CHARACTERIZATION OF THE COLANIC ACID BIOSYNTHESIS PATHWAY WITH A FLUORESCENT BACTOPRENYL PHOSPHATE ANALOGUE

by

Phillip Michael Scott

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

Charlotte

2015

Approved by:

_____
Dr. Jerry Troutman

_____
Dr. Daniel Rabinovich

_____
Dr. Juan Vivero-Escoto

_____
Dr. Irina Nesmelova

ABSTRACT

PHILLIP MICHAEL SCOTT. Characterization of the colanic acid biosynthesis pathway with a novel isoprenoid analogue. (Under the direction of DR. JERRY TROUTMAN)

When introduced to harsh conditions such as low pH, pathogenic *Escherichia coli* can secrete colanic acid, an exopolysaccharide that completely engulfs the organism, to establish a protective barrier between the organism and the acidic environment. The colanic acid polymer is made up of multiple six-sugar repeating units comprised of glucose, 2,3-acetylated fucose, fucose, galactose, glucoronic acid and 4,6-pyruvated galactose. Studies have shown that as pH of an environment is lowered, *E. coli* strains excreting colanic acid survive longer than strains that cannot produce the polymer. This action may contribute to the survival of virulent *E. coli* in contaminated food products and inside a host organism. Considering the contaminated food sources containing these bacteria are widespread and contribute to over 73,000 infections per year in the United States alone, it is vitally important to understand how colanic acid is synthesized to develop novel antibiotic therapeutics to target this pathway. While the region of the *E. coli* genome that encodes for colanic acid biosynthesis has been reported, only the first enzyme in the biosynthesis pathway has been characterized. Utilizing fluorescent analogues to bactoprenyl monophosphate, the isoprenoid anchor that colanic acid is built upon, the next three glycosyltransferase enzymes (WcaI, WcaE, WcaC) in the pathway have been characterized. Finally, the data presented here suggests that the acetylation of the first fucose residue is regulatory in colanic acid biosynthesis and therefore could be a potential antibiotic target for this pathway.

ACKNOWLEDGMENTS

There are honestly too many people to thank here and I could probably write an entire thesis about all of the people that have helped me along the way. First, I would like to thank Dr. Walsh for the continued tuition support and the Fellowship that have allowed me to do this work. Secondly, I would like to thank my thesis committee for their assistance when needed and guiding me through the process. Thirdly, I would like to acknowledge the Troutman lab members, all of whom have become my family the last two years. Specifically I would like to acknowledge one person in particular: Kate Erickson. I have often referred to Kate as "The Empress" because she is the leader of the Troutman Lab world, and it would certainly fall apart without her. The teasing often annoys her but it is true. Further, she is one of the best people that I have ever known and I challenge anyone to find someone more supportive and selfless. Fourthly, I would like to thank Dr. Troutman for being the best boss that I will certainly ever have. He has helped me tremendously and that is something I will always remember. Before joining his lab, I had minimal laboratory skills, but now I am confident that the techniques I have learned here will benefit me throughout my career. Mostly though, I'll miss our multi-hour long tangents that always start scientifically relevant but almost certainly end with abstract discussions concerning Quentin Tarantino movies.

Lastly, but certainly not least, I would like to acknowledge Rachel S. Helms. She is the most wonderful person and my life would be incomplete without her. Over the last couple years, she has changed my life for the better and I cannot wait to continue our journey together in Baltimore.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

2-CN-5Z-BP	Fluorescent BP analogue
Ac	Acetyl functional group
BP	Bactoprenyl monophosphate
BPP	Bactoprenyl diphosphate
CA	Colanic acid
DBU	1,8-Diazabicycloundec-7-ene
dNTPs	Deoxy-nucleotides
DPPA	Diphenylphosphorylazide
FLU	Arbitrary Fluorescence Units
FPP	Farnesyl diphosphate
Fuc	Fucose
Gal	Galactose
Glc	Glucose
GlcA	Glucuronic acid
GlcNAc	N-acetylglucosamine
IPP	Isopentenyl diphosphate
$NADP^+$	Nicotinamide adenine dinucleotide phosphate
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced form)
TMHMM	Transmembrane Hidden Markov Model
UDP	Uridine diphosphate
UPPS	Undecaprenyl pyrophosphate synthase

CHAPTER ONE: INTRODUCTION

Escherichia coli is a well-known food-borne pathogen. While not all strains of *E. coli* are harmful to humans, virulent strains are responsible for 73,480 infections, 2,168 hospitalizations and 61 deaths every year in the United States alone.[(1)] The virulence of these strains comes from the ability of the organism to produce and release Shiga toxin, a protein that enters a target cell and inhibits protein biosynthesis by disrupting a structural component of the ribosome, causing cell death.[(2, 3)] Interestingly, these same pathogenic bacteria 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 fecal excrement.[(3)] Various sources of *E. coli* outbreaks include contaminated meats, dairy products and vegetables. Considering that the number of virulent cells required to cause sickness is relatively low, the wide range of potential sources of contamination and the 5% mortality rate associated with symptoms related to *E. coli* infection—targeting these virulent *E. coli* is a significant health concern.

A major consideration for treating bacterial infections is the prevailing increase in antibiotic resistance in the environment. One clinical study reported an increase in *E. coli* antibiotic resistance to ampicillin, doxycycline and co-trimoxazole.[(4)] These findings highlight the significant problem that antibiotic resistance has in treating *E. coli* outbreaks, but do not fully elucidate the complexity of combating these organisms. In the

case of virulent *E. coli*, the specific target of an antibiotic contributes to the effectiveness of the therapeutic.[3] Antibiotics that interfere with bacterial DNA synthesis actually exacerbate the infection by increasing production and release of Shiga toxin. Conversely, disrupting other metabolic processes such as protein and cell wall biosynthesis prevent Shiga toxin production and are therefore, more appropriate targets for novel antibiotic development.[3]

One particular antibiotic target not yet utilized by modern antibiotics is the colanic acid biosynthesis pathway. Colanic acid (CA) is an encasing sugar polymer that *E. coli* excrete under acidic conditions to escape degradation.[5] The polymer is comprised of multiple hexasaccharide repeat units. Each repeat unit consists of glucose, 2,3-diacetylated fucose, fucose, galactose, glucuronic acid and 4,6-pyruvated galactose (Figure 1). [5, 6]

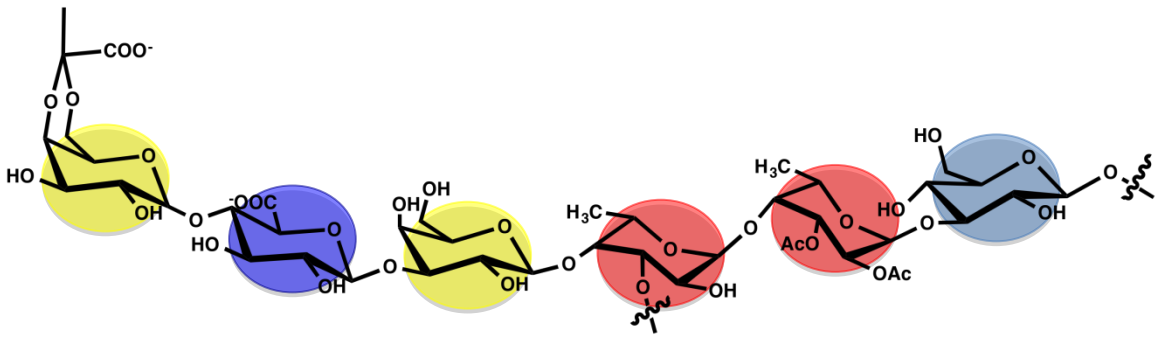


Figure 1: Structure of the colanic acid repeating unit. Colors are representative of the following sugars: blue = glucose, red = fucose, indigo = glucuronic acid, yellow = galactose.

Research has shown that when *E. coli* cannot synthesize colanic acid, they become highly susceptible to acidic pH. [7, 8]Chen *et al.* produced a mutated strain of *E. coli* that could not synthesize the sugar polymer.[7] As the pH was lowered from 6.5 to 4.5, there was a dramatic decline in cell populations in the mutated strains that could not produce colanic acid. However, *E. coli* cells producing colanic acid were unaffected after 15 hours, regardless of pH. This same group published a follow-up study containing similar experiments with simulated gastrointestinal fluid.[8] The results from this study further verified the protective effect that CA production displays in *E. coli*. Even at an extremely low pH of 2.0, cell lines producing colanic acid survived longer than those that could not. Based on these results, the authors suggest that colanic acid production allows virulent *E. coli* to survive in acidic foods and in stomach acid long enough to infect a host.[7, 8] These results elude to the potential for antibiotic targeting of virulent *E. coli*. By inhibiting colanic acid production, virulent *E. coli* would become completely vulnerable to acid degradation prior to host infection. Most importantly, as this type of inhibition would disrupt processes similar to cell wall biosynthesis, this type of antibiotic would be effective in targeting virulent *E. coli*. To determine potential targets for colanic acid biosynthesis inhibition, it is first important to understand how *E. coli* synthesize the polymer.

1.2: Biosynthesis of the Colanic Acid Repeating-Unit *In Vivo* and *In Vitro*

The colanic acid repeating unit is synthesized on the cytosolic face of the inner membrane (Figure 2).[5] Bactoprenyl phosphate (BP), embedded in the inner membrane, serving as an anchor for each sugar addition. An initiating hexose-1-phosphate transferase enzyme transfers the first glucose-1-phosphate to the BP anchor. Five

subsequent sugar additions by five individual glycosyltransferases complete the synthesis of the repeating unit. Two acetyltransferase enzymes acetylate the first fucose residue at the 2' and 3' positions of the sugar. Likewise, a pyruvyltransferase modifies the terminal galactose at the 4' and 6' positions of the sugar with a pyruvyl group.

The polymerization and export of colanic acid is anticipated to be a Wzy-dependent process, similar to that required in certain bacterial capsule biosyntheses. Once the repeat unit is synthesized, the six-sugar repeating unit is flipped across the inner membrane into the periplasmic space. Here, multiple repeat units are polymerized by Wzy and exported by Wza, a translocon protein, through the outer membrane of the bacterium. From here, the BP anchor is recycled for continued synthesis of colanic acid.

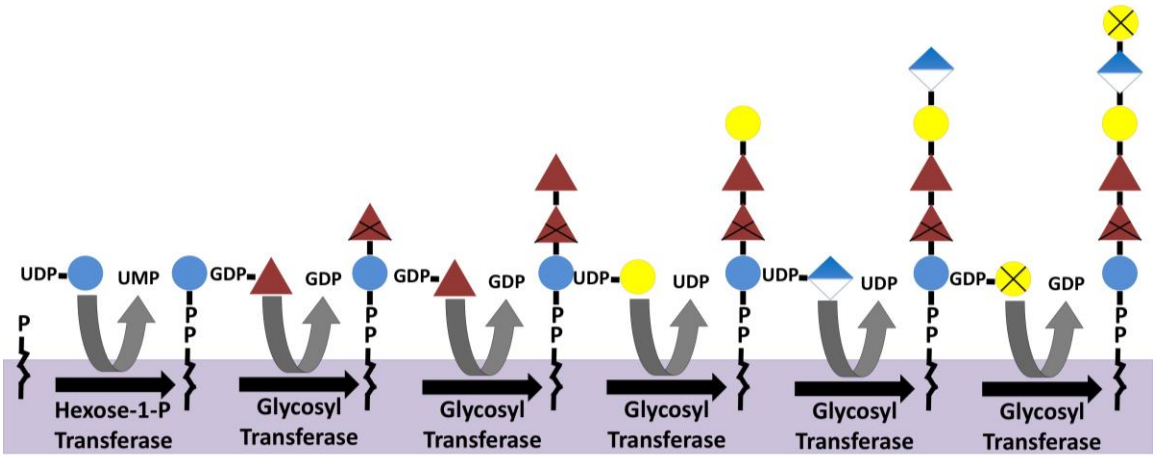


Figure 2: Biosynthesis of the colanic acid repeating unit *in vivo.* Colanic acid biosynthesis pathway. Black squiggle is representative of the bactoprenyl phosphate (BP) anchor. Arrows are representative of enzymatic activity. Colored circles follow same sugar scheme as in Figure 1. Crosses are indications of sugar modifications. The hexose-1-P transferase transfers the first glucose-1-phosphate to the BP anchor, followed by sequential addition of the remaining sugars by five different glycosyltransferase enzymes to build the complete repeating unit.

While the polymerization and export of colanic acid is somewhat understood, little to nothing is known regarding how the repeat unit is synthesized. The initiating-hexose-1-phosphate transferase has been functionally characterized,[9] but it is not known which glycosyltransferase transfers each additional sugar to complete the synthesis of the repeating unit. This glaring lack of knowledge greatly hinders the development of antibiotics targeted to the colanic acid biosynthesis pathway. Functionally characterizing the glycosyltransferases responsible for each step in the synthesis is required to provide potential targets for inhibition of colanic acid synthesis.

1.3: The Colanic Acid Biosynthesis Gene Cluster

With the eventual goal of finding antibiotic targets within the CA biosynthesis pathway, the first step is to utilize the region of the *E. coli* genome responsible for colanic acid production to determine which genes are required for synthesis *in vitro*. This region is comprised of a cluster of genes that have been previously identified by the Reeves group.[10] Utilizing bioinformatic approaches, this group was able to hypothesize potential functions of each of the genes in the cluster based upon protein sequence similarity to homologous proteins with known functions. A summary of their results is depicted in Figure 3.

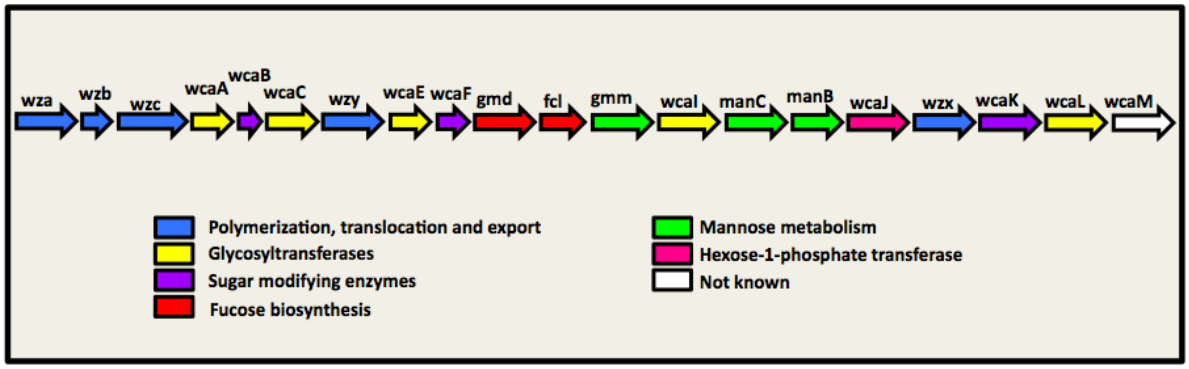


Figure 3: The colanic acid gene cluster. Arrows represent the genes located within the region of the *E. coli* genome responsible for encoding proteins that have a role in colanic acid biosynthesis. The sizes of each arrow are indicative of the size of the genes.

The biosynthesis of colanic acid begins with the transfer of a glucose-1-phosphate by the initiating hexose-1-phosphate transferase (Figure 3, pink). This reaction is catalyzed by WcaJ and is the only sugar-transferring enzyme in the colanic biosynthesis pathway that has been functionally characterized to confirm the bioinformatic results (Scheme 1). [(9)] The five glycosyltransferases that transfer the remaining sugars to complete the synthesis of the repeat unit are hypothesized to be WcaA, WcaC, WcaE, WcaI and WcaL (Figure 3, yellow). However, it is not known which enzyme transfers each individual sugar.

L-fucose, the second and third sugar of the repeating unit, is synthesized from GDP-mannose by Gmd and Fcl (Figure 3, red).[(11, 12)] There are two proposed acetyltransferases, WcaB and WcaF that acetylate the first fucose sugar at the 2' and 3' positions, but it is not known which enzyme catalyzes the transfer at each position (Figure 3, purple). It is also not understood which acetylation events are required or at what point in the synthesis of the repeating unit that these acetyl groups are transferred. Finally, WcaK is hypothesized to be the pyruvyltransferase enzyme that pyruvates the

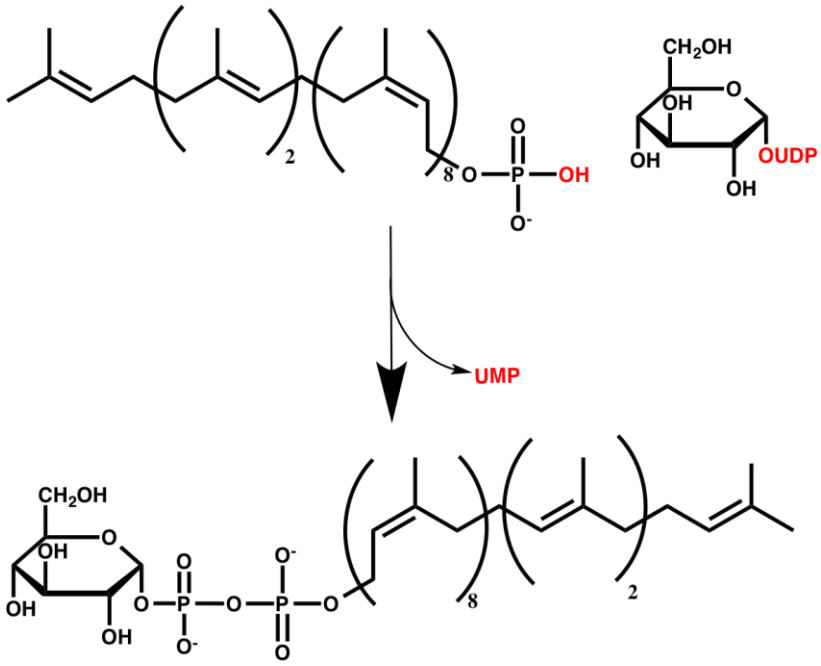
terminal galactose at the 4' and 6' positions (Figure 3, purple). Together, all of these genes provide the foundation for the complete synthesis of the CA repeat-unit.

