Novel methods in glycosaminoglycan and proteoglycan synthesis

A thesis submitted in partial fulfilment for the degree of Doctor of Philosophy at the University of Oxford

by **Arghya Modak**

Department of Chemistry, University of Oxford

October, 2015

Wolfson College

Declaration

I declare the work presented in this thesis was carried out at the Chemistry Research Laboratory and Inorganic Chemistry Laboratory, University of Oxford, UK. All the work is my own, except where otherwise stated, and has not been submitted for any other degree at this or any other university.

Arghya Modak
October 2015

Abstract

Novel methods in glycosaminoglycan and proteoglycan synthesis

Arghya Modak

Wolfson College

University of Oxford, Department of Chemistry

Submitted towards the degree of D.Phil. in Trinity Term 2015

Glycosaminoglycans (GAGs) in general and heparin/heparan sulfate in particular are implicated in various biological functions by interacting with various protein partners. To elucidate the structure activity relationship (SAR) of these carbohydrate chains, considerable efforts towards the synthesis of these complex molecules have been made. Because of the heterogonous structure of these highly negatively charged molecules, a general approach towards synthesis of these molecules has not been possible. The central objective of this thesis was to develop a generic modular approach for the synthesis of defined thiomimics of heparin and heparan sulfate. Specifically we sought to utilize the unsaturation developed at the non-reducing end of heparin/heparan sulfate oligosaccharides by the action of lyases, for thiol-ene radical reaction to generate defined thiomimics of heparin

Chapter -2 Thiol-ene radical reaction on model substrates

The thiol-ene radical reaction on endocyclic double bonds in sugar molecules was initially established on protected model substrates and more importantly on deprotected model substrates in aqueous conditions. This demonstrates the first example of radical reaction in aqueous conditions for glycosylation between two carbohydrate substrates.

Chapter-3 Towards building *S*-linked heparin and heparan sulfate mimics using thiol-ene chemistry

The thiol-ene radical glycosylation strategy established on model substrates was utilized on both non-sulfated and sulfated building blocks of heparin/heparan sulfate to build novel thiomimics of heparin/heparan sulfate. The labile sulfate groups of heparin were found to be compatible with the reaction conditions. Initial biological studies done with the synthesized thiomimics did not show significant binding activity, probably due to the small size of the synthesized mimics.

Chapter -4 Glycosyltransferase utilization for heparin thiomimic synthesis

Natural glycosyltransferases were investigated for extension of the carbohydrate backbone of the thio mimics. Specifically, it was shown by MS for the first time that glucuronyltransferase can accept a thio mimic with unnatural stereochemistry as an acceptor.

Acknowledgement

This thesis is dedicated to my teachers, family and friends who have supported me throughout this journey.

A special thank you to the teachers that helped me from the start. I thank them for their dedication in teaching with fairness, perseverance, patience, and kindness. They went beyond the duty of their job and they saw potential in every student. They are a true inspiration.

I want to thank Prof. Benjamin G. Davis for giving me the opportunity to work on the heparin project. Many thanks to my collaborator, Prof. Jeremy Turnbull and Dr. Yassir Ahmed for all the help regarding the heparin and heparan sulfate substrates and their preparation protocols, and also for their help in conducting the biological assays. Thanks to Tim and Barbara for their help in running all the NMRs and allowing me to use NMR times like crazy for the tiny amount of samples. James Wickens from Mass Spectrometry service also deserves a special mention for his help in running all the MS and MS/MS samples and for teaching me how to use the LC-MS systems. I would also like to thank Commonwealth Scholarship Commission for their funding for the 1st three years.

Secondly, I would like to thank my proof-reading, chemistry-talking, gossip-mongering, beer-swilling colleagues past and present for making this journey a memorable one. Including (but not limited to!) Charlie, Maria, 'Captain' Wei Min, 'Commander' Jenz, Seb, Ritu, Bala, Mahima, Bhaskar, Yiqun, Sylvain and all the wonderful members of LG 9 and ICL.

Wolfson College has made the last four years some of the most memorable ones in my life. Thanks so much for all the help and support which has been provided over the years.

My family members and friends also have much to thank for, they are responsible for who I am today and I am truly grateful for that.

Last, but not the least, thank you to my viva examiners Prof. Tom Brown (Oxford) and Dr. Seung S. Lee (Southampton) for making my viva enjoyable and insightful.

Publication

Modak, A.; Liu, W-M.; Ahmed, Y.; Turnbull, J.E.; Davis, B.G. 'GAG ligation: Thiyl-Glycosylation allows direct synthesis of homogenous heparin mimics and construction of synthetic proteoglycans', Manuscript in preparation for submission to **Nature Chemical Biology.**

Poster

'New methods in glycosaminoglycan synthesis' **Modak, A**.; Ahmed, Y.; Turnbull, J.E.; Davis, B.G., Poster, *Pfizer* Mini-symposium, 2013, Department of Chemistry, University of Oxford, Oxford, UK. (October 2013)

Presentation

Modak, A. 'Novel chemo-enzymatic methods in glycosaminoglycan synthesis' *Astra Zeneca* Final Year D.Phil. talks, Department of Chemistry, University of Oxford, Oxford, UK. (October 2014)

Abbreviations

AIBN	Azobisisobutyronitrile
Ala	Alanine
APP	Amyloid precursor protein
APS	Ammonium persulfate
Arg	Arginine
Asn	Asparagine
BACE-1	β-secretase-1
BSA	Bovine Serum Albumin
BSP	1-benzene sulfonyl piperidine
C4ST	4-*O*-chondroitin sulfotransferase
C6ST	6-*O*-chondroitin sulfotransferase
CD	Circular dichroism
COSY	Correlation spectroscopy
CS	Chondroitin sulfate
D4ST	4-*O*-dermatan sulfotransferase
Da	Dalton
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DCM	Dichloromethane
DEAE	Diethylaminoethyl
DMPU	1,3-Dimethyl-3,4,5,6-tetrahydro-2-pyrimidinone
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DPAP	2,2-Dimethoxy-2-phenylacetophenone
DSF	Differential scanning fluorimetry

ELISA	Enzyme linked immunosorbent assay
ESI	Electrospray ionization
EtOAc	Ethyl acetate
EtOH	Ethanol
EXT	Exostosin-like glycosyltransferase
FGF	Fibroblast growth factor
FGFR	Fibroblast growth factor receptor
FPLC	Fast protein liquid chromatography
Fuc	Fucose
GAG	Glycosaminoglycan
Gal	Galactose
GalNAc	*N*-acetylgalactosamine
GalNAcT	*N*-acetylgalactosaminyltransferase
GalT	Galactosyltransferase
Glc	Glucose
GlcA	Glucuronic acid
GlcAT	Glucuronyltransferase
GlcN	Glucosamine
GlcNAc	*N*-acetylglucosamine
GlcNAcT	*N*-acetylglucosaminyltransferase
GlcNTFA	*N*-trifluoroacetylglucosamine
Glu	Glutamic acid
Gly	Glycine
His	Histidine
HMPA	Hexamethylphosphoramide

HPLC	High Performance Liquid Chromatography
HS	Heparan sulfate
HSQC	Heteronuclear Single Quantum Coherence
Hz	Hertz
IdoA	Iduronic acid
IF2	Initiation factor 2 protein
IPTG	Isopropyl thiogalactoside
K4CP	*E. coli* K4 chondroitin synthase
KfiA	*E. coli* K5 *N*-acetylglucosaminyltransferase
KS	Keratan sulfate
KSAc	Potassium thioacetate
Lys	Lysine
Man	Mannose
MeOH	Methanol
MES	2-(*N*-morpholino)ethanesulfonic acid
MS	Mass spectrometry
NDST	*N*-deacetylase-*N*-sulfotransferase
NeuAc	Neuraminic acid
NMR	Nuclear Magnetic Resonance
OST	*O*-sulfotransferase
PAPS	3'-phosphoadenosine-5'-phosphosulfate
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
Ph_2SO	Diphenyl sulfoxide
Phe	Phenylalanine

pmHAS	*Pasteurella multocida* hyaluronan synthase
pmHS2	*Pasteurella multocida* heparosan synthase 2
ppm	Parts per million
Pro	Proline
R_f	Retention factor
SAX	Strong anion exchange
Ser	Serine
SUMO	Small ubiquitin-related modifier protein
TEMED	Tetramethylethylenediamine
TEMPO	(2,2,6,6-Tetramethyl-piperidin-1-yl) oxyl
Tf_2O	Triflic anhydride
Thr	Threonine
TLC	Thin-layer chromatography
T_m	Melting temperature
TMS	Tetramethylsilane
UDP	Uridine diphosphate
UV	Ultraviolet
VEGF	Vascular endothelial growth factor
WIPE	water:isopropanol:ethylacetate
Xyl	Xylose
XylT	Xylosyltransferase
β-GlcNAcSH	2-acetamido-2-deoxy-1-thio-β-D-glucopyranose

Table of Contents

Chapter 1

Introduction

Table of Contents

Chapter – 1 Introduction

1.1 Glycosaminoglycans

Glycosaminoglycans (GAGs) are long chain linear polysaccharides consisting primarily of a disaccharide repeating unit made up of a hexosamine and uronic acid[1-6]. In some instances, the uronic acid is replaced by galactose[2]. Often these oligosaccharides undergo dynamic modification by various enzymes during biosynthesis leading to structural heterogeneity in terms of *N*-and *O*-sulfations, epimerizations and *N*-acetylations[1, 7]. Glycosaminoglycans have been implicated in various biological functions for example blood anti-coagulation[8], fertilization[9] and reproduction[10] amongst others. Apart from playing an important role in the organization of extracellular matrices because of their high charge density (thus being able to retain water molecules)[11], glycosaminoglycans have been reported to interact with different proteins especially chemokines[12] during inflammation[13-16] and for promoting chemotaxis[17-18]. Several growth factors also interact with glycosaminoglycans such as fibroblast growth factor (FGF), vascular endothelial growth factor (VEGF), hepatocyte growth factor and platelet-derived growth factor among others[7, 19-22]. These interactions mainly help in extracellular signaling[23].

Glycosaminoglycans can be largely classified into four different categories, based on the core carbohydrate backbone[24] shown in **Figure 1**:

a) **Heparin/Heparan sulfate**: It consists of α-L-iduronic acid/β-D-glucuronic acid attached in a 1,4 linkage to α-D-glucosamine which in turn is attached to the next uronic acid through a 1,4 linkage as well. Sulfation occurs in 3 and 6 position of glucosamine and 2 position of the uronic acid. The amine group of glucosamine is either *N*-acetylated or *N*-sulfated. In

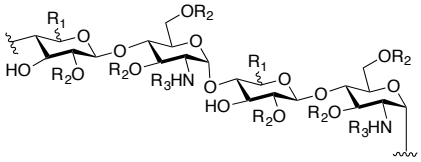
general, heparin has more sulfations on the carbohydrate units and more *N*-sulfated glucosamine units. Heparan sulfate contains relatively less sulfated carbohydrate units and the majority of glucosamine units are *N*-acetylated. The proportion of glucuronic acid is more in heparan sulfate whereas heparin contains more of iduronic acid. Heparan sulfate mostly exists as part of a proteoglycan where they are attached to a protein core. Thus they are mostly found on the surface of cells such as circulating leucocytes. Heparin in contrast can dissociate from the protein core to exist as free GAG chains[25]. Naturally available heparin ranges in molecular weight from 3 to 30 kDa depending upon the source it is isolated from. Commercial preparations of heparin, on the other hand, range in molecular weight from 12 to 15 kDa[26].

b) **Chondroitin sulfate/Dermatan sulfate:** It consists of α-L-iduronic acid/β-D-glucuronic acid attached via a 1,3 linkage to β-D-*N*-acetylgalactosamine while the hexosamine is attached to the next uronic acid through a 1,4 linkage. Sulfation is seen at position 4 and 6 of galactosamine and position 2 of the uronic acid. Chondroitin sulfate has a relatively higher number of glucuronic acid units as compared to iduronic acid units. In dermatan sulfate, the concentration of iduronic acid units is more as compared to the glucuronic acid units[24]. On an average, chondroitin sulfate chains have a molecular weight of around 50 kDa[27].

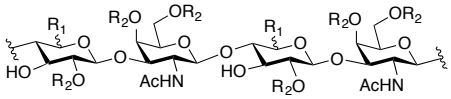
c) **Keratan:** Keratan consists of β-D-galactose attached in a 1,4 linkage to β-D-*N*-acetylglucosamine and the hexosamine in turn is attached to the next galactose unit through a 1,3 linkage. Sulfation can occur in the 6-position of both the *N*-acetylglucosamine unit and the galactose unit. The weight average molecular weight of keratan isolated from natural sources has been found to be around 20 kDa[27].

d) **Hyaluronic acid**: This glycosaminoglycan is completely unsulfated and it consists of β-D-glucuronic acid attached through a 1,3 linkage to β-D-*N*-acetylglucosamine. The hexosamine

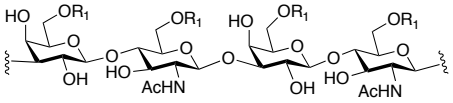
in turn is attached to the next glucuronic acid through a 1,4 linkage. Among all the glycosaminoglycans, hyaluronic acid has the largest molecular weight on average, with longer chains reaching a molecular weight of 10^4 kDa[27].



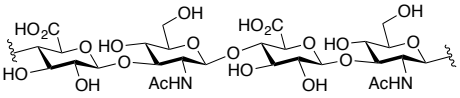
(A) Heparin/Heparan Sulfate



(B) Chondroitin Sulfate/Dermatan Sulfate



(C) Keratan



(D) Hyaluronic acid

Figure 1: Structures of the 4 categories of glycosaminoglycans. (A) Heparin/Heparan Sulfate: R_1 = axial CO_2H for iduronic acid, equatorial CO_2H for glucuronic acid, R_2 = H, SO_3^-, R_3 = Ac, SO_3^- (B) Chondroitin Sulfate/Dermatan Sulfate: R_1 = axial CO_2H for iduronic acid, equatorial CO_2H for glucuronic acid, R_2 = H, SO_3^- (C) Keratan: R_1 = H, SO_3^- (D) Hyaluronic acid

1.2 Glycosaminoglycan Biosynthesis

Biosynthesis of these glycosaminoglycans is carried out in different parts of the cell, with heparin/heparan sulfate and chondroitin sulfate/dermatan sulfate synthesis being carried out in the Golgi apparatus[28-29] and hyaluronic acid being synthesized by a plasma membrane synthase known as hyaluronan synthase, which has three different isoforms[30]. The corresponding oligosaccharides are assembled by an interplay of glycosyltransferases[3], sulfotransferases[31] and epimerases[32].

1.2.1 Chondroitin sulfate/dermatan sulfate biosynthesis

Chondroitin sulfate is *O*-linked to a protein backbone through a tetrasaccharide linker. The amino acid residue from the protein core, which is involved in *O*-linkage is serine and the tetrasaccharide linker consists of **–GlcA*β*1-3Gal*β*1-3Gal*β*1-4Xyl*β*1-*O*-(Ser)-**(GlcA = glucuronic acid, Gal = galactose, Xyl = xylose, Ser = serine). This whole protein carbohydrate complex is referred to as the proteoglycan. The protein core is synthesized in the cytosol and then translocated into the endoplasmic reticulum[33]. Thereafter, the tetrasaccharide linker gets assembled on the protein backbone in the endoplasmic reticulum and in the Golgi apparatus. Consequently, the synthesis of the remaining chondroitin sulfate structure takes place in the Golgi apparatus[29]. Assembly of the carbohydrate chain mostly happens on a serine amino acid which is part of a plausible consensus sequence, that has been reported as Ser-Gly/Ala-X-Gly, where 'X' is any other amino acid[24]. Indeed, Bourdon et al.[34] have also reported the plausible consensus sequence as being Ser-Gly-X-Gly, and this sequence is preceded by acidic residues such as glutamic and aspartic acid. Huber et al.[35] have reported on instances where the glycine in the second position of the amino acid consensus sequence has been replaced by alanine. However, contrary reports exist about the consensus amino acid sequence where the second glycine amino acid has been found to be replaced by other amino acids[36-37]. There are various chondroitin sulfate proteoglycans which have been reported such as neurocan[38], brevican[39], phosphoacan[40], decorin[41] and aggrecan[42], amongst others.

The tetrasaccharide linker is assembled using a combination of enzymes starting with xylosyltransferase, followed by galactosyltransferase I, galactosyltransferase II and glucuronyltransferase I[43-47] using the corresponding UDP-sugars. A *N*-acetylgalactosaminyltransferase then transfers a GalNAc unit to the glucuronic acid and this

event has been reported to determine whether the oligosaccharide chain is developing to be a chondroitin sulfate/dermatan sulfate or heparin/heparan sulfate[24]. There have been indications that the amino acids flanking the Ser-Gly consensus amino acid region might play a role in this event[48]. Chondroitin synthase (an enzyme having both *N*-acetylgalactosaminyltransferase and glucuronyltransferase activity) has been implicated in the elongation of the chain[49]. It should be noted though that a *N*-acetylgalactosaminyltransferase capable of adding both a GalNAc unit to glucuronic acid of the tetrasaccharide linker as well as extension of the chondroitin GAG chain, has been cloned too[50]. Therefore, it is plausible that both of these enzymes are responsible for transferring GalNAc during chain extension. The sulfations of the GAG chain are done by chondroitin 6-sulfotransferase[51], chondroitin 4-sulfotransferase[52] and 2-sulfotransferase that sulfates the 2-position of both iduronyl and glucuronyl residues[53]. All these sulfotransferases use PAPS (3'-phosphoadenosine-5'-phosphosulfate) as the sulfating substrate (**Figure 2**). In addition, a C-5-epimerase enzyme is implicated in converting glucuronic acid to iduronic acid for forming dermatan sulfate[54] (**Figure 3**).

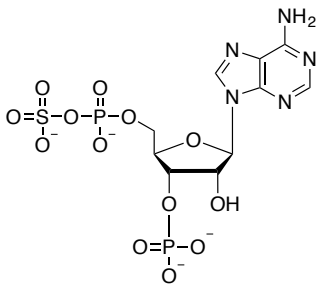


Figure 2: Structure of PAPS (3'-phosphoadenosine-5'-phosphosulfate). Since, the bond between the sulfate group and phosphate group is very labile, PAPS acts as the ideal sulfate donor.

1.2.2 Heparan sulfate/heparin biosynthesis

The tetrasaccharide linker connecting a heterogenous heparin/heparan sulfate (HS) chain to a serine amino acid residue of the protein core of a proteoglycan is the same as

reported for chondroitin sulfate i.e. **–GlcA*β*1-3Gal*β*1-3Gal*β*1-4Xyl*β*1-*O*-(Ser)-**[48]. This indicates that the biosynthesis of both the carbohydrate chains take place in the same way. When α-*N*-acetylglucosamine is added to the glucuronic acid of the tetrasaccharide linker, the chain is ultimately synthesized as heparan sulfate/heparin. In contrast when β-*N*-acetylgalactosamine is added to the linker glucuronic acid, it gets primed for chondroitin sulfate synthesis[55]. The transfer of this particular α-*N*-acetylglucosamine is undertaken by α-*N*-acetylglucosaminyltranferase I (GlcNAcT-I)[56] (**Figure 3**). As mentioned before, the tetrasaccharide linker is synthesized by four enzymes, xylosyltransferase (XylT), galactosyltransferase I (GalT-I), galactosyltransferase II (GalT-II) and glucuronyltransferase I (GlcAT-I) in sequence. Whether a tetrasaccharide sequence leads to the build up of a heparin or a chondroitin sulfate chain has been speculated to be affected by the amino acids, which flank the tetrasaccharide attachment site to the protein core. Acidic amino acid clusters as well as hydrophobic residues near the plausible consensus site for GAG attachment has been reported to favour the buildup of heparin/HS over chondroitin sulfate[28]. Indeed, the presence of a tryptophan residue in direct neighbourhood of the Ser-Gly GAG attachment site as well as Ser-Gly dipeptide repeats, have been reported as additional factors promoting HS assembly[57-59]. The GlcNAcT-1 activity is reported to be carried out by two EXT gene family proteins, EXTL2 and EXTL3[60-61]. An oligomeric complex composed of the gene products of EXT1 and EXT2 then extends the GAG chain further[62]. EXT1 and EXT2 have been implicated to have both glucuronyltransferase and *N*-acetylglucosaminyltransferase activities on their own[63-64]. As part of the modification of the carbohydrate backbone, *N*-deacetylase/*N*-sulfotransferase (NDST) catalyzes the removal of acetyl groups from α-*N*-acetylglucosamine and subsequent replacement by sulfate groups[65]. This function of *N*-deacetylation and consequent *N*-sulfation is a prerequisite for all other modifications such as *O*-sulfation and C-5 epimerization to take place[28,65]. This is followed by the action of C-5-epimerase, which is

implicated in the modification towards formation of both heparin and heparan sulfate. The inversion of the C-5 configuration has been reported to be reversible and therefore the enzyme can convert glucuronic acid to iduronic acid and vice versa. Also it has been observed that the action of this epimerase is distinct from the one involved in chondroitin sulfate, resulting in the inability of this enzyme to act on chondroitin sulfate GAGs and also the inability of chondroitin sulfate epimerase to act on heparin/HS[66]. 2-*O*-sulfation of the uronic acids is carried out in the next step.

Apart from the natural substrate, iduronic acid, the enzyme 2-*O*-sulfotransferase (2-OST) can also accept glucuronic acid, albeit with lower efficiency[67-68]. 3-*O*-sulfotransferase (3-OST) and 6-*O*-sulfotransferase (6-OST) follow next which sulfate the 3-position[69] and 6-position of glucosamine units, respectively[70]. The 6-*O*-sulfation occurs only on the 6-position of *N*-sulfoglucosamine and does not appear to happen on the primary position of *N*-acetylglucosamine[28,70]. All of these sulfotransferases use PAPS (3'-phosphoadenosine-5'-phosphosulfate) as the sulfate source as is the case for chondroitin sulfate sulfotransferases[71]. There are plenty of heparin/HS proteoglycans known too for eg. perlecan, agrin, syndecan and serglycin amongst others. They vary in terms of their localization, number of heparin sulfate chains and molecular weight of the core protein[2].

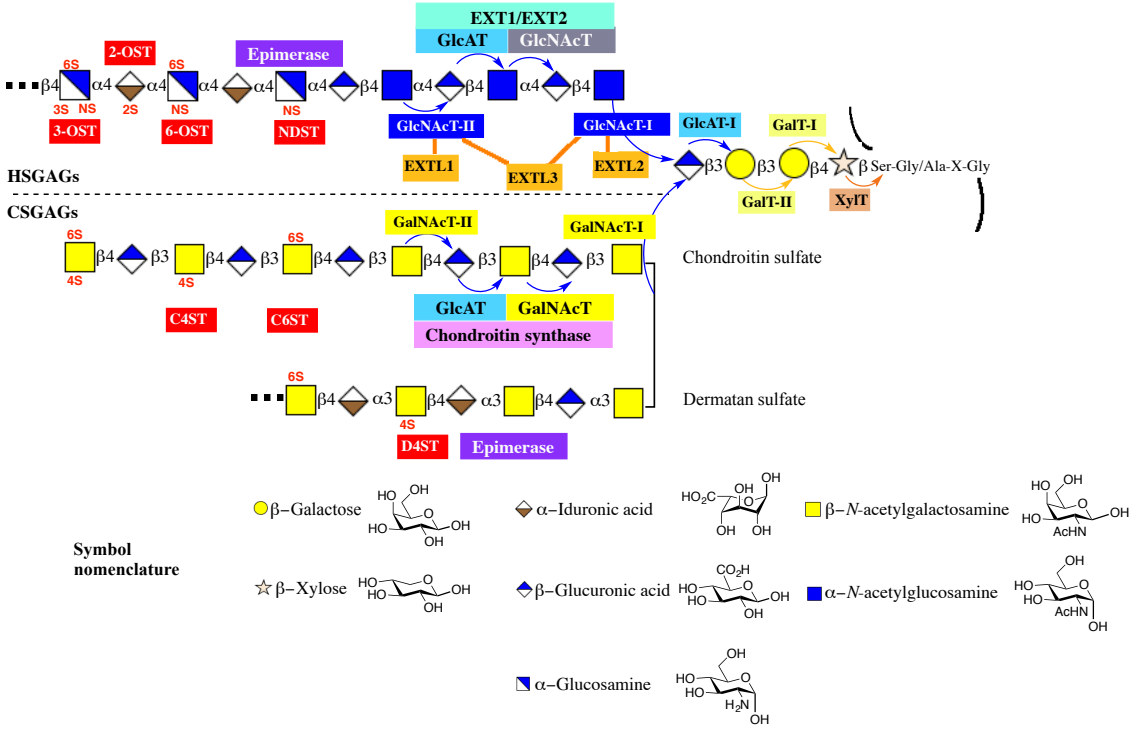


Figure 3: Chondroitin sulfate/dermatan sulfate and heparin/heparan sulfate biosynthesis. The different enzymes involved and the position of attachment to the core protein is shown[24].

1.2.3 Keratan sulfate biosynthesis

Keratan sulfate consists of disaccharide repeating units, which is composed of *N*-acetylglucosamine and galactose monosaccharides. Based on the nature of the linker region, keratan sulfate has been classified under three categories[72]:

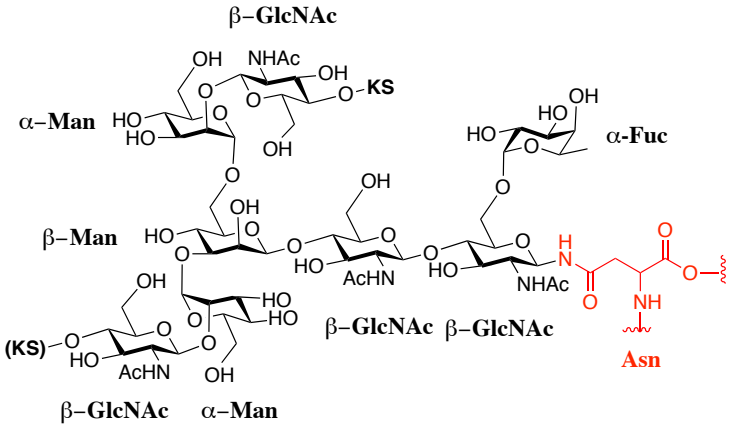
1. **Keratan sulfate I**: Keratan sulfate GAG residues are *N*-linked to an asparagine residue of the core protein through a branched polysaccharide, listed in **Figure 4**.

2. **Keratan sulfate II**: The GAG residues in this case are *O*-linked to either a serine or a threonine residue of the core protein via a mucin core-2 type carbohydrate residue (**Figure 4**). Mucin is a family of high molecular weight proteins, which have extensive glycosylations in the form of mostly *O*-linked oligosaccharides. The glycosylations mostly occur through a α-*N*-acetylgalactosamine monosaccharide, which in turn is attached to a serine or threonine

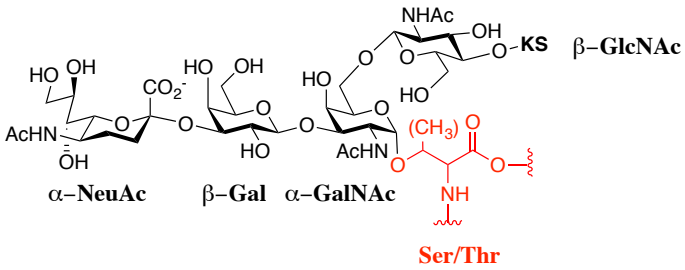
amino acid residue on the protein backbone[73-75]. Since, in case of keratin sulfate II, the linkage occurs through a α-*N*-acetylgalactosamine monosaccharide to a serine/threonine residue on the protein, the linkage residue is termed as a mucin type carbohydrate residue.

3. **Keratan sulfate III**: In this case, the sulfated sugar is *O*-linked to a serine residue through mannose (**Figure 4).**

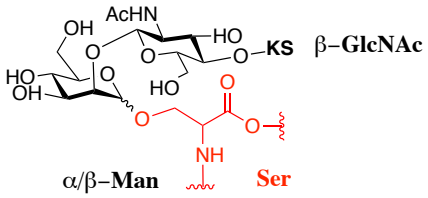
The consensus sequence for the assembly of keratan sulfate GAGs on protein backbone is believed to be a series of hexapeptides repeated one after another. The hexapeptide sequence is Glu-Glu/Lys-Pro-Phe-Pro-Ser[76]. The synthesis of the linker region for keratan sulfate I has been attributed to the biosynthetic machinery involved in high-mannose precursor oligosaccharide processing[72]. As for keratan sulfate III, an important enzyme protein-*O*-mannose β-1,2-*N*-acetylglucosaminyltransferase, POMGnT1 is needed for transferring *N*-acetylglucosamine units to mannosylated proteins[77]. The extension of the GAG chain is carried out by two enzymes, acting in tandem, a β-1,4-galactosyltransferase[78] and a β-1,3-*N*-acetylglucosaminyltransferase (β3GlcNAcT), which has been identified even though a number of plausible candidates have been enlisted[79-80] (**Figure 5**). More recently, Kitayama et al. have demonstrated the use of two specific isoforms of the two glycosyltransferases, β1,3-*N*-acetylglucosaminyltransferase-7 (β3GnT7) and β1,4-galactosyltransferase-4 (β4GalT4) for the synthesis of keratan sulfate *in vitro*, thus indicating that these are the enzymes responsible for keratan sulfate chain extension[81].



Keratan Sulfate I



Keratan Sulfate II



Keratan Sulfate III

Figure 4: Linkage oligosaccharides of Keratan Sulfate linking KS to protein core[72]. **KS** = keratan sulfate, **Man** = mannose, **GlcNAc** = *N*-acetylglucosamine, **Asn** = asparagine, **GalNAc** = *N*-acetylgalactosamine, **Gal** = galactose, **Ser** = serine, **Thr** = threonine, **Fuc** = fucose, **NeuAc** = neuraminic acid., mostly present as a capping sugar. **KS** in brackets refers to positions where keratan sulfate chains may not be present in all cases. **Asn, Ser** and **Thr** are amino acids which are part of the core protein backbone.

Both galactose and *N*-acetylglucosamine are sulfated in keratan sulfate. While galactose is sulfated by KS-Gal6ST[82], sulfation of GlcNAc residues is carried out by *N*-acetylglucosaminyl-6-sulfotransferase (GlcNAc6ST). One specific isoform of this enzyme,

C-GlcNAc6ST has been found to be involved in corneal keratan sulfate biosynthesis[83]. One important aspect of this enzyme is the ability to only sulfate GlcNAc sugars at the non-reducing terminus of the chain. Since almost all of the GlcNAc units in a keratan sulfate chain are normally sulfated, it has been inferred that chain elongation and sulfation occurs simultaneously for these GAGs[84-85]. Examples of keratan sulfate proteoglycans are keratocan, lumican, mimecan and fibromodulin among others[2].

1.2.4 Hyaluronic acid biosynthesis

Hyaluronic acid or hyaluronan is synthesized in the inner side of the plasma membrane as opposed to other GAGs, which are synthesized in the Golgi apparatus[86]. Unlike other GAGs, this glycosaminoglycan is not constructed as a carbohydrate chain attached to a protein core. Rather it is secreted as a carbohydrate chain into the extracellular space by the enzymes responsible for hyaluronic acid synthesis, hyaluronan synthases[86]. It has been determined that hyaluronan synthase has two active sites which enable it to transfer both glucuronic acid and *N*-acetylglucosamine to the growing chain of GAG[87-89]. There are several isoforms of hyaluronan synthase which have been discovered.

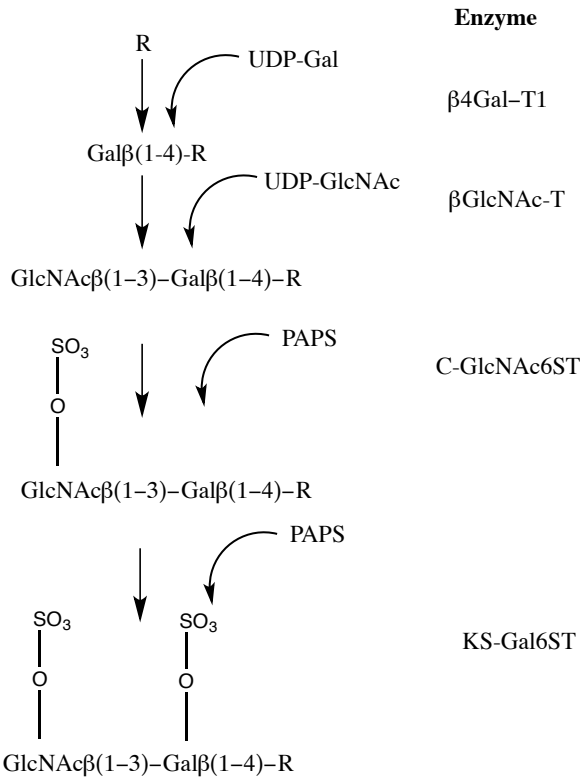


Figure 5: Keratan sulfate biosynthesis. R represents the growing chain. βGal-T1 = β-1,4-galactosyltransferase, βGlcNAc-T = β-1,3-*N*-acetylglucosaminyltransferase, C-GlcNAc6ST = corneal *N*-acetylglucosaminyl-6-sulfotransferase, KS-Gal6ST = keratan sulfate galactosyl-6-sulfotransferase, PAPS = 3'-phosphoadenosine-5'-phosphosulfate[72].

1.3 Heparin interaction with proteins

Heparin, the most electronegatively charged biomolecule[90], is implicated in various biological functions such as blood anticoagulation, viral and bacterial infection and host cell entry, angiogenesis, cancer and inflammation among others[6, 91-94].

Study of the structure of heparin for a long time has revealed quite a lot about the structure-activity relationship in the context of interaction with various protein partners. In particular, the sequence of heparin, which is responsible for interacting with antithrombin, resulting in its activity as an anticoagulant, has been studied in detail. A specific pentasaccharide sequence, constituting about one third of the complete heparin chain, is

responsible for the anticoagulant activity. This domain, which binds to antithrombin (**Figure 6**), is referred to as the major sequence while the rest of the sequence, termed minor sequence is also involved in other biological activities[95].

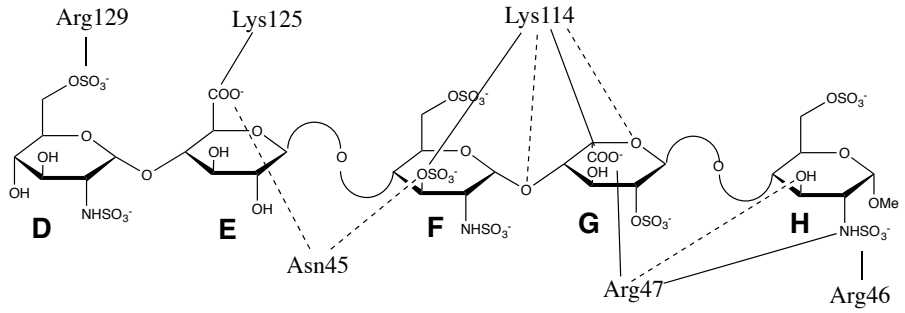


Figure 6: Schematic representation of antithrombin-heparin (pentasaccharide region) interaction. Solid lines denote salt bridges whereas dashed lines represent hydrogen bonds. Individual monosaccharide units are labeled as D to H[96].

Apart from interacting with antithrombin for its anticoagulant activity, heparin has been found to be interacting with fibroblast growth factors (FGFs). FGFs are a family of signaling polypeptides involved in a wide variety of physiological processes. Heparin interacts specifically with FGFs, which results in oligomerization of FGFs[97]. These oligomers in turn interact with the FGF receptors (FGFR) to induce cellular response through FGFR dimerization[98]. Thus, heparin helps by facilitating the complex formation between FGF and its receptor (FGF-FGFR complex) and also stabilizing the FGF-oligomer complex. Crystal structures solved for FGF-heparin-FGFR complex have shown that there are different stoichiometries of interaction among FGF, heparin and FGFR depending on the FGF protein[99-100]. Heparin and heparan sulfate are also known to interact with cytokines[101]. Due to the presence of sulfations and carboxylic groups on other GAGS such as chondroitin sulfate and dermatan sulfate, they have also been reported to interact with heparin binding proteins, with different affinities[102-103].

Heparin and heparan sulfate are also known to interact with chemokines[101]. Even though these protein molecules normally remain in monomer state, interaction with negatively charged heparin or heparan sulfate changes their oligomerization state to dimer or even higher orders[12]. Though the significance of this heparin assisted oligomerization is not completely understood, it is speculated that this might lead to better protection of proteins from proteolytic digestion and therefore increases its stability[12]. As chemokines are known to induce migration of cells by chemotaxis, it is plausible that this heparin chemokine interaction represents a building up of higher oligomerization state of chemokines. This increases the chemokine's local concentration and serves to offer a directional gradient for the cells to migrate[104].

The interaction of heparan sulfate and to a minor extent of heparin with amyloid precursor proteins (APP) is also similar to other protein binding partners. This precursor for amyloid-β, which is implicated in Alzheimer's disease[105], interacts with heparan sulfate to form dimers in a protein:sugar ratio of 2:1[106]. This interaction has been found to influence neurite outgrowth[107].

1.4 Heparin structural features

Heparin and heparan sulfate consist of alternating residues of an amino sugar in the form of D-glucosamine or D-*N*-acetylglucosamine and an uronic acid in the form of D-glucuronic acid or L-iduronic acid. The hydroxyl groups and free amino groups of glucosamine are heterogeneously sulfated to varying degrees. The major disaccharide sequence constituting heparin is a trisulfated sequence, $IdoA2SO_3$-$GlcNSO_36SO_3$. These sulfates (i.e. *N*-, 2-*O*-, and 6-*O*-sulfates) together with the carboxylate of uronic acid are responsible for the polyelectrolytic properties of heparin and also for electrostatic interaction,

hydrogen bonding or salt bridge formation with basic amino acids of the target protein[92]. Most of the heparin binding proteins bear increased levels of basic amino acid residues such as arginine, lysine and histidine allowing them to interact with the negatively charged GAG saccharides[104, 108]. Both GlcA and GlcN are found to be in the classic $^4C_1$ chair conformation[109]. The IdoA pyranose, however adopts either a $^4C_1$, $^1C_4$ chair or a $^2S_0$ skew boat conformation, all of which have been found to be equi-energetic (**Figure 7**).

Iduronic acid as a monosaccharide normally assumes the $^1C_4$ conformation because it is in this conformation that the bulky carboxylic group can be positioned in the equatorial position thus minimising steric hindrance. However, this uronic acid residue remains in a dynamic equilibrium among the three states ($^4C_1$, $^1C_4$ or a $^2S_0$)[110], when it is part of a glycosaminoglycan chain, so that most of the bulky substituents can be positioned in the equatorial position. The relative population of each conformer is further dependant on the degree of sulfation of the uronic acid itself and also of the adjacent GlcN residue[110-111]. Apart from this, the relative population has been observed to be affected by extrinsic factors, such as the type of counter ions present in the medium[111-112]. It is speculated that the ability of iduronic acid to adopt any of the three conformations confers a degree of flexibility to IdoA containing GAG chains. This proves valuable for binding to proteins and other biological activities[113]. Investigations concerning interactions of heparin octasaccharides with antithrombin, have shown that the non-sulfated iduronic acid unit located just before the antithrombin binding pentasaccharide region adopts a $^1C_4$ conformation when bound to antithrombin and $^2S_0$ conformation in the protein's absence[114]. Even cases in which the iduronic acid is directly involved in protein binding, the sugar populates the $^1C_4$ conformation[115]. Thus this flexibility of iduronic acid helps in more efficient binding of heparin to its protein binding partners, as it can simply adopt the required conformation. However the glucosamine unit, being more rigid, remains in $^4C_1$ conformation throughout.

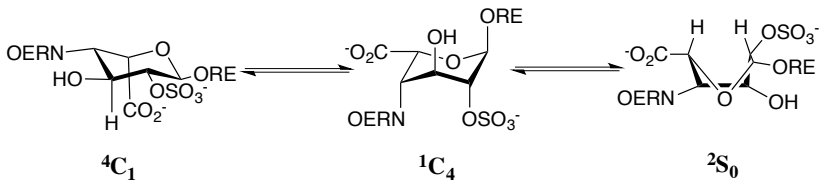
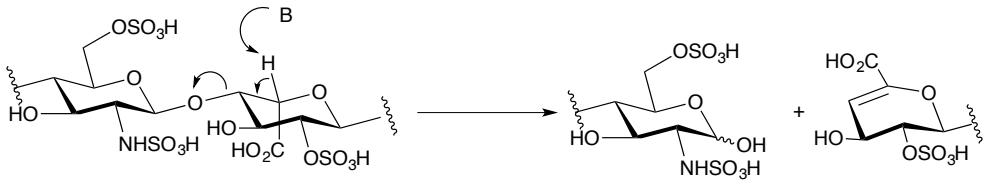


Figure 7: Three conformations in which iduronic acid exists, $^{4}C_{1}$,$^{1}C_{4}$ and $^{2}S_{0}$. OERN = non-reducing end glycosidic bond, ORE = reducing end glycosidic bond[5].

1.5 Glycosaminoglycan lyases

Glycosaminoglycan lyases are mostly synthesized by bacteria to break down GAGs. They have proved to be useful in elucidating GAG sequences as they are highly specific towards the GAG sequence on which they act[116-117]. These enzymes have also been put to therapeutic use where their GAG degrading ability has been exploited to cure various medical complications[118-119]. The mechanism by which lyases cleave an oligosaccharide is depicted in **Scheme 1**. A basic residue in the enzyme active site abstracts the C-5 ring proton of an uronic acid followed by elimination of the C-4 substituent. This leads to the formation of an unsaturation between C-4 and C-5. At the same time the linkage between C-4 and the glycosidic oxygen is cleaved and the resulting free anomeric oxygen on the non-reducing side of the uronic acid is protonated to form the hemi-acetal. The enzymes are therefore specific for cleaving the bond on the non-reducing side of uronic acids as the carboxylic group also takes part in the reaction. It makes the C-5 proton acidic enough to be abstracted by the basic residue of the enzyme.



Scheme 1: Mechanism of action of heparin lyases. Representative mechanism shown for Heparinase-1. **B** represents a basic amino residue in the enzyme.

There are mainly three categories of GAG lyases[120]:

a) Heparinase

b) Chondroitinase

c) Hyaluronidase

1.5.1 Heparinase

Heparinases are GAG lyase enzymes acting on heparin and heparan sulfate residues. They are expressed by several soil bacteria which includes *Flavobacterium heparinum*, *Bacillus sp*., *Bacteroides heparinaolyticus* etc.[121] Depending on their substrate specificities, they have been further classified as Heparinase I (acting mainly on heparin), Heparinase II (acting both on heparin and heparan sulfate) and Heparinase III (acting mainly on heparan sulfate). The specificities of these enzymes are known.

(i) Heparinase I cleaves the linkage $H_{NS,6X}$-I_{2S}

(ii) Heparinase II cleaves the linkages $H_{NY,6X}$-G/I_{2X}

(iii) Heparinase III cleaves the linkage H_{NAc}-I and $H_{NY,6X}$-G

(H = hexosamine, I = iduronic acid, G = glucuronic acid, Y = sulfated or acetylated and X = sulfated or unsubstituted).

Glucuronic acid containing linkages are not degraded by Heparinase I. This enzyme acts selectively on iduronic acid-containing units. It has also been seen that the bond between *N*-acetylated residue and iduronic acid is less susceptible to cleavage by Heparinase I as compared to the bond between *N*-sulfated residue and iduronic acid[122]. Furthermore, the activity of Heparinase I seems to increase with longer substrates[123].

Heparinase II has been shown to have much higher activity on heparan sulfate as compared to heparin[124]. Since oligosaccharides containing both iduronic acid and glucuronic acid are susceptible to degradation in the presence of this enzyme, Heparinase II has the broadest substrate specificity. It is also a matter of speculation as to whether it has a single active site or two active sites to accommodate iduronic acid and glucuronic acid substrates as the proton needs to be abstracted from C-5, which is pointing in two different directions in the two uronic acids[121].

Heparinase III is specific for heparan sulfate and is inactive towards linkages containing iduronic acid[125]. Like Heparinase I its activity increases with increased substrate length.

1.5.2 Chondroitinase

Chondroitinase enzymes are GAG lyase enzymes acting on chondroitin sulfates. They have been mostly isolated from genera of bacteria such as *Arthrobacter*, *Flavobacterium*, *Aeromonas, Bacillus* etc. Both soil and intestinal bacteria produce these enzymes. There are four different types of chondroitinase enzymes, which have been isolated and they have been

classified as Chondroitinase ABC, Chondroitinase AC, Chondroitinase B and Chondroitinase C based on their substrate. Chondroitinase enzymes cannot degrade heparin or heparan sulfate which have 1,4 linkages specifically even though they cleave a 1,4 linkage in chondroitin sulfates. This indicates that the 1,3 linkage in chondroitin sulfate also has a role to play in determining substrate specificity for this enzyme. However, it can act on hyaluronic acid[121].

1.5.3 Hyaluronidase

Hyaluronidase has been isolated from both bacterial sources such as those belonging to *Propionibacterium*, *Staphylococcus*, *Streptococcus* genera as well as animal sources[120]. Apart from hyaluronic acid, this enzyme can also act on chondroitin 4-sulfate and chondroitin 6-sulfate albeit with a lower activity[126].

1.6 Synthesis strategies for glycosaminoglycans

There has been numerous strategies devised over the years for the artificial synthesis of glycosaminoglycans in general and heparin/heparan sulfate in particular. Even though heparin has been used as an anticoagulant drug for a considerable amount of time, the total chemical synthesis of heparin remains an underdeveloped field mainly because of the heterogeneity in the sulfation pattern on the oligosaccharide and the configuration of the uronic acids[127-128]. Thus, a modular approach for the synthesis of these irregular structures has proved to be very difficult.

The approaches towards synthesis of GAGs can be broadly classified under the following sub headings:

a) Chemical synthesis of GAGs

b) Chemoenzymatic synthesis of GAGs

1.6.1 Chemical synthesis of GAGs

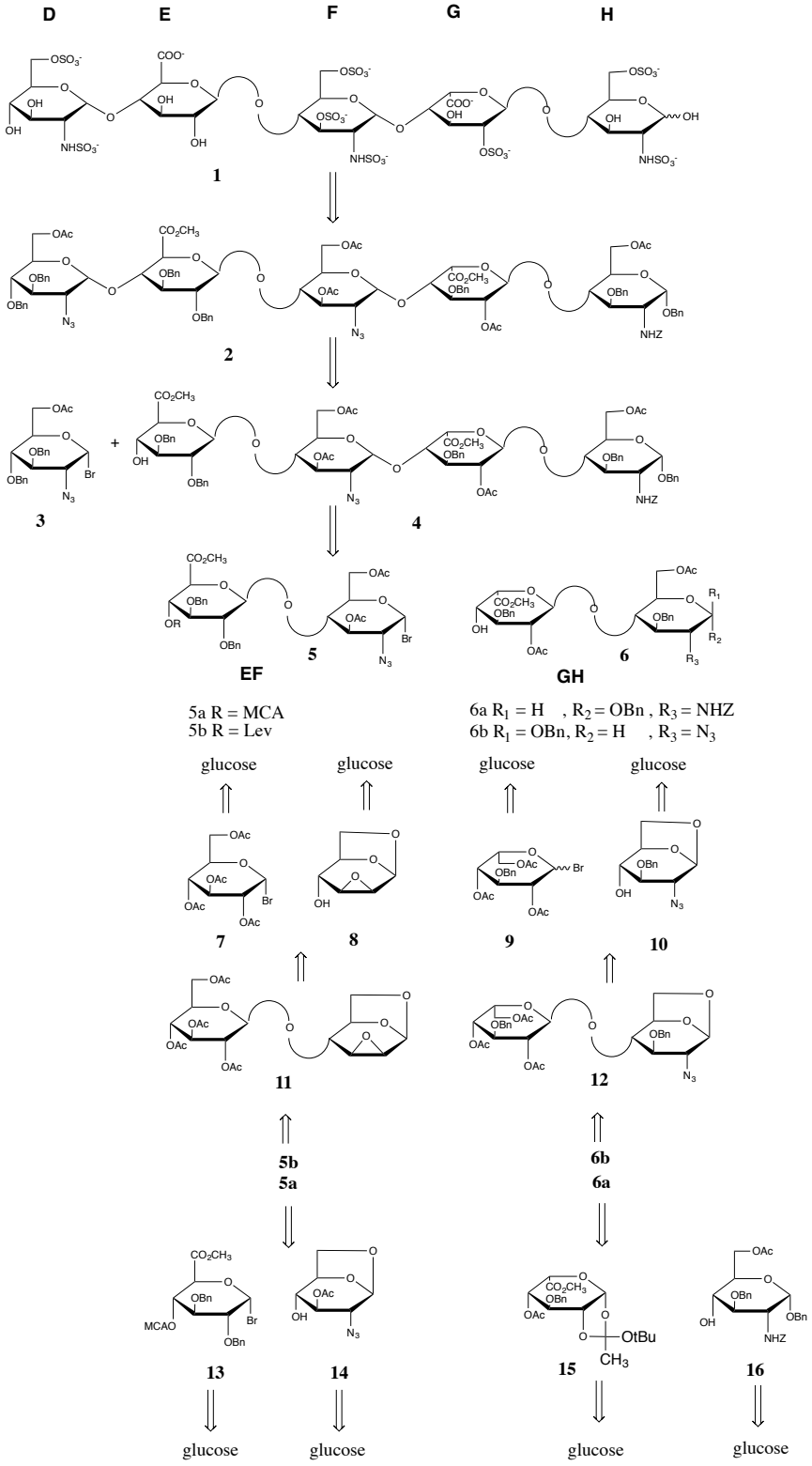
Due to its heterogenous structure and its various biological functions, it is not surprising that numerous synthesis strategies have been devised previously. The designed synthetic approaches have to account for appropriately protected building blocks (uronic acid and hexosamine units). Thus, orthogonal temporary protecting groups need to be installed on the C-4 hydroxyl group of each sugar for selective deprotection and glycosylation at a later stage in order to extend the sugar chain. The functional group repertoire also needs to be accessed in order to differentiate hydroxyl groups on the carbohydrate chain, which will be sulfated, from those that will be free. Focus has also been put on the facile synthesis of L-iduronic acid derivatives as the unnatural and therefore non-accessible isomer. In this regard, the desired derivatives have been synthesized using D-glucuronic acid glycals, D-glucuronolactone and D-diacetone glucose as starting materials[129-139].

Jaurand et al .have reported the synthesis of the basic trisulfated disaccharide building block of heparin in modest yields[140]. Incidentally, this is the most common disaccharide component of heparin. It has been shown by Haller et al.[141] that the uronic acid moiety can be introduced at a later stage of the synthesis by oxidation of the primary position using TEMPO. This circumvents the problem of low yield of glycosylation because of the deactivating nature of the carboxylic acid, epimerization of C-5 position and also protecting

group manipulations[141]. Diacetone glucose has been demonstrated to be a good precursor for incorporation of iduronic acid in the synthesized disaccharide through the intermediate 1,2,3,5-di-*O*-isopropylidene-β-L-idofuranose[142-143]. Prabhu et al. have demonstrated the judicious design of six monosaccharide precursors which can be used to build up all of the twenty combination of disaccharides found in heparan sulfate[144].

Boeckel et al.[145] and Sinay et al.[146] have reported on the synthesis of the heparin pentasaccharide region responsible for binding with antithrombin. Incidentally, both have reported on a 2+2+1 retrosynthetic approach for the total synthesis (**Scheme 2**).

While Sinay et al. synthesized the uronic acid derivatives prior to coupling with the protected amino sugar, Boeckel et al. synthesized the protected disaccharides first, followed by oxidation of the primary position of the non amino sugar to access the uronic acid unit. In both cases, they used similar orthogonal protecting groups to take into consideration the following details (i) sulfate groups on some hydroxyl groups and also on the amino group (ii) free hydroxyl groups and carboxylate groups in proper positions (iii) use of non participating protecting groups on the amino sugar for α-glycosylation of the amino sugars (iv) use of participating protecting group on the C-2 hydroxyl group of the uronic acid for β-glycosylation.



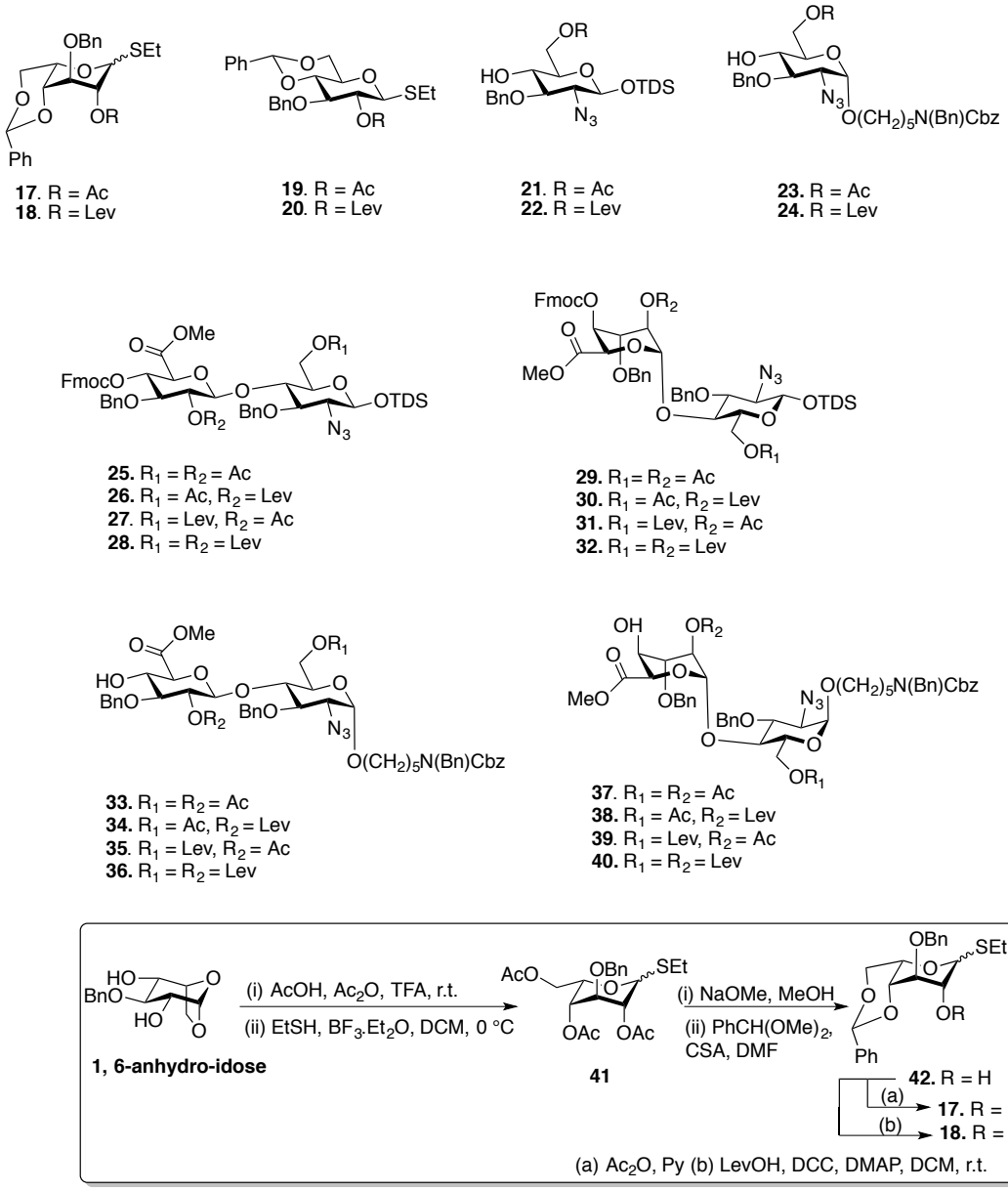
Scheme 2: Retrosynthetic analysis of the antithrombin binding pentasaccharide region as followed by Boeckel et al.[145] and Sinay et al.[146] Synthesis of the disaccharide building blocks **EF** and **GH** was done in two ways (1) both glucose and idose derivatives are oxidized at the monosacharide level resulting in **13** and **15.** Coupling is then done with the glucosamine derivatives, **14** and **16** to give the

disaccharides **5a** and **6a**[146]**.** (2) Glucose and idose derivative monosaccharides **7** and **9** are coupled initially to the amino sugar monosaccharides or precursor derivatives, **8** and **10,** to yield disaccharides **11** and **12** followed by the oxidation of the uronic acid precursors[145]**.** Final synthesis in both cases is then done with the disaccharide synthons and monosaccharide **3**.

The synthesis of this pentasaccharide region was important as it served as the foundation for the design of the anticoagulant drug Fondaparinux[147].

In addition, other approaches towards the synthesis of heparin analogues, lacking the sulfate groups in key hydroxyl positions, to study the effects of these groups on its interactions with protein binding partners, have been reported[148-149]. *N*-sulfates[150] and some carboxylate groups[151] have been found to play a key role in the interaction of heparin pentasaccharide region to antithrombin. More recently, Arungundram et al.[152] have reported on the modular synthesis of heparan sulfate oligosaccharides. Judiciously designed core disaccharide scaffolds were used for the construction of multiple tetrasaccharides and a hexasaccharide (**Scheme 3**). The idosyl donors in this case were synthesized from 1,6-anhydro-idose, which is readily available. These oligosaccharides were then used for binding studies with BACE-1 (an enzyme known to process amyloid precursor proteins in Alzheimer's disease pathway) to identify key charged groups on the carbohydrate chain responsible for binding.

A similar strategy has been adopted by Schwörer et al. where they have targeted variously sulfated hexa-, octa-, deca- and dodecasaccharides for binding activity studies against BACE-1. In this case, the disaccharides were synthesized from more reactive glucose and idose units followed by oxidation to the corresponding uronic acid only in the later stages to avoid the poor reliability of carboxylic acid containing donors[153]. There have been other approaches based on the disaccharide core structure for the design of heparin scaffolds[154-157].

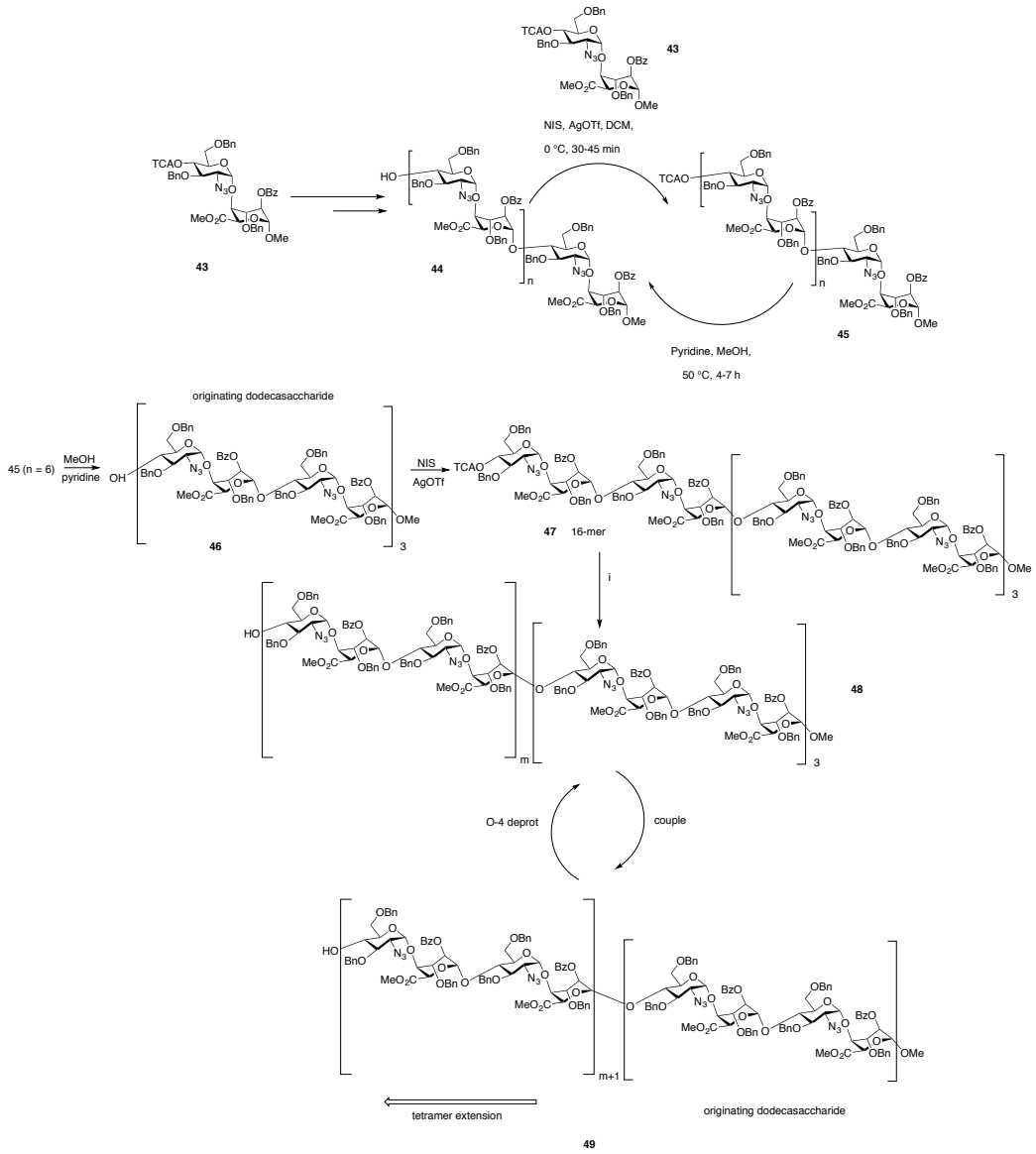


Scheme 3: Disaccharide building blocks used for synthesizing heparin oligosaccharides (**25-40**) and the monosaccharide synthons (**17-24**) for construction of the disaccharide blocks as reported by Arungundram et al.[152] The synthesis of iduronic acid from 1,6 anhydro-idose is shown in the inset.

A recent report by Hansen et al. has demonstrated the synthesis of heparin like oligosaccharides upto a length of 20 monosaccharide units for the first time[158]. This strategy is based on the coupling of tetrasaccharide units, which was repeated for the required number of times to assemble the required length of the target oligosaccharide. Protected

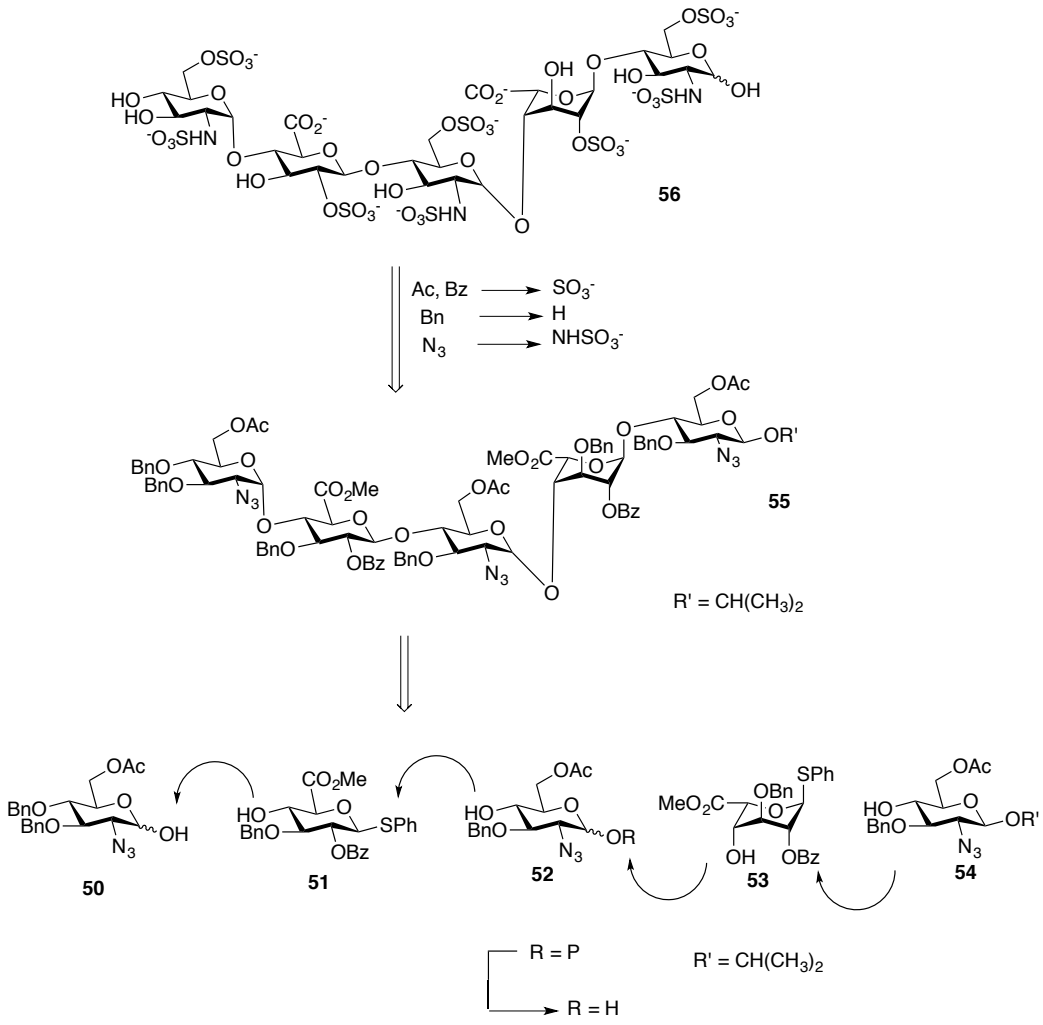
oligosaccharide synthesis by the iterative addition of the tetrasaccharide block has been successfully utilized to synthesize oligosaccharides upto a length of 40 carbohydrate units (**Scheme 4**). Deprotection and sulfation at selected hydroxyl and amino groups then yielded the *N*- and *O*- sulfated product in good yields. Although the strategy is capable of producing larger sugars, it is a complex time consuming methodology involving iterative synthesis. Nevertheless, this approach represents a useful methodology for the construction of challenging heparin and heparan sulfate units of larger sizes.

A notable exception to the iterative use of core disaccharide or tetrasaccharide building blocks was the work reported by Codée et al.[159] This strategy is based on monomeric building blocks to synthesize a protected heparin pentasaccharide. It utilized 1-thio uronic acid synthons, thus requiring the use of potent electrophilic activators in the form of Ph_2SO (diphenyl sulfoxide)/Tf_2O (triflic anhydride) and BSP (1-benzene sulfonyl piperidine)/Tf_2O. These were reportedly used for the first time in heparin and HS synthesis. The 1-thio uronic acid synthons were reacted with other monosaccharides as shown in **Scheme 5** to synthesize the target pentasaccharide. Even though this approach is amenable for the synthesis of pentasaccharides, it would require highly complex strategies to synthesize appropriate monomer units for building larger GAG chains.



Scheme 4: $[4]_n$ repeating strategy as reported by Hansen et al.[158] The tetrasaccharide core **45** (n = 1) was made by an iterative synthesis from the disaccharide **43**. This was further extended to form the originating dodecasaccharide **45 (n = 6)**[160]. The originating dodecasaccharide was then iteratively used for synthesis by *O*-4 deprotection to synthesize oligosaccharides upto 40-mers.

Attempts have also been made to synthesize heparin like oligomers on solid support. Synthesis on the solid phase has several advantages as excess reagent can be easily employed, no need for time consuming chromatographic separations and the reaction can be monitored by NMR at each step easily, to verify the progress without having to purify the mixture by a lengthy purification process.



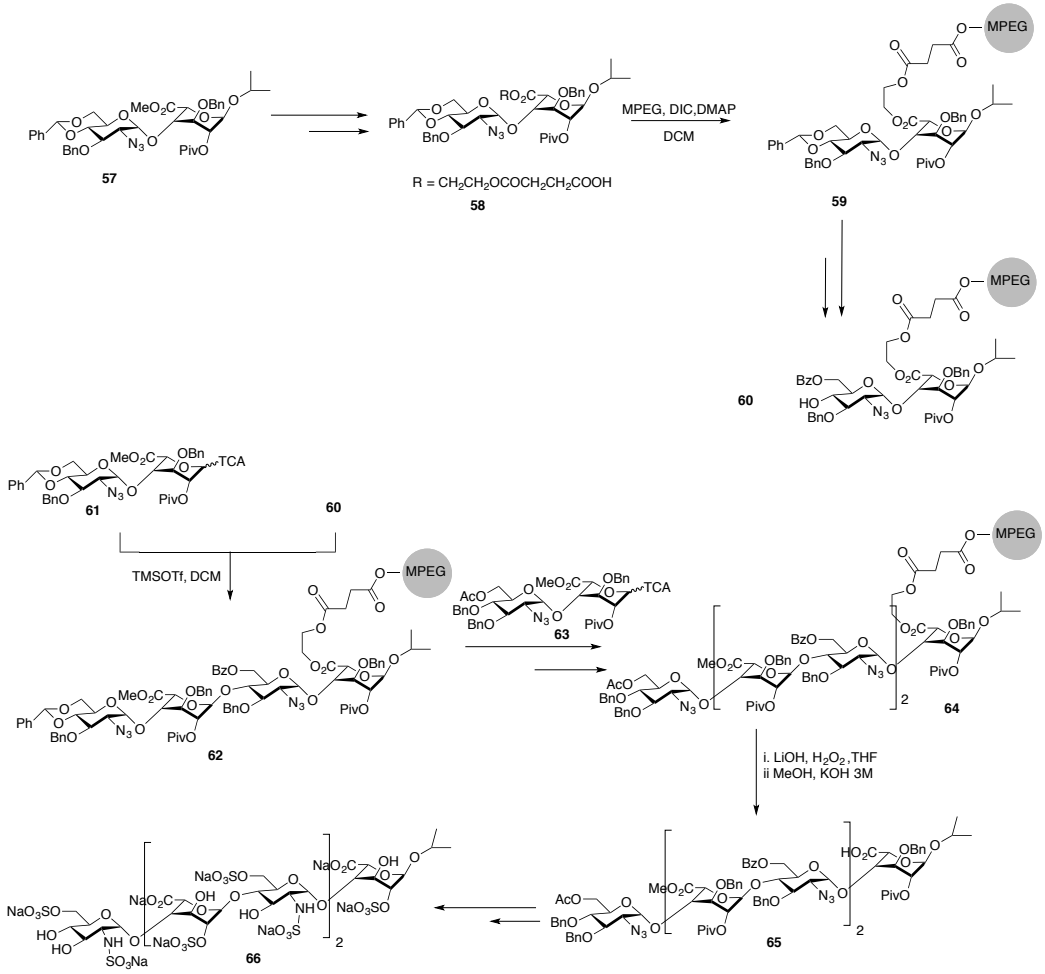
Scheme 5: Retrosynthetic analysis as reported by Codée et al.[159]. The target pentasaccharide **56** is the sequence of heparin to which antithrombin binds. **50-54** represent the monosaccharide synthons used to build up the pentasaccharide **55**. **51** and **53** represent the thio uronic acid synthons requiring use of a potent electrophilic activator. **P** = orthogonal protecting group

Dreef-Tromp et al.[161] reported using simple orthogonally protected disaccharide units (with one of the disaccharides being PEGylated) to prepare oligomers upto a length of 12 units. The PEG polymer-carbohydrate unit is reported to be soluble during glycosylation conditions and was in the solid form during work up[162]. Thus, this allows glycosylations to proceed with similar efficiencies achieved in solution phase reactions. It should be mentioned here that the oligomers did not contain any amino groups, neither was any sulfation or

general deprotection carried out. Therefore, these results are only preliminary and the applicability to more complex systems remains unaddressed.

