

# **The Effect of Epimers, Glycosidic Linkage and the Sulfation Pattern of Glycosaminoglycan Components in Glycomics Research**

**A Thesis**

**Submitted in partial fulfilment of the requirements**

**for the degree of  
Doctor of Philosophy**

**By  
Balamurugan Subramani  
20143334**

**Under the guidance of  
Dr. Raghavendra Kikkeri**



**INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH  
PUNE**

**2022**

*I would like to dedicate my thesis to everyone in the world*  
(Who are supporting me directly or indirectly)



தமிழ்: யாதும் ஊரே யாவரும் கேளிர் - கணியன் பூங்குன்றனார்

English: All the places on earth are our town and all the people are our relatives.

*Kaniyan Poongunranar*

# **CERTIFICATE**

This is to certify that the work incorporated in this thesis entitled “**The Effect of Epimers, Glycosidic Linkage and the Sulfation Pattern of Glycosaminoglycan Components in Glycomics Research**” submitted by **Balamurugan Bubramani** was carried out by candidate at Indian Institute of Science Education and Research, Pune under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other university or institution.

Date: 31/05/2022



**Dr. Raghavendra V. Kikkeri**  
Professor  
IISER, Pune

## **DECLARATION**

I hereby declare that the thesis entitled “**The Effect of Epimers, Glycosidic Linkage and the Sulfation Pattern of Glycosaminoglycan Components in Glycomics Research**” submitted for Doctor of Philosophy in Chemistry at Indian Institute of Science Education and Research, Pune, has not been submitted by me to any other university or institution. This work presented here was carried out at the, Indian Institute of Science Education and Research, Pune, India under the supervision of **Dr. Raghavendra Kikkeri**.

Date: 31/05/2022



**Balamurugan S**  
**20142014**



## **Acknowledgements**

This is the hardest thing to articulate heartfully who came directly or indirectly into my life to encourage and support me without any expectations. The words won't be satiated to address all the people. First and utmost gratitude to my thesis supervisor Prof. Raghavendra V. Kikkeri for giving me an opportunity to work in the lab, in the field of glycobiology and also interdisciplinary research. Without his help, my thesis would not be completed. He is there with me throughout the up and down of this Ph. D journey with his continuous support and motivated me during defeats. He is an exceptionally amiable and exalted thinking person toward research, he is highly devoted and passionate about science.

I like to thank my RAC members, Prof. H. N. Gopi, from IISER Pune, and Dr. Udaya Kiran Marelli, from NCL Pune, for their constructive suggestions and motivations during the meeting. I would also like to thank my collaborator Dr. Dhanasekaran Shanmugam, NCL, Pune for his time and efforts. I would like to thank Prof. Jayant B. Udgaonkar (Director, IISER Pune) and Prof. K. N. Ganesh (Former director, IISER Pune) for the excellent infrastructure. My sincere thanks to Prof. H.N. Gopi, chair of chemistry, and Prof. M. Jayakannan (Former chair of chemistry) for providing outstanding research facilities. I also like to thank all the faculty members of IISER Pune and the technical team and instrument operators of IISER Pune. My worm thank to CSIR for the five-year fellowship and Scivic Engineering India Pvt Ltd and Innoplexus Consulting Services Pvt Ltd for the short time financial support.

I am an extremely blessed person to have exceptionally good colleagues, in the absence of their teaching and sharing pieces of knowledge my Ph.D. will not be finished. I have been pleased to have labmates like Rohan, Harikrishna, Sivakoti, Madhuri (Kaka), and Raghvendra. They had trained me as a carbohydrate chemist and friendly approach without any hesitance they assisted. Special thanks to Preeti mam for her endless support and also as a mentor to teach all the technical parts as well as biological aspects. Very big thanks to Chathan, Suraj, Prashant, Trimbak, and Vijendra for sharing their knowledge and also being wonderful companions on this journey. Thanks to loquacious lady Sandhya for fighting with me and reconciling immediately as a nice buddy. Immense thanks to my young battler Rakesh, Saurabh, Deepak, and Ankita for their support. Thanks to Sharath and Virendra for their incredible support. Great thanks to BSMS lab alumni Phani, Cathrine, Akil, Keerthana, Haritha, and current members Prasanna,

Remya, and Dipti. A huge thanks to the entire RK group for the extraordinary expeditions, parties, and get-together and it is an unforgettable memory in my life.

My sincere thanks to Ashok, Sanjit, Ganesh, Shatrughan, Mohan, Pavan, Chandan, Ajay, Udaya, and Ph.D. batch 2014 friends in IISER Pune for their cheerful companionship, they made me very comfortable and joyful movements.

A special thanks to my intimates Dharmaraja, Anantharaj, Manikandan, Manoharan, Suresh Kumar (IT), Ravikumar, Sankar, Vijayakanth, Alagar, Veeresh, Siva, and Thamarai for their debated and gleeful memories with them. Thanks to the Alchemist team for giving me delighting moments to refresh the weekends. I soulful thanks to all the housekeepers and all the worker in IISER Pune for keeping campus tidy and security guards their friendly approach.

My deepest gratitude to my senior Ponnurangan and Karupasamy for their kindness and pushing me to Pune for the interview and for supporting me in all the possible ways. A special thanks to Mohammed Imran Khan, who helped me a lot while working in Bangalore.

I would like to express my profound and heartfelt gratitude to my M.Sc project supervisor, Dr. K. Srinivasan, Assistant Professor, School of Chemistry, Bharathidasan University, Tiruchirappalli. He had engaged me to qualify for the CSIR exam. My special thanks to my senior Sathishkannan, who is the starting point of my research career, and also like to thank Sakthivel, Akbar, and Selvi, Research scholars for their cooperation during my project. My immense pleasure is to express heartfelt thanks to my dear friends and M.Sc classmates Srinivasan, Ramamoorthi, Govindharaj, Balaji, Arjun, Manibalan, Yuvaraj, Sakthinathan, Radhika and Nirmala a for their support during the M.Sc course.

My profound thanks to many writers who gave me enormous confidence in life during the downfall. I deeply believe that the Books will dismiss my ignorance. Thanks to all the authors who gave positive thoughts in their writing to tackle any troubles and problems which motivate me to live with the commune.

My earnest thanks to my cousin G. Gopi Assistant Professor, Department of Physics, Govt. Arts College, for endless encouragement and his assistance for my entire career.

I would like to thank my beloved parents Subramani and Indirani, brother Sakthivel, sister Suruthi and family members for their love and devotion to supporting me all the period, without their dedication it would not be possible to pursue my degree.

Finally, I don't have the proper words to express my gratitude to all the ancestors for the revolution, wisdom, enlightenment, blessings and knowledge that transferred us to improve the qualities of life. My genuine thank to Nature for the benevolent environment which keeps me very pleasant.

*With love*

*Balamurugan ெ*

## Table of Contents

Abbreviations .....	i
Abstract .....	vi
List of Publications .....	viii

### **Chapter-1: Driving Force Behind Specific Carbohydrate-protein Interactions**

Abstract: .....	1
1. 1 Introduction: .....	2
1.2 Glycocalyx: .....	2
1.3 N-glycan and O-glycan: .....	3
1.4 Glycosaminoglycans: .....	3
1.5 Discovery of Antiviral Heparan Sulfate Mimics: .....	8
1.6 Carbohydrate-protein interactions: .....	10
1.7 Glyconanoparticles: .....	11
1.8 Glycopolymers: .....	12
1.9 Liposomes: .....	12
1.10 Conclusion. ....	13
1.10 References: .....	14

### **Chapter-2: Demystifying a Hexuronic Acid Ligand that Recognizes *Toxoplasma Gondii* and Blocks its Invasion Into Host Cells**

Abstract: .....	22
2.1 Introduction: .....	23
2.2 Results and discussion: .....	24
2.2.1 Synthesis of D-glucuronic acid derivatives: .....	24
2.2.2 Synthesis of L-iduronic acid derivatives: .....	25
2.2.3 Functionalization of Hexuronic Acids on Glass slides: .....	25
2.2.4 XPS analysis: .....	26
2.3 Conclusions: .....	33
2.4 Experimental section: .....	33
2.4.1 General information: .....	33
2.4.2 Immobilization of hexuronic acids on glass slides: .....	38
2.4.3 XPS analysis: .....	38
2.4.4 Quantification of sugar on glass slides: .....	38

2.4.5 Gelatin conjugation: . . . . .	39
2.4.6 Toxoplasma gondii culture protocol: . . . . .	39
2.4.7 Binding assays with <i>T. gondii</i> : . . . . .	39
2.4.8 Invasion assays: . . . . .	40
2.5 References: . . . . .	41
2. 6 NMR Spectra: . . . . .	44

## **Chapter-3: A Cell-Culture Technique to Encode Glyco-Nanoparticles Selectivity**

Abstract: . . . . .	58
3.1 Introduction: . . . . .	59
3.2 Results and Discussion: . . . . .	60
3.3 Conclusions: . . . . .	64
3.4 General Instructions: . . . . .	64
3.4.1 Synthesis procedure of $\alpha$ and $\beta$ -GalNAc: . . . . .	65
3.4.2 Synthesis of Fluorescein modified linker: . . . . .	69
3.4.3 Synthesis of fluorescein functionalized glyco-AuNPs (S-1, S-2, R-1 and R-2): . . . . .	69
3.4.4 UV profile of $\alpha$ and $\beta$ GalNAc functionalized AuNPs: . . . . .	70
3.4.5 Fluorescence spectra of fluorescein functionalized glyco-AuNPs: . . . . .	70
3.4.6 Zeta potential studies: . . . . .	70
3.4.7 Phenol-sulfuric acid method to quantify sugars on AuNPs . . . . .	70
3.4.8 Cytotoxicity assay: . . . . .	71
3.4.8.1 Upright condition: . . . . .	71
3.4.8.2 Inverted condition: . . . . .	71
3.4.9 Upright cell uptake studies using ICP-MS: . . . . .	71
3.4.10 Inverted uptake Studies using ICP-MS: . . . . .	72
3.4.11 Confocal microscopy studies: . . . . .	72
3.5 References: . . . . .	73
3.6 NMR Spectra: . . . . .	76

## **Chapter-4a: Synthesis of Chondroitin Sulfate Analogs**

Abstract: . . . . .	88
4a.1 Introduction: . . . . .	89
4a.2 Synthesis of Chondroitin sulfate: . . . . .	90
4a.3 Synthesis of galactosamine building block: . . . . .	90
4a.4 Synthesis of Galactosamine building block . . . . .	91
4a.5 Synthesis of Glucose building block: . . . . .	91

4a.6 Synthesis of chondroitin sulfate disaccharide analogs: . . . . .	92
4a.7 Synthesis of CS. Tetrasaccharide analog: . . . . .	93
4a.8 Synthesis of CS Hexasaccharide analog: . . . . .	94
4a. 9. Conclusion. . . . .	96
4a.10.1 General information: . . . . .	96
4a.10.2 General Procedure for sulfation reaction: . . . . .	106
4a.10.3 General procedure for hydrolysis and hydrogenolysis reaction: . . . . .	106
4a.11 References: . . . . .	119
4a.10 NMR Spectra: . . . . .	122

## **Chapter-4b: Decoding Sulfation Patterns Dependant Endocytosis of Glycosaminoglycan Functionalized Gold Nanoparticles to Rationalize Targeting Strategies**

Abstract: . . . . .	168
4b.1 Introduction: . . . . .	169
4b.2 Results and Discussion: . . . . .	172
4b.3 Synthesis of tripod: . . . . .	175
4b.4. Conclusion: . . . . .	176
4b.5. Experimental procedure: . . . . .	176
4b.5.1 General procedure for glycan conjugation on tripod: . . . . .	176
4b.5 References: . . . . .	178
4b.6 NMR Spectra . . . . .	180

## Abbreviations

### A

ACE2	Angiotensin-Converting Enzyme 2
NH <sub>4</sub> Cl	Ammonium Chloride
AcOH	Acetic Acid
Ac <sub>2</sub> O	Acetic Anhydride
AIBN	Azobisisobutyronitrile
ACN	Acetonitrile

### B

BAIB	Bis(Acetoxy)Iodobenzene
BnBr	Benzyl Bromide
BF <sub>3</sub> .Et <sub>2</sub> O	Boron Trifluoride Diethyl Etherate
Ph(OMe) <sub>2</sub>	Benzaldehyde Dimethyl Acetal

### C

CS	Chondroitin Sulfate
CAM	Cerium Ammonium Molybdate Stain
CDCl <sub>3</sub>	Deuterated Chloroform
CPIs	Carbohydrate-Protein Interactions
CAN	Ceric Ammonium Nitrate
CsCO <sub>3</sub>	Cesium Carbonate
CsF	Cesium Fluoride
CA	Chloroacetyl
(ClAcO) <sub>2</sub> O	Chloroacetic Anhydride
CSPGs	Chondroitin Sulfate Proteoglycans
CNS	Central Nervous System

### D

DS	Dermatan Sulfate
DCM	Dichloromethane

DMF	Dimethylformamide
DSS	Disuccinimidylsuberate
DMSO	Dimethyl Sulfoxide
DAPI	(4',6-Diamidino-2-Phenylindole)
H <sub>2</sub> NaPO <sub>4</sub>	Dihydrogen Monosodium Phosphate
DMAPA	Dimethylaminopropylamine
DBU	1,8-Diazabicyclo(5.4. 0)Undec-7-Ene
DCC	N,N'-Dicyclohexylcarbodiimide
DMAP	4-Dimethylaminopyridine

## E

EGFR	Epidermal Growth Factor Receptor
eV	Electron Volt
EtOAc	Ethyl Acetate
ECM	Extracellular Matrix

## F

FT-IR	Fourier-Transform Infrared Spectroscopy
-------	---

## G

GlcA	Glucuronic Acid
GAGs	Glycosaminoglycans
GalNAc	Galactosamine
GOPTMS	[(3-Glycidyloxypropyl)Trimethoxysilane]
AuNPs	Gold Nanoparticles

## H

HS	Heparan Sulfate
HP	Heparin
HA	Hyaluronic Acid
HRMS	High-Resolution Mass Spectroscopy



HFF	Human Foreskin Fibroblasts
HBr	Hydrogen Bromide
HCA	Hierarchical Clustering Analysis
HCl	Hydrochloric Acid
HOBt	Hydroxybenzotriazole
NH <sub>2</sub> NH <sub>2</sub> .H <sub>2</sub> O	Hydrazine Monohydrate
<b>I</b>	
IdoA	Iduronic Acid
ICP-MS	Inductively Coupled Plasma Mass Spectrometry
<b>K</b>	
KS	Keratan Sulfate
<b>L</b>	
LiOH	Lithium Hydroxide
Lev	Levulinic Acid
<b>M</b>	
μM	Micro Molar
mM	Mill Molar
MRI	Magnetic Resonance Imaging
MS	Molecular Sieves
MeOH	Methanol
M	Molar
Mg	Milli Gram
μg	Micro Gram
MHz	Megahertz
MTT	Methylthiazolyldiphenyl-Tetrazolium Bromide
<b>N</b>	
NPs	Nanoparticles

NU	<i>N</i> -Unsubstituted
NA	<i>N</i> -Acetylated
nm	Nano Meter
NPC	Nasopharyngeal Carcinoma
NIS	N- N-Iodosuccinimide
NMR	

## O

OA	Osteoarthritis
----	----------------

## P

PTSA	P-Toluenesulfonic Acid
PLL	Poly-Lysine
PBS	Phosphate-Buffered Saline
PEG	Polyethylene Glycol
Py	Pyridine
PNNs	Perineuronal Nets

## R

RFU	Relative Fluorescent Units
RBD	Receptor-Binding Domain
RT	Room Temperature

## S

SEM	Scanning Electron Microscopy
SARS-Cov-2	Severe Acute Respiratory Syndrome Coronavirus 2
NaHCO <sub>3</sub>	Sodium Bicarbonate
SO <sub>3</sub> .TEA	Sulfur Trioxide Trimethylamine Complex
Na <sub>2</sub> SO <sub>4</sub>	Sodium Sulfate
Na <sub>2</sub> S <sub>2</sub> O <sub>3</sub>	Sodium Thiosulfate
NaN <sub>3</sub>	Sodium Azide

NaOAc	Sodium Acetate
NaOMe	Sodium Methoxide
AgOTf	Silver Trifluoromethanesulfonate

## T

<i>T. gondii</i>	Toxoplasma Gondii
TfOH	Trifluoromethanesulfonic Acid
TEMPO	(2,2,6,6-Tetramethylpiperidin-1-Yl)Oxyl
THF	Tetrahydrofuran
TEA	Triethylamine
TLC	Thin-Layer Chromatography
TMSOTf	Trimethylsilyl Trifluoromethanesulfonate
TFA	Trifluoroacetic Acid
CCl <sub>3</sub> CN	Trichloroacetonitrile
TCA	Trichloroacetic Anhydride
TEM	Transmission Electron Microscopes

## U

UV	Ultraviolet–Visible
----	---------------------

## X

XPS	X-Ray Photoelectron Spectroscopy
-----	----------------------------------

## Abstract

Every mammalian cell surface is decorated with 30 nm thickness of carbohydrates, namely glycans from the plasma membrane. These glycans are expressed either in conjugated systems (glycoproteins, proteoglycans, and glycolipids) or without conjugation (glycosaminoglycans such as hyaluronan). These glycans play a fundamental role in cell biology, such as cell-cell interactions that govern many normal and pathological processes. Consequently, a better understanding of structure-activity relationships of glycan will provide new insights into the specific biological process. In my thesis, I have investigated how the conformation plasticity of uronic acid, glycosidic linkage with galactosamine, and sulfation pattern on chondroitin sulfate (CS) regulate specific biological functions.

**Chapter 1** summarizes various factors influencing the structure-activity relation of carbohydrates and thereby significances in biological activities. More specifically, we address basic structures and functions of glycocalyx. Followed by multivalent probes to study carbohydrate-protein interactions. Finally, we talk about the sulfation patterns of GAGs, particularly HS influence the carbohydrate-protein interactions.

**Chapter 2** describes a systematic investigation of preferential binding and therapeutic potential of two crucial non-sulfated and sulfated hexuronic acid derivatives (**GlcA** and **IdoA**) abundantly found in the glycosaminoglycan (GAG) family. We discovered 2,4-disulfated glucuronic acid (Di-S-GlcA) residue as a potential ligand for *T. gondii* recognition. The multivalent display of Di-S-GlcA significantly inhibited the interaction of the parasite with the host cell leading to decreased invasion of host cells. This finding paves the way for future use of Di-S-GlcA in therapeutic studies of *T. gondii* infection in *in vivo* model.

**Chapter 3** reports inverted cell culture technique to study glyconanoparticles selectivity. Briefly, glyco-nanoparticles are indispensable tool to target specific cell lines using the receptor-mediated exchange. However, nanoparticles have tendency to undergo sedimentation with increasing time and resulting in false biological results, including cytotoxicity and cellular internalization process. Here, provides a rational platform to improve the specificity, selectivity, and avoid results from sedimentation of NPs. Validated by a series of cytotoxicity assay and *in vitro* uptake studies, we showed that the inverted cell culture method could rapidly improve our knowledge of cell specificity. This simplicity and effectiveness of the system underscore its potential to accelerate nanoparticles research in biology.

**Chapter 4** deals with the late and preoxidized disaccharides building block approach for synthesizing chondroitin sulfate oligosaccharides. Regioselective benzylidene group opening or deprotection and sulfation yielded oligo-CS-E/A/C analogs in moderate overall yields. Later, the disaccharide CS analogs were functionalized on gold nanoparticles via tripodal conjugation to targeting cancer and glioblastoma cells. We found that the specific sulfation code regulate CD44 mediated uptake in MDA-MB-231 and U87 glioblastoma cell lines, showing the nanomedical application of these glyco-nanostructures.

## List of Publications

1. R. Yadav, S. L. Ben-Arye, **B. Subramani**, V. Padler-Karavani and R. Kikkeri., Screening of Neu5Ac $\alpha$ (2–6)gal isomer preferences of siglecs with a sialic acid microarray. *Org. Biomol. Chem*, 2016, **14**, 10812.
2. S. Sangabathuni, R. V. Murthy, P. M. Chaudhary, **B. Subramani**, S. Toraskar and R. Kikkeri., Mapping the Glyco-Gold Nanoparticles of Different Shapes Toxicity, Biodistribution and Sequestration in Adult Zebrafish. *Sci. Rep.*, 2017, **7**, 4239.
3. C. D. Shanthamurthy, P. Jain, S. Yehuda, J. Monteiro, S. L. Ben-Arye, **B. Subramani**, B. Lepenies, V. Padler-karavani and R. Kikkeri., ABO Antigens Active Tri- and Disaccharides Microarray to Evaluate C-type Lectin Receptor Binding Preferences. *Sci. Rep.*, 2018, **8**, 6603.
4. R. Yadav, P. M. Chaudhary, **B. Subramani**, S. Toraskar, H. Bavireddi, R. V. Murthy, S. Sangabathuni and R. Kikkeri., Imaging and Targeting of the  $\alpha$ (2–6) and  $\alpha$ (2–3) Linked Sialic Acid Quantum Dots in Zebrafish and Mouse Models. *ACS Appl. Mater. Interfaces*, 2018, **10**, 28322.
5. **B. Subramani**, C. D. Shantamurthy, P. Maru, M. A. Belekar, S. Mardhekar, D. Shanmugam, and R. Kikkeri., Demystifying a hexuronic acid ligand that recognizes *Toxoplasma gondii* and blocks its invasion into host cells. *Org. Biomol. Chem*, 2019, **17**, 4535.
6. **B. Subramani**, P. M. Chaudhary and R. Kikkeri., Cell-Culture Technique to Encode Selectivity, *Chem.: Asian J.* 2021, **16**, 1–6.

# **Chapter-1**

## **Driving Force Behind Specific Carbohydrate-Protein Interactions**

---

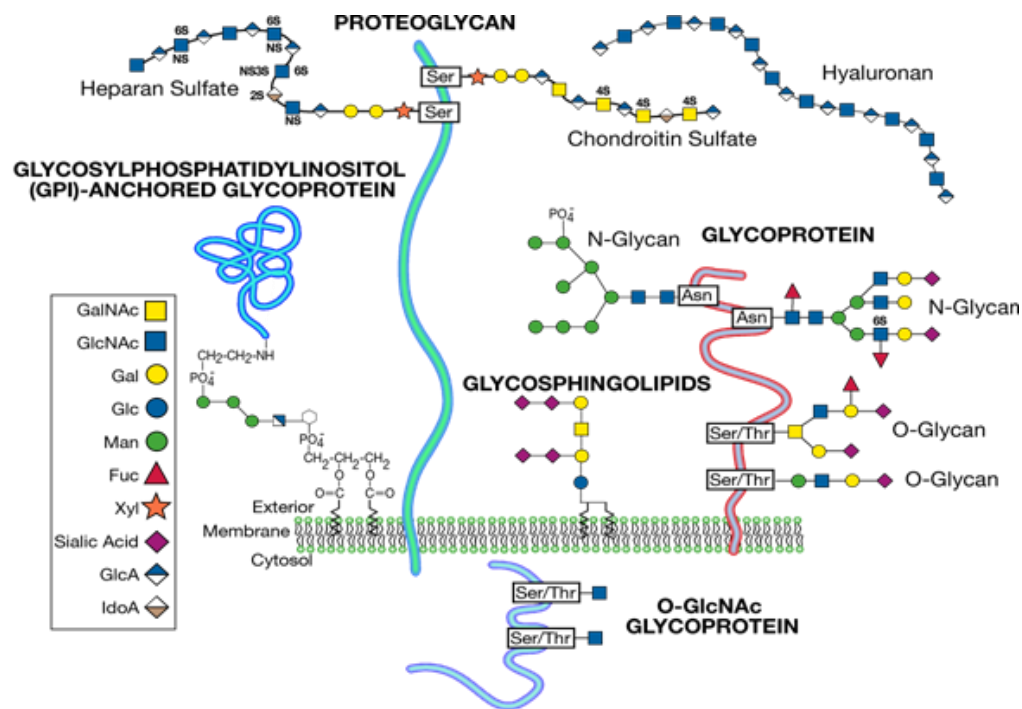
**Abstract:**

Carbohydrate-protein interactions play an important role in living cells due to its involvement in all kind of regulation, recognition and signal transduction processes. It requires that proteins bind to specific epitope of carbohydrates on the cell surface. Chapter 1 briefly summary of the literature study on glycocalyx and how the multivalent probes modulate specific-carbohydrate protein interaction and theirby influenc specific biological event.



## 1. 1 Introduction:

All mammalian cells and viruses are densely covered with carbohydrates on their surface. These carbohydrates bind to proteins or cell surface receptors to regulate several physiological and pathological processes.<sup>1-5</sup> However, carbohydrate structures are highly heterogeneous and complex in nature. Deciphering carbohydrate signature essential to activity specific proteins/receptor is crucial for drug discovery and biomedical applications.<sup>5-8</sup> Many research groups synthesized cell surface carbohydrate with various modifications to understand structure-function relationship. Detailed studies of these carbohydrate structures suggested several key factors for modulating carbohydrate-mediated interactions.<sup>9</sup> This chapter pays special attention to the carbohydrate-based structural variation that shows unique binding specificity with isoforms proteins. We first present a systematic overview of different types of carbohydrates structures on mammalian cell surfaces, followed by the importance structural modification to improve the carbohydrate-protein and development of inhibitors. More specifically, we check the role of multivalency and sulfation patterns in altering the molecular recognition (Figure 1).



**Figure 1.** Mammalian cell surface

## 1.2 Glycocalyx:

Cell surface of human body are densely covered by sugar molecules, commonly called glycocalyx. These sugars are conjugated in various linkages forming glycan, which are expressed as free form or protein/lipid conjugate resulting glycoproteins and

glycosphingolipids. The glycocalyx is known to regulate various fundamental cellular events, such as cell proliferation, differentiation and multiplications. Given the importance of the glycocalyx in all human body functions, understating the structure-activity relation is important for drug designing. Glycocalyx classified in to various types according to their monosaccharides, oligosaccharide or polysaccharide structures.<sup>10-11</sup>

### 1.3 *N*-glycan and *O*-glycan:

Glycoconjugation of protein is the most abundant post-translation modification, which modulate the functions of proteins, including cellular localization, protein quality and 3D-structure of protein. There are two major glycoconjugation observed in pro-translation modification. *N*-glycans are attached to the sequence of Asn-Xaa-Ser/Thr (where Xaa is any amino acid except proline).<sup>12</sup> Whereas, *O*-glycans are attached to the OH groups of serine and threonine residue. The *N*-glycans takes place in endoplasmic reticulum where, a block of sugar (Glc<sub>3</sub>Man<sub>9</sub>GlcNAc<sub>2</sub>), which get attached to the protein. In contrast, *O*-glycans initiated by attachment of *N*-acetyl glucosamine or galactosamine attachment to serine and threonine residue, followed by the extension of the glycan structure in step-by-stem manner.

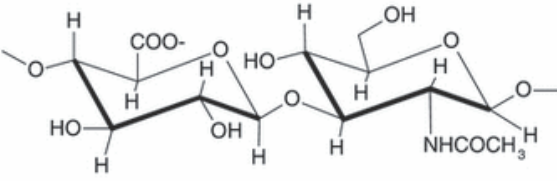
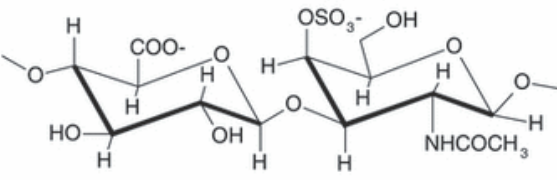
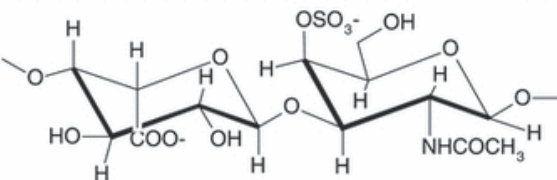
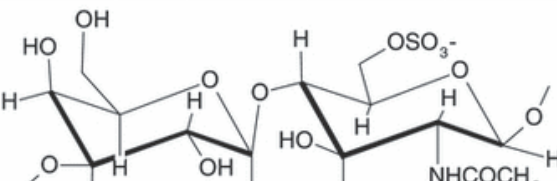
### 1.4 Glycosaminoglycans:

Glycosaminoglycans are a family of liner, highly sulfated polysaccharides. GAGs on glycocalyx has been broadly classified into Non-sulfated GAGs such as hyaluronic acid (HA) and sulphated GAGs include heparan sulfate(HS), chondroitin sulfate (CS), dermatan sulfate (DS) and keratan sulfate (KS). All GAGs contain disaccharide repeating units of uronic acid (l-iduronic acid or D-glucuronic acid) and amino sugar (Glucosamine or galactosamine).<sup>13-15</sup> Heparin (HP) and heparan sulfate (HS) have a common structural base of (1-4) glycosidic linkage of uronic acid and GlcNAc residue but differ in the percentage of uronic acid composition. In HP, the ratio between GlcA vs IdoA is 1:9, whereas in HS, GlcA composition is predominant over IdoA and the amount of GlcA is ranged from 10% to 90%. In addition, HS is low sulfated compared to HP; on average, HS contains a single sulphate group per disaccharides, whereas HP has 2-3 sulfate groups per disaccharides. In general, the glucosamine moiety is *O*-sulfated at C2, C3 or C6 and *N*-sulfated and the uronic acid residue is *O*-sulfated at C2 position. Due to their structure heterogeneity, HS/HP binds to a wide range of plasma proteins, which modulate various biological activities. For example, HS/HP binds to growth factors and triggers proteolytic degradation, resulting in angiogenesis. Similarly, growth factors-HS/HP interaction regulates cell adhesion, tumor development, and

metastasis.<sup>16-19</sup>

Chondroitin sulfate (CS) or dermatan sulfate (DS) are the third most abundant GAG molecules, which has  $\beta(1-3)$  GalNAc and uronic acid disaccharide units. CS has been classified into five groups depending on their sulfation patterns and uronic acid compositions. CS displayed O-sulfation patterns at C4 and C6 of GalNAc and C2 and C3 of uronic acid. CS-B is an old name of DS, which has IdoA as a uronic acid component. Similar to HS/HP, sulfation patterns of CS modulate various biological functions by interacting with growth factors, cytokines and cell adhesion molecules.<sup>20-23</sup>

**Table 1.** Repeating disaccharide units of various glycosaminoglycans

Glycosaminoglycan	Disaccharide units	Features
Hyaluronic acid	 <p style="text-align: center;">D-GlcA- <math>\beta(1 \rightarrow 4)</math>-D-GlcNAc-<math>\alpha(1 \rightarrow 4)</math></p>	<p>Molecular weight 4–8000 kDa</p> <p>Non-sulphated non-covalently attached to proteins in the ECM; also found in bacteria</p> <p>Usually found in synovial fluid, vitreous humour, ECM of loose connective tissue</p> <p>Excellent lubricators and shock absorbers</p>
Chondroitin sulphate	 <p style="text-align: center;">D-GlcA-<math>\beta(1 \rightarrow 3)</math>-D-GalNAc4S-<math>\beta(1 \rightarrow 4)</math></p>	<p>Molecular weight 5–50 kDa</p> <p>Most abundant GAG in the body</p> <p>Found in cartilage, tendon, ligament, aorta</p> <p>Bind to proteins (like collagen) to form proteoglycan aggregates</p>
Dermatan sulphate	 <p style="text-align: center;">L-IdoA-<math>\alpha(1 \rightarrow 3)</math>-D-GalNAc4S-<math>\beta(1 \rightarrow 4)</math></p>	<p>Molecular weight 15–40 kDa</p> <p>Found in skin, blood vessels, heart valves</p>
Keratan sulphates I and II	 <p style="text-align: center;">D-Gal-<math>\beta(1 \rightarrow 4)</math>-D-GalNAc6S-<math>\beta(1 \rightarrow 3)</math></p>	<p>Molecular weight 4–19 kDa</p> <p>Most heterogeneous GAG</p> <p>KS I is found in the cornea</p> <p>KS II is found in cartilage aggregated with CS</p>

Keratan sulfate is the last family of GAGs, which was identified in 1939 by Suzuki while extracting cornea. Structurally, KS contains  $\beta(1-3)$  linkage of Gal-GalNAc unit. Unlike other GAGs, KS replaced uronic acid unit by Galactose and the polymeric chain length is much

shorter than other GAGs.<sup>24-25</sup>

Hyaluronic acid (HA) exists as a glycoconjugation free GAGs, containing long chain  $\beta(1-3)$  of GlcA-GlcNAc units. HA play important roles in the tissue homeostasis, embryogenesis, inflammation, inflammation etc. HA is important of skin smoothness and wound healing and repairing. It is even used in moisturizers and eye drops to treat dry skin. HA is important for tumor growth and inhibition. Hence, it is one of the key ligand for tumor therapy.<sup>26-27</sup>

Overall, Heparan sulfate (HS), chondroitin sulfate (CS) and keratan sulfate (KS), carbohydrate polymer accompanied by a broad range of biological activities, is undisputedly the most complex polysaccharide present in living systems (Table 1). Despite decades of research investigating on these glycosaminoglycans (GAGs) structure-activity relationship, our current understanding is limited by its extensive microheterogeneity. Identification of the structural epitopes of HS, CS and KS that are responsible for orchestrating various cellular processes could enable us to fine-tune the interactions among those processes that modulate biological activities. Inspired by this, we have synthesized a combinatorial library of HS oligosaccharides and HS mimics to investigated their structural relationships with proteins.<sup>28-</sup>

31

## Discovery of Rare Sulfated *N*-Unsubstituted Glucosamine Based Heparan Sulfate Analogs Selectively Activating Chemokines.

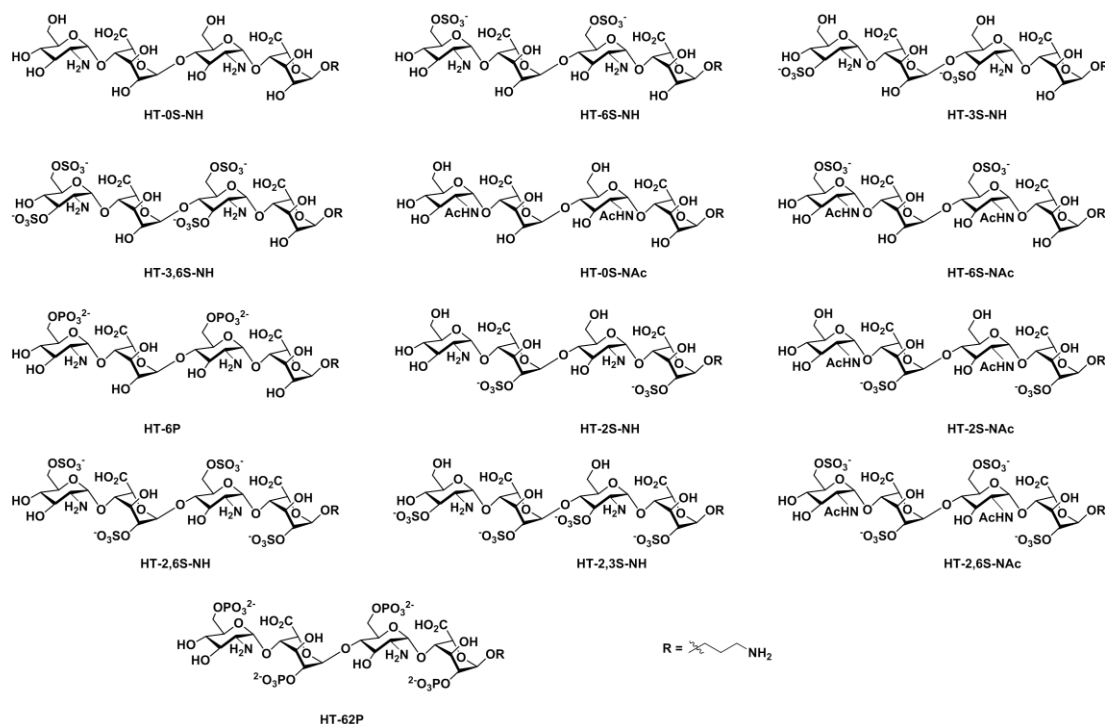


Figure 2. Structures of heparan sulfate tetrasaccharide analogs.

Chemokines are endogenous signaling peptides essential for immuno-surveillance, homeostasis, inflammation, infection and tissue repair.<sup>32-36</sup> Chemokines and their receptor activities depend on how they bind and oligomerize in the presence of glycosaminoglycans (GAGs).<sup>37-41</sup> Humans express 47 chemokines and 20 receptors, and most of the chemokines are highly basic proteins. It has been therefore hypothesized elsewhere that chemokine-GAG binding is non-specific. However, after the discovery of acidic CCL3 and CCL4 chemokines binding to GAGs, it had become clearer that their interaction proceeds in a sequence-dependent manner.<sup>42,43</sup> Moreover, chemokines have shown different interaction strengths with various GAGs, including heparan sulfate, chondroitin sulfate and dermatan sulfate, illustrating that microheterogeneity in GAG structures can modulate binding patterns, for example through uronic acid composition, sulfation patterns and oligosaccharide length.<sup>44</sup> This has prompted us to investigate HS oligosaccharides-chemokine interaction. However, HS is highly heterogeneous in its structure and the majority of the HS libraries that have been used for chemokine studies are composed of *N*-sulfated and *N*-acetylated (NA) glucosamine domains. Native HS also expresses an *N*-unsubstituted (NU) domain, but this region has not been fully investigated in existing studies. To address this gap, and to decipher the sulfation code of chemokine heparin binding, we have synthesized 13 new HS ligands, displaying different sulfate/phosphate patterns with NU/NA glucosamine residue (Figure 2). The binding interactions between HS ligands and chemokines on microarray platform displayed several cryptic binding pockets for sulfation patterns with NU domain, which was not identified with previous HS synthetic ligands. Among them, **HT-3,6S-NH** ligand displayed a marked selectivity and sensitivity to CCL2 chemokine. The biological relevance of such structural binding studies was illustrated by incubating **HT-3,6S-NH** ligand with cancer cells: the HS ligand inhibited cancer cells proliferation, migration and invasion. Thus, NU domain is important to regulate specific chemokine biological activities, thereby demonstrating potential novel therapeutic applications of HS ligands. To the best of our knowledge, we have identified CCL2 and CCL5 chemokines as only the fourth and fifth proteins currently known to recognize NU-domain HS ligands with different sulfation patterns.<sup>45</sup>

Binding was tested at 3 serial dilutions, then detected with the relevant biotinylated secondary antibody (1 µg/ml) followed by Cy3-streptavidin (1.5 µg/ml). Arrays were scanned, relative fluorescent units (RFU) were obtained, and maximum RFU determined was set as 100% binding. Then rank binding (per printed glycan per concentration, per each growth factor dilution, per printed block) was determined. Since each glycans was printed at 2

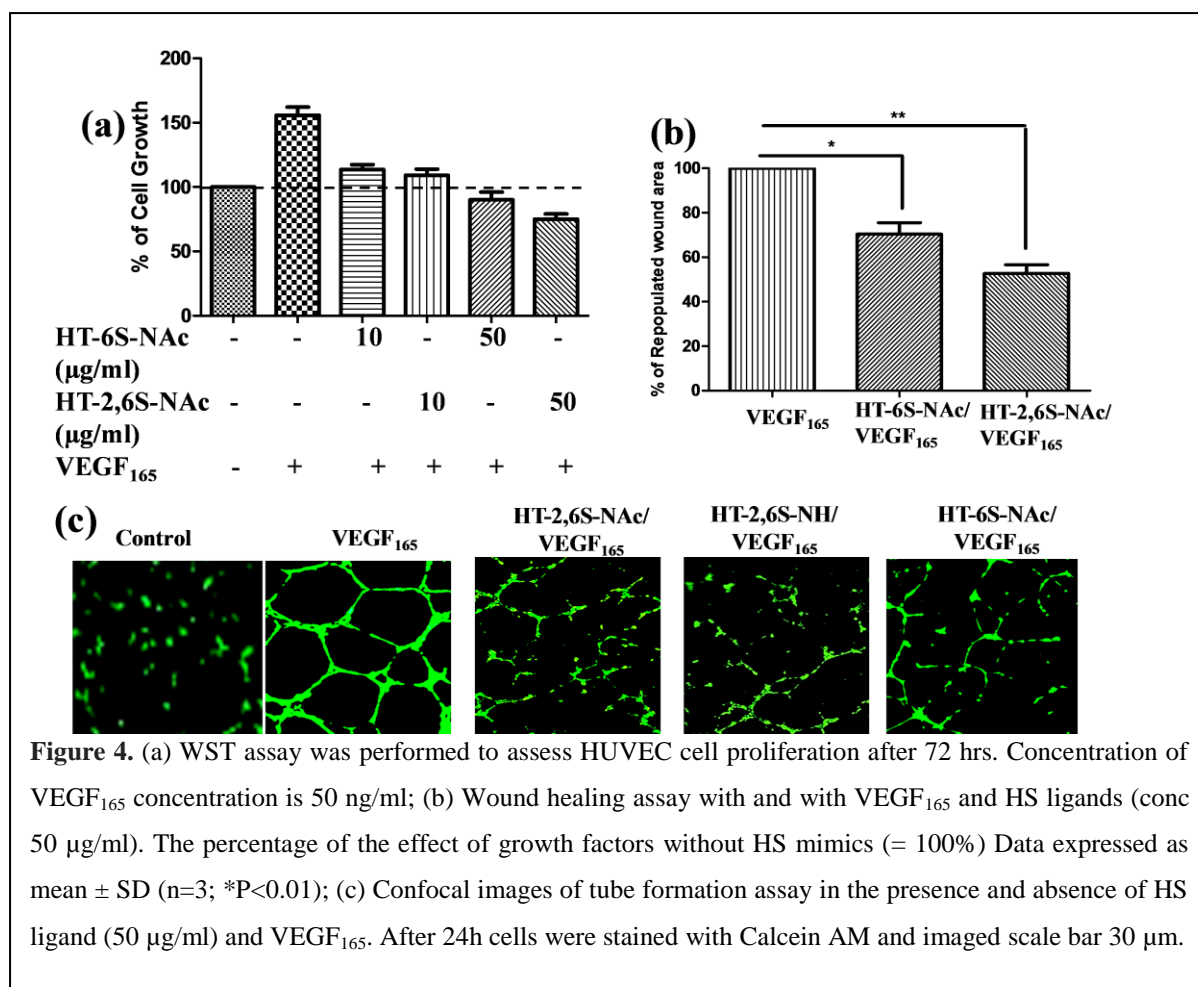
Sulfation/ Phosphatation	Substituent	ID	Amphiregulin	BMP2	VEGF <sub>165</sub>	HB-EGF	FGF2	% Rank
None	None	HT-0S-NH	36	42	32	32	46	100
	None	HT-0S-NAc	24	6	2	11	6	75
Di-Sulfateted	IdoA(2S)	HT-2S-NH	43	72	69	69	59	50
	IdoA(2S)	HT-2S-NAc	18	5	2	11	6	25
	GlcNH(6S)	HT-6S-NH	21	8	5	16	46	0
	GlcNAc(6S)	HT-6S-NAc	48	94	66	73	72	
	GlcNH(3S)	HT-3S-NH	47	68	68	62	68	
Tetra-Sulfateted	GlcNH(6S)-IdoA(2S)	HT-2,6S-NH	38	78	87	82	71	
	GlcNAc(6S)-IdoA(2S)	HT-2,6S-NAc	33	85	88	84	73	
	GlcNH(3S, 6S)	HT-3,6S-NH	64	74	82	83	68	
	GlcNH(3S)-IdoA(2S)	HT-2,3S-NH	72	41	19	42	54	
Phosphated	GlcNAc(6P)-IdoA(2P)	HT-6,2P	44	80	68	72	54	
	GlcNAc(6P)	HT-6P	48	67	22	42	38	

**Figure 3.** Growth Factors glycan microarray binding assay.

concentrations, 100% binding was set separately for each concentration. Then, binding to all the other glycans at the same concentration was ranked in comparison to the maximal binding, and the average rank binding and SEM for each glycan across the two glycan concentrations and three examined dilutions of each growth factor was calculated ( $n=6$ , 2 glycan concentrations across 3 growth factors dilutions). This analysis allowed comparing the glycan binding profiles of the different growth factors and dissecting their binding preferences. The mean rank is shown as a heat map of all the examined binding assays together (red highest, blue lowest and white 50<sup>th</sup> percentile of ranking).<sup>46-49</sup>

Many growth factors, bind to the extracellular matrix of target tissues by interaction with HS oligosaccharides. Identifying minimum HS epitope required to binding these growth factors can be used as drug molecules to modulate its activity.<sup>50-58</sup> However, a systematic study of HS analogs with growth factors, particularly with rare *N*-unsubstituted (NU) domain and *N*-acetate domain HS ligands are not reported. We report the discovery of a potential heparan sulfate (HS) ligand to target several growth factors using 13 unique HS tetrasaccharide ligands.<sup>59-61</sup> By employing an HS microarray (Figure 3) and SPR, we deciphered the crucial structure-binding relationship of these glycans with the growth factors amphiregulin, BMP2, VEGF<sub>165</sub>, HB-EGF, and FGF2. Notably, amphiregulin showed high affinity and selective binding to GlcNH(3-OSO<sub>3</sub>)-IdoA(2-O-SO<sub>3</sub>) (**HT-2,3S-NH**) ligand, whereas, GlcNHAc(6-O-SO<sub>3</sub>)-IdoA(2-O-SO<sub>3</sub>) (**HT-2,6S-NAc**) tetrasaccharide showed strong binding with VEGF<sub>165</sub> growth factor. Active binding between VEGF<sub>165</sub> and its native receptor VEGFR-2 is known to trigger several cellular events, including vascular cell proliferation, cell migration and angiogenesis (Figure 4). These results clearly showed that **2,6-disulfated HS ligands** are potential ligands for targeting VEGF<sub>165</sub>.<sup>62</sup>

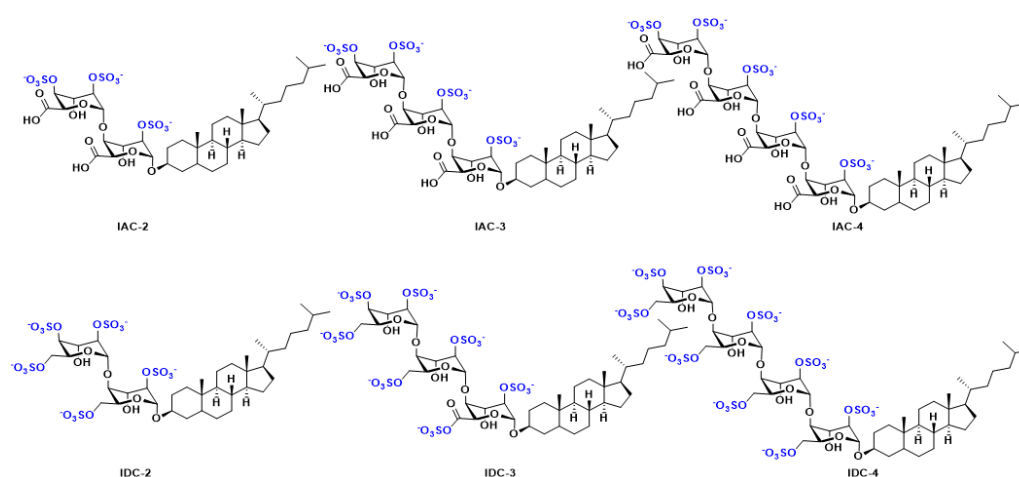




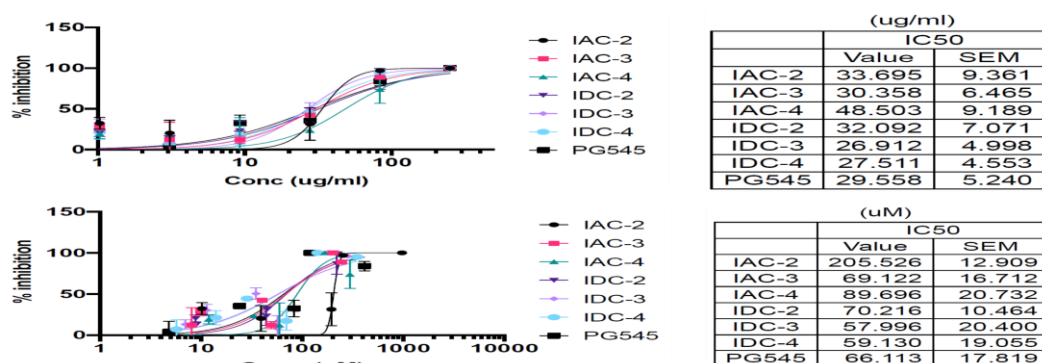
## 1.5 Discovery of Antiviral Heparan Sulfate Mimics:

Coronavirus is a major public health concern of 20<sup>th</sup> century, resulting in worldwide social and economic downfall and unsustainable health care challenges. This includes the development of effective vaccine and therapeutic drug molecules. The virulent activity of Severe acute respiratory syndrome coronavirus 2 (SARS-Cov-2) is principally mediated by the receptor-binding domain (RBD) of the spike protein (S1) and its binding to cell surface ACE2 receptor to transmit RNA to the host cells, causing series of pathological processes.<sup>63-65</sup> Recently, Prof. Esko showed that SARS-CoV-2 spike protein interacts with both cellular heparan sulfate (HS) and angiotensin-converting enzyme 2 (ACE2) through its receptor-binding domain (RBD). Docking studies showed a heparin/heparan sulfate-binding site adjacent to the ACE2-binding site.<sup>66-68</sup> Both ACE2 and heparin can bind independently to spike protein.<sup>6</sup> The binding of heparin to SARS-CoV-2 S protein shifts the structure to favor the RBD open conformation that binds ACE2, suggesting that HS or HS mimics with self-assembly nature could be a potential inhibitor. To achieve self-assembly in HS mimics, we

have incorporated cholesterol moiety at the reducing end of the sulfated oligo-L-iduronic acid and L-Idose and synthesized a library of HS-amphiphiles (Figure 4). We then performed inhibition studies SARS-CoV-2 infection of Vero Cells in the presence of HS mimics amphiphiles. As expected, all HS amphiphiles exhibited strong inhibition of SARS-CoV-2 infection. However, among the L-iduronic acid series, we observed **IAC-2** displayed poor inhibition, whereas **IAC-3** displayed strong inhibition of 69  $\mu\text{M}$  as compared to **IAC-4** 89  $\mu\text{M}$ . In contrast, L-idose analogs showed nearly the same range of SARS-CoV-2 inhibition (60-70  $\mu\text{M}$ ). Among them, **IDC-3** displayed strong inhibition at  $\sim 57 \mu\text{M}$ , in comparison to **IDC-4** of 69  $\mu\text{M}$ . These results are comparable to PG545 of 66  $\mu\text{M}$ . These results confirmed that the trisaccharide scaffold is optimal sugar length essential to inhibit the SARS-CoV-2 infection. Motivated by these results, we are now synthesizing **IDC-3** analogs with different sulfation and phosphate patten to fine-tune the binding affinity and antiviral activity (Figure 6).



**Figure 5.** Chemical structures of heparin sulfate (HS) amphiphiles.

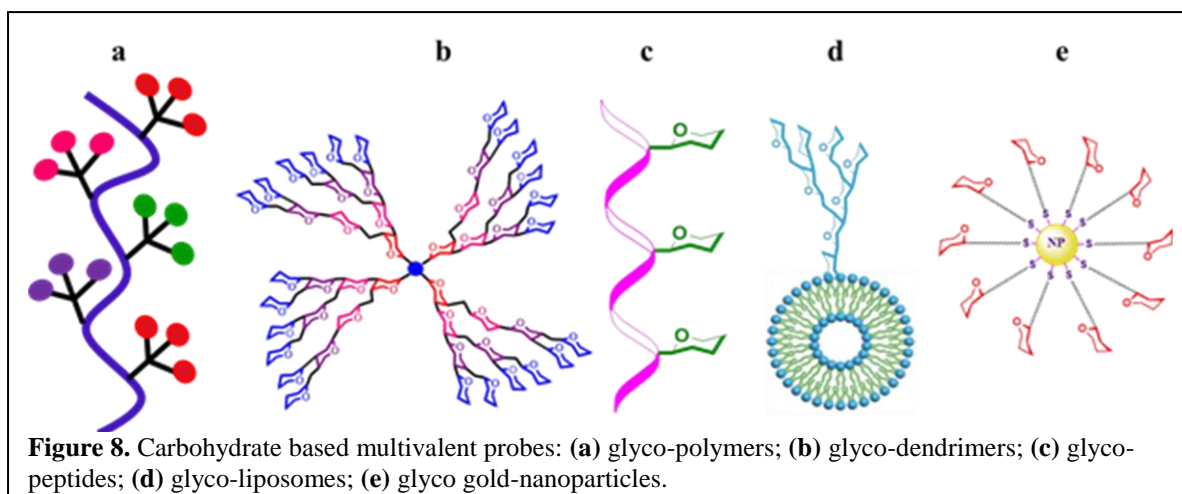
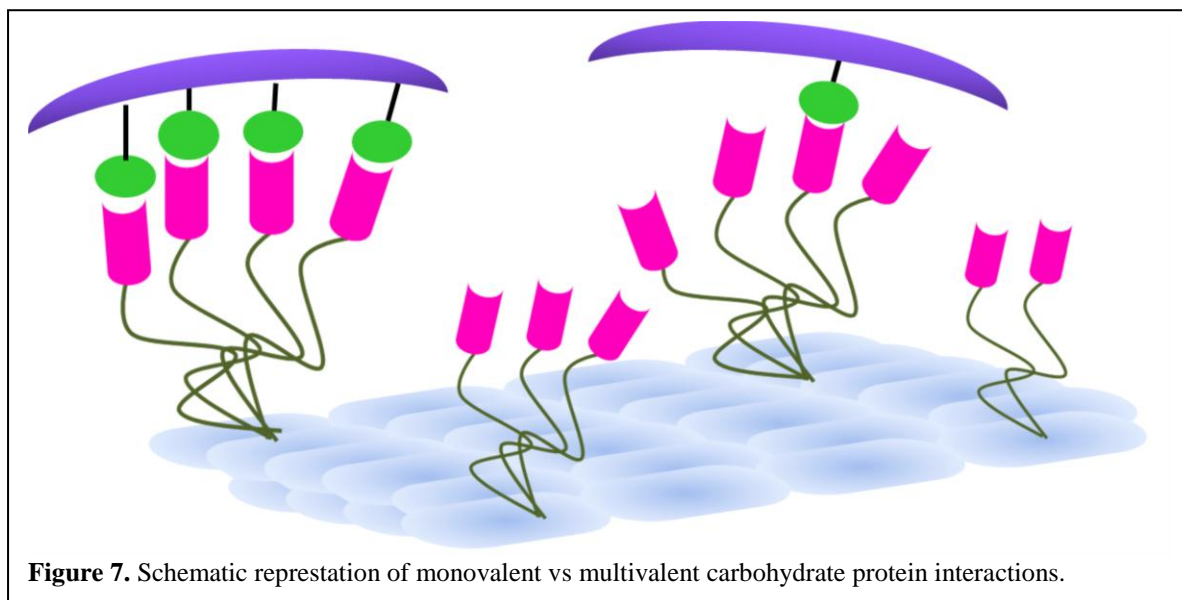


**Figure 6.** HS amphiphiles inhibits attachment and invasion of Vero cells by live SARS-CoV-2 virus isolates.



## 1.6 Carbohydrate-protein interactions:

Carbohydrate-protein interactions (CPIs) are essential in many biological processes. However, interactions between monomeric ligand and single binding site of protein has  $\mu\text{M}$  to  $\text{mM}$  range weak binding. Hence, many biological system choose multivalent probes to increase the binding avidity. In addition to multivancy, carbohydrate orientation, spacing is also important to enhance the glycoside cluster effect or multivalent effect to improve the binding (Figure 7). Inspired by the nature, several multivalent probes have been synthesized to modulate the carbohydrate-protein interactions, which include, glycodendrimers, glycopolymers, glycolipids and glyconanoparticles (Figure 8).<sup>69-81</sup>



## 1.7 Glyconanoparticles:

Nanoparticles are usually referred to the particles size of between 1 and 100 nm in diameter. At this tiny size, particles start behaving differently from that of bulk materials. Researchers have exploited the nanoparticles inherent optical, electronic, magnetic properties to generate imaging and biomedical applications.<sup>82-86</sup> Several types of nanoparticles have been synthesized to target carbohydrate-mediated biological interaction.<sup>87-90</sup> Among them, iron oxide nanoparticles are the most predominant one, due to its inherent magnetic resonance imaging (MRI) contrast agent to develop the real-time imaging of carbohydrate-mediated interactions.<sup>91-94</sup> These magnetic nanoparticles have several advantages over Gadolinium based contrast agent such as Gd-DTPA, which has a low T1 relaxivity. The development of iron oxide nanoparticles-based MRI imaging displayed high T2 relaxivity, essential to visualize deep neural networks and tissues. Previously, Ben Davis et al. and Seeberger et al. employed sialyl lewisx conjugated silica iron oxide nanoparticles to study neural inflammation in the ischemic stroke mice model.<sup>95-97</sup> MRI imaging of the brain demonstrated the feasibility of imaging endothelial activation *in vivo* after acute stroke. However, the surgical procedure to develop a stroke model enhances P-selectin expression, and therefore the exact level of P-selectin expression is difficult to interpret. Furthermore, the T2 relativity of iron oxide nanoparticles that produce dark MRI images induces a high background interface. Hence, IONPs are taken off the market and looking for an alternative NPs based contrast agent. Recently, ultra small iron oxide nanoparticles (USIONPs) of size <4 nm were used as a positive T1 relativity MRI contrast agent, as they offer bright MRI imaging and high-resolution, resulting in easy detection of the target.<sup>98-101</sup> Recently, Rao et al. reported the synthesis of quinic acid (QA)-coated USIONPs to target P-selectin overexpressed aggressive cancer cells. Working under the same notion, Liu et al. reported the synthesis of USIONPs-PEG-SLex nanoparticles to assess *in vivo* E-selectin expression level in nasopharyngeal carcinoma (NPC) metastasis.<sup>102</sup> On considering of the inherent magnetic properties of USIONPs, this is a valuable tool for MRI imaging of ischemic stroke.

## 1.8 Glycopolymers:

Polymers are powerful multivalent probes, as they offer a wide range of molecular weight, cheaper, biocompatible and readily scalable materials. There are two types of polymeric systems readily used in selectin-mediated studies.<sup>103-106</sup> In the first set of carbohydrate appended polymers were functionalized on acrylate or dopamine residue and employed polymerization strategy to obtain desired polymeric nanoparticles. However, the functional

groups such as carboxylic acid and hydroxy groups quench the free radical polymerization, resulting in low efficacy and high polydispersity. Alternatively, polymeric nanoparticles were formulated by the emulsion-solvent evaporation method encapsulated with drug molecules and selectin binding ligands. This method offers several advantages including the mild condition to formulate biocompatible and biodegradable nanoparticles without compromising the activity of drug. Moreover, the nanoparticles prepared by this method offer control delivery of drug molecules. Nguyen and co-workers reported the synthesis of drug (dexamethasone) and fluorescent tag (6-coumarin dye) and glycoprotein Ib (GPIb; p-selectin binding ligand)-loaded poly(lactide-co-glycolide) nanoparticles to target P-selectin overexpressed endothelial cells and drug delivery.<sup>107</sup> Similarly, Ran and co-workers synthesized polymeric nanoparticles composed of photothermal agents such as nanocarbons, doxorubicin and perfluoropentane encapsulated polymeric nanoparticles coated with platelet membrane.<sup>108</sup> Overall, these results all show the huge potential of polymer-based nanoparticles in real-time applications of cancer therapy.<sup>109-116</sup>

## **1.9 Liposomes:**

Liposomes are one of the attractive nanocarriers for controlled cargo delivery, composed of lipid bilayers in a discrete aqueous environment. They can host both hydrophilic drug molecules in the aqueous centre and hydrophobic molecules between the lipid bilayers and display a large surface area to functionalize biological ligands, including carbohydrates, peptides, and proteins.<sup>117-120</sup> Liposomes are extensively used to design carbohydrate-mediated drug delivery and imaging system to target various diseases. Matsumura et al. synthesized sialyl lewis X-modified doxo-liposomes to target injured vessel walls to prevent stenosis after angioplasty.<sup>121</sup> While, Zalipsky et al. reported silyl lewis X liposomes to develop anti-adhesion molecules.<sup>122</sup> Azab et al., prepared bone marrow microenvironment destructing inhibitor modified P-selectin glycoprotein ligand-1 conjugated liposome to target multiple myeloma-associated endothelium. However, single-ligands often fail to target the dynamic microenvironment of the tumor, particularly metastasis cancer cells. To improve the accuracy in targeting metastasis cancer cells, multi-ligand embedded liposomal nanoparticles have been synthesized. Here, P-selectin-specific ligands, integrin-targeting peptides, fibronectin targeting peptides, and epidermal growth factor receptor (EGFR) targeting peptides were assembled on a single liposome to target more than one receptor overexpressed on cancer cells. By employing such multi-ligand strategies, highly sensitive and precise imaging of early-stage cancer cells metastasis was achieved. Overall, selectin-ligands conjugated

liposomes have shown enormous potential in theranostics. However, the poor cost-effectiveness of liposome-based drug-delivery limited its clinical translation.

### **1.10 Conclusion.**

Glycocalyx is highly heterogeneous and multifunctional natural substrate. GAGs are important component of glycocalyx composed of heparan sulfate and chondroitin sulfate ligands. These molecules utilize selective sulfation patterns, uronic acid composition and the nature of glycosidic linkage to modulate the biological activity. Further more, when these molecules are expressed on multivalent systems such as nanoparticles, polymers and liposomes enabled their application in the areas of biosensing, imaging, and therapeutics.

## 1.10 References:

1. *Essentials of Glycobiology.*, A. Varki, R. Cummings, J. Esko, H. Freeze, G. Hart and J. Marth, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2009.
2. A. Varki, Biological roles of glycans. *Glycobiology*. 2017, **27**, 3-49.
3. L. Cipolla, A. C. Araujo, D. Bini, L. Gabrielli, L. Russo and N. Shaikh, *Expert Opin. Drug Discov.*, 2010, **5**, 721-737.
4. K. D. Schutter and E. J. M. Van Damme, *Molecules*, 2015, **20**, 15202-15205.
5. Y. Kamiya, M. Yagi-Utsumi, H. Yagi and K. Kato, *Current Pharmaceutical Design*, 2011, **17**, 1672-1684.
6. I. Pramudya and H. Chung, *Biomater. Sci.*, 2019, **7**, 4848–4872.
7. P. Valverde, A. Ardá, N. Reichardt, J. Jiménez-Barbero and A. Gimeno, *Med. Chem. Commun.*, 2019, **10**, 1678–1691.
8. S. Gim, Y. Zhu, P. H. Seeberger and M. Delbianco, *WIREs Nanomed Nanobiotechnol.*, 2019, **11**, 1-29.
9. S. Park, M. Lee and I. Shin, *Chem. Commun.*, 2008, 4389–4399.
10. L. Pan, C. Cai, C. Liu, D. Liu, G. Li, R. J. Linhardt and G. Yu, *Curr. Opin. Biotechnol.* 2021, **69**, 191–198.
11. I. Bucior and M. M. Burger, *Curr. Opin. Struct. Biol.* 2004, **14**, 631–637.
12. S. Reitsma, D. W. Slaaf, H. Vink, M. A. M. J. van Zandvoort, M. G. A. oude Egbrink, *Pflugers Arch - Eur J Physiol.*, 2007, **454**, 345–359.
13. C. I. Gama and L. C. Hsieh-Wilson, *Curr. Opin. Chem. Biol.* 2005, **9**:609–619.
14. M. Mende, C. Bednarek, M. Wawrzych, P. Sauter, M. B. Biskup, U. Schepers and S. Bräse, *Chem. Rev.*, 2016, **116**, 8193–8255.
15. A. Köwitsch, G. Zhou and T. Groth, *J Tissue Eng Regen Med.*, 2018, **12**, 23–41.
16. L. Fu, M. Suflita and R. J. Linhardt, *Adv. Drug Deliv. Rev.*, 2016, **97**, 237–249.
17. S. Sarrazin<sup>1</sup>, W.C. Lamanna<sup>1</sup> and J. D. Esko, *Cold Spring Harb. Perspect Biol.*, 2011, **3**, 1-33.
18. A. G. Toledo, J. T. Sorrentino, D. R. Sandoval, J. Malmström, N. E. Lewis, and J. D. Esko, *J. Histochem. Cytochem.* 2021, **69**, 105–119.

19. M. D. Stewart and R. D. Sanderson, *Matrix Biol.*, 2014, **35**, 56–59.
20. N. Maeda, M. Ishii, K. Nishimura and K. Kamimura, *Neurochem. Res.*, 2011, **36**, 1228–1240.
21. N. Sugiura, T. Shioiri, M. Chiba, T. Sato, H. Narimatsu, K. Kimata, and H. Watanabe, *J. Biol. Chem.*, 2012, 287, 43390–43400.
22. Kazuyuki Sugaharaya, Tadahisa Mikami, Toru Uyama,
23. S. Mizuguchiz, K. Nomuraz and H. Kitagawa, *Curr. Opin. Struct. Biol.*, 2003, **13**, 612–620.
24. T. M. Hering, J. A. Beller, C. M. Calulot and D. M. Snow, *Biology*, 2020, **9**, 29.
25. B. Caterson and J. Melrose, *Glycobiology*, 2018, **28**, 182–206.
26. M. Litwiniuk, A. Krejner and T. Grzela, *Wounds*, 2016, 28, 78–88.
27. H. Li, Z. Yu, Z. Qi, X. Yang, S. Zheng, and S. Pan, *Adv. Mater. Sci. Eng.*, 2019, 1–12. (<https://doi.org/10.1155/2019/3027303>).
28. S. Ricard-Blum and S. Perez, *Curr. Opin. Struct. Biol.*, 2022, **74**, 102355.
29. N. Afratis, C. Gialeli, D. Nikitovic, T. Tsegenidis, E. Karousou, A. D. Theocharis, M. S. Pavao, G. N. Tzanakakis and N. K. Karamanos, *FEBS Journal*, 2012, **279**, 1177–1197.
30. S. Morla, *Int. J. Mol. Sci.*, 2019, **20**, 1963.
31. Q. Liu, G. Chen and H. Chen, *Polym. Chem.*, 2019, **10**, 164–171.
32. A. Mantovani, R. Bonecchi and M. Locati, *Nat. Rev. Immunol.*, 2006, **6**, 907–918.
33. C. R. Mackay, *Nat. Immunol.*, 2001, **2**, 95–101.
34. R. J. Nibbs and G. J. Graham, *Nat. Rev. Immunol.*, 2013, **13**, 815–829.
35. B. Moser and P. Loetscher, *Nat. Immunol.*, 2001, **2**, 123–128.
36. A. Mortier, J. Van Damme and P. Proost, *Immunol. Lett.*, 2012, **145**, 2–9.
37. K. Jöhrer, L. Pleyer, A. Olivier, E. Maizner, C. Zelle-Rieser and R. Greil, *Expert Opin. Biol. Ther.*, 2008, **8**, 269–290.
38. W. G. Liang, C. G. Triandafillou, T. Y. Huang, M. M. L. Zulueta, S. Banerjee, A. R. Dinner, S.-C. Hung and W.-J. Tang, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 5000–5005.
39. D. P. Dyer, C. L. Salanga, B. F. Volkman, T. Kawamura and T. M. Handel, *Glycobiology*, 2016, **26**, 312–326.

40. L. Wang, M. Fuster, P. Sriramaraao and J. D. Esko, *Nat. Immunol.*, 2005, **6**, 902–910.
41. C. R. Parish, *Nat. Rev. Immunol.*, 2006, **6**, 633–643.
42. A. E. Proudfoot, T. M. Handel, Z. Johnson, E. K. Lau, P. LiWang, I. Clark-Lewis, F. Borlat, T. N. Wells and M. H. Kosco-Vilbois, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 1885–1890.
43. Z. Johnson, A. Proudfoot and T. Handel, *Cytokine Growth Factor Rev.*, 2005, **16**, 625–636.
44. K. M. Sepuru and K. Rajarathnam, *J. Biol. Chem.*, 2019, **294**, 15650–15661.
45. P. Jain, C. D. Shanthamurthy, S. L. Ben-Arye, R. J. Woods, R. Kikkeri and V. Padler-Karavani, *Chem. Sci.*, 2021, **12**, 3674–3681.
46. V. Padler-Karavani, X. Song, H. Yu, N. Hurtado-Ziola, S. Huang, S. Muthana, H. A. Chokhawala, J. Cheng, A. Verhagen and M. A. Langereis, *J. Biol. Chem.*, 2012, **287**, 22593–22608.
47. S. L. Ben-Arye, H. Yu, X. Chen and V. Padler-Karavani, *J. Visualized Exp.*, 2017, e56094.
48. M. Gade, C. Alex, S. L. Ben-Arye, J. T. Monteiro, S. Yehuda, B. Lepenies, V. Padler-Karavani and R. Kikkeri, *Chembiochem*, 2018, **19**, 1170–1177.
49. C. D. Shanthamurthy, P. Jain, S. Yehuda, J. T. Monteiro, S. L. Ben-Arye, B. Subramani, B. Lepenies, V. Padler-Karavani and R. Kikkeri, *Sci. Rep.*, 2018, **8**, 1–7.
50. J. Kreuger, D. Spillmann, J.-P. Li and U. Lindahl, *J. Cell Biol.*, 2006, **174**, 323–327.
51. L. E. Collins and L. Troeberg, *J. Leukocyte Biol.*, 2019, **105**, 81–92.
52. S. J. Allen, S. E. Crown and T. M. Handel, *Annu. Rev. Immunol.*, 2007, **25**, 787–820.
53. C. Zong, A. Venot, X. Li, W. Lu, W. Xiao, J.-S. L. Wilkes, C. L. Salanga, T. M. Handel, L. Wang, M. A. Wolfert and G. J. Boons, *J. Am. Chem. Soc.*, 2017, **139**, 9534–9543.
54. J. L. de Paz, E. A. Moseman, C. Noti, L. Polito, U. H. von Andrian and P. H. Seeberger, *ACS Chem. Biol.*, 2007, **2**, 735–744.
55. T. Handel, Z. Johnson, S. Crown, E. Lau, M. Sweeney and A. Proudfoot, *Annu. Rev. Biochem.*, 2005, **74**, 385–410.



56. M. Nonaka, X. Bao, F. Matsumura, S. Go'tze, J. Kandasamy, A. Kononov, D. H. Broide, J. Nakayama, P. H. Seeberger and M. Fukuda, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 8173–8178.
57. G. J. Sheng, Y. I. Oh, S.-K. Chang and L. C. Hsieh-Wilson, *J. Am. Chem. Soc.*, 2013, **135**, 10898–10901.
58. G. C. Jayson, S. U. Hansen, G. J. Miller, C. L. Cole, G. Rushton, E. Avizienyte and J. M. Gardiner, *Chem. Commun.*, 2015, **51**, 13846–13849.
59. P. Jain, C. D. Shanthamurthy, P. M. Chaudhary and R. Kikkeri, *Chem. Sci.*, 2021, **12**, 4021–4027.
60. C. D. Shanthamurthy and R. Kikkeri, *Eur. J. Org. Chem.*, 2019, 2950–2953.
61. S. Anand, S. Mardhekar, R. Raigawali, N. Mohanta, P. Jain, C. D. Shanthamurthy, B. Gnanaprakasam and R. Kikkeri, *Org. Lett.*, 2020, **22**, 3402–3406.
62. P. Jain, C. D. Shanthamurthy, S. L. Ben-Arye, S. Yehuda, S. S. Nandikol, H. V. Thulasiram, V. Padler-Karavani and R. Kikkeri, *Chem. Commun.*, 2021, **57**, 3516–3519.
63. J. Gao, Z. Tian and X. Yang, *Biosci. Trends*, 2000, **14** 1–2.
64. X. Xu, M. Han, T. Li, W. Sun, D. Wang, B. Fu, Y. Zhou, X. Zheng, Y. Yang, X. Li, X. Zhang, A. Pan and H. Wei, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 10970–10975.
65. C. A. Devaux, J. Rolai, P. Colson and D. Raoult, *Int. J. Antimicrob. Agents*, 2020, **55**, 105938.
66. T. M. Clausen, D. R. Sandoval, C. B. Spliid, J. Pihl, H. R. Perrett, C. D. Painter, A. Narayanan, S. A. Majowicz, E. M. Kwong, R. N. McVicar, B. E. Thacker, C. A. Glass, Z. Yang, J. L. Torres, G. J. Golden, P. L. Bartels, R. N. Porell, A. F. Garretson, L. Laubach, J. Feldman, X. Yin, Y. Pu, B. M. Hauser, T. M. Caradonna, B. P. Kellman, C. Martino, P. L.S.M. Gordts, S. K. Chanda, A. G. Schmidt, K. Godula, A. B. Ward, A. F. Carlin, and J. D. Esko, *Cell*, 2020, **183**, 1043–1057.
67. S. Y. Kim, W. Jin, A. Sood, D. W. Montgomery, O. C. Grant, M. M. Fuster, L. Fu, J. S. Dordick, R. J. Woods, F. Zhang and R. J. Linhardt, *Antiviral Res.*, 2020, **181**, 104873.
68. W. Hao, B. Ma, Z. Li, X. Wang, X. Gao, Y. Li, B. Qin, S. Shang, S. Cui and



- Z. Tan, *Sci. Bull.*, 2021, **66**, 1205–1214.
69. J. M. Hearn, I. Romero-Canelon, B. Qamar, Z. Liu, I. Hands-Portman, P. J. Sadler, *ACS Chem. Biol.*, 2013, **8**, 1335.
70. C. Santini, M. Pellei, V. Gandin, M. Porchia, F. Tisato, C. Marzano, *Chem. Rev.*, 2014, **114**, 815.
71. L. Larry Tso-Lun, C. Wing-Kin, T. Chun-Yat, Y. Shek-Man, K. Chi-Chu, C. Sung-Kay, *Organometallics*, 2011, **30**, 5873.
72. M. Ahmed, S. Mamba, X. H. Yang, J. Darkwa, P. Kumar, R. Narain, *Bioconjugate Chem.*, 2013, **24**, 979.
73. J. S. Butler, J. A. Woods, N. J. Farrer, M. E. Newton, P. J. Sadler, *J. Am. Chem. Soc.*, 2012, **134**, 16508.
74. T. H. Vien, P. Samuel, N. Janina-Miriam, A. Amanda, H. U. Robert, L. Hongxu, H. S. Martina, *ACS Macro Lett.*, 2013, **2**, 246.
75. A. Bergamo, G. Sava, *Dalton Trans.*, 2011, **40**, 7817.
76. E. S. Antonarakis, A. Emadi, *Cancer Chemother. Pharmacol.*, 2010, **66**, 1.
77. A. Amin, M. A. Buratovich, *Mini. Rev. Med. Chem.*, 2009, **9**, 1489.
78. I. Kostova, *Curr. Med. Chem.*, 2006, **13**, 1085.
79. Y. K. Yan, M. Melchart, A. Habtemariam, P. J. Sadler, *Chem. Commun.*, 2005, **38**, 4764.
80. G. C. Smith, B. Therrien, *Dalton Trans.*, 2011, **40**, 10793.
81. G. Gasser, I. Ott, N. Metzler-Nolte, *J. Med. Chem.*, 2011, **54**, 3.
82. K. McNamara and S. A. M. Tofail, *Adv. Phys.: X*, 2017 **2**, 54–88.
83. J. Shannahan, *Nanotechnol. Rev.*, 2017, **6**, 345–353.
84. N. T.K. Thanh and L. A.W. Green, *Nano Today.*, 2010, **5**, 213-230.
85. N. Elahi, M. Kamali and M. H. Baghersad, *Talanta.*, 2018, **184**, 537–556.
86. M. De, P. S. Ghosh, and V. M. Rotello, *Adv. Mater.*, 2008, **20**, 4225–4241.
87. R. Sunasee and R. Narain, *Macromol. Biosci.*, 2013, **13**, 9–27.
88. A. K. Adak, H. Lin and C. Lin, *Org. Biomol. Chem.*, 2014, **12**, 5563–5573.
89. M. Marradi, F. Chiodo, I Garcia and Soledad Penades, *Chem. Soc. Rev.*, 2013, **42**, 4728-4745.
90. I. García, J. Mosquera, J. Plou, and L. M. Liz-Marzán, *Adv. Optical Mater.*, 2018, **6**, 1800680.
91. L. Li, W. Jiang, K. Luo, H. Song, F. Lan, Y Wu and Z. Gu, *Theranostics*, 2013, **3**, 595-615.

92. M. Iv, N. Telischak, D. Feng, S. J. Holdsworth, K. W. Yeom and H. E. Daldrup-Link, *Nanomedicine (Lond.)*, 2015, **10**, 993–1018.
93. J. S. Weinstein, C. G Varallyay, E. Dosa, S. Gahramanov, B. Hamilton, W. D. Rooney, L. L. Muldoon and E. A. Neuwelt, *J. Cereb. Blood Flow Metab.*, 2010, **30**, 15–35.
94. R. Jin, B. Lin, D. Li and H Ai, *Curr Opin Pharmacol.*, 2014, **18**, 18–27.
95. D. P. Gamblin, S. I. van Kasteren, J. M. Chalker and B. G. Davis, *FEBS Journal*, 2008, **275**, 1949–1959.
96. M. Guberman, M. Brautigam and P. H. Seeberger, *Chem. Sci.*, 2019, **10**, 5634–5640.
97. S. Penadés, B. G. Davis, and P. H. Seeberger, Glycans in Nanotechnology. In: *Essentials of Glycobiology*. 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY); 2015. PMID: 28876841.
98. Y. Wang, C. Xu, Y. Chang, L. Zhao, K. Zhang, Y. Zhao and X. Gao, *ACS Appl. Mater. Interfaces*, 2017, **9**, 28959–28966.
99. L. Zeng, W. Ren, J. Zheng, P. Cui and A. Wu, *Phys. Chem. Chem. Phys.*, 2012, **14**, 2631–2636.
100. C. Song, W. Sun, Y. Xiao and X. Shi, *Drug Discov.*, 2019, **24**, 835-844.
101. J. S. Stiller, A. Hoth, L. Kaufner, U. Pison, and R. Cartier, *Biomacromolecules*, 2006, **7**, 3132-3138.
102. L. Liu, L. Liu, Y. Li, X. Huang, D. Gu, B. Wei, D. Su and G. Jin, *Int. J. Nanomedicine*, 2019, **14**, 4517–4528.
103. M. Bartneck, C. T. Schloßer, M. Barz, R. Zentel, C. Trautwein, T. Lammers and F. Tacke, *ACS Nano*, 2017, **11**, 9689–9700.
104. S. M. Rele, W. Cui, L. Wang, S. Hou, G. Barr-Zarse, D. Tatton, Y. Gnanou, J. D. Esko, and E. L. Chaikof, *J. Am. Chem. Soc.*, 2005, **127**, 10132-10133.
105. Z. Cai, Y. Yan, J. Zhou, Y. Yang, Y. Zhang, and J. Chen, *Biomacromolecules*, 2021, **22**, 1177–1185.
106. L. L. Kiessling and J. C. Grim, *Chem. Soc. Rev.*, 2013, **42**, 4476-4491.
107. B. Shah, S. Kona, T. A. Gilbertson and K. T. Nguyen, *J. Nanosci. Nanotechnol.*, 2011, **11**, 3533–3542.
108. Z. Liu, Y. Ran, J. Xi and J. Wang, *Soft Matter*, 2020, **16**, 9160-9175.
109. F. Mainini and M. R. Eccles, *Molecules*, 2020, **25**, 2692.
110. R. Grosso and M. Violante de-Paz, *Pharmaceutics*, 2021, **13**, 854.

111. B. T. Luk and L. Zhang, *ACS Appl. Mater. Interfaces*, 2014, **6**, 21859–21873.
112. S. Vrignaud, J. Benoit and P. Saulnier, *Biomaterials*, 2011, **32**, 8593-8604.
113. H. Idrees, S. Z. J. Zaidi, A. Sabir, R. U. Khan, X. Zhang and S. Hassan, *Nanomaterials*, 2020, **10**, 1970.
114. Y. Wang, P. Li, T. T. Tran, J. Zhang and L. Kong, *Nanomaterials*, 2016, **6**, 26.
115. N. Yang, C. Cao, H. Li, Y. Hong, Y. Cai, X. Song, W. Wang, X. Mou, and X. Dong, *Small Struct.*, 2021, **2**, 2100110.
116. H. Zhao, Z. Y. Lin, L. Yildirimer, A. Dhinakar, X. Zhao and J. Wu, *J. Mater. Chem. B*, 2016, **4**, 4060-4071.
117. V. Sihorkar, S.P. Vyas, *J. Pharm. Pharm. Sci.*, 2001, **4**, 138-158.
118. R. K. Shukla and A. Tiwari, *Crit. Rev. Ther. Drug Carr. Syst.*, 2011, **28**, 255–292.
119. G. Sharma, S. Anabousi, C. Ehrhardt and M. N. V. R. Kumar, *J. Drug Target.*, 2006, **14**, 301-310.
120. M. N. Jones, *Adv. Drug Deliv. Rev.*, 1994, **13**, 215-250.
121. W. Tsuruta, H. Tsurushima, T. Yamamoto, K. Suzuki, N. Yamazaki and A. Matsumura, *Biomaterials*, 2009, **30**, 118–125.
122. S. A. DeFrees, L. Phillips, L. Guo, and S. Zalipsky, *J. Am. Chem. Soc.*, 1996, **118**, 6101-6104.
123. A. K. Azab, J. Hu, P. Quang, F. Azab, C. Pitsillides, R. Awwad, B. Thompson, P. Maiso, J. D. Sun, C. P. Hart, A. M. Roccaro, A. Sacco, H. T. Ngo, C. P. Lin, A. L. Kung, R. D. Carrasco, K. Vanderkerken, and I. M. Ghobrial, *Blood*, 2012, **119**, 5782-5794.

## **Chapter-2**

# **Demystifying a Hexuronic Acid Ligand that Recognizes *Toxoplasma Gondii* and Blocks its Invasion Into Host Cells**

---

## Abstract:

*Toxoplasma gondii* is a ubiquitous eukaryotic pathogen responsible for toxoplasmosis in humans and animals. This parasite is an obligate intracellular pathogen and actively invades susceptible host cells, a process which is mediated by specific receptor–ligand interactions. Here, we have identified an unnatural 2,4-disulfated D-glucuronic acid (**Di-S-GlcA**), a hexuronic acid composed of heparin/heparansulfate, as a potential carbohydrate ligand that can selectively bind to *T. gondii* parasites. More importantly, the gelatin conjugated **Di-S-GlcA** multivalent probe displayed strong inhibition of parasite entry into host cells. These results open perspective for the future use of **Di-S-GlcA** epitopes in biomedical applications against toxoplasmosis.

## 2.1 Introduction:

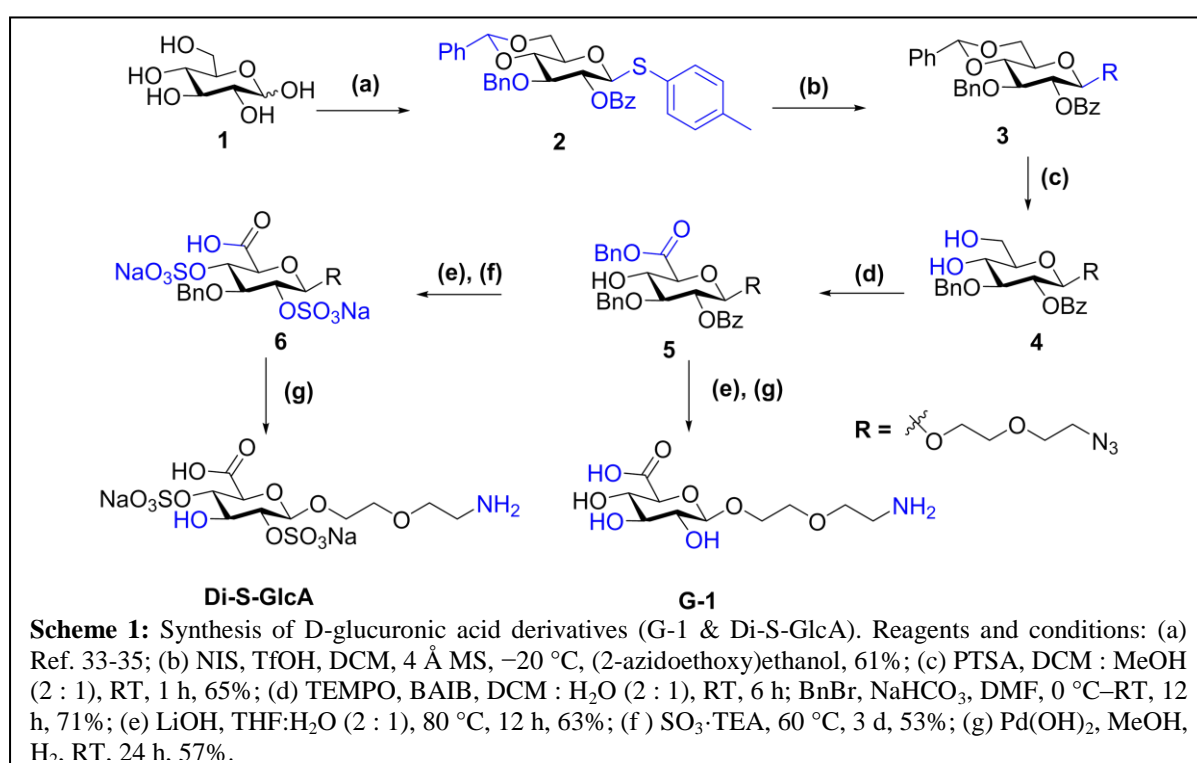
Cell surface carbohydrates play a prominent role in cell–cell interaction, especially in the case of host–pathogen interaction, to modulate biological functions via specific interactions.<sup>1-6</sup> Recent progress in identifying specific carbohydrate epitopes involved in these processes has enabled the development of structure–function relationships of carbohydrates in cell–cell and host–pathogen interactions, which ultimately aid in designing next-generation innovative bioactive materials for clinical applications.<sup>7-13</sup> For example, selective recognition of FimH receptors by mannose saccharide resulted in the development of various high-mannose multivalent scaffolds designed as biosensors and therapeutic molecules for *Escherichia coli* mediated urinary tract infections.<sup>14-19</sup> Similarly, the strong binding preference of PIM-2 receptors expressed on *Pseudomonas aeruginosa* for L-fucose saccharide, and the hemagglutinin protein of influenza virus for sialic acid glycans resulted in the development of various biomaterials for targeting and preventing host invasion by bacterial/viral pathogens.<sup>20-24</sup> *Toxoplasma gondii* (*T. gondii*) is a ubiquitous pathogen capable of infecting all warm blooded animals.<sup>25-26</sup> Human *T. gondii* infection is widespread, and while infection in healthy adults is asymptomatic, infection of newborn and immunocompromised persons can result in lethal toxoplasmosis. The *T. gondii* cell-surface proteins are known to interact with host cells by recognizing and binding to glycosaminoglycans (GAGs) such as heparin/heparin sulfate (HS) present on the host cells.<sup>27-31</sup> Hence, identifying the key HS epitope(s) that selectively recognize the parasite is extremely important for developing therapeutic applications. Seeberger et al. employed a synthetic heparin/HS microarray to identify the crucial HS structural preference of SAG1, ROP, GRA2 proteins (present in the secretory organelles rhoptries and dense granules), which are highly expressed on the cell surface of tachyzoite stage *T. gondii*. Interestingly, it was observed that unnatural 2,4-disulfated iduronic acid (Di-S-IdoA) showed remarkable binding preference to recombinant-SAG1, ROP2 and ROP4 proteins.<sup>32</sup> Moreover, the observed fluorescence intensity of **Di-S-IdoA** was comparable to HS hexasaccharide binding with *T. gondii* surface proteins. Taken together, we wondered whether a minimal sulfated-hexuronic acid derived epitope can preferentially recognize *T. gondii* parasites. Further, we also wanted to find out whether the multivalent display of such a carbohydrate residue can have some therapeutic value. Herein, we performed a systematic investigation of preferential binding and therapeutic potential of two crucial nonsulfated and sulphated hexuronic acid derivatives (GlcA and IdoA) abundantly found in the glycosaminoglycan (GAG) family. We

discovered 2,4-disulfated glucuronic acid (Di-S-GlcA) residue as a potential ligand for *T. gondii* recognition. The multivalent display of Di-S-GlcA significantly inhibited the interaction of the parasite with the host cells, leading to the decreased invasion of the host cells. This observation highlights the therapeutic value of **Di-S-GlcA** in toxoplasmosis.