1.4: Characterized Proteins Within the Colanic Acid Biosynthesis Pathway

1.4.1: WcaJ, The Initiating Hexose-1-Phosphate Transferase

WcaJ is the protein responsible for transferring glucose-1-phosphate to the BP anchor to begin the synthesis of colanic acid (Scheme 1). While there is no crystal structure, the protein has large hydrophobic domains that likely allows for insertion into the inner-membrane, predicted by the TMHMM server (Transmembrane Hidden Markov Model) that determines the presence of transmembrane domains based upon the presence of hydrophobic amino acids in the primary sequence (Figure 4).[13] This hydrophobicity profile is nearly identical to WbaP, the initating-galactose-1-phosphate transferase in the O-antigen biosynthesis pathway. [14, 15] Interestingly, it was determined that only the C-terminal domain of the enzyme (region beyond the last transmembrane domain) was required for catalytic function.[14] The activity of WcaJ was previously characterized using autoradiography of bactoprenols modified with labeled glucose.[9] The *in vitro* transferase assay consisted of incubating cellular membrane fractions containing overexpressed WcaJ and embedded BP with radiolabelled UDP-[^{14}C] glucose. Successful transfer of glucose to the BP anchor yields a very hydrophobic product that can be separated by TLC. Any glucose that was not transferred remains at the origin due to the polarity of the nucleotide-bound sugar. A limitation in the assay is assuming that BP remains prevelant in the microenvironment throughout the membrane fraction isolation to be utilized by WcaJ. Without any mass analysis of the final products, it is impossible to guarantee that the glucose sugar is being transferred to BP and not some other

hydrophobic compound in the membrane fractions. Further, by assuming that the BP anchor was native to the membrane fraction, there is no way to accurately quantify the activity of WcaJ because the amount of total BP present could not be known.



Scheme 1: Catalyzed transfer of glucose to bactoprenyl phosphate by WcaJ.

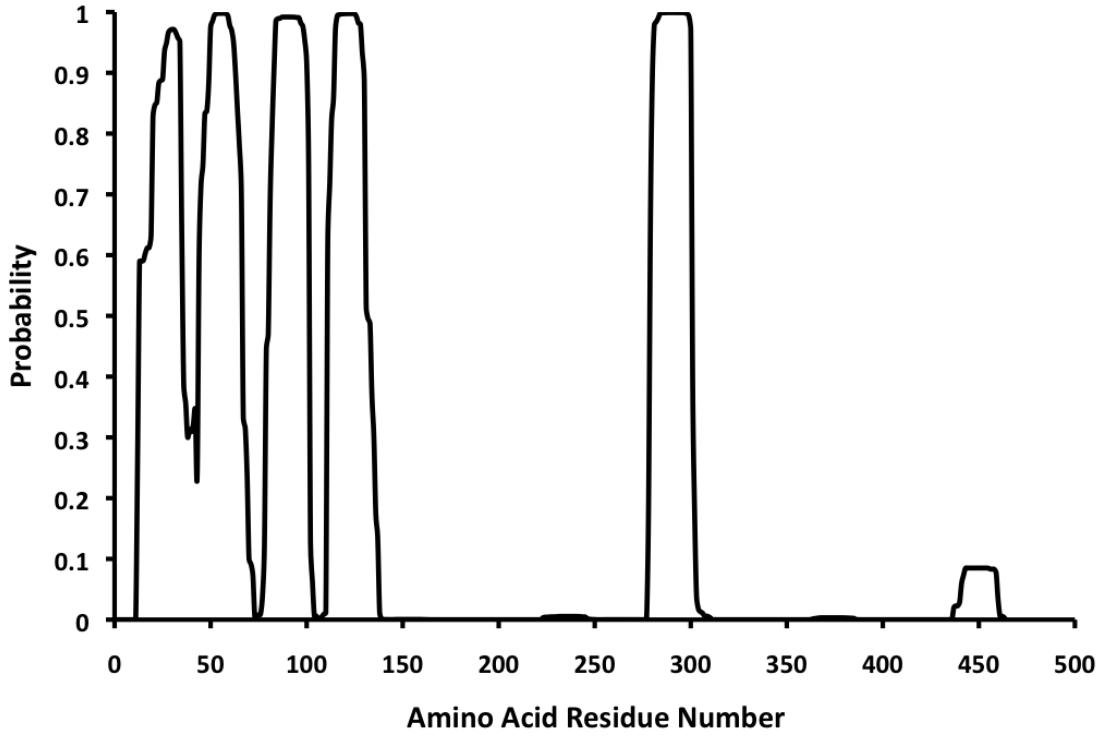
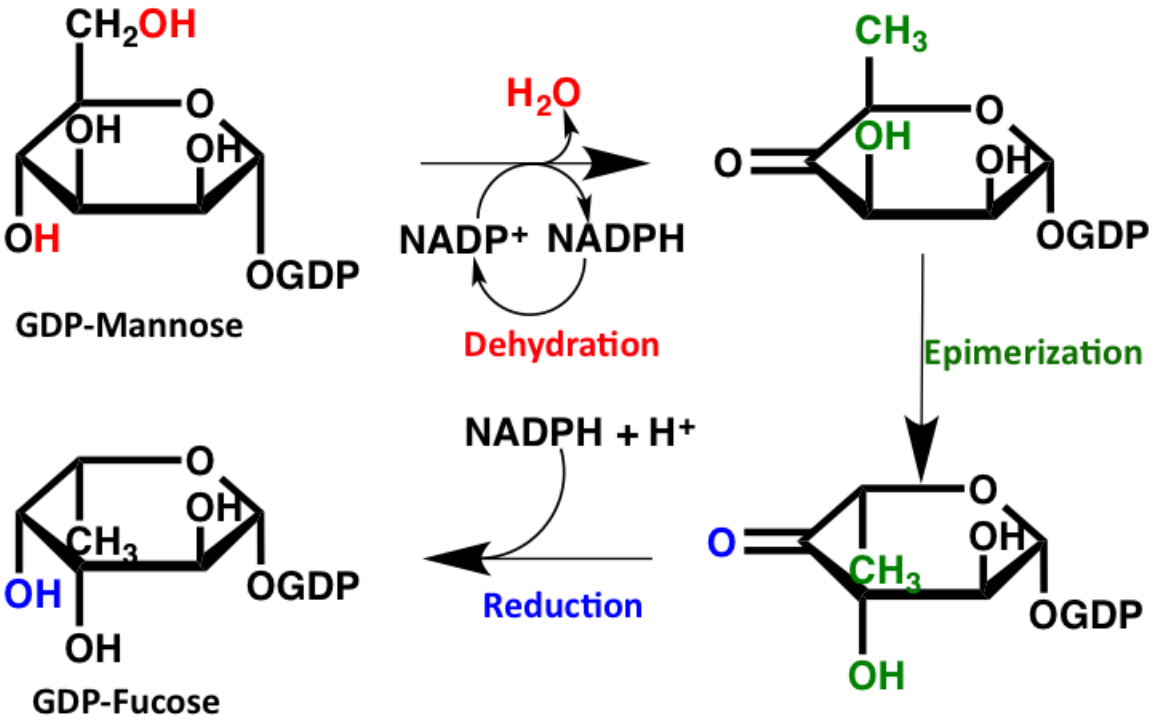


Figure 4: Solubility profile for WcaJ. The TMHMM server predicts the location and number of transmembrane domains based upon hydrophobic residues within the amino acid sequence. The x-axis represents the amino acid sequence for WcaJ, with the origin representing the start of the sequence. The y-axis represents the probability that a specific amino acid residue is part of a transmembrane domain.

The functional characterization of an initiating-glucose-1-phosphate transferase unrelated to the colanic acid biosynthetic pathway has been reported. Cps2E is the initating-glucose-1-phosphate that is critical for capsule biosynthesis in *Streptococcus pneumoniae*.[16] In the study reporting its characterization, Cps2E was shown to transfer glucose-1-phosphate to isoprenoid monophosphates of various lengths. This group also utilized radiolabeled TLC, but provided isoprenoid substrates to the membrane fractions containing Cps2E. These results are promising because they provide a potential alternative to WcaJ, should there be issues observing activity of WcaJ.

1.4.2: Gmd and Fcl, GDP-Fucose Biosynthesis Enzymes

The biosynthesis of GDP-fucose is a well-understood process as the production of the GDP-fucose is not novel to this pathway.[10-12] GDP-fucose is synthesized from GDP-mannose by two enzymes, Gmd and Fcl. Gmd is a dehydratase that removes water from GDP-mannose by converting the 4' position on the sugar to a keto group (Scheme 2, red).[11] The product of this reaction is then epimerized (Scheme 2, green) and the 4-keto group is reduced by Fcl to form GDP-fucose (Scheme 2, blue).[12]



Scheme 2: Biosynthesis of GDP-fucose from GDP-mannose.

Various aspects regarding the synthesis of GDP-fucose by these two enzymes are well known. Both Gmd and Fcl have crystal structures and solved mechanisms. [11, 17]Further, NADP+ and GDP-mannose binding affinities and catalytic efficiency values

for Gmd have been reported.[11] Fucose has various biological roles in many organisms including other bacterial species and even humans. Fucose is a structural component of the lipopolysaccharide repeat units for various organisms including *Helicobacter pylori* [18]and *Bacteroides fragilis*.[19, 20] Interestingly, fucose also plays a role in *H. pylori* similar to that in the colanic acid pathway.[11, 18] One group identified an Fcl homologue (WbcJ) that was required for the biosynthesis of O-antigen for protection against acid degradation by that organism.[18] Mutant strains of the organism were genetically altered so that they could not express WbcJ and found that at pH 7.0, there was no significant difference in the number of colonies of wild type and mutant strains. However, when the pH was lowered to 4.0, there was a drastic decline, by two log units, in the number of mutant strain colonies observed. At the final pH tested of 3.5, there were no observable mutant colonies present. In humans, fucose is a component of the blood-type antigens. [11, 21] For each of the antigens (A, B, O), fucose is bound to a GlcNAc sugar within the antigen.[21] For all of these reasons, the study of fucose is widespread and the demand for GDP-fucose is critical to these studies. One group reported the biosynthesis of GDP-fucose in one-liter batches with a maximum production of 55.2 mg/L in culture with cells overexpressing Gmd, Fcl and G6PDH that acts as a NADPH regenerator.[22]

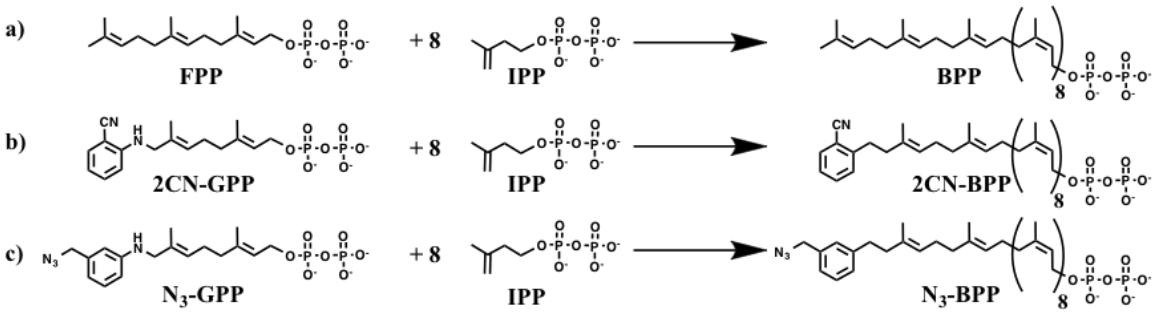
1.5: Approach for Functional Characterization of the Colanic Acid Biosynthesis Pathway

The use of radioactive materials is a common approach for characterizing bacterial polysaccharide synthesis.[9] While this method has provided support for characterizing activity of biosynthetic enzymes, this method has serious limitations including the requirement of expensive materials that: require special transport, are difficult to use for multiple steps in enzymatic synthesis and generate hazardous waste.

As a means to circumvent these issues, the Troutman lab has synthesized fluorescent BP analogues to study various bacterial capsular polysaccharide biosynthesis pathways. [23, 24] To do this, fluorescent analogues of one of the small-isoprenoid precursors to BP, farnesyl diphosphate (FPP) were synthesized. Fluorescent FPP analogues have also been utilized to study protein farnesylation, a post-translational modification made to a wide range of enzymes including G-coupled protein receptors and nuclear lamin proteins. [25, 26] *In vivo*, bacteria use FPP as a precursor to the BP isoprenoid anchor that is required for all capsular and exopolysachharide biosynthesis. Bactoprenyl diphosphate (BPP) is synthesized enzymatically by undecaprenyl pyrophosphate synthase (UPPS), which catalyzes the condensation of eight isopentenyl diphosphate (IPP) molecules to farnesyl diphosphate (FPP) (Scheme 3a).[27] A cellular phosphatase cleaves a phosphate group prior to form the active monophosphate form (BP). The Troutman group found that fluorescent analogues of FPP could be used as substrates by UPPS to generate fluorescent BP anchors to study the synthesis of various capsular polysaccharides (Scheme 3b). Specifically, a fluorescent analogue was used to characterize WcfS, the initiating hexose-1-phosphate transferase responsible for transferring acetamido-4-amino-6-deoxygalactopyranose-phosphate (AAD-Gal-P) to BP in the capsular polysachharide A (CPSA) biosynthesis pathway.[28] This same method will be utilized to characterize the glycosyltransferases of the colanic acid biosynthesis pathway using a shorter BP analogue (2-CN-5Z-BP, Scheme 3b for full length).

This thesis also focuses on the development of a multi-functional BPP probe that could improve detection of sugar transfer reactions, but also become a tool for unique applications beyond what has been done before in the Troutman laboratory (Scheme 3c).

Utilizing well-understood Huisgen chemistry, specifically a 1,3-dipolar cycloaddition, any molecule with a terminal alkyne can attach to species with an azide group.[29] The goal was to develop an analogue to FPP with an azide moiety that UPPS could utilize as a substrate for BPP synthesis. Previously, a group studying protein farnesylation published the synthesis of the FPP analogue with an azide moiety.[30]



Scheme 3: UPPS utilization of FPP (a) and analogues (b & c).

1.6: Summary of Thesis Project

Due to the prevalence of virulent *E. coli* outbreaks and the rise of antibiotic resistance, the development of new, targeted antibiotics is of major importance. The production of colanic acid by virulent *E. coli* to escape acid degradation is a potential pathway for novel antibiotic development. To elucidate antibiotic targets involved in colanic acid biosynthesis, a much better understanding of how this sugar polymer is built is required. Beyond the initiating-hexose-1-phosphate transferase, no other CA sugar-transferring enzyme has been characterized. The use of fluorescent analogues of the BP anchor to monitor each glycosyltransferase reaction will overcome serious limitations of established radiolabelled protocols and are key to the success of this project.

Understanding how *E. coli* synthesize colanic acid is more complex than just characterizing the glycosyltransferase responsible for each specific step in the pathway. Noting the structure of the repeat-unit, the first fucose residue is acetylated at the 2’ and

3’ hydroxyls (Figure 1). These acetylation events may be critical to the biosynthesis of the repeat unit as either acetylation or both acetylations could be regulatory. There are three hypothesized models for when these acetylation events could occur in CA biosynthesis. The first model suggests that free GDP-fucose is acetylated prior to incorporation into the repeat unit (Figure 5a). Another possibility is that both acetylations occur on the isoprenoid, after the formation of the disaccharide (BPP-Glc-Fuc) (Figure 5b). The third model suggests that monoacetylation or diacetylation of fucose occuring on the isoprenoid-bound repeat unit do not occur at the same time (Figure 5c). If acetylation of the fucose residue is indeed regulatory, then understanding which model is the true representation is critical to targeting this process. If regulatory acetylation events occur early in the synthesis (Figure 5b), then the acetyltransferases become great antibiotic targets, as this would prevent the majority of the repeat unit from being synthesized. However, if the events occur late in the process, after the incorporation of all the sugars, then export may not be prevented and targeting the acetyltransferases may not be the appropriate route to pursue.

Elucidating the specifics of colanic acid biosynthesis including the role of acetylation in its formation requires entering largely uncharted territory. As previously stated, the enzyme responsible for transferring the initial glucose-1-phosphate has been characterized. [(9)] To properly determine the relevance of these acetylation events, both fucosyltransferase enzymes will need to be functionally characterized and the function of each of the potential acetyltransferases must be verified. Therefore, the genes that encode each glycosyltransferase, both acetyltransferases, and the fucose biosynthesis proteins

must all be isolated from the *E.coli* genome for eventual protein expression and functional assay—all of which will be explained in Chapter 2.

