CHEMICAL SYNTHESIS AND BIOLOGICAL EVALUATION OF CHONDROITIN SULFATE DISACCHARIDES

POH ZHONG WEI

(B. Sc. (Hons.), NUS)

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

NUS GRADUATE SCHOOL FOR INTEGRATIVE SCIENCES AND ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2015

Declaration Page

I hereby declare that this thesis is my original work and it has been written by me in its entirety.

I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Poh Zhong Wei

01/07/2015

Acknowledgements

I would like to express my sincere gratitude to my supervisors A/P Lam Yulin (NUS Chemistry) and A/P George Yip (NUS Anatomy) for the invaluable guidance and insight they have given throughout my PhD study. Thank you for the patience and encouragement given, the positive lab environment greatly nurtured scientific inquiry. I would also like to thank NGS for providing a generous scholarship, and my TAC committee members A/P Chan Woon Khiong and A/P Tong Yen Wah for their support and helpful advice. Special mention also goes to Prof. Phillip K. Moore and the NGS support staff, for their unwavering service and dedication.

In addition, I would like to thank the members of the lab: Dr. Wong Lingkai, Dr. Woen Susanto, Dr. Sanjay Samanta, Hadhi Wijaya, Ang Wei Jie, Ng Cheng Yang, Eric Lee Jing Xiang, Alex Gan Chin Heng, Niu Zilu, Dr. Zhang Ting, Mario Octavianus Ihsan, Dr. Sen Yin Ping, Sharen Lim See Wee, Guo Suxian, Saritha Adepu and Dr. Tan Xing Fei for the interesting scientific discussions and technical assistance rendered, this has been greatly appreciated. I am also grateful to the research staff of the CMMAC facilities: Dr. Wu Ji'En, Mdm Han Yanhui, Mdm Wong Lai Kwai and Dr. Liu Qiping for their help in the NMR characterization and mass analysis of the synthesized compounds.

Last but not least, I would like to thank my family and friends for their unconditional love, support and understanding throughout my PhD study, without which this journey would have been impossible. Thank you so much!

TABLE OF CONTENTS

ABSTRACT

This thesis describes the development of a synthetic strategy to obtain all the sulfation pattern isomers possible in the repeating unit of chondroitin sulfate (CS). The strategy incorporates orthogonal protecting groups in four glycosyl donor and four glycosyl acceptor building blocks which are synthesized in a divergent mode using 2 common intermediates. Through the judicious choice of donor and acceptor building blocks, any sulfation pattern required in the final CS disaccharide can be obtained by modular glycosylation and selective transformations. Specifically, any sulfation pattern required in the final product is obtained by regioselective sulfation of either the benzyl ether protected hydroxyl groups or the ester protected hydroxyl groups to furnish all 16 theoretically plausible CS disaccharides; these include analogues currently available as well as novel sulfation motifs.

To determine if the sulfate groups present in CS encode important functional information for the regulation of physiological processes such as cancer progression, biological evaluation of all 16 CS disaccharides was next conducted on breast cancer cells. Cell viability assay results indicated that CS sulfation patterns had differing effects for different breast cancer cell types. The greatest inhibitory effect was observed for the most aggressive, triple negative breast cancer cell line MDA-MB-231, while low grade breast cancer cells (MCF-7 and T47D) were not affected. Apoptosis assays further indicated that some of these CS disaccharides could induce apoptosis in MDA-MB-231 cells.

LIST OF FIGURES

LIST OF SCHEMES

LIST OF TABLES

LIST OF ABBREVIATIONS

δ	Chemical shift
Ac	Acetyl
Å	Angstrom
AcOH	Acetic acid
ABCN	1,1'-Azobis(cyclohexanecarbonitrile)
AIBN	Azobisisobutyronitrile
ANOVA	One-way analysis of variance
ATCC	American type culture collection
BAIB	Bisacetoxyiodobenzene
Bn	Benzyl
Bu	Butyl
Bz	Benzoyl
*CHST*3	Chondroitin sulfotransferase 3
*CHST*11	Chondroitin sulfotransferase 11
Conc.	Concentrated
DABCO	1,4-Diazabicyclo[2.2.2]octane
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DCC	N,N'-Dicyclohexylcarbodiimide
CAN	Ceric ammonium nitrate
$CDCl_3$	Deuterated chloroform
ClAc	Chloroacetyl
CSA	Camphorsulfonic acid

CS	Chondroitin sulfate
CSPG	Chondroitin sulfate proteoglycan
d	Doublet
DCC	N,N'-Dicyclohexylcarbodiimide
DCIS	Ductal carcinoma in situ
DCM	Dichloromethane
DDQ	2,3-Dichloro-5,6-dicyano-1,4-benzoquinone
DMAP	4-Dimethylaminopyridine
DMEM	Dulbecco's Modified Eagle's Medium
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DS	Dermatan sulfate
ER	Estrogen receptor
ESI	Electron spray ionization
Et	Ethyl
EtCN	Propionitrile
EtOAc	Ethyl acetate
Et_2O	diethyl ether
EtOH	Ethanol
Equiv.	Equivalent
FBS	Fetal bovine serum
GAG	Glycosaminoglycan
GalNAc	N-acetyl galactosamine

GlcA	Glucuronic acid
h	Hour
HA	Hyaluronic acid
HCl	Hydrochloric acid
HER2	Human epidermal receptor 2
Hex	Hexane
HRMS	High resolution mass spectroscopy
HS	Heparan sulfate
LCIS	Lobular carcinoma in *situ*
Lev	Levulinyl
LiOH	Lithium hydroxide
m	Multiplet
Me	Methyl
MeCN	Acetonitrile
MeOD	Deuterated methanol
MeOH	Methanol
MS	Molecular sieves
mw	microwave
m/z	Mass-to-charge ratio
mmol	Millimole
NaOH	Sodium hydroxide
NaOAc	Sodium acetate
Nap	Naphthyl

NIS	*N*-iodosuccinimide
NMR	Nuclear magnetic resonance
PBS	Phosphate-buffered saline
PMB	Para-methoxy benzyl
PMP	Para-methoxy phenyl
Pd/C	Palladium on carbon
PG	Proteoglycan
Ph	Phenyl
PR	Progesterone receptor
q	Quartet
rpm	Revolutions per minute
r.t.	Room temperature
RPMI	Roswell Park Memorial Institute medium
s	Singlet
SD	Standard deviation
t	triplet
TBAI	Tetrabutylammonium iodide
TBDPS	*tert*-Butyldiphenylsilyl
TBDMS	*tert*-Butyldimethylsilyl
TCAHN	Trichloroacetamide
TEA	Triethylamine
TES	Triethylsilane
TFA	Trifluoroacetic acid

TEMPO	(2,2,6,6-Tetramethylpiperidin-1-yl)oxyl
TfOH	Trifluoromethanesulfonic acid
THF	Tetrahydrofuran
TLC	Thin-layer chromatography
TMA	Trimethylamine
TMS	Trimethylsilyl
TMSOTf	Trimethylsilyl trifluoromethanesulfonate
TNBC	Triple negative breast cancer
TsOH	Tosylic acid
UV	Ultraviolet

LIST OF PUBLICATIONS

1) Poh, Z. W., Gan, C. H., Lee, E. J., Guo, S., Yip, G. W., & Lam, Y.

Divergent Synthesis of Chondroitin Sulfate Disaccharides and Identification of Sulfate Motifs that Inhibit Triple Negative Breast Cancer.

Scientific Reports **5**, 14355 (2015).

2) Guo, S., Poh, Z. W., Adepu, S., Lam, Y. & Yip, G. W.

Chondroitin Sulfate Sulfation Patterns Regulate Triple Negative Breast Cancer.

(Pending submission)

CONFERENCE PRESENTATIONS

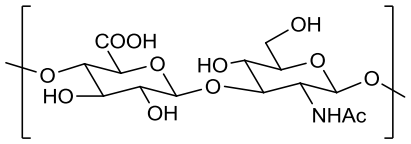
1) Poh, Z. W., Gan, C. H., Yip, G. W., & Lam, Y. "Divergent Synthesis and Biological Evaluation of Chondroitin Sulfate Disaccharides", poster presentation, *Gordon Research Conferences*: *Carbohydrates* (2015), Vermont, United States

2) Poh, Z. W., Guo, S., Yip, G. W., & Lam, Y. "Glycosaminoglycans in Cancer Biology", poster presentation, *19th European Symposium of Organic Chemistry* (2015), Lisbon, Portugal

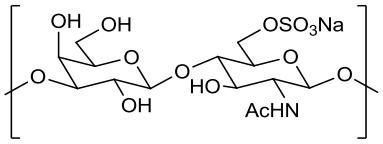
CHAPTER 1: INTRODUCTION

1.1 Introduction to Glycosaminoglycans

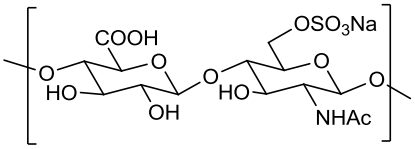
Glycosaminoglycans (GAGs) are a distinct class of carbohydrates widely distributed in the human body; these complex polysaccharides consist of long, unbranched and heterogeneous polysaccharide chains which play diverse roles in many key physiological processes.[1-3] GAGs are ubiquitous in the extracellular matrix, and can be covalently attached to core proteins to form proteoglycan side chains, or are located within intracellular secretory granules.[4-6] GAG chains have a common feature: they are constituted by core disaccharide repeat units which consist of both the uronic acid and the amino sugar; these are hexose derivatives of common monosaccharides.[7] The uronic acid has the C-6 alcohol group oxidized to form a carboxylic acid, while amino sugar has the C-2 hydroxyl group substituted with an amino group.



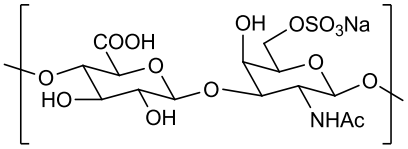
Hyaluronic acid (HA)



Keratan sulfate (KS)



Heparan sulfate (HS)



Chondroitin sulfate (CS)

Figure 1-1. Classes of glycosaminoglycans

Four main classes of GAGs have been identified, and they can be classified based on the structures of their repeat units[8] (Figure 1-1). These include hyaluronic acid (HA), keratan sulfate (KS), heparan sulfate (HS) and chondroitin sulfate (CS). In each GAG polysaccharide, the monosaccharide units are connected together by glycosidic bonds, which may either be of α- or β- stereochemical configuration. This linkage occurs between the C-1 (anomeric) hydroxyl group of the first unit and the C-3 or C-4 position of the adjacent sugar unit. GAGs are highly polar and negatively charged macromolecules, and usually contain sulfate groups attached at specific sites.[7]

Hyaluronic acid (HA) is only class of GAG polysaccharides which have no sites of sulfation. It is one of the principle components of the extracellular matrix, where it functions as a shock absorber and also a lubricant.[9] The repeating unit of HA consists of D-glucuronic acid and D-*N*-acetyl glucosamine. In addition, HA is a major component of the skin epidermis, and play crucial roles during the wound healing process.[10] Similarly, keratan sulfate (KS) can be found in the extracellular matrix, although it is concentrated mainly in the cornea, cartilage and the central nervous system.[11, 12] KS GAGs consists of *N*-acetyl glucosamine and D-galactose repeating units, and play key roles in several degenerative disorders such as macular corneal dystrophy and osteoarticular diseases.[13] Unlike other GAGs, KS contains D-galactose sugars in the repeat unit, these are non-uronic acid moieties.

Much emphasis has been placed on the study of heparan sulfate (HS) GAGs with the discovery of heparin as an injectable anticoagulant drug.[14, 15] Known as the biomolecule with the highest negative charge density, heparin is structurally similar to HS and produced exclusively in mast cell secretory granules.[16-18] As compared to heparin, HS GAGs are generally longer in length and have fewer sulfation sites. The major sequences of heparan sulfate consist of D-glucuronic acid and D-*N*-acetyl glucosamine repeat units, and HS GAGs play critical roles in the regulation of signaling processes important for growth development, immune system response and oncogenesis.[19-21]

1.2 Chondroitin Sulfate Glycosaminoglycans

Although relatively less studied than heparan sulfate, chondroitin sulfate (CS) polysaccharides are the most prevalent GAGs in the body, consisting of D-glucuronic acid and D-*N*-acetyl galactosamine repeat units. CS GAGs form important structural components of the human connective tissues in cartilage and bone joints.[22] In addition, CS GAGs can be covalently bonded to cell membrane core proteins to form chondroitin sulfate proteoglycans (CSPG); these CS side chains are attached to the protein scaffold via a serine residue, through a tetrasaccharide linkage.[23] CS GAGs have been demonstrated to play major supportive roles in developmental signaling in neural stem cells,[24] and are effective for the treatment of osteoporosis and osteoarthritis.[25]

During the biosynthesis of the CS GAGs, modifications such as epimerization and sulfation can occur, which give rise to structural diversity of GAG compounds. For instance, CS may be converted to dermatan sulfate (DS) GAGs by DS epimerases (Figure 1-2).[26] In this case, the orientation of the C-6 carboxylate group in CS is switched from equatorial to axial, and the glucuronic acid moiety becomes modified into an epimer, the L-iduronic acid. The L-iduronic acid and *N*-acetyl galactosamine pair constitutes the repeat unit found in DS GAGs.

Figure 1-2. Epimerization of chondroitin sulfate

In addition, sulfate groups may be attached to specific sites in the CS sequence by various sulfotransferase enzymes,[27] such as chondroitin 4-sulfotransferase and chondroitin 6-sulfotransferase, which attach sulfate groups to the C-4' and C-6' position of *N*-acetylgalactosamine respectively. These modifications create structural diversity of CS GAGs which enable specific binding interactions with different ligands.[28-31] The myriad of signaling effects generated thus allows for the modulation of many different cellular events and physiological processes, which include morphogenesis, viral invasion, cancer metastasis, spinal cord injury and stem cell proliferation.[32-41] For example, it has been demonstrated that CS GAGs can interact with key bioactive growth factors and neurotrophic factors implicated in various neuro-generative functions, via distinct sulfated sequences

embedded within the CS GAGs[42-44]. Specifically, it was shown that the CS-E sulfation motif can act as molecular recognition elements for growth factors which promote neuronal growth.[45, 46] This indicates that the sulfation patterns on CS chains do not occur randomly, but a "sulfation code" exists to mediate the physiological functions of these sequences.[47, 48] Important functional information can thus be encoded in CS GAGs via the attachment of position-specific sulfate groups.

There are 4 possible sites of sulfation in the CS repeat unit. Sulfate groups may be attached on the C-2, C-3 positions of glucuronic acid or the C-4', C-6' positions of *N*-acetyl galactosamine (Figure 1-3). A total of 16 sulfation patterns are thus theoretically possible; some of these CS sulfation patterns have already been identified and given conventional nomenclature.[26, 49-53]

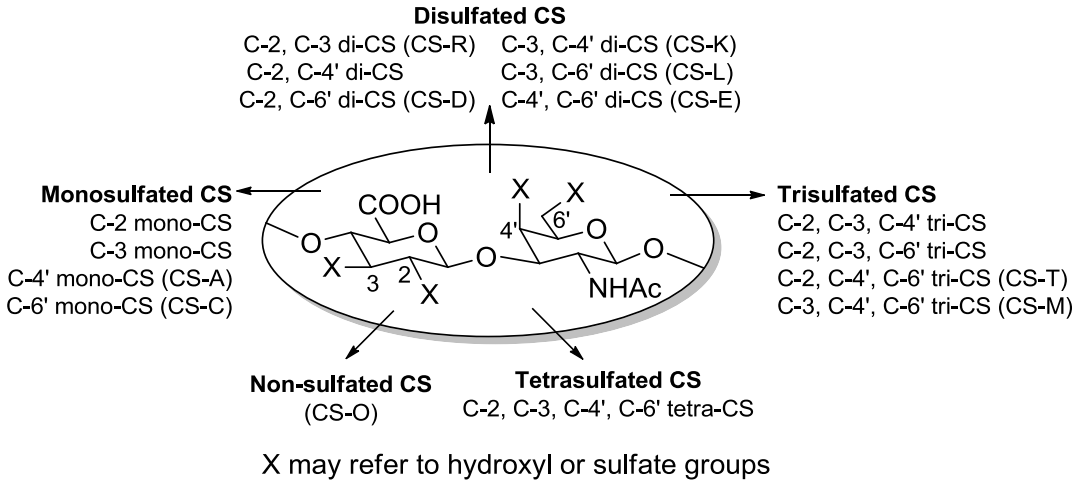


Figure 1-3. Sulfation sites in CS

The commonly occurring forms include a single sulfation on C-4' and C-6' of the *N*-acetyl galactosamine unit, known as CS-A and CS-C respectively. Over sulfation occurs when both C-4' and C-6' are sulfated on the *N*-acetyl galactosamine moiety to form CS-E, or when sulfation occurs on D-glucuronic acid. Slight changes in the sulfation pattern can have a major impact on biological processes. Most commonly occurring forms CS-A and CS-C have been well studied for their cell-cell recognition properties[54], osteoarthritis[55] and inhibition of complement factor Clq[56], while over-sulfated CS-D and CS-E have been proven to promote neurite outgrowth toward embryonic rat mesencephalic and hippocampal neurons.[57] Although sulfation on the C-2 and C-3 positions of glucuronic acid is relatively less common, some of these CS isomers such as CS-D, CS-L and CS-M have also been isolated.[58, 59]

1.3 Sulfation Code in Chondroitin Sulfate

To investigate the molecular level interactions of CS, techniques have to be employed to obtain pure and well-defined sulfated motifs to determine how CS sulfation patterns affect biological activity. However, procurement of pure, homogenous CS motifs which are not commonly expressed is a challenging task due to isolation difficulty and structural complexity. Current approaches to probe the "sulfation code" include genetic modifications which involve knocking out sulfotransferase genes to probe phenotypic development[60-62] or knocking down sulfotransferase genes to observe the effects on tumor cell proliferation and adhesiveness.[63] These methods are confined by the scope of the sulfotransferase

gene sequences available, which can result in the generation of a limited range of CS isomers. Moreover, the genetically modified CS sequences have diverse and heterogeneous sulfation profiles, thus it becomes difficult to ascertain the biological effects of any particular sulfation motif.[64, 65]

Biochemical methods of acquiring CS from natural sources also face challenges in purification and separation, and often result in mixtures of differently sulfated polysaccharides with poor linear definition due to the structural complexity and heterogeneity of GAGs.[66, 67] Since unique CS sulfation patterns have differing therapeutic potentials for different disease states[68], the possible confounding and contradictory effects exerted by different CS isomers could hamper systematic study.[69]

To circumvent limitations faced by current methods, chemical synthesis emerges as a viable alternative to prepare CS sequences via control of the site(s) of sulfation required in the final product. Through this approach, pure CS sequences with homogenous sulfation profiles can be obtained, to systematically probe the effect of CS sulfation patterns on various physiological processes. To achieve this, the organic chemist must first devise a synthetic strategy to obtain CS GAGs via chemical transformation of monosaccharides to form derivatized glycosides. Commonly known as carbohydrate chemistry, this sub-field of organic chemistry deals with the synthesis and chemical reactions for the preparation of natural and unnatural carbohydrate compounds.

1.4 Chemical Synthesis of Carbohydrates

Carbohydrates are the most abundant class of biomolecules found in nature, and can be broadly defined as polyhydroxy aldehydes and ketones with the general formula $C_n(H_2O)_n$. Research in carbohydrate chemistry began in the nineteenth century, pioneered by Nobel laureate Emil Fischer who first elucidated the basic structure and stereochemistry of monosaccharides. He created the *Fischer projection* system to depict the spatial orientation of the various hydroxyl groups present in monosaccharides, through a two-dimensional representation.[70] General interest in synthetic carbohydrate chemistry however developed only later in the 1960s, primarily motivated by the identification of biologically active carbohydrate compounds isolated from natural products and micro-organisms, such as unusual sugar fragments in antibiotics. The initial discovery of streptomycin as an effective antibiotic for the treatment of *mycobacterium tuberculosis* infections[71] sparked the interest in aminoglycoside drugs; these compounds consisted of an aminocyclitol fragment connected to one or two carbohydrate units.

By the 1990s, the boom in modern reagent methodology and asymmetric synthesis further encouraged the use of monosaccharides as chiral building blocks for the synthesis of complex glycosides.[72] However, despite the advancements in synthetic carbohydrate chemistry, this area of research remains a challenging field for organic chemists, due to the structural complexity of carbohydrates. The chemical synthesis of saccharides has been known to be a tedious and arduous

process, involving more than 40 steps of linear synthesis to obtain a single oligosaccharide.[73] Hans Paulson succinctly summarized this in 1982, which largely still holds true till date: “Each oligosaccharide synthesis remains an independent problem which resolution requires considerable systematic research and a good deal of know-how. There are no universal reaction conditions for oligosaccharide synthesis”.[74]

1.5 Protecting Group Chemistry

In carbohydrate sugars, the presence of multiple hydroxyl groups with similar chemical reactivities and other reactive functional groups such as carboxyl and amino groups have posed many challenges to the synthesis. Protecting group chemistry thus became a cornerstone in carbohydrate research, and the concept of orthogonal protecting groups was developed, to protect neighbouring functional groups from undesired side reactions during the course of synthesis. This involved selective cleavage of individual protecting groups under different reaction conditions to release specific groups for reaction and modification. Protection was essential as neighbouring functional groups would otherwise undergo similar reactions to form unwanted side products.

In any synthetic strategy, the type of protecting group used must be carefully chosen to ensure reaction compatibility. The protecting groups used must be robust to the reaction conditions to keep the functionalities sufficiently protected, but be labile to removal when reaching specific points during synthesis,

preferably under facile conditions. In addition, regioselectivity is required during protection to discern between different hydroxyl groups present in the carbohydrate backbone to enable functionalization of specific positions in the monosaccharide to obtain the target compound. When glycosidic bond formation is needed, only the anomeric C-1 hydroxyl group should be exposed for attachment to the neighbouring fragment, with the other functionalities suitably protected. Numerous hydroxyl protecting groups have been developed in the past few decades, and are now commonly used by organic chemists.[75-78] These orthogonal protecting groups can be cleaved under different reaction conditions, greatly benefiting the synthesis of complex carbohydrate scaffolds.[79-82]

1.5.1 Hydroxyl protecting groups: esters

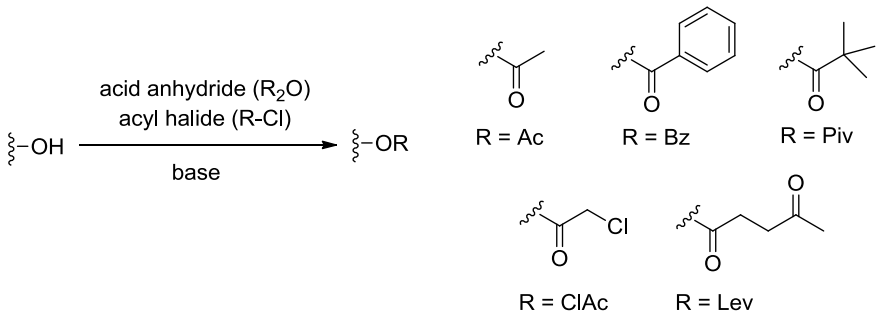


Figure 1-4. Protection of hydroxyl groups as esters

Esters are the most conventional modes of protection utilised for hydroxyl groups (Figure 1-4), and can be cleanly cleaved via basic hydrolysis, using catalytic sodium methoxide in methanol. Acetates (Ac) are generally more labile than benzoates (Bz), while pivaloyl esters (Piv) are the most stable to hydrolysis. Some

esters can also be cleaved via other orthogonal methods. These include the chloroacetates (ClAc), which can be selectively cleaved using DABCO as the dechloroacetylation reagent[83], and the *levulinoyl* ester (Lev), which can be selectively cleaved using hydrazine buffered with acetic acid.[84]

1.5.2 Hydroxyl protecting groups: ethers

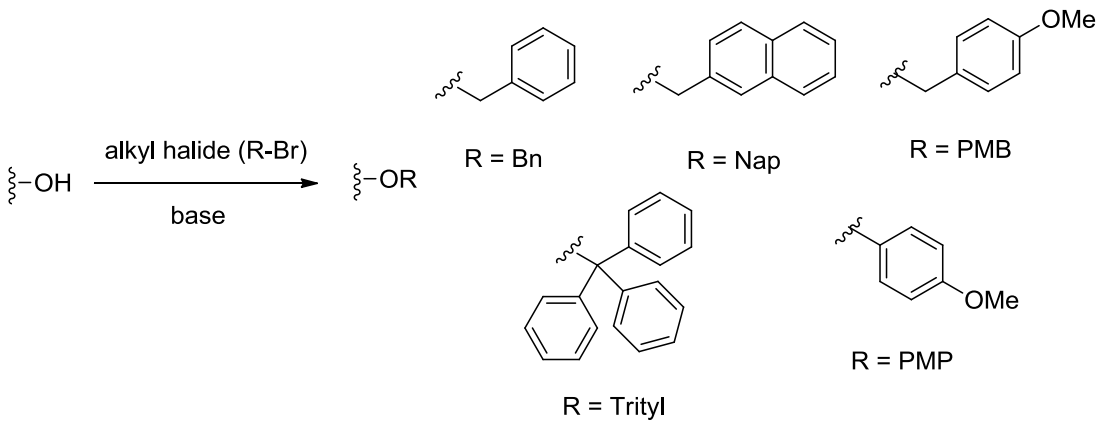


Figure 1-5. Protection of hydroxyl groups as ethers

Ethers are also a popular choice of protection for hydroxyl groups (Figure 1-5), as these groups are cleaved under reaction conditions orthogonal to that of esters. For the benzyl ether (Bn), cleavage occurs via catalytic hydrogenation, in the presence of palladium on charcoal (10% Pd/C) as the catalyst. On the other hand, the naphthyl ether (Nap), p-methoxybenzyl ether (PMB) and p-methoxylphenyl ethers (PMP) can all be cleaved via oxidative conditions, using either ceric ammonium nitrate (CAN) or 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) as the oxidant[76]; these do not affect the benzyl ethers.

For regioselective protection of the primary C-6 alcohol, the sterically bulky trityl group can be employed; this trityl group can be cleaved via acidic hydrolysis or hydrogenolysis when required.

1.5.3 Hydroxyl protecting groups: silyl ethers

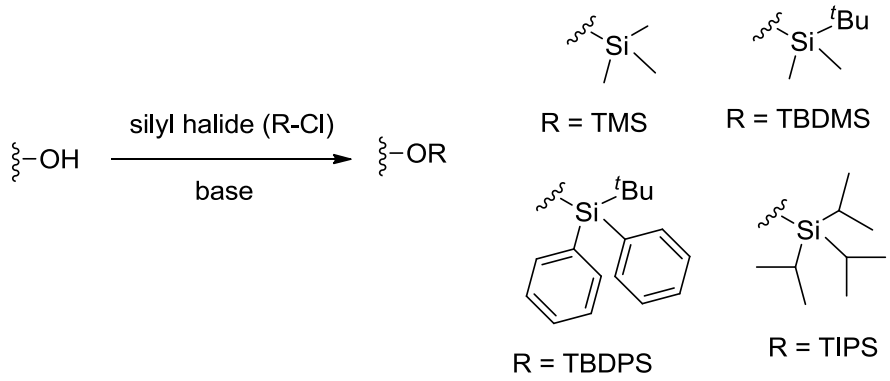


Figure 1-6. Protection of hydroxyl groups as silyl ethers

Silyl ether protecting groups are introduced under mild conditions by reacting the hydroxyl functionally with silyl chlorides in the presence of bases such as imidazole or pyridine (Figure 1-6). These protecting groups can be cleaved with acid, or with fluoride ions such tetrabutylammonium fluoride. The order of acid stability is proportional to the steric size of the silyl ether, and increases in the following order: TMS < TBDMS < TIPS < TBDPS. Regioselective silylation of the primary C-6 hydroxyl group is similarly possible by reacting pyranoses with bulky *tert*-butyldiphenylsilyl chloride in pyridine.[85]

1.5.4 Hydroxyl protecting groups: cyclic acetals

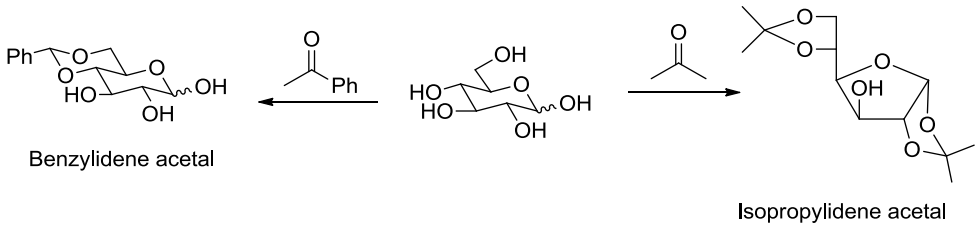


Figure 1-7. Protection of hydroxyl groups as cyclic acetals

Acetal protecting groups are also commonly used in carbohydrate chemistry, and can be removed via acidic hydrolysis to regenerate the hydroxyl group. To direct regioselective protection of a pair of hydroxyl groups in monosaccharides, cyclic acetals are commonly employed, which are more stable than acyclic variants (Figure 1-7). Isopropylidene acetals enable simultaneous protection of *cis*-positioned 1,2-diols groups in the sugar unit to obtain the furanose derivative. Selective cleavage of the C-5,C-6 isopropylidene acetal is possible using a weaker acid, as both isopropylidene acetals have different stabilities.

Alternatively, benzylidene acetals can be used for protection of both the C-4 and C-6 hydroxyl groups (1,3-diols) to form the stable bicylic ring system in pyranoses. This benzylidene acetal can be regioselectively opened up to generate an aryl ether and a free hydroxyl group at either C-4 or C-6 position, by varying the reaction conditions.[86-88]

1.6 Glycosylation

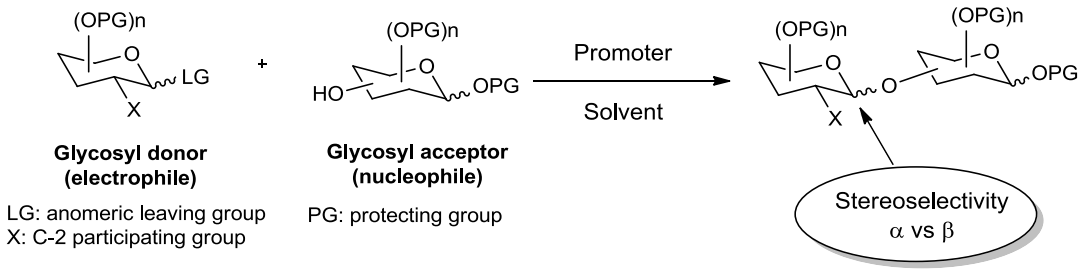


Figure 1-8. Glycosylation between a glycosyl donor and acceptor

To synthesize oligosaccharides, glycosylation is required: this involves connecting 2 suitably protected saccharide units to form a longer fragment, generating the glycosidic linkage at anomeric C-1 carbon of the glycosyl donor (Figure 1-8). For glycosylation to proceed smoothly, a suitably protected glycosyl donor with a leaving group (LG) attached to the anomeric positon is required, which is replaced by the hydroxyl group of the acceptor during glycosylation. The glycosidic linkage formed can either be 1,2-*cis* or 1,2-*trans* (alternatively classified as α- or β-configuration). Depending on the target molecule, synthetic methods have to be devised to obtain the required stereochemistry so as to probe the biological effects conferred by the carbohydrate motif in chiral living organisms.

Many types of glycosyl donors have been devised and optimized for the synthesis of different oligosaccharides (Figure 1-9).[89] To activate the glycosyl donor, various glycosyl auxiliaries can be attached to the anomeric C-1 carbon.

Examples include substituents like halides, thioglycosides, sulfoxides, selenoglycosides, and the trichloroacetimidate auxiliary, which is most commonly used. These substituents are activated by Lewis acids such as trimethylsilyl trifluoromethanesulfonate (TMSOTf), boron trifluoride diethyl etherate ($BF_3.OEt_2$), silver triflate and triflic acid/*N*- iodosuccinimide.

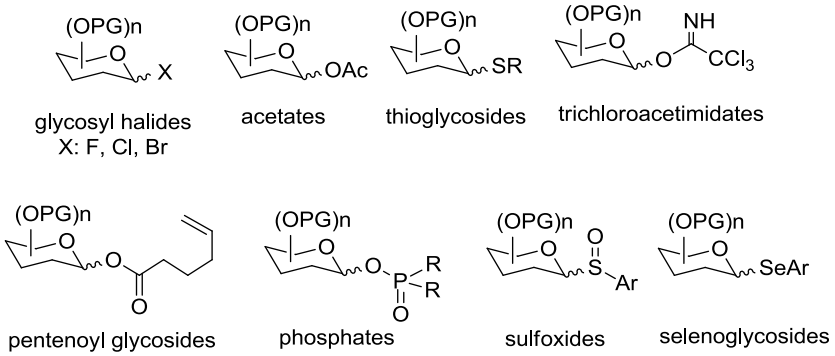


Figure 1-9. Examples of different glycosyl donors

1.6.1 Stereoselectivity of Glycosylation

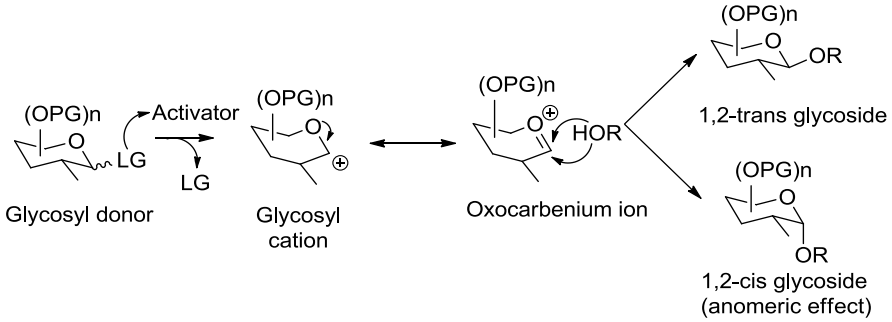


Figure 1-10. Glycosylation reaction mechanism

During glycosylation, nucleophilic displacement occurs at the anomeric carbon, involving the attack of a weak nucleophile (the glycosyl acceptor) on the

secondary carbon center via a unimolecular S_N1 mechanism.[90] When promoted by an activator, the glycosyl auxiliary departs as a stable leaving group which generates the reactive oxocarbenium ion, facilitating attack by the hydroxyl group of the glycosyl acceptor (Figure 1-10). The acceptor can either attack from the top or bottom of the plane, to form a mixture of 1,2-*cis* and 1,2-*trans* glycosides. Although the 1,2-*cis* product is thermodynamically favoured due to stabilization by the anomeric effect, 1,2-*trans* product is also kinetically formed due to irreversible glycosidic bond formation. Both the α and β anomers are difficult to separate by chromatographic techniques as they have very similar physical characteristics; formation of the undesired anomer not only contaminates the product, but also reduces yield of the glycosylation step. Hence there is a need for stereochemical control during glycosylation to maximize yield of the required anomer. Different methods have been developed to direct the stereoselectivity of glycosylation, these include the use of C-2 participating groups and reaction solvent effects.

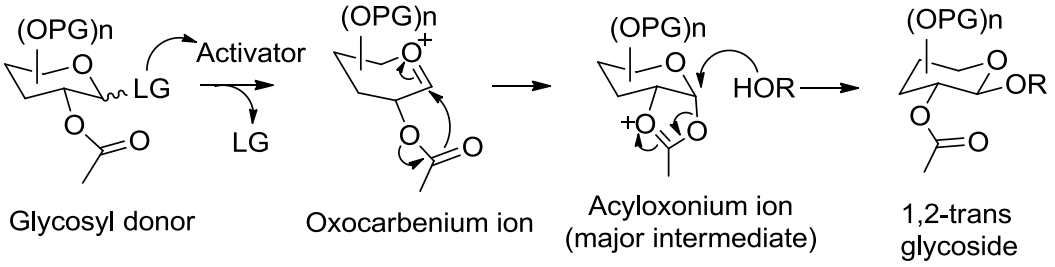


Figure 1-11: Synthesis of 1,2-*trans* glycosides by neighbouring group participation

To direct 1,2-*trans* stereoselective glycosylation, neighbouring group participation can be employed by using protecting groups such as esters and amides on the C-2 position of the glycosyl donor[90] (Figure 1-11). This anchimeric effect occurs when the electrons from the C-2 participating group delocalizes into the oxocarbenium ion ring to stabilize the positive charge. The C-2 O-acyl group forms a stable 1,2-*cis* fused ring system with the sugar unit via axial attack from the bottom face to generate the acyloxonium ion. This exposes the top plane to attack by the glycosyl acceptor which generates 1,2-*trans* glycosides in high stereoselectivity, with the new group attached on the equatorial position.

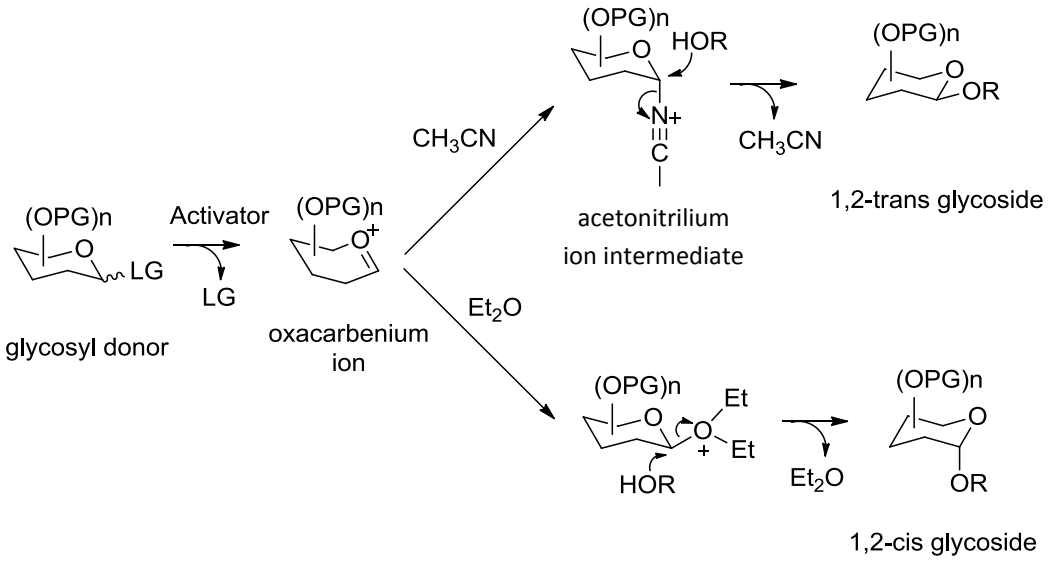


Figure 1-12. Solvent participation for stereoselective glycosylation

In the absence of C-2 participating group, solvent effects may also be utilised to direct stereoselectivity at the anomeric position[91] (Figure 1-12). Participating

solvents such as acetonitrile and diethyl ether can be used to direct the formation of *β*-glucosides and *α*-glucosides respectively. When glycosylations are carried out in acetonitrile, an acetronitrilium ion intermediate is formed *in situ*, which adopts the axial configuration.[92-94] Subsequent displacement of the acetonitrile group via S_N2 attack by the acceptor generates the glycosidic linkage at the equatorial position to obtain the 1,2-*trans* glycoside, even in the absence of the C-2 participating group. Conversely, diethyl ether forms the oxonium ion at the equatorial position due to electronic stabilization via anomeric effects, thus the acceptor attacks from the axial position to form the 1,2-*cis* glycoside.[95] Alternatively, pre-activation of thioglycosyl donors with DMF prior to the addition of the promoter has also been shown to direct α-stereoselective glycosylation[96].

In naturally occurring CS polysaccharides, the glycosidic bonds have *β* stereoselectivity. Hence, the protection strategy designed should enable stereoselective formation of the desired 1,2-*trans* glycosidic linkage, preferably through the use of neighbouring group participation on C-2 of the donor unit.

1.7 Chemoenzymatic Synthesis of Heparin Sulfate Construct

Due to the structural complexity of glycosaminoglycans, commercially available heparin drugs such as Arixtra are highly costly, requiring more than 50 steps to obtain via chemical synthesis.[97] The importance of anti-coagulant drugs has prompted the development of new methods to complement existing methods. To

improve overall yields, chemoenzymatic approaches were devised, these methods mimic biosynthesis, by utilizing heparan sulfate biosynthetic enzymes such as HS polymerases, epimerases and sulfotransferases.

Recently, Xu *et al.* demonstrated the chemoenzymatic synthesis of ultralow molecular weight heparin construct **1-4** (Figure 1-13).[98] Core disaccharide unit **1-1** was utilised as the starting material, which was readily available from heparosan, obtainable via fermentation. Through the use of two bacterial glycosyl transferases (*N*-acetyl glucosaminyl transferase of Escherichia coli K5 (KfiA) and heparosan synthase-2 (pmHS2) from Pasteurella multocida), tetrasaccharide **1-2** was prepared by attaching the GlcNTFA and GlcUA units to disaccharide **1-1**. Similarly, **1-2** was further extended by attaching 3 additional monosaccharides to obtain heptasaccharide **1-3**.

Subsequent epimerization of gluA to IdoA (residue D, **1-4**) was accompanied by 2-*O*-sulfation, 6-*O*-sulfation and 3-*O*-sulfation (residue C, **1-4**) to obtain heparin construct **1-3**, ultilizing 3'-phosphoadenosine-5'-phosphosulfate (PAPS) as the sulfo group donor. It was noted that selective epimerization/2-*O*-sulfation of residue D (instead of residue B) and 3-*O*-sulfaton of residue C (instead of residue A and E) required careful structural control of the substrate, which was dependent on the polysaccharide sequence. To obtain *N*-sulfo groups in **1-4**, trifluoroacetyl groups were attached in **1-3**, which were subsequently converted to *N*-sulfate groups.

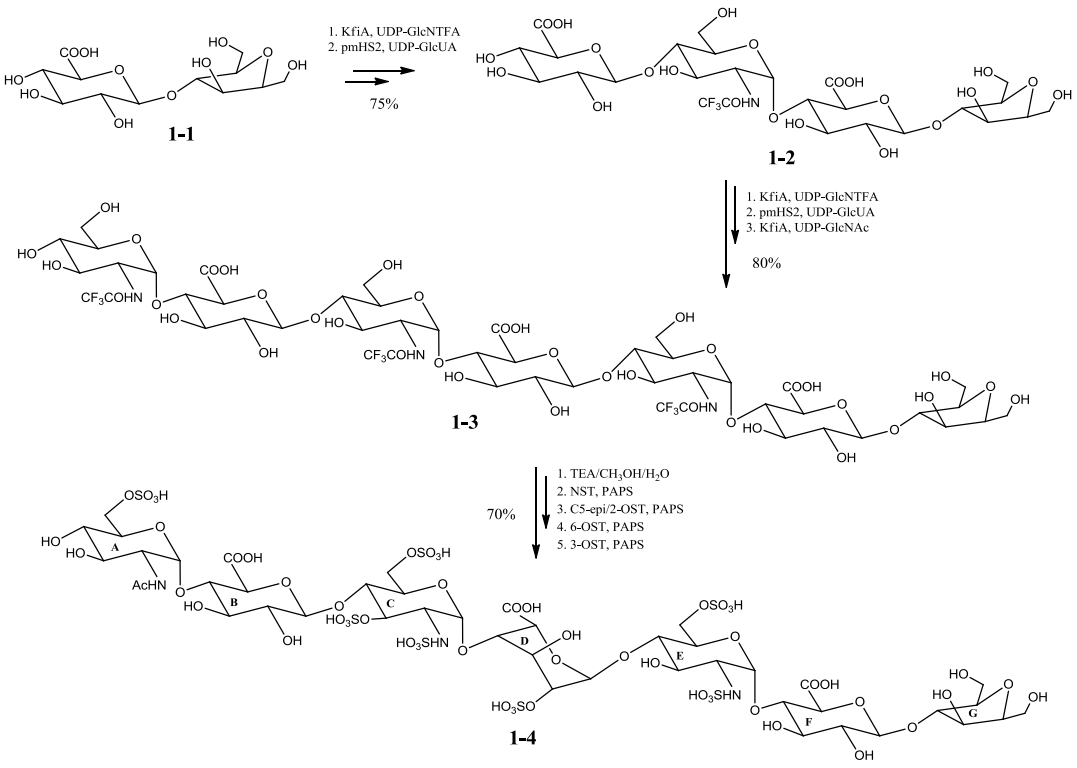
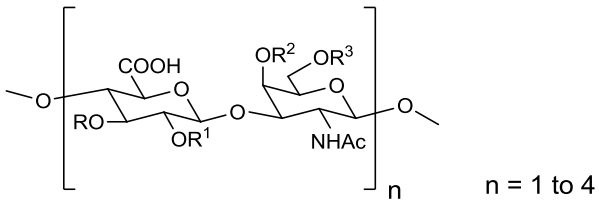


Figure 1-13. Chemoenzymatic synthesis of heparin sulfate construct

Unlike conventional chemical synthesis[99], the need to introduce and remove protecting groups could be averted through the use of chemoenzymatic methods. In addition, stereoselectivity of the glycosylation step could be readily controlled by using appropriate enzymes, without the need for glycosyl mediators to effect stereochemical outcome. This reduces the number of synthetic steps required, which improves overall reaction yields. However, a major limitation in chemoenzymatic methods is high substrate specificity; the constructs have to be carefully selected to ensure compatibility with enzymes used, and also to prevent the formation of other side products. In fact, to ensure compatibility with this chemoenzymatic approach, the heparin constructs synthesized via this method

were not the actual structure of Arixtra, but modified variants. Chemoenzymatic methods may thus not be viable for synthesis of a library of analogues.

1.8 Chemical Synthesis of Chondroitin Sulfate



CS-A: R = R^1 = R^3 = H; R^2 = OSO_3H **CS-C**: R = R^1 = R^2 = H; R^3 = OSO_3H
CS-D: R = R^2 = H; R^1 = R^3 = OSO_3H **CS-E**: R = R^1= H; R^2 = R^3 = OSO_3H
CS-K: R = R^3 = H; R = R^2 = OSO_3H **CS-L**: R^1 = R^2 = OH; R = R^3 = OSO_3H
CS-M: R^1 = H; R = R^2 = R^3 = OSO_3H **CS-R**: R^2 = R^3 = OH; R = R^1 = OSO_3H

Figure 1-14. CS sulfate motifs obtained via chemical synthesis

Although relatively less studied than heparan sulfate GAGs, notable work has been achieved by various groups to synthesize CS GAGs of varying chain length[45, 46, 51, 52, 100-108], with a primary focus on commonly occurring CS sulfation patterns such as CS-A, CS-C and CS-E (Figure 1-14). Much effort has been expended on the synthesis of longer CS sequences, including tetra-, hexa- and octa- saccharides, and some of these analogues have been studied for their biological effects.[45, 109, 110] The synthetic precursors used usually involve derivatized glucose and galactose monosaccharides, furnished with appropriate protecting groups to allow for chain elongation and site specific sulfation required in the final compounds. Typically, these protecting groups are manipulated after glycosylation, which involve regioselective deprotection and sulfation to obtain

the required CS analogue. Recently, some rare CS isomers have also been obtained through an alternative approach; these include CS-K, CS-L and CS-M.[52, 111] This method employs a different synthetic precursor: non-sulfated CS disaccharides derived from the acidic hydrolysis of commercially available CS polymers,[112, 113] which are further modified and protected to enable preparation of additional CS sulfation motifs.

However, based on the current synthetic strategies reported, not all the sulfation patterns possible in the CS repeating unit can be obtained, particularly CS sequences where the sulfate groups occur on the C-2 or C-3 position of the glucuronic acid moiety. We opined that these CS sulfate motifs that are not commonly expressed could encode important regulatory information for various physiological processes, but glycobiological research has been impeded due to the difficulty in acquiring these CS sequences. To address this, there is therefore a need to design new methods to facilitate access to all 16 unique CS sulfation patterns, for the evaluation of their potential biological effects.

1.9 Objective of Research Work

Thus, the aim of this project is to devise a synthetic strategy which would allow for the chemical synthesis of all 16 possible sulfation patterns present in the CS repeating unit. Ideally, the synthetic route should utilize common intermediates to obtain all 16 isomers via a divergent approach, providing access to novel CS sulfation motifs not hitherto reported. The library of synthesized compounds empowers us to systematically evaluate the effect of CS sulfation patterns on important biological events. In this project, some preliminary biological evaluation will be conducted on different cancer cell types, with a focus on breast cancer cells.

1.10 References

1. Theocharis, A.D., Skandalis, S.S., Tzanakakis, G.N. & Karamanos, N.K. Proteoglycans in health and disease: novel roles for proteoglycans in malignancy and their pharmacological targeting. *FEBS Journal* **277**, 3904-3923 (2010).
2. Sugahara, K. *et al.* Recent advances in the structural biology of chondroitin sulfate and dermatan sulfate. *Curr. Opin. Struct. Biol.* **13**, 612-620 (2003).
3. Maeda, N., Ishii, M., Nishimura, K. & Kamimura, K. Functions of chondroitin sulfate and heparan sulfate in the developing brain. *Neurochem. Res.* **36**, 1228-1240 (2011).
4. Kjellen, L. & Lindahl, U. Proteoglycans: structures and interactions. *Annu. Rev. Biochem.* **60**, 443-475 (1991).
5. Iozzo, R.V. Matrix proteoglycans: From molecular design to cellular function. *Annu. Rev. Biochem.* **67**, 609-652 (1998).
6. Kolset, S.O., Prydz, K. & Pejler, G. Intracellular proteoglycans. *Biochem. J.* **379**, 217-227 (2004).
7. Varki, A., Cummings, R.D., Esko, J.D. & Freeze, H.H. *Essentials of Glycobiology*, Edn. 2nd. (Cold Spring Harbor Laboratory Press, New York; **2009**).
8. Yip, G.W., Smollich, M. & Gotte, M. Therapeutic value of glycosaminoglycans in cancer. *Mol. Cancer Ther.* **5**, 2139-2148 (2006).

9. Papakonstantinou, E., Roth, M. & Karakiulakis, G. Hyaluronic acid: A key molecule in skin aging. *Dermatoendocrinol.* **4**, 253-258 (2012).

10. Kogan, G., Šoltés, L., Stern, R. & Gemeiner, P. Hyaluronic acid: a natural biopolymer with a broad range of biomedical and industrial applications. *Biotechnol. Lett.* **29**, 17-25 (2007).

11. Funderburgh, J.L., Caterson, B. & Conrad, G.W. Distribution of proteoglycans antigenically related to corneal keratan sulfate proteoglycan. *J. Biol. Chem.* **262**, 11634-11640 (1987).

12. Hopwood, J.J. & Robinson, H.C. The structure and composition of cartilage keratan sulphate. *Biochem. J.* **141**, 517-526 (1974).

13. Klintworth, G.K. *et al.* Macular corneal dystrophy. Lack of keratan sulfate in serum and cornea. *Ophthalmic Paediatr. Genet.* **7**, 139-143 (1986).

14. Gandhi, N.S. & Mancera, R.L. Heparin/heparan sulphate-based drugs. *Drug Discov. Today* **15**, 1058-1069 (2010).

15. Lever, R. & Page, C.P. Novel drug development opportunities for heparin. *Nat. Rev. Drug Discov.* **1**, 140-148 (2002).

16. Zhang, L. Glycosaminoglycan (GAG) Biosynthesis and GAG-Binding Proteins. *Prog. Mol. Biol. Transl. Sci.* **93**, 1-17 (2010).

17. Humphries, D.E. *et al.* Heparin is essential for the storage of specific granule proteases in mast cells. *Nature* **400**, 769-772 (1999).

18. Forsberg, E. *et al.* Abnormal mast cells in mice deficient in a heparin-synthesizing enzyme. *Nature* **400**, 773-776 (1999).

19. Bao, X. *et al.* Endothelial Heparan Sulfate Controls Chemokine Presentation in Recruitment of Lymphocytes and Dendritic Cells to Lymph Nodes. *Immunity* **33**, 817-829.

20. Sasisekharan, R., Shriver, Z., Venkataraman, G. & Narayanasami, U. Roles of heparan-sulphate glycosaminoglycans in cancer. *Nat. Rev. Cancer* **2**, 521-528 (2002).

21. Lin, X. Functions of heparan sulfate proteoglycans in cell signaling during development. *Development* **131**, 6009-6021 (2004).

22. Watanabe, H., Yamada, Y. & Kimata, K. Roles of Aggrecan, a Large Chondroitin Sulfate Proteoglycan, in Cartilage Structure and Function. *J. Biochem.* **124**, 687-693 (1998).

23. Silbert, J.E. & Sugumaran, G. Biosynthesis of Chondroitin/Dermatan Sulfate. *IUBMB Life* **54**, 177-186 (2002).

24. Bandtlow, C.E. & Zimmermann, D.R. Proteoglycans in the developing brain: new conceptual insights for old proteins. *Physiol. Rev.* **80**, 1267-1290 (2000).

25. Clegg, D.O. *et al.* Glucosamine, Chondroitin Sulfate, and the Two in Combination for Painful Knee Osteoarthritis. *N.Engl. J. Med.* **354**, 795-808 (2006).

26. Akatsu, C. *et al.* Dermatan sulfate epimerase 2 is the predominant isozyme in the formation of the chondroitin sulfate/dermatan sulfate hybrid structure in postnatal developing mouse brain. *Glycobiology* **21**, 565-574 (2011).

27. Kusche-Gullberg, M. & Kjellén, L. Sulfotransferases in glycosaminoglycan biosynthesis. *Curr. Opin. Struct. Biol.* **13**, 605-611 (2003).

28. Bernfield, M. *et al.* Functions of cell surface heparan sulfate proteoglycans. *Annu. Rev. Biochem.* **68**, 729-777 (1999).

29. Harris, E.N. & Weigel, P.H. The ligand-binding profile of HARE: hyaluronan and chondroitin sulfates A, C, and D bind to overlapping sites distinct from the sites for heparin, acetylated low-density lipoprotein, dermatan sulfate, and CS-E. *Glycobiology* **18**, 638-648 (2008).