## 2.2 Results and discussion:

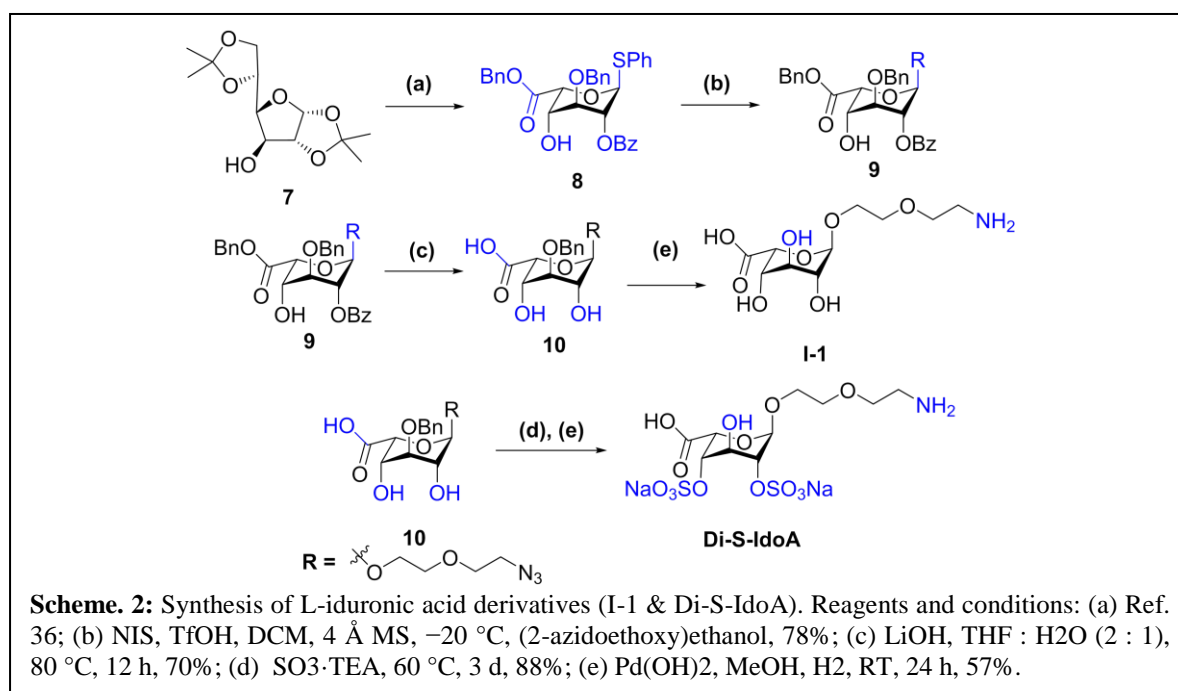
### 2.2.1 Synthesis of D-glucuronic acid derivatives:

To investigate the binding of hexuronic acid to tachyzoite stage *T. gondii*, we first synthesized sulfated and non-sulfated glucuronic and iduronic acid derivatives using different forms of D-glucose building blocks employing standard protection/deprotection strategies.



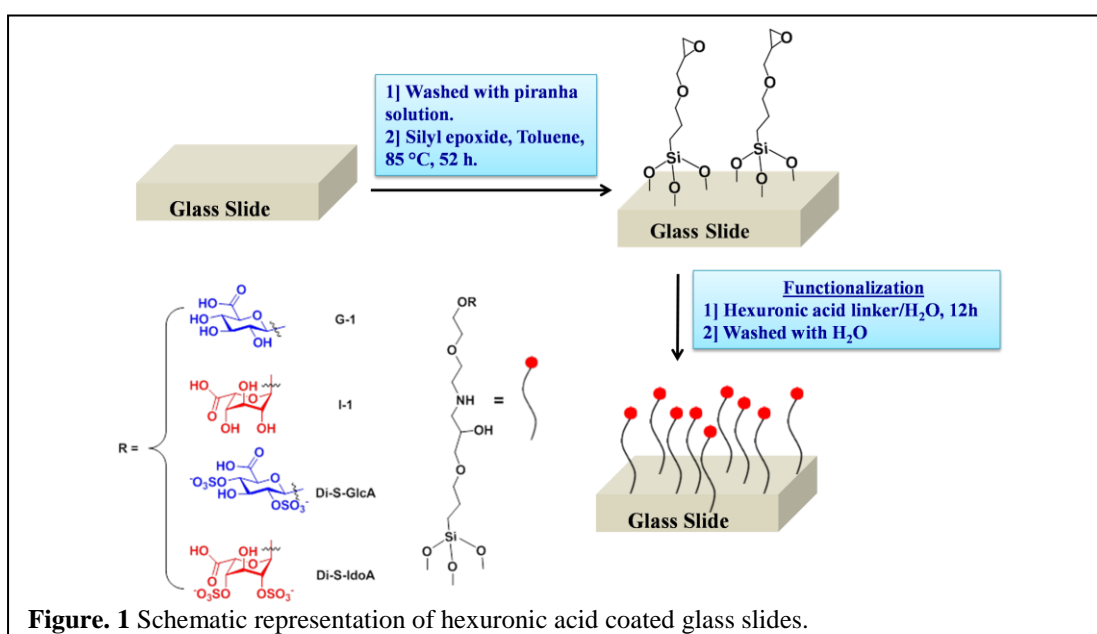
Briefly, D-glucuronic acid derivatives (G-1 and Di-S-GlcA) were synthesized from D-glucose, which was converted into thiotoluene donor **2** using six straight forward chemical reactions.<sup>33-35</sup> Then, **2** was glycosylated with a (2-azidoethoxy)ethanol linker in the presence of NIS (N-iodosuccinamide) and TfOH (trifluoromethanesulphonic acid) promoters to yield 61% of compound **3**. Benzylidene protection in the presence of PTSA (p-toluene sulfonic acid), followed by one-pot TEMPO (2,2,6,6-tetramethyl-1-piperidinyloxy free radical) mediated oxidation and benzyl esterification of compound **3** yielded 71% of **5**. Finally, lithium hydroxide mediated hydrolysis, followed by hydrogenolysis yielded **G-1**, whereas hydrolysis of **5**, and subsequent sulfation, in the presence of a sulfur trioxide triethylamine complex and hydrogenation yielded **Di-S-GlcA** (Scheme 1).

## 2.2.2 Synthesis of L-iduronic acid derivatives:



L-Iduronic acid derivatives (**I-1** and **Di-S-IdoA**) were synthesized from 1,2:5,6-Di-O-isopropylidene- $\alpha$ -D-glucopyranose (**7**) starting material. **7** was converted to the iduronic acid donor (**10**) through straightforward eight-step reactions with 36% of total yield.<sup>36</sup> **10** was glycosylated with the (2-azidoethoxy) ethanol linker, followed by hydrolysis, sulfation and hydrogenolysis to yield **Di-S-IdoA**. While hydrogenolysis of compound **10** yielded 50% of **I-1** (Scheme 2).

## 2.2.3 Functionalization of Hexuronic Acids on Glass slides:

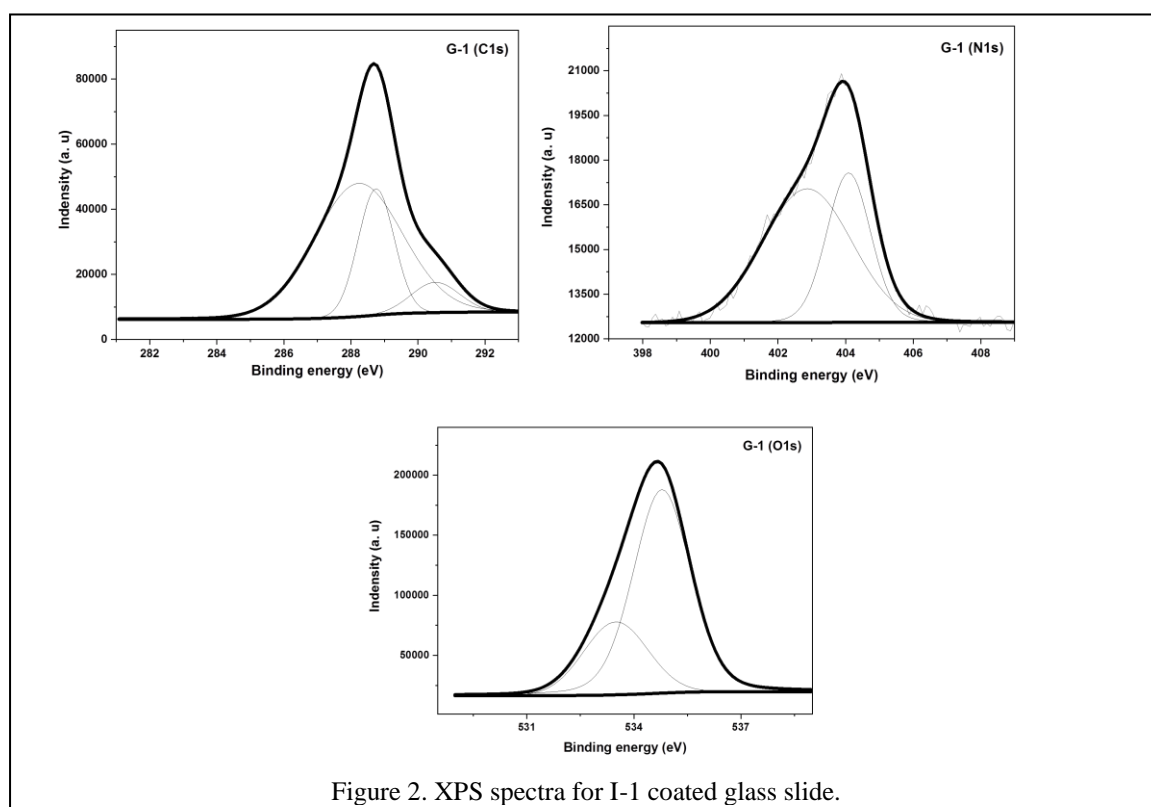


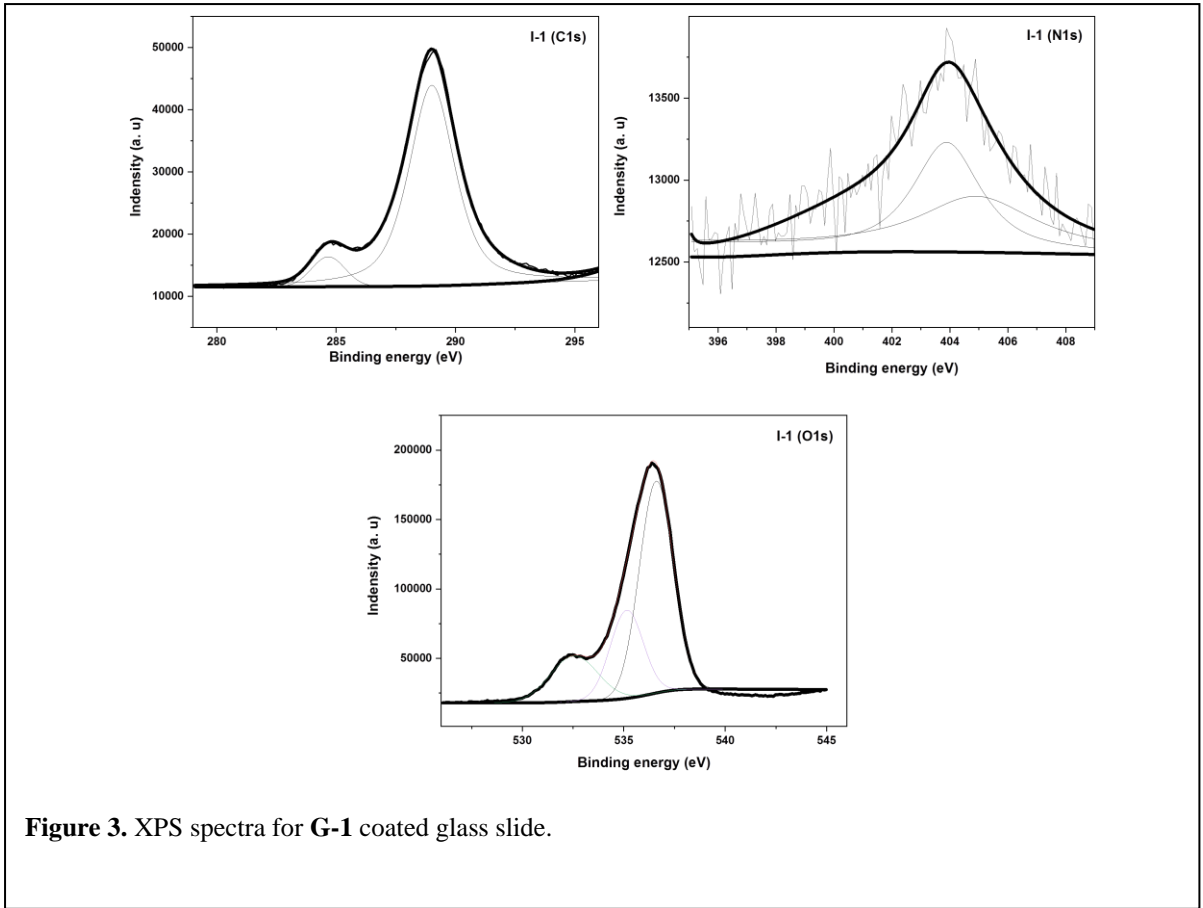


Next, we prepared robust monolayers of hexuronic acid derivatives by attaching **G-1**, **Di-S-GlcA**, **I-1** and **Di-S-IdoA** to epoxide-functionalized glass slides.<sup>37</sup> This was done by first washing glass slides (approx. 1 × 1 cm in size) with piranha solution. These activated slides were dipped in a toluene solution of GOPTMS [(3-glycidyloxypropyl)trimethoxysilane] and heating at 85 °C for 2 days in a pressure tube. After rinsing thoroughly with toluene, the slides were dipped in 0.02 M water solutions of the different sugars for 24 h. Hexuronic acid derivative coated glass slides were rinsed with water and dried (Figure 1).

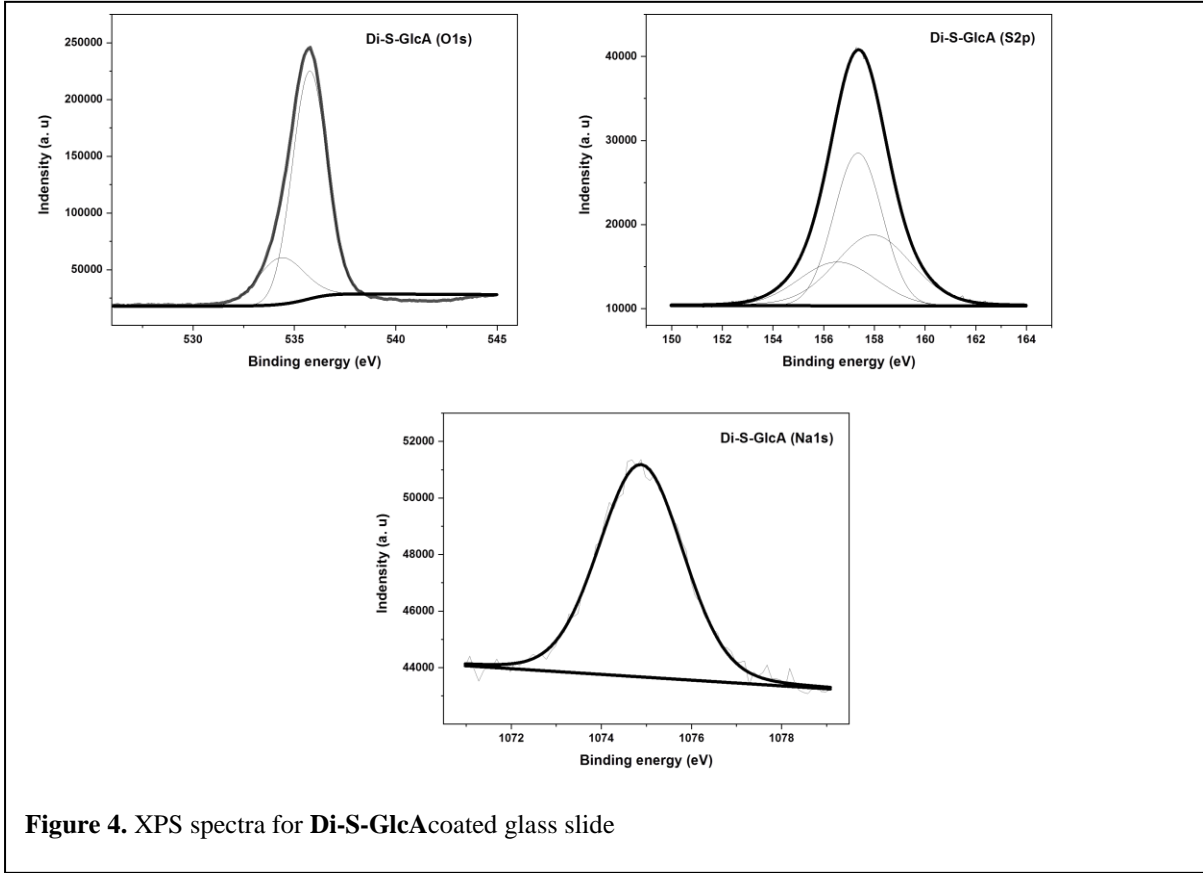
#### 2.2.4 XPS analysis:

XPS analysis of hexuronic acid coated slides showed relative abundance of carbon 1s, oxygen 1s, nitrogen 1s, sulphur 2p and sodium 1s atoms at 288.54 eV, 286.73 eV, 535.74 eV, 157.38 eV and 1074.88 eV peaks,<sup>38</sup> which confirmed the immobilization of sulfated analogs (Fig. S1, S2, S3 and S4†). In contrast, the absence of sulfate and sodium peaks in **G-1** and **I-1** coated glass slides confirmed the presence of non-sulfated hexuronic acid derivatives (Figure 2-5).

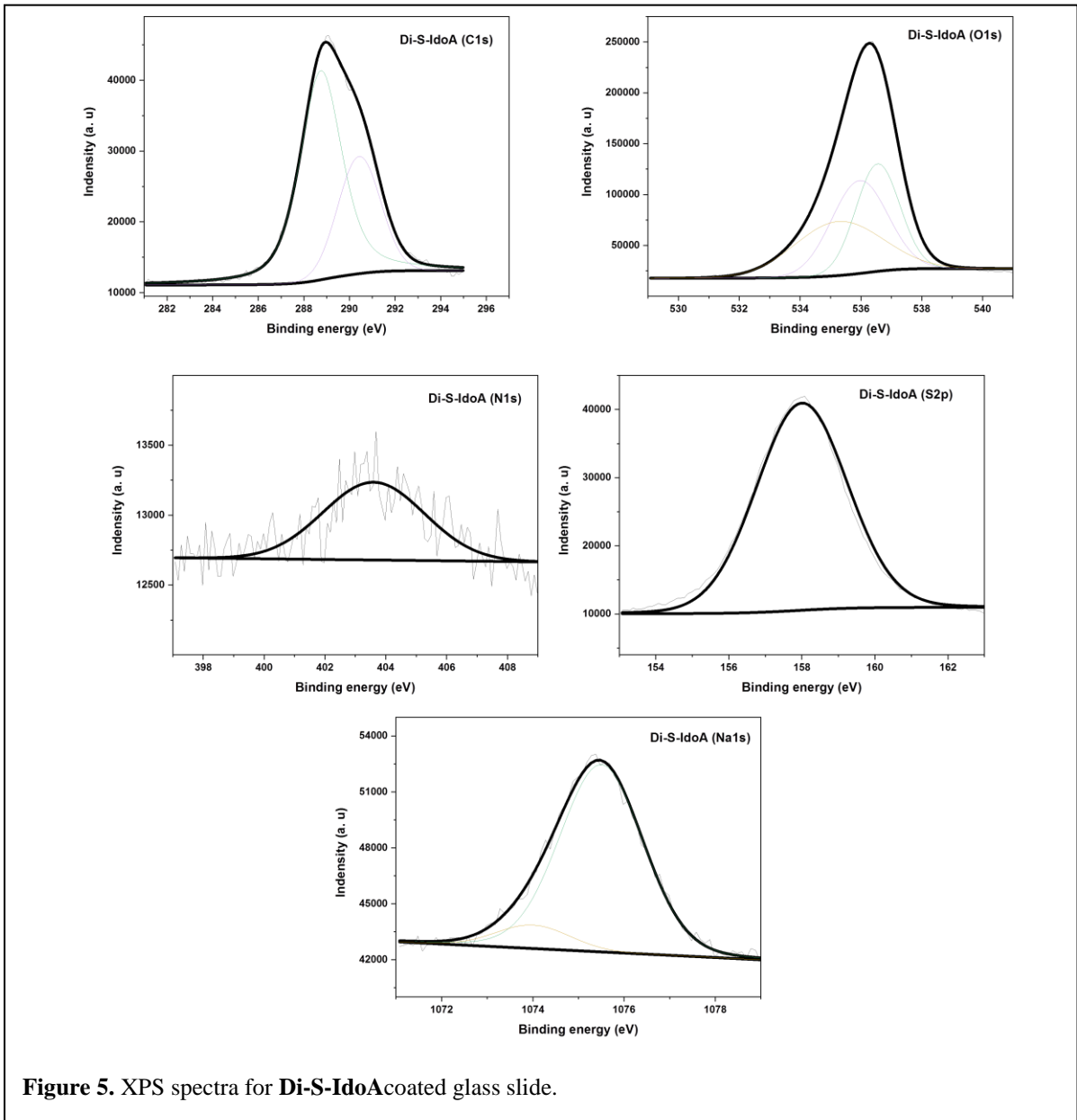
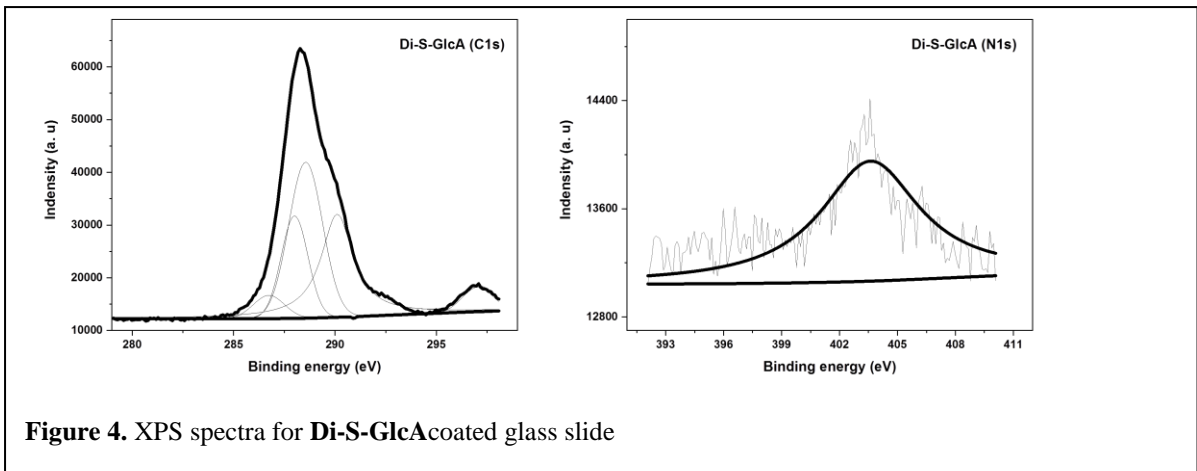




**Figure 3.** XPS spectra for G-1 coated glass slide.



**Figure 4.** XPS spectra for Di-S-GlcA coated glass slide



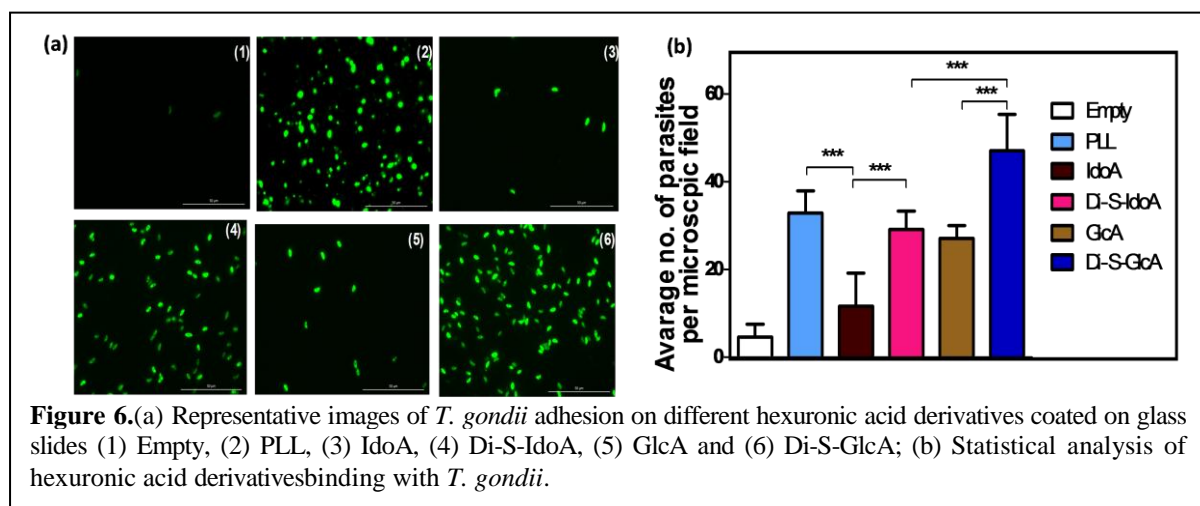
The number of sugar molecules on glass slides was quantified by using a phenol–sulfuric acid

method. As expected, the sulphated hexuronic acid modified slides displayed an almost equal concentration of sugarligands (Table 1).

**Table 1.** The estimated concentration of sugar on 1 X 1 cm<sup>2</sup> glass slides.

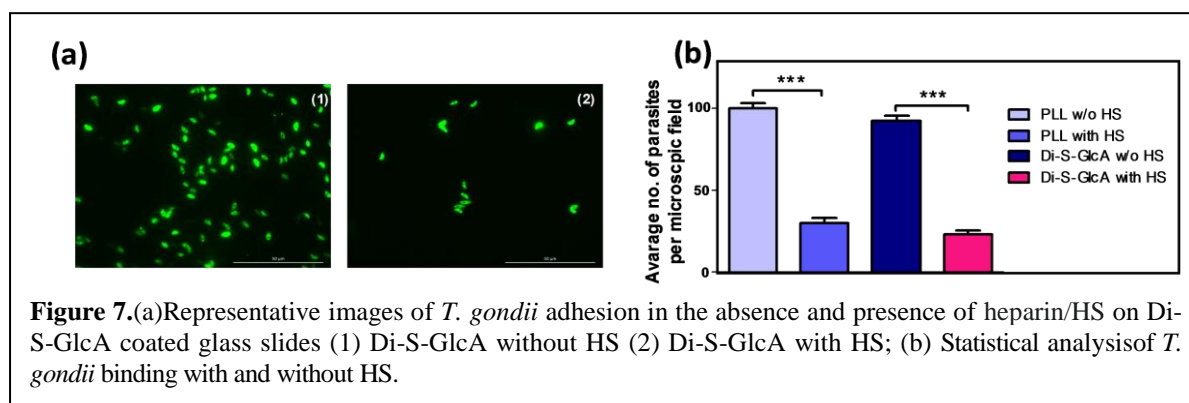
<b>G-1</b>	<b>I-1</b>	<b>Di-S-IdoA</b>	<b>Di-S-GlcA</b>
0.39 ± 0.01 μM	0.61 ± 0.01 μM	0.33 ± 0.02 μM	0.28 ± 0.04 μM

The glass slides modified with hexuronic acid derivatives were used for *T. gondii* binding studies. Our first experiment was to establish the ability of different hexuronic acid derivatives to bind tachyzoite stage *T. gondii*. In this experiment, we incubated the parasites expressing a fluorescent protein on glass slides coated with hexuronic acid derivatives. Polylysine (PLL) coated glass slides and untreated glass slides were used as positive and negative controls respectively. After 1 h incubation of the slides with host cell-free extracellular parasites, the glass slides were rinsed with PBS buffer and bound parasites were fixed with paraformaldehyde solution and imaged. Non-sulfatediduronic acid coated glass coverslips showed (**I-1**) significantly lower binding by the parasites compared to glucuronic acid **G-1**. Among sulfated-hexuronic acid analogs, **Di-S-GlcA** showed strong binding to parasites (Figure 6). We observed nearly 2- fold better binding with **Di-S-GlcA** when compared to **Di-S-idoA**. These results were reproduced in multiple experiments (n = 3).

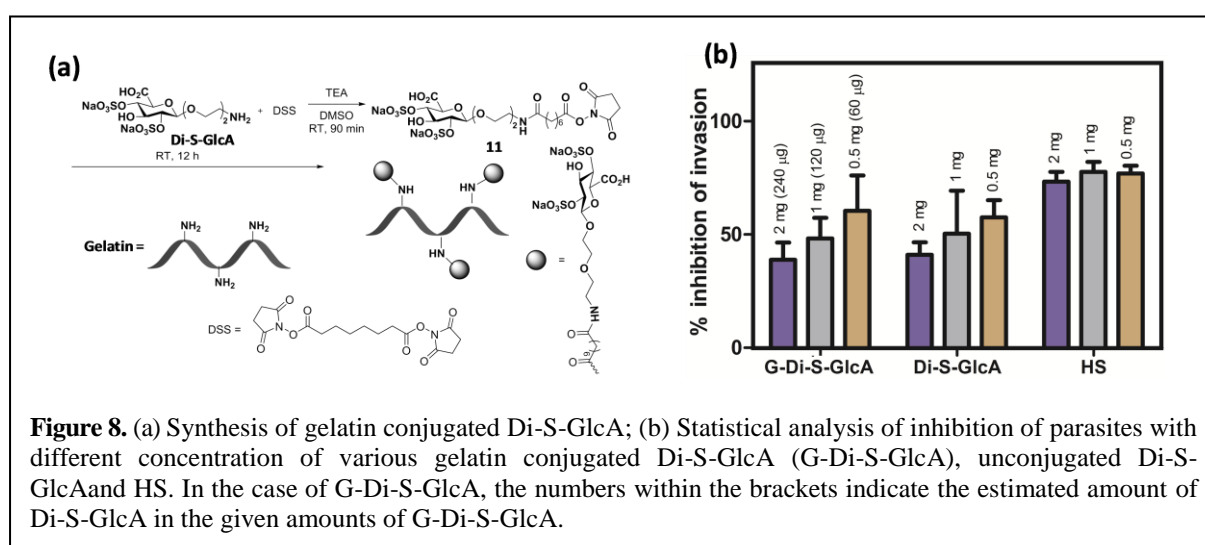


In order to demonstrate that the interaction between Di-S-GlcA and *T. gondii* is indeed heparin/HS associated binding, the binding experiments were performed in the presence and absence of heparin/HS. For this, extracellular parasites were first incubated with heparin/HS at 1 mg ml<sup>-1</sup> concentration in the culture medium. After 1 h, parasites were washed free of

excess heparin/HS and allowed to bind to **Di-S-GlcA** coated glass slides. After 2 h incubation, the glass slides were rinsed with PBS and fixed with paraformaldehyde before microscopy imaging. We observed that heparin treated *T. gondii* showed decreased binding to **Di-S-GlcA** coated slides. This confirmed that the interaction between **Di-S-GlcA** and *T. gondii* is heparin/HS mediated carbohydrate–protein interaction (Figure 7).

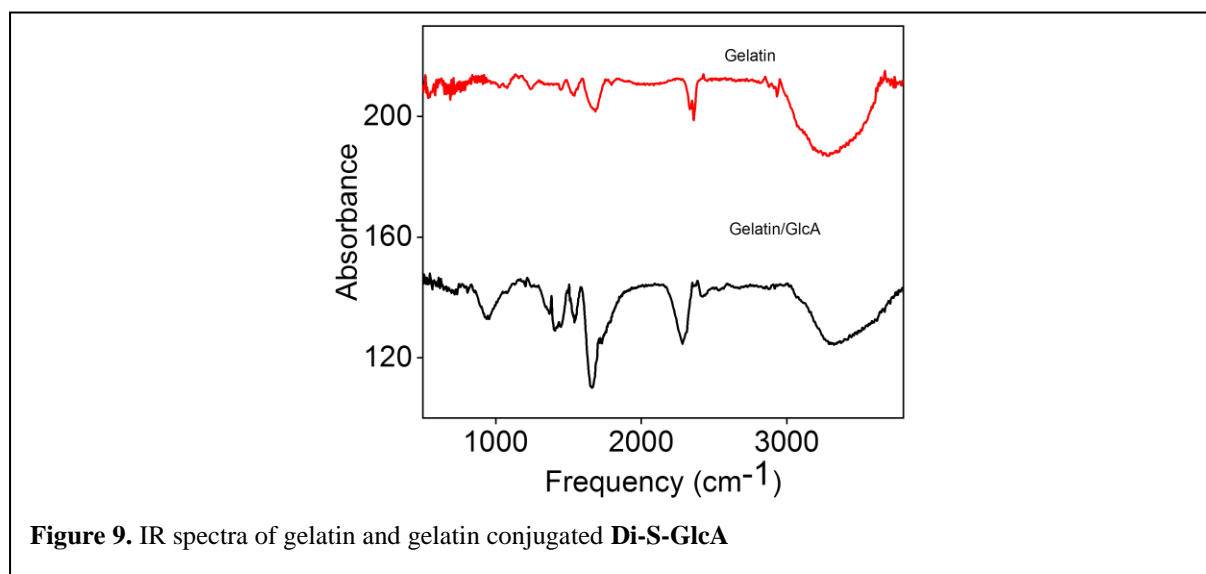


To demonstrate the potential biomedical applications of the **Di-S-GlcA** ligand, a gelatin-based **Di-S-GlcA** multivalent system<sup>39</sup> was assembled. Gelatin is a biocompatible, multivalent and commercially available multivalent probe. It has been extensively used to decorate ligands such as acrylate, thioligands to generate hydrogel materials.<sup>40-42</sup> Hence, it was of interest to synthesize a **Di-S-GlcA–gelatin** conjugate to study the multivalent effect of **Di-S-GlcA** in host–parasite interactions. The glyco–gelatin conjugate was synthesized<sup>43</sup> by mixing a 1: 13 ratio of **Di-S-GlcA** and the DSS linker (disuccinimidylsuberate) in DMSO for 1 h to obtain mono-substitution of the **Di-S-GlcA–DSS** (**11**) linker (Figure 8a).

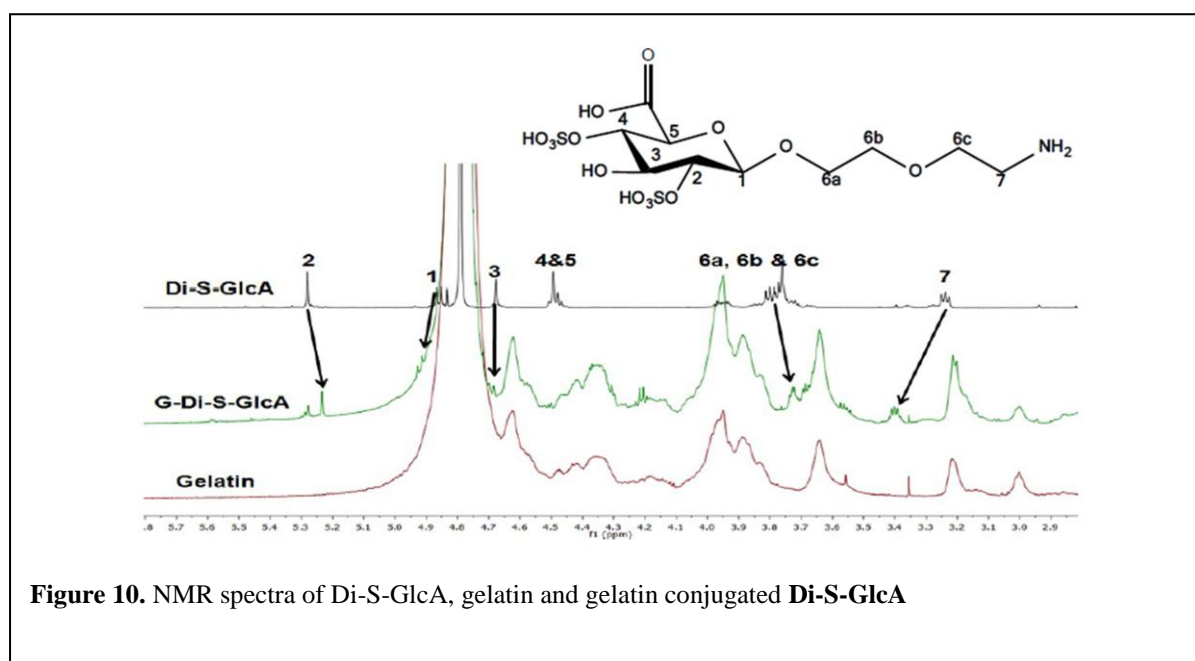


Compound **11** was extracted and characterized by high-resolution mass spectroscopy

(HRMS, Figure 11). Finally, compound **11** was mixed with the gelatin substrate in PBS buffer for 24 h and purified by dialysis against a 10–12 kDa cut-off membrane.

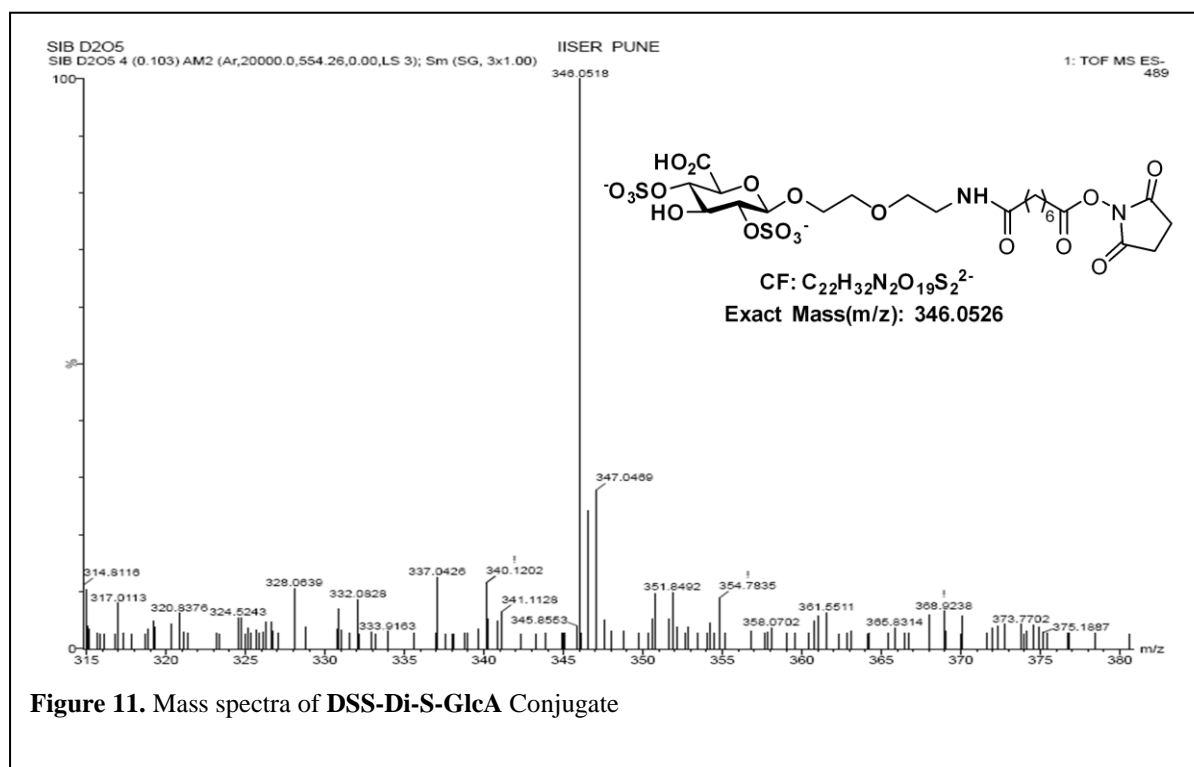


FT-IR peaks observed at 1380 and 1220  $\text{cm}^{-1}$  correspond to the S–O bond symmetric and asymmetric stretching of the sulfate group of **Di-S-GlcA**, which confirmed glyco-conjugation (Figure 9).<sup>44</sup> Similarly, the  $^1\text{H-NMR}$  peak at  $\delta$  4.84 corresponds to the anomeric proton of the **Di-S-GlcA** ligand (Figure 10).

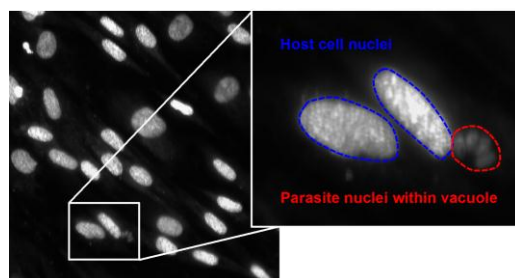


Using  $^1\text{H-NMR}$ , we also quantified the degree of **Di-S-GlcA** substitution. It was observed that 1 mg of gelatin contains approximately 120  $\mu\text{g}$  of **Di-S-GlcA**. The gelatin conjugated **Di-**

**S-GlcA** was then used to test its ability to inhibit host cell invasion by *T. gondii*. Host cell invasion assays were performed after incubating *T. gondii* with 2 mg, 1 mg and 0.5 mg each of **Di-S-GlcA** conjugated gelatin, **Di-S-GlcA** ligand and heparin (HS), respectively.



After 1 h incubation, the parasites were washed free of excess conjugated molecules by centrifugation and allowed to invade host HFF cells grown in 96-well culture plates. After allowing the parasites to invade and replicate within the host cells for 24 h, the infected monolayer was examined for the number of parasites that were able to enter the host cells and establish a productive infection. This was achieved by counting the number of parasitophorous vacuoles in infected cells with actively replicating parasites (Figure 12. for a representative microscopy image). Since only parasites that have invaded the host cells can replicate, this will help in determining the ability of the gelatin-based **Di-S-GlcA** multivalent system to inhibit parasite invasion of host cells. It was observed that the **G-Di-S-GlcA** substrate induced strong inhibition of parasite invasion compared to monovalent **Di-S-GlcA** and heparin, respectively, highlighting the multivalent effect of **G-Di-S-GlcA**. We observed that the **G-Di-S-GlcA** substrate was approximately 10–12 times more potent in inhibiting host cell invasion by the parasite in comparison with the monovalent **Di-S-GlcA** ligand (Figure 8b). These results demonstrate the therapeutic potential of **Di-S-GlcA**, with respect to toxoplasmosis.



**Figure 12.** Microscopic image showing DAPI stained *T. gondii* infected host cell monolayers imaged for the invasion assays (See Figure 6b in main text). This is a representative image shown for control in which untreated parasites were allowed to invade the host cells and replicate for 24 h before fixing and imaging the infected cells. The detailed description of the experiment and imaging method is given in the Experimental Section of the main text. The microscopic image shows fluorescent DAPI staining in both host cell nuclei (which are bigger in size) and parasite nuclei (which are smaller in size and clustered together depending on the number of parasites within a single parasitophorous vacuole). The inset shows an expanded view of the host and parasite nuclei. The images were collected from glass bottom 96 well plates using a 63X water emulsion objective fitted to the Operetta CLS High-Content Analysis System.

## 2.3 Conclusions:

We have identified **Di-S-GlcA** as a potential carbohydrate scaffold that can recognize and bind to surface exposed proteins on the virulent tachyzoite stage of *T. gondii* parasites. The interaction between the parasites and carbohydrate moiety is heparin-mediated carbohydrate–protein interaction. The multivalent display of **Di-S-GlcA** on gelatin scaffolds demonstrated effective inhibition of host–parasite invasion. The possibility of using **Di-S-GlcA** multivalent scaffolds in the therapeutic management of toxoplasmosis is currently under study.

## 2.4 Experimental section:

### 2.4.1 General information:

All chemicals were reagent grade and used as supplied except where noted. Analytical thin layer chromatography (TLC) was performed on Merck silica gel 60 F254 plates (0.25 mmol). Compounds were visualized by UV irradiation or dipping the plate in CAM/ninhydrin solution followed by heating. Column chromatography was carried out using force flow of the indicated solvent on a FlukaKieselgel 60 (230–400 mesh). <sup>1</sup>H and <sup>13</sup>C NMR spectra were recorded on a Jeol 400 MHz spectrometer, with a cryoprobe using residual solvents signals as an internal reference (CDCl<sub>3</sub> δH, 7.26 ppm, δC 77.3 ppm, CD<sub>3</sub>OD δH 3.31 ppm, δC 49.0 ppm and D<sub>2</sub>O δH 4.79 ppm). The chemical shifts (δ) are reported in ppm and coupling constants (J) in Hz. XPS experiments were performed on a VG Micro Tech ESCA3000 instrument at a pressure of < 1 × 10<sup>−9</sup> Torr. The overall resolution was limited to the bandwidth of the X-ray source (~1 eV). The spectra were recorded with monochromatic



Al K $\alpha$  radiation at a pass energy of 50 eV and an electron take off angle of 60°.

**Synthesis of compound 3:** Compound **2** (1.5 g, 2.64 mmol) and (2-azidoethoxy)ethanol (0.34 g, 2.64 mmol) were dissolved in anhydrous DCM and cooled to -20 °C. To this solution, NIS (0.65 g, 2.90 mmol) and TMSOTf (0.095 mL, 0.52 mmol) were added and stirred for 1 h. After the completion of the reaction, the reaction mixture was quenched with triethylamine, filtered on a Celite bed and the solvent was evaporated. The crude mixture was purified by column chromatography (EtOAc/hexane = 0.5/1) to afford **3** (0.92 g, 61%) as a white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  8.07–8.04 (m, 2H), 7.66–7.61 (m, 1H), 7.57–7.39 (m, 7H), 7.21–7.11 (m, 5H), 5.65 (s, 1H), 5.34 (t, J = 8.4 Hz, 1H), 4.87 (d, J = 12.1 Hz, 1H), 4.76–4.73 (m, 2H), 4.43 (dd, J = 10.5, 5.0 Hz, 1H), 4.01–3.96 (m, 1H), 3.94–3.86 (m, 3H), 3.77 (m, 1H), 3.63–3.53 (m, 3H), 3.49–3.44 (m, 2H), 3.15–3.01 (m, 2H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  165.18, 137.95, 137.33, 133.31, 129.94, 129.16, 128.50, 128.40, 128.29, 128.16, 127.67, 126.14, 101.99, 101.39, 81.79, 77.94, 74.14, 73.52, 70.46, 70.17, 69.64, 68.80, 66.40, 50.62. IR (cm<sup>-1</sup>, CH<sub>2</sub>Cl<sub>2</sub>), 2875, 2103, 1727, 1602. HRMS m/z calculated for C<sub>24</sub>H<sub>29</sub>N<sub>3</sub>O<sub>7</sub>: 471.2006; found: 471.2054.

**Synthesis of compound 4:** Compound **3** (0.9 g, 1.56 mmol) and p-toluenesulfonic acid (0.35 g, 1.87 mmol) were dissolved in DCM/MeOH [1/1 (v/v), 70 mL] and stirred at room temperature for 1 h. After the completion of the reaction, the mixture was quenched with triethylamine. The mixture was extracted with DCM and washed with NaHCO<sub>3</sub> and brine solution, dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was evaporated under reduced pressure. The residue was washed with hexane and filtered to afford **4** (0.5 g, 65%) as a white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  8.10–8.07 (m, 2H), 7.64–7.60 (m, 1H), 7.52–7.47 (m, 2H), 7.25–7.21 (m, 5H), 5.27 (dd, J = 9.3, 7.9 Hz, 1H), 4.78–4.64 (m, 3H), 4.00–3.93 (m, 2H), 3.87–3.71 (m, 4H), 3.63–3.54 (m, 2H), 3.50–3.45 (m, 3H), 3.17–3.02 (m, 2H), 2.07 (bs, 1H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  165.29, 137.90, 133.38, 129.91, 129.85, 128.58, 128.11, 128.03, 101.48, 82.37, 75.44, 74.67, 73.73, 70.52, 70.49, 70.13, 69.48, 62.40, 50.62. HRMS m/z calculated for C<sub>17</sub>H<sub>25</sub>N<sub>3</sub>O<sub>7</sub>: 383.1693; found: 383.1604.

**Synthesis of compound 5:** Compound **4** (0.5 g, 1.02 mmol) is dissolved in water and DCM [1/2 (v/v), 15 mL] and the 2,2,6,6-tetramethyl-1-piperidinyloxy free radical (TEMPO, 32 mg, 0.2 mmol) and bis(acetoxy)iodobenzene (BAIB, 0.82 g, 2.56 mmol) were added at room temperature and stirred for 6 h. After the completion of the reaction, the mixture was quenched with saturated NH<sub>4</sub>Cl solution and extracted with DCM. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered, and the solvent was evaporated under reduced pressure. The residue was taken as such for the next step without further purification. The crude mixture

was dissolved in DMF (10 mL) and NaHCO<sub>3</sub> (0.51 g, 2.99 mmol) was added. The reaction mixture was cooled to 0 °C and then benzyl bromide (0.35 mL, 2.99 mmol) was added. After stirring for 12 h, the reaction mixture was diluted with ethyl acetate and washed with brine solution, dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was evaporated under reduced pressure. The residue was purified by column chromatography (EtOAc/hexane = 1/5) to afford **5** (0.42 g, 71%) as a white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.05–8.03 (m, 2H), 7.64–7.59 (m, 1H), 7.50–7.43 (m, 2H), 7.44–7.34 (m, 5H), 7.23–7.16 (m, 5H), 5.33–5.26 (m, 3H), 4.77 (d, J = 2.8 Hz, 2H), 4.72 (d, J = 7.8 Hz, 1H), 4.10 (ddd, J = 9.7, 8.8, 2.7 Hz, 1H), 4.02–3.93 (m, 2H), 3.79–3.70 (m, 2H), 3.64–3.51 (m, 2H), 3.44 (t, J = 5.1 Hz, 2H), 3.13–3.07 (m, 1H), 3.06–2.97 (m, 1H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 169.06, 165.15, 137.85, 135.03, 133.33, 129.88, 129.84, 128.78, 128.70, 128.51, 128.43, 128.14, 127.84, 101.73, 80.82, 74.62, 74.45, 73.02, 72.09, 70.46, 70.10, 69.65, 67.61, 50.63. IR (cm<sup>-1</sup>, CH<sub>2</sub>Cl<sub>2</sub>) 3428, 2877, 2104, 1724, 1601. HRMS m/z calculated for C<sub>24</sub>H<sub>29</sub>N<sub>3</sub>O<sub>8</sub>: 487.1955; found: 487.1969.

**Synthesis of compound G-1:** Compound **5** (0.11 g, 0.18 mmol) and lithium hydroxide (0.42 g, 3.72 mmol) were dissolved in a mixture of THF and water [2/1 (v/v), 6 mL] and refluxed at 80 °C for 12 h. After the reaction completion, the mixture was cooled to room temperature and neutralized with Amberlite® IR 120 resin, filtered, and the solvent was evaporated under reduced pressure. The residue was purified by bond elution using water and then lyophilized to afford the crude product (0.13 g, 80%) as a white solid. The crude product was subjected to hydrogenolysis using Pd(OH)<sub>2</sub> on charcoal with hydrogen gas in methanol (3 mL). After 12 h, the mixture was filtered on Whatman 42 filter paper. The residue was purified by bond elution using water and then lyophilized to afford **G-1** (37 mg, 63%) as a white solid. <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) δ 4.44 (d, J = 7.9 Hz, 1H), 4.03–3.98 (m, 1H), 3.82–3.77 (m, 1H), 3.74–3.63 (m, 5H), 3.47–3.41 (m, 2H), 3.31–3.24 (m, 1H), 3.17–3.09 (m, 2H). <sup>13</sup>C NMR (100 MHz, D<sub>2</sub>O) δ 175.79, 102.13, 76.03, 75.54, 72.90, 71.75, 69.53, 68.99, 66.74, 39.16. HRMS m/z calculated for C<sub>10</sub>H<sub>19</sub>NO<sub>8</sub>: 281.1111; found: 281.1084.

**Synthesis of compound 6:** Compound **5** (0.24 g, 0.4 mmol) and lithium hydroxide (0.42 g, 1.01 mmol) were dissolved in the mixture of THF and water [2/1 (v/v), 6 mL] and refluxed at 80 °C for 12 h. After the reaction completion, the mixture was cooled to room temperature and neutralized with Amberlite® IR 120 resin, filtered, and the solvent was evaporated under reduced pressure. The residue was purified by bond elution using water and then lyophilized to afford (0.13 g, 80%) as a white solid. The crude product was subjected to the next step without purification. The crude product (0.12 g, 0.3 mmol) was dissolved in DMF (30 mL) and the SO<sub>3</sub>·TEA complex (1.09 g, 6.04 mmol) was added. The reaction mixture was

refluxed at 60 °C for 48 h. The solvent was evaporated and purified by column chromatography and lyophilized to afford **6** (0.09 g, 53%) as a white solid. <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) δ 7.52–7.43 (m, 2H), 7.41–7.33 (m, 3H), 5.17 (s, 1H), 4.85 (d, J = 9.1 Hz, 1H), 4.73 (s, 1H), 4.69–4.62 (m, 2H), 4.32 (dd, J = 9.2, 4.6 Hz, 1H), 4.24 (d, J = 4.7 Hz, 1H), 3.87–3.78 (m, 1H), 3.70–3.58 (m, 5H), 3.37–3.30 (m, 2H). <sup>13</sup>C NMR (100 MHz, D<sub>2</sub>O) δ 174.62, 137.05, 128.87, 128.62, 128.33, 106.94, 82.38, 80.76, 79.36, 74.49, 72.79, 69.41, 67.78, 50.16. IR (cm<sup>-1</sup>, CH<sub>2</sub>Cl<sub>2</sub>) 2884, 2108, 1726, 1601, 1496. HRMS m/z calculated for C<sub>17</sub>H<sub>21</sub>N<sub>3</sub>Na<sub>2</sub>O<sub>14</sub>S<sub>2</sub>: 601.0260; found: 601.0383.

**Synthesis of compound Di-S-GlcA:** Compound **6** (60 mg, 0.10 mmol) and 10% Pd(OH)<sub>2</sub> on charcoal (5 mg) were dissolved in methanol (3 mL) and stirred at room temperature for 12 h in the presence of hydrogen gas. After hydrogenolysis, the mixture was filtered on Whatman 42 filter paper. The residue was purified by bond elution using water and then lyophilized to afford **Di-S-GlcA** (27 mg, 57%) as a white solid. <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) δ 5.20 (s, 1H), 4.76 (d, J = 7.5 Hz, 1H), 4.60 (s, 1H), 4.44–4.38 (m, 2H), 3.91–3.84 (m, 1H), 3.74–3.67 (m, 5H), 3.21–3.12 (m, 2H). <sup>13</sup>C NMR (100 MHz, D<sub>2</sub>O) δ 174.63, 106.48, 85.20, 82.42, 76.17, 73.04, 69.57, 67.64, 66.31, 39.20, 22.42. HRMS m/z calculated for C<sub>10</sub>H<sub>17</sub>NNa<sub>2</sub>O<sub>14</sub>S<sub>2</sub>: 484.9886; found: 484.9871.

**Synthesis of compound 9:** Compound **8** (1.07 g, 1.87 mmol) and (2-azidoethoxy)ethanol (0.29 g, 2.25 mmol) were dissolved in anhydrous DCM (20 mL) with freshly dried 4 Å MS and stirred at room temperature for 1 h. NIS (0.63 g, 2.81 mmol) and TMSOTf (0.067 ml, 0.375 mmol) were added at -5 °C and stirred for 30 min. After the completion of the reaction, the reaction mixture was quenched with triethylamine and filtered through Celite. The organic layer was washed with Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> and brine solution, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and then the solvent was evaporated under reduced pressure. The residue was purified by column chromatography (EtOAc/hexane = 1/4) to afford **9** (0.86 g, 78%) as a white sticky solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.00 (dd, J = 8.3, 1.4 Hz, 2H), 7.62–7.57 (m, 1H), 7.48–7.43 (m, 2H), 7.40–7.28 (m, 10H), 5.32 (d, J = 12.3 Hz, 1H), 5.27–5.26 (m, 1H), 5.24 (d, J = 12.3 Hz, 1H), 5.17 (s, 1H), 5.02 (d, J = 1.7 Hz, 1H), 4.84 (d, J = 11.6 Hz, 1H), 4.65 (d, J = 11.6 Hz, 1H), 4.16–4.12 (m, 1H), 3.98–3.93 (m, 1H), 3.89 (td, J = 3.0, 1.3 Hz, 1H), 3.76–3.67 (m, 3H), 3.63–3.53 (m, 2H), 3.19 (t, J = 5.0 Hz, 2H), 2.80 (d, J = 11.6 Hz, 1H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 169.54, 165.11, 137.70, 135.49, 133.84, 129.89, 129.14, 128.77, 128.73, 128.55, 128.52, 128.44, 128.02, 127.88, 99.00, 74.75, 72.16, 70.33, 70.25, 68.48, 68.29, 67.96, 67.47, 67.17, 50.83. IR (cm<sup>-1</sup>, CH<sub>2</sub>Cl<sub>2</sub>) 2875, 2104, 1722, 1607. HRMS m/z calculated for C<sub>31</sub>H<sub>33</sub>N<sub>3</sub>O<sub>9</sub>Na: 614.2214; found: 614.2120.

**Synthesis of compound 10:** Compound **5** (0.11 g, 0.18 mmol) was dissolved in the mixture of THF and water [2/1 (v/v), 6 mL] and lithium hydroxide (0.42 g, 3.72 mmol) was added and then refluxed at 80 °C. After stirring for 12 h the mixture was cooled to room temperature and neutralized with Amberlite® IR 120 resin, filtered and the solvent was evaporated under reduced pressure and then purified through a Sephadex C-18 column to afford compound **10** (0.05 g, 70%). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 7.41–7.25 (m, 5H), 4.75–4.71 (m, 2H), 4.65–4.60 (m, 1H), 4.08 (s, 1H), 3.93–3.87 (m, 1H), 3.81–3.80 (m, 1H), 3.71–3.66 (m, 5H), 3.64–3.56 (m, 1H), 3.23 (t, J = 4.9 Hz, 2H). <sup>13</sup>C NMR (100 MHz, CD<sub>3</sub>OD) δ 173.88, 139.69, 129.32, 128.86, 128.67, 102.77, 78.04, 72.92, 71.28, 71.21, 69.92, 69.67, 69.24, 68.10, 51.75. IR (cm<sup>-1</sup>, CH<sub>2</sub>Cl<sub>2</sub>) 3427, 2875, 2104, 1724, 1604. HRMS m/z calculated for C<sub>17</sub>H<sub>23</sub>N<sub>3</sub>O<sub>8</sub>: 397.1485; found: 397.1389.

**Synthesis of compound I-1:** Compound **9** (0.200 g, 0.34 mmol) was dissolved in THF (2 mL) and water (2 mL). Lithium hydroxide (1 M solution in water, 1 mL, 1.35 mmol) was added at room temperature. After 16 h, the reaction was neutralized by adding DOWEX50WX-200 H<sup>+</sup> resin and filtered and the solvent was evaporated under reduced pressure and then the crude product proceeded further without purification. The crude product was dissolved in anhydrous methanol (2 mL), and 20% Pd(OH)<sub>2</sub> on carbon (0.017 g, 0.125 mmol) was added under a hydrogen atmosphere, and then the mixture was stirred at room temperature for 24 h. The mixture was filtered and the solvent was evaporated under reduced pressure and purified through a Sephadex C-18 column using water and then lyophilized to afford **I-1** (0.031 g, 88%). <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) δ 4.86 (d, J = 4.2 Hz, 1H), 4.62 (d, J = 3.6 Hz, 1H), 3.91–3.86 (m, 2H), 3.74 (dd, J = 6.9, 5.1 Hz, 2H), 3.71–3.68 (m, 4H), 3.51 (dd, J = 6.4, 4.2 Hz, 1H), 3.14 (t, J = 5.1 Hz, 2H). <sup>13</sup>C NMR (100 MHz, D<sub>2</sub>O) δ 173.78, 101.00, 71.07, 70.15, 69.96, 69.85, 69.67, 67.92, 66.34, 39.08. HRMS m/z calculated for C<sub>10</sub>H<sub>19</sub>NO<sub>8</sub>: 281.1111; found: 281.1108.

**Synthesis of compound Di-S-IdoA:** Compound **10** (0.12 g, 0.3 mmol) was dissolved in DMF (30 mL) and the SO<sub>3</sub>·TEA complex (1.09 g, 6.04 mmol) was added. The reaction mixture was refluxed at 60 °C for 48 h. After the completion of the reaction, the solvent was evaporated and it proceeded further. The sulfated product was dissolved in methanol (3 mL) and Pd(OH)<sub>2</sub> on charcoal was added under hydrogen atmosphere. After 12 h the mixture was filtered on Whatman 42 filter paper. The residue was purified by bondelution using water and then lyophilized to afford **Di-S-IdoA** (27 mg, 57%) as a white solid. <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) δ 5.14 (s, 1H), 4.59 (d, J = 2.0 Hz, 1H), 4.57 (d, J = 2.8 Hz, 1H), 4.48 (dt, J = 2.8, 1.4 Hz, 1H), 4.21 (dt, J = 2.5, 1.2 Hz, 1H), 3.84 (ddd, J = 13.9, 8.6, 4.4 Hz, 2H), 3.73 (q, J = 4.7, 4.2

Hz, 4H), 3.20–3.18 (m, 2H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{D}_2\text{O}$ )  $\delta$  174.88, 98.67, 74.60, 73.28, 69.86, 67.51, 66.46, 66.37, 66.26, 39.18. HRMS  $m/z$  calculated for  $\text{C}_{10}\text{H}_{17}\text{NNa}_2\text{O}_{14}\text{S}_2$ : 484.9886; found: 484.9873.

#### 2.4.2 Immobilization of hexuronic acids on glass slides:

Glass slides (approx.  $1 \times 1$  cm in dimension) were washed with piranha solution (Caution: Use carefully as piranha solution reacts violently) and immediately dipped in a 0.1 M solution of GOPTMS [(3-glycidyloxypropyl)trimethoxysilane] in toluene (2 mL). The substrates were heated at  $85^\circ\text{C}$  for 2 days in a pressure tube, cooled to room temperature and rinsed with toluene to remove excess of GOPTMS. The coated glass slides were dipped in 0.02 M water solutions of compounds **G-1**, **I-1**, **Di-S-GlcA** and **Di-S-IdoA** for 12 hours followed by rinsing with water to remove excess of sugars and to block free epoxide groups.

#### 2.4.3 XPS analysis:

The XPS technique was used to confirm the binding of molecules **G-1**, **I-1**, **Di-S-GlcA** and **Di-S-IdoA** on the glass slides. The spectra of molecule **G-1** contain the following peaks: carbon 1s – 288.74 eV, and 290.51 eV; nitrogen 1s – 402.87 eV; oxygen 1s – 533.49 eV (Fig. S1<sup>†</sup>). Similarly, for molecule **I-1**: carbon 1s – 284.66 eV and 288.73; nitrogen 1s – 403.89 eV; oxygen 1s – 532.54 eV (Fig. S2<sup>†</sup>). The spectra of molecule **Di-S-GlcA** contain the following peaks: carbon 1s – 288.54 eV and 290.02 eV; nitrogen 1s – 286.73 eV; oxygen 1s – 535.74 eV; sulfur 2p – 157.38 eV; sodium 1s – 1074.88 eV (Fig. S3<sup>†</sup>). Similarly, for molecule **Di-S-IdoA**: carbon 1s – 288.74 eV and 290.41 eV; nitrogen 1s – 403.6 eV; oxygen 1s – 535.95 eV; sulfur 2p – 158.0 eV; sodium 1s – 1075.5 eV (Fig. S4<sup>†</sup>). These data indicate clearly that molecules **G-1**, **I-1**, **Di-S-GlcA** and **Di-S-IdoA** are bound on the glass slides.

#### 2.4.4 Quantification of sugar on glass slides:

The concentration of hexuronic acid derivatives on the glass slide was determined by the phenol–sulfuric acid method. Hexuronic acid derivative coated glass slides were dipped in concentrated sulfuric acid (750  $\mu\text{L}$ , 100%) and aqueous phenol solution (5% w/v, 100  $\mu\text{L}$ ) in a pressure tube and heated to  $80^\circ\text{C}$  for 10 min. The glass slides were removed and the absorbance coefficient at 490 nm was measured. Blank slides and epoxide slides were used as a control. The hexuronic acid derivative concentration was estimated by comparing the absorption of the sample with a standard curve. The quantified values are given in Table S1.<sup>†</sup>

#### 2.4.5 Gelatin conjugation:

The DSS linker (19.7 mg, 56.6  $\mu\text{mol}$ ) was dissolved in dimethylsulfoxide (DMSO, 1 mL) and 15  $\mu\text{L}$  of triethylamine was added. **Di-S-GlcA** (2 mg, 4.12  $\mu\text{mol}$ ) was added dropwise which was dissolved in 0.5 mL of DMSO. The mixture was stirred at room temperature for 90 min and PBS buffer was added, which was washed with chloroform (3 $\times$ ) and then centrifuged at 3000g. The PBS buffer layer was stirred at room temperature and gelatin was added. The reaction mixture was stirred for 12 h and the PBS buffer was transferred to a 10–12 kDa dialysis membrane, and then dialyzed in water for three days.

#### 2.4.6 Toxoplasma gondii culture protocol:

Tachyzoite stage *T. gondii* parasites (RH strain), expressing the cytoplasmic yellow fluorescent protein (YFP), were used in all experiments. RH-YFP parasites were cultured as previously reported.<sup>45</sup> Briefly, confluent monolayers of human foreskin fibroblast (HFF) cells (obtained from ATCC, USA) were used to propagate parasites in DMEM supplemented with 1% heat inactivated fetal bovine serum (FBS), 2 mM glutamax and 50  $\mu\text{g ml}^{-1}$  gentamicin. Parasites were harvested after 48 h of intracellular replication by scraping the host cell monolayer and passing the infected cell suspension through a 25-gauge needle to mechanically disrupt the HFF cells to release the parasites. The lysed cell suspension was filtered through a 3  $\mu\text{m}$  nucleopore membrane (Whatman, GE Healthcare, USA) to obtain an intact parasite suspension devoid of host cell debris. The purified parasites were further used in different experiments.

#### 2.4.7 Binding assays with *T. gondii*:

Empty (uncoated) and coated (polylysine (PLL), **IdoA**, **Di-S-IdoA**, **GlcA** and **Di-S-GlcA**) glass coverslips were prepared and placed in 6 well plates, in duplicate. RH-YFP parasites were harvested, counted and suspended in culture medium to obtain 10<sup>6</sup> parasites per ml. To each glass coverslip, 100  $\mu\text{L}$  of the parasite suspension was added and the parasites were allowed to bind to the glass coverslips for 2 h in a 37 °C humidified incubator maintaining 5% CO<sub>2</sub>. In assays where heparin was used to compete for binding, the parasite suspension was divided into two equal portions and to one portion heparin was added to a final concentration of 1 mg ml<sup>-1</sup>. The parasite suspensions were then incubated for 1 h before adding to the glass coverslips as described above. After incubation, the glass coverslips were washed, fixed and mounted on glass slides, and YFP fluorescence was imaged using the EVOS Cell Imaging System (Thermo Fisher Scientific, USA).

#### 2.4.8 Invasion assays:

Extracellular *T. Gondii* tachyzoites were incubated with 2 mg, 1 mg and 0.5 mg of **Di-S-GlcA** conjugated gelatin, heparin and monovalent **Di-S-GlcA** ligands respectively for 1 h under optimal conditions. Following incubation, the parasites were washed and added to 96-well black (flat glass bottom) tissue culture plates (Corning, USA), containing confluent HFF cells. In each well, 1000 parasites were added and allowed to invade and replicate within the HFF host cells for 24 h under optimal growth conditions. In parallel, as a positive control for invasion and replication, 1000 untreated parasites were also allowed to invade and replicate within the host cells in the same plate. After 24 h, the culture plate was washed once with DMEM and the infected monolayer was fixed with 4% paraformaldehyde. The parasites were identified using the fluorescent dye 4',6-diamidino-2-phenylindole (DAPI) which specifically stains nuclear DNA within the cell. The plate was then imaged using a 63× water emulsion objective fitted to the Operetta CLS High-Content Analysis System (PerkinElmer, USA) available at the PerkinElmer Center of Excellence facility, IISER-Pune. Analysis of microscopy images In assays where the parasites were allowed to bind to glass coverslips coated with the sulfated and non-sulfated hexuronic acids, the coverslips were imaged using a 100× oil immersion objective and the number of parasites bound to the glass coverslip in each microscopic field was quantified. At least 10 microscopic fields were imaged per coverslip and an average parasite count was determined. This value, along with its standard deviation, was used for statistical analysis and plotting. For the invasion assays, at least 50 microscopic fields were imaged per well of the 96-well plate. In each microscopic field, the number of vacuoles with actively replicating parasites was quantified. For each treatment, data obtained from 4 replicate wells were used to derive the average and standard deviation values, which were further used for statistical analysis and plotting.

## 2.5 References:

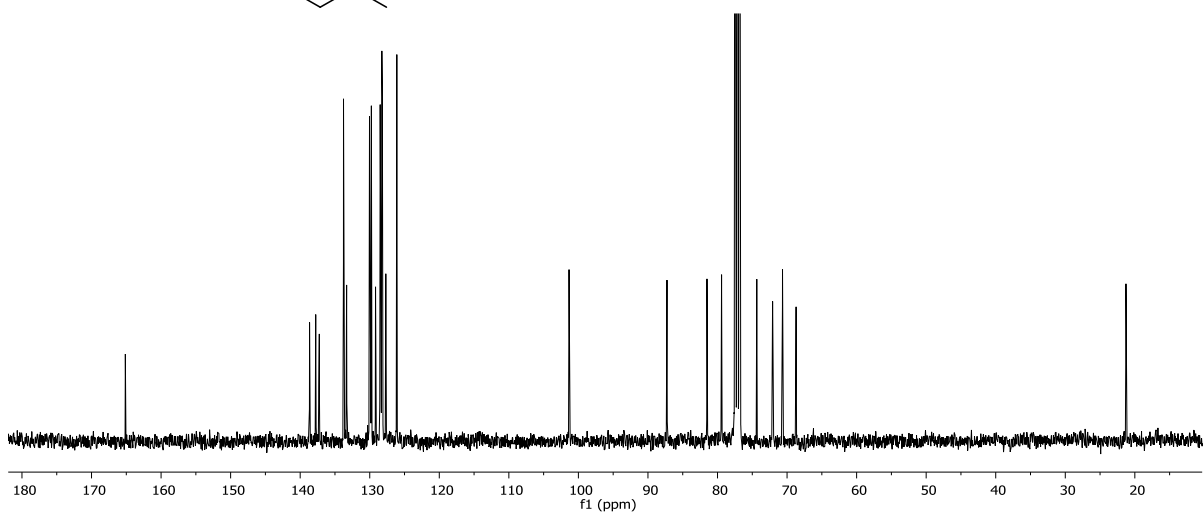
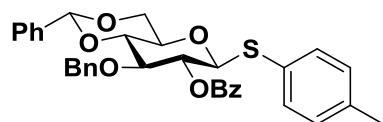
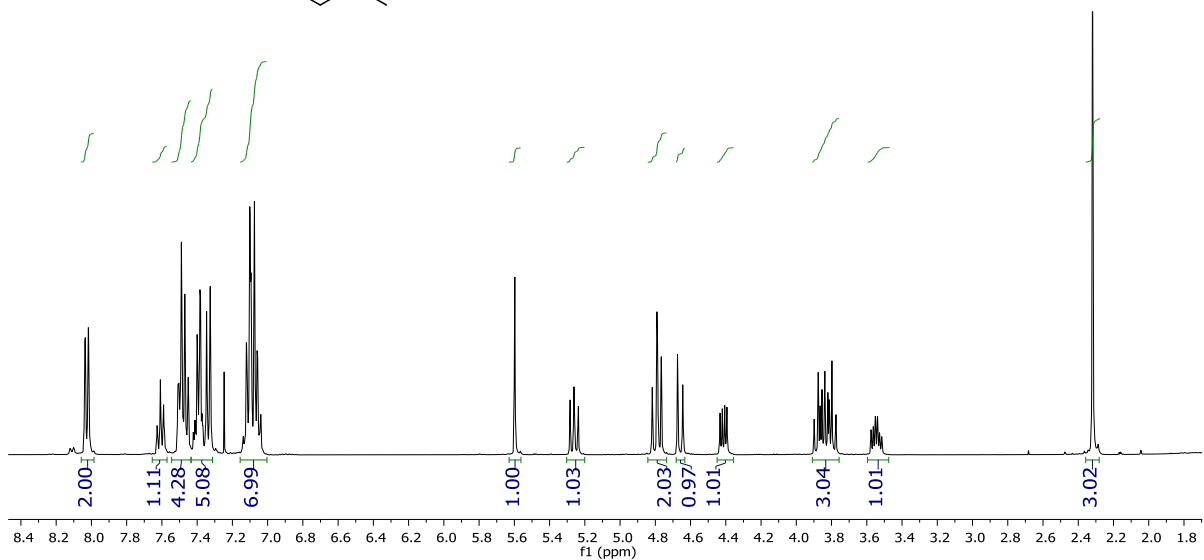
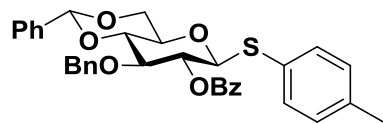
1. *Essentials of Glycobiology.*, ed. A. Varki, R. Cummings, J. Esko, H. Freeze, G. Hart and J. Marth, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2009.
2. S. J. Williams and G. J. Davies, *Trends Biotechnol.*, 2001, **19**, 356-362.
3. N. Jayaraman, *Chem. Soc. Rev.*, 2009, **38**, 3463-3483.
4. A. K. Adak, H. Lin and C. Lin, *Org. Biomol. Chem.*, 2014, **12**, 5563-5573.
5. B. Ernst and J. L. Magnani, *Nat Rev Drug Discov.*, 2009, **8**, 661-677.
6. X. Zeng, C. A. S. Andrade, M. D. L. Oliveira, X. Sun, *Anal Bioanal Chem.*, 2012, **402**, 3161-3176.
7. C. R. Bertozzi and L. L. Kiessling, *Science.*, 2001, **291**, 2357-2364.
8. M. D. Disney and P. H. Seeberger, *Chem. Biol.*, 2004, **11**, 1701-1707.
9. I. Bucior and M. M Burger, *Curr. Opin. Struct. Biol.*, 2004, **14**, 631-637.
10. M. L Huang, C. J Fisher and K. Godula, *Exp. Biol. Med.*, 2016, **0**, 1-12.
11. A. Varki, *Glycobiology.*, 2017, **27**, 3-49
12. A. H. Bartlett and P. W. Park, *Expert Rev Mol Med.*, 2010, **12**, 1-25.
13. S. Sampaolesi, F. Nicotra and L. Russo, *Future Med. Chem.*, 2019, **11**, 43-60.
14. M. Scharenberg, O. Schwardt, S. Rabbani, and B. Ernst, *J. Med. Chem.*, 2012, **55**, 9810-9816.
15. M. Totsika, D. G. Moriel, A. Idris, B. A. Rogers, *Curr. Drug. Targets.*, 2012, **13**, 1386-1399.
16. S. G. Gouin, G. Roos, and J. Bouckaert, *Top. Med. Chem.*, 2014, **12**, 123-168.
17. D. Grunstein, M. Maglinao, R. Kikkeri, M. Collot, K. Barylyuk, B. Lepenies, F. Kamena, R. Zenobi and P. H. Seeberger, *J. Am. Chem. Soc.*, 2011, **133**, 13957-13966.
18. M. Touaibia, E. Krammer, T. C. Shiao, N. Yamakawa, Q. Wang, A. Glinschert, A. Papadopoulos, L. Mousavifar, E. Maes, S. Oscarson, G. Vergoten, M. F. Lensink, R. Roy, and J. Bouckaert, *Molecules.*, 2017, **22**, 1101-1118.
19. A. A. Karelin, N. E. Ustyuzhanina, Y. E. Tsvetkov, N. E. Nifantiev, *Carbohydr. Res.*, 2019, **471**, 39-42.
20. A. Bernardi, J. Jimenez-Barbero, A. Casnati, C. D. Castro, T. Darbre, F. Fieschi, J. Finne, H. Funken, K. Jaeger, M. Lahmann, T. K. Lindhorst, M.

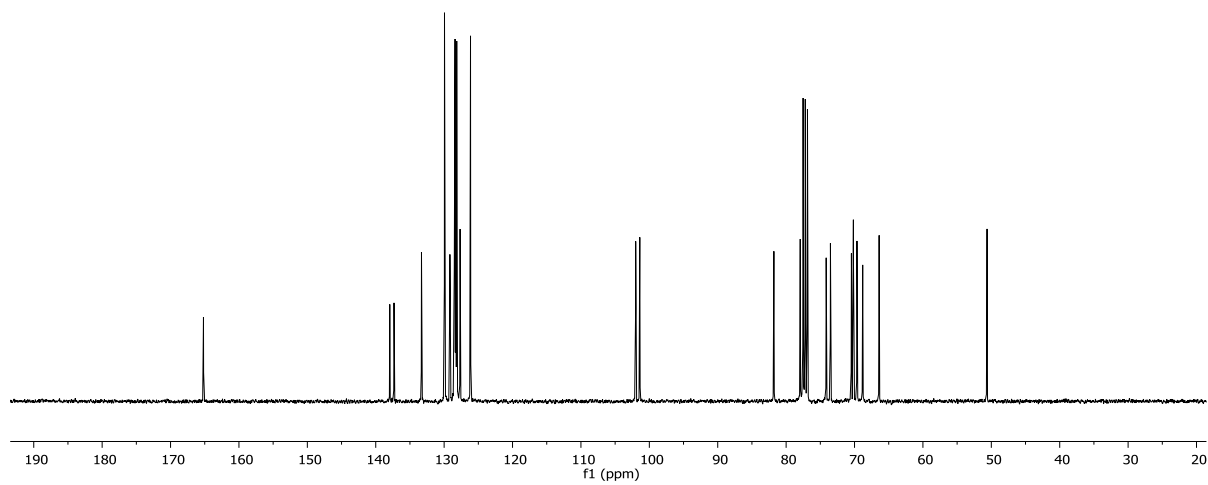
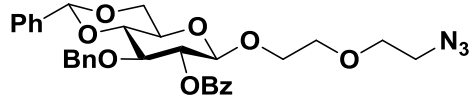
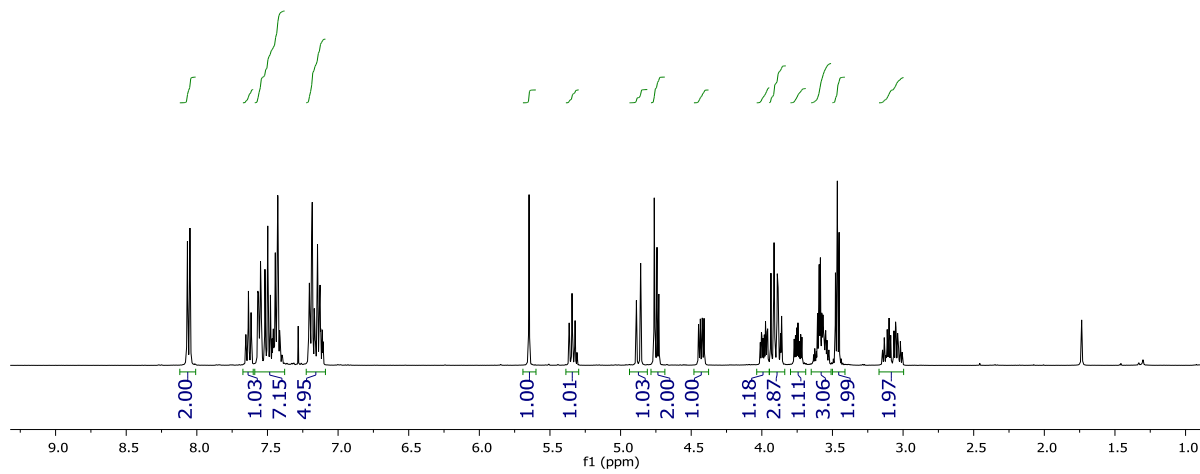
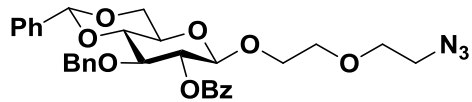


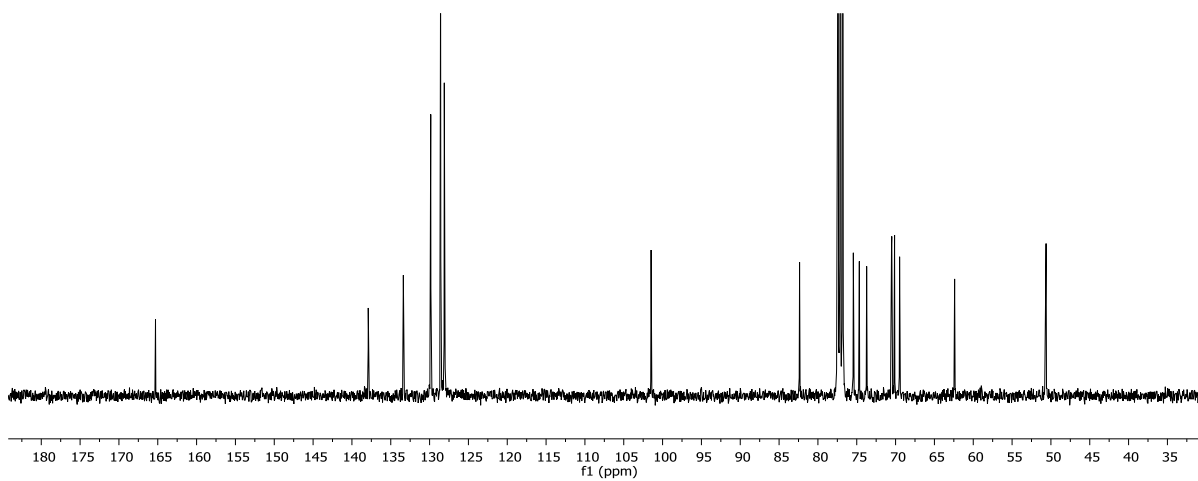
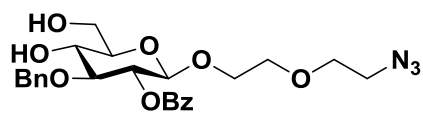
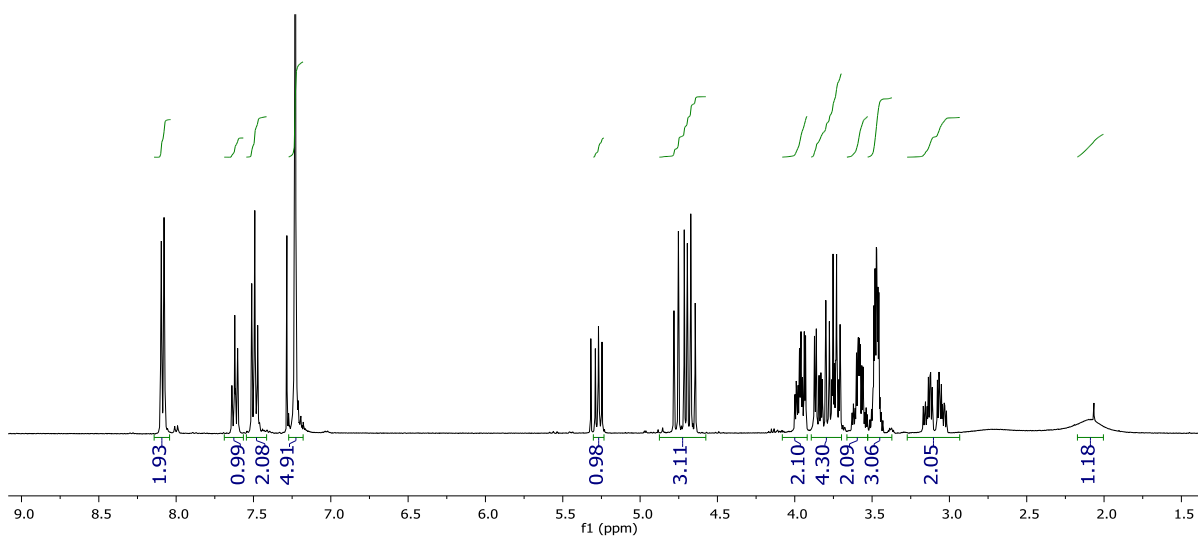
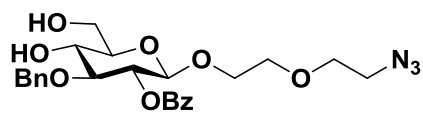
- Marradi, P. Messner, A. Molinaro, P. V. Murphy, C. Nativi, S. Oscarson, S. Penades, F. Peri, R. J. Pieters, O. Renaudet, J. Reymond, B. Richichi, J. Rojo, F. Sansone, C. Schaffer, W. B. Turnbull, T. Velasco-Torrijos, S. Vidal, S. Vincent, T. Wennekes, H. Zuilhofxy and A. Imberty, *Chem. Soc. Rev.*, 2013, **42**, 4709-4727.
21. G. Michaud, R. Visini, M. Bergmann, G. Salerno, R. Bosco, E. Gillon, B. Richichi, C. Nativi, A. Imberty, A. Stocker, T. Darbre and J. Reymond, *Chem. Sci.*, 2016, **7**, 166-182.
22. M. Gade, A. Paul, C. Alex, D. Choudhury, H. V. Thulasiram and R. Kikkeri, *Chem. Commun.*, 2015, **51**, 6346-6349.
23. R. Yadav, P. M. Chaudhary, B. Subramani, S. Toraskar, H. Bavireddi, R. V. Murthy, S. Sangabathuni and R. Kikkeri, *ACS Appl. Mater. Interfaces.*, 2018, **10**, 28322-28330.
24. M. M. Sauer, R. P. Jakob, T. Lubber, F. Canonica, G. Navarra, *J. Am. Chem. Soc.*, 2019, **141**, 936-944.
25. S. Götze, N. Azzouz, Y. Tsai, U. Groß, A. Reinhardt, C. Anish, P. H. Seeberger and D. V. Silva, *Angew. Chem. Int. Ed.*, 2014, **53**, 1-6.
26. A. Bahl, P. H. Davis, M. Behnke, F. Dziarszinski, M. Jagalur, F. Chen, D. Shanmugam, M. W. White, D. Kulp, D. S. Roos, *BMC Genomics.*, 2010, **11**, 603.
27. E. Ortega-Barria and J. C. Boothroyd, *J. Biol. Chem.*, 1999, **274**, 1267-1276.
28. V. B. Carruthers, S. Hakansson, O. K. Giddings and L. D. Sibley, *Infect. Immun.*, 2000, **68**, 4005-4011.
29. E. Kamhi, E. J. Joo, J. S. Dordick and R. J. Linhardt, *Biol. Rev.*, 2013, **88**, 928-943.
30. E. C. Mastrantonio, C. D. Lopes, C. G. Pereira, N. M. Silva, B. B. Fonseca, E. A. V. Ferro, J. R. Mineo and J. D. O. Pena, *Parasitol. Int.*, 2014, **63**, 337-340.
31. A. Jacquet, L. Coulon, J. D. Neve, V. Daminet, M. Haumont, L. Garcia, A. Bollen, M. Jurado, R. Biemans, *Mol. Biochem. Parasitol.*, 2001, **116**, 35-44.
32. N. Azzouz, F. Kamena, P. Laurino, R. Kikkeri, C. Mercier, M. Cesbron-Delauw, J. Dubremetz, L. D. Cola and P. H. Seeberger, *Glycobiology.*, 2013, **23**, 106-120.
33. L. Huang and X. Huang, *Chem. Eur. J.*, 2007, **13**, 529-540
34. T. Polat and C. Wong, *J. Am. Chem. Soc.*, 2007, **129**, 12795-12800.

