DOCTOR OF PHILOSOPHY

The Isolation, Synthesis and Enzymatic Testing of Naturally Occurring Biologically Active Carbohydrates

Nelligan, Liam Matthew

Award date:
2015

Awarding institution:
Bangor University

The Isolation, Synthesis and Enzymatic Testing of Naturally Occurring Biologically Active Carbohydrates

Liam Matthew Nelligan

A thesis submitted for the degree of
Doctor of Philosophy
in the School of Chemistry

PRIFYSGOL
BANGOR
UNIVERSITY

Prifysgol Bangor University

For Margaret and Ian,

Abstract

This thesis describes 3 projects relating to bioactive carbohydrates. The first project describes the preparation of a library of synthetic standards for use in the industrial manufacture of galacto-oligosaccharides (GOS). Using several carbohydrate building blocks, we attempted to synthesise linear and branched trisaccharides in the least amount of synthetic steps possible. Using one donor and an acceptor with two or more free hydroxyl groups, we attempted "random" glycosylations.

In the second project the isolation and purification of therapeutic carbohydrate analogues from the seeds of the common British bluebell, *Hyacinthoides non-scripta,* is described. We aimed to further investigate the iminosugar content of the seeds of the common British bluebell using a series of chromatographic extraction methods.

In the third project, carbohydrates were tested for biological activity. Using commercially available human therapeutic enzymes (Myozyme, Replagal and Fabrazyme) and a range of enzymes from prokaryotes, plants and eukaryotes, synthetic glycosides were tested for evidence of enzyme inhibition and activation.

Acknowledgements

I would like to acknowledge funding from KESS and ESF schemes, without which I would not be able to complete my studies.

I'd firstly like to thank my supervisor Martina Lahmann for the guidance, support and opportunity to work in her research group for the last 4 years. I'd also like to thank both Prof. Eoin Scanlan and Prof. Stefan Oscarson for their helpful discussions and kind hospitality during our meetings in Dublin. Thanks also to Dr. Vera Thoss and Dr. Igor Perepichka for their guidance, encouragement and support throughout my 3 years.

I would give my regards to Prof. Robert Nash and all the staff members at PhytoQuest Ltd.; Hazel, Jackie and Yana for their kindness during my time in Aberystwyth. I'd also like to specifically thank Barbara Bartholomew at PhytoQuest for her help in running inhibition assays.

For their contributions and assistance with regards to acquiring recombinant human enzymes, I'd like to thank Dr. Tarek Hiwot and Ms. Katherine Peers of Queen Elizabeth hospital Birmingham.

I would like to thank both Dr. Mark Long and Dr. Farah Oulaidi-Lamiable, not only for their expertise over the years, but also for friendships which I hold dear. On the same note, I'd like to thank Mat Buck, Phil Harper and Dr. Elliot Bennett for their input, and kindness.

Special mention is reserved for Samuel Fitch, Natasha Powell and Lorna O'Farrell for preparing building blocks in the GOS project and reaffirming my love of Chemistry. In particular, I'd like to also thank James McNaughton and Steffi Rohde, who both helped tremendously.

I'd also like to thank two students who brought a tremendous amount of care, support and warmth into the lab; Kashmala Carys, Angus Coyne-Grell. Thank you for your help in the final hours, wrapping up unanswered questions in the GOS project.

I'd like to thank the technical staff, in particular Sam Page and Gwynfor Davies for their consistent dedication and support. Thanks to Tom Wilson, who not only defatted the bluebell seeds on which the extraction was based, but also for his help in Aberystwyth with the NMR service. I'd like to thank Dr. Dave Preskett for his help during the bluebell project, and concentrating samples on site at Mona. I'd also like to mention Juliana Morganti for her operation of the MALDI-TOF on occasion. Siobhan, Caroline, Tracey for all the help with organisation and administration.

All of the above have made me a better chemist in one way or another. Without them, I would not have come this far.

Abbreviations

BMT	-	bone marrow transplant
BSTFA	-	(*N*,*O*-bis(trimethylsilyl)trifluoroacetamide)
COSY	-	correlation spectroscopy
CNS	-	central nervous system
D-AB1	-	1,4-dideoxy-1,4-imino-D-arabinitol
D/M	-	dichloromethane/methanol
DCE	-	1,2-dichloroethane
DCM	-	dichloromethane
DMDP	-	2,5-dideoxy-2,5-imino-D-mannitol
DMTST	-	dimethyl(methylthio)sulfonium triflate
DMF	-	dimethylformamide
DNJ	-	1-deoxynojirimycin
DVB	-	divinyl benzene
E/M	-	ethyl acetate/methanol
E/A/M/H	-	ethyl acetate/acetic acid/methanol/water
ELISA	-	enzyme-linked immunosorbent assay
ER	-	endoplasmic reticulum
ERT	-	enzyme replacement therapy
ESI-MS	-	electrospray ionization mass spectrometry
equiv.	-	equivalents
Fuc	-	fucose
Gal	-	galactose
GC-MS	-	gas chromatography – mass spectrometry
Glc	-	glucose
GlcNAc	-	*N*-acetylglucosamine
GOS	-	galacto-oligosaccharides
GSL	-	glycosphingolipidose
h	-	hours
HMO	-	human milk oligosaccharides
HNJ	-	homonojirimycin

homoDMDP	-	2,5-dideoxy-2,5-imino-glycero-D-manno-heptitol
HPAEC-PAD	-	high performance anion exchange chromatography with pulsed amperometric detection
HPLC	-	high pressure liquid chromatography
HSQC	-	heteronuclear single quantum coherence
m/z	-	mass/charge ratio
MALDI-TOF	-	matrix assisted laser desorption/ionisation – time of flight
MS	-	mass spectrometry
Mw	-	molecular weight
NB-DNJ	-	*N*-butyldeoxynojirimycin
NI	-	no inhibition
NMR	-	nuclear magnetic resonance
NBS	-	*N*-bromosuccinimide
Neu5Ac	-	sialic acid
OAc	-	acetate
OBn	-	benzyl ether
OBz	-	benzoyl ester
OpMP	-	*p*-methoxyphenol ether
OTs	-	tosylate
PHA	-	polyhydroxylated alkaloids
ppm	-	parts per million
rt	-	room temperature
SRT	-	substrate reduction therapy
STol	-	thiotolyl
TFA	-	trifluoroacetic acid
T/E	-	toluene/ethyl acetate
THF	-	tetrahydrofuran
T_H	-	T helper
TLC	-	thin layer chromatography
TMCS	-	trimethylchlorosilane
TMSOTf	-	trimethylsilyl trifluoromethanesulfonate
TMS-Cl	-	trimethylsilyl chloride
Tol.	-	Toluene
XOS	-	xylo-oligosaccharide

Table of Contents

1. Introduction to Carbohydrates

1.1 Introduction

1.1.1 What are carbohydrates?

Carbohydrates are naturally occurring organic molecules, which all have the same general formula of $C_n(H_2O)_n$[1]. One of the most abundant sugars in nature is glucose. Plants produce glucose via photosynthesis[2], taking in water from their roots and carbon dioxide from the air, synthesising glucose and oxygen from sunlight as energy[3]. Animals consume glucose and use it for energy production via respiration in the cells.

There are 4 categories of carbohydrates; monosaccharides, disaccharides, oligosaccharides and polysaccharides[4]. Monosaccharides are the base unit of all carbohydrates. They are the simplest form of sugar, and usually have a sweet taste. Glucose, galactose, and fructose are common examples of monosaccharides.

1.1.2 Classification

1.1.2.1 Simple Carbohydrates

Monosaccharides are conveniently represented using Fisher projections (Figure 1)[2]. Each of the carbon atoms in the ring system, which are secondary alcohols, are stereogenic centres affording stereoisomers. Stereoisomers are molecules with the same formula, differing only in the absolute configuration of their stereogenic centres. Many carbohydrates are diastereomers of each other; stereoisomers having different configurations at one or more stereocentres, and are not mirror images of one another. They have the same general structure and react in similar ways, but due to their stereochemistry have different properties.

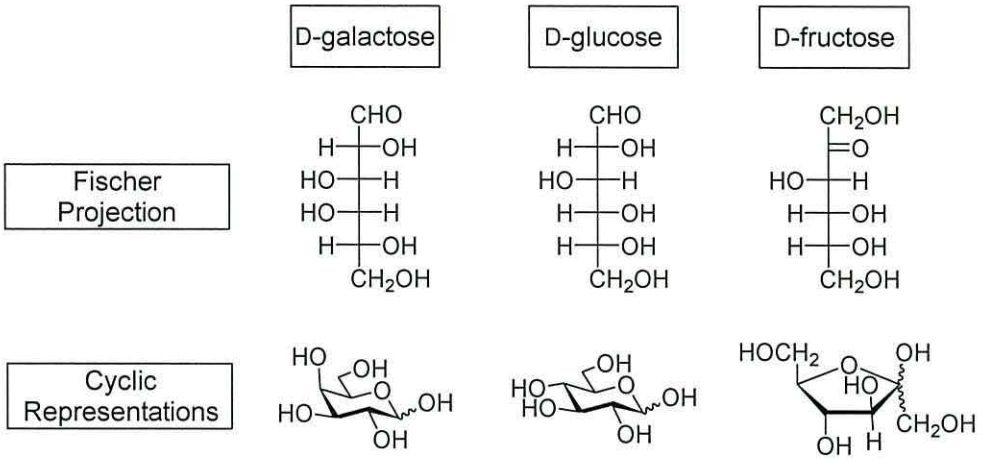


Figure 1. Fisher projections and ring structures of simple monosaccharides.

The Fisher convention of drawing structures simplifies the representation of the stereochemistry of the molecules. The general backbone is the same for all carbohydrates. The most common form of carbohydrates are polyhydroxylated aldehydes, referred to as aldoses (Figure 2). Aldoses contain an aldehyde group at the top of the chain, a number of secondary alcohols as the sugar backbone, and a primary alcohol at the opposing end. Ketoses are much less common, arranged with primary alcohols at each end of a chain containing a ketone. Fructose is an example of a ketose.

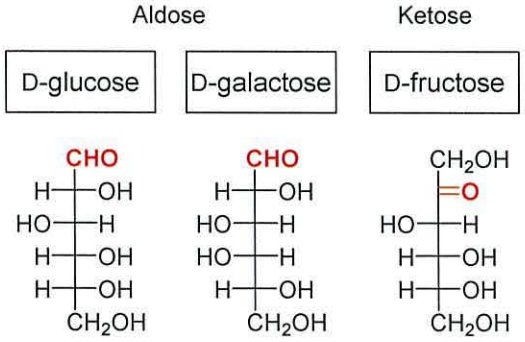


Figure 2. Aldohexoses and ketohexoses have 6 carbons and are defined by the position of the carbonyl group.

A monosaccharide can also be classified by the number of carbon atoms it contains[5]. Sugars with four carbons are tetroses, five carbons are pentoses, whilst six and seven carbons are hexoses and heptoses respectively. Glucose is an example of a polyhydroxylated aldehyde with six carbons, and is therefore an aldohexose, whilst fructose, a polyhydroxylated ketone with the same number of carbons, is a ketohexose[4].

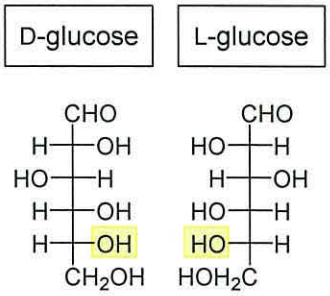


Figure 3. D- and L- notation are used to describe stereochemistry. This is assigned by the configuration at the highest numbered centre of chirality.

The notations D- and L- are also used to classify the configurations of sugars (Figure 3)[1]. In a Fisher projection, the bottom most asymmetric centre denotes whether the sugar is in the D- or L- form[6]. If the OH group is on the right, the molecule is a D-sugar, and if it is on the left, it is an L-sugar[5].

1.1.2.2 Complex carbohydrates

Disaccharides are made up of two monosaccharide units connected by a glycosidic linkage (Figure 4). Lactose, found in milk, and sucrose, in table sugar, are examples of common disaccharides. Oligosaccharides are between three to nine monosaccharide units attached to one another in either a branched or linear formation. If a monomer has more than one glycosidic linkage, it serves as a branching point. The naming of these long chain structures is usually done from the non-reducing end to the reducing end[7].

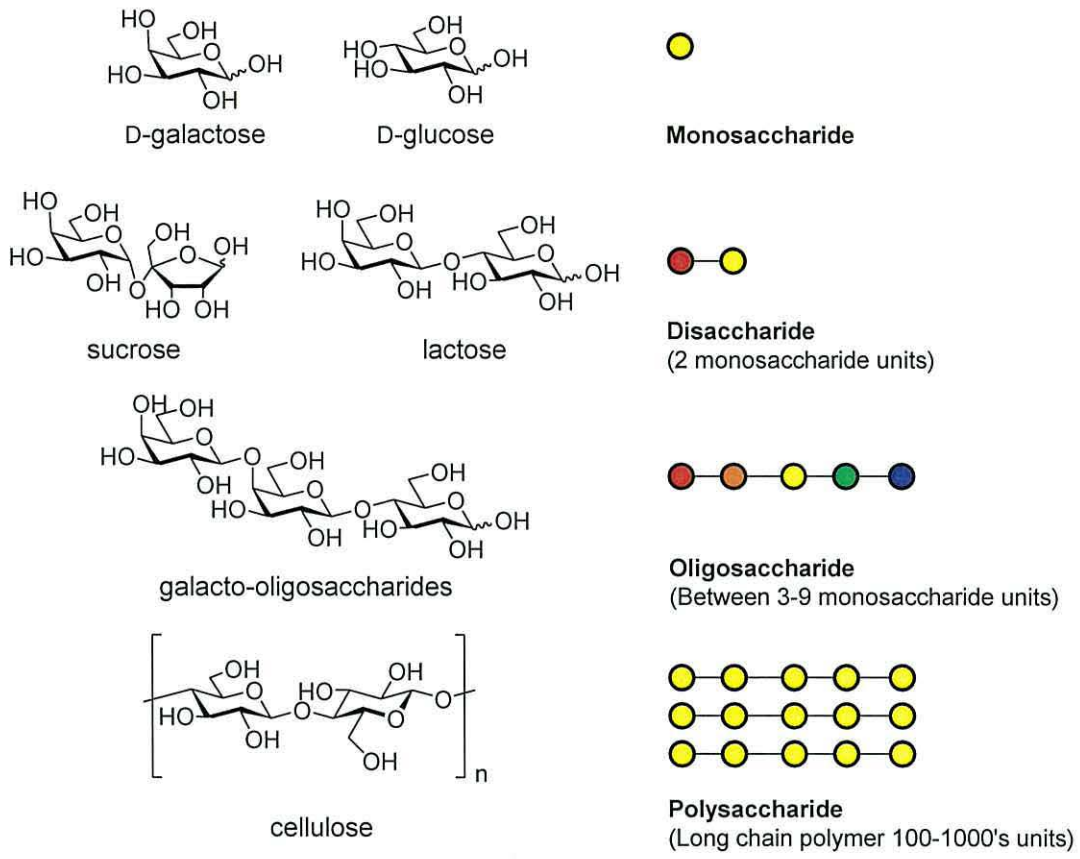


Figure 4. The 4 categories of sugars, with common examples and "stick and ball" annotations of each.

Polysaccharides contain as little as 10 monosaccharide units in a chain, to as many as several thousand units, joined by glycosidic linkages. Some of the most common polysaccharides in nature are starch, glycogen, chitin and cellulose. In cellular respiration plants oxidise glucose to form ATP, supplying their energy needs[8]. When animals have too much glucose for their energy requirements, they convert the excess into glycogen[9]. Glycogen is a branched structure consisting of α-linked glucose. The branched nature of the compound allows it to be broken down rapidly, meaning it is used a "quick" source of energy essential for animals as they move and respire[8]. In comparison, plants convert excess glucose into starch, which is insoluble in water and can be stored compactly[10].

Carbohydrates also have structural properties. Cellulose is a long chain polymer of β-D-glucose and is found in the cell walls of plants. It is a major component of both cotton (~90 %) and wood (~50 %)[2]. Another example is chitin, a polymer of *N*-acetylglucosamine, found in the exoskeletons of anthropods. All complex carbohydrates can be hydrolysed to their respective monosaccharide base units.

1.2 Carbohydrate Structure and Properties

1.2.1 Anomeric centre

Carbohydrates exist almost exclusively in the form of closed ring systems due to their low angle and eclipsing strain. Depending on which hydroxyl group of the open chain-form of the carbohydrate reacts with the carbonyl function, 5- or 6- membered rings are formed (Figure 5). Furanose forms are 5-membered ring systems, whilst pyranose are 6-membered. This naming system is in correlation to their unfunctionalised analogues.

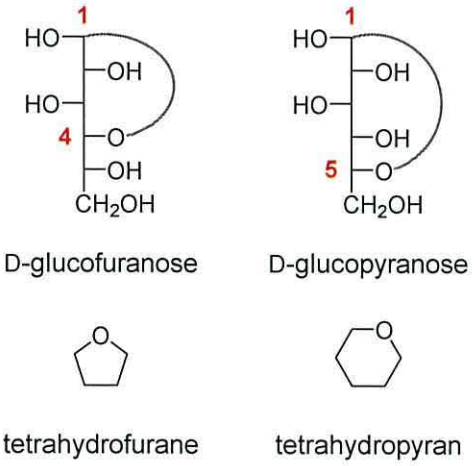


Figure 5. Formation of furanose and pyranose rings is dependent on the intramolecular attack from hydroxyl groups.

These cyclic structures are both the most kinetically and thermodynamically stable forms. When the open chain aldehyde form of a sugar cyclises, it produces a hemi-acetal or lactol[4]. This new stereocentre at C-1 is referred to as the anomeric position (Figure 6). The two anomers at this position are denoted as α- (cis) and β- (trans), and are both diastereomers of each other.

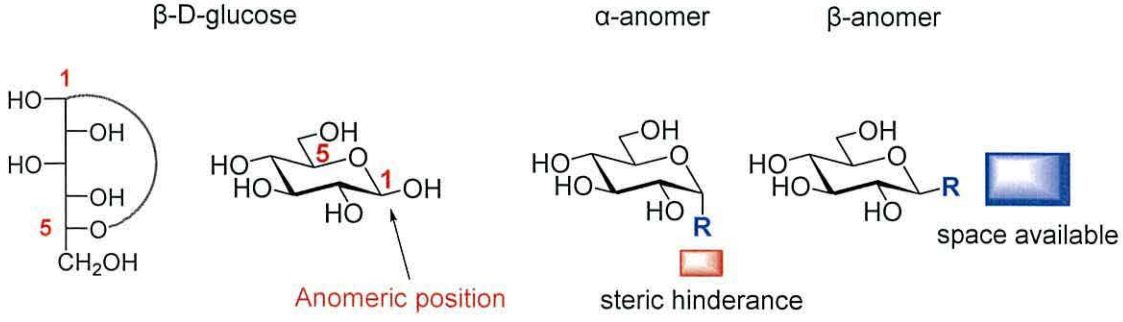
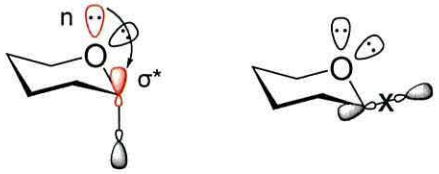


Figure 6. The anomeric position annotated as #1 in both Fisher and pyranose projections. α- and β-anomers of D-glucose with visual representation of steric effects.

1.2.2 Anomeric effect

Steric effects arise from the fact that each atom occupies a certain amount of space. Steric hindrance would describe the β-anomer as being more stable than its α-counterpart, as there is more room for a substituent in the equatorial position (Figure 6). However, due to the anomeric effect, the α-anomer is more thermodynamically stable. When an electronegative substituent is in the axial position at the anomeric centre, the endocyclic oxygen interacts with the anti-perpendicular electronegative functional group[1]. A stabilising interaction termed hyperconjugation occurs; a partial donation of the unshared electron pair on the endocylic oxygen and the σ* orbital for the axial C-X bond lowers the overall energy of the system, stabilising the molecule (Figure 7).

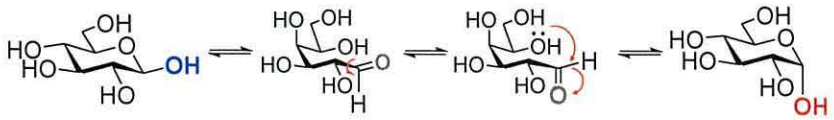


overlapping orbitals

Figure 7. Preference of certain substituents bonded to the anomeric carbon towards the axial position is called the anomeric effect[6]. Electron delocalisation across overlapping orbitals stabilise the α-anomer.

1.2.3 Mutarotation

In aqueous solutions, glucose exists in equilibrium at 36% of the α- and 64% of the β-form. In water, the hemiacetal of the sugar opens up, and upon re-cyclisation of the chain, both α- and β- forms are produced (Scheme 1). Dissolving either pure α-D-glucose or pure β-D-glucose will result in the same equilibration. The term mutarotation refers to the equilibrium of these α- and β- forms. Mutarotation can be defined as a reversible epimerisation which occurs via free rotation around a carbon-carbon bond.



β-D-glucose α-D-glucose

Scheme 1. The process of mutarotation; the hemiacetal of the sugar opens to the aldehyde form and re-cyclises to equilibrate as either α/ β anomer.

The fact that the β-anomer is the major part of glucose present at equilibrium is accredited to solvation[1]; defined as the attraction of solvent molecules with the molecules or ions of a solute. Solvation causes the anomeric hydroxyl group to become heavily solvated, increasing the steric bulk of the anomeric substituent.

Mutarotation allows interconversion between α, β- stereoisomers, however it can also allow interconversion between pyranose (6-ring) and furanose (5-ring) forms. The pyranose form is more stable. The furanose constituents form at a more rapid pace and are less stable.

1.2.4 Reducing sugar

Reducing sugars are sugars which have, or are able to form an aldehyde, through isomerisation (Figure 8). Hemiacetals are in equilibrium with the open-chain sugars in solution, and can therefore freely form an aldehyde. An acetal, such as the glycosidic linkage in a disaccharide, cannot easily form an aldehyde, is therefore a non-reducing sugar.

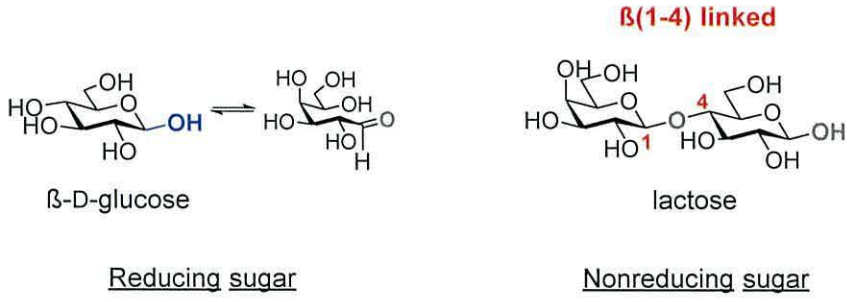


Figure 8. Reducing sugars contain a hemiacetal, and are able to freely form an aldehyde. Acetals in disaccharides have a fixed α- or β-position, and are non-reducing sugars. The linkage in lactose is β (1-4), as it is in the equatorial position and also connected from C-1 on one unit, to C-4 on the next.

1.3 Synthetic Pathways

1.3.1 Protecting groups

Functional groups on the molecule can be protected at any time using either permanent or temporary protecting groups. "Permanent protecting groups" will remain in position throughout the synthesis, only to be removed in the last synthetic steps, whilst "temporary protecting groups" can be easily added and removed[11]. The majority of synthetic steps will introduce temporary protecting stages to gradually develop the structure (Figure 9). For example, in oligosaccharide synthesis, permanent protecting groups will be removed once the target molecule has been synthesised, which temporary protecting groups have helped develop.

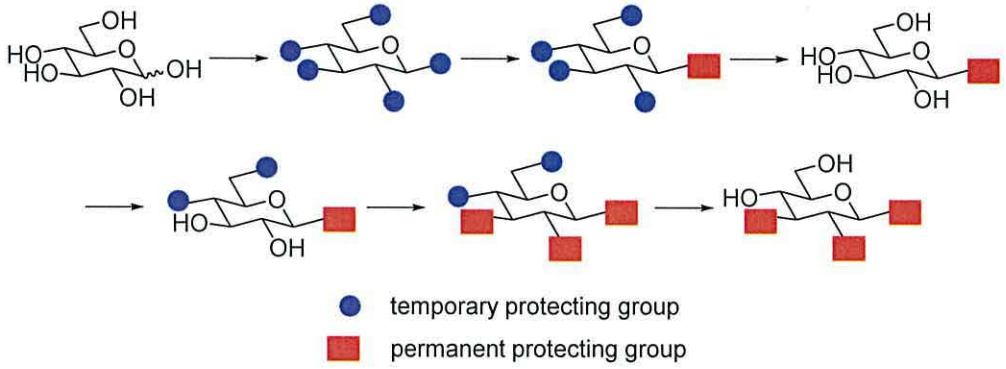


Figure 9. Example of synthesis pathway utilising "permanent" and "temporary" protecting groups. The "permanent" protecting groups will be removed in the last steps of the synthesis.

1.3.2 Permanent Protecting Groups

1.3.2.1 Benzyl ethers

Benzyl ethers are a very common ether protecting group used in carbohydrate chemistry[2]. Benzyl ethers are commonly prepared by treatment of an alcohol with a benzyl halide in the presence of a base. These protecting groups are considered to be more "permanent", remaining intact in a wide range of reaction conditions. General removal of benzyl ethers is achieved via catalytic hydrogenation, employing palladium on activated charcoal[12].

1.3.3 Temporary Protecting Groups

1.3.3.1 Acetylation

Acetates are the most common esters to be used in protecting group synthesis and can be thought of as "temporary protecting groups"[12]. Alcohols can be readily acetylated by stirring in acetic anhydride and a base. The base not only acts as an acceptor for any excess acetic acid formed, but will also catalyse the reaction. This is very important, as the uncatalysed reaction of hydroxyl groups and acetic anhydride at room temperature is very slow.

An acetylation using pyridine produces an anomeric mixture[13]. This is due to the fact that the free hydroxyl at the anomeric position is acetylated faster than competing mutarotation can occur[1] (Figure 10).

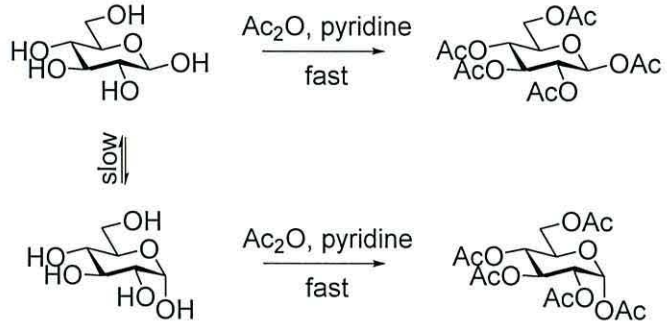


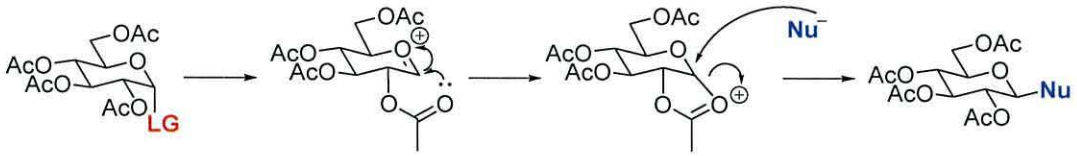
Figure 10. Acetylation with pyridine yields an α/β mixture[1]

Acetylation with sodium acetate gives a higher amount of the β-anomer. At the high temperatures required by the reaction, mutarotation occurs faster than the acetylation reaction itself. As the β-anomer is more nucleophilic than the α-form, it reacts in preference[1].

Anomeric acetates may also act as a leaving group under acidic conditions, allowing the addition of other moieties via nucleophilic substitution[12]. With this in mind, one of the first steps in many synthetic pathways will be the per-acetylation of the starting molecule.

1.3.4 Neighbouring group participation

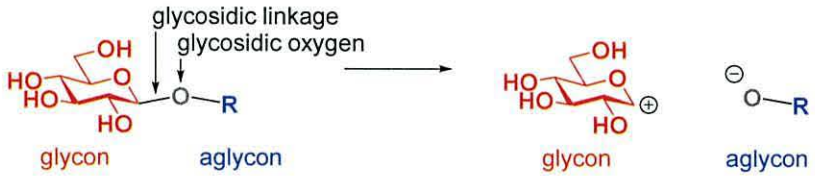
In the production of oligosaccharides, stereochemical control is crucial. A way of controlling the stereochemistry of the anomeric bond is through a process of neighbouring group participation[1] (Scheme 2). Protection of the C-2 position with an ester protecting group allows for an interaction between the ester and intermediate cation. This participation stabilises the intermediate glycosyl cation by cyclisation, and directs the attack by the nucleophile.



Scheme 2. Neighbouring group participation example with acetylated glucose, containing leaving group (LG) and addition of nucleophile (Nu).

1.4 Glycosides

A glycoside is defined as any class of compound that, once hydrolysed, yields a carbohydrate and an aglycone[12]. Bonds which connect sugars to other compounds, which may or may not be another sugar, are considered to be a glycosidic bond (Scheme 3). The classification of glycosides depends on this connection between the glycone and the aglycone. The glycosidic linkage between the two moieties may be connected by oxygen (*O*-glycosides), nitrogen (glycosylamines), sulphur (thioglycosides), or carbon atoms (*C*-glycosides)[14].



Scheme 3. Glycosides; cleavage at the glycosidic linkage gives glycone and aglycone (adapted from Essentials of Carbohydrate Chemistry[12])

1.4.1 Glycosylation Reactions

Oligosaccharide synthesis requires the conversion of building blocks into donor and acceptor precursors[15]. A glycosidic bond is formed between the two sugars via displacement of the leaving group on the anomeric position of the donor by a promoter[14], with the unprotected hydroxyl position of an acceptor. This mechanism is referred to as a glycosylation reaction, and, when one hydroxyl group of one sugar is used to form a glycoside, a disaccharide is produced (Scheme 4). This process can be repeated in the case of the synthesis of larger oligosaccharides.

Scheme 4. Glycosylation reaction of donor and acceptor forms a glycoside joined by glycosidic bond.

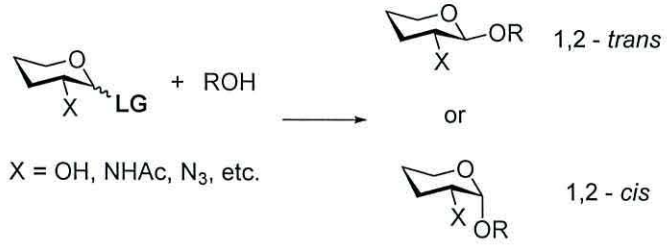
As water can act as a competitive nucleophile, hydrolysis of products can occur if the reaction conditions are not anhydrous.

1.4.2 Glycosyl acceptors

Glycosyl acceptors are carbohydrates with an unprotected hydroxyl position (Scheme 4). Usually, it is preferential to only have one hydroxyl position free, as to avoid the formation of mixtures of regioisomeric products. However, it is also possible to synthesise acceptors with multiple unprotected positions. The anomeric position of a glycosyl acceptor should not contain a leaving group which could be activated in the glycosylation reaction to avoid the possibility of it reacting with itself.

1.4.3 Glycosyl donors

The defining characteristic of a glycosyl donor is a good leaving group (Scheme 5). Many different techniques are used for glycosylations, and each leaving group has its pros and cons with stability and activation to be regarded. A glycosyl donor should be suitably protected with electron donating protecting groups (ethers) to increase reactivity, and to also avoid the possibility of self-reaction. Protection of the C-2 carbon of the glycosyl donor must also be considered. If one wishes to generate 1, 2-trans glycosides in a stereospecific reaction, the C-2 position should be protected with an ester. With a non-participating group at the C-2 position, generation of both 1, 2-cis (α-linked) and 1, 2-trans glycosides are possible depending on stereospecificity of the promoter (Scheme 5). It is much harder to synthesise 1, 2-cis glycosides than it is to generate 1, 2-trans glycosides.



Scheme 5. The nature of the protecting groups on the glycosyl donor has an effect on the outcome of the glycosylation.

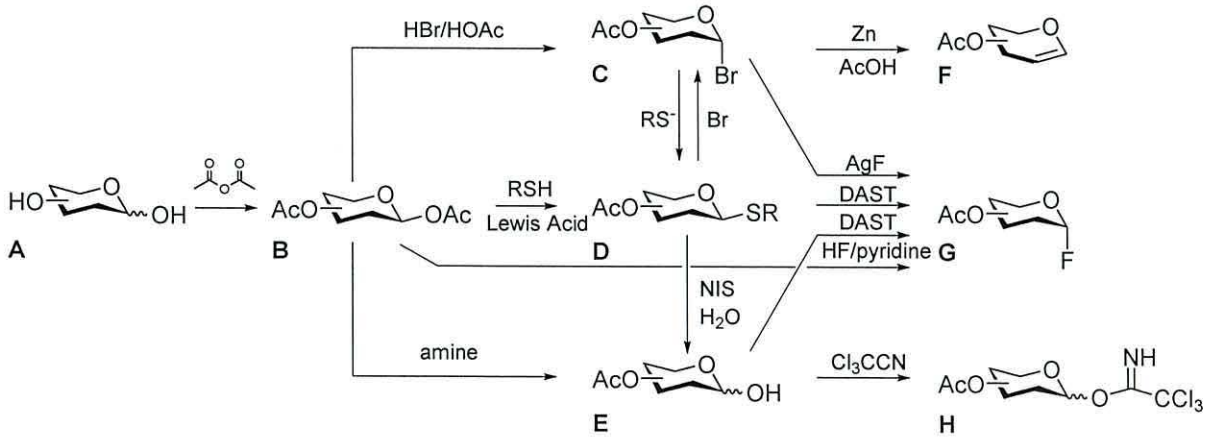
1.4.3.1 Glycosyl Halides

Glycosyl bromides are one of the oldest leaving groups[12]. They are very easily generated via a one-step synthesis using HBr in acetic acid from a per-acetylated sugar (Scheme 6 C). Due to their reactive nature, they are usually synthesised and used directly in glycosylations. Their ability to interconvert to several other functional groups is also very appealing (Scheme 6). The classical Koenigs-Knorr glycosylation procedure uses silver salts as promoters, whilst the Helferich procedure uses mercury salts[1,16].

Glycosyl chlorides and fluorides are also used as leaving groups in glycosyl donors[12,14]. The latter are the most modern form of halide donors (Scheme 6 G), and are much more stable than their bromide counterparts.

1.4.4 Thioglycosides

Thioglycosides are considered to be very effective temporary protecting groups for glycosylation reactions (Scheme 6 D). They are very stable species which can be stored for long periods of time and will not react until activated under specific circumstances[17]. The conditions used for the regioselective activation of these thioglycosides do not affect other glycosidic bonds or protecting groups in the molecule. Despite all of the positives, thioglycosides generally give lower yields than other leaving groups[18].



Scheme 6. Synthesis pathways of leaving groups used in glycosyl donors (adapted from Jacobssen et al[18]).

1.4.4.1 Promoters

For couplings with thioglycosyl donors, *N*-iodosuccinimide (NIS) and *N*-bromosuccinimide (NBS) are regularly used in conjunction with a catalytic amount of soluble triflic salts, usually silver trifluoromethanesulfonic acid (AgOTf) (Figure 11). These compounds are stable, crystalline, commercially available reagents. The fact that these thiophilic reagents can promote both "armed" and "disarmed" donors in mild conditions with high yields means that this method of activation is one of the most practical procedure in oligosaccharide synthesis[19].

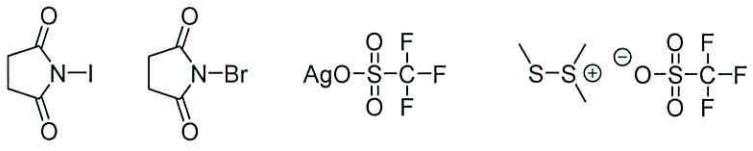


Figure 11. Common activators of thioglycosides.

Dimethyl(methylthio)sulfonium triflate (DMTST) is used as a promoter of thioglycosides in oligosaccharide synthesis[15](Figure 11). It has been reported to be the most efficient promoter of thioglycosides[20], although reaction rates are slower than NIS/AgOTf. This promoter is synthesised from dimethyl disulfide and methyl triflate[21]. Both DMTST and the reagents used to make it are unstable, toxic and expensive.

1.4.5 Modern Variations

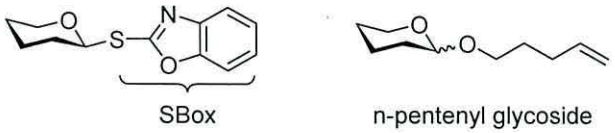


Figure 12. Examples of modern glycosyl donors

Modern glycosyl donors are utilised for their efficacy, accessibility, high stabilities and mild activation conditons[22] (Figure 12). *N*-pentenyl glycosides are activated in the presence of NIS, or other sources of iodine[23–25]. S-Benzoxazolyl (SBox) can be activated by MeOTf, and are useful for selective glycosylations as they are able to withstand other glycosyl donor activation conditions[22,26].

1.4.6 Trichloroacetimidates

The trichloroacetimidate method was developed by R.R. Schmidt[16]. Activation of trichloroacetimidates can be performed at low temperatures using a Lewis acid, and is therefore very useful for delicate or complicated molecules[18]. They are usually synthesised from free anomeric alcohols using trichloroacetonitrile (Scheme 6 H), and are more stable than glycosyl bromides[12]. If the C-2 position of the donor has a non-participating group, the stereochemistry of the leaving group has an outcome on the glycosidic linkage. In the correct conditions, α-trichloroacetimidates produce β-glycosides, and β-trichloroacetimidates give α-glycosides[12]. In general, glycosylations with trichloroacetimidates give good yields in both small and large scales.

1.4.7 Armed-disarmed donors[23]

The armed-disarmed concept stems from the idea that acyl protecting groups (i.e. acetates) reduce the reactivity at the anomeric centre, whilst ether protecting groups (benzyl ethers) increase it. The method describes oligosaccharide synthesis with two glycosyl donors of different reactivity, equipped with the same anomeric leaving group. Under specific conditions, it is possible to solely activate the more reactive donor, effectively allowing the less reactive donor to act as a glycosyl acceptor (Figure 13).

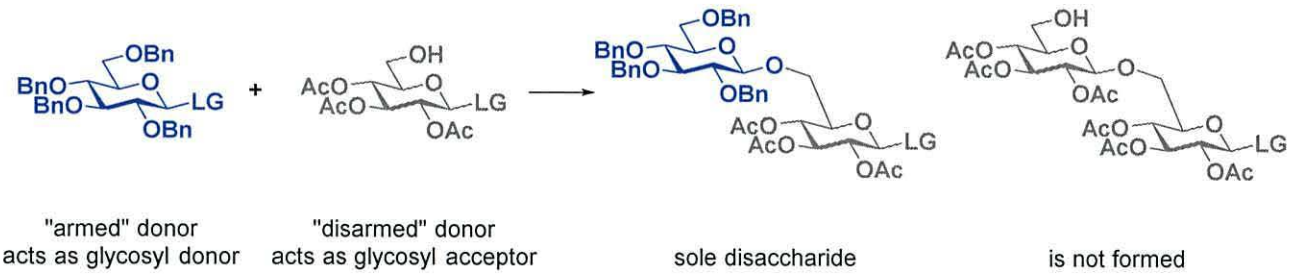


Figure 13. The armed-disarmed concept allows the promotion of a more reactive donor in the presence of a lesser, effectively allowing the less reactive donor to act as an acceptor.

2. Synthesis of Galacto-Oligosaccharide (GOS) Standards for Functional Food Studies

2.1 Introduction

Despite the widespread availability of formula milk and its convenience and ease of use, it is generally accepted that breast milk is the best option available to mothers raising new born children. Mothers' milk is an innate immune supplement, which complements the infant's developing immune system. In addition to traditional nutrients, there are several compounds in human milk which have documented immunological properties[27].One of the most important groups of compound in mother's milk are the Human Milk Oligosaccharides (HMOs). HMOs are the 3rd largest fraction in human breast milk[28] behind lactose and lipids [27] (Figure 14).

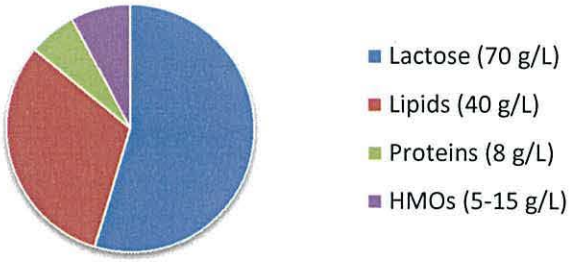


Figure 14. Composition of Human milk.

HMOs form a family of structurally diverse, unconjugated glycans with currently over 200 isomers isolated, varying in composition and glycosidic linkages[29]. HMOs are comprised of five monosaccharides; D-glucose (Glc), D-galactose (Gal), *N*-acetyl glucosamine (GlcNAc), L-fucose (Fuc), and sialic acid (Neu5Ac)[30] (Figure 15). All HMOs carry a lactose unit at the reducing end of the sugar[31]. It is thought that the biosynthesis of HMOs may be an extension of lactose biosynthesis which occurs in the Golgi apparatus, but the specific pathway remains largely uncertain[32]

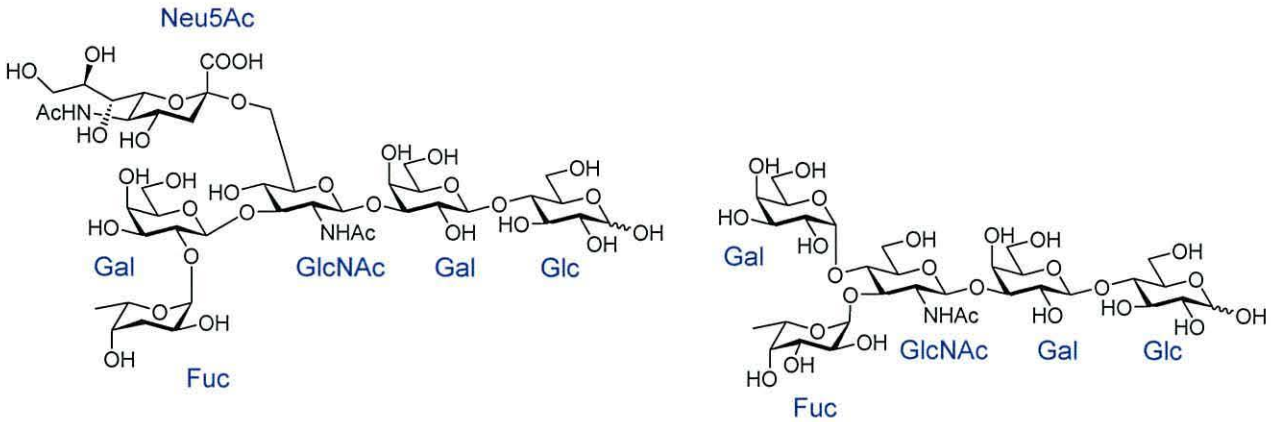


Figure 15. Examples of HMOs. Almost all HMOs contain a lactose unit at the reducing end. The non-reducing end contains fucose (Fuc) and sialic acid (Neu5Ac), and is generally the biologically active binding region of the sugar.

2.1.1 Function

HMOs are associated with a wide range of biological functions[32] (Figure 16).

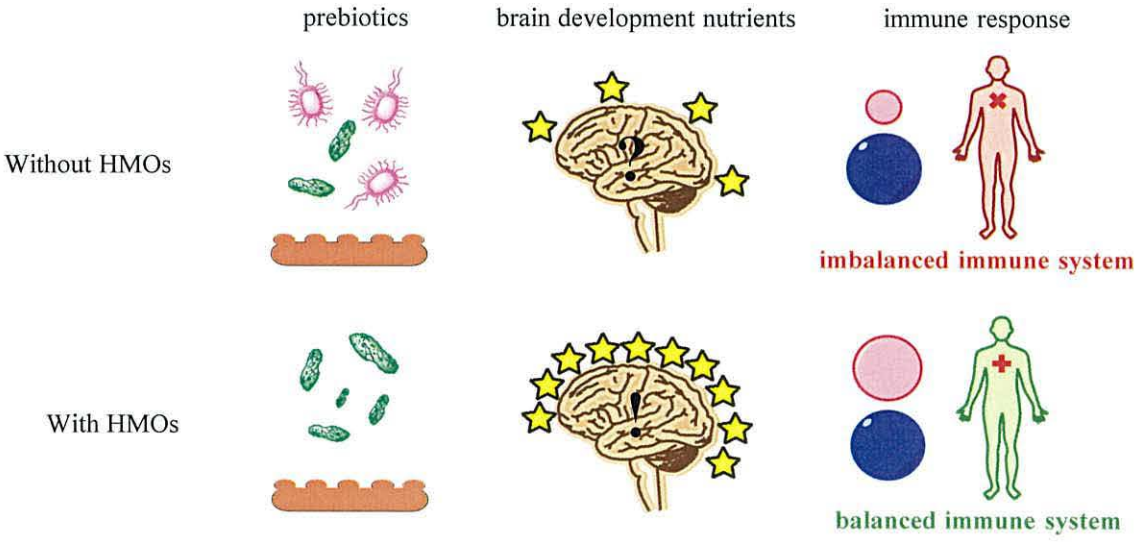


Figure 16. Examples of biological functions exhibited by HMOs.

