Progress Towards Identifying The Gal*f*NAc-Transferase In The Biosynthesis Of *Campylobacter jejuni* Strain NCTC11168H Capsular Polysaccharide

by

Anushka B. Jayasuriya

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry
University of Alberta

Abstract

Campylobacter jejuni (*C. jejuni*) is currently the leading cause of food-borne gastroenteritis, and is a precursor to Guillian–Barré and Miller–Fischer syndromes. *C. jejuni* produces a number of unusual carbohydrates on its cell surface that are essential to viability and pathogenicity. Of particular interest is the *C. jejuni* strain 11168H capsular polysaccharide (CPS), which, in addition to having a structurally interesting tetrasaccharide repeating unit, is a virulence factor. The tetrasaccharide consists of β-D-ribofuranose (Rib*f*), 2-acetamido-2-deoxy-β-D-galactofuranose (Gal*f*NAc), α-D-glucopyranosiduron-(2-amino-2-deoxyglycerol)-amide (GlcANGro), and 6-*O*-methyl-D-*glycero*-α-L-*gluco*-heptopyranose (Hep) residues. Inhibitors of CPS biosynthesis are attractive therapeutic targets in circumventing *C. jejuni* associated diseases. However, there remain many questions in regard to the mechanism by which the *C. jejuni* CPS is biosynthesized, and without a proper understanding and identification of the enzymes that catalyze the synthesis of this tetrasaccharide, producing effective inhibitors remains problematic.

The aim of this project was to address some of these deficiencies. The main focus was to identify the Gal*f*NAc transferase that catalyzes the coupling of the Gal*f*NAc onto the GlcANGro in the tetrasaccharide repeating unit. This involved cloning and expressing putative GTs in the cps gene cluster (*cj*1438, 1440, and 1442), as well as chemically synthesizing potential donor and acceptor substrates to test with these enzymes. It concluded with an enzymatic assay designed to probe for Gal*f*NAc transferase activity.

Preface

Part of Chapter 2 has been published: Peng, W.; Jayasuriya, A. B.; Imamura, A.; Lowary, T. L. *Org. Lett.* 2011, *13* (19), 5290–5293. My contributions included the scale up and optimization of key reaction conditions, as well as developing a synthetic route toward the heptose-BSA glycoconjugate. The manuscript was written jointly between me and Dr. Todd L. Lowary.

The rest of the work presented in this thesis is unpublished at the time of writing and was carried out solely be me, unless otherwise indicated in the text.

The cloning experiments presented in Chapter 4 were completed with lab technician Cory Q. Wenzel of Dr. Christine Szymanski's laboratory in the Department of Chemistry at the University of Alberta.

Acknowledgements

First, I would like to acknowledge my supervisor, Dr. Todd L. Lowary. His guidance and support over the years has meant a great deal to me, and I am forever thankful he decided to take a last minute meeting with a curious undergraduate, and gave him the chance to work in his group. I feel a debt of gratitude that I am afraid will never be fully paid. Todd created a wonderful environment in which to work: it was intellectually stimulating, and we were encouraged to come up with our own ideas, and try new things. I got the chance to attend conferences that were inspiring, where I learned a great deal, and met amazing people. Most importantly, Todd helped foster a culture of creative thinking, mentorship, and nurture, that passed on to everyone in the lab: knowledgeable advice and a helping hand were never far out of reach.

I am also grateful for my Supervisory Committee and my PhD Examination Committee members: Professor Jeffrey Stryker, Christopher Cairo, Derrick Clive, Rylan Lundgren, Brian Rempel, and Suzanne Walker, for their guidance and help throughout the thesis research. I would also like to thank Professors Christine Brzezowski and Jeffrey Stryker for making me fall in love with organic chemistry in the first place, and inspiring me to pursue Organic Chemistry in Graduate School.

I would like to thank Dr. Wenjie Peng, from whom I took over the heptose project. His work set a foundation for this research. As well, a special thanks go to Roger Ashmus, who in addition to enduring being my roommate for two years, collaborated with me for the second heptose route; I am lucky to have him as a friend. Special thanks also goes out

to Dr. Myles Poulin for his early guidance in in this project, and Ryan Snitynsky for making some of the precursors I used in my synthesis of enzyme substrates. I would also like to thank Cory Wenzel, whose help and insight were invaluable. I learned a great deal working under him, and I don't think any of the cloning and expression work would have been nearly as fruitful without his aid. I would also like to thank Dr. Christine Szymanski, who welcomed me into her wonderful lab to do a portion of the biochemical work for this thesis. Lastly, I would like to thank Justin Thuss for helping me proofread and edit this thesis.

I consider myself very lucky to have had the chance to work with amazing lab mates throughout my studies, and every member of the Lowary group is included in that list. Roger, Sweeney, Mickey, Peter, Myles, Lily, Jeff, Lu, Amira, Aki, you made coming into the lab every day enjoyable. You are all an amazing group of people, and like a second family to me, never lacking in good friendship, advice, and collaboration. I would like to give special thanks to those that impacted me in my formative years, Dr. Maju Joe and Yu Bai, who took me under their reigns from day one.

In addition, my research has tremendously benefited from the research services in the Department. I am grateful to the NMR staff and all the people in the Spectral Services Labs, the Mass Spectrometry Lab, the machine shop, the storeroom and the glass shop. I am also thankful for all the help from our secretary Lynne Lechelt and the graduate program assistant Anita Weiler and Jillian Fell for keeping me in line with all important paper work and supporting instructions.

I have met some of the very best people I know during my time in the Department of Chemistry, and for that I consider myself very fortunate and humbled. I would like to thank some of my closest friends for their support, laughs, and adventures. Colin, Jinyue, Sadek, Downey, Thuss, Owen, Tom, I'm lucky to have you in my life, and your friendship has made me a better person.

Lastly, I would like to thank my friends and family for their enduring support throughout my studies. To my mother and father, Nandanie Jayasuriya and Ananda Jayasuriya, none of this would have been possible without you being there for me.

Table of Contents

List of Figures

List of Schemes

List of Tables

List of Abbreviations

ADDP	1,1'-(azodicarbonyl)dipiperidine
BAIB	[bis(acetoxy)iodo]benzene
BLAST	Basic Local Alignment Search Tool
BSA	bovine serum albumin
C. jejuni	*Campylobacter jejuni*
CAZy	Carbohydrate-Active Enzyme
CE-PDA	capillary electrophoresis with a photodiode array detector
Cm	chloramphenicol
CPS	capsular polysaccharide
DEAD	diethyl azodicarboxylate
DIAD	diisopropyl azodicarboxylate
DMP	dimethoxypropane
EDTA	Ethylenediaminetetraacetic acid
GAGs	glycosaminoglycans
GalE	UDP-Glc*p*NAc–UDP-Gal*p*NAc by the epimerase
GalE-MBP	UDP-Gal*p*NAc-4-epimerase–maltose binding fusion protein
Gal*f*NAc	2-acetamido-2-deoxy-β-D-galactofuranose
GlcANGro	α-D-glucopyranosiduron-(2-amino-2-deoxyglycerol)-amide
GlfT2	UDP-galacofuranosyl transferase
GT	glycosyltransferase

Hep	6-*O*-methyl-D-*glycero*-α-L-*gluco*-heptopyranose
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMW	high molecular weight
HPLC	High-performance liquid chromatography
ImIm	2-imidazolyl-1,3-dimethylimidazolinium chloride
IPTG	isopropyl 1-thio-β-D-galactopyranoside
Kan	Kanamycin
Kdo	2-keto-3-deoxy-octulosonic acid
LB	Luria Bertani
LDH	lactate dehydrogenase
LMW	low molecular weight
LOS	lipooligosaccharide
LPS	lipopolysaccharide
m-CPBA	meta-Chloroperoxybenzoic acid
MALDI	matrix-assisted laser desorption/ionization
Me-OPN	methyl phosphoramidate
MHC	major histocompatibility complex
MOPS	3-(N-Morpholino)propanesulfonic acid
MS	Mass Spectrometry
MW	molecular weight
NADH	β-nicotinamide adenine dinucleotide
NBS	N-bromosuccinimide
NCBI	National Centre for Biotechnology Institute

Ni^{2+}-NTA	Nickel-Nitriloacetic acid
NIS	N-iodosuccinimide
NMR	Nuclear Magnetic Resonance
OD	optical density
PCR	Polymerase chain reaction
PK	pyruvate kinase
R-CBS	(R)-(+)-2-Methyl-CBS-oxazaborolidine
Rib	Ribose
TB	Teriffic Broth
TBAF	tetrabutyl- ammonium fluoride
TBDPS	*t*-butyldiphenylsilyl
TBTU	N,N,N',N'-tetramethyluronium tetrafluoroborate
TEMPO	2,2,6,6-tetramethylpiperidin-1-yl)oxyl
TLC	thin-layer chromatography
TMAD	N,N,N',N'-tetramethylazodicarboxamide
TMSOTf	Trimethylsilyl trifluoromethanesulfonate
Tris	Tris(hydroxymethyl)aminomethane
UDP-Gal*f*NAc	uridine disphospho-2-acetamido-2-deoxy-α-D-galactofuranose
UDP-Gal*p*NAc	uridine disphospho-2-acetamido-2-deoxy-α-D-galactopyranose
UGM	UDP-galactopyranose mutase
UMP	uridine-5′-monophosphate

UNGM	UDP-N-acetylgalactopyranose mutase
UV	ultraviolet

1 Introduction

1.1 Research Overview and Objectives

Campylobacter jejuni (*C. jejuni*) is currently the leading cause of food-borne gastroenteritis, and is a precursor to Guillian–Barré and Miller–Fischer syndromes.[1–3] *C. jejuni* produces a number of unusual carbohydrates that are essential to viability and pathogenicity. Of particular interest is the *C. jejuni* serotype NCTC11168H capsular polysaccharide (CPS), which, in addition to having a structurally unique tetrasaccharide repeating unit (Figure 1-1), is an important virulence factor.[4,5] Inhibitors of CPS biosynthesis are possible therapeutic targets in circumventing diseases associated with *C. jejuni*. However, there remain many questions in regard to the mechanism of CPS biosynthesis in *C. jejuni*. Without a proper understanding and identification of the enzymes that catalyze the synthesis of this tetrasaccharide, producing effective inhibitors remains problematic. This thesis aims to carry out investigations to identify the enzymes (glycosyltransferases, GTs) that catalyze the coupling of the carbohydrate residues in the tetrasaccharide repeating unit. The enzyme that introduces the 2-acetamido-2-deoxy-β-D-galactofuranose (Gal*f*NAc) residue (Figure 1-1, ring B), a Gal*f*NAc-transferase, will be the primary focus.

The structure of *C. jejuni* NCTC 11168 (HS:2) CPS is shown below. There are, however, many different CPS structures for *C. jejuni* that are strain dependent, and a full review has been published.[3]

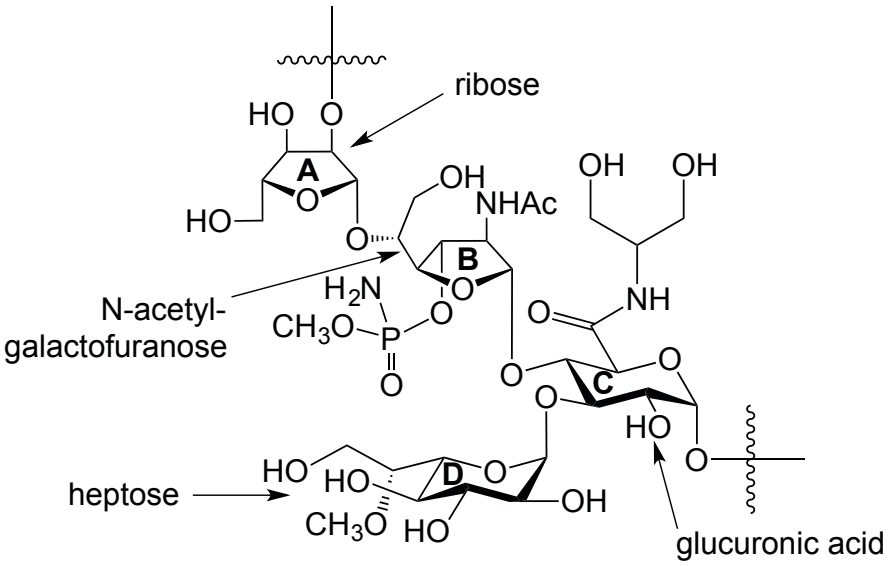


Figure 1-1: Structure of *C. jejuni* NCTC 11168 (HS:2) CPS repeating unit. **(A)** β-D-ribofuranose, **(B)** 2-acetamido-2-deoxy-β-D-galactofuranose, (**C**) α-D-glucopyranosiduron-(2-amino-2-deoxyglycerol)-amide, **(D)** 6-*O*-methyl-D-*glycero*-α-L-*gluco*-heptopyranose.

This introduction will begin with an overview of the organism, *C. jejuni*, and *C. jejuni* associated disease. Focus will then shift to the virulence factors that contribute to disease progression, with a sole focus on the CPS of serostrain 11168 (HS:2), and its initial discovery. CPS assembly and export will be examined, with special attention to the GTs responsible for elongating the CPS repeat unit. An overview of the genome and assignment of putative function based on sequence homology will lead to the first aim of the project: the cloning and expression of the most likely Gal*f*NAc-transferase gene candidates. Further in depth structural analysis will lead to the second aim: the chemical synthesis of potential donor and acceptors to test with the putative GTs. Lastly, visualization and characterization methods for the potential products of the Gal*f*NAc-transferase will be discussed.

1.2 *Campylobacter jejuni*: Associated Diseases and Virulence Factors

Campylobacter, consisting of 16 species and an additional 6 subspecies, is a member of the epsilon proteobacteria.[1] It is closely related to Gram-negative mucosal pathogens such as *Haemophilus influenzae* and *Neisseria meningitides* due to it being microaerophilic (requiring oxygen for growth, but at a lower amount than the atmosphere), naturally transformable (can actively take up free DNA and readily incorporate its genetic information), and encapsulated. *C. jejuni* exist as commensals in the gastrointestinal tract of domestic and wild birds, as well as mammals. Moreover, *C. jejuni* is currently the leading cause of food-borne gastroenteritis, making it a subject of research interest. Approximately 1% of Europeans suffer from campylobacteriosis, resulting in an annual cost of €2.4 billion; in the United States costs amount to $2.9 billion each year.[6] Over 90% of these cases are attributable to *C. jejuni*. Consumption or handling of contaminated poultry products is the primary source of human infection.[3,7] In addition to bacterial diarrhea, less than 1 in 1000 of those infected with campylobacter enteritis will develop Guillan–Barré or Miller–Fischer syndrome, which are chronic, and sometimes fatal, paralytic disorders.[2] The cause of paralysis is due to the host immune response, which, when targeting the *C. jenuni* lipopolysaccharide (LPS), produces antibodies with the same surface recognition moiety as those found on native peripheral nerves.[8] Consequently, the host mounts an autoimmune response that destroys myelin, and causes nerve damage.

Although *C. jejuni* is a causative agent of gastroenteritis, little is known as to the mechanism of *C. jejuni* survival and infection. This is due to the lack of non-primate models for diarrhoeal disease.[13] Thus, the route from environmental contaminant, to chicken ceca, to poultry carcass

contaminant, and finally human disease agent, contains many hurdles for bacterial survival and infectivity.

It is postulated that *C. jejuni* is capable of triggering various adaptive responses to different environmental stresses, due to its persistence in the gastrointestinal tract of poultry, a hostile environment. It is also known that *C. jejuni* colonizes the lower gastrointestinal tract after ingestion. Colonization alone, however, can be asymptomatic, and disease outcome is more dependent on the immune status of the host as well as the virulence characteristics of the campylobacter strain. It has been demonstrated that capsules are virulence factors in other pathogens. However, the role of CPS in *C. jejuni* disease, specifically, has not been well demonstrated except for descriptive studies demonstrating serum resistance and onset of diarrhea in a ferret model of disease.[3] One compelling piece of evidence giving significance to the CPS as a virulence factor in *C. jejuni* infections, and warranting the further study of its role as such, is that conjugate vaccines against CPS protect against diarrhoeal disease in non-human primates.[1]

As is a major area of concern for a wide variety of bacterial pathogens, *C. jejuni* has developed antibiotic resistance.[9–11] This has been attributed to the fact that, acting as a commensal in poultry and cattle, *C. jejnui* has been exposed to sub-therapeutic antibiotics for years as a measure of increasing output.[12] Thus, the exposure of *C. jejuni* to antibiotic treated commensal hosts allowed it to develop antibiotic resistance before transfer to humans.

Despite the growing contributions to understanding the progression and onset of campylobacteriosis, and its mechanism of virulence, there remain large areas for which, at the

molecular level, our understanding is lacking. Of specific concern is the biosynthesis and export machinery of polysaccharides in Gram-negative bacteria. Consequently, the need for investigations to address this gap in knowledge is ever increasing. Such studies will not only aid in the understanding of how these bacteria produce such diverse and elaborate virulence factors, but also lead to the production of novel chemotherapeutic agents that selectively target and disrupt capsule synthesis and export, to combat infection.

1.3 Bacterial pathogens and virulence factors

C. jejuni, like all bacterial pathogens, have evolved an array of physiological adaptations that allow them to survive under the harsh conditions in a host organism. These conditions arise from an environment devoid of natural elements required for bacterial growth, antibacterial factors in secretions covering mucosal surfaces, and phagocytic cells recognizing bacteria as foreign.[13]

For bacteria to cause disease, they need to meet two key criteria before colonizing: the capacity to infect, and to invade the host organism.[14] There are a number of cellular components required to overcome the host resistance mechanisms that threaten bacterial survival, and these are referred to as virulence factors. The functions of virulence factors are threefold; they aid the bacteria in (1) invading the host, (2) causing disease, and (3) evading host defenses.[1] In regard to Gram-negative bacteria such as *C. jejuni*, these virulence factors include cell surface and extracellular components—such as flagella and associated chemotaxis proteins—as well as the lipopolysaccharide (LPS) and capsular polysaccharide (CPS).[14] Bacterial infectivity results when there is an imbalance between bacterial virulence factors and host resistance. This thesis will focus on one of these virulence factors: the CPS.

1.4 Capsular Polysaccharides

Surface polysaccharides represent the predominant structures of all bacterial cell surfaces, and function to maintain surface charge, phase variation (are displayed in an on-off fashion in response to environment), as well as providing serum resistance.[15] Polysaccharide capsules, in particular, are essential for mediating the interaction between pathogens, their hosts, and the environment, because they are located on the outermost layer of the cell. As a result, they are often the first bacterial structure encountered by the host immune system. CPSs are widely dispersed and found in a diverse class of pathogens, including *Escherichia coli,* and *Neisseria meningitides,* and are also found in important Gram-positive pathogens such as *Staphylococcus aureus* and *Streptococcus pneumonia*. CPSs form an extensive layer around the cell, and are therefore often observable by electron microscopy (Figure 1-2). In *E. coli*, for example, the CPS can extend anywhere from 100–400 nm, and is formed by a series of glycan chains that are over 200 sugars long.[16] A single species of bacteria can synthesize a range of CPS; therefore, they act as a basis for serotyping schemes.

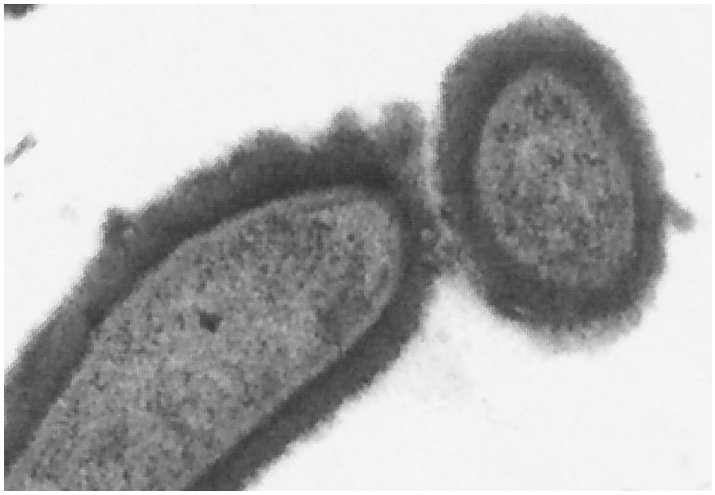


Figure 1-2: *C. jejuni* visualized by Electron Microscopy using Alcain Blue stain. Reprinted with permission form American Society for Microbiology: *Infect. Immun.,* **69**, 5921–5924, copyright 2001.[17]

These sugars also play a crucial role in bacterial survival and virulence by preventing desiccation, assisting in adherence, providing resistance to non-specific and specific host immunity, as well as mediating the diffusion of molecules through the cell surface.[18] Extracellular polysaccharides are classified into CPS and slime polysaccharides, where the former is intimately associated with the cell surface through a phospholipid linker, and the later loosely associated (Figure 1-3).[18] A CPS-lipid linker has been identified as the most likely means of association to the bacterial cell surface, but other possibilities exist, including ionic interactions between the CPS and the core region of the LPS.[19] Capsule genes have a significantly higher A+T composition than the rest of the chromosome, suggesting a common ancestry for these genes in Gram-negative bacteria, with the most likely candidate being horizontal gene transfer.[18]

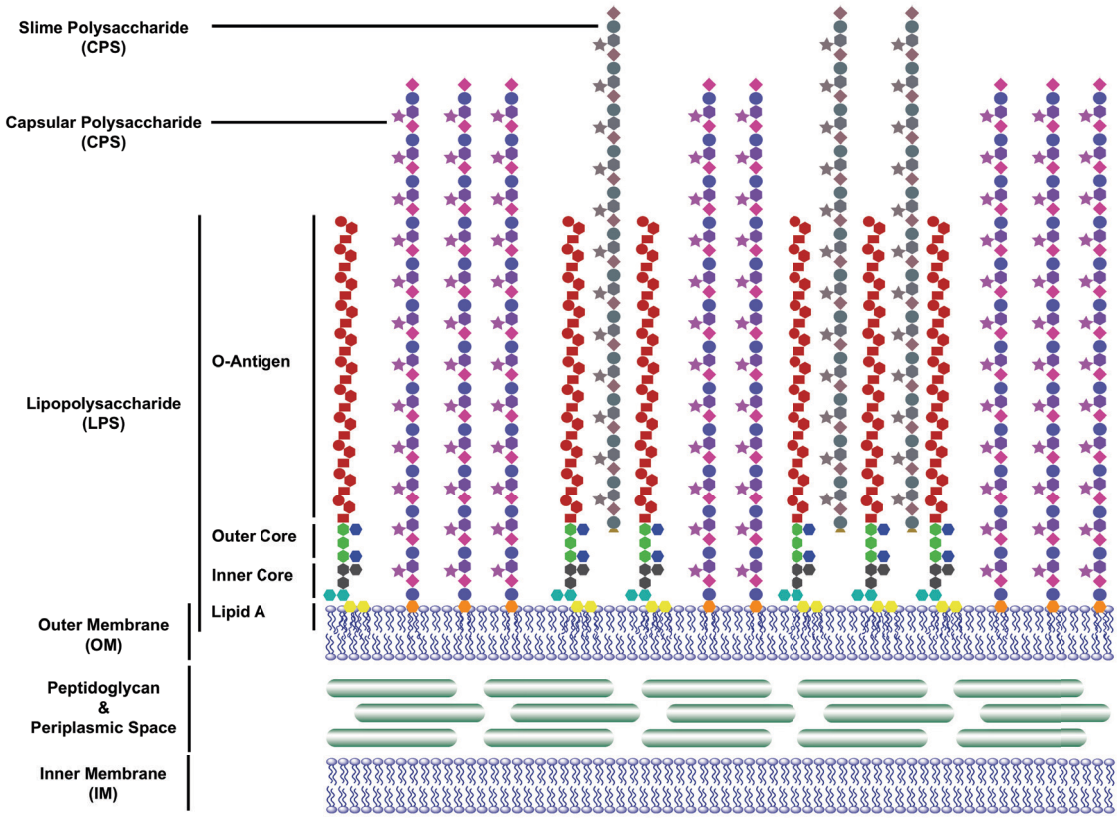


Figure 1-3: Cell Surface Glycans: General depiction of Gram-negative bacterial cell wall with Inner Membrane (IM), Periplasmic space with peptidoglycan, Outer Membrane (OM), Lipopolysaccharide (LPS, containing the O-Antigen, Outer Core, Inner Core, and Lipid A), and Capsular Polysaccharide (CPS).

1.4.1 Prevention of Desiccation

CPSs are highly hydrated molecules that are over 95% water. As a result, they form a hydrated gel around the cell surface that acts to protect the bacteria from the harmful effects of desiccation. This results in the increased survival of encapsulated bacteria outside of the host compared to non-encapsulated bacteria, which, in turn, promotes the transmission of bacteria from one host to the other.[18,20,21] Some evidence suggests that bacteria have evolved a mechanism by which they regulate capsule expression in response to desiccation; when external osmotic concentration is altered, an increase in capsule biosynthesis is triggered.[20,21]

1.4.2 Adherence

CPSs have been shown to facilitate the formation of biofilms, as well as the ability to colonize in various ecological niches.[22] This has been postulated to be due to their ability to promote adherence, both to each other and various surfaces. The formation of a biofilm is advantageous to individual bacteria in that it provides protection from phagocytic protozoa, infection by bacteriophages, as well as nutritional advantages, due to biofilm nutrient trapping.[22] The biofilm has also been shown to establish a microbial consortium, and the close interaction between individual bacteria is believed to arise as a result of colonizing bacteria providing bridges of attachment for subsequent bacteria introduced in the environment.[23,24] In addition to the adherent properties in bacterial biofilms, it has also been suggested that some CPSs have lubricant properties that allow their swarming over substrata by reducing friction.[25]

1.4.3 Resistance to Host Immunity

Because the CPS often constitutes the outermost layer of the bacterial cell surface, it is the first bacterial structure encountered by the host immune system. Therefore, it plays an essential role in evasion of the host immune system, and so it is not surprising that many acapsular bacterial isolates are non-pathogenic.[14] Consequently, CPS structures have been used to develop vaccines against *N. meningitides.*[1] There is, however, a fear in using more commonly seen CPS structures as vaccine candidates, due to their similarity to human glycans such as glycosaminoglycans (GAGs). For example, *Pasteurella multocida* and *E. coli* make CPSs that are either similar or identical to the nonsulfated backbone of hyaluronan, chondroitin sulfate, and heparosan. In contrast *C. jejuni,* the subject of this thesis, offers an advantageous feature in its CPS: its CPS contains very unusual sugars.

The CPS interferes with host innate immunity at several levels. It acts as a permeability barrier for complement components, and interferes with the attachment of the bacterial cell surface to phagocytes. Although the complement system can penetrate the CPS, the CPS effectively masks the underlying cell surface from activating the complement system.[26] In turn, the CPS is effectively responsible for conferring resistance to non-specific immunity of the host.[18,26,27] There is also evidence that suggests that the CPS confers resistance toward complement and non-complement mediated opsonophagocytosis.[21,28] Although the precise mechanism is not known, it has been shown that an increase in net negative charge on the CPS is correlated with enhanced complement resistance.[28] The capsule usually acts alongside other cell surface carbohydrates, such as the O-antigen, to confer complement resistance.[29] Additionally, some capsules modulate the release of cytokines by the host, thereby mediating its immune response by disrupting the coordination of the cell-mediated immune response.[30]

CPSs, like most carbohydrates, are poor immunogens, and this feature is advantageous for bacterial survival. These CPSs are referred to as T lymphocyte-independent antigens due to their inability to stimulate major histocompatibility complex (MHC) presentation on helper T cells.[31] As a result, the T cells are unable to generate immunological memory.

1.4.4 Identification of Lipopolysaccharide on *Campylobacter jejuni*

Differentiation and identification of CPS from other forms of bacterial extracellular polysaccharides remains problematic for a variety of reasons. These include the fact that the CPS may be released from the cell surface in extraction steps as a consequence of the susceptible

phosphodiester linkage connecting the CPS to phospholipid on the outer membrane, giving the appearance of slime polysaccharides. As well, distinguishing between CPS and O-antigenic moieties of the LPS remains challenging, because the CPS is found to be associated with the LPS (See Figure 1-3). CPS was not initially identified as a *C. jejuni* surface polysaccharides, and instead, it was classified as a high molecular weight (HMW) LPS.[32]

Early studies into the identification of surface polysaccharides in *C. jejuni* showed the presence of a low molecular weight (LMW) LPS lacking O-antigen repeats (See Figure 1-3); this was referred to as the lipooligosaccharide (LOS).[33] Further structural and chemical characterization revealed inner and outer core regions linked to Lipid A by 2-keto-3-deoxy-octulosonic acid (Kdo), which, being a conserved component of the LPS structure, further validated its classification as such.[34] Additionally, one third of *C. jejuni* serotype reference strains showed an additional HMW repeat unit structure, thought to be another LPS; this was detected by western blot analysis with Penner typing antisera.[35] The HMW LPS was referred to as O-antigenic, and both the sugar composition and the structure of the glycans isolated from different strains were observed. The classification of *C. jejuni* surface polysaccharides into HMW and LMW LPS, however, was insufficient in explaining a key a set of observations obtained in further studies of surface polysaccharide structure, export, and assembly. It was not until many years after that HMW LPS was classified as CPS (More in **Section 1.5**).

1.5 CPS Assembly & Export: Reclassifying HMW LPS as CPS

The biosynthesis and transport of LPS is divided into two classes: wzy-dependent (Figure 1-4) and wzy-independent (Figure 1-5). The wzy-dependent pathway is characterized by en block transfer, in which a repeat unit structure of carbohydrate is polymerized in the periplasm by the

wzy polymerase. The wzy-independent pathway is characterized by processive elongation and polymerization of the repeat unit carbohydrate in the cytosol by glycosyltransferases. The latter, having been identified in *C. jejuni* CPS biosynthesis and export, will be the focus of this section. Further, examining the wzy-independent pathway (also referred to as the ABC transport system), and its characteristic markers, will help elucidate why HMW LPS was classified as CPS.

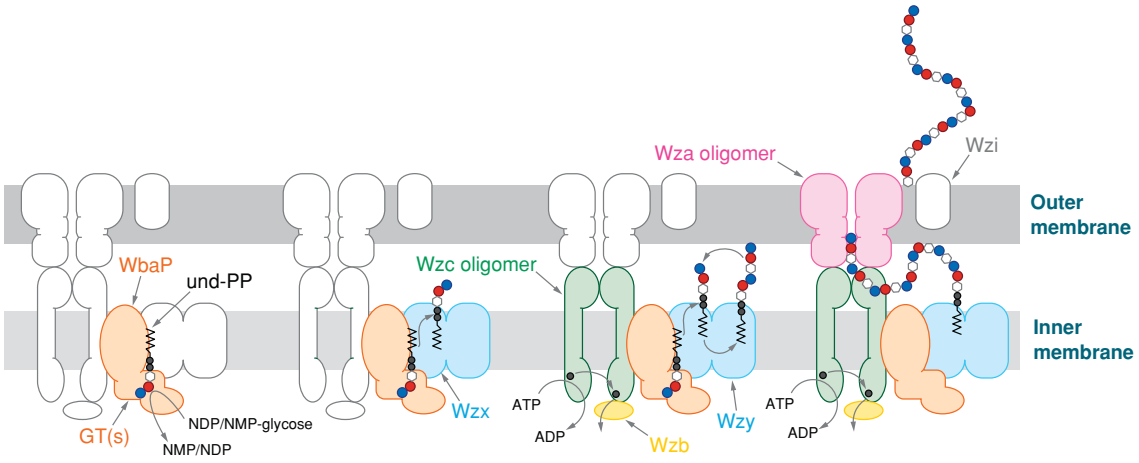


Figure 1-4: Wzy-dependent Transport:[15] Und-PP repeat units are synthesized by glycosyltransferases (GTs) at the inner membrane, with WbaP acting as the initiating transferase, attaching the first sugar unit to the und-PP. Wzx is responsible for flipping the synthesized und-PP-linked repeats to the periplasm, where Wzy-dependent polymerization occurs by en bloc transfer (polymer grows by transfer of growing chain to the incoming und-PP-linked repeat unit). Wzc and Wzb aid in continued polymerization; because the periplasm does not contain ATP as an energy source, transphosphorylation of Wzc on the inner membrane occurs, with dephosphorylation by the Wzb phosphatase. The polymer is then translocated to the outer membrane by Wza, which acts as a channel. Reprinted with permission form Annual Review of Biochemistry: *Annu. Rev. Biochem.*, ***75***, 39–68, copyright 2006.

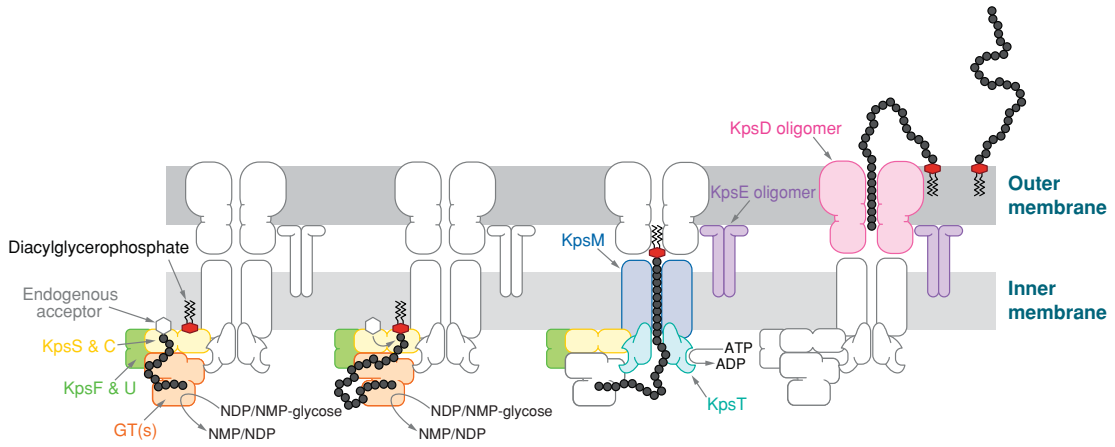


Figure 1-5: Wzy-independent ABC Transport:[15] Polymer formation is initiated on an unknown acceptor (hexagon), and is processively elongated by glycosyltransferases (GTs) on the inner membrane. The product is linked to a lipid phosphate (or lipid phosphate-Kdo), though at what point of the polymerization step (initiation, during, or after polymerization) is unknown. The polymer is exported by the ABC transporter KpsM and KpsT. KpsS and KpsC are essential for capsule transport, however the precise function has not been elucidated. KpsE and KpsD are required for translocation from the periplasm to outer membrane. KpsE is a putative membrane-fusion protein, which couples the ABC transporter to later translocation steps, with KpsD being a candidate for capsular polysaccharide translocation channel. KpsD requires KpsE for its proper localization. **Reprinted with permission form** Annual Review of Biochemistry**:** *Annu. Rev. Biochem.*, ***75***, 39–68, **copyright** 2006**.**

1.5.1 ABC Transport

The ABC transport system consists of four main components: two identical nucleotide binding domain polypeptides (KpsT); two integral membrane polypeptides (KpsM); a polysaccharide co-polymerase (KpsE) that transports the polysaccharide chain from the periplasm to the outer membrane and an outer membrane polysaccharide protein (KpsD). This group of proteins, referred to collectively as KpsTMED, are responsible for the synthesis and transport of the growing polysaccharide chain from the cytosol to the outer membrane of the bacteria in a coordinated manner (Figure 1-5). There are other components as well: the repeat unit structure of the CPS is made by GTs (**Glycosyltransferases 1.7**). The GTs catalyze the transfer of sugars

from an activated donor—in this case a sugar nucleotide—to a substrate—another sugar (though protein and lipid linkages are also possible with this class of enzyme). Additionally, two conserved genes, KpsS and KpsC, have been shown to transfer Kdo from CMP-Kdo to mono- and di-acyl phosphatidylglycerol *in vitro*. That these two genes are conserved across all CPS-containing bacteria, which points to a process specific to the biosynthesis of CPS rather than LPS, remained pivotal to the characterization of HMW LPS as CPS.[36]

Genome sequencing of ~100 clones by Karlyshev and coworkers identified genes encoding proteins similar to those that biosynthesize the CPS in other bacteria.[32] The mutational analysis of KpsC and KpsS, as well as KpsM, showed that they were integral in forming HMW LPS and not LMW LPS.[32] Identifying the presence of a wzy-independent pathway/ABC transporter-dependent pathway, alone, did not provide sufficient evidence for the presence of a CPS, because this pathway is responsible for the assembly and export of both LPS and CPS onto bacterial cell surfaces.

One of the characteristic features of the model CPS assembled by the ABC transporter-dependent pathway is the presence of a phospholipid on the reducing end of the polysaccharide chain.[14] This acts in contrast to LPS systems, where the defining anchor is Lipid A **(**Figure 1-6**)**.[15]

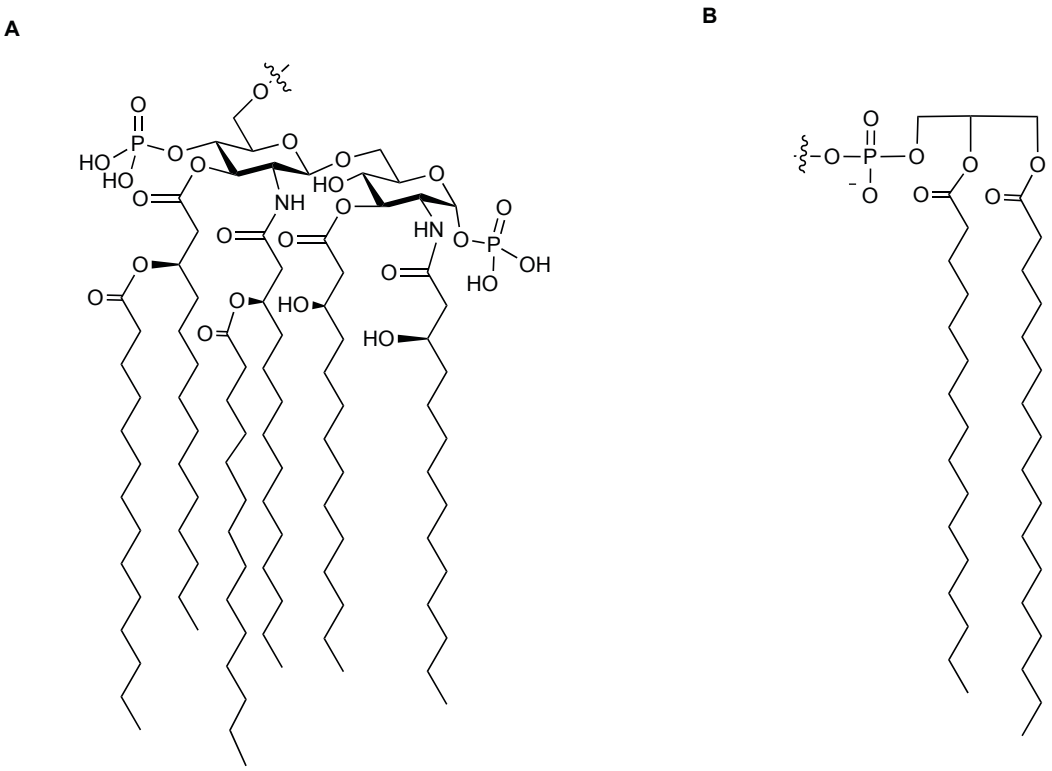


Figure 1-6: (A) Lipid A anchor for LPS, (B) Dipalmitoyl glycerophosphate anchor from *C. jejuni* 81-176 (HS:23 and HS:3) and 81-116 (HS:6).

In order to distinguish between the presence of a Lipid A anchor—which would suggest that HMW LPS was indeed LPS—or phospholipid anchor—which would suggest that HMW LPS was a CPS—*C. jejuni* cells were subjected to phospholipase treatment. An absence of HMW LPS on the bacterial cell surface suggested the anchor for this polysaccharide was a phospholipid, which is susceptible to such a treatment, and not Lipid A. This observation provided support that HMW LPS was indeed a CPS.[32] Subsequent studies have identified the lipid anchor of *C. jejuni* 81-176 (HS:23), serostrain HS:3, as well as *C. jejuni* 81-116 (HS:6) CPS as dipalmitoyl-glycerophosphate.[37]

Further support for the presence of a CPS in *C. jejuni* was provided by electron microscopy with Alcian blue staining, which showed the presence of a capsule-like structure.[38] Electron microscopy coupled with mutational studies of KpsM showed retention of the polysaccharide capsule in the wild type *C. jejuni*, whereas the KpsM mutant showed an absence of this structure on the cell surface. This further validated that the CPS was exported by the ABC-transport dependent pathway.[38]

After having identified the presence of a CPS on *C. jejuni,* its mechanism of export, and general structural features (phospholipid anchor vs Lipid A linkage), focus shifted to a more detailed structural understanding of the repeat unit polysaccharide chain and its biosynthesis. This thesis will focus on the structural features and biosynthesis of one strain of *C. jejuni* in particular: 11168 (HS:2)

1.6 *C. jejuni* NCTC 11168 (HS:2) CPS: Structure

1.6.1 Repeat Unit Tetrasaccharide Structure of *C. jejuni* NCTC 11168 (HS:2) CPS

The structure of the *C. jejuni* NCTC 11168 (HS:2) CPS was first elucidated by Monteiro and coworkers in 2002, relying primarily on tandem capillary electrophoresis-electrospray ionization-mass spectrometry (CE-ESI-MS) and High Resolution Magic Angle Spinning NMR (HR-MAS).[39] The structure of the CPS repeating unit was determined to consist of a β-D-ribofuranose (β-D-Rib*f*, ribose), 2-acetamido-2-deoxy-β-D-galactofuranose (β-D-Gal*f*NAc), α-D-glucopyranosiduron-(2-amino-2-deoxyglycerol)-amide linear chain (GlcANGro) with a 6-*O*-methyl-D-*glycero*-α-L-*gluco*-heptopyranose (heptose) residue as a side-branch (Figure 1-1).

1.6.2 Phase Variable Modifications

A large degree of variability in CPS structure was found upon examination of *C. jejnui* 11168 cells by HR-MAS. Genome sequencing suggests this may be attributable to the large proportion of phase-variable genes that are found within the CPS gene locus.[7] Further analysis led to the identification of a variable 6-*O*-methyl group on the heptose, high levels of the *N*-ethanolamine modification on glucuronic acid instead of 2-amino-2-deoxyglycerol, and lastly a unique and structurally interesting *O*-methyl phosphoramidate (MeOPN) moiety on the Gal*f*NAc and heptose, which at this point, has not been described anywhere else in nature.[40] Subsequent detection has identified the *O*-methyl phosphoramidate in 70% of strains from various clinical presentations, geographical locations, and animal sources, expressing different capsule structures. This moiety is versatile in that it is added to different sugars in different strains of *C. jejuni*. Such examples include: the β-fructofuranoses in the CPS of strain *C. jejuni* HS:1;[41] N-acetylglucosamine in *C. jejuni* HS:19 CPS;[42] galactose in *C. jejuni* 81–176 CPS;[43] 6-deoxy-β-D-ido-heptose in *C. jejuni* CG8486 CPS (HS:4);[44] O-3 of Gal*f*NAc and O-4 of 3,6-di-*O*-methyl-D-*glycer*o-α-L-*gluco*-heptose in *C. jejuni* 11168 CPS.[45]

Through western blot analysis, differences in immunoblotting patterns have been observed as a result of changes in capsule modification. Therefore, it has been postulated that phase variable modifications have the ability to affect the immunogenicity of the pathogen.[40,46] Variation is not only limited to the phase variable modifications that decorate the CPS, for *C. jejuni* strains 81-176 and 11168 have demonstrated an on–off variation in total capsule expression. The ability to shut off capsule expression has been suggested to be advantageous in the life cycle of the

pathogen. The mechanism of on–off expression is most likely attributable to slip strand mismatch repair of yet unknown gene(s).[1]

1.6.3 Gal*f*NAc

Of particular interest is the Gal*f*NAc residue on the repeat tetrasaccharide chain, which will be a focus of this thesis. Hexose sugars are found predominantly in the thermodynamically favourable pyranose form in nature. In contrast, the five-membered ring furanose form is seen more scarcely, and is often observed in bacteria, fungi, and parasites.[47,48] Due to the absence of hexofuranosides in mammalian glycan structures, interest has risen in identifying inhibitors of furanoside biosynthesis to target pathogenic organisms selectively. More specifically, galactofuranose (Gal*f*), contained on cell surface polysaccharides, has been shown to be essential for cell viability, in addition to playing a crucial role in cell physiology, of pathogenic microorganisms such as *Aspergillus fumigatus, Trypanosoma cruzi, Mycobacterium tuberculosis,* and *Klebsiella pneumonia.*[49,50]

The *C. jejuni* NCTC 11168 CPS features a Gal*f*NAc moiety, modified with a MeOPN group on O-3.[39] This is one of the few examples of Gal*f*NAc observed in nature, and little is known of the biosynthesis of these sugars. Previous work in our group has identified a pyranose–furanose mutase (cj1439c, UNGM) that interconverts uridine disphospho-2-acetamido-2-deoxy-α-D-galactopyranose (UDP-Gal*p*NAc) and uridine disphospho-2-acetamido-2-deoxy-α-D-galactofuranose (UDP-Gal*f*NAc).[51] However, no Gal*f*NAc transferase has been identified to date, and indeed, at the time the structure of *C. jejuni* NCTC 11168 (HS:2) CPS was elucidated, it was the first time Gal*f*NAc was reported in nature.

1.7 Glycosyltransferases

1.7.1 Overview

The complexity found in various oligosaccharide structures in nature arises from the enzymatic formation and breakdown of glycosidic linkages catalyzed by GTs, glycosidases, glycan phosphorylases, and polysaccharide lyases.[52] GTs are essential for producing large, diverse, and complex glycoconjugates by catalyzing glycosidic bond formation between an activated carbohydrate donor and acceptor molecule (Figure 1-7). CPSs are included in this category. Any two monosaccharides can be joined in a number of configurations, owing to multiple hydroxyl groups on each monosaccharide; as well, they are capable of forming these glycosidic bonds stereoselectively. Another layer of complexity is introduced in the form of branching in the polysaccharide chain, thus GT-catalyzed biosynthesis is responsible for the incredible amount of structural complexity observed in the CPS.[21]

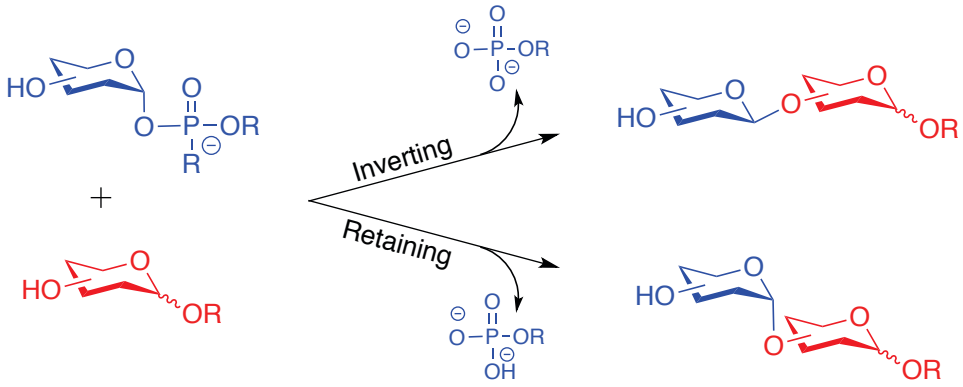


Figure 1-7: General glycosyltransferase reaction for sugar nucleotide with sugar acceptor.

GTs use an activated sugar donor as a substrate. Donors most often contain a phosphate leaving group in the form of a nucleoside diphosphate sugar (most common, **A**, Figure 1-8), nucleoside

monophosphate (e.g., cytidine monophosphate-N-acetyl-β-neuraminic acid, **B**, Figure 1-8), and lipid phosphates (e.g., dolichol phosphate oligosaccharides, **C**, Figure 1-8).

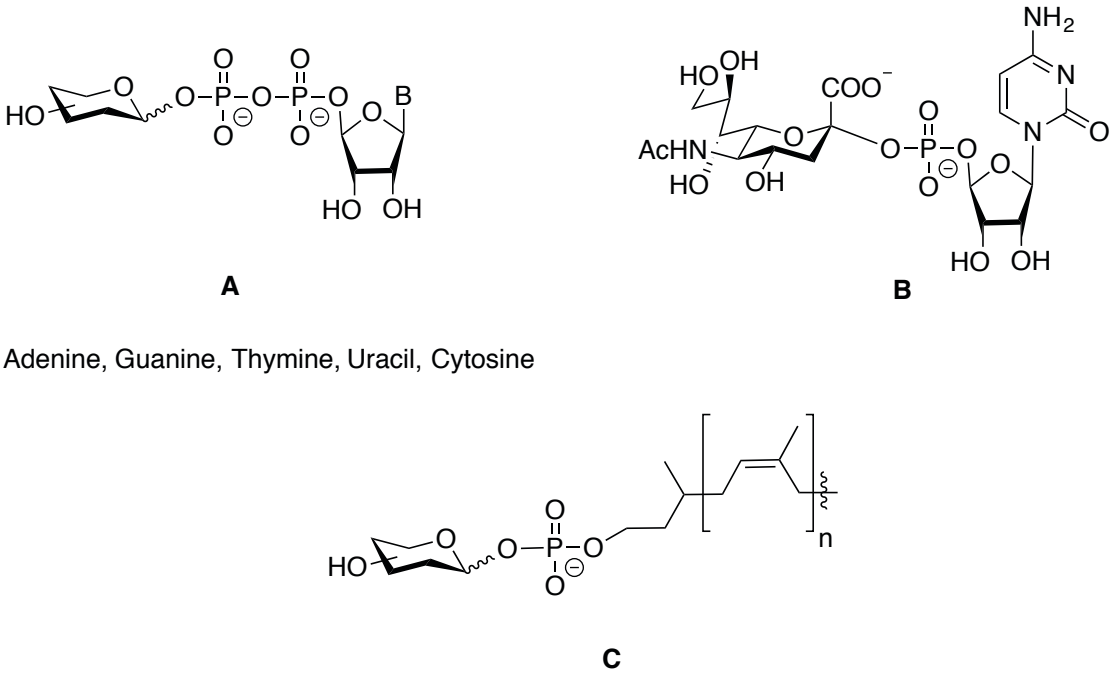


Figure 1-8: Most common sugar phosphate donors in the form of (**A**) nucleoside diphosphates, (**B**) nucleoside monophosphates (cytidine monophosphate-N-acetyl-β-neuraminic acid), and (**C**) lipid phosphates (dolichol phosphate oligosaccharides).

Currently, over 97 families containing over 160 000 individual GTs have been classified in the Carbohydrate-Active Enzyme (CAZy) database (http://www.cazy.org). GTs typically account for 1–2% of gene products of a given organism including eukaryotes, bacteria, and archea, as well as double stranded DNA viruses.[52] Two families, GT2 and GT4, account for over half the total number of GTs contained in the CAZy database. Plants and bacteria have many GTs (e.g., *Arabidopis* encodes ~450 GTs), as a result of their need to synthesize complex cell walls. Of the GTs classified by putative function in the CAZy database, only a small number have been

functionally characterized. This highlights the lack of efficient tools and methods to elucidate GT function.[53] To functionally classify, as well as identify both the donor and acceptor substrates for GTs, remains a challenging problem, and is an area in need of new and innovative systematic strategies.

Though sugars are the most common acceptor substrates used by GTs, lipids, proteins, nucleic acids, antibiotics, and other small nucleophiles also form glycosidic linkages catalyzed this class of enzyme.[52] Moreover, the type of linkages formed by GT's also vary: in addition to forming O-linkages with the hydroxyl group of acceptor substrates, which will be the focus of this thesis, it is also important to note that nitrogen (e.g., *N*-linked glycoproteins), sulfur (e.g., thioglycosides in plants), and carbon (e.g., *C*-glycoside antibiotics) can also act as nucleophiles. In fact, *N*-linked glycosylation is the most ubiquitous protein modification, and affects protein folding, oligomerization, sorting, and transport of secretory and membrane proteins.[54] Not only is it widely spread in eukaryotes, but also in archaeal and eubacterial organisms.[55]

1.7.2 Glycosyltransferase Folds: GT-A, GT-B, and GT-C

Two general folds have been observed for all solved structures of sugar nucleotide-dependent GTs, called GT-A and GT-B. In addition, a GT-C fold has been predicted for transmembrane GTs.[52,56–58] The limited number of folds observed suggests that the GT family of enzymes evolved from a small number of progenitor sequences, unlike the glycosidases, which have a remarkable degree of diversity in overall fold.

The GT-A fold was first described for SpsA in *Bacillus subtilis*.[59] SpsA is an inverting GT whose topology is characterized by two closely abutting β/α/β Rossman domains. These

domains lead to the formation of a central continuous β sheet, which is typical of nucleotide binding proteins, for which a distinct nucleotide binding domain and acceptor binding domains are present (Figure 1-9A).[60] It is important to note that not all enzymes displaying a GT-A fold are GTs, as is the case with the sugar-1-phosphopyrophosphorylase and nucleotidyltransferase superfamilies.[52,61] Though not a conserved motif, most GT-A enzymes posses an Asp-X-Asp (DXD signature), in which the carboxylates of the Asp residues coordinate with a divalent cation such as Mn^{2+}.

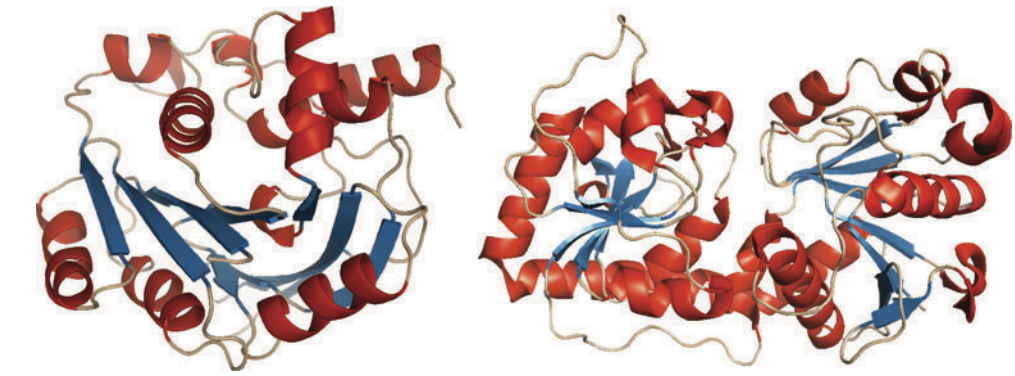


Figure 1-9: Glycosyltransferase folds.[52] (A) GT-A fold for SpsA in *Bacillus subtilis* (B) GT-B fold for bacteriophage T4 β glucosyltransferase. Reprinted with permission form Annual Review of Biochemistry: *Annu. Rev. Biochem.*, *77*, 521-555, copyright 2008.

A DNA-modifying β-glucosyltransferase from bacteriophage T4 was the first 3-D X-ray structure determined for a nucleoside diphosphate-utilizing GT (Figure 1-9B).[62] The overall fold of this structure, homologous to glycogen phosphorylase, was assigned the GT-B fold designation. Similar to the GT-A fold, the architecture of GT-B exhibited two β/α/β Rossman-like domains. However, the later has these domains less tightly associated, with an active site lying in the cleft formed when the β/α/β domains face each other. As is the case with GT-A fold,

the domains of the GT-B fold are associated with acceptor and donor substrate binding sites. Non-glycosyltransferases, such as UDP-GlcNAc 2-epimerase, for example, are also known to adopt the GT-B fold.[63] Unlike the GT-A fold, however, the GT-B lacks the DXD signature, and consequently lacks a divalent cation in the active site.

Iterative sequence searches have predicted a third GT fold, termed GT-C, using programs such as Basic Local Alignment Search Tool (BLAST, http://blast.ncbi.nlm.nih.gov/Blast.cgi).[58] The predicted architecture identifies a large hydrophobic integral membrane protein containing 8–13 transmembrane helices, either as part of the endoplasmic reticulum or plasma membrane, with an active site on the long loop region[52]. To date, 12 GT families have been predicted to adopt the GT-C fold and 10 of these use a lipid phosphate-activated donor. Therefore, they do not display structures that are constrained to bind nucleotides, unlike GT-A and GT-B folds. Most recently, a new structural fold, named GT-D has been reported for the GT DUF1792.[64] This family of GTs, similar to the GT-A and GT-B fold, have been found to bind UDP, and like that of the GT-A, requires a divalent metal ion to transfer the carbohydrate, though its sequence and structural similarity to known GTs remains distinct.[65]

1.7.3 Glycosyltransferases: Classification

The formation of a glycosidic bond will result in one of two stereochemical outcomes: the anomeric configuration about the newly formed bond can either be retained or inverted with respect to the anomeric configuration of the original activated donor substrate (Figure 1-7). Consequently, GTs are classified into inverting or retaining, to reflect the outcome of the

reaction.[52] The stereochemical outcome arises due to the mechanism by which the two classes of enzymes act (Figure 1-7).

