

**GENERAL APPROACHES TO CAGING
SULFATION IN BIOMOLECULES**

by

CHAO LIU

A dissertation
submitted to the Faculty of
the department of Chemistry
in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Boston College
Morrissey School of Arts and Sciences
Graduate School

May 2023

GENERAL APPROACHES TO CAGING SULFATION IN BIOMOLECULES

Chao Liu

Thesis Advisor: Jia Niu, Ph.D.

Abstract

O-Sulfation is an important chemical code existing widely in nature, participating in a variety of biological activities including immune response, hemostasis, hormone regulation, cell signaling, and viral invasion. The heterogeneous nature, high polarity with negative charge, and the chemical lability of the sulfate modification have created significant challenges in the synthesis and structure–function studies of *O*-sulfated biomolecules. It is therefore highly desirable to achieve caging and selective release of the *O*-sulfated biomolecules. Inspired by sulfur (VI) fluoride exchange reaction, our group developed a series of general approaches to caged *O*-sulfated biomolecules and their selective deprotection. First, an *O*-sulfation strategy is developed by coupling aromatic fluorosulfate with silylated target molecules. Scalable synthesis was demonstrated on monosaccharides, disaccharides, amino acid, and steroid. Selective hydrolytic and hydrogenolytic removal of the aryl masking groups yielded the corresponding *O*-sulfated products in excellent yields. Furthermore, a complete knowledge gap was noted in biocompatible caging of sulfate. With the rational design and systematic optimizations, we discovered that fluorosulfotyrosine in peptides and proteins was an ideal precursor for sulfotyrosine (sY), which can be efficiently converted into the anionic active form by hydroxamic acid activators under physiologically relevant conditions. Photocaging the hydroxamic acid activators further allowed for light-controlled activation of functional sulfopeptides. This system featuring fast kinetics, high selectivity, excellent robustness, and on-demand release provides a valuable tool for probing functional roles of sulfation in the peptides and proteins.

Dedication

I dedicate this thesis to my parents,
Junyi Liu and Xianju Du,
as well as my grandma Xuping Xiao.

Acknowledgements

I can still clearly recall the excitement when I got the offer from Boston College and the moment when I arrive here. Time flies so quick that I am now reaching to the last lap in my journey pursuing a PhD degree. All I experienced and learned throughout these five year has become an indispensable part of myself and will benefit me for the rest of life. I feel so lucky to have the wonderful mentors, friends, and family around me and support me through this invaluable process.

First, I want to express my deep appreciation to my advisor, Prof. Jia Niu. Jia is such a good mentor in academic research and a good friend in life. When I first got here, Jia provided a very welcoming environment and took care of me in and out of the lab. I can't thank him enough for the help. In our group, you can always feel safe to come up with naïve questions, doubts, and even disagreement. I remember the frustration when I failed the first proposal defense and how he calmed me down by pointing out what's wrong with my proposal and how I can improve. That became a wake-up call for me. With a plan to learn more about biology like auditing a Cell Biology class, Jia has been nothing but supportive to me. Along the five years' scientific exploration with Jia, the lab is always open-minded and aiming for excellence in the challenging topics. Jia inspired me to work as an academic researcher and showed me how to be a good one.

The collaboration environment in our department has been a valuable fortune to me. I can always stop by any professor's office seeking their wisdom for the doubts and questions. Specifically, I want thank Prof. Eranthie Weerapana and Abhishek Chatterjee for being my committee in both my proposal and thesis defense. Their critical spirits in scientific research, thoughtful questions, and heart-warming help made me aware of their strong academic commitment and encouraged me to adopt chemical tools into uncharted biology. I deeply appreciate that Prof. Chatterjee for kindly being the reference in my job application. And I want to thank Prof. Jianmin Gao and Jeffery B. Byers for the

illuminating discussion. I also thank Prof. Lin Cheng from Boston University for the mass spectrometry identification.

All the achievement I have made could not be done without the help and support from my friends and colleagues. I want to thank good friends Haiyang Zhao, Cangjie Yang, Xiang Ma, Wei Li, Mi Zhou, Lianqian Wu, Qian Zhu, Yichuan Zhang, Liming Qi, Qiwen Su, Zefeng Zhou, Yu Deng, Mengmeng Zheng, Wenqi Wang, Shibo Xu, Kyrsten Weissheier, Brayan Rondon, Jingchen Yang, Jiangwei Liu, Quan Pham, Jennifer Lin, Xiaonan Li, Rong Chen and many more for the heart-warming support through the blue days. I appreciate the collaboration and helpful discussion from Cangjie Yang, Xueyi Liu, Mi Zhou, Chaoshuang Xia, Chintan Soni, Yuhuan Lyu, Jiayi Fan, Conor Lyond, Fan Zhang, Zeyi Huang, and Soumya Jyoti Singha Roy.

Last but not least, I am so grateful to my family. As the first-generation college student in my family, my parents don't really teach me what to do but show me to be an honest person with integrity in their actions. They give me the very life and their selfless support is always the safest harbor for me. My grandma passed away in the same year I enrolled BC. As her heritage in the living world, I will cherish the opportunity the life given and carry on the journey with the happy memories.

Table of Contents

| | |
|---|-----|
| Abstract | III |
| Dedication..... | IV |
| Acknowledgements..... | V |
| List of Figures and Tables | IX |
| List of Abbreviations | XIX |
| Chapter 1 | 1 |
| INTRODUCTION | 1 |
| 1.1 The Significance of Sulfation in Nature | 2 |
| 1.2 Sulfation Detection Method | 9 |
| 1.2.1 MS..... | 9 |
| 1.2.2 Anti-Sulfotyrosine Antibody..... | 13 |
| 1.2.3 Isotope Labeling..... | 14 |
| 1.2.4 Fluorescent assay | 16 |
| 1.3 Access to Sulfated Biomolecules..... | 17 |
| 1.3.1 Enrichment from Biomaterials..... | 18 |
| 1.3.2 Synthesis | 21 |
| 1.4 Our work..... | 29 |
| 1.5 References..... | 31 |
| Chapter 2..... | 37 |
| A General Approach to <i>O</i> -Sulfation via Sulfur(VI) Fluoride Exchange Reaction | 37 |
| 2.1 INTRODUCTION..... | 38 |
| 2.1.1 Overview..... | 38 |
| 2.1.2 Significance of sulfation and major challenge..... | 38 |
| 2.1.3 Current approaches | 39 |
| 2.1.4 Our strategy..... | 41 |
| 2.2 RESULTS AND DISSCUSSION | 41 |
| 2.2.1 Initial screening with template reaction | 41 |
| 2.3 CONCLUSION | 64 |
| 2.4 EEPERIMENTAL | 64 |
| 2.4.1 Materials..... | 64 |
| 2.4.2 Characterization..... | 65 |
| 2.4.3 General Procedures..... | 65 |