T helper (T_H) cells help in a range of functions in the adaptive immune system, including activation of T cells and macrophages[32]. HMOs directly influence these immune system type responses, leading to a healthier, balanced immune system (Figure 16). They reduce low-level immunity responses, such as a reduction in allergies[33].

There is evidence that the nutritional value in HMOs plays a part in brain development[32]. Lucas et al. conducted a study on pre-term breast-fed infants at 18 months post-term[34]. Over the course of lactation for a mother, there are documented changes in the HMO type and concentration. The total amount of sugar is highest in the colostrum prior to birth, and decreases over time[35]. Milk produced by mothers with pre-term infants has been found to contain higher concentrations of oligosaccharides than that of term milk.

Lucas et. al. revealed that both motor and mental development were more advanced in infants fed a pre-term mixture as a supplement to mothers' milk. In addition, a later study by the same group recognised that breast-fed children were found to have superior IQ scores at 7-8 years, than those which were bottle fed[36].

HMOs are known to have anti-adhesive and anti-inflammatory activity[37]. Epithelial cells line the cavities and surfaces of structures and also form many glands in the body. HMOs structurally mimic these cells, acting as decoys to which pathogens bind, preventing infection[37–39] (Figure 17). In addition, HMOs can inhibit pathogens through competitive binding at the cell-surface receptor[40].

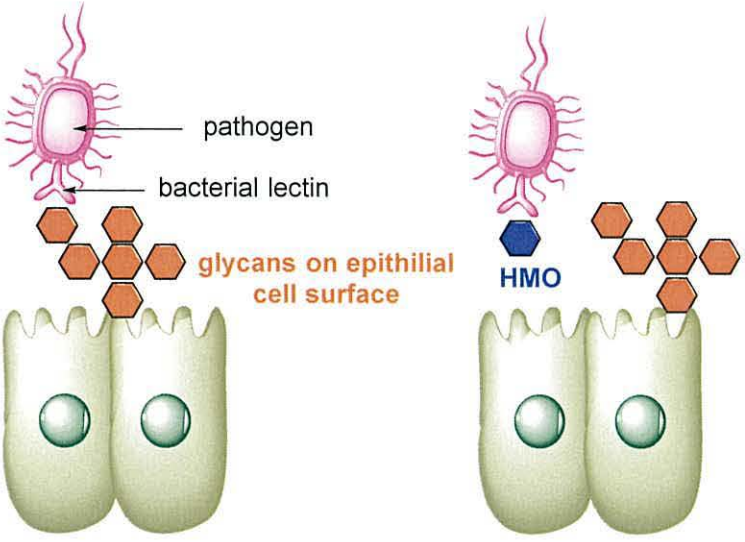


Figure 17. Anti-adhesive effects of HMOs. Most pathogens contain glycan-binding proteins which bind to glycans on the epithelial cell surface. HMOs structurally mimic these glycans, acting as lectin-ligand analogues, preventing binding by the pathogen. Adapted from Bode[32].

2.1.2 Role in metabolism

In 2001, Chaturvedi et al. studied the destination of HMOs in the intestines of infants[41]. Sugars were extracted from samples of infant's urine and stool, and compared to oligosaccharide samples obtained from their breast-feeding mothers. Concentrations of oligosaccharides were found to be higher in the faeces of infants than their mother's milk, and were also higher in the stool than compared with urine, indicating that they were not digested. This result confirmed that the HMO structures were not catabolised in the digestive system.

2.1.3 Prebiotics

Prebiotics are a class of compounds which, when consumed in sufficient quantities, stimulate the growth and/or activity of microorganisms in the colon, resulting in various health benefits[42] (Figure 16). Generally, prebiotics are short-chain carbohydrates (Figure 18).

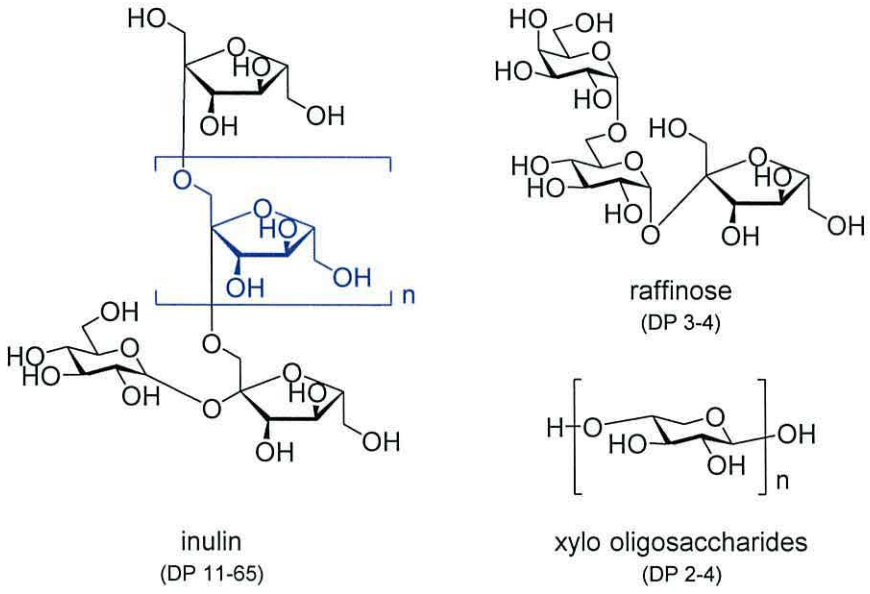


Figure 18. Structures of common non-digestible oligosaccharides.

Inulin consists of β(2-1) fructans and is manufactured by extracting from chicory root[43]. Both inulin and GOS are the most common commercially available prebiotics as they are relatively cheap to produce[44]. Raffinose is an example of soya-oligosaccharides, a class of compounds extracted from soya bean whey[44,45]. Xylo-oligosaccharides (XOS) can be chemically synthesised or produced via the enzymatic hydrolysis of xylan[46].

Prebiotics are minimally digested in the upper gastrointestinal tract, partially absorbed into the circulation, and remain there long enough to impose systemic effects[30] (Figure 19).

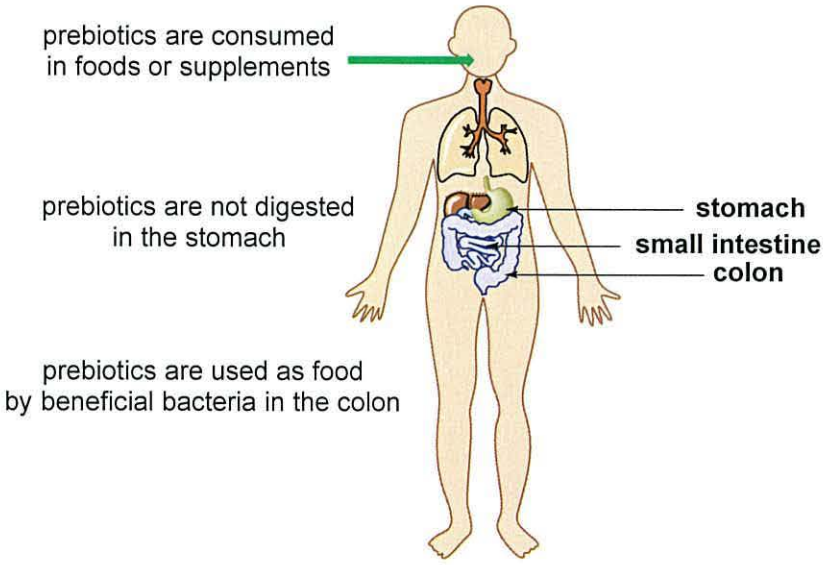


Figure 19. Prebiotics stimulate bacteria in the colon to give a variety of health benefits.

The prebiotic effect has been studied since the middle of the 20th century[47]. Gyorgy et al. found that human milk stimulated the growth of '*Lactobacillus bifidus*', which cow's milk could not[47]. The growth of this bacteria, now '*Bifidobacterium bifidum*', is known to be activated by *N*-acetyl glucosamine-containing oligosaccharides[28]. Lactobacilli are a large component of the microbacterial flora in the gastrointestinal tract of newborn infants[48]. This fact is thought to contribute to the lower incidences of gastrointestinal tract disorders in breast-fed than bottle-fed infants.

Whilst HMOs are regarded as exemplary standards, their potential is limited by the difficulties involved in manufacturing[49]. This has led to the search for functional alternatives, which can be used as food ingredients. Galacto-oligosaccharides (GOS) are considered to be the most prominent candidate for both digestive and immune health, owing to their structural similarity to HMOs[50]. GOS exhibit prebiotic activity, which as stated earlier, means they selectively enhance the growth and/or activity of microbes in the colon. Like HMOs, GOS also act as receptor decoys; exhibiting the highest adhesion inhibition of all prebiotics tested[51].

2.1.4 GOS

Commercially, GOS molecules are produced in a transgalactosylation reaction, by the enzymatic activity of β-galactosidase on lactose from whey[52]. Dietary carbohydrates are classified by size, degree of polymerisation (DP; the number of monosaccharide blocks combined), and physiological properties (digestible or non-digestible)[53]. GOS produced by enzymatic hydrolysis usually take the form of di-, tri-, tetra-, or penta-saccharides composed mainly of galactose, or for example, DP n, where n ranges from 2 - 5 respectively (Figure 20)[52].

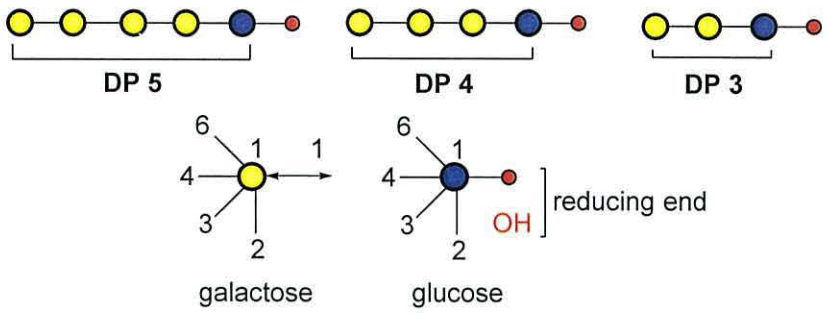


Figure 20. Dietary carbohydrates are classed by degree of polymerisation (DP n)

The starting material for GOS manufacture is lactose. It is present in the milk of all mammals except seals[54]. In bovine milk, it is found at a concentration of roughly 5 % w/v[55]. Rennet is an enzyme complex present in the stomachs of mammals having a diet of mostly milk, which is used in the cheese making industry[56]. Rennet causes milk to coagulate and separate into solid curds and liquid whey. The curds are filtered, pressed and then moulded into the blocks of cheese that are available commercially (Figure 21).

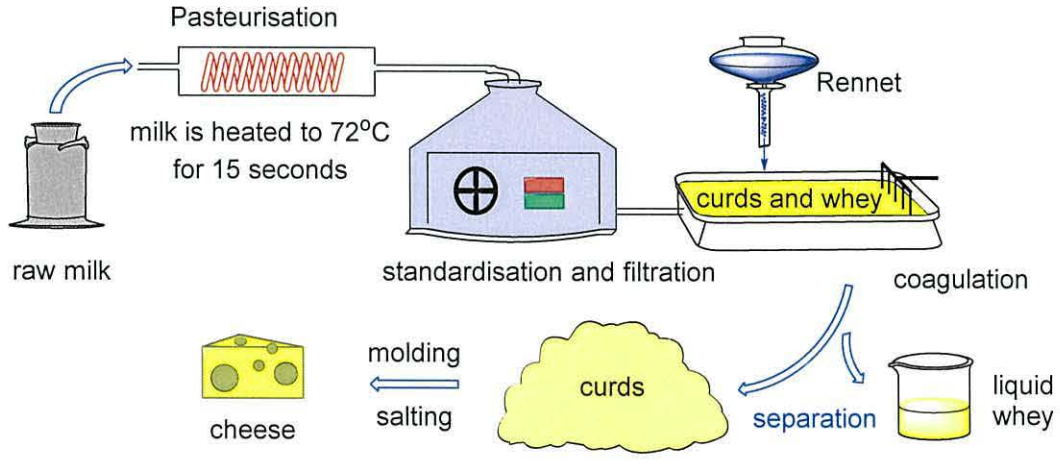


Figure 21. Processes in cheese production.

The liquid whey by-product makes up between 85 – 95% of the milk volume; to make 1 kg of cheese, about 9 kg of whey are produced[57]. Until recently, whey was considered to be a waste product in cheese production, and thus disposed. Whey is condensed to a powdered form to facilitate preservation, storage, transport, and manipulation[57]. Lactose can be extracted from whey powder using 70% ethanol[58].

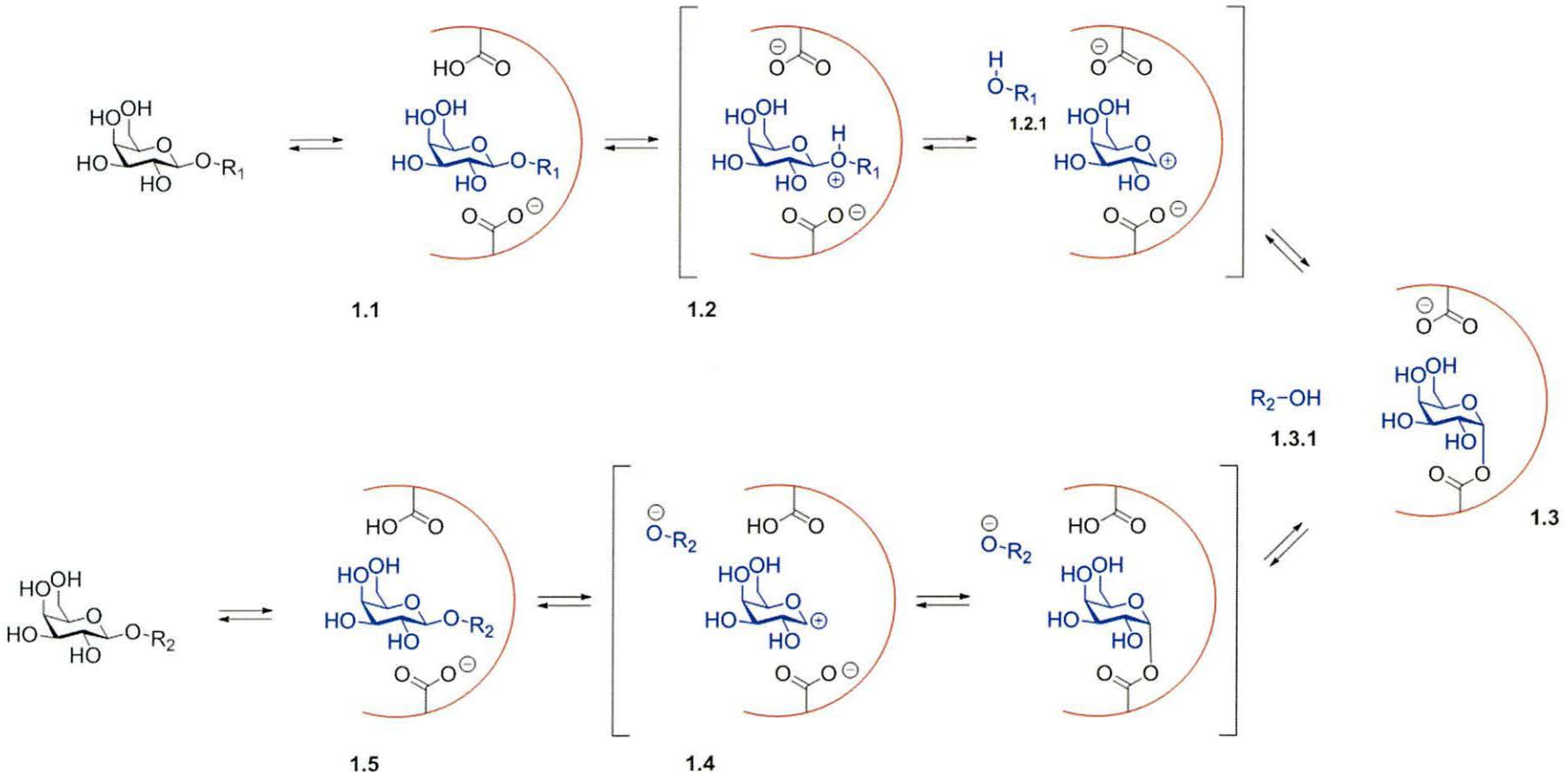
Commercial enzymes with β-galactosidase activity are used in the dairy industry to hydrolyse lactose into glucose and galactose molecules[59]. This process sweetens the milk whilst reducing the concentration of lactose, which is beneficial for consumers who are lactose-intolerant[55].

2.1.5 β-Galactosidase enzyme activity

β-Galactosidase is well known for its hydrolytic activity. However, the enzyme also polymerises galactosides under kinetic control. Polymerisation occurs when the sugar substrate attaches to the enzyme (Scheme 7).

If the enzyme accepts a lactose substrate, it can be broken down into both galactosyl and glucosyl (R_1 group in Scheme 7.1.3) components. The R_1 group (Scheme 7.1.3.1) may also represent an oligosaccharide, when the galactose substrate is part of a GOS molecule. It is currently hypothesised that a covalent bond forms between the enzyme and galactosyl unit, before attack from an acceptor nucleophile (Scheme 7.1.4)[60,61].

The products of this mechanism are dependent on the identity of the acceptor R_2 (Scheme 7.1.3.1). If this is a H_2O molecule, hydrolysis occurs, degrading the substrate. However, if the R_2 group (Scheme 7.1.3) is another form of saccharide (galactose, lactose, GOS), then galactosyl transfer occurs.



Scheme 7. Reaction mechanism for β-galactosidase activity on galactoside. Adapted from Gosling[54]

It is not known if the interactions in the active site differ between H_2O and saccharide acceptors (R_2). However, it is known that enzymes display different selectivities between these two acceptors, as different enzymes produce different GOS mixtures and yields at the same concentration of lactose[62]. The variations in structures of GOS produced enzymatically may also be due to differences in both the structure and mechanisms of different β-galactosidases from different sources[52].

The maximum concentration of GOS by enzymatic synthesis is dependent on temperature[63] and lactose concentration[64]. Hydrolysis by *E. coli* β-galactosidase is affected by pH, suggesting that it may be possible to control the rate of the reaction. GOS synthesis is also dependent on the type of enzyme used. Typically, GOS structures are a mixture of β (1→3), β (1→4), β (1→6) linked structures with DP 2 - 5[64].

Using the β-galactosidase derived from *Bacillus circulans* gives largely β (1→3) linkages, whilst on the other hand, *Kluyveromyces lactis* gives predominately β (1→4) oligosaccharides[49].

It has been shown that GOS produced from *Bifidobacterium bifidum* had a different structure and significantly higher prebiotic and bifidogenic efficacy, enhancing the growth of Bifidobacteria, when compared to a commercial GOS mixture[65]. This higher efficacy came from the presence of novel β(1→3) linked GOS structures synthesised by the bacteria, as well as the more common β(1→4) and β(1→6) linkages[65]. This opens up the possibility of manufacturing prebiotic GOS molecules with high selectivity towards specific bacterial groups.

2.1.6 Functional Food Potential

Functional foods can be described as foods which are marketed as promoting health or decreasing the risk of disease while enhancing our well-being[66]. GOS are used in a wide range of food products as non-digestible functional food ingredients[67]. The stability of GOS under acidic conditions, their high taste quality, and relatively low sweetness, make GOS suitable synthetic sweeteners. As they pass through the small intestine without being digested, they serve to be of low calorific value[67].

GOS are ideal for incorporation as a supplement into healthy food ingredients for specific demographics, such as children and infants, women, and the elderly due to these benefits. They are already marketed in infant formula[68], but are also an ideal ingredient for dairy products, fruit juice or soft drinks, and bakery[50].

2.1.7 Characterisation

GOS are produced enzymatically in the dairy industry using a variety of enzymes and large scale amounts of crude mixtures of sugars are produced (Figure 22). For example, in the enzymatic synthesis, glycosyl hydrolases assist in the hydrolysis of the glycosidic bonds in the lactose, giving both glucose and galactose components. Glycosyltransferases, on the other hand are enzymes which form natural glycosidic linkages, aiding in the biosynthesis of oligosaccharides.

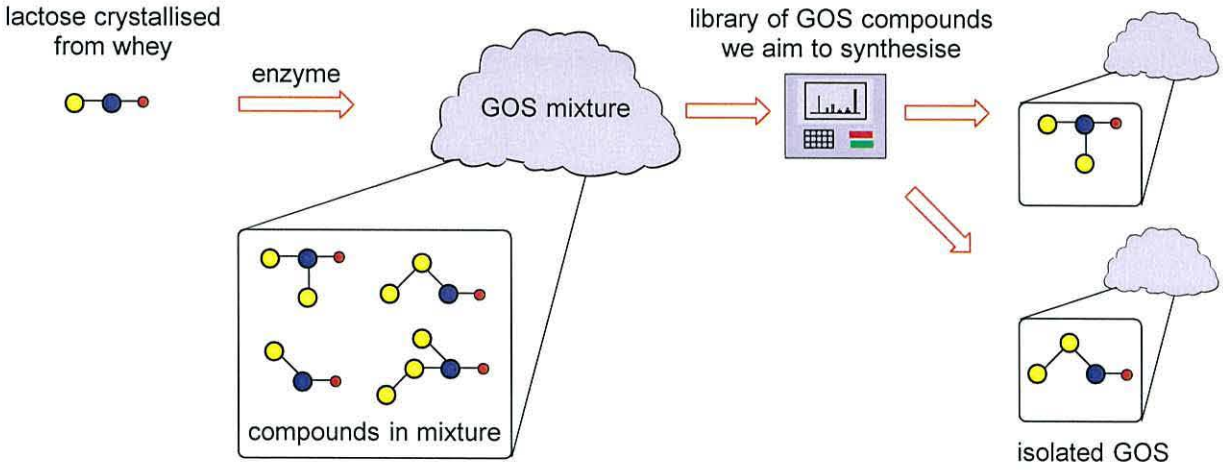


Figure 22. Visual representation of analysis/purification intended by production of GOS catalogue.

Analysis of GOS samples on line is often done by High Performance Anion Exchange Chromatography with Pulsed Amperometric Detection (HPAEC-PAD) fingerprinting. HPAEC-PAD is widely used analytical technique for the determination of a large range of carbohydrates, according to size, composition, anomericity and linkage isomerism[69]. HPAEC takes advantage of the fact that at high pHs, the hydroxyl groups of carbohydrates are partially or totally converted into oxyanions, allowing selective elution as anions[69]. Carbohydrates are oxidised at the surface of a gold electrode, generating an electric current which is detected. Pulsed amperometric detector (PAD) allows the analysis of underivatised carbohydrates.

This is a "crude" method of analysis; the resolution of the technique is very poor especially with complex mixtures. There is no separation of peaks, making isolation and characterisation impossible[67,70].

This technique is further limited due to the fact that there is a low availability of commercial standards for oligosaccharides with a high DP[69]. Peak identification is considered to be difficult, and peak assignment is generally based on assumption. We aimed to synthesise sets of analytically pure compounds to be used as standards for HPAEC, enabling more detailed analysis of commercially produced GOS.

The compounds synthesised in this project are a small section of what will be a large database. The collaborators in this project are committed to creating an expanding base of GOS standards. They plan on using the DP 3 structures synthesised in this project for characterisation using a number of analytical techniques such as HPLC and NMR[71].

The potential applications of such a database are widespread. It could, for example, improve the analytical aspects of research regarding the enzymatic production of GOS.

Analytical standards would also enable rapid evaluation and automated purification of GOS in a large scale industrial production. The Dutch dairy company Friesland Campina has already expressed interest in such a process, in order to make functional food products more readily available on the market[71].

2.2 Results and Discussion

2.2.1 Aim

The proposed target database would need to cover all possible combinations of branched and linear trisaccharides which could be produced enzymatically from lactose. Efforts were focussed on the formation of DP-3 trisaccharides (Figure 23).

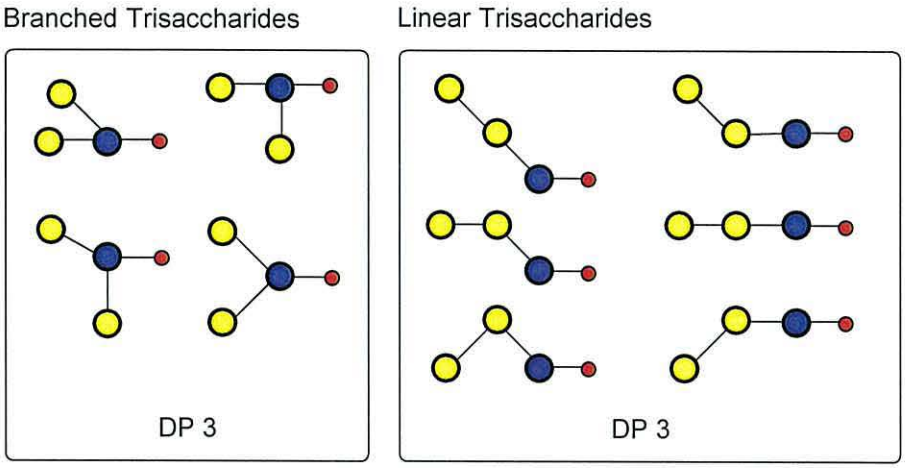
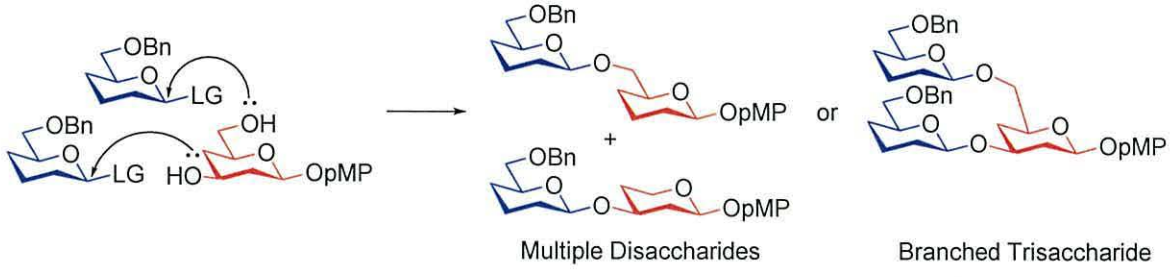


Figure 23. Target compounds for analytical database of standards.

2.2.2 Method

It was the aim to synthesise multiple disaccharides, and later trisaccharides, in potentially one step, using "random" glycosylations to couple our monosaccharide units together (Scheme 8). As described previously, in glycosylation couplings a promoter activates the leaving group on the anomeric position of the donor, which then allows a free hydroxyl on acceptor to attacks the newly formed carbocation. In a "random" glycosylation, more than one position on the acceptor is left free, and therefore allows attack by both positions.



Scheme 8. "Random" glycosylation may produce complex oligosaccharides in one synthetic step.

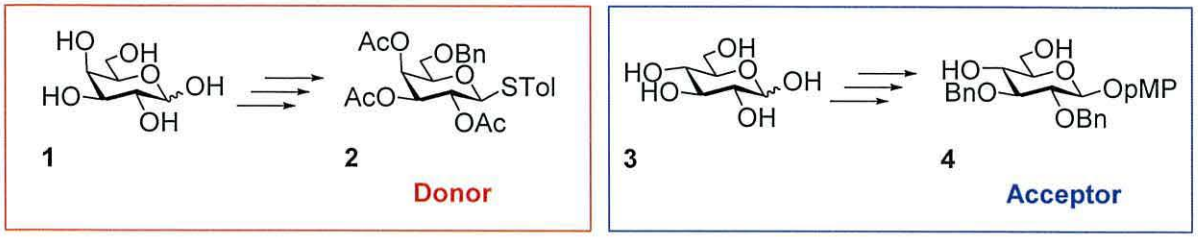
Using this random" glycosylation method, multiple complex carbohydrates were to be synthesised in the least amount of synthetic steps possible.

This was possible as GOS compounds all contain lactose at the reducing end of the sugar. Lactose consists of one galactose unit and one glucose unit in a β (1→4) linkage. The synthesis of oligosaccharides was carried out via various protecting group manipulation steps on two specific analogues, termed the donor and acceptor. The donor makes up the non-reducing end of the compound, galactose in this instance, whilst the acceptor forms the reducing end and therefore had to be a protected analogue of glucose. The donor and acceptor are synthesised separately and combined in a glycosylation coupling.

Using these donor and acceptor molecules, we anticipated that the same donor and acceptor building blocks could be used across a range of target molecules, increasing the overall efficiency of the synthesis.

2.2.3 Building Block Synthesis

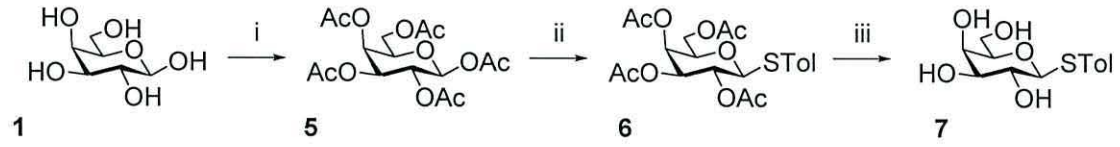
At first, building blocks **2** and **4** were synthesised via protecting group manipulation, from D-galactose **1**, and D-glucose **3** respectively (Scheme 9).



Scheme 9. Initial donor and acceptor pathways.

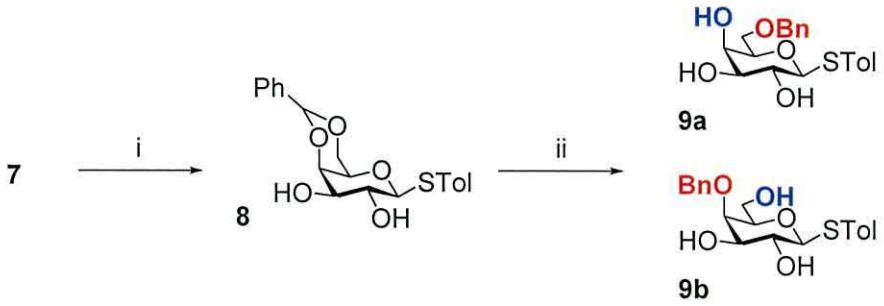
2.2.3.1 Donor

Compound **2** was prepared in 6 steps. Anhydrous D-galactose **1** was reacted with sodium acetate and acetic anhydride for 90 minutes at 140 °C, yielding the per-acetylated compound **5**[72]. Glycosylation of **5** in the presence of *p*-thiocresol and $BF_3{\cdot}(Et)_2O$ in dry CH_2Cl_2 gave compound **6**[73]; the β-selectivity of the product due to neighbouring group participation of the acetate group in the 2-position (Scheme 10).



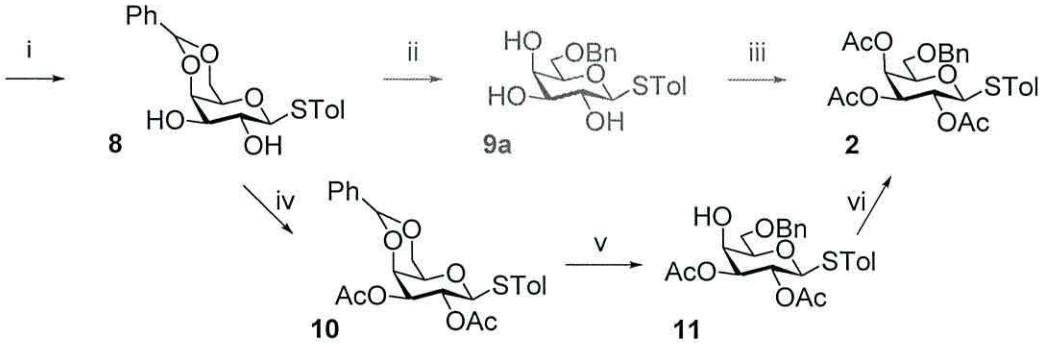
Scheme 10. Assembling the donor. Reaction conditions: (i) CH_3COONa, $(CH_3CO)_2O$, 1.5 h, 140 °C, 95 %; (ii) CH_2Cl_2, BF_3Et_2O, *p*-thiocresol, 3 h, 88 %; (iii) CH_3OH, Na,1.5 h, 98 %.

A Zemplén deacetylation afforded compound **7** in 98% yield[74]. The 4, 6-*O* positions of the tetrol **7** were protected with a benzylidene acetal using benzaldehyde and TFA. A characteristic peak at 5.5 ppm can be seen in the ^1H spectrum, which correlated to the CH of the benzylidene acetal. Reductive regioselective opening of **8**[75], employing $NaBH_3CN$ in 2N HCl/Et_2O, led to the formation of the 6-O benzylated derivative **9a** in 88% yield[13] (Scheme 11). However, in repeating this procedure varying ratios of regioisomers were obtained in low yields. Determination of regioisomers is done via NMR analysis. Once the ring is opened, either the '4' or '6' protons protected by the electron withdrawing benzyl ether can be seen to shift upfield.



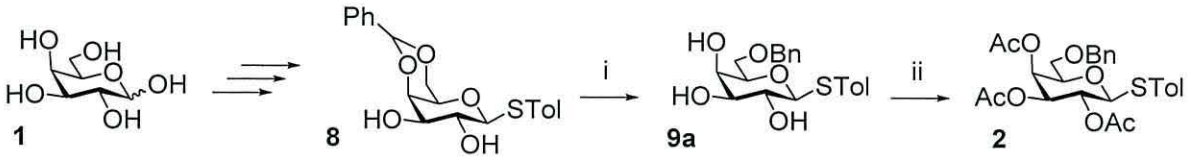
Scheme 11. Reductive ring opening with $NaBH_3CN$ gave varying isomers of **6**. Reaction conditions: Reaction conditions: (i) C_6H_5CHO, TFA, 88 % (ii) $NaBH_3CN$, THF, 2N HCl/Et_2O

As a work around, the diol of compound **8** was also acetylated, producing fully protected **9**[13] (Scheme 12). Compound **11** was produced using $NaBH_3CN$ and 2N HCl/Et_2O, and the intermediate immediately acetylated, yielding compound **12** in 92% yield.



Scheme 12. Work around to reductive regioselective ring opening issues regarding solubility. Reaction conditions: (i) C_6H_5CHO, TFA, 88 %; (ii) $NaBH_3CN$, THF, 2N HCl/Et_2O, 88 %; (iii) C_5H_5N, $(CH_3CO)_2O$, 1.5 h, 94 %; (iv) C_5H_5N, $(CH_3CO)_2O$, 1.5h, 97 %; (v) $NaBH_3CN$, THF, 2N HCl/Et2O 67%; (vi) C_5H_5N, $(CH_3CO)_2O$, 1.5h, 92 %.

As before, repeating the ring opening on **8** gave isomeric mixtures. The method of ring opening employing $NaBH_3CN$ in combination with HCl in diethyl ether is a standard procedure which has been utilised for a long time, dating back to its original discovery in 1981 by Garegg[76]. The use of triethylsilane in the presence of $BF_3 \cdot Et_2O$ as a regioselective reducing agent in a range of carbohydrates was first reported by Debenham et al[77] in 2000. By changing to this method, **9a** was obtained in 78 % yield as sole regioisomer. Acetylation with acetic anhydride in pyridine of triol **9a** produced **2** in 96 % yield (Scheme 13).



Scheme 13. Triethylsilane allowed regioselective ring opening of the benzylidene acetal, giving solely **9a**. Reaction conditions: (i) $C_6H_{16}Si$, $BF_3.Et_2O$, 3 h, rt, 78 %; (ii) C_5H_5N, $(CH_3CO)_2O$, 1.5 h, 96 %

Once the method of producing donor **2** was optimised and synthesis completed on the gram scale, our attention turned to the acceptor.

2.2.3.2 Acceptor

The 4,6-*O*-diol acceptor **4**[78] was synthesised in order to allow trisaccharide production to be completed via "random glycosylations" (Scheme 9). GOS and HMOs mainly contain lactose (Gal-β (1→4)-Glc) at the reducing end but in order to investigate the β (1→6) analogues and the competition of 1→6 vs 1→4 acceptors in subsequent glycosylations both linkages were desired.

Per-acetylated glucose **3** was prepared by acetylation of D-glucose **12** with acetic anhydride in pyridine[72]. The anomeric position of **12** was protected with a methoxyphenol ether[79]. This was prepared via glycosylation conditions, with 4-methoxyphenol in CH_2Cl_2, using either TMSOTf or $BF_3{\cdot}Et_2O$ as the promoter at 0 °C. Here, the anomeric acetate of **12** acts as a leaving group in the presence of a Lewis acid (Figure 24).

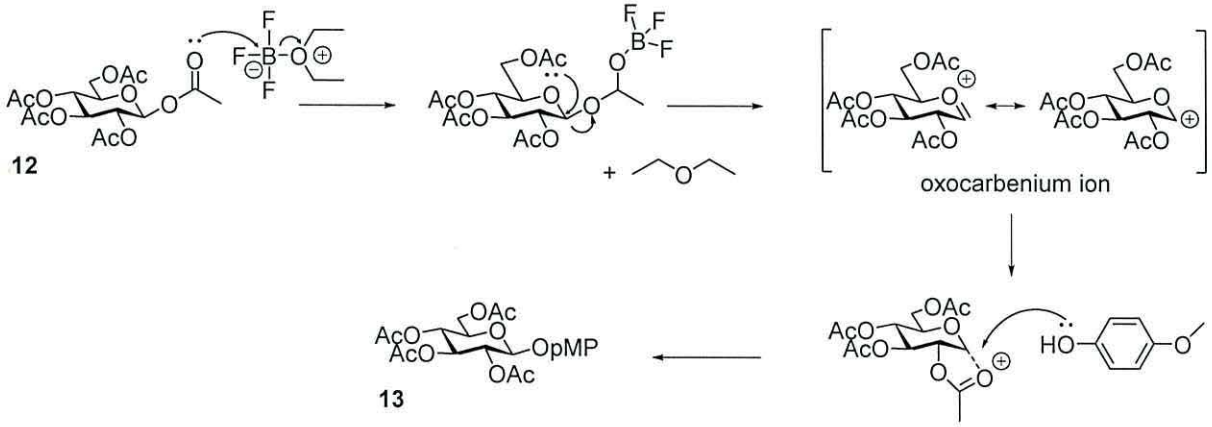
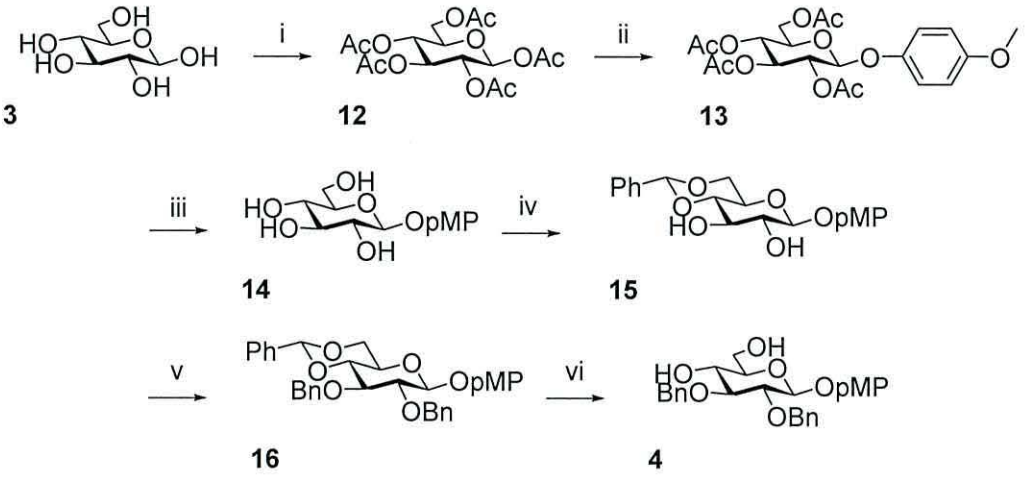


Figure 24. In the presence of a Lewis acid, the anomeric acetate acts as a leaving group producing an oxocarbenium ion.

Compound **13**[79] was produced in 86 % yield before deprotection under Zemplén deprotection was performed (Scheme 14). The 4 and 6-positions of compound **14** were protected with a benzylidene acetal using benzaldehyde and TFA. Protection of the free 3, 4-positions of **15** with benzyl ethers was performed using benzyl bromide and NaH in DMF. The benzylidene acetal of **16** was cleaved under acidic conditions in 79% yield, resulting in formation of acceptor building block **4**.



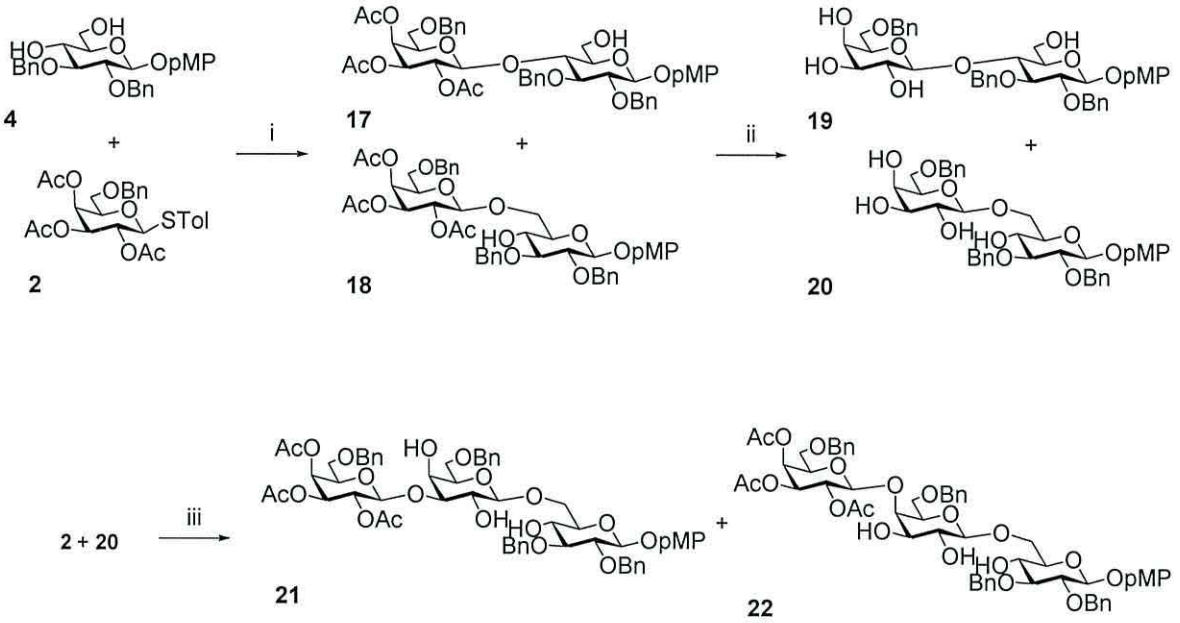
Scheme 14. First pathway for acceptor **4** synthesis. Reaction conditions: (i) C_5H_5N, $(CH_3CO)_2O$, 1.5 h, rt, 96% (ii) $CH_3OC_6H_4OH$, CH_2Cl_2, 0.75 h, TMSOTf, 0 °C, 3 h, 86%; (iii) Na, CH_3OH, 1.5 h, 98%; (iv) C_6H_5CHO, TFA, 0.5 h, 67 %; (v) BnBr, NaH, DMF, 18 h, 76 %; (vi) TFA, H_2O, rt, 78 %

2.2.4 Linear synthesis

The glycosylation coupling (Scheme 15) of donor **2** and **4** was carried out using 1.2:1 equivalents of donor to acceptor at -40 °C. *N*-iodosuccinimide, and AgOTf were used as activator system. At first, a test coupling was attempted with 10 mg of both donor and acceptor. Once TLC had revealed that the reaction had gone to completion with no presence of acceptor, the products were immediately purified via flash chromatography to avoid degradation. Analysis of NMR revealed the major product **18** was the β (1→6) linked glycoside. The '6a' proton is clearly shifted from the expected position at around 3.7 ppm, to a low field 4.1 ppm. The regioselective β(1→6) linkage was expected as the 6-position of the acceptor is a primary alcohol.

The reaction was repeated once more, this time at room temperature, using the same conditions. The reaction was quenched after 0.75 h, once TLC had revealed

the reaction had gone to completion. However, unlike the coupling at -40 °C, at room temperature, the coupling yielded only one major product, the 2,3,4-Tri-*O*-acetyl-6-*O*-benzyl-D-galactopyranose-β(1→6)-2,3-Di-*O*-benzyl-1-methoxyphenyl-β-D-glucopyranoside **18**.



Scheme 15 The successfully produced mixture of linear trisaccharides in 3 steps with donors, **7** and **14**. Reaction conditions: (i) NIS/AgOTf, CH_2Cl_2, -40 °C, 1.5 h, 86 %; (ii) CH_3OH, NaOMe, 1.5 h, rt, 94 %; (iii) NIS/AgOTf, CH_2Cl_2, rt, 1.5 h.

The two disaccharides, **17** and **18**, were subject to Zemplén deacetylation, yielding compounds **19** and **20**. After purification and subsequent deprotection, the 1→4 disaccharide **19** was obtained in negligible amounts. The β(1→6) linked compound **20** was isolated in 82 % yield.

The deprotected disaccharide **20** was subjected to glycosylation coupling with donor **2**. The reaction produced as anticipated two new spots visible on TLC. Several attempts at separation via flash chromatography, including automated purification did not give separation of the two compounds. The mixture was subject to NMR and MALDI-TOF analysis, which appeared to be inconclusive. MALDI-TOF data suggests the presence of two trisaccharides (Figure 25) and a tetrasaccharide (Figure 26).