GT-A and GT-B folds contain both retaining and inverting GTs, thus the overall fold does not dictate the stereochemical outcome.[56] In contrast, all enzymes predicted to adopt the GT-C fold belong to the inverting GT family. To deconvolute and simplify the identification and classification of GTs, Coutinho *et al.*[56] have proposed a system that takes into account both the fold and activity of the GT, and combines these two criteria to classify GTs into clans (I–IV, Figure 1-10). The family number (according to the CAZy database) is displayed on the far right, with solved 3-D structures annotated in red.

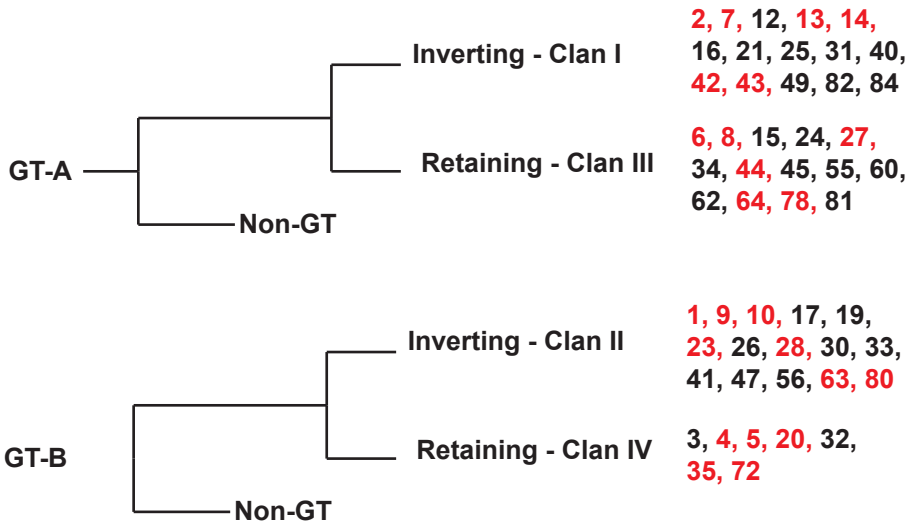


Figure 1-10: Glycosyltransferase classification system: Classification is based, from left to right, on fold (GT-A or GT-B), classification as a GT or non-GT, GT activity (inverting or retaining), clan number (I, II, III, IV), GT family assigned on CAZy database (red = solved 3-D structure). Proposed by Coutinho *et al.*[56]

1.7.4 Glycosyltransferases: Mechanism

Inverting GTs use a direct displacement S_N2-like mechanism (**A**, Figure 1-11). The mechanism by which retaining glycosyltransferases act, however, is less clear, and two models have been proposed: a double displacement mechanism (**B**, Figure 1-11), and a $D_N{*}A_{Nss}$ ion pair mechanism **(C,** Figure 1-11). In contrast to the more well characterized glycosidases,

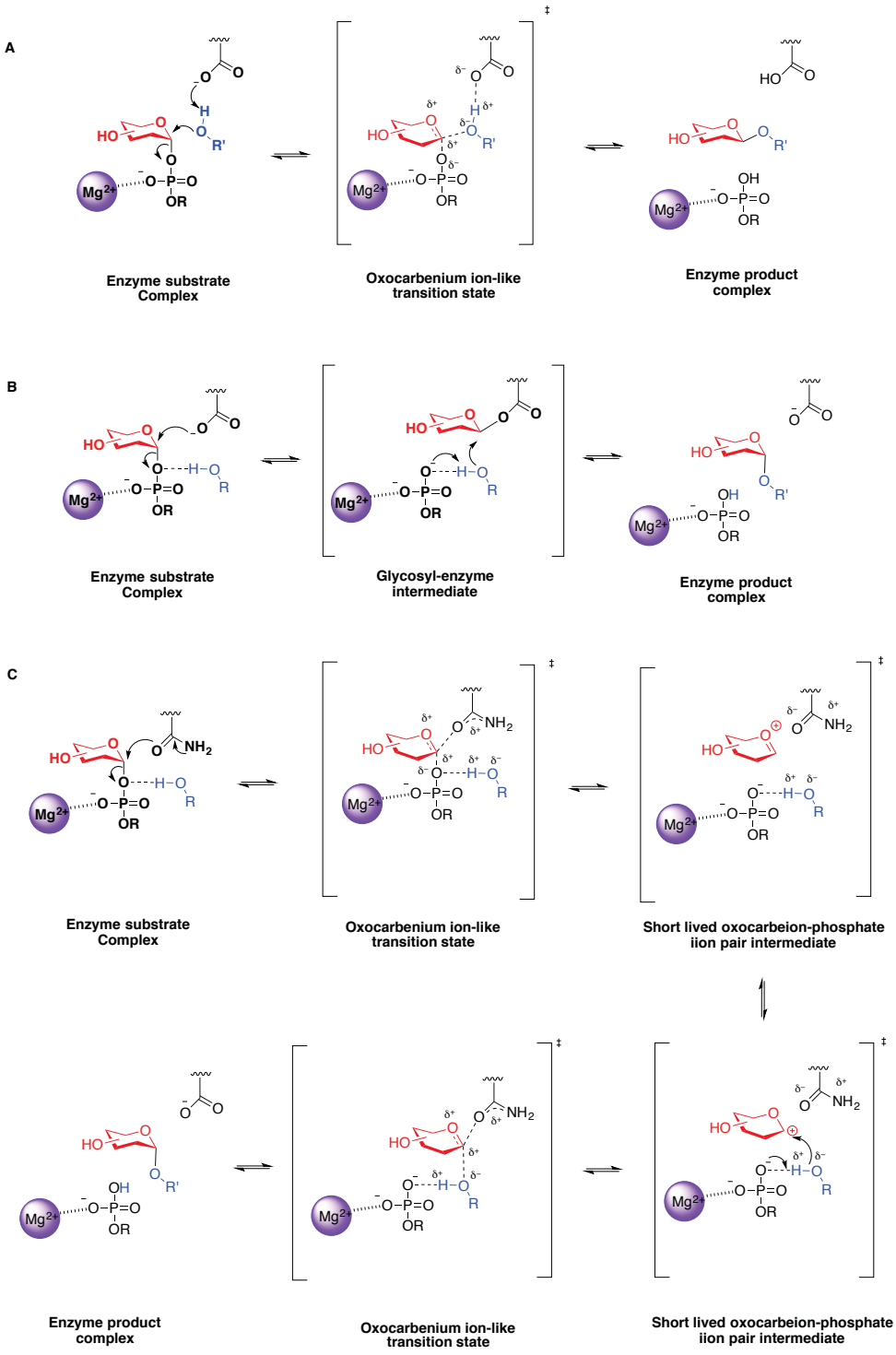


Figure 1-11: (A) Inverting glycosyltransferase direct displacement S_N2-like mechanism, (B) Retaining glycosyltransferase double displacement mechanism, (C) Retaining glycosyltransferase $D_N{*}A_{Nss}$ ion pair mechanism.

mechanistic studies for retaining GTs have failed to provide a clear or definitive answer. Most likely a mechanistic continuum exists in this class of enzyme, involving an ion pair formation with the kinetic properties of a double S_N2 mechanism, with the stereochemical outcome of an S_N1 reaction. This is comparable to the S_Ni mechanism, which is characteristic for the chlorination of alcohols using thionyl chloride, where the carbenium ion forms an intimate ion pair before racemization, thereby maintaining stereochemical integrity (Figure 1-12).

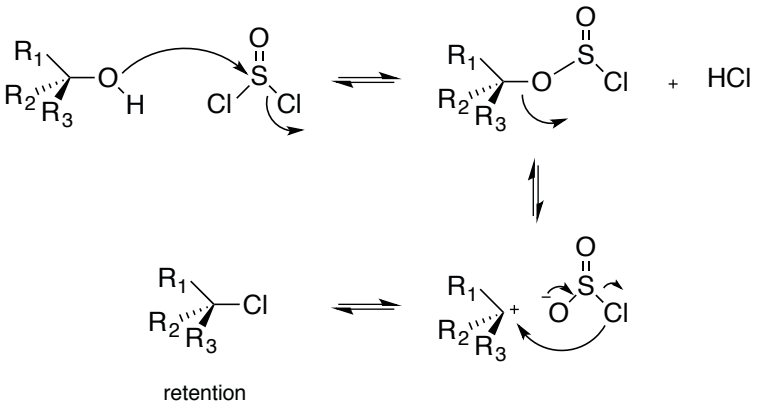


Figure 1-12: Retention in stereochemistry for chlorination of alcohol with thionyl chloride via S_Ni mechanism.

1.8 Genome of *C. jejuni* serotype 11168H

There remain many questions with regard to the mechanism of CPS biosynthesis in *C. jejuni*, and without a proper understanding and identification of the enzymes that catalyze the synthesis of this tetrasaccharide and its polymerization, synthesizing inhibitors with therapeutic potential is difficult. Currently, the GTs responsible for synthesizing the main chain of the CPS repeat unit have not been elucidated. The current thesis will involve addressing some of these deficiencies. In particular, the aim was to carry out investigations to identify the GTs that catalyze the joining

of the carbohydrate residues in the tetrasaccharide repeating unit, as well as probing the specificity and structure of one of these enzymes, particularly the Gal*f*NAc transferase.

The genome of C. *jejuni* serotype 11168 has been sequenced and putative functions have been assigned to many of the genes (Figure 1-13).[66] This work has led to the identification of five possible gene candidates that could encode for the four GTs that are responsible for installing the heptose, glucuronic acid, *N*-acetyl-galactofuranose and ribofuranose residues (Figure 1-1, rings A–D, respectively) in the tetrasaccharide. Note: it is believed that all of the modifications, e.g., addition of methyl and phosphoramidate groups are added after the polysaccharide is assembled.

However, it is difficult to predict the exact function of a putative GT based on sequence homology alone. There are many examples of enzymes with closely related sequences having different catalytic activities.[67] Only the heptosyl transferase, cj1431, responsible for adding the side branch heptose, has been identified through insertional mutagenesis. Mutagenesis of any of the other putative genes—*cj1432, cj1434, cj1438, cj1440,* and *cj1442*—shows a complete failure of CPS synthesis,[32,68] therefore an *in vitro* study is necessary to elucidate the CPS biosynthetic pathway.

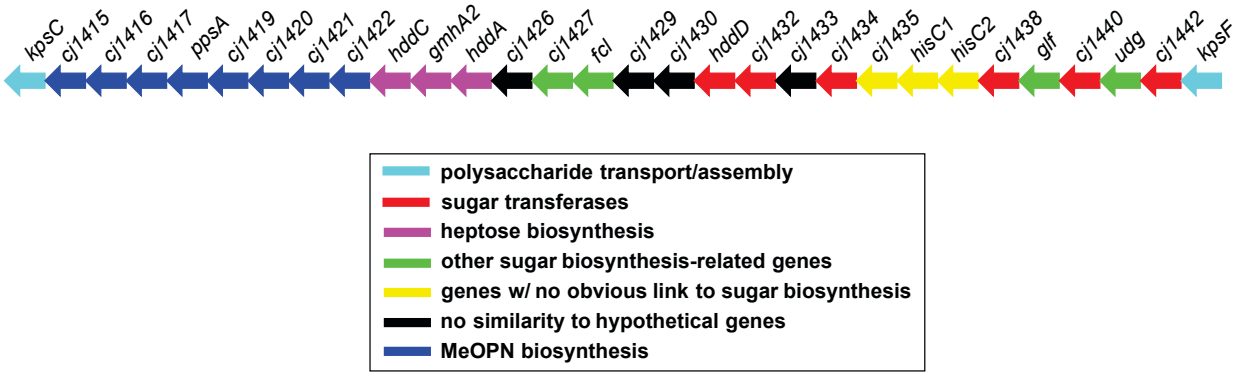


Figure 1-13: Genome of *C. jejuni* serotype 11168H, and assigned putative function.[32]

Gene Knockout	CPS Phenotype	Functional Annotation	Reference
Specific gene function modelled as unknown			
Cj1416	loss of OMePN	OMePN nucleotidyltransferase	McNally *et al.*[45]
Cj1417	loss of OMePN	OMePN biosynthesis	McNally *et al.*[45]
Cj1418	loss of OMePN	OMePN biosynthesis	McNally *et al.*[45]
Cj1432	loss of CPS	sugar transferase	Sternberg *et al.*[68]
Cj1434	loss of CPS	sugar transferase	Sternberg *et al.*[68]
Cj1438	loss of CPS	sugar transferase	Sternberg *et al.*[68]
Cj1440	loss of CPS	sugar transferase	Sternberg *et al.*[68]
Cj1442	loss of CPS	sugar transferase	Sternberg *et al.*[68]
Specific Gene function modelled as known			
cj1421	loss of MeOPN on Gal*f*NAc	MeOPN transferase	McNally *et al.*[45]
cj1422	loss of MeOPN on Hep	MeOPN transferase	McNally *et al.*[45]
cj1423	loss of Hep	Heptosyl transferase	Karlyshev *et al.*[32]
cj1424	loss of Hep	Sedoheptulose isomerase	Karlyshev *et al.*[32]
cj1425	loss of Hep	Heptose kinase	Karlyshev *et al.*[32]
cj1426	loss of 6-OMe on Hep	Methyltransferase	Sternberg *et al.*[68]
cj1427	loss of Hep	GDP-heptose epimerase	Sternberg *et al.*[68]
cj1428	loss of Hep	GDP-heptose epimerase	St. Michael *et al.*[69]
cj1430	loss of Hep	GDP-heptose epimerase	Sternberg *et al.*[68]
cj1431	loss of Hep	GDP-heptosyltransferase	Karlyshev *et al.*[32]
cj1439	loss of CPS	UDP-GalNac pyranose-furonse mutase	Poulin *et al.*[51] *pou*
cj1441	loss of CPS	UDP-Glc 6-dehydrogenase	Sternberg *et al.*[68] St. Michael *et al.*[69]

Table 1-1: *C. jejuni* single-knockout mutants and associated CPS glycan phenotypes.

Based on sequence homology, it is hypothesized that the four glycosyltransferases required for the assembly of the tetrasaccharide repeating unit in the *C. jejuni* 11168 CPS are encoded by one of the following genes: *cj1432*, *cj1434*, *cj1438*, *cj1440*, *cj1442.* More recent *in silico* work by Muggleton and coworkers has proposed the most likely candidate for *N*-acetyl-galactosamine-transferase is *cj1440*, with *cj1442* being a strong alternate possibility.[70] Although the probability of predicting gene function from the models assigned from *in silico* work may be questioned, the plausibility of the final result gives a starting point to aid *in vitro* investigations.

1.9 Research Objectives

The aim of the current thesis is threefold:

1) To clone each of the five putative glycosyltransferase genes into *E. coli* and then express and purify the resulting proteins (*cj1432*, *cj1434*, *cj1438*, *cj1440*, *cj1442*).

2) To synthesize possible substrates for the putative glycosyltransferases encoded by these genes, and then use the molecules in assays with the purified enzymes to establish their function.

3) To carry out more detailed structural and mechanistic studies on the enzyme that adds the *N*-acetyl-galactofuranose (Gal*f*NAc) residue.

The cloning of each of the genes will be carried out using standard approaches. Briefly, each of the five genes described above will be incorporated into the pWM1007 plasmid and then the plasmid DNA will be isolated from XL10-gold cells and sequenced before being transformed

into *E. coli* BL21:DE3 cells for protein expression. Each protein will be expressed with a His_6-tag to facilitate purification by affinity chromatography on a Ni^{2+}-NTA column. We anticipate the major challenge will be preparing the protein mutants as some may be hard to produce or express in soluble form. Should this be a problem, we will look at alternate plasmids and *E. coli* cell types, as well as expression conditions. It is impossible at the outset to determine which, if any, of these proteins will be difficult to express.

Each of the GTs will catalyze the formation of a glycosidic linkage by the reaction of an acceptor molecule with a donor species. It is suspected that the donor species are sugar nucleotides and that small oligosaccharides will serve as the acceptors. Thus, we will synthesize what we consider to be the most likely donor and acceptor substrates for each enzyme, and test each of the proteins for their ability to carry out the reaction with these substrates. This is illustrated below for the enzyme that adds the Gal*f*NAc residue (Figure 1-1, ring C). The sugar nucleotide donor substrate for the GT that introduces this residue is uridine diphospho-*N*-acetyl galactofuranose (UDP-Gal*f*NAc, **1-1**, Figure 1-14). Possible acceptors are either monosaccharides **1-2** or **1-3**, or disaccharides **1-4** or **1-5**. Sugar nucleotide **1-1** will be obtained by an enzyme-catalyzed reaction, carried out in our group, starting from the pyranose form of this sugar nucleotide.[51]

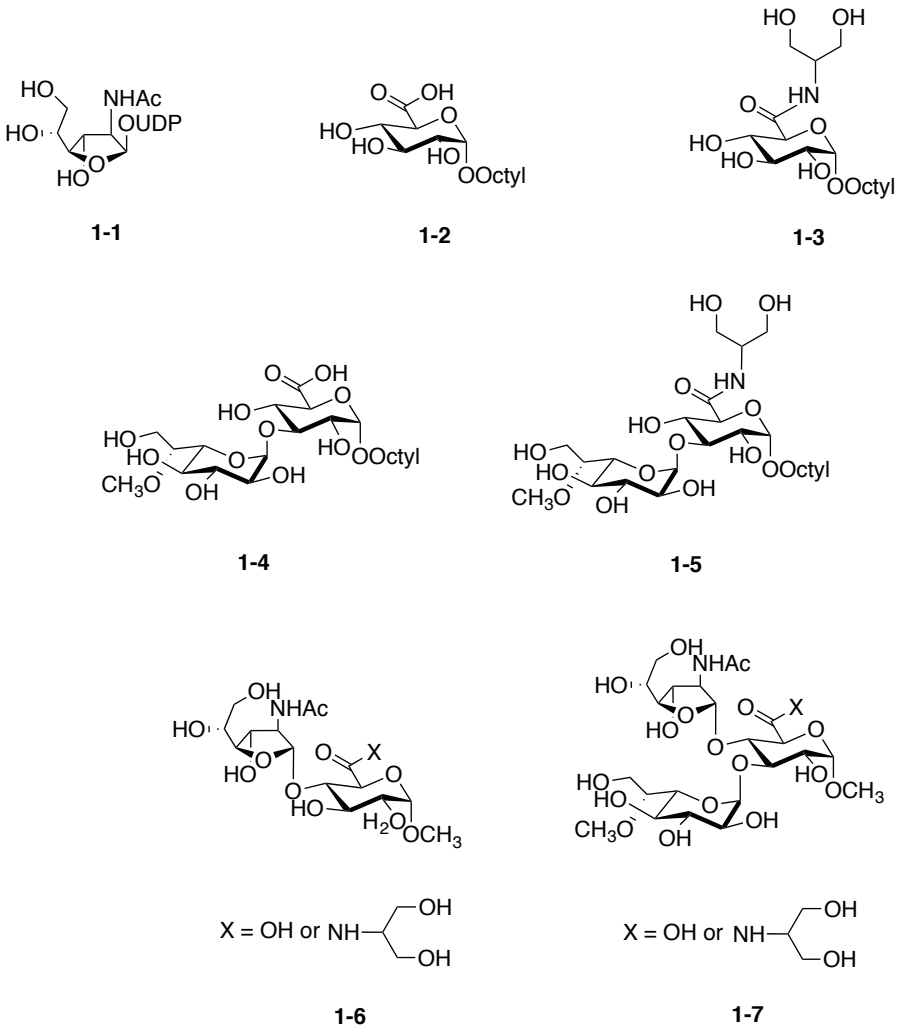


Figure 1-14: (1) Predicted donor substrate for putative GTs, UDP-Gal*f*NAc, **(2-4)** potential acceptor substrates, **(6)** predicted product for positive activity of Gal*f*NAc transferase with **1** and **3**, **(7)** predicted product for positive activity of Gal*f*NAc transferase with **1** and **5**.

Compounds **1-2** to **1-5 (**Figure 1-14**)**, will be prepared using chemical synthesis. As an example, the general approach for the preparation of **1-5** is shown in Figure 1-15. Similar approaches will be used for the preparation of **1-4**.

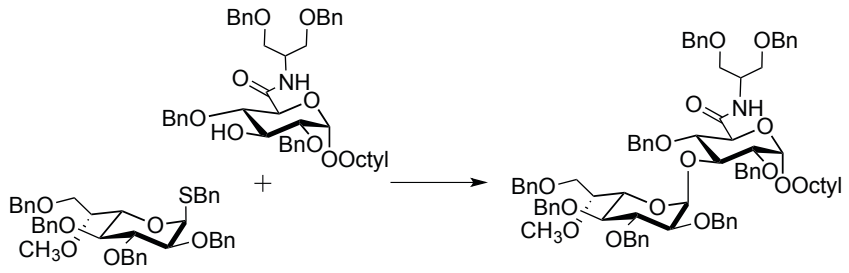


Figure 1-15: General approach to synthesizing disaccharide acceptor.

With these compounds in hand, they will be incubated with each expressed protein using a coupled spectrophotometric GT assay, which we have adapted and implemented for galactofuranosyltransferases (Figure 1-16).[71,72]

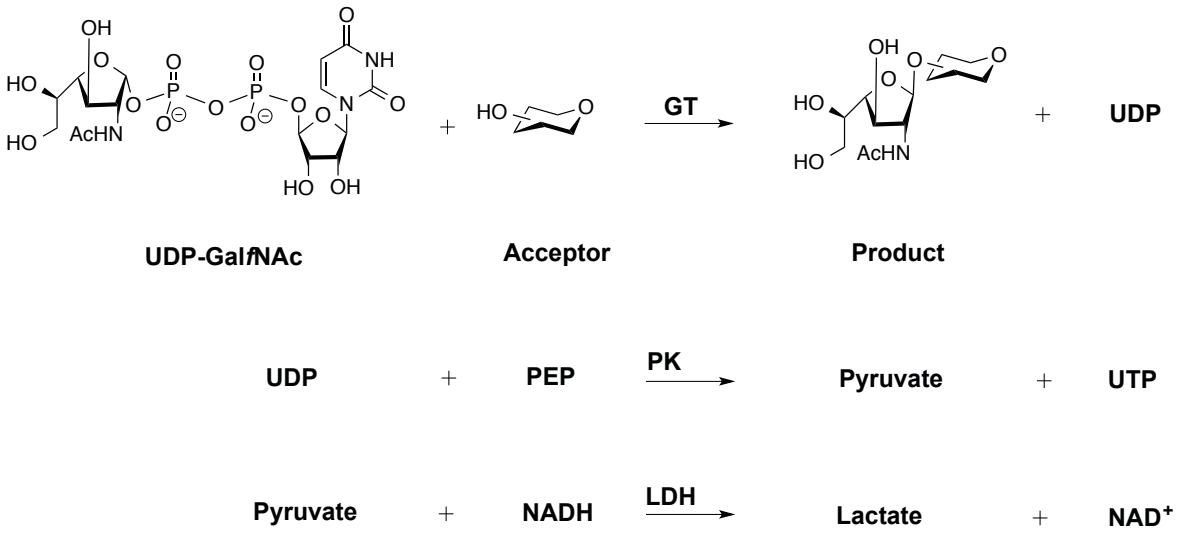


Figure 1-16: Glycosyltransferase Continuous Spectrophotometric Assay

Briefly, this assay is dependent upon transfer of the Gal*f*NAc onto an acceptor substrate. The UDP produced in the reaction is coupled to the oxidation of NADH by way of two enzymes, pyruvate kinase (PK), and lactate dehydrogenase (LDH). The decrease in absorption at 340 nm, resulting from NADH oxidation, will be proportional the activity of the GT.

The protein that shows activity with these substrates in this assay will correspond to the GT that installs the Gal*f*NAc residue. Further proof of enzymatic function will be achieved by performing larger scale enzymatic reactions, purifying the products and their characterization by mass spectrometry and NMR spectroscopy. As standards, authentic samples of the reactions products (e.g., disaccharide **1-6** or trisaccharide **1-7**, Figure 1-14) will also be synthesized for purposes of comparison. The same approach can be applied to future work in identifying the enzymes responsible for the formation of the other three glycosidic linkages, using different substrates.

These studies will represent the first detailed investigations of GTs involved in the assembly of campylobacter CPS. In addition, the carbohydrate residues in this polysaccharide are unusual, and so it is possible that new protein recognition motifs, involved in transferring these sugars, may be identified.

1.10 References

(1) Guerry, P.; Szymanski, C. M. *Trends Microbiol.* **2008**, *16* (9), 428–435.

(2) Jacobs, B. C.; Rothbarth, P. H.; van der Meché, F. G.; Herbrink, P.; Schmitz, P. I.; de Klerk, M. A.; van Doorn, P. A. *Neurology* **1998**, *51* (4), 1110–1115.

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

(4) Monteiro, M. a.; Baqar, S.; Hall, E. R.; Chen, Y. H.; Porter, C. K.; Bentzel, D. E.; Applebee, L.; Guerry, P. *Infect. Immun.* **2009**, *77* (3), 1128–1136.

(5) Coward, C.; Grant, A. J.; Swift, C.; Philp, J.; Towler, R.; Heydarian, M.; Frost, J. A.; Maskell, D. J. *Appl. Environ. Microbiol.* **2006**, *72* (7), 4638–4647.

(6) Bolton, D. J. *Food Microbiol.* **2015**, *48*, 99–108.

(7) Fouts, D. E.; Mongodin, E. F.; Mandrell, R. E.; Miller, W. G.; Rasko, D. A.; Ravel, J.; Brinkac, L. M.; DeBoy, R. T.; Parker, C. T.; Daugherty, S. C.; Dodson, R. J.; Durkin, A. S.; Madupu, R.; Sullivan, S. A.; Shetty, J. U.; Ayodeji, M. A.; Shvartsbeyn, A.; Schatz, M. C.; Badger, J. H.; Fraser, C. M.; Nelson, K. E. *PLoS Biol* **2005**, *3* (1), e15.

(8) Poly, F.; Read, T. D.; Chen, Y. H.; Monteiro, M. a.; Serichantalergs, O.; Pootong, P.; Bodhidatta, L.; Mason, C. J.; Rockabrand, D.; Baqar, S.; Porter, C. K.; Tribble, D.; Darsley, M.; Guerry, P. *Infect. Immun.* **2008**, *76* (12), 5655–5667.

(9) Lindow, J. C.; Poly, F.; Tribble, D. R.; Guerry, P.; Carmolli, M. P.; Baqar, S.; Porter, C. K.; Pierce, K. K.; Darsley, M. J.; Sadigh, K. S.; Dill, E. A.; Team, the C. S.; Kirkpatrick, B. D. *J. Clin. Microbiol.* **2010**, *48* (8), 3012–3015.

(10) Kittl, S.; Kuhnert, P.; Hächler, H.; Korczak, B. M. *J. Appl. Microbiol.* **2011**, *110* (2), 513–520.

(11) Zhao, S.; Young, S. R.; Tong, E.; Abbott, J. W.; Womack, N.; Friedman, S. L.; McDermott, P. F. *Appl. Environ. Microbiol.* **2010**, *76* (24), 7949–7956.

(12) Luangtongkum, T.; Jeon, B.; Han, J.; Plummer, P.; Logue, C. M.; Zhang, Q. *Future Microbiol.* **2009**, *4* (2), 189–200.

(13) Peterson, J. W. University of Texas Medical Branch at Galveston 1996.

(14) Willis, L. M.; Whitfield, C. *Carbohydr. Res.* **2013**, *378*, 35–44.

(15) Whitfield, C. *Annu. Rev. Biochem.* **2006**, *75*, 39–68.

(16) Pelkonen, S.; Hayrinen, J.; Finne, J. *J. Bacteriol.* **1988**, *170* (6), 2646–2653.

(17) Karlyshev, A. V; McCrossan, M. V; Wren, B. W. *Infect. Immun.* **2001**, *69* (9), 5921–5924.

(18) Reckseidler-zenteno, S. L. **2012**.

(19) Jiménez, N.; Senchenkova, S. N.; Knirel, Y. A.; Pieretti, G.; Corsaro, M. M.; Aquilini, E.; Regué, M.; Merino, S.; Tomás, J. M. *J. Bacteriol.* **2012**, *194* (13), 3356–3367.

(20) Ophir, T.; Gutnick, D. L. *Appl. Environ. Microbiol.* **1994**, *60* (2), 740–745.

(21) Roberts, I. S. *Annu. Rev. Microbiol.* **1996**, *50*, 285–315.

(22) Costerton, J. W.; Cheng, K. J.; Geesey, G. G.; Ladd, T. I.; Nickel, J. C.; Dasgupta, M.; Marrie, T. J. *Annu. Rev. Microbiol.* **1987**, *41* (1), 435–464.

(23) Jenkinson, H. F. *Trends Microbiol.* **1994**, *2* (6), 209–212.

(24) Kolenbrander, P. E. *J. Appl. Bacteriol.* **1993**, *74*, 79S – 86S.

(25) Gygi, D.; Rahman, M. M.; Lai, H.-C.; Carlson, R.; Guard-Petter, J.; Hughes, C. *Mol. Microbiol.* **1995**, *17* (6), 1167–1175.

(26) Horwitz, M. A.; Silverstein, S. C. *J. Clin. Invest.* **1980**, *65* (1), 82–94.

(27) Howard, C. J.; Glynn, A. A. *Immunology* **1971**, *20* (5), 767–777.

(28) Moxon, E.; Kroll, J. *Curr Top Microbiol Immunol.* **1990**, *150*, 65–85.

(29) Kim, K. S.; Kang, J. H.; Cross, A. S. *FEMS Microbiol. Lett.* **1986**, *35* (2-3), 275–278.

(30) Cross, A. *Curr Top Microbiol Immunol.* **1990**, *150*, 87–95.

(31) Weintraub, A. *Carbohydr. Res.* **2003**, *338* (23), 2539–2547.

(32) Karlyshev, A. V.; Linton, D.; Gregson, N. a.; Lastovica, A. J.; Wren, B. W. *Mol. Microbiol.* **2000**, *35* (3), 529–541.

(33) Logan, S. M.; Trust, T. J. *Infect. Immun.* **1984**, *45* (1), 210–216.

(34) Moran, A. P.; Penner, J. L. *J. Appl. Microbiol.* **1999**, *86* (3), 361–377.

(35) Mandatori, R.; Penner, J. L. *Infect. Immun.* **1989**, *57* (11), 3506–3511.

(36) Willis, L. M.; Whitfield, C. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (51), 20753–20758.

(37) Corcoran, A. T.; Annuk, H.; Moran, A. P. *FEMS Microbiol. Lett.* **2006**, *257* (2), 228–235.

(38) Karlyshev, A. V; Wren, B. W. *J. Clin. Microbiol.* **2001**, *39* (1), 279–284.

(39) 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* (21), 5119–5136.

(40) Szymanski, C. M.; Michael, F. St.; Jarrell, H. C.; Li, J.; Gilbert, M.; Larocque, S.; Vinogradov, E.; Brisson, J.-R. *J. Biol. Chem.* **2003**, *278* (27), 24509–24520.

(41) McNally, D. J.; Jarrell, H. C.; Li, J.; Khieu, N. H.; Vinogradov, E.; Szymanski, C. M.; Brisson, J. R. *FEBS J.* **2005**, *272* (17), 4407–4422.

(42) 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* (17), 3975–3989.

(43) Kanipes, M. I.; Tan, X.; Akelaitis, A.; Li, J.; Rockabrand, D.; Guerry, P.; Monteiro, M. A. *J. Bacteriol.* **2008**, *190* (5), 1568–1574.

(44) Chen, Y. H.; Poly, F.; Pakulski, Z.; Guerry, P.; Monteiro, M. a. *Carbohydr. Res.* **2008**, *343* (6), 1034–1040.

(45) 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* (39), 28566–28576.

(46) Kanipes, M. I.; Papp-Szabo, E.; Guerry, P.; Monteiro, M. A. *J. Bacteriol.* **2006**, *188* (9), 3273–3279.

(47) Peltier, P.; Euzen, R.; Daniellou, R.; Nugier-Chauvin, C.; Ferrières, V. *Carbohydr. Res.* **2008**, *343* (12), 1897–1923.

(48) Richards, M. R.; Lowary, T. L. *ChemBioChem* **2009**, *10* (12), 1920–1938.

(49) Lee, R. E.; Smith, M. D.; Nash, R. J.; Griffiths, R. C.; McNeil, M.; Grewal, R. K.; Yan, W.; Besra, G. S.; Brennan, P. J.; Fleet, G. W. J. *Tetrahedron Lett.* **1997**, *38* (38), 6733–6736.

(50) Pan, F.; Jackson, M.; Ma, Y.; McNeil, M. *J. Bacteriol.* **2001**, *183* (13), 3991–3998.

(51) Poulin, M. B.; Nothaft, H.; Hug, I.; Feldman, M. F.; Szymanski, C. M.; Lowary, T. L. *J. Biol. Chem.* **2010**, *285* (1), 493–501.

(52) Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. *Annu. Rev. Biochem.* **2008**, *77* (1), 521–555.

(53) Ban, L.; Pettit, N.; Li, L.; Stuparu, A. D.; Cai, L.; Chen, W.; Guan, W.; Han, W.; Wang, P. G.; Mrksich, M. *Nat Chem Biol* **2012**, *8* (9), 769–773.

(54) Burda, P.; Aebi, M. *Biochim. Biophys. Acta - Gen. Subj.* **1999**, *1426* (2), 239–257.

(55) Szymanski, C. M.; Logan, S. M.; Linton, D.; Wren, B. W. *Trends Microbiol.* **2003**, *11* (5), 233–238.

(56) Coutinho, P. M.; Deleury, E.; Davies, G. J.; Henrissat, B. *J. Mol. Biol.* **2003**, *328* (2), 307–317.

(57) Ünligil, U. M.; Rini, J. M. *Curr. Opin. Struct. Biol.* **2000**, *10* (5), 510–517.

(58) Ghafoor, A.; Jordens, Z.; Rehm, B. H. A. *Appl. Environ. Microbiol.* **2013**, *79* (9), 2968–2978.

(59) Charnock, S. J.; Davies, G. J. *Biochemistry* **1999**, *38* (20), 6380–6385.

(60) Keenleyside, W. J.; Clarke, A. J.; Whitfield, C. *J. Bacteriol.* **2001**, *183* (1), 77–85.

(61) Brown, K.; Pompeo, F.; Dixon, S.; Mengin-Lecreulx, D.; Cambillau, C.; Bourne, Y. *EMBO J.* **1999**, *18* (15), 4096–4107.

(62) Vrielink, A.; Rüger, W.; Driessen, H. P.; Freemont, P. S. *EMBO J.* **1994**, *13* (15), 3413–3422.

(63) Campbell, R. E.; Mosimann, S. C.; Tanner, M. E.; Strynadka, N. C. J. *Biochemistry* **2000**, *39* (49), 14993–15001.

(64) Zhu, F.; Zhang, H.; Wu, H. *J. Dent. Res.* **2015**, *94* (5), 659–665.

(65) Zhang, H.; Zhu, F.; Yang, T.; Ding, L.; Zhou, M.; Li, J.; Haslam, S. M.; Dell, A.; Erlandsen, H.; Wu, H. *Nat Commun* **2014**, *5*.

(66) 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* (1), 90–103.

(67) Goubet, F.; Ström, A.; Quéméner, B.; Stephens, E.; Williams, M. A. K.; Dupree, P. *Glycobiol.* **2006**, *16* (1), 29–35.

(68) 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* (1), 186–197.

(69) Michael, F. St.; 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* (21), 5119–5136.

(70) 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* (1), 186–197.

(71) Gosselin, S.; Alhussaini, M.; Streiff, M. B.; Takabayashi, K.; Palcic, M. M. *Anal. Biochem.* **1994**, *220* (1), 92–97.

(72) Rose, N. L.; Zheng, R. B.; Pearcey, J.; Zhou, R.; Completo, G. C.; Lowary, T. L. *Carbohydr. Res.* **2008**, *343* (12), 2130–2139.

2 Synthetic routes toward 6-*O*-methyl-D-*glycero*-L-*gluco*-heptose

2.1 Introduction

Heptose sugars are present in a number of bacterial species, particularly in the inner core structures of lipopolysaccharides from Gram-negative bacteria (Figure 1-3, labelled in black). These heptoses display various stereochemical configurations. The most common of these are the D-*glycero*-D-*manno*-heptopyranosides, which are common constituents of the inner core of Gram-negative bacteria such as *Salmonella, Escherichia coli, Shigella, Klebsiella, Pseudomonas, Neisseria, Vibrio, Bordetella, Yersinia,* and *Campylobacter*.[1] There are also rarer forms found in the bacterial surface polysaccharides of *Campylobacter, Yersinia, Burkholderia, Eubacterium, Plesiomonas,* and *Escherichia*, in both the pyranose and furanose form.[2]

The first naturally occurring glycoconjugate reported to contain a heptose with the D-*glycero*-L-*gluco*- stereochemistry was the CPS from *C. jejuni* NCTC11168 (HS:2) (Figure 2-1, ring D). It is also the only known example of a naturally occurring sugar in the L-*gluco* configuration.[3]

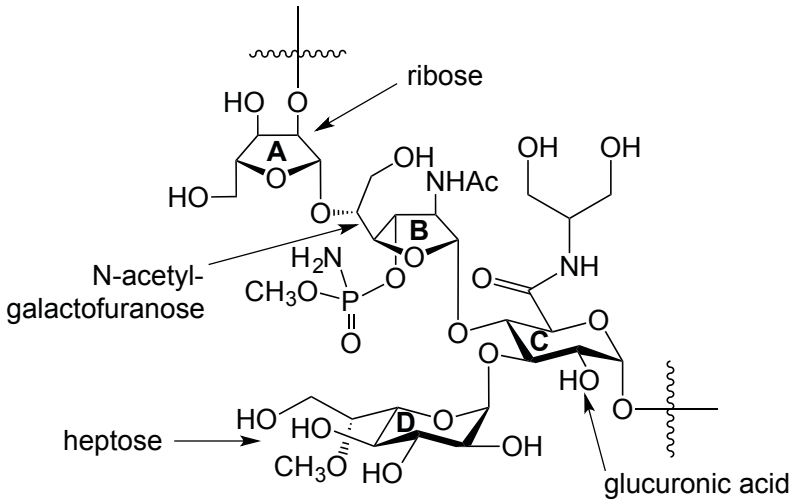


Figure 2-1: Structure of *C. jejuni* NCTC 11168 (HS:2) CPS repeating unit. **(A)** β-D-ribofuranose, **(B)** 2-acetamido-2-deoxy-β-D-galactofuranose, (**C**) α-D-glucopyranosiduron-(2-amino-2-deoxyglycerol)-amide, **(D)** 6-*O*-methyl-D-*glycero*-α-L-*gluco*-heptopyranose.

Of particular interest to this thesis is the chemical synthesis of the heptose moiety in *C. jejuni* 11168 (HS:2), as a glycoconjugate—for the generation of antibodies specific to this novel motif—as well as part of potential substrates to probe the various biosynthetic GTs involved in assembling the full-length glycan (**1-4** and **1-5**, Figure 2-2).

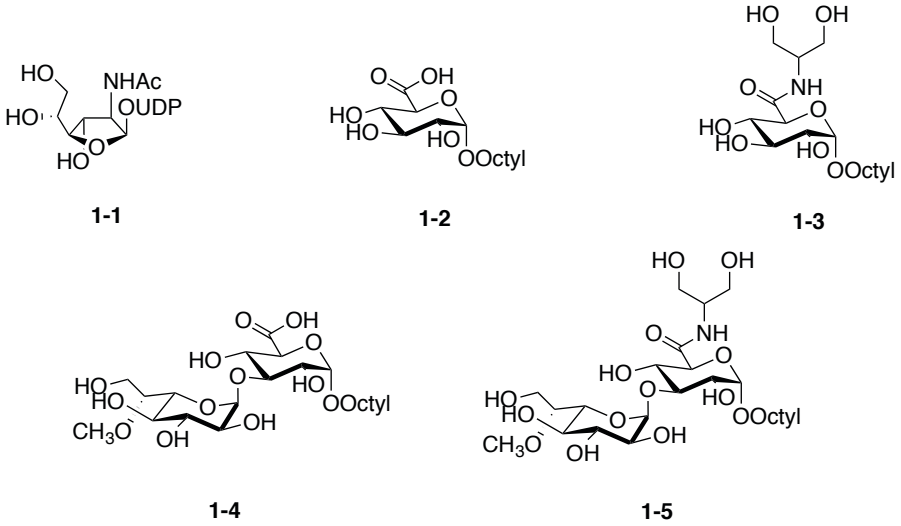


Figure 2-2: Potential donor and acceptor substrates for putative Gal*f*NAc transferases.

Heptoses are not widely available from natural sources; therefore, their chemical synthesis has been the topic of a number of investigations.[1,4–6] The majority of heptose syntheses involve chain extension from C-6 of a hexose. This is typically achieved by oxidizing the C-6 hydroxyl group to an aldehyde, and chain extension via Grignard addition or Wittig olefination reactions. Subsequent functionalization of this homologated product provides the target molecule (Figure 2-3, **A**).[7]

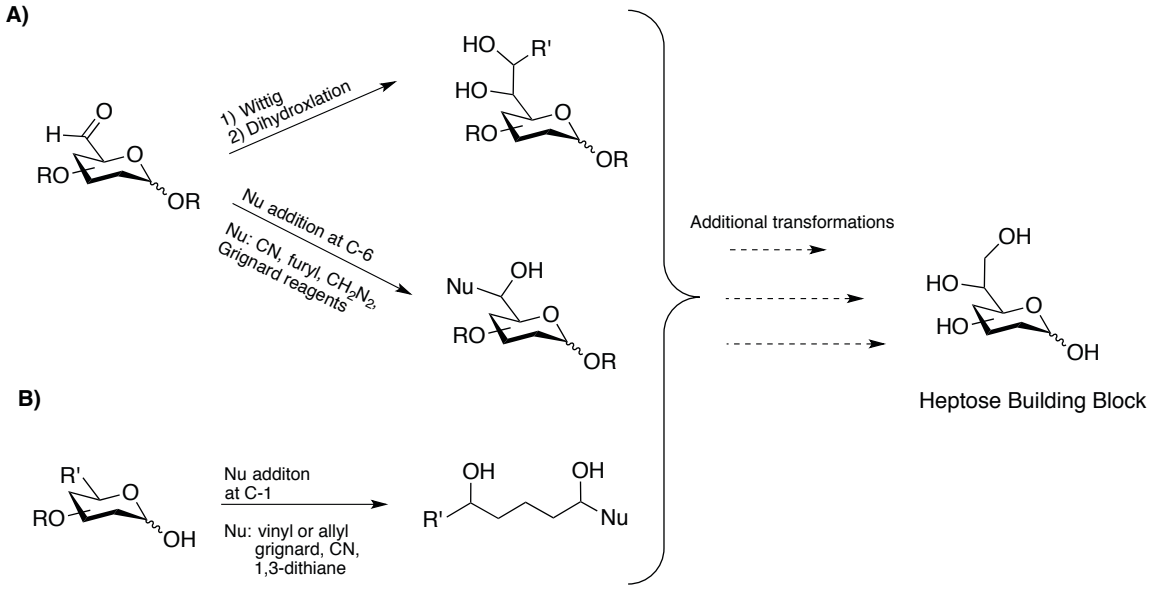


Figure 2-3: Synthetic strategies toward constructing heptoses.[7]

In the case of D-*glycero*-L-*gluco*-heptose, this approach would require access to significant quantities of L-glucose, which is not readily available. Therefore, a less widely used strategy was investigated. This approach involved chain extension from the C-1 aldehyde functionality of a hexose (Figure 2-3, **B**).

Accessing heptoses by this methodology is not a completely novel idea, and has been reported as far back as investigations by Fischer and Kilani on the synthesis of cyanohydrins.[8] This approach has not seen wide application in modern organic chemistry due in part to generally low stereoselectivities.[1]

Mindful of this, we envisioned that the heptose backbone, **2-2**, could be synthesized from galactofuranosyl thioglycoside **2-6** (Figure 2-4). The two key transformations in this route involved a stereoselective C-1 elongation to give **2-4**, followed by ozonolysis mediated ring closure, to provide the D-*glycero*-L-*gluco*-heptose backbone, **2-3**.

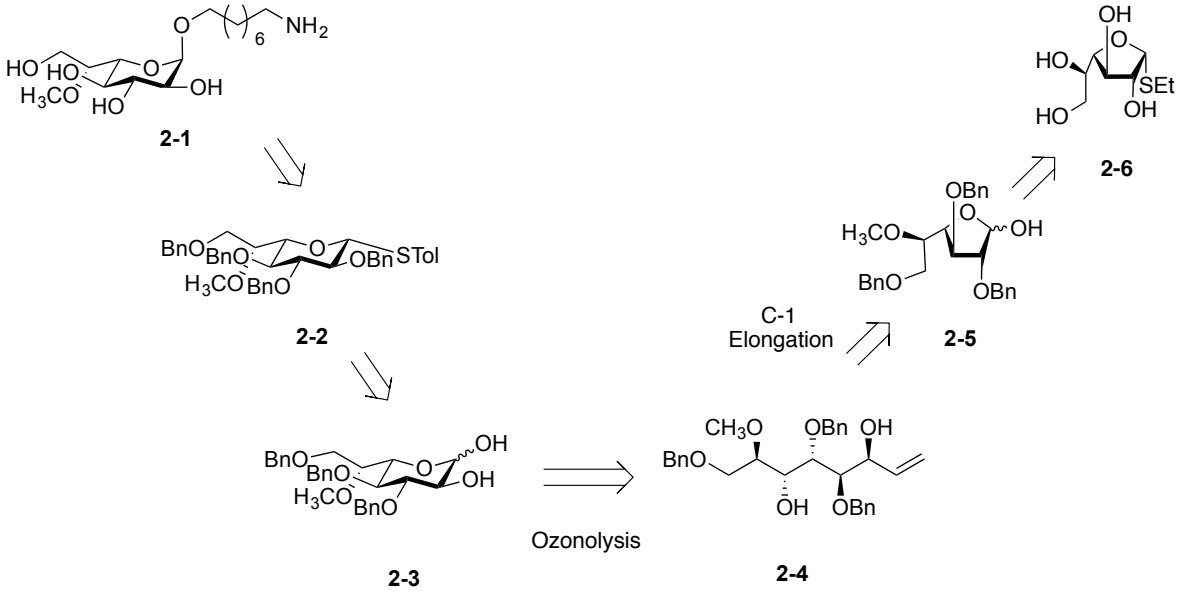


Figure 2-4: Retrosynthetic analysis for heptose.

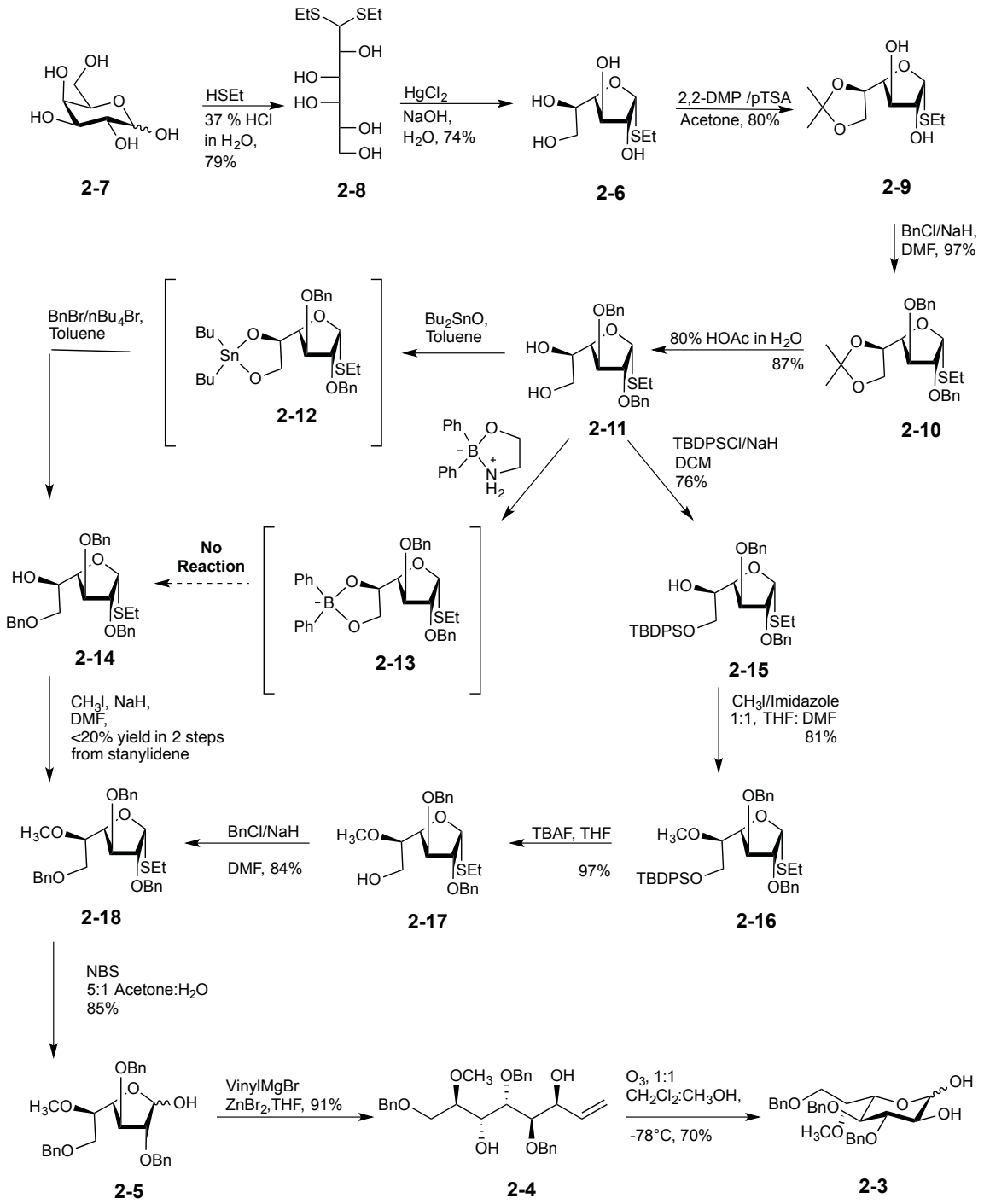
Having this backbone allowed access to monosaccharide donor **2-2**, which had not been synthesized previously. Thioglycoside **2-2** can be used for the synthesis of acceptors **1-4** and **1-5** (Figure 2-2), as well as for making a heptose glycoconjugate. In regard to the latter, **2-2** was

functionalized with an aminooctyl spacer arm **2-1** (Figure 2-4), and its attachment to a protein carrier (bovine serum albumin, BSA) yielded glyconconjugate **2-30** for the generation of antibodies (see below).

2.2 Synthesis of D-*glycero*-L-*gluco*-heptose backbone (2-3)

The synthetic route towards the D-*glycero*-L-*gluco*-heptose backbone, **2-3**, as illustrated in Scheme 2-1, requires chain extension from a galactofuranosyl thioglycoside. In nature, Gal*f*-residues are biosynthesized at the sugar nucleotide level by an enzymatic ring contraction of uridine diphospho-galactopyranose (UDP-Gal*p*) to UDP-Gal*f*.[9] A number of synthetic methods have been developed for accessing the thermodynamically less stable Gal*f* ring form of galactose, which exists predominantly in the pyranose ring form. These include kinetically controlled Fischer glycosylations,[10,11] high temperature perbenzoylation of galactose in pyridine,[12] reduction of D-galactonolactones with borane reagents,[13] reduction of furanoside derivatives of D-galacturonic acid with sodium borohydride and iodine,[14] as well as electrophile induced cyclization of dithioacetals or *O*,*S*-mixed acetals.[15–18]

The first key intermediate, thioglycoside **2-6**, was thus obtained from D-galactose (**2-7**) via a two-step route involving formation of the corresponding diethyl dithioacetal (**2-8**) and cyclization promoted by mercuric chloride, to afford the target in 59% yield over two steps. [16,19] This method was originally developed in 1937 by Pacsu and Green[16] and then later refined by Wolfrom and coworkers.[20]



Scheme 2-1: Construction of D-*glycero*-L-*gluco*-heptose backbone.

After obtaining **2-6**, it was then treated with 2,2-dimethoxypropane (DMP) and *p*-toluenesulfonic acid to afford the isopropylidene acetal **2-9**. This reaction afforded the expected product in 80%

yield. Benzylation of the two hydroxyl groups using standard conditions (giving **2-10**), followed by hydrolysis of the acetal provided **2-11** in 84% yield from **2-9**. The resulting diol was then protected at the primary position as a *t*-butyldiphenylsilyl (TBDPS) ether affording **2-15** in 76% yield.

With a route to secondary alcohol **2-15** in place, treatment with methyl iodide and sodium hydride resulted in introduction of the methyl group to afford **2-16**. The silyl ether was replaced with a benzyl ether via a process involving desilylation with tetrabutyl- ammonium fluoride (TBAF) and alkylation using benzyl chloride/sodium hydride. The product, **2-18**, was obtained in 66% yield over the three steps from **2-15**. Attempts to prepare **2-18** in two steps from diol **2-11**—by selective O-6 benzylation via formation of a stannylene acetal (**2-12**)[21], followed by methylation—led to a lower overall yield (20% over two steps). Selective O-6 benzylation of **2-11** was also attempted by forming a tetracoordinated borinate complex (**2-13**),[22] however no product was detected by thin-layer chromatography (TLC).

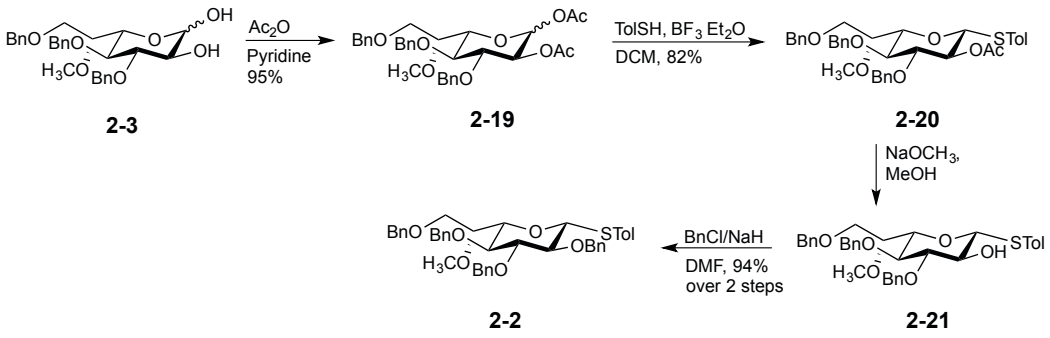
Hydrolysis of thioglycoside **2-18** was achieved by reaction with *N*-bromosuccinimide (NBS) and water in acetone affording hemiacetal **2-5** in 85% yield. The key step in the route was the chain extension of **2-5**, which was successfully achieved upon reaction with divinyl zinc, generated in situ from vinylmagnesium bromide with $ZnBr_2$. The divinyl zinc addition led to the formation of **2-4** in 91% yield with excellent stereoselectivity due to chelation control; none of the other isomer could be detected.

To explore the importance of the $ZnBr_2$ to the stereocontrol of the reaction, **2-4** was also treated with vinyl magnesium bromide. This reaction provided the product with poor selectivity and reduced yield. The structure of **2-4** was corroborated by ^{1}H NMR (Nuclear Magnetic Resonance) spectroscopy of a subsequently synthesized cyclic intermediate and the preparation of a crystalline derivative, by former lab member Dr. Akihiro Imamura.[23] Ozonolysis of the alkene in **2-4** afforded the heptose backbone, **2-3,** in 70% yield.

2.3 Completion of Heptose Thioglycoside Donor

Having the desired D-*glycero*-L-*gluco*-heptose backbone, our attention turned to completing the synthesis of the corresponding heptose thioglycoside donor, as a functional handle for further modification at the anomeric center.

Acetylation of the hydroxyl groups in **2-3**, followed by treatment with *p*-toluenethiol and $BF_3 \cdot OEt_2$, yielded thioglycoside **2-20** in 78% yield over two steps. The β isomer of **2-20** could be separated and the stereochemistry was confirmed by ^{1}H NMR spectroscopy; the $^3J_{1,2}$ was 10.1 Hz, indicative of a *trans* relationship between H-1 and H-2. The acetate ester on O-2 was converted into a nonparticipating benzyl group because the final target, **2-1**, possesses an axially oriented aglycone. This was achieved in two steps by deacetylation (yielding **2-21**) and benzylation resulting in a 94% yield of **2-2**.



Scheme 2-2: Completion of heptose thioglycoside donor synthesis.