Another example of solid phase synthesis for the construction of heparin hexasaccharide was reported by Ojeda et al.[163] The solid support in this case was attached to the carboxylic group on the uronic acid at the non-reducing end. The whole disaccharide, of which the uronic acid is a constituent monosaccharide, was then attached to two other appropriately protected disaccharides to ultimately yield the hexasaccharide. Cleavage of the solid support was then carried out followed by deprotection and addition of sulfate groups at appropriate positions (**Scheme 6**).

Heparin mimics in the form of polymer supported disaccharides have been reported to serve as potent GAG mimetics. This approach from the Hsieh-Wilson lab relies on the construction of appropriately sulfated heparin disaccharides. These disaccharides are then attached to a norbornene monomer and are subsequently polymerized using Grubbs catalyst. The resulting glycopolymers were shown to be effective mimetics of heparin against various chemokines, in some cases even surpassing the activity of heparin itself[164]. Effective anticoagulants have also been developed using this approach[165]. Constructing chondroitin sulfate mimetics has further validated the versatility of this approach. These glycopolymers have been shown to mimic the activity of natural chondroitin sulfates in affecting neurite outgrowth[166].



Scheme 6: Synthesis of heparin like oligosaccharides on solid support as reported by Ojeda et al.[163] The core disaccharide structure was added to the solid support by means of a succinyl linker. The chain was thereafter extended on the solid support and at the end the oligosaccharide was cleaved off from the support. Deprotection and installation of sulfate groups were then done at appropriate positions. (MPEG = methyl ether PEG)

In this case, ring-opening metathesis reaction was used with cis-cyclooctene attached chondroitin sulfate disaccharides, as the basis for the polymerization reaction.

The chemical synthesis of other GAGs such as chondroitin sulfate has also been reported. Hsieh-Wilson and coworkers synthesized chondroitin sulfate tetrasaccharides, where a 2+2 retrosynthetic approach was followed to access the target molecules[167-168]. The approach is similar to the one followed for heparin where appropriately protected

disaccharides were glycosylated before deprotection and installation of the sulfate groups. Jacquinet et al. reported a novel synthesis of chondoritin sulfate from commercial chondroitin sulfate. Desulfation during hydrolysis of commercial chondroitin sulfate enabled the synthesis of non-sulfated disaccharides which was then protected orthogonally to synthesize both the donor as well as the acceptor[169]. A hexasaccharide of desulfated chondroitin sulfate was then synthesized using the disaccharide synthons. Automated solid phase synthesis of chondroitin sulfate has been achieved by Seeberger and coworkers using a photolabile linker. A protected and sulfated chondroitin sulfate hexasaccharide was prepared in this way[170].

The synthesis of hyaluronic acid decasaccharide units were reported by Lu et al.[171] The strategy involved using two monosaccharide building blocks, which was ultimately expanded to utilize three building blocks when problems during deprotection of the carbohydrate moieties were encountered.

Even though there have been significant advances in chemical approaches towards the synthesis of heparin/heparan sulfate oligosaccharides, most of the approaches have been tailormade for the target protein binding partner thus precluding a generic approach which can be followed to synthesize complex GAGs. While the disaccharide based core motif has been central to many approaches, multiple protection and deprotection steps make it difficult to synthesize a reasonable length GAG oligosaccharide in a short time.

1.6.2 Chemoenzymatic synthesis of GAGs

Chemoenzymatic synthesis of glycosaminoglycans in general and heparin/heparan sulfate in particular relies on using biosynthetic enzymes for carrying out the synthesis of the

glycosaminoglycan backbone as well as the sulfation and epimerization of the carbohydrate chain. Essentially this approach mimics the synthesis of GAGs in biological systems. Even though this approach is relatively recent, major advances have been made in this field for the synthesis of natural as well as unnatural GAG structures. Since, these enzymes have evolved to provide good regioselectivity and are adept in working on deprotected substrates, no protection/deprotection strategy is required, which significantly decreases the number of steps. Also, because the enzymes have high specificities, the likelihood of generating byproducts is reduced with this approach, which allows to cut down on several purification steps[172].

The K5 strain of *Escherichia coli* (*E. coli)* is known to produce heparosan as part of its capsular polysaccharide. Heparosan is a long chain polysaccharide consisting of alternate units of *N*-acetylglucosamine and glucuronic acid having the same linkage stereochemistry as found in natural heparin/heparan sulfate [-(4)βGlcA(1)→(4)αGlcNAc(1)-]$_n$[173-174]. It is the precursor of heparin before epimerization and sulfation takes place. This backbone can be used as a template for the enzymatic synthesis of heparin-like polysaccharides. Kuberan et al. have demonstrated the use of heparosan as a starting material for the subsequent synthesis of heparin oligosaccharides using a combination of different enzymes[175]. Naturally available heparosan was partially depolymerized using lyase enzymes. Thereafter, it was subjected to base-assisted *N*-deacetylation followed by sulfation of the resulting free amino groups using trimethylamine sulfur trioxide complex. Epimerization of the uronic acids was then undertaken using C-5 epimerase. The activity of C-5 epimerase and a lot of other enzymes are dependant on *N*-deacetylation and *N*-sulfation of the polysaccharide. Hence, it was necessary to perform the step of *N*-deacetylation-*N*-sulfation initially. Epimerization was subsequently carried out only on residues located at the reducing side of *N*-sulfated glucosamine. Moreover, uronic acids of the substrate had to be devoid of *O*-sulfation and

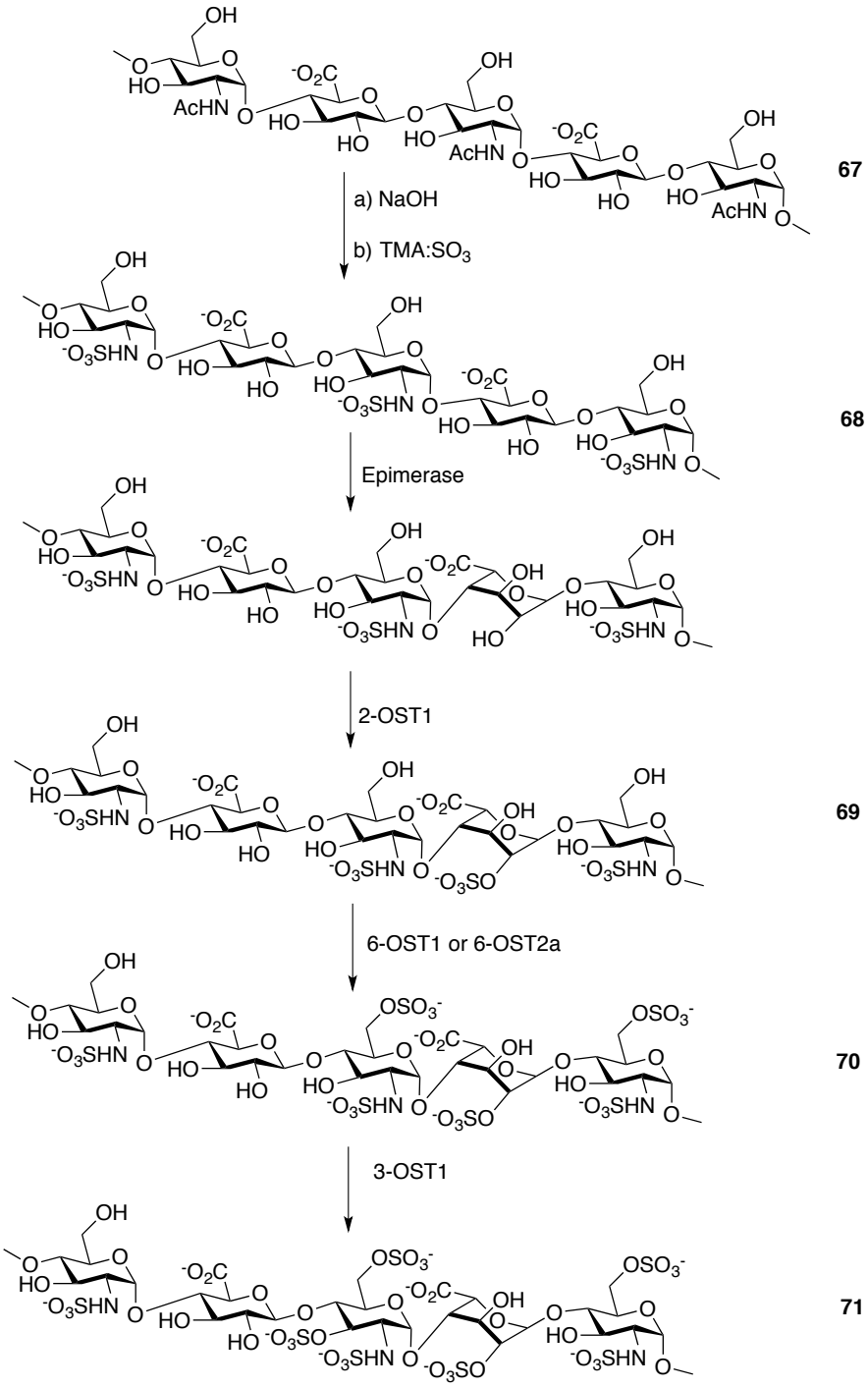
must not be adjacent to *O*-sulfated glucosamine residues for them to be epimerized by C-5 epimerase[173, 176]. Thus, the use of a judicious sequence of enzymes had to be followed to prepare a substrate with given sulfation and iduronic acid distribution. Next, the 2-*O*-sulfotransferase, 6-*O*-transferase and 3-*O*-sulfotransferase enzymes were used in tandem to sulfate the hydroxyl groups of C-2 of the uronic acid, and C-6 and C-3 of the amino sugar of selected monosaccharide residues in the carbohydrate chain (**Scheme 7**). The sulfotransferases use PAPs as the sulfate source. Many of these enzymes have got different isoforms and some of them, especially isoforms of 3-*O*-sulfotransferases, have been shown to have different substrate specifities. This can be used to our advantage to prepare tailored oligosaccharides designed to bind selectively to one type of heparin binding protein.

Even though this approach proved to be radically new to produce GAG oligosaccharides, structurally heterogeneous products with different oligosaccharide size and sulfation patterns were obtained. This also leads to reduced yields of the desired product. The causative factors for this had been speculated to be

(i) incomplete conversion of starting material to product during each enzymatic step and

(ii) structural heterogeneity of the starting material itself[172].

In order to circumvent the second problem, a defined hexasaccharide starting material was used in another study, which was then enzymatically modified[177].

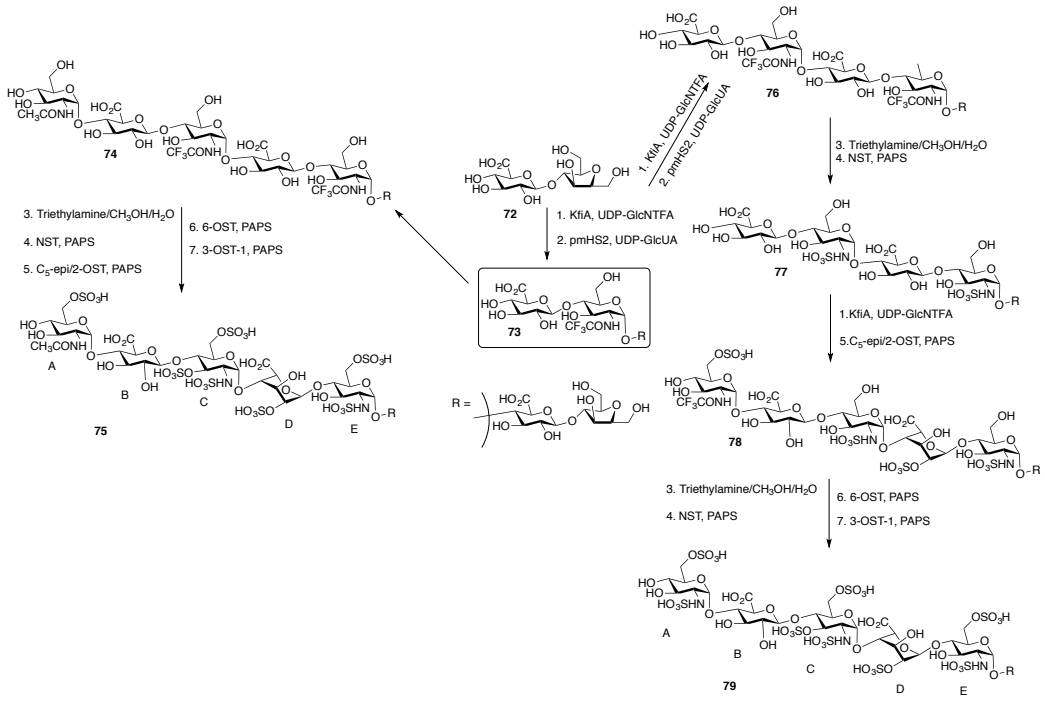


Scheme 7: Chemoenzymatic synthesis of anticoagulant heparin oligosaccharide as reported by Kuberan et al.[175] The synthesis was done with a heparosan backbone **67**.

Recently, Liu and coworkers have demonstrated an improved enzymatic synthesis of heparin pentasaccharides capable of binding to antithrombin[178] (**Scheme 8**). The synthesis of

heparin oligosaccharide was started from a disaccharide obtained by exhaustive degradation of heparosan using nitrous acid followed by reduction with sodium borohydride[179-180].

The carbohydrate backbone chain was extended using a combination of a *N*-acetylglucosaminyltransferase, in the form of K5 *N*-acetylglucosaminyltransferase (KfiA) and *Pasteurella multocida* heparosan synthase 2 (pmHS2) as a glucuronyltransferase. While KfiA is a monofunctional glycosyltransferase involved in the construction of the capsular polysaccharide of K5 strain of *E. coli*[181], pmHS2 is a bifunctional glycosyltransferase responsible for construction of heparosan structures in the cell wall of *Pasteurella multocida*[182]. Thus, pmHS2 can also transfer *N*-acetylglucosamine to an oligosaccharide. KfiA has been used in this work to transfer GlcNTFA (*N*-trifluoroacetylglucosamine) instead of GlcNAc. Being structurally similar to GlcNAc, GlcNTFA can be used by KfiA as a substrate. Moreover, the trifluoroacetyl protecting group can be removed after incorporation into the oligosaccharide by treating with triethylamine in the presence of water and methanol, so that the resulting free amino group can be sulfated[183]. After extension of the backbone chain, the deprotected amino group was sulfated using *N*-sulfotransferase[184], followed by epimerization of selected residues to form iduronic acid, 2-*O*-sulfation and finally 6-*O*-sulfation. The specifities of the different modifying enzymes were exploited to install the sulfation and have epimerizations at only selected positions. The authors also ensured that the reaction proceeds to completion at each step of the enzymatic reactions. Following the approach, they were able to synthesize multi milligram quantities of pure biologically active heparin heptasaccharides.



Scheme 8: Chemoenzymatic synthesis of two heparin oligosaccharides **75** and **79** as reported by Liu et al.[178] Synthesis was initiated from the heparosan derived disaccharide **72.**

Chondroitin sulfates have also been synthesized chemoenzymatically. In the works reported by Kobayashi and coworkers[185-186], oxazoline containing non-sulfated and sulfated disaccharide building blocks were used for enzymatic transglycosylation to synthesize both non-sulfated as well as 4-*O*-sulfated chondroitin sulfate using a hyaluronidase. Even though hyaluronidase is an endo-β-*N*-acetylhexosaminidase responsible for hydrolyzing glycosidic bonds[187-188], it has also been found to catalyse transglycosylation. Substrates upto a chain length of 74 monosaccharide units were prepared. However, 6-*O*-sulfated and 4,6-*O*-sulfated monomeric substrates could not be transformed by the enzyme in this reaction. Later, the same group demonstrated the use of this enzyme for synthesis of various natural and unnatural GAGs[189].

Enzymatic synthesis of chondroitin sulfate has also been achieved by using chondroitin synthase enzyme (K4CP), expressed by *E. coli* K4[190]. Since K4CP is a bifunctional glycosyltransferase[191] just as pmHS2 mentioned earlier, it is capable of transferring both glucuronic acid as well as GalNAc. Two mutant variants of the enzyme were prepared, featuring either glucuronyl transfer or *N*-acetylgalactosaminyl transfer selectively. These mutants were then immobilized for ease of purification of products and were used in tandem to construct GAG chains upto a length of 16 sugar units.

A similar approach has also been followed for the enzymatic synthesis of hyaluronan oligosaccharides, where bifunctional *Pasteurella multocida* HA synthase, pmHAS has been converted into two single functional mutants capable of transferring either *N*-acetylglucosamine or glucuronic acid. After immobilization of these single action mutants on solid support, pure hyaluronan oligosaccharides were synthesized from the corresponding UDP sugars[192].

Nevertheless, these promising approaches for the enzymatic synthesis of GAG structures notwithstanding, enzymatic synthesis of GAG oligosaccharides do have certain disadvantages, notably:

(i) Reactions can be done only in water, which is difficult to remove from the reaction mixture.

(ii) Scalability is a problem for most enzymatic reactions.

(iii) It is generally more expensive to carry out enzymatic glycosylations since the cost of expressing enzymes is high.

(iv) In addition, enzymes are not stable for long periods, thus they are not the most robust glycosylating reagents.

1.7 Thiol-ene radical reaction in synthesis

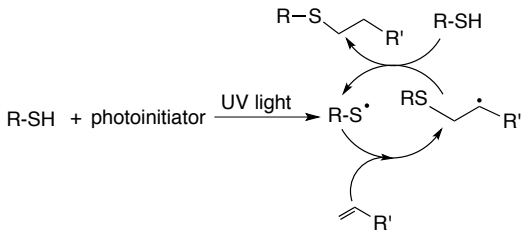
Thiol-ene chemistry is a highly efficient reaction of thiols with carbon-carbon double bonds. It is a well-known reaction from the early 1900s[193] and can be broadly described under two headings:

1. Free-radical addition to electron–rich/electron-poor carbon-carbon double bonds.

2. Catalyzed thio-Michael addition to carbon-carbon double bonds using a variety of catalysts ranging from strong bases, metals and organometallics to Lewis acids[193].

Thiol-ene radical reactions can be carried out by irradiating the thiol-ene reaction mixture with UV light in the optional presence of a radical initiator. The reaction proceeds by the generation of thiyl radicals from thiols that is attached by the alkene in an anti-Markownikoff manner[194]. This results in the generation of a carbon radical that abstracts a proton from another thiol to generate another thiyl radical and the process continues. It is represented by the scheme depicted in **Scheme 9**.

Because of characteristics such as quantitative yields, requirement of small amounts of catalysts and faster reaction times, this reaction is appropriately categorized under the heading of click chemistry.



Scheme 9: Mechanism of thiol-ene free radical reaction[194].

The thiol-ene click reaction has long been used in the field of polymer and materials chemistry for the preparation of microfluidic devices[195], nanoimprint lithography[196], hydrogel bio materials[197] etc.

Thiol-ene click reaction gradually percolated into the domain of organic and bio organic chemistry with applications such as modification of synthetic polymers with bioorganic molecules in the form of amino acids, peptides and carbohydrates[198]. Examples include the work carried out by Schlaad and co-workers where they grafted the surface of polybutadiene polymers by amino acids using the ene functionalities from the polymer resulting in amphiphilic substances with greatly altered physical properties[199]. These materials have immense applications as drug delivery agents or biologically relevant vesicles. Other noteworthy works in this domain include chain elongation of allyl glycosides using different thiols to provide sulfide spacer glycosides[200]. Heidecke and Lindhorst prepared glycodendrons that had mannoside fragments attached by means of thio ether linkages accessible through thiol-ene click reactions[201]. The glycodendrons were also tested for potential activity against bacterial adhesion. Specifically, it was shown that one of these synthesized glycodendrons ranked among the best inhibitors of the type 1 fimbriae-mediated adhesion of *E. coli.*

Work has previously been carried out in our group and in Dondoni's group towards the coupling of sugar thiols with amino acids in order to prepare thio-linked glycopeptides. While Dondoni's group demonstrated the thiol-ene click reaction between peracetylated lactosyl thiols and *N*-protected allyl and vinyl glycinates[202,203], our group focused on the synthesis of *S*-glycosyl amino acids through addition of glycosyl thiols to homoallylglycine[204]. Interestingly, this work also demonstrated how the thiol-ene click reaction can be carried out in water using Vazo-44, a water-soluble photo initiator. In fact,

this increases the scope of thiol-ene click reaction significantly, in a bioorganic setting as reactions can be carried out with peptides and carbohydrates in a benign solvent, water.

Kunz and co-workers showed a remarkable application of this reaction when they synthesized vaccines targeted against tumor associated glycopeptide antigens[205]. The thiol-ene click reaction was used to introduce preformed glycopeptides into a carrier protein BSA (Bovine Serum Albumin) to yield the BSA-glycopeptide vaccines. The thiol-ene click reaction did not lead to racemization of amino acids, in the reaction, as demonstrated by Waldmann and co-workers when they synthesized various S-alkylated cysteines using AIBN initiation[206].

During this work, investigations included suitability tests of common sugar protecting groups in the thiol-ene click reaction. In the course of a work aimed at demonstrating the utility of *N*-pentenyl group as a convenient precursor for various spacers to study biological and other activities of carbohydrates, it has been shown that the yield of the thio linked *N*-pentenyl glucoside product is more when the sugar is acetylated or unprotected compared to benzylated sugars[207]. The probable reason for this low yield was attributed to radical side reactions of the benzylic carbon.

1.8 Aim of the present work

Considering the advantages of thiol-ene click reaction, it is of great advantage if the said reaction can be utilized for the construction of heparin and heparan sulfate oligosaccharides in particular and GAGs in general. The unsaturation produced on the uronic acid at the non-reducing end, when GAGS are broken down by lyase enzymes, can be

utilized for this purpose. In this work we propose a thiol-ene click approach for the synthesis of heparin and heparan sulfate oligosaccharide mimics in a modular way.

This strategy utilizes oligosaccharides of heparin and heparan sulfate obtained by enzymatic break down of naturally available heterogeneous heparin and heparan sulfate, as building blocks. After enzymatic breakdown of heterogeneous heparin and heparan sulfate chains, they can be purified by size exclusion and strong anion exchange chromatography to yield defined oligosaccharides. The anomeric position of these GAG building blocks can be thiolated potentially, followed by addition to the unsaturation at the non-reducing end of another unit of oligosaccharide, produced by lyase degradation. This approach can therefore potentially be used to build defined heparin and heparan sulfate thio-mimics of any oligosaccharide length. Also it can be utilized for the synthesis of thio-mimics of other GAGs where an unsaturation is produced by the action of the corresponding GAG lyase.

Accordingly, in chapter 2, we talk about using thiol-ene radical reaction for the addition of a thiosugar to an unsaturated monosaccharide for small model substrates as part of an optimization process. Apart from protected substrates, the reaction has also been optimized for deprotected substrates in aqueous conditions. This sets the stage for the utilization of sulfated heparin oligosaccharides in aqueous conditons for the thiol-ene coupling reaction.

In chapter 3, we lay out the strategy for the synthesis of heparin and heparan sulfate thio mimics from building blocks derived by lyase degradation of naturally available heparin. The synthesis of these building blocks from heparin is discussed. The synthesis of simple thio heparin and heparan sulfate mimics utilizing these building blocks has also been

described. Finally, biological studies with these synthesized thio mimics of heparin have been discussed.

In chapter 4, the use of a glucuronoyltransferase pmHS2 and a *N*-acetylglucosaminyltransferase for further enzymatic extension of the thio-mimics have been described. Specifically, the enzymes have been used on commercially available and readily synthesized substrates to demonstrate the applicability of this approach for GAG extension. Potentially, the approach can be coupled with the thiol-ene radical strategy to synthesize novel thio GAG mimics as well as proteoglycan mimics. A plausible strategy for the design and synthesis of such proteoglycans using both the radical thiol-ene reaction and consequent enzymatic extension has been laid out.

1.9 References

1. Antonio, J. D. S.; Iozzo, R. V. Glycosaminoglycans: Structure and Biological Functions. In *eLS*, John Wiley & Sons, Ltd: **2001**.

2. Esko, J. D.; Kimata, K.; Lindahl, U. Proteoglycans and Sulfated Glycosaminoglycans. In *Essentials of Glycobiology*, 2nd Edition ed.; Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E., Eds. Cold Spring Harbour Laboratory Press: New York, **2009**.

3. DeAngelis, P. L. Evolution of glycosaminoglycans and their glycosyltransferases: Implications for the extracellular matrices of animals and the capsules of pathogenic bacteria. *Anat Rec* **2002,** *268* (3), 317-326.

4. Taylor, K. R.; Gallo, R. L. Glycosaminoglycans and their proteoglycans: host-associated molecular patterns for initiation and modulation of inflammation. *FASEB J*. **2006,** *20* (1), 9-22.

5. Gandhi, N. S.; Mancera, R. L. The Structure of Glycosaminoglycans and their Interactions with Proteins. *Chem Biol Drug Des* **2008,** *72* (6), 455-482.

6. Raman, R.; Sasisekharan, V.; Sasisekharan, R. Structural insights into biological roles of protein-glycosaminoglycan interactions. *Chem Biol* **2005,** *12* (3), 267-277.

7. Sasisekharan, R.; Venkataraman, G. Heparin and heparan sulfate: biosynthesis, structure and function. *Curr. Opin. Chem. Biol.* **2000,** *4* (6), 626-631.

8. Bourin, M. C.; Lindahl, U. Glycosaminoglycans and the Regulation of Blood-Coagulation. *Biochem. J* **1993,** *289*, 313-330.

9. Ariga, T.; Miyatake, T.; Yu, R. K. Role of proteoglycans and glycosaminoglycans in the pathogenesis of Alzheimer's disease and related disorders: Amyloidogenesis and therapeutic strategies—A review. *J. Neurosci. Res.* **2010,** *88* (11), 2303-2315.

10. Kano, K.; Miyano, T.; Kato, S. Effects of glycosaminoglycans on the development of in vitro matured and fertilized porcine oocytes to the blastocyst stage in vitro. *Biol Reprod* **1998,** *58* (5), 1226-1232.

11. Schnabelrauch, M.; Scharnweber, D.; Schiller, J. Sulfated Glycosaminoglycans As Promising Artificial Extracellular Matrix Components to Improve the Regeneration of Tissues. *Curr Med Chem* **2013,** *20* (20), 2501-2523.

12. Salanga, C. L.; Handel, T. M. Chemokine oligomerization and interactions with receptors and glycosaminoglycans: The role of structural dynamics in function. *Exp Cell Res* **2011,** *317* (5), 590-601.

13. Middleton, J.; Neil, S.; Wintle, J.; ClarkLewis, I.; Moore, H.; Lam, C.; Auer, M.; Hub, E.; Rot, A. Transcytosis and surface presentation of IL-8 by venular endothelial cells. *Cell* **1997,** *91* (3), 385-395.

14. Middleton, J.; Patterson, A. M.; Gardner, L.; Schmutz, C.; Ashton, B. A. Leukocyte extravasation: chemokine transport and presentation by the endothelium. *Blood* **2002,** *100* (12), 3853-3860.

15. Wang, L. C.; Fuster, M.; Sriramarao, P.; Esko, J. D. Endothelial heparan sulfate deficiency impairs L-selectin- and chemokine-mediated neutrophil trafficking during inflammatory responses. *Nat Immunol* **2005,** *6* (9), 902-910.

16. Bishop, J. R.; Schuksz, M.; Esko, J. D. Heparan sulphate proteoglycans fine-tune mammalian physiology. *Nature* **2007,** *446* (7139), 1030-1037.

17. Luster, A. D.; Greenberg, S. M.; Leder, P. The Ip-10 Chemokine Binds to a Specific Cell-Surface Heparan-Sulfate Site Shared with Platelet Factor-4 and Inhibits Endothelial-Cell Proliferation. *J Exp Med* **1995,** *182* (1), 219-231.

18. Webb, L. M. C.; Ehrengruber, M. U.; Clarklewis, I.; Baggiolini, M.; Rot, A. Binding to Heparan-Sulfate or Heparin Enhances Neutrophil Responses to Interleukin-8. *Proc Natl Acad Sci USA* **1993,** *90* (15), 7158-7162.

19. Kreuger, J.; Salmivirta, M.; Sturiale, L.; Gimenez-Gallego, G.; Lindahl, U. Sequence analysis of heparan sulfate epitopes with graded affinities for fibroblast growth factors 1 and 2. *J. Biol. Chem.* **2001,** *276* (33), 30744-30752.

20. Teran, M.; Nugent, M. A. Synergistic Binding of Vascular Endothelial Growth Factor-A and Its Receptors to Heparin Selectively Modulates Complex Affinity. *J. Biol. Chem.* **2015,** *290* (26), 16451-16462.

21. Cecchi, F.; Pajalunga, D.; Fowler, C. A.; Uren, A.; Rabe, D. C.; Peruzzi, B.; MacDonald, N. J.; Blackman, D. K.; Stahl, S. J.; Byrd, R. A.; Bottaro, D. P. Targeted

Disruption of Heparan Sulfate Interaction with Hepatocyte and Vascular Endothelial Growth Factors Blocks Normal and Oncogenic Signaling. *Cancer Cell* **2012,** *22* (2), 250-262.

22. Majack, R. A.; Cook, S. C.; Bornstein, P. Platelet-Derived Growth-Factor and Heparin-Like Glycosaminoglycans Regulate Thrombospondin Synthesis and Deposition in the Matrix by Smooth-Muscle Cells. *J. Cell Biol.* **1985,** *101* (3), 1059-1070.

23. Kim, S. H.; Turnbull, J.; Guimond, S. Extracellular matrix and cell signalling: the dynamic cooperation of integrin, proteoglycan and growth factor receptor. *J. Endocrinol.* **2011,** *209* (2), 139-151.

24. Sasisekharan, R.; Raman, R.; Prabhakar, V. Glycomics approach to structure-function relationships of glycosaminoglycans. *Annu Rev Biomed Eng* **2006,** *8*, 181-231.

25. Lever, R.; Page, C. P. Novel drug development opportunities for heparin. *Nat Rev Drug Discov* **2002,** *1* (2), 140-148.

26. Francis, C. W.; Kaplan, K. L. Principles of Antithrombotic Therapy. In *Williams Hematology*, 7th ed.; Lichtman, M. A.; Beutler, E.; Kipps, T. J., Eds. **2006**.

27. Charles, E.; Carraher, J. *Carraher's Polymer Chemistry*. 9th ed.; CRC Press: **2013**.

28. Sugahara, K.; Kitagawa, H. Heparin and heparan sulfate biosynthesis. *Iubmb Life* **2002,** *54* (4), 163-175.

29. Silbert, J. E.; Sugumaran, G. Biosynthesis of chondroitin/dermatan sulfate. *Iubmb Life* **2002,** *54* (4), 177-186.

30. Prabhakar, V.; Raman, R.; Capila, I.; Bosques, C. J.; Pojasek, K.; Sasisekharan, R. Biochemical characterization of the chondroitinase ABC I active site. *Biochem. J* **2005,** *390*, 395-405.

31. Habuchi, O. Sulfotransferases Involved in Sulfation of Glycosaminoglycans. In *Experimental Glycoscience*, Taniguchi, N.; Suzuki, A.; Ito, Y.; Narimatsu, H.; Kawasaki, T.; Hase, S., Eds. Springer Japan: **2008**; pp 87-93.

32. Feyerabend, T. B.; Li, J. P.; Lindahl, U.; Rodewald, H. R. Heparan sulfate C5-epimerase is essential for heparin biosynthesis in mast cells. *Nat Chem Biol* **2006,** *2* (4), 195-196.

33. Sugumaran, G.; Vertel, B. M. Biosynthesis of chondroitin sulfate and dermatan sulfate proteoglycans. In *Oligosaccharides in Chemistry and Biology: A Comprehensive Handbook*, Ernst, B.; Hart, G.; Sinay, P., Eds. Wiley VCH: **2000**; pp 375-394.

34. Bourdon, M. A.; Krusius, T.; Campbell, S.; Schwartz, N. B.; Ruoslahti, E. Identification and Synthesis of a Recognition Signal for the Attachment of Glycosaminoglycans to Proteins. *Proc Natl Acad Sci USA* **1987,** *84* (10), 3194-3198.

35. Huber, S.; Winterhalter, K. H.; Vaughan, L. Isolation and Sequence-Analysis of the Glycosaminoglycan Attachment Site of Type-IX Collagen. *J. Biol. Chem.* **1988,** *263* (2), 752-756.

36. Zimmermann, D. R.; Ruoslahti, E. Multiple domains of the large fibroblast proteoglycan, versican. *The EMBO Journal* **1989,** *8* (10), 2975-2981.

37. Noonan, D. M.; Horigan, E.; Ledbetter, S.; Peterson, T.; Vogeli, G.; Sasaki, M.; Yamada, Y.; Hassell, J. R. Structure of the Basement-Membrane Heparan-Sulfate Proteoglycan Core Protein Deduced from cDNA Clones. *Collagen Rel Res* **1988,** *8* (6), 477-478.

38. Rauch, U.; Feng, K.; Zhou, X. H. Neurocan: a brain chondroitin sulfate proteoglycan. *Cell Mol Life Sci* **2001,** *58* (12-13), 1842-1856.

39. Yamada, H.; Fredette, B.; Shitara, K.; Hagihara, K.; Miura, R.; Ranscht, B.; Stallcup, W. B.; Yamaguchi, Y. The brain chondroitin sulfate proteoglycan brevican associates with astrocytes ensheathing cerebellar glomeruli and inhibits neurite outgrowth from granule neurons. *J. Neurosci.* **1997,** *17* (20), 7784-7795.

40. Jones, L. L.; Margolis, R. U.; Tuszynski, M. H. The chondroitin sulfate proteoglycans neurocan, brevican, phosphacan, and versican are differentially regulated following spinal cord injury. *Exp Neurol* **2003,** *182* (2), 399-411.

41. Yamaguchi, Y.; Mann, D. M.; Ruoslahti, E. Negative Regulation of Transforming Growth-Factor-Beta by the Proteoglycan Decorin. *Nature* **1990,** *346* (6281), 281-284.

42. Doege, K. J.; Sasaki, M.; Kimura, T.; Yamada, Y. Complete Coding Sequence and Deduced Primary Structure of the Human Cartilage Large Aggregating Proteoglycan, Aggrecan - Human-Specific Repeats, and Additional Alternatively Spliced Forms. *J. Biol. Chem.* **1991,** *266* (2), 894-902.

43. Rodén, L. Structure and metabolism of connective tissue proteoglycans. In *The Biochemistry of Glycoproteins and Proteoglycans*, Lennarz, W. J., Ed. Plenum Publishers New York, **1980**; pp 269-314.

44. Okajima, T.; Yoshida, K.; Kondo, T.; Furukawa, K. Human homolog of Caenorhabditis elegans sqv-3 gene is galactosyltransferase I involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. *J. Biol. Chem.* **1999,** *274* (33), 22915-22918.

45. Almeida, R.; Levery, S. B.; Mandel, U.; Kresse, H.; Schwientek, T.; Bennett, E. P.; Clausen, H. Cloning and expression of a proteoglycan UDP-galactose :beta-xylose beta 1,4-galactosyltransferase I - A seventh member of the human beta 4-galactosyltransferase gene family. *J. Biol. Chem.* **1999,** *274* (37), 26165-26171.

46. Bai, X. M.; Zhou, D. P.; Brown, J. R.; Crawford, B. E.; Hennet, T.; Esko, J. D. Biosynthesis of the linkage region of glycosaminoglycans - Cloning and activity of galactosyltransferase II, the sixth member of the beta 1,3-galactosyltransferase family (beta 3GalT6). *J. Biol. Chem.* **2001,** *276* (51), 48189-48195.

47. Kitagawa, H.; Tone, Y.; Tamura, J.-i.; Neumann, K. W.; Ogawa, T.; Oka, S.; Kawasaki, T.; Sugahara, K. Molecular Cloning and Expression of Glucuronyltransferase I Involved in the Biosynthesis of the Glycosaminoglycan-Protein Linkage Region of Proteoglycans. *J. Biol. Chem.* **1998,** *273* (12), 6615-6618.

48. Esko, J. D.; Zhang, L. J. Influence of core protein sequence on glycosaminoglycan assembly. *Curr Opin Struc Biol* **1996,** *6* (5), 663-670.

49. Kitagawa, H.; Uyama, T.; Sugahara, K. Molecular cloning and expression of a human chondroitin synthase. *J. Biol. Chem.* **2001,** *276* (42), 38721-38726.

50. Uyama, T.; Kitagawa, H.; Tamura, J.; Sugahara, K. Molecular cloning and expression of human chondroitin N-acetylgalactosaminyltransferase - The key enzyme for chain initiation and elongation of chondroitin/dermatan sulfate on the protein linkage region tetrasaccharide shared by heparin/heparan sulfate. *J. Biol. Chem.* **2002,** *277* (11), 8841-8846.

51. Fukuta, M.; Uchimura, K.; Nakashima, K.; Kato, M.; Kimata, K.; Shinomura, T.; Habuchi, O. Molecular-Cloning and Expression of Chick Chondrocyte Chondroitin 6-Sulfotransferase. *J. Biol. Chem.* **1995,** *270* (31), 18575-18580.

52. Yamauchi, S.; Mita, S.; Matsubara, T.; Fukuta, M.; Habuchi, H.; Kimata, K.; Habuchi, O. Molecular cloning and expression of chondroitin 4-sulfotransferase. *J. Biol. Chem.* **2000,** *275* (12), 8975-8981.

53. Kobayashi, M.; Sugumaran, G.; Liu, J. A.; Shworak, N. W.; Silbert, J. E.; Rosenberg, R. D. Molecular cloning and characterization of a human uronyl 2-sulfotransferase that sulfates iduronyl and glucuronyl residues in dermatan chondroitin sulfate. *J. Biol. Chem.* **1999,** *274* (15), 10474-10480.

54. Maccarana, M.; Olander, B.; Malmstrom, J.; Tiedemann, K.; Aebersold, R.; Lindahl, U.; Li, J. P.; Malmstrom, A. Biosynthesis of dermatan sulfate - Chondroitin-glucuronate C5-epimerase is identical to SART2. *J. Biol. Chem.* **2006,** *281* (17), 11560-11568.

55. Sugahara, K.; Kitagawa, H. Recent advances in the study of the biosynthesis and functions of sulfated glycosaminoglycans. *Curr Opin Struc Biol* **2000,** *10* (5), 518-527.

56. Fritz, T. A.; Gabb, M. M.; Wei, G.; Esko, J. D. 2 N-Acetylglucosaminyltransferases Catalyze the Biosynthesis of Heparan-Sulfate. *J. Biol. Chem.* **1994,** *269* (46), 28809-28814.

57. Zhang, L. J.; David, G.; Esko, J. D. Repetitive Ser-Gly Sequences Enhance Heparan-Sulfate Assembly in Proteoglycans. *J. Biol. Chem.* **1995,** *270* (45), 27127-27135.

58. Dolan, M.; Horchar, T.; Rigatti, B.; Hassell, J. R. Identification of sites in domain I of perlecan that regulate heparan sulfate synthesis. *J. Biol. Chem.* **1997,** *272* (7), 4316-4322.

59. Zhang, L. J.; Esko, J. D. Amino-Acid Determinants That Drive Heparan-Sulfate Assembly in a Proteoglycan. *J. Biol. Chem.* **1994,** *269* (30), 19295-19299.

60. Kitagawa, H.; Shimakawa, H.; Sugahara, K. The tumor suppressor EXT-like gene EXTL2 encodes an alpha 1, 4-N-acetylhexosaminyltransferase that transfers N-acetylgalactosamine and N-acetylglucosamine to the common glycosaminoglycan-protein linkage region - The key enzyme for the chain initiation of heparan sulfate. *J. Biol. Chem.* **1999,** *274* (20), 13933-13937.

61. Kim, B. T.; Kitagawa, H.; Tamura, J.; Saito, T.; Kusche-Gullberg, M.; Lindahl, U.; Sugahara, K. Human tumor suppressor EXT gene family members EXTL1 and EXTL3 encode alpha 1,4-N-acetylglucosaminyltransferases that likely are involved in heparan sulfate/heparin biosynthesis. *Proc Natl Acad Sci USA* **2001,** *98* (13), 7176-7181.

62. McCormick, C.; Duncan, G.; Goutsos, K. T.; Tufaro, F. The putative tumor suppressors EXT1 and EXT2 form a stable complex that accumulates in the Golgi apparatus and catalyzes the synthesis of heparan sulfate. *Proc Natl Acad Sci USA* **2000,** *97* (2), 668-673.

63. Lind, T.; Tufaro, F.; McCormick, C.; Lindahl, U.; Lidholt, K. The putative tumor suppressors EXT1 and EXT2 are glycosyltransferases required for the biosynthesis of heparan sulfate. *J. Biol. Chem.* **1998,** *273* (41), 26265-26268.

64. Senay, C.; Lind, T.; Muguruma, K.; Tone, Y.; Kitagawa, H.; Sugahara, K.; Lidholt, K.; Lindahl, U.; Kusche-Gullberg, M. The EXT1/EXT2 tumor suppressors: catalytic activities and role in heparan sulfate biosynthesis. *Embo Rep* **2000,** *1* (3), 282-286.

65. Kjellen, L. Glucosaminyl N-deacetylase/N-sulphotransferases in heparan sulphate biosynthesis and biology. *Biochem. Soc. Trans.* **2003,** *31*, 340-342.

66. Crawford, B. E.; Olson, S. K.; Esko, J. D.; Pinhal, M. A. S. Cloning, Golgi localization, and enzyme activity of the full-length heparin/heparan sulfate-glucuronic acid C5-epimerase. *J. Biol. Chem.* **2001,** *276* (24), 21538-21543.

67. Pinhal, M. A. S.; Smith, B.; Olson, S.; Aikawa, J.; Kimata, K.; Esko, J. D. Enzyme interactions in heparan sulfate biosynthesis: Uronosyl 5-epimerase and 2-O-sulfotransferase interact in vivo. *Proc Natl Acad Sci USA* **2001,** *98* (23), 12984-12989.

68. Rong, J. H.; Habuchi, H.; Kimata, K.; Lindahl, U.; Kusche-Gullberg, M. Substrate specificity of the heparan sulfate hexuronic acid 2-O-sulfotransferase. *Biochemistry* **2001,** *40* (18), 5548-5555.

69. Liu, J.; Shworak, N. W.; Fritze, L. M. S.; Edelberg, J. M.; Rosenberg, R. D. Purification of heparan sulfate D-glucosaminyl 3-O-sulfotransferase. *J. Biol. Chem.* **1996,** *271* (43), 27072-27082.

70. Habuchi, H.; Habuchi, O.; Kimata, K. Purification and Characterization of Heparan-Sulfate 6-Sulfotransferase from the Culture-Medium of Chinese-Hamster Ovary Cells. *J. Biol. Chem.* **1995,** *270* (8), 4172-4179.

71. Kusche-Gullberg, M.; Kjellen, L. Sulfotransferases in glycosaminoglycan biosynthesis. *Curr Opin Struc Biol* **2003,** *13* (5), 605-611.

72. Funderburgh, J. L. Keratan sulfate biosynthesis. *Iubmb Life* **2002,** *54* (4), 187-194.

73. Marin, F.; Luquet, G.; Marie, B.; Medakovic, D. Molluscan Shell Proteins: Primary Structure, Origin, and Evolution. In *Current Topics in Developmental Biology*, Academic Press: **2007**; Vol. Volume 80, pp 209-276.

74. Hang, H. C.; Bertozzi, C. R. The chemistry and biology of mucin-type O-linked glycosylation. *Biorg. Med. Chem.* **2005,** *13* (17), 5021-5034.

75. Jensen, P. H.; Kolarich, D.; Packer, N. H. Mucin-type O-glycosylation – putting the pieces together. *FEBS Journal* **2010,** *277* (1), 81-94.

76. Barry, F. P.; Neame, P. J.; Sasse, J.; Pearson, D. Length Variation in the Keratan Sulfate Domain of Mammalian Aggrecan. *Matrix Biol.* **1994,** *14* (4), 323-328.

77. Yoshida, A.; Kobayashi, K.; Manya, H.; Taniguchi, K.; Kano, H.; Mizuno, M.; Inazu, T.; Mitsuhashi, H.; Takahashi, S.; Takeuchi, M.; Herrmann, R.; Straub, V.; Talim, B.; Voit, T.; Tapaloglu, H.; Toda, T.; Endo, T. Muscular dystrophy and neuronal migration disorder caused by mutations in a glycosyltransferase, POMGnT1. *Dev Cell* **2001,** *1* (5), 717-724.

78. Christner, J. E.; Distler, J. J.; Jourdian, G. W. Biosynthesis of Keratan Sulfate - Purification and Properties of a Galactosyltransferase from Bovine Cornea. *Arch. Biochem. Biophys.* **1979,** *192* (2), 548-558.

79. Ujita, M.; Misra, A. K.; McAuliffe, J.; Hindsgaul, O.; Fukuda, M. Poly-N-acetyllactosamine extension in N-glycans and core 2-and core 4-branched O-glycans is differentially controlled by i-extension enzyme and different members of the beta 1,4-galactosyltransferase gene family. *J. Biol. Chem.* **2000,** *275* (21), 15868-15875.

80. Shiraishi, N.; Natsume, A.; Togayachi, A.; Endo, T.; Akashima, T.; Yamada, Y.; Imai, N.; Nakagawa, S.; Koizumi, S.; Sekine, S.; Narimatsu, H.; Sasaki, K. Identification and characterization of three novel beta 1,3-N-acetylglucosaminyltransferases structurally related to the beta 1,3-galactosyltransferase family. *J. Biol. Chem.* **2001,** *276* (5), 3498-3507.

81. Kitayama, K.; Hayashida, Y.; Nishida, K.; Akama, T. O. Enzymes responsible for synthesis of corneal keratan sulfate glycosaminoglycans. *J. Biol. Chem.* **2007,** *282* (41), 30085-30096.

82. Fukuta, M.; Inazawa, J.; Torii, T.; Tsuzuki, K.; Shimada, E.; Habuchi, O. Molecular cloning and characterization of human keratan sulfate Gal-6-sulfotransferase. *J. Biol. Chem.* **1997,** *272* (51), 32321-32328.

83. Bartes, A.; Bhakta, S.; Hemmerich, S. Sulfation of endothelial mucin by corneal keratan N-acetylglucosamine 6-O-sulfotransferase (GST-4 beta). *Biochem. Biophys. Res. Commun.* **2001,** *282* (4), 928-933.

84. Degroote, S.; LoGuidice, J. M.; Strecker, G.; Ducourouble, M. P.; Roussel, P.; Lamblin, G. Characterization of an N-acetylglucosamine-6-O-sulfotransferase from human respiratory mucosa active on mucin carbohydrate chains. *J. Biol. Chem.* **1997,** *272* (47), 29493-29501.

85. Uchimura, K.; Muramatsu, H.; Kadomatsu, K.; Fan, Q. W.; Kurosawa, N.; Mitsuoka, C.; Kannagi, R.; Habuchi, O.; Muramatsu, T. Molecular cloning and characterization of an N-acetylglucosamine-6-O-sulfotransferase. *J. Biol. Chem.* **1998,** *273* (35), 22577-22583.

86. Itano, N.; Kimata, K. Mammalian hyaluronan synthases. *Iubmb Life* **2002,** *54* (4), 195-199.

87. Yoshida, M.; Itano, N.; Yamada, Y.; Kimata, K. In vitro synthesis of hyaluronan by a single protein derived from mouse HAS1 gene and characterization of amino acid residues essential for the activity. *J. Biol. Chem.* **2000,** *275* (1), 497-506.

88. Deangelis, P. L.; Weigel, P. H. Immunochemical Confirmation of the Primary Structure of Streptococcal Hyaluronan Synthase and Synthesis of High-Molecular-Weight Product by the Recombinant Enzyme. *Biochemistry* **1994,** *33* (31), 9033-9039.

89. Jing, W.; DeAngelis, P. L. Dissection of the two transferase activities of the Pasteurella multocida hyaluronan synthase: two active sites exist in one polypeptide. *Glycobiology* **2000,** *10* (9), 883-889.

90. Cox, M.; Nelson, D. *Lehninger, Principles of Biochemistry*. **2004**.

91. Xu, D.; Esko, J. D. Demystifying Heparan Sulfate-Protein Interactions. *Annual Review of Biochemistry, Vol 83* **2014,** *83*, 129-157.

92. Capila, I.; Linhardt, R. J. Heparin - Protein interactions. *Angew. Chem. Int. Ed.* **2002,** *41* (3), 391-412.

93. Muñoz, E. M.; Linhardt, R. J. Heparin-Binding Domains in Vascular Biology. *Arteriosclerosis, Thrombosis, and Vascular Biology* **2004,** *24* (9), 1549-1557.

94. Meneghetti, M. C. Z.; Hughes, A. J.; Rudd, T. R.; Nader, H. B.; Powell, A. K.; Yates, E. A.; Lima, M. A. Heparan sulfate and heparin interactions with proteins. *J.R. Soc. Interface* **2015,** *12* (110), 1-13.

95. Garg, H. G.; Linhardt, R. J.; Hales, C. A. *Chemistry and Biology of Heparin and Heparan Sulfate*. Elsevier: **2005**.

96. Johnson, D. J. D.; Li, W.; Adams, T. E.; Huntington, J. A. Antithrombin-S195A factor Xa-heparin structure reveals the allosteric mechanism of antithrombin activation. *EMBO J*. **2006,** *25* (9), 2029-2037.

97. Waksman, G.; Herr, A. B. New insights into heparin-induced FGF oligomerization. *Nature Structural Biology* **1998,** *5* (7), 527-530.

98. Spivakkroizman, T.; Lemmon, M. A.; Dikic, I.; Ladbury, J. E.; Pinchasi, D.; Huang, J.; Jaye, M.; Crumley, G.; Schlessinger, J.; Lax, I. Heparin-Induced Oligomerization of Fgf Molecules Is Responsible for Fgf Receptor Dimerization, Activation, and Cell-Proliferation. *Cell* **1994,** *79* (6), 1015-1024.

99. Pellegrini, L.; Burke, D. F.; von Delft, F.; Mulloy, B.; Blundell, T. L. Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. *Nature* **2000,** *407* (6807), 1029-1034.

100. DiGabriele, A. D.; Lax, I.; Chen, D. I.; Svahn, C. M.; Jaye, M.; Schlessinger, J.; Hendrickson, W. A. Structure of a heparin-linked biologically active dimer of fibroblast growth factor. *Nature* **1998,** *393* (6687), 812-817.

101. Peysselon, F.; Ricard-Blum, S. Heparin-protein interactions: From affinity and kinetics to biological roles. Application to an interaction network regulating angiogenesis. *Matrix Biol*. **2014,** *35*, 73-81.

102. Sugahara, K.; Mikami, T.; Uyama, T.; Mizuguchi, S.; Nomura, K.; Kitagawa, H. Recent advances in the structural biology of chondroitin sulfate and dermatan sulfate. *Curr Opin Struc Biol* **2003,** *13* (5), 612-620.

103. Trowbridge, J. M.; Gallo, R. L. Dermatan sulfate: new functions from an old glycosaminoglycan. *Glycobiology* **2002,** *12* (9), 117R-125R.

104. Hamel, D. J.; Sielaff, I.; Proudfoot, A. E. I.; Handel, T. M. Interactions of Chemokines with Glycosaminoglycans. *Method Enzymol* **2009,** *461*, 71-102.

105. Thinakaran, G.; Koo, E. H. Amyloid Precursor Protein Trafficking, Processing, and Function. *J. Biol. Chem*. **2008,** *283* (44), 29615-29619.

106. Xue, Y.; Lee, S. W.; Wang, Y. C.; Ha, Y. Crystal Structure of the E2 Domain of Amyloid Precursor Protein-like Protein 1 in Complex with Sucrose Octasulfate. *J. Biol. Chem*. **2011,** *286* (34), 29748-29757.

107. Small, D. H.; Nurcombe, V.; Reed, G.; Clarris, H.; Moir, R.; Beyreuther, K.; Masters, C. L. A Heparin-Binding Domain in the Amyloid Protein-Precursor of Alzheimers-Disease Is Involved in the Regulation of Neurite Outgrowth. *J. Neurosci*. **1994,** *14* (4), 2117-2127.

108. Wettreich, A.; Sebollela, A.; Carvalho, M. A.; Azevedo, S. P.; Borojevic, R.; Ferreira, S. T.; Coelho-Sampaio, T. Acidic pH modulates the interaction between human granulocyte-macrophage colony-stimulating factor and glycosaminoglycans. *J. Biol. Chem.* **1999,** *274* (44), 31468-31475.

109. Casu, B. *Molecular Biophysics of the Extracellular Matrix*. Humana Press: **1984**; p 69-93.

110. Ferro, D. R.; Provasoli, A.; Ragazzi, M.; Torri, G.; Casu, B.; Gatti, G.; Jacquinet, J. C.; Sinay, P.; Petitou, M.; Choay, J. Evidence for Conformational Equilibrium of the Sulfated L-Iduronate Residue in Heparin and in Synthetic Heparin Monosaccharides and Oligosaccharides - Nmr and Force-Field Studies. *J. Am. Chem. Soc.* **1986,** *108* (21), 6773-6778.

111. Ferro, D. R.; Provasoli, A.; Ragazzi, M.; Casu, B.; Torri, G.; Bossennec, V.; Perly, B.; Sinay, P.; Petitou, M.; Choay, J. Conformer Populations of L-Iduronic Acid Residues in Glycosaminoglycan Sequences. *Carbohydr. Res.* **1990,** *195* (2), 157-167.

112. van Boeckel, C. A. A.; van Aelst, S. F.; Wagenaars, G. N.; Mellema, J. R.; Paulsen, H.; Peters, T.; Pollex, A.; Sinnwell, V. Conformational analysis of synthetic heparin-like oligosaccharides containing α-L-idopyranosyluronic acid. *Recl. Trav. Chim. Pays-Bas* **1987,** *106* (1), 19-29.

113. Casu, B.; Petitou, M.; Provasoli, M.; Sinay, P. Conformational Flexibility - a New Concept for Explaining Binding and Biological Properties of Iduronic Acid-Containing Glycosaminoglycans. *Trends Biochem. Sci* **1988,** *13* (6), 221-225.

114. Guerrini, M.; Guglieri, S.; Beccati, D.; Torri, G.; Viskov, C.; Mourier, P. Conformational transitions induced in heparin octasaccharides by binding with antithrombin III. *Biochem. J* **2006,** *399*, 191-198.

115. Hricovini, M.; Guerrini, M.; Bisio, A.; Torri, G.; Naggi, A.; Casu, B. Active conformations of glycosaminoglycans. NMR determination of the conformation of heparin sequences complexed with antithrombin and fibroblast growth factors in solution. *Semin Thromb Hemost* **2002,** *28* (4), 325-334.

116. Linhardt, R. J.; Rice, K. G.; Kim, Y. S.; Lohse, D. L.; Wang, H. M.; Loganathan, D. Mapping and quantification of the major oligosaccharide components of heparin. *Biochem. J* **1988,** *254* (3), 781-7.

117. Habuchi, H.; Suzuki, S.; Saito, T.; Tamura, T.; Harada, T.; Yoshida, K.; Kimata, K. Structure of a Heparan-Sulfate Oligosaccharide That Binds to Basic Fibroblast Growth-Factor. *Biochem. J* **1992,** *285*, 805-813.

118. Sawaguchi, S.; Yue, B. Y. J. T.; Yeh, P.; Tso, M. O. M. Effects of Intracameral Injection of Chondroitinase Abc in vivo. *Arch Ophthalmol-Chic* **1992,** *110* (1), 110-117.

119. Sasisekharan, R.; Moses, M. A.; Nugent, M. A.; Cooney, C. L.; Langer, R. Heparinase Inhibits Neovascularization. *Proc Natl Acad Sci USA* **1994,** *91* (4), 1524-1528.

120. Linhardt, R. J.; Galliher, P. M.; Cooney, C. L. Polysaccharide Lyases. *Appl. Biochem. Biotechnol.* **1986,** *12* (2), 135-176.

121. Ernst, S.; Langer, R.; Cooney, C. L.; Sasisekharan, R. Enzymatic Degradation of Glycosaminoglycans. *Crit. Rev. Biochem. Mol. Biol.* **1995,** *30* (5), 387-444.

122. Desai, U. R.; Wang, H. M.; Linhardt, R. J. Specificity Studies on the Heparin Lyases from Flavobacterium-Heparinum. *Biochemistry* **1993,** *32* (32), 8140-8145.

123. Linhardt, R. J.; Turnbull, J. E.; Wang, H. M.; Loganathan, D.; Gallagher, J. T. Examination of the Substrate-Specificity of Heparin and Heparan-Sulfate Lyases. *Biochemistry* **1990,** *29* (10), 2611-2617.

124. Nader, H. B.; Porcionatto, M. A.; Tersariol, I. L. S.; Pinhal, M. A. S.; Oliveira, F. W.; Moraes, C. T.; Dietrich, C. P. Purification and Substrate-Specificity of Heparitinase-I and

Heparitinase-II from Flavobacterium-Heparinum - Analyses of the Heparin and Heparan-Sulfate Degradation Products by C-13 Nmr-Spectroscopy. *J. Biol. Chem.* **1990,** *265* (28), 16807-16813.

125. Rice, K. G.; Linhardt, R. J. Study of Structurally Defined Oligosaccharide Substrates of Heparin and Heparan Monosulfate Lyases. *Carbohydr. Res.* **1989,** *190* (2), 219-233.

126. Tam, Y. C.; Chan, E. C. S. Purification and Characterization of Hyaluronidase from Oral Peptostreptococcus Species. *Infect Immun* **1985,** *47* (2), 508-513.

127. Karst, N. A.; Linhardt, R. J. Recent chemical and enzymatic approaches to the synthesis of glycosaminoglycan oligosaccharides. *Curr Med Chem* **2003,** *10* (19), 1993-2031.

128. Avci, F. Y.; Karst, N. A.; Linhardt, R. J. Synthetic oligosaccharides as heparin-mimetics displaying anticoagulant properties. *Curr Pharm Design* **2003,** *9* (28), 2323-2335.

129. Gavard, O.; Hersant, Y.; Alais, J.; Duverger, V.; Dilhas, A.; Bascou, A.; Bonnaffe, D. Efficient preparation of three building blocks for the synthesis of heparan sulfate fragments: Towards the combinatorial synthesis of oligosaccharides from hypervariable regions. *Eur. J. Org. Chem.* **2003,** (18), 3603-3620.

130. Yu, H. N.; Furukawa, J.; Ikeda, T.; Wong, C. H. Novel efficient routes to heparin monosaccharides and disaccharides achieved via regio- and stereoselective glycosidation. *Org. Lett.* **2004,** *6* (5), 723-726.

131. Lee, J. C.; Chang, S. W.; Liao, C. C.; Chi, F. C.; Chen, C. S.; Wen, Y. S.; Wang, C. C.; Kulkarni, S. S.; Puranik, R.; Liu, Y. H.; Hung, S. C. From D-glucose to biologically potent L-hexose derivatives: Synthesis of alpha-L-iduronidase fluorogenic detector and the disaccharide moieties of bleomycin A(2) and heparan sulfate. *Chem. Eur. J.* **2004,** *10* (2), 399-415.

132. Ke, W.; Whitfield, D. M.; Gill, M.; Larocque, S.; Yu, S. H. A short route to L-iduronic acid building blocks for the syntheses of heparin-like disaccharides. *Tetrahedron Lett*. **2003,** *44* (42), 7767-7770.

133. Dilhas, A.; Bonnaffe, D. Efficient selective preparation of methyl-1,2,4-tri-O-acetyl-3-O-benzyl-beta-L-idopyranuronate from methyl 3-O-benzyl-L-iduronate. *Carbohydr. Res*. **2003,** *338* (7), 681-686.

134. Schell, P.; Orgueira, H. A.; Roehrig, S.; Seeberger, P. H. Synthesis and transformations of D-glucuronic and L-iduronic acid glycals. *Tetrahedron Lett*. **2001,** *42* (23), 3811-3814.

135. Das, S. K.; Mallet, J. M.; Esnault, J.; Driguez, P. A.; Duchaussoy, P.; Sizun, P.; Herault, J. P.; Herbert, J. M.; Petitou, M.; Sinay, P. Synthesis of conformationally locked L-iduronic acid derivatives: Direct evidence for a critical role of the skew-boat (2)S(0) conformer in the activation of antithrombin by heparin. *Chem. Eur. J*. **2001,** *7* (22), 4821-4834.

136. Lubineau, A.; Gavard, O.; Alais, J.; Bonnaffe, D. New accesses to L-iduronyl synthons. *Tetrahedron Lett*. **2000,** *41* (3), 307-311.

137. Ojeda, R.; de Paz, J. L.; Martin-Lomas, H.; Lassaletta, J. M. A new route to L-iduronate building-blocks for the synthesis of heparin-like oligosaccharides. *Synlett* **1999,** (8), 1316-1318.

138. Orgueira, H. A.; Bartolozzi, A.; Schell, P.; Litjens, R. E. J. N.; Palmacci, E. R.; Seeberger, P. H. Modular synthesis of heparin oligosaccharides. *Chem. Eur. J*. **2003,** *9* (1), 140-169.

139. Lohman, G. J. S.; Hunt, D. K.; Hogermeier, J. A.; Seeberger, P. H. Synthesis of iduronic acid building blocks for the modular assembly of glycosaminoglycans. *J. Org. Chem*. **2003,** *68* (19), 7559-7561.

140. Jaurand, G.; Tabeur, C.; Petitou, M. Synthesis of the Basic Disaccharide Unit of Heparin. *Carbohydr. Res.* **1994,** *255*, 295-301.

141. Haller, M.; Boons, G. J. Towards a modular approach for heparin synthesis. *J Chem Soc Perk T 1* **2001,** (8), 814-822.

142. Hung, S. C.; Thopate, S. R.; Chi, F. C.; Chang, S. W.; Lee, J. C.; Wang, C. C.; Wen, Y. S. 1,6-anhydro-beta-L-hexopyranoses as potent synthons in the synthesis of the disaccharide units of bleomycin A(2) and heparin. *J. Am. Chem. Soc.* **2001,** *123* (13), 3153-3154.

143. Hung, S. C.; Puranik, R.; Chi, F. C. Novel synthesis of 1,2 : 3,5-di-O-isopropylidene-beta-L-idofuranoside and its derivatives at C6. *Tetrahedron Lett.* **2000,***41* (1), 77-80.

144. Prabhu, A.; Venot, A.; Boons, G. J. New set of orthogonal protecting groups for the modular synthesis of heparan sulfate fragments. *Org. Lett.* **2003,** *5* (26), 4975-4978.

145. Vanboeckel, C. A. A.; Beetz, T.; Vos, J. N.; Dejong, A. J. M.; Vanaelst, S. F.; Vandenbosch, R. H.; Mertens, J. M. R.; Vandervlugt, F. A. Synthesis of a Pentasaccharide Corresponding to the Antithrombin-III Binding Fragment of Heparin. *J. Carbohydr. Chem.* **1985,** *4* (3), 293-321.

146. Sinay, P.; Jacquinet, J. C.; Petitou, M.; Duchaussoy, P.; Lederman, I.; Choay, J.; Torri, G. Total Synthesis of a Heparin Pentasaccharide Fragment Having High-Affinity for Antithrombin-III. *Carbohydr. Res.* **1984,** *132* (2), C5-C9.

147. Petitou, M.; van Boeckel, C. A. A. A synthetic antithrombin III binding pentasaccharide is now a drug! What comes next? *Angew. Chem. Int. Ed.* **2004,** *43* (24), 3118-3133.

148. Petitou, M.; Duchaussoy, P.; Lederman, I.; Choay, J.; Sinay, P. Binding of Heparin to Antithrombin-III - a Chemical Proof of the Critical Role Played by a 3-Sulfated 2-Amino-2-Deoxy-D-Glucose Residue. *Carbohydr. Res.* **1988,** *179*, 163-172.

149. Petitou, M.; Lormeau, J. C.; Choay, J. Interaction of Heparin and Antithrombin-III - the Role of O-Sulfate Groups. *Eur. J. Biochem.* **1988,** *176* (3), 637-640.

150. Riesenfeld, J.; Thunberg, L.; Hook, M.; Lindahl, U. The Antithrombin-Binding Sequence of Heparin - Location of Essential N-Sulfate Groups. *J. Biol. Chem.* **1981,** *256* (5), 2389-2394.

151. Vanboeckel, C. A. A.; Lucas, H.; Vanaelst, S. F.; Vandennieuwenhof, M. W. P.; Wagenaars, G. N.; Mellema, J. R. Synthesis and Conformational-Analysis Of an Analog of the Antithrombin-Binding Region of Heparin - the Role of the Carboxylate Function of Alpha-L-Idopyranuronate. *Recl Trav Chim Pay B* **1987,** *106* (11), 581-591.

152. Arungundram, S.; Al-Mafraji, K.; Asong, J.; Leach, F. E., III; Amster, I. J.; Venot, A.; Turnbull, J. E.; Boons, G. J. Modular synthesis of heparan sulfate oligosaccharides for structure-activity relationship studies. *J. Am. Chem. Soc.* **2009,** *131* (47), 17394-405.

153. Schworer, R.; Zubkova, O. V.; Turnbull, J. E.; Tyler, P. C. Synthesis of a Targeted Library of Heparan Sulfate Hexa- to Dodecasaccharides as Inhibitors of beta-Secretase: Potential Therapeutics for Alzheimer's Disease. *Chem. Eur. J.* **2013,** *19* (21), 6817-6823.

154. Baleux, F.; Loureiro-Morais, L.; Hersant, Y.; Clayette, P.; Arenzana-Seisdedos, F.; Bonnaffe, D.; Lortat-Jacob, H. A synthetic CD4-heparan sulfate glycoconjugate inhibits CCR5 and CXCR4 HIV-1 attachment and entry. *Nat Chem Biol* **2009,** *5* (10), 743-748.

155. Hu, Y. P.; Lin, S. Y.; Huang, C. Y.; Zulueta, M. M. L.; Liu, J. Y.; Chang, W.; Hung, S. C. Synthesis of 3-O-sulfonated heparan sulfate octasaccharides that inhibit the herpes simplex virus type 1 host-cell interaction. *Nat Chem* **2011,** *3* (7), 557-563.

156. Lubineau, A.; Lortat-Jacob, H.; Gavard, O.; Sarrazin, S.; Bonnaffe, D. Synthesis of tailor-made glycoconjugate mimetics of heparan sulfate that bind IFN-gamma in the nanomolar range. *Chem. Eur. J.* **2004,** *10* (17), 4265-4282.

157. Orgueira, H. A.; Bartolozzi, A.; Schell, P.; Litjens, R. E. J. N.; Palmacci, E. R.; Seeberger, P. H. Modular Synthesis of Heparin Oligosaccharides. *Chem. Eur. J.* **2003,** *9* (1), 140-169.

158. Hansen, S. U.; Miller, G. J.; Cliff, M. J.; Jayson, G. C.; Gardiner, J. M. Making the longest sugars: a chemical synthesis of heparin-related [4]n oligosaccharides from 16-mer to 40-mer. *Chem. Sci.* **2015,** 6, 6158-6164.

159. Codee, J. D. C.; Stubba, B.; Schiattarella, M.; Overkleeft, H. S.; van Boeckel, C. A. A.; van Boom, J. H.; van der Marel, G. A. A modular strategy toward the synthesis of heparin-like oligosaccharides using monomeric building blocks in a sequential glycosylation strategy. *J. Am. Chem. Soc.* **2005,** *127* (11), 3767-3773.

160. Hansen, S. U.; Miller, G. J.; Jayson, G. C.; Gardiner, J. M. First Gram-Scale Synthesis of a Heparin-Related Dodecasaccharide. *Org. Lett.* **2013,** *15* (1), 88-91.

161. DreefTromp, C. M.; Willems, H. A. M.; Westerduin, P.; vanVeelen, P.; vanBoeckel, C. A. A. Polymer-supported solution synthesis of heparan sulphate-like oligomers. *Biorg. Med. Chem. Lett.* **1997,** *7* (9), 1175-1180.

162. Douglas, S. P.; Whitfield, D. M.; Krepinsky, J. J. Polymer-Supported Solution Synthesis of Oligosaccharides. *J. Am. Chem. Soc.* **1991,** *113* (13), 5095-5097.

163. Ojeda, R.; de Paz, J. L.; Martin-Lomas, M. Synthesis of heparin-like oligosaccharides on a soluble polymer support. *Chem. Commun.* **2003,** (19), 2486-2487.