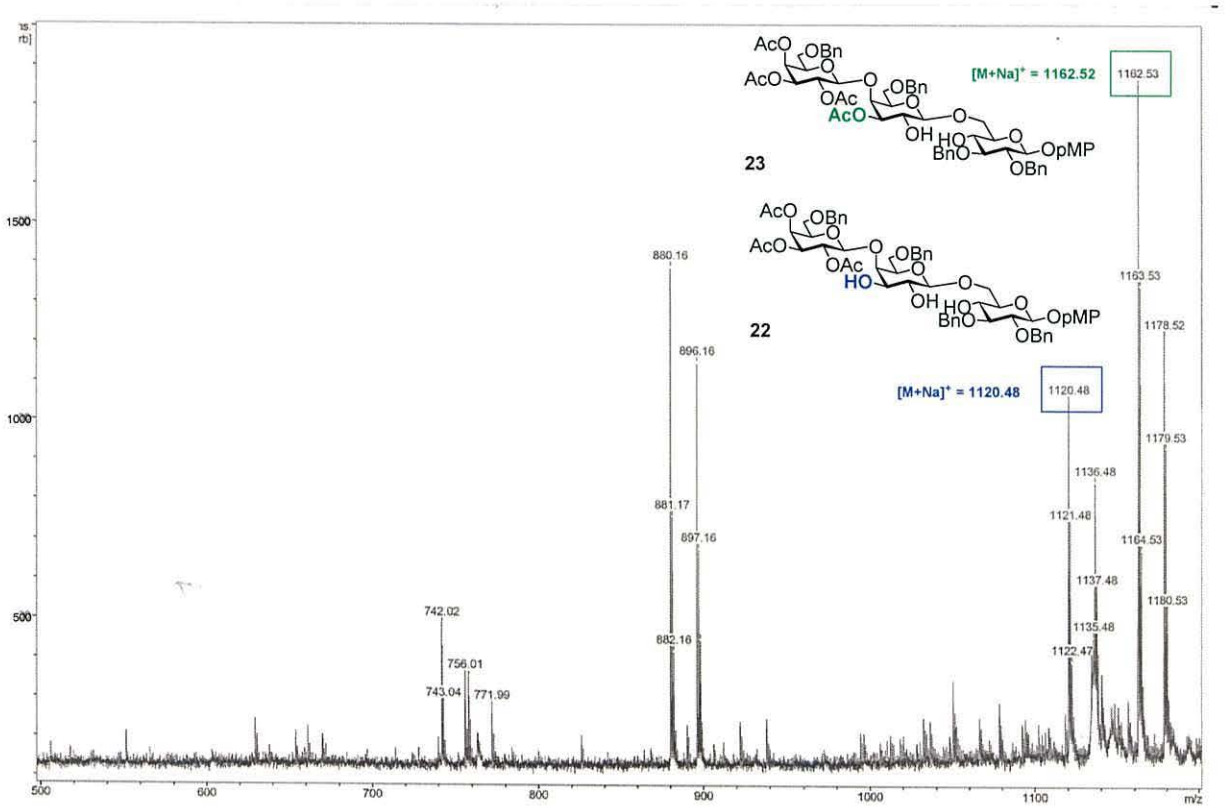


Figure 25. Potential structures hypothesised by MALDI analysis in the glycosylation reaction mixture.

An expected peak for trisaccharide **22** or its 3'-*O*-regioisomer was found at 1120.48 Da. Acetyl migration can occur in dilute alkalis, or when the pH of the reaction falls below 3[80]. In this instance, the peak at 1162.53 Da accounts for the addition of one migrated acetate, forming trisaccharide **23** or a regioisomer of it. Peaks were also observed further along the spectra at 1498.93 Da, suggesting formation of a branched tetrasaccharide **24** (Figure 26).

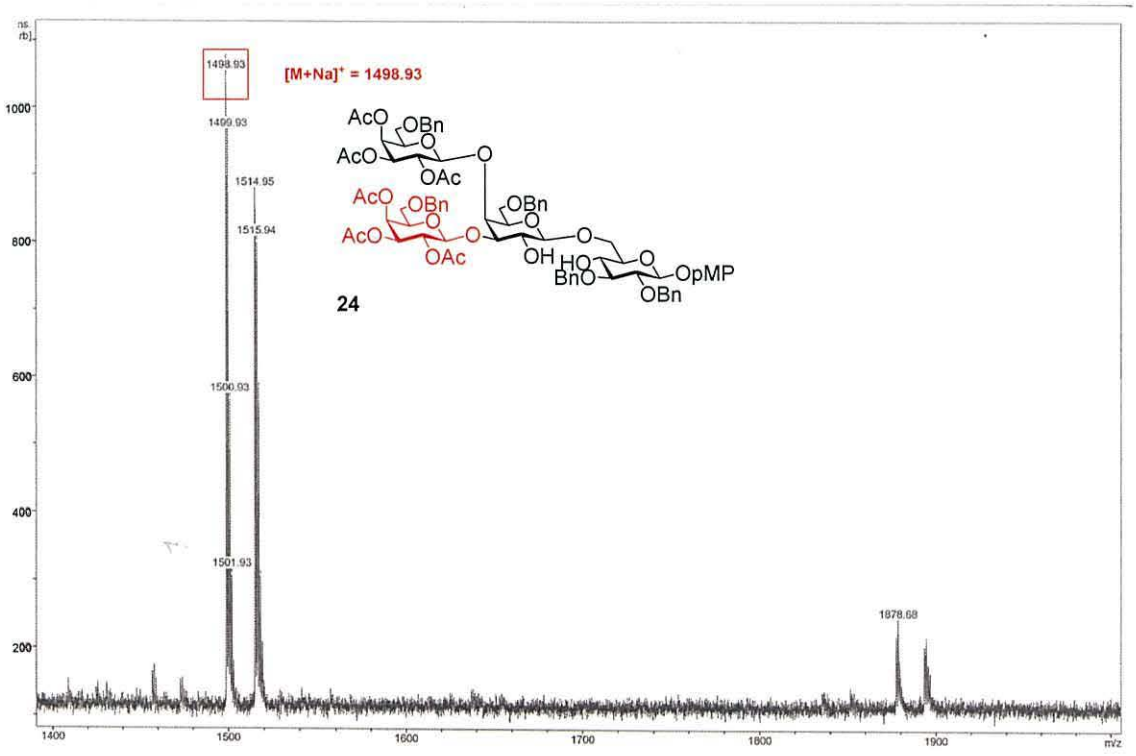


Figure 26. Potential tetrasaccharide structure hypothesised by MALDI analysis.

The two dimensional NMR technique HSQC (**H**eteronuclear **S**ingle **Q**uantum **C**oherence) also gives insight into the nature of the mixture (Figure 27). HSQC is a correlation technique between both ^{1}H and ^{13}C spectra. HSQC is a very valuable technique, especially in oligosaccharide synthesis, where it is possible to pick apart compacted and overlapping signals in spectra. The circular plots in the centre of the spectra denotes coupling between hydrogen atoms (^{1}H spectra above) to neighbouring carbons (^{13}C spectra to the left), one bond away.

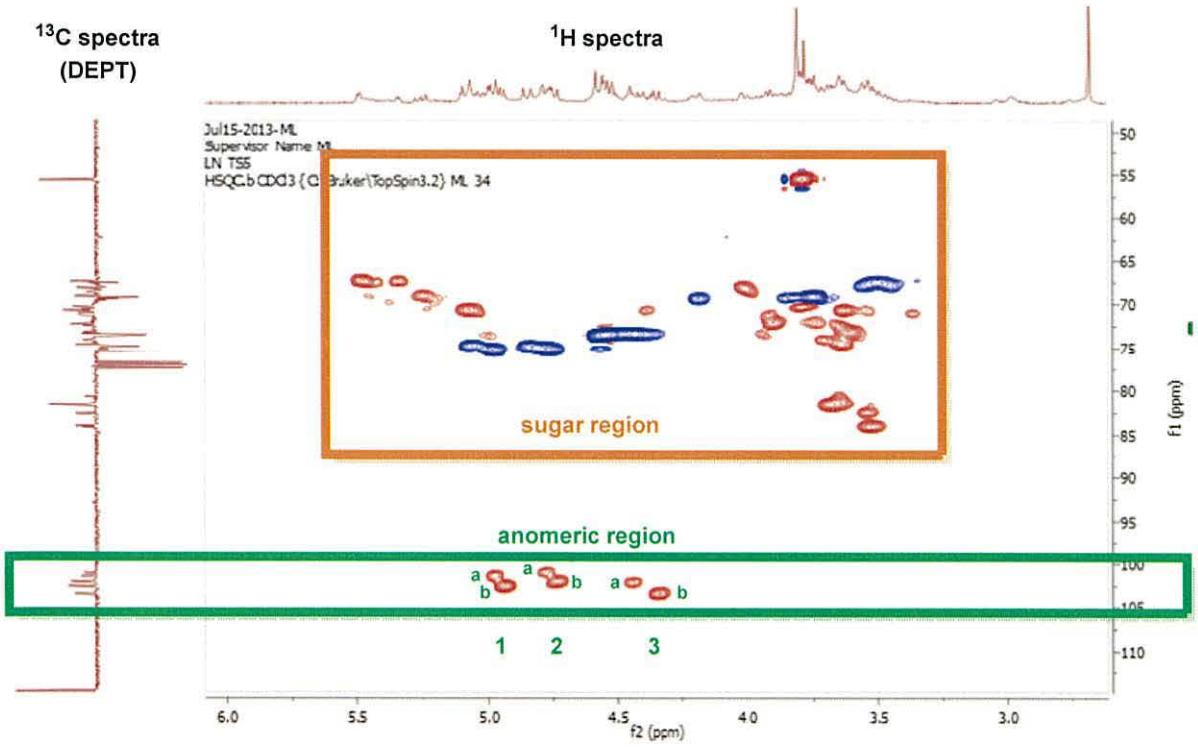


Figure 27. 6 sets of anomeric peaks are visible in the HSQC spectra of the trisaccharide mixture.

The "general sugar region" highlighted in orange, can be seen between 50 – 90 ppm on the ^{13}C spectra. The anomeric carbons are characteristically shifted to between 100-105 ppm, highlighted in green. Here, one can see 6 different correlation signals with the ^{1}H spectra in similar regions. The two sets of signals labelled **1** and **2** are close to the same position, whereas the set of signals labelled 3 are slightly shifted apart. It seemed as the less intense signals **1a**, **2a** and **3a** all belonged to one system, whilst **1b**, **2b** and **3b** belonged to a second product.

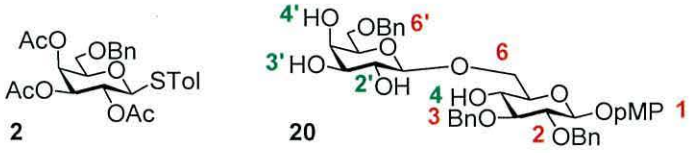


Figure 28. Positions on disaccharide acceptor **20** available (in green) for coupling to donor **2**.

Different shifts are usually associated with changes in the electron density, such as in different linkages and positions on the sugar. Purification would allow determination of stereochemistry and the identification of possible regioisomers. However, the most likely structures can be considered. As the primary alcohol is protected by a benzyl ether, the 4' position should be the most reactive, and therefore primarily the position where the glycosidic linkage would form (Figure 28).

2.2.5 Branched synthesis

2.2.5.1 Production of Acceptors

With the partial success of the linear synthesis, the attention was turned to the branched disaccharides (Figure 29). With a large amount of donor still available after the linear synthesis, we decided to first focus on the synthesis of the required acceptors. In the case of the branched trisaccharides, the acceptors had to have a minimum of two free positions to allow attack. The rest of the sugar must be suitably protected by activating groups, as to improve the reactivity.

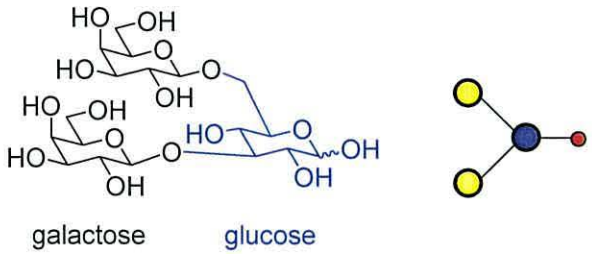
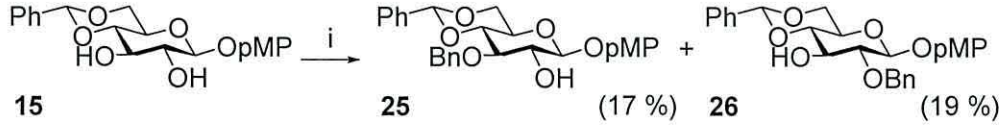


Figure 29. Branched GOS structures have two galactose units stemming from one glucose monomer.

Much of the synthesis of the acceptors used in the linear trisaccharide synthesis was repeated on a large scale. 1-Methoxyphenyl-β-D-glucopyranoside **14**[81] was prepared in bulk, as this compound was an extremely suitable starting point for further protection. To start, the 4 and 6-positions of compound **14** were again protected with a benzylidene acetal using benzaldehyde and TFA to afford **15**.

The phase transfer catalyst tetrabutylammonium hydrogen sulfate was used in the monobenzylation of the diol **15**[82], using a procedure derived by Garegg et al[4] (Scheme 16). Using a phase transfer catalyst reduces the rate at which dibenzylation occurs, and gives a mixture of monobenzylated compounds **25** and **26**[83].



15 **25** **26**

Scheme 16. The use of a phase transfer catalyst allowed monobenzylation of the diol, **12**. Reaction conditions: (i) $Bu_4N(HSO_4)$, BnBr, 5% NaOH (aq.), reflux, 68 h.

Depending on the reagents used, the benzylidene acetal on these two monobenzylated glycosides can be opened to either 4- or 6-position. Using this technique, it is possible to synthesise four acceptors in one step each, from the two monobenzylated compounds (Figure 30).

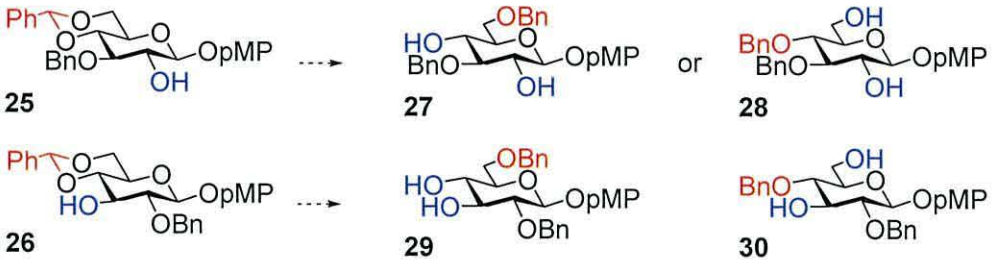
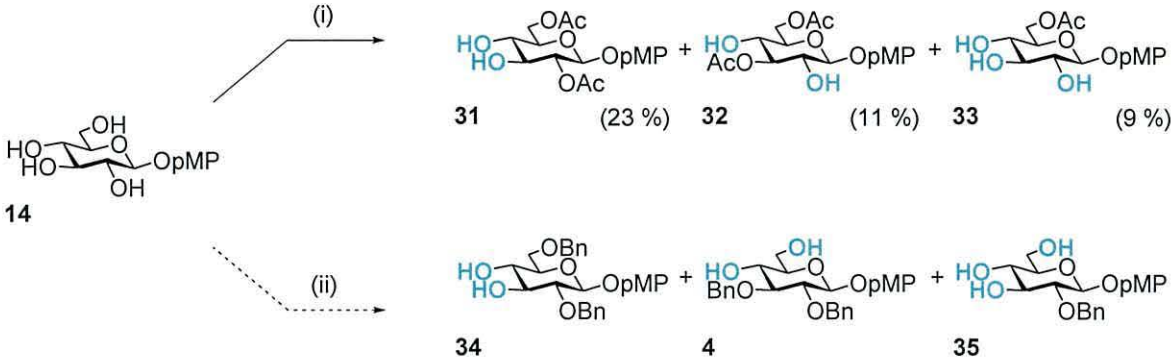


Figure 30. The benzylidene acetal is a very versatile protecting group. Depending on the reagents used, 4 acceptors can be produced from 2 precursors.

This method of ring opening can produce multiple isomers as discussed earlier in the "linear synthesis". However, the monobenzylation is slow taking place over 72 hours and did not produce very high yields. Alternative routes were researched to produce the desired acceptors in a smaller time frame.

Acetates are very valuable protecting groups as they are very easily removed in quantitative yields making the later deprotection steps much simpler. However, acetates also have disadvantages in that they can migrate quite easily in acidic conditions, and are considered deactivating. Several acetylated regioisomers were synthesised in a one step process using a convenient procedure described by Dong et al[84]. The regioselective acetylation of compound **14**, in the presence of dibutyltin oxide, proceeds via a stannylene acetal intermediate. Depending on the polarity of solvent and equivalent of acyl reagent used, regioselectivity can be achieved in different positions of the molecule.

Thus, selective acetylation using this method was performed with tetrol **14**. Three isomers were isolated from the reaction mixture. The 2, 6-protected diol **31** was the major product at 23 % yield; the 3, 6-protected diol **32** and mono substituted acceptor **33** were isolated at 11 and 9 % respectively (Scheme 17). Using this route, three additional alkylated acceptors (**34**, **4** and **35**) may also be synthesised[84]. The selective benzylation could be carried out according to a procedure by Zhou et al[85].



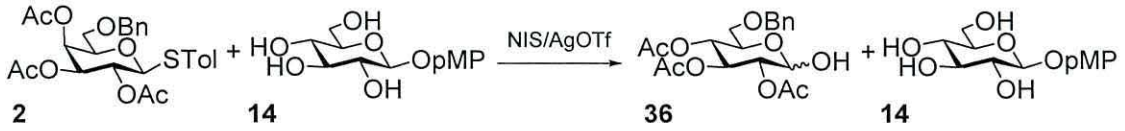
Scheme 17. Regioselective control over multiple protections in an organotin-mediated selective acylation[84,85]. Reaction conditions: (i) (a) Bu_2SnO (2.2 equiv.), CH_3OH, reflux, 3 h; $(CH_3CO)_2O$ (2.2 equiv.), DMF, rt, 18 h; (ii) Bu_2SnO (3.3 equiv.), Tol., reflux, 1 h; BnBr (4 equiv.), 100 °C, 12 h.

Similarly to the organotin-mediated selective protection described by Dong[84], the tetrol **14** along with dibutlytin oxide was refluxed in toluene. Upon the removal of toluene, benzylation can be performed using benzyl bromide in DMF.

These procedures are extremely useful in producing a wide range of acceptors in the minimal amount of synthesis steps, although the cost of efficiency was at the expense of yield and separation.

2.2.6 Coupling Issues

As the unprotected acceptor, 1-methoxyphenyl-β-D-glucopyranoside **14**, was produced in bulk, a glycosylation coupling with donor **2** was attempted, despite expected solubility issues. As anticipated, the overall reactivity was too low and no reaction was observed. The thiotolyl leaving group of **2** was activated by the NIS promoter, but the acceptor was not reactive enough to attack the carbocation formed (Scheme 18). The hydrolysed donor **36** (56 %) was recovered by flash chromatography.



Scheme 18. Hydrolysis of donor during glycosylation couplings

Unfortunately, donor **2** did not perform well in combination with other acceptors and varied glycosylation conditions (Table 1).

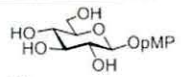
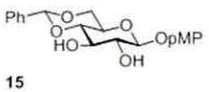
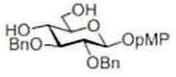
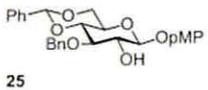
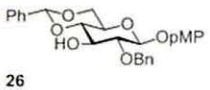
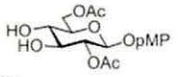
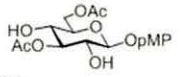
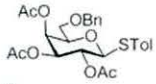
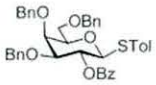
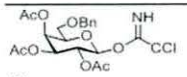
Donors	Promoter	Acceptors						
		 14	 15	 4	 25	 26	 31	 32
 2	NIS AgOTf (2 equiv.)	N Dioxane, rt (D/A, 1.2:1) Et_2O, rt (D/A, 1.5:1) DMF, rt (D/A, 1.5:1)	N Dioxane, rt (D/A, 1.5:1) CH_2Cl_2, rt (D/A, 1.5:1)	rt = 1 product -40 °C = 2 products CH_2Cl_2, 2 temp. (D/A, 1.2:1)	N CH_2Cl_2, rt (D/A 2:1)	N CH_2Cl_2, rt (D/A 2:1)	N CH_2Cl_2, rt (D/A, 1.2:1) Dioxane, rt (D/A, 3:1)	N CH_2Cl_2, rt (D/A, 1.2:1) Dioxane, rt (D/A, 1.2:1)
	DMTST (3 equiv.)	N Dioxane, rt (D/A, 3:1)	N Dioxane, rt (D/A, 3:1)				N CH_2Cl_2, -15 °C (D/A, 3:1)	
 37	NIS AgOTf (2 equiv.)	N Dioxane, rt (D/A, 3:1)	N Dioxane, rt (D/A, 3:1) CH_2Cl_2, rt (D/A, 1.5:1)				N Dioxane, rt (D/A, 3:1)	N Dioxane, rt (D/A, 3:1)
	DMTST (3 equiv.)	N Dioxane, rt (D/A, 3:1)	N Dioxane, rt (D/A, 3:1)				N CH_2Cl_2, -15 °C (D/A, 3:1)	N CH_2Cl_2, -15 °C (D/A, 3:1)
 38	$BF_3 \cdot (Et)_2O$ (catalytic)	N CH_2Cl_2, rt (D/A 3:1)					N CH_2Cl_2, rt (D/A, 3:1)	

Table 1. Annotation of the different donor/acceptor combinations tested. Reactions were repeated multiple times with different reagents in order to test efficacy of material. Donor was recovered in all cases for reuse in future couplings. Donors (on the left) were reacted with acceptors (along the top). Promoter system is listed next to each donor, whilst reaction conditions (solvent system, temperature, donor and acceptor ratio), and products are listed at the point of intersection.

Various acceptors were tested under different glycosylation conditions in combination with donor **2**. However, no successful coupling conditions were found (Table 1). In the majority of cases, hydrolysis was seen within the first 5 minutes, with no oligosaccharide production visible after 1 hour.

To address the issues, a risk analysis approach was used to thoroughly scrutinise all methods and reagents. Reaction conditions were investigated to see what effects solvent, temperature and equivalency of reagents had on the success of the coupling.

By altering the solvent system, potential problems with the solubility of material were ruled out. Solvents tested included DMF, CH_2Cl_2, Et_2O and 1, 4-dioxane. The amount of solvent used in the reaction was decreased, in case the reaction was too dilute to proceed. The effect of temperature was also examined, experimenting with temperatures between -40 °C to 40 °C. We increased the temperature of the reaction to increase the reaction rate, in an effort to "kick-start" the reaction. Increasing the amounts of promoter and catalyst was also tested.

Once these tests were deemed to be unsuccessful, attention was focussed towards validating the methods, in an effort to address the issues with hydrolysis. Reaction vessels were heated to 140 °C for 24 hours prior to use to remove any moisture. Heated glassware and instruments were cooled to room temperature in a desiccator under vacuum, before being flushed with nitrogen several times. All solvents were dried over sieves for a minimum of 12 hours prior to use.

The inquiry into the problems did not yield any issues with regards to reaction conditions. As the majority of acceptors were protected with acetates, benzylidene acetals, or largely unprotected, the reactivity of donor/acceptor combinations were mismatched.

In the midst of the issues faced, we were able to reproduce the successful coupling of donor **2** and acceptor **4**, as seen in the linear synthesis. This acted as a standard, to which we could test our reaction conditions; if this continued to work, then the problem should not originate from our handling. The rest of the donor and acceptor

combinations gave no coupling, with the hydrolysed donor **37**[86] recovered in all cases.

It was therefore essential to look for reactive donors. In oligosaccharide preparation, it is not possible to think of the building blocks as simply reactive or unreactive; reactivity of each species has to be considered as a scale. If donor and acceptor molecules are on opposite ends of said scale, then the reaction will not proceed. It is therefore very important to be able to "tune" the reactivity of both donor and acceptor to be able to "match" the activity[87].

2.2.7 Reactivity

Many factors affect the reactivity of a donor, such as protecting groups, leaving group, and nature of the carbohydrate skeleton. Douglas et al[87] introduced the idea of deactivation factors of protecting groups on thiomannoside donors. Using mannose as a comparative system, we extrapolated from Douglas' methods of calculation to predict the reactivity of our galactoside donors. We approximated that the inclusion of a benzyl ether on the 6-position had made donor **6** around 8 times more reactive than per acetylated donor **6** (Figure 31).

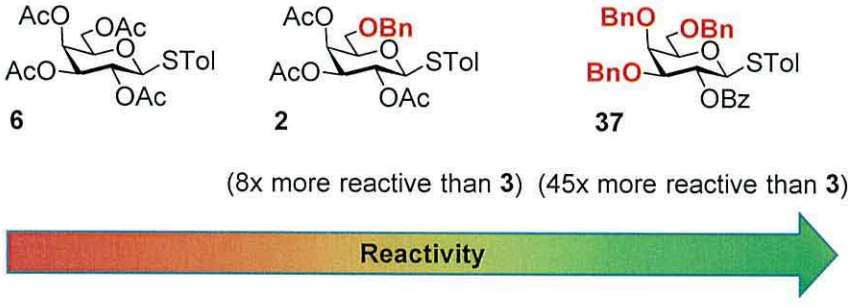
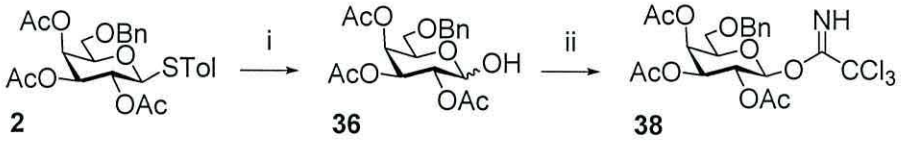


Figure 31. Projected reactivities of glycosyl donors used according to values by Douglas et al[87].

Another donor, *p*-tolyl 2-*O*-benzoyl-3, 4, 6-tri-*O*-benzyl-1-thio-β-D-galactopyranoside **37** had been synthesised previously in the lab for a different project, and was therefore readily available. The three activating benzyl groups on the 3, 4, 6-positions of **37** indicate that it is at least 5.5 times more reactive than the already reactive donor **2**.

In addition to using this more reactive donor, hydrolysed material recovered from all previous couplings was also recycled. Using standard NIS glycosylation conditions, with the addition of a small amount of H_2O, we purposefully hydrolysed some donor **7** as a standard for TLC and NMR. The addition of the more reactive trichloroacetimidate as leaving group at the anomeric position of **36**[86] was successfully carried out using trichloroacetonitrile with NaH, yielding donor **38** in 76 % yield (Scheme 19)[88].



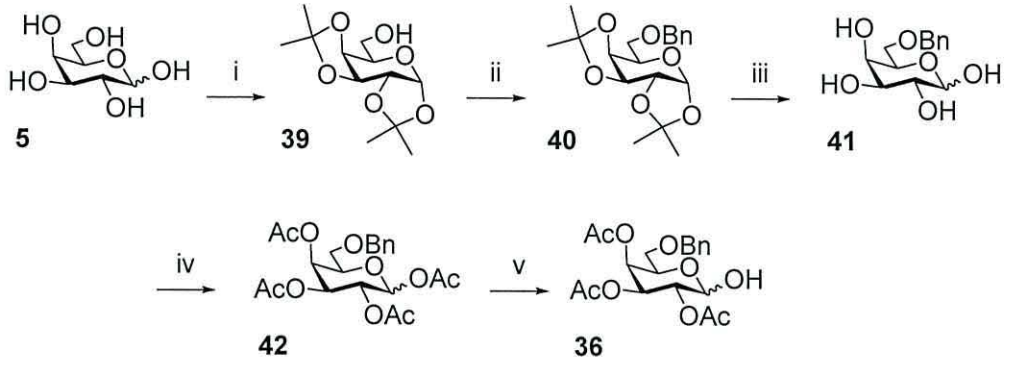
Scheme 19. Trichloroacetimidate leaving group synthesis from purposefully hydrolysed donor. Reaction conditions: (i) NIS, CH_2Cl_2/H_2O (9:1); (ii) CCl_3CN, NaH, CH_2Cl_2, 1 h, rt, 76%.

The amount of **36** recovered from previous glycosylation reactions was minimal; in order to make this donor viable, it was also prepared on a large scale. Using methods first described by Levene and Meyer[89] and more recently by Sugimoto[90]. Thus, 1,2:3,4-Di-*O*-isopropylidene-D-galactpyranose **39** was synthesised by addition of a catalytic amount of H_2SO_4 into a solution of anhydrous D-galactose **1** in acetone. As this reaction produces H_2O as a by-product, and as H_2O acts as a competitive nucleophile in this reaction, the procedure must be carried out in dry conditions in the presence of the drying agent $CuSO_4$.

We tailored the method to increase yields. Increasing the amount of the catalytic H_2SO_4 (acetone/H_2SO_4, 20:1) gave improved yields. Anhydrous $CuSO_4$ was heated in an oven at 130 °C for 48 hours prior to use in order to remove any excess moisture. The reaction must then be monitored closely by TLC over a 24 – 48 hour period. The reactants have to be left long enough to react, although stirring too long in the presence of the harsh acid can lead to degradation of products. We found the optimum reaction time to be between 24 – 30 hours, with the highest reported yield of **39** to be 92 %. We did not at any point observe full conversion of reactants.

The 6 position of this intermediate was then immediately protected with a benzyl ether in basic conditions give 6-*O*-benzyl-diisopropylidene-D-galactopyranose **39**[89] in 22 % yield[90]. The isopropylidene groups were cleaved by refluxing in a mixture of 1N

HCl in 1, 4-dioxane[91] before acetylation to give 6-*O*-benzyl-1, 2, 3, 4-tetra-*O*-acetyl-D-galactopyranose **42**[91] in a 92 % yield. Selective anomeric deprotection with was carried out immediately, refluxing with morpholine in CH_2Cl_2 for 3 hours to give **36** in 82 % yield (Scheme 20).

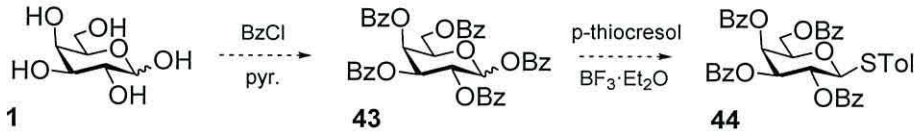


Scheme 20. Bulk synthesis of trichloroacetimidate donor using method described by Sugimoto[90]. Reaction conditions: (i) acetone, conc. H_2SO_4, $CuSO_4$, 26 h, rt, 92 %; (ii) BnBr, NaH, DMF 18 h, rt, 22 %; (iii) 1N HCl, 1,4-dioxane, 4h, reflux, 89 %, (iv) C_5H_5N, $(CH_3CO)_2O$, 1.5 h, 50 %; (v) morpholine, CH_2Cl_2, 90 min, reflux, 82 %.

Donors **37** and **38** were subject to the same glycosylation conditions as previously stated, with promoters, equivalencies, and temperatures varied (Table 1). No reaction was observed with any donor/acceptor combination. At this point, we decided to look for another donor that would suit the reactivity of the acceptor molecules.

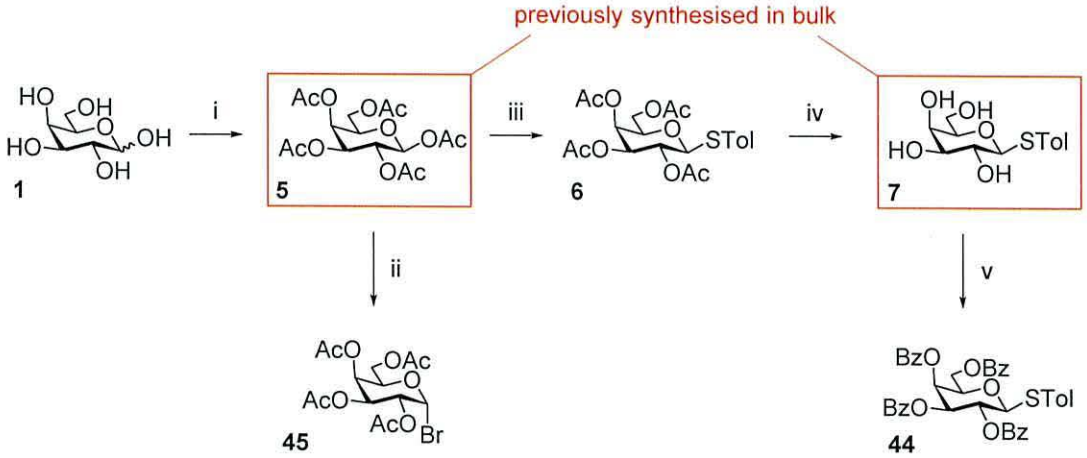
When planning the donor synthesis, we had to ensure that a carbonyl was present on the 2 position of the galactose. This ensured selective β-formation of the glycosidic bond by neighbouring group participation. We also have to take into account the comparative deactivation factors extrapolated from the calculations of Douglas et al.[87], who used benzyl ethers as a standard. Benzyl ethers are highly activating protecting groups, due to their ability to donate electrons into the ring. The failure of any successful coupling between the highly reactive donor **37** and our acceptors suggests that we need a donor which is not overly reactive to match the acceptors.

The synthesis of donor **44**[13] would be prepared in 2 steps (Scheme 21). Benzoyl esters are substituted onto the sugar by stirring in benzoyl chloride and pyridine for 3 hours, before the addition of a thiotolyl leaving group under glycosylation conditions.



Scheme 21. Potential pathway for the preparation of donor **44**.

However, as compounds **5** and **7** had been previously prepared in bulk, donors **45**[92] and **44**[93] were synthesised in one step each from their respective analogues (Scheme 22). Penta-*O*-acetyl-β-D-galactose **5** and *p*-methylphenyl 1-thio-β-D-galactopyranose **7** were prepared as previously stated. A bromide leaving group was substituted onto the anomeric position of **5** by stirring in 33% HBr in acetic acid for 90 minutes, producing **45**[94] in 82% yield. The benzoylation of compound **7** with benzoyl chloride in pyridine gave the white crystalline solid **45** in 70 % yield.



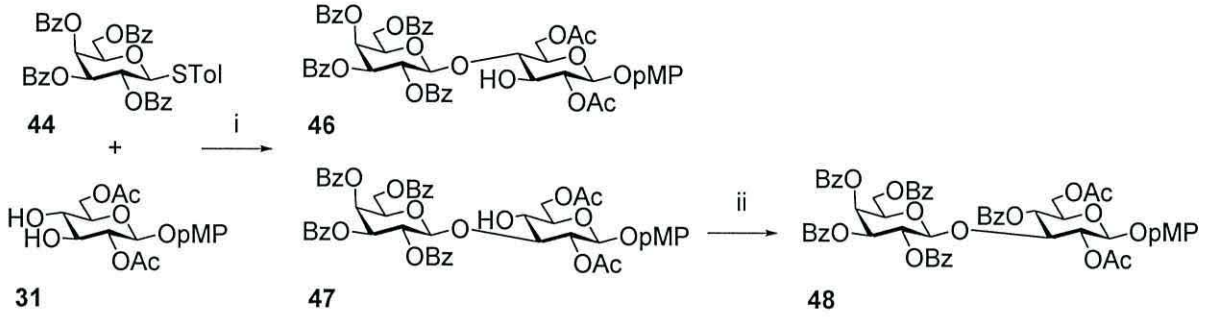
Scheme 22 Synthesis of donors based on deactivation factors by Douglas[87]. Reaction conditions: (i) CH_3COONa, $(CH_3CO)_2O$, 1.5 h, rt, 96%; (ii) 33% HBr in AcOH, rt, 90 min, 82 %; (iii) CH_2Cl_2, $BF_3.Et_2O$, *p*-thiocresol, 3 h, 88 %; (iv) CH_3OH, NaOMe,1.5 h, 98 %; (v) C_5H_5N, BzCl, rt, 16 h, 70 %.

Although the acetates are thought of as deactivating, we wanted to test the option of the bromide leaving group in addition to our other synthesised donors. Bromides are one of the most commonly used amongst donors due to their high reactivity, although they are also one of the least stable[18]. The stability of the leaving group of donor **44**, plus deactivation factors of benzoyl protecting groups being comparably lower made this the more suitable donor.

2.2.8 Successful disaccharide synthesis

Test couplings were carried out on a milligram scale with the newly synthesised donors. Compound **44** was observed on TLC to successfully couple with acceptors **15**, **31**, **32** and **33**. Even with the donor in excess, the donor did not couple to the acceptor in multiple positions, and only disaccharide formation was observed. When an acceptor is less reactive than its donor counterpart, selective glycosidation is observed, with specific disaccharides synthesised[87]. It may be possible to form a trisaccharide with a more reactive acceptor.

With the success of these tests, the synthesis was scaled up. Acceptor **31** (224 mg, 0.6 mmol) and 1.5 equivalents of donor **44** (650 mg, 0.9 mmol) were coupled together under NIS/AgOTf glycosylation conditions and immediately purified (Scheme 23).



Scheme 23. Successful synthesis of disaccharides, with further investigation into linkage by benzoylation of product. Reaction conditions: (i) NIS/AgOTf, CH_2Cl_2, rt, 3 h, 37 %; (ii) C_5H_5N, BzCl, rt, 16 h, crude.

A mixture of compounds was visible on both TLC and NMR. Integration of ^1H spectra revealed the expected single peaks for the acceptor, along with multiple sets signals correlated to the donor. This suggested that coupling produced both regioisomers of the same disaccharide. The primary product occurred in a 3:1 ratio to the secondary, as suggested by integration of ^1H signals.

The simplest way to distinguish between the two products was to protect the remaining hydroxyl group with a benzoyl ester. Benzoyl esters were chosen as they would give a clear downfield shift in any NMR spectra, away from the compacted signals up field. The shifted signal would indicate which position on the acceptor had been protected via comparison with the original spectra.

When comparing the HSQC spectra of the two disaccharides, we observed a downfield shift of a signal correlating to the 4-position. This signal was shifted downfield from an original 4.44 ppm for **47**, to 4.55 ppm for the newly benzoylated compound **48** (Figure 32). One of the anomeric signals was also shifted, signalling an increase in the electron density of the region. This complements the notion that the glycosidic linkage is next to the newly benzoylated position at the 4 position.

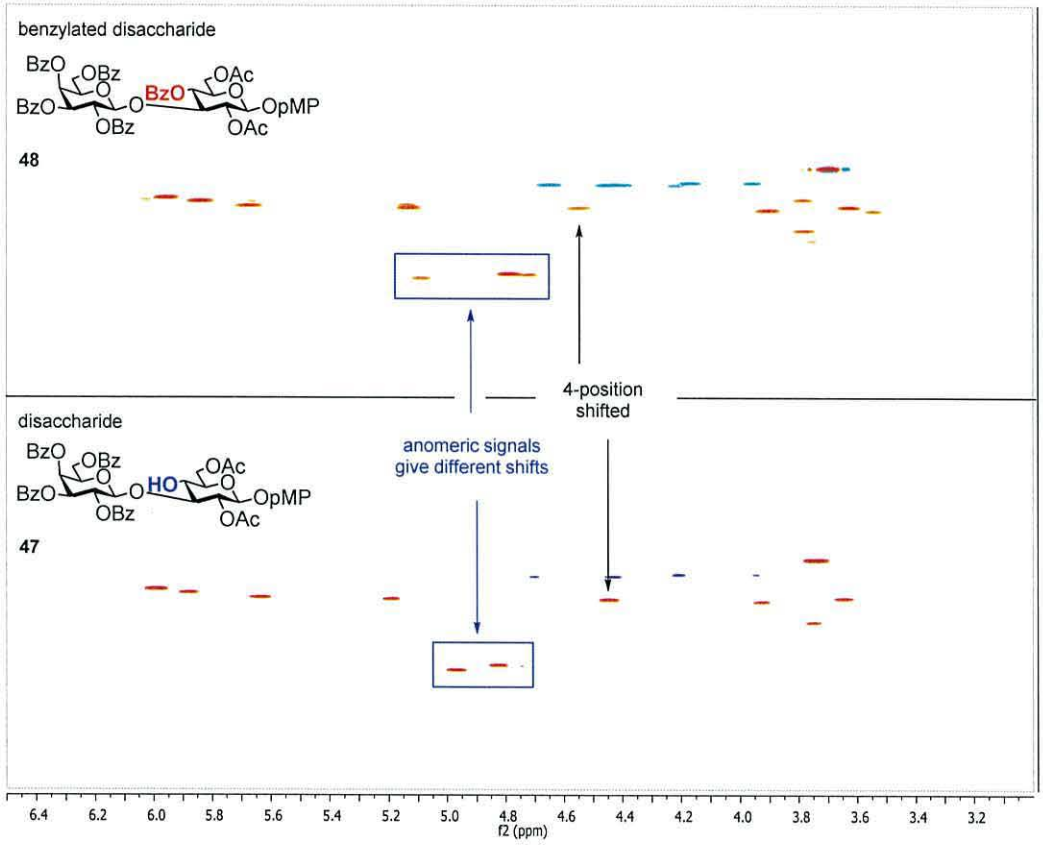
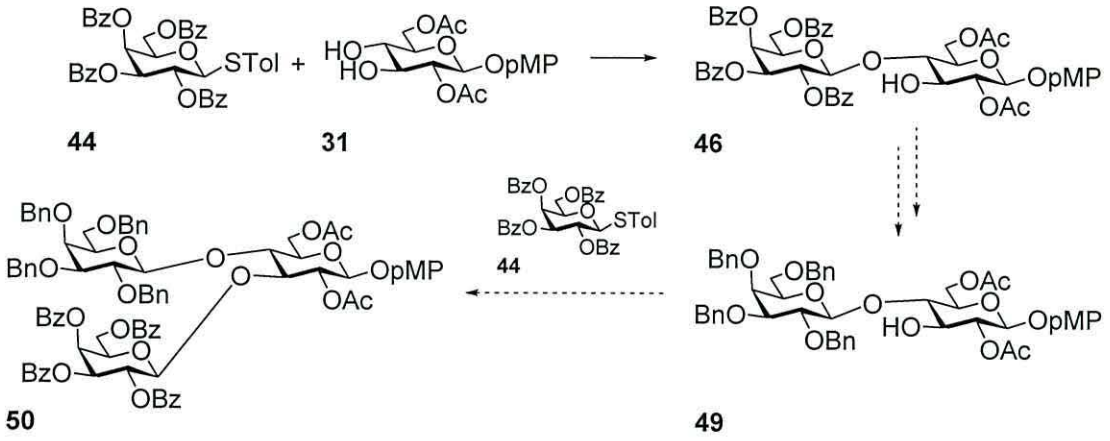


Figure 32. HSQC showing comparison of compounds **48** and **47**.

Once disaccharide **47** was purified, we experienced several unsuccessful coupling attempts with donor **44**. The formation of a trisaccharide could not be achieved, either due to reactivity or steric hindrance of the disaccharide acceptor. If the problem is reactivity, a simple work around for this would be replace the benzoyl protecting groups on the newly formed acceptor with benzyl ethers (Scheme 24). These electron "donating" groups would increase the overall reactivity of the system.

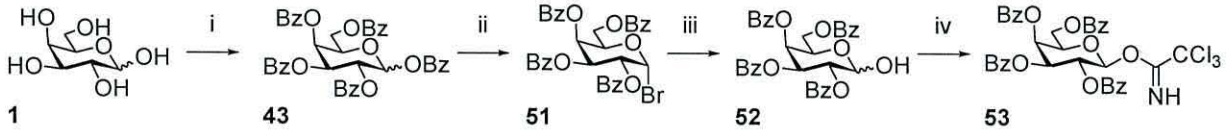


Scheme 24. Protecting the disaccharide with more electron donating groups may allow trisaccharide formation.

With each protection and deprotection, a loss of yield is experienced. Additional protection steps would not be desirable. However, further protection with electron donating groups, such as benzyl ethers, may "match" the reactivity of the systems once again, allowing glycosidic bond formation.

2.2.9 Optimising the conditions

In order to improve yields, the more reactive trichloroacetimidate donors were synthesised[95] (Scheme 25). A simple pathway was revised for the benzoylated donor to add a trichloroacetimidate leaving group.