Thioglycoside **2-2** was expected to be a suitable donor for the preparation of glycoconjugates containing this unusual heptose. It should be noted that **2-2** was obtained from **2-6** in a comparable number of steps to that reported for the *de novo* synthesis of heptose glycosyl donors.[4] As mentioned previously, support for the stereochemistry of the heptose ring substituents was obtained by the preparation of a crystalline derivative of **2-2** by Dr. Akihiro Imamura (Figure 2-5).[23] This was achieved by first treatment of **2-2β** with hydrogen over a palladium on carbon in ethanol. This reaction afforded a product in which not only had the four benzyl ethers been cleaved, but the thioglycoside moiety had been replaced by an ethoxy group, from the ethanol used as the reaction solvent (Figure 2-5). Subsequent treatment of this product with *p*-nitrobenzoyl chloride in pyridine gave the **2-23** in 37% yield over the two steps (Figure 2-5). The product could be recrystallized from ethyl acetate–hexane and X-ray analysis of the resulting crystals confirmed the D-*glycero*-L-*gluco*- stereochemistry.[23]

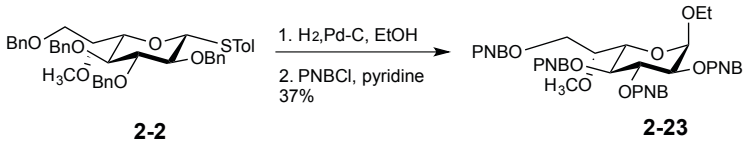
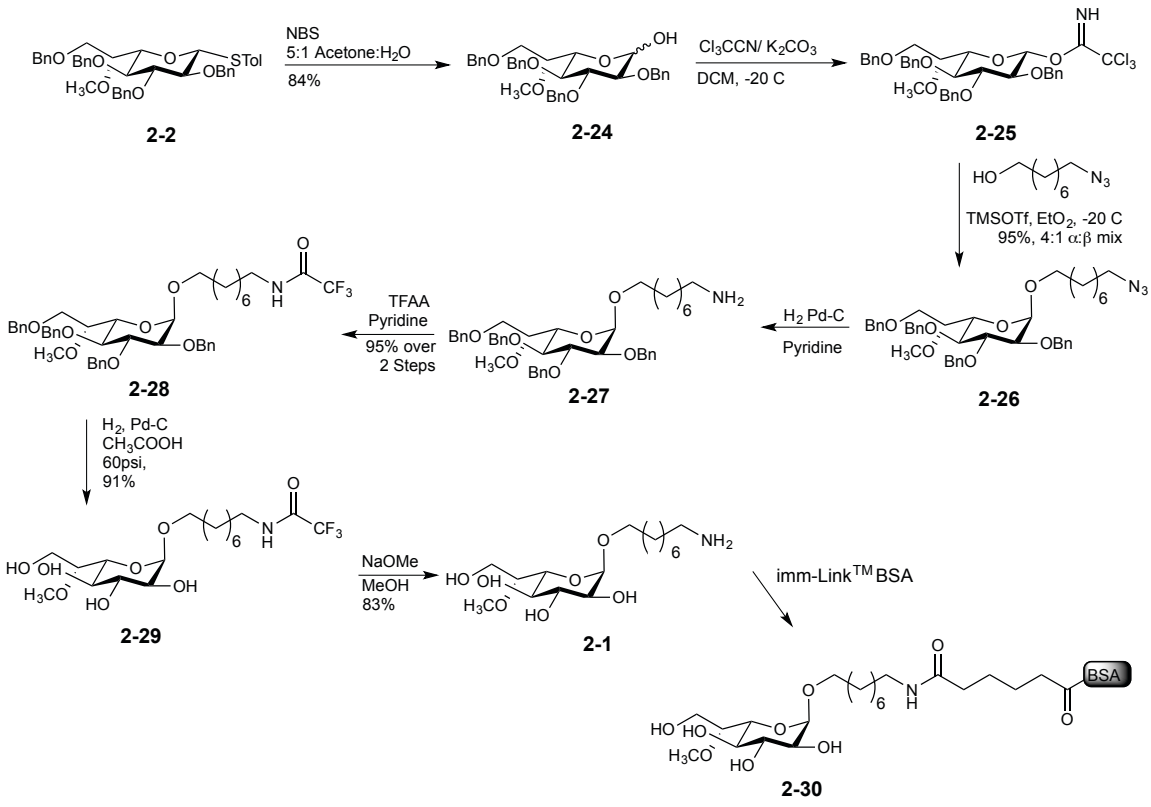


Figure 2-5: Preparation of crystalline intermediate for X-ray analysis

2.4 Attachment of Aminooctyl Linker for Preparation of BSA Glycoconjugate

Having unequivocally established the stereochemistry of **2-2**, the glycosylation of the thioglycoside donor was attempted with 8-azidooctanol in the presence of *N*-iodosuccinimide (NIS) and silver triflate.[24] Unexpectedly, this reaction led to a mixture of products in which the undesired β-(1,2-*trans)*-glycoside predominated. This problem could be circumvented (Scheme 2-3) by conversion of **2-2** into the corresponding reducing sugar **2-24** and, in turn, the β trichloroacetimidate **2-25** by using a mild base (K_2CO_3).

It has been postulated that the β imidate is favoured by using bases such as NaH and K_2CO_3 due to kinetic factors, which favour the reaction of the more reactive β oxyanion. In contrast, the use of $CsCO_3$ is proposed to increase the nucleophilicity of the α oxanion, leading to the α imidate formation by thermodynamic control.[25,26] Therefore, the β imidate was formed under kinetic conditions (–20 °C, <30 min reaction time) to influence stereoselectivity.



Scheme 2-3: Preparation of amino octyl linker on heptose for preparation of glycocongugates.

Glycosylation of **2-25** with 8-azidooctanol promoted by trimethylsilyl trifluoromethanesulfonate (TMSOTf) led to the formation of **2-26** (95% yield) as a 4:1 α:β mixture of glycosides. The mixture of anomers was separable. Although the preference in selectivity is not known, it has been postulated that using a mild activator such as TMSOTf promotes an S_N2-like mechanism (Figure 2-6). This is a result of the activator forming a complex with the donor (**2-31**), which is favoured over the oxocarbenium ion (**2-32**), which leads to stereochemical scrambling.[27] In addition, this reaction was performed with diethyl ether as the solvent, which has been postulated to form an intimate ion pair with the oxocarbenium ion, with the ether occupying the equatorial (β) position. The β-intermediate (**2-33**), being more reactive than its α-counterpart (**2-34**), which is stabilized by the endo-anomeric effect, promotes the formation of the α-product **2-26**.

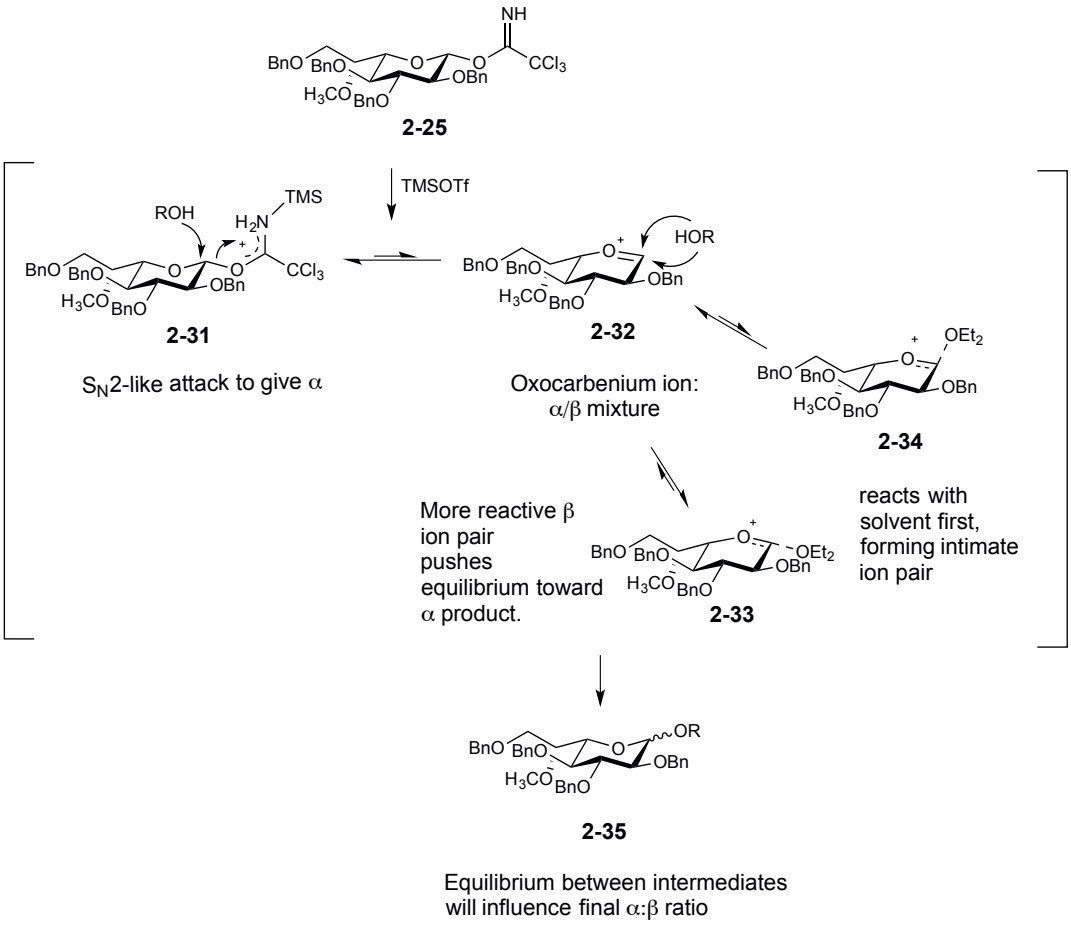


Figure 2-6: Rationalization for stereoselectivity of heptose glycosylation.

The conversion of **2-26** into **2-1** in a single step by treatment with hydrogen over a palladium on carbon catalyst proved problematic. Very low yields of impure product were obtained, as mixture of the mono, di, tri, and tetra-benzylated products, likely as a result of the reduced amine poisoning the palladium catalyst. Therefore, we explored a stepwise approach, which we had used when a similar problem arose in the course of another investigation.[28] Consequently, the azido group in **2-26** was selectively reduced by hydrogenation over a palladium catalyst poisoned by pyridine. The resulting amine **2-27** was protected as the corresponding *N*-trifluoroacetamide derivative **2-28** in 95% overall yield. Hydrogenolysis of the benzyl ethers in **2-28**, under high pressure (60 psi) and using acetic acid as solvent, gave **2-29**, which was

obtained in 91% yield. Finally, cleavage of the amide in **2-29** was carried out by reaction with sodium methoxide yielding an 83% yield of **2-1**. Subsequent reaction of **2-1** with amine-reactive bovine serum albumim, using the imm-Link^{TM} immunogen linking kits from Innova Biosciences, led to the expected BSA conjugate **2-30**, which was confirmed by matrix-assisted laser desorption/ionization (MALDI).

This section describes the first synthesis of the D-*glycero*-L-*gluco*-heptose monosaccharide present in the CPS of *C. jejuni* NCTC11168. The route involved the chain extension of a readily-obtained galactofuranose derivative (**2-5**), which was then converted into the target **2-1**. This approach differs from the usual approach used for the synthesis of heptoses, where homologation proceeds via a C-6 oxidized hexose derivative.

2.5 Alternate Heptose Route

A key drawback of the heptose route discussed in Section 2.2 was that this approach required a number of purification steps via flash chromatography. The large number of purification steps became very tedious and time consuming, which led us to look into other synthetic approaches that require less time and effort. This work is detailed below and was conducted together with Roger A. Ashmus, another graduate student in the group.

The synthesis of various D- and L-hexopyranoses from furfural has been previously investigated by Ogasawara and O'Doherty.[29–33] Ogasawara and coworkers' approach involved the synthesis of enone **2-39**, which was accessed from diol **2-37** through an Achmatowicz reaction, followed

by acid catalyzed ring closure (Figure 2-7).[29] O'Doherty and coworkers were able to access a variety of D- and L-sugars through the stereoselective manipulation of pyrenone **2-38**.[30–33]

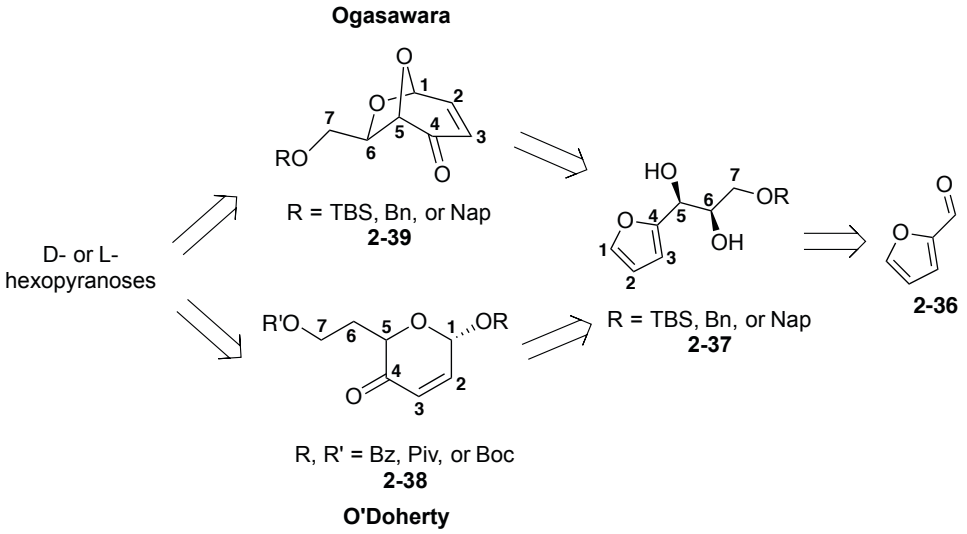


Figure 2-7: Retrosynthetic analyses for synthesis of heptose from furfural from the Ogasawara and O'Doherty routes.

After considering these approaches, we attempted to apply a similar strategy for the synthesis of this heptose from furfural. The final heptose donor, **2-2**, was to be accessed by the Ogasawara intermediate **2-39**, a dioxabicyclooctane framework

in which the stereochemistry at the 5 and 6 carbons are set in accordance with the D-*glycero*-L-*gluco*- configurations. The desired stereochemistry at C-4 would be accessed by reduction and stereoinversion, while introduction of the hydroxyl groups at C-2 and C-3 would result by epoxidation of the alkene followed by hydrolysis (Figure 2-8).

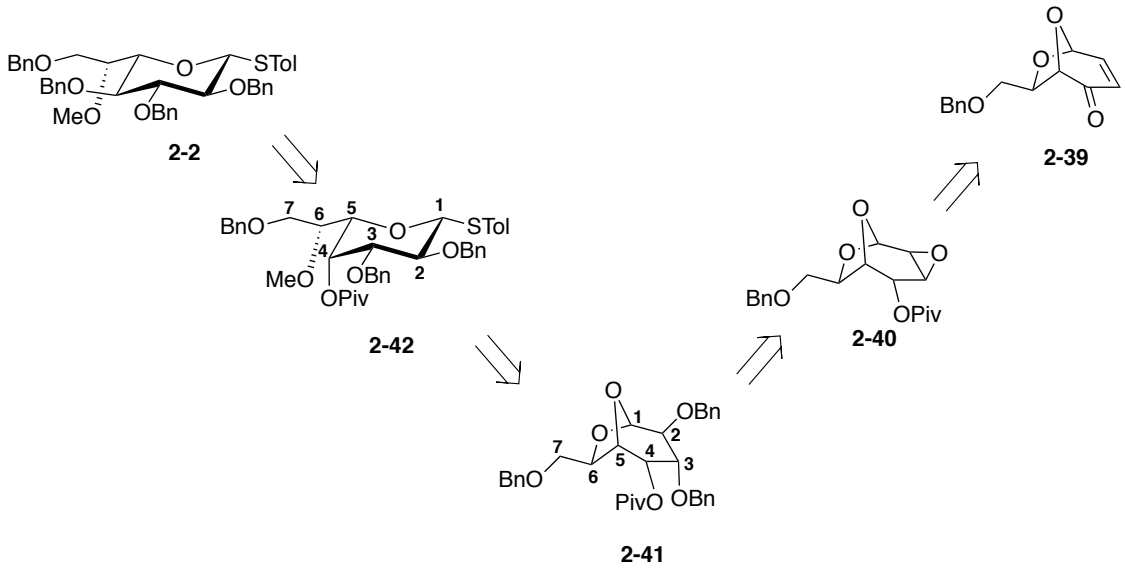
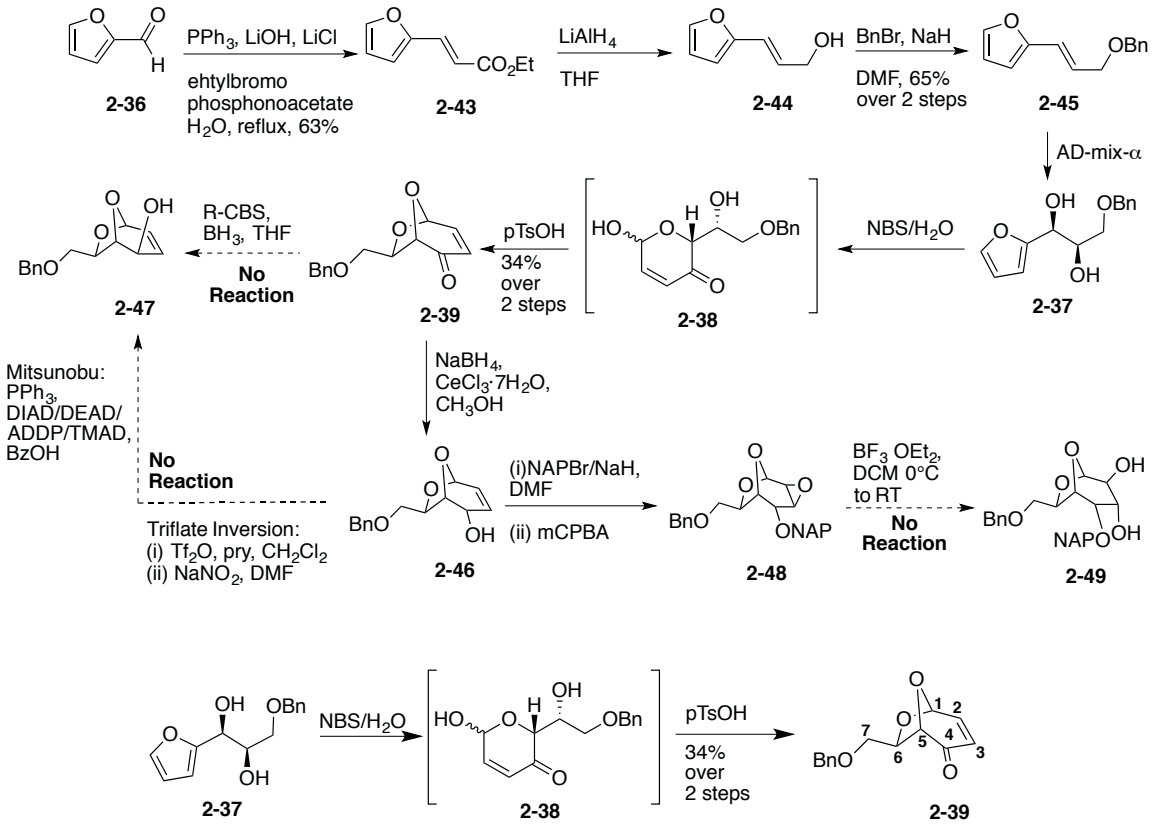


Figure 2-8: Retrosynthetic route to heptose from Ogawasara intermediate.

Scheme 2-4 shows the attempted route towards a more facile synthesis of heptoses based on Ogasawara's synthesis from furfural. Starting from furfural **2-36**, a Horner–Wadsworth–Emmons olefination with bromophosphonoacetate yields **2-43** in 63% yield, as a 99:1 mixture of *E*:*Z* alkenes. The ester in **2-43** then undergoes reduction with $LiAlH_4$ and benzylation of the resulting alcohol (**2-44**) to give **2-45** in 65% over two steps. Diol **2-37** was then obtained after asymmetric dihydroxylation of **2-45** with AD-mix α in 42% yield.* The enone, **2-39**, was obtained after Achmatowicz reaction of **2-37** with NBS, followed by acid catalyzed ring closure with *p*-toluenesulfonic acid, though the yield of this two-step reaction was low (34% over two steps).

* This step was completed by Roger A. Ashmus

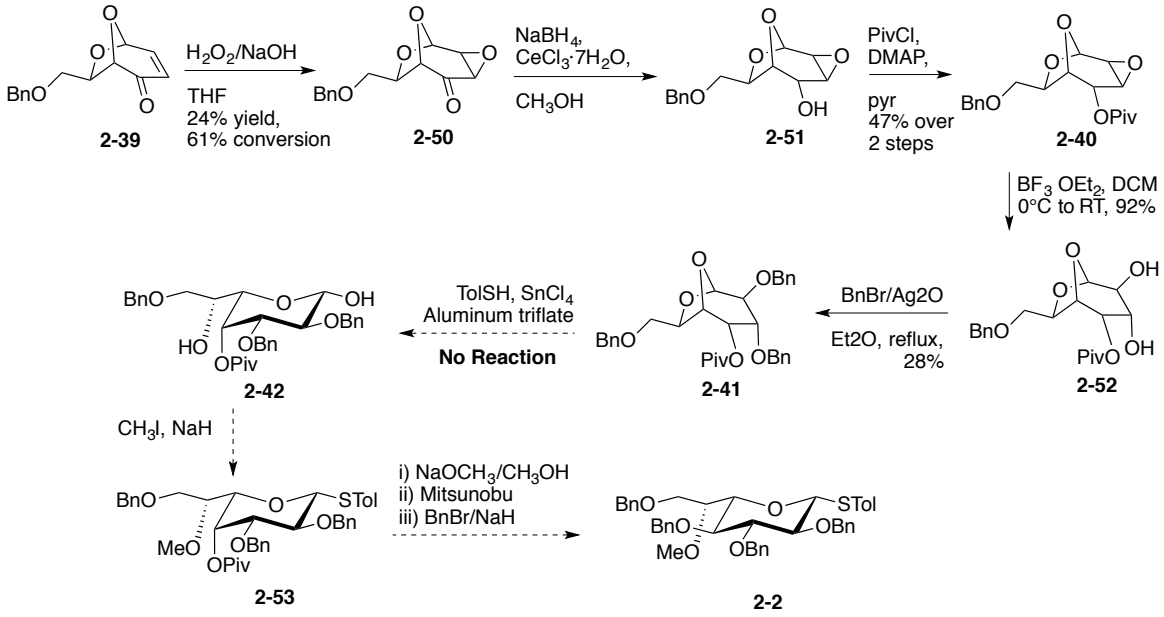


Scheme 2-4: Route toward Ogawasara dioxabicyclooctane framework, and stereocontrolled introduction of hydroxyl groups.

Reduction of the carbonyl group in **2-39** under Luche conditions provided **2-46**, which had the opposite C-4 stereochemistry to that desired; therefore, an inversion step was necessary at this position. However, this reaction proved very difficult. Inversion of the C-4 stereocentre to yield **2-47** from **2-46** by Mitsunobu conditions using diisopropyl azodicarboxylate (DIAD), diethyl azodicarboxylate (DEAD), 1,1'-(azodicarbonyl)dipiperidine (ADDP)[34], *N*,*N*,*N*',*N*'-tetramethylazodicarboxamide (TMAD)[35], proved unsuccessful. Similarly, triflate displacement with sodium nitrate failed. The stereoselective reduction of the carbonyl group with (*R*)-(+)-2-Methyl-CBS-oxazaborolidine (*R*-CBS)[36] to attempt the direct conversion of **2-39** to **2-47** also failed. Additionally, installing a naphthylmethyl protecting group on the 4-OH followed by

epoxidation of the alkene yielded a stable intermediate **2-48**, whose epoxide ring could not be opened by using BF_3-OEt_2, thus failing to yield **2-49**.

To circumvent the problem of 4-OH inversion, we investigated performing this step on the pyranose ring (**2-42)**. This was done because we suspected that steric impedance on the bottom face of the bicyclic ring in **2-46** prohibited to formation of the oxyphosphonium ion the Mitsunobu reaction. In implementing this approach, it was necessary to find conditions by which the epoxide moiety in molecule **2-48** could be opened to provide a diol. To address problems associated with ring opening of the epoxide, we chose to change the naphthylmethyl protecting group to a pivalate ester, which we expected would help open the epoxide by neighbouring group participation. This strategy proved to be successful (Scheme 2-5). After epoxidation of the electrophilic alkene in **2-39** using H_2O_2/NaOH, the ketone in the resulting product, **2-50**, was reduced and protected with a pivaloate ester to yield **2-40**. Epoxide **2-40** was successfully opened using BF_3-OEt_2 in 92% yield to give **2-52**. After protection of the diol in **2-52** as benzyl ethers, however, compound **2-41** was found to be unreactive towards opening of the bicyclic ring. Thus, the L-sugar **2-42** could not be obtained by thioglycosylation to break the seven-membered component of the bicyclic ring. After experimenting with stronger Lewis acids such as $SnCl_4$, the reaction still did not progress. Attempts at pushing the reaction with higher temperatures lead to the degradation of **2-41**, and so this route was abandoned. Given time limitations, I did not further pursue this route.



Scheme 2-5: Route toward heptose from Ogawasara intermediate **2-39**.

2.6 Current Work & Future Directions

With the route shown in Scheme 2-5 abandoned, current work by other group members is exploring the synthesis of the heptose from furfural by functionalizing intermediate enone **2-38**, via O'Doherty's approach (Figure 2-9). Briefly, enone **2-54** will be reduced to yield the desired stereochemistry at C-4. This is followed by dihydroxylation of the alkene, protecting group chemistry so that the 2-OH can be selectively inverted, followed by the installation of a *p*-toluenethio group at the anomeric position.

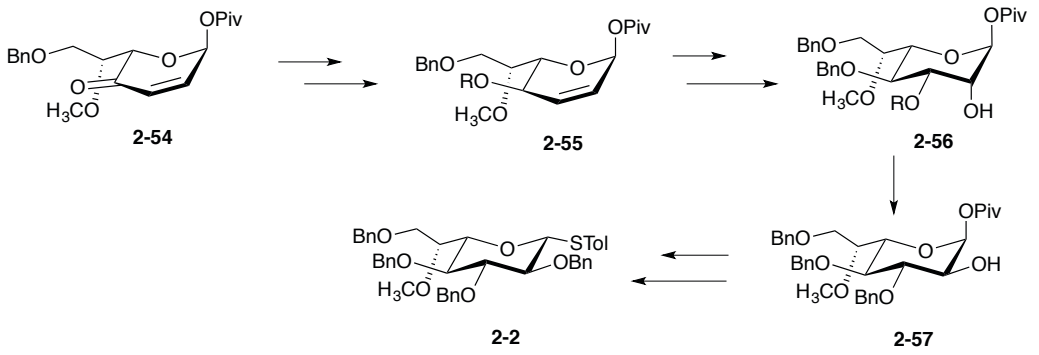


Figure 2-9: Current approach to heptose from O'Doherty enone **2-54**.

2.7 Experimental

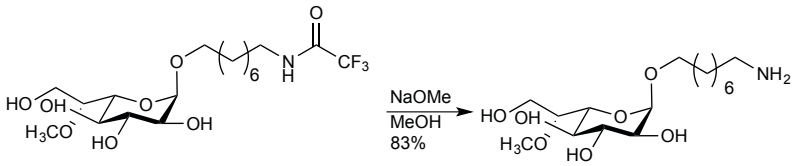
2.7.1 General Experimental Methods

All reagents were purchased from commercial sources and were used without further purification unless noted. Solvents used in reactions were purified by successive passage through columns of alumina and copper under argon. Unless stated otherwise, all reactions were carried out at room temperature and under a positive pressure of argon and monitored by TLC on Silica Gel G-25 F254 (0.25 mm). TLC spots were detected under UV light and/or by charring with a solution of anisaldehyde in ethanol, acetic acid and H_2SO_4. Column chromatography was performed on Silica Gel 60 (40–60 μm). Organic solutions were dried using anhydrous Na_2SO_4 and solvents were evaporated under reduced pressure and below 50 °C (water bath) on a rotary evaporator. ^{1}H NMR and ^{13}C NMR spectra were recorded at 400, 500 or 600 MHz. ^{1}H NMR chemical shifts are referenced to $CHCl_3$ (7.26 ppm, $CDCl_3$). ^{13}C NMR chemical shifts are referenced to $CDCl_3$ (77.0 ppm, $CDCl_3$). ^{1}H NMR data are reported as though they were first order, and the peak assignments were made on the basis of 2D-NMR (^{1}H–^{1}H COSY and HSQC) experiments. ESI-MS spectra were recorded on samples suspended in THF or CH_3OH and added NaCl. Optical

rotations were measured at 22 ± 2 °C at the sodium D line (589 nm) and are in units of deg·mL(dm·g)$^{-1}$.

2.7.2 Experimental, spectroscopic, and analytical data

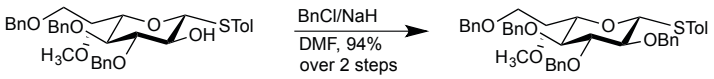
8-Aminooctyl 6-*O*-methyl-D-*glycero*-*α*-L-*gluco*-heptopyranoside (2-1)



2-29 **2-1**

To a solution of **2-29** (4 mg, 0.0086 mmol) in CH_3OH (0.5 mL) was added 1M $NaOCH_3$ until the pH of the solution was 9. After stirring for 48 h, Amberlite IR120 H+ resin was added to neutralize the solution; after filtration the solution was concentrated. The residue was dissolved in H_2O and passed through a C_{18}-SepPak cartridge and eluted with a gradient of 0 to 30% CH_3OH in H_2O to give **2-1** (2.5 mg, 83%) as a white solid. Rf = 0.23 (5:1 CH_3OH–CH_2Cl_2); $[\alpha]_D$ –48.9 (*c* 0.3, CH_3OH); ^{1}H NMR (500 MHz, CD_3OD) δ 4.71 (d, 1H, *J* = 3.8 Hz, H-1), 3.76 (dd, 1H, *J* = 6.0, 10.3 Hz, octyl OC*H*$_2$), 3.57–3.70 (m, 5H, H-7a, H-6, H-5, octyl OC*H*$_2$, H-3), 3.52 (s, 3H, C*H*$_3$O), 3.50–3.44 (m, 1H, H-4), 3.43–3.33 (m, 2H, H-7b, H-2), 2.83 (t, 2H, *J* = 7.6 Hz, octyl NC*H*$_2$), 1.90 (s, 3H, acetate salt OAc) 1.69–1.51 (m, 2H, octyl), 1.36 (s, 10H, octyl); ^{13}C NMR (125 MHz, CD_3OD): δ 100.2 (C-1), 80.5 (C-5), 75.4 (C-3), 73.5 (C-2), 72.0 (C-6), 71.1 (C-4), 69.0 (C-7), 62.4 (octyl O*C*H$_2$), 60.3 (C*H*$_3$O), 40.8 (octyl N*C*H$_2$), 30.5 (octyl *C*H$_2$), 30.3 (octyl *C*H$_2$), 30.2 (octyl *C*H$_2$), 27.5 (octyl *C*H$_2$), 27.2 (octyl *C*H$_2$). HR ESIMS: *m*/*z* [M+Na$^+$] calcd for $C_{16}H_{33}NO_7Na$: 374.2149. Found 374.2148.

***p*-Tolyl 3,4,7-Tri-*O*-benzyl-6-*O*-methyl-1-thio-D-*glycero*-*α*/*β*-L-*gluco*-heptopyranoside (2-2)**

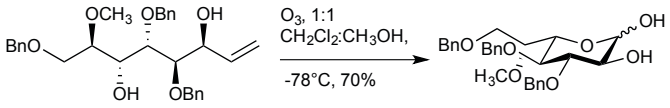


2-21 **2-2**

Compound **2-20**, as a α/β mixture of isomers (628 mg, 0.90 mmol), was dissolved in 3:1 CH_2Cl_2–CH_3OH (6 mL) and $NaOCH_3$ in CH_3OH was added until the pH of the solution was above 10. After 4 h, the solution was neutralized by the addition of IR-120 (H+), filtered, and concentrated. To a solution of crude 2-20 α/β (538 mg, 2.2 mmol) in anhydrous DMF (5 mL) at 0 °C were added BnBr (0.25 mL, 1.5 equiv) and NaH (61 mg, 2.0 equiv). After stirring for 6 h, CH_3OH (0.5 mL) was added and the solution was concentrated. The residue was purified by chromatography (9:1 hexane–EtOAc) to give **2-2** (1680m g, 94% over two steps) as a syrup. **2-2β** isomer: R_f = 0.54 (4:1 hexane–EtOAc); **2-2β** isomer: $[\alpha]_D$ –20.1 (*c* 0.8, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.42 (d, 2H, *J* = 8.1 Hz, Ph), 7.39–7.21 (m, 15H, Ph), 7.05 (d, 2H, *J* = 7.9 Hz, Ph), 4.95 (d, 1H, *J* = 11.0 Hz, PhC*H*$_2$), 4.94 (d, 1H, *J* = 11.3 Hz, PhC*H*$_2$), 4.86 (d, 1H, *J* = 11.3 Hz, PhC*H*$_2$), 4.66 (d, 1H, *J* = 11.0 Hz, PhC*H*$_2$), 4.46 (d, 1H, *J* = 11.8 Hz, PhC*H*$_2$), 4.43 (d, 1H, *J* = 9.7 Hz, H-1), 4.42 (d, 1H, *J* = 11.8 Hz, PhC*H*$_2$), 3.88–3.75 (m, 3H, H-4, H-6, H-7a), 3.67–3.60 (m, 2H, H-3, H-7b), 3.56 (app t, 1H, *J* = 9.1 Hz, H-2), 3.47 (dd, 1H, *J* = 9.8, 1.6 Hz, H-5), 3.46 (s, 3H, C*H*$_3$O), 2.42 (s, 1H, OH), 2.29 (s, 3H, C*H*$_3$Ph); ^{13}C NMR (125 MHz, $CDCl_3$): *δ* 138.7, 138.5, 138.3, 138.2, 133.1, 129.8, 129.2, 128.6, 128.6, 128.5, 128.1, 127.9, 127.8, 127.7(5), 127.6, 89.9 (C-1), 86.6 (C-3), 78.9 (C-5), 76.8 (C-4 or C-6), 76.5 (C-6 or C-4), 75.3 (Ph*C*H$_2$), 75.0 (Ph*C*H$_2$), 73.5 (Ph*C*H$_2$), 73.4 (C-2), 69.3 (C-7), 58.6 (*C*H$_3$O), 21.2 (*C*H$_3$Ph). HR ESIMS: *m*/*z* [M+Na$^+$] calcd for $C_{36}H_{40}O_6SNa$: 623.2438. Found: 623.2435. **2-2α** isomer: $[\alpha]_D$ –198.3 (*c* 0.6, $CHCl_3$). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.49–7.20 (m, 17H, Ph), 7.07 (d, 2H, *J* = 7.9 Hz, Ph), 5.66 (d, 1H, *J* = 5.5 Hz, H-1), 4.96 (d, 1H, *J* = 11.0 Hz, PhC*H*$_2$), 4.91 (s, 2H,

PhCH_2), 4.68 (d, 1H, J = 11.0 Hz, PhCH_2), 4.43 (d, 1H, J = 11.9 Hz, PhCH_2), 4.38 (d, 1H, J = 11.9 Hz, PhCH_2), 4.27–4.23 (m, 1H, H-5), 4.04 (dd, 1H, J = 9.4, 5.5 Hz, H-2), 3.87 (app t, 1H, J = 5.7 Hz, H-6), 3.80 (app t, 1H, J = 10.2 Hz, H-4), 3.71 (app t, 1H, J = 9.4 Hz, H-3), 3.64 (dd, 1H, J = 9.9, 6.4 Hz, H-7a), 3.49 (s, 3H, CH_3O), 3.44 (dd, 1H, J = 9.9, 5.0 Hz, H-7b), 2.33 (s, 3H, CH_3Ph); ^{13}C NMR (125 MHz, $CDCl_3$): δ 138.4, 138.2, 138.1, 137.2, 131.2, 130.4, 129.8, 128.6, 128.5, 128.3, 127.9, 127.8(7), 127.8(1), 127.7, 127.6, 127.5, 89.9 (C-1), 83.9 (C-3), 77.1 (C-4), 76.6 (C-6), 75.4 (PhCH$_2$), 75.0 (PhCH$_2$), 73.4 (PhCH$_2$), 72.3 (C-2), 72.2(5) (C-5), 70.5 (C-7), 59.0 (CH$_3$O), 21.1 (CH$_3$Ph); HR ESIMS: m/z [M+Na$^+$] calcd for $C_{36}H_{40}O_6SNa$: 623.2438. Found: 623.2430.

3,4,7-Tri-*O*-benzyl-6-*O*-methyl-D-*glycero*-α/β-L-*gluco*-heptopyranose (2-3)

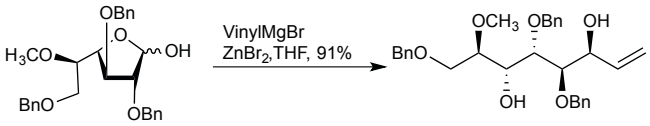


2-4 **2-3**

A stream of O_2/O_3 was passed through a solution of **2-4** (2.48 g, 5.0 mmol) in CH_2Cl_2 (10 mL) and CH_3OH (10 mL) at –78 °C with gentle stirring until the solution took on a blue color. Dimethyl sulfide was then added (2.6 mL, 7.0 equiv). After warming to rt, the mixture was concentrated and the residue was purified on silica gel column (2:1 → 1:1 hexane–EtOAc) to afford a 2:1 α/β mixture of **2-3** (1.72 mg, 70%) as a syrup, Rf = 0.38 (1:1 hexane–EtOAc). ^{1}H NMR (600 MHz, $CDCl_3$): δ = 7.44–7.22 (m, 22.5H, *Ph*-), 5.19 (d, J = 3.8 Hz, 1H, H-1$_\alpha$), 4.99–4.83 (m, 4.5H, PhCH_2), 4.68–4.61 (m, 1.5H, PhCH_2-), 4.59–4.47 (m, 3H, PhCH_2), 4.39 (d, J = 7.7 Hz, 0.5H, H-1$_\beta$), 4.01 (dd, J = 9.9, 1.4 Hz, 1H, H-4$_\alpha$), 3.86–3.81 (m, 2H, H-3$_\alpha$, H-6$_\alpha$), 3.81–3.72 (m, 3.5H, H-5$_\alpha$, H-7a$_\alpha$, H-4$_\beta$, H-6$_\beta$, H-7a$_\beta$), 3.73–3.65 (m, 2.5H, H-7b$_\beta$, H-2$_\alpha$, H-7b$_\alpha$), 3.59

(t, J = 9.1 Hz, 0.5H, H-3$_\beta$), 3.48 (m, 0.5H, H-5$_\beta$), 3.47 (s, 3H, CH_3O$_\alpha$), 3.46–3.43 (m, 0.5H, H-2$_\beta$), 3.42 (s, 1.5H, CH_3O$_\beta$). ESI-MS: m/z [M + Na$^+$] 517.2.

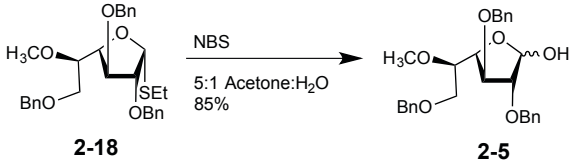
(2*R*,3*R*,4*R*,5*R*,6*S*)-1,4,5-tris-benzyloxy-2-methoxy-oct-7-ene-3,6-diol (2-4)



2-5 **2-4**

Fresh vinyl magnesium bromide was prepared by adding magnesium (0.66g, 27 mmol) to 1M vinyl bromide (41 mL, 41.0 mmol) in THF, with a crystal of NaI and heating at reflux for 1 h. The solution (5 equiv) was cooled and then added to dried $ZnBr_2$ (3.08 g, 13.7 mmol, 2.5 equiv). After 20 min, the fresh divinylzinc solution was slowly added to a solution of hemiacetal **2-3** (2.5 g, 5.5 mmol) in anhydrous THF (15 mL). The mixture was stirred for 4 h and concentrated. The residue was diluted with EtOAc and washed with 1N HCl, water, and brine, before being concentrated. The resulting residue was purified by chromatography (1:1 hexane–EtOAc) to give **2-4**(2.46 g, 91%) as a syrup. R_f = 0.48 (1:1 hexane–EtOAc); $[\alpha]_D$ –10.2 (*c* 0.5, $CHCl_3$); ^{1}H NMR (600 MHz, $CDCl_3$): δ 7.45–7.18 (m, 15H, *Ph*CH$_2$), 5.97 (ddd, 1H, J = 17.1, 10.5, 5.6 Hz, H-7), 5.37 (dt, 1H, J = 17.2, 1.6 Hz, H-8a), 5.21 (1H, dt, J = 10.5, 1.5 Hz, H-8b), 4.78 (d, 1H, J = 11.4 Hz, PhCH_2), 4.70 (d, 1H, J = 11.4 Hz, PhCH_2), 4.67 (d, 1H, J = 11.4 Hz, PhCH_2), 4.62–4.52 (m, 3H, PhCH_2), 4.51–4.37 (m, 1H, H-6), 3.98 (dd, 1H, J = 8.1, 1.6 Hz, H-3), 3.85–3.74 (m, 3H, H-1a, H-4, H-5), 3.73–3.59 (m, 2H, H-1b, H-2), 3.43 (s, 3H, CH_3O), 3.05 (brs, 2H, O*H*); ^{13}C NMR (100 MHz, $CDCl_3$): δ 138.6 (C-7), 138.3, 138.2, 138.0, 128.6, 128.5, 128.5, 128.3, 128.0, 127.8(6), 127.7(9), 116.2 (C-8), 820 (C-5), 78.1 (C-4), 77.7 (C-2), 74.8 (Ph*C*H$_2$), 73.9 (Ph*C*H$_2$), 73.7 (Ph*C*H$_2$), 72.6 (C-3), 72.3 (C-6), 70.2 (C-1), 58.0 (*C*H$_3$O). HR ESIMS: m/z [M+Na$^+$] calcd for $C_{30}H_{36}O_6Na$: 515.2404. Found: 515.2397.

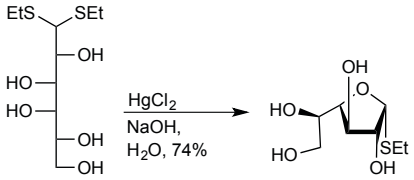
2,3,6-Tri-*O*-benzyl-5-*O*-methyl-α/β-D-galactofuranoside (2-5)



2-18 **2-5**

To a solution of **2-18** (4.44 g, 8.73 mmol) in acetone (100 mL) and water (20 mL) at 0 °C was added NBS (3.11, 2.0 equiv). After 1 h, the mixture was concentrated and the residue was diluted with CH_2Cl_2. This solution was washed with $NaHCO_3$ (satd), dried, and concentrated. Chromatography of the crude product (2:1 hexane–EtOAc) gave **2-5** (3.44 g, 85%) as a syrup. Rf = 0.29 (2:1 hexane–EtOAc);); ^{1}H NMR (600 MHz, $CDCl_3$): δ 7.56–7.15 (m, 15H), 5.43 (d, 0.6H, $J_{1,2}$ = 0 Hz, $J_{1,\text{-}OH}$ = 5.5 Hz, H-1$_\beta$), 5.28 (dd, 0.4H, $J_{1,2}$ = 4.5 Hz, $J_{1,\text{-}OH}$ = 10.9 Hz, H-1$_\alpha$), 4.77–4.45 (m, 6H, PhCH_2), 4.34 (app t, 0.6H, J = 5.2 Hz, H-4$_\beta$), 4.18 (app t, 0.4H, J = 5.0 Hz, H-3$_\alpha$), 4.10 (dd, 0.4H, J = 5.0, 2.5 Hz, H-4$_\alpha$), 4.06 (dd, 0.6H, J = 5.2, 2.4 Hz, H-3$_\beta$), 4.04 (dd, 0.4H, J = 5.5, 4.6 Hz, H-2$_\alpha$), 4.00 (dd, 0.6H, J = 2.4, 1.0 Hz, H-2$_\beta$), 3.69 (dd, 0.4H, J = 10.0, 6.3 Hz, H-6$_\alpha$), 3.64 (dd, 0.4H, J = 10.0, 5.4 Hz, H-6$_\alpha$'), 3.63 (dd, 0.6H, J = 10.3, 4.5 Hz, H-6$_\beta$), 3.59 (dd, 0.6H, J = 10.4, 5.8 Hz, H-6$_\beta$'), 3.54-3.50 (m, 0.6H, H-5$_\beta$), 3.47 (s, 1.8H, CH_3O), 3.45 (s, 1.2H, CH_3O), 3.38–3.33 (m, 0.4H, H-5$_\alpha$), 3.17 (d, 0.6H, J = 5.8 Hz, OH$_\beta$); ^{13}C NMR (100 MHz, $CDCl_3$): δ 138.2, 138.0, 137.8, 137.5(3), 137.5(0), 128.4(8), 128.4(5), 128.3(9), 128.3(4), 128.1, 128.0, 127.9(0), 127.8(7), 127.6(3), 127.6(2), 100.9 (C-1$_\beta$), 95.9 (C-1$_\alpha$), 87.3 (C-2$_\beta$), 84.6 (C-2$_\alpha$), 82.5 (C-4$_\beta$), 82.3 (C-3$_\beta$), 81.3(5) (C-3$_\alpha$), 81.3(0) (C-4$_\alpha$), 79.7 (C-5$_\beta$), 79.5 (C-5$_\alpha$), 73.5, 73.4, 72.1, 72.0, 71.94, 71.8, 70.1 (C-6$_\beta$), 70.0 (C-6$_\alpha$), 59.5 (OCH$_{3\alpha}$), 59.1 (OCH$_{3\beta}$). HR ESIMS: *m*/*z* [M+Na$^+$] calcd for $C_{28}H_{32}O_6Na$: 487.2091. Found: 487.2084.

Ethyl 1-thio-α-D-galactofuranoside (2-6)



2-8 **2-6**

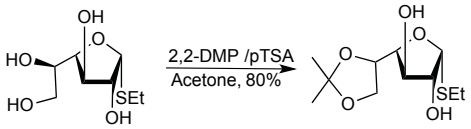
To a solution of D-galactose diethyl dithioacetal (20.0 g, 0.0699 mol) in water (400 mL) was added a solution of mercury (II) chloride (24.7 g, 0.0909 mol) in water (290 mL) and then 1.0M NaOH was added portion-wise, to maintain a solution pH of 6–7. Once disappearance of the starting material was observed by TLC, along with stabilization of the pH of the solution,[†] the excess mercury (II) chloride was filtered, and the filtrate was concentrated under reduced pressure.[‡] The resulting residue was purified by column chromatography (6:1 CH_2Cl_2–CH_3OH) to give **2-6** (11.2 g, 71%) as an oil. ^{1}H NMR (600 MHz, D_2O) δ 5.40 (d, 1H, *J* = 5.4 Hz, H-1), 4.27 (app t, 1H, *J* = 5.4 Hz, H-2), 4.18 (dd, 1H, *J* = 5.4, 5.1 Hz, H–3), 3.85–3.87 (m, 1H, H-5), 3.80 (app t, 1H, *J* = 5.1 Hz, H-4), 3.71 (dd, 1H, *J* = 4.2, 11.4 Hz, H-6a), 3.62 (dd, 1H, *J* = 7.2, 11.4 Hz, H-6b), 2.74 (q, 2H, *J* = 7.2 Hz, CH_3CH_2), 1.28 (t, 1 H, *J* = 7.2 Hz, CH_3CH_2). ^{13}C NMR (125 MHz, $CDCl_3$, δ_C): 87.5 (C-1), 83.7 (C-4), 77.6 (C-2), 76.8 (C-3), 71.7 (C-5), 62.6 (C-6), 25.0 (CH_2), 14.5 (CH_3).

---

† Once the mercury (II) chloride finishes reacting with the dithioacetal, there is a cessation of HCl production.

‡ 1 mL of octanol was added to prevent foaming of the solution when concentrating under reduced pressure.

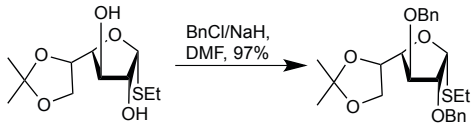
Ethyl 5,6-*O*-Isopropylidene-1-thio-α-D-galactofuranoside (2-9)



2-6 **2-9**

To a solution of **2-6** (14.81 g, 66.0 mmol) and *p*-TsOH (0.57 g, 10 mmol, 0.1 equiv) in dry 4:1 acetone:DMF (200 mL) at 0 °C was added 2,2-dimethoxypropane (40 mL, 330 mmol, 5.0 equiv). The reaction mixture was stirred overnight, neutralized with $NaHCO_3$ and filtered. The filtrate was dried, concentrated and the residue was purified by chromatography (2:1 hexane–EtOAc) to afford **2-9** (13.9 g, 80%) as a white solid. R_f = 0.48 (2:1 hexane–EtOAc); R_f = 0.48 (1:1 CH_2Cl_2–EtOAc); $[\alpha]_D$ +103.0 (*c* 0.9, $CHCl_3$); ^{1}H NMR (600 MHz, $CDCl_3$) δ 5.31 (d, 1H, *J* = 4.2 Hz, H-1), 4.27 (ddd, 1H, *J* = 8.1 6.7, 3.2 Hz, H-5), 4.18 (app t, 1H, *J* = 3.1 Hz, H-3), 4.08–4.00 (m, 2H, H-6a, H-2), 3.97 (app t, 1H, *J* = 8.1 Hz, H-6b), 3.85 (app t, 1H, *J* = 3.1 Hz, H-4), 2.70 (q, 2H, *J* = 7.4 Hz, CH_3CH_2), 1.47 (s, 3H, $(CH_3)_2C$), 1.37 (s, 3H, $(CH_3)_2C$), 1.29 (t, 3H, *J* = 7.4 Hz, CH_3CH_2); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 110.3 ($C(CH_3)_2$), 89.2 (C-1), 84.0 (C-4), 78.7 (C-2), 78.6 (C-3, C-5), 65.6 (C-6), 26.3 ($CH_3)_2C$), 25.9 (CH_3CH_2), 25.8 ($CH_3)_2C$), 15.6 (CH_3CH_2). HR ESIMS: *m*/*z* [M+Na^+] calcd for $C_{11}H_{20}O_5SNa$: 287.0924. Found: 287.0927.

Ethyl 2,3-Di-*O*-benzyl-5,6-*O*-isopropylidene-1-thio-α-D-galactofuranoside (2-10)

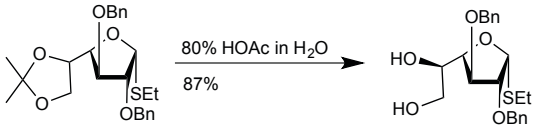


2-9 **2-10**

To a solution of **2-9** (8.1 g, 30.6 mmol) in anhydrous DMF (75 mL) at 0 °C were added BnBr (9.5 mL, 2.6 equiv) and NaH (3.7 g, 3.0 equiv). The mixture was slowly warmed to rt and stirred for another 4 h, before CH3OH (2 mL) was added and the solution was concentrated. The

residue was dissolved in CH_2Cl_2 and washed with HCl (1 N), dried over Na_2SO_4, and concentrated under reduced pressure. Purification by chromatography (7:1 hexane–EtOAc) gave **2-10** (13.3 g, 97%) as a syrup. R_f = 0.35 (6:1 hexane–EtOAc); . R_f = 0.35 (6:1 hexane–EtOAc); $[\alpha]_D$ +69.5 (*c* 0.9, $CHCl_3$); ^{1}H NMR (600 MHz, $CDCl_3$): δ 7.52–7.18 (m, 10H, *Ph*CH_2), 5.45 (d, 1H, *J* = 4.8 Hz, H-1), 4.72 (d, 1H, *J* = 11.6 Hz, PhC*H*$_2$), 4.61 (d, 1H, *J* = 11.6 Hz, PhC*H*$_2$), 4.53 (d, 1H, *J* = 11.4 Hz, PhC*H*$_2$), 4.51 (d, 1H, *J* = 12.0 Hz, PhC*H*$_2$), 4.40 (q, 1H, *J* = 7.3 Hz, H-5), 4.24 (app t, 1H, *J* = 4.8 Hz, H-2), 4.01 (app t, 1H, *J* = 4.8 Hz, H-3), 3.93 (dd, 1H, *J* = 7.3, 4.8 Hz, H-4), 3.86 (dd, 1H, *J* = 8.5, 7.3 Hz, H-6a), 3.80 (dd, 1H, *J* = 8.5, 7.0 Hz, H-6b), 2.75 (q, 2H, *J* = 7.4 Hz, CH_3CH_2), 1.43 (s, 3H, $(CH_3)_2C$), 1.36 (s, 3H, $(CH_3)_2C$), 1.32 (t, *J* = 7.4 Hz, 3H, CH_3CH_2); ^{13}C NMR (100 MHz, $CDCl_3$): *δ* 137.5, 137.3, 128.5, 128.1, 128.9(8), 127.9(6), 127.9(4), 109.6 ($(CH_3)_2C$), 86.5 (C-1), 84.4 (C-2), 83.5 (C-4), 82.5 (C-3), 77.2 (C-5), 72.5 (Ph*C*H_2), 72.1 (Ph*C*H_2), 65.26 (C-6), 26.6 ($(CH_3)_2C$), 25.2 ($(CH_3)_2C$), 24.6 (CH_3CH_2), 15.0 (CH_3CH_2). HR ESIMS: *m*/*z* [M+Na^+] calcd for $C_{25}H_{32}O_5SNa$: 467.1863. Found: 467.1861.

Ethyl 2,3-Di-*O*-benzyl-1-thio-α-D-galactofuranoside (2-11)

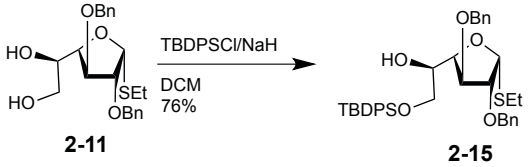


2-10 **2-11**

Compound **2-10** (12.0 g, 27.0 mmol) was dissolved in 70% aq HOAc (50 mL) and the solution was stirred at 60 °C overnight. The solution was cooled, concentrated and the residue was purified by chromatography (2:1, hexane–EtOAc) to afford **2-11** (9.5 g, 87%) as a syrup. R_f = 0.42 (2:1 hexane–EtOAc); $[\alpha]_D$ +68.6 (*c* 0.6, $CHCl_3$); ^{1}H NMR (600 MHz, $CDCl_3$): δ 7.53–7.07 (m, 10H, *Ph*CH_2), 5.41 (d, 1H, *J* = 4.7 Hz, H-1), 4.67 (d, 1H, *J* = 11.6 Hz, PhC*H*$_2$), 4.63 (d, 1H, *J* = 11.7 Hz, PhC*H*$_2$), 4.55 (d, 1H, *J* = 11.4 Hz, PhC*H*$_2$), 4.53 (d, 1H, *J* = 12.0 Hz, PhC*H*$_2$), 4.26

(app t, 1H, J = 4.7 Hz, H-3), 4.22 (app t, 1H, J = 4.7 Hz, H-2), 4.04 (app t, 1 H, J = 4.7 Hz, H-4), 3.79-3.74 (m, 1H, H-5), 3.66-3.60 (m, 2H, H-6ab), 3.00 (brs, 1H, OH), 2.71 (q, 2H, J = 7.4 Hz, CH_3CH_2), 2.30 (brs, 1H, OH), 1.33 (t, 3H, J = 7.4 Hz, CH_3CH_2); ^{13}C NMR (100 MHz, $CDCl_3$): δ 137.5, 137.1, 128.5(1), 128.4(9), 128.0(4), 127.9(9), 127.8, 87.4 (C-1), 83.8 (C-2), 83.6 (C-4), 82.4 (C-3), 72.5 ($PhCH_2$), 72.4 ($PhCH_2$), 71.3 (C-5), 64.4 (C-6), 25.2 (CH_3CH_2), 15.2 (CH_3CH_2). HR ESIMS: m/z [M+Na$^+$] calcd for $C_{22}H_{28}O_5SNa$: 427.1550. Found: 427.1542.