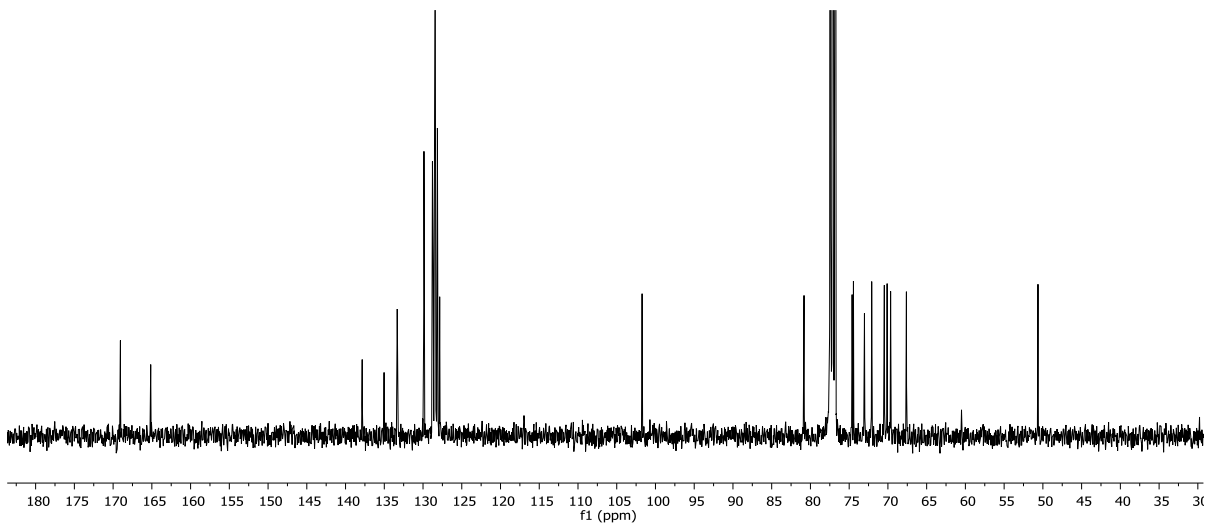
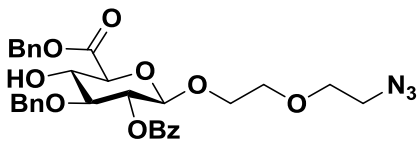
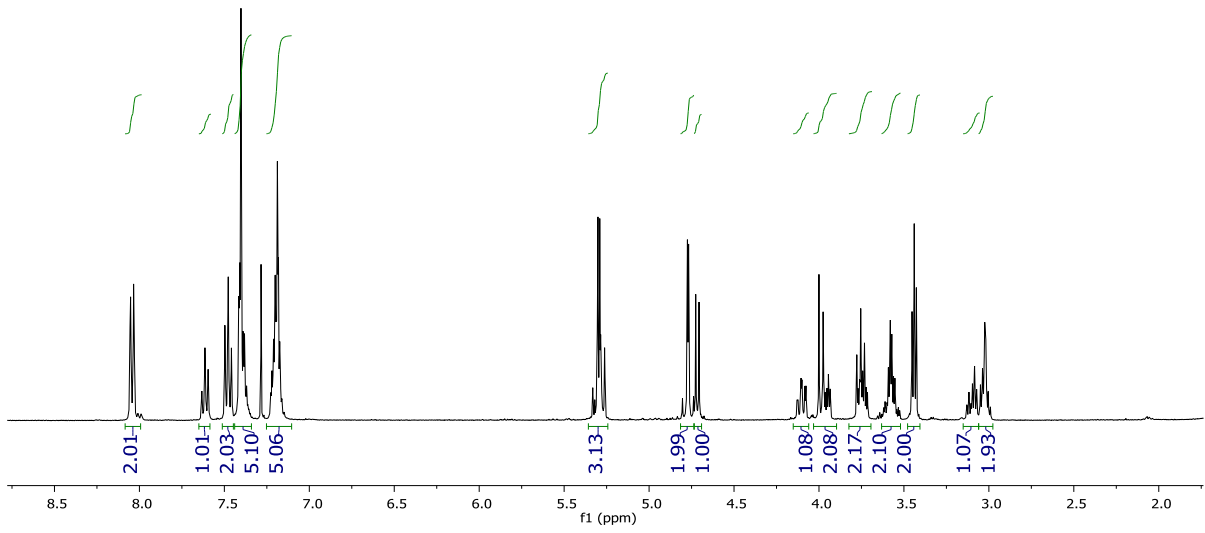
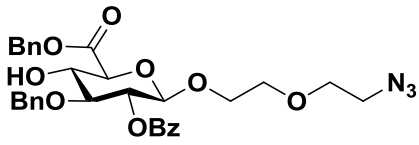
35. T. Li, H. Ye, X. Cao, J. Wang, Y. Liu, L. Zhou, Q. Liu, W. Wang, J. Shen, W. Zhao and P. Wang, *ChemMedChem.*, 2014, **9**, 1071-1080.
36. N. Guedes, P. Czechura, B. Echeverria, A. Ruiz, O. Michelena, *J. Org. Chem.*, 2013, **78**, 6911-6934.
37. M Gade, P. M. Chaudhary, H. V. Thulasiram and R. Kikkeri, *ChemistrySelect.*, 2017, **2**, 8865-8869.
38. M. Gade, P. Khandelwal, S. Sangabathuni, H. Bavireddi, R. V. Murthy, P. Poddar and R. Kikkeri, *Analyst.*, 2016, **141**, 2250-2258.
39. R. S. Loka, F. Yu, E. T. Sletten and H. M. Nguyen, *Chem. Commun.*, 2017, **53**, 9163-9166.
40. K. Xu, D. A. Cantu, Y. F. J. Kim, X. Zheng, P. Hematti, and W. J. Kao, *Acta Biomaterialia.*, 2013, **9**, 8802-8814.
41. S. M. H. Bukhari, S. Khan, M. Rehanullah, and N. M. Ranjha, *Int J Polym Sci.*, 2015, **2015**, 1-15.
42. L. Li, C. Lu, L. Wang, M. Chen, J. White, X. Hao, K. M. McLean, H. Chen, and T. C. Hughes, *ACS Appl. Mater. Interfaces.*, 2018, **10**, 13283-13292.
43. K. Brzezicka, U. Vogel, S. Serna, T. Johannssen, B. Lepenies, and N. Reichard, *ACS Chem. Biol.*, 2016, **11**, 2347-2356.
44. Z. Li, T. Qu, C. Ding, C. Mab, H. Sun, S. Li, X. Liu, *Acta Biomaterialia.*, 2015, **13**, 88-100.
45. D. S. Roos, R. G. K. Donald, N. S. Morrisette and A. L. C. Moulton, *Methods Cell Biol.*, 1994, **45**, 27-63.

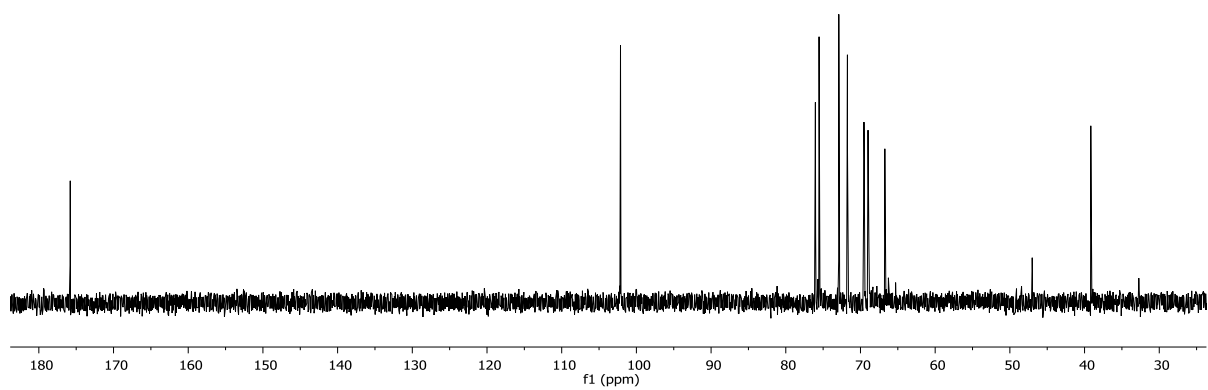
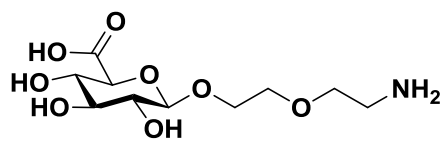
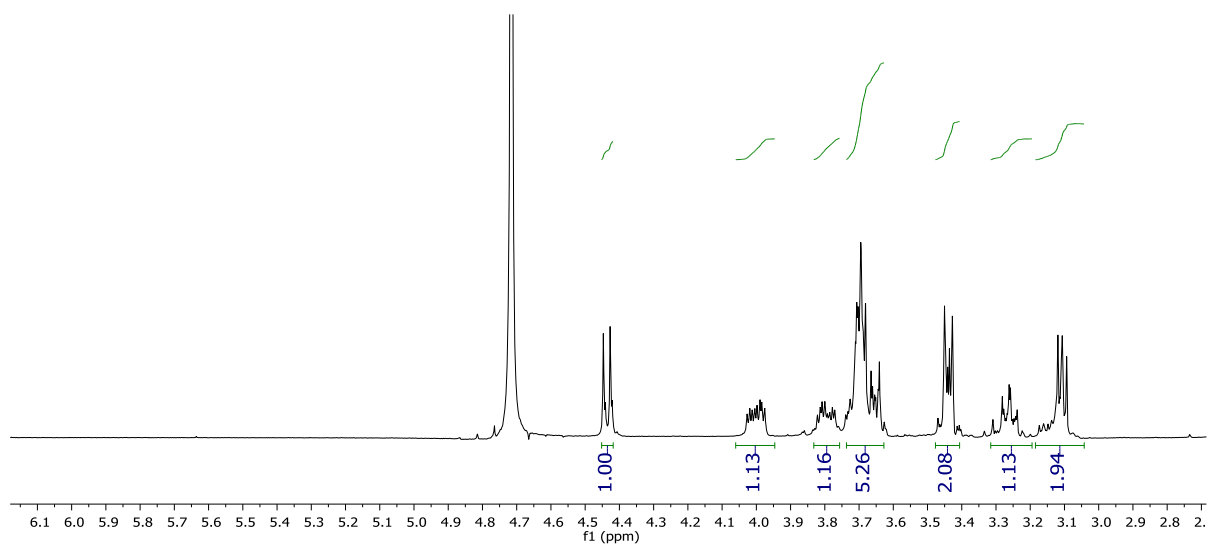
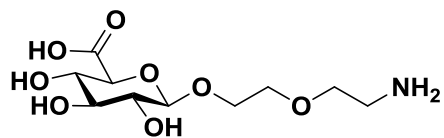
## 2. 6 NMR Spectra:

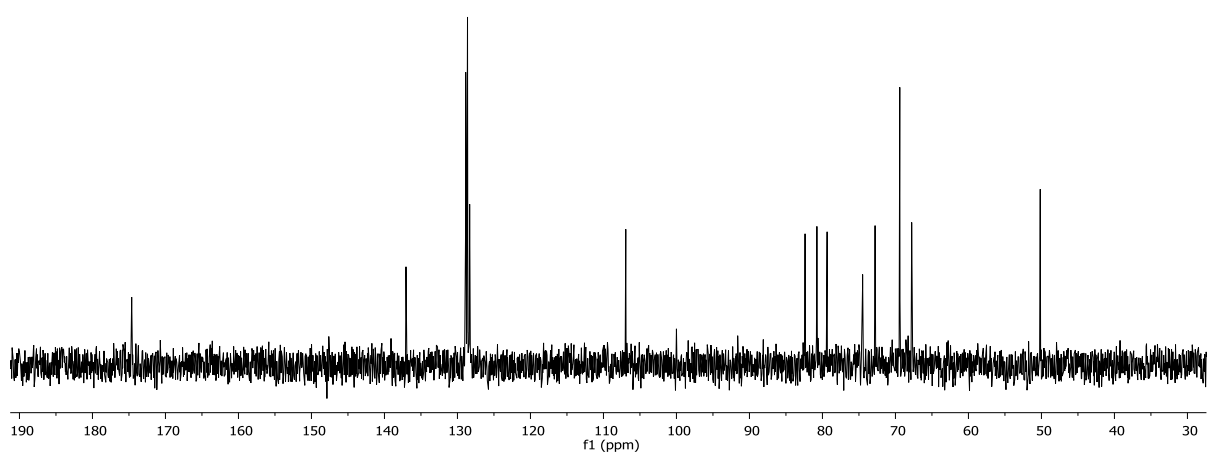
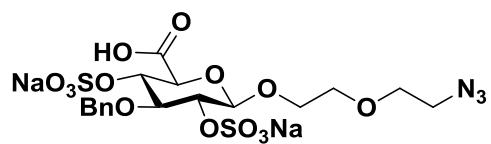
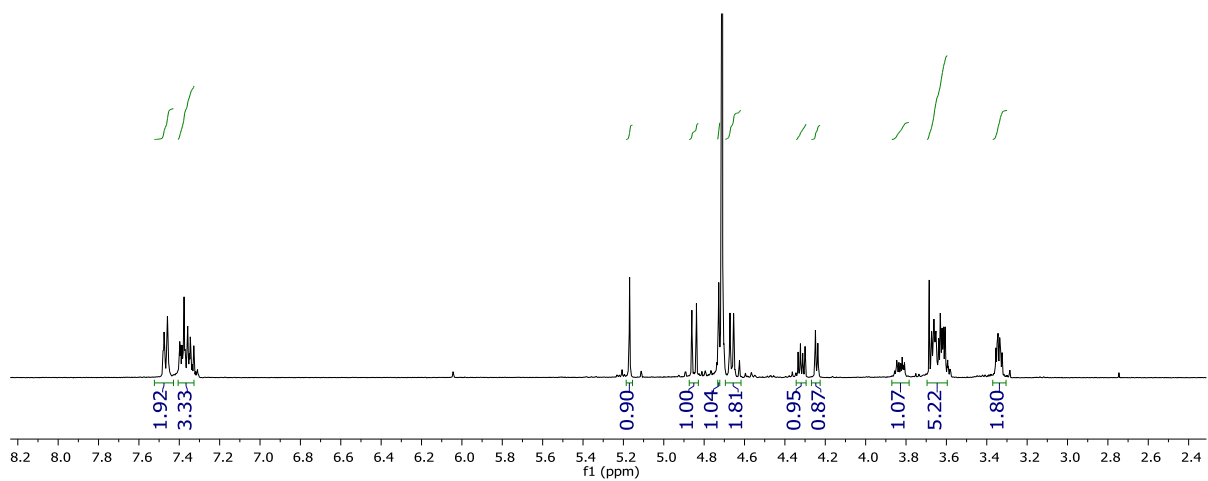
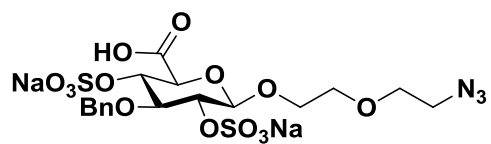




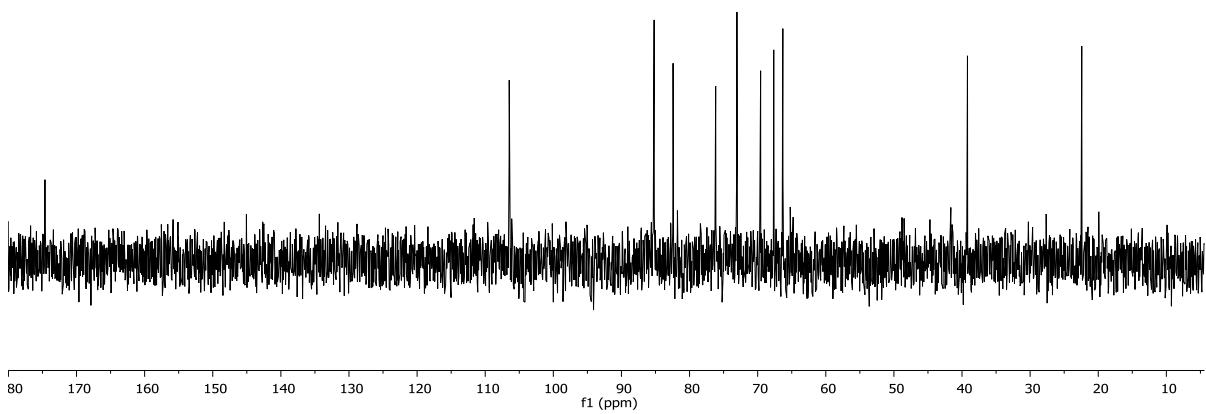
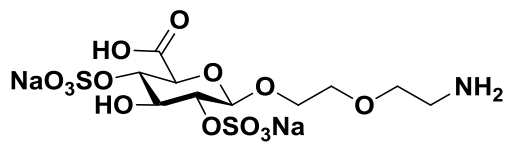
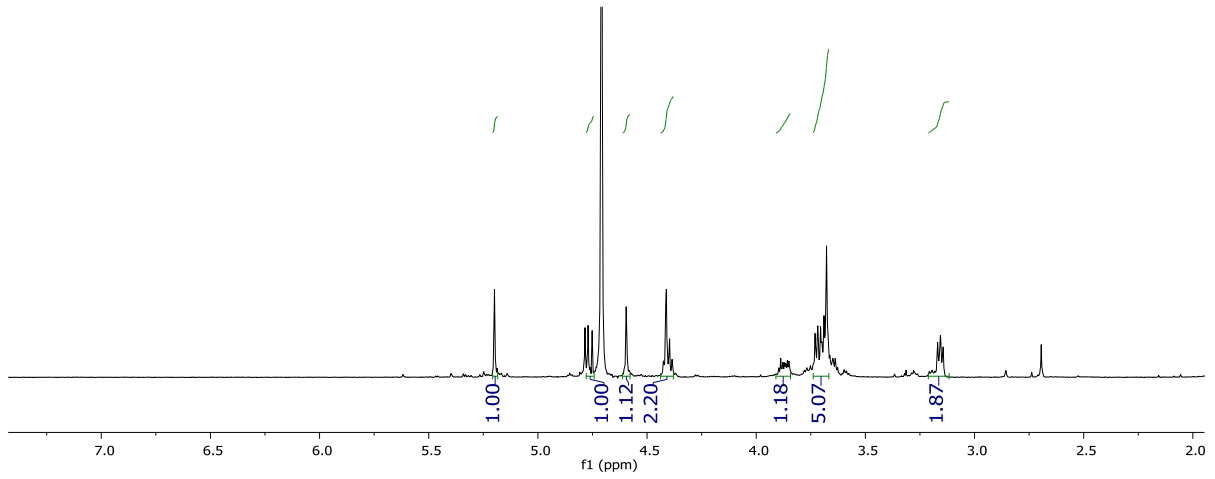
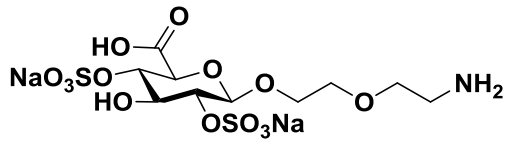


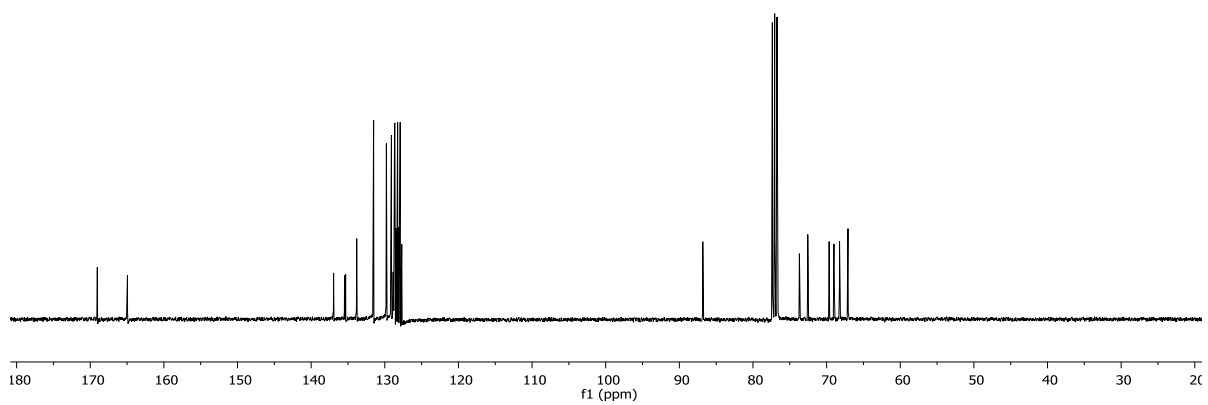
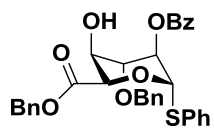
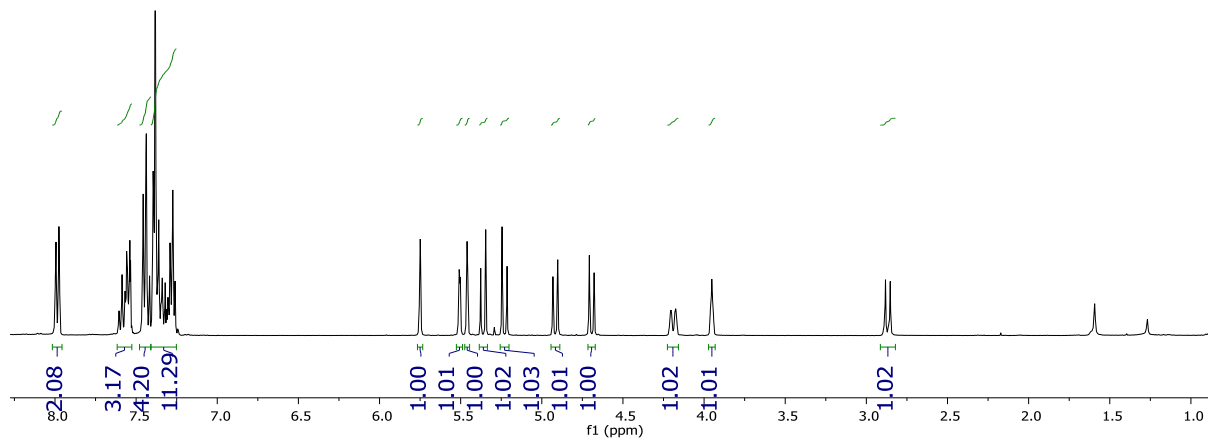
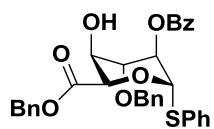


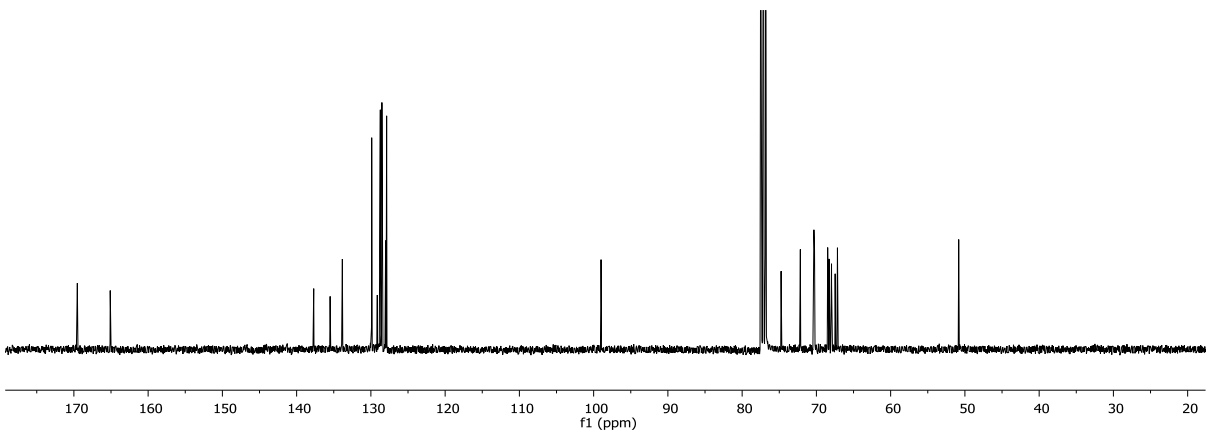
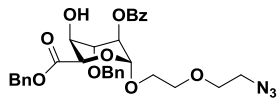
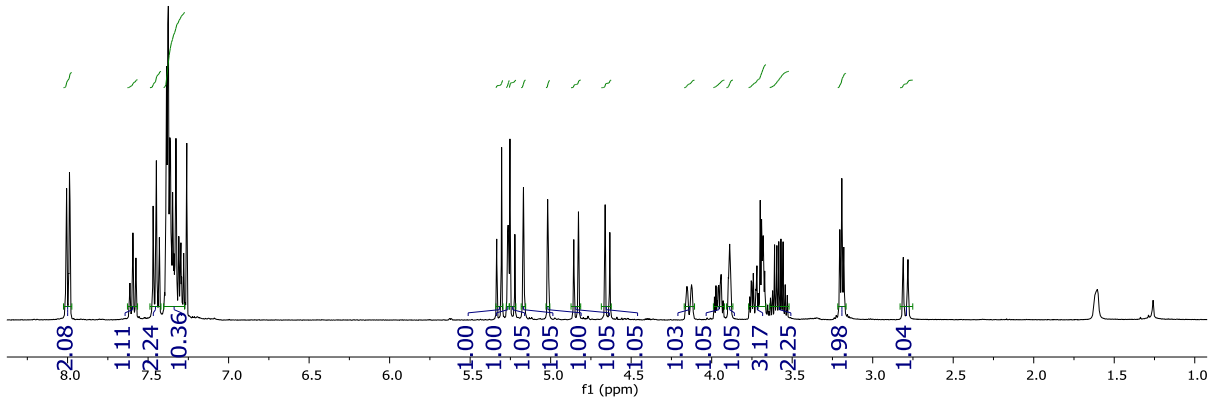
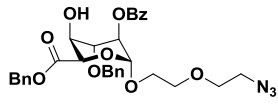


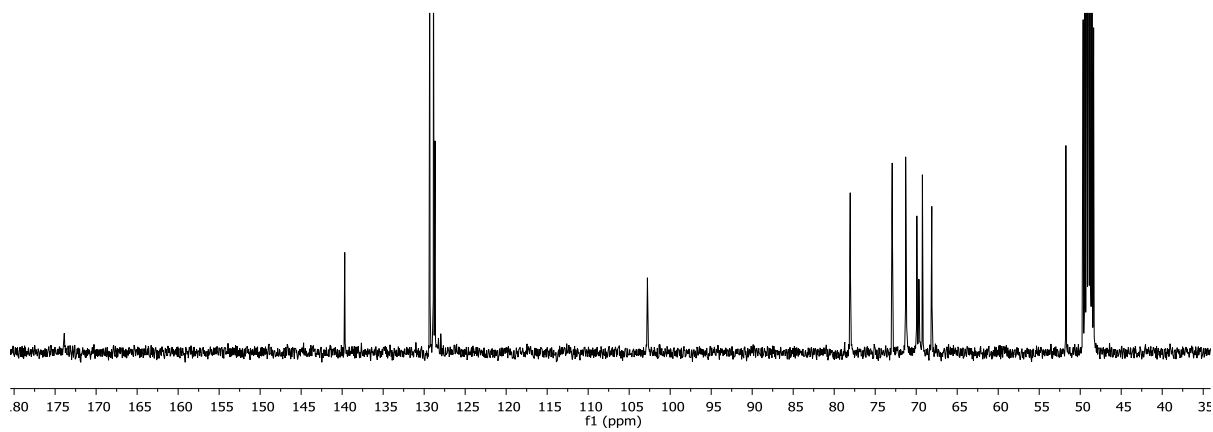
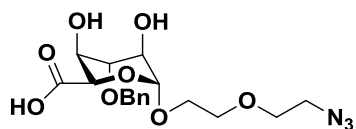
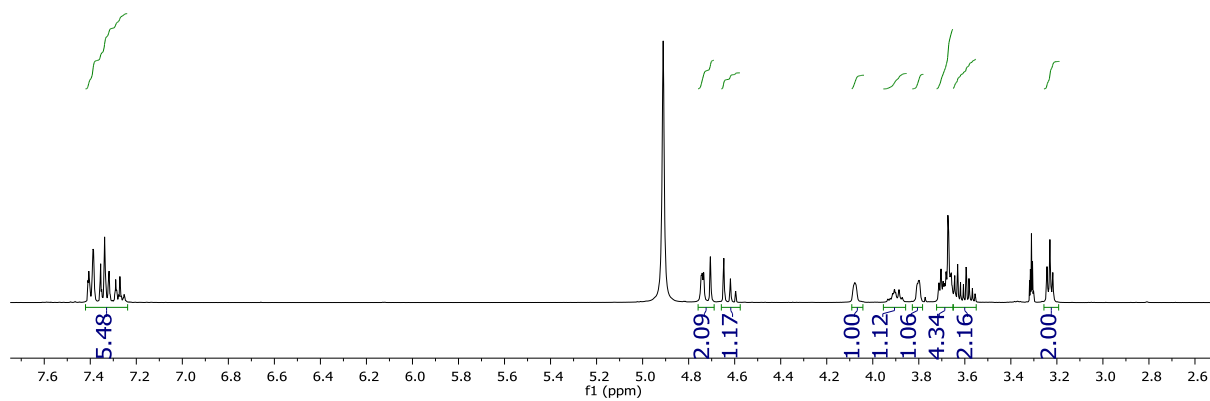
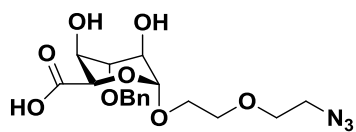


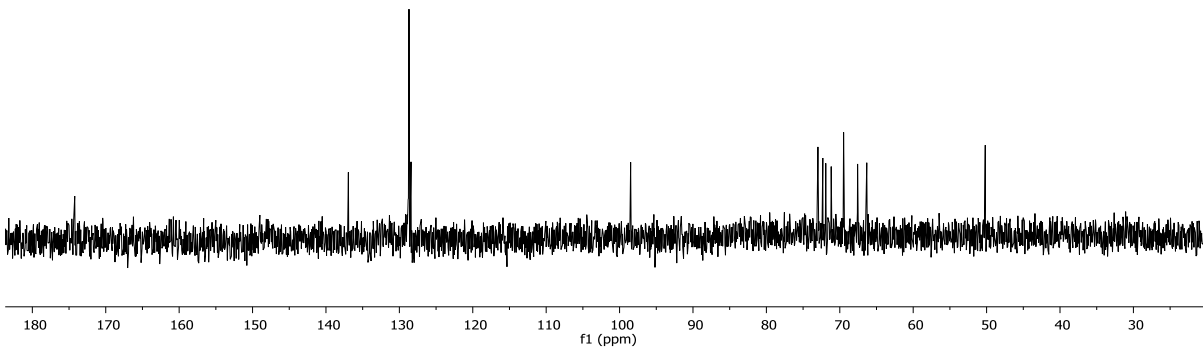
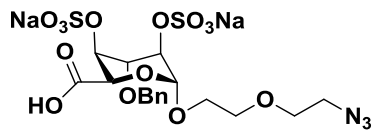
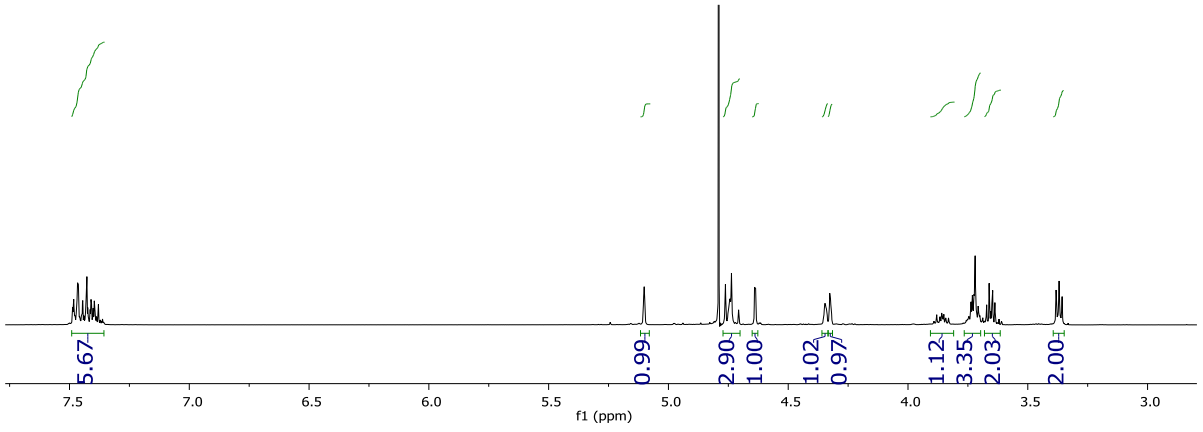
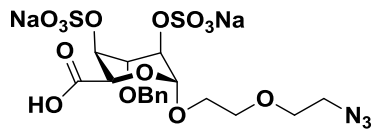


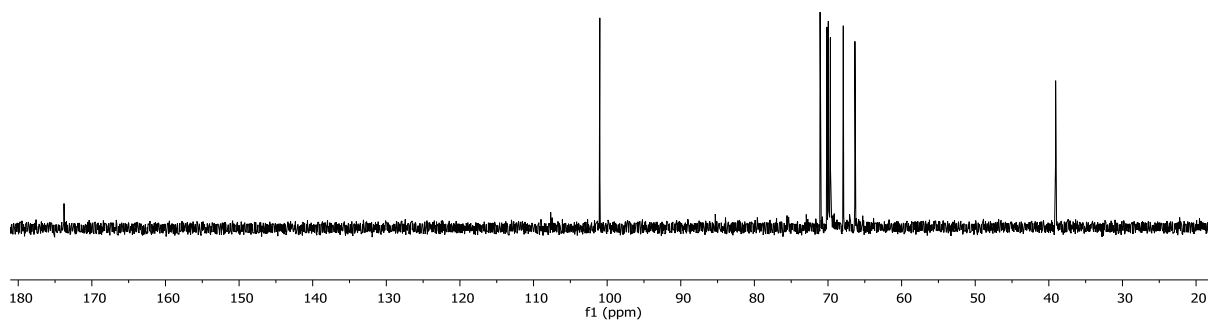
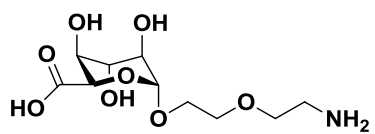
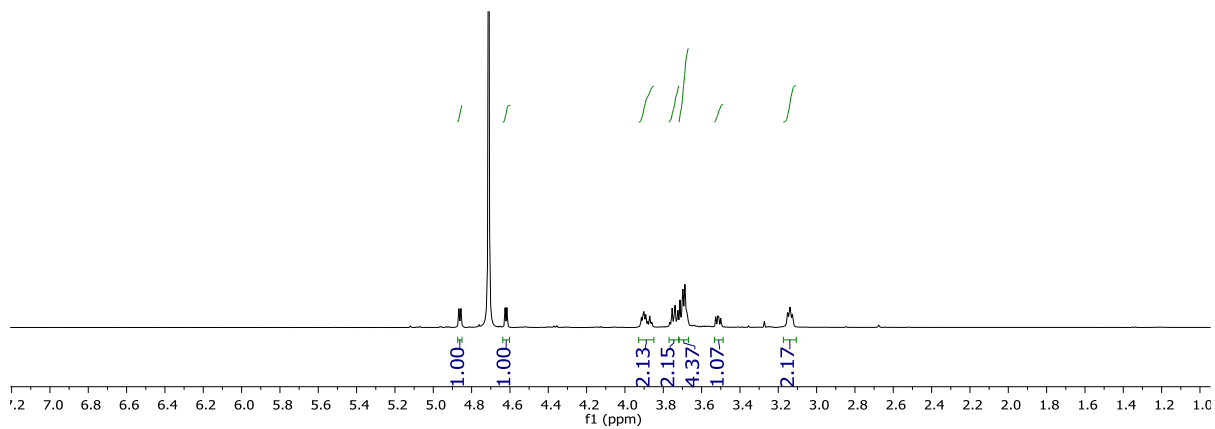
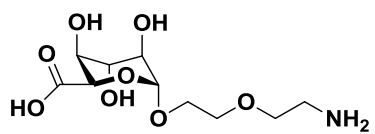


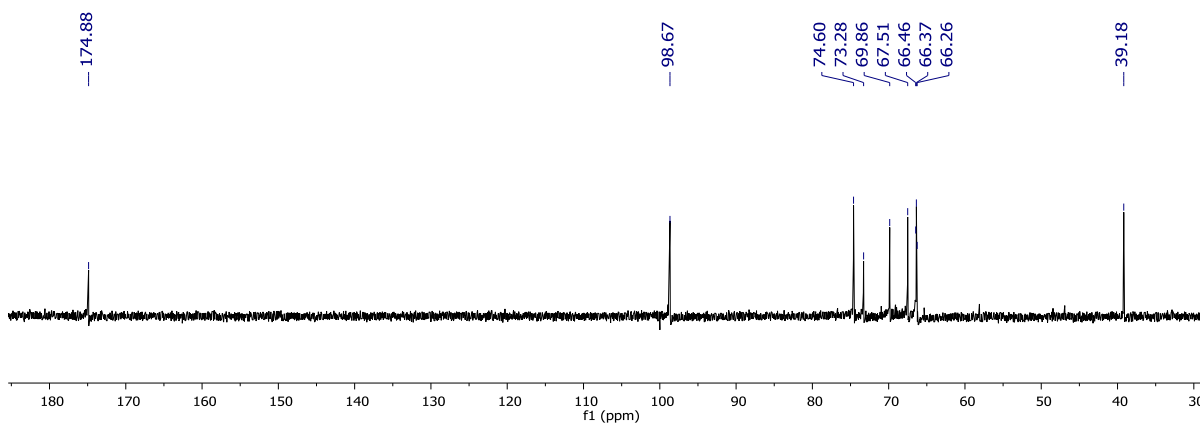
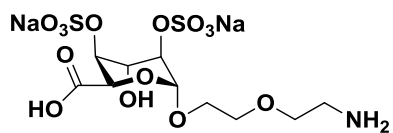
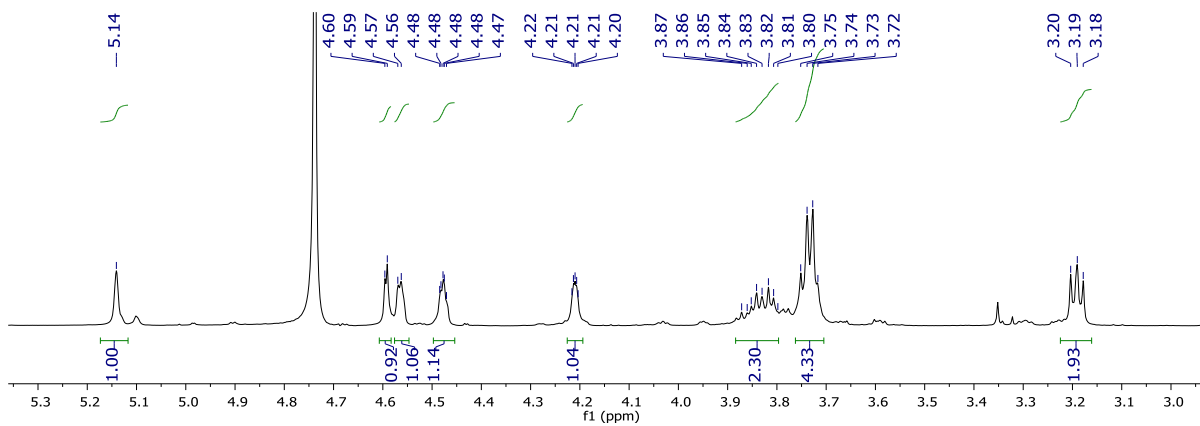
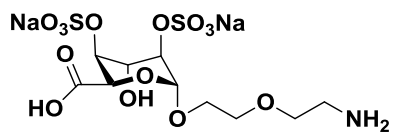












## **Chapter-3**

# **A Cell-Culture Technique to Encode Glyco-Nanoparticles Selectivity**

---



## **Abstract:**

Nanoparticles (NPs) embedded with bioactive ligands such as carbohydrates, peptides, and nucleic acid have emerged as a potential tool to target biological processes. Traditional in vitro assays performed under static conditions may result in non-specific outcomes sometimes, mainly because of the sedimentation and self-assembly nature of NPs. Inverted cell-culture assay allows for flexible and accurate detection of the receptor-mediated uptake and cytotoxicity of NPs. By combining this technique with glyco-gold nanoparticles, cellular internalization and cytotoxicity were investigated. Regioselective glycosylation patterns and shapes of the NPs could tune the receptors' binding affinity, resulting in precise cellular uptake of gold nanoparticles (AuNPs). Two cell lines HepG2 and HeLa were probed with galactosamine-embedded fluorescent AuNPs, revealing significant differences in cytotoxicity and uptake mechanism in upright and inverted in vitro cell-culture assay, high-specificity toward uptake, and allowing for a rapid screening and optimization technique.

### 3.1 Introduction:

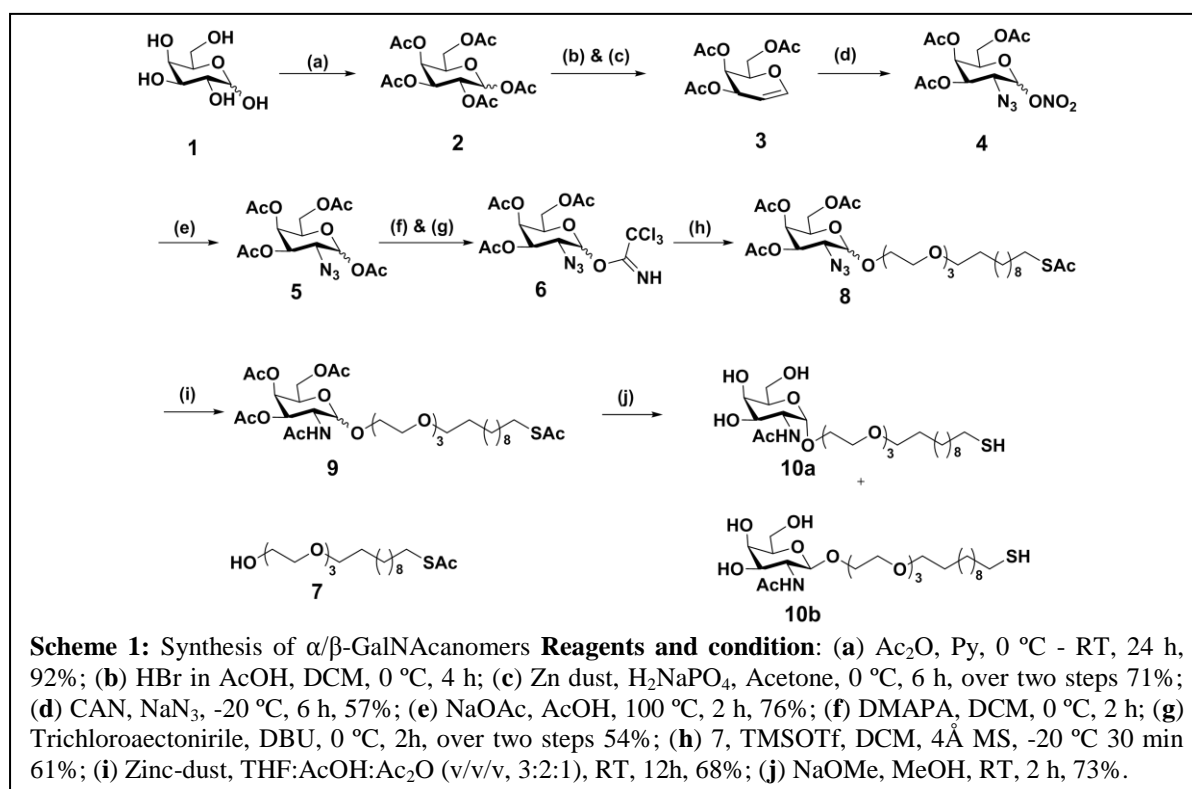
Carbohydrates represent the most abundant biopolymers, and they are ubiquitously found as a part of every mammalian cell surface, the so-called glycocalyx. Recent studies have confirmed that the glycocalyx plays a vital role in a plethora of biological processes, such as cell-cell interaction, immune responses, viral and bacterial infection, and inflammation.<sup>1-5</sup> All these events are regulated by specific carbohydrate-protein interactions (CPIs), in which the inherent structural complexity of carbohydrates synergistically modulate the specificity of CPIs; characteristics that affect CPIs are as follows (a) the nature of the glycosidic linkage, (b) the H-bonding network and (c) the conformation plasticity of sugars.<sup>6-</sup><sup>9</sup> However, the carbohydrate- protein binding affinity is weak (in the mili- to micromolar range), resulting in low specificity.<sup>10</sup> To improve the avidity and specificity of this binding event, carbohydrates are decorated on multivalent probes such as nanoparticles (NPs), polymers, dendrimers, bacteriophages and liposomes.<sup>11-18</sup> These probes offer a multivalent spatial assembly of sugar units to enhance avidity. Among them, glyco-NPs are indispensable tools for developing biosensors, biomarkers, and drug delivery probes.<sup>19,20</sup> They offer optical, magnetic and electrochemical modes of detecting biological events and have become a standard tool in discovering receptor-mediated targeting of cell lines.<sup>21</sup> By applying this tool, various cell surface events have already been successfully monitored over the last several years.

Since many CPIs govern cellular mechanisms and endocytosis pathways, investigating such interactions represents a basic research in the field. However, NPs can undergo self assembly and sediment on the cell surfaces, which increases the concentration of NPs, and results in inaccurate output in cytotoxicity and cellular internalization process.<sup>22,23</sup> To avoid such aggregation, NPs have been modified with inert substrates such as polyethylene glycol (PEG) to increase solubility and mono-dispersity.<sup>24-28</sup> This method provides good qualitative results regarding specificity and selectivity, although PEGylation of NPs results in unfavorable physicochemical properties.<sup>29</sup> Alternatively, in vitro analysis under flow conditions can provide nanotoxicity in more realistic biological conditions.<sup>30,31</sup> However, this approach usually requires a special setup and a trained person to perform this assay. Given the limitations to the approaches mentioned above, a promising strategy is to generate a static cell-culture assay in which cells only react with well-dispersed NPs. This should allow rapid and accurate in vitro data. Here, we present a facile inverted in vitro assay, using glyco-gold NPs to establish specific CPIs. Since our approach does not require any

special setup, it offers a rapid and flexible way to conduct nanotechnology research.

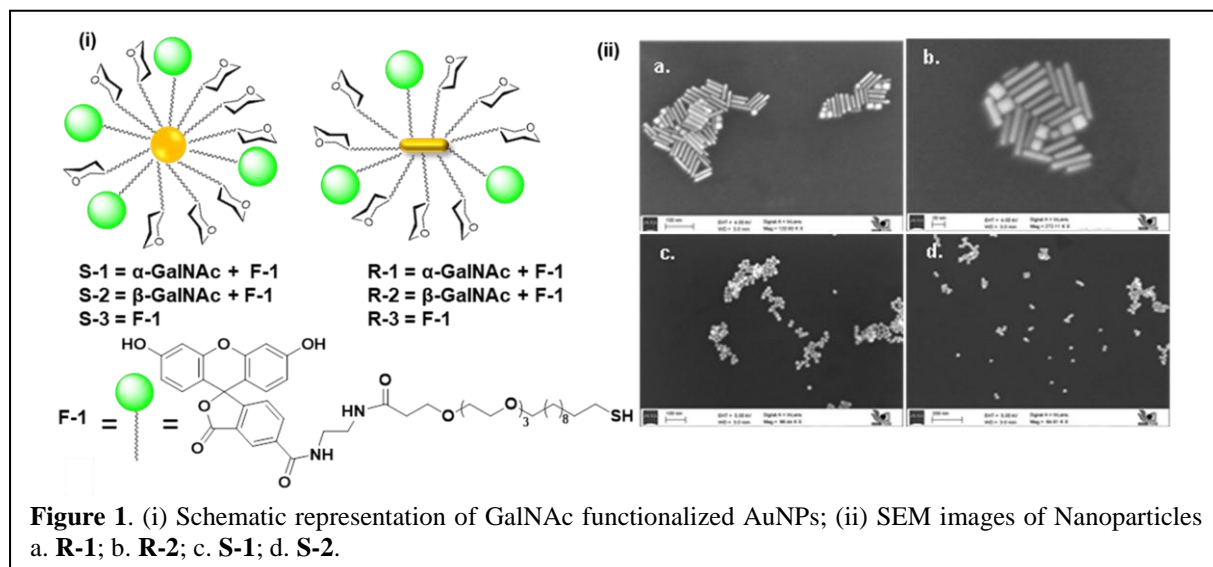
### 3.2 Results and Discussion:

Alpha and beta N-acetyl D-galactosamine-linker (10a and 10b) was synthesized from D-galactose using standard protection and deprotection strategies as reported in the literature with some modifications.<sup>32,33</sup> Briefly, galactose **1** was converted to 2-azidogalactose derivative by peracetylated galactose **2** followed by subsequent anomeric bromination, and zinc-dust based dehydrobromination to obtain galactal derivative **3**. Peracetylated galactal **3** upon treatment with ceric ammonium nitrate (CAN), sodium azide (NaN<sub>3</sub>), and sodium acetate (NaOAc) furnished 2-azido-2-deoxy-D-peracetylated galactose **5**. Compound **5** was converted to trichloroacetimidate donor **6**, which on glycosylation with amphiphilic-PEG linker **7** afforded the corresponding 1:0.6 mixtures of  $\alpha/\beta$ -anomer **8**. The azide group of compound **8** was then reduced by using Zn/AcOH, followed by N-acetylation and O-deacetylation by using sodium methoxide (NaOMe) resulted in the desired final  $\alpha/\beta$ -glycosyl **10a** and **10b** as a mixture of anomers (Scheme 1).



Finally, two anomers were separated by preparative thin layer chromatography and characterized by nuclear magnetic resonance (NMR) spectroscopy. The <sup>1</sup>H NMR spectra of compound **10a** exhibited doublet 4.73 ppm with couple constant of J = 3.7 Hz. Whereas, **10b**

showed doublet 4.36 ppm with  $J = 8.4$  Hz, which are consistent with the vicinal coupling of adjacent proton of anomeric group. Synthesis of Fluorescent linker (F-1) was done as per the reported procedure (Scheme 2).<sup>34</sup>



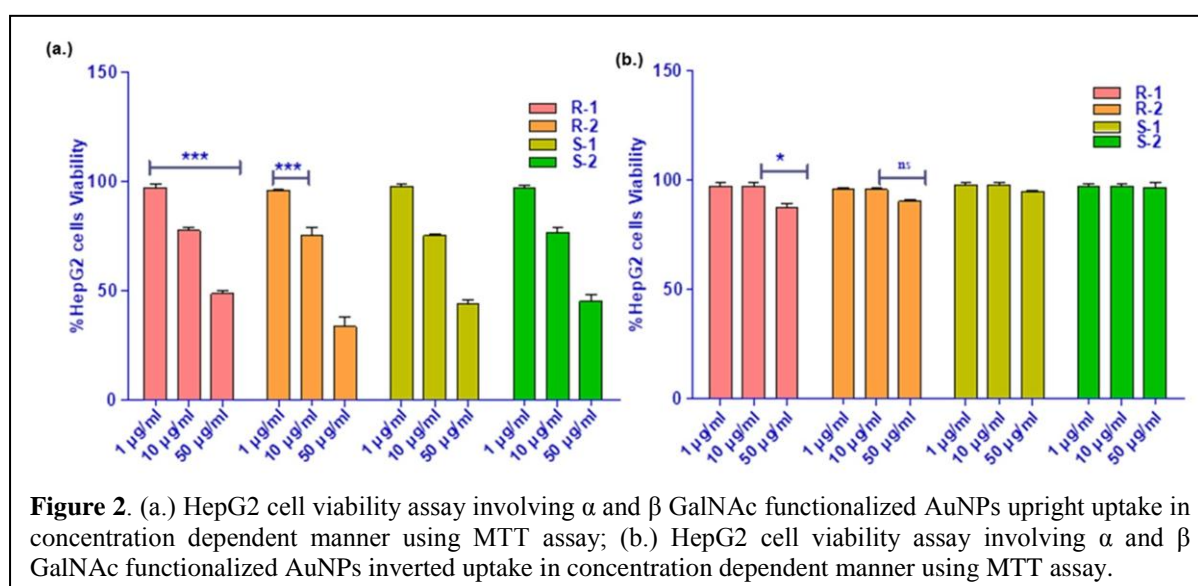
These ligands were functionalized on AuNPs of two different shapes [spherical AuNPs ( $17.78 \pm 1.2$  nm), and rod AuNPs ( $57.54 \pm 3.2 \times 12.05 \pm 1.3$  nm)] using previously published protocol.<sup>35,36</sup> The size and shape of the AuNPs were confirmed by scanning electron microscopy (SEM, Figure 1) and UV-visible spectra.

**Table 1.** Physicochemical properties of  $\alpha$  and  $\beta$  galctosamine functionalized AuNPs of different shapes.

Pariticle	Diameter(s) (nm)	$\lambda_{\max}$ (nm)	$\zeta$ -potential (mV) Water	Concentration (mg/ml)
Rod	$57.54 \pm 3.2 \times 12.05 \pm 1.3$	985	$33.7 \pm 1.5$	---
Sphere	$17.77 \pm 1.2$	521	$-20.2 \pm 0.9$	---
R-1	$58.96 \pm 2.8 \times 12.14 \pm 1.9$	989	$-12.5 \pm 1.2$	$0.22 \pm 0.11$
R-2	$57.87 \pm 3.1 \times 13.91 \pm 1.3$	987	$-14.4 \pm 0.3$	$0.27 \pm 0.12$
S-1	$17.88 \pm 1.5$	522	$-23.1 \pm 1.1$	$0.22 \pm 0.51$
S-2	$18.22 \pm 0.7$	524	$-21.2 \pm 0.8$	$0.20 \pm 0.20$

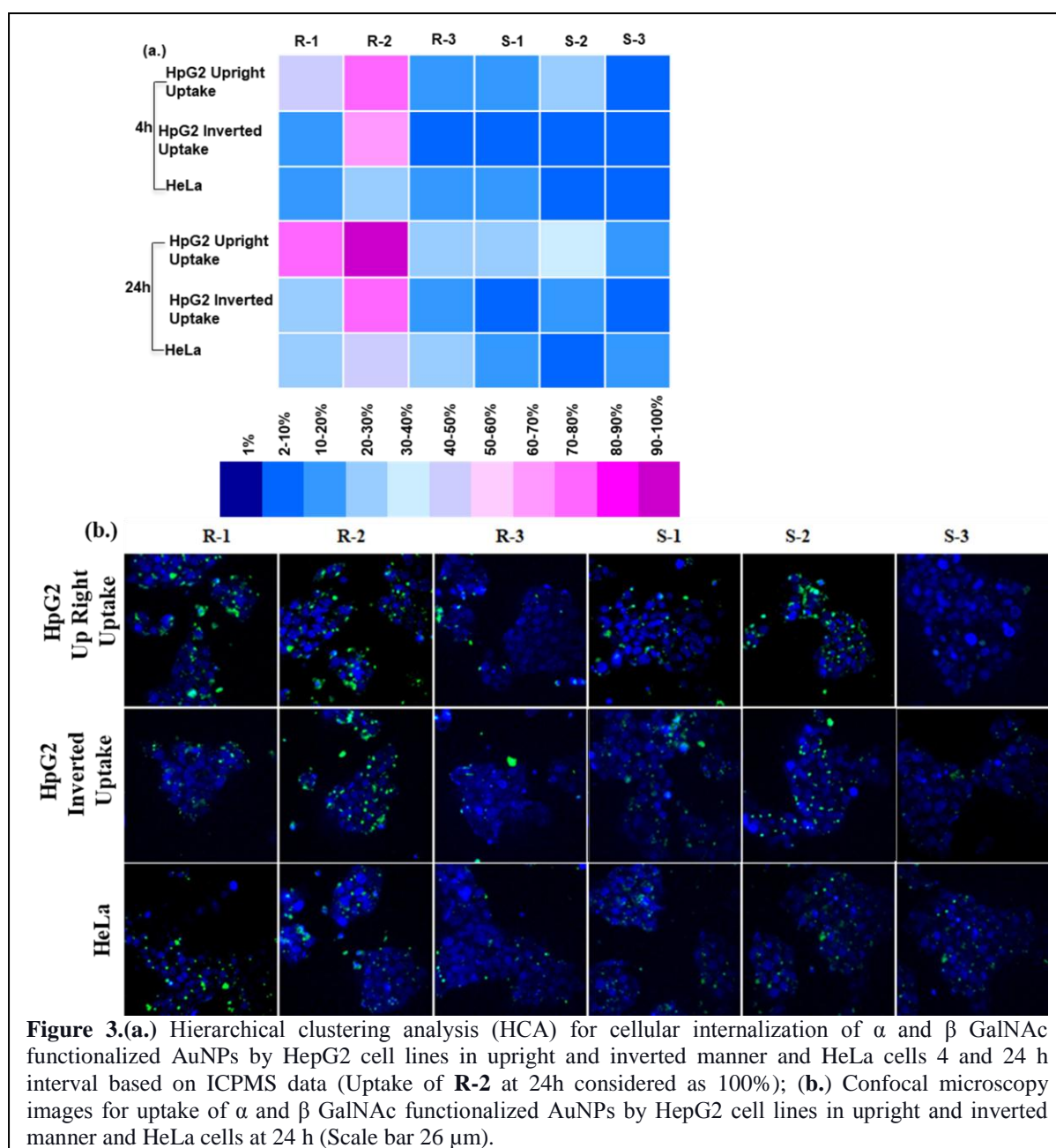
Notably, sphere shaped AuNPs showed a localized surface plasmon resonance (LSPR) peak at 521 nm. The rod shaped AuNPs displayed absorption at near infrared (NIR) region 985 nm (Figure 4). These AuNPs were then treated with F-1 and **10a** or **10b** ligand in ratio 1:5 (Figure 1). The conjugation of fluorescent tag was confirmed by fluorescence spectroscopy (Figure 5) and quantification of N-acetyl galactosamine was done using the phenol-sulfuric acid method, respectively. (Table 1)

Next, we investigated the cytotoxicity of these AuNPs in HepG2 cells using a standard in vitro viability assay (methylthiazolyldiphenyl-tetrazolium bromide or MTT assay).



HepG2 cells are a well-established cell line for galactosamine-based NPs cytotoxicity, uptake, and specificity because they ubiquitously express asialoglycoprotein receptor (ASGP-R) on their cell surfaces.<sup>37-41</sup> We performed a direct comparison of cytotoxicity with static upright and inverted conditions. HepG2 cells were grown on coverslips in a 26-well plate for 24 h. Then, coverslips were placed either in an upright position or inverted setup and cells were exposed to different concentrations of glyco-AuNPs. After 24 h, MTT assay was performed and cell viability quantified. All glyco-AuNPs showed some degree of toxicity in upright conditions with increasing concentrations of glyco-AuNPs (Figure 2a). While in inverted setup, we did not see any toxicity even at a concentration of 50 µg/ml (Figure 2b).

These results revealed that sedimentation of the nanoparticles on the cell surface is one of the likely cause for the discrepancy between upright and inverted cytotoxicity assays.<sup>42</sup>



**Figure 3.**(a.) Hierarchical clustering analysis (HCA) for cellular internalization of  $\alpha$  and  $\beta$  GalNAc functionalized AuNPs by HepG2 cell lines in upright and inverted manner and HeLa cells 4 and 24 h interval based on ICPMS data (Uptake of **R-2** at 24h considered as 100%); (b.) Confocal microscopy images for uptake of  $\alpha$  and  $\beta$  GalNAc functionalized AuNPs by HepG2 cell lines in upright and inverted manner and HeLa cells at 24 h (Scale bar 26  $\mu$ m).

It has generally been shown that toxicity is dependent on NPs amount on the cell surface and exposure time. Under inverted conditions, the NPs exposed to cell are only those that are well dispersed in cell culture media and effectively interact with the cell-surface receptors. To investigate cellular interaction and uptake, we collected the cell pellet after 4 h and 24 h time intervals of exposure to AuNPs (0.1  $\mu$ g/ml) in static upright and inverted conditions. Then, we quantified the amount of AuNPs per cell by using inductively coupled plasma mass spectrometry (ICP-MS). We used HeLa cells as a negative control, as they did not express galactosamine-specific ASGP-R receptors. Based on the ICP-MS data (Table 3), we

developed a hierarchical clustering analysis (HCA) Figure 3a). The HCA showed several cryptic binding pockets. The shape and glycosidic linkage displayed disparities in cellular internalization and specificity in both the upright and inverted condition. Figure 3 confirms that HeLa cells did not show any specific uptake of galactosamine coated AuNPs in either condition, indicating that the AuNPs displayed specific carbohydrate protein interaction. Among two different shapes, rod-shaped AuNPs showed higher uptake than spherical nanostructures after 4 h of exposure to HepG2 cells. This trend continued even after 24 h. Between glycosidic linkages,  $\beta$ -glycoside (**R-2**) displayed strong uptake compared with  $\alpha$ -glycoside (**R-1**). However, in upright condition the difference between **R-1** and **R-2** was ~30% after 4 h of exposure and this difference was reduced to 20–30% after 24 h of exposure. In contrast, **S-1** and **S-2** showed moderate to poor uptake (10–30%) in both upright and inverted condition and this trend continues even after 24 h. One possible explanation for this is that, rod shaped AuNPs may be sedimented on the cell surface with time under upright conditions, possibly resulting in false signals as AuNPs exposure time increases.<sup>43,44</sup>

### 3.3 Conclusions:

In summary, this work reported a systematic investigation of glyco-gold nanoparticles cytotoxicity and cellular uptake under static upright and inverted conditions. Using  $\alpha$  and  $\beta$ -GalNAc appended fluorescent glycoAuNPs of two different shapes, we demonstrated how sedimentation of nanoparticles modulate the receptor-mediated and nanoparticles shape-dependent cellular internalization and toxicity and how simple and readily usable inverted static cell-culture technique is advantageous in such circumstance. Overall, this platform may be useful in various other applications, in which static up-right conditions give false results and sensitive receptor-mediated cellular internalization needs to be established.

### 3.4 General Instructions:

All chemicals were reagent grade and used as supplied except where noted. Analytical thin layer chromatography (TLC) was performed on Merck silica gel 60 F254 plates (0.25 mmol). Compounds were visualized by UV irradiation or dipping the plate in CAM/ninhydrin solution followed by heating. Column chromatography was carried out using force flow of the indicated solvent on Fluka Kieselgel 60 (230–400 mesh).  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra were recorded on Jeol 400 MHz, using residual solvent signals as an internal reference  $\text{CDCl}_3$   $\delta\text{H}$ , 7.23 ppm, and  $\text{CD}_3\text{OD}$   $\delta\text{H}$  3.31 ppm,  $\delta\text{C}$  49.0 ppm). The chemical shifts ( $\delta$ ) are reported in ppm and coupling constants (J) in Hz. UV-visible measurements were performed with



Evolution 300 UV-visible spectrophotometer (Thermo Fisher Scientific, USA). Fluorescence spectra were recorded in FluoroMax-4 spectrofluorimeter (Horiba Scientific, U.S.A.).

### 3.4.1 Synthesis procedure of $\alpha$ and $\beta$ -GalNAc:

**Synthesis of compound 2:** D-Galactose (5 g, 27.77 mmol) was dissolved in pyridine (30 mL), and acetic anhydride (26.25 mL, 277.77 mmol) dropwise. After stirring 24 h at room temperature the mixture was concentrated in high vacuum and co-evaporated with toluene (3x). The residue was extracted with ethyl acetate and washed with HCl (1 M). The organic phase was dried over  $\text{Na}_2\text{SO}_4$ , filtered, and concentrated. The crude residue recrystallized by diethyl ether to afford **2** as a white solid, which was directly used for next reaction without further purification.

**Synthesis of compound 3:** Compound **2** (10 g, 25.64 mmol) was dissolved in DCM (60 mL), and 33% HBr in acetic acid solution (7.41 mL, 128.2 mmol) mixture. After stirring 2 h, the reaction mixture was extracted with DCM and washed with ice-cold water. The organic layer was dried over  $\text{Na}_2\text{SO}_4$ , filtered, and concentrated to afford galactopyranosyl bromide as colorless liquid, which was directly used for next reaction without further purification. The galactopyranosyl bromide was dissolved in acetone (50 mL) and satd.  $\text{NaH}_2\text{PO}_4$  (100 mL), Zinc dust (20.96 g, 320.5 mmol) were added at 0 °C. After stirring for 6 h the zinc dust was filtered off on celite bed and resulting filtrate diluted with EtOAc and washed with 10%  $\text{NaHCO}_3$ , brine and dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated. The crude residue was purified by column chromatography (EtOAc/Hexane, 1/2) to afford **3** (5 g, 71%) as a colorless syrup.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  6.47 (dd,  $J = 6.3, 1.8$  Hz, 1H), 5.57 – 5.55 (m, 1H), 5.44 – 5.42 (m, 1H), 4.73 (ddd,  $J = 6.3, 2.7, 1.5$  Hz, 1H), 4.35 – 4.31 (m, 1H), 4.30 – 4.20 (m, 2H), 2.13 (s, 3H), 2.09 (s, 3H), 2.03 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  170.55, 170.27, 170.13, 145.40, 98.84, 72.79, 63.89, 63.75, 61.92, 20.80, 20.75, 20.65. HRMS (ESI)  $m/z$ : calc'd for  $\text{C}_{12}\text{H}_{16}\text{O}_7$ : 272.0896; Found: 272.0872.

**Synthesis of compound 4:** Compound **3** (5 g, 18.38 mmol) was dissolved in acetonitrile (35 mL) and stirred at -20 °C with 4Å molecular sieves (1.0 g) for 30 min. Later  $\text{NaN}_3$  (1.79 g, 27.57 mmol) and ceric ammonium nitrate (20.15 g, 36.76 mmol) were added. The reaction mixture was stirred at -20 °C for 6 h. The mixture was filtered on celite, diluted with ethyl acetate, and washed with water and brine. The organic layer was dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated. The residue was purified by column chromatography (EtOAc/hexane, 1:1) to afford **4** (3.95 g, 57 %) as a mixture of  $\alpha$ ,  $\beta$  and talo-stereoisomers in the ratio of 1:0.7:0.2



(as determined by  $^1\text{H}$  NMR spectroscopy).  $\alpha$ -isomer  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  6.36 (d,  $J = 4.1$  Hz, 1H), 5.52 (dd,  $J = 3.3, 1.3$  Hz, 1H), 5.27 (dd,  $J = 11.4, 3.2$  Hz, 1H), 4.40 – 4.36 (m, 1H), 4.16 – 4.13 (m, 3H), 2.19 (s, 3H), 2.09 (s, 3H), 2.05 (d,  $J = 0.9$  Hz, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  170.29 - 169.18 (C=O, acyl), 96.87, 69.49, 68.58, 66.62, 60.91, 55.93, 20.58 – 20.36 (3x CH<sub>3</sub>).  $\beta$ -isomer  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  5.60 (d,  $J = 8.8$  Hz, 1H), 5.41 (dd,  $J = 3.3, 1.1$  Hz, 1H), 4.98 (dd,  $J = 10.6, 3.3$  Hz, 1H), 4.13 – 4.09 (m, 2H), 4.10 – 4.05 (m, 1H), 3.84 (dd,  $J = 10.6, 8.8$  Hz, 1H), 2.19 (s, 3H), 2.10 (s, 3H), 2.06 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  170.29 – 169.18 (C=O, acyl), 98.08, 71.87, 71.71, 65.82, 60.83, 57.49, 20.58 – 20.36 (3x CH<sub>3</sub>). Talo-isomer  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  6.23 (d,  $J = 1.5$  Hz, 1H), 5.47– 5.45 (m, 1H), 5.33 – 5.30 (m, 1H), 4.23 – 4.20 (m, 3H), 4.01 – 3.98 (m, 1H), 2.23 (s, 3H), 2.12 (s, 3H), 2.06 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  170.29 – 169.18 (C=O, acyl).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  97.74, 69.51, 67.20, 64.81, 61.14, 55.27, 20.58 – 20.36 (3x CH<sub>3</sub>). HRMS (ESI)  $m/z$ : calc'd for  $\text{C}_{12}\text{H}_{16}\text{N}_4\text{O}_{10}$ : 376.0866; Found: 376.0839.

**Synthesis of compound 5:** A mixture of compound **4** (17.77 g, 47.6 mmol) and NaOAc (7.53 g, 94.52 mmol) in AcOH (130 mL) was refluxed at 100 °C. After 1 h the mixture was allowed to cool to room temperature, which was diluted with DCM, washed with water, sat.  $\text{NaHCO}_3$  (50 mL), and brine respectively. The combined organic layer was dried over  $\text{Na}_2\text{SO}_4$ , filtered, and concentrated. The crude product was purified by column chromatography (EtOAc-hexane, 1:2) to afford **5** (13.6 g, 76%) as a white solid as a mixture of  $\alpha$  and  $\beta$  anomers ( $\alpha/\beta$  1:0.5 as determined by  $^1\text{H}$  NMR spectroscopy). The individual isomer was not isolated from crude but NMR data for each is reported separately.  $\alpha$ -isomer  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  6.34 (d,  $J = 3.6$  Hz, 1H), 5.49 (dd,  $J = 3.2, 1.2$  Hz, 1H), 5.33 (dd,  $J = 11.2, 3.2$  Hz, 1H), 4.33 – 4.28 (m, 1H), 4.12 – 4.06 (m, 2H), 3.95 (dd,  $J = 11.0, 3.6$  Hz, 1H), 2.19 (s, 3H), 2.18 (s, 3H), 2.09 (s, 3H), 2.05 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  170.32 – 168.52 (C=O, acyl), 90.41, 68.74, 68.66, 66.85, 61.08, 56.83, 20.91 – 20.55 (3x CH<sub>3</sub>).  $\beta$ -isomer  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  5.56 (d,  $J = 8.6$  Hz, 1H), 5.39 (dd,  $J = 3.2, 0.8$  Hz, 1H), 4.91 (dd,  $J = 10.8, 3.4$  Hz, 1H), 4.16 – 4.12 (m, 2H), 4.04 – 4.00 (m, 1H), 3.86 (dd,  $J = 10.8, 8.4$  Hz, 1H), 2.22 (s, 3H), 2.18 (s, 3H), 2.08 (s, 3H), 2.05 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  170.32 – 168.52 (C=O, acyl), 92.86, 71.72, 71.29, 66.16, 60.93, 59.65, 20.91 – 20.55 (3x CH<sub>3</sub>). HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{Na}]^+$   $\text{C}_{14}\text{H}_{19}\text{N}_3\text{O}_9\text{Na}$ : 396.1019; Found: 396.1010.

**Synthesis of compound 6:** The mixture of compound **5** (1.0 g, 2.68 mmol) and DMAPA (5 equiv, 1.68 mL, 13.40 mmol) were dissolved in THF (15 mL) and stirred at room temperature

till complete disappearance of starting material on TLC. The reaction mixture was diluted in DCM and washed with 1M hydrochloric acid and brine respectively. The organic layer was dried over  $\text{Na}_2\text{SO}_4$ , filtered, and concentrated to afford hemiacetal. The hemiacetal was directly used for next reaction without further purification. Hemiacetal (0.88 g, 2.65 mmol) was dissolved in dry DCM (15 mL) and were added trichloroacetonitrile (2.66 mL, 26.58 mmol) and DBU (0.16 mL, 1.06 mmol). After stirring 2 h, the reaction mixture was concentrated and purified by column chromatography (EtOAc-hexane, 0.4/1) to afford **6** (0.65 g, 54%) with trace amount of  $\beta$  isomer ( $\alpha/\beta$  1:0.2 as determined by  $^1\text{H}$  NMR spectroscopy) as a white foam.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.81 (s, 1H), 6.50 (d,  $J = 3.5$  Hz, 1H), 5.53 (dd,  $J = 3.4, 1.5$  Hz, 1H), 5.36 (dd,  $J = 11.1, 3.0$  Hz, 1H), 4.41 (t,  $J = 6.6$  Hz, 1H), 4.14 (m, 1H), 4.04 (m, 2H), 2.16 (s, 3H), 2.06 (s, 3H), 1.99 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  170.39, 170.03, 169.80, 160.78, 94.55, 69.18, 68.77, 66.99, 61.28, 57.10, 20.71. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{Na}]^+$   $\text{C}_{14}\text{H}_{17}\text{Cl}_3\text{N}_4\text{O}_8\text{Na}$ : 497.0010, Found: 497.0015.

**Synthesis of compound 7:** Compound **7** was synthesized by following synthetic reported protocol.<sup>45</sup>

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  3.77 – 3.73 (m, 2H), 3.72 – 3.66 (m, 6H), 3.66 – 3.59 (m, 4H), 3.47 (t,  $J = 6.8$  Hz, 2H), 2.88 (t,  $J = 7.3$  Hz, 2H), 2.34 (s, 3H), 1.63 – 1.53 (m, 4H), 1.33 – 1.28 (s, 14H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  196.10, 72.52, 71.59, 70.62, 70.37, 70.02, 61.78, 30.65, 29.58, 29.55, 29.50, 29.45, 29.16, 29.11, 28.81, 26.06. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{H}]^+$   $\text{C}_{19}\text{H}_{39}\text{O}_5\text{S}$ : 379.2518, Found: 379.2526.

**Synthesis of compound 8:** The compound **6** (0.5 g, 1.05 mmol) and **8** (0.47 g, 1.26 mmol) were kept in under vacuum overnight and dissolved in dry DCM and added 4Å MS. The mixture was stirred at room temperature for 1 h and cooled at  $-20$  °C then added TMSOTf (0.038 mL, 0.21 mmol). After 1 h the reaction mixture was quenched with triethylamine and concentrated. The crude was purified by column chromatography (EtOAc/Hexane, 0.5/1) to afford **8** (0.45 g, 61%) with mixture of  $\alpha$  and  $\beta$  anomers ( $\alpha/\beta$  1:0.6 as determined by  $^1\text{H}$  NMR spectroscopy) as a white foam. The individual isomer was not isolated from the crude.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  5.49 – 5.33 (m, 2H), 4.78 (dd,  $J = 10.9, 3.2$  Hz, 1H), 4.50 (d,  $J = 8.0$  Hz, 1H), 4.17 – 4.10 (m, 2H), 4.10 – 4.02 (m, 1H), 3.88 – 3.82 (m, 2H), 3.78 – 3.73 (m, 2H), 3.71 – 3.64 (m, 6H), 3.64 – 3.57 (m, 2H), 3.46 (t,  $J = 6.8$  Hz, 2H), 2.87 (t,  $J = 7.3$  Hz, 2H), 2.33 (s, 3H), 2.16 (s, 3H), 2.06 (s, 6H), 1.61 – 1.53 (m, 4H), 1.43 – 1.27 (m, 14H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  196.08, 170.37, 170.09, 170.07, 169.82, 102.52, 98.25, 71.55,

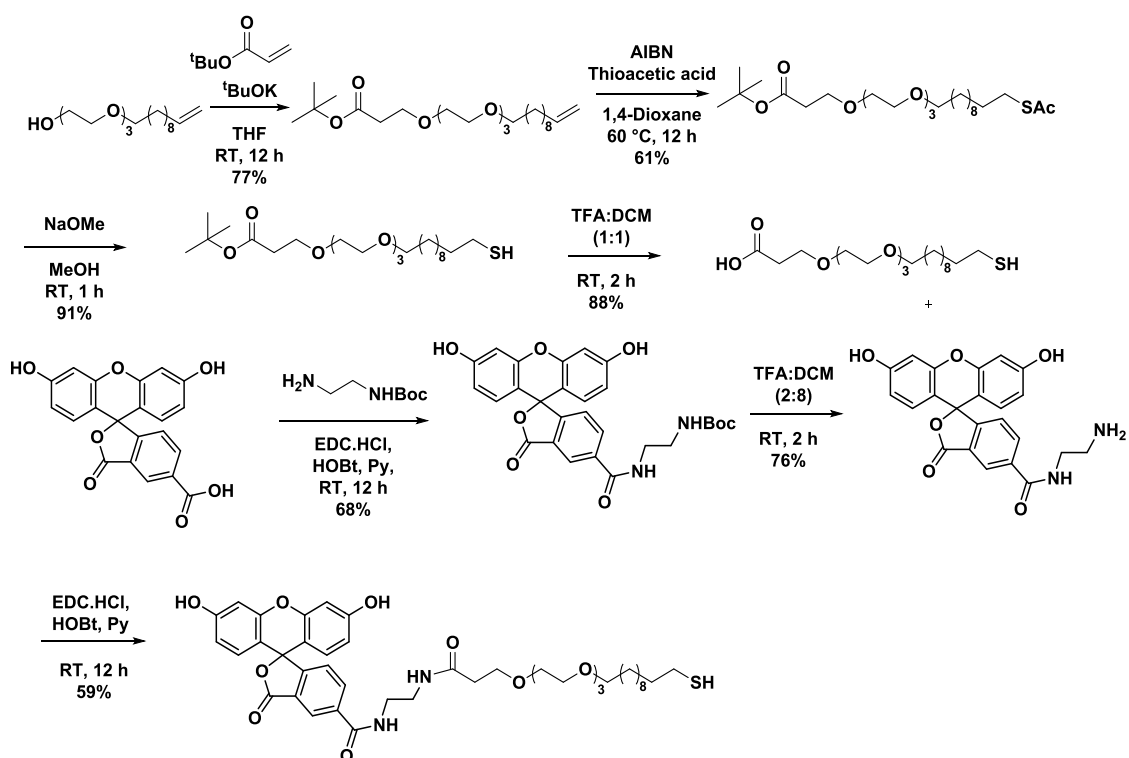
70.99, 70.72, 70.64, 70.61, 70.40, 70.08, 70.04, 69.41, 68.17, 67.74, 67.64, 66.57, 66.38, 61.59, 61.28, 60.72, 57.41, 30.64, 29.62, 29.55, 29.50, 29.47, 29.45, 29.16, 29.11, 28.81, 26.08, 20.68, 20.63. IR ( $\text{cm}^{-1}$ ,  $\text{CH}_2\text{Cl}_2$ ) 2925, 2854, 2111, 1747, 1690. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{H}]^+\text{C}_{31}\text{H}_{53}\text{N}_3\text{O}_{12}\text{S}$ : 692.3383; Found: 692.3432.

**Synthesis of compound 9:** The mixture of anomers **8** (0.4 g, 0.57 mmol) were dissolved in mixture of THF:AcOH:Ac<sub>2</sub>O [v/v/v, (3:2:1), 14 mL] and added Zn dust (0.7 g, 10.7 mmol). The mixture was stirred for overnight. The zinc dust was filtered off on celite bed and the filtrate was diluted with DCM then washed with sat NaHCO<sub>3</sub> solution. The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated. The crude was purified by column chromatography (EtOAc/Hexane, 0.7/1) to afford **9** (0.28 g, 68%) with mixture of  $\alpha$  and  $\beta$  anomers ( $\alpha/\beta$  1:0.6 as determined by <sup>1</sup>H NMR spectroscopy) as a white foam. The individual isomer was not isolated from the crude. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  6.63 (d,  $J = 9.5$  Hz, 1H), 5.33 (d,  $J = 2.8$  Hz, 1H), 4.98 (dd,  $J = 11.2, 3.4$  Hz, 1H), 4.81 (d,  $J = 8.6$  Hz, 1H), 4.33 – 4.03 (m, 3H), 3.95 – 3.80 (m, 2H), 3.78 – 3.55 (m, 10H), 3.46 – 3.40 (m, 2H), 2.87 (t,  $J = 7.4$  Hz, 2H), 2.33 (s, 3H), 2.17 (s, 3H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  196.08, 170.82, 170.77, 170.48, 170.47, 170.45, 170.41, 170.38, 102.40, 98.07, 71.74, 71.56, 71.42, 71.23, 70.67, 70.56, 70.54, 70.51, 70.44, 70.01, 69.91, 69.89, 68.55, 67.39, 67.35, 66.73, 66.65, 61.91, 61.63, 50.38, 47.53, 30.64, 29.58, 29.56, 29.54, 29.49, 29.46, 29.44, 29.14, 29.10, 28.80, 26.06, 26.05, 23.18, 23.16, 20.77, 20.72. IR ( $\text{cm}^{-1}$ ,  $\text{CH}_2\text{Cl}_2$ ) 3305, 2925, 2854, 1745, 1688. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{H}]^+\text{C}_{33}\text{H}_{57}\text{NO}_{13}\text{S}$ : 708.3629; Found: 708.3619.

