the base sequence. *Proc Natl Acad Sci U S A* **1986,** *83* (11), 3746-50.

123. Chang, Y. C.; Yang, C. Y.; Sun, R. L.; Cheng, Y. F.; Kao, W. C.; Yang, P. C., Rapid single cell detection of Staphylococcus aureus by aptamer- conjugated gold nanoparticles. *Sci Rep-Uk* **2013,** *3*.

124. Zou, Y.; Duan, N.; Wu, S.; Shen, M.; Wang, Z., Selection, Identification, and Binding Mechanism Studies of an ssDNA Aptamer Targeted to Different Stages of E-coli O157:H7. *J Agr Food Chem* **2018,** *66* (22), 5677-5682.

125. Damase, T. R.; Ellington, A. D.; Allen, P. B., Purification of single-stranded DNA by co-polymerization with acrylamide and electrophoresis. *Biotechniques* **2017,** *62* (6), 275-282.

126. CR, E.; L, B.; C-C, H.; Betteken, M.; H, A.-P.; C., A., Salmonella Pathogenicity Island One is Expressed in the Chicken Gut and Promotes Bacterial Outgrowth. *Infect Immun* **2018**.

127. Sahler, J. M.; Eade, C. R.; Altier, C.; March, J. C., Salmonella enterica Serovar Typhimurium Increases Functional PD-L1 Synergistically with Gamma Interferon in Intestinal Epithelial Cells via Salmonella Pathogenicity Island 2. *Infect Immun* **2018,** *86* (5).

128. Taylor, D. L.; Herriott, I. C.; Long, J.; O'Neill, K., TOPO TA is A-OK: a test of phylogenetic bias in fungal environmental clone library construction. *Environ Microbiol* **2007,** *9* (5), 1329-1334.

129. Zadeh, J. N.; Steenberg, C. D.; Bois, J. S.; Wolfe, B. R.; Pierce, M. B.; Khan, A. R.; Dirks, R. M.; Pierce, N. A., NUPACK: Analysis and Design of Nucleic Acid Systems. *J Comput Chem* **2011,** *32* (1), 170-173.

130. Ren, G.; Wang, Z.; Li, Y.; Hu, X.; Wang, X., Effects of Lipopolysaccharide Core Sugar Deficiency on Colanic Acid Biosynthesis in Escherichia coli. *J Bacteriol* **2016,** *198* (11), 1576-1584.

131. Chen, J.; Lee, S. M.; Mao, Y., Protective effect of exopolysaccharide colanic acid of Escherichia coli O157:H7 to osmotic and oxidative stress. *Int J Food Microbiol* **2004,** *93* (3), 281-6.

132. Cohenford, M. A.; Abraham, A.; Abraham, J.; Dain, J. A., Colorimetric assay for free and bound L-fucose. *Anal Biochem* **1989,** *177* (1), 172-7.