Ethyl 2,3-Di-*O*-benzyl-6-*O*-tert-butyldiphenylsilyl-1-thio-α-D-galactofuranoside (2-15)

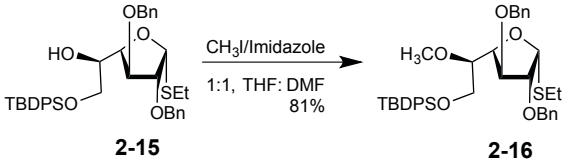


2-11 **2-15**

A solution of diol **2-11**(6.4 g, 15.8 mmol) in anhydrous DMF (70 mL) at 0 °C was treated with TBDPSCl (4.3 mL, 1.05 equiv) and imidazole (2.2 g, 2.0 equiv). The mixture was warmed to rt while stirring overnight, and then CH_3OH (0.5 mL) was added and the solution was concentrated. The residue was dissolved in CH_2Cl_2 and washed with HCl (1N), dried, and concentrated under reduced pressure. Purification by chromatography (10:1 hexane–EtOAc) afforded **2-15** (9.6 g, 76%) as a syrup. R_f = 0.30 (8:1 hexane–EtOAc); $[\alpha]_D$ +39.7 (*c* 0.9, $CHCl_3$); ^{1}H NMR (600 MHz, $CDCl_3$): δ 7.71–7.64 (m, 4H, Ph), 7.46–7.23 (m, 16H, Ph), 5.44 (d, 1H, J = 4.8 Hz, H-1), 4.67 (d, 1H, J = 11.7 Hz, $PhCH_2$), 4.62 (d, 1H, J = 11.6 Hz, $PhCH_2$), 4.55 (d, 1H, J = 11.4 Hz, $PhCH_2$), 4.52 (d, 1H, J = 11.4 Hz, $PhCH_2$), 4.32-4.28 (m, 2H, H-3, H-4), 4.26 (app t, 1H, J = 4.8 Hz, H-2), 3.82–3.77 (m, 1H, H-5), 3.73–3.68 (m, 2H, H-6ab), 2.90 (d, 1H, J = 7.8 Hz, OH), 2.67 (dq, 2H, J = 7.4, 2.5 Hz, CH_3CH_2), 1.30 (t, 3H, J = 7.4 Hz, CH_3CH_2), 1.08 (s, 9H, $(CH_3)_3C$); ^{13}C NMR (100 MHz, $CDCl_3$): δ 137.8, 137.3, 135.6, 135.6, 133.5, 129.6(9), 129.6(5), 128.5, 128.0, 127.8, 127.7, 127.6(6), 87.0 (C-1), 83.9 (C-2), 82.5 (C-3,4), 72.4 (2 $PhCH_2$), 71.3

(C-5), 64.7 (C-6), 26.9 $(CH_3)_3C$), 25.0 (CH_3CH_2), 19.3 ($C(CH_3)_3C$), 15.1 (CH_3CH_2). HR ESIMS: *m*/*z* [M+Na^+] calcd for $C_{38}H_{46}O_5SSiNa$: 665.2727. Found: 665.2717.

Ethyl 2,3-Di-*O*-benzyl-6-*O*-tert-butyldiphenylsilyl-5-*O*-methyl-1-thio-α-D-galactofuranoside (2-16)

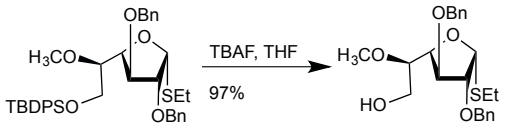


2-15 **2-16**

A solution of **2-15** (9.5 g, 14.8 mmol) in anhydrous DMF (70 mL) at 0 °C was treated with CH_3I (2.6 mL, 1.3 equiv) and NaH (5.5g, 1.5 equiv) was added slowly batch-wise. The mixture was slowly warmed to rt and, after the starting material disappeared, CH_3OH (1 mL) was added and the solution was concentrated. The residue was dissolved in CH_2Cl_2 and washed with HCl (1N), dried, concentrated and the residue was purified by chromatography (13:1 hexane–EtOAc) to give **2-16** (7.4 g, 76%) as a syrup. R_f = 0.47 (9:1 hexane–EtOAc); $[\alpha]_D$ +49.0 (*c* 0.7, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.86–7.65 (m, 4H), 7.55–7.06 (m, 16H), 5.44 (d, 1H, *J* = 5.2 Hz, H-1), 4.70 (d, 1H, *J* = 11.7 Hz, PhCH_2), 4.55 (d, 1H, *J* = 12.3 Hz, PhCH_2), 4.51 (d, 1H, *J* = 12.1 Hz, PhCH_2), 4.38 (d, 1H, *J* = 11.5 Hz, PhCH_2), 4.23 (app t, 1H, *J* = 5.2 Hz, H-2), 4.12 (app t, 1H, *J* = 5.0 Hz, H-3), 3.97 (dd, 1H, *J* = 7.2, 5.7 Hz, H-4), 3.87 (dd, 1H, *J* = 11.2, 4.0 Hz, H-6a), 3.76 (dd, 1H, *J* = 11.2, 6.0 Hz, H-6b), 3.66-3.60 (m, 1H, H-5), 3.50 (s, 3H, CH_3O), 2.76 (dd, *J* = 7.4, 4.6 Hz, 2H, CH_2CH$_3$), 1.35 (t, 3H, *J* = 7.4 Hz, CH$_2$CH_3), 1.11 (s, 9H, (CH_3)$_3$C); ^{13}C NMR (100 MHz, $CDCl_3$): δ 137.7, 137.5, 135.6(7), 135.6(4), 135.5, 134.8, 133.5, 133.4, 129.6, 129.5(9), 129.5(5), 128.4, 128.3, 127.9, 127.8, 127.7, 127.6, 86.3 (C-1), 84.4 (C-2), 83.1 (C-5), 83.0 (C-3), 82.3 (C-4), 72.3 (PhCH_2), 72.0 (PhCH_2), 63.6 (C-6), 59.3 (CH_3O), 26.8 ((CH$_3$)$_3$C), 24.5

(CH_2CH_3), 19.2 ($(CH_3)_3C$), 15.0 (CH_2CH_3); HR ESIMS: *m*/*z* [M+Na^+] calcd for $C_{39}H_{48}O_5SSiNa$: 679.2878. Found: 679.2884.

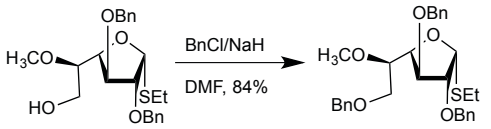
Ethyl 2,3-Di-*O*-benzyl-5-*O*-methyl-1-thio-α-D-galactofuranoside (2-17)



2-16 **2-17**

To a solution of **2-16** (7.64 g, 11.2 mmol) in THF (80 mL) was added 1M TBAF in THF (14.5 mL). The mixture was stirred and after the starting material disappeared, the solution was diluted with CH_2Cl_2 and washed with 1M HCl (satd), brine, dried, and concentrated. Chromatography of the crude product (2:1 hexane–EtOAc) gave **2-17** (4.55 g, 97%) as a syrup. R_f = 0.38 (2:1 hexane–EtOAc; $[\alpha]_D$ +94.0 (*c* 0.6, $CHCl_3$). ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.47–7.19 (m, 10H, *Ph*CH_2), 5.38 (d, 1H, *J* = 4.5 Hz, H-1), 4.69 (d, 1H, *J* = 11.7 Hz, PhCH_2), 4.58 (d, 1H, *J* = 11.6 Hz, PhCH_2), 4.53 (d, 2H, *J* = 11.6 Hz, PhCH_2), 4.22–4.15 (m, 2H, H-2, H-3), 4.05 (dd, 1H, *J* = 7.1, 5.2 Hz, H-4), 3.72 (ddd, 1H, *J* = 11.6, 7.6, 3.9 Hz, H-6a), 3.61-3.56 (m, 1H, H-6b), 3.55-3.53 (m, 1H, H-5), 3.53 (s, 3H, CH_3O), 2.73 (q, 2H, *J* = 7.4 Hz, CH_3CH_2), 2.29 (dd, 1H, *J* = 7.6, 5.7 Hz, OH), 1.32 (t, 3H, *J* = 7.4 Hz, CH_3CH_2); ^{13}C NMR (125 MHz, $CDCl_3$): δ 137.4, 137.2, 128.5, 128.1, 128.0, 86.8 (C-1), 84.2 (C-2), 83.2 (C-4), 82.4 (C-3), 81.9 (C-5), 72.5 (PhCH_2), 72.0 (PhCH_2), 61.0 (C-6), 58.7 (CH_3O), 25.0 (CH_3CH_2), 15.1 (CH_3CH_2). HR ESIMS: *m*/*z* [M+Na^+] calcd for $C_{23}H_{30}O_5SNa$: 441.1706. Found: 441.1700.

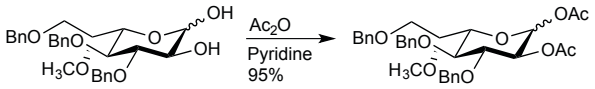
Ethyl 2,3,6-Tri-*O*-benzyl-5-*O*-methyl-1-thio-*α*-D-galactofuranoside (2-18)



2-17 **2-18**

To a solution of **2-17** (4.69 g, 11.2 mmol) in anhydrous DMF (40 mL) at 0 °C were added BnBr (1.7 mL, 1.3 equiv) and NaH (584 mg, 1.5 equiv). The mixture was slowly warmed to rt and stirred overnight before CH_3OH (~1 mL) was added. The solution was concentrated and the residue was purified by chromatography (12:1→8:1 hexane–EtOAc) to give **2-18** (4.74 g, 84%) as a syrup. R_f = 0.72 (2:1 hexane–EtOAc); $[\alpha]_D$ +72.3 (*c* 0.6, CH_2Cl_2). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.50–7.12 (m, 15H, Ph), 5.42 (d, 1H, *J* = 4.8 Hz, H-1), 4.67 (d, 1H, *J* = 11.7 Hz, PhC*H*$_2$), 4.57 (d, 1H, *J* = 11.7 Hz, PhC*H*$_2$), 4.55–4.45 (m, 4H, PhC*H*$_2$), 4.22–4.14 (m, 2H, H-2, H-3), 4.00 (dd, 1H, *J* = 7.1, 4.9 Hz, H-4), 3.74–3.67 (m, 1H, H-5), 3.64 (dd, 1H, *J* = 10.5, 3.9 Hz, H-6a), 3.55 (s, 3H, C*H*$_3$O), 3.51 (dd, 1H, *J* = 10.5, 5.8 Hz, H-6b), 2.75 (dq, 2H, *J* = 7.4, 1.8 Hz, OC*H*$_2$CH$_3$), 1.34 (t, 3H, *J* = 7.4 Hz, CH$_2$C*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$): *δ* 138.5, 138.1, 137.7, 128.7, 128.6, 128.5(6), 128.2, 128.1, 128.0, 127.9, 127.8, 86.8 (C-1), 84.6 (C-2), 83.4 (C-4), 82.7 (C-3), 81.3 (C-5), 73.6 (Ph*C*H$_2$O), 72.6 (Ph*C*H$_2$O), 72.1 (Ph*C*H$_2$O), 70.2 (C-6), 59.3 (*C*H$_3$O), 25.1 (*C*H$_2$CH$_3$), 15.3 (CH$_2$*C*H$_3$). HR ESIMS: *m*/*z* [M+Na$^+$] calcd for $C_{30}H_{36}O_5SNa$: 531.2170. Found: 531.2176.

1,2-Di-*O*-acetyl-3,4,7-tri-*O*-benzyl-6-*O*-methyl-D-*glycero*-*α*/*β*-L-*gluco*-heptopyranose (2-19)

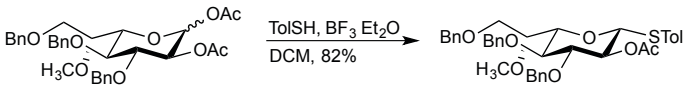


2-3 **2-19**

To a solution of **2-3** (692 mg, 1.4 mmol) in anhydrous pyridine (5 mL) at 0 °C was added Ac2O (1 mL). After stirring for 3 h, the mixture was concentrated. The residue was dissolved in CH_2Cl_2 and then washed with 1N HCl, water, and brine before being dried and concentrated. The residue was purified by chromatography (4:1 hexane–EtOAc) to afford a 2:1 α/β mixture of **2-19** (788 mg, 95%) as a syrup. R_f = 0.50 (2:1 hexane–EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.41–7.23 (m, 22.5H, *Ph*CH_2), 6.26 (d, 1H, J = 3.7 Hz, H-1$_\alpha$), 5.54 (d, 0.5H, J = 8.3 Hz, H-1$_\beta$), 5.16 (dd, 0.5H, J = 9.4, 8.4 Hz, H-2$_\beta$), 5.08 (dd, 1H, J = 10.0, 3.7 Hz, H-2$_\alpha$), 4.96 (d, 1H, J = 10.9 Hz, C*H*$_2$Ph$_\alpha$), 4.92 (d, 0.5H, J = 11.0 Hz, C*H*$_2$Ph$_\beta$), 4.85 (d, 1H, J = 11.4 Hz, C*H*$_2$Ph$_\alpha$), 4.82 (d, 0.5H, J = 11.4 Hz, C*H*$_2$Ph$_\beta$), 4.78 (d, 1H, J = 11.4 Hz, C*H*$_2$Ph$_\alpha$), 4.69 (d, 0.5H, J = 10.6 Hz, C*H*$_2$Ph$_\beta$), 4.68 (d, 1H, J = 10.9 Hz, C*H*$_2$Ph$_\alpha$), 4.55 (d, 0.5H, J = 12.5 Hz, C*H*$_2$Ph$_\beta$), 4.54 (d, 1H, J = 11.8 Hz, C*H*$_2$Ph$_\alpha$), 4.51 (d, 1H, J = 10.9 Hz, C*H*$_2$Ph$_\alpha$), 4.04 (app t, 1H, J = 9.4 Hz, H-3$_\alpha$), 3.99 (dd, 1H, J = 10.0, 1.6 Hz, H-5$_\alpha$), 3.95–3.89 (m, 1.5H, H-4$_\beta$, H-4$_\alpha$), 3.86-3.82 (m, 1H, H-6$_\alpha$), 3.82–3.73 (m, 2.5H, H-3$_\beta$, H-6$_\beta$, H-7$_\alpha$, H-7$_\beta$), 3.69–3.62 (m, 2H, H-5$_\beta$, H-7$_\beta$'), 3.60 (dd, 1H, J = 9.2, 7.0 Hz, H-7$_\alpha$'), 3.46 (s, 3H, OC*H*$_{3\alpha}$), 3.45 (s, 3H, OC*H*$_{3\beta}$), 2.08 (s, 3H, OAc$_\beta$), 2.00 (s, 3H, OAc$_\alpha$), 1.97 (s, 3H, OAc$_\alpha$), 1.95 (s, 3H, OAc$_\beta$); ^{13}C NMR (125 MHz, $CDCl_3$): δ 169.8, 169.4, 169.3, 169.0, 138.4, 138.2, 138.0, 138.0(0), 137.9(9), 128.5(4), 128.5(1), 128.4(8), 128.4(5), 128.4(0), 128.3(7), 127.9, 127.8(6), 127.7(8), 127.7(3), 127.6, 127.5, 127.4, 92.6 (C-1$_\beta$), 89.8 (C-1$_\alpha$), 83.2 (C-3$_\alpha$), 80.5 (C-3$_\beta$), 76.9 (C-4$_\beta$), 76.7 (C-4$_\alpha$), 76.0(3) (C-6$_\beta$), 75.9(8) (C-6$_\alpha$), 75.4 (*C*H$_2$Ph$_\alpha$), 75.2(7) (*C*H$_2$Ph$_\alpha$), 75.2(2) (C-5$_\beta$), 75.2 (*C*H$_2$Ph$_\beta$), 75.0 (*C*H$_2$Ph$_\beta$), 73.7 (*C*H$_2$Ph$_\beta$), 73.4 (*C*H$_2$Ph$_\alpha$), 72.2 (C-5$_\alpha$), 72.1 (C-2$_\beta$), 71.9 (C-2$_\alpha$), 68.9 (C-7$_\beta$), 68.6 (C-7$_\alpha$), 58.7 (O*C*H$_{3\beta}$), 58.6

(O*C*$H_{3\alpha}$), 20.9 (OAc_β), 20.7(7) (OAc_β), 20.7(6) (OAc_α), 20.6(7) (OAc_α). HR ESIMS: *m*/*z* [M+Na^+] calcd for $C_{33}H_{38}O_9Na$: 601.2408. Found: 601.2407.

_p_-Tolyl 2-_O_-Acetyl-3,4,7-tri-_O_-benzyl-6-_O_-methyl-1-thio-_α_/_β_-D-_glycero_-L-_gluco_-heptopyranose (2-20)

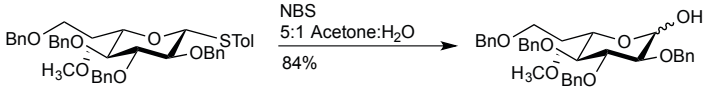


2-19 **2-20**

Compound **2-19** (691 g, 1.2 mol) and p-thiocresol (193 mg, 1.3 equiv) were dissolved in anhydrous CH_2Cl_2 (12 mL). After the addition of BF_3·Et2O (0.30 mL, 2.0 equiv) at 0 °C, the reaction mixture was stirred for 1 h while warming to room temperature. The solution was diluted with CH_2Cl_2, washed with $NaHCO_3$ (sat), dried over Na_2SO_4, and concentrated. The crude product was purified by chromatography (6:1 hexane–EtOAc) to afford **2-20** as a 1:9 α/β mixture **2-20** (628 mg, 82%) as a syrup. R_f = 0.39 (4:1 hexane–EtOAc). Data for major isomer **2-20β** $[\alpha]_D$ –26.5 (*c* 0.5, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.51–7.15 (m, 17H, Ph), 7.03 (d, 2H, *J* = 7.5 Hz, Ph), 5.05 (t, 1H, *J* = 10.1 Hz, H-2), 4.90 (d, 1H, *J* = 11.0 Hz, PhC*H*$_2$), 4.80 (d, 1H, *J* = 11.3 Hz, PhC*H*$_2$), 4.70 (d, 1H, *J* = 11.4 Hz, PhC*H*$_2$), 4.66 (d, 1H, *J* = 11.0 Hz, PhC*H*$_2$), 4.54 (d, 1H, *J* = 10.1 Hz, H-1), 4.43 (d, 1H, *J* = 11.8 Hz, PhC*H*$_2$), 4.39 (d, 1H, *J* = 11.8 Hz, PhC*H*$_2$), 3.92 (app t, 1H, *J* = 8.8 Hz, H-4), 3.83–3.76 (m, 2H, H-6, H-7a), 3.73 (app t, 1H, *J* = 8.8 Hz, H-3), 3.60 (app t, 1H, *J* = 9.7 Hz, H-7b), 3.52-3.45 (m, 1H, H-5), 3.44 (s, 3H, C*H*$_3$O), 2.28 (s, 3H, C*H*$_3$Ph), 2.01 (s, 3H, C*H*$_3$CO); ^{13}C NMR (125 MHz, $CDCl_3$): *δ* 169.7, 138.3(1), 138.2(7), 138.2(5), 137.9, 132.9, 130.1, 129.7, 128.6, 128.5(6), 128.5(0), 127.9(5), 127.8(9), 127.8(5), 127.7(6), 127.6, 87.7 (C-1), 84.9 (C-3), 78.7 (C-5), 77.2 (C-4), 76.3 (C-6), 75.3 (Ph*C*H$_2$), 75.1

(Ph*C*H_2), 73.5 (Ph*C*H_2), 72.3 (C-2), 68.9 (C-7), 58.5 (*C*H_3O), 21.2 (2C, *C*H_3Ph, *C*H_3CO). HR ESIMS: *m*/*z* [M+Na^+] calcd for $C_{38}H_{42}O_7SNa$: 665.2543. Found: 665.2538.

_p_-Tolyl 2,3,4,7-Tetra-_O_-benzyl-6-_O_-methyl-1-thio-D-_glycero-α/β_-L-_gluco_-heptopyranose (2-24)

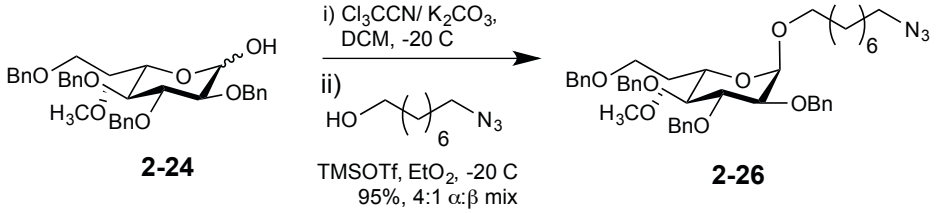


2-2 **2-24**

To a solution of **2-2** (203 mg, 0.29 mmol) in 5:1 acetone–H_2O (5 mL) was added N-bromosuccinimide (157 mg, 3 equiv). After stirring for 2 h, the solution was concentrated and the residue was purified by chromatography (2:1 hexane–EtOAc) to give **2-24** (145 mg, 1:0.3 α/β mixture, 84%) as a white solid. R_f = 0.35 (2:1 hexane–EtOAc); ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.45–7.22 (m, 26H, Ph), 5.12 (d, 1H, *J* = 3.6 Hz, H-1α), 4.98–4.91 (m, 3H, PhC$H_{2\alpha.}$ PhC$H_{2\beta}$), 4.84 (d, 1H, *J* = 10.9 Hz, PhC$H_{2\alpha}$), 4.81 (d, 0.3H, *J* = 11.0 Hz, PhC$H_{2\beta}$), 4.76 (d, 1H, *J* = 12.0 Hz, PhC$H_{2\alpha}$), 4.75 (d, 0.3H, *J* = 11.4 Hz, PhC$H_{2\beta}$), 4.67 (d, 1H, *J* = 11.9 Hz, PhC$H_{2\alpha}$), 4.64 (d, 1H, *J* = 11.1 Hz, PhC$H_{2\alpha}$), 4.63 (d, 0.3H, *J* = 11.1 Hz, PhC$H_{2\beta}$), 4.60 (d, 0.3H, *J* = 12.0 Hz, PhC$H_{2\beta}$), 4.58 (d, 0.3H, *J* = 10.2 Hz, H-1$_\beta$), 4.56 (d, 1H, *J* = 12.0 Hz, PhC$H_{2\alpha}$), 4.51 (d, 1H, *J* = 12.0 Hz, PhC$H_{2\alpha}$), 4.48 (d, 1H, *J* = 12.0 Hz, PhC$H_{2\alpha}$), 4.02–3.94 (m, 2H, H-3$_\alpha$, H-5$_\alpha$), 3.85–3.79 (m, 1.3H, H-6$_\beta$, H-6$_\alpha$), 3.79–3.66 (m, 3.2H, H-7a$_\alpha$, H-7a$_\beta$, H-4$_\alpha$, H-4$_\beta$), 3.72–3.66 (m, 0.6H, H-3$_\beta$, H-7b$_\beta$), 3.64 (dd, 1H, *J* = 6.0, 9.9 Hz, H-7b$_\alpha$), 3.57 (dd, 1H, *J* = 9.5, 3.6 Hz, H-2$_\alpha$), 3.49 (0.3H, td, *J* = 9.8, 4.7 Hz, H-5$_\beta$), 3.46 (s, 3H, CH_3O_α), 3.44 (1H, s, CH_3O_β), 3.36 (dd, 0.3H, *J* = 9.1, 7.7 Hz, H-2$_\beta$); ^{13}C NMR (125 MHz, $CDCl_3$): *δ* 138.6, 138.4, 138.2, 137.8, 128.5, 128.4, 128.3(8), 128.3(6), 128.2, 128.1, 127.9, 127.8(6), 127.8(0), 127.7(5), 127.7(2), 127.6(6), 127.6(4), 127.6(1), 127.6, 97.9 (C-1$_\beta$) 91.3 (C-1$_\alpha$), 84.8 (C-3$_\beta$), 83.8 (C-2$_\beta$), 82.0 (C-3$_\alpha$), 79.9 (C-2$_\alpha$), 76.8

(C-4_α), 76.5 (C-4_β), 75.7 (C-6_β), 75.0 (C-6_α), 74.5 (Ph*C*$H_{2\alpha}$), 73.4 (Ph*C*$H_{2\alpha}$), 73.3(4) (C-5_β), 73.3(2) (Ph*C*$H_{2\alpha}$), 70.4 (C-7_α), 70.2 (C-5_α), 68.9 (C-7_β), 59.0(2) (C*H*$_3O_\alpha$), 59.0(0) (C*H*$_3O_\beta$). HR ESIMS: *m*/*z* [M+Na^+] calcd for $C_{36}H_{40}O_7Na$: 607.2666. Found: 607.2657.

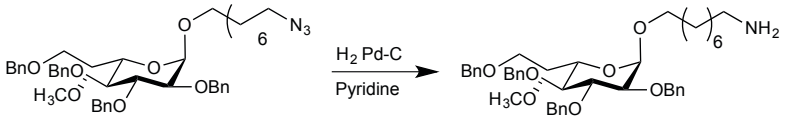
8-Azido-octyl 2,3,4,7-Tetra-*O*-benzyl-6-*O*-methyl-D-*glycero*-*α*-L-*gluco*-heptopyranoside (2-26)



To a solution of **2-24** (22 mg, 0.037 mmol) in anhydrous CH_2Cl_2 (1 mL) was added trichloroacetonitrile (40 μL, 10 equiv) and K_2CO_3 (26 mg, 5 equiv). After stirring for 4 h, the solution was filtered, concentrated and the residue was purified by chromatography (6:1 hexane–EtOAc with 0.5% Et_3N) to give **2-25** (12.7 mg major spot (presumably β), 4.9 mg minor spot (presumably α), 66%) as a yellow syrup. Rf (major isomer) = 0.71 (4:1 hexane–EtOAc). To a solution of **2-25** β-anomer (12.7 mg, 0.017 mmol), 1-azidooctanol (2 mg, 1.1 equiv) and 4 Å molecular sieves (10 mg) in Et2O (0.5 mL) was added 5% TMSOTf in Et_2O until the pH of the solution was 5. After stirring for 45 min, the solution was filtered, concentrated and the residue was purified by chromatography (9:1 hexane–EtOAc) to give **2-26** (12.3 mg, separable 4:1 α:β mixture, 95%) as a white solid. Data for α-isomer: R_f = 0.34 (9:1 hexane–EtOAc); $[\alpha]_D$ –24.5 (*c* 0.7, $CHCl_3$); ^{1}H NMR (500 MHz, CD_3OD) δ 7.33 (m, 20H, Ph), 5.04 (d, 1H, *J* = 10.9 Hz, PhC*H*$_2$), 4.99 (d, 1H, *J* = 11.0 Hz, PhC*H*$_2$), 4.84 (d, 1H, *J* = 10.9 Hz, PhC*H*$_2$), 4.79 (d, 1H, *J* = 12.1 Hz, PhC*H*$_2$), 4.74 (d, 1H, *J* = 3.6 Hz, H-1), 4.66 (d, 1H, *J* = 12.1 Hz, PhC*H*$_2$), 4.64 (d, 1H, *J*

= 11.0 Hz, PhCH_2), 4.55 (s, 2H, PhCH_2), 4.04 (app t, 1H, J = 9.4 Hz, H-3), 3.86–3.79 (m, 2H, H-6, H-7a), 3.79–3.72 (m, 2H, H-4, H-7b), 3.65 (dd, 1H, J = 9.5, 6.1 Hz, H-5), 3.61–3.54 (m, 2H, H-2, octyl OCH_2), 3.47 (s, 3H, CH_3O), 3.33 (td, 1H, J = 9.6, 6.7 Hz, octyl OCH_2), 3.26 (t, 2H, J = 7.0 Hz, octyl NCH_2), 1.66–1.53 (m, 2H, octyl), 1.40–1.21 (m, 10H, octyl); ^{13}C NMR ($CDCl_3$) δ 138.9, 138.5, 138.4, 138.1, 128.5, 128.4, 128.3(7), 128.1, 128.0, 127.8, 127.5, 96.8 (C-1), 82.3 (C-3), 80.0 (C-2), 77.2 (C-4), 76.52 (C-6), 75.6 (PhCH$_2$), 75.0 (PhCH$_2$), 73.4 (PhCH$_2$), 73.10 (PhCH$_2$), 70.0 (C-5), 69.9 (C-7), 67.9 (octyl OCH$_2$), 59.0 (CH_3O) 51.5 (octyl NCH$_2$), 29.7 (octyl), 29.4 (octyl CH$_2$), 29.1 (octyl CH$_2$), 28.9 (octyl CH$_2$), 26.7 (octyl CH$_2$), 26.1 (octyl CH$_2$). HR ESIMS: m/z [M+Na$^+$] calcd for $C_{44}H_{55}N_3O_7Na$: 760.3932. Found 760.3927.

8-Trifloroacetamido-octyl 2,3,4,7-Tetra-*O*-benzyl-6-*O*-methyl-D-*glycero*-α-L-*gluco*-heptopyranoside (2-28)

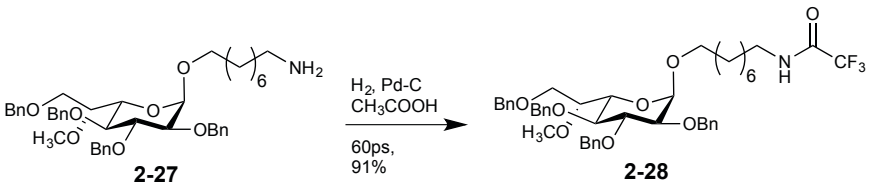


2-26 **2-27**

To a solution of **2-26** (12.5 mg, 0.017 mmol) in pyridine (2 mL) was added 10 mg Pd–C. The mixture was purged with 1 atm H_2 (g) for 2 h and filtered through Celite. Partial concentration yielded a clear colourless liquid to which were added trifluoroacetic anhydride (10 µL, 2 equiv) and pyridine (12 µL, 4 equiv). After stirring for 36 h, the solution was concentrated and the residue was purified by chromatography (4:1 hexane– EtOAc) to give **2-28** (13 mg, 95% over 2 steps) as a solid. R_f = 0.23 (4:1 hexane–EtOAc); $[\alpha]_D$ –2.5 (*c* 0.5, CH_2Cl_2); ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.44–7.20 (m, 20H, Ph), 6.30–6.12 (br s, 1 H, NH), 5.02 (d, 1H, J = 10.8 Hz, PhCH_2), 4.97 (d, 1H, J = 10.9 Hz, PhCH_2), 4.82 (d, 1H, J = 10.8 Hz, PhCH_2), 4.76 (d, 1H, J = 12.1 Hz, PhCH_2), 4.72 (d, 1H, J = 3.6 Hz, H-1), 4.65 (d, 1H, J = 12.1 Hz, PhCH_2), 4.64 (d, 1H, J = 10.9

Hz, PhCH_2), 4.52 (s, 2H, PhCH_2), 4.04 (app t, 1H, J = 9.4 Hz, H-3), 3.72–3.84 (m, 4H, H-6, H-5, H-7a, H-4), 3.67 (dd, 1H, J = 9.5, 6.2 Hz, H-7b), 3.58–3.52 (m, 2H, octyl OCH_2, H-2), 3.44 (s, 3H, OCH_3), 3.38–3.29 (m, 3H, octyl OCH_2, octyl NCH_2), 1.55 (m, 2H, octyl), 1.37–1.13 (m, 10H, octyl). ^{13}C NMR ($CDCl_3$) δ, 138.9, 138.4, 138.3(6), 138.1, 128.5, 128.4, 128.3(7), 128.0, 127.9, 127.8(5), 127.8(2), 127.7(5), 127.6, 127.5(6), 127.4(7), 96.8 (C-1), 82.3 (C-3), 80.0 (C-2), 76.5 (C-4), 76.4(9) (C-6), 75.6 (PhCH_2), 75.1 (PhCH_2), 73.4 (PhCH_2), 73.1 (PhCH_2), 70.0 (C-5), 69.9 (C-7), 67.9 (octyl OCH$_2$), 59.0 (CH_3O), 40.0 (octyl NCH$_2$), 29.3 (octyl CH$_2$), 29.2(7) (octyl CH$_2$), 29.1 (octyl CH$_2$), 29.0 (octyl CH$_2$), 26.6 (octyl CH$_2$), 26.1 (octyl CH$_2$). HR ESIMS: *m*/*z* [M+Na$^+$] calcd for $C_{46}H_{56}F_3NO_8Na$: 830.3850. Found 830.3846.

8-Trifloroacetamido-octyl 6-*O*-methyl-D-*glycero*-*α*-L-*gluco*-heptopyranoside (2-29)



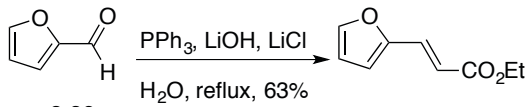
To a solution of **2-28** (18 mg, 0.024 mmol) in HOAc (1 mL) was added 12 mg Pd–C. The reaction was shaken under 60 psi H_2 (g) for 60 h, filtered through Celite, and concentrated to give **2-29** (9.4 mg, 91%) as a white solid. Rf = 0.69 (6:1 CH_2Cl_2–CH_3OH)$[\alpha]_D$ –60.0 (*c* 0.3, CH_3OH); ^{1}H NMR (500 MHz, CD_3OD) δ 4.72 (d, 1H, J = 3.7 Hz, H-1), 3.77 (dd, 1H, J = 10.3, 5.9 Hz, octyl OCH_2), 3.70–3.57 (m, 5H, H-7a, H-6, H-5, octyl OCH_2, H-3), 3.53 (s, 3H, CH_3O), 3.51–3.43 (m, 1H, H-4), 3.41–3.32 (m, 2H, H-7b, H-2), 3.28–3.23 (m, 2H, octyl NCH_2), 1.78–1.44 (m, 2H, octyl), 1.44–1.15 (m, 10H, octyl); ^{13}C NMR (125 MHz, CD_3OD): δ 100.2 (C-1), 80.5 (C-5), 75.4 (C-3), 73.5 (C-2), 72.0 (C-6), 71.1 (C-4), 69.1 (C-7), 62.5 (octyl OCH$_2$), 60.3 (CH_3O), 40.6 (octyl NCH$_2$), 30.6 (octyl), 30.5 (octyl CH$_2$), 30.2 (octyl CH$_2$), 29.8 (octyl CH$_2$),

27.8 (octyl *C*H_2), 27.3 (octyl *C*H_2). HR ESIMS: *m*/*z* [M+Na^+] calcd for $C_{18}H_{32}F_3NO_8Na$: 470.1972. Found 470.1973.

Synthesis of BSA conjugate (2-30).

To a solution of 2-1 (2.5 mg, 0.010 mmol) in 20 mM Na2PO4 buffer at pH 7.5(0.6 mL) was added imm-Link^{TM} Amine Modifier reagent (60 μL). The solution was mixed gently and added to imm-LinkTM BSA mix (2 mg), resuspended, and left for 16 h. The reaction mixture was then dialysed against deionized water mixed with dialysis buffer (1 L, changed twice, 4h each). The dialyzed conjugate was lyophilized to give the BSA conjugate of 2-1, corresponding to the incorporation of 4 ligands per molecule of BSA. MALDI-MS (positive mode, matrix sinapinic acid, H2O): imm-Link^{TM}BSA conjugate 1 (72528); imm-LinkTM BSA (71076).

Ethyl (*E*)-3-(2-furanyl)prop-2-enoate (2-43)

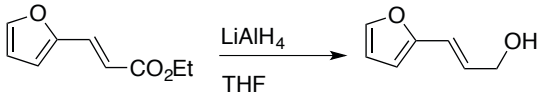


2-36 **2-43**

To a solution of **2-36** (4.3 mL, 52 mmol) in H_2O (260 mL) was added bromophosphonoacetate(10.3 mL, 93.6 mmol), triphenylphosphine (26 g, 93.6 mmol), LiOH (4.6 g, 109.2 mmol), and LiCl (13.3g, 312 mmol). The mixture was heated at refluxed in air for 40 min and then cooled to room temperature. The mixture was then extracted with EtOAc (3 x 60 mL) and the combined organic extracts were washed with brine, dried over Na_2SO_4 filtered, and concentrated under reduced pressure. The residue was then purified by column chromatography over silica gel (7:1 hexane–EtOAc) to give **2-43** (5.4 g, 63%) in a 99:1 mixture of *E*:*Z* isomers as an clear colourless oil. The product was confirmed by matching ^{1}H NMR

spectroscopic data to literature reports.[37] ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.47 (d, 1H, *J* =1.6 Hz), 7.42 (d, 1H, *J* =15.6Hz), 6.60 (d, 1H, *J* = 3.2Hz), 6.46 (dd, 1H, *J* = 3.2 Hz, 1.6 Hz, 1H), 6.31 (d, 1H, *J* =15.6 Hz), 4.24 (q, 2H, *J* = 7.2 Hz), 1.32 (3H, t, *J* = 7.2 Hz).

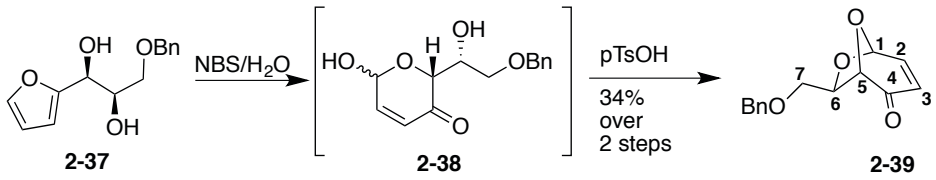
(*E*)-3-(2-Furfuryl)prop-2-en-1-ol (2-44)



2-43 **2-44**

To a solution of **2-43** (1.43 g, 7.93 mmol) in THF (35 mL) at 0 °C, was added 1M $LiAlH_4$ in THF (6.3 mL). After 1 h, a saturated solution of NH_4OH was added and then the mixture was filtered through Celite. The organic phase was separated, washed with brine, and dried over Na_2SO_4, and concentrated under reduced pressure. The residue was purified by chromatography (2:1 hexane–EtOAc) to give **2-44** (0.64 g, 65%) as a yellow oil. R_f = 0.33 (2:1 hexane–EtOAc). The product was confirmed by matching ^{1}H NMR spectroscopic data to literature reports.[29] ^{1}H NMR ($CDCl_3$, 300 MHz): δ 7.48 (d, 1H, *J* = 1.6 Hz), 7.43 (d, 1H, *J* = 15.9 Hz), 6.60 (d, 1H, *J* = 3.4 Hz), 6.46 (dd, 1H, *J* = 3.4, 1.6 Hz), 6.32 (d, 1H, *J* = 15.9 Hz), 4.25 (q, 2H, *J* = 7.1 Hz), 1.32 (t, 3H, *J* = 7.1 Hz).

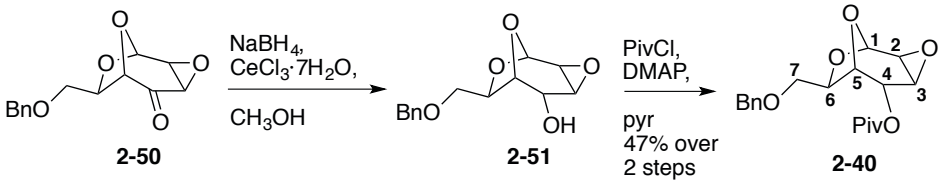
(+)-(1*S*,5*S*,6*R*)-7-*O*-benzyl-6,8-dioxabicyclo[3.2.1]oct-2-en-4-one (2-39)



2-37 2-38 2-39

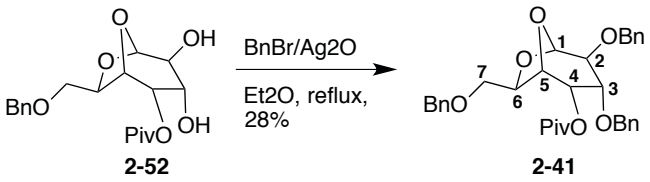
To a solution of **2-37**(2.08 g, 8.40 mmol) in 4:1 THF:H_2O (25 mL) at 0 °C was added $NaOAc \cdot 3H_2O$ (1.14 g, 8.40 mmol), $NaHCO_3$ (1.41 g, 16.8 mmol), and *N*-bromosuccinimide (1.50 g, 8.40 mmol). After dissappearance of the starting material, the a saturated solution of $NaHCO_3$ was added. The mixture was extracted with EtOAc (3 x 30 mL) and the combined organic layers were washed with brine, dried over Na_2SO_4, and filtered. Concentration of the organic layer yielded crude, **2-38**, which was dissolved in benzene (20 mL), before *p*-toluenesulfonic acid was added (0.10 g, 5%/wt). The solution was heated at reflux in a Dean–Stark apparatus overnight. After cooling, the mixture was diluted with EtOAc (100 mL) and the solution was washed with sat. aq $NaHCO_3$ (3 x 10 mL) and brine (10 mL), and dried over $MgSO_4$. After evaporation of the solvent under reduced pressure, the residue was purified by column chromatography (8:1 hexane–EtOAc) to give **2-39** (695 mg, 34% over two steps) as a yellow oil. The product was confirmed by matching ^{1}H NMR spectroscopic data to literature reports.[38] ^{1}H NMR ($CDCl_3$, 500 MHz): δ 7.30–7.19 (m, 5H), 7.05 (dd, 1H, J = 9.8, 3.1 Hz, H-2), 6.01 (d, 1H, J = 9.8 Hz, H-3), 5.75 (d, 1H, J = 3.1 Hz, H-1), 4.52 (m, 3H, PhCH_2), 3.96–3.93 (m, 1H, H-6), 3.56 (dd, 1H, J = 9.8, 6.1 Hz, H-7a), 3.46 (dd, 1H, J = 9.8, 6.8 Hz, H-7b). ^{13}C NMR ($CDCl_3$, 125 MHz): δ = 194.0, 147.3, 137.6, 128.5, 127.9, 127.8, 126.9, 96.6, 81.8, 73.5, 73.0, 69.8.

(+)-(1*S*,5*S*,6*R*)-7-*O*-benzyl-6,8-dioxabicyclo[3.2.1]oct-2,3-epoxy-4-pivaloyl-ane (2-40)



To a solution of **2-50** (695 mg, 2.82 mmol) in MeOH (8 mL) at 0 °C, was added $CeCl_3 \cdot 7H_2O$ (1.27 g, 3.38 mmol, 1.2 equiv) and then $NaBH_4$ (129 mg, 3.38 mmol, 1.2 equiv). After 0.5 h, the reaction mixture was dilluted with EtOAc, washed with H_2O, brine, dried over Na_2SO_4, then filtered, and the organic layer was concentrated under reduced pressure to give **2-49**, which was used without further purification. To a solution of **2-40** (161 mg crude, 0.609 mmol) in pyridine (3 mL) was added PivCl (150 μL, 2 equiv) and 4-dimethylaminopyridine (16 mg, 0.10 equiv) and the mixture was stirred overnight. The solution was diluted with CH_2Cl_2, washed with aq satd $NaHCO_3$, dried over Na_2SO_4, and concentrated. The crude product was purified by chromatography (9:1 hexane–EtOAc) to afford **2-39** (100 mg, 47% over two steps) as a syrup. Data for **2-50**: ^{1}H NMR ($CDCl_3$, 500 MHz): ^{1}H NMR ($CDCl_3$, 500 MHz): δ 7.23–7.12 (m, 5H), 5.63 (s, 1H, H-1), 4.60 (d, 1H, *J* = 6.7 Hz, Ph*CH*$_2$), 4.52 (ddd, 1H, *J* = 1.7, 5.3, 7.0 Hz, H-6), 4.24–4.20 (m, 1H, H-3), 4.19-4.17 (m, 1H, H-5), 3.59 (dd, 1H, *J* = 7.0, 5.3 Hz, H-7a), 3.45 (dd, 1 H, *J* = 7.0, 5.3 Hz, H-7b), 3.09–3.04 (m, 1H, H-2), 2.58 (app d, 1H, *J* = 3.8 Hz, H-4). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 137.6, 128.5, 127.9, 127.8, 97.7 (C-1), 75.9 (C-5), 73.6 (Ph*C*H$_2$), 72.8 (C-4), 70.7 (C-7), 63.2 (C-6), 52.9 (C-2), 48.9 (C-3). Data for **2-39**: ^{1}H NMR ($CDCl_3$, 500 MHz): δ 7.23–7.12 (m, 5H), 5.53 (s, 1H, H-1), 4.72 (dd, *J* = 2.6, 2.6 Hz, 1H, H-4), 4.68–4.40 (m, 4H, Ph*CH*$_2$, H-3, H-5), 4.12 (app t, *J* = 6.4, 1H, H-4), 3.52 (dd, 1H, *J* = 9.5, 5.9 Hz, H-7a), 3.44 (dd, 1H, *J* = 9.5, 5.9 Hz, H-7b), 3.38 (dd, *J* = 2.6, 2.6 Hz, 1H, H-2), 2.58 (app d, 1H, *J* = 3.8 Hz, H-4), 1.25 (s, 9H, –OC(*C*H$_3$)$_3$).

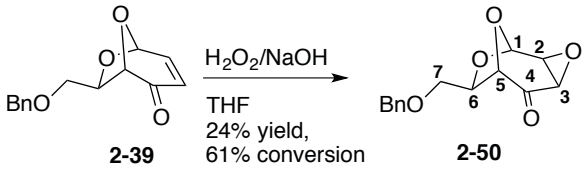
(+)-(1*S*,5*S*,6*R*)-7-*O*-benzyl-6,8-dioxabicyclo[3.2.1]oct-2,3-di-*O*-benzyl-4-pivaloyl-ane (2-41)



2-52 **2-41**

To a solution of **2-52** (95 mg, 0.260 mmol) in Et_2O (5 mL), freshly prepared Ag_2O (180 mg, 3.0 equiv) and BnBr (124 μL, 4.0 equiv) were added. After stirring for 2 d, the mixture was filtered through Celite and the filtrate was concentrated under reduced pressure. The concentrate was dissolved in EtOAc, and washed with brine, before the organic layer was dried over Na_2SO_4, and concentrated under reduced pressure. Purification by chromatography (10:1 hexane–EtOAc) gave **2-40** (40 mg, 28% yield) as a white solid, and unreacted **2-41** (50 mg). R_f = 0.35 (9:1 hexane–EtOAc); ^{1}H NMR ($CDCl_3$, 500 MHz): δ 7.19–7.14 (m, 19H), 5.40 (dd, 1H, *J* = 5.3, 1.3 Hz, H-4, 5.33 (app t, 1H, *J* = 1.5 Hz, H-1), 4.71 (app t, 1H, *J* = 6.6, H-6) 4.82–4.48 (m, 6H, 3 x Ph*CH*$_2$), 4.31 (d, 1H, *J* = 3.7 Hz, H-5), 4.05 (app t, 1H, *J* = 5.1 Hz, H-3), 3.48 (dd, 1H, *J* = 9.8, 7.1 Hz, H-7a), 3.38 (dd, 1H, *J* = 9.8, 7.1 Hz, H-7b), 3.52 (s, 1H, H-2), 2.40 (s, 9H, 3 x CH_3).

(+)-(1*S*,5*S*,6*R*)-7-*O*-benzyl-6,8-dioxabicyclo[3.2.1]oct-2,3-epoxy-4-one (2-50)

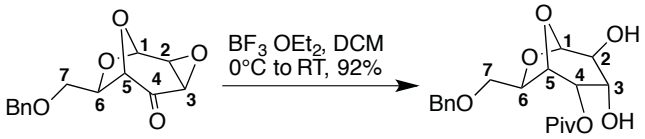


2-39 **2-50**

To a solution of **2-39** (613 mg, 2.49 mmol) in THF (25 mL) at 0 °C was added a 30% solution of H_2O_2 in H_2O (292 μL, 1.5 equiv) and 1M solution of NaOH in H_2O (1.25 mL). After 3 h, an additional amount of H_2O_2 in H_2O (150 μL, 0.75 equiv) and 1M NaOH (0.625 mL, 0.75 equiv) was added. After stirring overnight, the reaction mixture was extracted with EtOAc, washed with brine, and dried over Na_2SO_4. After evaporation of the solvent under reduced pressure, the

residue was purified by column chromatography (3:1 hexane–EtOAc) to give **2-50** (159 mg, 24% yield, 61% conversion) as a white solid, R_f = 0.29, 1:1 hexane–EtOAc) and unreacted starting material **2-39** (390 mg). R_f = 0.78, 1:1 hexane–EtOAc)

(+)-(1*S*,5*S*,6*R*)-7-*O*-benzyl-6,8-dioxabicyclo[3.2.1]oct-2,3-dihydroxy-4-pivaloyl-ane (2-52)



2-50 **2-52**

To a solution of **2-40** (103 mg, 0.296 mol) in wet CH_2Cl_2 (1.5 mL) at –65 °C was added $BF_3{\cdot}OEt_2$ (47 μL, 1.3 equiv). After 2.5 h, an additonal amount of $BF_3{\cdot}OEt_2$ (47 μL, 1.3 equiv) was added. After 1 h, the reaction mixture was neutralized with triethylamine. After concentration of the reaction mixture, the crude product was purified by column chromatography (1:1 hexane–EtOAc) to give **2-52** (99 mg, 92%) as a yellow oil. R_f = 0.31, 1:1 hexane–EtOAc; ^{1}H NMR ($CDCl_3$, 500 MHz): δ 7.20–7.13 (m, 5H), 5.40 (app s, 1H, H-1), 5.10 (d, *J* = 8.9 Hz, 1H, H-4), 4.64–4.51 (m, 4H, Ph*CH*$_2$, H-5, H-6), 4.31 (s, 1H, H-3), 3.70 (br s, 1H, H-2), 3.52 (dd, 1H, *J* = 9.6, 7.3 Hz, H-7a), 3.44 (dd, 1H, *J* = 9.6, 7.3 Hz, H-7b), 1.30 (s, 9H, 3xCH_3). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 178.8 (C=O –Piv), 137.9, 128.5, 128.4, 127.8, 127.8, 127.6, 101.2 (C-1), 76.2 (C-6), 74.1 (*C*H_2Ph), 73.4 (C-3), 72.8 (C-2), 70.7 (C-7), 70.2 (C-4), 64.4 (C-5), 39.1 (–O*C*$(CH_3)_3$), 27.1 (–OC(*C*$H_3)_3$).

2.8 References

(1) Kosma, P. *Curr. Org. Chem.* **2008**, *12* (12), 1021–1039.

(2) Pakulski, Z.; Poly, F.; Dorabawila, N.; Guerry, P.; Monteiro, M. A. **2014**, 1818–1845.

(3) Michael, F. St.; 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* (21), 5119–5136.

(4) Ohara, T.; Adibekian, A.; Esposito, D.; Stallforth, P.; Seeberger, P. H. *Chem. Commun.* **2010**, *46* (23), 4106–4108.

(5) Hansson and S. Oscarson, J. *Curr. Org. Chem.* **2015**, *4* (5).

(6) Oscarson, S. In *Glycoscience Synthesis of Oligosaccharides and Glycoconjugates*; Driguez, H., Thiem, J., Eds.; Topics in Current Chemistry; Springer Berlin Heidelberg: Berlin, Heidelberg, 1997; Vol. 186, pp 171–202.

(7) Tikad, A.; Vincent, S. P. In *Modern Synthetic Methods in Carbohydrate Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA, 2013; pp 29–65.

(8) Morrison, J. D.; Mosher, H. S. *Asymmetric organic reactions*; Prentice-Hall, 1971.

(9) Richards, M. R.; Lowary, T. L. *ChemBioChem* **2009**, *10* (12), 1920–1938.

(10) Arasappan, A.; Fraser-Reid, B. *Tetrahedron Lett.* **1995**, *36* (44), 7967–7970.

(11) Velty, R.; Benvegnu, T.; Gelin, M.; Privat, E.; Plusquellec, D. *Carbohydr. Res.* **1997**, *299* (1–2), 7–14.

(12) D'Accorso, N. B.; Thiel, I. M. E.; Schüller, M. *Carbohydr. Res.* **1983**, *124* (2), 177–184.

(13) Kohn, P.; Samaritano, R. H.; Lerner, L. M. *J. Am. Chem. Soc.* **1965**, *87* (23), 5475–5480.

(14) Bordoni, A.; de Lederkremer, R. M.; Marino, C. *Tetrahedron* **2008**, *64* (8), 1703–1710.

(15) Completo, G. C.; Lowary, T. L. *J. Org. Chem.* **2008**, *73* (12), 4513–4525.

(16) Green, J. W.; Pacsu, E. *J. Am. Chem. Soc.* **1937**, *59* (7), 1205–1210.

(17) McAuliffe, J. C.; Hindsgaul, O. *J. Org. Chem.* **1997**, *62* (5), 1234–1239.

(18) Madhusudan, S. K.; Misra, A. K. *Carbohydr. Res.* **2005**, *340* (3), 497–502.

(19) Jayasuiya, A.; Peng, W.; Guillotin, L.; Lowary, T. In *Carbohydrate Chemistry*; Carbohydrate Chemistry; CRC Press, 2015; pp 143–146.

(20) WOLFROM, M. L.; YOSIZAWA, Z.; JULIANO, B. O. *J. Org. Chem.* **1959**, *24* (10), 1529–1530.

(21) Bai, Y.; Lowary, T. L. *J. Org. Chem.* **2006**, *71* (26), 9672–9680.

(22) Lee, D.; Williamson, C. L.; Chan, L.; Taylor, M. S. *J. Am. Chem. Soc.* **2012**, *134* (19), 8260–8267.

(23) Peng, W.; Jayasuriya, A. B.; Imamura, A.; Lowary, T. L. *Org. Lett.* **2011**, *13* (19), 5290–5293.

(24) Konradsson, P.; Udodong, U. E.; Fraser-Reid, B. *Tetrahedron Lett.* **1990**, *31* (30), 4313–4316.

(25) Urban, F. J.; Moore, B. S.; Breitenbach, R. *Tetrahedron Lett.* **1990**, *31* (31), 4421–4424.

(26) Schmidt, R. R.; Michel, J.; Roos, M. *Liebigs Ann. der Chemie* **1984**, *1984* (7), 1343–1357.

(27) Stick, R. V. Carbohydrates: The Essential Molecules of Life, 2nd, Ed .; Academic Press: London, 2001; pp 113–178.

(28) Gadikota, R. R.; Callam, C. S.; Appelmelk, B. J.; Lowary, T. L. *J. Carbohydr. Chem.* **2003**, *22* (3-4), 149–170.

(29) Takeuchi, M.; Taniguchi, T.; Ogasawara, K. *Synthesis (Stuttg).* **1999**, *1999* (02), 341–354.

(30) Ahmed, M. M.; Berry, B. P.; Hunter, T. J.; Tomcik, D. J.; O'Doherty, G. A. *Org. Lett.* **2005**, *7* (4), 745–748.

(31) Harris, J. M.; Keränen, M. D.; Nguyen, H.; Young, V. G.; O'Doherty, G. A. *Carbohydr. Res.* **2000**, *328* (1), 17–36.

(32) Babu, R. S.; O'Doherty, G. A. *J. Am. Chem. Soc.* **2003**, *125* (41), 12406–12407.

(33) Shan, M.; Xing, Y.; O'Doherty, G. A. *J. Org. Chem.* **2009**, *74* (16), 5961–5966.

(34) Tsunoda, T.; Yamamiya, Y.; Itô, S. *Tetrahedron Lett.* **1993**, *34* (10), 1639–1642.

(35) Tsunoda, T.; Nagaku, M.; Nagino, C.; Kawamura, Y.; Ozaki, F.; Hioki, H.; Itô, S. *Tetrahedron Lett.* **1995**, *36* (14), 2531–2534.

(36) Cho, B. T. *Chem. Soc. Rev.* **2009**, *38* (2), 443–452.

(37) Wu, J.; Yue, C. *Synth. Commun.* **2006**, *36* (20), 2939–2947.

(38) Taniguchi, T. *Synthesis (Stuttg).* **1999**, *1999* (08), 1325–1330.

3 Synthesis of Donor and Acceptor Substrates For Testing With Putative Gal*f*NAc Transferases.

This chapter will discuss work towards the synthesis of the potential donor and acceptor substrates (**1-1** to **1-5**) to be tested with the putative Gal*f*NAc transferases (Figure 3-1, shown below). It will begin with two routes toward the donor, UDP-Gal*f*NAc (Figure 3-1, **1-1**). The first involves a small scale enzymatic synthesis through epimerization of the C-4 hydroxyl group of UDP-Glc*p*NAc to UDP-Gal*p*NAc using GalE, and then a conversion to the furanose ring form by using the pyranose–furanose mutase UNGM (cj1439, Scheme 3-1) characterized by our group.[1] The second route involves a preparative chemical synthesis,[2] in which Gal*f*NAc-1-phosphate (prepared by former graduate student Ryan Snitynsky) is subjected to a coupling reaction with uridine-5′-monophosphate (UMP) through the formation of an imidazole-activated UMP.