164. Sheng, G. J.; Oh, Y. I.; Chang, S. K.; Hsieh-Wilson, L. C. Tunable Heparan Sulfate Mimetics for Modulating Chemokine Activity. *J. Am. Chem. Soc.* **2013,** *135* (30), 10898-10901.

165. Oh, Y. I.; Sheng, G. J.; Chang, S. K.; Hsieh-Wilson, L. C. Tailored Glycopolymers as Anticoagulant Heparin Mimetics. *Angew. Chem. Int. Ed.* **2013,** *52* (45), 11796-11799.

166. Rawat, M.; Gama, C. I.; Matson, J. B.; Hsieh-Wilson, L. C. Neuroactive chondroitin sulfate glycomimetics. *J. Am. Chem. Soc.* **2008,** *130* (10), 2959-2961.

167. Gama, C. I.; Tully, S. E.; Sotogaku, N.; Clark, P. M.; Rawat, M.; Vaidehi, N.; Goddard, W. A.; Nishi, A.; Hsieh-Wilson, L. C. Sulfation patterns of glycosaminoglycans encode molecular recognition and activity. *Nat Chem Biol* **2006,** *2* (9), 467-473.

168. Tully, S. E.; Mabon, R.; Gama, C. I.; Tsai, S. M.; Liu, X. W.; Hsieh-Wilson, L. C. A chondroitin sulfate small molecule that stimulates neuronal growth. *J. Am. Chem. Soc.* **2004,** *126* (25), 7736-7737.

169. Lopin, C.; Jacquinet, J. C. From polymer to size-defined oligomers: An expeditious route for the preparation of chondroitin oligosaccharides. *Angew. Chem. Int. Ed.* **2006,** *45* (16), 2574-2578.

170. Eller, S.; Collot, M.; Yin, J.; Hahm, H. S.; Seeberger, P. H. Automated Solid-Phase Synthesis of Chondroitin Sulfate Glycosaminoglycans. *Angew. Chem. Int. Ed.* **2013,** *52* (22), 5858-5861.

171. Lu, X. W.; Kamat, M. N.; Huang, L. J.; Huang, X. F. Chemical Synthesis of a Hyaluronic Acid Decasaccharide. *J. Org. Chem.* **2009,** *74* (20), 7608-7617.

172. Peterson, S.; Frick, A.; Liu, J. Design of biologically active heparan sulfate and heparin using an enzyme-based approach. *Natural Product Reports* **2009,** *26* (5), 610-627.

173. Kusche, M.; Hannesson, H. H.; Lindahl, U. Biosynthesis of Heparin - Use of Escherichia-Coli K5 Capsular Polysaccharide as a Model Substrate in Enzymatic Polymer-Modification Reactions. *Biochem. J* **1991,** *275*, 151-158.

174. Lindahl, U.; Li, J. P.; Kusche-Gullberg, M.; Salmivirta, M.; Alaranta, S.; Veromaa, T.; Emeis, J.; Roberts, I.; Taylor, C.; Oreste, P.; Zoppetti, G.; Naggi, A.; Torri, G.; Casu, B. Generation of "Neoheparin" from E coli K5 capsular polysaccharide. *J. Med. Chem.* **2005,** *48* (2), 349-352.

175. Kuberan, B.; Beeler, D. L.; Lech, M.; Wu, Z. L. L.; Rosenberg, R. D. Chemoenzymatic synthesis of classical and non-classical anticoagulant heparan sulfate polysaccharides. *J. Biol. Chem.* **2003,** *278* (52), 52613-52621.

176. Li, J. P.; HagnerMcWhirter, A.; Kjellen, L.; Palgi, J.; Jalkanen, M.; Lindahl, U. Biosynthesis of heparin/heparan sulfate - DNA cloning and expression of D-glucuronyl C5-epimerase from bovine lung. *J. Biol. Chem.* **1997,** *272* (44), 28158-28163.

177. Kuberan, B.; Lech, M. Z.; Beeler, D. L.; Wu, Z. L. L.; Rosenberg, R. D. Enzymatic synthesis of antithrombin III-binding heparan sulfate pentasaccharide. *Nat. Biotechnol.* **2003,** *21* (11), 1343-1346.

178. Xu, Y.; Masuko, S.; Takieddin, M.; Xu, H.; Liu, R.; Jing, J.; Mousa, S. A.; Linhardt, R. J.; Liu, J. Chemoenzymatic Synthesis of Homogeneous Ultralow Molecular Weight Heparins. *Science* **2011,** *334* (6055), 498-501.

179. Shaklee, P. N.; Conrad, H. E. Hydrazinolysis of Heparin and Other Glycosaminoglycans. *Biochem. J* **1984,** *217* (1), 187-197.

180. Chen, M.; Bridges, A.; Liu, J. Determination of the substrate specificities of N-acetyl-D-glucosaminyltransferase. *Biochemistry* **2006,** *45* (40), 12358-12365.

181. Hodson, N.; Griffiths, G.; Cook, N.; Pourhossein, M.; Gottfridson, E.; Lind, T.; Lidholt, K.; Roberts, I. S. Identification that KfiA, a protein essential for the biosynthesis of the Escherichia coli K5 capsular polysaccharide, is an alpha-UDP-GlcNAc glycosyltransferase - The formation of a membrane-associated K5 biosynthetic complex requires KfiA, KfiB, and KfiC. *J. Biol. Chem.* **2000,** *275* (35), 27311-27315.

182. Otto, N. J.; Green, D. E.; Masuko, S.; Mayer, A.; Tanner, M. E.; Linhardt, R. J.; DeAngelis, P. L. Structure/Function Analysis of Pasteurella multocida Heparosan Synthases: toward defining enzyme specificity and engineering novel catalysts. *J. Biol. Chem.* **2012,** *287* (10), 7203-7212.

183. Liu, R. P.; Xu, Y. M.; Chen, M. A.; Weiwer, M.; Zhou, X. X.; Bridges, A. S.; DeAngelis, P. L.; Zhang, Q. S.; Linhardt, R. J.; Liu, J. A. Chemoenzymatic Design of Heparan Sulfate Oligosaccharides. *J. Biol. Chem.* **2010,** *285* (44), 34240-34249.

184. Johnson, A. H.; Baker, J. R. Enzymatic Sulfation of Heparan Sulfate by Hen's Uterus. *Biochim. Biophys. Acta* **1973,** *320* (2), 341-351.

185. Fujikawa, S.; Ohmae, M.; Kobayashi, S. Enzymatic synthesis of chondroitin 4-sulfate with well-defined structure. *Biomacromolecules* **2005,** *6* (6), 2935-2942.

186. Kobayashi, S.; Fujikawa, S.; Ohmae, M. Enzymatic synthesis of chondroitin and its derivatives catalyzed by hyaluronidase. *J. Am. Chem. Soc.* **2003,** *125* (47), 14357-14369.

187. Meyer, K. The Biological Significance of Hyaluronic Acid and Hyaluronidase. *Physiol Rev* **1947,** *27* (3), 335-359.

188. Vibert, A.; Jacquinet, J. C.; Lopin-Bon, C. Recent Advances in the Chemical and Enzymatic Chondroitin Sulfate Synthesis. *J. Carbohydr. Chem.* **2011,** *30* (7-9), 393-414.

189. Kobayashi, S.; Ohmae, M.; Ochiai, H.; Fujikawa, S. I. A hyaluronidase supercatalyst for the enzymatic polymerization to synthesize glycosaminoglycans. *Chem. Eur. J.* **2006,** *12* (23), 5962-5971.

190. Sugiura, N.; Shimokata, S.; Minamisawa, T.; Hirabayashi, J.; Kimata, K.; Watanabe, H. Sequential synthesis of chondroitin oligosaccharides by immobilized chondroitin polymerase mutants. *Glycoconjugate J.* **2008,** *25* (6), 521-530.

191. Sobhany, M.; Kakuta, Y.; Sugiura, N.; Kimata, K.; Negishi, M. The Structural Basis for a Coordinated Reaction Catalyzed by a Bifunctional Glycosyltransferase in Chondroitin Biosynthesis. *J. Biol. Chem.* **2012,** *287* (43), 36022-36028.

192. Deangelis, P. L.; Oatman, L. C.; Gay, D. F. Rapid chemoenzymatic synthesis of monodisperse hyaluronan oligosaccharides with immobilized enzyme reactors. *J. Biol. Chem.* **2003,** *278* (37), 35199-35203.

193. Posner, T. *Ber. Dtsch. Chem. Ges.* **1905,** *38*, 646-657.

194. Dondoni, A.; Marra, A. Recent applications of thiol-ene coupling as a click process for glycoconjugation. *Chem. Soc. Rev.* **2012,** *41* (2), 573-586.

195. Natali, M.; Begolo, S.; Carofiglio, T.; Mistura, G. Rapid prototyping of multilayer thiolene microfluidic chips by photopolymerization and transfer lamination. *Lab Chip* **2008,** *8* (3), 492-494.

196. Khire, V. S.; Yi, Y.; Clark, N. A.; Bowman, C. N. Formation and surface modification of nanopatterned thiol-ene substrates using Step and Flash Imprint Lithography. *Adv. Mater.* **2008,** *20* (17), 3308-3313.

197. Rydholm, A. E.; Reddy, S. K.; Anseth, K. S.; Bowman, C. N. Controlling network structure in degradable thiol-acrylate biomaterials to tune mass loss behavior. *Biomacromolecules* **2006,** *7* (10), 2827-2836.

198. Floyd, N.; Vijayakrishnan, B.; Koeppe, A. R.; Davis, B. G. Thiyl Glycosylation of Olefinic Proteins: S-Linked Glycoconjugate Synthesis. *Angew. Chem. Int. Ed.* **2009,** *48* (42), 7798-7802.

199. Geng, Y.; Discher, D. E.; Justynska, J.; Schlaad, H. Grafting short peptides onto polybutadiene-block-poly(ethylene oxide): A platform for self-assembling hybrid amphiphiles. *Angew. Chem. Int. Ed.* **2006,** *45* (45), 7578-7581.

200. van Seeventer, P. B.; van Dorst, J. A. L. M.; Siemerink, J. F.; Kamerling, J. P.; Vliegenthart, J. F. G. Thiol addition to protected allyl glycosides: An improved method for the preparation of spacer-arm glycosides. *Carbohydr. Res.* **1997,** *300* (4), 369-373.

201. Heidecke, C. D.; Lindhorst, T. K. Iterative synthesis of spacered glycodendrons as oligomannoside mimetics and evaluation of their antiadhesive properties. *Chem. Eur. J.* **2007,** *13* (32), 9056-9067.

202. Dondoni, A. Free-radical thiol-ene and thiol-yne coupling as selective ligation tools for glycoconjugation. *Abstr Pap Am Chem S* **2010,** *240*.

203. Fiore, M.; Lo Conte, M.; Pacifico, S.; Marra, A.; Dondoni, A. Synthesis of S-glycosyl amino acids and S-glycopeptides via photoinduced click thiol-ene coupling. *Tetrahedron Lett*. **2011,** *52* (3), 444-447.

204. Floyd, N.; Vijayakrishnan, B.; Koeppe, J. R.; Davis, B. G. Thiyl Glycosylation of Olefinic Proteins: S-Linked Glycoconjugate Synthesis. *Angew. Chem. Int. Ed*. **2009,** *48* (42), 7798-7802.

205. Wittrock, S.; Becker, T.; Kunz, H. Synthetic vaccines of tumor-associated glycopeptide antigens by immune-compatible thioether linkage to bovine serum albumin. *Angew. Chem. Int. Ed*. **2007,** *46* (27), 5226-5230.

206. Triola, G.; Brunsveld, L.; Waldmann, H. Racemization-free synthesis of S-alkylated cysteines via thiol-ene reaction. *J. Org. Chem*. **2008,** *73* (9), 3646-3649.

207. Buskas, T.; Soderberg, E.; Konradsson, P.; Fraser-Reid, B. Use of n-pentenyl glycosides as precursors to various spacer functionalities. *J. Org. Chem*. **2000,** *65* (4), 958-963.

Chapter 2

Thiol-ene radical reaction on model substrates

Table of Contents

Chapter –2 Thiol-ene radical reaction on model substrates

2.1 Introduction

As discussed in Chapter 1, thiol-ene radical chemistry has been extensively used for attaching peptides and amino acids to carbohydrate components, both in protected and deprotected conditions[1-3]. It has also been used for the preparation of glycopolymers and glycopeptide vaccine candidates, among other applications[4-5].

There have been reports of thiol-ene coupling between two sugar fragments[6]. The coupling reactions were carried out using several sugar thiol–sugar alkene pairs in a variety of solvents such as methanol, DCM and toluene. This work established thiol-ene "click" reaction[7] as an efficient way of preparing thiodisaccharides in high yields and diastereoselectivity. Furthermore, the authors also demonstrated the synthesis of the thio-isosteres of naturally occurring *O*-disaccharides. The main caveat is that the work focused exclusively on the utilization of exocyclic double bonds for thiol-ene reaction. No examples of endocyclic double bond utilization were depicted.

These examples suggest that this thiol-ene click reaction can be conveniently used to prepare thioglycoside conjugates such as thio-oligosaccharides and glycopeptides without significant racemization. Notably, the relatively mild UV light activation conditions are compatible with biomolecules such as peptides and carbohydrates[3]. Thio-linked glycopeptides

and glycoderivatives are known to show enhanced resistance towards enzymatic hydrolysis in comparison to natural oxygen-linked counterparts[8]. At the same time these thio-analogs have been found to be immunocompatible[5,9]. It can be safely said therefore, that thiol-ene click reactions can prove to be an invaluable tool in the bio-organic domain for construction of drug molecules, hitherto undiscovered.

Given the utility of this photocatalyzed thiol-ene reaction toward sugar synthesis, we have developed conditions suitable for coupling a broad scope of sugar thiols and alkenes. This work ultimately targets the construction of defined thio-heparin building blocks, a challenging goal that could be addressed by the scope and selectivity of this reaction.

2.2 Results and Discussion

In order to carry out the proposed coupling toward heparin analogs, it was necessary to establish thiol-ene coupling conditions on a model system. This enabled us to optimize the conditions toward suitable yields and the required stereochemistry of the glycosidic linkages. For this purpose we chose 2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-1-thio-β-D-glucopyranose and 2,3,4,6-tetra-*O*-acetyl-1-thio-β-D-glucopyranoside as the model thio systems and methyl-(methyl 2,3-di-*O*-acetyl-4-deoxy-α-L-threo-hex-4-enopyranosid)uronate as the unsaturated system. Successful conditions were extended to the corresponding deprotected derivatives as well as deprotected α-thioglucose.

2.2.1 Synthesis of model substrate precursors

2.2.1.1 1-Glycosyl thiol synthesis

A number of glycosyl thiols were synthesized to be used as model thio-substrates to carry out optimization of the thio-radical addition reaction to the model unsaturated system. These thiols varied in their anomeric configuration (e.g. α- and β-thioglucose) and modification (protected or unprotected) (**Figure 1**).

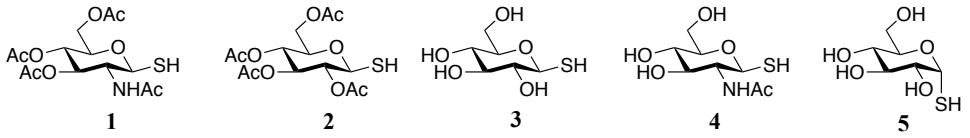
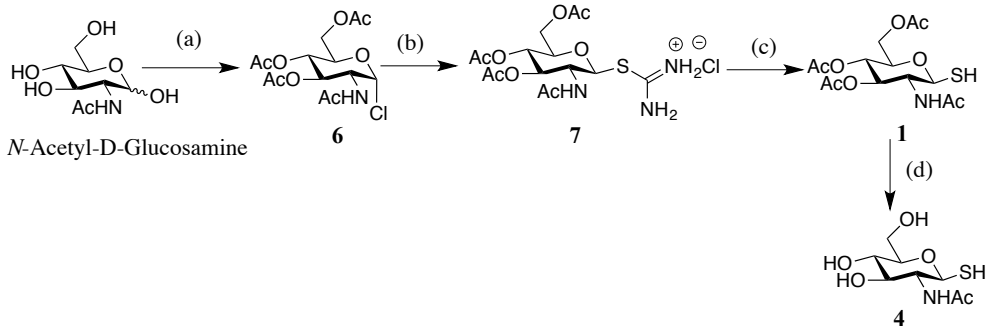


Figure 1: Protected and deprotected thiosugars used for model thio-radical addition reaction.

2.2.1.1.1 2-Acetamido-2-deoxy-1-thio-β-D-glucopyranose

The synthesis of 2-acetamido-2-deoxy-1-thio-β-D-glucopyranose (β-GlcNAcSH) was started from commercially available *N*-acetyl-D-glucosamine (**Scheme 1**). Glycosyl chloride **6** was prepared by treating *N*-acetyl-D-glucosamine with acetyl chloride[10-11]. **6** was then subsequently treated with thiourea to yield the isothiouronium salt **7**[12], which was then hydrolysed using sodium metabisulfite to yield the protected thiosugar **1**[13]. Zemplén deacetylation[14] was subsequently performed to yield the deprotected thiol **4**.



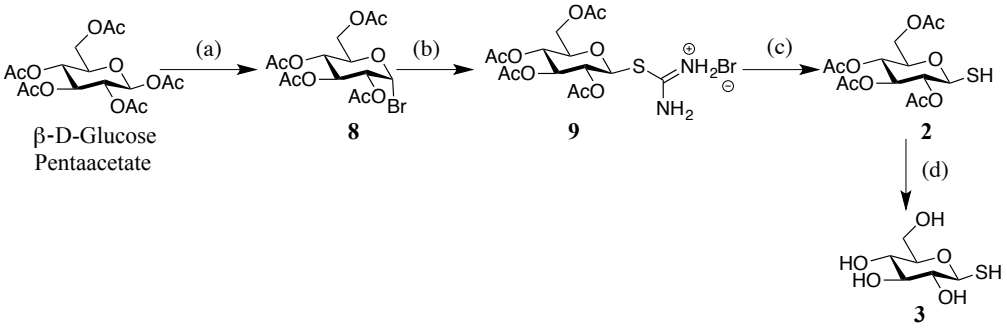
N-Acetyl-D-Glucosamine

6 **7** **1** **4**

Scheme 1. (a) Acetyl Chloride (49%) (b) Thiourea, Acetone, 60 °C (72%) (c) $Na_2S_2O_5$, DCM/H_2O, reflux (86%) (d) NaOMe, MeOH, r.t. (96%)

2.2.1.1.2 1-thio-β-D-glucopyranose

Synthesis of 1-thio-β-D-glucopyranose was similarly initiated from commercially available glucose pentaacetate (**Scheme 2**). Glycosyl bromide **8** was prepared by treating the pentaacetate with HBr in acetic acid[10-11]. This yielded the isothiouronium salt **9** when treated with thiourea[12], due to neighboring group participation. Sodium metabisulfite-assisted hydrolysis provided the protected thiosugar **2**[13], which when exposed to Zemplén deacetylation conditions gave the deprotected thiosugar **3**[14]. Alternately, commercially available 2,3,4,6 tetra-*O*-acetyl-1-thio-β-D-glucopyranose was deprotected using Zemplén deacetylation to give the product.



β-D-Glucose Pentaacetate

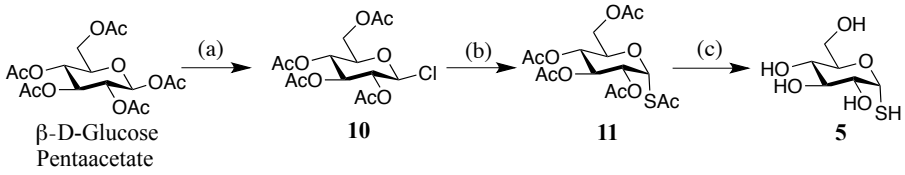
8 **9** **2** **3**

Scheme 2. (a) HBr (33% in acetic acid), DCM (84%) (b) Thiourea, Acetone, 60 °C (78%) (c) $Na_2S_2O_5$, DCM/H_2O, reflux (88%) (d) NaOMe, MeOH, r.t. (96%)

2.2.1.1.3 1-thio-α-D-glucopyranose

Glycosyl chloride **10** was prepared by reacting the pentaacetate with PCl_5[15] (**Scheme 3**). Subsequently, a S_N2 reaction was carried out using KSAc to install the thioacetate in α-configuration to give the corresponding protected thiosugar, **11**. The reaction used was a modification of a procedure reported by Blanc-Muesser et al.[16] where more carcinogenic HMPA was replaced by DMPU, without compromising on yield. This step was found to be particularly slow, taking 3 days to complete, which might be because KSAc takes considerable time to dissolve in DMPU, the solvent for the reaction. Next, Zemplén deacetylation[14] of the protected sugar was carried out to yield the deprotected thiol **5**.

NMR analysis (**Experimental Section 2.4.3**) of the thiosugar after deacetylation revealed a significant presence of β-sugar alongside the intended α-sugar. The desired α-thiosugar and the minor β-sugar were present in a 5:3 ratio.



Scheme 3. (a) PCl_5, $BF_3.Et_2O$, DCM (47%) (b) KSAc, DMPU, 0 °C (67%) (c) NaOMe, MeOH, r.t. (94%)

A search for literature precedence revealed that this result can be explained by the pathway shown in **Figure 2**, as reported by Ramström et al.[17]

In acidic conditions, the sugar spontaneously mutarotates to give the β-thiosugar. Also, in the presence of water, the sugar can spontaneously hydrolyze, yielding an anomeric mixture of oxo-sugars and releasing H_2S in the process. In our case, it was observed in NMR that the peak

corresponding to a β-sugar was the β-thiosugar (based on coupling constants). Therefore, we decided to neutralize the basic reaction mixture with minimum amount of DOWEX 50WX8-200 to keep the resultant pH slightly alkaline and also expose it to minimum amount of water during subsequent lyophilization. Fortunately, with the aforementioned precautions, significant hydrolysis or anomerisation was not observed for subsequent reactions.

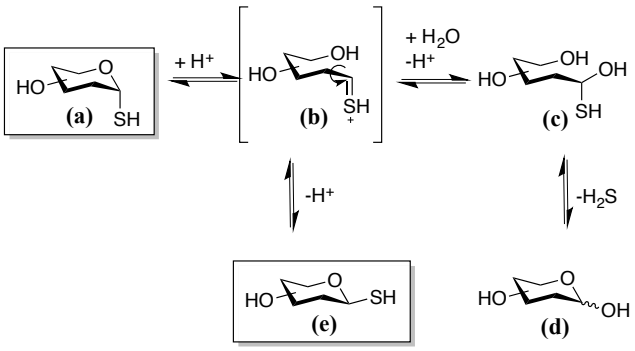
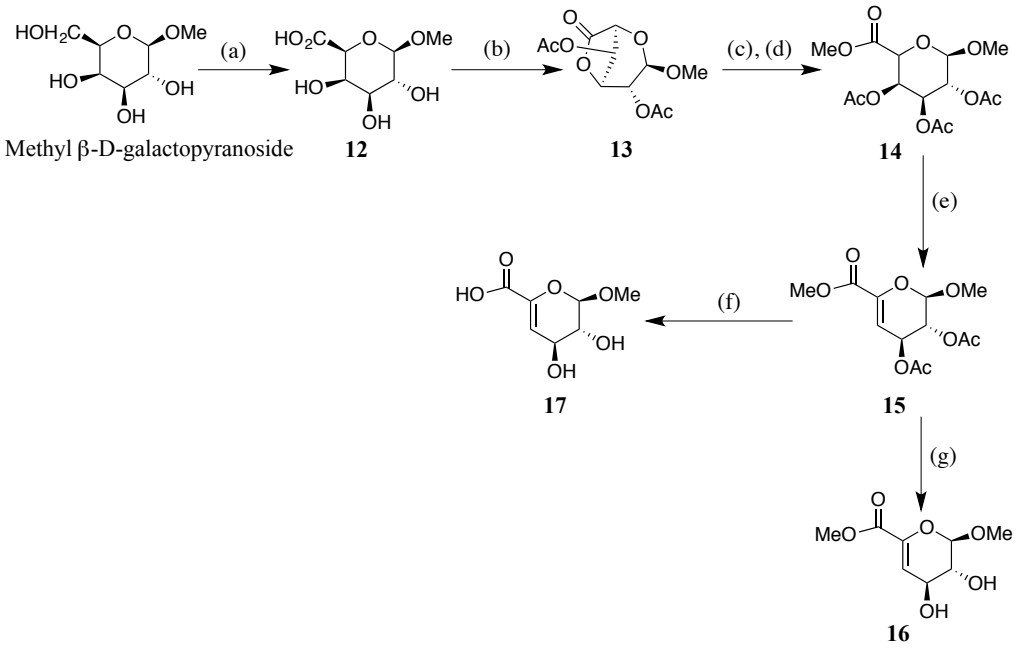


Figure 2: Proposed α-thiosugar anomerization and hydrolysis pathway in acidic condition, mechanism adopted from scheme proposed by Ramström et al.[17]

2.2.1.2 Unsaturated model sugar synthesis

The synthesis of methyl 4-deoxy-α-L-threo-hex-4-enopyranosiduronic acid was initiated from commercially available methyl β-D-galactopyranoside (**Scheme 4**)[18]. Oxidation of the starting material, using catalytic TEMPO and sodium hypochlorite/sodium bromide was carried out to yield the corresponding uronic acid **12**[19-21]. Without purification, the product was subsequently acetylated by refluxing with acetic anhydride and sodium acetate for 15 minutes[22]. This yielded the γ-lactone **13**[23], which was formed via esterification of the newly formed carboxylic acid group on position 5 with the hydroxyl group at position 3. The remaining two hydroxyl groups were acetylated in the process. Ring opening of the lactone was then performed

using DBU in methanol to yield the corresponding methyl ester. Partial deacetylation and acetyl group migration are observed as side reactions in this step. Without attempting to separate them, the mixture was reacetylated using acetic anhydride-pyridine and DBU as catalyst to yield the completely protected product **14**. Thereafter, unsaturation was introduced between position 4 and 5 by an elimination reaction using DBU in dry DCM[24-25]. The product from the previous step, **15** was partially deprotected to yield the product **16**[26]. Finally, base-catalysed hydrolysis using LiOH, gave the deprotected uronic acid **17**[18].

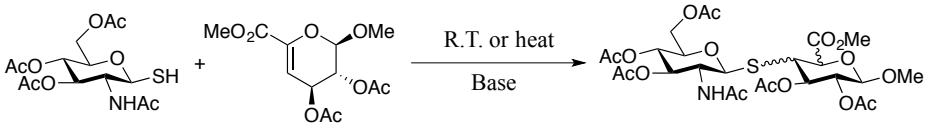


Scheme 4. (a) TEMPO, NaBr, NaOCl, 0 °C (b) Ac_2O, NaOAc, reflux (35%) (c) MeOH, DBU, r.t. (d) Ac_2O, Py, DBU, r.t. (92%) (e) DCM, DBU, r.t. (94%) (f) LiOH (MeOH-H_2O-THF 5:4:1), DOWEX 50 H^+, r.t. (98%) (g) NaOMe, MeOH, r.t. (94%)

2.2.2 Thiol-ene radical coupling between protected systems

With several thiosugars and an unsaturated acceptor in hand, both thiol-ene radical coupling reactions as well as thio-Michael addition reactions were attempted. Vazo-44, a water soluble radical initiator that decomposes at 44 °C or exposure to UV-light, was used to start the radical reactions[27].

Thio-Michael addition reactions were attempted first as summarized in **Table 1**. The initial condition used was a mixture of methanol and sodium carbonate buffer (pH 10) as the solvent system, and the reaction performed at room temperature. Water was used in the solvent system in order to dissolve the base while methanol was expected to dissolve the protected sugars. The two sugars **15** and **1** were used in a 1:1 ratio.



1 **15**

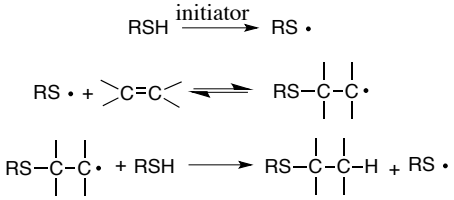
Table 1: Thio-Michael reaction conditions attempted using thiosugar donors and unsaturated acceptors.

Entry	Unsaturated sugar (Equiv.)	Thiosugar (Equiv.)	Condition	Time (h)	Yield (%)
1.	**15** (1)	**1** (1)	R.T. (Methanol-Na_2CO_3 buffer pH 10)	12	-
2.	**15** (1)	**1** (1)	Δ at 55 °C (Methanol-Na_2CO_3 buffer pH 10)	2	-
3.	**15** (1)	**1** (3)	Δ at 55 °C (Methanol-Na_2CO_3 buffer pH 10)	2	-
4.	**15** (1)	**1** (3)	Δ at 55 °C, triethylamine (methanol)	2	-

However, no reaction progress was observed by TLC after 3 h. A moderate formation of sugar disulfide species was observed due to dimerization of the thioradicals. Overnight stirring of the reaction mixture as well as heating it at 55 °C for two hours did not show any reaction progress (**Table 1**, Entry **1** and **2**). A subsequent column chromatographic separation of the reaction mixture led to the recovery of **15** entirely and **1** was recovered mostly as the disulfide. Increasing the equivalents of thiosugar to three equivalents did not have any considerable effect on the reaction outcome (**Table 1,** Entry **3**). Changing the base to triethylamine and the solvent system to methanol only, failed to show any formation of the desired product either (**Table 1**, Entry **4**). In these cases heating only hastened the formation of the disulfide.

Realizing that the unsaturation is not amenable to thio-Michael conditions, the same reaction was attempted with radical conditions. The propagation steps of a thiol-ene radical reaction are shown in **Scheme 5**. In presence of a radical initiator, a thiyl radical is generated from the thiol. Thereafter, the thiyl radical adds on to the unsaturated molecule resulting in the formation of a carbon-centered radical. Subsequently, the carbon-centered radical abstracts a proton from another molecule of thiol to generate a second thiyl radical, along with the formation of the desired product.

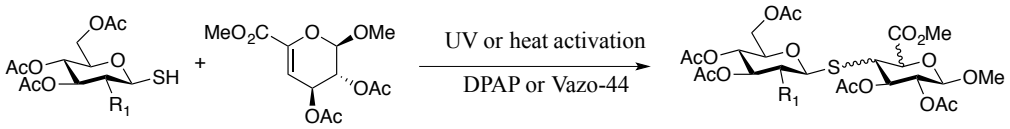
The reaction was performed using 1:1 equivalents of the thiosugar and unsaturated system, in the presence of Vazo-44 and a mixture of methanol and water (at neutral pH) as the solvent system, as summarized in **Table 2**, at room temperature. Water was used in the solvent system in order to dissolve Vazo-44 while methanol was expected to dissolve the protected sugars (**Table 2**, Entry **1**). Here again, **15** was recovered in its entirety. No progress of the reaction was observed even with heating (**Table 2**, Entry **2**).



Scheme 5: Steps involved in a thiol-ene click reaction[28].

Hypothesizing that the thiosugar **1** was more reactive than **15**, the thiosugar was thereafter added dropwise to a stirred solution of the unsaturated system heated at 55 °C (**Table 2**, Entry **4**). The radical initiator was also added to the heated solution of **15** so that the thiosugar would be converted into the radical form and consumed by excess **15** upon entering the heated solution. At the end of 2 hours, however, the unsaturated system and the disulfide were again recovered. Fortunately, UV irradiation of the reaction mixture under initial reaction conditions resulted in traces of product formation (**Table 2**, Entry **3**). An extensive literature review at this stage revealed that the radical addition of thiosugars to homoallylglycine[29] has been greatly facilitated by carrying out the reaction in pH 4 acetate buffer, wherein the low pH prevents significant formation of the disulfide. Accordingly, the next attempt was carried out in a solvent system of MeOH/pH 4 acetate buffer. The yield was observed to increase marginally (**Table 2**, Entry **5**).

Table 2: Optimization of the thiol-ene click reaction between thiosugar donors and unsaturated acceptors. Vazo-44 is 2,2'-azobis[2-(2-imidazolin-2-yl)propane]dihydrochloride), a water soluble photoinitiator[27].



R_1 = OAc (2), NHAc (1) **15**

Entry	Unsaturated sugar (Equiv.)	Thiosugar (Equiv.)	Condition	Photo initiator (Equiv.)	Time (h)	Yield (%)
1	**15** (1)	**1** (1)	R.T. (MeOH-H_2O)	Vazo-44 (0.2 eq.)	12	-
2	**15** (1)	**1** (1)	Δ at 55 °C (MeOH-H_2O)	Vazo-44 (0.2 eq.)	2	-
3	**15** (1)	**1** (1)	hν** (MeOH-H_2O)	Vazo-44 (0.2 eq.)	6	3%*
4	**15** (1)	**1** (1)	Δ at 55 °C, thiosugar added dropwise (MeOH-H_2O)	Vazo-44 (0.2 eq.)	2	-
5	**15** (1)	**1** (1)	hν** (MeOH-pH 4 acetate buffer)	Vazo-44 (0.2 eq.)	6	5%*
6	**15** (1)	**2** (1)	hν** (MeOH-pH 4 acetate buffer)	Vazo-44 (0.2 eq.)	6	5%*
7	**15** (1)	**1** (1)	hν (MeOH)	DPAP (0.2 eq.)	6	15%*
8	**15** (1)	**1** (1)	hν** (MeOH-pH 4 acetate buffer), N_2 atmosphere	Vazo-44 (0.2 eq.)	6	12%
9	**15** (1)	**1** (1)	Δ at 55 °C, (MeOH-pH 4 acetate buffer), N_2 atmosphere	Vazo-44 (0.2 eq.)	3	-
10	**15** (1)	**1** (1)	hν**(MeOH), quartz jacket, N_2 atmosphere	DPAP (0.2 eq.)	6	18%*
11	**15** (1)	**1** (1)	hν** (MeOH), quartz jacket, N_2 atmosphere	DPAP (0.2 eq.)	15	20%*
12	**15** (1)	**1** (2)	hν** (MeOH)	DPAP (0.2 eq.)	6	32%*
13	**15** (1)	**1** (2)	hν** (MeOH)	DPAP (0.3 eq.)	6	39%*
14	**15** (1)	**1** (4)	hν** (MeOH)	DPAP (0.6 eq.)	7	78%
15	**15** (1)	**1** (8)	hν**(EtOH-DCM 20:1)	DPAP (0.6 eq.)	3	56%
16	**15** (1)	**1** (8)	hν** (MeOH)	DPAP (0.6 eq.)	3	84%

* Yield calculated based on ^{1}H NMR, ** UV irradiation carried out with medium-pressure 125 W Hg lamp

Changing the thiosugar to **2** instead of **1** showed the same lack of reactivity (**Table 2**, Entry **6**). Thereafter, the reaction was attempted in MeOH with the radical initiator being DPAP (2,2-dimethoxy-2-phenyl acetophenone) instead of Vazo-44. Fortunately, formation of the

product was observed (**Table 2**, Entry **7**). Increasing the radical initiator concentration helped in increasing the yield, but complete consumption of **15** was not seen.

Since it is known that thioradicals are quenched in the presence of oxygen[30-31], the above reactions were also attempted after flushing the system with N_2, but this also failed to show any significant improvement. Though a glass jacket was used around the UV source, replacement with a quartz jacket also did not have a significant effect on the reaction (**Table 2**, Entries **8**, **9**, **10** and **11**).

The next step comprised of increasing the equivalents of thiosugar with respect to the unsaturation. Accordingly, the equivalents of thiosugar were increased to two equivalents, at which point we saw considerable consumption of **15** at the end of 6 h (**Table 2**, Entry **12**). Encouraged by the outcome, a further two equivalents were added after another hour and the unsaturated system was completely consumed at the end of 7 h, with DPAP equivalent increased to 0.6. Thus, four equivalents of the thiosugar were added in total (**Table 2**, Entry **14**).

During the course of these studies, Dondoni et al.[32] described a free-radical addition of thiosugars to glycals. Interestingly, their reaction condition entails using 20:1 EtOH:DCM as the solvent system to perform their reactions. In our experience, the use of DCM led to the formation of side products ascribed to radical formation by DCM itself[33].

The final reaction conditions entailed using 8 equivalents of the thiosugar, 0.6 equivalents of DPAP and 100% MeOH as the solvent (**Table 2**, Entry **16**). Under these conditions, the reaction was observed to be complete after 2-3 h. The relatively large amount of photoinitiator being used points to the fact that the chain propagation step does not proceed efficiently in this case. This is evident, as a significant amount of the thiyl radicals get converted to the disulfide.

2.2.3 Thiol-ene radical coupling between deprotected systems

To further the utility of this reaction, we sought to perform it using deprotected sugars. Success here could lead to its ultimate use on heparin-derived substrates. For this purpose, both the esterified substrate **16** and the completely deprotected substrate **17** were used. However, **16** did not offer any additional reactivity over **17,** which might have been the case if the reaction proceeded through a thio-Michael mechanism. This is because, in case of thio-Michael reaction, the approaching thiyl anion would be repelled by the negative charge on the free carboxylic group on **17**. With **16**, however, there would not be any charge repulsion since the carboxylic group is protected and thus it would be assumed to have greater reactivity.

With some optimization the deprotected substrates could be coupled in pH 4 acetate buffer, using 10 equivalents of the thiosugar, Vazo-44 as the water-soluble photoinitiator and UV irradiation. A longer reaction time of 8-10 h ensured complete consumption of **17**. Purification was carried out by column chromatography using a WIPE (water:isopropanol:ethylacetate) solvent system. Later, purification by anion exchange chromatography was found to give higher yields. For this purpose, the solvent was changed from pH 4 sodium acetate buffer to pH 4 ammonium formate buffer so that the salt could be removed by lyophilization before purification.

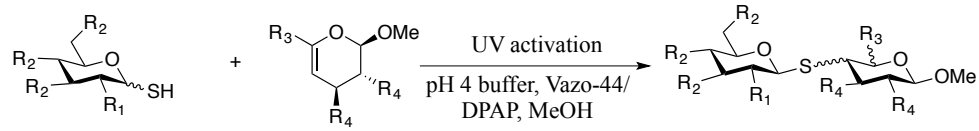
With optimized parameters in hand, a small library of protected and deprotected model disaccharides was prepared in moderate to good yields (**Table 3**). For the deprotected substrates, this reaction was also performed with an α-thiol **5,** further increasing the utility of this reaction.

It was observed, that the conversion was reduced with the α-thiol, with some unreacted **17** recovered after the reaction. This lower reactivity can be explained by the fact that the α-thiol

5 being a c*is*-1,2-hydroxythiol is much more sterically hindered to react through its anomeric position as compared to the β-*trans*-1,2-hydroxythiol[29].

Based on the stereo- and regiochemistry, even though four possible products are expected from this type of reaction, we have recovered only one isomer in isolable yields in all the reactions done. The steroisomer in all cases has been identified after NOESY studies (**Experimental Section 2.4.2**) as coupling constants alone were not enough to identify the correct structure.

This regioselectivity is preceded in literature[32,34]. The electrophilic anomeric thiyl radicals[35] preferentially attack the C-4 position of the unsaturated sugar. The free radical generated at C-5 by this addition can be stabilized by a captodative effect arising from the synergistic electron-withdrawing effect of the adjacent carboxylic group and the electron-donating effect of the adjacent ring oxygen. This contributes to the regioselectivity seen in the reactions described. In terms of the stereochemistry, the formation of the axial linkage over an equatorial linkage is consistent with results obtained from addition of thiyl radicals to cyclohexene derivatives[36-40]. Studies on glucal as the unsaturated counterpart have also yielded similar results[34]. Thus, the stereochemistry may be explained by the facts that attack of the thioradical must occur on the double bond in a perpendicular direction to the π-orbitals for maximum orbital overlap. Axial attack from the top face results in a chair shaped intermediate, which is very stable and also lacks considerable steric hindrance. This gives rise to the axially linked product. On the other hand, thiyl attack from the bottom face results in an unfavorable boat conformation, which can rearrange to form a chair conformation. The final equatorial product is formed when this final chair conformation abstracts a proton. Thus the axially linked product is favored over the equatorially linked one due to intermediate stability (**Scheme 6**).

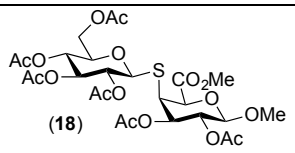
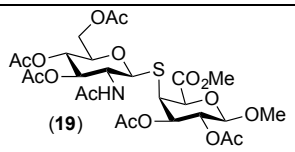
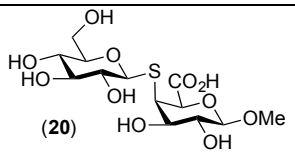
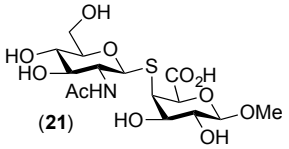
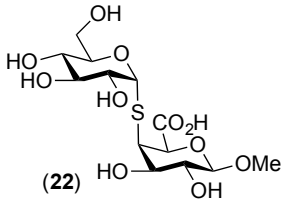


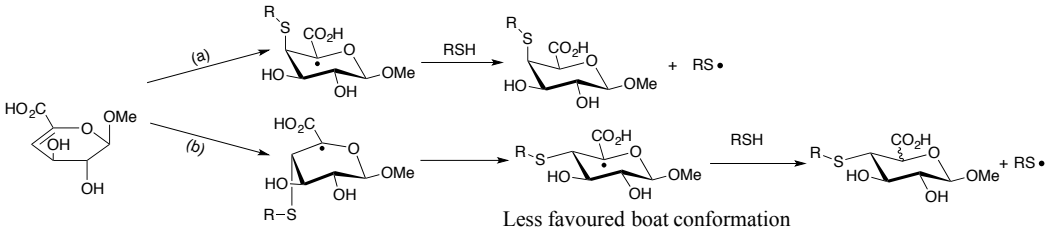
R_1 = R_2 = OAc, β-SH (2);
R_1 = NHAc, R_2 = OAc, β-SH (1);
R_1 = R_2 = OH, β-SH (3);
R_1 = NHAc, R_2 = OH, β-SH (4);
R_1 = R_2 = OH, α-SH (5)

R_3 = CO_2Me, R_4 = OAc (15);
R_3 = CO_2H, R_4 = OH (17)

R_1 = R_2 = R_4 = OAc, R_3 = CO_2Me, β-thiol (18)
R_1 = NHAc, R_2 = R_4 = OAc, R_3 = CO_2Me, β-thiol (19)
R_1 = R_2 = R_4 = OH, R_3 = CO_2H, β-thiol (20)
R_1 = NHAc, R_2 = R_4 = OH, R_3 = CO_2H, β-thiol (21)
R_1 = R_2 = R_4 = OH, R_3 = CO_2H, α–thiol (22)

Table 3: Model substrate library in both protected and deprotected conditions

Entry	Unsaturated sugar (Equiv.)	Thiosugar (Equiv.)	Condition	Photo Initiator (Equiv.)	Product	Yield
1	**15** (1)	**2** (8)	MeOH, 2 h. hν	DPAP (0.6 eq.)	 (**18**)	88%
2	**15** (1)	**1** (8)	MeOH, 3 h. hν	DPAP (0.6 eq.)	 (**19**)	84%
3	**17** (1)	**3** (10)	pH 4 acetate/formate buffer, 8 h. hν	Vazo-44 (0.8 eq.)	 (**20**)	62%
4	**17** (1)	**4** (10)	pH4 acetate/formate buffer, 8 h. hν	Vazo-44 (0.8 eq.)	 (**21**)	64%
5	**17** (1)	**5** (10)	pH 4 acetate/formate buffer, 8 h. hν	Vazo-44 (0.8 eq.)	 (**22**)	42%



RSH = glycosidic thiol

(a) Thiyl radical attack from top face (b) Thiyl radical attack from bottom face

Chair conformation of intermediate in case of thiol attack by step (a) is favoured over boat configuration of intermdediate in case of thiol attack by step (b)

Scheme 6: Attack of thiyl radical from top or bottom face of the unsaturated glycoside leads to formation of two types of intermediate.

2.3 Summary

In conclusion, we have established a thiol-ene radical approach for synthesis of *S*-linked disaccharides. In addition to protected sugars, deprotected sugars can also be directly coupled in aqueous conditions. This enables the formation of potentially biologically-relevant substrates, with the added advantage that such *S*-linked glycosides are bioisosteres that show enhanced stability towards enzymatic degradation over *O*-linked sugars[41].

The scope of this work can be further enhanced with other epimers of thioglucose and glucosamine to elucidate the effect of various structural variations on the reactivity. This would be especially interesting in the context of this work as other epimers of glucosamine can be used to potentially mimic other members of the glycosaminoglycan family other than heparin/heparan sulfate. Also the unsaturated sugar counterpart can be varied in terms of the position of the unsaturation inside the ring so that other linkages other than 1,4 can also be potentially established. Further optimization of these conditions would be needed for each of these epimers. Despite this, the effect of various functionalities on both the thiosugar and the unsaturated sugar

can be investigated to both explain reactivity trends and to readily build up various biologically relevant carbohydrate substrates.

2.4 Experimental Section

2.4.1 Synthesis of compounds

Melting points: Melting points were recorded on a Kofler hot block and are uncorrected.

NMR data: Proton nuclear magnetic resonance spectra (^{1}H) were recorded on a Bruker DPX 400 (400 MHz) and Bruker DQX 400 (400 MHz) spectrometer. Carbon nuclear magnetic resonance spectra (^{13}C) were recorded on a Bruker DQX 400 (101 MHz) spectrometer. Spectra were fully assigned based on chemical shifts using a combination of COSY, HSQC data and comparison with spectra of related compounds. All chemical shifts are referenced to the relevant residual solvent peak and are quoted in ppm with respect to the internal standard of tetramethylsilane (TMS). Resonances are described as s (singlet), d (doublet), t (triplet), q (quartet), quin (quintet), sxt (sextet), sept (septet), oct (octet), multiplet (m), app. (apparent) and br (broad). Coupling constants (*J*) are given in hertz (Hz).

*Mass spectra: M*ass spectra were recorded on a Fisons Platform VG Spectrometer (ESI) using electrospray ionisation in a positive (ESI+) or negative (ESI-) scan mode. *m/z* are reported with their percentage abundance.

Accurate mass Service

Analyses were performed using a Thermo Exactive mass spectrometer equipped with Waters Acquity liquid chromatography system. Instrument control and data processing were performed using Thermo Xcalibur Software. The system was calibrated on the day of the analysis and its mass accuracy with external calibration (as used for these experiments) is better than 5ppm for 24 hours following calibration. The mass spec was operated using the heated electrospray (HESI-II) probe and resolution was set to 50,000. Electrospray source conditions were adjusted to maximise sensitivity. A mixture of 10% water, 89.9% methanol and 0.1% formic acid was used to transport samples to the mass spectrometer at a flow rate of 0.2 mL/min.

Specific rotation: Specific rotations were measured on a Perkin Elmer 241 polarimeter with a path length of 1 dm with concentrations (*c*) given in g/100 mL.

Chromatography: Thin layer chromatography (TLC) was carried out on Merk Kieselgel $60F_{254}$ 0.2 mm precoated aluminium backed plates. Product spots were visualized with a combination of the following: 254 nm UV lamp, aqueous phosphomolybdic acid and $Ce^{(IV)}$ (2.5% phosphomolybdic acid hydrate, 1% cerium(IV) sulfate hydrate, and 6% H_2SO_4); ammonium molybdate (5% in 2M H_2SO_4). Retention factors (R_f) are reported with the solvent system used in parentheses. Flash column chromatography (FC) was carried out with Fluka Kiegselgel 60 220-440 mesh silica gel, with the solvent system used, in parentheses.

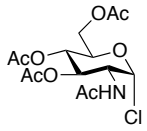
Solvents and reagents: Unless specified, reagents and starting materials were obtained from standard commercial sources and were used without further purification. DCM was dried by distillation over alumina. Remaining anhydrous solvents were purchased from Aldrich or Fluka. All other solvents were used as supplied (analytical or HPLC grade), without further purification.

'Petrol' refers to the fraction of petroleum ether boiling in the range 40-60 °C. Brine refers to a saturated aqueous solution of sodium chloride.

Reactions: All non-aqueous reactions were carried out in oven-dried glassware under an inert atmosphere of nitrogen where appropriate. Dowex ion exchange resins were activated by washing with methanol, water, 1 M HCl, and water until the filtrate was pH neutral.

Proton and carbon NMRs of compounds were assigned using a combination of proton, carbon, COSY, HSQC and 1D-TOCSY studies. Spectra of novel compounds are shown.

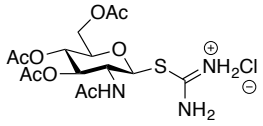
1. Synthesis of compound (6):



N-Acetyl-D-glucosamine (7.0 g, 31.64 mmol) was added to a solution of acetyl chloride (70.19 mL, 0.99 mol) under an atmosphere of N_2 and the reaction mixture was stirred for 19 h. at r.t. After this time, TLC (petrol/EtOAc 1:2) indicated the formation of product (R_f 0.4) with complete consumption of the starting material. The slightly discoloured solution mixture was then diluted with $CHCl_3$ (70 mL) and poured under strong stirring onto ice (28.0 g) and H_2O (20 mL). The phases were separated and the organic layer transferred to a beaker containing ice (14.0 g) and sat. $NaHCO_3$ solution (87.5 mL). The mixture was stirred, transferred to a separatory funnel and shaken until gas production ended. The phases were separated, the organic layer dried over Na_2SO_4, filtered and the concentrated *in vacuo*. The residue was then purified by flash column chromatography (EtOAc/petrol 3:2) giving the purified product in 51.4% yield. TLC: R_f 0.4 (ethyl acetate/petrol 2:1); m.p. = 130-132 °C [lit. value 133-134 °C][42]; $[\alpha]_D^{23}$ = +125.0 (c = 0.9, $CHCl_3$) [Lit. $[\alpha]_D^{18}$ = +127.0 (c = 1, $CHCl_3$)][43]; IR (neat): ν_{max} 3367, 3295, 3057,

2957, 1741, 1666, 1532, 1433, 1367, 1215, 1114, 1041, 909, 894, 735 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$):δ ppm 6.20 (d, $J_{1,2}$ = 3.54 Hz, 1H, H-1), 5.82 (d, $J_{NH,2}$ = 8.59 Hz, 1H, NH), 5.34 (dd, J = 9.6, $J_{3,2}$ = 12.38 Hz, $J_{3,4}$ = 9.6 Hz, 1H, H-3), 5.22 (dd, J = 9.85 Hz, $J_{4,3}$= 9.6 Hz, 1H, H-4), 4.54 (ddd, $J_{2,NH}$ = 8.59 Hz, $J_{2,1}$ = 3.54 Hz, $J_{2,3}$= 12.38 Hz, 1H, H-2), 4.32-4.25 (m,2H, CH/CH_2, H-5/H-6a), 4.14 (dd, J = 12.3 Hz, J = 2 Hz, 1H, H-6b), 2.11 (s, 3H, -OAc), 2.06 (s, 3H, -OAc), 2.06 (s, 3H, -OAc), 1.99 (s, 3H, $NHCOCH_3$); ^{13}C NMR (101 MHz, $CDCl_3$):δ ppm 171.51, 170.59, 170.08, 169.13 (4s, 3 x $COCH_3$, $NHCOCH_3$), 93.54 (C-1), 70.88 (C-5), 70.12 (C-3), 66.91 (C-4), 61.12 (C-6), 53.49 (C-2), 23.09 ($NHCOCH_3$), 20.67 ($COCH_3$), 20.54 ($COCH_3$); ESI^+ LRMS: calcd. for $C_{14}H_{20}{}^{35}ClNO_8$ ($M+Na^+$) 388.08, found 388.1, calcd. for $C_{14}H_{20}{}^{37}ClNO_8$ ($M+Na^+$) 390.08, found 390.1 ($C_{14}H_{20}{}^{35}ClNO_8$: $C_{14}H_{20}{}^{37}ClNO_8$ = 2.88:1)

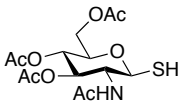
2. Synthesis of compound (7):



2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-glucopyranosyl chloride (5.94 g, 16.24 mmol) and thiourea (2.19 g, 28.88 mmol) were dissolved in acetone (65 mL) under N_2. Thereafter, the reaction mixture was heated to 60 °C. A white precipitate was observed at the end of 2 h. which was filtered off and the filtrate was returned to reflux. The process was repeated until the precipitate ceased to form. Thereafter, the solids were recrystallized from acetone/petrol to give 2-acetamido-3,4,6-tri-*O*-acetyl-β-D-glucopyranosyl-1-isothiouronium chloride in 65% yield. TLC: R_f 0.0 (100% EtOAc), m.p. = 147-153 °C [lit. value 135 °C][44]; $[\alpha]_D{}^{20}$ = -18.9 (c = 1.0, H_2O) [Lit. $[\alpha]_D{}^{25}$ = -29.3 (c =1.0, H_2O)][45]; IR (neat): $\nu_{max.}$ 3230, 3053, 1751, 1654, 1543, 1654, 1543, 1451, 1370, 1303, 1226, 1098, 1028, 909 cm^{-1}; ^{1}H NMR(400 MHz, DMSO-d_6): δ ppm 9.41, 9.21 (2s, 2 x NH_2), 8.43 (d, $J_{NH,2}$ = 9 Hz, 1H, NH), 5.68 (d, $J_{1,2}$ = 9.7 Hz,1H, H-1), 5.13 (t,

J = 9.8 Hz, 1H, H-3), 4.94 (t, J = 9.7 Hz, 1H, H-4), 4.26-4.15 (m, $J_{6a,\ 6b}$ = 12.2 Hz, $J_{6a,5}$ 4.8 Hz, 2H, H-5/H-6a), 4.09-3.97 (m, $J_{6b,5}$ = 1.7 Hz, $J_{6b,\ 6a}$ = 12 Hz, 2H, H-2/H-6b), 2.02, 1.98, 1.94 (3s, 3 x $COCH_3$), 1.81 (s, 3H, $NHCOCH_3$); ^{13}C NMR (101 MHz, DMSO-d_6):δ ppm 170.1, 169.9, 169.6, 169.3 (4s, 3 x $COCH_3$, $NHCOCH_3$), 80.6 (s, 1C., C-1), 74.7 (s, 1C, C-5), 72.7 (s, 1C, C-3), 67.9 (s, 1C, C-4), 61.5 (s, 1C, C-6), 52 (s, 1C, C-2), 22.5 (s, 1C, $NHCOCH_3$), 20.6, 20.4, 20.3 (3s, 3 x $COCH_3$); ESI^- LRMS: calcd. for $C_{15}H_{24}ClN_3O_8S$ ($M\text{-}Cl^-$) 406.1, found 406.1

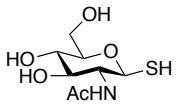
3. Synthesis of compound (1):



2-Acetamido-3,4,6-tri-*O*-acetyl-β-D-glucopyranosyl-1-isothiouronium chloride (4.39 g, 9.95 mmol) and sodium metabisulfite (3.22 g, 16.9 mmol) were added to a stirred mixture of DCM (70 mL) and water (30 mL). The mixture was heated to reflux under an atmosphere of N_2. After 2 h, when TLC (ethyl acetate, 100%) showed the formation of the product (R_f > 0) with complete consumption of the starting material (R_f 0.0), the reaction mixture was cooled to RT and the phases separated. The organic layer was re-extracted with DCM (2 x 50 mL). The combined organic layers were washed with water (50 mL), brine (50 mL), dried over Na_2SO_4, filtered and the solvent removed *in vacuo*. The resulting white solid was purified by flash column chromatography (EtOAc/petrol 7:3). The product was obtained in 80% yield. TLC: R_f 0.24 (ethyl acetate/petrol 7:3); m.p. = 52-153 °C [lit. value 166-168 °C][3]; $[\alpha]_D^{23}$ = -25.8 (c 1.0, $CHCl_3$) [Lit. $[\alpha]_D^{20}$ = -25.3 (c 1.0, $CHCl_3$)][3]; IR (neat): ν_{max} 3284, 3078, 2942, 2560, 2360, 1743, 1662, 1545, 1370, 1227, 1043, 960, 905 cm^{-1};^1H NMR (400 MHz, $CDCl_3$): δ ppm 5.73 (d, $J_{NH,2}$= 9.09 Hz, 1H, N<u>H</u>), 5.15-5.06 (m, 2H,H-3/H-4), 4.59 (dd, $J_{1,2}$ = 9.8 Hz, $J_{1,SH}$ = 9.6 Hz, 1H, H-1), 4.24 (dd, $J_{6a,6b}$ = 12.38 Hz, $J_{6a,\ 5}$= 4.8 Hz, 1H, H-6a), 4.16-4.09 (m, $J_{6a,6b}$= 12.38 Hz, $J_{5,6b}$= 2.27 Hz,

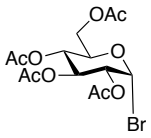
$J_{2,1}$ = 9.8 Hz, 2H, H-2/H-6b), 3.70 (ddd, $J_{5,6a}$ = 4.8 Hz, $J_{5,6b}$ = 2.27 Hz, $J_{5,4}$ = 9.6 Hz, 1H, H-5), 2.57 (d, $J_{SH,1}$ = 9.35 Hz, 1H, SH), 2.1 (s, 3H, -OAc), 2.04 (s, 3H, -OAc), 2.03 (s, 3H, -OAc), 1.99 (s, 3H, $NHCOCH_3$); ^{13}C NMR(101 MHz, $CDCl_3$) : δ ppm 171.25, 170.76, 170.42, 169.23 (4s, 3 x $COCH_3$, $NHCOCH_3$), 80.34 (C-1), 76.26 (C-5), 73.47, 68 (2s, 2C,C-3/C-4), 62.13 (C-6), 56.78 (C-2), 23.31 ($NHCOCH_3$), 20.79 ($-COCH_3$), 20.68 ($-COCH_3$), 20.60 ($-COCH_3$); ESI^+ LRMS: calcd. for $C_{14}H_{21}NO_8S$ ($M-H^+$) 362.09, found 362.0

4. Synthesis of compound (4):



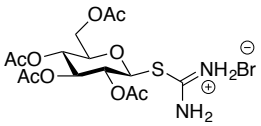
2-Acetamido-3,4,6-tri-*O*-acetyl-1-thio-β-D-glucopyranoside (700 mg, 1.93 mmol) was taken in a round bottomed flask in 7 mL MeOH and to it was added sodium methoxide (25% wt solution, 0.57 mL). The reaction mixture was stirred for 15 mins. When TLC (EtOAc, 100%) indicated complete consumption of the starting material ($R_f > 0$) and the formation of the product ($R_f = 0$), the reaction was stopped and neutralized with DOWEX® 50WX8-200. Thereafter, it was filtered and concentrated i*n vacuo* to afford 2-acetamido-2-deoxy-1-thio-β-D-glucopyranoside (439 mg, 96%) as a white solid. R_f = 0.5 (H_2O/propan-2-ol/ethyl acetate; 1:2:2); m.p. 158-160 °C [lit. value 167-168 °C][3]; $[\alpha]_D^{23}$= -12.9 (c 1.0, MeOH) [Lit. $[\alpha]_D^{22}$ =-10.4 (c 1.0, MeOH][46]; IR(neat): ν_{max} 3290, 2442, 1636, 1568, 1434, 1376, 1319, 1054, 998 cm^{-1}; ^1H NMR (400 MHz, D_2O): δ ppm 4.6 (d, $J_{1,2}$=10.2 Hz, 1H, H-1), 3.82 (dd, $J_{6b,\ 6a}$ = 12.3 Hz, $J_{6b,5}$ = 1 Hz, 1H, H-6b), 3.72-3.62 (m, 2H, H-2/H-6a), 3.40-3.30 (m, 3H, H-3/H-4/H-5), 2.00 (s, 3H, $NHCOCH_3$); ^{13}C NMR (101 MHz, D_2O):δ ppm 174.9 (s, 1C, $NHCOCH_3$), 80.5 (s, 1C, C-4 or C-5), 79.3 (s, 1C, C-1), 75.2 (s, 1C, C-3), 69.9 (s, 1C, C-4 or C-5), 61.0 (s, 1C, C-6), 58.2 (s, 1C, C-2), 22.5 (s, 1C, $NHCOCH_3$);ESI^-LRMS: calcd. for $C_8H_{15}NO_5S$ ($M+Cl^-$) 272.7, found 272.0

5. Synthesis of compound (8):



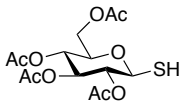
Pentaacetyl-β-D-glucopyranose (5.0 g, 12.81 mmol) was dissolved in DCM (37.5 mL) and to it was added hydrogen bromide (33% in acetic acid, 30 mL). The reaction mixture was stirred at room temperature for 2 h, under N_2. At the end of this time, TLC (EtOAc/petrol 1:1) indicated the complete consumption of the starting material (R_f 0.5) and formation of the product (R_f 0.6). Thereafter, the reaction mixture was partitioned between DCM (50 mL) and H_2O (50 mL), and the aqueous layer was re-extracted with DCM (3 x 50 mL). The combined organic layers were washed with sodium bicarbonate (250 mL), brine (125 mL), dried over Na_2SO_4 and concentrated *in vacuo* to give the product as a white powder. Yield: 84%. TLC: R_f 0.6 (ethyl acetate/petrol 1:1); m.p. = 83-84 °C [lit. value 89-91 °C][47]; $[\alpha]_D^{23}$= +192.7 (c 1.0, $CHCl_3$) [Lit. $[\alpha]_D^{25}$ = +197.0 (c 1.0, $CHCl_3$)][48]; IR (neat) : ν_{max}2964, 1749, 1368, 1216, 1111,1078, 1041, 975, 911 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$): δ ppm 6.62 (d, $J_{1,2}$ = 4.04 Hz, 1H, H-1), 5.56 (dd, $J_{3,4}$ = 9.6 Hz, $J_{3,2}$ = 9.85 Hz, 1H, H-3), 5.17 (dd, $J_{4,3}$ = 9.6 Hz, $J_{4,5}$ = 10.11 Hz, 1H, H-4), 4.84 (dd, $J_{2,1}$ = 4.04 Hz, $J_{2,3}$ = 9.85 Hz, 1H, H-2), 4.32 (m, 2H, CH/CH_2, H-5/H-6a), 4.14 (d, J = 10.86 Hz, 1H, H-6b), 2.11 (s, 3H, -OAc), 2.10 (s, 3H, -OAc), 2.06 (s, 3H, -OAc), 2.04 (s, 3H, -OAc); ^{13}C NMR (101 MHz, $CDCl_3$) :δ ppm 170.5, 169.8, 169.8, 169.4 (4s, 4 x $COCH_3$), 86.5 (C-1), 70.5 (C-2), 70.1 (C-3), 67.1 (C-4), 72.1 (C-5), 60.9 (C-6), 20.6 (-$COCH_3$), 20.6 (-$COCH_3$), 20.6 (-$COCH_3$), 20.5 (-$COCH_3$); ESI^+ LRMS: calcd. for $C_{14}H_{19}{}^{79}BrO_9$ ($M+Na^+$) 433.01, found 433.0, calcd. for $C_{14}H_{19}{}^{81}BrO_9$ ($M+Na^+$) 435.01, found 435.01 ($C_{14}H_{19}{}^{79}BrO_9$: $C_{14}H_{19}{}^{81}BrO_9$ = 1.07:1)

6. Synthesis of compound (9):



Thiourea (1.23 g, 16.12 mmol) and 2,3,4,6-Tetra-*O*-acetyl-α-D-glucopyranosyl bromide (4.42 g, 10.75 mmol) were dissolved in acetone (25 mL) under N_2. The reaction mixture was heated to 60 °C for 2 h. after which a white precipitate appeared. It was removed by filtration and the filtrate returned to reflux. The process was repeated until the precipitate ceased to form. It was then washed with acetone/petrol to give a white solid. Yield: 62% TLC: R_f 0.62 (ethyl acetate/2-propanol/water 5:4:1); m.p. = 201-202 °C [lit. value 191 °C][49]; $[\alpha]_D^{23}$= -7.2 (c. 1.0, H_2O) [Lit. $[\alpha]_D^{25}$ -7.6 (c 1.4, H_2O)][3]; IR (neat): ν_{max}3270, 3055, 1742, 1653, 1369, 1218, 1053, 1032, 906 cm^{-1};^{1}H NMR (400 MHz, DMSO-d_6):δ ppm 9.25 (br s, 2H, NH_2), 9.07 (br s, 2H, NH_2), 5.69 (d, $J_{1,2}$ = 10.11 Hz, 1H, H-1), 5.32 (dd, *J* 9.35 Hz, *J* 9.1 Hz,1H, H-3), 5.11 (m, 2H, H-2/H-4), 4.22-4.16 (m, 2H, H-5/ H-6a), 4.11-4.06 (m, 1H, H-6b), 2.05, 2.02, 2.00, 1.97 (4s, 4 x $COCH_3$); ^{13}C NMR(101 MHz, DMSO-d_6):δ ppm 170.8, 170.3, 170.2, 170.1, 167.1 (4 x $COCH_3$, 1 x C=N), 80.5 (C-1), 76.1 (C-5), 73.2 (C-3), 69.5, 68.2 (2s, 2C, C-2/C-4), 62.4 (C-6), 21.4 (-$COCH_3$), 21.2 (-$COCH_3$), 21.1 (-$COCH_3$), 21.1 (-$COCH_3$); ESI^+ LRMS: calcd. for $C_{15}H_{23}N_2O_9SBr$ ($M\text{-}Br^-$) 407.4, found 407.1

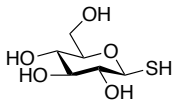
7. Synthesis of compound (2):



$Na_2S_2O_5$ (1.77 g, 9.34 mmol) and 2,3,4,6-Tetra-*O*-acetyl-β-D-glucopyranosyl-1-isothiouronium bromide (3.25 g, 6.67 mmol) were added to a stirred mixture of DCM (50 mL) and water (25 mL) and heated to reflux under N_2. After 3 h., TLC (EtOAc/petrol 1:1) indicated the formation

of product (R_f 0.6) with complete consumption of the starting material (R_f 0.0). Thereafter, the reaction mixture was cooled to RT and the phases separated. The aqueous layer was re-extracted with DCM (2 x 50 mL), dried over Na_2SO_4, filtered and concentrated *in vacuo* to give the product as a white crystalline solid. Yield: 94%. TLC: R_f 0.6 (ethyl acetate/petrol 1:1); m.p. = 113-116 °C [lit. value 113-115 °C][50]; $[\alpha]_D^{23}$= +3 (c. 1.0, $CHCl_3$) [Lit. $[\alpha]_D^{25}$ =+6.5 (c 1.0, $CHCl_3$)][51]; IR(neat): ν_{max} 2952, 2561, 1749, 1370, 1223, 1040 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$):δ ppm 5.20 (dd, $J_{3,2}$ = 9.35 Hz, $J_{3,4}$ = 9.35 Hz, 1H, H-3), 5.11 (dd, $J_{4,5}$ = 9.60 Hz, $J_{4,3}$ = 9.35 Hz, 1H, H-4), 4.98 (dd, $J_{2,3}$ = 9.35 Hz, $J_{2,1}$= 9.85 Hz, 1H, H-2), 4.55 (dd, $J_{1,2}$ = 9.85 Hz, $J_{1,SH}$= 10.11 Hz, 1H, H-1), 4.25 (dd, $J_{6a,6b}$ = 12.38 Hz, $J_{6a,5}$ = 4.8 Hz, 1H, H-6a), 4.13 (dd, $J_{6b,6a}$ = 12.38 Hz, $J_{6b,5}$ = 2.02 Hz, 1H, H-6b), 3.73 (ddd, $J_{5,4}$ = 9.85 Hz, $J_{5,6a}$ = 4.8 Hz, $J_{5,6b}$ = 2.02 Hz,1H, H-5), 2.32 (d, $J_{SH,1}$ = 10.11 Hz, 1H, SH), 2.10 (s, 3H, -OAc), 2.09 (s, 3H, -OAc), 2.03 (s, 3H, -OAc), 2.02 (s, 3H, -OAc); ^{13}C NMR (101 MHz, $CDCl_3$):δ ppm 170.6, 170.1, 169.6, 169.3 (4 x $COCH_3$), 78.7 (C-1), 76.3 (C-5), 73.5, 73.5 (2s, 2C, C-2/C-3), 68 (C-4), 61 (C-6), 20.7, 20.7, 20.5, 20.5 (4 x $COCH_3$); ESI^+ LRMS: calcd. for $C_{14}H_{20}O_9S$ ($M+Na^+$) 387.3, found 387.1

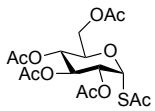
8. Synthesis of compound (3):



2,3,4,6-Tetra-*O*-acetyl-1-thio-β-D-glucopyranose (500 mg, 1.37 mmol) was taken in a round bottomed flask and dissolved in 10 mL MeOH. 0.36 ml NaOMe solution in MeOH (4.2 M) was then added to it dropwise and it was stirred for 15 mins. when TLC (EtOAc, 100%) indicated the complete consumption of the starting material (R_f > 0) and the formation of the product (R_f = 0). At the end of the reaction, it was neutralized using DOWEX® 50WX8-200, filtered and concentrated *in vacuo*. It was subsequently lyophilized to give a yellowish oil. Yield: 95%. TLC:

R_f0.3 (EtOAc/Ispropanol/Water 2:2:1); $[\alpha]_D^{23}$ = +33.2 (c. 1.0, H_2O) [Lit. $[\alpha]_D^{25}$ = +35.7 (c 1.0, H_2O)][3];IR (neat): ν_{max} 3264, 2360, 2341, 1636, 1033 cm^{-1};[1]H NMR (400 MHz, D_2O):δ ppm 4.50 (d, 1H, $J_{1,2}$ = 9.09 Hz, H-1), 3.84 (d, 1H, $J_{4,5}$ = 12.38 Hz, H-4), 3.64 (dd, 1H, $J_{5,4}$ = 12.38 Hz, $J_{5,6}$ = 5.56 Hz, H-5), 3.43-3.34 (m, 3H, H-3, H-6a/6b), 3.19 (t, 1H, $J_{1,2}$ = 9.09 Hz, H-2) ^{13}C NMR (101 MHz, D_2O):δ ppm 80.50 (C-1), 77.18 (C-3), 76.21 (C-2), 69.78 (C-6), 61.06 (C-4/C-5); ESI^+ LRMS: calcd. for $C_6H_{12}O_5S$ ($M+Na^+$) 219.2, found 219.1

9. Synthesis of compound (11):

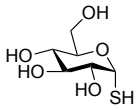


To pentaacetyl-β-D-glucopyranose (5.0 g, 12.8 mmol) and phosphorus pentachloride (2.9 g, 14.1 mmol) under an atmosphere of N_2 was added 20 µL of $BF_3.Et_2O$. After 10 mins., TLC (EtOAc/petrol 1:1) indicated the formation of the product (R_f = 0.6) with complete consumption of the starting material (R_f = 0.5). The reaction mixture was partitioned between DCM (40 mL) and water (70 mL) and the aqueous phase was re-extracted with DCM (2 x 50 mL). The combined organic phases were subsequently washed with sat. sodium bicarbonate, brine, dried over Na_2SO_4 and concentrated *in vacuo.* The crude product **10** was directly taken over to the next step without further purification.

