# University of Alberta

Synthesis of Furanose Sugar Nucleotides from *Mycobacterium tuberculosis* and *Campylobacter jejuni*

by

Ryan Blake Snitynsky

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Chemistry



Spring 2014
Edmonton, Alberta

*Dedicated to*

*Anne Snitynsky*

*and*

*Mita Dasog;*

*One gave me roots*

*And the other gave me wings.*

# Abstract

Five-membered ring furanose sugars, though absent in mammals, have been found in the cell walls of pathogenic bacteria such as *Mycobacterium tuberculosis* and *Campylobacter jejuni*. Because these sugars are only found in microorganisms, it has been proposed that inhibiting the enzymes these bacteria use to process furanose sugars could lead to the development of new targeted treatments against disease.

To study these enzymes, access to their natural substrates and analogues is essential. In this work, we explored methods of producing these sugar nucleotides. First, activated donors were used to attempt to make a glycosyltransferase enzyme (GlfT2) work in reverse to produce sugar nucleotides. Subsequently, we attempted to expand the substrate scope of a nucleotidyltransferase (GalPUT). Finally, a wholly synthetic approach was developed to synthesize two nitrogen-containing furanose sugar nucleotides.

# Acknowledgments

My sincere gratitude goes out to Dr. Todd L. Lowary, under whose supervision I carried out this work. Todd's guidance and support over the past couple years has been phenomenal, and his friendship will be forever valued. I would also like to thank Dr. M. Soledade C. Pedras from the University of Saskatchewan for allowing me to work in her lab for a number of years as an undergraduate, and first introducing me to research.

I would next like to thank Dr. Robert Campbell and Dr. Julianne Gibbs-Davis, my committee members, for their helpful suggestions and patience as my research progressed. Also, Mark Miskolzie, Wayne Moffat, Gareth Lambkin, and the staff of the Mass Spectrometry Laboratory were instrumental in helping me obtain the technical expertise and characterization details required for my project. Special thanks also to Anita Weiler for everything administration related; without you, I'm certain the Department would implode.

Every member of the Lowary group deserves special mention here, but for want of space all I can say is that you are all an amazing group of people. I have learned so much from all of you, and will forever treasure my memories of working with you. An extra-special thanks to three individuals who really helped me get my project started: Dr. Myles Poulin, Dr. Maju Joe, and Blake Ruixiang Zheng.

In addition to the Lowary group, I would also like to thank my friends from other groups that have helped cheer me up on even the darkest days: Taras, Ho-Yan, Burçin, Shuting, Samantha, and Melanie. I also owe a deep appreciation

to the Volya Ukrainian Dance Ensemble for helping to make Edmonton feel like home. Finally, I am grateful to my wife Mita for her love and support throughout my degree.

I am indebted to the following agencies for funding: Natural Sciences and Engineering Research Council of Canada, Alberta Innovates – Technology Futures, Alberta Glycomics Centre, and the U of A Faculty of Graduate Studies and Research.

# Table of Contents

# List of Figures

# List of Schemes

# List of Abbreviations

| | |
|---|---|
| $[\alpha]_D$ | Specific rotation (sodium D line) |
| Ac | Acetyl |
| $Ac_2O$ | Acetic anhydride |
| AcOH | Acetic acid |
| AcONa | Sodium acetate |
| ADP | Adenosine diphosphate |
| Alt | Altrose |
| Amp | Ampicillin |
| Ar | Aromatic |
| Ara*f* | Arabinofuranose |
| Asp | Aspartate |
| Bn | Benzyl |
| Bz | Benzoyl |
| *C. jejuni* | *Campylobacter jejuni* |
| CDI | 1,1′-Carbonyldiimidazole |
| CIAP | Calf intestinal alkaline phosphatase |
| CPS | Capsular polysaccharide |
| DAST | Diethylaminosulfur trifluoride |
| DCM | Dichloromethane |
| DIPEA | *N*,*N*-Diisopropylethylamine |
| DMC | 2-Chloro-1,3-dimethylimidazolinium chloride |
| DNA | Deoxyribonucleic acid |

| | |
|---|---|
| dTDP | Thymidine diphosphate |
| *E. coli* | *Escherichia coli* |
| ESI | Electrospray injection |
| $Et_3N$ | Triethylamine |
| EtOAc | Ethyl acetate |
| FCC | Flash column chromatography |
| FTIR | Fourier transform infrared |
| Fuc | Fucose |
| Gal/Gal*p* | Galactose/Galactopyranose |
| Gal*f* | Galactofuranose |
| Gal*f*NAc | 2-Acetamido-2-deoxy-α-D-galactofuranose |
| GalPUT | Galactose-1-phosphate uridyltransferase |
| GalU | UDP-glucose pyrophosphorylase |
| GBS | Guillain–Barré Syndrome |
| Glc/Glc*p* | Glucose/glucopyranose |
| GlcA6 | Glucuronic acid |
| GlcNAc/Glc*p*NAc | *N*-Acetylglucosamine |
| GlfT1 | UDP-Galactofuranosyltransferase I |
| GlfT2 | UDP-Galactofuranosyltransferase II |
| GT | Glycosyltransferase |
| Hep | Heptose |
| HEPES | 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid |
| HPLC | High performance liquid chromatography |

| | |
|---|---|
| HRMS | High resolution mass spectrometry |
| IPP | Inorganic pyrophosphatase |
| Kan | Kanamycin |
| Kdo | 3-Deoxy-D-*manno*-oct-2-ulosonic acid |
| LAM | Lipoarabinomannan |
| LB | Luria–Bertani |
| LDH | Lactate dehydrogenase |
| LM | Lipomannan |
| LOS | Lipooligosaccharide |
| *M. tuberculosis* | *Mycobacterium tuberculosis* |
| mAG | Mycolyl–Arabinogalactan |
| MBP | Maltose binding protein |
| MDR-TB | Multidrug-Resistant Tuberculosis |
| MeOH | Methanol |
| MeOPN | O-Methyl phosphoramidate |
| MHz | Megahertz |
| MOPS | 3-(*N*-Morpholino)propanesulfonic acid |
| MQ | Milli-Q ultrapure water |
| MS | Mass spectrometry |
| NADH/$NAD^+$ | Nicotinamide adenine dinucleotide |
| NaOMe | Sodium methoxide |
| NDP | Nucleotide diphosphate |
| NEtn | Ethanolamine |

| | |
|---|---|
| Neu5Ac | *N*-Acetylneuraminic acid |
| NGro | Aminoglycerol |
| NMP | Nucleotide monophosphate |
| NMR | Nuclear magnetic resonance |
| NT | Nucleotidyltransferase |
| OleD | Oleandomycin glycosyltransferase |
| PCR | Polymerase chain reaction |
| PEP | Phosphoenolpyruvate |
| PEtn | Phosphoethanolamine |
| PG | Peptidoglycan |
| PK | Pyruvate kinase |
| *p*-TsOH | *p*-Toluenesulfonic acid |
| Qui3NAc | *N*-Acetylquinovosamine |
| $R_f$ | Retention factor |
| Rha/Rha*p* | Rhamnose/Rhamnopyranose |
| Rib | Ribose |
| RNA | Ribonucleic acid |
| $S_N2$ | Substitution nucleophilic bi-molecular |
| $S_Ni$ | Substitution nucleophilic internal |
| TB | Tuberculosis |
| TBAF | *tert*-Butyl ammonium fluoride |
| TBS | *tert*-Butyldimethylsilyl |
| TEAA | Triethylammonium acetate |

| | |
|---|---|
| $TfN_3$ | Trifluoromethanesulfonic azide |
| THF | Tetrahydrofuran |
| TLC | Thin layer chromatography |
| TMS | Trimethylsilyl |
| Tos-$NHNH_2$ | *p*-Toluenesulfonyl hydrazide |
| Tris | Tris(hydroxymethyl)aminomethane |
| UDP | Uridine diphosphate |
| UGM | UDP-Galactopyranose mutase |
| UMP | Uridine monophosphate |
| UNGM | UDP-*N*-Acetylgalactopyranose mutase |
| UTP | Uridine triphosphate |
| UV | Ultraviolet |
| WbbI | β-(1→6)-Galactofuranosyltransferase from *E. coli* K-12 |
| WHO | World Health Organization |
| XDR-TB | Extensively Drug-Resistant Tuberculosis |
| Xul | Xylulose |

# Chapter 1 : Introduction

## 1.1 Overview

Carbohydrates are essential for the pathogenesis and growth of microorganisms that infect humans. Long oligosaccharide chains called glycans are present on the cell surface and within the cell wall, mediating the organism's interactions with its environment and maintaining the structural integrity of the cell.[1] These glycan structures can be incredibly diverse between species, but are generally synthesized by the stepwise addition of sugars onto a biomolecule through the action of glycosyltransferase (GT) enzymes.[2,3] The activated sugar substrates used by GTs are either sugar nucleotides or glycosylphospholipids.[2,3]

Several pathogenic organisms, such as *Mycobacterium tuberculosis* and *Campylobacter jejuni*, have been found to possess cell wall structures containing sugars in the five-membered (furanose) ring form.[4,5] The complete absence of furanose sugars in mammals suggests that the furanose-processing enzymes used by bacteria could be potential drug targets.[4b] Therefore, access to the natural furanose sugar nucleotides used by GTs, as well as their analogs, would facilitate biochemical studies of these enzymes and potentially lead to new therapies to treat diseases such as tuberculosis and campylobacteriosis.

## 1.2 Tuberculosis

Tuberculosis (TB) is a widespread infectious disease that has plagued humanity since ancient times.[6] Primarily a disease of the lungs, TB is spread from infected to uninfected individuals through aerosol transfer,[7] a process that is exacerbated in crowded and/or unhygienic living conditions. *Mycobacterium*

*tuberculosis* was identified as the causative agent of TB by Robert Koch in 1882,[8] and despite significant advances in treatment since that time the disease continues to be problematic around the world (Figure 1-1). The World Health Organization (WHO) estimates that one third of the world's population is infected with TB, with over 8 million new infections occurring each year.[9]

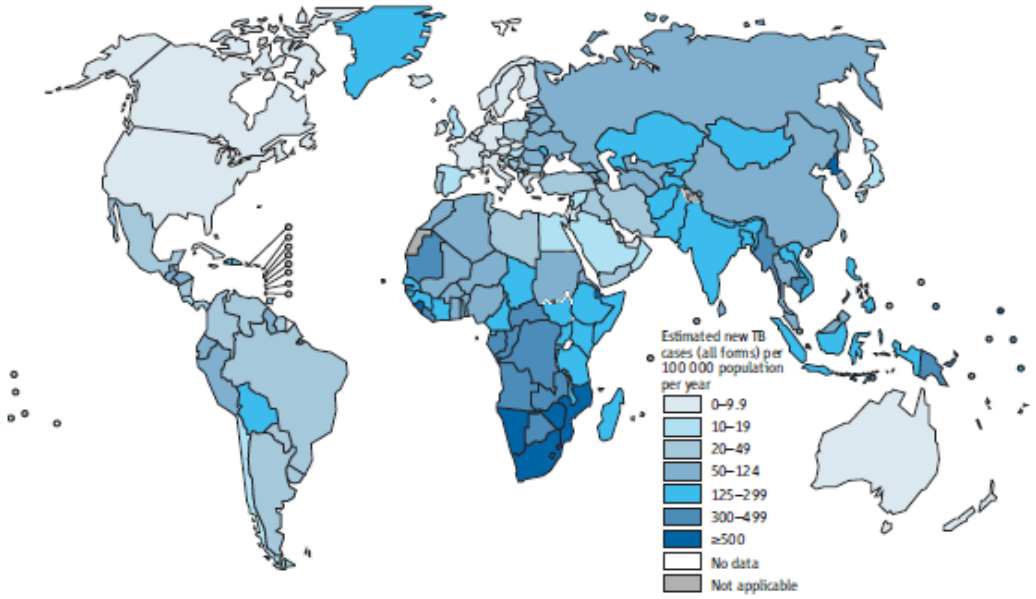



**Figure 1-1.** Estimated incidence of new TB cases in 2012. Source: WHO Global Tuberculosis Report 2013.[9]

There are two types of TB infection possible: Latent and active.[7] In a latent infection, living *M. tuberculosis* bacteria are present in the body, but the immune system is able to prevent the bacteria from multiplying excessively. Individuals with latent TB are not infectious. However, a latent TB infection may progress to an active infection, at which time the individual develops symptoms characteristic of TB (bad cough lasting longer than three weeks, pain in the chest, coughing up blood or sputum, weakness or feeling very tired, weight loss, lack of

appetite, chills, fever, and night sweats) and may require hospitalization. Individuals with latent infections are most at risk for developing an active infection within two years of the first infection.[9]

Treatment of TB involves a prolonged regimen of antibiotic drugs. The typical course of treatment for the most common forms of TB is a cocktail of isoniazid, a rifamycin (often rifampicin), pyrazinamide, and ethambutol for two months, followed by the continued use of isoniazid and a rifamycin for an additional four months (Figure 1-2).[9,10] Isoniazid and ethambutol inhibit the formation of the protective bacterial cell wall, while rifamycins inhibit RNA synthesis by binding to DNA-dependent RNA polymerase.[11] Although there has been some debate regarding the mechanism of action of pyrazinamide, a recent study suggests that it may inhibit the *trans*-translation of RNA.[12]

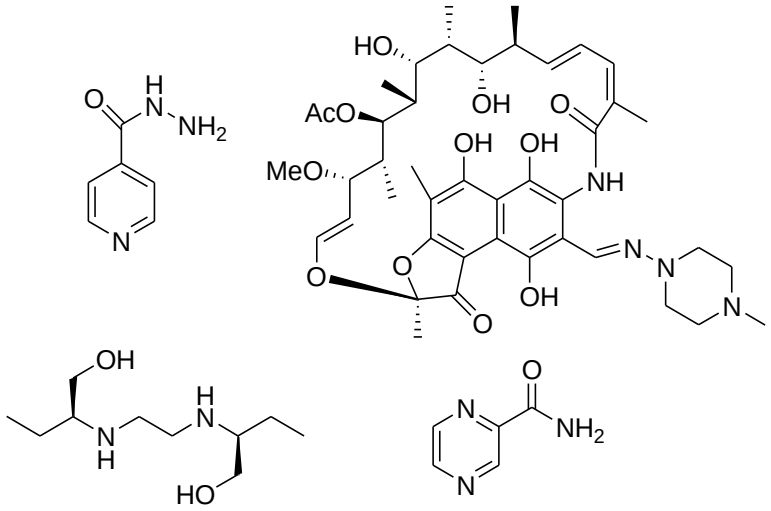



**Figure 1-2.** Front line drugs used to treat TB. Clockwise from top left: Isoniazid, rifampicin, pyrazinamide, and ethambutol.

While 95% of infected patients with drug-susceptible TB can be cured using the treatment regimen described above, the rise of multi-drug resistant TB (MDR-TB) and extensively drug resistant TB (XDR-TB) has raised concern in recent decades.[9,10] These strains can arise from antibiotic selection pressure in improperly managed patients with drug-susceptible TB, or can be transmitted from infected to non-infected individuals. MDR-TB is defined as TB that is resistant to two or more of the front-line antibiotics, while XDR-TB is also resistant to several second-line treatments (Figure 1-3), specifically all fluoroquinolones and one of capreomycin, kanamycin, or amikacin.[10,13] Treatment regimens for MDR/XDR-TB can last up to two years, and have lower cure rates than for drug-susceptible TB (50–70% for MDR-TB, and 30–50% for XDR-TB).[9,10] MDR/XDR-TB is a global concern (Figure 1-4), and within the last decade TB strains that are resistant to all currently used TB antibiotics have emerged.[14]

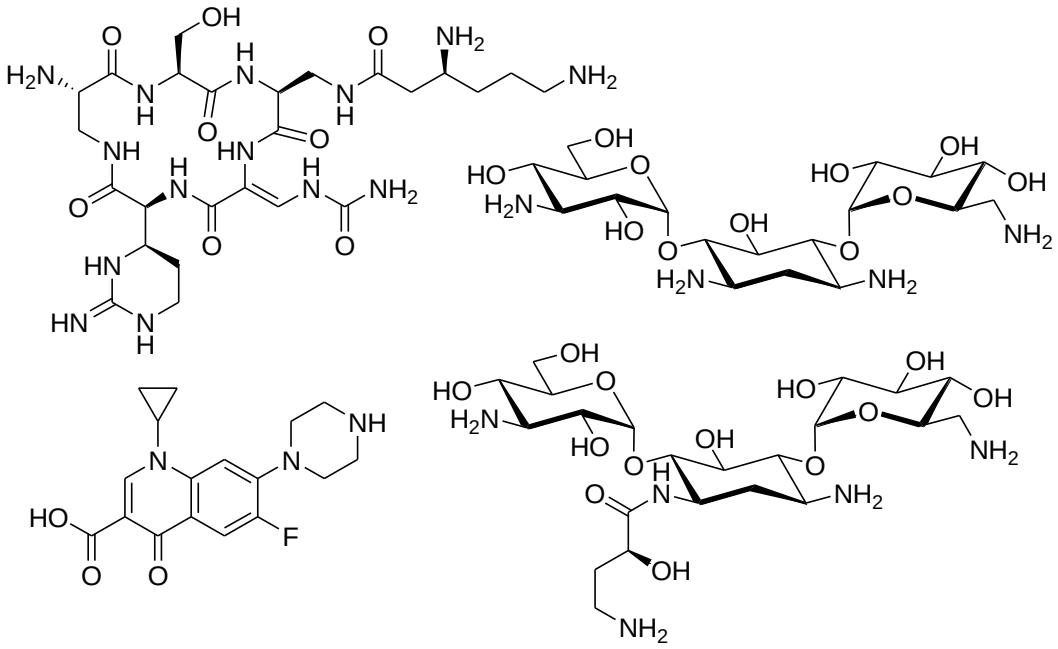


**Figure 1-3.** Second line drugs for the treatment of resistant strains of TB. Clockwise from top left: Capreomycin, kanamycin, amikacin, and ciprofloxacin.

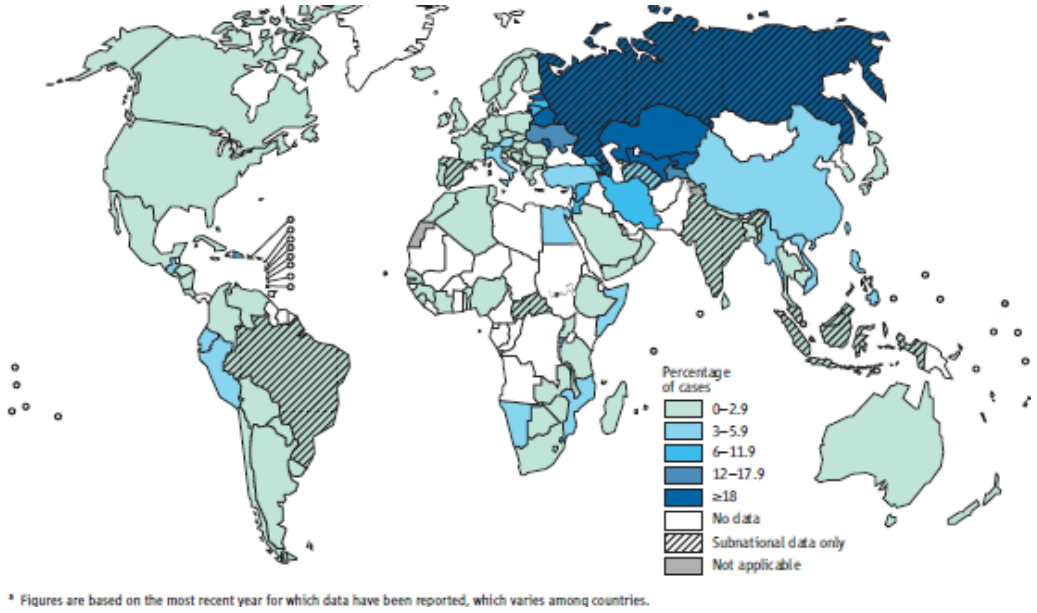



**Figure 1-4.** Percentage of new TB cases with MDR-TB. Source: WHO Global Tuberculosis Report 2013.[9]

### 1.3 Campylobacteriosis

*Campylobacter jejuni* is a zoonotic motile pathogen that is responsible for nearly half a billion cases of bacterial gastroenteritis (campylobacteriosis) worldwide each year.[15] *C. jejuni* is most often contracted from raw or undercooked poultry, and infection results in food poisoning-like symptoms such as abdominal pain, fever, nausea, and diarrhea.[16] In a small number of individuals, *C. jejuni* infection can lead to symptoms of Guillain–Barré syndrome (GBS), a neuropathic disease that can lead to paralysis. Although GBS is quite rare, 14–32% of cases can be traced back to *C. jejuni* infection.[17]

Disease symptoms begin to appear two to five days after infection, and typically last for three to six days. During this time, treatment most often consists of symptom alleviation and maintaining hydration and electrolyte balance against fluid loss, as the majority of infections are self-limiting. Severe, prolonged, or systemic infections may be treated with antibiotics such as macrolides (e.g. azithromycin and erythromycin) and fluoroquinolones (e.g. ciprofloxacin),[18] although resistance to both these classes of drugs has been reported.[15,18,19] Prevention of *C. jejuni* infection requires safe food handling practices at all levels of the production chain. Despite its often serious symptoms, campylobacteriosis is rarely fatal, with the most at-risk populations being the elderly, very young children, and individuals with compromised immune systems.[15] Disease management of campylobacteriosis is still of the utmost importance, as the sheer number of infections has a significant socio-economic impact.[17]

## 1.4 Cell Wall Glycans in *Mycobacterium tuberculosis* and *Campylobacter jejuni*

Both *M. tuberculosis* and *C. jejuni* contain a large number of glycans within their cell walls, structures that have been shown to be essential to the virulence of these pathogens.[5b,20] Both organisms also feature sugars in a five-membered ring (furanose) form that is not present in mammals.[4,5] For these reasons, the bacterial enzymes responsible for synthesizing these cell walls have been proposed as potential drug targets.[4b]

### 1.4.1 Structure of *M. tuberculosis* Cell Wall Glycans

The innermost carbohydrate-containing layer of the *M. tuberculosis* cell wall is the peptidoglycan (PG), consisting of alternating units of *N*-acetylglucosamine (GlcNAc) and *N*-glycolylmuramic acid in a β-(1→4) linkage (Figure 1-5, Figure 1-6).[21,22] A tetrapeptide side chain, L-alanyl-D-isoglutaminyl-*meso*-diaminopimelyl-D-alanine, is covalently attached to the *N*-glycolylmuramic acid moiety, and tetrapeptides on adjacent PG strands can cross-link to increase the strength of this layer. Although the PG glycan layer has traditionally been presented as parallel to the plasma membrane, recent NMR studies have suggested that these glycans may be coiled perpendicular to the membrane plane (Figure 1-5).[23]

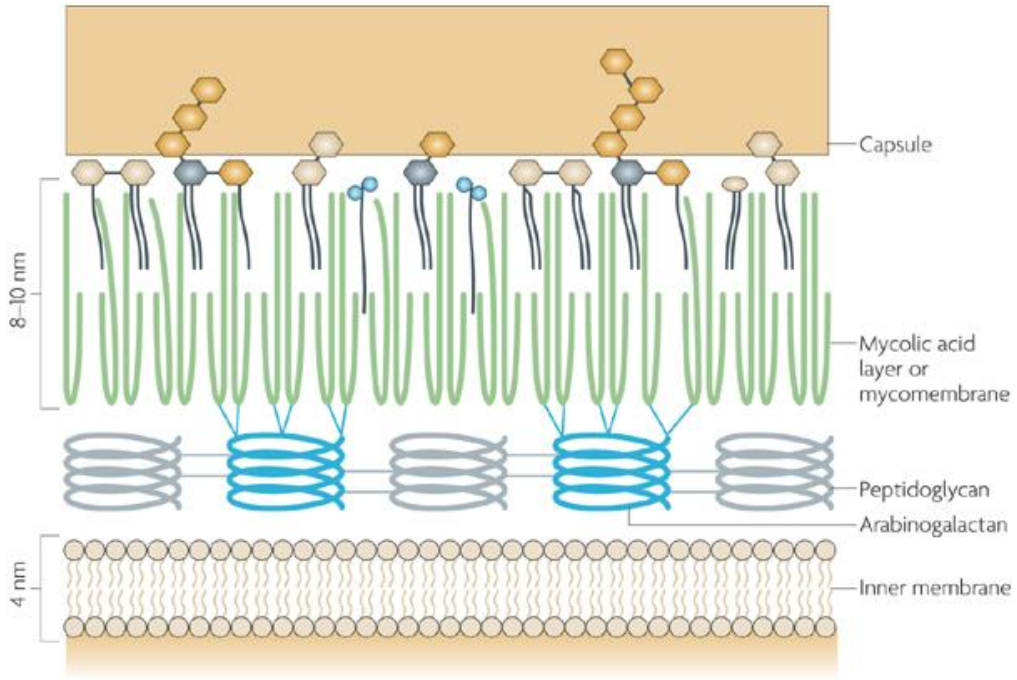



**Figure 1-5.** Structure of the *M. tuberculosis* cell wall. Reprinted by permission from Macmillan Publishers Ltd: *Nat. Rev. Microbiol*., **5**, 883-891, copyright 2007.

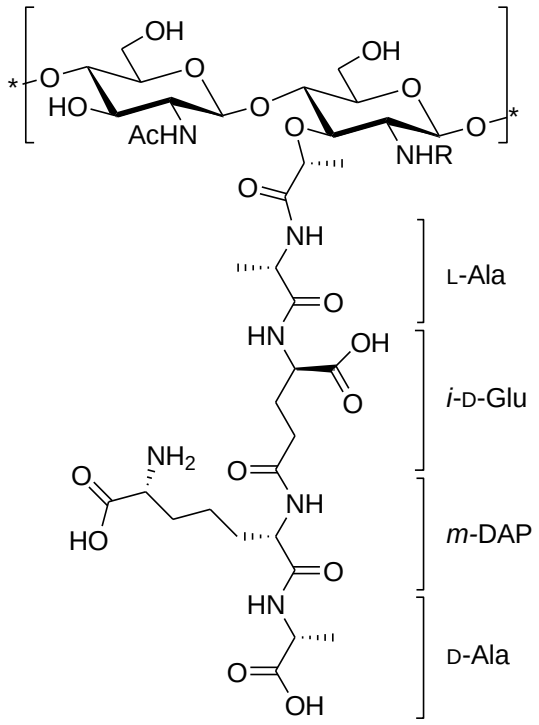



**Figure 1-6.** Repeating unit of *M. tuberculosis* PG. Crosslinking occurs between the *m*-DAP or terminal D-Ala units. R = acetyl or glycolyl.

The PG layer is linked to what is known as the mycolyl–arabinogalactan (mAG) layer, the bulkiest component of the *M. tuberculosis* cell wall. Its structure has been thoroughly reviewed.[4a,22,24] The galactan, consisting of approximately 30–35 β-D-galactofuranose (Gal*f*) residues linked in an alternating (1→5), (1→6) fashion, is attached to C-6 of the *N*-glycolylmuramic acid via a rhamnopyranose (Rha*p*)-*N*-acetylglucosamine disaccharide-phosphate bridge, (→4)-α-L-Rha*p*-(1→3)-α-D-Glc*p*NAc-(1→P). The arabinan portion of the mAG consists of branched chains of 31 arabinofuranose residues, with three arabinans attached to each galactan at positions 8, 10, and 12 of O-5 of (1→6) linked Gal*f* residues. Each arabinan consists primarily of α-(1→5) linked arabinofuranose (Ara*f*) residues, capped with a six-Ara*f* motif, β-Ara*f*-(1→2)-α-Ara*f*-(1→5)-[β-Ara*f*-

(1→2)-α-Ara*f*-(1→3)]-α-Ara*f*-(1→5)-α-Ara*f*-(1→), at the non-reducing ends. Mycolic acids, $C_{70}$–$C_{90}$ α-alkyl-β-hydroxy fatty acids, are anchored to position 5 of the last and second-to-last Ara*f* residues through an ester bond on about two-thirds of the Ara*f* hexasaccharides. The mycolic acids comprise the outermost layer of the cell wall, forming an impenetrable hydrophobic barrier that resists Gram-staining but is required for virulence.[25]

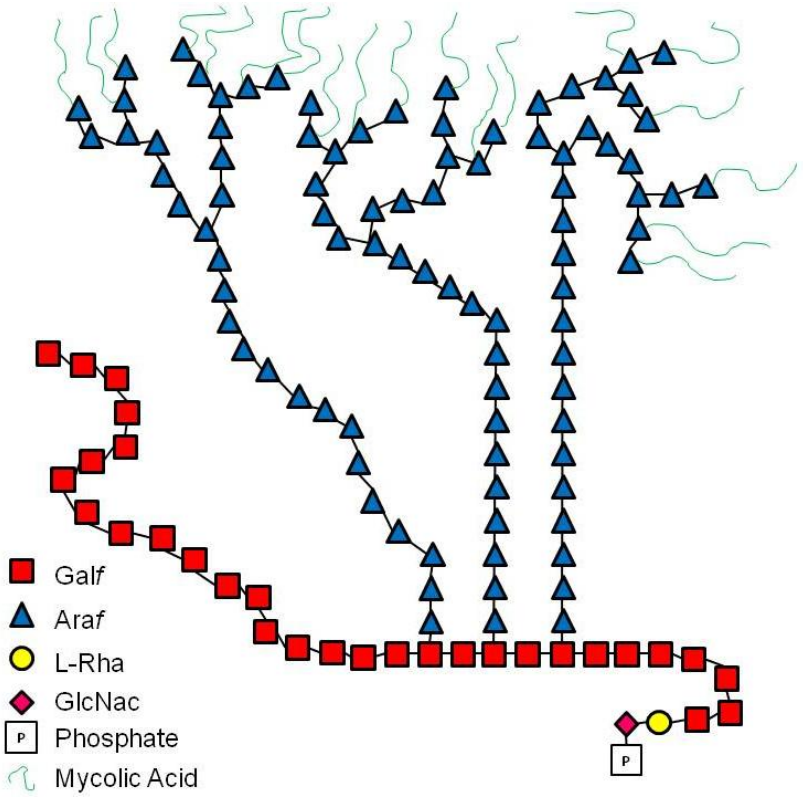



**Figure 1-7.** Stylized representation of mAG complex.

Within the mAG complex lie lipomannan (LM) and lipoarabinomannan (LAM), two structures that are found within the cell walls of many species of *Mycobacterium*.[22] Both structures feature a phosphatidyl-*myo*-inositol core,

responsible for anchoring these structures to cell membranes through non-covalent hydrophobic interactions, with a mannan extending from C-6 of the *myo*-inositol. Approximately 20–25 mannopyranose residues are linked together in an α-(1→6) fashion to form this mannan, with occasional C-2 substitution of one mannopyranose residue. In LAM, approximately 60 α-(1→5) Ara*f* residues are attached to the mannan core in structures that display some degree of branching, and are structurally similar to mAG arabinan.[22]

### 1.4.2 Structure of *C. jejuni* Cell Wall Glycans

The cell wall of *C. jejuni* consists of two major parts (Figure 1-8). Both the capsular polysaccharide (CPS) and the lipooligosaccharide (LOS) core structures are highly variable between different strains, and can undergo further phase-variable surface modifications; for example, the incorporation of methyl, ethanolamine, aminoglycerol, and methyl phosphoramidate moieties onto CPS sugars have all been reported.[5b,26] It has been suggested that such modifications may enable the bacterium to elude host defence mechanisms.[26b,c] The degree of hypervariability in the *C. jejuni* genome became evident after its first sequencing in 2000.[27,28]

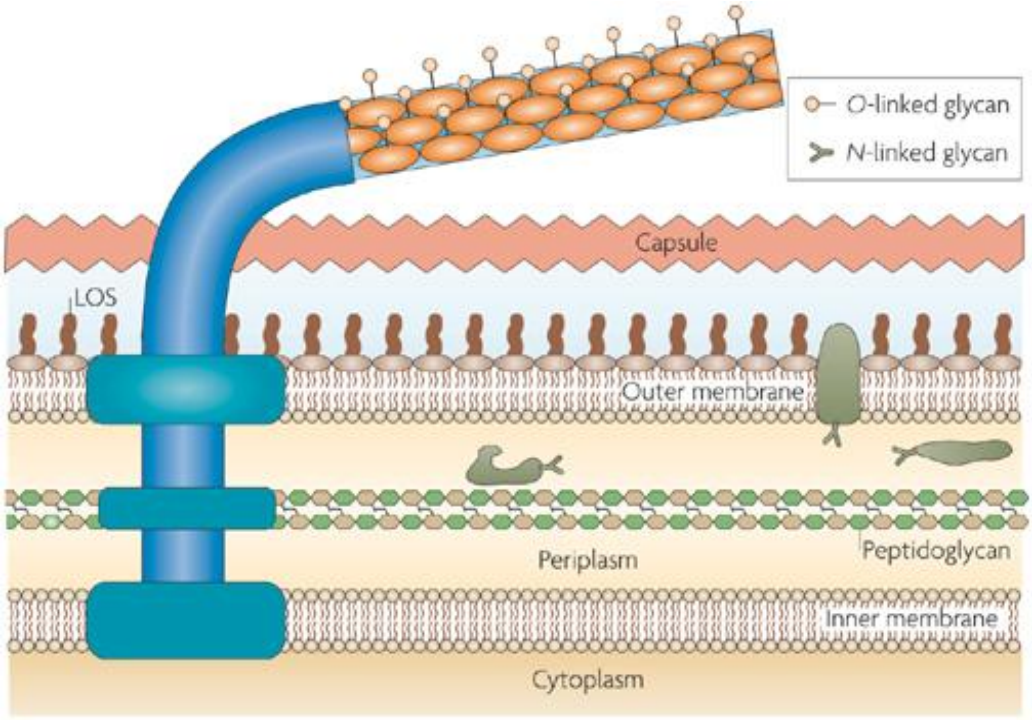
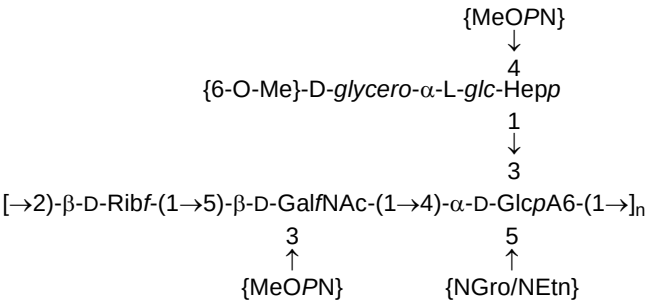



**Figure 1-8.** Structure of the *C. jejuni* cell wall. Reprinted by permission from Macmillan Publishers Ltd: *Nat. Rev. Microbiol.* **5**, 665-679, copyright 2007.