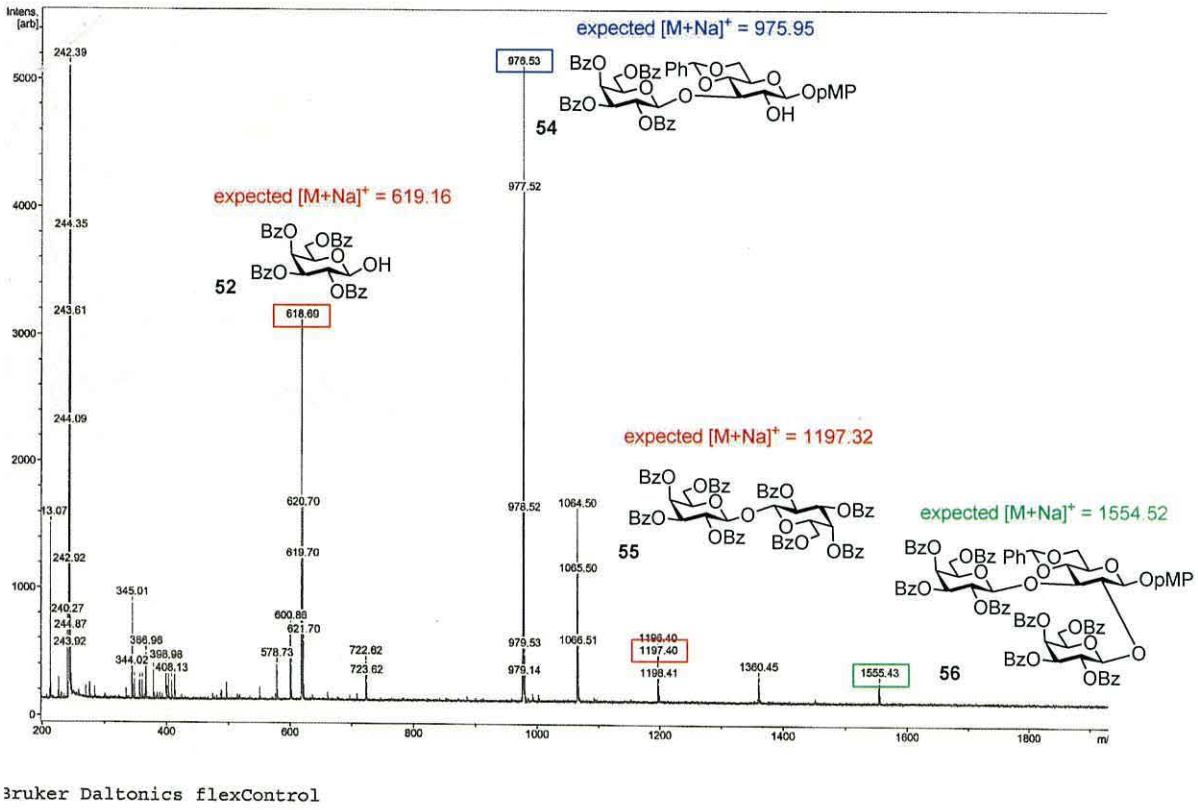


Scheme 25. Reaction conditions: (i) BzCl, C_5H_5N, 1.5 h, rt, 78 %;(ii) 33 % HBr in AcOH, CH_2Cl_2, 1.5 h, rt, 66 %;(iii) Ag_2CO_3, H_2O, acetone, 2 h, 62 %; (iv) CCl_3CN, NaH, CH_2Cl_2, 1 h, rt, 50 %.

Anhydrous D-galactose **1** was esterified by stirring in benzoyl chloride and pyridine for 2 hours. The resulting per-*O*-benzoylated galactose **43** was then stirred in 33% HBr in acetic acid for 90 minutes, yielding the brominated donor **51** in 87 %. As trichloroactimidates require a free hydroxyl position for their synthesis, hydrolysis of the bromine leaving group was carried out using the activator Ag_2CO_3 in the presence of H_2O[96]. Substitution of the trichloroacetimidate leaving group was carried out using trichloroacetonitrile in DMF affording **53** in 50 % yield[90].

2.2.9.1 *Crude Glycosylation Reaction Mixture*

Without quenching the reaction and purifying the mixture, NMR analysis often gave complicated spectra which did give any information into the crude mixture. We began to monitor glycosylation couplings by MALDI-TOF mass spectrometry. A test coupling using the tetra-*O*-benzoylated trichloroacetimidate donor **53** and the diol acceptor **15** was attempted using the promoter $BF_3{\cdot}(Et)_2O$. After stirring for 2 hours, some of the acceptor **15** was still present in the mixture (Figure 33).



Bruker Daltonics flexControl

Figure 33. MALDI-TOF spectrum of the reaction mixture between donor **45** and acceptor **12**.

Evidence of disaccharide **54** formation is supported by the $[M+Na]^+$ peak at 976.53 Da. However, hydrolysed donor **52** can be seen at 619.16 Da, suggesting the presence of H_2O in the reaction.

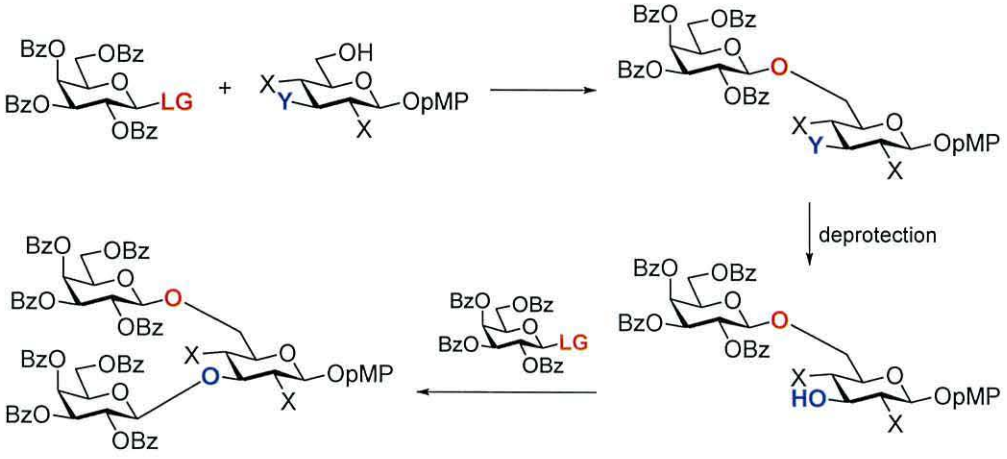
The free hydroxyl group on the hydrolysed donor **52** may have resulted in the molecule acting as an acceptor, leading to the disaccharide **55** being formed.

There was also a nominal amount of a possible trisaccharide **56** formed at 1555.43 m/z. In the test couplings that we carried out, large scale trisaccharide formation was not observed. In order to produce longer branched compounds on a large scale, it

seemed donor and acceptor must be reacted sequentially. Disaccharide formation seems to be very stable, and even with excess donor and promoter, no further large scale coupling was seen.

2.2.10 Alternatives

As trisaccharide formation was not seen on any viable scale during the "random" glycosylations, it would be appropriate to temporarily protect one of the acceptor's two free hydroxyls. It would be advisable to only have one hydroxyl position on the acceptor free at any point in time, with the option to deprotect a second position for attack in a "two-step" glycosylation (Scheme 26).



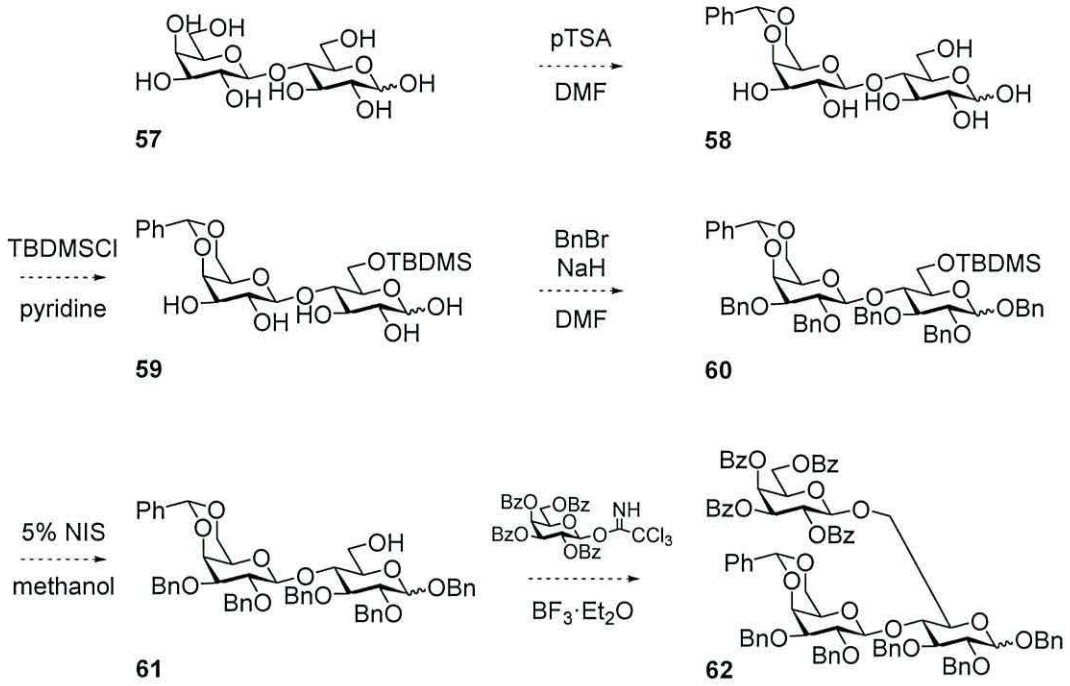
Scheme 26. Example of "two step" oligosaccharide synthesis - deprotection of a "temporary" leaving group to allow a second glycosylation.

2.2.11 Next Steps

Going forward, the large scale synthesis of the benzoylated donor **35** is advised. With each glycosylation excess donor is used, and on a few occasions, the donor had to be remade as the material became depleted. With the repetitive use of the same donor, it is possible to go through over 500 mg with each large scale coupling.

As lactose makes up the reducing end of most GOS structures, it may be suitable to utilise lactose as an acceptor. A quick and efficient synthesis pathway is depicted for the production of a disaccharide acceptor allowing formation of a new β (1→6) linkage (Scheme 27). This pathway was not initially employed as we intended to produce several di- and tri-saccharides from one "set" of acceptors.

At this point, acceptor **55** may only be used for the synthesis of one branched compounds. With more time, various other lactose acceptors may be synthesised, reducing the step wise.



Scheme 27. Efficient synthesis pathway for the production of a lactose acceptor.

With the issues faced and outcomes we have observed in the "random" glycosylations, I believe it is obvious that a more specific step-wise route would be a preferred pathway. We could not produce trisaccharides in any viable yield in one synthetic step. Their production firstly required the synthesis and purification of a disaccharide acceptor. By ensuring one product is formed in this initial step, overall yields will surely increase.

2.3 Conclusions

2.3.1 Linear Synthesis

Building blocks **2** and **4** were successfully synthesised from D-galactose **1** and D-glucose **3** monosaccharide units respectively. By coupling these building blocks, disaccharides **17** and **18** were successfully produced in a "random" glycosylation.

Building block **2** was further coupled via "random glycosylation" to the 3 and 4 positions of **20**. This resulted in a trisaccharide mixture, which could not be separated by means of chromatography. Structural determination of these compounds is purely speculative without further purification; however, we propose the structures **21** and **22** on the information gathered. We suggest the use of semi-preparative HPLC to isolate the two compounds visible on TLC and NMR.

2.3.2 Branched Synthesis

Issues began to arise during attempted coupling of the previously prepared donor **2**, with the "new" series of branched acceptors. We attempted to rectify the problems by reviewing our methods and the conditions of which we were using. We recovered material in all cases of hydrolysed donor **36**, and regenerated the compound by addition of another leaving group, in the form of a trichloroacetimidate of compound **38**.

Various solvents, promoters and temperature ranges were experimented with investigate the circumstances around the unsuccessful nature of our couplings.

We concluded the donor and acceptor combinations were mismatched; we have examined donor and acceptor relationships with regard to reactivity. After exhaustive attempts to "match" donors and acceptors, we clarified that benzoyl esters can, and should be used, to protect a donor for use in further oligosaccharide synthesis against a wide range of orthogonally protected acceptors.

After addressing issues with reactivity of donor/acceptor combinations, donor **44** and acceptor **31** were coupled together to give the major product **47** in 37 % yield. Disaccharide **47** was further benzoylated to yield **48**, confirming the structure of a β

(1→3) linked disaccharide by analysis of HSQC. Attempted trisaccharide synthesis, coupling **48** to **44** could not be produced on a large scale. Confirmation of trisaccharide synthesis can be seen by MALDI-TOF, but in low quantities.

2.3.3 Trisaccharide formation

We have recognised that trisaccharides cannot be synthesised in one "random" glycosylation step using the donor and acceptor combinations listed. Disaccharide formation seems to be very stable, with no larger scale coupling observed in a single step across all test couplings. Trace amounts of larger compounds were found in a reaction mixture under MALDI-TOF analysis.

Although synthesis of a branched trisaccharide was not achieved in high yields on a milligram scale, the groundwork has been set for successful synthesis in the near future. We have tailored both donor and acceptor building synthesis for production on a large scale in the minimal amount of steps possible, resolving any issues with the synthesis along the way. Acceptors have been synthesised in gram scale, which will fulfil all requirements for the production of target molecules. We have also produced and improved upon several types of donor, which should allow rapid completion of the project.

2.4. Experimental

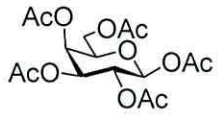
2.4.1. General Methods

Organic solvents were purchased from Fisher Scientific (Loughborough, Leics., UK) as analytical reagent grade. Sugar standards were purchased from Alfa Aesar (Heysham, Lancs., UK). NMR spectra were recorded with a Bruker Ultrashield 400 (Coventry, UK) spectrometer; ^{1}H NMR were run at 400 MHz, whilst ^{13}C was run at 101 MHz. All chemical shifts are quoted on the δ-scale, with $CDCl_3$, MeOD, D_2O, CD_3SOCD_3 and Me_3Si as an internal standard. Deuterated solvents were purchased from Cambridge Isotope Laboratories Inc. (Andover, MA, USA). Mass spectra were recorded on a Bruker Reflex VI using THAP MALDI matrix solution in methanol, and TOF technique. TLC was carried out on $60F_{254}$ glass silica plates (Merck), and visualised using 10 % H_2SO_4 in H_2O. Flash chromatography was performed on 40 - 60 μm silica gel 60 (Fluorochem). Products were concentrated on either Heidolph Laborata 4000 (Norfolk, UK) or Stuart RE300 (Stone, Staffordshire, UK) rotary evaporators. Samples fewer than 100 mg were concentrated under N_2 using a Techne (Stone, Staffordshire, UK) Dri-Block DB-3 and additional sample concentrator. Samples were dried using the hi-vac, operated by a Telestar TD3 double stage vacuum pump, and over P_2O_5 in a vacuum desiccator.

2.4.2 Donor Synthesis

2.4.2.1

1,2,3,4,6-Penta-O-acetyl-D-galactose[72] (5)

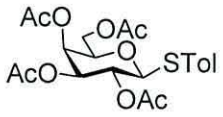


Sodium acetate (1.26 g, 15 mmol, 1 equiv.) was dissolved in acetic anhydride (15 mL, 154 mmol, 10 equiv.) and heated to reflux (140 °C). Anhydrous-D-galactose **1** (2.8 g, 15 mmol) was added portion wise to the mixture, to allow continuous reflux. The mixture was stirred until TLC (toluene/ethyl acetate, 3:2) showed completion (3 h). Upon completion, the reaction mixture is poured onto a slurry of crushed ice and stirred vigorously (1 h). The crude product is allowed to settle and the aqueous phase decanted. Ethyl acetate was added, and the organic phase washed with H_2O, sat. aq. $NaHCO_3$ and dried with $MgSO_4$. The solution was filtered and concentrated. Recrystallisation was carried out with ethanol (5.75 g, 15 mmol, 96 % yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.69 (d, J = 8.3 Hz, H-1β, 1H), 5.41 (d, J = 3.2 Hz, H-4, 1H), 5.32

(dd, *J* = 10.3, 8.4 Hz, H-2, 1H), 5.07 (dd, *J* = 10.4, 3.4 Hz, H-3, 1H), 4.18 (m, H-6a/b, 2H), 4.05 (dd, *J* = 14.8, 7.9 Hz, H-5, 1H), 2.15 (s, CO*CH*$_3$, 3H), 2.11 (s, CO*CH*$_3$, 3H), 2.03 (s, CO*CH*$_3$, 6H), 1.98 (s, COCH$_3$,3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.47, 170.25, 170.09, 169.51, 169.11, 92.29, 71.84, 70.97, 67.96, 66.93, 61.16, 20.93, 20.77, 20.75, 20.66; MS (MALDI-TOF), m/z: calcd for $C_{16}H_{22}O_{11}$: 413.11. $[M+Na]^+$; found: 413.31.

2.4.2.2

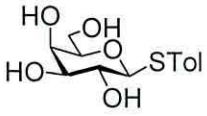
p-Methylphenyl 2,3,4,6-tetra-O-acetyl-1-thio-β-D-galactopyranoside[73] (6)



The per-acetylated sugar **5** (25 g, 64 mmol, 1 equiv.) was dissolved in a minimal amount of dry CH_2Cl_2. *p*-thiocresol (9 g, 72 mmol, 1.1 equiv.) was added and the solution cooled to 0 °C. After cooling, boron trifluoride diethyl etherate (12 mL, 96 mmol, 1.5 equiv.) was added and the solution was stirred (90 mins, rt). The reaction was monitored by TLC (toluene/ethyl acetate 1:1, R_f 0.65). Upon completion, ice water was added and the organic phase was washed with sat. aq. $NaHCO_3$, and dried over $MgSO_4$. Recrystallisation was carried out with petroleum ether/ethyl acetate (30.5 g, 57 mmol, 88 % yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.41 (d, *J* = 8.1 Hz, 2H, $CH_3C_6H_4S$), 7.12 (d, *J* = 8.0 Hz, $CH_3C_6H_4S$, 2H), 5.40 (d, *J* = 3.2 Hz, H-4, 1H), 5.22 (t, *J* = 10.0 Hz, H-2, 1H), 5.03 (dd, *J* = 9.9, 3.3 Hz, H-3, 1H), 4.64 (d, *J* = 10.0 Hz, H-1, 1H), 4.15 (dd, *J* = 11.3, 6.9 Hz, H-6, 1H), 3.91 (t, *J* = 6.6 Hz, H-5, 1H), 2.34 (s, $CH_3C_6H_4S$, 3H), 2.11 (s, CO*CH*$_3$, 3H), 2.10 (s, CO*CH*$_3$, 3H), 2.04 (s, CO*CH*$_3$, 3H), 1.97 (s, CO*CH*$_3$, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.38, 170.21, 170.08, 169.44, 138.47, 133.16, 129.64, 128.62, 86.96, 74.36, 72.04, 67.31, 67.23, 61.58, 21.16, 20.86, 20.67, 20.63, 20.58; MS (MALDI-TOF), m/z: calcd for $C_{21}H_{26}O_9S$: 477.12. $[M+Na]^+$; found: 477.13.

2.4.2.3

p-Methylphenyl 1-thio-β-D-galactopyranoside[74] (7)

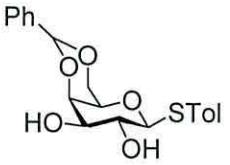


The starting material **6** (30.5 g, 57 mmol, 1 equiv.) was weighed into an oven dried round bottom flask. Dry methanol (25 mL) was syringed into the flask. Sodium (catalytic) was prepared, cut and washed with petroleum ether, and the mixture stirred under N_2. The reaction mixture was stirred to completion (1 h). Amberlite IR-120 H^+ form resin was added to the mixture and stirred gently (10 min). This process

was repeated until a neutral pH was observed. The solution was filtered and concentrated on the rotary evaporator (toluene/ethyl acetate 1:2, R_f 0.05) (16 g, 56 mmol, 98 %). ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.35 (d, J = 8.1 Hz, $CH_3C_6H_4S$, 2H), 7.11 (d, J = 8.0 Hz, $CH_3C_6H_4S$, 2H), 4.48 (d, J = 9.3 Hz, H-1, 1H), 3.70 (d, J = 2.7 Hz, H-5, 1H), 3.54 (m, H-6a/b, 2H), 3.41 (dd, J = 10.2, 4.4 Hz, H-2, 1H), 3.41 – 3.35 (m, H-3, 1H), 3.33 (dd, J = 9.0, 3.0 Hz, H-4, 1H), 2.26 (s, $CH_3C_6H_4S$, 3H); ^{13}C NMR (101 MHz, DMSO) δ 135.79, 131.51, 130.16, 129.43, 88.17, 79.14, 74.72, 69.25, 68.33, 60.54, 20.60; MS (MALDI-TOF), m/z: calcd for $C_{13}H_{18}O_5S$: 309.08. $[M+Na]^+$; found: 309.45.

2.4.2.4

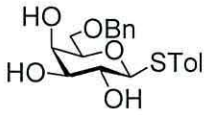
p-Methylphenyl 4,6-O-Benzylidene-1-thio-β-D-galactopyranoside[75] (8)



Compound **7** (12.17 g, 43 mmol, 1 equiv.) was dissolved in benzaldehyde (20 mL). Trifluoroacetic acid (10 mL, 128 mmol, 3 equiv.) was added to the reaction, whilst stirring. After 30 minutes, the reaction was quenched with triethylamine (12 mL, 2 equiv.). The solid which formed was broken apart and stirred for a further 10 minutes. After this time, the precipitate was filtered and washed with ether. Further precipitation, if needed, was carried out by first dissolving in methanol, and then by drop-wise addition of pyridine (toluene/ethyl acetate 1:2, R_f 0.52) (14.1 g, 38 mmol, 88 % yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.57 (m, $CH_3C_6H_4S$, 2H), 7.37 (m, C_6H_5, 5H), 7.11 (d, J = 8.0 Hz, $CH_3C_6H_4S$, 2H), 5.50 (s, C_6H_5CH, 1H), 4.47 (d, J = 9.2 Hz, H-1, 1H), 4.37 (dd, J = 12.4, 1.5 Hz, H-6a, 1H), 4.22 (m, H-4, 1H), 4.03 (dd, J = 12.5, 1.7 Hz, H-6b, 1H), 3.70 (d, H-3, 1H), 3.63 (t, J = 9.2 Hz, H-2, 1H), 3.55 (d, J = 1.1 Hz, H-5, 1H), 2.35 (s, $CH_3C_6H_4S$, 3H). MS (MALDI-TOF), m/z: calcd for $C_{20}H_{22}O_5S$: 397.11. $[M+Na]^+$; found: 397.35.

2.4.2.5

p-Methylphenyl 6-O-benzyl-1-thio-β-D-galactopyranoside[13] (9a)



Method A: Compound **8** (1.4 g, 3 mmol) was mixed with 3Å powdered molecular sieves in THF (20 mL) and stirred for a minimum of 2 hours under N_2. After this time, $NaBH_3CN$ (1.88 g, 18 mmol, 6 equiv.) was added, along with a second amount of sieves. After stirring for a further hour, 1N HCl/Et_2O was syringed drop-wise to the solution until gas production ceased. The reaction was stirred for a further 30 minutes

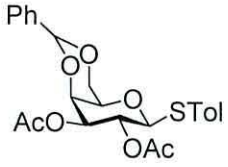
and then quenched with Et_3N. The mixture was diluted with CH_2Cl_2 (50 mL), filtered, washed with brine, dried with $MgSO_4$, and concentrated (toluene/ethyl acetate 3:2, R_f 0.75) (1.38g, 3 mmol, 88%).

Method B: Compound **8** (53 mg, 0.1 mmol) was dissolved in dry CH_2Cl_2 (2 mL) and stirred under dry conditions for 1 hour. The solution was cooled to 0 °C and triethylsilane (165 µL, 1 mmol, 10 equiv.), and $BF_3{\cdot}Et_2O$ (24.5 µL, 0.2 mmol, 2 equiv.) were added. The reaction was left to stir at room temperature until TLC showed completion (toluene/ethyl acetate 3:2, R_f 0.75). The reaction mixture was diluted with CH_2Cl_2, washed with H_2O and $NaHCO_3$, dried over CH_2Cl_2 and concentrated. Purification was carried out via flash chromatography (4:1, toluene/ethyl acetate).

^1H NMR (400 MHz, $CDCl_3$) δ 7.21 (d, J = 8.0 Hz, $CH_3C_6H_4S$, 2H), 7.10 (m, C_6H_5, 5H), 6.83 (d, J = 6.3 Hz, $CH_3C_6H_4S$, 2H), 4.59 (d, J = 11.7 Hz, CH_2Ph, 1H), 4.44 (d, J = 11.7 Hz, CH_2Ph, 1H), 4.30 (m, H-6a/b, 2H), 3.81 (d, J = 2.6 Hz, H-1, 1H), 3.54 (d, J = 5.2 Hz, H-5, 1H), 3.48 (m, H-4, 1H), 3.44 (m, H-2, 1H), 3.37 (d, J = 3.5 Hz, H-3, 1H), 2.07 (s, $CH_3C_6H_4S$, 3H).

2.4.2.6

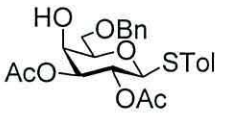
p-Methylphenyl 2,3-Di-O-acetyl-4,6-O-benzylidene-1-thio-β-D-galactopyranoside[13] (10)



Compound **9a** (3 g, 8 mmol) was dissolved in acetic anhydride (4.6 mL, 48 mmol, 6 equiv.) and pyridine (15 mL). The reaction mixture was stirred until TLC showed completion (1.5 h, rt) (toluene/ethyl acetate 1:1, R_f 0.42). Pyridine was removed *in vacuo* by co-evaporation with toluene (3.55 g, 8 mmol, 97%). ^1H NMR (400 MHz, $CDCl_3$) δ 7.53 (d, J = 8.1 Hz, $CH_3C_6H_4S$, 2H), 7.42 (m, C_6H_5, 5H), 7.10 (d, J = 7.9 Hz, $CH_3C_6H_4S$, 2H), 5.48 (s, CHC_6H_5, 1H), 5.31 (t, J = 9.8 Hz, H-2, 1H), 5.01 (dd, J = 9.9, 3.6 Hz, H-3, 1H), 4.67 (d, J = 9.8 Hz, H-1, 1H), 4.41 (d, J = 12.4 Hz, H-6a, 1H), 4.37 (d, J = 2.2 Hz, H-4, 1H) 4.03 (dd, J = 12.4 Hz, 6.8, H-6b, 1H), 3.59 (s, H-5, 1H), 2.36 (s, $CH_3C_6H_4S$, 3H), 2.12 (s, $COCH_3$, 3H), 2.05 (s, $COCH_3$, 3H). MS (MALDI-TOF), m/z: calcd for $C_{24}H_{26}O_7S$: 481.13. $[M+Na]^+$; found: 481.48.

2.4.2.7

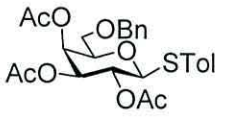
p-Methylphenyl 2,3-Di-O-acetyl-6-O-benzyl-1-thio-β-D-galactopyranoside (11)



Compound **10** (6.36 g, 14 mmol) was mixed with 3Å powdered molecular sieves in THF (10 mL) and stirred for a minimum of 2 hours under N_2. After this time, $NaBH_3CN$ (4.36 g, 69 mmol, 5 equiv.) was added, along with a second amount of sieves. After stirring for a further hour, 1N HCl/Et_2O was syringed drop-wise to the solution until gas production ceased. The reaction was stirred for a further 30 minutes and then quenched with Et_3N. The mixture was diluted with CH_2Cl_2 (50 mL), filtered, washed with brine, dried with $MgSO_4$, and concentrated. Not characterised – used as intermediate.

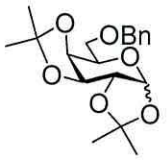
2.4.2.8

p-Methylphenyl 2,3,4-Tri-O-acetyl-6-O-benzyl-1-thio-β-D-galactopyranoside (2)



Compound **11** (1.46 g, 3 mmol) was dissolved in acetic anhydride (1.75 ml, 19 mmol, 6 equiv.) and pyridine (7.5 mL). The reaction mixture was stirred until TLC showed completion (1.5 h, rt) (toluene/ethyl acetate 1:1, 3:1, R_f 0.65). Pyridine was removed *in vacuo* by co-evaporation with toluene. Purification by column chromatography (toluene/ethyl acetate 40:1). (1.46 g, 3 mmol, 94 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40 (dd, *J* = 8.2, 1.8 Hz, $CH_3C_6H_4S$, 2H), 7.36 (m, C_6H_5, 5H), 7.08 (d, *J* = 7.9 Hz, $CH_3C_6H_4S$, 2H), 5.48 (d, *J* = 2.9 Hz, H-4, 1H), 5.21 (t, *J* = 9.9 Hz, H-2, 1H), 5.04 (m, H-3, 1H), 4.67 (d, *J* = 10.0 Hz, H-1, 1H), 4.53 (m, H-2, $CH_2C_6H_5$,1H), 3.86 (t, *J* = 6.3 Hz, H-5, 1H), 3.60 (dd, *J* = 9.7 Hz, 6.2 Hz, H-6a, 1H), 3.49 (dd, *J* = 9.7 Hz, 6.5 Hz, H-6b, 1H), 2.32 (s, $CH_3C_6H_4S$, 3H), 2.09 (s, $COCH_3$, 3H), 2.03 (s, $COCH_3$, 3H), 1.97 (s, $COCH_3$, 3H); 13C NMR (101 MHz, $CDCl_3$) δ 170.29, 170.15, 169.62, 138.31, 137.70, 133.07, 132.85, 129.78, 129.68, 129.18, 128.55, 128.17, 128.02, 127.95, 87.22, 76.11, 74.96, 73.66, 72.33, 67.88, 67.86, 67.67, 21.25, 20.98, 20.75, 20.73. MS (MALDI-TOF), m/z: calcd for $C_{26}H_{30}O_8S$: 525.15. $[M+Na]^+$; found: 525.47.

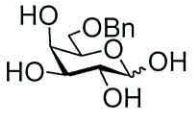
2.4.2.9



6-O-Benzyl-1,2:3,4-di-O-isopropylidene-α-D-galactopyranoside[89] (40)

Dry acetone (200 mL) and copper sulphate (20 g) were stirred in a round bottom flask for 2 hours under N_2 at room temperature. After this time, a catalytic amount of H_2SO_4 (10 mL) was added to the reaction mixture, and the solution cooled to 0 °C. Once cooled, a solution of D-galactose **1** (10.5 g, 58 mmol, 1 equiv.) in acetone (50 mL) was added portion-wise to the reaction mixture. The reaction was stirred at room temperature for 26 hours; TLC (petroleum ether/ethyl acetate 3:7 R_f 0.43). Upon completion, the product was filtered, solvent evaporated and the sugar dried *in vacuo*. Once dried, the intermediate **39** was dissolved in dry DMF with 3A molecular sieves and stirred for 1 hours. The solution was cooled to 0 °C and NaH (0.6 mg, 25 mmol, 3 equiv.) was added portion-wise. The solution was stirred vigorously for a further 30 minutes. After this time, benzyl bromide (1.19 ml, 10 mmol, 1.2 equiv.) was added. After stirring for 18 h, the remaining NaH was quenched by addition of methanol; TLC (dichloromethane/methanol 8:1 R_f 0.28). (4.40 g, 12 mmol, 22 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.35 (m, 5H), 5.55 (d, *J* = 5.0 Hz, 1H), 4.61 (d, *J* = 1.9 Hz, 1H), 4.60 (dd, *J* = 5.6, 2.4 Hz, 1H), 4.57 (d, *J* = 1.5 Hz, 1H), 4.31 (dd, *J* = 5.0, 2.4 Hz, 1H), 4.28 (dd, *J* = 7.9, 1.8 Hz, 1H), 4.04 (m, 1H), 3.67 (ddd, *J* = 26.2, 10.1, 6.3 Hz, 2H), 1.54 (s, 3H), 1.44 (s, 3H), 1.34 (s, 3H), 1.34 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 129.19, 128.46, 128.38, 127.94, 127.87, 127.69, 96.53, 73.47, 72.27, 71.34, 70.81, 70.76, 69.03, 67.04, 26.25, 26.13, 25.10, 24.60; MS (MALDI-TOF), m/z: calcd for $C_{19}H_{26}O_6$: 373.16. $[M+Na]^+$; found: 373.40.

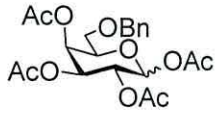
2.4.2.10



6-O-Benzyl-D-galactose[91] (41)

Compound **40** (4.4 g, 12 mmol) was dissolved in 1, 4-dioxane (20 mL) and heated to reflux. 1N HCl (2 mL) was added to the reaction and heated for 4 h until TLC showed completion (toluene/ethyl acetate 1:6, R_f 0.12). The solution was allowed to cool, before solvent was removed *in vacuo* (2.67 g, 6 mmol, 50 %). Not characterised – used as intermediate.

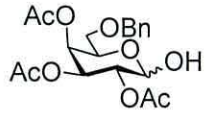
2.4.2.11



1,2,3,4-Tetra-O-acetyl-6-O-benzyl-D-galactopyranose[91] (42)

compound **41** (2.19 g, 8 mmol) was dissolved in acetic anhydride (6.2 ml, 65 mmol, 8 equiv.) and pyridine (15 mL). The reaction mixture was stirred until TLC showed completion (1.5 h, rt) (toluene/ethyl acetate 3:1, R_f 0.62). Pyridine was removed *in vacuo* by co-evaporation with toluene. Not characterised – used as intermediate.

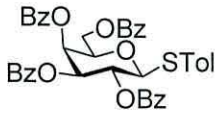
2.4.2.12



2,3,4-Tri-O-acetyl-6-O-benzyl-D-galactopyranose[86] (36)

Compound **33** (2.5 g, 6 mmol) was dissolved in dry CH_2Cl_2 (20 mL). Morpholine (2 mL, 23 mmol, 4 equiv.) was added, and the reaction heated to reflux. Once TLC showed completion (toluene/ethyl acetate 1:1, R_f 0.57), the mixture was cooled to room temperature, washed with 1N HCl and H_2O. The organic phase was dried with MgSO4, concentrated, and dried *in vacuo* (1.62 g, 6 mmol, 72 %).

2.4.2.13

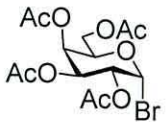


p Metyhlphenyl 2,3,4,6–Tetra-O-benzoyl-1-thio-β-D-galactopyranoside[93] (44)

p-Methylphenyl 1-thio-β-D-galactopyranoside **7** (4 g, 13 mmol) was dissolved in pyridine (40 mL) and cooled to 0 °C. To the already cooled solution, benzoyl chloride (12.3 mL, 106 mmol, 8 equiv.) is added and the reaction stirred at room temperature for 3 hours. Once TLC shows completion (toluene/ethyl acetate 3:1, R_f 0.65), the reaction is quenched with Et_3N. Pyridine is co-evaporated with toluene (x3), and the concentrated gel diluted in CH_2Cl_2. The sugar is applied directly to a column and purified via flash chromatography (8:1, toluene/ethyl acetate) (4.66 g, 6 mmol, 50 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.97 (m, C_6H_5, 5H), 7.65 (m, C_6H_5, 5H), 7.51 (m, C_6H_5, 10H), 7.23 (t, J = 7.8 Hz, $CH_3C_6H_4S$, 2H), 7.08 (d, J = 8.0 Hz, $CH_3C_6H_4S$, 2H), 6.00 (d, J = 3.2 Hz, H-4, 1H), 5.75 (t, J = 9.9 Hz, H-2, 1H), 5.60 (dd, J = 9.9, 3.3 Hz, H-3, 1H), 4.99 (d, J = 9.9 Hz, H-1, 1H), 4.66 (dd, J = 11.3, 6.8 Hz, H-6a, 1H), 4.44 (dd, J = 11.3, 5.8 Hz, H-6b, 1H), 4.38 (t, J = 6.3 Hz, H-5, 1H), 2.37 (s, $CH_3C_6H_4S$, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.19, 165.66, 165.53, 165.31, 138.79, 134.54, 133.68, 133.47, 133.40, 130.17, 129.98, 129.96, 129.91, 129.76, 129.61, 129.49, 129.18, 129.07, 128.88, 128.66, 128.58, 128.41, 128.37, 127.54, 86.27, 75.21, 73.17, 68.50,

68.06, 62.63, 21.48. MS (MALDI-TOF), m/z: calcd for $C_{41}H_{34}O_9S$: 725.18. $[M+Na]^+$; found: 724.78.

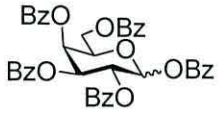
2.4.2.14



2,3,4,6-Tetra-O-acetyl-1-bromo-α-D-galactopyranoside[94] (45)

Per-acetylated galactose **5** (5.77 g, 15 mmol) was dissolved in HBr in acetic acid (18 mL, 33% w/w) in dry CH_2Cl_2 (20 mL). After 1 hour, the solution was reduced by co-evaporation with toluene. The solution was then diluted in ethyl acetate, and poured onto a slurry of crushed ice in sat. aq. $NaHCO_3$. The organic phase was separated and washed again with $NaHCO_3$. The product was dried over CH_2Cl_2, filtered and concentrated (5.66 g, 14 mmol, 93 %). TLC (toluene/ethyl acetate 1:1, R_f 0.65); ^1H NMR (400 MHz, $CDCl_3$) δ 6.69 (d, *J* = 3.9 Hz, H-1, 1H), 5.54 (m, H-4, 1H), 5.40 (dd, *J* = 10.6 3.3 Hz, H-3, 1H), 5.04 (dd, *J* = 10.6, 4.0 Hz, H-2, 1H), 4.48 (t, *J* = 6.6 Hz, H-5, 1H), 4.14 (dd, *J* = 11.4, 6.6 Hz, H-6a/b, 2H), 2.15 (s, CO*CH*$_3$, 3H), 2.11 (s, CO*CH*$_3$, 3H), 2.06 (s, CO*CH*$_3$, 3H), 2.01 (s, CO*CH*$_3$, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.46, 170.20, 170.03, 169.85, 88.25, 71.19, 68.13, 67.91, 67.11, 60.96, 20.87, 20.77, 20.71, 20.68.

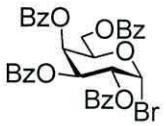
2.4.2.15



1,2,3,4,6-Penta-O-benzoyl-D-galactopyranoside[92] (43)

Anhydrous-D-galactose **1** (10 g, 53 mmol) was dissolved in pyridine (40 mL) and cooled to 0 °C. To the already cooled solution, benzoyl chloride (75 ml, 10 equiv.) is added and the reaction stirred at room temperature for 3 hours. Once TLC shows completion (petroleum ether/ethyl acetate 3:7, R_f 0.81), the reaction is quenched with Et_3N. Pyridine is co-evaporated with toluene (x3). The sugar is applied directly to a column and purified via flash chromatography (8:1, toluene/ethyl acetate) (12.36 g, 48 mmol, 86 %). ^1H NMR (400 MHz, $CDCl_3$) δ 8.17 (m, C_6H_5CO, 2H), 8.16 (m, C_6H_5CO, 2H), 8.14 (m, C_6H_5CO, 2H), 7.98 (m, C_6H_5CO, 2H), 7.84 (m, C_6H_5CO, 3H), 7.71 (m, C_6H_5CO, 5H), 7.53 (m, C_6H_5CO, 3H), 7.47 (m, C_6H_5CO, 3H), 7.32 (m, C_6H_5CO, 3H), 6.95 (d, *J* = 3.6 Hz, H-1, 1H), 6.19 (d, *J* = 2.3 Hz, H-4, 1H), 6.12 (d, *J* = 3.3 Hz, H-3, 1H), 6.03 (d, *J* = 3.6 Hz, H-2, 1H), 4.83 (t, *J* = 6.7 Hz, H-5, 1H), 4.63 (dd, *J* = 11.3, 6.8 Hz, H-6a, 1H), 4.42 (dd, *J* = 11.3, 5.8 Hz, H-6b, 1H); MS (MALDI-TOF), m/z: calcd for $C_{41}H_{32}O_{11}$: 725.18. $[M+Na]^+$; found: 725.19.

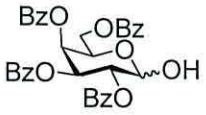
2.4.2.16



1-Bromo-2,3,4,6-tetra-O-benzoyl-D-galactopyranoside[97] (51)

Compound **43** (2.5 g, 4 mmol) was dissolved in HBr in acetic acid (1.6 ml, 20 mmol) in dry CH_2Cl_2 (20 mL). After 1 hour, the solution was reduced by co-evaporation with toluene. The solution was then diluted in ethyl acetate, and poured onto a slurry of crushed ice in sat. aq. $NaHCO_3$. The organic phase was separated and washed again with $NaHCO_3$. The product was dried over CH_2Cl_2, filtered and concentrated; TLC (toluene/ethyl acetate 3:1, R_f 0.74) (3.15 g, 3 mmol, 87 %). ^1H NMR (400 MHz, $CDCl_3$) δ 8.14 (m, C_6H_5CO, 2H), 8.12 (m, C_6H_5CO, 2H), 8.08 (m, C_6H_5CO, 3H), 7.87 (m, C_6H_5CO, 2H), 7.65 (m, C_6H_5CO, 3H), 7.58 (m, C_6H_5CO, 5H), 7.31 (m, C_6H_5CO, 3H), 7.01 (d, *J* = 3.9 Hz, H-1, 1H), 6.16 (d, *J* = 2.8 Hz, H-4, 1H), 6.11 (m, H-3, 1H), 5.71 (dd, *J* = 10.4, 4.0 Hz, H-2, 1H), 4.96 (t, *J* = 6.4 Hz, H-5, 1H), 4.69 (m, H-6a, 1H), 4.54 (m, H-6b, 1H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.58, 165.95, 165.58, 165.34, 133.80, 133.76, 133.69, 133.55, 133.40, 133.35, 130.21, 130.03, 129.96, 129.83, 129.77, 129.59, 129.33, 128.81, 128.75, 128.68, 128.58, 128.52, 128.49, 128.47, 128.35, 92.39, 88.29, 72.42, 71.85, 71.70, 68.91, 68.77, 68.62, 68.10, 67.92, 61.87.

2.4.2.17



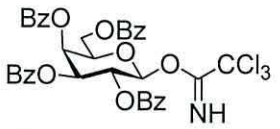
2,3,4,6–Tetra-O-benzoyl-D-galactopyranoside[98] (52)

1-Bromo-2, 3, 4, 6-tetra-*O*-benzoyl-D-galactopyranoside **51** (5.27 g, 8 mmol) was dissolved in a mixture of acetone/H_2O (20 mL, 5:1). Ag_2CO_3 (1.10 g, 4 mmol, 0.5 equiv.) was added to the solution and allowed to stir at room temperature to completion (toluene/ethyl acetate 6:1, R_f 0.80). The mixture was filtrated through a pad of celite and directly purified by flash chromatography (toluene/ethyl acetate 15:1) (3.85 g, 7 mmol, 81 %). ^1H NMR (400 MHz, $CDCl_3$) δ 8.10 (d, *J* = 7.2 Hz, C_6H_5CO, 3H), 8.05 (m, C_6H_5CO, 3H), 7.82 (m, C_6H_5CO, 2H), 7.62 (m, C_6H_5CO, 2H), 7.56 (m, C_6H_5CO, 3H), 7.45 (m, C_6H_5CO, 3H), 7.29 (m, C_6H_5CO, 2H), 7.19 (m, C_6H_5CO, 2H), 6.09 (m, H-3β, 1H), 6.07 (d, *J* = 3.5 Hz, H-4β, 1H), 6.01 (d, *J* = 2.7 Hz, H-4α, 1H), 5.86 (d, *J* = 3.5 Hz, H-1β, 1H), 5.73 (m, H-2, 1H), 5.63 (dd, *J* = 10.4, 7.9 Hz, H-2α, 1H), 5.06 (d, *J* = 7.9 Hz, H-1α, 1H), 4.88 (t, *J* = 6.5 Hz, H-5, 1H), 4.68 (dt, *J* = 11.1, 5.6 Hz, H-6bα 5H), 4.62 (dd, *J* = 11.3, 6.5 Hz, H-6a, 1H), 4.45 (dd, *J* = 11.2, 6.5 Hz, H-6aα 1H), 4.39 (dt, *J* = 11.4, 5.8 Hz, H-6b, 1H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.27, 166.26, 166.21, 165.74, 133.82, 133.69, 133.58, 133.37, 133.31, 130.13, 130.09, 129.99, 129.95, 129.86, 128.79, 128.61, 128.56, 128.47, 128.40, 128.37,

125.44, 96.52, 91.26, 72.57, 71.80, 71.09, 69.64, 69.40, 68.32, 68.11, 67.05, 62.54, 62.22, 21.59.

2.4.2.18

2,3,4,6–Tetra-O-benzoyl-D-galactopyranoside trichloroacetimidate[88] (53)

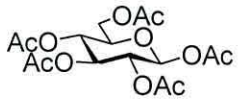


2,3,4,6–Tetra-*O*-benzoyl-D-galactopyranoside **52** (0.54 g, 0.9 mmol) was dissolved in DMF. Trichloroacetonitrile (905 ml, 9 mmol, 10 equiv.) was added, along with a catalytic amount of NaH (5 mg, 90 µmol, 0.1 equiv.). The reaction was stirred for 2 hours under N_2, until TLC showed completion (toluene/ethyl acetate 6:1, R_f 0.57). The reaction mixture was diluted with ethyl acetate and washed with H_2O and brine (450 mg, 0.6 mmol, 67 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.11 (m, C_6H_5CO, 3H), 8.00 (m, C_6H_5CO, 3H), 7.80 (m, C_6H_5CO, 2H), 7.64 (m, C_6H_5CO, 2H), 7.51 (m, C_6H_5CO, 5H), 7.44 (m, 5H), 6.92 (d, *J* = 3.6 Hz, H-1, 1H), 6.16 (d, *J* = 2.5 Hz, H-4, 1H), 6.06 (d, *J* = 3.4 Hz, H-3, 1H), 5.96 (dd, *J* = 10.7, 3.6 Hz, H-2, 1H), 4.87 (t, *J* = 6.8 Hz, H-5, 1H), 4.62 (m, H-6a, 1H), 4.46 (m, H-6b, 1H); MS (MALDI-TOF), m/z: calcd for $C_{36}H_{28}CCl_3NO_{10}$: 736.96. $[M+Na]^+$; found: 736.63.