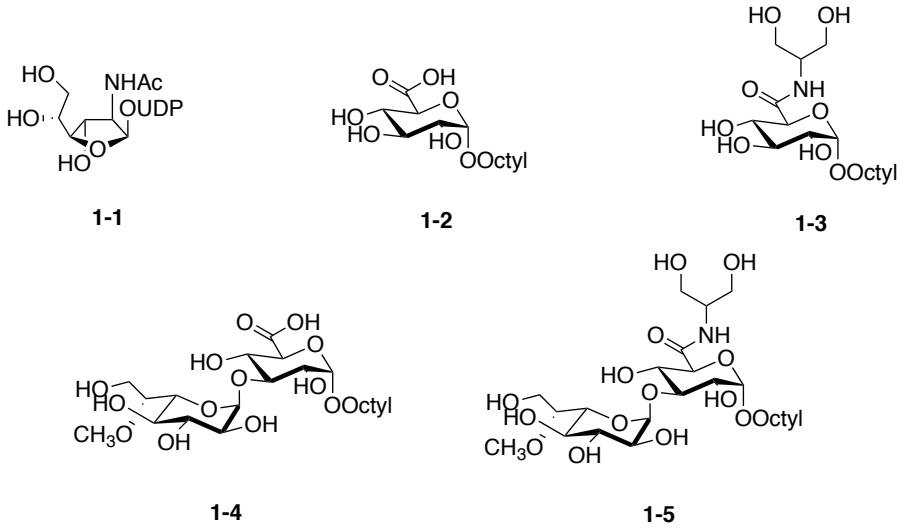
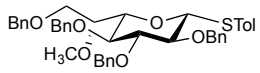


Figure 3-1: Potential donor and acceptor substrates for putative Gal*f*NAc transferases.

After presentation of the routes used to access the UDP-Gal*f*NAc donor, steps toward the chemical synthesis of acceptors **1-2** to **1-5** will be discussed. With heptose donor **2-2** in hand

(Figure 3-2), various glycosylation strategies toward completing the total synthesis of compounds **1-4** and **1-5** were explored. With these substrates, further biochemical work could elucidate the Gal*f*NAc transferase among the putative GTs (Chapter 4).



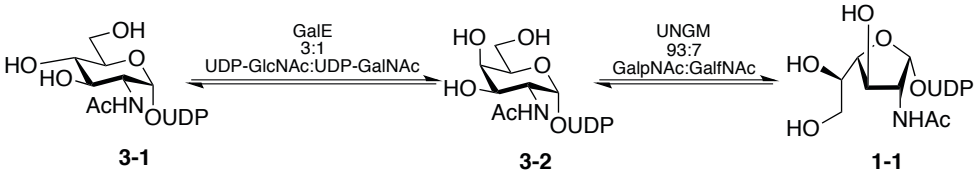
2-2

Figure 3-2: Heptose donor **2-2**, used for the synthesis of acceptors **1-4** and **1-5**.

3.1 Synthesis of UDP-Gal*f*NAc (1-1)

3.1.1 Enzymatic Approach

One route toward the synthesis of UDP-Gal*f*NAc is through the enzymatic route outlined in Scheme 3-1. It relies on the enzymatic conversion of UDP-Glc*p*NAc into UDP-Gal*p*NAc by the epimerase GalE and then isomerization of UDPGal*p*NAc into UDPGal*f*NAc using the pyranose–furanose mutase UNGM. Though both UDP-Glc*p*NAc and UDP-Gal*p*NAc are commercially available, the latter is significantly more expensive ($824.00 CAD /25 mg vs $67.00 CAD/25 mg, Sigma Aldrich), making the preparative scale synthesis using UDP-Gal*p*NAc as starting material much less cost effective.



3-1 **3-2** **1-1**

Scheme 3-1: Enzymatic route towards UDP-Gal*f*NAc.

3.1.1.1 Expression, evaluation and mechanism of GalE.

GalE is encoded for by the Gne gene (*cj1431*) in *C. jejuni* 11168, and is a bifunctional UDP-Glc*p*NAc/Glc*p* 4-epimerase.[3] To explore the possibility of using the route outlined in Scheme 3-1, a clone of GalE-MBP (UDP-Gal*p*NAc-4-epimerase–maltose binding protein fusion), obtained from Warren Wakarchuk's laboratory at Ryerson University, was overexpressed and purified. A protein of the expected molecular weight (76 kDa) was visible after purification on an amylose column.

The enzyme activity was confirmed using capillary electrophoresis with a photodiode array detector (CE-PDA), where two peaks in a 3:1 ratio indicated enzymatic activity (3 UDP-Glc*p*NAc : 1 UDP-Gal*p*NAc) (Figure 3-3). This matched closely results reported in the literature for the equilibrium for this enzyme.[3]

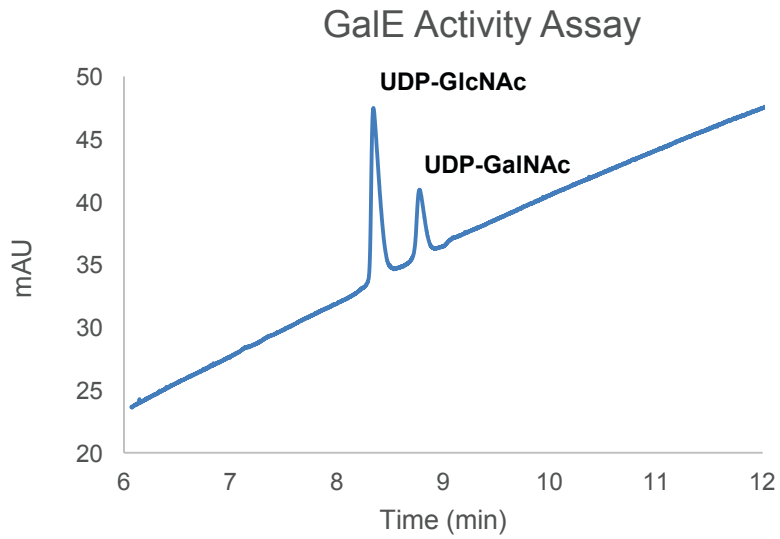


Figure 3-3: GalE activity monitored by CE-PDA.

Interestingly, the epimerization by GalE is thought to go through the UDP-4-keto-GlcNAc intermediate, as outlined in the mechanism in

Figure 3-4.[4]

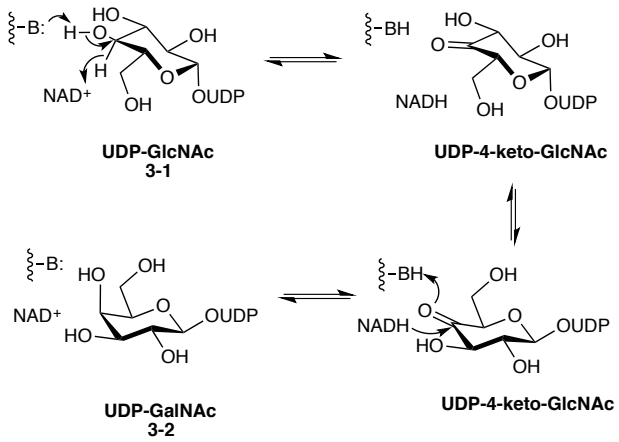


Figure 3-4: Mechanism of GalE through a 4-keto intermediate.

3.1.1.2 Expression, evaluation and mechanism of UNGM

A key step in the formation of UDP-Gal*f*NAc in nature is the transformation mediated by UNGM, a subclass of UDP-galactopyranose mutase (UGM) flavoproteins that catalyzes the reversible interconversion of UDP-Gal*p*NAc (3-2) to UDP-Gal*f*NAc (1-1,

Figure 3-5). In these enzymes, the flavin adenine dinucleotide (FAD) co-factor is non-covalently bound to the enzyme, and is directly involved in catalysis; furthermore, it must be in the reduced form for the enzyme to be active.[5] Previous work on UGM has shown that the reduced FADH acts as a nucleophile to displace the anomeric $UDPO^-$ in **3-2** and form a new covalent intermediate **3-3**.[6] Formation of the covalently-bound acyclic iminium ion **3-4** from **3-3**, involves breakage of the O5–C1 bond. Intermediate **3-4** is then cyclized through an exo-trig ring closure involving O-4 to form the furanose ring **3-5**. Displacement of the FAD adduct from **3-5** by $UDPO^-$ gives UDP-Gal*f*NAc **1-1** as the final product.

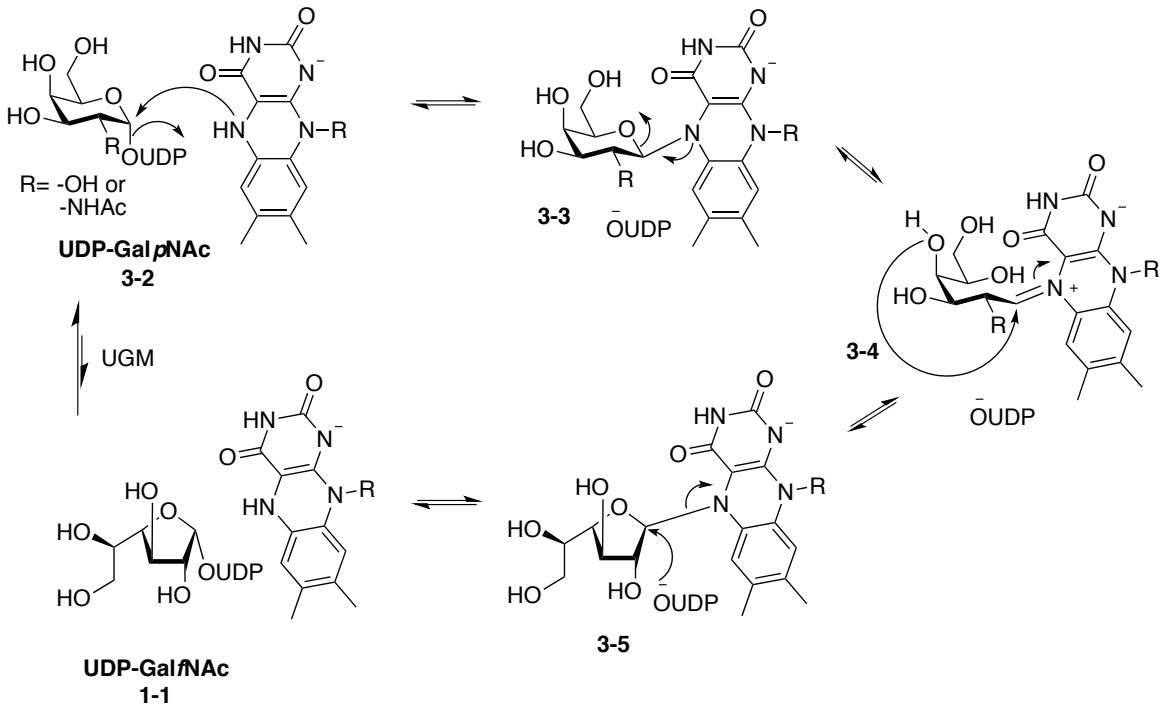


Figure 3-5: Proposed mechanism of UGM (R = OH) and UNGM (R= NHAc).

Based on sequence homology, the *cj1439c* gene from *C. jejuni* 11168 (HS:2) (Figure 1-13) was initially annotated as a *glf* gene.[7] The function of the protein was examined by Poulin *et al.,*[1] where the activity of recombinant *C. jejuni* UNGM (*cj1439c*) was confirmed by incubation with UDP-Gal*p*NAc under the reducing conditions modified from Liu and Zhang.[5] Product conversion was monitored by reversed-phase HPLC (high-performance liquid chromatography) using UV detection (254 nm). Formation of a longer retention time product—which was proved to be UDP-Gal*f*NAc by ^{1}H NMR spectroscopy and mass spectrometry—was observed at an equilibrium ratio of 7:93 (with respect to the UDP-Gal*p*NAc peak). This result agreed well with the equilibrium ratios observed for the *K. pneumonia*[8] and *E. coli*[5] UGM with UDP-Gal*p*, and highlights that enzymatic synthesis of UDP-Gal*f*NAc on a preparative scale is difficult to achieve. Though the reaction was attempted, and trace amounts of product was observed by HPLC (Figure 3-6) and MS, obtaining enough compound to carry out further investigation into

the putative GTs by this route would be both laborious and inefficient. Although coupling the enzymatic synthesis of **1-1** from **3-1** with the GT reaction of acceptors **1-2** to **1-5** would be advantageous in that GT product formation would push the equilibrium of the GalE and UNGM reactions forward, the introduction of more components into a one pot enzymatic synthesis would make analysis and product characterization more difficult. Our focus then shifted toward the chemical synthesis of UDP-Gal*f*NAc as a more facile route toward obtaining **1-1** in multi-milligram quantities.[2]

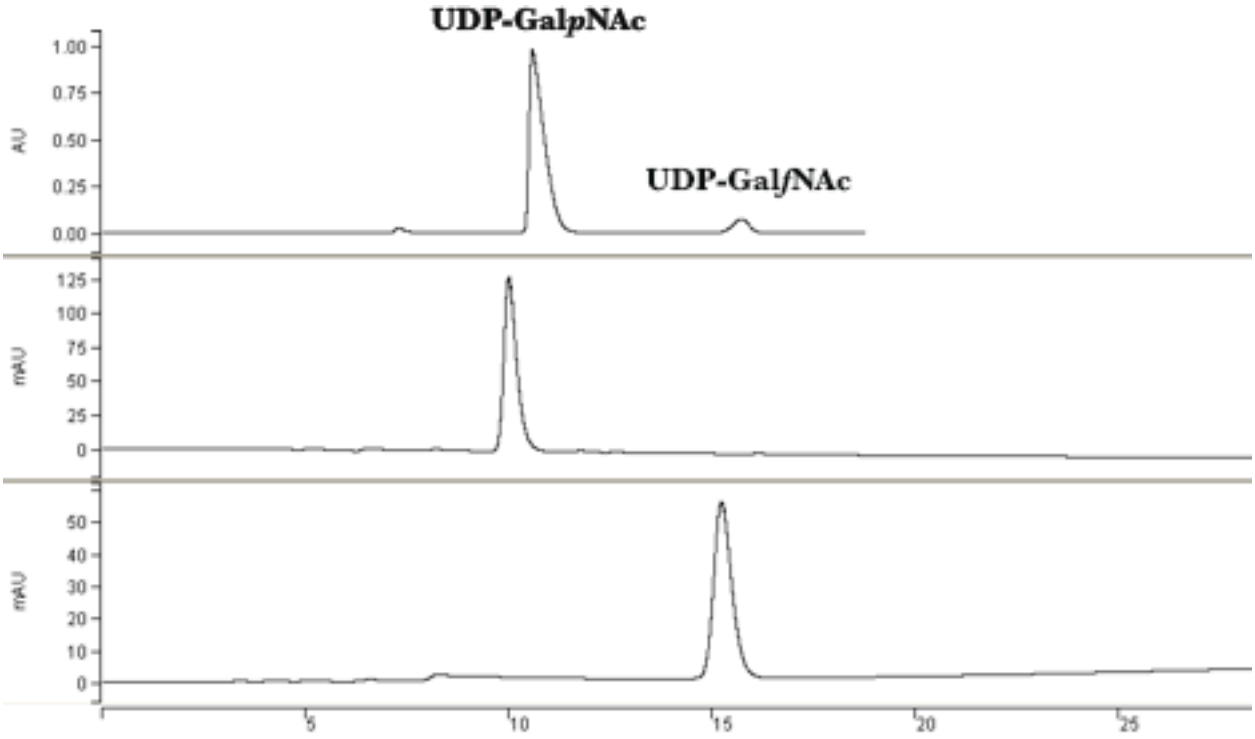


Figure 3-6: HPLC of UDP-Gal*f*NAc synthesis from GalE/UNGM one pot reaction (top). Internal standards of UDP-Gal*p*NAc and UDP-Gal*f*NAc (middle and bottom) confirm the newly formed peak was Gal*f*NAc. Note: UDP-Gal*p*NAc and UDP-Glc*p*NAc have same retention time on HPLC.

3.2 Chemical Synthesis

The synthesis of UDP-Gal*f*NAc began with the synthesis of 2-acetamido-2-deoxy-α-D-galactofuranose-1-phosphate, **3-10**, which was completed by former group graduate student Ryan Snitynsky in 2014.[2] The eight-step route used for the synthesis of **3-10** from 2-azido-2-deoxy-D-galactopyranose (**3-6**) is shown in Figure 3-7. Azidosugar **3-6** can be readly obtained by either azidonitration[9] or diazotransfer[10] of 3,4,6-tri-*O*-acetyl-D-galactal. Isopropylidene protection with 2,2-dimethoxypropane and *p*-TsOH afforded the furanose ring form with an anomeric methoxy group. Deprotection of the isopropylidene acetal and acetylation of the resulting triol yielded methyl glycoside **3-7**. Acetolysis of **3-7** yielded tetraacetate **3-8**, which was amenable to phosphorylation, first by conversion to the bromide with $TiBr_4$, and then reaction with dibenzylphosphate to obtain **3-9**. Debenzylation, azide reduction, and *N*-acetylation of **3-9** gave the required Gal*f*NAc-1-P **3-10** as a triethylammonium salt. The conversion of **3-9** into **3-10** was carried out in one step through hydrogenation over Pd/C using acetic anhydride as the solvent.

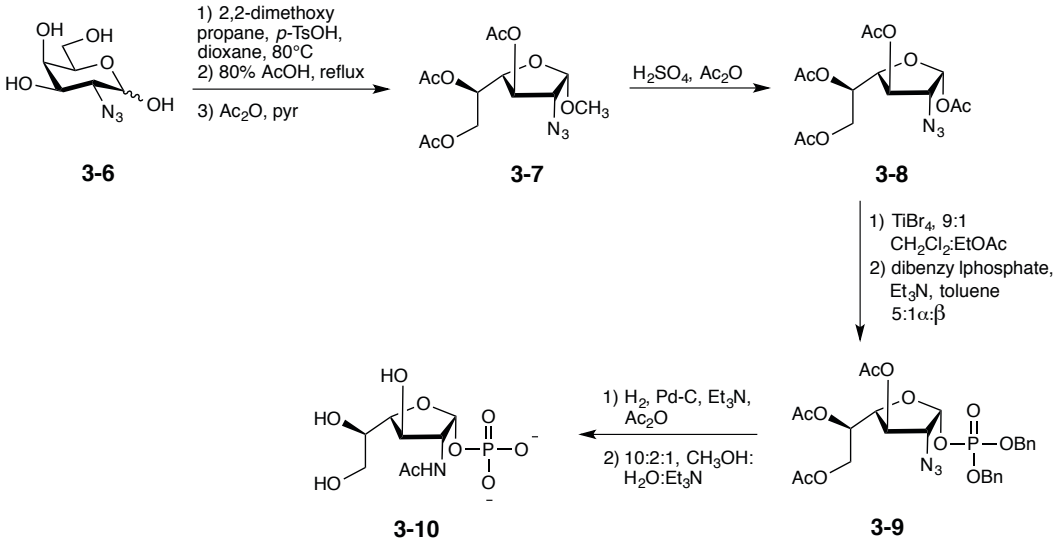


Figure 3-7 : Synthetic route to Gal*f*NAc-1-phosphate (3-10).[2]

With **3-10** in hand, its coupling with uridine-5′-monophosphate (UMP) was investigated, to complete the synthesis of UDP-Gal*f*NAc. This was most efficiently achieved by a route developed recently by Hindsgaul and coworkers.[11] This method first relies on the formation of an imidazole-activated uridine monophosphate (UMP) derivative. The activated UMP, UMP-Im, is formed from 2-imidazolyl-1,3-dimethylimidazolinium chloride (ImIm), which is generated *in situ* from imidazole and 2-chloro-1,3-dimethylimidazolinium chloride (DMC) (Figure 3-8). Subsequent addition of **3-10** yielded UDP-Gal*f*NAc in 18% yield, with a significant amount of starting material remaining (50%). Reaction times could not be elongated more than an hour, due to the formation of significant hydrolysis products and UMP dimerization. In addition, the reaction must be performed in D_2O because significantly higher rates of hydrolysis were observed when H_2O is used as the solvent (60–70% more hydrolysis reported).[11]

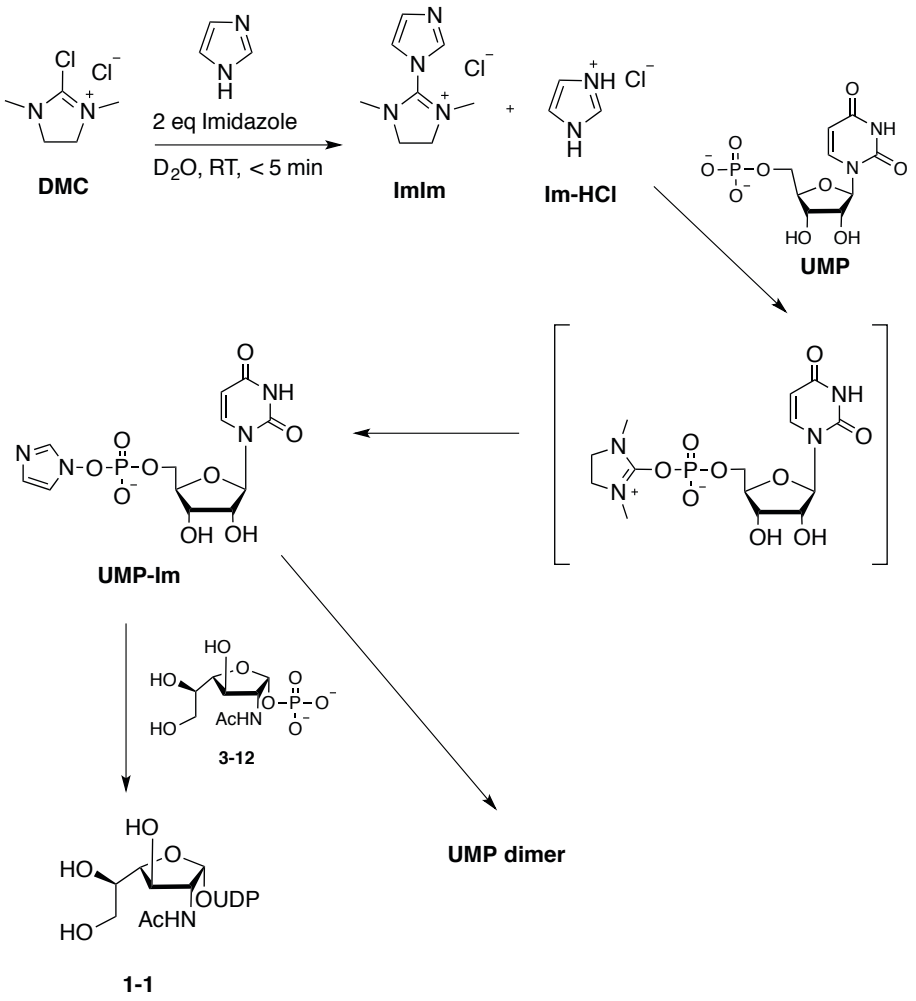


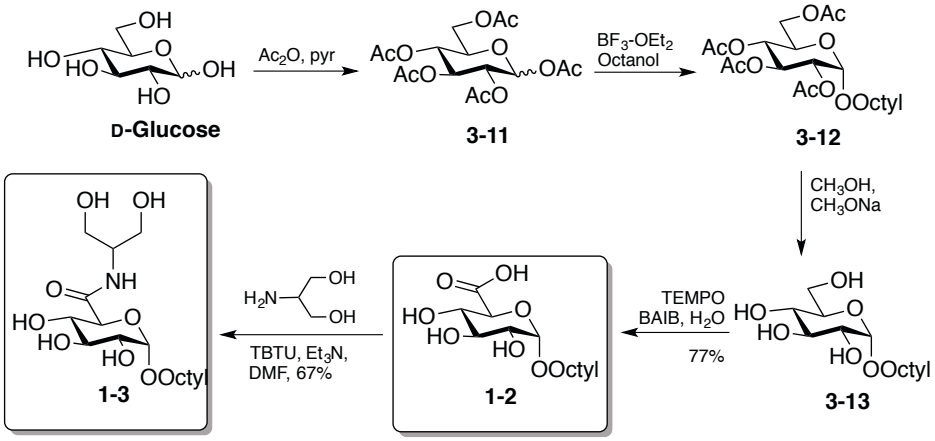
Figure 3-8: Chemical coupling of **3-10** with UMP via ImIm reagent.

3.3 Synthesis of Acceptor Substrates

3.3.1 Synthesis of Monosaccharide Acceptors 1-2 and 1-3

The synthesis of monosaccharide acceptors **1-2** and **1-3** was achieved using D-glucose as the starting material. The acceptors were made with an anomeric octyl linker, which eases the purification of the water soluble final product of enzymatic reactions, by acting as a 'handle' for performing column chromatography on reverse-phase (C_{18}) silica gel cartridges. Past work has demonstrated that this unnatural linker is well tolerated in GT reactions.[12,13]

Octyl α-D-glucopyranoside, **3-13**, was most efficiently obtained in three steps (Scheme 3-2). This involved acetylation, glycosylation with octanol activated by BF_3-OEt_2, to obtain 39% of the α anomer **3-12**, followed deprotection under Zemplen conditions in 84% to give **3-13**. Although neighbouring group participation by the C-2 acetoxy group leads to the β-octyl glucoside being formed, longer reaction times result in the formation of the more thermodynamically stable α-glucopyranoside. Preparation of **3-13** from D-glucose in a one step via Fischer glycosylation with *p*-toluenesulfonic acid and *n*-octanol as the solvent was also performed and the product could be obtained. However, this method was disadvantageous during scale up due to the cumbersome removal of large quantities of *n*-octanol (boiling point 195 °C).



Scheme 3-2: Route to monosaccharide acceptors **1-2** and **1-3**.

With a route to **3-15** in place, its conversion to the corresponding uronic acid, **1-2**, was achieved by selective oxidation of the primary hydroxyl group by treatment with catalytic amounts of (2,2,6,6-tetramethylpiperidin-1-yl)oxyl (TEMPO). This reaction employs a co-oxidant to regenerate TEMPO *in situ* (Figure 3-9)[14–16] and [bis(acetoxy)iodo]benzene (BAIB) was chosen over the cheaper KBr/NaOCl, because the final product, **1-2**, is water soluble, and its separation from the inorganic salts was expected to be cumbersome. Moreover, a simple dichloromethane

extraction removes all of the reaction byproducts, and the yields obtained in this reaction were more favourable than the KBr/NaOCl counterpart.

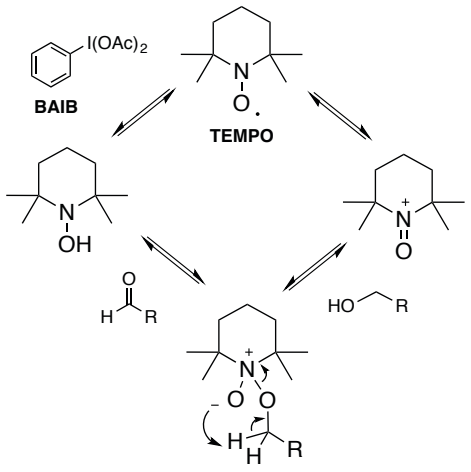
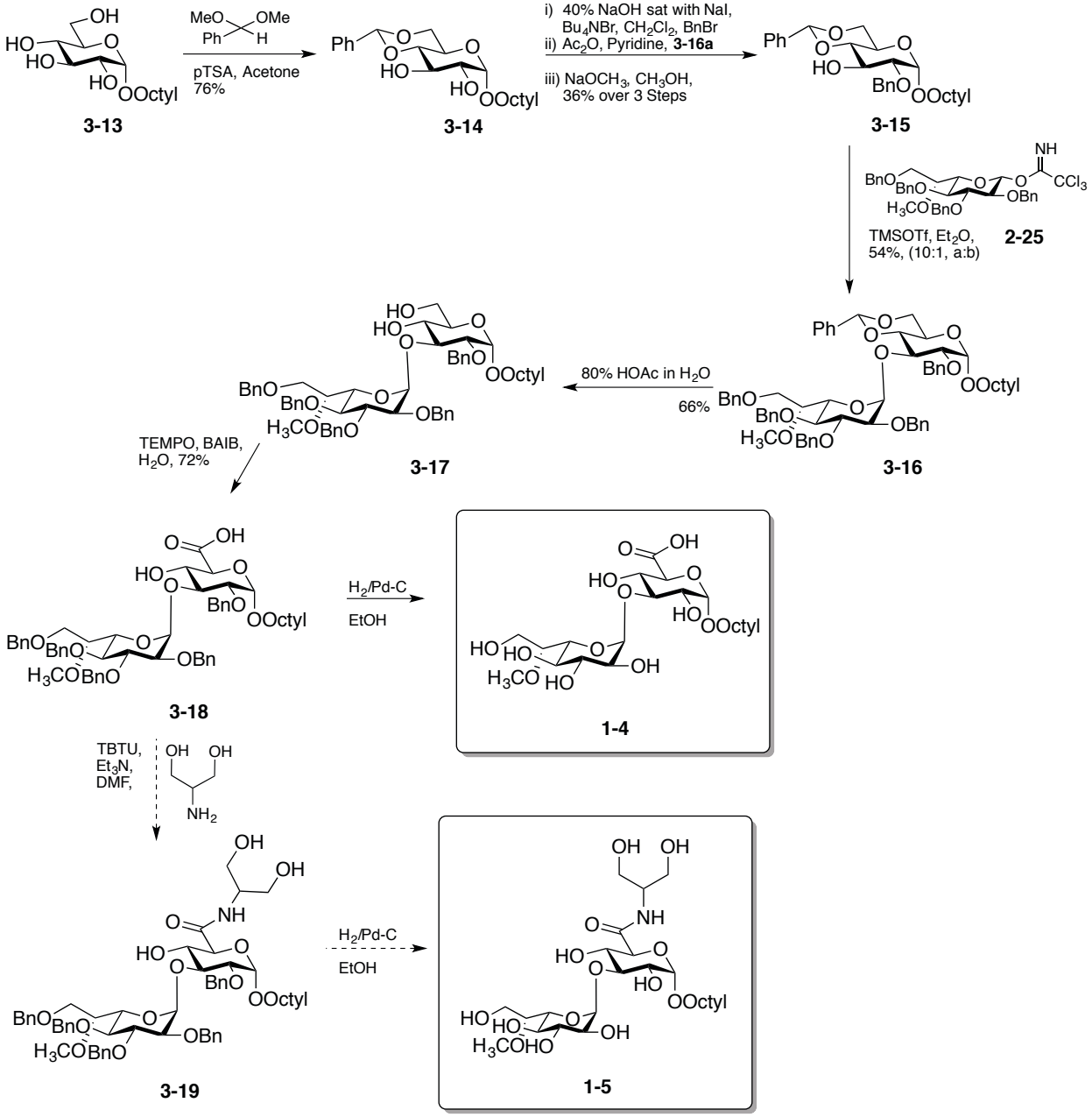


Figure 3-9: TEMPO oxidation mechanism.[14–16]

Reaction of the carboxylic acid in **1-2** with 2-amino-2-deoxy glycerol (serinol, available commercially) yielded the corresponding amide **1-3** in 67% yield. Selective amidation, via reaction with the amine, over esterification, via reaction with hydroxyl groups in the mixture (both **1-2** and the 2-amino-2-deoxy glycerol) was achieved by using *O*-(benzotriazol-1-yl)-*N*,*N*,*N*',*N*'-tetramethyluronium tetrafluoroborate (TBTU) as the coupling agent.[17] To validate the regioselectivity of the coupling strategy, 2-amino-1,3-di-*O*-benzyl-2-deoxy glycerol was also coupled with **1-2**, which would only result in amidation. After hydrogenolysis of the benzyl groups (56% yield over two steps), the ^{1}H NMR spectrum of the product was identical to that obtained from the TBTU-mediated coupling of **1-2** with 2-amino-2-deoxy glycerol. Additionally, the ^{1}H NMR spectrum showed H-2 of the serinol appears as symmetric quintet (see data in experimental section for **1-2**), which is only possible as a result of coupling with the N on C-2.

3.3.2 Synthesis of Disaccharide Acceptors 1-4 and 1-5.

With heptose donor **2-2** in hand, as well as an established route toward the glucuronic acid derivative **1-2** and the corresponding amino glycerol derivative **1-3**, attention was turned towards completing the synthesis of disaccharide acceptors **1-4** and **1-5**. The synthesis of these compounds began with octyl α-D-glucopyranoside, **3-13** (Scheme 3-3).



Scheme 3-3: Route to disaccharide acceptors **1-4** and **1-5**.

Benzylidination of **3-13** using benzaldehyde dimethyl acetal and *p*-toluenesulfonic acid gave **3-14** in 80% yield. The regioselective benzylation of C-2 hydroxyl group over the C-3 hydroxyl group was achieved by using a phase-transfer process (Figure 3-10).

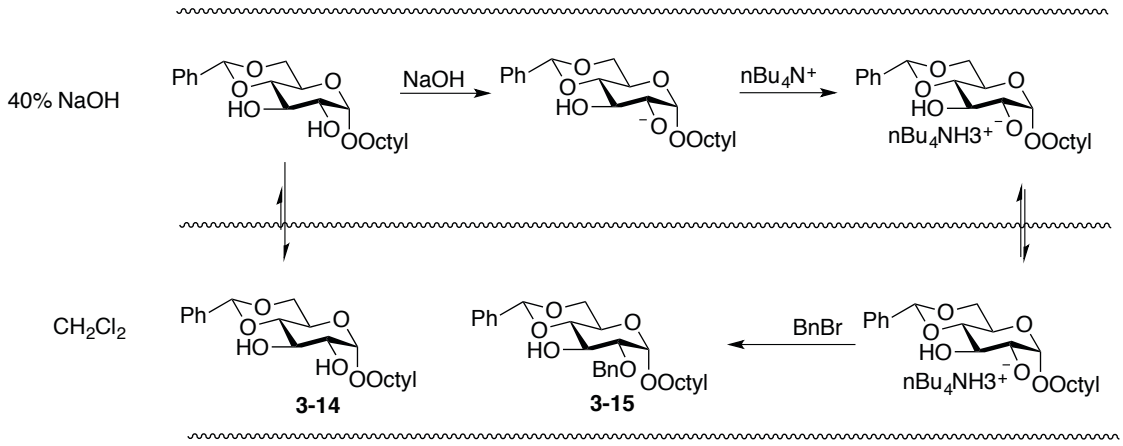
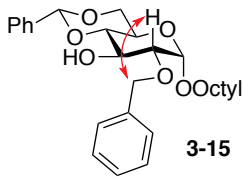


Figure 3-10: Phase transfer process to selectively benzylate the C-2 hydroyxl group in the conversion of **3-16** to **3-17**.

The reaction involved treating the substrate with sodium hydroxide saturated with sodium iodide, then adding tetrabutylammonium bromide, and benzyl bromide in a mixture of dichloromethane and water.[18] The *p*Ka of the C-2 hydroxyl group is lower than the *p*Ka of the C-3 hydroxyl group, which can be rationalized based on inductive effects and the proximity to the ring and anomeric oxygens. Therefore, under these phase-transfer conditions, the C-2 hydroxyl group is selectively deprotonated. The resulting alkoxide goes into the aqueous layer, complexes with the phase transfer catalyst, and then re-enters the organic phase where it is alkylated.[19] The resulting reaction gave a 10:1 mixture of the 2-*O*-benzylated and 3-*O*-benzylated products as determined from the ^{1}H NMR spectrum of **3-15**. However, the two compounds had very similar chromatographic mobilities and separation was therefore difficult. Thus, the combined mixture of products was acetylated, which facilitated purification. Deacetylation under Zemplen

conditions provided **3-15** in 36% yield from **3-14**. An HMBC experiment on **3-15** was used to confirm the presence of the desired 2-*O*-benzylation over 3-*O*-benzylation; a correlation was seen between H-2 and the methylene carbon on the benzyl ether.



3-15

Figure 3-11: HMBC correlation between H-2 and *C*H_2Ph suggesting 2-*O*-benzylation vs 3-*O*-benzylation.

In addition, the NMR spectrum of acetylated **3-15** showed a downfield shift of the H-3 proton compared to this resonance in **3-15** (δ 5.61 ppm vs 3.94 ppm). This would be expected upon acylation of the C-3 hydroxyl group. The lack of a downfield shift of H-2 upon acetylation further indicates that the C-2 hydroxyl group has been benzylated.

With acceptor **3-15** in hand, various heptose glycosyl donors were considered to obtain disaccharide **3-16** (Figure 3-12). The heptose is linked to the glucuronic acid residue in an α-1,2-*cis* manner, making this a synthetically challenging target. We first attempted the glycosylation using donor **2-2** directly, activated by NIS/AgOTf; however, this strategy provided poor stereoselectivity (2:1 α:β)

We also explored the use of a sulfoxide donor. Thus **2-2** was converted into sulfoxide **3-20** in 33% yield (plus 15% of the over-oxidized products) upon treatment with *m*-CPBA. However, attempted glycosylation of **3-15** with **3-20** produced the glycosyl sulfenate **3-22** as the only

observable product (in 16% yield), with the majority being unreacted donor and acceptor. Kahne and coworkers demonstrated that the energetic preference toward the sulfenate dramatically increases in the presence of an α oxygen on the carbon of a sulfoxide moiety.[20] Thus all anomeric sulfoxides have the ability to be converted to anomeric sulfenates, through the formation of a sulfonium ion (**3-21**), after the sulfoxide has been activated by TMSOTf (Figure 3-12). Byproducts of this type have previously been observed, particularly in reactions with unreactive hydroxyl groups.[21]

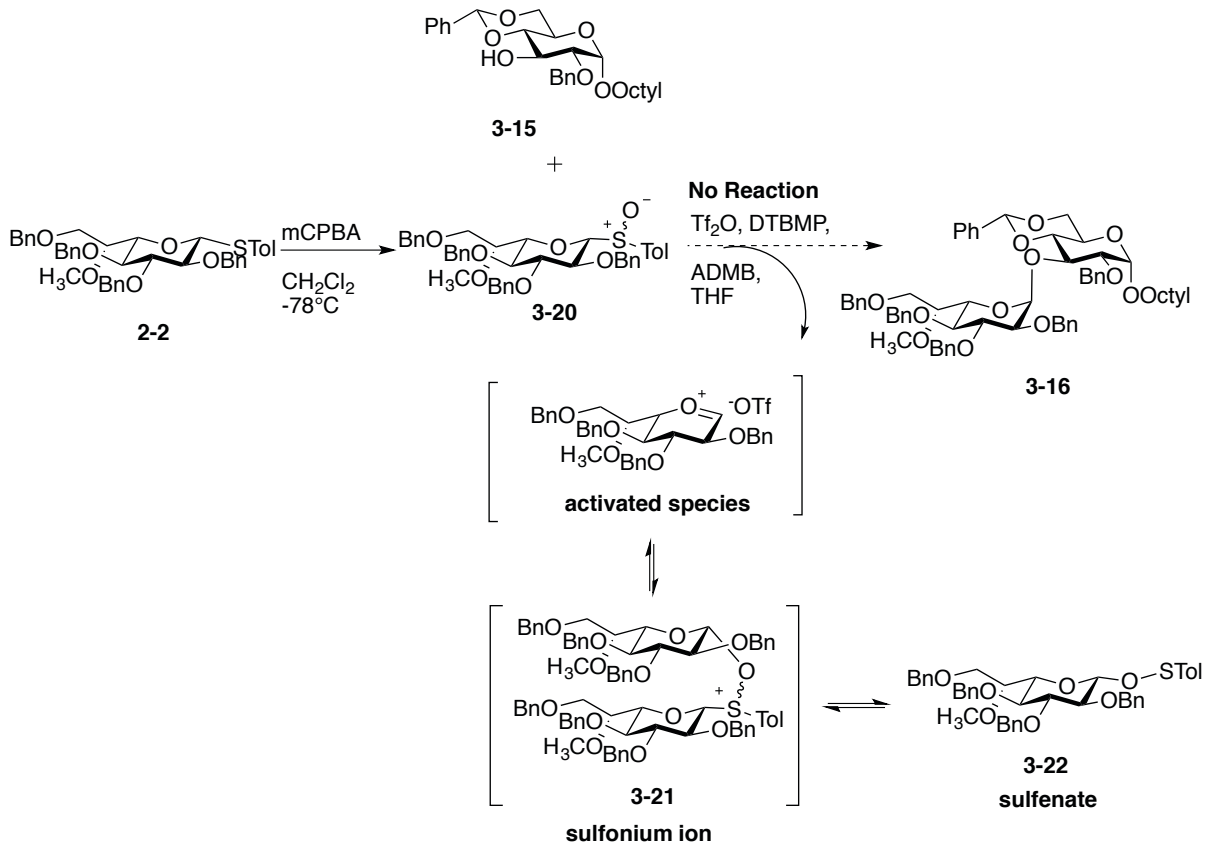


Figure 3-12: Attempt at accessing **3-16** from sulfoxide donor **3-20**.

After these failures, we found the most effective method for synthesizing **3-16** was through the imidate donor, **2-25**. Glycosylation of **2-25** with **3-15** resulted in a 10:1 α:β ratio of products, with overall modest yield (54%). After the generation **3-16**, the benzylidene acetal was cleaved by acid hydrolysis to afford diol **3-17**. Completion of the disaccharide acceptors was done using the same chemistry developed for the monosaccharide acceptors **1-2** and **1-3**. First oxidation of **3-17** with TEMPO/BAIB gave **3-18** in 80% yield. Finally, hydrogenolysis of the benzyl ethers in **3-18** provided the final acceptors **1-5** in 33% yield. Due to a lack of starting material **3-18**, the TBTU coupling and subsequent hydrogenolysis have yet to be completed.

3.4 Experimental

3.4.1 General Experimental Methods

All reagents were purchased from commercial sources and were used without further purification unless noted. Solvents used in reactions were purified by successive passage through columns of alumina and copper under argon. Unless stated otherwise, all reactions were carried out at room temperature and under a positive pressure of argon and monitored by TLC on Silica Gel G-25 F254 (0.25 mm). TLC spots were detected under UV light and/or by charring with a solution of anisaldehyde in ethanol, acetic acid and H_2SO_4. Column chromatography was performed on Silica Gel 60 (40–60 μm). Organic solutions were dried using anhydrous Na_2SO_4 and solvents were evaporated under reduced pressure and below 50 °C (water bath) on a rotary evaporator. ^{1}H NMR and ^{13}C NMR spectra were recorded at 400, 500 or 600 MHz. ^{1}H NMR chemical shifts are referenced to $CHCl_3$ (7.26 ppm, $CDCl_3$). ^{13}C NMR chemical shifts are referenced to CDC_{l3} (77.0 ppm, $CDCl_3$). ^{1}H NMR data are reported as though they were first order, and the peak assignments were made on the basis of 2D-NMR (^{1}H–^{1}H COSY and HSQC) experiments. ESI-

MS spectra were recorded on samples suspended in THF or CH_3OH and added NaCl. Optical rotations were measured at 22 ± 2 °C at the sodium D line (589 nm) and are in units of $deg \cdot mL(dm \cdot g)^{-1}$.

3.4.2 GalE Epimerase: Growth and Purification

Glycerol stock of GalE was used to inoculate a 2YT Broth (16g Bacto Tryptone + 10 g Bacto Yeast extract + 5 g NaCl + H_2O to final volume of 1 L, pH = 7.0) + 150 μg/mL ampicillin plate, which was then incubated at 37 °C overnight. The plate was used to inoculate 20 mL 2YT + 150 ug/ml ampicillin in a baffled flask. This was incubated at 37 °C x 200 rpm x overnight, and the culture was used to inoculate 1L 2YT + 150 μg/mL ampicillin in a baffled flask, incubated at 37 °C x for 1 hour x 200 rpm, until the OD_{600nm} = 0.400. Next, 0.5 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG) was added and the culture continued to incubate at 25 °C x 200 rpm x overnight. Cells were harvested by spinning at 5000 rpm x 30 min x 4 °C, and the supernatant was discarded. The cell paste was resuspended to 1:10 with 20 mM Tris pH 7.5 Emulsiflex with 1x protease inhibitor cocktail. Centrifuge was performed at 15k rpm x 30 min x 4 °C, and supernatant decanted, then filtered with a Millex 0.22 μm GV-13 filter. The supernatant was diluted 1:1 with running buffer. Purification was performed on an amylose column, where running buffer = 20 mM Tris pH 7.5 + 200 mM NaCl + 5 mM β-mercaptoethanol + 1mM EDTA, and the elution buffer = running buffer + 10% glycerol + 10 mM maltose.

Conditions for the Enzyme Assay:

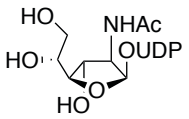
To test GalE activity, UDP-Glc*p*NAc (1 mM) was incubated with 10 mM $MnCl_2$, 50 mM Tris pH = 8 buffer, and 5 ul of GalE enzyme in a 10 μL reaction. The reaction was heated at 37 °C x 1 h,

and then visualized by CE-PDA using 25 mM $Na_2B_4O_7$ pH 9.4, 254 nm, 20 min run. Two peaks at 3:1 size ratio indicated enzymatic activity (**Figure 3-3**).

3.4.3 Mutase UNGM

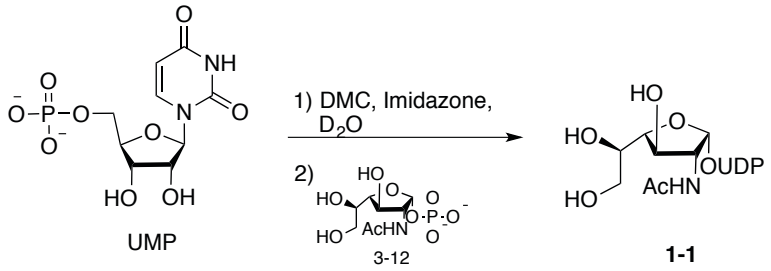
C-terminal hexa-histidine-tagged *C. jejuni* UNGM was expressed in *E. coli* DH5 from plasmid pWM1007:(*kan*)-Cj-*glf* after growth for 18 h at 28 °C using Lysogeny broth (LB 10 g Bacto Tryptone + 5 g Bacto Yeast extract + 10 g NaCl + H_2O to final volume of 1 L, pH = 7). Soluble UNGM-His_6 proteins were purified by Ni-NTA affinity chromatography, giving approximately 3 mg of >95% pure protein per liter of culture.

UDP-2-acetamido-2-deoxy-α-D-galactofuranose (1-1)



Enzymatic synthesis:

UDP-Glc*p*NAc (25 mg, 0.041mmol) was added to a solution of 500 mM phosphate buffer solution (500 μL), 50% glycerol (20 μL), 200 mM $MnCl_2$ (20 μL), 400 mM $Na_2S_2O_4$ (40 μL), GalE (40 μL), UNGM (20 μL), and H_2O (1.36 mL). The mixture was in incubated at 37 °C for 2 h, and then purified by HPLC using a C_{18} column (Microsorb, Varian, 21.4 250 mm) with 50 mM triethylamonium acetate (pH, 6.5) containing 1.5% acetonitrile at a flow rate of 7.0 mL/min, to give trace amounts of **1-1** (<1 mg, detected by HRMS (ESI)).

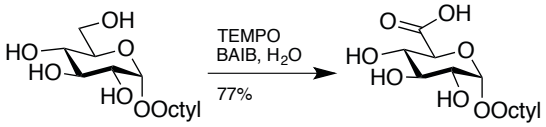


Chemical Coupling of UMP and GalfNAc-1-Phosphate:

UMP (Na+ form; 58.2 mg, 0.158 mmol), imidazole (43.2 mg, 0.636 mmol), and 2-chloro-1,3-dimethylimidazolinium chloride (DMC, 53.8 mg, 0.318 mmol) were dissolved in 60 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 **3-12** (20 mg, 0.040 mmol) was added, and the mixture was stirred for 18 h. The solution was diluted in 10 mL Tris-HCl buffer, pH = 8.0, 50 U of calf intestinal alkaline phosphatase 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 triethylammonium acetate 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 (x2), followed by freeze-drying to isolate the desired sugar nucleotide as a fluffy white solid (5.9 mg, 18%). ^{1}H NMR (500 MHz, D_2O) δ 7.98 (d, 1H, *J* = 8.1 Hz, Uridine H-6), 6.05–5.96 (m, 2H, Ribose H-1, Uridine H-5), 5.68 (dd, 1H, *J* = 5.5, 4.5 Hz, GalfNAc H-1), 4.53 (ddd, 1H, *J* = 9.2, 4.4, 2.5 Hz, Gal*f*NAc H-2), 4.42–4.37 (m, 2H, Ribose H-2, H-3), 4.33 (dd, 1H, *J* = 9.2, 7.7 Hz, GalfNAc H-3), 4.32–4.30 (m, 1H, Ribose H-4), 4.26 (ddd, 1H, *J* = 11.7, 4.6, 2.6 Hz, Ribose H-5), 4.21 (ddd, 1H, *J* = 11.8, 5.6, 3.0 Hz, Ribose H-5'), 3.95 (dd, 1H, *J* = 7.6, 5.0 Hz, GalfNAc H-4), 3.81 (ddd, 1H, *J* = 7.2, 5.0, 4.4 Hz, Gal*f*NAc H-5), 3.75 (dd, 1H, *J* = 11.8, 4.4 Hz, Gal*f*NAc H-6), 3.67 (dd, 1H, *J* = 11.8, 7.2 Hz, Gal*f*NAc H-6′), 3.23 (q,

12H, *J* = 7.3 Hz, NCH_2CH_3), 2.11 (s, 3H, GalfNAc $COCH_3$), 1.31 (t, 18H, *J* = 7.3 Hz, NCH_2CH_3). ^{13}C NMR (125 MHz, D_2O) δ 174.8 ($COCH_3$), 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, GalfNAc C-1), 88.4 (Ribose C-1), 83.3 (d, *J* = 9.0 Hz, Ribose C-4), 82.6 (GalfNAc 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 (NCH_2CH_3), 22.0 ($COCH_3$), 8.3 (NCH_2CH_3); ^{31}P NMR (162 MHz, D2O) 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: 606.0742.

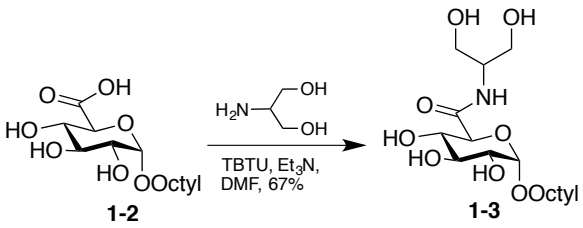
Octyl α-D-glucopyranosiduronic acid (1-2)



3-13 **1-2**

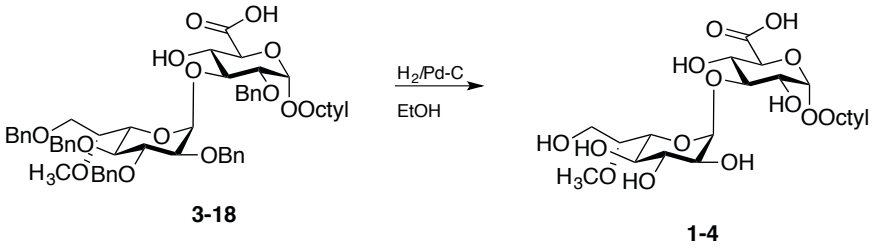
To a solution of **3-15** (169 mg, 0.578 mmol) in H_2O (15 mL) at 0 °C was added BAIB (205 mg, 1.1 equiv) and TEMPO (6 mg, 0.10 equiv). After 2 h, the reaction mixture was extracted with CH_2Cl_2, and the aqueous layer was partially concentrated under reduced pressure. The product was then passed through a C_{18}-SepPak cartridge and eluted with a gradient of 0 to 30% CH_3OH in H_2O. The eluant was lyophilized to give **1-2** (124 mg, 71%) as a white solid. R_f = 0.24 (7:1 CH_3OH –H_2O, C_{18} TLC); ^{1}H NMR (500 MHz, CD_3OD) δ 4.97 (d, 1H, *J* = 3.5 Hz, H-1), 4.06 (d, 1H, *J* = 9.5 Hz, H-5), 3.54−3.81 (m, 5H, H-2, H-3, H-4, OC*H*$_2$ octyl), 1.62−1.72 (m, 2H, octyl), 1.34−1.45 (br m, 10H, octyl), 0.92−0.95 (m, 3H, octyl); ^{13}C NMR (125 MHz) δ 173.4 (C=O), 100.6 (C1), 74.6, 73.4, 73.1, 72.9, 69.8, (6 C, C-2,C-3,C-4,C-5, O*C*H_2 octyl), 33.0, 30.5, 30.4, 27.3, 23.7, 14.4 (octyl). HRMS (ESI) m/z Calcd for $C_{14}H_{26}O_7$ [M– H] : 306.1679. Found 306.1923.

Octyl α-D-(*N*-2-deoxyglycero)glucopyranosiduronamide (1-3)



To a solution of **1-2** (55 mg, 0.180 mmol) in DMF (1 mL) was added TBTU (86 mg, 1.5 equiv), DIEA (94 μL) and lastly serinol (25 mg, 1.5 equiv). After 3 h, the reaction mixture was concentrated and the product was purified by column chromatography on Iatrobeads (3:1 CH_2Cl_2 –CH_3OH) to give **1-3** (46 mg, 67%) as a yellow solid. R_f = 0.18 (6:1 CH_2Cl_2 –CH_3OH); ^{1}H NMR (500 MHz, D_2O) δ 4.95 (d, 1H, *J* = 3.6 Hz, H-1), 4.07 (d, 1H, *J* = 9.9 Hz, H-3), 4.02 (quint., 1H, *J* = 5.7 Hz, N-C*H* of serinol), 3.77–3.62 (m, 8H, OCH_2 octyl, H-5, 2 x CH_2 serinol, H-6a, H-6b), 3.60 (dd, 1H, *J* = 9.7, 6.4 Hz, H-2), 3.54 (t, 1H, *J* = 10.8 Hz, OCH_2 octyl), 3.49 (dt, 1H, *J* = 9.9, 6.8 Hz, H-4), 2.68 (broad m, 1H, N*H*), 1.83–1.78 (m, 1H, octyl), 1.63–1.55 (m, 2H, octyl), 1.33–1.20 (m, 14H, octyl), 0.83 (t, 3H, *J* = 6.8 Hz, octyl); ^{13}C NMR (125 MHz) δ 172.2 (C=O), 99.6 (C-1), 73.6, 72.7, 72.5, 71.8, 69.8 (C-2, C-3, C-4, C-5, N-*C*H serinol), 61.20, 61.18 (*C*H_2 serinol), 53.6 (O*C*H_2 octyl), 29.8, 26.5, 23.2, 23.1, 20.7, 19.2, 14.6 (octyl); HRMS (ESI) m/z Calcd for $C_{17}H_{33}NO_8Na$ [M+Na]: 402.2098. Found: 402.2097.

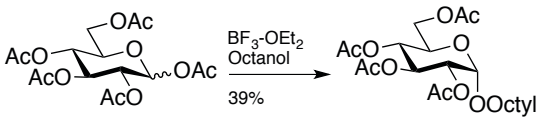
Octyl 6-*O*-methyl-D-*glycero*-α-L-*gluco*-heptopyranosyl-(1→3) α-D-glucopyranosiduronic acid (1-4)



3-18 1-4

To a solution of **3-20** (20 mg, 0.021 mmol) in EtOH (2 mL), was added Pd–C (2 mg, 10%/wt) and the mixture was stirred overnight under 1 atm H_2(g). Upon completion of the reaction, the reaction mixture was filtered through Celite, concentrated, and the resulting residue was purified under Iatrobead column chromatography to give **1-4** (3.5 mg, 33%) as a white solid. R*f* = 0.32 (2:1 CH_2Cl_2–CH_3OH); ^{1}H NMR (500 MHz, $CDCl_3$): δ 5.28 (d, 1H, *J* = 3.9 Hz, H-1 Glc), 5.07 (d, 1H, *J* = 3.9 Hz, H-1 Glc), 3.94–3.88 (m, 3H, H-3 Hep), 3.85–3.78 (m, 7H, H-3 Glc, OC*H*$_2$ octyl), 3.75–3.68 (m, 5H, H-2 Hep, OC*H*$_2$ octyl, H-5 Hep), 3.66–3.60 (m, 2H, H-2 Glc), 3.69 (s, 3H, OC*H*$_3$), 1.69–1.58 (m, 2H, octyl), 1.31–1.26 (m, 10H, octyl), 0.97 (t, 3H, *J* = 7.4 Hz, octyl); ^{13}C NMR (125 MHz, CDCl3): δ 101.3 (C-1 Glc), 98.4 (C-1 Hep), 79.0, 78.7, 72.9, 72.6, 72.57, 72.52, 71.9, 70.5, 70.1, 69.0, 61.4(O*C*H$_2$ octyl), 60.9 (C-7 Hep), 60.3 (O*C*H$_3$), 32.1, 29.5, 29.4, 26.4, 23.0, 14.4.