The CPS of *C. jejuni* contains a number of rather exotic sugars. Like *M. tuberculosis*, some strains contain sugars in the furanose form.[5] For example, the NCTC11168 (HS:2) strain contains a 2-acetamido-2-deoxy-α-D-galactofuranose (Gal*f*NAc) residue within its tetrasaccharide repeating unit, with O-3 being variably modified with a methyl phosphoramidate (Figure 1-9).[5b] Further, every CPS sugar of serostrain HS:41 (β-L-arabinose, 6-deoxy-β-D-*altro*heptose, 6-deoxy-β-L-altrose, and α-D-fucose) is found in the furanose form (Figure 1-9).[29]

**HS:2**




**HS:3**

[→3)-L-*glycero*-α-D-*ido*-Hep*p*-(1→4)-α-D-Gal*p*-(1→]$_n$

**HS:41**

[-2)-β-L-Ara*f*-(1→2)-β-D-6d-*altro*-Hep*f*-(1→2)-β-L-6d-Alt*f*-1-]$_n$

-α-D-Fuc*f*-1-]$_n$

**HS:53**

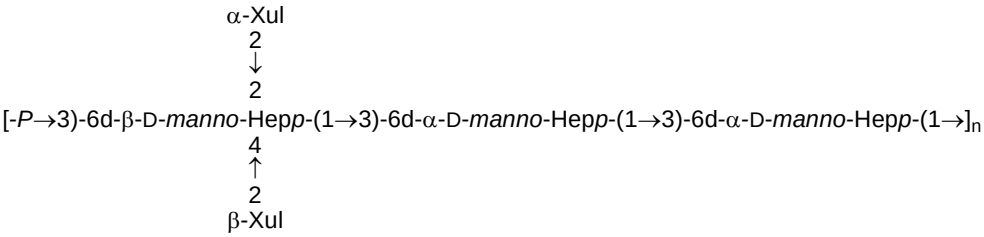


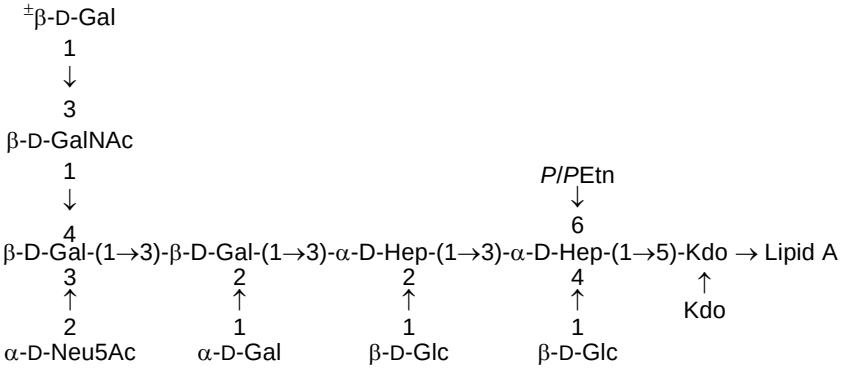

**Figure 1-9.** Representative repeating CPS units from several *C. jejuni* strains. Serostrains are indicated in bold. Hep, heptose; Rib, Ribose; GalNAc, *N*-acetylgalactosamine; Gal, galactose; Ara, arabinose; Alt, altrose; Fuc, fucose; Xul, xylulose; GlcA6, glucuronic acid; MeOPN, O-methyl phosphoramidate; NGro, aminoglycerol; NEtn, ethanolamine. Pyranose and furanose configurations denoted by *p* and *f*, respectively. Figure adapted from references 29 and 30**.**

In addition to furanoses, a number of unusual heptose sugars are present in several *C. jejuni* strains. NCTC11168 (HS:2) features a 6-*O*-methyl-D-*glycero*-α-L-*gluco*-heptopyranose moiety within its CPS tetrasaccharide,[5b,29] with some mutants containing a methyl phosphoramidate at O-3 and methylation at O-4

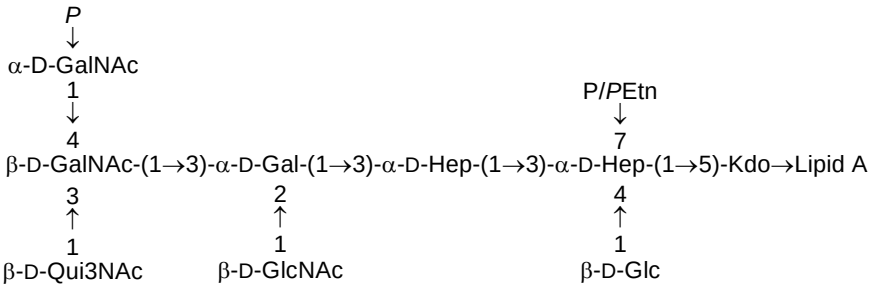
(Figure 1-9). Strain HS:41 contains 6-deoxy-β-D-*altro*-heptofuranose, a heptose in a furanose configuration,[29] while HS:3 and HS:53 contain L-*glycero*-α-D-*ido*-heptopyranose and 6-deoxy-α-D-*manno*-heptopyranose, respectively.[30] Other examples of *C. jejuni* CPS of varying complexity have been described.[29,30]

Unlike CPS, LOS repeating units are not known to contain any sugars in the furanose form, although heptose sugars abound (Figure 1-10). These repeating units most often feature galactose, *N*-acetylgalactosamine, 3-deoxy-D-*manno*-oct-2-ulosonic acid (Kdo), L-*glycero*-D-*manno*-heptose, and occasionally glucose in the core structure.[26a,28,31] This core is modified with phosphates and/or phosphoethanolamine and decorated with sialic acids (Neu5Ac). The presence of sialic acid residues allows many *C. jejuni* LOSs to mimic the oligosaccharides found on human gangliosides, in an attempt to evade detection by the host cell.[28] This strategy often has severe consequences for the human host, as the body will manufacture antibodies against both *C. jejuni* and its own nerve cells. Macrophages then begin destroying both types of cells, and the resultant nerve damage leads to the GBS discussed previously.[28] Not all *C. jejuni* LOS structures are ganglioside mimics; these structures may contain quinovosamine (*N*-acetyl 3-amino-3,6-dideoxyglucosamine).[31]

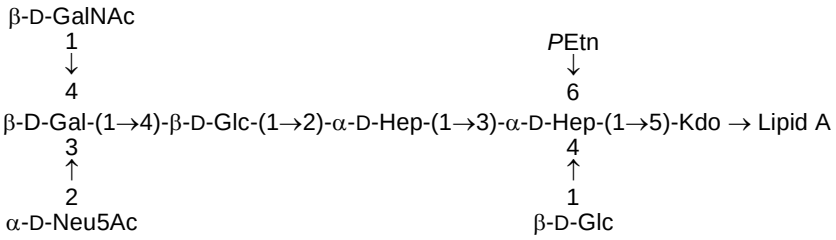
**HS:2**




**HS:3**




**HS:23**




**HS:41**

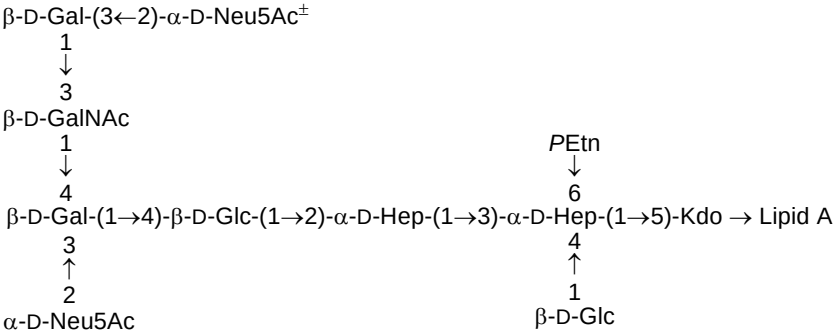
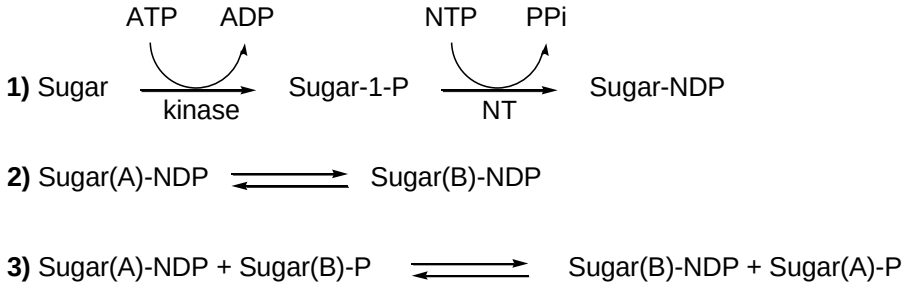


**Figure 1-10.** (Preceding page) LOS structures found in various *C. jejuni* species. Gal, galactose; GalNac, *N*-acetylgalactosamine; Hep, L-*glycero*-D-*manno*-heptose; Kdo, 3-deoxy-D-*manno*-oct-2-ulosonic acid; Neu5Ac, *N*-acetylneuraminic acid; Glc, glucose; Qui3NAc, *N*-acetylquinovosamine; GlcNac, *N*-acetylgalactosamine; *P*, phosphate; *P*Etn, phosphoethanolamine. All sugars are in the pyranose configuration. Variable residues denoted by ±. Figure adapted from references 26a and 31.

### 1.4.3 Biosynthesis of Cell Wall Glycans

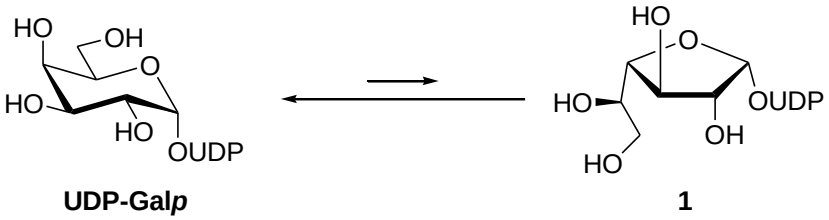
#### 1.4.3.1 Biosynthesis of Sugar Nucleotides

Sugar nucleotides are the most common form of sugar donor for glycosylation reactions *in vivo*. The *de novo* synthesis of sugar nucleotides in living systems begins with glycolysis intermediates;[32] kinase phosphorylation forms sugar-1-phosphates, which are then converted into sugar nucleotide donors by nucleotidyltransferase enzymes (Scheme 1-1, Reaction 1).[38] Glycosyl phospholipid donors are formed by the attack of a phospholipid onto a sugar nucleotide.




**Scheme 1-1.** Biosynthesis of sugar nucleotides. Scheme adapted from reference 38.

In addition to *de novo* synthesis, sugar nucleotides in biological systems can also be obtained through the interconversion of structural forms (Scheme 1-1, Reaction 2) or through nucleotide transfer (Scheme 1-1, Reaction 3). For example, UDP-galactopyranose mutase (UGM) is an enzyme found in all bacteria, enzymes, and fungi that contain Gal*f* within their glycoconjugates.[33] UGM catalyzes the ring contraction interconverting UDP-Gal*p* into UDP-Gal*f*, the donor species required to synthesize Gal*f*-containing glycoconjugates (Scheme 1-2). Alternatively, nucleotidyltransferases can transfer the nucleotide moiety from a sugar nucleotide onto a structurally different sugar-1-phosphate, generating a new sugar nucleotide (e.g. galactose-1-phosphate uridyltransferase, GalPUT) (Scheme 1-1, Reaction 3).[34]




**UDP-Gal*p*** **1**

**Scheme 1-2.** Ring conversion catalyzed by UGM.

In order to obtain quantities of natural and non-natural sugar nucleotides that are sufficient for biological experiments, a number of robust methods have been developed. The following sections outline some recent approaches to the production of sugar nucleotides, both chemically[35] and chemoenzymatically.[32,36]

### 1.4.3.2 Structure and Mechanism of Glycosyltransferases

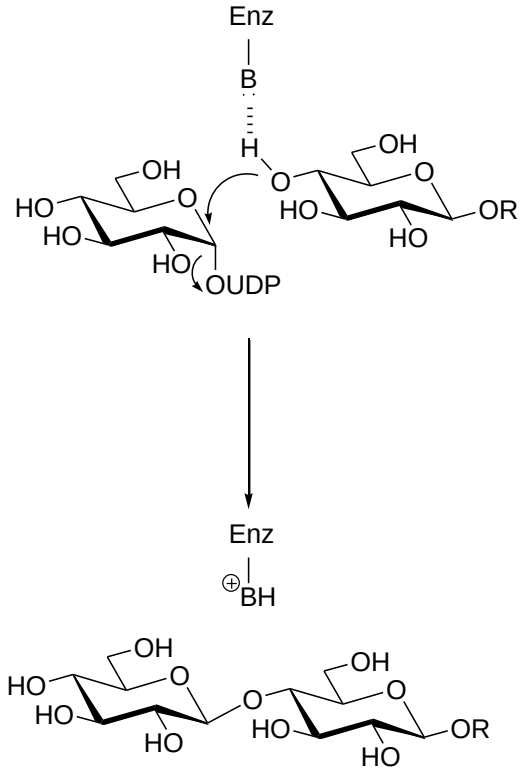
Glycans present in all living organisms are metabolized through the action of glycosyltransferase enzymes (GTs), responsible for the addition of individual sugars to other biomolecules,[2,3] and glycosidases, which hydrolyze glycosidic bonds.[37] Sugar donors used by GTs must be pre-activated for addition to a nucleophilic acceptor, which may be another carbohydrate or a protein, lipid, nucleic acid, or other small molecule. Donors are most often activated as sugar nucleotides, and can be either nucleotide diphosphates (NDPs) or nucleotide monophosphates (NMPs).[38] Two other types of activated donors, glycosylphospholipids and sugar-1-phosphates, are also known.[3] GTs that use sugar nucleotide donors are known as Leloir enzymes.

Two types of folds have been observed experimentally in GTs, allowing these proteins to be classified accordingly.[2,3] The first group, dubbed the GT-A family, possesses two β/α/β Rossman-type domains that lie very close to each other, resulting in a continuous central β-sheet. GT-A enzymes also contain divalent metal ions binding to an Asp–X–Asp motif. When binding sugar nucleotide substrates, the metal ion binds to the pyrophosphate moiety and makes it a more effective leaving group. Some examples of GT-A proteins that do not possess an Asp–X–Asp have come to light, and not all proteins that possess the Asp–X–Asp motif are GT-A enzymes.[3]

Another group of experimentally observed enzymes is known as the GT-B family.[2,3] This family also possesses two β/α/β Rossman-type domains, but these domains are not as tightly associated and face each other, leaving a binding

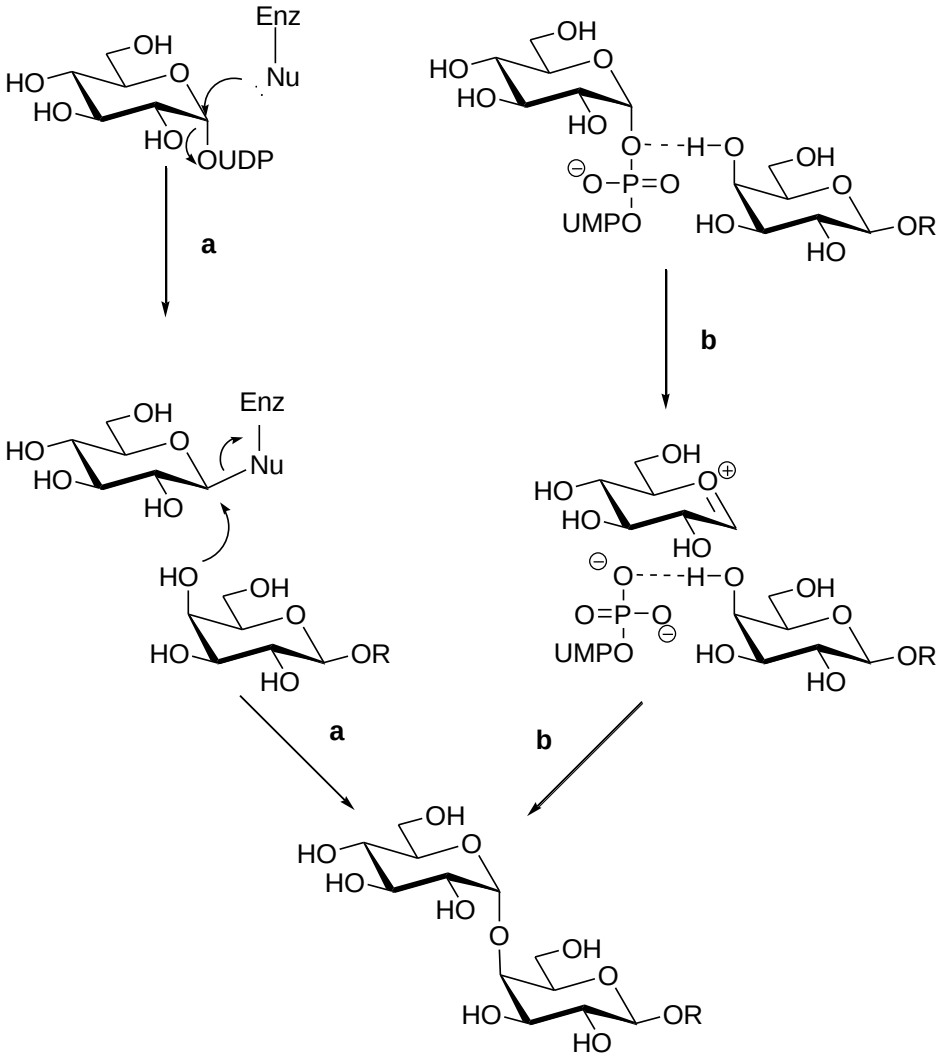
pocket between the two domains. One further type of enzyme fold, the GT-C family, has not been experimentally verified but is predicted to exist.[39] Putative enzymes containing this type of fold are believed to be transmembrane proteins that use sugar phospholipids as substrates.

GTs may also be classified based upon the mechanism of the enzyme. Sugar donors may be added with either retention or inversion of their initial anomeric configuration, allowing GTs to be classified as retaining or inverting.[2,3] The mechanism of inverting GTs is relatively straightforward (Scheme 1-3): The nucleophilic acceptor attacks the anomeric center of the sugar donor, displacing the leaving group in an $S_N2$-type of mechanism. A catalytic residue within the active site acts as a base to deprotonate the acceptor, leading to a transition state that has a large amount of oxocarbenium character on the donor.




**Scheme 1-3.** Mechanism of inverting glycosyltransferases.

The mechanism used by retaining GTs has been more controversial.[2,3,40,41] One proposal, termed the "double displacement" mechanism, creates a new glycosidic bond in two steps.[2,3,41] First, a nucleophilic residue within the enzyme active site attacks the sugar donor, creating a new covalent bond between the enzyme and the donor and inverting the initial anomeric configuration of the donor (Scheme 1-4a). This configuration returns to its initial stereochemistry upon attack of the acceptor, replacing the enzyme-donor bond with an acceptor-donor bond. Alternatively, an "$S_Ni$" mechanism has also been proposed,[2,3] where the acceptor and the leaving group form a four-centred transition state on the same face of the donor. Further, an $S_Ni$-like stepwise mechanism involving ion pair formation has also been put forward (Scheme 1-4b).[3,40] $S_Ni$ and $S_Ni$-type mechanisms were proposed to explain why no covalent intermediate could be isolated to support the double displacement mechanism, and both mechanisms are believed to have significant oxocarbenium character.[40] However, evidence for a covalent intermediate in GT mutants has been found through mass spectrometry.[42]



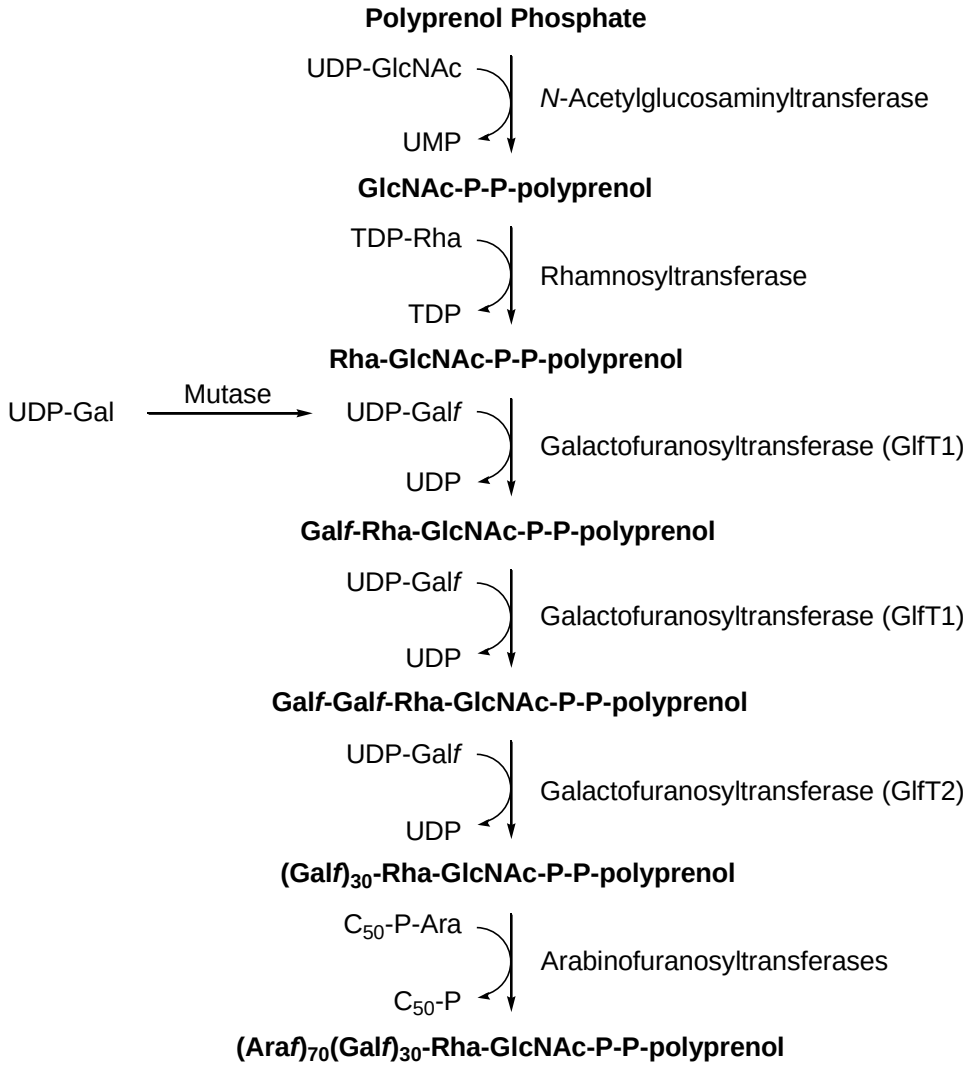

**Scheme 1-4.** Proposed mechanisms of retaining GTs. (a) Double displacement mechanism; and (b) $S_Ni$-like mechanism.

### 1.4.3.3 Biosynthesis of Cell Wall Glycans in *M. tuberculosis* and *C. jejuni*

Cell wall glycan biosynthesis is a complex, energy-intensive undertaking for microbes, emphasizing the importance of this layer to the viability of the organism. The construction of glycans and glycoconjugates from activated donors

is catalyzed by GTs,[2,3] and the resulting products are exported across the cell membrane for incorporation into the cell wall.

A description of arabinogalactan biosynthesis in *M. tuberculosis* will be used to illustrate this complex process.[4a,20,22,24a,43] The biosynthesis of the arabinogalactan layer of the mAG complex is initiated by the condensation of UDP-GlcNAc and polyprenol phosphate, followed by the addition of rhamnose from dTDP-Rha*p* in an α-(1→3) fashion. UDP-Gal*f*, formed from UDP-Gal*p* through the action of UDP-Gal*p* mutase, is then added to the Rha-GlcNAc-P-P-polyprenol substrate through the action of two GTs. The first, GlfT1, is responsible for the addition of the first and second Gal*f* residues, while GlfT2 adds the third and subsequent residues to form a galactan chain that is 30–35 residues in length. The only Ara*f* donor known in mycobacteria, polyprenolmonophosphoryl-β-D-arabinose, is used by various GTs to attach the first Ara*f* residue and extend the arabinan with its various branched structures. This arabinogalactan structure can then be transferred from its polyprenol carrier onto the PG.



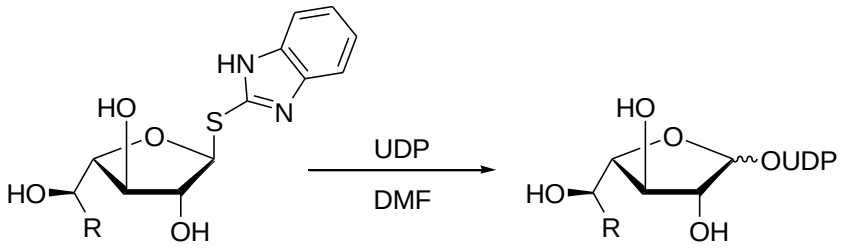

**Scheme 1-5.** Biosynthesis of mAG complex in *M. tuberculosis*. Adapted from reference 22.

## 1.5 Methods of Producing Sugar Nucleotides

### 1.5.1 Chemical Synthesis of Sugar-1-Phosphates

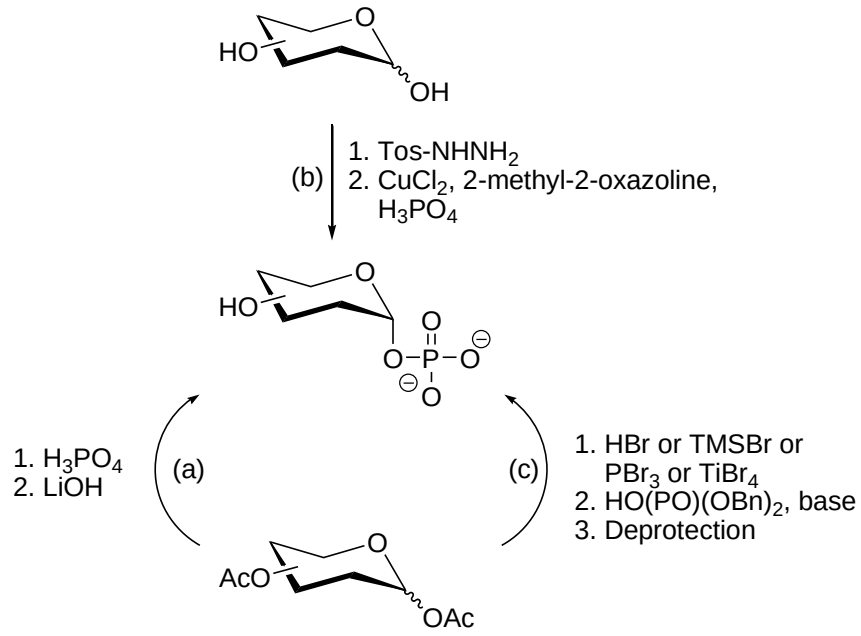
The chemical synthesis of sugar nucleotides via sugar-1-phosphates has proven to be the method of choice for most researchers. Although the glycosylation of sugar donors with a nucleotide acceptor has been accomplished

(Scheme 1-6),[44] the major drawback of this method is the production of α/β anomers that complicate the isolation of the pure sugar nucleotide products at a late stage of the synthesis. The production of sugar-1-phosphates, on the other hand, allows the anomers to be separated earlier, at a stage where the compounds are more stable. Subsequent formation of a pyrophosphate bond can then proceed without complications from anomeric mixtures.




**Scheme 1-6.** Glycosylation of a furanose thioimidate, forming an anomeric mixture of UDP–furanoses. R = H, $CH_2OH$, $CH_2F$, $CH_3$.[44]

One of the earliest methods of phosphorylation is the MacDonald reaction, where peracetylated sugars are reacted with phosphoric acid, followed by deprotection with LiOH to yield α-sugar-1-phosphates in 31–35% yield (Scheme 1-7a).[45] Later, Hanessian developed an α-selective method using phosphoric acid and methoxypyridine donors that proceeded without protecting groups;[46] Nitz later extended this method using a *p*-toluenesulfonyl hydrazine intermediate (Scheme 1-7b).[47] Although the MacDonald reaction is attractively simple, the harsh conditions employed by these methods make them incompatible with many modified sugar precursors.

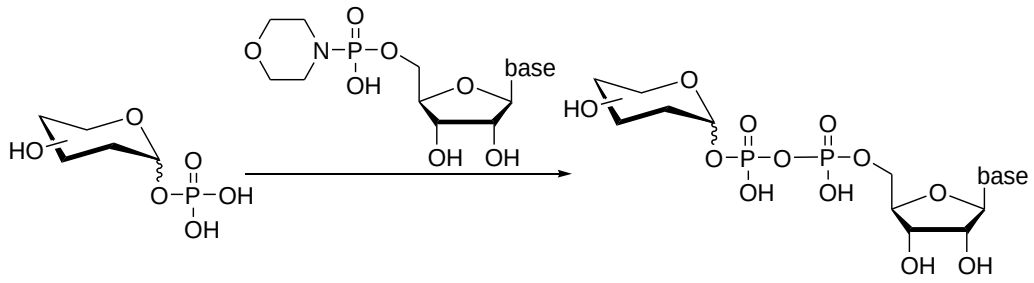



**Scheme 1-7.** Selected methods of synthesizing sugar-1-phosphates. (a) MacDonald reaction, (b) phosphorylation of unprotected sugars, and (c) phosphorylation via a bromide intermediate.