**Synthesis of compound 10a and 10b:** The mixture of anomers **9** (0.2 g, 0.28 mmol) was dissolved in methanol (10 mL) and NaOMe (61 mg, 1.13 mmol) was added. After stirring 2 h at room temperature the reaction mixture was neutralized with Amberlite® IR-120H resin and filtered, concentrated to afford mixture of **10a** and **10b** (0.11 g, 73%), which was purified by preparative TLC to separate the anomers (using 8:2 of DCM:MeOH).  $\alpha$ -isomer **10a** <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta$  4.73 (d,  $J = 3.7$  Hz, 1H), 4.04 (dd,  $J = 10.9, 3.7$  Hz, 1H), 3.83 (d,  $J = 2.9$  Hz, 1H), 3.81 – 3.73 (m, 2H), 3.72 – 3.66 (m, 1H), 3.62 – 3.56 (m, 4H), 3.56 – 3.49 (m, 7H), 3.48 – 3.44 (m, 2H), 3.33 (t,  $J = 6.6$  Hz, 2H), 2.36 (t,  $J = 7.1$  Hz, 2H), 1.89 (s, 3H), 1.50 – 1.43 (m, 4H), 1.26 – 1.17 (s, 14H). <sup>13</sup>C NMR (100 MHz, D<sub>2</sub>O)  $\delta$  173.95, 97.40, 71.18, 71.05, 69.94, 69.86, 69.81, 69.74, 68.53, 67.97, 66.50, 61.19, 49.86, 34.15, 29.92, 29.88, 29.72, 29.46, 29.39, 28.59, 26.14, 24.39, 22.18. HRMS (ESI)  $m/z$ : calc'd for  $\text{C}_{25}\text{H}_{49}\text{NO}_9\text{SNa}$ : 562.3026, Found: 562.3026.  $\beta$ -isomer <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta$  4.36 (d,  $J = 8.4$  Hz, 1H), 3.88 – 3.84

(m, 1H), 3.81 – 3.73(m, 2H), 3.66 – 3.59 (m, 4H), 3.55 – 3.50 (m, 9H), 3.51 – 3.46 (m, 2H), 3.33 (t, J = 6.8 Hz, 2H), 2.37 (t, J = 7.1 Hz, 2H), 1.89 (s, 3H), 1.51 – 1.44 (m, 4H), 1.26 – 1.18 (m, 14H). <sup>13</sup>C NMR (100MHz, D<sub>2</sub>O) δ 174.18, 101.60, 75.12, 71.22, 71.18, 69.94, 69.86, 69.72, 68.74, 67.82, 60.93, 52.44, 34.14, 29.93, 29.89, 29.87, 29.72, 29.47, 29.39, 28.59, 26.15, 24.41, 22.47. HRMS (ESI)m/z: calc'd for C<sub>25</sub>H<sub>49</sub>NO<sub>9</sub>SNa: 562.3026, Found: 562.3040.

### 3.4.2 Synthesis of fluorescein modified linker:



**Scheme 2.** Synthesis of fluorescein linker.

Fluorescein modified linker was synthesized by following synthetic reported protocol.<sup>46</sup>

### 3.4.3 Synthesis of fluorescein functionalized glyco-AuNPs (S-1, S-2, R-1 and R-2):

To the respective 0.5 ml gold nanoparticle solution in PBS buffer pH 7.4 (1 mL, 0.1 M), 0.5 mg of 10a/10b and 0.1 mg of F-1 linker were added simultaneously. The resulting solution was kept at room temperature for 12 h with constant shaking. Then the solution was centrifuged, and the pellet was washed three times with Milli Q water to remove the unbound sugar molecules as well as linker to get final fluorescent glyco-AuNPs. S-3 and R-3 were synthesized by using same protocol without sugar unit.

### 3.4.4 UV profile of $\alpha$ and $\beta$ GalNAc functionalized AuNPs:

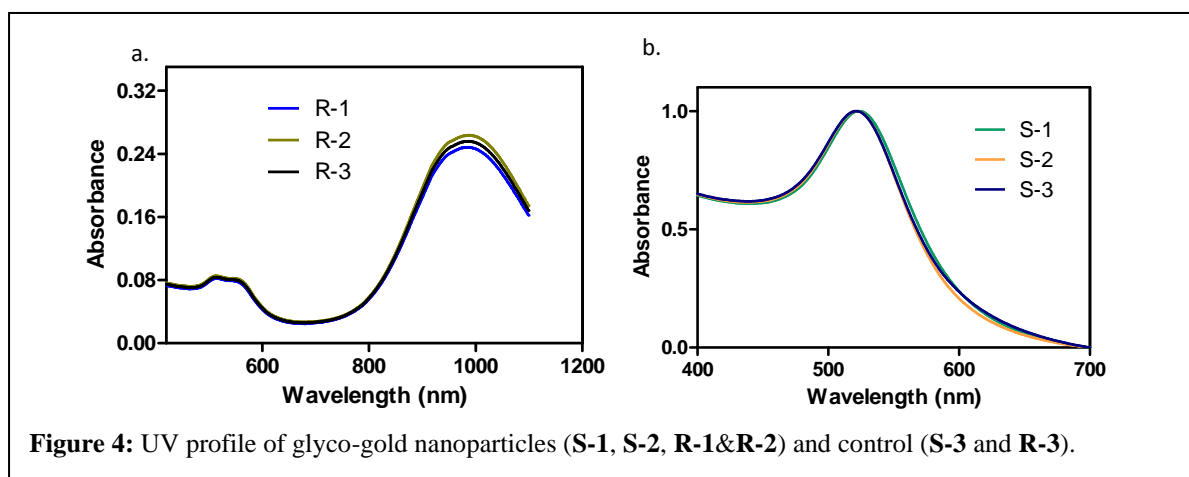


Figure 4: UV profile of glyco-gold nanoparticles (S-1, S-2, R-1&R-2) and control (S-3 and R-3).

### 3.4.5 Fluorescence spectra of fluorescein functionalized glyco-AuNPs:

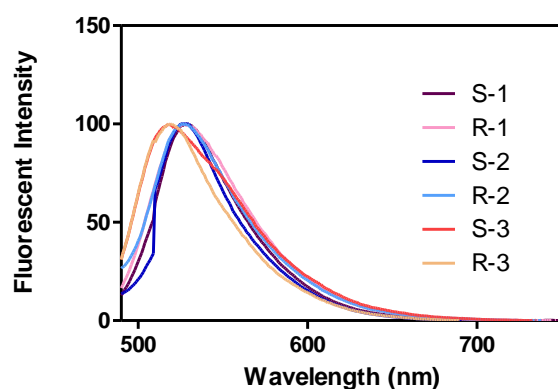


Figure 5: Fluorescence spectra of (S-1, S-2, R-1 & R-2) and control (S-3 and R-3).

Table 2. Aspect ratio of AuNPs:

Particle	Aspect ratio	Volume (nm <sup>3</sup> )
Rod	3.6: 1	4179 ± 454
Sphere	1 : 1	3945 ± 564

Aspect ratio of rod and sphere shape nanoparticles calculated by the ratio of major axis to the minor axis.<sup>47</sup>

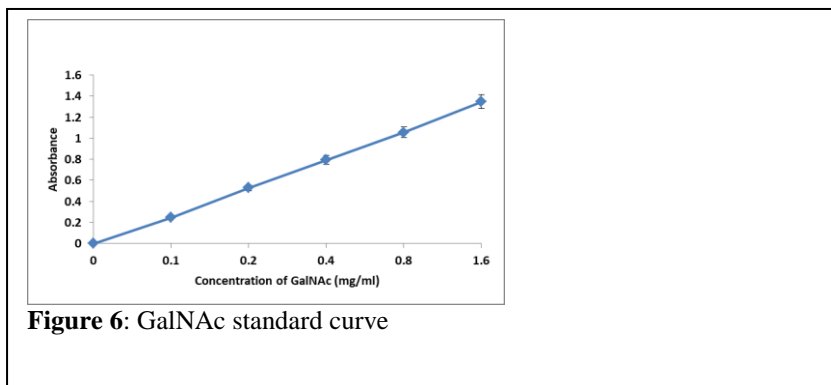
### 3.4.6 Zeta potential studies:

The electrostatic potential on the gold nanoparticles was measured by applying 1 Volt per meter to the AuNP solution.

### 3.4.7 Phenol-sulfuric acid method to quantify sugars on AuNPs:

The concentration of GalNAc sugars on AuNPs were determined by the phenol-sulfuric acid method. Briefly, 100  $\mu$ L sugar functionalized-AuNPs were mixed with concentrated sulfuric

acid (750  $\mu$ L, 100%) and aqueous phenol solution (5% w/v, 100  $\mu$ L) in the test tube and heated to 80°C. After 5 min, solution was cooled to room temperature and the absorbance coefficient at 490 nm was measured. The sugar concentration was estimated by comparing the absorption of the sample with a standard curve.



**Figure 6:** GalNAc standard curve

### 3.4.8 Cytotoxicity assay:

MTT assay was used to measure cytotoxicity in upright as well as inverted conditions.

**3.4.8.1 Upright condition:** HepG2 cells ( $1 \times 10^6$  per well) were plated in 24 well plate in DMEM medium and incubated overnight for adhering followed by treatment with different concentration of AuNPs and further incubation for 24 h. Afterward 50  $\mu$ l of MTT reagent was added to each well and incubated plates further for 4h at 37 °C. Purple precipitate formed was dissolved by adding 100  $\mu$ l of DMSO and plate was read at 570 nm.

**3.4.8.2 Inverted condition:** Same protocol was used only cells were grown on coverslip then coverslip was kept inverted on support of two pegs in DMEM media containing AuNPs. After 24 h incubation coverslip transfer in fresh well and MTT reagent was added and further plate incubated for 4h at 37 °C. Purple precipitate formed was dissolved by adding 100  $\mu$ l of DMSO and plate was read at 570 nm.

### 3.4.9 Upright cell uptake studies using ICP-MS:

HeLa and HepG2 cells ( $2 \times 10^4$  cells per plate) were plated on coverslips and allowed to adhere overnight. Cells were then treated with the rod and sphere AuNPs ( $5 \times 10^8$  nanoparticles) for 4h and 24 h at 37 °C. The medium was removed, and the cells were washed with PBS (3 times), then trypsinized and centrifuged. Cell pellets were digested at 85 °C with 200  $\mu$ l of fresh aquaregia for 4 h. Then each digested sample was diluted to 6 ml with Millipore water. The concentration of gold, determined by ICP-MS (Thermo-Fisher Scientific, Germany), was converted into the number of AuNPs per cell.

### 3.4.10 Inverted uptake Studies using ICP-MS:

**Table 3.** ICP-MS data for AuNPs cellular internalization (Average of triplicate).

Sr. No.	Compound	Concentration of AuNPs in ppb					
		4 h			24 h		
		HepG2		HeLa	HepG2		HeLa
		Upright uptake	Inverted uptake	Upright uptake	Upright uptake	Inverted uptake	Upright uptake
1	<b>R-1</b>	90.604	13.151	10.438	130.185	39.288	40.076
2	<b>R-2</b>	130.935	109.531	20.963	180.562	120.933	60.823
3	<b>R-3</b>	18.876	10.656	9.976	35.915	12.245	20.313
4	<b>S-1</b>	15.589	11.487	12.052	30.698	7.752	11.683
5	<b>S-2</b>	72.623	12.778	7.806	65.082	15.259	10.426
6	<b>S-3</b>	10.692	8.427	9.649	12.076	8.652	13.626

This experiment was performed using the reported procedure.<sup>48</sup> Briefly, In a 24 well tissue cultureplate, HepG2 cells were seeded (20,000 cells per well) on coverslips and incubated overnight foradhering. The coverslips were washed with the PBS buffer and placed inverted on the two equalsize (1mm) glass pegs and media was added until the coverslip will deep. Next, sphere or rodAuNPs ( $5 \times 10^8$  nanoparticles) were added and incubated for 4 h and 24 h. coverslips werewashed with PBS then cells were trypsinized and centrifuged. Cell pellets were digested at 85 °Cwith 200  $\mu$ l of fresh aqua regia for 4 h. Then each digested sample was diluted to 6 ml withMillipore water. The concentration of gold, determined by ICP-MS (Thermo-Fisher Scientific,Germany), was converted into the number of AuNPs per cell.

### 3.4.11 Confocal microscopy studies:

HepG2 cells ( $2 \times 10^4$  cells/well) were seeded on coverslips in 24 well plates and incubated overnight for adhering. For upright experiment, AuNPs were added to wells as such, while in case of inverted conditions, coverslips were kept inverted on the support of two glass pegs and the AuNPs were added and incubated for 24h at 37 °C. Finally, coverslips were washed with PBS and cells were fixed with paraformaldehyde solution followed by mounting on slides in the presence of mounting media for confocal microscopy.

### 3.5 References:

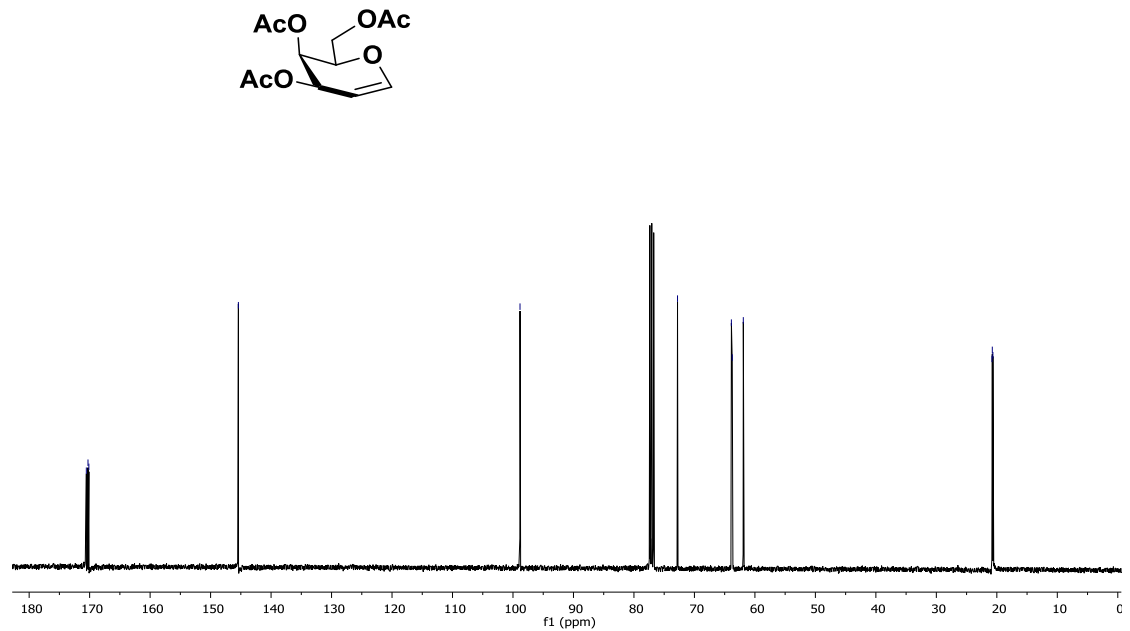
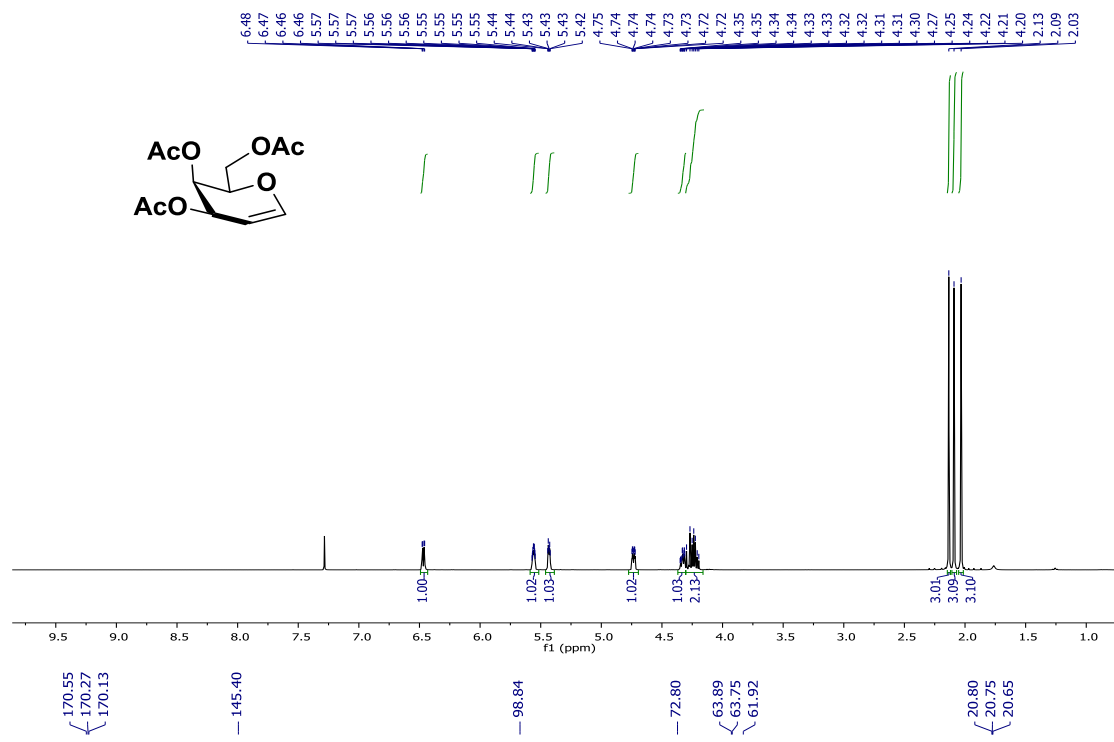
1. L. Möckl, *Front Cell Dev Biol.*, 2020, **8**, 253.
2. A. Buffone, V. M. Weaver, *J. Cell Biol.*, 2020, **219**, e201910070.
3. N. Sharon, *Sci. Am.*, 1980, **243**, 90-116.
4. N. Sharon, H. Lis, *Sci. Am.*, 1993, **268**, 82-89.
5. H. Lis, N. Sharon, *Chem. Rev.*, 1998, **98**, 637–674.
6. Sharma, M. Vijayan, *Glycobiology*, 2011, **21**, 23-33.
7. A. S. Norgren, M. Geitmann, U. H. Danielson, P. I. Arvidsson, *J. Mol. Recognit.*, 2007, **20**, 132-138.
8. C. D. Shanthamurthy, S. Leviatan Ben-Arye, N. V. Kumar, S. Yehuda, R. Amon, R. J. Woods, V. Padler-Karavani, R. Kikkeri, *J. Med. Chem.* 2021, **64**, 3367-3380.
9. A. Geissner, A. Reinhardt, C. Rademacher, T. Johannssen, J. Monteiro, B. Lepenies, M. Thépaut, F. Fieschi, J. Mrázková, M. Wimmerova, F. Schuhmacher, S. Götze, D. Grünstein, X. Guo, H. S. Hahm, J. Kandasamy, D. Leonori, C. E. Martin, S. G. Parameswarappa, S. Pasari, M. K.Schlegel, H.Tanaka, G.Xiao, Y.Yang, C. L.Pereira, C.Anish, P. H.Seeberger. *Proc. Natl. Acad. Sci. USA.*, 2019, **116**, 1958-1967.
10. T. K. Dam, R. Roy, D. Pagé, C. F. Brewer, *Biochemistry*, 2002, **41**, 1359-63.
11. P. M. Chaudhary, S. Toraskar, R. Yadav, A. Hande, R. A. Yellin, R. Kikkeri, *Chem. Asian J.*, 2019, **14**, 1344-1355.
12. M. Delbianco, P. Bharate, S. Varela-Aramburu, P. H. Seeberger, *Chem. Rev.* 2016, **116**, 1693-752.
13. M. Gade, C. Alex, S. Leviatan Ben-Arye, J. T. Monteiro, S. Yehuda, B. Lepenies, V. Padler-Karavani, R. Kikkeri, *Chembiochem.*, 2018, **19**, 1170-1177.
14. S. Toraskar, M. Gade, S. Sangabathuni, H. V. Thulasiram, R. Kikkeri, *ChemMedChem.*, 2017, **12**, 1116-1124.
15. P. Kumar, P. Kanjilal, R. Das, T. K.Dash, M.Mohanan, T.-N. LecN. V.Rao, B. Mukhopadhyay, R. Shunmugam, *Carbohydr. Res.*, 2021, **508**, 108397.
16. M. Sojitra, S. Sarkar, J. Maghera, E. Rodrigues, E. J. Carpenter, S. Seth, D. V. Ferrer, N. J. Bennett, R. Reddy, A. Khalil, X. Xue, M. R. Bell, R. B. Zheng, P. Zhang, C. Nycholat, J. J. Bailey, C. C. Ling, T. L. Lowary, J. C. Paulson, M. S. Macauley, R. Derda, *Nat. Chem. Biol.*, 2021, **17**, 806-816.
17. Z. Lyu, L. Ding, A. Tintaru, L. Peng *Acc. Chem. Res.* 2020, **53**, 2936-2949.
18. L. L. Kiessling, J. C. Grim, *Chem. Soc. Rev.* 2013, **42**, 4476–4491.

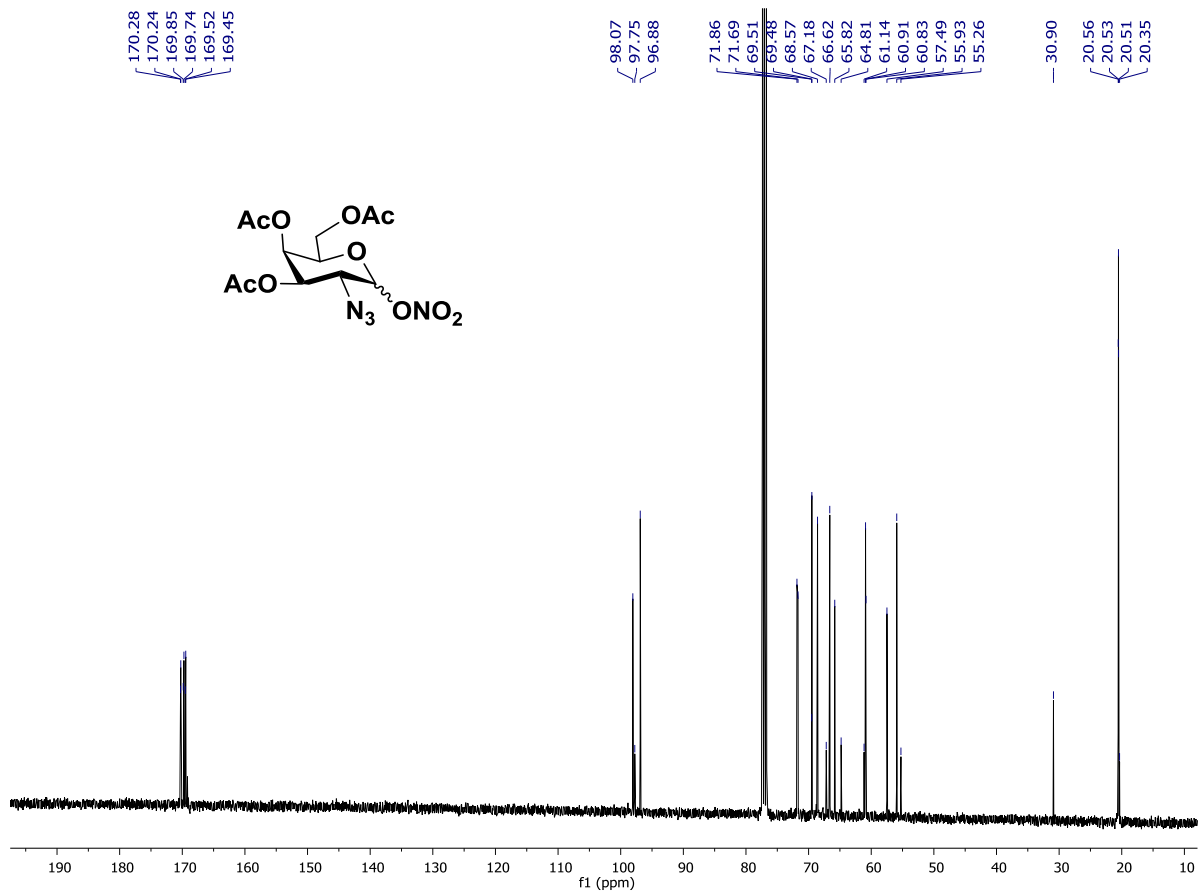
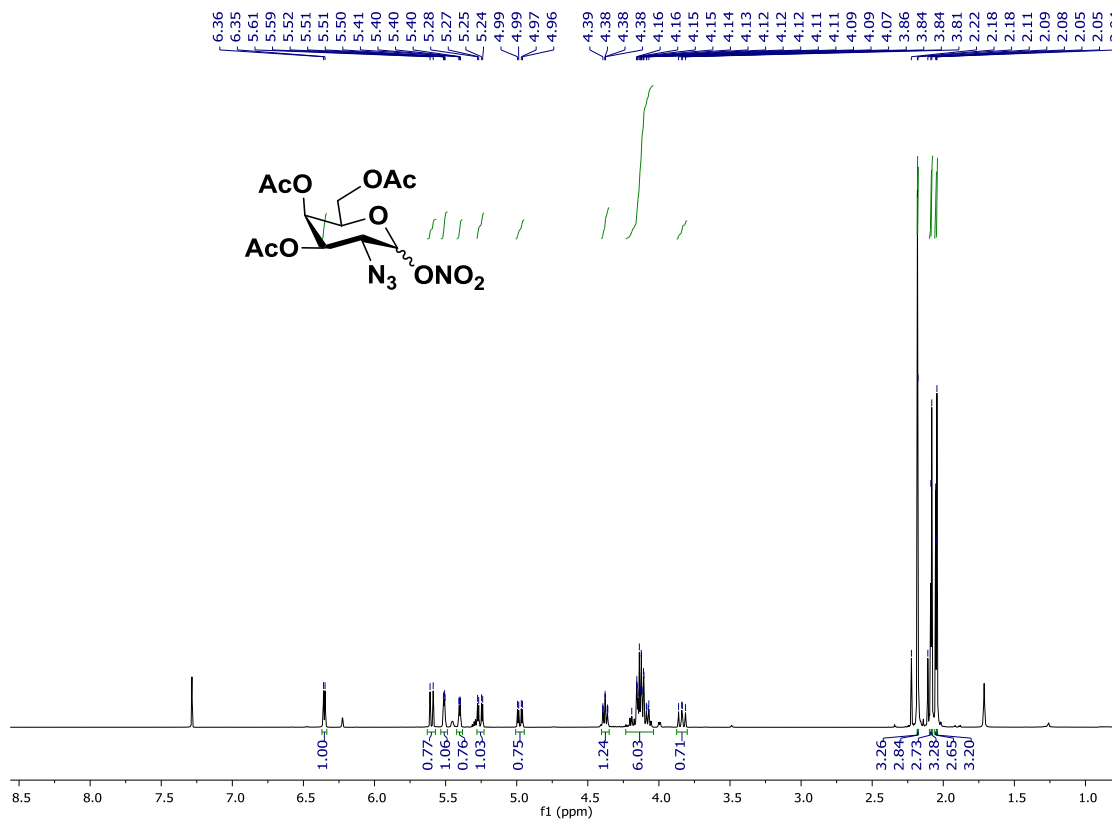


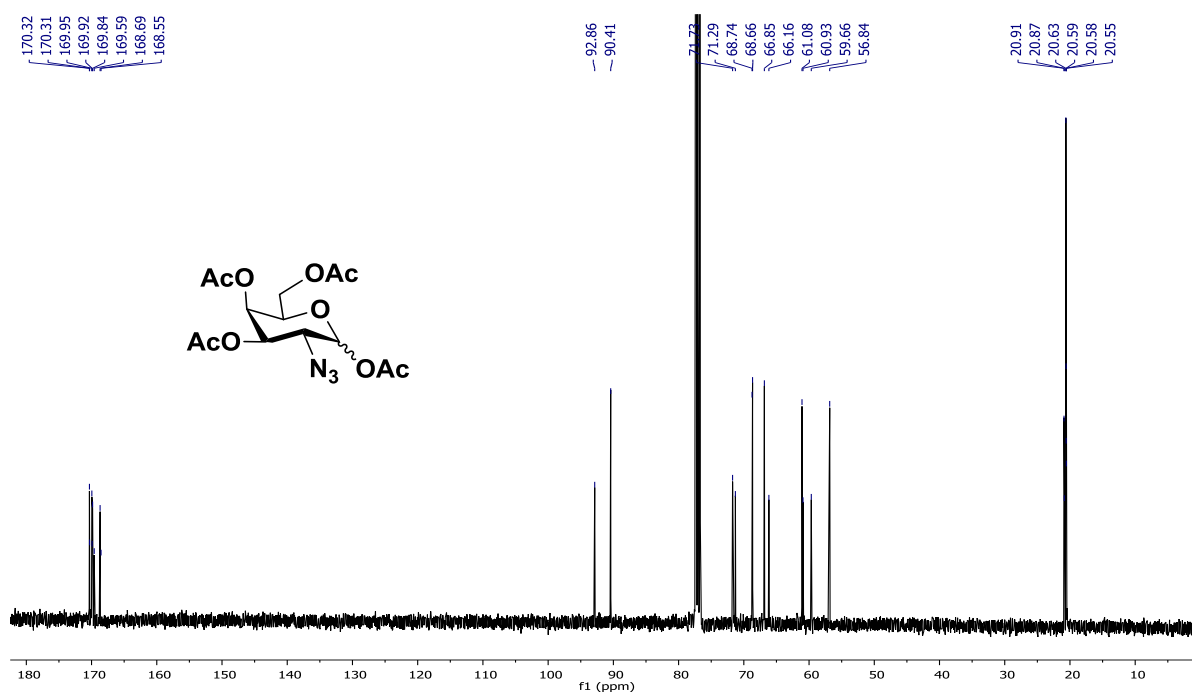
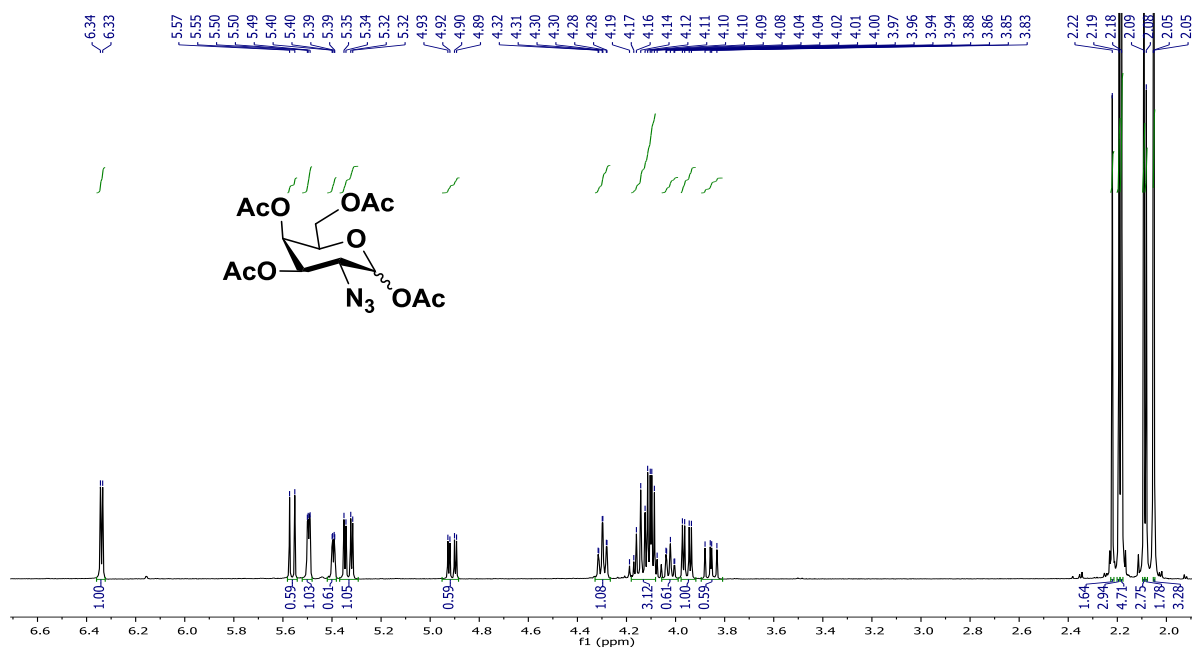
19. D. Kim, N. Rahhal, C. Rademacher, *Front. Chem.*, 2021, **9**, 669969.
20. R. Das, B. Mukhopadhyay, *Carbohydr. Res.*, 2021, 108394.
21. P. J. Hernando, S. Dedola, M. J. Marín, R. A. Field, *Front. Chem.*, 2021, **9**, 668509.
22. O. B. Adewale, H. Davids, L. Cairncross, S. Roux, *Int. J. Toxicol.* 2019, **38**, 357-384.
23. M. S. Park, J. Park, S. K. Jeon, T. H. Yoon, *Nanosci. Nanotechnol.* 2013, **13**, 7264-70.
24. J. Manson, D. Kumar, B. J. Meenan, D. Dixon, *Gold. Bull.*, 2011, 44, 99-105.
25. P. M. Valencia, M. H. Hanewich-Hollatz, W. Gao, F. Karim, R. Langer, R. Karnik, O. C. Farokhzad, *Biomaterials*, 2011, **32**, 6226-6233.
26. C. M. Yeh, M. C. Chen, T. C. Wu, J. W. Chen, C. H. Lai, *Chem Asian J.* 2021, doi: 10.1002/asia.202100865.
27. Y. H. Tang, H. C. Lin, C. L. Lai, P. Y. Chen, C. H. Lai, *Biosens Bioelectron.*, 2018, **116**, 100-107.
28. H. Y. Li, H. C. Lin, B. J. Huang, A. Z. Kai Lo, S. Saidin, C. H. Lai, *Langmuir*, 2020, **36**, 11374-11382.
29. L. Shi, J. Zhang, M. Zhao, S. Tang, X. Cheng, W. Zhang, W. Li, X. Liu, H. Peng, Q. Wang, *Nanoscale*, 2021, **13**, 10748-10764.
30. M. Ostermann, A. Sauter, Y. Xue, E. Birkeland, J. Schoelermann, B. Holst, M. R. Cimpan. *Sci. Rep.*, 2020, **10**, 142.
31. S. C. McCormick, F. H. Kriel, A. Ivask, Z. Tong, E. Lombi, N. H. Voelcker, C. Priest, *Micromachines*, 2017, **8**, 124.
32. J. Kalikanda, Z. Li, *J. Org. Chem.*, 2011, **76**, 5207-5218.
33. A. L. Parry, N. A. Clemson, J. Ellis, S. S. Bernhard, B. G. Davis, N. R. Cameron, *J. Am. Chem. Soc.*, 2013, **135**, 9362-9365.
34. S. Sangabathuni, R. V. Murthy, P. M. Chaudhary, B. Subramani, S. Toraskar, R. Kikkeri, *Sci. Rep.*, 2017, **7**, 4239.
35. S. Sangabathuni, R.V. Murthy, P. M. Chaudhary, M. Surve, A. Banerjee, R. Kikkeri, *Nanoscale*, 2016, **8**, 12729-12735.
36. P. M. Chaudhary, S. Sangabathuni, R. V. Murthy, A. Paul, H. V. Thulasiram, R. Kikkeri, *Chem. Commun.*, 2015, **51**, 15669-15672.
37. Y. Li, G. Huang, J. Diakur, L. I. Wiebe, *Curr. Drug. Deliv.*, 2008, **5**, 299-302.
38. J. Wu, M. H. Nantz, M. A. Zern. *Front. Biosci.*, 2002, **7**, 717-725.
39. M. M. Janas, M. K. Schlegel, C. E. Harbison, V. O. Yilmaz, Y. Jiang, R. Parmar, I.

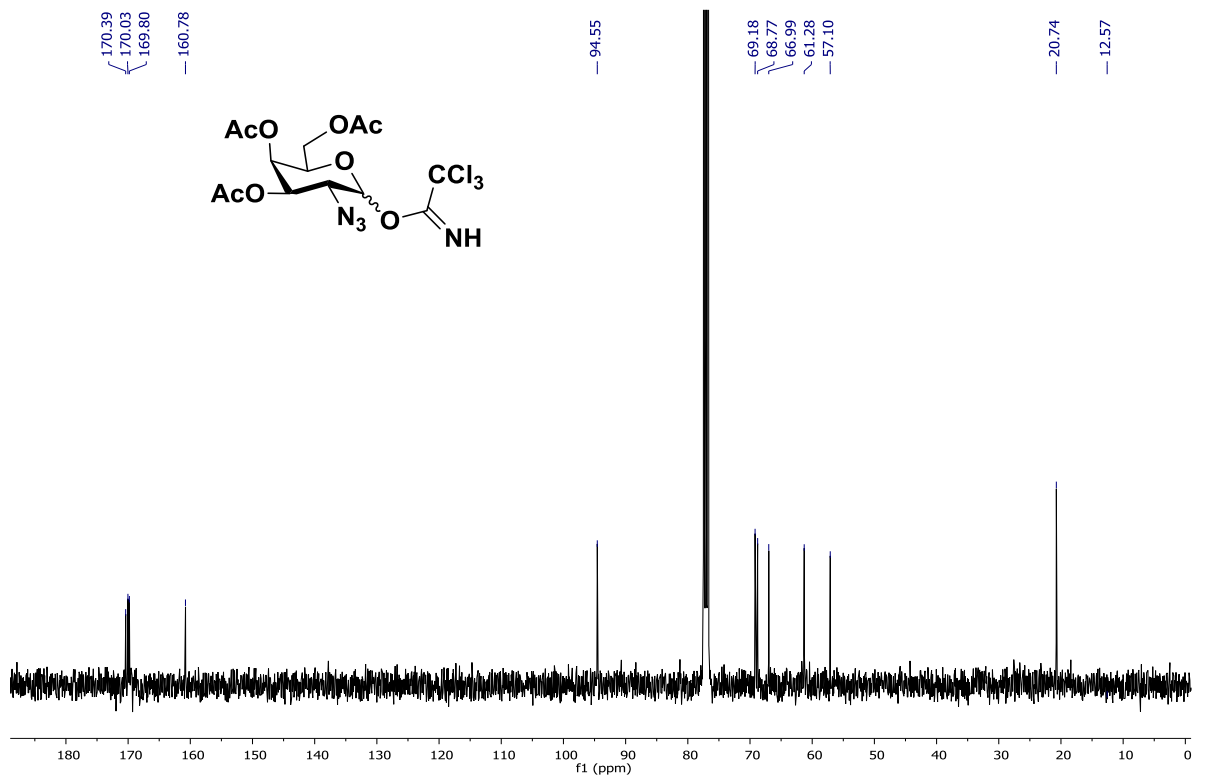
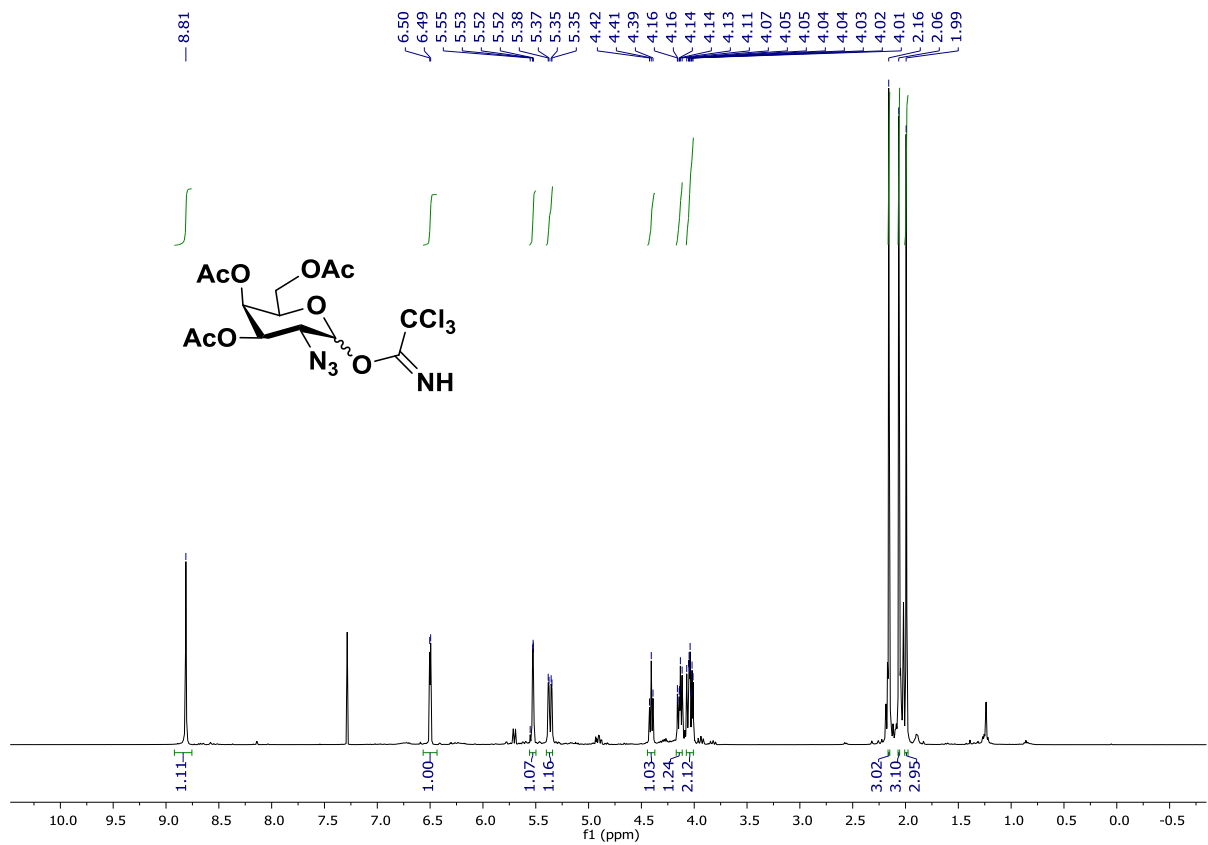
- Zlatev, A. Castoreno, H. Xu, S. Shulga-Morskaya, K. G. Rajeev, M. Manoharan, N. D. Keirstead, M. A. Maier, V. Jadhav, *Nat. Commun.*, 2018, **9**, 723.
40. C. A. Sanhueza, M. M. Baksh, B. Thuma, M. D. Roy, S. Dutta, C. Prévile, B. A. Chrnyk, K. Beaumont, R. Dullea, M. Ammirati, S. Liu, D. Gebhard, J. E. Finley, C. T. Salatto, A. King-Ahmad, I. Stock, K. Atkinson, B. Reidich, W. Lin, R. Kumar, M. Tu, E. Menhaji-Klotz, D. A. Price, S. Liras, M. G. Finn, V. Mascitti, *J. Am. Chem. Soc.*, 2017, **139**, 3528–3536.
41. D. B. Rozema, D. L. Lewis, D. H. Wakefield, S. C. Wong, J. J. Klein, P. L. Roesch, S. L. Bertin, T. W. Reppen, Q. Chu, A. V. Blokhin, J. E. Hagstrom, J. A. Wolff, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**, 12982–12987.
42. M. S. Park, J. Park, S. K. Jeon, T. H. Yoon, *J. Nanosci. Nanotechnol.*, 2013, **13**, 7264-7270.
43. R. Agarwal, V. Singh, P. Journey, L. Shi, S. V. Sreenivasan, K. Roy, *Proc. Natl. Acad. Sci.*, 2013, **110**, 17247-17252.
44. E. C. Cho, Q. Zhang, Y. Xia, *Nat. Nanotechnol.* 2011, **6**, 385-391.
45. G. Bellapadrona, A. B. Tesler, D. Grunstein, L. H. Hossain, R. Kikkeri, P. H. Seeberger, A. Vaskevich and I. Rubinstein, *Anal. Chem.*, 2012, **84**, 232–240.
46. S. Sangabathuni, R. V. Murthy, P. M. Chaudhary, B. Subramani, S. Toraskar and R. Kikkeri, *Sci. Rep.*, 2017, **7**, 4239.
47. P. M. Chaudhary, S. Sangabathuni, R. V. Murthy, A. Paul, H. V. Thulasiram and R. Kikkeri, *Chem. Commun.*, 2015, **51**, 15669.
48. R. Agarwal, V. Singh, P. Journey, L. Shi, S. V. Sreenivasan, and K. Roy, *Proc. Natl. Acad. Sci. USA*, 2013, **110**, 17247.

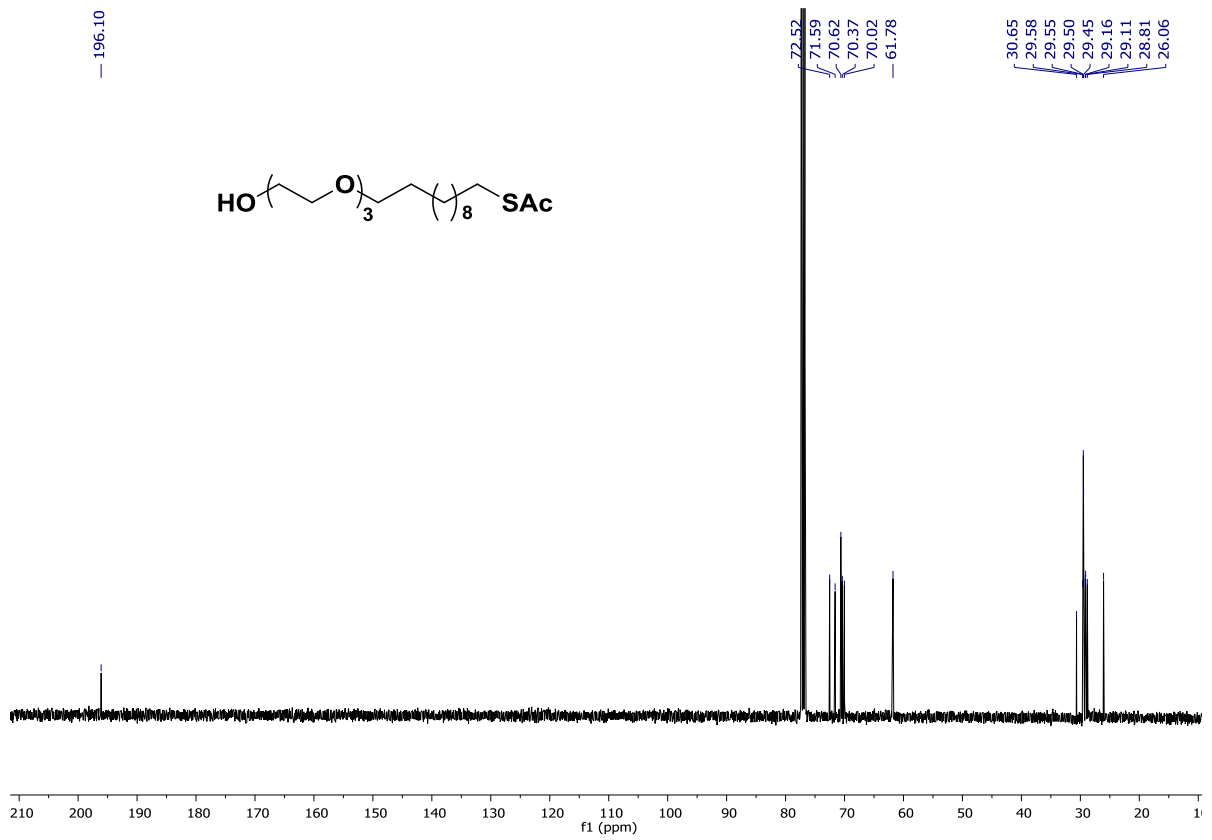
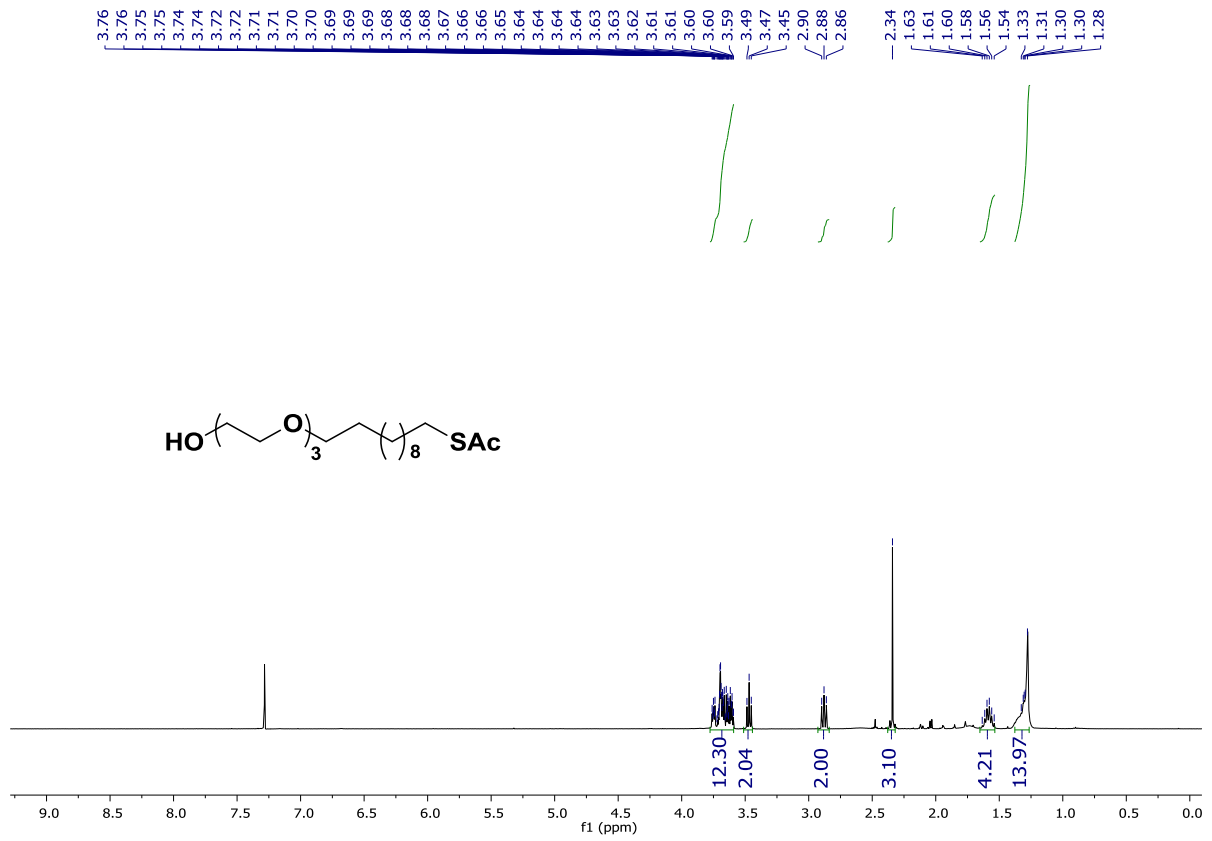
### 3.6 NMR Spectra:

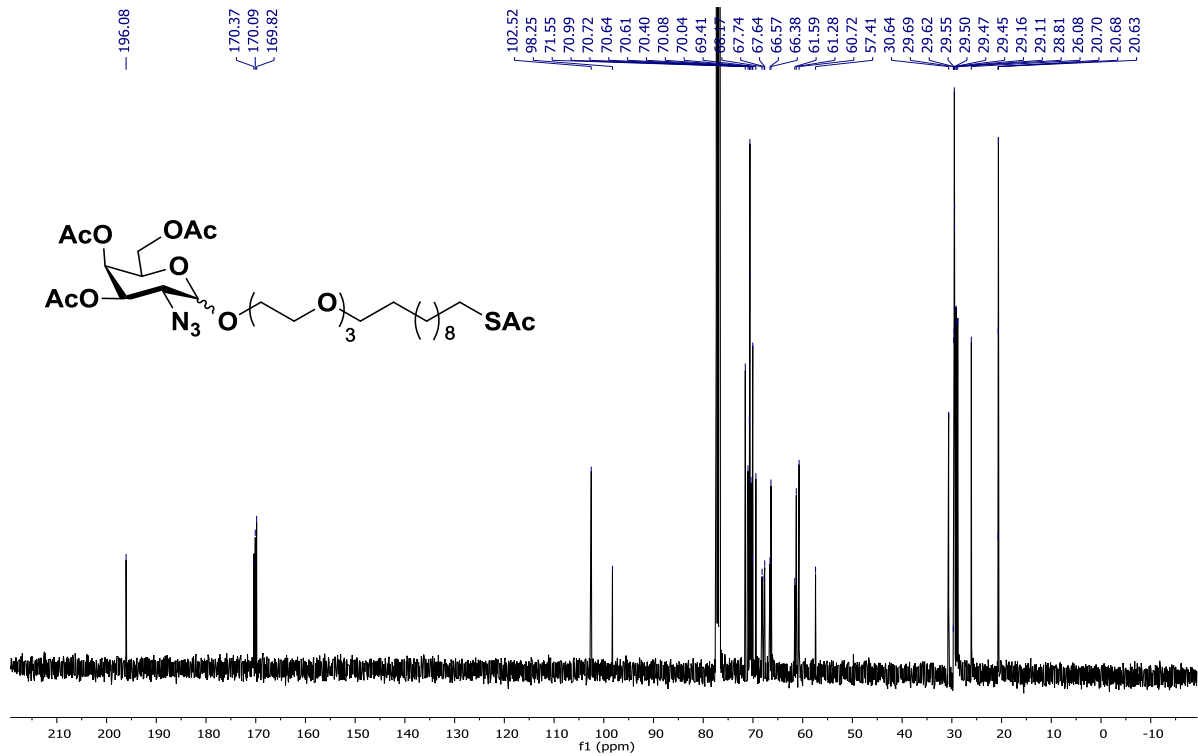
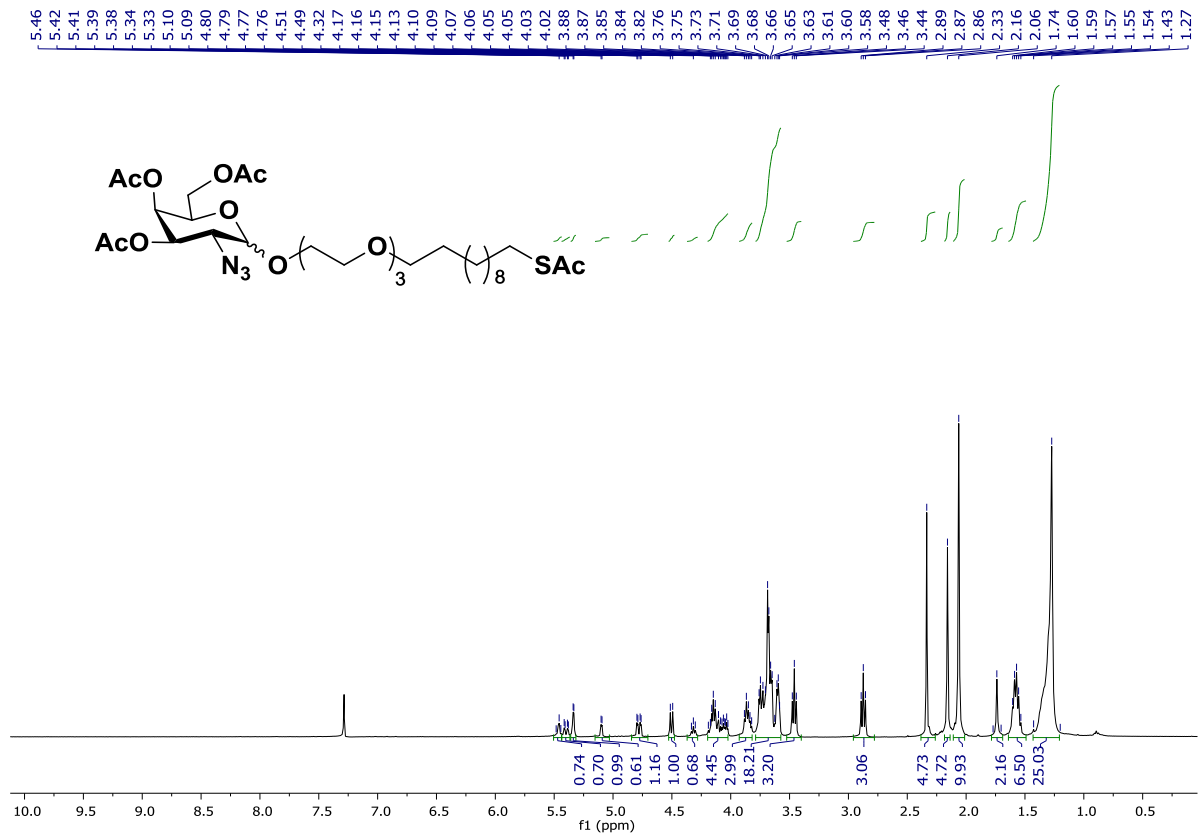




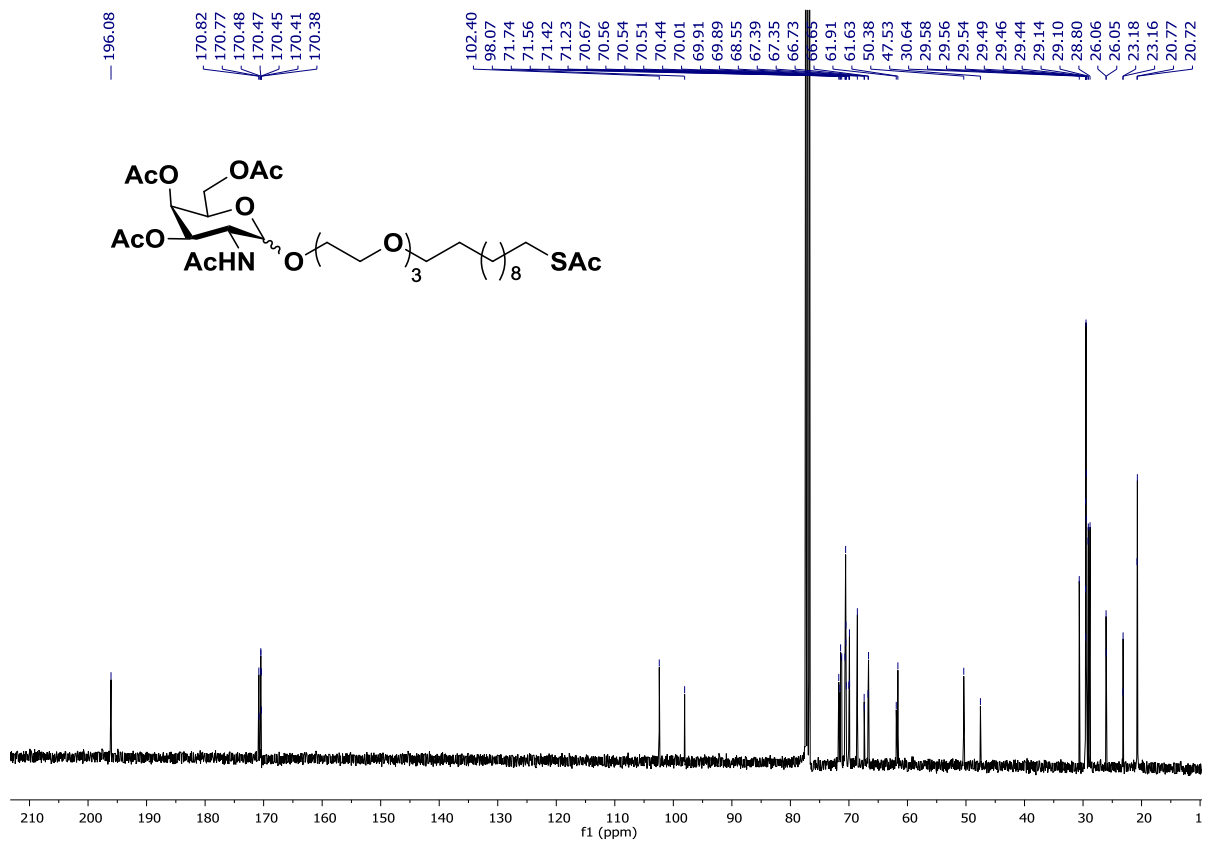
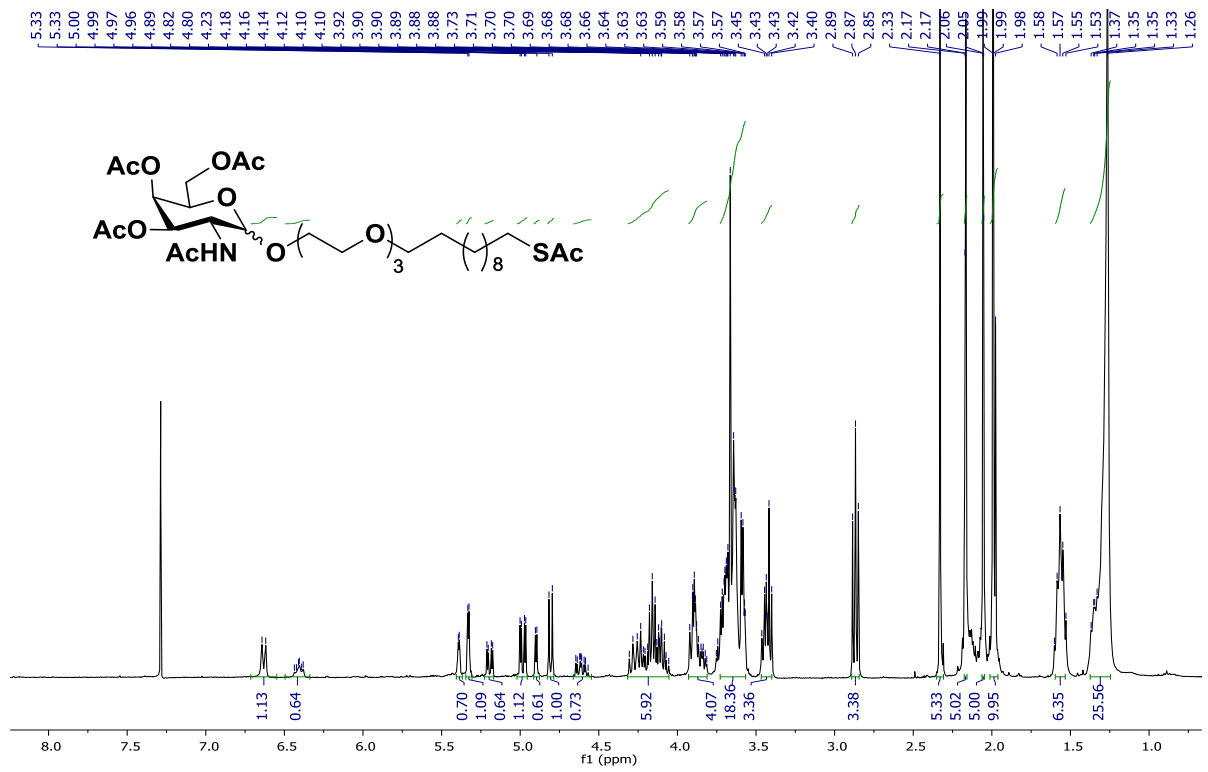


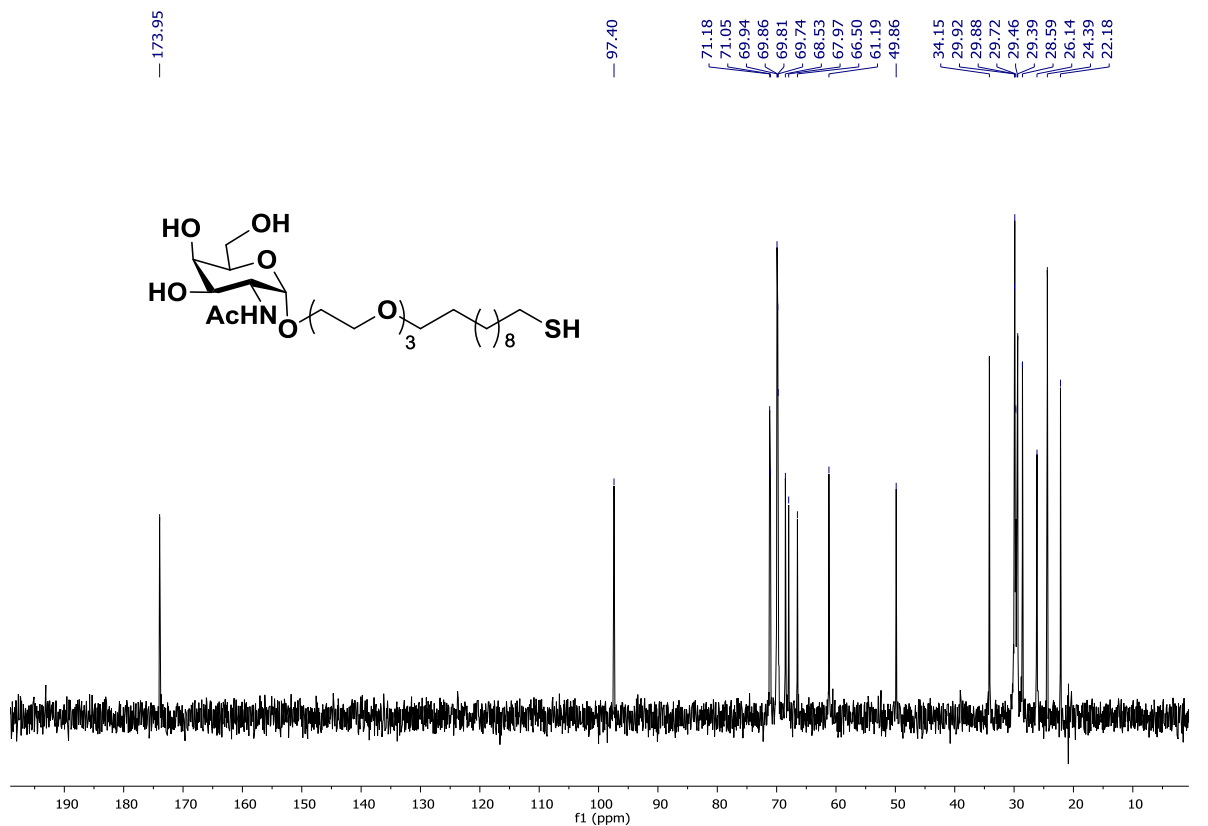
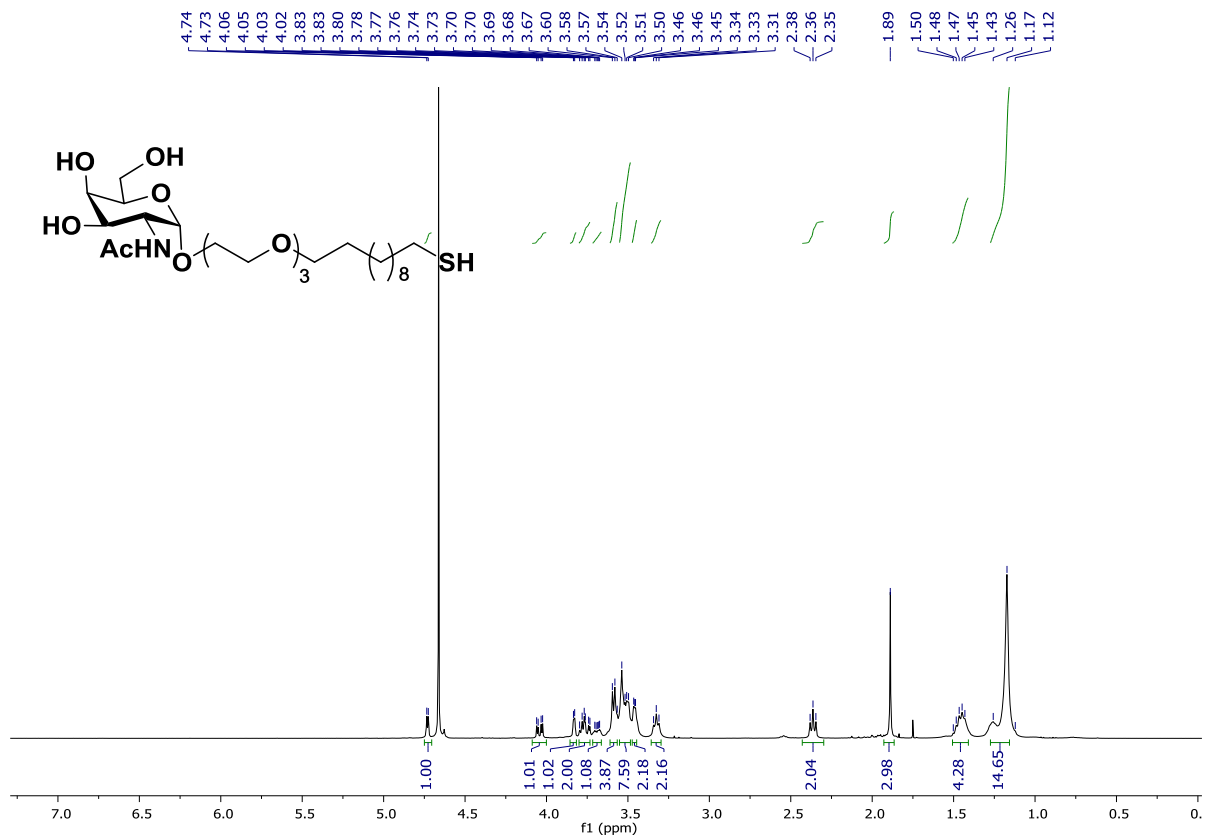










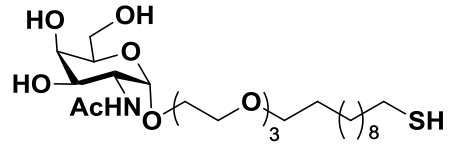
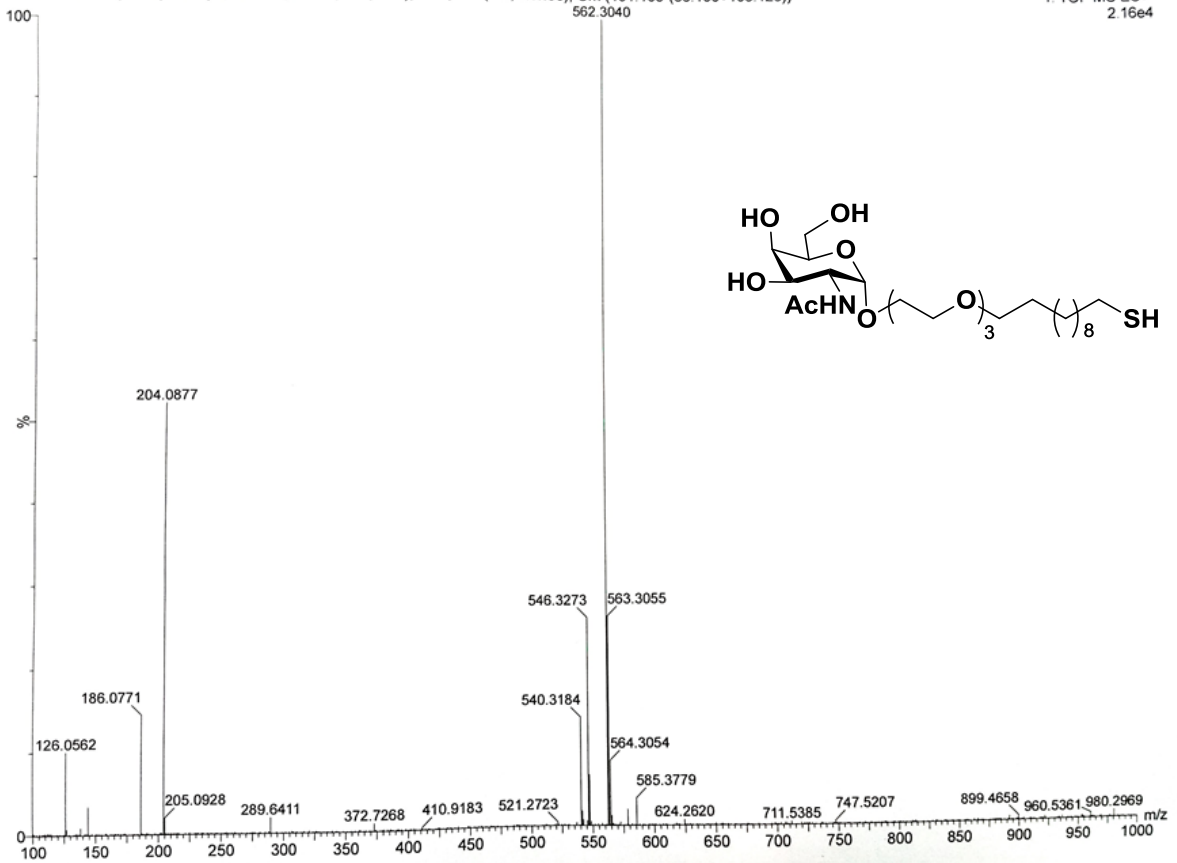


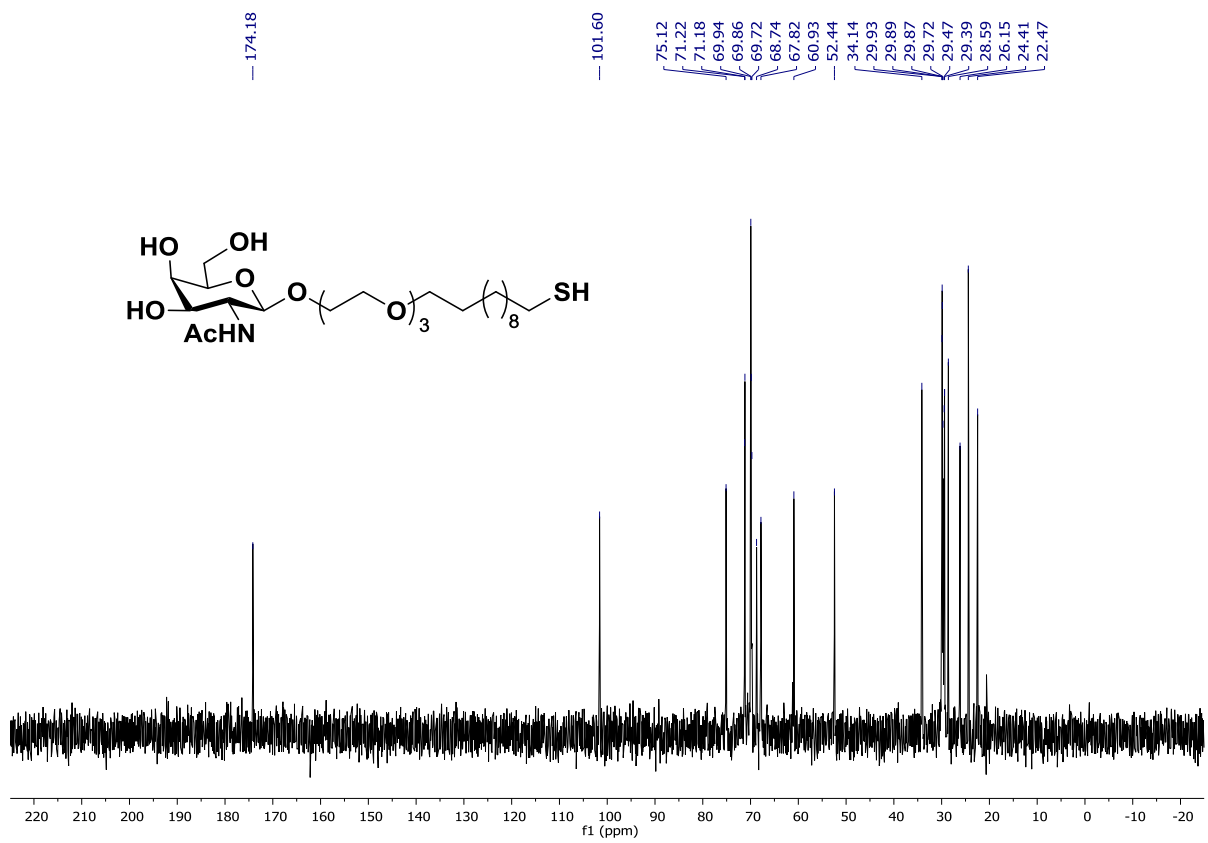
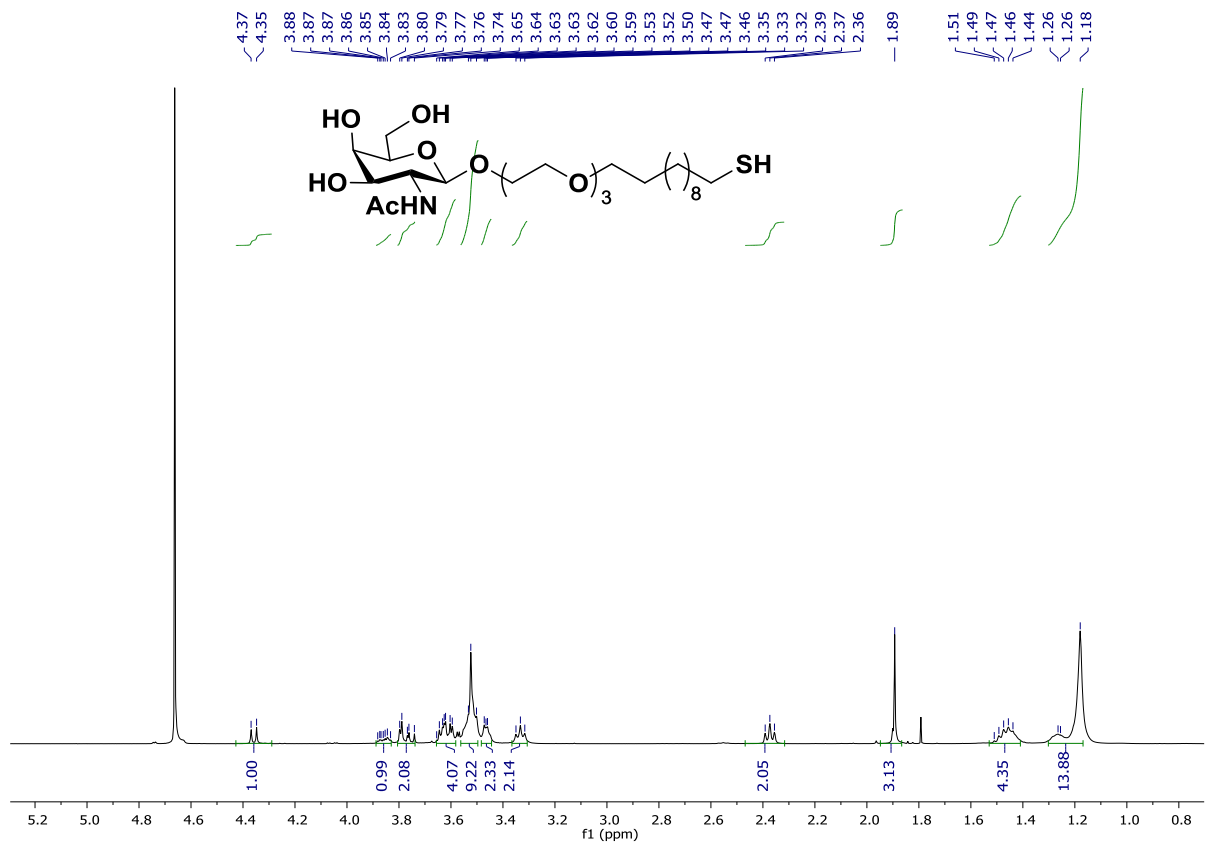
SIB BLS 3a

SIB BLS 3a 104 (1.909) AM2 (Ar,20000.0,556.28,0.00,LS 3); ABS; Sm (SG, 1x1.00); Cm (101:105-(88:100+106:126))

IISER PUNE

1: TOF MS ES+  
2.16e4



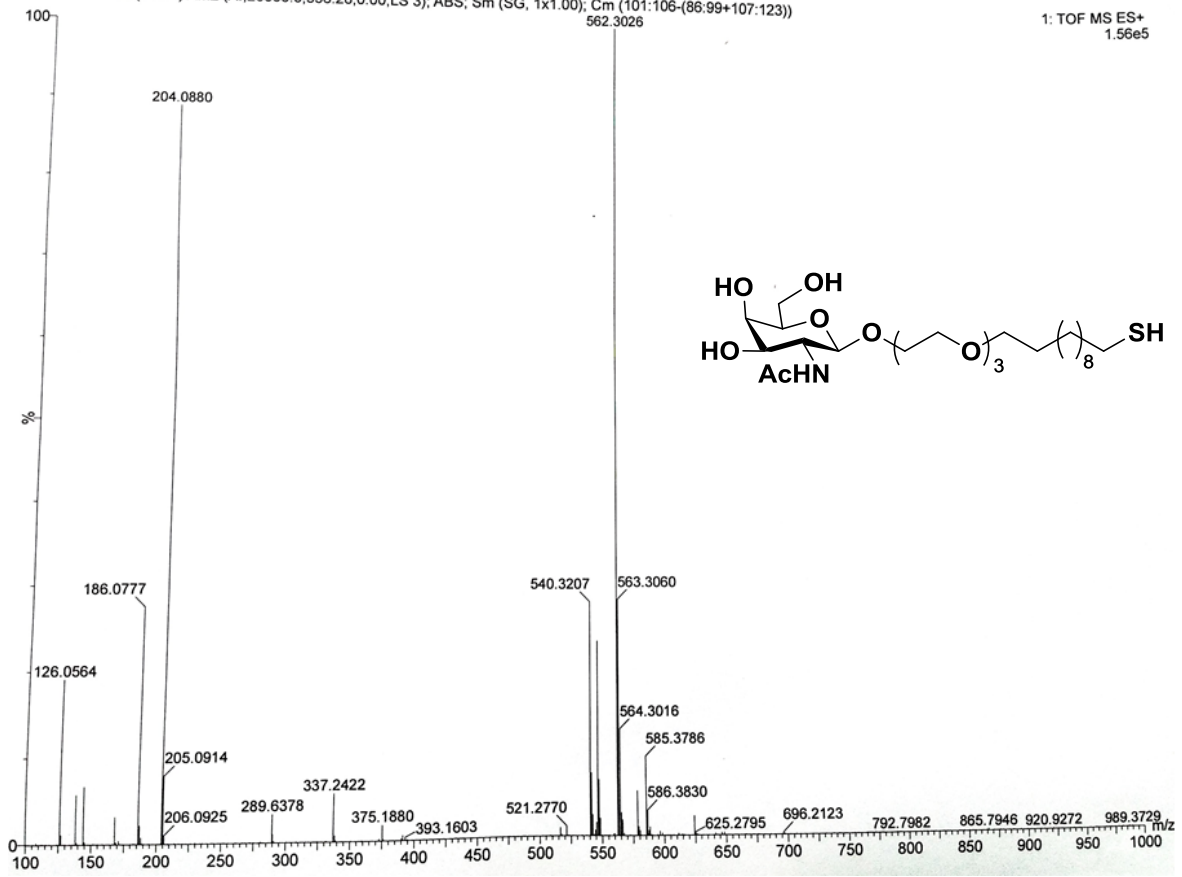


SIB BLS 3b

SIB BLS 3b 103 (1.892) AM2 (Ar,20000.0,556.28,0.00,LS 3); ABS; Sm (SG, 1x1.00); Cm (101:106-(86:99+107:123))

IISER PUNE

1: TOF MS ES+  
1.56e5



## **Chapter-4a**

# **Synthesis of Chondroitin Sulfate Analog**

---

**Abstract:**

Chondroitin sulfate (CS) exhibits exceptional multifunctional properties by its polymeric nature and sulfation patterns. Recent developments in the synthesis of CS oligosaccharides shed light on the structure-activity relationship (SAR) and its potential in biological relevance. Hence, it is essential to synthesize CS oligosaccharides with different sulfation patterns. This chapter summarizes the successful synthesis of CS di and tetra with varying sulfation patterns. In addition, we also reported all failed reactions toward hexasaccharide analogs.

## 4a.1 Introduction:

Structurally defined cell surface glycans are integral to modern medicine, vaccines and are routinely used as biomarkers for several diseases.<sup>1-4</sup> Nonetheless, the synthesis of such oligosaccharides in gram scales drives advancements in theranostics, imaging and biosensing applications.<sup>5-8</sup> Significantly much progress has recently been made in developing automated carbohydrate synthesis and chemoenzymatic strategies to synthesize a large pool of oligosaccharides and enhance our understanding of the key mechanisms of actions associated with oligosaccharide structures oligosaccharide chain length, glycosidic linkage, conformation plasticity of sugars in biological systems.<sup>9-12</sup>

Cell surface composed a wide range of carbohydrate structures, including O-glycans, N-glycans, glycosphingolipids, glycosaminoglycans (GAGs). These glycostructures play an important role in cell signaling and cell-cell interaction.<sup>13-16</sup> Among them, glycosaminoglycan are highly negatively charged polymers consists of disaccharide repeating units with heterogenous sulfation patterns, interact with positively charges biomolecules, including, growth factors, chemokines, plasma proteins and modulate physiological properties of cells.<sup>17,18</sup> The glycosaminoglycans are classified into four different groups based on the repeating disaccharides units. Namely: (a) Heparan sulfate/heparin: composed of D-glucosamine and uronic acid (L-iduronic acid or D-glucuronic acid). (b) Chondroitin sulfate (CS)/dermatan sulfate (DS): D-*N*-acetyl-galactosamine and D-glucuronic acid. (c) Keratan sulfate: D-galactose and D-*N*-acetyl-glucosamine. (d) Hyaluronan (HA): D-glucuronic acid and D-*N*-acetyl-glucosamine.<sup>19,20</sup>

Among these GAGs, CS is perhaps one of the second most important glycosaminoglycans after heparan sulfate (HS), ubiquitously found at the cell surface and extracellular matrix of many connective tissues, including cartilage, bone, skin, ligaments and central nervous system (CNS).<sup>21-25</sup> CS has been used in medical applications for decades before its structure was established. In 1914, Levena and Forge first confirmed the presence of sulfate and carboxylic acid groups and named Chondroitin Sulfuric acid.<sup>26</sup> Later Davidson et al. identified the structure of CS and its analogs such as CS-A and CS-C with different sulfation patterns.<sup>27</sup> The major breakthrough of CS medical relevance came when they were used alongside glucosamine to treat osteoarthritis (OA).<sup>28</sup> In 2004, FDA approved CS as a dietary supplement to treat Osteoarthritis (OA).<sup>29</sup> Currently, CS, collagen, glucosamine containing capsule such as Renovar are extensively used for osteoarthritis. CS inhibits various inflammatory intermediates, including nitric oxide synthase, cyclooxygenase (COX)-2, and



prostaglandin E2. However, some question remains about the molecular-level mechanism of these actions.<sup>30</sup>

The use of chemoenzymatic strategies and automated synthesis has enabled the synthesis of wide range of CS analogs to elucidate the structure-function relationship. This is exemplified of fucosylated-CS-E and its applications anti-coagulation activity.<sup>31-35</sup> In addition, the discovery of low sulfate CS interaction to placental malaria and CS-E binding the neural growth factors were a major boost to CS-based research, as they emerged as an important component of treatment for neural diseases and anti-malarial activity.<sup>36-38</sup> But, there is no active drug molecule currently available in the market as the synthesis CS oligosaccharides are difficult.