Crude 2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl chloride (4.75 g, 12.95 mmol) was taken in DMPU (24 mL) and potassium thioacetate (1.63 g, 14.24 mmol) was added to it at 0 °C under an atmosphere of N_2. After 3 days, TLC (EtOAc/petrol 1:1) indicated the formation of product (R_f = 0.4) with complete consumption of the starting material (R_f = 0.6). The reaction mixture was then partitioned between EtOAc (100 mL) and water (100 mL) and the organic phase was re

extracted with water (2 x 100 mL). The combined aqueous layers were washed with ethyl acetate (100 mL). Thereafter, the combined organic layers were washed with brine, dried over Na_2SO_4 and concentrated *in vacuo*. The resulting residue was purified by flash column chromatography (EtOAc/petrol 3:7) to yield the product in the form of a peach coloured solid. Yield: 62%; R_f = 0.4 (EtOAc/petrol 1:1); m.p. = 118-120 °C [lit. value 128-129 °C][3]; $[\alpha]_D^{23}$ = +126 (c 1.0, $CHCl_3$) [Lit. $[\alpha]_D^{20}$ = +135 (c 1.0, $CHCl_3$)][3]; IR (neat): ν_{max} 2946, 2360, 1748, 1710, 1367, 1218, 1069 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$): δ ppm 6.23 (d, $J_{1,2}$ = 5.31 Hz, 1H, H-1), 5.24 (dd, $J_{2,1}$ = 5.31 Hz, $J_{2,3}$ = 10.11 Hz, 1H, H-2), 5.18 (t, $J_{3,2}$ = 10.11 Hz, $J_{3,4}$ = 9.09 Hz, 1H, H-3), 5.10 (t, $J_{4,3}$ = 9.09 Hz, $J_{4,5}$ = 9.60 Hz, 1H, H-4), 4.28 (dd, $J_{6a,6b}$ = 12.38 Hz, $J_{6a,5}$ = 4.04 Hz, 1H, H-6a), 4.05 (dd, $J_{6b,6a}$ =12.38 Hz, $J_{6b,5}$ = 1.52 Hz, 1H, H-6b), 3.97 (ddd, $J_{5,4}$ = 9.6 Hz, $J_{5,6b}$ = 1.52 Hz, $J_{5,6a}$ = 4.10 Hz, 1H, H-5), 2.43 (s, 3H, -OAc), 2.09 (s, 3H, -OAc), 2.03 (s, 3H, -OAc), 2.03 (s, 3H, -OAc), 2.02 (s, 3H, -OAc); ^{13}C NMR (101 MHz, $CDCl_3$): δ ppm 191.3 (-SAc), 170.6 (-OAc), 170.0 (-OAc), 169.4 (-OAc), 169.3 (-OAc), 80.2 (C-1), 68.9 (C-2), 71.4 (C-3/C-5), 71.1 (C-3/C-5), 67.7 (C-4), 61.4 (C-6), 31.4 (-SAc), 20.6, 20.6, 20.6, 20.5 (4 x –OAc); ESI^+ LRMS: calcd. for $C_{16}H_{22}O_{10}S$ ($M+Na^+$) 429.3, found 429.1

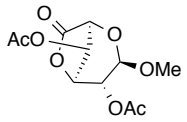
10. Synthesis of compound (5):



To a solution of 1-S-acetyl-2,3,4,6-tetra-*O*-acetyl-1-thio-α-D-glucopyranoside (1.0 g, 2.7 mmol, 1 eq) in MeOH was added sodium methoxide (25% wt. solution, 0.8 mL). The reaction mixture was stirred at RT. After 10 minutes, TLC (EtOAc; 100 %) indicated the formation of product (R_f 0.0) with complete consumption of the starting material (R_f > 0.0). The reaction mixture was

acidified with DOWEX® 50WX8-200, filtered and concentrated *in vacuo* to afford 1-thio-α-D-glucose (α-GlcSH) (0.54 g, 96 %) as a pale yellowish solid. TLC:R_f 0.3 (H_2O/propan-2-ol/EtOAc; 1:2:2); ^{1}H NMR (400 MHz, D_2O): δ ppm 3.3 (dd, $J_{4,3}$ = 9.3 Hz, $J_{4,5}$ = 9.9 Hz, 1H, H-4), 3.54 (dd, $J_{3,4}$ = 9.3 Hz, $J_{3,2}$ = 9.7 Hz, 1H, H-3), 3.71 (dd, $J_{6a,5}$ = 5.5 Hz, $J_{6a,6b}$ = 12.2 Hz, 1H, H-6a), 3.72 (dd, $J_{2,1}$ = 5.5 Hz, $J_{2,3}$ = 9.7 Hz, 1H, H-2), 3.79 (dd, $J_{6b,6a}$ = 12.2 Hz, $J_{6b,5}$ = 2.3 Hz, 1H, H-6b), 3.94 (ddd, $J_{5,4}$ = 9.9 Hz, $J_{5,6a}$ = 5.5 Hz, $J_{5,6b}$ = 2.3 Hz, $J_{5,4}$ = 9.9 Hz, 1H, H-5), 5.58 (d, $J_{1,2}$ = 5.5 Hz, 1H, H-1); ^{13}C NMR (101 MHz, $CDCl_3$): δ ppm 61.0 (C-6), 69.7 (C-4), 70.6 (C-2), 72.5 (C-5), 72.9 (C-3), 80.0 (C-1); ESI$^-$ LRMS: calcd. for $C_6H_{12}O_5S$ (M-H$^+$) 195.2, found 195.0; m.p. = 119 ºC; $[\alpha]_D^{23}$ = +104 (c=1.0, H_2O)

11. Synthesis of compound (13):

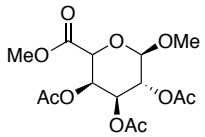


Methyl-β-D-galactopyranoside (1.0 g, 5.15 mmol), NaBr (265 mg, 2.58 mmol) and TEMPO (100 mg, 0.58 mmol) were dissolved in distilled water (50 mL) and cooled to 0 °C using an ice-bath. NaOCl soln. (12 mL, 12.5%) was added dropwise keeping the pH within 10-11 with 0.5 M NaOH. Thereafter the ice-bath was removed and the solution was stirred for 23 h. The water was then evaporated completely and the crude product **12**, along with the salts was taken over directly to the next step for acetylation. TLC: R_f 0.34 (DCM/MeOH/H_2O 6:4:1)

Crude galactopyranosiduronic acid (1.3 g) was added to a soln. of acetic anhydride (16 mL, 0.17 mmol) along with sodium acetate (500 mg, 6.09 mmol) and the reaction mixture was heated to reflux for 15 mins. After cooling to room temperature, the excess acetic anhydride was evaporated completely. Sat. sodium bicarbonate solution was then used for neutralizing the

reaction mixture and it was extracted with DCM. The organic phase was dried over Na_2SO_4, filtered and purified using flash column chromatography (EtOAc/Petrol 3:7). Yield (for 2 steps): 35%; TLC: R_f 0.31 (EtOAc/Petrol 3:7); m.p. = 107-108 °C [lit. value 112-113 °C][18]; $[\alpha]_D^{23}$ = -167 (c 0.2, $CHCl_3$) [Lit. $[\alpha]_D^{23}$= -172 (c 0.2, $CHCl_3$)][52]; IR(neat): ν_{max} 3400, 2361, 1805, 1746, 1370, 1219, 1153, 1058, 897 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$): δ ppm 5.35 (s, 1H, H-4), 5.25 (d, $J_{2,3}$ = 4.7 Hz, 1H, H-3), 4.9 (d, $J_{2,3}$ = 4.7 Hz, 1H, H-2), 4.7 (s,1H, H-1), 4.1 (s, 1H, H-5), 3.4 (s, 3H, -OC$\underline{H}_3$) 2.18 (s, 3H, -COC$\underline{H}_3$) 2.12 (s, 3H, -COC$\underline{H}_3$); ^{13}C NMR(101 MHz, $CDCl_3$): δ ppm 171.90 (C-6), 169.61 (-OAc), 169.11 (-OAc), 100.91 (C-1), 77.61 (C-2), 71.73 (C-4), 70.79 (C-3), 69.88 (C-5), 56.12 (-OMe), 20.67 (-OAc), 20.67 (-OAc); ESI^+ LRMS: calcd. for $C_{11}H_{14}O_8$ ($M+Na^+$) 297.2, found 297.1

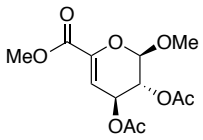
12. Synthesis of compound (14):



Methyl 2,4-di-*O*-acetyl-β-D-galactopyranosiduronic acid, γ-lactone (220 mg, 0.91 mmol) was taken and DBU (9.5 µL, 0.06 mmol) was added to it in MeOH (9 mL). The reaction mixture was stirred for 19 h. at room temperature. After the reaction was completed, the solvent was removed *in vacuo.* The crude product was dissolved in pyridine (3 mL, 37.16 mmol), Ac_2O (3 mL, 31.8 mmol), a catalytic amount of DBU was added and the reaction mixture stirred for 8 h. Thereafter, the compound was concentrated *in vacuo* and it was purified using flash column chromatography (EtOAc/Petrol 2:3). Yield: 92%; TLC: R_f0.2 (EtOAc/Petrol 3:7); m.p. = 114-115 °C [lit. value 118-120 °C][18]; $[\alpha]_D^{23}$= +14. 3 (c 2.8, $CHCl_3$) [Lit. $[\alpha]_D^{23}$ = +15.2 (c 3.0, $CHCl_3$)][53]; IR (neat): ν_{max} 3640, 2956, 2361, 1748, 1439, 1370, 1144, 975 cm^{-1}; ^1H NMR(400 MHz, $CDCl_3$):δ ppm 5.71 (s,1H, H-4), 5.26 (dd, J = 9.85 Hz, $J_{1,2}$ = 8.08 Hz, 1H, H-2), 5.06 (d, J

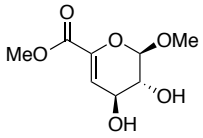
= 10.36 Hz, 1H, H-3), 4.42 (d, $J_{1,2}$ = 8.08 Hz, 1H, H-1), 4.3 (s, 1H, H-5), 3.77 (s, 3H, COOC$\underline{H}_3$), 3.58 (s, 3H, -OC$\underline{H}_3$), 2.12 (s, 3H, -COC$\underline{H}_3$) 2.07 (s, 3H, -COC$\underline{H}_3$), 2.0 (s, 3H, -COC$\underline{H}_3$);^{13}C NMR(101 MHz, $CDCl_3$): δ ppm 170.09 (-OAc), 169.88 (-OAc), 169.37 (-OAc), 166.48 (C-6), 101.99 (C-1), 72.35 (C-5), 70.56 (C-3), 68.39 (C-4), 68.32 (C-2), 57.36 (-OMe), 52.81 (-COOMe), 20.78 (-OAc), 20.61 (-OAc), 20.57 (-OAc);ESI$^+$ LRMS: calcd. for $C_{14}H_{20}O_{10}$ (M+Na^+) 371.2, found 371.0

13. Synthesis of compound (15):



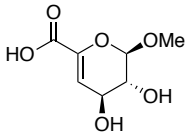
Methyl-(methyl 2,3,4-tri-*O*-acetyl-β-D-galactopyranosid)uronate from the previous step (220 mg, 0.63 mmol) was taken and DBU (1.02 mL, 6.82 mmol) was added to it after making a solution in 5 mL dry DCM. The reaction mixture was then stirred overnight. Subsequently, it was concentrated *in vacuo* and purified by flash column chromatography (EtOAc/Petrol 1:3). Yield: 94%. TLC: R_f 0.58 (EtOAc/Petrol 1:1); m.p. = 88-90 °C [lit. value 91-93 °C][26]; $[\alpha]_D^{20}$ = + 46.2 (c 0.9, $CHCl_3$);IR (neat): ν_{max} 2958, 1735, 1657, 1439, 1371, 1218, 1124, 1059, 1026, 887 cm^{-1}; ^{1}H NMR(400 MHz, $CDCl_3$): δ ppm 6.15 (dd, 1H, $J_{3,4}$ = 4.55 Hz, J = 1.01 Hz, H-4), 5.13 (dd,1H, $J_{3,4}$ = 4.55 Hz, J = 1.26 Hz, H-3), 5.06 (d, $J_{1,2}$ = 2.53 Hz, 1H, H-1), 5.01 (dd, 1H, $J_{1,2}$ = 2.53 Hz, J = 3.54 Hz, H-2), 3.77 (s, 3H, -COOC$\underline{H}_3$), 3.45 (s, 3H, 1-OC$\underline{H}_3$), 2.04 (s, 3H, -OCOC$\underline{H}_3$), 2.02 (s, 3H, -OCOC$\underline{H}_3$). ^{13}C NMR (101 MHz, $CDCl_3$):δ ppm 170.03 (-OAc), 169.44 (-OAc), 162.26 (C-6), 142.24 (C-5), 107.27 (C-4), 97.66 (C-1), 68.39 (C-2), 64.24 (C-3), 57.03 (-OMe), 52.63 (-COOMe), 20.94 (OAc), 20.76 (OAc); ESI$^+$ LRMS: calcd. for $C_{12}H_{16}O_8$ (M+K^+) 327.3, found 327.

14. Synthesis of compound (16):



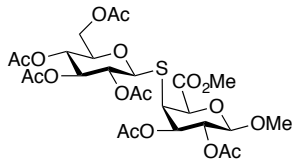
Methyl-(methyl 2,3-di-*O*-acetyl-4-deoxy-α-L-threo-hex-4-enopyranosid)uronate (670 mg, 2.32 mmol.) was taken in a round bottomed flask in 10 mL MeOH and to it was added sodium methoxide (25% wt. soln, 0.11 mL). The solution was stirred for 1 hr. When TLC analysis showed the complete consumption of the starting material and formation of the product, the reaction was stopped, neutralized with DOWEX® 50WX8-200, filtered and concentrated *in vacuo* to give methyl-(methyl 4-deoxy-α-L-threo-hex-4-enopyranosid)uronate as a colourless oil (465 mg, 98%); R_f = 0.38 (MeOH/$CHCl_3$ 1:9); $[\alpha]_D^{20}$= -48.4 (c 0.9, MeOH); IR : ν_{max} 3413, 2954, 2360, 2341, 1724, 1650, 1440, 1314, 1167, 1076991 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$):δ ppm 6.29 ppm (m, 1H, H-4), 5.13 (dd, J = 2.53 Hz, $J_{1,2}$ = 2.02 Hz, 1H, H-1), 4.00 (dd, J = 2.78 Hz, $J_{2,1}$ = 2.02 Hz, 2H, H-2/H-3), 3.85 (s, 3H, -COOMe), 3.53 (s, 3H, -OMe); ^{13}C NMR (101 MHz, $CDCl_3$):δ ppm 162.6 (s, 1C, -COOMe), 139.7 (s, 1C, C-5), 111.9 (s, 1C, C-4), 100.9 (s, 1C, C-1), 68.9, 65.3 (2s, 2C, C-2/C-3), 57.01 (s, 1C, -OCH_3), 52.5 (s, 1C,-COOCH_3); ESI^+ LRMS: calcd. for $C_8H_{12}O_6$ ($M+Na^+$) 227.1, found 227.1

15. Synthesis of compound (17):



Methyl-(methyl 2,3-di-*O*-acetyl-4-deoxy-α-L-threo-hex-4-enopyranosid)uronate (670 mg, 2.32 mmol.) was dissolved in LiOH solution (37 mL, 0.3 M, MeOH/water/THF = 5:4:1) at 0 °C and stirred for 4 h. After complete conversion, the reaction mixture was neutralized with DOWEX® 50WX8-200, filtered and concentrated *in vacuo*. It was subsequently lyophilized to give a yellowish liquid (433 mg, 98%) TLC: R_f = 0.15 (MeOH/DCM/H_2O5:10:1); $[\alpha]_D^{20}$ = -50.4 (c 0.4, MeOH);IR: ν_{max} 3291, 2942, 1590, 1401, 1261, 1131, 1044, 989, 960, 854, 791 cm^{-1}; ^{1}H NMR (400 MHz, D_2O):δ ppm 5.7 (s, 1H, H-4), 4.9 (s, 1H, H-1), 4.05 (s, 1H, H-3), 3.65 (s, 1H, H-2), 3.40 (s, 3H, OMe); ^{13}C NMR (101 MHz, D_2O):δppm 169.6 (s, 1C, C-6), 144.7 (s, 1C, C-5), 107.6 (s, 1C, C-4), 101.2 (s, 1C, C-1), 70.1 (s, 1C, C-3), 66.7 (s, 1C, C-2), 56.8 (s, 1C, OMe);ESI$^-$ LRMS: calcd. for $C_7H_{10}O_6$ (M-H^+) 189.1, found 189.0

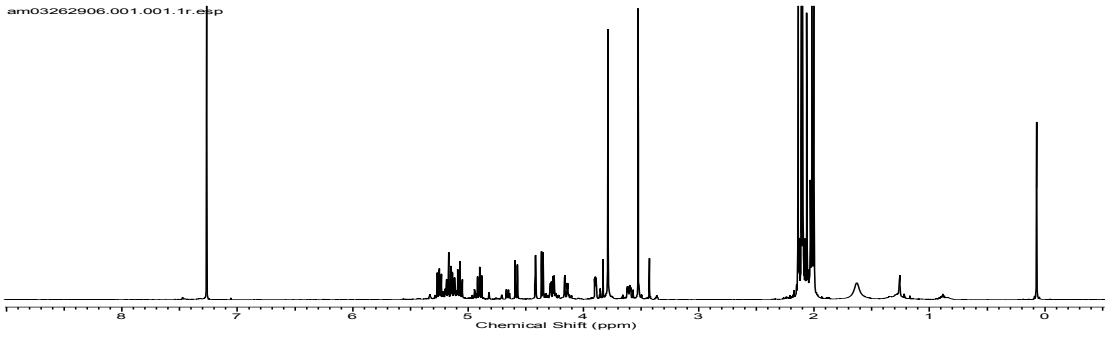
16. Synthesis of compound (18):



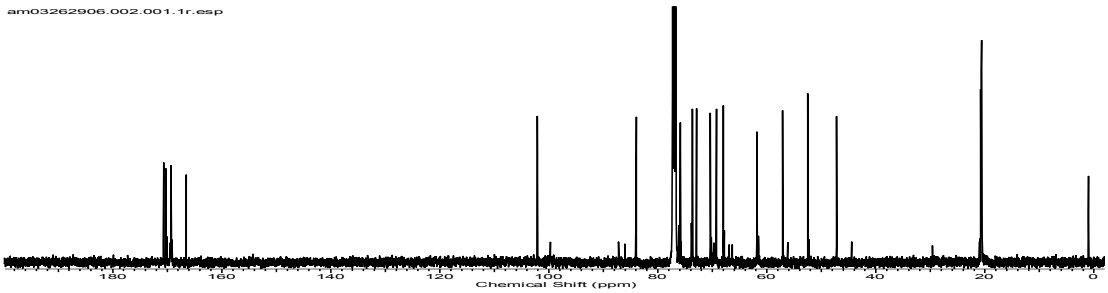
2,3,4,6-Tetra-*O*-acetyl-1-thio-β-D-glucopyranose (291.49 mg, 0.8 mmol) and methyl-(methyl 2,3-di-*O*-acetyl-4-deoxy-α-L-threo-hex-4-enopyranosid)uronate (30 mg, 0.1 mmol) were taken in a glass vial along with DPAP as a photo initiator (20.5 mg, 0.08 mmol). These were dissolved in 400 μL of 5:1 EtOH:DCM and irradiated under UV for 2 h. At the end of the reaction, the

product was purified by flash column chromatography (EtOAc/toluene 1:1) to give an amorphous solid. Yield: 84%; TLC:R_f 0.28 (EtOAc/toluene 1:1); $[\alpha]_D^{20}$ = -19.8 (c 0.3, $CHCl_3$); IR: ν_{max} 2957,2360, 2342, 1745, 1436, 1369, 1216, 1038, 913, 774 cm^{-1}**;** ^{1}H NMR (400 MHz, $CDCl_3$): δ ppm 5.25 (dd, $J_{2,1}$ = 7.83 Hz, $J_{2,3}$ = 9.85 Hz, 1H, H-2), 5.17 (t, $J_{3',4'}$ = 9.09 Hz, $J_{3',2'}$ = 9.35 Hz, 1H, H-3'), 5.12 (dd, $J_{3,2}$= 9.0 Hz, $J_{3,4}$ = 3.5 Hz, 1H, H-3), 5.07 (t, J = 9.85 Hz, J = 9.6 Hz, 1H, H-4'), 4.90 (t, $J_{2',3'}$ = 9.35 Hz, $J_{2',1'}$ = 10.36 Hz, 1H, H-2'), 4.57 (d, $J_{1',2}$ = 10.36 Hz, 1H, H-1'), 4.41 (d, J = 1.77 Hz, 1H, H-5), 4.35 (d, $J_{1,2}$ = 7.83 Hz, 1H, H-1), 4.27 (dd, $J_{6a',6b'}$ = 12.38 Hz, $J_{6a',5'}$ = 4.55 Hz, 1H, H-6a'), 4.14 (dd, $J_{6b',6a'}$ = 12.63 Hz, $J_{6b',5}$ = 2.02 Hz, 1H, H-6b'), 3.90 (m, 1H, H-4), 3.78 (s, 3H, -COOMe), 3.60 (m, 1H, H-5'), 3.52 (s, 3H, -OMe), 2.14 (s, 3H, -OAc), 2.11 (s, 3H, -OAc), 2.10 (s, 3H, -OAc), 2.06 (s, 3H, -OAc), 2.02 (s, 3H, -OAc), 2.00 (s, 3H, -OAc); ^{13}C NMR (101 MHz, $CDCl_3$): δ ppm 170.6 (s, 1C, -OAc), 170.3 (s, 1C, -OAc), 170.2 (s, 1C, -OAc), 169.3 (s, 1C, -OAc), 169.3, (s, 1C, -OAc), 169.2 (s, 1C, -OAc), 166.5 (-COOMe), 102.1 (C-1), 83.6 (C-1'), 75.6 (C-5'), 73.6 (C-5), 73.2 (C-3), 72.8 (C-3'), 70.3 (C-2'), 68.9 (C-2), 67.7 (C-4'), 61.6 (C-6), 56.8 (s, 1C, -OMe), 52.2 (-COOMe), 56.8 (-OMe), 46.9 (C-4), 20.7, 20.7, 20.6, 20.6, 20.5, 20.5 (6 x –OAc); ESI$^+$ HRMS:calcd. for $C_{26}H_{36}O_{17}S$ (M+Na^+) 675.1571, found 675.1573

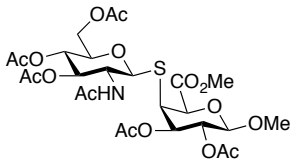
^{1}H spectra:



^{13}C spectra:



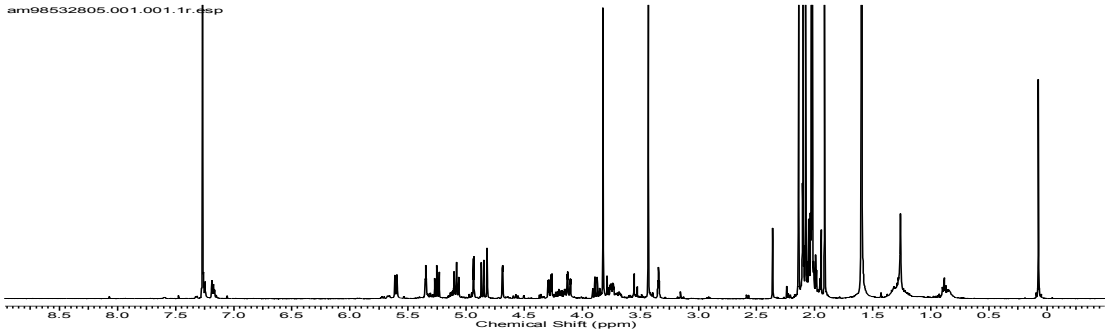
17. Synthesis of compound (19):



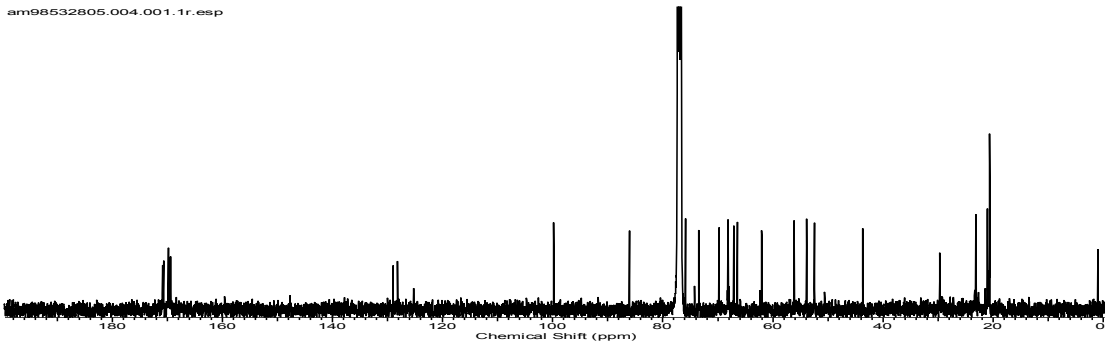
2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-1-thio-β-D-glucopyranose (290.7 mg, 0.8 mmol) and Methyl-(methyl 2,3-di-*O*-acetyl-4-deoxy-α-L-threo-hex-4-enopyranosid)uronate (30 mg, 0.1 mmol) were taken in a glass vial along with DPAP as a photo initiator (20.5 mg, 0.08 mmol). These were dissolved in 400 μL of 5:1 EtOH:DCM and irradiated under UV for 3 h. At the end of the reaction, the product was purified by flash column chromatography using 80% EtOAc-toluene to give an amorphous solid. Yield: 78%. TLC: R_f 0.26 (EtOAc/Toluene 4:1); $[\alpha]_D^{20}$ = -15.9 (c 0.3, $CHCl_3$); IR (neat): $\nu_{max.}$ 3282, 2935, 2360, 2341, 1745, 1666, 1572, 1437, 1370, 1230, 1046, 915 cm^{-1};^{1}H NMR (400 MHz, $CDCl_3$): δ ppm 5.57 (d, *J* = 8.83 Hz, 1H, NH), 5.35 (s, 1H, H-3), 5.26 (t, $J_{3',4'}$ = 9.46 Hz, $J_{3',2'}$ = 10.09 Hz, 1H, H-3'), 5.09 (t, $J_{4',2'}$ = 9.77 Hz, $J_{4',3'}$ = 9.46 Hz, 1H, H-4'), 4.94 (d, *J* = 2.84 Hz, 1H, H-5), 4.86 (d, $J_{1',2'}$ = 10.4 Hz, 1H, H-1'), 4.82 (s, 1H, H-1), 4.69 (t, $J_{2,1}$ = 1.26 Hz, *J* = 2.52 Hz, 1H, H-2), 4.28 (dd, $J_{6b',6a'}$ = 12.3 Hz, $J_{6b',5'}$ = 4.73 Hz, 1H, H-6b'), 4.12 (dd, $J_{6a',6b'}$ = 12.3 Hz, $J_{6a',5'}$ = 1.89 Hz, 1H, H-6a'), 3.89 (t, $J_{2',1'}$ = 10.4 Hz,

$J_{2',3'}$ = 8.83 Hz, 1H, H-2'), 3.82 (s, 3H, -COOMe), 3.74 (m, 1H, H-5'), 3.43 (s, 3H, -OMe), 3.35 (t, $J_{4,5}$ = 2.84 Hz, H-4), 2.14 (s, 3H, -OAc), 2.10 (s, 3H, -OAc), 2.08 (s, 3H, -OAc), 2.03 (s, 3H, -OAc), 2.02 (s, 3H, -OAc), 1.92 (s, 3H, -NHAc); ^{13}C NMR (101 MHz, $CDCl_3$): δ ppm 170.8 (s, 1C, -NHAc), 169.7 (s,1C, -OAc), 169.7 (s,1C, -OAc), 169.6 (s,1C, -OAc), 169.5 (s,1C, -OAc), 169.3 (s,1C, -OAc), 99.8 (C-1), 86.05 (C-1'), 75.8 (C-5'), 73.1 (C-3'), 69.5 (C-3), 67.8 (C-4'), 66.6 (C-2), 66.2 (C-5), 61.6 (C-6), 53.6 (C-2'), 55.8 (-OMe), 52.4 (-COOMe), 43.3 (C-4), 23.1 (s, 1C, -NHAc), 21.05, 20.9, 20.6, 20.6, 20.5 (5 x –OAc); ESI$^+$ HRMS: calcd. for $C_{26}H_{37}NO_{16}S$ (M+Na^+) 651.6341, found 674.1722

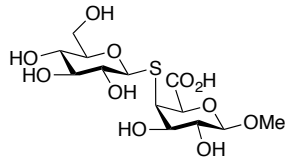
^{1}H spectra:



^{13}C spectra:

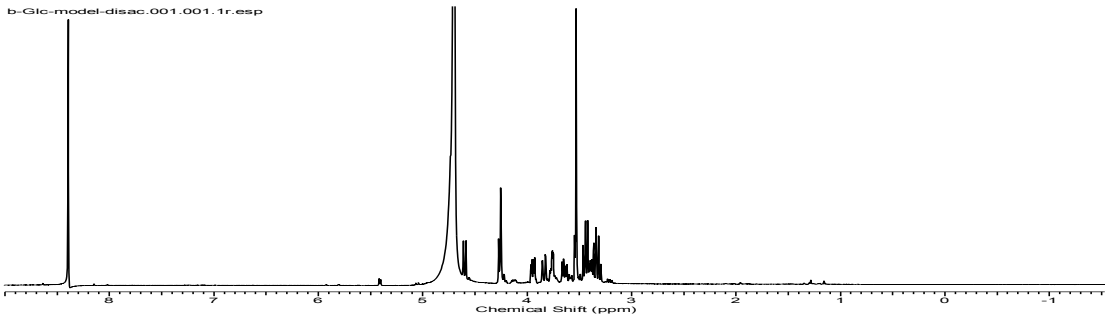


18. Synthesis of compound (20):

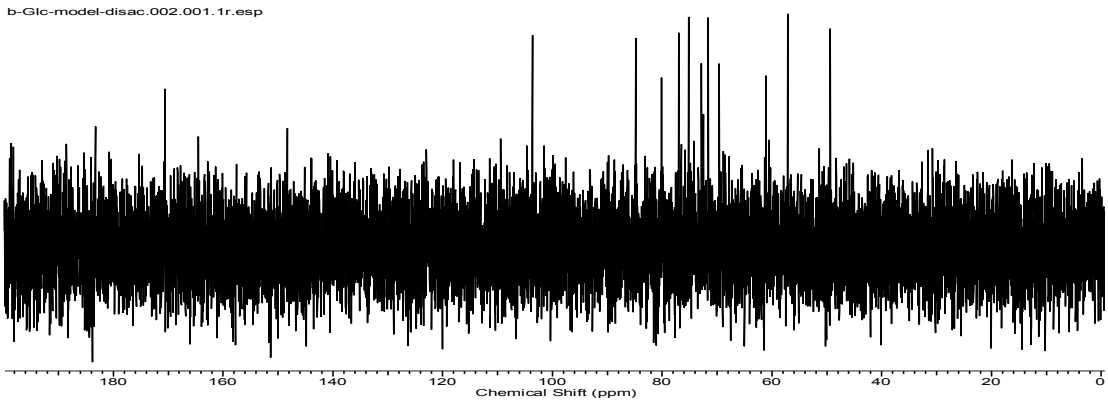


1-Thio-β-D-glucopyranose (204.06 mg, 1.04 mmol) and Methyl 4-deoxy-α-L-threo-hex-4-enopyranosiduronic acid (25 mg, 0.13 mmol) were taken in a glass vial along with Vazo-44 (33.6 mg, 0.10 mmol) as a water soluble photo initiator and these were dissolved in 0.25M pH 4 NaOAc buffer (400 μL). Consequently, the reaction mixture was irradiated under UV for 8 h. At the end of the reaction, the product was purified using 2:2:1 WIPE eluent system by flash column chromatography to give a white foam. Yield: 62%; TLC: R_f 0.29 (EtOAc/Isopropanol/Water 2:2:1); $[\alpha]_D^{20}$=-8.7 (c 0.4, MeOH);IR(neat): $\nu_{max.}$ 2963, 2360, 2341, 1749, 1699, 1507, 1220, 1037, 1013, 922 cm^{-1};^{1}H NMR (400 MHz, D_2O): δ ppm 4.60 (d, J = 9.85 Hz, 1H, H-1'), 4.27 (d, J = 7.57 Hz, 1H, H-1), 4.25 (d, $J_{5,4}$ = 1.89 Hz, 1H, H-5), 3.95 (dd, $J_{3,2}$ = 9.6 Hz, $J_{3,4}$ = 4.4 Hz, 1H, H-3), 3.84 (dd, $J_{6a',6b'}$ = 12.6 Hz, $J_{6a',5}$ = 1.89 Hz, 1H, H-6a'), 3.76 (dd, $J_{4,3}$ = 4.4 Hz, $J_{4,5}$ = 1.89 Hz, 1H, H-4), 3.65 (dd, $J_{6b',6a'}$ = 12.6 Hz, $J_{6b',5'}$ = 6.3 Hz, 1H, H-6b'), 3.53 (s, 3H, -OMe), 3.46-3.42 (m, 2H, H-2/H-3'), 3.42-3.38 (m, 1H, H-5'), 3.36-3.30 (m,2H, H-2'/H-4'); ^{13}C NMR (101 MHz, D_2O): δ ppm 174.7 (s, 1C, -COOH), 103.4 (C-1), 84.3 (C-1'), 74.9 (C-5), 72.7 (C-3), 60.7 (C-6'), 49.3 (C-4), 57.3 (-OMe), 69.1, 71.3 (C2/C-3'), 72.2 (C-5'), 72.7, 74.9 (C-2'/C-4'); ESI$^-$ HRMS: calcd. for $C_{13}H_{22}O_{11}S$ ($M-H^+$) 385.0810, found 385.0815

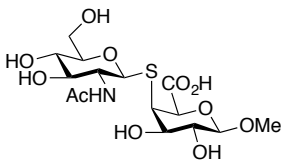
^{1}H spectra:



^{13}C spectra:



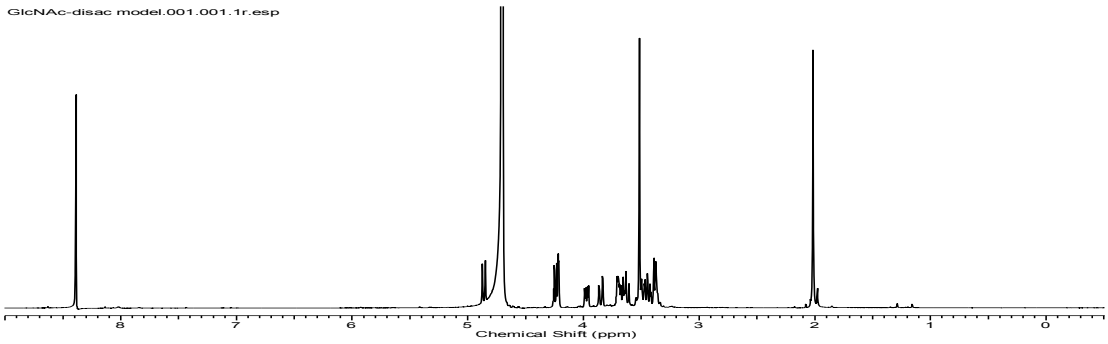
19. Synthesis of compound (21):



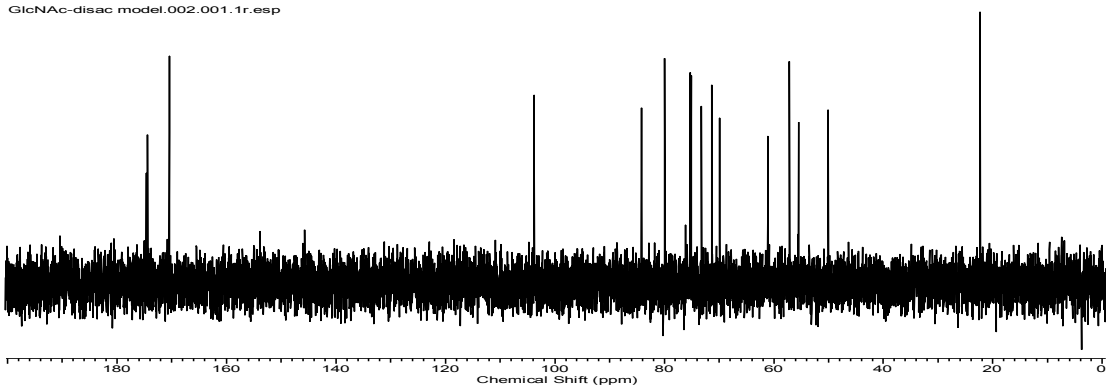
2-Acetamido-1-thio-β-D-glucopyranose (246.48 mg, 1.04 mmol) and Methyl 4-deoxy-α-L-threo-hex-4-enopyranosiduronic acid (25 mg, 0.13 mmol) were taken in a glass vial along with Vazo-44 (33.6 mg, 0.10 mmol) as a water soluble photo initiator and these were dissolved in 0.25M pH 4 NaOAc buffer (400 μL). Consequently, these were irradiated under UV for 8 h. At the end

of the reaction, the product was purified using 2:2:1 WIPE eluent system by flash column chromatography to give a white foam. Yield: 64%; TLC: R_f 0.21 (EtOAc/Isopropanol/Water 2:2:1); $[\alpha]_D^{20}$ = -9.4 (c 0.3, MeOH); IR (neat): $\nu_{max.}$3281, 2920, 2360, 2341, 1615, 1559, 1417, 1313, 1054, 948 cm^{-1}; ^1H NMR(400 MHz, D_2O): δ ppm 4.81 (d, $J_{1',2'}$ = 10.09 Hz, 1H, H-1'), 4.2 (d, $J_{1,2}$ = 7.88 Hz, 1H, H-1), 4.16 (d, $J_{5,4}$ = 1.89 Hz, 1H, H-5), 3.93 (dd, $J_{3,4}$ = 4.41 Hz, $J_{3,2}$ = 9.77 Hz. 1H, H-3), 3.8 (dd, $J_{6a',6b'}$ = 12.6 Hz, $J_{6a',5'}$ = 1.89 Hz, 1H, H-6a'), 3.66 (dd, $J_{4,5}$ = 1.89 Hz, $J_{4,3}$ = 4.41 Hz, 1H, H-4), 3.62 (dd, $J_{6b',6a}$= 12. 6 Hz, $J_{6b',5'}$ = 5.99 Hz, 1H, H-6b'), 3.58 (d, $J_{2',1'}$ = 10.09 Hz, 1H, H-2'), 3.43 (t, $J_{3',4'}$ = 9.77 Hz, $J_{3',2'}$ = 5.36 Hz, 1H, H-3'), 3.4 (dd, $J_{4',3'}$ = 9.77 Hz, $J_{2,1}$ = 7.88 Hz, 2H, H-2/H-4'), 3.34-3.32 (m, 1H, H-5'), 1.97 (s, 3H, -NHAc); ^{13}C NMR (101 MHz, D_2O): δ ppm 174.5 (s, 1C, -NHAc), 103.3 (C-1), 83.8 (C-1'), 79.5 (C-5'), 74.7 (C-5), 74.9 (C-3'), 72.7 (C-3), 70.8, 69.6 (2s, C-2/C-4'), 60.4 (C-6'), 56.6 (-OMe), 54.9 (C-2'), 49.4 (C-4), 21.9 (-NHAc); ESI^+ HRMS: calcd. for $C_{15}H_{25}NO_{11}S$ ($M+Na^+$) 450.1041, found 450.1025

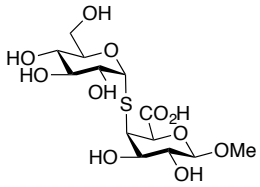
^{1}H spectra:



^{13}C spectra:



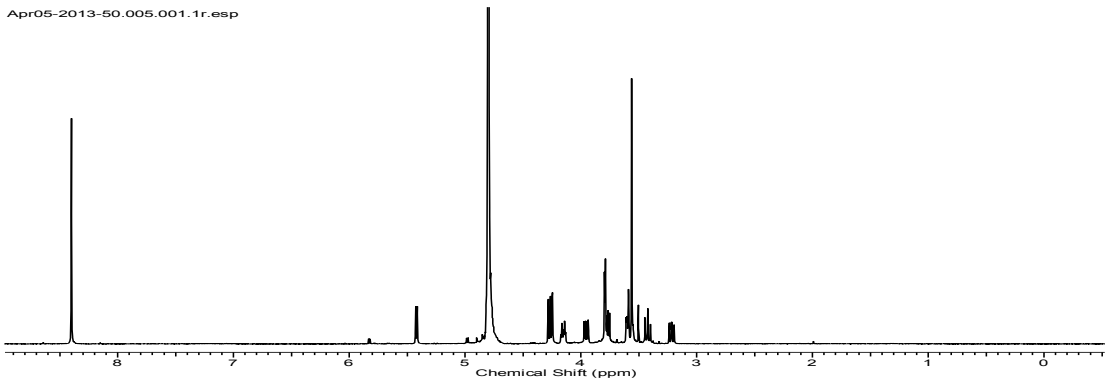
20. Synthesis of compound (22)



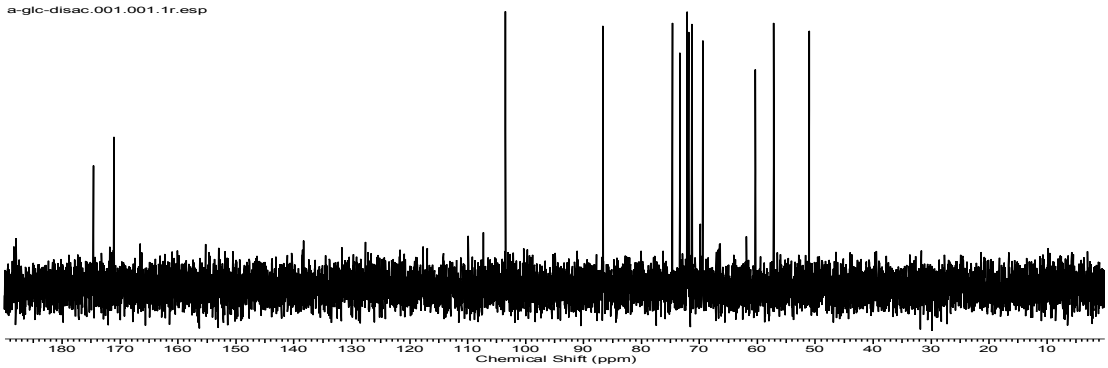
1-Thio-α-D-glucopyranose (204.06 mg, 1.04 mmol),methyl 4-deoxy-α-L-threo-hex-4-enopyranosiduronic acid (25 mg, 0.13 mmol) and Vazo-44 (33.6 mg, 0.10 mmol)were taken in a glass vial and dissolved in 250 mM ammonium formate buffer, pH 4 (400 μL). Subsequently, the reaction mixture was irradiated under UV for 8 hrs. Purification of the reaction mixture was done using a Mono Q column on an AKTA FPLC purifier system using 250 mM ammonium formate buffer and water (flow rate = 1 mL/min). Individual fractions were checked by TLC to locate the product. Pure fractions were then lyophilized multiple times to remove ammonium and obtain the product as a white foam in 56% yield. TLC:R_f 0.29 (EtOAc/Isopropanol/Water 2:2:1); IR (neat): $\nu_{max.}$ 629, 1061, 1454, 1585, 2921 cm^{-1}; ^{1}H NMR (400 MHz, D_2O): δ ppm 5.37 (d, J = 5.3 Hz, 1H, H-1'), 4.22 (d, J = 7.8 Hz, 1H, H-1), 4.19 (d, $J_{5,4}$ = 1.3 Hz, 1H, H-5), 4.08-4.12 (m, 1H,

H-5'), 3.9 (dd, $J_{3,4}$ = 10.1 Hz, $J_{3,2}$ = 9.9 Hz, 1H, H-3), 3.74 (m, 2H, H-6a'/H-6b'), 3.71 (dd,$J_{2',1'}$ = 5.4 Hz, $J_{2',3'}$ = 4.6 Hz, 1H, H-2'), 3.54-3.56 (m, 2H, H-4/H-3'), 3.51 (s, 3H, -OMe), 3.37 (dd, J = 9.9 Hz, J = 9.4 Hz, 1H, H-4'), 3.16 (dd, $J_{2,3}$ = 9.9 Hz, $J_{2,1}$ = 7.8 Hz, 1H, H-2); ^{13}C NMR (101 MHz, D_2O): δ ppm 174.6 (-COOH), 103.3 (C-1), 86.7 (C-1'), 74.7 (C-5), 73.4 (-OMe), 72.1 (C-3/C-2/C-5'), 72.0 (C-5'/C-3/C-2), 71.9 (C-2/C-3/C-5'), 71.3 (C-2'), 69.3 (C-4'), 60.3 (C-6'), 57.0 (C-3'/C-4), 50.8 (C-4/C-3') ; ESI$^-$ HRMS: calcd. for $C_{13}H_{22}O_{11}S$ ($M-H^+$) 385.0810, found 385.0806. ; $[\alpha]_D^{20}$ = +18.4 (c = 0.7, MeOH)

^{1}H spectra:

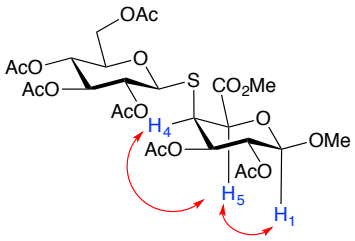


^{13}C spectra:

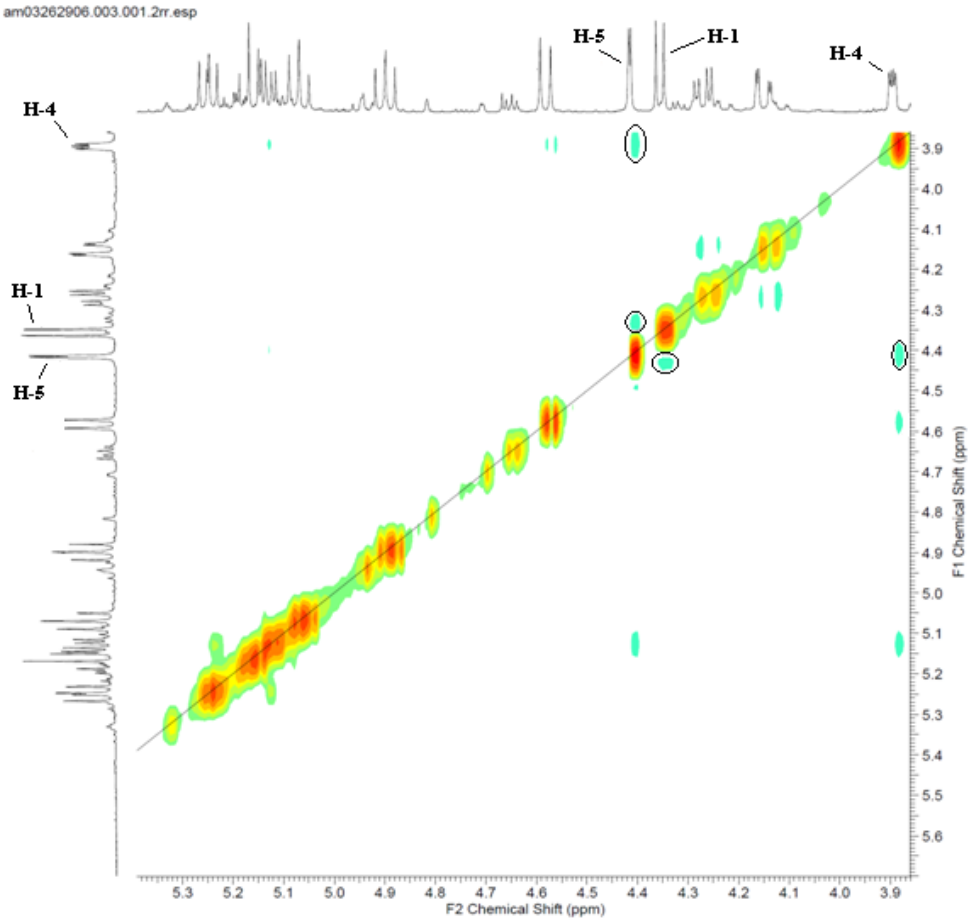


2.4.2 NMR analysis for determining stereochemistry of the disaccharides

(Example given for Methyl-2,3,4,6-tetra-*O*-acetyl-1-thio-β-D-glucopyranosyl-(1→4)-methyl-2,3-di-*O*-acetyl-β-D galacturonate) (**18**)



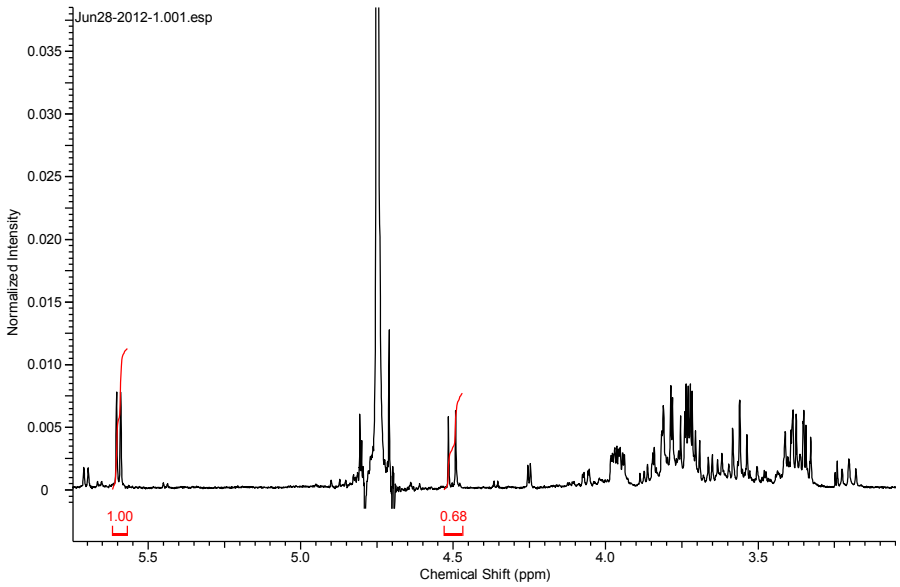
NOESY spectra for the disaccharide:



It is evident from the spectra that there is interaction between H-1 and H-5 which is possible only if H-5 is axial, thereby proving that the $–CO_2H$ group is equatorial. Thus the possibility of the acid being iduronic acid is ruled out. Next, we also see an interaction between H-4 and H-5. This points out to the fact that the uronic acid is galacturonic acid rather than glucuronic acid. In case of glucuronic acid, the H-4 proton of the uronic acid will be axial and will have no interaction with H-5 since it would be on the opposite side of the plane of the molecule. However, in case of galacturonic acid, H-4 and H-5 are on the same side of the molecule and should interact through space. Also, the coupling constant between H-4 and H-5 is very small (J = 1.77 Hz) which suggests that H-4 is equatorial. That means the uronic acid is in galacturonic form.

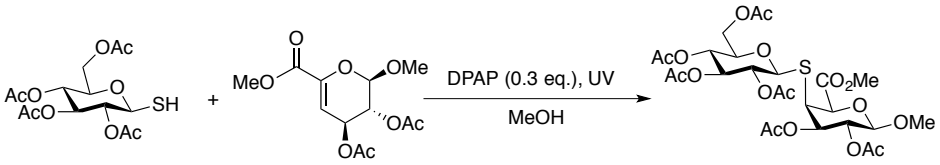
The same procedure has been followed for all the disaccharides.

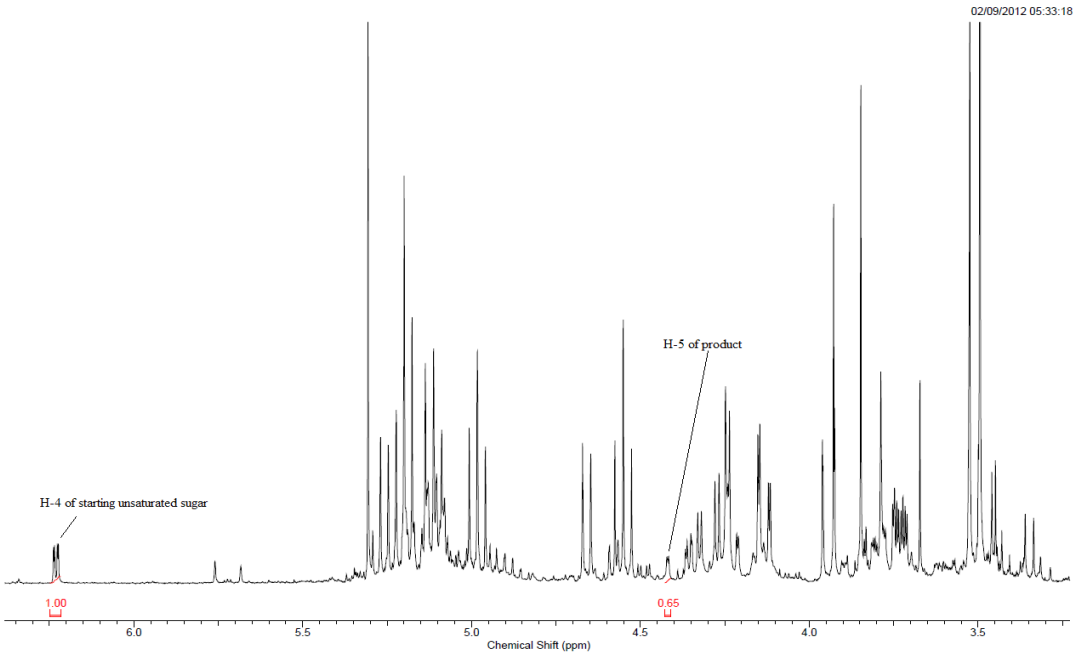
2.4.3 NMR showing anomerization and hydrolysis of α-thiosugar during deacetylation



Anomeric region of the NMR of the monosaccharide has been expanded. Peak at 5.6 ppm corresponds to anomeric proton of the α-thio sugar (based on coupling constant). Peak at 4.5 ppm corresponds to β-thio sugar (based on coupling constant). The ratio of α-thio sugar to β-thio sugar is 1:0.68 or roughly 5:3.

2.4.4 Crude NMR analysis for calculating yields





The above example is for entry 13 in Table 2. The yield can be consequently calculated to be 39%.

2.5 References

1. Dondoni, A. Free-radical thiol-ene and thiol-yne coupling as selective ligation tools for glycoconjugation. *Abstr Pap Am Chem S* **2010,** *240*.

2. Fiore, M.; Lo Conte, M.; Pacifico, S.; Marra, A.; Dondoni, A. Synthesis of S-glycosyl amino acids and S-glycopeptides via photoinduced click thiol-ene coupling. *Tetrahedron Lett*. **2011,** *52* (3), 444-447.

3. Floyd, N.; Vijayakrishnan, B.; Koeppe, A. R.; Davis, B. G. Thiyl Glycosylation of Olefinic Proteins: S-Linked Glycoconjugate Synthesis. *Angew. Chem. Int. Ed.* **2009,** *48* (42), 7798-7802.

4. Hu, Z. G.; Fan, X. S.; Zhang, G. S. Synthesis and characterization of glucose-grafted biodegradable amphiphilic glycopolymers P(AGE-glucose)-b-PLA. *Carbohydr. Polym.* **2010,** *79* (1), 119-124.

5. Wittrock, S.; Becker, T.; Kunz, H. Synthetic vaccines of tumor-associated glycopeptide antigens by immune-compatible thioether linkage to bovine serum albumin. *Angew. Chem. Int. Ed.* **2007,** *46* (27), 5226-5230.

6. Fiore, M.; Marra, A.; Dondoni, A. Photoinduced Thiol-Ene Coupling as a Click Ligation Tool for Thiodisaccharide Synthesis. *J. Org. Chem.* **2009,** *74* (11), 4422-4425.

7. Hoyle, C. E.; Bowman, C. N. Thiol-Ene Click Chemistry. *Angew. Chem. Int. Ed.* **2010,** *49* (9), 1540-1573.

8. Driguez, H. Thiooligosaccharides in glycobiology. *Glycoscience Synthesis of Substrate Analogs and Mimetics* **1997,** *187*, 85-116.

9. Dondoni, A. The Emergence of Thiol-Ene Coupling as a Click Process for Materials and Bioorganic Chemistry. *Angew. Chem. Int. Ed.* **2008,** *47* (47), 8995-8997.

10. Heidlas, J. E.; Lees, W. J.; Pale, P.; Whitesides, G. M. Gram-Scale Synthesis of Uridine 5'-Diphospho-N-Acetylglucosamine - Comparison of Enzymatic and Chemical Routes. *J. Org. Chem.* **1992,** *57* (1), 146-151.

11. Sen, A. K.; Banerji, N. Synthesis of 4-O-Beta-D-Xylopyranosyl-D-Glucopyranoses and 6-O-Beta-D-Xylopyranosyl-D-Glucopyranoses and Their Protein Conjugates. *Indian J Chem B* **1989,** *28* (10), 818-823.

12. Bonner, W. A.; Kahn, J. E. A Study of Some S-(Polyacetyl-β-D-glycopyranosyl)-thiuronium Halides. *J. Am. Chem. Soc.* **1951,** *73* (5), 2241-2245.

13. Johnston, B. D.; Pinto, B. M. Synthesis of Thio-Linked Disaccharides by 1→2 Intramolecular Thioglycosyl Migration: Oxacarbenium versus Episulfonium Ion Intermediates. *The Journal of Organic Chemistry* **2000,** *65* (15), 4607-4617.

14. Zemplén, G.; Kuntz, A. *Brit* **1923,** *56B*, 1705-10.

15. Ibatullin, F. M.; Selivanov, S. I. Reaction of 1,2-trans-glycosyl acetates with phosphorus pentachloride: new efficient approach to 1,2-trans-glycosyl chlorides. *Tetrahedron Lett.* **2002,** *43* (52), 9577-9580.

16. Blanc-Muesser, M.; Defaye, J.; Driguez, H. Syntheses stereoselectives de 1-thioglycosides. *Carbohydr. Res.* **1978,** *67* (2), 305-328.

17. Caraballo, R.; Deng, L. Q.; Amorim, L.; Brinck, T.; Ramstrom, O. pH-Dependent Mutarotation of 1-Thioaldoses in Water. Unexpected Behavior of (2S)-D-Aldopyranoses. *J. Org. Chem.* **2010,** *75* (18), 6115-6121.

18. Adorjan, I.; Jaaskelainen, A. S.; Vuorinen, T. Synthesis and characterization of the hexenuronic acid model methyl 4-deoxy-beta-L-threo-hex-4-enopyranosiduronic acid. *Carbohydr. Res.* **2006,** *341* (14), 2439-2443.

19. de Nooy, A. E. J.; Besemer, A. C.; van Bekkum, H. Highly selective nitroxyl radical-mediated oxidation of primary alcohol groups in water-soluble glucans. *Carbohydr. Res.* **1995,** *269* (1), 89-98.

20. Nooy, A. E. J. d.; Besemer, A. C.; Bekkum, H. v. On the Use of Stable Organic Nitroxyl Radicals for the Oxidation of Primary and Secondary Alcohols. *Synthesis* **1996,** *1996* (10), 1153-1176.

21. Li, K.; Helm, R. F. A practical synthesis of methyl 4-O-methyl-α-d-glucopyranosiduronic acid. *Carbohydr. Res.* **1995,** *273* (2), 249-253.

22. Llewellyn, J. W.; Williams, J. M. The synthesis of methyl 4-deoxy-β-l-threo-hex-4-enopyranoside. *Carbohydr. Res.* **1972,** *22* (1), 221-224.

23. Chernyak, A. Y.; Kononov, L. O.; Antonov, K. V. *Ser. Khim.* **1988,** *7*, 1660-1667.

24. Aspinall, G. O.; Krishnamurthy, T. N.; Mitura, W.; Funabashi, M. Base-Catalyzed Degradations of Carbohydrates .IX. Beta-Eliminations of 4-O-Substituted Hexopyranosiduronates. *Canadian Journal of Chemistry-Revue Canadienne De Chimie* **1975,** *53* (14), 2182-2188.

25. Bemiller, J. N.; Kumari, G. V. Beta-Elimination in Uronic Acids - Evidence for an Elcb Mechanism. *Carbohydr. Res.* **1972,** *25* (2), 419-428.

26. Gemma, E.; Hulme, A. N.; Jahnke, A.; Jin, L.; Lyon, M.; Muller, R. M.; Uhrin, D. DMT-MM mediated functionalisation of the non-reducing end of glycosaminoglycans. *Chem. Commun.* **2007,** (26), 2686-2688.

27. Hammond, G. S.; Neuman, R. C. The Mechanism of Decomposition of Azo Compounds. III. Cage Effects with Positively Charged Geminate Radical Pairs. *J. Am. Chem. Soc.* **1963,** *85* (10), 1501-1508.

28. Griesbau.K. Problems and Possibilities of Free-Radical Addition of Thiols to Unsaturated Compounds. *Angew. Chem. Int. Ed.* **1970,** *9* (4), 273-287.

29. Floyd, N.; Vijayakrishnan, B.; Koeppe, J. R.; Davis, B. G. Thiyl Glycosylation of Olefinic Proteins: S-Linked Glycoconjugate Synthesis. *Angew. Chem. Int. Ed.* **2009,** *48* (42), 7798-7802.

30. van de Linde, S.; Krstic, I.; Prisner, T.; Doose, S.; Heilemann, M.; Sauer, M. Photoinduced formation of reversible dye radicals and their impact on super-resolution imaging. *Photoch Photobio Sci* **2011,** *10* (4), 499-506.

31. Darmanyan, A. P.; Gregory, D. D.; Guo, Y. S.; Jenks, W. S.; Burel, L.; Eloy, D.; Jardon, P. Quenching of singlet oxygen by oxygen- and sulfur-centered radicals: Evidence for energy transfer to peroxyl radicals in solution. *J. Am. Chem. Soc.* **1998,** *120* (2), 396-403.

32. Staderini, S.; Chambery, A.; Marra, A.; Dondoni, A. Free-radical hydrothiolation of glycals: a thiol-ene-based synthesis of S-disaccharides. *Tetrahedron Lett.* **2012,** *53* (6), 702-704.

33. Trost, B. M.; Semmelhack, M. F.; Fleming, I. *Comprehensive Organic Synthesis*. Elsevier: **1991**; Vol. 4, p 1299.

34. Igarashi, K.; Honma, T. Addition Reactions of Glycals .IV. Free-Radical Addition of Thiolacetic Acid to D-Glucal Triacetate. *J. Org. Chem.* **1970,** *35* (3), 606-610.

35. Pryor, W. A. *Mechanisms of Sulfur Reactions*. McGraw-Hill Book Co., Inc., : New York, N.Y., **1962**; p 85.

36. Goering, H. L.; Relyea, D. I.; Larsen, D. W. The Stereochemistry of Radical Additions .III. The Radical Addition of Hydrogen Sulfide, Thiophenol and Thioacetic Acid to 1-Chlorocyclohexene. *J. Am. Chem. Soc.* **1956,** *78* (2), 348-353.

37. Bordwell, F. G.; Landis, P. S.; Whitney, G. S. The Stereochemistry of Free-Radical Addition of Thiolacetic Acid to Cyclohexenes. *J. Org. Chem.* **1965,** *30* (11), 3764-3768.

38. Huyser, E. S.; Benson, H.; Sinnige, H. J. Stereochemical Course of Free-Radical Additions of Mercaptans to Trans-Delta2-Octalin. *J. Org. Chem.* **1967,** *32* (3), 622-625.

39. Lebel, N. A.; DeBoer, A. Stereochemistry of Free-Radical Addition of Thiolacetic Acid to 2-Chloro-4-tert-Butylcyclohexene. *J. Am. Chem. Soc.* **1967,** *89* (11), 2784-2785.

40. Readio, P. D.; Skell, P. S. Stereochemistry of Free-Radical Addition of Methyl Mercaptan to 1-Chloro-4-T-Butylcyclohexene. *J. Org. Chem.* **1966,** *31* (3), 759-764.

41. Pachamuthu, K.; Schmidt, R. R. Synthetic routes to thiooligosaccharides and thioglycopeptides. *Chem. Rev.* **2006,** *106* (1), 160-187.

42. Lote, C. J.; Weiss, J. B. Identification of Digalactosylcysteine in a Glycopeptide Isolated from Urine by a New Preparative Technique. *FEBS Lett.* **1971,** *16* (2), 81-85.

43. Bernardes, G. J. L.; Grayson, E. J.; Thompson, S.; Chalker, J. M.; Errey, J. C.; El Oualid, F.; Claridge, T. D. W.; Davis, B. G. From disulfide- to thioether-linked glycoproteins. *Angew. Chem. Int. Ed.* **2008,** *47* (12), 2244-2247.

44. Zhdanov, Y. A.; Levitan, G. E. Cupric Complex of Thioaminoglucose. *Zh. Obshch. Khim.* **1976,** *46* (4), 937-938.

45. Horton, D.; Wolfrom, M. L. Thiosugars .1. Synthesis of Derivatives of 2-Amino-2-Deoxy-1-Thio-D-Glucose. *J. Org. Chem.* **1962,** *27* (5), 1794-1800.

46. Gamblin, D. P.; Garnier, P.; van Kasteren, S.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. Glyco-SeS: Selenenylsulfide-mediated protein glycoconjugation - A new strategy in post-translational modification. *Angew. Chem. Int. Ed.* **2004,** *43* (7), 828-833.

47. Latham Jr., H. G.; May, E. L.; Mosettig, E. Amino- and Guanidino-Phenylglucosides. *J. Org. Chem.* **1950,** *15* (4), 884-889.

48. Bredereck, H.; Hoschele, G. Hochvakuum-Destillation Acetylierter Zucker. *Chem. Ber. Recl.* **1953,** *86* (1), 47-54.

49. Beyer, H.; Schultz, U. Uber Thiazole .18. Uber Die Synthese Von 2-[D-Gluco-Pentaoxy-Pentyl]-Thiazolen. *Chem. Ber. Recl.* **1954,** *87* (1), 78-81.

50. Cerny, M.; Pacak, J. Darstellung Von 2,3,4,6-Tetra-O-Acetyl-Beta-D-Glucopyranosylmercaptan Und Von Natrium-Beta-D-Glucopyranosylmercaptid Und Gold-Beta-D-Glucopyranosylmercaptid. *Collect. Czech. Chem. Commun.* **1961,** *26* (8), 2084-2086.

51. Ferrier, R. J.; Furneaux, R. H. Chemistry of Some 1-Mercury(II)Thio-D-Glucose Compounds - New Synthesis of 1-Thio Sugars. *Carbohydr. Res.* **1977,** *57* (Aug), 73-83.

52. Usov, A. I.; Deryabin, V. V. Oxidation of the Methyl-3,6-Anhydro-D-Galactopyranoside Acetates and Agarose by Chromic Anhydride. *B Acad Sci Ussr* **1980,** *29* (2), 306-310.

53. Betaneli, V. I.; Ott, A. Y.; Brukhanova, O. V.; Kochetkov, N. K. Synthesis of 1,2-O-Cyanoethylidene Derivatives of Alkyl Glycopyranuronates by Oxidation of the 6-Trityl Ethers of Their Hexose Analogs. *Carbohydr. Res.* **1988,** *179*, 37-50.

