CHEMOENZYMATIC SYNTHESIS OF MULTIMODAL GLYCOLIGANDS

WITH BIO-ORTHOGONAL ALDEHYDE BASED FUNCTIONAL MOIETY

By

Antonio Goncalves

A thesis submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Master of Science

Graduate Program in Chemical and Biochemical Engineering

Written under the direction of

Dr. Shishir P. S. Chundawat

And approved

_____

_____

_____

New Brunswick, New Jersey

[October, 2019]

ABSTRACT OF THE THESIS

**CHEMOENZYMATIC SYNTHESIS OF MULTIMODAL GLYCOLIGANDS
WITH BIO-ORTHOGONAL ALDEHYDE BASED FUNCTIONAL MOIETY**

By Antonio Goncalves

Thesis Director:

Dr. Shishir P. S. Chundawat

Multimodal ligands can be used to achieve selective clearance of impurities in a single chromatography step by having multiple modes of interaction between the ligand and the targeted protein analyte. One strategy for designing effective multimodal ligands is to establish a large library of chemically diverse ligands and utilize high-throughput screening methods to test each ligands effectiveness. However, this method becomes ineffective as the number of ligands in the library increases, so other strategies should also be utilized. A strategy used to create a smaller but, more effective library, is to design ligands that closely mimic protein-ligand binding interactions found in natural systems. One such example would be design ligands that mimic the multimodal interactions between proteins and complex glycans. However, limited research has been conducted to design and synthesize multimodal glycan-based ligands (or glycoligands) for protein chromatography or other similar applications.

Here, we investigate the chemoenzymatic synthesis of glycan-based multimodal glycoligands using a glycosynthase enzyme engineered from a native glucuronidase

belonging to the family 2 glycosyl hydrolase. To the best of our knowledge, this work signifies the first reported attempt to chemoenzymatically synthesize a glucuronide with a bio-orthogonal aldehyde functional group allowing it to be covalently attached to a suitable resin or support. The four specific objectives that were completed to accomplish this goal were; (i) use *in-silico* docking simulations to establish and justify a library of hydrophobic alcohol acceptors containing an aldehyde functional group, (ii) synthesize and purify 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid using TEMPO oxidation and anion exchange chromatography, (iii) generate a library of glycosynthase mutants from the *E.coli* glucuronidase gene, *uidA*, and express, purify and characterize the wild-type and nucleophilic mutant enzymes, (iv) run glycoligand synthesis assays with an active glycosynthase mutant, activated glucopyranuronic acid donor and a library of alcohol acceptors to confirm if the engineered uidA glycosynthases can synthesize the desired multimodal glycoligands. Our results led further credence to our initial hypothesis and *in-silico* studies, that uidA glycosynthases can synthesize glucuronides with aldehyde containing bio-orthogonal functional groups that can then serve as potential multimodal ligands for protein chromatography. We also discuss future research directions into how we can utilize additional glycosynthases enzymes to create more complex multimodal charged glycans by using our currently synthesized glycoligand product as substrate.

Acknowledgements

I would like to thank my thesis advisor, Dr. Shishir Chundawat, for allowing me to conduct research in his lab for the past 3 years and for his guidance with this project. I would also like to thank Dr. Tsilomelekis and Dr. Schuster for serving on my committee. I would like to thank all other faculty who have influenced me in the Department of Chemical and Biochemical Engineering at Rutgers University. I am also very grateful for having great laboratory colleagues; without them my experiences working in lab wouldn't have been as enjoyable. I am especially thankful to Chandra Kanth Bandi for introducing me to biological research, for teaching me everything and for always being patient and helpful. I would also like to acknowledge Dr. Zoltan Szekely for allowing me to use his lyophilizer. Finally, I would like to thank the National Science Foundation (NSF Award 1704679) for supporting this research project.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1 Background

The main sources of impurities in the production of biotherapeutics come from the host cell itself, the growth media, and the denaturation and alteration of the protein product. To remove these impurities, biopharmaceutical companies utilize a combination of different chromatography more or techniques including affinity, hydrophobic interaction, ion exchange, and size exclusion chromatography.[1] Each of these chromatography techniques is based on a single mode of interaction between the resin ligands and the protein analyte, with each mode having certain limitations. For example, ion exchange chromatography cannot be used with a high conductivity feed since an increasing salt concentration is used to exchange sites with the protein bound onto the ligands during elution. Affinity chromatography cannot be used on certain proteins where the affinity tag affects the folding and *in vivo* efficacy of the protein drug etc. Hydrophobic interaction chromatography can cause strong protein adsorption and unfolding when the hydrophobic interactions are too strong.[2] To overcome these setbacks and to decrease purification costs from running multiple modes of chromatography in a sequential manner, researchers are looking to develop new purification strategies which includes the development of new multimodal ligands for protein bioseparations.

Multimodal chromatography works by using multiple modes of interaction between the ligand attached to the support resin and the binding protein analyte to achieve selective clearance of impurities in a single chromatography step.[2] The most common type multimodal purification is hydrophobic charge induction chromatography (HCIC). This

purification strategy works by having the protein selectively interact with the ligand through both hydrophobic and electrostatic interactions to separate it from the host cell proteins and then having the bound targeted protein eluted by raising the pH so that electrostatic repulsion exceeds the hydrophobic interaction.[3] The electrostatic repulsion in tandem with hydrophobic interactions can prevent the protein denaturation found in traditional single-mode hydrophobic interaction chromatography.[4] Also, the hydrophobic interactions on the HCIC ligands allows for proteins to bind onto the ligand at higher salt concentration levels than traditional single-mode ion exchange chromatography.[5]

HCIC ligands are designed to have at least one aromatic group to contribute to it hydrophobicity and at least one ionic group like a carboxyl or amino group. HCIC ligands also have other groups to provide secondary interactions. Hydrogen bond donating groups on the ligand also helps the protein bind in the presence of high salt concentrations.[6] However, with so many combinations of hydrophobic and ionic groups, design strategies must be used to efficiently create HCIC ligands. One common strategy is to design ligands so that its interactions with the proteins mimics interactions commonly found in nature.[2] Since most proteins that interact with cell surfaces do so by interacting with the cell surface immobilized glycans (e.g., glycocalyx cell surface of mammalian cells), this interaction can serve as inspiration for designing a glycan-based multimodal ligand or glycoligand.[7]

To synthesize an effective multimodal glycoligand, an aromatic alcohol acceptor needs to be chemically linked to a sugar moiety containing an ionic group like glucuronic acid or sialic acid. Chemical synthesis methods for synthesizing a glycan-based multimodal ligand can be very. *De novo* glycan synthesis also has very low yields due to the complex stereochemistry and regioselectivity of carbohydrate functional groups. Chemically

modifying already existing glycan polymers is also inefficient since they have multiple hydroxyl groups which makes it difficult to synthesis a multimodal ligand that is homogeneously functionalized. Enzymatic methods could be used instead for synthesizing a glycan-based multimodal ligand. Naturally, glycosyltransferases are used to synthesis glycans. These enzymes synthesize glycans by transferring nucleotide-sugar donors to the glycone or aglycone acceptor groups.[8] However, these enzymes are not often used for large scale *in-vitro* synthesis because they are membrane bound proteins that are difficult to express well, and also because nucleotide-sugar donor groups are expensive.[9] Another enzymatic strategy used to synthesis glycans is taking advantage of alternative enzymatic pathways of glycosyl hydrolases. Some glycosyl hydrolases can utilize the transglycosylation pathway, where a glycosyl acceptor group can be used as a nucleophile instead of water, to synthesize complex glycans using appropriate donor and acceptor groups. However, since the hydrolytic and transglycosylation pathway are always in competition with each other and since the transglycosylation pathway product can be used as a substrate for the hydrolytic pathway, this synthesis strategy typically commonly results in low yields.[10] To increase the yield of the synthesis product, some glycosyl hydrolases can be mutated into a glycosynthase by inactivating the nucleophilic amino acid residue by mutating it to either alanine, glycine and/or serine residue. Glycosynthases work by transferring an activated α-sugar donor to a glycone or aglycone acceptor to create a β glycosidic linkage.[11] Wilkinson et.al. found that by utilizing a glycosynthase derived from *E.coli* β-glucuronidase, a glucuronide can be synthesized containing various hydrophobic aromatic alcohols.[12]

While work by Wilkinson et.al. shows that glucuronic acid can be glycosidically

linked to hydrophobic alcohols through enzymatic synthesis with glycosynthases, another requirement must be met for the resulting glucuronide to be an effective chromatography ligand. The ligand product must contain some functional group chemistry capable of being linked to a chromatography resin. When choosing an appropriate resin-linking chemistry, certain criteria must be met. It should not require acidic conditions that could dehydrate the sugar, only require a minimal amount of solvent, if any, that can be easily removed, and create a covalent bond that isn't easily reduced. One such chemistry is to use an aldehyde functional group and a hydrazine resin to create a strong bio-orthogonally relevant hydrazone linkage with a high coupling affinity. We propose that the glycosynthase derived from the *E.coli* glucuronidase can synthesize glucuronides containing an aldehyde functional groups on the hydrophobic alcohol.