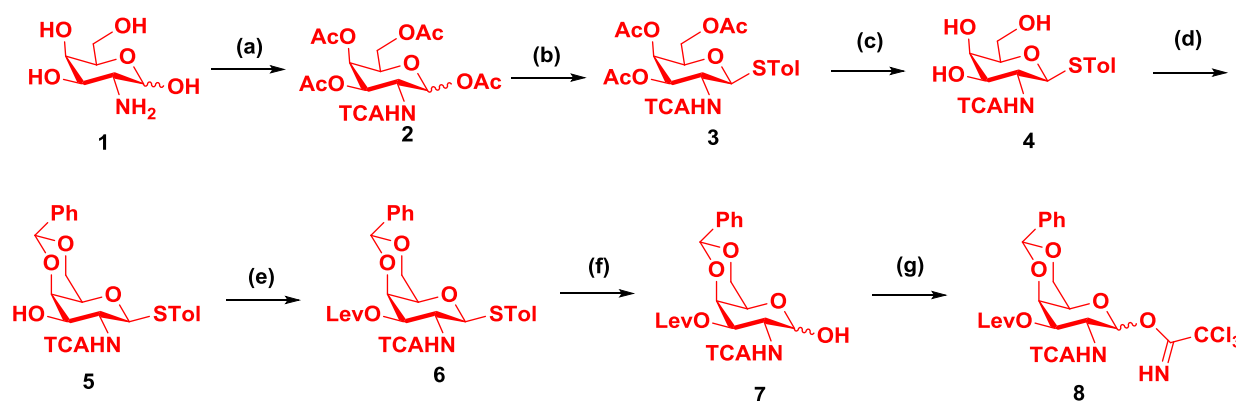
#### **4a.2 Synthesis of Chondroitin sulfate:**

Several methods have been proposed to synthesize structurally well-defined CS oligosaccharides.<sup>39-53</sup> However, most of these methods failed to yield a well-defined CS oligosaccharides library. Here, we reported the synthesis of CS oligosaccharides using a divergent strategy. Our synthetic strategy is based on a suitable protocol for synthesizing disaccharide donor and acceptor and performing [2+2] glycosylation. In the present study, we employed galactosamine donor **8** and the glucose acceptor **18** as important precursors to synthesize disaccharide moieties.

#### **4a.3 Synthesis of galactosamine building block:**

Commercially available D-galactosamine was used to prepare the required building block **8** in 7 steps. In the first step all the hydroxyl groups were protected with a labile acetyl group, and amine was protected with base sensitive TCA group in a one-pot strategy. The obtained compound **2** was dissolved in DCM and thioglycosylated with thiotoluene and  $\text{BF}_3 \cdot \text{Et}_2\text{O}$  to obtain a semi solid compound. The compound was purified by column chromatography. Then, the acetate group was deprotected in sodium methoxide, followed by benzylidene protection of 4 and 6-OH group of **4**. The 3-OH group of **5** was protected with levulinic acid group. The donor **6** was converted to a trichloroacetamidate donor. By deprotection of thiotoluene using NIS and TfOH, The compound **7** was dissolved in dry DCM, and  $\text{CsCO}_3$  &  $\text{CCl}_3\text{CN}$  were subsequently added to the reaction flask. The clear solution will turn into a blackish liquid indicating the reaction completion. Excess solvents were evaporated and purified on silica column chromatography which  $\text{Et}_3\text{N}$  basified.

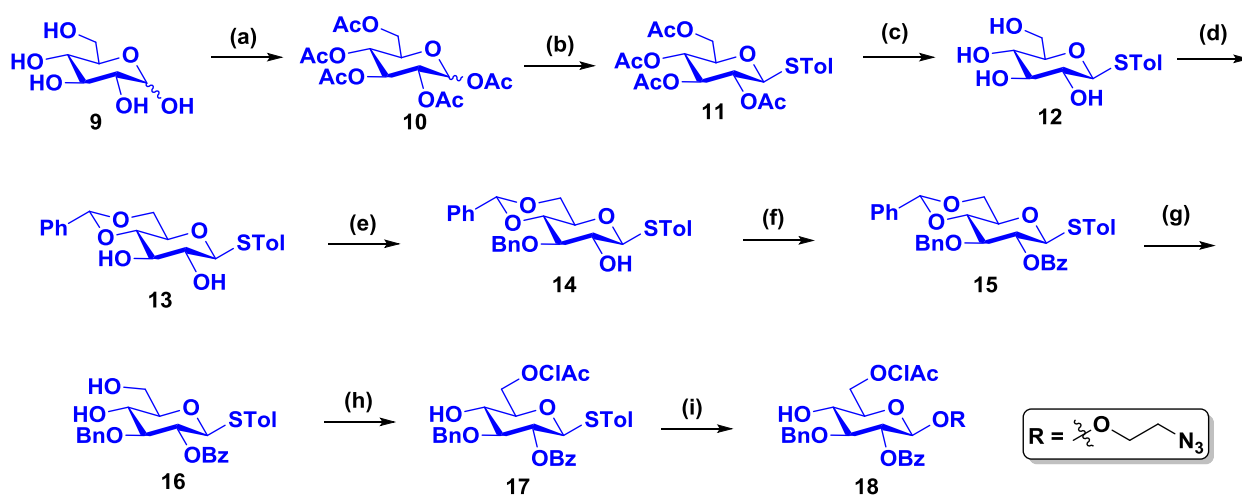
## 4a.4 Synthesis of Galactosamine building block



**Scheme 1.** (a) TCA, NaOMe, MeOH, 0 °C - RT, 1 h; Ac<sub>2</sub>O, Py, 0 °C - RT, 24 h 91 % (b) p-Thiocresol, BF<sub>3</sub>.OEt<sub>2</sub>, DCM, 0 °C - RT 12 h, 76 %; (c) NaOMe, MeOH, RT, 2 h, 74 %; (d) Ph(OMe)<sub>2</sub>, PTSA, ACN, RT, 12 h, 76 %; (e) Lev-OH, DCC, DCM, DMAP, 1 h, 80 %; (f) NIS, TfOH, DCM:H<sub>2</sub>O (10:1) RT, 30 min, 84 %; (g) Trichloroacetonitrile, DCM, CsCO<sub>3</sub>, RT, 12 h, 65%.

## 4a.5 Synthesis of Glucose building block:

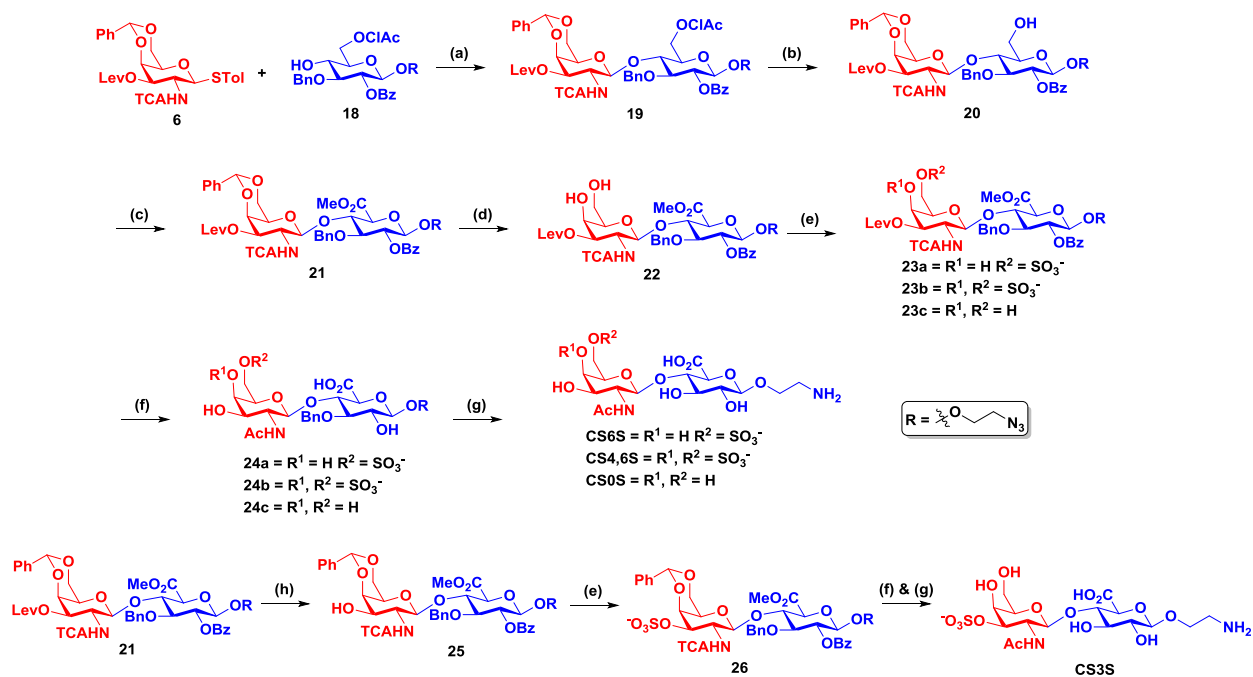
We synthesized D-glucose-based building block in 9 steps using some standard reported procedures. Briefly, all the hydroxyl groups of glucose were protected with base labile acetyl group using pyridine as both solvent & base for proton abstraction and acetic anhydride as a reagent. The compound was then dissolved in dry DCM, and p-thiocresol was added to the solution and stirred at 0 °C. Subsequently, mild Lewis acid BF<sub>3</sub>.Et<sub>2</sub>O was added dropwise to the reaction mixture. The reaction mixture was quenched and extracted with NaHCO<sub>3</sub> and purified by column chromatography to yield compound **11** as a white solid. The compound **11** was dissolved in methanol and basified with sodium methoxide to remove all acetate protecting groups. The selective 4 and 6-OH group protection using benzylidene and 3-OH by benzyl group was carried out as reported in standard butyl tin reaction, followed by benzyl bromide treatment in the presence of CsF. Finally, 2-OH of **14** was protected with benzoyl group, resulted fully protected glucose building block **15** in moderate to good yield. The compound **15** was later subjected to benzylidene deprotection and selective 6-OH chloroacetate protected and yielded desired glucose building block **17** one the acceptor for further CS synthesis. Compound **11** was glycosylated with azidoethanol linker glycosylation yielded desired another glucose **18** acceptor for further CS synthesis



**Scheme 2.** (a) Ac<sub>2</sub>O, Py 0 °C - RT 24 h, 95 %; (b) p-Thiocresol, BF<sub>3</sub>·OEt<sub>2</sub> DCM, 0 °C - RT 12 h, 72 %; (c) NaOMe MeOH, RT, 2 h, 84%; (d) Ph(OMe)<sub>2</sub> PTSA, ACN, RT 8 h, 69 %; (e) Bu<sub>2</sub>SnO, Bn-Br, CsF, Toluene, 120 °C, 18 h, 65 %; (f) Bz-Cl, DMAP, DCM:Py (4:1) 0 °C - RT 4 h, 81 %; (g) PTSA, DCM:MeOH (2:1) RT, 8 h, 82 %; (h) (ClAcO)<sub>2</sub>O, DCM:Py(4:1) -40 °C, 1 h, 78%; (i) Azidoethanol, NIS, TMSOTf, DCM, 4Å MS -20 °C, 68 %.

#### 4a.6 Synthesis of chondroitin sulfate disaccharide analogs:

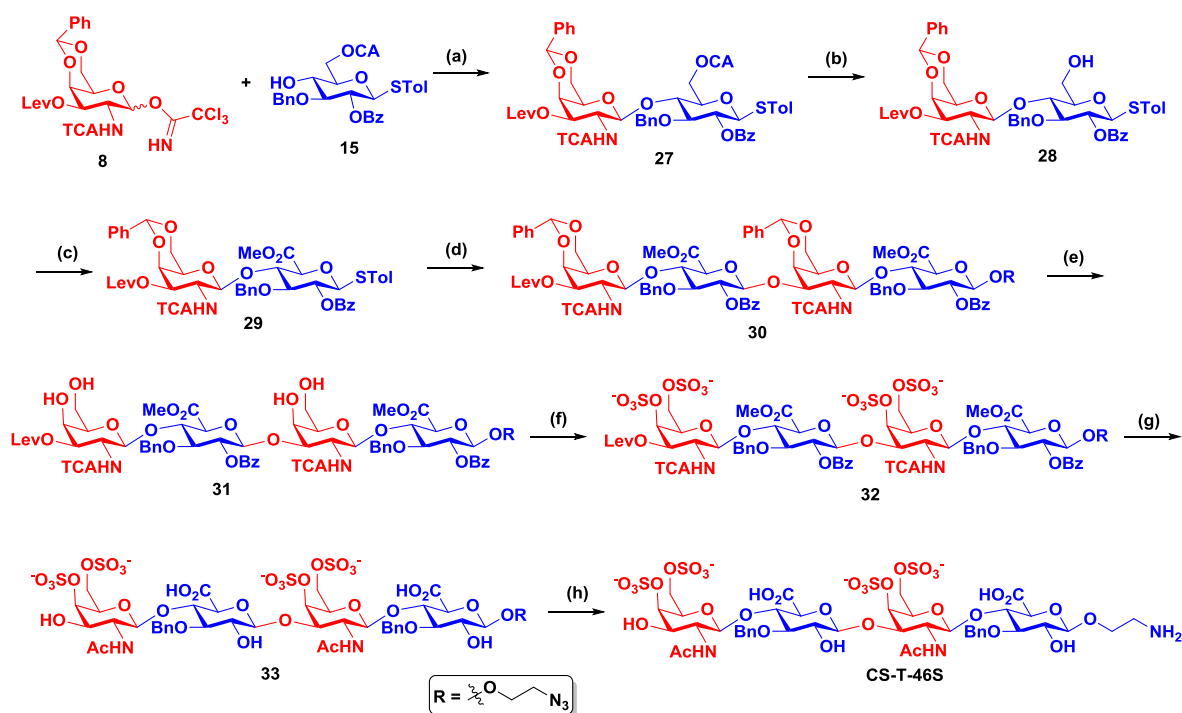
The CS disaccharide analogs have been synthesized from the previously synthesized building blocks **6**, and **18**. As shown in **scheme 3**, we first glycosylate the galactosamine donor **6** with glucose acceptor **18** using NIS and TMSOTf promotor at -40 °C yielded regioselective beta-glycosylated 81% of disaccharide precursor **19**. The compound **19** was subjected to selective chloroacetate deprotection using thiourea to convert glucose to glucuronic acid moiety. Then, oxidation of primary alcohol with a catalytic 2,2,6,6-tetramethyl-1-piperidinyloxy free radical (TEMPO) and [bis(acetoxy)iodo]benzene (BAIB) and methyl esterification yielded CS disaccharide precursor **21**. The compound **21** was subjected to benzylidene deprotection in the presence of PTSA yielded 4- and 6-OH free CS derivative **22**. The stoichiometric mixing of **22** and sulfation reagent sulfur trioxide:triethylamine complex in DMF and heating at 40 °C for several hours resulted in the final desired sulfation patterns (**CS-46S**, **CS-6S** and **CS-0S**). Finally, global deprotection and hydrogenolysis yielded desired disaccharide sulfation patterns. The **CS-3S** was synthesized by selective lev deprotection of **21** in the presence of hydrazine, followed by sulfation, global deprotection and hydrogenolysis.



**Scheme 3:** (a) NIS, TMSOTf, DCM, 4Å MS 0 °C, 81 %; (b) Thiourea, Py:MeOH (1:1) 80 °C, 2 h, 88 %; (c) (1) TEMPO, BAIB, DCM:BuOH:H<sub>2</sub>O, (4: 1:1), RT, 6 h; (2) MeI, K<sub>2</sub>CO<sub>3</sub> DMF, RT, 12 h, 82 %; (d) PTSA, DCM:MeOH (2:1), RT, 6 h, 88 %; (e) SO<sub>3</sub>.TEA, DMF, 60 °C, 72 h, (f) 1. LiOH.H<sub>2</sub>O, THF:H<sub>2</sub>O (2:1), 80 °C, 12 h; 2. Ac<sub>2</sub>O, TEA, MeOH, 0 °C - RT, 12 h; (g) Pd(OH)<sub>2</sub>, H<sub>2</sub>, MeOH, RT, 24 h; (h) AcOH:NH<sub>2</sub>NH<sub>2</sub>.H<sub>2</sub>O (2.5:1), THF:MeOH (10:1), RT, 1 h, 80%.

#### 4a.7 Synthesis of CS. Tetrasaccharide analog:

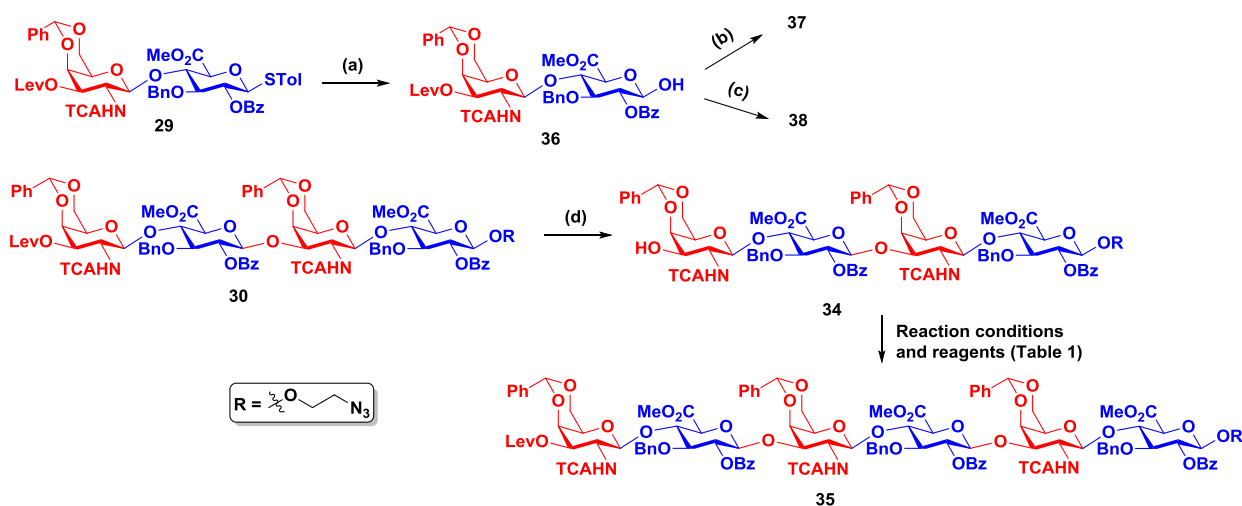
The synthesis of the CS tetrasaccharides has been achieved using **26** acceptor and new disaccharide donor prepared from galactosamine **20a** and glucose **18** building block respectively. The glycosylation between **20a** and **18** was performed by using AgOTf as promotor at room temperature to obtain moderate to good yield of CS disaccharide donor. The lev deprotected CS disaccharide acceptor **26** was glycosylated with **27** was dissolved in DCM solvent and activated by using NIS and TfoH yielded CS tetrasaccharide precursor in a regieselective manner. The selective benzylidene deprotection and sulfation, followed by global deprotection including hydrogenolysis yielded final CS-T-46S compounds in moderate yield. The final compound was characterized by <sup>1</sup>H, <sup>13</sup>C-NMRs and masspectrometry. .



**Scheme 4.** (a) AgOTf, DCM, 4Å MS, RT, 67 %; (b) Thiourea, Py:MeOH (1:1) 80 °C, 2 h, 80 %; (c) 1. TEMPO, BAIB, DCM:H<sub>2</sub>O, ( 2:1 ) RT, 6 h; 2. MeI, K<sub>2</sub>CO<sub>3</sub>, DMF, RT, 12 h, 69 %; (d) **25**, NIS, TMSOTf, DCM, 4Å MS, RT, 57 %; (e) PTSA, DCM:MeOH (2:1) RT, 1 h, 67 %; (f) SO<sub>3</sub>, TEA, DMF, 60 °C, 72 h, 59 %; (g) 1. LiOH.H<sub>2</sub>O, THF:H<sub>2</sub>O (2:1) 80 °C, 12h; 2. Ac<sub>2</sub>O, TEA, MeOH 0 °C - RT, 12 h, (h) Pd(OH)<sub>2</sub>, H<sub>2</sub>, MeOH, RT, 24 h, 37 % (over three steps).

#### 4a.8 Synthesis of CS Hexasaccharide analog:

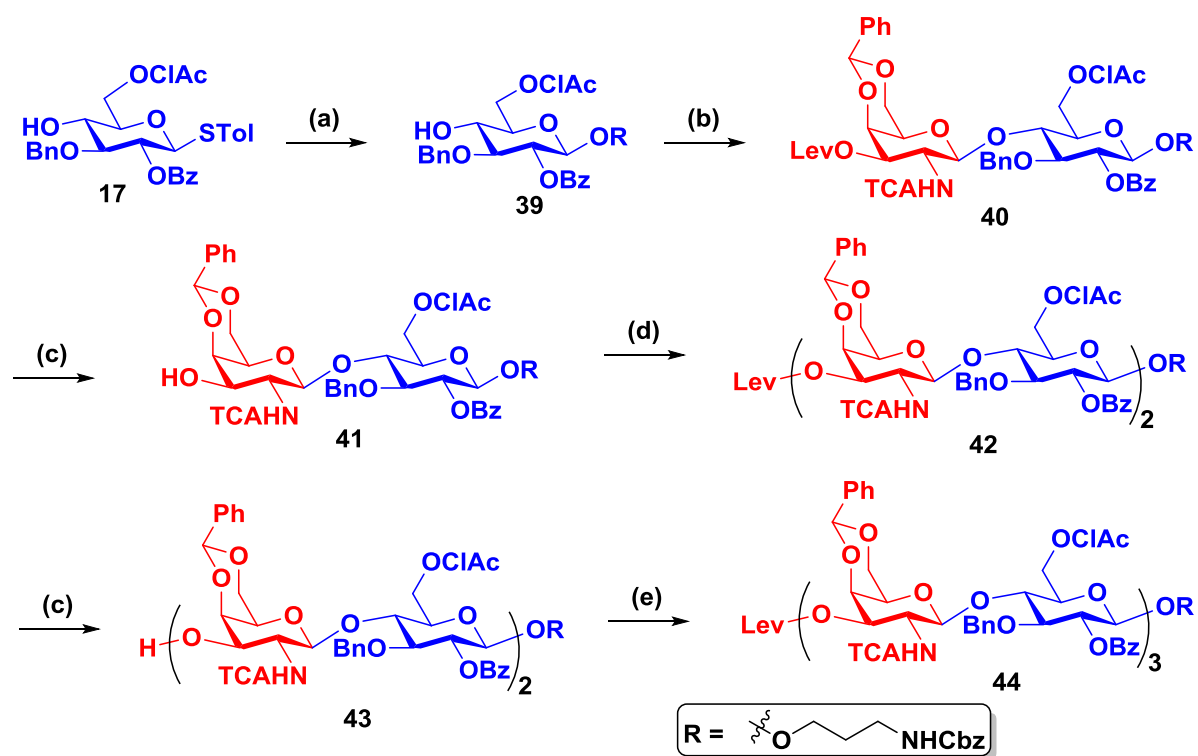
The synthesis of CS hexasaccharide was carried out using various CS disaccharide donors and acceptor. As shown in Table 1, we tried different donors and promotor reagent conditions to synthesize CS hexasaccharide. Unfortunately, we didn't find a suitable condition to get the hexasaccharides. Currently, we are modifying the protecting group strategy to synthesize the hexasaccharide. Alternatively, we have also tried synthesizing oligosaccharides using a non-oxidized glucose building block. As shown in scheme 5, we successfully got the final hexasaccharide. However, oxidation of hexasaccharide didn't go in the presence of TEMPO and BIAB. Currently, we are testing other oxidation reagents to obtain the final compound.



**Scheme 5.** (a) NIS, TfOH, DCM:H<sub>2</sub>O (10:1) RT, 30 min; (b) ClN(Ph)CCF<sub>3</sub>, DCM, CsCO<sub>3</sub>, RT, 12 h, 66 %; (c) Trichloroacetonitrile, DCM, CsCO<sub>3</sub>, RT, 12 h, 69 %; (d) AcOH:NH<sub>2</sub>NH<sub>2</sub>.H<sub>2</sub>O (2.5:1), THF:MeOH (10:1), RT, 1 h, 84 %.

**Table 1.** Reagents and conditions to synthesize CS hexasaccharide.

Sl.No	Donor	Reagent	Yield and remarks
1	<p style="text-align: center;">29</p>	NIS, TMSOTf or TfOH, 4Å MS, DCM -40 °C -RT	Free hydroxyl donor formation and Trace amount of product
2	<p style="text-align: center;">37</p>	TMSOTf, or TfOH or AgOTf, 4Å MS, DCM -78 °C -RT	Free hydroxyl donor formation
3	<p style="text-align: center;">38</p>	TMSOTf, or TfOH or AgOTf, 4Å MS, DCM -20 °C -RT	Free hydroxyl donor formation and Trace amount of product
4	<p style="text-align: center;">27</p>	NIS, TMSOTf or TfOH, 4Å MS, DCM -40 °C -RT	Free hydroxyl donor formation and Trace amount of product



**Scheme 6:** (a) NIS, TMSOTf, DCM, 4Å MS, -40 °C, 78 %; (b) **6**, NIS, TMSOTf, DCM, 4Å MS, -40 °C, 76 %; (c) AcOH:NH<sub>2</sub>NH<sub>2</sub>·H<sub>2</sub>O (2.5:1), THF:MeOH (10:1), RT, 1 h, 74 %; (d) **27**, NIS, TMSOTf, DCM, 4Å MS, RT, 69 %; (e) **27**, NIS, TMSOTf, DCM, 4Å MS, RT, 61 %.

## 4a. 9. Conclusion.

we reported the synthesis of chondroitin sulfate precursor using late oxidized and pre-oxidized of glucose building block. A systematic screening of beta-glycosylation of GalNAc and GluA building blocks using different promoters ensured good-reactivity and high-yield disaccharides building blocks to synthesize CS oligosaccharides using [2+2] strategy. Oxidation, selective deprotection and sulfation of such precursor provided high-yield, reliable and comprehensive strategies to synthesize CS di and tetrasaccharide analogs. However, we failed to obtain hexa-saccharide analog using this strategy. Currently, we are investigating other strategies to synthesize higher oligosaccharides.

## 4a.10 Experimental section:

### 4a.10.1 General information:

All chemicals were reagent grade and used as supplied except where noted. Analytical thin layer chromatography (TLC) was performed on Merck silica gel 60 F254 plates (0.25 mmol). Compounds were visualized by UV irradiation or dipping the plate in CAM/ninhydrin

solution followed by heating. Column chromatography was carried out using force flow of the indicated solvent on a FlukaKieselgel 60 (230–400 mesh). <sup>1</sup>H and <sup>13</sup>C NMR spectra were recorded on a Jeol 400 MHz spectrometer, with a cryoprobe using residual solvents signals as an internal reference (CDCl<sub>3</sub> δ<sub>H</sub>, 7.26 ppm, δ<sub>C</sub> 77.3 ppm, CD<sub>3</sub>OD δ<sub>H</sub> 3.31 ppm, δ<sub>C</sub> 49.0 ppm and D<sub>2</sub>O δ<sub>H</sub> 4.79 ppm). The chemical shifts (δ) are reported in ppm and coupling constants (J) in Hz.

**Synthesis of compound 2:** D-Galactosamine **1** (15 g, 69.76 mmol) was dissolved in MeOH (100 mL), NaOMe (5.6 g, 104.65 mmol) and Chloroacetic anhydride (15.29 mL, 83.72 mmol) were added at 0 °C drop wise. After 1 h the reaction mixture was neutralized with Amberlite® IR-120H resin and filtered, concentrated under reduced pressure. The crude product was (22.5 g, 55.55 mmol) was dissolved in pyridine (80 mL) and acetic anhydride (52.67 mL, 557.27 mmol) added at 0 °C. The reaction mixture allows stirring at room temperature for 24h and concentrated under reduced pressure. The residue was extracted with EtOAc and washed with 2 N HCl (3x). The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated. The crude product was purified by column chromatography (EtOAc/Hexane, 1/2) to afford as mixture of α and β diastereomers (α:β = 1:4) **2** (31.8 g, 91 %) as pale yellow viscous liquid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.67 (d, *J* = 8.7 Hz, 1H), 6.31 (d, *J* = 3.7 Hz, 1H), 5.45 (dd, *J* = 3.2, 1.1 Hz, 1H), 5.33 (dd, *J* = 11.4, 3.2 Hz, 1H), 4.61–4.55 (m, 1H), 4.27 – 4.24 (m, 1H), 4.13 – 4.04 (m, 2H), 2.18 (s, 3H), 2.16 (s, 3H), 2.02 (s, 3H), 2.01 (s, 3H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 171.12, 170.34, 170.04, 168.51, 162.07, 91.93, 90.29, 68.79, 67.57, 66.54, 61.15, 49.35, 20.78, 20.65, 20.64, 20.62. HRMS (ESI) *m/z*: calc'd for [M+H]<sup>+</sup> C<sub>16</sub>H<sub>21</sub>Cl<sub>3</sub>O<sub>10</sub>NS: 492.0231; Found: 492.0235.

**Synthesis of compound 3:** Compound **2** (20 g, 40.73 mmol) was dissolved in anhydrous DCM (100 mL) and *p*-thiocresol (7.58 g, 61.09 mmol) was added under nitrogen atmospheric condition. Borontrifluoride.diethyletherate (15.08 mL, 122.19 mmol) was added drop wise at 0 °C. After string 12 h the reaction mixture was quenched by triethylamine, and extracted with DCM (3x). The combined organic layer washed with NaHCO<sub>3</sub> and brine solution, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/2) to afford **3** (17.2 g, 76 %) as white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.46 – 7.43 (m, 2H), 7.18 – 7.13 (m, 2H), 6.81 (d, *J* = 9.0 Hz, 1H), 5.42 (dd, *J* = 3.3, 1.1 Hz, 1H), 5.31 (dd, *J* = 10.8, 3.3 Hz, 1H), 4.91 (d, *J* = 10.3 Hz, 1H), 4.24 – 4.11 (m, 3H), 3.99 – 3.95 (m, 1H), 2.37 (s, 3H), 2.15 (s, 3H), 2.07 (s, 3H), 2.00 (s, 3H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 170.56, 170.47, 170.22, 161.83, 138.83, 133.46, 129.84, 128.37, 92.35, 87.11, 74.69, 70.69, 66.93, 61.73, 51.41, 21.29, 20.79, 20.74, 20.63. HRMS



(ESI)  $m/z$ : calc'd for  $[M+Na]^+$   $C_{21}H_{24}Cl_3O_8NSNa$ : 578.0186; Found: 578.0182.

**Synthesis of compound 4:** Compound **3** (15.5 g, 27.92 mmol) was dissolved in methanol (150 mL) and NaOMe (4.52 g, 83.78 mmol) was added at RT. After string 2 h the reaction mixture was neutralized with Amberlite® IR-120H resin and filtered, concentrated under reduced pressure. The crude product was purified by column chromatography (MeOH/DCM, 1/8) to afford **4** (8.9 g, 74 %) as white solid.  $^1H$  NMR (400 MHz,  $CD_3OD$ )  $\delta$  7.39 – 7.37 (m, 2H), 7.08 (d,  $J = 8.1$  Hz, 2H), 4.80 (d,  $J = 10.4$  Hz, 1H), 4.03 (t,  $J = 10.3$  Hz, 1H), 3.87 (d,  $J = 3.0$  Hz, 1H), 3.78 – 3.68 (m, 3H), 3.53 – 3.50 (m, 1H), 2.27 (s, 3H).  $^{13}C$  NMR (100 MHz,  $CD_3OD$ )  $\delta$  162.87, 137.36, 131.71, 130.83, 129.25, 87.76, 79.42, 72.05, 68.53, 61.25, 53.45, 19.77. HRMS (ESI)  $m/z$ : calc'd for  $[M+Na]^+$   $C_{15}H_{18}Cl_3O_5NSNa$ : 451.9869; Found: 451.9871.

**Synthesis of compound 5:** Compound **4** (8.3 g, 19.34 mmol) was dissolved in acetonitrile (100 mL) and benzaldehyde dimethoxyacetal (4.35 mL, 29.02 mmol) and *p*-Toluenesulfonic acid (0.73 g, 3.86 mmol) were added at RT. After string 12 h, the reaction mixture was quenched by triethylamine (up to  $P^H \sim 7$ ) and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/1) to afford **5** (7.6 g, 76 %) as white solid.  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  7.56 – 7.53 (m, 2H), 7.43 – 7.34 (m, 5H), 7.10 (d,  $J = 7.9$  Hz, 1H), 6.78 (d,  $J = 7.6$  Hz, 1H), 5.53 (s, 1H), 5.04 (d,  $J = 10.0$  Hz, 1H), 4.39 (dd,  $J = 12.5, 1.5$  Hz, 1H), 4.23 – 4.22 (m, 1H), 4.19 – 4.09 (m, 1H), 4.04 (dd,  $J = 12.5, 1.6$  Hz, 1H), 3.72 – 3.65 (m, 1H), 3.60 – 3.58 (m, 1H), 2.64 (d,  $J = 10.4$  Hz, 1H), 2.35 (s, 3H).  $^{13}C$  NMR (100 MHz,  $CDCl_3$ )  $\delta$  161.95, 138.86, 137.45, 134.48, 129.85, 129.43, 128.24, 126.67, 126.55, 101.32, 92.46, 83.66, 74.98, 70.44, 70.09, 69.28, 54.15, 21.31. HRMS (ESI)  $m/z$ : calc'd for  $[M+H]^+$   $C_{22}H_{23}Cl_3O_5NS$ : 518.0363; Found: 518.0358.

**Synthesis of compound 6:** Compound **5** (9.5 g, 18.37 mmol) was dissolved in anhydrous DCM (80 mL). The reaction mixture was cooled to 0 °C and DCC (6.82 mL, 45.45 mmol) and DMAP (1.93 g, 11.36 mmol) and Levulinic acid (2.25 mL, 22.05 mmol) were added at the same temperature. After string 1 h, the reaction mixture was filtered on celite bed and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/1) to afford **6** (9.1 g, 80 %) as white solid.  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  7.56 – 7.53 (m, 2H), 7.43 – 7.34 (m, 5H), 7.10 (d,  $J = 7.9$  Hz, 1H), 6.78 (d,  $J = 7.6$  Hz, 1H), 5.53 (s, 1H), 5.04 (d,  $J = 10.0$  Hz, 1H), 4.39 (dd,  $J = 12.5, 1.5$  Hz, 1H), 4.23 – 4.22 (m, 1H), 4.19 – 4.09 (m, 1H), 4.04 (dd,  $J = 12.5, 1.6$  Hz, 1H), 3.72 – 3.65 (m, 1H), 3.60 – 3.58 (m, 1H), 2.64 (d,  $J = 10.4$  Hz, 1H), 2.35 (s, 3H).  $^{13}C$  NMR (100 MHz,  $CDCl_3$ )  $\delta$

161.95, 138.86, 137.45, 134.48, 129.85, 129.43, 128.24, 126.67, 126.55, 101.32, 92.46, 83.66, 74.98, 70.44, 70.09, 69.28, 54.15, 21.31. HRMS (ESI)  $m/z$ : calc'd for  $[M+H]^+$   $C_{27}H_{29}Cl_3O_7NS$ : 616.0730; Found: 616.0732.

**Synthesis of compound 7:** Compound **6** (5 g, 8.13 mmol) was dissolved in mixture of DCM/H<sub>2</sub>O [(15/1), (v/v), 70 mL] and NIS (2.19 g, 9.75 mmol) and TfOH (0.14 mL, 1.62 mmol) were added at RT. After 15 min the reaction mixture was quenched with triethylamine and washed with sodium thiosulfate solution (2x). The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/1) to afford **7** (3.5 g, 84 %) as white foam. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  7.47 – 7.44 (m, 2H), 7.33 – 7.28 (m, 3H), 6.94 (d,  $J$  = 9.3 Hz, 1H), 5.47 (s, 1H), 5.39 (t,  $J$  = 3.6 Hz, 1H), 5.27 (dd,  $J$  = 11.2, 3.3 Hz, 1H), 4.57 (td,  $J$  = 11.2, 10.5, 3.4 Hz, 1H), 4.20 – 4.14 (m, 2H), 4.02 – 3.96 (m, 2H), 3.87 (s, 2H), 2.73 – 2.59 (m, 2H), 2.56 – 2.47 (m, 2H), 2.25 (s, 1H), 2.00 (s, 3H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  206.83, 172.83, 162.11, 137.39, 129.79, 129.18, 128.54, 128.26, 126.23, 100.75, 92.32, 91.85, 73.50, 69.25, 68.84, 62.54, 49.91, 37.75, 29.75, 29.72, 28.21. HRMS (ESI)  $m/z$ : calc'd for  $[M+Na]^+$   $C_{20}H_{22}Cl_3O_8NSNa$ : 532.0309; Found: 532.0311.

**Synthesis of compound 8:** Compound **7** (3.5 g, 6.87 mmol) was dissolved in anhydrous DCM (70 mL) and Cesium carbonate (0.89 g, 2.75 mmol) and trichloroacetonitrile (5.51 mL, 55 mmol) were added at RT. After 12 h the reaction mixture was concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/3) to afford **8** (2.9 g, 65 %) as white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  8.77 (s, 1H), 7.54 – 7.52 (m, 2H), 7.41 – 7.35 (m, 3H), 6.95 (d,  $J$  = 8.6 Hz, 1H), 6.62 (d,  $J$  = 3.3 Hz, 1H), 5.58 (s, 1H), 5.45 (dd,  $J$  = 11.4, 3.2 Hz, 1H), 4.87 (ddd,  $J$  = 11.7, 8.7, 3.4 Hz, 1H), 4.44 (d,  $J$  = 3.0 Hz, 1H), 4.33 (dd,  $J$  = 12.7, 1.4 Hz, 1H), 4.07 (dd,  $J$  = 12.7, 1.4 Hz, 1H), 3.94 (s, 1H), 2.73 – 2.65 (m, 2H), 2.62 – 2.57 (m, 2H), 2.08 (s, 3H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  206.30, 173.38, 162.04, 160.13, 137.16, 129.29, 128.30, 126.33, 100.98, 95.32, 92.03, 90.81, 73.08, 68.81, 68.42, 65.16, 49.76, 37.72, 29.70, 28.10. HRMS (ESI)  $m/z$ : calc'd for  $[M+Na]^+$   $C_{22}H_{22}Cl_6O_8N_2SNa$ : 674.9405; Found: 674.9409.

**Synthesis of compound 10:** D-glucose **9** (50 g, 277.7 mmol) was dissolved in pyridine (200 mL) and cooled the reaction at 0 °C. Ac<sub>2</sub>O (262 mL, 2777.7 mmol) was added to the reaction mixture dropwise and slowly allowed to room temperature. After 24 h stirring, the solvent concentrated. The mixture was diluted with EtOAc and washed with 1 N HCl solution (3x).

The organic layer dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/3) to afford **10** (103.4 g, 95 %) as a white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.34 (d, *J* = 3.7 Hz, 1H), 5.51 – 5.46 (m, 1H), 5.20 – 5.07 (m, 2H), 4.33 – 4.23 (m, 1H), 4.18 – 4.06 (m, 2H), 2.19 (s, 3H), 2.10 (s, 3H), 2.05 (s, 3H), 2.04 (s, 3H), 2.03 (s, 3H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 170.61, 170.61, 170.21, 169.64, 169.38, 168.74, 89.05, 69.81, 69.18, 67.88, 61.45, 20.87, 20.69, 20.66, 20.56, 20.44. HRMS (ESI) *m/z*: calc'd for [M+Na]<sup>+</sup> C<sub>16</sub>H<sub>22</sub>O<sub>11</sub>Na: 413.1060; Found: 413.1059.

**Synthesis of compound 11:** Compound **10** (30 g, 76.9 mmol) was dissolved in anhydrous DCM (150 mL) and *p*-thiocresol (11.4 g, 92.3 mmol) was added then cooled the reaction at 0 °C. BF<sub>3</sub>.OEt (28.4 mL, 230 mmol) was added to the reaction mixture dropwise and allowed to room temperature. After 12 h stirring, the reaction mixture was washed with aq. NaHCO<sub>3</sub> solution (2x). The combined organic layer dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/3) to afford **11** (25.3 g, 72 %) as a white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.38 (dd, *J* = 8.2, 2.0 Hz, 2H), 7.11 (d, *J* = 8.0 Hz, 2H), 5.20 (t, *J* = 9.4 Hz, 1H), 5.01 (t, *J* = 9.8 Hz, 1H), 4.92 (t, *J* = 9.7 Hz, 1H), 4.63 (d, *J* = 10.1 Hz, 1H), 4.24 – 4.13 (m, 2H), 3.69 (ddt, *J* = 9.0, 4.2, 2.0 Hz, 1H), 2.34 (s, 3H), 2.08 (s, 3H), 2.07 (s, 3H), 2.00 (s, 3H), 1.97 (zs, 3H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 170.51, 170.50, 170.13, 170.11, 169.35, 169.19, 138.74, 133.80, 129.65, 127.51, 85.74, 75.71, 73.99, 69.89, 68.17, 62.09, 21.17, 20.74, 20.71, 20.57, 20.55. HRMS (ESI) *m/z*: calc'd for [M+Na]<sup>+</sup> C<sub>21</sub>H<sub>26</sub>O<sub>9</sub>SNa: 413.1195; Found: 477.1195.

**Synthesis of compound 12:** Compound **11** (26 g, 57.2 mmol) was dissolved in MeOH (250 mL) and NaOMe (12.3 g, 229.0 mmol) was added at room temperature. After 2 h stirring, the reaction mixture was quenched with amberlite<sup>TM</sup> IR-120(H) resin. The resin was filtered and the solvent was concentrated under reduced pressure. The crude product was purified by column chromatography (MeOH/DCM, 1/9) to afford **12** (13.9 g, 84 %) as a white solid. <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 7.53 – 7.42 (m, 2H), 7.20 – 7.07 (m, 2H), 4.53 (d, *J* = 9.7 Hz, 1H), 3.87 (dd, *J* = 12.1, 1.7 Hz, 1H), 3.72 – 3.64 (m, 1H), 3.38 (d, *J* = 9.3 Hz, 1H), 3.30 (ddd, *J* = 5.5, 3.9, 2.1 Hz, 2H), 3.20 (dd, *J* = 9.7, 8.7 Hz, 1H), 2.33 (s, 3H). <sup>13</sup>C NMR (100 MHz, CD<sub>3</sub>OD) δ 141.30, 136.05, 133.71, 133.08, 92.18, 84.54, 82.19, 76.21, 73.89, 65.43, 23.66. HRMS (ESI) *m/z*: calc'd for [M+Na]<sup>+</sup> C<sub>13</sub>H<sub>18</sub>O<sub>5</sub>SNa: 309.0773; Found: 309.0777.

**Synthesis of compound 13:** Compound **12** (20 g, 69.9 mmol) was dissolved in ACN (150

mL) and benzylidene dimethyl acetal (15.7 mL, 104.8 mmol) and PTSA (2.6 g, 13.9 mmol) were added at room temperature. After 8 h stirring, the reaction mixture was quenched with Et<sub>3</sub> and washed with aq. NaHCO<sub>3</sub> solution (2x). The combined organic layer dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (MeOH/DCM, 1/9.5) to afford **13** (18.2 g, 69 %) as a white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.53 – 7.41 (m, 4H), 7.42 – 7.34 (m, 3H), 7.20 – 7.13 (m, 2H), 5.52 (s, 1H), 4.55 (d, *J* = 9.7 Hz, 1H), 4.37 (dd, *J* = 10.5, 4.3 Hz, 1H), 3.86 – 3.69 (m, 2H), 3.55 – 3.37 (m, 3H), 3.21 (d, *J* = 2.3 Hz, 1H), 2.97 (d, *J* = 2.6 Hz, 1H), 2.38 (s, 3H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 138.82, 136.90, 133.65, 129.90, 129.34, 128.38, 127.36, 126.33, 101.91, 88.68, 80.22, 74.51, 72.51, 70.49, 68.58, 21.21. HRMS (ESI) *m/z*: calc'd for [M+Na]<sup>+</sup> C<sub>20</sub>H<sub>22</sub>O<sub>5</sub>SNa: 397.1086; Found: 397.1085.

**Synthesis of compound 14:** Compound **13** (15 g, 40.1 mmol) was dissolved in toluene (150 mL) and dibutyltin oxide (10.9 g, 44.1 mmol) was added at room temperature then reflux at 120 °C. After 6 h stirring, the reaction mixture was cooled to ambient temperature and concentrated under reduced pressure. The crude product was kept in high vacuum for 2h and dissolved in DMF:Toluene (v/v 1/1, 200ml ). Benzyl bromide (5.2 mL, 44.1 mmol) and CsF (6.7 g, 44.1 mmol) were added at room temperature then reflux at 120 °C. After 12 h reflux the reaction mixture was cooled to ambient temperature. The solvents were evaporated then diluted with EtOAc and washed with brine solution (3x). The organic layer dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/5) to afford **14** (12.1 g, 65 %) as a white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.52 – 7.41 (m, 4H), 7.41 – 7.26 (m, 8H), 7.13 (d, *J* = 8.0 Hz, 2H), 5.56 (s, 1H), 4.94 (d, *J* = 11.5 Hz, 1H), 4.79 (d, *J* = 11.5 Hz, 1H), 4.56 (d, *J* = 9.6 Hz, 1H), 4.38 (dd, *J* = 10.5, 4.9 Hz, 1H), 3.78 (t, *J* = 10.3 Hz, 1H), 3.73 – 3.58 (m, 2H), 3.49 (td, *J* = 9.7, 4.8 Hz, 2H), 2.62 (bs, 1H), 2.34 (s, 3H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 138.86, 138.29, 137.29, 133.93, 129.94, 129.13, 128.58, 128.38, 128.24, 128.00, 127.30, 126.11, 101.32, 88.66, 81.71, 81.22, 74.93, 72.20, 70.80, 68.73, 21.30. HRMS (ESI) *m/z*: calc'd for [M+H]<sup>+</sup> C<sub>27</sub>H<sub>29</sub>O<sub>5</sub>S: 465.1736; Found: 465.1735.

**Synthesis of compound 15:** Compound **14** (10 g, 21.5 mmol) was dissolved in mixture of DCM:Pyridine (4/2, v/v, 100 mL) and cooled the reaction mixture at 0 °C. Dimethyl aminopyridine (DMAP) (0.52 g, 4.3 mmol) and benzoyl chloride (7.5 mL, 64.6 mmol) were added at the same room temperature. After 4 h stirring, the reaction mixture was washed with

aq. NaHCO<sub>3</sub> solution (2x) and 1 N HCl solution (3x). The combined organic layer dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/5) to afford **15** (9.9 g, 81 %) as a white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.07 – 7.97 (m, 2H), 7.66 – 7.55 (m, 1H), 7.54 – 7.42 (m, 4H), 7.45 – 7.34 (m, 5H), 7.15 – 7.02 (m, 7H), 5.60 (s, 1H), 5.26 (dd, *J* = 9.9, 8.7 Hz, 1H), 4.82 – 4.75 (m, 2H), 4.66 (d, *J* = 11.9 Hz, 1H), 4.41 (dd, *J* = 10.5, 5.0 Hz, 1H), 3.93 – 3.73 (m, 3H), 3.55 (td, *J* = 9.7, 5.0 Hz, 1H), 2.32 (s, 1H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 165.13, 138.65, 137.77, 137.25, 133.77, 133.32, 130.03, 129.91, 129.77, 129.16, 128.50, 128.40, 128.26, 128.18, 127.68, 126.10, 101.34, 87.26, 81.51, 79.42, 74.35, 72.08, 70.65, 68.70, 21.28. HRMS (ESI) *m/z*: calc'd for [M+H]<sup>+</sup> C<sub>34</sub>H<sub>33</sub>O<sub>6</sub>S: 569.1998; Found: 569.1998.

**Synthesis of compound 16:** Compound **15** (9.7 g, 17.0 mmol) was dissolved in mixture of DCM:MeOH (2/1, v/v, 90 mL) and PTSA (4.8 g, 25.6 mmol) was added at room temperature. After 8 h stirring, the reaction mixture was quenched with Et<sub>3</sub>N then washed with aq. NaHCO<sub>3</sub> solution (2x). The combined organic layer dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/3) to afford **16** (6.7 g, 82 %) as a white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.09 – 8.04 (m, 2H), 7.62 – 7.58 (m, 1H), 7.50 – 7.44 (m, 2H), 7.34 – 7.29 (m, 2H), 7.21 – 7.06 (m, 5H), 7.07 (d, *J* = 7.9 Hz, 2H), 5.25 – 5.19 (m, 1H), 4.80 – 4.52 (m, 3H), 3.92 (dd, *J* = 11.9, 3.3 Hz, 1H), 3.79 (dd, *J* = 12.0, 5.0 Hz, 1H), 3.74 – 3.66 (m, 2H), 3.46 – 3.42 (m, 1H), 2.80 (bs, 0H), 2.30 (s, 3H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 165.24, 138.39, 137.69, 133.36, 133.14, 129.90, 129.76, 129.73, 128.68, 128.52, 128.05, 127.99, 86.65, 83.87, 79.44, 74.81, 72.40, 70.31, 62.56, 21.15. HRMS (ESI) *m/z*: calc'd for [M+Na]<sup>+</sup> C<sub>27</sub>H<sub>28</sub>O<sub>6</sub>SNa: 503.1504; Found: 503.1501.

**Synthesis of compound 17:** Compound **16** (5.6 g, 11.6 mmol) was dissolved in anhydrous DCM (50 mL) and cooled the reaction mixture at -40 °C. Chloroacetic anhydride (1.79 g, 10.5 mmol) was added at the same room temperature. After 1 h stirring, the reaction mixture was allowed to ambient temperature and washed with 1 N HCl solution (3x). The combined organic layer dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/5) to afford **17** (5.1 g, 78 %) as a white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.15 – 8.07 (m, 2H), 7.68 – 7.60 (m, 1H), 7.56 – 7.46 (m, 2H), 7.41 – 7.33 (m, 2H), 7.28 – 7.16 (m, 5H), 7.16 – 7.08 (m, 2H), 5.24 (dd, *J* = 10.0, 8.9 Hz, 1H), 4.79 – 4.58 (m, 3H), 4.57 (dd, *J* = 11.9, 1.7 Hz, 1H), 4.54 – 4.43 (m,

1H), 4.13 – 4.08 (m, 10H), 3.76 – 3.65 (m, 1H), 3.69 – 3.58 (m, 2H), 2.62 (d,  $J = 2.2$  Hz, 1H), 2.35 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  167.45, 165.18, 138.45, 137.52, 133.51, 133.41, 129.89, 129.72, 129.58, 128.59, 128.55, 128.43, 128.11, 128.10, 86.53, 83.73, 74.97, 72.15, 69.77, 64.78, 40.74, 21.17. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{Na}]^+$   $\text{C}_{29}\text{H}_{29}\text{O}_7\text{SNa}$ : 579.1220; Found: 579.1218.

**Synthesis of compound 18:** Compound **17** (1.5 g, 2.69 mmol) and azidoethanol (0.35 g, 4.04 mmol) were dissolved in anhydrous DCM (20 mL) with 4Å molecular sieves and stirred at room temperature for 2 h. The reaction mixture cooled at  $-40$  °C and NIS (0.72 g, 3.23 mmol) and TMSOTf (97  $\mu\text{L}$ , 0.53 mmol) were added at the same room temperature. After 30 min, the reaction mixture was quenched  $\text{Et}_3\text{N}$  then allowed to ambient temperature and the molecular sieves was filtered on celite bed. The reaction mixture was washed with aq.  $\text{Na}_2\text{S}_2\text{O}_3$  solution (2x). The combined organic layer dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/5) to afford **18** (0.8 g, 57 %) as a white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.08 – 8.04 (m, 2H), 7.62 – 7.56 (m, 1H), 7.48 – 7.44 (m, 1H), 7.24 – 7.18 (m, 5H), 5.29 – 5.25 (m, 1H), 4.78 – 4.59 (m, 3H), 4.56 – 4.44 (m, 2H), 4.12 (s, 2H), 3.99 – 3.94 (m, 1H), 3.72 – 3.57 (m, 4H), 3.42 – 3.36 (m, 1H), 3.30 – 3.25 (m, 1H), 2.64 (bs, 1H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  167.59, 165.17, 137.59, 133.29, 129.80, 129.71, 128.59, 128.46, 128.09, 101.21, 82.12, 74.75, 73.51, 73.20, 69.79, 68.10, 64.49, 50.67, 40.76. IR ( $\text{cm}^{-1}$ ,  $\text{CH}_2\text{Cl}_2$ ) 2882, 2103, 1723, 1601, 1451. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{Na}]^+$   $\text{C}_{24}\text{H}_{26}\text{O}_8\text{ClN}_3\text{SNa}$ : 542.1306; Found: 542.1310.

**Synthesis of compound 19:** Compound (Glycosyl donor) **6** (1 g, 1.62 mmol) and (glycosyl acceptor) **18** (0.75 g, 1.46 mmol) were dissolved in anhydrous DCM (20 mL) with freshly dried 4 Å MS and stirred at RT for 1 h. Cooled the reaction mixture at  $-40$  °C and NIS (0.43 g, 1.95 mmol) and TMSOTf (58  $\mu\text{L}$ , 0.32 mmol) were added at the same temperature. After consumption of donor, the reaction mixture was quenched with  $\text{Et}_3\text{N}$  and filtered on celite bed and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/1) to afford **19** (1.2 g, 81 %) as white foam.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.00 – 7.92 (m, 2H), 7.59 – 7.53 (m, 1H), 7.48 – 7.39 (m, 4H), 7.31 – 7.29 (m, 1H), 7.17 – 7.11 (m, 3H), 7.08 – 6.97 (m, 3H), 5.50 (s, 1H), 5.29 – 5.23 (m, 1H), 5.19 – 5.15 (m, 1H), 5.00 (d,  $J = 11.8$  Hz, 1H), 4.84 (d,  $J = 8.2$  Hz, 1H), 4.75 (d,  $J = 11.8$  Hz, 1H), 4.62 (d,  $J = 7.7$  Hz, 1H), 4.57 – 4.45 (m, 2H), 4.43 – .35 (m, 1H), 4.28 – 4.19 (m, 2H), 4.13 (d,  $J = 2.1$  Hz, 2H), 4.02 (dd,  $J = 9.6, 8.2$  Hz, 1H), 3.96 – 3.89 (m, 3H), 3.69 – 3.60 (m,

2H), 3.39 – 3.32 (m, 2H), 3.29v – 3.23 (m, 1H), 2.73 – 2.66 (m, 2H), 2.58 – 2.54 (m, 1H), 2.07 (s, 3H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 206.51, 172.60, 167.45, 165.14, 162.18, 138.14, 137.54, 133.12, 129.83, 129.74, 129.08, 128.45, 128.33, 128.23, 128.15, 128.04, 127.99, 127.93, 127.83, 127.27, 126.36, 126.33, 126.25, 100.93, 100.90, 100.60, 92.46, 80.51, 77.26, 74.64, 72.85, 72.82, 72.79, 70.38, 68.70, 68.12, 66.45, 63.98, 52.99, 50.61, 40.84, 37.76, 29.71, 28.11. IR (cm<sup>-1</sup>, CH<sub>2</sub>Cl<sub>2</sub>) 3341, 2917, 2104, 1716, 1524. HRMS (ESI) m/z: calc'd for [M+H]<sup>+</sup> C<sub>44</sub>H<sub>47</sub>O<sub>15</sub>Cl<sub>4</sub>N<sub>4</sub>: 1011.1792; Found: 1011.1787.

**Synthesis of compound 20:** Compound **19** (0.85 g, 0.84 mmol) was dissolved in mixture of Pyridine:MeOH [(v/v), (1:1), 20 mL] and thiourea (0.12 g, 1.68 mmol) was added. The reaction mixture was reflux at 80 °C for 2 h. The mixture was allow to ambient temperature and diluated with EtOAc and washed with 1N HCl solution (3x) and brine solution (2x). The combined organic layer dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/0.5) to afford **20** (0.69 g, 88 %) as a white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.01 – 7.94 (m, 2H), 7.60 – 7.54 (m, 1H), 7.48 – 7.40 (m, 5H), 7.33 – 7.26 (m, 3H), 7.17 – 7.12 (m, 2H), 7.05 – 7.00 (m, 1H), 6.93 (dd, *J* = 8.2, 6.8 Hz, 2H), 5.44 (s, 1H), 5.25 – 5.17 (m, 2H), 5.09 (d, *J* = 11.8 Hz, 1H), 4.95 (d, *J* = 8.4 Hz, 1H), 4.74 (d, *J* = 11.8 Hz, 1H), 4.60 (d, *J* = 7.9 Hz, 1H), 4.48 (dt, *J* = 11.3, 8.8 Hz, 1H), 4.28 – 4.10 (m, 3H), 3.97 – 3.81 (m, 5H), 3.63 (ddd, *J* = 10.8, 6.9, 3.9 Hz, 1H), 3.50 – 3.26 (m, 4H), 3.12 (s, 1H), 2.77 – 2.52 (m, 4H), 2.04 (s, 3H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 206.46, 172.64, 165.44, 162.19, 138.38, 137.74, 133.13, 129.89, 129.80, 128.99, 128.33, 128.30, 128.10, 127.97, 127.16, 126.39, 101.11, 100.85, 100.74, 92.72, 80.32, 76.39, 75.22, 75.05, 73.03, 72.83, 70.84, 68.80, 68.07, 66.26, 60.73, 52.83, 50.71, 37.78, 29.70, 28.26. IR (cm<sup>-1</sup>, CH<sub>2</sub>Cl<sub>2</sub>) 3331, 2924, 2104, 1711, 1527. HRMS (ESI) m/z: calc'd for [M+H]<sup>+</sup> C<sub>42</sub>H<sub>45</sub>O<sub>14</sub>Cl<sub>3</sub>N<sub>4</sub>: 957.1896; Found: 957.1899.

**Synthesis of compound 21:** Compound **20** (0.65 g, 0.69 mmol) is dissolved in mixture of DCM:<sup>t</sup>BuOH:H<sub>2</sub>O [(4/1/1), (v/v/v), 20 mL] were added 2,2,6,6-tetramethyl-1-piperidinyloxy free radical (TEMPO, 89.9 mg, 0.17 mmol) and bis(acetoxy)iodobenzene (BAIB, 0.31 g, 1.39 mmol) at room temperature. After 6 h, the mixture was quenched with saturated NH<sub>4</sub>Cl solution and extracted with DCM. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated under reduced pressure. The residue was taken as such for next step without further purification. The crude was dissolved in DMF (10 mL) and K<sub>2</sub>CO<sub>3</sub> (0.28 g, 2.08 mmol) methyl iodide (0.38 mL, 6.25 mmol) were added at the room temperature. After string 12 h, the reaction mixture diluted with ethyl acetate and washed with brine solution (3x), the

combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/1) to afford **21** (0.55 g, 82 %) as white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.97 – 7.92 (m, 2H), 7.59 – 7.54 (m, 1H), 7.50 – 7.39 (m, 4H), 7.30 – 7.28 (m, 2H), 7.20 – 7.15 (m, 2H), 7.07 – 6.93 (m, 4H), 5.52 (s, 1H), 5.23 (dd, *J* = 8.8, 6.8 Hz, 1H), 5.15 – 5.00 (m, 3H), 4.76 – 4.61 (m, 2H), 4.48 (dt, *J* = 11.2, 8.7 Hz, 1H), 4.35 – 4.18 (m, 3H), 4.08 – 3.86 (m, 4H), 3.84 (s, 3H), 3.67 (ddd, *J* = 11.0, 7.0, 4.2 Hz, 1H), 3.51 (q, *J* = 1.5 Hz, 1H), 3.42 – 3.26 (m, 2H), 2.83 – 2.55 (m, 4H), 2.05 (s, 3H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 206.43, 172.47, 170.06, 165.18, 161.97, 138.02, 137.68, 133.14, 129.83, 129.65, 129.02, 128.36, 128.33, 128.12, 128.00, 127.32, 126.41, 101.34, 100.95, 100.74, 92.64, 79.74, 78.56, 75.44, 74.17, 73.16, 72.96, 71.54, 68.66, 68.38, 66.70, 53.11, 52.40, 50.56, 37.73, 29.70, 28.20. IR (cm<sup>-1</sup>, CH<sub>2</sub>Cl<sub>2</sub>) 2916, 2104, 1715, 1525. HRMS (ESI) *m/z*: calc'd for [M+Na]<sup>+</sup> C<sub>43</sub>H<sub>45</sub>O<sub>15</sub>Cl<sub>3</sub>N<sub>4</sub>Na: 985.1845; Found: 985.1841.

**Synthesis of compound 25:** Compound **21** (0.1 g, 0.1 mmol) was dissolved in mixture of THF/MeOH [(v/v), (10/1), xx mL]. Hydrazine hydrate/acetic acid [(v/v), (2.5/1)] were added in mixture THF:MeOH (5:1) [(v/v), (10/1)], the resulting solution was added to reaction mixture at room temperature. After 1 h the reaction mixture was diluted with EtOAc and washed with aq. NaHCO<sub>3</sub> solution (3x) and brine solution (2x). The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/0.7) to afford **25** (72 mg, 80 %) as white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.99 – 7.91 (m, 2H), 7.61 – 7.50 (m, 1H), 7.50 – 7.36 (m, 5H), 7.34 – 7.24 (m, 3H), 7.19 (d, *J* = 6.6 Hz, 2H), 7.07 – 6.95 (m, 2H), 5.55 (s, 1H), 5.23 (dd, *J* = 8.8, 7.0 Hz, 1H), 5.09 (d, *J* = 11.4 Hz, 1H), 4.90 (d, *J* = 8.4 Hz, 1H), 4.72 – 4.61 (m, 2H), 4.28 – 4.15 (m, 4H), 4.08 – 3.90 (m, 4H), 3.89 – 3.79 (m, 4H), 3.80 – 3.74 (m, 1H), 3.74 – 3.61 (m, 1H), 3.48 (s, 1H), 3.38 – 3.27 (m, 2H), 2.77 (d, *J* = 10.6 Hz, 1H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 170.60, 165.30, 163.22, 138.05, 137.65, 133.30, 129.99, 129.91, 129.68, 129.29, 128.50, 128.46, 128.36, 128.32, 128.13, 127.50, 126.43, 126.37, 101.42, 101.30, 100.80, 92.81, 79.99, 79.02, 75.78, 74.91, 74.18, 73.25, 72.46, 68.77, 68.59, 66.98, 55.64, 53.28, 50.64. HRMS (ESI) *m/z*: calc'd for [M+H]<sup>+</sup> C<sub>38</sub>H<sub>40</sub>O<sub>13</sub>Cl<sub>3</sub>N<sub>4</sub>: 865.1657; Found: 865.1651.

**Synthesis of compound 22:** Compound **21** (0.5 g, 0.51 mmol) was dissolved in mixture of DCM/MeOH [(2/1), (v/v), 12 mL] and p-Toluenesulfonic acid (PTSA, 0.14 g, 0.77 mmol)



was added at RT. After 6 h stirring, the mixture was quenched by triethylamine (up to pH ~ 7) and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/0.5) to afford **22** (0.4 g, 88 %) as white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.04 – 7.92 (m, 2H), 7.62 – 7.54 (m, 1H), 7.50 – 7.40 (m, 2H), 7.27 – 7.10 (m, 5H), 7.03 (d, *J* = 9.0 Hz, 1H), 5.27 – 5.21 (m, 1H), 4.98 – 4.88 (m, 2H), 4.75 – 4.65 (m, 2H), 4.42 (dt, *J* = 11.0, 8.7 Hz, 1H), 4.28 (t, *J* = 8.9 Hz, 1H), 4.12 – 4.04 (m, 2H), 3.99 (ddt, *J* = 10.7, 5.6, 3.5 Hz, 1H), 3.88 (d, *J* = 3.5 Hz, 1H), 3.85 (d, *J* = 3.5 Hz, 3H), 3.74 – 3.58 (m, 3H), 3.58 – 3.50 (m, 1H), 3.37 (qdd, *J* = 13.3, 6.3, 3.9 Hz, 2H), 3.14 (s, 1H), 2.85 – 2.71 (m, 2H), 2.67 – 2.52 (m, 2H), 2.22 (s, 3H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 207.86, 172.25, 169.72, 165.25, 162.17, 137.70, 133.26, 129.79, 129.55, 128.50, 128.41, 128.27, 128.24, 128.00, 127.95, 127.89, 127.76, 101.27, 100.58, 92.65, 79.66, 77.50, 75.42, 74.69, 74.12, 73.79, 72.94, 68.39, 67.12, 62.23, 60.42, 53.10, 52.43, 50.54, 38.04, 29.84, 28.14, 21.07, 14.20. IR (cm<sup>-1</sup>, CH<sub>2</sub>Cl<sub>2</sub>) 3337, 2925, 2104, 1714, 1526. HRMS (ESI) *m/z*: calc'd for [M+H]<sup>+</sup> C<sub>36</sub>H<sub>41</sub>O<sub>15</sub>Cl<sub>3</sub>N<sub>4</sub>: 897.1532; Found: 897.1533.

#### 4a.10.2 General Procedure for sulfation reaction:

The respective compound (1 eq) was dissolved in DMF (2 mL) and added SO<sub>3</sub>.TEA complex (10 eq. per OH group). The reaction mixture was refluxed at 60 °C for 48 h. The solvent was evaporated and purified by column chromatography and lyophilized to afford corresponding sulphated compound.

#### 4a.10.3 General procedure for hydrolysis and hydrogenolysis reaction:

The respective compound (1 eq.) and lithium hydroxide (5 eq per ester) were dissolved in the mixture of THF and water [2/1 (v/v), 6 mL] and refluxed at 80 °C for 12 h. After the reaction completion, the mixture was cooled to room temperature and neutralized with amberlite® IR 120 resin, filtered, and concentrated under reduced pressure. The residue was purified by bond elute using water then lyophilized to get corresponding hydrolysis compounds which was proceeded for N-acetylation reaction. The crude product was dissolved in MeOH (3 mL) and cooled at 0 °C. Triethylamine (10 eq per amine) and Ac<sub>2</sub>O (10 eq per amine) were added. After 12 h, solvent was concentrated and purified by bond elute using water then lyophilized to get corresponding compounds which also proceeded for the hydrogenolysis, using Pd(OH)<sub>2</sub> on charcoal with hydrogen gas in H<sub>2</sub>O (3 mL). After 12 - 24 h (depend on compounds), the mixture was filtered on Whatman 42 filter paper. The residue was purified by bond elute using water then lyophilized to get corresponding compounds.

**Synthesis of compound 26:** Compound **25** (40 mg, 0.04 mmol) followed the general procedure for sulfation reaction to afford **26** (32 mg, 73%) as a white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.99 – 7.94 (m, 2H), 7.62 (ddt,  $J = 8.8, 7.8, 1.3$  Hz, 1H), 7.52 – 7.44 (m, 4H), 7.27 – 7.18 (m, 3H), 7.17 – 7.12 (m, 2H), 7.04 – 6.98 (m, 1H), 6.93 – 6.88 (m, 2H), 5.66 (s, 1H), 5.18 – 5.09 (m, 2H), 4.98 (d,  $J = 8.2$  Hz, 1H), 4.77 – 4.64 (m, 3H), 4.41 – 4.30 (m, 3H), 4.20 – 4.11 (m, 2H), 4.00 – 3.90 (m, 2H), 3.86 (s, 3H), 3.75 – 3.62 (m, 2H), 3.35 (d,  $J = 3.5$  Hz, 1H), 3.28 (ddd,  $J = 13.4, 5.6, 3.5$  Hz, 1H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  169.21, 165.54, 163.22, 138.39, 138.24, 133.00, 129.89, 129.55, 128.35, 128.30, 128.16, 127.62, 127.54, 126.96, 126.26, 100.95, 100.77, 100.20, 80.07, 76.85, 75.03, 74.26, 73.75, 72.73, 68.62, 68.50, 66.77, 52.90, 52.23, 50.38, 48.55. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{H}]^+$   $\text{C}_{38}\text{H}_{39}\text{O}_{16}\text{Cl}_3\text{N}_4\text{S}$ : 944.1153; Found: 944.1157.

**Synthesis of compound 23a:** Compound **22** (35 mg, 0.03 mmol) followed the general procedure for sulfation reaction to afford **23a** (25 mg, 66 %) as a white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  8.07 – 7.98 (m, 2H), 7.67 – 7.59 (m, 1H), 7.54 – 7.46 (m, 2H), 7.29 – 7.22 (m, 2H), 7.20 – 7.10 (m, 3H), 5.15 (t,  $J = 7.6$  Hz, 1H), 5.10 (dd,  $J = 11.2, 3.2$  Hz, 1H), 4.95 – 4.90 (m, 3H), 4.68 (d,  $J = 11.4$  Hz, 1H), 4.40 (t,  $J = 7.6$  Hz, 1H), 4.36 – 4.30 (m, 1H), 4.27 – 4.12 (m, 4H), 4.01 – 3.92 (m, 2H), 3.89 (ddd,  $J = 7.0, 5.8, 1.1$  Hz, 1H), 3.83 (s, 3H), 3.69 (ddd,  $J = 11.0, 7.4, 3.6$  Hz, 1H), 3.44 – 3.33 (m, 2H), 3.32 – 3.27 (m, 1H), 2.87 – 2.74 (m, 2H), 2.67 – 2.54 (m, 2H), 2.18 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  208.05, 172.16, 169.19, 165.53, 162.97, 138.01, 132.96, 129.78, 129.53, 128.13, 127.99, 127.77, 127.70, 127.10, 100.52, 99.45, 92.69, 79.24, 76.05, 74.58, 73.78, 72.83, 72.78, 72.72, 68.36, 65.27, 64.89, 52.43, 52.10, 50.31, 37.14, 29.35, 28.36, 27.53. IR ( $\text{cm}^{-1}$ ,  $\text{CH}_2\text{Cl}_2$ ) 3354, 2919, 2849, 2354, 2106, 1719, 1659. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{H}]^+$   $\text{C}_{36}\text{H}_{41}\text{O}_{14}\text{Cl}_3\text{N}_4\text{S}$ : 954.1208; Found: 954.1215.

**Synthesis of compound 23b:** Compound **22** (30 mg, 0.034 mmol) followed the general procedure for sulfation reaction to afford **23b** (29 mg, 82 %) as a white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  8.06 – 7.99 (m, 2H), 7.67 – 7.59 (m, 1H), 7.54 – 7.47 (m, 2H), 7.28 – 7.22 (m, 2H), 7.17 – 7.09 (m, 3H), 5.18 (dd,  $J = 8.1, 3.3$  Hz, 1H), 5.16 – 5.13 (m, 1H), 5.00 – 4.89 (m, 5H), 4.68 (d,  $J = 11.4$  Hz, 1H), 4.48 – 4.40 (m, 2H), 4.31 – 4.19 (m, 3H), 4.11 – 4.05 (m, 1H), 4.03 – 3.92 (m, 2H), 3.83 (s, 3H), 3.73 – 3.66 (m, 2H), 3.41 – 3.35 (m, 2H), 3.32 – 3.25 (m, 1H), 2.91 – 2.70 (m, 2H), 2.62 (t,  $J = 6.5$  Hz, 2H), 2.16 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  208.32, 172.59, 169.12, 165.58, 162.93, 138.02, 132.98, 129.79, 129.51, 128.16,

127.74, 127.00, 100.63, 99.11, 92.65, 79.23, 76.00, 74.49, 73.89, 72.76, 72.60, 71.42, 70.51, 68.38, 52.51, 52.13, 50.32, 37.14, 31.67, 29.50, 29.35, 29.07, 28.38, 27.72, 22.34, 13.05. IR ( $\text{cm}^{-1}$ ,  $\text{CH}_2\text{Cl}_2$ ) 3452, 2991, 2105, 1715, 1636, 1539. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}]^{2-}$   $\text{C}_{36}\text{H}_{39}\text{O}_{21}\text{Cl}_3\text{N}_4\text{S}_2$ : 516.0312; Found: 516.0317.

**Synthesis of compound CS0S:** Compound **23c** (30 mg, 0.034 mmol) followed the general procedure for hydrolysis and hydrogenolysis reactions to afford **CS0S** [7 mg, 46 %, (over all three reaction)] as a white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  4.46 (d,  $J = 7.9$  Hz, 1H), 4.41 (d,  $J = 8.4$  Hz, 1H), 4.00 (dt,  $J = 10.9, 5.2$  Hz, 1H), 3.90 (t,  $J = 5.1$  Hz, 1H), 3.86 – 3.83 (m, 1H), 3.82 – 3.78 (m, 1H), 3.75 – 3.69 (m, 3H), 3.66 – 3.60 (m, 2H), 3.57 (ddd,  $J = 8.8, 6.5, 2.2$  Hz, 1H), 3.33 (dd,  $J = 9.4, 7.9$  Hz, 1H), 3.18 (t,  $J = 5.2$  Hz, 2H), 1.97 (s, 3H).  $^{13}\text{C}$  NMR (150 MHz,  $\text{D}_2\text{O}$ )  $\delta$  174.97, 173.80, 102.09, 101.08, 79.46, 75.88, 75.21, 73.74, 72.48, 70.90, 67.69, 65.90, 61.02, 52.30, 39.36, 22.39. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{H}]^+$   $\text{C}_{16}\text{H}_{29}\text{O}_{12}\text{N}_2$ : 441.1720; Found: 441.1723.

**Synthesis of compound CS3S:** Compound **26** (30 mg, 0.031 mmol) followed the general procedure for hydrolysis and hydrogenolysis reactions to afford **CS3S** [7 mg, 42 %, (over all three reaction)] as a white solid.  $^1\text{H}$  NMR (600 MHz,  $\text{D}_2\text{O}$ )  $\delta$  4.51 (d,  $J = 8.4$  Hz, 1H), 4.43 (d,  $J = 7.9$  Hz, 1H), 4.31 – 4.27 (m, 1H), 4.14 (d,  $J = 3.1$  Hz, 1H), 3.99 – 3.92 (m, 2H), 3.87 (ddd,  $J = 11.7, 5.7, 4.0$  Hz, 1H), 3.74 – 3.68 (m, 3H), 3.67 – 3.64 (m, 2H), 3.54 (t,  $J = 9.1$  Hz, 1H), 3.31 (dd,  $J = 9.5, 7.9$  Hz, 1H), 3.21 – 3.14 (m, 2H), 1.94 (s, 3H).  $^{13}\text{C}$  NMR (150 MHz,  $\text{D}_2\text{O}$ )  $\delta$  174.81, 174.33, 102.10, 100.53, 79.51, 77.78, 76.36, 74.77, 73.76, 72.54, 66.18, 65.90, 60.92, 50.40, 39.40, 22.44. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{H}]^+$   $\text{C}_{16}\text{H}_{27}\text{O}_{15}\text{N}_2\text{S}$ : 519.1138; Found: 519.1142.

**Synthesis of compound CS6S:** Compound **23c** (20 mg, 0.034 mmol) followed the general procedure for hydrolysis and hydrogenolysis reactions to afford **CS6S** [6 mg, 50 %, (over all three reaction)] as a white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  4.44 (dd,  $J = 8.2, 6.2$  Hz, 2H), 4.18 – 4.13 (m, 2H), 4.02 – 3.96 (m, 1H), 3.92 – 3.86 (m, 3H), 3.83 (dd,  $J = 10.9, 8.4$  Hz, 1H), 3.68 – 3.63 (m, 3H), 3.55 (ddd,  $J = 8.9, 5.4, 3.2$  Hz, 1H), 3.33 (dd,  $J = 9.4, 7.9$  Hz, 1H), 3.17 (t,  $J = 5.1$  Hz, 2H), 1.96 (s, 3H).  $^{13}\text{C}$  NMR (150 MHz,  $\text{D}_2\text{O}$ )  $\delta$  174.96, 174.07, 102.03, 101.46, 80.78, 76.37, 73.86, 72.69, 72.47, 70.73, 67.37, 67.14, 65.80, 52.10, 39.38, 22.43. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{H}]^+$   $\text{C}_{16}\text{H}_{27}\text{O}_{15}\text{N}_2\text{S}$ : 519.1138; Found: 519.1136.

**Synthesis of compound CS4,6S:** Compound **23c** (25 mg, 0.024 mmol) followed the general

procedure for hydrolysis and hydrogenolysis reactions to afford **CS4,6S** [6 mg, 41 %, (over all three reaction)] as a white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  4.61 – 4.58 (m, 1H), 4.55 (d,  $J = 8.0$  Hz, 1H), 4.33 (dd,  $J = 11.3, 3.4$  Hz, 1H), 4.29 – 4.24 (m, 1H), 4.15 – 4.06 (m, 2H), 4.01 – 3.95 (m, 1H), 3.93 – 3.89 (m, 2H), 3.77 (d,  $J = 3.2$  Hz, 2H), 3.66 (ddd,  $J = 8.8, 5.5, 3.0$  Hz, 1H), 3.43 (dd,  $J = 9.4, 7.9$  Hz, 2H), 3.26 (t,  $J = 5.1$  Hz, 2H), 2.05 (s, 3H).  $^{13}\text{C}$  NMR (150 MHz,  $\text{D}_2\text{O}$ )  $\delta$  174.95, 173.86, 101.99, 101.56, 81.38, 76.26, 75.32, 73.98, 72.41, 72.26, 69.76, 67.74, 65.78, 52.39, 39.38, 22.44. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{H}]^{2-}$   $\text{C}_{16}\text{H}_{26}\text{O}_{18}\text{N}_2\text{S}_2$ : 299.0317; Found: 299.0323.

**Synthesis of compound 27:** Compound (Glycosyl donor) **8** (1.8 g, 2.76 mmol) and (glycosyl acceptor) **15** (1.38 g, 2.48 mmol) were dissolved in anhydrous DCM (30 mL) with freshly dried 4 Å MS and stirred at RT for 1 h. AgOTf (3.55 g, 13.82 mmol) was added at the same temperature. After consumption of donor, the reaction mixture was quenched with  $\text{Et}_3\text{N}$  and filtered on celite bed and the organic layer was washed with water (3x), the combined organic layer was dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/3) to afford **27** (1.76 g, 67 %) as white foam.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.01 – 7.94 (m, 2H), 7.63 – 7.53 (m, 1H), 7.49 – 7.39 (m, 4H), 7.36 – 7.23 (m, 4H), 7.19 (dd,  $J = 9.1, 3.0$  Hz, 1H), 7.16 – 7.04 (m, 4H), 7.05 – 6.92 (m, 2H), 5.50 (s, 1H), 5.25 – 5.14 (m, 2H), 5.02 (d,  $J = 11.7$  Hz, 1H), 4.82 (d,  $J = 8.2$  Hz, 1H), 4.70 (dd,  $J = 10.9, 3.8$  Hz, 2H), 4.55 (dd,  $J = 11.9, 2.1$  Hz, 1H), 4.50 – 4.33 (m, 2H), 4.26 (d,  $J = 3.4$  Hz, 1H), 4.20 (dd,  $J = 12.6, 1.5$  Hz, 1H), 4.09 (s, 2H), 3.99 – 3.84 (m, 3H), 3.63 (ddd,  $J = 9.4, 4.8, 2.1$  Hz, 1H), 3.37 (s, 1H), 2.76 – 2.66 (m, 1H), 2.59 – 2.54 (m, 1H), 2.32 (s, 3H), 2.05 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  206.64, 172.69, 167.47, 165.19, 162.26, 138.56, 138.14, 137.61, 133.78, 133.28, 129.97, 129.87, 129.64, 129.16, 128.47, 128.23, 128.08, 128.01, 127.33, 126.42, 100.99, 100.71, 92.52, 86.28, 81.95, 76.64, 75.13, 72.90, 71.91, 70.42, 68.75, 66.54, 64.28, 53.07, 40.89, 37.83, 29.79, 28.19, 21.25. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{H}]^+$   $\text{C}_{49}\text{H}_{50}\text{O}_{14}\text{Cl}_4\text{NS}$ : 1048.1706; Found: 1048.1706.

**Synthesis of compound 28:** Compound **27** (1.7 g, 1.62 mmol) was dissolved in mixture of Pyridine:MeOH [(v/v), (1:1), 20 mL] and thiourea (0.24 g, 3.24 mmol) was added. The reaction mixture was reflux at 80 °C for 2 h. The mixture was allow to ambient temperature and diluated with EtOAc and washed with 1N HCl solution (3x) and brine solution (2x). The combined organic layer dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/0.5) to afford

**28** (0.69 g, 88 %) as a white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.01 – 7.94 (m, 2H), 7.63 – 7.54 (m, 1H), 7.53 – 7.40 (m, 5H), 7.33 – 7.22 (m, 4H), 7.16 – 7.02 (m, 4H), 7.05 – 6.86 (m, 3H), 5.45 (s, 1H), 5.26 – 5.16 (m, 2H), 5.09 (d,  $J = 11.7$  Hz, 1H), 4.92 (d,  $J = 8.2$  Hz, 1H), 4.72 (dd,  $J = 10.9, 7.7$  Hz, 2H), 4.47 (dt,  $J = 11.1, 8.7$  Hz, 1H), 4.23 (dd,  $J = 12.4, 1.5$  Hz, 1H), 4.18 – 4.08 (m, 2H), 3.94 (d,  $J = 12.6$  Hz, 1H), 3.91 – 3.79 (m, 3H), 3.48 – 3.35 (m, 2H), 2.91 (s, 1H), 2.79 – 2.58 (m, 2H), 2.56 (qd,  $J = 6.5, 4.7$  Hz, 2H), 2.29 (s, 3H), 2.02 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  206.73, 172.74, 165.54, 162.31, 138.50, 138.39, 137.79, 133.31, 133.17, 130.04, 129.87, 129.08, 128.84, 128.47, 128.43, 128.19, 128.03, 127.24, 126.48, 100.94, 100.84, 92.79, 87.05, 81.72, 79.22, 76.42, 75.36, 72.88, 72.22, 70.94, 68.84, 66.34, 61.05, 52.86, 37.86, 29.80, 28.35, 21.23. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{Na}]^+$   $\text{C}_{47}\text{H}_{48}\text{O}_{13}\text{Cl}_3\text{NSNa}$ : 994.1810; Found: 994.1983.

**Synthesis of compound 29:** Compound **28** (1.2 g, 1.23 mmol) is dissolved in mixture of DCM: $t$ BuOH: $\text{H}_2\text{O}$  [(4/1/1), (v/v/v), 20 mL] were added 2,2,6,6-tetramethyl-1-piperidinyloxy free radical (TEMPO, 159 mg, 0.30 mmol) and bis(acetoxy)iodobenzene (BAIB, 0.55 g, 2.47 mmol) at room temperature. After 6 h, the mixture was quenched with saturated  $\text{NH}_4\text{Cl}$  solution and extracted with DCM. The organic layer was dried over  $\text{Na}_2\text{SO}_4$ , filtered, and concentrated under reduced pressure. The residue was taken as such for next step without further purification. The crude was dissolved in DMF (10 mL) and  $\text{K}_2\text{CO}_3$  (0.49 g, 3.56 mmol) methyl iodide (0.66 mL, 10.69 mmol) were added at the room temperature. After string 12 h, the reaction mixture diluted with ethyl acetate and washed with brine solution (3x), the combined organic layer was dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/1) to afford **29** (0.82 g, 69 %) as white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.00 – 7.94 (m, 2H), 7.60 – 7.55 (m, 1H), 7.49 – 7.41 (m, 4H), 7.35 – 7.32 (m, 2H), 7.28 (hept,  $J = 1.8$  Hz, 2H), 7.16 – 7.05 (m, 4H), 7.01 – 6.88 (m, 4H), 5.51 (s, 1H), 5.19 (dd,  $J = 9.9, 9.0$  Hz, 1H), 5.08 (dd,  $J = 11.3, 3.1$  Hz, 2H), 4.98 (d,  $J = 8.4$  Hz, 1H), 4.75 (d,  $J = 9.9$  Hz, 1H), 4.64 (d,  $J = 11.4$  Hz, 1H), 4.46 (dt,  $J = 11.1, 8.6$  Hz, 1H), 4.30 – 4.16 (m, 3H), 4.02 – 3.95 (m, 2H), 3.84 (d,  $J = 5.4$  Hz, 4H), 3.49 (q,  $J = 1.5$  Hz, 1H), 2.80 – 2.63 (m, 2H), 2.60 – 2.56 (m, 2H), 2.31 (s, 3H), 2.04 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  206.47, 172.45, 169.57, 165.02, 162.00, 138.47, 137.90, 137.67, 133.35, 133.18, 129.88, 129.76, 129.66, 129.00, 128.44, 128.38, 128.33, 128.17, 128.10, 127.94, 127.28, 126.39, 100.94, 100.58, 92.62, 87.20, 81.39, 78.48, 75.76, 72.96, 71.57, 71.40, 68.62, 66.73, 53.18, 52.37, 37.73, 29.70, 28.20, 21.16. IR ( $\text{cm}^{-1}$ ,  $\text{CH}_2\text{Cl}_2$ ) 2919, 1716, 1512, 1452, 731, 628. HRMS (ESI)

m/z: calc'd for  $[M+Na]^+$   $C_{48}H_{48}Cl_3O_{14}NSNa$ : 1022.1759; Found: 1022.1771.