2.4.3 Acceptor

2.4.3.1

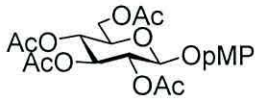
Penta-O-acetyl-D-glucose[72] (12)



Sodium acetate (14.5 g, 175 mmol, 1 equiv.) was dissolved in acetic anhydride (180 ml, 2 mol, 10 equiv.) and heated to reflux (140 °C). Anhydrous-D-glucose **3** (31.5 g, 175 mmol) was added portion wise to the mixture, to allow continuous reflux. The mixture was stirred until TLC (toluene/ethyl acetate 3:2, R_f 0.30) showed completion (3 h). Upon completion, the reaction mixture is poured onto a slurry of crushed ice and stirred vigorously (1 h). The crude product is allowed to settle and the aqueous phase decanted. Ethyl acetate was added, and the organic phase washed with H_2O, sat. aq. $NaHCO_3$ and dried with $MgSO_4$. The solution was filtered and concentrated. Recrystallisation was carried out with ethanol (65.45 g, 167 mmol, 98 %) ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.32 (d, J = 3.7 Hz, H-1α, 4H), 5.71 (d, J = 8.3 Hz, H-1β, 16H), 5.46 (t, J = 9.9 Hz, H-4α, 1H), 5.24 (t, J = 9.4 Hz, H-4β,1H), 5.17 (m, H-2,1H), 4.31 – 4.25 (m, H-3, 1H), 4.10 (dd, J = 12.5, 2.2 Hz, H-6a/b, 2H), 3.82 (dd, J = 8.4, 6.1 Hz, H-5, 1H), 2.11 (s, $COCH_3$, 3H), 2.08 (s, $COCH_3$, 3H), 2.02 (s, $COCH_3$, 3H), 2.02 (s,$COCH_3$, 3H), 2.00 (s, $COCH_3$, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.47, 170.25, 170.09, 169.51, 169.11, 92.29, 77.48, 77.16, 76.84, 71.84, 70.97, 67.96, 66.93, 61.16, 20.93, 20.77, 20.75,. MS (MALDI-TOF), m/z: calcd for $C_{16}H_{22}O_{11}$: 413.11. $[M+Na]^+$; found: 413.02.

2.4.3.2

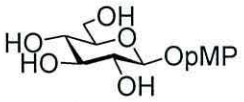
2,3,4,6-Tetra-O-acetyl-1-methoxyphenyl-β-D-glucopyranoside[79] (13)



Per-acetylated glucose **12** (10 g, 26 mmol) and 4-methoxyphenol (6 g, 50 mmol, 1.8 equiv.) were dissolved in dry CH_2Cl_2 (150 mL), with 4Å powdered sieves. After stirring for 45 minutes, TMSOTf (0.5 mL, 3 mmol, 0.1 equiv.) was added, and the solution stirred at 0 °C. Upon completion (toluene/ethyl acetate 1:1) (3 h), the solution was filtered under vacuum. CH_2Cl_2 was added, and the organic phase washed with H_2O, sat. aq. $NaHCO_3$, H_2O, dried with $MgSO_4$ and concentrated. Recrystallisation was carried out using isopropanol to give a white solid (toluene/ethyl acetate 3:1, R_f 0.32) (30 g, 68 mmol, 94 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.41 (d, J = 8.1 Hz, $CH_3OC_6H_4O$-, 2H), 7.14 (m, $CH_3OC_6H_4O$-, 2H), 5.40 (d, J = 3.0 Hz, H-3, 1H), 5.21 (t,

J = 10.0 Hz, H-2, 1H), 5.03 (dd, J = 3.3 Hz, H-4, 1H), 4.64 (d, J = 10.0 Hz, H-1, 1H), 4.21 – 4.09 (m, H-6a/b, 2H), 3.91 (t, J = 6.6 Hz, H-5, 1H), 2.34 (s, $CH_3C_6H_4O_3$, 3H), 2.11 (s, $COCH_3$, 3H), 2.10 (s, $COCH_3$, 3H), 2.04 (s, $COCH_3$, 3H), 1.97 (s, $COCH_3$, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.60, 170.28, 169.40, 155.82, 150.92, 129.04, 128.23, 125.30, 118.74, 114.57, 100.34, 72.77, 71.96, 71.25, 68.33, 61.95, 55.66, 20.71, 20.66, 20.62, 20.60. MS (MALDI-TOF), m/z: calcd for $C_{21}H_{26}O_{11}$: 477.14. $[M+Na]^+$; found: 476.98.

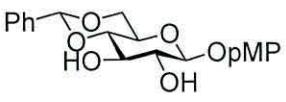
2.4.3.3



1-Methoxyphenyl-β-D-glucopyranoside[81] (14)

Compound **13** (31.12 g, 69 mmol) was weighed into an oven dried round bottom flask. Dry methanol (25 mL) was syringed into the flask. Sodium (catalytic) was prepared, cut and washed with petroleum ether, and the mixture stirred under N_2. The reaction mixture was stirred to completion (1 h). Amberlite IR-120 H^+ form resin was added to the mixture and stirred gently (10 min). This process was repeated until a neutral pH was observed. The solution was filtered and concentrated on the rotary evaporator. Purification of **14** was carried out by hot filtration in ethanol. Recrystallisation in ethanol gave a white powder (19.23 g, 67 mmol, 98 %). ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.01 (m, $CH_3OC_6H_4O$-, 2H), 6.87 (m, $CH_3OC_6H_4O$-, 2H), 4.70 (d, J = 7.4 Hz, H-1, 1H), 3.70 (s, $CH_3OC_6H_4O$-, 3H), 3.68 (m, H-6a, 1H), 3.46 (dd, J = 11.7, 5.8 Hz, H-6b, 1H), 3.29 (m, H-5, 1H), 3.24 (d, J = 8.4 Hz, H-2, 1H), 3.22 (m, H-4, 1H), 3.17 (m, H-3, 1H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 154.29, 151.52, 117.61, 115.73, 114.41, 101.55, 77.03, 76.65, 73.32, 69.81, 60.80, 55.37, 55.34, MS (MALDI-TOF), m/z: calcd for $C_{13}H_{18}O_7$: 309.10. $[M+Na]^+$; found: 309.12.

2.4.3.4



4,6-O-Benzylidene-1-methoxyphenyl-β-D-glucopyranoside[82] (15)

Compound **14** (7.15 g, 25 mmol) was dissolved in benzaldehyde (20 mL). Trifluoroacetic acid (3.8 ml, 50 mmol, 2 equiv.) was added to the reaction, whilst stirring. After 30 minutes, the reaction was quenched with triethylamine (5 mL). The solid which formed was broken apart and stirred for a further 10 minutes. After this time, the precipitate was filtered and washed with ether. Further precipitation, if needed, was carried out by first dissolving in methanol, and then by drop-wise

addition of pyridine (toluene/ethyl acetate 1:2, R_f 0.61) (6.31g, 17 mmol, 67%). ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 7.92 (m, $CH_3OC_6H_4O$-, 2H), 7.76 (m, C_6H_5, 5H), 7.62 (m, *J* = 7.6 Hz, $CH_3OC_6H_4O$-, 2H), 6.76 (m, C_6H_5CH, 1H), 4.91 (d, *J* = 3.6 Hz, H-1, 1H), 3.67 (dd, *J* = 11.2, 2.2 Hz, H-5, 1H), 3.60 (dd, *J* = 11.2, 2.1 Hz, H-4, 1H), 3.47 (m, H-6a/b, 2H), 3.10 (d, *J* = 2.4 Hz, H-2, 1H), 3.03 (d, *J* = 8.1 Hz, H-3, 1H), 1.18 (s, $CH_3OC_6H_4O$-, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.19, 169.60, 155.84, 150.88, 136.73, 129.21, 128.29, 126.17, 118.70, 114.65, 101.60, 100.91, 76.71, 72.26, 71.78, 68.55, 66.52, 55.66; 20.81, 20.71 MS (MALDI-TOF), m/z: calcd for $C_{20}H_{22}O_7$: 397.13. $[M+Na]^+$; found: 397.42.

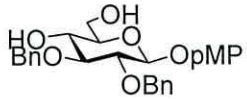
2.4.3.5

2,3-Di-O-Benzyl-4, 6-O-benzylidene-1-methoxyphenyl-β-D-glucopyranoside[99] (16)

Compound **15** (8.19 g, 21 mmol) was dissolved in dry DMF with 3A molecular sieves and stirred for 1 hours. The solution was cooled to 0 °C and NaH (7.5 g, 315 mmol, 15 equiv.) was added portion-wise. The solution was stirred vigorously for a further 30 minutes. After this time, benzyl bromide (12.47 mL, 0.1 mol, 5 equiv.) was added. After stirring for 18 h, the remaining NaH was quenched by addition of methanol. The mixture was diluted with CH_2Cl_2, washed with H_2O, dried with $MgSO_4$, and concentrated (toluene/ethyl acetate 5:1, R_f 0.81) (6.31 g, 17 mmol, 76 %). Not characterised – used as intermediate. MS (MALDI-TOF), m/z: calcd for $C_{34}H_{34}O_7$: 577.22. $[M+Na]^+$; found: 577.50.

2.4.3.6

2,3-Di-O-Benzyl-1-methoxyphenyl-β-D-glucopyranoside[78] (4)



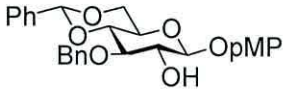
A mixture of **16** (6 g, 11 mmol) and 70% TFA (aq.) (10 mL) was stirred at room temperature for 0.5 h. The solvent was removed under reduced pressure, and the residue was purified by flash chromatography to give the 4,6-diol (toluene/ethyl acetate 3:1, R_f 0.30) (4.1 g, 9 mmol, 78 %); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36 (m, $C_6H_5CH_2$, 5H), 7.31 (m, $C_6H_5CH_2$, 5H), 7.01 (m, $CH_3OC_6H_4O$-, 2H), 6.87 (m, $CH_3OC_6H_4O$-, 2H), 5.06 (d, *J* = 10.9 Hz, $C_6H_5CH_2$, 1H), 5.00 (d, *J* = 11.9 Hz, $C_6H_5CH_2$, 1H), 4.97 (d, *J* = 7.8 Hz, H-1, 1H), 4.82 (d, *J* = 10.9 Hz, $C_6H_5CH_2$, 1H), 4.71 (d, *J* = 11.5 Hz, $C_6H_5CH_2$, 1H), 3.94 (m, H-6a, 1H), 3.79 (s, $CH_3OC_6H_4O$-, 3H), 3.75 (m, H-6b, 1H), 3.70 (m, H-2, 1H), 3.62 (dd, *J* = 5.4. 3.7 Hz, H-4, 1H), 3.53 (t, *J* = 9.0

Hz, H- 3, 1H), 3.46 (m, H- 5, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 155.55, 151.19, 138.34, 137.20, 128.41, 128.36, 128.29, 128.25, 128.10, 127.86, 127.74, 126.03, 118.56, 118.08, 114.73, 114.65, 103.22, 102.64, 101.22, 81.86, 81.28, 80.87, 75.58, 75.22, 68.75, 66.16, 55.69. MS (MALDI-TOF), m/z: calcd for $C_{27}H_{30}O_7$ $[M+Na]^+$: 489.18 ; found: 489.61.

2.4.3.7

3-O-Benzyl-4,6-O-benzylidene-1-methoxyphenyl-β-D-glucopyranoside (25) and 2-O-Benzyl-4, 6-O-benzylidene-1-methoxyphenyl-β-D-glucopyranoside (26) [83]

Compound **15** (1.0 g, 3 mmol), tetrabutyl ammonium hydrogen sulphate (0.18 g, 0.5 mmol, 0.2 equiv.) and benzyl bromide (555 µL, 5 mmol, 1.75 equiv.), were dissolved in CH_2Cl_2. To this, 5% aqueous NaOH was added, and the mixture stirred under reflux for 68 h. The reaction mixture was cooled and the organic phase separated. The organic phase was washed with H_2O, dried with $MgSO_4$, filtered and concentrated



Compound **25** (215 mg, 0.5 mmol, 18 %) was obtained after purification via flash chromatography (10:1 toluene/ethyl acetate: (toluene/ethyl acetate 6:1, R_f 0.32). ^1H NMR (400 MHz, $CDCl_3$) δ 7.53 (m, C_6H_5CH, 2H), 7.42 (m, C_6H_5CH, 3H), 7.38 (m, $C_6H_5CH_2$, 5H) 7.05 (m, $CH_3OC_6H_4O$-, 2H), 6.88 (m, $CH_3OC_6H_4O$-, 2H), 5.55 (s, C_6H_5CH, 1H), 5.05 (d, J = 11.7 Hz, 1H), 5.02 (d, J = 7.9 Hz, 1H), 4.87 (m, 1H), 4.37 (dd, J = 10.5, 5.0 Hz, H-6a, 1H), 3.92 (t, J = 9.1 Hz, H-5, 1H), 3.82 (t, J = 7.2 Hz, H-6b, 1H), 3.80 (m, $CH_3OC_6H_4O$-, 3H), 3.64 (dd, J = 8.6, 2.4 Hz, H-2, 1H), 3.63 (m, H-4, 1H), 3.52 (td, J = 9.7, 4.9 Hz, H-3, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 155.71, 151.34, 138.57, 138.29, 137.38, 129.13, 128.68, 128.52, 128.46, 128.40, 128.34, 128.18, 118.69, 115.99, 114.77, 103.38, 101.34, 82.03, 81.45, 81.04, 75.68, 75.31, 72.26, 70.85, 68.89, 66.30, 55.79; MS (MALDI-TOF), m/z: calcd for $C_{27}H_{28}O_7$: 487.17. $[M+Na]^+$; found: 487.44.

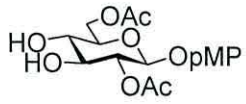
Compound **26** (198 mg, 43 mmol, 16 %) was obtained after purification carried out via flash chromatography (toluene/ethyl acetate 15:1). TLC: (toluene/ethyl acetate 6:1 R_f 0.54). ^1H NMR (400 MHz, $CDCl_3$) δ 7.53 (m, C_6H_5CH, 2H), 7.40 (m, C_6H_5CH,

3H), 7. 33 (m, $C_6H_5CH_2$, 5H) 7.05 (m, $CH_3OC_6H_4O$-, 2H), 6.88 (m, $CH_3OC_6H_4O$-, 2H), 5.55 (s, C_6H_5CH, 1H), 5.05 (d, J = 11.7 Hz, $C_6H_5CH_2$, 1H), 5.02 (d, J = 7.9 Hz, H-1, 1H), 4.87 (d, J = 11.7 Hz, $C_6H_5CH_2$, 1H), 4.37 (dd, J = 10.5, 5.0 Hz, H-6a, 1H), 3.92 (t, J = 9.1 Hz, H-5, 1H), 3.82 (t, J = 7.2 Hz, H-6b, 1H), 3.80 (s, $CH_3OC_6H_4O$-, 3H), 3.64 (dd, J = 8.6, 2.4 Hz, H-4, 1H), 3.63 (m, H-2, 1H), 3.52 (td, J = 9.7, 4.9 Hz, H-3, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 155.76, 151.20, 138.16, 137.07, 129.39, 128.72, 128.62, 128.47, 128.18, 127.88, 127.12, 118.62, 114.80, 114.71, 103.07, 101.99, 101.46, 81.80, 81.31, 80.36, 75.17, 74.84, 74.20, 73.41, 68.82, 66.36, 55.81. ; MS (MALDI-TOF), m/z: calcd for $C_{27}H_{28}O_7$: 487.17. $[M+Na]^+$; found: 487.46.

2.4.3.8

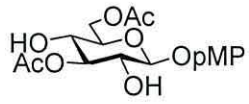
2,6-Di-O-acetyl-1-methoxyphenyl-β-D-glucopyranoside[84] (31), 3,6-Di-O-acetyl-1-methoxyphenyl-β-D-glucopyranoside[84] (32) and 6-O-Acetyl-1-methoxyphenyl-β-D-glucopyranoside[84] (33)

1-methoxyphenyl-β-D-glucopyranoside **14** (885 mg, 4 mmol) and dibutyltin oxide (2.3 g, 9 mmol, 2.2. equiv.) were dissolved in methanol (15 mL) and refluxed for 3 hours, until TLC showed completion (dichloromethane/methanol 1:4) After evaporation of the solvent, the residue was dried under vacuum, and then dissolved in DMF (100 mL). The solution was cooled to 0 °C, and a solution of acetic anhydride (870 µL, 9 mmol, 2.2 equiv.) in the aprotic solvent was added dropwise. Once addition was complete the solution was left to stir at room temperature for 18 hours. After this time, DMF was removed by co-evaporation with toluene. The resulting mixture was diluted in ethyl acetate, washed with brine and the layers separated. The organic phase was dried with $MgSO_4$, concentrated and applied directly to a column.

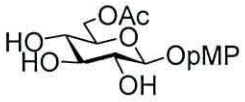


Compound **31;** (toluene/ethyl acetate 1:6, R_f 0.50) (358 mg, 1 mmol, 23 %) was purified by flash chromatography (1:1 – 1:3 petroleum ether/ethyl acetate). ^1H NMR (400 MHz, $CDCl_3$) δ 6.96 – 6.92 (m, $CH_3OC_6H_4O$-, 2H), 6.82 – 6.79 (m, $CH_3OC_6H_4O$-, 2H), 5.02 (dd, J = 9.5, 8.0 Hz, H-2, 1H), 4.87 (d, J = 7.9 Hz, H-1, 1H), 4.47 (dd, J = 12.2, 4.4 Hz, H-6a, 1H), 4.35 (dd, J = 12.1, 1.3 Hz, H-6b, 1H), 3.77 (s, $CH_3OC_6H_4O$-, 3H), 3.69 (dd, J = 5.7 Hz, H-3, 1H), 3.59 (s, H-5, 1H), 3.55 (d, J = 6.5 Hz, H-4, 1H), 2.15 (s, $COCH_3$, 3H), 2.11 (s, $COCH_3$, 3H); δ c (101 MHz, $CDCl_3$) 171.96, 171.10,

155.75, 151.29, 118.84, 118.69, 114.67, 100.44, 75.24, 73.98, 73.77, 70.58, 63.18, 55.80, 21.10, 20.99, 20.89; MS (MALDI-TOF), m/z: calcd for $C_{17}H_{22}O_9$: 393.12. $[M+Na]^+$; found: 393.03.



Compound **32**; (171 mg, 0.5 mmol, 11 %) (toluene/ethyl acetate 1:6, R_f 0.31) was purified via flash chromatography (1:6 – 1:8 petroleum ether/ethyl acetate). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.02 – 6.97 (m, $CH_3OC_6H_4O$-, 2H), 6.84 – 6.80 (m, $CH_3OC_6H_4O$-, 2H), 5.04 – 4.97 (m, H-3, 1H), 4.81 (d, *J* = 7.8 Hz, H-1, 1H), 4.45 (dd, *J* = 12.1, 4.4 Hz, H-6a, 1H), 4.37 – 4.31 (dd, *J* = 12.1, 1.3 Hz, H-6b, 1H), 3.77 (s, $CH_3OC_6H_4O$-, 3H), 3.74 (d, *J* = 7.6 Hz, H-2, 1H), 3.60 (d, *J* = 1.6 Hz, H-5, 1H), 3.59 (d, *J* = 4.8 Hz, H-4, 1H), 2.18 (s, CO*CH*$_3$, 3H), 2.10 (s, CO*CH*$_3$, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.57, 171.68, 155.83, 151.06, 118.82, 114.69, 102.31, 78.44, 78.34, 77.63, 74.38, 72.05, 69.14, 63.18, 55.78, 21.18, 20.99. MS (MALDI-TOF), m/z: calcd for $C_{17}H_{22}O_9$: 393.12. $[M+Na]^+$; found: 393.04.



Compound **33**; TLC (toluene/ethyl acetate 1:6, R_f 0.16) (140 mg, 0.1 mmol, 10 %) was purified via flash chromatography (1:10 – 1:15 petroleum ether/ethyl acetate). ^{1}H NMR (400 MHz, DMSO-d_6) δ 6.97 – 6.91 (m, $CH_3OC_6H_4O$-, 2H), 6.88 – 6.82 (m, $CH_3OC_6H_4O$-, 2H), 4.75 (d, *J* = 7.5 Hz, H-1, 1H), 4.27 (dd, *J* = 11.7, 1.7 Hz, H-6a, 1H), 4.08 (dd, *J* = 11.8, 6.8 Hz, H-6b, 1H), 3.70 (s, $CH_3OC_6H_4O$-, 3H), 3.53 (t, *J* = 7.4 Hz, H-3, 1H), 3.26 (dd, *J* = 8.6, 5.0 Hz, H-2, 1H), 3.23 (dd, *J* = 7.7, 5.2 Hz, H-4 1H), 3.20 – 3.12 (m, H-5, 1H), 2.01 (s, CO*CH*$_3$, 3H). . ^{13}C NMR (101 MHz, DMSO-d_6) δ 170.25, 154.36, 151.23, 117.56, 114.39, 101.22, 76.32, 73.68, 73.55, 73.20, 69.90, 63.45, 55.36; MS (MALDI-TOF), m/z: calcd for $C_{15}H_{20}O_8$: 351.11. $[M+Na]^+$; found: 351.05.

2.4.4 Couplings

2.4.4.1

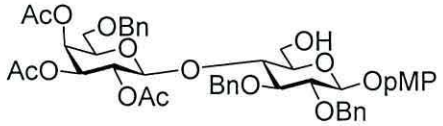
General procedure for glycosylation coupling[96]

The donor (1 equiv.) and acceptor (1 equiv.) were mixed together with molecular sieves (2 g) in the aprotic solvent (30 mL), and the solution stirred for 1 h under N_2. Whilst maintaining the temperature, the promoter (minimum 1.8 equiv. to donor) and catalyst (if viable) were added. The solution was stirred until TLC (toluene/ethyl acetate 3:2, toluene/ethyl acetate 1:6) indicated completion. The reaction was quenched with Et_3N and filtered through Celite. For large scale synthesis, the reaction mixture was washed with $Na_2S_2O_3$, dried with $MgSO_4$ and concentrated. For small scale reactions, the aqueous work-up was omitted. The product was directly purified by flash chromatography (8:1, toluene/ethyl acetate).

2.4.4.2

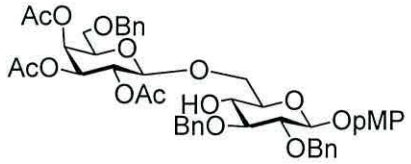
(2,3,4-Tri-O-acetyl-6-O-benzyl-β-D-galactopyranosyl)-β(1→4)-2,3-Di-O-benzyl-1-methoxyphenyl-β-D-glucopyranoside (17) and (2,3,4-Tri-O-acetyl-6-O-benzyl-β-D-galactopyranosyl)-β(1→6)2,3-Di-O-benzyl-1-methoxyphenyl-β-D-glucopyranoside (18)

Donor **2** (500 mg, 1 mmol, 1 equiv.) and acceptor **4** (600 mg, 1 mmol, 1 equiv.) were mixed together with molecular sieves (200 mg) in dry DCM (5 mL), and the solution stirred for 1 h under N_2. Whilst maintaining the temperature, NIS (450 mg, 2 mmol, 2 equiv.) and a catalytic amount of AgOTf were added. The solution was stirred until TLC (toluene/ethyl acetate 3:2) indicated completion. The reaction was quenched with Et_3N and filtered through Celite. The reaction mixture was washed with $Na_2S_2O_3$, dried with $MgSO_4$ and concentrated; purification via flash chromatography (toluene/ethyl acetate 8:1).



Compound **17**; TLC (toluene/ethyl acetate 3:2, R_f 0.44) (180 mg, 5 mmol, 52%). ^1H NMR (400 MHz, $CDCl_3$) δ 7.40 (dd, *J* = 8.1, 1.8 $CH_3OC_6H_4O$-, 2H), 7.36 – 7.27 (m, C_6H_5, 15H), 7.08 (m, $CH_3OC_6H_4O$-, 2H), 5.48 (d, *J* = 2.9 Hz, H-4', 1H), 5.38 (t, *J* = 9.9 Hz, H-2, 1H), 5.21 (t, *J* = 9.9 Hz, H-2', 1H), 5.04 (dd, *J* = 9.9, 3.3 Hz, H-3', 1H), 4.99

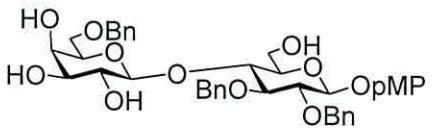
(dd, J = 9.9, 2.9 Hz, H-3, 1H), 4.74 (d, J = 11.6 Hz, CH_2Ph, 2H), 4.67 (d, J = 10.0 Hz, H-1', 1H), 4.61 (d, J = 10.0 Hz, H-1, 1H), 4.54 (d, J = 5.2 Hz, CH_2Ph, 2H), 4.51 (d, J = 5.0 Hz, CH_2Ph , 2H), 4.41 (d, J = 11.8 Hz, CH_2Ph, 1H), 4.31 (d, J = 5.1 Hz, CH_2Ph , 1H), 4.29 (m, H-6a, 1H), 4.09 (m, H-6b, 1H), 4.07 (m, J = 5.1 Hz, CH_2Ph, 1H) 3.93 (d, J = 2.6 Hz, H-4, 1H), 3.87 (t, J = 6.2 Hz, H-5', 1H), 3.72 (t, J = 6.4 Hz, H-5, 1H), 3.59 (dd, J = 9.7, 6.2 Hz, H-6a, 1H), 3.49 (dd, J = 9.7, 6.5 Hz, H-6b, 1H), 2.32 (s, CH_3C$_6$H$_4$O, 3H), 2.03 (s, COCH_3, 3H), 2.01 (s, COCH_3, 3H), 2.00 (s, COCH_3, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.54, 170.49, 170.31, 170.17, 169.64, 169.55, 138.31, 138.22, 137.68, 137.54, 133.05, 132.83, 129.78, 129.68, 129.22, 129.04, 128.55, 128.54, 128.18, 128.03, 127.95, 87.23, 86.86, 76.07, 75.95, 75.13, 74.94, 73.74, 73.65, 72.31, 68.07, 67.85, 67.83, 67.64, 62.64, 21.28, 21.00, 20.95, 20.89, 20.77. MS (MALDI-TOF), m/z: calcd for $C_{46}H_{52}O_{15}$: 867.88. $[M+Na]^+$; found: 868.21



Compound **18**; TLC (toluene/ethyl acetate 3:2, R_f 0.52) (110 mg, 1 mmol, 30%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40 (m, C_6H_5, 15H), 7.05 (m, $CH_3OC_6H_4O$-, 2H), 6.92 (m, $CH_3OC_6H_4O$-, 2H), 5.44 (d, J = 3.1 Hz, H-4', 1H), 5.21 (dd, J = 10.4, 8.0 Hz, H-2', 1H), 5.06 (d, J = 10.9 Hz, CH_2Ph, 1H), 4.99 (m, H-3', 1H), 4.97 (m, CH_2Ph', 1H), 4.90 (d, J = 7.8 Hz, H-1, 1H), 4.82 (d, J = 10.9 Hz, CH_2Ph, 1H), 4.73 (d, J = 11.5 Hz, CH_2Ph', 1H), 4.59 (d, J = 8.0 Hz, H-1', 1H), 4.53 (d, J = 12.0 Hz, CH_2Ph", 1H), 4.41 (d, J = 12.0 Hz, CH_2Ph", 1H), 4.19 (dd, J = 11.1, 2.3 Hz, H-6a, 1H), 3.82 (d, J = 6.6 Hz, H-4, 1H), 3.79 (s, $CH_3OC_6H_4O$-, 3H), 3.76 (m, H-6b, 1H), 3.67 (m, H-2, 1H), 3.58 (dd, J = 5.3 Hz, H-5, 1H), 3.56 (m, H-5', 1H), 3.51 (m, H-6a/b' 2H), 3.49 (m, H-3, 1H), 2.06 (s, COCH_3, 3H), 1.98 (s, COCH_3, 3H), 1.90 (s, COCH_3, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.35, 170.20, 169.67, 155.45, 151.50, 138.50, 138.18, 137.54, 133.66, 130.24, 128.76, 128.59, 128.57, 128.38, 128.12, 128.05, 127.99, 118.05, 114.87, 102.72, 101.14, 83.84, 81.59, 75.46, 75.04, 73.65, 72.20, 71.20, 70.60, 69.07, 68.86, 67.68, 67.51, 55.78, 20.77. MS (MALDI-TOF), m/z: calcd for $C_{46}H_{52}O_{15}$: 867.88. $[M+Na]^+$; found: 868.24

2.4.4.4

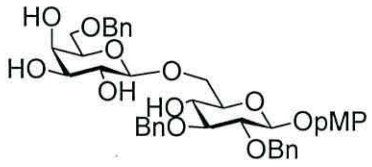
(6-O-benzyl-D-galactopyranosyl)-β(1→4)-2,3-Di-O-benzyl-1-methoxyphenyl-β-D-glucopyranoside (19)



Compound **17** (86 mg, 0.01 mmol) was weighed into an oven dried round bottom flask. Dry methanol (2.5 mL) was syringed into the flask. Sodium (catalytic) was prepared, cut and washed with petroleum ether, and the mixture stirred under N_2. The reaction mixture was stirred to completion (30 min). Amberlite IR-120 H^+ form resin was added to the mixture and stirred gently (10 min). This process was repeated until a neutral pH was observed. The solution was filtered and concentrated on the rotary evaporator. (70 mg, 0.98 mmol, 95 %) (toluene/ethyl acetate 1:6, R_f 0.54). MS (MALDI-TOF), m/z: calcd for $C_{40}H_{46}O_{12}$: 741.29. $[M+Na]^+$; found: 740.87.

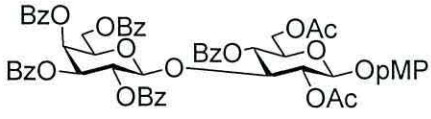
2.4.4.5

(6-O-benzyl -D-galactopyranosyl)-β(1→6)-2,3-Di-O-benzyl-1-methoxyphenyl-β-D-glucopyranoside (20)



Compound **18** (40 mg, 5 µmol) was weighed into an oven dried round bottom flask. Dry methanol (1 mL) was syringed into the flask. Sodium (catalytic) was prepared, cut and washed with petroleum ether, and the mixture stirred under N_2. The reaction mixture was stirred to completion (30 min). Amberlite IR-120 H^+ form resin was added to the mixture and stirred gently (10 min). This process was repeated until a neutral pH was observed. The solution was filtered and concentrated on the rotary evaporator. (30 mg, 5 µmol, 86 %) See 2.4.2.3 for general glycosylation procedure (toluene/ethyl acetate 1:6, R_f 0.63). MS (MALDI-TOF), m/z: calcd for $C_{40}H_{46}O_{12}$: 741.29. $[M+Na]^+$; found: 740.87.

(2,3,4,6-tetra-O-benzoyl-D-galactopyranosyl)-β(1→3)-3-O-benzoyl-2,6-Di-O-acetyl-1-methoxyphenyl-β-D-glucopyranoside (48)



Donor **44** (640 mg, 0.9 mmol, 1.5 equiv.) and acceptor **31** (224 mg, 0.6 mmol, 1 equiv.) were mixed together with molecular sieves (2 g) in DMF (5 mL), and the solution stirred for 1 h under N_2. Whilst maintaining the temperature, NIS (408.6, 1.8 mmol 3 equiv.) and a catalytic amount of AgOTf (5 mg) were added. The solution

was stirred until TLC (toluene/ethyl acetate 3:1) indicated completion. The reaction was quenched with Et_3N and filtered through Celite. For large scale synthesis, the reaction mixture was washed with $Na_2S_2O_3$, dried with $MgSO_4$ and concentrated, yielding disaccharide **47** (221 mg, 0.223 mmol, 38 %). The disaccharide **47** (5 mg, 5.3 mmol) was dissolved in pyridine (1 L) and cooled to 0 °C. To the already cooled solution, benzoyl chloride (20 µl) is added and the reaction stirred at room temperature for 3 hours. Once TLC shows completion (toluene/ethyl acetate 3:1, R_f 0.42), the reaction is quenched with Et_3N. Pyridine is co-evaporated with toluene. The sugar is applied directly to a column and purified via flash chromatography (6:1, toluene/ethyl acetate) (3 mg, 2.8 mmol, 47 %). ^1H NMR (400 MHz, $CDCl_3$) δ 8.09 (dd, *J* = 7.5 Hz, C_6H_5, 3H), 7.94 (m, C_6H_5, 2H), 7.76 (d, *J* = 7.4 Hz, C_6H_5, 2H), 7.67 (d, *J* = 7.7 Hz C_6H_5, 3H), 7.50 (m, C_6H_5, 5H), 7.39 (m, C_6H_5, 5H), 7.25 (m, C_6H_5, 5H), 6.91 (m, $CH_3OC_6H_4O$-, 2H), 6.78 (m, $CH_3OC_6H_4O$-, 2H), 5.99 (d, *J* = 3.5 Hz, H-4, 1H), 5.88 (m, H-2, 1H), 5.62 (dd, *J* = 9.7, 3.4 Hz, H-3, 1H), 5.21 (m, H-2', 1H), 4.97 (d, *J* = 8.0 Hz, H-1, 1H), 4.82 (d, *J* = 8.1 Hz, H-1', 1H), 4.73 (m, H-6a, 1H), 4.45 (d, *J* = 5.2 Hz, H-5, 1H), 4.43 (d, *J* = 8.5 Hz, H-6b, 1H), 4.21 (d, *J* = 10.4 Hz, H-6a', 1H), 3.98 (m, H-3', 1H), 3.92 (d, *J* = 8.8 Hz, H-6b', 1H), 3.76 (d, *J* = 5.8 Hz, H-4', 1H), 3.74 (s, $CH_3OC_6H_4O$-, 3H), 3.64 (dd, *J* = 10.7, 3.7 Hz, H-5', 1H), 2.06 (s, $COCH_3$, 3H), 1.87 (s, $COCH_3$, 3H); MS (MALDI-TOF), m/z: calcd for $C_{59}H_{68}O_{20}$: 1120.16. $[M+Na]^+$; found: 1120.48.

3. Investigation into the isolation of therapeutic carbohydrates from the seeds of the common British bluebell, *Hyacinthoides non-scripta*

3.1 Introduction

3.1.1 Renewable Resources

Glucose is the most abundant carbohydrate on the planet and is a natural, renewable resource. It is produced in plants by photosynthesis, and stored as cellulose. Materials containing cellulose have been examined as natural renewable feed stocks, particularly since the 1970's[100].

Ethanol is produced from plants using cellulose enzymes, followed by fermentation of the glucose to produce ethanol (Figure 34). It forms a central role as a primary chemical feedstock or starting material from which industrial materials are produced. It is used as a starting point for the production of plastics, detergents and pharmaceuticals may be produced industrially, to serve our everyday needs[101].

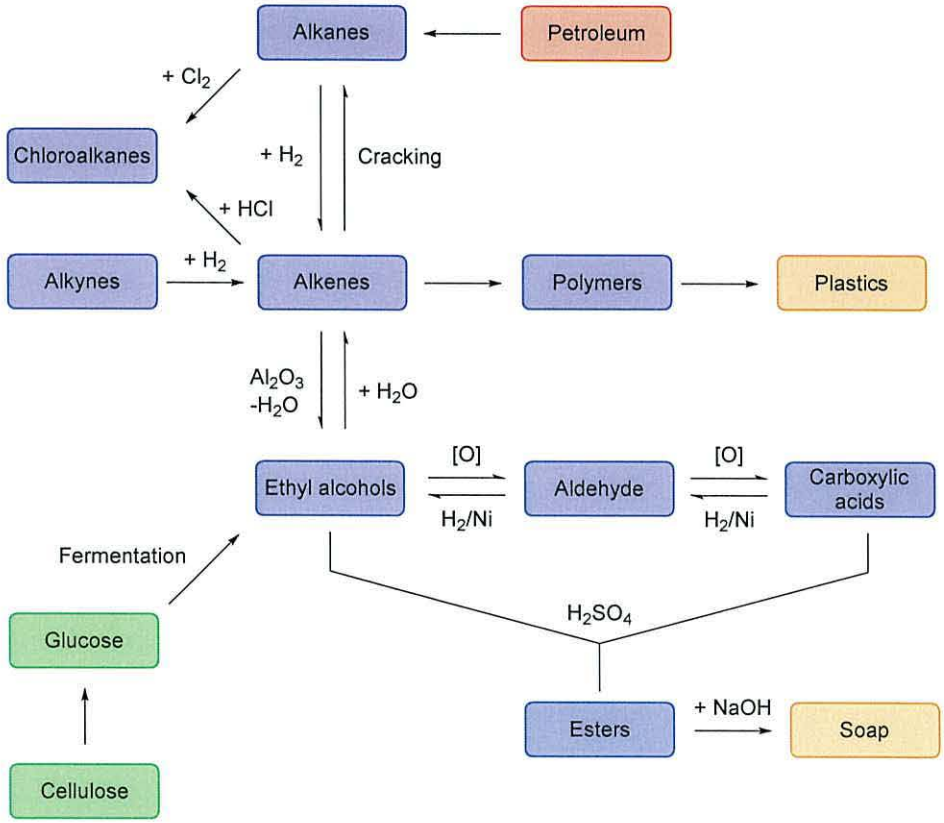


Figure 34. Synthetic routes in organic chemistry using alcohol as a feedstock.

Plants produce many compounds from glucose and have complex carbohydrate content, with many other water soluble components making it difficult to characterise each compound[102]. While it may be possible to extract one major material efficiently from plants, isolation of several chemicals from plant materials in a systematic

manner which is commercially viable is more difficult. The bluebell, *Hyacinthoides non-scripta*, is particularly interesting in that it contains not only many sugars but also alkaloids which are sugar mimics with potent biological activity. The extraction of high value plant based therapeutic chemical substances from the rest of the carbohydrate content of the bluebell could serve as one of the first milestones in research towards a carbohydrate economy[101].

3.1.2 Taxonomy

The common bluebell, *Hyacinthoides non-scripta* (*syn. Endymion non-scriptum*, *Scilla non-scripta*), is a member of the *Hyacinthaceae* monocotyledon family (Table 2). It is abundant in areas of woodland, rough grassland and bracken communities in the UK[103]. It is thought that almost 50 % of the world's population of the *Hyacinthoides non-scripta* can be found in the United Kingdom[104].

Kingdom	**Plantae**
Phylum	Angiosperms
Order	Asparagales
Family	*Hyacinthaceae*
Genus	*Hyacinthoides*
Species	*Hyacinthoides non-scripta*

Table 2. Classification of the common British bluebell

The flower has recently been placed under the protection of the Wildlife and Countryside Act in response to years of over exploitation and the invasion of two new species[104]; the Spanish bluebell, *Hyacinthoides hispanica*, and *Hyacinthoides x massartiana*, hybrid species (Figure 35). These alien species produce highly fertile seeds that threaten the existence of the native *Hyacinthoides non-scripta*.

Figure 35. The *Hyacinthoides non-scripta* (left) has darker flowers growing on one side of a stem, with sharply curved petals. The *Hyacinthoides hispanica (right) has paler flowers produced on all sides of the stem.*

Bluebells propagate by seed germination. Germination occurs when temperatures fall below 10 °C[105]. In the first year of growth one leaf and a bulblet is formed. Bluebells grow quickly, with each year of growth resulting in an increase in biomass above and below ground[106]. Each year, contractile roots pull the bulb deeper into the soil, up to 25 cm in depth[107]. After five years of growth flowers may be formed.

3.1.3 Sugars in the seeds

The seeds of any plant are particularly interesting, as seeds are a high energy source. In order for the new seedling to survive, seeds are packed with components such as carbohydrates, lipids, DNA and enzymes. In addition to these primary metabolites, they also contain protective chemicals such as insecticides, fungicides and other bioactive secondary compounds to protect themselves[100]. Although most of these structures can be found in other parts of the plant, they can be found at particularly high concentrations in the seeds.

Considering the high concentration of natural products in seeds, extraction of compounds from seeds is very appealing industrially, especially during exploratory programmes.

3.1.4 Plants with Biological Potential

Only a very small percentage of plant species have been investigated for any possible biological activity[100]. Plants have adapted to their surroundings and

developed biological activity mechanisms in order to survive. Millions of years of evolution have redirected excess biochemical intermediates to create compounds known as secondary metabolites. These secondary metabolites do not take part in the main life process of the organism, but can give the plant a chance to survive or compete in a specific area or niche of nature[108]. Plants have included secondary metabolites into plant defence mechanisms which although often produced in small quantities in plants, but can nonetheless be potent compounds. Secondary metabolites can be divided into nitrogen containing compounds and non-nitrogen containing compounds[109] (Table 3).

Secondary Metabolites	
Nitrogen containing	Non-Nitrogen containing
Alkaloids	Terpenoids
	Phenolics

Table 3. Secondary metabolite structures in plants, with and without nitrogen.

Terpenoids and phenolics are secondary metabolites which do not contain nitrogen. Terpenoids are lipophilic compounds based on multiple isoprene units which are found in all living organisms and are the largest group of naturally occurring products. Phenolics are a class of chemical compounds consisting of a hydroxyl group bonded directly to an aromatic hydrocarbon group.

Alkaloids have the most numerous pharmacological effects on humans and other animals. Many alkaloids can inhibit or activate biological enzymes. These properties are of great interest to pharmaceutical manufacturers, as many commercial therapeutic drugs affect enzyme function[109]. The biochemical basis of enzyme function is described in more detail in Chapter 4.

3.1.5 Iminosugars

Certain survival mechanisms exhibited by plants are thought to be indicators of therapeutic potential[110]. It is widely noted that bluebells are poisonous. This, in part, explains why the flower does not have a history of applications. The flower was used as a medieval leprosy treatment, but more traditionally as a glue to bind books[101]. It wasn't until 1997 that the toxic principles of the plant were characterised[111]. Nitrogen

analogues of monosaccharides and disaccharides were discovered in cases of livestock poisoning, resulting in the first report of iminosugars in the *Hyacinthaceae* family. These alkaloids had only previously been reported in tropical plants and *Streptomyces* strains of bacteria[112]. Iminosugars are a class of alkaloids, which have also been referred to as polyhydroxylated alkaloids (PHAs).

3.1.6 Iminosugars as Glycosidase inhibitors

The *Hyacinthaceae* family is known to be a rich source of glycosidase inhibitors of structural diversity[113] (Figure 36). Glycosidase inhibitors have aroused interest since the initial isolation of 1-deoxynojirimycin (DNJ) from mulberry leaves, used as a traditional Chinese medicine for the treatment of diabetes[114]. Two iminosugar drugs which are in full production and used as glucoside inhibitors are[115]:

- Migitol (Glyset®- Pfizer)

  Treatment of type II diabetes.
- *N*-Butyl-1-deoxynojirimycin (Zavesca®- Actelion)

  Treatment of Gaucher disease, a common lysosomal storage disease.

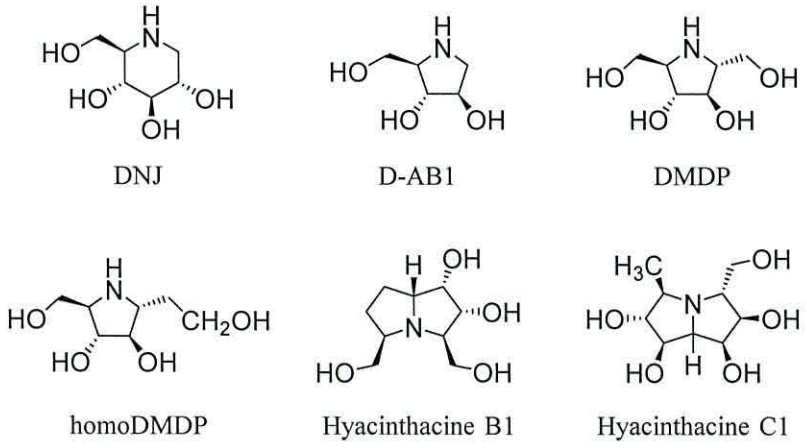


Figure 36. Structures more commonly associated with the British bluebell, *Hyacinthoides non-scripta.*

Potential applications in the clinical practice include therapeutic agents for human lysosomal storage diseases[116], anti-diabetic, anti-viral[117,118], anti-cancer [119] and anti-insect [118,120] agents.

3.1.7 Iminosugars in the Hyacinthaceae family

Over one hundred naturally occurring iminosugars have been discovered over a range of families. A range of iminosugars have been found in the Hyacinthaceae family (Table 4)[121–124].