The most common approach, compatible with a larger number of substrates, is to phosphorylate an acylated sugar via a bromide intermediate (Scheme 1-7c).[48] Common brominating agents include HBr, TMSBr, and $PBr_3$, while $TiBr_4$ has been used for sugars containing azide substituents.[49] A nucleophilic phosphorus reagent, commonly dibenzyl phosphate, is added with a non-nucleophilic base to displace the bromide.[48] Deprotection of the resulting sugar then leads to the desired α-sugar-1-phosphate. Glycosyl halides have also been used to form β-sugar-1-phosphates;[50,51] these sugar-1-phosphates are not stabilized by an anomeric effect, and are therefore more difficult to prepare. Other methods of producing sugar-1-phosphates, including the use of P(III) reagents followed by oxidation to P(V), have been reviewed.[35]

### 1.5.2 Chemical Synthesis of Sugar Nucleotides from Sugar-1-Phosphates

The earliest method of coupling a sugar-1-phosphate with a nucleotide to form a pyrophosphate bond was developed by Khorana and co-workers in the late 1950s.[52] In this method, a NMP is activated as a phosphoramidate (later as a phosphoromorpholidate[53]) for coupling with the sugar-1-phosphate (Scheme 1-8). This procedure and its variants were used for decades to synthesize sugar nucleotides. Despite relatively high yields, the Khorana procedure suffers from lengthy reaction times; however, catalyzed versions using either tetrazole[54] or $Mn^{2+}$ (Ref. 55) have overcome this problem somewhat. Tetrazole, it should be noted, is explosive, while $Mn^{2+}$ catalysis appears to be limited to purine nucleobases.[35]

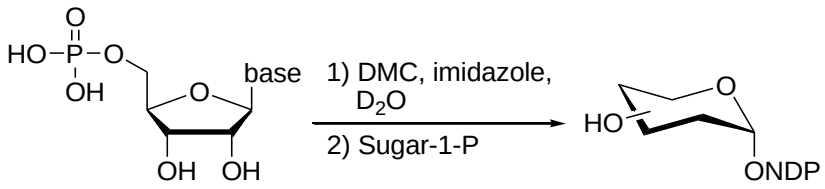



**Scheme 1-8.** Synthesis of NDP-sugars using the Khorana protocol. Catalyzed variants using $Mn^{2+}$ or tetrazole are also known. Scheme adapted from reference 35.

Nucleotide monophosphates can also be activated with *N.N*′-carbonyldiimidazole (CDI).[56] This gentle alternative to the Khorana method typically proceeds on the order of 1–2 days with moderate yields for both purine and pyrimidine nucleobases. A related method, introduced by Bogachev for the

synthesis of deoxynucleoside triphosphates,[57] activates (and protects) the NMP sugars with trifluoroacetic anhydride, followed by conversion into the active phosphoroimidazolide species with *N*-methylimidazole. Marlow and Kiessling used this method to synthesize UDP-Gal*f* in 35% yield,[58] while Timmons and Jakeman used it to synthesize a number of UDP- and ADP-sugars.[51]

One further development, recently published by Hindsgaul and co-workers, has found conditions for pyrophosphate bond formation under aqueous conditions (Scheme 1-9).[59] In this method, imidazole reacts with 2-chloro-1,3-dimethylimidazolinium chloride (DMC) to form a diimidazole reagent in $D_2O$, which goes on to further react with an NMP to form an NMP-imidazolide. Subsequent addition of a sugar-1-phosphate leads to formation of the desired sugar nucleotide. Although the product yields from this reaction were moderate and significant dimerization of the NMP occurred, the operational simplicity of this method makes it attractive to pursue.




**Scheme 1-9.** Sugar nucleotide synthesis developed by Tanaka *et al.*[59]

### 1.5.3 Chemoenzymatic Synthesis of Sugar Nucleotides

Due to the low yields observed during the chemical syntheses of many sugar nucleotides, enzymatic[32] and chemoenzymatic[36] approaches have become

popular in recent years. Three types of enzymes, nucleotidyltransferases, glycosyltransferases, and pyranose–furanose mutases have risen to prominence in this area.

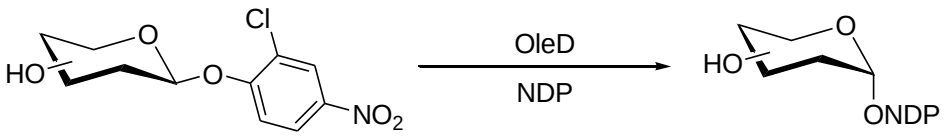
Nucleotidyltransferases (NTs), as discussed above, are a natural choice of enzyme toward the synthesis of sugar nucleotides, and examples of their use in sugar nucleotide synthesis are widespread.[32,36] Wild-type enzymes can be used to catalyze the transfer of a nucleotide unit from a NTP onto a sugar-1-phosphate that has been obtained either chemically or enzymatically. Often, the innate promiscuity of some wild-type NTs will catalyze the transfer of non-natural sugars and/or nucleotides,[34,60] although the efficiency of these transfers may be compromised. Active site engineering has been carried out to improve upon the substrate flexibility displayed by some NTs, improving the catalytic efficiency by several orders of magnitude.[61,62]

Initially, glycosyltransferases (GTs) may seem like unusual candidates for the synthesis of sugar nucleotides, as in the forward direction GTs break down sugar nucleotides to form glycosylated products.[2,3] However, the observation that some GTs catalyze a reversible transfer has opened up a new avenue for exploration.[63] The initial reports of reversible GTs allowed for sugar exchange between various glycosylated products, but provided only a fleeting glimpse of sugar nucleotide products due to the inherent thermodynamic bias against their formation.

To overcome such a bias, activated donors were explored. Galactosyl fluorides were first employed by Withers to produce UDP-Gal*p*, although high

substrate concentrations were required for product formation.[64] Subsequent experiments using *p*-nitrophenyl and 2,4-dinitrophenyl glycosides with both inverting and retaining GTs generated NDP-sugars *in situ* to form non-natural oligosaccharides.[65] The sugar nucleotides formed by these methods were not isolable, and instead were immediately consumed to produce oligosaccharides. Other methods were clearly needed to form substantial amounts of sugar nucleotides.

In recent years, Thorson and co-workers have improved upon the natural reversibility of several GTs through directed evolution, and used these mutants to create libraries of differentially glycosylated natural products.[66,67,68] One enzyme in particular, OleD from *Streptomyces antibioticus*,[69] was found to have exceptional active site tolerance toward both donors and acceptors. This GT was used in a spectrophotometric high-throughput assay to convert aromatic sugar donors into a variety of isolable sugar nucleotides (Scheme 1-10).[66] Predictably, the thermodynamics of the reaction relied upon the leaving group ability of the phenolic aglycon, with the 2-chloro-4-nitrophenyl derivative performing best.




**Scheme 1-10.** Sugar nucleotide synthesis developed by Thorson *et al*, using an engineered variant of the glycosyltransferase OleD.[66]

While NTs and GTs have been used to great success for the synthesis of pyranoside sugar nucleotides, the synthesis of furanose sugar nucleotides has been less well studied. Many of the initial efforts to synthesize these compounds have focused on pyranose–furanose mutases, enzymes that interconvert pyranose and furanose sugar nucleotides.[33] However, six-membered ring pyranose sugars are typically more thermodynamically favored over their five-membered ring furanose counterparts, resulting in low isolated yields of the furanose sugars. Generally, chemically synthesizing a furanose sugar-1-phosphate followed by nucleotide coupling using a NT leads to higher yields, when a suitable NT is available.[34]

### 1.6 Project Goals

The research described in this thesis focuses on methods of producing furanose sugar nucleotides, specifically UDP-Gal*f* **1** and UDP-Gal*f*NAc **2**, to aid in the biochemical investigation of the cell wall biosynthesis in *M. tuberculosis* and *C. jejuni*.

UDP-Gal*f* **1**, the donor substrate used by *M. tuberculosis* to form the galactan portion of its cell wall, is a particularly valuable sugar nucleotide. Due to the lack of furanose sugars and furanosyltransferases in mammalian systems, the enzyme responsible for processing **1** in *M. tuberculosis* (GlfT2) has been proposed as a potential target for new anti-TB drugs.[4b] Routine access to **1** and analogues would facilitate biochemical investigations into the enzyme's mechanism of action and help identify new therapeutics for TB. Unfortunately,

current methods of chemically synthesizing this compound suffer from low yields,[58] while chemoenzymatic methods involve a long and tedious preparation.[70]

The recent work of Thorson and co-workers[66] has suggested that highly activated donors can be used to thermodynamically drive GTs to work in reverse, forming sugar nucleotides in a fast and facile manner. The goals of this project were therefore set as follows:

1. Synthesis of activated donors: A range of phenyl galactofuranosides will be synthesized using chemical methods. The substituents on the phenyl ring will be varied to balance leaving group ability with resistance to hydrolysis in aqueous solution.
2. UDP-Gal*f* (**1**) formation: The synthesized phenyl precursors will be incubated with UDP and GlfT2 and monitored for the formation of **1** (Figure 1-11). The reaction may be monitored in several ways: Upon cleavage, the phenyl substituent could be detected by fluorescence, absorbance, or color change; alternatively, **1** could be directly detected using mass spectrometry, or be coupled to another enzyme.
3. Enzyme optimization and directed evolution: Upon the successful formation of UDP-Gal*f* (**1**), GlfT2 will be optimized using the techniques of directed evolution. Error-prone PCR will be used to create gene libraries that will be expressed in *E. coli*, where each colony will contain a single mutant. These colonies will then be screened with a fluorogenic substrate and optimized for production of **1**. This approach could also be

extended to produce a promiscuous enzyme capable of producing sugar nucleotides varying in both the sugar and nucleotide moieties.

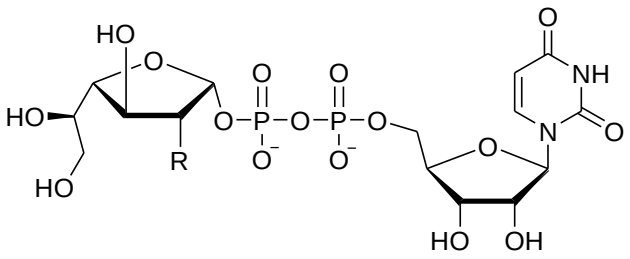



**Figure 1-11.** Structure of UDP-Gal*f* (**1**, R = OH) and UDP-Gal*f*NAc (**2**, R = NHAc).

UDP-Gal*f*NAc **2**, a sugar nucleotide from *C. jejuni*, was also targeted for synthesis for reasons identical to UDP-Gal*f* (**1**). Unlike **1**, this sugar nucleotide from *C. jejuni* has never been synthesized chemically, and is currently only obtainable through a mutase-mediated reaction that interconverts the pyranose and furanose forms of UDP-GalNAc.[33] This reaction heavily favors the pyranose form (93%) to the furanose form (7%) at equilibrium, resulting in low yields. Clearly, a more robust method of synthesizing this donor is needed. To this end, a synthesis of a sugar-1-phosphate precursor will be devised, and then incubated with galactose-1-phosphate uridyltransferase (GalPUT) and UTP using an enzymatic method optimized for the production of UDP-Gal*f* .[34,70] The chemoenzymatic synthesis of potentially useful analogues of **2** (e.g. replacing the –NHAc with an azide) will also be explored. Should this enzyme system fail to produce UDP-Gal*f*NAc, synthetic methodology could also be used to chemically synthesize this compound.

## 1.7 Bibliography


(1) Springer, S. A.; Gagneux, P. *J. Biol. Chem.* **2013**, *288*, 6904-6911.

(2) Breton, C.; Šnajdrová, L.; Jeanneau, C.; Koča, J.; Imberty, A. *Glycobiology* **2006**, *16*, 29R-37R.

(3) Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. *Annu. Rev. Biochem.* **2008**, *77*, 521-555.

(4) (a) Brennan, P. J. *Tuberculosis* **2003**, *83*, 91−97. (b) Richards, M. R.; Lowary, T. L. *ChemBioChem* **2009**, *10*, 1920−1938. (c) Chlubnova, I.; Legentil, L.; Dureau, R.; Pennec, A.; Almendros, M.; Daniellou, R.; Nugier-Chauvin, C.; Ferrieres, V. *Carbohydr. Res.* **2012**, *356*, 44−61.

(5) (a) St. Michael, F.; Szymanski, C. M.; Li, J.; Chan, K. H.; Khieu, N. H.; Larocque, S.; Wakarchuk, W. W.; Brisson, J. R.; Monteiro, M. A. *Eur. J. Biochem*. **2002**, *269*, 5119−5136. (b) McNally, D. J.; Lamoureux, M. P.; Karlyshev, A. V.; Fiori, L. M.; Li, J.; Thacker, G.; Coleman, R. A.; Khieu, N. H.; Wren, B. W.; Brisson, J.-R.; Jarrell, H. C.; Szymanski, C. M. *J. Biol. Chem.* **2007**, *282*, 28566−28576.

(6) Smith, N. H.; Hewinson, R. G.; Kremer, K.; Brosch, R.; Gordon, S. V. *Nat. Rev. Microbiol.* **2009**, *7*, 537-544.

(7) Ernst, J. D. *Nat. Rev. Immunol.* **2012**, *12*, 581-591.

(8) Kaufmann, S. H. E. *N. Engl. J. Med.* **2005**, *353*, 2423-2426.

(9) World Health Organization. *Global Tuberculosis Report 2013*. Geneva, Switzerland: World Health Organization, 2013.

http://www.who.int/tb/publications/global_report/en/index.html (Accessed February 5, 2014).

(10) Koul, A.; Arnoult, E.; Lounis, N.; Guillemont, J.; Andries, K. *Nature* **2011**, *469*, 483-490.

(11) Blanchard, J. S. *Annu. Rev. Biochem.* **1996**, *65*, 215-239.

(12) Shi, W.; Zhang, X.; Jiang, X.; Yuan, H.; Lee, J. S.; Barry, C. E. III; Wang, H.; Zhang, W.; Zhang, Y. *Science* **2011**, *333*, 1630-1632.

(13) Gandhi, N. R.; Nunn, P.; Dheda, K.; Schaaf, H. S.; Zignol, M.; van Soolingen, D.; Jensen, P.; Bayona, J. *Lancet* **2010**, *375*, 1830-1843.

(14) Klopper, M.; Warren, R. M.; Hayes, C.; van Pittius, N. C. G.; Streicher, E. M.; Müller, B.; Sirgel, F. A.; Chaluba-Nxiweni, M.; Hoosain, E.; Coetzee, G.; van Helden, P. D.; Victor, T. C.; Trollip, A. P. *Emerg. Infect. Dis.* **2013**, *19*, 449-455.

(15) Wieczorek, K.; Osek, J. *Biomed. Res. Int.* **2013**, *2013*, 340605.

(16) Young, K. T.; Davis, L. M.; DiRita, V. J. *Nat. Rev. Microbiol.* **2007**, *5*, 665−679.

(17) Guerry, P.; Poly, F.; Riddle, M.; Maue, A. C.; Chen, Y.-H.; Monteiro, M. A. *Front. Cell. Infect. Microbiol.* **2012**, *2*, 1-11.

(18) Moore, J. E.; Barton, M. D.; Blair, I. S.; Corcoran, D.; Dooley, J. S. G.; Fanning, S.; Kempf, I.; Lastovica, A. J.; Lowery, C. J.; Matsuda, M.; McDowell, D. A.; McMahon, A.; Millar, B. C.; Rao, J. R.; Rooney, P. J.; Seal, B. S.; Snelling, W. J.; Tolba, O. *Microbes Infect.* **2006**, *8*, 1955-1966.

(19) Koluman, A.; Dikici, A. *Crit. Rev. Microbiol.* **2013**, *39*, 57-69.

(20) Pedersen, L. L.; Turco, S. J. *Cell. Mol. Life Sci.* **2003**, *60*, 259-266.

(21) Lovering, A. L.; Safadi, S. S.; Strynadka, N. C. J. *Annu. Rev. Biochem.* **2012**, *81*, 451-478.

(22) Kaur, D.; Guerin, M. E.; Škovierová, H.; Brennan, P. J.; Jackson, M. *Adv. Appl. Microbiol.* **2009**, *69*, 23-78.

(23) Meroueh, S. O.; Bencze, K. Z.; Hesek, D.; Lee, M.; Fisher, J. F.; Stemmler, T. L.; Mobashery, S. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 4404–4409.

(24) (a) Dover, L. G.; Cerdeño-Tárraga, A. M.; Pallen, M. J.; Parkhill, J.; Besra, G. S. *FEMS Microbiol. Rev.* **2004**, *28*, 225-250. (b) Lopez-Martin, L. M. *Clin. Dev. Immunol.* **2012**, *2012*, 917860.

(25) Neyrolles, O.; Guilhot, C. *Tuberculosis* **2011**, *91*, 187-195.

(26) (a) Szymanski, C. M.; St. Michael, F.; Jarrell, H. C.; Li, J.; Gilbert, M.; Larocque, S.; Vinogradov, E.; Brisson, J. R. *J. Biol. Chem.* **2003**, *278*, 24509-24520. (b) McNally, D. J.; Jarrell, H. C.; Khieu, N. H.; Vinogradov, E.; Szymanski, C. M.; Brisson, J. R. *FEBS J.* **2005**, *272*, 4407-4422. (c) McNally, D. J.; Jarrell, H. C.; Khieu, N. H.; Li, J.; Vinogradov, E.; Whitfield, D. M.; Szymanski, C. M.; Brisson, J. R. *FEBS J.* **2006**, *273*, 3975-3989.

(27) Parkhill, J.; Wren, B. W.; Mungall, K.; Ketley, J. M.; Churcher, C.; Basham, D.; Chillingworth, T.; Davies, R. M.; Feltwell, T.; Holroyd, S.; Jagels, K.; Karlyshev, A. V.; Moule, S.; Pallen, M. J.; Penn, C. W.; Quail, M. A.; Rajandream, M.-A.; Rutherford, K. M.; van Vliet, A. H. M.; Whitehead, S.; Barrell, B. G. *Nature* **2000**, *403*, 665-668.

(28) Guerry, P.; Szymanski, C. M. *Trends Microbiol.* **2008**, *16*, 428-435.

(29) Karlyshev, A. V.; Champion, O. L.; Churcher, C.; Brisson, J.-R.; Jarrell, H. C.; Gilbert, M.; Brochu, D.; St. Michael, F.; Li, J.; Wakarchuk, W. W.; Goodhead, I.; Sanders, M.; Stevens, K.; White, B.; Parkhill, J.; Wren, B. W.; Szymanski, C. M. *Mol. Microbiol.* **2005**, *55*, 90-103.

(30) Karlyshev, A. V.; Wren, B. W.; Moran, A. P. In *Campylobacter (Third Edition)*; Nachamkin, I., Szymanski, C. M., Blaser, M. J., Eds.; American Society for Microbiology Press: Washington, DC, 2008; pp 505-521.

(31) Gilbert, M.; Parker, C. T.; Moran, A. P. In *Campylobacter (Third Edition)*; Nachamkin, I., Szymanski, C. M., Blaser, M. J., Eds.; American Society for Microbiology Press: Washington, DC, 2008; pp 483-504.

(32) Cai, L. J. *Carbohydr. Chem.* **2012**, *31*, 535−552.

(33) Poulin, M. B.; Nothaft, H.; Hug, I.; Feldman, M. F.; Szymanski, C. M.; Lowary, T. L. *J. Biol. Chem.* **2010**, *285*, 493−501.

(34) Errey, J. C.; Mukhopadhyay, B.; Kartha, K. P. R.; Field, R. A. *Chem. Commun.* **2004**, *23*, 2706−2707.

(35) Wagner, G. K.; Pesnot, T.; Field, R. A. *Nat. Prod. Rep.* **2009**, *26*, 1172−1194.

(36) Engels, L.; Elling, E. In *Carbohydrate-Modifying Biocatalysts*; Grunwald, P., Ed.; Pan Stanford Publishing: Boca Raton, FL, 2012; pp 237−254.

(37) Vocadlo, D. J.; Davies, G. J. *Curr. Opin. Chem. Biol.* **2008**, *12*, 539-555.

(38) Recent reviews: (a) Thibodeaux, C. J.; Melançon, C. E.; Liu, H. W. *Angew. Chem., Int. Ed.* **2008**, *47*, 9814−9859. (b) Singh, S.; Phillips, G. N., Jr.;

Thorson, J. S. *Nat. Prod. Rep*. **2012**, *29*, 1201−1237. (c) Gantt, R. W.; Peltier-Pain, P.; Thorson, J. S. *Nat. Prod. Rep.* **2011**, *28*, 1811−1853.

(39) Rosén, M. L.; Edman, M.; Sjöström, M.; Wieslander, A. *J. Biol. Chem.* **2004**, *279*, 38683-38692.

(40) Lee, S. S.; Hong, S. Y.; Errey, J. C.; Izumi, A.; Davies, G. J.; Davis, B. G. *Nat. Chem. Biol*. **2011**, *7*, 631-638.

(41) Lairson, L. L.; Chiu, C. P. C.; Ly, H. D.; He, S.; Wakarchuk, W. W.; Strynadka, N. C. J.; Withers, S. G. *J. Biol. Chem.* **2004**, *279*, 28339-28344.

(42) Soya, N.; Fang, Y.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2011**, *21*, 547-552.

(43) (a) Crick, D. C.; Mahapatra, S.; Brennan, P. J. *Glycobiology* **2001**, *11*, 107R-118R. (b) Alderwick, L. J.; Birch, H. L.; Mishra, A. K.; Eggeling, L.; Besra, G. S. *Biochem. Soc. T.* **2007**, *35*, 1325-1328.

(44) Peltier, P.; Daniellou, R.; Nugier-Chauvin, C.; Ferrières, V. *Org. Lett.* **2007**, *9*, 5227–5230.

(45) MacDonald, D. L. *J. Org. Chem.*, **1962**, *27*, 1107–1109.

(46) Hanessian, S.; Lu, P.-P.; Ishida, H. *J. Am. Chem. Soc.* **1998**, *120*, 13296–13300.

(47) Edgar, L. J. G.; Dasgupta, S.; Nitz, M. *Org. Lett.* **2012**, *14*, 4226-4229.

(48) de Lederkremer, R. M.; Nahmad, V. B.; Varela, O. *J. Org. Chem.* **1994**, *59*, 690−692.

(49) Lowary, T. L.; Hindsgaul, O. *Carbohydr. Res.* **1994** , *251* , 33–67.

(50) Oberthür, M.; Leimkuhler, C.; Kahne, D.; *Org. Lett.* **2004**, *6*, 2873–2876.

(51) Timmons, S. C.; Jakeman, D. L. *Carbohydr. Res.* **2008**, *343*, 865–874.

(52) Moffatt, J. G.; Khorana, H. G. *J. Am. Chem. Soc.* **1958**, *80*, 756.

(53) Roseman, S.; Distler, J. J.; Moffatt, J. G.; Khorana, H. G. *J. Am. Chem. Soc.* **1961**, *83*, 659–663.

(54) Wittmann, V.; Wong, C.-H. *J. Org. Chem.* **1997**, *62*, 2144–2147.

(55) Collier; A.; Wagner, G. K. *Chem. Commun.* **2008**, 178–180.

(56) Baisch, G.; Oehrlein, R. *Bioorg. Med. Chem.* **1997**, *5*, 383–391.

(57) Bogachev, V. S. *Russ. J. Bioorg. Chem.* **1996**, *22*, 699.

(58) Marlow, A. L.; Kiessling, L. L. *Org. Lett.* **2001**, *3*, 2517–2519.

(59) Tanaka, H.; Yoshimura, Y.; Jørgensen, M. R.; Cuesta-Seijo, J. A.; Hindsgaul, O. *Angew. Chem., Int. Ed.* **2012**, *51*, 11531−11534.

(60) Langenhan, J. M.; Griffith, B. R.; Thorson, J. S. *J. Nat. Prod.* **2005**, *68*, 1696–1711

(61) Moretti, R.; Thorson, J. S. *J. Biol. Chem.* **2007**, *282*, 16942-16947.

(62) Moretti, R.; Chang, A.; Peltier-Pain, P.; Bingman, C. A.; Philips, G. N., Jr.; Thorson, J. S. *J. Biol. Chem.* **2011**, *286*, 13235-13243.

(63) Zhang, C.; Griffith, B. R.; Fu, Q.; Albermann, C.; Fu, X.; Lee, I.-K.; Li, L.; Thorson, J. S. *Science* **2006**, *313*, 1291-1294.

(64) Lougheed, B.; Ly, H.D.; Wakarchuk, W.W.; Withers, S.G. *J. Biol. Chem.* **1999**, *274*, 37717–37722.

(65) Lairson, L.L., Wakarchuk, W.W. & Withers, S.G. *Chem. Commun.* **2007**, 365–367.

(66) Gantt, R. W.; Peltier-Pain, P.; Cournoyer, W. J.; Thorson, J. S. *Nat. Chem. Biol.* **2011**, *7*, 685-695.

(67) Gantt, R. W.; Peltier-Pain, P.; Singh, S.; Zhou, M.; Thorson, J. S. *P. Natl. Acad. Sci. USA* **2013**, *110*, 7648-7653.

(68) Williams, G. J.; Zhang, C.; Thorson, J. S. *Nat. Chem. Biol.* **2007**, *3*, 657-662.

(69) Yang, M.; Proctor, M. R.; Bolam, D. N.; Errey, J. C.; Field, R. A.; Gilbert, H. J.; Davis, B. G. *J. Am. Chem. Soc.* **2005**, *127*, 9336-9337.

(70) Rose, N. L.; Zheng, R. B.; Pearcey, J.; Zhou, R.; Completo, G. C.; Lowary, T. L. *Carbohydr. Res.* **2008**, *343*, 2130-2139.

# Chapter 2 : Attempted Chemoenzymatic Synthesis of UDP-Gal*f* Using Activated Donors[1]

[1] W309S and W399S mutants were prepared by Ruixiang Blake Zheng. Gal*f*-containing trisaccharide **17** was prepared by Simon Byrns.

## 2.1 Introduction and Project Aim

The work of Thorson and co-workers has demonstrated that some glycosyltransferase enzymes can be made to work in reverse to synthesize sugar nucleotides from activated sugar donors.[1,2] This work used exclusively sugars in the six-membered pyranose ring form, leaving room for determining if a similar system could be developed to synthesize five-membered ring furanose sugar nucleotides.

The key to Thorson's system was the use of an appropriate enzyme to carry out the conversion of the activated donor into a sugar nucleotide.[3] Several glycosyltransferases are known to have relaxed substrate specificity, accepting a wide range of substrates for glycosylation.[4] A recently isolated and characterized galactofuranosyltransferase (GlfT2) from *M. tuberculosis*[5] was chosen to probe for substrate promiscuity; this enzyme catalyzes the addition of galactofuranose residues into bacterial polysaccharides using UDP-Gal*f* **1** as the donor. It was proposed that, given a suitable highly activated substrate, this enzyme could be induced to work in reverse, thereby producing **1**.

The chemoenzymatic approach to producing UDP-Gal*f* outlined above requires several steps. First, a robust synthesis of the required activated sugar donors needs to be devised. With these compounds in hand, the wild-type GlfT2 could be screened for activity with a range of substrates. Once it has been demonstrated that GlfT2 can work in reverse, and the ideal substrate identified, error-prone PCR can be used to carry out directed evolution[6] of the enzyme expressed in *E. coli* with the aim of engineering an enzyme that produces **1**

quickly in good yield. This enzyme could also be screened with other sugars and other nucleotides.

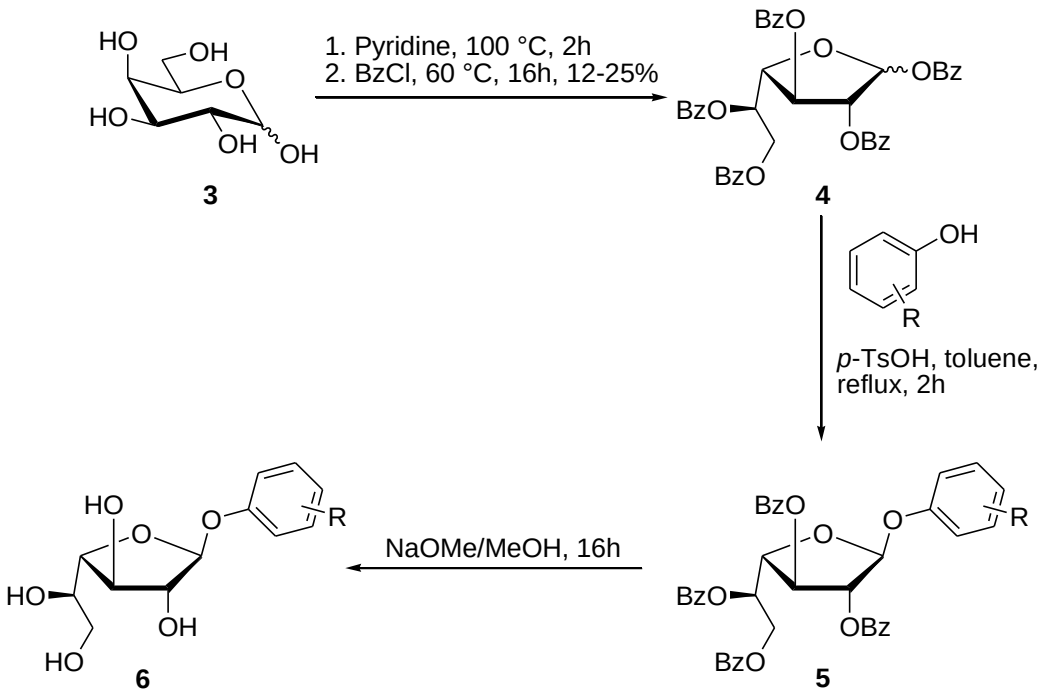
## 2.2 Synthesis of Activated Donors

Similar to the work of Thorson,[1] phenyl-based furanosides were chosen to be the activated donors for this project. Phenol derivatives are readily available, are relatively inexpensive, and can be substituted with a wide variety of groups, depending on how activated the substrate needs to be. Further, phenols, both when attached to a carbohydrate and when liberated from them, can also be detected by UV light, making high-throughput fluorogenic or UV/colorimetric screening of substrates or mutants possible.