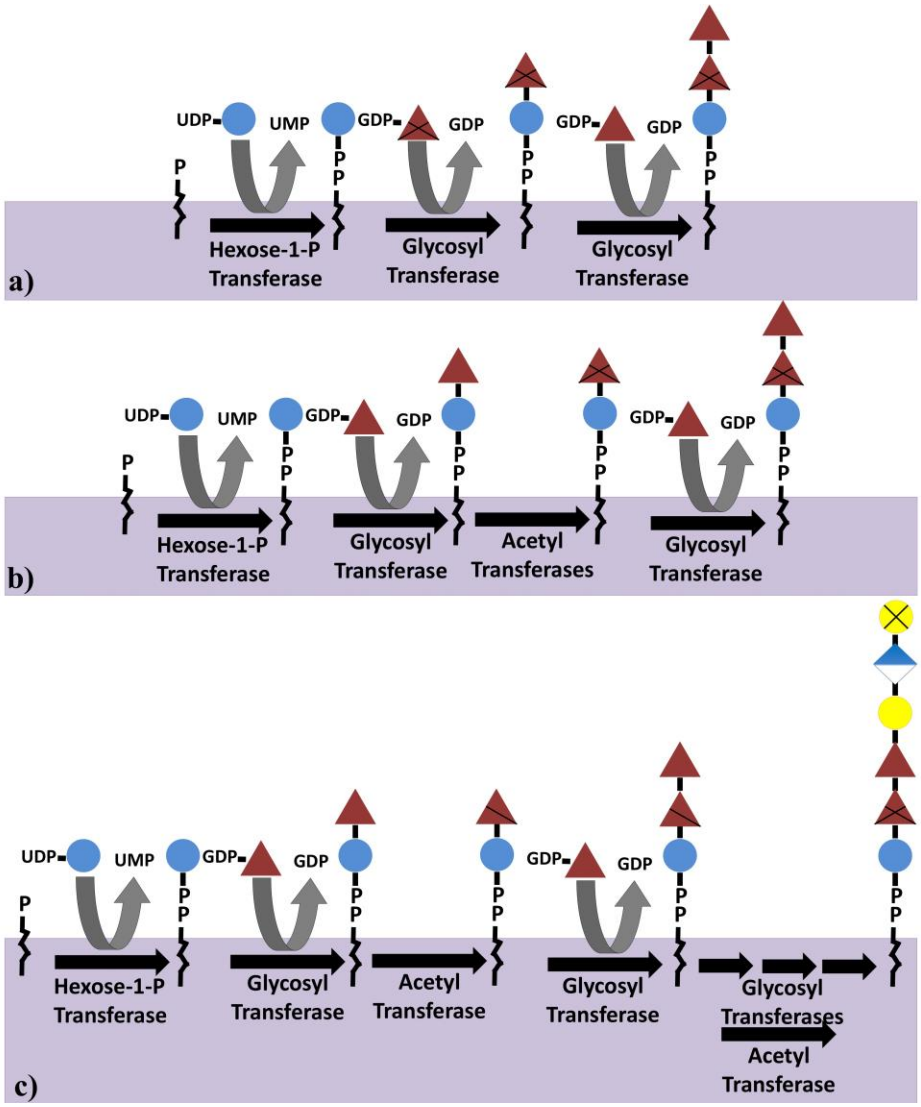


Figure 5: Potential models for acetylation in CA biosynthesis. a) Free GDP-fucose is acetylated prior to incorporation into the repeat unit. b) Both acetylations occur on the isoprenoid prior to formation of trisaccharide. c) One acetylation is required for continued synthesis beyond the disaccharide with the second occurring at some undetermined point afterwards.

The fluorescent BP anchor (2CN-5Z-BP) has significant value, as they allow for easy monitoring and characterization of biosynthetic reactions by reverse high-performance liquid chromatography (HPLC). Starting with BP-bound glucose (the product of WcaJ), each glycosyltransferase will be tested to determine which enzyme transfers fucose to form the disaccharide product. Utilizing reverse phase-HPLC with a C_8 column and a fluorescence detector, successful transfer of a fucose residue will result in a product with an earlier retention time (Figure 6). Each sequential sugar addition product will have earlier retention times, as more polar molecules interact less with the hydrocarbons within the column.

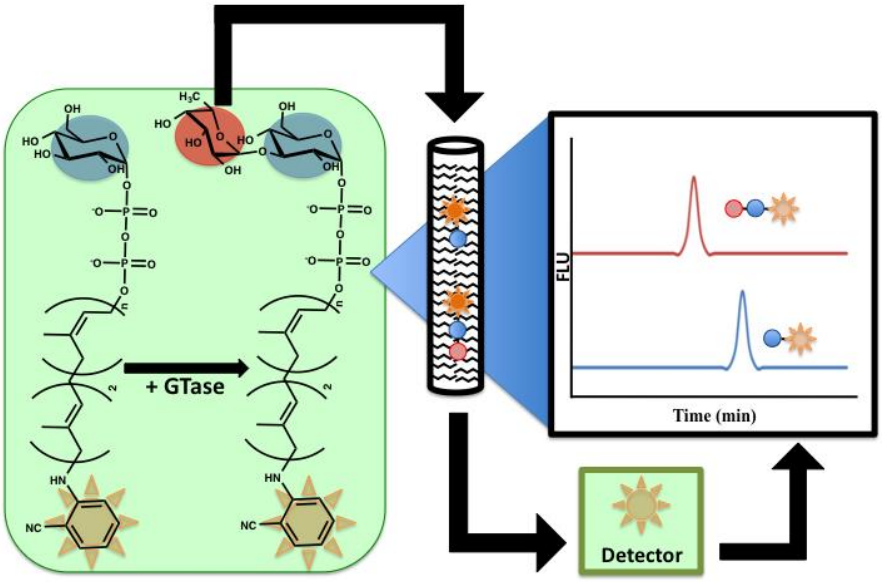


Figure 6: Reverse HPLC analysis of fucosyltransferase reaction. Each glycosyltransferase will be tested with BP-bound glucose and GDP-fucose to determine which transfers the sugar to the isoprenoid bound monosaccharide (green oval). Once reactions have incubated, they will be analyzed through reverse-phase HPLC with a C_8 column. The disaccharide product (BP-Glc-Fuc) will have an earlier retention time than the starting material (BP-Glc) due to the increased polarity of the added sugar that decreases interactions with the hydrocarbons of the column.

CHAPTER TWO: METHODS

2.1: Ligation-Independent Cloning (LIC):

Polymerase Chain Reaction (PCR) was utilized to amplify all the genes that encode the colanic acid biosynthesis proteins. The PCR for each gene was performed on an Eppendorf MasterCycler thermocycler. All PCR consisted of a 50 μL volume with the following components: 1x HF Buffer, 0.2 mM dNTPs, 0.2 μM Primer Mix, 0.5 μL *E. coli* template DNA (90.2 ng in reaction, 180.4 ng/μL stock), 1 unit Phusion High Fidelity DNA Polymerase and 37.5 μL of H_2O. All amplified gene products were purified using the ProMega Wizard SV Gel and PCR Clean Up kit.

A pET-24a vector was mutated at the nde1 and xhoI restriction sites to allow for digest by Sca1 and Msc1 (New England BioLabs). The typical digest reaction mixture consisted of 50 μL of vector DNA (concentration was always within restriction enzyme activity capability), 5 μL of Buffer 3.1 (New England BioLabs), 0.5 μL Bovine Serum Albumin (BSA) (100x) and 1 μL of each restriction enzyme (20,000 units/mL). The Sca1/Msc1 double digest was performed for ~3 hrs at 37 °C. All digested DNA was purified through a 1% agarose gel and subsequent use of the ProMega Wizard SV Gel and PCR Clean Up kit.

Once both amplified gene product and doubly digested DNA vector were isolated, each was treated with a T4 DNA Polymerase (ThermoScientific). In the presence of a

sole dNTP, the polymerase exhibits 3'→5' exonuclease activity that creates long single-stranded overhangs when the enzyme reaches a nucleotide on the DNA strand that matches the free dNTP in solution. Reactions for the gene products and isolated vector were performed separately, as each requires the presence of a different dNTP in solution. A typical reaction mixture consisted of a 20 μL volume with the following components: 1x T4 Pol. Buffer (5x, 4 μL), X ng DNA (reaction specific, 15.3 μL for the less concentrated between vector and insert), 2.5 mM of dNTP (dCTP for vector reaction, dGTP for gene insert reaction), 1 unit of T4 DNA Polymerase (5 units/uL, 0.2μL) and X μL of H_2O (up to 20 μL total volume). Reactions were incubated for at least half an hour and products were purified using the ProMega Wizard SV Gel and PCR Clean Up kit.

Once the T4-treated gene products and vectors were purified, products were mixed and incubated together at room temperature for half an hour. Because equal nanogram amounts of gene and vector DNA were treated with the T4 DNA polymerase, volumetric ratios between each can be set up to incubate. Most incubations were 1:1 or 1:3 vector to insert by volume. After the incubation period, reaction mixtures were transformed into DH5α cells, spread on LB/kanamycin plates and allowed to incubate overnight at 37 °C. A negative control consisting of DH5α cells transformed with T4-treated vector but no insert were also grown. Any colonies growing on plates containing gene inserts were grown in an overnight starter culture. The following day, these cultures were pelleted and plasmid DNA was extracted with the ThermoScientific GenJET Miniprep kit. Finally, the size of each isolated plasmid was analyzed by an agarose gel against a pET-24a vector without an inserted gene. Sequencing of the isolated vectors is done in an outside laboratory by Eurofins.

2.2: Expression and Induction of the Colanic Acid Biosynthesis Proteins

All pET-24a vectors containing each of the colanic acid biosynthesis genes were transformed into C41 expression cells, from which overnight cultures and glycerol stocks were made. Protein was typically expressed in 500 mL cultures. Once cells from the overnight culture were added, they incubated in a shaker at 37 °C until they reached an optical density (OD) of at least 0.4. Once an appropriate OD was reached, 1 mL of 0.5 M IPTG was added to each culture flask and the temperature of the incubator shaker was lowered to 30 °C. After 3 hours, the cultures were pelleted with an Avanti TM J-25I centrifuge (Eppendorf 5810 R centrifuge for 75 mL expression) for 15 minutes at 5000xg. The supernatant was then discarded and the pelleted cells were resuspended in a 0.9% NaCl solution. The resuspended cells were transferred to a 50 mL Falcon Tube and pelleted with an Eppendorf 5810 R centrifuge for 15 minutes at 5000 xg. The supernatant was again discarded and the pelleted cells were stored at -80 °C until needed.

2.3: Induction and Overexpression of Cps2E

A 500 mL culture containing cells with the vector to express Cps2E incubated while shaking at 37 °C until an optical density (OD) of at least 0.4 was reached. The vector containing Cps2E requires induction by arabinose rather than IPTG. After 1mL of arabinose (50 mg/mL stock) was added to the culture flask, the temperature of the incubator was lowered to 16 °C and allowed to shake overnight. The next day, cells were pelleted, resuspended in salt and pelleted for storage at -80 °C as described above.

2.4: Isolation and Purification of Soluble Colanic Acid Biosynthesis Proteins

Pelleted cells were resuspended in lysis buffer (50 mM Tris (pH 8), 300 mM NaCl, 20 mM imidazole). The resuspended cells were lysed by sonication with a Fischer Scientific Sonic Dismembrator Model 50b sonicator for two minutes. The lysed contents were pelleted by Eppendorf 5810 R centrifuge for an hour at 18514 xg. The supernatant containing the overexpressed protein was purified through affinity chromatography between a Ni^{2+}-NTA resin and a histidine tag encoded by the pET-24a vector. The column resin was 1-2 mL in volume. Between 50 and 100 mL of lysis buffer was poured through the column to equilibrate the resin by removing the ethanol that the resin is stored. The supernatant from the ultracentrifuge was poured through the column and collected as flow-through below. Next, 12 mL of wash buffer (50 mM Tris (pH 8), 50 mM imidazole, 300 mM NaCl) was poured through the column and collected as wash. Lastly, six 1mL aliquots of elution buffer (50 mM Tris (pH 8), 500 mM imidazole, 300 mM NaCl) were added by pipette to the column. All fractions were collected and stored at 4 °C and dialyzed. Elutions containing expressed proteins were pooled and dialyzed in 1000 mL of dialysis buffer (50 mM Tris, 300 mM NaCl) at 4 °C using SnakeSkin dialysis tubing. After three rounds of dialysis, glycerol was added (10% of the total volume) to preserve protein function when storing at -80 °C.

2.5: Isolation and Purification of Membrane-Bound Colanic Acid Biosynthesis Proteins

When expressing Cps2E, a short spin (2500 xg, 30 min) was done to remove cellular debris prior to ultracentrifugation. The remaining supernatant was spun down at 18514 xg for an hour. The pellet containing Cps2E was then homogenized in 1 mL of 10 mM MES (pH 6.0). This sample became the membrane fraction used in activity assays.

2.6: Verification of Protein

Both SDS-PAGE and Western Blot analysis were used to detect the purity and presence of the protein of interest, respectively. Fifteen μL of loading buffer was added to 30 μL of sample. All samples were boiled and spun down before being loaded. All gels were run at 180V for 40 minutes. Gels stained with Coomassie incubated for around half an hour before destaining overnight. SDS-PAGE gels designated for Western Blots were transferred onto a nitrocellulose membrane by running at 380 mA for 75 minutes. Anti-His tag Western Blotting was done to detect each protein of interest.

2.7: Cps2E Activity Assay

The reaction mixture contained 20.8 mM MES (pH 6.0), 1 mM $MnCl_2$, 0.1% Triton, 5.7 μM 2-CN-5Z-BP, 19 mM UDP-Glc, 0.000315 mg Cps2E membrane fraction. Reactions were done in 48 μL at room temperature. Reactions were analyzed by HPLC on a C_8 column with an isocratic mobile phase consisting of 49% propanol and 51% 100 mM NH_4HCO_3.

2.8: GDP-Fucose Biosynthesis Assay

General assay conditions consisted of 50 mM Tris buffer (pH 8), 10 mM $MgCl_2$, 0.5 mM GDP-mannose, 1 mM NADPH, 1 mM $NADP^+$, 0.053 mg total protein within the Gmd lysate and 0.037 mg total protein within the Fcl lysate. The final volume of each reaction was 200 μL. The fluorescence decrease as NADPH is oxidized to $NADP^+$ was monitored on a SpectraMax M5 spectrophotometer at an excitation wavelength of 340 nm and an emission wavelength of 450 nm in a 96-well plate. Once reactions reached completion, each was pelleted to remove lysate. Remaining supernatants were pooled, frozen and lyophilized. The resulting dried material was resuspended in water and GDP-

fucose was isolated by HPLC with 50 mM $NH_4C_2H_3O_2$ (pH 4.5) with an aminopropylsilane column with guard.

2.9: BPP-Glc Fucosyltransferase Assay

Each reaction mixture consisted of 50 mM Tris (pH 8), 10 mM $MgCl_2$, 0.1% Triton, 2 μM 2-CN-B5PP-Glc, 60 μM GDP-fucose and 1 μM WcaI. The final volume of each was 100 μL. The reaction was compared against a 2 μM 2-CN-B5PP-Glc standard, containing 50 mM Tris (pH 8), 10 mM $MgCl_2$, and 0.1% Triton. The reaction components were separated by HPLC on a C_8 column with a mobile phase consisting of 49% propanol, 51% 100mM NH_4HCO_3.

2.10: BPP-Glc-Fuc Acetyltransferase Assay

This assay was a one-pot reaction with the fluorescent BPP-Glc starting material. Each reaction contained 50 mM Tris (pH 8), 10 mM $MgCl_2$, 0.1% Triton, 2 μM 2-CN-B5PP-Glc, 60 μM GDP-fucose, 600 μM Acetyl-CoA, 1 μM WcaI and 4.5 μM WcaE. The presence of each acetyltransferase varied depending on which control was performed. When only one of the two acetyltransferases was present in a reaction, 4.7 μM WcaB or 1.9 μM WcaF was added. These same concentrations of each acetyltransferase were added to reactions containing both enzymes. The reaction was separated by HPLC on a C_8 column with a mobile phase consisting of 49% propanol, 51% 100 mM NH_4HCO_3.

2.11: BPP-Glc-AcFuc Fucosyltransferase Assay

This assay was a one-pot reaction with the fluorescent 2-CN-5Z-BPP-Glc starting material. Each reaction contained of 50 mM Tris (pH 8), 10 mM $MgCl_2$, 0.1% Triton, 2 μM 2-CN-B5PP-Glc, 60 μM GDP-fucose, 600 μM Acetyl-CoA, 1 μM WcaI, 4.7 μM

WcaB, 1.9 μM WcaF and 4.5 μM WcaE. The reaction was separated by HPLC on a C_8 column with a mobile phase consisting of 49% propanol, 51% 100mM NH_4HCO_3.

2.12: BPP-Glc-AcFuc-Fuc Galactosyltransferase Assay

This assay was a one-pot reaction with the fluorescent BPP-Glc starting material (2-CN-5Z-BPP-Glc). Each reaction contained of 50 mM Tris (pH 8), 10 mM $MgCl_2$, 0.1% Triton, 2 μM 2-CN-5Z-BPP-Glc, 60 μM GDP-fucose, 600 μM Acetyl-CoA, 1 μM WcaI, 4.7 μM WcaB, 1.9 μM WcaF, 4.5 μM WcaE and 8.9 μM WcaC. The reaction was separated by HPLC on a C_8 column with a mobile phase consisting of 49% propanol, 51% 100mM NH_4HCO_3.

2.13: Synthesis of the FPP Analogue

The synthesis protocol for the FPP analogue has been adapted and modified from Chehade *et. al* and Labadie *et. al*[25, 30].

2.13.1: 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.

2.13.2: 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%.

2.13.3: 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 R_f. 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. (^{1}H, $CDCl_3$, 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). R_f = 0.3. Yield = 33%.

2.13.4: 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 25mM 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). Yield =16%.

2.14: UPPS Utilization Assay with the 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphate

A reverse-phase HPLC assay was utilized to determine the ability for UPPS to use the synthesized azido-FPP as a substrate. Assay conditions consisted of 49.4 mM Bicine buffer, 0.5 mM $MgCl_2$, 4.9 mM KCl, 1.2% DDM, 1 mM azido-FPP analogue, 2.5 mM

IPP and 1.75 μL of UPPS with a total reaction volume of 101.25 μL. After incubation at room temperature, reactions were analyzed by HPLC by a propanol gradient method with 100 mM NH_4HCO_3 and monitor the absorbance of the associated aromatic ring at 250 nm. The gradient starts with 15% propanol that increases for half an hour until 80% propanol is reached. 80% propanol is pumped through the system for five minutes before the propanol concentration decreases back down to 15% over a five minute time period. The assay runs on a reverse-phase C_{18} column with a 40-minute total runtime.

CHAPTER THREE: RESULTS & DISCUSSION

3.1: Isolation and Molecular Cloning of the Colanic Acid Biosynthesis Genes

The first step in characterizing the first fucosyltransferase in the CA biosynthesis pathway is to isolate and clone each individual gene within the CA cluster to allow for transformation into protein expression cell lines (Table 1). All of the colanic acid biosynthesis genes responsible for the construction of the repeating-unit have successfully been amplified by PCR (Figure 7). Amplified PCR products were separated by electrophoresis through an agarose gel. A molecular gene ladder is also separated on the gel as a standard for size comparison.