Iminosugar	Abbreviation	Source
1-Deoxynojirimycin	DNJ	Bulb◊
α-Homonojirimycin	HNJ	Bulb◊
β-Homonojirimycin	HNJ	Bulb◊
1,4-Dideoxy-1,4-imino-D-arabinitol	D-AB1	Bulbs/Leaves*†
2,5-Dihydroxymethyl-3R,4R-dihydroxy-pyrrolidine	DMDP	Bulbs/Leaves*†◊
2,5-Dideoxy-2,5-imino-DL-glycero-D-manno-hepitol	homoDMDP	Fruit/Stalks*†◊
1S,2R,3R,5R,7aR)-1,2-Dihydroxy-3,5-dihydroxymethylpyrrolizidine	Hyacinthacine B1	Fruit/Stalks*†
(1S,2R,3R,5R,6R,7R,7aR)-3-Hydroxymethyl-5-methyl-1,2,6,7-tetrahydroxypyrrolizidine	Hyacinthacine C1	Fruit/Stalks*
6-Deoxy-6-C-(2,5-dihydroxyhexyl)-DMDP		Fruits/Stalks*
Homo-DMDP-7-*O*-apioside		Bulbs/Leaves*
Homo-DMDP-7-*O*-β-D-xylopyranoside		Fruits/Stalks*

**Hyacinthoides non-scripta*

◊*Hyacinthus orientalis*

†*Scilla campanulata*

Table 4. The distribution of iminosugars in the Hyacinthaceae family[122].

3.1.8 Aim

The aim of this work was to investigate the iminosugar content of the seeds of the common British bluebell, *Hyacinthoides non-scripta* by classical water extraction. Ion exchange chromatography will be employed in sequential purification steps[125,126]. Charged iminosugars can be separated by their differing affinities for the resins using column chromatography (Figure 37).

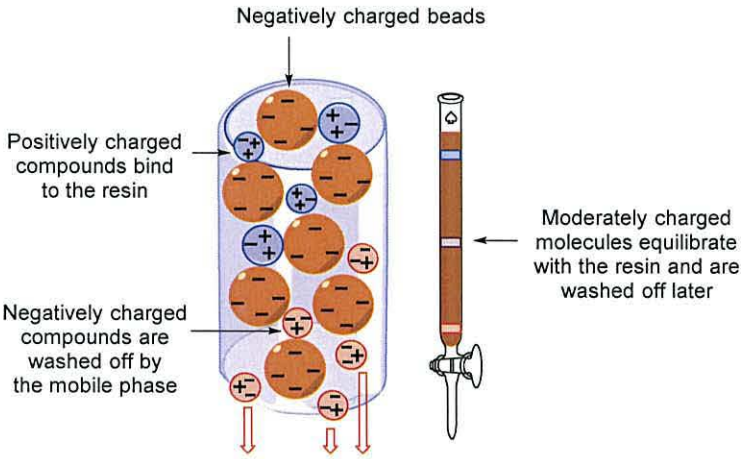


Figure 37. Charged compounds are separated by differing affinities for the ion exchange resins.

Ion exchange can be useful for a range of applications, varying from water purification, protein purification and analysis. The ion exchange resin beads are typically made up of a polystyrenic matrix. Styrene or vinyl benzene is polymerised into long chains (Figure 38). Divinyl benzene (DVB) molecules act as cross linkages between different polymeric chains giving strength and structure to the beads[127]. The amount of DVB in the polymerisation accounts for the pore size of the resin.

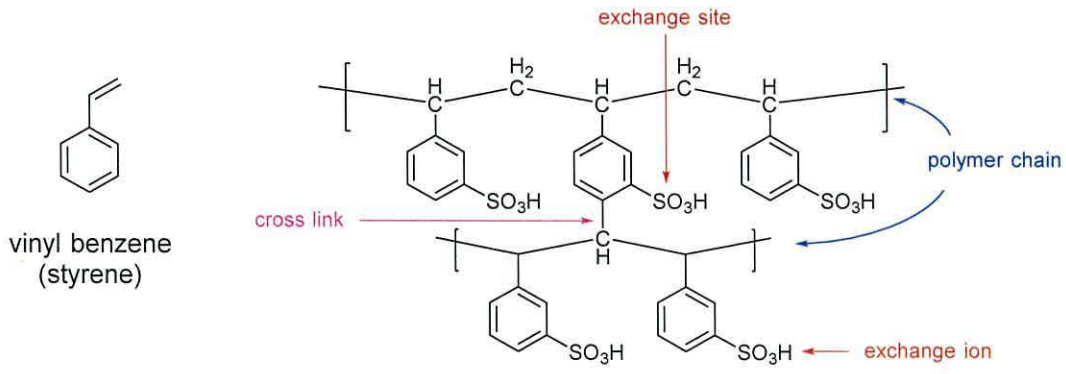


Figure 38. Structure of Amberlite IR-120 cation exchange resin.

3.2 Results and Discussion

3.2.1 Extraction

In previous work, the extraction of approximately 6 grams of a crude iminosugar extract from seeds via a classical water extraction method was described[128].

In general, extraction of PHAs from plant material is carried out with aqueous methanol or ethanol. A great advantage of alcohol over water is the ability to remove the solvent under reduced pressure. However, alcohols can also extract other less polar material, whilst water extracts proteins. Generally, alcohol/water mixtures are used to increase the efficiency of the extraction. The iminosugars of interest we aimed to extract were all water soluble and bluebell seeds contain oils and so only water was used in the extraction method described here.

Once the seeds/large particles were extracted from the extract using both gravity and vacuum filtration, a series of ion exchange purification columns steps were performed.

3.2.2 Ion Exchange Chromatography

3.2.2.1 Unbound Extract

An *unbound* water extract was collected in the first cation exchange (strongly acidic Amberlite IR-120 H+ form) purification column. As iminosugars bind to the cation exchange resin, the material which eluted with water contained amino acids, simple sugars, glycosides, and other material which was not of interest to the study. The *unbound* material was analysed by GC-MS (of trimethylsily-derivatieves [TMS]) to determine whether the initial purification had been successful (Figure 39).

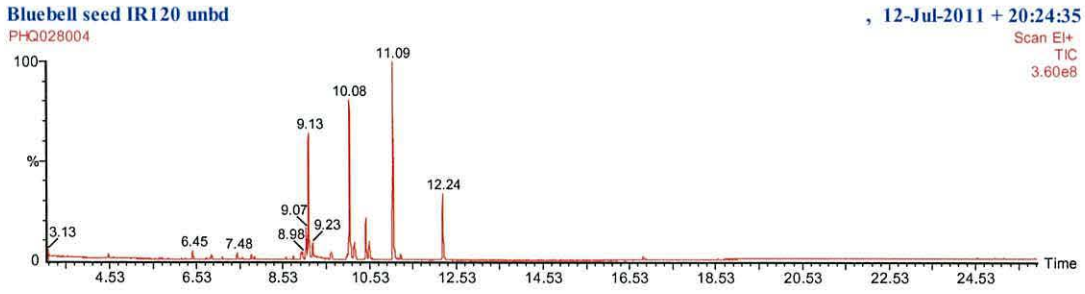


Figure 39. GC-MS (TMS) chromatogram of *unbound* column extract.

The main components were small sugars and sugar alcohols (mainly fructose, glucose and myo-inositol). A small amount of the iminosugar DMDP (8.78 min) was observed in the *unbound* extract (Figure 40). The DMDP could have been lost due to the resin capacity being exceeded, or the DMDP being weakly bound to the resin.

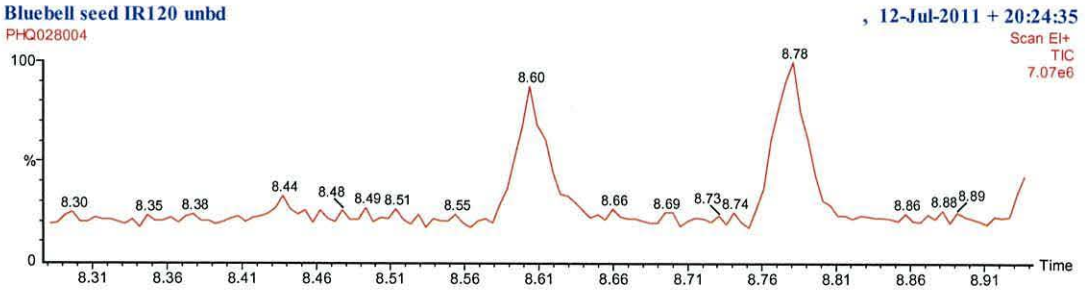


Figure 40. Expansion of chromatogram, revealing a small amount of DMDP within the *unbound* material at 8.78 min.

3.2.2.2 Bound Extract

Once the resin had been washed sufficiently with water to remove compounds which did not bind to the resin, 2M NH_4OH was flushed through the column. By washing with a strong base, *bound* material would be displaced. All of the NH_4OH fractions were concentrated by evaporation, combined with a minimal amount of water, and freeze dried.

The main compound with a retention time of 8.79 min, indicates the presence of DMDP (Figure 41), as determined by coupled mass spectrometry (Figure 42).

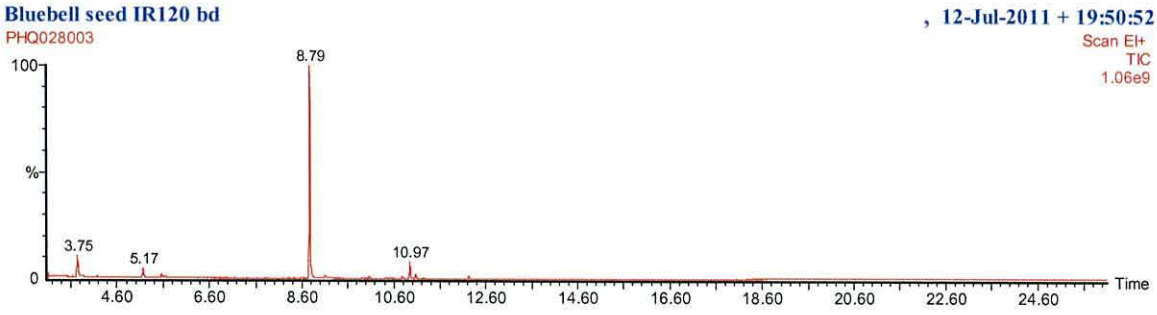


Figure 41. GC-MS Chromatogram containing DMDP and homoDMDP.

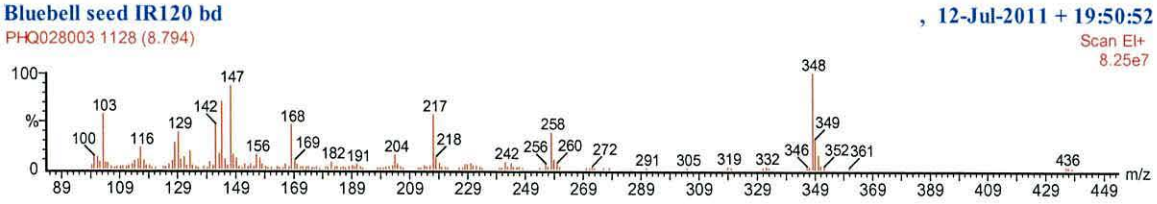


Figure 42. Mass Spectrum of DMDP (TMS derivative).

HomoDMDP was also observed at 10.97 min (Figure 43).

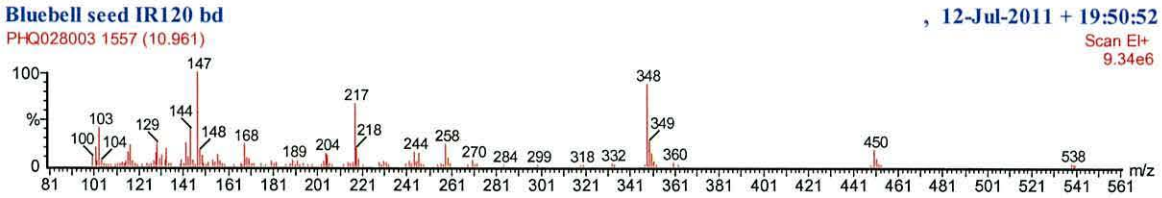


Figure 43. Mass spectrum of homoDMDP (TMS derivative).

Aspartic acid can be seen in the *bound* extract, eluting at 3.75 min (Figure 44).

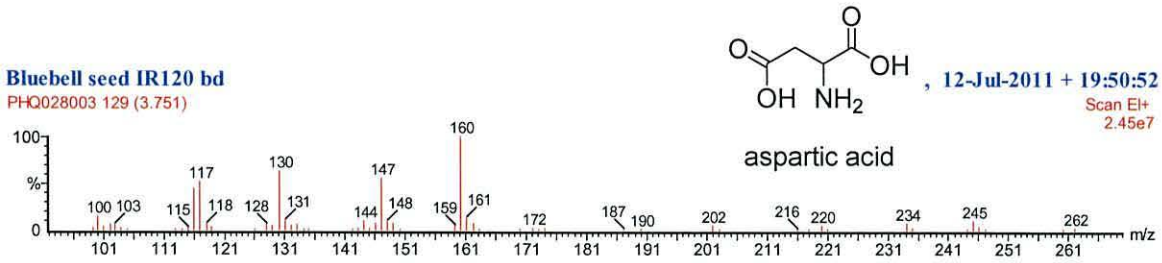


Figure 44. Mass spectrum of aspartic acid (TMS derivative).

3.2.3 Further Purification

The process of extracting target iminosugars from material which was not of interest was completed via a series of ion exchange chromatography techniques (Figure 45).

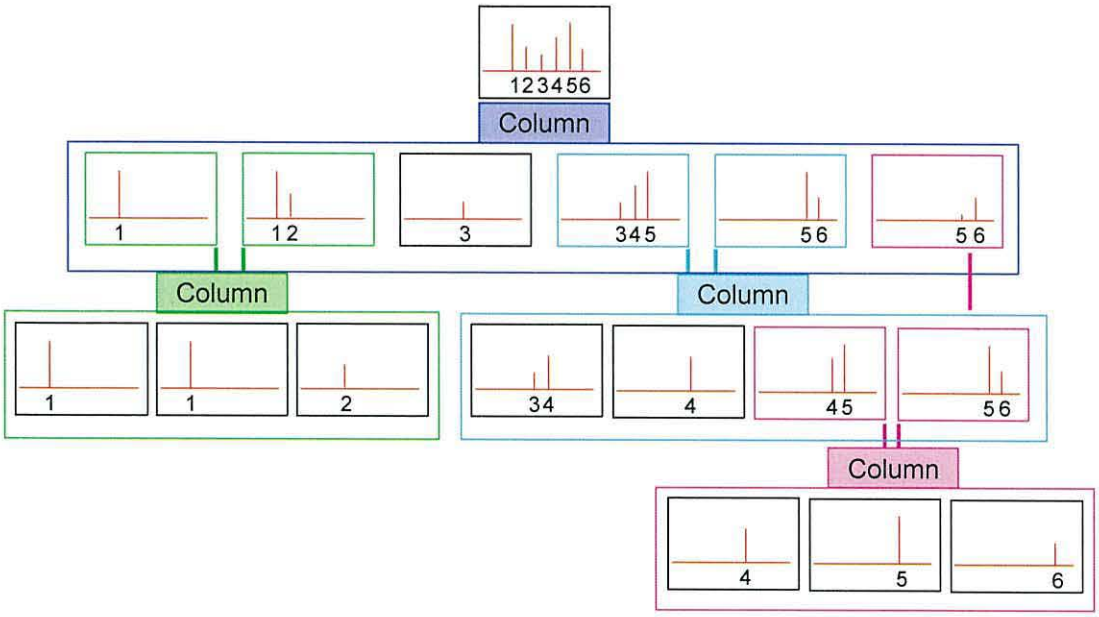


Figure 45. An example of purification by ion exchange chromatography in series.

Aqueous iminosugar solutions were purified into a series of fractions which were analysed by GC-MS. Fractions with similar contents were pooled together and purified. Combining fractions increased productivity by reducing the amount of columns which would be required for purification of all compounds.

We were able to slowly purify the initial crude mixture into several mixtures containing select compounds by repeating this step wise chromatography method (Figure 46). Each fraction was analysed by GC-MS.

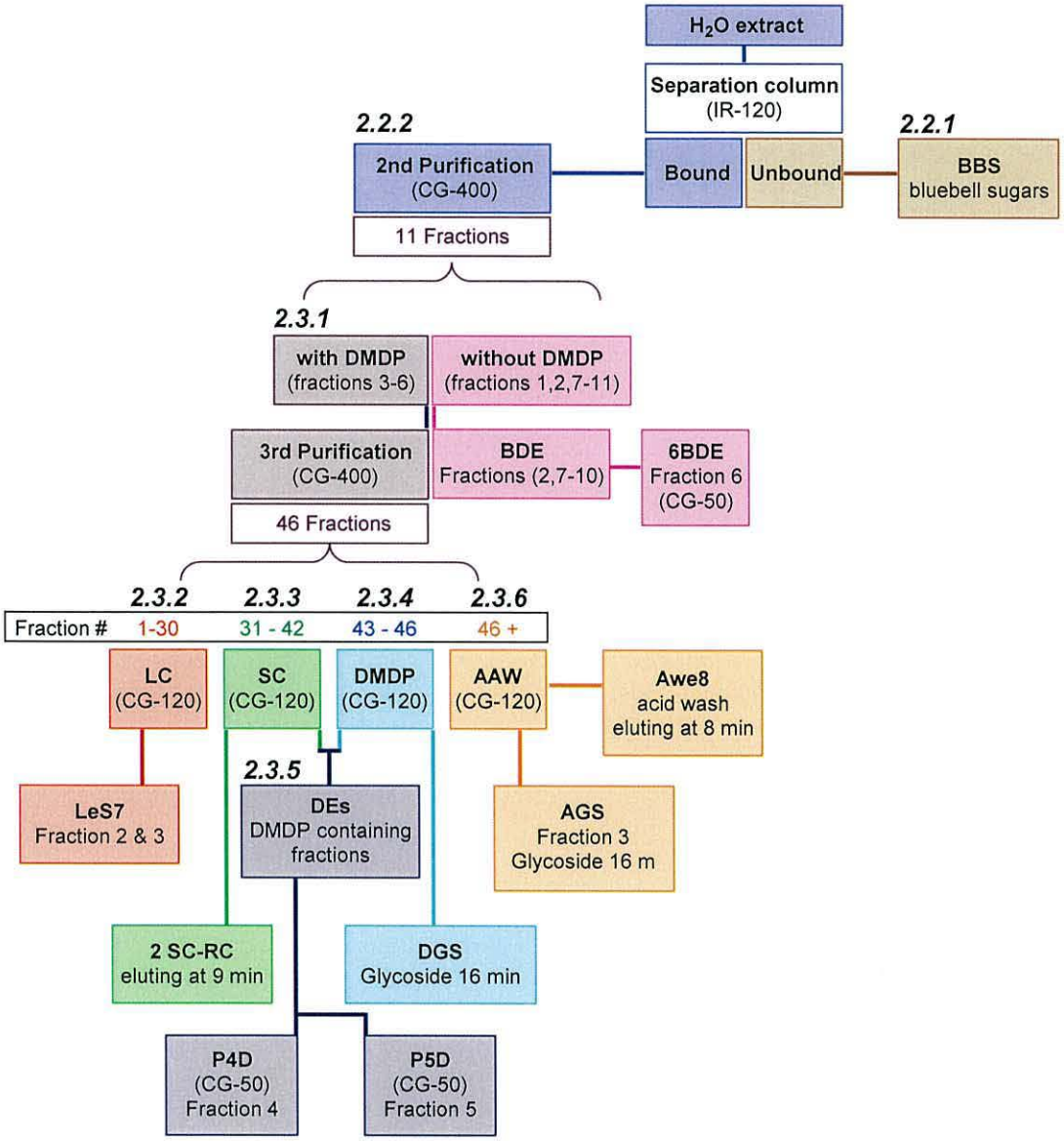


Figure 46. Flow chart of purification carried out by ion exchange chromatography with corresponding sections in bold.

The physical size of a resin bead is annotated by mesh size. Resin mesh size range from 20 - 400. The higher the mesh size, the higher the solute uptake for the resin. As the purification progressed, the resin mesh size decreased, so as to improve separation.

3.2.3.1 Bound Material

A column was prepared with Serdolit CG-400 anion exchange resin for the third stage. The crude bound material was loaded onto the top of the resin and run down the column. Fractions were taken at specific volumetric intervals (8 x 15 mL) and

analysed by GC-MS. For exhaustive elution water and acetic acid fractions were collected into round bottom flasks.

The presence of iminosugars other than DMDP and homo-DMDP in the seeds is seen by GC-MS (Figure 47). DNJ (Figure 48) and DAB (Figure 49) are seen in small quantities, eluting at 5.58 min and 10.97 min, along with myo-inositol at 12.24 min, amongst other alkaloids.

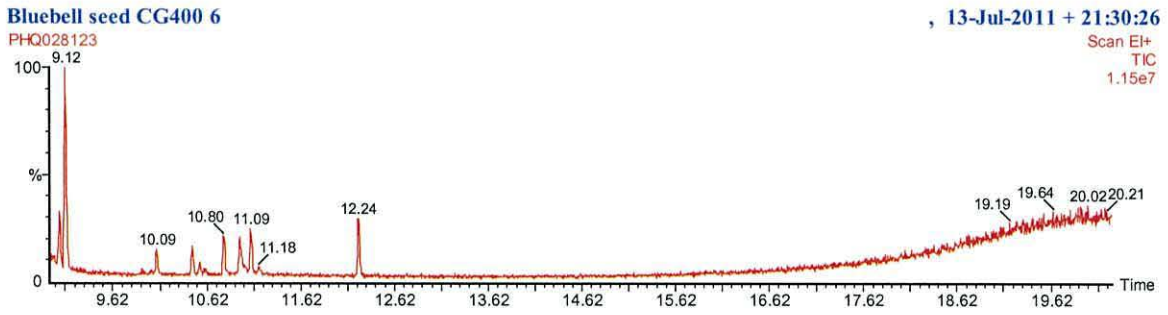


Figure 47. GC-MS data of first fraction.

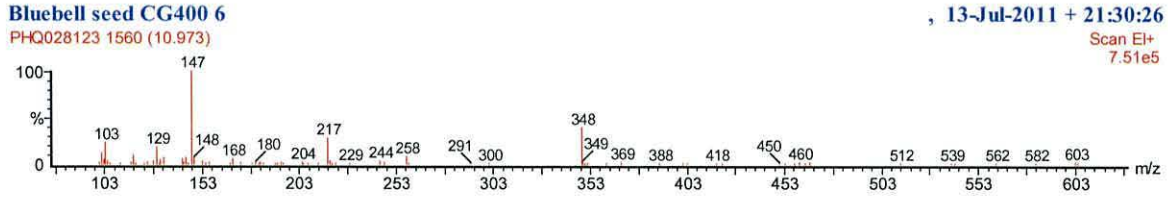


Figure 48. Mass spectrum shows traces of DNJ **1** at 10.97 min

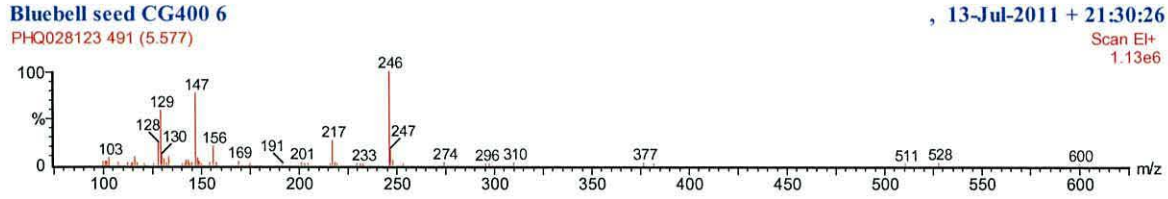


Figure 49. Mass spectrum shows compound eluting at 5.58 min to be the iminosugar DAB **2**.

DAB has one less hydroxylmethyl group than DMDP. This change in structure gives DAB a more potent selective inhibition against α-glucosidases, with no effect towards other β-glucosidases[126].

GC analysis revealed fractions 3 - 6 contained a mixture of mainly DMDP and homoDMDP. These fractions were combined and subjected to a further purification.

Column size was determined by the size of the sample and the binding capacity of the resin. The sample was loaded slowly by pipette onto the top of the Serdolit anion exchange resin. The material was run down the column, eluted with water, and collected in fractions. Two larger fractions were collected and concentrated on the rotary evaporator.

3.2.3.2 Fractions 1-30

In the first 30 fractions of the purification of initial bound extract, a large amount of material other than iminosugars is observed in the GC-MS spectra. DMDP appears to elute at 8.79 min (Figure 50), although in small quantities. To separate the DMDP from other material, a large column (500 mm) was prepared with Serdolit CG-120. With the amount of crude material being loaded onto the resin, a long column was chosen as to not exceed the binding capacity, and give sufficient time for the iminosugars to bind.

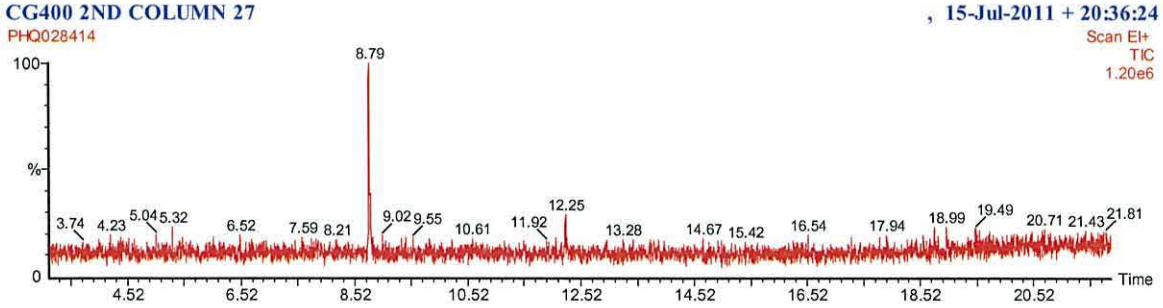


Figure 50. DMDP appears to elute in the first 30 fractions of the general purification column. Other compounds which may be of taxonomic interest are also present in smaller quantities.

The column, given the abbreviation *LC* (

Figure 46), was eluted to general procedure. The aim of this purification was to identify other compounds in the crude iminosugar mixture. Unidentified compounds were observed in the mixture, however without further purification, structural analysis could not be determined (Figure 51).

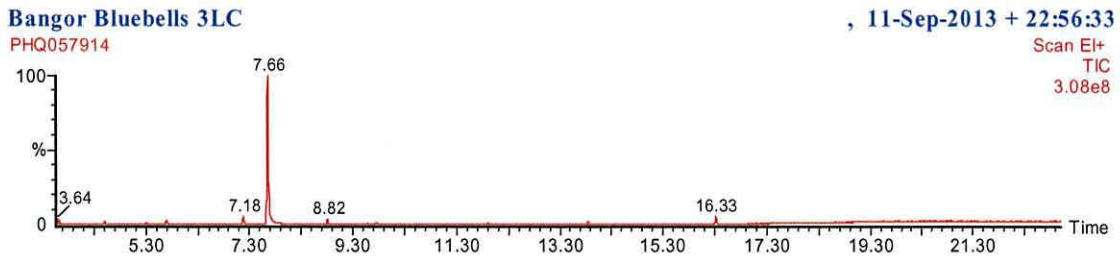


Figure 51. Unidentified compound eluting at 7.66 min.

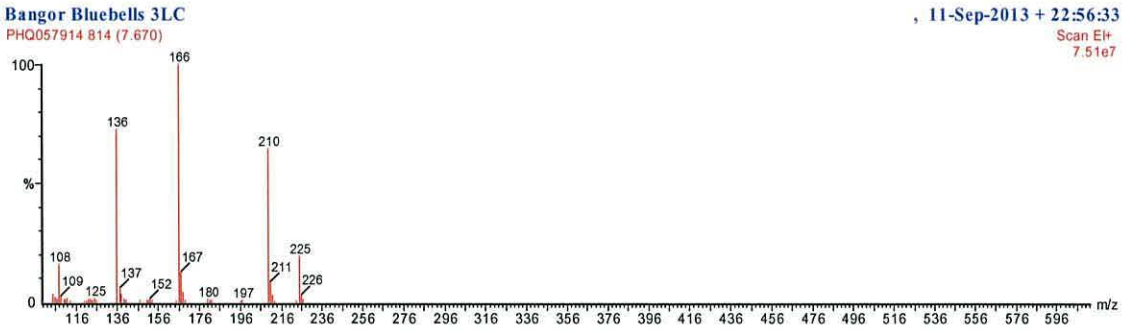


Figure 52. Mass spectra of unidentified compound at 7.66 min.

LC fractions 2 - 4 contained the unidentified material and were pooled together and purified in a column titled "*Les7*". Purification of these *LC* fractions did not yield information into the identification of the compound eluting at 7 minutes; further work must be done in order to carry out this objective. Despite this, this purification column led to the isolation of DMDP when a finer meshed resin was employed (Figure 53). The comparatively pure iminosugar was attained as a brown caramel totalling 97 mg.

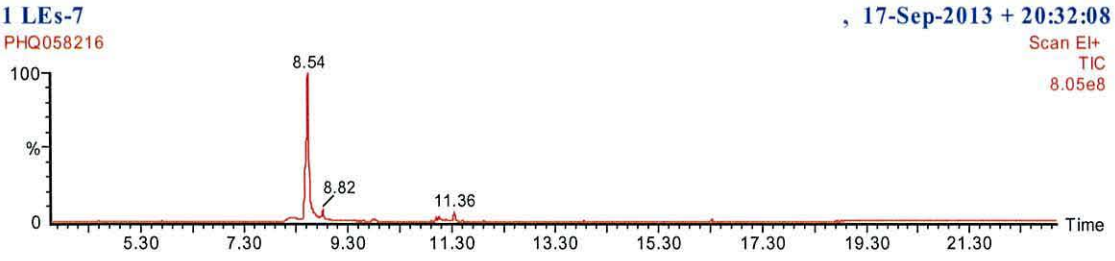


Figure 53. GC spectrum of purified DMDP via ion exchange chromatography.

3.2.3.3 Fractions 31-42

Between fractions 31 - 42, a large amount of seemingly isolated DMDP appeared (Figure 54). Due to the intensity being so high, purification was needed to remove any possible "background" noise. A small column (150cm^3) was prepared with Serdolit CG-120.

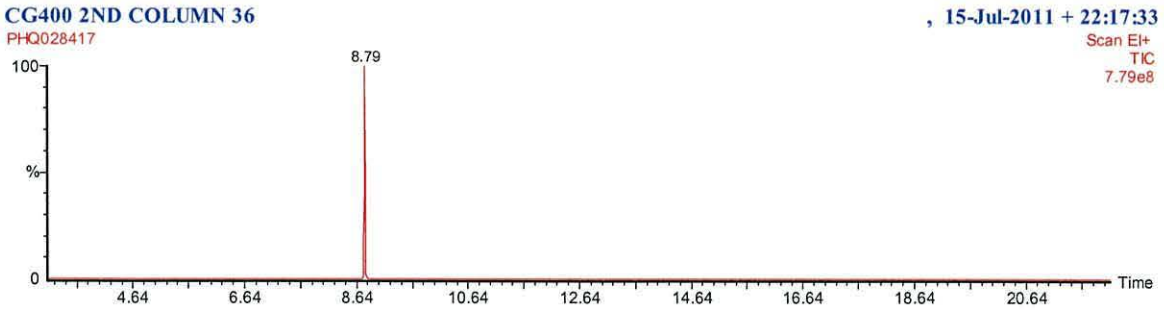


Figure 54. DMDP observable in fractions 31-42

As expected, once the DMDP containing fractions had been purified through a finer meshed CG-120 resin, other compounds could be seen to elute via GC. In the early fractions, a range of "background" material could be seen in small quantities. Some of these compounds could be potentially novel, although further purification and structural determination must be carried out before any conclusions can be made.

In total, 12 fractions were collected from the purification column *SCR*C. In the first 6 fractions, a glycoside could be seen to elute at 16.33 min (Figure 55). This glycoside would also be later observed in *DMDP* fractions 4 and 5. These glycoside containing fractions would be combined and purified under the acronym *DGS*.

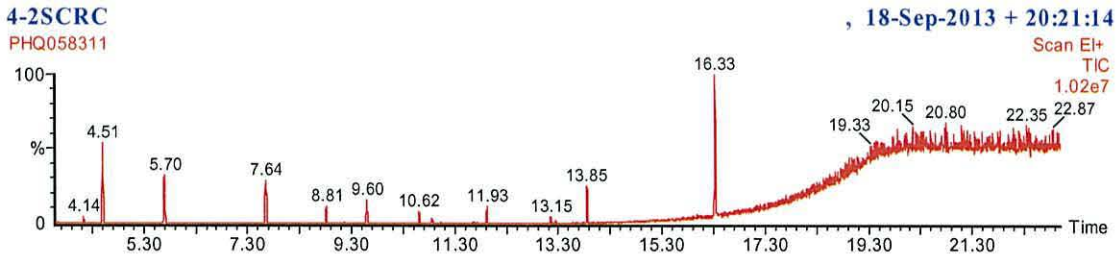


Figure 55. Purification of fractions 31-42

DMDP and homo-DMDP which eluted in the second half of collected fractions were analysed. As these compounds were not analytically pure, they were stored for further purification.

3.2.3.4 Fractions 43-46

A potentially novel iminosugar was observed in fraction 44 of the 2nd purification column (Figure 56 and Figure 57). An epimer of this unknown iminosugar was also observed (Figure 58). Further isolation will allow structural determination of these compounds.

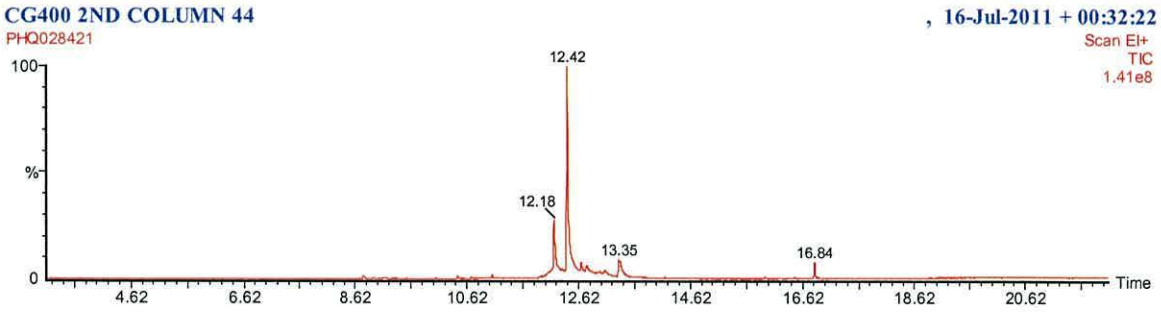


Figure 56. GC spectra of fraction 44 containing potentially novel iminosugar.

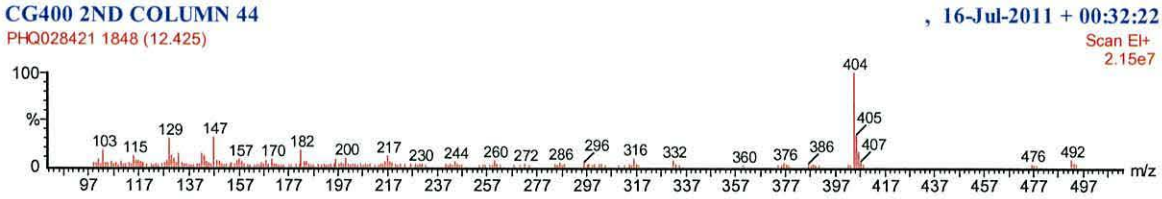


Figure 57. Mass spectra of a potentially novel iminosugar

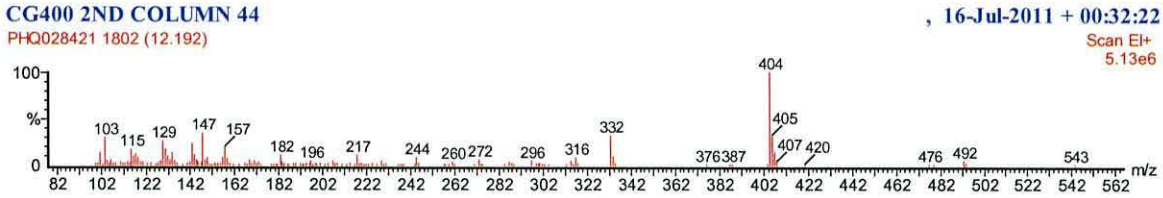


Figure 58. Epimer of the above iminosugar

This novel compound exists in such small quantities, that when combined with other fractions to purify, the signal intensity was so poor that it was suppressed by other components in the mixture. In order to isolate such novel compounds, the extraction must be carried out on a larger scale, with much more focus on each individual fraction.

Very weak signals at 10 and 11 min were also determined to be glucose derivatives, whilst a disaccharide or glycoside could be seen at 16 min.

DMDP containing fractions 43 - 46 were pooled together and purified using Dowex CG-120 anion exchange resin. Compounds were eluted as by previously stated, eluting first with water, then with NH_4OH. The majority compounds eluted in the first 9 fractions were not of interest for further studies. DMDP and homo-DMDP, appearing at 8.90 and 11.09 min respectively, began to elute in fraction 10 through 16 (Figure 59).

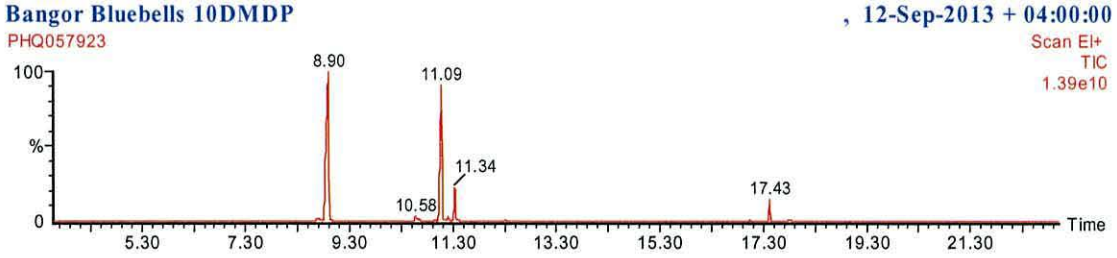


Figure 59. DMDP and homo-DMDP appears in fraction 10 at 8.90 and 11.09 min respectively.

3.2.3.5 DMDP containing fractions

Fractions containing the iminosugar DMDP from both "SC" (31 – 42) and "DMDP" (43-46) were combined for purification using a 200 cm^3 column.

DMDP and homo-DMDP were observed by GC-MS analysis to have eluted in the first fraction containing water (Figure 60). This would suggest that the column was overloaded. Further purification by ion exchange chromatography must be carried out to obtain fractions pure DMDP and homo-DMDP.

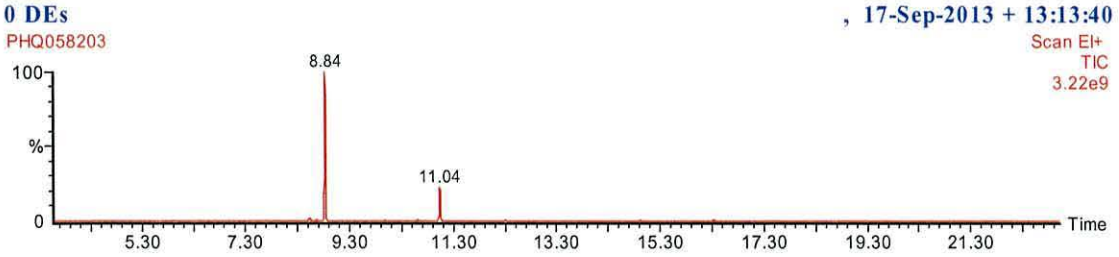


Figure 60. Presence of DMDP (8.84 min) and homo-DMDP (11.04 min) in the first fraction shows overloading of the column.

1.63 grams of iminosugar extract were calculated to be present in the "Des" iminosugar mixture. Pure DMDP can retail for as much as £35 per 1 mg (figures according to Carbosynth Ltd.). With over 1.6 g, the estimated retail price which was extracted can be calculated to be between £35,000 - £50,000 of DMDP. Scaling up the extraction and purification of bluebell seeds would be a lucrative project.

Producing these naturally occurring iminosugars to chemical suppliers would increase availability to the scientific community for further research, whilst also reducing costs.

3.2.3.6 Fractions 46+

Starting with fraction 46, exhaustive elution of the CG-400 purification column was performed with acetic acid. A relatively small amount of DMDP was found to elute in these "washing" fractions. This may be due to possible equilibration with the resin, resulting in retention on the column.

The acetic acid wash fractions were pooled together and subjected to further cation exchange with IR-120.

The glycoside eluting at 8.52 min was investigated further (Figure 61). Fractions containing this compound were pooled together for further purification.

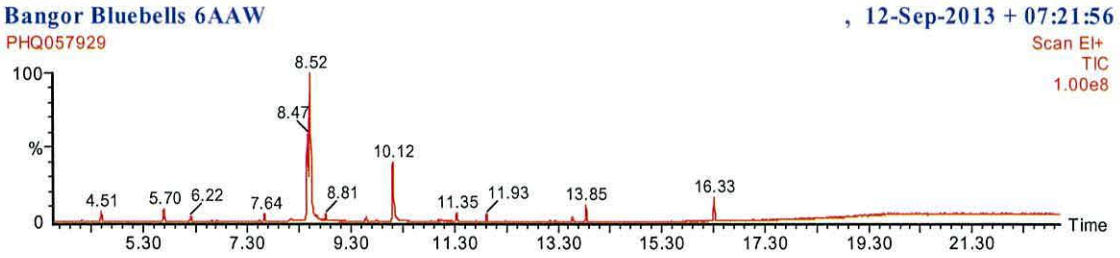


Figure 61. Glycoside found to elute at 8.52, with DMDP at 8.81.

An IR-50 column eluting with water gave good separation. This gave purification of a glycoside (Figure 62).

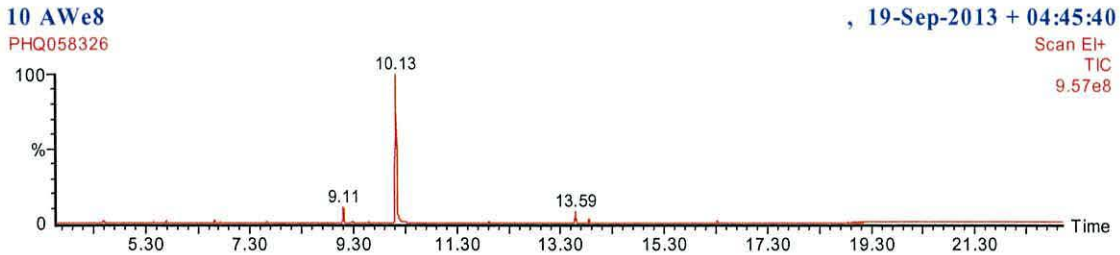
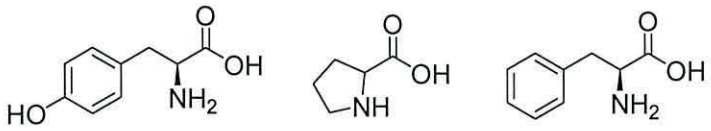


Figure 62. Example of purified glycoside in fraction 10.

3.2.4 Bound Material - Without DMDP

The bound material was purified again. Once sufficient elution using water had been performed, exhaustive elution with acetic acid was carried out. In this acid wash, we intended to wash all material which would have bound to the column.

As expected, three amino acids (Figure 63) were all observed to elute in a final acetic acid wash of the very first separation column.



tyrosine proline L-phenylalanine

Figure 63. Amino acids present in the unbound extract.

As amino acids are amphoteric, their charges can be positive, negative. Their elution pattern is based solely on ion exchange considerations. The aromatic nature of tyrosine and phenylalanine provides a negative electrostatic potential, which allows non-covalent bonding to cations in nature[129].

Although these common amino acids are not of relevance to this project, the parallel isolation of these compounds alongside the more biologically active substances is of interest. Amino acids are produced industrially using microorganisms and enzymatic processes in protein hydrolysis, and also by chemical synthesis[130]. Although amino acids are low in value, the extraction of amino acids from a natural source could also be taken as an additional "green" commercial opportunity.

3.3 Conclusions

3.3.1 Current Methods

Using a crude analytical method such as TLC during purification could help identify fractions of interest. This would increase productivity by decreasing the amount of time needed for concentration, followed by analysis of all fractions by GC-MS.

However, although TLC may improve the efficiency, it will not fix the issues with purification and identification of other compounds. The crude extraction method used in the laboratory was to study compounds obscured in the crude mixture's spectra. Although extraction of the lucrative iminosugars DMDP and homo-DMDP are of interest to us, it is the presence of novel compounds which must be investigated next.

3.3.2 Novel Compounds

The likes of such novel iminosugar compounds observed in some of the first purification fractions occur in small quantities. Without further purification and analysis on the preserved fractions, it is not possible to estimate the amount (w/w %) of novel compounds present. A scale up of the procedure would be required to have any significant form of structural analysis and biological testing on these compounds.

In this project, 1.2 Kg of seeds gave a yield of just over 6.2 g of crude iminosugar material, equating to about 0.5 % of the initial defatted weight. Hypothetically, the unknown iminosugars may make up around 0.1 % of this crude material; without future purification, any yield is unknown. This example would theoretically equate to around 6 mg of any novel compound. With a 10 fold scale up, the amounts of material which would obtainable would give more thorough analysis.

The process of commercialisation of iminosugars is much simpler than their synthetic counterparts, as they are of natural origin. The stricter GMP regulations applied to laboratory grade pharmaceuticals do not apply to as they do not need to pass the same rigorous mammalian testing.

Advancing our knowledge of the composition of these plants at various times of year and under specific circumstances is the next important step in the research of naturally occurring iminosugars. By understanding the factors that influence the production of these "defence mechanisms", organisms could be tailored to produce higher percentages of iminosugars. This in turn would increase the potential both in large scale extraction and marketability of these sugars.