Chapter 3

Towards building *S*-linked heparin and heparan sulfate mimics using thiol-ene chemistry

Table of Contents

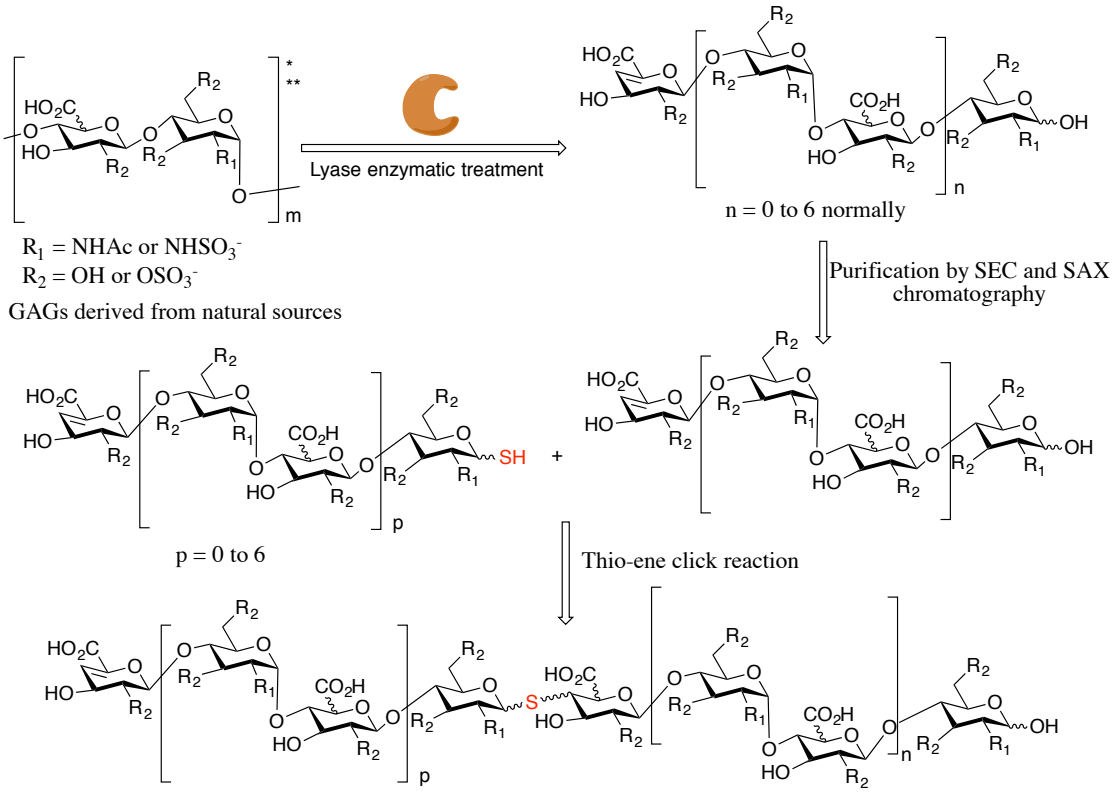
Chapter – 3 Towards building *S*-linked heparin and heparan sulfate mimics using thiol-ene chemistry

3.1 Introduction

Chapter 1 introduced the complexity of glycosaminoglycan (GAG) structures such as heparin and heparan sulfate and the various approaches, which have been taken to prepare these oligosaccharides in a defined, homogenous form. Here we present an approach based on the thiol-ene chemistry developed in Chapter 2. This approach is especially useful for building heparin thio-mimics, as heparin-degrading lyases reveal unsaturation between position 4 and position 5 of the uronic acid at the non-reducing end of the carbohydrate chain[1-2]. This presents an opportunity for addition of desired thiosugars using a thiol-ene radical reaction–a modular approach for the synthesis of defined glycosaminoglycan structures.

This strategy entails degradation of naturally available heterogeneous heparin or other GAGs with GAG lyases in the first step. The resulting smaller fragments, containing an olefin at the non-reducing end, can be separated using size-exclusion chromatography. The fractions resulting from size exclusion purification can then be subjected to strong anion exchange purification, which will result in further separation based on the charge of the fragments. This sequence should provide defined glycosaminoglycan fragments, which can be thiolated in the anomeric position to prepare thio-donors. Finally, these donors can then be utilized in conjugation reactions to the unsaturation of another unit of purified glycosaminoglycan, resulting in the synthesis of defined glycosaminoglycan structures (**Scheme 1**).

Thus, we envisioned using this technique to synthesize glycosaminoglycan structures with different patterns of sulfate distribution. These products are potentially useful in elucidating structure activity relationships (SAR) of GAGs with proteins. Theoretically this approach can also enable us to prepare novel GAG hybrid structures with different GAG entities as the constituents, to screen for recognition and potential bioactivity.



* In some GAGs, the amino sugar can be galactosamine instead of glucosamine
** In case of heparin and heparan sulfate the linkage between the uronic acid and glucosamine unit is β-1,4 instead of β-1,3 as seen for other GAGs.

Scheme 1: Proposed methodology for synthesis of defined glycosaminoglycan thio-mimics.

The work towards preparation of these defined heparin fragments was done in collaboration with Prof. Jeremy Turnbull's laboratory in University of Liverpool. Additionally, the protein substrates used for elucidating the biological activity of the synthesized heparin mimics were also obtained from the Turnbull lab.

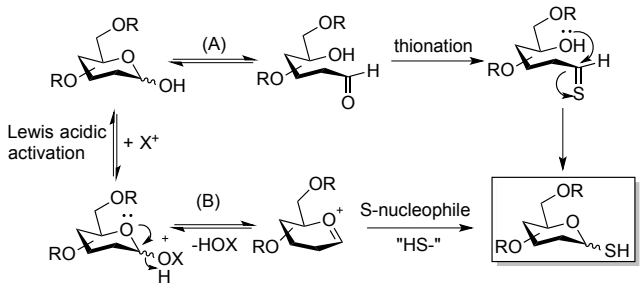
3.2 Results and Discussion

The methodology developed for coupling deprotected monosaccharides to prepare disaccharides in aqueous solution using a thiol-ene radical reaction was described in Chapter 2. This strategy requires efficient anomeric thiolation of one of the carbohydrate-coupling partners of interest. Here, we discuss approaches towards the development of conditions suitable for anomeric thiolation of highly negatively charged sugars derived from heparin digestion. These substrates allow the synthesis of heparin mimics using our approach.

3.2.1 Attempts at anomeric thiolation of deprotected sugars

There has been a previous report from our laboratory describing anomeric thiolation of deprotected sugars[3]. Essentially, this involves treating deprotected sugars at high temperature with Lawesson's reagent[4]. Lawesson's reagent has been widely reported as a convenient thiolating agent for synthesizing thio-carbonyl functionalities from oxo-carbonyl functionalities[5-8]. There has also been a report of synthesizing thiols from alcohols using this reagent[9]. **Scheme 2** suggests two possible mechanisms by which sugars can be thiolated at the anomeric position. These are (A) through the open-chain intermediate, which converts the aldehyde into a

thiocarbonyl group by Lawesson's reagent or (B) through an oxocarbenium intermediate that is attacked by a thiol nucleophile from Lawesson's reagent. Both provide an anomeric mixture of products with the α/β product ratios varying from1:6 to 1:4 depending on the monosaccharide.

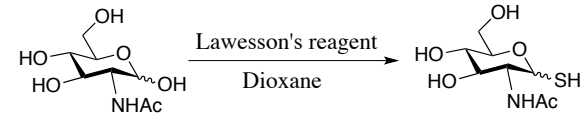


Scheme 2: Proposed mechanisms for anomeric thiolation of deprotected sugars using Lawesson's reagent. (A) Through an open chain intermediate (B) through an oxocarbenium intermediate[3].

N-acetylglucosamine was initially used as a model substrate to establish this reaction, as the report does not include this substrate. A variety of reaction conditions including the reported reaction conditions[3] were tried, as summarized in **Table 1**. The solvent in all cases was kept as dioxane, since other solvents were found to be unsuitable for dissolving the deprotected sugar. The most common problem faced was that Lawesson's reagent degraded at higher temperatures, while at lower temperatures the sugar remained undissolved. To avoid this, the sugar in dioxane was brought to reflux before the addition of Lawesson's reagent to minimize the degradation of Lawesson's reagent (**Table 1**, Entry **4**). Alternatively, aqueous HCl was added to facilitate the formation of both the oxocarbenium intermediate or the open chain form (**Table 1**, Entry **5)**. Unfortunately, the desired product was not observed by TLC or mass spectrometry under any of these conditions.

Investigating this, it is reported that amides react with Lawesson's reagent readily[10-11]. Thus, a major concern was that the carbonyl group of the amide might also undergo conversion

to a thiocarbonyl, resulting in either the undesired thioamide or dual functionalization to the anomeric thiol/thioamide. However, no thiolations were observed under the reaction conditons explored, as observed by TLC and mass spec.



N-acetyl-D-glucosamine

Table 1: Optimization conditions attempted for one-step thiolation of *N*-acetyl glucosamine

Entry 1	Equivalent of LR	Temperature	Time (h)	Yield
1	1.2	R.T.	48	-
2	1.2	reflux	6	-
3	2.4	reflux	12	-
4	1.2*	reflux	12	-
5	1.2**	reflux	12	-

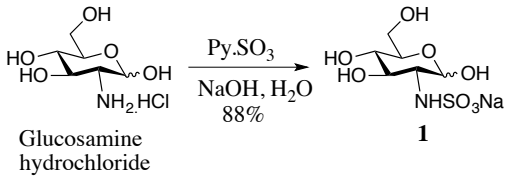
LR = Lawesson's reagent, *LR added after sugar had dissolved in dioxane under refluxing conditions, **A drop of aq. HCl added to the reaction mixture to facilitate acidic activation

At this stage, we decided to take a step back and do the reaction on glucose, which has been reported in the paper. All the reaction conditions described in **Table 1** including the optimized reaction conditions reported were again tried using glucose as the substrate. Unfortunately, we were again unable to observe any desired product. Hypothesizing that the product might be lost during work-up, a crude NMR was taken of the reaction mixture after carrying it out under the reported optimized conditions, which showed only Lawesson's reagent degradation products. A new batch of Lawesson's reagent did not help either. Thus, unprotected sugars were not good substrates for this reaction in our hands.

3.2.2 Synthesis of *N*-sulfated glucosamine

Since we were unsuccessful in optimizing the reaction for thiolating the anomeric position in deprotected sugars, we next attempted to thiolate the anomeric position of a sulfated sugar using the classical route of protection and deprotection. Two classical routes are known– attack of the anomeric position of the protected sugar having a good leaving group with either potassium thioacetate or thiourea followed by reduction with sodium metabisulfite (discussed in Chapter 2). For this purpose, we synthesized *N*-sulfated glucosamine, which was used as a model substrate to evaluate any difficulties due to the presence of the sulfate group as well as the overall yield for the whole reaction. These concerns are especially important since the sulfated glycosaminoglycan fragments obtained by heparin lyase treatment are only obtained in limited amounts.

N-sulfated glucosamine was synthesized from glucosamine hydrochloride as reported[12]. Briefly, glucosamine hydrochloride was treated with sulfur trioxide-pyridine complex at pH 9.5 to transfer the sulfate functionality from pyridine to the amino group of glucosamine to provide **1** (**Scheme 3**).



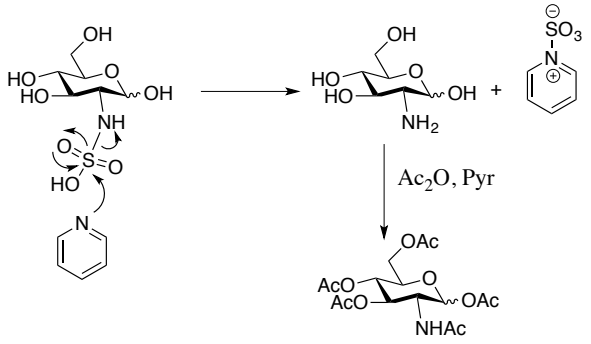
Glucosamine hydrochloride

1

Scheme 3: *N*-sulfated glucosamine synthesis[12].

With **1** in hand, attempts were made towards acetylating the substrate with acetic anhydride and pyridine for further reaction towards preparation of sulfated glucosamine with an anomeric thiol. However, when an aqueous workup was done, it was realized that a substantial

amount of the sugar entered the aqueous layer because the product behaves as an amphiphile. In addition, a considerable amount of acetylated *N*-acetylglucosamine was recovered, verified by crude NMR. The formation of *N*-acetylglucosamine can be explained by the plausible mechanism shown in **Scheme 4**. Essentially, the reverse reaction of **1** formation might have occurred, followed by acetylation of the free amine group.



Scheme 4: Plausible mechanism for formation of acetylated *N*-acetylglucosamine.

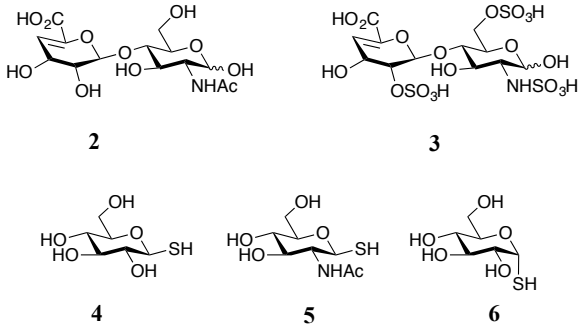
To overcome this problem, alternative acetylation conditions were attempted. Acetic anhydride/NaOAc-assisted acetylation had similar problems during aqueous work up with considerable loss in yield while acid catalysed acetylation reaction conditons[13] were seen to lead to desulfation. It is to be noted that acid catalysed *N*-desulfation has been reported for heparin[14].

At this point we abandoned such a late-stage thiolation approach for the synthesis of anomeric thiols from sulfated carbohydrates. We therefore, set out to establish a proof of concept approach using available thiols **4**, **5**, and **6** (**Figure 1**). We chose to use β-thio GlcNAc, **5,** instead of the naturally present α-thiosugar as we had observed low yield of formation in Chapter-2 from α-thiosugars. Moreover, α-thiosugar, **6** was still chosen to verify if low yield is also observed in coupling with heparin/HS derived disaccharides.

Also, two heparin- and heparan sulfate-derived disaccharides **2** and **3** (**Figure 1**) were chosen as the unsaturated counterpart for the preparation of heparin trisaccharide mimics.

3.2.3 Expression of Heparinase-1

Unsaturated heparin-derived oligosaccharides were prepared by the enzymatic digestion of heparin and heparan sulfate GAG chains. There are three different heparinases, each of which act on different heparin/HS substrates[15]. The substrate selectivity varies with the sulfate distribution of GAG chains[16-17]. Sometimes, a combination of two or all three is needed to break down a particular heparin/HS oligosaccharide owing to the structural heterogeneity of these glycosaminoglycans. The substrate **3** is the dominant disaccharide structure in heparin, whereas **2** is obtained from heparan sulfate[18-19]. In our work, **3** can be obtained by the action of Heparinase-1 (Hep-1) on heparin alone, while **2** is obtained by the degradation of heparan sulfate using Heparinase-3 (Hep-3). Both substrates were initially obtained from our collaboration with the Turnbull group.



2 **3**

4 **5** **6**

Figure 1: Heparin/HS disaccharide structures selected for thiol-ene radical conjugation. **2** represents non-sulfated derivative obtained mostly from HS and to a lesser extent from heparin. **3** is the most sulfated derivative obtained from heparin. **4**, **5** and **6** correspond to the available thiols used for preparation of heparin trisaccharide mimics.

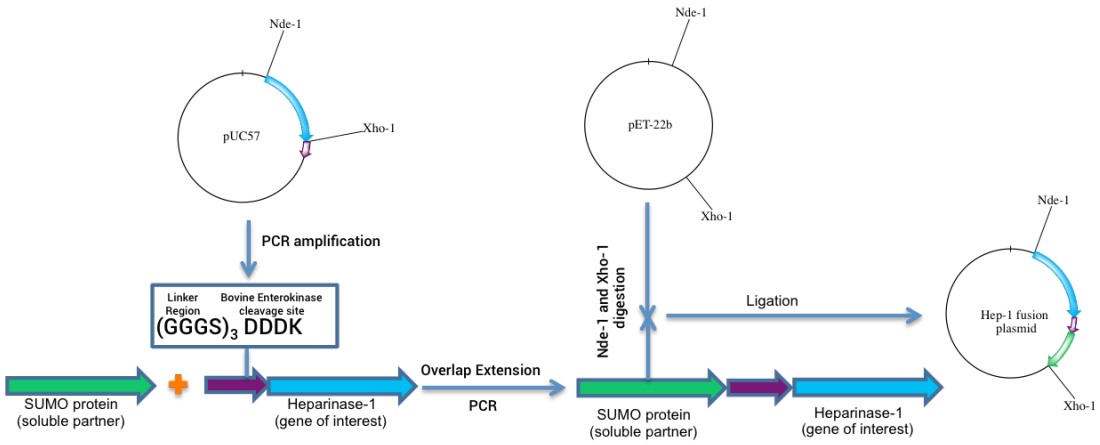
We also expressed Heparinase-1 for continued use in our laboratory. Few examples of expression of Heparinase-1 can be found in literature. In two such studies, Heparinase-1 was found to be expressed in the inclusion bodies[20-21]. This necessitates an extra step of refolding and also affects the yield of the enzyme obtained. To avoid this, in such cases the protein is normally fused with a soluble protein partner, which increases its percentage of expression in *E. coli.* as a soluble protein[22]. Accordingly, there has been a recent report of expression of Heparinase-1 after fusing it with various soluble protein partners[23]. After screening through various soluble protein partners, the study found SUMO (small ubiquitin related modifier protein) and IF2 (initiation factor 2 protein) to be the most promising soluble protein partners. The activity of the enzyme was preserved without having to cleave off the soluble protein tag after expression. The construct was designed such that Hep-1 gene was attached to the fusion protein gene through a linker.

For our expression of Heparinase-1, we decided to express it as a fusion protein with SUMO. The fused gene was incorporated in pET-22b vector[24] to construct the plasmid, as depicted in **Scheme 5.**

3.2.3.1 Design and construction of Heparinase-1 fused construct with SUMO

The Hep-1 gene construct including the linker region (which consists of a glycine rich peptide sequence followed by a bovine enterokinase restriction site as mentioned by Huang et al.[23]) at the *N*-terminus (without the SUMO tag), was procured from Genscript (**Scheme 5**). We intended to ligate the SUMO tag ourselves with the Hep-1 gene construct using overlap-extension PCR[25-27]. Accordingly, four different primers, SUMO_F and SUMO_R (forward and reverse primers, respectively for PCR amplification of SUMO gene sequence) and Hep-1_F and

Hep-1_R (forward and reverse primers, respectively for PCR amplification of Hep-1 gene sequence) were designed (**Experimental Section 3.4.3**). Notably they were designed in such a way that the SUMO_R also had the reverse compliment of Hep-1_F in the form of extra few bases at the end, so that overlap PCR can be performed subsequently. The Hep-1 gene sequence was PCR amplified from pUC57 plasmid using the Hep-1_F and Hep-1_R primers and the SUMO gene sequence was amplified using the SUMO_F and SUMO_R primers respectively, from a pET-15b plasmid in which it was originally present. The reaction mixture was then loaded onto a DNA gel and the appropriate bands were extracted (**Figure 2**) using a QIAquick Gel Extraction kit (a commercially available kit which is used to extract DNA from a gel by dissolving the agarose gel).



Scheme 5: Strategy for construction of Hep-1-SUMO fusion protein. Hep-1 (in blue), along with the linker region (in purple) in pUC57 vector was ordered from Genscript, with *Nde*1 and *Xho*1 as the restriction sites. The linker region consisted of a glycine rich region $(GGGS)_3$ and a site for bovine enterokinase cleavage as reported in literature[23]. SUMO gene was PCR amplified along with HisTag region on the N-terminus from a pET-15 vector in which it was originally present, using the SUMO_F and SUMO_R primers.

Thereafter, overlap-extension PCR[25-27] was done to fuse the two gene products together. The product from this reaction was amplified further by PCR using the SUMO forward and Hep-1 reverse primer. The success of the PCR was confirmed by running it on an agarose DNA gel where the expected mass increase of Hep-1 was observed (**Figure 3**).

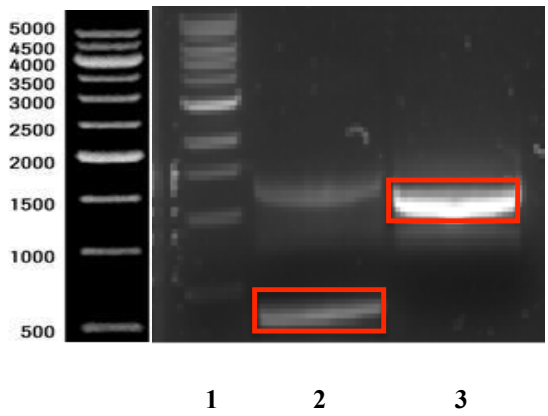


Figure 2: 0.8% Agarose DNA gel run for the PCR products of Hep-1 and SUMO. **1** represents 1 kb DNA ladder. **2** represents band corresponding to SUMO gene after PCR (enclosed band). **3** represents band corresponding to Hep-1 PCR product (enclosed band).

The fusion gene product was subsequently extracted from the gel and digested with *Nde*1[28] and *Xho*1[29] restriction enzymes as we wanted to use the corresponding restriction sites for inserting the fusion gene into pET-22 vector (pET vector is a vector which has got a T7 bacteriophage promoter sequence and *lac* operon upstream of the multiple cloning site of the vector. Thus expression of this system can be induced using isopropyl thiogalactoside or IPTG).

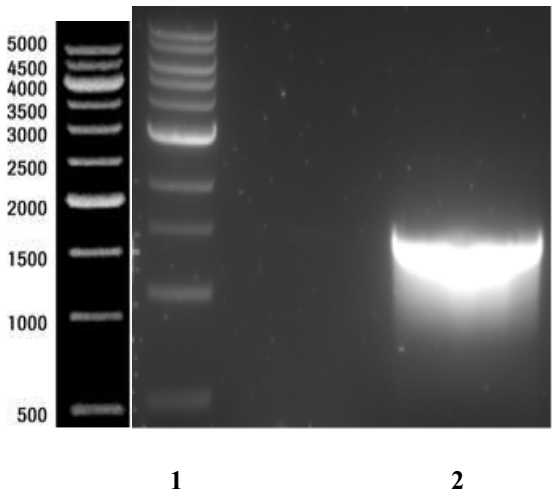


Figure 3: 0.8% Agarose DNA gel for PCR product of fusion gene. **1** represents 1kb DNA ladder. **2** represents Hep-1-SUMO fusion gene.

Simultaneously, a plasmid containing the pET-22 vector was also digested with the above-mentioned restriction enzymes to introduce the sticky ends and release the gene it contained previously. The digested mixture was then run on a DNA gel (**Figure 4**) and the corresponding DNA band was extracted with a gel extraction kit. Consequently, the fusion gene was inserted into pET-22b vector using the Quick Ligation Protocol (a commercially available kit which is used to ligate a gene into an empty vector to prepare the corresponding plasmid with the help of a DNA ligase enzyme).

3.2.3.2 Expression of fused Heparinase-1-SUMO protein

The constructed plasmid (Hep-1-SUMO fused in pET-22b) was subsequently transformed into XL10-Gold ultracompetent cells. XL10-Gold cells are commercially available cell lines that are competent and therefore they are more likely to incorporate foreign DNA. These cell lines are optimized for plasmid cloning as opposed to corresponding protein

production. The cell culture was then plated on ampicillin resistant plates and a single colony was grown overnight. The plasmid DNA was thereby extracted from the single cell colony and sent for sequencing. After confirmation of the correct gene sequence by sequencing, the plasmid was transformed into BL21(DE3) cells, which are cell lines optimized for protein expression. The cell culture was plated on ampicillin resistant plates and a single cell colony was picked and cultured overnight. Consequently, small-scale protein expression was started, expression was induced using IPTG and the soluble protein was extracted using BugBuster mix (a detergent based commercial solution that can perforate cell wall and liberate soluble protein without denaturing it). A protein gel was run, which does not indicate high expression (**Figure 5**)

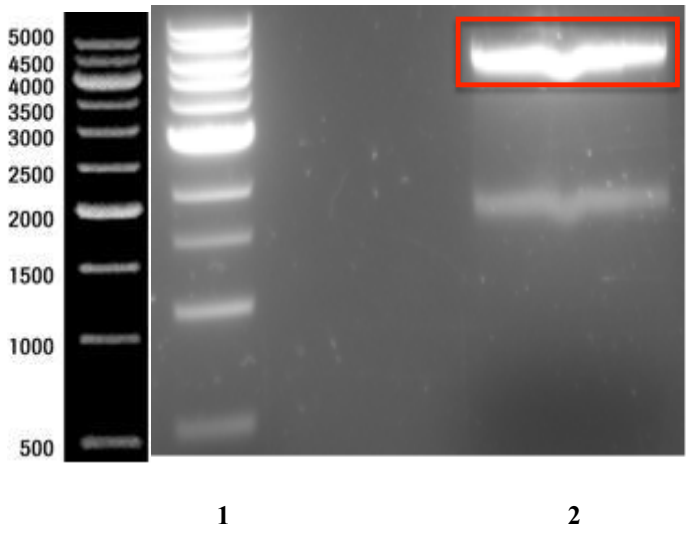


Figure 4: Agarose DNA gel of the restriction enzyme digested pET-22 plasmid. **1** represents 1kb DNA ladder. **2** corresponds to the empty pET-22 vector (enclosed band) while the lower band corresponds to the gene contained in the plasmid previously.

Nonetheless, the expression was scaled up to a 2 L culture and expression was induced by IPTG. The cells were lysed by sonication and the lysate was purified using HisTrap column, as the desired protein carries a His-Tag (HisTrap columns are Nickel Sepharose High Performance

commercial columns for purification of histidine-tagged protein by immobilized metal ion affinity chromatography). The purified protein showed bands in SDS-gel corresponding to the expected mass of 56 KDa (**Figure 6A**).

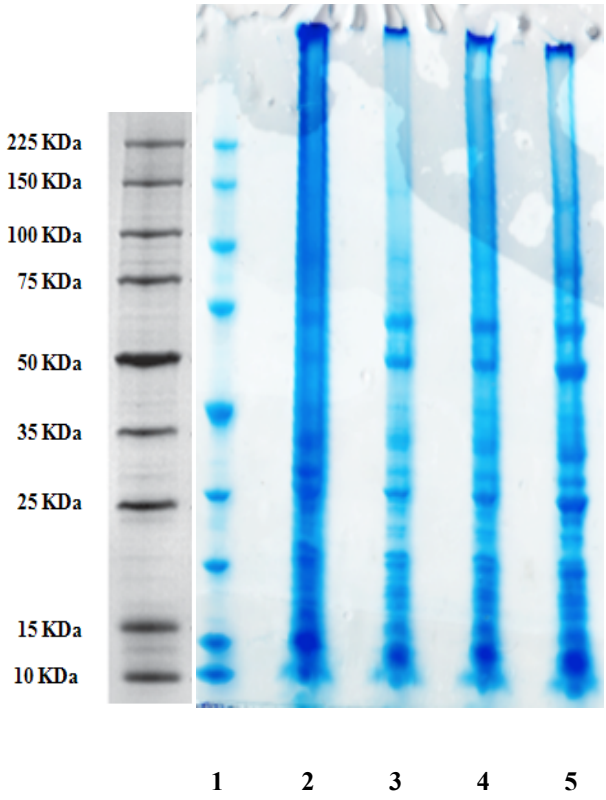
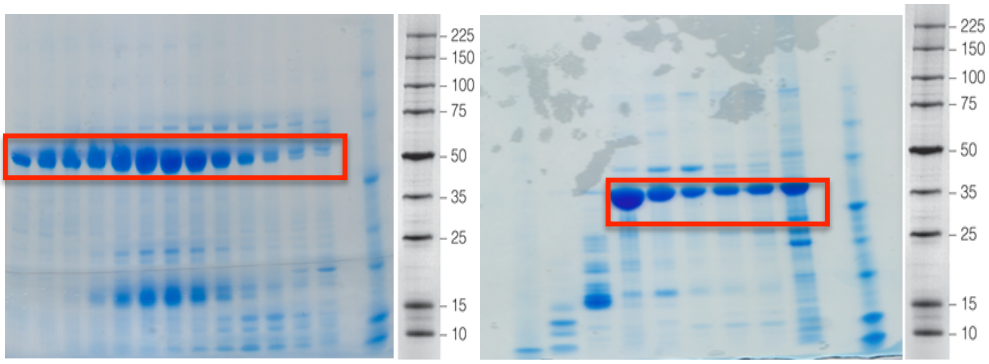


Figure 5: Small scale expression of Hep-1-SUMO fusion protein. **1** represents 10-250 KDa ladder, **2** shows induced cell expression, **3**, **4** and **5** correspond to uninduced cell expression.

The protein could be further purified with size exclusion chromatography (**Figure 6B**). This gave us sufficiently pure Hep-1 fusion protein to begin biochemical studies. The expected mass was also confirmed by MS (**Figure 8**).



(A) **(B)**

Figure 6: (A) His-Trap purification fractions of Hep-1-SUMO fusion protein. **(B)** Size exclusion purification fractions of Hep-1-SUMO fusion protein on S-200 column.

The yield of the expression was found to be moderate (3 mg/litre). In order to investigate the low yield, another gel was run with the flowthrough from HisTrap purification and the inclusion bodies (dissolved in 8M urea) (**Figure 7**). It was revealed that a significant amount of the enzyme was still being expressed in the inclusion bodies. This also explained the absence of a prominent band corresponding to the desired molecular weight in the small-scale expression tests**.** Though the amount was lower than reported, it was sufficient for preparative heparin digestion.

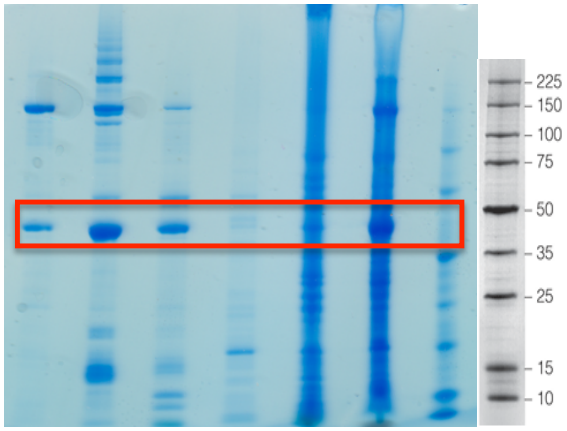


Figure 7: Protein gel showing significant amount of expression as insoluble inclusion bodies. **1**, **2**, **3** and **4** corresponds to HisTrap purified fractions, **5** is flowthrough fraction, **6** represents inclusion bodies (dissolved in 8M urea) and **7** corresponds to 10-250 KDa protein ladder.

Mass spectrometry confirmed the protein to be of the correct mass. (**Figure 8**) The protein was stored with 10% glycerol in 1 ml aliquots at a final concentration of 1 mg/ml and they were found to be stable and exhibit lyase activity for at least 6 months.

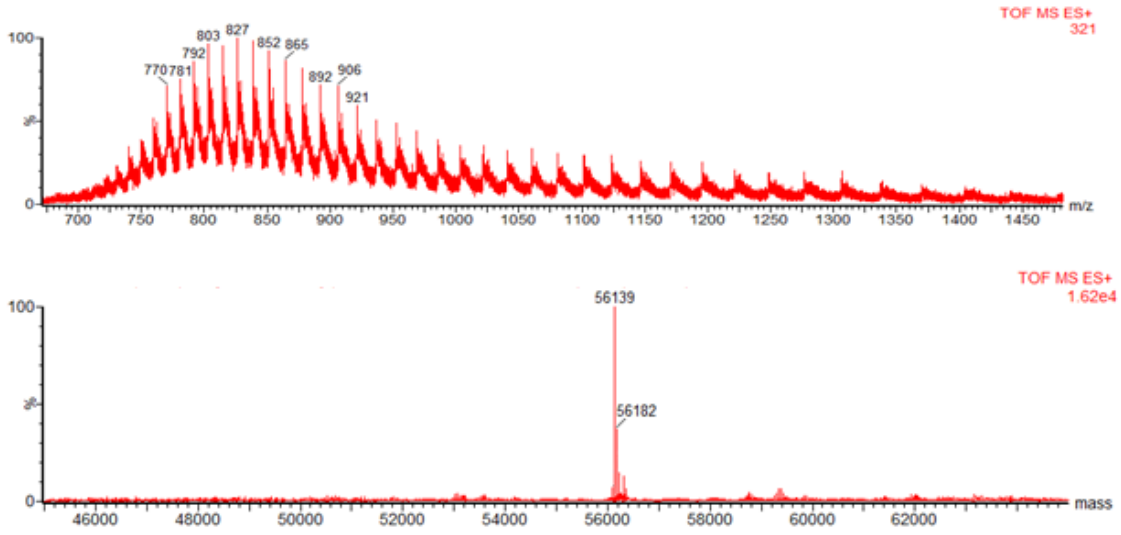


Figure 8: Expected mass of the protein = 56138.8 Da, found 56139 Da.

3.2.4 Heparin oligosaccharide preparation by lyase action on heparin

With the heparinase-1 enzyme in hand, degradation of heparin was carried out as reported in literature[30]. Briefly, 1 mg of enzyme was added to 10 mg of porcine intestinal mucosa-derived heparin dissolved in heparinase digestion buffer (0.1 M sodium acetate and 0.1 mM calcium acetate, pH 7). The reaction was incubated at 37 °C. The extent of degradation was determined at regular time intervals by running a high-density polyacrylamide gel (sugar gel) (see **Experimental Section** for preparation procedure). After sufficient degradation had occurred, the sample was lyophilized and size exclusion purification was done of the mixture on a Superdex-30 column (**Figure 9**).

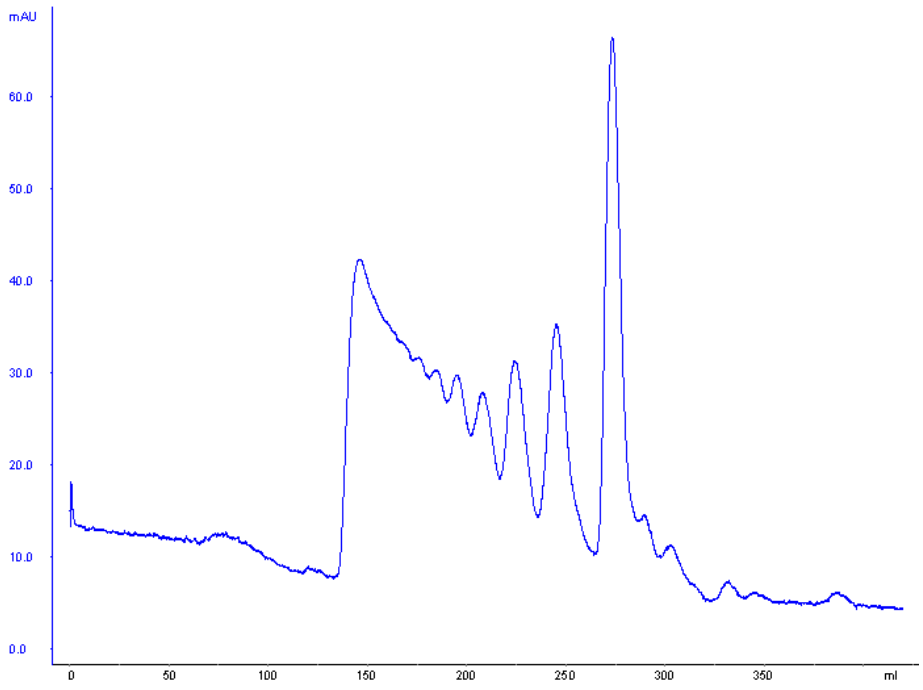


Figure 9: Size exclusion purification of Heparinase-1 treated heparin on S30 column. **1** represents heparin octasaccharide, **2** is heparin hexasaccharide, **3** is heparin tetrasaccharide and **4** is heparin disaccharide**.** Oligosaccharides larger than octasaccharide are seen before the octasaccharide peak with decreasing resolution.

The peak corresponding to disaccharide was collected separately, lyophilized and then taken for the next step of strong anion exchange (SAX) purification. For this purpose, standards of heparin disaccharide were purchased from Dextra Laboratories, UK, which were used to provide the reference retention time for purification of the required disaccharide from the size exclusion purified mixture.

SAX purification was carried out using a ProPac PA1 column (resin based strong anion exchange column from Thermo Scientific) with 2 M sodium chloride. Since the trisulfated disaccharide **2** is highly charged, it eluted much later on the chromatogram (**Figure 10**).

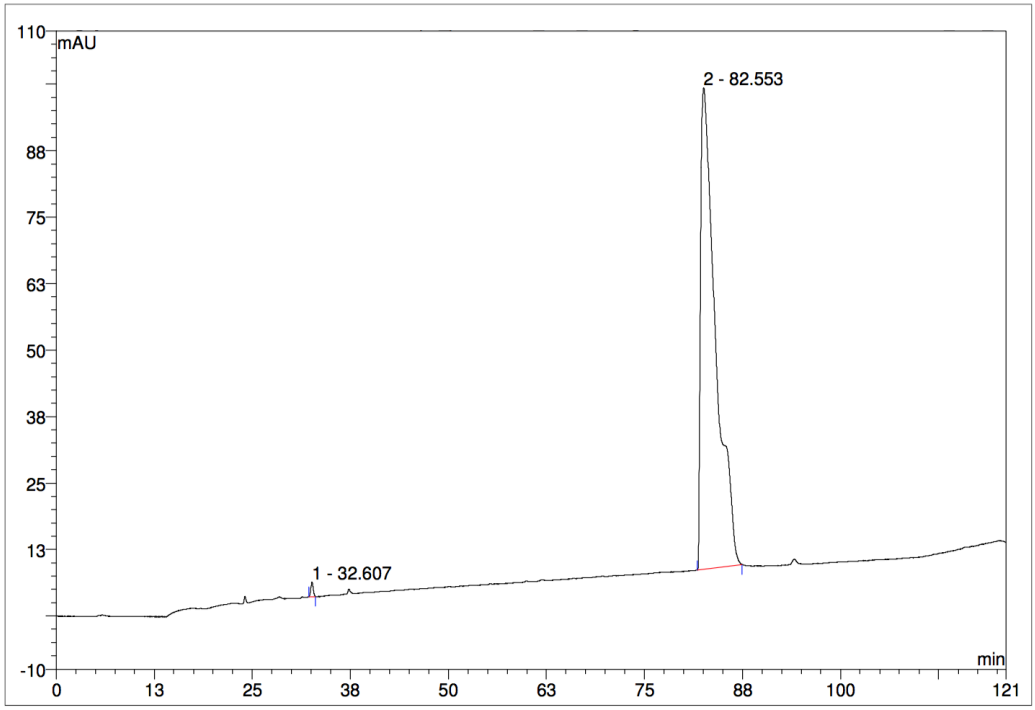


Figure 10: Strong anion exchange purification of disaccharide fraction from size exclusion purification. Desired product peak seen at 82.5 min retention time. Product absorbs at 232 nm because of the presence of unsaturation.

The appropriate peak was then collected, lyophilized and desalted. (**Figure 11**)

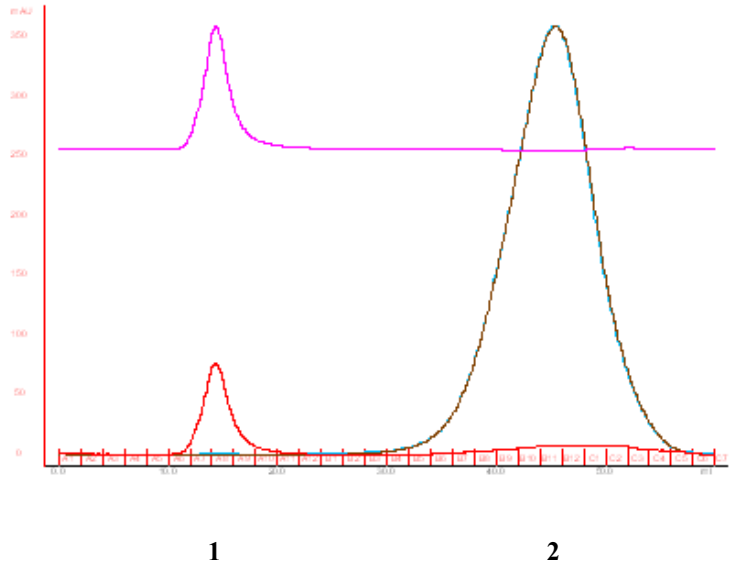


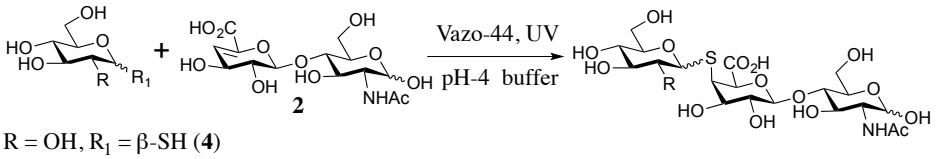
Figure 11: Desalting of HPLC purified fraction on Sephadex G-25 column. **1** represents peak corresponding to heparin disaccharide, monitored at both 210 nm and 232 nm, **2** represents peak corresponding to salt detected by its conductivity.

Consequently, fractions corresponding to the disaccharide were lyophilized to afford the pure disaccharide **3**. It should be noted that the complete preparation and purification procedure for these heparin derived building blocks is pretty elaborate and time consuming (taking more than two weeks as reported)[30]. Also limited amount of oligosaccharides are obtained at a time.

With these in hand, we went forward toward the synthesis of the thio-trisaccharide heparin/HS mimics.

3.2.5 Reaction of thio-monosaccharides with a non-sulfated HS disaccharide

We started with the reaction of thiosugars with the simpler non-sulfated HS disaccharide **2**. For this purpose, the three monosaccharide thiols **4**, **5** and **6** were used to create a library.



R = OH, R_1 = β-SH (**4**)
R = NHAc, R_1 = β-SH (**5**)
R = OH, R_1 = α-SH (**6**)

Scheme 6: Reaction of HS derived non-sulfated disaccharide with monosaccharide thiols to prepare trisaccharides. Yields shown in **Table 2**.

The conditions developed for thiol-ene radical reaction in aqueous conditions were used for the coupling reactions. With a few optimizations it was found that 10-12 equivalents of the thiosugar were needed to completely consume the unsaturated disaccharide, with the photoinitiator amount as 1 equivalent. The yield for the coupling reaction with the α-thiosugar was much less as compared to the other thiosugars. This observation corroborated the trend seen for reaction with the model disaccharides.

Since the amount of the starting disaccharide available per reaction was limited (~5 mg) owing to the elaborate preparation and purification process, efforts were made towards developing an efficient purification strategy. For all the purification optimizations, the coupling product formed between **5** and **2** was used. Purification attempts by normal silica chromatography led to severe loss of product. Attempts to purify using reverse phase chromatography with a Jupiter Proteo column (C12 based reverse phase column), was unsuccessful in separating the excess thiosugar and the product. Reverse phase conditions with a

volatile ion-pairing reagent has been reported for separation of heparin oligosaccharides[31]. Separation between the product and the excess thiosugar could not be achieved with this condition. Use of an amino column, which is known to furnish efficient separation of deprotected sugars, in some cases, (in the form of Luna amino column) proved to be unsuccessful too[32]. Next a Phenomenex STRATA strong anion exchange cartridge was used. A protocol using several washing steps with water followed by a 250 mM ammonium formate buffer wash was developed to purify the product (**Experimental Section 3.4.2**). The neutral thiosugar would be washed off with the water wash while the charged product would be eluted out with the ammonium formate buffer wash. Pure product was obtained by this method as determined by TLC and mass spectrometry. However, NMR taken after repeated lyophilisation to remove the ammonium formate, showed the presence of several large unrelated peaks (**Figure 12**). On further investigation, the four unrelated peaks observed were found to have an integration ratio of 2:2:2:9. These were suspected to be from the resin leaching out from the cartridge column. The peaks corresponded to the six methylene protons present on the linker attached to the resin and nine methyl protons present on the methyl carbons of the quarternary ammonium nitrogen (**Figure 13).** Prewashing the column with ammonium formate and sodium chloride did not stop the leaching process from contaminating the sample.

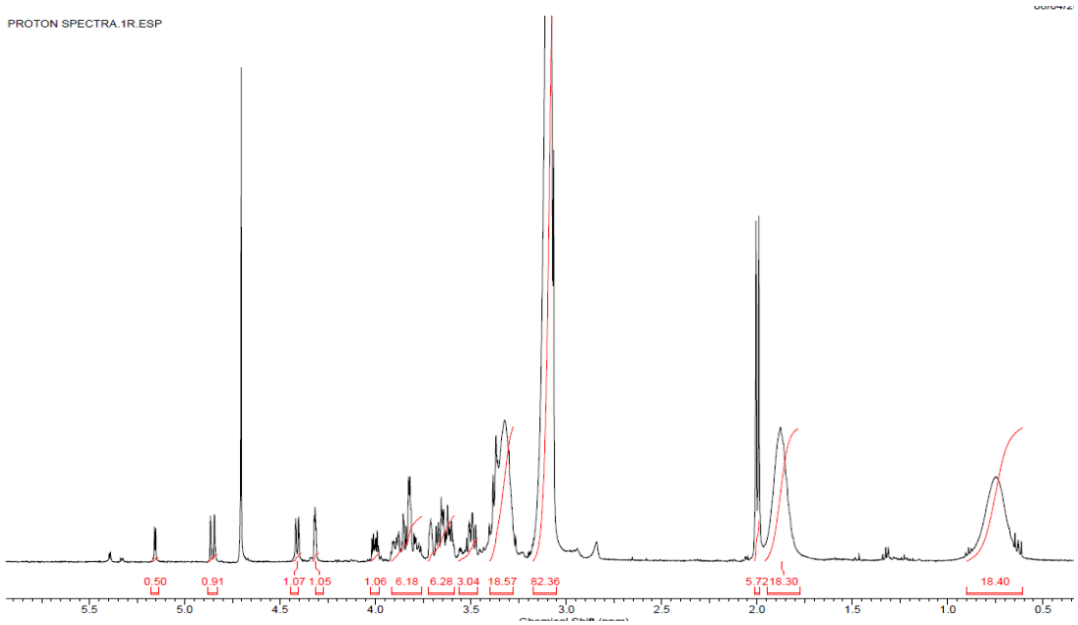


Figure 12: NMR spectra for the trisaccharide formed by coupling **5** with **2** and purified using Phenomenex SAX cartridges show extra peaks. Peaks with integration 18.57,18.30,18.40 and 82.36 (present in the ratio of 2:2:2:9) are suspected to be from the leached products from the cartridge.

Thereafter, we resorted to purification of the product using strong anion exchange column on HPLC using ammonium formate as the buffer. Even though the product lacks a strong chromophore, the limited absorption of the –NHAc group at 220 nm[33-34], allowed product detection by UV.

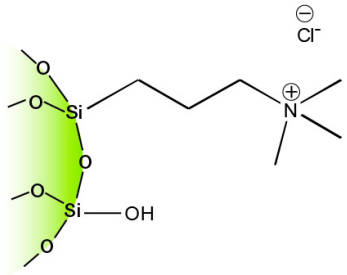
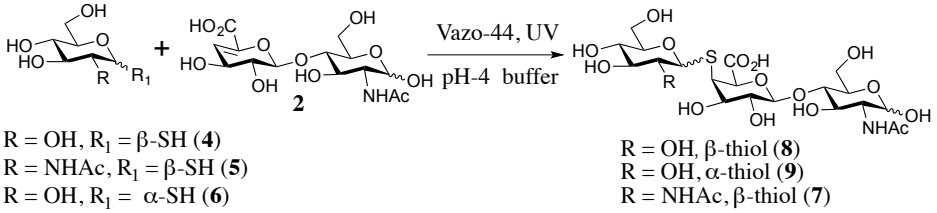


Figure 13: Structure of the quarternary ammonium resin on the SAX cartridge (structure adopted from Phenomenex product website)[35]. This is suspected to have leached out and co-eluted with product.

The recovery yield was still low because of manual collection from the HPLC system. Consequently, the purification run was re-optimized on a Mono Q SAX (strong anion exchange) column on FPLC. In addition to the UV chromatogram obtained, the individual fractions were also checked by mass spectrometry. This increased the recovery yield greatly. Ultimately a small set of trisaccharides was prepared with these optimized conditions as enlisted in **Table 2**.



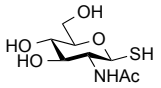
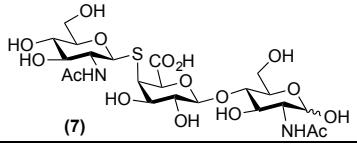
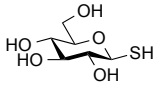
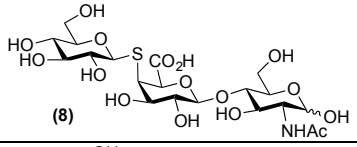
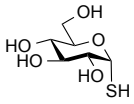
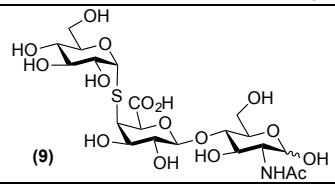
Entry	Thiosugar	Product	Yield
1		 (7)	78%
2		 (8)	73%
3		 (9)	48%

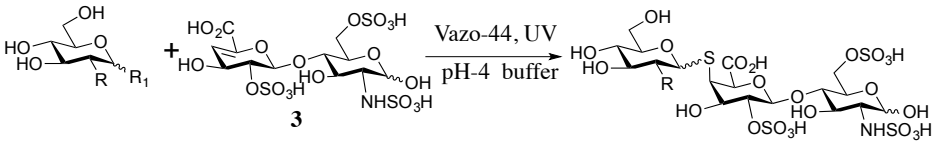
Table 2: Survey of trisaccharides formed by coupling HS disaccharide with monosaccharide thiols.

3.2.6 Reaction of thio monosaccharides with sulphated heparin disaccharide

Before carrying out UV-induced radical thiocoupling reactions with the trisulfated disaccharide **3,** it was necessary to verify that the sulfates were stable under the reaction conditions. The lability of sulfate functional groups has been widely reported in literature, with the *N*-sulfate group being most labile followed by the *O*-sulfates[36]. The desulfation occurs in a wide range of conditions, including exposure to highly acidic as well as basic conditions[37-40]. To the best of our knowledge, however, there has not been any study undertaken to investigate the stability of the sulfate groups under radical conditions in a mildly acidic buffer. For this purpose,

3 was used in control reactions, where it was exposed to identical conditions as for the non sulfated HS disaccharide coupling reactions, namely Vazo-44, pH 4 buffer and UV irradiation, but without the thiosugar. HPLC and mass spectrometry analysis, however revealed the sulfates to be quite stable with no smaller mass or change in retention time in the HPLC chromatogram, being visible which might indicate desulfation.

Thereafter, reactions were performed with the three thiosugars as in the case of the non-sulfated disaccharide as depicted in **Scheme 7**.



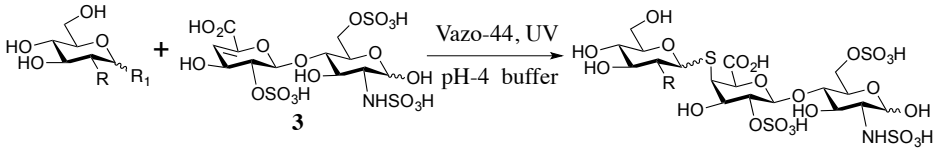
R = OH, R_1 = β-SH (**4**)
R = NHAc, R_1 = β-SH (**5**)
R = OH, R_1 = α-SH (**6**)

Scheme 7: Reaction of heparin derived trisulfated disaccharide with monosaccharide thiols to prepare trisaccharides.

Compared with the reaction conditions, optimized for the non-sulfated HS derived disaccharide **2,** it was realized that a higher amount of thiosugar was required (16-18 equivalents for the β-sugars and 20 equivalents for the α-thiosugar) to completely consume the disaccharide. The photoinitiator amount used was 1.2 equivalent. The higher thiosugar requirement is presumably due to the deactivating influence exerted by the highly electronegative sulfate groups, which renders the disaccharide much less reactive. In this context, it is worthwhile to note that that the deactivating influence of sulfate groups in case of coupling between *N*-methylaminooxy groups and hemiacetal functionality of the sugars has been reported[41]. When the starting disaccharide was completely consumed, as determined by the disappearance of the alkene peak in crude NMR, the reaction mixture was lyophilized several times to remove ammonium formate from the buffer. The presence of the product was verified by high resolution

mass spectrometry. Thereafter, purification was attempted with the optimized conditons found for the non-sulfated HS trisaccharide products. Unfortunately, almost the entire product was lost on the column as the sulfate groups bind too strongly with the SAX column to be dislocated using 250 mM ammonium formate buffer. Increasing the buffer concentration to 2M ammonium formate, improved the yield marginally. Thereafter, the column was changed to a DEAE (weak anion exchange) column at which point, the yield increased substantially to allow NMR characterization. With these conditons in hand, a similar trisaccharide set as the one obtained from **2**, was synthesized[a] (**Table 3**).

[a] The final reactions for synthesis of trisaccharides **10**, **11** and **12** with the optimized conditions were done by Dr. Wei-Min Liu. Hence the procedure for synthesis is mentioned only with the HRMS data instead of the full characterization.



R = OH, R_1 = β-SH (**4**)
R = NHAc, R_1 = β-SH (**5**)
R = OH, R_1 = α-SH (**6**)

R = OH, β-thiol (**11**)
R = OH, α-thiol (**12**)
R = NHAc, β-thiol (**10**)

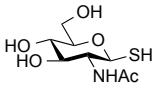
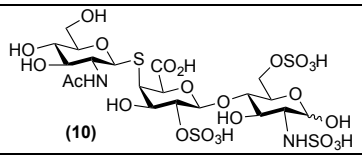
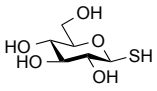
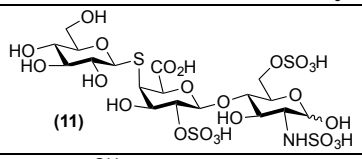
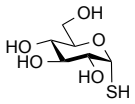
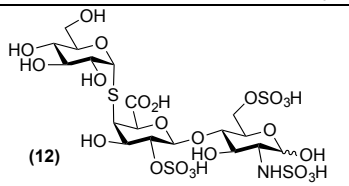
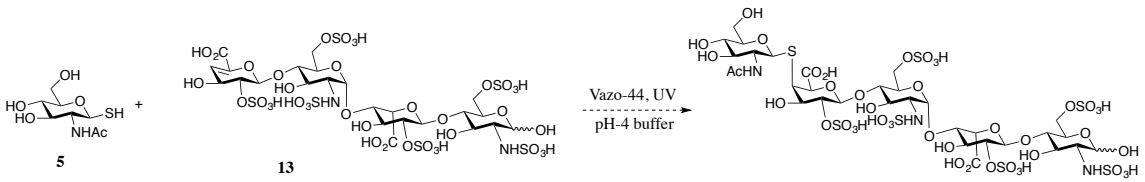
Entry	Thiosugar	Product	Yield
1.		 **(10)**	**57%**
2.		 **(11)**	**52%**
3.		 **(12)**	**17%**

Table 3: Library of trisaccharides formed by coupling heparin disaccharide with monosaccharide thiols.

3.2.7 Toward making a thio-linked heparin pentasaccharide mimic

With the library of heparin trisaccharides synthesized, we next turned our attention toward synthesizing heparin thio pentasaccharide mimic using a mono and tetrasaccharide coupling strategy. Here a heparin derived tetrasaccharide and a thio monosaccharide were utilized as shown in **Scheme 8**. For this purpose, heparin was depolymerized into smaller fragments using heparinase-1 as mentioned in **Section 3.2.4.** Size exclusion purification was carried out, and the fractions containing heparin tetrasaccharide was collected (**Figure 9**). Strong

anion exchange purification was carried out consequently with 2M sodium chloride using a Propac PA1 column. After lyophilization, the sample was desalted using a HiPrep 26/10 desalting column.



5 13

Scheme 8: Proposed reaction of β-thioGlcNAc with heparin derived tetrasaccharide.

Reactions were then carried out using the optimized conditions obtained for the synthesis of trisaccharides from the heparin-derived disaccharide, **3.** Unsurprisingly, a substantial amount of the tetrasaccharide was left unreacted, as determined from crude NMR, as the number of sulfate groups in the tetrasaccharide is double that in the disaccharide. However, increasing the thiosugar to even 40 equivalents and the photo initiator concentration to 4 equivalents did not lead to full consumption of the tetrasaccharide. Increasing the reaction time to 36 h did not help either. It was therefore realized that because of the deactivating influence of the sulfates, highly-sulfated oligosaccharides larger than disaccharides are difficult substrates for this coupling reaction.

In order to verify that the reaction did not proceed because of the deactivating nature of the sulfate groups, the next step would be to prepare a desulfated version of tetrasaccharide **13** and perform the thio-coupling reaction with it. Chemical approaches for 2-*O*-desulfation[42], 6-*O*-desulfation[43] and *N*-desulfation[39] have been reported and these could be used to generate the desulfated tetrasaccharide. The product thereby generated can also be sulfated at specific positions using sulfotransferases to prepare heparin thio-mimics.

Nevertheless, with the trisaccharide mimics in our hand, we proceeded towards the biological studies of these molecules. It should be mentioned here that for biological studies we are using the unnatural anomer (β as compared to α) with an unnatural stereochemistry of linkage (axial as compared to equatorial). We persisted with the unnatural anomer of GlcNAc because of synthetic convenience. It was found during experimentation that the α anomer (which is present in native heparin) gave extremely low yield for the thiol-ene reaction. The exclusive axial stereochemistry of linkage during the radical reaction was also found during experimentation and it can be explained by the reason given in chapter-2. We persisted with the wrong stereochemistry of linkage since it was reasoned that one wrong linkage stereochemistry in a long polysaccharide of heparin should not affect the biological activity of heparin to a great extent (only long polysaccharide chains of heparin have been shown to have significant biological activity). Of course, this can be verified only by extensive modelling studies and co-crystallization studies of this unnatural heparin mimic with protein binding partners, where the electrostatic interaction between the electronegative groups on the polysaccharide and the basic amino acids in the protein can be studied. Also, it is plausible that this distortion in the resultant polysaccharide chain because of the unnatural linkage stereochemistry can give rise to exclusivity in terms of the interaction of this unnatural polysaccharide with its protein-binding partners. This can in turn exclude side effects of heparin when it is being used for medicinal purposes. It would therefore be interesting to study the biological activity of heparin with an unnatural linkage stereochemistry. In addition we also wanted to explore if the thio-linkage or monosaccharides in the form of glucose instead of GlcNAc have any interesting effect in terms of their interaction with the protein-binding partners.

3.2.8 Biological studies with the synthesized trisaccharides

Heparin is known to interact with a plethora of proteins mainly through electrostatic interaction and hydrogen bonding interaction[44-45]. Amongst other proteins, heparin is known to interact with fibroblast growth factors[46], coagulation factors and their inhibitors[47] and chemokines[48]. Apart from this, β-secretase-1 (BACE-1), an enzyme implicated in the formation of amyloid β-peptide from amyloid precursor protein (APP), which leads to Alzheimer's disease, has been reported to interact with heparin[49]. BACE-1 has been known to interact with HS-derived oligosaccharides containing *N*-acetyl groups as well as heparin-derived oligosaccharides containing *N*-sulfates, along with *O*-sulfates[50-51]. Thus BACE-1 (expressed in the Turnbull laboratory) was chosen as the target, to investigate the interaction of our synthesized oligosaccharides.

Various methods have been used to study ligand protein interaction, which includes enzyme linked immunosorbent assays (ELISA)[52], surface plasmon resonance[53], isothermal titration calorimetry[54], differential scanning calorimetry[55] and circular dichroism[56] among others. However, most of these approaches suffer from the fact that they are either not high throughput or require relatively large amounts of sample. Differential scanning fluorimetry (DSF)[57-58], on the other hand, is highly sensitive and also it requires low amounts of compound. Its main drawback is that it is not a quantitative assay. This approach is based on the assumption that proteins experience an increase in stability on binding to their ligands. Consequently this stability is reflected in an increase in the denaturation or melting temperature of the protein-ligand complex as compared to the protein itself. In order to detect the denaturation temperature, a hydrophobic dye (in the form of SYPRO orange) is most commonly used as it can bind to the less exposed hydrophobic regions of a protein that is revealed when it denatures in aqueous

solution. A reporter function is therefore executed by the dye, which indicates the denaturation temperature by an increase in fluorescence as it gets to bind with a hydrophobic substrate. Plotting the intensity of fluorescence of the dye against the temperature to which the protein-complex is exposed, generally gives a sigmoidal curve (**Figure 14**). The temperature at which the protein unfolds is referred to as the melting temperature (T_m). A change in the stability of the protein-ligand complex, as compared to the protein, is indicated by a shift in T_m while the degree of shifting indicates the stabilization extent and consequently the relative affinity of the ligand for the protein.

Since the amounts of the trisaccharides obtained from the coupling of HS disaccharide **2** and heparin-derived disaccharide **3** with various thio-sugars were quite limited, we decided to go for DSF for initial screening.

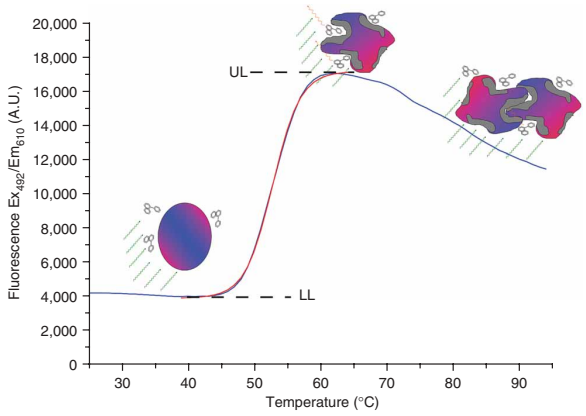


Figure 14: Typical sigmoidal curve obtained for plot of fluorescence intensity against temperature. The state of the protein under study, at each stage of the transition is also depicted. (Adapted from Niesen et al., Nat. Protoc., 2, 2007, 2212-2221[58])

3.2.8.1 Binding studies with BACE-1

Before the actual trisaccharides were used, a control study was set up with standards of heparin disaccharides, tetrasaccharides, hexasaccharides, octasaccharides and decasaccharides (dp2, dp4, dp6, dp8 and dp10) obtained by size exclusion purification of heparinase-1 degraded heparin. DSF studies indicated however, that there was only a marginal difference in binding affinity amongst the various standards. The binding was weak, even for a heparin-derived decasaccharide normally reported to have a half maximal effective concentration against BACE-1 as 10 μg/ml[59]. Considering the reported versus our observed results, the lack of binding was unexpected, as it is known that the binding affinity of heparin oligosaccharides for BACE-1 increases with increase in length of the oligosaccharide[50]. Moreover, dp2 did not seem to bind with BACE-1 at all, since the melting temperature shift was found to be less than the error bar (thus proving the data to be insignificant).

To troubleshoot this result, we investigated different concentrations. A titrimetric assay was carried out with different concentrations of heparin, as a control. It was observed that changing the concentration of heparin had no significant effect in the binding of heparin with BACE-1. The binding of heparin with BACE-1 was not found to be significant either. Using BACE-1 obtained from commercial sources (R&D labs) did not improve the situation. Consequently, we decided to change the protein substrate to fibroblast growth factors.

3.2.8.2 Binding studies with fibroblast growth factors

The interaction of fibroblast growth factors (FGFs) especially FGF-1 and FGF-2 with various heparin oligosaccharides has been reported previously[60-62]. Heparin is known to facilitate dimerization of FGF-1 on binding to the sugar, which then binds to two units of fibroblast growth factor receptor 2c (FGFR2c). Bound FGFR2c then initiates a signaling cascade leading to a strong mitogenic response[63]. Similarly, FGF-2 has been reported to interact with fibroblast growth factor receptor 1 (FGFR-1) inducing mitogenic activities[64-65]. For both of these binding events, DSF studies have also been reported using differently-sized heparin oligosaccharides[66].

Accordingly, we initiated binding studies with both of FGF-1 and FGF-2. Control studies were done initially with heparin to verify that there was a concentration dependent change in stability of the protein-heparin complex with changes in heparin concentration. With FGF-1, the shift in melting temperature was significant, with a change of 12.9 °C and 22.75 °C when 2:1 and 5:1 molar ratio of heparin was used (since heparin is heterogeneous, the average molecular weight provided by the supplier was used for calculating molar values) respectively (**Figure 15**). However, with these and additional concentrations of heparin, FGF-2 was stabilized to such an extent by heparin that the reading went beyond the measurable parameters of the assay. Therefore, FGF-1 was subsequently used in all the studies.

We then proceeded to perform binding studies with the three sulfated trisaccharides synthesized, at a 5:1 concentration of the trisaccharides to FGF-1. This is because heparin has been reported to interact with growth factors through sulfate groups and carboxylic groups with the basic amino acid residues of the protein[67]. However, no changes in melting temperature were observed indicating no significant binding of the synthesized trisaccharides with FGF-1 (**Figure**

16). The starting sulfated disaccharide **3** showed no binding with FGF-1 either. Increasing the concentrations of the sulfated oligosaccharides did not have any effect on the melting temperature shift. Despite the fact that there are three sulfate groups able to interact with the basic amino acid residues of the protein in the case of **3**, **10, 11** and **12**, the lack of significant binding is presumably because of the small size of the sulfated domain. The minimal length of heparin oligosaccharide binding to FGF-1 has been reported to be a hexasaccharide[68]. Thus the small disaccharide unit may not be interacting effectively with the active cleft of the growth factor protein.

However, there are reports of sulfated heparin mimetic disaccharides binding to FGF-1 with K_d in the range of 48-77 μM as measured by surface plasmon resonance solution affinity assay[69]. The binding constant has been observed to decrease with increase in sulfate groups. Also, Hu et al. has reported binding of synthesized similar heparin disaccharides with FGF-1 with K_d ranging from 9 to 21 μM. The binding constant has been reported to follow the same trend with increasing sulfates. The study was done using isothermal titration calorimetry[70].

Nonetheless, since longer sulfated oligosaccharides could not be synthesized by this approach, we sought to measure the binding activities with the synthesized trisaccharides.

To further probe any effect the thio-linkage and thioglucose, has, in terms of binding with FGF-1, the synthesized non-sulfated trisaccharides and starting non-sulfated disaccharide **2** were also studied for their binding activity with FGF-1. No melting temperature shift was observed in this case too. This observation is presumably because in the absence of sulfate groups, there was not a significant electrostatic interaction between the sugar and the basic amino acid residues in the protein.

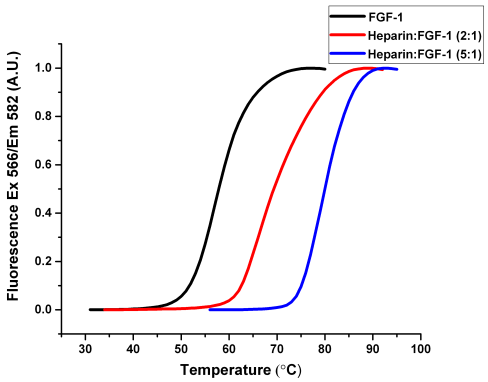
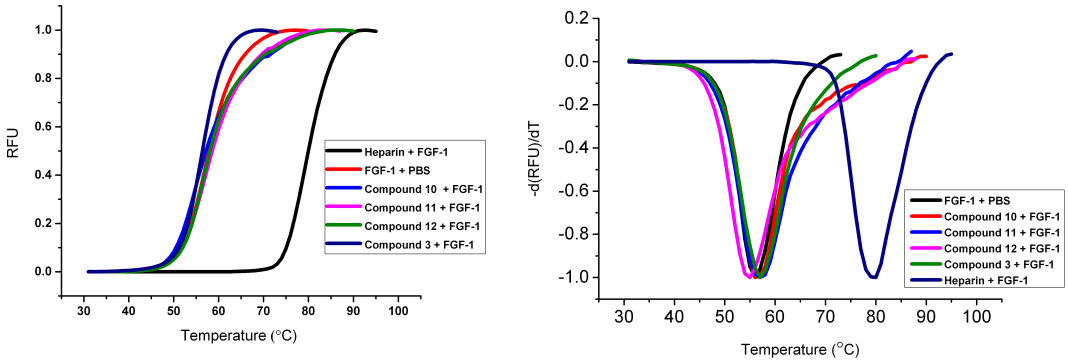


Figure 15: Heparin used in molar ratio of 2:1 and 5:1 w.r.t. FGF-1. Significant melting temperature shift is observed w.r.t. just FGF-1 (control)



(A) **(B)**

Figure 16: **(A)** DSF studies done with sulfated disaccharide **3** and its trisaccharide derivatives. Heparin has been used as a reference **(B)** 1st order derivative of the same plots does not show significant difference in melting temperature on being treated with the compounds.

It was therefore realized that in order to observe significant binding activity, larger sulfated oligosaccharides were needed. Since we had experienced problems extending the coupling reaction to heparin derived tetrasaccharide (as explained before), because of the deactivating nature of the electronegative sulfate groups, it is necessary to complement this chemical approach with an enzymatic one. This is the topic of discussion in Chapter-4 in which

we have investigated the possibility of utilizing enzymes to extend on the oligosaccharide chain towards building larger heparin analogues.

3.3 Summary

We have succeeded in using the thiol-ene radical reaction, discussed previously, to couple thiosugars with heparin- and heparan sulfate (HS)-derived oligosaccharides. This serves as a proof of concept towards building larger glycosaminoglycan oligosaccharides. Importantly, the reactions could be carried out in completely aqueous conditions, thus enabling coupling reactions to be carried out directly on heparin and HS natural products without the need for protective group manipulation.

Sulfated thiosugars would be an interesting reactant with which to carry out these coupling reactions, where it would be important to investigate potential deactivating effects the sulfates might have on the electrophilic nature of the thiosugar. Also, by virtue of being sulfated, the products would be better thiomimics of heparin/HS. We envision that this reaction can also be used for tagging of various types of probes to the non-reducing end of glycosaminoglycan derived oligosaccharides, leading to potentially interesting substrates for screening and diagnostic assays.

In addition, the non-sulfated HS analogues can also be potentially used for further modification by enzymes such as *N*-deacetylase/N-sulftotransferase[71], C-5 epimerase[72] and *O*-sulfotransferases[73] to synthesize different heparin/HS thio-mimics. Since, this strategy involved coupling two carbohydrate units, the approach might also find utility in coupling different

glycosaminoglycans together, leading to the novel hybrid glycosaminoglycans with hitherto undiscovered activities.

The synthesized compounds can be further tested for biological activities using other assays, such as displacement assays in which the synthesized oligosaccharides would be used to compete with radiolabelled heparin for binding with FGF-1. This should give more quantitative information about the interaction between the oligosaccharides and FGF-1[74]. Alternately, isothermal titration calorimetry may also be used to yield quantitative information about the carbohydrate protein interaction[61].

Our further efforts include extending these heparin and HS thio-mimics chemoenzymatically. This would enable us to get better binding results. A chemoenzymatic approach in this context would be useful because it should not be affected by the deactivating nature of the sulfate groups. Moreover, sulfate groups can potentially be added enzymatically to the neutral thiosugar, after coupling with the heparin derived unsaturated oligosaccharide. This would overcome the potential deactivating influence of the sulfates on the thio sugar towards the thiol-alkene radical reaction. The chemoenzymatic strategy for the extension of the thio mimics is discussed in Chapter-4.