1.2 Aims of Thesis

1.) Use in-silico docking simulations to establish and justify a library of hydrophobic alcohol acceptors containing an aldehyde functional group.

2.) Synthesize and purify 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid using TEMPO oxidation and anion exchange chromatography.

3.) Generate a library of glycosynthase mutants from the *E.coli* glucuronidase gene, uidA, and express, purify and characterize the wild-type and mutant enzymes,

4.) Run glycoligand synthesis assays with an active glycosynthase mutant, activated glucopyranuronic acid donor and a library of alcohol acceptors to confirm if the uidA glycosynthases can synthesize multimodal ligands capable of being covalently linked to a resin.

To achieve these aims, previous work by Wilkinson et.al. and docking simulations were used to give insight on what possible alcohol acceptor structures should be modified to introduce conjugation chemistry like an aldehyde group onto the ligand.[12] Next, TEMPO oxidation of α-D-glucopyranosyl fluoride and anion exchange chromatography with an ammonium carbonate elution buffer was used to purify the activated glucuronic acid donor. Then, the glycosynthase library was designed using site-directed mutagenesis and purified using immobilized metal affinity chromatography. Finally, glycosynthase assays were performed to test the GS multimodal ligand synthesis hypothesis.

Chapter 2. *In-silico* modeling for developing aglycone library

To develop our alcohol acceptor library, we built off previous research by Wilkinson et.al. which used uidA derived glycosynthases to synthesize different glucuronides using a variety of primary, secondary and aryl alcohols. We selected aryl alcohols that were monosubstituted and reacted without the need of any additional additive like DMSO or DDM. This left us with phenylmethanol which had a maximum yield of 59%, and 2-phenylethanol which had a maximum yield of 93%.[12] Para-positioned aldehyde functional groups were added to the alcohols to create 4-(Hydroxymethyl)benzaldehyde and 4-(2-Hydroxyethyl)benzaldehyde. Aldehyde functional groups were selected so that in future work they can be easily linked to protein chromatography affinity supports like Ultralink[TM] Hydrazide Resin (Thermo Scientific). The resin's hydrazide groups can form stable hydrazone bonds with aldehydes with a >90% coupling affinity based on supplier specifications.

Protein and ligand docking simulations were performed to further justify the functional group additions to the alcohols, Phenylmethanol and 2-Phenylethanol. The simulations predict the position and the orientation of the target ligands when it is bound within the active site and comparisons between the binding free energies are used to justify the likelihood of certain positionings. AutoDock vina software was used to perform the ligand docking simulations. AutoDock vina is based on a rapid gradient-optimizing conformational search method to find the minimal free energy of interaction between the ligand and protein by accounting for ionic bonds, hydrogen bonds and van der Waals interactions.[13] It uses spherical symmetric hydrogen bond potentials, implicit hydrogens

and ignores electrostatic charge contribution which decreases the computational intensity of the program and uses gradient optimized search to guide the search to the energy minimum.[13]

The binding free energy of interaction calculated on AutoDock vina considers intermolecular interaction energies between the protein and the alcohol accounting for hydrogen bonds, van der Waals interactions and hydrophobic interactions. It also considers the total changes in the total internal energy of the system caused by docking of the alcohol in the active site and torsional energy caused by rotational restrictions on the amino acid side chains residues and the docked alcohol.

2.1 Method

The uidA crystal structure, pdb 3LPF, was saved in pdbqt file format which gives information about the coordinates and types of atoms in the crystal structure. The enzyme was crystalized by Wallace et.al. to study the inhibition of symbiotic bacterial β-glucuronidases in the gut. In this study, the crystal structure was determined to assist in a high throughput inhibitor design study for inhibiting the reactivation of colon cancer chemotherapeutic CPT-11 by the β-glucuronidase in the gut which causes several side effects.[14] The crystal structure was then cleaned up in PyMol to remove to the inhibitor and to only have one monomer of the crystallized tetramer. The nucleophilic amino acid in the enzyme active site was then mutated into glycine using the mutagenesis option in PyMol to better represent the uidA derived glycosynthase. The protein's grid box was set to (20,20,20) dimensions in the x, y and z direction centered around the location of the catalytic nucleophile (-26.302, 4,975, -3.666). This was done to confine the search space

to the catalytic pocket of the enzyme which allows for faster simulations and more relevant protein-ligand interaction data. Before docking the alcohol acceptors into the protein, the activated 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid was first docked into the active site to make sure that the orientation of the hydroxyl group of the alcohol positions itself near the anomeric carbon of the sugar moiety. The protein docked with the sugar donor was then used as the receptor in the alcohol docking simulations. The data obtained from the docking studies was analyzed using PyMol for 3D structural viewing and LigPlot to visualize the various types of interactions in a 2D view.

2.2 Phenylmethanol and 4-(Hydroxymethyl)benzaldehyde Docking Analysis

Shown here are overlaid PyMol images of the lowest binding free energy docked structures of Phenylmethanol (in blue) and its para-aldehyde variant 4-(2-Hydroxymethyl)benzaldehyde (in pink) within uidA, with dashed lines representing the distance between the acceptor's hydroxyl and anomeric carbon of the docked sugar.
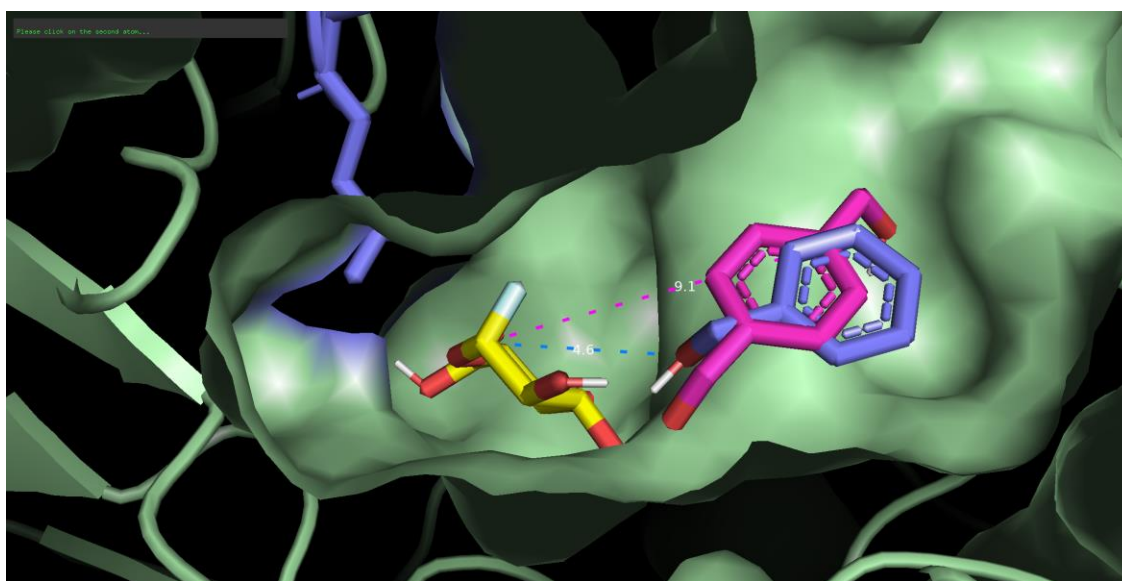
Figure 1. PyMol image of docking of Phenylmethanol and 4-(Hydroxymethyl)benzaldehyde with the lowest binding free energy. Dashed lines show the distance between the anomeric carbon of the docked sugar and the hydroxyl of each alcohol acceptor.

Binding Conformation	Binding Free Energy (kcal/mol)		Binding Conformation	Binding Free Energy (kcal/mol)
1	-5.0		1	-4.3
2	-5.0		2	-4.3
3	-4.9		3	-4.3
4	-4.9		4	-4.2
5	-4.9		5	-4.2
6	-4.8		6	-4.2
7	-4.8		7	-4.2
8	-4.8		8	-4.2
9	-4.8		9	-4.2

Table 1. Binding free energies of 4-(Hydroxymethyl)benzaldehyde (left) and Phenylmethanol (right) docked on uidA receptor containing 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid in the active site. Highlighted binding conformation shows binding free energy of docking shown in PyMol image above.