3.3.3 Future work

All the groundwork has been laid in order to successfully complete the isolation of iminosugars from the defatted seeds of the common British bluebell, *Hyacinthoides non-scripta*. We expect to obtain isolated DMDP and homoDMDP in gram quantities within a few further purification columns.

In addition to this, we have identified the presence of potentially novel iminosugars. These compounds exist in such minute quantities that their signals on analytical spectra are suppressed by other components in the mixture. Extraction on a larger scale will lead to better separations of such compounds.

3.4 Experimental

Resins were purchased from Sigma Aldrich (Gillingham, Dorset, UK). Organic solvents were purchased from Fisher Scientific (Loughborough, Leics., UK). All samples were freeze dried before derivatisation. Trimethylsilyl (TMS) derivatives were prepared *in situ* using a mixture of hexamethyldisilazane and trimethylchlorosilane in pyridine (Pierce 'Tri-Sil' silylation reagent, HMDS:TMCS:pyridine in a ratio of 2:1:10). Analysis was carried out by GC-MS using a Perkin Elmer Autosystem XL gas chromatograph with a high polarity fused-silica column (Varian 'Factor Four' VF-5ms column, 25 m x 0.25 mm i.d., 0.25 µm phase thickness). The carrier gas (helium) flow rate was 1 ml min-1. Trimethylsilyl- (TMS) derivatives were separated using a temperature programme that started at 160 °C (5 min), followed by a linear increase to 300 °C at a rate of 10 °C min-1. The temperature was held at 300 °C for an additional 10 min. Total analysis time equated to 29 min. Electron impact mass spectrometry of the column eluent was carried out using a Perkin Elmer TurboMass Gold mass spectrometer with quadrupole ion filter system. This was run at 250 °C constantly during analysis. The detector mass range was set to 100 to 650 amu. The temperature of the transfer line (GC to MS) was held at 250 °C. Samples were injected onto the column via a split vent (split ratio 50:1) through a fused silica narrow bore injection liner packed with deactivated quartz wool; injection port temperature was maintained at 200 °C. Injection volume was 1 µl. System control, data collection and mass spectral analysis was carried out using Perkin Elmer TurboMass software (TurboMass v. 4.4).

Defatted seeds of the common British bluebell, *Hyacinthoides non-scripta,* were provided by Dr. Vera Thoss, Bangor University/Vera Bluebell, Llanberis, Gwynedd, Wales. Approximately 1.54 Kg of seeds were defatted by hexane extraction, to give 1.27 Kg of material.

3.4.1 Extraction Method

Water soluble sugars were extracted classically from the seeds by soaking in water (2 L) for 18 h at room temperature. Filtration of seed material was carried out by first passing the material over a cotton mesh to remove large, coarse particles, before further vacuum filtration. The soaking procedure was repeated twice more (2 x 1.5 L

H_2O) with the addition semi-permeable Visking tubing filled with Amberlite IR-120 cation exchange resin, to ensure full extraction of iminosugars. The second and third extracts were concentrated, and all three extracts were combined (ca. 6 L). An Amberlite IR-120 column was prepared; the resin packed inside Visking tubing was applied to the top of the column. The combined H_2O extracts were eluted through the column and the H_2O eluate concentrated.

The crude iminosugar material which had bound to the column was eluted with 2M NH_4OH. This eluate was concentrated to about 50 mL, and passed through a Serdolit CG 400 (OH^- form) anion exchange resin. GC-MS analysis revealed 4 fractions contained DMDP and homo DMDP. These 4 fractions were pooled together and purified by further anion exchange column chromatography.

3.4.2 Resin Preparation

3.4.2.1 Cation exchange resin activation (H^+ form)

The Na^+ form resin (200 g) was prepared by first soaking in 1M HCl (1 L) for 2 h. The aqueous layer was decanted, and the resin sequentially soaked with H_2O (x 2) for 1 h. After this time, the resin was loaded onto the column, washing with H_2O until a neutral pH was achieved.

3.4.2.2 Anion exchange resin activation (OH^- form)

The Cl^- form resin (200 g) was soaked in 1N NaOH (1 L) and left to soak for 2 h. The resin was loaded onto a column and washed with H_2O until a pH of 8 - 9 was reached and a yellow discolouration had ceased to elute.

3.4.3 Column Operation

3.4.3.1 General Procedure for Operation of Cation Exchange Column

Cation exchange resin was loaded onto a column and washed with a small amount of H_2O, in order to neutralise the resin at the top of the column. The product was carefully loaded onto the bed of resin by a Pasteur pipette, so as not to disturb the surface. H_2O was run through the column and collected into a 250 mL round bottom flask until a slight discolouration could be observed in the eluent. Fractions were

eluted with NH_4OH collected into round bottom flasks. Excess NH_4OH was removed by evaporation. Samples were transferred to pre-weighed glass vials with H_2O (2 mL) and freeze dried overnight. Freeze dried samples were analysed by GC-MS.

3.4.3.2 General Procedure for Operation of Anion Exchange Column

Cation exchange resin was loaded onto a column and washed with a small amount of H_2O, in order to neutralise the resin at the top of the column. The product was carefully loaded onto the bed of resin by Pasteur pipette, as not to disturb the surface. H_2O was run through the column. Fractions were eluted with H_2O into pre-weighed glass vials and freeze dried overnight. For exhaustive elution, 1M acetic acid was eluted into round bottom flasks. Excess acetic acid was removed by evaporation. Samples were transferred to pre-weighed glass vials with H_2O (2 mL) and freeze dried overnight. Freeze dried samples were analysed by GC-MS.

Column "LC"

Procedure	Anion Exchange
Resin	Serdolit CG-120
Starting material	Fractions 1-30 (3rd Purification column)
Column diameter (mm)	20
Column height (mm)	400

Column "SC"

Procedure	Anion Exchange
Resin	Serdolit CG-120
Starting material	Fractions 31-42 (3rd Purification column)
Column diameter (mm)	20
Column height (mm)	200

Column "DMDP"

Procedure	Anion Exchange
Resin	Serdolit CG-120
Starting material	Fractions 43-46 (3rd Purification column)
Column diameter (mm)	20
Column height (mm)	400

Column "AAW"

Procedure	Anion Exchange
Resin	Serdolit CG-120
Starting material	Fractions 46+ (3rd Purification column)
Column diameter (mm)	25
Column height (mm)	500

Column "Les7"

Procedure	Anion Exchange
Resin	Serdolit CG-120
Starting material	Fractions 2 & 3 (column *LC*)
Column diameter (mm)	20
Column height (mm)	400

Column "DEs"

Procedure	Anion Exchange
Resin	Serdolit CG-120
Starting material	Fractions 10-16 (column *DMDP*) Fractions 4-11 (column *SC*)
Column diameter (mm)	25
Column height (mm)	500

Column "AGS"

Procedure	Cation Exchange
Resin	Serdolit IR-120
Starting material	Fraction 3 (column *AGS*)
Column diameter (mm)	20
Column height (mm)	400

Column "Awe8"

Procedure	Cation Exchange
Resin	Serdolit CG-120
Starting material	Fractions 4, 5, 6 (Column *AAW*)
Column diameter (mm)	20
Column height (mm)	400

Column "BBS"

Procedure	Cation Exchange
Resin	Amberlite IR-120
Starting material	(unbound extract)
Column diameter (mm)	25
Column height (mm)	500

Column "BDE"

Procedure	Anion Exchange
Resin	Amberlite CG-50
Starting material	Fraction 2, 7, 8, 9, 10 (*2nd purification*)
Column diameter (mm)	20
Column height (mm)	400

Column "SCRC"

Procedure	Anion Exchange
Resin	Serdolit CG-120
Starting material	Fractions 1-3 (SC)
Column diameter (mm)	20
Column height (mm)	400

Column "DGS"

Procedure	Anion Exchange
Resin	Serdolit CG-120
Starting material	Fractions 1 – 7 (column *DMDP*)
Column diameter (mm)	20
Column height (mm)	200

Column "P4D"

Procedure	Anion Exchange
Resin	Serdolit CG-50
Starting material	Fraction 4 (column *DEs*)
Column diameter (mm)	20
Column height (mm)	200

Column "P5D"

Procedure	Anion Exchange
Resin	Serdolit CG-50
Starting material	Fraction 5 (column *DEs*)
Column diameter (mm)	20
Column height (mm)	200

4. Testing the Enzymatic Activity of Simple Glycosides

4.1 Introduction

4.1.1 Enzymes

Biological enzymes belong to a group of compounds called proteins[7]. Proteins are produced from genetic material which resides in the nucleus of living cells. In eukaryotic cells, proteins are actually produced at the ribosomes in the endoplasmic reticulum (ER). The proteins are transported to the Golgi apparatus from the ER in protective lipid vesicles (Figure 64).

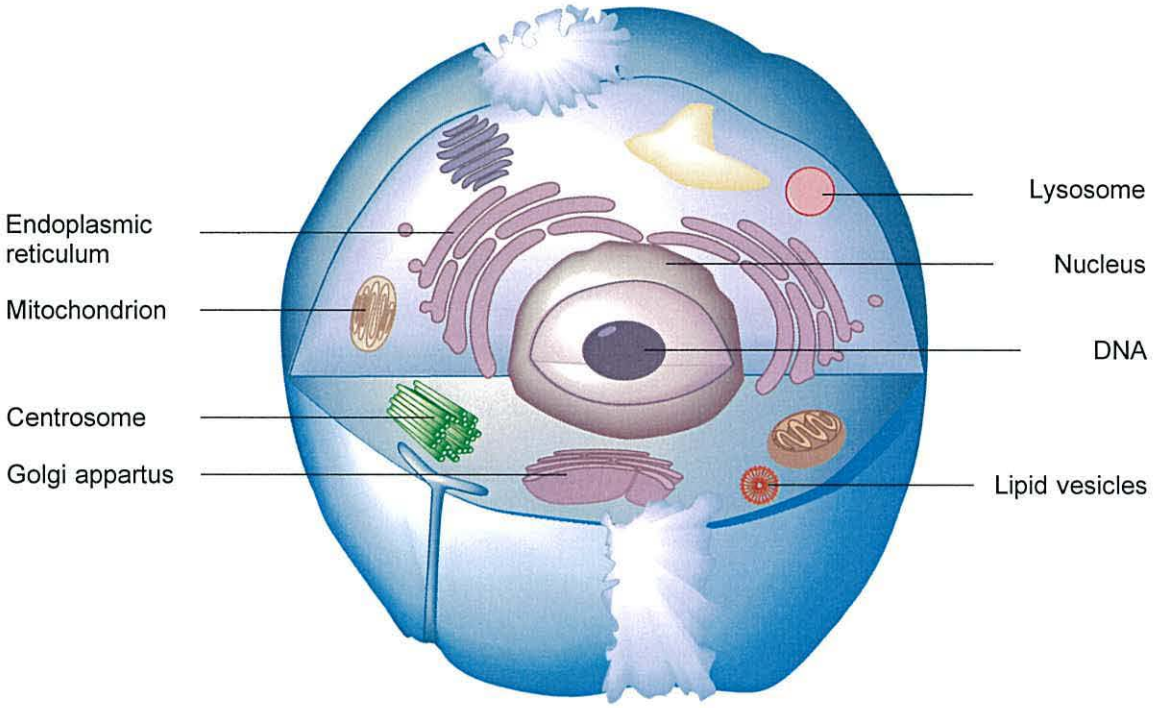


Figure 64. Major components of a typical animal cell.

At the Golgi apparatus, some proteins are provided a function, in a process called post-translational protein modification[131].

4.1.2 Glycoprotein Enzymes - Glycosidases

Sugars on the surface of proteins can affect the overall properties of the molecule; changing the solubility and protecting against peptidases. They also instruct the molecule how to fold correctly, and in the case of enzymes, assign its specific function[132].

Powerful catalysts for enabling these post translational protein modifications in the cell are enzymes called glycosidases[133]. Glycosidases have found wide spread applicability as antifungal agents, insecticides, anti-virals and anti-bacterials[134]. Glycosidases have also been used in the human healthcare industry as therapeutic agents for a wide range of genetic disorders. It is in this last area in which we are focused; specifically on lysosomal genetic disorders.

4.1.3 Lysosomal Enzymes

Lysosomes are sacks of digestive enzymes in the cell which are responsible for the disposal of waste material outside the cell (Figure 65)[135].

Lysosomal enzymes are synthesised on membrane bound ribosomes of the ER. They are carried across the cell cytoplasm to the Golgi apparatus along with other types of proteins by mannose 6 phosphate receptors which are recycled (Figure 65 A).

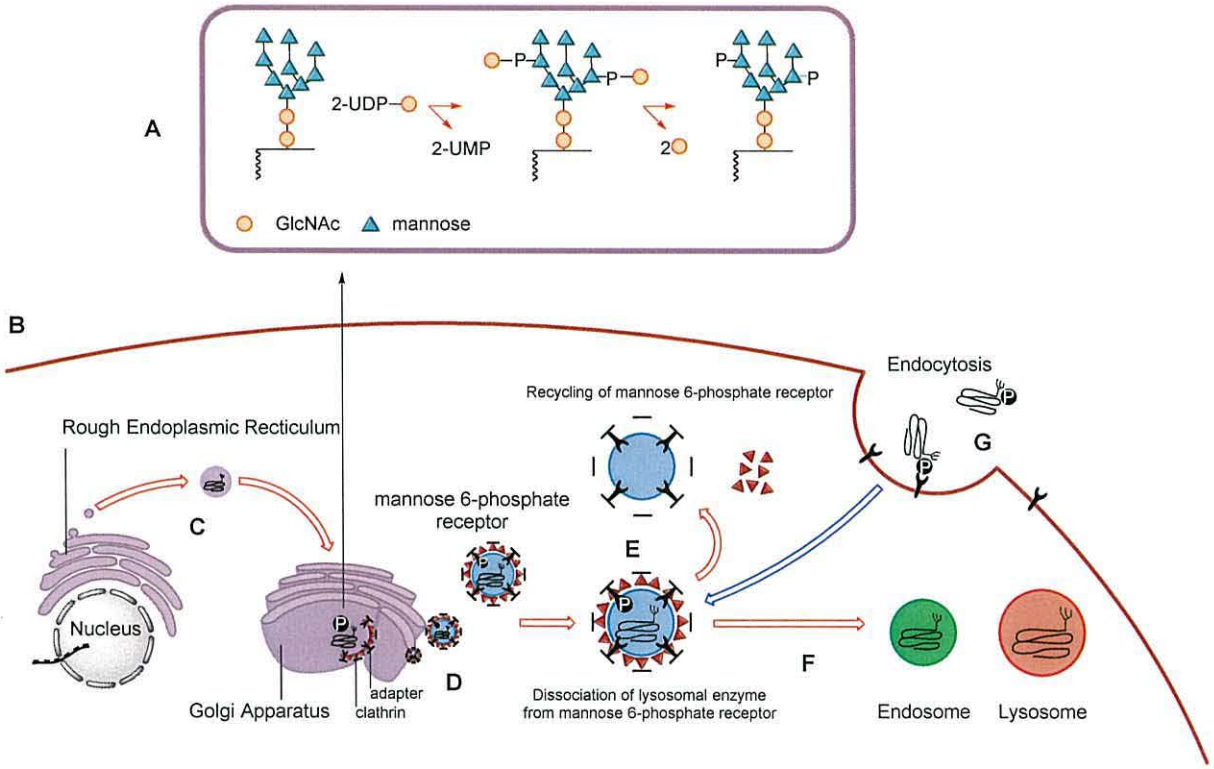


Figure 65. Schematic diagram depicting the biological pathway lysosomal enzymes undertake. Adapted from 'Cell and Molecular Biology: Concepts and Experiments'[136]

4.1.4 Glycosphingolipidoses

Glycosphingolipids are found extensively in the cell membranes of eukaryotic cells. They are produced by mechanisms similar to that used in glycoprotein synthesis; sugars are attached to proteins, instead of lipids. There is a large group of metabolic diseases that stem from defects in glycosphingolipid catabolism. These are sphingolipid diseases, also referred to as glycosphingolipidoses (GSL), and are typically inherited (Table 5).

Disease	Enzyme Deficiency	Clinical Effects	Possible Therapies
Gaucher's	β-glucocerebrosidase	Enlarged liver and spleen, mental retardation in infants	ERT (approved) SRT (clinical trials)
Fabry's	α-galactosidase	Skin rash, kidney failure	ERT (approved)
Krabbe's	β-galactosidase	Mental retardation	BMT
Tay-Sachs	Hexosaminidase A	Mental retardation, blindness	Gene therapy (pending)
Farber's	Ceramidase	Skeletal deformation, mental retardation; fatal in early life	BMT
Niemann-Pick	Sphingomyelinase	Enlarged liver and spleen, mental retardation; fatal in early life.	ERT (clinical trials)

Table 5. Properties of sphingolipid diseases. ERT, enzyme replacement therapy, SRT, substrate replacement therapy, BMT, bone marrow transplantation. Adapted from Platt[137].

Lipids undergo catabolism via acid hydrolysis in the endosomes and lysosomes. GSL are a complex type of sphingolipids which are modified by the addition of a carbohydrate group to the head of a ceramide backbone (Figure 66).

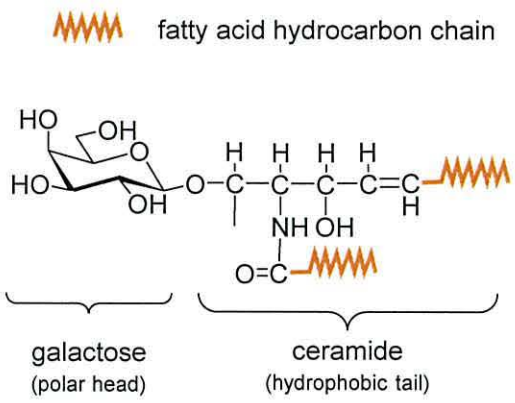


Figure 66. Structure of glycosphingolipid galactocerobrosidase. Adapted from Platt[137].

GSLs are a subset of lysosomal storage disorders which share certain features with other lysosomal disorders[138].

4.1.5 Orphan Diseases - Lysosomal Storage Diseases (LSD)

Rare diseases which affect a small percentage of the population, around 5 in 10,000 people, are referred to as orphan diseases. Orphan diseases are often chronic or life threatening. The majority have genetic components and will therefore be present throughout a person's entire life, even if symptoms do not appear immediately. Lysosomal storage diseases are an example of an Orphan disease.

The majority of LSDs stem from genetic mutations, causing a loss of or absence in lysosomal enzymes, leading to an accumulation of material in the tissue of the body[139]. The generation of this waste material is a slow but continual process, which explains why lysosomal diseases are progressive, protracted diseases[139]. Gaucher's and Fabry's diseases belong to the GSL subset of LSDs.

4.1.5.1 Gaucher's Disease

Gaucher's disease is the most common inherited sphingolipid metabolic storage disorder[140]. A deficiency of the enzyme glucocerebrosidase causes a build-up of the glycolipid glucocerebroside in macrophages[141]. The glycolipid originates with the biodegradation of red and white blood cells in the bodily organs[142]. Glucocerebrosidase in the brain stems from the production of complex lipids during brain development, including the formation of the myelin sheath of nerves.

There are 3 forms of Gaucher's; type 1 is the most common with organs effected including the spleen, liver, lungs, bone marrow[143]. Type 2 is a very rare infantile disease[144]; patients suffer from of liver and spleen, with severe brain damage[142]. As the disease is prominently linked with brain development, the life expectancy for a child with type 2 Gaucher's is 2 year of age. Type 3 Gaucher's has much better prognosis than type 2. It includes the possibility of liver and spleen enlargement and signs of brain and CNS deterioration become gradually prominent[144].

4.1.5.2 Fabry's Disease

Fabry's disease is another inherited sphingolipid storage disorder. The condition stems from the absence or lack of activity from the enzyme α-galactosidase in lysosomes, which results in the accumulation of phospholipids such as globotriaosylceramide in several different types of cell[145]. Unlike most other lysosomal diseases, Fabry's is an inherited condition which is specifically linked to the X-chromosome[146]. Both female and males suffer the same severity of symptoms and a reduction in life expectantcy[147]. Fabry's patients exhibit a wide range of symptoms affecting the central and peripheral nervous systems, skin, heart, kidneys, and eyes[148].

4.1.5.3 Pompe's Disease

Pompe's disease is another inherited condition, but it is not a sphingolipid disease.
It is caused by an α-glucosidase deficiency which results in an inability to break down glycogen in the cells[149]. Pompe's disease is an example of autosomal recessive metabolic myopathy. Metabolic myopathies are a group of hereditary muscle disorders which originate due to specific enzymatic defects because of defective genes.

4.1.6 Orphan Diseases and ERT Treatment

The only large-scale successful treatment to date is enzyme replacement therapy (ERT)[150].

ERT can be administered to patients whose enzyme activity is absent or lower than normal. Examples of drugs available on the market are 'Replagal' (Shire) and

'Fabrazyme' (Genzyme), which are used for the treatment of Fabry's[151], whilst 'Myozyme' (Genzyme) is used as a treatment for the Pompe's disease[152]. Drugs to combat Gaucher's disease also include Cerezyme (Genzyme), VPRIV (Shire) and Elelyso (Pfizer)[153].

However, as lysosomal enzymes do not cross the blood-brain barrier, ERT cannot be used to treat CNS symptoms in patients with type 1 or type 2 Gaucher's disease[137]. There is also a chance of immune responses in patients to the infused enzyme[151].

4.1.7 Orphan Diseases; Alternative Therapies (SRT)

Miglustat is an example of a drug approved for the specific treatment of type 1 Gaucher's that works using "substrate deprivation" or "substrate reduction therapy" (SRT). It works by inhibiting GSL biosynthesis[154]. In this way, the number of GSLs requiring lysosomal catabolism is reduced, thereby reducing accumulation of products in the lysosomes. Miglustat inhibits hydrolytic enzymes (α-glucosidases) which are involved in the biosynthesis of glycoproteins inside the endoplasmic reticulum[155]. It can cross the blood-brain barrier, and is used as treatment, and preventative for CNS depression in patients with Gaucher's[137].

There are several other treatments for patients with LSDs. Gene therapy is currently being investigated in clinical trials as a potential treatment for the Tay Sachs GSL disease (Table 5). Bone marrow transplantation is another possibility[156].

4.1.8 Alternative Biochemical Therapies

An alternative to drug therapy for the treatment of genetic diseases is biochemical therapy[157]. Biochemical therapies offer the possibility of further cost effective and simple Orphan disease treatments. They work at the molecular level by modification of the active sites of enzymes which can alter the metabolism of cells. A brief introduction to the biochemical nature of enzyme function and dynamics is provided next.

4.1.9 Enzyme Function and Dynamics

4.1.9.1 Inhibition of Enzyme Activity

Inhibition of enzymes by chemicals or analogues can occur *in vivo*[158]. There are 4 types of inhibition; reversible (Figure 67) and irreversible (Figure 68), and competitive and non-competitive (Figure 69).

Competitive inhibition occurs when an inhibitor, which has a similar structure to the substrate, blocks the active site of an enzyme and cannot undergo any chemical change. The effect of reversible inhibition is temporary. The blocking of the active site can be overcome as the inhibitor can be removed (Figure 67). Competitive inhibition is usually also reversible[158]

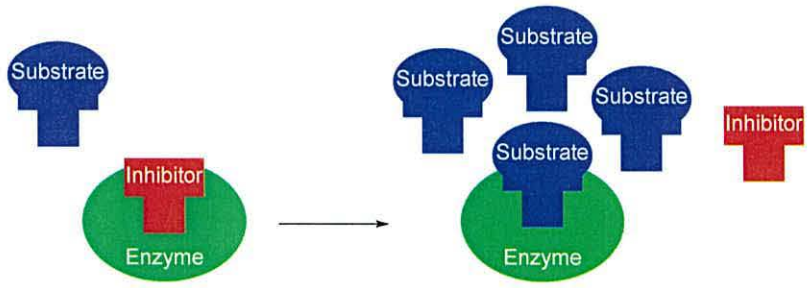


Figure 67. Reversible competitive inhibition

Iminosugars are known to inhibit glycosidases because of their structural resemblance to the sugar moiety of the enzyme substrate[159]. *N*-butyldeoxynojirimycin (NB-DNJ) was found to be an inhibitor of GSL biosynthesis in addition to its known activity as an inhibitor of α-glucosidase[154].

Inhibition can also occur when an inhibitor blocks an active site which in turn alters the conformation of the enzyme (Figure 68). This conformational change results in the substrate no longer "fitting" the enzyme.

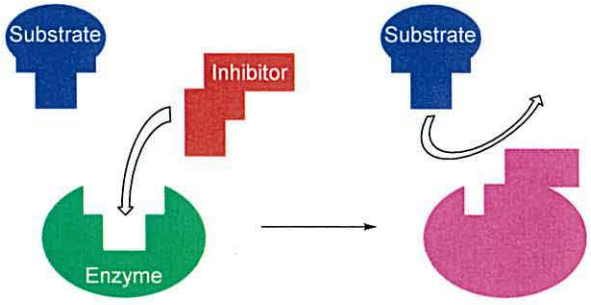


Figure 68. Irreversible Inhibition.

Non-competitive inhibition occurs when an inhibitor binds to an allosteric site of the enzyme, which in turn alters the active site (Figure 69).

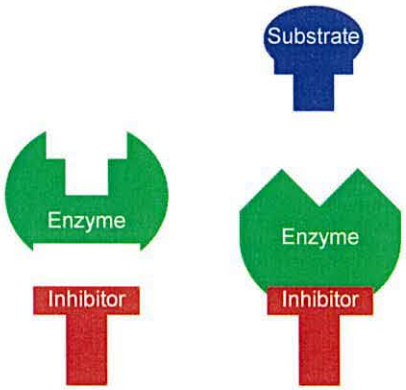


Figure 69. Non-competitive inhibition by allosteric site modification

The non-competitive inhibition of GSL biosynthesis by iminosugars was reported by Platt et al. in 1994[154]. Examples of iminosugars which have shown to directly or indirectly inhibit of glycosidases by allosteric site modification include NB-DNJ, DNJ and DMDP (Figure 70)[156].

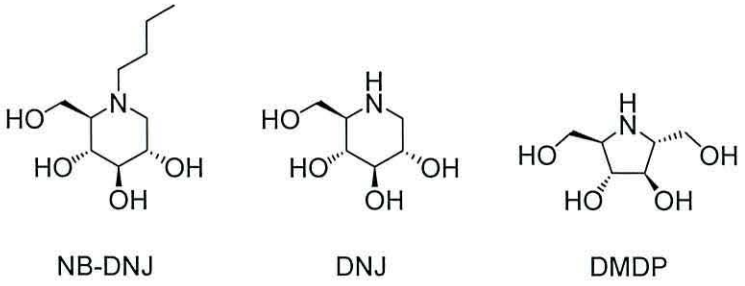


Figure 70. Iminosugars are shown to be potent glycoside inhibitors[160].

4.1.10 Activation

The active sites of enzymes can also be activated by biochemicals. Enzyme activators can stimulate enzyme activity above its normal level. When an enzyme activator binds to an allosteric site, the conformation of the active site can change, allowing the active site to adjust to the substrate (Figure 71).

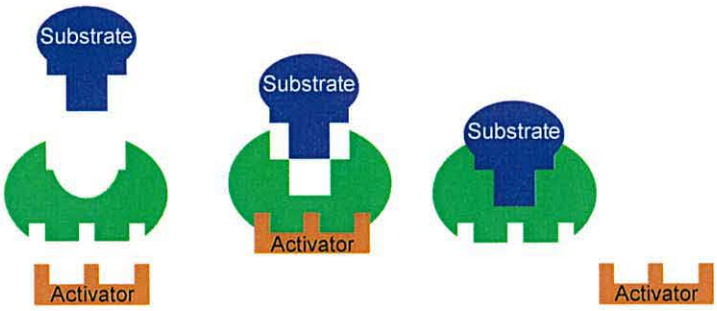


Figure 71. Activation of an enzyme by non-competitive activation at an allosteric site.

In enzyme kinetics, V_{max} is the maximal speed an enzyme can convert a substrate to product. C_m is the concentration of substrate required to produce 50% of the V_{max} value. Competitive and non-competitive inhibitors alter the values of both V_{max} and C_m values[161] (Figure 72).

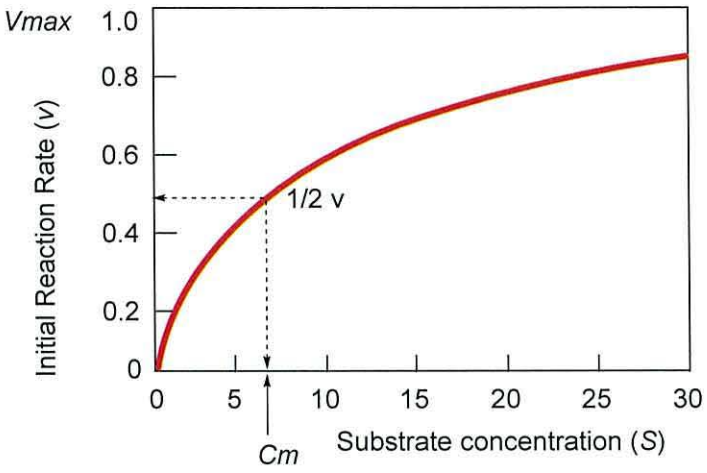


Figure 72. Michaelis-Menten rate law - Relationship between initial rate of reaction and substrate concentration. Adapted from Sauro[161].

In a normal lysosome, glycolipid substrate levels rarely appear in concentrations above the Cm value. When they do, excess catalytic activity ensures that all of the glycolipid material is hydrolysed[159]. It is only when all enzymes are bound to substrates, when concentration exceeds the rate of reaction, does accumulation takes place. (Figure 73)

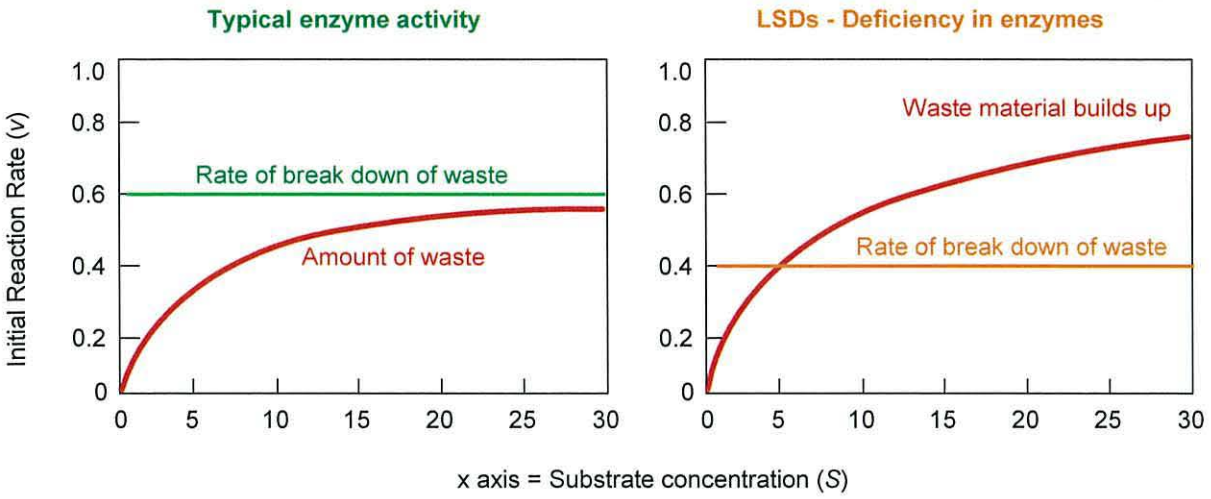


Figure 73. In LSDs the rate of hydrolysis cannot manage the build-up of waste material.

For example, in LSDs, when there is a deficiency of hydrolytic enzymes, accumulation of waste occurs, and the waste material is stored in the cells (Figure 73).

Two solutions which can be achieved by SRT are either activation of hydrolytic enzymes, or inhibition of the enzymatic processes involved in the biosynthesis of waste material (Figure 74). By increasing the folding of the hydrolytic enzymes, catalytic activity can be boosted to sufficient levels ensuring all waste material is hydrolysed correctly. Only small increases in the activation energy of the hydrolytic enzymes may be needed to allow efficient hydrolysis of waste[162]. On the other hand, by inhibiting the processes involved in GSL biosynthesis, the amount of waste material produced is maintained below harmful levels.

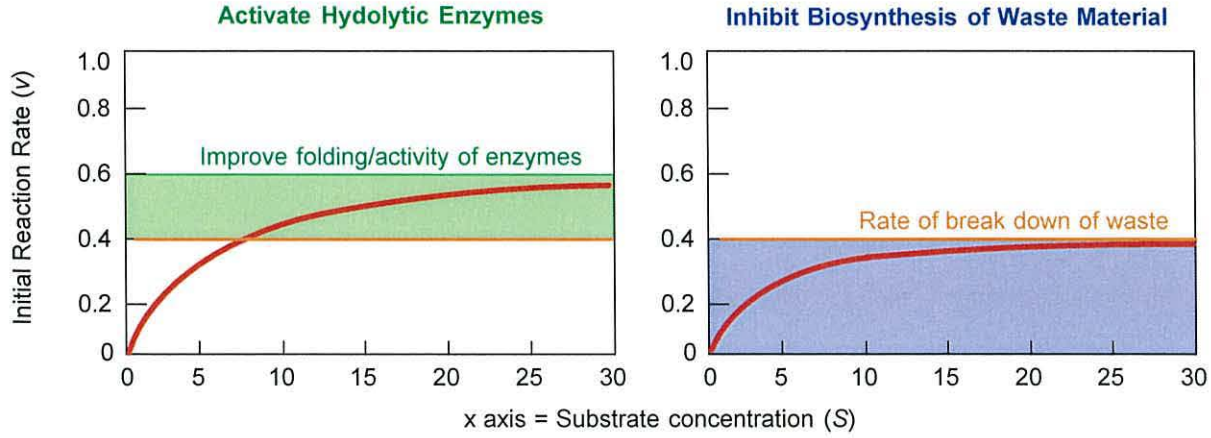


Figure 74. Either activation of hydrolytic enzymes, or inhibition of the processes leading to GSL biosynthesis are solutions in SRT.

4.1.11 Hypotheses

Thioglycosides have similar properties to iminosugars, and could play an important role in glycosidase and GSL inhibition and activation. Using commercially available enzymes as model systems, we set out to test a library of previously untested synthetic compounds for evidence of enzyme inhibition and activation. It was our hypothesis that the aromatic thioglycosides would be the best candidates for further study. The presence of an aromatic glycoside may stabilise the transition state via hydrogen bonding or pi-complexation at the active site of the enzyme[163].

Depending on the selectivity, any activity including activation could be of interest for further studies.

Enzymes are usually expected to interact with their specific substrates. However, it has also recently been found that α-D-glucosidase inhibitors may also exhibit β-D-glucosidase inhibitory activity[164,165]. We would test our compounds, looking for such non-specific activity.

4.2 Results and Discussion

4.2.1 Glycosides

The project involved the organisation, testing and analysis of the therapeutic compounds. Sampling was carried out within the laboratory, with a number of previously untested water soluble glycoside compounds selected from current research projects (Figure 75).

Preliminary purity tests of the sugars were carried out via TLC to check for any potential degradation. Whenever necessary, acetylated compounds were deprotected and purified via reverse phase column chromatography.

These compounds were tested as potential inhibitor/activator candidates against commercially available enzymes as a model of the human LSD diseases (Table 6).

4.2.2 Enzymes

The chosen enzymes acted as models of the GSL diseases against which inhibition or activation could be tested.

Bacterial enzymes included α-D-glucosidase and β-D-mannosidase from the Gram-positive bacteria *Geobacillus stearothermophilus* and *Cellullomonas fimi* respectively. α-D-glucosidase from the yeast, *Saccharomyces cerevisia*, was also tested. This eukaryotic cell is one of the most studied model organisms and is used as a standard in molecular biology.

Plant enzymes tested included α-D-galactosidase from the *Coffea* species of the Green coffee bean, β-D-glucosidase from the *Prunus* species of the Almond, and *N*-acetyl-β-D-glucosaminidase and α-D-mannosidase from *Canavalia ensiformis*, more commonly known as the Jack bean. Jack bean α-D-mannosidase has been shown to have similar mechanics to the human equivalent lysosomal enzyme and will provide a useful model in inhibition testing[166].

N-acetyl-β-D-glucosaminidase from bovine kidney was also tested, alongside α-L-fucosidase from the same mammalian source. In addition to β-D-galactosidase and β-glucuronidase from bovine liver, the activities of recombinant human enzymes were also investigated; Myozyme (α-D-glucosidase) Fabrazyme and Replagal (α-galactosidase). Either inhibition or activation of these enzymes may potentially be profitable in future.

As part of the work, we collaborated with Dr. Tarekegn Hiwot and Mrs. Katherine Peers of Queen Elizabeth Hospital, Birmingham, who kindly supplied vials of Replagal, Myozyme and Fabrazyme. The vials contained remnants of the respective enzyme after successful ERT infusion. Over the space of a few months, several hundred milligrams of three different recombinant enzymes were accumulated for testing.

4.2.3 Assays

In inhibition testing, the glycoside is pipetted into a well on the surface of a microplate. The protein is then added and the mixture incubated for 15 minutes at 30 °C. Finally, the enzyme's substrate is added in a pre-made pH buffer. The substrate changes colour upon reaction with the enzyme. The stronger the colour change, the weaker the inhibition values. A spectrophotometer quantifies the colour strength into percentage values. As the enzyme can continue reacting and producing colour, the reaction is stopped by addition of glycine.

Enzyme activity was tested at PhytoQuest laboratories, Aberystwyth, via colourimetric analysis. Enzyme activity values over ± 15 % were taken as an indicator of significant inhibition or activation, whilst values below this threshold were considered to have no inhibition (NI) (Table 6). Enzyme assays were carried out in triplicate, with mean values reported.

In total, 5 enzymes did not show any interaction with any of the glycosides tested; α-D-galactosidase (Fabrazyme, Coffee) and β-D-galactosidase (Replagal), as well as β-D-mannosidase (*Cellullomonas fimi*) and β-glucuronidase. No activity suggests that the sugar substrate is neither complementary to the enzyme as part of the lock and

key mechanism, nor drawn into the transition state configuration needed to accommodate the induced fit model.

4.2.3.1 Aglycon Structure

It is well established that phenols react with lipid radicals due to their hydrogen donating ability to form stable phenoxy radicals[167]. These phenoxy radicals are themselves effective chain breaking antioxidants[168]. Increasing the electron density by substitution of alkyl groups into either ortho or para position of the ring increases the overall reactivity with lipid radicals[169]. It is also generally well accepted that the presence of alkyl groups and other lipophilic moieties can increase the activity of glycosidase inhibitors[170]. Aglycon moieties with high polarities are detrimental to the overall reactivity. With this in mind, it is not surprising to see that neither allyl α-D-mannopyranoside **64** nor isopropyl 1-thio-β-D-glucopyranoside **68** shows any interaction with the enzymes tested, as these molecules are some of the more polar tested.

In a similar vein, the low inhibition values of methyl 4, 6-*O*-benzylidene-β-D-glucopyranoside **63**, methyl 4, 6-*O*-benzylidene-β-D-galactopyranoside **66** and 2-aminoethyl-β-D-lactopyranoside **79** can also be attributed to this factor. The linear natures of these protecting groups restrict the overall activity of the glycoside. Inversely, cyclohexyl 1-thio-β-D-glucopyranoside **70** exhibited the highest inhibition of β-D-glucosidase. Cyclohexane is a less polar substituent that the benzene structures we have analysed. It is clear that activity of the glycoside is strongly dependent on the structure of aglycon moiety.

4.2.4 *O*-glycosides

The *O*-glycosides showed very little activity across the range of various enzymes. allyl α-D-mannopyranoside **64** did not show any activity, whilst methyl 4,6-*O*-benzylidene-β-D-glucopyranoside **63** and methyl 4,6-*O*-benzylidene-β-D-galactopyranoside **66**, gave low specific inhibition of α-D-glucosidase from yeast and bacteria respectively. 1, 2, 5, 6-di-isopropylidene-α-D-glucofuranose **65** gave low activation against plant α-D-mannosidase with 32.4 % activation of the enzyme *N*-acetyl-β-D-glucosaminidase fairing a little higher.

		Thioglycosides																	
		O-glycosides				thioglucopyranosides					thiogalactopyranosides				other			lactosides	
		Compound																	
Type	Origin	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
α-D-glucosidase	*Saccharomyces cerevisiae*	18	NI	-4.5	13	30.7	6.6	-24.8	NI	4.5	22.4	40.8	-7.8	25.9	NI	-4.3	-11	11.2	6
α-D-glucosidase	*Bacillus sterothermophilus*	7	5.6	-12.7	18	38.9	11.8	-3.2	7.6	24.8	18.4	29.6	-6	23.1	24	NI	-56.5	41	19
α-D-glucosidase	'Myozyme'	NI	3.3	-1.9	NI	7.5	4.6	-9.4	3.4	5.4	7.8	-1.2	NI	5.2	-4	8.3	-3	-20.1	NI
β-D-glucosidase	Almond	NI	NI	11.9	NI	3.6	-2.1	-8.5	54.2	4.2	9.1	3.6	-6	4.3	-1	-4.1	-15.4	-1.5	NI
α-D-galactosidase	Coffee	NI	NI	-4.9	NI	NI	1.7	-1.4	-3.7	-9.7	-1	NI	-4.9	-5.1	-9.7	-8.6	-6.6	10.6	NI
β-D-galactosidase	Bovine liver	NI	-1	NI	NI	5.8	-4.2	2.6	18	15.5	-3.4	NI	5.6	18.4	NI	5.6	44.6	-3.3	NI
α-D-galactosidase	'Fabrazyme'	NI	2.8	-4.4	NI	NI	5.4	6.5	NI	-1.3	-1.5	4.6	1.5	1.4	1.9	2.2	NI	2.3	NI
α-D-galactosidase	'Replagal'	6	1.3	-1.3	6	-5.1	-3	-5.4	-1.3	-2.6	-5.1	-1.4	1.5	NI	-1.1	2.7	NI	NI	NI
α-D-mannosidase	Jack bean	NI	2.5	-20.4	NI	-13.7	NI	-25.8	-14.9	-13.8	-35.3	-18.1	-19.5	-13.1	-20.9	-22	-44.9	2.6	NI
β-D-mannosidase	*Cellullomonas fimi*	NI	2.4	-1.1	NI	-5.2	-5.5	-2.3	-3.4	NI	-5.4	3	1.1	1.9	-4.8	2.8	NI	1.7	NI
N-acetyl-β-D-gluc	Bovine Kidney	NI	-1.3	2.7	NI	1.2	-2.3	3.3	2.9	5.3	NI	7.4	NI	8.5	10.6	-9.4	-17.5	-3.8	NI
N-acetyl-β-D-gluc	Jack bean	10	-4.9	-32.4	14	26.1	5.4	23.4	5.9	9	10.6	-4.9	-37.4	23.2	-18	-23.6	-57.4	-8.1	8
β-glucuronidase	Bovine Liver	NI	1.2	NI	NI	-2.1	-1	NI	NI	NI	-3	NI	1.5	-1.4	NI	NI	NI	-2	NI

Activation	Low	High (> 40 %)
Inhibition	Low	High (> 40 %)

No Interaction (NI) *(less than 1 %)*

Table 6. Mean % inhibition at 143 µg/ml

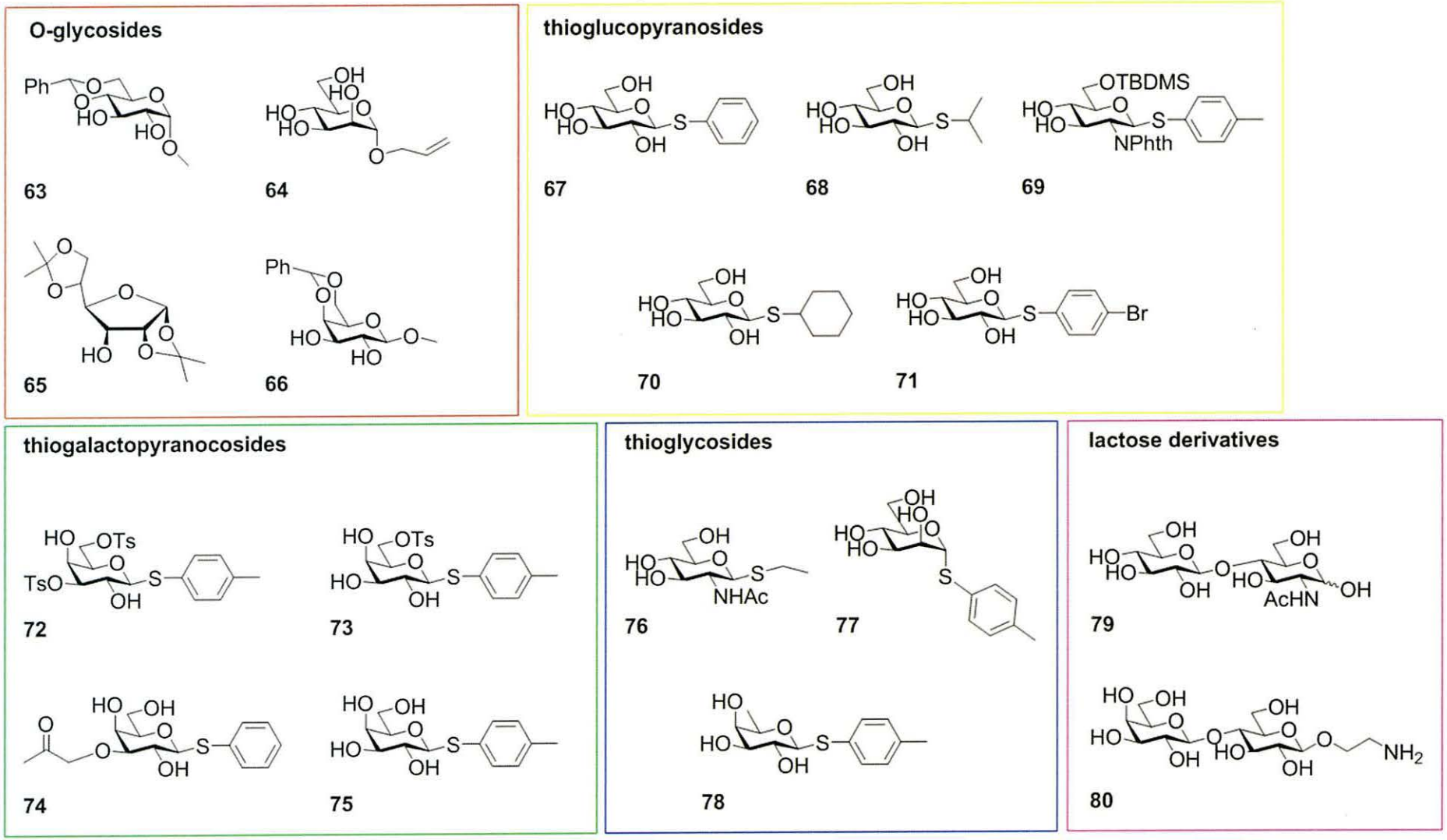


Figure 75. List of compounds assayed for inhibition/activation of enzymes

4.2.5 Thioglycosides

4.2.5.1 Confirming the Hypothesis

The hypothesis that thioglycosides would prove to be the most effective inhibitory compounds was proved to be correct. The range of *O*-glycosides and lactose derivatives did not fare as well as the sulphur containing compounds.