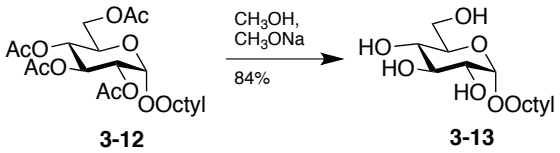
Octyl 2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranoside (3-12)



3-11 **3-12**

To a solution of commerciually available 1,2,3,4,6-penta-*O*-acetyl-β-D-glucopyranose (4.8 g, 13.0 mmol) in dry CH_2Cl_2 at 0 °C, was added 1-octanol (6.20 mL, 3 equiv), and $BF_3{\cdot}OEt_2$. After stirring 1.5 d at room temperature, CH_3OH (10 mL) was added, and the mixture was stirred for 10 min. The solution was diluted with EtOAc (100 mL) and washed with water, brine, and the organic layer was dried over Na_2SO_4. After evaporation under reduced pressure, the crude residue was purified by column chromatography (4:1 hexane–EtOAc) to give **3-14** (2.30 g, 39% yield) as a clear oil. R_f = 0.27 (4:1 hexane–EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$): δ 5.48 (app t, 1H, *J* = 10.2 Hz, H-3), 5.05 (d, 1H, *J* = 3.7 Hz, H-1), 5.04 (dd, 1H, *J* = 9.8, 10.2 Hz, H- 4), 4.84 (dd, 1H, *J* = 10.2, 3.7 Hz, H-2), 4.25 (dd, 1H, *J* = 12.2, 4.5 Hz, H-6a), 4.08 (dd, 1H, *J* = 12.2, 2.2 Hz, H-6b), 4.01 (ddd, 1H, *J* = 10.2, 4.5, 2.2 Hz, H-5), 3.67 (td, 1H, *J* = 9.7, 6.5 Hz, OC*H*$_2$ octyl), 3.42 (td, 1H, *J* = 9.7, 6.5 Hz, OC*H*$_2$ octyl), 1.96, 1.93, 1.91, 1.89 (4s, 12H,4 x CH3), 1.58 (quint, 2H, J = 6.5 Hz, octyl), 1.36–1.22 (m, 10H, octyl), 0.88 (t, 3H, *J* = 6.8 Hz, octyl); ^{13}C NMR (125 MHz, $CDCl_3$): δ 170.2, 169.70, 169.68, 169.2 (4 x C=O), 95.3 (C-1), 70.6 (C-2), 69.9 (C-3), 68.3 (O*C*H$_2$ octyl), 66.8 (C-5), 61.6 (C-6), 31.5, 28.9, 25.7, 22.3 (octyl), 20.3 (2 *C*H$_3$–CO), 20.24, 20.22 (2 *C*H$_3$–CO), 13.7 (octyl).

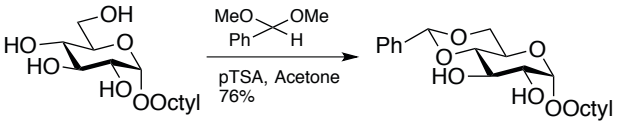
Octyl α-D-glucopyranoside (3-13)



3-12 **3-13**

To a solution of **3-14** (2.3 g, 5.0 mmol) in 3:1 CH_2Cl_2–CH_3OH (20 mL), 1M $NaOCH_3$ in CH_3OH was added until the pH of the solution was above 10. After 4 h, the solution was neutralized by the addition of IR-120 (H+), filtered, and concentrated. The crude product was purified by column chromatography (6:1 CH_2Cl_2–CH_3OH) to give **3-15** as a white solid (1.23 g, 84%). R_f = 0.39 (6:1 CH_2Cl_2–CH_3OH); ^{1}H NMR (500 MHz, $CDCl_3$): δ 4.98 (d, 1H, *J* = 3.7 Hz, H-1), 3.92 (dd, 1H, *J* = 11.8, 2.4 Hz, H-6a), 3.81 (td, 1H, *J* = 9.6, 6.9 Hz, OC*H*$_2$ octyl), 3.80 (dd, 1H, *J* = 11.8, 5.4 Hz, H-6b), 3.75 (dd, 1H, *J* = 9.7, 9.3 Hz, H-3), 3.72 (ddd, 1H, *J* = 9.8 Hz, 5.4, 2.4 Hz, H-5), 3.62 (td, 1H, *J* = 9.6 Hz, 6.4 Hz, OC*H*$_2$ octyl), 3.60 (dd, 1H, *J* = 9.7, 3.7 Hz, H-2), 3.47 (dd, 1H, *J* = 9.8, 9.3 Hz, H-4), 1.72–1.65 (m, 2H, octyl), 1.48-1.30 (m, 10H, octyl), 0.92 (t, 3H, *J* = 6.8 Hz, octyl); ^{13}C NMR (125 MHz, CD_3OD): δ 99.0 (C-1), 74.1 (C-3), 72.7 (C-2), 72.3 (C-5), 70.5 (C-4), 69.3 (O*C*H$_2$ octyl), 61.9 (C-6), 33.0, 30.6, 30.5, 30.3, 27.3, 23.7, 14.5 (octyl).

Octyl 4,6-*O*-benzylidene-α-D-glucopyranoside (3-14)

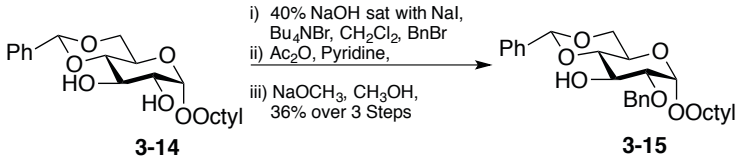


3-13 **3-14**

To a solution of **3-15** (2.22 g, 7.60 mmol) in acetonitrile (50 mL) was benzaledhyde dimethyl acetal (13.7 mL, 12 equiv) and *p*-toluenesulfonic acid (11 mg, 5% wt). After stirring overnight, the solution was cooled to 0 °C and neutralized with triethylamine and concentrated. The resulting residue was purified by chromatography (1:1 hexane–EtOAc) to give **3-16** (2.06 g,

76%) as a white solid. R_f = 0.35 (1:1 hexane–EtOAc); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.35− 7.51 (m, 5 H, Ph), 5.54 (s, 1 H, C*H*-Ph), 4.89 (d, 1H, *J* = 4.0 Hz, H-1), 4.28 (dd, 1H, *J* = 10.0, 4.5, Hz, H-6a), 3.94 (app t, 1H, J 9.5 Hz, H-3), 3.86−3.78 (m, 1H, H-5), 3.78−3.70 (m, 2H, H-6b, OC*H*$_2$ octyl), 3.65−3.58 (m, 1H, H-2), 3.51−3.43 (m, 2H, H-4, OC*H*2), 2.78 (br s, 1H, OH), 2.20 (d, 1 H, *J* =1.8 Hz, OH), 1.71−1.60 (m, 4H, octyl), 1.40−1.23 (br m, 10H, octyl), 0.88−0.91 (m, 3H, octyl); ^{13}C NMR (125 MHz, $CDCl_3$) δ 137.3, 129.4, 128.5, 126.5, 102.1 (Ph*C*H), 98.9 (C-1), 81.2 (C-3), 73.2 (C-2), 72.1 (C-5), 69.2 (C-4), 69.0 (C-6), 62.7(O*C*H$_2$ octyl), 31.9, 29.6, 29.5, 29.3, 26.2, 22.8, 14.4 (octyl).

Octyl 2-*O*-benzyl-4,6-*O*-benzylidene-α-D-glucopyranoside (3-15).

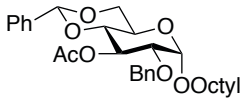


3-14 **3-15**

To a solution of **3-16** (498 mg, 1.31 mmol) in anhydrous dichloromethane (3.3 mL) and 40% NaOH in water saturated with NaI (0.26 mL), was added tetrabutylammonium iodide (49 mg, 0.13 eq) and benzyl bromide (0.16 mL, 1 eq). After 2 h, the reaction mixture was poured into a separatory funnel with 10 mL of H_2O. This mixture was extracted with dichloromethane (3 x 5 mL) and the organic later was washed with brine before being dried with $NaHCO_3$ and concentrated. To a solution of the crude in anhydrous pyridine (5 mL) at 0 °C was added Ac_2O (1 mL). After stirring for 3 h, the mixture was concentrated. The residue was dissolved in CH_2Cl_2 and then washed with 1N HCl, water, and brine before being dried and concentrated. The residue was purified by chromatography (9:1 hexane–EtOAc). After concentration, the product was dissolved in CH_3OH (16 mL) and $NaOCH_3$ in CH_3OH was added until the pH of the solution was 9. After stirring for 4 h, Amberlite IR120 H+ resin was added to neutralize the

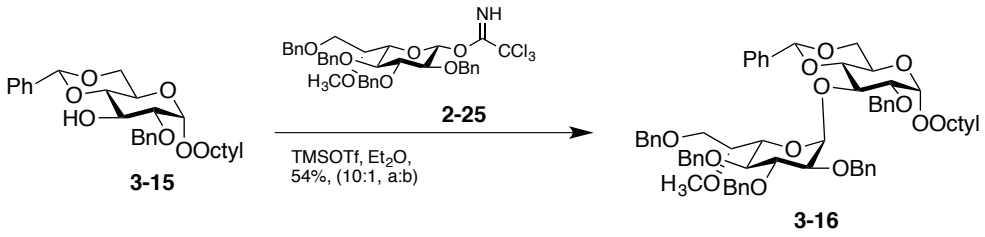
solution. After filtration, the solution was concentrated. The crude product was purified by chromatography (6:1 hexane–EtOAc) to give **3-17** (221 mg, 36% over 3 steps) as a white solid. R*f* = 0.43 (4:1 hexane– EtOAc). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.45−7.28 (m, 10H, Ph), 5.60 (s, 1H, C*H*-Ph), 4.99 (d, 1H, *J* = 11.0 Hz, PhCH2), 4.93 (d, 1H, *J* = 3.6 Hz, H-1), 4.87 (d, 1H, *J* = 10.9 Hz, PhCH2), 4.32(dd, 1H, *J* = 10.1, 4.8, Hz, H-3), 3.90-3.72 (m, 5H, H-4, H-6a, H-2, OCH_2 octyl, H-6b), 3.66 (t, 1H, app, *J* = 9.4 Hz, H-5), 3.51 (dt, 1H, *J* = 9.6, 6.7 Hz, OCH_2 octyl), 2.27 (d, 1H, *J* = 8.5 Hz, OH), 1.67 (quint, 2H, *J* = 7.0 Hz, octyl), 1.40–1.25 (m, 10H, octyl), 0.92 (t, 3H, *J* = 7.2 Hz, octyl); ^{13}C NMR (125 MHz, $CDCl_3$) δ 126.8, 125.9, 125.6, 123.9, 123.6, 99.6 (Ph*C*H), 95.0 (C-1), 79.0 (C-5), 77.4 (C-4), 70.1 (C-2) , 67.7 (C-3), 66.7, 66.2 (O*C*H_2 octyl), 59.8 (C-6), 29.5, 27.0, 23.8, 11.7 (octyl).

Octyl 3-*O*-acetyl-2-*O*-benzyl-4,-*O*-benzylidene-D-glucopyranoside (Isolated from intermediate towards synthesis of 2-15).



^{1}H NMR (400 MHz, $CDCl_3$) δ 7.45−7.30 (10H, m, Ph), 5.61 (1H, app t, *J* = 9.68 Hz, H-3), 5.49 (1H, s, C*H*-Ph), 4.84 (1H, d, *J* = 3.6 Hz, H-1), 4.68 (2H, d, *J* = 8.0 Hz, PhCH_2), 4.29 (1H, dd, *J* = 9.7, 4.9, Hz, H-6a), 3.95 (1H, ddd, *J* = 10.0, 9.9, 4.9 Hz, H-6b), 3.72 (1H, t, app, *J* = 10.36 Hz, H-4), 3.68 (1H, app t, *J* = 6.8 Hz, OCH_2 octyl), 3.61(1H, dd, *J* = 9.8, 3.6, Hz, H-2), 3.56 (1H, t, app, *J* = 9.6 Hz, H-5), 3.44 (1H, dt, *J* = 9.7, 6.7 Hz, OCH_2 octyl), 2.10, (3H, s, CH_3), 2.73-2.62 (2H, m, octyl), 1.48–1.34 (10H, m, octyl), 1.90 (3H, t, *J* = 6.8 Hz, octyl).

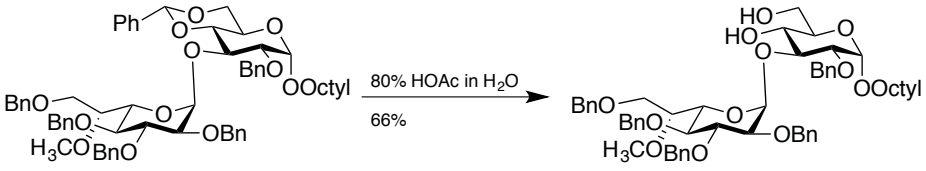
Octyl 2,3,4,7-tetra-*O*-benzyl-6-*O*-methyl-D-*glycero*-α-L-*gluco*-heptopyranosyl (1→3)-2-*O*-benzyl-4,6-*O*-benzylidene-α-D-glucopyranoside (3-16)



To a solution of imidate donor 2-24 (119 mg, 0.165 mmol), 3-17 (86 mg, 1.1 equiv) in Et_2O (1 mL and 4 Å molecular sieves (10 mg) in Et_2O (0.5 mL) was added 5% TMSOTf in Et_2O until the pH of the solution was 5. After stirring for 45 min, the solution was filtered, concentrated and the residue was purified by chromatography (9:1 hexane–EtOAc) to give **3-18 α** (90 mg, 54%) + **3-18 β** (5 mg, 3%) both as a white solids. Data for α-anomer: R_f = 0.43 (4:1 hexane–EtOAc); $[\alpha]_D$ 1.68 (*c* 1.0, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.25–7.15 (m, 30H, Ph), 5.57, (s, 1H, C*H*Ph), 5.48 (d, 1H, *J* = 3.5 Hz, H-1 Glc), 5.06 (d, 1H, *J* = 10.8 Hz, CH_2Ph), 5.03 (d, 1H, *J* = 10.8 Hz, CH_2Ph), 4.98 (d, 1H, *J* = 10.5 Hz, CH_2Ph), 4.77 (d, 1H, *J* = 3.5 Hz, H-1 Hep), 4.82 (d, 1H, *J* = 10.8 Hz, CH_2Ph), 4.78 (d, 1H, *J* = 10.8 Hz, CH_2Ph), 4.64-4.58 (m, 3H, CH_2Ph), 4.52 (d, 1H, *J* = 11.9 Hz, CH_2Ph), 4.43(d, 1H, *J* = 12.5 Hz, CH_2Ph), 4.30 (m, 3H, H-5 Hep), 4.18 (app t, 1H, *J* = 9.4 Hz, H-3 Hep), 4.08 (app t, 1H, *J* = 9.4 Hz, H-3 Glc), 3.88–3.68 (m, 10H, H-6 Hep, H-5 Glc, H-4 Hep, H-4 Glc, H-7a/b Hep, H-6 a/b Glc, H-2 Hep), 3.56–3.51 (m, 2H, OCH_2 octyl, H-2 Glc), 3.38 (s, 3H, OCH_3), 3.33 (dt, 1 H, *J* = 9.9, 6.8 Hz, OCH_2 octyl), 1.40–1.20 (m, 20H, octyl), 0.95 (t, 3H, *J* = 7.2 Hz, octyl); ^{13}C NMR (125 MHz, $CDCl_3$) δ 139.0, 138.7, 138.6, 137.9, 137.7, 137.4, 128.5, 128.44, 128.40, 128.38, 128.36, 128.31, 128.27, 128.23, 128.21, 128.17, 127.88, 127.81, 127.74, 127.69, 127.59, 127.54, 127.45, 126.78, 126.1, 101.3 (*C*HPh), 98.31 (C-1 Hep), 98.27 (C-1 Glc), 91.9, 83.1, 81.7, 79.3, 79.3, 78.6, 78.6, 76.3, 76.1, 76.3 (*C*H_2PH), 75.5 (*C*H_2PH), 74.9 (*C*H_2PH), 74.8 (*C*H_2PH), 71.8 (*C*H_2PH), 70.2, 69.3, 68.2 (OCH_2 octyl), 67.8,

62.1, 58.8 (O*C*H$_3$), 31.8, 29.4, 29.2, 26.1, 24.0, 22.7, 14.1 (CH$_3$ octyl); HRMS (ESI) m/z Calcd for $C_{64}H_{76}O_{12}Na$ [M+Na]: 1036.5337. Found: 1036.5339.

Octyl-2,3,4,7-Tetra-*O*-benzyl-6-*O*-methyl-D-*glycero*-α-L-*gluco*-heptopyranosyl (1→3)-2-*O*-benzyl-α-D-glucopyranoside (3-17)

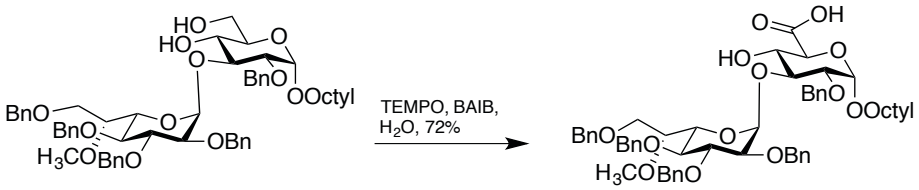


3-16 **3-17**

To a solution of **3-18** (47 mg, 0.045 mmol) in 6:1 CH_2Cl_2–CH_3OH (3.5 mL) was added *p*-toluenesulfonic acid (10 mg, 0.058 mmol) and the reaction was heated at 30 °C overnight. The mixture was cooled, and triethylamine was added until the was pH = 7–8. The solution was then concentrated, and the resulting residue was purified by column chromatography (1:1 hexane–EtOAc) to give **3-19** (27 mg, 63%) as a pale yellow oil. Rf = 0.42 (4:1 hexane–EtOAc); $[\alpha]_D$ – 3.2 (*c* 0.9, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$) δ7.40–7.20 (m, 25H, Ph), 5.19 (d, 1H, *J* = 3.3 Hz, H-1 Glc), 5.20 (d, 1H, *J* = 12.3 Hz, C*H*$_2$Ph), 5.02 (d, 1H, *J* = 10.8 Hz, C*H*$_2$Ph), 5.01 (d, 1H, *J* = 11.4 Hz, C*H*$_2$Ph), 4.90 (d, 1H, *J* = 3.5 Hz, H-1 Hep), 4.70 (d, 1H, *J* = 12.4 Hz, C*H*$_2$Ph), 4.66 (d, 1H, *J* = 11.4 Hz, C*H*$_2$Ph), 4.65 (d, 1H, *J* = 12.1Hz, C*H*$_2$Ph), 4.61 (d, 1H, *J* = 12.1 Hz, C*H*$_2$Ph), 4.56 (s, 1H, C*H*$_2$Ph), 4.14 (app t, 1H, *J* = 9.4 Hz, H-3 Glc), 3.87 (dd, 1H, *J* = 9.9, 1.1 Hz, H-5 Glc), 3.84–3.68 (m, 7H, H-6a/b Glc, H-3 Hep, H-4 Glc, H-7a/b Hep, H-5 Hep), 3.62–2.58 (m, 2H, H-2 Hep, H-6 Hep), 3.55 (dd, 1H, *J* = 9.9, 3.5 Hz, H-2 Glc), 3.52–3.45 (m, 2H, H-4 Hep, OC*H*$_2$ octyl), 3.44 (s, 3H, OC*H*$_3$), 3.23 (dt, 1 H, *J* = 9.7, 6.8 Hz, OC*H*$_2$ octyl), 1.65–1.55 (m, 7H, octyl), 1.28–1.22 (m, 15H, octyl), 0.95 (t, 3H, *J* = 7.2 Hz, octyl); ^{13}C NMR (125 MHz, $CDCl_3$) δ 139.1, 128.7, 128.6, 128.0, 127.96, 128.6, 128.4, 128.36, 128.35, 128.26, 127.8,

127.74, 127.68, 127.3, 127.6, 127.58, 127.56, 127.54, 127.49, 127.4, 99.6 (C-1 Glc), 97.8 (C-1 Hep), 81.7 (H-3 Glc), 81.1, 79.4, 79.3, 77.5, 75.5 (CH$_2$PH), 74.9 (CH$_2$PH), 74.8 (CH$_2$PH), 73.5 (CH$_2$PH), 73.3 (CH$_2$PH), 71.2, 71.1, 69.2, 7.7 (H-6 Glc or H-7 Hep), 67.7 (H-6 Glc or H-7 Hep), 62.8 (OCH$_2$ octyl), 59.1 (OCH$_3$), 31.9, 29.6, 29.5, 29.3, 26.2, 22.7, 14.1 (octyl); HRMS (ESI) m/z Calcd for $C_{57}H_{72}O_{12}Na$ [M+Na]: 948.5024. Found: 948.502.

Octyl 2,3,4,7-Tetra-*O*-benzyl-6-*O*-methyl-D-*glycero*-α-L-*gluco*-heptopyranosyl (1→3)-2-*O*-benzyl-α-D-glucopyranosiduronic acid (3-18)

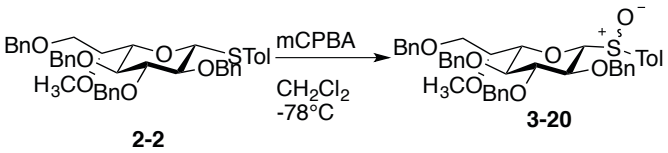


3-17 **3-18**

To a solution of **3-19** (9 mg, 0.00948 mmol) in wet CH_2Cl_2 (0.5 mL) at 0 °C was added BAIB (3.5 mg, 1.1 equiv) and TEMPO (1 mg, 0.10 equiv). After 2 h, the reaction mixture was concentrated under reduced pressure and the product was purified by column chromatography (20:1 CH_2Cl_2–CH_3OH with 1% AcOH) to give **3-20** (8 mg, 80%) as a white solid. R_f = 0.33 (20:1 CH_2Cl_2–CH_3OH); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.23 (1H, s, COOH), 7.40–7.20 (m, 25H, Ph), 5.89 (d, 1H, J = 3.3 Hz, H-4 Glc), 5.32 (d, 1H, J = 3.7 Hz, H-1 Hep), 5.21 (d, 1H, J = 2.6 Hz, H-1 Glc), 5.03(d, 1H, J = 11.0 Hz, CH_2Ph), 4.99 (d, 1H, J = 11 Hz, CH_2Ph), 4.86 (d, 1H, J = 12.3 Hz, CH_2Ph), 4.81 (d, 1H, J = 12.3 Hz, CH_2Ph), 4.75 (d, 1H, J = 11.9 Hz, CH_2Ph), 4.70 (d, 1H, J = 11.9 Hz, CH_2Ph), 4.65–4.51 (m, 10H, 2 x CH_2Ph, H-5 Glc), 4.06–4.03 (m, 2H, H-3Hep), 4.01(dd, 1H, J = 9.9, 2.6 Hz, H-2 Glc), 3.82–3.66 (m, 8H, H-4 Hep), 3.58 (dd, 1H, J = 9.4, 3.7 Hz, H-2 Hep), 3.56–3.53 (m, 2H, OCH_2 octyl), 3.51, (s, 3H, OCH_3), 3.50–3.48 (m, 1H,

OCH_2 octyl), 1.69–1.58 (m, 2H, octyl), 1.31–1.26 (m, 10H, octyl), 0.97 (t, 3H, J = 7.4 Hz, octyl); HRMS (ESI) m/z Calcd for $C_{57}H_{70}O_{13}Na$ [M+Na]: 962.4816. Found: 948.4818.

2,3,4,7-Tetra-*O*-benzyl-6-*O*-methyl-D-*glycero*-α/β-L-*gluco*-heptopyranoyl *p*-toluyl (*R*/*S*)-sulfoxide (3-20)



2-2 **3-20**

To a solution of **2-21** (66 mg, 0.0955 mmol) in CH_2Cl_2 (2 mL) cooled to –70 °C was added *m*-CPBA (18 mg, 0.105 mmol, 1.1 equiv). After stirring 3 h, the mixture was concentrated under reduced pressure and purified by column chromatography (2:1 hexane–EtOAc) to give **3-22** (35 mg, 52%) as an white solid and the over oxidized sulfone product (10 mg, 15%) as an white solid. R_f = 0.12 (4:1 hexane–EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.88 (d, 2H, J = 8.0 Hz, Ph), 7.43 (d, 2H, J = 8.0 Hz, Ph), 7.48–7.22 (m, 20H, Ph), 7.17 (d, 2H, J = 8.0 Hz, Ph), 5.18 (d, 1H, J = 9.7 Hz, PhCH_2), 5.05 (d, 1H, J = 10.0 Hz, PhCH_2), 4.94–4.85 (m, 3H, PhCH_2), 4.62 (d, 1H, J = 10.0 Hz, PhCH_2), 4.34 (d, 1H, J = 9.9 Hz, H-1), 3.87 (app t, 1H, J = 8.8 Hz, H-2), 4.13 (d, 1H, J = 11.7 Hz, PhCH_2), 3.98 (d, 1H, J = 11.7 Hz, PhCH_2), 3.88–3.80 (m, 2H, H-3, H-5), 3.67 (ddd, 1H, J = 6.4, 6.2, 1.8 Hz, H-6), (s, 3H, OCH_3), 3.36–3.32 (m, 2H, H-4, H-7b), 2.88 (dd, 1H, J = 9.4, 6.8 Hz, H-5), 2.46 (s, 3H, CH_3Ph); ^{13}C NMR (125 MHz, $CDCl_3$): δ 144.8, 138.3, 138.1, 138.0, 137.6, 134.8, 129.7, 129.5, 128.8, 128.5, 128.46, 128.44, 128.42, 128.0, 127.8, 127.73, 127.7, 127.66, 127.6, 127.2, 91.3 (C-1), 86.7 (C-3), 79.3 (C-4), 77.8 (C-2), 76.8 (C-5), 76.4 (C-6), 75.84 (*C*H_2PH), 75.8 (*C*H_2PH), 75.6 (*C*H_2PH), 75.0 (*C*H_2PH), 73.0 (*C*H_2PH), 68.6 (C-7) 58.3 (O*C*H_3), 21.7 (CH_3Ph); HRMS (ESI) m/z Calcd for $C_{43}H_{46}O_7Na$ [M+Na]: 729.2856. Found:

729.2856. For the sulfone product: HRMS (ESI) m/z Calcd for $C_{43}H_{46}O_8Na$ [M+Na]: 745.2806. Found: 745.2798.

3.5 References

(1) Poulin, M. B.; Nothaft, H.; Hug, I.; Feldman, M. F.; Szymanski, C. M.; Lowary, T. L. *J. Biol. Chem.* **2010**, *285* (1), 493–501.

(2) Snitynsky, R. B.; Lowary, T. L. *Org. Lett.* **2014**, *16* (1), 212–215.

(3) Bernatchez, S.; Szymanski, C. M.; Ishiyama, N.; Li, J.; Jarrell, H. C.; Lau, P. C.; Berghuis, A. M.; Young, N. M.; Wakarchuk, W. W. *J. Biol. Chem.* **2005**, *280* (6), 4792–4802.

(4) Dalrymple, S. a.; Ko, J.; Sheoran, I.; Kaminskyj, S. G. W.; Sanders, D. a R. *PLoS One* **2013**, *8* (10), 1–15.

(5) Zhang, Q.; Liu, H. *J. Am. Chem. Soc.* **2000**, *122* (38), 9065–9070.

(6) Soltero-Higgin, M.; Carlson, E. E.; Gruber, T. D.; Kiessling, L. L. *Nat. Struct. Mol. Biol.* **2004**, *11* (6), 539–543.

(7) Karlyshev, a. V.; Quail, M. a.; Parkhill, J.; Wren, B. W. *Gene* **2013**, *522* (1), 37–45.

(8) Errey, J. C.; Mann, M. C.; Fairhurst, S. A.; Hill, L.; McNeil, M. R.; Naismith, J. H.; Percy, J. M.; Whitfield, C.; Field, R. A. *Org. Biomol. Chem.* **2009**, *7* (5), 1009–1016.

(9) Lemieux, R. U.; Ratcliffe, R. M. *Can. J. Chem.* **1979**, *57* (10), 1244–1251.

(10) Liu, Q.; Tor, Y. *Org. Lett.* **2003**, *5* (14), 2571–2572.

(11) Tanaka, H.; Yoshimura, Y.; Jørgensen, M. R.; Cuesta-Seijo, J. A.; Hindsgaul, O. *Angew. Chem. Int. Ed. Engl.* **2012**, *51* (46), 11531–11534.

(12) Xia, L.; Lowary, T. L. *J. Org. Chem.* **2013**, *78* (7), 2863–2880.

(13) Palcic, M. M.; Heerze, L. D.; Pierce, M.; Hindsgaul, O. *Glycoconj. J.* **1988**, *5* (1), 49–63.

(14) Sen, V. D.; Golubev, V. a. *J. Phys. Org. Chem.* **2009**, *22* (April 2008), 138–143.

(15) Ma, Y.; Loyns, C.; Price, P.; Chechik, V. *Org. Biomol. Chem.* **2011**, *9* (15), 5573.

(16) De Mico, A.; Margarita, R.; Parlanti, L.; Vescovi, A.; Piancatelli, G. *J. Org. Chem.* **1997**, *62* (20), 6974–6977.

(17) Montalbetti, C. a G. N.; Falque, V. *Tetrahedron* **2005**, *61* (46), 10827–10852.

(18) Garegg, P. J.; Iversen, T.; Oscarson, S. *Carbohydr. Res.* **1976**, *50* (2), C12–C14.

(19) Boons, G.-J.; Hale, K. J. *Organic Synthesis with Carbohydrates*; John Wiley & Sons, 2008.

(20) Gildersleeve, J.; Pascal, R. A.; Kahne, D. *J. Am. Chem. Soc.* **1998**, *120* (24), 5961–5969.

(21) Wang, Z. In *Comprehensive Organic Name Reactions and Reagents*; John Wiley & Sons, Inc., 2010.

4 Cloning, Expression, and Preliminary Studies on Enzyme Function with Putative Glycosyltransferases

4.1 *C. jejuni* CPS genes: Initial studies and annotated gene function

The genome sequence of *C. jejuni* was elucidated by Parkhill and coworkers in 2000,[1] and the presence of homologues of conserved KPS genes led Karlyshev and coworkers to identify the presence of a CPS in *C. jejuni* shortly thereafter (Section 1.4.4).[2] The genes involved in CPS production are organized into gene clusters, consisting of a highly variable internal biosynthetic region, which is flanked by conserved regions involved in transport and assembly (Figure 4-1).[2] Despite overall variability, Karlyshev postulated that the internal region may contain some conserved genes, and he and coworkers developed a polymerase chain reacion (PCR)-based procedure for comparative sequence analysis of *C. jejuni* sequence clusters.[3]

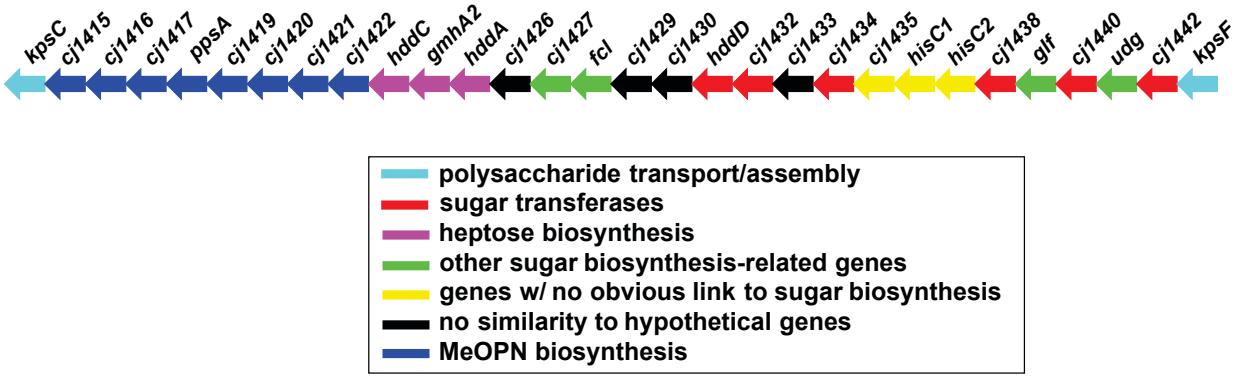


Figure 4-1: *C. jejuni* HS:2 CPS gene cluster.[2]

Since this discovery, extensive work has gone into annotating the individual genes in the *C. jejuni* CPS gene cluster. [1,2,4–7] However, biochemical studies in this area have been lacking. Insertional mutagenesis studies have identified only a specific subset of genes, and of those, the only GT identified was the heptosyl transferase (*cj1431*). A summary of the knockout studies conducted, and functional annotations based on CPS phenotype, are listed below:

Gene Knockout	CPS Phenotype	Functional Annotation	Reference
cj1421	loss of MeOPN on Gal*f*NAc	MeOPN transferase	McNally *et al.*[8]
cj1422	loss of MeOPN on Hep	MeOPN transferase	McNally *et al.*[8]
cj1423	loss of Hep	Heptose guanosyltransferase	Karlyshev *et al.*[2]
cj1424	loss of Hep	Sedoheptulose isomerase	Karlyshev *et al.*[2]
cj1425	loss of Hep	Heptose kinase	Karlyshev *et al.*[2]
cj1426	loss of 6-OMe on Hep	Methyltransferase	Sternberg *et al.*[9]
cj1427	loss of Hep	GDP-heptose epimerase	Sternberg *et al.*[9]
cj1428	loss of Hep	GDP-heptose epimerase	St. Michael *et al.*[10]
cj1430	loss of Hep	GDP-heptose epimerase	Sternberg *et al.*[9]
cj1431	loss of Hep	Heptosyltransferase	Karlyshev *et al.*[2]
cj1439	loss of CPS	UDP-Gal/GalNAc pyranose-furanose mutase	Poulin *et al.*[11]
cj1441	loss of CPS	UDP-Glc 6-dehydrogenase	Sternberg *et al.*[9] St. Michael *et al.*[10]

Table 4-1: Proposed functions of CPS genes based on phenotypic analysis of gene knockouts

In addition to the *cj1431* gene, sequence homology analysis of capsule-related genes has led to the identification of five additional GT candidates: *cj1432*, *cj1434*, *cj1438*, *cj1440*, and *cj1442* (Table 4-2).

ORF	Homologue or source	Score	Activity
cj1442c	SacD *Neisseria meningitidis*	2e-7	Glycosyltransferase
cj1440c	EpsI *Streptococcus agalactiae*	7e-19	Galactosyltransferase
cj1438c	EpsI *Streptococcus agalactiae*	4e-16	Galactosyltransferase
cj1434c	EpsI *Streptococcus pneumoniae*	4e-21	Galactosyltransferase
cj1432c	KfaA *Escherichia coli*	2e-4	Glycosyltransferase

Table 4-2: Putative glycosyltransferases by sequence homology. (Note: homology score < 10^{-4} shows significant sequence alignment.)[9]

Mutagenesis of any of these genes results in complete loss of CPS, which would be expected if they are involved in adding the sugar residues in the backbone of the repeating unit. Therefore, identifying the specific transformation attributable to the protein produced by each gene requires a biochemical approach. In my case, having chemically synthesized UDP-Gal*f*NAc (**1-1**) and acceptor substrates **1-2**–**1-4**, allows us to probe the specificity of each putative GT. The ability of a particular enzyme to transfer Gal*f*NAc from UDP-Gal*f*NAc to an acceptor, would allow its identification as the Gal*f*NAc transferase.

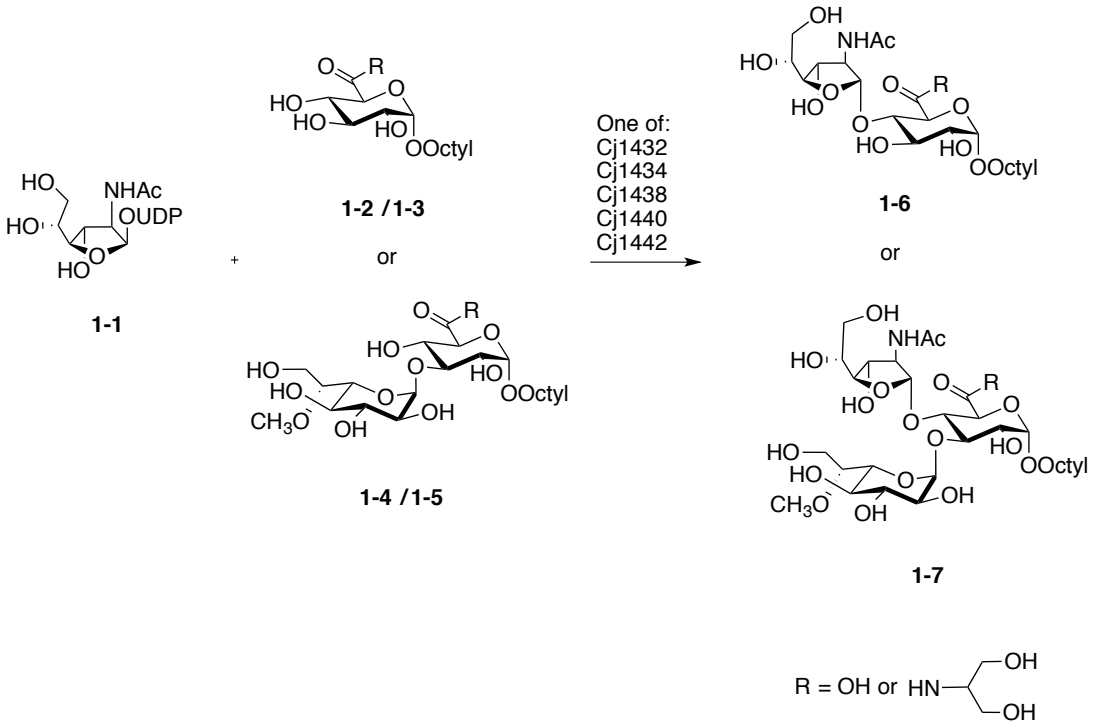


Figure 4-2: Reaction of donor **1-1** with acceptors **1-2**–**1-5** with putative GTs to monitor UDP-Gal*f*NAc transferase activity.

In order to carry out this biochemical work, however, the putative GT candidates must be cloned and overexpressed in soluble form. The current chapter will describe progress towards this end, and initial results on the assays performed with the over-expressed and purified gene products.

4.2 Cloning *cj1432*, *cj1434*, *cj1438*, *cj1440*, and *cj1442* into expression vectors [§]

Our aim was to make the N and C-terminal His-6 tagged clones for each of the five putative GTs. The addition of His_6 tags allows for the purification of the expressed gene products by immobilized metal affinity chromatography (using a Ni^{2+}–NTA (Nickel–Nitriloacetic acid)

[§] The cloning work was completed under the aid and supervision of Cory Q. Wenzel, a technician in Dr. Christine Szymanski's laboratory in the Department of Biological Sciences at the University of Alberta.

agarose column). The reason for attempting to make both N- and C-terminally tagged versions was to give us insurances against the possibility of the His_6 tag affecting the expression, function, or solubility of the protein, or accessibility of the tag when on one end vs the other. The varying effect of N vs C terminal His_6 tags has been observed in the past with WbpO.[12] Table 4-3 provides a list of the putative GTs, the size of the genes that encode them (in base pairs), their predicted sizes (in kilodaltons), and their theoretical *p*I values.

Gene	Number of Base Pairs (bp)	Size of Expected Protein (kDa)	*p*I
cj1432	3096	121.8	8.95
cj1434	1338	52.9	9.17
cj1438	2331	92.5	8.84
cj1440	1224	49.2	8.53
cj1442	1635	65.1	8.67

Table 4-3: Number of base pairs, molecular weight of final protein, and corresponding *p*I of putative GTs to be cloned and expressed. ProtParam (http://web.expasy.org/protparam/).

The pET28a vector was the backbone for the N-terminal expression constructs, and C-terminal expression constructs were generated using the pET30a vector. The *cj1432*, *cj1434*, *cj1438*, *cj1440* and *cj1442* genes were amplified from genomic DNA (*C. jejuni* strain NCTC11168) by PCR using the following primers and restriction sites (note that the restriction site sequence is underlined in each primer):

Primer	Vector	Sequence (5'→3')	Restriction Site	T_m (°C)
cj1432	pET28/30a	GAAAGTCTACCATGGTAAAAAAAATTCTTC	NcoI	62
His_6-*cj1432*	pET28a	GAATTACGAGCTCTTATCCTTTCTTAAATTTC	SacI	62
cj1421-His_6	pET30a	GAATTATCGCGGCCGCTCCTTTCTTAAATTTC	NotI	60
cj1434	pET28/30a	GGATGAAACCATGGATTATAATACTCC	NcoI	62
His_6-*cj1434*	pET28a	CAAACACCCACTCGAGTTATATCATCTTATTC	XhoI	62
cj1434-His_6	pET30a	CAAATCTCGAGTATCATCTTATTCTTATATTC	NcoI	62
cj1438	pET28/30a	CAAGGATCATATGATGAATTATAATAC	NdeI	62
His_6-*cj1438*	pET28a	GTTTTTCACTCGAGTTATTGGCTTGCATTATC	XhoI	62
cj1438-His_6	pET30a	GTTTTTGATCGCTCGAGTTTGCTTGCATTATC	XhoI	62
cj1440	pET28/30a	GGAAATTCATATGAAAACTGTAGGTG	NdeI	62
His_6-*cj1440*	pET28a	CTAATACATGCTCGAGTTACCTGAAAAATAAATATC	XhoI	62
cj1440-His_6	pET30a	CTAATTATATTCTCTCGAGCCTGAAAAATAAATATC	XhoI	62
cj1442	pET28/30a	GAAATGTCATATGCCTCTTTTATCTG	NdeI	62
His_6-*cj1442*	pET28a	CAATGACTCGAGTCATATTTTACCTTTC	XhoI	62
cj1442-His_6	pET30a	CAACAATGACATCTCGAGTATTTTACCTTT	XhoI	62

Table 4-4: Primers used to construct N and C terminal clones for *cj1432*, *cj1434*, *cj1438*, *cj1440* and *cj1442*.
*Nucleotides underlined indicate restriction endonuclease sites.

Initial failures in obtaining clean PCR products—due to multiple bands observed when screening the clones—were overcome by using a combination of hot start with touchdown PCR (see Section 4.4.2). This is a modified form of the PCR that aims to avoid non-specific binding of the primer oligonucleotide sequence, and therefore non-specific DNA amplification. The initial annealing temperature (Step 2, Figure 4-3) is higher than the optimal Tm of the primers, and this is incrementally lowered over subsequent cycles until a "touchdown temperature," or optimal

annealing temperature, is reached. By using temperatures higher than the calculated T_m (general procedures section 4.4.1), only high primer-template complementarity amplicons are extended, thereby decreasing the incidence of off-target binding of the primers. The subsequent lowering of the annealing temperature, over iterative cycles, increases the yield of the PCR reaction, where now the desired amplicons out-compete any non-specific gene products

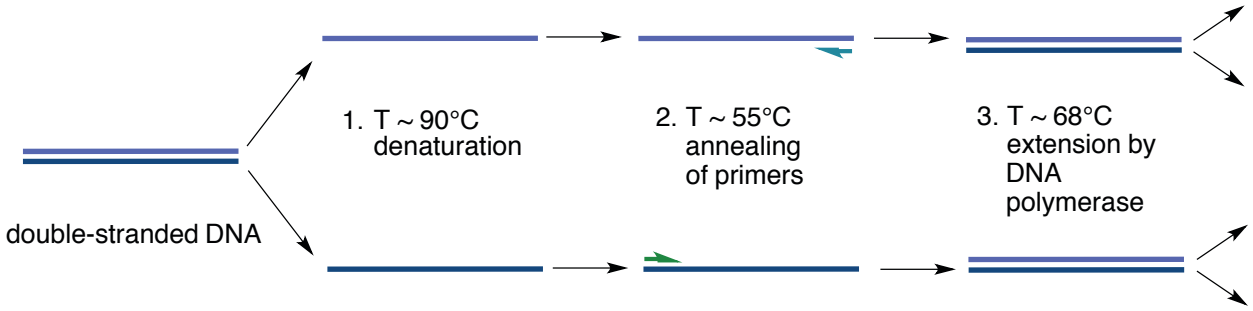


Figure 4-3: Steps in the polymerase chain reaction (PCR): 1. Denaturation of double stranded DNA into single strands, 2. Lowering of temperature allowing annealing of primers (at this stage, primers are prone to non specific binding), 3. Raising of the temperature allowing DNA chain extension.

Following PCR, digestion, band isolation, ligation, transformation, and isolation of plasmid DNA from resultant colonies was performed (Sections 4.6.1–4.6.3). Analytical digests were performed on plasmids isolated from resultant colonies after transformation (Section 4.6.3), and putative expression constructs for *His_6-cj1438* , *His_6-cj1440*, *cj1440-His_6*, *His_6-cj1442*, and *cj1442-His_6* were identified that correlated with the expected size of the gene inserts (Figure 4-4). DNA sequencing (Section 4.4.4) confirmed the correct sequences for *His_6-cj1438* *His_6-cj1440*, *cj1440-His_6*, *His_6-cj1442*, expression constructs, while all of the *cj1442-His_6* constructs isolated contained a premature stop codon (Note: His_6-tag before the gene name indicates an N terminal His_6 tag, and His_6-tag after the gene name indicates a C terminal His_6-tag*)*.[13]

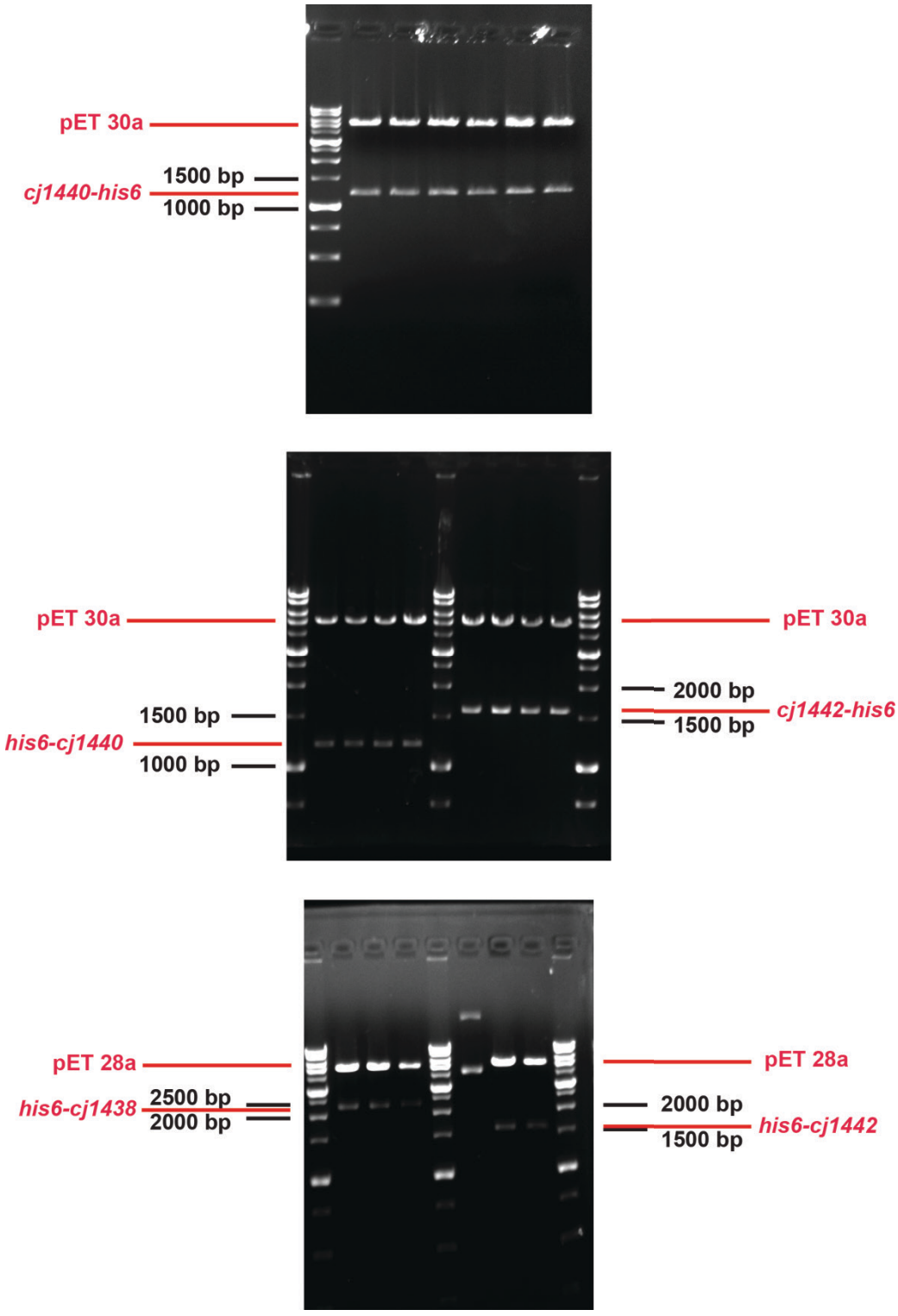


Figure 4-4: Analytical digests for clones of *cj1440-His*$_6$/pET30a, *His*$_6$*-cj1440*/pE28a, *cj1442-His*$_6$/pET30a, *His*$_6$*-cj1438*/pET28a, and *His*$_6$*-cj1442*/pET28a (separated by DNA ladders when on same gel; multiple bands reflect each set of PCR reaction).

Unfortunately, cloning attempts for *cj1432* and *cj1334* cloning were unsuccessful, regardless of the position of this His6-tag, and sequencing showed point mutations in each. It is possible that *cj1432*, being significantly larger than the other genes (>3000 bp), would benefit from longer extension times during the PCR reaction, as well as the use of a less error-prone polymerase. Therefore, if constructing clones for this gene are attempted again, it may be more advantageous to use a high fidelity proofreading polymerase such as Q5, Pfu, or platinum Pfx, versus PWO Polymerase used in these studies.

4.3 Expression of *cj1440* and *cj1442*.

With clones for *cj1438, cj1440,* and *cj1442* in hand, our focus turned towards expressing their gene products. As described in Chapter 1, a recent *in silico* study by Muggleton and coworkers has proposed that the most likely candidate for Gal*f*NAc transferase is *cj1440*, with *cj1442* being a strong alternate possibility.[14] For this study, they combined bioinformatics analysis with the insertional mutagenesis studies done previously, to give a hypothesized annotation of gene function. Given these findings, we chose to first focus on the expression of *cj1440* and *cj1442*.

Initial attempts at expressing both these proteins after transformation into BL21(DE3) competent *E. coli* cells were unsuccessful. The expected protein band was not observed on the SDS–PAGE (sodium dodecyl sulfate–polyacrylamide gel electrophoresis) gel or anti-His_6 Western immunoblots. Experimenting with growth and purification conditions including growth media (going from Luria–Bertani (LB) to the more nutrient dense Terrific Broth (TB)), increasing induction times, and altering the optical density (OD) at which cultures were induced, failed to increase expression levels. Under closer scrutiny, it was noted that these genes contained several

rare codons (codons for which there are few tRNAs in *E. coli* due to low frequency of usage). The use of rare codons in the *C. jejuni* cps gene cluster, up to this point, had not been reported.

As a result of these findings, we chose to express *cj1440-His$_6$* and *His$_6$-cj1442* in the BL21(DE3)pLacIRARE competent cell line. This cell line contains an additional vector that encodes for rare tRNAs and, when it was used, protein expression increased significantly. Moreover, it increased to the point that inclusion bodies were observed when using TB broth and growth temperatures of 37 °C (Figure 4-5). As a result, expression optimization studies were carried out on *cj1440-His$_6$*, and the following factors were compared: growth media (TB vs LB), incubation temperature (37 °C vs 20 °C), and induction time (1, 2, 3, and 24 h). Cells were collected after each of the growth cycle, lysed, and were subjected to Western immunoblotting (using an anti-His$_6$ monoclonal antibody (mAb) as the primary antibody).

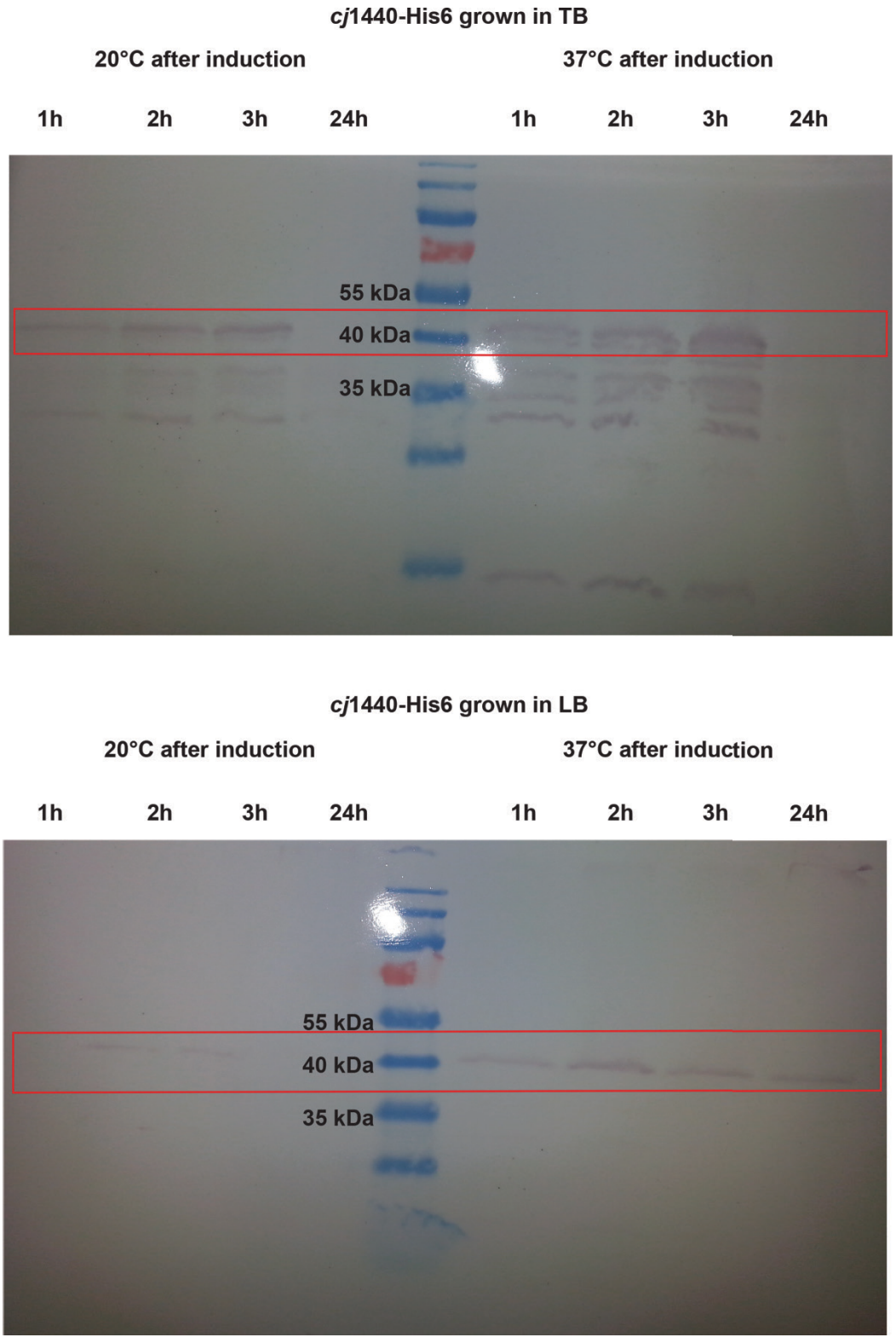


Figure 4-5: Anti-His6 Western immunoblotting for *cj1440-His6* expression optimization: Growth media (TB vs LB), incubation temperature (20 °C vs 37 °C), induction time (1, 2, 3, and 24 h).

To summarize, the optimized conditions decided upon for the expression of *cj1440-His$_6$* were growing the cells in LB growth medium at 37 °C until OD_{600nm}= 0.4–0.5, followed by induction with isopropyl 1-thio-β-D-galactopyranoside (IPTG) followed by a 2 h induction time at room temperature. In addition, a cold shock was applied to the cell culture directly before induction. It has been shown that this procedure increases the production of molecular chaperones, aiding in the proper folding of the protein, and overall reduction in the number of inclusion bodies.[15] Therefore, it is expected that protein solubility and stability would increase significantly.

The results of the protein expression are shown below (Figure 4-6). Using the pLacIRARE expression strain, in addition to our optimized conditions, significantly increased levels of soluble protein (prior to this, no bands at the expected molecular weight (MW) for the desired protein were seen). The optimized expression conditions for *cj1440-His$_6$*, when applied to the expression of *His$_6$-cj1442*, also displayed an increase in expression amounts.