PyMol images analysis revealed that the lowest binding free energy docking structure does not have the alcohols in the proper orientation and position for reaction. All other higher or equal binding free energy structures in this simulation also do not place the alcohol in the correct position. However, since Phenylmethanol still reacts with a maximum yield of 59%, this supports the possibility that there may be some other binding orientations that have a lower binding free energy not fully captured in these simulations but still will be reacting with the donor.[12] Therefore, both Phenylmethanol and 4-(Hydroxymethyl)benzaldehyde were included in the alcohol acceptor library.

2.3 2-Phenylethanol and 4-(2-Hydroxyethyl)benzaldehyde Docking Analysis

LigPlot analysis showing the hydrogen bonding and the hydrophobic interactions on 2-Phenylethanol and its para-aldehyde variant 4-(2-Hydroxyethyl)benzaldehyde can be seen below with the binding free energies of each binding conformation.
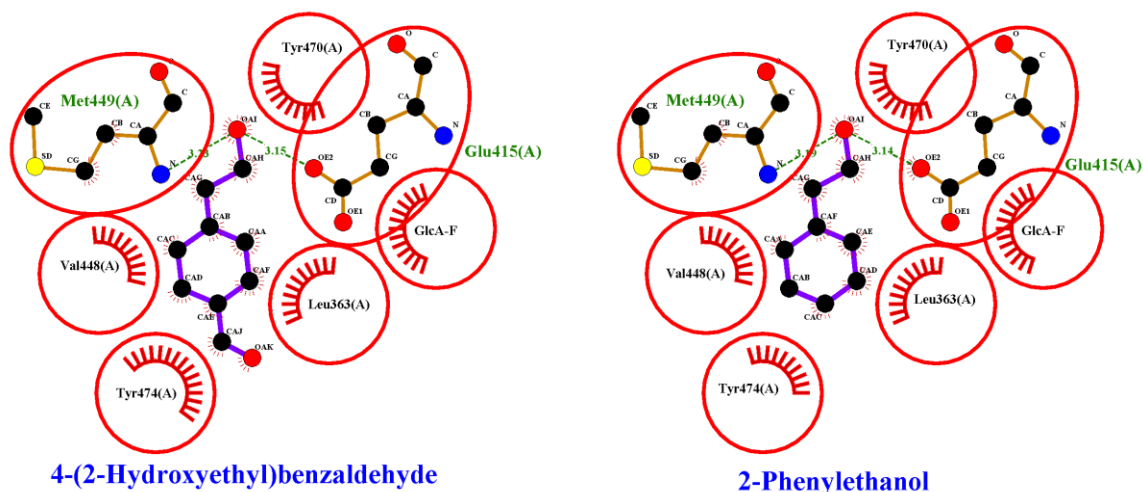
Figure 2. LigPlot analysis of 4-(2-Hydroxyethyl)benzaldehyde and 2-Phenylethanol docked on uidA receptor containing 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid in the active site.

Binding Conformation	Binding Free Energy (kcal/mol)	Binding Conformation	Binding Free Energy (kcal/mol)
1	-5.3	1	-4.9
2	-5.3	2	-4.7
3	-5.1	3	-4.7
4	-5.0	4	-4.7
5	-5.0	5	-4.7
6	-4.9	6	-4.6
7	-4.9	7	-4.6
8	-4.9	8	-4.6
9	-4.9	9	-4.5

Table 2. Binding free energies of 4-(2-Hydroxyethyl)benzaldehyde and 2-Phenylethanol docked on uidA receptor containing 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid in the active site. Highlighted binding conformation shows binding free energy of docking shown in LigPlot analysis image above.

The binding free energies of both ligands are similar with minimal differences between the free energy of these orientations and that of the orientation with the minimum free energy. The LigPlot analysis shows that both alcohols have similar interactions with the protein. They both form hydrophobic interactions with residues Tyrosine 470, Leucine 363, Tyrosine 474 and Valine 448. The both also form hydrogen bonds with the nitrogen in the Methionine 449 residue and with the carboxylic acid on the side chain of Glutamic Acid 415 which is the acid-base amino acid of the glucuronidase. The alcohol's hydroxyl group

also forms a hydrophobic interaction with the anomeric carbon of the 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid. The positioning near the acid-base residue and the anomeric carbon supports the possibility that the aldehyde variant may also be able to be glycosidically linked to the glucuronic acid donor. Further experimental work was done to justify these simulations.

Chapter 3. Synthesis and Purification of 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid

1-Deoxy-1-fluoro-α-D-glucopyranuronic acid is the activated sugar donor required for the glycosynthase reaction with the alcohol acceptor library discussed in the previous chapter. To synthesize this activated sugar donor, the sixth carbon of the readily available α-D-glucopyranosyl fluoride must be oxidized to a carboxylic acid. The selection of oxidizing agents is important when properly synthesizing 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid because, the oxidizing agent must be regioselective to only oxidize the primary hydroxyl on the sixth carbon of α-D-glucopyranosyl fluoride and not be reactive with any secondary hydroxyls on the sugar. One common class of oxidizing agent used when primary alcohol regioselectivity is desired are the nonconjugated nitroxyl radical oxidizing agents. In this research, 2,2,2,6-tetramethylpiperidine-1-oxyl (TEMPO) was used as the oxidizing agent with the addition of a sodium hypochlorite secondary oxidant. Finally, to purify the oxidized product from the unreacted substrate, anion exchange chromatography was used with an ammonium carbonate mobile phase to elute the bound oxidized product.

3.1 TEMPO Oxidation

TEMPO is a stable nitroxyl radical that can be activated to make an unstable oxoammonium salt oxidizing agent either by acid-catalyzed disproportionation or through the presence secondary oxidation like sodium hypochlorite.[15] This oxoammonium salt can then be utilized to oxidize primary and secondary alcohols. However, certain experimental

conditions can cause an increase in the reaction rate and selectivity towards primary alcohols. Under basic, aqueous conditions with the secondary oxidizing agent activation method, the oxidation mechanism favors the oxidation of only primary alcohols allowing for the oxidation of α-D-glucopyranosyl fluoride to 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid shown in the reaction scheme below.[15]
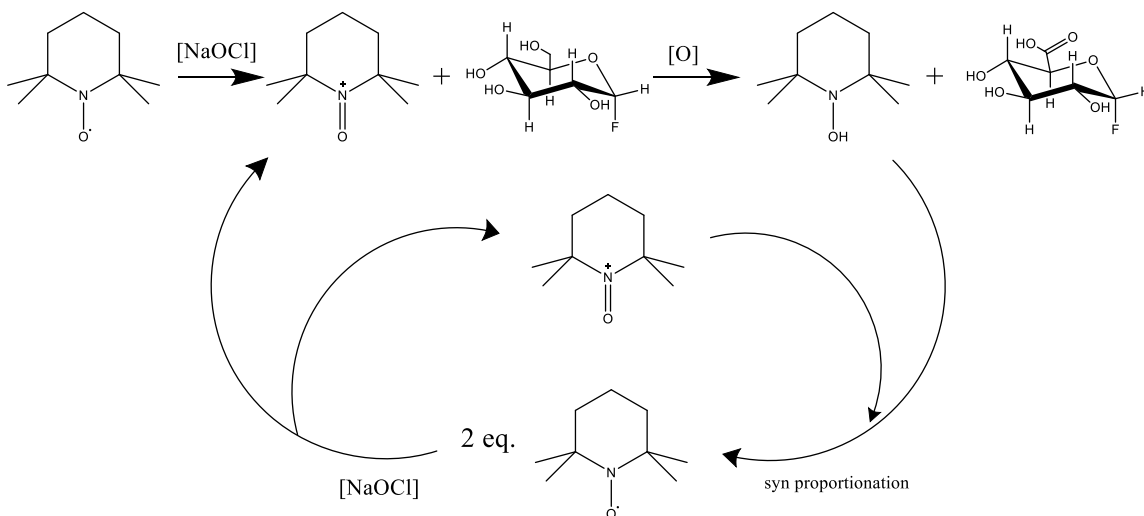
Figure 3. Oxidation reaction scheme of α-D-glucopyranosyl fluoride to 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid using TEMPO and sodium hypochlorite secondary oxidizing agent in basic, aqueous conditions.

The reaction scheme starts with the in-situ oxidation of TEMPO by sodium hypochlorite to generate the initial oxoammonium salt needed to begin the reaction cycle. The oxoammonium salt then oxidizes the α-D-glucopyranosyl fluoride and gets converted to a hydroxylamine. This reaction will initially generate the aldehyde product but, since the reaction is carried out in aqueous conditions the aldehyde product will be quickly converted to its hydrate form and get oxidized again to form the desired carboxylic acid product, 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid.[15] To restart this process, the oxidized oxoammonium salt and reduced hydroxylamine must react through syn proportionation to create two equivalents of TEMPO, which can then be oxidized by sodium hypochlorite to

create two equivalents of the oxoammonium salt oxidizing agent.[15] One of those equivalents can be used to restart the cycle while the other can be used for syn proportionation.