| | |
|---|-----|
| 2.4.4 Detailed Experimental Procedure..... | 68 |
| 2.5 REFERENCES | 127 |
| Chapter 3..... | 131 |
| Fluorosulfate as a Latent Sulfate in Peptides and Proteins..... | 131 |
| 3.1 INTRODUCTION..... | 132 |
| 3.1.1 Overview..... | 132 |
| 3.1.2 Significance of tyrosine-O-sulfation and current challenge..... | 132 |
| 3.1.3 Our Solution..... | 133 |
| 3.2 RESULTS AND DISCUSSION | 135 |
| Figure 3-22..... | 158 |
| Figure 3-24..... | 159 |
| 3.3 CONCLUSION | 162 |
| 3.4 EXPERIMENTAL..... | 163 |
| 3.4.1 Material and Equipment..... | 163 |
| 3.4.2 Characterization..... | 163 |
| 3.4.3 General Procedures..... | 166 |
| 3.4.4 Detailed Procedure for Supplementary Results | 170 |
| 3.4.5 Small Molecule Synthesis | 183 |
| 3.4.6 Peptide Synthesis and Decaging..... | 189 |
| 3.4.7 Protein Expression via Unnatural Amino Acid (fsY) Incorporation | 217 |
| 3.5 REFERENCES | 222 |
| Appendix..... | 226 |
| Original NMR spectra of new compounds in Chapter 2..... | 226 |
| Original NMR spectra of new compounds in Chapter 3..... | 297 |

List of Figures and Tables

Figure 1-1. Post-translational modifications are key mechanisms to increase proteomic diversity

Figure 1-2. Significance of sulfate modification

Figure 1-3. Sulfation pathway in human cells

Figure 1-4. Sulfated sugar on different positions and skeletons

Figure 1-5. Sulfated glycopeptides, proteoglycans, and their functions

Figure 1-6. Sulfotyrosyl proteins and function

Figure 1-7. Function of sulfotyrosine-containing proteins

Figure 1-8. Br-tag installation of phosphotyrosine and sulfotyrosine residues

Figure 1-9. Antisulfotyrosine antibody development

Figure 1-10. Fluorescent assay probing sulfation pathway

Figure 1-11. Schematic representation of an industrial heparin purification process

Figure 1-12. Enrichment for analysis systems

Figure 1-13. WAX system for sulfotyrosylpeptide enrichment

Figure 1-14. Chemical synthesis of sulfated glycan

Figure 1-15. Solution-phase synthesis of sulfated glycan

Figure 1-16. On-resin synthesis of KS oligosaccharides

Figure 1-17. Site-selective SPPS of sulfated CCR5 library

Figure 1-18. Chemoenzymatic synthesis of glycosaminoglycans

Figure 1-19. Chemoenzymatic synthesis of ULMW heparin

Figure 1-20. Direct sulfoprotein expression via ncAA mutagenesis

Figure 1-21. Our work

Figure 2-1. SuFEx-based rapid access to sulfated molecules

Figure 2-2. Diverse bioactive natural products consist of *O*-sulfation

Figure 2-3. Existing approaches to access *O*-sulfation

Figure 2-4. Our design: early-stage *O*-sulfation via SuFEx

Table 2-1. Preparation of phenyl fluorosulfates with various electron-withdrawing and electron-donating substitutions

Table 2-2. Screening reaction factors for the coupling of **2-1** with **2-2a**

Table 2-3. The coupling of **1** with various fluorosulfate

Table 2-4. Catalyst screening for the coupling of **2-2i** with **2-1**

Table 2-5. Hydrolysis of **2-3a**

Table 2-6. Hydrolysis of a variety of sulfate diester **2-3(a-j)**

Figure 2-6. Effect of aromatic substitutions on the kinetics of the hydrolysis of sulfate diesters

Figure 2-7. Proposed mechanism of hydrolysis

Table 2-7. Optimization of the hydrolysis of **2-3a**

Table 2-8. Hydrolysis of sulfate diester **2-3(a-j)**

Figure 2-8. GC-MS analysis of the hydrogenolysis of **2-3h**

Figure 2-9. GC-MS analysis of the hydrogenolysis of sulfate diesters with various substitutions

Figure 2-10. GC-MS analysis of the hydrogenolysis of **2-3g**

Figure 2-11. Our proposed mechanism of hydrogenolysis

Figure 2-12. ¹H-NMR spectra of **4** from hydrolysis and hydrogenolysis process

Figure 2-13. Substrate scope of the established method

Figure 2-14. Side reactions in the SuFEx coupling to synthesize **2-10** and **2-12a**

Figure 2-15. Compatibility of aryl sulfate diesters to common reagents used in carbohydrate and peptide chemistries

Table 2-9. Robustness screening for tyrosine sulfate diester

Figure 2-16. Glycosylation coupling with the sulfate masks on

Figure 2-17. Decaging via hydrogenolysis

Figure 2-18. NMR quantification of disulfated idose **2-37** with β -glycerophosphate disodium tetrahydrate

Figure 3-1. Sulfation widely exists in diverse bioactive peptides and proteins

Figure 3-2. Our solution

Table 3-1. Initial exploration of the release conditions of fsY

Table 3-2. Stability test of fluorosulfotyrosine (fsY) and fluorosulfohexapeptide **1** stability in cell lysate

Figure 3-3. Analytical RP-HPLC verified the stability of fsY and **3-1** in cell lysate. See the caption in next page.

Figure 3-4. HPLC analysis of the yield of the decaging of **1** shown in Table 3-1

Table 3-3. Decaging reaction mediated by NHS in Tris buffer

Table 3-4. The pH dependence of the fsY decaging reaction mediated by NHS

Figure 3-5. Hydrolysis of NHS in deuterated water by titrating NaOH

Figure 3-6. A variety of hydroxamic acid activators were investigated for their ability to activate fluorosulfate of model peptide **3-1**

Table 3-5. Decaging of fsY in **3-1** mediated by silyl-protected activator **3-11**

Figure 3-7. Extended activator screening for the fsY decaging using **3-1** as the substrate

Figure 3-8. Analytical HPLC trace of the fsY decaging mediated by silyl-protected activators

Table 3-6. Compatibility test of the decaging reaction of **1** in the presence of free amino acids

Figure 3-9. Real-time LC-MS analysis to probe reaction mechanism

Figure 3-10. Mass spectra of H₂¹⁸O isotope labeling experiment

Figure 3-11. No ¹⁸O-labeled products were found from the reaction in H₂¹⁸O buffer

Figure 3-12. Real-time LC-MS reaction monitoring identified two activator **3-7** adducts **3-12** and **3-13**, suggesting a Lossen rearrangement mechanism

Table 3-7. Fluorosulfate activation and decaging in fsY-containing synthetic peptides