The first challenge in the synthesis of galactofuranosides is to convert galactose from the pyranose to furanose form. A recent review[7] outlines some of the methods available for catalyzing this transformation, including classical Fischer glycosylation with simple alcohols and the cyclization of dithioacetals. D-Galactose diethyl dithioacetal can be cyclised using $Hg^{2+}$ salts to form the β-ethyl galactofuranoside.[8] To avoid the use of toxic heavy metals, $I_2$ can also be used as a catalyst for the cyclisation, forming an anomeric mixture of methyl galactofuranosides.[9]

One further strategy uses a high temperature acylation to convert D-galactose directly into its perbenzoylated derivative.[10] This method proved to be the most desirable for this project for a number of reasons: It employs no toxic heavy metals or odorous reagents, the setup and workup are facile and fast, and

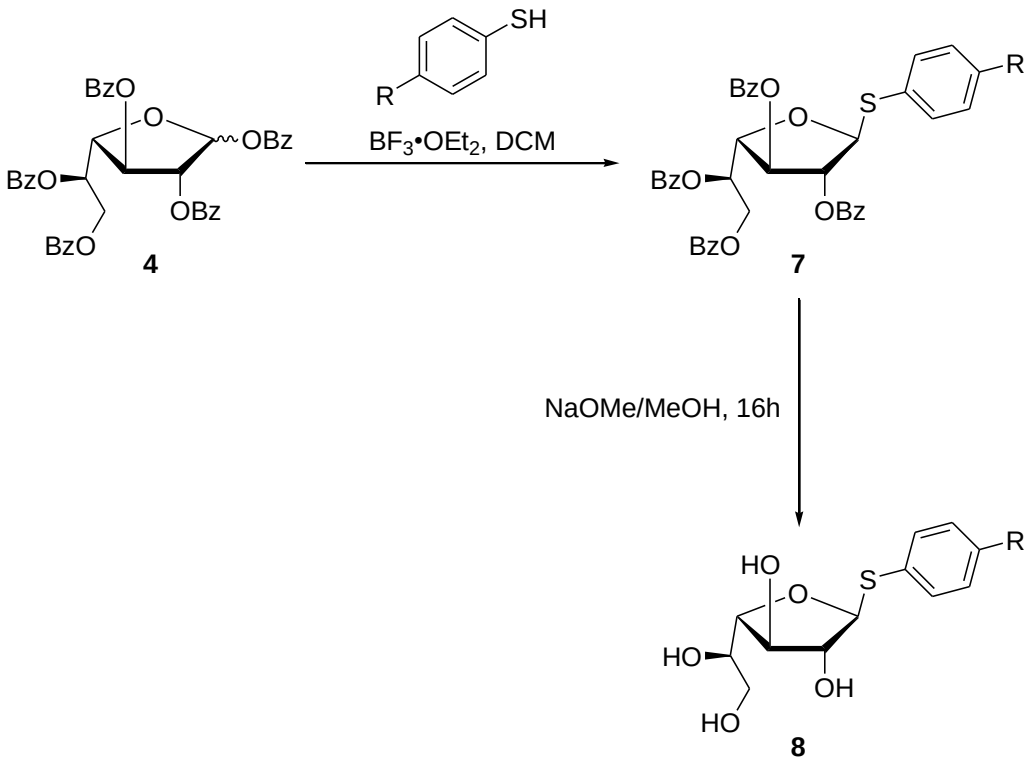
the reagents are inexpensive. D-Galactose **3** in pyridine was heated in a boiling water bath for 2 h to furnish a pyranose/furanose mixture (Scheme 2-1), then the temperature was lowered to 60 ˚C and benzoyl chloride added. After stirring overnight, the product was worked up and recrystallized from ethanol. Although previous reports have stated that separation of the perbenzoylated furanose **4** and pyranose forms can be difficult,[9] in this instance no such problems were encountered. More than one recrystallization was typically needed, but often even α and β furanose forms of **4** could be separated from each other. Typical yields were 12–25%.




**Scheme 2-1.** Synthesis of phenyl galactofuranosides **6a-e**. (**6a**, R = H; **6b**, R = *p*-$NO_2$; **6c**, R = *p*-Br; **6d**, R = *p*-I; **6e**, R = *o*-I).

A wide range of galactofuranoside derivatives could now be synthesized from this perbenzoylated galactofuranose derivative, and two glycosylation methods were explored. The first method heats galactofuranose **4** in toluene at reflux with a substituted phenol and a catalytic amount of *p*-toluenesulfonic acid to yield the desired β-galactofuranoside **5**.[11] Alternatively, boron trifluoride diethyl etherate could be used to form the glycoside. In each case, neighbouring group participation by the C-2 benzoyl group ensured that only a single anomer (the 1,2-*trans* glycoside) of **5** was obtained. The crude product of either reaction was then deacylated using NaOMe/MeOH, furnishing the desired β-galactofuranoside **6**. Using phenol as an acceptor, similar yields (78–84%) of **6a** were obtained for each method. Due to the shorter time required for the reaction catalyzed by *p*-toluenesulfonic acid, this method was employed for the synthesis of the remaining phenyl galactofuranosides **6b-e**.

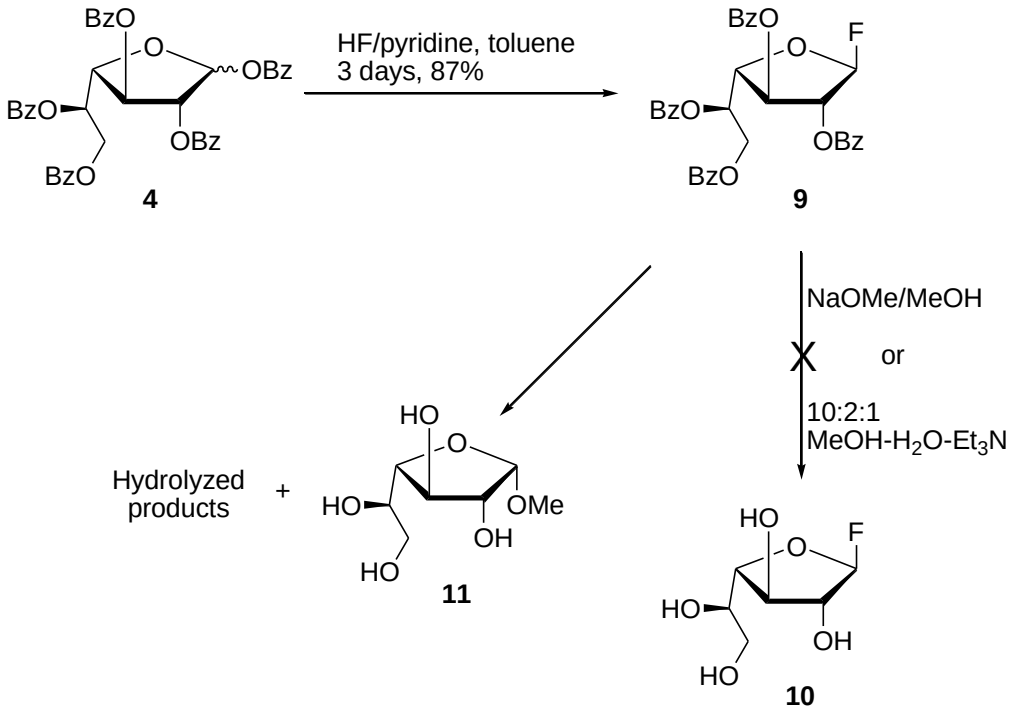
Two thioglycosides **8a** and **8b** were also synthesized (Scheme 2-2). Compared to a carbon–oxygen bond, a carbon–sulfur bond is longer (143 vs. 182 pm) and weaker (360 vs. 272 kJ/mol) due to less favourable orbital overlap between the C and S atoms. As a result, thioglycosides were predicted to be more highly activated donors. *p*-Tolyl 1-thio-β-D-galactofuranoside **8a** and *p*-chlorophenyl 1-thio-β-D-galactofuranoside **8b** were synthesized using $BF_3 \cdot OEt_2$ for the glycosylation of perbenzoylated Gal*f* with *p*-thiocresol or *p*-chlorothiophenol, followed by deprotection with NaOMe/MeOH. Overall yields of 64% and 54% were achieved for compounds **8a** and **8b**, respectively.




**Scheme 2-2.** Synthesis of thiophenyl galactofuranosides **8a** (R = $CH_3$) and **8b** (R = Cl).

In addition to phenyl and thiophenyl glycosides **6** and **8**, galactofuranosyl fluorides were also explored as potential activated donors. Galactopyranosyl fluorides have been used in the past as substrates for Leloir glycosyltransferases,[12] where the fluoride group was chosen for its small size and relative stability compared to other halides. A straightforward synthesis of 2,3,5,6-tetra-*O*-benzoyl-β-D-galactofuranosyl fluoride **9** based on a previous literature procedure was developed (Scheme 2-3),[13] where perbenzoylated galactofuranose **4** was treated with a 70% solution of HF in pyridine in a polyethylene flask. This simple synthesis produced the protected β-glycosyl fluoride **9** in good yield (87%) after

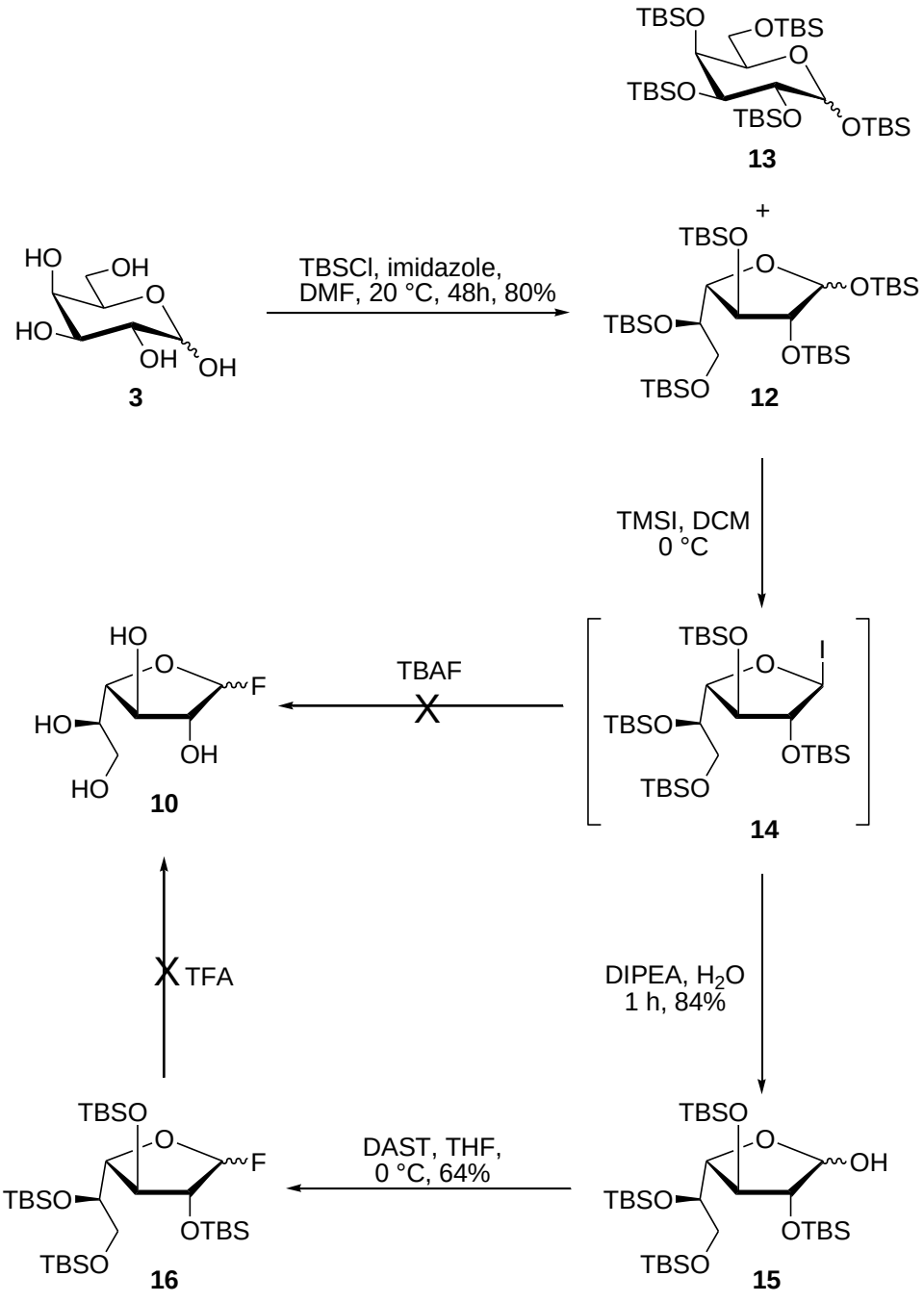
reacting for three days. Unfortunately, efforts to deprotect this compound with NaOMe/MeOH or a methanol–water–triethylamine solution to produce **10** resulted in either methanolysis or hydrolysis of the desired fluoride.




**Scheme 2-3.** Attempted synthesis of galactofuranosyl fluoride **10**.

In search of an orthogonal deprotection strategy, TBS-protected galactofuranose derivatives were also explored. 1,2,3,5,6-Penta-*O*-*tert*-butyldimethylsilyl-β-D-galactofuranose **12** has been previously synthesized from D-galactose **3** and TBSCl in DMF/pyridine, with imidazole acting as base.[14,15] This substrate could be activated as a donor with TMSI to form a galactofuranosyl iodide **14**, and then reacted with an acceptor to form glycosides in a one-pot

reaction. Using this method, pure persilylated β-D-galactofuranose **12** was isolated in 21% yield, although another fraction containing an inseparable β-furanose and β-pyranose mixture was also obtained (59%;Scheme 2-4).




**Scheme 2-4.** Synthesis of fluorinated galactofuranosides.

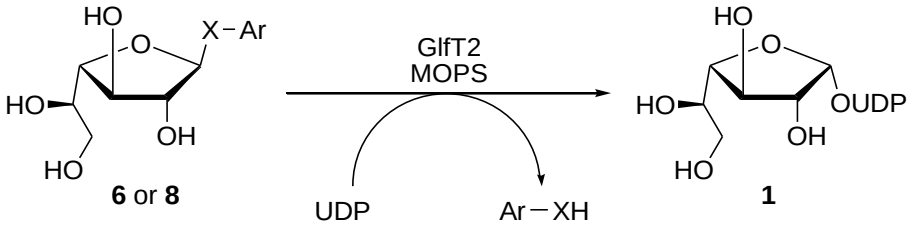
Initially, it was thought that galactofuranosyl fluoride **10** could be synthesized in a two-step, one pot reaction by TMSI activation[14] followed by the addition of TBAF. In addition to providing a source of fluoride ions to displace the iodide, TBAF would also remove the TBS protecting groups. Sadly, this strategy resulted in only hydrolyzed product, and no fluorinated product was visible in $^{19}F$ NMR spectra. A sequential fluorination–deprotection strategy was then employed by first hydrolyzing **12** to a reducing sugar **15** followed by reaction with DAST[16] to form a tetrasilylated fluoride **16**. Although this compound could be formed, TLC analysis indicated that the fluoride group was readily hydrolyzed, and subsequent deprotection attempts using TFA did not produce the target molecule. Any compound to be used as a donor for coupling reactions with GlfT2 would necessarily need to be stable in an aqueous environment. Based on these investigations, it was determined that galactofuranosyl fluorides were simply too unstable to be of use for enzymatic experiments, and this line of investigation was discontinued.

In total, five phenyl and two thiophenyl galactofuranosides were synthesized for potential use as activated donors with GlfT2. Yields ranged from 22% to 78%, but each compound could be obtained in sufficient amount for subsequent experiments.

### 2.3 Engineering of GlfT2

With substrates in hand, the possibility of using GlfT2 in reverse to form UDP-Gal*f* was explored (Scheme 2-5). GlfT2 was expressed and purified

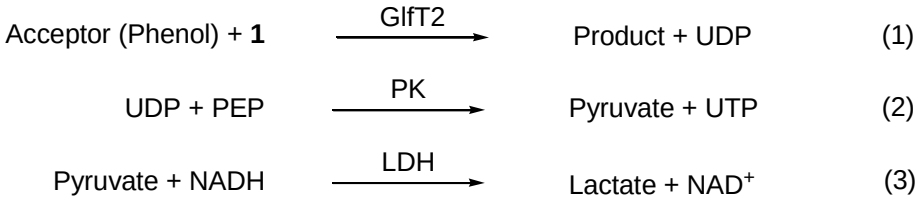
according to a previously reported literature procedure,[17] then incubated in MOPS buffer (pH = 7.6) with one of the phenyl galactofuranosides **6a**, **6b**, or **6c** and UDP. Thiophenyl galactofuranosides **8a** and **8b** were also investigated. When HPLC analysis of the reaction mixture after 1 h of incubation failed to show any UDP-Gal*f* (**1**), the mixtures were allowed to continue reacting for 24 h. Even after this time, no trace of **1** was observed. Using **6b** as the substrate, the effect of the reaction mixture pH was investigated, but no product formation was observed with pH = 7.0, 8.0, or 8.6. Increasing the reaction temperature to 37 °C also had no effect.



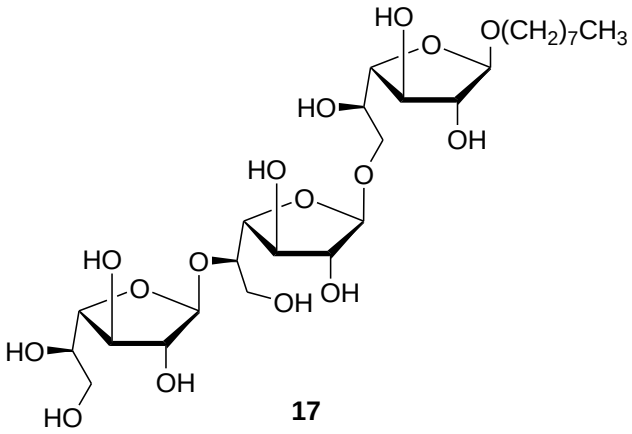

**Scheme 2-5.** Attempted chemoenzymatic synthesis of 1. X = O (**6**), S (**8**).

There could be two reasons for the inability of GlfT2 to work in reverse to produce UDP-Gal*f* (**1**): Either the substrates were not sufficiently activated for the reaction to proceed in the reverse direction, or the substrate was not fitting into the active site as assumed. To investigate these possibilities, a spectrophotometric assay was used to verify the efficiency of GlfT2 in catalyzing the forward reaction,[17] using **1** as donor and phenol as acceptor to form the phenyl glycoside. An overview of the assay is shown in Scheme 2-6; briefly, **1** and an acceptor undergo a glycosylation catalyzed by GlfT2 at 37 °C to form a glycoside with

UDP as a side product (Reaction 1). The UDP goes on to be phosphorylated by pyruvate kinase, converting phosphoenolpyruvate to pyruvate in the process (Reaction 2). Finally, pyruvate is reduced to lactate through the action of lactate dehydrogenase (Reaction 3). The rate of NADH oxidation to $NAD^+$ produced by this final reaction is monitored at $\lambda$ = 340 nm. Using this assay, no activity was observed using phenol as the acceptor. Control reactions incubating **1** and GlfT2 with an analogue (**17**)[9] (Figure 2-1) of the natural substrate showed that the enzyme used was functional.




**Scheme 2-6.** Coupled spectrophotometric GlfT2 assay. Scheme adapted from reference 17.




**17**

**Figure 2-1.** Structure of the (1→5),(1→6) trisaccharide **17** described in this study.

The failure of the forward reaction to produce phenyl galactofuranosides led to the conclusion that the active site specificity would not allow for alternative substrates to be used. It was proposed that mutagenesis of GlfT2, converting large amino acids in the active site to smaller alanine residues, could provide extra room for the substrate. Fortunately the crystal structure of GlfT2 was published in 2012,[5a] enabling a detailed analysis of the enzyme–substrate interactions at the active site. Based upon the proposed binding interactions, three residues were targeted for mutagenesis: E300, Y344, and H413 (Figure 2-2). The initial approach was to create a triple alanine mutant, E300A_Y344A_H413A, which would hollow out the area of the active site where the phenyl group was predicted to sit. Although none of these residues is the catalytic residue of this enzyme, three double mutants (E300A_Y344A, E300A_H413A, and H413A_Y344A) were also desirable in case one of the targeted residues was required for activity.

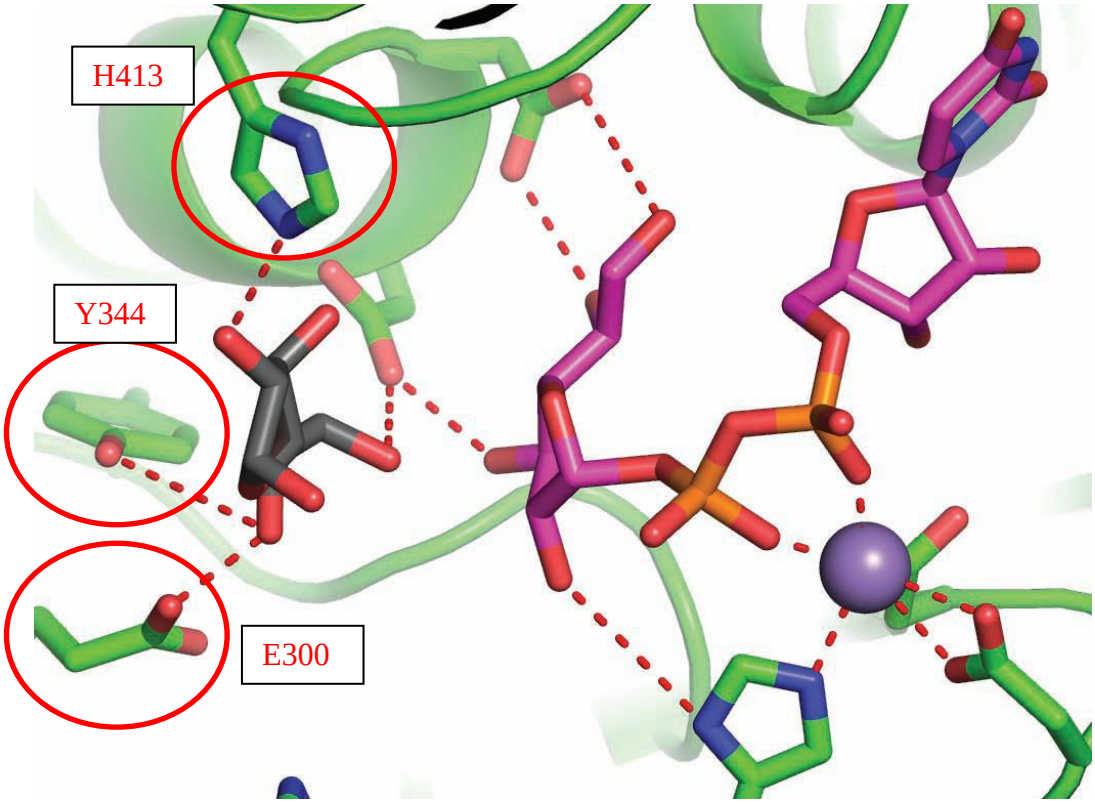



**Figure 2-2.** UDP-Gal*f* (**1**) addition to a Gal*f* acceptor in the GlfT2 active site, with residues targeted for mutagenesis highlighted in red. Reprinted by permission; this image was originally published in *Journal of Biological Chemistry*. Wheatley, R. W.; Zheng, R. B.; Richards, M. R.; Lowary, T. L.; Ng, K. K. S. Tetrameric Structure of the GlfT2 Galactofuranosyltransferase Reveals a Scaffold for the Assembly of Mycobacterial Arabinogalactan. *J. Biol. Chem.* 2012; 287: 28132−28143. © the American Society for Biochemistry and Molecular Biology.

The four mutants described above were created using the Stratagene Quikchange II mutagenesis kit, and mutations confirmed by genetic sequencing. Each of the four proteins was expressed and purified as described for GlfT2,[17] and enzymatic reactions set up using 1 mM concentrations of donor **6b**, **6c**, or **8a** and UDP, and 200 μg of enzyme in MOPS buffer (pH = 7.6). Control reactions that

did not contain any enzyme were also set up using the same conditions. Aliquots were analyzed by HLPC after an incubation time of 1 h and 24 h.

None of the phenyl glycosides were converted to UDP-Gal*f*. To assess if these compounds were simply not activated enough, each of the four mutants was tested in the forward direction with UDP-Gal*f* (**1**) as a donor and phenol as an acceptor, using the enzymatic spectrophotometric assay described above. This experiment also failed to show any activity with the mutant enzymes, implying that the substrate may still be unable to bind in the active site. The mutants were tested next with **1** and a (1→5),(1→6) Gal*f* trisaccharide **17**[9] to determine if they retained any activity with their natural substrate. The assay results indicated no activity with any of these mutant enzymes, suggesting that these residues were required for GlfT2 to have any activity.

At this point, the choice of mutation was re-evaluated by examining the crystal structure of GlfT2. The enzyme was viewed using PyMol software, and potential steric interactions were analyzed. By positioning a phenyl ring on the anomeric position of the putative Gal*f* donor, steric clash could be seen with two tryptophan residues, W309 and W399 (Figure 2-3). Mutants had been recently expressed and purified by R. B. Zheng, where each tryptophan was replaced by a serine to form W309S and W399S GlfT2 mutants. These two mutants were created using the methods outlined above; in addition, the double mutant W309S_W399S was also expressed and purified by me. These three mutants were tested with phenyl galactofuranoside donors **6a**, **6b**, and **8a**, but failed to show

activity with any of these compounds. At this point, further research into modifying GlfT2 specificity was discontinued.

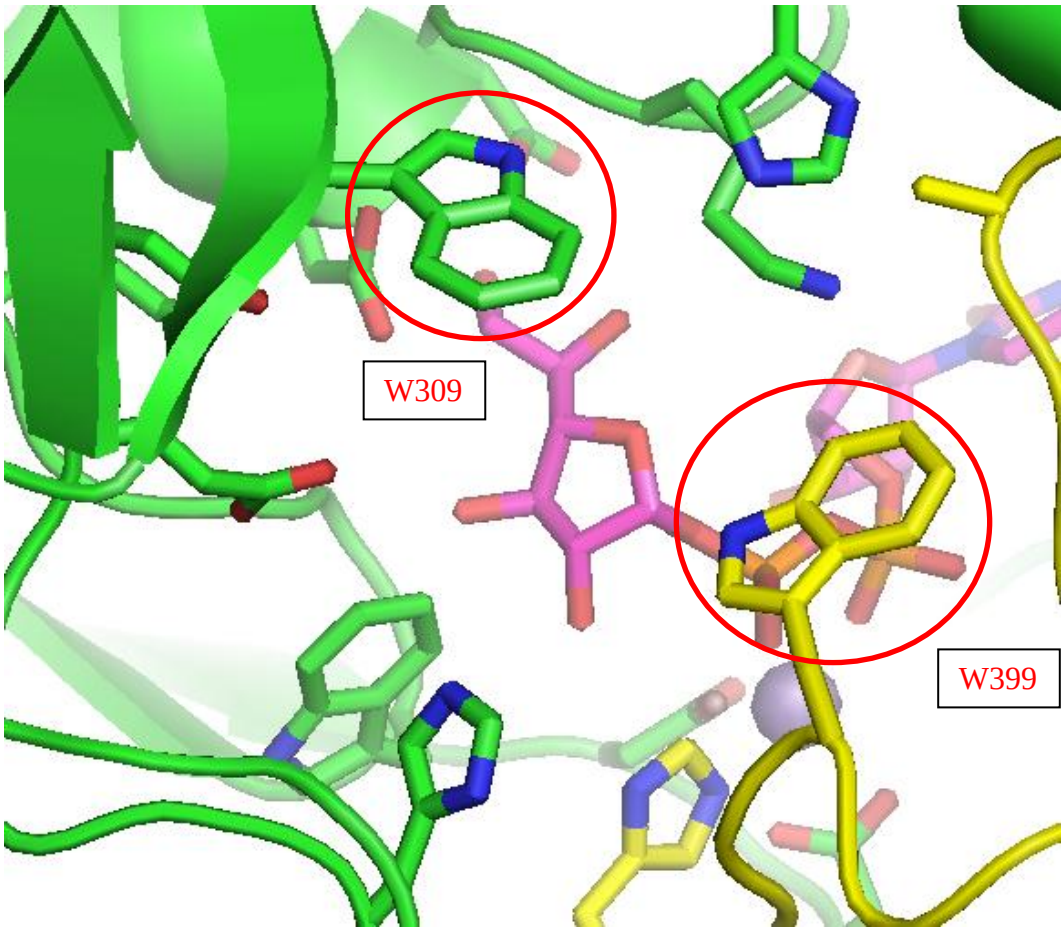
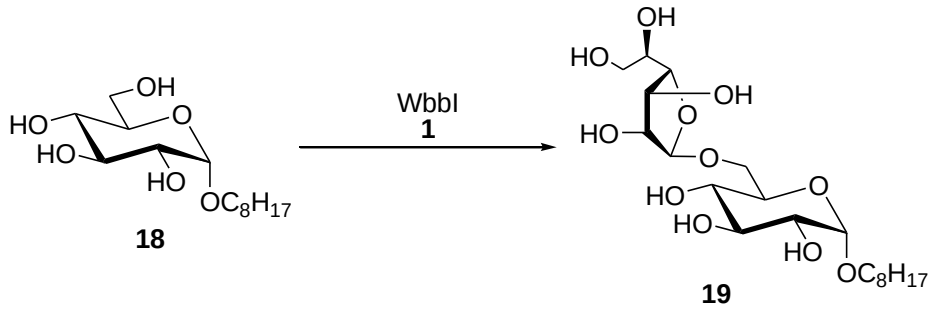



**Figure 2-3.** Position of UDP-Gal*f* (**1**) relative to W309 and W399 in the GlfT2 active site. Image generated using PyMol software: The PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC.

Aside from GlfT2, only one other galactofuranosyltransferase enzyme has been well characterized. WbbI was isolated from *E. coli* K-12 in 2006 by Field and coworkers,[18] where it is involved in the biosynthesis of bacterial O antigen. This enzyme catalyzes the transfer of a Gal*f* moiety to form a β-(1→6) linkage

with a glucose acceptor (Scheme 2-7). No sequence homology was observed between WbbI and GlfT2.




**Scheme 2-7.** Glycosylation catalyzed by WbbI.

WbbI was expressed and purified as a maltose-binding protein, as reported previously.[18] The enzyme's activity was verified by incubating **1** and WbbI with octyl α-D-glucopyranoside (**18**) and monitoring the product formation by TLC, where the disaccharide **19** could be easily seen when stained with an orcinol solution. Galactofuranosides **6b**, **6c**, and **8a** were screened against the enzyme, but no sugar nucleotide product formation was detectable by either TLC or HPLC after 3 h of incubation.

## 2.4 Conclusions and Future Outlook

In this project, a series of phenyl galactofuranosides were synthesized and tested as substrates of the galactofuranosyltransferase enzyme GlfT2. The ultimate goal was to develop a chemoenzymatic synthesis of the sugar nucleotide UDP-Gal*f* (**1**), but sadly GlfT2 did not accept the phenyl donors as substrates. Although many glycosyltransferase enzymes possess relaxed substrate

specificity,[1,4] the work presented above suggests that GlfT2 does not. While the activated phenyl donor molecules **6** and **8** could be synthesized in reasonable yields, galactofuranose modified in this way is clearly not a substrate for GlfT2.

In hindsight, a comparison of GlfT2 with Thorson's original OleD enzyme reveals several critical differences that could rationalize the observed lack of activity in GlfT2. OleD is an inverting glycosyltransferase belonging to the GT-B superfamily that inactivates macrolide antibiotics through glycosylation.[1,19] It has been noted that even wild-type OleD displays a broad substrate scope that appears to favor "planar, cyclic, hydrophobic molecules."[20] Directed evolution of this enzyme resulted in an exceptionally promiscuous mutant,[3a] setting the stage for later experiments.

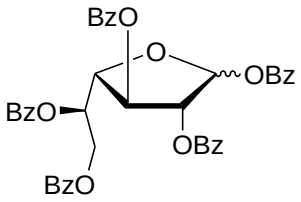
While OleD appears to have evolved to be able to glycosylate a number of different natural products, GlfT2 has only been shown to carry out the addition of galactofuranose to galactan in an alternating (1→5),(1→6) fashion.[5] Like other members of the GT-A family of glycosyltransferases, GlfT2 has a Rossman fold that is much tighter than those found in the GT-B family.[5a] This enzyme presumably evolved only to carry out the biosynthesis of a specific cell wall component.

Future attempts to carry out syntheses using glycosyltransferases may benefit from a careful examination of the wild-type enzyme's natural substrate scope. For example, a method currently in use to produce **1** relies upon the natural relaxed substrate specificity of a wild-type uridyltransferase, GalPUT.[4b] Expanding the substrate scope of GTs that are known to be inherently permissive

could most easily identify mutants capable of carrying out a desired transformation.