30. Rogerson, S.J., Chaiyaroj, S.C., Ng, K., Reeder, J.C. & Brown, G.V. Chondroitin sulfate A is a cell surface receptor for Plasmodium falciparum-infected erythrocytes. *J. Exp. Med.* **182**, 15-20 (1995).

31. Kawashima, H. *et al.* Oversulfated chondroitin/dermatan sulfates containing GlcAbeta1/IdoAalpha1-3GalNAc(4,6-O-disulfate) interact with L- and P-selectin and chemokines. *J. Biol. Chem.* **277**, 12921-12930 (2002).

32. Sugahara, K. & Mikami, T. Chondroitin/dermatan sulfate in the central nervous system. *Curr. Opin. Struct. Biol.* **17**, 536-545 (2007).

33. Iaci, J.F., Vecchione, A.M., Zimber, M.P. & Caggiano, A.O. Chondroitin sulfate proteoglycans in spinal cord contusion injury and the effects of chondroitinase treatment. *J Neurotrauma* **24**, 1743-1759 (2007).

34. Tanaka, M., Maeda, N., Noda, M. & Marunouchi, T. A chondroitin sulfate proteoglycan PTPzeta /RPTPbeta regulates the morphogenesis of Purkinje

cell dendrites in the developing cerebellum. *J. Neurosci.* **23**, 2804-2814 (2003).

35. Cooney, C.A. *et al.* Chondroitin sulfates play a major role in breast cancer metastasis: a role for CSPG4 and CHST11 gene expression in forming surface P-selectin ligands in aggressive breast cancer cells. *Breast Cancer Res.* **13**, R58 (2011).

36. Monzavi-Karbassi, B. *et al.* Chondroitin sulfate glycosaminoglycans as major P-selectin ligands on metastatic breast cancer cell lines. *Int. J. Cancer* **120**, 1179-1191 (2007).

37. Sakko, A.J. *et al.* Immunohistochemical Level of Unsulfated Chondroitin Disaccharides in the Cancer Stroma Is an Independent Predictor of Prostate Cancer Relapse. *Cancer Epidemiol. Biomarkers Prevent.* **17**, 2488-2497 (2008).

38. Yang, J. *et al.* Melanoma chondroitin sulfate proteoglycan enhances FAK and ERK activation by distinct mechanisms. *J. Cell Biol.* **165**, 881-891 (2004).

39. Kim, E. *et al.* Paradoxical effects of chondroitin sulfate-E on Japanese encephalitis viral infection. *Biochem. Biophys. Res. Commun.* **409**, 717-722 (2011).

40. Yip, G.W., Smollich, M. & Götte, M. Therapeutic value of glycosaminoglycans in cancer. *Mol. Cancer Ther.* **5**, 2139-2148 (2006).

41. Brittis, P.A., Canning, D.R. & Silver, J. Chondroitin sulfate as a regulator of neuronal patterning in the retina. *Science* **255**, 733-736 (1992).

42. Sugahara, K. & Mikami, T. Chondroitin/dermatan sulfate in the central nervous system. *Curr. Opin. Struct. Biol.* **17**, 536-545 (2007).

43. Miyata, S., Komatsu, Y., Yoshimura, Y., Taya, C. & Kitagawa, H. Persistent cortical plasticity by upregulation of chondroitin 6-sulfation. *Nat. Neurosci.* **15**, 414-422 (2012).

44. Wang, H. *et al.* Chondroitin-4-sulfation negatively regulates axonal guidance and growth. *J. Cell Sci.* **121**, 3083-3091 (2008).

45. Tully, S.E. *et al.* A chondroitin sulfate small molecule that stimulates neuronal growth. *J. Am. Chem. Soc.* **126**, 7736-7737 (2004).

46. Gama, C.I. *et al.* Sulfation patterns of glycosaminoglycans encode molecular recognition and activity. *Nat. Chem. Biol.* **2**, 467-473 (2006).

47. Habuchi, H., Habuchi, O. & Kimata, K. Sulfation pattern in glycosaminoglycan: Does it have a code? *Glycoconjugate J.* **21**, 47-52 (2004).

48. Shipp, E.L. & Hsieh-Wilson, L.C. Profiling the Sulfation Specificities of Glycosaminoglycan Interactions with Growth Factors and Chemotactic Proteins Using Microarrays. *Chemistry & Biology* **14**, 195-208 (2007).

49. Sugahara, K. *et al.* Structural studies on the chondroitinase ABC-resistant sulfated tetrasaccharides isolated from various chondroitin sulfate isomers. *Carbohydr. Res.* **255**, 145-163 (1994).

50. Sugahara, K. *et al.* Chondroitinase ABC-resistant sulfated trisaccharides isolated from digests of chondroitin/dermatan sulfate chains. *Carbohydr. Res.* **255**, 165-182 (1994).

51. Karst, N. & Jacquinet, J.-C. Chemical synthesis of the disaccharide repeating unit of shark cartilage chondroitin sulfate D, and of its methyl D-glycoside derivative. *J. Chem. Soc., Perkin Trans. 1*, 2709-2717 (2000).

52. Jacquinet, J.-C., Lopin-Bon, C. & Vibert, A. From polymer to size-defined oligomers: a highly divergent and stereocontrolled construction of chondroitin sulfate A, C, D, E, K, L, and M oligomers from a single precursor: part 2. *Chem. - Eur. J.* **15**, 9579-9595 (2009).

53. Gargiulo, V., Lanzetta, R., Parrilli, M. & De Castro, C. Structural analysis of chondroitin sulfate from Scyliorhinus canicula: a useful source of this polysaccharide. *Glycobiology* **19**, 1485-1491 (2009).

54. Jalkanen, S. & Jalkanen, M. Lymphocyte Cd44 Binds the Cooh-Terminal Heparin-Binding Domain of Fibronectin. *J. Cell Biol.* **116**, 817-825 (1992).

55. Carney, S.L. *et al.* Changes in proteoglycan turnover in experimental canine osteoarthritic cartilage. *Matrix* **12**, 137-147 (1992).

56. Silvestri, L., Baker, J.R., Roden, L. & Stroud, R.M. The C1q inhibitor in serum is a chondroitin 4-sulfate proteoglycan. *J. Biol. Chem.* **256**, 7383-7387 (1981).

57. Hikino, M. *et al.* Oversulfated Dermatan Sulfate Exhibits Neurite Outgrowth-promoting Activity toward Embryonic Mouse Hippocampal Neurons. *J. Biol. Chem.* **278**, 43744-43754 (2003).

58. Nadanaka, S., Clement, A., Masayama, K., Faissner, A. & Sugahara, K. Characteristic hexasaccharide sequences in octasaccharides derived from

shark cartilage chondroitin sulfate D with a neurite outgrowth promoting activity. *J. Biol. Chem.* **273**, 3296-3307 (1998).

59. Kinoshita, A. *et al.* Novel tetrasaccharides isolated from squid cartilage chondroitin sulfate E contain unusual sulfated disaccharide units GlcA(3-O-sulfate)beta1-3GalNAc(6-O-sulfate) or GlcA(3-O-sulfate)beta1-3GalNAc. *J. Biol. Chem.* **272**, 19656-19665 (1997).

60. Uchimura, K. *et al.* Functional analysis of the chondroitin 6-sulfotransferase gene in relation to lymphocyte subpopulations, brain development, and oversulfated chondroitin sulfates. *J. Biol. Chem.* **277**, 1443-1450 (2002).

61. Mizumoto, S. & Yamada, S. Human Genetic Disorders and Knockout Mice Deficient in Glycosaminoglycan. *Biomed. Res. Int.* **2014**, 495764 (2014).

62. Izumikawa, T., Sato, B. & Kitagawa, H. Chondroitin sulfate is indispensable for pluripotency and differentiation of mouse embryonic stem cells. *Sci. Rep.* **4**, 3701 (2014).

63. Mizumoto, S., Watanabe, M., Yamada, S. & Sugahara, K. Expression of N-acetylgalactosamine 4-sulfate 6-O-sulfotransferase involved in chondroitin sulfate synthesis is responsible for pulmonary metastasis. *Biomed. Res. Int.* **2013**, 656319 (2013).

64. Granovsky, M. *et al.* Suppression of tumor growth and metastasis in Mgat5-deficient mice. *Nat. Med.* **6**, 306-312 (2000).

65. Marth, J.D. Will the transgenic mouse serve as a Rosetta Stone to glycoconjugate function? *Glycoconj. J.* **11**, 3-8 (1994).

66. Pomin, V.H., Piquet, A.A., Pereira, M.S. & Mourao, P.A. Residual keratan sulfate in chondroitin sulfate formulations for oral administration. *Carbohydr. Polym.* **90**, 839-846 (2012).

67. Pan, J. *et al.* Oversulfated chondroitin sulfate is not the sole contaminant in heparin. *Nat. Biotech.* **28**, 203-207 (2010).

68. Yamada, S. & Sugahara, K. Potential therapeutic application of chondroitin sulfate/dermatan sulfate. *Curr. Drug Discov. Technol.* **5**, 289-301 (2008).

69. Lauder, R.M. Chondroitin sulphate: A complex molecule with potential impacts on a wide range of biological systems. *Complement. Ther. Med.* **17**, 56-62 (2009).

70. Scherz, H. & Bonn, G. *Analytical Chemistry of Carbohydrates*. (Thieme, Stuttgart; 1998).

71. Zumla, A., Nahid, P. & Cole, S.T. Advances in the development of new tuberculosis drugs and treatment regimens. *Nat. Rev. Drug Discov.* **12**, 388-404 (2013).

72. Nicolaou, K.C. & Mitchell, H.J. Adventures in Carbohydrate Chemistry: New Synthetic Technologies, Chemical Synthesis, Molecular Design, and Chemical Biology. *Angew. Chem. Int. Ed.* **40**, 1576-1624 (2001).

73. Karst, N.A. & Linhardt, R.J. Recent chemical and enzymatic approaches to the synthesis of glycosaminoglycan oligosaccharides. *Curr. Med. Chem.* **10**, 1993-2031 (2003).

74. Paulsen, H. Advances in Selective Chemical Syntheses of Complex Oligosaccharides. *Angew. Chem. Int. Ed.* **21**, 155-173 (1982).

75. Wuts, P.G.M. & Greene, T.W. *Greene's Protective Groups in Organic Synthesis*, Edn. 4th. (John Wiley & Sons, New Jersey; 2007).

76. Levy, D.E. & Fugedi, P. *The Organic Chemistry of Sugars*. (Taylor & Francis Group, Florida; 2006).

77. Robertson, J. *Protecting Group Chemistry*. (Oxford University Press, New York; 2000).

78. Kocienski, P. *Protecting Groups*, Edn. 3rd. (Georg Thieme Verlag, Stuttgart; 2003).

79. Nicolaou, K.C. *et al.* Total Synthesis of Apoptolidin: Part 1. Retrosynthetic Analysis and Construction of Building Blocks. *Angew. Chem. Int. Ed.* **40**, 3849-3854 (2001).

80. Nicolaou, K.C., Li, Y., Fylaktakidou, K.C., Mitchell, H.J. & Sugita, K. Total Synthesis of Apoptolidin: Part 2. Coupling of Key Building Blocks and Completion of the Synthesis. *Angew. Chem. Int. Ed.* **40**, 3854-3857 (2001).

81. Yu, B., Xie, J., Deng, S. & Hui, Y. First Synthesis of a Bidesmosidic Triterpene Saponin by a Highly Efficient Procedure. *J. Am. Chem. Soc.* **121**, 12196-12197 (1999).

82. McDonald, F.E. & Reddy, K.S. Convergent Synthesis of Digitoxin: Stereoselective Synthesis and Glycosylation of the Digoxin Trisaccharide Glycal. *Angew. Chem. Int. Ed.* **40**, 3653-3655 (2001).

83. Lefeber, D.J., Kamerling, J.P. & Vliegenthart, J.F.G. The Use of Diazabicyclo[2.2.2]octane as a Novel Highly Selective Dechloroacetylation Reagent. *Org. Lett.* **2**, 701-703 (2000).

84. Koeners, H.J., Verhoeven, J. & van Boom, J.H. Application of levulinic acid ester as a protective group in the synthesis of oligosaccharides. *Recl. Trav. Chim. Pays-Bas* **100**, 65-72 (1981).

85. Haines, A.H. Relative Reactivities of Hydroxyl Groups in Carbohydrates, in *Adv. Carbohydr. Chem. Biochem.* **33**, 11-109 (1976).

86. Ohlin, M., Johnsson, R. & Ellervik, U. Regioselective reductive openings of 4,6-benzylidene acetals: synthetic and mechanistic aspects. *Carbohydr. Res.* **346**, 1358-1370 (2011).

87. Rao, K.V., Patil, P.R., Atmakuri, S. & Kartha, K.P.R. Iodine–sodium cyanoborohydride-mediated reductive ring opening of 4,6-O-benzylidene acetals of hexopyranosides. *Carbohydr. Res.* **345**, 2709-2713 (2010).

88. Sakagami, M. & Hamana, H. A selective ring opening reaction of 4,6-O-benzylidene acetals in carbohydrates using trialkylsilane derivatives. *Tetrahedron Lett.* **41**, 5547-5551 (2000).

89. Lindhorst, T.K. *Essentials of Carbohydrate Chemistry and Biochemistry*, Edn. 3rd. (Wiley-VCH, Darmstadt; 2007).

90. Demchenko, A.V. *Handbook of Chemical Glycosylation*. (Wiley-VCH, Morlenbach; 2008).

91. Eby, R. & Schuerch, C. The use of 1-O-tosyl-d-glucopyranose derivatives in α-d-glucoside synthesis. *Carbohydr. Res.* **34**, 79-90 (1974).

92. Vankar, Y.D., Vankar, P.S., Behrendt, M. & Schmidt, R.R. Synthesis of β-O-glycosides using enol ether and imidate derived leaving groups. Emphasis on the use of nitriles as a solvent. *Tetrahedron* **47**, 9985-9992 (1991).

93. Schmidt, R.R., Behrendt, M. & Toepfer, A. Nitriles as Solvents in Glycosylation Reactions: Highly Selective *β*-Glycoside Synthesis. *Synlett.* **1990**, 694-696 (1990).

94. Chao, C.-S., Lin, C.-Y., Mulani, S., Hung, W.-C. & Mong, K.-k.T. Neighboring-Group Participation by C-2 Ether Functions in Glycosylations Directed by Nitrile Solvents. *Chem. - Eur. J.* **17**, 12193-12202 (2011).

95. Wulff, G. & Röhle, G. Results and Problems of O-Glycoside Synthesis. *Angew. Chem. Int. Ed.* **13**, 157-170 (1974).

96. Lu, S.-R., Lai, Y.-H., Chen, J.-H., Liu, C.-Y. & Mong, K.-K.T. Dimethylformamide: An Unusual Glycosylation Modulator. *Angew. Chem. Int. Ed.* **50**, 7315-7320 (2011).

97. Dulaney, S.B. & Huang, X. STRATEGIES IN SYNTHESIS OF HEPARIN/HEPARAN SULFATE OLIGOSACCHARIDES: 2000–PRESENT. *Adv. Carbohydr. Chem. Biochem.* **67**, 95-136 (2012).

98. Xu, Y. *et al.* Chemoenzymatic synthesis of homogeneous ultralow molecular weight heparins. *Science* **334**, 498-501 (2011).

99. Yasomanee, J.P. & Demchenko, A.V. From Stereocontrolled Glycosylation to Expeditious Oligosaccharide Synthesis. *Trends Glycosci. Glycotechnol.* **25**, 13-42 (2013).

100. Jacquinet, J.-C., Rochepeau-Jobron, L. & Combal, J.-P. Multigram syntheses of the disaccharide repeating units of chondroitin 4- and 6-sulfates. *Carbohydr. Res.* **314**, 283-288 (1998).

101. Tamura, J.-i., Neumann, K.W. & Ogawa, T. A regio- and stereoselective synthesis of 4-O-sulfated chondroitin di- and tetrasaccharides based on the strategy designed for the elongation of the repeating unit. *Bioorg. Med. Chem. Lett.* **5**, 1351-1354 (1995).

102. Tamura, J.-i., Neumann, K.W., Kurono, S. & Ogawa, T. Synthetic approach towards sulfated chondroitin di-, tri- and tetrasaccharides corresponding to the repeating unit. *Carbohydr. Res.* **305**, 43-63 (1997).

103. Tamura, J., Nakada, Y., Taniguchi, K. & Yamane, M. Synthesis of chondroitin sulfate E octasaccharide in a repeating region involving an acetamide auxiliary. *Carbohydr. Res.* **343**, 39-47 (2008).

104. Lubineau, A. & Bonnaffé, D. Access to molecular diversity in glycosaminoglycans: combinatorial synthesis of eight chondroitin sulfate disaccharides. *Eur. J. Org. Chem.* **1999**, 2523-2532 (1999).

105. Vibert, A., Jacquinet, J.-C. & Lopin-Bon, C. Recent advances in the chemical and enzymatic chondroitin sulfate synthesis. *J. Carbohydr. Chem.* **30**, 393-414 (2011).

106. Bedini, E. *et al.* Semi-synthesis of unusual chondroitin sulfate polysaccharides containing GlcA(3-O-sulfate) or GlcA(2,3-di-O-sulfate) units. *Chem. – Eur. J.* **18**, 2123-2130 (2012).

107. Bedini, E. & Parrilli, M. Synthetic and semi-synthetic chondroitin sulfate oligosaccharides, polysaccharides, and glycomimetics. *Carbohydr. Res.* **356**, 75-85 (2012).

108. Karst, N. & Jacquinet, J.-C. Stereocontrolled total syntheses of shark cartilage chondroitin sulfate D-related tetra- and hexasaccharide methyl glycosides. *Eur. J. Org. Chem.* **2002**, 815-825 (2002).

109. Tamura, J. *et al.* Synthesis and interaction with midkine of biotinylated chondroitin sulfate tetrasaccharides. *Bioorg. Med. Chem. Lett.* **22**, 1371-1374 (2012).

110. Despras, G. *et al.* Toward libraries of biotinylated chondroitin sulfate analogues: from synthesis to in vivo studies. *Chem. - Eur. J.* **19**, 531-540 (2013).

111. Vibert, A., Lopin-Bon, C. & Jacquinet, J.-C. From polymer to size-defined oligomers: a step economy process for the efficient and stereocontrolled construction of chondroitin oligosaccharides and biotinylated conjugates thereof: part 1. *Chem. - Eur. J.* **15**, 9561-9578 (2009).

112. Davidson, E.A. & Meyer, K. Structural Studies on Chondroitin Sulfuric Acid.1,2 I. The Nature of Chondrosine. *J. Am. Chem. Soc.* **76**, 5686-5689 (1954).

113. Levene, P.A. ON CHONDROSIN. *J. Biol. Chem.* **140**, 267-277 (1941).

CHAPTER 2: CHEMICAL SYNTHESIS OF CHONDROITIN SULFATE DISACCHARIDE LIBRARY

2.1 Introduction

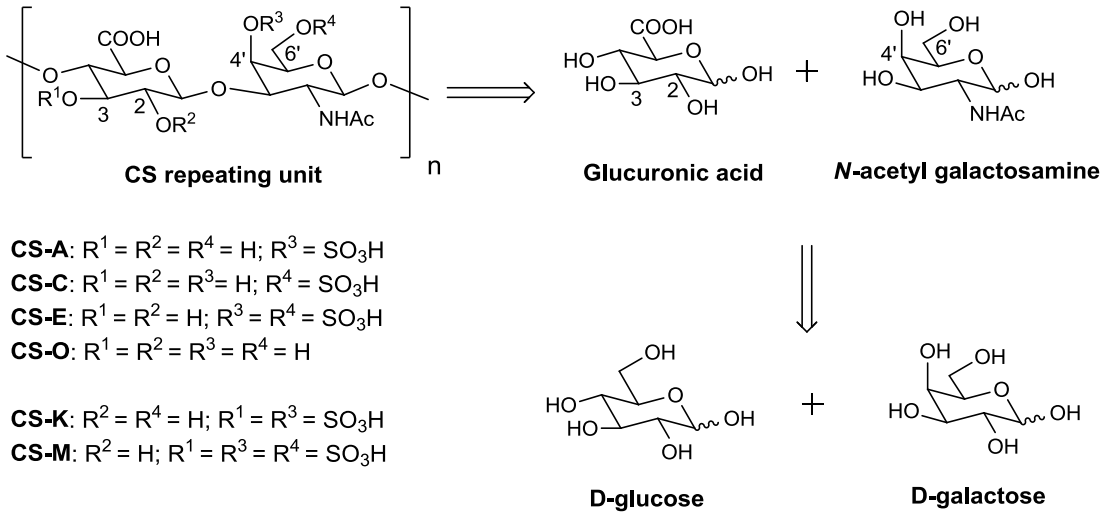


Figure 2-1. Retrosynthetic analysis for CS disaccharide synthesis

A retrosynthetic analysis of chondroitin sulfate (CS) is depicted in Figure 2-1. To obtain the sulfation patterns theoretically possible in the CS repeating unit, the required monomeric building blocks include the glucuronic acid and *N*-acetyl galactosamine precursors which could be obtained from D-glucose and D-galactose respectively; these monosaccharides are readily available and of low cost. We first embarked on the synthesis of sulfation patterns CS-A, CS-C, CS-E, CS-O, CS-K and CS-M which were more well studied[1-12] to familiarize with the chemical synthesis of CS before devising the synthetic strategy to obtain other novel sulfation patterns. These six analogues mainly involve sulfation at the C-4'

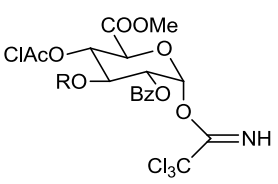
and/or C-6’ positions on the *N*-acetyl galactosamine moiety, or have sulfate groups on C-3 of the glucuronic acid moiety. For the efficient synthesis of the CS disaccharide library, a divergent approach was adopted, through the use of common reaction intermediates.

In naturally occurring CS polysaccharides, the glycosidic bonds have *β*-stereoselectivity. Hence, stereoselective formation of the 1,2-*trans* glycosidic linkage is required, which can be obtained by attaching participating groups on the C-2 position of the donor unit, such as ester protecting groups. In the accepter unit, the *N*-acetyl group is required on the C-2’ position. However, *N*-acetyl groups reduce solubility and reactivity of the precursor[6] hence temporary protecting groups such as trichloroacetyl, trichloroethoxycarbonyl and phthaloyl groups may be used[13-15], these can be converted into *N*-acetyl groups during later steps of synthesis.

For the synthesis of each CS analogue, different protection strategies have to be utilised to direct regioselective sulfation of the required sites in the presence of other chemically equivalent hydroxyl groups. Upon attachment of the O-sulfate groups in the CS disaccharide, the increase in negative charges decrease solubility of the compound in common organic solvents, making isolation and characterization difficult.[16] These acidic groups are also reactive and labile to various pH conditions, making them incompatible with many of the modification steps. Thus, the sulfate groups should only introduced during the final steps of

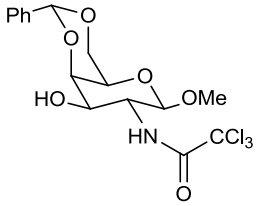
synthesis upon selective cleavage of orthogonal protecting groups. To introduce the sulfate groups, selectively deprotected hydroxyl groups are first reacted with the sulfating agent, prior to cleavage of the remaining protecting groups.

By adapting some of the currently available methods[3], we improved on the strategy to prepare CS-A, CS-C, CS-E, CS-O, CS-K and CS-M by reducing reaction time and increasing reaction yield through a divergent synthetic approach. Two protected glucuronic acid donors and one protected *N*-acetyl galactosamine acceptor were prepared after careful retrosynthetic analysis (Figure 2-2). The protecting groups utilised were carefully chosen to allow for stereoselective glycosylation and sulfation of appropriate sites to obtain the required CS analogues.



Glucuronic acid donor

D1: R = Bz
2-17: R = Nap



***N*-acetylgalactosamine acceptor**

2-30

Figure 2-2. Glycosyl donor and acceptor for the synthesis of CS disaccharides

D-glucose was derivatized to obtain protected glucuronic acids **D1**[17] and **2-17**, which were used as the glycosyl donors. To form the glycosidic linkage, the trichloroacetimidate auxiliary[18] was required at C-1 to function as the leaving group and facilitate attachment of the neighboring sugar unit. The C-2 and C-3

positions were possible sites of sulfation and hence depending on the target sulfation sites, appropriate protecting groups which could be selectively cleaved were required to enable regioselective sulfation. Since a C-2 ester protecting group can also act as a participating group during glycosylation to direct the formation of the desired *β*-glycosidic bond[1], the benzoyl ester protecting group was chosen for C-2. In CS polysaccharides, the C-4 position is the site of elongation. Hence another orthogonal protecting group was required, to enable selective cleavage and synthesis of longer CS sequences when required. To obtain the C-6 carboxyl functional group, selective oxidation of the C-6 primary alcohol in D-glucose was required, followed by protection of the C-6 carboxylate as a methyl ester.

Similarly, D-galactose was derivatized by the introduction of appropriate protecting groups to obtain protected *N*-acetyl galactosamine **2-30,** which was used as the glycosyl acceptor during synthesis. The terminal C-1' position was protected as methyl ether, this served as a bioisosteric group and also kept the anomeric proton in the axial conformation to aid in compound characterization.[19] In *N*-acetylgalactosamine, a nitrogen was required in C-2' position and thus the C-2' oxygen atom in D-galactose was replaced. The trichloroacetyl (TCAHN) group was chosen as the temporary C-2' amide protecting group[15]; this acted as the precursor to form the C-2' *N*-acetyl functionality. This TCAHN group could also direct *β*-stereoselective glycosylation, and promote substrate solubility. In CS, the possible sites of sulfation on *N*-acetyl galactosamine include C-4' and C-6',

hence both positions were simultaneously protected as a benzylidene acetal, which acts as an orthogonal protecting group that could be cleaved when required for selective attachment of sulfate groups. A free hydroxyl group was also required on the C-3' position for reaction with the glycosyl donor during glycosylation.

2.1.1 Protection Strategy to Obtain CS-A, CS-C, CS-E and CS-O

Glucuronic acid donor **D1** and *N*-acetyl galactosamine acceptor **2-30** were utilised for the synthesis of CS-A, CS-C, CS-E and CS-O disaccharides (Figure 2-1).[1] In these four CS sulfation patterns, there are no sulfate groups on the glucuronic acid moiety, hence the C-2 and C-3 hydroxyl groups can be protected orthogonally as benzoyl esters to protect against sulfation, and cleaved only in the final step. The benzylidene acetal protecting group can first be cleaved by acidic hydrolysis to regenerate the C-4' and C-6' hydroxyl groups, to enable sulfation on the *N*-acetyl galactosamine moiety. To obtain regioselective sulfation of the C-6' primary alcohol, milder reaction conditions and shorter reaction times can be used (to obtain CS-C). Similarly when sulfation on C-4' site is required, the more reactive C-6' primary alcohol can first be preferentially protected[20], before subjecting the free C-4' alcohol to the sulfating agent (to obtain CS-A). If both C-4' and C-6' sites require sulfation, a prolonged sulfation step would proceed upon removal of the benzylidene acetal (to obtain CS-E). For non-sulfated CS disaccharide, global deprotection of all remaining ester protecting groups would proceed after glycosylation without any sulfation step (to obtain CS-O). In this way, all 4

possible sulfation combinations on the *N*-acetyl galactosamine moiety can be obtained.

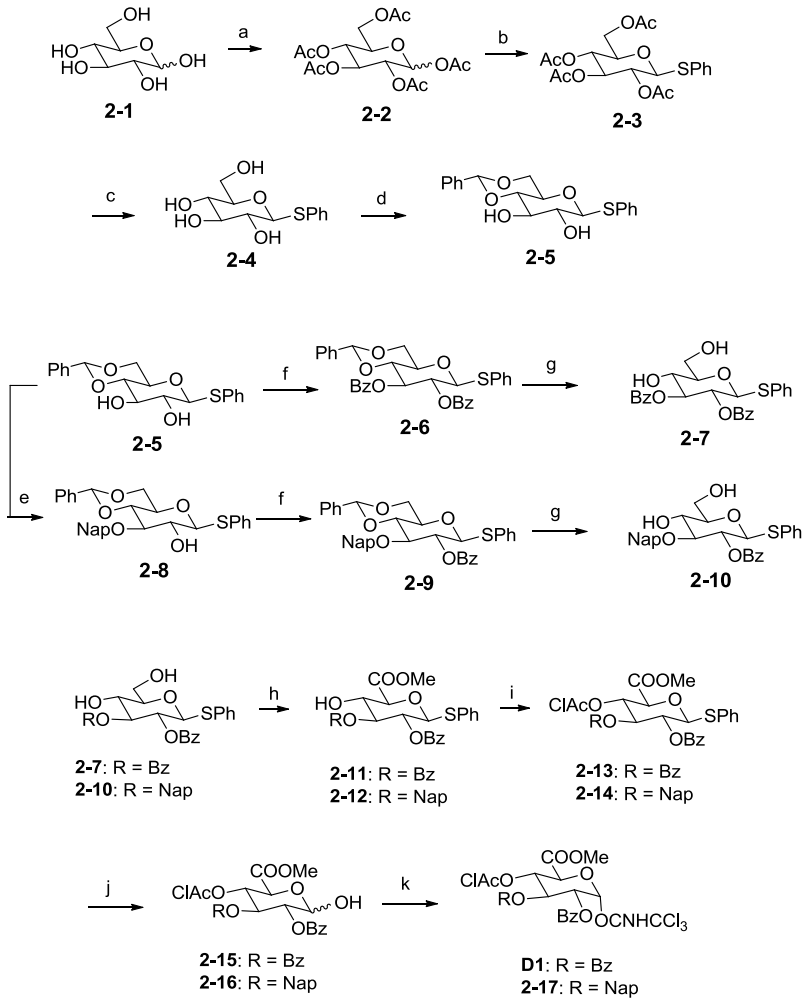
2.1.2 Protection Strategy to Obtain CS-K and CS-M

Glucuronic acid donor **2-17** and *N*-acetyl galactosamine acceptor **2-30** were used to prepare CS-K and CS-M analogues.[4] These analogues are sulfated at the C-3 position of the glucuronic acid moiety, in addition to sulfation on *N*-acetyl galactosamine. Hence, an additional orthogonal protecting group was required at C-3 to allow for selective deprotection and sulfation in later steps of synthesis. To achieve this, the naphthyl ether protecting group was introduced at the C-3 position. This protecting group could be selectively removed via oxidative cleavage with DDQ to release the C-3 hydroxyl group for reaction with the sulfating agent prior to the global deprotection step.

2.2 Synthesis of Glucuronic Acid Donors D1 and 2-17

D-glucose was used as the starting material to prepare common intermediate **2-5**, which was required for the synthesis of glycosyl donors **D1** and **2-17** (Scheme 2-1). The five hydroxyl groups in D-glucose were first protected using acetic anhydride and pyridine to form penta-acetylated glucose **2-2**. This provided solubility of the substrate in organic solvents, and intermediate **2-2** was next reacted with thiophenol and boron trifluoride as the Lewis acid. This step introduced the thiophenol protecting group exclusively at the anomeric C-1 carbon to form thioglycoside **2-3**.

Scheme 2-1. Synthesis of glycosyl donors **D1**, **2-17**[a]



[a]Reagents and conditions: (a) Ac_2O, DMAP, pyridine, 1 h, 85%; (b) PhSH, $BF_3.OEt_2$, 35^oC, 14 h, 70%; (c) NaOMe, MeOH, 1 h, 90%, (d) $PhCH(OMe)_2$, CSA, CH_3CN/DMF, 16 h, 60^oC, 72%; (e) Bu_2SnO, toluene, 4 h, 110^oC, then NapBr, CsF, DMF, 12 h, 47%; (f) Bz_2O, TEA, DMAP, CH_2Cl_2, 1 h, **2-6**: 80%, **2-9**: 86%; (g) TFA, CH_2Cl_2/H_2O, 4 h, **2-7**:85%, **2-10**: 88%; (h) TEMPO, BAIB, CH_2Cl_2/H_2O, 4 h, then MeI, $NaHCO_3$, TBAI, DMF, 1 h, 50^oC, **2-11**: 51%, **2-12**: 48%; (i) Cl_2Ac_2O, pyridine, 1 h, CH_2Cl_2, **2-13**: 89%, **2-14**: 82%; (j) NIS, TFA, 14 h, CH_2Cl_2/H_2O, **2-15**: 72%, **2-16**: 70%; (k) Cl_3CCN, DBU, CH_2Cl_2, 6 h, **D1**: 74%, **2-17**: 69%

With the anomeric carbon protected, the remaining acetyl groups could then be cleaved by basic hydrolysis to form **2-4**, followed by regioselective protection of the C-4 and C-6 hydroxyl group as a benzylidene acetal via the formation of a stable double chair conformation to obtain common intermediate **2-5**.[21] By varying some of the synthetic steps to establish different protecting groups on the C-2 and C-3 positions, both glucuronic acid donors **D1** and **2-17** could be prepared by using **2-5** as the common precursor.

To obtain glycosyl donor **D1**, the hydroxyl groups on C-2 and C-3 were first benzoylated by reacting intermediate **2-5** with benzoic anhydride and DMAP as the nucleophilic catalyst to form intermediate **2-6**. For donor **2-17**, regioselective naphthylation at the C-3 position was required. This was achieved by refluxing **2-5** with dibutyltin oxide in toluene to first generate the dibutyl tin complex *in situ*, before reacting the intermediate with naphthyl bromide and caesium fluoride to direct naphthylation at the C-3 position to form intermediate **2-8**.[22] The free hydroxyl group on C-2 was next protected with benzoic anhydride to form **2-9**. With the key protecting groups established on C-2 and C-3 positions, intermediates **2-6** and **2-9** were next subjected to acidic hydrolysis using trifluoracetic acid in the presence of water to cleave the benzylidene acetal and release the C-4 and C-6 hydroxyl groups.[23] This formed intermediates **2-7** and **2-10** as the respective products.

The C-6 primary alcohols in intermediates **2-7** and **2-10** were then selectively oxidized using TEMPO as the radical initiator and BAIB co-oxidant to form the C-6 carboxylate group[24], which was subsequently protected as a methyl ester to form intermediates **2-11** and **2-12**. The C-4 hydroxyl groups were next protected as chloroacetyl esters to form intermediates **2-13** and **2-14**. This chloroacetate could serve as an orthogonal protecting group which can be selectively cleaved to release the C-4 hydroxyl group for further elongation when required.[25] Thioglycoside **2-13** was initially used as the glycosyl donor as the anomeric thiophenol protecting group could be activated by a promoter to function as the leaving group during glycosylation.

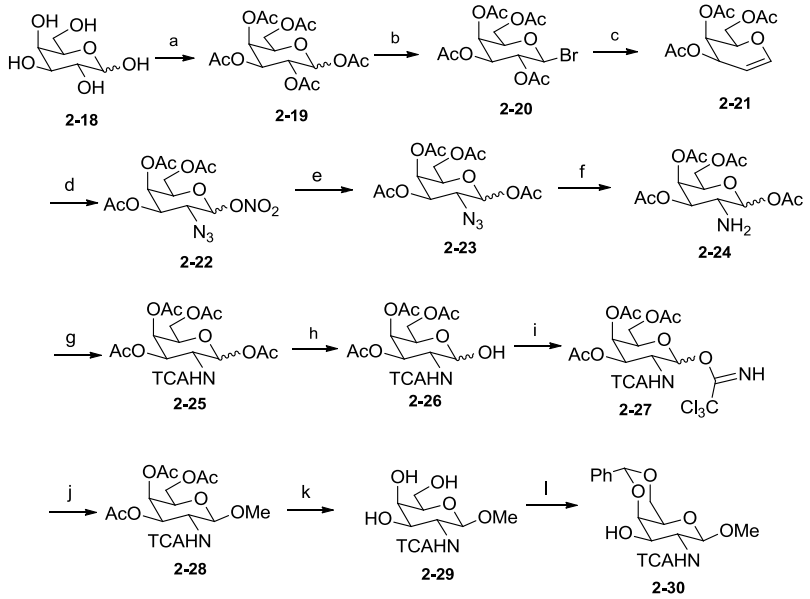
To further increase the reactivity of **2-13** and **2-14**, two additional steps were introduced to obtain glucuronic acid donors **D1** and **2-17**. This involved cleavage of the thiophenol protecting group by *N*-iodosuccinimide (NIS) and trifluoroacetic acid (TFA) to regenerate the anomeric hydroxyl group in intermediates **2-15** and **2-16**[26], prior to attachment of the trichloroacetimidate auxiliary to form **D1** and **2-17.**

2.3 Synthesis of *N*-acetyl galactosamine Acceptor 2-30

To obtain *N*-acetyl galactosamine acceptor **2-30**, D-galactose was used as the starting material (Scheme 2-2). D-galactose was first protected with acetic anhydride and pyridine to form penta-acetylated galactose **2-19**. Intermediate **2-19** was next reacted with HBr in acetic acid to replace the anomeric group with a

bromine atom followed by reductive elimination using zinc and sodium acetate to form the double bond in D-galactal **2-21**.[27, 28] Azidonitration via a radical mediated approach enabled attachment of the azide group onto the C-2' position in intermediate **2-22**, by reacting intermediate **2-21** with excess cerium ammonium nitrate and sodium azide.[29] This introduced the nitrogen atom on the C-2' position of galactose as an azide group; however only moderate yields of 45% could be obtained.

Scheme 2-2. Synthesis of glycosyl acceptor **2-30**[a]



[a]Reagents and conditions: (a) Ac_2O, DMAP, pyridine, 1 h, 85%; (b) HBr, AcOH, EtOAc, 1 h, 76%; (c) Zn, $CuSO_4$, NaOAc, EtOAc, AcOH/H_2O, 4 h, 85%; (d) NaN_3, CAN, MeCN, -15^{o}C, 14 h, 45%; (e) NaOAc, AcOH, 100^{o}C, 1 h, 82%; (f) H_2, Pd/C, TsOH.H_2O, THF, 14 h, 90%; (g) Cl_3CCOCl, TEA, THF, 30 mins, 71%; (h) 33% $MeNH_2$ in EtOH, THF, 4 h, 71%; (i) Cl_3CCN, DBU, CH_2Cl_2, 6 h, 72%; (j) MeOH, TMSOTf, CH_2Cl_2, 3Å MS, 1 h, -10^{o}C, 84% (k) NaOMe, MeOH, 1 h, 90%; (l) $PhCH(OMe)_2$, CSA, MeCN, 14 h, 55^{o}C, 73%

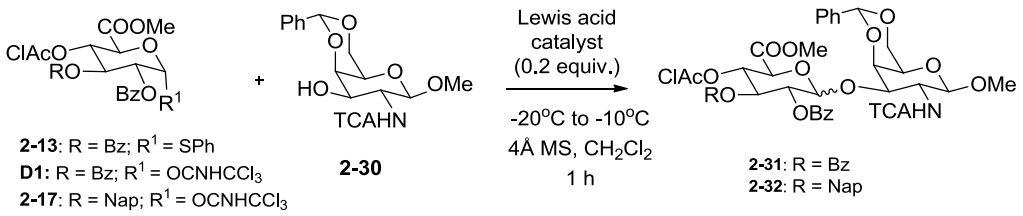
The anomeric group was next substituted with a stable acetate group, and the C-2' azide in **2-23** was hydrogenated to obtain amine **2-24**. To obtain the C-2' *N*-acetyl group, direct acetylation of the amine group by reaction with acetyl chloride and pyridine was initially attempted, but solubility of the product was very poor and isolation was difficult. Thus, the amine group was protected as an *N*-trichloroacetyl group (TCANH) to form **2-25**[30, 31]; this group could be reduced to an *N*-acetyl group when required.

With the amine functionality suitably protected, the anomeric acetate group was next selectively deprotected using methylamine to release the hydroxyl group, before attachment of the trichloroacetimidate auxiliary to form **2-27**. Glycosylating intermediate **2-27** with methanol using TMSOTf catalyst formed the methyl ether with *β*-stereoselectivity due to the C-2' TCANH directing group. Deacetylation followed by regioselective protection of the C-4' and C-6' hydroxyl groups with benzylidene acetal formed protected *N*-acetyl galactosamine acceptor **2-30**.

Glycosyl acceptor **2-30** was next reacted with glycosyl donors **D1** and **2-17** to obtain protected disaccharides **2-31** and **2-32**, which was utilised for the synthesis of CS-A, CS-C, CS-E, CS-O, CS-K and CS-M.

2.4 Synthesis of Protected Disaccharides 2-31 and 2-32 via Glycosylation

Table 2-1. Synthesis of protected disaccharides **2-31** and **2-32**



2-13: R = Bz; R^1 = SPh
D1: R = Bz; R^1 = $OCNHCCl_3$
2-17: R = Nap; R^1 = $OCNHCCl_3$

2-30

2-31: R = Bz
2-32: R = Nap

Entry	Glycosyl donor	Lewis acid	Product	Yield (%)	Stereoselectivity (% β-product)
1	**2-13**	NIS, TfOH	**2-31**	52	80
2	**D1**	TMSOTf	**2-31**	78	>99
3	**D1**	CuOTf	**2-31**	68	>99
4	**D1**	AgOTf	**2-31**	60	>99
5	**2-17**	TMSOTf	**2-32**	74	>99

During the initial glycosylation attempts, thioglycoside **2-13** was used as the donor, by utilising the anomeric thiophenol group as the glycosyl auxiliary, activated by NIS and triflic acid. The presence of the C-2 participating group in donor **2-13** directed β-stereoselective glycosylation with acceptor **2-30** to form mainly the β-product **2-31**; but a fraction of the desired product was lost due to formation of the α-anomer (Table 2-1, entry 1). In addition, it was difficult to separate the anomers via chromatographic techniques as both anomers were structurally very similar, and some desired product had to be sacrificed for purity. To avert this, thioglycoside **2-13** was modified to donor **D1,** which employed the more reactive trichloroacetimidate auxiliary. In the presence of the C-2

participating group, donor **D1** formed almost exclusively the β-product (table 2-1, entry 2)[1]; thus the trichloroacetimidate auxiliary was utilised for subsequent glycosylation reactions. To further improve glycosylation yield, several Lewis acids such as TMSOTf, CuOTf and AgOTf were attempted, but it was noted that TMSOTf gave the highest reaction yields (Table 2-1, entries 2 – 4).

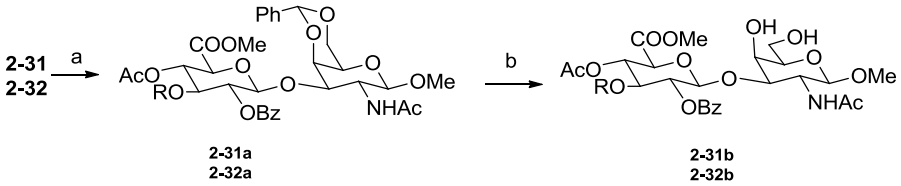
Due to the high reactivity of donor **D1**, a low temperature of -40^{o}C was used initially. However, glycosylation of **D1** with acceptor **2-30** using TMSOTf could not proceed at -40^{o}C after 24 hours, while raising the temperature to -10^{o}C allowed completion of the reaction within 1 hour. A further increase to room temperature however caused the formation of side products due hydrolysis of the glycosyl donor, and a drop in β-stereoselectivity. Thus, we decided to add the TMSOTf Lewis acid at -20^{o}C and allow the reaction to proceed at -10^{o}C for 1 hour. The optimized reaction conditions were applied for the synthesis of protected disaccharides **2-31** and **2-32**, using glycosyl donors **D1** and **2-17** respectively (Table 2-1, entries 2 and 5).

2.5 Synthesis of CS-A, CS-C, CS-E, CS-O, CS-K, CS-M

After the glycosylation step, protected dimers **2-31** and **2-32** were subjected to reduction with tributyl tin hydride to convert the *N*-trichloroacetyl group to *N*-acetyl group required in the *N*-acetylgalactosamine moiety[15] (Scheme 2-3). The benzylidene acetal was next cleaved via acidic hydrolysis, by refluxing in aqueous acetic acid to liberate the C-4’ and C-6’ hydroxyl groups in disaccharides **2-31b**

and **2-32b**. It was noted that prolonged heating in acid caused an unwanted side reaction: acid catalyzed ring opening of the anomeric methyl ether linkage, and a mixture of α- and β- anomers (via subsequent ring closure) were isolated at the end of the reaction. Hence, this reaction had to be monitored carefully and quenched immediately upon completion to reduce side product formation.

Scheme 2-3. Synthesis of disaccharides **2-31b** and **2-32b**[a]



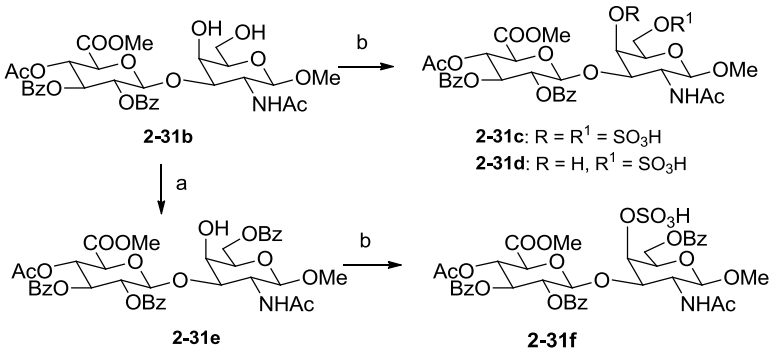
[a]Reagents and conditions: (a) Bu_3SnH, ABCN, toluene, 4 h, **2-31a**: 68%; **2-32a**: 73%; (b) 80% AcOH, 1 h, 80^{o}C, **2-31b**: 85%; **2-32b**: 88%

With disaccharides **2-31b** and **2-32b** on hand, the free hydroxyl groups on the *N*-acetyl galactosamine moiety were next subjected to regioselective sulfation with the other functionalities suitably protected. **2-31b** was utilised for the synthesis of CS-A, CS-C, CS-E and CS-O, while **2-32b** was used for the synthesis of CS-K and CS-M.

For the sulfation of both C-4' and C-6' hydroxyl groups, **2-31b** was reacted with a large excess of sulfating agent SO_3.TMA (30 equiv), and the reaction was heated for a prolonged period (48 hours), to obtain disulfated product **2-31c** (Scheme 2-4)[1]. Mono-sulfation of the more reactive C-6' primary hydroxyl group required

milder conditions, employing a smaller amount of sulfating agent (3 equiv.), and a shorter reaction period of 4 hours to obtain C-6' monosulfated disaccharide **2-31d**. For the mono-sulfation of the C-4' hydroxyl group, the more reactive C-6' hydroxyl group was first protected as a benzoyl ester by reacting **2-31b** with benzoyl cyanide in pyridine to form **2-31e**.[11] Sulfation of intermediate **2-31e** furnished C-4' monosulfated analogue **2-31f**.

Scheme 2-4. Synthesis of disaccharides **2-31c**, **2-31d**, **2-31f**[a]

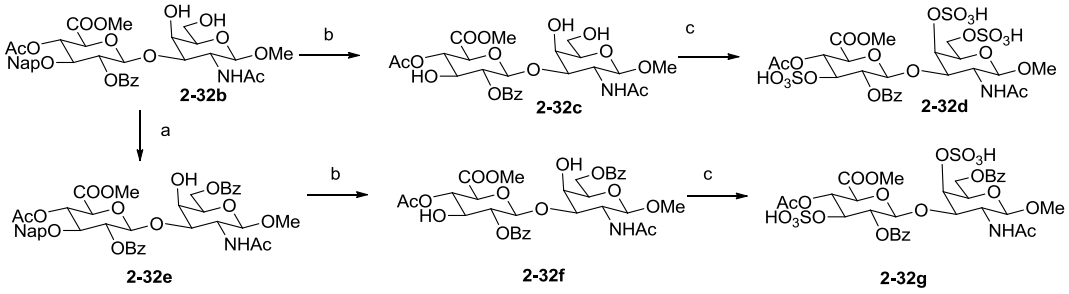


[a]Reagents and conditions: (a) BzCN, pyridine, 16 h, 75%; (b) **2-31c**: SO_3.TMA (30 equiv.), DMF, 50°C, 48 h, 60%; **2-31d**: SO_3.TMA(3 equiv.), DMF, 50°C, 4 h, 67%; **2-31f**: SO_3.TMA (10 equiv.), DMF, 50°C, 48 h, 64%

In the structure of CS-M, both the C-4' and C-6' hydroxyl groups in the *N*-acetylgalactosamine moiety are sulfated, in addition to the C-3 hydroxyl group on the glucuronic acid moiety. Hence, the C-3 protecting group in **2-32b** was first cleaved by DDQ oxidation to form **2-32c**[32] before sulfation step of all three hydroxyl groups to form tri-sulfated analogue **2-32d** (Scheme 2-5). To obtain the sulfation pattern on CS-K, where the C-3 and C-4' positions are sulfated, the

more reactive hydroxyl group on the C-6' position of **2-32b** was first selectively protected as a benzoyl ester to form **2-32e**. The C-3 naphthyl ether on the glucuronic acid moiety was next removed by oxidative cleavage with DDQ to release the hydroxyl group to form **2-32f**. The two unprotected hydroxyl groups were then subjected to sulfation to obtain **2-32g**.[4] It was noted that a stronger sulfating agent SO_3.TEA was required for the sulfation of the C-3 hydroxyl group on glucuronic acid to obtain **2-32d** and **2-32g**, as SO_3.TMA was not sufficiently strong to ensure reaction completion.

Scheme 2-5. Synthesis of disaccharides **2-32d**, **2-32g**[a]

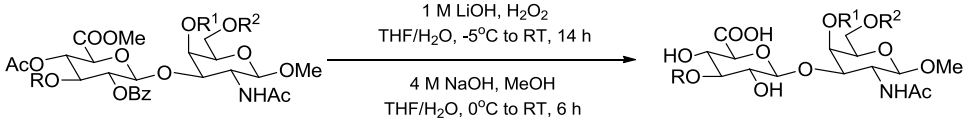


[a]Reagents and conditions: (a) BzCN, pyridine, 16 h, 42%; (b) DDQ, CH_2Cl_2/MeOH, 12 h, **2-32c**: 64%, **2-32f**: 33%; (c) SO_3.TEA (20 equiv.), DMF, 48 h, 50^{o}C, **2-32d**: 57%, **2-32g**: 80%

With the sulfate groups attached at the desired positions, global deprotection of the ester protecting groups proceeded in the final step to release all the masked hydroxyl groups to form the target molecules (Scheme 2-6). In this basic hydrolysis step, the intermediates were first subjected to LiOH/H_2O_2 before

saponification with NaOH; this sequential LiOOH-NaOH treatment was to minimize *β*-elimination at the C-4 position.[33] Through this approach, CS-A, CS-C, CS-E, CS-O, CS-K and CS-M were successfully synthesized.

Scheme 2-6. Synthesis of CS-A, CS-C, CS-E, CS-O, CS-K, CS-M



2-31b: R = Bz; $R^1 = R^2$ = H
2-31c: R = Bz; $R^1 = R^2 = SO_3H$
2-31d: R = Bz; R^1 = H; $R^2 = SO_3H$
2-31f: $R = R^2$ =Bz; $R^1 = SO_3H$
2-32d: $R = R^1 = R^2 = SO_3H$
2-32g: R^2 = Bz; $R = R^1 = SO_3H$

CS-O: $R = R^1 = R^2$ = H (80%);
CS-E: R = H; $R^1 = R^2 = SO_3H$ (78%);
CS-C: $R = R^1$ = H; $R^2 = SO_3H$ (75%);
CS-A: $R = R^2$ = H; $R^1 = SO_3H$ (68%);
CS-M: $R = R^1 = R^2 = SO_3H$ (72%);
CS-K: R^2 = H; $R = R^1 = SO_3H$ (78%)

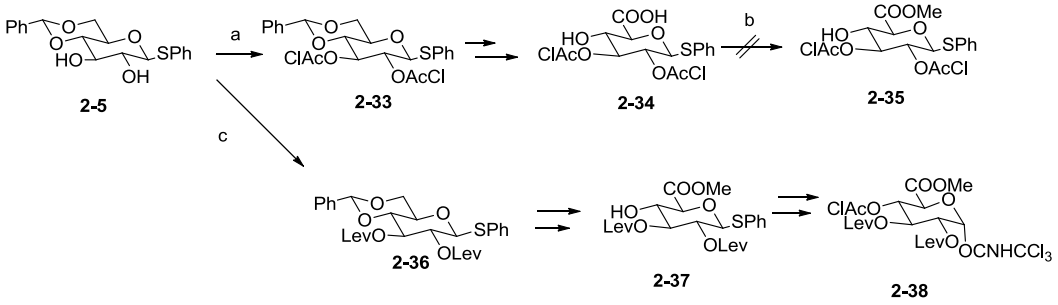
2.6 Protection Strategy to Obtain CS-R

Upon successful synthesis of the 6 CS disaccharides, we proceeded to prepare CS-R, which has sulfation on both C-2 and C-3 of glucuronic acid, but no sulfate groups on *N*-acetyl galactosamine.[1] An orthogonal protecting group was thus required on both C-2 and C-3 positions of the glucuronic acid monomer for selective deprotection and sulfation. Several orthogonal protecting groups were attempted, which included the chloroacetate ester and levulinate ester. Ester protecting groups were favored on C-2 as they can also serve as participating groups to direct *β*-stereoselective glycosylation.

2.6.1 Synthesis of Glucuronic Acid Donors 2-35 and 2-38

The proposed glycosyl donors were prepared using similar synthetic steps employed for the synthesis of donors **D1** and **2-17**, by replacing the C-2 benzoyl ester with other ester groups which can be cleaved in an orthogonal fashion.