3.1.1 Materials and Methods

The composition of reaction mixture was similar to that described previously by Wilkinson et.al. with adjustments to a lower concentration of α-D-glucopyranosyl fluoride.[12] First, 100 mg of α-D-glucopyranosyl fluoride (0.55 mmol) was dissolved in 11 mL of PCR grade water and 4.6 mg of TEMPO (0.293 mmol) was added to the solution. To begin the reaction, the solution was placed in an ice bath and 1.564 mL of sodium hypochlorite (1.043 mmol) was added. The pH of the reaction was maintained around 10 with dropwise addition of 1 M NaOH. Next, 5 µL samples were taken from the reaction mixture every 15 minutes for the first hour and then every hour for the next 3 hours to be used for Thin Layer Chromatography (TLC) analysis. TLC analysis was done using a 250 µm thick silica gel GHLF TLC plate with a 3:2:1:1 ethyl acetate: 2-propanol: acetic acid: water mobile phase. The plate was visualized using a 0.1% orcinol, 10% sulfuric acid in ethanol stain and heating the plate at 100 °C for 10 minutes.

3.1.2 Results

A representative image of the stained TLC plate with a 1 µL loading of 10-fold diluted reaction mixture sample can be seen below.
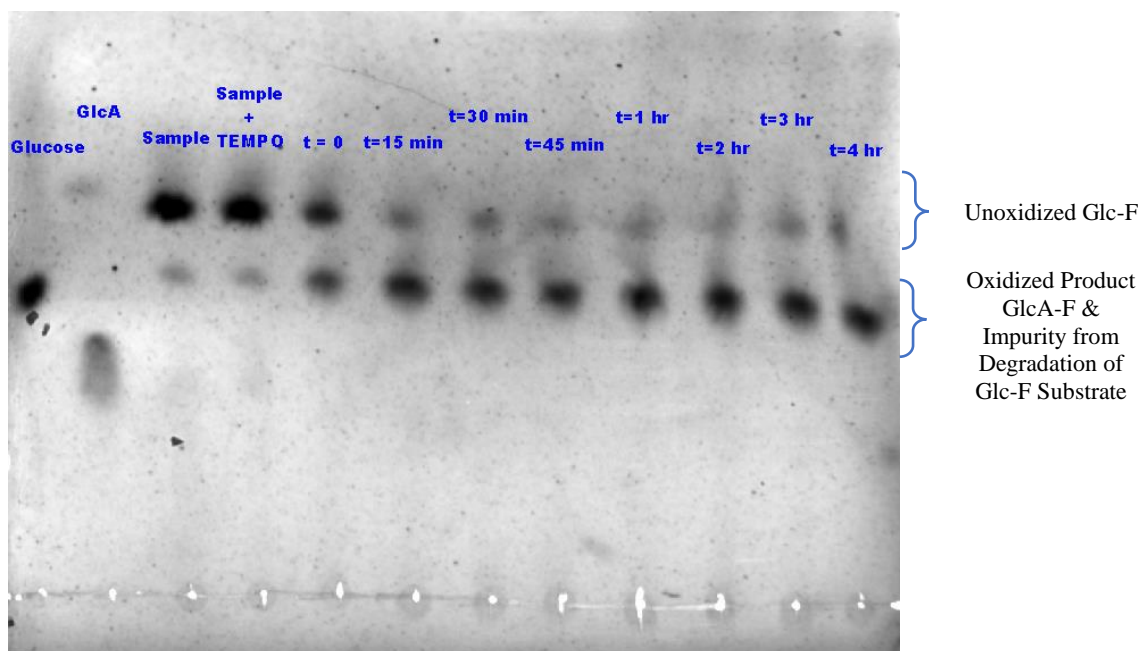
Figure 4. Stained TLC of the progression of the TEMPO Oxidation of α-D-glucopyranosyl fluoride over the course of 4 hours with glucose and glucuronic acid standards. TLC analysis was done using a 250 μm thick silica gel GHLF TLC plate with a 3:2:1:1 ethyl acetate: 2-propanol: acetic acid: water mobile phase and a 0.1% orcinol, 10% sulfuric acid in ethanol stain. TLC standards are a 1 μL sample of 50 mM glucose and a 1 μL sample of 50 mM glucuronic acid. TLC samples are a 1 μL loading of a 10-fold dilution of the reaction mixture before adding TEMPO, after adding TEMPO, after adding sodium hypochlorite (0 time point), after 15 minutes for the first hour and then every hour for an additional 3 hours.

Before adding sodium hypochlorite, there is some minor degradation of the α-D-glucopyranosyl fluoride substrate caused due to nucleophilic attach by water on the anomeric carbon to create glucose. Upon addition of the sodium hypochlorite, there was near instantaneous oxidation of α-D-glucopyranosyl fluoride seen at its corresponding C6-hydroxymethyl position. The oxidation product, 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid, runs identically like glucose under these TLC conditions. After about 15 minutes, the changes in the intensity of the oxidized product on the TLC plate is negligible. This reaction mixture was then frozen and lyophilized for anion exchange purification of the oxidized product.

3.2 Anion Exchange Chromatography

Anion exchange chromatography utilizes positively charged resins to form electrostatic interactions with the negatively charged analytes. This allows for uncharged and positively charged compounds to flow through the column and, separate from the desired negatively charged compound. To elute the negatively charged analyte, a salt buffer is passed through the column. The anion of the salt will exchange sites with the negatively charged analyte causing the analyte to flow through the column.
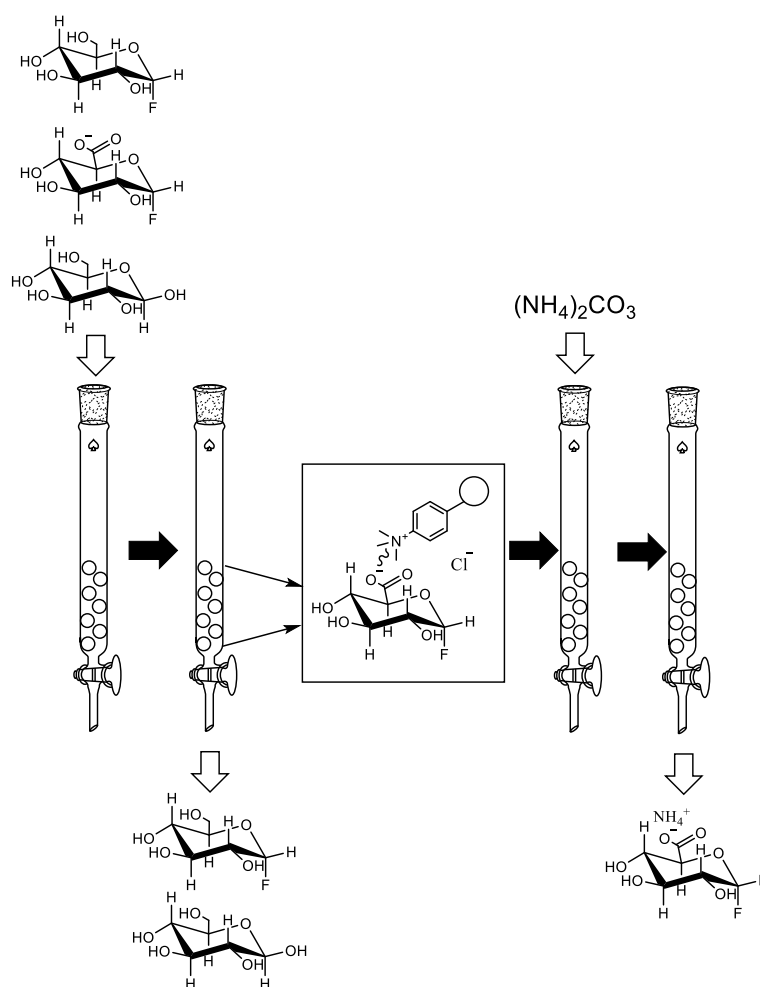

Figure 5. Diagram of 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid purification strategy using a DOWEX 1x8 anion exchange resin and an ammonium carbonate elution buffer.

This method can be used to separate 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid, which has a negatively charged carboxylic acid group, from the unreacted α-D-

glucopyranosyl fluoride and glucose impurity. A diagram of the purification strategy can be seen above.

In this experiment, a DOWEX 1x8 anion exchange resin, which has a positively charged trimethyl amine functional group, was used to bind to 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid and an ammonium carbonate buffer was used to elute the analyte from the column. Ammonium carbonate was selected as the elution buffer because during lyophilization, an ammonium carbonate solution will sublime into ammonia, water and carbon dioxide which will increase the purity of the lyophilized sample in comparison to using other salt solutions.