Figure 3-13. Light-mediated release of C5a receptor 22mer peptide from caged activator **3-18**

Figure 3-14. Compatibility of semicarbazide with fsY-containing substrates **3-1**

Figure 3-15. TTI peptide sequences and sulfation patterns

Figure 3-16. Thrombin inhibition by in situ decaged TTI peptides

Figure 3-17. Light-mediated activation and decaging of fluorosulfate-containing TTI peptide

TTI04(fsY) regulates its sulfation-dependent thrombin inhibitory activity

Figure 3-18. Tandem MS analysis of sfGFP-151-fsY after trypsin digestion and alkylation with iodoacetamide. X in the sequence logo represents fsY

Figure 3-19. High-resolution Orbitrap mass spectrometry analysis of the decaging of sfGFP-151-fsY

Figure 3-20. Fluorosulfate activation and decaging in fsY-containing protein sfGFP-151-fsY and its corresponding high-resolution mass spectrometry

Figure 3-21. sfGFP-151-fsY and sfGFP-151-sY remained intact in decaging conditions by Anti-His western blot

Figure 3-22. Fluorescence intensity of sf-GFP-151-fsY after being treated with various conditions

Figure 3-23. Plating assay to determine *E. coli* (BL21) viability under various fluorosulfate decaging conditions

Figure 3-24. Free peptide 1 decaging in the presence of live *E. coli* cells to mimic the soluble sulfated small molecules and peptides

Figure 3-25. Cytocompatibility of the reagents

Figure 3-26. Fluorescent micrograph of the *S. aureus* cells before and after ligation of the 5(6)-carboxyfluorescein (FAM)-labeled peptide **3-20**

Figure 3-27. Flow cytometry of the *S. aureus* cells before and after ligation of the FAM-labeled peptide **3-20**

Figure 3-28. Plating assay to determine *E. coli* (BL21) viability under the decaging reaction

Figure 3-29: Analytical RP-HPLC (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~33% CH₃CN from 0-20 min, 20-21 min 33%~95% CH₃CN, and 95% CH₃CN washing for another 2 min) chromatogram of purified peptide **DADEfsYL-NH₂ (3-1)**

Figure 3-30. ESI-TOF mass spectrum for peptide **DADEfsYL-NH₂ (3-1)**

Figure 3-31: Analytical RP-HPLC (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~33% CH₃CN from 0-20 min, 20-21 min 33%~95% CH₃CN, and 95% CH₃CN washing for another 2 min) chromatograms of the decaging reaction mixture (middle) and purified peptide **DADEsYL-NH₂ (3-2, bottom)**

Figure 3-32. ESI-TOF mass spectrum for peptide **DADEsYL-NH₂ (3-2)**

Figure 3-33. Analytical RP-HPLC (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~60% CH₃CN over 40 min) chromatogram of purified peptide **fsYEfsYLDfsYDF-NH₂ (3-14)**

Figure 3-34. ESI-TOF mass spectrum for peptide **fsYEfsYLDfsYDF-NH₂ (3-14)**

Figure 3-35: Analytical RP-HPLC (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~60% CH₃CN over 40 min) chromatogram of reaction mixture after 3 h and passed through a PD-10 column (middle) purified peptide **sYEsYLDsYDF-NH₂ (3-15, bottom)**

Figure 3-36. ESI-TOF mass spectrum for peptide **sYEsYLDsYDF-NH₂ (3-15)**

Figure 3-37. ESI-TOF mass spectrum for peptide **DDKDTLNLNTPVVK-NH₂**

Figure 3-38. Analytical RP-HPLC chromatogram (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~60% CH₃CN over 40 min) of purified peptide **Biotin-TTPDsYGHfsYDDKDTLNLNTPVVK-NH₂ (3-16)**

Figure 3-39. ESI-TOF mass spectrum for peptide **Biotin-TTPDsYGHfsYDDKDTLNLNTPVVK-NH₂ (3-16)**

Figure 3-40. Analytical RP-HPLC chromatogram of (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~60% CH₃CN over 40 min) purified reaction mixture (middle) peptide **Biotin-TTPDsYGHsYDDKDTLNLNTPVVK-NH₂ (3-17, bottom)**

Figure 3-41. ESI-TOF mass spectrum for peptide **Biotin-TTPDsYGHsYDDKDTLNLNTPVVK-NH₂ (3-17)**

Figure 3-42. Analytical RP-HPLC chromatogram of **a)** photolytic reaction mixture (middle, CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~33% CH₃CN from 0-20 min, 20-21 min 33%~95% CH₃CN, and 95% CH₃CN washing for another 2 min, monitored under 230 nm wavelength) and **b)** comparison of purified **3-17** from free (top) or photocaged (bottom) activator (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~60% CH₃CN over 40 min, monitored under 230 nm wavelength)

Figure 3-43. ESI-TOF mass spectrum for peptide **Biotin-**

TTPDsYGHsYDDKDTLNLNTPVDK-NH₂ (3-17) from light-mediated decaging reaction

Figure 3-45. Analytical RP-HPLC chromatogram of (A (95% H₂O, 5% CH₃CN, and 0.1% formic acid): B (5% H₂O, 95% CH₃CN, and 0.1% formic acid) as mobile phase, linear gradient from 5%~95% B over 12 min) purified peptide **TTI01**

Figure 3-46. ESI-TOF mass spectrum for peptide **TTI01**

Figure 3-47. Analytical RP-HPLC chromatogram of (A (95% H₂O, 5% CH₃CN, and 0.1% formic acid): B (5% H₂O, 95% CH₃CN, and 0.1% formic acid) as mobile phase, linear gradient from 5%~95% B over 12 min) purified peptide **TTI02(fsY)**

Figure 3-48. ESI-TOF mass spectrum for peptide **TTI02(fsY)**

Figure 3-49. Analytical RP-HPLC chromatogram of (A (95% H₂O, 5% CH₃CN, and 0.1% formic acid): B (5% H₂O, 95% CH₃CN, and 0.1% formic acid) as mobile phase, linear gradient from 5%~95% B over 12 min) purified peptide **TTI03(fsY)**

Figure 3-50. ESI-TOF mass spectrum for peptide **TTI03(fsY)**

Figure 3-51. Analytical RP-HPLC chromatogram of (A (95% H₂O, 5% CH₃CN, and 0.1% formic acid): B (5% H₂O, 95% CH₃CN, and 0.1% formic acid) as mobile phase, linear gradient from 5%~95% B over 12 min) purified peptide **TTI04(fsY)**

Figure 3-52. ESI-TOF mass spectrum for peptide **TTI04(fsY)**

Figure 3-53. Analytical RP-HPLC chromatogram of **a)** decaging reaction mixture (bottom, CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~35% CH₃CN over 20

min) and **b**) purified **TTI02(sY)** (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~35% CH₃CN over 30 min)