Scheme 2-7. Synthesis of glycosyl donors **2-35** and **2-38**



Reagents and conditions: (a) Cl_2Ac_2O, pyridine, CH_2Cl_2, 1 h, 85%; (b) MeI, $NaHCO_3$, DMF, 1 h, 55^oC; (c) levulinic acid, DCC, DMAP, CH_2Cl_2, 4 h, 73%

Common intermediate **2-5** was first protected at the C-2 and C-3 position using chloroacetic anhydride to obtain intermediate **2-33** (Scheme 2-7). Following similar modification steps of acid hydrolysis and selective primary oxidation, intermediate **2-34** was obtained, and was next subjected to methylation for the protection of the C-6 carboxyl group. However, the chloroacetyl protecting group was not stable under the methylation reaction conditions and many side products were formed. This was partly attributed to the acidic nature of the chloroacetate protons, which reacted with the methylating agent in the presence of weak bases such as carbonates. Thus, we decided to modify the C-2 and C-3 protecting groups to levulinate esters by coupling **2-5** with levulinic acid, in the presence of

DCC and DMAP catalyst to form **2-36**. **2-36** was modified by similar methods which included benzylidene acetal cleavage, selective C-6 oxidation and methylation of the C-6 acid to obtain intermediate **2-37**. The C-4 hydroxyl group was next protected as the orthogonal chloroacetate ester before anomeric deprotection and attachment of the trichloroacetimidate auxiliary to form protected glucuronic acid donor **2-38**.

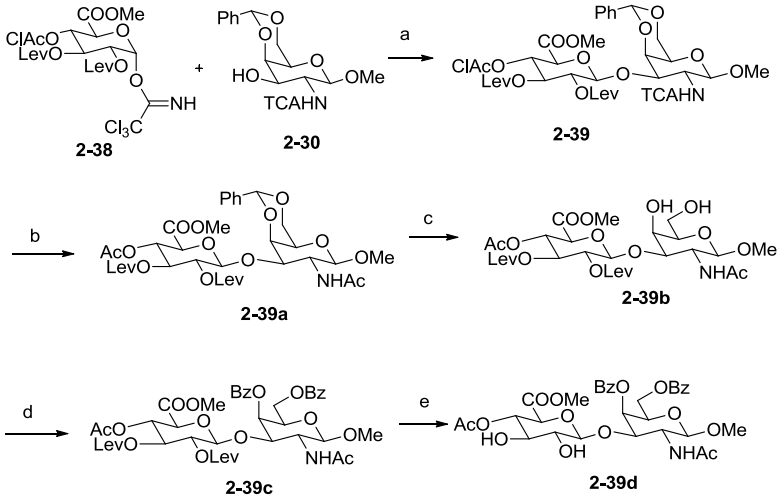
2.6.2 Synthesis of Disaccharide 2-39d

Using the previously optimized glycosylation conditions, **2-38** was used as the glycosyl donor to react with acceptor **2-30** to obtain disaccharide **2-39** (Scheme 2-8). A relatively low glycosylation yield was obtained when the C-2 levulinate ester was used (donor **2-38**), this could be attributed to long length of the C-2 protecting group, which could interfere with glycosylation at the anomeric position. **2-39** was next subjected to tributyltin hydride reduction to obtain the C-2' *N*-acetyl group, before acidic hydrolysis to cleave the benzylidene acetal. Both the C-4' and C-6' hydroxyl groups in **2-39b** were then protected as benzoyl esters; these serve as temporary protecting groups during the sulfation step, which can be cleaved via basic hydrolysis in the global deprotection step.

The levulinate esters on C-2 and C-3 in intermediate **2-39c** were finally subjected to selective cleavage by hydrazine acetate in acetic acid to release the hydroxyl groups for reaction with the sulfating agent, but this step had very poor yields of 30%. There was insufficient product (intermediate **2-39d**) to proceed to sulfation

and global deprotection step to obtain CS-R, even after several repeated attempts. Extensive modification of the disaccharide was not very feasible as the losses were compounded from the already low yields obtained during glycosylation. We realized that it was extremely important to reduce the number of modification steps after glycosylation to cut losses, and a new protection strategy was required to obtain CS-R.

Scheme 2-8. Synthesis of disaccharide **2-39d**[a]



[a]Reagents and conditions: (a) TMSOTf, CH_2Cl_2, 4Å MS, 1 h, 41%; (b) Bu_3SnH, AIBN, toluene, 80^oC, 4 h, 70%; (c) 80% AcOH, 80^oC, 1 h, 65%; (d) Bz_2O, DMAP, pyridine, 1 h, 30%; (e) $H_2NNH_2.CH_3COOH$, EtOH/toluene, 6 h, 30%

In our exploratory attempts, C-2 ester protecting groups were favoured as these groups could direct *β*-stereoselectivity during glycosylation. However, both the levulinate and chloroacetate esters were unsuccessful, and there were no other viable ester protecting groups which could be cleaved orthogonally in the

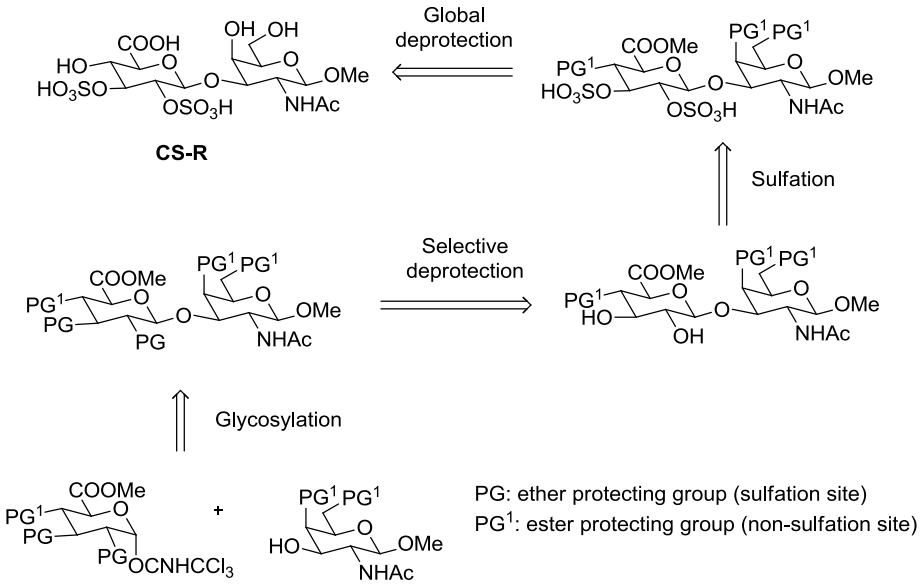
presence of other esters, for regioselective deprotection and sulfation of C-2 and C-3 hydroxyl groups for the synthesis of CS-R. Thus, other non-ester protecting groups had to be employed for the C-2 hydroxyl group. The loss of the C-2 ester to direct β-stereoselective glycosylation meant that other reaction conditions would have to be modified to control stereoselectivity. The failed synthesis of CS-R prompted us to re-evaluate our synthetic approach: it was too time consuming to explore and develop a new protection strategy for each subsequent analogue, and we needed to devise a common method for the synthesis of all 16 sulfation patterns possible in CS.

2.7 Sulfation Specific Protection Strategy

To devise a general method to obtain all 16 isomers possible in the CS repeating unit, it would be ideal if we had a pair of orthogonal protecting groups specifically introduced for the 4 sulfation sites, to either release the hydroxyl groups for sulfation to occur, or protect the hydroxyl groups against sulfation. In addition, the required protecting groups should be introduced onto the monomeric units to reduce the number of modification steps required in the disaccharide stage so as to improve reaction yields.

To achieve this, we decided to utilize the ester and ether protecting groups, which can be cleaved via different reaction conditions; this enables regioselective deprotection and sulfation of the required sites. Ether groups were first designated for the protection of hydroxyl sites which require sulfation, since the C-6 carboxyl

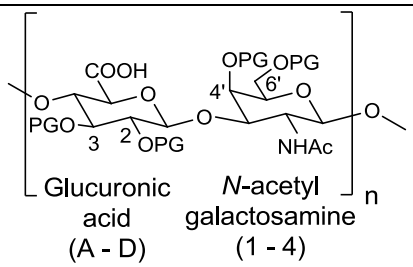
group (non-sulfation site) was protected as a methyl ester. In this approach, ether protected hydroxyl groups would first be liberated upon glycosylation to allow for sulfation to occur, while the ester protecting groups remain attached. After the sulfation step, the remaining protecting groups can then be cleaved in a global deprotection step to obtain the desired product.



Scheme 2-9. Retrosynthetic analysis for synthesis of CS-R

The retrosynthetic analysis of CS-R is depicted in scheme 2-9. In this strategy, CS-R would require ether protecting groups on C-2 and C-3 positions of the glucuronic acid moiety; these are selectively cleaved after glycosylation to enable sulfation on C-2 and C-3. The C-4' and C-6' hydroxyl groups (on *N*-acetyl galactosamine) which do not require sulfation would be orthogonally protected as esters, and only removed in the final global deprotection step, after sulfation has occurred.

Table 2-2. Sulfation specific strategy to obtain all 16 CS disaccharides



PG: **ether** or ester protecting group

ether PG: **sulfation** sites

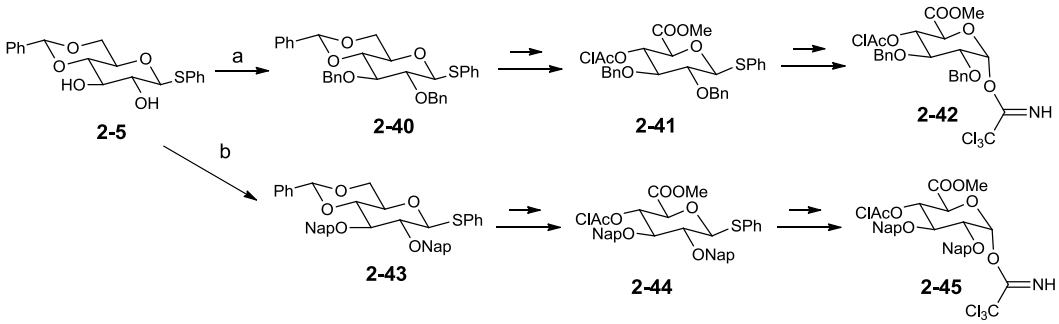
Glucuronic acid donor / *N*-acetyl galactosamine acceptor		**A**	**B**	**C**	**D**
		C-2 **ether**; C-3 **ether**	C-2 ester; C-3 **ether**	C-2 **ether**; C-3 ester	C-2 ester; C-3 ester
1	C-4’ **ether**; C-6’ **ether**	C-2, C-3, C-4’, C-6’ tetrasulfated	C-3, C-4’, C-6’ trisulfated	C-2, C-4’, C-6’ trisulfated	C-4’, C-6’ disulfated
2	C-4’ ester; C-6’ **ether**	C-2, C-3, C-6’ trisulfated	C-3, C-6’ disulfated	C-2, C-6’ disulfated	C-6’ monosulfated
3	C-4’ **ether**; C-6’ ester	C-2, C-3, C-4’ trisulfated	C-3, C-4’ disulfated	C-2, C-4’ disulfated	C-4’ monosulfated
4	C-4’ ester; C-6’ ester	C-2, C-3 disulfated	C-3 monosulfated	C-2 monosulfated	non-sulfated

To extend this approach, all 16 CS sulfation patterns can be obtained by manipulating the protecting groups present in the donor and acceptor building blocks. 4 possible protecting group combinations are possible for the glycosyl donor and acceptor, which give rise to a total of 16 possible CS disaccharides (Table 2-2). Preferably, a common intermediate should be used for the synthesis of the set of donor and acceptor building blocks, through a divergent approach.

To obtain the sulfation pattern required in CS-R which has sulfate groups on C-2 and C-3 of glucuronic acid, the monomeric precursors required are glycosyl donor **A** (C-2 and C-3 ether protection) and acceptor **4** (C-4' and C-6' ester protection). Unlike glycosyl donors **B** and **D** which have C-2 ester protecting groups to direct *β*-stereoselective glycosylation, glycosyl donors **A** and **C** utilize ether protecting groups on C-2 which are non-participating groups. Thus, other factors such as solvent effects would be required to direct *β*-stereoselectivity.

2.7.1 Synthesis of Glycosyl Donors 2-42, 2-45

Scheme 2-10. Synthesis of glycosyl donors **2-42** and **2-45**



Reagents and conditions: (a) BnBr, NaH, DMF, 3 h, 85%; (b) NapBr, NaH, DMF, 3 h, 80%

To prepare the glycosyl donors for the synthesis of CS-R, both the C-2 and C-3 hydroxyl groups were either protected as benzyl ethers (**2-42**) or naphthyl ethers (**2-45**) (Scheme 2-10). Protecting both the C-2 and C-3 of the glucuronic acid donor with benzyl ether or naphthyl ether functionality enables regioselective deprotection by catalytic hydrogenation or DDQ oxidation respectively to free the

hydroxyl groups to direct site specific sulfation at C-2 and C-3 during later steps of synthesis.

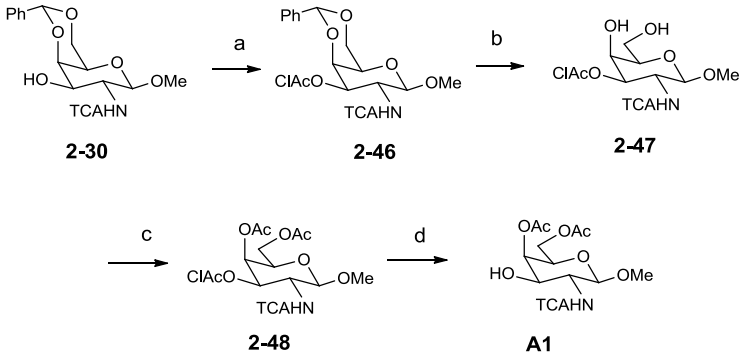
Intermediate **2-5** was first reacted with benzyl bromide and sodium hydride to introduce the benzyl ether at the C-2 and C-3 position in intermediate **2-40**. Similarly, both the C-2 and C-3 hydroxyl groups on **2-5** could be protected as naphthyl ethers by reaction with naphthyl bromide and sodium hydride to obtain intermediate **2-43**. Both **2-40** and **2-43** were next subjected to acetal hydrolysis, primary alcohol oxidation, methylation and C-4 choloroacetylation to obtain thioglycosides **2-41** and **2-44**. An ester protecting group was utilised for the C-4 hydroxyl group as C-4 is a non-sulfation site which required orthogonal protection. Deprotection of the anomeric thiophenol group and attachment of the trichloroacetimidate auxiliary furnished glycosyl donors **2-42** and **2-45**. The loss of C-2 ester participating groups in these donors required a change in the glycosylation solvent to direct β-stereoselectivity through the use of participating solvents such as acetonitrile (MeCN).

2.7.2 Synthesis of Glycosyl Acceptor A1

In order to minimize the number of modification steps required after glycosylation, the required ester protecting groups on C-4' and C-6' were now introduced into glycosyl acceptor **2-30** (Scheme 2-11). The C-3' hydroxyl group was first protected as a chloroacetyl ester in **2-46**, prior to cleavage of the benzylidene acetal via acidic hydrolysis to release the C-4' and C-6' hydroxyl

groups; these were both acetylated to obtain **2-48**. The acidic cleavage of the benzylidene acetal however, was not very favorable due to the acid catalyzed ring opening at the anomeric methyl ether linkage, which converted some *β*-product into the *α* form and the reaction had to be carefully monitored (discussed later). To obtain glycosyl acceptor **A1**, the C-3' chloroacetyl ester in **2-48** was selectively deprotected by reaction with thiourea in ethanol/pyridine to regenerate the free hydroxyl group.[17]

Scheme 2-11. Synthesis of glycosyl acceptor **A1**[a]



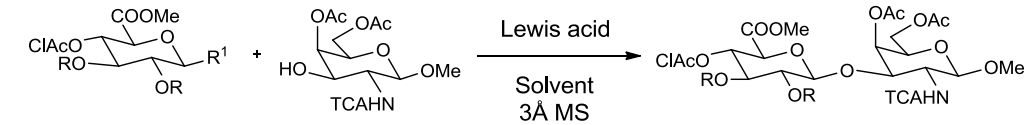
[a]Reagents and conditions: (a) Cl_2Ac_2O, pyridine, CH_2Cl_2, 1 h, 85%; (b) 80% AcOH, 80°C, 1 h, 60%; (c) Ac_2O, DMAP, pyridine, 1 h, 82%; (d) thiourea, pyridine/EtOH, 4 h, 80°C, 71%

2.7.3 Synthesis of Protected Disaccharides 2-49 and 2-50 via Glycosylation

Glycosyl donors **2-41** and **2-42** were first utilised for glycosylation to obtain disaccharide **2-49** (Table 2-3) as benzyl ether protecting groups could be cleanly cleaved via catalytic hydrogenation when required. This was preferred to naphthyl

ether protecting groups (in donors **2-44** and **2-45**) which require oxidative cleavage, resulting in the formation of water soluble salts that are difficult to remove.

Table 2-3. Glycosylation reaction to obtain protected disaccharides **2-49** and **2-50**



2-41: R = Bn; R^1 = SPh
2-42: R = Bn; R^1 = $OCNHCCl_3$
2-44: R = Nap; R^1 = SPh
2-45: R = Nap; R^1 = $OCNHCCl_3$

A1

2-49: R = Bn
2-50: R = Nap

Entry	Donor	Lewis acid (equiv.)	Solvent	Temp.	Product (%)
1	**2-41**	NIS,TfOH (1.68,0.63)	DCM	-20 to 0^{o}C	**2-49 (**43%) (α:β) =1:1
2	**2-41**	NIS,TfOH (1.68,0.63)	MeCN	-20 to 0^{o}C	-
3	**2-41**	NIS,TMSOTf (1.68,0.24)	DCM/MeCN/ EtCN (1:2:1)	-78 to -40 to -20 ^{o}C to r.t.	-
4	**2-42**	TMSOTf (1.1)	EtCN	-78 to -40 to -20 ^{o}C to r.t.	-
5	**2-42**	TMSOTf (1.1)	DCM/MeCN (1:3)	-40 to -20 to 0 ^{o}C	-
6	**2-44**	NIS,TfOH (1.68,0.63)	DCM	-20 to 0 ^{o}C	-
7	**2-45**	TMSOTf (0.2)	DCM	-20 to -10 ^{o}C	**2-50 (**10%) (α:β) =1:1

Since the C-2 benzyl ether was not a participating group, solvent effects were required to direct β-stereoselectivity of the glycosylation step. To ascertain reactivity of the donor and acceptor, we first reacted thioglycoside **2-41** and acceptor **A1** using the previously optimized solvent system dichloromethane (DCM), and disaccharide **2-49** was obtained in 43% yield after 1 hour (Table 2-3, entry 1). As expected, there was no stereoselectivity in the glycosylation step due to the loss of the C-2 ester group and both the α- and β- product were obtained in approximately equal amounts. Both anomers were structurally very similar and it was difficult to separate them by chromatography without compromising on the yield. Moreover, too much product was lost from the formation of the undesired anomer. To overcome this, we proceeded to direct β-stereoselectivity by modifying the reaction conditions, through the use of participating solvents.

Glycosylation was re-attempted with acetonitrile (MeCN), this solvent was postulated to participate by first displacing the auxiliary group on the donor to form a stable leaving group at the anomeric carbon of the glycosyl donor, in the axial position. Acceptor **A1** thus attacks the nitrilium intermediate from the equatorial position to displace the nitrile solvent, to form the β-glycosidic linkage through the solvent effect.[34-37] However, when donor **2-41** was reacted with acceptor **A1** using MeCN, no product was observed (Table 2-3, entry 2). Monitoring of the reaction by mass spectrometry did indicate formation of the nitrilium intermediate, but the nitrilium intermediate formed was too stable and the acceptor **A1** could not displace the nitrile group to form the glycosidic linkage.

Upon reaction workup, only the hydrolyzed donor with anomeric hydroxyl group could be identified, together with the unreacted acceptor **A1**, and no disaccharide was obtained.

To facilitate attack of the glycosyl acceptor to the nitrilium intermediate to form the glycosidic linkage, a stronger Lewis acid TMSOTf was used, to coordinate to the nitrile group and generate a stable leaving group. The reaction was re-attempted at -78^{o}C and a solvent mixture of DCM/MeCN/EtCN (1:2:1) was used; DCM and propionitrile (EtCN) were required as co-solvents as MeCN freezes at -45^{o}C. The reaction unfortunately again, did not proceed. The reaction temperature was gradually increased from -40^{o}C to -20^{o}C to room temperature, but no disaccharide could be detected after 24 hours (Table 2-3, entry 3).

We next modified the glycosyl donor to **2-42** to increase the donor reactivity for glycosylation to proceed, by replacement of the anomeric thiophenol group with the trichloroacetimidate auxiliary. More Lewis acid (1.1 equivalent TMSOTf) was also added to promote completion of the reaction; EtCN was used as the solvent since MeCN was not a suitable at -78^{o}C. Unfortunately, no disaccharide was obtained, even when the reaction temperature was raised from -40^{o}C to -20^{o}C to room temperature (Table 2-3, entry 4). Upon prolonged stirring, the nitrilum intermediate hydrolysed to form the hydroxyl group at the anomeric carbon. To push the acceptor to react and displace the nitrilium intermediate, the concentration of the nitrile solvent was subsequently reduced by changing the

solvent system to DCM/MeCN (1:3). However, no disaccharide was observed even when the reaction was warmed from -40^{o}C to -20^{o}C to 0^{o}C (Table 2-3, entry 5).

Since C-2 benzyl ether protecting groups were not suitable, we re-attempted using the C-2 naphthyl ether by utilising glycosyl donors **2-44** and **2-45**. Similarly, we first reacted glycosyl donors **2-44** and **2-45** with acceptor **A1** using previously optimized conditions in dichloromethane solvent to ascertain the donor reactivity (Table 2-3, entries 6 and 7). Unfortunately, the glycosylation steps gave very low reaction yields, and only trace amount of the dimer could be obtained when donor **2-45** was used. The low reaction yield was further compounded by the lack of β-stereoselectivity during glycosylation due to the loss of C-2 participating group. Poor reactivity of the donors **2-44** and **2-45** could be attributed to the bulky naphthyl ether at the C-2 position, which caused steric hindrance, blocking the attack of the glycosyl acceptor at the anomeric position. Thus naphthyl ether protecting group was also not suitable for C-2 hydroxyl group protection.

The use of nitrile solvents to direct β-stereoselectivity have been reported to be highly substrate dependent, and different reaction yields and stereoselectivities have been reported for different donors, depending on the structure of the protected glycosyl donors.[34, 37] Unfortunately, this method could not be applied to the protection strategy utilised for our monomeric building blocks. The nitrilium intermediate generated was too stable for glycosylation to proceed; no

disaccharide product could be obtained when glycosyl donors **2-41** and **2-42** were reacted with acceptor **A1** using nitrile solvents. The sulfation specific protection strategy thus had to be further fine-tuned to account for β-stereoselective glycosylation, which is the required stereochemistry in CS GAGs.

2.8 Modifying the Sulfation Specific Protection Strategy

To obtain all the 16 sulfation patterns possible in CS through a more general approach, a sulfation specific protection strategy was devised, which incorporated the use of ether and ester protecting groups as orthogonal handles to direct sulfation. One of these protecting groups was denoted as sulfation sites, and cleaved first to release the free hydroxyl groups for sulfation, prior to global deprotection of the remaining protecting groups.

From our failed glycosylation attempts, it was noted that ether protecting groups were not suitable for the C-2 position of the glycosyl donor as we were not able to direct β-stereoselective glycosylation though the use of nitrilium solvents. Due to the lack of options to direct β-stereoselectivity, we had to revisit the use of C-2 participating groups for the glucuronic acid donor. Since our attempts at identifying orthogonal ester protecting groups (such as chloroacetyl and levulinyl) were also not successful, we decided to return to the use of benzoyl ester for C-2 protection as this group was able to direct β-stereoselective glycosylation and also make the donor sufficiently reactive for glycosylation to proceed in good yields, as shown from the synthesis of CS-A, CS-C, CS-E, CS-O, CS-K and CS-M.

However in CS, the C-2 position of glucuronic acid might require attachment of sulfate groups. Thus when we fixed the C-2 protecting group as the benzoyl ester to direct β-stereoselective glycosylation, esters now need to serve as both sulfation and non-sulfation protection handles, depending on which CS isomer is synthesized. In the case when C-2 sulfation is required, esters will be designated as sites of sulfation (since C-2 benzoyl ester is mandatory). On the other hand, when C-2 sulfation is not required, ethers now become the designated as sites of sulfation.

2.8.1 Protection Strategy for Glycosyl Donors

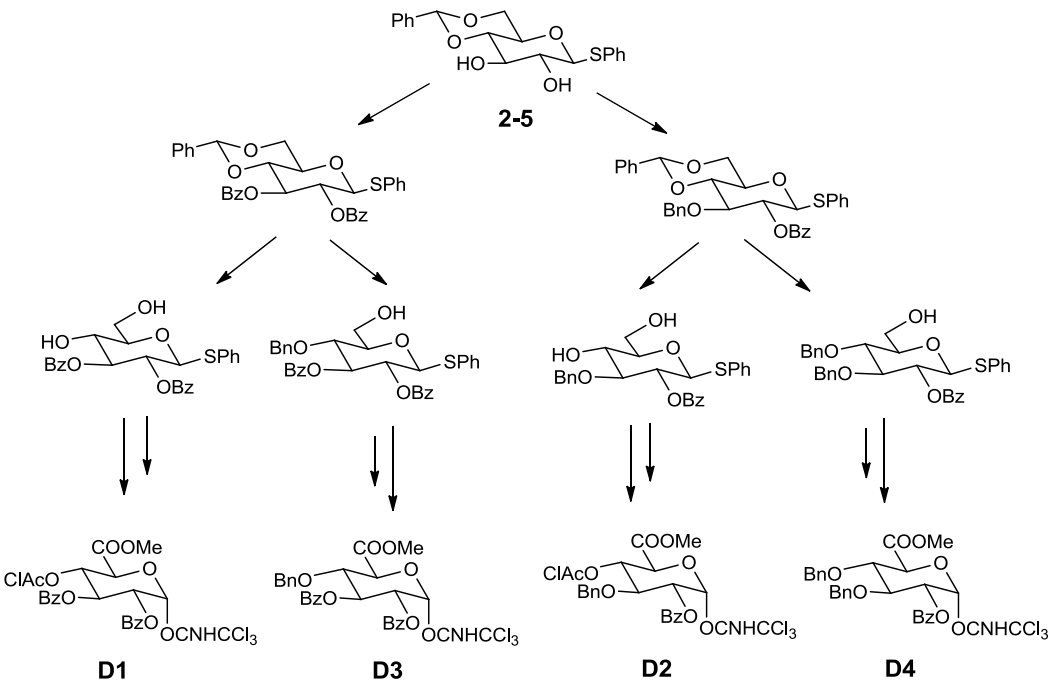


Figure 2-3. Glycosyl donors for CS disaccharide library synthesis

Four different glycosyl donors were thus proposed, which can be synthesized from common intermediate **2-5** (Figure 2-3). In all four glycosyl donors, C-2 is always protected as a benzoyl ester, this serves to direct β-stereoselective glycosylation. On the other hand, the C-3 hydroxyl group can either be protected as a benzyl ether or benzoyl ester.

The C-1 position retains the trichloroacetimidate auxiliary which functions as the leaving group during glycosylation, and the C-6 carboxyl group will still require protection as a methyl ester.

The C-4 position of the glucuronic acid donor is the site of elongation and is never sulfated, thus the protecting group chosen for C-4 must be orthogonal to protecting groups utilised for sulfation sites. If ethers are used to mask the sulfation sites, ester protecting group would be required on C-4, but if ester groups denote sulfation sites, ether protection would be needed on C-4.

For donors **D1** and **D2**, benzyl ether protected hydroxyl groups are marked as sulfation sites, hence C-4 was protected with the orthogonal chloroacetyl ester. Donors **D1** and **D2** form analogues which do not have sulfate groups attached at C-2.

For donors **D3** and **D4**, ester protected hydroxyl groups are denoted as sulfation sites, hence C-4 was protected with the orthogonal benzyl ether. Donors **D3** and **D4** form CS analogues which have sulfate groups attached at C-2.

2.8.2 Protection Strategy for Glycosyl Acceptors

Four different glycosyl acceptors were also proposed, these can be prepared from intermediate **2-30** which was previously synthesized. We decided to retain the C-1' methyl ether and the C-2' *N*-trichloroacetyl group (which assisted in solubility and reactivity of the acceptor during glycosylation), but modify the benzylidene acetal since the C-4' and C-6' positions were possible sites of sulfation. Through regio-reductive ring opening of the benzylidene acetal, we could direct formation of the benzyl ether on either C-4' or C-6'. This introduces sulfation specific protecting groups on the glycosyl acceptor at the C-4' and C-6' positions: either as esters, benzyl ethers or a combination of both (Figure 2-4).

The C-3' hydroxyl group should first be protected prior to modification of the C-4' and C-6' groups. Since C-3' is the site of elongation, this temporary protecting group must be cleaved in the final step to generate the free C-3' hydroxyl group in acceptors **A1 – A4** for reaction with the glycosyl donor to form the glycosidic linkage.

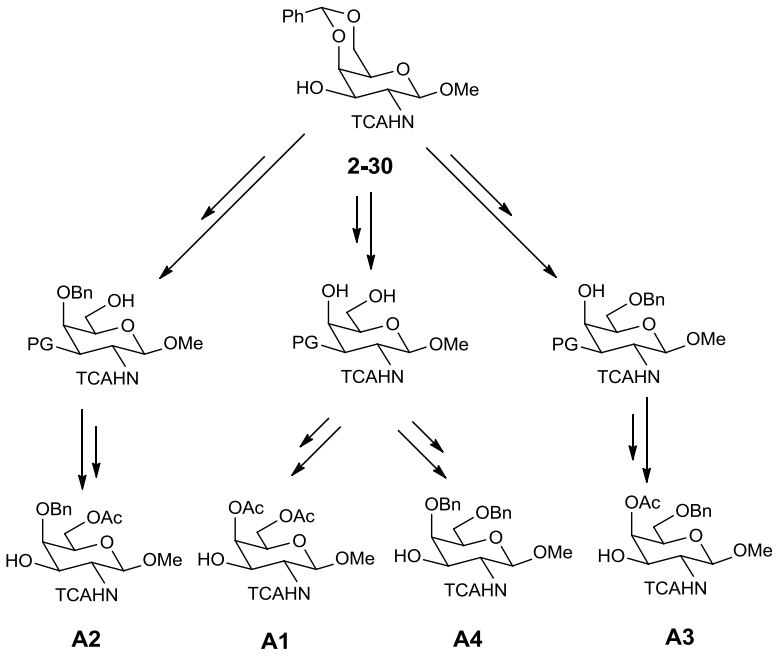


Figure 2-4. Glycosyl acceptors for CS disaccharide library synthesis

2.8.3 Synthetic Strategy to Obtain CS Disaccharide Library

In this strategy, sulfation specific protecting groups are first incorporated into the monomeric units; these precursors were synthesized in a divergent manner using common intermediates. With glycosyl donors **D1**–**D4** and glycosyl acceptors **A1**–**A4** on hand, any sulfation pattern required in the final CS disaccharide can be obtained by appropriate combination of donor and acceptor building blocks, through modular glycosylation and selective final transformations. Due to the need for the C-2 ester group to direct β-stereoselective glycosylation, both the ester and benzyl ether protecting groups may be required as sulfation sites. The choice is dependent on whether C-2 sulfation is required in the final product.

As shown in Table 2-4, eight plausible CS disaccharides (**P1 – P8**) can be obtained when benzyl ether protecting groups are designated the sulfation sites, by employing **D1** and **D2** as the glycosyl donors and selectively coupling with the required acceptors **A1** – **A4**. In addition, another 8 unique CS disaccharides (**P9 – P16**) can be prepared when ester protecting groups are designated as the sulfation sites, through the use of glycosyl donors **D3** and **D4** and acceptors **A1** – **A4** (Table 2-5). Both approaches enable the preparation of all 16 theoretically plausible CS disaccharides; these include previously synthesized CS sulfation patterns CS-A, CS-C, CS-E, CS-O, CS-K and CS-M, and novel sulfation patterns which were not isolated or synthesized earlier. Our synthetic strategy thus expands the scope of CS motifs obtainable via chemical synthesis, allowing access to rare sulfate motifs which may encode important functional information in biological systems.

Table 2-4. Structures of CS disaccharides when benzyl ether PGs are sulfation sites

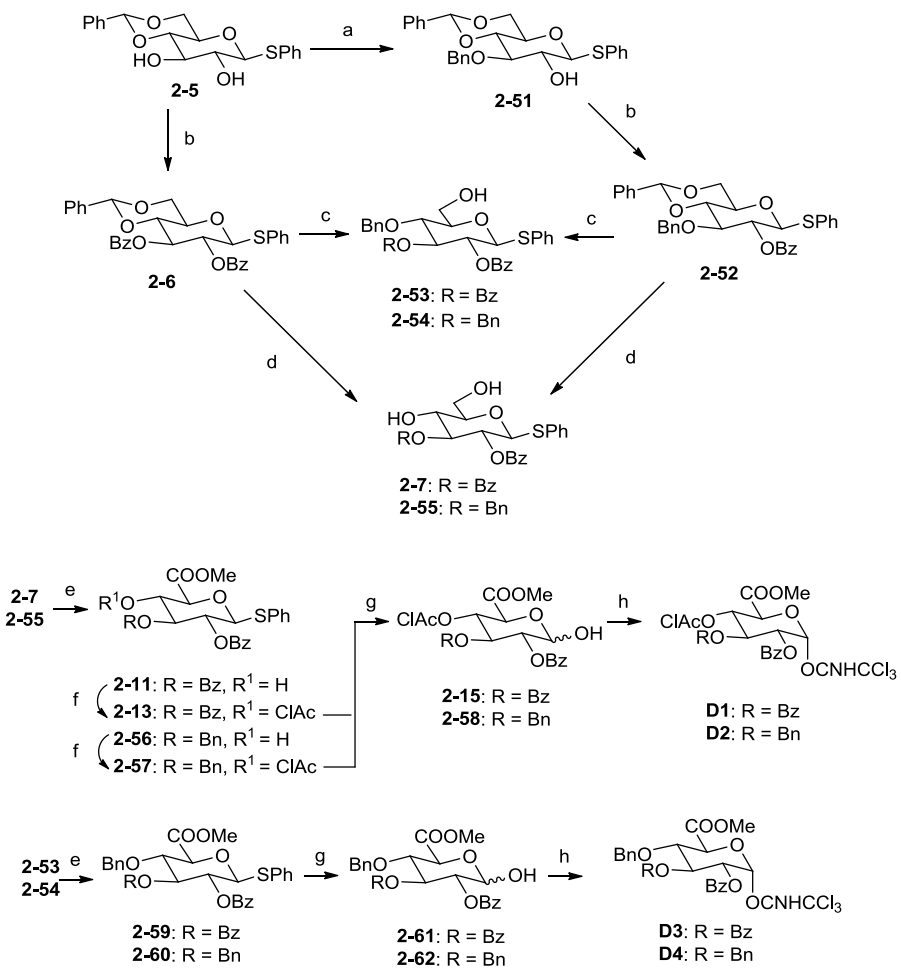
Donor / Acceptor	D1	D2
A1	P1	P5
A2	P2	P6
A3	P3	P7
A4	P4	P8

Table 2-5. Structures of CS disaccharides when ester PGs are sulfation sites

Donor / Acceptor	D3	D4
A1	P9	P13
A2	P10	P14
A3	P11	P15
A4	P12	P16

2.9 Synthesis of Glycosyl Donors D1, D2, D3, D4

Scheme 2-12. Synthesis of glycosyl donors **D1, D2, D3, D4**[a]



[a]Reagents and conditions: (a) Bu_2SnO, toluene, 3 h, 90°C, then BnBr, CsF, DMF, 14 h, 65%; (b) Bz_2O (1.2 equiv. per –OH), DMAP, CH_2Cl_2, 1 h, **2-6**: 88%, **2-52**: 85%; (c) BH_3.THF, $CoCl_2$, 2 h, **2-53**: 72%, **2-54**: 75%; (d) TFA, CH_2Cl_2/H_2O, 12 h, **2-7**: 84%, **2-55**: 88%; (e) TEMPO, BAIB, CH_2Cl_2/H_2O, 4 h, then MeI, $NaHCO_3$, TBAI, DMF, 1 h, 50°C, **2-11**: 48%, **2-56**: 55%, **2-59**: 50%, **2-60**: 52%; (f) Cl_2Ac_2O, pyridine, CH_2Cl_2, 1 h, **2-13**: 89%, **2-57**: 92% (g) NIS, TFA, CH_2Cl_2/H_2O, 14 h, **2-15**: 74%, **2-58**: 70%, **2-61**: 77%, **2-62**: 75%; (h) Cl_3CCN, DBU, CH_2Cl_2, 8 h, **D1**: 68%, **D2**: 75%, **D3**: 70%, **D4**: 78%

To obtain glycosyl donors **D1** – **D4**, common intermediate **2-5** was used as the synthetic precursor (Scheme 2-12). Through a divergent approach, the protecting groups present on C-2, C-3 and C-4 sites could be varied by modifying some of the key synthetic steps. C-2 ester protection was required for all 4 donors to direct *β*-stereoselective glycosylation, and the C-3 hydroxyl group could either be protected as a benzoyl ester (to obtain **2-6**) or selectively protected as a benzyl ether via the dibutyl tin oxide mediated approach[22] (to obtain **2-52**). The benzylidene acetal protecting group in intermediates **2-6** and **2-52** were next subjected to different cleavage methods, either by regio-reductive ring opening using to obtain the C-4 benzyl ether in **2-53** and **2-54**, or completely removed via acidic hydrolysis to liberate both C-4 and C-6 hydroxyl groups in **2-7** and **2-55**.

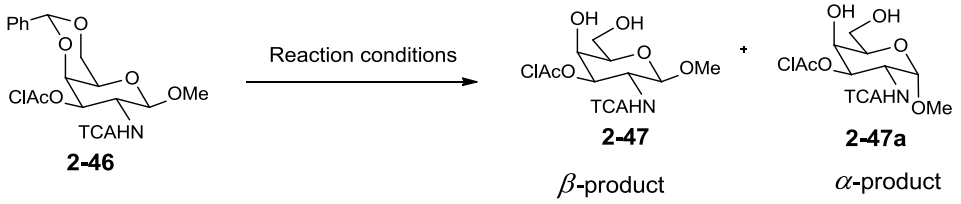
For **D3** and **D4**, ester protected hydroxyl groups were denoted as sulfation sites; hence an orthogonal benzyl ether protecting group was required on C-4, which was achieved by regio-reductive ring opening of the benzylidene acetal protecting group using $CoCl_2$ and BH_3.THF as the reductant[38, 39] to form **2-53** and **2-54**. Complete regioselective ring opening was achieved to obtain the benzyl ether on the C-4 position; retaining the benzylidene group on C-4 reduced the number of steps required which improved overall synthetic yield. Conversely, **D1** and **D2** mark benzyl ether protected hydroxyl groups as sulfation sites and hence require the orthogonal ester protection at C-4; chloroacetyl ester was attached to form **2-13** and **2-57**, this protection can be selectively cleaved to allow for the synthesis of longer CS fragments when required[12, 25]. Using synthetic methods highlighted

previously, intermediates **2-7**, **2-53**, **2-54**, **2-55** were subjected to C-6 oxidation and carboxylate methylation, followed by C-4 chloroacetylation (for intermediates **2-11**, **2-56**) prior to anomeric thiophenol deprotection and attachment of the trichloroacetimidate glycosyl auxiliary to furnish donors **D1**–**D4**.

2.10 Synthesis of Glycosyl Acceptors A1, A2, A3, A4

To obtain the four glycosyl acceptors, previously synthesized intermediate **2-46** was used as the common intermediate. Modification of the benzylidene acetal protecting group enabled attachment of sulfation specific protecting groups on C-4' and C-6' positions. The C-3' chloroacetyl ester could be selectively cleaved in the final step to regenerate the acceptor.

Table 2-6. Hydrolysis of benzylidene acetal in **2-46**



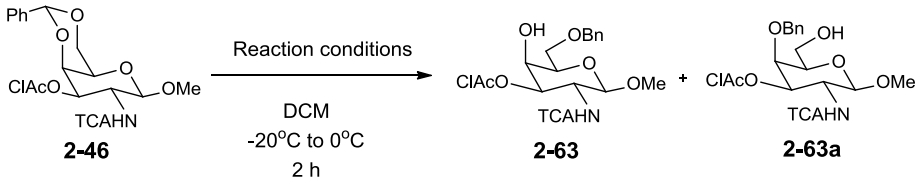
Entry	Reaction conditions	Yield (β:α)	Time	Temp. (^{o}C)
1	80% AcOH	70:30	1 h	80
2	TFA, CH_2Cl_2/H_2O	-	24 h	r.t.
3	Amberlyst H^+ resin, CH_2Cl_2	-	24 h	r.t.
4	H_2, Pd/C, THF	-	12 h	r.t.
5	60% AcOH	90:10	15 min	110, mw

For acceptors **A1** and **A4**, the benzylidene acetal was first removed by acidic hydrolysis (Table 2-6, entry 1) to release the free hydroxyl groups on C-4' and C-6' positions. Both hydroxyl groups can next be protected as esters (to obtain **A1**) or benzyl ether groups (to obtain **A4**). However, acidic hydrolysis of the benzylidene acetal by refluxing **2-46** in 80% acetic acid led to the unwanted side reaction: acid catalyzed ring opening of the cyclic sugar **2-46** at anomeric carbon to form the open chain (as mentioned previously). Upon ring cyclization, both the α- and β- anomers were obtained, and some β- anomer was converted into the α-form (**2-47a**). Although this side reaction could be minimized by reducing the reaction time, some α- anomer was still formed as the side reaction proceeded concurrently with the deprotection step. Milder reaction conditions were next attempted, by reacting **2-46** with trifluoroacetic acid in CH_2Cl_2/H_2O at room temperature, and by using by H^+ resins, but deprotection did not proceed (Table 2-6, entries 2 and 3). In addition, hydrogenation of the benzylidene acetal via palladium catalyst was unsuccessful (Table 2-6, entry 4). Hence, the reaction was repeated by reducing the reaction time to 15 minutes, by refluxing **2-46** in 60% acetic acid at 110^oC through the use of microwave irradiation (Table 2-6 entry 5). This reduced α- anomer formation to smaller amounts which improved reaction yield.

We next proceeded to modify **2-46** for the synthesis of acceptors **A2** and **A3**, which required the benzyl ether protecting group on either the C-4' or C-6' positions. To achieve this, the benzylidene acetal protecting group was modified

via regio-reductive ring opening to retain the benzyl ether on either the C-4' or C-6' position, before protecting the other hydroxyl group with an ester group.

Table 2-7. Regio-reductive ring open of benzylidene ring in **2-46**



Entry	Reaction conditions	Product
1	sodium cyanoborohydride, iodine	**2-63** and **2-63a**
2	triethylsilane, iodine	**2-47**
3	triethylsilane, TfOH	-

Regio-reductive ring opening of the benzylidene acetal protecting group was first attempted using sodium cyanoborohydride and molecular iodine (Table 2-7, entry 1), which was reported to yield C-6 benzyl ether in hexopyranosides[40]. However when the conditions were applied to intermediate **2-46**, no regioselectivity was observed, and both C-4 and C-6 benzyl ethers were formed in equal amounts. Other stronger reducing agents such as $NaBH_4$ and $LiAlH_4$ were not suitable as they were not compatible with the ester and amide protecting groups present in intermediate **2-46**.

Hence, we varied the reducing agent to triethylsilane to determine if regioselective reduction can be obtained, through the similar activation by iodine.

Unfortunately, complete cleavage of the benzylidene ring proceeded to form the diol **2-47** (table 2-7, entry 2), and no benzyl ether could be obtained. This reaction was also attempted using triethylsilane/triflic acid system but was unsuccessful (table 2-7, entry 3); side products were formed as a result of the incomplete reduction of the chloro groups on C-2' and C-3' positions. It was noted that the chloroacetyl groups present were incompatible with the triethylsilane reducing agent. Since intermediate **2-46** was not versatile as a synthetic precursor, we had to redesign the synthetic strategy for the divergent synthesis of the four glycosyl acceptors.

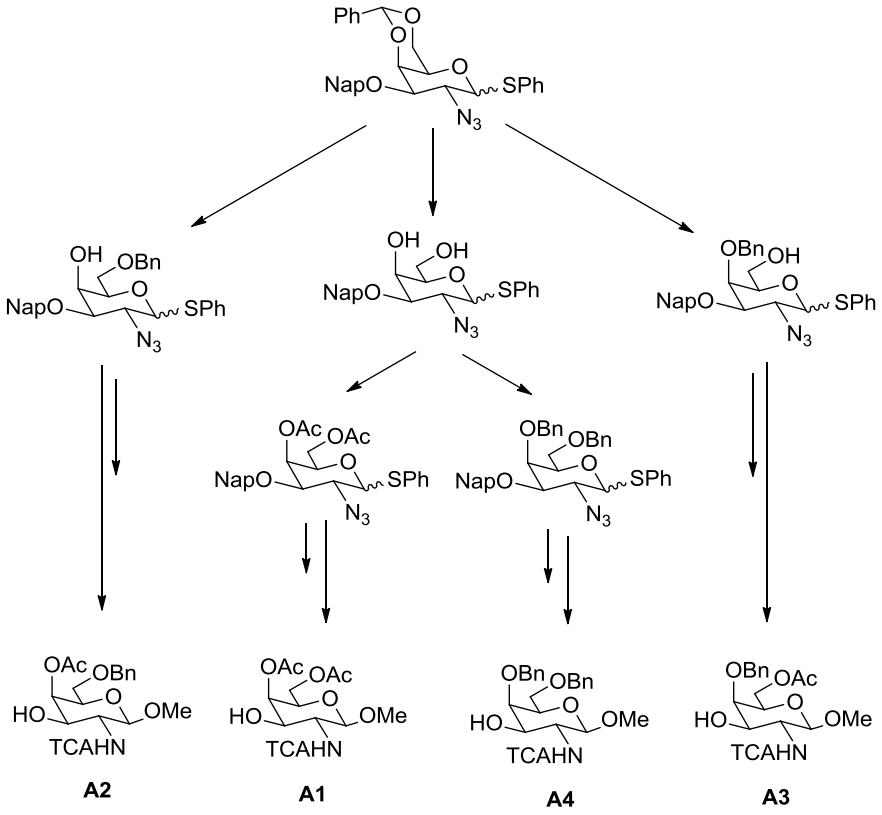
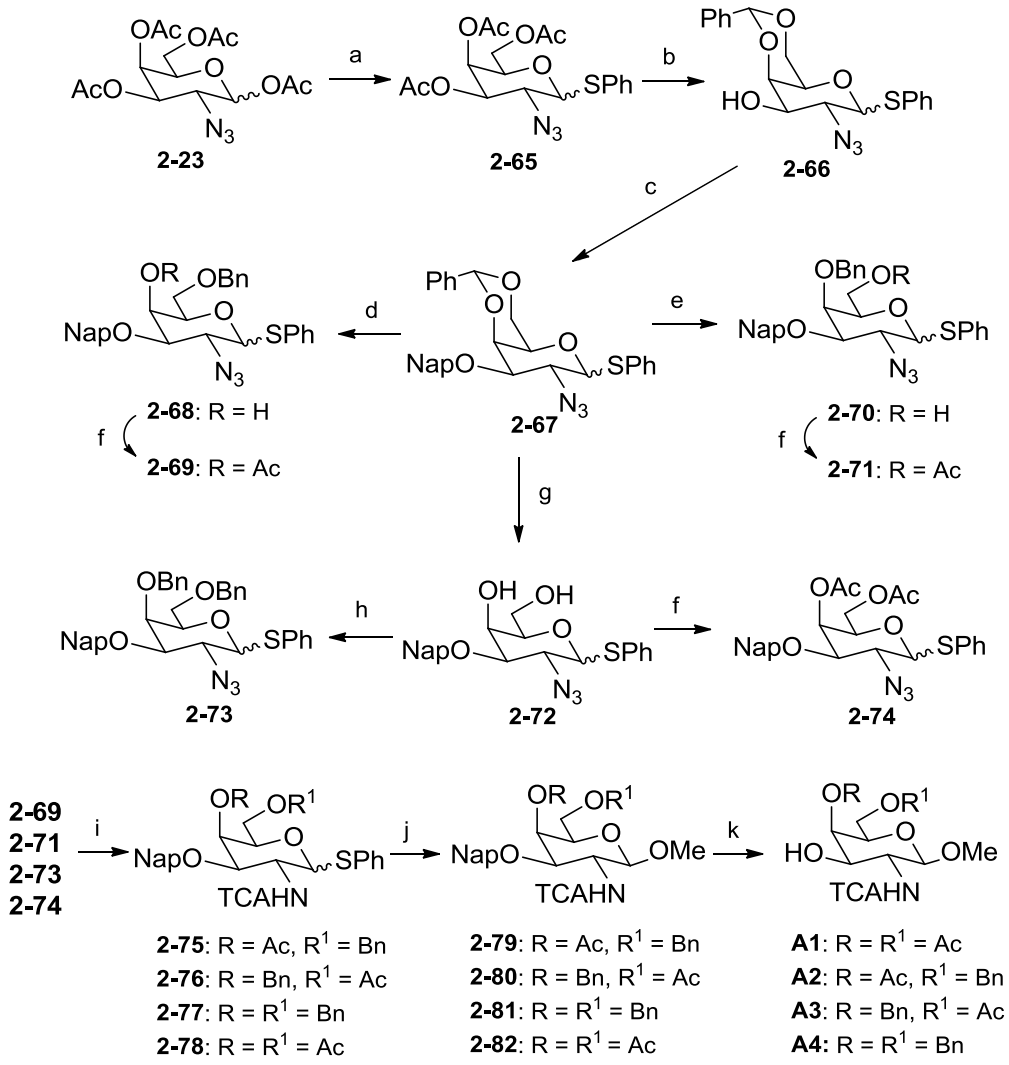


Figure 2-5. Proposed new route for the synthesis of acceptors **A1 – A4**

Several factors were identified from the failed attempts and these had to be addressed in the new route. Firstly, the C-1' methyl ether group was not stable to acidic reflux which was required for benzylidene acetal hydrolysis (to obtain acceptor **A1** and **A4**). Secondly, the C-2' trichloroacetyl group and C-3' chloroacetyl ester groups were not compatible with the reducing agents required for regio-reductive ring opening (to obtain acceptor **A2** and **A3**).

Thus, a new intermediate was proposed to obtain acceptors **A1** – **A4** (Figure 2-5). The required protecting groups on the C-4' and C-6' positons was first introduced via modification of the benzylidene acetal (acidic hydrolysis/regio-reductive ring opening), prior to the attachment of the C-1' methyl ether group and the C-2' trichloroacetyl group. In addition, the chloroacetyl protecting group on C-3' was also replaced with a naphthyl ether to ensure reaction compatibility. In the revised strategy, the inert protecting groups chosen for the C-1', C-2' and C-3' positions were the thiophenol, azide and naphthyl ether protecting groups respectively; these were used as temporary protecting groups to prepare glycosyl acceptors **A1** – **A4**. To achieve this, **2-23** was used as the synthetic precursor, which was obtained from D-galactose over 5 steps as previously described. **2-23** was first subjected to anomeric thiophenol protection to form **2-65** (Scheme 2-13), prior to removal of the acetyl protecting groups and attachment of the benzylidene acetal on C-4' and C-6' to form **2-66**. With C-1' protected as a thiophenol group and C-2' protected as an azide group, the benzylidene acetal was further modified to obtain the protecting groups required in acceptors **A1** – **A4**.

Scheme 2-13. Synthesis of glycosyl acceptors **A1 – A4**[a]



[a]Reagents and conditions: (a) PhSH, $BF_3.OEt_2$, CH_2Cl_2, 14 h, 78%; (b) NaOMe, MeOH, 30 mins, then $PhCH(OMe)_2$, CSA, MeCN, 60°C, 14 h, 74%; (c) NapBr, NaH, DMF, 3 h, 81%; (d) TES, TfOH, CH_2Cl_2, -78°C, 3 h, 78%; (e) TES, $PhBCl_2$, CH_2Cl_2, -78°C, 3 h, 68%; (f) Ac_2O, DMAP, CH_2Cl_2, 1 h, **2-69**: 88%, **2-71**: 89%, **2-74**: 85%; (g) 80% AcOH, 80°C, 1 h, 90%; (h) BnBr, NaH, DMF, 75%; (i) Zn, NH_4Cl, $EtOH/H_2O$, 80°C, 1 h then Cl_3COCl, TEA, THF, 30 mins, **2-75**: 54%, **2-76**: 58%, **2-77**: 55%, **2-78**: 60%; (j) NIS/TMSOTf, MeOH, CH_2Cl_2, -10°C, 1 h, **2-79**: 75%, **2-80**: 70%, **2-81**: 74%, **2-82**: 68%; (k) DDQ, H_2O, CH_2Cl_2, 12 h, **A1**: 70%, **A2**: 69%, **A3**: 72%, **A4**: 68%

The naphthyl ether was first attached on the C-3' hydroxyl group by reacting **2-66** with naphthyl bromide and sodium hydride to obtain common intermediate **2-67**. This bulky protecting group was chosen to assist in the regio-reductive ring opening of benzylidene acetal protecting group. The phenyl ring from the benzylidene acetal and naphthyl groups provided a steric block at C-4' site during ring opening of the benzylidene acetal. Coordination to the C-4' oxygen is not thus favorable if a bulky Lewis acid is used, enabling regioselective benzylidene ring opening by steric control.

To obtain the C-4' benzyl ether, we first attempted the regio-reductive ring opening method utilised for synthesis of glycosyl donors **D3** and **D4**. However, using the BH_3.THF/$CoCl_2$ system[39] gave very low yields of product **2-70** (about 30%) and thus we had to modify our approach. Through the use of triethylsilane reducing agent, we proceeded to direct regio-selectivity of the benzylidene ring opening by varying the Lewis acids used.

To obtain the C-4' benzyl ether **2-70**, common intermediate **2-67** was first reacted with a sterically demanding electrophile $PhBCl_2$ in the presence of triethylsilane reductant.[41] Coordination of $PhBCl_2$ occurred at the less hindered C-6' oxygen and the triethylsilane hydride was regioselectively added at C-6' to form the C-6' hydroxyl group. The benzyl ether was thus retained on C-4' to form intermediate **2-70**.