3.2.1 Materials and Methods

The lyophilized TEMPO oxidation sample was dissolved in 1 mL of DI water and was divided amongst two anion exchange chromatography runs. 10 grams of DOWEX 1x8 resin was wet-loaded onto a glass chromatography column and was cleaned with one column volume (20 mL) of DI water. Next, 0.5 mL of TEMPO oxidation sample was loaded directly onto the column. Two column volumes of water were run through the column followed by two column volumes of 0.1 M ammonium carbonate and then two column volumes of 0.2 M ammonium carbonate. During all elution steps, 1 mL fractions were collected.

3.2.2 Results

To identify which fractions contained sugars, 5 μL of each elution fraction was spotted onto a TLC plate and stained with a 0.1% orcinol, 10% sulfuric acid in ethanol based

staining solution. The appropriate fractions were then pooled together and analyzed again with TLC. Orcinol stained TLC images from anion exchange chromatography runs can be seen below.
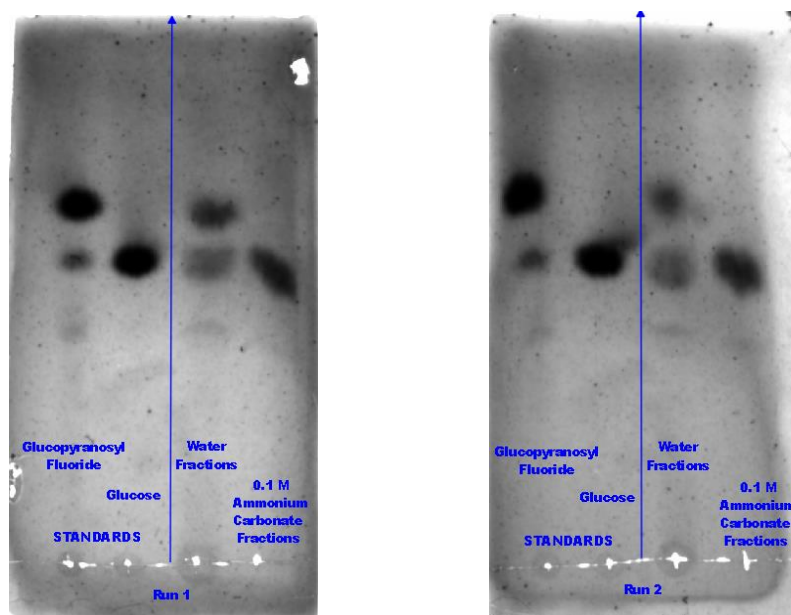
Figure 6. TLC analysis of water and 0.1 M ammonium carbonate fractions after anion exchange chromatograph of TEMPO reaction mixture. TLC analysis was done using a 250 μm thick silica gel GHLF TLC plate with a 3:2:1:1 ethyl acetate: 2-propanol: acetic acid: water mobile phase and a 0.1% orcinol, 10% sulfuric acid in ethanol stain. TLC standards are a 1 μL sample of 25 mM α-D-glucopyranosyl fluoride and a 1 μL sample of 25 mM glucose. TLC samples are a 2 μL loading of the pooled water fractions and a 2 μL of loading of the pooled 0.1 M ammonium carbonate fractions.

The water fractions contained α-D-glucopyranosyl fluoride and glucose since they both do not contain any negatively charged groups to strongly interact with the resin. While, the 0.1 M ammonium carbonate fractions contained mostly 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid. The 0.1 M ammonium carbonate elution fractions were then frozen and lyophilized to purify 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid for the subsequent glycosynthase reactions.

Chapter 4. Engineering, Expression, and Purification of uidA Glycosynthase Mutants

4.1 Site-Directed Mutagenesis to create uidA nucleophilic mutant library

To generate a library of glycosynthases from the *uidA* wild type glucuronidase gene, a point mutation was introduced to inactivate the nucleophilic amino acid, glutamic acid at 504 protein sequence position. To create a glycosynthase, this glutamic acid residue was mutated to either alanine, glycine or serine residue. Assays are then conducted to determine which, if any, mutations were effective at creating an active glycosynthase. Correlations between which mutation works and the structure of the enzyme is still being researched but, it has been found that almost all successful glycosynthase mutants are generated from point mutations at the nucleophilic site to either alanine, glycine or serine.[11]

To design effective primers for site directed mutagenesis, certain design conditions must be met. The primer length must be between 18 to 35 base pairs long with a compositional percentage of guanine & cytosine (%GC) between 40 to 60%. It is also preferred if the primers can start and end with guanine and cytosine to increase the strength of binding at the ends and that there are no complementary regions as to avoid the formation of a hairpin. Primer sequences to introduce alanine, glycine and serine can be seen below.

Point Mutation	Forward Primer	Reverse Primer
E504A	5'- ATCATTATCACCGCA TACGGTGTTGA -3'	5'- GTCAACACCGTATGC GGTGATAATGAT -3'
E504S	5'- ATCATTATCACCTCA TACGGTGTTGAC -3'	5'- GTCAACACCGTATGA GGTGATAATGAT -3'
E504G	5'- ATCATTATCACCGGA TACGGTGTTGAC -3'	5'- GTCAACACCGTATCC GGTGATAATGAT -3'

Table 3. Primers used for site directed mutagenesis of uidA wild type gene. Highlighted region represents the codon containing the point mutation.

4.1.1 Materials and Methods

The wildtype plasmid DNA stock, pET28a vector with the *uidA* gene insert, was custom synthesized by Genscript and then diluted into a 20 ng/ µL stock. The reaction volume for site directed mutagenesis includes 10 µL of Phusion High-Fidelity PCR Master Mix from ThermoFisher, which contains a high-fidelity DNA polymerase, nucleotides and a $MgCl_2$ buffer for optimal PCR conditions. As well as 20 ng of wildtype plasmid DNA, 0.5 µM of both the forward and reverse primers, and an additional amount of PCR water to make up the volume to 20 µL. The reaction was setup by adding each of the aforementioned reagents in the reverse order into a chilled PCR tube to ensure that the DNA doesn't get damaged and to prevent the polymerase from starting the reaction early. The PCR reaction conditions used are summarized below.

	Process	Temperature (°C)	Duration (sec)	
1	Initial Denaturation	98	30	
2	Denaturation	98	10	30 Cycles
	Annealing	E504A: 64 E504S: 61 E504G: 63	30	
	Extension	72	210	
3	Final Extension	72	300	
4	Hold	10	∞	

Table 4. Thermocycler conditions used in site-directed mutagenesis of glutamic acid 504 to alanine, serine and glycine.

Next, 5 µL of PCR reaction product was then run on a 0.7 % agarose 1x SYBR safe DNA gel at 120 V for 40 minutes. The presence of a prominent single band on the DNA gel at the proper location shows that the site directed mutagenesis was successful. Following the PCR and gel, the leftover PCR product then underwent Dpn1 digestion. Dpn1 is a methylation-dependent endonuclease used to eliminate all parental template plasmid DNAs that is necessary to remove most wildtype DNA before transformation of

the PCR synthesized product. The reacted PCR product was then used for heat shock transformation into E. cloni competent cells, which is a strain of E. coli optimized for plasmid propagation and cloning. The transformation mixture was streaked onto a kanamycin containing LB agar plate and incubated at 37 °C for 16 hours. Individual bacterial colonies on the agar plate were isolated and screened for the desired mutation using a MiniPrep Plasmid Extraction kit from IBI Scientific followed by DNA sequencing (Genscript). The sequence was confirmed by sending the extracted plasmid DNA to Genscript for sequencing using T7 forward and T7 reverse primers, which were synthesize by IDTDNA. After confirming proper DNA mutation sequence, the extracted plasmid DNA was transformed into BL21 (DE3) competent cells for protein expression.

4.1.2 Results

Proper analysis of DNA sequencing results is essential to ensure that the extracted DNA is the intended product. First, the forward and reverse DNA sequencing results must overlap to show that there is no possibility of additional errors or inserts in the DNA. Then, the sequencing results must match the consensus DNA sequence except at the location of the desired point mutations. Finally, the chromatogram must be reliable and clean to be sure that the final read sequence is indeed correct. The DNA chromatograms of the 15 bp stretch around the point mutation for alanine, glycine and serine can be seen below.
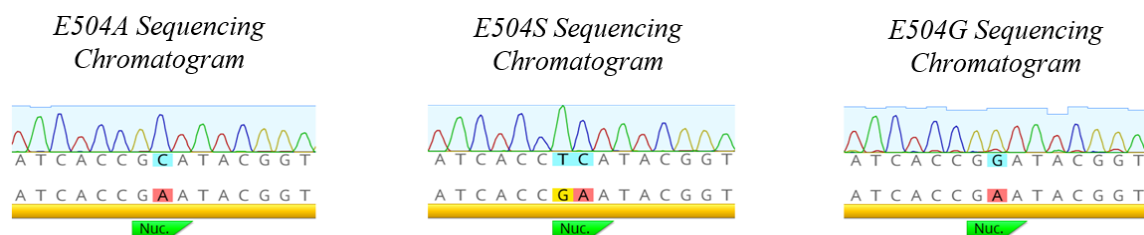

Figure 7. Sequencing chromatogram showing 15 bp stretch of DNA around the successful glutamic acid 504 point mutations to alanine, serine and glycine. Viewed using the *Geneious* software.