## 2.5 Experimental

### 1,2,3,5,6-Penta-*O*-benzoyl-α,β-D-galactofuranose (4)



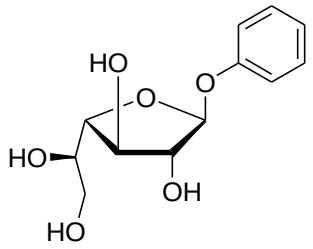

**4**

Synthesized as described previously.[10] D-Galactose (**3**) (5.0 g, 27.7 mmol) was stirred in dry pyridine (70 mL) over a gently boiling water bath for 2 h. The flask was then removed from the water bath and allowed to cool to 60–65 °C. Benzoyl chloride (20 mL, 170 mmol) was added portion-wise while maintaining the temperature around 65 °C, followed by heating at 60 °C for 1.5 h. Water (10 mL) was slowly added and the mixture stirred at room temperature for 30 min, then poured with stirring into 500 mL of ice water. Immediately after the ice had melted, the liquid was decanted and the solid washed with 500 mL ice water while rubbing the gummy solid with a glass rod. The solid was vacuum filtered and washed with several portions of ice cold water, then recrystallized from 800 mL ethanol to obtain 2.25 g (3.21 mmol, 12%) of white crystals as a mixture of α and β anomers of **4**. The NMR data obtained on the product was consistent with previously published spectra.[10] $R_f$ = 0.58 (9:1 toluene–EtOAc); $^1$H NMR (500 MHz, $CDCl_3$) δ 8.12–7.13 (m, Ar), 6.85 (d, $J$ = 4.9 Hz, 1H, H-1α), 6.78 (s, 1H, H-1β), 6.29 (app t, $J$ = 7.2 Hz, 1H, H-3α), 6.13 (ddd, $J$ = 7.0, 4.0. 4.0 Hz, 1H, H-5β), 5.91–5.87 (m, 2H, H-2α,5α), 5.80 (d, $J$ = 4.1, 1H, H-3β), 5.77 (s, 1H, H-2β), 4.86 (app t, $J$ = 4.0 Hz, 1H, H-4β), 4.82–4.74 (m, 4H, H-4α,6aα,H-6aβ,6bβ) 4.70 (dd, $J$

= 12.0, 6.2 Hz, 1H, 6bα); $^{13}$C NMR (125 MHz, $CDCl_3$) δ 166.1, 166.0, 165.8, 165.6, 165.5, 165.5, 165.3, 165.0, 164.5, 159.6 (10 x *CO*Ph), 133.8–133.1 (Ar), 130.1 – 128.2 (Ar), 99.8 (C-1β), 94.2 (C-1α), 84.4 (C-4β), 81.1 (C-2β), 79.6 (C-4α), 77.5 (C-3β), 76.1 (C-2α), 73.9 (C-3α), 70.4 (C-5α), 70.3 (C-5β), 63.6 (C-6β), 63.1 (C-6α).

**General synthesis of phenyl β-D-galactofuranosides 6:** Perbenzoylated galactofuranose, the phenol (5 eq.), and *p*-toluenesulfonic acid (5 mg, catalytic) were dissolved in toluene and heated to reflux for 2 h. After this time, toluene was removed under reduced pressure and the residue dissolved in DCM (30 mL). The solution was extracted 2 x 30 mL with 2% aq. NaOH and 1 x 30 mL with brine, and then the organic layer was dried over $Na_2SO_4$, filtered, and concentrated. Solid NaOMe was added to the crude benzoylated product stirring in methanol (2 mL) until the solution was cloudy and strongly basic. After stirring overnight, the solution was neutralized using Amberlite IR-120 ($H^+$ form) resin and filtered over Celite. The solvent was evaporated and the crude material was purified using column chromatography (FCC on silica, 100% EtOAc to 95:5 EtOAc–MeOH).

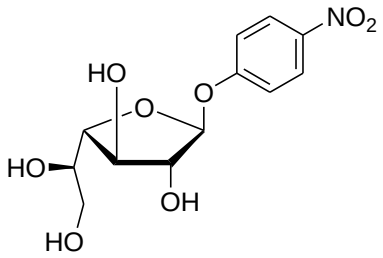
**Phenyl β-D-galactofuranoside (6a)**




**6a**

Synthesized using general synthetic procedure (57.7 mg, 0.23 mmol, 78%). $R_f$ = 0.53 (9:1 EtOAc–MeOH); $[\alpha]_D$ –154 (*c* 0.56, MeOH); $^1$H NMR (500 MHz, $D_2O$) δ 7.38 (ddd, *J* = 8.0, 7.5, 0.5 Hz, 2H, Ar), 7.14–7.10 (m, 3H, Ar), 5.65 (s, *J* < 1.0 Hz, 1H, H-1), 4.33 (dd, *J* = 4.0, 1.0 Hz, H, H-2), 4.18 (dd, *J* = 6.0, 4.0 Hz, 1H, H-3), 4.10 (dd, *J* = 6.0, 4.0 Hz, 1 H, H-4), 3.83 (ddd, *J* = 8.0, 4.0, 4.0 Hz, 1H, H-5), 3.66 (dd, *J* = 12.0, 4.5 Hz, 1H, H-6), 3.61 (dd, *J* = 12.0, 8.0 Hz, 1H, H-6′). $^{13}$C NMR (125 MHz, $D_2O$) δ 155.8 (Ar), 130.0 (Ar), 123.2 (Ar), 117.3 (Ar), 106.0 (C-1), 83.6 (C-4), 81.1 (C-2), 76.4 (C-3), 70.7 (C-5), 62.7 (C-6). HRMS (ESI) *m/z* Calcd for $C_{12}H_{16}O_6Na$ $[M+Na]^+$: 279.0839. Found: 279.0835.

***p*-Nitrophenyl β-D-galactofuranoside (6b)**

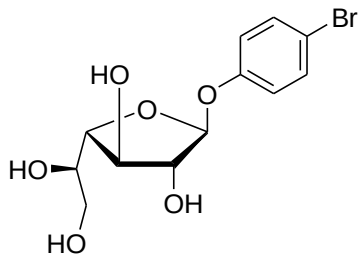



**6b**

Synthesized using general synthetic procedure (27.2 mg, 0.09 mmol, 31%). NMR data was consistent with previously published spectra.[11] $R_f$ = 0.31 (9:1 EtOAc–

MeOH); $[\alpha]_D$ –168 (*c* 0.48, MeOH); $^1$H NMR (500 MHz, $D_2O$) δ 8.23 (d, *J* = 9.0 Hz, 2H, Ar), 7.21 (d, *J* = 9.0 Hz, 2H, Ar), 5.79 (s, *J* < 1.0 Hz, 1H, H-1), 4.41 (dd, *J* = 4.0, 1.0 Hz, 1H, H-2), 4.22 (dd, *J* = 6.0, 4.0 Hz, 1H, H-3), 4.10 (dd, *J* = 6.0, 4.5 Hz, 1H, H-4), 3.86 (ddd, *J* = 8.0, 4.5, 4.5 Hz, 1H, H-5), 3.69–3.61 (m, 2H, H-6,6′). $^{13}$C NMR (125 MHz, $D_2O$) δ 161.4 (Ar), 142.2 (Ar), 126.1 (Ar), 116.7 (Ar), 105.5 (C-1), 84.2 (C-4), 81.2 (C-2), 76.5 (C-3), 70.6 (C-5), 62.7 (C-6). HRMS (ESI) *m/z* Calcd for $C_{12}H_{15}NO_8Na$ $[M+Na]^+$: 324.069. Found: 324.0687.

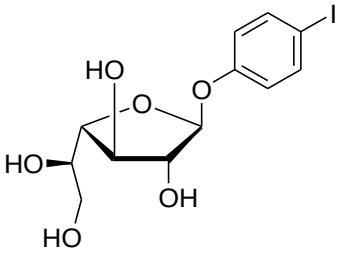
***p*-Bromophenyl β-D-galactofuranoside (6c)**




**6c**

Synthesized using general synthetic procedure (14.5 mg, 0.043 mmol, 60%). $R_f$ = 0.49 (9:1 EtOAc–MeOH); $[\alpha]_D$ –134 (*c* 0.50, MeOH); $^1$H NMR (500 MHz, $CD_3OD$) δ 7.38 (d, *J* = 8.5 Hz, 2H, Ar), 6.98 (d, *J* = 8.5 Hz, 2H, Ar), 5.47 (d, *J* = 1.5 Hz, 1H, H-1), 4.21 (dd, *J* = 4.0, 1.5 Hz, 1H, H-2), 4.13 (dd, *J* = 6.5, 4.0 Hz, 1H, H-3), 4.06 (dd, *J* = 6.5, 3.0 Hz, 1H, H-4), 3.73 (ddd, *J* = 6.5, 6.5, 3.0 Hz, 1H, H-5), 3.59 (d, *J* = 6.5 Hz, 2H, H-6,6′); $^{13}$C NMR (125 MHz, $CD_3OD$) δ 157.7 (Ar), 133.3 (Ar), 119.8 (Ar), 115.2 (Ar, C-Br), 108.0 (C-1), 85.0 (C-4), 83.5 (C-2), 78.2 (C-3), 72.2 (C-5), 64.3 (C-6); HRMS (ESI) *m/z* Calcd for $C_{12}H_{15}BrO_6Na$ $[M+Na]^+$: 356.9944. Found: 356.9934.

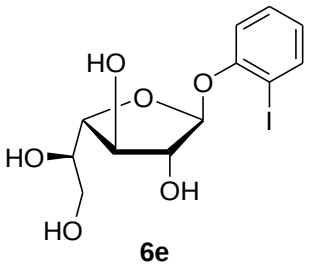
***p*-Iodophenyl β-D-galactofuranoside (6d)**




**6d**

Synthesized using general synthetic procedure (55.4 mg, 0.145 mmol, 50%). $R_f$ = 0.40 (9:1 EtOAc–MeOH); $[\alpha]_D$ –127 (*c* 0.53, MeOH); $^1$H NMR (500 MHz, $CD_3OD$) δ 7.56 (d, *J* = 9.0 Hz, 2H, Ar), 6.87 (d, *J* = 9.0 Hz, 2H, Ar), 5.47 (d, *J* = 2.0 Hz, 1H, H-1), 4.21 (dd, *J* = 4.0, 2.0 Hz, 1H, H-2), 4.13 (dd, *J* = 6.5, 4.5 Hz, 1H, H-3), 4.05 (dd, *J* = 6.5, 3.0 Hz, 1H, H-3), 3.73 (ddd, *J* = 6.5, 6.5, 3.0 Hz, 1H, H-5), 3.59 (d, *J* = 6.0 Hz, 2H, H-6,6′); $^{13}$C NMR (125 MHz, $CD_3OD$) δ 158.4 (Ar), 139.4 (Ar), 120.2 (Ar), 107.8 (C-1), 85.0 (C-4), 84.9 (Ar, C-I), 83.5 (C-2), 78.2 (C-3), 72.2 (C-5), 64.3 (C-6); HRMS (ESI) *m/z* Calcd for $C_{12}H_{15}IO_6Na$ $[M+Na]^+$: 404.9806. Found: 404.9799.

***o*-Iodophenyl β-D-galactofuranoside (6e)**



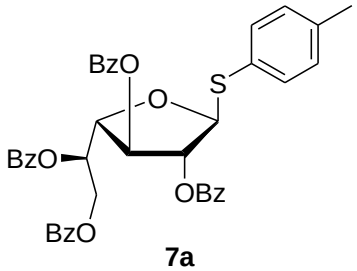

**6e**

Synthesized using general synthetic procedure (19.2 mg, 0.050 mmol, 36%). $R_f$ = 0.44 (9:1 EtOAc–MeOH); $[\alpha]_D$ –84 (*c* 0.48, MeOH); $^1$H NMR (500 MHz, $D_2O$) δ 7.93 (d, *J* = 8.0 Hz, 1H, Ar), 7.46 (dd, *J* = 8.0, 7.5 Hz, 1H, Ar), 7.23 (d, *J* = 8.0

Hz, 1H, Ar), 6.97 (dd, *J* = 8.0, 7.5 Hz, 1H, Ar), 5.69 (app s, 1H, H-1), 4.52 (dd, *J* = 4.0, 1.5 Hz, 1H, H-2), 4.30 (dd, *J* = 6.0, 4.0 Hz, 1H, H-4), 4.25 (dd, *J* = 6.0, 4.0 Hz, 1H, H-3), 3.91 (ddd, *J* = 7.5, 4.5, 4.0 Hz, 1H, H-5), 3.74 (dd, *J* = 11.5, 4.5 Hz, 1H, H-6), 3.67 (dd, *J* = 11.5, 7.5 Hz, 1H, H-6′); $^{13}$C NMR (125 MHz, $D_2O$) δ 155.1 (Ar), 139.7 (Ar), 130.1 (Ar), 125.3 (Ar), 117.8 (Ar), 107.5 (C-1), 88.0 (Ar, C-I), 83.8 (C-4), 81.4 (C-2), 76.5 (C-3), 70.5 (C-5), 62.8 (C-6); HRMS (ESI) *m/z* Calcd for $C_{12}H_{15}IO_6Na$ $[M+Na]^+$: 404.9806. Found: 404.9797.

.

**_p_-Tolyl 2,3,5,6-tetra-_O_-benzoyl-1-thio-β-D-galactofuranoside (7a)**

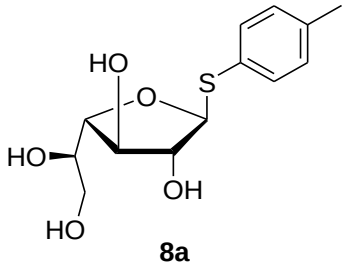



**7a**

Synthesized as described previously.[9] Perbenzoylated galactofuranose (**4**) (590 mg, 0.843 mmol) and *p*-thiocresol (137 mg, 1.10 mmol) were dissolved in DCM (5 mL) and cooled to 0 °C. $BF_3{\cdot}OEt_2$ (700 μL, 5.5 mmol) was slowly added, and the mixture was stirred overnight while warming to room temperature. $Et_3N$ (500 μL) was added to quench the reaction, and then the mixture was diluted in DCM (45 mL) and sequentially washed with $NaHCO_3$ saturated solution (50 mL), water (50 mL), and brine (25 mL). Column chromatography (6:1 hexane–EtOAc) yielded *p*-tolyl 2,3,5,6-tetra-*O*-benzoyl-1-thio-β-D-galactofuranoside **7a** (467.8 mg, 0.65 mmol, 77%) as a clear syrup which solidified upon standing. NMR data

was consistent with previously published spectra.[9] $R_f$ = 0.20 (4:1 hexane–EtOAc); $^1$H NMR (500 MHz, $CDCl_3$) δ 8.09–7.06 (m, 24H, Ar), 6.10 (ddd, *J* = 6.5, 4.0, 4.0 Hz, 1H, H-5), 5.77 (s, 1H, H-1), 5.70 (d, *J* = 5.0 Hz, 1H, H-3), 5.65 (s, 1H, H-2), 4.95 (dd, *J* = 5.0, 4.0 Hz, 1H, H-4), 4.76 (dd, *J* = 11.5, 4.5 Hz, 1H, H-6), 4.70 (dd, *J* = 11.5, 7.0 Hz, 1H, H-6′), 2.31 (s, 3H, Ar-$CH_3$); $^{13}$C NMR (125 MHz, $CDCl_3$) δ 166.1, 165.7, 165.6, 165.4 (4 x PhC=O), 138.3-128.3 (aromatic C), 91.6 (C-1), 82.4 (C-2), 81.5 (C-4), 78.0 (C-3), 70.4 (C-5), 63.5 (C-6), 21.2 (Ar-$CH_3$).

***p*-Tolyl 1-thio-β-D-galactofuranoside (8a)**

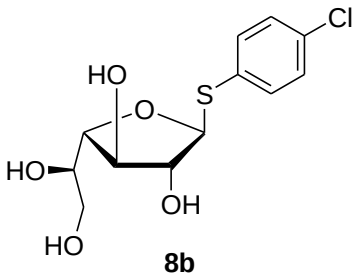



**8a**

Synthesized as described previously.9 Compound **7a** was dissolved in MeOH (2.5 mL) and catalytic NaOMe added until the solution became strongly basic. Once TLC analysis indicated the complete disappearance of the starting material, the mixture was neutralized with Amberlite IR-120 ($H^+$ form) resin, filtered through Celite, and concentrated to dryness. Column purification (100% EtOAc to 95:5 EtOAc–MeOH) yielded the deprotected glycoside **8a** as a white solid (26.2 mg, 0.0915 mmol, 83%). NMR data was consistent with previously published spectra.[9] $R_f$ = 0.44 (9:1 EtOAc–MeOH); $^1$H NMR (500 MHz, $D_2O$) δ 7.50 (d, *J* = 8.0 Hz, 2H, Ar), 7.30 (d, *J* = 8.0 Hz, 2H, Ar), 5.23 (d, *J* = 6.0 Hz, 1H, H-1), 4.15 (dd, *J* = 7.5, 6.0 Hz, 1H, H-3), 4.05 (dd, *J* = 6.0, 6.0 Hz, 1H, H-2), 3.90 (dd, *J* =

8.0, 3.5 Hz, 1H, H-4), 3.83 (ddd, *J* = 7.5, 4.5, 3.5 Hz, 1H, H-5), 3.69 (dd, *J* = 12.0, 5.0 Hz, 1H, H-6), 3.65 (dd, *J* = 12.0, 7.5 Hz, 1H, H-6′), 2.23 (s, 3H, Ar-$CH_3$); $^{13}$C NMR (125 MHz, $D_2O$) δ 139.4 (Ar), 133.4 (Ar), 130.1 (Ar), 128.0 (Ar), 90.3 (C-1), 81.4 (C-4), 79.7 (C-2), 75.5 (C-3), 70.3 (C-5), 62.8 (C-6), 20.2 (Ar-$CH_3$).

***p*-Chlorophenyl 1-thio-β-D-galactofuranoside (8b)**

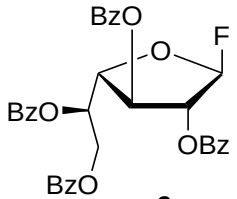



**8b**

Perbenzoylated galactofuranose (**4**) (200 mg, 0.286 mmol) and 4-chlorothiophenol (53.7 mg, 0.371 mmol) were stirred under argon in DCM (5 mL) at 0 °C. $BF_3{\cdot}OEt_2$ (250 μL, 1.6 mmol) was added, and the reaction allowed to warm to room temperature overnight. The reaction was quenched by the addition of 10 mL $NaHCO_3$ saturated solution, and the solution stirred vigorously until gas evolution ceased. The solution was diluted in DCM (50 mL) and washed successively with $NaHCO_3$ solution (50 mL), $H_2O$ (50 mL), and brine (25 mL), then dried over $Na_2SO_4$, filtered, and solvent evaporated. The crude material was dissolved in MeOH (5 mL) and NaOMe added until the solution turned strongly basic. Once complete, the solution was neutralized with Amberlite IR-120 ($H^+$ form) resin, filtered through Celite, and sovent evaporated. Column purification (100% EtOAc to 95:5 EtOAc–MeOH) afforded the pure compound **8b** as a white solid (61.1 mg, 0.20 mmol, 54%). $R_f$ = 0.51 (9:1 EtOAc–MeOH); $[\alpha]_D$ –116 (*c*

0.37, $H_2O$); $^1H$ NMR (600 MHz, $CDCl_3$) δ 7.55 (d, *J* = 8.4 Hz, 2H, Ar), 7.45 (d, *J* = 8.4 Hz, 2H, Ar), 5.29 (d, *J* = 6.0 Hz, 1H, H-1), 4.16 (dd, *J* = 7.8, 6.0 Hz, 1H, H-3), 4.06 (dd, *J* = 6.0, 6.0 Hz, 1H, H-2), 3.93 (dd, *J* = 7.8, 3.0 Hz, 1H, H-4), 3.83 (ddd, *J* = 7.8, 4.8, 3.0 Hz, 1H, H-5), 3.68 (dd, *J* = 11.4, 4.8 Hz, 1H, H-6), 3.64 (dd, *J* = 11.4, 7.2 Hz, 1H, H-6′); $^{13}C$ NMR (125 MHz, $CDCl_3$) δ 134.1 (Ar), 134.1 (Ar), 130.7 (Ar), 129.3 (Ar), 90.1 (C-1), 81.6 (C-4), 79.9 (C-2), 75.5 (C-3), 70.3 (C-5), 62.7 (C-6); HRMS (ESI) *m/z* Calcd for $C_{12}H_{15}O_5SClNa$ $[M+Na]^+$: 329.0221. Found: 329.0219.

**2,3,5,6-Tetra-*O*-benzoyl-β-D-galactofuranosyl fluoride (9)**

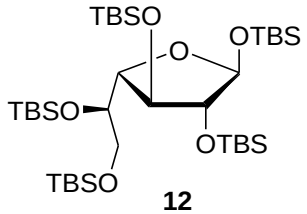



**9**

70% HF–pyridine solution (1.5 mL) was added to perbenzoylated galactofuranose (**4**) (200 mg, 0.286 mmol) stirring in toluene (3 mL) at room temperature. After 3 days, the mixture was quenched by the addition of 10 mL $NaHCO_3$ saturated solution. DCM (20 mL) was added and the solution washed with $CaCl_2$ saturated solution (2 x 20 mL), $NaHCO_3$ saturated solution (2 x 20 mL), and brine (1 x 20 mL). Drying over $Na_2SO_4$, filtration, and concentration under reduced pressure produced a clear oil, which was purified by column chromatography (9:1 hexane–EtOAc) to obtain 149 mg of the desired product **9** (0.248 mmol, 87%, 1:4 α:β) as separable anomers. NMR data was consistent with previously published

spectra.[21] Data for β anomer: $R_f$ = 0.28 (9:1 EtOAc–MeOH); $^1$H NMR (600 MHz, $CDCl_3$) δ 8.07–7.87 (m, 8H, Ar), 7.61–7.27 (m, 12H, Ar), 6.12–6.08 (m, 1H, H-5), 6.04 (d, *J* = 60 Hz, 1H, H-1), 5.69 (d, *J* = 4.5 Hz, 1H, H-2), 5.67 (d, *J* = 6.5 Hz, 1H, H-3), 4.91 (td, *J* = 4.5, 1.5 Hz, 1H, H-5), 4.80 (dd, *J* = 12.0, 4.5 Hz, 1H, H-6), 4.73 (dd, *J* = 12.0, 7.0 Hz, 1H, H-6′); $^{13}$C NMR (125 MHz, $CDCl_3$) δ 166.1, 165.7, 165.6, 165.2 (4 x *CO*$CH_3$), 133.7–128.7 (Ar), 112.5 (d, *J* = 225 Hz, C-1), 84.9 (C-4), 81.0 (d, *J* = 40 Hz, C-2), 76.6 (C-3), 70.3 (C-5), 63.4 (C-6); $^{19}$F NMR (469 MHz, $CDCl_3$) δ –122.1 (dd, *J* = 58.0, 6.3 Hz). HRMS (ESI) *m/z* Calcd for $C_{34}H_{27}O_9FNa$ $[M+Na]^+$: 621.1531. Found: 621.1530.

**1,2,3,5,6-Penta-*O*-*tert*-butyldimethylsilyl-β-D-galactofuranose (12)**

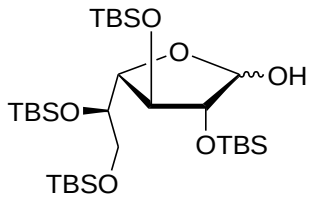



**12**

Synthesized as described previously.[14,15] TBSCl (3.0 g, 19.9 mmol) and imidazole (2.74 g, 40.2 mmol) were added to D-galactose (**3**) (500 mg, 2.77 mmol) stirring in dry DMF (14 mL) at 30 °C. The reaction continued for 2 days, and then the solution was poured into ice water (50 mL). DCM (50 mL) was added, and the solution washed with 50 mL portions of 1M HCl, $H_2O$, $NaHCO_3$ saturated solution, $H_2O$, and brine. The organic layer was dried over $Na_2SO_4$, filtered, and concentrated to a clear oil containing a mixture of furanoses and hexoses. The crude mixture was purified by column chromatography (100:1 hexane–EtOAc) to

yield 443 mg of pure β-furanose **12** (0.59 mmol, 21%) plus another fraction containing a mixture of β-furanose **12** and β-pyranose **13** (1.23 g, 1.64 mmol, 59%). NMR data was consistent with previously published spectra.[14] $R_f$ = 0.71 (9:1 hexane–EtOAc); $^1$H NMR (500 MHz, $CDCl_3$) δ 5.15 (d, *J* = 2.4 Hz, 1H, H-1), 4.09 (dd, *J* = 4.7, 2.9 Hz, 1H, H-3), 4.00 (dd, *J* = 4.7, 3.5 Hz, 1H, H-4), 3.92 (dd, *J* = 3.0, 3.0 Hz, 1H, H-2), 3.74 (td, *J* = 6.0, 3.5 Hz, 1H, H-5), 3.67 (dd, *J* = 9.9, 6.3 Hz, 1H, H-6), 3.55 (dd, *J* = 9.9, 5.7 Hz, 1H, H-6′), 0.93–0.84 (m, 45H, -SiC(*CH*$_3$)), 0.12–0.03 (m, 28H, -Si(*CH*$_3$)); $^{13}$C NMR (125 MHz, $CDCl_3$) δ 102.9 (C-1), 86.0 (C-2), 84.5 (C-4), 79.6 (C-3), 74.2 (C-5), 64.8 (C-6), 26.1–25.8 (SiC(*CH*$_3$)), 18.4–17.9 (Si*C*($CH_3$)), –3.7–(–5.3) (Si*CH*$_3$).

**2,3,5,6-Tetra-*O*-*tert*-butyldimethylsilyl-α,β-D-galactofuranose (15)**




**15**

Compound **12** (120 mg, 0.160 mmol) was stirred in DCM with 4 Å molecular sieves for 30 min. at room temperature then cooled to 0 °C. TMSI (25 μL, 0.20 mmol) was added via syringe and stirred until the reaction mixture turned a bright salmon pink color, indicative of the formation of **14**. A solution of DIPEA (32 μL, 0.20 mmol) and water (120 μL) were added, and the mixture stirred a further 1 h. The mixture was diluted in DCM and washed twice with $NaHCO_3$ saturated solution (50 mL), once with brine (50 mL), then dried over $Na_2SO_4$, filtered, and

concentrated to dryness. The crude material was purified using 100:1 hexane–EtOAc to obtain 85.6 mg of a 1.2:1 α:β mixture of **15** (0.13 mmol, 84%). NMR data was consistent with previously published spectra.[14] $R_f$ = 0.60, 0.51 (9:1 hexane–EtOAc); $^1$H NMR (600 MHz, $CDCl_3$) δ 5.07 (d, *J* = 11.5 Hz, 1H, H-1β), 5.06 (dd, *J* = 11.8, 4.6 Hz, 1H, H-1α), 4.21-4.19 (m, 2H, 2β, 3α), 4.09 (dd, *J* = 8.4, 1.3 Hz, 1H, H-4β), 4.03 (dd, *J* = 4.7, 4.7 Hz, 1H, H-2α), 4.01 (dd, *J* = 3.8, 2.4 Hz, 1H, H-4α), 3.97 (d, *J* = 1.0 Hz, 1H, H-3β), 3.85 (d, *J* = 11.8 Hz, 1H, OH-α), 3.77 (ddd, *J* = 8.2, 5.3, 4.7 Hz, 1H, H-5β), 3.74-3.69 (m, 2H, H-5α, H-6α), 3.66 (dd, *J* = 10.7, 5.5 Hz, 1H, H-6β), 3.62-3.55 (m, 2H, H-6α′, H-6β′), 3.51 (d, *J* = 11.5 Hz, 1H, OH-β), 0.94 – 0.88 (m, 72H, -SiC(*CH*$_3$)), 0.20 – 0.05 (m, 48H, -Si(*CH*$_3$)); $^{13}$C NMR (126 MHz, $CDCl_3$) δ 103.4 (C-1β), 97.2 (C-1α), 90.0 (C-2β), 84.2 (C-2α), 82.0 (C-4β), 81.7 (C-4α), 78.3 (C-3α), 77.6 (C-3β), 73.9 (C-5β), 73.5 (C-5α), 66.3 (C-6β), 63.9 (C-6α), 26.2–25.7 (SiC(*CH*$_3$)), 18.6–17.8 (Si*C*($CH_3$)), –3.9–(–5.3) (Si*CH*$_3$).

**Site-Directed Mutagenesis and Expression of GlfT2 Mutants**

The wild-type *M. tuberculosis Rv3808c* gene inserted into a pET-15b vector and containing a C-terminal $His_6$-tag was used as a template.[5c] Site-directed mutagenesis to form the H413A_Y344A, E300A_Y344A, E300A_H413A, H413A_Y344A_E300A, and W309S_W399S mutants was carried out using the QuikChange II XL kit (Stratagene, USA) and mutations confirmed by DNA sequencing. The W309S and W399S mutants had been previously expressed (Zheng, unpublished results). The following primers were used:

H413A: 5′-GCAATTGATTGGCAGGCTTATTTTGCTCTGCGTAACCGC-3′

5′-GCGGTTACGCAGAGCAAAATAAGCCTGCCAATCAATTGC-3′

E300A: 5′-CTGCACATCATGGGTGCAGTTGTCGACCGTTC-3′

5′-GAACGGTCGACAACTGCACCCATGATGTGCAG-3′

Y344A: 5′-CGTCGCATCGACGTCGATGCTAATGGTTGGTGGACCTG-3′

5′-CAGGTCCACCAACCATTAGCATCGACGTCGATGCGACG-3′

W309S: 5′-CGACCGTTCTATTTTTATGTCGACCGCAGCTCCGCATGCG-3′

5′-CGCATGCGGAGCTGCGGTCGACATAAAAATAGAACGGTCG-3′

W399S: 5′-CTGGCACATGGCGTCGTCTGACAAAGACG-3′

5′-CGTCTTTGTCAGACGACGCCATGTGCCAG-3′

Mutant plasmids were transformed into *E. coli* BL21 DE3 competent cells, and the protein expressed and purified using a Ni-NTA agarose column as described previously.[17]

**Enzymatic Reactions of GlfT2 Mutants**

Mutant GlfT2 enzymes (200 μg) were incubated with 1 mM UDP and 1 mM phenyl donor in 200 μL MOPS (100 mM, pH = 7.6) at room temperature with rotation for 1 h or 24 h. Aliquots (25 μL) were taken, diluted 5x in MOPS, centrifuged through an Amicon centrifugal filter unit (MW cutoff = 10K), and analyzed by HPLC (see conditions below).

**Analytical HPLC Conditions**

Buffer A: 200 mM triethylammonium acetate (TEAA), pH = 6.6

Buffer B: 200 mM TEAA, pH = 6.6, containing 5% acetonitrile

Column: Analytical reverse phase C18 (4.6 x 250 mm), Phenomenex, P/N: 00G-4252-E0

Flow Rate: 0.8 mL/min

Detection: Photodiode array, $\lambda$ = 262 nm

| **Time (min)** | **Flow Rate (mL/min)** | **%A** | **%B** |
|---|---|---|---|
| 0 | 0.8 | 96 | 4 |
| 10 | 0.8 | 96 | 4 |
| 25 | 0.8 | 0 | 100 |
| 35 | 0.8 | 0 | 100 |
| 36 | 0.8 | 96 | 4 |
| 5 | 0.8 | 96 | 4 |

**Enzymatic Reaction with WbbI**

Expression and purification of WbbI-maltose binding fusion protein, as well as the enzyme-mediated reaction, were carried out as described previously.[18] WbbI–MBP (40 μg) was incubated in 50 mM MOPS (100 μL) with UDP (4 mM) and one of donors **6b**, **6c**, or **8a** (4 mM) and monitored at room temperature with rotation over the course of 3 h. After this time, aliquots were filtered as described above for the GlfT2 mutants and analyzed by HPLC (200 mM TEAA, pH = 6.8, containing 1.5% acetonitrile, flow rate = 0.8 mL/min.).

## 2.6 Bibliography


(1) Gantt, R. W.; Peltier-Pain, P.; Cournoyer, W. J.; Thorson, J. S. *Nat. Chem. Biol.* **2011**, *7*, 685–691.

(2) Gantt, R. W.; Peltier-Pain, P.; Singh, S.; Zhou, M.; Thorson, J. S. *P. Nat. Acad. Sci. USA* **2013**, *110*, 7648–7653.