Once each gene was amplified by PCR, ligation-independent cloning (LIC) was utilized to successfully clone each gene into a bacterial DNA vector. This is required for future expression of the CA biosynthesis enzymes. While LIC was an effective cloning technique, optimization could be done to reduce the largely qualitative rather than quantitative process. A significant limitation in how LIC was done disregards molar comparisons between the amount of vector and insert used in the annealing step. This was done because the large excess of free nucleotide from polymerase reaction (dCTP and dGTP) skewed accurate concentrations of the T4-treated products. Instead, equal nanograms of vector and insert were added to the polymerase reaction for each round of cloning. Considering that vector DNA is about five times larger than any gene being

inserted, it was roughly assumed that a 1:1 volumetric ratio between vector and insert DNA was one vector molecule to five gene insert molecules.

Table 1: Genes required for amplification and cloning within the CA gene cluster.

Protein Type	Role in CA biosynthesis	Gene(s) that Encode This Protein Type
GDP-fucose biosynthesis	Enzymatically synthesizing GDP-fucose from GDP-mannose	*gmd*, *fcl*
Glycosyltranferases	Attaching each sugar in the construction of the developing repeat-unit	*wcaA*, *wcaC*, *wcaE*, *wcaI*, *wcaL*
Acetyltransferases	Acetylating the first BPP-Glc bound fucose at the 2' and 3' position	*wcaB*, *wcaF*
Pyruvaltransferase	Pyruvating the terminal galactose residue in the repeat unit at the 4' and 6' positions	*wcaK*

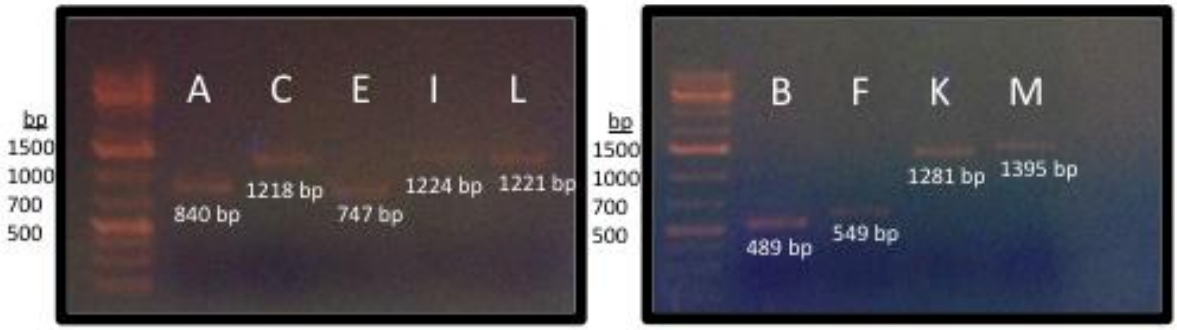


Figure 7: Amplification of CA biosynthesis genes. Lane designations: L = ladder, A,C,E,I,L,B,F,K,M = Wca_ genes. All of the genes fall between two bright orange bands on the ladder, 500bp and 1500 bp respectively. The majority of the amplified CA biosynthesis genes are displayed here including the glycosyltransferases (left gel), acetyltransferases and pyruvyltransferase (right gel).

A more accurate method would be to use the molecular weights for the vector and gene insert sequences to make solutions of equal concentration to incubate prior to transformation. The most accurate method would be to calculate an extinction coefficient for both the DNA insert and vector, but this process would certainly be tedious and

ultimately not worth the effort considering the minimal effect it would have on the process. Another weakness with this protocol is the use of the clean-up kit after the polymerase treatment. The kit was used to isolate the DNA from the polymerase, as mixing the vector and insert reactions would certainly result in 5' $\rightarrow$ 3' polymerization and disrupt the formation of the overhangs. However, assuming that equal amounts of DNA elute from each column has the potential to further skew the volumetric ratios of insert to vector during the incubation prior to transformation. Other groups who have used LIC tend to denature the polymerase through heat-shock before incubating the treated vector and insert DNA. While this method has not been attempted for this work, other members of the Troutman lab have tried to heat denature the polymerase but yielded no success with cloning. Regardless, elimination of the clean-up kit would improve the qualitative nature of the LIC method.

3.2: Isolation of the Colanic Acid Biosynthesis Proteins

All colanic acid biosynthesis proteins have been overexpressed, purified by affinity chromatography and analyzed by Coomassie staining and Western Blotting (Figure 8). Protein samples are separated by size through a polyacrylamide gel. A protein molecular weight ladder is used as a standard for size comparison. Staining the polyacrylamide gel with Coomassie allows for purity determination of a particular sample (Figure 6, Top). A pure sample should only have one blue band in that respective lane. Therefore, multiple bands are indicative of multiple proteins. However, Coomassie staining does not confirm the identity of a desired protein—even if the molecular weights match. For this reason, Western Blot analysis is done (Figure 8, bottom). A Western Blot utilizes antibodies specific to a structural feature on the protein that can be detected

through chemical means. All of the colanic acid proteins have an extra six histidine residues encoded at the end of each amino acid sequence that is recognized by a primary antibody. A secondary antibody conjugated with alkaline phosphatase recognizes and binds to the primary antibody. When the substrate for alkaline phosphatase is poured over the Western blot, a color change occurs and confirms the presence of the protein of interest.

While the majority of the CA biosynthesis proteins were expressed and isolated, WcaJ, the initiating hexose-1-phosphate transferase could not be successfully expressed. The structure of the enzyme contains five transmembrane domains that allow for insertion into the inner membrane of *E. coli*. Generally, membrane-bound proteins are more difficult to express due to the tendency to misfold and aggregate prior to membrane insertion. Although expressing WcaJ was unsuccessful, the Troutman lab had access to a homologous protein, Cps2E from *S. pneumoniae* to transfer glucose-1-phosphate to the fluorescent BP analogue. Using WcaJ to catalyze this reaction would have been more ideal because the enzyme is part of the colanic acid biosynthesis pathway. However, the function of WcaJ has already been characterized, and therefore substituting Cps2E for WcaJ is acceptable in terms of the goals being accomplished by this work.

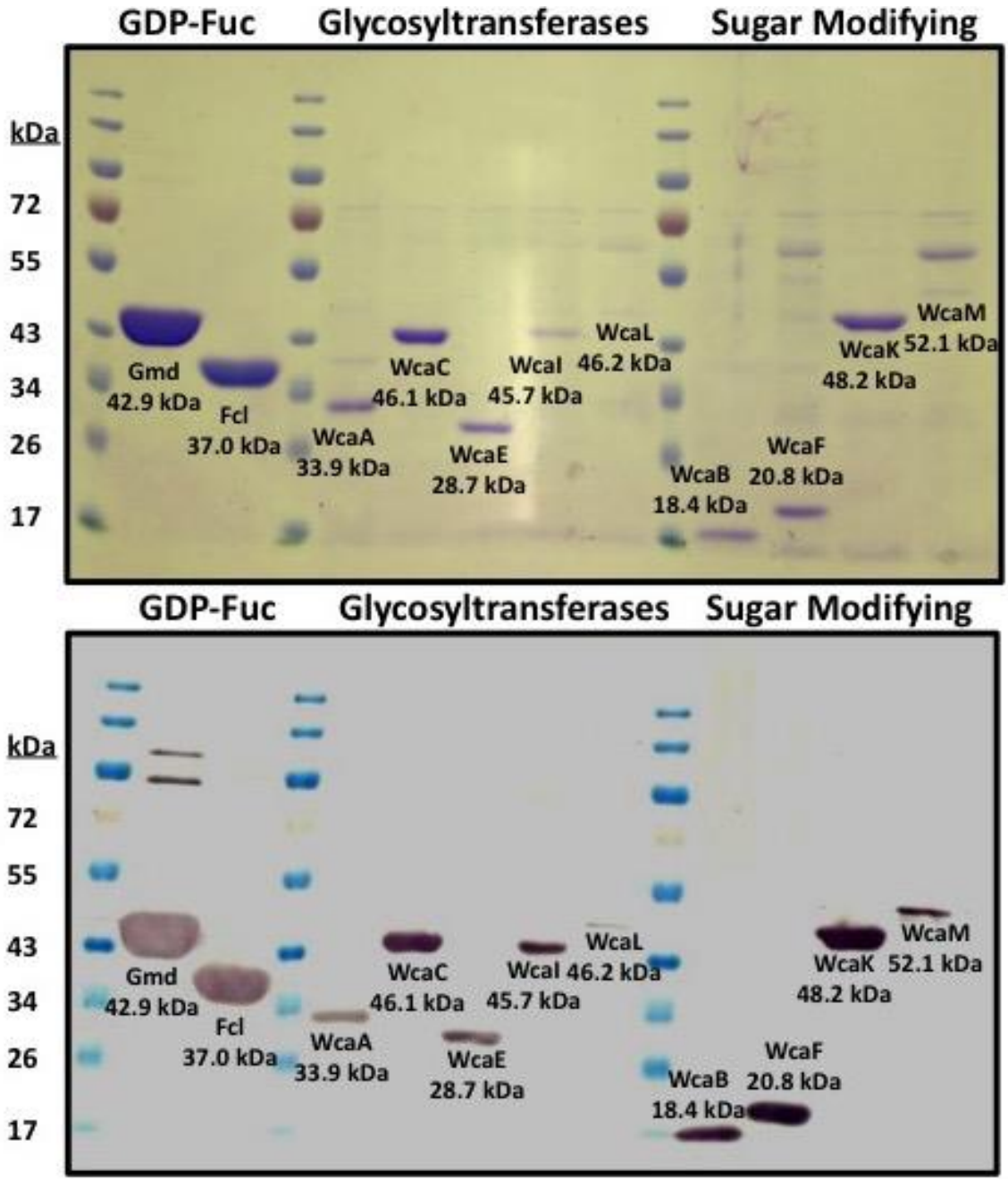


Figure 8: Expression of the CA Biosynthesis Proteins. Top) Coomassie stain. Each protein class is labeled above. Each individual protein is labeled with their respective molecular weights. Shows relative purity of each respective colanic acid protein. Bottom) Western blot. Same as above. Confirms the identity of the colanic acid biosynthesis proteins.

3.3: Cps2E Transfers Glucose-1-Phosphate to BP

As mentioned, Cps2E was utilized to transfer glucose-1-phosphate to the bactoprenyl anchor to circumvent the significant issues expressing WcaJ (Figure 7). Reactions containing 2CN-5Z-BP, UDP-Glc and Cps2E yielded a product with a retention time at 4 min (Figure 9, red line). This retention time is expected, considering that the addition of a polar sugar would cause less interaction with the hydrophobic C_8 column. Unfortunately, even though activity was observed with Cps2E, the enzyme did not completely turn over the starting material, which is evidenced by the peak at 9 minutes that is identical to that of the BPP-Glc control (Figure 9, blue line). The exact reason for this lack of activity is unknown. A potential reason for the poor function may be due to the isolation of Cps2E as a membrane fraction. Because of the required membrane-bound environment for Cps2E, the enzyme is isolated within the cell membrane along with every other membrane-bound protein that the organism expresses. Further, when Cps2E is added to a reaction, all proteins within the membrane are being added. The only way to quantify Cps2E is to report the amount of total protein added. Therefore, the exact concentration of Cps2E added is unknown and can contribute to a low product yield even if the *total* protein concentration is high. Cps2E can be extracted from the membrane in a detergent-rich environment that would simulate a membrane and could then be purified through affinity chromatography. However, optimizing this process is time consuming and does not guarantee that purified protein within the detergent will function as it would in a membrane fraction. Also, any unreacted starting material can be isolated by HPLC and reused. Even though full turnover was not

achieved, the isolation of the 2-CN-5Z-BPP-Glc product from the Cps2E became the starting point for further characterization of the colanic acid biosynthesis pathway.

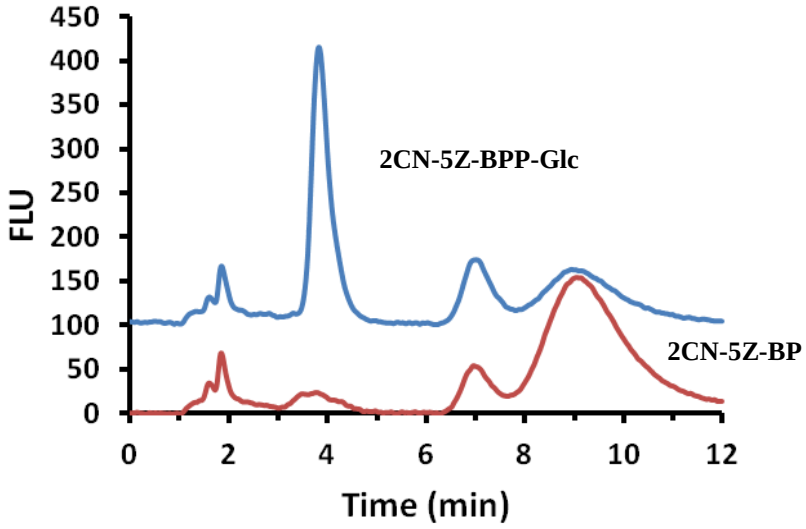


Figure 9: Cps2E transfers Glc-1-P to 2CN-5Z-BP. Reverse-phase HPLC analysis of Cps2E reactions. After 2CN-5Z-BP is treated with Cps2E and UDP-Glc, the reaction mixture is separated with a C_8 column in 49% propanol and 51% 100mM NH_4HCO_3. The successful Cps2E reaction (blue) yields two significant peaks: one at 9 minutes, which is unused 2CN-5Z-BP and one is at 4 minutes which is representative of 2CN-5Z-BPP-Glc.

3.4: Biosynthesis of GDP-Fucose

The isolation of the Cps2E product (2CN-5Z-BPP-Glc) yields the first of two substrates needed to characterize the first CA fucosyltransferase. The second substrate, GDP-fucose, can be synthesized enzymatically by a dehydratase (Gmd) and a synthetase (Fcl). The process is NADPH dependent, the oxidation of which causes a fluorescence decrease that can be monitored by spectroscopy in a 96-well plate. The fluorescence assay was an efficient approach to test a wide variety of reaction conditions for GDP-fucose biosynthesis without wasting materials. Initially, no activity was observed with this assay, even as conditions such as pH and temperature were changed (data not shown). Eventually, it was hypothesized that Gmd was not functional in the assay because a significant amount of the protein precipitated out of solution during the

purification process, potentially due to the high concentration of imidazole required for elution from the Ni-NTA resin. To circumvent this issue, Gmd was re-expressed and the cells containing the protein were lysed, but the protein was not purified. Instead, this lysate was added directly into a reaction. In doing so, there was an observed decrease in fluorescence of NADPH as the substrate is oxidized to the non-fluorescent NADP+ (Figure 10, light blue, purple).

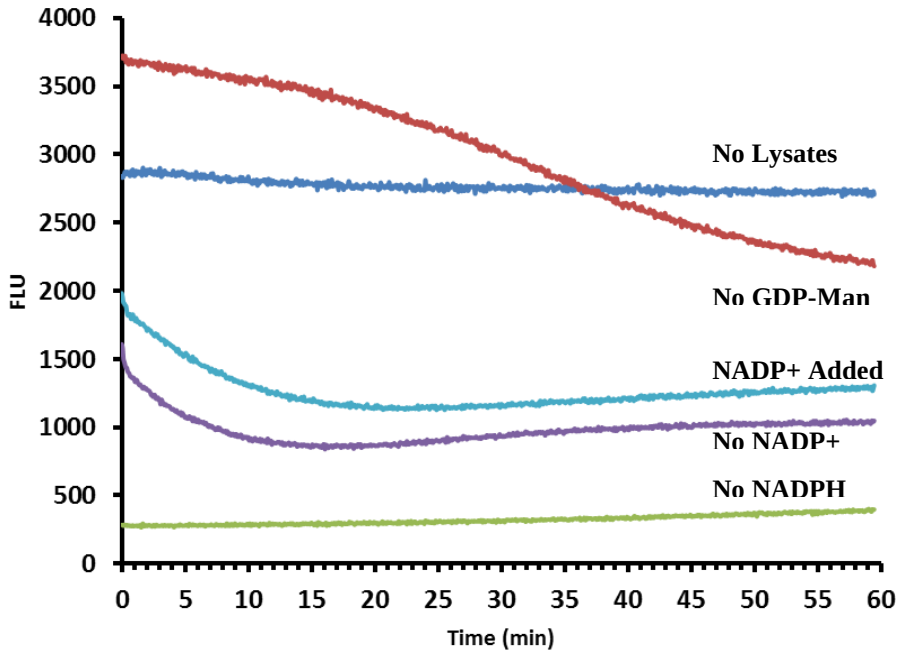


Figure 10: The biosynthesis of GDP-Fucose. GDP-mannose is converted to GDP-fucose using lysates of Gmd and Fcl. The process is NADPH-dependent. The utilization of NADPH results in decrease in fluorescence that can be monitored at excitation wavelength of 340 nm and an emission wavelength of 450 nm. The controls for this experiment were reactions without GDP-mannose (red), NADP+ (light blue) and NADPH (green).

The lack of fluorescence change in reactions without lysate and NADPH was not surprising. However, without the addition of GDP-mannose, there was still a decrease in fluorescence. This was not always observed but is indicative that addition of the lysates causes oxidation of NADPH at a much slower rate than if lysates were present.The GDP-

fucose was then purified from these reaction mixtures by HPLC and used in fucosyltransferase reactions.