4.2 Expression of uidA protein library

The pET28a vector uses an IPTG induction system with a T7 promoter control and an N-terminal histidine tag. To express proteins of interest in large scale culture, IPTG is added to remove the lac repressor from the lac operator adjacent downstream of the T7 promoter site once the optical density of the culture reaches between 0.6 to 0.8 which indicates that the culture is in the mid to late log phase. Then after 24 hours, the cells are spun down and harvested for sonication.

4.2.1 Materials and Methods

Each starter culture was a 100 mL LB media culture with kanamycin antibiotic which was inoculated by its respective BL21 glycerol stock. The starter culture was then grown at 37 °C for 16 hours at 200 rpm in a New Brunswick Scientific Excella E24 series incubator shaker. Next, 50 mL of starter culture was transferred to a 1 L large culture containing kanamycin antibiotic. The larger culture was then grown at 37 °C until the optical density is between 0.6 – 0.8 and then IPTG was added so the final concentration is 0.5 mM in the large culture. The induced culture was then grown at 25 °C for 24 hours. The cell pellets are harvested by spinning down the media at 10000 rpm for 15 minutes in a Beckman Coulter centrifuge with a JA-14 fixed angle rotor.

4.3 Protein Purification

The protein purification strategy involves using a combination of chemical and mechanical methods for cell lysis, followed by centrifugation and syringe filtration to isolate the cell lysate supernatant. Next, the his-tagged proteins are purified by immobilized metal affinity

chromatography and buffer exchanged into a stable low salt concentration buffer.

4.3.1 Materials and Methods

To recover the expressed proteins from the harvested cell pellets, the pellets were first suspended in cell lysis buffer consisting of 20 mM sodium phosphate buffer at pH 7.4 with 500 mM NaCl and 20% v/v glycerol. Then in order to prevent the degradation of the expressed protein by proteases released during cell lysis, a protease inhibitor cocktail is added. Finally, lysozyme is also added to assist in cell lysis through enzymatic degradation of the cell wall. For every 3 grams of cell pellet, 15 ml of lysis buffer, 15 µL of lysozyme and 200 µL of PIC were added. The suspended cell lysis was then sonicated using a Misonix Sonicator 3000 at an output amplitude level of 25 for a 10 second pulse on time and a 30 second pulse off time for a total period of 4 minutes with a 1-minute break after every minute of pulse on time. Following sonication, the lysate was centrifuged at 11500 rpm for 1 hour with a Beckman Coulter centrifuge with a JA-10 fixed angle rotor to remove all intact cells, cell walls, plasma membranes and genomic DNA. Then the supernatants were syringe filtered with a 0.22 µm filter to remove all leftover large impurities and to prepare the supernatant for IMAC purification.

In immobilized metal affinity chromatography, the column is initially equilibrated in a low imidazole concentration buffer, IMAC A, which consists of 100 mM MOPS, 10 mM imidazole, and 500 mM NaCl. Then the supernatant is passed through the column where the histidine tags on the protein form complexes with the nickel resin. All other proteins are washed out of the column using an IMAC A wash. Then a 95 % IMAC A + 5% IMAC B buffer is passed through the column. IMAC B is a high imidazole

concentration buffer which consists of 100 mM MOPS, 500 mM imidazole, and 500 mM NaCl. Using a buffer of only 5% IMAC B will clean the column by removing all nonspecifically bound proteins interacting with the nickel resin. Then a 100 % IMAC B elution buffer is used to elute out the his-tag protein. Finally, to transfer the proteins into a more stable lower salt storage buffer, a Sephadex G-25 in PD-10 desalting column is used to buffer exchange the proteins into a 10 mM sodium phosphate buffer at pH 7.5.

4.3.2 Results

The purity of the protein was analyzed using an SDS-PAGE protein gel with a Coomassie blue stain. The stained protein gel can be seen below. This protein gel shows that the IMAC purification was effective at removing all impurities and there is no cleavage leaving a truncated protein with a histidine tag.
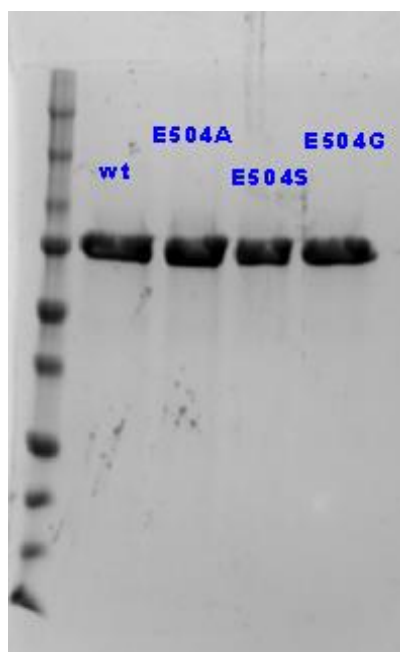
Figure 8. Coomassie blue stained SDS-PAGE protein gel of uidA wildtype enzyme and glycosynthase mutants.

To measure the concentration of the enzymes in solution, a Molecular Devices Spectramax M5E microplate reader is used to measure the absorbance of 4 µL of protein solution at 280 nm across a 0.1 cm pathlength. Then Beer-Lamberts' law is used to relate the average absorbance of the protein to its concentration. A table average absorbances and concentrations of the proteins can be seen below.

	Average Absorbance	Molecular Weight (Da)	Concentration (mg/mL)
WT	0.324	70610	1.625
E504A	0.478	70552	2.4
E504S	0.328	70568	1.64
E504G	0.390	70538	1.95

Table 5. Protein concentrations of uidA wild type enzyme and glycosynthase mutants using a 10 mM sodium phosphate blank, protein extinction coefficient of 140760 and 0.1 cm path length.

Chapter 5. Enzymatic Synthesis of Aldehyde Containing Glucuronides

Now that the alcohol acceptor library has been designed, the sugar donor substrate has been synthesized and purified, and the glycosynthase enzymes have been expressed and purified all that is left is to do before the reactions is to assess which glycosynthase mutant should be used. To assess which glycosynthase mutant is the most effective, multiple strategies have been developed. One strategy is to use the addition of an external nucleophile, like formate or azide, to attempt to recover the hydrolytic activity of the glycosynthase mutants. In research, a correlation has been found between if the mutant has its hydrolysis activity returned in the presence of an external nucleophile and the glycosynthetic activity of that mutant.[16] Another strategy is to use small scale assays with the glycosynthase mutants and the appropriate donors and acceptors to observe which mutant generates a larger amount of product. In our case, previous research by Wilkinson et.al. found that the nucleophilic mutation to glycine generated the most active uidA glycosynthase with a large acceptor library consisting of primary and cyclic secondary aliphatic alcohols, substitute benzyl alcohols, and isomeric naphthalene methanols. With the alanine mutant having much lower yield of β-glucuronide product and the serine mutant being inactive in all cases.[12] Therefore, we chose to exclude the alanine and serine glycosynthase mutants in our assays and skip running chemical rescue assays as well.