Figure 3-54. ESI-TOF mass spectrum for peptide **TTI02(sY)**

Figure 3-55. Analytical RP-HPLC chromatogram of **a**) decaging reaction mixture (bottom, CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~35% CH₃CN over 20 min) and **b**) purified **TTI02(sY)** (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~35% CH₃CN over 30 min)

Figure 3-56. ESI-TOF mass spectrum for peptide **TTI03(sY)**

Figure 3-57. Analytical RP-HPLC chromatogram of **a**) decaging reaction mixture (bottom, CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~35% CH₃CN over 20 min) and **b**) purified **TTI04(sY)** (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~35% CH₃CN over 30 min)

Figure 3-58. ESI-TOF mass spectrum for peptide **TTI04(sY)**

Figure 3-59. Analytical RP-HPLC chromatogram of purified fluorosulfopeptide with fluorescent label **3-20**

Figure 60. ESI-TOF mass spectrum for peptide **DfsYDHfsYDDKGSO(FAM)SENLYFQGSLPETGGS-NH₂ (22)**

Figure 3-61. Normalized sfGFP fluorescence. Left: sfGFP-wt expression in the presence/absence of IPTG; Right: sfGFP-151-fsY expression in the presence of IPTG and in the presence/absence of fsY

Figure 3-62. SDS-PAGE analysis of purified sfGFP-wt (lane 1), sfGFP-151-sY (lane 2) and sfGFP-151-fsY (lane 3)

List of Abbreviations

Standard 3- letter and 1-letter codes are used for the 20 natural amino acids

Human Genome Project (HGP)

post-translational modification (PTM)

sulfur (S)

adenosine 5'-phosphosulfate (APS)

adenosine triphosphate sulfurylase (ATPS)

adenosine triphosphate (ATP)

3'-phosphoadenosine-5'-phosphosulfate synthase (PAPSS)

3'-phosphoadenosine-5'-phosphosulfate (PAPS)

PAPSS1 (3'-phosphoadenosine 5'-phosphosulfate synthase 1)

PAPSS2 (3'-phosphoadenosine 5'-phosphosulfate synthase 2)

glycosaminoglycans (GAGs)

Basic fibroblast growth factor (bFGF)

heparan sulfate proteoglycans (HSPGs)

spike protein (protein S)

angiotensin converting enzyme 2 (ACE2)

phosphotyrosyl (pY)

sulfotyrosine (sY)

complement component 5a receptor (C5aR)

Chemotaxis Inhibitory Protein of Staphylococcus aureus (CHIPS)

complement component 5a (C5a)

spectrometry (MS)

chondroitin sulfate (CS)

dermatan sulfate (DS)

decorin (DCN)

reversed-phase high-performance liquid chromatography (RP-HPLC)

glycosaminoglycan (GAGs)

human leukocyte antigen (HLA)

matrix-assisted laser/desorption ionization time-of-flight (MALDI-TOF)

hydrogen bromide (HBr)

Src Homology 2 (SH2)

Normal rat kidney epithelial (NRK)

sodium dodecyl sulfate (SDS)

polyvinylidene difluoride (PVDF)

polyacrylamide gel electrophoresis (PAGE)

fourth component of human complement (C4)

human tyrosylprotein sulfotransferase (hTPST)

L-2,3-diaminopropionic acid (L-DAP)

pyrene (Pyr)

tyrosylprotein sulfotransferase (TPST)

phenol sulfotransferase (PST)

4-methylumbelliferone (MU)

P-selectin glycoprotein ligand-1 (PSGL-1)

Molecular weight (MW)

immobilized metal ion affinity chromatography (IMAC-Ga)

thyroid stimulating hormone (TSH)

bovine luteinizing hormone (bLH)

sulfate emerging (SE)

strong anion exchange chromatography (SAX)

bovine luteinizing hormone (bLH)

liquid chromatography/electrospray ionization mass spectrometry (LC/ESI-MS)

weak anion exchange (WAX)

ultraviolet photodissociation (UVPD)

automated glycan assembly (AGA)

keratin sulfate (KS)

heparan sulfate (HS)

solid-phase peptide synthesis (SPPS)

C-C chemokine receptor type 5 (CCR5)

o-nitrobenzyl (o-NB)

sulfur (VI) fluoride exchange (SuFEx)

1,8-diazabicyclo[5.4.0]undec-7-ene (DBU)

1,5,7-triazabicyclo[4.4.0]dec-5-ene (TBD)

1,5-diazabicyclo[4.3.0]non-5-ene (DBN)

1,5,7-triazabicyclo[4.4.0]dec-5-ene (MTBD)

2-tert-butylimino-2-diethylamino-1,3-dimethylperhydro-1,3,2-diazaphosphorine (BEMP)

tert-Butylimino-tri(pyrrolidino)phosphorene (BTTP)

gas chromatography–mass spectrometry (GC-MS)

phosphate buffer saline (PBS)

tris(hydroxymethyl)aminomethane (Tris)

triisopropylsilyl (TIPS)

potassium fluoride (KF)

Chapter 1

INTRODUCTION

1.1 The Significance of Sulfation in Nature

Since the completion of Human Genome Project (HGP), scientists have discovered that the diversities of human proteome are enormously larger than human genome.¹ While most of the human gene estimations so far fall in the range of 25,000-30,000 counts, the total number of proteins is approximately over 1,000,000.² Among all the biological events that account for such discrepancy, post-translational modification (PTM) is a major one accounted for various functional group addition onto amino acid residues (Figure 1-1). To date, there are more than 620 types of PTMs have been identified experimentally which are forged by nature in an enzymatic or non-enzymatic manner.³ Phosphorylation, acetylation, glycosylation, methylation, and ubiquitination are the most abundant by enzymatic addition approaches. Meanwhile, non-enzymatic additions such as glycation, oxidation, citrullination, palmitoylation, and formylation also frequently occur in nature. Moreover, similar PTM incorporation pathways also occur on carbohydrate and nucleotide residues, thereby significantly expanding the biomolecule libraries.