(3) (a) Williams, G. J.; Zhang, C.; Thorson, J. S. *Nat. Chem. Bio.* **2007**, *3*, 657–662. (b) Gantt, R. W.; Goff, R. D.; Williams, G. J.; Thorson, J. S. *Angew. Chem. Int. Edit.* **2008**, *47*, 8889–8892. (c) Yang, M.; Proctor, M. R.; Bolam, D. N.; Errey, J. C.; Field, R. A.; Gilbert, H. J.; Davis, B. G. *J. Am. Chem. Soc.* **2005**, *127*, 9336–9337.

(4) a) Zhang, C.; Griffith, B. R.; Fu, Q.; Albermann, C.; Fu, X.; Lee, I.-K.; Li, L.; Thorson, J. S. *Science* **2006**, *313*, 1291–1294. b) Errey, J. C.; Mukhopadhyay, B.; Kartha, K. P. R.; Field, R. A. *Chem. Commun.* **2004**, *23*, 2706−2707. c) Lairson, L. L.; Watts, A. G.; Wakarchuk, W. W.; Withers, S. G. *Nat. Chem. Biol.* **2006**, *2*, 724–728. d) Oberthuer, M.; Leimkuhler, C.; Kruger, R. G.; Lu, W.; Walsh, C. T.; Kahne, D. *J. Am. Chem. Soc.* **2005**, *127*, 10747–10752.

(5) (a) Wheatley, R. W.; Zheng, R. B.; Richards, M. R.; Lowary, T. L.; Ng, K. K. S. *J. Biol. Chem.* **2012**, *287*, 28132−28143. (b) May, J. F.; Levengood, M. R.; Splain, R. A.; Brown, C. D.; Kiessling, L. L. *Biochemistry* **2012**, *51*, 1148−1159. (c) Rose, N. L.; Completo, G. C.; Lin, S.-J.; McNeil, M.; Palcic, M. M.; Lowary, T. L. *J. Am. Chem. Soc.* **2006**, *128*, 6721−6729.

(6) Recent reviews: (a) Kwan, D. H.; Withers, S. G. *J. Carbohyd. Chem.* **2011**, *30*, 181–205. (b) Shaikh, F. A.; Withers, S. G. *Biochem. Cell Biol.* **2008**, *86*, 169–177.

(7) Imamura, A.; Lowary, T. L. *Trends Glycosci. Glyc.* **2011**, *23*, 134–152.

(8) Szarek, W. A.; Zamojski, A.; Tiwari, K. N.; Ison, E. R. *Tetrahedron Lett.* **1986**, *27*, 3827–3830.

(9) Completo, G. C.; Lowary, T. L. *J. Org. Chem.* **2008**, *73*, 4513–4525.

(10) Marino, C.; Gandolfi-Donadio, L.; Rodriguez, C. G.; Bai, Y.; de Lederkremer, R. M. In *Carbohydrate Chemistry: Proven Synthetic Methods*; Kováč, P., Ed.; CRC Press: Boca Raton, FL, 2012; pp 231–238.

(11) Varela, O.; Marino, C.; de Lederkremer, R. M. *Carbohyd. Res.* **1986**, *155*, 247–251.

(12) Lougheed, B.; Ly, H. D.; Wakarchuk, W. W.; Withers, S. G. *J. Biol. Chem.* **1999**, *274*, 37717–37722.

(13) Hayashi, M.; Hashimoto, S.; Noyori, R. *Chem. Lett.* **1984**, 1747–1750.

(14) Baldoni, L.; Marino, C. *J. Org. Chem.* **2009**, *74*, 1994–2003.

(15) Dureau, R.; Legentil, L.; Daniellou, R.; Ferrières, V. *J. Org. Chem*. **2012**, *77*, 1301–1307.

(16) Posner, G. H.; Haines, S. R. *Tetrahedron Lett.* **1985**, *26*, 5–8.

(17) Rose, N. L.; Zheng, R. B.; Pearcey, J.; Zhou, R.; Completo, G. C.; Lowary, T. L. *Carbohydr. Res.* **2008**, *343*, 2130–2139.

(18) Wing, C.; Errey, J. C.; Mukhopadhyay, B.; Blanchard, J. S.; Field, R. A. *Org. Biomol. Chem.* **2006**, *4*, 3945−3950.

(19) Bolam, D. N.; Roberts, S.; Proctor, M. R.; Turkenburg, J. P.; Dodson, E. J.; Martinez-Fleites, C.; Yang, M.; Davis, B. G.; Davies, G. J.; Gilbert, H. J. *P. Natl. Acad. Sci. USA* **2007**, *104*, 5336–5341.

(20) Yang, M.; Proctor, M. R.; Bolam, D. N.; Errey, J. C.; Field, R. A.; Gilbert, H. J.; Davis, B. G. *J. Am. Chem. Soc.* **2005**, *127*, 9336–9337.

(21) Bock, K.; Pedersen, C.; Wiebe, L. *Acta Chem. Scand.* **1973**, *27*, 3586–3590.

# Chapter 3 : Synthesis of Nitrogen-Containing Furanose Sugar Nucleotides for Use as Enzymatic Probes[1,2]

[1] A version of this chapter has been published:
Snitynsky, R. B.; Lowary, T. L. *Org. Lett.* **2014**, *16*, 212-215.
[2] H413S and W309S GlfT2 mutant enzymes were prepared by Ruixiang Blake Zheng. K311A and E317A mutant GalPUT enzymes were prepared and assayed by Mengyao (Aaron) Dong.

## 3.1 Introduction and Project Aim

The inhibition of cell wall biosynthesis to treat bacterial infections remains a promising area of study, given that a number of bacterial pathogens contain within their cell walls five-membered ring (furanose) sugars not found in mammals.[1,2,3] One such pathogen, *Campylobacter jejuni*, is known to produce capsular polysaccharides (CPS) containing unusual sugars,[4,5] although the genes responsible for cell wall biosynthesis in this organism have not been isolated. One well-studied isolate of *C. jejuni*, NCTC11168 (serostrain HS:2), contains the repeating tetrasaccharide unit shown in Figure 3-1.[6,7]

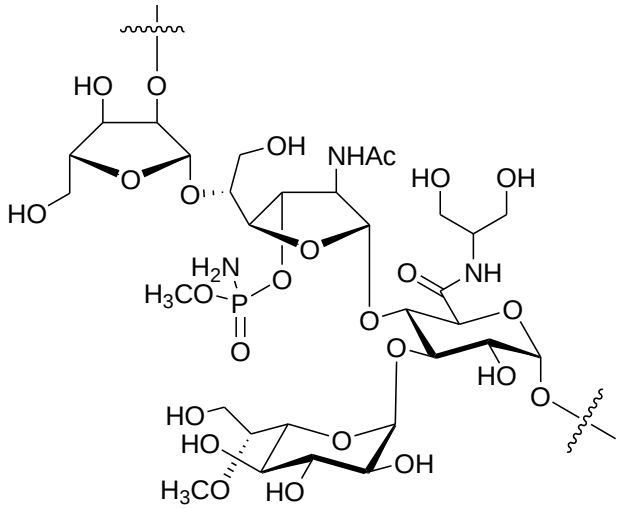
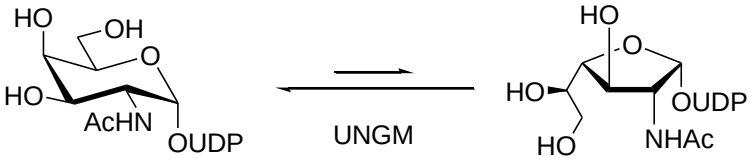



**Figure 3-1.** Structure of a repeating tetrasaccharide from *C. jejuni* NCTC 11168 (HS:2).

Presumably, the biosynthetic precursors of most of the individual residues that compose this tetrasaccharide are sugar nucleotides, incorporated into the final structure through the action of glycosyltransferase enzymes. To aid in the characterization of these and other putative enzymes involved in cell wall biosynthesis, milligram quantities of the sugar nucleotide precursors would be

beneficial. For example, UDP-2-acetamido-2-deoxy-α-D-galactofuranose (UDP-Gal*f*NAc, **2**) was enzymatically synthesized during the characterization of a pyranose-furanose mutase from *C. jejuni*.[8] However, the equilibrium ratio strongly favors the pyranose form (UDP-Gal*p*NAc), resulting in a 93:7 ratio of pyranose–furanose and a mere 4% overall isolated yield of the furanose (Scheme 3-1).




**UDP-Gal*p*NAc** **2**

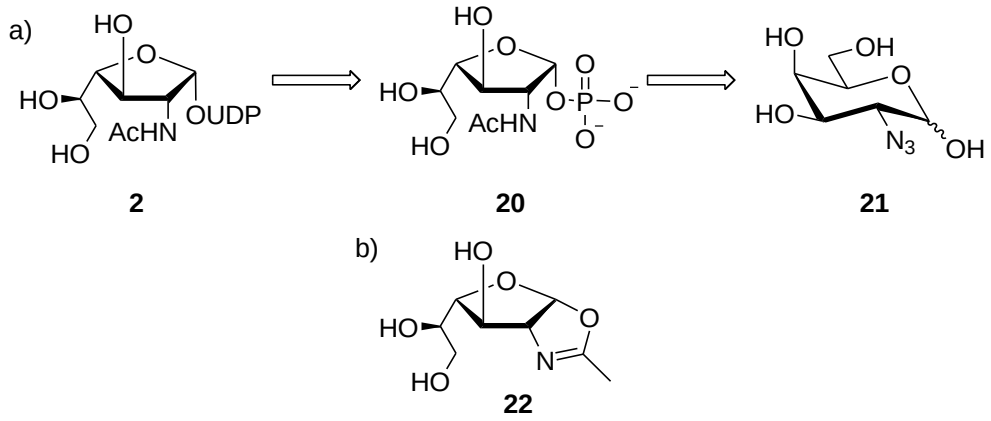
**Scheme 3-1.** Reaction catalyzed by UDP-galactopyranose mutase (UNGM). The equilibrium lies 93:7 in favor of the pyranose form.

Clearly, a more robust method of obtaining UDP-Gal*f*NAc (**2**) would be beneficial for future biosynthetic studies, and its synthesis was set as the primary goal of this project. Additionally, UDP-Gal*f*NAc (**2**) analogues that could be useful for biochemical investigations were also explored.

### 3.2 Synthesis of Nitrogen-Containing Furanose Sugar-1-Phosphates

Sugar nucleotides have most often been obtained through sugar-1-phosphates. Both synthetic[9] and chemoenzymatic[10] methods make use of these compounds, as the stereochemistry about the anomeric position is set earlier in the synthesis. Gal*f*-1-phosphate has been synthesized previously by converting D-

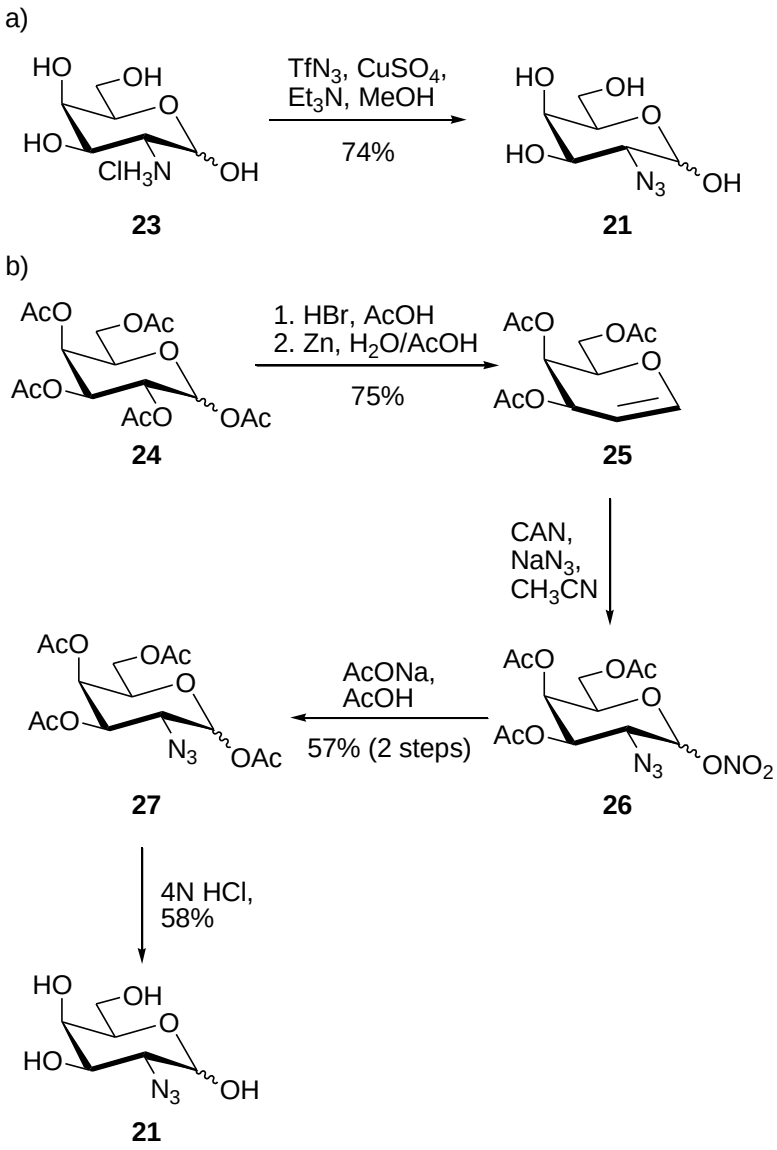
galactose to its furanose form, followed by bromination of the perbenzoylated sugar to form the galactofuranosyl bromide.[11] This compound can then be phosphorylated with dibenzyl phosphate and deprotected to yield Gal*f*-1-phosphate. A similar synthesis was envisioned to produce Gal*f*NAc-1-phosphate (**20**) (Scheme 3-2a), which could then be coupled to UMP to form **2**. However, the synthesis of **20** is complicated by the fact that the 2-acetamido group will readily form an oxazoline **22** following activation of the anomeric position (Scheme 3-2b). To avoid this obstacle, an azide could be used, which is non-participatory in glycosylation reactions.




**Scheme 3-2.** a) Retrosynthetic analysis of **2**. b) Oxazoline (**22**), resulting from attack of the 2-acetamido group on the anomeric position.

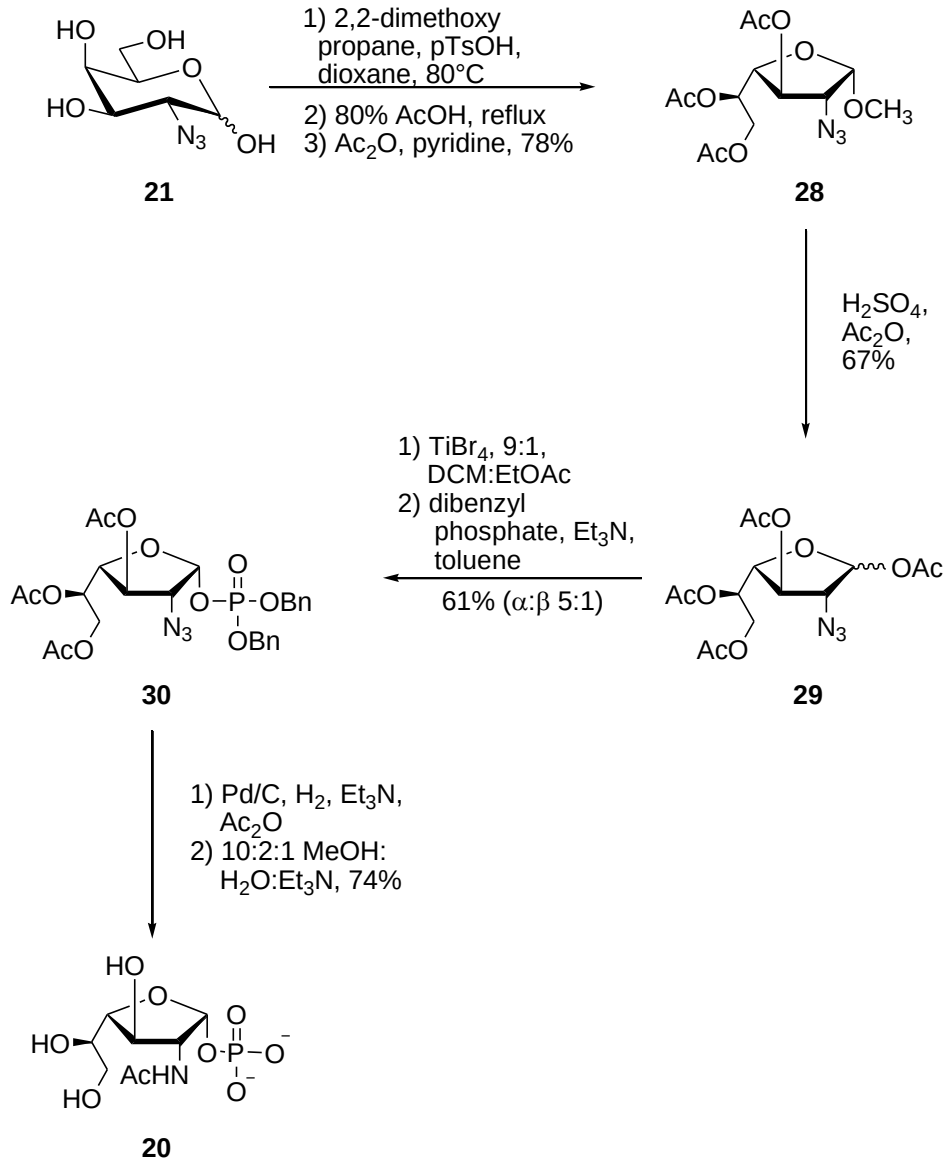
2-Azido-2-deoxy-D-galactopyranose (**21**) can be synthesized from the corresponding amine **23** by way of a copper-catalyzed diazotransfer reaction.[12] Sodium azide and triflic anhydride react to form triflic azide, which is then added to galactosamine (**23**) under basic conditions to form the desired azide (Scheme

3-3a). Although this reaction proceeds with a relatively good yield (74%), the expense of the starting amine salt makes this method impractical for the first step of a multi-step synthesis. Instead, the same compound can be made on a multi-gram scale through the azidonitration of tri-*O*-acetyl-D-galactal (**25**) as reported by Lemieux and co-worker in 1979 (Scheme 3-3b).[13]




**Scheme 3-3.** Synthesis of **21** by a) diazotransfer and b) azidonitration.

Once the azido functionality had been successfully introduced, the next major challenge was to convert the pyranose form of the sugar into the furanose form. Several methods of carrying out this transformation exist, including Fischer glycosylation, high temperature acylation, dithioacetal cyclization, and others, reviewed by Imamura and Lowary in 2011.[14] While low yields and the presence of pyranosides complicates the first two methods, the latter method is extremely slow when an electron-withdrawing acetamido group is positioned at C-2. A literature search uncovered a method that proved to be quite successful, where the starting compound **21** was heated with 2,2-dimethoxypropane and *p*-TsOH to form an isopropylidene-protected methyl furanoside.[15] After removal of the isopropylidene and subsequent acetylation, the desired compound **28** could be isolated in 78% yield over three steps (Scheme 3-4).

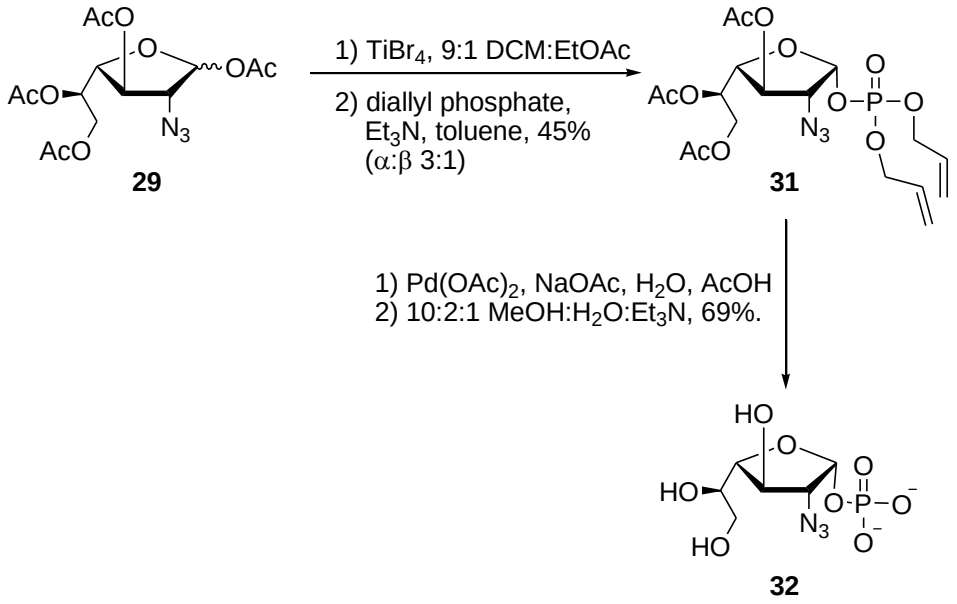



**Scheme 3-4.** Synthesis of Gal*f*NAc-1-phosphate **20**.

Acetolysis of the methyl furanoside gave the fully acetylated compound **29** in 67% yield, which is a suitable substrate for phosphorylation. Typically, galactofuranoses have been phosphorylated by first forming the glycosyl bromide using either HBr or TMSBr, then adding dibenzyl phosphate under basic conditions to displace the bromine.[11] Unfortunately, the azide moiety is not stable

to such harsh conditions, and $TiBr_4$ was chosen as a suitable alternative.[16] The desired glycosyl bromide was formed after stirring for four days with $TiBr_4$, at which time the solvent was changed and dibenzyl phosphate added with triethylamine. Workup and purification yielded the protected phosphate **30** in 61% yield over two steps, with an anomeric ratio of 5:1 α:β. Reduction of the azide and benzyl groups, as well as acetylation of the resultant amine, was carried out in one pot with $H_2$ and Pd/C in acetic anhydride. Finally, the acetyl groups were removed in a methanol–water–triethylamine (10:2:1) solution to yield **20** in 74% over two steps.

While synthesizing **20**, it was thought that retention of the azide moiety in the final sugar nucleotide could prove useful for future biochemical investigations. To retain the azide, orthogonal protection of the phosphate nucleophile was required, and diallyl phosphate was synthesized using a known literature method.[17] This compound was then used in place of dibenzyl phosphate in the phosphorylation of **29** to form the protected sugar-1-phosphate in 45% yield and 3:1 α:β selectivity (Scheme 3-5). Deallylation was carried out with $Pd(OAc)_2$, followed by removal of the acetate groups with methanol-water-triethylamine solution (10:2:1) to yield 2-azido-2-deoxy-α-D-galactofuranose-1-phosphate **32** in 69% yield.

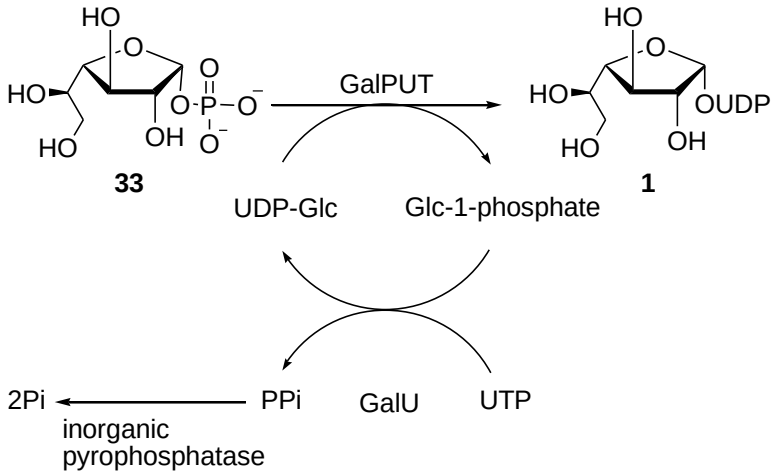



**Scheme 3-5.** Synthesis of Gal*f*$N_3$-1-phosphate **32**.

### 3.3 Attempted Enzymatic Synthesis of UDP-Gal*f*NAc (2)

It is well established that, in many cases, the enzymatic or chemoenzymatic synthesis of sugar nucleotides can result in increased yields and selectivity over chemical synthesis.[3,10,18] For example, the highest reported yield from the chemical synthesis of UDP-Gal*f* (**1**) is 35%,[19] whereas a yield of 79% was achieved by Errey *et al.* using a three enzyme chemoenzymatic approach (Scheme 3-6).[20,21] Briefly, galactose-1-phosphate uridyltransferase (GalPUT) converts Gal*f*-1-phosphate (**33**) to UDP-Gal*f* (**1**) by transferring a UMP unit from catalytic UDP-glucose, which is regenerated by nucleotide transfer from UTP by UDP-glucose pyrophosphorylase (GalU). The pyrophosphate group formed as a side product is then hydrolyzed by inorganic pyrophosphatase (IPP), providing a thermodynamic driving force to sustain the reaction. It was reasoned that a similar

chemoenzymatic synthesis could provide **2** and UDP-Gal*f*$N_3$ **34** in good yield from their corresponding sugar-1-phosphates.[21,22]




**Scheme 3-6.** Field's chemoenzymatic synthesis of **1** using GalPUT.

HPLC analysis of the reaction mixtures following overnight incubation revealed that no conversion of the sugar-1-phosphates **20** or **32** into UDP-sugars took place, and that large amounts of UTP remained in the reaction solution. This result was somewhat surprising, given that GalPUT is known to have relaxed substrate specificity.[20] As well, a previous study has shown that 2-OMe substituted Gal*f*-1-phosphate is a weak substrate of GalPUT.[23] To probe this result further, 2-amino-2-deoxy-α-D-galactofuranosyl-1-phosphate was prepared from the azide by reduction with $H_2$ and Pd/C (see Experimental), and then incubated with the three-enzyme coupled system. Again, no sugar nucleotide formation was observed.

It was suggested that the lack of GalPUT activity could be the result of palladium contamination (Martin Rejzek, personal communication), arising from the deprotection reactions (e.g., conversion of **31** into **32**, Scheme 3-5). Thus, sugar-1-phosphates **20** and **32** were re-purified using a strong anion exchange column (Poros HQ 50 μm resin), washing with water and eluting with 50 mM ammonium bicarbonate buffer. Despite this extra purification step, similar negative results were obtained with both **20** and **32**.

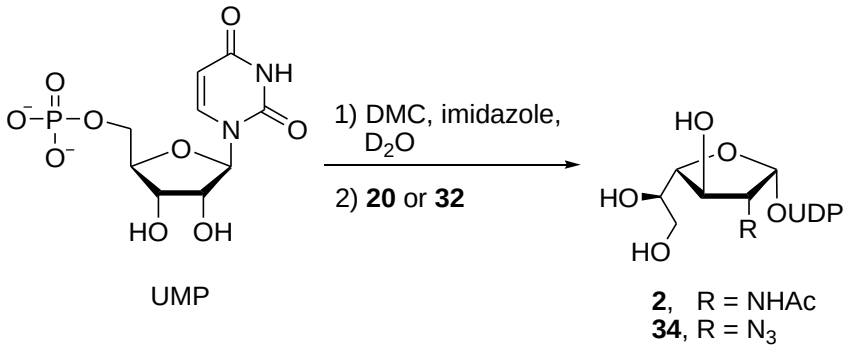
To further probe the active site, several GalPUT mutants were made by site-directed mutagenesis. Based upon the GalPUT crystal structure,[24] three mutants (K311A, F312A, and E317A) targeting residues predicted to be near the C-2 substituent were created (Aaron Dong, unpublished work) and tested with both substrates **20** and **32**. No activity with either substrate was detected by HPLC, and control reactions containing Gal*f*-1-phosphate (**33**) showed significantly reduced activity for each mutant.

### 3.4 Chemical Synthesis of UDP-Gal*f*NAc (2) and UDP-Gal*f*$N_3$ (34)

After the enzymatic synthesis of UDP-Gal*f*NAc (**2**) using GalPUT was shown to be unfeasible, methods of synthesizing this compound chemically were explored. Several methods of coupling nucleotides to sugar-1-phosphates exist as summarized by Wagner *et al.*[9] Using sugar-1-phosphates as the substrates sets the desired anomeric configuration before the coupling reaction, eliminating selectivity issues, but even so yields can be highly variable with lengthy reaction

times. For instance, the yield of the final coupling step between UMP and **33** to form UDP-Gal*f* (**1**) varies between 20–35%, depending upon the method used.[9]

Recently, a method of synthesizing sugar nucleoside diphosphates in $D_2O$ was developed (Scheme 3-7), where UMP is activated as an imidazolide for coupling with sugar-1-phosphates.[25] Although this method is simple to set up, takes place in an aqueous environment, and has yields comparable to other sugar nucleotide coupling methods, it has not yet seen widespread application. To evaluate its potential, this method was selected for the synthesis of UDP-Gal*f*NAc (**2**) and UDP-Gal*f*$N_3$ (**34**).




**Scheme 3-7.** Chemical synthesis of UDP-Gal*f*NAc (**2**) and UDP-Gal*f*$N_3$ (**34**).

This coupling method employs imidazole and 2-chloro-1,3-dimethylimidazolinium chloride to form an imidazole dimer, 2-imidazolyl-1,3-dimethylimidazolinium chloride. The dimer goes on to react with UMP to form UMP-imidazolide, at which point the sugar-1-phosphate is added and UDP-sugar formation begins. After 18 h, all of the UMP-imidazolide had reacted, although starting material was still present in a 1:1 ratio with the product. An alkaline phosphatase is used during the workup to cleave terminal phosphate groups; this

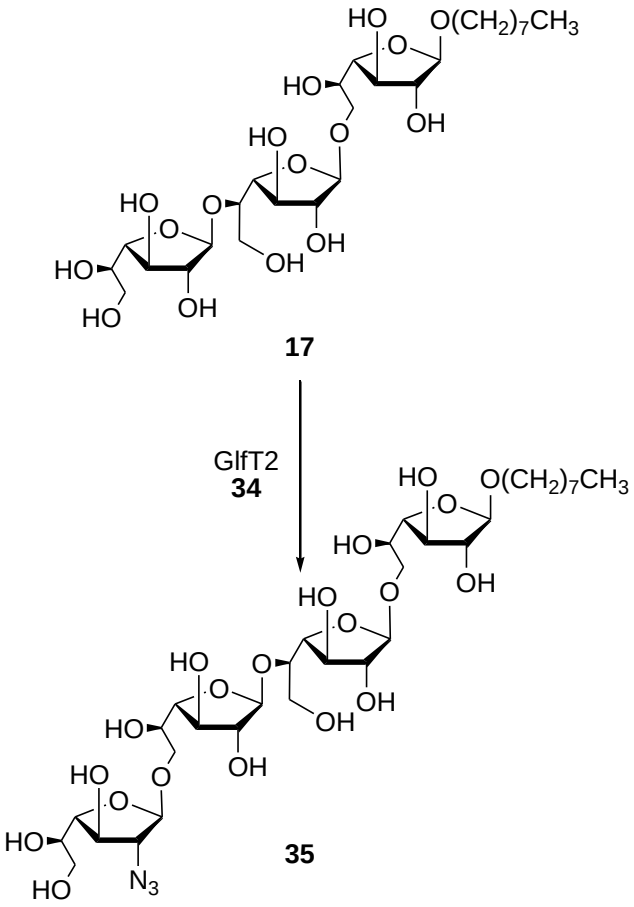
treatment is essential to obtain a clean end product, which is purified by preparatory HPLC. Using this method, **2** was obtained in 16% yield while **34** was obtained in 23% yield.