3.5: WcaI Transfers Fucose to BPP-Glc

Once both substrates (2CN-5Z-BPP-Glc and GDP-Fuc) for the fucosyltransferase had been isolated, functional characterization of the enzyme was now possible. Because any of the five hypothesized glycosyltransferase enzymes (WcaA, WcaC, WcaE, WcaI, WcaL) could potentially transfer fucose to BPP-Glc, each has to be tested for activity. In doing so, it was determined that WcaI is the first fucosyltransferase in the CA biosynthesis pathway (Figure 11). Like the Cps2E reactions described in 3.3, addition of a sugar should produce a product with an earlier retention time than the starting material. Here, four of the glycosyltransferase reactions (WcaA, WcaC, WcaE, WcaL) yield products with retention times no different than that of the BPP-Glc starting material (data not shown). However, the WcaI reaction has a retention time earlier than the starting material, indicating that fucose has been transferred to the starting BPP-Glc (Figure 11, red). This observation has not been reported previously. Interestingly, the location of *wcaI* within the gene cluster (Figure 3) actually supports these results as the gene is near the fucose biosynthesis genes and between the GDP-mannose synthesis genes. *(10)*

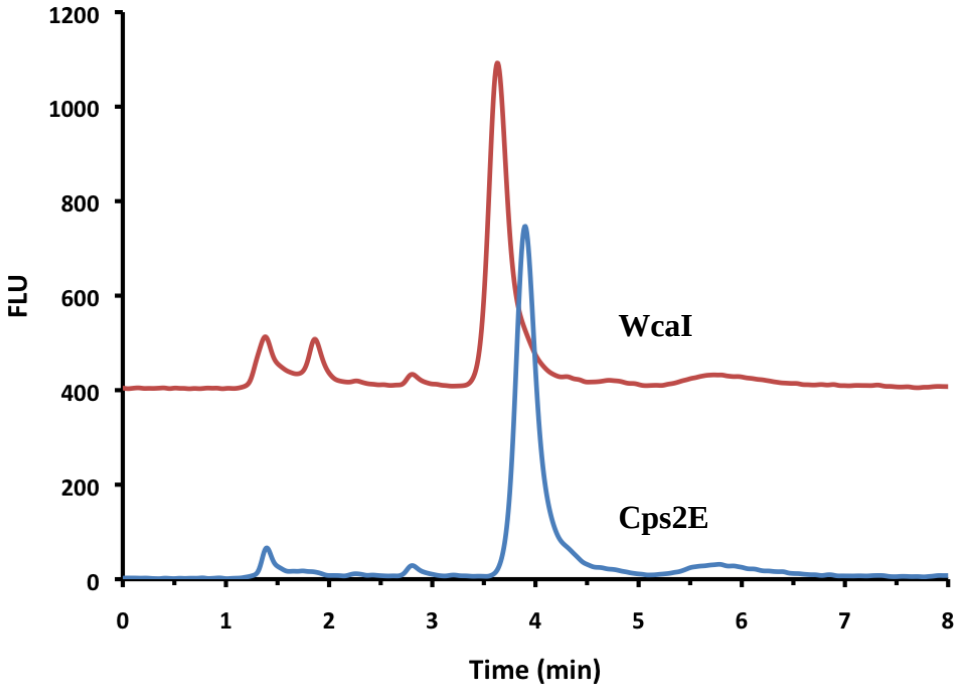


Figure 11: Characterization of the first Fucosyltransferase in CA biosynthesis. When the product of Cps2E (2-CN-5Z-BPP-Glc, blue) is treated with GDP-fucose and WcaI, there is an observed difference in retention indicating that the fucose has been transferred (3.75 min, red). Reaction mixtures were separated by reverse-phase HPLC with a C_8 column in 49% propanol and 51% 100mM NH_4HCO_3.

3.6: WcaE Transfers Fucose to BPP-Glc-Ac-Fuc

The isolated product from WcaI represents the first two sugars of the colanic acid repeating unit. The next sugar to be transferred in the repeat unit synthesis is another fucose sugar. To characterize this reaction, the product from WcaI and GDP-fucose are tested as substrates for the remaining glycosyltransferases (WcaA, WcaC, WcaE, WcaL). Because the first fucose sugar in the repeating unit is acetylated at the 2' and 3' positions, each reaction is coupled with the two proposed acetyltransferases (WcaB, WcaF) and Acetyl-CoA. In doing so, it was observed that WcaE transfers the second fucose sugar to BPP-Glc-Ac-Fuc, as WcaE is the only glycosyltransferase that gives a product with a

retention time earlier than that of the disaccharide starting material (Figure 12). The product of WcaE is a trisaccharide consisting of BPP-Glc-Ac-Fuc-Fuc.

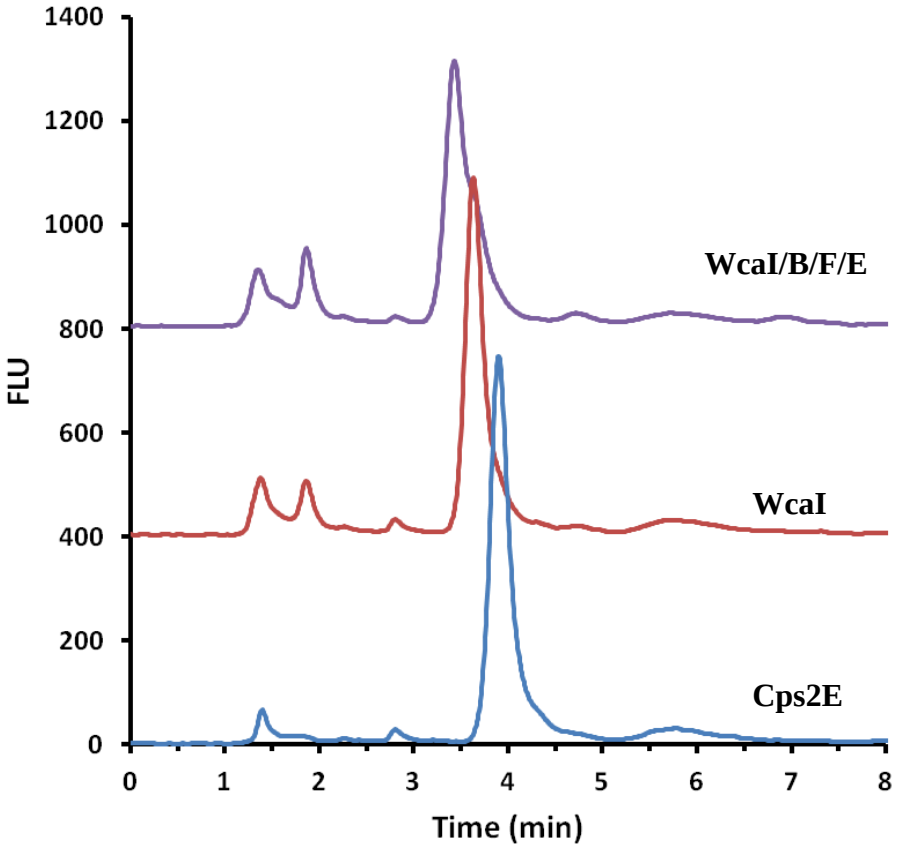


Figure 12: Characterization of the second fucosyltransferase in CA biosynthesis. When the WcaI product (2CN-5Z-BPP-Glc-Fuc, red) is treated with GDP-fucose, Acetyl-CoA, acetyltransferases (WcaB, WcaF) and WcaE, there is a change in retention time indicating that WcaE transfers the second fucose, forming a trisaccharide product (2CN-5Z-BPP-Glc-AcFuc-Fuc, purple). This is all done in a one-pot reaction starting with 2CN-5Z-BPP-Glc (blue). Reaction mixtures were separated by reverse-phase HPLC with a C_8 column in 49% propanol and 51% 100mM NH_4HCO_3.

3.7: Acetylation is Required for WcaE activity

When characterizing WcaE as the second fucosyltransferase in the colanic acid biosynthesis pathway, it was observed that the acetylation of the previous fucose residue was critical to the activity of WcaE. When WcaE was given unacetylated WcaI product (2CN-5Z-BPP-Glc-Fuc) as a substrate, the next fucose was not transferred (Figure 13, green). This observation suggested that acetylation is necessary for WcaE activity.

Because the reported structure of colanic acid consists of two acetylations on the first fucose sugar, the next step was to determine whether both acetylations are required for WcaE to function. To test this, unacetylated BPP-Glc-Fuc was treated with only one of the two proposed acetyltransferases (WcaB or WcaF) or treated with both acetytransferases (WcaB and WcaF), coupled with WcaE (Figure 14). Interestingly, while the addition of only WcaB did not promote fucosyltransferase activity by WcaE (Figure 14, light blue), the addition of only WcaF was enough to observe activity (Figure 14, purple). Considering the three acetylation models presented in the introduction (Figure 5), this data suggests that there is an aceylation by WcaF once the disaccharide product is released by WcaI and is required for further synthesis of the CA repeat unit.

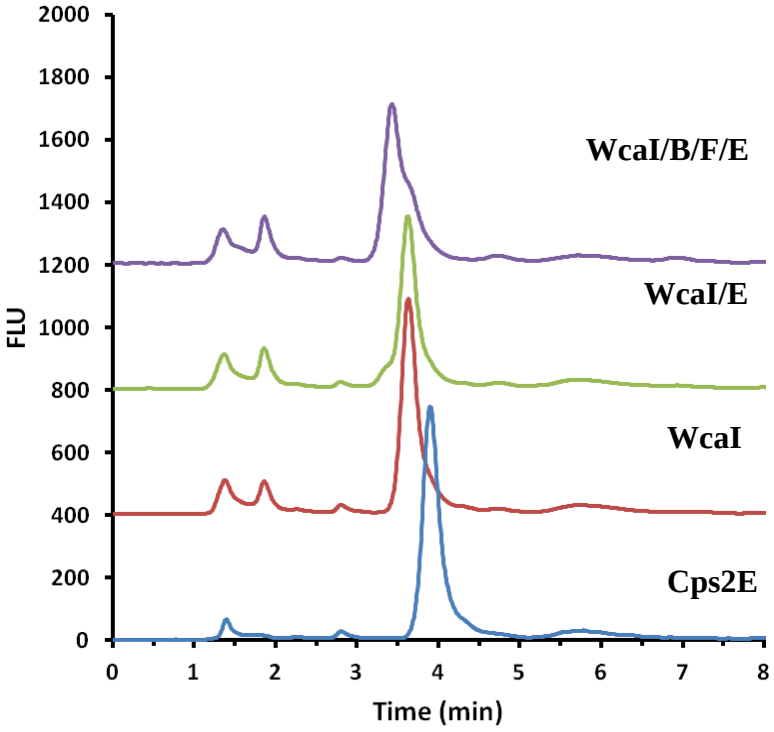


Figure 13: Acetylation is required for WcaE Activity. The formation of the trisaccharide product (2CN-5Z-BPP-Glc-AcFuc-Fuc, purple) requires the presence of acetyltransferases (WcaB and WcaF) or no activity will occur (green). This is all done in a one-pot reaction starting with 2CN-5Z-BPP-Glc (blue). The product with a retention time of 3.75 min resembles the disaccharide product (2CN-5Z-BPP-Glc-Fuc, red).

Further, these results were significant because they suggest that acetylation is regulatory for synthesis of colanic acid. From an application standpoint, these results provide potential antibiotic targets for new therapeutic design. A common antibiotic target related to this pathway involves the inhibition of UPPS, the protein that synthesizes the BP anchor. [31] The issue with these types of antibiotics is that they are not selective, they will kill both pathogenic and nonpathogenic bacteria. BP is involved in biosynthetic processes common to most bacteria including peptidoglycan biosynthesis and several types of capsular polysaccharide and peptidoglycan biosynthesis. Targeting colanic acid production would promote selection against the virulent bacteria that are evading acid degradation.

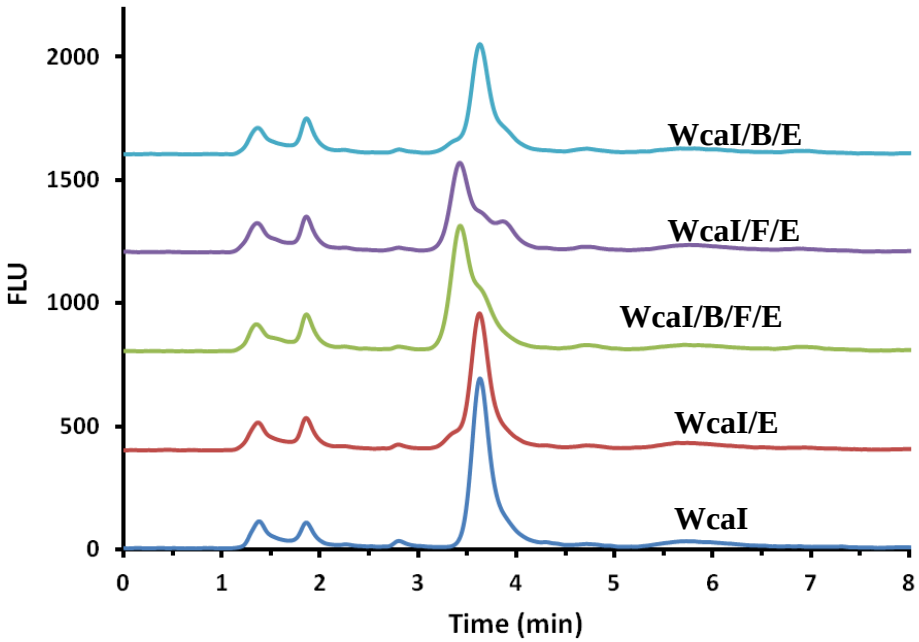


Figure 14: The role of acetylation events in the formation of 2CN-5Z-BPP-Glc-AcFuc-Fuc. The formation 2CN-5Z-BPP-Glc-AcFuc-Fuc, is an acetyltransferase-dependent process. When both WcaB and WcaF are present, WcaE transfers fucose to the developing repeat unit (green). This same result was observed when treated with WcaF alone (purple) but not when treated with WcaB alone (light blue). Each reaction was one-pot starting with 2CN-5Z-BPP-Glc. The product with a retention time of 3.75 min resembles the disaccharide product (2CN-5Z-BPP-Glc-Fuc, blue) Reaction mixtures were separated by reverse-phase HPLC with a C_8 column in 49% propanol and 51% 100mM NH_4HCO_3.

3.8: Further Characterization of the Colanic Acid Biosynthesis Pathway

The work thus far has described the biosynthesis of a trisaccharide product resembling half of the colanic acid repeating unit (BPP-Glc-2,3 Ac-Fuc-Fuc). Some preliminary experiments have been done with the remaining glycosyltransferase enzymes (WcaC, WcaA, WcaL) to potentially characterize further biosynthesis of the colanic acid repeating unit. By reacting the product of WcaE (BPP-Glc-2,3 Ac-Fuc-Fuc) with UDP-galactose and each remaining glycosyltransferase enzyme, it was determined that WcaC is responsible for the formation of the tetrasaccharide product (BPP-Glc-Ac-Fuc-Fuc-Gal, Figure 15). These results further suggest the regulatory role that aceytlation has in this pathway considering that all of the WcaE product within the WcaC reaction has been turned over with the only intermediate left in the reaction is the WcaI product (light blue).

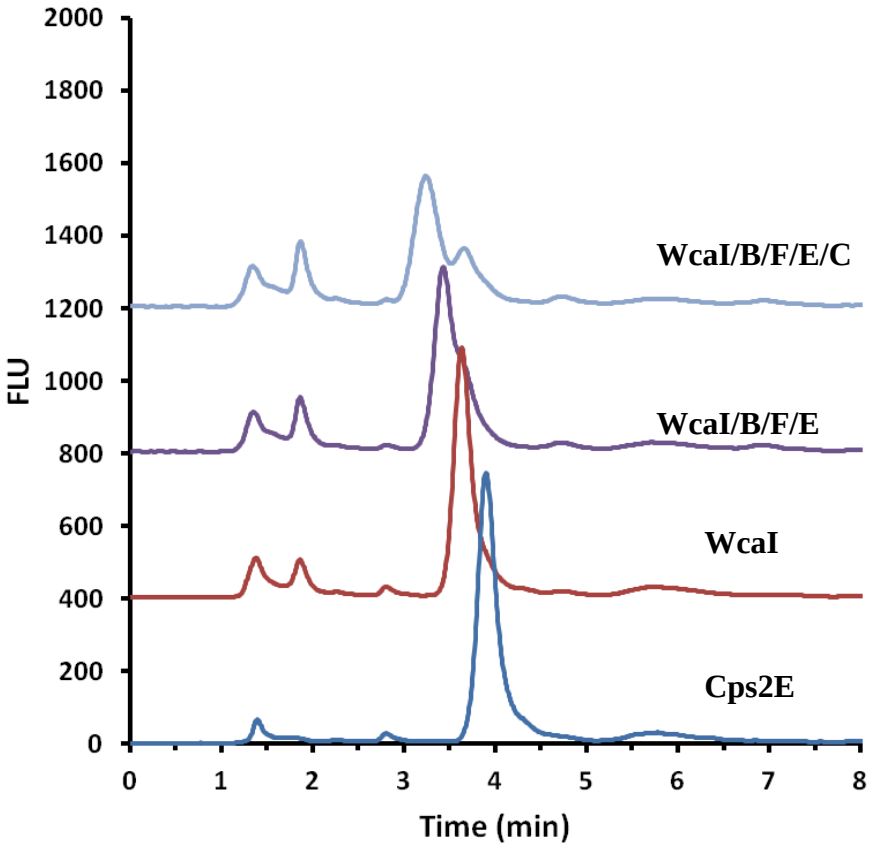
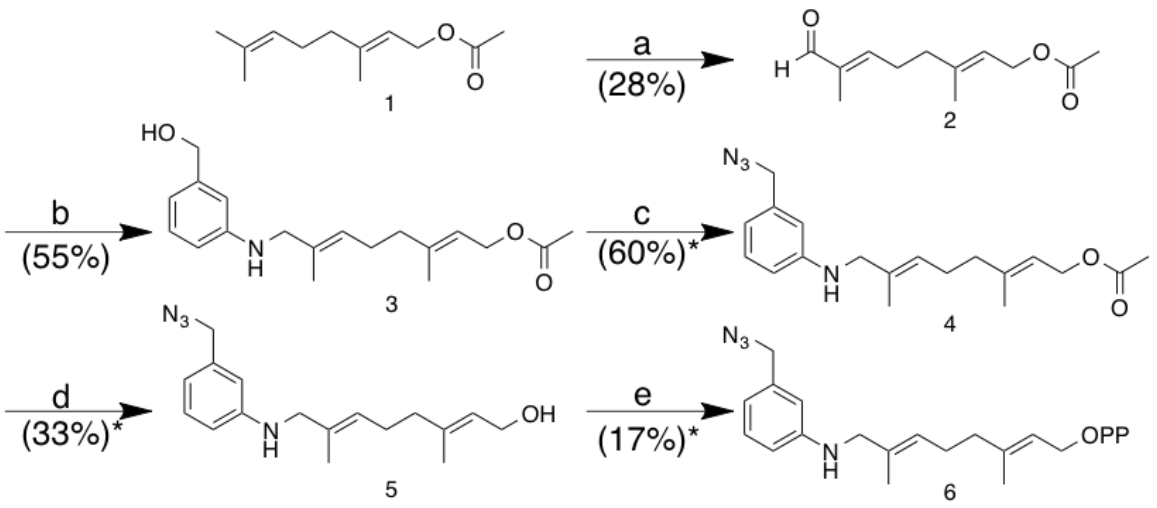