To obtain the C-6’ benzyl ether in **2-68**, common intermediate **2-67** was next reacted with triflic acid and triethylsilane.[41] Coordination of the sterically less hindered electrophile occurred at the C-4’ oxygen and the triethylsilane reductant was regioselectively added at C-4’ to form the C-4’ hydroxyl group; this generated the benzyl ether on C-6’ in intermediate **2-68**. Selectivity was favored by placing the bulky benzyl substituent on the equatorial position at C-6’ as opposed to the axial position at C-4’; this reduced ring strain arising from the 1,3-diaxial interactions.

Through this approach, we managed to direct regio-reductive ring opening of the benzylidene acetal in common intermediate **2-67** to obtain the benzyl ether on either the C-4’ or C-6’ position. The remaining hydroxyl groups were then protected as esters, to obtain the C-4’ and C-6’ protection required in acceptors **A2** and **A3**. Alternatively, the benzylidene acetal in common intermediate **2-67** could also be cleaved by acidic hydrolysis to form diol **2-72**. Both resulting hydroxyl groups are then protected as esters or ethers to obtain the C-4' and C-6' protection required in acceptors **A1** and **A4**.

With the key protecting groups established on C-4’ and C-6’, the C-2’ azide group in intermediates **2-69**, **2-71**, **2-73** and **2-74** were next reduced to a primary amine and the *N*-trichloroacetyl group (TCAHN) introduced onto the C-2’ position by reacting with trichloroacetyl chloride in the presence of TEA. This C-2’ TCAHN participating group next directed stereoselective glycosylation of methanol at the

anomeric position to form the β-methyl ether linkage. In the final step, cleavage of the C-3' naphthyl ether in **2-79**, **2-80**, **2-81** and **2-82** via DDQ oxidation furnished acceptors **A1**–**A4**.

2.11 Glycosylation of Monomeric Building Blocks

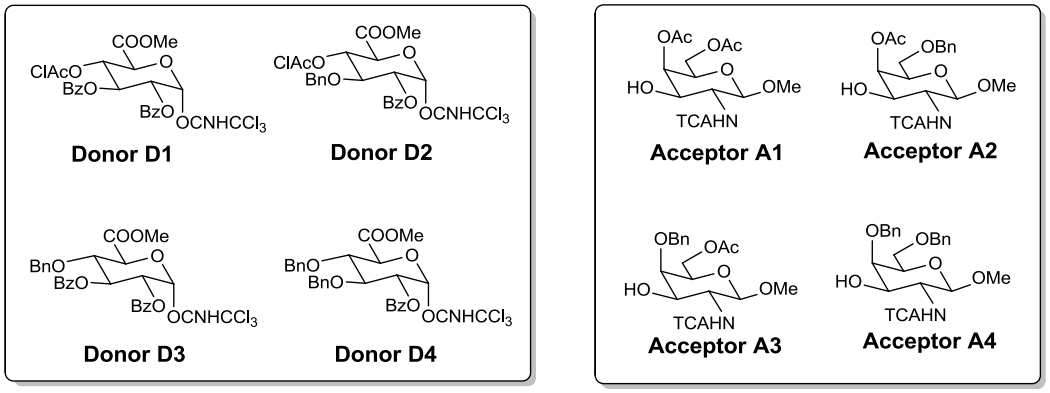


Figure 2-6. Glycosyl donors and acceptors designed for CS disaccharide library synthesis

With glycosyl donors **D1**–**D4** and glycosyl acceptors **A1**–**A4** on hand (Figure 2-6), any sulfation pattern required in the final CS disaccharide can be obtained by the judicious choice of donor and acceptor building blocks. This strategy can be classified into 2 main categories depending on the sulfation pattern required in the final CS disaccharide, which is illustrated in the synthesis of CS-R and CS-L.

For the synthesis of CS-R, sulfate groups are attached on the C-2 and C-3 positions of glucuronic acid moiety (Figure 2-7). Since glucuronic acid C-2 sulfation is required, ester protecting groups are denoted as sulfation sites (C-2 benzoyl ester mandatory to direct β-stereoselectivity). Hence only the C-2 and C-

3 hydroxyl groups of glucuronic acid require ester protection, with the other sites protected as benzyl ethers. Based on this analysis, donor **D3** and acceptor **A4** precursors are required to obtain CS-R.

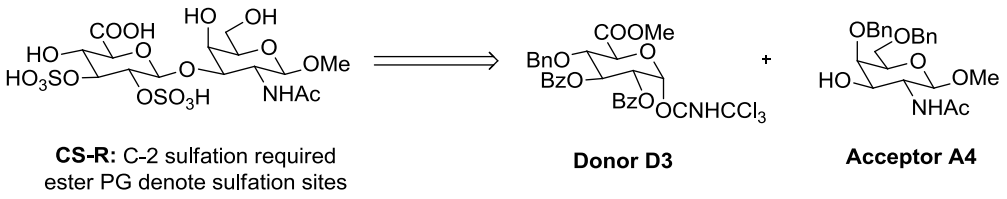


Figure 2-7. Retrosynthetic analysis for CS-R disaccharide

In CS-L, sulfation occurs on C-3 of the glucuronic acid and C-6' of the *N*-acetyl galactosamine (Figure 2-8). Since glucuronic acid C-2 sulfation is not required, benzyl ethers are now denoted as sulfation sites. Hence only C-3 glucuronic acid and C-6' *N*-acetyl galactosamine require benzyl ether protection, while the other sites are protected as esters. Donor **D2** and acceptor **A2** are the required precursors for CS-L.

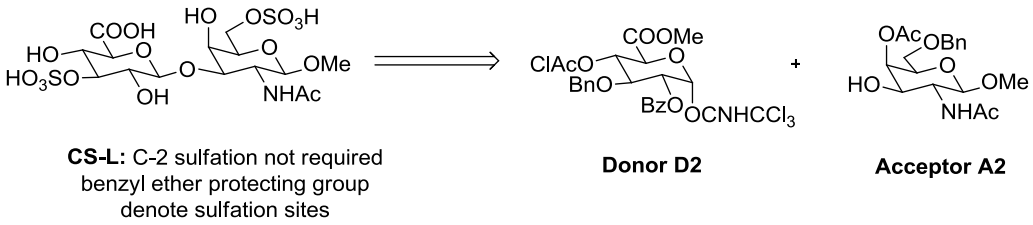


Figure 2-8. Retrosynthetic analysis for CS-L disaccharide

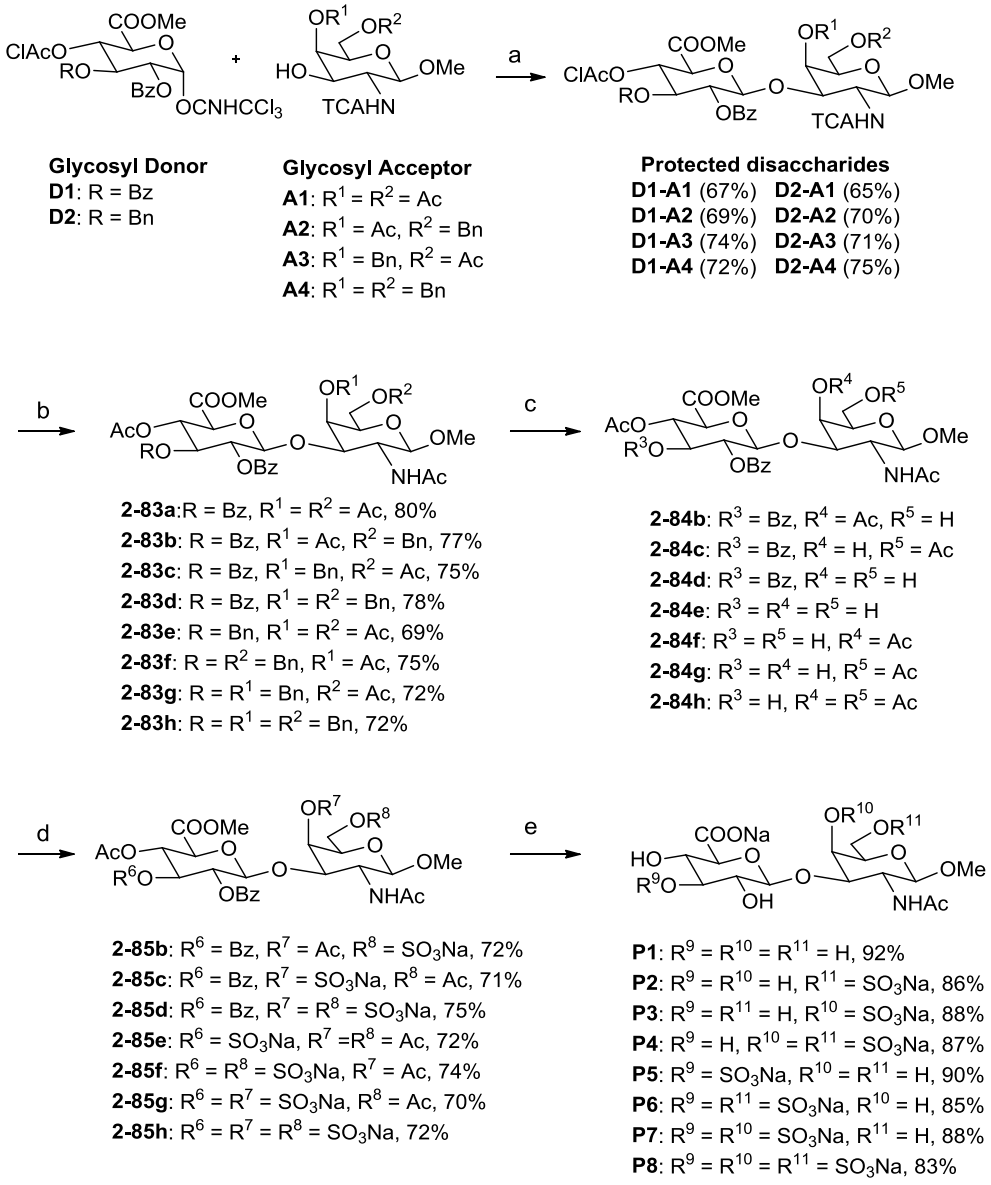
2.12 Synthesis of CS Disaccharide Analogues P1 to P8

In eight of the CS disaccharide analogues, C-2 sulfation is not required. To synthesize these eight compounds, glycosyl donors **D1** and **D2** were utilized, with benzyl ether protected hydroxyl groups denoted as sulfation sites. **D1** and **D2** employ the orthogonal chloroacetyl ester protecting group on the C-4 position.

Glycosyl donors **D1** and **D2** were first glycosylated with acceptors **A1**–**A4** using TMSOTf catalyst to form the protected disaccharides (Scheme 2-14). The C-2 directing group present in donors **D1** and **D2** enabled exclusive formation of the *β*-product[1]. After glycosylation was completed, the *N*-trichloroacetyl group was reduced to an *N*-acetyl group by radical mediated tributylstannane reduction to obtain **2-83a** to **2-83h**. Direct basic hydrolysis of **2-83a** formed non-sulfated CS disaccharide **P1**.

For the seven remaining intermediates **2-83b** to **2-83h**, hydrogenation of the benzyl ether groups liberated the free hydroxyl groups to form **2-84b** to **2-84h**, which were next reacted with sulfating agent SO_3.TEA to sulfate at the required sites to form **2-85b** to **2-85h**. Following this, global deprotection by basic hydrolysis liberated all the remaining ester protecting groups to form **P2** to **P8**. This approach formed a total of eight distinct CS disaccharides **P1** to **P8**.

Scheme 2-14. Synthesis of CS disaccharide analogues **P1** to **P8**[a]



[a]Reagents and conditions: (a) TMSOTf, 4Å MS, CH_2Cl_2, -10°C, 1 h; (b) Bu_3SnH, ABCN, toluene, 100°C, 4 h; (c) H_2 gas, Pd/C, CH_2Cl_2/MeOH 14 h; (d) SO_3.TEA, DMF, 50°C 40 h; (e) LiOH, H_2O_2, THF/H_2O, -5°C to r.t., 14 h, then NaOH, MeOH, 0°C to r.t, 6 h.

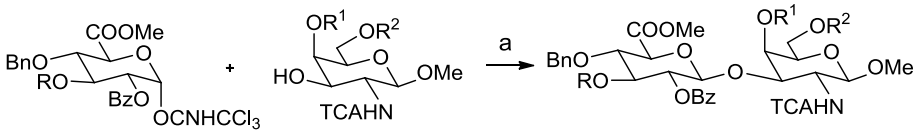
2.13 Synthesis of CS Disaccharide Analogues P9 to P16

In the other eight CS disaccharide analogues, C-2 sulfation is required. To obtain these eight analogues, glycosyl donors **D3** and **D4** were utilised, with ester protected hydroxyl groups denoted as sulfation sites. **D3** and **D4** employ the orthogonal benzyl ether protecting group on the C-4 position.

Glycosylation of **D3** and **D4** with the glycosyl acceptors **A1**–**A4** via TMSOTf catalyst formed another eight distinct protected disaccharides (Scheme 2-15). Presence of the C-2 directing group on **D3** and **D4** again enabled stereoselective formation of the *β*-glycosidic linkage. The *N*-trichloroacetyl groups present in the protected disaccharides were similarly reduced to the *N*-acetyl group in **2-86a** to **2-86h**. This was next followed by basic hydrolysis of the ester protecting group to release the hydroxyls groups[11] in **2-87a** to **2-87h**.

With the benzyl ether protecting groups intact, the free hydroxyl groups in **2-87b** to **2-87h** were then reacted with sulfating agent SO_3.TEA to sulfate at the desired sites to form intermediates **2-88b** to **2-88h**. Fortunately, the C-6 carboxylate group generated from the ester deprotection step did not affect the sulfation reaction. With the sulfate groups attached at the required positions, global deprotection by hydrogenation of the remaining benzyl ether groups proceeded in the final step to obtain CS disaccharides **P10** to **P16**.

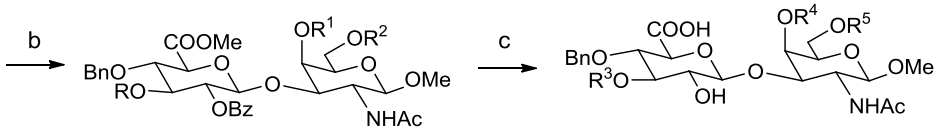
Scheme 2-15. Synthesis of CS disaccharide analogues **P10** to **P16**[a]



Glycosyl Donor
D3: R = Bz
D4: R = Bn

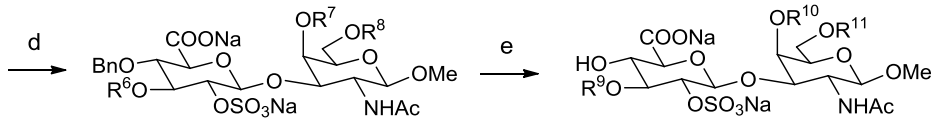
Glycosyl Acceptor
A1: $R^1 = R^2$ = Ac
A2: R^1 = Ac, R^2 = Bn
A3: R^1 = Bn, R^2 = Ac
A4: $R^1 = R^2$ = Bn

Protected dimer
D3-A1 (65%) **D4-A1** (64%)
D3-A2 (72%) **D4-A2** (68%)
D3-A3 (74%) **D4-A3** (69%)
D3-A4 (70%) **D4-A4** (72%)



2-86a: R = Bz, $R^1 = R^2$ = Ac, 78%
2-86b: R = Bz, R^1 = Ac, R^2 = Bn, 75%
2-86c: R = Bz, R^1 = Bn, R^2 = Ac, 77%
2-86d: R = Bz, $R^1 = R^2$ = Bn, 78%
2-86e: R = Bn, $R^1 = R^2$ = Ac, 69%
2-86f: $R = R^2$ = Bn, R^1 = Ac, 75%
2-86g: $R = R^1$ = Bn, R^2 = Ac, 79%
2-86h: $R = R^1 = R^2$ = Bn, 72%

2-87a: $R^3 = R^4 = R^5$ = H, 85%
2-87b: $R^3 = R^4$ = H, R^5 = Bn, 84%
2-87c: $R^3 = R^5$ =H, R^4 = Bn, 88%
2-87d: R^3 = H, $R^4 = R^5$ = Bn, 86%
2-87e: R^3 = Bn, $R^4 = R^5$ = H, 86%
2-87f: $R^3 = R^5$ = Bn, R^4 = H, 83%
2-87g: $R^3 = R^4$ = Bn, R^5 = H, 88%
2-87h: $R^3 = R^4 = R^5$ = Bn, 85%



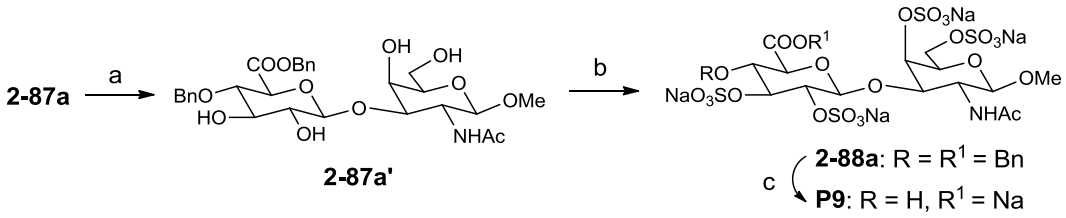
2-88b: $R^6 = R^8 = SO_3Na$, R^7 = Bn, 77%
2-88c: $R^6 = SO_3Na$, $R^7 = R^8$ = Bn, 75%
2-88d: $R^6 = R^7 = SO_3Na$, R^8 = Bn, 74%
2-88e: R^6 = Bn, $R^7 = R^8 = SO_3Na$, 68%
2-88f: $R^6 = R^8$ = Bn, $R^7 = SO_3Na$, 72%
2-88g: $R^6 = R^7$ = Bn, $R^8 = SO_3Na$, 73%
2-88h: $R^6 = R^7 = R^8$ = Bn, 75%

P10: $R^9 = R^{11} = SO_3Na$, R^{10} = H, 91%
P11: $R^9 = R^{10} = SO_3Na$, R^{11} = H, 94%
P12:$R^9 = SO_3Na$, $R^{10} = R^{11}$ = H, 92%
P13: R^9 = H, $R^{10} = R^{11} = SO_3Na$, 95%
P14: $R^9 = R^{11}$ = H, $R^{10} = SO_3Na$, 93%
P15: $R^9 = R^{10}$ = H, $R^{11} = SO_3Na$, 95%
P16: $R^9 = R^{10} = R^{11}$ = H, 92%

[a]Reagents and conditions: (a) TMSOTf, 4Å MS, CH_2Cl_2, -10°C, 1 h; (b) Bu_3SnH, ABCN, toluene, 100°C, 4 h; (c) LiOH, H_2O_2, THF, -5°C to r.t., 14 h, then NaOH, MeOH, 0°C to r.t., 6 h; (d) SO_3.TEA, DMF, 50°C, 40 h; (e) H_2 gas, Pd/C, MeOH/H_2O, 14 h.

For intermediate **2-87a**, sulfation of all four free hydroxyl groups made the tetrasulfated product very polar and difficult to isolate. Hence, an extra step was included where the carboxylate group was first protected as benzyl ester **2-87a'** (Scheme 2-16). This protecting group facilitated purification of the highly negatively charged tetrasulfated product. Sulfation of the free hydroxyl groups in **2-87a'** formed **2-88a**, and the benzyl groups were similarly cleaved by hydrogenation in the final step to form **P9**. A total of eight CS disaccharide analogues **P9** to **P16** were thus synthesized via this approach.

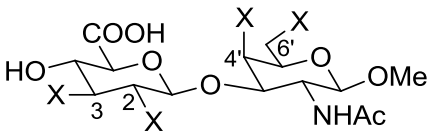
Scheme 2-16. Synthesis of CS disaccharide **P9**



[a]Reagents and conditions: (a) BnBr, $NaHCO_3$, DMF, 50°C, 89%, 1 h; (b) SO_3.TEA, DMF, 50°C, 68%, 48 h; (c) H_2 gas, Pd/C, MeOH/H_2O, 90%, 14 h.

The combination of both methods enabled the synthesis of all the 16 CS disaccharides which can be categorized based on the number sulfate groups present (Table 2-8). These included disaccharides which have already been reported and given nomenclature, in addition to novel sulfation motifs. The synthesized CS disaccharide library empowered us to systematically probe the "sulfation code" of CS in biological systems via structural activity relationship studies.

Table 2-8. CS disaccharide library



X = hydroxyl or sulfate groups

CS analogue		Position of sulfation	Conventional name
Non-sulfated	**P1**		CS-O
Mono-sulfated	**P16**	C-2	
	P5	C-3	
	P3	C-4’	CS-A
	P2	C-6’	CS-C
Di-sulfated	**P12**	C-2, C-3	CS-R
	P14	C-2, C-4’	
	P15	C-2, C-6’	CS-D
	P7	C-3, C-4’	CS-K
	P6	C-3, C-6’	CS-L
	P4	C4’, -C6’	CS-E
Tri-sulfated	**P11**	C-2, C-3, C-4’	
	P10	C-2, C-3, C-6’	
	P13	C-2, C-4’, C-6’	CS-T
	P8	C-3, C-4’, C-6’	CS-M
Tetrasulfated	**P9**	C-2, C-3, C-4’, C-6’	

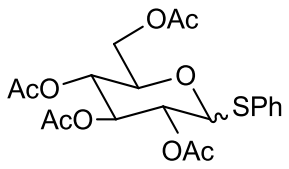
2.14 Conclusion

A versatile synthetic strategy has been devised for the chemical synthesis of all the sulfation patterns possible in the CS repeating unit. A total of 16 different CS disaccharides have been synthesized; these include analogues currently available as well as novel sulfation motifs. This strategy incorporated orthogonal protecting groups in four glycosyl donor and four glycosyl acceptor building blocks which were synthesized in a divergent mode using common intermediates. Through the glycosylation of appropriate donor and acceptor building blocks, any sulfation pattern required in the final CS disaccharide can be obtained by regioselective sulfation of either the benzyl ether protected hydroxyl groups or the ester protected hydroxyl groups to obtain the CS disaccharide library.

2.15 Experimental

General

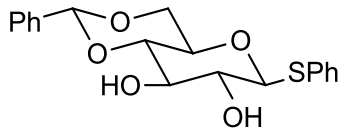
All commercially available reagents purchased from Sigma Aldrich and Alfa Aesar were used without further purification. Solvents such as hexane, ethyl acetate, dichloromethane, and methanol were pre-distilled prior to usage. Moisture-sensitive reactions were conducted under N_2 environment using commercially obtained anhydrous solvents and activated molecular sieves. Thin layer chromatography (TLC) was performed using Merck silica gel 60 F_{254} pre-coated glass plates and visualised under UV light prior to charring with 10% concentrated sulphuric acid in ethanol. Flash column chromatography was carried out with silica gel (Merck, 230 – 400 mesh); gel filtration chromatography was carried out using Sephadex® G-10 and Sephadex® C-25. ^{1}H and ^{13}C NMR spectrums were recorded on Bruker Avance 500 (DRX500) and Bruker Avance 500 (AV500) at 298K. All *J*-values are reported in Hz and chemical shift (δ) reported in parts per million (ppm) relative to tetramethylsilane (TMS). Mass spectra were determined by high resolution mass spectrometry (HRMS) electrospray ionization (ESI).



2-3

Compound 2-3: D-(+)-glucose (Compound **2-1**) (6 g, 33.3 mmol) and DMAP (101.8 mg, 0.83 mmol) were added to pyridine (16.1 mL, 200 mmol) and Ac_2O (18.9 mL, 200 mmol) and stirred at room temperature for 2 h. When TLC analysis (hexane/EtOAc, 1:1 v/v) showed completion of reaction, the reaction mixture was diluted with EtOAc and washed with 5M HCl (3x), followed by saturated $NaHCO_3$ (3x) and brine (1x). The organic layer was dried over Na_2SO_4, filtered and concentrated to a white solid (crude compound **2-2**), which was redissolved in CH_2Cl_2 (80 mL), followed by the addition of $BF_3{\cdot}OEt_2$ (12.3 mL, 100 mmol) and PhSH (10.3 mL, 100 mmol). The reaction mixture was stirred for 16 h at 35°C. When TLC analysis (hexane/EtOAc, 2:1 v/v) showed completion of the reaction, the reaction mixture was diluted with CH_2Cl_2, washed with $NaHCO_3$ (3x) and brine (1x), dried over Na_2SO_4, filtered, and concentrated. The product was purified by flash column chromatography (hexane/EtOAc, 4:1 towards 3:1 towards 2:1) to afford compound **2-3** as a white solid (11.88 g, 81%, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.52 – 7.46 (m, 2H), 7.35 – 7.27 (m, 3H), 5.22 (t, *J* = 9.4 Hz, 1H), 5.03 (t, *J* = 9.8 Hz, 1H), 4.97 (t, *J* = 9.7 Hz, 1H), 4.70 (d, *J* = 10.1 Hz, 1H), 4.19 (qd, *J* = 12.2, 3.8 Hz, 2H), 3.72 (ddd, *J* = 10.0, 5.1, 2.5 Hz, 1H), 2.07 (d, *J* = 3.8 Hz, 6H), 1.99 (d, *J* = 14.4 Hz, 6H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.5, 170.1, 169.3, 169.2, 133.1, 131.6, 128.9, 128.4, 85.7, 75.8, 73.9, 69.9, 68.2, 62.1,

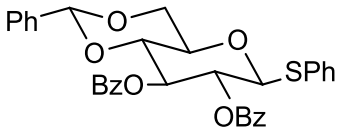
20.7, 20.7, 20.5, 20.5. HRMS (ESI): *m/z*: calcd for $C_{20}H_{24}NaO_9S$ [M + Na]: 463.1033; found: 463.1041.



2-5

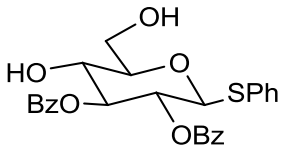
Compound 2-5: To a solution of compound **2-3** (10.77 g, 24.48 mmol) in methanol (100 mL) and CH_2Cl_2 (10 mL) was added NaOMe (529.4 mg, 9.8 mmol) and the mixture was stirred at room temperature for 1 h. When TLC analysis (hexane/EtOAc, 2:1 v/v) showed completion of reaction, the reaction mixture was quenched with Amberlyst-15 and stirred for another 15 min. When the pH was about 3 – 4, the mixture was filtered and concentrated to form a yellow oil (crude compound **2-4**), which was redissolved in MeCN (87.4 mL) and DMF (8.7 mL), followed by the addition of $PhCH(OMe)_2$ (7.4 mL, 48.96 mmol) and camphorsulfonic acid (1.86 g, 9.80 mmol). The reaction mixture was stirred for 16 h at 55°C. When TLC analysis (hexane/EtOAc, 1:1 v/v) showed completion of reaction, the reaction mixture was quenched with TEA (2.8 mL), followed by evaporation of solvent. The product was purified by flash column chromatography (hexane/EtOAc, 1.5:1 towards 1:1 towards 1:2) to afford compound **2-5** as a white solid (7.1 g, 80%, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.58 – 7.51 (m, 2H), 7.52 – 7.45 (m, 2H), 7.41 – 7.31 (m, 6H), 5.52 (s, 1H), 4.62 (dd, *J* = 10.0, 0.8 Hz, 1H), 4.37 (dd, *J* = 10.4, 3.1 Hz, 1H), 3.88 – 3.72 (m, 2H), 3.48 (dq, *J* = 22.0, 9.1 Hz, 3H), 2.99 (s, 1H), 2.81 (s, 1H). ^{13}C NMR (500 MHz,

$CDCl_3$) δ 136.8, 133.0, 131.3, 129.3, 129.1, 128.4, 128.3, 126.3, 101.9, 88.6, 80.2, 74.6, 72.6, 70.5, 68.5. HRMS (ESI): *m/z*: calcd for $C_{19}H_{20}NaO_5S$ [M + Na]: 383.0924; found: 383.0923.



2-6

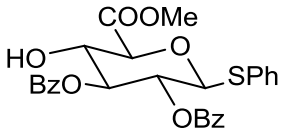
Compound 2-6: To a solution of compound **2-5** (6.52 g, 18.1 mmol) in CH_2Cl_2 (70 mL) was added Bz_2O (10.2 g, 45.3 mmol), DMAP (5.53 g, 45.3 mmol) and the reaction mixture was stirred for 2 h at room temperature. When TLC analysis (hexane/EtOAc, 6:1 v/v) showed completion of reaction, the reaction mixture was diluted with CH_2Cl_2 and washed with 5M HCl (3x), followed by saturated $NaHCO_3$ (3x) and brine (1x). The organic layer was dried over Na_2SO_4, filtered and concentrated to give crude compound **2-6** as a white solid which was immediately used for the next reaction to obtain compound **2-7**.



2-7

Compound 2-7: Crude compound **2-6** was redissolved in CH_2Cl_2 (100 mL) and TFA (2.75 mL) and water (5 mL) was added. The reaction mixture was stirred for 4 h at room temperature. When TLC analysis (hexane/EtOAc, 1:1 v/v) showed completion of reaction, the reaction mixture was diluted with CH_2Cl_2, washed

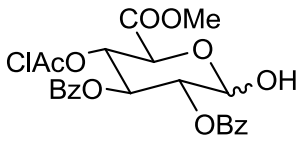
saturated $NaHCO_3$ (3x) and brine (1x). The organic layer was dried over Na_2SO_4, filtered, and concentrated. The product was purified by flash column chromatography (hexane/EtOAc, 3:1 towards 1:1) to afford compound **2-7** as a white solid (6.4 g, 74%, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.95 (dd, *J* = 13.0, 7.9 Hz, 4H), 7.55 – 7.42 (m, 4H), 7.37 (dt, *J* = 12.6, 7.7 Hz, 4H), 7.33 – 7.28 (m, 3H), 5.43 (dd, *J* = 7.4, 4.0 Hz, 2H), 5.01 – 4.91 (m, 1H), 4.02 (dd, *J* = 12.0, 3.1 Hz, 1H), 3.91 (ddd, *J* = 16.9, 13.7, 7.0 Hz, 2H), 3.67 – 3.55 (m, 1H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 167.5, 165.2, 133.6, 133.3, 132.7, 132.0, 130.0, 129.8, 129.2, 129.0, 128.8, 128.4, 128.4, 128.3, 86.1, 80.0, 78.3, 70.2, 69.8, 62.4. HRMS (ESI): *m/z*: calcd for $C_{26}H_{24}NaO_7S$ [M + Na]: 503.1135; found: 503.1130.



2-11

Compound 2-11: Compound **2-7** (5.06 g, 10.54 mmol) was dissolved in CH_2Cl_2 (70 mL) followed by the addition of TEMPO (201 mg, 1.29 mmol), BAIB (5.18 g, 16.1 mmol) and H_2O (35 mL), and the reaction mixture was stirred at room temperature for 4 h. When TLC analysis (hexane/EtOAc, 1:1, v/v) showed completion of reaction, the reaction mixture was quenched with saturated $Na_2S_2O_3$, diluted with CH_2Cl_2 and washed with saturated $Na_2S_2O_3$ (3x), brine (1x). The organic layer was dried over Na_2SO_4, filtered and concentrated to form a brown oil, which was redissolved in DMF (40 mL), followed by the addition of MeI (0.6 mL, 9.64 mmol), $NaHCO_3$ (2.7 g, 32.2 mmol) and TBAI (59.4 mg, 0.16

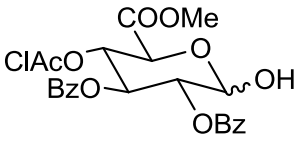
mmol), and stirred at 50°C for 1 h. When TLC analysis (hexane/EtOAc, 3:1 v/v) showed completion of reaction, the reaction mixture was diluted with diethyl ether, washed with water (3x), brine (1x), dried over Na_2SO_4, filtered, and concentrated. The product was purified by flash column chromatography (hexane/EtOAc, 5:1 towards 3:1) to afford compound **2-11** as a white solid (2.68 g, 50%, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.00 – 7.88 (m, 4H), 7.58 – 7.45 (m, 4H), 7.42 – 7.33 (m, 4H), 7.33 – 7.28 (m, 3H), 5.55 (t, *J* = 9.2 Hz, 1H), 5.42 (t, *J* = 9.7 Hz, 1H), 4.97 (d, *J* = 10.0 Hz, 1H), 4.18 (td, *J* = 9.4, 2.1 Hz, 1H), 4.11 (d, *J* = 9.6 Hz, 1H), 3.89 (s, 3H), 3.32 (s, 1H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 168.8, 166.6, 165.1, 133.4, 133.3, 133.0, 131.9, 129.9, 129.8, 129.1, 129.0, 128.9, 128.4, 128.4, 87.0, 78.1, 76.2, 70.4, 69.7, 53.0. HRMS (ESI): *m/z*: calcd for $C_{27}H_{24}NaO_8S$ [M + Na]: 531.1084; found: 531.1092.



2-15

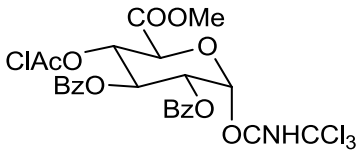
Compound 2-13: Compound **2-11** (2.80 g, 5.27 mmol) was dissolved in CH_2Cl_2 (30 mL) followed by the addition of chloroacetic anhydride (1.8 g, 10.54 mmol) and pyridine (4.25 mL, 52.7 mmol) and the reaction mixture was stirred at room temperature for 1 h. When TLC analysis (hexane/EtOAc, 2:1, v/v) showed completion of reaction, the reaction mixture was diluted with CH_2Cl_2 and washed with 5M HCl (3x), followed by saturated $NaHCO_3$ (3x) and brine (1x). The

organic layer was dried over Na_2SO_4, filtered and concentrated to crude compound **2-13** as a white solid which was immediately used for the next reaction.



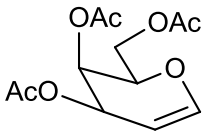
2-15

Compound 2-15: Crude compound **2-13** was redissolved in CH_2Cl_2 (50 mL) followed by the addition of NIS (1.77 g, 7.91 mmol) and TFA (807 μL, 8.99 mmol) and H_2O (8 mL). The reaction was stirred at room temperature for 16 h. When TLC analysis (hexane/EtOAc, 2:1, v/v) showed completion of reaction, the reaction mixture was diluted with CH_2Cl_2 and washed with saturated $Na_2S_2O_3$ (3x) and brine (1x). The organic layer was dried over anhydrous Na_2SO_4, filtered and concentrated. The product was purified by flash column chromatography (hexane/EtOAc, 4:1 towards 2:1) to afford compound **2-15** as a white solid (1.72 g, 68%, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.95 (dt, *J* = 16.5, 8.1 Hz, 4H), 7.56 – 7.48 (m, 2H), 7.42 – 7.34 (m, 4H), 6.08 (t, *J* = 9.7 Hz, 1H), 5.80 (t, *J* = 4.5 Hz, 1H), 5.57 – 5.46 (m, 1H), 5.30 – 5.20 (m, 1H), 4.78 (d, *J* = 10.0 Hz, 1H), 3.97 (t, *J* = 6.7 Hz, 2H), 3.80 – 3.75 (m, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 168.0, 166.2, 165.6, 133.6, 133.5, 129.9, 129.8, 129.8, 128.8, 128.7, 128.5, 128.5, 90.5, 71.4, 70.8, 69.2, 68.0, 53.1, 40.3. HRMS (ESI): *m/z*: calcd for $C_{23}H_{21}ClNaO_{10}$ [M + Na]: 515.0715; found: 515.0721.



D1

Donor D1: Compound **2-15** (250 mg, 0.49 mmol) was dissolved in CH_2Cl_2 (5 mL) followed by the addition of DBU (30.5 μL, 0.2 mmol) and trichloroacetonitrile (511.4 μL, 5.1 mmol) and the reaction mixture was stirred for 4 hours at room temperature. When TLC analysis (hexane/EtOAc, 4:1, v/v) showed completion of reaction, the solvent was evaporated and the compound was purified by flash column chromatography (hexane/EtOAc, 4:1) to afford donor **D1** as a white foam (220.9 mg, 68%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.69 (s, 1H), 7.96 – 7.89 (m, 4H), 7.55 – 7.48 (m, 2H), 7.37 (dt, *J* = 15.8, 7.7 Hz, 4H), 6.85 (d, *J* = 3.5 Hz, 1H), 6.13 (t, *J* = 10.0 Hz, 1H), 5.64 – 5.49 (m, 2H), 4.68 (d, *J* = 10.2 Hz, 1H), 3.99 (q, *J* = 14.9 Hz, 2H), 3.80 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 166.9, 166.2, 165.5, 165.2, 160.2, 133.7, 129.9, 129.8, 128.5, 128.5, 128.4, 128.3, 92.7, 90.4, 70.4, 70.3, 70.0, 69.2, 53.2, 40.2. HRMS (ESI): *m/z*: calcd for $C_{25}H_{21}Cl_4NNaO_{10}$ [M + Na]: 657.9812; found: 657.9817.

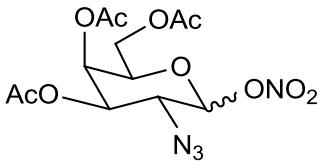


2-21

Compound 2-21: D-(+)-galactose (Compound **2-18**) (6.00 g, 33.33 mmol) and DMAP (101.80 mg, 0.83 mmol) were added to pyridine (16.1 mL, 200 mmol) and

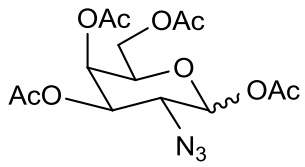
Ac_2O (18.9 mL, 200 mmol) and stirred at room temperature for 2 h. When TLC analysis (hexane/EtOAc, 1:1 v/v) showed completion of reaction, the reaction mixture was diluted with EtOAc and washed with 5M HCl (3x), followed by saturated $NaHCO_3$ (3x) and brine (1x). The organic layer was dried over Na_2SO_4, filtered and concentrated to a white solid (crude compound **2-19**), which was redissolved in EtOAc (80 mL), cooled to 0 °C and treated with HBr (41.1 mL). The reaction mixture was allowed to warm to room temperature and stirred for 1 h. When TLC analysis (hexane/EtOAc, 2:1 v/v) showed completion of reaction, the mixture was quenched with saturated $NaHCO_3$, diluted with EtOAc and washed with saturated $NaHCO_3$ (3x) and brine (1x). The organic layer was dried over anhydrous Na_2SO_4, filtered and concentrated. The crude product **2-20** was redissolved in EtOAc (44 mL), followed by the addition of Zn (11.96 g, 183 mmol), NaOAc (13.04 g, 150 mmol), $CuSO_4$ (259.0 mg, 1.04 mmol), AcOH (31 mL) and H_2O (12.2 mL). The mixture was stirred for 2 h at room temperature. When TLC analysis (hexane/EtOAc, 2:1 v/v) showed completion of reaction, the mixture was filtered, diluted with EtOAc, washed with saturated $NaHCO_3$ (3x) and brine (1x). The organic layer was dried over Na_2SO_4, filtered and concentrated. The product was purified by flash column chromatography (hexane/EtOAc, 3:1 towards 2.5:1) to afford compound **2-21** as a pale yellow oil (4.99 g, 55%, 3 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.40 (dd, *J* = 6.3, 1.7 Hz, 1H), 5.54 – 5.45 (m, 1H), 5.36 (dd, *J* = 4.4, 1.7 Hz, 1H), 4.66 (ddd, *J* = 6.3, 2.4, 1.3 Hz, 1H), 4.26 (d, *J* = 6.2 Hz, 1H), 4.18 (ddd, *J* = 17.0, 11.6, 6.3 Hz, 2H), 2.06 (s, 3H), 2.02 (s, 3H), 1.96 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.3, 170.1,

169.9, 145.2, 98.7, 72.6, 63.7, 63.6, 61.7, 20.6, 20.5, 20.4. HRMS (ESI): *m/z*: calcd for $C_{12}H_{16}NaO_7$ [M + Na]: 295.0788; found: 295.0781.



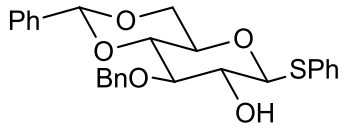
2-22

Compound 2-22: A solution of compound **2-21** (5.60 g, 20.59 mmol) in MeCN (50 mL) was added dropwise to a solution of NaN_3 (2.01 g, 65.05 mmol) and CAN (33.87 g, 61.77 mmol) in MeCN (50 mL) at -25°C. The reaction mixture was allowed to warm to -15°C and stirred for 18 h. When TLC analysis (hexane/EtOAc, 2:1 v/v) showed completion of reaction, the mixture was diluted with diethyl ether, washed with water (2x) and brine (1x). The organic layer was dried over Na_2SO_4, filtered and concentrated. The product was purified by flash column chromatography (hexane/EtOAc, 4:1 towards 3:1) to afford **2-22** as a yellow solid (3.49 g, 45%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.28 (d, *J* = 4.1 Hz, 1H), 5.57 (d, *J* = 8.8 Hz, 1H), 5.41 (d, *J* = 2.4 Hz, 1H), 5.30 (d, *J* = 3.2 Hz, 1H), 5.15 (dd, *J* = 11.3, 3.2 Hz, 1H), 4.93 (dd, *J* = 10.6, 3.3 Hz, 1H), 4.30 (t, *J* = 6.5 Hz, 1H), 4.10 – 3.99 (m, 6H), 3.75 (dd, *J* = 10.5, 8.9 Hz, 1H), 2.08 (s, 6H), 1.97 (d, *J* = 1.2 Hz, 6H), 1.94 (d, *J* = 5.0 Hz, 6H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.0, 170.0, 169.7, 169.5, 169.3, 169.2, 97.8, 96.8, 71.6, 71.4, 69.2, 68.3, 66.4, 65.7, 60.7, 60.7, 57.3, 55.7, 20.2, 20.2. HRMS (ESI): *m/z*: calcd for $C_{12}H_{16}N_4NaO_{10}$ [M + Na]: 399.0759; found: 399.0765.



2-23

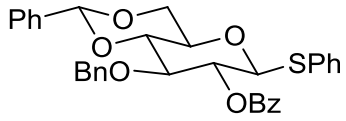
Compound 2-23: Compound **2-22** (2.76 g, 7.34 mmol) was dissolved in AcOH (20 mL), followed by addition of NaOAc (1.2 g, 1.47 mmol). The mixture was stirred for 1 h at 100°C. When TLC analysis (hexane/EtOAc, 2:1 v/v) showed completion of reaction, the mixture was diluted with EtOAc, washed with saturated $NaHCO_3$ (3x) and brine (1x). The organic layer was dried over Na_2SO_4, filtered and concentrated to form a yellow oil, which was used immediately for the next step to obtain **2-65**.



2-51

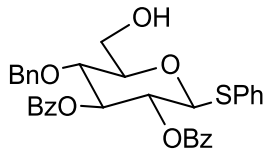
Compound 2-51: To a solution of compound **2-5** (9.50 g, 26.4 mmol) in toluene (92.4 mL) was added dibutyltin oxide (7.23 g, 29 mmol) and the reaction mixture was stirred for 3 h at 110°C. The solvent was next evaporated off to form a yellow oil which was redissolved in DMF (100 mL), followed by the addition of benzyl bromide (3.77 mL, 31.7 mmol) and cesium fluoride (5.21 g, 34.31 mmol). The reaction mixture was stirred for 16 h at room temperature. After TLC analysis (hexane/EtOAc, 4:1 v/v) indicated completion of the reaction, the mixture was diluted with CH_2Cl_2 and washed with water (3x) and brine (1x), dried over Na_2SO_4, filtered, and concentrated. The product was purified by flash column

chromatography (hexane/CH_2Cl_2, 6:1 towards 100% CH_2Cl_2) to afford compound **2-51** as a white solid (7.72 g, 65%, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.52 (ddd, *J* = 27.8, 6.9, 2.5 Hz, 4H), 7.42 – 7.27 (m, 11H), 5.58 (s, 1H), 4.88 (dd, *J* = 80.6, 11.5 Hz, 2H), 4.64 (d, *J* = 9.7 Hz, 1H), 4.40 (dd, *J* = 10.5, 5.0 Hz, 1H), 3.80 (t, *J* = 10.3 Hz, 1H), 3.73 – 3.62 (m, 2H), 3.60 – 3.46 (m, 2H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 138.1, 137.1, 133.4, 131.3, 129.0, 128.4, 128.3, 128.2, 128.1, 127.9, 126.0, 101.2, 88.4, 81.6, 81.1, 74.8, 72.2, 70.7, 68.6. HRMS (ESI): *m/z*: calcd for $C_{26}H_{26}NaO_5S$ [M + Na]: 473.1393; found: 473.1388.



2-52

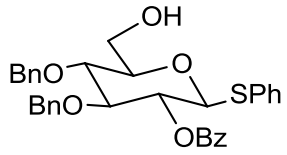
Compound 2-52: Compound **2-51** (7.72 g, 17.1 mmol) was treated as described for the preparation of compound **2-6** except that 1.2 equivalents of Bz_2O and DMAP were used. Compound **2-52** was obtained as a white solid which was immediately used for the next reaction to obtain compound **2-54** or compound **2-55**.



2-53

Compound 2-53: Crude compound **2-6** was dissolved in 1M BH_3·THF (54.3mL) and $CoCl_2$ (7.05g, 54.3mmol) was added in 3 portions. The reaction was stirred

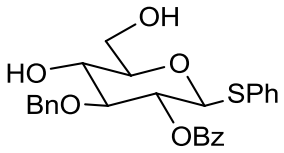
for 2 h at room temperature. When TLC analysis (hexane/EtOAc, 3:1 v/v) showed completion of reaction, the reaction was diluted with EtOAc and quenched with saturated $NaHCO_3$. The solid precipitates were filtered off, and the organic layer washed with saturated $NaHCO_3$ (3x) and brine (1x). The organic layer was dried over anhydrous Na_2SO_4, filtered and concentrated. The product was purified by flash column chromatography (hexane/EtOAc, 5:1 towards 3:1) to afford compound **2-53** as a white solid (6.51 g, 63%, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.04 (dd, *J* = 29.4, 7.4 Hz, 4H), 7.59 – 7.55 (m, 2H), 7.48 (dd, *J* = 14.2, 7.2 Hz, 2H), 7.33 (ddd, *J* = 14.9, 10.9, 5.1 Hz, 7H), 7.23 – 7.16 (m, 5H), 6.01 – 5.92 (m, 1H), 5.58 (t, *J* = 9.7 Hz, 1H), 5.19 (d, *J* = 10.0 Hz, 1H), 4.71 (dd, *J* = 27.9, 11.1 Hz, 2H), 4.10 – 4.01 (m, 2H), 3.93 (dd, *J* = 12.1, 3.4 Hz, 1H), 3.81 – 3.74 (m, 1H), 3.04 (s, 1H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 165.4, 165.0, 136.9, 132.9, 132.8, 132.3, 131.9, 129.4, 129.3, 129.0, 128.8, 128.7, 128.0, 128.0, 127.8, 127.6, 127.5, 85.8, 79.4, 76.0, 75.1, 74.4, 70.7, 61.0. HRMS (ESI): *m/z*: calcd for $C_{33}H_{30}NaO_7S$ [M + Na]: 593.1604; found: 593.1609.



2-54

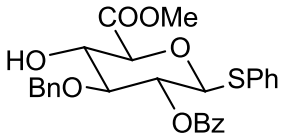
Compound 2-54: Crude compound **2-52** was treated as described for the preparation of compound **2-53** to afford compound **2-54** as a white solid (6.1 g, 64%, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.11 – 8.07 (m, 2H), 7.62 (t, *J* = 7.5 Hz, 1H), 7.53 – 7.43 (m, 4H), 7.40 – 7.28 (m, 8H), 7.20 – 7.10 (m, 5H), 5.38 –

5.27 (m, 1H), 4.93 – 4.85 (m, 2H), 4.79 (s, 1H), 4.72 (d, *J* = 11.0 Hz, 2H), 3.99 – 3.82 (m, 2H), 3.82 – 3.70 (m, 2H), 3.57 (ddd, *J* = 9.6, 4.7, 2.5 Hz, 1H), 2.17 (s, 1H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 165.1, 137.7, 137.5, 133.1, 132.6, 132.3, 129.7, 129.7, 128.9, 128.4, 128.3, 128.2, 128.0, 127.9, 127.9, 127.9, 127.6, 86.1, 83.9, 79.6, 77.4, 75.2, 75.0, 72.4, 61.9. HRMS (ESI): *m/z*: calcd for $C_{33}H_{32}NaO_6S$ [M + Na]: 579.1812; found: 579.1814.



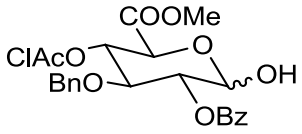
2-55

Compound 2-55 Crude compound **2-52** was treated as described for the preparation of compound **2-7** to afford compound **2-55** as a white solid (5.98 g, 75%, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.08 (d, *J* = 8.0 Hz, 2H), 7.61 (dd, *J* = 10.9, 4.2 Hz, 1H), 7.52 – 7.36 (m, 4H), 7.31 – 7.24 (m, 3H), 7.22 – 7.11 (m, 5H), 5.35 – 5.20 (m, 1H), 4.85 (d, *J* = 10.1 Hz, 1H), 4.67 (dd, *J* = 57.9, 11.4 Hz, 2H), 3.94 (d, *J* = 11.9 Hz, 1H), 3.85 – 3.76 (m, 1H), 3.76 – 3.67 (m, 2H), 3.48 (dt, *J* = 8.4, 4.4 Hz, 1H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 165.2, 137.6, 133.3, 132.7, 132.3, 129.8, 128.9, 128.5, 128.0, 128.0, 86.4, 83.8, 79.4, 74.8, 72.3, 70.3, 62.5. HRMS (ESI): *m/z*: calcd for $C_{26}H_{26}NaO_6S$ [M + Na]: 489.1342; found: 489.1346.



2-56

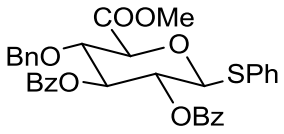
Compound 2-56: Compound **2-55** (3.22 g, 6.9 mmol) was treated as described for the preparation of **compound 2-11** to afford compound **2-56** as a white solid (1.77 g, 52%, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.05 (d, *J* = 7.2 Hz, 2H), 7.63 (t, *J* = 7.5 Hz, 1H), 7.49 (dd, *J* = 9.2, 5.9 Hz, 4H), 7.33 – 7.24 (m, 3H), 7.21 – 7.12 (m, 5H), 5.33 – 5.23 (m, 1H), 4.85 (d, *J* = 10.1 Hz, 1H), 4.76 (q, *J* = 11.6 Hz, 2H), 4.05 (d, *J* = 9.2 Hz, 1H), 3.97 (d, *J* = 9.7 Hz, 1H), 3.88 (s, 3H), 3.76 (t, *J* = 8.9 Hz, 1H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 169.3, 165.0, 137.6, 133.3, 132.7, 132.4, 129.8, 129.6, 128.9, 128.4, 128.3, 128.1, 128.0, 127.7, 87.2, 82.2, 77.6, 74.7, 71.9, 71.3, 52.9. HRMS (ESI): *m/z*: calcd for $C_{27}H_{26}NaO_7S$ [M + Na]: 517.1291; found: 517.1298.



2-58

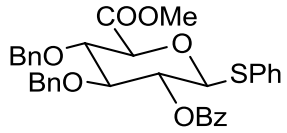
Compound 2-58: Compound **2-56** (1.77 g, 3.58 mmol) was treated as described for the preparation of compound **2-15** to afford compound **2-58** as a white solid (1.22 g, 69%, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.12 – 7.95 (m, 2H), 7.59 (dd, *J* = 12.5, 4.3 Hz, 1H), 7.45 (q, *J* = 7.5 Hz, 2H), 7.28 – 7.17 (m, 5H), 5.67 (d, *J* = 3.4 Hz, 1H), 5.25 (d, *J* = 9.2 Hz, 1H), 5.14 (dd, *J* = 9.4, 3.3 Hz, 1H), 4.77 (d, *J* = 11.8 Hz, 1H), 4.65 – 4.56 (m, 2H), 4.27 (q, *J* = 9.0 Hz, 1H), 4.13 – 4.04 (m, 1H),

3.90 – 3.86 (m, 1H), 3.71 (s, 3H). 13C NMR (500 MHz, CDCl3) δ 168.6, 166.2, 165.7, 137.6, 133.7, 133.5, 129.8, 129.8, 128.5, 128.4, 127.9, 127.9, 127.8, 90.2, 75.9, 75.0, 72.9, 72.0, 68.3, 52.9, 40.3. HRMS (ESI): m/z: calcd for C23H23ClNaO9 [M + Na]: 501.0923; found: 501.0922.



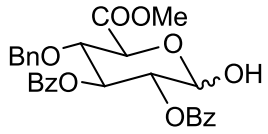
2-59

Compound 2-59: Compound **2-53** (3.67 g, 6.43 mmol) was treated as described for the preparation of compound **2-11** to afford compound **2-59** as a white solid (1.85 g, 48%, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.99 (dd, *J* = 21.7, 7.9 Hz, 4H), 7.56 – 7.48 (m, 4H), 7.41 – 7.30 (m, 7H), 7.17 (dt, *J* = 7.1, 3.9 Hz, 5H), 5.86 (t, *J* = 8.9 Hz, 1H), 5.50 (t, *J* = 9.6 Hz, 1H), 5.09 (d, *J* = 9.8 Hz, 1H), 4.62 (dd, *J* = 31.8, 11.1 Hz, 2H), 4.27 (dd, *J* = 14.3, 9.1 Hz, 2H), 3.80 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 167.9, 165.3, 164.9, 136.8, 133.1, 132.6, 131.7, 129.6, 129.5, 128.9, 128.8, 128.2, 128.1, 128.1, 127.8, 127.7, 86.6, 77.8, 77.0, 75.3, 74.5, 70.2, 52.5. HRMS (ESI): *m/z*: calcd for $C_{34}H_{30}NaO_8S$ [M + Na]: 621.1554; found: 621.1556.