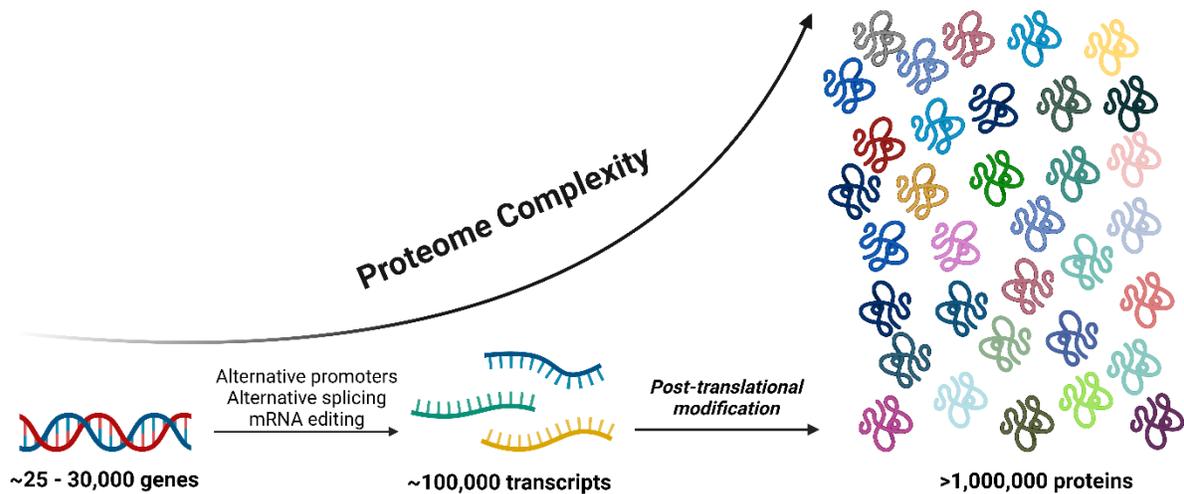


Figure 1-1. Post-translational modifications are key mechanisms to increase proteomic diversity

PTMs are crucial in regulating functional proteomics, influencing biological functions, molecular localization, and interactions between cellular components. These modifications govern essential cellular processes, including cell growth and proliferation, signal transduction, intermediate metabolism,⁴ and a variety of pathological developments like organ diseases (cardiovascular systems, pancreas, and kidney), neurodegeneration, and cancer.⁵ In particular, mature proteins with sulfate modification have been an evolutionary mark of multicellular living systems which participates actively in physiological functions, encompassing cell-cell signaling pathways, viral invasion, immune responses like inflammation, and hemostasis (Figure 1-2).

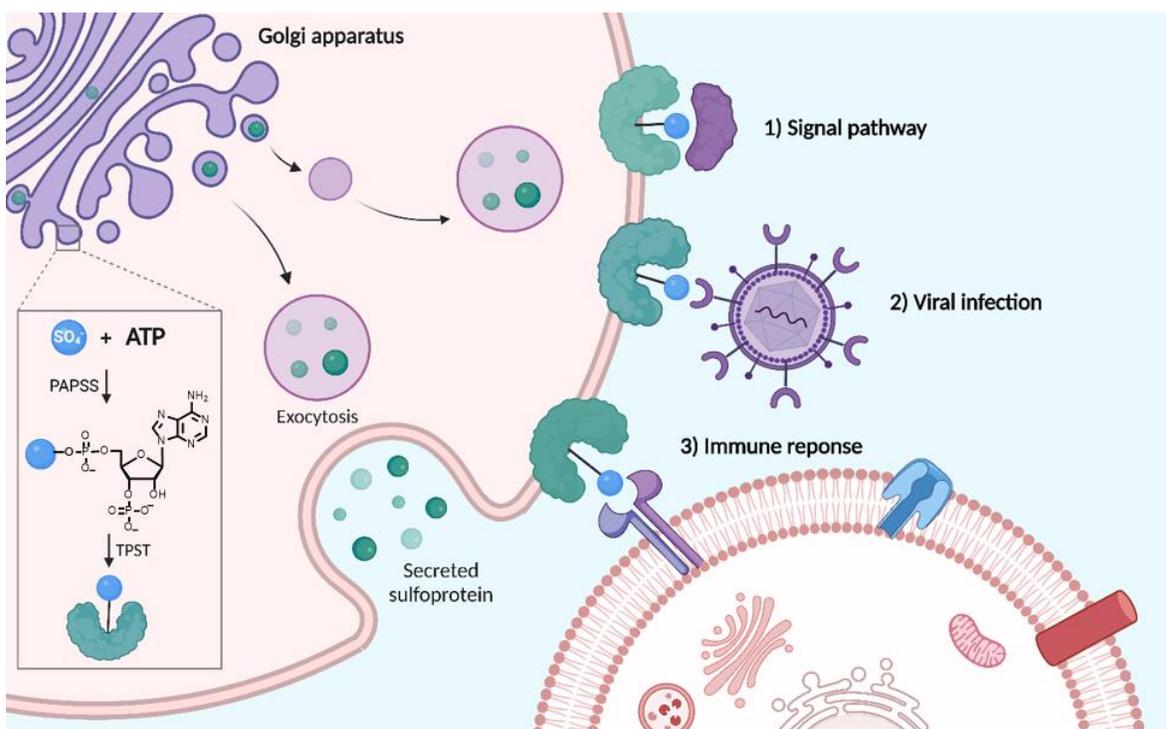


Figure 1-2. Significance of sulfate modification

Sulfate, as the oxidized form of sulfur (S) metabolite in nature, is an active player in sulfur metabolism processes. Once the cell takes in the inorganic sulfate ion, it is first converted to adenosine 5'-phosphosulfate (APS) by adenosine triphosphate sulfurylase (ATPS), and the subsequent phosphorylation of APS at its 3'-OH by APS kinase eventually furnishes the universal sulfate donor

3'-phosphoadenosine-5'-phosphosulfate (PAPS). During this process, the two involved enzymes are either fused into a single dual-function enzyme 3'-phosphoadenosine-5'-phosphosulfate synthase (PAPSS) in animals or present by two separate enzymes in prokaryotes, fungi, and plants.⁶ Till now, two isoforms of PAPSS are identified in human cells: PAPSS1 and PAPSS2.⁷ While PAPSS1 is localized in the cellular nucleus, PAPSS2 is predominantly found in the cytoplasm (Figure 1-3).⁸ Different medical conditions are associated with the deficiencies of the two isoforms. For instance, PAPSS1-deficient A549 spheroids exhibited a 58% reduction in size when treated with low-dose cisplatin.⁹ Additionally, PAPSS2 deficiency was reported to cause androgen excess.¹⁰