**Synthesis of compound 30:** Compound (Glycosyl donor) **29** (0.5 g, 0.50 mmol) and (glycosyl acceptor) **25** (0.3 g, 0.35 mmol) were dissolved in anhydrous DCM (20 mL) with freshly dried 4 Å MS and stirred at RT for 1 h. NIS (0.13 g, 0.6 mmol) and TMSOTf (18  $\mu$ L, 0.32 mmol) were added at the same temperature. After consumption of donor, the reaction mixture was quenched with  $Et_3N$  and filtered on celite bed and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/0.5) to afford **30** (0.35 g, 57 %) as white foam.  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  7.94 (td,  $J = 8.6, 1.4$  Hz, 4H), 7.56 – 7.52 (m, 2H), 7.44 (dd,  $J = 7.6, 1.9$  Hz, 3H), 7.42 – 7.39 (m, 4H), 7.38 (s, 1H), 7.29 – 7.27 (m, 3H), 7.22 (d,  $J = 2.2$  Hz, 1H), 7.20 – 7.19 (m, 2H), 7.12 – 7.09 (m, 2H), 7.09 – 6.98 (m, 5H), 6.96 – 6.91 (m, 3H), 6.66 (d,  $J = 9.0$  Hz, 1H), 5.51 (s, 1H), 5.48 (s, 1H), 5.29 (t,  $J = 6.3$  Hz, 1H), 5.23 (dd,  $J = 8.5, 7.0$  Hz, 1H), 5.13 (d,  $J = 8.2$  Hz, 1H), 5.04 (d,  $J = 6.0$  Hz, 1H), 4.99 – 4.89 (m, 3H), 4.82 – 4.74 (m, 2H), 4.70 – 4.63 (m, 2H), 4.49 – 4.35 (m, 5H), 4.31 – 4.24 (m, 2H), 4.19 (dd,  $J = 5.2, 2.5$  Hz, 2H), 4.16 – 4.08 (m, 3H), 4.01 – 3.94 (m, 4H), 3.92 (dd,  $J = 3.2, 1.8$  Hz, 1H), 3.89 – 3.85 (m, 2H), 3.85 – 3.83 (m, 1H), 3.80 (s, 3H), 3.78 (s, 3H), 3.63 (ddd,  $J = 11.0, 7.4, 3.9$  Hz, 1H), 3.44 (s, 1H), 3.39 – 3.25 (m, 3H), 3.24 – 3.21 (m, 1H), 2.79 – 2.62 (m, 2H), 2.60 – 2.54 (m, 2H), 2.17 (s, 2H), 2.04 (s, 3H).  $^{13}C$  NMR (100 MHz,  $CDCl_3$ )  $\delta$  206.42, 172.37, 169.82, 169.31, 165.17, 165.09, 161.97, 161.76, 138.08, 137.99, 137.96, 137.54, 133.40, 133.08, 129.99, 129.81, 129.73, 129.49, 129.29, 129.00, 128.82, 128.43, 128.30, 128.17, 128.09, 128.06, 127.94, 127.40, 127.19, 126.34, 126.21, 101.20, 100.86, 100.66, 100.32, 99.94, 98.75, 92.56, 92.48, 80.51, 79.65, 78.19, 74.74, 74.50, 74.44, 74.35, 73.39, 72.84, 72.32, 71.39, 68.89, 68.64, 68.17, 66.74, 66.51, 60.42, 54.48, 53.05, 53.02, 52.29, 50.54, 37.69, 36.64, 33.71, 31.94, 31.44, 30.17, 29.37, 28.15, 26.71, 22.70, 21.06, 14.20, 14.13. IR ( $cm^{-1}$ ,  $CH_2Cl_2$ ) 3029, 2883, 2349, 2336, 2104, 1721, 1657. HRMS (ESI) m/z: calc'd for  $[M+Na]^+$   $C_{79}H_{79}O_{27}N_5SCl_6Na$ : 1762.2991; Found: 1762.3411.

**Synthesis of compound 31:** Compound **30** (0.25 g, 0.14 mmol) was dissolved in mixture of DCM/MeOH [(2/1), (v/v), 12 mL] and p-Toluenesulfonic acid (PTSA, 41 mg, 0.21 mmol) was added at RT. After 6 h stirring, the mixture was quenched by triethylamine (up to PH ~ 7) and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/0.5) to afford **31** (0.15 g, 67 %) as white solid.  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  8.04 – 7.88 (m, 4H), 7.62 – 7.49 (m, 2H), 7.42 (p,  $J = 7.9, 7.4$  Hz, 4H),

7.21 – 7.01 (m, 10H), 5.19 – 5.13 (m, 1H), 4.99 – 4.79 (m, 5H), 4.67 – 4.56 (m, 2H), 4.49 (d,  $J = 11.1$  Hz, 1H), 4.42 – 4.35 (m, 1H), 4.29 (q,  $J = 8.4$  Hz, 1H), 4.19 – 4.14 (m, 1H), 4.09 – 4.01 (m, 2H), 3.96 – 3.92 (m, 1H), 3.85 – 3.79 (m, 2H), 3.79 – 3.70 (m, 6H), 3.64 (tdd,  $J = 15.0, 7.9, 4.1$  Hz, 4H), 3.54 (dt,  $J = 12.4, 5.6$  Hz, 2H), 3.35 (ddq,  $J = 14.5, 6.9, 3.3$  Hz, 1H), 3.29 (q,  $J = 4.8, 4.0$  Hz, 1H), 3.11 (s, 1H), 2.75 (t,  $J = 6.6$  Hz, 2H), 2.65 – 2.51 (m, 2H), 2.18 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  208.11, 172.27, 169.64, 169.23, 165.49, 165.22, 162.38, 161.84, 137.67, 137.36, 133.57, 133.18, 130.09, 129.86, 129.78, 129.62, 129.17, 128.54, 128.36, 128.27, 128.24, 128.20, 128.11, 128.02, 127.99, 127.82, 127.57, 101.18, 100.97, 100.48, 92.66, 92.38, 79.57, 75.14, 75.02, 74.65, 74.36, 73.87, 73.65, 72.84, 68.24, 54.44, 53.40, 53.03, 52.38, 50.52, 38.05, 37.48, 29.85, 28.17, 28.10, 1.03, 0.01. IR ( $\text{cm}^{-1}$ ,  $\text{CH}_2\text{Cl}_2$ ) 3352, 2955, 2917, 2850, 2105, 1721, 1527. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{Na}]^+$   $\text{C}_{65}\text{H}_{71}\text{O}_{27}\text{N}_5\text{Cl}_6\text{Na}$ : 1586.2365; Found: 1586.2773.

**Synthesis of compound 32:** Compound **31** (35 mg, 0.02 mmol) followed the general procedure for sulfation reaction to afford **32** (25 mg, 59 %) as a white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.95 – 7.85 (m, 4H), 7.53 – 7.47 (m, 2H), 7.45 – 7.33 (m, 5H), 7.14 – 7.09 (m, 4H), 7.04 – 7.00 (m, 2H), 7.00 – 6.95 (m, 3H), 5.21 (t,  $J = 7.6$  Hz, 1H), 5.06 (dd,  $J = 11.4, 3.3$  Hz, 1H), 5.01 – 4.92 (m, 2H), 4.84 – 4.77 (m, 3H), 4.71 – 4.65 (m, 2H), 4.58 – 4.42 (m, 3H), 4.34 – 4.24 (m, 4H), 4.18 – 4.07 (m, 5H), 4.05 – 4.00 (m, 2H), 3.98 – 3.94 (m, 1H), 3.88 (t,  $J = 6.1$  Hz, 1H), 3.84 – 3.80 (m, 1H), 3.79 (d,  $J = 4.5$  Hz, 3H), 3.73 (d,  $J = 6.5$  Hz, 1H), 3.69 (d,  $J = 3.9$  Hz, 3H), 3.57 – 3.52 (m, 1H), 3.45 (dd,  $J = 5.5, 2.2$  Hz, 1H), 3.15 (ddd,  $J = 13.5, 5.7, 3.8$  Hz, 1H), 2.77 – 2.60 (m, 2H), 2.51 (dd,  $J = 13.8, 7.3$  Hz, 2H), 2.05 (d,  $J = 1.9$  Hz, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  208.44, 172.59, 169.80, 169.23, 166.30, 165.60, 163.00, 162.76, 137.97, 137.53, 132.95, 130.17, 129.52, 128.71, 128.24, 128.15, 127.90, 127.75, 127.70, 127.33, 126.96, 100.94, 100.49, 99.64, 99.30, 92.61, 92.57, 79.24, 79.13, 76.01, 75.55, 75.38, 75.25, 75.03, 74.53, 74.39, 73.79, 73.16, 73.15, 72.99, 72.85, 72.77, 72.77, 72.69, 71.29, 68.35, 66.92, 65.07, 63.88, 62.64, 53.74, 52.59, 52.05, 50.30, 37.16, 29.36, 28.39, 27.73. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}/\text{z}]^{4+}$   $\text{C}_{65}\text{H}_{67}\text{O}_{39}\text{N}_5\text{Cl}_6\text{S}$ : 470.2605; Found: 470.2602.

**Synthesis of compound CS-T-4,6S:** Compound **32** (20 mg, 0.01 mmol) followed the general procedure for hydrolysis and hydrogenolysis reactions to afford **CS-T-4,6S** [4.5 mg, 37 %, (over all three reaction)] as a white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  4.80 (d,  $J = 10.1$  Hz, 1H), 4.66 – 4.64 (m, 1H), 4.61 (s, 1H), 4.52 (d,  $J = 7.5$  Hz, 1H), 4.49 – 4.37 (m, 3H), 4.25 –

4.13 (m, 4H), 4.07 – 3.95 (m, 5H), 3.94 – 3.81 (m, 3H), 3.71 – 3.65 (m, 3H), 3.61 – 3.55 (m, 2H), 3.54 – 3.49 (m, 1H), 3.38 – 3.33 (m, 1H), 3.32 – 3.26 (m, 1H), 3.20 – 3.15 (m, 2H), 1.98 (s, 3H), 1.96 (s, 3H). HRMS (ESI)  $m/z$ : calc'd for  $[M/z]^{4+}$  C<sub>30</sub>H<sub>45</sub>O<sub>35</sub>N<sub>3</sub>S: 283.7685; Found: 283.7680.

**Synthesis of compound 34:** Compound **30** (0.15 g, 0.086 mmol) was dissolved in mixture of THF/MeOH [(v/v), (10/1), xx mL]. Hydrazine hydrate/acetic acid [(v/v), (2.5/1)] were added in mixture THF:MeOH (5:1) [(v/v), (10/1)], the resulting solution was added to reaction mixture at room temperature. After 1 h the reaction mixture was diluted with EtOAc and washed with aq. NaHCO<sub>3</sub> solution (3x) and brine solution (2x). The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/0.7) to afford **34** (120 mg, 84 %) as white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.94 (ddd,  $J = 8.1, 6.4, 1.3$  Hz, 4H), 7.59 – 7.52 (m, 2H), 7.46 – 7.37 (m, 8H), 7.33 – 7.28 (m, 3H), 7.23 – 7.19 (m, 3H), 7.12 – 7.09 (m, 2H), 7.08 – 7.02 (m, 3H), 7.02 – 6.91 (m, 5H), 5.52 (d,  $J = 10.8$  Hz, 2H), 5.31 – 5.20 (m, 2H), 5.15 (d,  $J = 8.2$  Hz, 1H), 5.03 (d,  $J = 6.2$  Hz, 1H), 4.96 (d,  $J = 11.5$  Hz, 2H), 4.74 – 4.63 (m, 4H), 4.47 – 4.38 (m, 3H), 4.36 – 4.26 (m, 2H), 4.19 – 4.07 (m, 5H), 3.99 (d,  $J = 8.8$  Hz, 2H), 3.96 – 3.89 (m, 3H), 3.87 – 3.82 (m, 2H), 3.80 (s, 6H), 3.63 (ddd,  $J = 10.9, 7.3, 4.0$  Hz, 2H), 3.45 (s, 1H), 3.34 (dd,  $J = 7.4, 3.9$  Hz, 1H), 3.31 – 3.23 (m, 2H), 2.78 (d,  $J = 11.0$  Hz, 1H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 170.23, 169.38, 165.27, 165.22, 163.18, 161.86, 138.15, 138.06, 137.97, 137.51, 133.51, 133.19, 130.07, 129.89, 129.79, 129.38, 129.29, 128.82, 128.53, 128.39, 128.30, 128.24, 128.16, 128.13, 128.02, 127.54, 127.28, 126.37, 126.29, 101.28, 100.68, 100.45, 100.10, 98.71, 92.72, 92.54, 80.60, 79.73, 78.56, 77.32, 76.70, 74.91, 74.85, 74.83, 74.57, 74.53, 73.49, 72.90, 72.57, 72.23, 68.97, 68.75, 68.27, 66.84, 66.78, 60.50, 55.61, 54.59, 50.61, 36.71, 29.79, 24.77, 21.14, 14.28. HRMS (ESI)  $m/z$ : calc'd for  $[M+Na]^+$  C<sub>74</sub>H<sub>73</sub>Cl<sub>6</sub>N<sub>5</sub>O<sub>25</sub>Na: 1664.2623; Found: 1664.2643.

**Synthesis of compound 36:** Compound **29** (0.2 g, 0.2 mmol) was dissolved in mixture of DCM/H<sub>2</sub>O [(15/1), (v/v), 10 mL] and NIS (54 mg, 0.24 mmol) and TfOH (3.5 μL, 0.04 mmol) were added at RT. After 15 min the reaction mixture was quenched with triethylamine and washed with sodium thiosulfate solution (2x). The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/1) to afford **36** (0.13 g, 77 %) as white foam. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.00 – 7.94 (m, 2H), 7.62 – 7.47 (m, 3H), 7.46 –



7.37 (m, 2H), 7.35 – 7.28 (m, 3H), 7.24 (d,  $J = 1.7$  Hz, 1H), 7.10 – 6.97 (m, 4H), 5.56 (d,  $J = 3.6$  Hz, 1H), 5.52 (d,  $J = 1.8$  Hz, 1H), 5.21 – 4.98 (m, 4H), 4.73 (d,  $J = 11.1$  Hz, 1H), 4.57 – 4.41 (m, 2H), 4.32 – 4.15 (m, 4H), 4.03 – 3.88 (m, 1H), 3.82 (s, 3H), 3.58 – 3.41 (m, 2H), 2.81 – 2.63 (m, 2H), 2.62 – 2.53 (m, 2H), 2.04 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  206.71, 172.50, 170.95, 165.95, 162.18, 138.29, 137.78, 133.41, 130.03, 129.94, 129.48, 129.09, 128.53, 128.38, 128.19, 128.15, 127.46, 126.51, 126.48, 101.02, 100.73, 96.05, 92.69, 90.37, 78.90, 76.04, 73.04, 72.56, 71.69, 70.10, 68.70, 66.83, 60.53, 53.14, 52.46, 37.81, 29.78, 28.28, 21.15, 14.28. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{H}]^+$   $\text{C}_{41}\text{H}_{43}\text{O}_{15}\text{NCl}_3$ : 894.1698; Found: 894.1671.

**Synthesis of compound 37:** Compound **36** (0.12 g, 0.13 mmol) was dissolved in anhydrous DCM (10 mL) and Cesium carbonate (48 mg, 2.75 mmol) and 2,2,2-Trifluoro-N-phenylacetimidoyl chloride (43  $\mu\text{L}$ , 0.26 mmol) were added at RT. After 12 h the reaction mixture was concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/3) to afford **37** (95 mg, 66 %) as white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.00 – 7.91 (m, 2H), 7.67 – 7.59 (m, 1H), 7.55 – 7.43 (m, 4H), 7.32 (dd,  $J = 5.1, 1.9$  Hz, 3H), 7.26 – 7.22 (m, 2H), 7.17 – 6.98 (m, 7H), 6.42 (d,  $J = 7.7$  Hz, 1H), 5.55 (s, 1H), 5.32 (dd,  $J = 9.8, 3.5$  Hz, 1H), 5.20 (d,  $J = 11.3$  Hz, 1H), 5.13 (dd,  $J = 11.2, 3.4$  Hz, 1H), 5.01 (d,  $J = 8.4$  Hz, 1H), 4.75 (d,  $J = 11.3$  Hz, 1H), 4.54 – 4.41 (m, 2H), 4.34 – 4.16 (m, 4H), 4.00 (dd,  $J = 12.6, 1.8$  Hz, 1H), 3.88 (s, 3H), 3.51 (q,  $J = 1.4$  Hz, 1H), 2.82 – 2.65 (m, 2H), 2.62 – 2.56 (m, 2H), 2.05 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  206.57, 172.51, 169.78, 165.35, 162.19, 142.84, 138.09, 137.72, 133.67, 129.95, 129.14, 129.01, 128.69, 128.34, 128.21, 127.54, 126.50, 124.45, 119.11, 101.06, 100.54, 92.67, 92.08, 78.15, 76.18, 73.03, 72.17, 71.50, 70.88, 68.72, 66.92, 53.45, 52.52, 37.83, 29.79, 28.28. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{Na}]^+$   $\text{C}_{49}\text{H}_{46}\text{O}_{15}\text{N}_2\text{Cl}_3\text{F}_3\text{Na}$ : 1087.1814; Found: 1087.1819.

**Synthesis of compound 38:** Compound **36** (0.15 g, 0.16 mmol) was dissolved in anhydrous DCM (10 mL) and Cesium carbonate (21 mg, 0.067 mmol) and trichloroacetonitrile (0.13 mL, 1.34 mmol) were added at RT. After 12 h the reaction mixture was concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/3) to afford **38** (0.12 g, 69 %) as white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.00 – 7.91 (m, 2H), 7.67 – 7.59 (m, 1H), 7.55 – 7.43 (m, 4H), 7.32 (dd,  $J = 5.1, 1.9$  Hz, 3H), 7.26 – 7.22 (m, 2H), 7.17 – 6.98 (m, 7H), 6.42 (d,  $J = 7.7$  Hz, 1H), 5.55 (s, 1H), 5.32 (dd,  $J = 9.8, 3.5$  Hz, 1H), 5.20 (d,  $J = 11.3$  Hz, 1H), 5.13 (dd,  $J = 11.2, 3.4$  Hz, 1H), 5.01 (d,  $J$

= 8.4 Hz, 1H), 4.75 (d,  $J$  = 11.3 Hz, 1H), 4.54 – 4.41 (m, 2H), 4.34 – 4.16 (m, 4H), 4.00 (dd,  $J$  = 12.6, 1.8 Hz, 1H), 3.88 (s, 3H), 3.51 (q,  $J$  = 1.4 Hz, 1H), 2.82 – 2.65 (m, 2H), 2.62 – 2.56 (m, 2H), 2.05 (s, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  206.57, 172.51, 169.78, 165.35, 162.19, 142.84, 138.09, 137.72, 133.67, 129.95, 129.14, 129.01, 128.69, 128.34, 128.21, 127.54, 126.50, 124.45, 119.11, 101.06, 100.54, 92.67, 92.08, 78.15, 76.18, 73.03, 72.17, 71.50, 70.88, 68.72, 66.92, 53.45, 52.52, 37.83, 29.79, 28.28. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{Na}]^+$   $\text{C}_{43}\text{H}_{42}\text{O}_{15}\text{N}_2\text{Cl}_6\text{Na}$ : 1059.0614; Found: 1059.0618.

**Synthesis of compound 39:** Compound (Glycosyl donor) **17** (1.07 g, 1.87 mmol) and (glycosyl acceptor) benzyl(3-hydroxypropyl)carbamate (0.29 g, 2.25 mmol) were dissolved in anhydrous DCM (20 mL) with freshly dried 4 Å MS and stirred at RT for 1 h. Cooled the reaction mixture at  $-40\text{ }^\circ\text{C}$  and NIS and TMSOTf (0.63 g, 2.81 mmol) were added at the same temperature. After consumption of donor, the reaction mixture was quenched with  $\text{Et}_3\text{N}$  and filtered on celite bed and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/3) to afford **39** (0.86 g, 78%) as white foam.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.05 (d,  $J$  = 7.6 Hz, 2H), 7.60 (t,  $J$  = 7.4 Hz, 1H), 7.46 (t,  $J$  = 7.6 Hz, 2H), 7.42 – 7.28 (m, 5H), 7.23 (q,  $J$  = 3.8, 3.2 Hz, 5H), 5.26 (t,  $J$  = 8.3 Hz, 1H), 5.07 (s, 3H), 4.80 – 4.62 (m, 2H), 4.60 – 4.43 (m, 3H), 4.09 (s, 2H), 3.96 – 3.84 (m, 1H), 3.76 – 3.52 (m, 4H), 3.17 (dh,  $J$  = 26.1, 6.5 Hz, 2H), 2.77 (s, 1H), 1.82 – 1.61 (m, 2H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  167.64, 165.26, 156.46, 137.63, 136.76, 133.41, 129.75, 129.57, 128.57, 128.49, 128.08, 128.06, 128.03, 101.14, 82.04, 74.74, 73.45, 73.43, 69.87, 67.38, 66.46, 64.54, 40.73, 38.04, 29.71, 29.45. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{H}]^+$   $\text{C}_{33}\text{H}_{37}\text{O}_{10}\text{NCl}$ : 642.2106; Found: 642.2101.

**Synthesis of compound 40:** Compound (Glycosyl donor) **6** (2 g, 3.25 mmol) and (glycosyl acceptor) **39** (1.87 g, 2.92 mmol) were dissolved in anhydrous DCM (20 mL) with freshly dried 4 Å MS and stirred at RT for 1 h. Cooled the reaction mixture at  $-40\text{ }^\circ\text{C}$  and NIS (0.87 g, 3.90 mmol) and TMSOTf (117  $\mu\text{L}$ , 0.65 mmol) were added at the same temperature. After consumption of donor, the reaction mixture was quenched with  $\text{Et}_3\text{N}$  and filtered on celite bed and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/1) to afford **40** (2.8 g, 76 %) as white foam.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.92 (d,  $J$  = 7.9 Hz, 2H), 7.55 – 7.43 (m, 4H), 7.40 – 7.36 (m, 2H), 7.33 – 7.26 (m, 7H), 7.14 – 6.95 (m, 5H), 5.48 (s, 1H), 5.20 (t,  $J$  = 8.5 Hz, 2H), 5.10 – 5.02 (m, 4H), 4.86 – 4.84 (m, 1H), 4.75 (d,  $J$  = 11.9 Hz, 1H), 4.52 – 4.40 (m, 4H), 4.23 – 4.17 (m,

2H), 4.11 – 3.97 (m, 3H), 3.91 – 3.79 (m, 3H), 3.63 (d,  $J = 8.4$  Hz, 1H), 3.52 – 3.47 (m, 1H), 3.35 (s, 1H), 3.17 – 3.03 (m, 2H), 2.74 – 2.51 (m, 4H), 2.02 (s, 3H), 1.68 – 1.57 (m, 1H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  206.83, 172.62, 167.64, 165.38, 162.37, 156.57, 138.33, 137.70, 136.85, 133.37, 129.87, 129.66, 129.16, 128.60, 128.54, 128.24, 128.12, 127.95, 127.35, 126.44, 100.94, 100.67, 92.67, 80.60, 74.85, 73.13, 72.83, 70.60, 68.80, 67.49, 66.49, 64.24, 52.96, 40.95, 38.15, 37.87, 29.81, 29.52, 28.29. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{Na}]^+$   $\text{C}_{53}\text{H}_{56}\text{O}_{17}\text{N}_2\text{Cl}_4\text{Na}$ : 1155.2231; Found: 1155.2239.

**Synthesis of compound 41:** Compound **40** (2 g, 1.76 mmol) was dissolved in mixture of THF/MeOH [(v/v), (10/1), xx mL]. Hydrazine hydrate/acetic acid [(v/v), (2.5/1)] were added in mixture THF:MeOH (5:1) [(v/v), (10/1)], the resulting solution was added to reaction mixture at room temperature. After 1 h the reaction mixture was diluted with EtOAc and washed with aq.  $\text{NaHCO}_3$  solution (3x) and brine solution (2x). The combined organic layer was dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/0.7) to afford **41** (1.35 g, 74 %) as white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.93 (d,  $J = 7.4$  Hz, 2H), 7.54 (t,  $J = 7.4$  Hz, 1H), 7.45 – 7.29 (m, 12H), 7.16 (d,  $J = 6.8$  Hz, 2H), 7.06 – 7.01 (m, 3H), 5.53 (s, 1H), 5.21 (t,  $J = 8.3$  Hz, 1H), 5.09 – 5.01 (m, 4H), 4.70 (d,  $J = 11.7$  Hz, 1H), 4.62 (dd,  $J = 17.1, 6.0$  Hz, 1H), 4.51 (d,  $J = 7.8$  Hz, 1H), 4.41 (d,  $J = 11.7$  Hz, 1H), 4.20 – 4.15 (m, 2H), 4.10 – 4.04 (m, 2H), 3.93 – 3.81 (m, 4H), 3.62 – 3.49 (m, 2H), 3.34 (s, 1H), 3.18 – 3.04 (m, 2H), 2.64 (d,  $J = 11.3$  Hz, 1H), 1.71 – 1.65 (m, 2H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  167.85, 165.35, 162.85, 156.53, 138.20, 137.44, 136.79, 133.37, 129.85, 129.63, 129.42, 128.59, 128.54, 128.38, 128.15, 128.13, 128.03, 127.44, 126.43, 101.39, 100.97, 100.75, 92.60, 80.54, 74.96, 74.67, 73.05, 72.92, 70.20, 68.80, 67.40, 66.74, 66.53, 63.98, 56.17, 40.85, 38.08, 29.46. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}/z]^+$   $\text{C}_{48}\text{H}_{50}\text{O}_{15}\text{N}_2\text{Cl}_4\text{Na}$ : 1057.1863; Found: 1057.1871.

**Synthesis of compound 42:** Compound (Glycosyl donor) **27** (1.6 g, 1.52 mmol) and (glycosyl acceptor) **41** (1.26 g, 1.22 mmol) were dissolved in anhydrous DCM (20 mL) with freshly dried 4 Å MS and stirred at RT for 1 h. NIS (0.41 g, 1.83 mmol) and TMSOTf (55  $\mu\text{L}$ , 0.30 mmol) were added at the same temperature. After consumption of donor, the reaction mixture was quenched with  $\text{Et}_3\text{N}$  and filtered on celite bed and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/1) to afford **42** (1.65 g, 69 %) as white foam.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )

$\delta$  8.03 – 7.91 (m, 4H), 7.59 – 7.53 (m, 2H), 7.46 (d,  $J = 7.0$  Hz, 2H), 7.42 – 7.29 (m, 12H), 7.24 – 20 (m, 3H), 7.16 – 7.09 (m, 5H), 7.07 – 6.95 (m, 7H), 5.52 (s, 1H), 5.43 (s, 1H), 5.26 – 5.20 (m, 3H), 5.05 – 4.94 (m, 7H), 4.86 (d,  $J = 11.4$  Hz, 1H), 4.71 (dd,  $J = 11.8, 6.5$  Hz, 2H), 4.60 (d,  $J = 11.0$  Hz, 1H), 4.51 (d,  $J = 8.1$  Hz, 1H), 4.41 – 4.18 (m, 8H), 4.14 – 3.78 (m, 13H), 3.70 – 3.51 (m, 3H), 3.40 – 3.08 (m, 4H), 2.74 – 70 (m, 2H), 2.61 – 2.58 (m, 2H), 2.10 (s, 3H), 1.72 – 1.67 (m, 2H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  206.45, 172.62, 167.53, 167.03, 165.33, 165.21, 162.20, 162.11, 156.45, 138.24, 137.90, 137.67, 137.42, 136.79, 133.41, 133.16, 129.99, 129.75, 129.67, 129.24, 129.12, 128.66, 128.47, 128.42, 128.39, 128.16, 128.13, 128.08, 128.02, 127.97, 127.93, 127.85, 127.81, 127.75, 127.67, 127.38, 127.15, 126.59, 126.36, 126.28, 126.08, 100.90, 100.73, 100.51, 100.30, 99.79, 98.79, 92.56, 92.34, 80.62, 76.18, 74.71, 74.66, 74.61, 73.60, 73.06, 72.91, 72.83, 72.75, 71.43, 70.02, 68.79, 68.67, 67.27, 66.75, 66.49, 66.39, 64.29, 63.53, 54.31, 53.31, 41.00, 40.87, 38.09, 37.74, 29.71, 29.38, 28.05. HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{Na}]^+$   $\text{C}_{90}\text{H}_{91}\text{O}_{29}\text{N}_3\text{Cl}_8\text{Na}$ : 1982.3144; Found: 1982.3529.

**Synthesis of compound 43:** Compound **42** (1.5 g, 0.76 mmol) was dissolved in mixture of THF/MeOH [(v/v), (10/1), xx mL]. Hydrazine hydrate/acetic acid [(v/v), (2.5/1)] were added in mixture THF:MeOH (5:1) [(v/v), (10/1)], the resulting solution was added to reaction mixture at room temperature. After 1 h the reaction mixture was diluted with EtOAc and washed with aq.  $\text{NaHCO}_3$  solution (3x) and brine solution (2x). The combined organic layer was dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/0.7) to afford **43** (1.25 g, 84 %) as white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.84 (t,  $J = 6.9$  Hz, 4H), 7.49 – 7.44 (m, 2H), 7.37 – 7.21 (m, 14H), 7.15 – 7.09 (m, 3H), 7.07 – 7.02 (m, 5H), 6.99 – 6.85 (m, 7H), 5.47 (s, 1H), 5.33 (s, 1H), 5.17 – 5.10 (m, 2H), 4.95 – 4.84 (m, 6H), 4.78 – 4.70 (m, 2H), 4.59 (dd,  $J = 11.8, 4.8$  Hz, 2H), 4.51 – 4.48 (m, 1H), 4.41 (d,  $J = 7.8$  Hz, 1H), 4.35 – 4.26 (m, 3H), 4.19 – 3.73 (m, 20H), 3.59 – 3.39 (m, 3H), 3.28 (s, 1H), 3.18 (s, 1H), 3.09 – 2.99 (m, 1H), 2.63 – 2.37 (m, 2H), 1.62 – 1.54 (m, 2H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  167.60, 167.07, 165.35, 165.23, 162.67, 162.10, 156.44, 138.23, 137.89, 137.68, 137.26, 133.41, 133.18, 129.98, 129.75, 129.66, 129.36, 129.28, 128.65, 128.47, 128.43, 128.40, 128.29, 128.08, 128.02, 127.97, 127.94, 127.86, 127.80, 127.42, 127.16, 126.31, 126.07, 101.33, 100.74, 100.49, 100.00, 99.76, 98.91, 92.54, 92.44, 80.61, 80.60, 76.18, 74.79, 74.60, 73.67, 73.07, 72.90, 72.71, 71.62, 71.60, 69.37, 68.78, 68.72, 68.68, 67.27, 66.75, 66.39, 64.33, 63.58, 56.68, 54.36, 49.45, 49.15, 40.97, 40.86, 38.08, 29.71, 29.38. HRMS (ESI)  $m/z$ : calc'd for

$[M+Na]^+$   $C_{85}H_{85}O_{27}N_3Cl_8Na$ : 1882.2776; Found: 1882.2780.

**Synthesis of compound 44:** Compound (Glycosyl donor) **6** (0.5 g, 0.47 mmol) and (glycosyl acceptor) **39** (0.71 g, 0.38 mmol) were dissolved in anhydrous DCM (20 mL) with freshly dried 4 Å MS and stirred at RT for 1 h. NIS (0.12 g, 0.57 mmol) and TMSOTf (17  $\mu$ L, 0.095 mmol) were added at the same temperature. After consumption of donor, the reaction mixture was quenched with  $Et_3N$  and filtered on celite bed and concentrated under reduced pressure. The crude product was purified by column chromatography (EtOAc/Hexane, 1/1) to afford **44** (0.65 g, 61 %) as white foam.  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  7.91 – 7.86 (m, 5H), 7.55 – 7.49 (m, 4H), 7.44 – 7.42 (m, 2H), 7.38 – 7.35 (m, 5H), 7.33 – 7.30 (m, 7H), 7.28 – 7.27 (m, 3H), 7.21 – 7.16 (m, 4H), 7.12 – 7.07 (m, 9H), 7.05 – 6.97 (m, 8H), 6.94 – 6.89 (m, 3H), 5.49 (s, 1H), 5.41 (s, 1H), 5.35 (s, 1H), 5.23 – 5.12 (m, 5H), 5.07 – 5.01 (m, 4H), 4.98 – 4.85 (m, 8H), 4.67 (dd,  $J = 11.7, 6.7$  Hz, 2H), 4.57 (d,  $J = 11.6$  Hz, 2H), 4.46 (d,  $J = 8.0$  Hz, 1H), 4.37 – 4.23 (m, 9H), 4.21 – 3.75 (m, 25H), 3.66 – 3.63 (m, 1H), 3.59 – 3.55 (m, 2H), 3.50 – 3.45 (m, 1H), 3.29 – 3.24 (m, 2H), 3.15 – 3.04 (m, 2H), 2.71 – 2.67 (m, 2H), 2.57 – 2.52 (m, 2H), 2.06 (s, 3H), 1.68 – 1.60 (m, 2H).  $^{13}C$  NMR (100 MHz,  $CDCl_3$ )  $\delta$  206.40, 172.63, 167.50, 167.45, 167.00, 165.38, 165.31, 165.22, 162.20, 162.08, 156.43, 138.26, 137.91, 137.88, 137.63, 137.39, 133.31, 133.14, 129.97, 129.75, 129.31, 129.20, 128.67, 128.46, 128.43, 128.37, 128.29, 128.17, 128.11, 128.02, 127.95, 127.89, 127.85, 127.81, 127.56, 127.41, 127.25, 127.14, 126.27, 126.16, 126.06, 126.03, 100.90, 100.73, 100.56, 100.47, 100.26, 99.98, 99.44, 98.90, 98.43, 92.66, 92.50, 92.33, 80.58, 76.16, 75.44, 74.95, 74.76, 74.67, 74.66, 74.62, 74.08, 73.63, 73.05, 72.94, 72.84, 72.66, 69.90, 68.77, 68.74, 68.68, 67.26, 66.77, 66.49, 64.30, 53.39, 41.09, 40.99, 38.08, 37.73, 29.71, 29.38, 28.03, 1.04. HRMS (ESI)  $m/z$ : calc'd for  $[M+H]^+$   $C_{127}H_{126}O_{41}N_4Cl_{12}Na$ : 2813.4036; Found: 2813.5869.

#### 4a.11 References:

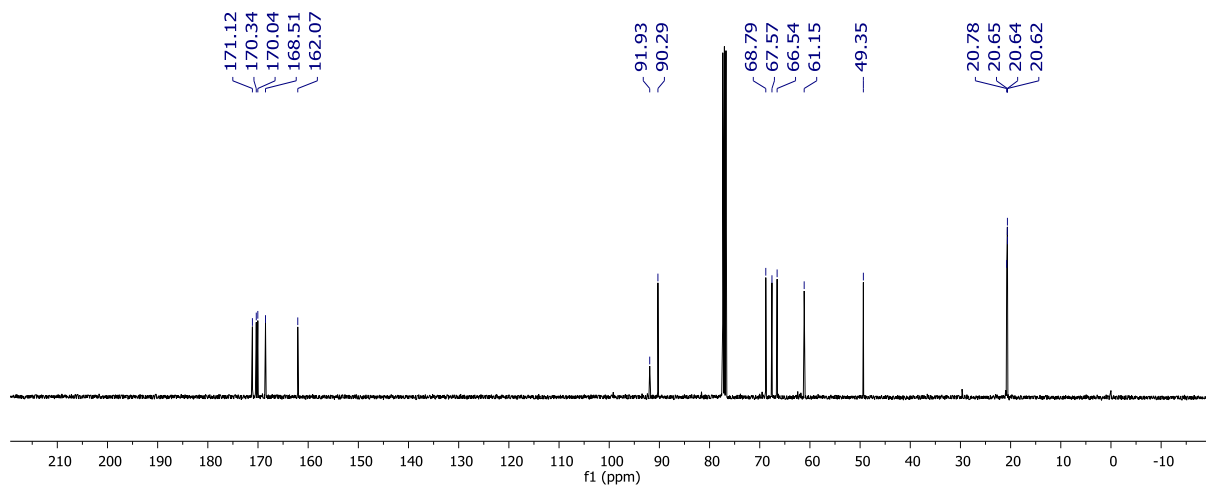
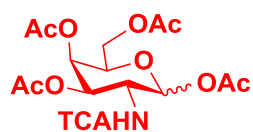
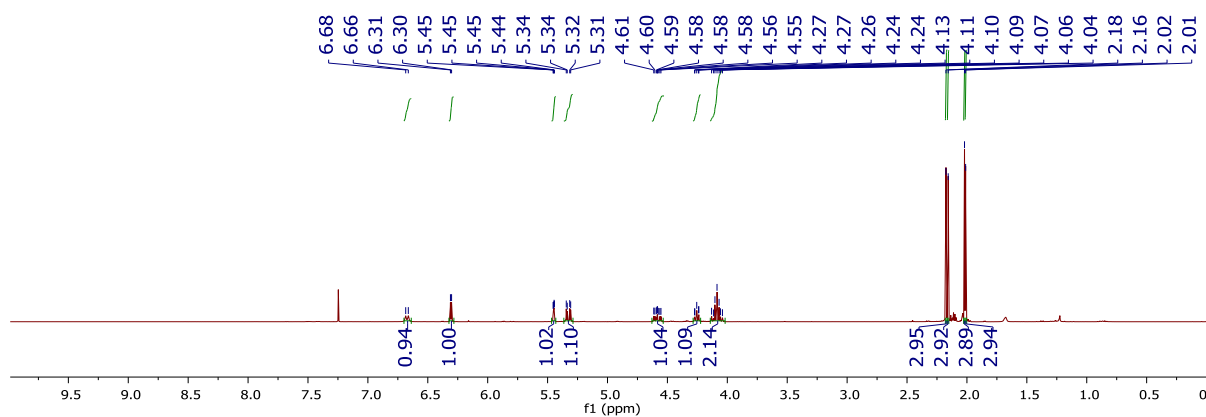
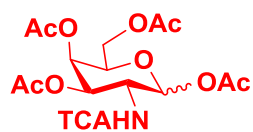
1. Y. Mechref, Y. Hu, A. Garcia, S. Zhou, J. L. Desantos-Garcia and A. Hussein, *Bioanalysis*, 2012, **4**, 2457–2469.
2. Y. Lan, C. Hao, X. Zeng, Y. He, P. Zeng, Z. Guo and L. Zhang, *Am. J. Cancer Res.*, 2016, **6**, 2390-2415.
3. E. Scott and J. Munkley, *Int. J. Mol. Sci.* 2019, **20**, 1389.
4. A. Varki, *Glycobiology*, 1993, **3**, 97-130.
5. Y. H. Chung, H. Cai and N. F. Steinmetz, *Adv. Drug Deliv. Rev.*, 2020, **156**, 214–235.
6. E. Dosekova, J. Filip, T. Bertok, P. Both, P. Kasak, and J. Tkac, *Med. Res. Rev.* 2017, **37**, 514–626.
7. Z. Qiao, J. Zhang, X. Hai, Y. Yan, W. Song and S. Bi, *Biosens. Bioelectron.* 2021, **176**, 112898.
8. Y. Yan, P. Shi, W. Song and S. Bi, *Theranostics*, 2019, **9**, 4047-4065.
9. D. Albesa-Jove and M. E. Guerin, *Curr. Opin. Struct. Biol.* 2016, **40**, 23–32.
10. A. Gimeno, P. Valverde, A. Arda and J. Jimenez-Barbero, *Curr. Opin. Struct. Biol.*, 2020, **62**, 22-30.
11. R. Hurtado-Guerrero, and G. J. Davies, *Curr. Opin. Chem. Biol.*, 2012, **16**, 479-487.
12. C. Breton, S. Fournel-Gigleux and M. M. Palcic, *Curr. Opin. Struct. Biol.*, 2012, **22**, 54-549.
13. Y. Zhao, M. Takahashi, J. Gu, E. Miyoshi, A. Matsumoto, S. Kitazume and N. Taniguchi, *Cancer Sci.*, 2008, **99**, 1304-1310.
14. I. Bucior and M. M. Burger, *Curr. Opin. Struct. Biol.*, 2004, **14**, 631-637.
15. P. Scheiffele, *Annu. Rev. Neurosci.*, 2003, **26**, 485–508.
16. B. K. Brandley and R. L. Schnaar, *J. Leukoc. Biol.*, 1986, **40**, 97-111.
17. A. J. Hayes and J. Melrose, *Adv. Therap.*, 2020, **3**, 2000151.
18. D. S. Costa, R. L. Reis, and I. Pashkuleva, *Annu. Rev. Biomed. Eng.*, 2017, **19**, 1–26.
19. M. Mende, C. Bednarek, M. Wawryszyn, P. Sauter, M. B. Biskup, U. Schepers and S. Brase, *Chem. Rev.*, 2016, **116**, 8193–8255.
20. A. Köwitsch, G. Zhou and T. Groth, *J Tissue Eng Regen Med.*, 2018, **12**, 23–41.

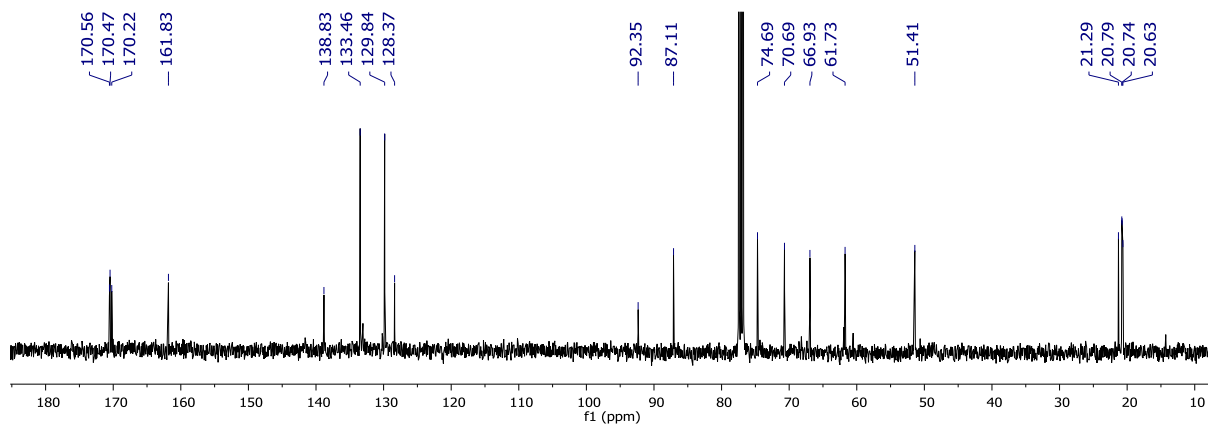
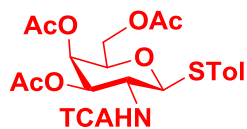
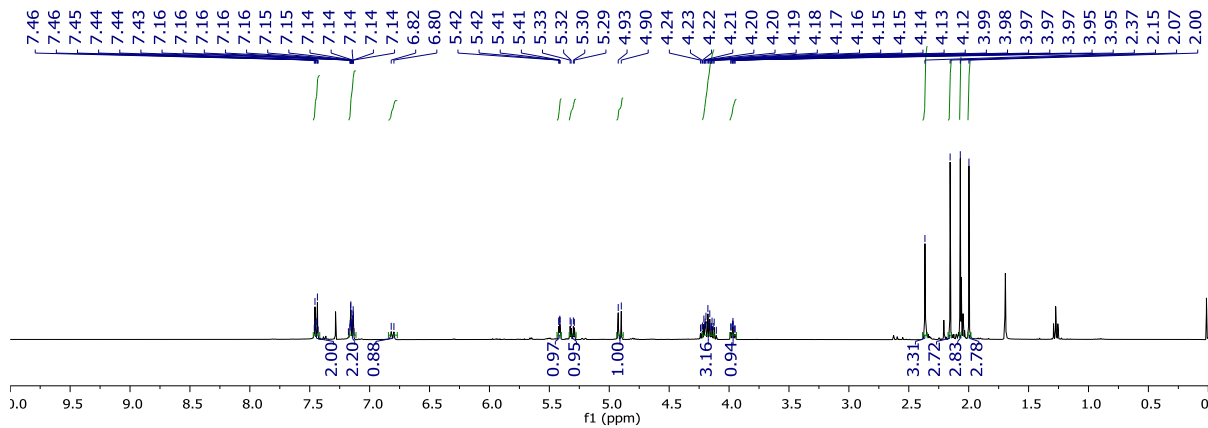
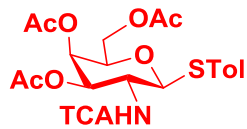
21. B. L. Farrugia, M. S. Lord, J. M. Whitelock and J. Melrose, *Biomater. Sci.*, 2018, **6**, 947–957.
22. S. Mishra and M. Ganguli, *Drug Discov. Today*, 2021, **26**, 1185-1199.
23. C. J. Handley, T. Samiric and M. Z. Ilic, *Adv. Pharmacol.*, 2006, **53**, 220-232.
24. M. Bishnoi, A. Jain, P. Hurkat and S. K. Jain, *Glycoconj. J.*, 2016, **33**, 693–705.
25. P. Souich, A. G. García, Z. Vergés and E. Montell, *J. Cell. Mol. Med.*, 2009, **13**, 1451–1463.
26. P. A. Levene and F. B. L. Forge, *Scientific proceedings*, 1914, **58**, 124-125.
27. E. A. Davidson and W. Small, *Biochim. Biophys. Acta.*, 1963, **69**, 459-463.
28. M. T. Bayliss, D. Osborne, S. Woodhouse and C. Davidson, *J. Biol. Chem.*, 1999, **274**, 15892-15900.
29. C. E. Argoff, *Current Pain and Headache Reports.*, 2004, **8**, 199–204.
30. Y. Henrotin, M. Mathy, C. Sanchez and C. Lambert, *Ther Adv Musculoskel Dis.*, 2010, **2**, 335348.
31. V. B. Krylov, A. A. Grachev, N. E. Ustyuzhanina, N. A. Ushakova, M. E. Preobrazhenskaya, N. I. Kozlova, M. N. Portsel, I. N. Konovalova, V. Yu. Novikov, H.C.. Siebert, A. S. Shashkov, and N. E. Nifantiev, *Russ. Chem. Bull. Int.Ed.*, 2011, **60**, 731-738.
32. T. Maruyama, T. Toida, T. Imanari, G. Yu, R. J. Linhardt, *Carbohydr. Res.*, 1998, **306**, 35-43.
33. S. Chen, C. Xue, L. Yin, Q. Tang, G. Yu and W. Chai, *Carbohydr. Polym.*, 2011, **83**, 688-696.
34. V. H. Pomin, *Mar. Drugs*, 2014, **12**, 232-254.
35. N. M. Mestechkina and V. D. Shcherbukhin, *Appl. Biochem. Microbiol.*, 2010, **46**, 267-273.
36. N.P. Aditya, P.G. Vathsala, V. Vieira, R.S.R. Murthy and E.B. Souto, *Adv. Colloid Interface Sci.*, 2013, **201-202**, 1-17.
37. N. J. White, *White Malar J.*, 2017, **16**, 88.
38. C. L. Mackintosh, J. G. Beeson and K. Marsh, *Trends Parasitol.*, 2004, **20**, 597-603.
39. K. Aït-Mohand, C. Lopin-Bon and J. Jacquinet, *Carbohydr. Res.*, 2012, **353**, 33-48.
40. S. Eller, M. Collot, J. Yin, H. S. Hahm, and P. H. Seeberger, *Angew. Chem.*

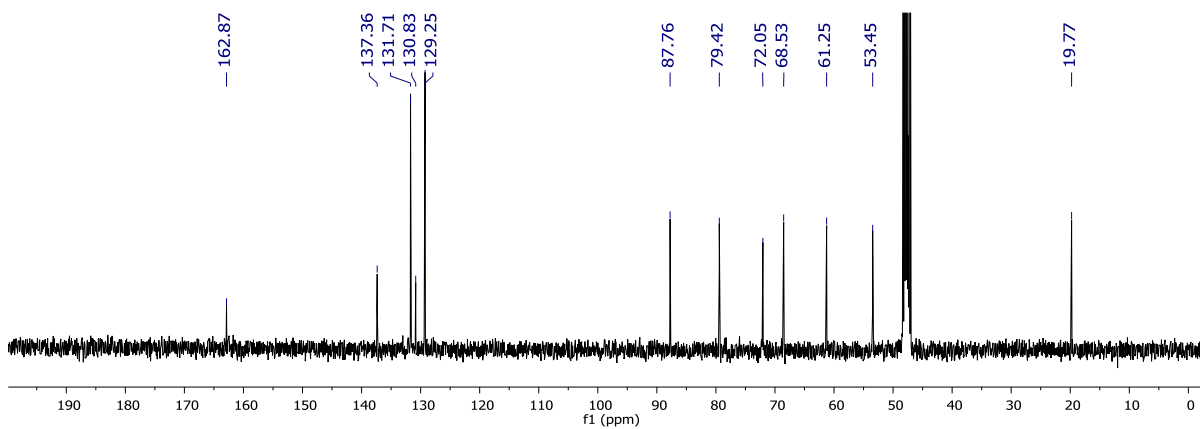
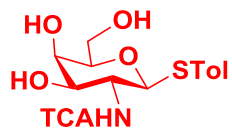
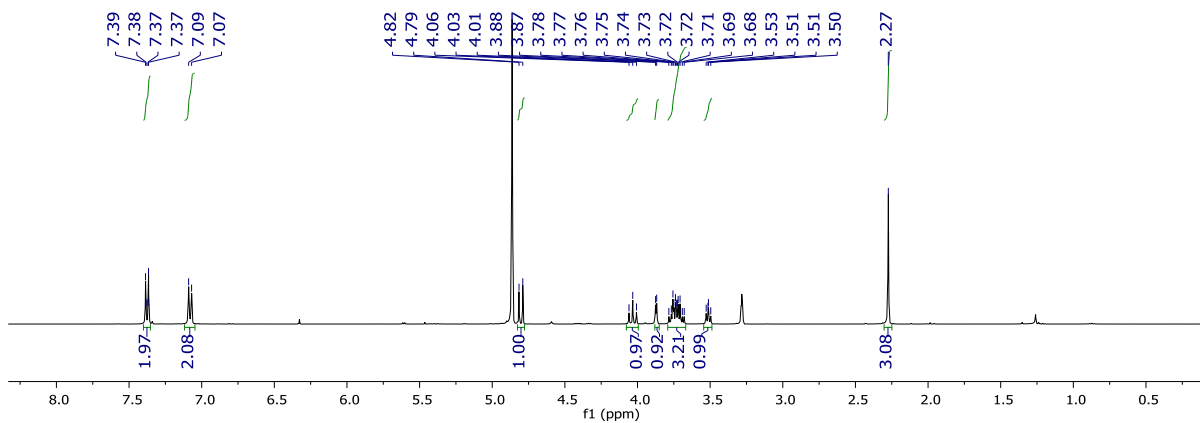
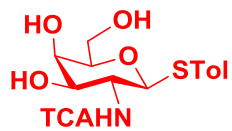
- Int. Ed.*, 2013, **52**, 5858–5861.
41. S. Maza, M. M. Kayser, G. Macchione, J. López-Prados, J. Angulo, J. L. Paz and P. M. Nieto, *Org. Biomol. Chem.*, 2013, **42**, 3510–3525.
  42. G. Macchione, Susana Maza, M. M. Kayser, J. L. Paz, and P. M. Nieto, *Eur. J. Org. Chem.*, 2014, 3868–3884.
  43. J. Jacquinet and C. Lopin-Bon, *Carbohydrate Res.*, 2015, **402**, 35–43.
  44. Z. W. Poh, C. H. Gan, E. J. Lee, S. Guo, G. W. Yip and Y. Lam, *Sci. Rep.*, 2015, **5**, 14355.
  45. S. Ramadan, W. Yang, Z. Zhang, and X. Huang, *Org. Lett.*, 2017, **19**, 4838–4841.
  46. S. Jadhav, V. Gulumkar, P. Deshpande, E. T. Coffey, H Lonnberg, and P. Virta, *Bioconjugate Chem.*, 2018, **29**, 2382–2393.
  47. S. Yang, Q. Liu, G. Zhang, X. Zhang, Z. Zhao and P. Lei, *J. Org. Chem.*, 2018, **83**, 5897–5908.
  48. L. Zhang, P. Xu, B. Liu, and Biao Yu, *J. Org. Chem.*, 2020, **85**, 15908–15919.
  49. S. Lee, J. M. Brown, C. J. Rogers, J. B. Matson, C. Krishnamurthy, M. Rawat and L. C. Hsieh-Wilson, *Chem. Sci.*, 2010, **1**, 322–325.
  50. M. Rawat, C. I. Gama, J. B. Matson, and L. C. Hsieh-Wilson, *J. Am. Chem. Soc.* 2008, **130**, 2959–2961.
  51. C. I. Gama, S. E. Tully, N. Sotogaku, P. M. Clark, M. Rawat, N. Vaidehi, W. A Goddard, A. Nishi and L. C. Hsieh-Wilson, *Nat. Chem. Biol.*, 2006, **2**, 467–473.
  52. S. E. Tully, R. Mabon, C. I. Gama, S. M. Tsai, X. Liu, and L. C. Hsieh-Wilson, *J. Am. Chem. Soc.* 2004, **126**, 7736–7737.
  53. S. Yang, A. Wang, G. Zhang, X. Di, Z. Zhao and P. Lei, *Tetrahedron*, 2016, **72**, 5659–5670.

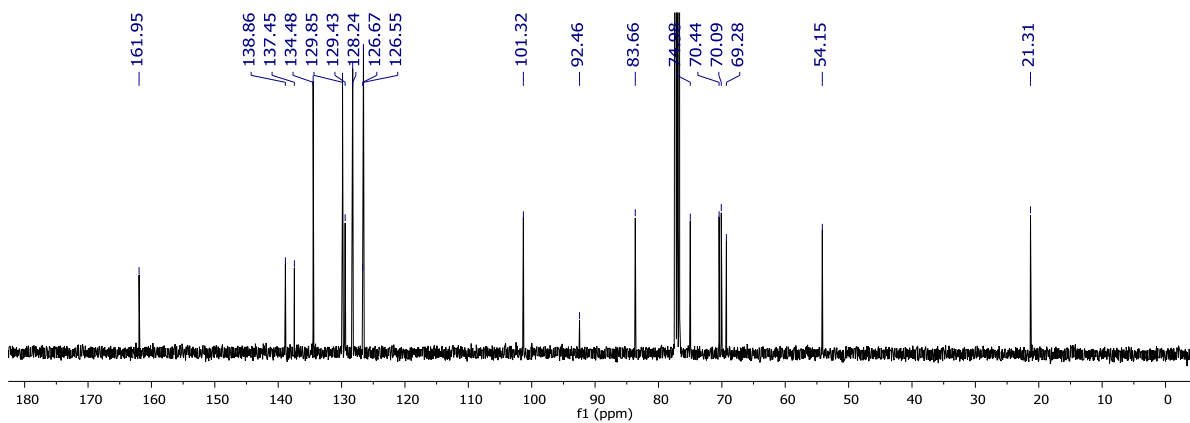
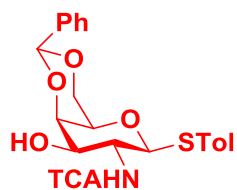
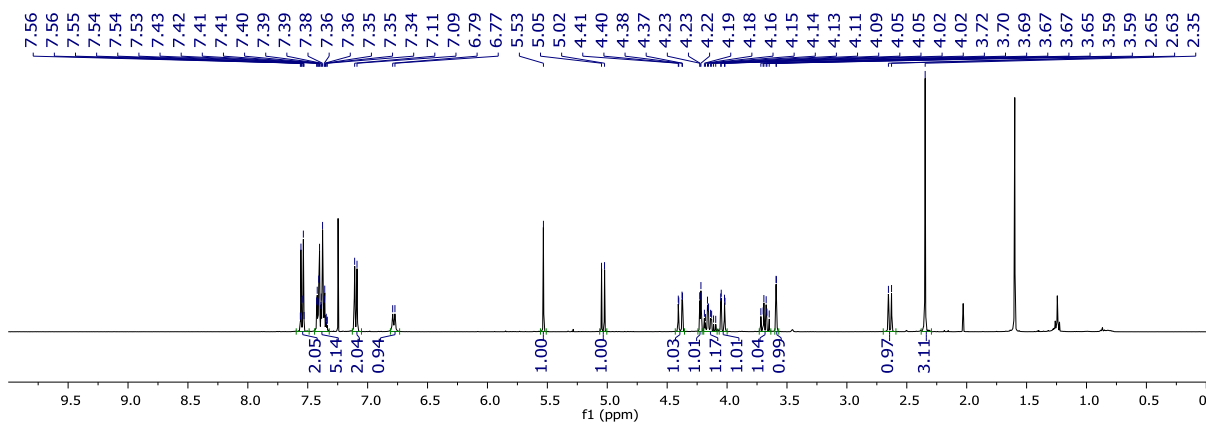
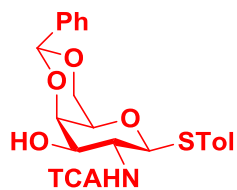


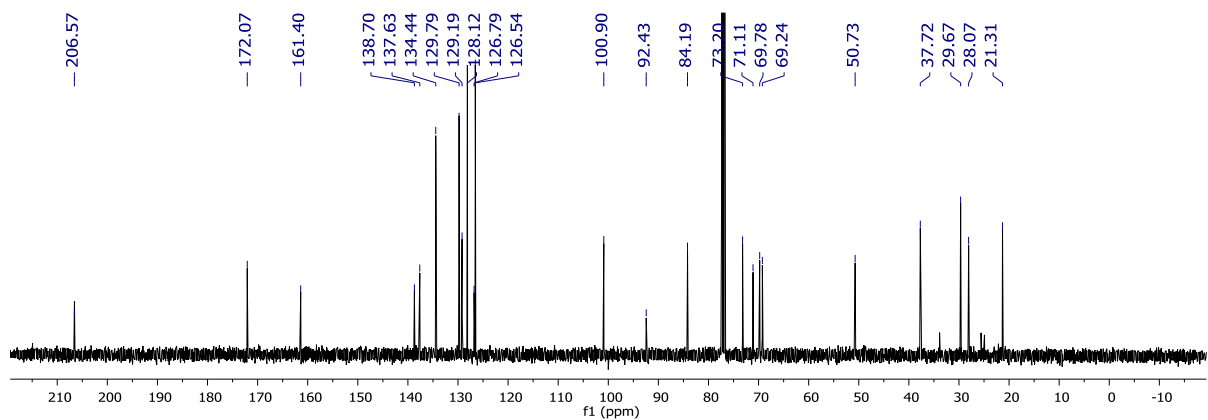
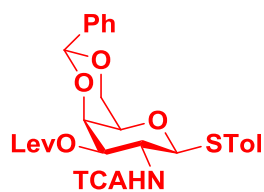
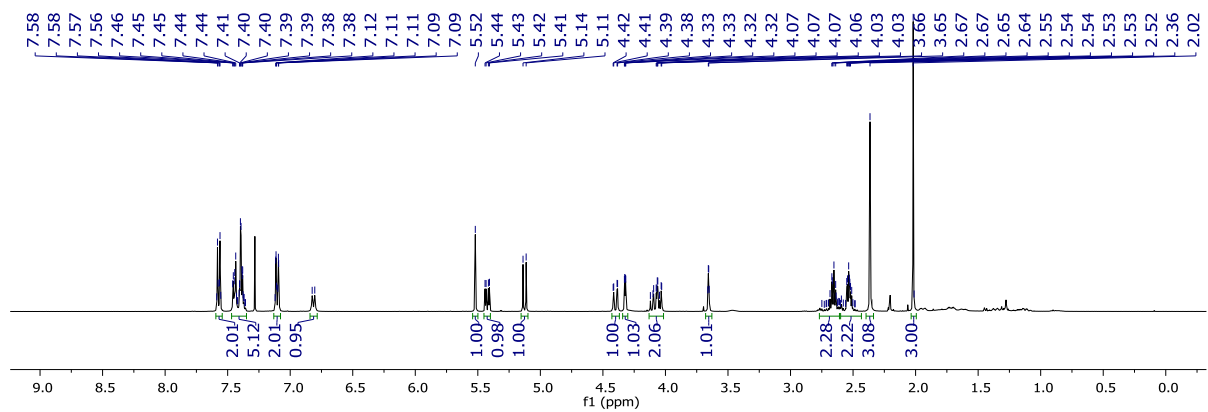
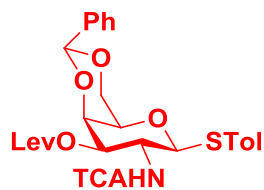
## 4a.10 NMR Spectra:

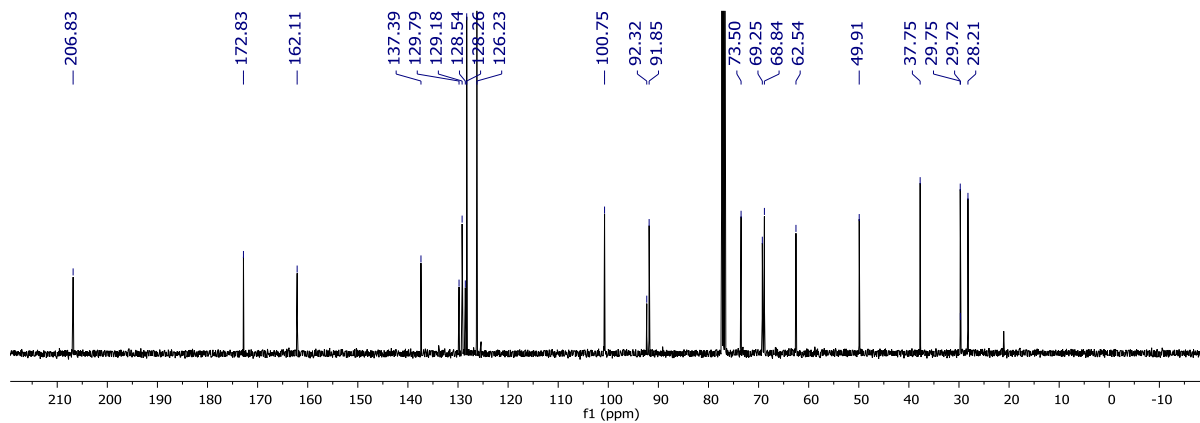
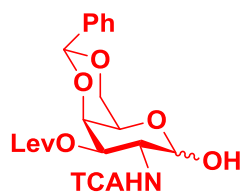
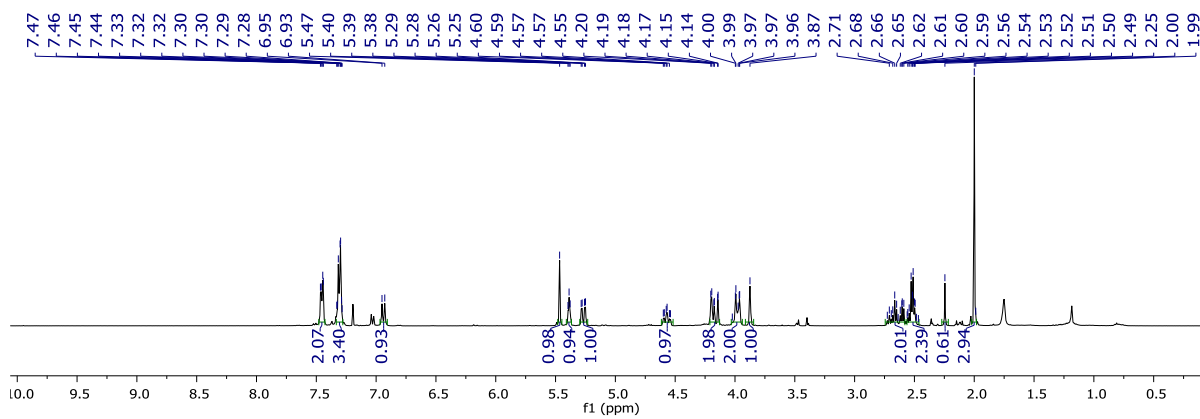
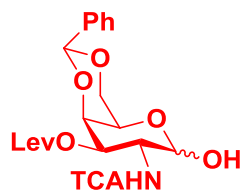


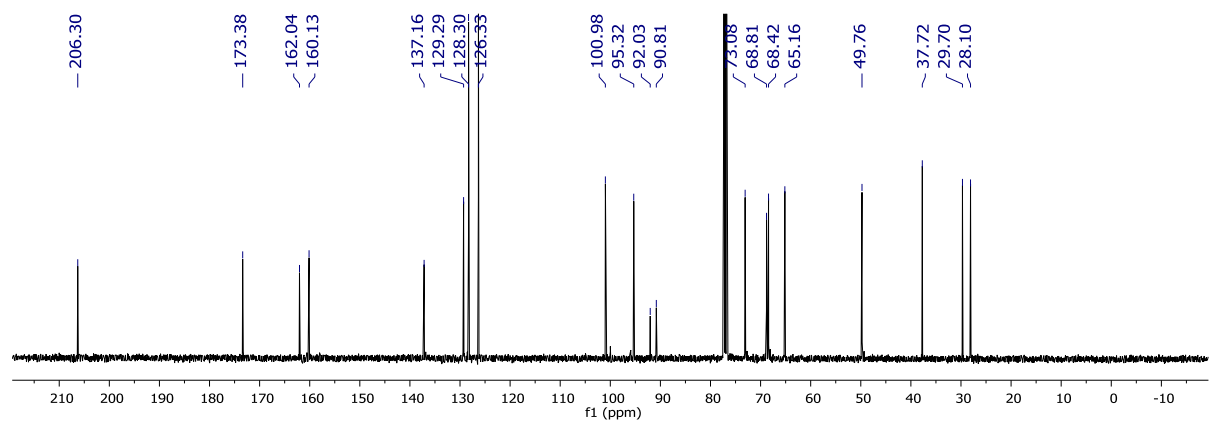
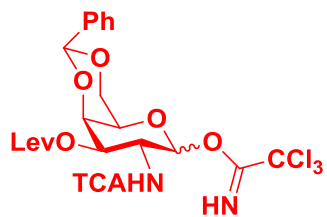
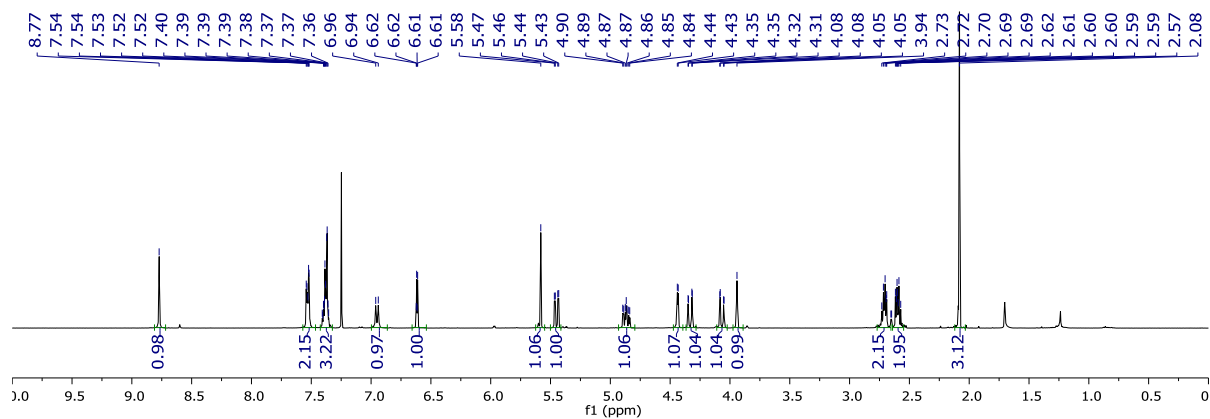
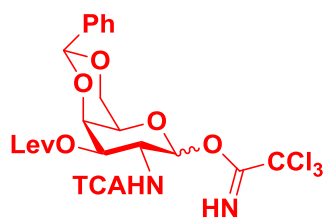


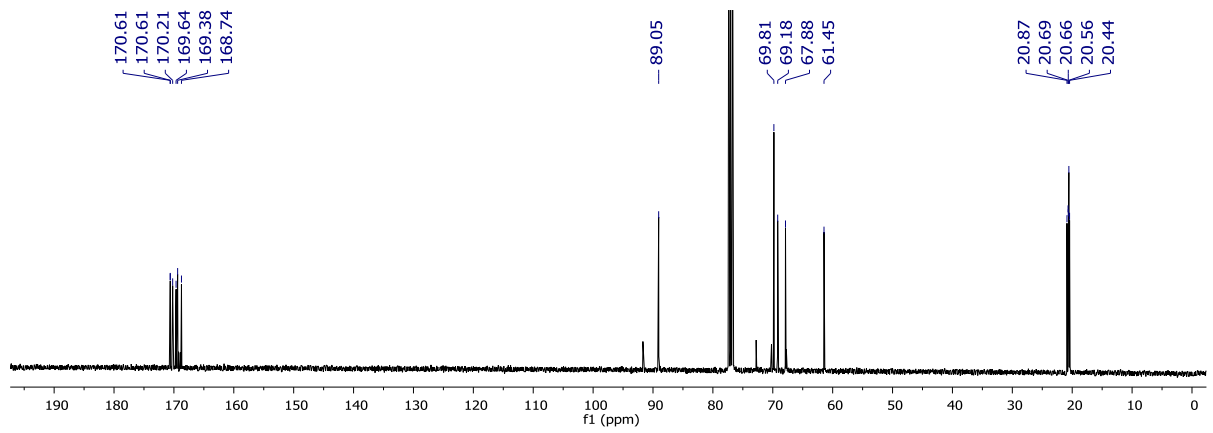
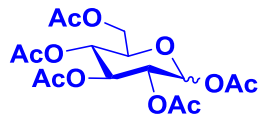
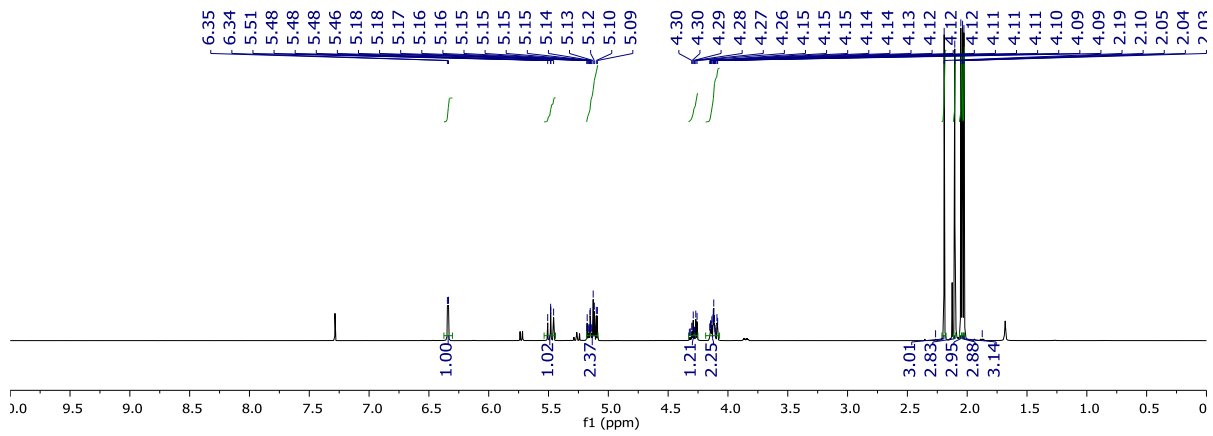
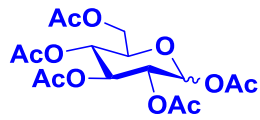




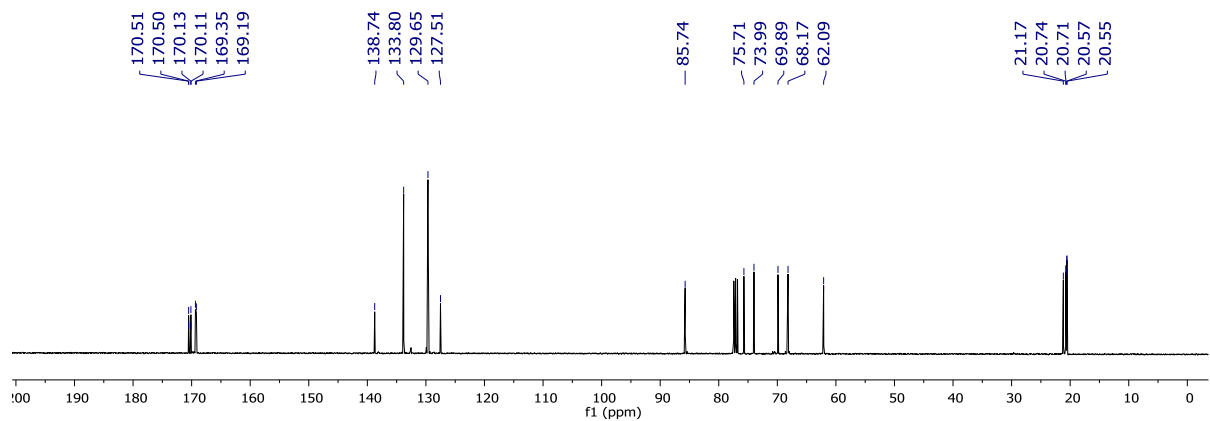
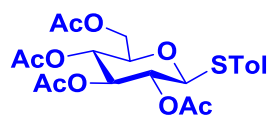
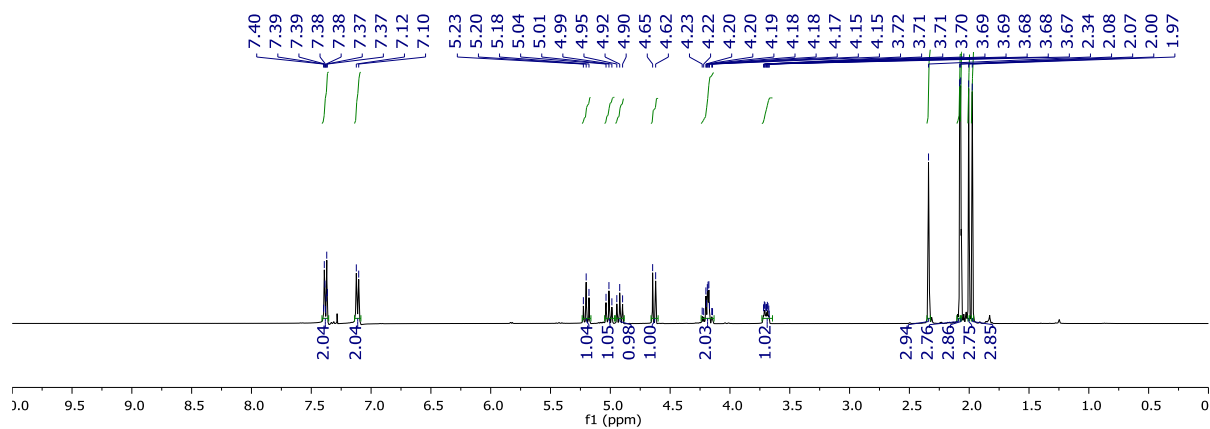
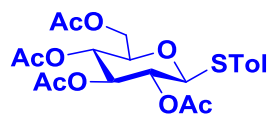


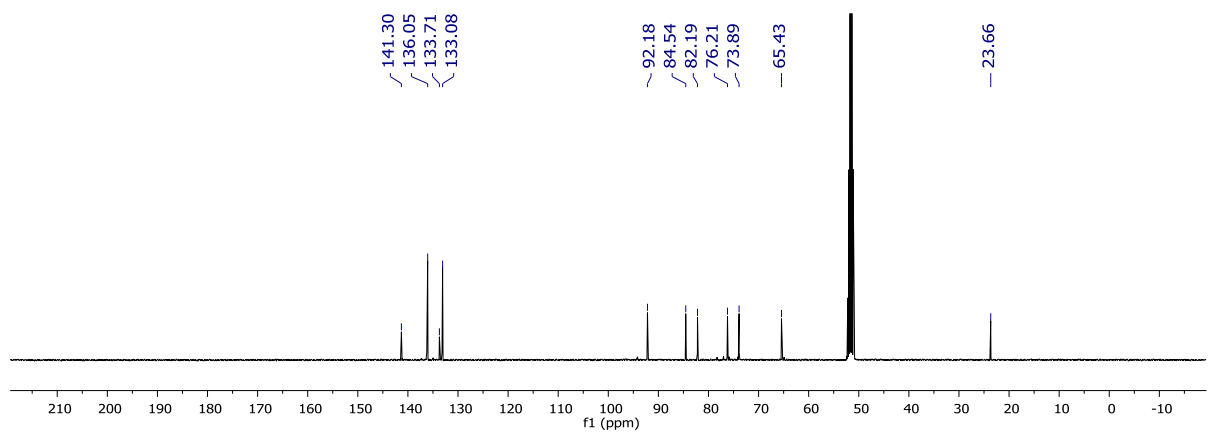
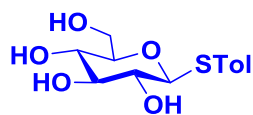
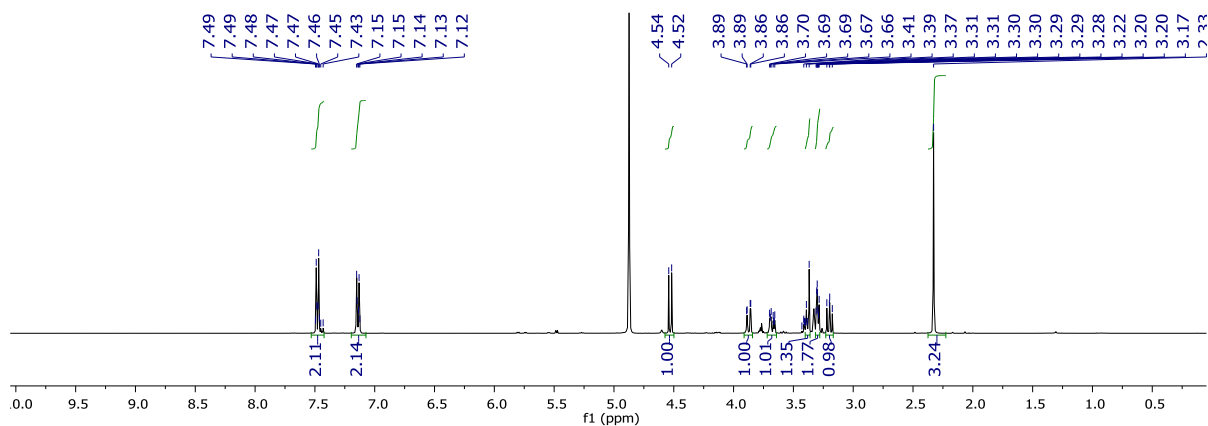
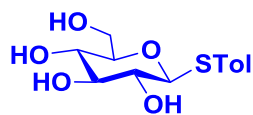


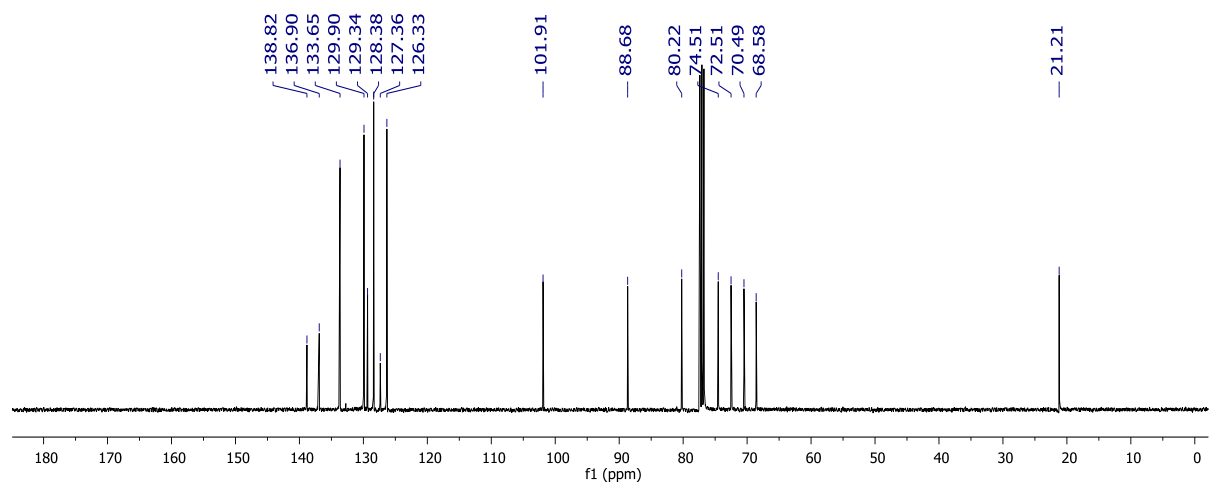
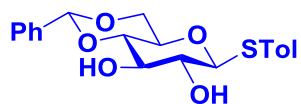
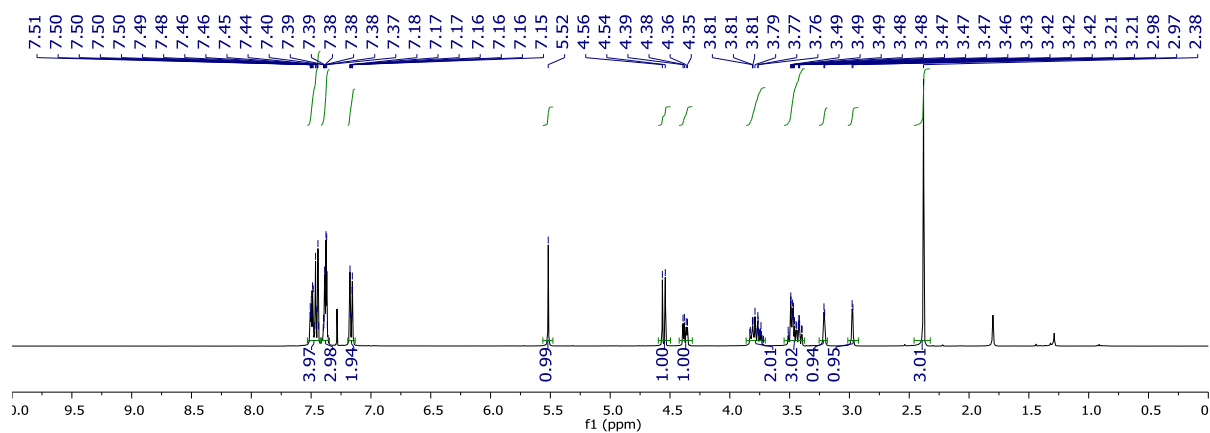
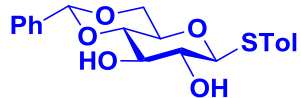


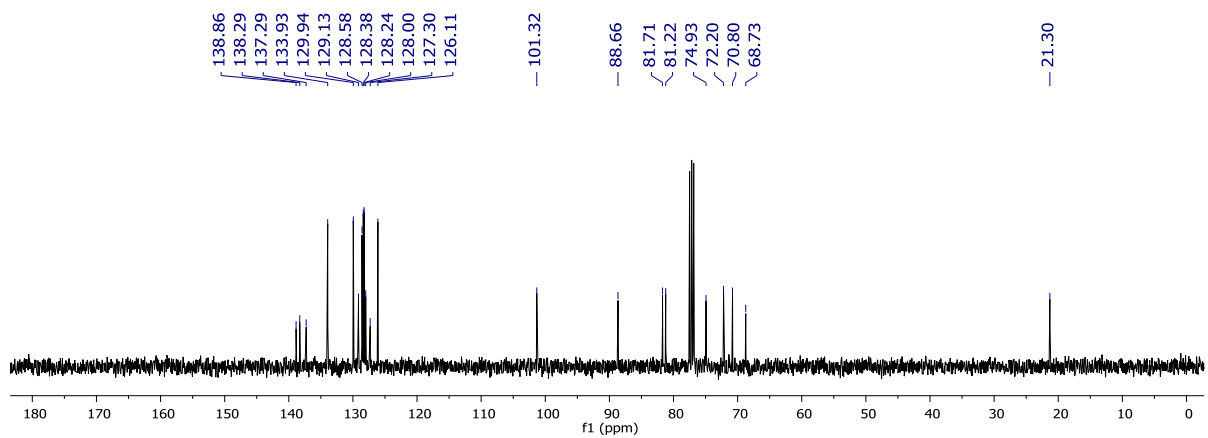
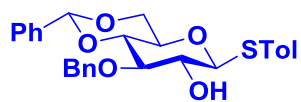
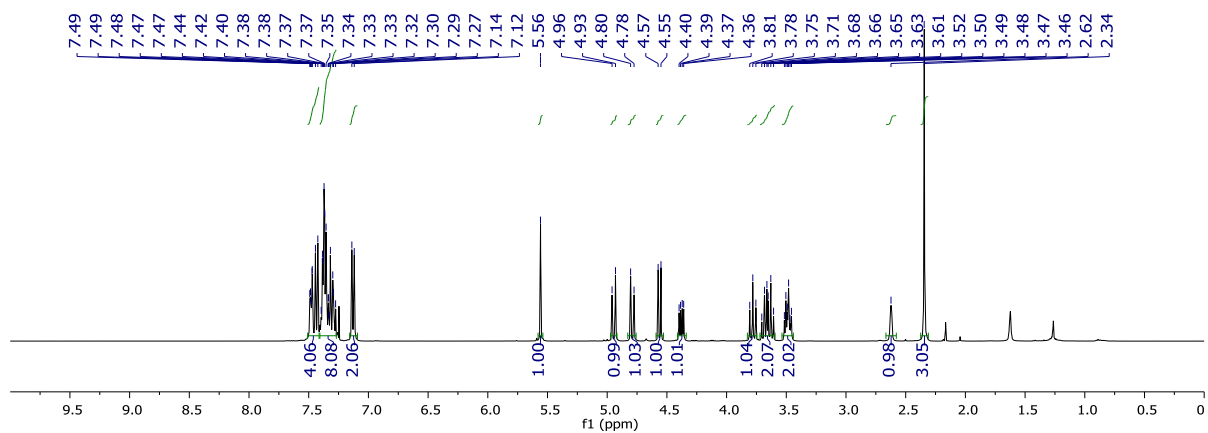
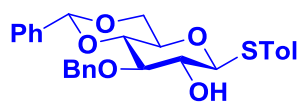