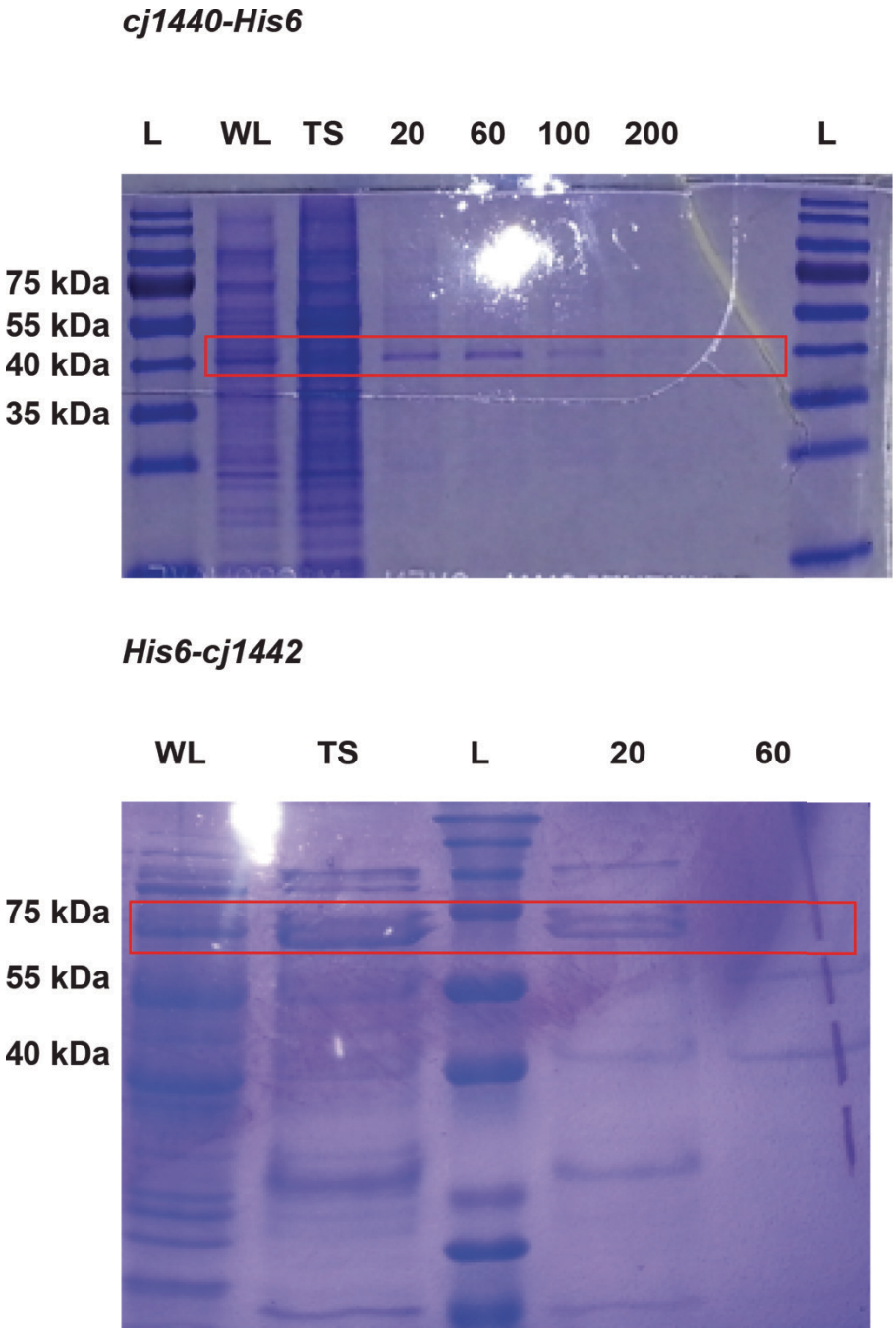


Figure 4-6: Top: Expression of *cj1440-His*$_6$, Bottom: Expression of *His*$_6$*-cj1442*. L = Ladder, WL = whole cell lysate, TS = total soluble fraction/flow through fraction, 20 = elution with 20 mM imidazole, 60 = elution with 60 mM imidazole, 100 = elution with 100 mM imidazole, 200 = elution with 200 mM imidazole.

Though the total amount of soluble protein expressed improved markedly, binding of the His6 tag to our affinity resin was not as successful. The proteins expressed from *His*$_6$*-cj1440* and

cj1440-His$_6$ (not shown) showed moderate retention on the Ni^{2+}–NTA column, with most of the protein eluting at 60 mM of imidazole in 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer, pH 7.5, containing 10% glycerol, and 150 mM NaCl. Despite this, a significant portion of total protein was lost in the flow through fraction, as well as low concentration eluting buffer (20 mM imidazole). In the case of *His$_6$-cj1442*, almost all of the desired protein was eluted in the flow through and 20 mM imidazole fractions, and none contained in the elution buffers, possibly due to the His_6 tag not being presented on the protein surface, thereby preventing it from binding to the Ni^{2+}–NTA resin.

In addition to losses in the purification steps, protein was also lost during the ultracentrifugation step. This step removes cell debris from the whole cell lysate, and these include membrane fractions as well as membrane associated proteins. This suggests that perhaps these proteins are membrane associated. *In silico* work (http://molbiol-tools.ca/ Online Analysis Tools, using secondary structural prediction analysis) showed no trans membrane helical domains. However, it is not unreasonable to assume that these proteins are in a complex near the cell membrane where the CPS is synthesized (more in Section 4.5).

Due to these findings, we decided to test our potential substrates on both the total soluble fractions as well as the crude whole cell lysate, to ensure we had adequate concentrations of protein required for our GT assay.

4.4 GT Assay: Preliminary Results

The spectrophotometric assay we used for proteins produced from *cj1440-His$_6$* and *His$_6$-cj1442* were adapted from Rose *et al.* and is summarized below and in Figure 4-7.[16]

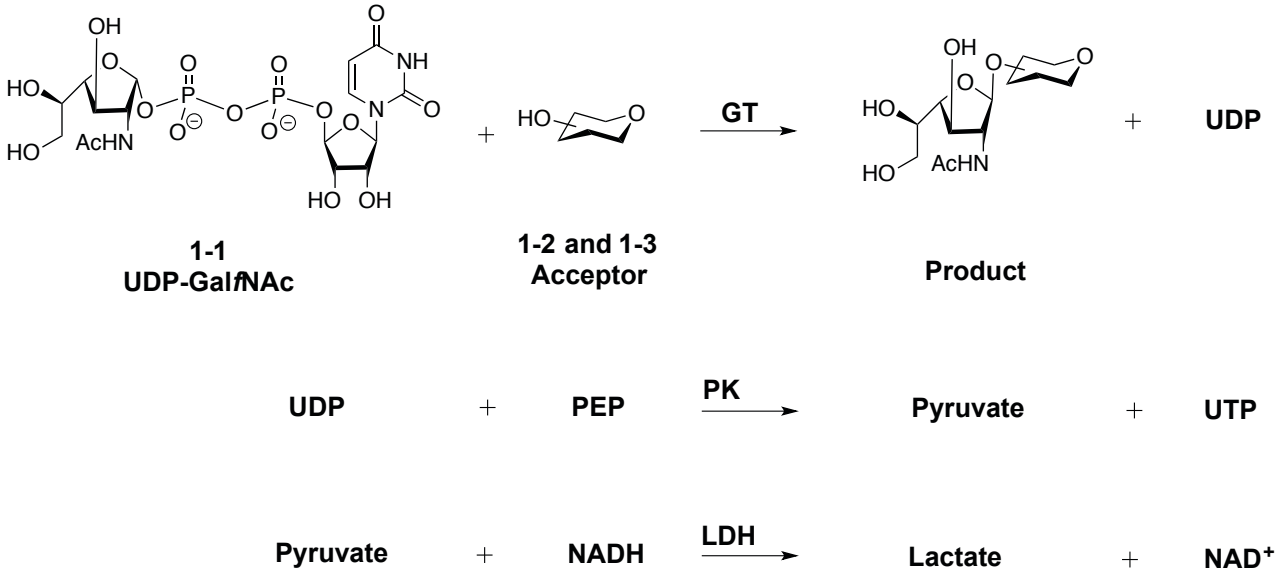


Figure 4-7: Spectophotometric assay used to monitor Gal*f*NAc-transferase activity; adapted from *Rose et al*.[16]

This assay is dependent upon the transfer of a monosaccharide (here Gal*f*NAc) from a sugar nucleotide (**1-1**) onto one of our acceptor substrates (**1-2** or **1-3**). It should be noted that we used only the monosaccharide substrates in these preliminary studies because at the time, the valuable disaccharides were not made in adequate amounts, and the conditions of the assay were not yet optimized. The UDP produced in the reaction is coupled to the oxidation of β-nicotinamide adenine dinucleotide (NADH) by two enzymes, pyruvate kinase (PK), which converts phosphoenolpyruvate (PEP) into pyruvate, and lactate dehydrogenase (LDH), which converts pyruvate into lactate, while oxidizing NADH. The decrease in absorption at 340 nm, resulting from NADH oxidation to NAD^+, is proportional to the activity of the GT. The protein that shows

activity with these substrates in this assay will correspond to the GT that installs the Gal*f*NAc residue.

The assay was designed as detailed in Table 4-5 for both Cj1440-His_6 and His_6-Cj1442 (cells correspond to the wells in the microplate reader):

Controls	**Reaction with 1-2**	**Reaction with 1-3**
No Acceptor Control (Enz)	Enz + **1-2**	Enz + **1-3**
1-1 + **1-2**	Enz + **1-2**	Enz + **1-3**
1-1 + **1-3**	Enz + **1-2**	Enz + **1-3**
Positive Control (UDP)		
No Acceptor Control (CL)	CL + **1-2**	CL + **1-3**
	CL + **1-2**	CL + **1-3**
	CL + **1-2**	CL + **1-3**

Table 4-5: Setup for spectrophotometric assay. This setup was run for both the *cj1440-His_6* and *His_6-cj1442* gene products. Enz = Total Soluble Fraction, CL = Crude Whole Cell Lysate, No Acceptor Control = Enz + **1-1**, Positive Control = UDP

No acceptor controls, in which the donor (**1-1**) + enzyme were added, as well as negative controls (no enzyme added), were included. In addition, a positive control (where UDP was added) was included to test the assay. Triplicates were run with each of the enzyme preparations (whole cell lysate and soluble fraction) with both of the acceptor monosaccharides (**1-2** and **1-3**).

A decrease in UV absorption (340 nm) was not seen for any of the samples with crude whole cell lysate. The increased cell debris is most likely the cause, increasing the baseline reading, therefore failing to show an adequate drop in UV absorption due to positive activity. Further

supporting this is that the positive control, with the whole cell lysate, also failed to show a reading.

The soluble enzyme fractions, however, showed the expected negative results with both the no acceptor controls, as well as the no enzyme controls. In addition, the assay showed the expected activity with the positive control. The protein produced from *cj1440-His$_6$* did not show a significant decrease in absorption compared to our negative controls. However, the protein produced from *His$_6$-cj1442,* showed a decrease in absorbance, correlating to UDP release, and therefore Gal*f*NAc transferase activity, with both **1-2** and **1-3**. Furthermore, the decrease in absorbance seen with donor **1-1** + *cj1442 His$_6$* + **1-2** and **1-1** + *His$_6$-cj1442* + **1-3**, mimicked our positive control values, further supporting UDP release, and positive enzyme activity (Figure 4-8).

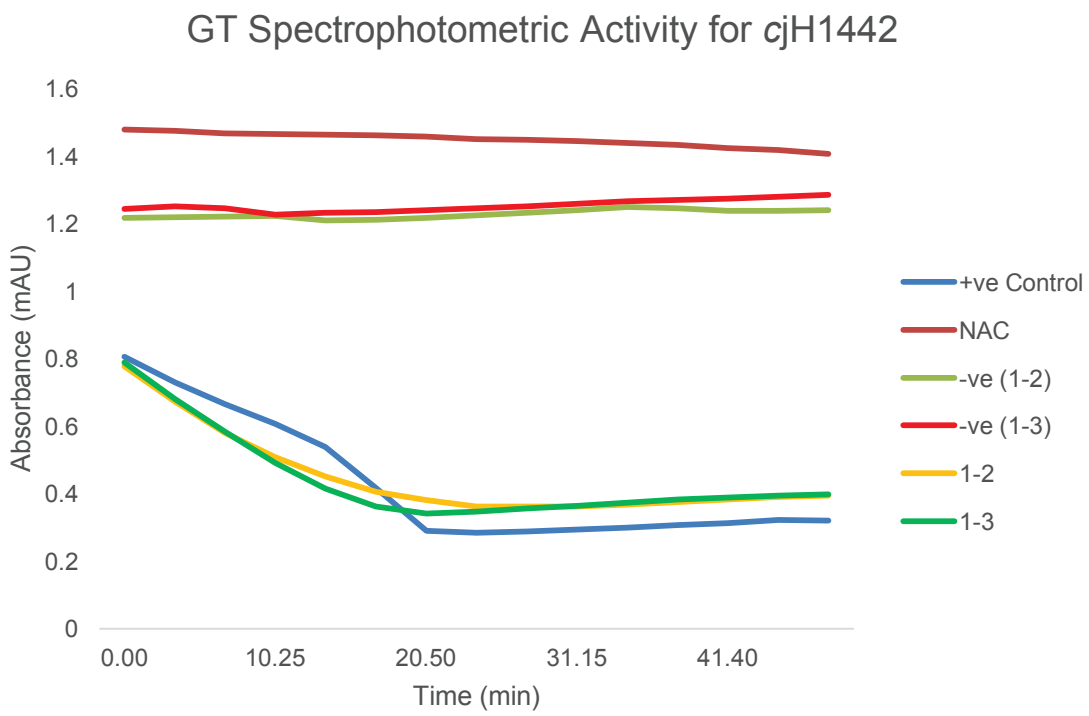


Figure 4-8: Spectophotometric activity for *cjH1442*. +ve Control = positive control with UDP added; NAC = no acceptor control (*cj1442* + donor **1-1**); –ve (**1-2**) = negative control (no enzyme) with **1-2**; –ve (**1-3**) = negative control (no enzyme) with **1-3**; **1-2** = donor **1-1** + enzyme + **1-2**; **1-3** = donor **1-1** + enzyme + **1-3**.

Although positive activity was observed seen by the spec assay for *cj1442-His*$_6$ with **1-3**, upon submitting the isolated assay mixtures for MS, the desired product **1-6** was not observed. In the future, this reaction will need to be scaled up to obtain sufficient amounts of product, after C_{18} Sep-pack purification, to confirm the presence of product. LC-MS should also be investigated at as an alternative to manual C_{18} Sep-pack purification, to detect the presence of the desired product with minimal losses at the purification stage.

4.5 Conclusions and Future Directions

This chapter highlights progress towards identifying the putative Gal*f*NAc transferase involved in the biosynthesis of the CPS in *C. jejuni* 11168. Progress has been made toward obtaining the clones of the putative GTs, and those that have been completed and correctly sequenced are *His_6-cj1338*, *His_6-cj1440*, *cj1440-His_6*, and *His_6-cj1442*. We experienced a difficult hurdle in not detecting any levels of expression for *cj1440- His_6* and *His_6-cj1442*. However, growth condition optimization, use of cold shock induction, in addition to switching to a rare codon usage strain, pLacRARE, overcame this problem. Though these proteins were observed in the soluble fraction, there was a significant amount seen in the membrane fraction, and, of the total soluble protein, only a fraction bound with high affinity to the Ni^{2+}–NTA column. The later made purification a challenge; moreover, a large proportion of total protein expressed was lost in the purification step. Both of these issues will have to be addressed in optimizing the expression of *cj1440* and *cj1442*, as obtaining pure protein will be required if any kinetic evaluation is conducted for positive Gal*f*NAc transferase activity.

GTs are a notoriously difficult class of enzymes to express in soluble form. One particular hurdle is that those involved with transmembrane export machinery, in this case the ABC transport system, often form a coordinated and associated complex localized on the cell membrane. The likelihood of these proteins being membrane associated is high, and this fact may require the co-expression of accessory proteins for them to adopt their native fold and to show activity.[17] Therefore, expecting CPS-associated GTs to express individually in soluble form, and show high levels of activity, may be a tall order in that they may not be able to adopt their native fold. Despite the difficulty that this class of enzyme pose, there are examples where glycosyltransferases have been found to express and show activity in soluble form.[18–21]

In addition to investigating methods to express the putative GTs with their native folded conformations, the enzyme assay conditions will also need to be optimized. Those used in this study were adapted from the study of GlfT2 (a UDP-galactofuranosyltransferase from mycobacteria)[16], and more studies will be needed regarding experimental conditions, such as optimal buffers and salt concentrations, temperature, and time. If this fails, there are a number of other strategies for probing glycosyltransferase activity that could be explored, such as colorimetric assays,[22] phosphatase coupled assays,[23,24] glycosyl transferase activity kits (R&D Systems Inc., Minneapolis, MN), or bioluminescent assay using UDP-Glo^{TM} (Promega, Madison, WI). Ultimately, however, these reactions must be carried out on miligram scale (including appropriate purification), to confirm the presence of GT products by MS and ^{1}H NMR spectroscopy.

In addition, we have not yet assessed the activity of the *cj1432*, *cj1434*, and *cj1438* gene products. It may very well be the case that the proteins expressed from *cj1440* and *cj1442* do not recognize UDP-Gal*f*NAc as a substrate. Clones for *cj1432* and *cj1434* must still be sequenced (constructs have already been made), and *cj1438* expression conditions must still be optimized. When complete, the disaccharides, which were not shown in these studies, may prove to be better substrates than the monosaccharides to facilitate turnover.

Despite these shortfalls, a considerable amount of progress has been made in attempting to obtain the putative GTs. The enzyme that have been expressed, and though not conclusively showing activity for Gal*f*NAc transferase, may still be used to elucidate the ribosyltransferase or

GlcA transferase activity required for CPS biosynthesis, with their appropriate substrates. For example, other members of the group are currently working on substrates for identifying the ribosyltransferase activity, and the clones that I have constructed, as well as optimized expression conditions can aid in this endeavor. In addition, the substrates for Gal*f*NAc have been synthesized and are ready for testing once the three remaining putative GTs are expressed.

4.6 Experimental

4.6.1 General Cloning Procedures

Plasmid DNA was isolated using Qiaprep Spin Miniprep Kit (Qiagen Inc., Mussussauga, ON) as described by manufacturers. DNA restriction endonucleases were used as recommended by the manufacturer (Invitrogen Canada Inc.). Ligation of restriction endonuclease-digested vectors and insert DNA was performed using T4 DNA ligase (New England Biolabs, Ltd., Mississauga, ON) at 15 °C for 16 h prior to transformation into *E. coli* strains. Cloning of PCR products into pET28 and pET30a was performed using the corresponding cloning kits (Novagen, Madison, WI) as described by the manufacturer. Purification of DNA from agarose gels was performed using the UltraClean 15 DNA Purification Kit (Mo Mio Laboratories, Inc., Solana Beach, CA).

4.6.2 PCR Amplification:

The oligonucleotide primers were designed and analyzed using DNA analysis software GENERUNNER for Windows (Hastings Software, NY; www.generunner.com), based on the nucleotide sequences from the complete *C. jejuni* NCTC11168 genome. The oligonucleotide primers used are listed in Table 4-4. All restriction endonucleases were obtained from New

England Biolabs (Mississauga, ON, Canada). PCR-amplification reactions were performed with a Gene Amp 2400 PCR System thermocycler (PerkinElmer Canada, Woodbridge, ON) using Taq DNA polymerase (Roche Diagnostics Corporation, Indianapolis IN) according to the manufacturers specifications. All PCR reactions were performed with HPLC-grade water (Fischer Scientific Ltd, Nepean, ON).

Hot start with touchdown PCR reaction: The template DNA was denatured by incubation at 95 °C for 5 min with the forward primer, prior to addition of Taq DNA polymerase. The reaction then cycled through a series of denaturation (90 °C), annealing (68 °C), and polymerization/extension (68 °C) steps for 3 cycles (2.5 min extension time), before adding the reverse primer. Lastly, the reverse primer was added, and the reaction cycled through a series of denaturation (90 °C), annealing (68 °C to 50 °C), and polymerization/extension (68 °C) steps for 27 cycles (2.5 min extension time), followed by a final 7 min polymerization/extension (68 °C) step.

PCR reactions were analyzed by agarose gel electrophoresis as follows: 5 μL of DNA tracking dye (0.1% bromophenol blue, 15% ficoll) was loaded into a 0.8%–1.0% agarose-TAE gel (0.8%–1.0% agarose, 20 mM acetic acid, 1 mM ethylenediaminetetraacetic acid (EDTA), 40 mM tris(hydroxymethyl)aminomethane (Tris) and electrophoresed at 80–100 V. Gels were stained in a 0.002% ethidium bromide solution and visualized under ultraviolet (UV) light. Products were purified using QIAquick PCR Purification Kit (Qiagen Inc) prior to restriction endonuclease digestion.

4.6.3 Screening Potential Clones

Screening of potential clones for those containing plasmid DNA with the desired insert was performed by the E-lyse method.[25] Colonies to be screened were patched onto the appropriate selective media and incubated overnight at 37 °C. Patched cells were scraped using a sterile loop and resuspended in 15 μL of TE buffer (1 mM EDTA, 10 mM Tris-HCl, pH 8.0) and combined with 15 μL of SL lysis solution (25% sucrose, 2 units/mL RNAse, and 1 mg/mL lysozyme in TBE buffer (90 mM boric acid, 2.5 mM EDTA, 90 mM Tris-HCl). Samples were then loaded into a 1% agarose–TBE gel (1% agarose, 0.2% SDS in TBE buffer) that was submerged in TBE buffer (45 mM Tris-borate, 1 mM EDTA, pH = 8.3), and allowed to stand in the wells for 10 min. Electrophoresis was performed at 120 V for 2–3 h. The gel was stained in a 0.002% ethidium bromide bath for 20 min, and then visualized under UV light.

4.6.4 DNA Sequencing

DNA samples were prepared as described above (4.6.1) and DNA sequencing was performed by The Molecular Biology Service Unit (MBSU), Department of Biological Sciences, University of Alberta. The resultant DNA sequences were analyzed using the analyzed using DNA analysis software GENERUNNER for Windows (Hastings Software, NY; htttp://www.generunner.com), and DNA- and protein-database searches were performed using the National Centre for Biotechnology Institute (NCBI) Basic Local Alignment Search Tool (BLAST) network server.

4.6.5 General Protein Methods

Unless otherwise, stated all reagents were obtained from Sigma–Aldrich Canada Ltd. (Oakville, ON), all media were obtained from Difco Laboratories (Detroit, MI), and all solutions were

prepared using water purified by a Milli-Q PF Ultra-Pure Water System (Millipore (Canada) Ltd., Missisauga, ON). The bacterial strains and plasmids were used as listed in Table 4-4. Where appropriate, media were supplemented with antibiotics at the following concentrations: chloramphenicol (Cm) 34 μg/mL, kanamycin (Km) 50 μg/mL.

4.6.6 Expression of *cj1440-His$_6$* and *His$_6$-cj1442*

Newly transformed E. coli Rosetta^{TM} (DE3)2pLacI^{TM} competent cells (Novagen Biosciences, La Jolla, CA) carrying recombinant plasmids were cultured while shaking at 200 rpm overnight at 30 °C in 5 mL of LB broth supplemented with Kan and Cm at the appropriate concentrations. Cultures were expanded to a volume of 1 L and the OD_{600nm} regularly monitored.

Overexpression of the target protein was achieved by inducing the *E. coli* culture with 0.1 mM IPTG (Invitrogen, Burlington, ON) at OD_{600nm} of 0.4–0.5 after the starter culture was laid to rest in an ice bath for 10 min. This was followed by incubating the culture for a further 2 h at 23 °C. Cells were harvested at 4 °C by centrifugation at 11,300g_{max} and the pellets were stored –20 °C until further use. Upon use, the cell pellets were thawed on ice and resuspended at a ratio of 1:10 in ice-cold 50 mM HEPES buffer, pH 7.5, containing 10% glycerol, and 150 mM NaCl. All subsequent steps were performed at 4 °C.

Lysates were prepared by a double passage of the resuspension mix through a pre-chilled bench top cell disrupter (Constant Systems Incorporated, Sanford, NC) set to 35,000 psi. The cell lysate was centrifuged at 105,000g for 60 min to remove cellular debris and other insoluble proteins. The supernatant fraction was applied directly to a pre-equilibrated Ni^{2+}–NTA-agarose affinity

column using gravity flow. The crude soluble fraction was combined with the resin Ni^{2+}–NTA-agarose resin and gently mixed for 30 min. Unbound proteins were washed from the column using 15 bed volumes of binding buffer, and weakly eluting proteins were washed using 15 bed volumes of 20mM imidazole. The overexpressed protein was eluted with 4 bed volumes each of 60 mM, 100 mM, and 200mM imidazole. Evaluation by sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS–PAGE) indicated the presence of a band with the expected protein MW (Figure 4-6).

4.6.7 Spectrophotometric assays

Potassium chloride (KCl), NADH as the reduced disodium salt, PEP, PK (type III, lyophilized powder, rabbit muscle), and LDH (type XI, salt free, rabbit muscle) were from Sigma–Aldrich (St. Louis, MO). 3-(*N*-morpholino)propanesulfonic acid (MOPS), BioUltraPure grade, was from BioShop (Burlington, ON), and magnesium chloride hexahydrate was obtained from EMD Biosciences (La Jolla, CA).

Spectrophotometric assays were performed in 384-array microtiter plate wells containing 50 mM MOPS, pH 7.9, 50 mM KCl, 20 mM $MgCl_2$, 1.1 mM NADH, 3.5 mM PEP, 7.5 U pyruvate kinase (PK), and 16.8 U lactate dehydrogenase (LDH). MOPS, pH 7.9, was added as a 20-fold stock solution. KCl, $MgCl_2$, and the acceptor and donor substrates were dissolved in deionized distilled water. All other assay components were prepared in 50 mM MOPS, pH 7.9, with the exception of PEP, which, due to its acidity, was buffered in 250 mM MOPS, pH 7.9. Stock solutions of NADH, PEP, PK, LDH, and donor were made fresh on the day of use. A standard assay reaction contained UDP-Gal*f*NAc (**1-1**) at a final concentration of 3 mM and acceptor

monosaccharide (**1-2** or **1-3**) at 2 mM. Each reaction was initiated by the addition of 0.75 μL *cj*1442 and *cj*1440 to the assay mixture. The final assay volume was 40 μL.

Reactions were continuously monitored at 37 °C using an Epoch 2 Microplate Spectophotometer (BioTek, VN, USA) in the kinetics read mode. The wavelength was set at 340 nm and each reaction was monitored at 3.43 s intervals for up to 50 min. Data were graphed and analyzed using GraphPad Prism Version 4.

4.7 References

(1) 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.; Whitehead, S.; Barrell, B. G. *Nature* **2000**, *403* (6770), 665–668.

(2) Karlyshev, A. V.; Linton, D.; Gregson, N. a.; Lastovica, A. J.; Wren, B. W. *Mol. Microbiol.* **2000**, *35* (3), 529–541.

(3) 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* (1), 90–103.

(4) Champion, O. L.; Karlyshev, A. V; Senior, N. J.; Woodward, M.; La Ragione, R.; Howard, S. L.; Wren, B. W.; Titball, R. W. *J. Infect. Dis.* **2010**, *201* (5), 776–782.

(5) Karlyshev, a. V.; Quail, M. a.; Parkhill, J.; Wren, B. W. *Gene* **2013**, *522* (1), 37–45.

(6) Karlyshev, A. V; Wren, B. W. *J. Clin. Microbiol.* **2001**, *39* (1), 279–284.

(7) Guerry, P.; Szymanski, C. M. *Trends Microbiol.* **2008**, *16* (9), 428–435.

(8) 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* (39), 28566–28576.

(9) 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* (1), 186–197.

(10) Michael, F. St.; 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* (21), 5119–5136.

(11) Poulin, M. B.; Nothaft, H.; Hug, I.; Feldman, M. F.; Szymanski, C. M.; Lowary, T. L. *J. Biol. Chem.* **2010**, *285* (1), 493–501.

(12) Zhao, X.; Creuzenet, C.; Bélanger, M.; Egbosimba, E.; Li, J.; Lam, J. S. *J. Biol. Chem.* **2000**, *275* (43), 33252–33259.

(13) Korbie, D. J.; Mattick, J. S. *Nat. Protoc.* **2008**, *3* (9), 1452–1456.

(14) 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* (1), 186–197.

(15) Qing, G.; Ma, L.-C.; Khorchid, A.; Swapna, G. V. T.; Mal, T. K.; Takayama, M. M.; Xia, B.; Phadtare, S.; Ke, H.; Acton, T.; Montelione, G. T.; Ikura, M.; Inouye, M. *Nat. Biotechnol.* **2004**, *22* (7), 877–882.

(16) Rose, N. L.; Zheng, R. B.; Pearcey, J.; Zhou, R.; Completo, G. C.; Lowary, T. L. *Carbohydr. Res.* **2008**, *343* (12), 2130–2139.

(17) Wenzel, C. Q.; St Michael, F.; Stupak, J.; Li, J.; Cox, A. D.; Richards, J. C. *J. Bacteriol.* **2010**, *192* (1), 208–216.

(18) Rose, N. L.; Completo, G. C.; Lin, S.-J.; McNeil, M.; Palcic, M. M.; Lowary, T. L. *J. Am. Chem. Soc.* **2006**, *128* (20), 6721–6729.

(19) Lauber, J.; Handrick, R.; Leptihn, S.; Dürre, P.; Gaisser, S. *Microb. Cell Fact.* **2015**, *14*, 3.

(20) Kaulpiboon, J.; Prasong, W.; Rimphanitchayakit, V.; Murakami, S.; Aoki, K.; Pongsawasdi, P. *J. Basic Microbiol.* **2010**, *50* (5), 427–435.

(21) Engel, J.; Schmalhorst, P. S.; Krüger, A. T.; Müller, C. T.; Buettner, F. F. R.; Routier, F. H. *Glycobiol.* **2015**.

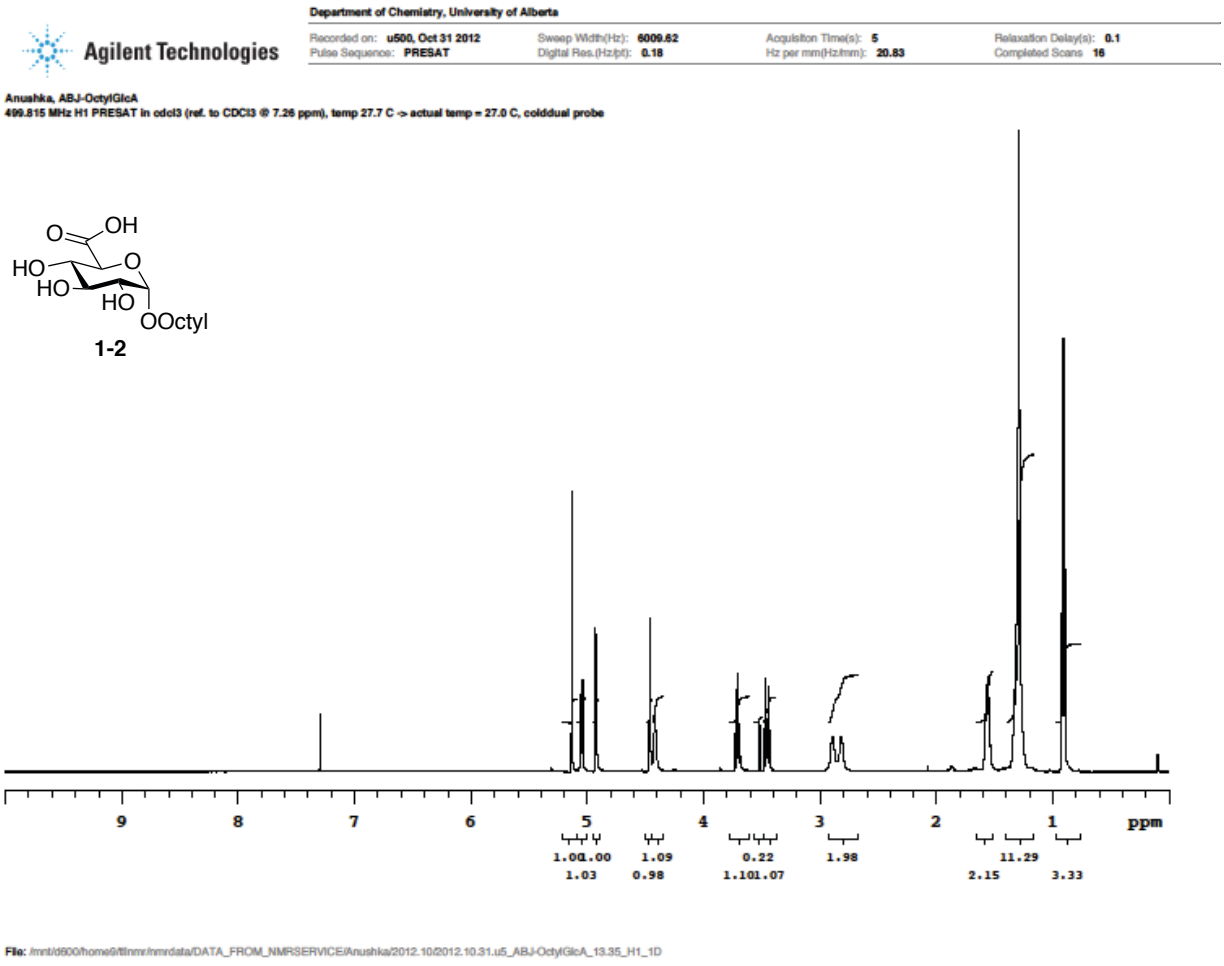
(22) Shen, R.; Wang, S.; Ma, X.; Xian, J.; Li, J.; Zhang, L.; Wang, P. *Biochem. Biokhimiia* **2010**, *75* (7), 944–950.

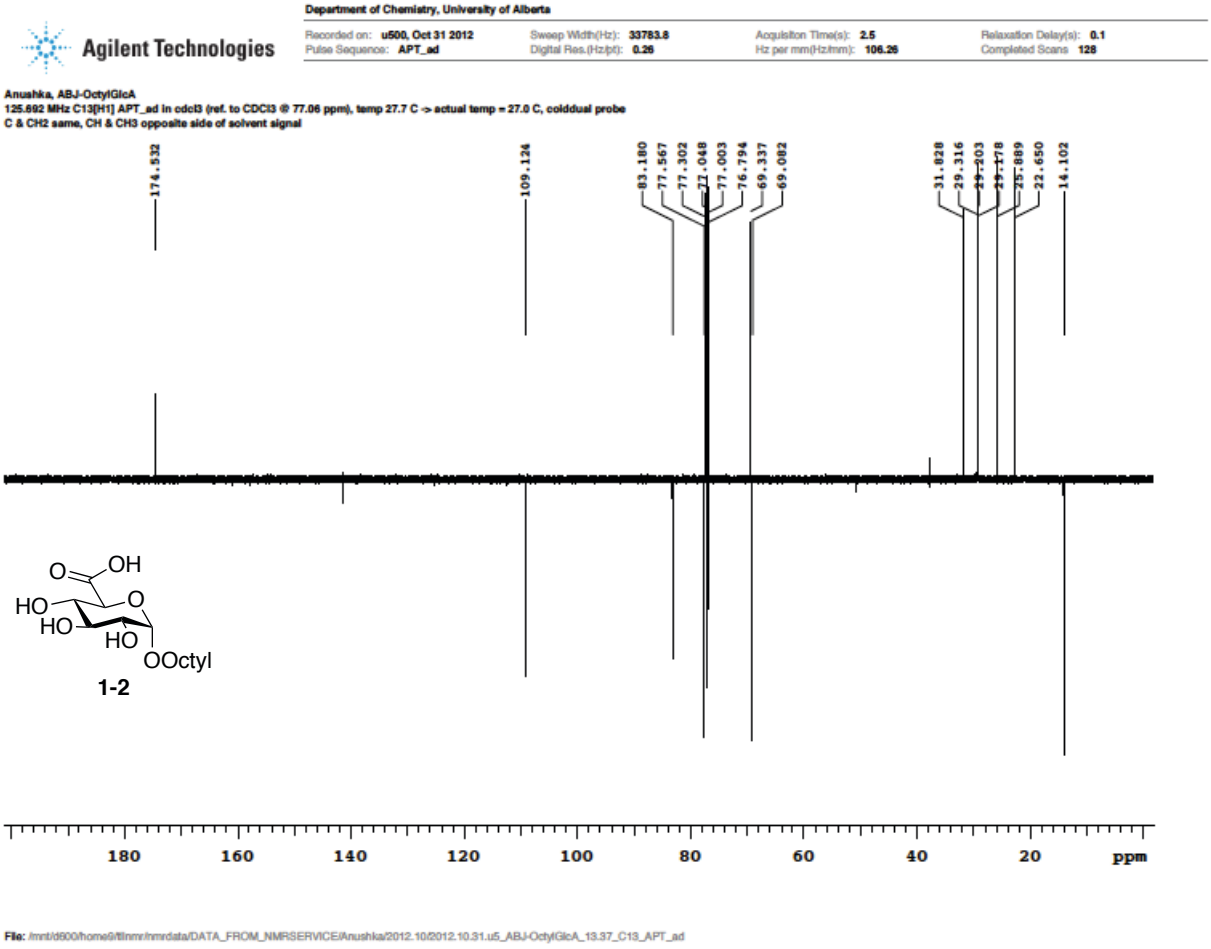
(23) Wu, Z. L.; Ethen, C. M.; Prather, B.; Machacek, M.; Jiang, W. *Glycobiology* **2011**, *21* (6), 727–733.

(24) Bubner, P.; Czabany, T.; Luley-Goedl, C.; Nidetzky, B. *Anal. Biochem.* **2015**, *490*, 46–51.

(25) Eckhardt, T. *Plasmid* **1978**, *1* (4), 584–588.

Appendix: Selected Copies of NMR Spectra





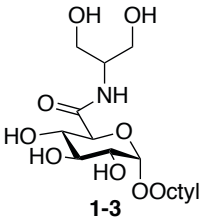
File: /mnt/d600/home9/tfinmr/nmrdata/DATA_FROM_NMRSERVICE/Anushka/2012.10/2012.10.31.u5_ABJ-OctylGlcA_13.37_C13_APT_ad

Department of Chemistry, University of Alberta

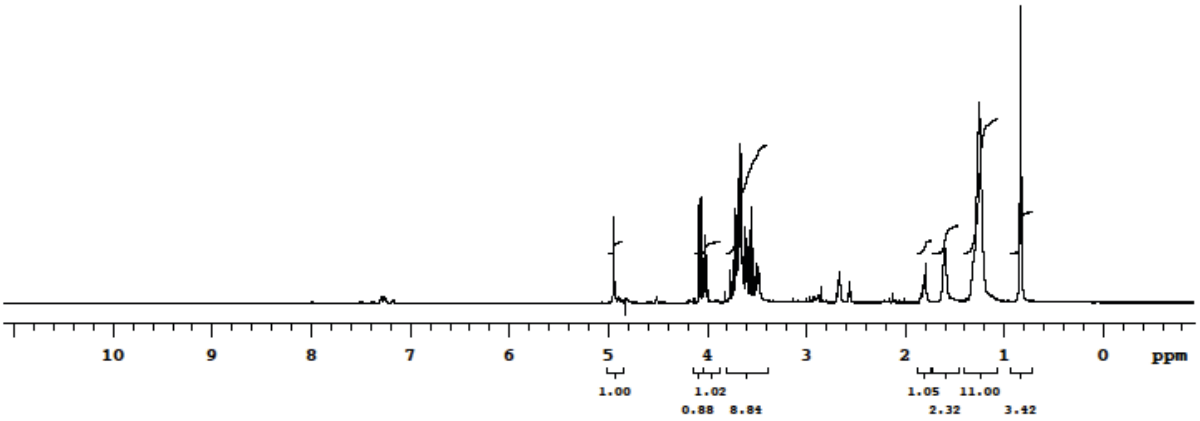
Agilent Technologies

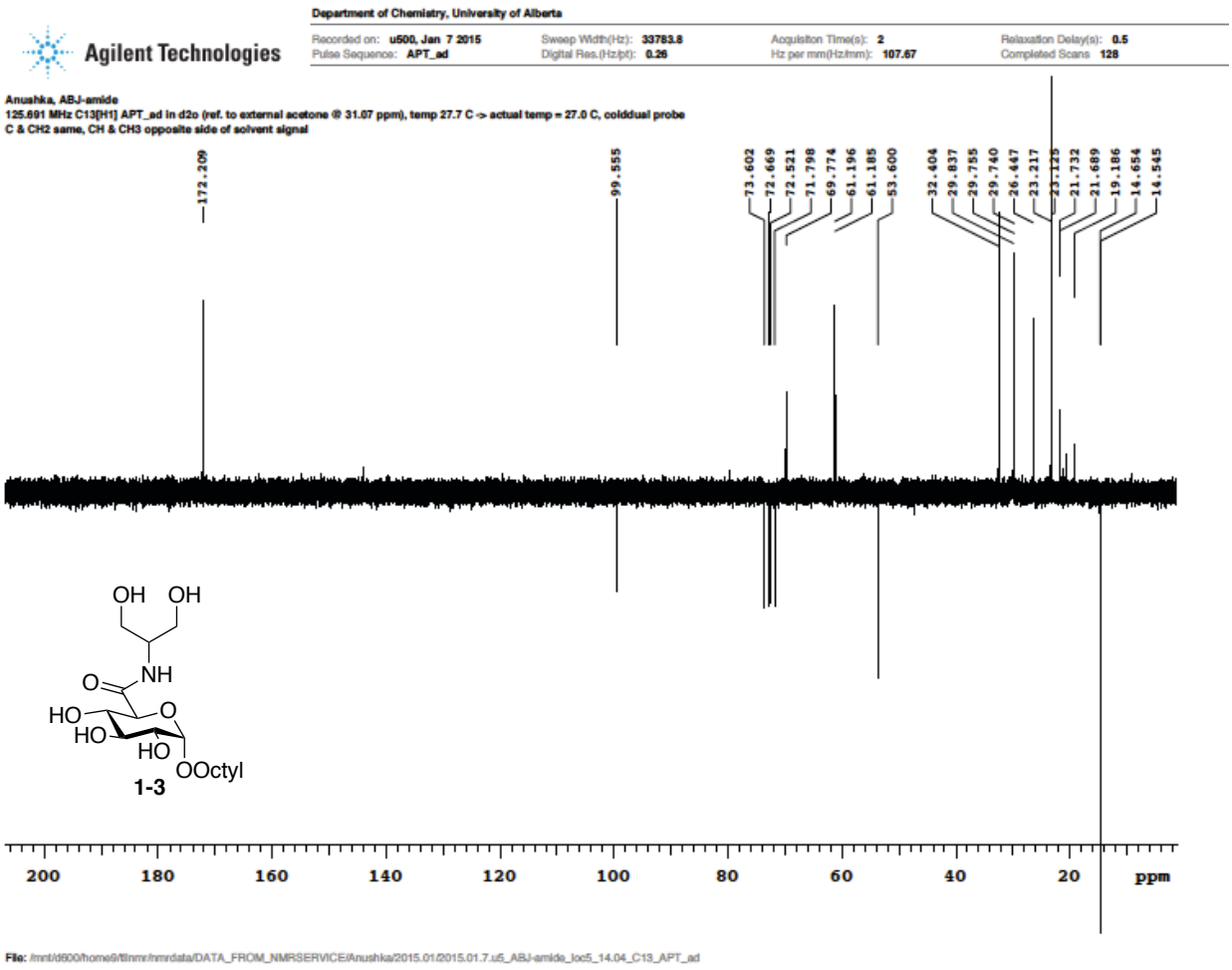
Recorded on: u500, Jan 7 2015	Sweep Width(Hz): 6009.62	Acquisiton Time(s): 5	Relaxation Delay(s): 0.1
Pulse Sequence: PRESAT	Digital Res.(Hz/pt): 0.18	Hz per mm(Hz/mm): 25.04	Completed Scans 16

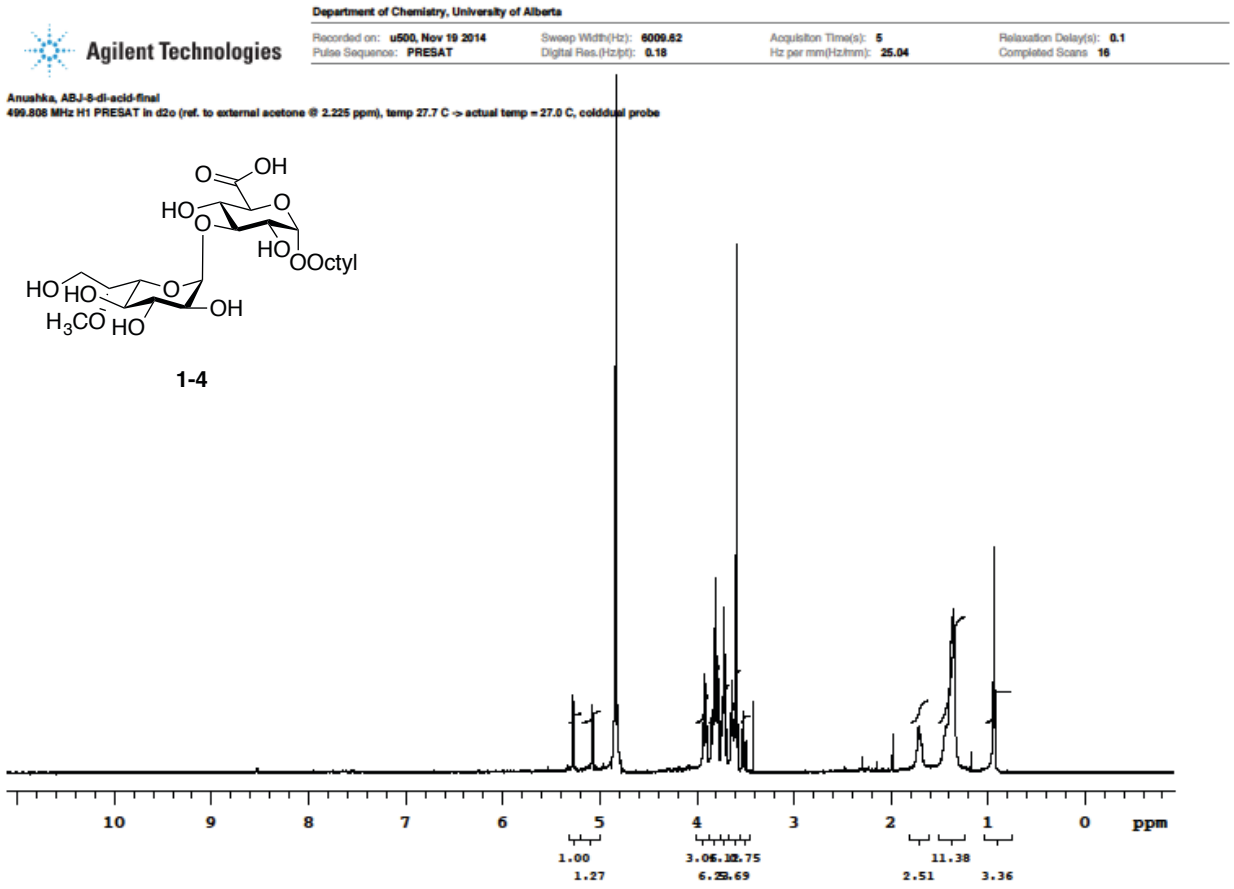
Anushka, ABJ-amide
499.808 MHz H1 PRESAT in d2o (ref. to external acetone @ 2.225 ppm), temp 27.7 C -> actual temp = 27.0 C, colddual probe



1-3




Department of Chemistry, University of Alberta
Agilent Technologies
Recorded on: u500, Jan 7 2015
Pulse Sequence: APT_ad
Sweep Width(Hz): 33783.8
Digital Res.(Hz/pt): 0.26
Acquisition Time(s): 2
Hz per mm(Hz/mm): 107.67
Relaxation Delay(s): 0.5
Completed Scans 128
Anushka, ABJ-amide
125.691 MHz C13[H1] APT_ad in d2o (ref. to external acetone @ 31.07 ppm), temp 27.7 C -> actual temp = 27.0 C, colddual probe
C & CH2 same, CH & CH3 opposite side of solvent signal
172.209
99.555
73.602
72.669
72.521
71.798
69.774
61.196
61.185
53.600
32.404
29.837
29.755
29.740
26.447
23.217
23.125
21.732
21.689
19.186
14.654
14.545
OH
OH
NH
O
HO
HO
HO
O
OOctyl
1-3
200
180
160
140
120
100
80
60
40
20
ppm
File: /mnt/d600/home9/ftlnmr/nmrdata/DATA_FROM_NMRSERVICE/Anushka/2015.01/2015.01.7.u5_ABJ-amide_loc5_14.04_C13_APT_ad

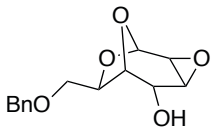

Department of Chemistry, University of Alberta
Agilent Technologies
Recorded on: u500, Nov 19 2014
Pulse Sequence: PRESAT
Sweep Width(Hz): 6009.62
Digital Res.(Hz/pt): 0.18
Acquisiton Time(s): 5
Hz per mm(Hz/mm): 25.04
Relaxation Delay(s): 0.1
Completed Scans 16
Anushka, ABJ-8-di-acid-final
499.808 MHz H1 PRESAT in d2o (ref. to external acetone @ 2.225 ppm), temp 27.7 C -> actual temp = 27.0 C, colddual probe
OH
O
HO
O
O
HO
OOctyl
HO
HO
O
OH
H3CO
HO
1-4
10
9
8
7
6
5
4
3
2
1
0
ppm
1.00
1.27
3.06
6.25
2.75
5.69
11.38
2.51
3.36

Department of Chemistry, University of Alberta

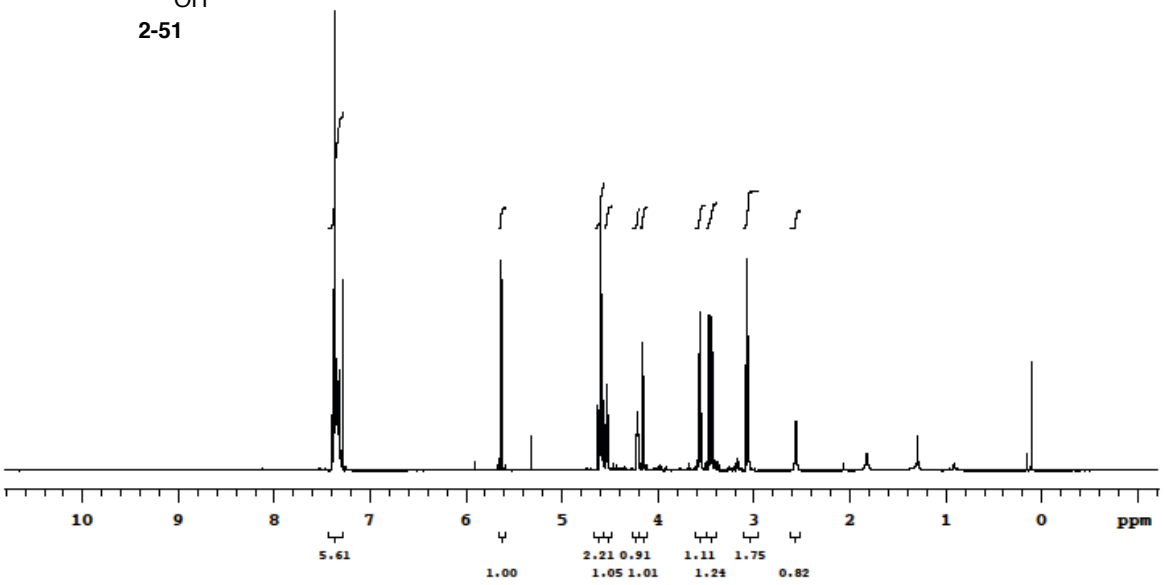
Agilent Technologies

Recorded on: u500, Apr 7 2014	Sweep Width(Hz): 6009.62	Acquisition Time(s): 5	Relaxation Delay(s): 0.1
Pulse Sequence: s2pul	Digital Res.(Hz/pt): 0.09	Hz per mm(Hz/mm): 25.04	Completed Scans 4

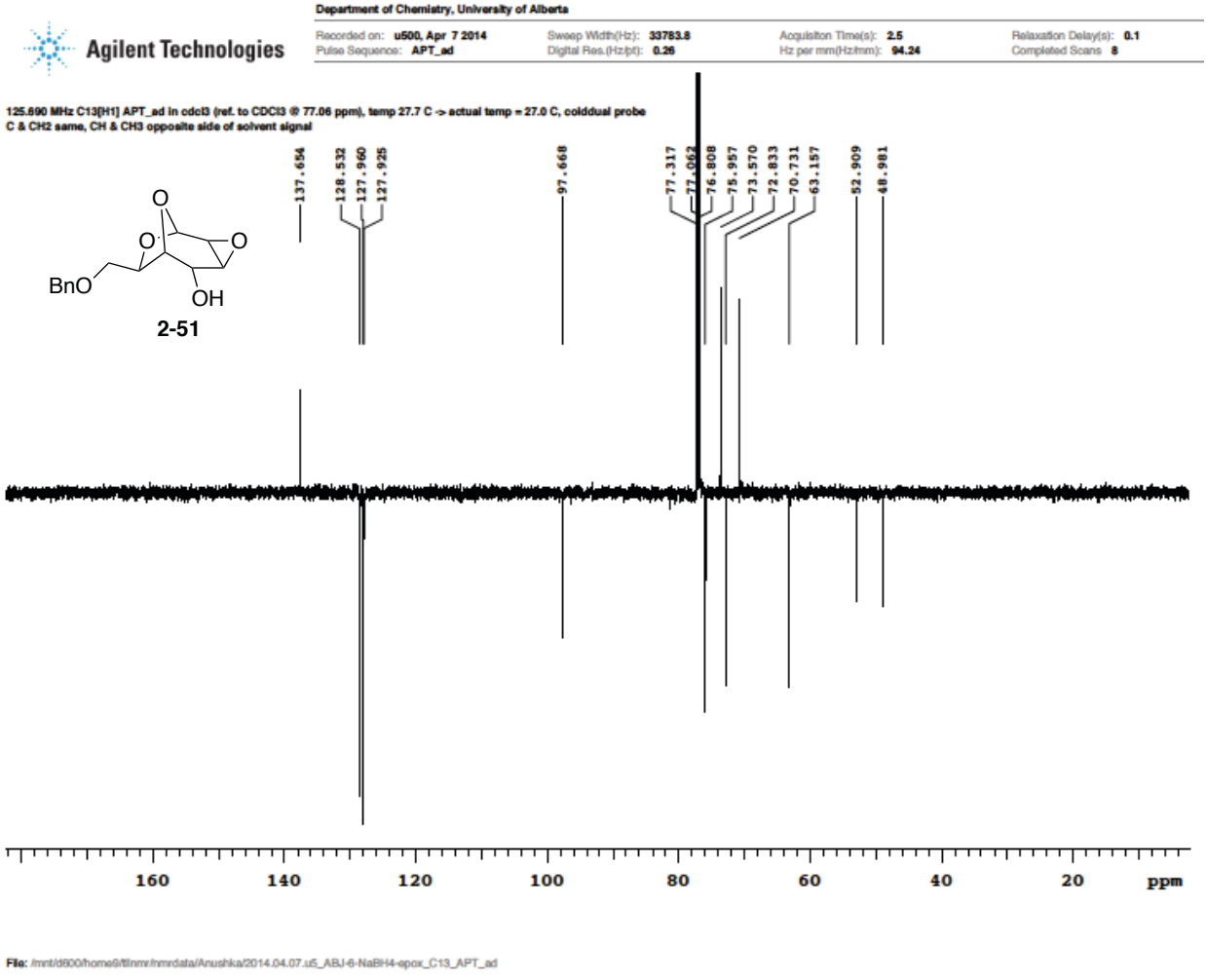
499.806 MHz H1 1D in cdcl3 (ref. to CDCl3 @ 7.26 ppm), temp 27.7 C -> actual temp = 27.0 C, colddual probe