5.1 Glycosynthase Reactions

A diagram of the glucuronide synthesis reactions with the 1-Deoxy-1-fluoro-α-D-

glucopyranuronic acid sugar donor and the alcohol acceptor library consisting of the substituted benzyl alcohols, Phenylmethanol and 2-Phenylethanol, and the aldehyde-containing substituted benzyl alcohols, 4-(Hydroxymethyl)benzaldehyde and 4-(2-Hydroxyethyl)benzaldehyde can be seen below.
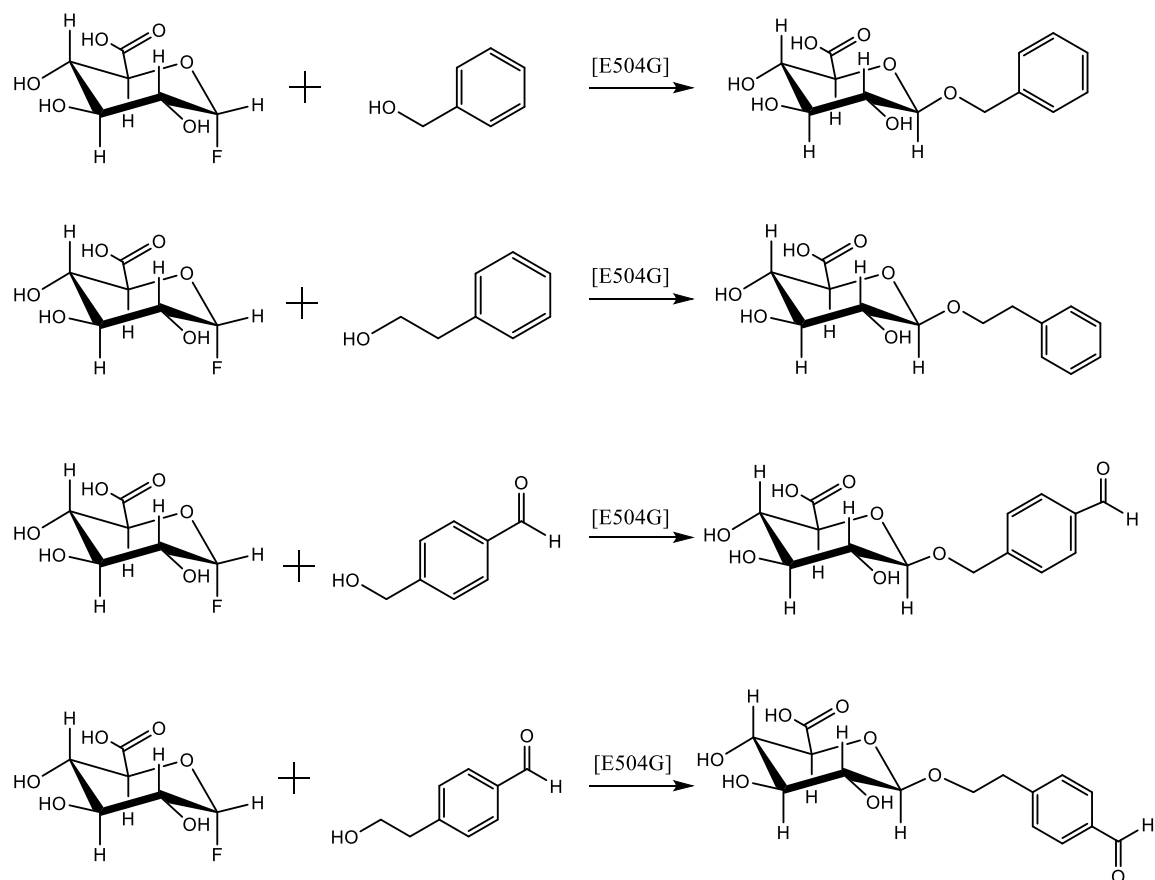

Figure 9. Diagram of the glucuronide synthesis reactions with 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid as the activated sugar donor, with Phenylmethanol, 2-Phenylethanol, 4-(Hydroxymethyl)benzaldehyde and 4-(2-Hydroxyethyl)benzaldehyde as the library of alcohol acceptors and uidA_E504G as the enzyme catalyst.

5.1.1 Materials and Methods

Reaction mixture contained 10 mM of 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid with 20 mM of either Phenylmethanol, 2-Phenylethanol, 4-(Hydroxymethyl)benzaldehyde or 4-(2-Hydroxyethyl)benzaldehyde and 1000 picomoles of uidA E504G glycosynthase in a 1 mL reaction volume containing 50 mM sodium phosphate pH 7.5 buffer. Reaction was run

at 25 °C for 24 hours in an Eppendorf Thermomixer C at 300 RPM. TLC analysis was done using a 250 μm thick silica gel GHLF TLC plate with a 3:2:1:1 ethyl acetate: 2-propanol: acetic acid: water mobile phase. The plate was visualized using a 0.1% orcinol, 10% sulfuric acid in ethanol stain and heating the plate at 100 °C for 10 minutes.

5.1.2 Results

TLC images of orcinol-stained reaction plates before the addition of the glycosynthase and after the reaction ran for 24 hours can be seen below.
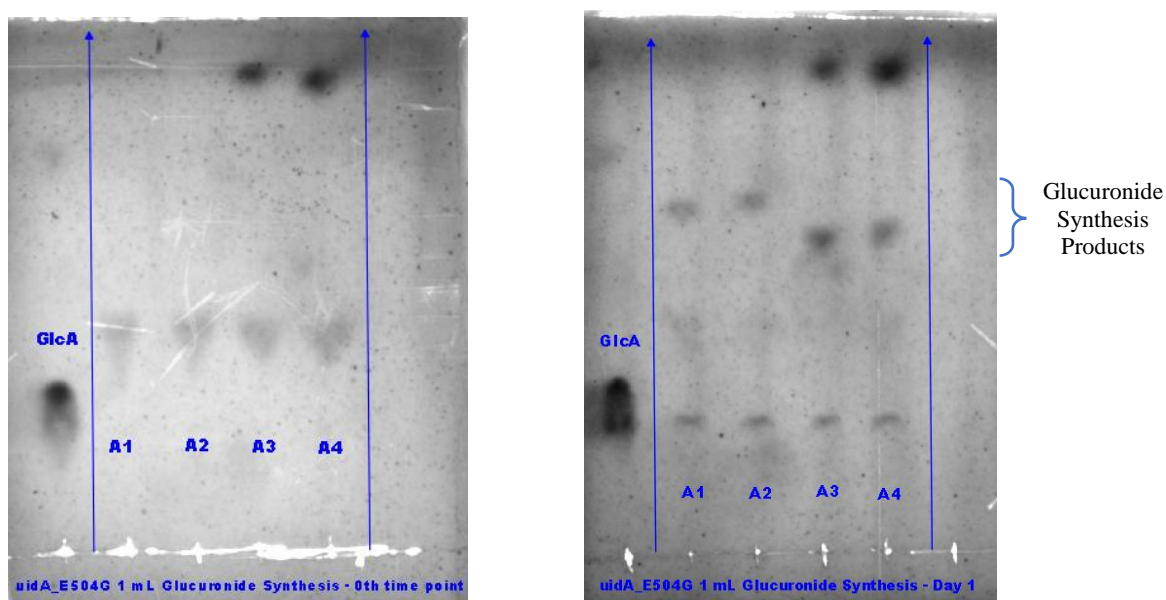

Figure 10. TLC images of glycosynthase reactions before addition of protein and after 24 hour reaction time with 10 mM of 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid and 20 mM of alcohol acceptors, A1 = Phenylmethanol, A2 = 2-Phenylethanol, A3 = 4-(Hydroxymethyl)benzaldehyde, A4 = 4-(2-Hydroxyethyl)benzaldehyde.

Of all the alcohol acceptors, only the aldehyde containing alcohols, 4-(Hydroxymethyl)benzaldehyde and 4-(2-Hydroxyethyl)benzaldehyde, represented by A3 and A4 respectively, can be seen on the top of the stained TLC plates. The explanation for this can be seen in Appendix Section A which discusses the orcinol staining mechanism.

After 24 hours, there is complete conversion of 1-Deoxy-1-fluoro-α-D-

glucopyranuronic acid into the glucuronide products and to glucuronic acid. This supports our hypothesis that the uidA glycosynthase could synthesize glycans with aldehyde-containing alcohol acceptors which can then be used as multimodal ligands.

Chapter 6: Future Work

After showing that it is possible to synthesize a multimodal ligand using the uidA based glycosynthase, the product must be purified using column chromatography and lyophilized for further reactions and binding studies. One property that is important for binding strength in multimodal ligands is charge density. In HCIC, increasing the charge density would cause the electrostatic repulsion to be greater, making the elution stronger. Placing the charged groups in different sequences can also possibly change the way proteins interact with the ligand potentially effecting the elution of proteins. To increase the charge density on the ligand, different activated sugar donors need to be enzymatically linked to the multimodal ligand which would act as the sugar donor. Müllegger et.al. showed that is possible to use a glycosynthase derived from a *Thermotoga maritima* glucuronidase to enzymatically link 1-Deoxy-1-fluoro-α-D-glucopyranuronic acid donor to 4-Nitrophenyl-β-D-glucuronic acid acceptor.[17] Using this enzyme, we should be able to extend the length of the multimodal ligand with more glucuronic acid sugars with the goal of creating a library of ligands with different charge density. After establishing a library of ligands with different degrees of polymerization and charge density, they will be conjugated to hydrazine resin and used in binding studies. This will allow us to assess the binding efficiency of our library in comparison to commercially available multimodal ligands.

References

1. Kallberg, K., Johansson, H.O., Bulow, L., "Multimodal Chromatography: An efficient tool in downstream processing of proteins," *Journal of Biotechnology*, vol. 7, 2012, pp. 1485-1495.