### 3.5 Enzymatic Reactions of UDP-Gal*f*NAc (2) and UDP-Gal*f*$N_3$ (34)

To demonstrate their biochemical utility, **2** and **34** were further tested as substrates in glycosyltransferase-catalyzed reactions. While UDP-Gal*f*NAc (**2**) is known already to be a substrate of UNGM,[8] a mutase from *C. jejuni*, the glycosyltransferase that incorporates this compound into *C. jejuni* CPS has yet to be isolated. Therefore, **2** and **34** were tested as substrates of two glycosyltransferases that normally use UDP-Gal*f* (**1**) as a substrate: GlfT2, an enzyme contributing to the biosynthesis of cell wall galactan in *M. tuberculosis*,[26] and WbbI, a β-(1→6)-Gal*f* transferase from *Escherichia coli* K-12 involved in O-antigen biosynthesis[27] (see Scheme 2-7).

No product formation was observed in WbbI-catalyzed reactions using either donor **2** or **34** and octyl α-D-glucopyranose as acceptor. The reaction was monitored over a period of three days, and more donor and enzyme were added in an effort to drive the reaction forward. Likewise, no product formation was observed in the GlfT2-mediated reaction using UDP-Gal*f*NAc (**2**) as the donor and a trisaccharide acceptor **17** (see Figure 2-1). However, MS analysis of the reaction of UDP-Gal*f*$N_3$ (**34**) with the trisaccharide **17** detected the presence of a tetrasaccharide, terminated by one unit of Gal*f*$N_3$ **35** (Scheme 3-8). Although GlfT2 is known to produce polymeric products,[28] no further Gal*f*$N_3$ additions

were detected. Subsequent incubation of the Gal*f*$N_3$-terminated tetrasaccharide **35** with GlfT2 and UDP-Gal*f* (**1**) did not result in the addition of any further Gal*f* residues, suggesting that **34** could be used as a chain-terminating substrate for biochemical investigations of *M. tuberculosis* galactan.[23,29]




**Scheme 3-8.** Addition of **34** to trisaccharide **17**, catalyzed by GlfT2.

Three GlfT2 mutants were used to explore possible interactions of the C-2 substituent with other nearby active site residues. However, incubating **34** with H413S, W309S, or K369S did not produce any tetrasaccharide.

### 3.6 Discussion and Conclusions

This straightforward synthesis relies on well-established methodologies to synthesize the sugar-1-phosphates **20** and **34**, and provides a validation of a recently introduced method of coupling sugars to nucleosides through a pyrophosphate bond.[25] Although modest, yields of 16% for **2** and 23% for **34** are comparable to other methods using similar substrates,[9] and certainly represent an increase over the 4% yield achieved using the mutase-mediated system.[8] As Tanaka *et al.* note, the crude reaction mixture may be used directly as a source of sugar nucleotides for the investigation of glycosyltransferase enzymes and their oligosaccharide products.[25] Ready access to these sugar nucleotides will also facilitate studies of *C. jejuni* CPS biosynthesis.[8,30]

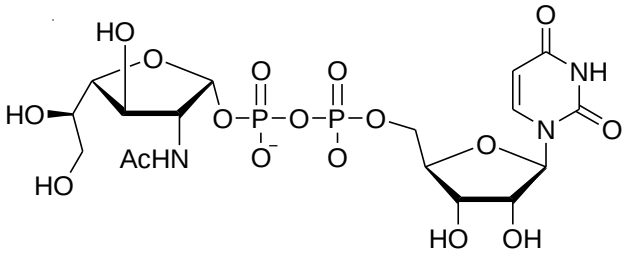
The recognition of UDP-Gal*f*$N_3$ (**34**) by GlfT2 and its subsequent incorporation into a galactan demonstrates that these compounds could prove useful for investigations of oligosaccharide biosynthesis mediated by glycosyltransferases. The incorporation of a non-natural substrate onto the galactan further demonstrates that glycosyltransferases may tolerate functional group substitutions and other modifications in certain circumstances.[31] This statement is far from being general for all glycosyltransferases, however.

In addition, the inclusion of an azide moiety into the galactan will permit further chemical modification at this site. Azides are very useful chemical handles, and may be chemically modified in a number of different ways; for instance, Bertozzi and co-workers have demonstrated that azide-containing sugars

can be incorporated into cell surface glycoconjugates both *in vitro*[32] and *in vivo*.[33] The azide can then be converted to an amide by mild Staudinger ligation, followed by tagging with a fluorescent antibody and subsequent detection in a high-throughput assay. Azides are also useful substrates for "click chemistry" type reactions, where they undergo cycloaddition with alkynes to form triazoles.[34]

## 3.7 Experimental

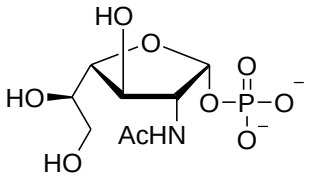
**UDP-2-acetamido-2-deoxy-α-D-galactofuranose triethylammonium salt (2)**




Carried out using a previously reported method.[1] UMP ($Na^+$ form; 29.1 mg, 0.079 mmol), imidazole (21.6 mg, 0.318 mmol), and 2-chloro-1,3-dimethylimidazolinium chloride (DMC, 26.9 mg, 0.159 mmol) were dissolved in 40 μL of $D_2O$ in a small Eppendorf tube. A stir bar was added, and the mixture rotated in a 37 °C incubator for 1 h. After this time, Gal*f*NAc-1-P triethylammonium salt **20** (10 mg, 0.020 mmol) was added, and the mixture was stirred for 18 h. The solution was diluted in 5 mL Tris-HCl buffer, pH = 8.0, 50 U of calf intestinal alkaline phosphatase (CIAP) added, and the solution shaken at 30 °C for 24 h. This solution was centrifuged to remove protein and concentrated under vacuum before purification by HPLC (mobile phase: 50 mM TEAA buffer, pH = 6.8, containing 1.5% acetonitrile; column: Varian Microsorb C18, 21.4 x 250 mm; flow rate: 7.0 mL/min; detection: PDA, λ = 262 nm). Further purification was carried out on a Sephadex G-15 column, eluting with MQ water, followed by freeze-drying to isolate the desired sugar nucleotide **2** as a fluffy white solid (2.6 mg, 3.21 μmol, 16%). $^1$H NMR (500 MHz, $D_2O$) δ 7.98 (d, *J* = 8.1 Hz, 1H, Uridine H-6), 6.05–5.96 (m, 2H, Ribose H-1, Uridine H-5), 5.68 (dd,

*J* = 5.5, 4.5 Hz, 1H, Gal*f*NAc H-1), 4.53 (ddd, *J* = 9.2, 4.4, 2.5 Hz, 1H, Gal*f*NAc H-2), 4.42–4.37 (m, 2H, Ribose H-2, H-3), 4.33 (dd, *J* = 9.2, 7.7 Hz, 1H, Gal*f*NAc H-3), 4.32–4.30 (m, 1H, Ribose H-4), 4.26 (ddd, *J* = 11.7, 4.6, 2.6 Hz, 1H, Ribose H-5), 4.21 (ddd, *J* = 11.8, 5.6, 3.0 Hz, 1H, Ribose H-5'), 3.95 (dd, *J* = 7.6, 5.0 Hz, 1H, Gal*f*NAc H-4), 3.81 (ddd, *J* = 7.2, 5.0, 4.4 Hz, 1H, Gal*f*NAc H-5), 3.75 (dd, *J* = 11.8, 4.4 Hz, 1H, Gal*f*NAc H-6), 3.67 (dd, *J* = 11.8, 7.2 Hz, 1H, Gal*f*NAc H-6′), 3.23 (q, *J* = 7.3 Hz, 12H, N*CH*$_2$CH$_3$), 2.11 (s, 3H, Gal*f*NAc CO*CH*$_3$), 1.31 (t, *J* = 7.3 Hz, 18H, NCH$_2$*CH*$_3$). $^{13}$C NMR (125 MHz, D$_2$O) δ 174.8 (*CO*CH$_3$), 166.4 (Uridine C-4), 151.9 (Uridine C-2), 141.8 (Uridine C-6), 102.7 (Uridine C-5), 97.1 (d, *J* = 5.9 Hz, Gal*f*NAc C-1), 88.4 (Ribose C-1), 83.3 (d, *J* = 9.0 Hz, Ribose C-4), 82.6 (Gal*f*NAc C-4), 73.8 (Ribose C-2), 72.6 (Gal*f*NAc C-3), 72.0 (Gal*f*NAc C-5), 69.7 (Ribose C-3), 65.0 (d, *J* = 5.6 Hz, Ribose C-5), 62.3 (Gal*f*NAc C-6), 58.3 (d, *J* = 7.6 Hz, Gal*f*NAc C-2), 46.7 (N*CH*$_2$CH$_3$), 22.0 (CO*CH*$_3$), 8.3 (NCH$_2$*CH*$_3$); $^{31}$P NMR (162 MHz, D$_2$O) δ –11.51 (d, *J* = 19.9 Hz), –13.30 (d, *J* = 19.9 Hz); HRMS (ESI) *m/z* Calcd for $C_{17}H_{26}N_3O_{17}P_2$ $[M–H]^-$: 606.0743. Found: 606.0737.

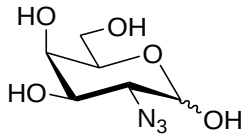
**2-Acetamido-2-deoxy-α-D-galactofuranosyl phosphate (20)**




Dibenzyl 3,5,6-tri-*O*-acetyl-2-azido-2-deoxy-α-D-galactofuranosyl phosphate (**30**) (59 mg, 0.10 mmol), Et$_3$N (85 μL, 0.6 mmol), and 10% Pd/C (12 mg, 20% w/w)

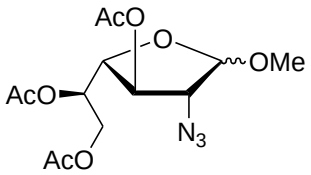
were stirred under an $H_2$ atmosphere in $Ac_2O$ (4 mL) for 3 days at room temperature. The solution was filtered, concentrated under reduced pressure as an azeotrope with toluene, and re-dissolved in 10:2:1 MeOH–$H_2O$–$Et_3N$ (3 mL). This solution stirred 7 days at room temperature, then the solvent was removed under reduced pressure and the residue passed through a C-18 Sep-Pak cartridge, eluting with $H_2O$. Fractions containing the desired compound were pooled, frozen, and freeze-dried. The desired triethylammonium salt **20** (43.4 mg, 0.072 mmol, 72%) was obtained as a dark yellow-brown oil. $[\alpha]_D$ +34 (*c* 0.5, $H_2O$); FTIR (thin film): 1673 $cm^{-1}$; $^1H$ NMR (500 MHz, $D_2O$) δ 5.57 (dd, *J* = 5.0, 4.5 Hz, 1H, H-1), 4.46 (ddd, *J* = 9.0, 4.5, 2.0 Hz, 1H, H-2), 4.34 (dd, *J* = 9.0, 7.5 Hz, 1H, H-3), 3.93 (dd, *J* = 7.5, 4.5 Hz, 1H, H-4), 3.80 (dd, *J* = 7.5, 4.5, 4.5 Hz, 1H, H-5), 3.74 (dd, *J* = 12.0, 4.5 Hz, 1H, H-6), 3.67 (dd, *J* = 12.0, 7.5 Hz, 1H, H-6′), 3.23 (q, *J* = 7.5 Hz, 9H, N*CH*$_2$CH$_3$), 2.09 (s, 3H, CO*CH*$_3$), 1.31 (t, *J* = 7.5, 14H, NCH$_2$*CH*$_3$); $^{13}C$ NMR (125 MHz, $D_2O$) δ 174.8 (s, *CO*CH$_3$), 96.1 (d, *J* = 5.4 Hz, C-1), 82.2 (s, C-4), 72.9 (s, C-3), 71.7 (s, C-5), 62.5 (s, C-6), 58.3 (d, *J* = 7.6 Hz, C-2), 46.7 (s, N*CH*$_2$CH$_3$), 22.0 (s, CO*CH*$_3$), 8.3 (s, NCH$_2$*CH*$_3$); $^{31}P$ NMR (200 MHz, $D_2O$) δ 0.50 (s); HRMS (ESI) *m/z* Calcd for $C_8H_{15}NO_9P$ $[M–H]^-$: 300.0490. Found: 300.0489.

**2-Azido-2-deoxy-D-galactose (21)**




Synthesized using a previously reported method.[2] Trifluoromethanesulfonic anhydride (235 μL, 1.4 mmol) was added dropwise to sodium azide (546 mg, 8.4 mmol) stirring vigorously in 2:1 $H_2O$:$CH_2Cl_2$ (3 mL) at 0 ˚C. After 2 h, the organic layer was separated and the aqueous layer extracted with $CH_2Cl_2$ (1 mL). The organic layers were combined, washed with $NaHCO_3$ saturated solution, and added dropwise to a solution of galactosamine hydrochloride (**23**) (100 mg, 0.464 mmol), $CuSO_4 \cdot 5H_2O$ (6 mg, 0.023 mmol), and triethylamine (200 μL, 1.4 mmol) stirring in MeOH (2 mL). After 2 h, solid $NaHCO_3$ was added and stirred 15 min., and then the solution was filtered through Celite, dried over $Na_2SO_4$, filtered, and concentrated to dryness. Purification (9:1 EtOAc-MeOH) yielded the purified compound as a brown solid mixture of anomers (70.4 mg, 0.343 mmol, 74%, 1:2 α:β). $R_f$ 0.43, 0.35 (9:1 EtOAc-MeOH); FTIR (Nujol mull): 2128 $cm^{-1}$. $^1H$ NMR (500 MHz, $D_2O$) δ 5.41 (d, $J$ = 4.0 Hz, 1H, H-1α), 4.66 (d, $J$ = 8.0 Hz, 1H, H-1β), 4.12 (dd, $J$ = 6.0, 6.0 Hz, 1H, H-5α), 4.04–4.01 (m, 2H, H-3α, H-4α), 3.94 (d, $J$ = 3.5 Hz, 1H, H-4β), 3.82–3.75 (m, 4H, H-6,6′α, H-6,6′β), 3.72–3.67 (m, 3H, H-2α, H-3β, H-5β), 3.49 (dd, $J$ = 10.0, 8.0 Hz, 1H, H-2β). $^{13}C$ NMR (125 MHz, $D_2O$) δ 95.6 (C-1β), 91.4 (C-1α), 75.3 (C-5β), 71.7 (C-3β), 70.7 (C-5α), 68.9 (C-4α), 68.1 (C-3α), 68.0 (C-4β), 64.7 (C-2β), 61.2 (C-6α), 61.0 (C-6β), 60.5 (C-2α). HRMS (ESI) *m/z* Calcd for $C_6H_{11}N_3NaO_5$ $[M+Na]^+$: 228.0591. Found: 228.0586.

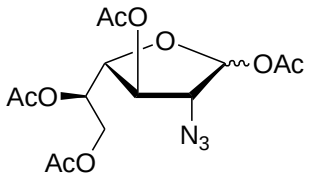
**Methyl 3,5,6-tri-*O*-acetyl-2-azido-2-deoxy-D-galactofuranoside (28)**




2-Azido-2-deoxy-D-galactose (**21**) (4.67 g, 22.8 mmol) was heated to 80 ˚C in 1,4-dioxane (50 mL), then 2,2-dimethoxypropane (15.1 mL, 123 mmol) and *p*-toluenesulfonic acid monohydrate (1.1 g, 8.0 mmol) were sequentially added. After stirring at 80 ˚C for 2 h, the solution was cooled, neutralized with Amberlite IR-410 ($HO^-$) resin, and filtered through Celite, washing with MeOH. After the filtrate was concentrated under reduced pressure, the dark oil was diluted in 20 mL of 80% aqueous AcOH and heated at reflux for 1.5 h to remove the isopropylidene moiety. The solvent was removed as an azeotrope with toluene and the crude material dissolved in pyridine (13 mL). $Ac_2O$ (13 mL) was added, and the solution stirred at room temperature for 16 h. Saturated $NaHCO_3$ solution (50 mL) was added, then the solution extracted with 50 mL $CH_2Cl_2$. After washing the organic phase with 1 M HCl (1 x 100 mL), $NaHCO_3$ saturated solution (1 x 100 mL), and brine (1 x 50 mL), the organic phase was dried over $Na_2SO_4$, filtered, and the filtrate was concentrated under reduced pressure. Column chromatography (3:1 hexane–EtOAc) provided a honey-yellow oil (6.11 g, 17.7 mmol, 78%, 1.4:1 α:β). The β component could not be fully purified, and the mixture was typically carried forward with minor impurities. Data for α anomer: $R_f$ 0.27 (2:1 hexane–EtOAc); $[\alpha]_D$ +107 (*c* 0.89, $CH_2Cl_2$); FTIR (thin film): 2115 $cm^{-1}$; $^1H$ NMR (500 MHz, $CDCl_3$) δ 5.54 (dd, *J* = 8.5, 6.5 Hz, 1H, H-3), 5.22 (ddd, *J* = 6.5, 4.5, 4.5 Hz, 1H, H-5), 4.92 (d, *J* = 4.5 Hz, 1H, H-1), 4.30

(dd, *J* = 12.0, 4.5 Hz, 1H, H-6), 4.16 (dd, *J* = 12.0, 6.5 Hz, 1H, H-6′), 4.13 (d, *J* = 6.5 Hz, 1H, H-4), 3.67 (dd, *J* = 8.5, 4.5 Hz, 1H, H-2), 3.45 (s, 3H, O*CH*$_3$), 2.12 (s, 3H, CO*CH*$_3$), 2.10 (s, 3H, CO*CH*$_3$), 2.05 (s, 3H, CO*CH*$_3$); $^{13}$C NMR (125 MHz, $CDCl_3$) δ 170.5, 170.0, 169.7 (3 x *CO*$CH_3$), 102.7 (C-1), 78.8 (C-4), 73.1 (C-3), 70.4 (C-5), 65.1 (C-2), 62.2 (C-6), 55.3 (O*CH*$_3$), 20.7, 20.7, 20.7 (3 x CO*CH*$_3$); HRMS (ESI) *m/z* Calcd for $C_{13}H_{19}N_3NaO_8$ $[M+Na]^+$: 368.1064. Found: 368.1059. Data for β anomer: $R_f$ 0.27 (2:1 hexane–EtOAc); $^1$H NMR (500 MHz, $CDCl_3$) δ 5.40–5.32 (m, 1H, H-5), 4.90 (m, 2H, H-1, H-3), 4.35 (dd, *J* = 11.8, 4.7 Hz, 1H, H-6), 4.26 (dd, *J* = 5.9, 3.2 Hz, 1H, H-4), 4.22 (dd, *J* = 11.8, 7.1 Hz, 1H, H-6′), 3.90 (d, *J* = 1.2 Hz, 1H, H-2), 3.39 (s, 3H, O*CH*$_3$), 2.15 (s, 3H, CO*CH*$_3$), 2.11 (s, 3H, CO*CH*$_3$), 2.06 (s, 3H, CO*CH*$_3$); $^{13}$C NMR (125 MHz, $CDCl_3$) δ 170.6, 170.3, 170.2 (3 x *CO*$CH_3$), 107.0 (C-1), 79.9 (C-4), 77.20 (C-3), 70.2 (C-2), 69.2 (C-5), 62.6 (C-6), 55.2 (O*CH*$_3$), 20.8, 20.8, 20.7 (3 x CO*CH*$_3$).

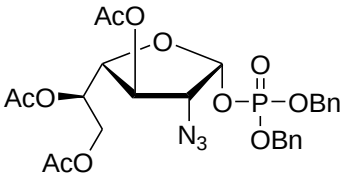
**1,3,5,6-Tetra-*O*-acetyl-2-azido-2-deoxy-D-galactofuranose (29)**




Concentrated $H_2SO_4$ (5 drops) was added to **28** (6.11 g, 17.7 mmol) stirring in $Ac_2O$ (120 mL), and the mixture stirred 3 h. Saturated aqueous $NaHCO_3$ solution (200 mL) was added, and then extracted with $CH_2Cl_2$ (2 x 100 mL) and the combined organic layers washed with $NaHCO_3$ saturated solution (1 x 200 mL) and brine (1 x 100 mL). The organic layer was dried over $Na_2SO_4$, filtered, and concentrated under reduced pressure. Column purification (4:1 hexane–EtOAc)

yielded the product as a slightly yellow oil (4.41 g, 11.8 mmol, 67%, 1.75:1 α:β). $R_f$ 0.42 (1:1 hexane–EtOAc); $[\alpha]_D$ +25 (*c* 0.61, $CH_2Cl_2$); FTIR (thin film): 2116 $cm^{-1}$; $^1H$ NMR (500 MHz, $CDCl_3$) δ 6.27 (d, *J* = 4.5 Hz, 1H, H-1α), 6.13 (s, *J* < 1.0 Hz, 1H, H-1β), 5.54 (dd, *J* = 8.5, 6.5 Hz, 1H, H-3α), 5.35 (ddd, *J* = 6.5, 4.5, 3.5 Hz, 1H, H-5β), 5.25 (ddd, *J* = 6.0, 6.0, 4.5 Hz, 1H, H-5α), 4.96 (dd, *J* = 5.0, 2.5 Hz, 1H, H-3β), 4.39 (dd, *J* = 5.0, 3.5 Hz, 1H, H-4β), 4.32 (dd, *J* = 12.0, 4.5 Hz, 1H, H-6β), 4.28 (dd, *J* = 12.5, 4.5 Hz, 1H, H-6α), 4.20 (dd, *J* = 12.0, 6.5 Hz, 1H, H-6β′), 4.20 (dd, *J* = 6.5, 6.0 Hz, 1H, H-4α), 4.15 (dd, *J* = 12.0, 6.0 Hz, 1H, H-6α′), 4.05 (dd, *J* = 2.5, <1.0 Hz, 1H, H-2β), 3.92 (dd, *J* = 8.5, 4.5 Hz, 1H, H-2α), 2.17 (s, 3H, CO*CH*$_3$ α), 2.14 (s, 3H, CO*CH*$_3$ β), 2.13 (s, 3H, CO*CH*$_3$ α), 2.12 (s, 3H, CO*CH*$_3$ β), 2.11 (s, 3H, CO*CH*$_3$ α), 2.11 (s, 3H, CO*CH*$_3$ β), 2.05 (s, 3H, CO*CH*$_3$ α), 2.05 (s, 3H, CO*CH*$_3$ β); $^{13}C$ NMR (126 MHz, $CDCl_3$) δ 170.6, 170.4, 170.2, 170.1, 169.8, 169.7, 169.3, 169.2 (8 x *CO*$CH_3$), 100.1 (C-1β), 94.7 (C-1α), 82.6 (C-4β), 79.8 (C-4α), 77.1 (C-3β), 72.9 (C-3α), 70.1 (C-5α), 69.9 (C-2β), 69.1 (C-5β), 64.9 (C-2α), 62.6 (C-6β), 62.1 (C-6α), 21.2, 21.1, 20.8, 20.8, 20.8, 20.7, 20.7, 20.7 (8 x CO*CH*$_3$); HRMS (ESI) *m/z* Calcd for $C_{14}H_{19}N_3NaO_9$ $[M+Na]^+$: 396.1014. Found: 396.1009.

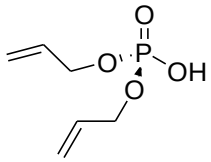
**Dibenzyl 3,5,6-tri-*O*-acetyl-2-azido-2-deoxy-α-D-galactofuranosyl phosphate (30)**




$TiBr_4$ (790 mg, 2.14 mmol) was added to 1,3,5,6-tetra-*O*-acetyl-2-azido-2-deoxy-D-galactofuranose (**29**) (500 mg, 1.34 mmol) stirring in 25 mL of 9:1 $CH_2Cl_2$–EtOAc, and the solution stirred at room temperature for 4 days. After this time, the solvent was removed under reduced pressure, and 5 mL of toluene was added. A solution of dibenzyl phosphate (1.9 g, 6.8 mmol) dissolved in 5 mL toluene and 5 mL $Et_3N$ was then added to the reaction mixture, and the solution was stirred for 3 h. The dark solution was filtered and concentrated under reduced pressure, followed by column purification of the residue (3:2 hexane–EtOAc) to yield **30** (357 mg, 0.604 mmol, 45%) as a yellow oil. A small amount of the β anomer (153 mg, 0.259 mmol, 19%) was also obtained. Data for α anomer: $R_f$ 0.17 (1:1 hexane–EtOAc); $[\alpha]_D$ +42 (*c* 0.55, $CH_2Cl_2$); FTIR (thin film): 2116 $cm^{-1}$; $^1H$ NMR (600 MHz, $CDCl_3$) δ 7.37–7.32 (m, 10H, Ar), 5.92 (dd, *J* = 6.0, 4.8 Hz, 1H, H-1), 5.51 (dd, *J* = 8.4, 7.2 Hz, 1H, H-3), 5.21 (ddd, *J* = 6.0, 4.8, 4.2 Hz, 1H, H-5), 5.15–5.08 (m, 4H, -C*H*$_2$Ph), 4.30 (dd, *J* = 11.4, 4.8 Hz, 1H, H-6), 4.23 (dd, *J* = 7.2, 4.2 Hz, 1H, H-4), 4.12 (dd, *J* = 11.4, 6.0 Hz, 1H, H-6′), 3.91 (ddd, *J* = 8.4, 4.8, 2.4 Hz, 1H, H-2), 2.13 (s, 3H, CO*CH*$_3$), 2.04 (s, 3H, CO*CH*$_3$), 1.97 (s, 3H, CO*CH*$_3$). $^{13}C$ NMR (125 MHz, $CDCl_3$) δ 170.4, 170.0, 169.6 (3 x *CO*$CH_3$), 128.69, 128.66, 128.64, 128.0, 127.9 (10 x aryl C), 98.9 (d, *J* = 5.1 Hz, C-1), 80.0

(C-4), 72.1 (C-3), 69.7 (d, $J$ = 5.4 Hz, -$CH_2$Ph), 69.4 (d, $J$ = 5.4 Hz, -$CH_2$Ph), 69.3 (C-5), 65.4 (d, $J$ = 7.8 Hz, C-2), 62.0 (C-6), 20.7, 20.6, 20.5 (3 x CO$CH_3$); $^{31}$P NMR (162 MHz, $CDCl_3$) δ –2.6 (s); HRMS (ESI) *m/z* Calcd for $C_{26}H_{30}N_3O_{11}PNa$ $[M+Na]^+$: 614.1510. Found: 614.1503. Data for β anomer: $R_f$ 0.30 (1:1 hexane–EtOAc); $[\alpha]_D$ –33 (*c* 0.75, $CH_2Cl_2$); FTIR (thin film): 2113 $cm^{-1}$; $^1$H NMR (400 MHz, $CDCl_3$) δ 7.41–7.30 (m, 10H, Ar), 5.68 (d, $J$ = 4.7 Hz, 1H, H-1), 5.33 (ddd, $J$ = 7.0, 4.6, 3.7 Hz, 1H, H-5), 5.13–5.01 (m, 4H, -$CH_2$Ph), 4.86 (d, $J$ = 4.7 Hz, 1H, H-3), 4.40 (dd, $J$ = 4.7, 3.7 Hz, 1H, H-4), 4.22 (dd, $J$ = 11.8, 4.6 Hz, 1H, H-6), 4.12 (dd, $J$ = 11.8, 7.0 Hz, 1H, H-6′), 3.89 (m, 1H, H-2), 2.10 (s, 3H, CO$CH_3$), 2.07 (s, 3H, CO$CH_3$), 2.01 (s, 3H, CO$CH_3$); $^{13}$C NMR (100 MHz, $CDCl_3$) δ 170.4, 170.2, 170.1 (3 x *CO*$CH_3$), 135.4, 135.4, 128.8, 128.7, 128.7, 128.1 (10 x aryl C), 103.1 (d, $J$ = 5.4 Hz, C-1), 82.6 (C-4), 76.9 (C-3), 70.5 (d, $J$ = 9.6 Hz, C-2), 69.7 (d, $J$ = 6.0 Hz, -$CH_2$Ph), 69.7 (d, $J$ = 6.0 Hz, -$CH_2$Ph), 69.0 (C-5), 62.3 (C-6), 20.7, 20.7, 20.7 (3 x CO$CH_3$); $^{31}$P NMR (162 MHz, $CDCl_3$) δ –3.0 (s); HRMS (ESI) *m/z* Calcd for $C_{26}H_{30}N_3O_{11}PNa$ $[M+Na]^+$: 614.1510. Found: 614.1508.

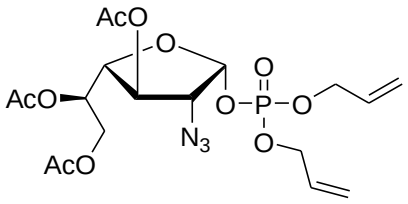
**Diallyl phosphate**




Synthesized according to a known procedure.[3] Methyl dichlorophosphate (85%, 1.2 mL, 10 mmol) was added dropwise to pyridine (10 mL) stirring at 0 ˚C. The solution stirred a further 15 min at this temperature, then allyl alcohol (3.4 mL, 50 mmol) was added and the solution was warmed to room temperature and stirred

overnight. $NaHCO_3$ saturated solution (50 mL) was added, then excess alcohol removed by extraction with ether (2 x 50 mL). The aqueous fraction was acidified to pH 1 with conc. HCl, and then extracted with 7:3 $CH_2Cl_2$–*n*-BuOH (2 x 50 mL). The organic phase was washed with 1 M HCl (1 x 50 mL) and brine (1 x 50 mL), then dried over $Na_2SO_4$, filtered, and concentrated under reduced pressure. The desired phosphate (1.24 g, 7.0 mmol, 70%) was isolated as a clear oil and used without further purification. $^1$H NMR (500 MHz, $CDCl_3$) δ 5.95 (ddd, *J* = 17.0, 10.5, 5.5 Hz, 2H, =C*H*), 5.36 (ddd, *J* = 17.0, 2.5, 1.0 Hz, 2H, =C$H_2$), 5.24 (dd, *J* = 10.5, 1.5 Hz, 2H, =C$H_2$), 4.53 (ddt, *J* = 8.5, 5.5, 1.5 Hz, 4H, -C$H_2$); $^{13}$C NMR (125 MHz, $CDCl_3$) δ 132.4 (d, *J* = 7.3 Hz, =*CH*), 118.2 (=$CH_2$), 68.1 (d, *J* = 5.4 Hz, -$CH_2$); $^{31}$P NMR (162 MHz, $CDCl_3$) δ 0.63 (s); HRMS (ESI) *m/z* Calcd for $C_6H_{10}O_4P$ $[M–H]^-$: 177.0322. Found: 177.0320.