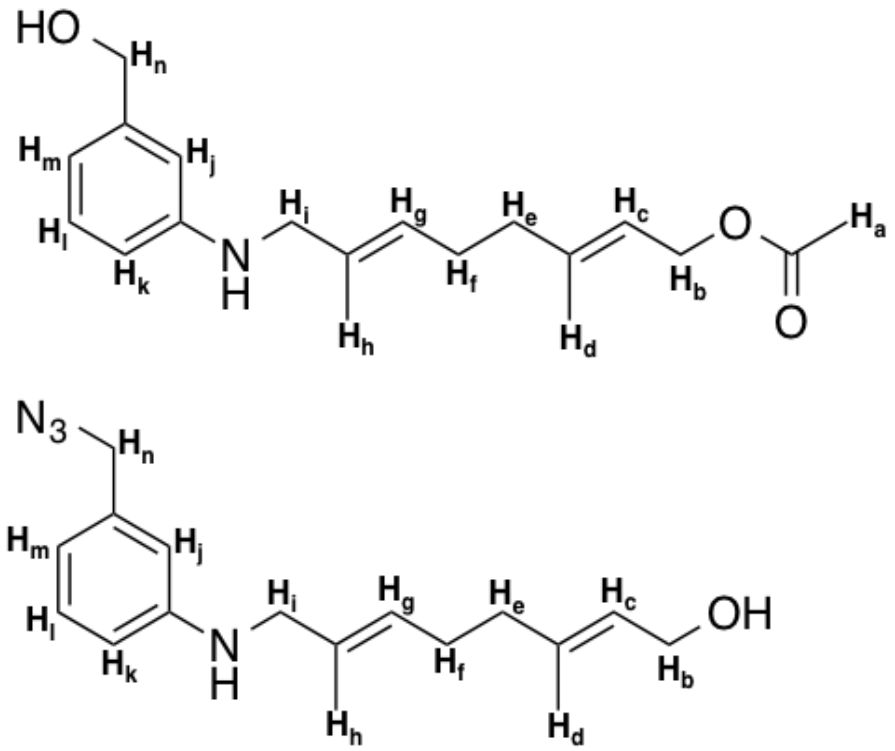
Figure 15: Characterization of the first galactosyltransferase in CA biosynthesis. When 2CN-5Z-BPP-Glc-AcFuc-Fuc (purple), is treated with UDP-Gal and WcaC, the formation of a tetrasaccharide product is formed (2CN-5Z-BPP-Glc-AcFuc-Fuc-Gal, light blue). This data shows all steps solved in CA biosynthesis including the disaccharide (red), trisaccharide (purple) These reactions were done in one-pot starting with 2CN-5Z-BPP-Glc (blue). Reaction mixtures were separated by reverse-phase HPLC with a C_8 column in 49% propanol and 51% 100mM NH_4HCO_3.

3.9: Synthesis of a Multi-Functional Probe for Glycan Biosynthesis

At this time, the synthesis of the azido FPP analogue has been somewhat crude with significant improvement needed at various steps in the synthesis (asterisks, Scheme 4). The yields for each are provided in the parentheses in Scheme 4. Notably, many of them are low. However, the NMR-associated chemical shifts obtained for each intermediate and final product are very similar (Table 2-3).



Scheme 4: Synthesis of 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.



Scheme 5: Labeled protons for ^{1}H-NMR for acetoxygeranyl benzyl alcohol (top) and 8-N-m-benzyl alcohol-amino-3,7-dimethyl-2,6 octadien-1-ol (bottom)

Table 2: Comparison of ^{1}H-chemical shifts for intermediates of 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphate synthesis between reported and observed

Compound Name	^{1}H-Chemical Shifts Reported in Literature[30]	^{1}H-Chemical Shifts Observed
8-N-*m*- benzyl alcohol-amino-3,7-dimethyl-1-acetoxy-2,6 octadiene (Scheme 4, 3) (Scheme 5, top)	($CDCl_3$, δ): 7.15 (t, 1H), 6.67 (d, 1H), 6.53 (d, 1H), 6.62 (s, 1H), 5.38 (t, 1H), 5.31 (t, 1H), 4.61 (s, 2H), 4.55 (d, 2H), 3.65 (s, 2H), 2.20-2.00 (m, 4H), 2.06 (s, 3H), 1.70 (s, 3H), 1.67 (s, 3H)	($CDCl_3$, δ): 7.19 (t, 1H, H_l), 6.73-6.59 (m, 3H, $H_{j, k, m}$), 5.44 (m, 2H, $H_{c, g}$), 4.61-4.65 (m, 4H, $H_{b, n}$), 3.70 (s, 2H, H_i), 2.22-2.11 (m, 7H, $H_{a, f, e}$), 1.78 (s, 3H, H_d), 1.75 (s, 3H, H_h)
8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadien-1-ol (Scheme 4, 5) (Scheme 5, bottom)	($CDCl_3$, δ): 7.16 (t, 1H), 6.62 (d, 1H), 6.56 (d, 1H), 6.53-6.50 (m, 1H), 5.39 (q, 1H), 4.24 (s, 2H), 4.13 (d, 2H), 3.65 (s, 2H), 2.20-2.00 (m, 4H), 1.67 (s, 6H)	($CDCl_3$, δ): 7.19 (m, 1H, H_l), 6.61 (m, 3H, $H_{j, k, m}$), 5.41 (m, 2H, $H_{c, g}$), 4.23-4.12 (m, 4H, $H_{b, n}$), 3.70 (s, 2H, H_i), 2.20-2.04 (m, 4H, $H_{f, e}$), 1.75(m, 6H, $H_{d, h}$)

3.9.1: Synthesis of Acetoxy Geranyl Aldehyde

The first step in the synthesis of 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphate is the oxidation of acetoxygeranyl acetate to (E,E)-3,7-Dimethyl-1-acetoxy-2,6-octadien-8-al (Scheme 4, a). Unfortunately, this reaction does not go to completion from the alcohol intermediate, which contributes to the 28% yield for this reaction. Another issue with this reaction is that stored aldehyde (Scheme 4, 2) can potentially convert back to the alcohol intermediate, even at -20°C. This becomes problematic when purifying the reductive amination product (Scheme 4, 3) because it has an R_f value close to that of the alcohol intermediate, which contaminates the yield of the subsequent reductive amination.

3.9.2 Synthesis of 8-N-m- benzyl alcohol-amino-3,7-dimethyl-1-acetoxy-2,6 octadiene

A reductive amination of acetoxy geranyl aldehyde with 3-amino benzyl alcohol produces 8-N-m- benzyl alcohol-amino-3,7-dimethyl-1-acetoxy-2,6 octadiene. This reaction yielded a fairly respectable yield of 55%. Further, the observed ^{1}H-chemical shifts matched well with what was reported in the literature (Table 2, Appendix Figure 1). The reported spectra as described by Labadie *et al.* seems to have better resolved peaks, especially within the 6.73-6.59 ppm region. *(30)*

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

8-N-m- benzyl alcohol-amino-3,7-dimethyl-1-acetoxy-2,6 octadiene is converted to acetoxy 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphate(Scheme 4 c) and then undergoes acetate hydrolysis to form 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadien-1-ol (Scheme 4, 5). Most alarmingly, the 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadien-1-ol product yield was only 33% after acetate hydrolysis (Scheme 4, d), which is also uncharacteristically low. The removal of the acetate group should have a nearly 100% yield but is instead drastically lower. The conversion of the 8-N-m- benzyl alcohol-amino-3,7-dimethyl-1-acetoxy-2,6 octadiene to the azide (Scheme 4, c, Figure 14, left), yields the formation of a second spot with an R_f value very close to the desired product (Figure 14, middle). Instead of trying to separate the two spots by column chromatography, the two spots are purified together. During the acetate hydrolysis, the desired product's R_f value shifts due to removal of the acetate group, which makes the next purification much easier (Figure 14, right).

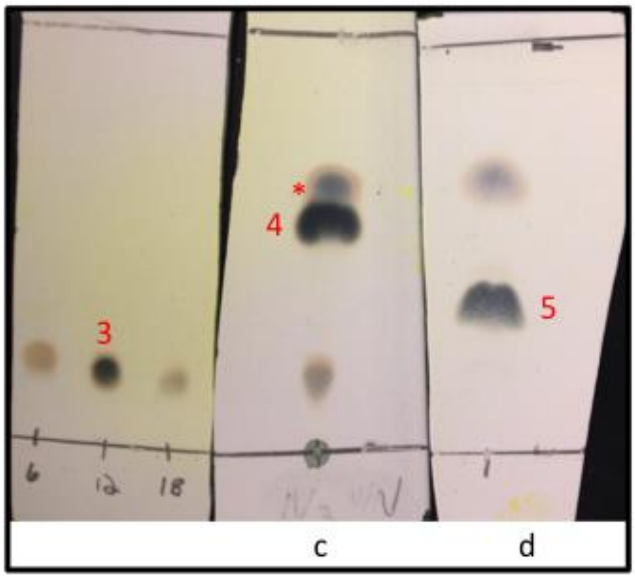
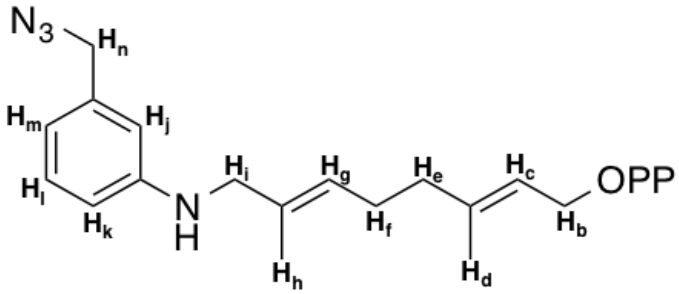


Figure 14: TLC of Azide Conversion and Acetate Hydrolysis. Solvent was 30% EtOAc/Hexanes. Numbers and letters designate compounds and reactions as labeled in Scheme 4. The acetyoxygeranyl benzyl alcohol (3, left plate) is converted to an azide (4, center plate) before hydrolyzing the acetate group to form the benzyl azido geranyl alcohol (5, right plate).

While no specific reason is provided, Lambadie *et al.* perform acetate hydrolysis on acetoxy 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphatewithout any purification whatsoever.[(30)] Therefore, it is not known whether this group observed a similar side product formation by TLC and utilized the subsequent reaction to make purification of the 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadien-1-ol easier. Increasing the yields is not the most significant concern considering that this synthesis has already been reported, strengthening the process is important for efficiency especially for future projects that will utilize this compound. Regardless of technique used, the reported proton chemical shifts are close to what is represented in the literature (Table 2, Appendix Figure 2).

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

8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphate was fully synthesized and isolated by HPLC. ^{31}P-NMR, ^{13}C-NMR and ^{1}H-NMR were performed to characterize the structure of the analogue (Appendix, Figures 3-5). The chemical shifts observed by NMR are very similar all three types of NMR (Table 3). The ^{13}C-NMR had three extra peaks at 215.1, 29.9 and -14.2 ppm. When acetone is introduced to D_2O, it displays shifts at 30.89 and 215.54.[32] The peak at -14.2 ppm is likely background. The ^{31}P-NMR spectra showed that there were two *distinct* phosphorus atoms present in the structure, which was expected considering the structure. Mass spectrometry analysis of the compound yields a mass of 459.17 g/mol (Appendix, Figure 6), which is nearly identical to the expected mass (459.12 g/mol) of the compound, indicating successful synthesis of the compound. Further, the mass spectrometry data suggests that the azide moiety is still intact on the molecule, something that cannot be detected through the other characterization techniques performed.



Scheme 6: Labeled protons for ^{1}H-NMR for 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphatediphosphate

Table 3: Various NMR for 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphate.

NMR-Type	Chemical Shifts Reported in Literature[30]	^{1}Chemical Shifts Observed
^{1}H-NMR (Scheme 6, top)	(D_2O, δ): 7.14 (m, 1H), 6.64 (m, 3H), 5.32 (q, 2H), 4.17 (s, 2H), 3.56 (s, 2H), 2.07 (m, 2H), 1.98 (m, 2H), 1.61 (s, 3H), 1.54 (s, 3H)	(D_2O, δ): 7.13 (t, 1H, H_l), 6.67 (d, 3H, $H_{j, k, m}$), 5.25 (q, 2H, $H_{c, g}$), 4.30 (t, 2H H_b), 4.17 (s, 2H H_n), 3.54 (s, 2H, H_i), 2.07-1.94 (m, 4H, $H_{f, e}$), 1.52 (s, 3H, H_d), 1.47 (s, 3H, H_h)
^{13}C-NMR (Scheme 6, bottom)	(D_2O, δ): 149.4, 143.1, 137.2, 133.2, 130.4, 127.0, 120.6, 118.8, 115.2, 114.7, 63.2, 55.2, 51.7, 39.4, 26.2, 16.4, 14.5	(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-NMR	(D_2O, δ): -7.30 (1P), -10.5 (1P)	(D_2O, δ): -8.91 (1P), -10.14 (1P)

3.10 Azido-FPP Analogue is a Substrate for UPPS

As stated in the previous section, the synthesis of the 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphate was not novel. However, the utilization of this compound as a substrate for UPPS has not been reported. The analogue is in fact a substrate for UPPS (Figure 16). Interestingly, there is variation in product formation

between different species of UPPS. The enzyme from *B. fragilis* makes multiple Z-addition products while the enzyme from *E. coli* makes one Z-addition product. Further, the sole product released by UPPS from *E. coli* seems to be one isoprene longer than the products released by that of *V. vulnificus.* This trend is similar to data with other FPP analogues published by the Troutman lab.[(24)] Further, the enzyme from *V. vulnificus* does not seem to utilize this analogue as a substrate, as indicated by the large amount of diphosphate still present. Because the addition of enzyme was consistent by volume and not concentration, differing levels of enzyme could affect the incorporation of the analogue. However, the purpose for this assay was to rudimentarily determine whether the analogue could be utilized by UPPS as a substrate, not by any means quantify its activity.

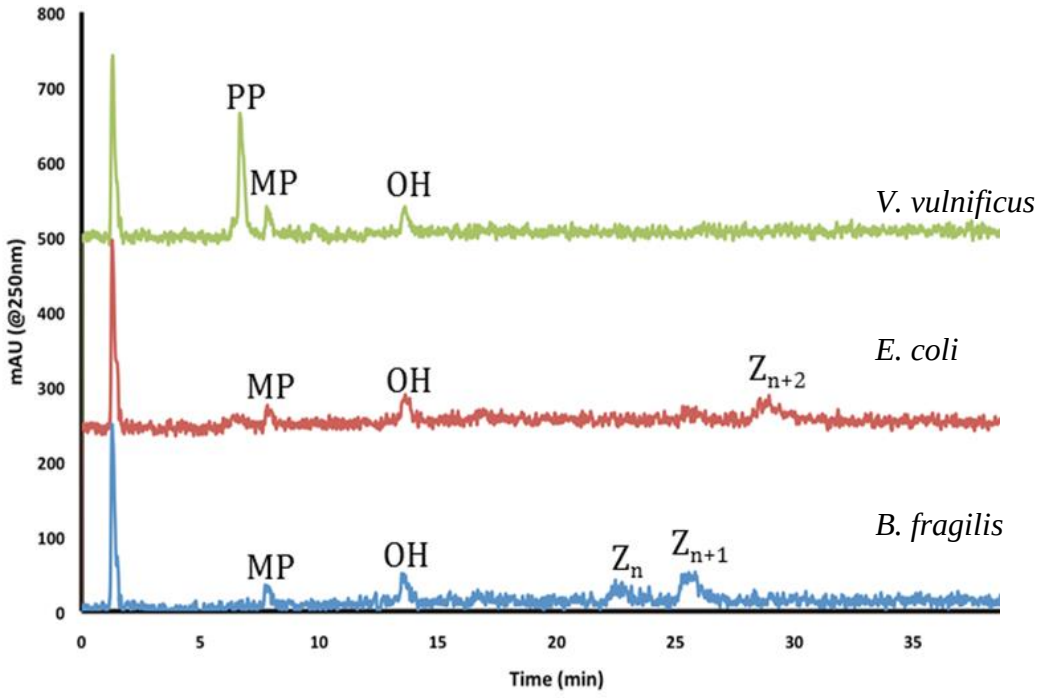


Figure 16: 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphate is a substrate for UPPS. The potential activity of UPPS with the 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphate was tested with UPPS from three species: *B. fragilis* (blue), *E. coli* (red), *V. vulnificus* (green). PP = diphosphate, MP = monophosphate, OH = alcohol, Z= *cis*-addition products.

CHAPTER FOUR: CONCLUSIONS & FUTURE DIRECTIONS

4.1: Characterization of the Colanic Acid Biosynthesis Pathway

The work described here details the characterization of three glycosyltransferases (WcaI, WcaE, WcaC) and one acetyltransferase (WcaF) of the colanic acid biosynthesis pathway (Figure 17). Each of these characterizations has not been previously reported and is significant progress with regards to understanding how *E. coli* synthesize colanic acid. To accomplish this, each individual gene product in the locus was amplified by PCR and inserted into a vector through LIC. Next, each vector was transformed into C41 cell lines for protein overexpression. Each protein was isolated and purified by Ni-NTA affinity chromatography and dialyzed. Finally, every characterization assay utilized the fluorescent 2-CN-5Z-BP analogue and each reaction mixture was analyzed by reverse-phase HPLC. Not only did this process allow for the characterization of three glycosyltransferases responsible for the biosynthesis of the tetrasaccharide component of the CA repeat-unit, the data displayed here suggests that WcaF may be a valuable antibiotic target.