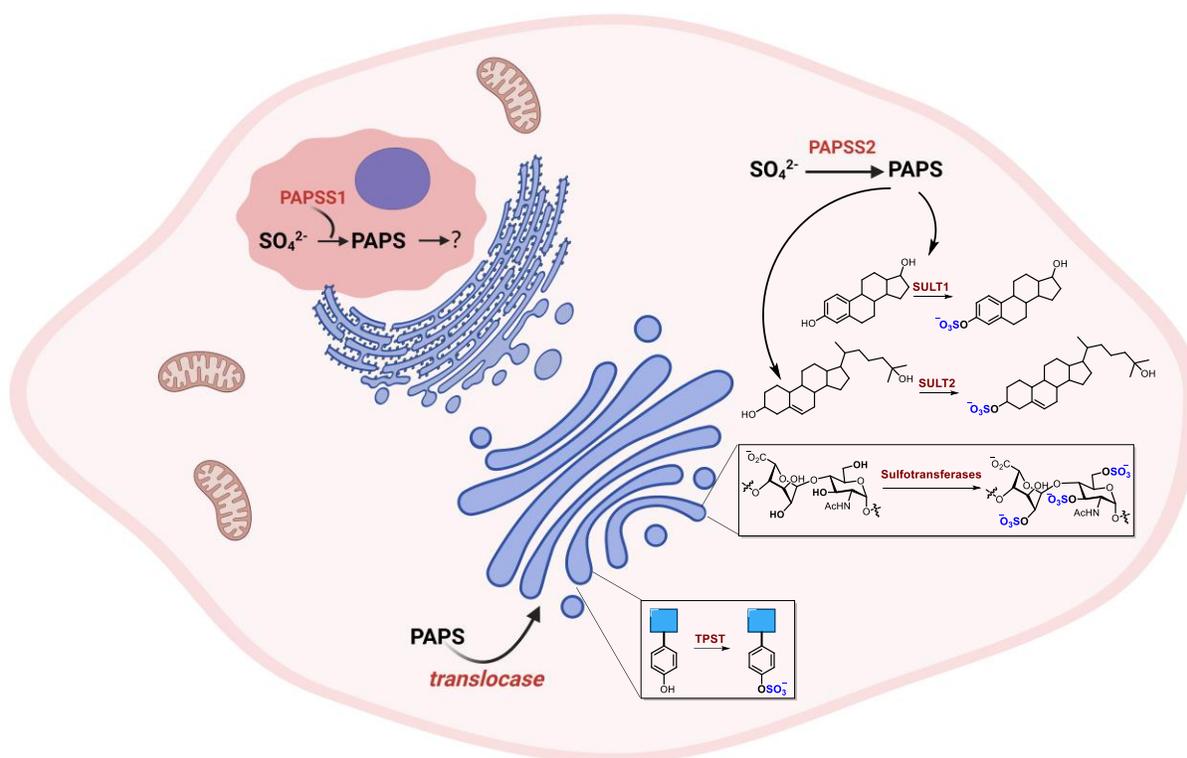


Figure 1-3. Sulfation pathway in human cells

The activated sulfate donor PAPS is utilized by a variety of sulfotransferases to transfer sulfate to the hydroxyl or amino groups of their specific substrates. These transferases can be primarily

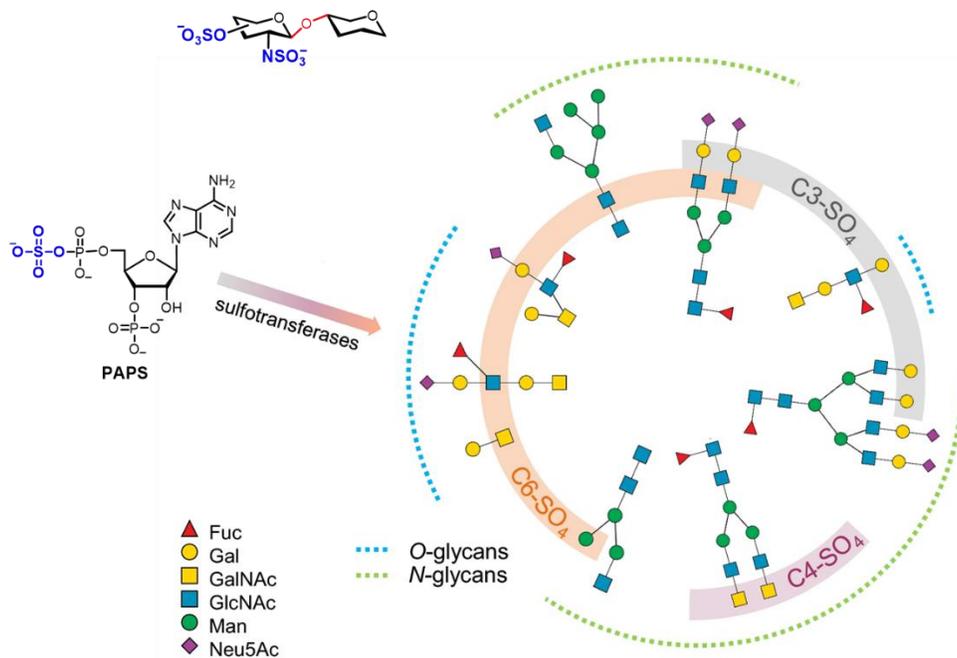


Figure 1-4. Sulfated sugar on different positions and skeletons

categorized into two groups: cytosolic (soluble) and membrane-bound (Golgi-located).¹¹ For the cytosolic ones, it is involved in the sulfation of small endogenous compounds like xenobiotics, neurotransmitters, and hormones which constitute the enzyme superfamily known as SULTs. On the other hand, the membrane-bound sulfotransferases are in charge of the post-translational sulfation of endogenous macromolecules like carbohydrates, lipids, and proteins. Specifically, multiple sulfotransferases such as HS2ST, HS6ST, and HsGAL3ST are discovered, and responsible for sulfation of different sugar positions and skeletons, thus resulting in an extensive with immense diversity (Figure 1-4).¹² The resulting sulfated sugars are found on glycopeptides such as luteinizing hormone LH β participating in the clearance of hormone,¹³ proteoglycans like heparan sulfate-proteoglycans on cell surface playing important roles in molecular recognition,¹⁴ and sulfated glycosaminoglycans (GAGs) like well-known heparin and heparan sulfate, which are actively involved in brain development, blood clotting, and cancer progression¹⁵ (Figure 1-4). Studies have shown that both free (or named cytosolic/soluble) heparin and immobilized heparan sulfate proteoglycans