2-60

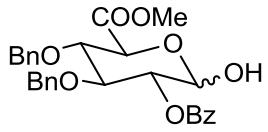
Compound 2-60: Compound **2-54** (3.54 g, 6.36 mmol) was treated as described for the preparation of compound **2-11** to afford compound **2-54** as a white solid (2.04 g, 55%, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.06 (dd, *J* = 8.1, 1.0 Hz, 2H), 7.65 – 7.57 (m, 1H), 7.48 (ddd, *J* = 6.2, 5.0, 3.0 Hz, 4H), 7.41 – 7.22 (m, 8H), 7.20 – 7.08 (m, 5H), 5.44 – 5.27 (m, 1H), 4.90 – 4.72 (m, 3H), 4.68 (dd, *J* = 11.0, 4.6 Hz, 2H), 4.08 (d, *J* = 9.5 Hz, 1H), 4.01 (t, *J* = 9.2 Hz, 1H), 3.90 (t, *J* = 8.9 Hz, 1H), 3.78 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 168.2, 164.9, 137.5, 137.3, 133.2, 132.6, 132.2, 129.7, 129.6, 128.8, 128.3, 128.3, 128.1, 128.0, 127.9, 127.9, 127.8, 127.6, 86.9, 83.1, 79.0, 78.0, 75.1, 75.0, 71.7, 52.5. HRMS (ESI): *m/z*: calcd for $C_{34}H_{32}NaO_7S$ [M + Na]: 607.1761; found: 607.1767.



2-61

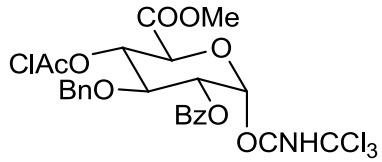
Compound 2-61: Compound **2-59** (1.85 g, 3.09 mmol) was treated as described for the preparation of compound **2-15** as a white solid to afford compound **2-61** as a white solid (1.16 g, 74%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.97 (ddd, *J* = 14.0, 10.1, 5.4 Hz, 4H), 7.58 – 7.42 (m, 2H), 7.43 – 7.27 (m, 4H), 7.21 – 7.07 (m, 5H), 6.13 (t, *J* = 9.6 Hz, 1H), 5.74 (d, *J* = 3.5 Hz, 1H), 5.24 (dd, *J* = 10.0, 3.6 Hz, 1H), 4.75 (d, *J* = 9.7 Hz, 1H), 4.69 – 4.47 (m, 2H), 4.30 – 4.06 (m, 1H), 3.73 (t, *J* = 4.9

Hz, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 169.7, 168. 8, 166.3, 165.7, 165.5, 137.0, 136.8, 133.3, 133.2, 133.1, 129.7, 129.6, 129.6, 128.3, 128.2, 128.1, 127.9, 95.8, 90.5, 74.6, 74.4, 74.3, 73.7, 73.6, 71.8, 71.5, 69.8, 52.5. HRMS (ESI): *m/z*: calcd for $C_{28}H_{26}NaO_9$ [M + Na]: 529.1469; found: 529.1473.



2-62

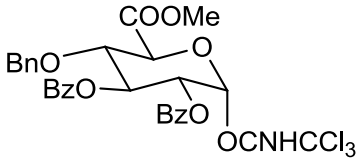
Compound 2-62: Compound **2-60** (2.04 g, 3.49 mmol) was treated as described for the preparation of compound **2-15** to afford compound **2-62** as a white solid (1.2 g, 70%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.08 – 8.02 (m, 2H), 7.57 (t, *J* = 7.4 Hz, 1H), 7.43 (t, *J* = 7.7 Hz, 2H), 7.37 – 7.15 (m, 10H), 5.59 (d, *J* = 3.6 Hz, 1H), 5.15 (dd, *J* = 9.7, 3.5 Hz, 1H), 4.87 – 4.76 (m, 3H), 4.62 (dd, *J* = 32.8, 10.2 Hz, 2H), 4.25 (t, *J* = 9.2 Hz, 1H), 3.96 – 3.87 (m, 1H), 3.73 (s, 4H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 169.8, 165.8, 137.8, 137.7, 133.3, 129.8, 128.4, 128.3, 128.3, 127.9, 127.8, 127.8, 127.7, 90.7, 79.2, 78.7, 75.4, 75.0, 73.4, 70.3, 52.5. HRMS (ESI): *m/z*: calcd for $C_{28}H_{28}NaO_8$ [M + Na]: 515.1676; found: 515.1681.



D2

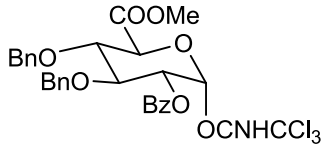
Donor D2: Compound **2-58** (250 mg, 0.522 mmol) was treated as described for the preparation of donor **D1** to afford donor **D2** as a white foam (227.7 mg, 70%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.61 (s, 1H), 7.95 (dd, *J* = 8.1, 1.0 Hz, 2H), 7.56 (d, *J* = 7.3 Hz, 1H), 7.41 (t, *J* = 7.8 Hz, 2H), 7.25 – 7.21 (m, 3H), 7.18 – 7.14 (m, 2H), 6.71 (d, *J* = 3.6 Hz, 1H), 5.42 (dd, *J* = 9.9, 3.6 Hz, 1H), 5.38 – 5.31 (m, 1H), 4.69 (dd, *J* = 45.7, 11.7 Hz, 2H), 4.46 (d, *J* = 10.2 Hz, 1H), 4.29 (t, *J* = 9.6 Hz, 1H), 3.94 (s, 1H), 3.81 (d, *J* = 14.7 Hz, 1H), 3.72 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 167.2, 166.0, 165.1, 160.0, 137.2, 133.6, 129.8, 128.8, 128.5, 128.4, 128.1, 128.0, 93.0, 90.6, 75.6, 75.1, 71.9, 71.6, 70.6, 53.1, 40.3. HRMS (ESI): *m/z*: calcd for $C_{25}H_{23}Cl_4NNaO_9$ [M + Na]: 644.0019; found: 644.0021.



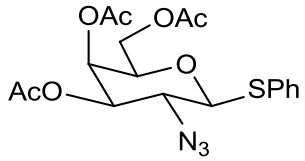
D3

Donor D3: Compound **2-61** (250 mg, 0.494 mmol) was treated as described for the preparation of donor **D1** to afford donor **D3** as a white foam (241.1 mg, 75%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.63 (s, 1H), 8.00 – 7.89 (m, 4H), 7.57 – 7.47 (m, 2H), 7.40 (t, *J* = 7.8 Hz, 2H), 7.34 (d, *J* = 7.8 Hz, 2H), 7.17 (dd, *J* = 6.5, 3.5 Hz, 3H), 7.12 (dd, *J* = 6.5, 3.0 Hz, 2H), 6.78 (d, *J* = 3.5 Hz, 1H), 6.15 (t, *J* = 9.8 Hz, 1H), 5.49 (dd, *J* = 10.2, 3.7 Hz, 1H), 4.68 – 4.54 (m, 3H), 4.25 (t, *J* = 9.7 Hz, 1H), 3.78 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 168.5, 165.4, 165.0, 160.4, 136.7, 133.5, 133.3, 129.8, 129.7, 129.2, 128.4, 128.4, 128.3, 128.2, 128.0, 93.1, 90.5, 77.1, 75.0, 72.4, 71.4, 70.4, 52.8. HRMS (ESI): *m/z*: calcd for $C_{30}H_{26}Cl_3NNaO_9$ [M + Na]: 672.0565; found: 672.0558.



D4

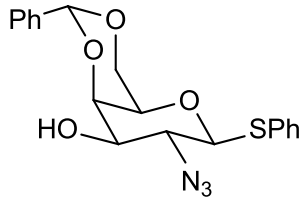
Donor D4: Compound **2-62** was treated as described for the preparation of donor **D1** to afford donor **D4** as a white foam (220.9 mg, 68%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.62 (s, 1H), 8.00 (d, *J* = 7.6 Hz, 2H), 7.59 (t, *J* = 7.5 Hz, 1H), 7.44 (t, *J* = 7.8 Hz, 2H), 7.40 – 7.29 (m, 5H), 7.29 – 7.17 (m, 5H), 6.71 (d, *J* = 3.6 Hz, 1H), 5.46 (dd, *J* = 9.9, 3.5 Hz, 1H), 4.88 (dt, *J* = 20.0, 8.5 Hz, 3H), 4.70 (d, *J* = 10.7 Hz, 1H), 4.55 (d, *J* = 10.1 Hz, 1H), 4.34 (t, *J* = 9.5 Hz, 1H), 4.07 (t, *J* = 9.7 Hz, 1H), 3.77 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 168.5, 165.2, 160.1, 137.4, 137.2, 133.3, 129.6, 128.9, 128.4, 128.3, 128.2, 128.1, 128.0, 127.7, 93.4, 90.6, 78.9, 78.5, 75.4, 75.4, 72.6, 71.9, 52.5. HRMS (ESI): *m/z*: calcd for $C_{30}H_{28}Cl_3NNaO_8$ [M + Na]: 658.0773; found: 658.0775.



2-65

Compound 2-65: Crude compound **2-23** was redissolved in CH_2Cl_2 (30 mL), followed by the addition of $BF_3{\cdot}OEt_2$ (2.72 mL, 22 mmol) and PhSH (2.56 mL, 25 mmol). The reaction mixture was stirred for 16 h at 35°C. When TLC analysis (hexane/EtOAc, 3:1 v/v) showed completion of reaction, the mixture was diluted with CH_2Cl_2, washed with $NaHCO_3$ (3x) and brine (1x). The organic layer was dried over Na_2SO_4, filtered, and concentrated. The product was purified by flash

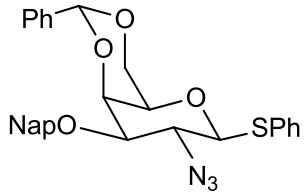
column chromatography (hexane/EtOAc, 6:1 towards 3:1) to afford compound **2-65** as a yellow oil (2.42 g, 78%, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.59 (dd, *J* = 6.6, 2.9 Hz, 2H), 7.48 (dd, *J* = 7.5, 1.7 Hz, 2H), 7.36 – 7.25 (m, 6H), 5.67 (d, *J* = 5.4 Hz, 1H), 5.45 (d, *J* = 2.8 Hz, 1H), 5.32 (d, *J* = 3.0 Hz, 1H), 5.15 (dd, *J* = 11.1, 3.2 Hz, 1H), 4.85 (dd, *J* = 10.3, 3.2 Hz, 1H), 4.72 (d, *J* = 6.5 Hz, 1H), 4.51 (d, *J* = 10.1 Hz, 1H), 4.29 (dd, *J* = 11.1, 5.5 Hz, 1H), 4.18 – 4.11 (m, 1H), 4.11 – 4.02 (m, 3H), 3.87 (d, *J* = 6.6 Hz, 1H), 3.62 (t, *J* = 10.2 Hz, 1H), 2.12 (s, 3H), 2.04 (d, *J* = 10.5 Hz, 6H), 2.00 (d, *J* = 5.4 Hz, 6H), 1.95 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.2, 170.2, 169.8, 169.8, 169.5, 169.4, 133.4, 132.4, 132.3, 130.9, 129.0, 128.8, 128.4, 127.9, 86.7, 86.2, 74.2, 72.8, 70.0, 67.4, 67.3, 66.4, 61.5, 61.4, 59.2, 57.9, 20.5, 20.4, 20.4. HRMS (ESI): *m/z*: calcd for $C_{18}H_{21}N_3NaO_7S$ [M + Na]: 446.0992; found: 446.0999.



2-66

Compound 2-66: To a solution of compound **2-65** (5.28 g, 12.28 mmol) in methanol (80 mL) and CH_2Cl_2 (8 mL) was added NaOMe (265.3 mg, 4.91 mmol) and the mixture was stirred at room temperature for 1 h. When TLC analysis (hexane/EtOAc, 2:1 v/v) showed completion of reaction, the reaction mixture was quenched with Amberlyst-15 and stirred for another 15 min. When the pH was about 3 – 4, the mixture was filtered and concentrated to form a pale brown oil, which was redissolved in MeCN (100 mL), followed by the addition of

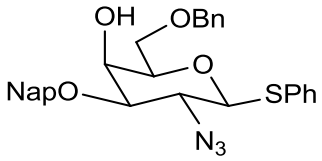
$PhCH(OMe)_2$ (3.69 mL, 24.56 mmol) and camphorsulfonic acid (934 mg, 4.91 mmol). The reaction mixture was stirred for 16 h at 55°C. When TLC analysis (hexane/EtOAc, 3:1 v/v) showed completion of reaction, the mixture was quenched with TEA (1.5 mL), followed by evaporation of solvent. The product was purified by flash column chromatography (hexane/EtOAc, 6:1 towards 3:1 towards 1:1) to afford compound **2-66** as a yellow oil (3.5 g, 74%, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.53 – 7.45 (m, 4H), 7.42 – 7.38 (m, 3H), 7.32 – 7.24 (m, 3H), 5.75 (d, *J* = 5.2 Hz, 1H), 5.58 (s, 1H), 4.27 (d, *J* = 3.6 Hz, 1H), 4.19 (dd, *J* = 11.2, 5.9 Hz, 3H), 4.08 (dd, *J* = 12.9, 1.8 Hz, 1H), 4.00 (dd, *J* = 10.1, 2.4 Hz, 1H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 137.1, 133.5, 131.0, 129.3, 129.0, 128.2, 127.3, 126.1, 101.2, 87.1, 75.0, 69.4, 69.0, 63.5, 61.1. HRMS (ESI): *m/z*: calcd for $C_{19}H_{19}N_3NaO_4S$ [M + Na]: 408.0988; found: 408.0996.



2-67

Compound 2-67: Compound **2-66** (2.44 g, 6.32 mmol) and NapBr (1.68 g, 7.59 mmol) were dissolved in DMF (24 ml) before NaH (304 mg, 7.59 mmol) was added in 3 portions to the reaction mixture. Reaction mixture was stirred for 3 h at room temperature. When TLC analysis (hexane/EtOAc, 3:1 v/v) showed completion of reaction, the mixture was diluted in diethyl ether and washed with H_2O (3x) and brine (1x). The organic layer was dried over Na_2SO_4, filtered, and concentrated. The product was purified by flash column chromatography

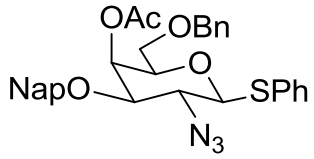
(hexane/CH_2Cl_2, 3:1 towards 1:1 towards 100% CH_2Cl_2) to afford compound **2-67** as a white solid (2.69 g, 81%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.85 (ddd, *J* = 10.6, 9.4, 7.8 Hz, 4H), 7.62 – 7.45 (m, 7H), 7.40 (q, *J* = 5.5 Hz, 3H), 7.32 – 7.25 (m, 3H), 5.80 (d, *J* = 5.3 Hz, 1H), 5.50 (s, 1H), 4.95 (dd, *J* = 31.0, 12.2 Hz, 2H), 4.55 (dd, *J* = 10.6, 5.4 Hz, 1H), 4.27 (d, *J* = 3.1 Hz, 1H), 4.19 (dd, *J* = 12.6, 1.0 Hz, 1H), 4.11 (s, 1H), 4.02 (dd, *J* = 12.6, 1.4 Hz, 1H), 3.92 (dd, *J* = 10.6, 3.4 Hz, 1H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 137.5, 135.1, 133.6, 133.2, 133.1, 131.0, 129.0, 128.2, 128.2, 127.9, 127.7, 127.3, 126.4, 126.2, 126.0, 125.6, 100.9, 87.5, 76.5, 72.9, 71.5, 69.2, 63.7, 59.3. HRMS (ESI): *m/z*: calcd for $C_{30}H_{27}N_3NaO_4S$ [M + Na]: 548.1614; found: 548.1615.



2-68

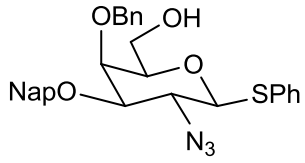
Compound 2-68: To a solution of compound **2-67** (2 g, 3.8 mmol) in CH_2Cl_2 (40 mL) was added 4Å molecular sieves (1 g). The mixture was stirred at room temperature for 30 mins, before cooling to -78°C. Triethylsilane (1.97 mL, 12.35 mmol) and TfOH (962 µL, 10.9 mmol) were then added dropwise into the reaction mixture, and the solution was stirred for another 2 h at -78°C. When TLC analysis (hexane/EtOAc, 3:1 v/v) showed completion of reaction, the mixture was quenched with TEA (2 mL) and MeOH (3 mL), followed by evaporation of solvent. The product was purified by flash column chromatography (hexane/EtOAc, 4:1 towards 3:1) to afford compound **2-68** as a yellow oil (1.56 g,

78%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.88 (dd, *J* = 9.7, 6.7 Hz, 4H), 7.59 – 7.44 (m, 5H), 7.39 – 7.28 (m, 5H), 7.26 (dd, *J* = 5.0, 2.0 Hz, 3H), 5.64 (d, *J* = 5.5 Hz, 1H), 4.91 (d, *J* = 4.1 Hz, 2H), 4.56 – 4.50 (m, 3H), 4.36 (dd, *J* = 10.4, 5.5 Hz, 1H), 4.21 (d, *J* = 2.3 Hz, 1H), 3.84 – 3.76 (m, 2H), 3.73 (dd, *J* = 10.3, 5.9 Hz, 1H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 137.7, 134.4, 133.2, 133.2, 133.2, 133.1, 132.4, 129.0, 128.5, 128.4, 127.9, 127.7, 127.6, 127.6, 126.9, 126.3, 126.2, 125.6, 87.4, 77.8, 73.5, 72.1, 69.7, 69.5, 66.7, 59.6. HRMS (ESI): *m/z*: calcd for $C_{30}H_{29}N_3NaO_4S$ [M + Na]: 550.1771; found: 550.1777.



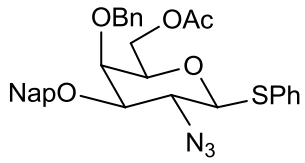
2-69

Compound 2-69: To a solution of compound **2-68** (3 g, 5.69 mmol) in CH_2Cl_2 (40 mL) was added DMAP (1.04 g, 8.54 mmol) and Ac_2O (643 µL, 6.82 mmol) and reaction mixture was stirred at room temperature for 2 h. When TLC analysis (hexane/EtOAc, 3:1 v/v) showed completion of reaction, the reaction mixture was diluted with CH_2Cl_2 and washed with 5M HCl (3x), followed by saturated $NaHCO_3$ (3x) and brine (1x). The organic layer was dried over Na_2SO_4, filtered and concentrated to form compound **2-69** as a yellow oil which was immediately used for the next reaction to obtain compound **2-75**.



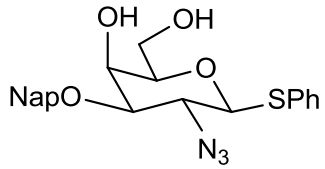
2-70

Compound 2-70: Compound **2-68** (2.6 g, 4.95 mmol) was treated as described for the preparation of compound **2-69** except that triethylsilane (2.57 mL, 16.1 mmol) and $PhBCl_2$ (1.84 mL, 14.2 mmol) were used, to afford compound **2-70** as a yellow oil (1.78 g, 68%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.94 – 7.83 (m, 4H), 7.58 (dd, *J* = 8.5, 1.3 Hz, 1H), 7.57 – 7.49 (m, 4H), 7.36 – 7.25 (m, 8H), 5.70 (d, *J* = 5.4 Hz, 1H), 5.02 – 4.89 (m, 3H), 4.63 (d, *J* = 11.5 Hz, 1H), 4.51 (dd, *J* = 10.6, 5.5 Hz, 1H), 4.29 (t, *J* = 6.0 Hz, 1H), 4.00 (s, 1H), 3.88 (dd, *J* = 10.5, 2.7 Hz, 1H), 3.75 (dd, *J* = 11.4, 6.8 Hz, 1H), 3.55 (dd, *J* = 11.4, 5.2 Hz, 1H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 137.7, 134.7, 133.2, 133.1, 132.4, 129.0, 128.5, 128.4, 128.2, 128.0, 127.9, 127.7, 127.7, 126.7, 126.2, 126.1, 125.6, 87.3, 79.2, 74.6, 73.3, 72.8, 71.9, 62.0, 60.4. HRMS (ESI): *m/z*: calcd for $C_{30}H_{29}N_3NaO_4S$ [M + Na]: 550.1771; found: 550.1780.



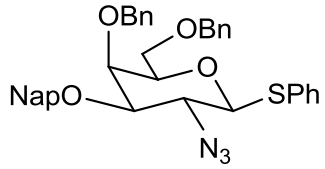
2-71

Compound 2-71: Compound **2-70** (1.78 g, 3.37 mmol) was treated as described for the preparation of **2-69** to form compound **2-76** as a yellow oil which was immediately used for the next reaction to obtain compound **35**.



2-72

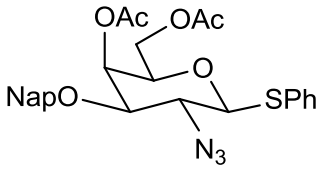
Compound 2-72: Compound **2-67** (1.5 g, 2.86 mmol) was dissolved in AcOH/H_2O (16 mL/3 mL) and stirred at 100°C for 2 h. When TLC analysis (hexane/EtOAc, 1:1 v/v) showed completion of reaction, the mixture was diluted with EtOAc and washed with saturated $NaHCO_3$ (3x) and brine (1x). The organic layer was dried over Na_2SO_4, filtered, and concentrated. The product was purified by flash column chromatography (hexane/EtOAc, 3:1 towards 1:1) to afford compound **2-72** as a white solid (1.13 g, 90%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.96 – 7.77 (m, 4H), 7.62 – 7.41 (m, 5H), 7.38 – 7.22 (m, 3H), 5.64 (d, *J* = 5.4 Hz, 1H), 4.89 (dd, *J* = 31.2, 11.5 Hz, 2H), 4.31 (dd, *J* = 10.0, 5.6 Hz, 2H), 4.16 (s, 1H), 3.89 (dd, *J* = 11.8, 5.6 Hz, 1H), 3.87 – 3.68 (m, 2H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 134.3, 133.2, 132.4, 129.1, 128.6, 127.9, 127.8, 127.7, 127.0, 126.4, 126.3, 125.6, 87.2, 77.6, 72.3, 70.5, 67.6, 62.8, 59.6. HRMS (ESI): *m/z*: calcd for $C_{23}H_{23}N_3NaO_4S$ [M + Na]: 460.1301; found: 460.1300.



2-73

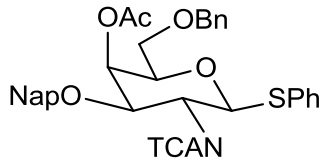
Compound 2-73: To a solution of compound **2-72** (2.2 g, 5.02 mmol) in DMF (30 mL) was added BnBr (1.44 mL, 12 mmol) and NaH (480 mg, 12 mmol). The

reaction was stirred at room temperature for 3 h. When TLC analysis (hexane/EtOAc, 6:1 v/v) showed completion of reaction, the mixture was diluted with diethyl ether and washed with water (3x) and brine (1x). The organic layer was dried over Na_2SO_4, filtered, and concentrated. The product was purified by flash column chromatography (hexane/EtOAc, 12:1 towards 8:1) to afford compound **2-73** as a yellow oil (2.33 g, 75%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.93 – 7.85 (m, 4H), 7.59 – 7.48 (m, 5H), 7.39 – 7.24 (m, 13H), 5.67 (d, *J* = 5.5 Hz, 1H), 5.00 – 4.92 (m, 3H), 4.62 (d, *J* = 11.3 Hz, 1H), 4.56 – 4.50 (m, 2H), 4.50 – 4.42 (m, 2H), 4.12 (d, *J* = 1.7 Hz, 1H), 3.88 (dd, *J* = 10.6, 2.7 Hz, 1H), 3.68 (dd, *J* = 9.4, 7.0 Hz, 1H), 3.59 (dd, *J* = 9.4, 6.1 Hz, 1H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 138.2, 137.8, 134.8, 133.4, 133.2, 133.0, 132.2, 128.9, 128.4, 128.3, 128.3, 127.9, 127.7, 127.7, 127.5, 126.5, 126.2, 126.0, 125.6, 87.6, 79.2, 74.9, 73.5, 73.4, 72.5, 70.5, 68.6, 60.4. HRMS (ESI): *m/z*: calcd for $C_{37}H_{35}N_3NaO_4S$ [M + Na]: 640.2240; found: 640.2245.



2-74

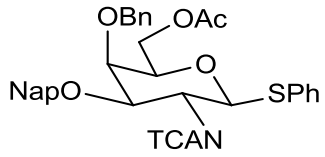
Compound 2-74: Compound **2-72** (2 g, 4.57 mmol) was treated as described for the preparation of **2-69** except that 2.4 equivalents of Ac_2O and DMAP were used, to afford compound **2-74** as a yellow oil which was immediately used for the next reaction to obtain compound **2-78**.



2-75

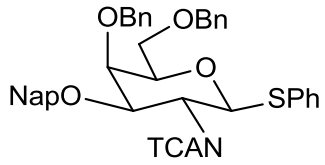
Compound 2-75: To a solution of crude compound **2-69** in EtOH/H_2O (40 mL/13 mL) was added Zn (492 mg, 7.57 mmol) and NH_4Cl (703 mg, 13.26 mmol) and the mixture stirred was at 80°C for 2 h. When TLC analysis (hexane/EtOAc, 6:1 v/v) showed completion of reaction, the reaction was diluted with EtOAc, quenched with diluted saturated $NaHCO_3$ and the solid precipitates were filtered off. The filtrate was washed with saturated $NaHCO_3$ (3x) and brine (1x), and the organic layer was dried over Na_2SO_4, filtered, and concentrated. The yellow oil was redissolved in THF (40 mL) and cooled to 0°C, before TEA (2.5 mL, 17.1 mmol) and CCl_3COCl (1.9 ml, 17.1 mmol) were added. The mixture was stirred for 30 min at 0°C. When TLC analysis (hexane/EtOAc, 4:1 v/v) showed completion of reaction, water was added and THF was evaporated. The compound was extracted using EtOAc, washed with saturated $NaHCO_3$ (3x) and brine (1x), dried over Na_2SO_4, filtered and concentrated. The product was purified by flash column chromatography (hexane/EtOAc, 6:1 towards 4:1) to afford compound **2-75** as a yellow oil (2.09 g, 48%, 3 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.84 (dd, *J* = 7.8, 3.9 Hz, 3H), 7.79 (s, 1H), 7.55 – 7.47 (m, 2H), 7.47 – 7.29 (m, 8H), 7.26 – 7.20 (m, 3H), 6.75 (d, *J* = 7.0 Hz, 1H), 5.83 (d, *J* = 5.3 Hz, 1H), 5.79 (d, *J* = 2.1 Hz, 1H), 4.94 (d, *J* = 12.1 Hz, 1H), 4.67 (d, *J* = 6.4 Hz, 1H), 4.64 – 4.55 (m, 3H), 4.49 (d, *J* = 11.7 Hz, 1H), 3.73 (dd, *J* = 11.3, 3.0 Hz, 1H), 3.60 (d, *J* = 6.4 Hz, 2H), 2.13 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.3, 161.7, 137.6,

134.3, 133.2, 133.2, 132.8, 132.6, 129.1, 128.7, 128.5, 128.1, 128.0, 127.9, 127.7, 127.2, 126.4, 126.2, 125.7, 92.2, 88.4, 73.7, 73.6, 71.0, 69.3, 68.1, 65.8, 51.6, 20.8. HRMS (ESI): *m/z*: calcd for $C_{34}H_{32}Cl_3NNaO_6S$ [M + Na]: 710.0908; found: 710.0911.



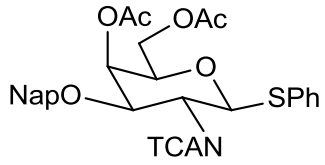
2-76

Compound 2-76: Crude compound **2-71** was treated as described for the preparation of compound **2-75** to afford compound **2-76** as a yellow oil (1.21 g, 52%, 3 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.94 – 7.82 (m, 4H), 7.58 – 7.48 (m, 3H), 7.45 (dd, *J* = 6.7, 2.9 Hz, 2H), δ 7.38 – 7.26 (m, 8H), 6.86 (d, *J* = 7.4 Hz, 1H), 5.86 (d, *J* = 5.1 Hz, 1H), 5.00 (d, *J* = 11.5 Hz, 1H), 4.96 – 4.88 (m, 2H), 4.75 (d, *J* = 12.0 Hz, 1H), 4.67 (d, *J* = 11.6 Hz, 1H), 4.52 – 4.46 (m, 1H), 4.29 (dd, *J* = 11.3, 7.4 Hz, 1H), 4.22 (dd, *J* = 11.4, 5.0 Hz, 1H), 4.10 (s, 1H), 3.80 (dd, *J* = 11.1, 2.3 Hz, 1H), 2.00 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.5, 161.5, 137.6, 134.2, 133.2, 133.1, 132.8, 132.3, 129.0, 128.7, 128.4, 128.2, 127.9, 127.9, 127.7, 126.7, 126.4, 126.3, 125.4, 92.3, 88.2, 77.1, 74.4, 71.8, 71.7, 70.1, 63.2, 51.4, 20.7. HRMS (ESI): *m/z*: calcd for $C_{34}H_{32}Cl_3NNaO_6S$ [M + Na]: 710.0908; found: 710.0906.



2-77

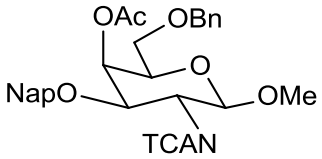
Compound 2-77: Compound **2-73** (2.51g, 4.07 mmol) was treated as described for the preparation of compound **2-75** to afford compound **2-77** as a yellow oil (1.65 g, 55%, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.90 – 7.80 (m, 4H), 7.54 – 7.44 (m, 5H), 7.40 – 7.21 (m, 13H), 6.87 (d, *J* = 7.3 Hz, 1H), 5.85 (d, *J* = 5.1 Hz, 1H), 4.99 (d, *J* = 11.4 Hz, 1H), 4.96 – 4.85 (m, 2H), 4.68 (t, *J* = 11.3 Hz, 2H), 4.57 – 4.47 (m, 3H), 4.22 (s, 1H), 3.81 – 3.71 (m, 2H), 3.66 (dd, *J* = 9.3, 5.9 Hz, 1H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 161.5, 138.1, 137.8, 134.5, 133.3, 133.2, 133.1, 132.3, 129.1, 128.6, 128.4, 128.3, 128.0, 127.9, 127.8, 127.8, 127.7, 126.6, 126.4, 126.2, 125.40, 92.4, 88.7, 77.1, 74.7, 73.5, 72.1, 71.3, 71.0, 68.5, 51.6. HRMS (ESI): *m/z*: calcd for $C_{39}H_{36}Cl_3NNaO_5S$ [M + Na]: 758.1272; found: 758.1281.



2-78

Compound 2-78: Crude compound **2-74** was treated as described for the preparation of compound **2-75** to afford compound **2-78** as a yellow oil (1.49 g, 51%, 3 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.78 – 7.71 (m, 4H), 7.44 – 7.34 (m, 5H), 7.21 (dd, *J* = 4.9, 1.6 Hz, 3H), 5.83 (d, *J* = 5.2 Hz, 1H), 5.65 (d, *J* = 2.4 Hz,

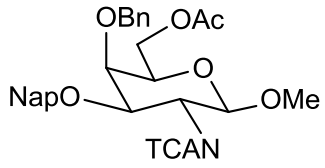
1H), 4.83 (d, *J* = 11.4 Hz, 1H), 4.71 – 4.63 (m, 1H), 4.60 – 4.49 (m, 2H), 4.17 (dd, *J* = 11.4, 5.2 Hz, 1H), 4.08 (dd, *J* = 11.4, 7.5 Hz, 1H), 3.87 (dd, *J* = 11.3, 3.0 Hz, 1H), 3.22 (dq, *J* = 14.3, 7.2 Hz, 1H), 2.07 (s, 3H), 1.97 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.1, 169.9, 161.5, 134.2, 132.8, 132.7, 132.4, 132.1, 128.8, 128.1, 127.7, 127.5, 127.3, 126.7, 125.9, 125.8, 125.5, 87.5, 84.2, 73.2, 70.8, 67.7, 65.4, 62.0, 51.3, 20.4, 20.4. HRMS (ESI): *m/z*: calcd for $C_{29}H_{27}Cl_3NO_7S$ [M + Na]: 638.0579; found: 638.0584.



2-79

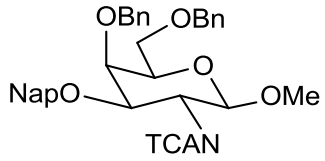
Compound 2-79: Compound **2-75** (3.8 g, 5.51 mmol) was dissolved in CH_2Cl_2 (70 mL) and cooled to -20°C. NIS (2.48g, 11 mmol) pre-dissolved in 10 mL MeOH was added, followed by TMSOTf (1.99 mL, 11 mmol). The reaction was stirred for 2 h at -10°C. When TLC analysis (hexane/EtOAc, 3:1 v/v) showed completion of reaction, the reaction was quenched with TEA (2.5 mL), diluted with CH_2Cl_2 and washed with saturated $Na_2S_2O_3$ (3x), followed by brine (1x). The organic layer was dried over Na_2SO_4, filtered, and concentrated. The product was purified by flash column chromatography (hexane/EtOAc, 5:1 towards 3:1) to afford compound **2-79** as a yellow oil (2.52 g, 75%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.80 (dd, *J* = 12.3, 4.3 Hz, 3H), 7.76 (s, 1H), 7.53 – 7.40 (m, 3H), 7.42 – 7.29 (m, 5H), 6.94 (d, *J* = 7.5 Hz, 1H), 5.73 (d, *J* = 3.0 Hz, 1H), 4.91 (dd, *J* = 10.9, 4.8 Hz, 1H), 4.80 (dd, *J* = 8.3, 5.3 Hz, 1H), 4.66 – 4.54 (m, 2H), 4.50 (dd, *J* =

11.9, 5.0 Hz, 1H), 4.24 (dd, *J* = 10.9, 3.2 Hz, 1H), 3.86 (t, *J* = 6.2 Hz, 1H), 3.72 (dd, *J* = 8.0, 2.7 Hz, 1H), 3.64 (dd, *J* = 9.4, 5.6 Hz, 1H), 3.58 (dd, *J* = 9.3, 7.1 Hz, 1H), 3.53 (s, 3H), 2.07 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.1, 161.9, 137.4, 134.4, 133.0, 133.0, 128.3, 128.2, 127.9, 127.8, 127.8, 127.5, 127.3, 126.1, 126.0, 125.9, 100.5, 92.3, 74.4, 73.5, 72.1, 71.7, 67.7, 65.7, 57.1, 55.5, 20.6. HRMS (ESI): *m/z*: calcd for $C_{29}H_{30}Cl_3NNaO_7$ [M + Na]: 632.0980; found: 632.0988.



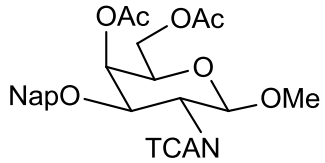
2-80

Compound 2-80: Compound **2-76** (3.28 g, 4.76 mmol) was treated as described for the preparation of compound **2-79** to afford compound **2-80** as a yellow oil (2.04 g, 70%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.96 – 7.73 (m, 4H), 7.56 – 7.40 (m, 3H), 7.34 – 7.19 (m, 5H), 7.03 (d, *J* = 7.1 Hz, 1H), 4.95 (d, *J* = 11.5 Hz, 1H), 4.86 (dd, *J* = 9.9, 3.4 Hz, 2H), 4.77 (d, *J* = 11.4 Hz, 1H), 4.65 (d, *J* = 11.7 Hz, 1H), 4.39 (dd, *J* = 10.9, 2.6 Hz, 1H), 4.27 (dd, *J* = 11.1, 6.6 Hz, 1H), 4.11 (dd, *J* = 11.2, 6.3 Hz, 1H), 3.91 – 3.88 (m, 1H), 3.89 – 3.82 (m, 1H), 3.67 (t, *J* = 6.4 Hz, 1H), 3.51 (s, 3H), 1.97 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.4, 162.1, 137.8, 134.7, 133.2, 133.0, 128.4, 128.3, 128.2, 127.8, 127.8, 127.6, 126.8, 126.2, 126.1, 125.8, 100.0, 92.5, 77.3, 74.4, 72.8, 72.2, 71.9, 62.8, 57.0, 56.0, 20.6. HRMS (ESI): *m/z*: calcd for $C_{29}H_{30}Cl_3NNaO_7$ [M + Na]: 632.0980; found: 632.0985.



2-81

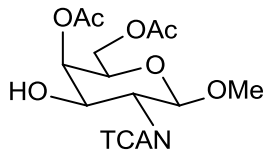
Compound 2-81: Compound **2-77** (2.5 g, 3.4 mmol) was treated as described for the preparation of compound **2-79** to afford compound **2-81** as a yellow oil (1.66 g, 74%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.93 – 7.75 (m, 4H), 7.59 – 7.43 (m, 3H), 7.42 – 7.24 (m, 10H), 7.02 (d, *J* = 7.3 Hz, 1H), 4.94 (d, *J* = 11.5 Hz, 1H), 4.85 (dd, *J* = 20.0, 9.9 Hz, 2H), 4.69 (dd, *J* = 26.7, 11.5 Hz, 2H), 4.51 (q, *J* = 11.8 Hz, 2H), 4.34 (dd, *J* = 11.0, 2.7 Hz, 1H), 4.09 (d, *J* = 2.6 Hz, 1H), 3.92 (dt, *J* = 10.8, 7.9 Hz, 1H), 3.78 – 3.62 (m, 3H), 3.53 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 162.0, 138.2, 137.7, 134.8, 133.1, 133.0, 128.4, 128.3, 128.1, 128.1, 127.8, 127.8, 127.6, 127.5, 126.7, 126.1, 126.0, 125.8, 100.2, 92.5, 77.3, 74.5, 73.5, 73.4, 72.3, 72.2, 68.4, 56.9, 56.0. HRMS (ESI): *m/z*: calcd for $C_{34}H_{34}Cl_3NNaO_6$ [M + Na]: 680.1344; found: 680.1346.



2-82

Compound 2-82: Compound **2-78** (1.8 g, 2.81 mmol) was treated as described for the preparation of compound **2-79** to afford compound **2-82** as a yellow oil (1.07 g, 68%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.83 – 7.76 (m, 3H), 7.73 (s, 1H), 7.50 – 7.43 (m, 2H), 7.40 (dd, *J* = 8.5, 1.2 Hz, 1H), 6.93 (d, *J* = 7.2 Hz, 1H), 5.62

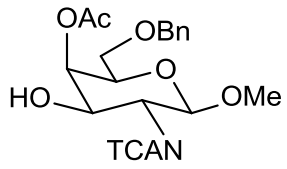
(d, *J* = 2.9 Hz, 1H), 4.88 (d, *J* = 8.9 Hz, 2H), 4.56 (d, *J* = 10.8 Hz, 1H), 4.34 (dd, *J* = 10.9, 3.3 Hz, 1H), 4.20 (d, *J* = 6.6 Hz, 2H), 3.90 (t, *J* = 6.6 Hz, 1H), 3.62 (dt, *J* = 10.9, 7.9 Hz, 1H), 3.52 (s, 3H), 2.15 (s, 3H), 2.09 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.5, 170.3, 162.0, 134.3, 133.2, 133.1, 128.4, 127.9, 127.6, 127.6, 126.2, 126.1, 126.1, 100.2, 92.4, 74.0, 72.1, 70.9, 65.6, 61.9, 57.2, 55.8, 20.8, 20.7. HRMS (ESI): *m/z*: calcd for $C_{24}H_{25}Cl_3NO_8$ [M – H]: 560.0651; found: 560.0655.



A1

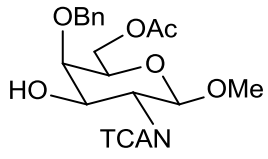
Acceptor A1: To a solution of compound **2-82** (1.51 g, 2.3 mmol) dissolved in CH_2Cl_2 (50 mL) was added DDQ (626 mg, 2.76 mmol) and 3 mL H_2O, and the reaction was stirred for 14 h at room temperature. When TLC analysis (hexane/EtOAc, 2:1 v/v) showed completion of reaction, the reaction was diluted with CH_2Cl_2 and washed with saturated $NaHCO_3$ (3x) and brine (1x). The organic layer was dried over Na_2SO_4, filtered and concentrated. The product was purified by flash column chromatography (hexane/EtOAc, 4:1 towards 2:1) to afford compound **A1** as a white foam (920 mg, 77%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.22 (d, *J* = 7.8 Hz, 1H), 5.31 (d, *J* = 3.1 Hz, 1H), 4.62 (d, *J* = 8.3 Hz, 1H), 4.37 – 4.19 (m, 1H), 4.13 (d, *J* = 6.4 Hz, 2H), 3.87 (t, *J* = 6.3 Hz, 1H), 3.83 – 3.72 (m, 1H), 3.70 (d, *J* = 6.2 Hz, 1H), 3.49 (s, 3H), 2.14 (s, 3H), 2.04 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 171.3, 170.6, 162.8, 101.0, 92.3, 70.9, 69.4, 68.7, 62.0, 57.1,

56.1, 20.7, 20.7. HRMS (ESI): *m/z*: calcd for $C_{13}H_{17}Cl_3NO_8$ [M – H]: 420.0025; found: 420.0022.



A2

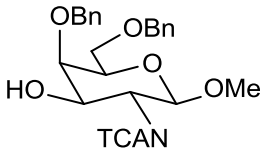
Acceptor A2: Compound **2-79** (2 g, 3.27 mmol) was treated as described for the preparation of **A1** to afford **A2** as a white foam (1.14 g, 74%). ^{1}H NMR (500 MHz, MeOD) δ 7.33 (d, *J* = 4.2 Hz, 4H), 7.29 – 7.25 (m, 1H), 5.35 (d, *J* = 2.8 Hz, 1H), 4.54 (d, *J* = 11.8 Hz, 1H), 4.48 (dd, *J* = 9.7, 7.9 Hz, 2H), 4.00 (dd, *J* = 10.8, 3.2 Hz, 1H), 3.95 – 3.88 (m, 1H), 3.85 (t, *J* = 6.0 Hz, 1H), 3.58 – 3.52 (m, 2H), 3.49 (d, *J* = 14.5 Hz, 3H), 2.09 (s, 3H). ^{13}C NMR (500 MHz, MeOD) δ 172.3, 164.5, 139.2, 129.3, 129.0, 128.8, 103.5, 94.1, 74.5, 73.6, 71.4, 70.3, 69.6, 57.4, 56.4, 20.9. HRMS (ESI): *m/z*: calcd for $C_{18}H_{22}Cl_3NNaO_7$ [M + Na]: 492.0354; found: 492.0361.



A3

Acceptor A3: Compound **2-80** (1.1 g, 1.8 mmol) was treated as described for the preparation of **A1** to afford **A3** as a white foam (610 mg, 72%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.41 – 7.21 (m, 5H), 6.98 (d, *J* = 7.2 Hz, 1H), 4.73 (dd, *J* = 26.3,

11.6 Hz, 2H), 4.51 (d, *J* = 8.3 Hz, 1H), 4.30 (dd, *J* = 11.2, 6.6 Hz, 1H), 4.07 (ddd, *J* = 10.9, 7.9, 4.8 Hz, 2H), 3.85 – 3.74 (m, 2H), 3.67 (t, *J* = 6.5 Hz, 1H), 3.46 (s, 3H), 2.00 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.5, 162.8, 137.5, 128.5, 128.2, 128.0, 100.8, 92.4, 75.4, 72.3, 71.1, 62.6, 56.9, 56.8, 20.7. HRMS (ESI): *m/z*: calcd for $C_{18}H_{22}Cl_3NNaO_7$ [M + Na]: 492.0354; found: 492.0362.



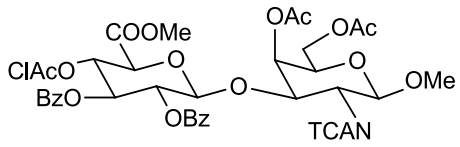
A4

Acceptor A4: Compound **2-81** (1 g, 1.78 mmol) was treated as described for the preparation of **A1** to afford **A4** as a white foam (563 mg, 75%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.39 – 7.25 (m, 10H), 6.90 (d, *J* = 7.3 Hz, 1H), 4.71 (d, *J* = 1.4 Hz, 2H), 4.61 – 4.43 (m, 3H), 4.02 (dd, *J* = 10.8, 3.3 Hz, 1H), 3.91 (d, *J* = 3.2 Hz, 1H), 3.76 (dd, *J* = 7.8, 2.8 Hz, 1H), 3.72 – 3.66 (m, 3H), 3.49 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 162.6, 137.9, 137.6, 128.5, 128.4, 128.0, 127.9, 127.8, 127.8, 100.9, 92.5, 75.7, 75.3, 73.5, 73.5, 70.9, 68.1, 57.1, 56.8. HRMS (ESI): *m/z*: calcd for $C_{23}H_{26}Cl_3NNaO_6$ [M + Na]: 540.0718; found: 540.0710.

General procedure for the synthesis of protected dimers

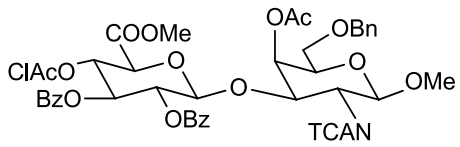
Glycosyl donor (0.385 mmol) and glycosyl acceptor (0.321 mmol) was dissolved in CH_2Cl_2 (8 mL) and 4Å molecular sieves (400 mg) was added. The mixture was stirred at room temperature for 30 min, before cooling to -20°C. TMSOTf (0.064 mmol) was added and the reaction stirred for 2 h at -10°C. When TLC analysis

showed completion of reaction, TEA (0.1 mmol) was added and the reaction filtered, followed by evaporation of solvent. The product was purified by flash column chromatography to afford the expected dimers.



D1-A1

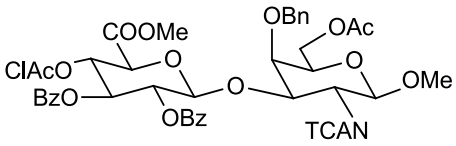
Protected dimer D1-A1: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.86 (dd, *J* = 15.8, 7.6 Hz, 4H), 7.49 (q, *J* = 7.3 Hz, 2H), 7.34 (q, *J* = 7.8 Hz, 4H), 6.77 (d, *J* = 6.7 Hz, 1H), 5.63 (t, *J* = 9.5 Hz, 1H), 5.56 – 5.46 (m, 2H), 5.39 (dd, *J* = 9.4, 7.7 Hz, 1H), 4.88 (dd, *J* = 22.1, 7.9 Hz, 2H), 4.81 (dd, *J* = 10.9, 3.4 Hz, 1H), 4.22 (d, *J* = 10.0 Hz, 1H), 4.13 (ddd, *J* = 18.5, 11.5, 6.4 Hz, 2H), 3.92 (dd, *J* = 15.9, 6.7 Hz, 3H), 3.81 (s, 3H), 3.48 (s, 3H), 3.43 (dd, *J* = 7.3, 2.8 Hz, 1H), 2.15 (s, 3H), 2.07 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.5, 169.8, 166.5, 165.9, 165.5, 164.7, 162.3, 133.6, 133.5, 129.8, 129.8, 128.7, 128.4, 128.4, 100.5, 99.3, 91.9, 73.1, 72.1, 71.9, 71.5, 71.2, 70.7, 68.8, 62.2, 57.4, 56.4, 53.0, 40.2, 20.7, 20.6. HRMS (ESI): *m/z*: calcd for $C_{36}H_{37}Cl_4NNaO_{17}$ [M + Na]: 918.0708; found: 918.0701.



D1-A2

Protected dimer D1-A2: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.86 (dd, *J* = 15.0, 7.6 Hz, 4H), 7.49 (dd, *J* = 14.1, 7.3 Hz, 2H), 7.38 – 7.27 (m, 9H), 6.76 (d, *J* = 6.7 Hz,

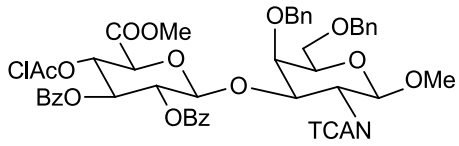
1H), 5.66 – 5.56 (m, 2H), 5.52 (t, *J* = 9.6 Hz, 1H), 5.40 (dd, *J* = 9.4, 7.7 Hz, 1H), 4.88 (dd, *J* = 24.4, 8.0 Hz, 2H), 4.78 (dd, *J* = 10.9, 3.3 Hz, 1H), 4.52 (dd, *J* = 26.6, 11.7 Hz, 2H), 4.20 (d, *J* = 9.9 Hz, 1H), 3.94 (q, *J* = 14.9 Hz, 2H), 3.86 (t, *J* = 5.9 Hz, 1H), 3.75 (s, 3H), 3.56 (d, *J* = 5.9 Hz, 2H), 3.49 (s, 3H), 3.46 – 3.38 (m, 1H), 2.13 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 169.9, 166.5, 165.9, 165.5, 164.7, 162.2, 137.7, 133.5, 133.5, 129.8, 129.8, 128.7, 128.4, 128.4, 127.8, 127.7, 100.4, 99.3, 91.9, 73.6, 73.4, 72.8, 72.1, 72.0, 71.5, 70.7, 69.2, 68.7, 57.4, 56.5, 53.0, 40.2, 20.7. HRMS (ESI): *m/z*: calcd for $C_{41}H_{41}Cl_4NNaO_{16}$ [M + Na]: 966.1072; found: 966.1068.



D1-A3

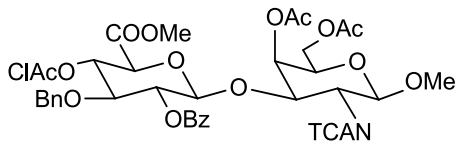
Protected dimer D1-A3: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.87 (dd, *J* = 20.2, 7.3 Hz, 4H), 7.51 – 7.44 (m, 4H), 7.39 – 7.30 (m, 7H), 6.84 (d, *J* = 6.5 Hz, 1H), 5.71 (t, *J* = 9.6 Hz, 1H), 5.55 – 5.44 (m, 2H), 5.03 (d, *J* = 11.4 Hz, 1H), 4.95 (d, *J* = 7.8 Hz, 1H), 4.82 – 4.76 (m, 2H), 4.73 (d, *J* = 11.4 Hz, 1H), 4.28 (d, *J* = 9.9 Hz, 1H), 4.22 (dd, *J* = 11.1, 6.6 Hz, 1H), 4.11 (d, *J* = 2.6 Hz, 1H), 4.04 (dd, *J* = 11.2, 6.3 Hz, 1H), 3.96 (q, *J* = 14.9 Hz, 2H), 3.80 (s, 3H), 3.71 (t, *J* = 6.4 Hz, 1H), 3.59 – 3.50 (m, 1H), 3.43 (s, 3H), 2.00 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.5, 166.6, 166.2, 165.5, 164.8, 162.5, 138.1, 133.6, 133.5, 129.9, 129.8, 129.0, 128.7, 128.5, 128.4, 128.3, 128.3, 127.7, 101.2, 99.1, 91.9, 76.6, 74.9, 74.9, 71.9, 71.8,

71.4, 70.9, 62.6, 57.2, 56.3, 53.0, 40.2, 20.7. HRMS (ESI): *m/z*: calcd for $C_{41}H_{41}Cl_4NNaO_{16}$ [M + Na]: 966.1072; found: 966.1068.



D1-A4

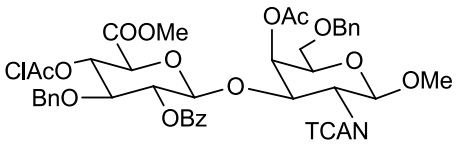
Protected dimer D1-A4: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.88 (dd, *J* = 20.3, 7.6 Hz, 4H), 7.46 (dd, *J* = 29.2, 7.3 Hz, 4H), 7.39 – 7.27 (m, 12H), 6.80 (d, *J* = 6.5 Hz, 1H), 5.68 (d, *J* = 9.7 Hz, 1H), 5.49 (dd, *J* = 5.7, 4.1 Hz, 2H), 5.01 (d, *J* = 11.5 Hz, 1H), 4.95 (d, *J* = 7.9 Hz, 1H), 4.81 (d, *J* = 8.3 Hz, 2H), 4.68 (d, *J* = 11.4 Hz, 1H), 4.56 – 4.50 (m, 1H), 4.43 (dd, *J* = 34.7, 11.6 Hz, 2H), 4.25 (d, *J* = 10.0 Hz, 1H), 4.19 (d, *J* = 1.9 Hz, 1H), 4.00 – 3.91 (m, 2H), 3.74 – 3.66 (m, 4H), 3.62 – 3.53 (m, 2H), 3.43 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 166.6, 166.1, 165.5, 164.8, 162.4, 138.5, 137.9, 133.6, 133.4, 129.9, 129.7, 128.8, 128.5, 128.4, 128.4, 128.4, 128.2, 128.1, 128.0, 127.9, 127.8, 127.7, 127.6, 101.2, 99.2, 92.0, 75.7, 75.5, 75.0, 73.5, 73.3, 71.9, 71.8, 71.4, 70.8, 68.6, 57.2, 56.5, 52.9, 40.2. HRMS (ESI): *m/z*: calcd for $C_{46}H_{45}Cl_4NNaO_{15}$ [M + Na]: 1014.1436; found: 1014.1441.