The characterization of WcaI as the first fucosyltransferase within the pathway was the first significant discovery with regards to this project. Prior to this, it was not known which glycosyltransferase was responsible for this transfer with any of the five (WcaA, WcaC, WcaE, WcaI and WcaL) having the potential to do so. However the data in Figure 11 clearly shows that WcaI transfers fucose to the developing repeat unit.

Likewise, WcaE and WcaC were shown to exhibit fucosyltransferase and galactosyltransferase activity, respectively (Figures 12 and 15). At this point, with the tetrasaccharide product synthesized, there are only two glycosyltransferases remaining to be characterized (WcaA and WcaL). Although the repeat-unit is not complete, the progress made greatly narrows the focus for the characterization of the remainder of the pathway because of the decrease in glycosyltransferases left to catalyze the transfer of glucuronic acid to form the pentasaccharide product.

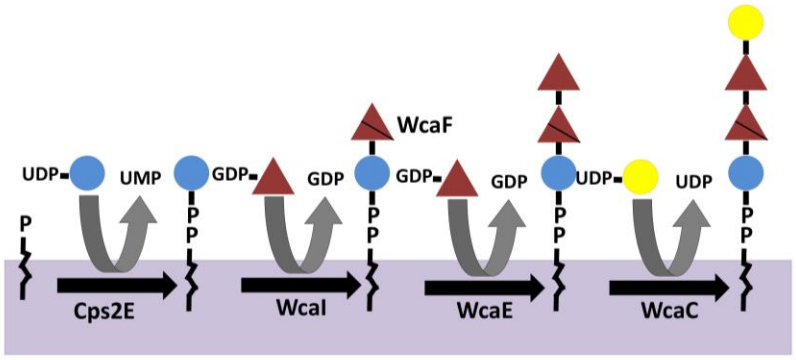


Figure 17: Summary of CA biosynthesis characterization. Cps2E was the initiating glucose-1-phosphate transferase that adds glucose-1-P to the fluorescent BP anchor (2-CN-5Z-BPP). WcaI is the first fucosyltransferase in the pathway. WcaF acetylates the WcaI product (2-CN-5Z-BPP-Glc-Fuc) prior to addition of another fucose by WcaE. Finally, WcaC transfers galactose to form a tetrasaccharide product (2-CN-5Z-BPP-Glc-AcFuc-Fuc-Gal). Blue circles represent glucose, red triangles represent fucose, and yellow circles represent galactose. Acetylaction by WcaF is represented by a black line in the triangle.

Perhaps the most interesting element of this work is that the CA repeat-unit was not synthesized as hypothesized with regards to acetylation. When considering the potential models for repeat-unit acetylation (Figure 5), it was originally believed that both acetylations would occur together after the formation of the disaccharide product (2-CN-5Z-BPP-Glc-Fuc). However, this was not the case. Instead, only WcaF was required for the continuation of CA biosynthesis. Prior to these experiments, the successful transfer of

fucose to BPP-Glc by WcaI eliminated the possibility that free GDP-fucose was acetylated before being incorporation into the repeat unit (Figure 5a). The data presented in Figure 14 suggests that in the presence of acetyl-CoA and WcaF, the fucose residue is acetylated and is required for WcaE to transfer the next fucose to the developing repeat unit. This would suggest that the acetylation model presented in Figure 5c reflects what is being observed *in vitro*. However, this observation based on the HPLC chromatogram does not confirm that acetylation is actually happening—only that WcaF and acetyl-coA are required for WcaE to function. Also, it would be wrong to assume that only WcaF is the only acetyltransferase functioning at this point in the synthesis is that it blindly ignores the possibility that WcaB also acetylates this fucose residue, but its activity is not required for the continuation of the synthesis. It is entirely possible that WcaE only recognizes the position on the fucose residue that WcaF acetylates, but not WcaB. If this is true, then it would explain why there seems to no activity when the disaccharide is only treated with WcaB, but displays turnover when treated solely with WcaF. Mass spectrometry analysis of the disaccharide treated with each acetyltransferase individually and the two together needs to be done to verify that acetylation is actually occurring at this step and determine which acetyltransferase(s) are required to progress the synthesis by WcaE. Because mass spectrometry has not been done, either remaining acetylation model presented in Figure 5 is entirely possible.

The first future goal regarding this project is finishing the characterization of the remaining glycosyltransferases in the colanic acid biosynthesis pathway. Hypothesized characterizations for the remaining two glycosyltransferases (WcaA, WcaL) can be made based upon bioinformatics and sequence similarity to homologous proteins. For example,

WcaL has a 52% amino acid homology to AmsK from *Erwinia amylovora*, a protein that has been found to be a galactosyltransferase.[10, 33] Therefore, WcaL may be the terminal galactosyltransferase, making WcaA the glucuronic acid transferase enzyme. To test this, isolated WcaC product (BPP-Glc-2,3-AcFuc-Fuc-Gal) will be treated with glucuronic acid and WcaA or WcaL to determine which generates the pentasaccharide product.

Beyond characterizing the remaining glycosyltransferases in the colanic acid biosynthesis pathway, there is nothing reported concerning the substrate specificities for glycosyltransferase enzymes characterized within this project (WcaI, WcaE, WcaC). For example, the effect that isoprenoid length has on enzymatic activity of these enzymes is not known. Recently, the Troutman lab published a study that showed the effects that surfactants have on product release by UPPS resulting in varied length fluorescent BPP analogues.[34] These varied length substrates will be utilized to determine the relationship between isoprenoid length and the activity of each protein. For example, if the active site of a glycosyltransferase binds to four isoprenoid units and the substrate provided only has three, then it is unlikely that the glycosyltransferase will attach a sugar. Now considering that *in vivo*, the isoprenes are completely embedded in the membrane and glycosyltransferses are localized in the cytoplasm, the variable length substrates may not have too much of an effect on activity. The specificity of WcaI for BPP-Glc can also be probed by using WbaP in place of Cps2E. As mentioned, WbaP is an initiating galactose-1-phosphate transferase. The product of WbaP, BPP-Gal, would probe the stereochemical requirements for WcaI at the 4’ of the sugar, considering that glucose and galactose are epimers at that 4’ carbon position.

A significant focus of this work was to identify potential antibiotic targets to inhibit CA production. It was suggested that WcaF would be an appropriate target considering that it seems that this acetyltransferase is required for the continued synthesis of the repeat unit (Figure 14). The most practical approach to target this acetyltransferase is to use acetyltransferase inhibitors to prevent enzyme function. To do this, knowing the specific active site residues are required for substrate binding is required. Utilizing pBLAST homology sequnce alignment bioinformatic techniques, it was determined that putative active site residues for WcaF are W79, G81, H111, H113, W132, A134, T135, A140, G152, A153, F158, K159, R168, G169 and P171.[(35)] Interestingly, both the "chemical substrate" binding site and the acetyl-CoA binding site share the 111H residue. Unfortunately, other acetyltransferases critical to human biochemical metabolism such as citrate synthase and carnitine acetyltransferase also require the presence of a catalytic histidine for function. [(36, 37)] Therefore, an inhibitor for WcaF needs to be specific enough that it does not bind to secondary targets and cause detriment to a human host.

One potential method to accomplish WcaF inhibition other than inhibitor use would be to hijack the *E. coli* immune response to target and degrade the mRNA coding for these acetyltransferases. Because *E. coli* are consistently in DNA rich environments due to transformation of DNA between organisms or interactions with bacteriophages, each cell utilizes a CAS/CRISPR system for protection against DNA that will ultimately cause harm.[(38)] This works by a CAS protein complex binding to foreign DNA and transcribing short repeat sequences (clustered regularly interspaced palindromic repeats, CRISPR) that incorporate into a region of the bacterial genome with other short recognition sequences. Once the host is infected again with the same foreign DNA, the

short repeat sequence within the genome is transcribed into RNA that is used by the CAS protein system to recognize and bind intruding DNA that ultimately causes degradation of the genetic material.

Recently, a group published a study utilizing this immune response to target antibiotic resistance genes within certain strains of multi-drug resistant bacteria.[39] This group used bacteriophages to infect bacteria with the genes encoding a CAS system and CRISPR sequences targeting the antibiotic resistance for streptomycin and gentamycin. Specifically, the designed CRISPR sequences targeted native plasmids containing these antibiotic resistance genes and ultimately reversed drug-resistance as the plasmids were recognized and destroyed by the CAS protein complex. A similar method could be used to target the acetyltransferase activity required for colanic acid production. By incorporating CRISPR sequences inhibiting the production of *wcaB* and *wcaF* mRNA, colanic acid biosynthesis would cease after the formation of the disaccharide product by WcaI. One limitation with this proposal is that it targets mRNA production rather than plasmid DNA as reported in the literature. However, recently a group was able to target mRNA with a CAS/CRISPR system by employing small DNA sequences that bind to mRNA.[40] If this system could prevent translation of CA acetyltransferase mRNA, then evaluating each gene for antibiotic targeting can be done.

4.2: Development of the Azido-FPP Analogue

The synthesis of the azido-FPP analogue was not novel as it was previously published by Labadie *et al*.[30] However, this is the first study utilizing this multi-functional probe as a substrate for UPPS. As this was a secondary project, there is not

much to conclude based upon the progress made thus far. However, there is significant progress regarding what can be done with the elongated species in the future.

Now that the azido-FPP analogue has been shown to be a substrate for UPPS, the next step is enhancing the applicability of the compound through Huisgen chemistry. To test the conditions of 1,3-dipolarcycloaddition reaction, the first alkyne that should be tested with the FPP analogue is a fluorophore. Attaching a fluorophore that is more easily detected than the compounds readily available in the Troutman laboratory will allow for smaller-scale BP reactions and save material over time. Further, detecting the attached fluorophore prior to cycloadditions with other terminal alkynes is important to determine whether the conditions of the reaction are working.

Until this point, the only function of the BPP analogues has been fluorescence detection. However, this azido-FPP probe can be multi-functional, with the potential only limited by what terminal alkynes are attached. One potential application would be to use this compound to identify unknown protein targets in cell lysates. Biotin is a biomolecule that has a strong binding interaction with the protein streptavidin. Utilizing streptavidin coated beads, a resin could be generated with the azido-BP analogue clicked to a biotin molecule. Pouring cell lysates through a column with this resin will allow the binding of any protein that recognizes that particular substrate. To test whether this works, a resin should be generated with the Cps2E product (BPP-Glc) protruding from the beads. Then, lysates of C41 cells expressing WcaI should be poured through the column. In doing so, WcaI should bind to the resin. WcaI can be eluted from the resin by rinsing the column volume with free biotin. Any isolated protein can be reverse sequenced to determine the success of the experiment.

REFERENCES

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

2. Sandvig, K. (2001) Shiga toxins, *Toxicon 39*, 1629-1635.

3. Pacheco, A. R., and Sperandio, V. (2012) Shiga toxin in enterohemorrhagic E.coli: regulation and novel anti-virulence strategies, *Frontiers in cellular and infection microbiology 2*, 81.

4. Raum, E., Lietzau, S., von Baum, H., Marre, R., and Brenner, H. (2008) Changes in Escherichia coli resistance patterns during and after antibiotic therapy: a longitudinal study among outpatients in Germany, *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases 14*, 41-48.

5. Whitfield, C. (2006) Biosynthesis and assembly of capsular polysaccharides in Escherichia coli, *Annual review of biochemistry 75*, 39-68.

6. Sutherland, I. W. (1969) Structural studies on colanic acid, the common exopolysaccharide found in the enterobacteriaceae, by partial acid hydrolysis. Oligosaccharides from colanic acid, *The Biochemical journal 115*, 935-945.

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

8. Mao, Y., Doyle, M. P., and Chen, J. (2006) Role of colanic acid exopolysaccharide in the survival of enterohaemorrhagic Escherichia coli O157:H7 in simulated gastrointestinal fluids, *Letters in applied microbiology 42*, 642-647.

9. Patel, K. B., Toh, E., Fernandez, X. B., Hanuszkiewicz, A., Hardy, G. G., Brun, Y. V., Bernards, M. A., and Valvano, M. A. (2012) Functional characterization of UDP-glucose:undecaprenyl-phosphate glucose-1-phosphate transferases of Escherichia coli and Caulobacter crescentus, *Journal of bacteriology 194*, 2646-2657.

10. Stevenson, G., Andrianopoulos, K., Hobbs, M., and Reeves, P. R. (1996) Organization of the Escherichia coli K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid, *Journal of bacteriology 178*, 4885-4893.

11. Somoza, J. R., Menon, S., Schmidt, H., Joseph-McCarthy, D., Dessen, A., Stahl, M. L., Somers, W. S., and Sullivan, F. X. (2000) Structural and kinetic analysis of Escherichia coli GDP-mannose 4,6 dehydratase provides insights into the enzyme's catalytic mechanism and regulation by GDP-fucose, *Structure 8*, 123-135.

12. Andrianopoulos, K., Wang, L., and Reeves, P. R. (1998) Identification of the fucose synthetase gene in the colanic acid gene cluster of Escherichia coli K-12, *Journal of bacteriology 180*, 998-1001.

13. Sonnhammer, E. L., von Heijne, G., and Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences, *Proceedings / ... International Conference on Intelligent Systems for Molecular Biology ; ISMB. International Conference on Intelligent Systems for Molecular Biology 6*, 175-182.

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

15. Wang, L., Liu, D., and Reeves, P. R. (1996) C-terminal half of Salmonella enterica WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis, *Journal of bacteriology 178*, 2598-2604.

16. Cartee, R. T., Forsee, W. T., Bender, M. H., Ambrose, K. D., and Yother, J. (2005) 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, *Journal of bacteriology 187*, 7425-7433.

17. Somers, W. S., Stahl, M. L., and Sullivan, F. X. (1998) GDP-fucose synthetase from Escherichia coli: structure of a unique member of the short-chain dehydrogenase/reductase family that catalyzes two distinct reactions at the same active site, *Structure 6*, 1601-1612.

18. McGowan, C. C., Necheva, A., Thompson, S. A., Cover, T. L., and Blaser, M. J. (1998) Acid-induced expression of an LPS-associated gene in Helicobacter pylori, *Molecular microbiology 30*, 19-31.

19. Edwards, N. J., Monteiro, M. A., Faller, G., Walsh, E. J., Moran, A. P., Roberts, I. S., and High, N. J. (2000) Lewis X structures in the O antigen side-chain promote adhesion of Helicobacter pylori to the gastric epithelium, *Molecular microbiology 35*, 1530-1539.

20. Comstock, L. E., Coyne, M. J., Tzianabos, A. O., Pantosti, A., Onderdonk, A. B., and Kasper, D. L. (1999) Analysis of a capsular polysaccharide biosynthesis locus of Bacteroides fragilis, *Infection and immunity 67*, 3525-3532.

21. Hollan, S. (2000) Molecular structure and function of erythrocyte blood group antigens, *Haematologia 30*, 237-252.

22. Byun, S. G., Kim, M. D., Lee, W. H., Lee, K. J., Han, N. S., and Seo, J. H. (2007) Production of GDP-L-fucose, L-fucose donor for fucosyloligosaccharide synthesis, in recombinant Escherichia coli, *Applied microbiology and biotechnology 74*, 768-775.

23. Lujan, D. K., Stanziale, J. A., Mostafavi, A. Z., Sharma, S., and Troutman, J. M. (2012) Chemoenzymatic synthesis of an isoprenoid phosphate tool for the analysis of complex bacterial oligosaccharide biosynthesis, *Carbohydrate research 359*, 44-53.

24. Mostafavi, A. Z., Lujan, D. K., Erickson, K. M., Martinez, C. D., and Troutman, J. M. (2013) Fluorescent probes for investigation of isoprenoid configuration and size discrimination by bactoprenol-utilizing enzymes, *Bioorganic & medicinal chemistry 21*, 5428-5435.

25. Chehade, K. A., Andres, D. A., Morimoto, H., and Spielmann, H. P. (2000) Design and synthesis of a transferable farnesyl pyrophosphate analogue to Ras by protein farnesyltransferase, *The Journal of organic chemistry 65*, 3027-3033.

26. Novelli, G., and D'Apice, M. R. (2012) Protein farnesylation and disease, *Journal of inherited metabolic disease 35*, 917-926.

27. Chang, S. Y., Ko, T. P., Chen, A. P., Wang, A. H., and Liang, P. H. (2004) Substrate binding mode and reaction mechanism of undecaprenyl pyrophosphate synthase deduced from crystallographic studies, *Protein science : a publication of the Protein Society 13*, 971-978.

28. Mostafavi, A. Z., and Troutman, J. M. (2013) Biosynthetic assembly of the Bacteroides fragilis capsular polysaccharide A precursor bactoprenyl diphosphate-linked acetamido-4-amino-6-deoxygalactopyranose, *Biochemistry 52*, 1939-1949.

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

30. Labadie, G. R., Viswanathan, R., and Poulter, C. D. (2007) Farnesyl diphosphate analogues with omega-bioorthogonal azide and alkyne functional groups for protein farnesyl transferase-catalyzed ligation reactions, *The Journal of organic chemistry 72*, 9291-9297.

31. Zhu, W., Zhang, Y., Sinko, W., Hensler, M. E., Olson, J., Molohon, K. J., Lindert, S., Cao, R., Li, K., Wang, K., Wang, Y., Liu, Y. L., Sankovsky, A., de Oliveira, C. A., Mitchell, D. A., Nizet, V., McCammon, J. A., and Oldfield, E. (2013) Antibacterial drug leads targeting isoprenoid biosynthesis, *Proceedings of the National Academy of Sciences of the United States of America 110*, 123-128.