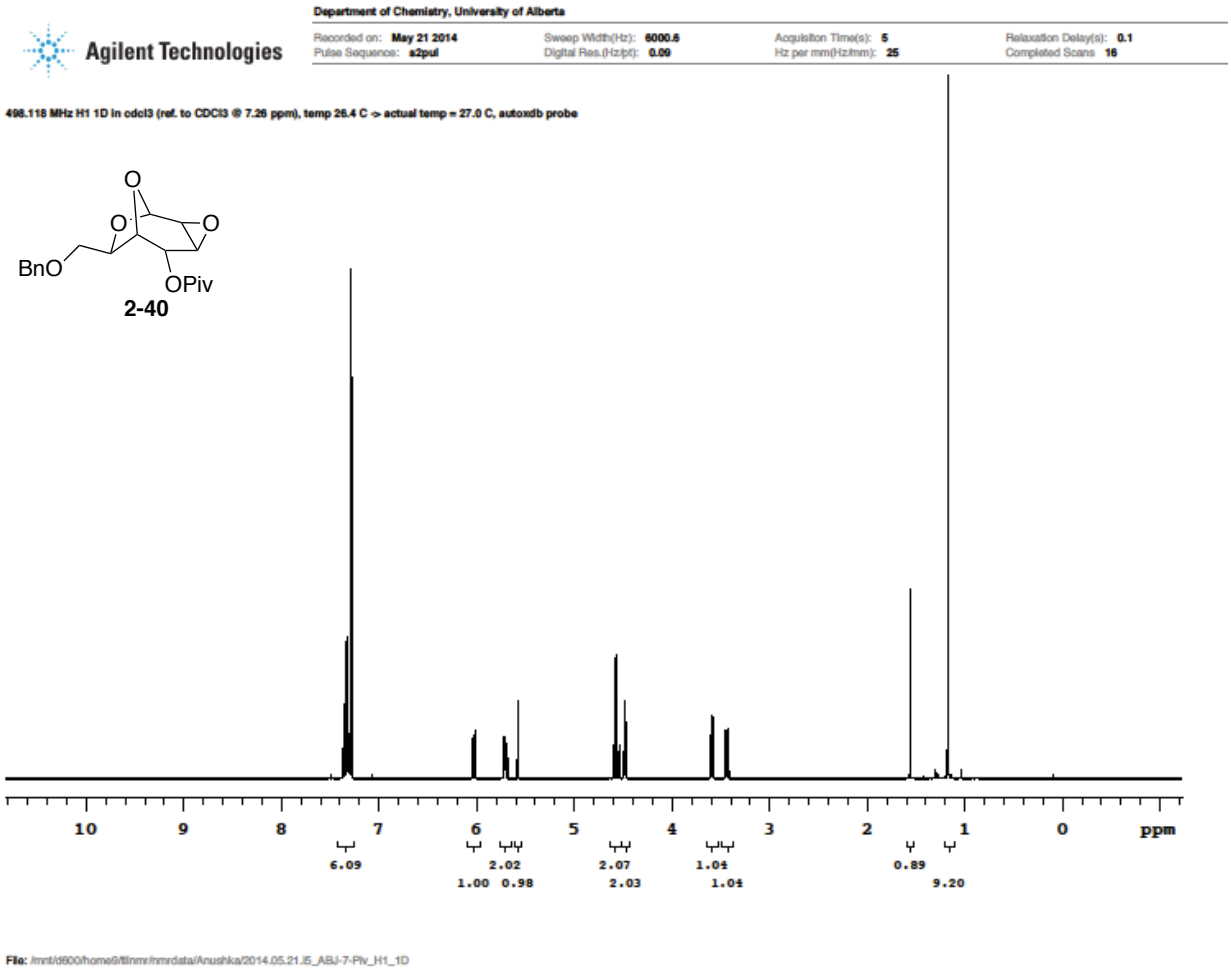


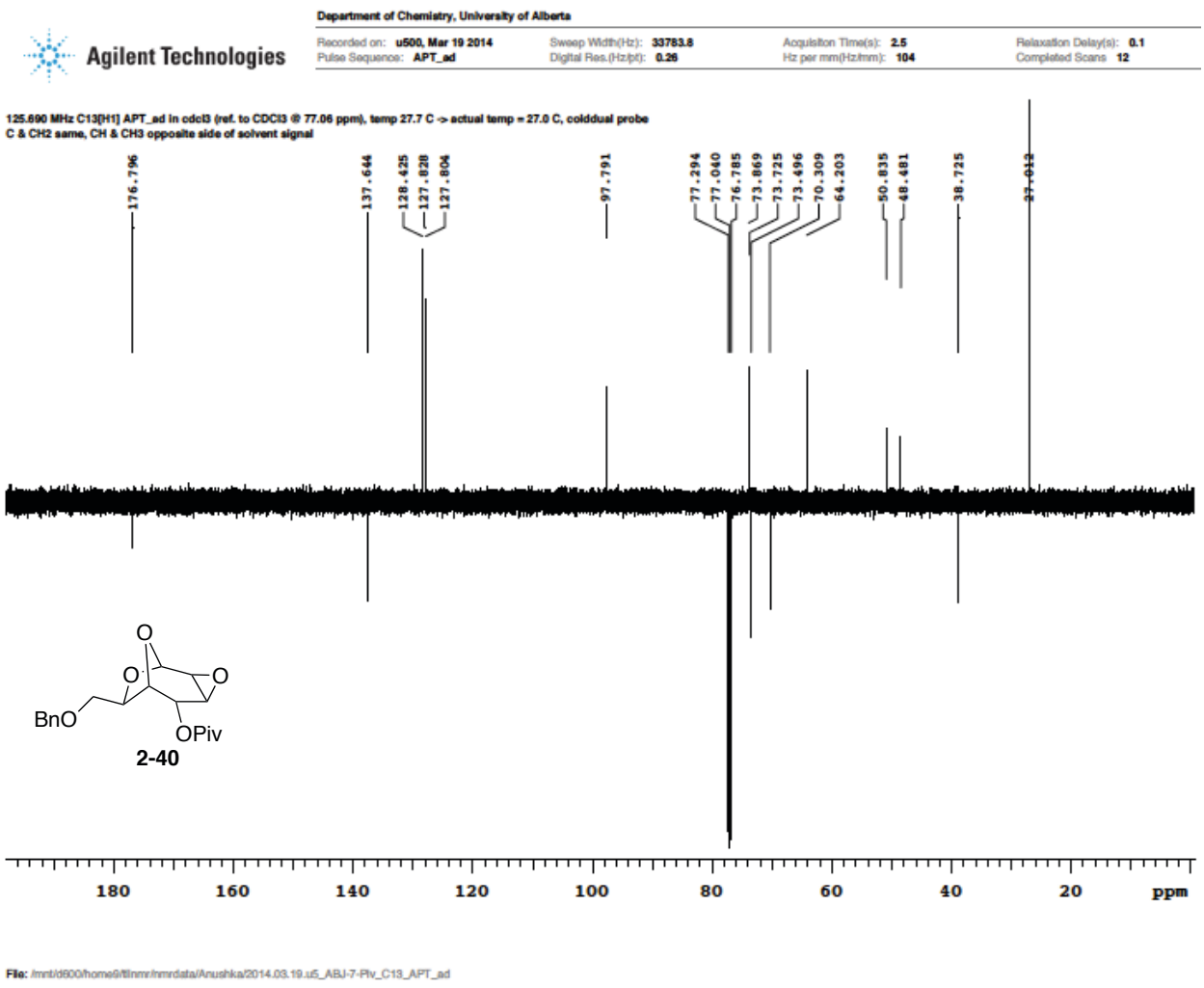
2-51

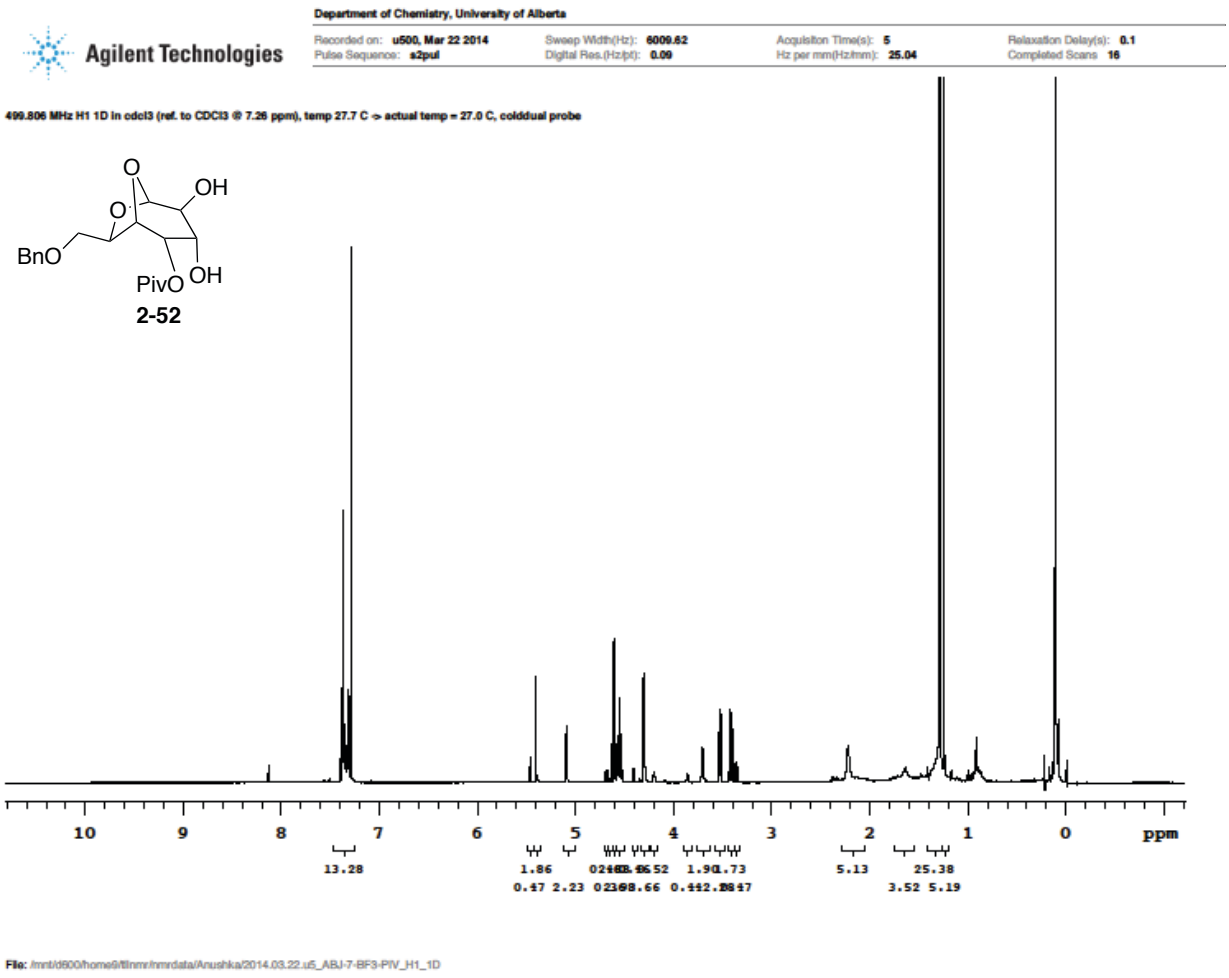


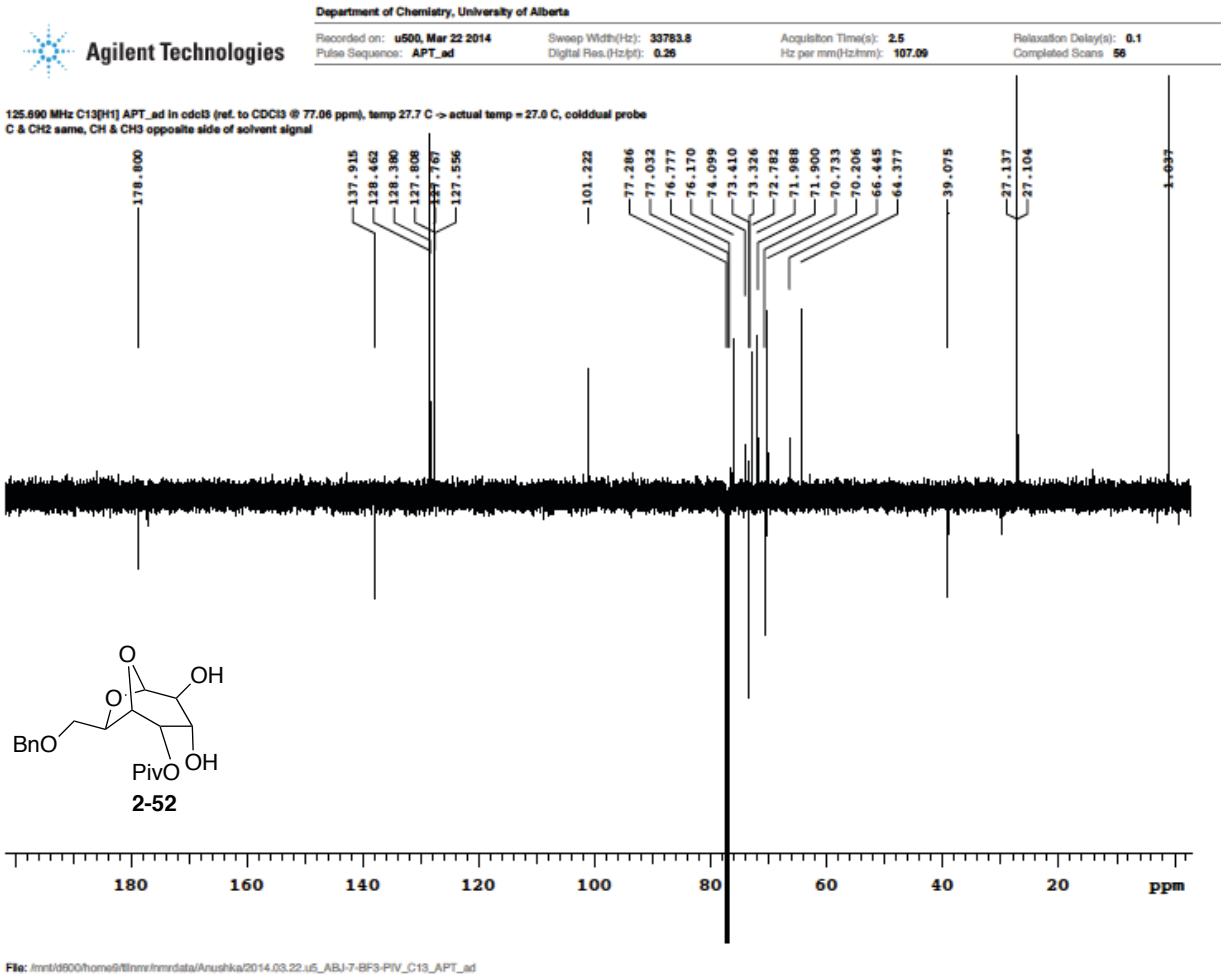
File: /mnt/d600/home9/tlnmr/nmrdata/Anushka/2014.04.07.u5_ABJ-6-NaBH4-epox_H1_1D

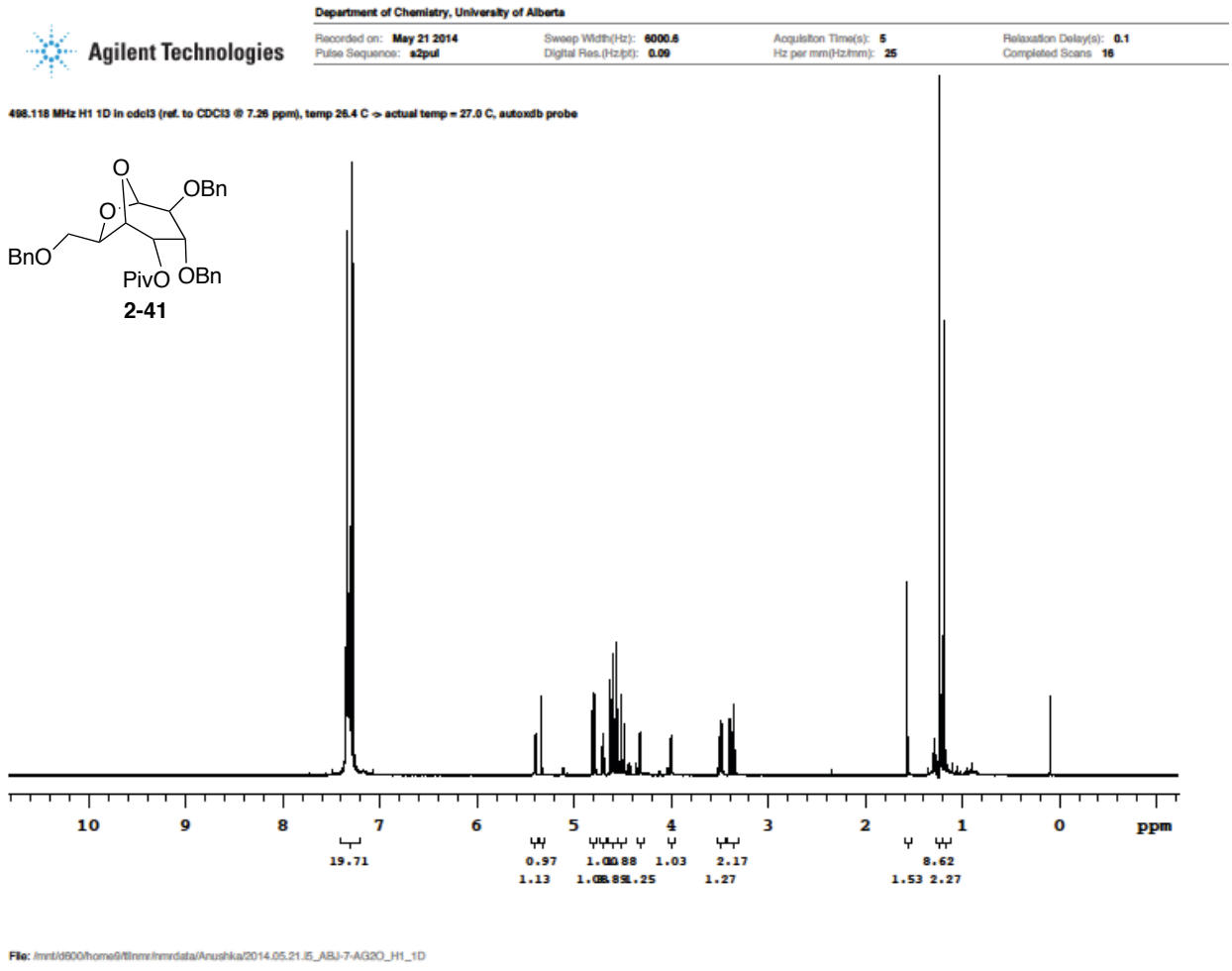

Agilent Technologies
Department of Chemistry, University of Alberta
Recorded on: u500, Apr 7 2014
Pulse Sequence: APT_ad
Sweep Width(Hz): 33783.8
Digital Res.(Hz/pt): 0.26
Acquisiton Time(s): 2.5
Hz per mm(Hz/mm): 94.24
Relaxation Delay(s): 0.1
Completed Scans 8
125.690 MHz C13[H1] APT_ad in cdcl3 (ref. to CDCl3 @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, colddual probe
C & CH2 same, CH & CH3 opposite side of solvent signal
137.654
128.532
127.960
127.925
97.668
77.317
77.062
76.808
75.957
73.570
72.833
70.731
63.157
52.909
48.981
O
O
O
BnO
OH
2-51
160
140
120
100
80
60
40
20
ppm
File: /mnt/d600/home9/tlnmr/nmrdata/Anushka/2014.04.07.u5_ABJ-6-NaBH4-epox_C13_APT_ad


Department of Chemistry, University of Alberta
Agilent Technologies
Recorded on: May 21 2014
Pulse Sequence: s2pul
Sweep Width(Hz): 6000.6
Digital Res.(Hz/pt): 0.09
Acquisition Time(s): 5
Hz per mm(Hz/mm): 25
Relaxation Delay(s): 0.1
Completed Scans 16
498.118 MHz H1 1D in cdcl3 (ref. to CDCl3 @ 7.26 ppm), temp 26.4 C -> actual temp = 27.0 C, autoxdb probe
O
O
O
BnO
OPiv
2-40
10
9
8
7
6
5
4
3
2
1
0
ppm
6.09
2.02
1.00
0.98
2.07
2.03
1.04
1.04
0.89
9.20
File: /mnt/d600/home9/tlnmr/nmrdata/Anushka/2014.05.21.I5_ABJ-7-Piv_H1_1D


Department of Chemistry, University of Alberta
Agilent Technologies
Recorded on: u500, Mar 19 2014
Pulse Sequence: APT_ad
Sweep Width(Hz): 33783.8
Digital Res.(Hz/pt): 0.26
Acquisition Time(s): 2.5
Hz per mm(Hz/mm): 104
Relaxation Delay(s): 0.1
Completed Scans 12
125.690 MHz C13[H1] APT_ad in cdcl3 (ref. to CDCl3 @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, colddual probe
C & CH2 same, CH & CH3 opposite side of solvent signal
176.796
137.644
128.425
127.828
127.804
97.791
77.294
77.040
76.785
73.869
73.725
73.496
70.309
64.203
50.835
48.481
38.725
27.013
O
O
O
BnO
OPiv
2-40
180
160
140
120
100
80
60
40
20
ppm
File: /mnt/d600/home9/tlnmr/nmrdata/Anushka/2014.03.19.u5_ABJ-7-Piv_C13_APT_ad


Agilent Technologies
Department of Chemistry, University of Alberta
Recorded on: u500, Mar 22 2014
Pulse Sequence: s2pul
Sweep Width(Hz): 6009.62
Digital Res.(Hz/pt): 0.09
Acquisition Time(s): 5
Hz per mm(Hz/mm): 25.04
Relaxation Delay(s): 0.1
Completed Scans 16
499.806 MHz H1 1D in cdcl3 (ref. to CDCl3 @ 7.26 ppm), temp 27.7 C -> actual temp = 27.0 C, colddual probe
O
OH
O
BnO
PivO
OH
2-52
10
9
8
7
6
5
4
3
2
1
0
ppm
13.28
1.86
5.13
25.38
File: /mnt/d600/home9/tlnmr/nmrdata/Anushka/2014.03.22.u5_ABJ-7-BF3-PIV_H1_1D


Agilent Technologies
Department of Chemistry, University of Alberta
Recorded on: u500, Mar 22 2014
Pulse Sequence: APT_ad
Sweep Width(Hz): 33783.8
Digital Res.(Hz/pt): 0.26
Acquisition Time(s): 2.5
Hz per mm(Hz/mm): 107.09
Relaxation Delay(s): 0.1
Completed Scans 56
125.690 MHz C13[H1] APT_ad in cdcl3 (ref. to CDCl3 @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, colddual probe
C & CH2 same, CH & CH3 opposite side of solvent signal
178.800
137.915
128.462
128.380
127.808
127.767
127.556
101.222
77.286
77.032
76.777
76.170
74.099
73.410
73.326
72.782
71.988
71.900
70.733
70.206
66.445
64.377
39.075
27.137
27.104
1.037
O
OH
O
BnO
PivO
OH
2-52
180
160
140
120
100
80
60
40
20
ppm
File: /mnt/d600/home9/fllnmr/nmrdata/Anushka/2014.03.22.u5_ABJ-7-BF3-PIV_C13_APT_ad

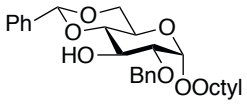

Department of Chemistry, University of Alberta
Agilent Technologies
Recorded on: May 21 2014
Pulse Sequence: s2pul
Sweep Width(Hz): 6000.6
Digital Res.(Hz/pt): 0.09
Acquisition Time(s): 5
Hz per mm(Hz/mm): 25
Relaxation Delay(s): 0.1
Completed Scans 16
498.118 MHz H1 1D in cdcl3 (ref. to CDCl3 @ 7.26 ppm), temp 26.4 C -> actual temp = 27.0 C, autoxdb probe
O
OBn
O
BnO
PivO
OBn
2-41
10
9
8
7
6
5
4
3
2
1
0
ppm
19.71
0.97
1.13
1.03
2.17
1.27
8.62
1.53
2.27
File: /mnt/d600/home9/tlnmr/nmrdata/Anushka/2014.05.21.I5_ABJ-7-AG2O_H1_1D

Agilent Technologies

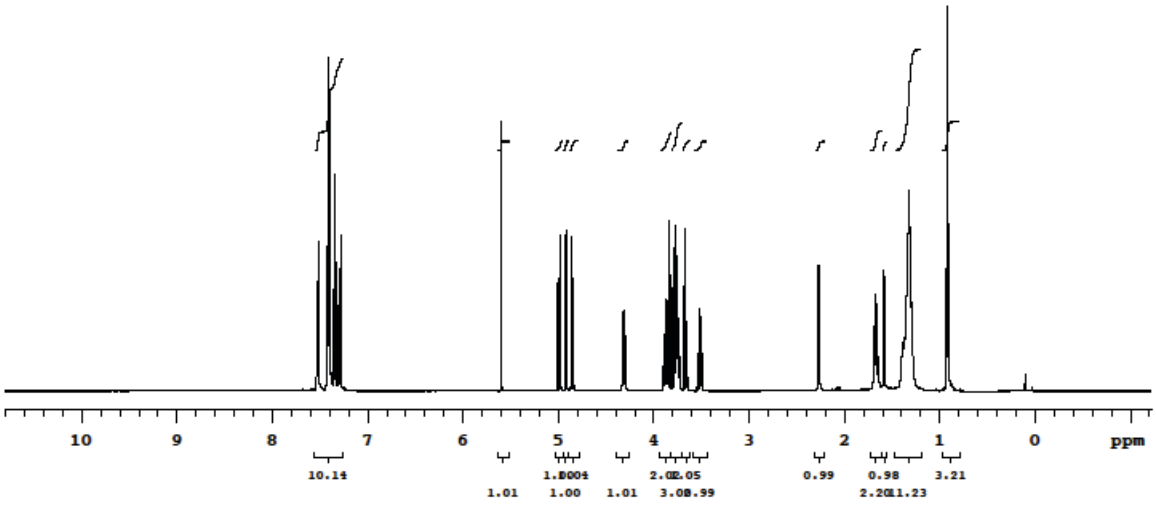
Department of Chemistry, University of Alberta

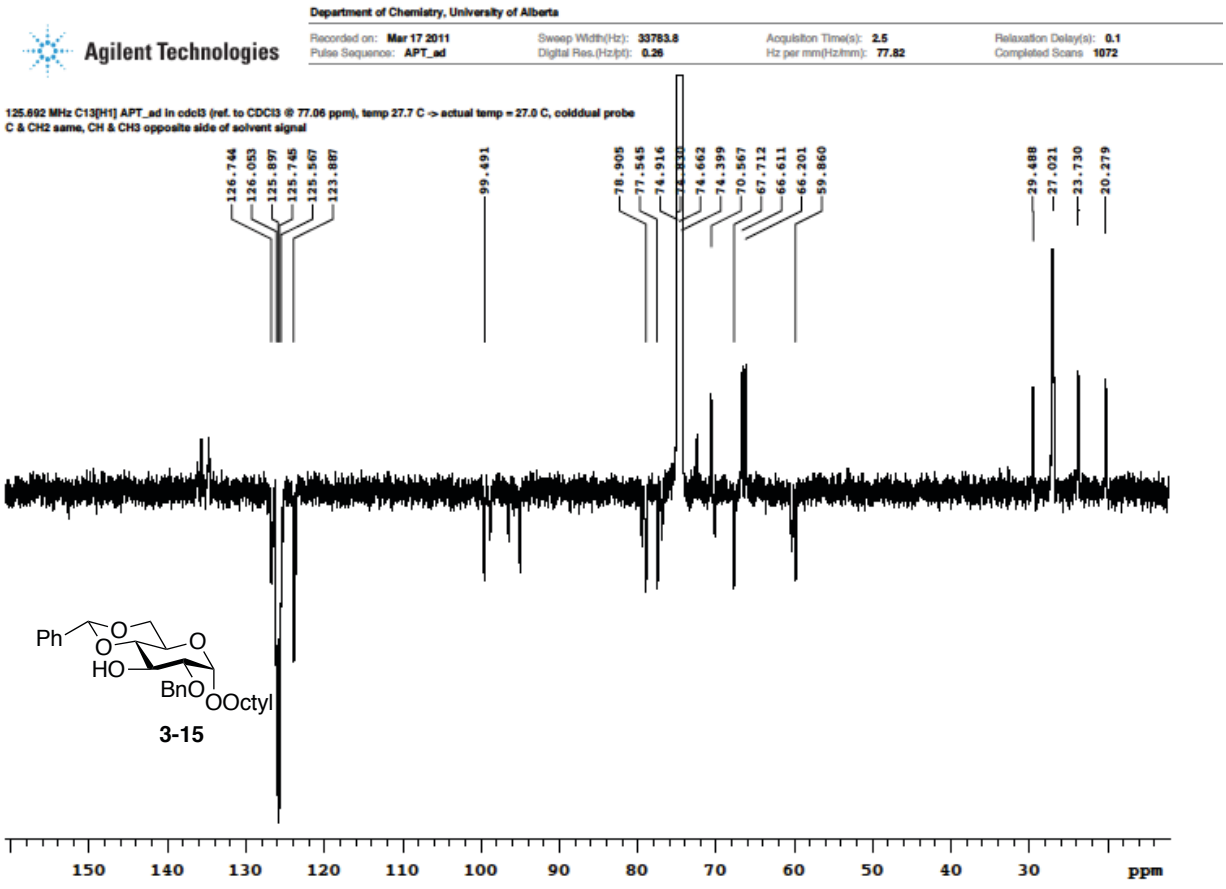
Recorded on:	Sep 15 2011	Sweep Width(Hz):	6009.62	Acquisition Time(s):	5	Relaxation Delay(s):	0.1
Pulse Sequence:	s2pul	Digital Res.(Hz/pt):	0.09	Hz per mm(Hz/mm):	25.04	Completed Scans	8

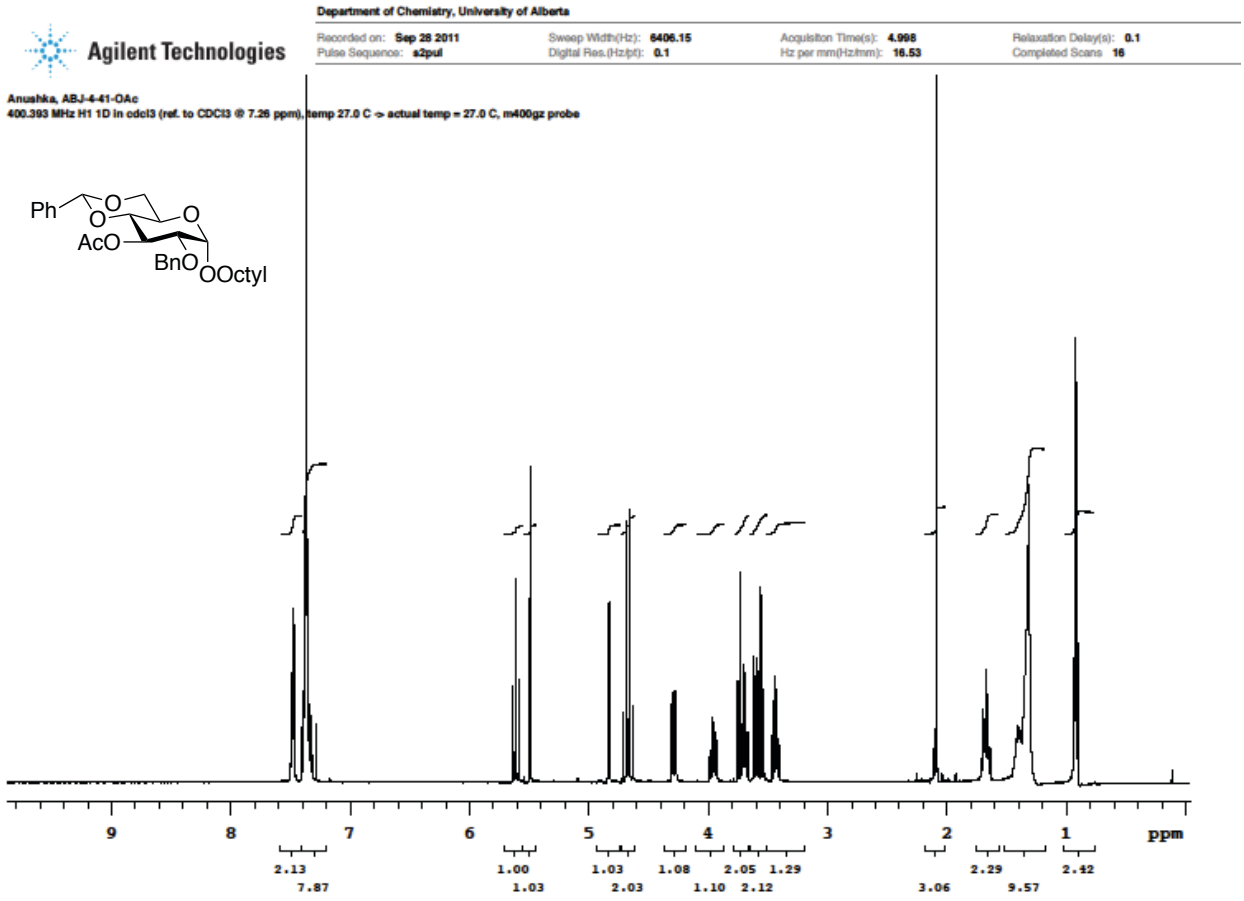
499.815 MHz H1 1D in cdcl3 (ref. to CDCl3 @ 7.26 ppm), temp 27.7 C -> actual temp = 27.0 C, colddual probe

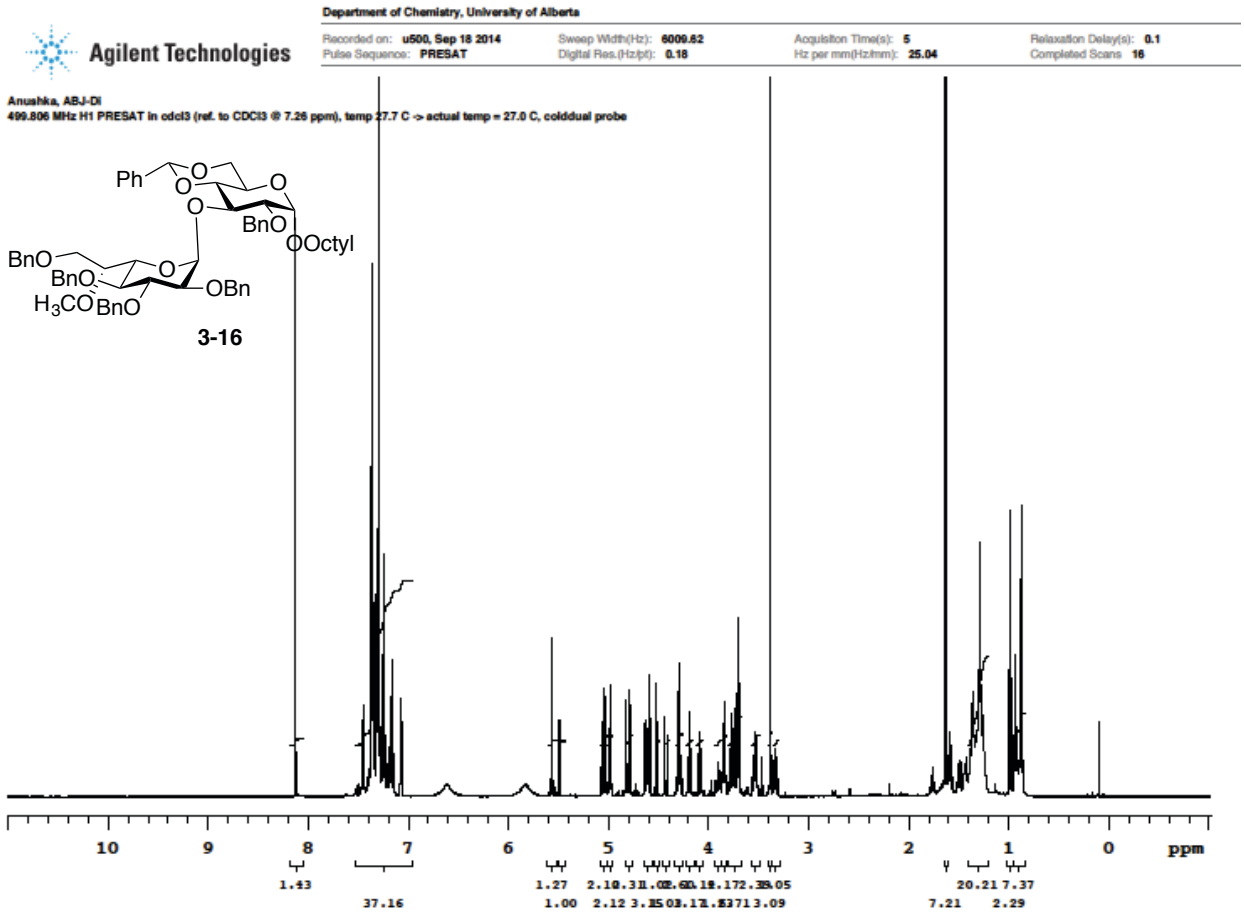


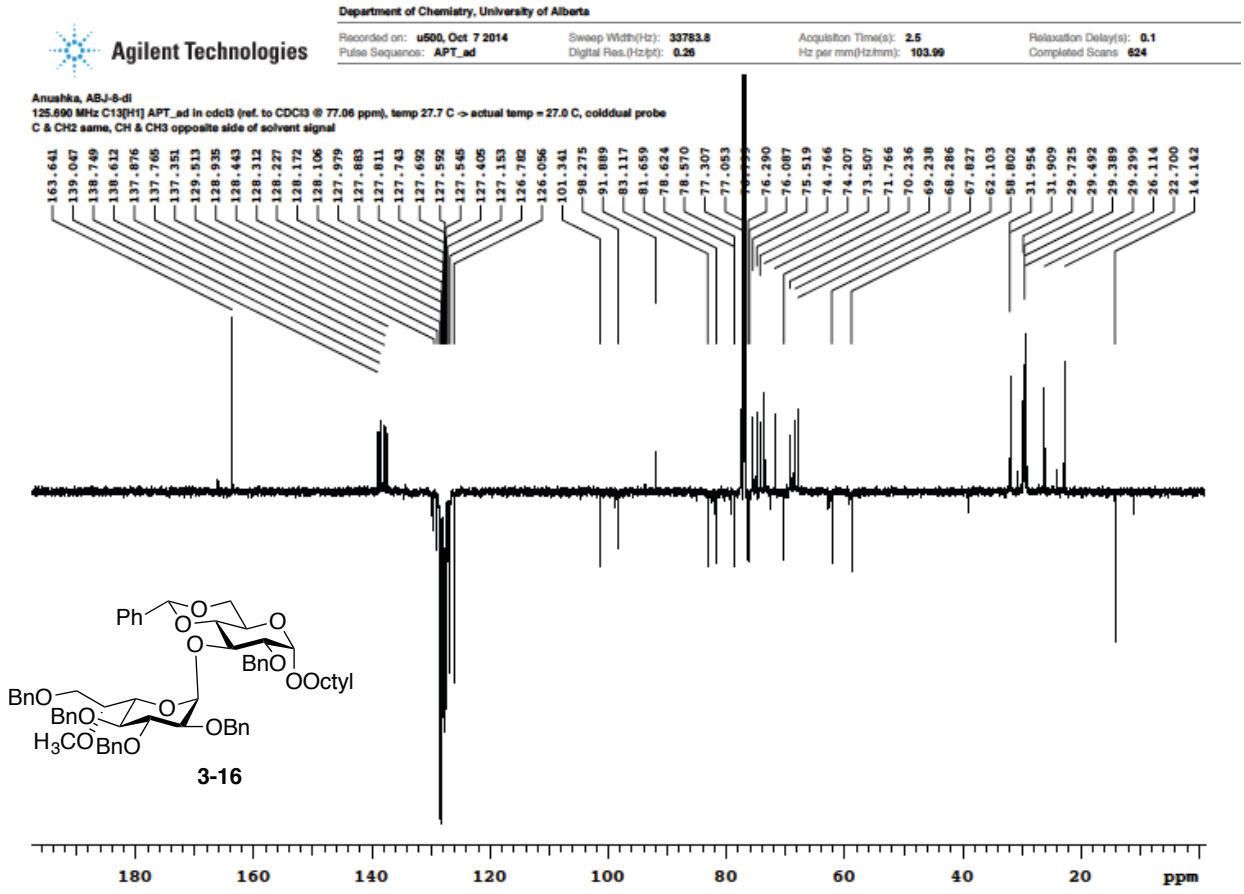
3-15

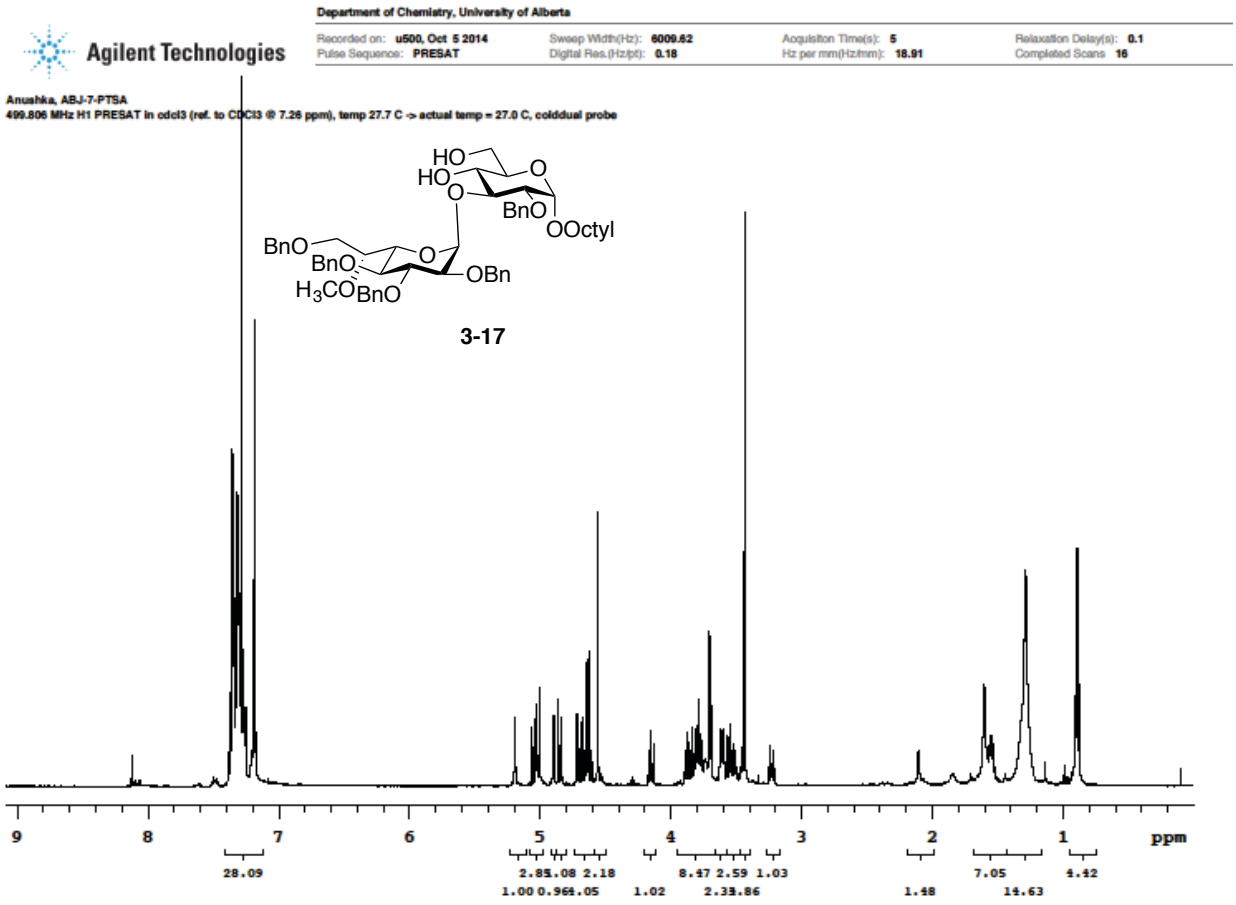


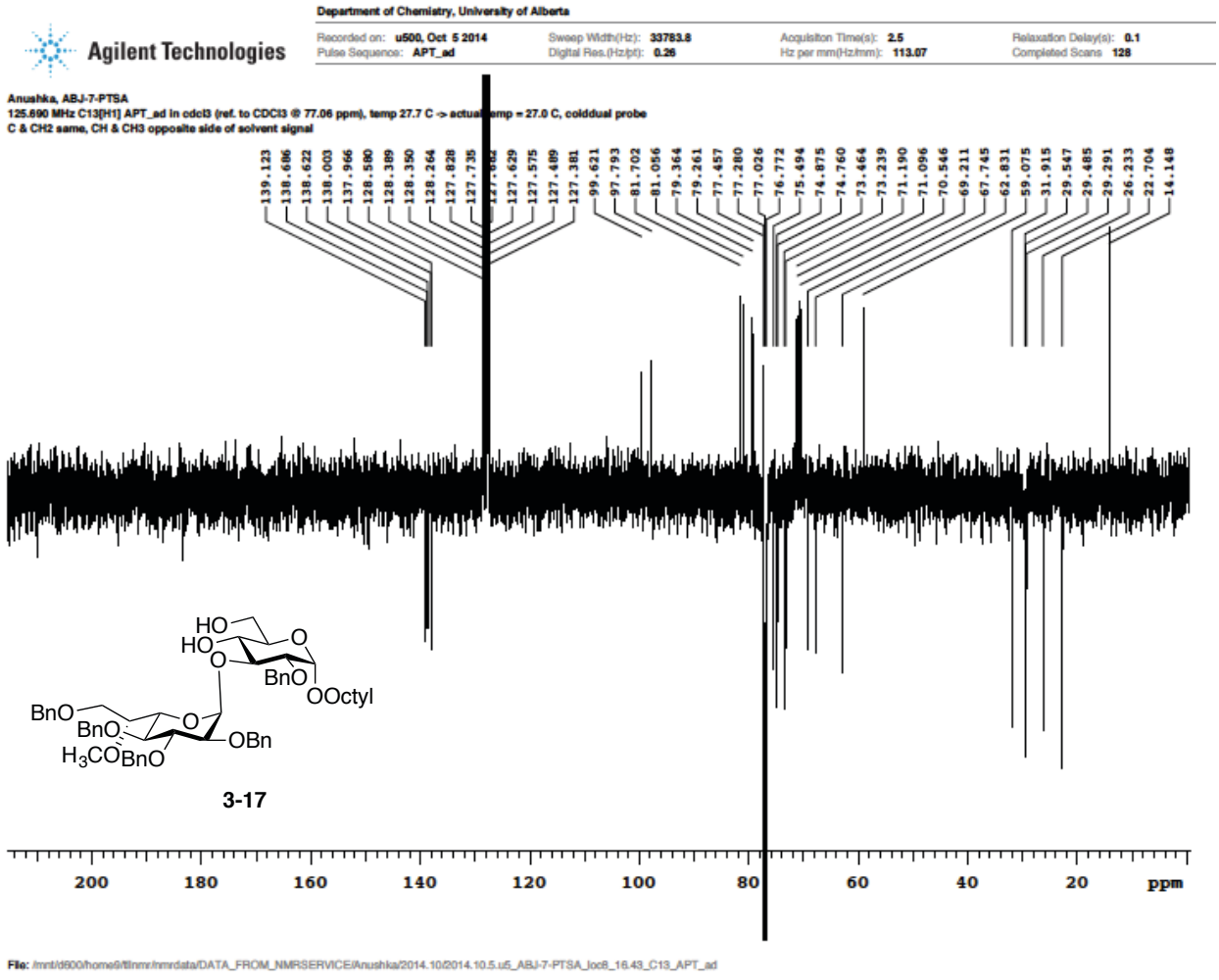

Department of Chemistry, University of Alberta
Agilent Technologies
Recorded on: Mar 17 2011
Pulse Sequence: APT_ad
Sweep Width(Hz): 33783.8
Digital Res.(Hz/pt): 0.26
Acquisiton Time(s): 2.5
Hz per mm(Hz/mm): 77.82
Relaxation Delay(s): 0.1
Completed Scans 1072
125.692 MHz C13[H1] APT_ad in cdcl3 (ref. to CDCl3 @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, colddual probe
C & CH2 same, CH & CH3 opposite side of solvent signal
126.744
126.053
125.897
125.745
125.567
123.887
99.491
78.905
77.545
74.916
74.662
74.399
70.567
67.712
66.611
66.201
59.860
29.488
27.021
23.730
20.279
Ph
O
O
O
HO
BnO
OOctyl
3-15
150
140
130
120
110
100
90
80
70
60
50
40
30
ppm


Department of Chemistry, University of Alberta
Agilent Technologies
Recorded on: Sep 28 2011
Pulse Sequence: s2pul
Sweep Width(Hz): 6406.15
Digital Res.(Hz/pt): 0.1
Acquisition Time(s): 4.998
Hz per mm(Hz/mm): 16.53
Relaxation Delay(s): 0.1
Completed Scans 16
Anushka, ABJ-4-41-OAc
400.393 MHz H1 1D in cdcl3 (ref. to CDCl3 @ 7.26 ppm), temp 27.0 C -> actual temp = 27.0 C, m400gz probe
Ph
O
O
O
AcO
BnO
OOctyl
9
8
7
6
5
4
3
2
1
ppm
2.13
7.87
1.00
1.03
1.03
2.03
1.08
1.10
2.05
2.12
1.29
3.06
2.29
9.57
2.42


Department of Chemistry, University of Alberta
Agilent Technologies
Recorded on: u500, Sep 18 2014
Pulse Sequence: PRESAT
Sweep Width(Hz): 6009.62
Digital Res.(Hz/pt): 0.18
Acquisition Time(s): 5
Hz per mm(Hz/mm): 25.04
Relaxation Delay(s): 0.1
Completed Scans 16
Anushka, ABJ-DI
499.806 MHz H1 PRESAT in cdcl3 (ref. to CDCl3 @ 7.26 ppm), temp 27.7 C -> actual temp = 27.0 C, colddual probe
Ph
BnO
OOctyl
BnO
BnO
OBn
H3CO
BnO
3-16
10
9
8
7
6
5
4
3
2
1
0
ppm
1.43
37.16
1.27
1.00
7.21
20.21
7.37
2.29

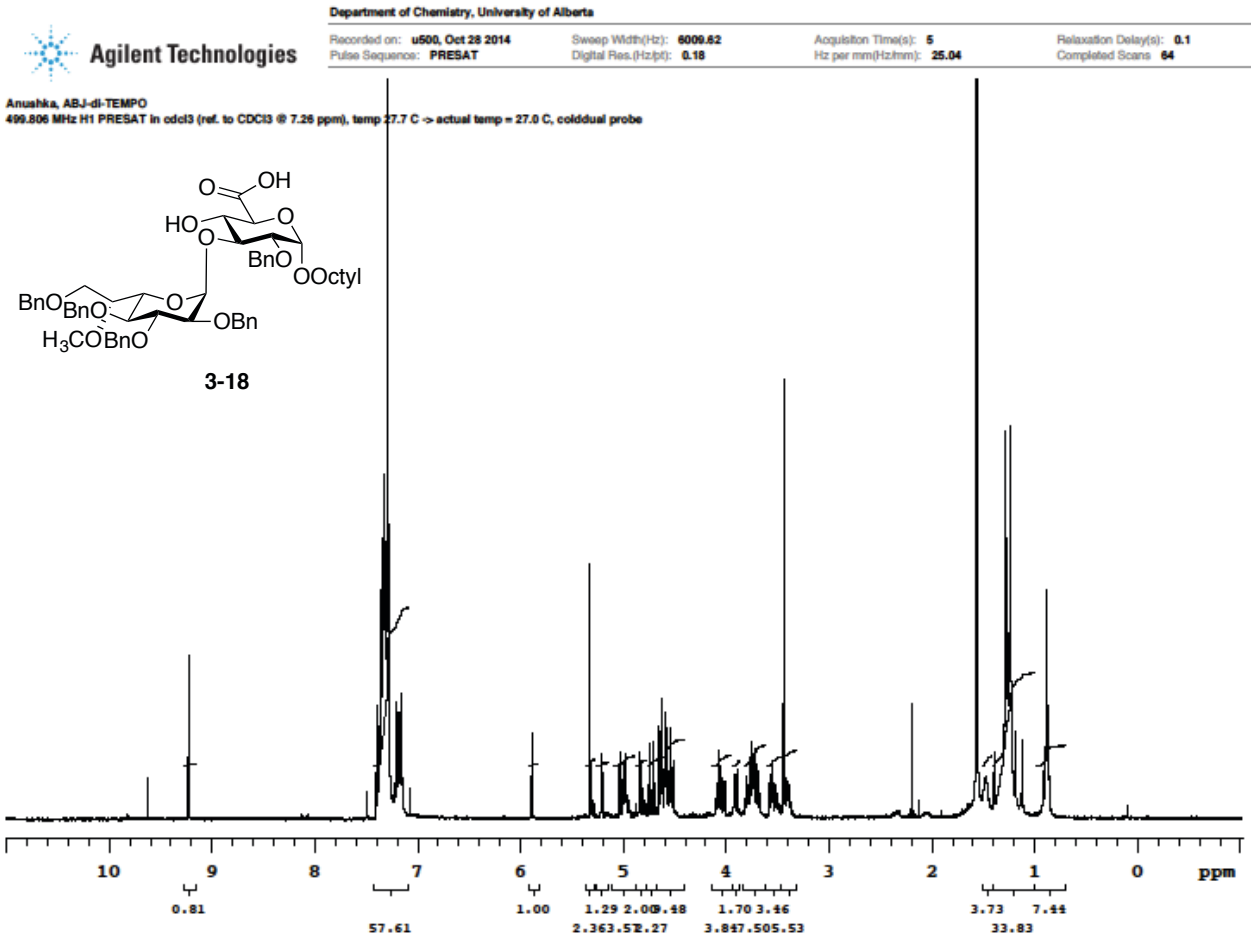

Department of Chemistry, University of Alberta
Agilent Technologies
Recorded on: u500, Oct 7 2014
Pulse Sequence: APT_ad
Sweep Width(Hz): 33783.8
Digital Res.(Hz/pt): 0.26
Acquisiton Time(s): 2.5
Hz per mm(Hz/mm): 103.99
Relaxation Delay(s): 0.1
Completed Scans 624
Anushka, ABJ-8-dl
125.690 MHz C13[H1] APT_ad in cdcl3 (ref. to CDCl3 @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, colddual probe
C & CH2 same, CH & CH3 opposite side of solvent signal
163.641
139.047
138.749
138.612
137.876
137.765
137.351
129.513
128.935
128.443
128.312
128.227
128.172
128.106
127.979
127.883
127.811
127.743
127.692
127.592
127.545
127.405
127.153
126.782
126.056
101.341
98.275
91.889
83.117
81.659
78.624
78.570
77.307
77.053
76.290
76.087
75.519
74.766
74.207
73.507
71.766
70.236
69.238
68.286
67.827
62.103
59.802
31.954
31.909
29.725
29.492
29.389
29.299
26.114
22.700
14.142
Ph
O
O
O
O
BnO
OOctyl
BnO
BnO
O
OBn
H3CO
BnO
3-16
180
160
140
120
100
80
60
40
20
ppm


Department of Chemistry, University of Alberta
Agilent Technologies
Recorded on: u500, Oct 5 2014
Pulse Sequence: PRESAT
Sweep Width(Hz): 6009.62
Digital Res.(Hz/pt): 0.18
Acquisition Time(s): 5
Hz per mm(Hz/mm): 18.91
Relaxation Delay(s): 0.1
Completed Scans 16
Anushka, ABJ-7-PTSA
499.806 MHz H1 PRESAT in cdcl3 (ref. to CDCl3 @ 7.26 ppm), temp 27.7 C -> actual temp = 27.0 C, colddual probe
HO
HO
O
O
BnO
OOctyl
BnO
BnO
O
OBn
H3CO
BnO
3-17
9
8
7
6
5
4
3
2
1
ppm
28.09
2.85 1.08 2.18
8.47 2.59 1.03
7.05
4.12
1.00 0.96 4.05
1.02
2.33 .86
1.48
14.63



3-17

File: /mnt/d600/home9/tlnmr/nmrdata/DATA_FROM_NMRSERVICE/Anushka/2014.10/2014.10.5.u5_ABJ-7-PTSA_loc8_16.43_C13_APT_ad


Department of Chemistry, University of Alberta
Agilent Technologies
Recorded on: u500, Oct 28 2014
Pulse Sequence: PRESAT
Sweep Width(Hz): 6009.62
Digital Res.(Hz/pt): 0.18
Acquisition Time(s): 5
Hz per mm(Hz/mm): 25.04
Relaxation Delay(s): 0.1
Completed Scans 64
Anuahka, ABJ-di-TEMPO
499.806 MHz H1 PRESAT in cdcl3 (ref. to CDCl3 @ 7.26 ppm), temp 27.7 C -> actual temp = 27.0 C, colddual probe
O
OH
HO
O
O
BnO
OOctyl
BnO
BnO
O
OBn
H3CO
BnO
3-18
10
9
8
7
6
5
4
3
2
1
0
ppm
0.81
57.61
1.00
1.29
2.06
9.48
2.36
3.52
2.27
1.70
3.46
3.84
7.50
5.53
3.73
7.44
33.83