3.4 Experimental Section

3.4.1 Synthesis of compounds

NMR data: Proton nuclear magnetic resonance spectra (^{1}H) were recorded on a Bruker DPX 400 (400 MHz) and Bruker DQX 400 (400 MHz) spectrometer. Carbon nuclear magnetic resonance spectra (^{13}C) were recorded on a Bruker DQX 400 (101 MHz) spectrometer. Spectra were fully

assigned based on chemical shifts using a combination of COSY, HSQC data and comparison with spectra of related compounds. All chemical shifts are referenced to the relevant residual solvent peak and are quoted in ppm with respect to the internal standard of tetramethylsilane (TMS). Resonances are described as s (singlet), d (doublet), t (triplet), q (quartet), quin (quintet), sxt (sextet), sept (septet), oct (octet), multiplet (m), app. (apparent) and br (broad). Coupling constants (*J*) are given in hertz (Hz).

*Mass spectra: M*ass spectra were recorded on a Fisons Platform VG Spectrometer (ESI) using electrospray ionisation in a positive (ESI+) or negative (ESI-) scan mode. *m/z* are reported with their percentage abundance.

Accurate mass Service

Analyses were performed using a Thermo Exactive mass spectrometer equipped with Waters Acquity liquid chromatography system. Instrument control and data processing were performed using Thermo Xcalibur Software. The system was calibrated on the day of the analysis and its mass accuracy with external calibration (as used for these experiments) is better than 5 ppm for 24 hours following calibration. The mass spec was operated using the heated electrospray (HESI-II) probe and resolution was set to 50,000. Electrospray source conditions were adjusted to maximise sensitivity. A mixture of 10% water, 89.9% methanol and 0.1% formic acid was used to transport samples to the mass spectrometer at a flow rate of 0.2 mL/min.

LC samples parameters

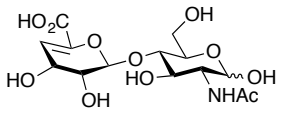
Analyses were performed using a Waters Xevo G2-S QToF mass spectrometer equipped with Waters Acquity liquid chromatography system. Instrument control and data processing were performed using Waters Masslynx 4.1 Software. The system was calibrated throughout

the analysis using Lockspray (TM) and mass accuracy is better than 5 ppm. The mass spec was operated using electrospray, source conditions were adjusted to maximise sensitivity.

Chromatography: Thin layer chromatography (TLC) was carried out on Merk Kieselgel $60F_{254}$ 0.2 mm precoated aluminium backed plates. Product spots were visualized with a combination of the following: 254 nm UV lamp, aqueous phosphomolybdic acid and $Ce^{(IV)}$ (2.5% phosphomolybdic acid hydrate, 1% cerium(IV) sulfate hydrate, and 6% H_2SO_4); ammonium molybdate (5% in 2M H_2SO_4). Retention factors (R_f) are reported with the solvent system used in parentheses. Flash column chromatography (FC) was carried out with Fluka Kiegselgel 60 220-440 mesh silica gel, with the solvent system used, in parentheses.

Proton and carbon NMRs of compounds were assigned using a combination of proton, carbon, COSY, HSQC and 1D-TOCSY studies.

1. Synthesis of compound (2):

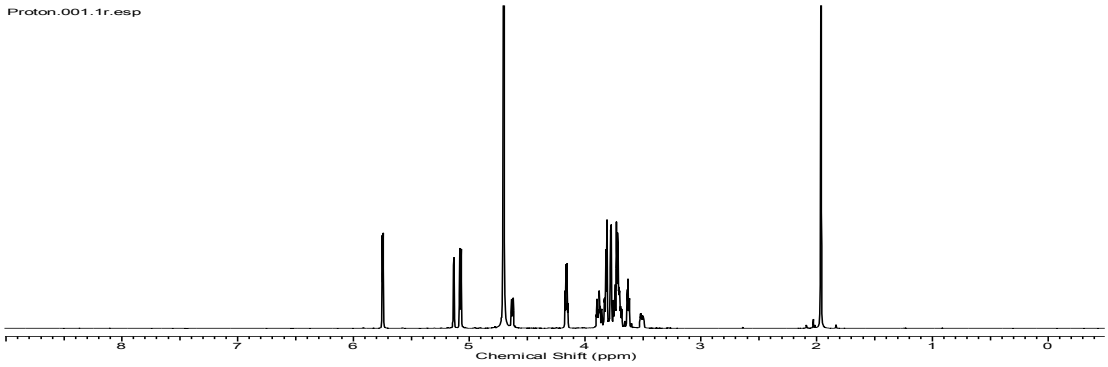


Disaccharide **2** derived by Heparinase-3 degradation of heparan sulfate, was obtained from the Turnbull laboratory. Hence its preparation is not reported here. Characterization data is as mentioned below:

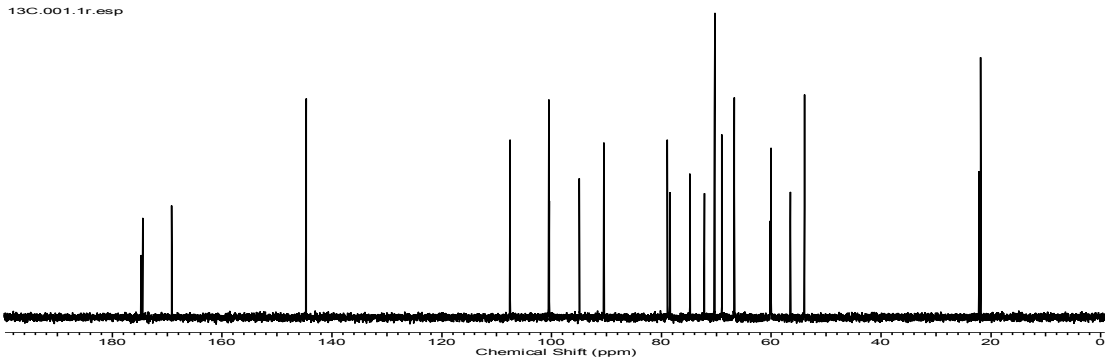
^{1}H NMR (400 MHz, D_2O): δ ppm 2.01 (s, 3H, -NHAc), 3.73-3.8 (m, 3H, H-2', H-3, H-4), 3.85 (dd, $J_{5,6a}$ = 3.5 Hz, $J_{6a,6b}$ = 12.6 Hz, 1H, H-6b), 3.87 (d, $J_{1,2}$ = 6.6 Hz, 1H, H-2), 3.89 (dd, $J_{5,6a}$ = 2.2 Hz, $J_{6a,6b}$ = 12.6 Hz, 1H, H-6a), 3.93 (ddd, $J_{5,6a}$ = 3.5 Hz, $J_{5,6b}$ = 2.2 Hz, $J_{5,4}$ = 10.1 Hz, 1H, H-5), 4.21 (dd, J = 3.5 Hz, J = 3.8 Hz, 1H, H-3'), 4.68 (d, J = 7.9 Hz, H-1β), 5.12 (d, $J_{1',2'}$ = 6 Hz, 1H, H-1'), 5.18 (d, J = 1.9 Hz, H-1α), 5.79 (d, $J_{4',3'}$ = 3.5 Hz, 1H, H-4'); ^{13}C NMR (101

MHz, D_2O): δ ppm 22.1 (NHCO<u>C</u>H_3), 53.9 (C-2), 56.5 (C-2β), 60.2 (C-6), 66.6 (C-3'), 69.0 (C-3/C-4/C-2'), 70.3 (C-5), 72.1 (C-4/C-3/C-2'), 78.4 (C-2'/C-3/C-4), 90.5 (C-1α), 94.9 (C-1β), 100.4 (C-1'), 107.6 (C-4'), 144.7 (C-5'), 169.1 (NH<u>C</u>OCH_3), 174.7 (<u>C</u>O_2H); ESI$^-$HRMS: calcd. for $C_{14}H_{21}NO_{11}$ (M-H^+) 378.1042, found 378.1047.

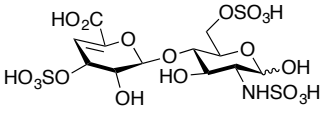
^{1}H spectra



^{13}C spectra



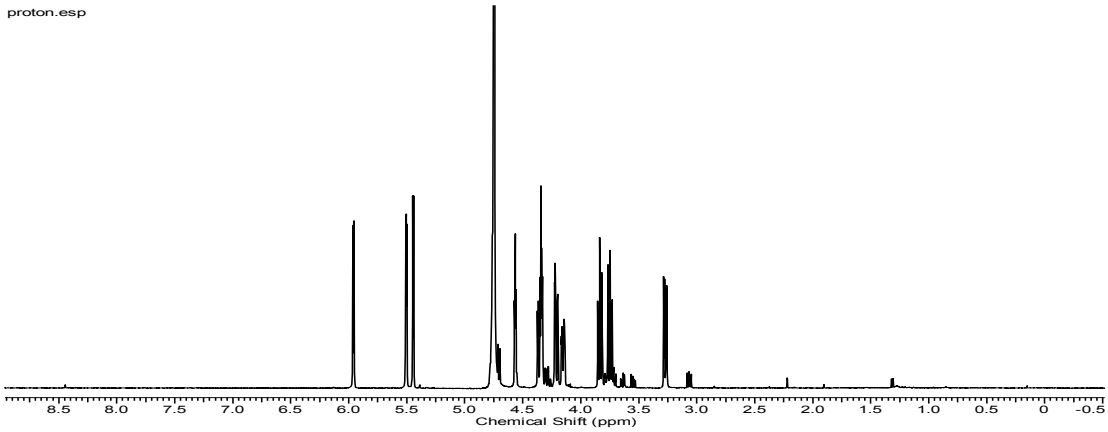
2. Synthesis of compound (3):



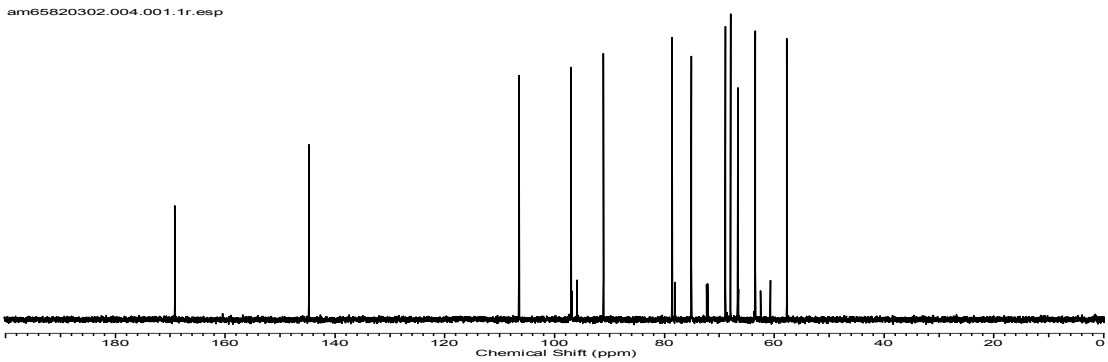
Heparin sodium salt from porcine intestinal mucosa (from Sigma Aldrich) (50 mg) was treated with 5 mg of Heparinase-1, in 5 mL heparin digestion buffer (0.1 M sodium acetate and 0.1 mM calcium acetate, pH 7) at 37 °C. The extent of degradation was determined at regular time intervals by running a sugar gel. The progress of the reaction was monitored by running sugar gels at regular intervals as well as by monitoring the UV absorbance of the reaction mixture at 232 nm. At the end of 72 hrs. the reaction was terminated by heating the mixture at 100 °C for 2 minutes. Thereafter, the protein was removed via vivaspin (10 KDa) and the flow-through was lyophilized and purified on a Superdex 30 size exclusion column (16 mm X 200 cm) using 0.5 M ammonium bicarbonate as buffer, running at 0.5 ml/min. The fraction corresponding to the disaccharide was then lyophilized and purified by means of SAX chromatography on a Propac PA1 column, using 2M NaCl and 100% water as buffer, running at 1 ml/min. The peak corresponding to the disaccharide was collected and lyophilized. Thereafter, desalting was done on a HiPrep 26/10 (Sephadex G-25) desalting column running at 5ml/min flowrate. The peak corresponding to the disaccharide was collected and lyophilized to yield the final pure compound.^{1}H NMR (400 MHz, D_2O): δ ppm 3.27 (dd, *J* = 3.6 Hz, *J* = 10.3 Hz, 1H, H-2), 3.75 (dd, $J_{3,4}$ = 8.9 Hz, $J_{3,2}$ = 10.3 Hz, 1H, H-3), 3.84 (dd, $J_{4,5}$ = 9.8 Hz, $J_{4,3}$ = 8.9 Hz, 1H, H-4), 4.15 (ddd, $J_{5,6a}$ = 2.1 Hz, $J_{5,6b}$ = 3.5 Hz, $J_{5,4}$ = 9.8 Hz, 1H, H-5), 4.21 (dd, $J_{6a,5}$ = 2.1 Hz, $J_{6a,6b}$ = 11.2 Hz, 1H, H-6a), 4.33 (dd, $J_{3',4'}$ = 4.4 Hz, $J_{3',2'}$ = 3.4 Hz, 1H, H-3'), 4.36 (dd, $J_{6b,5}$ = 3.5 Hz, *J6b,6a* = 11.2 Hz, 1H, H-6b), 4.57 (dd, *J2',1'* = 2.9 Hz, *J2',3'* = 3.4 Hz, 1H, H-2'), 4.7 (d, *J* = 8.3 Hz, H-1β), 5.45 (d, *J* = 3.5 Hz, 1H, H-1α), 5.5 (d, *J* = 3.5 Hz, 1H, H-1'), 5.9 (d, *J* = 4.4 Hz, 1H, H-

4'); ^{13}C NMR (101 MHz, D_2O): δ ppm 57.8 (C-2), 63.5 (C-3'), 66.7 (C-6), 67.9 (C-5), 68.8 (C-3), 75.2 (C-2'), 78.5 (C-4), 91.1 (C-1α), 95.9 (C-1β), 97.0 (C-1'), 106.5 (C-4'), 144.7 (C-5'), 169.1 (CO_2H); ESI-HRMS: calcd. for $C_{12}H_{19}NO_{19}S_3$ ($M-4H^+ +3Na^+$) 641.9099, found 641.9111

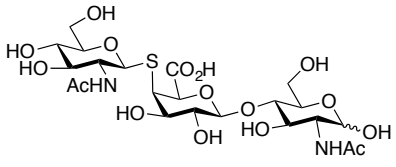
^{1}H spectra



^{13}C spectra



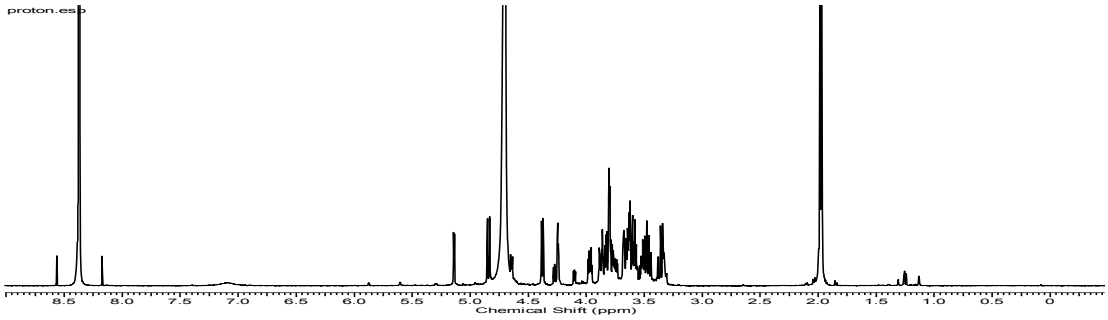
3. Synthesis of compound (7):



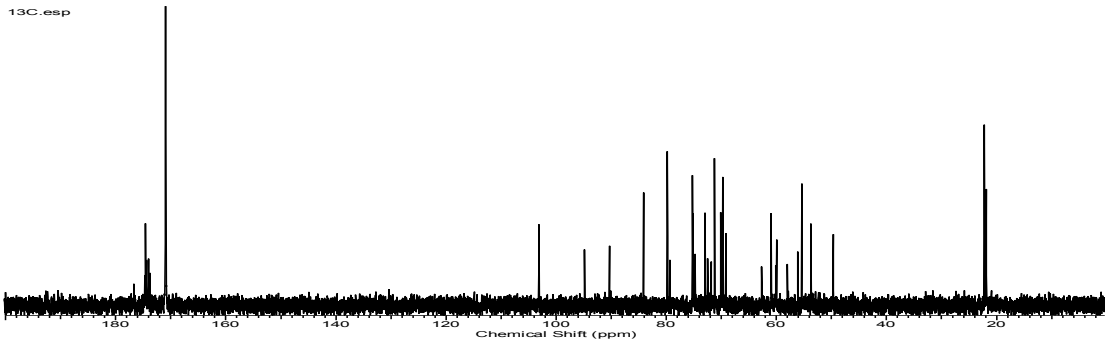
Compound **2** (5 mg, 0.02 mmol), was dissolved in 300 µL ammonium formate buffer (250 mM, pH-4). 2-Acetamido-2-deoxy-1-thio-β-D-glucopyranose (38 mg, 0.16 mmol) and Vazo-44 (6.5 mg, 0.02 mmol) were then added to it and the reaction mixture was irradiated under UV for 8 hrs. Completion of the reaction was verified using crude NMR. Thereafter, the reaction mixture was lyophilized several times to get rid of ammonium formate. Purification of the reaction mixture was done using a Mono Q column on an AKTA FPLC purifier system using 250 mM ammonium formate and water (flow rate = 1 mL/min). Individual fractions were checked by mass spec. to locate the product. Pure fractions were then lyophilized multiple times to remove ammonium formate salt and the product as a white powder was obtained in 78% yield. ^{1}H NMR (500 MHz, D_2O): δ ppm 2.01 (s, 3H, -NHAc), 2.03 (s, 3H, -NHAc), 3.36-3.43 (m, 2H, H-5''/H-3''), 3.51 (dd, $J_{4'',3''}$ = 9.05 Hz, J = 1.7 Hz, 1H, H-4''), 3.55 (dd, $J_{2',3'}$ = 9.8 Hz, $J_{2',1'}$ = 7.4 Hz, 1H, H-2'), 3.57-3.61 (m, 2H, H-3/H-4), 3.63-3.71 (m, 2H, H-6b''/H-2''), 3.68 (dd, $J_{4',3'}$ = 4.2 Hz, $J_{4',5'}$ = 2 Hz, 1H, H-4'), 3.78 (d, J = 3.4 Hz, 1H, H-2), 3.8 (dd, $J_{6b,6a}$ = 8.3 Hz, $J_{6b,5}$ = 3.4 Hz, 1H, H-6b), 3.84 (dd, $J_{6a,6b}$ = 8.3 Hz, $J_{6a,5}$ = 2.5 Hz, H-6a), 3.87 (dd, $J_{6a'',6b''}$ = 11.3 Hz, $J_{6a'',5}$ = 1.2 Hz, 1H, H-6a''), 3.88-3.9 (m, 1H, H-5), 3.97 (dd, $J_{3',2'}$ = 9.8 Hz, $J_{3',4'}$ = 4.2 Hz, 1H, H-3'), 4.25 (dd, $J_{5',4'}$ = 2 Hz, J = 3.4 Hz, 1H, H-5'), 4.42 (d, $J_{1',2'}$ = 7.9 Hz, 1H, H-1'), 4.69 (d, $J_{1,2}$ = 8.2 Hz, H-1β), 4.89 (d, $J_{1'',2''}$ = 10.4 Hz, 1H, H-1''), 5.18 (d, J = 3.3 Hz, 1H, H-1α); ^{13}C NMR (126 MHz, D_2O): δ ppm 21.8 (NHCO<u>C</u>H$_3$), 22.3 (NHCO<u>C</u>H$_3$), 49.6 (C-4'), 53.7 (C-2), 55.3 (C-2''), 55.4 (C-4/C-3), 59.8 (C-6), 60.9 (C-6''), 69.7 (C-3''/C-5''), 70.1 (C-5), 71.2 (C-4''), 72.9 (C-3'),

74.6 (C-2'), 75.1 (C-5'), 79.7 (C-3/C-4), 79.8 (C-5''/C-3''), 84.0 (C-1''), 90.1 (C-1α), 103.0 (C-1'); IR (neat) ν_{max} 760, 1058, 1379, 1574, 3084 cm^{-1}; ESI$^-$HRMS: calcd. for $C_{22}H_{36}N_2O_{16}S$ ($M-H^+$) 615.1707, found 615.1713

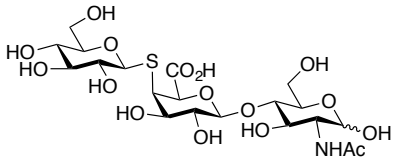
^{1}H spectra



^{13}C spectra



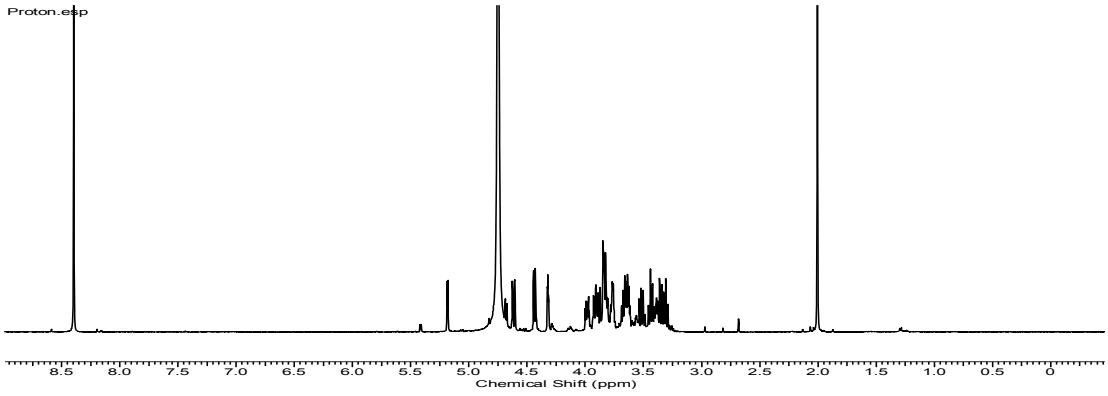
4. Synthesis of compound (8)



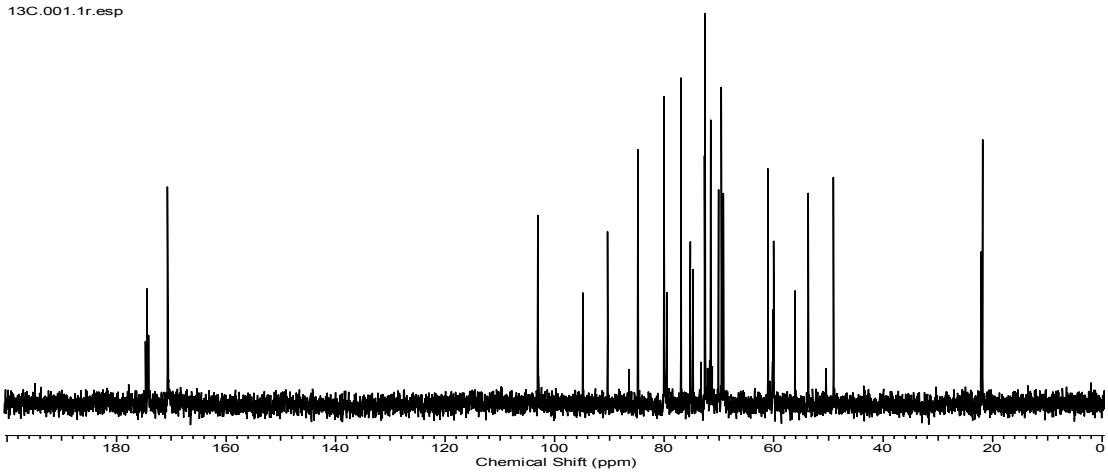
Compound **2** (5 mg, 0.02 mmol) was dissolved in 300 µL ammonium formate buffer (250 mM, pH-4). 1-Thio-β-D-glucopyranose (32 mg, 0.16 mmol) and Vazo-44 (6.5 mg, 0.02 mmol) were then added to it and the reaction mixture was irradiated under UV for 8 hrs. Completion of the reaction was verified using crude NMR. Thereafter, the reaction mixture was lyophilized several times to get rid of ammonium formate. Purification of the reaction mixture was done using a Mono Q column on an AKTA FPLC purifier system using 250 mM ammonium formate and water (flow rate = 1 mL/min). Individual fractions were checked by mass spec. to locate the product. Pure fractions were then lyophilized multiple times to remove ammonium formate salt and obtain the product as a white foam in 73% yield. ^{1}H NMR (500 MHz, D_2O): δ ppm 2.01 (s, 3H, -NHAc), 3.32-3.52 (m, 2H, H-5''/H-3''), 3.57 (dd, *J* = 9.5 Hz, *J* = 8.8 Hz, 1H, H-2''), 3.65 (dd, *J* = 10.8 Hz, *J* = 7.6 Hz, 1H, H-2'), 3.78 (dd, *J* = 16.3 Hz, *J* = 1.7 Hz, 1H, H-6b''), 3.87 (dd, *J* = 8.9 Hz, *J* = 8.1 Hz, 1H, H-2), 3.93 (dd, *J* = 5.6 Hz, *J* = 3.2 Hz, 1H, H-3), 4.02 (dd, *J* = 12.8 Hz, *J* = 5.4 Hz, 1H, H-4''), 4.13 (dd, *J* = 2.5 Hz, *J* = 4.8 Hz, 1H, H-2), 4.15-4.23 (m, 2H, H-2', H-6a), 4.31 (d, *J* = 2.8 Hz, 1H, H-6b), 4.33 (d, *J* = 11.3 Hz, *J* = 1.2 Hz, 1H, H-6a''), 4.42 (dt, *J* = 10.1 Hz, *J* = 3.2 Hz, 1H, H-4), 4.46 (ddd, $J_{5,6a}$ = 2.1 Hz, $J_{5,6b}$ = 3.5 Hz, $J_{5,4}$ = 9.8 Hz, 1H, H-5), 4.53 (dd, *J* = 3.7 Hz, *J* = 8.3 Hz, 1H, H-3'), 4.58 (d, $J_{1',2'}$ = 7.7 Hz, 1H, H-1'), 4.62 (d, *J* = 3.5 Hz, 1H, H-4'), 4.67 (dd, *J* = 2.3 Hz, *J* = 5.6 Hz, 1H, H-5'), 4.69 (d, $J_{1,2}$ = 8.2 Hz, H-1β), 4.89 (d, $J_{1'',2''}$ = 10.4 Hz, 1H, H-1''), 5.18 (d, *J* = 5.2 Hz, 1H, H-1α); ^{13}C NMR (126 MHz, D_2O): δ ppm 22.1 (NHCO<u>C</u>H$_3$), 54.5 (C-2), 56.4 (C-2β), 58.5 (C-2''), 60.1 (C-6), 61.5 (C-6''), 66.6 (C-3'), 68.0 (C-3), 70.3 (C-4''/C-5''), 71.6 (C-5), 72.1 (C-4), 75.6 (C-3''), 79.4 (C-2'), 79.8 (C-1''),

80.4 (C-5''/C-4''), 90.2 (C-1α), 95.9 (C-1β), 100.6 (C-1'), 67.6 (C-4'), 74.8 (C-5'), 169.1 ($NH\underline{C}OCH_3$), 174.7 ($\underline{C}O_2H$); IR (neat) ν_{max} 618, 633, 1072, 1596, 3044 cm^{-1}; ESI^-HRMS: calcd. for $C_{20}H_{33}NO_{16}S$ ($M-H^+$) 574.1447, found 574.1457

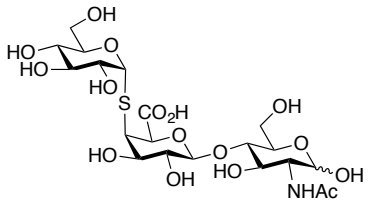
^{1}H spectra



^{13}C spectra



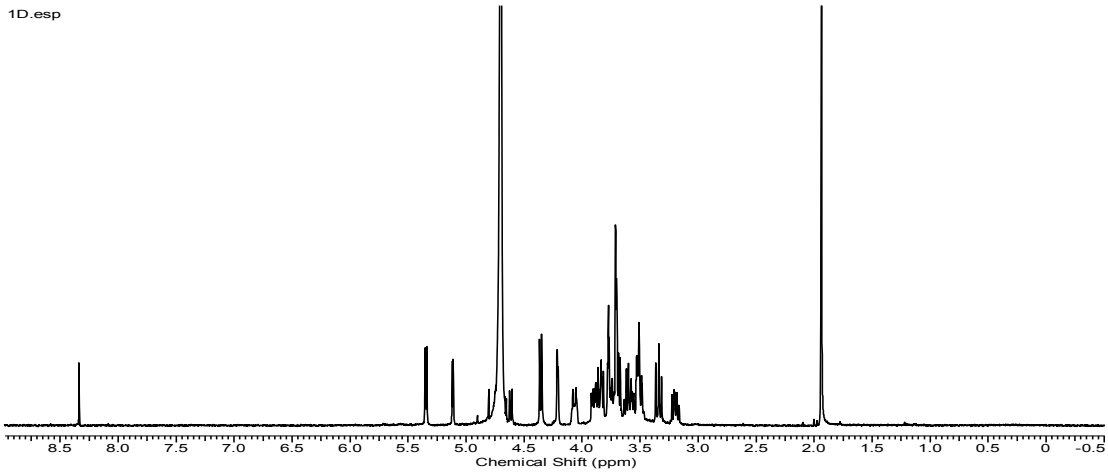
5. Synthesis of compound (9):



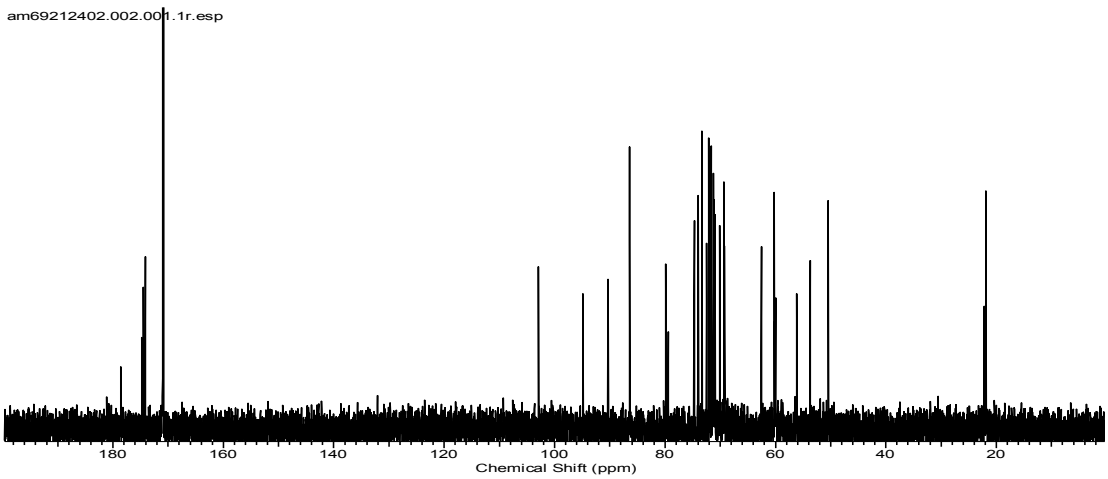
Compound **2** (5 mg, 0.02 mmol) was dissolved in 300 µL ammonium formate buffer (250 mM, pH-4).1-Thio-α-D-glucopyranose (32 mg, 0.16 mmol) and Vazo-44 (6.5 mg, 0.02 mmol) was then added to it and the reaction mixture was irradiated under UV for 9 hrs. Thereafter, the reaction mixture was lyophilized several times to get rid of ammonium formate. Purification of the reaction mixture was done using a Mono Q column on an AKTA FPLC purifier system using 250 mM ammonium formate and water (flow rate = 1 mL/min). Individual fractions were checked by mass spec. to locate the product. Pure fractions were combined and lyophilized multiple times to remove ammonium formate salt yielding desired product as a white foam in 48% yield. ^{1}H NMR (500 MHz, D_2O): δ ppm 2.00 (s, 3H, -NHAc), 3.35 (dd, *J* = 9.3 Hz, *J* = 9.9 Hz, 1H, H-4''), 3.57 (dd, *J* = 9.3 Hz, *J* = 9.7 Hz, 1H, H-3''), 3.58 (dd, *J* = 8.7 Hz, *J* = 7.5 Hz, 1H, H-2'), 3.59-3.62 (m, 2H, H-3/H-4), 3.64 (dd, *J* = 4.1 Hz, *J* = 1.8 Hz, 1H, H-4'), 3.74 (dd, *J* = 5.5 Hz, *J* = 12.2 Hz, 1H, H-6a''), 3.76 (dd, *J* = 5.5 Hz, *J* = 9.7 Hz, 1H, H-2''), 3.77 (d, *J* = 3.5 Hz, 1H, H-2), 3.78 (dd, *J* = 12.2 Hz, *J* = 2.3 Hz, 1H, H-6b''), 3.81 (dd, $J_{6b,6a}$ = 8.5 Hz, $J_{6b,5}$ = 3.2 Hz, 1H, H-6b), 3.85 (dd, $J_{6a,6b}$ = 8.4 Hz, $J_{6a,5}$ = 2.7 Hz, 1H, H-6a), 3.89-3.91 (m, 1H, H-5), 3.94-3.96 (m,1H, H-5''), 3.98 (dd, *J* = 9.6 Hz, *J* = 4.5 Hz, 1H, H-3'), 4.27 (dd, $J_{5',4'}$ = 1.9 Hz, *J* = 3.2 Hz, 1H, H-5'), 4.40 (d, $J_{1',2'}$ = 7.6 Hz, 1H, H-1'), 4.65 (d, $J_{1,2}$ = 7.8 Hz, H-1β), 5.16 (d, $J_{1'',2''}$ = 3.4 Hz, 1H, H-1''), 5.39 (d, *J* = 5.4 Hz, 1H, H-1α); ^{13}C NMR (126 MHz, D_2O): δ ppm 22.1 ($NHCOCH_3$), 55.5 (C-2), 56.4 (C-2β), 58.5 (C-2''), 60.8 (C-6), 61.8 (C-6''), 65.6 (C-3'), 68.7

(C-3), 70.9 (C-4''/C-5''), 74.6 (C-5), 79.1 (C-4), 81.6 (C-3''), 83.4 (C-2'), 85.8 (C-1''), 87.4 (C-5''/C-4''), 90.1 (C-1α), 94.9 (C-1β), 100.1 (C-1'), 68.6 (C-4'), 75.6 (C-5'), 158.1 (NH<u>C</u>OCH$_3$), 164.7 (<u>C</u>O$_2$H); IR (neat) ν_{max} 615, 633, 649, 772, 1058, 1097, 1418, 1590, 3082 cm^{-1}; ESI$^-$ HRMS: calcd. for $C_{20}H_{33}NO_{16}S$ (M-H$^+$) 574.1447, found 574.1452

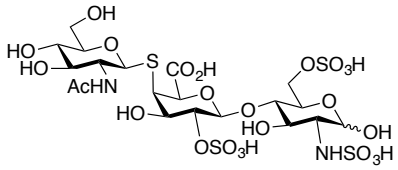
^{1}H spectra



^{13}C spectra

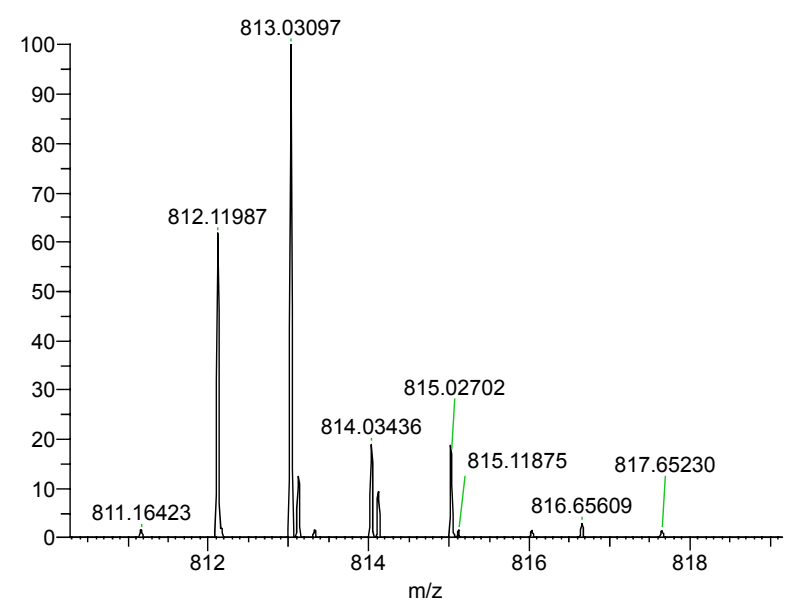


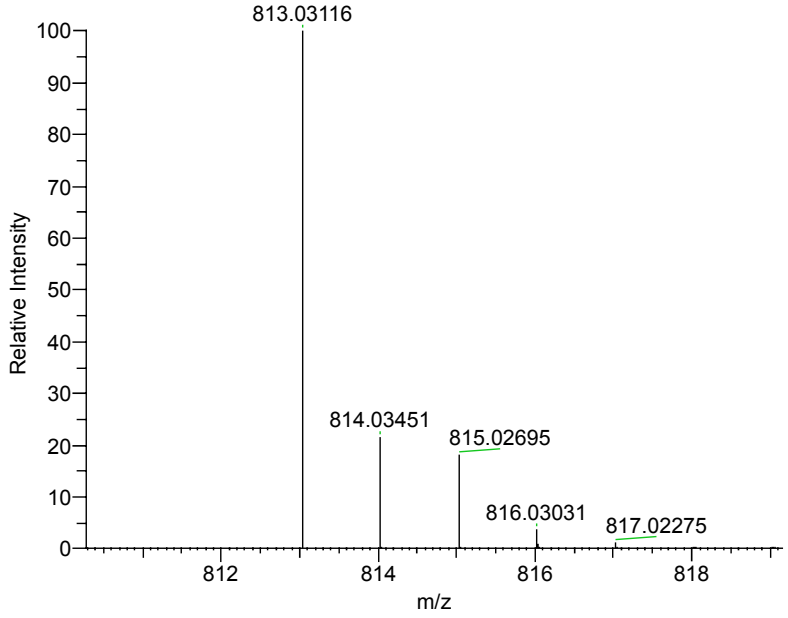
6. Synthesis of compound (10)



Compound **3** (3 mg, 0.005 mmol) was dissolved in 300 µL ammonium formate buffer (250 mM, pH-4). Subsequently, 2-Acetamido-2-deoxy-1-thio-β-D-glucopyranose (20 mg, 0.1 mmol) and Vazo-44 (3 mg, 0.01 mmol) were added to the solution mixture and it was irradiated under UV for 9 hrs. Completion of the reaction was verified using crude NMR where the unsaturation peak from compound **3** coming at around 5.9 ppm was monitored. Thereafter, the reaction mixture was lyophilized to remove ammonium formate.

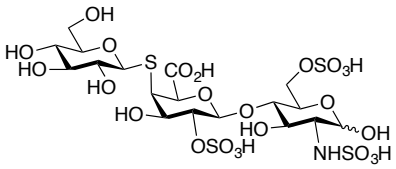
HRMS data:





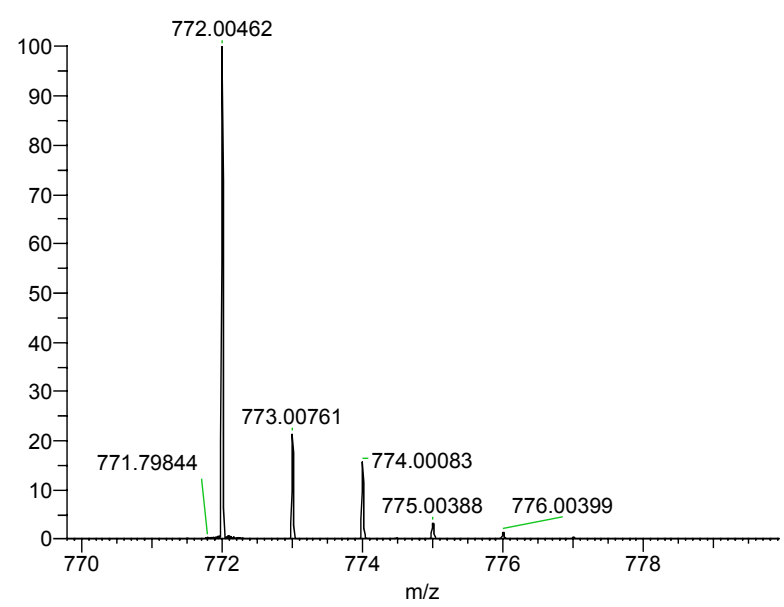
m/z	Formula	RDB	Delta ppm	Theo. Mass
813.03094	$C_{20}H_{33}O_{24}N_2{}^{32}S_4$	5.5	-0.26	813.03116

7. Synthesis of compound (11)

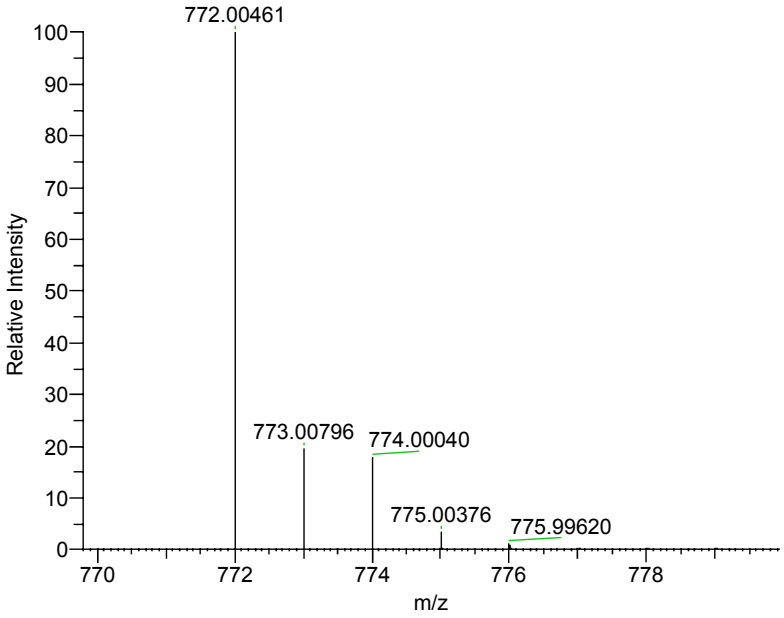


Compound **3** (3 mg, 0.005 mmol) was dissolved in 300 μL ammonium formate buffer (250 mM, pH-4). Subsequently, 1-thio-β-D-glucose (20 mg, 0.1 mmol) and Vazo-44 (3 mg, 0.01 mmol) were added to the solution mixture and it was irradiated under UV for 9 hrs. Completion of the reaction was verified using crude NMR. Thereafter, the reaction mixture was lyophilized to remove ammonium formate.

HRMS data



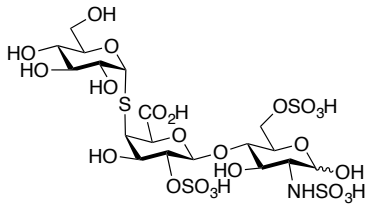
NL: 2.91E6
BGD_AM_07Oct2015_4 #11-16 RT: 0.13-0.17 AV: 3 NL: 7.88E6
T: FTMS {1,2} - p ESI Full ms [80.00-1600.00]



NL: 6.27E5
C18H30O24N1S4: $C_{18}H_{30}O_{24}NS_4$ Chrg -1 R: 1000000 Res. Pwr. @FWHM

m/z	Formula	RDB	Delta ppm	Theo. Mass
772.00464	$C_{18}H_{30}O_{24}N^{32}S_4$	4.5	0.04	772.00461

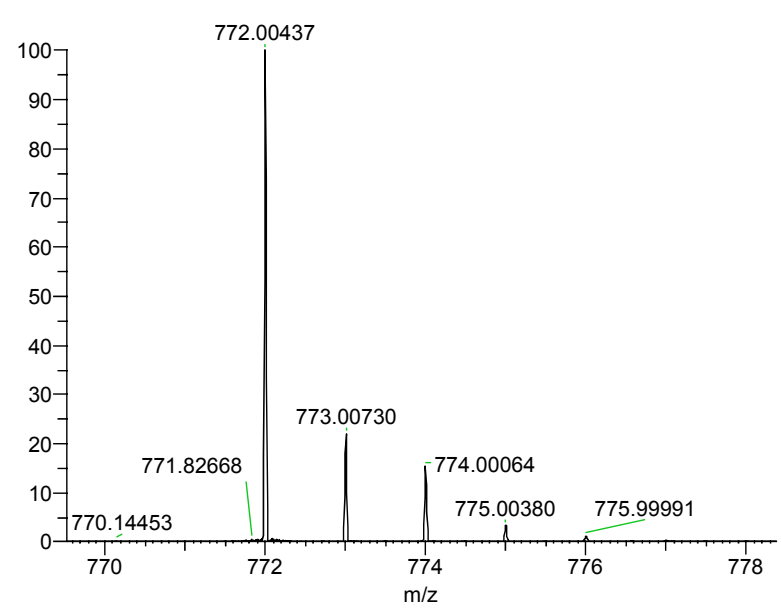
8. Synthesis of compound 12



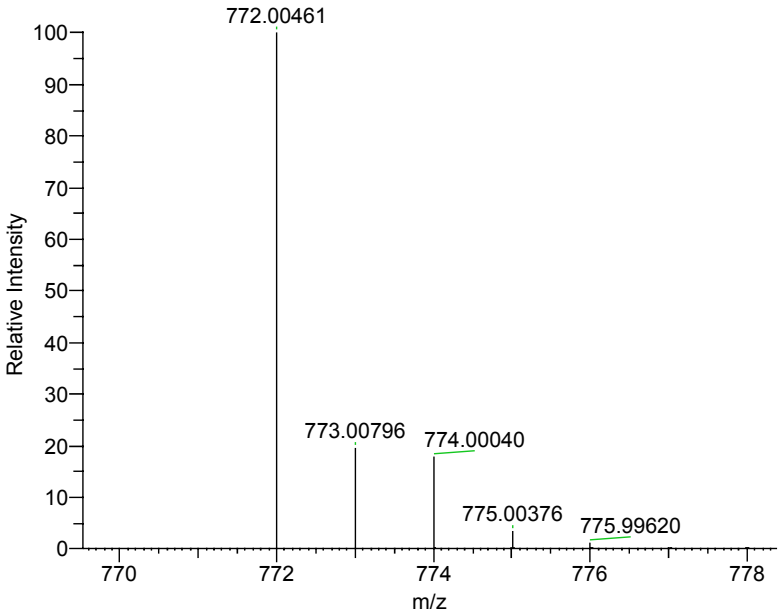
4-Deoxy-α-L-threo-hex-4-enopyranosyl-2-sulfato-(1→4)-6-sulfato-glucose-N-sulfate (5 mg, 0.009 mmol) was dissolved in 300 μL ammonium formate buffer (250 mM, pH-4). Subsequently, 1-thio-α-D-glucose (35 mg, 0.18 mmol) and Vazo-44 (6 mg, 0.018 mmol) were added to the solution mixture and it was irradiated under UV. After 9 hrs, additional 1-thio-α-D-glucose (35 mg, 0.18 mmol) was added into the reaction mixture and the reaction mixture was irradiated for further 4 hrs. The crude reaction mixture was checked with ^{1}H NMR and the crude NMR spectrum showed that the conversion yield was 17%. The reaction mixture was lyophilized thereafter.

HRMS data:

S:\data\Oct 15\BGD_AM_07Oct2015_5.raw 07/10/2015 4:33 pm



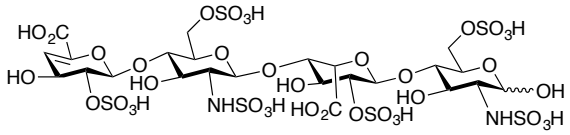
NL: 3.05E6
BGD_AM_07Oct2015_5 #10-26 RT: 0.11-0.29 AV: 9 NL: 8.03E6
T: FTMS {1,2} - p ESI Full ms [80.00-1600.00]



NL: 6.27E5
C18H30O24N1S4: $C_{18}H_{30}O_{24}NS_4$ Chrg -1 R: 1000000 Res. Pwr. @FWHM

m/z	Formula	RDB	Delta ppm	Theo. Mass
772.00439	$C_{18}H_{30}O_{24}N^{32}S_4$	4.5	-0.27	772.00461

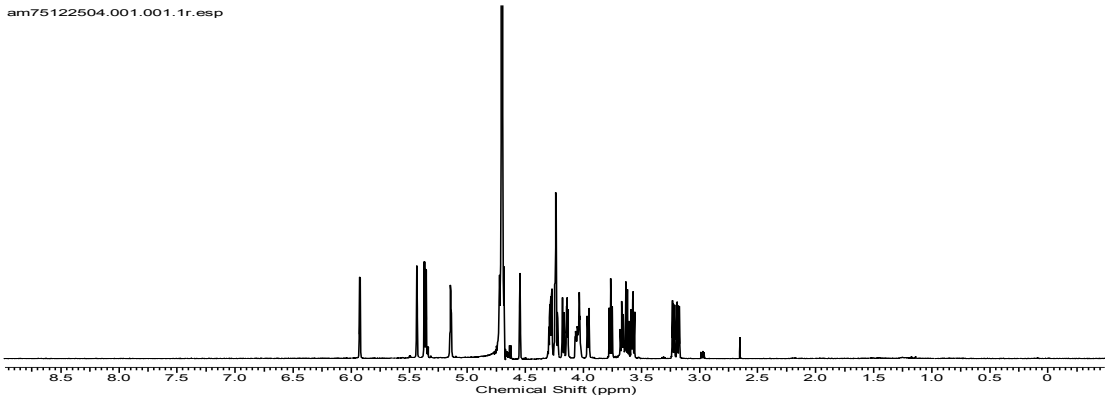
9. Synthesis of compound (13)



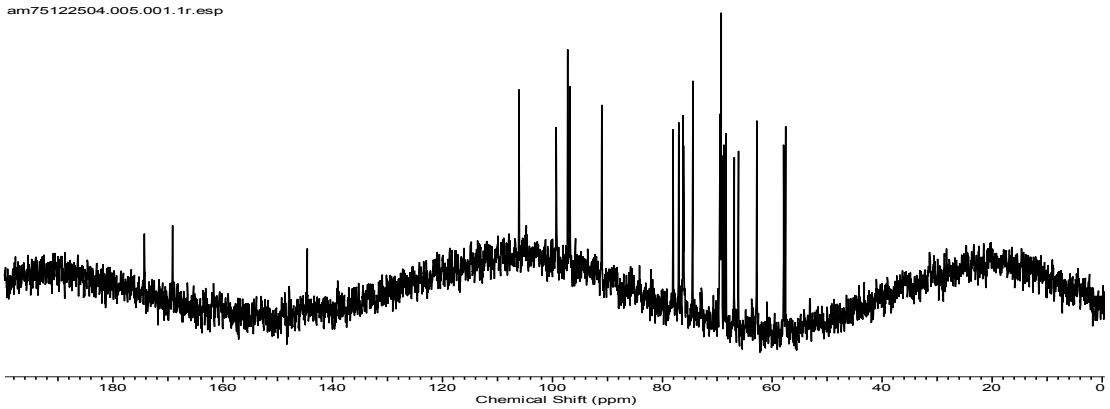
Heparin sodium salt from porcine intestinal mucosa (from Sigma Aldrich) (50 mg) was treated with 5 mg of Heparinase-1, in 5 mL heparin digestion buffer (0.1 M sodium acetate and 0.1 mM calcium acetate, pH 7) at 37 °C. The extent of degradation was determined at regular time intervals by running a sugar gel. The progress of the reaction was monitored by running sugar gels at regular intervals as well as by monitoring the UV absorbance of the reaction mixture at 232 nm. At the end of 72 hrs. the reaction was terminated by heating the mixture at 100 °C for 2 minutes. Thereafter, the protein was removed via vivaspin (10 KDa) and the flow-through was lyophilized and purified on a Superdex 30 size exclusion column (16 mm X 200 cm) using 0.5 M ammonium bicarbonate as buffer, running at 0.5 ml/min. The fraction corresponding to the tetrasaccharide was then lyophilized and purified by means of SAX chromatography on a Propac PA1 column, using 2M NaCl and 100% water as buffer, running at 1 ml/min. The peak corresponding to the disaccharide was collected and lyophilized. Thereafter, desalting was done on a HiPrep 26/10 (Sephadex G-25) desalting column running at 5ml/min flowrate. The peak corresponding to the disaccharide was collected and lyophilized to yield the final pure compound. ^{1}H NMR (700 MHz, D_2O): δ ppm 3.19 (dd, $J_{2'',1''}$ = 3.5 Hz, $J_{2'',3''}$ = 10.3 Hz, 1H, H-2''), 3.23 (dd, $J_{2,1}$ = 3.7 Hz, $J_{2,3}$ = 10.6 Hz, 1H, H-2), 3.57 (dd, $J_{3,2}$ = 10.3 Hz, $J_{3,4}$ = 9 Hz, 1H, H-3), 3.62 (dd, $J_{3'',2''}$ = 10.1 Hz, $J_{3'',4''}$ = 9.2 Hz, 1H, H-3''), 3.67 (dd, $J_{4'',3''}$ = 8.8 Hz, *J4'',5''* = 9.7 Hz, 1H, H-4''), 3.77 (dd, *J4,5* = 9.7 Hz, *J4,3* = 9.2 Hz, 1H, H-4), 3.95-3.97 (m, 1H, H-5), 4.04 (dd, $J_{4',3'}$ = 3.7 Hz, $J_{4',5'}$ = 3.1 Hz, 1H, H-4'), 4.05-4.07 (m, 1H, H-5''), 4.14 (dd, $J_{3',2'}$ = 6.2 Hz,

$J_{3',4'}$ = 3.7 Hz, 1H, H-3'), 4.17 (dd, $J_{6a,6b}$ = 11.2 Hz, *J6a,5* = 1.5 Hz, 1H, H-6a), 4.23 (dd, $J_{6b'',5''}$ = 1.9 Hz, 1H, H-6b''), 4.24 (dd, $J_{3''',2'''}$ = 3.3 Hz, $J_{3''',4'''}$ = 4.8 Hz, 1H, H-3'''), 4.24 (dd, $J_{2',1'}$ = 2.8 Hz, $J_{2',3'}$ = 6.2 Hz, 1H, H-2'), 4.27 (dd, $J_{6a,6b}$ = 11.2 Hz, 1H, H-6b), 4.28 (dd, 1H, H-6a''), 4.55 (dd, $J_{2''',3'''}$ = 3.3 Hz, $J_{2''',1'''}$ = 2.4 Hz, 1H, H-2'''), 4.72 (d, $J_{5',4'}$ = 3.1 Hz, 1H, H-5'), 5.14 (d, $J_{1',2'}$ = 2.8 Hz, 1H, H-1'), 5.36 (d, $J_{1,2}$ = 3.74 Hz, 1H, H-1), 5.37 (d, $J_{1'',2''}$ = 3.52 Hz, 1H, H-1''), 5.44 Hz, (d, $J_{1''',2'''}$ = 2.4 Hz, 1H, H-1'''), 5.93 (d, $J_{4''',3'''}$ = 4.8 Hz, 1H, H-4'''); ^{13}C NMR (176 Hz, D_2O): δ ppm 57.3 (C-2), 57.5 (C-2''), 62.4 (C-3'''), 65.7 (C-6), 66.6 (C-6''), 67.8 (C-5''), 68.5 (C-5), 68.8 (C-3'), 69.2 (C-3), 69.3 (C-3''), 69.5 (C-5'), 74.1 (C-2'''), 75.8 (C-2'),76 (C-4'), 76.5 (C-4''),77.7 (C-4), 90.7 (C-1''), 96.8 (C-1), 97.1 (C-1'''), 99.1 (C-1'), 106.2 (C-4'''); ESI- HRMS: calcd. For $C_{24}H_{36}O_{38}N_2S_6$ $[M-2H]^{2-}$ 575.9641, found 575.9648.

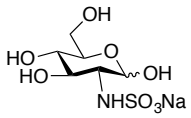
^{1}H spectra



^{13}C spectra



10. Synthesis of compound (1)



Glucosamine hydrochloride (1 g, 4.63 mmol) was dissolved in 30 mL water. Sulfur trioxide pyridine complex (0.88 g, 5.56 mmol) was added in three equal portions of 35 mins. intervals at room temperature. During this whole process, the pH was maintained at 9.5 using sat. $NaHCO_3$. Thereafter, it was left to be stirred for 5 hrs. At the end of the reaction, it was concentrated *in vacuo* and the product was purified using flash chromatography using a solvent system of EtOAc:MeOH:H_2O = 5:2:1. (Yield = 88%) TLC: R_f 0.17 (EtOAc:MeOH:H_2O = 5:2:1) ^{1}H NMR (400 MHz, D_2O): δ 5.34 (d, J=3.54 Hz, 0.7H), 4.60 (d, J= 8.34 Hz, 0.26 H), 3.81-3.61 (m, 3H) , 3.5-3.47 (m, 1H), 3.40-3.34 (m, 1H), 3.12 (dd, J=3.54 Hz, J=6.82 Hz, 0.8 H), 2.90 (dd, J= 8.34 Hz, J=1.52 Hz, 0.2 H), ^{13}C NMR (D_2O) : δ 91.62 (C-1α), 71.68, 71.47, 70.43, 61, 58.35, IR: 1177.06 cm^{-1}, 1350 cm^{-1}

3.4.2 General Chemical methods

Protocol for purification of charged sugars on Phenomenex SAX cartridge

1. Column wash with 2 CV methanol (Conditioning)
2. Column wash with 3 CV water (Equilibrating)
3. Column wash with 300 mM ammonium formate (3 CV)
4. Column wash with 1 M NaCl (2 CV)
5. Column wash with 3 CV water.
6. Crude product loading (10 mgs. in 1 mL water)
7. Washing with 4 CV water (to wash out neutral thio sugar)
8. Washing with 250 mM ammonium formate (3 CV) to wash out charged product.

(CV= column volume)

HPLC time program for purification of compound **3** and compound **13**

0-1 min @ 100% A

1-101 min @ 0% - 75% B (with the gradient increasing from 0 to 75% over 100 mins)

101-111 min @75% - 100%B (with gradient increasing from 75% B to 100% B over 10 mins)

111-121 min @100% B-0% B (with gradient decreasing from 100% B to 0% B over 10 mins)

Protocol for sugar gel preparation

Acrylamide for resolving gel (33%):

297 ml of water is added to 40 % 19:1 acrylamide powder (bought from Sigma) to make 60 %/ 19:1 ratio ie T60 % (total) C5 % (crosslinker).

Acrylamide for loading gel (9%):

29.7 ml of water is added to 40 % 49:1 acrylamide powder to make 60 %/ 49:1 ratio i.e. T60 % (total) C2 % (crosslinker).

GlycoMap buffer:

- 200 ml 895 mM Tris Acetate pH 7.0

  pH is adjusted to 7.0 using 50% acetic acid

  Made up to 1 litre with water to give 895 mM tris base.

  It can be stored at 4 °C until use.

1. Resolving gel: For 1 gel 6 ml solution needed: 3.15 ml GlycoMap Buffer, 3.85 ml Acrylamide 60 % 19:1, 20 µl 10 % APS and 5 µl Temed are mixed.
2. Overlay buffer is composed of 3.15 ml GlycoMap buffer and 3.85 ml water.
3. Loading Gel: For 1 gel: 0.9 ml GlycoMap buffer, 0.3 ml acrylamide, 0.8 ml water, 10 µl 10 % APS, 5 µl TEMED are needed.
4. Running buffer is made with 36.3g Tris, 40g MES and in 1 litre of water. This corresponds to 10X concentration.

5. Loading buffer is 1:1 running buffer: glycerol if dry sample or 80 % glycerol/ 20 % running buffer if diluting into sample. 1 µl phenol red is dissolved in 50 % glycerol/ water to each sample. Loading buffer (x2 strength) is composed of 895 mM tris pH 7.0, 550 µl + 80% glycerol 250 µl + 200 µl H_2O.
6. 200 V run takes about 4 hours, it can then be run overnight at 100 V so that phenol red has run approx 2/3 down the gel.
7. Destaining at the end of the run is done with 0.08 % azure A (dissolved in water) for 10 minutes minimum and then destained using water.

3.4.3 Biological Experimental

Heparinase-1 gene sequence (UniProt ID: B3U3D9)

CAGCAAAAAAAATCCGGTAACATCCCTTACCGGGTAAATGTGCAGGCCGACAGTGC
TACAGAGCGAGATTATTGACAACAAATGGGTGGCAGTAGGCATCAATAAACCTTAT
GCATTACAATATGACGATAAACTGCG CTTTAATGGA AAACCATCCT ATCGCTTTGA
GCTTAAAGCCGAAGACAATTCGCTTGAAGGTTATGCTGCAGGAGAAACAAAGGGCC
GTATAGAATTGTCGTACAGCTATGCAACCACCAATGATTTTAAGAAATTTCCCCCAA
GCGTATACCAAAATGCGCAAAAGCTAAAAACCGTTTATCATTACGGCAAAGGGATT
TGTGAACAGGGGAGCTCCCGCAGCTATACCTTTTCAGTGTACATACCCTCCTCCTTC
CCCGACAATGCGACTACTATTTTTGCCCAATGGCATGGTGCACCCAGCAGAACGCTT
GTAGCTACACCAGAGGGAGAAATTAAAACACTGAGCATAGAAGAGTTTTTGGCCTT
ATACGACCGCATGATCTTCAAAAAAAATATCGCCCATGATAAAGTTGAAAAAAAAG
ATAAGGACGGAAAAATTACTTATGTAGCCGGAAAGCCAAATGGCTGGAAGGTAGAA
CAAGGTGGTTATCCACCGCTGGCCTTTGGTTTTTCTAAAGGGTATTTTTACATCAAGG
CAAACTCCGACCGGCAGTGGCTTACCGACAAAGCCGACCGTAACAATGCCAATCCC
GAGAATAGTGAAGTAATGAAGCCCTATTCCTCGGAATACAAAACTTCTACCATTGCC
TATAAAATGCCCTTTGCCCAGTTCCCTAAAGATTGCTGGATTACTTTTGATGTCGCCA
TAGACTGGACGAAATATGGAAAAGAGGCCAATACAATTTTGAAACCCGGTAAGCTG
GATGTGATGATGACTTATACCAAGAATAAGAAACCACAAAAAGCGCATATCGTAAA
CCAGCAGGAAATCCTGATCGGACGTAACGATGACGATGGCTATTACTTCAAATTTGG

AATTTACAGGGTCGGTAACAGCACGGTCCCGGTTACTTATAACCTGAGCGGGTACAG
CGAAACTGCCAGA

SUMO protein gene sequence:

AGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGT
CAAGCCTGAGACTCACATCAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTT
CAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGAC
AGGGTAAGGAAATGGACTCCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTG
ATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTATTGAGGCTCACAGA
GAACAGATTGGTGGTATC

Hep-1 gene sequence with linker region (as procured from Genscript)

CATATGGGTGGTGGCGGCAGTGGCGGCGGTGGTTCAGGCGGTGGCGGTTCAGACGA
CGATGACAAACAACAGAAAAAATCAGGCAACATTCCGTATCGTGTGAACGTTCAGG
CAGACAGTGCTAAACAAAAAGCTATCATCGATAACAAATGGGTCGCGGTGGGCATT
AATAAACCGTATGCCCTGCAGTACGATGACAAACTGCGTTTTAACGGTAAACCGAG
CTATCGCTTCGAACTGAAAGCGGAAGACAATTCTCTGGAAGGCTACGCGGCCGGTG
AAACCAAAGGCCGTACGGAACTGAGCTATTCTTACGCAACCACGAACGATTTTAAA
AAATTCCCGCCGAGCGTGTATCAGAATGCCCAAAAACTGAAAACGGTTTATCATTAC
GGCAAAGGTATTTGCGAACAGGGCAGCTCTCGCAGTTATACCTTTTCCGTTTACATT
CCGAGTTCCTTCCCGGATAACGCGACCACGATCTTTGCCCAATGGCATGGTGCACCG
TCTCGTACCCTGGTCGCGACGCCGGAAGGTGAAATTAAAACCCTGAGCATCGAAGA
ATTCCTGGCGCTGTACGACCGCATGATCTTCAAGAAAAACATCGCCCATGATAAAGT
GGAAAAGAAAGATAAAGATGGTAAAATCACGTATGTTGCCGGTAAACCGAATGGCT
GGAAAGTCGAACAGGGCGGTTACCCGACCCTGGCATTTGGTTTCTCAAAAGGCTATT
TCTACATCAAAGCTAATTCGGATCGTCAATGGCTGACCGACAAAGCTGATCGCAACA
ATGCGAACCCGGAAAATAGTGAAGTGATGAAACCGTATTCATCGGAATACAAAACC
AGCACGATTGCCTATAAAATGCCGTTTGCACAGTTCCCGAAAGATTGTTGGATTACC
TTTGACGTTGCAATCGATTGGACGAAATATGGCAAAGAAGCTAACACCATCCTGAA
ACCGGGTAAACTGGATGTCATGATGACCTACACGAAAAATAAAAAACCGCAGAAAG
CGCACATTGTGAACCAGCAAGAAATTCTGATCGGTCGTAACGATGACGATGGCTACT
ACTTCAAATTCGGTATCTATCGCGTGGGTAACAGCACGGTTCCGGTCACCTACAATC
TGAGCGGCTACTCCGAAACGGCACGCTAACTCGAG

Primers for Hep-1-SUMO fusion protein gene synthesis

SUMO_F: AGTCGTCATATGGGCAGCAGCCATCATCATCATCATCA

SUMO_R: CCGCCACTGCCGCCACCACCACCACCAATCTGTTCTCTGTGAGCCTC

HepI_F: GGTGGTGGCGGCAGTGGCGG

HepI_R: AGGCGACTCGAGTTAGCGTGCCGTTTCGGAGTAG

Fusion gene synthesis

Hep-1 gene sequence procured from Genscript and SUMO are subjected to PCR amplification. Pfu ULTRA II Hotstart PCR mastermix used.

PCR cycle used:

Denaturation = 95 °C for 2 mins. 20 secs

Hybridization = 55 °C for 20 secs.

Elongation = 72 °C for 3 mins. 45 secs.

It was repeated for 25 cycles.

Thereafter, a 0.8% agarose DNA gel was run for 35 mins. at 110 V with the PCR products. The purified products were then extracted using the QIAquick Gel Extraction kit. For overlap extension PCR, 3.5 µl Hep-1 at a concentration of 0.03 pmol/µl was added with 1.25 µl of SUMO at a concentration of 0.08 pmol/µl., alongwith Pfu Mastermix of 25 µl and 20.25 µl water. The total 50 µl reaction mixture was subjected to PCR with the following conditions:

Denaturation = 94 °C for 5 mins. 30 sec.

Hybridization = 60 °C for 30 secs.

Elongation = 72 °C for 3 mins. 45 secs.

The cycle was repeated 30 times.

Thereafter, it was run with SUMO_F (1 µl at 10 µM concentration) and Hep-1_R (1 µl at 10 µM concentration) primers for 20 more cycles.

The ligated mass was purified by running a DNA gel and extraction with QIAquick Gel Extraction kit.

The restriction enzyme digest was subsequently performed as follows:

To 20 µl pET 22b plasmid at 116 ng/ µl concentration, 2 µl each of *Nde*1 and *Xho*1 were added, alongwith 2.7 µl CutSmart buffer. They were then incubated at 37 °C for 3 h.

To a 50 µl fused gene at 15.3 ng/ µl, 2 µl each of Nde-1 and Xho-1 were added, alongwith 6 µl CutSmart buffer. They were then incubated at 37 °C for 3 h.

Purification of the enzymatic mixture was done by running a DNA gel. Quick ligation of the fused gene was performed in the now empty pET-22 vector.

5.32 µl of pET-22 vector at 9.4 ng/µl was added to 3.5 µl of fused gene at 15.3 ng/µl concentration, alongwith 1.18 µl water, 10 µl of 2X ligation buffer and 1 µl of Quick T4 DNA ligase. They were thoroughly mixed by flicking and incubated at room temperature for 5 mins. They were then chilled on ice and this gene mixture was then transformed into XL10-Gold Cells.

General protocol for transformation of plasmid into XL-10 Gold cells

1. 10 µl of cells are taken out from -80°C freezer and thawed on ice.

2. 2 µl of β-mercaptoethanol is added to the cells in a tube.

3. The tube is flicked to mix the contents and then incubated on ice for 10 mins.

4. 4 µl of ligation mixture is added to the culture.

5. The tube is flicked again and incubated on ice for further 30 mins.

6. The tube is then heat pulsed in 42 °C water bath for 30 secs. And immediately put in ice and incubated for 10 mins.

7. 300 µl of SOC media is added to the tube and thereafter it is placed in 37 °C incubator for 1 h (at 225-250 r.p.m.)

8. Ampicillin-LB-Agar plates is used for plating the culture.

9. Three small scale cultures are started with three colonies of cells from the plates, the next day, with 10 ml LB media each, alongwith 10 µl of (1000X stock) Ampicillin.

10. The DNA is extracted by miniprep kit the next day and is sent for sequencing.

Thereafter, they are transformed into BL21(DE3) cells for corresponding protein expression.

General protocol for transformation of plasmid into BL21(DE3) cells

1. 10 µl of cells are taken out from the freezer and thawed on ice.

2. 4 µl of plasmid DNA then added to the cell culture. Tube is flicked a few times to mix the contents.

3. Mixture is kept on ice for 30 mins.

4. Heat shock provided for 30 secs at 42 °C

5. Tube is incubated on ice for 10 mins.

6. 200 µl SOC media is added to it.

7. Thereafter it is placed in a 37 °C incubator for 1 h (at 225-250 r.p.m.)

8. The cell cultures are then plated on LB-Agar-Amp plates.

9. Next day, four small case cultures are started from individual colonies.

Small scale expression

1. Four small scale cultures of 10 ml LB each are started.

2. 10 µl of Amp (from 1000X stock) added to the tubes.