32. Fulmer, G. R., Miller, A. J. M., Sherden, N. H., Gottlieb, H. E., Nudelman, A., Stoltz, B. M., Bercaw, J. E., and Goldberg, K. I. (2010) NMR Chemical Shifts of Trace Impurities: Common Laboratory Solvents, Organics, and Gases in Deuterated Solvents Relevant to the Organometallic Chemist, *Organometallics 29*, 2176-2179.

33. Langlotz, C., Schollmeyer, M., Coplin, D. L., Nimtz, M., and Geider, K. (2011) Biosynthesis of the repeating units of the exopolysaccharides amylovoran from Erwinia amylovora and stewartan from Pantoea stewartii, *Physiol Mol Plant P 75*, 163-169.

34. Troutman, J. M., Erickson, K. M., Scott, P. M., Hazel, J. M., Martinez, C. D., and Dodbele, S. (2015) Tuning the production of variable length, fluorescent polyisoprenoids using surfactant-controlled enzymatic synthesis, *Biochemistry 54*, 2817-2827.

35. Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S. N., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M., Hurwitz, D. I., Lanczycki, C. J., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z. X., Yamashita, R. A., Zhang, D. C., Zheng, C. J., and Bryant, S. H. (2015) Cdd: Ncbi's Conserved Domain Database, *Nucleic Acids Res 43*, D222-D226.

36. Voet, D., Voet, J.G. (2011) Enzymes of the citric acid cycle, In *Biochemistry* (Recta, P., Ed.) 4th ed., pp 806-807, John Wiley & Sons, Inc., Hoboken, NJ.

37. Jogl, G., and Tong, L. (2003) Crystal structure of carnitine acetyltransferase and implications for the catalytic mechanism and fatty acid transport, *Cell 112*, 113-122.

38. Horvath, P., and Barrangou, R. (2010) CRISPR/Cas, the Immune System of Bacteria and Archaea, *Science 327*, 167-170.

39. Yosef, I., Manor, M., Kiro, R., and Qimron, U. (2015) Temperate and lytic bacteriophages programmed to sensitize and kill antibiotic-resistant bacteria, *Proceedings of the National Academy of Sciences of the United States of America 112*, 7267-7272.

40. O'Connell, M. R., Oakes, B. L., Sternberg, S. H., East-Seletsky, A., Kaplan, M., and Doudna, J. A. (2014) Programmable RNA recognition and cleavage by CRISPR/Cas9, *Nature 516*, 267.

APPENDIX: FIGURES

Figure 1: ^{1}H-NMR of 8-N-m- benzyl alcohol-amino-3,7-dimethyl-1-acetoxy-2,6 octadiene

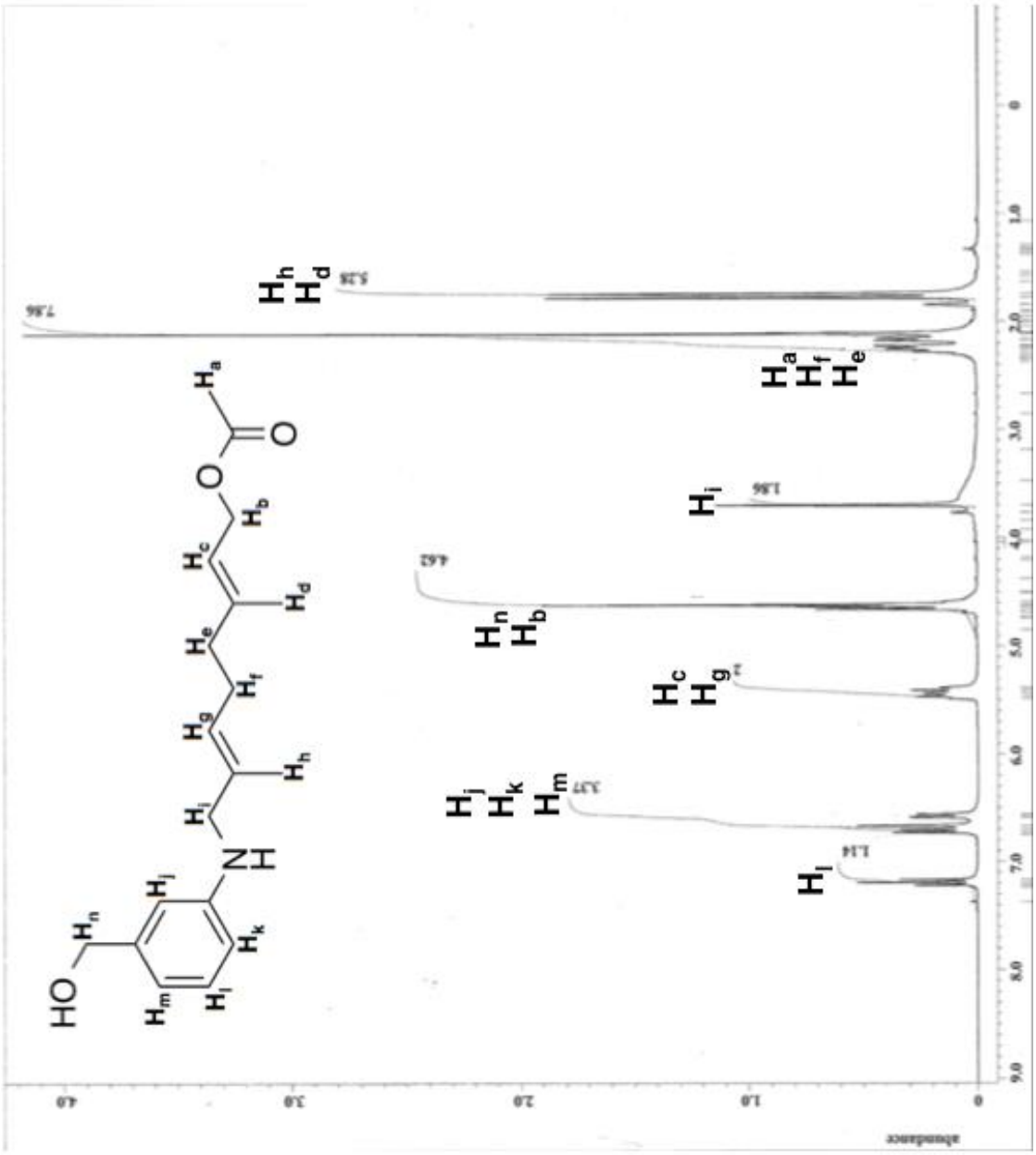


Figure 2: [1]H-NMR of 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadien-1-ol

Figure 3: ^{1}H-NMR of 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphate

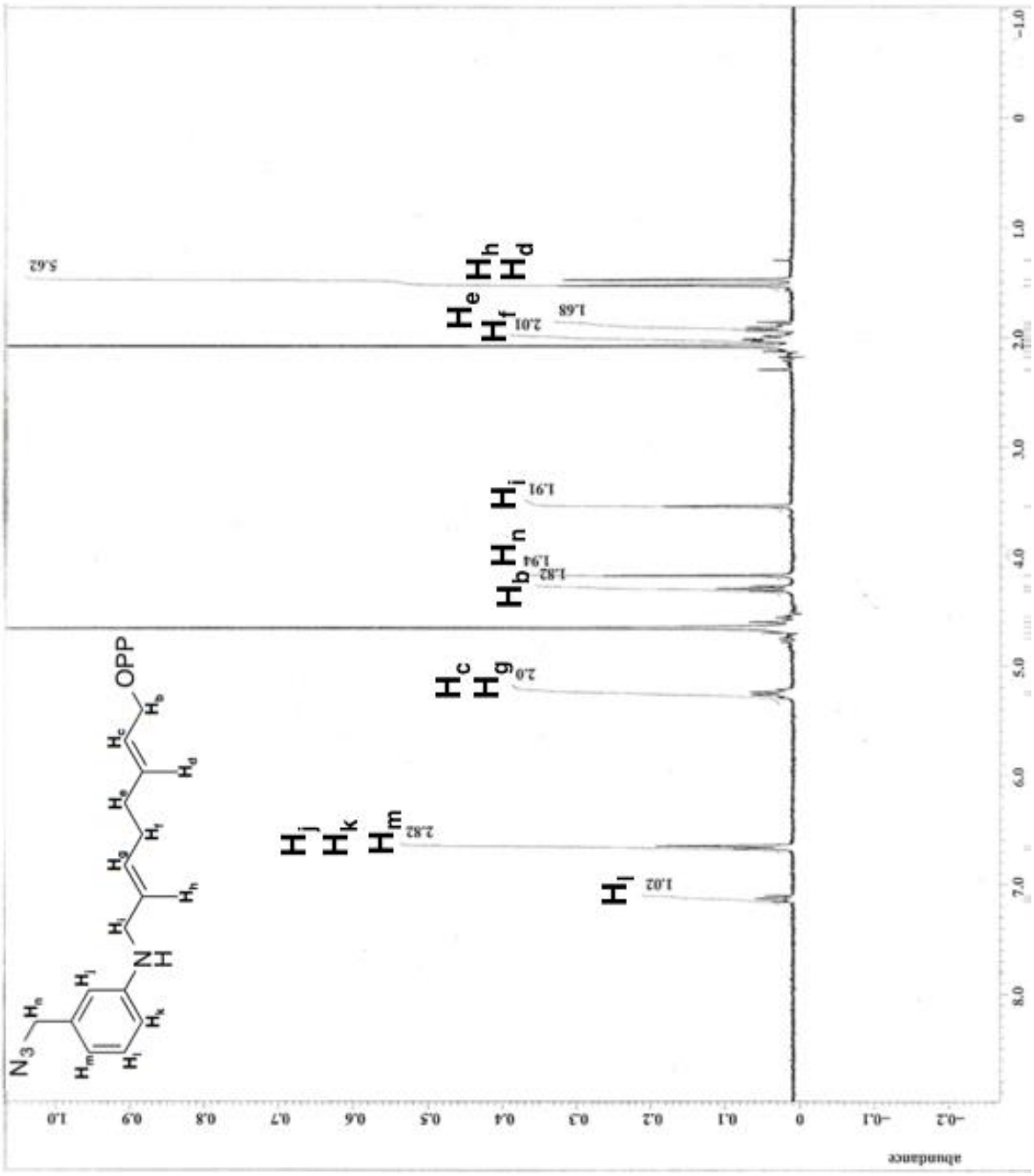


Figure 4: ^{13}C-NMR of 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphate

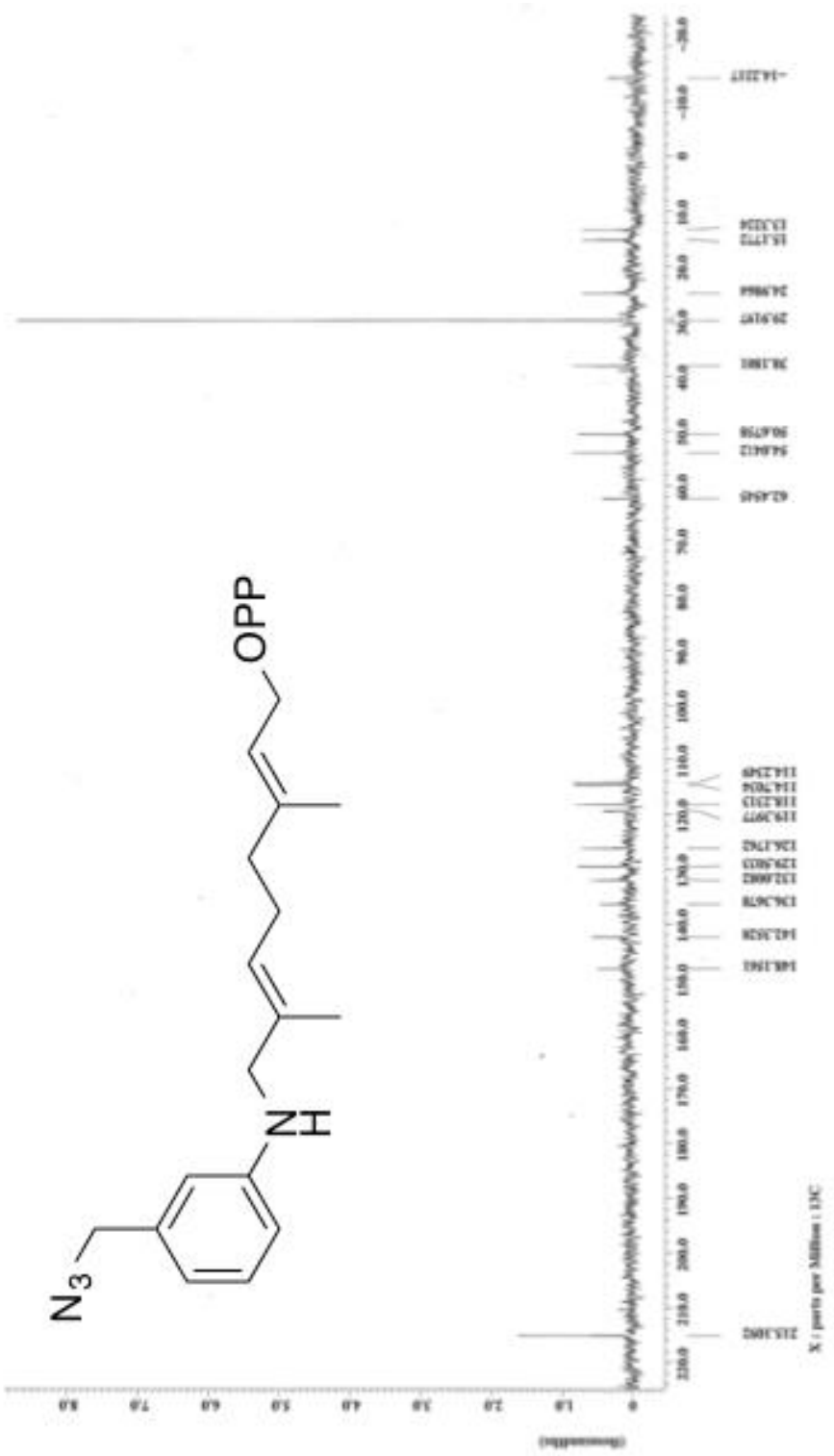


Figure 5: ^{31}P-NMR of 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphate

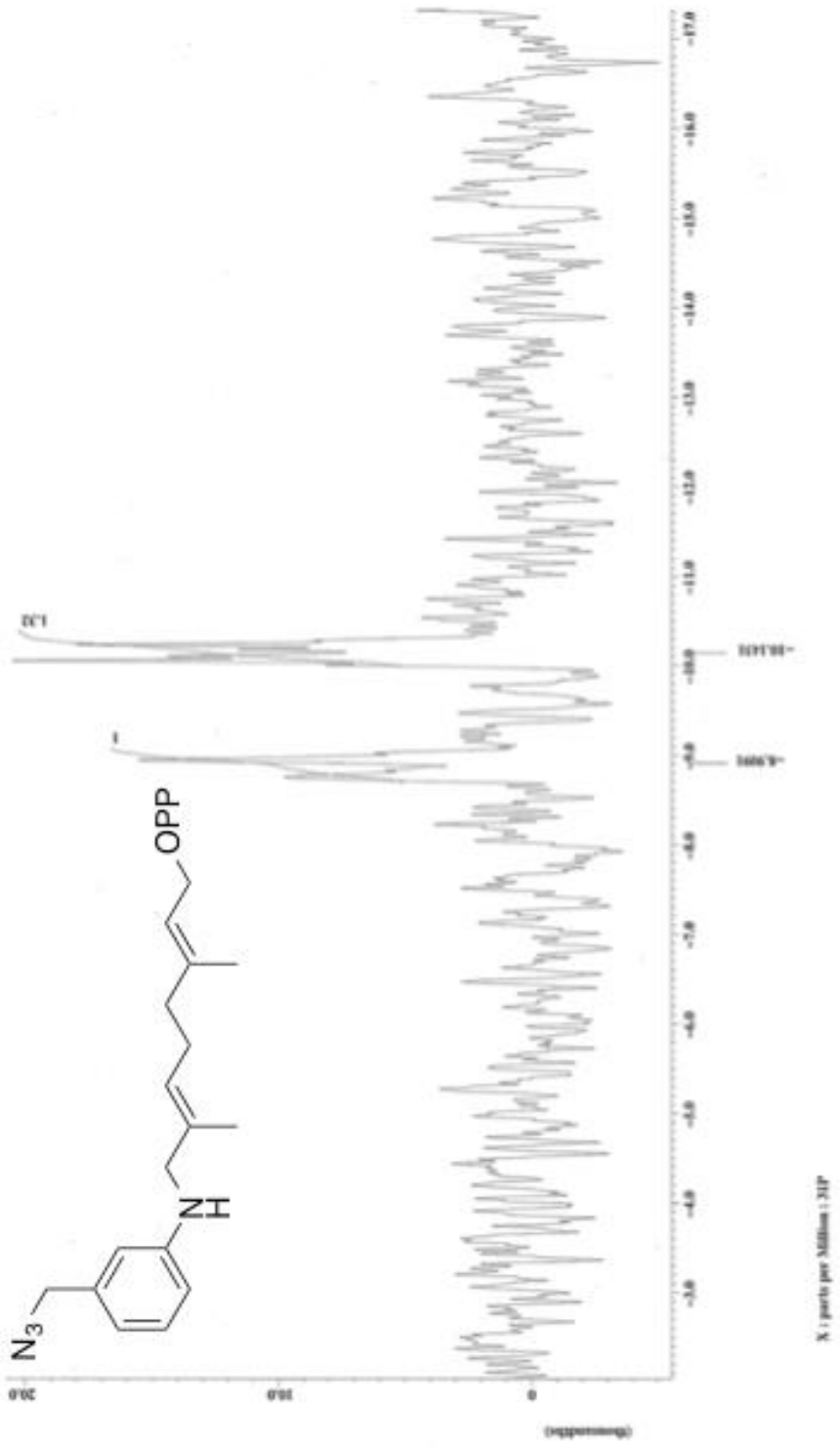


Figure 6: Mass Spectrometry Analysis of 8-N-m- benzyl alcohol-amino-3,7-dimethyl-2,6 octadiene diphosphate.