D2-A1

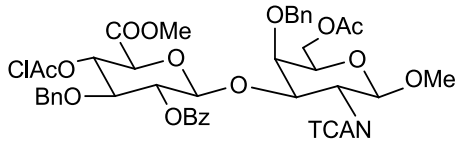
Protected dimer D2-A1: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.99 – 7.88 (m, 2H), 7.59 (t, *J* = 7.5 Hz, 1H), 7.44 (t, *J* = 7.8 Hz, 2H), 7.17 (dd, *J* = 5.0, 1.7 Hz, 3H),

7.08 (dd, *J* = 6.4, 3.0 Hz, 2H), 6.80 (s, 1H), 5.50 (d, *J* = 3.4 Hz, 1H), 5.36 – 5.19 (m, 2H), 4.83 (d, *J* = 8.4 Hz, 1H), 4.80 – 4.71 (m, 2H), 4.55 (dd, *J* = 32.2, 11.7 Hz, 2H), 4.20 – 4.06 (m, 2H), 4.00 (d, *J* = 9.9 Hz, 1H), 3.85 (ddd, *J* = 31.9, 11.8, 6.3 Hz, 4H), 3.76 (s, 3H), 3.48 – 3.41 (m, 4H), 2.12 (s, 3H), 2.06 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.5, 169.8, 167.0, 165.8, 164.6, 162.2, 137.1, 133.5, 129.8, 129.2, 128.5, 128.3, 127.9, 127.9, 100.3, 99.3, 92.0, 78.9, 74.1, 72.9, 72.8, 72.2, 72.0, 71.3, 68.7, 62.3, 57.3, 56.3, 52.9, 40.3, 20.7, 20.6. HRMS (ESI): *m/z*: calcd for $C_{36}H_{38}Cl_4NO_{16}$ [M – H]: 880.0950; found: 880.0946.



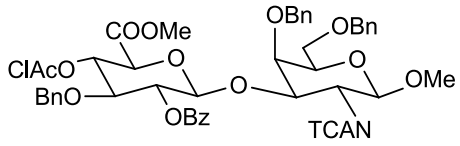
D2-A2

Protected dimer D2-A2: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.97 (dd, *J* = 8.0, 0.9 Hz, 2H), 7.58 (dd, *J* = 10.8, 4.2 Hz, 1H), 7.44 (t, *J* = 7.8 Hz, 2H), 7.36 – 7.25 (m, 5H), 7.21 – 7.14 (m, 3H), 7.08 (dd, *J* = 6.3, 3.0 Hz, 2H), 6.81 (d, *J* = 6.8 Hz, 1H), 5.54 (d, *J* = 3.3 Hz, 1H), 5.38 – 5.24 (m, 2H), 4.80 (dd, *J* = 20.7, 7.9 Hz, 2H), 4.73 – 4.66 (m, 1H), 4.59 – 4.42 (m, 4H), 4.00 (d, *J* = 9.9 Hz, 1H), 3.89 (d, *J* = 14.8 Hz, 1H), 3.85 – 3.77 (m, 3H), 3.71 (s, 3H), 3.55 (dd, *J* = 5.8, 3.1 Hz, 2H), 3.47 (d, *J* = 4.0 Hz, 4H), 2.10 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 169.8, 167.0, 165.7, 164.6, 162.1, 137.8, 137.1, 133.5, 129.8, 129.2, 128.5, 128.3, 128.3, 127.9, 127.8, 127.8, 127.7, 100.2, 99.4, 92.0, 78.9, 74.1, 73.6, 73.2, 72.9, 72.8, 72.2, 72.0, 69.1, 68.8, 57.3, 56.4, 52.8, 40.3, 20.6. HRMS (ESI): *m/z*: calcd for $C_{41}H_{43}Cl_4NNaO_{15}$ [M + Na]: 952.1279; found: 952.1285.



D2-A3

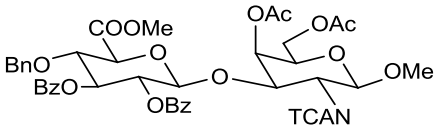
Protected dimer D2-A3: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.00 (d, *J* = 7.6 Hz, 2H), 7.59 (t, *J* = 7.4 Hz, 1H), 7.44 (t, *J* = 7.1 Hz, 4H), 7.35 (t, *J* = 7.4 Hz, 2H), 7.28 (dd, *J* = 13.5, 6.4 Hz, 1H), 7.21 – 7.15 (m, 3H), 7.09 (dd, *J* = 6.2, 2.4 Hz, 2H), 6.87 (d, *J* = 6.5 Hz, 1H), 5.41 (dd, *J* = 8.9, 8.2 Hz, 1H), 5.30 (t, *J* = 9.5 Hz, 1H), 5.01 (d, *J* = 11.5 Hz, 1H), 4.82 – 4.68 (m, 4H), 4.59 (dd, *J* = 32.3, 11.7 Hz, 2H), 4.20 (dd, *J* = 11.1, 6.6 Hz, 1H), 4.08 (d, *J* = 10.1 Hz, 2H), 4.02 (dd, *J* = 11.1, 6.4 Hz, 1H), 3.95 – 3.79 (m, 3H), 3.75 (s, 3H), 3.69 (dd, *J* = 11.5, 5.0 Hz, 1H), 3.58 – 3.48 (m, 1H), 3.42 (d, *J* = 5.5 Hz, 3H), 1.99 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.5, 167.0, 166.0, 164.6, 162.3, 138.1, 137.1, 133.5, 129.9, 129.1, 128.9, 128.5, 128.3, 128.2, 127.9, 127.9, 127.6, 101.1, 99.2, 92.0, 79.1, 76.2, 74.9, 74.8, 74.6, 72.9, 72.4, 71.9, 62.6, 57.1, 56.3, 52.8, 40.3, 20.6. HRMS (ESI): *m/z*: calcd for $C_{41}H_{43}Cl_4NNaO_{15}$ [M + Na]: 952.1279; found: 952.1276.



D2-A4

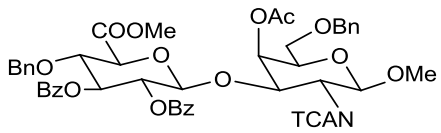
Protected dimer D2-A4: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.02 – 7.95 (m, 2H), 7.63 – 7.56 (m, 1H), 7.48 – 7.39 (m, 4H), 7.37 – 7.26 (m, 8H), 7.23 – 7.14 (m, 3H), 7.10 – 7.04 (m, 2H), 6.79 (d, *J* = 6.5 Hz, 1H), 5.39 (dd, *J* = 9.2, 7.9 Hz, 1H), 5.33 – 5.23 (m, 1H), 4.99 (d, *J* = 11.5 Hz, 1H), 4.79 (dd, *J* = 14.1, 8.0 Hz, 2H),

4.71 (dd, *J* = 11.0, 2.9 Hz, 1H), 4.67 – 4.60 (m, 2H), 4.55 (d, *J* = 11.6 Hz, 1H), 4.41 (dd, *J* = 34.8, 11.7 Hz, 2H), 4.16 (d, *J* = 2.6 Hz, 1H), 4.04 (d, *J* = 10.0 Hz, 1H), 3.96 – 3.74 (m, 3H), 3.68 (t, *J* = 6.3 Hz, 1H), 3.65 (s, 3H), 3.54 (dddd, *J* = 14.5, 10.7, 8.7, 6.0 Hz, 3H), 3.43 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 167.1, 166.0, 164.6, 162.3, 138.5, 137.9, 137.2, 133.5, 130.0, 129.2, 128.8, 128.5, 128.47, 128.1, 127.9, 127.8, 127.7, 127.5, 101.1, 99.2, 92.1, 79.1, 76.4, 75.5, 74.9, 74.5, 73.5, 73.5, 72.9, 72.4, 71.9, 68.7, 57.2, 56.6, 52.8, 40.3. HRMS (ESI): *m/z*: calcd for $C_{48}H_{46}Cl_4NO_{14}$ [M + Na]: 1000.1667; found: 1000.1660.



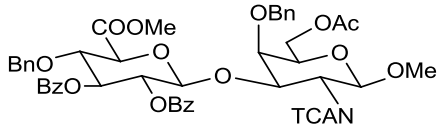
D3-A1

Protected dimer D3-A1: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.94 – 7.82 (m, 4H), 7.48 (td, *J* = 7.2, 1.2 Hz, 2H), 7.34 (t, *J* = 7.6 Hz, 4H), 7.18 – 7.10 (m, 3H), 7.10 – 7.02 (m, 2H), 6.77 (d, *J* = 6.8 Hz, 1H), 5.61 (t, *J* = 9.1 Hz, 1H), 5.50 (d, *J* = 3.4 Hz, 1H), 5.31 (dd, *J* = 9.3, 7.5 Hz, 1H), 4.87 (d, *J* = 7.9 Hz, 2H), 4.73 (dd, *J* = 10.8, 3.5 Hz, 1H), 4.53 (dd, *J* = 33.5, 11.1 Hz, 2H), 4.13 (ddt, *J* = 18.7, 11.6, 8.2 Hz, 4H), 3.90 (t, *J* = 6.3 Hz, 1H), 3.79 (s, 3H), 3.49 – 3.40 (m, 4H), 2.14 (s, 3H), 2.07 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.5, 169.8, 168.0, 165.3, 165.0, 162.2, 136.9, 133.3, 133.3, 129.8, 129.6, 129.0, 128.9, 128.4, 128.3, 128.2, 128.0, 127.9, 100.6, 99.3, 91.8, 74.7, 74.4, 74.0, 73.0, 71.8, 71.2, 68.8, 62.3, 57.3, 56.3, 52.7, 20.7, 20.6. HRMS (ESI): *m/z*: calcd for $C_{41}H_{42}Cl_3NNaO_{16}$ [M + Na]: 932.1461; found: 932.1458.



D3-A2

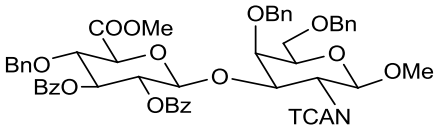
Protected dimer D3-A2: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.95 – 7.81 (m, 4H), 7.48 (ddd, *J* = 7.3, 3.5, 1.4 Hz, 2H), 7.37 – 7.31 (m, 8H), 7.30 – 7.25 (m, 1H), 7.16 – 7.12 (m, 3H), 7.09 (dd, *J* = 6.7, 2.9 Hz, 2H), 6.77 (d, *J* = 7.0 Hz, 1H), 5.63 (t, *J* = 9.1 Hz, 1H), 5.57 (d, *J* = 3.4 Hz, 1H), 5.33 (dd, *J* = 9.1, 7.4 Hz, 1H), 4.89 (dd, *J* = 30.2, 7.8 Hz, 2H), 4.71 (dd, *J* = 10.8, 3.5 Hz, 1H), 4.55 (dt, *J* = 23.4, 11.4 Hz, 4H), 4.22 (t, *J* = 9.3 Hz, 1H), 4.13 (d, *J* = 9.5 Hz, 1H), 3.86 (t, *J* = 6.0 Hz, 1H), 3.75 (s, 3H), 3.58 (dd, *J* = 5.9, 2.3 Hz, 2H), 3.53 – 3.44 (m, 4H), 2.12 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 169.8, 168.0, 165.3, 164.9, 162.0, 137.8, 137.0, 133.2, 133.1, 129.8, 129.6, 129.1, 129.0, 128.3, 128.3, 128.2, 127.9, 127.7, 127.7, 127.6, 100.5, 99.5, 91.9, 77.0, 74.6, 74.4, 74.1, 73.5, 73.3, 72.8, 72.0, 69.3, 68.7, 57.2, 56.3, 52.5, 20.6. HRMS (ESI): *m/z*: calcd for $C_{46}H_{46}Cl_3NNaO_{15}$ [M + Na]: 980.1825; found: 980.1820.



D3-A3

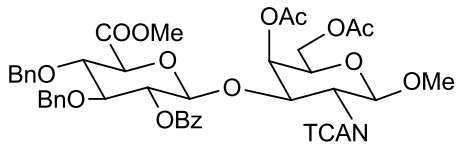
Protected dimer D3-A3: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.90 (dd, *J* = 12.5, 4.4 Hz, 4H), 7.55 – 7.42 (m, 4H), 7.40 – 7.28 (m, 7H), 7.20 – 7.14 (m, 3H), 7.10 (dd, *J* = 6.5, 2.9 Hz, 2H), 6.81 (d, *J* = 6.5 Hz, 1H), 5.70 (dd, *J* = 9.3, 8.6 Hz, 1H), 5.48 (dd, *J* = 9.6, 8.0 Hz, 1H), 5.02 (d, *J* = 11.4 Hz, 1H), 4.92 (d, *J* = 7.8 Hz, 1H), 4.82

(d, *J* = 8.3 Hz, 1H), 4.74 (dd, *J* = 10.9, 2.9 Hz, 1H), 4.69 (d, *J* = 11.4 Hz, 1H), 4.56 (dd, *J* = 29.3, 11.1 Hz, 2H), 4.19 (dt, *J* = 11.3, 8.5 Hz, 3H), 4.00 (dd, *J* = 11.3, 5.9 Hz, 2H), 3.79 (s, 3H), 3.69 (t, *J* = 6.4 Hz, 1H), 3.56 – 3.47 (m, 1H), 3.43 (s, 3H), 1.98 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.4, 168.3, 165.3, 165.1, 162.4, 138.0, 136.8, 133.3, 133.3, 129.9, 129.6, 129.0, 128.9, 128.4, 128.3, 128.3, 128.2, 128.0, 127.9, 127.7, 101.5, 99.1, 91.9, 77.4, 76.3, 74.7, 74.6, 74.3, 73.8, 71.9, 71.8, 62.7, 57.1, 56.4, 52.6, 20.6. HRMS (ESI): *m/z*: calcd for $C_{46}H_{45}Cl_3NO_{15}$ [M – H]: 956.1860; found: 956.1867.



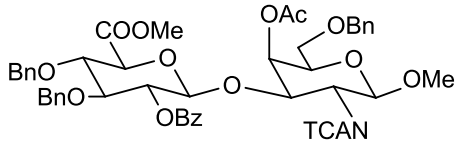
D3-A4

Protected dimer D3-A4: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.94 – 7.84 (m, 4H), 7.54 – 7.43 (m, 2H), 7.44 – 7.37 (m, 2H), 7.40 – 7.23 (m, 12H), 7.18 – 7.13 (m, 3H), 7.09 (dd, *J* = 6.5, 2.9 Hz, 2H), 6.73 (d, *J* = 6.6 Hz, 1H), 5.72 – 5.60 (m, 1H), 5.44 (dd, *J* = 9.6, 7.8 Hz, 1H), 5.00 (d, *J* = 11.4 Hz, 1H), 4.91 (d, *J* = 7.7 Hz, 1H), 4.82 (d, *J* = 8.3 Hz, 1H), 4.71 (dd, *J* = 11.0, 2.8 Hz, 1H), 4.66 – 4.47 (m, 3H), 4.45 – 4.34 (m, 2H), 4.21 – 4.13 (m, 2H), 4.07 (d, *J* = 2.6 Hz, 1H), 3.78 – 3.67 (m, 4H), 3.59 – 3.44 (m, 3H), 3.43 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 168.3, 165.4, 165.1, 162.3, 138.5, 136.9, 133.3, 129.9, 129.7, 129.0, 128.9, 128.4, 128.4, 128.3, 128.2, 128.0, 127.9, 127.8, 127.7, 127.6, 101.5, 99.2, 92.0, 77.4, 76.5, 75.3, 74.8, 74.7, 74.3, 73.9, 73.4, 73.3, 71.8, 68.7, 57.2, 56.6, 52.6. HRMS (ESI): *m/z*: calcd for $C_{51}H_{49}Cl_3NO_{14}$ [M – H]: 1004.2224; found: 1004.2217.



D4-A1

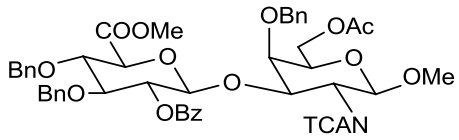
Protected dimer D4-A1: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.94 (dd, *J* = 8.1, 0.9 Hz, 2H), 7.56 (t, *J* = 7.5 Hz, 1H), 7.42 (t, *J* = 7.8 Hz, 2H), 7.33 – 7.21 (m, 5H), 7.14 – 7.05 (m, 5H), 6.84 (d, *J* = 7.0 Hz, 1H), 5.47 (d, *J* = 3.3 Hz, 1H), 5.22 (dd, *J* = 8.3, 7.4 Hz, 1H), 4.84 (d, *J* = 8.4 Hz, 1H), 4.78 – 4.59 (m, 6H), 4.16 (dd, *J* = 11.6, 5.4 Hz, 1H), 4.07 (dd, *J* = 11.6, 7.2 Hz, 1H), 4.02 – 3.98 (m, 2H), 3.87 (t, *J* = 6.3 Hz, 1H), 3.76 (d, *J* = 6.0 Hz, 3H), 3.74 – 3.70 (m, 1H), 3.52 – 3.40 (m, 4H), 2.08 (s, 3H), 2.06 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.5, 169.8, 168.4, 164.8, 162.0, 137.5, 137.3, 133.3, 129.8, 129.3, 128.4, 128.3, 128.2, 127.9, 127.9, 127.7, 100.4, 99.4, 91.9, 81.4, 78.8, 74.9, 74.6, 74.2, 73.1, 72.6, 71.2, 68.7, 62.3, 57.2, 56.2, 52.5, 20.6, 20.5. HRMS (ESI): *m/z*: calcd for $C_{41}H_{44}Cl_3NNaO_{15}$ [M + Na]: 918.1669; found: 918.1673.



D4-A2

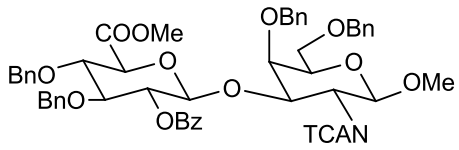
Protected dimer D4-A2: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.58 (dd, *J* = 5.9, 1.9 Hz, 2H), 7.35 – 7.26 (m, 16H), 7.16 – 7.12 (m, 2H), 6.71 (d, *J* = 7.4 Hz, 1H), 6.05 (d, *J* = 4.8 Hz, 1H), 5.63 (d, *J* = 3.1 Hz, 1H), 4.86 (d, *J* = 8.3 Hz, 1H), 4.69 (d, *J* = 11.7 Hz, 1H), 4.53 (dt, *J* = 20.2, 11.4 Hz, 5H), 4.44 (d, *J* = 11.4 Hz, 1H), 4.38 (dd,

J = 10.9, 3.3 Hz, 1H), 3.95 (d, *J* = 7.9 Hz, 1H), 3.83 – 3.71 (m, 3H), 3.65 (s, 3H), 3.59 – 3.53 (m, 3H), 3.51 (d, *J* = 5.2 Hz, 3H), 2.05 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 169.9, 169.8, 162.2, 137.6, 137.6, 137.5, 137.4, 129.8, 129.4, 128.4, 128.3, 128.2, 127.9, 127.9, 127.8, 127.8, 127.7, 126.0, 122.4, 100.2, 97.7, 92.3, 78.0, 77.7, 75.7, 73.7, 73.2, 72.7, 72.5, 71.8, 69.5, 68.9, 68.4, 57.3, 55.2, 52.3, 20.8. HRMS (ESI): *m/z*: calcd for $C_{46}H_{48}Cl_3NNaO_{14}$ [M + Na]: 966.2033; found: 966.2038.



D4-A3

Protected dimer D4-A3: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.03 – 7.94 (m, 2H), 7.61 – 7.54 (m, 1H), 7.47 – 7.38 (m, 4H), 7.36 – 7.23 (m, 8H), 7.18 – 7.06 (m, 5H), 6.83 (d, *J* = 6.7 Hz, 1H), 5.41 (dd, *J* = 9.0, 7.9 Hz, 1H), 5.00 (d, *J* = 11.4 Hz, 1H), 4.83 – 4.72 (m, 4H), 4.67 (ddd, *J* = 14.4, 11.9, 5.3 Hz, 4H), 4.19 (dd, *J* = 11.2, 6.7 Hz, 1H), 4.10 – 3.94 (m, 4H), 3.87 – 3.80 (m, 1H), 3.77 (s, 3H), 3.68 (t, *J* = 6.4 Hz, 1H), 3.54 (ddd, *J* = 10.9, 8.2, 6.8 Hz, 1H), 3.43 (s, 3H), 1.98 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.4, 168.5, 164.9, 162.2, 138.0, 137.4, 137.2, 133.3, 129.9, 129.3, 128.9, 128.4, 128.4, 128.2, 128.1, 127.9, 127.9, 127.7, 127.6, 101.3, 99.1, 92.0, 81.4, 79.4, 75.8, 75.1, 75.0, 74.55, 74.5, 74.2, 73.2, 71.9, 62.7, 57.0, 56.3, 52.5, 20.6. HRMS (ESI): *m/z*: calcd for $C_{46}H_{48}Cl_3NNaO_{14}$ [M + Na]: 966.2033; found: 966.2039.

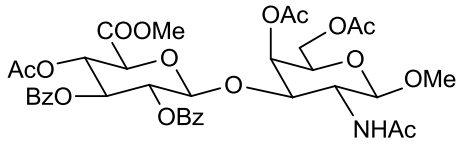


D4-A4

Protected dimer D4-A4: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.03 (d, *J* = 7.3 Hz, 2H), 7.60 (t, *J* = 7.5 Hz, 1H), 7.51 – 7.40 (m, 4H), 7.40 – 7.26 (m, 13H), 7.21 – 7.10 (m, 5H), 6.90 (d, *J* = 6.7 Hz, 1H), 5.48 – 5.37 (m, 1H), 5.02 (d, *J* = 11.4 Hz, 1H), 4.90 – 4.75 (m, 4H), 4.69 (ddd, *J* = 23.2, 12.9, 7.2 Hz, 4H), 4.43 (dd, *J* = 39.1, 11.7 Hz, 2H), 4.14 – 4.07 (m, 2H), 4.09 – 4.02 (m, 1H), 3.85 (t, *J* = 8.8 Hz, 1H), 3.74 – 3.67 (m, 4H), 3.64 – 3.52 (m, 3H), 3.45 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 168.4, 164.9, 162.1, 138.4, 137.8, 137.4, 137.2, 133.2, 129.8, 129.3, 128.7, 128.4, 128.3, 128.2, 128.1, 128.0, 127.9, 127.8, 127.7, 127.6, 127.6, 127.4, 101.2, 99.2, 92.0, 81.4, 79.3, 76.0, 75.2, 75.0, 74.9, 74.7, 74.1, 73.3, 73.2, 68.6, 57.0, 56.3, 52.4. HRMS (ESI): *m/z*: calcd for $C_{51}H_{52}Cl_3NNaO_{13}$ [M + Na]: 1014.2396; found: 1014.2401.

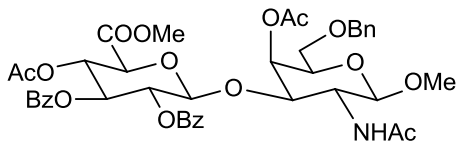
General procedure for the synthesis of compounds 2-83a – 2-83h

The protected dimer (0.25 mmol) was dissolved in toluene (8 mL) and Bu_3SnH (2 mmol) and ABCN (145 mg) was added. The reaction was heated to 100°C for 4 h. When TLC analysis showed completion of reaction, the solvent was evaporated off and the product was purified by flash column chromatography to afford **2-83a – 2-83h**.



2-83a

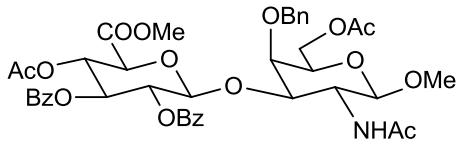
Compound 2-83a: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.99 – 7.80 (m, 4H), 7.57 – 7.42 (m, 2H), 7.40 – 7.29 (m, 4H), 5.65 (t, *J* = 9.6 Hz, 1H), 5.50 – 5.34 (m, 4H), 4.93 (d, *J* = 8.3 Hz, 1H), 4.85 (d, *J* = 7.6 Hz, 1H), 4.80 (dd, *J* = 10.8, 3.4 Hz, 1H), 4.20 – 4.12 (m, 2H), 4.04 (dd, *J* = 11.6, 7.3 Hz, 1H), 3.89 – 3.84 (m, 1H), 3.79 (s, 3H), 3.44 (s, 3H), 3.12 (dt, *J* = 10.7, 7.8 Hz, 1H), 2.12 (s, 3H), 2.05 (s, 3H), 1.93 (s, 3H), 1.44 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 171.2, 170.5, 169.9, 169.2, 166.8, 165.5, 164.7, 133.5, 133.4, 129.8, 129.7, 129.0, 128.6, 128.6, 128.4, 101.0, 99.7, 75.1, 72.6, 72.2, 71.6, 71.2, 69.2, 68.7, 62.6, 57.0, 55.5, 52.8, 23.0, 20.7, 20.4. HRMS (ESI): *m/z*: calcd for $C_{36}H_{41}NNaO_{17}$ [M + Na]: 782.2267; found: 782.2271.



2-83b

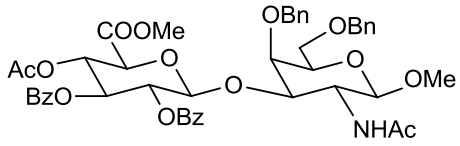
Compound 2-83b: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.90 (dd, *J* = 24.9, 7.5 Hz, 4H), 7.59 – 7.47 (m, 2H), 7.40 – 7.22 (m, 9H), 5.64 (t, *J* = 9.6 Hz, 1H), 5.53 – 5.34 (m, 4H), 4.93 (d, *J* = 8.2 Hz, 1H), 4.85 (d, *J* = 7.6 Hz, 1H), 4.78 (dd, *J* = 10.8, 3.3 Hz, 1H), 4.51 (q, *J* = 11.7 Hz, 2H), 4.15 (d, *J* = 9.9 Hz, 1H), 3.84 (t, *J* = 5.8 Hz, 1H), 3.76 (s, 3H), 3.55 (ddd, *J* = 17.0, 10.2, 5.9 Hz, 2H), 3.45 (s, 3H), 3.13 (dt, *J* = 10.5, 7.8 Hz, 1H), 2.10 (s, 3H), 1.93 (s, 3H), 1.43 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$)

δ 171.1, 170.0, 169.2, 166.8, 165.5, 164.7, 137.9, 133.5, 133.4, 129.8, 129.7, 129.0, 128.6, 128.6, 128.4, 128.4, 127.7, 127.6, 100.9, 99.7, 75.3, 73.5, 72.9, 72.6, 72.3, 71.6, 69.2, 57.1, 55.6, 52.8, 34.0, 23.0, 20.8, 20.4. HRMS (ESI): *m/z*: calcd for $C_{41}H_{45}NNaO_{16}$ [M + Na]: 830.2631; found: 830.2640.



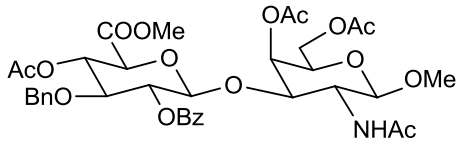
2-83c

Compound 2-83c: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.93 (dd, *J* = 23.4, 7.5 Hz, 4H), 7.57 – 7.26 (m, 11H), 5.74 (t, *J* = 9.7 Hz, 1H), 5.52 – 5.39 (m, 2H), 5.30 (d, *J* = 6.6 Hz, 1H), 5.01 (d, *J* = 11.6 Hz, 1H), 4.90 (d, *J* = 7.9 Hz, 1H), 4.81 (dd, *J* = 13.0, 5.2 Hz, 2H), 4.71 (d, *J* = 11.6 Hz, 1H), 4.23 (d, *J* = 10.0 Hz, 1H), 4.16 (dd, *J* = 11.1, 6.7 Hz, 1H), 4.02 (d, *J* = 1.6 Hz, 1H), 3.97 (dd, *J* = 11.2, 6.1 Hz, 1H), 3.78 (s, 3H), 3.67 (t, *J* = 6.3 Hz, 1H), 3.40 (s, 3H), 3.19 (dt, *J* = 10.7, 7.8 Hz, 1H), 1.96 (d, *J* = 3.1 Hz, 6H), 1.49 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 171.4, 170.5, 169.4, 167.0, 165.6, 164.6, 138.4, 133.5, 129.8, 129.6, 129.2, 129.0, 128.6, 128.5, 128.2, 127.6, 102.0, 99.6, 78.2, 74.9, 74.6, 72.3, 72.0, 71.8, 70.6, 69.4, 62.9, 56.8, 55.6, 52.8, 23.3, 20.7, 20.4. HRMS (ESI): *m/z*: calcd for $C_{41}H_{45}NNaO_{16}$ [M + Na]: 830.2631; found: 830.2626.



2-83d

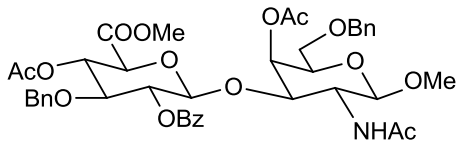
Compound 2-83d: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.92 (dd, *J* = 24.8, 7.5 Hz, 4H), 7.55 – 7.46 (m, 2H), 7.43 – 7.25 (m, 14H), 5.72 (t, *J* = 9.7 Hz, 1H), 5.46 (dd, *J* = 13.6, 6.1 Hz, 2H), 5.28 (d, *J* = 6.5 Hz, 1H), 5.00 (d, *J* = 11.7 Hz, 1H), 4.90 (d, *J* = 7.8 Hz, 1H), 4.84 (d, *J* = 8.2 Hz, 1H), 4.78 (d, *J* = 2.7 Hz, 1H), 4.66 (d, *J* = 11.7 Hz, 1H), 4.39 (dd, *J* = 36.8, 11.7 Hz, 2H), 4.20 (d, *J* = 10.0 Hz, 1H), 4.10 (s, 1H), 3.68 (d, *J* = 5.8 Hz, 4H), 3.56 (dd, *J* = 9.5, 5.7 Hz, 1H), 3.49 (dd, *J* = 9.4, 6.8 Hz, 1H), 3.40 (s, 3H), 3.19 (dd, *J* = 7.2, 2.8 Hz, 1H), 1.96 (s, 3H), 1.47 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 171.3, 169.4, 167.0, 165.6, 164.7, 138.8, 138.1, 133.5, 129.8, 129.7, 129.3, 128.8, 128.7, 128.6, 128.5, 128.3, 128.1, 127.8, 127.6, 127.4, 102.0, 99.6, 78.4, 75.6, 74.7, 73.4, 73.3, 72.3, 72.1, 71.8, 69.4, 69.0, 56.8, 55.7, 52.7, 23.3, 20.4. HRMS (ESI): *m/z*: calcd for $C_{46}H_{49}NNaO_{15}$ [M + Na]: 878.2994; found: 878.3002.



2-83e

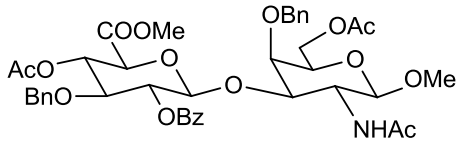
Compound 2-83e: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.97 (d, *J* = 7.6 Hz, 2H), 7.59 (t, *J* = 7.4 Hz, 1H), 7.45 (t, *J* = 7.7 Hz, 2H), 7.19 – 7.11 (m, 3H), 7.08 (dd, *J* = 6.4, 2.5 Hz, 2H), 5.53 (d, *J* = 6.7 Hz, 1H), 5.38 (d, *J* = 3.2 Hz, 1H), 5.31 – 5.21 (m, 2H), 4.91 (d, *J* = 8.2 Hz, 1H), 4.71 (dd, *J* = 14.3, 5.5 Hz, 2H), 4.56 (s, 2H), 4.15

(dd, *J* = 11.6, 4.9 Hz, 1H), 4.01 (dd, *J* = 11.6, 7.4 Hz, 1H), 3.95 (d, *J* = 9.8 Hz, 1H), 3.84 (dd, *J* = 11.4, 6.9 Hz, 2H), 3.75 (s, 3H), 3.42 (s, 3H), 3.14 (dt, *J* = 10.4, 7.8 Hz, 1H), 2.08 (s, 3H), 2.04 (s, 3H), 1.97 (s, 3H), 1.38 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 171.1, 170.5, 170.0, 169.2, 167.2, 164.7, 137.2, 133.5, 129.7, 128.6, 128.3, 127.8, 127.8, 100.7, 99.7, 79.0, 75.1, 73.9, 72.8, 71.3, 70.7, 68.4, 62.7, 57.0, 55.3, 52.7, 23.0, 20.7, 20.6. HRMS (ESI): *m/z*: calcd for $C_{36}H_{42}NO_{16}$ [M – H]: 744.2509; found: 744.2503.



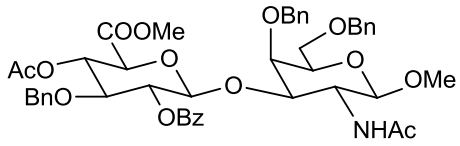
2-83f

Compound 2-83f: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.98 (d, *J* = 7.9 Hz, 2H), 7.59 (t, *J* = 7.5 Hz, 1H), 7.45 (t, *J* = 7.7 Hz, 2H), 7.32 (dd, *J* = 7.2, 6.3 Hz, 4H), 7.26 (dd, *J* = 8.4, 4.9 Hz, 1H), 7.18 – 7.03 (m, 5H), 5.52 (d, *J* = 6.7 Hz, 1H), 5.41 (d, *J* = 3.2 Hz, 1H), 5.28 (dd, *J* = 18.3, 9.0 Hz, 2H), 4.93 (d, *J* = 8.2 Hz, 1H), 4.78 – 4.65 (m, 2H), 4.57 (d, *J* = 12.2 Hz, 2H), 4.54 – 4.44 (m, 2H), 3.95 (d, *J* = 9.7 Hz, 1H), 3.88 – 3.79 (m, 2H), 3.73 (s, 3H), 3.57 (dd, *J* = 10.2, 4.8 Hz, 1H), 3.50 (dd, *J* = 10.3, 6.9 Hz, 1H), 3.44 (s, 3H), 3.15 (dt, *J* = 10.5, 7.7 Hz, 1H), 2.07 (s, 3H), 1.98 (s, 3H), 1.38 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 171.0, 170.1, 169.2, 167.2, 164.6, 137.9, 137.2, 133.5, 129.7, 129.3, 128.5, 128.3, 128.2, 127.8, 127.7, 127.7, 127.6, 100.6, 99.7, 79.0, 75.4, 73.8, 73.5, 72.9, 72.8, 72.7, 70.7, 69.3, 68.8, 57.1, 55.4, 52.6, 22.9, 20.7, 20.6. HRMS (ESI): *m/z*: calcd for $C_{41}H_{47}NNaO_{15}$ [M + Na]: 816.2838; found: 816.2833.



2-83g

Compound 2-83g: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.02 – 7.92 (m, 2H), 7.60 (t, *J* = 7.5 Hz, 1H), 7.48 – 7.38 (m, 4H), 7.34 (t, *J* = 7.4 Hz, 2H), 7.30 – 7.24 (m, 1H), 7.19 – 7.09 (m, 5H), 5.42 – 5.31 (m, 2H), 5.27 (t, *J* = 9.5 Hz, 1H), 4.99 (d, *J* = 11.6 Hz, 1H), 4.78 (dd, *J* = 13.2, 8.1 Hz, 2H), 4.74 – 4.65 (m, 2H), 4.62 (s, 2H), 4.15 (dd, *J* = 11.2, 6.7 Hz, 1H), 4.01 (t, *J* = 13.3 Hz, 2H), 3.95 (dt, *J* = 18.5, 7.7 Hz, 2H), 3.73 (d, *J* = 8.0 Hz, 3H), 3.65 (t, *J* = 6.3 Hz, 1H), 3.39 (d, *J* = 5.6 Hz, 3H), 3.19 (dt, *J* = 10.9, 7.9 Hz, 1H), 2.01 (s, 3H), 1.95 (s, 3H), 1.48 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 171.3, 170.5, 169.5, 167.4, 164.5, 138.4, 137.3, 133.5, 129.7, 129.5, 128.9, 128.6, 128.3, 128.2, 127.8, 127.5, 101.9, 99.6, 79.0, 77.9, 75.0, 74.6, 74.2, 73.2, 72.3, 71.7, 71.1, 62.9, 56.7, 55.4, 52.7, 23.4, 20.7, 20.6. HRMS (ESI): *m/z*: calcd for $C_{41}H_{47}NNaO_{15}$ [M + Na]: 816.2838; found: 816.2846.



2-83h

Compound 2-83h: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.00 (d, *J* = 7.3 Hz, 2H), 7.60 (t, *J* = 7.4 Hz, 1H), 7.50 – 7.36 (m, 4H), 7.33 (dd, *J* = 11.0, 5.2 Hz, 4H), 7.26 (dd, *J* = 7.0, 5.2 Hz, 4H), 7.20 – 7.06 (m, 5H), 5.35 (dd, *J* = 16.2, 7.9 Hz, 2H), 5.27 (t, *J* = 9.6 Hz, 1H), 4.99 (d, *J* = 11.7 Hz, 1H), 4.82 (d, *J* = 8.3 Hz, 1H), 4.77 (d, *J* = 7.8 Hz, 1H), 4.71 (dd, *J* = 11.0, 2.6 Hz, 1H), 4.68 – 4.55 (m, 3H), 4.38 (dd, *J* =

36.7, 11.7 Hz, 2H), 4.10 (d, *J* = 2.1 Hz, 1H), 4.01 (d, *J* = 9.9 Hz, 1H), 3.92 (t, *J* = 9.2 Hz, 1H), 3.72 – 3.58 (m, 4H), 3.55 (dd, *J* = 9.4, 5.7 Hz, 1H), 3.50 – 3.45 (m, 1H), 3.39 (s, 3H), 3.19 (dt, *J* = 10.8, 7.8 Hz, 1H), 2.01 (s, 3H), 1.46 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 171.3, 169.5, 167.4, 164.5, 138.9, 138.1, 137.3, 133.5, 129.7, 129.6, 128.8, 128.6, 128.3, 128.1, 127.8, 127.8, 127.7, 127.6, 127.4, 101.9, 99.6, 79.0, 78.2, 75.6, 74.7, 74.2, 73.3, 73.3, 73.2, 72.3, 71.0, 69.0, 56.8, 55.6, 52.6, 23.3, 20.6. HRMS (ESI): *m/z*: calcd for $C_{46}H_{51}NNaO_{14}$ [M + Na]: 864.3202; found: 864.3206.

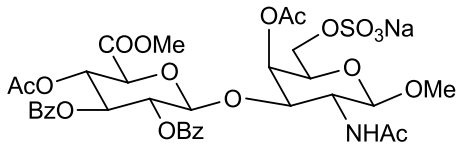
General procedure for the synthesis of compounds 2-84b – 2-84h

Precursor **2-83b** – **2-83h** (0.2 mmol) was dissolved in CH_2Cl_2/methanol (2 mL/2 mL) and 10% Pd on carbon (50 mg) was added. The reaction was placed under H_2 atmosphere and stirred for 16 h at room temperature. When TLC analysis showed completion of reaction, the mixture was filtered through Celite and **2-84b** – **2-84h** was afforded as a white solid which was immediately used for the next reaction to obtain analogues **2-85b** – **2-85h**.

General procedure for the synthesis of compounds 2-85b – 2-85h

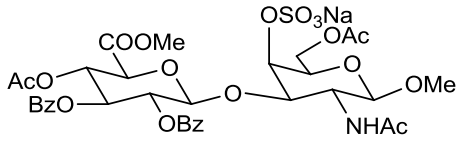
Precursor **2-84b** – **2-84h** (0.08 mmol) was dissolved in DMF (1 mL) and SO_3·TEA (5 equivalent per –OH present in precursor) was added. The reaction was stirred at 50°C for 36 h. When TLC analysis showed completion of reaction, the reaction was quenched with methanol (0.5 mL) and concentrated. The crude product was first purified using Sephadex LH-20 (CH_2Cl_2/methanol, 1:1) before

purification by flash column chromatography. Purified product was next passed through Sephadex C-25 Na^+ resin (methanol/H_2O, 1:1) to afford **2-85b** – **2-85h**. **2-85b** and **2-85c** required shorter reaction periods of 4 h and 12 h respectively.



2-85b

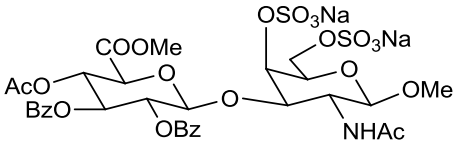
Compound 2-85b: ^{1}H NMR (500 MHz, DMSO) δ 7.78 (d, *J* = 8.1 Hz, 4H), 7.63 – 7.55 (m, 2H), 7.45 (dd, *J* = 16.3, 8.0 Hz, 4H), 5.77 – 5.68 (m, 1H), 5.31 – 5.22 (m, 3H), 4.58 – 4.49 (m, 2H), 4.29 (dd, *J* = 12.1, 2.9 Hz, 1H), 4.22 – 4.15 (m, 2H), 3.85 (m, 3H), 3.70 (s, 3H), 3.11 (s, 3H), 1.98 (s, 3H), 1.90 (s, 3H), 1.59 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 170.0, 169.1, 168.8, 167.4, 165.0, 164.5, 133.7, 133.2, 129.3, 129.2, 129.1, 128.7, 128.5, 128.3, 101.2, 99.5, 76.0, 73.0, 72.6, 72.2, 71.8, 71.1, 69.1, 64.6, 54.9, 52.6, 50.1, 22.6, 20.7, 20.2. HRMS (ESI): *m/z*: calcd for $C_{34}H_{38}NO_{19}S$ [M – Na]: 796.1764; found: 796.1760.



2-85c

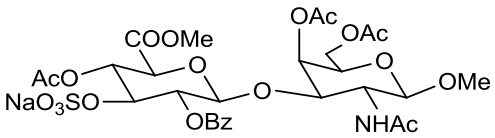
Compound 2-85c: ^{1}H NMR (500 MHz, DMSO) δ 7.78 (d, *J* = 6.9 Hz, 4H), 7.66 – 7.53 (m, 2H), 7.44 (dt, *J* = 15.7, 7.7 Hz, 4H), 5.73 (t, *J* = 8.9 Hz, 1H), 5.34 – 5.18 (m, 3H), 4.61 – 4.48 (m, 2H), 4.29 (dd, *J* = 12.0, 2.7 Hz, 1H), 4.23 – 4.09 (m, 2H), 3.85 (dd, *J* = 24.3, 7.7 Hz, 3H), 3.70 (s, 3H), 3.11 (s, 3H), 1.98 (s, 3H), 1.90

(s, 3H), 1.57 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 170.0, 169.1, 168.8, 167.5, 164.9, 164.5, 133.7, 133.2, 129.3, 129.2, 129.1, 128.7, 128.5, 128.3, 101.1, 99.6, 76.2, 73.0, 72.2, 71.7, 71.1, 69.1, 64.5, 54.9, 52.6, 50.1, 22.5, 20.7, 20.1. HRMS (ESI): *m/z*: calcd for $C_{34}H_{38}NO_{19}S$ [M – H]: 796.1764; found: 796.1771.



2-85d

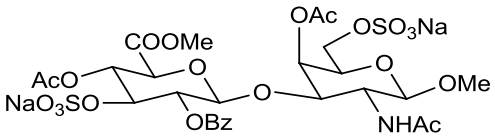
Compound 2-85d: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.81 – 7.71 (m, 4H), 7.58 (ddd, *J* = 20.9, 13.8, 8.1 Hz, 2H), 7.49 – 7.36 (m, 4H), 5.70 (t, *J* = 9.4 Hz, 1H), 5.36 – 5.24 (m, 2H), 5.20 (d, *J* = 7.8 Hz, 1H), 4.58 – 4.50 (m, 2H), 4.16 (d, *J* = 7.6 Hz, 1H), 4.09 (dd, *J* = 12.0, 2.0 Hz, 1H), 3.90 – 3.75 (m, 3H), 3.71 (s, 3H), 3.66 (d, *J* = 8.0 Hz, 1H), 3.25 (s, 3H), 1.90 (s, 3H), 1.44 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 169.1, 168.7, 167.8, 165.0, 164.5, 133.7, 133.2, 129.3, 129.2, 129.1, 128.7, 128.5, 128.3, 101.6, 100.6, 77.4, 74.1, 73.5, 73.2, 71.8, 71.0, 69.0, 66.9, 55.1, 52.8, 52.7, 22.5, 20.1. HRMS (ESI): *m/z*: calcd for $C_{32}H_{36}NO_{21}S_2$ [M – H]: 834.1227; found: 834.1221.



2-85e

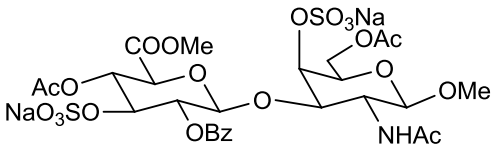
Compound 2-85e: ^{1}H NMR (500 MHz, DMSO) δ 7.95 (d, *J* = 7.3 Hz, 2H), 7.59 (t, *J* = 7.4 Hz, 1H), 7.46 (t, *J* = 7.6 Hz, 2H), 5.22 (d, *J* = 3.2 Hz, 1H), 4.99 (d, *J* =

8.1 Hz, 1H), 4.89 – 4.78 (m, 2H), 4.58 (t, *J* = 9.4 Hz, 1H), 4.29 – 4.19 (m, 2H), 4.05 (dd, *J* = 11.4, 3.4 Hz, 1H), 3.95 – 3.87 (m, 1H), 3.80 (ddd, *J* = 11.3, 9.6, 5.7 Hz, 2H), 3.71 (d, *J* = 9.5 Hz, 1H), 3.64 (s, 3H), 3.24 (s, 3H), 1.99 (d, *J* = 2.5 Hz, 6H), 1.92 (s, 3H), 1.18 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 170.1, 169.9, 169.2, 168.8, 167.5, 164.5, 132.7, 130.1, 130.0, 128.0, 101.6, 99.7, 77.4, 75.6, 71.8, 71.7, 70.8, 70.0, 68.4, 62.9, 55.7, 52.2, 22.2, 20.7, 20.6. HRMS (ESI): *m/z*: calcd for $C_{29}H_{36}NNa_2O_{19}S$ [M + Na]: 780.1392; found: 780.1397.



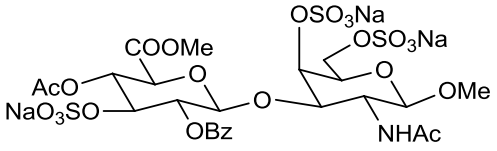
2-85f

Compound 2-85f: ^{1}H NMR (500 MHz, MeOD) δ 8.09 (d, *J* = 7.2 Hz, 2H), 7.54 (dd, *J* = 10.5, 4.1 Hz, 1H), 7.41 (t, *J* = 7.8 Hz, 2H), 5.30 (dd, *J* = 18.1, 9.0 Hz, 2H), 5.00 – 4.89 (m, 2H), 4.38 (dd, *J* = 11.8, 4.4 Hz, 1H), 4.35 – 4.20 (m, 3H), 3.96 (dd, *J* = 10.5, 8.2 Hz, 1H), 3.84 (d, *J* = 10.7 Hz, 1H), 3.78 (dd, *J* = 8.0, 5.1 Hz, 1H), 3.73 (s, 3H), 3.27 (dd, *J* = 3.1, 1.7 Hz, 3H), 2.03 (d, *J* = 4.7 Hz, 3H), 2.01 (s, 3H), 1.32 (s, 3H). ^{13}C NMR (500 MHz, MeOD) δ 173.7, 172.6, 171.8, 170.0, 167.8, 134.3, 131.8, 131.2, 129.2, 103.3, 103.1, 79.8, 78.9, 76.6, 73.7, 73.5, 73.3, 71.0, 65.3, 56.7, 53.4, 52.2, 22.4, 21.0, 20.8. HRMS (ESI): *m/z*: calcd for $C_{27}H_{33}NNa_3O_{21}S_2$ [M + Na]: 840.0674; found: 840.0681.



2-85g

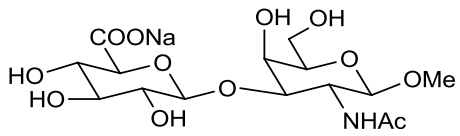
Compound 2-85g: ^{1}H NMR (500 MHz, MeOD) δ 8.09 (d, *J* = 8.1 Hz, 2H), 7.55 (t, *J* = 7.4 Hz, 1H), 7.41 (t, *J* = 7.6 Hz, 2H), 5.29 (t, *J* = 8.9 Hz, 2H), 4.99 – 4.92 (m, 2H), 4.88 – 4.81 (m, 1H), 4.38 (dd, *J* = 11.8, 4.1 Hz, 1H), 4.28 (ddd, *J* = 17.6, 12.6, 7.9 Hz, 3H), 3.95 (d, *J* = 8.2 Hz, 1H), 3.84 (d, *J* = 10.9 Hz, 1H), 3.78 (dd, *J* = 7.2, 4.5 Hz, 1H), 3.73 (s, 3H), 3.27 (s, 3H), 2.04 (s, 3H), 2.01 (s, 3H), 1.32 (s, 3H). ^{13}C NMR (500 MHz, MeOD) δ 173.7, 172.7, 171.9, 170.0, 167.8, 134.3, 131.8, 131.1, 129.2, 103.4, 103.1, 79.8, 78.9, 76.7, 73.7, 73.4, 73.3, 71.1, 65.2, 56.7, 53.5, 52.2, 22.4, 21.0, 20.8. HRMS (ESI): *m/z*: calcd for $C_{27}H_{33}NNa_3O_{21}S_2$ [M + Na]: 840.0674; found: 840.0682.



2-85h

Compound 2-85h: ^{1}H NMR (500 MHz, DMSO) δ 7.95 (d, *J* = 7.6 Hz, 2H), 7.56 (s, 1H), 7.43 (t, *J* = 7.5 Hz, 2H), 5.15 – 5.01 (m, 2H), 4.92 (d, *J* = 8.0 Hz, 1H), 4.59 (t, *J* = 9.2 Hz, 1H), 4.53 (s, 1H), 4.28 (d, *J* = 10.1 Hz, 1H), 4.11 – 4.02 (m, 2H), 3.88 – 3.80 (m, 1H), 3.78 – 3.67 (m, 5H), 3.62 (d, *J* = 8.3 Hz, 1H), 3.23 (s, 3H), 1.94 (s, 3H), 1.16 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 169.2, 168.8, 168.7, 164.9, 132.3, 130.6, 130.0, 127.7, 101.8, 101.7, 78.6, 76.1, 74.6, 73.4, 71.8,

71.6, 69.7, 66.8, 55.0, 52.6, 49.9, 22.1, 20.7. HRMS (ESI): *m/z*: calcd for $C_{25}H_{30}NNa_4O_{23}S_3$ [M + Na]: 899.9956; found: 899.9953.

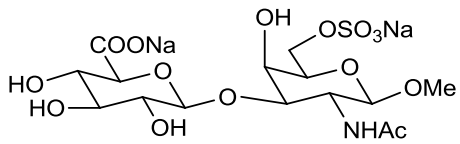


P1

Compound P1: Precursor **2-83a** (22.6 mg, 0.03 mmol) was dissolved in THF/H_2O (0.9 mL/0.3mL) and cooled to 0°C, before addition of H_2O_2 (151 µL) and LiOH (303 µL, 1 M). The reaction was stirred for 16 h at room temperature, after which, methanol (0.5 mL) was added and the reaction was cooled to 0°C before addition of NaOH (365 µL). The reaction was stirred at room temperature for 6 h. When the reaction was complete, it was quenched with Amberlyst-15 and stirred for another 15 min. At pH 3 – 4, the mixture was filtered and solvent was evaporated off. The product was purified using Sephadex G-10 (100% H_2O). Purified product was passed through Sephadex C-25 Na resin (100% H_2O) to afford **P1** (11.9 mg, 92%). ^{1}H NMR (500 MHz, D_2O) δ 4.43 (d, *J* = 7.9 Hz, 1H), 4.36 (d, *J* = 8.5 Hz, 1H), 4.11 (d, *J* = 3.0 Hz, 1H), 3.95 (dd, *J* = 10.8, 8.7 Hz, 1H), 3.78 – 3.70 (m, 3H), 3.62 (dd, *J* = 8.3, 4.1 Hz, 2H), 3.45 (s, 3H), 3.41 (dd, *J* = 5.7, 3.3 Hz, 2H), 3.29 – 3.22 (m, 1H), 1.95 (s, 3H). ^{13}C NMR (500 MHz, D_2O) δ 175.9, 174.9, 104.1, 102.2, 80.3, 76.2, 75.3, 74.9, 72.7, 71.8, 67.8, 61.1, 57.0, 51.0, 22.2. HRMS (ESI): *m/z*: calcd for $C_{15}H_{24}NNa_2O_{12}$ [M + Na]: 456.1088; found: 456.1083.

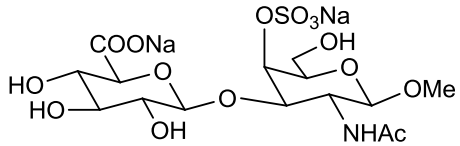
General procedure for the synthesis of compounds P2 – P8

Precursor **2-85b** – **2-85h** was treated as described for the preparation of **P1** to afford **P2** – **P8**.



P2

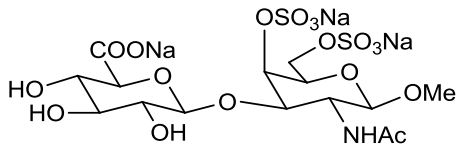
Compound P2: ^{1}H NMR (500 MHz, D_2O) δ 4.46 (d, *J* = 7.9 Hz, 1H), 4.40 (d, *J* = 8.6 Hz, 1H), 4.22 – 4.13 (m, 3H), 3.99 (dd, *J* = 10.8, 8.7 Hz, 1H), 3.89 (dd, *J* = 7.6, 4.4 Hz, 1H), 3.81 (dd, *J* = 10.9, 3.2 Hz, 1H), 3.68 – 3.63 (m, 1H), 3.47 (d, *J* = 3.7 Hz, 3H), 3.46 – 3.41 (m, 2H), 3.32 – 3.25 (m, 1H), 1.97 (s, 3H). ^{13}C NMR (500 MHz, D_2O) δ 175.95, 174.91, 104.09, 102.13, 80.08, 76.16, 75.40, 72.75, 72.68, 71.82, 67.70, 67.63, 57.10, 50.95, 22.27. HRMS (ESI): *m/z*: calcd for $C_{15}H_{23}NNa_3O_{15}S$ [M + Na]: 558.0476; found: 558.0467.