The secondary hypothesis that tetrol thioglycosides would give the best inhibition potential amongst other tested compounds was also proved to be correct. Compounds **67**, **71** and **75**, identified as prime candidates prior to the experiment, proved to exhibit good inhibition across the board; neither **67**, **71** nor **75** exhibited any activation.

Compound **75** exhibited the widest range of inhibition of the 3 identified thioglycosides. This may be due to the fact that **75** is a thiogalactoside, whilst **67** and **71** are both thioglucosides.

The activity of **75** may also be related to the presence of the methyl group on the aglycon part of the structure. This would indicate that the positioning of electron density around the aglycon substituent directly affects the reactivity of the glycoside with the lipid as hypothesised. To determine which factor affects the reactivity, a list of glucose and galactose derivatives should be synthesised to compare results.

4.2.6 Aromatic Protecting Groups

With the addition of the tosyl groups in compounds **72** and **73**, the steric bulk of the compound is increasing. With this increase, comes a decrease in the rate of competitive inhibition. However, compound **73** exhibited some of the highest inhibition values observed amongst α-D-glucosidase. This compound also exhibited low activation of α-D-mannosidase, which the unprotected analogue did not. Addition of a second tosyl group to the 4 position, yielding compound **72**, effectively doubled this activation value to 35 %.

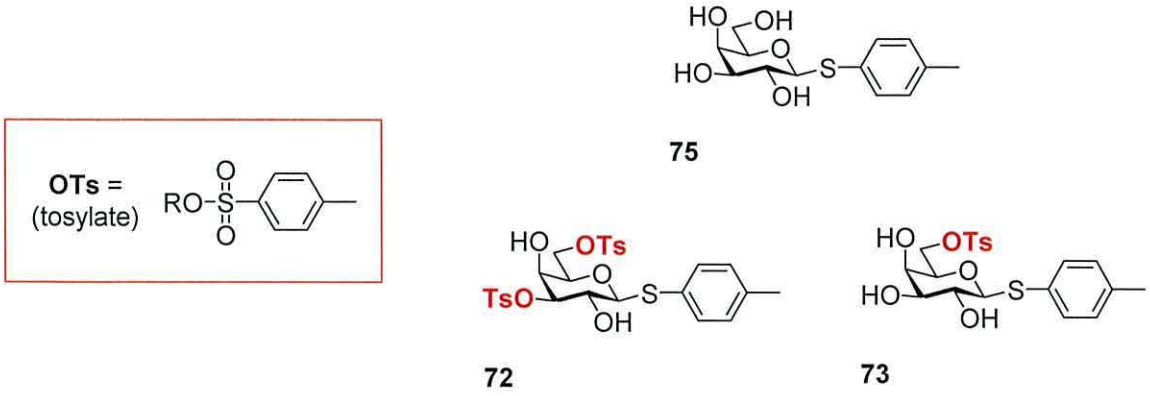


Figure 76. Compounds **72** and **73** differ only by the positions of the tosyl protecting group(s).

The mechanism of activation is currently unknown. One possible explanation may be that the tosylate may interact similarly to the aglycon with lipids of the glycoprotein improving folding of the protein. The results suggest that the activation and inhibition values for compounds significantly affected by the position and multiplicity of protecting groups on the carbohydrate skeleton.

4.2.7 Fucoside

The majority of the thioglycosides showed very weak activation towards α-D-mannosidase from Jack bean, with the exception of *p*-methylphenyl 1-thio-β-D-galactopyranoside **75** and isopropyl 1-thio-β-D-glucopyranoside **68**. Compounds **78** were exhibited the highest activation value of 45 %.

The most active compound to be analysed was *p*-methylphenyl 1-thio-β-D-fucopyranoside **78**. The D-fucose derivative gave the highest values of both inhibition and activation across the range of enzymes. This entailed activation of α-D-glucosidase (bacterial), α-D-mannosidase and *N*-acetyl-β-D-glucosaminidase (plant), whilst β-D-galactosidase (mammalian) was inhibited. Going forward, a range of fucoside compounds should be prepared in the laboratory for testing. Synthesising various sulphur containing glycosides may lead to the development of a glycoside with superior activation than what we have observed.

There are reports of fucosides being used to improve binding of oligosaccharide moieties in literature[171]. Huang et al. studied the interaction of cell-surface carbohydrates with monoclonal antibodies in a fluorescence-based binding assay.

The authors also determined that presence of the fucose moiety on the oligosaccharide chain was required for binding[171].

Huang also found that generally, the complexity of oligosaccharides was directly related with binding capacity; as one increased, the other did too[171].

4.2.8 Lactose Derivatives

In addition to the glycosides, *N*-acetyl lactosamine **79** and 2-aminoethyl-β-D-lactopyranoside **80** were also tested. Both lactose derivatives gave partial inhibition of the bacterial α-D-glucosidase, with **79** exhibiting a high inhibition of 40 %. *N*-acetyl lactosamine is widely used in studies of lectin binding, has been previously reported to be a fairly potent inhibitor of galectins[172].

The recombinant human enzyme replacement drug Myozyme was also partially activated by **79**. Although this was a very low activation value (~ 20 %) this sugar could prove to be interesting if used in conjunction with the drug. Further testing may lead to better insights to whether *N*-acetyl lactosamine **79** could be used as a supplement.

4.3 Conclusion

4.3.1 Glycoside activity

Our results show simple glycosides can be used as good inhibitors and activators of enzymatic activity. We have identified thioglycosides, especially those with cyclic aglycons, as potent inhibitors of a range of model enzymes. We observed an increase in inhibition with cyclic aglycons as predicted, as well as an unexpected increase in activity with the addition of cyclic protecting groups around the glycon.

Despite differences in values, all compounds exhibiting activity above 15 % can be said to be of interest to further studies. These compounds can now be added to the growing collection of structures that have some sort of biological activity.

4.3.1 Enzymes

Both allyl α-D-mannopyranoside **64** and isopropyl 1-thio-β-D-glucopyranoside **68** did not show any activity. On the other hand, methyl 4, 6-*O*-benzylidene-β-D-glucopyranoside **63** and *N*-acetyl lactosamine **79** showed very low, yet specific inhibition against α-D-glucosidase from *Saccharomyces cerevisiae and Bacillus sterothermophilus* respectively. This low activity may be of further interest, as specificity could potentially be useful.

In general, *N*-acetyl-β-D-glucosaminidase from Jack bean was activated by a range of glycosides; the highest percentages falling to compounds **72** and **78**. This may be due to either improved folding of the enzyme *in vitro* or allosteric binding[173].

In comparison, α-D-glucosidase from *Bacillus sterothermophilus* was inhibited by a wide range of glycosides. Phenyl 1-thio-β-D-glucopyranoside **67** and *N*-acetyl lactosamine **74** gave the highest inhibitory results at around 40 %, whilst *p*-methylphenyl 1-thio-6-*O*-*p*-toluenesulfonyl-β-D-galactopyranoside **73** gave 30 % inhibition. The fucoside **78** showed good activation at 56.5 %.

4.3.2 Fucoside

The most promising results of this study are that of cyclohexyl 1-thio-β-D-glucopyranoside **70** and *p*-methylphenyl 1-thio-β-D-fucopyranoside **78**, which exhibit over 50% activation and inhibition respectively.

N-acetyl-β-D-glucosaminidase extracted from bovine kidney gave no activity apart from a low, yet specific activation by *p*-methylphenyl 1-thio-β-D-fucopyranoside **78**. This thiofucoside showed the highest activation and inhibition percentages across the board, exhibiting 4 of the 5 highest results. Good inhibition was observed against the mammalian β-D-galactosidase from bovine liver, whilst also showing very good activation towards the bacterial α-D-glucosidase, and both α-D-mannosidase and *N*-acetyl-β-D-glucosaminidase from Jack bean.

4.3.4 Future Work

Thorough research has yet to be done on the biological activity of thioglycosides. Firstly, an investigation into the factors which affect reactivity should be completed. A comprehensive library of glucose and galactose derivatives should be synthesised. By analysing the activities the different aglycon components implement, the characteristics of a successful inhibitor could be discovered, allowing glycoside tuning for inhibitory effects.

Ideally, comparisons should be made between either ortho- or para- substituted phenyl structures or cyclohexyl rings connected to the anomeric position.

In addition to protecting the anomeric position, substituting protecting groups around the skeleton could be of interest. The activation of thiogalactoside **75** increased with the addition of cyclic tosyl groups. Further studies into the positioning and reactivity of protecting groups on the glycon may give us a better understanding about these various “activation” factors.

N-acetyl lactosamine was the only sample tested which had a positive effect on the recombinant human α-D-glucosidase. Further testing should be carried out to see whether this compound could be used in conjunction with the drug.

4.4 Experimental

4.4.1 General Methods

Classical deprotection was carried out *in situ*. To a stirred solution of the protected sugar in methanol, a catalytic amount of NaOMe was added. The reaction was stirred for 1 hour, until TLC (toluene/ethyl acetate 1:1, dichloromethane/methanol 7:1) showed complete removal of acetates. Amberlite IR-120 ion exchange resin was added until a neutral pH was reached.

Purification of water soluble monosaccharides was performed via reverse phase column chromatography (water/methanol, 19:1 – 1:4). Samples were concentrated and freeze-dried overnight.

As all samples were previously synthesised and analysed in the lab, crude purity analysis of samples was carried out by TLC. If necessary, purification of non-water soluble compounds was performed via flash chromatography (ethyl acetate/methanol, 30:1). Samples were concentrated and dried *in vacuo*.

4.4.2 General Method for Enzyme Assays[174]

All enzymes and para-nitrophenyl substrates were purchased from Sigma Aldrich (Dorset, UK), with the exception of beta-mannosidase which was purchased from Megazyme (Bray, Wicklow, Rep. Ireland). 'Replagal' (Shire) 'Fabrazyme' (Genzyme) and 'Myozyme' (Genzyme) were supplied by Ms. Katherine Peers and Dr. Tarekegn Hiwot, Queen Elizabeth Hospital, Birmingham, UK.

Activity/inhibition against a range of available glycosides was assayed at a pH optimum for each enzyme. The activities screened were α-D-glucosidase (*Saccharomyces cerevisiae, Bacillus sterothermophilus, rice*), β-D-glucosidase (Almond (*Prunus sp.*)), α-D-galactosidase (Green coffee bean (*Coffea sp.*)), β-D-galactosidase (Bovine liver), α-L-fucosidase (Bovine kidney), α-D-mannosidase (Jack bean (*Canavalia ensiformis*)), β-D-mannosidase (*Cellullomonas fimi*), *N*-acetyl-β-D-glucosaminidase (Bovine kidney, Jack bean), β-glucuronidase (Bovine liver), α-D-glucosidase (Myozyme), α-galactosidase (Replagal), and α-galactosidase (Fabrazyme).

Enzyme assays were carried out at 27 °C in 0.1 M citric acid/0.2 M disodium hydrogen phosphate buffers at the optimum pH for the enzyme. An incubation mixture of 10 µl enzyme solution, 10 µl (1 mg/mL aq. soln.) of extract and 50 µl of the appropriate para-nitrophenyl substrate (5 mM) was prepared in buffer at the optimum pH for the enzyme. Enzyme and inhibitor were pre-incubated for 15 min at 30 °C before starting the reaction. Reactions were stopped by addition of 70 µl of 0.4 M glycine (pH 10.4) during the exponential phase of the reaction. This phase had been determined prior to the test, using blank uninhibited assays in which water replaced inhibitor. Final absorbances were read at 405 nm using a Versamax microplate reader (Molecular Devices). Assays were carried out in triplicate; values given are means of replicates.

4.4.3 Glycosides

4.4.3.1 Methyl 4,6-O-benzylidene-α-D-glucopyranoside[175] (63)

TLC (methanol/ethyl acetate 1:4, R_f 0.62).

4.4.3.2 Allyl-α-D-mannopyranoside[176] (64)

TLC (methanol/ethyl acetate 1:4, R_f 0.3).

4.4.3.3 1, 2, 5, 6-Di-isopropylidene-α-D-glucofuranose[177] (65)

TLC (toluene/ethyl acetate 2:1, R_f 0.45).

4.4.3.5 Phenyl 1-thio-β-D-glucopyranoside[178] (67)

^{1}H NMR (400 MHz, D_2O) δ 7.59 (d, *J* = 6.8 Hz, 2H, C_6H_5), 7.42 (m, *J* = 5.9 Hz, 3H, C_6H_5), 4.81 (s, 1H, H-1), 3.90 (d, *J* = 11.7 Hz, H-6a, 1H), 3.72 (dd, *J* = 12.4, 5.6 Hz, H-6b, 1H), 3.53 (t, *J* = 8.9 Hz, H-3, 1H), 3.48 (d, *J* = 4.0 Hz, H-5, 1H), 3.43 (d, *J* = 9.2 Hz, H-4, 1H), 3.36 (t, *J* = 9.4 Hz, H-2, 1H); ^{13}C NMR (101 MHz, D_2O) δ 131.97, 131.59, 129.30, 128.08, 87.24, 79.86, 77.19, 71.69, 69.32, 60.76.

4.4.3.8 Cyclohexyl 1-thio-β-D-glucopyranoside (70)

^{1}H NMR (400 MHz, D_2O) δ 4.63 (d, *J* = 9.9 Hz, H-1, 1H), 3.90 (d, *J* = 12.3 Hz, H-6a, 1H), 3.70 (dd, *J* = 12.4 Hz, H-b, 1H), 3.48 (d, *J* = 8.9 Hz, H-3, 1H), 3.47 (m, H-5, 1H), 3.42 (d, *J* = 8.9 Hz, H-4, 1H), 3.28 (t, *J* = 9.3 Hz, H-2, 1H), 2.02 (m, C_6H_{12}, 2H), 1.78

(m, C_6H_{12}, 2H), 1.64 (m, C_6H_{12}, 2H), 1.42 (m, C_6H_{12}, 2H), 1.36 (m, C_6H_{12}, 2H), 1.33 (m, C_6H_{12}, 2H).

4.4.3.13 p-Methylphenyl 1-thio-β-D-galactopyranoside[179] (75)

^{1}H NMR (400 MHz, DMSO-d_6) δ 7.35 (d, *J* = 8.1 Hz, $CH_3C_6H_4S$, 2H), 7.11 (d, *J* = 8.0 Hz, $CH_3C_6H_4S$, 2H), 4.48 (d, *J* = 9.3 Hz, H-1, 1H), 3.70 (d, *J* = 2.7 Hz, H-5, 1H), 3.54 – 3.44 (m, H-6a/b, 2H), 3.41 (dd, *J* = 4.4 Hz, H-2, 1H), 3.41 – 3.35 (m, H-3, 1H), 3.33 (dd, *J* = 9.0 Hz, H-4, 1H), 2.26 (s, $CH_3C_6H_4S$, 3H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 136.24, 131.96, 130.61, 129.88, 88.62, 79.59, 75.17, 69.70, 68.78, 60.99, 40.60, 40.39, 40.18, 39.97, 39.76, 39.55, 39.34, 21.05.

4.4.3.16 p-Methylphenyl 1-thio-β-D-fucopyranoside[180] (78)

^{1}H NMR (400 MHz, D_2O) δ 7.38 (d, *J* = 8.0 Hz, $CH_3C_6H_4S$, 2H), 7.17 (d, *J* = 7.9 Hz, $CH_3C_6H_4S$, 2H), 4.58 (d, *J* = 9.8 Hz, H-1, 1H), 3.74 (m, H-4, 1H), 3.69 (d, *J* = 5.2 Hz, H-5, 1H), 3.58 (dd, *J* = 9.5 Hz, H-3, 1H), 3.47 (t, *J* = 9.6 Hz, H-2, 1H), 2.25 (s, $CH_3C_6H_4S$, 3H), 1.15 (d, *J* = 6.5 Hz, $CH_3C_5H_9O_5$, 3H); ^{13}C NMR (101 MHz, D_2O) δ 131.76, 129.94, 88.21, 74.95, 74.10, 71.42, 68.92, 61.48, 27.80, 20.09, 15.68.

References

1 B. G. Davis and A. J. Fairbanks, *Carbohydrate chemistry*, Oxford University Press, Oxford, 2002.
2 J. F. Robyt, *Essentials of Carbohydrate Chemistry*, Springer Science & Business Media, 1998.
3 W. Z. Hassid, *Science*, 1969, **165**, 137–144.
4 P. Y. Bruice, *Organic chemistry*, Upper Saddle River N.J. : Prentice Hall c2001, 3rd edn.
5 M. Miljkovic, *Carbohydrates: Synthesis, Mechanisms, and Stereoelectronic Effects*, Springer Science & Business Media, 2009.
6 M. Sinnott, *Carbohydrate Chemistry and Biochemistry: Structure and Mechanism*, Royal Society of Chemistry, 2007.
7 A. Varki, *Essentials of glycobiology*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2009.
8 S. C. Zeeman, J. Kossmann and A. M. Smith, *Annu. Rev. Plant Biol.*, 2010, **61**, 209–234.
9 T. Cox, in *eLS*, John Wiley & Sons, Ltd, 2001.
10 A. Bahaji, J. Li, Á. M. Sánchez-López, E. Baroja-Fernández, F. J. Muñoz, M. Ovecka, G. Almagro, M. Montero, I. Ezquer, E. Etxeberria and J. Pozueta-Romero, *Biotechnol. Adv.*, 2014, **32**, 87–106.
11 R. A. Meyers, *Molecular Biology and Biotechnology: A Comprehensive Desk Reference*, John Wiley & Sons, 1995.
12 D. J. Vocadlo, *ChemBioChem*, 2007, **8**, 1886–1887.
13 U. Ellervik, H. Grundberg and G. Magnusson, *J. Org. Chem.*, 1998, **63**, 9323–9338.
14 D. Crich, *Acc. Chem. Res.*, 2010, **43**, 1144–1153.
15 F. Andersson, P. Fúgedi, P. J. Garegg and M. Nashed, *Tetrahedron Lett.*, 1986, **27**, 3919–3922.
16 R. R. Schmidt, *Angew. Chem. Int. Ed. Engl.*, 1986, **25**, 212–235.
17 J. D. C. Codée, R. E. J. N. Litjens, L. J. van den Bos, H. S. Overkleeft and G. A. van der Marel, *Chem. Soc. Rev.*, 2005, **34**, 769–782.
18 M. Jacobsson, J. Malmberg and U. Ellervik, *Carbohydr. Res.*, 2006, **341**, 1266–1281.
19 F. Koichi, H. Atsushi and K. Shoichi, *Tetrahedron Lett.*, 1993, **34**, 2187–2190.
20 P. Fügedi, P. J. Garegg, H. Lönn and T. Norberg, *Glycoconj. J.*, 1987, **4**, 97–108.
21 V. Martichonok and G. M. Whitesides, *J. Org. Chem.*, 1996, **61**, 1702–1706.
22 A. V. Demchenko, N. N. Malysheva and C. De Meo, *Org. Lett.*, 2003, **5**, 455–458.
23 D. Mootoo, P. Konradsson, U. Udodong and B. Fraserreid, *J. Am. Chem. Soc.*, 1988, **110**, 5583–5584.
24 B. Fraser-Reid, G. Anilkumar, M. R. Gilbert, S. Joshi, R. Kraehmer, B. Ernst, G. W. Hart and P. Sinaý, in *Carbohydrates in Chemistry and Biology*, Wiley-VCH Verlag GmbH, 2000, pp. 135–154.
25 J. C. Lopez and B. Fraser-Reid, *J. Chem. Soc. Chem. Commun.*, 1991, 159–161.
26 M. N. Kamat, C. De Meo and A. V. Demchenko, *J. Org. Chem.*, 2007, **72**, 6947–6955.
27 R. M. Hickey, *Int. Dairy J.*, 2012, **22**, 141–146.
28 S. Thurl, M. Munzert, J. Henker, G. Boehm, B. Müller-Werner, J. Jelinek and B. Stahl, *Br. J. Nutr.*, 2010, **104**, 1261–1271.

29 J. B. German, S. L. Freeman, C. B. Lebrilla and D. A. Mills, in *Personalized Nutrition for the Diverse Needs of Infants and Children*, eds. D. M. Bier, J. B. German and B. Lonnerdal, Karger, Basel, 2008, vol. 62, pp. 205–222.
30 C. Kunz and S. Rudloff, *Int. Dairy J.*, 2006, **16**, 1341–1346.
31 N. S. Han, T.-J. Kim, Y.-C. Park, J. Kim and J.-H. Seo, *Biotechnol. Adv.*, 2012, **30**, 1268–1278.
32 L. Bode, *Glycobiology*, 2012, **22**, 1147–1162.
33 T. Eiwegger, B. Stahl, P. Haidl, J. Schmitt, G. Boehm, E. Dehlink, R. Urbanek and Z. Szepfalusi, *Pediatr. Allergy Immunol.*, 2010, **21**, 1179–1188.
34 A. Lucas, R. Morley, T. J. Cole, P. J. Lucas, S. M. Gore, P. Crowle, R. Pearse, A. J. Boon and R. Powell, *The Lancet*, 1990, **335**, 1477–1481.
35 L. Bode and E. Jantscher-Krenn, *Adv. Nutr. Int. Rev. J.*, 2012, **3**, 383S–391S.
36 A. Lucas, R. Morley, T. Cole, G. Lister and C. Leesonpayne, *Lancet*, 1992, **339**, 261–264.
37 D. S. Newburg, *J. Mammary Gland Biol. Neoplasia*, 1996, **1**, 271–283.
38 C. Kunz, S. Rudloff, W. Baier, N. Klein and S. Strobel, *Annu. Rev. Nutr.*, 2000, **20**, 699–722.
39 M. Roberfroid, G. R. Gibson, L. Hoyles, A. L. McCartney, R. Rastall, I. Rowland, D. Wolvers, B. Watzl, H. Szajewska, B. Stahl, F. Guarner, F. Respondek, K. Whelan, V. Coxam, M.-J. Davicco, L. Léotoing, Y. Wittrant, N. M. Delzenne, P. D. Cani, A. M. Neyrinck and A. Meheust, *Br. J. Nutr.*, 2010, **104**, S1–S63.
40 A. L. Morrow, G. M. Ruiz-Palacios, X. Jiang and D. S. Newburg, *J. Nutr.*, 2005, **135**, 1304–1307.
41 P. Chaturvedi, C. D. Warren, C. R. Buescher, L. K. Pickering and D. S. Newburg, *Bioact. Compon. Hum. Milk*, 2001, **501**, 315–323.
42 A. Rijnierse, P. V. Jeurink, B. C. A. M. van Esch, J. Garssen and L. M. J. Knippels, *Eur. J. Pharmacol.*, 2011, **668, Supplement 1**, S117–S123.
43 G. Flamm, W. Glinsmann, D. Kritchevsky, L. Prosky and M. Roberfroid, *Crit. Rev. Food Sci. Nutr.*, 2001, **41**, 353–362.
44 S. Macfarlane, G. T. Macfarlane and J. H. Cummings, *Aliment. Pharmacol. Ther.*, 2006, **24**, 701–714.
45 C. Martínez-Villaluenga, J. Frías, C. Vidal-Valverde and R. Gómez, *J. Food Prot.*, 2005, **68**, 1246–1252.
46 A. A. Aachary and S. G. Prapulla, *Compr. Rev. Food Sci. Food Saf.*, 2011, **10**, 2–16.
47 P. Gyorgy, R. Norris and C. Rose, *Arch. Biochem. Biophys.*, 1994, **48**, 193–01.
48 A. Bezkorovainy, D. Grohlich and J. H. Nichols, *Am. J. Clin. Nutr.*, 1979, **32**, 1428–1432.
49 D. Barile and R. A. Rastall, *Curr. Opin. Biotechnol.*, 2013, **24**, 214–219.
50 V. Sangwan, S. K. Tomar, R. R. B. Singh, A. K. Singh and B. Ali, *J. Food Sci.*, 2011, **76**, R103–R111.
51 K. Shoaf, G. L. Mulvey, G. D. Armstrong and R. W. Hutkins, *Infect. Immun.*, 2006, **74**, 6920–6928.
52 A. Gosling, G. W. Stevens, A. R. Barber, S. E. Kentish and S. L. Gras, *Food Chem.*, 2010, **121**, 307–318.
53 T. Sako, K. Matsumoto and R. Tanaka, *Int. Dairy J.*, 1999, **9**, 69–80.
54 Y. W. Park, M. Juárez, M. Ramos and G. F. W. Haenlein, *Small Rumin. Res.*, 2007, **68**, 88–113.

55 D. P. M. Torres, M. P. Goncalves, J. A. Teixeira and L. R. Rodrigues, *Compr. Rev. Food Sci. Food Saf.*, 2010, **9**, 438–454.
56 P. F. Fox and L. Stepaniak, *Int. Dairy J.*, 1993, **3**, 509–530.
57 M. I. G. Siso, *Bioresour. Technol.*, 1996, **57**, 1–11.
58 R. K. Singh, S. S. Nielsen and J. V. Chambers, *J. Food Sci.*, 1991, **56**, 935–937.
59 R. R. Mahoney, *Food Chem.*, 1998, **63**, 147–154.
60 D. L. Zechel and S. G. Withers, *Acc. Chem. Res.*, 2000, **33**, 11–18.
61 A. Vasella, G. J. Davies and M. Böhm, *Curr. Opin. Chem. Biol.*, 2002, **6**, 619–629.
62 J. E. Prenosil, E. Stuker and J. R. Bourne, *Biotechnol. Bioeng.*, 1987, **30**, 1019–1025.
63 C. A. Hsu, S. L. Lee and C. C. Chou, *J. Agric. Food Chem.*, 2007, **55**, 2225–2230.
64 V. Dumortier, C. Brassart and S. Bouquelet, *Biotechnol. Appl. Biochem.*, 1994, **19**, 341–354.
65 F. Depeint, G. Tzortzis, J. Vulevic, K. I'Anson and G. R. Gibson, *Am. J. Clin. Nutr.*, 2008, **87**, 785–791.
66 A. Annunziata and R. Vecchio, *Food Qual. Prefer.*, 2013, **28**, 348–355.
67 B. Splechtna, T. H. Nguyen, M. Steinbock, K. D. Kulbe, W. Lorenz and D. Haltrich, *J. Agric. Food Chem.*, 2006, **54**, 4999–5006.
68 A. M. Bakker-Zierikzee, M. S. Alles, J. Knol, F. J. Kok, J. J. M. Tolboom and J. G. Bindels, *Br. J. Nutr.*, 2005, **94**, 783–790.
69 C. Corradini, A. Cavazza and C. Bignardi, *Int. J. Carbohydr. Chem.*, 2012, **2012**, e487564.
70 P. Urrutia, B. Rodriguez-Colinas, L. Fernandez-Arrojo, A. O. Ballesteros, L. Wilson, A. Illanes and F. J. Plou, *J. Agric. Food Chem.*, 2013, **61**, 1081–1087.
71. Personal Communication with Prof. J. P. Kamerling
72 C. R. Noller and W. C. Rockwell, *J. Am. Chem. Soc.*, 1938, **60**, 2076–2077.
73 Z. Zhang, I. R. Ollmann, X.-S. Ye, R. Wischnat, T. Baasov and C.-H. Wong, *J. Am. Chem. Soc.*, 1999, **121**, 734–753.
74 F. Dasgupta and L. Anderson, *Carbohydr. Res.*, 1994, **264**, 155–160.
75 H. H. Jensen, L. U. Nordstrøm and M. Bols, *J. Am. Chem. Soc.*, 2004, **126**, 9205–9213.
76 P. J. Garegg and H. Hultberg, *Carbohydr. Res.*, 1981, **93**, C10–C11.
77 S. D. Debenham and E. J. Toone, *Tetrahedron Asymmetry*, 2000, **11**, 385–387.
78 M. Emmadi and S. S. Kulkarni, *Org. Biomol. Chem.*, 2013, **11**, 3098–3102.
79 K. M. Halkes, H. J. Vermeer, T. M. Slaghek, P. A. . van Hooft, A. Loof, J. P. Kamerling and J. F. . Vliegenthart, *Carbohydr. Res.*, 1998, **309**, 175–188.
80 A. Varki and S. Diaz, *Anal. Biochem.*, 1984, **137**, 236–247.
81 M. Mori, Y. Ito and T. Ogawa, *Carbohydr. Res.*, 1989, **192**, 131–146.
82 T. M. Slaghek, Y. Nakahara, T. Ogawa, J. P. Kamerling and J. F. G. Vliegenthart, *Carbohydr. Res.*, 1994, **255**, 61–85.
83 P. J. Garegg, T. Iversen and S. Oscarson, *Carbohydr. Res.*, 1976, **50**, C12–C14.
84 H. Dong, Z. Pei, S. Byström and O. Ramström, *J. Org. Chem.*, 2007, **72**, 1499–1502.

85 Y. Zhou, J. Li, Y. Zhan, Z. Pei and H. Dong, *Tetrahedron*, 2013, **69**, 2693–2700.
86 N. Touisni, A. Maresca, P. C. McDonald, Y. Lou, A. Scozzafava, S. Dedhar, J.-Y. Winum and C. T. Supuran, *J. Med. Chem.*, 2011, **54**, 8271–8277.
87 N. L. Douglas, S. V. Ley, U. Lücking and S. L. Warriner, *J. Chem. Soc. [Perkin 1]*, 1998, 51–66.
88 X. Wu, J.-R. Mariño-Albernas, F.-I. Auzanneau, V. Verez-Bencomo and B. M. Pinto, *Carbohydr. Res.*, 1998, **306**, 493–503.
89 P. A. Levene and G. M. Meyer, *J. Biol. Chem.*, 1931, **92**, 257–262.
90 T. Sugimoto, T. Fujii, Y. Hatanaka, S. Yamamura and M. Ueda, *Tetrahedron Lett.*, 2002, **43**, 6529–6532.
91 J. Xia, T. Srikrishnan, J. L. Alderfer, R. K. Jain, C. F. Piskorz and K. L. Matta, *Carbohydr. Res.*, 2000, **329**, 561–577.
92 D. Sail and P. Kováč, *Carbohydr. Res.*, 2012, **357**, 47–52.
93 Y.-H. Wang, H.-W. Yeh, H.-W. Wang, C.-C. Yu, J.-H. Guh, D.-Z. Liu and P.-H. Liang, *Carbohydr. Res.*, 2013, **375**, 118–135.
94 R. Rodebaugh and B. Fraser-Reid, *Tetrahedron*, 1996, **52**, 7663–7678.
95. Personal communication with Prof. S. Oscarson
96 S. Oscarson and H. Ritzén, *Carbohydr. Res.*, 1994, **254**, 81–90.
97 S. M. Andersen, C.-C. Ling, P. Zhang, K. Townson, H. J. Willison and D. R. Bundle, *Org. Biomol. Chem.*, 2004, **2**, 1199–1212.
98 B. N. A. Mbadugha and F. M. Menger, *Org. Lett.*, 2003, **5**, 4041–4044.
99 A. Meijer and U. Ellervik, *J. Org. Chem.*, 2004, **69**, 6249–6256.
100 R. G. Powell, *J. Nat. Prod.*, 2009, **72**, 516–523.
101 V. Thoss, *British Bluebells: The potential of using a protected species as a provider of fine chemicals to enhance its conservation*, Nova, 2011.
102. Personal communication Prof. R. J. Nash
103 G. E. Blackman and A. J. Rutter, *J. Ecol.*, 1954, **42**, 629–638.
104 D. D. Kohn, P. E. Hulme, P. M. Hollingsworth and A. Butler, *Biol. Conserv.*, 2009, **142**, 61–74.
105 P. Thompson and S. Cox, *Ann. Bot.*, 1978, **42**, 51–62.
106 V. Thoss, P. J. Murphy, R. Marriott and T. Wilson, *RSC Adv.*, 2012, **2**, 5314–5322.
107 P. Grabham and J. Packham, *Biol. Conserv.*, 1983, **26**, 105–126.
108 C. T. Calam, *Process Development in Antibiotic Fermentations*, Cambridge University Press, Cambridge, 1987.
109 M. Wink, *Phytochemical Diversity of Secondary Metabolites*, 2004.
110 R. J. Molyneux, S. T. Lee, D. R. Gardner, K. E. Panter and L. F. James, *Phytochemistry,* 2007, **68**, 2973–2985.
111 A. A. Watson, R. J. Nash, M. R. Wormald, D. J. Harvey, S. Dealler, E. Lees, N. Asano, H. Kizu, A. Kato, R. C. Griffiths, A. J. Cairns and G. W. J. Fleet, *Phytochemistry*, 1997, **46**, 255–259.
112 A. Welter, J. Jadot, G. Dardenne, M. Marlier and J. Casimir, *Phytochemistry*, 1976, **15**, 747–749.
113 T. Yamashita, K. Yasuda, H. Kizu, Y. Kameda, A. A. Watson, R. J. Nash, G. W. J. Fleet and N. Asano, *J. Nat. Prod.*, 2002, **65**, 1875–1881.
114 N. Asano, R. J. Nash, R. J. Molyneux and G. W. J. Fleet, 2000, *Phytochemistry,* **11**, 1645–1680.
115 E. Borges de Melo, A. da Silveira Gomes and I. Carvalho, *Tetrahedron*, 2006, **62**, 10277–10302.

116 H. R. Mellor, A. Adam, F. M. Platt, R. A. Dwek and T. D. Butters, *Anal. Biochem.,* 2000, **284**, 136–142.
117 P. Greimel, J. Spreitz, A. E. Stutz and T. M. Wrodnigg, *Curr Top Med Chem*, 2003, **3**, 513–523.
118 T. M. Wrodnigg, *Monatsh. Chem.,* 2002, **133**, 393–426.
119 T. M. Wrodnigg, A. J. Steiner and B. J. Ueberbacher, *Anti-Cancer Agents Med. Chem.*, 2008, **8**, 77–85.
120 M. S. J. Simmonds, W. M. Blaney and L. E. Fellows, *Chem. Ecol.,* 1990, **16**, 3167–3196.
121 M. J. Egan, E. A. Porter, G. C. Kite, M. S. J. Simmonds, J. Barker and S. Howells, *Rapid Commun. Mass Spectrom.*, 1999, **13**, 195–200.
122 A. A. Watson, G. W. J. Fleet, N. Asano, R. J. Molyneux and R. J. Nash, *Phytochemistry*, 2001, **56**, 265–295.
123 C. Y. Yu, N. Asano, K. Ikeda, M. X. Wang, T. D. Butters, M. R. Wormald, R. A. Dwek, A. L. Winters, R. J. Nash and G. W. J. Fleet, *Chem. Commun.,* 2004, 1936–1937.
124 A. Kato, I. Adachi, M. Miyauchi, K. Ikeda, T. Komae, H. Kizu, Y. Kameda, A. A. Watson, R. J. Nash, M. R. Wormald, G. W. J. Fleet and N. Asano, *Carbohydr. Res.*, 1999, **316**, 95–103.
125 N. Asano, R. J. Nash, R. J. Molyneux and G. W. J. Fleet, *J. Nat. Prod.,* 2000, **11**, 1645–1680.
126 A. Kato, N. Kato, S. Miyauchi, Y. Minoshima, I. Adachi, K. Ikeda, N. Asano, A. A. Watson and R. J. Nash, *Phytochemistry,* 2008, **69**, 1261–1265.
127 C. E. Harland, *Ion Exchange: Theory and Practice*, Royal Society of Chemistry, 1994.
128 L. Nelligan, MSc thesis, Bangor University, 2011.
129 D. A. Dougherty, *J. Nutr.*, 2007, **137**, 1504S–1508S; discussion 1516S–1517S.
130 Y. Izumi, I. Chibata and T. Itoh, *Angew. Chem. Int. Ed. Engl.*, 1978, **17**, 176–183.
131 B. G. Davis and A. J. Fairbanks, *Carbohydrate chemistry*, Oxford University Press, Oxford, 2002.
132 N. Asano, *Glycobiology*, 2003, **13**, 93R–104R.
133 G. Lauc, I. Rudan, H. Campbell and P. M. Rudd, *Mol Biosyst.*, 2010, **6**, 329–335.
134 E. S. H. El Ashry, E. S. H. El Tamany, M. E. D. A. El Fattah, A. T. A. Boraei and H. M. Abd El-Nabi, *Eur. J. Med. Chem.*, 2013, **66**, 106–113.
135 G. M. Cooper, *The Cell*, Sinauer Associates, 2nd edn., 2000.
136 G. Karp, *Cell and Molecular Biology: Concepts and Experiments*, John Wiley & Sons, 2009.
137 F. M. Platt, *Nature*, 2014, **510**, 68–75.
138 E. B. Vitner, F. M. Platt and A. H. Futerman, *J. Biol. Chem.*, 2010, **285**, 20423–20427.
139 G. F. Hoffmann, J. Zschocke and W. L. Nyhan, *Inherited Metabolic Diseases: A Clinical Approach*, Springer Science & Business Media, 2009.
140 N. W. Barton, R. O. Brady, J. M. Dambrosia, A. M. Di Bisceglie, S. H. Doppelt, S. C. Hill, H. J. Mankin, G. J. Murray, R. I. Parker, C. E. Argoff, R. P. Grewal and K.-T. Yu, *N. Engl. J. Med.*, 1991, **324**, 1464–1470.
141 C. E. Hollak, S. van Weely, M. H. van Oers and J. M. Aerts, *J. Clin. Invest.*, 1994, **93**, 1288–1292.

142 E. Beutler, *PLoS Med.*, 2004, **1**.
143 A. Zimran, A. Kay, T. Gelbart, P. Garver, D. Thurston, A. Saven and E. Beutler, *Medicine (Baltimore)*, 1992, **71**, 337–353.
144 S. Mohindroo, *Cases J.*, 2009, **2**, 9380.
145 D. P. Germain, *Orphanet J. Rare Dis.*, 2010, **5**, 30.
146 K. D. MacDermot, A. Holmes and A. H. Miners, *J. Med. Genet.*, 2001, **38**, 769–775.
147 A. Mehta, M. Beck, F. Eyskens, C. Feliciani, I. Kantola, U. Ramaswami, A. Rolfs, A. Rivera, S. Waldek and D. P. Germain, *QJM*, 2010, **103**, 641–659.
148 G. Felisati, E. Salvatici, C. Pipolo, S. Portaleone, E. Riva and M. Giovannini, *J. Med. Case Reports*, 2012, **6**, 112.
149 A. T. van der Ploeg and A. J. Reuser, *The Lancet*, 2008, **372**, 1342–1353.
150 N. W. Barton, F. S. Furbish, G. J. Murray, M. Garfield and R. O. Brady, *Proc. Natl. Acad. Sci.*, 1990, **87**, 1913–1916.
151 G. Altarescu, D. Elstein and M. Beck, *Fabry Disease*, Springer Science & Business Media, 2010.
152 A. E. Mulberg, D. Murphy, J. Dunne and L. L. Mathis, *Pediatric Drug Development*, John Wiley & Sons, 2013.
153 A. H. Futerman and A. Zimran, *Gaucher Disease*, CRC Press, 2006.
154 F. M. Platt, G. R. Neises, G. B. Karlsson, R. A. Dwek and T. D. Butters, *J. Biol. Chem.*, 1994, **269**, 27108–27114.
155 L. A. Sorbera, J. Castañer and M. Bayés, *Drugs Future*, 2003, **28**, 229.
156 N. Asano, R. J. Nash, R. J. Molyneux and G. W. J. Fleet, *Tetrahedron: Asymmetry*, 2000, **11**, 1645–1680.
157 A. R. Sawkar, W.-C. Cheng, E. Beutler, C.-H. Wong, W. E. Balch and J. W. Kelly, *Proc. Natl. Acad. Sci.*, 2002, **99**, 15428–15433.
158 V. K. J. J. P. Sharma, *Comprehensive Objective Biology*, Golden Bells.
159 T. D. Butters, H. R. Mellor, K. Narita, R. A. Dwek and F. M. Platt, *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 2003, **358**, 927–945.
160 N. Asano, R. J. Nash, R. J. Molyneux and G. W. J. Fleet, *Tetrahedron: Asymmetry*, 2000, **11**, 1645–1680.
161 H. M. Sauro, *Enzyme Kinetics for Systems Biology*, Future Skill Software, 2012.
162 A. R. Sawkar, S. L. Adamski-Werner, W.-C. Cheng, C.-H. Wong, E. Beutler, K.-P. Zimmer and J. W. Kelly, *Chem. Biol.*, 2005, **12**, 1235–1244.
163 P. H. Scudder, *Electron Flow in Organic Chemistry: A Decision-Based Guide to Organic Mechanisms*, John Wiley & Sons, 2013.
164 I. A. Dotsenko, M. Curtis, N. M. Samoshina and V. V. Samoshin, *Tetrahedron*, 2011, **67**, 7470–7478.
165 B. Brazdova, N. S. Tan, N. M. Samoshina and V. V. Samoshin, *Carbohydr. Res.*, 2009, **344**, 311–321.
166 S. Howard, C. Braun, J. McCarter, K. W. Moremen, Y. F. Liao and S. G. Withers, *Biochem. Biophys. Res. Commun.*, 1997, **238**, 896–898.
167 G. W. Burton and K. U. Ingold, *J. Am. Chem. Soc.*, 1981, **103**, 6472–6477.
168 M. Valoti, H. J. Sipe Jr., G. Sgaragli and R. P. Mason, *Arch. Biochem. Biophys.*, 1989, **269**, 423–432.
169 F. Ponticelli, A. Trendafilova, M. Valoti, S. Saponara and G. Sgaragli, *Carbohydr. Res.*, 2001, **330**, 459–468.
170 A. V. Samoshin, I. A. Dotsenko, N. M. Samoshina, A. H. Franz and V. V. Samoshin, *Int. J. Carbohydr. Chem.*, 2014, **2014**, e941059.

171 C.-Y. Huang, D. A. Thayer, A. Y. Chang, M. D. Best, J. Hoffmann, S. Head and C.-H. Wong, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 15–20.
172 R. M. Hyland, T. P. Griener, G. L. Mulvey, P. I. Kitov, O. P. Srivastava, P. Marcato and G. D. Armstrong, *J. Med. Microbiol.*, 2006, **55**, 669–675.
173 T. M. Wrodnigg and A. E. Stutz, *Curr. Enzyme Inhib.*, 2012, **8**, 47–99.
174 A. A. Watson, R. J. Nash, M. R. Wormald, D. J. Harvey, S. Dealler, E. Lees, N. Asano, H. Kizu, A. Kato, R. C. Griffiths, A. J. Cairns and G. W. J. Fleet, *Phytochemistry*, 1997, **46**, 255–259.
175 B. Elchert, J. Li, J. Wang, Y. Hui, R. Rai, R. Ptak, P. Ward, J. Y. Takemoto, M. Bensaci and C.-W. T. Chang, *J. Org. Chem.*, 2004, **69**, 1513–1523.
176 Y. A. Lin, J. M. Chalker and B. G. Davis, *J. Am. Chem. Soc.*, 2010, **132**, 16805–16811.
177 D. Horton and H. S. Prihar, *Carbohydr. Res.*, 1967, **4**, 115–125.
178 M. S. Motawia, C. E. Olsen, K. Enevoldsen, J. Marcussen and B. L. Møller, *Carbohydr. Res.*, 1995, **277**, 109–123.
179 C. Wang, Q. Li, H. Wang, L.-H. Zhang and X.-S. Ye, *Tetrahedron*, 2006, **62**, 11657–11662.
180 J. Wang, J. Li, H.-N. Chen, H. Chang, C. T. Tanifum, H.-H. Liu, P. G. Czyryca and C.-W. T. Chang, *J Med Chem*, 2005, **48**, 6271–6285.