2. Zhao, G., Dong, X., Sun, Y., " Ligands for mixed-mode protein chromatography: Principles, characteristics and design," *Journal of Biotechnology*, vol. 144, 2009, pp. 3-11.

3. Burton, S.C., Harding, D.R.K., "Hydrophobic charge induction chromatography: salt independent protein adsorption and facile elution with aquoes buffers," *Journal of Chromatography A*, vol. 814, 1998, pp. 71-81.

4. Yon, R.J., Simmonds, R.J., "Protein Chromatography on Adsorbents with Hydrophobic and Ionic Groups," *Journal of Biochemistry*, vol. 151, 1975, pp. 281-290.

5. Halan, V., Maity, S., Bhambure, R., Rathore, A.S., "Multimodal Chromatography for Purification of Biotherapeutics – A Review," *Current Protein and Peptide Science*, vol. 20, 2019, pp. 4-13.

6. Johansson, B.L., Belew, M., Eriksson, S., Glad, G., Lind, O., Maloisel, J.L., Norman, N., "Preparation and characterization of prototypes for multi-modal separation media aimed for capture of negatively charged biomolecules at high salt conditions," *Journal of Chromatography A*, vol. 1016, 2003, pp. 21-33.

7. Varki, A., Lowe, J.B., "Chapter 6: Biological Roles of Glycans," *Essentials of Glycobiology. 2nd edition.,* Cold Spring Harbor Laboratory Press, 2009.

8. Lairson, L.L., Henrissat, B., Davies, G.J., Withers, S.G., "Glycosltransferases: Structures, Functions and Mechanisms," *Annual Review of Biochemistry*, vol. 77, 2008, pp. 521-555.

9. Boltje, T.J., Buskas, T., Boons, G.J., "Opportunites and challenges in synthetic oligosaccharide and glycoconjugate research," *Nature Chemistry*, vol. 1, 2009, pp. 611-622.

10. Bissaro, B., Monsan, P., Fauré, R., O'Donohue, M.J., "Glycosynthesis in a waterworld: new insight into the molecular basis of transglycosylation in retaining glycoside hydrolases," *The Biochemical Journal*, vol. 467, 2015, pp. 17-35.

11. Cobucci-Ponzano, B., Strazzulli, A., Rossi, M., Moracci, M., "Glycosynthases in biocatalysis", *Advanced Synthesis and Catalysis*, vol. 353, 2011, pp. 2284-2300.

12. Wilkinson, S.M., Liew, C.W., Salleh, H.M., Withers, S.G., McLeod, M.D., "*Escherichia coli* Glucuronylsynthase: An Engineered Enzyme for the Synthesis of β-Glucuronides," *Organic Letters*, vol. 10, 2008, pp. 1585-1588.

13. Forli, S., Huey, R., Pique, M.E., Sanner, M., Goodsell, D.S., Olson, A.J., "Computational protein-ligand and virtual drug screening with the AutoDock Suite," *Nature Protocols*, vol. 11, 2016, pp. 905-919.

14. Wallace, B.D., Wang, H., Lane, K.T., Scott, J.E., Orans, J., Koo, J.S., Venkatesh, M., Jobin, C., Yeh, L.A., Mani, S., Redinbo, M.R., "Alleviating Cancer Drug Toxicity by Inhibiting a Bacterial Enzyme," *Science*, vol. 330, 2010, pp. 831-836.

15. de Nooy, A.E.J, Besemer, A.C., van Bekkum, H., "On the Use of Stable Organic Nitroxyl for the Oxidation of Primary and Secondary Alcohol," *Synthesis*, 1996, pp. 1153-1174.

16. Zechel, D.L., Reid, S.P., Stoll, D., Nashiru, O., Warren, R.A.J., Withers. S.G., "Mechanism, Mutagenesis, and Chemical Rescue of a β-Mannosidase from *Cellulomonas fimi*," *Journal of Biochemistry*, vol. 42, 2003, pp. 7195-7204.

17. Müllegger, J., Chen, H.M., Chan, W.Y., Reid, S.P., Jahn, M., Warren, R.A.J., Salleh, H.M., Withers, S.G., "Thermostable Glycosynthases and Thioglycoligases Derived from *Thermotoga maritima* β-Glucuronidase," *ChemBioChem*, vol. 7, 2006, pp. 1028-1030.

18. Thetford, D., "Triphenylmethane and Related Dyes," Van Nostrand's Encyclopedia of Chemistry, 2005, pp. 1-21.

Appendix

A. Orcinol Staining Mechanism

The orcinol stain used in this research was a 0.1 %w/v orcinol, 10 %v/v sulfuric acid in ethanol stain. This stain is traditionally used to stain glycans on a TLC plate by spraying the plate and then heating it up at 100 °C for 10 minutes. However, it can also be used to visualize the aldehyde-containing alcohols used in our glucuronide synthesis assays. An example of the orcinol staining mechanism with 4-(2-Hydroxyethyl)benzaldehyde can be seen below.
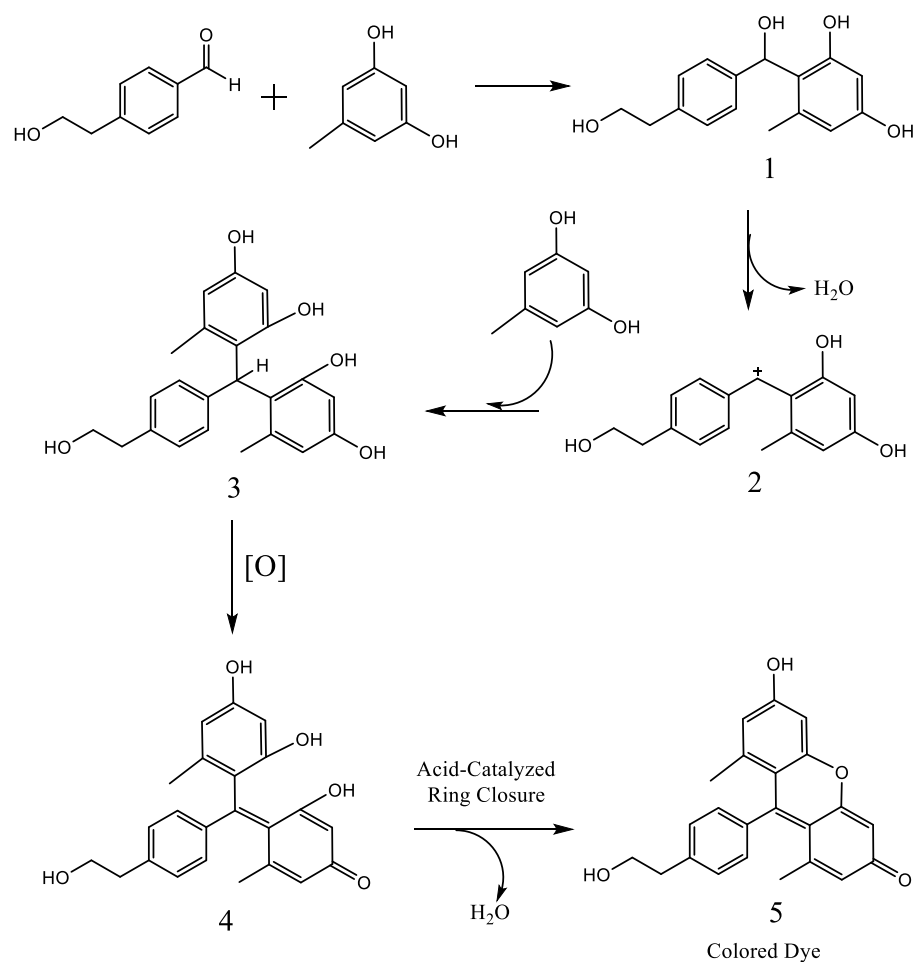

Figure 11. Orcinol staining mechanism for 4-(2-Hydroxyethyl)benzaldehyde.

This mechanism requires an aromatic aldehyde to first be protonated by sulfuric acid and then undergo nucleophilic substitution at the electrophilic carbonyl carbon by the reactive para-position of the orcinol reagent. This generates intermediate 1 which undergoes acid-catalyzed dehydration to make intermediate 2. Intermediate 2 has is a resonance stabilized carbocation which can undergo nucleophilic substitution with the reactive para-position of another equivalent of orcinol to make intermediate 3. Intermediate 3 is a triphenylmethane, which are commonly used as dyes upon oxidation.[18] Intermediate 3 is oxidized, which is facilitated by heating the TLC plate at 100 °C for 10 minutes, to make intermediate 4 which undergoes quick acid-catalyzed ring closure to make the final product which is the colored dye seen on the TLC plate. Glycans also undergo the same mechanism after going through acid-catalyzed dehydration with the sulfuric acid to make a furfural derivative which has the aldehyde required for the orcinol staining mechanism.