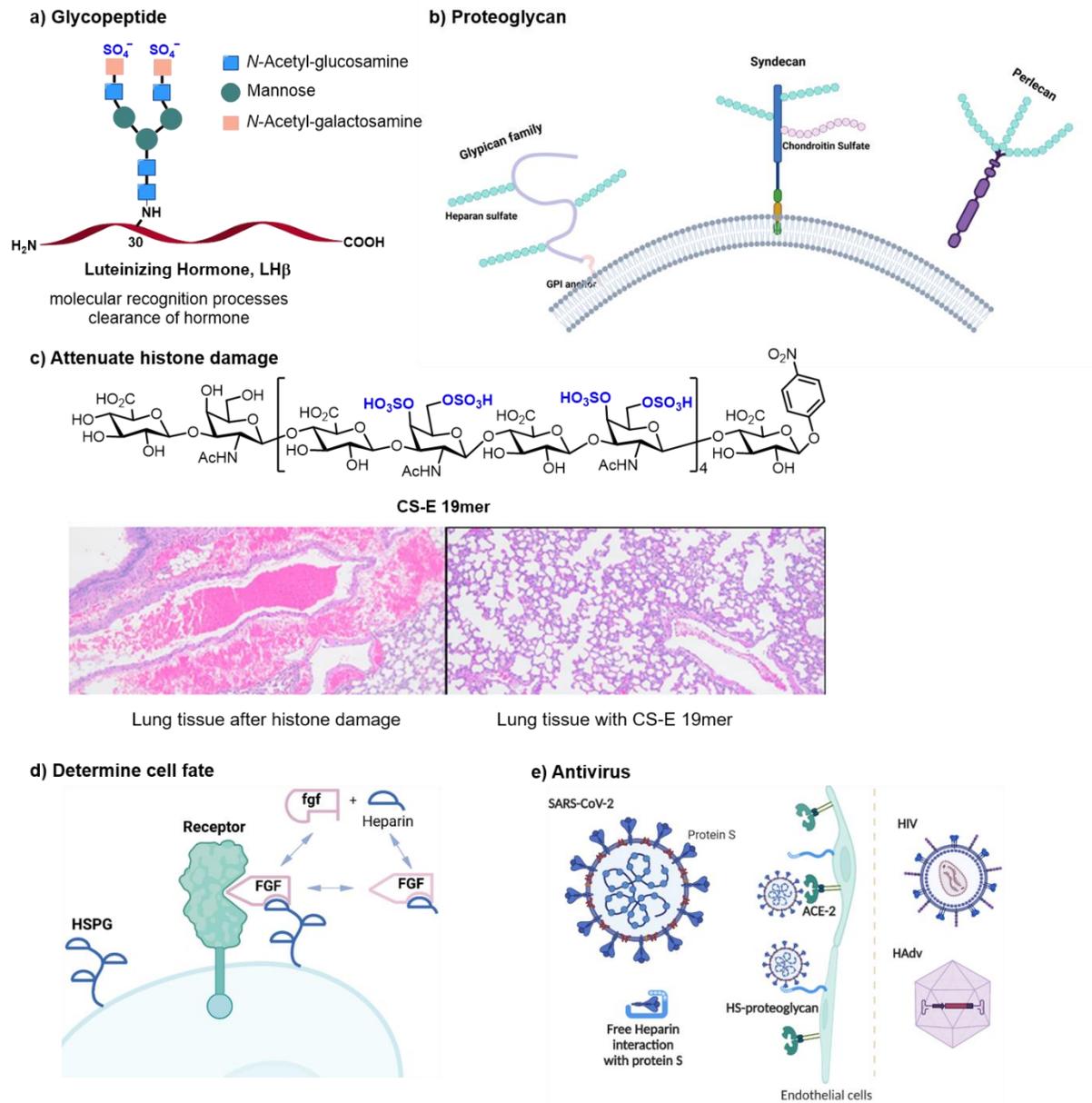


Figure 1-5. Sulfated glycopeptides, proteoglycans, and their functions

(HSPGs) interchangeably participated in reconstituting a stable and receptor-compatible conformation of bFGF, subsequently stimulating cell proliferation, migration, and differentiation.¹⁶⁻¹⁷ Meanwhile, the secreted sulfated glycosaminoglycan polysaccharides are reported to serve as the protecting front line of cells (Figure 1-5). Among them, free chondroitin sulfate-E (CS-E) was reported to attenuate bacteria lipopolysaccharide-induced organ damage (Figure 1-5c).¹⁸ Free heparin

receptor (C5aR) extracellular *N*-terminus with phosphoryl modification (Figure 1-6b), comparable binding affinity and similar binding mode towards Chemotaxis Inhibitory Protein of *Staphylococcus aureus* (CHIPS) was confirmed by isothermal titration calorimetry (ITC) and ¹⁵N-heteronuclear single quantum coherence spectroscopy (¹⁵N-HSQC) spectroscopy. However, the pY-containing peptides were much less active in the calcium mobilization assay as compared to the sulfated ones, potentially due to mobility differences caused by doubly charged cations and the presence of fast-acting phosphatases. Owing to its aromatic module and negative charge, sulfotyrosylproteins have been implicated as a determinant factor of protein-protein interactions related to chemokine signaling for leukocyte moving and adhesion, blood coagulation, and pathogen evasion.²⁴

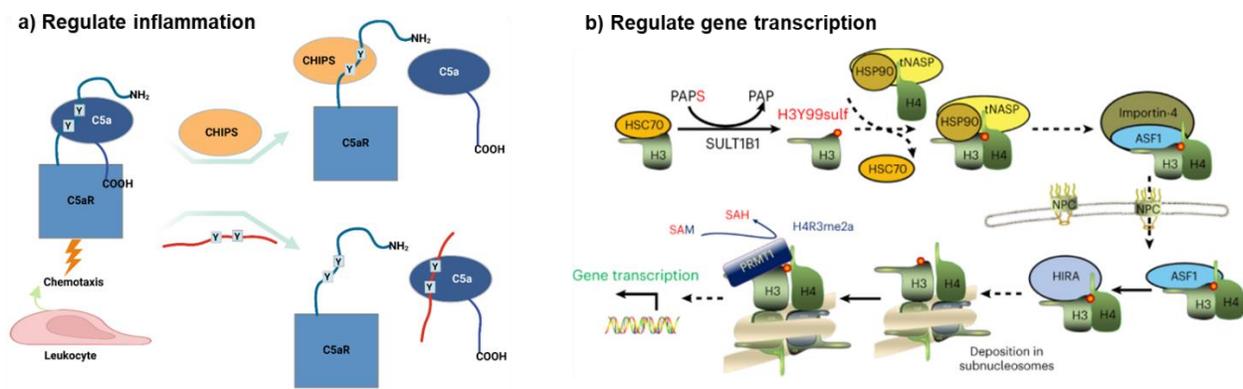


Figure 1-7. Function of sulfotyrosine-containing proteins (figure 1-7b was reproduced from reference 27 with permission, copyright @ Springer Nature America, Inc)

A series of sulfotyrosylproteins were demonstrated to play crucial roles in both intra- and extracellular biological process regulation. In order to protect the host from invading pathogens caused by infection or injury, an innate defense system is developed in human, in which complement component 5a (C5a) is a potent pro-inflammatory agent.²⁵ C5a binds to the *N*-terminus of C5aR (residues 1-35) inducing the chemotaxis of leukocytes for exotic ingredients clearance and highly excessive levels of C5a is whereas deleterious, causing numerous inflammatory diseases. Sulfation at

residue 11 and 14 of C5aR was verified to have a substantial influence on its binding affinity towards C5a. CHIPS, which has a nanomolar affinity to C5aR, and C5aR₇₋₂₈ analogs synthesis with different sulfation patterns, provide lead sequences for pharmaceutical antagonist innovation.²⁶

Although the majority of the sulfated molecules were reported to be exclusively secreted outside of the cell, onto the cell membrane, or into the cytoplasm, a very recent paper by Yugang Wang *et al.* described the sulfation of tyrosine 99 residue in nascent histone H3 by SULT1B1 as a regulator of gene transcription.²⁷ With the first solid proof of sulfotyrosine (sY) existing in nuclear proteins and extension of the chromatin regulator, drawing increasing attention to the undervalued sulfation pathway from both academia and the pharmaceutical industry.