3. They are left at 37 °C for 3 h until O.D. reaches 0.6

4. Thereafter, IPTG (1mM) is added.

5. Incubated at 37 °C for another 3 h.

6. They are then centrifuged at 3750 rpm for 10 mins.

7. Supernatant is drained while cell pellet is suspended in 0.5 ml BugBuster mix.

8. They are then shaken at r.t. for 15 mins.

9. Centrifugation then done at 20K for 10 mins to separate out the supernatant and pellet.

10. They are then run on SDS gel.

General Protocol for Large scale expression

1. Two 250 ml flasks with 25 grams each of LB and 1 litre water were autoclaved.

2. After the small scale culture was complete, 10 ml of the culture was added to 1 litre of the large scale culture. 1 ml of (1000X stock) Amp. is also added.

3. They were shaken at 180 rpm for 3 h until O.D. reaches 0.6.

4. 1 mM IPTG (from 1M stock Amp) was added to each of the 2 litre cultures.

5. They were further shaken at 37 °C for further 3 h.

6. Thereafter, the media was centrifuged for 10 mins at 8K and the supernatant poured out.

7. A minimum amount of lysis buffer (25 mM Tris, pH – 7.2, 500 mM NaCl, 30 mM imidazole) was then added to the cell pellet, along with 25 mg of lysozyme, 5 mg of DNAse and 1 protease inhibitor tablet. The cell pellet was resuspended in this reaction mixture for 3 h.

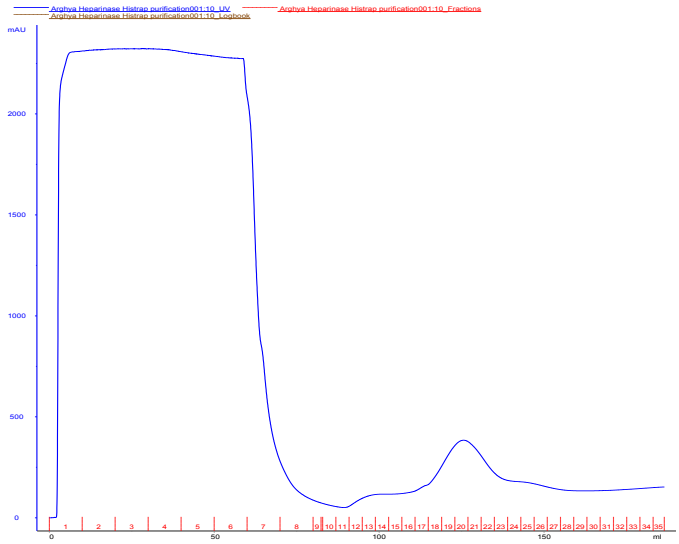
8. Cells were then lysed by sonication. A sonication pulse cycle of 0.5 sec off and 0.5 sec on time was run for 5 mins.

9. The lysed cell mixture was then centrifuged at 20K for 30 mins. to separate out the supernatant from the inclusion bodies.

10. The supernatant was loaded onto a Histrap column for purification using binding buffer (25 mM Tris, pH – 7.2, 500 mM NaCl, 30 mM imidazole) and elution buffer (25 mM Tris, pH – 7.2, 500 mM NaCl, 250 mM imidazole) running at 5 ml/min.

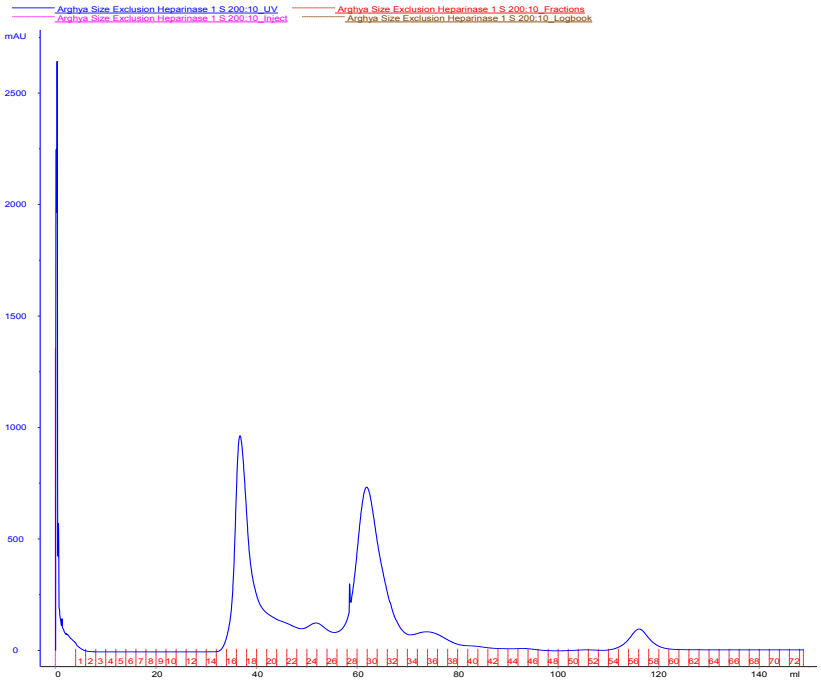
11. The fractions were then dialysed against the storage buffer (25 mM Tris, pH – 7.2, 150 mM NaCl) and stored with 10% glycerol.

HisTrap purification chromatogram of Hep-1.



It was further purified by size exclusion chromatography on a S200 column.

Size Exclusion purification chromatogram of Hep-1

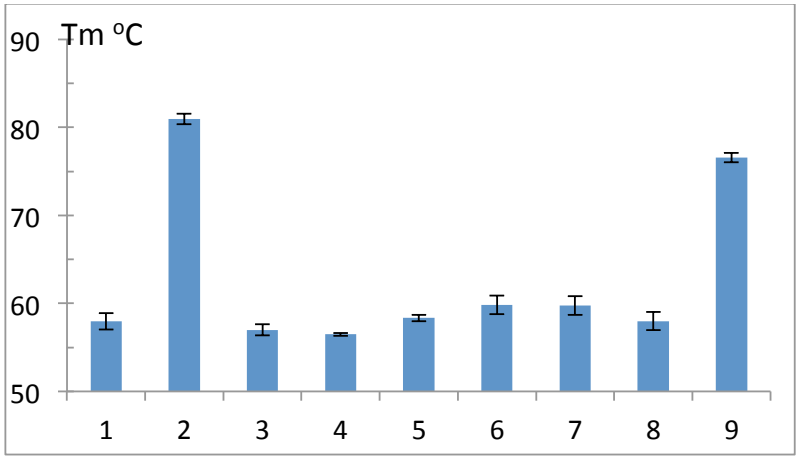


2nd peak was found to be of the product while the first peak remained unidentified.

DSF studies with BACE-1, FGF-1 and FGF-2

A BioRad RT-PCR machine was used for the DSF studies, in 96 well plates. Experiments were done in triplicates. Typically, 10 µl reaction volumes were setup. Heparin disaccharide, heparin and synthesized compounds were added in 5:1 molar ratio w.r.t the protein (with the protein concentration as 10 µM). 1 µl of Sypro Orange was added from a 100X stock concentration to make the final concentration of the dye as 10X. (This was found to be optimal for giving the best reading). The rest of the solution was made up with PBS buffer upto 10 µl. The plate was sealed with transparent foil and the thermocycler was set up for an experiment with the following

parameters (0.3 °C/min rise in temperature with a 3 sec hold time, excitation at 483 nm and emission at 568 nm, scanning from 31 °C to 81 °C). The data was subjected to a Excel-based processing using Frank Niesens (SGC Oxford) analysis tool[58].



A typical data shown from a run with 10:1 concentration of sugar:protein taken (1) FGF-1 control (2) heparin:FGF-1 (10:1), (3) compound **2:** FGF-1 (10:1) (4) compound **3:** FGF-1 **(**10:1**)** (5), compound **10**: FGF-1 (5:1) (6) compound **11**: FGF-1 (10:1), (7) compound **12**: FGF-1 (10:1), (8) compound **13:** FGF-1 (10:1) (9) heparin dp16:FGF-1 (10:1)

The results show that even heparin tetrasaccharide (compound **13**) does not significant binding with FGF-1 as measured by DSF. However, heparin derived dp16 has got a quite significant binding.

3.5 References

1. Ernst, S.; Langer, R.; Cooney, C. L.; Sasisekharan, R. Enzymatic Degradation of Glycosaminoglycans. *Crit. Rev. Biochem. Mol. Biol.* **1995,** *30* (5), 387-444.

2. Shriver, Z.; Sundaram, M.; Venkataraman, G.; Fareed, J.; Linhardt, R.; Biemann, K.; Sasisekharan, R. Cleavage of the antithrombin III binding site in heparin by heparinases and its implication in the generation of low molecular weight heparin. *Proc. Natl. Acad. Sci.USA* **2000,** *97* (19), 10365-10370.

3. Bernardes, G. J. L.; Gamblin, D. P.; Davis, B. G. The Direct Formation of Glycosyl Thiols from Reducing Sugars Allows One-Pot Protein Glycoconjugation. *Angew. Chem. Int. Ed.* **2006,** *45* (24), 4007-4011.

4. Lecher, H. Z.; Greenwood, R. A.; Whitehouse, K. C.; Chao, T. H. The Phosphonation of Aromatic Compounds with Phosphorus Pentasulfide. *J. Am. Chem. Soc.* **1956,** *78* (19), 5018-5022.

5. Pedersen, B. S.; Scheibye, S.; Nilsson, N. H.; Lawesson, S. O. Studies on Organophosphorus Compounds .10. Syntheses of Thioketones. *Bull. Soc. Chim. Belg.* **1978,** *87* (3), 223-228.

6. Jones, B. A.; Bradshaw, J. S. Synthesis and Reduction of Thiocarboxylic O-Esters. *Chem. Rev.* **1984,** *84* (1), 17-30.

7. Scheibye, S.; Kristensen, J.; Lawesson, S. O. Studies on Organophosphorus Compounds .27. Synthesis of Thionolactones, Thiololactones and Dithiolactones. *Tetrahedron* **1979,** *35* (11), 1339-1343.

8. Shabana, R.; Scheibye, S.; Clausen, K.; Olesen, S. O.; Lawesson, S. O. Studies on Organo-Phosphorus Compounds .31. Synthesis of Thiolactams and Thioimides. *Nouv. J. Chim.* **1980,** *4* (1), 47-51.

9. Nishio, T. A novel transformation of alcohols to thiols. *J. Chem. Soc., Chem. Commun.* **1989,** (4), 205-206.

10. Ori, M.; Nishio, T. Sulfur-containing heterocycles: Facile synthesis of 4H-1,3-thiazines by the reaction of 3-N-acylamino ketones with Lawesson's reagent. *Heterocycles* **2000,** *52* (1), 111-116.

11. Sanz-Cervera, J. F.; Blasco, R.; Piera, J.; Cynamon, M.; Ibanez, I.; Murguia, M.; Fustero, S. Solution versus Fluorous versus Solid-Phase Synthesis of 2,5-Disubstituted 1,3-Azoles. Preliminary Antibacterial Activity Studies. *J. Org. Chem.* **2009,** *74* (23), 8988-8996.

12. Lau, K.; Thon, V.; Yu, H.; Ding, L.; Chen, Y.; Muthana, M. M.; Wong, D.; Huang, R.; Chen, X. Highly efficient chemoenzymatic synthesis of β1-4-linked galactosides with promiscuous bacterial β1-4-galactosyltransferases. *Chem. Commun.* **2010,** *46* (33), 6066-6068.

13. Fritz, J. S.; Schenk, G. H. Acid-Catalyzed Acetylation of Organic Hydroxyl Groups. *Anal. Chem.* **1959,** *31* (11), 1808-1812.

14. Bradshaw, R. *Heparin: Structure, Function and Clinical Implications*. Springer Science & Business Media: **2013**.

15. Desai, U. R.; Wang, H. M.; Linhardt, R. J. Substrate-Specificity of the Heparin Lyases from Flavobacterium-Heparinum. *Arch. Biochem. Biophys.* **1993,** *306* (2), 461-468.

16. Desai, U. R.; Wang, H. M.; Linhardt, R. J. Specificity Studies on the Heparin Lyases from Flavobacterium-Heparinum. *Biochemistry* **1993,** *32* (32), 8140-8145.

17. Wei, Z.; Lyon, M.; Gallagher, J. T. Distinct substrate specificities of bacterial heparinases against N-unsubstituted glucosamine residues in heparan sulfate. *J. Biol. Chem.* **2005,** *280* (16), 15742-15748.

18. Hileman, R. E.; Fromm, J. R.; Weiler, J. M.; Linhardt, R. J. Glycosaminoglycan-protein interactions: definition of consensus sites in glycosaminoglycan binding proteins. *Bioessays* **1998,** *20* (2), 156-167.

19. Rabenstein, D. L. Heparin and heparan sulfate: structure and function. *Nat. Prod. Rep.* **2002,** *19* (3), 312-331.

20. Ernst, S.; Venkataraman, G.; Winkler, S.; Godavarti, R.; Langer, R.; Cooney, C. L.; Sasisekharan, R. Expression in Escherichia coli, purification and characterization of heparinase I from Flavobacterium heparinum. *Biochem. J.* **1996,** *315*, 589-597.

21. Sasisekharan, R.; Bulmer, M.; Moremen, K. W.; Cooney, C. L.; Langer, R. Cloning and Expression of Heparinase-I Gene from Flavobacterium-Heparinum. *Proc. Natl. Acad. Sci. USA* **1993,** *90* (8), 3660-3664.

22. Costa, S.; Almeida, A.; Castro, A.; Domingues, L. Fusion tags for protein solubility, purification, and immunogenicity in Escherichia coli: the novel Fh8 system. *Front. Microbiol.* **2014,** *5*.

23. Huang, J.; Cao, L.; Guo, W.; Yuan, R.; Jia, Z.; Huang, K. Enhanced soluble expression of recombinant Flavobacterium heparinum heparinase I in Escherichia coli by fusing it with various soluble partners. *Protein Expres. Purif.* **2012,** *83* (2), 169-176.

24. Roy, L. M.; Jaruga, P.; Wood, T. G.; McCullough, A. K.; Dizdaroglu, M.; Lloyd, R. S. Human polymorphic variants of the NEIL1 DNA glycosylase. *J. Biol. Chem.* **2007,** *282* (21), 15790-15798.

25. Shuldiner, A. R.; Scott, L. A.; Roth, J. PCR-Induced (Ligase-Free) Subcloning - a Rapid Reliable Method to Subclone Polymerase Chain-Reaction (PCR) Products. *Nucleic Acids Res.* **1990,** *18* (7), 1920-1920.

26. Shuldiner, A. R.; Tanner, K.; Scott, L. A.; Moore, C. A.; Roth, J. Ligase-Free Subcloning - a Versatile Method to Subclone Polymerase Chain-Reaction (PCR) Products in a Single Day. *Anal. Biochem.* **1991,** *194* (1), 9-15.

27. Bryksin, A. V.; Matsumura, I. Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. *Biotechniques* **2010,** *48* (6), 463-465.

28. Watson, R. J.; Schildkraut, I.; Qiang, B. Q.; Martin, S. M.; Visentin, L. P. NdeI - a Restriction Endonuclease from Neisseria denitrificans Which Cleaves DNA at 5'-CATATG-3' Sequences. *FEBS Lett.* **1982,** *150* (1), 114-116.

29. Theriault, G.; Roy, P. H.; Howard, K. A.; Benner, J. S.; Brooks, J. E.; Waters, A. F.; Gingeras, T. R. Nucleotide-Sequence of the PaeR7 Restriction Modification System and Partial Characterization of Its Protein Products. *Nucleic Acids Res.* **1985,** *13* (23), 8441-8461.

30. Powell, A. K.; Ahmed, Y. A.; Yates, E. A.; Turnbull, J. E. Generating heparan sulfate saccharide libraries for glycomics applications. *Nat. Protoc.* **2010,** *5* (5), 821-33.

31. Thanawiroon, C.; Linhardt, R. J. Separation of a complex mixture of heparin-derived oligosaccharides using reversed-phase high-performance liquid chromatography. *J. Chromatogr A.* **2003,** *1014* (1), 215-223.

32. Cramer, H. M., & Cone, R. Improved Analysis of Simple Sugars Using apHera™ NH2 HPLC Columns. *Reporter US*, Sigma-Aldrich Corporation, 25 (2), http://www.sigmaaldrich.com/technical-documents/articles/reporter-us/improved-analysis.html (accessed 10 August 2015).

33. Blumberg, K.; Liniere, F.; Pustilnik, L.; Bush, C. A. Fractionation of Oligosaccharides Containing N-Acetyl Amino-Sugars by Reverse-Phase High-Pressure Liquid-Chromatography. *Anal. Biochem.* **1982,** *119* (2), 407-412.

34. Muraki, E.; Yaku, F.; Iyoda, J.; Kojima, H. Measurement of Degree of Deacetylation in D-Glucosamine Oligosaccharides by UV Absorption. *Biosci. Biotechnol. Biochem.* **1993,** *57* (11), 1929-1930.

35. http://www.phenomenex.com/Products/SPDetail/Strata/SAX.

36. Jaseja, M.; Rej, R. N.; Sauriol, F.; Perlin, A. S. Novel Regioselective and Stereoselective Modifications of Heparin in Alkaline-Solution - Nuclear Magnetic-Resonance Spectroscopic Evidence. *Can. J. Chem.* **1989,** *67* (9), 1449-1456.

37. Wolfrom, M. L.; Gibbons, R. A.; Huggard, A. J. Derivatives of D-Glucose Containing the Sulfoamino Group. *J. Am. Chem. Soc.* **1957,** *79* (18), 5043-5046.

38. Wolfrom, M. L.; Thomas, G. H. S.; Vercellotti, J. R. Carboxyl-Reduced Heparin - Monosaccharide Components. *J. Org. Chem.* **1964,** *29* (3), 536-&.

39. Ayotte, L.; Perlin, A. S. NMR Spectroscopic Observations Related to the Function of Sulfate Groups in Heparin - Calcium-Binding Vs Biological-Activity. *Carbohydr. Res.* **1986,** *145* (2), 267-277.

40. Inoue, Y.; Nagasawa, K. Selective N-Desulfation of Heparin with Dimethyl-Sulfoxide Containing Water or Methanol. *Carbohydr. Res.* **1976,** *46* (1), 87-95.

41. Huang, M. L.; Smith, R. A. A.; Trieger, G. W.; Godula, K. Glycocalyx Remodeling with Proteoglycan Mimetics Promotes Neural Specification in Embryonic Stem Cells. *J. Am. Chem. Soc.* **2014,** *136* (30), 10565-10568.

42. Ishihara, M.; Kariya, Y.; Kikuchi, H.; Minamisawa, T.; Yoshida, K. Importance of 2-O-sulfate groups of uronate residues in heparin for activation of FGF-1 and FGF-2. *J. Biochem.* **1997,** *121* (2), 345-349.

43. Kariya, Y.; Kyogashima, M.; Suzuki, K.; Isomura, T.; Sakamoto, T.; Horie, K.; Ishihara, M.; Takano, R.; Kamei, K.; Hara, S. Preparation of completely 6-O-desulfated heparin and its ability to enhance activity of basic fibroblast growth factor. *J. Biol. Chem.* **2000,** *275* (34), 25949-25958.

44. Xu, D.; Esko, J. D. Demystifying Heparan Sulfate-Protein Interactions. *Annu. Rev. Biochem.* **2014,** *83*, 129-157.

45. Capila, I.; Linhardt, R. J. Heparin-Protein Interactions. *Angew. Chem. Int. Ed.* **2002,** *41* (3), 390-412.

46. Brown, A.; Robinson, C. J.; Gallagher, J. T.; Blundell, T. L. Cooperative Heparin-Mediated Oligomerization of Fibroblast Growth Factor-1 (FGF1) Precedes Recruitment of FGFR2 to Ternary Complexes. *Biophys. J.* **2013,** *104* (8), 1720-1730.

47. Huntington, J. A. Thrombin inhibition by the serpins. *J. Thromb. Haemost.* **2013,** *11*, 254-264.

48. Kufareva, I.; Salanga, C. L.; Handel, T. M. Chemokine and chemokine receptor structure and interactions: implications for therapeutic strategies. *Immunol. Cell Biol.* **2015,** *93* (4), 372-383.

49. Cui, H.; Hung, A. C.; Freeman, C.; Narkowicz, C.; Jacobson, G. A.; Small, D. H. Size and sulfation are critical for the effect of heparin on APP processing and Aβ production. *J. Neurochem.* **2012,** *123* (3), 447-457.

50. Patey, S. J.; Edwards, E. A.; Yates, E. A.; Turnbull, J. E. Heparin Derivatives as Inhibitors of BACE-1, the Alzheimer's β-Secretase, with Reduced Activity against Factor Xa and Other Proteases. *J. Med. Chem.* **2006,** *49* (20), 6129-6132.

51. Arungundram, S.; Al-Mafraji, K.; Asong, J.; Leach, F. E., 3rd; Amster, I. J.; Venot, A.; Turnbull, J. E.; Boons, G. J. Modular synthesis of heparan sulfate oligosaccharides for structure-activity relationship studies. *J. Am. Chem. Soc.* **2009,** *131* (47), 17394-405.

52. Jiang, W. X.; Wang, Z. H.; Beier, R. C.; Jiang, H. Y.; Wu, Y. N.; Shen, J. Z. Simultaneous Determination of 13 Fluoroquinolone and 22 Sulfonamide Residues in Milk by a

Dual-Colorimetric Enzyme-Linked Immunosorbent Assay. *Anal. Chem.* **2013,** *85* (4), 1995-1999.

53. Bellapadrona, G.; Tesler, A. B.; Grunstein, D.; Hossain, L. H.; Kikkeri, R.; Seeberger, P. H.; Vaskevich, A.; Rubinstein, I. Optimization of Localized Surface Plasmon Resonance Transducers for Studying Carbohydrate-Protein Interactions. *Anal. Chem.* **2012,** *84* (1), 232-240.

54. Wang, X.; Matei, E.; Gronenborn, A. M.; Ramstrom, O.; Yan, M. D. Direct Measurement of Glyconanoparticles and Lectin Interactions by Isothermal Titration Calorimetry. *Anal. Chem.* **2012,** *84* (10), 4248-4252.

55. Timmer, C. M.; Michmerhuizen, N. L.; Witte, A. B.; Van Winkle, M.; Zhou, D. J.; Sinniah, K, An Isothermal Titration and Differential Scanning Calorimetry Study of the G-Quadruplex DNA-Insulin Interaction. *J. Phys. Chem. B* **2014,** *118* (7), 1784-1790.

56. Siligardi, G.; Hussain, R.; Patching, S. G.; Phillips-Jones, M. K. Ligand- and drug-bindin studies of membrane proteins revealed through circular dichroism spectroscopy. *BBA-Biomembranes* **2014,** *1838* (1), 34-42.

57. Senisterra, G.; Chau, I.; Vedadi, M. Thermal Denaturation Assays in Chemical Biology. *Assay Drug Dev. Techn.* **2012,** *10* (2), 128-136.

58. Niesen, F. H.; Berglund, H.; Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2007,** *2* (9), 2212-2221.

59. Patey, S. J.; Edwards, E. A.; Yates, E. A.; Turnbull, J. E. Heparin derivatives as inhibitors of BACE-1, the Alzheimer's β-secretase, with reduced activity against factor Xa and other proteases. *J. Med. Chem.* **2006,** *49* (20), 6129-6132.

60. Zulueta, M. M. L.; Lin, S. Y.; Hu, Y. P.; Hung, S. C. Synthetic heparin and heparan sulfate oligosaccharides and their protein interactions. *Curr. Opin. Chem. Biol.* **2013,** *17* (6), 1023-1029.

61. Hu, Y.-P.; Zhong, Y.-Q.; Chen, Z.-G.; Chen, C.-Y.; Shi, Z.; Zulueta, M. M. L.; Ku, C.-C.; Lee, P.-Y.; Wang, C.-C.; Hung, S.-C. Divergent Synthesis of 48 Heparan Sulfate-Based Disaccharides and Probing the Specific Sugar-Fibroblast Growth Factor-1 Interaction. *J. Am. Chem. Soc.* **2012,** *134* (51), 20722-20727.

62. Xu, R. Y.; Ori, A.; Rudd, T. R.; Uniewicz, K. A.; Ahmed, Y. A.; Guimond, S. E.; Skidmore, M. A.; Siligardi, G.; Yates, E. A.; Fernig, D. G. Diversification of the Structural Determinants of Fibroblast Growth Factor-Heparin Interactions Implications for binding specificity. *J. Biol. Chem.* **2012,** *287* (47), 40061-73.

63. Robinson, C. J.; Harmer, N. J.; Goodger, S. J.; Blundell, T. L.; Gallagher, J. T. Cooperative dimerization of fibroblast growth factor 1 (FGF1) upon a single heparin saccharide may drive the formation of 2:2:1 FGF1•FGFR2c•Heparin ternary complexes. *J. Biol. Chem.* **2005,** *280* (51), 42274-42282.

64. Caldwell, M. A.; Svendsen, C. N. Heparin, but not other proteoglycans potentiates the mitogenic effects of FGF-2 on mesencephalic precursor cells. *Exp. Neurol.* **1998,** *152* (1), 1-10.

65. Plotnikov, A. N.; Schlessinger, J.; Hubbard, S. R.; Mohammadi, M. Structural basis for FGF receptor dimerization and activation. *Cell* **1999,** *98* (5), 641-650.

66. Uniewicz, K. A.; Ori, A.; Xu, R.; Ahmed, Y.; Wilkinson, M. C.; Fernig, D. G.; Yates, E. A., Differential Scanning Fluorimetry Measurement of Protein Stability Changes upon Binding to Glycosaminoglycans: A Screening Test for Binding Specificity. *Anal. Chem.* **2010,** *82* (9), 3796-3802.

67. Raman, R.; Venkataraman, G.; Ernst, S.; Sasisekharan, V.; Sasisekharan, R., Structural specificity of heparin binding in the fibroblast growth factor family of proteins. *Proc Natl Acad Sci USA* **2003,** *100* (5), 2357-2362.

68. Kreuger, J.; Prydz, K.; Pettersson, R. F.; Lindahl, U.; Salmivirta, M., Characterization of fibroblast growth factor 1 binding heparan sulfate domain. *Glycobiology* **1999,** *9* (7), 723-9.

69. Liu, L. G.; Bytheway, I.; Karoli, T.; Fairweather, J. K.; Cochran, S.; Li, C. P.; Ferro, V., Design, synthesis, FGF-1 binding, and molecular modeling studies of conformationally flexible heparin mimetic disaccharides. *Biorg. Med. Chem. Lett.* **2008,** *18* (1), 344-349.

70. Hu, Y. P.; Zhong, Y. Q.; Chen, Z. G.; Chen, C. Y.; Shi, Z. H.; Zulueta, M. M. L.; Ku, C. C.; Lee, P. Y.; Wang, C. C.; Hung, S. C., Divergent Synthesis of 48 Heparan Sulfate-Based Disaccharides and Probing the Specific Sugar-Fibroblast Growth Factor-1 Interaction. *J. Am. Chem. Soc.* **2012,** *134* (51), 20722-20727.

71. Dou, W. F.; Xu, Y. M.; Pagadala, V.; Pedersen, L. C.; Liu, J., Role of Deacetylase Activity of N-Deacetylase/N-Sulfotransferase 1 in Forming N-Sulfated Domain in Heparan Sulfate. *J. Biol. Chem.* **2015,** *290* (33), 20427-20437.

72. Qin, Y.; Ke, J. Y.; Gu, X.; Fang, J. P.; Wang, W. C.; Cong, Q. F.; Li, J.; Tan, J. Z.; Brunzelle, J. S.; Zhang, C. H.; Jiang, Y.; Melcher, K.; Li, J. P.; Xu, H. E.; Ding, K., Structural and Functional Study of D-Glucuronyl C5-epimerase. *J. Biol. Chem.* **2015,** *290* (8), 4620-4630.

73. Fukuda, M.; Hiraoka, N.; Akama, T. O.; Fukuda, M. N., Carbohydrate-modifying sulfotransferases: Structure, function and pathophysiology. *J. Biol. Chem.* **2001,** *276* (51), 47747-47750.

74. Ornitz, D. M.; Herr, A. B.; Nilsson, M.; Westman, J.; Svahn, C. M.; Waksman, G., Fgf Binding and Fgf Receptor Activation by Synthetic Heparan-Derived Disaccharides and Trisaccharides. *Science* **1995,** *268* (5209), 432-436.

Chapter 4

Glycosyltransferase utilization for heparin thiomimic synthesis

Table of Contents

Chapter – 4 Glycosyltransferase utilization for heparin thiomimic synthesis

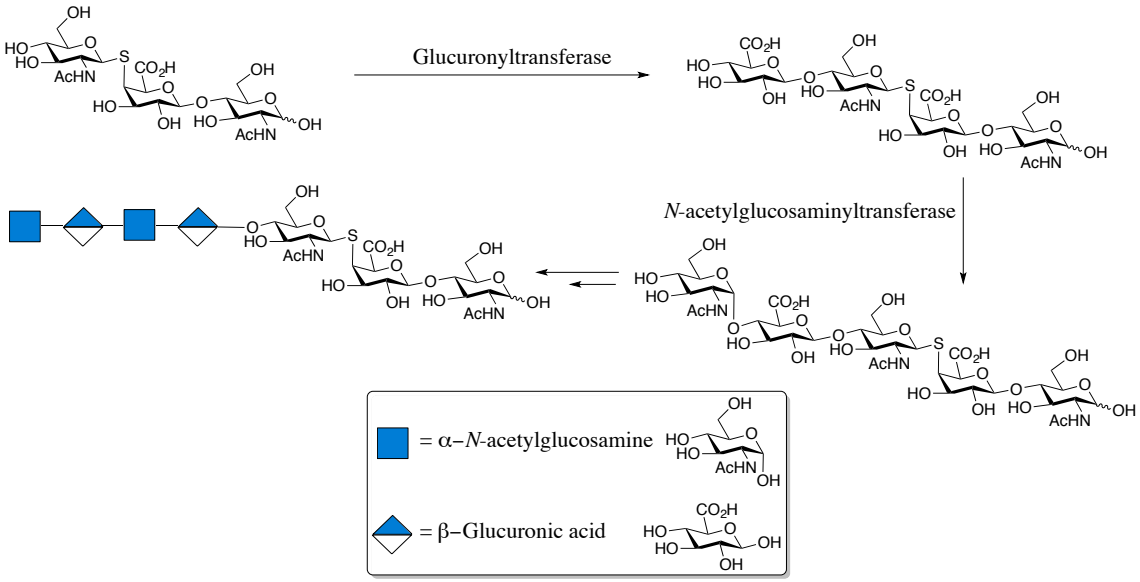
4.1 Introduction

Glycosyltransferase enzymes are involved in biosynthesis of glycosaminoglycans in natural systems. Normally, an *N*-acetylglucosaminyltransferase and a glucuronyltransferase enzyme are responsible for alternately transferring an *N*-acetylglucosamine unit and glucuronic acid unit, respectively, to build up the backbone chain[1-3], known as heparosan[4]. Thereafter, the heparin carbohydrate backbone can be modified by various enzymes such as *O*-sulfotransferases[5], which facilitate transfer of sulfate groups to various hydroxyl positions in the sugar. Another important enzyme, *N*-deacetylase *N*-sulfotransferase[6-7] (NDST), facilitates the deacetylation of the acetamido group in *N*-acetyl glucosamine and installs a sulfate group in its position. Finally, C-5 epimerase[8] acts to reverse the stereochemistry at C-5 to convert glucuronic acid to iduronic acid.

Thio-analogues of sugars rather than the common oxo-sugars have been used successfully as substrates for glycosyltransferase reactions. Tsuruta et al. have reported the usage of GDP-5-thiomannose and GDP-5-thiofucose as substrates for mannosyltransferase and fucosyltransferase respectively[9]. Nishio et al. have reported using 1-thio-α-D-rhamnopyranoside as a glycosyl acceptor for jack bean α-mannosidase catalyzed glycosylation[10]. The usage of a β-galactosyltransferase on a thio acceptor has been demonstrated by Renaudie et al. in the synthesis of lacto-N-neotetraose[11].

All these examples prompted us to explore if a similar enzymatic approach could be followed to perform an enzymatic extension of the thio-mimics of heparin discussed previously. Since the deactivating influence of the sulfate groups does not allow heparin-derived oligosaccharides larger than disaccharides to be used for the thio-coupling reaction, we hypothesized that a combination of *N*-acetylglucosaminyltransferase and glucuronyltransferase can be used to extend the sugar chain to the desired lengths.

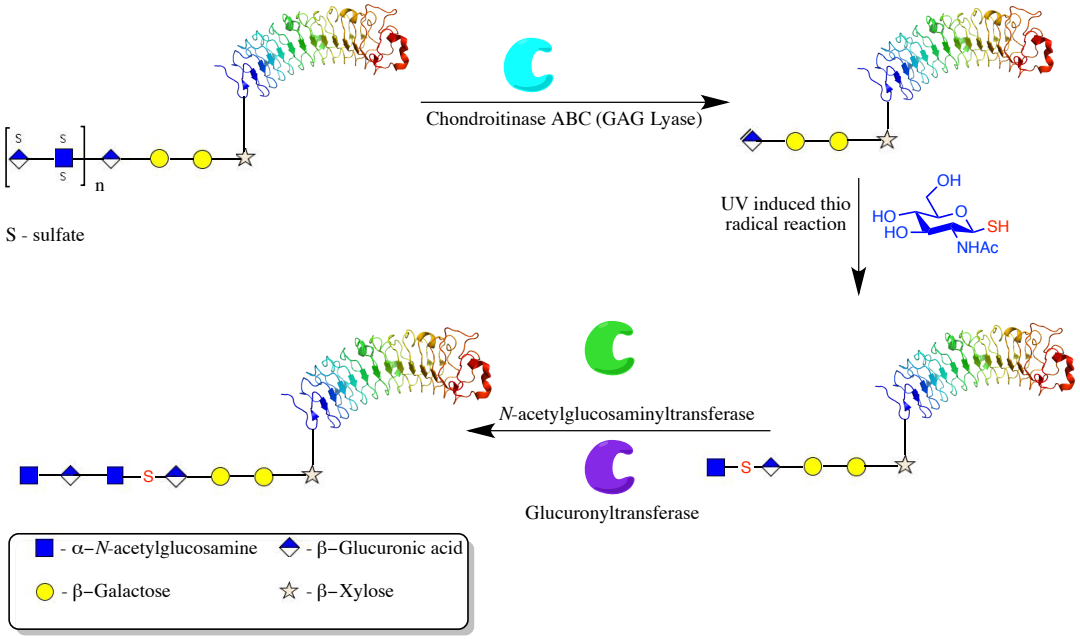
As shown in **Scheme 1**, glucuronyltransferase should be able to use a typical thio-heparin mimic building block with a thio-*N*-acetylglucosamine (thio-GlcNAc) at the non-reducing end, as an acceptor, to transfer a glucuronic acid on to thio-GlcNAc. This can then be further extended using an *N*-acetylglucosaminyltransferase to attach a *N*-acetylglucosamine at the non-reducing end.



Scheme 1: Scheme for usage of glycosyltransferases on thio-mimics of heparin for chain extension

The two enzymes can thereafter be used in tandem to extend on the sugar backbone with the thio linkage already incorporated. An additional advantage of using this approach is the hydrolysis-resistant thio linkage. This allows stable analogs of the heparin GAG backbone to be extended to have biological activities comparable to naturally available heparin.

A potentially useful application of the enzymatic extension approach on thiosugars is for building up of synthetic proteoglycans. It is known that in nature, proteoglycans consist of heterogeneous glycosaminoglycan chains attached to the core protein through a tetrasaccharide linker consisting of xylose-galactose-galactose-glucuronic acid[12-14]. Exhaustive lyase treatment on such a proteoglycan should be able to chop off the heterogeneous chain of GAG, leaving the tetrasaccharide linker behind with an unsaturation on the non-reducing end glucuronic acid. A thio-*N*-acetylglucosamine monosaccharide can then be attached to the unsaturation using aqueous phase radical chemistry already developed. Subsequently, the chain can be further extended using pmHS2 and KfiA as shown in **Scheme 2**.



Scheme 2: Plausible application of combined thiyl-alkene radical reaction and enzymatic extension for synthesis of defined proteoglycans

In addition to yielding synthetic proteoglycans, which has not been reported before, it also can be used to design mixed proteoglycans where the core protein originally linked to a specific type of glycosaminoglycan, in nature, can be used to link to a different type of glycosaminoglycan. These mixed proteoglycans might have novel bioactivities.

4.2 Results and Discussion

The first consideration towards this approach was the selection of appropriate transferases for both *N*-acetylglucosamine and glucuronic acid transfer. Glycosaminoglycans are biosynthesized in various organisms, and in each case a specific set of enzymes are involved. Recent chemoenzymatic synthesis approaches towards heparin synthesis have focused on

utilization of KfiA as an *N*-acetylglucosaminyltransferase and pmHS2 as a glucuronyltransferase.

As discussed in the introduction, the activity of pmHS2 therefore needs to be evaluated with respect to a thiosugar-containing oligosaccharide. Since pmHS2 has been reported to have broad substrate specificity, wherein it can use various unnatural UDP-sugars[15], it was hypothesized that it might show a similar tolerance towards unnatural acceptors too. In order to evaluate the ability of a glucuronyltransferase to use an oligosaccharide containing thio-GlcNAc at the non-reducing end, as an acceptor, several criteria needs to be kept in mind:

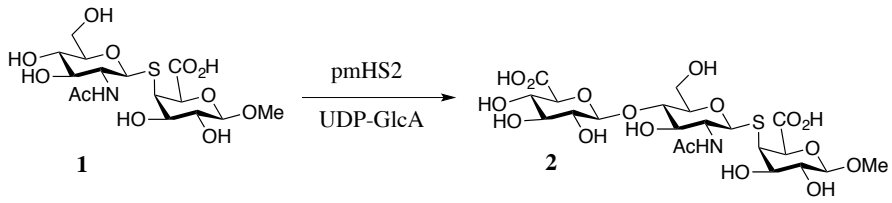
a) The substrate should be easily available.

b) In terms of chain length it should be a disaccharide, at the minimum. This is because it has been reported that pmHS2 is inefficient in using a glucosamine monosaccharide as an acceptor[16].

c) The linkage stereochemistry between the thiosugar and the next sugar should be the same as is obtained in the thiol-ene radical coupling reaction for preparation of heparin and heparan sulfate mimics. This requirement is specifically because we wanted to utilize the unnatural heparin and heparan sulfate mimics developed by thiol-ene reaction as acceptors for enzymatic extension.

Keeping all these considerations in mind, the substrate chosen is depicted as **1** in **Scheme 3**. Incidentally, it is one of the deprotected model substrates prepared before. Since it can be produced easily, is a disaccharide and also has the same linkage stereochemistry (for the thiosugar) as found in the synthesized heparin thio-mimics, it seemed like a good substrate for testing and optimizing the activity of pmHS2.

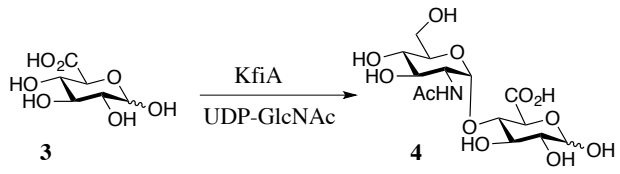
The specific difference regarding the linkage stereochemistry was the axial-linkage of the model, which differs from natural substrates. In order to investigate the substrate tolerance of pmHS2 towards unnatural stereochemical linkages, substrate **1** is ideal as it is composed of β-*N*-acetylglucosamine linked to the glucuronide in axial conformation. In contrast, naturally available heparosan contains α-*N*-acetylglucosamine linked to glucuronic acid in an equatorial conformation. Also, since α-sugars were found to have less reactivity towards thiol-ene radical reaction, which led to lowered yield of the desired thio-heparin mimic and model disaccharide, the β-sugar disaccharide model was chosen for synthetic convenience. Therefore, it was imperative to test and subsequently optimize pmHS2 reactivity against β-GlcNAc-containing acceptors.



Scheme 3: Proposed reaction for utilization of the model substrate **1** as acceptor for glucuronyltransferase, pmHS2.

On the other hand, the minimum length of oligosaccharide on which KfiA is active has been reported to be a disaccharide[17]. The disaccharide in this case was isolated from heparosan by nitrous acid treatment after a considerable number of steps. Li et al.[18] has also reported the usage of a glucuronic acid attached to a fluorophore through the reducing end as a substrate for KfiA. However, to date no investigation has been carried out as to whether commercially available glucuronic acid (a monosaccharide) can be used as an acceptor for KfiA. Thus investigating the reaction of using glucuronic acid as an acceptor by KfiA, as depicted in

Scheme 4, would help us in re-establishing the minimum substrate specificity for KfiA as well as enable us to extend the heparin mimic chain as depicted in **Scheme 1**.



Scheme 4: Proposed reaction for utilization of glucuronic acid as acceptor for KfiA.

We sought to profile the unnatural substrate acceptability and corresponding activity of these glycosyltransferases. Accordingly, the two enzymes needed to be expressed in appropriate vectors in abundant quantities for carrying out the biocatalytic reactions.

4.2.1 Expression of KfiA

The expression of KfiA has been reported by Chen et al.[17] The gene from the host organism *E. coli* K5 has been cloned into pET 21b vector using *Bam*H1 and *Xho*1 as the restriction enzymes. Consequently, the protein was expressed as a *C*-terminal six-histidine tagged protein. The yield has been reported to be moderate with 10 mg being recovered from 1 litre of culture. Consequently, we decided to synthesize the gene in a similar way. The synthesized plasmid (KfiA in pET 21b vector) was procured from GenScript. Thereafter, the plasmid was transformed into XL10-Gold cells for cloning. The cloned gene was expressed in BL21(DE3), and grown in LB medium supplemented with 100 μg/ml ampicillin, at 37 °C. Induction was carried out at an A_{600} of 0.6, when isopropyl β-D-thiogalactopyranoside (IPTG)

was added to induce the expression. The culture was then subjected to overnight shaking at 22 °C. After harvesting the bacteria, the cells were sonicated in lysis buffer (25 mM Tris HCl, pH-7.5, 500 mM NaCl, 30 mM imidazole) and purified using Nickel Sepharose High Performance columns (His Trap, GE) using a linear gradient of elution buffer (25 mM Tris HCl, pH-7.5, 500 mM NaCl, 250 mM imidazole). The protein was further purified using size-exclusion chromatography. Thereafter, the protein was dialysed into the storage buffer (25 mM Tris HCl, pH-7.5, 150 mM NaCl).

The identity of the expressed protein was verified using gel electrophoresis (**Figure 1**) and further by mass spectrometry (**Figure 2**).

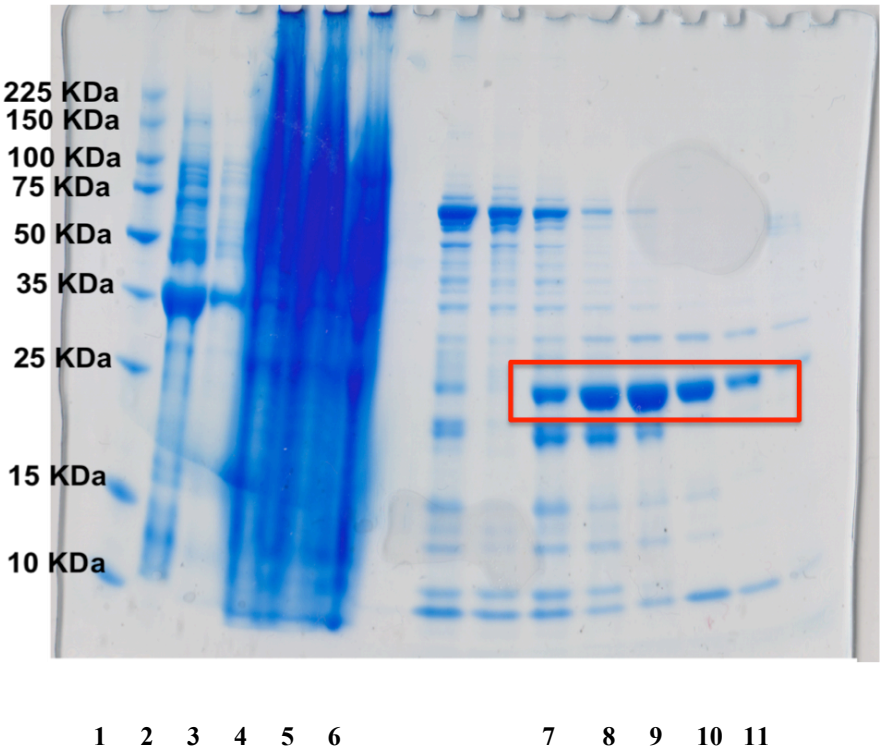


Figure 1: Gel electrophoresis of fractions obtained by HisTrap purification of KfiA. **1** represents 10-225 KDa ladder, **2** represents inclusion bodies (dissolved in 8M urea), **3** shows inclusion bodies (5x dilution) **4**, **5** and **6** flowthrough, **7**, **8**, **9**, **10** and **11** correspond to purified fractions.

The final yield of the protein was quite low (2 mg/litre) as against the reported value. As can be seen from **Figure 1**, there was no presence of the protein in either inclusion bodies (which would indicate that the protein was precipitating during expression) or the HisTrap flow-through (indicating loss of protein because of insufficient binding to the column). Efforts were therefore made to further optimize the expression conditions. Carrying out the induction using IPTG at an O.D. (A_{600}) of 0.8 was seen to increase the yield slightly. The yield decreased on inducing the cells at an O.D. of 1.0. This is presumably because at this higher cell concentration the nutrients in the medium become a limiting factor. The induced stress on the cells leads to less

production of protein. Changing the amount of IPTG (both doubling as well as reducing the amount) for the induction did not have any noticeable effect on the expression yield. The other parameter, which was changed, was the induction phase temperature. Decreasing the temperature below 22 °C led to decrease in yield of the protein (with respect to **Entry 1, Table 1**) while increasing it led to precipitation of the protein in the form of inclusion bodies. Lastly, when the induction phase was increased to 22 h from 16 h, the yield was observed to have increased slightly.

Sonication during the process of lysis has been sometimes reported to lead to formation of protein aggregates, causing soluble protein loss[19]. Therefore, cells were also lysed using repeated freeze thaw cycles. The yield however was found to be the same.

Serial Number	Expression Parameter Changed	Yield
1.	Induction at O.D. 0.8	3 mg/litre
2.	Induction at O.D. 1.0	1.8 mg/litre
3.	IPTG concentration (2 mM)	2.5 mg/litre
4.	IPTG concentration (0.5 mM)	2.3 mg/litre
5.	Induction phase temperature (15 °C) (at O.D. 0.8)	2.7 mg/litre
6.	Induction phase temperature (37 °C)	1.3 mg/litre
7.	Induction phase time (22 h)	4.3 mg/litre

Table 1: Optimization attempts to increase yield of expressed KfiA. Beneficial conditions were carried forward for the next parameter change, and therefore the yield of 4.3 mg/litre reflects a combined effect of induction at increased O.D. coupled with longer induction time.

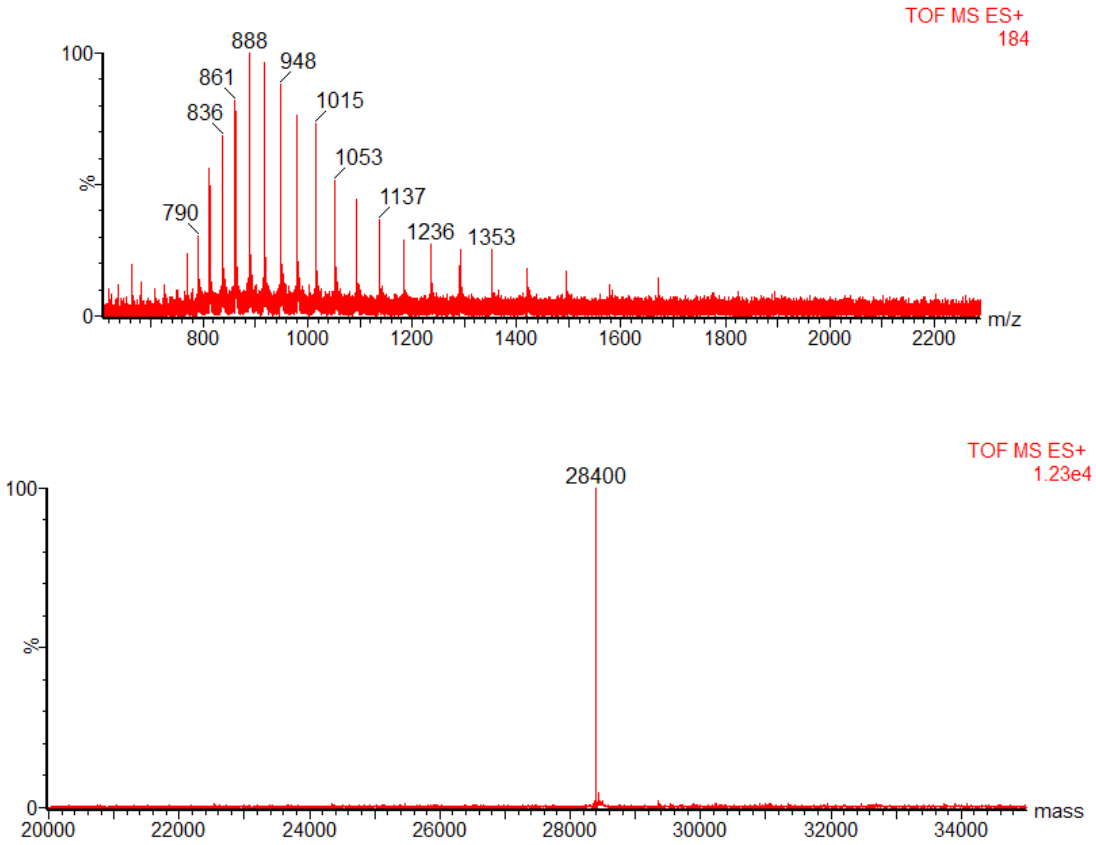


Figure 2: Expected mass of KfiA = 28397 Da, found 28400 Da

As can be seen from **Table 1**, the yield of the protein could not be increased to a large extent by these changes. The reported procedures for expression for KfiA mention the utilization of BL21 star (DE3) cell line along with co-expression of these enzymes with chaperone plasmid pGro7 (Takara), expressing GroEL and GroES chaperone proteins[17]. This cell line differs from normal BL21(DE3) in that it contains an additional mutation in a gene coding for RNase E, which degrades mRNA within the cell[20-22]. Usage of this strain therefore results in increased mRNA stability and consequently increased protein expression and therefore this maybe one of the contributing factors behind the low expression yield observed.

Another reason might be the lack of pGro7 plasmid being co-expressed. Chaperone proteins GroEL and GroES are known to bind to hydrophobic domains, which are exposed to solvents and also to each other and assist in folding of their substrates by ATP-driven conformational changes. Thus, it prevents aggregation of nascent peptides[23-24]. In the absence of these folding mechanisms, however, the misfolded proteins are exposed to partial or complete proteolytic cleavage[25]. Presumably, the KfiA protein in absence of the chaperones is also being exposed to proteolytic cleavage. This hypothesis of KfiA degradation by internal proteases, is supported by the observation of additional bands as seen from the gel in **Figure 1**. Apart from the pure protein, bands of probable cleaved protein can also be seen on the gel, which might also contain the *C*-terminal His tag region. These bands were obtained in all cases of the optimization trials. However, in the absence of chaperones, during over-expression of the protein, some precipitation in the inclusion bodies is also expected. But as can be seen from **Figure 1**, no desired protein is visible in the inclusion bodies. Nonetheless in the absence of this cell line and the chaperone plasmid pGro7, we decided to go ahead with the expression conditions giving highest yield (**Entry 7, Table 1**).

In order to investigate the folding state of the protein, a CD experiment was done on the protein (**Figure 3**).

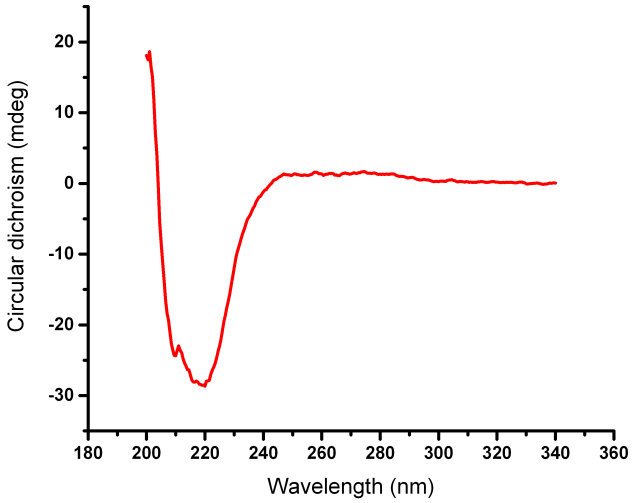


Figure 3: CD spectrum of KfiA

While the spectrum reveals that the protein has got a definite 3D structure, instead of a random coil, it is not possible to say if all of the protein molecules are folded in the correct conformation. It is also not possible to say if the entire protein molecule is folded in the correct conformation. This is because the CD structure of the protein has not been reported. This is in spite of the fact that the spectrum is similar to the theoretically predicted CD spectrum based on the protein function (**Experimental Section 4.4.2)**

KfiA was initially dialysed into a storage buffer comprising of 20 mM Tris, pH-7.5 and 20% glycerol as reported[18] and stored at -80 °C after flash freezing. However, it was found that a considerable amount of the protein precipitated during dialysis, after thawing, to remove the glycerol before activity tests. Even though usage of 0.1% Triton-X detergent during the lysis stage of the cells during protein purification did stabilize the protein and lowered the

precipitation considerably, presence of the detergent was found to be detrimental for the activity of the protein and also hampered in obtaining mass spectra. Triton-X is normally used since being an amphiphilic detergent, it helps in solubilizing the protein by shielding the hydrophobic regions of the protein from water. Passing the protein through a detergent removal column (Pierce^{TM} Detergent Removal Spin Column) or running an extended gradient on a C-18 column (in LC stage) before injecting into the mass spectrometer was not beneficial either. Consequently, we stopped using Triton-X in the lysis buffer and instead the storage buffer was changed to 25 mM Tris, pH-7.5 and 150 mM NaCl[26]. When the protein was stored with 10% glycerol in this buffer, protein precipitation was observed to have stopped after thawing and dialysis. Presumably, the increased stability of the protein is due to the charge interaction stabilization between the salt and the charged amino acid residues of the protein.

4.2.2 Expression of pmHS2

pmHS2, is a bacterial enzyme involved in the extension of heparosan chains from *Pasteurella multocida*[27]. Heparosan is the unepimerized and unsulfated precursor of heparin. Li et al. has reported that the expression level of *N*-His_6-tagged pmHS2 is higher than the *C*-His_6-tagged construct[18]. Accordingly, we generated our pmHS2 construct as a *N*-His_6-tagged protein sub-cloned in pET-15b vector using *Nde*1 and *Bam*H1 as the restriction enzymes. The described gene construct in pET-15b was procured from Genscript and transformed into XL10-Gold cells in a similar way as done for KfiA, for cloning. After verifying the gene sequencing, expression was carried out using standard procedures in BL21(DE3) cells using 1mM IPTG for induction and 25 °C as the induction phase temperature. Induction was done at an O.D. of 0.8. Purification

of the *N*-terminal His tagged protein using Histrap column, again yielded the protein in moderate yields (7 mg/litre) as against the reported yield of 17 mg/litre reported by Li et al. It is clear from the gel obtained after purification of the protein that there was expression of the protein as inclusion bodies apart from the soluble fractions (**Figure 4**). Similarly, there was loss of the protein in HisTrap nickel column flow-through (lane **4**,**5** and **6**, **Figure 4**).

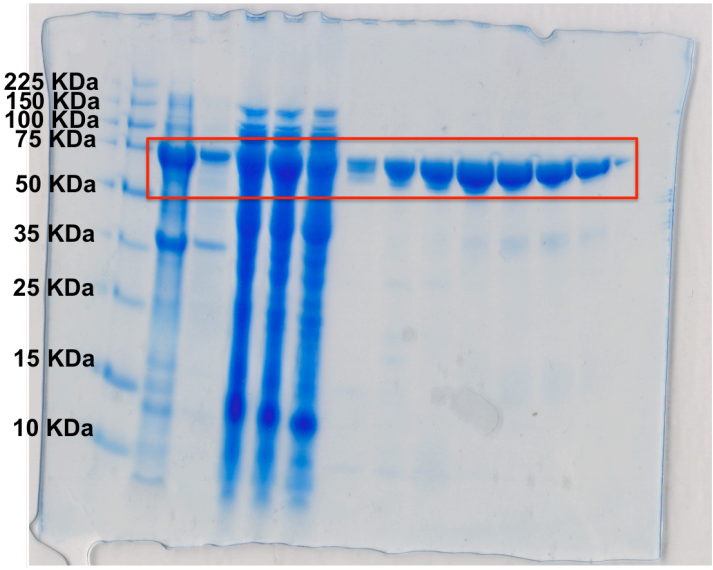


Figure 4: Gel electrophoresis of fractions obtained by HisTrap purification of pmHS2. **1** represents 10-225 KDa ladder, **2** shows inclusion bodies (dissolved in 8M urea), **3** is inclusion bodies (5x dilution), **4**, **5** and **6** correspond to flowthrough, while **7**, **8**, **9**, **10**, **11**, **12** and **13** correspond to purified fractions.

The loss of protein in the flow-through was prevented by loading a lower volume of the lysate on the HisTrap column. This increased the yield to 9 mg/litre. Therefore, similar optimization conditions used for KfiA were tried for pmHS2 (**Table 2**). Here, a combination of 0.5 mM IPTG and reduced induction phase temperature of 15 °C along with an increased

induction phase time of 22 h were found to be beneficial, yielding an increase to 12 mg/litre. This is consistent with reported means of increasing the expression yield of large sized proteins as lowering the induction phase temperature increases solubility of the expressed protein[28-29]. This is because, at lower temperature, the rate of expression of the protein is reduced thus providing more time for better folding of the protein, reflected in its increased solubility. Also proteases are less active at lower temperature and therefore loss of protein due to proteolytic digestion is reduced[30]. Thus the induction phase temperature was kept at 15 °C throughout.

Serial Number	Expression Parameter Changed	Yield
1.	Induction at O.D. 1.0	7.3 mg/litre
2.	IPTG concentration (2 mM)	6.2 mg/litre
3.	IPTG concentration (0.5 mM)	9.4 mg/litre
4.	Induction phase temperature (15 °C)	10.1 mg/litre
5.	Induction phase temperature (37 °C)	6.4 mg/litre
6.	Induction phase time (22 h)	12 mg/litre

Table 2: Optimization of expression conditions for pmHS2. Beneficial conditions were carried forward for the next parameter change, and therefore the yield of 12 mg/litre reflects a combined effect of increase in induction time, decrease in induction phase temperature and decrease in IPTG concentration.

The structure of expressed pmHS2 was investigated with the CD spectrum too (**Figure 5**).

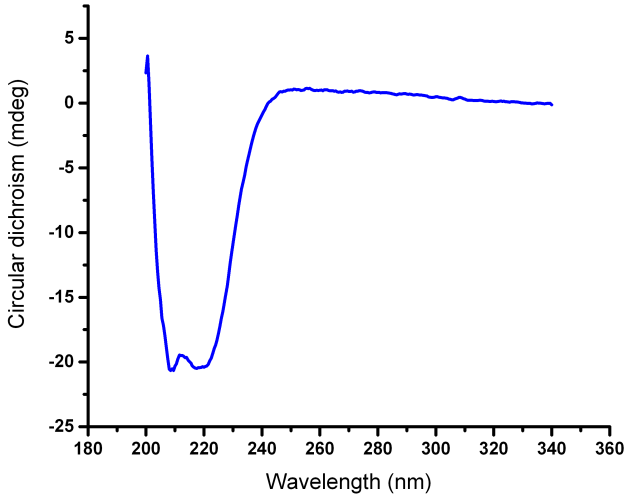


Figure 5: CD spectrum of pmHS2

However, as with KfiA, since the CD spectrum of pmHS2 is yet to be reported, the experimentally obtained spectrum could not be compared with a reference spectrum, to have an insight into the folding state of the entire protein. The predicted CD spectrum was similar with the experimentally obtained spectrum though (**Experimental Section 4.4.2**).

When Xu et al. reported the preparative scale synthesis of heparin oligosaccharides, they utilized the pGro7 chaperone plasmid for co-expression of GroEL and GroES chaperone proteins with pmHS2 as they did for KfiA[16]. In addition, pmHS2 was also expressed in BL21 star (DE3) cell line. As explained for the expression of KfiA, these factors might have an important role to play regarding the relatively low yield of the expressed protein. As reported for KfiA, storage of the protein in 20 mM Tris buffer and 20% glycerol without salt resulted in precipitation of pmHS2 as well when it was thawed from -80 °C and dialysed to remove glycerol. Subsequently, introduction of 150 mM NaCl in storage buffer helped in stabilization of the expressed protein.

Use of Triton-X in the lysis buffer did stabilize the protein to a great extent but at the cost of adversely affecting the activity and interfering with the ionization process during mass spectrometric identification. As a result the use of the detergent in subsequent trials was discontinued.

The mass spectra of the protein (**Figure 6**) revealed a second peak apart from the expected peak at +178 Da.

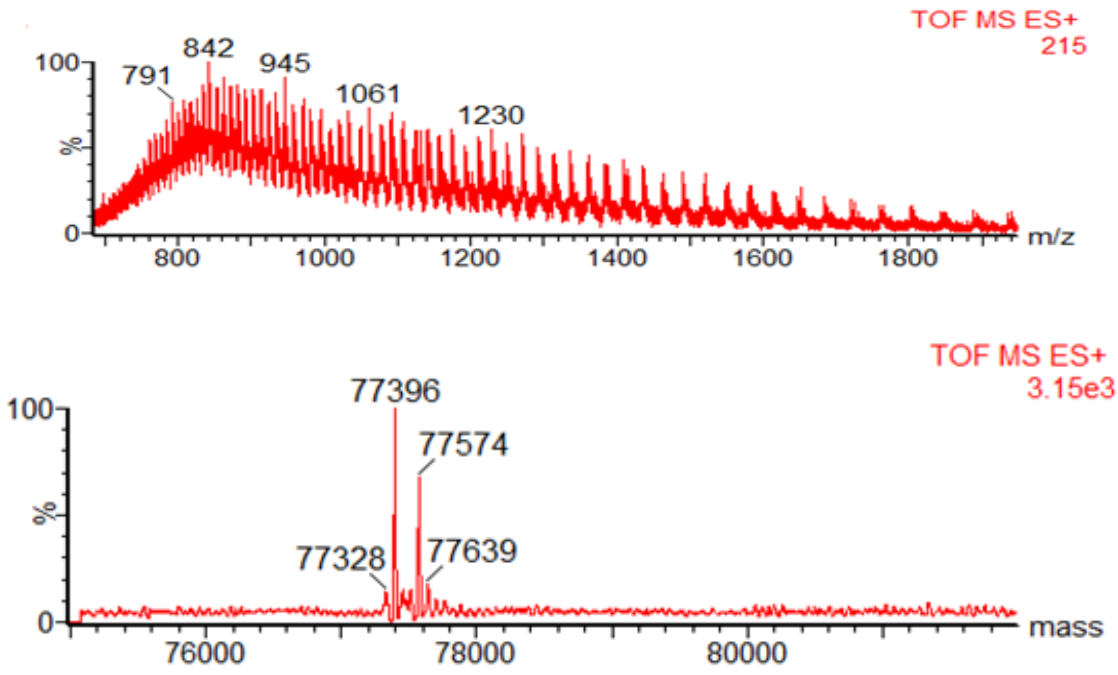
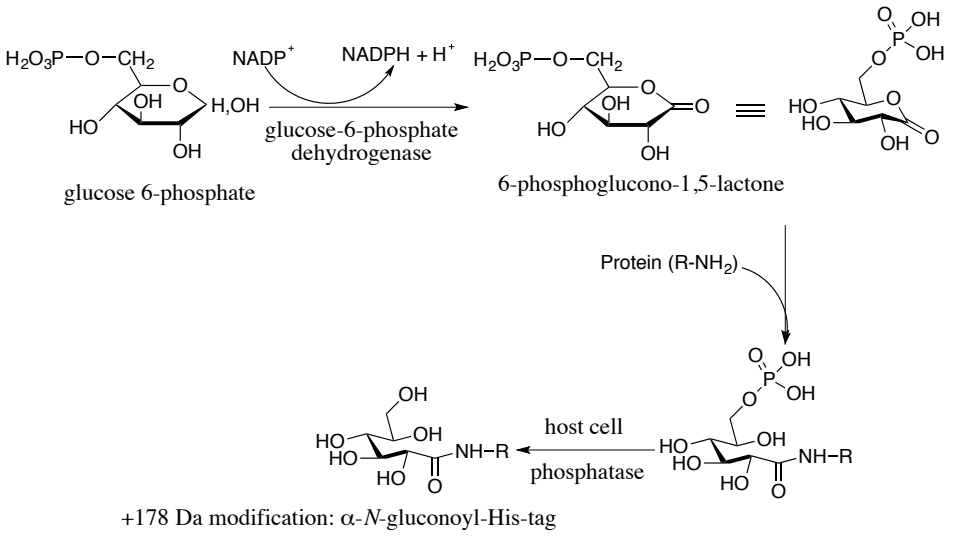


Figure 6: Expected mass of pmHS2 = 77399 Da, found = 77396 Da

On further search through literature, it was found that the second peak is most probably due to *N*-terminal gluconoylation of the *N*-His_6 tagged protein. α-*N*-terminal gluconoylation has been reported on protein residues containing Gly-Ser-Ser-$[His]_6$ on the *N*-terminus, wherein it has

been proposed that the amino group at the *N*-terminus reacts with 6-phosphoglucono-1,5-lactone to give rise to the gluconoylated product as shown in **Scheme 5**[31].



Scheme 5: Proposed biochemical pathway for α-*N* gluonoylation of *N*-terminal His tagged proteins[31-32]

As can be seen from the protein sequence of expressed pmHS2 (**Experimental Section 4.4.1**), the *N*-terminus of the expressed protein has Gly-Ser-Ser-$[His]_6$, so this supports our hypothesis that the second peak comes from *N*-terminal gluconoylation. Incidentally, Geoghegan et al. has reported using pET-15b vector (the same vector used by us for the expression of pmHS2) as one of the vectors for the study of α-*N*-gluconoylation[31]. This post-translational modification is presumed to occur because of the accumulation of 6-phosphogluconolactone in the absence of phosphogluconolactonase in the phosphate pentose pathway[33]. *N*-terminal gluconolyation is known to interfere with crystallization of proteins[34]. However, even though general glycation of amino acid residues of the side chains of non-terminal amino acid residues have been reported to adversely affect the activity of enzymes[35-36], any adverse activity of this

particular post translational modification on enzymatic activity has not been reported, to the best of our knowledge. Hypothesizing that this might be because the modification takes place at the terminus of the protein chain, leaving the active site residues intact, it was decided to go ahead with the expressed pmHS2 having the modification.

4.2.3 Using KfiA and pmHS2 for activity studies

With both enzymes in hand, optimization of conditions for the glycosyltransferase reaction of both enzymes was initiated.

We sought to use KfiA for transferring *N*-acetylglucosamine onto glucuronic acid (**Scheme 4**). The reaction conditions described by Liu et al. for transferring a UDP-GlcNAc analogue (UDP-GlcNTFA) onto a disaccharide derived from nitrous acid degradation of heparosan, were used initially[37]. Briefly, UDP-GlcNAc and glucuronic acid (5 μM) were added in the ratio of 1:1.2, along with 10 mM $MgCl_2$ as co factor and 0.1 mg of the enzyme in 1ml pH-7.2 Tris buffer at 25 °C. The reaction progress was estimated using HPLC, where the conversion of UDP-GlcNAc to UDP was noted at regular intervals. This was not a direct readout of product formation, but informed the timescale for purification and analysis.

Divalent metal co-factors have been reported be essential to the activity of several glycosyltransferases (both inverting and retaining glycosyltransferase)[38] where it is involved in coordinating with oxygen atoms of the two phosphate groups of UDP-sugars as well as one of the aspartate groups of the DXD conserved motif in most glycosyltransferases. However, the metal ions are a requirement only for the glycosyltransferase A (GT-A) family of enzymes. This induces a conformational change, paving the way for the acceptor to bind to the active site[39].

With the aforementioned reaction conditions, however, we could not detect the formation of product even after overnight reaction. Mg^{2+} was used as the cofactor here, but studies by Li et al.[18] and Sugiura et al.[26] has suggested that Mn^{2+} is a much more effective cofactor for this enzyme. Therefore, the co-factor was changed to $MnCl_2$ for all subsequent reactions. Fortunately, we were able to observe the expected product this time. The identity of the product was extensively verified by mass spectrometry and further by MS/MS (**Experimental Section 4.4.4**)

As shown in **Table 3**, lowering the pH from 7.2 to 6.5 and through to 5.5 had a significant positive effect on the reaction yield as seen from the UDP-GlcNAc to UDP conversion, with lower pH leading to greater yield. However, the stability of the protein was significantly affected as the protein precipitated much more quickly with decreasing pH. The increase in yield is consistent with what has been reported in literature[18].

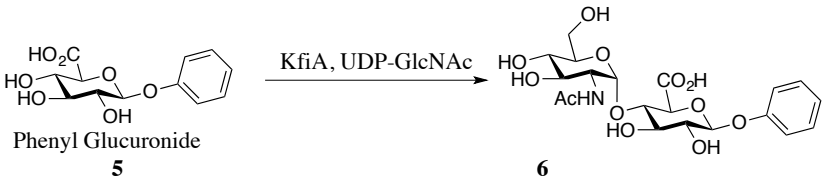
A CD melting experiment was done to determine the maximum temperature that KfiA can be subjected to before it denatures. Melting temperature for KfiA was determined to be 47 °C (**Experimental Section 4.4.3**). Accordingly, we decided to expose the protein reaction to a temperature maximum of 37 °C. Increasing the temperature from 25 °C to 30 °C and thereafter to 37 °C led to increase in conversion at the expense of protein stability. While, increasing the equivalents of either UDP-GlcNAc or glucuronic acid did not have any significant effect, increasing the concentration of the protein did increase the conversion, albeit to a small extent.