P3

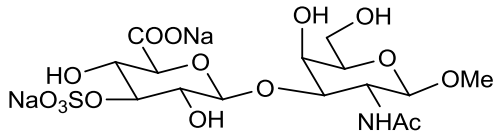
Compound P3: ^{1}H NMR (500 MHz, D_2O) δ 4.43 (dd, *J* = 7.7, 3.7 Hz, 2H), 4.05 – 3.96 (m, 2H), 3.83 – 3.71 (m, 3H), 3.62 (d, *J* = 9.6 Hz, 1H), 3.54 – 3.46 (m, 4H), 3.42 (t, *J* = 9.1 Hz, 1H), 3.36 – 3.29 (m, 1H), 1.98 (s, 3H). ^{13}C NMR (500 MHz, D_2O) δ 175.9, 174.9, 103.3, 102.1, 76.5, 76.4, 75.2, 74.9, 74.5, 72.5, 71.8, 61.1,

57.2, 51.7, 22.3. HRMS (ESI): *m/z*: calcd for $C_{15}H_{23}NNa_3O_{15}S$ [M + Na]: 558.0476; found: 558.0483.



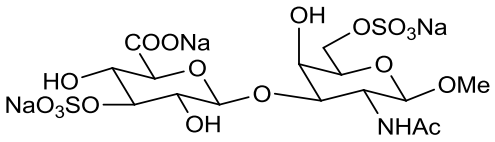
P4

Compound P4: ^{1}H NMR (500 MHz, D_2O) δ 4.45 (dd, *J* = 7.8, 1.5 Hz, 2H), 4.29 (dd, *J* = 11.4, 3.0 Hz, 1H), 4.19 (dd, *J* = 11.4, 8.7 Hz, 1H), 4.05 (qd, *J* = 10.6, 8.1 Hz, 3H), 3.64 (d, *J* = 9.8 Hz, 1H), 3.54 – 3.48 (m, 4H), 3.43 (d, *J* = 9.0 Hz, 1H), 3.36 – 3.30 (m, 1H), 2.00 (s, 3H). ^{13}C NMR (500 MHz, D_2O) δ 175.9, 174.9, 103.3, 102.0, 76.4, 76.3, 75.2, 74.7, 72.5, 72.4, 71.9, 68.0, 57.3, 51.6, 22.3. HRMS (ESI): *m/z*: calcd for $C_{15}H_{22}NNa_4O_{18}S_2$ [M + Na]: 659.9864; found: 659.9873.



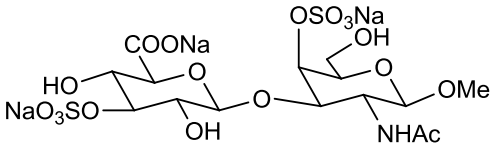
P5

Compound P5: ^{1}H NMR (500 MHz, D_2O) δ 4.52 (d, *J* = 7.9 Hz, 1H), 4.36 (d, *J* = 8.6 Hz, 1H), 4.21 (t, *J* = 9.0 Hz, 1H), 4.11 (d, *J* = 3.0 Hz, 1H), 3.95 (dd, *J* = 10.8, 8.6 Hz, 1H), 3.78 – 3.67 (m, 5H), 3.64 – 3.57 (m, 2H), 3.46 – 3.40 (m, 4H), 1.94 (s, 3H). ^{13}C NMR (500 MHz, D_2O) δ 175.5, 174.9, 103.7, 102.2, 83.9, 80.7, 76.1, 74.9, 71.6, 70.5, 67.6, 61.1, 57.0, 50.9, 22.3. HRMS (ESI): *m/z*: calcd for $C_{15}H_{23}NNa_3O_{15}S$ [M + Na]: 558.0476; found: 558.0471.



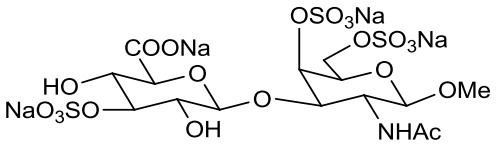
P6

Compound P6: ^{1}H NMR (500 MHz, D_2O) δ 4.51 (d, *J* = 7.9 Hz, 1H), 4.42 (d, *J* = 7.8 Hz, 1H), 4.22 (t, *J* = 9.0 Hz, 1H), 4.03 – 3.95 (m, 2H), 3.80 – 3.72 (m, 4H), 3.68 (t, *J* = 9.4 Hz, 1H), 3.50 – 3.41 (m, 4H), 1.96 (s, 3H). ^{13}C NMR (500 MHz, D_2O) δ 175.0, 174.6, 103.3, 102.0, 83.6, 76.4, 75.9, 75.7, 74.5, 71.3, 70.4, 61.0, 57.2, 51.6, 22.3. HRMS (ESI): *m/z*: calcd for $C_{15}H_{22}NNa_4O_{18}S_2$ [M + Na]: 659.9864; found: 659.9861.



P7

Compound P7: ^{1}H NMR (500 MHz, D_2O) δ 4.52 (d, *J* = 7.8 Hz, 1H), 4.44 (d, *J* = 7.9 Hz, 1H), 4.24 (t, *J* = 8.6 Hz, 1H), 4.06 – 3.95 (m, 2H), 3.81 – 3.74 (m, 3H), 3.69 (q, *J* = 10.1 Hz, 2H), 3.50 – 3.47 (m, 4H), 1.98 (s, 3H). ^{13}C NMR (500 MHz, D_2O) δ 175.2, 175.0, 103.2, 102.1, 83.7, 76.5, 76.3, 75.5, 74.5, 71.3, 70.5, 61.1, 57.2, 51.6, 22.3. HRMS (ESI): *m/z*: calcd for $C_{15}H_{22}NNa_4O_{18}S_2$ [M + Na]: 659.9864; found: 659.9872.

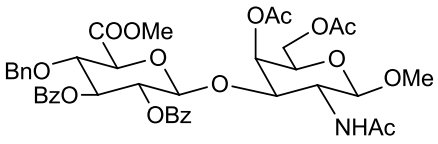


P8

Compound P8: ^{1}H NMR (500 MHz, D_2O) δ 4.50 (d, *J* = 7.9 Hz, 1H), 4.43 (d, *J* = 7.7 Hz, 1H), 4.29 – 4.20 (m, 2H), 4.18 – 4.11 (m, 1H), 4.06 – 3.96 (m, 3H), 3.67 (q, *J* = 10.1 Hz, 2H), 3.53 – 3.42 (m, 4H), 1.96 (s, 3H). ^{13}C NMR (500 MHz, D_2O) δ 175.4, 175.0, 103.1, 102.0, 83.8, 76.4, 76.2, 75.2, 72.4, 71.3, 70.5, 68.0, 57.2, 51.5, 22.3. HRMS (ESI): *m/z*: calcd for $C_{15}H_{21}NNa_5O_{21}S_3$ [M + Na]: 761.9251; found: 761.9258.

General procedure for the synthesis of compounds 2-86a – 2-86h

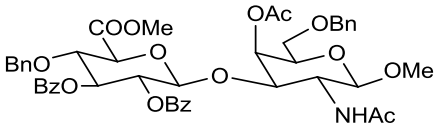
The protected dimer (0.25 mmol) was dissolved in toluene (8 mL) and Bu_3SnH (1.5 mmol) and ABCN (109 mg) was added. The reaction was heated to 100°C for 4 h. When TLC analysis showed completion of reaction, the solvent was evaporated off and the product was purified by flash column chromatography to afford of **2-86a** – **2-86h**.



2-86a

Compound 2-86a: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.90 (dd, *J* = 19.5, 7.9 Hz, 4H), 7.55 – 7.45 (m, 2H), 7.43 – 7.31 (m, 4H), 7.18 – 7.10 (m, 3H), 7.10 – 7.04 (m, 2H), 5.65 (t, *J* = 9.2 Hz, 1H), 5.41 (dd, *J* = 20.3, 5.1 Hz, 2H), 5.33 (dd, *J* = 9.3,

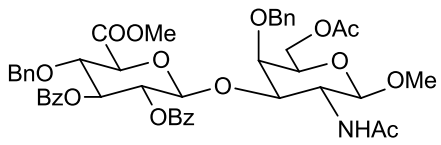
7.7 Hz, 1H), 4.95 (d, *J* = 8.2 Hz, 1H), 4.81 (d, *J* = 7.5 Hz, 1H), 4.80 – 4.72 (m, 1H), 4.53 (dd, *J* = 33.4, 11.1 Hz, 2H), 4.21 – 4.13 (m, 2H), 4.10 (d, *J* = 9.5 Hz, 1H), 4.04 (dd, *J* = 11.5, 7.4 Hz, 1H), 3.90 – 3.83 (m, 1H), 3.79 (s, 3H), 3.44 (s, 3H), 3.10 (dt, *J* = 10.5, 7.7 Hz, 1H), 2.10 (s, 3H), 2.05 (s, 3H), 1.39 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 171.1, 170.5, 169.9, 168.0, 165.4, 164.9, 137.0, 133.5, 133.3, 129.7, 129.1, 128.6, 128.4, 128.2, 128.0, 127.9, 101.2, 99.6, 75.0, 74.7, 74.3, 74.1, 71.9, 71.2, 68.9, 62.7, 57.0, 55.5, 52.7, 22.9, 20.7, 20.7. HRMS (ESI): *m/z*: calcd for $C_{41}H_{45}NNaO_{16}$ [M + Na]: 830.2631; found: 830.2635.



2-86b

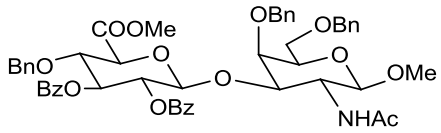
Compound 2-86b: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.02 – 7.83 (m, 4H), 7.51 (dd, *J* = 13.7, 7.1 Hz, 2H), 7.41 – 7.22 (m, 9H), 7.18 – 7.12 (m, 3H), 7.08 (dt, *J* = 5.9, 3.1 Hz, 2H), 5.65 (t, *J* = 9.2 Hz, 1H), 5.46 (d, *J* = 3.4 Hz, 1H), 5.40 – 5.21 (m, 2H), 4.95 (d, *J* = 8.3 Hz, 1H), 4.80 (d, *J* = 7.5 Hz, 1H), 4.73 (dd, *J* = 10.7, 3.4 Hz, 1H), 4.60 – 4.45 (m, 4H), 4.18 (t, *J* = 9.3 Hz, 1H), 4.09 (d, *J* = 9.6 Hz, 1H), 3.90 – 3.80 (m, 1H), 3.77 (d, *J* = 7.8 Hz, 3H), 3.55 (ddd, *J* = 17.0, 10.3, 5.9 Hz, 2H), 3.46 (d, *J* = 7.8 Hz, 3H), 3.11 (dt, *J* = 10.6, 7.8 Hz, 1H), 2.09 (s, 3H), 1.41 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 171.0, 170.1, 168.1, 165.4, 165.0, 138.0, 137.1, 133.5, 133.3, 129.7, 129.3, 128.6, 128.4, 128.3, 128.0, 127.9, 127.7, 127.7, 101.2, 99.8, 77.1, 75.2, 74.7, 74.3, 74.2, 73.6, 72.9, 72.0, 69.4, 69.3, 57.2, 55.7, 52.7,

23.0, 20.9. HRMS (ESI): *m/z*: calcd for $C_{46}H_{49}NNaO_{15}$ [M + Na]: 878.2994; found: 878.2998.



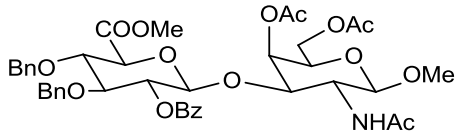
2-86c

Compound 2-86c: ^{1}H NMR (500 MHz, CDCl3) δ 7.93 (dd, J = 16.6, 7.8 Hz, 4H), 7.51 (dt, J = 11.0, 7.5 Hz, 2H), 7.37 (tdd, J = 25.2, 21.1, 7.2 Hz, 9H), 7.16 (dd, J = 4.9, 1.7 Hz, 3H), 7.12 – 7.08 (m, 2H), 5.81 – 5.69 (m, 1H), 5.44 (dd, J = 9.4, 8.2 Hz, 1H), 5.27 (d, J = 6.6 Hz, 1H), 4.98 (d, J = 11.6 Hz, 1H), 4.86 (dd, J = 14.3, 8.0 Hz, 2H), 4.77 (dd, J = 11.0, 2.6 Hz, 1H), 4.68 (d, J = 11.6 Hz, 1H), 4.57 (dd, J = 28.7, 11.1 Hz, 2H), 4.19 (d, J = 3.9 Hz, 2H), 4.15 – 4.10 (m, 1H), 3.94 (dd, J = 9.9, 4.3 Hz, 2H), 3.78 (s, 3H), 3.65 (t, J = 6.2 Hz, 1H), 3.39 (s, 3H), 3.18 (dt, J = 10.7, 7.7 Hz, 1H), 1.94 (s, 3H), 1.48 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 171.4, 170.5, 168.4, 165.5, 164.9, 138.4, 137.0, 133.5, 133.4, 129.8, 129.7, 129.4, 129.2, 128.6, 128.4, 128.3, 128.3, 128.0, 128.0, 127.7, 102.5, 99.6, 78.3, 77.5, 74.8, 74.7, 74.4, 74.3, 74.00, 72.1, 71.8, 63.1, 56.8, 55.7, 52.7, 23.3, 20.7. HRMS (ESI): *m/z*: calcd for $C_{46}H_{49}NNaO_{15}$ [M + Na]: 878.2994; found: 878.2996



2-86d

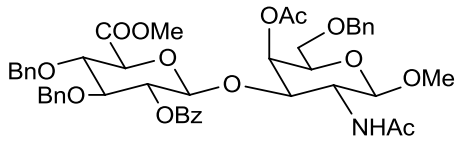
Compound 2-86d: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.04 – 7.85 (m, 4H), 7.59 – 7.45 (m, 2H), 7.43 – 7.24 (m, 14H), 7.20 – 7.14 (m, 3H), 7.15 – 7.07 (m, 2H), 5.74 (dt, *J* = 9.2, 4.4 Hz, 1H), 5.43 (dd, *J* = 9.7, 7.9 Hz, 1H), 5.26 (d, *J* = 6.7 Hz, 1H), 4.97 (d, *J* = 11.5 Hz, 1H), 4.87 (dd, *J* = 8.0, 2.3 Hz, 2H), 4.76 (dd, *J* = 10.9, 2.8 Hz, 1H), 4.67 – 4.49 (m, 3H), 4.39 (dd, *J* = 40.2, 11.7 Hz, 2H), 4.17 (t, *J* = 5.2 Hz, 2H), 4.02 (d, *J* = 2.5 Hz, 1H), 3.77 – 3.62 (m, 4H), 3.55 (dd, *J* = 9.5, 5.7 Hz, 1H), 3.46 (dd, *J* = 9.5, 6.5 Hz, 1H), 3.40 (s, 3H), 3.21 – 3.14 (m, 1H), 1.47 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 171.3, 168.3, 165.4, 164.9, 138.8, 138.1, 137.0, 133.4, 133.3, 129.7, 129.6, 129.4, 129.1, 128.9, 128.6, 128.4, 128.3, 128.3, 128.1, 128.0, 127.9, 127.7, 127.6, 127.5, 102.4, 99.6, 78.4, 77.4, 75.3, 74.7, 74.5, 74.2, 74.00, 73.3, 73.2, 72.1, 69.0, 56.8, 55.8, 52.6, 23.3. HRMS (ESI): *m/z*: calcd for $C_{51}H_{53}NNaO_{14}$ [M + Na]: 926.3358; found: 926.3355.



2-86e

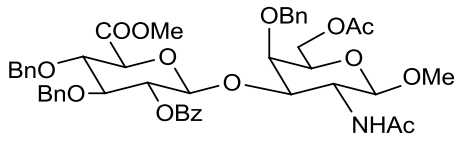
Compound 2-86e: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.03 – 7.93 (m, 2H), 7.60 (t, *J* = 7.5 Hz, 1H), 7.46 (t, *J* = 7.8 Hz, 2H), 7.35 – 7.21 (m, 5H), 7.17 – 7.06 (m, 5H), 5.58 (d, *J* = 6.8 Hz, 1H), 5.40 (d, *J* = 3.4 Hz, 1H), 5.26 (dd, *J* = 8.7, 7.7 Hz, 1H), 4.97 (d, *J* = 8.3 Hz, 1H), 4.73 (ddd, *J* = 14.7, 12.3, 7.3 Hz, 4H), 4.64 (dd, *J* = 11.1,

3.8 Hz, 2H), 4.19 (dd, *J* = 11.6, 4.8 Hz, 1H), 4.06 – 3.95 (m, 3H), 3.86 (dd, *J* = 7.2, 5.1 Hz, 1H), 3.82 – 3.74 (m, 4H), 3.45 (s, 3H), 3.15 (dt, *J* = 10.6, 7.9 Hz, 1H), 2.07 (s, 3H), 2.06 (s, 3H), 1.36 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 171.0, 170.5, 170.0, 168.3, 164.8, 137.5, 137.4, 133.4, 129.7, 129.4, 128.5, 128.3, 128.2, 127.9, 127.9, 127.6, 101.0, 99.6, 81.4, 79.0, 74.9, 74.8, 74.2, 73.2, 71.2, 68.6, 62.8, 56.9, 55.3, 52.5, 22.8, 20.7, 20.6. HRMS (ESI): *m/z*: calcd for $C_{41}H_{47}NNaO_{15}$ [M + Na]: 816.2838; found: 816.2835.



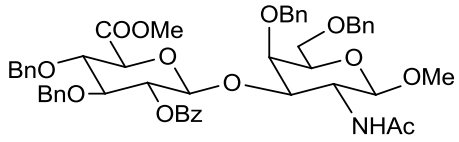
2-86f

Compound 2-86f: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.60 (dd, *J* = 7.4, 1.9 Hz, 2H), 7.41 – 7.21 (m, 16H), 7.19 – 7.09 (m, 2H), 6.10 (d, *J* = 4.8 Hz, 1H), 5.59 (d, *J* = 3.0 Hz, 1H), 4.95 (d, *J* = 7.9 Hz, 1H), 4.68 (dd, *J* = 23.7, 10.1 Hz, 2H), 4.52 (tdd, *J* = 14.4, 10.2, 4.5 Hz, 5H), 4.42 (d, *J* = 11.5 Hz, 1H), 4.19 (dd, *J* = 10.9, 3.2 Hz, 1H), 3.93 (d, *J* = 8.4 Hz, 1H), 3.85 – 3.72 (m, 3H), 3.67 (s, 3H), 3.58 – 3.47 (m, 3H), 3.45 (d, *J* = 4.8 Hz, 3H), 2.05 (s, 3H), 1.78 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.6, 170.3, 169.7, 137.7, 137.6, 137.4, 137.2, 129.4, 128.4, 128.4, 128.2, 128.0, 127.9, 127.9, 127.8, 127.7, 127.7, 126.1, 121.8, 101.3, 98.0, 76.6, 76.0, 73.6, 72.9, 72.4, 72.2, 71.1, 69.8, 68.4, 56.9, 53.3, 52.4, 23.8, 20.9. HRMS (ESI): *m/z*: calcd for $C_{46}H_{51}NNaO_{14}$ [M + Na]: 864.3202; found: 864.3205.



2-86g

Compound 2-86g: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.99 (d, *J* = 7.5 Hz, 2H), 7.60 (t, *J* = 7.5 Hz, 1H), 7.48 – 7.38 (m, 4H), 7.38 – 7.21 (m, 8H), 7.19 – 7.06 (m, 5H), 5.36 (dd, *J* = 9.0, 8.2 Hz, 1H), 5.30 (d, *J* = 6.6 Hz, 1H), 4.96 (d, *J* = 11.6 Hz, 1H), 4.86 – 4.60 (m, 8H), 4.14 (dd, *J* = 11.2, 6.8 Hz, 1H), 4.07 – 3.98 (m, 2H), 3.94 (dd, *J* = 11.1, 6.0 Hz, 2H), 3.87 (d, *J* = 8.6 Hz, 1H), 3.75 (s, 3H), 3.64 (t, *J* = 6.3 Hz, 1H), 3.40 (s, 3H), 3.18 (dd, *J* = 7.1, 3.0 Hz, 1H), 1.93 (d, *J* = 6.5 Hz, 3H), 1.50 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 171.2, 170.4, 168.6, 164.7, 138.4, 137.6, 137.5, 133.4, 129.6, 129.0, 128.6, 128.4, 128.3, 128.2, 127.9, 127.9, 127.7, 127.6, 102.4, 99.5, 81.6, 79.5, 77.9, 75.1, 75.1, 74.6, 74.3, 74.1, 73.6, 71.7, 63.0, 56.7, 55.6, 52.5, 23.4, 20.7. HRMS (ESI): *m/z*: calcd for $C_{46}H_{51}NNaO_{14}$ [M + Na]: 864.3202; found: 864.3207.



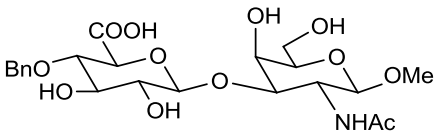
2-86h

Compound 2-86h: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.00 (d, *J* = 7.2 Hz, 2H), 7.59 (t, *J* = 7.3 Hz, 1H), 7.53 – 7.19 (m, 17H), 7.21 – 7.08 (m, 5H), 5.36 (dd, *J* = 14.8, 6.6 Hz, 2H), 4.97 (d, *J* = 11.7 Hz, 1H), 4.86 (d, *J* = 8.3 Hz, 1H), 4.83 – 4.74 (m, 3H), 4.74 – 4.58 (m, 4H), 4.38 (dd, *J* = 41.6, 11.7 Hz, 2H), 4.15 – 3.95 (m, 3H), 3.88 (t, *J* = 8.9 Hz, 1H), 3.74 – 3.60 (m, 4H), 3.55 (dd, *J* = 9.5, 5.8 Hz, 1H), 3.46

(dd, *J* = 9.6, 6.6 Hz, 1H), 3.40 (d, *J* = 4.6 Hz, 3H), 3.19 (dt, *J* = 10.8, 7.8 Hz, 1H), 1.48 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 171.1, 168.5, 164.7, 138.8, 138.0, 137.5, 137.5, 133.5, 129.6, 128.8, 128.5, 128.4, 128.3, 128. 2, 128.0, 127.9, 127.84, 127.6, 127.5, 127.3, 102.38, 99.5, 81.6, 79.4, 78.1, 75.3, 75.0, 75.0, 74.4, 74.0, 73.6, 73.2, 73.2, 69.0, 56.7, 55.6, 52.3, 23.3. HRMS (ESI): *m/z*: calcd for $C_{51}H_{55}NNaO_{13}$ [M + Na]: 912.3566; found: 912.3572.

General procedure for the synthesis of compounds 2-87a – 2-87h

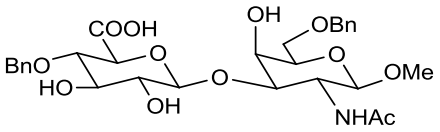
Precursor **2-86a** – **2-86h** (0.2 mmol) was dissolved in THF (3.5 mL) and cooled to 0°C, before addition of H_2O_2 (0.4 mL) and LiOH (0.8 mL, 1 M). The reaction was stirred for 16 h at room temperature, after which, methanol (2.26 mL) was added and the reaction was cooled to 0°C before addition of NaOH (0.9 mL). The reaction was stirred at room temperature for 6 h. When TLC analysis showed completion of reaction, the reaction was quenched with Amberlyst-15 and stirred for another 15 min. At pH 3 – 4, the mixture was filtered and solvent was evaporated off. The product was purified by flash column chromatography to afford **2-87a** – **2-87h**.



2-87a

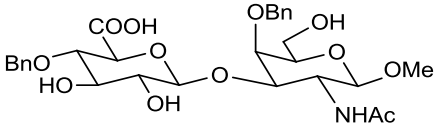
Compound 2-87a: ^{1}H NMR (500 MHz, DMSO) δ 7.69 (d, *J* = 8.9 Hz, 1H), 7.36 – 7.22 (m, 5H), 4.72 (d, *J* = 11.2 Hz, 1H), 4.57 (d, *J* = 11.2 Hz, 1H), 4.34 (d, *J* =

7.6 Hz, 1H), 4.25 (d, *J* = 8.4 Hz, 1H), 3.90 – 3.78 (m, 2H), 3.62 – 3.52 (m, 3H), 3.52 – 3.39 (m, 2H), 3.33 (dd, *J* = 9.9, 4.0 Hz, 2H), 3.31 (d, *J* = 6.1 Hz, 3H), 3.14 (dd, *J* = 16.2, 8.1 Hz, 1H), 1.78 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 171.7, 170.0, 139.1, 127.9, 127.6, 127.6, 127.1, 104.5, 101.8, 80.5, 80.5, 76.3, 75.5, 75.0, 73.3, 73.2, 66.7, 60.1, 55.3, 50.6, 23.1. HRMS (ESI): *m/z*: calcd for $C_{22}H_{31}NNaO_{12}$ [M + Na]: 524.1738; found: 524.1745.



2-87b

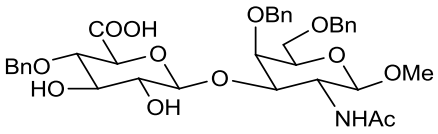
Compound 2-87b: ^{1}H NMR (500 MHz, DMSO) δ 7.75 (d, *J* = 8.7 Hz, 1H), 7.38 – 7.15 (m, 10H), 4.65 (s, 2H), 4.49 (d, *J* = 13.5 Hz, 3H), 4.34 (d, *J* = 7.3 Hz, 1H), 4.28 (d, *J* = 8.4 Hz, 1H), 3.89 – 3.78 (m, 2H), 3.67 – 3.43 (m, 6H), 3.30 (s, 3H), 3.14 (t, *J* = 7.1 Hz, 1H), 1.79 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 173.7, 170.1, 139.5, 138.5, 128.2, 128.2, 127.8, 127.7, 127.7, 127.5, 127.4, 126.9, 104.3, 101.6, 80.9, 79.6, 75.4, 73.3, 73.1, 72.7, 72.3, 69.5, 67.6, 56.0, 55.4, 50.6, 23.1. HRMS (ESI): *m/z*: calcd for $C_{29}H_{36}NO_{12}$ [M – H]: 590.2243; found: 590.2248.



2-87c

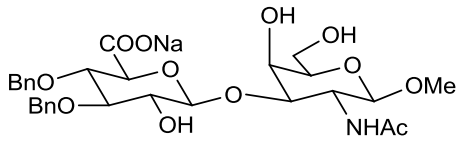
Compound 2-87c: ^{1}H NMR (500 MHz, DMSO) δ 7.71 (d, *J* = 7.8 Hz, 1H), 7.42 – 7.12 (m, 10H), 4.86 (d, *J* = 11.5 Hz, 1H), 4.79 (d, *J* = 11.2 Hz, 1H), 4.58 (d, *J* =

11.2 Hz, 1H), 4.52 (d, *J* = 11.5 Hz, 1H), 4.38 (d, *J* = 7.7 Hz, 1H), 4.28 (d, *J* = 4.2 Hz, 1H), 3.86 (s, 1H), 3.81 (d, *J* = 7.4 Hz, 2H), 3.70 (d, *J* = 9.5 Hz, 2H), 3.49 – 3.34 (m, 4H), 3.29 (s, 3H), 3.11 (t, *J* = 8.3 Hz, 1H), 1.79 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 170.9, 170.1, 139.2, 139.0, 128.1, 128.0, 127.6, 127.3, 127.2, 105.0, 101.7, 80.5, 80.0, 75.8, 75.4, 75.3, 74.7, 73.9, 73.6, 60.1, 55.5, 51.4, 23.2. HRMS (ESI): *m/z*: calcd for $C_{29}H_{37}NNaO_{12}$ [M – H]: 614.2208; found: 614.2201.



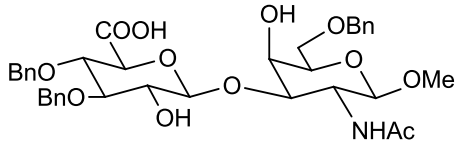
2-87d

Compound 2-87d: ^{1}H NMR (500 MHz, DMSO) δ 7.72 (d, *J* = 7.1 Hz, 1H), 7.37 – 7.21 (m, 14H), 4.82 (dd, *J* = 23.6, 11.4 Hz, 2H), 4.60 (d, *J* = 11.2 Hz, 1H), 4.47 (d, *J* = 11.6 Hz, 1H), 4.42 – 4.33 (m, 3H), 4.30 (d, *J* = 7.2 Hz, 1H), 3.87 – 3.76 (m, 3H), 3.71 (d, *J* = 9.5 Hz, 1H), 3.61 (t, *J* = 5.9 Hz, 1H), 3.51 (t, *J* = 9.3 Hz, 1H), 3.45 – 3.34 (m, 3H), 3.33 – 3.30 (m, 1H), 3.28 (s, 3H), 3.12 (t, *J* = 8.3 Hz, 1H), 1.80 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 170.8, 170.0, 139.0, 138.9, 138.2, 128.3, 128.2, 128.1, 127.9, 127.6, 127.4, 127.3, 127.1, 104.9, 101.6, 80.3, 80.1, 75.8, 75.3, 73.9, 73.7, 73.5, 72.6, 72.2, 68.8, 60.00, 55.5, 51.3, 23.1. HRMS (ESI): *m/z*: calcd for $C_{36}H_{42}NO_{12}$ [M – H]: 680.2712; found: 680.2716.



2-87e

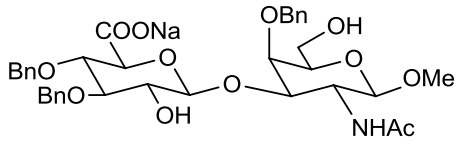
Compound 2-87e: ^{1}H NMR (500 MHz, DMSO) δ 7.75 (d, *J* = 8.7 Hz, 1H), 7.34 – 7.21 (m, 9H), 5.45 (s, 1H), 5.13 (s, 1H), 4.81 (d, *J* = 11.7 Hz, 1H), 4.69 (dd, *J* = 11.1, 6.3 Hz, 2H), 4.61 (s, 1H), 4.53 (d, *J* = 10.9 Hz, 1H), 4.38 (d, *J* = 6.3 Hz, 1H), 4.28 (d, *J* = 8.4 Hz, 1H), 3.96 (s, 1H), 3.81 (d, *J* = 9.7 Hz, 1H), 3.74 (t, *J* = 7.4 Hz, 1H), 3.60 (dd, *J* = 18.8, 10.8 Hz, 3H), 3.46 – 3.42 (m, 2H), 3.30 (s, 3H), 1.80 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 172.9, 170.0, 139.2, 138.9, 127.9, 127.9, 127.4, 127.1, 127.1, 104.5, 101.8, 83.7, 80.7, 80.4, 77.6, 74.9, 73.5, 73.0, 72.8, 66.5, 59.9, 55.3, 50.7, 23.1. HRMS (ESI): *m/z*: calcd for $C_{29}H_{36}NO_{12}$ [M – Na]: 590.2243; found: 590.2238.



2-87f

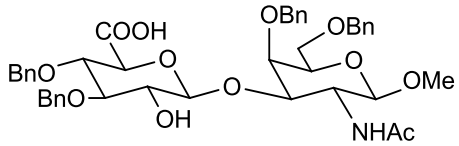
Compound 2-87f: ^{1}H NMR (500 MHz, DMSO) δ 7.69 (d, *J* = 8.9 Hz, 1H), 7.38 – 7.18 (m, 14H), 4.87 (d, *J* = 11.5 Hz, 1H), 4.82 (s, 1H), 4.68 (dd, *J* = 18.3, 11.2 Hz, 2H), 4.58 (d, *J* = 10.9 Hz, 1H), 4.51 (s, 2H), 4.46 (d, *J* = 7.6 Hz, 1H), 4.30 (d, *J* = 8.5 Hz, 1H), 3.86 (dd, *J* = 20.1, 10.9 Hz, 2H), 3.77 (d, *J* = 9.2 Hz, 1H), 3.69 – 3.58 (m, 4H), 3.54 (dd, *J* = 11.3, 8.4 Hz, 1H), 3.48 – 3.43 (m, 1H), 3.40 – 3.33 (m, 2H), 3.32 (s, 3H), 1.82 (s, 3H). ^{13}C NMR (500 MHz, $CDCl_3$) δ 170.5, 169.9, 138.9, 138.5, 138.3, 128.2, 128.0, 127.6, 127.6, 127.4, 127.4, 127.3, 127.2, 104.2, 101.6,

83.7, 79.7, 79.3, 73.7, 73.6, 73.3, 73.2, 72.2, 69.4, 67.6, 59.9, 55.3, 50.5, 23.1. HRMS (ESI): *m/z*: calcd for $C_{36}H_{43}NNaO_{12}$ [M + Na]: 704.2677; found: 704.2678.



2-87g

Compound 2-87g: ^{1}H NMR (500 MHz, DMSO) δ 7.71 (d, *J* = 7.5 Hz, 1H), 7.41 – 7.19 (m, 14H), 4.93 – 4.82 (m, 2H), 4.71 (d, *J* = 11.5 Hz, 1H), 4.68 – 4.56 (m, 2H), 4.54 (d, *J* = 11.5 Hz, 1H), 4.48 (d, *J* = 7.5 Hz, 1H), 4.31 (d, *J* = 7.4 Hz, 1H), 3.90 (s, 1H), 3.83 (dd, *J* = 12.5, 8.8 Hz, 3H), 3.64 (t, *J* = 9.3 Hz, 1H), 3.53 – 3.46 (m, 2H), 3.44 – 3.32 (m, 3H), 3.31 (s, 3H), 1.83 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 170.5, 170.0, 139.2, 139.0, 138.4, 128.0, 127.9, 127.7, 127.5, 127.4, 127.2, 104.7, 101.7, 83.8, 80.5, 79.2, 75.3, 74.7, 73.9, 73.9, 73.8, 73.8, 60.0, 55.5, 51.3, 23.2. HRMS (ESI): *m/z*: calcd for $C_{36}H_{43}NNaO_{12}$ [M + Na]: 704.2677; found: 704.2670.



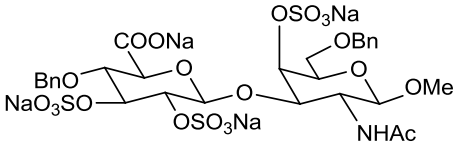
2-87h

Compound 2-87h: ^{1}H NMR (500 MHz, DMSO) δ 7.74 (d, *J* = 7.3 Hz, 1H), 7.43 – 7.19 (m, 19H), 4.90 (dd, *J* = 21.2, 11.5 Hz, 2H), 4.77 – 4.67 (m, 2H), 4.61 (d, *J* = 10.8 Hz, 1H), 4.52 (dd, *J* = 16.7, 9.6 Hz, 2H), 4.38 (dt, *J* = 17.4, 8.7 Hz, 3H),

3.97 – 3.79 (m, 4H), 3.66 (dt, *J* = 12.2, 7.6 Hz, 2H), 3.55 (t, *J* = 8.9 Hz, 1H), 3.45 (dd, *J* = 9.6, 6.3 Hz, 1H), 3.40 – 3.33 (m, 2H), 3.31 (s, 3H), 1.85 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 170.4, 170.1, 139.0, 138.9, 138.3, 138.2, 128.3, 128.1, 128.1, 127.7, 127.6, 127.5, 127.4, 127.3, 127.3, 104.8, 101.7, 83.8, 80.6, 79.1, 75.4, 74.4, 73.9, 73.9, 73.8, 72.6, 72.3, 68.8, 55.5, 51.3, 48.6, 23.2. HRMS (ESI): *m/z*: calcd for $C_{43}H_{49}NNaO_{12}$ [M + Na]: 794.3147; found: 794.3140.

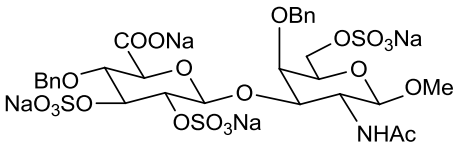
General procedure for the synthesis of compounds 2-88b – 2-88h

Precursor **2-87b – 2-87h** was treated as described for the preparation of **2-85b – 2-85h** with reaction time of 36 h to afford **2-88b – 2-88h** respectively.



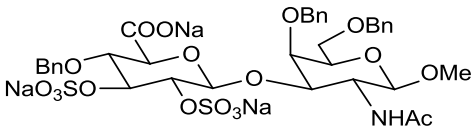
2-88b

Compound 2-88b: ^{1}H NMR (500 MHz, DMSO) δ 7.69 (d, *J* = 7.9 Hz, 1H), 7.41 (d, *J* = 7.2 Hz, 2H), 7.36 – 7.19 (m, 8H), 4.86 (s, 2H), 4.74 (d, *J* = 11.2 Hz, 1H), 4.59 (s, 1H), 4.53 (d, *J* = 4.7 Hz, 1H), 4.49 (s, 2H), 4.44 (d, *J* = 11.2 Hz, 1H), 4.38 (d, *J* = 8.0 Hz, 1H), 4.11 (s, 1H), 3.92 (d, *J* = 3.4 Hz, 1H), 3.81 (d, *J* = 10.6 Hz, 1H), 3.77 – 3.61 (m, 3H), 3.58 (dd, *J* = 10.5, 7.8 Hz, 1H), 3.32 (s, 3H), 1.89 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 173.5, 170.9, 138.8, 138.6, 128.2, 127.9, 127.9, 127.4, 127.3, 127.2, 101.8, 101.6, 79.0, 77.4, 76.7, 75.9, 73.4, 73.2, 72.1, 70.5, 70.3, 55.5, 51.1, 23.1. HRMS (ESI): *m/z*: calcd for $C_{29}H_{33}NNa_5O_{21}S_3$ [M + Na]: 942.0190; found: 942.0183.



2-88c

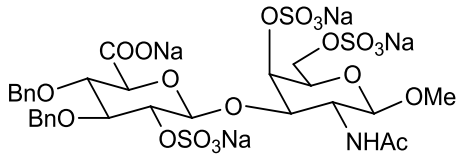
Compound 2-88c: ^{1}H NMR (500 MHz, DMSO) δ 7.77 (d, *J* = 8.6 Hz, 1H), 7.46 (d, *J* = 7.4 Hz, 2H), 7.34 – 7.25 (m, 4H), 7.24 – 7.14 (m, 4H), 5.16 (d, *J* = 11.5 Hz, 1H), 5.05 (d, *J* = 2.2 Hz, 1H), 4.82 (d, *J* = 11.1 Hz, 1H), 4.63 – 4.51 (m, 2H), 4.47 (d, *J* = 11.2 Hz, 2H), 4.35 (d, *J* = 8.3 Hz, 1H), 4.30 (dd, *J* = 8.9, 5.9 Hz, 1H), 4.17 (s, 1H), 4.04 – 3.97 (m, 1H), 3.93 (dd, *J* = 10.4, 5.7 Hz, 1H), 3.87 (dd, *J* = 10.4, 6.3 Hz, 1H), 3.78 (d, *J* = 8.9 Hz, 1H), 3.73 – 3.67 (m, 1H), 3.63 (t, *J* = 6.0 Hz, 1H), 3.28 (s, 3H), 1.83 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 172.9, 170.2, 139.8, 139.1, 127.9, 127.9, 127.9, 127.8, 127.0, 126.9, 101.8, 100.1, 78.3, 77.2, 76.7, 76.3, 76.1, 74.1, 73.0, 71.6, 65.2, 55.6, 51.5, 23.1. HRMS (ESI): *m/z*: calcd for $C_{29}H_{34}NNa_4O_{21}S_3$ [M + H]: 920.0371; found: 920.0375.



2-88d

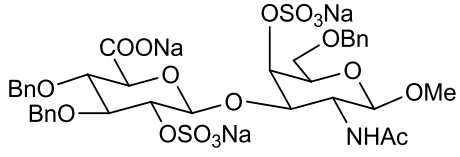
Compound 2-88d: ^{1}H NMR (500 MHz, DMSO) δ 7.78 (d, *J* = 8.3 Hz, 1H), 7.44 – 7.14 (m, 15H), 5.11 (d, *J* = 11.8 Hz, 1H), 5.00 (d, *J* = 2.9 Hz, 1H), 4.79 (d, *J* = 11.1 Hz, 1H), 4.61 – 4.53 (m, 2H), 4.52 – 4.42 (m, 4H), 4.34 (d, *J* = 8.3 Hz, 1H), 4.29 – 4.22 (m, 1H), 4.14 (s, 1H), 4.03 (d, *J* = 11.5 Hz, 1H), 3.78 – 3.70 (m, 2H), 3.63 (t, *J* = 5.5 Hz, 1H), 3.54 (t, *J* = 8.4 Hz, 1H), 3.48 (dd, *J* = 10.0, 4.4 Hz, 1H), 3.27 (s, 3H), 1.83 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 173.2, 170.1, 139.7,

139.3, 138.3, 128.2, 128.0, 127.8, 127.8, 127.7, 127.6, 127.4, 126.8, 102.0, 100.4, 77.8, 77.6, 76.8, 76.7, 76.5, 76.0, 74.0, 73.6, 72.2, 71.4, 69.6, 55.4, 51.4, 23.1. HRMS (ESI): *m/z*: calcd for $C_{36}H_{41}NNa_3O_{18}S_2$ [M + H]: 908.1453; found: 908.1450.



2-88e

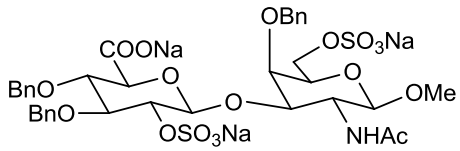
Compound 2-88e: ^{1}H NMR (500 MHz, DMSO) δ 7.91 (d, *J* = 7.4 Hz, 1H), 7.35 (d, *J* = 6.8 Hz, 2H), 7.28 – 7.16 (m, 8H), 4.91 (d, *J* = 11.4 Hz, 1H), 4.74 (d, *J* = 8.2 Hz, 1H), 4.55 (dd, *J* = 18.5, 9.2 Hz, 3H), 4.48 (s, 2H), 4.44 (t, *J* = 6.2 Hz, 1H), 4.04 (d, *J* = 9.9 Hz, 1H), 3.88 (d, *J* = 10.8 Hz, 1H), 3.81 – 3.66 (m, 5H), 3.33 (s, 3H), 1.90 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 172.6, 171.1, 138.9, 138.7, 128.0, 127.8, 127.7, 127.5, 127.1, 103.2, 100.5, 79.9, 78.6, 78.3, 78.2, 77.7, 74.8, 72.9, 72.2, 71.7, 66.6, 55.5, 52.3, 23.1. HRMS (ESI): *m/z*: calcd for $C_{29}H_{33}NNa_3O_{21}S_3$ [M + Na]: 896.0406; found: 896.0402.



2-88f

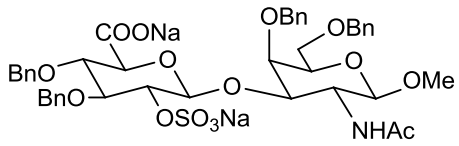
Compound 2-88f: ^{1}H NMR (500 MHz, DMSO) δ 7.95 (d, *J* = 7.2 Hz, 1H), 7.37 – 7.20 (m, 14H), 4.91 (d, *J* = 11.4 Hz, 1H), 4.79 (d, *J* = 8.2 Hz, 1H), 4.63 – 4.37

(m, 9H), 3.95 (d, *J* = 10.9 Hz, 1H), 3.74 (d, *J* = 11.4 Hz, 3H), 3.68 (d, *J* = 7.9 Hz, 1H), 3.61 (t, *J* = 8.5 Hz, 2H), 3.38 (s, 3H), 1.89 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 172.1, 170.9, 138.9, 138.8, 138.8, 128.1, 127.9, 127.8, 127.8, 127.6, 127.3, 127.0, 102.8, 100.7, 80.2, 78.6, 77.8, 74.2, 73.3, 72.2, 72.0, 71.9, 70.3, 55.5, 52.8, 48.6, 23.2. HRMS (ESI): *m/z*: calcd for $C_{36}H_{40}NNa_2O_{18}S$ [M - Na]: 884.1488; found: 884.1480.



2-88g

Compound 2-88g: ^{1}H NMR (500 MHz, DMSO) δ 7.93 (d, *J* = 8.0 Hz, 1H), 7.35 (dd, *J* = 6.5, 2.7 Hz, 2H), 7.32 – 7.28 (m, 2H), 7.24 – 7.17 (m, 10H), 5.09 (d, *J* = 11.1 Hz, 1H), 4.93 (d, *J* = 11.3 Hz, 1H), 4.77 (d, *J* = 4.9 Hz, 1H), 4.66 (d, *J* = 11.3 Hz, 1H), 4.57 – 4.51 (m, 3H), 4.48 (d, *J* = 11.1 Hz, 1H), 4.29 (t, *J* = 5.0 Hz, 1H), 4.18 (s, 1H), 4.04 – 3.93 (m, 2H), 3.91 (dd, *J* = 10.3, 6.5 Hz, 1H), 3.82 (dd, *J* = 10.3, 5.9 Hz, 1H), 3.70 – 3.57 (m, 4H), 3.29 (s, 3H), 1.87 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 173.0, 170.2, 139.5, 139.3, 139.1, 128.0, 127. 9, 127.8, 127.8, 127.5, 127.0, 126.9, 101.6, 100.9, 82.9, 79.0, 77.8, 77.4, 76.7, 75.4, 73.9, 72.9, 72.7, 71.9, 64.7, 55.6, 52.1, 23.1. HRMS (ESI): *m/z*: calcd for $C_{36}H_{40}NNa_2O_{18}S_2$ [M – Na]: 884.1488; found: 884.1489.

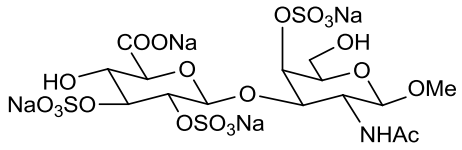


2-88h

Compound 2-88h: ^{1}H NMR (500 MHz, DMSO) δ 7.94 (d, *J* = 8.0 Hz, 1H), 7.47 – 7.17 (m, 19H), 4.99 (d, *J* = 11.3 Hz, 1H), 4.95 – 4.80 (m, 2H), 4.61 – 4.51 (m, 3H), 4.44 (dt, *J* = 18.0, 11.7 Hz, 4H), 4.37 – 4.29 (m, 1H), 4.11 – 3.94 (m, 3H), 3.89 – 3.74 (m, 3H), 3.62 (s, 1H), 3.43 – 3.58 (m, 2H), 3.27 (s, 3H), 1.87 (s, 3H). ^{13}C NMR (500 MHz, DMSO) δ 172.2, 170.3, 139.2, 138.8, 138.8, 138.3, 128.2, 128.2, 128.0, 127.9, 127.6, 127.5, 127.4, 127.1, 127.0, 101.7, 100.5, 82.4, 78.5, 77.7, 77.4, 75.3, 73.9, 73.1, 73.0, 72.3, 71.8, 69.1, 55.6, 51.8, 23.1. HRMS (ESI): *m/z*: calcd for $C_{43}H_{48}NNa_2O_{15}S$ [M + Na]: 896.2535; found: 896.2539.

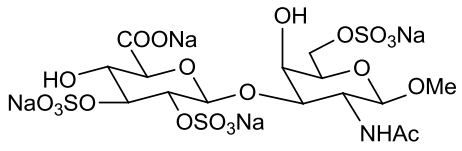
General procedure for the synthesis of compounds P10 – P16

Precursor **2-88b** – **2-88h** (0.03 mmol) was dissolved in H_2O/MeOH (1 mL/1 mL) and 10% Pd on carbon (100 mg) was added. The reaction was placed under H_2 atmosphere and stirred for 16 h at room temperature. When the reaction was complete, the mixture was filtered through Celite and purified using Sephadex G-10 (100% H_2O). Purified product was passed through Sephadex C-25 Na resin (100% H_2O) to afford **P10** – **P16**.



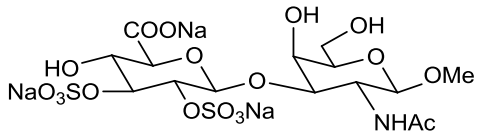
P10

Compound P10: ^{1}H NMR (500 MHz, D_2O) δ 4.84 (d, *J* = 2.7 Hz, 1H), 4.38 (dd, *J* = 13.3, 5.5 Hz, 1H), 4.27 (t, *J* = 7.2 Hz, 1H), 4.20 – 4.12 (m, 1H), 3.92 (t, *J* = 9.0 Hz, 1H), 3.85 – 3.73 (m, 5H), 3.69 (dd, *J* = 11.1, 8.8 Hz, 1H), 3.46 (s, 3H), 2.00 (s, 3H). ^{13}C NMR (500 MHz, D_2O) δ 175.4, 175.0, 101.9, 101.8, 81.7, 77.4, 76.5, 76.4, 76.1, 74.7, 70.3, 61.1, 57.1, 52.6, 22.6. HRMS (ESI): *m/z*: calcd for $C_{15}H_{21}NNa_5O_{21}S_3$ [M + Na]: 761.9251; found: 761.9256.



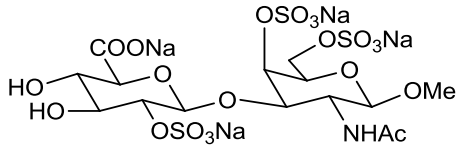
P11

Compound P11: ^{1}H NMR (500 MHz, D_2O) δ 4.84 (d, *J* = 6.4 Hz, 1H), 4.45 (d, *J* = 8.4 Hz, 1H), 4.39 – 4.34 (m, 1H), 4.25 (t, *J* = 6.8 Hz, 1H), 4.21 – 4.14 (m, 3H), 3.93 – 3.76 (m, 5H), 3.44 (s, 3H), 1.98 (s, 3H). ^{13}C NMR (500 MHz, D_2O) δ 175.6, 175.0, 102.4, 100.7, 81.2, 79.6, 77.2, 76.1, 72.6, 70.4, 67.8, 67.1, 57.1, 51.1, 22.5. HRMS (ESI): *m/z*: calcd for $C_{15}H_{21}NNa_5O_{21}S_3$ [M + Na]: 761.9251; found: 761.9246.



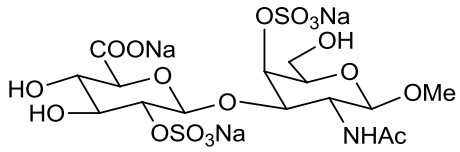
P12

Compound P12: ^{1}H NMR (500 MHz, D_2O) δ 4.83 (d, *J* = 6.5 Hz, 1H), 4.44 (d, *J* = 8.4 Hz, 1H), 4.39 – 4.32 (m, 1H), 4.24 (t, *J* = 6.8 Hz, 1H), 4.14 (d, *J* = 2.8 Hz, 1H), 3.81 (dddd, *J* = 24.4, 9.7, 7.6, 2.6 Hz, 6H), 3.63 (dd, *J* = 7.4, 4.7 Hz, 1H), 3.43 (s, 3H), 1.98 (s, 3H). ^{13}C NMR (500 MHz, D_2O) δ 175.6, 174.9, 102.4, 100.7, 81.1, 79.7, 77.2, 76.1, 74.8, 70.4, 67.2, 61.1, 57.0, 51.2, 22.5. HRMS (ESI): *m/z*: calcd for $C_{15}H_{22}NNa_2O_{18}S_2$ [M – Na]: 614.0079; found: 614.0077.



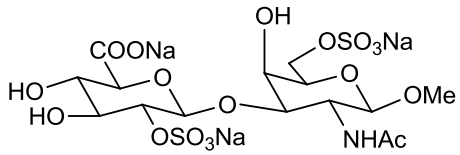
P13

Compound P13: ^{1}H NMR (500 MHz, D_2O) δ 4.87 (d, *J* = 2.6 Hz, 1H), 4.62 (d, *J* = 8.0 Hz, 2H), 4.28 (dd, *J* = 11.4, 2.7 Hz, 1H), 4.20 – 4.09 (m, 3H), 4.03 (dd, *J* = 8.9, 2.7 Hz, 1H), 3.80 (dd, *J* = 10.8, 8.9 Hz, 1H), 3.70 – 3.57 (m, 3H), 3.49 (s, 3H), 2.00 (s, 3H). ^{13}C NMR (500 MHz, D_2O) δ 175.7, 174.9, 101.9, 101.8, 80.1, 76.3, 76.2, 76.1, 75.1, 72.5, 71.4, 68.1, 57.2, 52.3, 22.4. HRMS (ESI): *m/z*: calcd for $C_{15}H_{21}NNa_5O_{21}S_3$ [M + Na]: 761.9251; found: 761.9257.



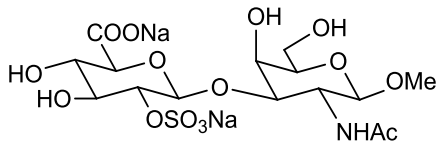
P14

Compound P14: ^{1}H NMR (500 MHz, D_2O) δ 4.83 (d, *J* = 2.8 Hz, 1H), 4.62 (dd, *J* = 8.0, 4.1 Hz, 2H), 4.13 (dt, *J* = 15.1, 5.2 Hz, 2H), 3.76 (dt, *J* = 8.2, 5.8 Hz, 4H), 3.72 – 3.61 (m, 4H), 3.47 (s, 3H), 2.00 (s, 3H). ^{13}C NMR (500 MHz, D_2O) δ 175.7, 174.9, 101.9, 101.8, 80.2, 76.4, 76.3, 76.2, 75.1, 74.7, 71.3, 61.1, 57.1, 52.5, 22.4. HRMS (ESI): *m/z*: calcd for $C_{15}H_{22}NNa_2O_{18}S_2$ [M – Na]: 614.0079; found: 614.0076.