**Diallyl 3,5,6-tri-*O*-acetyl-2-azido-2-deoxy-α-D-galactofuranosyl phosphate (31)**

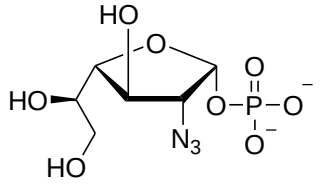



$TiBr_4$ (790 mg, 2.14 mmol) was added to 1,3,5,6-tetra-*O*-acetyl-2-azido-2-deoxy-D-galactofuranose (**29**) (500 mg, 1.34 mmol) stirring in 25 mL of 9–1 $CH_2Cl_2$–EtOAc, and the solution stirred at room temperature for 4 days. After this time the solvent was removed under reduced pressure, and 5 mL of toluene was added. A solution of diallyl phosphate (1.0 g, 5.61 mmol) dissolved in 5 mL toluene and 5

mL $Et_3N$ was then added to the reaction mixture, and the solution was stirred 3 h. The dark mixture was filtered and concentrated under reduced pressure, followed by column purification (3:2 hexane–EtOAc) to yield **31** (228 mg, 0.460 mmol, 35%) as a yellow oil. A small amount of the β anomer (59 mg, 0.12 mmol, 9%) was also obtained. Data for α anomer: $R_f$ 0.20 (1:1 hexane–EtOAc); $[\alpha]_D$ +41 (*c* 0.9, $CH_2Cl_2$); FTIR (thin film): 2116 $cm^{-1}$; $^1H$ NMR (500 MHz, $CDCl_3$) δ 6.00–5.92 (m, 2H, 2 x =C*H*–), 5.90 (dd, *J* = 5.5, 4.5 Hz, 1H, H-1), 5.52 (dd, *J* = 8.5, 7.5 Hz, 1H, H-3), 5.39 (d, *J* = 17.5 Hz, 2H, =C$H_2$), 5.30–5.22 (m, 3H, =C$H_2$, H-5), 4.64–4.59 (m, 4H, 2 x -C$H_2$), 4.32 (dd, *J* = 12.0, 5.0 Hz, 1H, H-6), 4.24 (dd, *J* = 7.5, 4.0 Hz, 1H, H-4), 4.18 (dd, *J* = 12.0, 6.0 Hz, 1H, H-6′), 3.91 (ddd, *J* = 8.5, 4.5, 2.0 Hz, 1H, H-2), 2.13 (s, 3H, CO$CH_3$), 2.12 (s, 3H, CO$CH_3$), 2.05 (s, 3H, CO$CH_3$); $^{13}C$ NMR (125 MHz, $CDCl_3$) δ 170.4, 169.9, 169.7 (3 x *CO*$CH_3$), 132.3 (d, *J* = 7.8 Hz, =*CH*–), 132.2 (d, *J* = 7.3 Hz, =*CH*–), 118.6 (=$CH_2$), 118.5 (=$CH_2$), 98.9 (d, *J* = 5.1 Hz, C-1), 80.0 (C-4), 72.0 (C-3), 69.4 (C-5), 68.6 (d, *J* = 5.4 Hz, -$CH_2$), 68.4 (d, *J* = 5.1 Hz, -$CH_2$), 65.4 (d, *J* = 7.6, C-2), 62.0 (C-6), 20.7, 20.7, 20.6 (3 x CO$CH_3$); $^{31}P$ NMR (200 MHz, $CDCl_3$) δ –2.5 (s); HRMS (ESI) *m/z* Calcd for $C_{18}H_{26}N_3O_{11}PNa$ $[M+Na]^+$: 514.1197. Found: 514.1190. Data for β anomer: $R_f$ 0.26 (1:1 hexane–EtOAc); $[\alpha]_D$ –39 (*c* 0.5, $CH_2Cl_2$); FTIR (thin film): 2114 $cm^{-1}$; $^1H$ NMR (500 MHz, $CDCl_3$) δ 5.99–5.91 (m, 2H, 2 x =C*H*–), 5.75 (d, *J* = 4.5 Hz, 1H, H-1), 5.41–5.37 (m, 3H, 2 x =$CH_2$, H-5), 5.29 (dd, *J* = 10.5, 1.0 Hz, 2H, 2 x =$CH_2$), 4.94 (dd, *J* = 5.0, 1.5 Hz, 1H, H-3), 4.60–4.57 (m, 4H, 2 x –$CH_2$-), 4.50 (dd, *J* = 5.0, 3.5 Hz, 1H, H-4), 4.32 (dd, *J* = 11.5, 4.5 Hz, 1H, H-6), 4.19 (dd, *J* = 11.5, 7.0 Hz, 1H, H-6′), 4.16 (d, *J* = 1.5 Hz, 1H, H-2), 2.13 (s, 3H,

CO*CH*$_3$), 2.12 (s, 3H, CO*CH*$_3$), 2.06 (s, 3H, CO*CH*$_3$); $^{13}$C NMR (125 MHz, $CDCl_3$) δ 170.5, 170.3, 170.2 (3 x *CO*$CH_3$), 132.1 (d, *J* = 6.4 Hz, 2 x =*CH*–), 118.9 (=*CH*$_2$), 103.2 (d, *J* = 5.4 Hz, C-1), 82.6 (C-4), 76.9 (C-3), 70.7 (d, *J* = 9.5 Hz, C-2), 69.0 (C-5), 68.6 (d, *J* = 6.6, -*CH*$_2$-), 68.6 (d, *J* = 5.4, -*CH*$_2$-), 62.4 (C-6), 20.8, 20.7, 20.7 (3 x CO*CH*$_3$); $^{31}$P NMR (162 MHz, $CDCl_3$) δ –3.6 (s); HRMS (ESI) *m/z* Calcd for $C_{18}H_{26}N_3O_{11}PNa$ $[M+Na]^+$: 514.1197. Found: 514.1186.

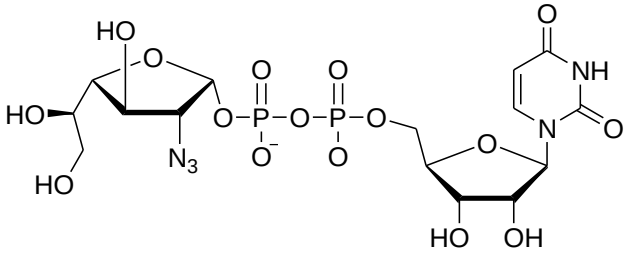
**2-Azido-2-deoxy-α-D-galactofuranosyl phosphate (32)**




$Pd(OAc)_2$ (170 mg, 0.76 mmol) and NaOAc (135 mg, 1.65 mmol) were added to the azide **31** (250 mg, 0.51 mmol) stirring in AcOH (2.5 mL) with $H_2O$ (5 drops), and the solution stirred for 16 h at room temperature. The solution was filtered and concentrated under reduced pressure as an azeotrope with toluene. The residue was dissolved in 10 mL of 10:2:1 MeOH–$H_2O$–$Et_3N$ and stirred at room temperature for 6 days, then the solvent was removed under reduced pressure and the residue passed through a C-18 Sep-Pak cartridge, eluting with $H_2O$. Fractions containing the desired compound were pooled, frozen, and freeze-dried. 172.1 mg of the desired triethylammonium salt form of **32** (172.1 mg, 0.353 mmol, 69%) was obtained as an oil. A small amount was converted to the $Na^+$ salt form for characterization using Amberlite IR-140 resin ($Na^+$ form), and appeared as a light brown solid. $[\alpha]_D$ +43 (*c* 0.2, $H_2O$); FTIR (thin film): 2118 $cm^{-1}$; $^1$H NMR

(500 MHz, $D_2O$) δ 5.62 (dd, *J* = 5.0, 5.0 Hz, 1H, H-1), 4.49 (dd, *J* = 9.0, 7.0 Hz, 1H, H-3), 3.88 (dd, *J* = 7.0, 4.5 Hz, 1H, H-4), 3.84–3.78 (m, 2H, H-2, H-5), 3.73 (dd, *J* = 11.5, 4.5 Hz, 1H, H-6), 3.68 (dd, *J* = 11.5, 7.0 Hz, 1H, H-6′); $^{13}$C NMR (125 MHz, $D_2O$) δ 96.5 (d, *J* = 5.2 Hz, C-1), 81.8 (C-4), 72.3 (C-3), 71.7 (C-5), 66.8 (d, *J* = 7.6 Hz, C-2), 62.6 (C-6); $^{31}$P NMR (200 MHz, $D_2O$) δ 2.2 (s); HRMS (ESI) *m/z* Calcd for $C_6H_{11}N_3O_8P$ $[M–H]^-$: 284.0289. Found: 284.0285.

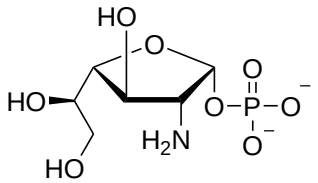
**UDP-2-azido-2-deoxy- α-D-galactofuranose (34)**




Carried out similarly to the synthesis of **1**, above. UMP ($Na^+$ form; 30.0 mg, 0.082 mmol), imidazole (22 mg, 0.328 mmol), and DMC (28 mg, 0.164 mmol) were dissolved in 40 μL of $D_2O$ in a small Eppendorf tube. A stir bar was added, and the mixture rotated in a 37 °C incubator for 1 h. After this time, Gal*f*$N_3$-1-P triethylammonium salt **9** (10 mg, 0.021 mmol) was added, and the mixture was stirred for 18 h. The solution was diluted in 5 mL Tris-HCl buffer, pH = 8.0, 50 U of CIAP added, and the solution shaken at 30 °C for 24 h. This solution was centrifuged to remove protein and concentrated under vacuum before purification by HPLC using the system described for **1** above. Further purification was carried out on a Sephadex G-15 column, eluting with MQ water, followed by freeze-drying to isolate the desired sugar nucleotide. Compound **2**, (3.7 mg, 4.66 μmol,

23%), was obtained as a fluffy white solid. $^{1}$H NMR (500 MHz, $D_2O$) δ 8.00 (d, *J* = 8.1 Hz, 1H, Uridine H-6), 6.04–5.95 (m, 2H, Uridine H-5, Ribose H-1), 5.79 (dd, *J* = 5.7, 4.6 Hz, 1H, Gal*f*$N_3$ H-1), 4.47 (dd, *J* = 9.0, 7.6 Hz, 1H, Gal*f*$N_3$ H-3), 4.43–4.37 (m, 2H, Ribose H-2, H-3), 4.34–4.31 (m, 1H, Ribose H-4), 4.28 (ddd, *J* = 12.0, 4.5, 2.5 Hz, 1H, Ribose H-5), 4.24 (ddd, *J* = 12.0, 5.5, 3.0 Hz, 1H, Ribose H-5'), 3.98–3.92 (m, 2H, Gal*f*$N_3$ H-2, H-4), 3.82 (ddd, *J* = 7.2, 4.6, 4.6 Hz, 1H, Gal*f*$N_3$ H-5), 3.75 (dd, *J* = 11.8, 4.5 Hz, 1H, Gal*f*$N_3$ H-6), 3.68 (dd, *J* = 11.8, 7.2 Hz, 1H, Gal*f*$N_3$ H-6′), 3.23 (q, *J* = 7.3 Hz, 13H, N*C*$H_2$$CH_3$ x 3), 1.31 (t, *J* = 7.3 Hz, 19.5H, $NCH_2$*C*$H_3$ x 3). $^{13}$C NMR (125 MHz, $D_2O$) δ 166.3 (Uridine C-4), 151.9 (Uridine C-2), 141.7 (Uridine C-6), 102.7 (Uridine C-5), 97.7 (d, *J* = 5.6 Hz, Gal*f*$N_3$ C-1), 88.3 (Ribose C-1), 83.4 (d, *J* = 9.1 Hz, Ribose C-4), 82.4 (Gal*f*$N_3$ C-4), 73.8 (Ribose C-2), 71.9 (Gal*f*$N_3$ C-3), 71.6 (Gal*f*$N_3$ C-5), 69.8 (Ribose C-3), 66.6 (d, *J* = 7.9 Hz, Gal*f*$N_3$ C-2), 65.0 (d, *J* = 5.3 Hz, Ribose C-5), 62.2 (Gal*f*$N_3$ C-6), 46.72 (N*C*$H_2$$CH_3$), 8.26 ($NCH_2$*C*$H_3$); $^{31}$P NMR (162 MHz, $D_2O$) δ –11.4 (d, *J* = 16.4 Hz), –12.98 (d, *J* = 16.4 Hz); HRMS (ESI) *m/z* Calcd for $C_{15}H_{22}N_5O_{16}P_2$ $[M–H]^-$: 590.0542. Found: 590.0538.

**2-Amino-2-deoxy-α-D-galactofuranosyl phosphate**




Gal*f*$N_3$-1-P (**32**, 7 mg, 0.025 mmol) was dissolved in 3:1 MeOH:$H_2O$ (2 mL), then 10% Pd/C (2 mg) added and the solution stirred under an $H_2$ atmosphere for 3

hours. The solution was filtered and the solvent removed under reduced pressure. The residue was diluted in a minimal amount of $H_2O$ and passed through a C-18 Sep-Pak cartridge, eluting with $H_2O$. Fractions containing the desired compound were pooled, frozen, and freeze-dried to give the desired amine (4.5 mg, 0.017 mmol, 68%). FTIR (solid): 3237, 2945, 1101 $cm^{-1}$; $^1H$ NMR (500 MHz, $D_2O$) δ 5.73 (dd, *J* = 5.0, 5.0 Hz, 1H, H-1), 4.54 (dd, *J* = 7.5, 7.0 Hz, 1H, H-3) 3.95 (dd, *J* = 7.0, 4.0 Hz, 1H, H-4), 3.83 – 3.79 (m, 2H, H-2, H-5), 3.74 (dd, *J* = 11.5, 4.5 Hz, 1H, H-6), 3.70 (dd, *J* = 11.5, 7.0 Hz, 1H, H-6'); $^{13}C$ NMR (125 MHz, $D_2O$) δ 95.1 (d, *J* = 4.6 Hz, C-1), 82.5 (C-3), 72.0 (C-4), 71.1 (C-5), 62.5 (C-6), 58.3 (d, *J* = 7.1 Hz, C-2); $^{31}P$ NMR (200 MHz, $D_2O$) δ 1.9 (s);. HRMS (ESI) *m/z* Calcd for $C_6H_{14}NO_8P$ $[M-H]^-$: 258.0384. Found: 258.0381.

**Attempted Enzymatic Synthesis of UDP-Gal*f*NAc (2) or UDP-Gal*f*$N_3$ (34)**

Carried out as previously described.[8] UTP (14.8 μmol), Gal*f*NAc-1-phosphate (**3**) or Gal*f*$N_3$-1-phosphate (**32**) (15.2 μmol), GalU (5 U), IPP (1.25 U), and immobilized GalPUT (300 μL of $Ni^{2+}$-NTA-agarose resin containing 5 U enzyme) were added to HEPES buffer (50 mM, pH = 8.0) containing $MgCl_2$ (10 mM) and KCl (5 mM), with a total reaction volume of 500 μL. UDP-Glc (0.05 μmol) was added to initiate the reaction, and then the mixture was flushed with $N_2$ and incubated overnight at room temperature with agitation. Reaction progress was monitored by reverse phase HPLC (see HPLC conditions, below). Control reactions containing Gal*f*-1-phosphate were incubated alongside the test reactions, and samples were spiked with authentic standards of UTP and UDP-Gal*f*NAc (**2**)

or UDP-Gal*f*$N_3$ (**34**) to aid in peak identification. The control reactions showed that Gal*f*-1-phosphate was converted into UDP-Gal*f*, but that neither Gal*f*NAc-1-phosphate nor Gal*f*$N_3$-1-phosphate was converted into the corresponding UDP sugar.

**Analytical HPLC Conditions**

Buffer A: 200 mM triethylammonium acetate (TEAA), pH = 6.6

Buffer B: 200 mM TEAA, pH = 6.6, containing 5% acetonitrile

Column: Analytical reverse phase C18 (4.6 x 250 mm), Phenomenex, P/N: 00G-4252-E0

Flow Rate: 0.8 mL/min

Detection: Photodiode array, $\lambda$ = 262 nm

| **Time (min)** | **Flow Rate (mL/min)** | **%A** | **%B** |
|---|---|---|---|
| 0 | 0.8 | 96 | 4 |
| 10 | 0.8 | 96 | 4 |
| 25 | 0.8 | 0 | 100 |
| 35 | 0.8 | 0 | 100 |
| 36 | 0.8 | 96 | 4 |
| 5 | 0.8 | 96 | 4 |

**Site-Directed Mutagenesis and Expression of GalPUT Mutants**

Origami BL-21 cells containing the GalPUT gene in a pET-15b plasmid vector were grown on LB agar plates containing 100 μg/mL ampicillin (LB Amp-100) at 37 °C overnight. One colony was grown overnight in 3.5 mL LB media containing 100 μg/mL ampicillin and its plasmid isolated using a GeneJET PCR

Purification Kit (Thermo Scientific, USA). This plasmid was subsequently transformed into DH5α competent cells, plated on LB Amp-100 plates at 37 °C overnight, and one colony grown overnight in 3.5 mL LB media containing 100 μg/mL ampicillin. The plasmid was again isolated using a GeneJET kit, and sequenced to ensure the plasmid was free of mutations. The F312A mutant was made using the Quikchange II XL (Stratagene, USA) site-directed mutagenesis kit, while the K311A and E317A mutants used the Q5 (New England Biolabs, USA) site-directed mutagenesis kit. Mutations were confirmed by sequencing. The following primers were used:

K311A: 5′-CACCGTACGTGCATTTATGGTTGGTTATGAAATGCTGG-3′

5′- GCGGAGCGCAGCAGAGGC-3′

F312A: 5′-CCACCGTACGTAAAGCTATGGTTGGTTATGAAATGCTGG-3′

5′-CCAGCATTTCATAACCAACCATAGCTTTACGTACGGTGG-3′

E317A: 5′-GGTTGGTTATGCAATGCTGGCAG-3′

5′- ATAAATTTACGTACGGTGGC -3′

Mutant plasmids were transformed into *E. coli* Origami BL21-DE3 competent cells and stored at –80 °C. Proteins were expressed and purified as previously described,[21] and reactions set up as described above.

**GlfT2-Mediated Coupling of Trisaccharide 17 and UDP-Gal*f*$N_3$ (34)**

Carried out as described previously;[21,22] GlfT2 was expressed and purified as described elsewhere.[23] Reactions containing UDP-Gal*f*$N_3$ (**34**, 2 mM), trisaccharide acceptor **17** (0.5 mM), GlfT2 (50 μg), CIAP (2 U), $MgCl_2$ (20 mM),

and β-mercaptoethanol (5 mM) in 100 μL MOPS (pH = 7.6) were flushed with nitrogen and incubated at room temperature with rotation. CIAP was added to degrade the UDP by-product, which is known to inhibit GlfT2. Reaction progress was monitored by TLC, eluting with $CHCl_3$-$CH_3OH$-$NH_4OH$-$H_2O$ (65:35:0.5:3.6). To drive the reaction to completion, additional UDP-Gal*f*$N_3$ (**34**) was added on day 2 to bring the final concentration to 4 mM, and additional GlfT2 (2 x 50 μg) was added on days 1 and 2 to bring the final mass to 150 μg. After 3 days, the solution was passed through a Sep-Pak $C_{18}$ cartridge, eluting with 10 mL of $H_2O$. Products were eluted with 4 mL of $CH_3OH$, and then concentrated by rotary evaporator, suspended in 1 mL MQ $H_2O$, and filtered through a 0.22 μm Millex-GV filter. From this solution, 100 μL was lyophilized and analyzed by ESI-MS. HRMS (ESI) *m/z* Calcd for $C_{32}H_{57}N_3NaO_{20}$ $[M+Na]^+$: 826.3428. Found: 826.3417.

**Site-Directed Mutagenesis and Expression of GlfT2 K369S**

The mutagenesis was carried out as described in Chapter 2. The following primers were used:

5'-GCTGTTTATTAGCTGGGACGATGCCGATTATGGTC-3'

5'-GGCAGCGGTTGACCCAGT-3'

The enzyme was expressed and purified as described in Chapter 2.

**Attempted WbbI-Mediated Coupling of Octyl α-D-Glucopyranoside and UDP-Gal*f*$N_3$ (34)**

Expression and purification of WbbI-maltose binding fusion protein, as well as the enzyme-mediated reaction, were carried out as described previously.[27] Reactions containing UDP-Gal*f*$N_3$ (**34**, 2 mM), octyl α-D-glucopyranoside (**37**) (0.5 mM), WbbI-MBP (50 μg), $MgCl_2$ (2 mM), and $Na_2S_2O_4$ (5 mM) in 100 μL MOPS (pH = 7.6) were flushed with nitrogen and incubated at 37 ˚C with rotation. Reaction progress was monitored by TLC, eluting with $CHCl_3$-$CH_3OH$-$H_2O$ (65:35:4.1). After 3 hours, no products were observed. The reaction was allowed to continue for an additional 3 days, with addition of UDP-Gal*f*$N_3$ (**34**) added on day 2 to bring the final concentration to 4 mM and WbbI-MBP (2 x 50 μg) on days 1 and 2 to bring the final mass to 150 μg. After this time, the solution was passed through a Sep-Pak $C_{18}$ cartridge, eluting with 10 mL of $H_2O$. Products were eluted with 4 mL of $CH_3OH$, and then concentrated by rotary evaporator, suspended in 1 mL MQ $H_2O$, and filtered through a 0.22 μm Millex-GV filter. From this solution, 100 μL was lyophilized and analyzed by ESI-MS.

### 3.8 Bibliography


(1) Brennan, P. J. *Tuberculosis* **2003**, *83*, 91−97.

(2) Richards, M. R.; Lowary, T. L. *ChemBioChem* **2009**, *10*, 1920−1938.

(3) Chlubnova, I.; Legentil, L.; Dureau, R.; Pennec, A.; Almendros, M.; Daniellou, R.; Nugier-Chauvin, C.; Ferrieres, V. *Carbohydr. Res.* **2012**, *356*, 44−61.

(4) Karlyshev, A. V.; Champion, O. L.; Churcher, C.; Brisson, J.-R.; Jarrell, H. C.; Gilbert, M.; Brochu, D.; St. Michael, F.; Li, J.; Wakarchuk, W. W.; Goodhead, I.; Sanders, M.; Stevens, K.; White, B.; Parkhill, J.; Wren, B. W.; Szymanski, C. M. *Mol. Microbiol.* **2005**, *55*, 90-103.

(5) Guerry, P.; Szymanski, C. M. *Trends Microbiol.* **2008**, *16*, 428-435.

(6) St. Michael, F.; Szymanski, C. M.; Li, J.; Chan, K. H.; Khieu, N. H.; Larocque, S.; Wakarchuk, W. W.; Brisson, J. R.; Monteiro, M. A. *Eur. J. Biochem*. **2002**, *269*, 5119−5136.

(7) McNally, D. J.; Lamoureux, M. P.; Karlyshev, A. V.; Fiori, L. M.; Li, J.; Thacker, G.; Coleman, R. A.; Khieu, N. H.; Wren, B. W.; Brisson, J.-R.; Jarrell, H. C.; Szymanski, C. M. *J. Biol. Chem*. **2007**, *282*, 28566−28576.

(8) Poulin, M. B.; Nothaft, H.; Hug, I.; Feldman, M. F.; Szymanski, C. M.; Lowary, T. L. *J. Biol. Chem*. **2010**, *285*, 493−501.

(9) Wagner, G. K.; Pesnot, T.; Field, R. A. *Nat. Prod. Rep*. **2009**, *26*, 1172−1194.

(10) Recent reviews: (a) Cai, L. *J. Carbohydr. Chem.* **2012**, *31*, 535−552. (b) Engels, L.; Elling, E. In *Carbohydrate-Modifying Biocatalysts*; Grunwald, P., Ed.; Pan Stanford Publishing: Boca Raton, FL, 2012; pp 237−254.

(11) de Lederkremer, R. M.; Nahmad, V. B.; Varela, O. *J. Org. Chem.* **1994**, *59*, 690−692.

(12) Liu, Q.; Tor, Y. *Org. Lett.* **2003**, *5*, 2571−2572.

(13) Lemieux, R. U.; Ratcliffe, R. M. *Can. J. Chem.* **1979**, *57*, 1244−1251.

(14) Imamura, A.; Lowary, T. L. *Trends Glycosci. Glyc.* **2011**, *23*, 134-152.

(15) Tanahashi, E.; Kiso, M.; Hasegawa, A. *J. Carbohydr. Chem.* **1983**, *2*, 129−137.

(16) Paulsen, H.; Stenzel, W. *Chem. Ber.* **1978**, *111*, 2348−2357.

(17) Rubinstein, M.; Patchornik, A. *Tetrahedron* **1975**, *31*, 2107−2110.

(18) (a) Timmons, S. C.; Hui, J. P. M.; Pearson, J. L.; Peltier, P.; Daniellou, R.; Nugier-Chauvin, C.; Soo, E. C.; Syvitski, R. T.; Ferrières, V.; Jakeman, D. L. *Org. Lett.* **2008**, *10*, 161−163. (b) Timmons, S. C.; Mosher, R. H.; Knowles, S. A.; Jakeman, D. L. *Org. Lett.* **2007**, *9*, 857−860.

(19) Marlow, A. L.; Kiessling, L. L. *Org. Lett.* **2001**, *3*, 2517−2519.

(20) Errey, J. C.; Mukhopadhyay, B.; Kartha, K. P. R.; Field, R. A. *Chem. Commun.* **2004**, *23*, 2706−2707.

(21) Rose, N. L.; Zheng, R. B.; Pearcey, J.; Zhou, R.; Completo, G. C.; Lowary, T. L. *Carbohydr. Res.* **2008**, *343*, 2130−2139.

(22) Poulin, M. B.; Lowary, T. L. *Methods Enzymol.* **2010**, *478*, 389−411.

(23) Poulin, M. B.; Zhou, R.; Lowary, T. L. *Org. Biomol. Chem.* **2012**, 10, 4074−4087.

(24) Thoden, J. B.; Ruzicka, F. J.; Frey, P. A.; Rayment, I.; Holden, H. M. *Biochemistry*. **1997**, *36*, 1212−1222.

(25) Tanaka, H.; Yoshimura, Y.; Jørgensen, M. R.; Cuesta-Seijo, J. A.; Hindsgaul, O. *Angew. Chem., Int. Ed.* **2012**, *51*, 11531−11534.

(26) (a) Wheatley, R. W.; Zheng, R. B.; Richards, M. R.; Lowary, T. L.; Ng, K. K. S. *J. Biol. Chem.* **2012**, *287*, 28132−28143. (b) May, J. F.; Levengood, M. R.; Splain, R. A.; Brown, C. D.; Kiessling, L. L. *Biochemistry* **2012**, *51*, 1148−1159. (c) Rose, N. L.; Completo, G. C.; Lin, S.-J.; McNeil, M.; Palcic, M. M.; Lowary, T. L. *J. Am. Chem. Soc.* **2006**, *128*, 6721−6729.

(27) Wing, C.; Errey, J. C.; Mukhopadhyay, B.; Blanchard, J. S.; Field, R. A. *Org. Biomol. Chem.* **2006**, *4*, 3945−3950.

(28) Levengood, M. R.; Splain, R. A.; Kiessling, L. L. *J. Am. Chem. Soc.* **2011**, *133*, 12758−12766.

(29) Brown, C. D.; Rusek, M. S.; Kiessling, L. L. *J. Am. Chem. Soc.* **2012**, *134*, 6552−6555.

(30) Sternberg, M. J. E.; Tamaddoni-Nezhad, A.; Lesk, V. I.; Kay, E.; Hitchen, P. G.; Cootes, A.; van Alphen, L. B.; Lamoureux, M. P.; Jarrell, H. C.; Rawlings, C. J.; Soo, E. C.; Szymanski, C. M.; Dell, A.; Wren, B. W.; Muggleton, S. H. *J. Mol. Biol.* **2013**, *425*, 186−197.

(31) Recent reviews: (a) Thibodeaux, C. J.; Melançon, C. E.; Liu, H. W. *Angew. Chem., Int. Ed.* **2008**, *47*, 9814−9859. (b) Singh, S.; Phillips, G. N., Jr.;

Thorson, J. S. *Nat. Prod. Rep.* **2012**, *29*, 1201−1237. (c) Gantt, R. W.; Peltier-Pain, P.; Thorson, J. S. *Nat. Prod. Rep.* **2011**, *28*, 1811−1853.

(32) Goon, S.; Schilling, B.; Tullius, M. V.; Gibson, B. W.; Bertozzi, C. R. *Proc. Natl. Acad. Sci.* **2003**, *100*, 3089-3094.

(33) Prescher, J. A.; Dube, D. H.; Bertozzi, C. R. *Nature* **2004**, *430*, 873−877.

(34) Anderson, C. T.; Wallace, I. S.; Somerville, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 1329−1324.

# Chapter 4 : Summary and Outlook

## 4.1 Summary

In this work, methods of producing furanose sugar nucleotides using chemoenzymatic and synthetic methodology were explored. In Chapter 2, an attempt to produce UDP-Gal*f* (**1**) using a chemoenzymatic strategy similar to that employed by Gantt *et al.* is described.[1] In this approach, phenyl galactofuranosides **6a-e** and thiophenyl galactofuranosides **8a-b** were synthesized and used as donors in an attempt to force the galactofuranosyltransferase enzyme GlfT2 to work in reverse to produce UDP-Gal*f*. Unfortunately, the wild type enzyme was not capable of turning over either phenyl or thiophenyl derivatives, prompting a redesign of the GlfT2 active site using site-directed mutagenesis.

Seven GlfT2 mutants were created, targeting two regions around the active site. By replacing various bulky residues with alanines or serines, it was thought that poor steric interactions between the enzyme and the phenyl substituent within the active site could be relieved. Sadly, none of the mutants were able to carry out the desired transformation. Another galactofuranosyltransferase, WbbI from *E. coli* K-12,[2] was also examined for activity with these substrates; no activity was observed.

Galactofuranosyl fluorides were then explored as alternative substrates. Although fluorides protected with either benzoyl or TBS groups could be synthesized, the deprotected compound proved to be too unstable to isolate. In light of these negative results, this project was terminated without attempting to optimize the enzyme through directed evolution.

Because UDP-Gal*f* (**1**) can be chemoenzymatically synthesized in other ways,[3] another target (UDP-Gal*f*NAc, **2**) was chosen that had never been produced synthetically or in high yield before. Two nitrogen-containing sugar-1-phosphates (**20** and **32**) were synthesized chemically and then incubated with a uridylyltransferase enzyme (GalPUT) in an attempt to produce the corresponding sugar nucleotides. When none of the desired compound was formed, a newly reported method[4] of coupling phosphates together to form a pyrophosphate bond was used to chemically synthesize UDP-Gal*f*NAc (**2**) and its azido analogue UDP-Gal*f*$N_3$ (**34**).

### 4.2 Outlook

The most immediate use of this work is centred on compounds **2** and **34**. An ongoing collaborative project between the Lowary and Szymanski research groups is the identification and characterization of the enzymes used by *C. jejuni* to form its cell wall polysaccharides. UDP-Gal*f*NAc (**2**) produced synthetically using the method outlined above can be used with these enzymes to determine kinetic parameters and further characterize the proteins. Access to analogues of **2**, such as **34**, will allow the enzyme's active site specificity to be investigated in detail. Knowledge of the active site architecture will aid in the design of new therapeutics against *C. jejuni*.

The demonstration of azide incorporation into the galactan of *M. tuberculosis* also presents attractive metabolic labeling opportunities. If the azide-containing galactofuranose moiety is incorporated into the bacterial cell wall, a

fluorescent label can then be attached *in vivo* through either a cycloaddition with an alkyne[5] or Staudinger ligation.[6] Galactofuranosyltransferase enzymes from *M. tuberculosis* and *C. jejuni* may be specific enough that they take up only certain azide-containing galactofuranose residues, in which case a bacteria-specific diagnostic assay could be developed.[7]

## 4.3 Bibliography


(1) Gantt, R. W.; Peltier-Pain, P.; Cournoyer, W. J.; Thorson, J. S. *Nat. Chem. Biol.* **2011**, *7*, 685–691.

(2) Wing, C.; Errey, J. C.; Mukhopadhyay, B.; Blanchard, J. S.; Field, R. A. *Org. Biomol. Chem.* **2006**, *4*, 3945−3950.

(3) (a) Wagner, G. K.; Pesnot, T.; Field, R. A. *Nat. Prod. Rep*. **2009**, *26*, 1172−1194. (b) Rose, N. L.; Zheng, R. B.; Pearcey, J.; Zhou, R.; Completo, G. C.; Lowary, T. L. *Carbohydr. Res.* **2008**, *343*, 2130−2139. (c) Poulin, M. B.; Lowary, T. L. *Methods Enzymol*. **2010**, *478*, 389−411. (d) Errey, J. C.; Mukhopadhyay, B.; Kartha, K. P. R.; Field, R. A. *Chem. Commun.* **2004**, *23*, 2706−2707.

(4) Tanaka, H.; Yoshimura, Y.; Jørgensen, M. R.; Cuesta-Seijo, J. A.; Hindsgaul, O. *Angew. Chem., Int. Ed.* **2012**, *51*, 11531−11534.

(5) Anderson, C. T.; Wallace, I. S.; Somerville, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 1329−1324.

(6) Prescher, J. A.; Dube, D. H.; Bertozzi, C. R. *Nature* **2004**, *430*, 873−877.

(7) Pons, J. M.; Dumont, A.; Sautejeau, G.; Fugier, E.; Baron, A.; Dukan, S.; Vauzeilles, B. *Angew. Chem. Int. Ed.* **2014**, *53*, 1275–1278.