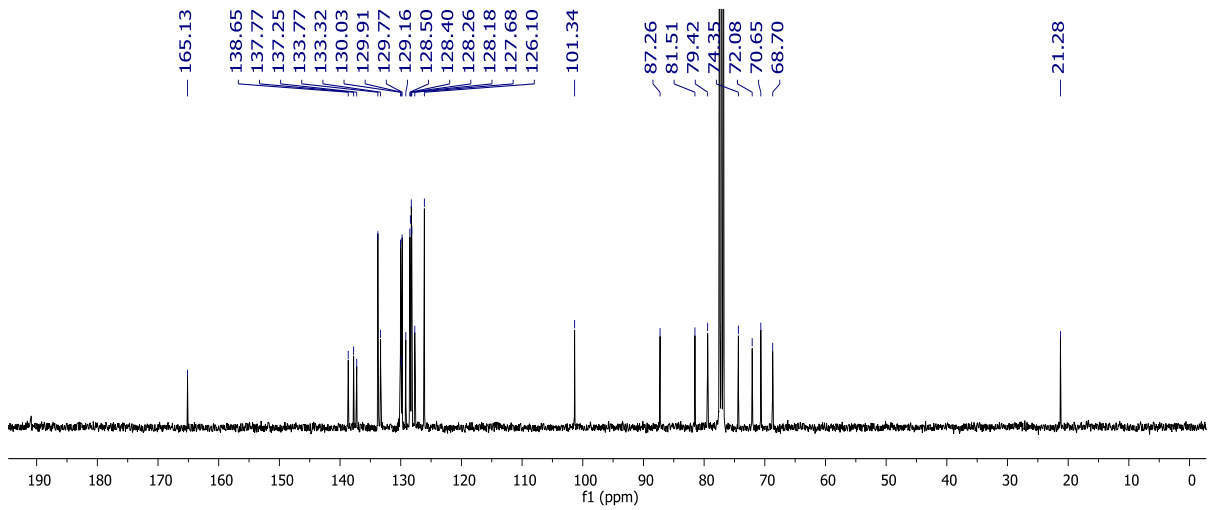
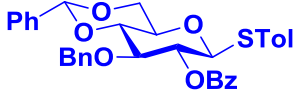
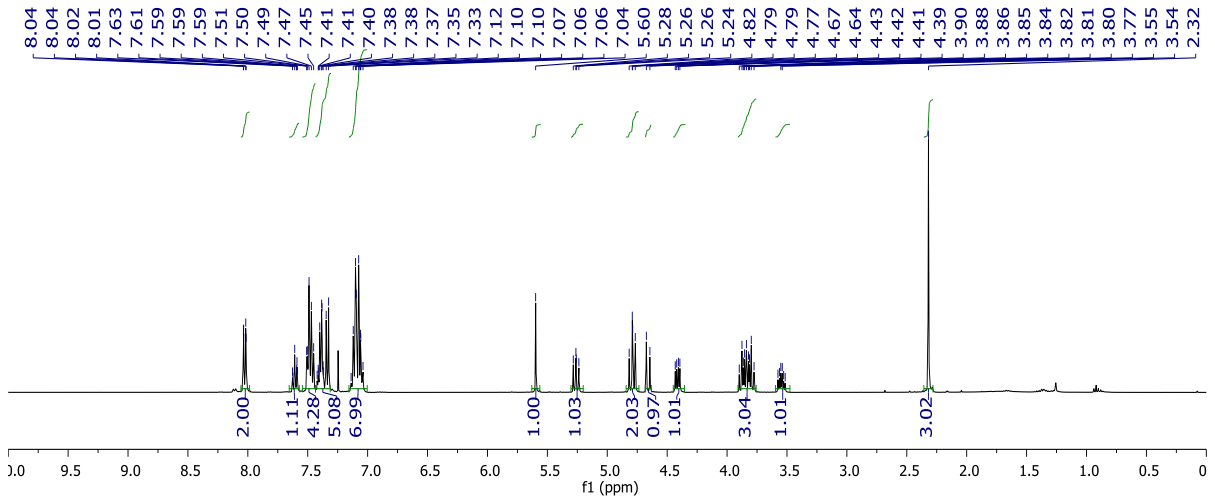
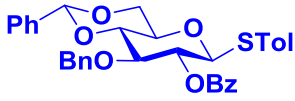


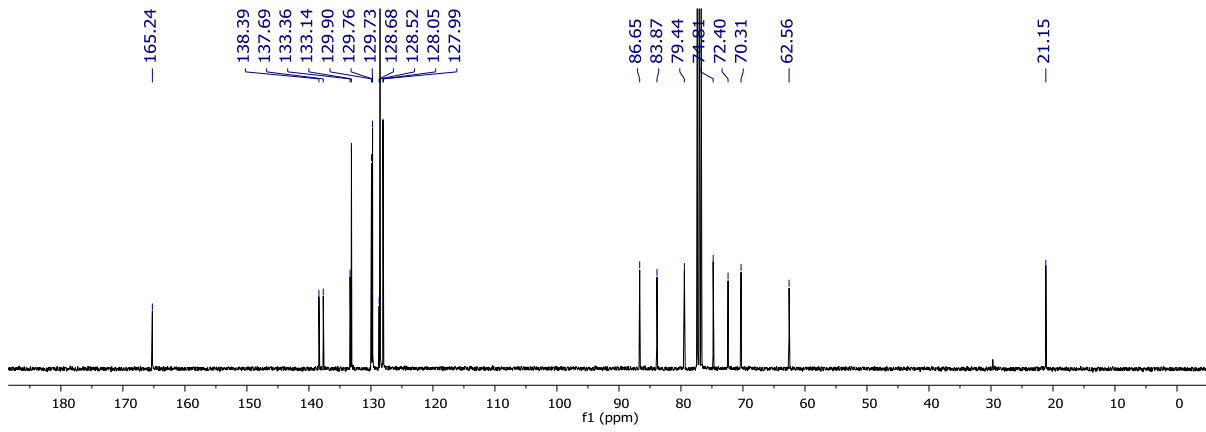
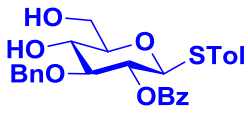
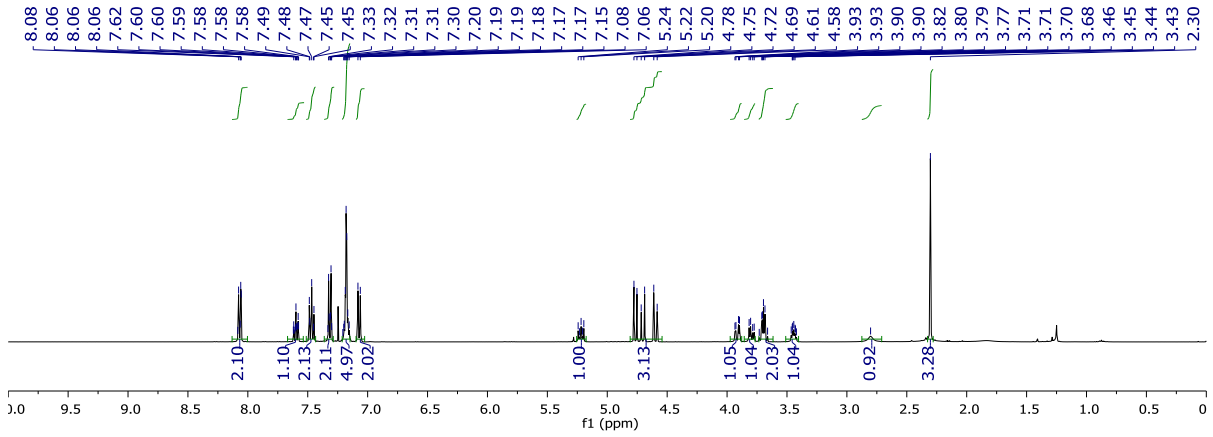
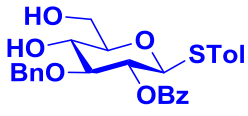


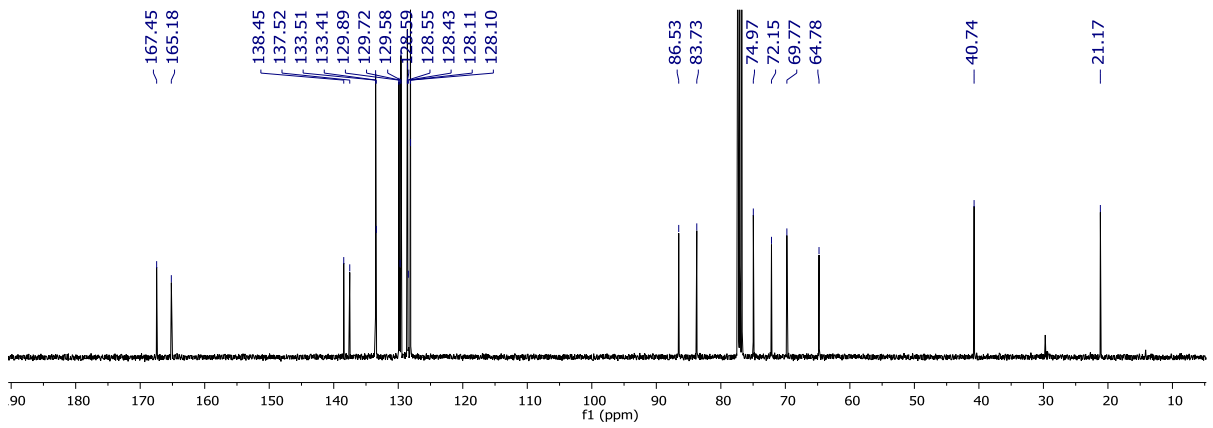
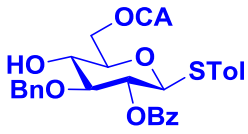
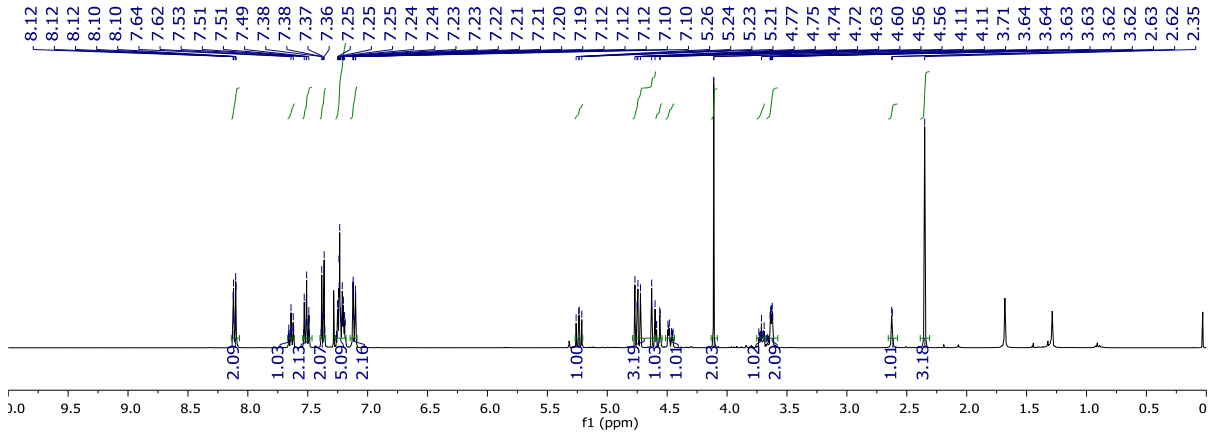
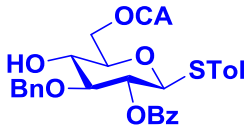


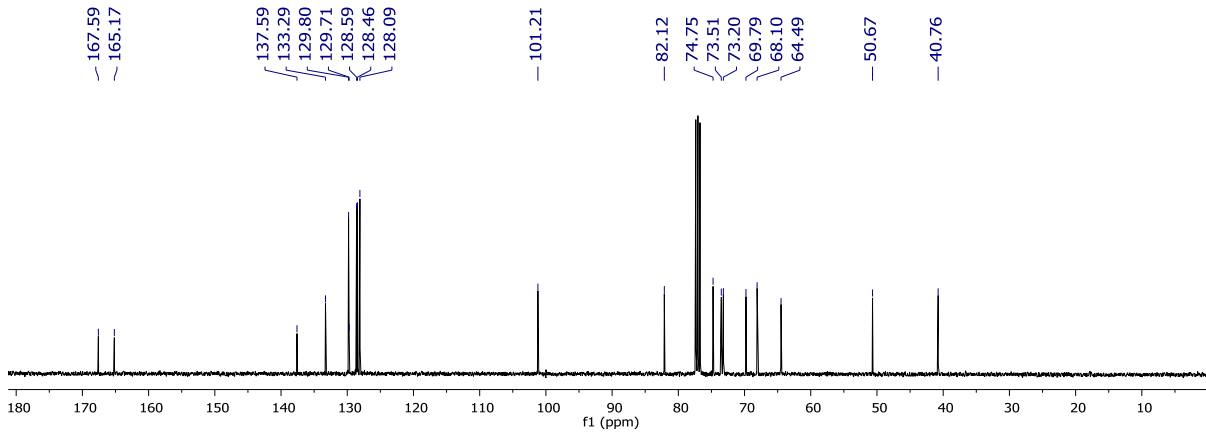
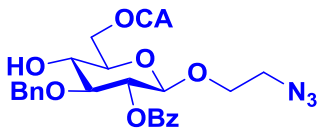
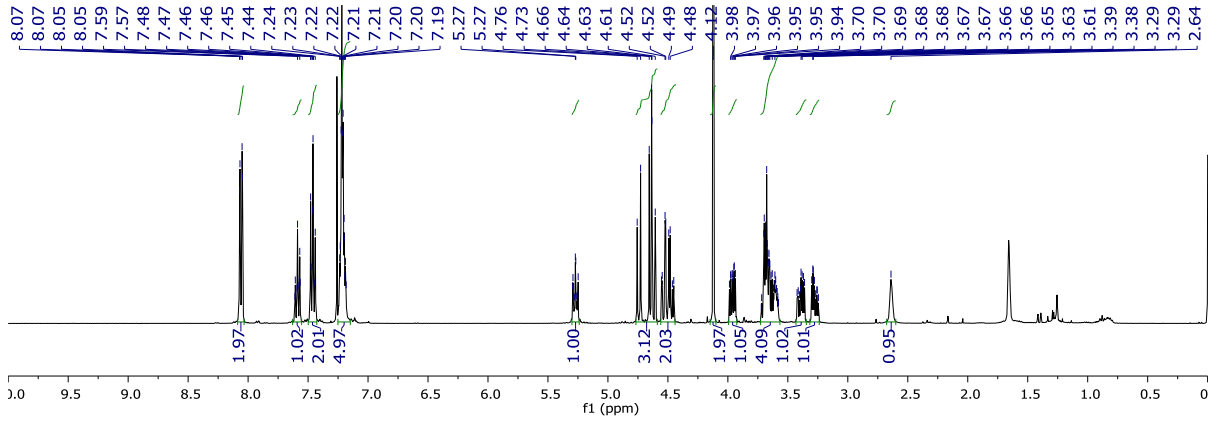
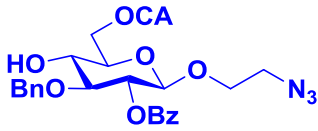




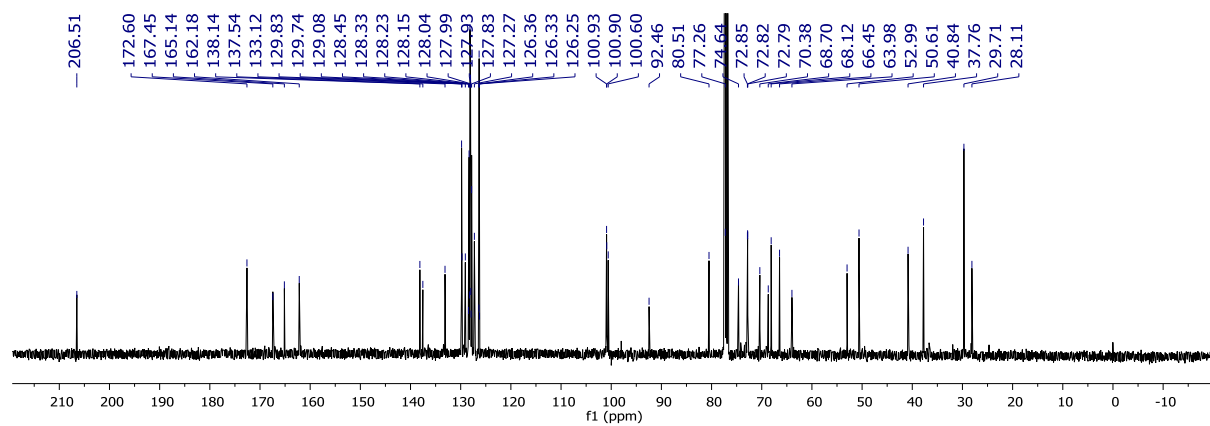
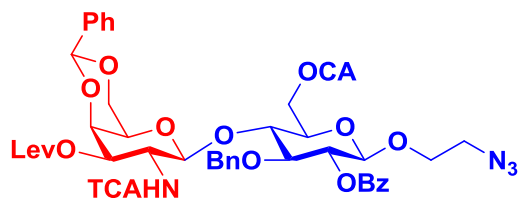
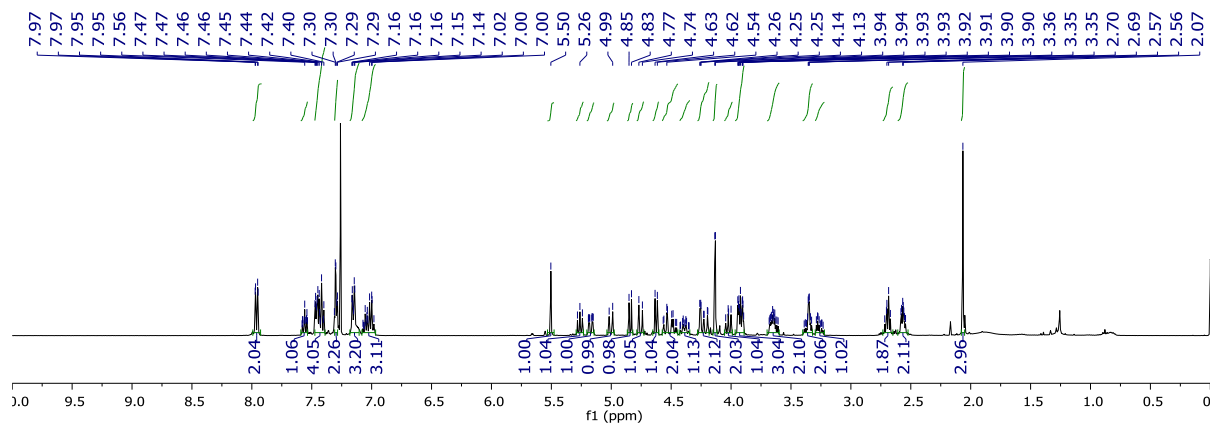
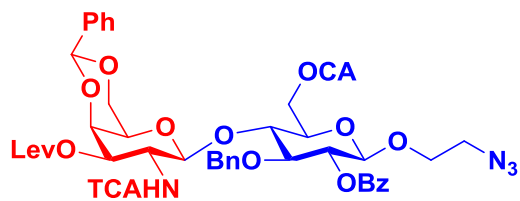


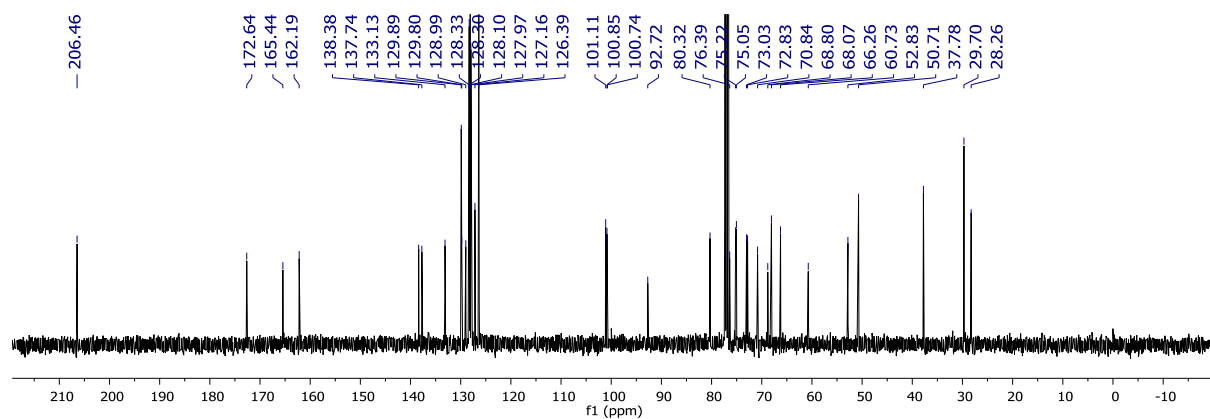
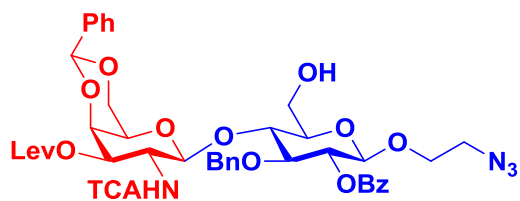
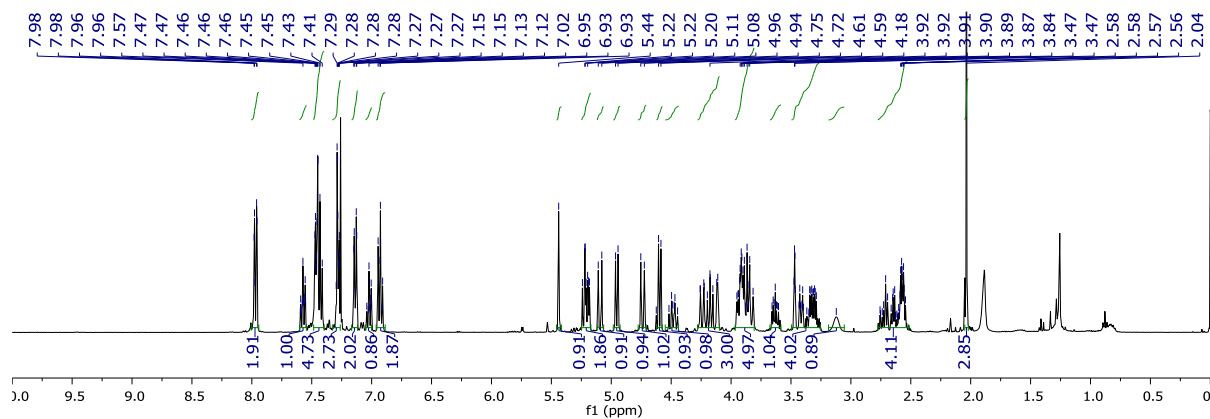
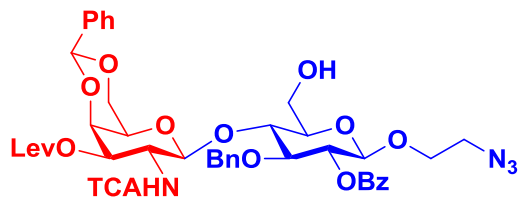


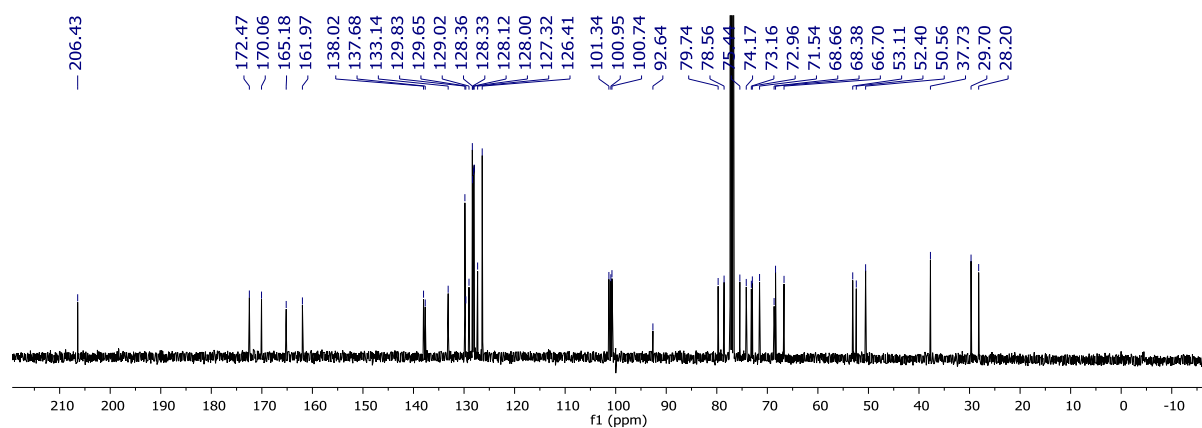
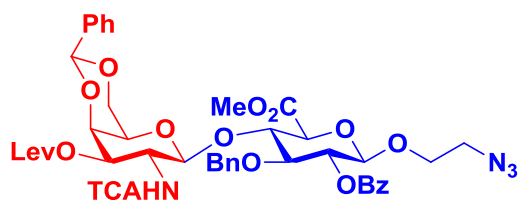
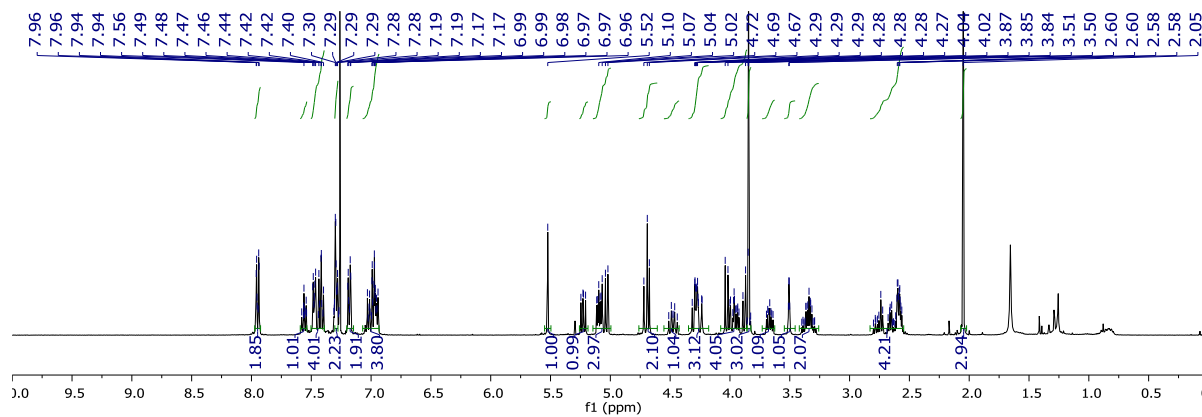
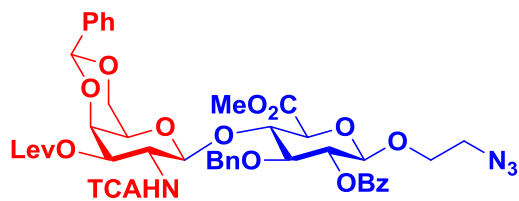


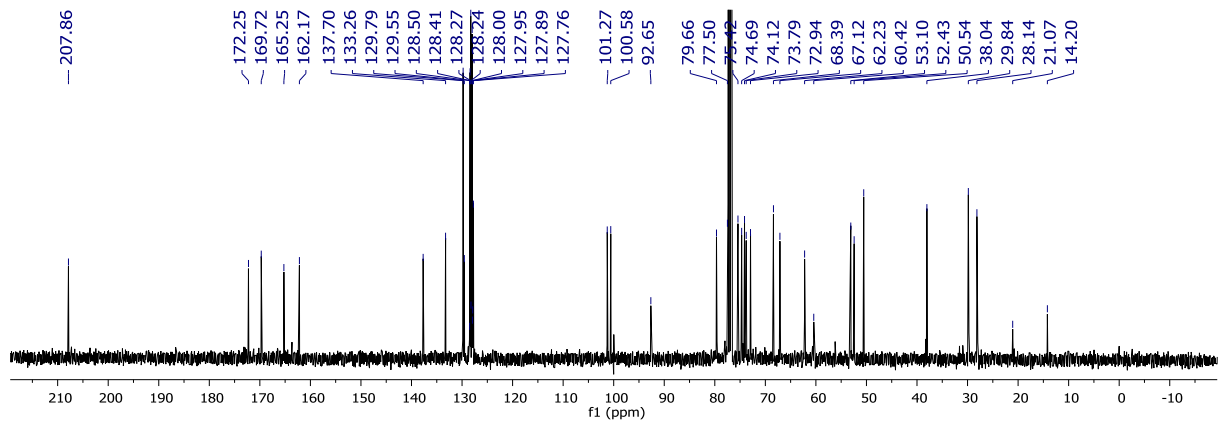
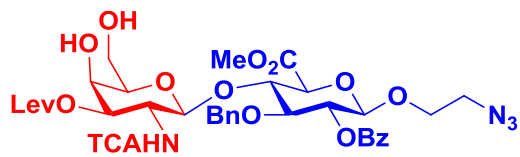
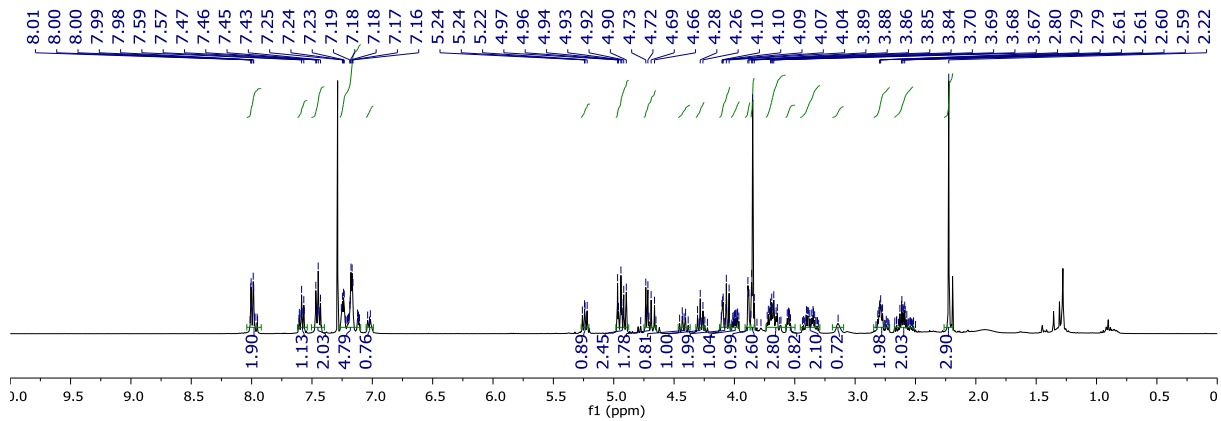
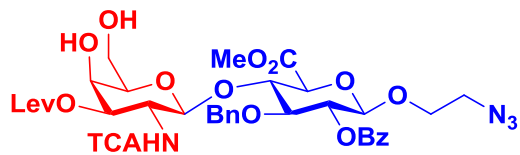


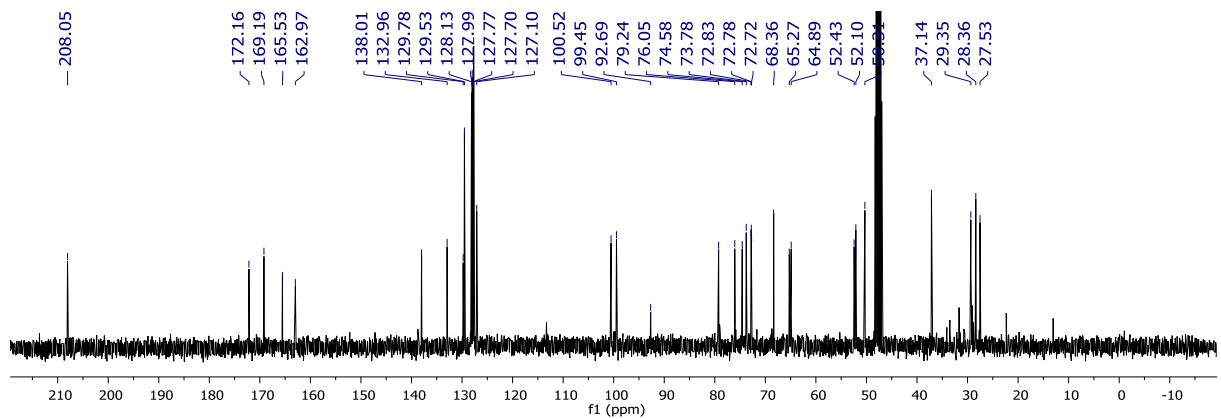
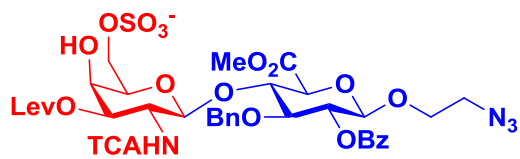
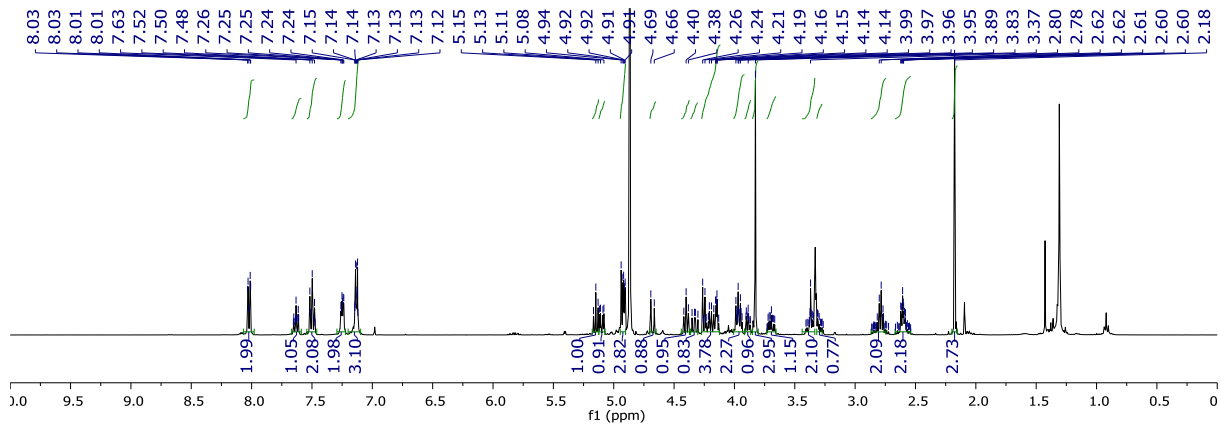
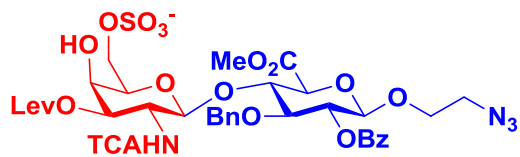


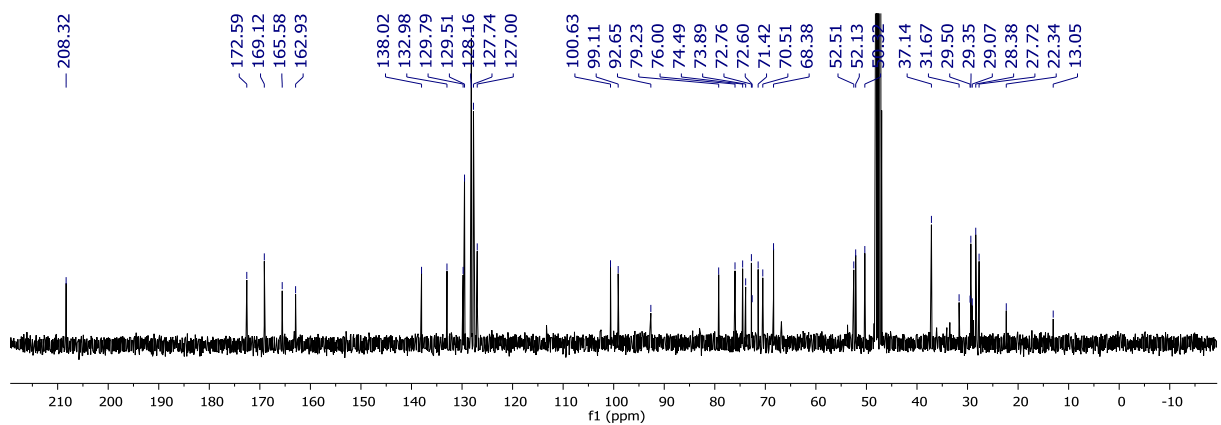
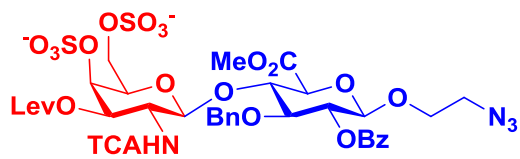
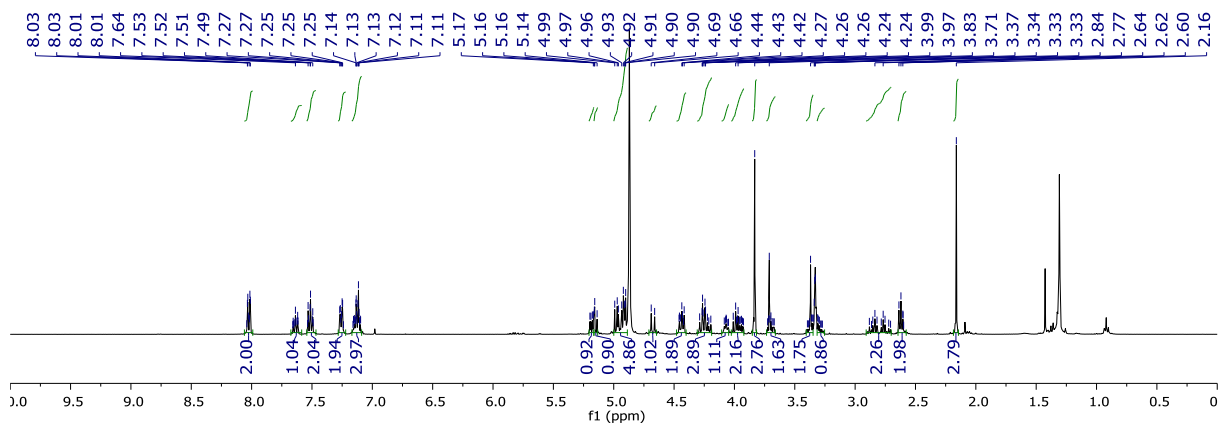
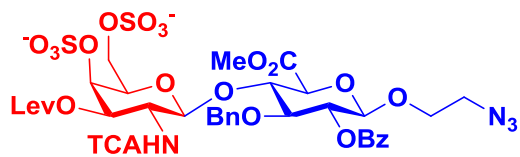


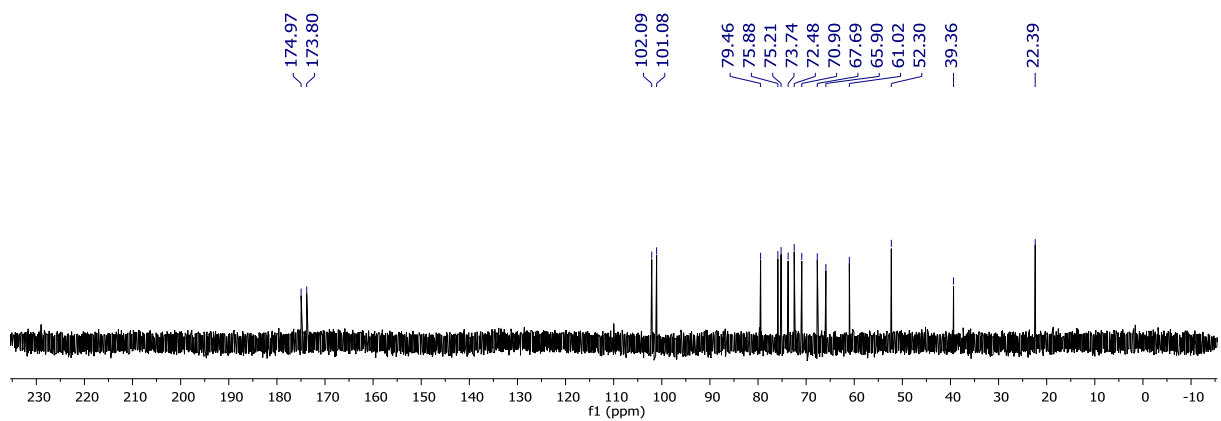
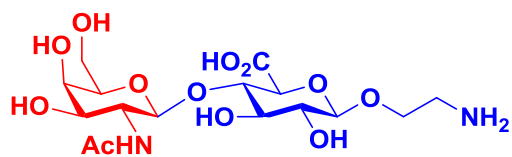
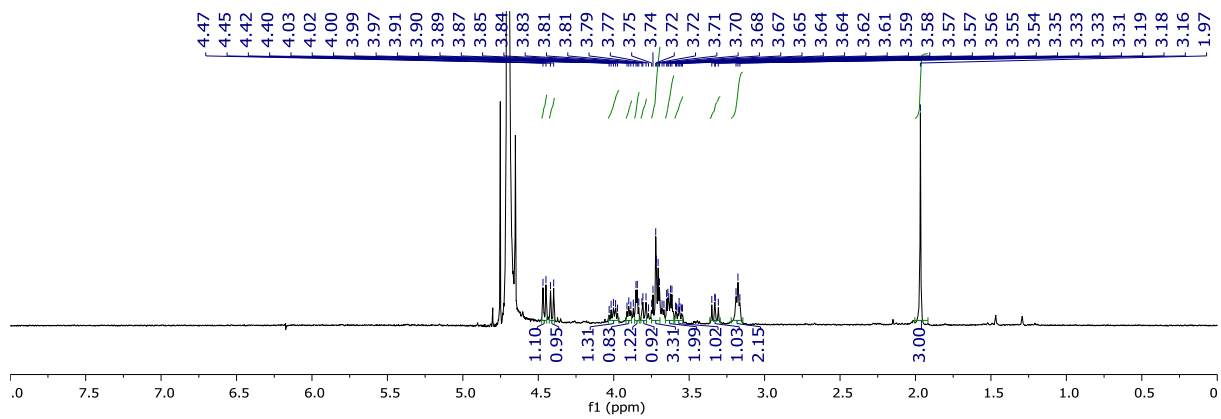
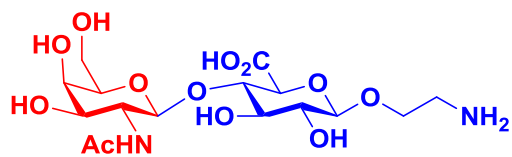


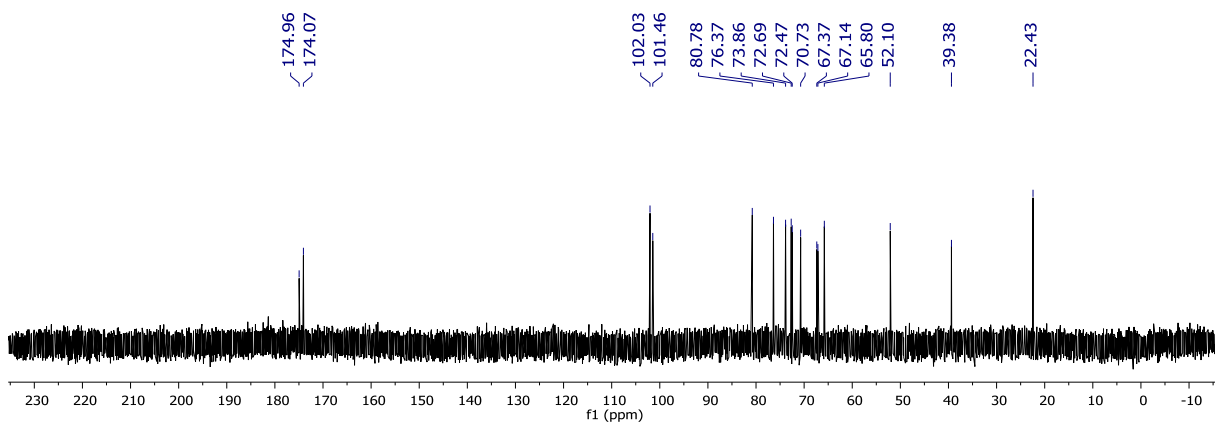
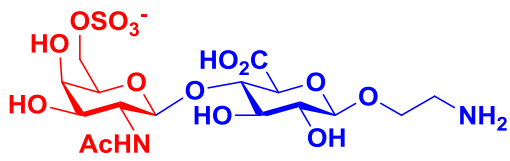
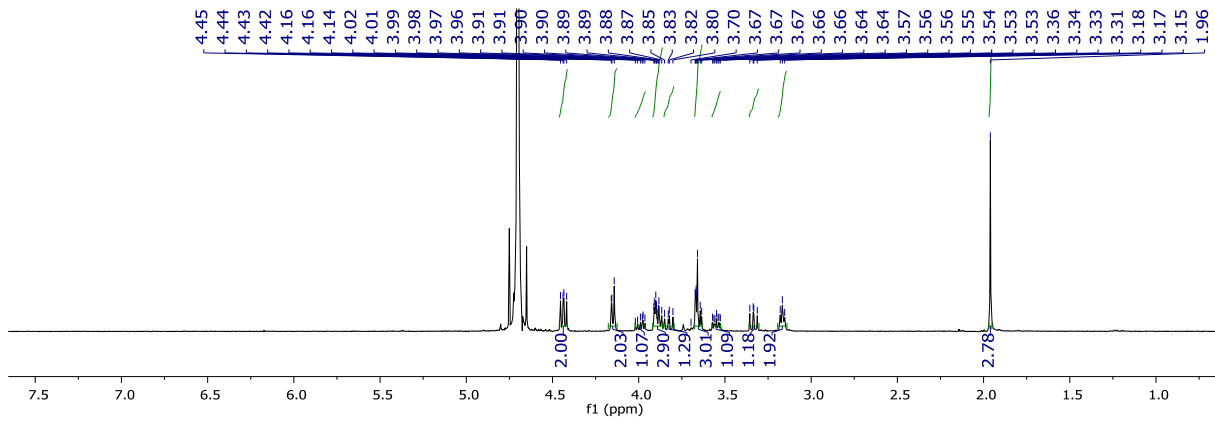
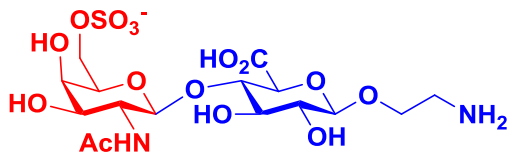




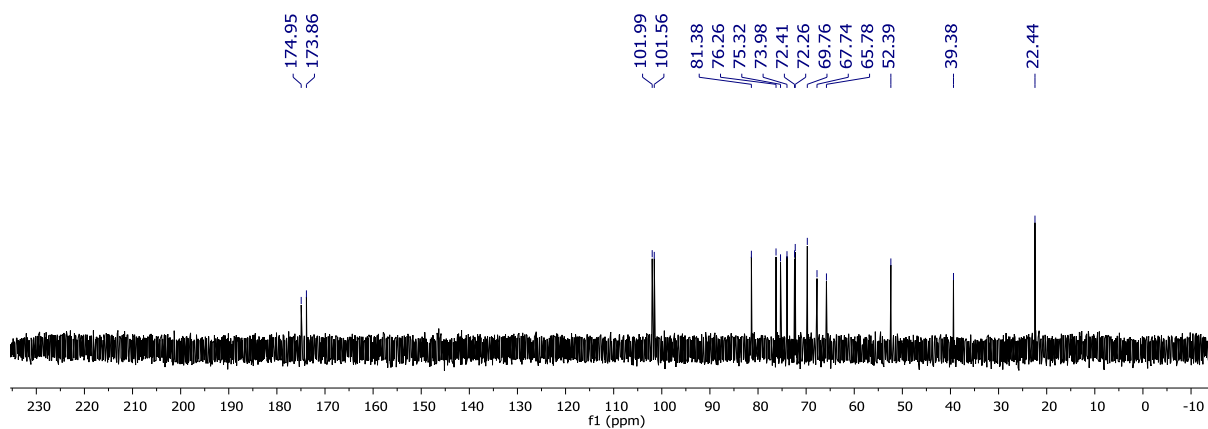
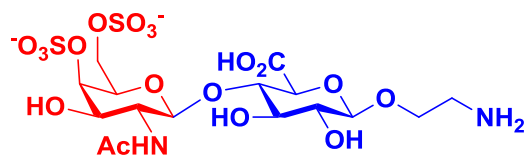
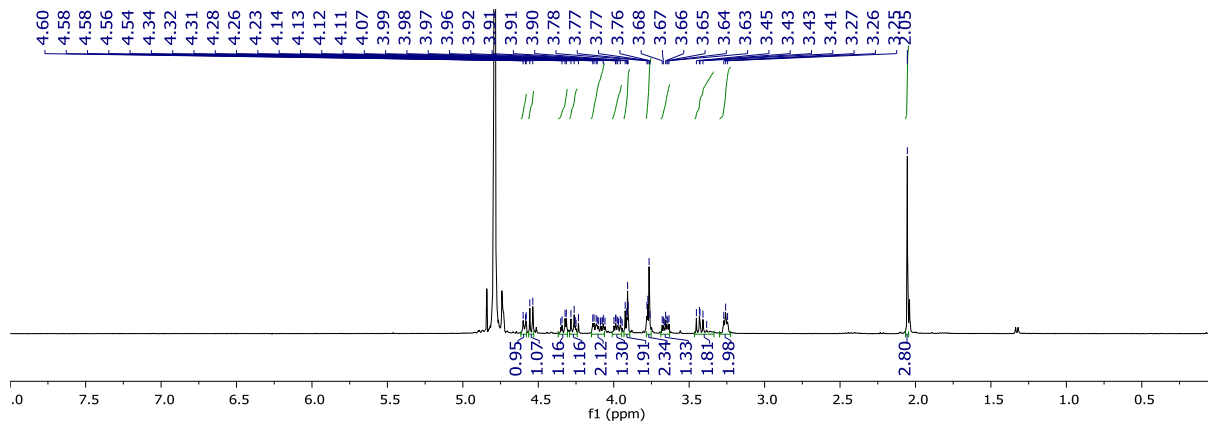
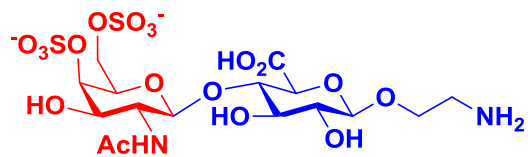


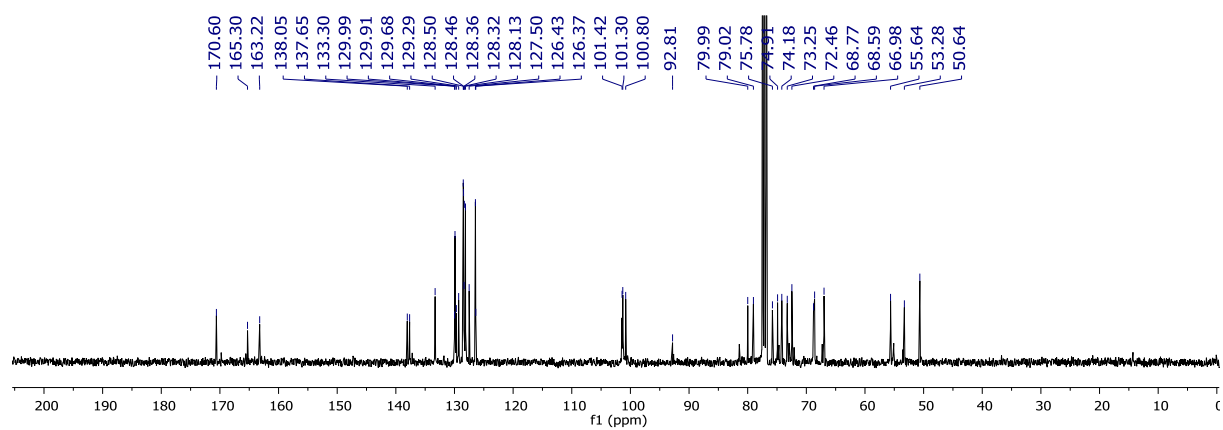
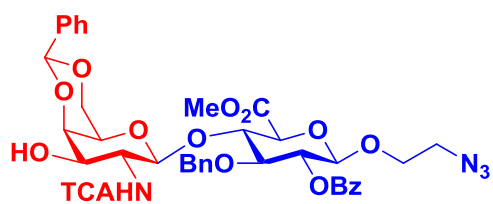
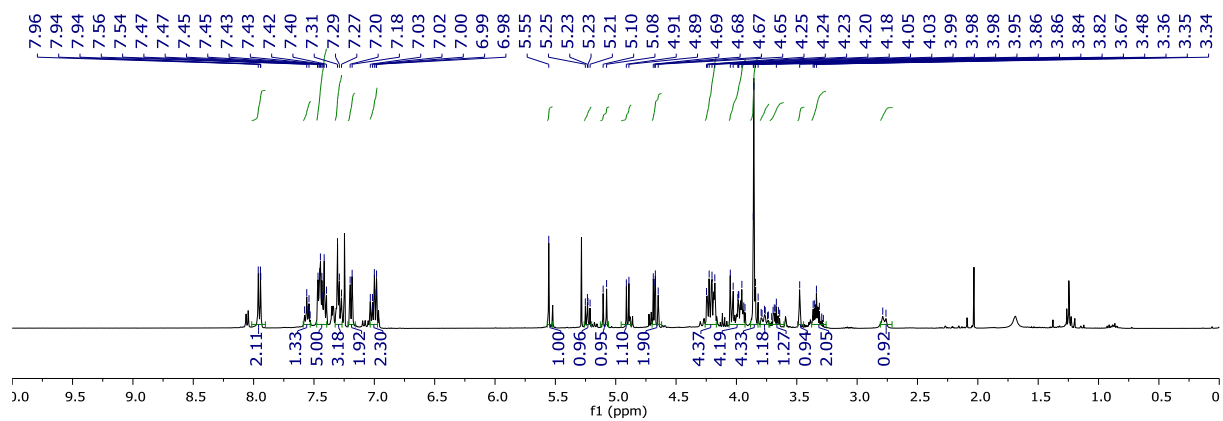
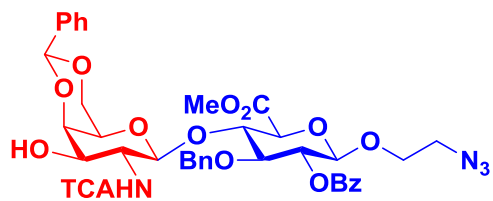


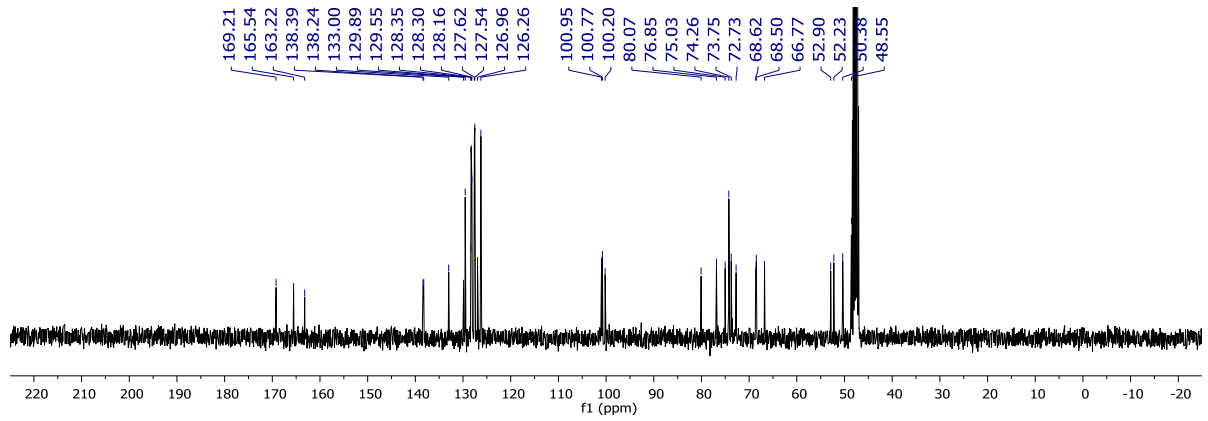
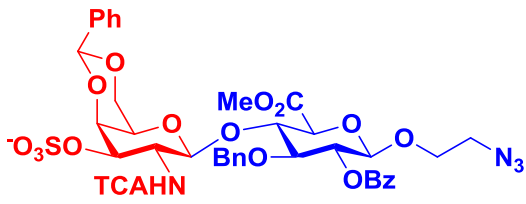
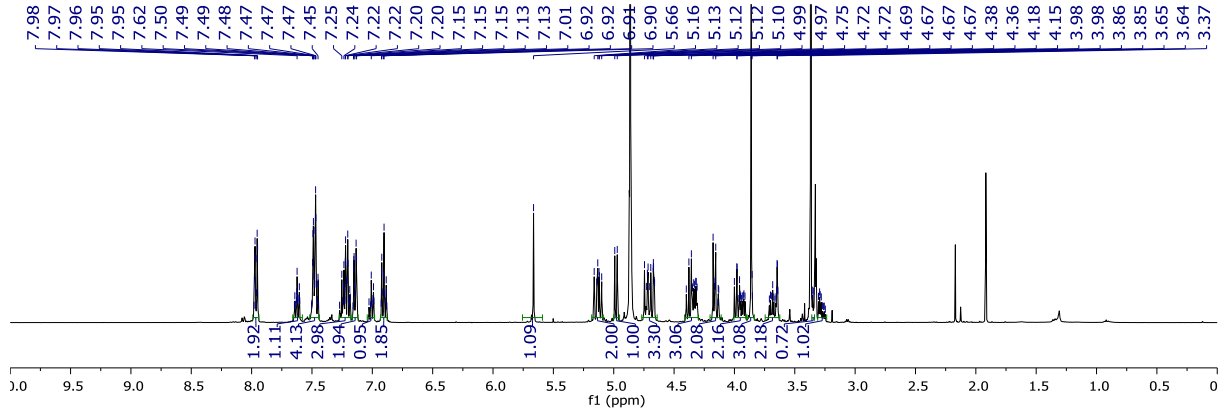
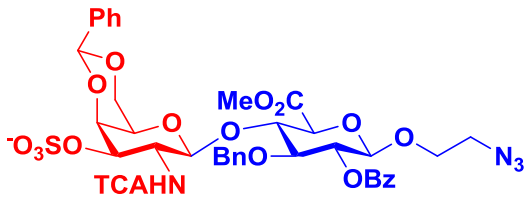


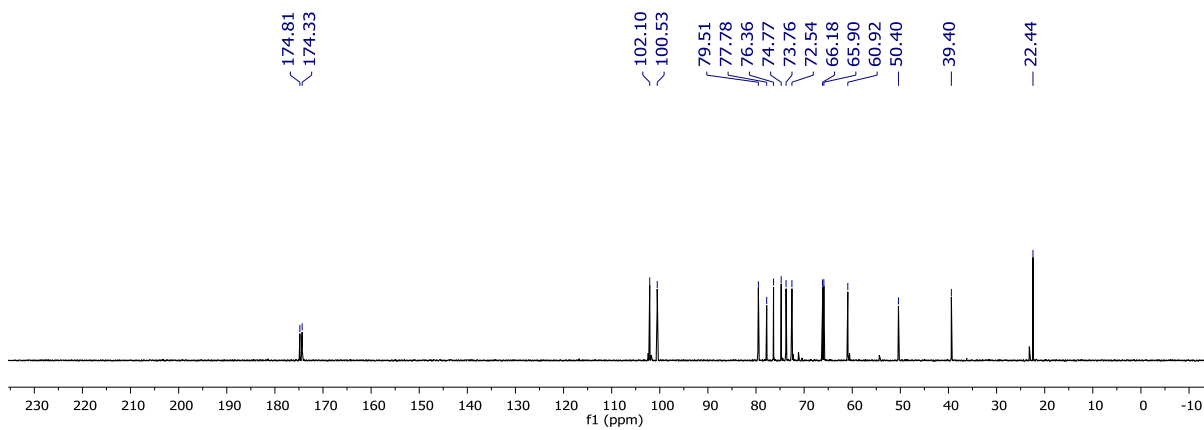
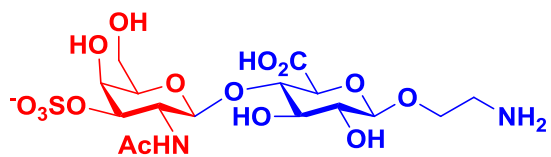
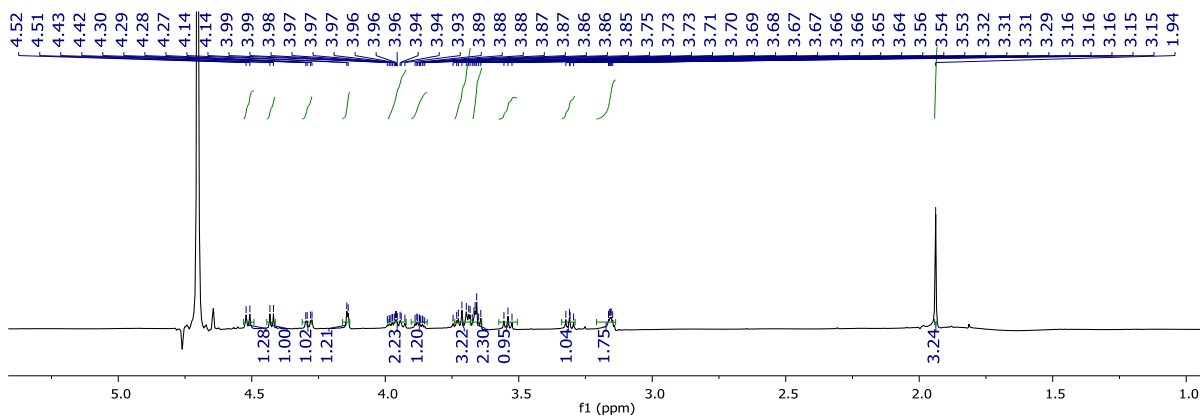
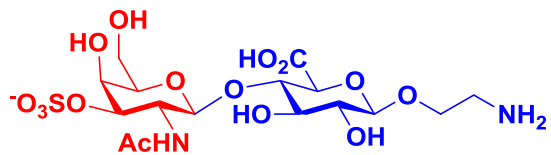


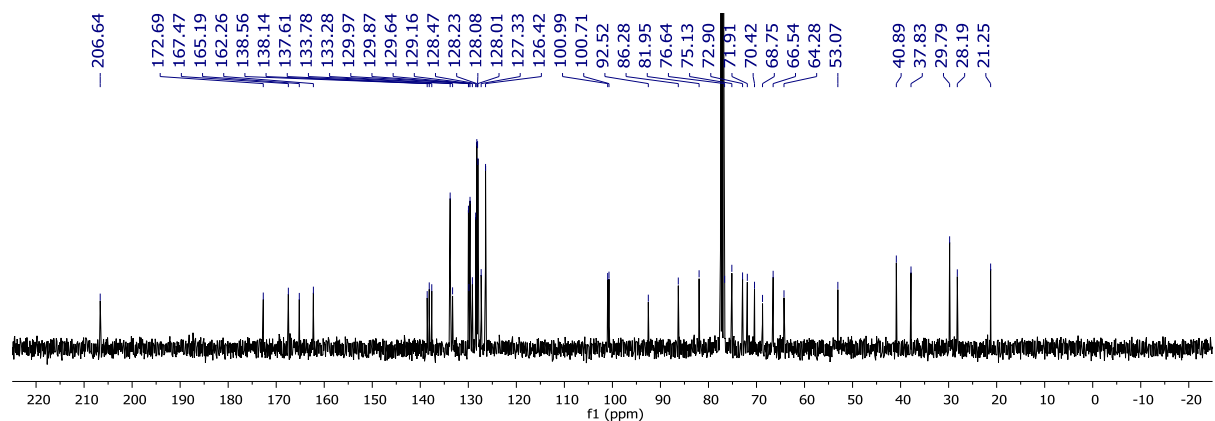
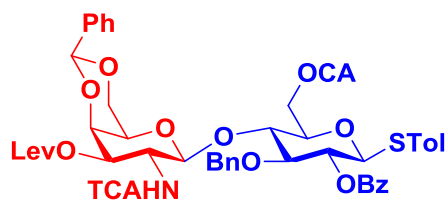
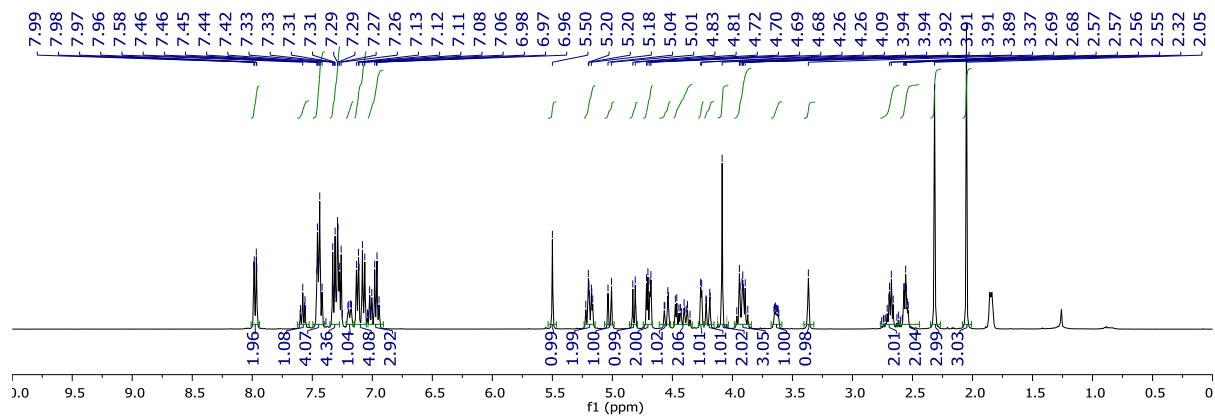
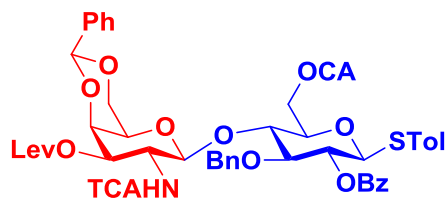


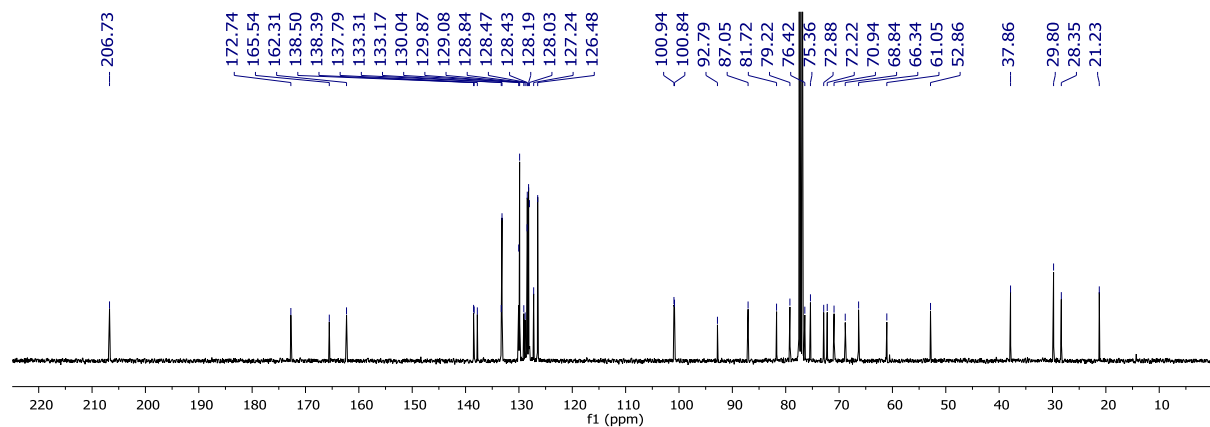
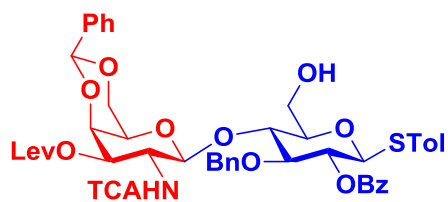
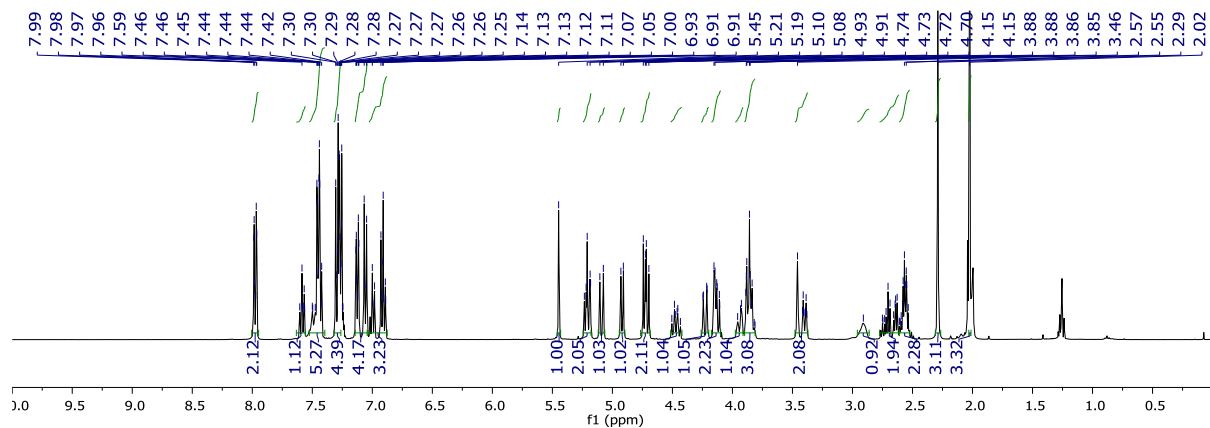
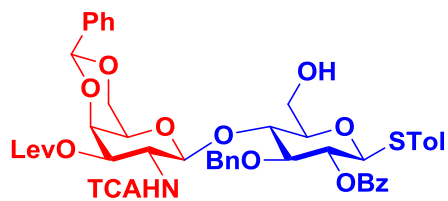


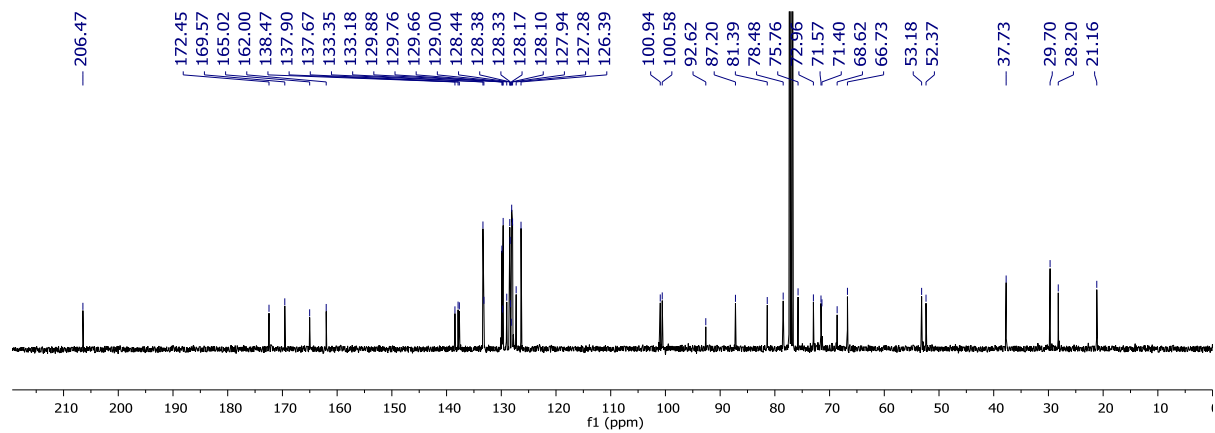
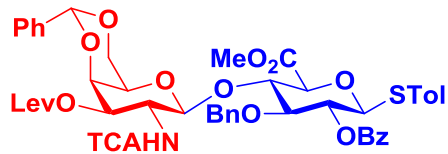
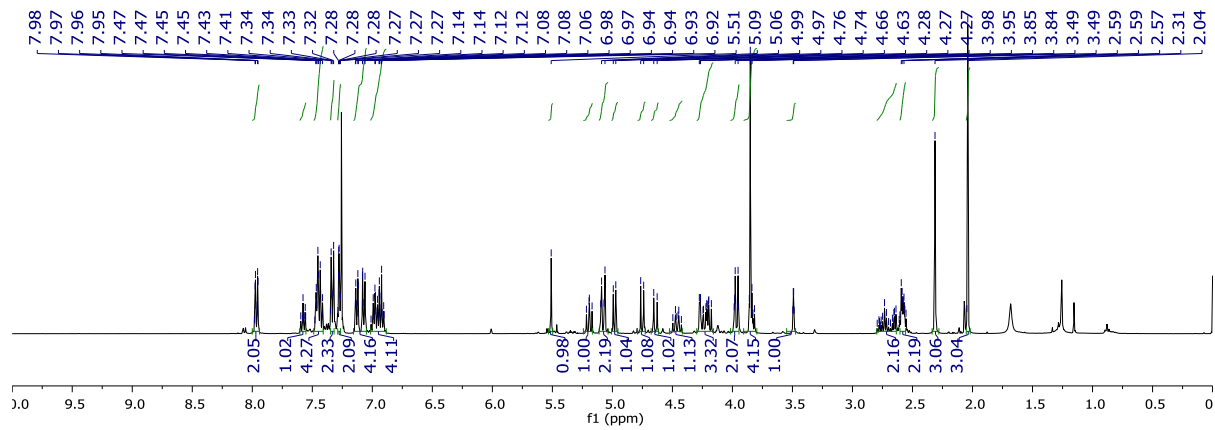
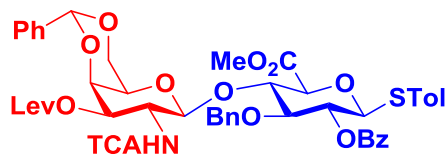


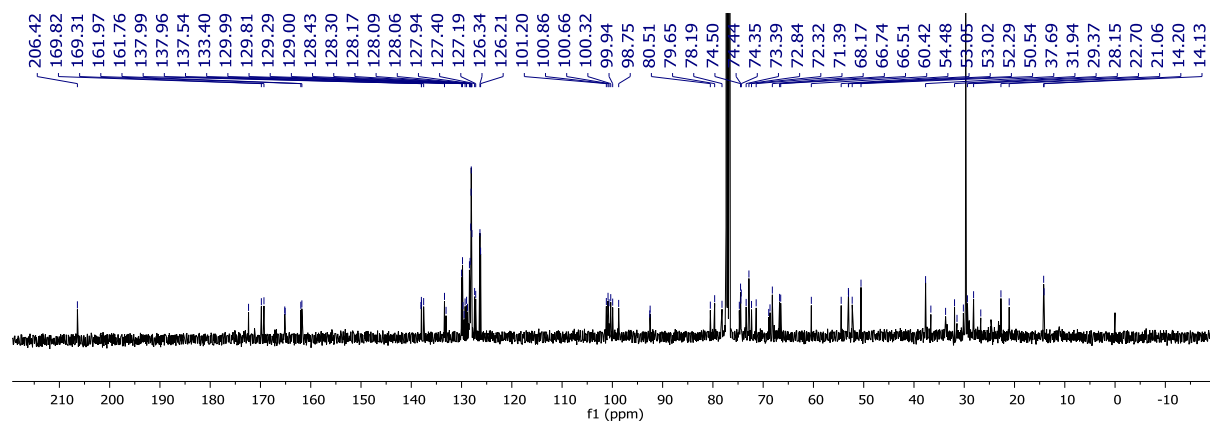
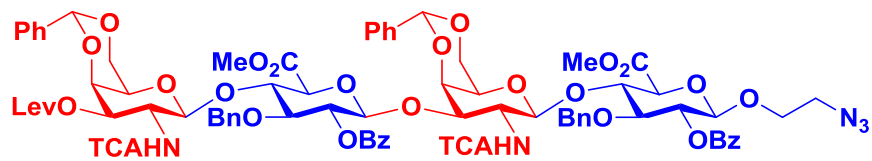
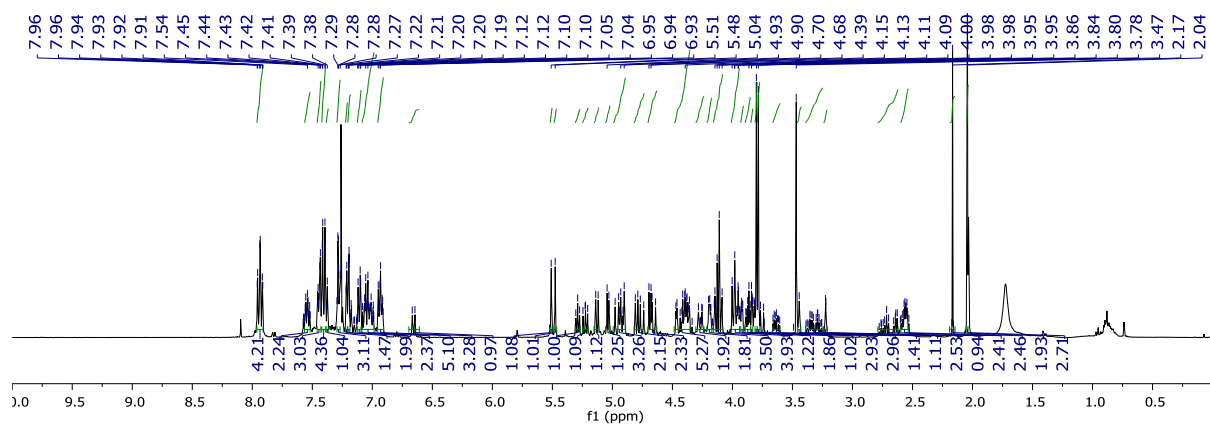
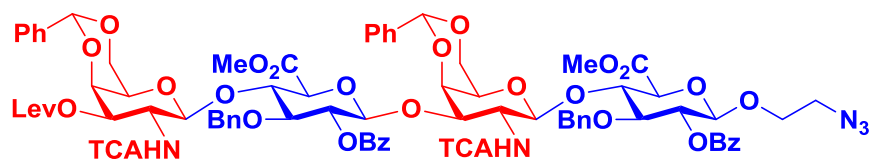




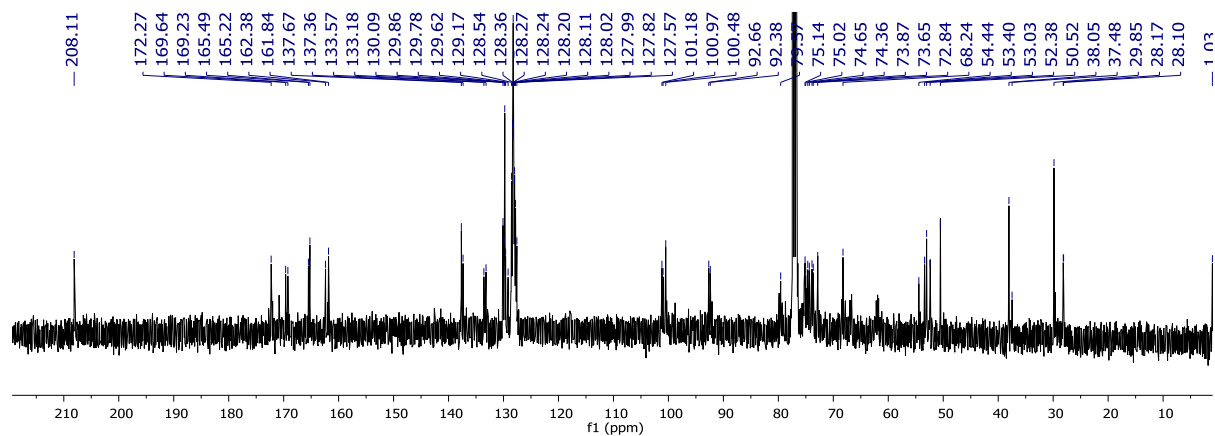
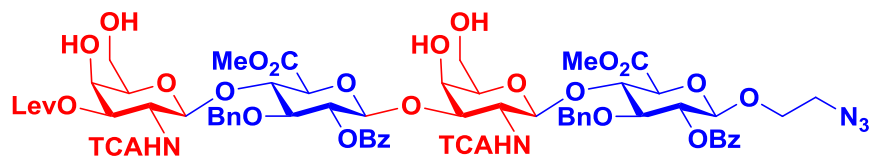
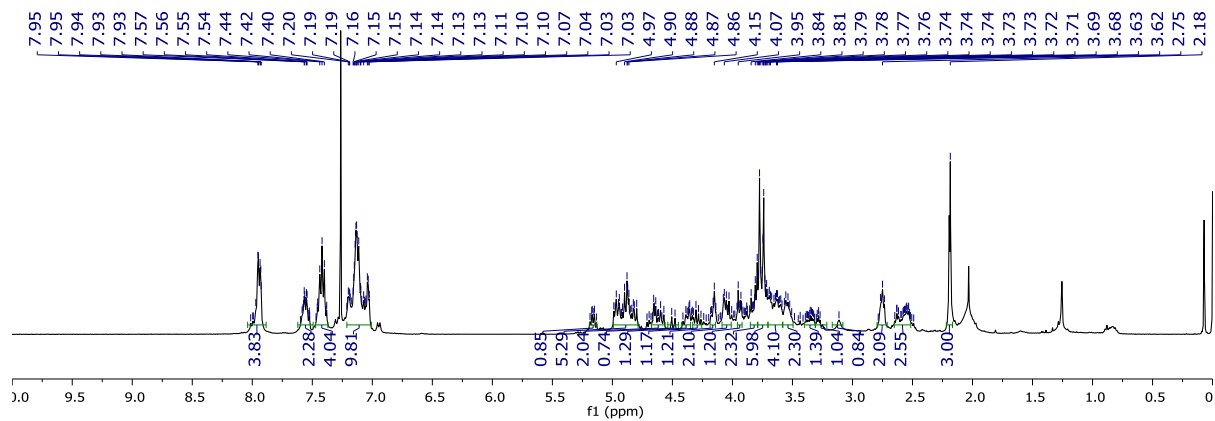
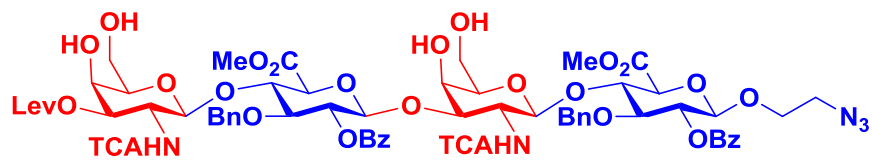