Serial Number	Reaction condition	Conversion (from UDP GlcNAc to UDP)
1.	Mg^{2+} as cofactor	-
2.	Mn^{2+} as cofactor	2 %
3.	pH – 6.5 reaction buffer	8.3 %
4.	pH – 5.5 reaction buffer	9.2 %
5.	UDP-GlcNAc:glucuronic acid (4:1)	3.4 %
6.	UDP-GlcNAc:glucuronic acid (1:4)	3.5 %
7.	Reaction temperature 30 °C	6.6 %
8.	Reaction temperature 37 °C	7.6 %
9.	Enzyme concentration (0.4 mg/ml)	4.5 %
10.	Enzyme concentration (0.6 mg/ml)	5.8 %
11.	Enzyme concentration (0.8 mg/ml)	11.2 %*

Table 3: Reaction condition optimization for KfiA reaction using glucuronic acid as acceptor. Conversions calculated from standard curves of UDP-GlcNAc and UDP. (individual parameters as stated in the entries changed from initial condition mentioned in text)

* Conversion after 48 h of reaction. 0.3 mg/ml of enzyme added at the beginning of the reaction and a further 0.5 mg/ml added after 24 h. Reaction buffer pH – 6.5. In all other cases, conversion recorded after overnight reaction.

Exploring the substrate scope of the enzyme, phenyl glucuronide was also used as a potential acceptor for KfiA as represented in **Scheme 6**. This had the added benefit that the conversion of the starting material to product could be monitored directly by HPLC, without relying on UDP-GlcNAc to UDP conversion.



Scheme 6: Phenyl glucuronide as acceptor for KfiA

Again, the expected product could be observed on MS and the product identity was verified by HRMS. The conversion as calculated from absorbance of phenyl glucuronide and the product was found to be comparable with that obtained with glucuronic acid when similar reaction conditions were used.

In order to further characterize the product, attempts were made to purify the product. Unfortunately, a crude NMR could not be obtained because of presence of paramagnetic Mn^{2+} in the solution. Thus, efforts to remove the metal ions using a cation exchange column were attempted. Though a TLC system comprising of *n*-butanol/formic acid/water has been reported for the analysis of glycosaminoglycan oligosaccharides[40], it did not prove to be particularly effective in separating out the high concentration of UDP-GlcNAc and glucuronic acid from the low concentration of the product. Reverse phase column chromatography was not effective either as the compounds did not bind to the column at all. Acetylation of the reaction mixture for further purification was not effective because a distinct spot corresponding to the product could not be observed due to streaking of the electronegative phosphate and carboxylic groups and low concentration of the product. Conducting the reaction on a larger scale and adding more enzyme (up to 8 mgs in 10 ml of reaction mixture) with longer reaction time was not effective in increasing the product yield. Presumably, this is because of the fact that the most of the enzyme

was not properly folded due to it being expressed in the absence of the chaperone proteins, and the residual activity being seen was because of the minor fraction of the enzyme expressed in the correctly folded state. It should also be noted at this point that it has been reported that the catalytic efficiency of the enzyme increases with increase in chain length of the acceptor[17]. Hence, an acceptor with longer chain length might experience higher rate of glycosyltransferase activity by KfiA.

Glycosyltransferase reactions with pmHS2 were similarly initiated with the disaccharide **1** as shown in **Scheme 3**. The reaction conditions started with were the ones reported for the transfer of glucuronic acid onto a trisaccharide of heparosan[37]. The reaction conditons are 10 mM $MgCl_2$, disaccharide **1** and UDP-GlcA in the ratio of 1:1.2 and 0.2 mg of the enzyme in 1ml pH-7.2 Tris buffer at 25 °C. Monitoring the conversion of UDP-GlcA to UDP using a polyamino column proved to be effective for the optimization reactions.

After not observing any product after overnight reaction, the cofactor was changed to 10 mM $MnCl_2$ and the buffer pH lowered. As noted in **Table 4**, fortunately product formation was observed. The product was characterized by MS and further by MS/MS to reveal that it was indeed the expected product (**Experimental Section 4.4.4**). Thus, this was the first time that pmHS2 had been shown to perform glycosyltransferase reaction on a thio-linked acceptor. Also, as mentioned before, this also proves that the active site of pmHS2 has sufficient tolerance for an unnatural stereochemistry of linkage.

Nevertheless attempts were made to increase the yield of the reaction. The increased GlcA transferase activity of pmHS2 at lower pH is precedented[15], with a pH range of 4-5 being

reported as optimal. In our hands, the stability of the enzyme was also seen to decrease with lowered pH.

The melting temperature of pmHS2 was determined to be 41 °C (**Experimental Section 4.4.3**) and consequently the temperature max for the reaction was fixed at 37 °C.

Increased reaction temperature also led to a moderate increase in yield, again at the cost of stability of the protein. Increasing the equivalents of either the acceptor disaccharide or UDP-GlcA had no observable effect in increasing the yield of the reaction as no UDP formation was detected. Adding more enzyme after regular intervals, on the other hand seemed to be more promising in this case.

Accordingly, a large scale reaction was set up using 10 mgs of the acceptor disaccharide **1**, in 5 ml MES buffer at pH – 6.5, 1 mg of pmHS2, 10 mM $MnCl_2$ and 1.2 equivalents of UDP-GlcA with respect to the acceptor, at 30 °C. After overnight reaction, a very low conversion of 3.2 % was seen, with protein precipitation. Since an excess of UDP-GlcA was still left behind in the reaction mixture, another 3 mgs. of the enzyme was added and setup for another overnight reaction. Observing the conversion to be still low at 5.3%, the reaction was further left for another overnight after addition of another 6 mgs of the enzyme. The conversion yield after 3 days was estimated at 8.8%.

Serial Number	Reaction condition	Conversion (from UDP GlcA to UDP)
1.	Mg^{2+} as cofactor	-
2.	pH – 6.5 reaction buffer (with Mn^{2+})	3.4 %
3.	pH – 5.5 reaction buffer	3.8 %
4.	UDP-GlcNAc:glucuronic acid (4:1) (pH – 7.2)	-
5.	UDP-GlcNAc:glucuronic acid (1:4) (pH – 7.2)	-
6.	Reaction temperature 30 °C (pH – 7.2)	4.2 %
7.	Reaction temperature 37 °C (pH – 7.2)	4.7 %
8.	Enzyme concentration (0.4 mg/ml)	2.1 %
9.	Enzyme concentration (1 mg/ml)	2.9 %
10.	Enzyme concentration (2 mg/ml) (pH – 6.5)	8.8 %*

Table 4: Reaction condition optimization for pmHS2 reaction using disaccharide **1** as acceptor. Conversions calculated from standard curves of UDP-GlcA and UDP after overnight reaction.

* Enzyme added sequentially as mentioned in discussion with final concentration of enzyme after three days being (2 mg/ml)

Attempts to purify the product using normal chromatography with a combination of water/butanol/formic acid was not successful because of the low concentration of the product and only UDP-GlcA and starting disaccharide could be recovered. Acetylation of the reaction mixture to purify the protected product again led to no recovery of the same because of streaking on silica due to the unprotected negative charges. Size exclusion purification has been reported as a means to purify this kind of mixtures[37]. However, no observable product could be detected when run with 250 mM ammonium bicarbonate, possibly due to dilution. It was apparent that the

yield of the product needed to be increased before purification of the product could be carried out for NMR studies.

The low yield of the product can be presumably due to the following reasons:

1. Expression of the enzyme in the absence of chaperone proteins had led to only a minor fraction of the enzyme being folded in the correct conformation and therefore functionally active.

2. The stereochemistry of linkage of the acceptor is different from that present in natural sources and therefore, the acceptor was not accommodated in the active site as easily as its naturally available counterpart leading to lower conversion.

3. The thio linkage between glucuronide derivative and *N*-acetylglucosamine is unnatural too as in naturally available counterparts the linking atom is oxygen. This might have led to lower acceptance of the acceptor by the enzyme because of size difference.

4. It has been reported that pmHS2 is inefficient in extending the chain from sugars as small as glucosamine monosaccharide[16]. It is probable therefore, that the activity of the enzyme increases with increase in chain length of the acceptor as has been observed for KfiA. Using a longer oligosaccharide as acceptor instead of a disaccharide, therefore, might result in more activity of pmHS2.

5. Enzyme activity can also be inhibited by the products being formed. Thus inhibition of the enzyme activity by the products needs to be investigated too. This holds true for the KfiA reaction too, discussed previously. In this respect, it should be mentioned that Chen et al. has reported inhibition of KfiA activity by UDP (one of the products of the reaction)[17].

Performing analytical size exclusion chromatography can validate the speculation, regarding improper folding of the protein in absence of chaperones, for both KfiA and pmHS2. This will give an idea about the proportion of protein in the unfolded state, if any. In case, there is unfolded protein, it will affect the activity of the enzyme leading to low yield of the reaction. Moreover, kinetic experiments should be performed with the expressed enzymes, on substrates used by other research groups for the same set of enzymes and for which kinetic parameters are already available. This will enable us to compare those parameters and determine if the reason for the low yield of the reaction is because of unnatural substrates (in the form of thio substrates and unnatural stereochemistry of linkage) being used. Kinetic parameter comparison with experiments using longer more native substrates will also enable us to determine how the usage of shorter acceptors is affecting the yield of the enzymatic reactions. Regarding product inhibition, it is necessary to probe the effect of the products on KfiA activity. This holds true for the pmHS2 reaction too.

We think the next step to try would be to co-express both the enzymes with chaperone plasmids. Presumably this would lead to higher activity of both the enzymes. It would be useful to further characterize the enzymatic products especially using NMR, which would give an insight to the stereochemistry of the linkage being formed.

4.3 Summary

We expressed two bacterial enzymes to attempt a biocatalytic extension to heparin-like building blocks. Coupled with the thio-alkene radical reaction, this would thus comprise a chemo-enzymatic approach to defined heparin analogs. This required the expression and

optimization of both KfiA and pmHS2. This sets the stage for further development work in our laboratory.

Unfortunately, we encountered extremely low yield of both of the proteins, which proved to be a hindrance in reaction optimization. The activity of both the enzymes was also found to be low. This in turn would require further optimization of the expression and reaction conditons for the two enzymes. Co-expression with chaperone proteins should be useful in properly folding enzymes, thus displaying more activity as reported by Chen et al.[17] & Xu et al.[16] It would also be useful to utilize the disaccharide isolated from the nitrous acid degradation of heparosan, GlcA-anhydromannitol, as a control to test the activity of the enzymes. This disaccharide has been used as an acceptor for KfiA and subsequent extension of the sugar backbone using pmHS2 and KfiA in tandem for the chemo-enzymatic synthesis of heparin[16, 37].

In conclusion, we have successfully demonstrated the use of pmHS2 for extending on disaccharide acceptors containing unnatural thio linkages along with two unnatural stereochemistry of linkage by MS and MS/MS. To the best of our knowledge, this is the first reported observation of the action of pmHS2 on thio substrates. Owing to the tolerance of the enzyme towards unnatural linkages as observed in this work, and reports that it can accept unnatural UDP-sugars as substrates[15], it is plausible to synthesize novel glycosaminoglycans with this technique whose biological activities would be interesting to study. The approach can also be used to extend on the heparin trisaccharide mimics discussed previously, which can be further modified using C-5 epimerase and sulfotransferases to prepare interesting glycosaminoglycan mimics.

KfiA has also been shown to utilize glucuronic acid and phenyl glucuronide as acceptors, which shows that the enzyme is capable of accepting substrates with unnatural substituents at the

reducing end. This fact can potentially be utilized for analysis and purification later, such as using fluorogenic substrates attached to the reducing end to help in monitoring. Thus, the use of this enzyme for synthesis of interesting carbohydrate functionalized materials can be considered beyond the extension of GAG chains from glucuronic acid, such as extending from substrates immobilized on solid surfaces or nanoparticles through unnatural linkages. Such kind of structures would be useful in mimicking cell surface GAGs, and could give further insight into how they interact with chemokines and fibroblast growth factors for downstream cell signaling.

The next step in this project is to optimize the expression of the two enzymes and use them to build up a small library of model substrates for NMR characterization. This should include both α- and β-thio-*N*-acetylglucosamine-containing acceptors to probe the effect of stereochemistry on enzyme activity. Thereafter, the thio heparin trisaccharide mimics (both sulfated and non-sulfated) can be extended by the tandem use of the two enzymes. Finally with the optimized conditions they can be used to build up defined proteoglycan as well as novel hybrid proteoglycan structures.

4.4 Experimental Section

4.4.1 Expression of KfiA and pmHS2

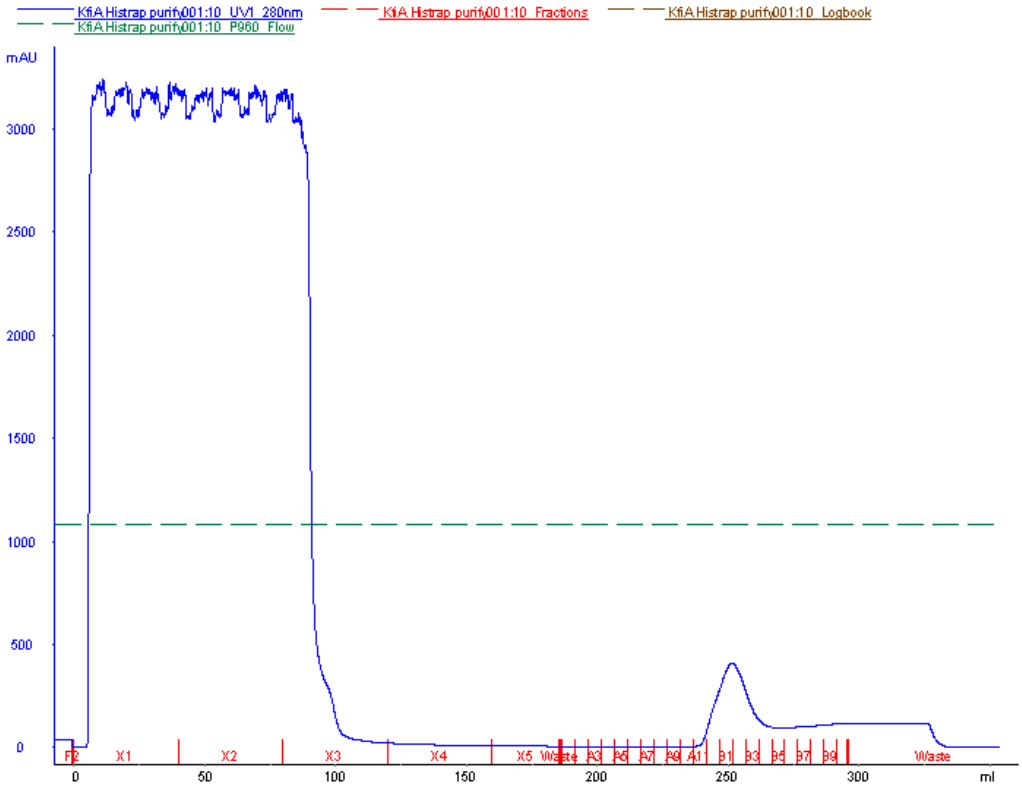
KfiA gene sequence (UniProt Accession No.: Q47332)

ATGATTGTTGCAAATATGTCATCATACCCACCTCGAAAAAAAGAGTTGGTGCATTCT
ATACAAAGTTTACATGCTCAAGTAGATAAAATTAATCTTTGCCTGAATGAGTTTGAA
GAAATTCCTGAGGAATTAGATGGTTTTTCAAAATTAAATCCAGTTATTCCAGATAAA

GATTATAAGGATGTGGGCAAATTTATATTTCCTTGCGCTAAAAATGATATGATCGTA
CTTACAGATGATGATATTATTTACCCTCCCGATTATGTAGAAAAAATGCTCAATTTTT
ATAATTCCTTTGCAATATTCAATTGCATTGTTGGGATTCATGGCTGTATATACATAGA
TGCATTTGATGGAGATCAGTCTAAAAGAAAAGTATTTTCATTTACTCAAGGGCTATT
GCGACCGAGAGTTGTAAATCAATTAGGTACAGGGACTGTTTTTCTTAAGGCAGATCA
ATTACCATCTTTAAAATATATGGATGGTTCTCAACGATTCGTCGATGTTAGATTTTCT
CGCTATATGTTAGAGAATGAAATTGGTATGATATGTGTTCCCAGAGAAAAAAACTG
GCTAAGAGAGGTCTCATCAGGTTCAATGGAAGGACTTTGGAACACATTTACAAAAA
AATGGCCTTTAGACATCATAAAAGAAACACAAGCAATCGCAGGATATTCAAAACTT
AACCTCGAATTAGTGTATAATGTGGA AGGGTAA

KfiA was expressed in pET-21b vector. The whole construct was procured from Genscript. The plasmid was scaled up by Miniprep according to the general procedure described in **Experimental Section** of **Chapter-3**.

Expression of the protein on a large scale was also done according to the general procedure described in **Experimental Section** of **Chapter-3**.



HisTrap purification of KfiA

pmHS2 gene sequence (Uniprot Accession No.: Q5SGE1)

ATGAAGGGAAAAAAAGAGATGACTCAAATTCAAATAGCTAAAAATCCACCCCAACA

TGAAAAAGAAAATGAACTCAACACCTTTCAAAATAAAATTGATAGTCTAAAAACAA

CTTTAAACAAAGACATCATTTCTCAACAAACTTTATTGGCAAAACAGGACAGTAAAC

ATCCGCTATCCGCATCCCTTGAAAACGAAAATAAACTTTTATTAAAACAACTCCAAT

TGGTTCTGCAAGAATTTGAAAAAATATATACCTATAATCAAGCATTAGAAGCAAAG

CTAGAAAAAGATAAGCAAACAACATCAATAACAGATTTATATAATGAAGTCGCTAA

AAGTGATTTAGGGTTAGTCAAAGAAACTAACAGCGCAAATCCATTAGTCAGTATTAT

CATGACATCTCACAATACAGCGCAATTTATCGAAGCTTCTATTAATTCATTATTGTTA

CAAACATATAAAAACATAGAAATTATTATTGTAGATGATGATAGCTCGGATAATAC

ATTTGAAATTGCCTCGAGAATAGCGAATACGACAAGCAAAGTCAGAGTATTTAGAT

TAAATTCAAACCTAGGAACTTACTTTGCGAAAAATACAGGCATATTAAAATCTAAAG

GTGACATTATTTTCTTTCAAGATAGTGATGATGTATGTCATCATGAAAGAATAGAAA

GATGTGTAAATATATTATTAGCTAATAAAGAAACTATTGCTGTTCGTTGTGCATACT

CAAGACTAGCACCAGAAACACAACATATCATTAAAGTCAATAATATGGATTATAGA

TTAGGTTTTATAACCTTGGGTATGCACAGAAAAGTATTTCAAGAAATTGGTTTCTTC

AATTGTACGACTAAAGGCTCAGATGATGAGTTTTTTCATAGAATTGCGAAATATTAT

GGAAAAGAAAAAATAAAAAATTTACTCTTGCCGTTATACTACAACACAATGAGAGA

AAACTCTTTATTTACTGATATGGTTGAATGGATAGACAATCATAACATAATACAGAA

AATGTCTGATACCAGACAACATTATGCAACCCTGTTTCAAGCGATGCATAACGAAAC

TGCCTCACATGATTTCAAAAATCTTTTTCAATTCCCTCGTATTTACGATGCCTTACCA

GTACCACAAGAAATGAGTAAGTTGTCCAATCCTAAGATTCCTGTTTATATCAATATT

TGTTCTATTCCCTCAAGAATAGCGCAATTACGACGTATTATCGGCATACTAAAAAAT

CAATGTGATCATTTTCATATTTATCTTGATGGCTATGTAGAAATCCCTGACTTCATAA

AAAATTTAGGTAATAAAGCAACCGTTGTTCATTGCAAAGATAAAGATAACTCCATTA

GAGATAATGGCAAATTCATTTTACTGGAAGAGTTGATTGAAAAAAATCAAGATGGA

ATTATATAACCTGTGATGATGACATTATCTATCCAAGCGATTACATCAATACGATGA

TCAAGAAGCTGAATGAATACGATGATAAAGCGGTTATTGGTTTACACGGCATTCTCT

TTCCAAGTAGAATGACCAAATATTTTTCGGCGGATAGACTGGTATATAGCTTCTATA

AACCTCTGGAAAAAGACAAAGCGGTCAATGTATTAGGTACAGGAACTGTTAGCTTT

AGAGTCAGTCTCTTTAATCAATTTTCTCTTTCTGACTTTACCCATTCAGGCATGGCTG

ATATCTATTTCTCTCTCTTGTGTAAGAAAAATAATATTCTTCAGATTTGTATTTCAAG
ACCAGCAAACTGGCTAACAGAAGATAATAGAGACAGCGAAACACTCTATCATCAAT
ATCGAGACAATGATGAGCAACAAACTCAGCTGATCATGGAAAACGGTCCATGGGGA
TATTCAAGTATTTATCCATTAGTCAAAAATCATCCTAAATTTACTGACCTTATCCCCT
GTTTACCTTTTTATTTTTTATAA

pmHS2 was expressed in pET-15b vector. The whole construct was procured from Genscript. Corresponding protein sequence, when expressed in pET-15b vector with N-terminal His-Tag:

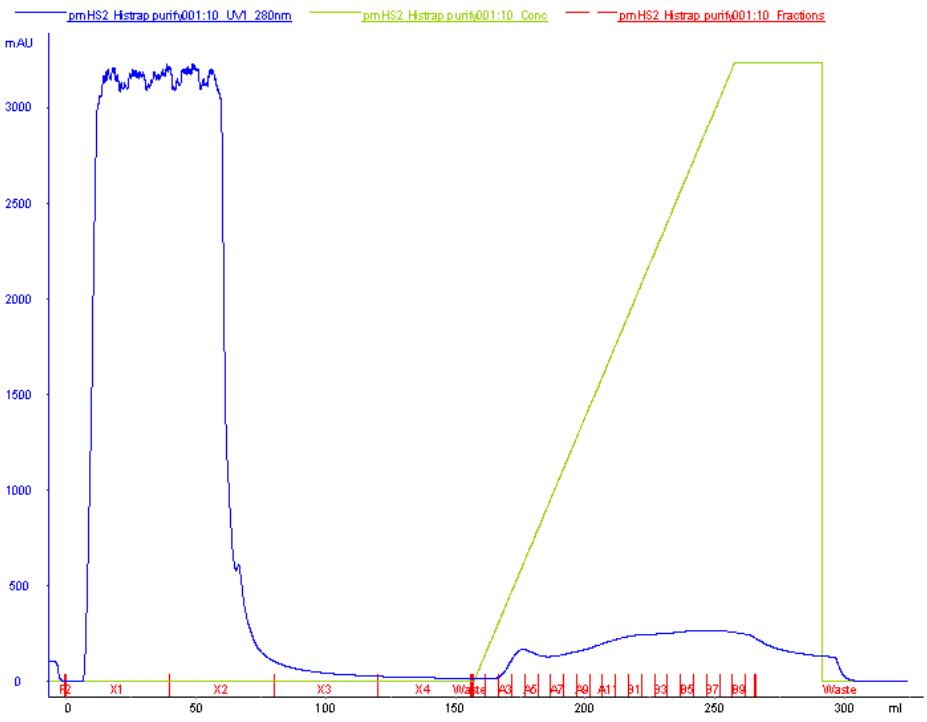
MGSSHHHHHHSSGLVPRGSMKGKKEMTQIQIAKNPPQHEKENELNTFQNKIDSLKTTLN
KDIISQQTLLAKQDSKHPLSASLENENKLLLKQLQLVLQEFEKIYTYNQALEAKLEKDKQ
TTSITDLYNEVAKSDLGLVKETNSANPLVSIIMTSHNTAQFIEASINSLLLQTYKNIEIIIVD
DDSSDNTFEIASRIANTTSKVRVFRLNSNLGTYFAKNTGILKSKGDIIFFQDSDDVCHHER
IERCVNILLANKETIAVRCAYSRLAPETQHIIKVNNMDYRLGFITLGMHRKVFQEIGFFNC
TTKGSDDEFFHRIAKYYGKEKIKNLLLPLYYNTMRENSLFTDMVEWIDNHNIIQKMSDT
RQHYATLFQAMHNETASHDFKNLFQFPRIYDALPVPQEMSKLSNPKIPVYINICSIPSRIA
QLRRIIGILKNQCDHFHIYLDGYVEIPDFIKNLGNKATVVHCKDKDNSIRDNGKFILLEELI
EKNQDGYYITCDDDIIYPSDYINTMIKKLNEYDDKAVIGLHGILFPSRMTKYFSADRLVY
SFYKPLEKDKAVNVLGTGTVSFRVSLFNQFSLSDFTHSGMADIYFSLLCKKNNILQICISR
PANWLTEDNRDSETLYHQYRDNDEQQTQLIMENGPWGYSSIYPLVKNHPKFTDLIPCLP
FYFL

The sequence highlighted in red is from the pET- vector. That consists of the sequence Gly-Ser-Ser (methionine at the beginning has been observed to fall off during expression of the protein),

which is susceptible to *N*-terminal gluconoylation. Also there exists a thrombin cleavage site in the sequence after His-Tag and before the pmHS2 sequence starts. Cleavage of the tag in this position should remove the His-Tag as well as the gluconoylated residue. This can be an additional verification that the second peak is indeed due to *N*-terminal gluconoylation.

pmHS2 plasmid scaling up was done by Miniprep according to the general procedure described in **Experimental Section** of **Chapter-3**.

Expression of the protein on a large scale was also done according to the general procedure described in **Experimental Section** of **Chapter-3**.



HisTrap purification of pmHS2

4.4.2 Structure and CD spectrum prediction of KfiA and pmHS2

CD analysis of the proteins

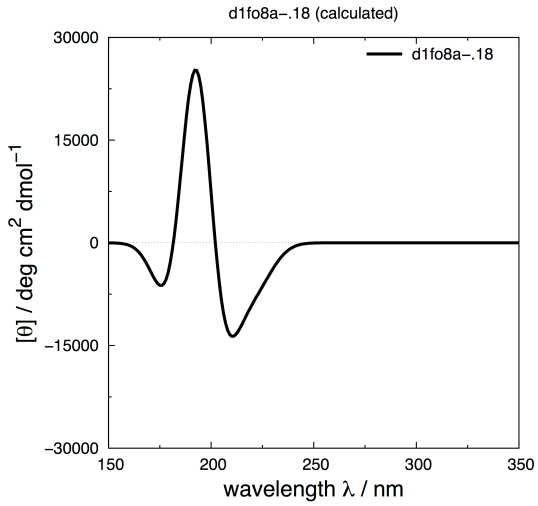
For CD analysis, both KfiA and pmHS2 were buffer exchanged into 10 mM $NaHPO_4$, pH 7.5, 150 mM NaCl. For the measurement 180 µl of protein solution (10 µM) was placed in a micro cuvette (1 mm path length) was used. CD spectra were recorded from 200–300 nm with a step size of 0.5 nm and 1 s measurement per point.

Prediction of CD structure

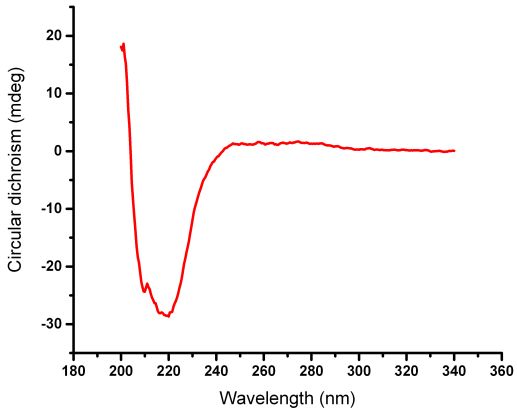
The amino acid sequences of KfiA (UniProt Accession No.: Q47332) and Heparosan synthase B (pmHS2) (UniProt Accession No.: Q5SGE1) were retrieved from the UniProtKB database and were used to generate 3D structural models (in PDB format) using the intensive mode in the Phyre2 webserver[41] PHYRE 2 (Protein Homology/AnalogY Recognition Engine version 2.0) is a widely used webserver that uses homology modelling for structure prediction i.e. constructing the structural model of a query protein sequence (target protein) based on the experimental structure of a homologous protein (template protein).

The predicted structures were then used to calculate the theoretical circular dichroism (CD) spectrum of a protein using the DichroCalc web interface[42] using the default parameters.

For KfiA a homology model belonging to the fold, Nucleotide-diphospho-sugar transferase, Superfamily-Nucleotide-diphospho-sugar transferase and family, *N*-acetylglucosaminyltransferase was chosen to yield the corresponding CD structure, based on similarity in function.



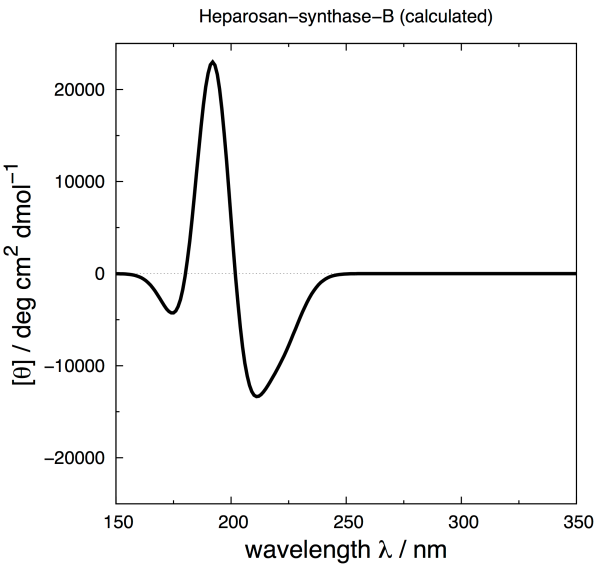
Predicted CD for KfiA



Experimentally obtained CD for KfiA

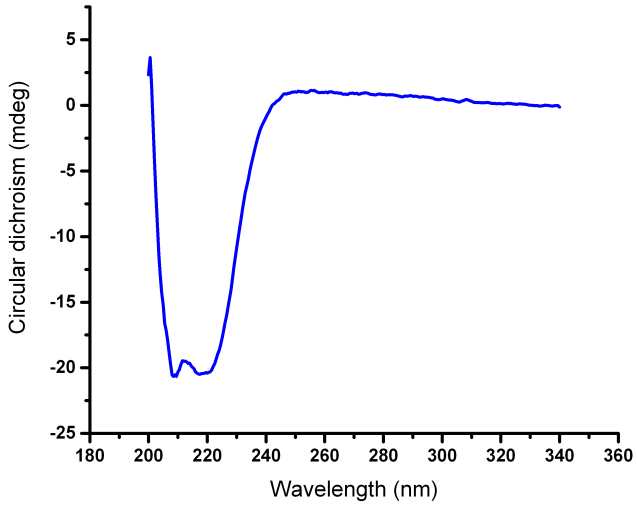
The predicted CD structure is similar to what has been obtained experimentally, with the exception of a few folds. However, since this is a theoretical prediction, not much information can be obtained from this.

Similarly for pmHS2, a homology model with the PDB header as transferase, PDB molecule as the chrondroitin polymerase from *E. coli* strain K4 (K4CP) was chosen to generate the predicted CD structure, based on similarity in function.



Predicted CD structure of pmHS2

For pmHS2 too, even though the predicted structure is similar to the experimentally obtained one, there are some differences in the folds.

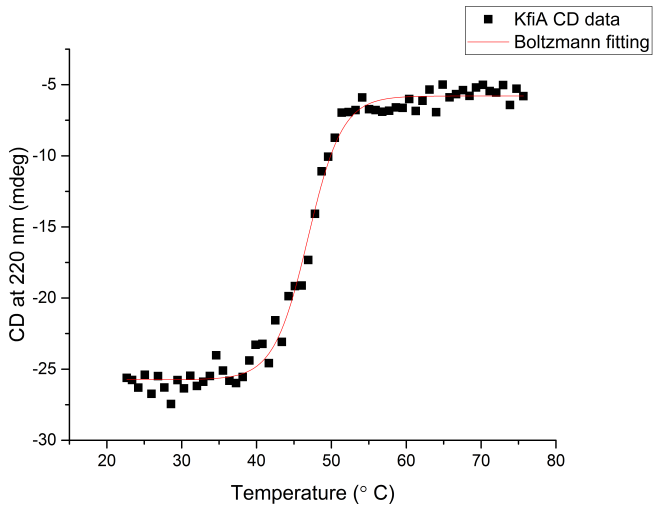


Experimentally obtained CD spectrum of pmHS2

4.4.3 Melting temperature determination of KfiA and pmHS2

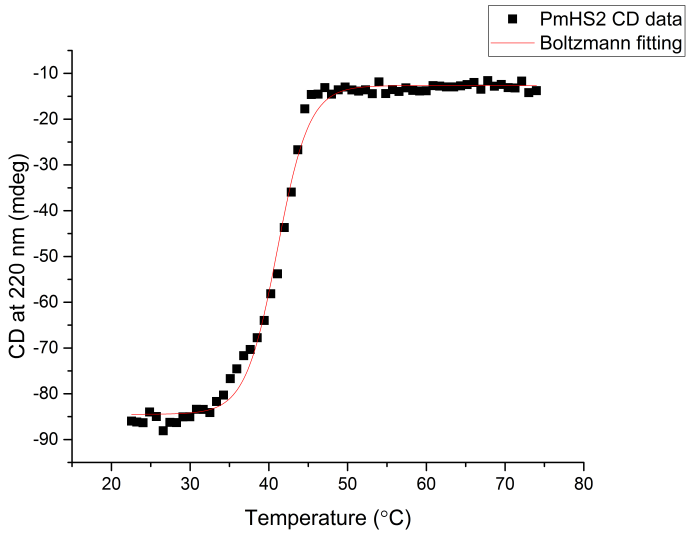
For melting temperature experiments, the protein was dialysed into phosphate buffer as mentioned before. CD spectra were recorded from 200–300 nm with a step size of 0.5 nm and 1 s measurement per point. Temperature was ramped stepwise from 25–85 °C in 1 °C steps with 120 s at each temperature.

Melting temperature curve for KfiA:



The melting temperature was found to be 46.87 +/- 0.16 °C

Melting temperature curve for pmHS2:



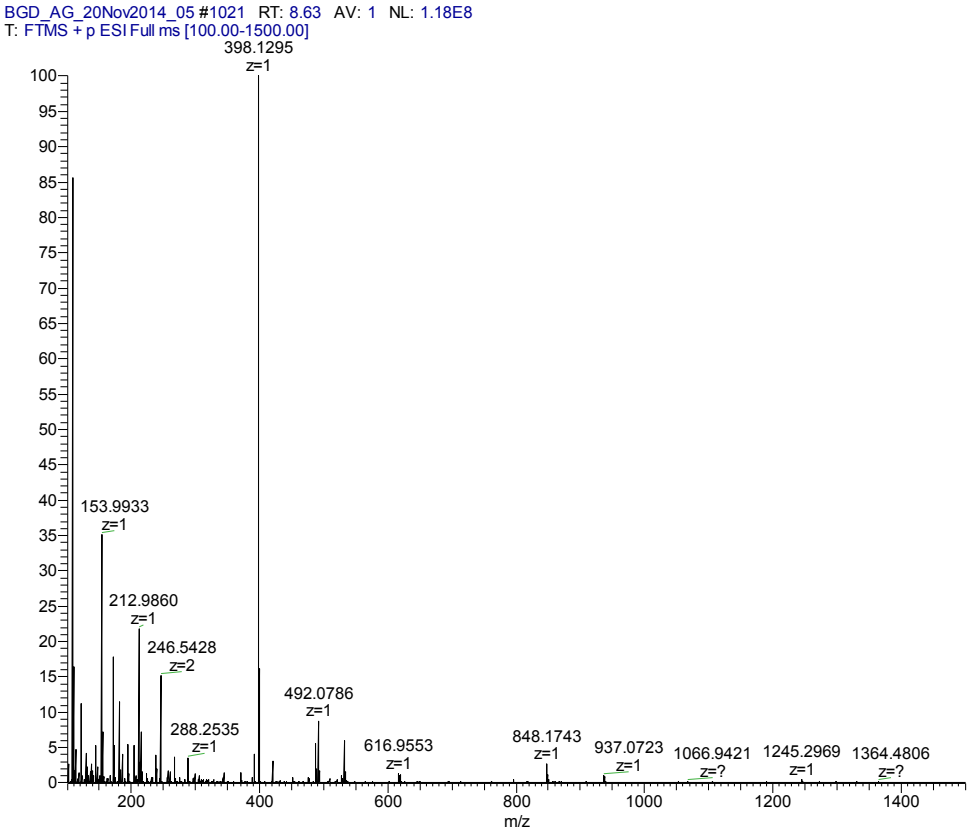
Melting temperature of the protein was found to be 41.07 +/- 0.11 °C.

4.4.4 MS and MS/MS data for KfiA and pmHS2 reactions

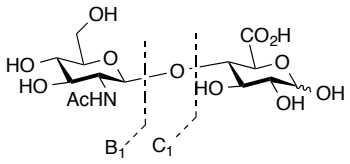
Reactions with KfiA

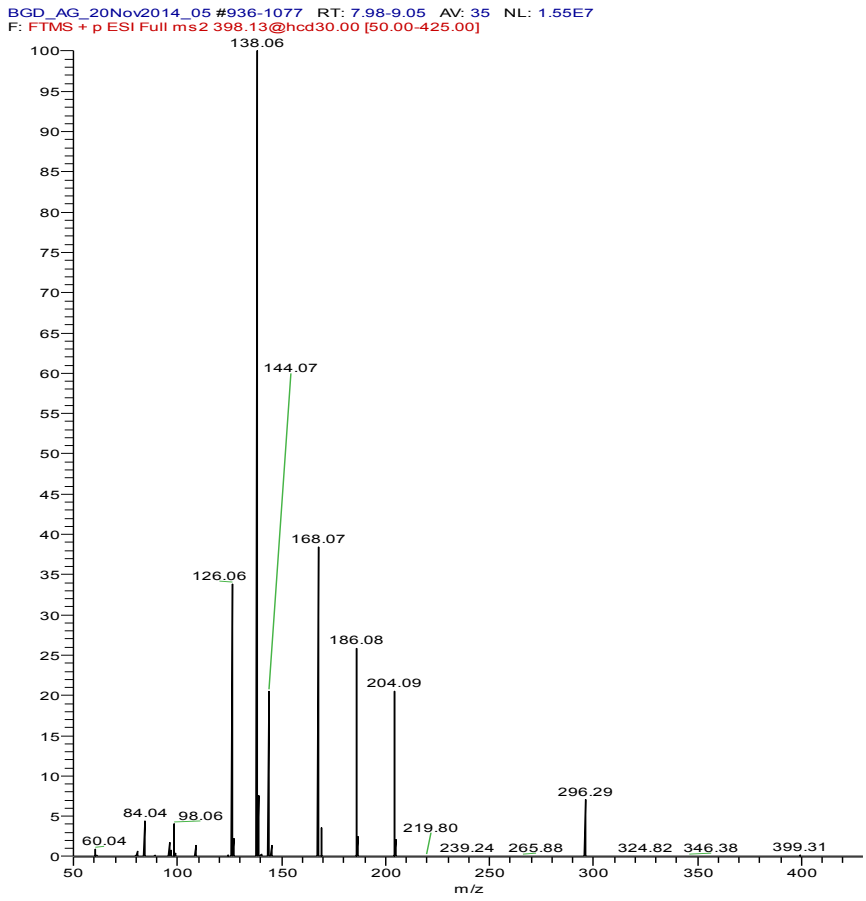
Mass predicted for compound **4** is 398.1299 (M+H) (in positive mode)

Found = 398.1295



Subsequent MS/MS of this peak was done.

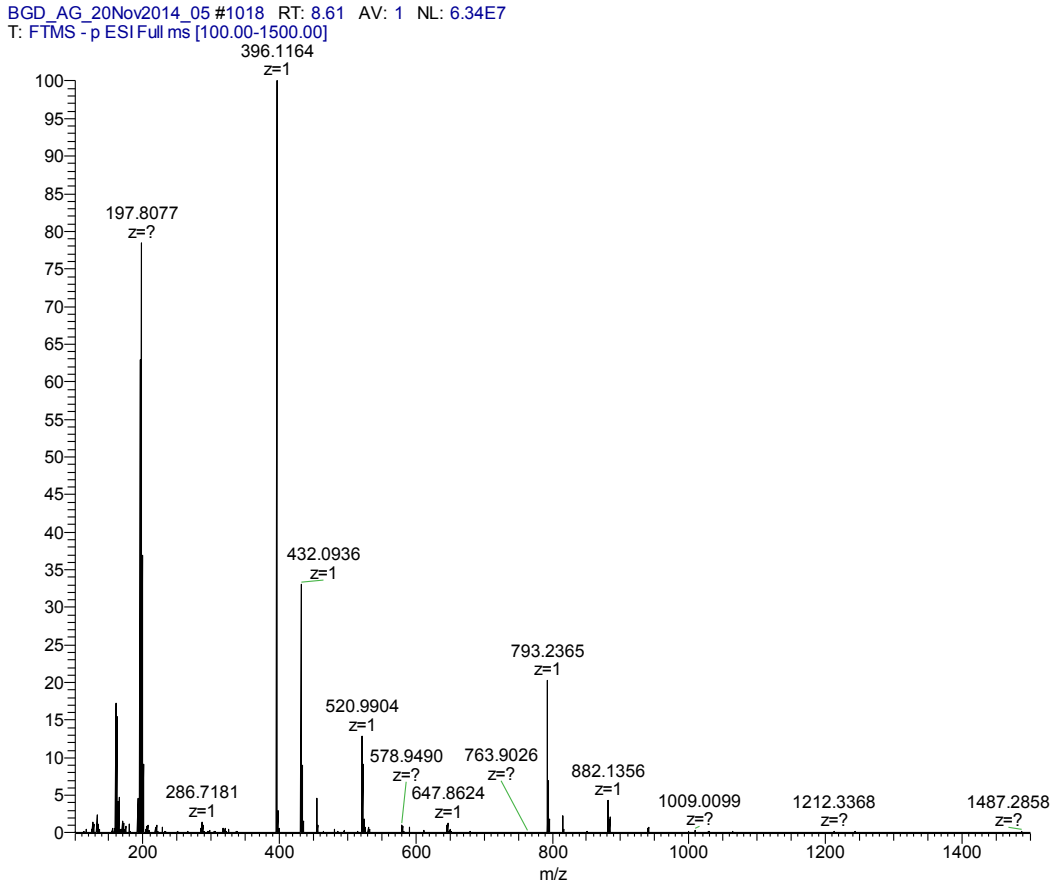




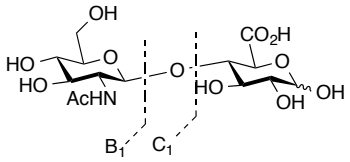
$B_1 - 2H_2O = 168.07$, $B_1 - H_2O = 186.08$, $B_1 = 204.09$ (Nomenclature of fragments done according to the Domon and Costello proposed method)[43]

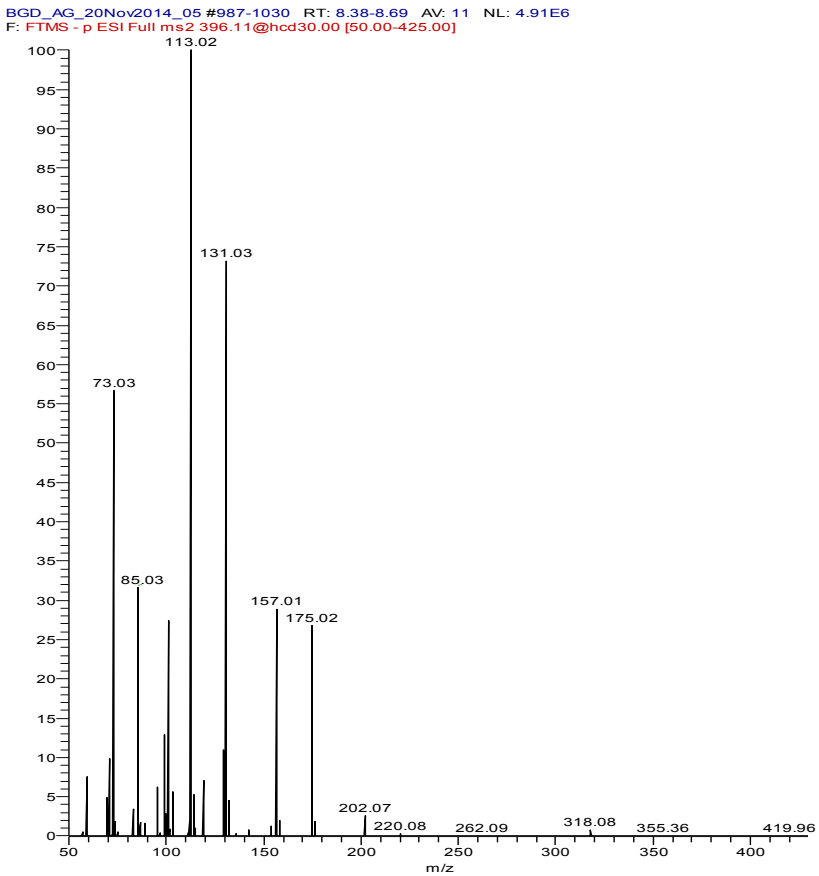
Expected mass of compound **4** = 396.1142 (M-H, in negative mode)

Found = 396.1164



MS/MS for this peak was done:



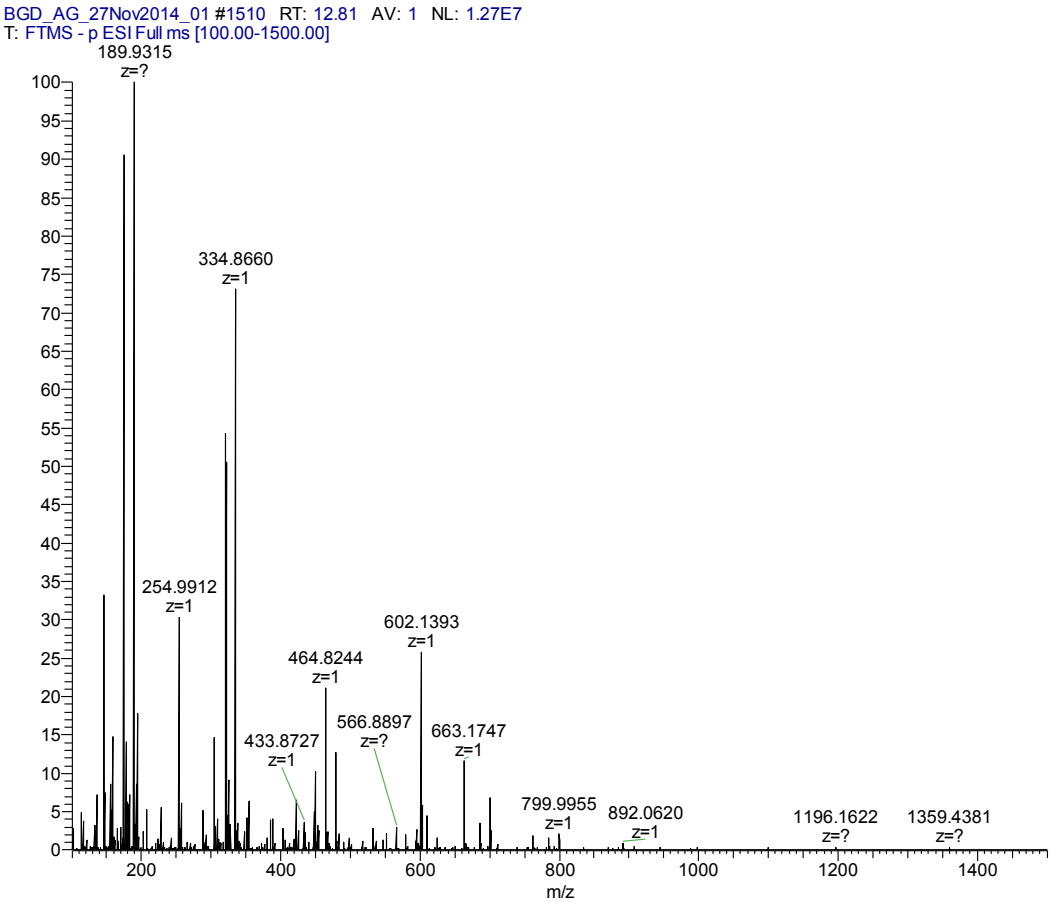


C_1-H_2O = 202.07

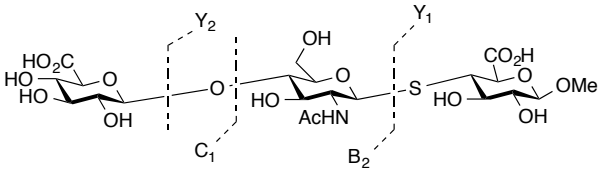
Reactions with pmHS2

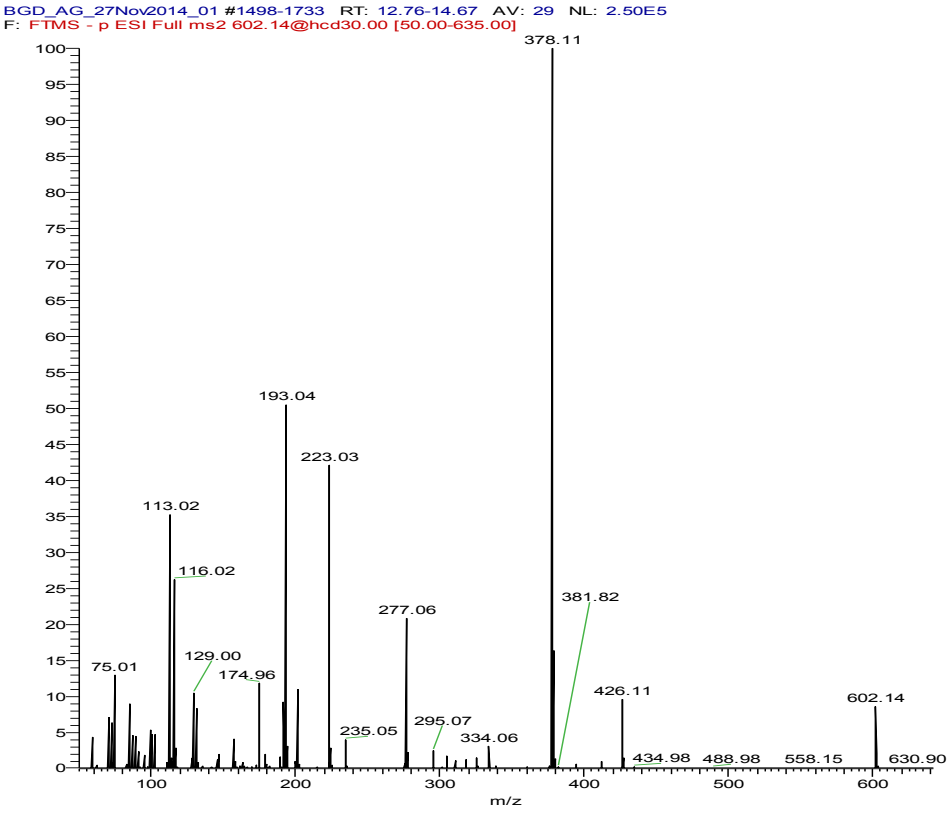
Predicted mass of compound **2** = 602.1391 (M-H, negative mode)

Found = 602.1393



MS/MS of this peak was done.





C_1 = 193.04, Y_1 = 223.03, B_2 = 378.11, Y_2 = 426.11

4.4.5 Reaction procedure of KfiA and pmHS2

Before the reaction, the protein (KfiA or pmHS2) was dialysed against the reaction buffer (25 Tris, 150 mM NaCl, pH -7) to remove glycerol. Thereafter typical reactions were set up in 1 ml aliquots. The acceptor and UDP-sugar was added to a final concentration of 1mM. $MnCl_2$ was added to a final concentration of 10 mM. Thereafter the protein was added to the required concentration (typically 1 mg/ml final concentration and the rest of the volume was filled up with the reaction buffer. The reaction mixture was set up in the thermoshaker and progress of the reaction was monitored at regular intervals by running a HPLC to check for the conversion of UDP-sugar to UDP. At the end of the reaction, 10% acetonitrile was added to precipitate out the protein, while the supernatant was lyophilized to obtain the crude reaction mixture.

4.5 References

1. Sugahara, K.; Kitagawa, H. Heparin and heparan sulfate biosynthesis. *IUBMB Life* **2002,** *54* (4), 163-175.

2. Sugahara, K.; Kitagawa, H. Recent advances in the study of the biosynthesis and functions of sulfated glycosaminoglycans. *Curr. Opin. Struc. Biol.* **2000,** *10* (5), 518-527.

3. Prydz, K.; Dalen, K. T. Synthesis and sorting of proteoglycans - Commentary. *J. Cell Sci.* **2000,** *113* (2), 193-205.

4. Chavaroche, A. A. E.; van den Broek, L. A. M.; Eggink, G. Production methods for heparosan, a precursor of heparin and heparan sulfate. *Carbohydr. Polym.* **2013,** *93* (1), 38-47.

5. Habuchi, O. Diversity and functions of glycosaminoglycan sulfotransferases. *BBA-Gen. Subjects* **2000,** *1474* (2), 115-127.

6. Kjellen, L. Glucosaminyl N-deacetylase/N-sulphotransferases in heparan sulphate biosynthesis and biology. *Biochem. Soc. Trans.* **2003,** *31*, 340-342.

7. Kusche-Gullberg, M.; Kjellen, L. Sulfotransferases in glycosaminoglycan biosynthesis. *Curr. Opin. Struc. Biol.* **2003,** *13* (5), 605-611.

8. Li, J. P.; Gong, F.; El Darwish, K.; Jalkanen, M.; Lindahl, U. Characterization of the D-glucuronyl C5-epimerase involved in the biosynthesis of heparin and heparan sulfate. *J. Biol. Chem.* **2001,** *276* (23), 20069-20077.

9. Tsuruta, O.; Yuasa, H.; Hashimoto, H.; Sujino, K.; Otter, A.; Li, H.; Palcic, M. M. Synthesis of GDP-5-thiosugars and their use as glycosyl donor substrates for glycosyltransferases. *J. Org. Chem.* **2003,** *68* (16), 6400-6406.

10. Nishio, T.; Hoshino, S.; Kondo, A.; Ogawa, M.; Matsuishi, Y.; Kitagawa, M.; Kawachi, R.; Oku, T. α-mannosidase-catalyzed synthesis of a (1->2)-α-D-rhamnodisaccharide derivative. *Carbohydr. Res.* **2004,** *339* (7), 1389-1393.

11. Renaudie, L.; Daniellou, R.; Auge, C.; Le Narvor, C. Enzymatic supported synthesis of lacto-N-neotetraose using dendrimeric polyethylene glycol. *Carbohydr. Res.* **2004,** *339* (3), 693-698.

12. Sarrazin, S.; Lamanna, W. C.; Esko, J. D. Heparan Sulfate Proteoglycans. *CSH Perspect. Biol.* **2011,** *3* (7).

13. Kreuger, J.; Kjellen, L. Heparan Sulfate Biosynthesis: Regulation and Variability. *J. Histochem. Cytochem.* **2012,** *60* (12), 898-907.

14. Hacker, U.; Nybakken, K.; Perrimon, N. Heparan sulphate proteoglycans: The sweet side of development. *Nat. Rev. Mol. Cell Bio.* **2005,** *6* (7), 530-541.

15. Sismey-Ragatz, A. E.; Green, D. E.; Otto, N. J.; Rejzek, M.; Field, R. A.; DeAngelis, P. L. Chemoenzymatic synthesis with distinct Pasteurella heparosan synthases - Monodisperse polymers and unnatural structures. *J. Biol. Chem.* **2007,** *282* (39), 28321-28327.

16. Xu, Y.; Masuko, S.; Takieddin, M.; Xu, H.; Liu, R.; Jing, J.; Mousa, S. A.; Linhardt, R. J.; Liu, J. Chemoenzymatic Synthesis of Homogeneous Ultralow Molecular Weight Heparins. *Science* **2011,** *334* (6055), 498-501.

17. Chen, M.; Bridges, A.; Liu, J. Determination of the substrate specificities of N-acetyl-D-glucosaminyltransferase. *Biochemistry* **2006,** *45* (40), 12358-12365.

18. Li, Y. H.; Yu, H.; Thon, V.; Chen, Y.; Muthana, M. M.; Qu, J. Y.; Hie, L. N.; Chen, X. Donor substrate promiscuity of the N-acetylglucosaminyltransferase activities of Pasteurella multocida heparosan synthase 2 (PmHS2) and Escherichia coli K5 KfiA. *Appl. Microbiol. Biotechnol.* **2014,** *98* (3), 1127-1134.

19. Stathopulos, P. B.; Scholz, G. A.; Hwang, Y. M.; Rumfeldt, J. A. O.; Lepock, J. R.; Meiering, E. M. Sonication of proteins causes formation of aggregates that resemble amyloid. *Protein Sci.* **2004,** *13* (11), 3017-3027.

20. Lopez, P. J.; Marchand, I.; Joyce, S. A.; Dreyfus, M. The C-terminal half of RNase E, which organizes the Escherichia coli degradosome, participates in mRNA degradation but not rRNA processing in vivo. *Mol. Microbiol.* **1999,** *33* (1), 188-199.

21. Grunberg-Manago, M. Messenger RNA stability and its role in control of gene expression in bacteria and phages. *Annu. Rev. Genet.* **1999,** *33*, 193-227.

22. Carpousis, A. J. The RNA degradosome of Escherichia coli: An mRNA-degrading machine assembled on RNase E. *Annu. Rev. Microbiol.* **2007,** *61*, 71-87.

23. Francis, D. M.; Page, R. Strategies to Optimize Protein Expression in E. coli. In *Current Protocols in Protein Science*, John Wiley & Sons, Inc.: 2001.

24. Baneyx, F.; Mujacic, M. Recombinant protein folding and misfolding in Escherichia coli. *Nat. Biotechnol.* **2004,** *22* (11), 1399-1408.

25. Kraft, C.; Peter, M.; Hofmann, K. Selective autophagy: ubiquitin-mediated recognition and beyond. *Nat. Cell Biol.* **2010,** *12* (9), 836-841.

26. Sugiura, N.; Baba, Y.; Kawaguchi, Y.; Iwatani, T.; Suzuki, K.; Kusakabe, T.; Yamagishi, K.; Kimata, K.; Kakuta, Y.; Watanabe, H. Glucuronyltransferase Activity of KfiC from Escherichia coli Strain K5 Requires Association of KfiA KfiC and KfiA Are Essential Enzymes For Production Of K5 Polysaccharide, N-Acetylheparosan. *J. Biol. Chem.* **2010,** *285* (3), 1597-1606.

27. Chavaroche, A. A. E.; van den Broek, L. A. M.; Springer, J.; Boeriu, C.; Eggink, G. Analysis of the Polymerization Initiation and Activity of Pasteurella multocida Heparosan Synthase PmHS2, an Enzyme with Glycosyltransferase and UDP-sugar Hydrolase Activity. *J. Biol. Chem.* **2011,** *286* (3), 1777-1785.

28. Volonte, F.; Marinelli, F.; Gastaldo, L.; Sacchi, S.; Pilone, M. S.; Pollegioni, L.; Molla, G. Optimization of glutaryl-7-aminocephalosporanic acid acylase expression in E-coli. *Protein Expres. Purif.* **2008,** *61* (2), 131-137.

29. Shirano, Y.; Shibata, D. Low-Temperature Cultivation of Escherichia-Coli Carrying a Rice Lipoxygenase L-2 cDNA Produces a Soluble and Active Enzyme at a High-Level. *FEBS Lett.* **1990,** *271* (1-2), 128-130.

30. Hunke, S.; Betton, J. M. Temperature effect on inclusion body formation and stress response in the periplasm of Escherichia coli. *Mol. Microbiol.* **2003,** *50* (5), 1579-1589.

31. Geoghegan, K. F.; Dixon, H. B. F.; Rosner, P. J.; Hoth, L. R.; Lanzetti, A. J.; Borzilleri, K. A.; Marr, E. S.; Pezzullo, L. H.; Martin, L. B.; LeMotte, P. K.; McColl, A. S.; Kamath, A. V.; Stroh, J. G. Spontaneous alpha-N-6-phosphogluconoylation of a "His tag" in Escherichia coli: The cause of extra mass of 258 or 178 Da in fusion proteins. *Anal. Biochem.* **1999,** *267* (1), 169-184.

32. Sasaki, E.; Zhang, X.; Sun, H. G.; Lu, M. Y. J.; Liu, T. L.; Ou, A.; Li, J. Y.; Chen, Y. H.; Ealick, S. E.; Liu, H. W. Co-opting sulphur-carrier proteins from primary metabolic pathways for 2-thiosugar biosynthesis. *Nature* **2014,** *510* (7505), 427-431.

33. Aon, J. C.; Caimi, R. J.; Taylor, A. H.; Lu, Q.; Oluboyede, F.; Dally, J.; Kessler, M. D.; Kerrigan, J. J.; Lewis, T. S.; Wysocki, L. A.; Patel, P. S. Suppressing posttranslational gluconoylation of heterologous proteins by metabolic engineering of Escherichia coli. *Appl. Environ. Microb.* **2008,** *74* (4), 950-958.

34. Kim, K. M.; Yi, E. C.; Baker, D.; Zhang, K. Y. J. Post-translational modification of the N-terminal His tag interferes with the crystallization of the wild-type and mutant SH3 domains from chicken src tyrosine kinase. *Acta Crystallogr. D* **2001,** *57*, 759-762.

35. Beranek, M.; Drsata, J.; Palicka, V. Inhibitory effect of glycation on catalytic activity of alanine aminotransferase. *Mol. Cell. Biochem.* **2001,** *218* (1-2), 35-39.

36. Sakurai, T.; Matsuyama, M.; Tsuchiya, S. Glycation of Erythrocyte Superoxide-Dismutase Reduces Its Activity. *Chem. Pharm. Bull.* **1987,** *35* (1), 302-307.

37. Liu, R. P.; Xu, Y. M.; Chen, M. A.; Weiwer, M.; Zhou, X. X.; Bridges, A. S.; DeAngelis, P. L.; Zhang, Q. S.; Linhardt, R. J.; Liu, J. A. Chemoenzymatic Design of Heparan Sulfate Oligosaccharides. *J. Biol. Chem.* **2010,** *285* (44), 34240-34249.

38. Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. Glycosyltransferases: Structures, functions, and mechanisms. *Annu. Rev. Biochem* **2008,** *77*, 521-555.

39. Qasba, P. K.; Ramakrishnan, B.; Boeggeman, E. Substrate-induced conformational changes in glycosyltransferases. *Trends Biochem. Sci.* **2005,** *30* (1), 53-62.

40. Zhang, Z. Q.; Xie, J.; Zhang, F. M.; Linhardt, R. J. Thin-layer chromatography for the analysis of glycosaminoglycan oligosaccharides. *Anal. Biochem.* **2007,** *371* (1), 118-120.

41. Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015,** *10* (6), 845-858.

42. Bulheller, B. M.; Hirst, J. D. DichroCalc-circular and linear dichroism online. *Bioinformatics* **2009,** *25* (4), 539-540.

43. Domon, B.; Costello, C. E. A Systematic Nomenclature for Carbohydrate Fragmentations in FAB-MS/MS Spectra of Glycoconjugates. *Glycoconjugate J.* **1988,** *5* (4), 397-409.

Chapter 5

Conclusion

Chapter – 5 Conclusion

In conclusion we have established an efficient thiol-ene radical approach for synthesis of *S*-linked disaccharides. Apart from protected sugars, deprotected carbohydrates can also be directly coupled in aqueous conditions. This constitutes the first example of thiol–ene mediated glycosylation between two carbohydrate moieties in aqueous conditions, to the best of our knowledge. This approach can enable us to synthesize potentially biologically relevant substrates which are resistant to enzymatic degradation at the same time, because of the thio-linkage. With the help of other thiosugar epimers, the scope for these reactions can be further extended to mimic more biologically relevant molecules.

We have also succeeded in using the thiol-ene radical reaction to couple thiosugars with heparin- and heparan sulfate (HS)-derived oligosaccharides. This serves as a proof of concept towards building larger glycosaminoglycan oligosaccharides. Importantly, the reactions could be carried out in completely aqueous conditions, thus enabling coupling reactions to be carried out directly on heparin and HS derived oligosaccharides without the need for protective group manipulation. Also, the labile sulfate groups of the heparin carbohydrates were found to be stable in the reaction conditions used. This approach can also be utilized for tagging of the non-reducing end of the glycosaminoglycan oligosaccharides by various molecules such as fluorophores etc. for diagnostic purposes. Thus, apart from the reducing end of the oligosaccharide, which can be accessed by means of Schiff base formation, our developed strategy also enables accessing the non-reducing end for tagging purposes. This can be useful in instances where it is known that the reducing end of the oligosaccharide is involved in direct interaction with its binding partners and therefore modifying the reducing end can hamper the interaction.

Interestingly, such an approach of thiol-ene radical reaction addition to the unsaturation of the oligosaccharide has already been pursued, albeit unsuccessfully by the Nitz group as part of their continuing endeavor towards developing protecting group free glycosylation strategies[1]. In this unpublished work, as a thesis, the author reports the addition of *N*-acetylcysteamine to an unsaturated model monosaccharide. They have reported the isolation of only one isomer in this case in quite good yields. Specifically, a galacturonic configuration of the product was observed by them. When the reactions were tried on a chondroitin sulfate building block, with an unsaturation at the non-reducing end, they could not obtain the desired product even with different reaction conditions and photoinitiators. However, when increased thiol concentration was used, they reported the formation of several side products in the reaction mixture. These could not be characterized, as the author was unable to separate the products. Use of Michael acceptors by esterifying the carboxylic group or use of electrophiles for attack on the unsaturation did not increase the yield.

The reversal of enzymatic activity by mutagenesis has been reported in some cases especially in the context of Endo enzymes[2]. It would be interesting, therefore to perform site directed mutagenesis on Heparinase-1 in an attempt to induce a Michael type reaction. Specifically this might involve reversing the acidic and basic amino acid residues in the active site, which are involved in the normal reaction, in the first instance (**Figure 1 and 2**). More specifically, if thio substrates can be accepted by the mutated enzyme, it would lead to the formation of stable thio linkages which should not be cleaved by the enzyme's forward activity.

Unfortunately, larger oligosaccharides from heparin could not be used for thiol-ene radical reaction because of the deactivating influence of the sulfate groups. Binding studies with the synthesized small mimics did not show any measurable binding activity. Therefore the enzymatic extension of the thio mimics for definite binding activity was sought.

As part of the enzymatic approach, we expressed two bacterial enzymes to attempt a biocatalytic extension of heparin thio mimics. We showed for the first time by MS and MS/MS conclusively, that the glucuronyl transferase involved in enzymatic backbone synthesis of heparin, can also accept thio analogs, even with unnatural stereochemistry. However, the yield of this reaction could not be increased sufficiently to allow NMR characterization of the product. It is possible that with correct folding of the enzyme, the activity, and thereby the yield will increase. The enzymatic approach coupled with the thiol-ene radical reaction can also be utilized to build defined proteoglycans and even novel hybrid proteoglycan structures. This sets the stage for further development work in our laboratory.

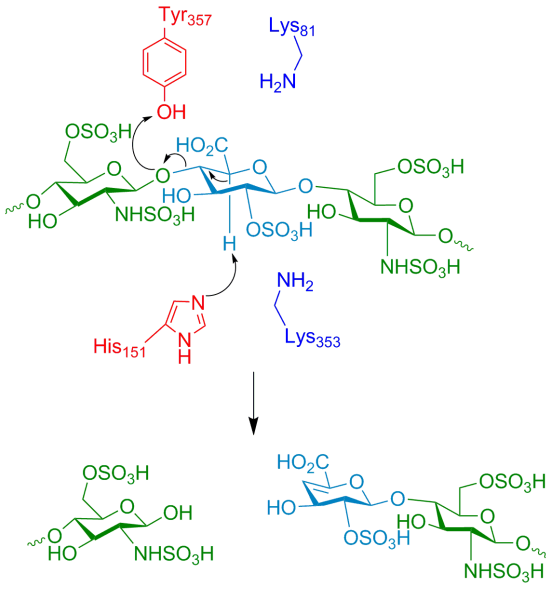


Figure 1: Mechanism of action of wild type Heparinase-1

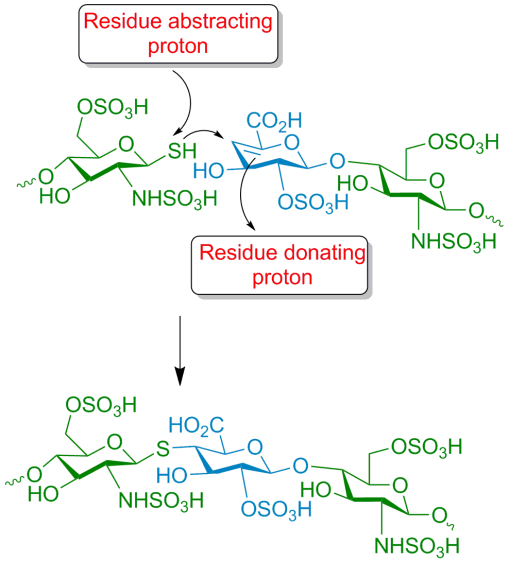


Figure 2: Proposed mutagenesis study on the enzyme to impart a Michael addition type reaction

References

1. Gudmundsdottir, V. A. Protecting Group-Free Chemical Modifications on Carbohydrates. University of Toronto, Toronto, 2009.

2. Fan, S. Q.; Huang, W.; Wang, L. X., Remarkable Transglycosylation Activity of Glycosynthase Mutants of Endo-D, an Endo-beta-N-acetylglucosaminidase from Streptococcus pneumoniae. *J. Biol. Chem.* **2012,** *287* (14), 11272-11281.