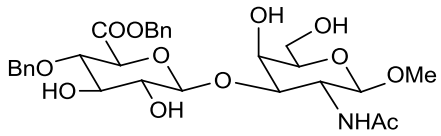
P15

Compound P15: ^{1}H NMR (500 MHz, D_2O) δ 4.48 (d, *J* = 8.2 Hz, 1H), 4.25 – 4.14 (m, 3H), 4.09 – 4.03 (m, 1H), 3.91 (tdd, *J* = 16.7, 11.2, 5.6 Hz, 3H), 3.73 – 3.66 (m, 3H), 3.56 (t, *J* = 9.5 Hz, 1H), 3.47 (s, 3H), 2.01 (s, 3H). ^{13}C NMR (500 MHz, D_2O) δ 175.7, 174.9, 102.4, 101.2, 79.9, 79.7, 76.0, 74.6, 72.6, 71.5, 67.8, 67.2, 57.1, 51.1, 22.5. HRMS (ESI): *m/z*: calcd for $C_{15}H_{22}NNa_2O_{18}S_2$ [M – Na]: 614.0079; found: 614.0085.



P16

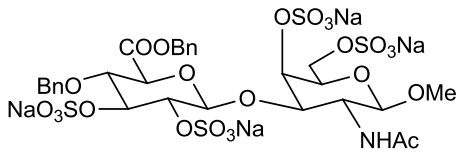
Compound P16: ^{1}H NMR (500 MHz, D_2O) δ 4.44 (d, *J* = 8.3 Hz, 1H), 4.14 (d, *J* = 2.7 Hz, 1H), 4.03 (dd, *J* = 8.8, 7.8 Hz, 1H), 3.87 (ddd, *J* = 19.2, 11.0, 5.7 Hz, 2H), 3.76 – 3.70 (m, 2H), 3.70 – 3.65 (m, 2H), 3.63 (dd, *J* = 7.3, 4.6 Hz, 1H), 3.53 (t, *J* = 9.5 Hz, 1H), 3.44 (s, 3H), 1.98 (s, 3H). ^{13}C NMR (500 MHz, D_2O) δ 175.7, 174.9, 102.4, 101.2, 79.9, 79.8, 76.0, 74.8, 74.6, 71.5, 67.4, 61.1, 57.0, 51.2, 22.5. HRMS (ESI): *m/z*: calcd for $C_{15}H_{23}NNa_3O_{15}S$ [M + Na]: 558.0476; found: 558.0483.



2-87a'

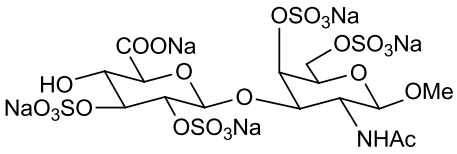
Compound 2-87a': To a solution of **2-87a** (37 mg, 0.071 mmol) in DMF (0.5 mL) was added BnBr (12.7μL, 0.107 mmol) and $NaHCO_3$ (30 mg, 0.355 mmol) and the reaction stirred at 50°C for 2 h. When the reaction was complete, the solvent was evaporated off. The product was purified by Sephadex LH-20 (CH_2Cl_2/methanol, 1:1) to afford **2-87a'** (37.7 mg, 90%). ^{1}H NMR (500 MHz, MeOD) δ 7.42 – 7.16 (m, 10H), 5.15 (d, *J* = 2.1 Hz, 1H), 4.60 (s, 1H), 4.46 (dd, *J* = 21.7, 9.4 Hz, 2H), 4.36 (d, *J* = 8.3 Hz, 1H), 4.03 (dd, *J* = 10.7, 8.6 Hz, 1H), 3.99 – 3.93 (m, 2H), 3.82 – 3.67 (m, 3H), 3.64 – 3.53 (m, 3H), 3.51 – 3.48 (m, 1H), 3.47 (s, 3H), 3.28 – 3.21 (m, 2H), 1.94 (s, 3H). ^{13}C NMR (500 MHz, MeOD) δ

174.4, 170.4, 139.6, 136.8, 129.6, 129.5, 129.2, 129.0, 128.6, 106.4, 103.6, 82.0, 80.8, 77.4, 76.4, 75.7, 74.7, 69.7, 68.3, 62.5, 59.5, 56.9, 52.8, 23.2. HRMS (ESI): *m/z*: calcd for $C_{29}H_{37}NNaO_{12}$ [M + Na]: 614.2208; found: 614.2214.



2-88a

Compound 2-88a: Precursor **2-87a'** (37.7 mg, 0.064 mmol) was treated as described for the preparation of **2-85b** – **2-85h** with a longer reaction period of 48 h to afford **2-88a** (43.5 mg, 68%). ^{1}H NMR (500 MHz, MeOD) δ 7.45 – 7.27 (m, 7H), 7.23 (dd, *J* = 8.1, 4.2 Hz, 3H), 5.22 (dd, *J* = 35.0, 12.1 Hz, 2H), 5.12 (t, *J* = 4.9 Hz, 2H), 4.95 (s, 1H), 4.70 (d, *J* = 5.6 Hz, 1H), 4.62 – 4.56 (m, 2H), 4.54 – 4.46 (m, 2H), 4.39 (dd, *J* = 11.7, 3.7 Hz, 1H), 4.25 (dd, *J* = 11.6, 7.8 Hz, 1H), 3.97 (d, *J* = 3.3 Hz, 2H), 3.93 (dd, *J* = 7.6, 3.7 Hz, 1H), 3.46 (s, 3H), 2.01 (s, 3H). ^{13}C NMR (500 MHz, MeOD) δ 174.9, 172.2, 139.0, 136.6, 129.7, 129.6, 129.5, 129.2, 128.7, 103.0, 102.3, 79.4, 78.5, 77.9, 76.6, 76.5, 76.5, 74.1, 73.4, 69.0, 68.9, 57.1, 53.1, 23.3. HRMS (ESI): *m/z*: calcd for $C_{29}H_{33}NNa_3O_{24}S_4$ [M – Na]: 975.9974; found: 975.9978.



P9

Compound P9: Compound **2-88a** (43.5 mg, 0.044 mmol) was treated as described for the preparation of **P10** – **P16** to afford **P9** (33 mg, 90%). ^{1}H NMR (500 MHz, D_2O) δ 4.88 (d, *J* = 2.6 Hz, 1H), 4.66 (d, *J* = 8.6 Hz, 1H), 4.38 (t, *J* = 7.9 Hz, 1H), 4.30 – 4.24 (m, 2H), 4.20 – 4.11 (m, 2H), 4.02 (dd, *J* = 8.8, 2.5 Hz, 1H), 3.92 (t, *J* = 8.9 Hz, 1H), 3.75 (dt, *J* = 11.0, 6.1 Hz, 2H), 3.47 (s, 3H), 2.01 (s, 3H). ^{13}C NMR (500 MHz, D_2O) δ 175.4, 175.0, 101.9, 101.7, 81.6, 77.3, 76.5, 76.4, 75.9, 72.5, 70.4, 68.1, 57.2, 52.4, 22.5. HRMS (ESI): *m/z*: calcd for $C_{15}H_{20}NNa_4O_{24}S_4$ [M – Na]: 817.8854; found: 817.8857.

2.16 References

1. Gama, C.I. *et al.* Sulfation patterns of glycosaminoglycans encode molecular recognition and activity. *Nat. Chem. Biol.* **2**, 467-473 (2006).
2. Tully, S.E. *et al.* A chondroitin sulfate small molecule that stimulates neuronal growth. *J. Am. Chem. Soc.* **126**, 7736-7737 (2004).
3. Vibert, A., Jacquinet, J.-C. & Lopin-Bon, C. Recent advances in the chemical and enzymatic chondroitin sulfate synthesis. *J. Carbohydr. Chem.* **30**, 393-414 (2011).
4. Jacquinet, J.-C., Lopin-Bon, C. & Vibert, A. From polymer to size-defined oligomers: a highly divergent and stereocontrolled construction of chondroitin sulfate A, C, D, E, K, L, and M oligomers from a single precursor: part 2. *Chem. - Eur. J.* **15**, 9579-9595 (2009).
5. Yeung, B.K.S., Chong, P.Y.C. & Petillo, P.A. Synthesis of glycosaminoglycans. *J. Carbohydr. Chem.* **21**, 799-865 (2002).
6. Karst, N.A. & Linhardt, R.J. Recent chemical and enzymatic approaches to the synthesis of glycosaminoglycan oligosaccharides. *Curr. Med. Chem.* **10**, 1993-2031 (2003).
7. Vibert, A., Lopin-Bon, C. & Jacquinet, J.-C. From polymer to size-defined oligomers: a step economy process for the efficient and stereocontrolled construction of chondroitin oligosaccharides and biotinylated conjugates thereof: part 1. *Chem. - Eur. J.* **15**, 9561-9578 (2009).

8. Tamura, J. *et al.* Synthesis and interaction with midkine of biotinylated chondroitin sulfate tetrasaccharides. *Bioorg. Med. Chem. Lett.* **22**, 1371-1374 (2012).

9. Tamura, J.-i., Neumann, K.W., Kurono, S. & Ogawa, T. Synthetic approach towards sulfated chondroitin di-, tri- and tetrasaccharides corresponding to the repeating unit. *Carbohydr. Res.* **305**, 43-63 (1997).

10. Lubineau, A. & Bonnaffé, D. Access to molecular diversity in glycosaminoglycans: combinatorial synthesis of eight chondroitin sulfate disaccharides. *Eur. J. Org. Chem.* **1999**, 2523-2532 (1999).

11. Jacquinet, J.C., Rochepeau-Jobron, L. & Combal, J.P. Multigram syntheses of the disaccharide repeating units of chondroitin 4- and 6-sulfates. *Carbohydr. Res.* **314**, 283-288 (1998).

12. Lopin, C. & Jacquinet, J.C. From polymer to size-defined oligomers: an expeditious route for the preparation of chondroitin oligosaccharides. *Angew. Chem. Int. Ed.* **45**, 2574-2578 (2006).

13. Slaghek, T.M., Hyppönen, T.K., Ogawa, T., Kamerling, J.P. & Vliegenthart, J.F.G. Synthesis of a tetrasaccharide fragment of hyaluronic acid having a glucuronic acid at the reducing end. *Tetrahedron Lett.* **34**, 7939-7942 (1993).

14. Yeung, B.K.S., Hill, D.C., Janicka, M. & Petillo, P.A. Synthesis of Two Hyaluronan Trisaccharides. *Org. Lett.* **2**, 1279-1282 (2000).

15. Coutant, C. & Jacquinet, J.-C. 2-Deoxy-2-trichloroacetamido-D-glucopyranose derivatives in oligosaccharide synthesis: from hyaluronic

acid to chondroitin 4-sulfate trisaccharides. *J. Chem. Soc., Perkin Trans. 1*, 1573-1581 (1995).

16. J, S. *et al.* Synthesis and Characterization of Chemically Modified Hyaluronan and Chondroitin Sulfate. *Mini-Rev. Org. Chem.* **7**, 290-299 (2010).

17. Karst, N. & Jacquinet, J.-C. Stereocontrolled total syntheses of shark cartilage chondroitin sulfate D-related tetra- and hexasaccharide methyl glycosides. *Eur. J. Org. Chem.* **2002**, 815-825 (2002).

18. Schmidt, R.R. & Michel, J. Facile Synthesis of α- and β-O-Glycosyl Imidates; Preparation of Glycosides and Disaccharides. *Angew. Chem. Int. Ed.* **19**, 731-732 (1980).

19. Jacquinet, J.-C. Syntheses of the methyl glycosides of the repeating units of chondroitin 4- and 6-sulfate. *Carbohydr. Res.* **199**, 153-181 (1990).

20. Abbas, S.A. & Haines, A.H. Benzoyl cyanide as a selective acylating agent. *Carbohydr. Res.* **39**, 358-363 (1975).

21. Ye, X.-S. & Wong, C.-H. Anomeric Reactivity-Based One-Pot Oligosaccharide Synthesis: A Rapid Route to Oligosaccharide Libraries. *J. Org.Chem.* **65**, 2410-2431 (2000).

22. Hada, N., Sonoda, Y. & Takeda, T. Synthesis of a novel glycosphingolipid from the millipede, Parafontaria laminata armigera, and the assembly of its carbohydrate moiety into multivalent structures. *Carbohydr. Res.* **341**, 1341-1352 (2006).

23. Gold, H. *et al.* A Concise Synthesis of Globotriaosylsphingosine. *Eur. J. Org. Chem.* **2011**, 1652-1663 (2011).

24. Codée, J.D.C. *et al.* A Modular Strategy Toward the Synthesis of Heparin-like Oligosaccharides Using Monomeric Building Blocks in a Sequential Glycosylation Strategy. *J. Am. Chem. Soc.* **127**, 3767-3773 (2005).

25. Blatter, G. & Jacquinet, J.C. The use of 2-deoxy-2-trichloroacetamido-D-glucopyranose derivatives in syntheses of hyaluronic acid-related tetra-, hexa-, and octa-saccharides having a methyl beta-D-glucopyranosiduronic acid at the reducing end. *Carbohydr. Res.* **288**, 109-125 (1996).

26. Dinkelaar, J., Witte, M.D., van den Bos, L.J., Overkleeft, H.S. & van der Marel, G.A. NIS/TFA: a general method for hydrolyzing thioglycosides. *Carbohydr. Res.* **341**, 1723-1729 (2006).

27. Pilgrim, W. & Murphy, P.V. α-glycosphingolipids via chelation-induced anomerization of O- and S-glucuronic and galacturonic acid derivatives. *Org. Lett.* **11**, 939-942 (2009).

28. Banaag, A.R. & Tius, M.A. Design of chiral auxiliaries for the allene ether nazarov cyclization. *J. Am. Chem. Soc.* **129**, 5328-5329 (2007).

29. Lemieux, R.U. & Ratcliffe, R.M. The azidonitration of tri-O-acetyl-D-galactal. *Can. J. Chem.* **57**, 1244-1251 (1979).

30. Belot, F. & Jacquinet, J.C. Syntheses of chondroitin 4- and 6-sulfate pentasaccharide derivatives having a methyl beta-D-glucopyranosiduronic acid at the reducing end. *Carbohydr. Res.* **326**, 88-97 (2000).

31. Karst, N. & Jacquinet, J.-C. Chemical synthesis of the disaccharide repeating unit of shark cartilage chondroitin sulfate D, and of its methyl D-glycoside derivative. *J. Chem. Soc., Perkin Trans. 1*, 2709-2717 (2000).

32. Xia, J. *et al.* Use of 1,2-dichloro 4,5-dicyanoquinone (DDQ) for cleavage of the 2-naphthylmethyl (NAP) group. *Tetrahedron Lett.* **41**, 169-173 (2000).

33. Lucas, H. *et al.* Syntheses of heparin - like pentamers containing "opened" uronic acid moieties. *Tetrahedron* **46**, 8207-8228 (1990).

34. Chao, C.-S., Li, C.-W., Chen, M.-C., Chang, S.-S. & Mong, K.-K.T. Low-Concentration 1,2-trans β-Selective Glycosylation Strategy and Its Applications in Oligosaccharide Synthesis. *Chem. - Eur. J.* **15**, 10972-10982 (2009).

35. Vankar, Y.D., Vankar, P.S., Behrendt, M. & Schmidt, R.R. Synthesis of β-O-glycosides using enol ether and imidate derived leaving groups. Emphasis on the use of nitriles as a solvent. *Tetrahedron* **47**, 9985-9992 (1991).

36. Chao, C.-S., Lin, C.-Y., Mulani, S., Hung, W.-C. & Mong, K.-k.T. Neighboring-Group Participation by C-2 Ether Functions in Glycosylations Directed by Nitrile Solvents. *Chem. - Eur. J.* **17**, 12193-12202 (2011).

37. Schmidt, R.R., Behrendt, M. & Toepfer, A. Nitriles as Solvents in Glycosylation Reactions: Highly Selective *β*-Glycoside Synthesis1. *Synlett.* **1990**, 694-696 (1990).

38. Ohlin, M., Johnsson, R. & Ellervik, U. Regioselective reductive openings of 4,6-benzylidene acetals: synthetic and mechanistic aspects. *Carbohydr. Res.* **346**, 1358-1370 (2011).

39. Tani, S., Sawadi, S., Kojima, M., Akai, S. & Sato, K.-i. A novel method for regioselective ring-opening reduction of 4,6-O-benzylidene hexopyranoside derivatives using CoCl2 and BH3·THF. *Tetrahedron Lett.* **48**, 3103-3104 (2007).

40. Rao, K.V., Patil, P.R., Atmakuri, S. & Kartha, K.P.R. Iodine–sodium cyanoborohydride-mediated reductive ring opening of 4,6-O-benzylidene acetals of hexopyranosides. *Carbohydr. Res.* **345**, 2709-2713 (2010).

41. Sakagami, M. & Hamana, H. A selective ring opening reaction of 4,6-O-benzylidene acetals in carbohydrates using trialkylsilane derivatives. *Tetrahedron Lett.* **41**, 5547-5551 (2000).

CHAPTER 3: BIOLOGICAL EVALUATION OF CHONDROITIN SULFATE IN BREAST CANCER

3.1 Introduction to Breast Cancer

Breast cancer is the top diagnosed cancer for women worldwide, accounting for 29% of the new cancer cases, and is ranked second for the number of female cancer-related deaths after lung cancer in 2013.[1] It is estimated that 1 in 8 women will develop breast cancer in her lifetime in the United States.[2] In Singapore, breast cancer is the most common cancer in the female population and has the highest mortality rate across all ethnic groups, accounting for up to 18.1% of cancer deaths from 2009-2013.[3]

Breast cancer arises from the uncontrolled division of cells in the breast tissue, which can either occur in the lobules (glands involved in milk production), or the ducts (which connect the lobules to the nipple). In the *in situ* stage of breast cancer, the abnormal cells are localized to form lobular carcinoma *in situ* (LCIS) or the ductal carcinoma *in situ* (DCIS).[1] This could progress to the invasive forms of breast cancer, which is classified into various stages, depending on the primary tumor size (T), extend of tumor spread to nearby lymph nodes (N), and whether distant metastasis (M) to other organs has occurred. The TNM staging system[4] was established by the Union for International Cancer Control (UICC) to categorize the degree of cancer spread (Table 3-1), and is currently recognized as the global standard.

Table 3-1. TNM staging system

Stage	Characteristics
0	*in situ* carcinoma: cancer cells localized at original site
I	Tumor size ≤ 2 cm; negative lymph node involvement
II	Tumor size between 2-5cm; metastasis to nearby axillary lymph nodes
	Tumor size ≥ 5 cm; negative lymph node involvement
IIIA	Tumor ≥ 5 cm; metastasis to nearby axillary lymph nodes
IIIB	Tumor of any size; spread to chest wall, breast skin and axillary lymph nodes
IV	Tumor of any size; metastasis to distant organs

The common risk factors of breast cancer include an increase in age, family history of breast cancer, use of menopausal hormone therapy, physical inactivity and alcohol intake.[5] Detection of breast cancer in early stages via mammography is often recommended to improve survival outcomes, reducing the risk of death by more than 30%.[6, 7] In addition, early detection provides more treatment options such as less extensive surgery, radiotherapy and the use of milder chemotherapeutic drugs with less severe side effects.[1] For instance, the patient may opt for lumpectomy instead of mastectomy to conserve the breast tissue if the detected tumor is small and has yet to spread. To reduce the risk of a relapse, a combination of chemotherapeutic drugs is commonly prescribed, which is more effective than single drug treatment.[8] Examples include taxotere with

cyclophosphamide (TC chemotherapy), adriamycin with cyclophosphamide (AC chemotherapy) and epirubicin with cyclophosphamide (EC chemotherapy).[9]

Breast cancer is a complicated disease with a heterogeneous molecular profile. Thus, this may lead to unpredictable clinical responses when treated with the same therapeutic approach, making it difficult to predict survival outcome.[10-12] Many unique molecular subtypes of breast cancer have been identified through gene expression profiling, and these subtypes may be characterized based on the tumor grade, histological type, lymph node involvement and expression of distinct biological markers such as estrogen receptors (ER), progesterone receptors (PR) and human epidermal growth factor receptor 2 (HER2).[13, 14] The five molecular subtypes of breast cancer have been classified as luminal A, luminal B, basal, claudin-low and HER2.[14]

For each molecular subtype, the therapy indicated primarily depends on the therapeutic targets available.[15] For example, the expression of ER enables treatment via hormone/endocrine therapy. This approach targets breast cancer cells which require estrogen to multiply, through the use of estrogen receptor antagonists such as tamoxifen.[16] This drug binds competitively to the estrogen receptor but does not activate it, thereby arresting the development of the cancer cells. In addition, tumor subtypes which over-express HER2 can be targeted using trasuszumab, a monoclonal antibody which binds to the HER2 receptor; this drug blocks off the HER2 pathway responsible for the uncontrolled proliferation of the

cancer cells.[17] However, some breast cancer subtypes (such as basal tumors) do not express any known therapeutic targets such as ER, PR, and HER2, and are otherwise known as triple negative breast cancer (TNBC).[18] These tumor subtypes are more challenging to treat as they are endocrine non-responsive and show low response to chemotherapy. TNBC subtypes are biologically more aggressive, and usually results in higher patient mortality. Thus, new treatment options are required to diagnose and treat TBNC molecular subtypes, so as to improve survival outcomes.

3.2 Chondroitin Sulfate in Breast Cancer

Glycosaminoglycans (GAGs) have gained interest as potential therapeutic agents in cancer treatment, with studies showing their involvement in various pathobiological cancer stages, including the modification of cellular characteristics such as proliferation, invasion, migration and angiogenesis.[19, 20] GAGs have also been shown to interact with various effective molecules such as growth factors, cytokines and chemokines, thereby regulating their biological effects.[21]

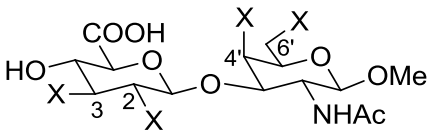
Recently, a correlation was established between chondroitin sulfate proteoglycan (CSPG) expression and various cancer phenotypes, with the overexpression of CSPGs versican and decorin in prostate, testicular, gastric, pancreatic and breast cancer.[22-26] For instance, compositional analysis of proteoglycan side chains obtained from neoplastic breast cancer tissues indicate an elevation in CS

expression[27, 28], and these sequences have been shown to regulate many important cellular signaling processes affecting cell proliferation, adhesion and migration.[29] The increased expression of specific CS sulfation patterns such as CS-E on the surface of breast cancer cells has also been associated with their metastatic spread. These CS GAGs were reported to act as P-selectin ligands[30-32], which can cause blood borne metastasis.[33, 34] Specifically, there is a down regulation of *CHST3* gene which codes for biosynthesis of CS-C and CS-D sulfation sequences and an increase in expression of *CHST11* which encodes for biosynthesis of CS-A and CS-E.[35-37]

The dysregulation of CS expression has the potential to serve as prognostic and diagnostic markers for targeted cancer therapy.[38] In addition, structural changes to the CS proteoglycans of malignant breast tissues indicate that sulfate groups present on CS might play an important role in the cellular processes involved in the progression of breast cancer.[22, 39-41]

With our complete library of CS disaccharides (Table 3-2), we proceeded to evaluate their biological effects in breast cancer by systematic structural activity relationship studies to investigate the CS "sulfation code". The aim is to determine if position specific sulfate groups encode important functional information for the cellular processes involved in breast cancer.

Table 3-2. CS disaccharide library



X = hydroxyl or sulfate groups

CS analogue		Position of sulfation	Conventional name
Non-sulfated	**P1**		CS-O
Mono-sulfated	**P16**	C-2	
	P5	C-3	
	P3	C-4’	CS-A
	P2	C-6’	CS-C
Di-sulfated	**P12**	C-2, C-3	CS-R
	P14	C-2, C-4’	
	P15	C-2, C-6’	CS-D
	P7	C-3, C-4’	CS-K
	P6	C-3, C-6’	CS-L
	P4	C4’, -C6’	CS-E
Tri-sulfated	**P11**	C-2, C-3, C-4’	
	P10	C-2, C-3, C-6’	
	P13	C-2, C-4’, C-6’	CS-T
	P8	C-3, C-4’, C-6’	CS-M
Tetrasulfated	**P9**	C-2, C-3, C-4’, C-6’	

3.3 Biological Evaluation of Chondroitin Sulfate Disaccharide Library

3.3.1 Cell Viability Assay

The effect of CS sulfation patterns on breast cancer cell viability was first assessed via the MTS assay; this colorimetric method determined the number of viable cells present after drug treatment. The MTS solution contained a tetrazolium compound (Owen's reagent), which was reduced by dehydrogenase enzymes in metabolically active cells to form a colored formazan product.[42] The number of viable cells was thus proportional to the absorbance readings measured.

In the preliminary MTS assays conducted, the synthesized CS disaccharides were tested on 4 different human breast cell lines. This included the non-tumorigenic breast epithelial cell line MCF-12A, to evaluate compound cytotoxicity, and 3 breast cancer cell lines: MCF-7, T47D and MDA-MB-231. MCF-7 and T47D are low grade breast cancer cells which express the estrogen receptor; while MDA-MB-231 cells are high grade triple negative breast cancer cells (TNBC) which do not express the ER, PR, and HER2.[14]

The biological effect of each CS disaccharide was investigated by incubating the cells with the CS disaccharide for 72 hours, prior to the addition of MTS reagent to determine the number of viable cells after treatment period. Absorbance readings were measured after 3 hours; 4 different CS disaccharide concentrations were tested (0.1μg/mL, 1μg/mL, 10μg/mL and 100μg/mL).

We first screened the 16 CS disaccharides on normal breast cells (MCF-12A), and there was no significant change in cell viability even at the highest tested concentration of 100µg/mL (Figure 3-1). Hence these 16 CS disaccharides did not appear to be cytotoxic to the normal breast cells MCF-12A.

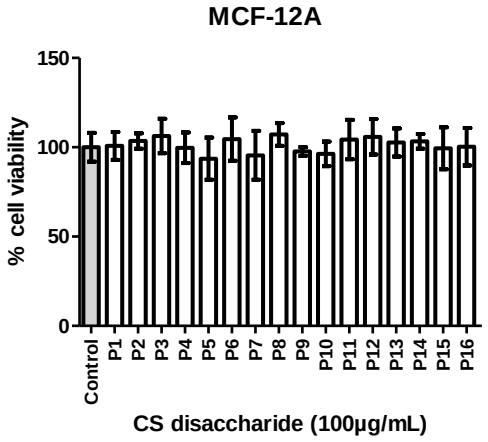


Figure 3-1. MTS assay results for MCF-12A breast cell line. Data represents the mean ± SD (n=6) with reference to non-treatment group (control), analyzed using one-way ANOVA with post-hoc Dunnett's test.

Interestingly, when the 16 CS disaccharides were tested on the more aggressive MDA-MB-231 cell line, a statistically significant decrease in cell viability was observed for CS disaccharides **P2**, **P13** and **P14** (Figure 3-2). This effect was only observed at the highest drug concentration of 100µg/mL, but not at the other lower drug concentrations (Figure 3-3).

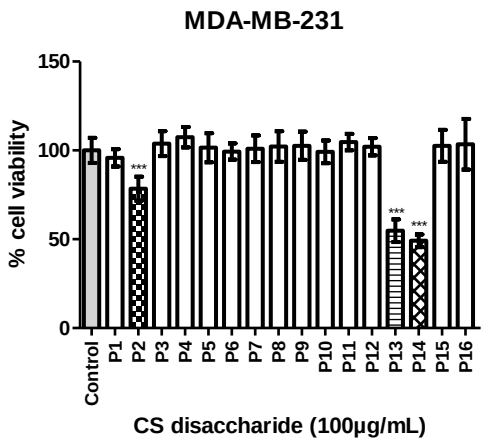


Figure 3-2. MTS assay results for MDA-MB-231 breast cancer cell line. Data represents the mean ± SD (n=6) with reference to non-treatment group (control), analyzed using one-way ANOVA with post-hoc Dunnett's test, ***p < 0.001.

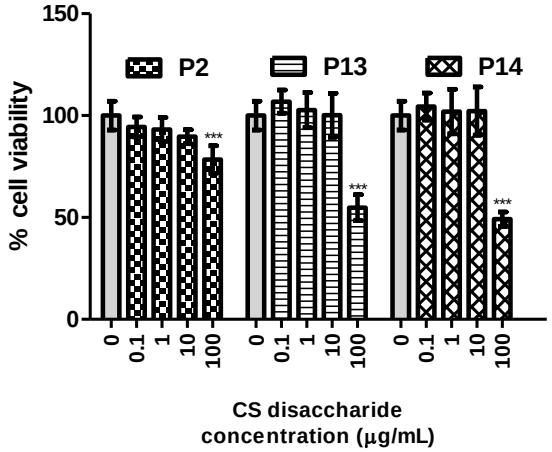


Figure 3-3. MDA-MB-231 MTS assay results for CS disaccharides **P2**, **P13**, **P14**. Data represents the mean ± SD (n=6) with reference to non-treatment group (control), analyzed using one-way ANOVA with post-hoc Dunnett's test, ***p < 0.001.

We next proceeded to screen the CS disaccharides on low grade breast cancer cells. The results showed no change in the number of viable cells after treatment with the CS disaccharides, indicating that all 16 CS disaccharides had no significant effect on both (a) MCF-7 cells and (b) T47D cells (Figure 3-4). This suggested differences in the pathological and biological features of MCF-7 and T47D cell types when compared to the MDA-MB-231 MTS assay results.

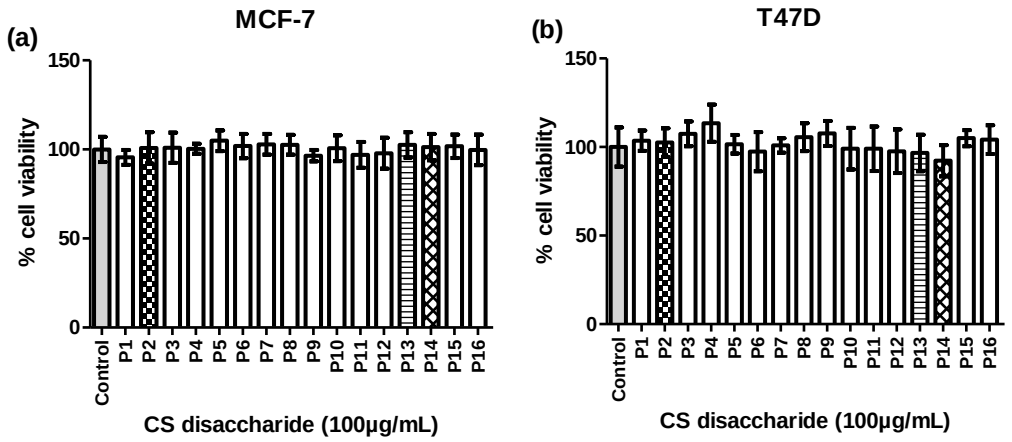


Figure 3-4. MTS assay results for (a) MCF-7 and (b) T47D breast cancer cell line. Data represents the mean ± SD (n=6) with reference to non-treatment group (control), analyzed using one-way ANOVA with post-hoc Dunnett's test

Based on the cell viability assay results obtained for CS disaccharides **P2**, **P13** and **P14** (Figure 3-5), both the number of sulfate groups present in the CS disaccharide and the position of these negative charges appeared to have an effect on MDA-MB-231 cell viability, but did not affect MCF-7, T47D and MCF-12A cells.

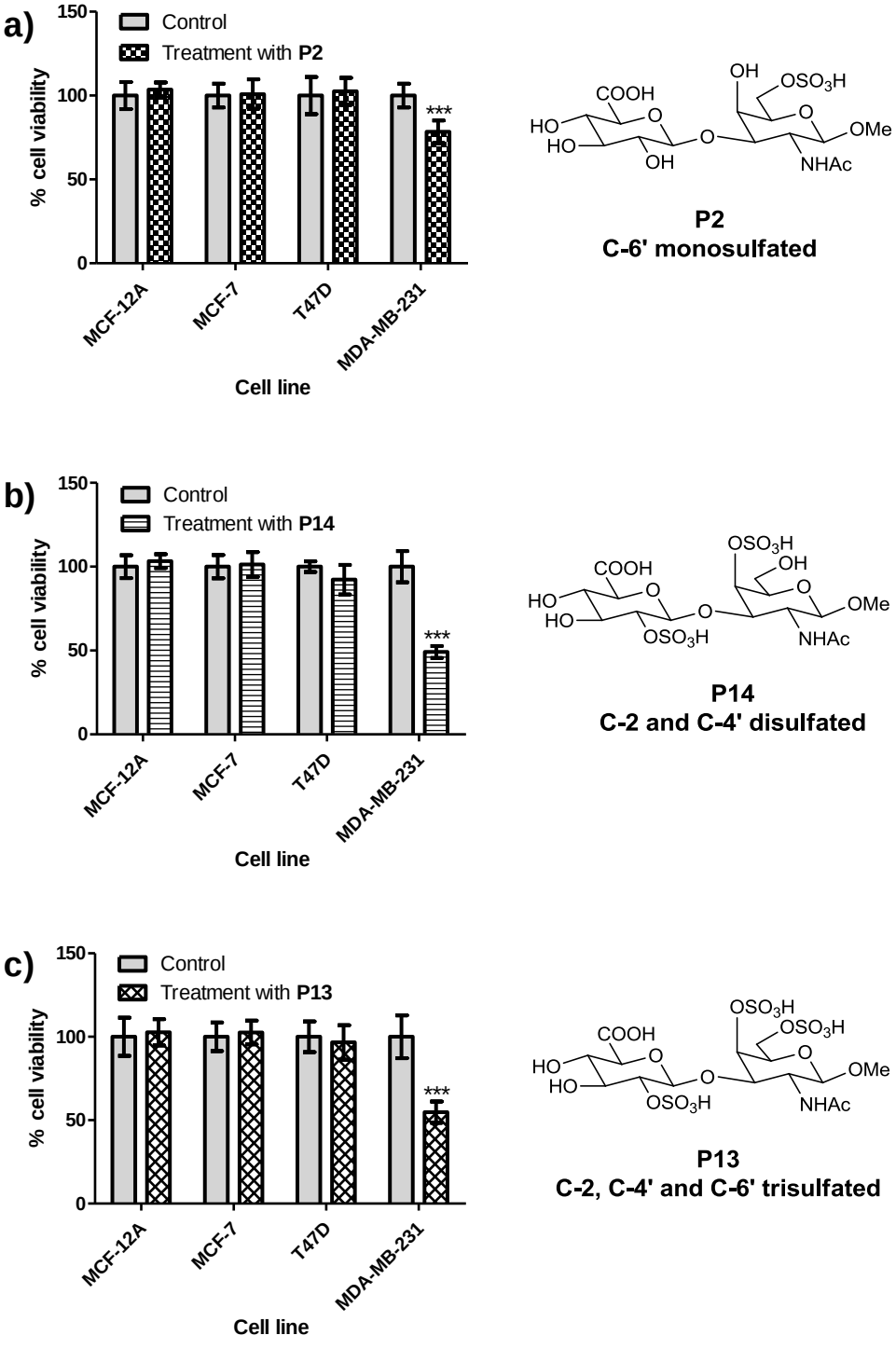


Figure 3-5. Cell viability assay results for MCF-12A, MCF-7, T47D and MDA-MB-231 cells at 100µg/mL CS disaccharide concentration; (a): **P2**, (b): **P14**, (c): **P13**. Data represents the mean ± SD (n=6) with reference to non-treatment group (control), ***$p < 0.001$.

Disaccharide **P2** was the only monosulfated CS analogue which elicited a 20% decrease in cell viability when treated on MDA-MB-231 cells (at 200μM drug concentration). **P2** has the sulfation pattern of CS-C, where the C-6' position is sulfated (Figure 3-5a). This sulfation sequence has previously been shown to be down regulated in breast cancer tumours.[35] Since our results indicate that CS-C could decrease cancer cell viability, malignant cells could thus have escaped the regulatory control exerted of CS by decreasing the expression of CS-C.

In the case of the disulfated CS analogues, a 50% decrease in cell viability was observed when MDA-MB-231 cells were treated with **P14** (at 150μM drug concentration). Disaccharide **P14** has sulfate groups attached on both the C-2 and C-4' positions (Figure 3-5b). This sulfation sequence is not commonly occurring and thus has not been extensively studied. Access to this CS sulfation motif through chemical synthesis enabled us identify the inhibitory effects of **P14** on MDA-MB-231.

For the trisulfated CS analogues, there was a 45% decrease in cell viability when CS disaccharide **P13** was tested (at 175μM drug concentration). Disaccharide **P13** has the sulfation pattern of CS-T, where the C-2, C-4' and C-6' positions are sulfated (Figure 3-5c). Again, this is a rarely occurring CS sulfation sequence which has not been previously evaluated in breast cancer. Our results indicate that non-commonly occurring CS sulfate motifs might also be important in regulating the progression of breast cancer.

The non-sulfated and fully sulfated CS disaccharides, **P1** and **P9**, had no effect on cell viability suggesting that the presence of some sulfate groups was required for CS to elicit an inhibitory effect on MDA-MB-231 cells but saturating all the possible sulfation sites would lead to a loss of activity.

3.3.2 Apoptosis Assay

To further evaluate the active CS disaccharides **P2**, **P13** and **P14**, apoptosis assays were subsequently conducted with the Caspase-Glo 3/7 kit which monitored the amount of caspase-3 and -7 activities present in the MDA-MB-231 cells after treatment with the respective CS disaccharides. Caspase (cysteine aspartic acid-specific protease) enzymes have been shown to play key effector roles in apoptosis in mammalian cells.[43, 44] Addition of the Caspase-Glo reagent first induced cell lysis, which released the caspase 3/7 enzymes for reaction with the proluminescent substrate. This generated the free aminoluciferin product, which was subsequently cleaved by luciferase to generated a "glow-type" luminescent signal[44]; the luminescent intensity was proportional to the degree of caspase activity 3/7 present.

Results from the caspase assay showed an increase in luminescence when MDA-MB-231 cells were treated with CS disaccharides **P2**, **P13** and **P14** (Figure 3-6), indicating an increase in caspase-3 and -7 activities. This suggested that the CS disaccharides could induce death of breast cancer cells via apoptosis, which

corroborated with the drop in cell viability observed in the previously conducted MTS assays.

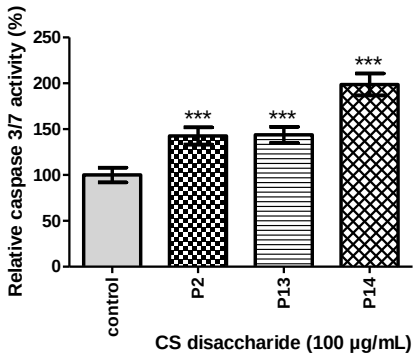


Figure 3-6. Caspase-Glo 3/7 assay results for MDA-MB-231. Data represents the mean ± SD (n=6) with reference to non-treatment group (control), ***$p < 0.001$.

Interestingly, the largest decrease in cancer cell viability and highest amount of caspase activity were seen in the CS disaccharide **P14**–treated group. On the other hand, **P13** elicited similar assay readings as **P2**, this indicates that **P13** could induce the death of breast cancer cells by other non-apoptotic pathways such as the inhibition of cancer cell proliferation. Preliminary biological evaluation thus indicates that sulfate groups present on CS could encode important regulatory information for cellular processes involved in breast cancer survival.

3.4 Conclusion

Biological evaluation of all 16 CS disaccharides was conducted on 4 different breast cell types, and results indicated that CS sulfation patterns had differential effects on different types of breast cancer cells. High grade breast tumor cells (MDA-MB-231) showed significant reduction in cell viability upon treatment with CS disaccharides **P2**, **P13** and **P14** while low grade breast tumor cells (MCF-7, T47D) and normal breast cells (MCF-12A) were unaffected. Apoptosis assays suggested that **P2**, **P13** and **P14** could induce apoptosis in MDA-MB-231 cells.

3.5 Experimental

General

Cell culture medium RPMI 1640 containing 2.05mM L-Glutamine, Dulbecco's Modified Eagle Medium (DMEM), Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12 (DMEM-F12) and Fetal Bovine Serum (FBS) were purchased from GE Healthcare Life Sciences (Buckinghamshire, United Kingdom). Human mammary and breast cancer cell lines MCF-12A, MCF-7, MDA-MB-231 and T47D were obtained from American Type Culture Collection (ATCC, USA), and cultured at 37^{o}C with 5% CO_2. The growth medium used was DMEM supplemented with 10% fetal bovine serum for MCF-7, RPMI-1640 supplemented with 10% fetal bovine serum for T47D and MDA-MB-231; and DMEM-F12 supplemented with 5% FBS, 0.5µg/mL hydrocortisone, 20ng/mL EGF, 10µg/mL insulin, 40µg/mL gentamicin and 100ng/mL cholera toxin for

MCF-12A. CellTiter 96® AQueous One Solution Cell Proliferation Assay (MTS assay) kit and Caspase-Glo® 3/7 Assay System were purchased from Promega (Wisconsin, USA). Absorbance readings were taken using the Tecan Infinite F200 Pro plate reader with the Tecan i-control software, while the luminescence readings were taken using the SpectraMax M5 plate reader using the SpectraMax software (integration time= 1 sec). The readings obtained were normalized with respect to the control data and statistical analysis calculated using GraphPad Prism software; level of statistical significance was set at $p < 0.05$.

Cell viability assay (MTS)

The cells were plated onto a 96-well plate and cultured for 24 hours at 37^{o}C (5% CO_2) with 100µL complete medium/well. The seeding density used was 2500 cells/well for MCF-12A; 4000 cells/well for MCF-7; 5000 cells/well for MDA-MB-231 and 6000 cells/well for T47D. After 24h growth period, the cells were treated with the desired CS disaccharide at 4 different concentrations: 0.1µg/mL, 1µg/mL, 10µg/mL and 100µg/mL. A control group was included where only the drug vehicle was used. Six biological replicates were set up for each group. Upon addition of the drug, the cells were incubated at 37^{o}C (5% CO_2) for 72 hours. After removal of the medium and washing of the cells with phosphate-buffered saline (PBS), the CellTiter 96® AQueous One Solution (MTS reagent) was next added to each well (100µL complete medium + 20µL MTS reagent per well) and the cells were incubated at 37^{o}C (5% CO_2) in the absence of light. Absorbance

readings (λ=490nm) taken at the 3rd hour were analyzed using one-way Analysis of Variance (ANOVA) with post-hoc Dunnett's test.

Apoptosis assay

MDA-MB-231 cells were plated on a 6-well plate with a seeding density of 200,000 cells/well and cultured for 24 hours at 37^{o}C (5% CO_2) with 2mL complete medium/well. After 24h, the cells were treated with the desired CS disaccharide at 100µg/mL drug concentration and incubated for a further 48 hours. A control set was included where only the drug vehicle was used. To facilitate luminescence measurements, the cells were then reseeded into a white opaque 96-well plate at a seeding density of 20,000 cells/well and incubated for an additional 24 hours. Thereafter, 100µL of Caspase-Glo® 3/7 reagent was added to each well (n=6). The contents in the wells were gently mixed and were incubated at room temperature in the dark. Luminescence readings taken at after 1 hr were analyzed using one-way Analysis of Variance (ANOVA) with post-hoc Dunnett's test.

3.6 References

1. Alteri, R., Barnes, C., Burke, A. & Gansler, T. *Breast Cancer Facts and Figures*. (American Cancer Society, Atlanta; 2013-2014).
2. Howlader N, Noone AM, Krapcho M & J, G. *SEER Cancer Statistics Review*. (National Cancer Institute; 1975-2012).

3. Lee Hin Peng, Ling, C., Yew, C.K. & Tin, K.T. *Trends in Cancer Incidence in Singapore*. (Singapore: National Registry of Diseases Office; 2009-2013).

4. Edge SB *et al. AJCC Cancer Staging Manual*, Edn. 7th. (New York: Springer; 2010).

5. Singletary, S.E. Rating the risk factors for breast cancer. *Ann. Surg.* **237**, 474-482 (2003).

6. Marmot, M., Altman, D., Cameron, D. & Dewar, J. The benefits and harms of breast cancer screening: an independent review. *Lancet* **380**, 1778-1786 (2012).

7. Tabar, L. *et al.* Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology* **260**, 658-663 (2011).

8. Hortobagyi, G.N. Treatment of breast cancer. *N. Engl. J. Med.* **339**, 974-984 (1998).

9. *British National Formulary*, Edn. 65th. (British Medical Association and Royal Pharmaceutical Society of Great Britain; 2013).

10. Perou, C.M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747-752 (2000).

11. Viale, G. The current state of breast cancer classification. *Ann. Oncol.* **23**, 207-210 (2012).

12. Bertucci, F. & Birnbaum, D. Reasons for breast cancer heterogeneity. *J. Biol.* **7**, 6-6 (2008).

13. Perou, C.M. *et al.* Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc. Natl. Acad. Sci.* **96**, 9212-9217 (1999).

14. Holliday, D.L. & Speirs, V. Choosing the right cell line for breast cancer research. *Breast Cancer Res.* **13**, 215 (2011).

15. Reis-Filho, J.S. & Pusztai, L. Gene expression profiling in breast cancer: classification, prognostication, and prediction. *Lancet* **378**, 1812-1823 (2011).

16. Gallo, M.A. & Kaufman, D. Antagonistic and agonistic effects of tamoxifen: significance in human cancer. *Semin. Oncol.* **24**, S1-71-S71-80 (1997).

17. Hudis, C.A. Trastuzumab — Mechanism of Action and Use in Clinical Practice. *N. Engl. J. Med.* **357**, 39-51 (2007).

18. Foulkes, W.D., Smith, I.E. & Reis-Filho, J.S. Triple-Negative Breast Cancer. *N. Engl. J. Med.* **363**, 1938-1948 (2010).

19. Iozzo, R.V. Basement membrane proteoglycans: from cellar to ceiling. *Nat. Rev. Mol. Cell Biol.* **6**, 646-656 (2005).

20. Yang, J. *et al.* Melanoma chondroitin sulfate proteoglycan enhances FAK and ERK activation by distinct mechanisms. *J. Cell Biol.* **165**, 881-891 (2004).

21. Asimakopoulou, A.P., Theocharis, A.D., Tzanakakis, G.N. & Karamanos, N.K. The biological role of chondroitin sulfate in cancer and chondroitin-based anticancer agents. *In Vivo* **22**, 385-389 (2008).

22. Svensson, K.J. *et al.* Chondroitin sulfate expression predicts poor outcome in breast cancer. *Int. J. Oncol.* **39**, 1421-1428 (2011).

23. Ricciardelli, C. *et al.* Elevated stromal chondroitin sulfate glycosaminoglycan predicts progression in early-stage prostate cancer. *Clin. Cancer Res.* **3**, 983-992 (1997).

24. Labropoulou, V.T. *et al.* Versican but not decorin accumulation is related to metastatic potential and neovascularization in testicular germ cell tumours. *Histopathology* **49**, 582-593 (2006).

25. Theocharis, A.D., Tsara, M.E., Papageorgacopoulou, N., Karavias, D.D. & Theocharis, D.A. Pancreatic carcinoma is characterized by elevated content of hyaluronan and chondroitin sulfate with altered disaccharide composition. *Biochim. Biophys. Acta* **1502**, 201-206 (2000).

26. Theocharis, A.D., Vynios, D.H., Papageorgakopoulou, N., Skandalis, S.S. & Theocharis, D.A. Altered content composition and structure of glycosaminoglycans and proteoglycans in gastric carcinoma. *Int. J. Biochem. Cell Biol.* **35**, 376-390 (2003).

27. Alini, M. & Losa, G.A. Partial characterization of proteoglycans isolated from neoplastic and nonneoplastic human breast tissues. *Cancer Res.* **51**, 1443-1447 (1991).

28. Olsen, E.B., Trier, K., Eldov, K. & Ammitzboll, T. Glycosaminoglycans in human breast cancer. *Acta Obstet. Gynecol. Scand.* **67**, 539-542 (1988).

29. Yin, L. Chondroitin synthase 1 is a key molecule in myeloma cell-osteoclast interactions. *J. Biol. Chem.* **280**, 15666-15672 (2005).

30. Cooney, C.A. *et al.* Chondroitin sulfates play a major role in breast cancer metastasis: a role for CSPG4 and CHST11 gene expression in forming surface P-selectin ligands in aggressive breast cancer cells. *Breast Cancer Res.* **13**, R58 (2011).

31. Herman, D. *et al.* CHST11 gene expression and DNA methylation in breast cancer. *Int. J. Oncol.* **46**, 1243-1251 (2015).

32. Monzavi-Karbassi, B. *et al.* Chondroitin sulfate glycosaminoglycans as major P-selectin ligands on metastatic breast cancer cell lines. *Int. J. Cancer* **120**, 1179-1191 (2007).

33. Garcia, J., Callewaert, N. & Borsig, L. P-selectin mediates metastatic progression through binding to sulfatides on tumor cells. *Glycobiology* **17**, 185-196 (2007).

34. Kim, Y.J., Borsig, L., Varki, N.M. & Varki, A. P-selectin deficiency attenuates tumor growth and metastasis. *Proc. Natl. Acad. Sci.* **95**, 9325-9330 (1998).

35. Potapenko, I.O. *et al.* Glycan gene expression signatures in normal and malignant breast tissue; possible role in diagnosis and progression. *Mol. Oncol.* **4**, 98-118 (2010).

36. Iida, J. *et al.* Role for chondroitin sulfate glycosaminoglycan in NEDD9-mediated breast cancer cell growth. *Exp. Cell Res.* **330**, 358-370 (2015).

37. Willis, C.M. & Klüppel, M. Chondroitin sulfate-E is a negative regulator of a pro-tumorigenic Wnt/Beta-catenin-collagen 1 axis in breast cancer cells. *PLoS ONE* **9**, e103966 (2014).

38. Yip, G.W., Smollich, M. & Gotte, M. Therapeutic value of glycosaminoglycans in cancer. *Mol. Cancer Ther.* **5**, 2139-2148 (2006).

39. Prinz, R.D., Willis, C.M., Viloria-Petit, A. & Kluppel, M. Elimination of breast tumor-associated chondroitin sulfate promotes metastasis. *Genet. Mol. Res.* **10**, 3901-3913 (2011).

40. Afratis, N. *et al.* Glycosaminoglycans: key players in cancer cell biology and treatment. *FEBS J.* **279**, 1177-1197 (2012).

41. Theocharis, A.D., Tsolakis, I., Tzanakakis, G.N. & Karamanos, N.K. Chondroitin Sulfate as a Key Molecule in the Development of Atherosclerosis and Cancer Progression. *Adv. Pharmacol.* **53**, 281-295 (2006).

42. Mosmann, T. Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. *J. of Immunol. Methods* **65**, 55-63 (1983).

43. Thornberry, N.A. & Lazebnik, Y. Caspases: enemies within. *Science* **281**, 1312-1316 (1998).

44. Promega Caspase-Glo® 3/7 Assay Technical Bulletin, USA, Madison; 2015.

CHAPTER 4: CONCLUSION AND PERSPECTIVES

4.1 Conclusion

In this project we devised a synthetic strategy to obtain 16 differently sulfated chondroitin sulfate (CS) disaccharides; these isomers encompass all the sulfation patterns theoretically plausible in the CS repeating unit. The strategy incorporated orthogonal protecting groups in glucuronic acid and *N*-acetyl galactosamine precursor units obtained through a divergent approach, prior to modular glycosylation and selective final transformations to prepare the library of 16 CS disaccharides (Figure 4-1). The CS disaccharides were then biologically evaluated on breast cancer cells. Cell viability assays indicated that high grade breast tumor MDA-MB-231 could be inhibited by disaccharides **P2**, **P13** and **P14**, while low grade breast tumor (MCF-7, T47D) and normal breast cells (MCF-12A) were unaffected. Apoptosis assays further suggested that disaccharides **P2**, **P13** and **P14** could induce apoptosis in MDA-MB-231 cells.

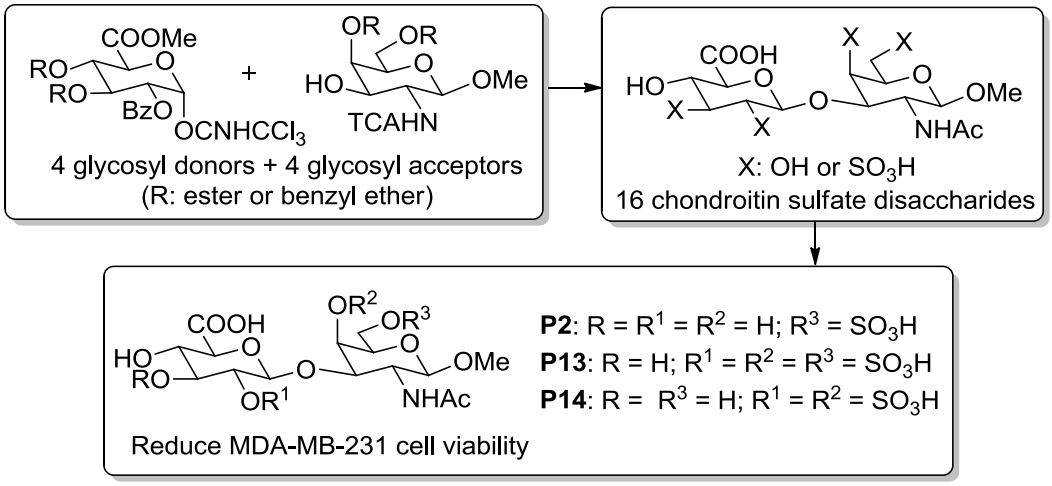


Figure 4-1. Chemical synthesis and biological evaluation of CS disaccharides

4.2 Future Perspectives

With the potent CS sulfate motifs identified, the mechanism of inhibition of MDA-MB-231 cells can next be elucidated by conducting additional phenotypic assays which include adhesion, migration and invasion assays to determine how disaccharides **P2**, **P13** and **P14** may affect cellular characteristics.

In addition, a relatively high concentration of 100μg/mL was required to elicit a decrease in MDA-MB-231 cell viability. Since longer CS sequences could provide stronger activities than disaccharides, further studies may include the synthesis and evaluation of CS tetra-/hexa- saccharides with the active sulfation profiles to determine if inhibition can be enhanced, using a lower concentration of CS compounds.

To extend the scope of study, the CS "sulfation code" may also be probed on other cancer cell lines which have dysregulated GAG expression, these include gastric and prostate cells. This enables us to investigate the molecular effects of CS sulfation in different forms of cancer biology.