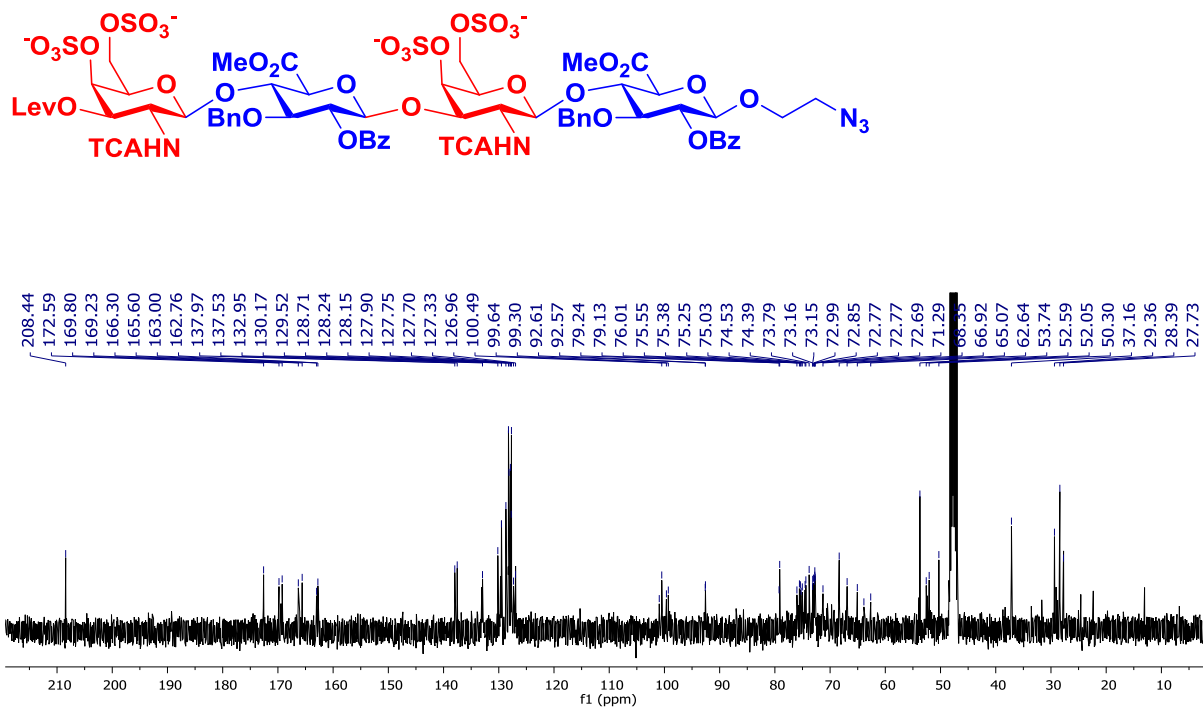
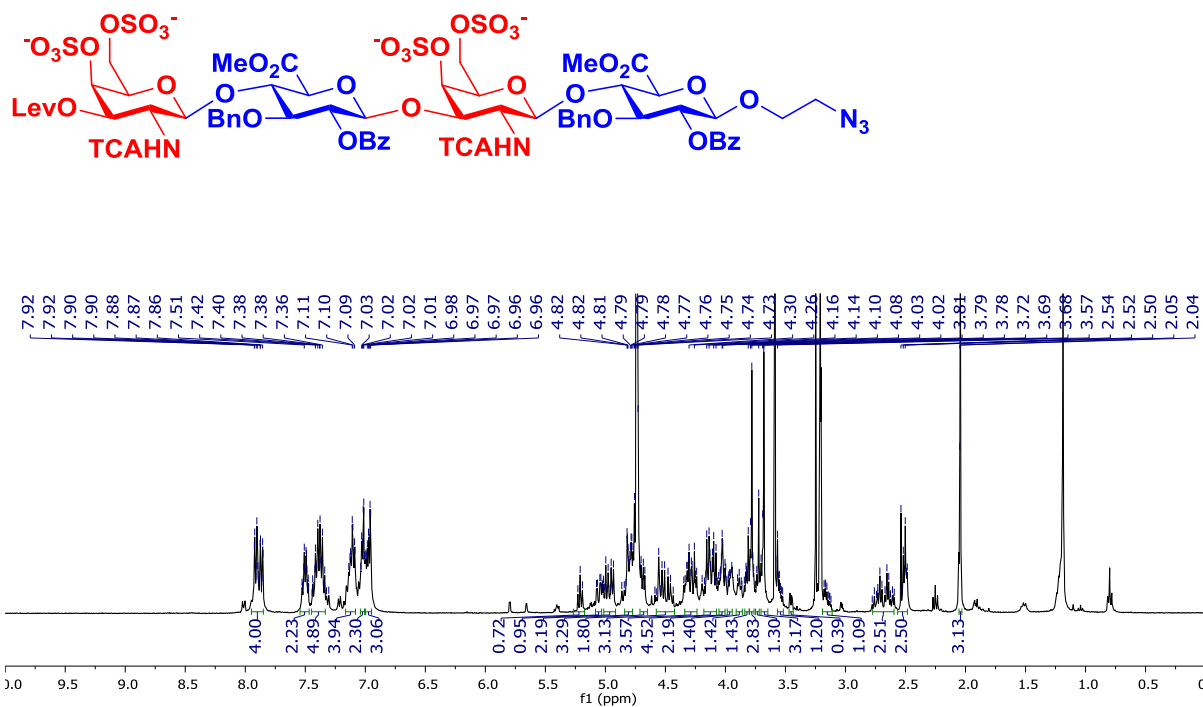


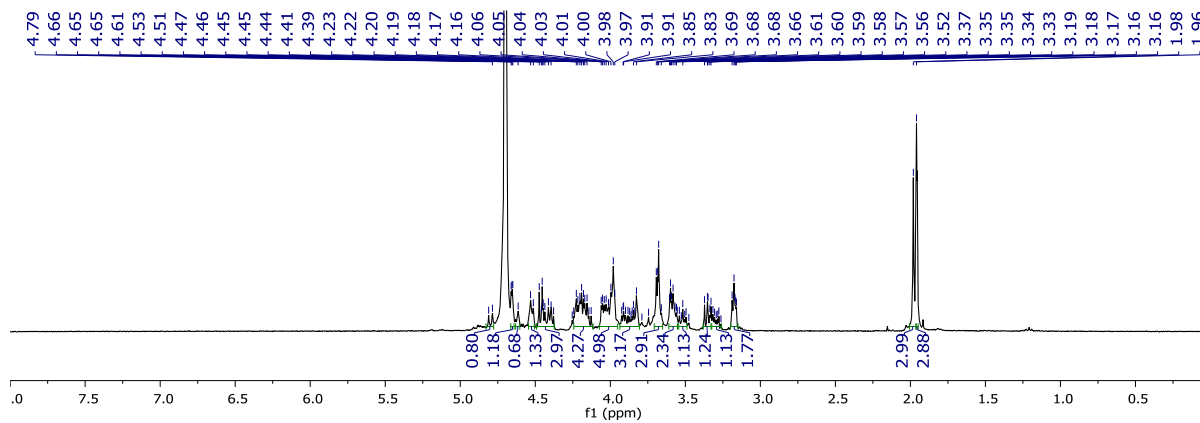
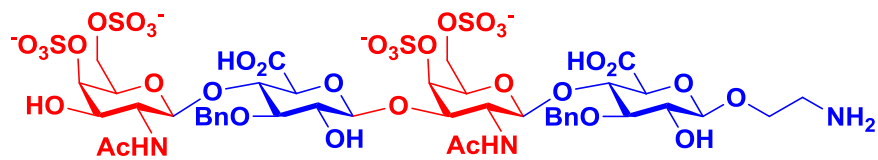


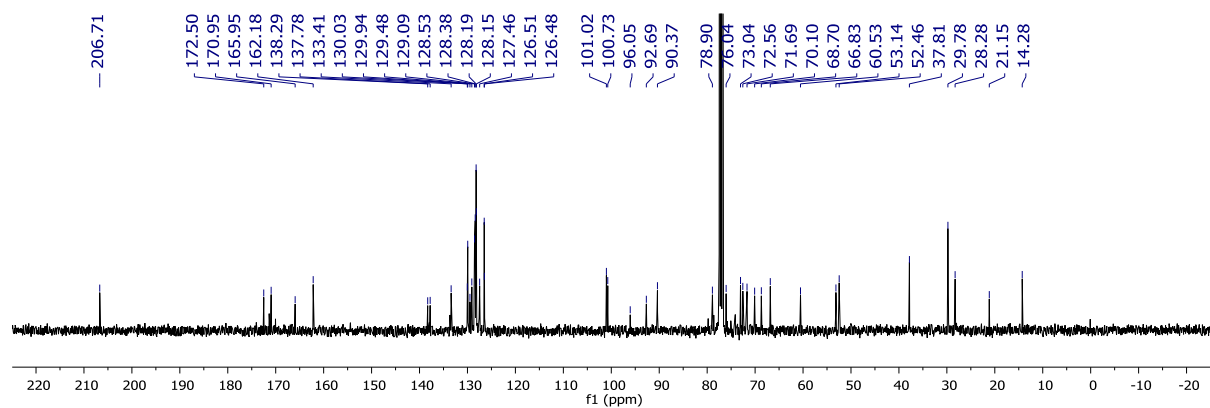
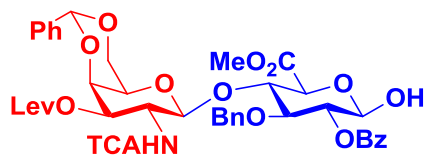
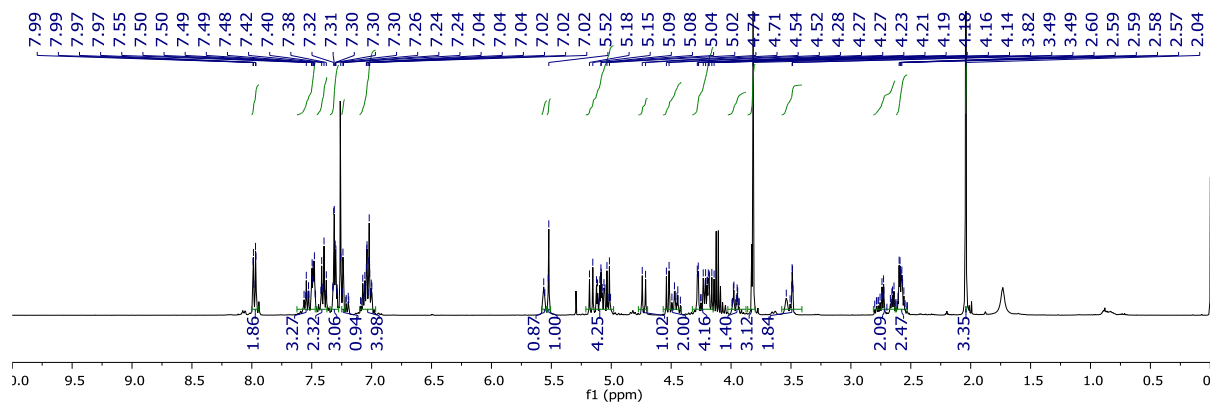
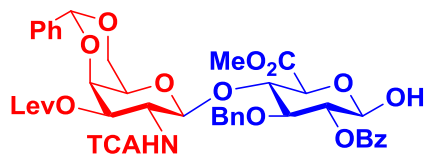


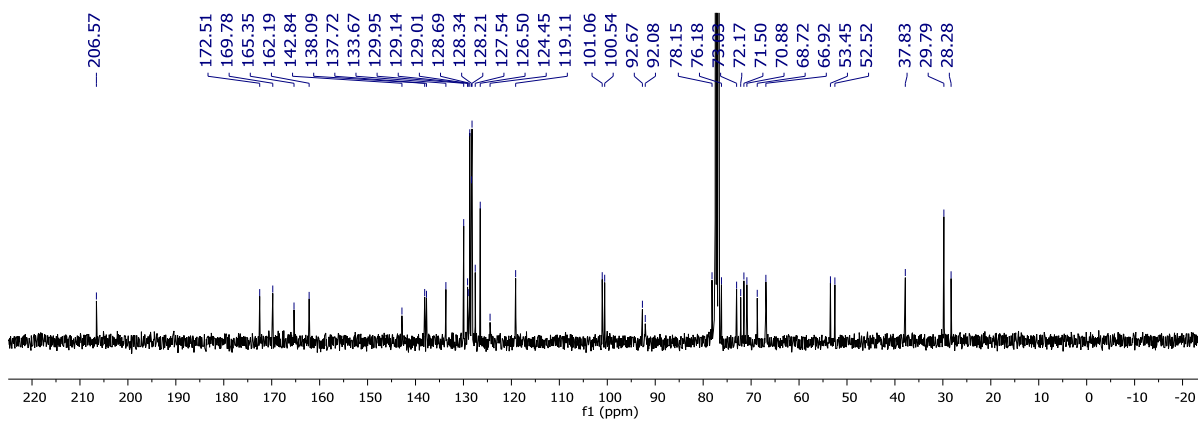
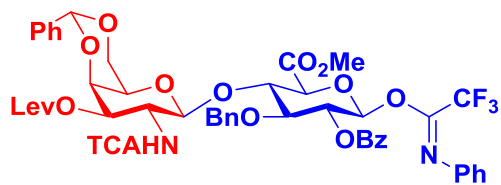
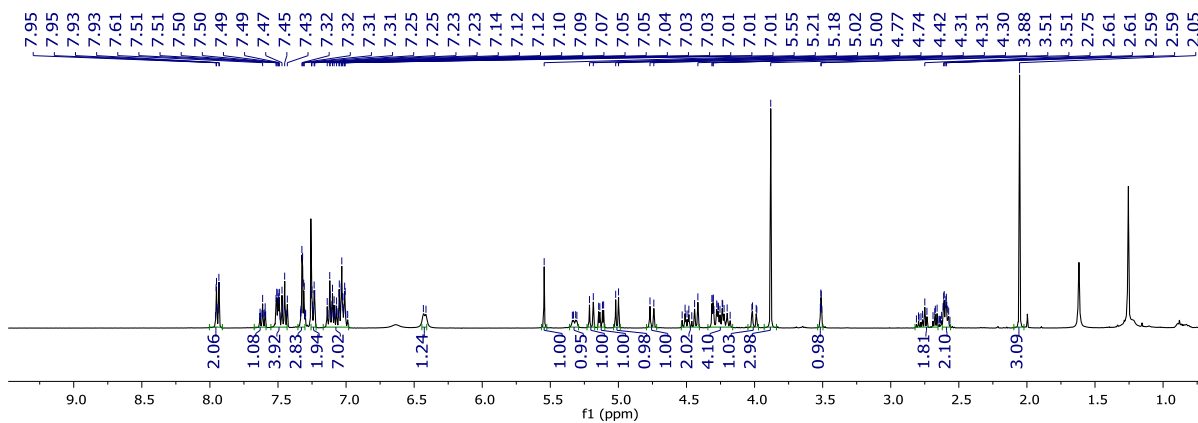
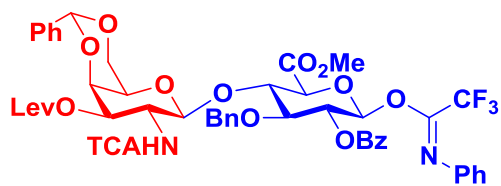


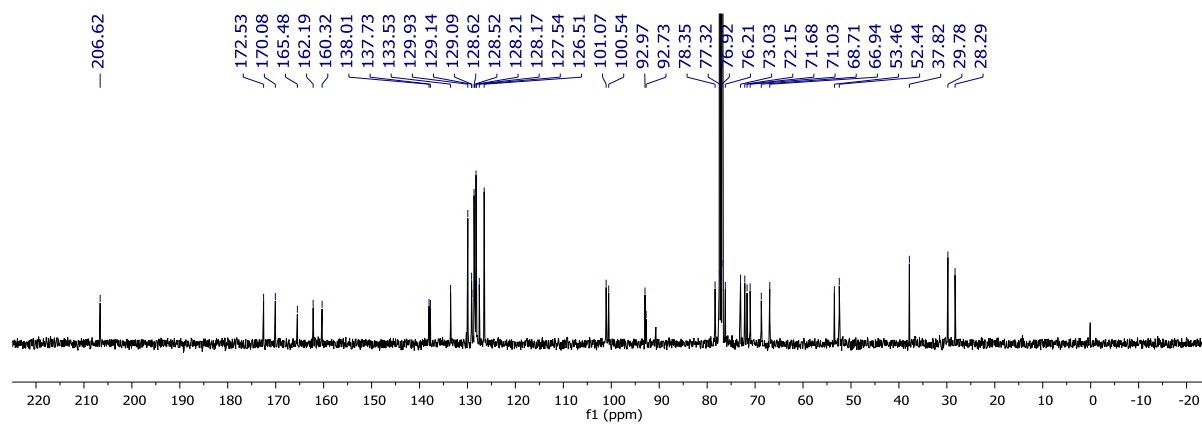
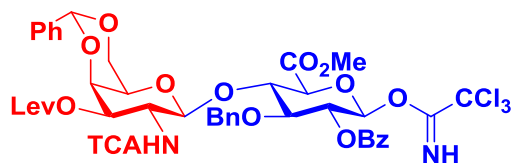
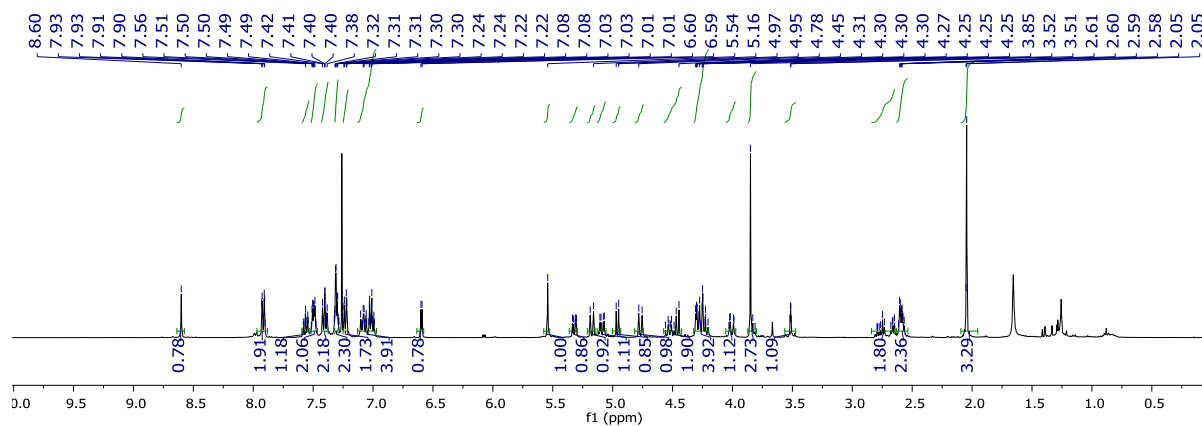
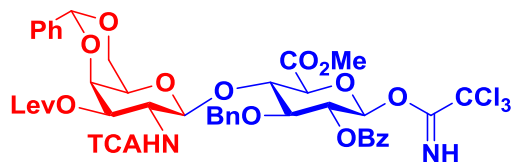


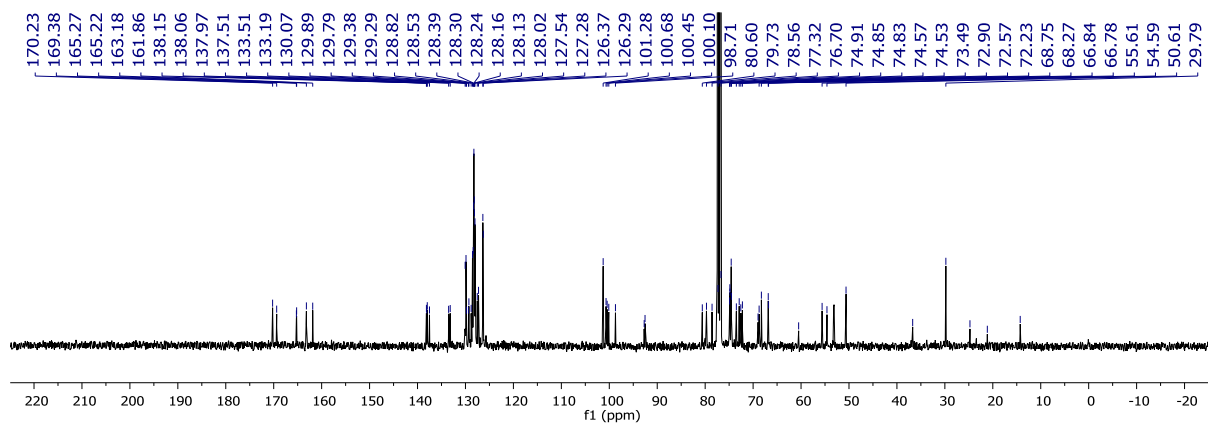
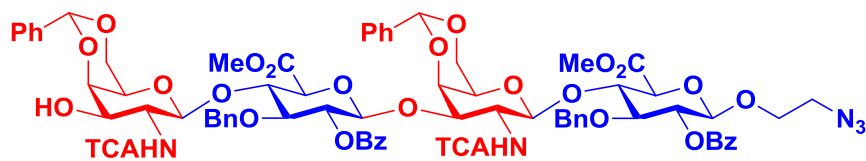
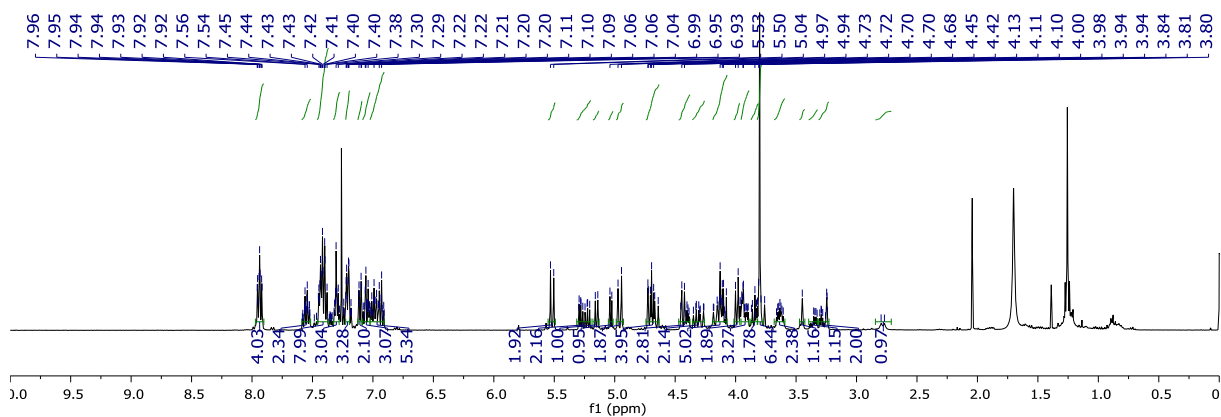
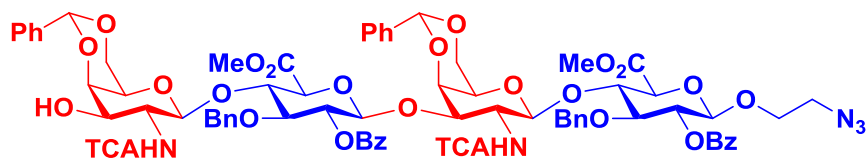


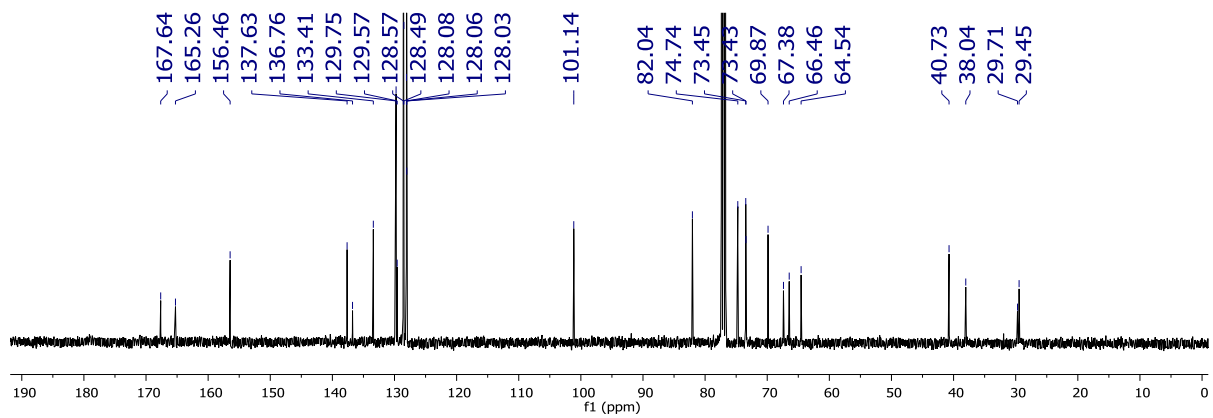
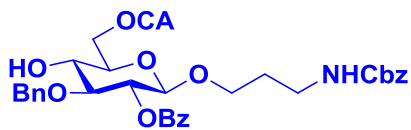
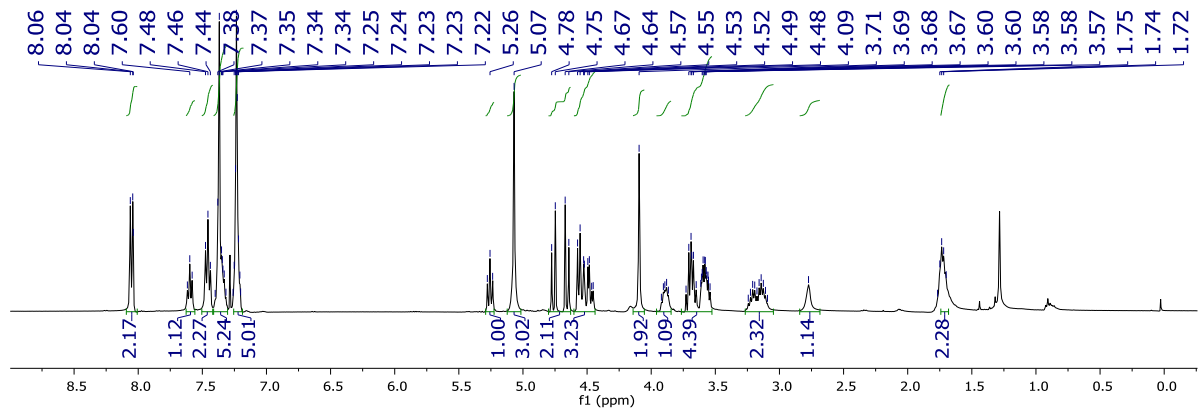
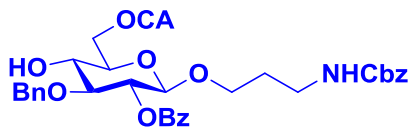




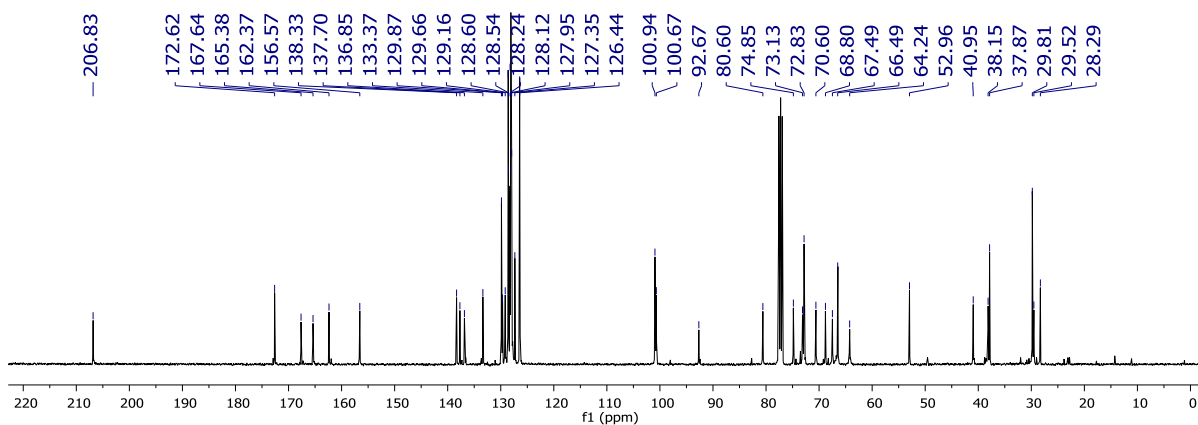
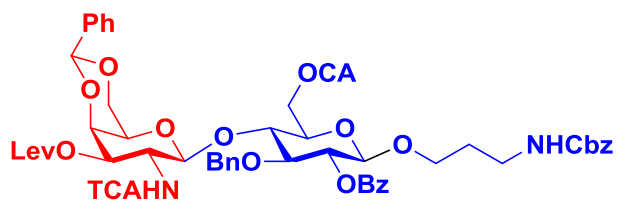
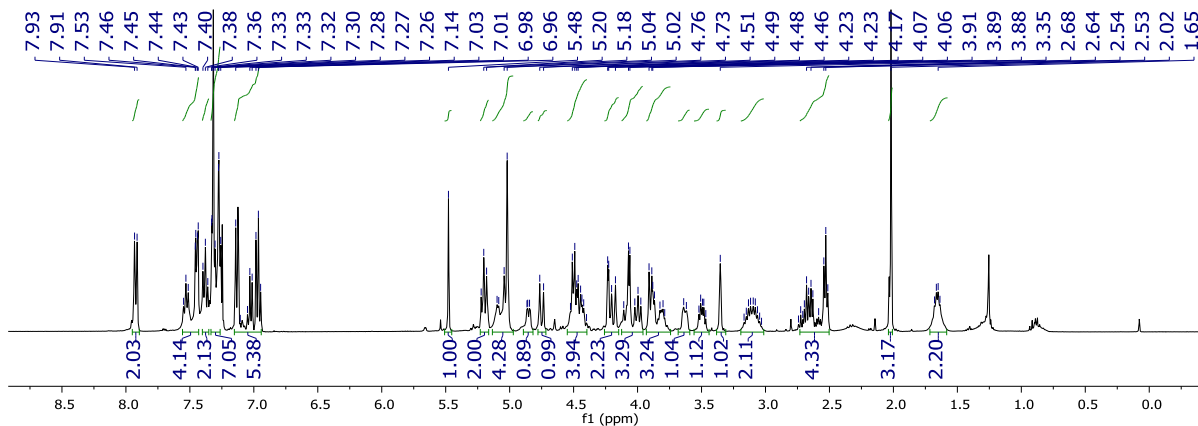
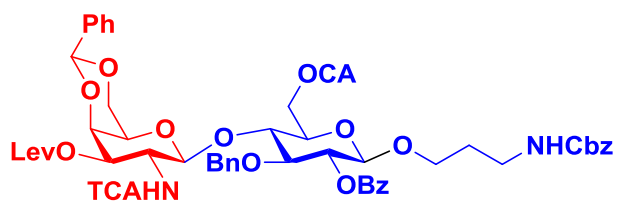


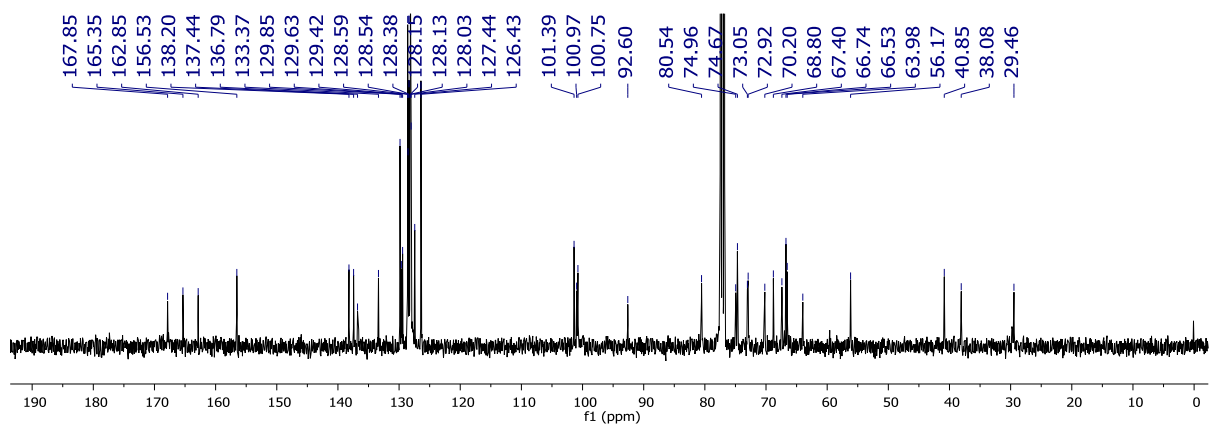
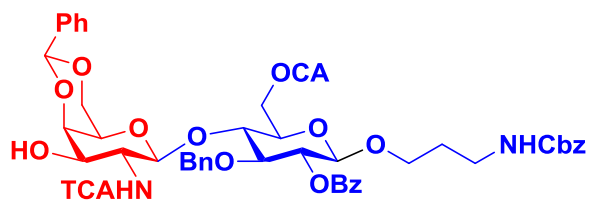
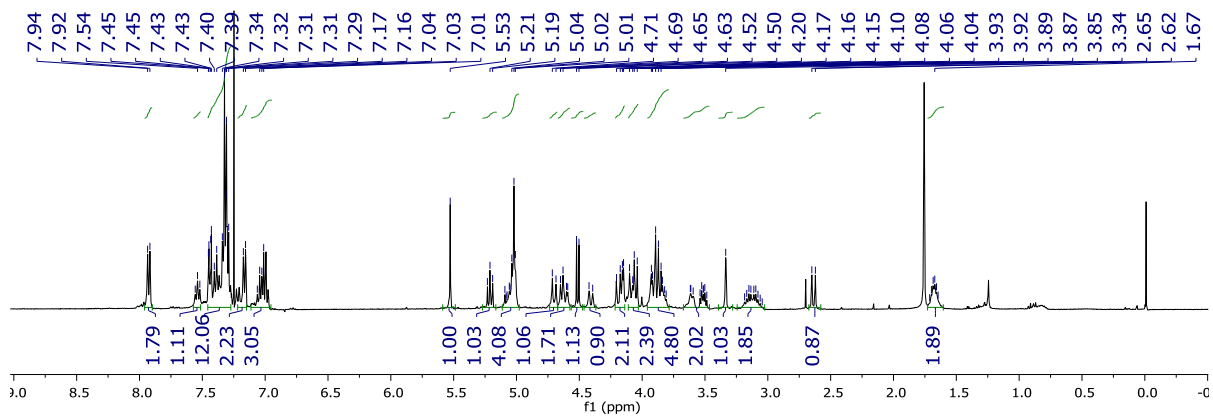
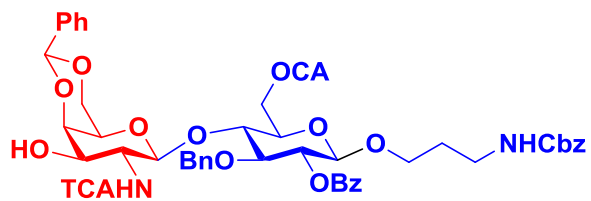


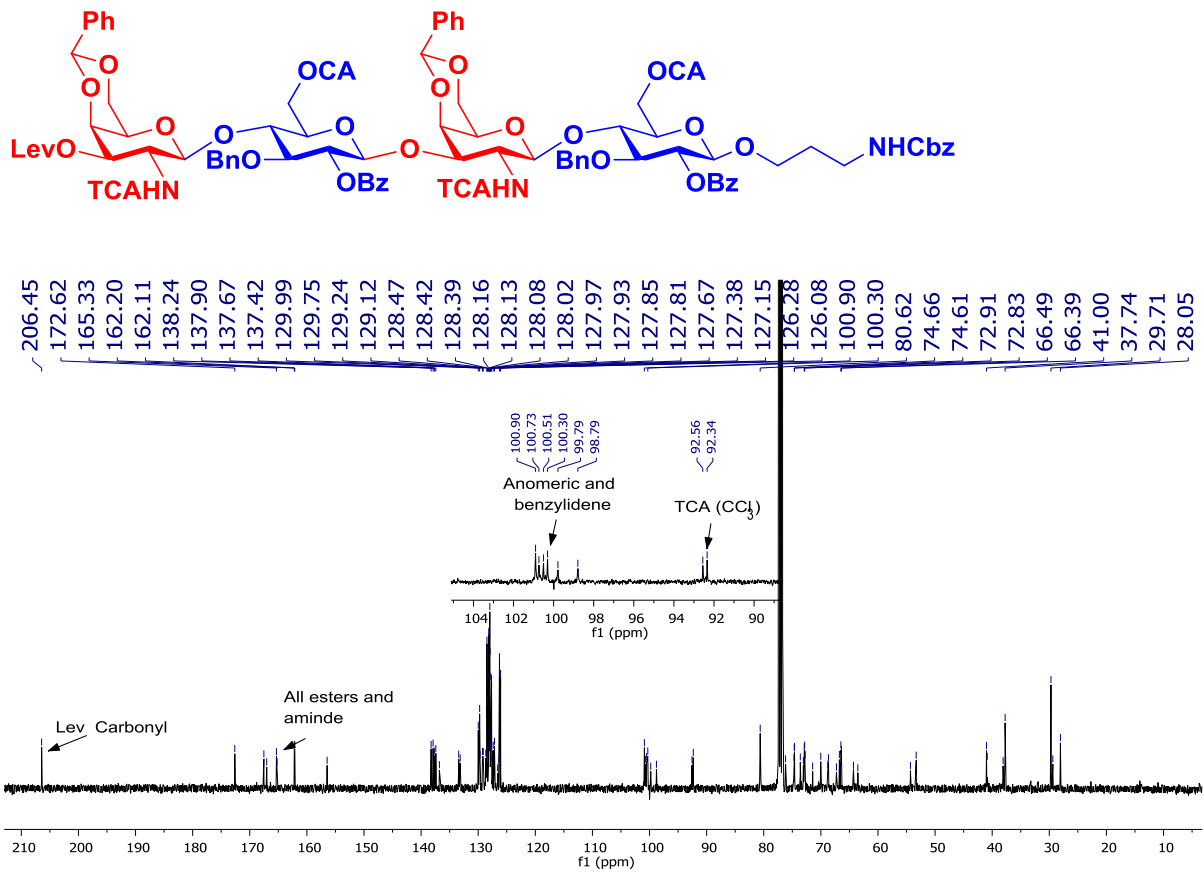
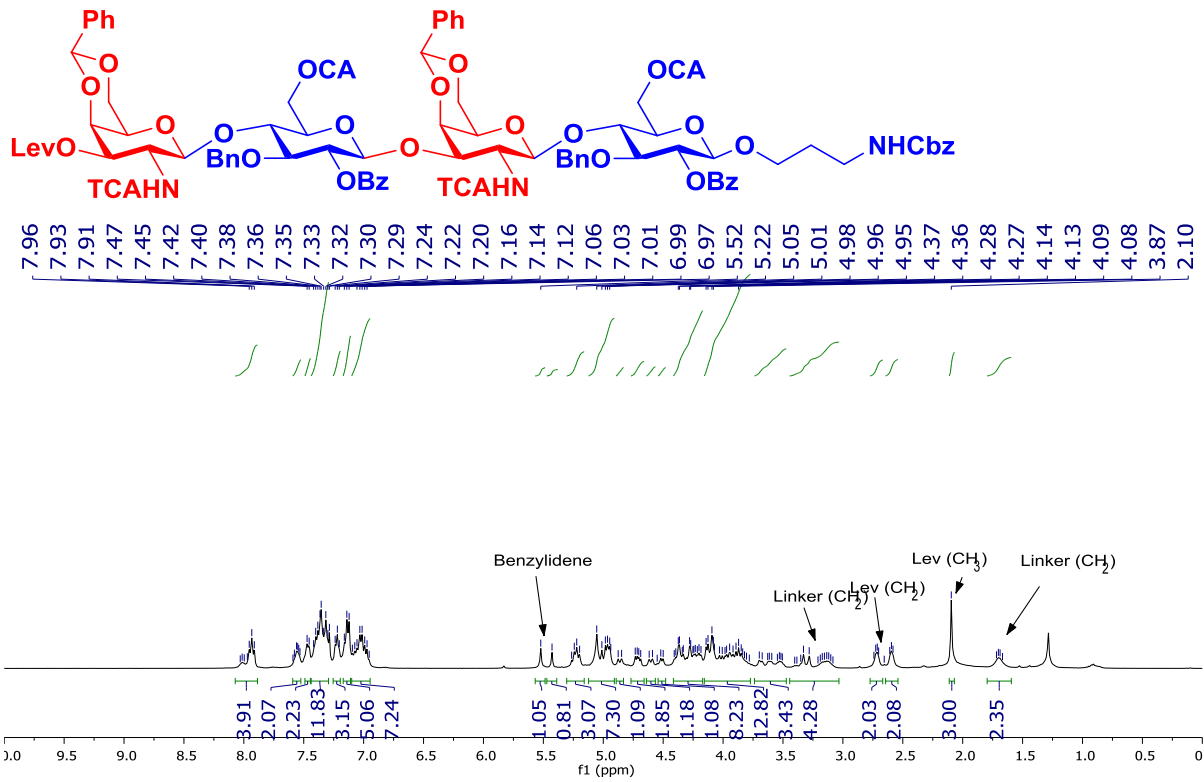


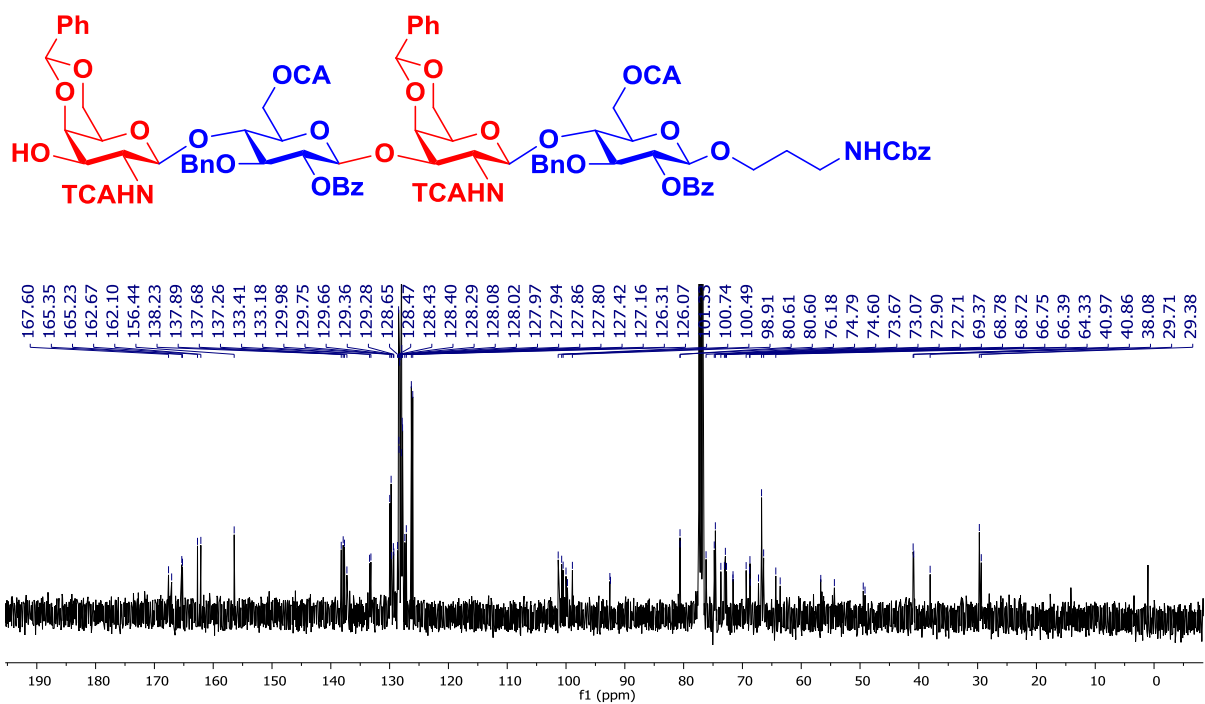
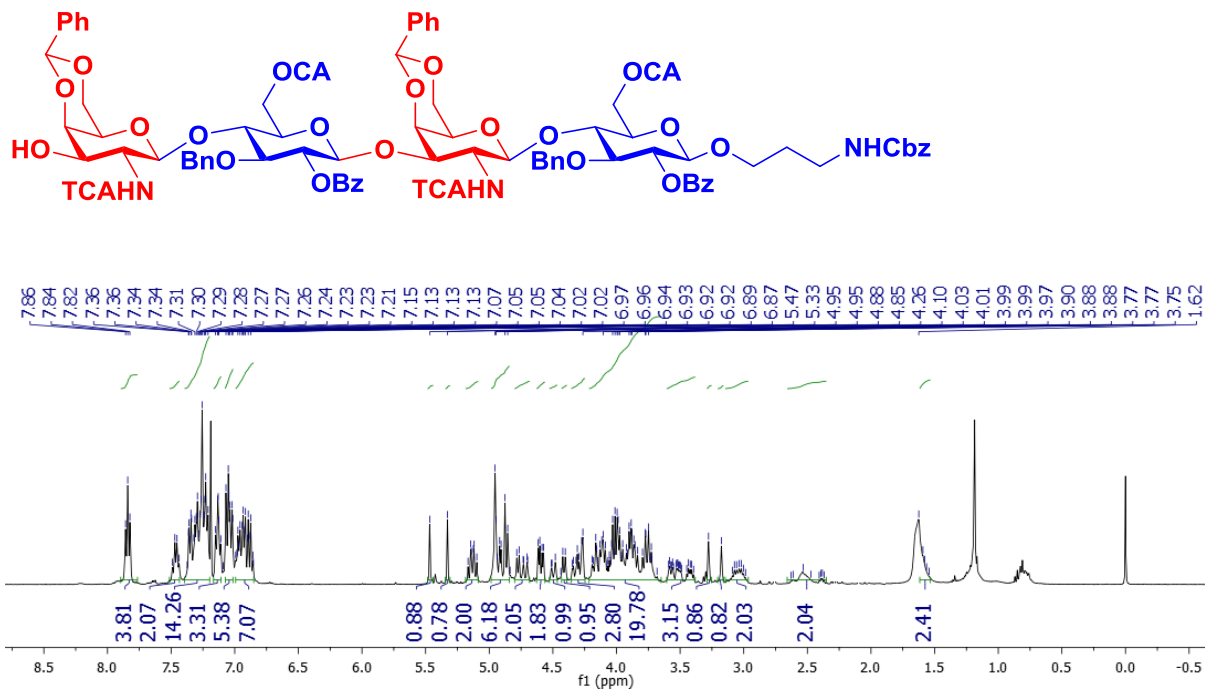


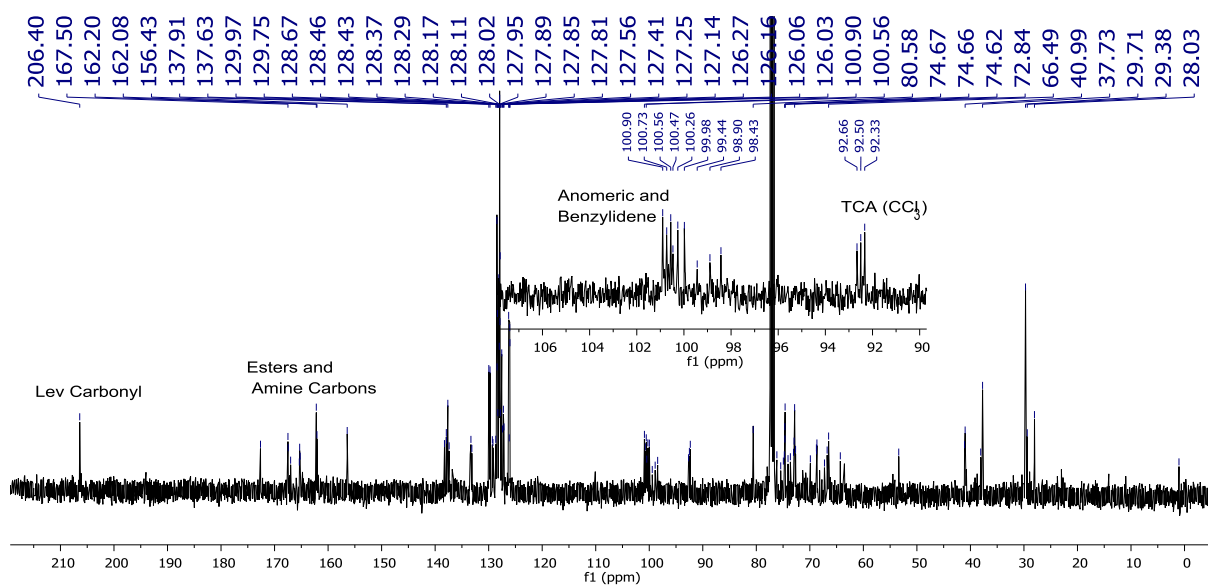
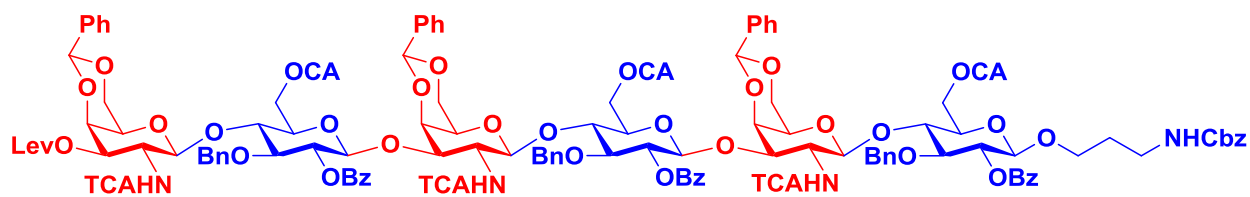
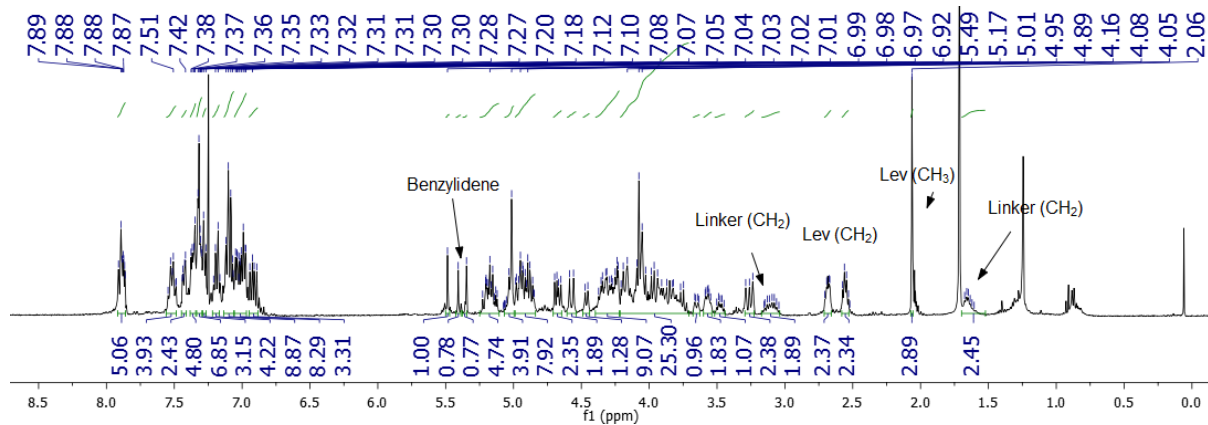
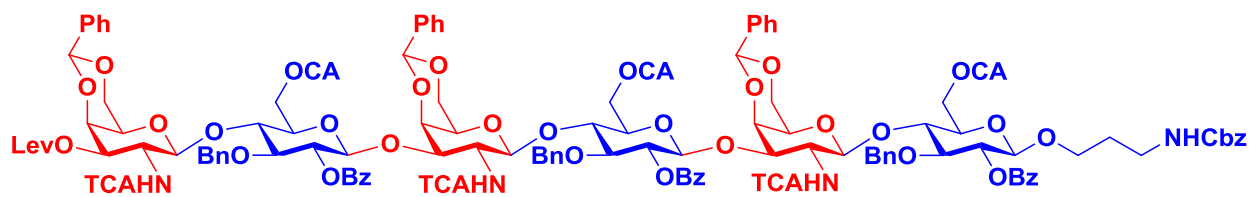












## **Chapter-4b**

# **Decoding Sulfation Patterns Dependant Endocytosis of Glycosaminoglycan Functionalized Gold Nanoparticles to Rationalize Targeting Strategies**

---

## **Abstract:**

Glycosaminoglycans (GAGs) are highly complex and information-rich biomacromolecules that bind to a wide range of proteins, including endocytosis cell surface receptors, in a sequence-specific manner. Hence, delineating the relation between GAGs-proteins interactions has a great potential for diagnostic and therapeutic applications. Here, we report the first systematic profiling of synthetic GAGs nanoparticles targeting cancer and glioblastoma cells. Our results showed that specific sulfation patterns on GAGs sequence binds to CD44 and regulate nanoparticles delivery to MDA-MB-231 and U87 glioblastoma cell lines. Further mechanism studies revealed that GAGs-nanostructures uptake is facilitated by clathrin-dependent endocytosis. These results imply the potentials of GAGs-based nanoprobe in nanomedicine.

## 4b.1 Introduction:

The extracellular matrix is an endogenous component of cells that comprise proteoglycans, collagen fibers and matrix proteins like laminin and fibronectin. An omnipresent component of the extracellular matrix is proteoglycans, which encompass linear, anionic glycosaminoglycans (GAGs) namely heparin/heparan sulfate (HS), chondroitin sulfate (CS)/dermatan sulfate (DS), keratan sulfate (KS) and hyaluronic acid. These glycosaminoglycans are composed of one amino sugar and either hexuronic acid or galactose as a dimeric unit.<sup>1,2</sup> Chondroitin sulfate plays inevitable biological events like cell division, cell-cell communication angiogenesis, inflammatory processes, neural development and spinal cord injury.<sup>3-5</sup> These functions are regulated by binding interactions with specific biomolecules such as cytokines, growth factors and cell adhesion molecules. CS is composed of N-Acetyl galactosamine (GalNAc) and Uronic acids [D-Glucuronic acid(GlcA) and L-Idouronic acid (IdoA)] as disaccharide repeating units. Divers sulfation at C4 and C2 on GalNAc and C2 position of uronic acid make the molecule more complex and heterogeneous. The definite sequence of CSPG is a potential molecule for a wide range of therapeutic applications. Scrutiny towards chondroitin sulfate proteoglycan (CSPG) pays less attention owing to its complexity and heterogeneity. The specific sequence and precise sulfation pattern of CSPG is a challenging task. This motivates us to aim for a prompt synthetic strategy for the preparation of chondroitin sulfate higher oligosaccharides for the structure-function relationship and their potential therapeutic applications.<sup>6-11</sup> However, synthesizing the precise sequences in CSPGs is a challenging task due to its complexity and heterogeneity. So, in this work, we aim to achieve the precise CSPGs oligosaccharide sequence with a specific sulfated microdomain to envision the structure-function relationship.<sup>12-15</sup>

At the structural level, Hyaluronic acid and CS form the macromolecular aggregates in PNNs and the effects of PNNs on neural function can be attributed to the properties of the GAG molecules. The development of PNNs is directly promoted by neural activity and is known to occur at least partly through changes in calcium and potassium conductance and the activation of NMDA and calcium-permeable AMPA receptors for the glutamate neurotransmitter.<sup>16,17</sup> A key concept regarding PNNs is that they limit plasticity in neurons once established, marking the end of the ‘critical period’ of development and that they can be dispersed or degraded under special conditions to reinstate juvenile-like states of plasticity to produce the sprouting of axons and regeneration of function in damaged areas of the brain.<sup>18</sup>



Notably, there are essential changes that occur in the sulfation patterns of CS chains during the development of PNNs. Particularly, 6-O-sulfation is found to be dominant in the juvenile brain to produce CS-C during periods of enhanced plasticity, whereas 4-O-sulfation becomes prominent in the adult brain to produce CS-A.<sup>19,20</sup> The influence of sulfation patterns of CS chains in regulating major events of plasticity has been overlooked in many previous studies, and the ability of sulfation to selectively direct the binding preferences of the HS and CS proteoglycans to cell surface receptors and other biomolecular entities in the ECM is not completely understood. Hence, there is a strong requirement for molecular mimics of heterogeneously sulfated proteoglycans – in the form of sulfated synthetic GAG oligomers – to study the binding activity of the proteoglycans and decode the structure-function relationship of the various GAG domains which are actively involved in physiological processes. We focused on synthesizing the disaccharides of CS and performed biological assays to study the structure-function relationship and also the differential regulatory role of the sulfation patterns.

The programmed release and accurate delivery of drugs/macromolecules for targeting specific physiological areas is a key challenge for therapeutics.<sup>21</sup> In the past, the modification of gold nanoparticle (AuNP) surfaces with biomolecules of interest has been established as an efficient mode of transport of biomacromolecules with minimal cytotoxicity, and provides a biocompatible way to achieve specific targeting of regions without conjugation of additional moieties.<sup>22</sup> Glyco-AuNPs present multi-modal and versatile nanoplatforms to develop carbohydrate-based technology for applications as biosensors, drug carriers, and effective tools in the field of cancer and immunological research.<sup>23</sup> In our study, we utilized spherical AuNPs conjugated with a FAM fluorophore as carriers for laboratory synthesized oligomer mimics of our GAGs of interest. The glyco-AuNPs were used in cellular targeting and molecular uptake studies in primary neuronal cell cultures and also secondary cell lines in vitro. We also utilized antibodies and blockers for certain receptors or pathways to study the mechanism of uptake/targeting which is utilized by different sulfated CS glycans in vitro.

CD44 is an integral membrane glycoprotein involved in intercellular interactions, cell adhesion, and migration for multiple cell types, and it serves as the primary cell surface receptor for Hyaluronic acid (HA) and as the recognition site for CSPGs. Researchers in the past have shown that malignant cells synthesize increased levels of CD44-related CSPG as a matrix receptor to mediate the invasion and migration of tumors.<sup>24</sup> Further, a CSPG found on the cell surfaces of melanomas contains a core protein that is recognized by CD44 antibodies,

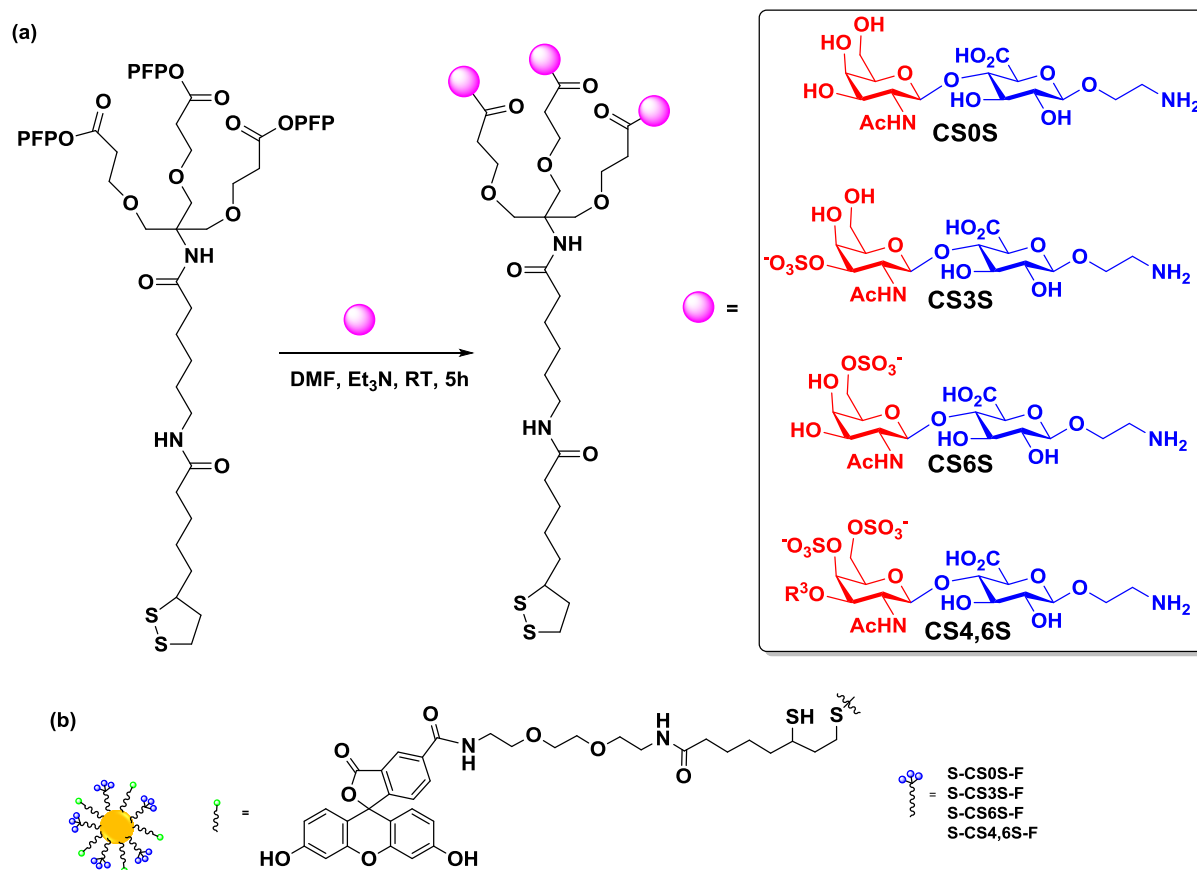
and helps in mediating cell motility and invasion of the melanoma into type I collagen gels.<sup>25</sup> In non-malignant cells, studies have shown that a CD44-related CSPG cell surface matrix receptor is capable of binding to fibrinogen/fibrin, in turn mediating the microvascular endothelial cell migration of wounds on fibrinogen and invasion into a fibrin matrix.<sup>26</sup> The various isoforms of CD44 are differentially modified by the GAGs and the assembly of these macromolecules occurs at sites that have a serine unit followed by glycine (SG).

As described before, synthetic access to GAG molecules or mimetics of defined length and sulfation pattern is very important for systematic investigations of SFR. Historically, considerable attention has been focused on HS and its functions, while much less is known about the CS class of GAGs. CS has been shown to inhibit the growth of axons in neural networks; yet it is also found in developing, growth-permissive regions of the brain and also in the recovering areas of the brain and spinal cord.<sup>28</sup> Similarly, in-vitro studies have shown that CS polysaccharides can both attenuate and stimulate the growth of primary neuronal cells in culture.<sup>29,30</sup>

Overall, was aimed toward understanding the role of sulfation in regulating the binding, uptake, and receptor interactions of the HS and CS GAGs and studying the functionality of the CS-E mimetics in the paradigm of neural plasticity and modulation of neurite outgrowth. Our study presents a novel attempt to characterize the targeting of secondary cell lines (U87 & SH-SY5Y) and primary murine cells of the olfactory bulb and the hippocampus by using heterogeneously sulfated libraries of HS and CS mimetics functionalized to AuNPs. For our research, we utilized p0/p1 neonatal mice which are present in the critical period of development of the CNS, for which the PNNs surrounding the PV cells are not developed.<sup>32</sup> Herein, we report the use of fluorescent glycosamino glycan conjugated AuNPs for detection of sulfation code in neural cell recognition. We have synthesized chondroitin sulphate disaccharides with various sulfation patterns conjugated to tripod functionalized gold nanoparticles.

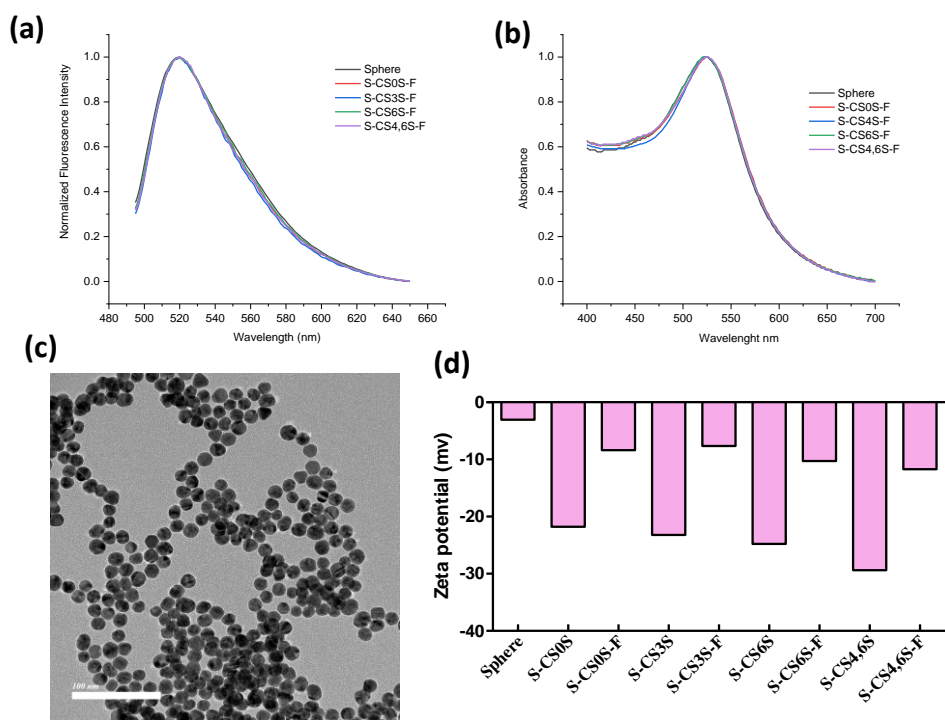
## 4b.2 Results and Discussion:

Further to assess the efficacy of different sulfation patterns in neural cell recognition we functionalized CS/HS conjugated tripods on spherical shape AuNPs. Spherical AuNPs of 25 nm were synthesized using previously reported procedure.



**Figure 1.** (a) CS conjugation to tripod; (b) Schematic representation of CS-tripod functionalized fluorescent AuNPs

CS and HS Tripod functionalization on AuNPs was carried out using simple ligand exchange method by mixing tripod with AuNPs in PBS buffer at RT for 12h. AuNPs functionalization was confirmed by shift in zeta potential value towards more negative side. Further these tripod functionalized AuNPs were mixed with fluorescent linker F-1 in PBS buffer at RT for 12h to get fluorescent AuNPs (**S-CS0S-F**, **S-CS3S-F**, **S-CS6S-F**, **S-CS4,6S-F**). Physical characterizations of all AuNPs were done using transmission electron microscopy (TEM), UV-visible and fluorescence spectroscopy and zeta potential measurements.

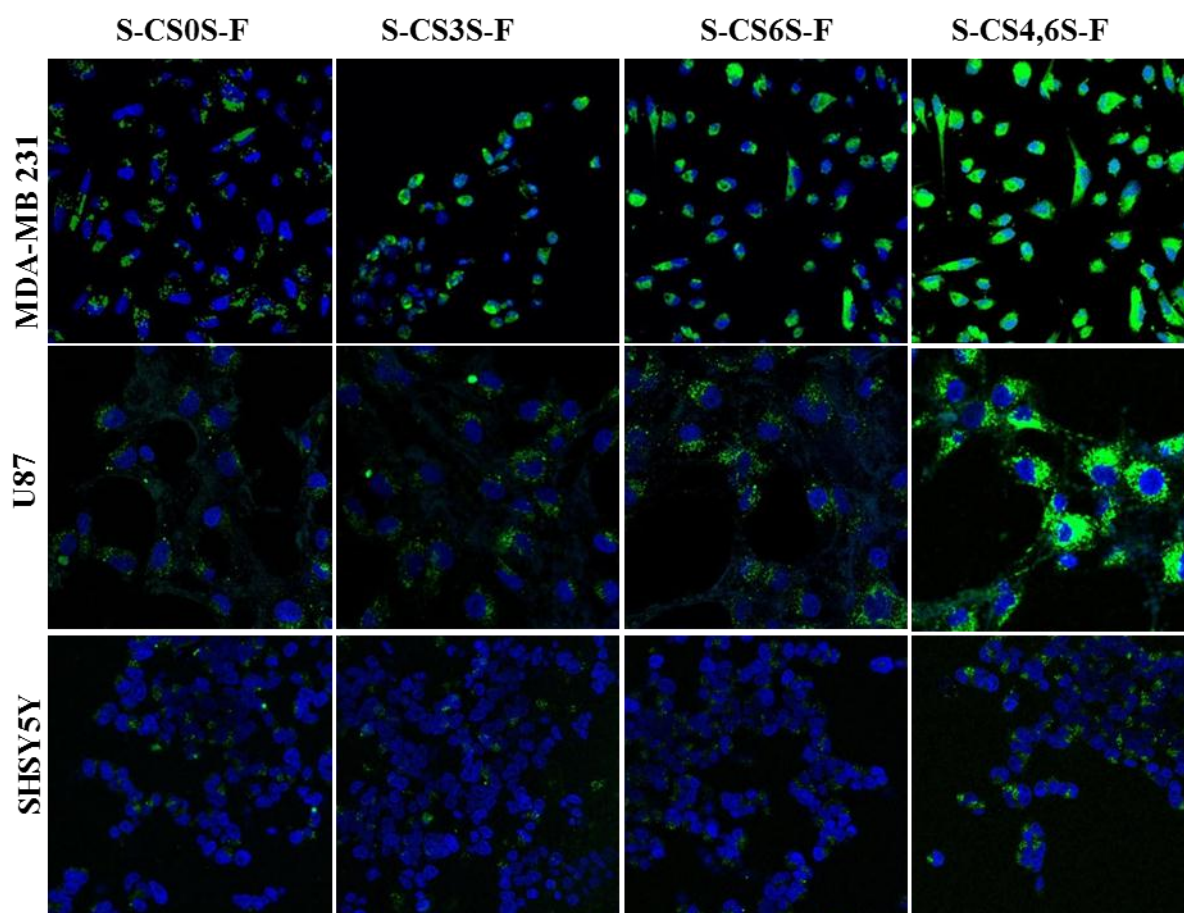


**Figure 2:** Physical characterization of fluorescent and CS functionalized AuNPs; (a) Fluorescent spectra; (b) UV Spectra; (c) TEM image; (d) Zeta potential.

Further a cellular uptake assay was performed using different neural cell lines. We have selected secondary neural cell lines U87 and SHSY-5Y, primary neural cell culture olfactory bulb and hippocampal. Cells were seeded on coverslip and treated with AuNPs for 4h. Among different AuNPs CS tripod functionalized AuNPs showed moderate to high uptake. Non sulphated CS tripod was unable to be taken up by any cell line asserting the importance of sulfation code for uptake mechanism. Compared to mono sulphate CS tripod the 4,6 disulfate tripod AuNPs (**S-CS4,6S-F**) exhibited maximum uptake in U87 cell lines. Similar trend of uptake was observed in case of olfactory bulb culture and hippocampal culture. On contrary, SHSY-5Y cells did not show uptake of CS tripod functionalized AuNPs, only **S-CS4,6S-F** showed less uptake.

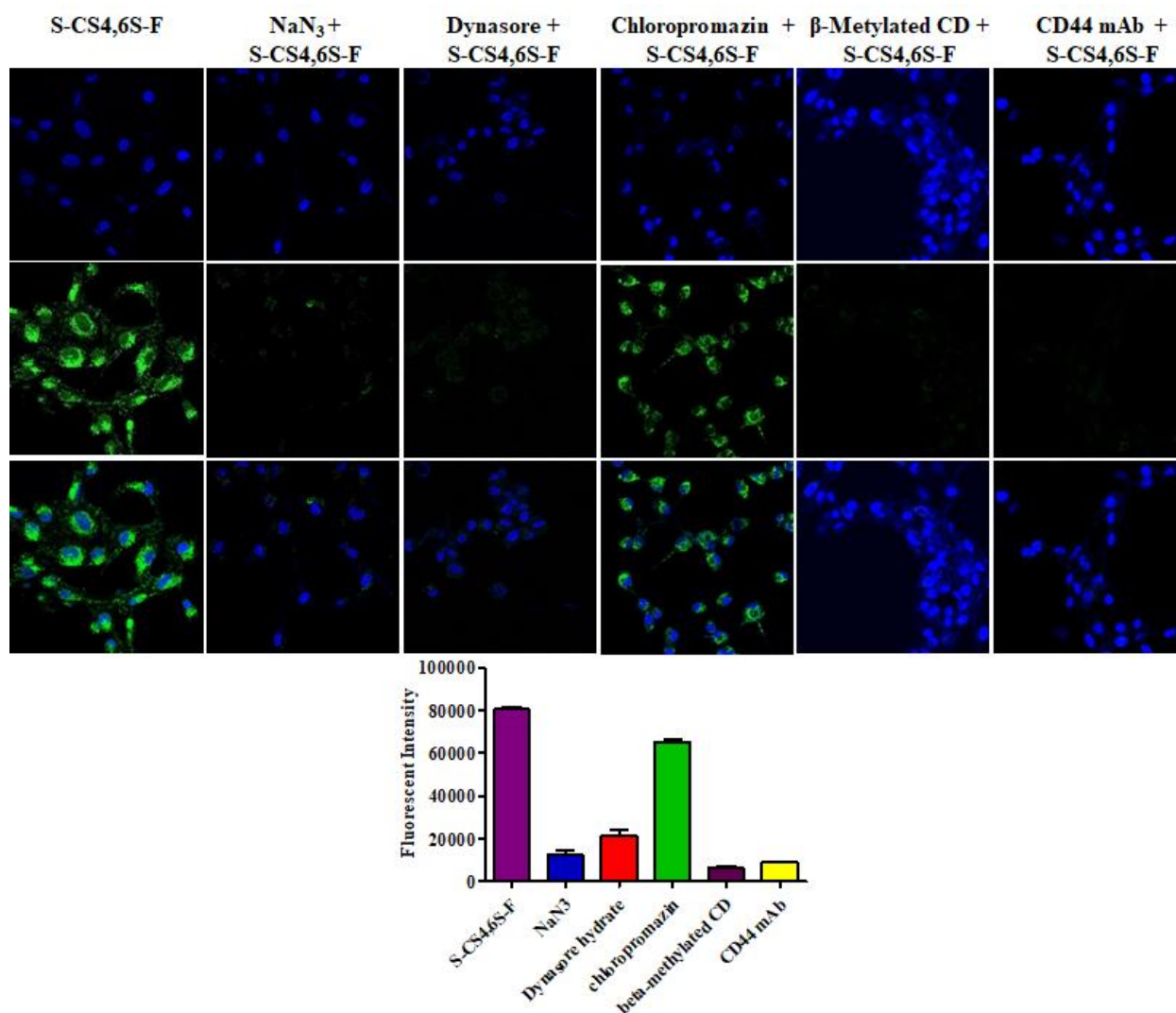
The expression level of CD44 receptor on cell surface was checked using immunostaining using CD44-FITC antibody, it can be seen clearly that U87 and MDA-MD 231 as well as olfactory bulb culture and hippocampal culture showed high level of CD44 expression. While in case of SHSY-5Y CD44 receptor expression was absent, which is indicative of AuNPs uptake might be CD44 receptor mediated. To prove CD44 depended uptake mechanism cells were pre-treated with monoclonal CD44 antibody for 1 h followed by treatment with **S-CS4,6S-F**. It was observed that there was no uptake of **S-CS4,6S-F** in case

of pre-treated U87 and primary neural cultures.



**Figure 3:** Confocal microscopy images for uptake of CS functionalized AuNPs by SHSY5Y, U87 and MDA-MB 231 cell lines at 4 h.

To analyse the mechanism of cellular internalization, cellular uptake of **S-CS4,6S-F** was evaluated in presence of different known inhibitors for dynamin, clathrin and caveolae pathways. Cells pre-treated with sodium azide for 30 min showed strong decrease in the uptake of **S-CS4,6S-F** indicating the active endocytosis in U87 cells. Further pre-treatment of dynasore hydrate, a dynamin inhibitor resulted in reduced uptake of **S-CS4,6S-F** indicative of conventional trend of a clathrin or caveolae pathway. Chlorpromazin pre-treated cells showed negligible reduction in uptake of **S-CS4,6S-F** while addition of  $\beta$ -methylated CD to cells showed significant reduction in uptake of **S-CS4,6S-F** validated caveolae mediated endocytosis mechanism.



**Figure 4:** Confocal microscopy images for mechanism of cellular internalization of CS AuNPs using U87 cell lines and respective fluorescence quantification.

### 4b.3 Synthesis of tripod:

Tripod-active pentafluorophenol ester was synthesized from tris-base utilizing a previously reported procedure. Briefly, Michael addition of tris-base with acrylonitrile in the presence of KOH. The nitrile functional group was hydrolyzed and ethyl esterified in the presence of concentrated HCl and ethanol and coupling with 5-(Boc-amino)pentanoic acid and Boc deprotection followed by another coupling with lipoic acid succinimide ester resulted in the tripod-active pentafluorophenol ester.

#### 4b.4. Conclusion:

To understand CS's structural-functional relations of disaccharide heterogeneity, we have synthesized tripodal-CS analogs and functionalized them on fluorescent gold nanoparticles. Using secondary neuroblastoma cells such as U87 and SH-SY-5Y, showed for the first time that specific carbohydrate-protein interactions, modulate selective delivery of nanoparticles to the neuroblastoma cells. Our results showed that 4,6-*O*-disulfated-CS nanoprobe selectively target CD44 overexpressed cell lines *via* caveolin-dependent endocytosis pathway. These results imply new opportunities to utilize GAGs-nanoprobes in nanomedicine.

#### 4b.5. Experimental procedure:

##### 4b.5.1 General procedure for glycan conjugation on tripod:

Respective glycan (3.5 eq) and pentafluorophenol tripod (1 eq) were dissolved in anhydrous DMF (400  $\mu$ l) and triethylamine was added (3 eq per PFP active ester). The resulting mixture was stirred at room temperature. After 5 h the reaction mixture was concentrated and purified by bond elute column (water as eluent) followed by lyophilisation to afford corresponding glycan conjugated tripod.

**Synthesis of CS0S tripod:** The general procedure for glycan conjugation on tripod was followed to afford **CS0S** (1.1 mg, 45 %) as a white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  4.52 (d,  $J = 8.5$  Hz, 1H), 4.41 (t,  $J = 7.8$  Hz, 2H), 4.27 (d,  $J = 3.2$  Hz, 1H), 3.99 – 3.47 (m, 30H), 3.41 – 3.26 (m, 5H), 3.20 – 3.00 (M, 7H), 2.57 – 2.39 (m, 8H), 2.16 (q,  $J = 8.2, 7.8$  Hz, 5H), 2.03 – 1.85 (m, 7H), 1.71 – 1.64 (M, 2H), 1.58 – 1.41 (m, 8H), 1.36 – 1.31 (m, 2H), 1.27 – 1.23 (m, 2H). HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}+\text{H}]^+$   $\text{C}_{75}\text{H}_{124}\text{O}_{44}\text{N}_8\text{S}_2$ : 1906.5031; Found: 1906.5028.

**Synthesis of CS3S tripod:** The general procedure for glycan conjugation on tripod was followed to afford **CS3S** (1.5 mg, 47 %) as a white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  4.55 (d,  $J = 8.4$  Hz, 1H), 4.45 – 4.37 (m, 2H), 4.35 – 4.26 (m, 2H), 4.19 (dd,  $J = 17.6, 3.2$  Hz, 2H), 4.06 (dd,  $J = 11.0, 8.4$  Hz, 1H), 3.99 – 3.92 (m, 2H), 3.87 – 3.83 (m, 2H), 3.80 – 3.48 (m, 18H), 3.38 – 3.26 (m, 4H), 3.19 – 3.06 (m, 4H), 2.55 (t,  $J = 5.9$  Hz, 2H), 2.46 – 2.38 (m, 4H), 2.19 – 2.12 (m, 3H), 2.01 – 1.85 (m, 7H), 1.71 – 1.64 (m, 2H), 1.60 – 1.42 (m, 6H), 1.37 – 1.29 (m, 2H), 1.25 – 1.20 (m, 3H). HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}]^{3-}$   $\text{C}_{75}\text{H}_{121}\text{O}_{53}\text{N}_8\text{S}_5$ : 713.8546; Found: 713.8541.

**Synthesis of CS6S tripod:** The general procedure for glycan conjugation on tripod was followed to afford **CS6S** (0.9 mg, 50 %) as a white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  4.54 (d,  $J = 8.5$  Hz, 1H), 4.46 (dd,  $J = 11.2, 8.2$  Hz, 2H), 4.27 (d,  $J = 3.3$  Hz, 1H), 4.16 (dd,  $J = 6.3, 3.1$  Hz, 2H), 3.97 – 3.79 (m, 7H), 3.78 – 3.51 (m, 18H), 3.40 – 3.26 (m, 5H), 3.25 – 3.04 (m, 7H), 2.56 (t,  $J = 5.9$  Hz, 3H), 2.47 – 2.38 (m, 4H), 2.19 – 2.11 (m, 4H), 2.02 – 1.86 (m, 7H), 1.71 – 1.64 (m, 2H), 1.59 – 1.52 (m, 3H), 1.46 (dt,  $J = 20.5, 8.1$  Hz, 5H), 1.36 – 1.31 (m, 2H), 1.26 – 1.21 (m, 2H). HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}]^{3-}$   $\text{C}_{75}\text{H}_{121}\text{O}_{53}\text{N}_8\text{S}_5$ : 713.8546; Found: 713.8549.

**Synthesis of CS4,6S tripod:** The general procedure for glycan conjugation on tripod was followed to afford **CS4,6S** (1.2 mg, 52 %) as a white solid.  $^1\text{H}$  NMR (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  4.60 (d,  $J = 8.1$  Hz, 1H), 4.47 (d,  $J = 8.1$  Hz, 2H), 4.31 – 4.13 (m, 4H), 4.04 (d,  $J = 8.2$  Hz, 1H), 3.99 – 3.80 (m, 5H), 3.77 – 3.56 (m, 13H), 3.31 (t,  $J = 18.8$  Hz, 4H), 3.21 – 3.06 (m, 7H), 2.60 – 2.52 (m, 3H), 2.43 (dt,  $J = 12.7, 5.6$  Hz, 4H), 2.16 (q,  $J = 7.6$  Hz, 4H), 1.99 – 1.88 (m, 7H), 1.68 (s, 2H), 1.58 – 1.42 (m, 7H), 1.38 – 1.26 (m, 5H). HRMS (ESI)  $m/z$ : calc'd for  $[\text{M}]^{6-}$   $\text{C}_{75}\text{H}_{118}\text{O}_{62}\text{N}_8\text{S}_8$ : 396.4021; Found: 396.4027.



## 4b.5 References:

1. L. Kjellen and U. Lindahl, *Curr. Opin. Struct. Biol.*, 2018, **50**, 101-108.
2. W. Jin, F. Zhang and R. J. Linhardt, *Systems Microbiology and Biomanufacturing*, 2021, **1**, 123-130.
3. L. Pan, C. Cai, C. Liu, D. Liu, G. Li, R. J. Linhardt and G. Yu, *Curr. Opin. Biotechnol.*, 2021, **69**, 191-198.
4. X. Zhang, L. Lin, H. Huang, and R. J. Linhardt, *Acc. Chem. Res.*, 2020, **53**, 335–346.
5. J. Gottschalk and L. Elling, *Curr. Opin. Chem. Biol.*, 2021, **61**, 71-80.
6. L. Djerbal H. Lortat-Jacob and J. C. F. Kwok, *Glycoconj J.*, 2017, **34**, 363–376.
7. K. Sugahara and T. Mikami, *Curr. Opin. Struct. Biol.*, 2007, **17**, 536-545.
8. J. Shin, E. H. Kang, S. Choi, E. J. Jeon, J. H. Cho, D. Kang, H. Lee, I. S. Yun and S. Cho, *ACS Biomater. Sci. Eng.*, 2021, **7**, 4230–4243.
9. V. P. Swarup, T. W. Hsiao, J. Zhang, G. D. Prestwich, B. Kuberan and V. Hlady, *J. Am. Chem. Soc.*, 2013, **135**, 13488–13494.
10. J. Egea, A.G. García, J. Verges, E. Montell and M.G. López, *Osteoarthr. Cartil.*, 2010, **18**, 24-27.
11. C. P. Mencio, R. K. Hussein, P. Yu and H. M. Geller, *J. Histochem. Cytochem.*, 2021, **69**, 61-80.
12. C. Malavaki and S. Mizumoto, *Connect. Tissue Res.*, 2008, **49**, 133-139.
13. A. V. Maksimenko and R. S. Beabealashvili, *Russ. Chem. Bull.*, 2018, **67**, 636-646.
14. J. E. Sadler, S. R. Lentz, J. P. Sheehan, M. Tsiang and Q. Wu, *Haemostasis*, 1993, **23**, 183-193.
15. R. Sasisekharan, R. Raman, and V. Prabhakar, *Annu. Rev. Biomed. Eng.*, 2006, **8**, 181–231.
16. A. Dityatev, G. Bruckner, G. Dityateva, J. Grosche, R. Kleene, M. Schachner, *Periodicals, Inc. Develop Neurobiol.*, 2007, **67**, 570–588.
17. S. Hockfield, R.G. Kalb, S. Zaremba, and H. Fryer, *Cold Spring Harb. Symp. Quant. Biol.*, 1990, **55**, 505-514.
18. B. A. Sorg, S. Berretta, J. M. Blacktop, J. W. Fawcett, H. Kitagawa, J. C.F. Kwok, and M. Miquel, *J. Neurosci.*, 2016, **36**, 11459–11468.

19. S. Miyata, Y. Komatsu, Y. Yoshimura, C. Taya and H. Kitagawa, *Nat. Neurosci.*, 2012, **15**, 414-424.
20. F. R. Lin, *J Gerontol A Biol Sci Med Sci.*, 2011, **10**, 1131–1136.
21. E. Boisselier and D. Astruc, *Chem. Soc. Rev.*, 2009, **38**, 1759–1782.
22. S. Rana, C. E. Powe, S. Salahuddin, S. Verlohren, F. H. Perschel, R. J. Levine, K. Lim, J. B. Wenger, R. Thadhani, and S. A. Karumanchi, *Circulation.*, 2012, **125**, 911-919.
23. M. Bortolotti, D. Mercatelli and L. Polito, *Front. Pharmacol.*, 2019, 10, 486.
24. H. J. Garrigues, M. W. Lark, S. Lara, I. Hellstrim, K. E. Hellstrim and T. N. Wight, *J. Cell Biol.*, 1986, **103**, 1699-1710.
25. A. E. Faassen, J. A. Schrage, D. J. Klein, T. R. Oegema, J. R. Couchman and J. B. McCarthy, *J. Cell Biol.*, 1992, **116**, 521-531.
26. C. A. Henke, U. Roongta, D. J. Mickelson, J. R. Knutson and J. B. McCarthy, *J. Clin. Invest.*, 1996, **97**, 2541-2552.
27. N. Malik, B. W. Greenfield, A. F. Wahl and P. A. Kiener, *J. Immunol.*, 1996, **156**, 3952-3960.
28. E. J. Bradbury, L. D. F. Moon, R. J. Popat, V. R. King, G. S. Bennett, P. N. Patel, J. W. Fawcett and S. B. McMahon, *Nature*, 2002, **416**, 636-640.
29. K. N. Sugahara, T. Teesalu, P. P. Karmali, V. R. Kotamraju, L. Agemy, O. M. Girard, D. Hanahan, R. F. Mattrey, and E. Ruoslahti, *Cancer Cell*, 2009, **16**, 510-520.
30. C. Dou and J. M. Levine, *J. Neurosci.*, 1995, **15**, 8053-8066.
31. N. Khidekel, S. B. Ficarro, E. C. Peters, and L. C. Hsieh-Wilson, *Proc. Natl. Acad. Sci. U.S.A.*, 2004, **101**, 13132-13137.
32. D. Wang and J. Fawcett, *Cell Tissue Res.*, 2012, **349**, 147–160.

## 4b.6 NMR Spectra