1.2 Sulfation Detection Method

Regarding a particular PTM, its corresponding detection techniques are essential for pinpointing the occurring sites and profiling substrate scopes, as well as for understanding the biological functions and providing starting points for drug or antagonist design. Nevertheless, sulfation detection method are still in its infancy in comparison with a handful of PTMs, which are extensively investigated like phosphorylation of serine/threonine/tyrosine and lysine modification by acetylation/methylation.²⁸ Currently, only a limited number of tools are generally available for the study of post-translational sulfation on glycan and tyrosine residues, including mass spectrometry (MS), antibody-based techniques, radioactive isotope labeling, and fluorescence sensing.²⁹

1.2.1 MS

Nowadays, the advancing of cutting-edge mass spectrometry techniques featuring extraordinary accuracy, sensitivity, and automatic analysis software have enabled sequence dissection and sulfated site determination.³⁰ A notable work by Alina D. Zamfir *et al.* in 2022 demonstrated a remarkable progress in chondroitin sulfate (CS) and dermatan sulfate (DS) structural analysis (Figure

1-7).³¹ Aiming to profile CS/DS hexasaccharide domains in human embryonic kidney HEK293 cells decorin (DCN), the target glycan chains were firstly released by β -elimination and then depolymerized into oligosaccharide utilizing chondroitin AC I lysate followed by rapid size-exclusion chromatography. The resulting hexamers were subjected to ion mobility separation mass spectrometry (IMS-MS) and tandem collision-induced dissociation MS/MS (CID-MS/MS). In the optimal MS conditions, a series of under- or oversulfated hexasaccharide species were observed, such as trisulfated-4,5-D-GlcAGalNAc[IdoAGalNAc]2 sequence with unusual sulfation sites, which indicated diagnostic fragment ions crucial for the characterization of misregulations. However, the possible existing regioisomers raised questions about the conclusive assignment of the sulfation sites. Therefore, chemical or chemoenzymatic synthesis of glycosaminoglycans (GAGs) is often employed to provide reliable standards for decisive result verification with reversed-phase high-performance liquid chromatography (RP-HPLC).³² Although the similar principle seems applicable for the sulfotyrosyl peptides and proteins identification, the labile desulfurization processes under ionization circumstances pose unique challenges with false negative readings.³³⁻³⁴ By operating matrix-assisted laser/desorption ionization time-of-flight (MALDI-TOF) MS examination in both negative and positive ion modes, Dick Heinegård group identified variable sulfotyrosyl sites in extracellular class II leucine-rich repeat protein/proteoglycans in 2004.³⁵ A clear trend was observed in this work that the peptides was normally showed up in non-sulfated form in positive mode, whereas sulfated and non-sulfated species both appeared in negative mode.

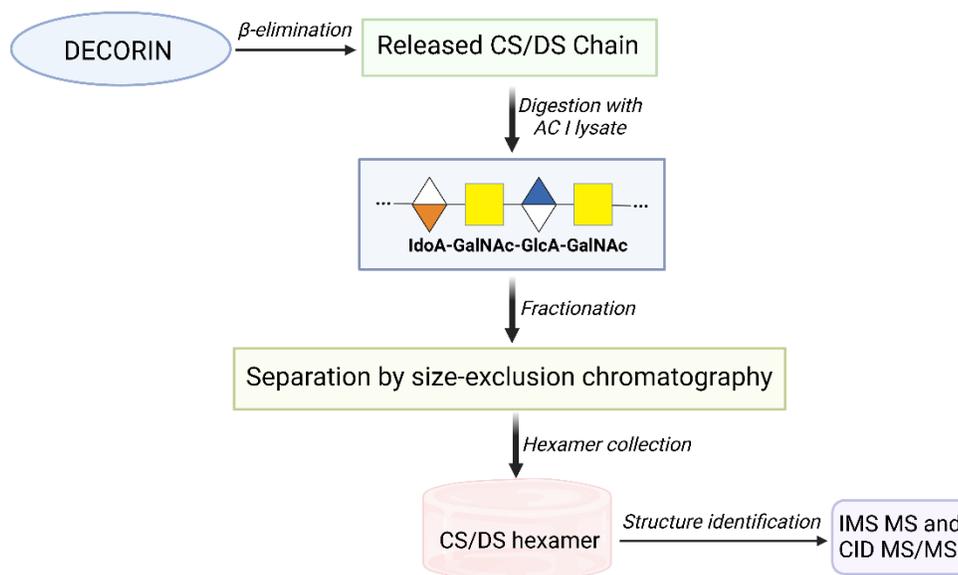


Figure 1-8. Mass spectrometry for sulfated glycan

Moreover, while no evidence has proved the natural eraser of tyrosyl sulfation modification *in vivo*, Kim *et al.* showcased the simultaneous determination of tyrosine phosphorylation and sulfation sites with PTMs enzymatic removal *in vitro* followed by tyrosine-specific bromination (Figure 1-9a).³⁶ By using phosphatase or sulfatase to eliminate the corresponding phosphate or sulfate group, the researchers incubated the pre-treated tyrosine-containing sequences with 32% hydrogen bromide (HBr), introducing the unique Br tag onto the released tyrosine residues. Subsequent MALDI-TOF analysis enabled unambiguously determining the original phosphate and sulfated positions. With the advent of new techniques, Ghosh *et al.* described the unexpected sulfation in chrombacin by analyzing low-level peptide samples direct from cell culture media via microfluidic separation and high-resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometer (FT-ICR MS) with complementary techniques.³⁷

While sulfation is predominantly found in glycan and tyrosine residues, Burlingame *et al.* documented the very first case of sulfoserine- and sulfothreonine-containing proteins from a wide

range of eukaryotes, encompassing freshwater snail *Lymnaea stagnalis*, malaria parasite *Plasmodium falciparum*, and human beings.³⁸ The separated protein from SDS-PAGE was digested with trypsin and analyzed by coupling RP-HPLC to electrospray ionization tandem mass spectrometry (ESI MS/MS). High-accuracy FT-ICR mass measurement and characteristic fragmentation patterns differentiated the serine/threonine sulfation from the phosphorylation (Figure 1-9b).

The continuous improvement of novel fragmentation techniques in MS/MS will undoubtedly enhance the accuracy, convenience, and in-depth understanding of the sulfation realm.

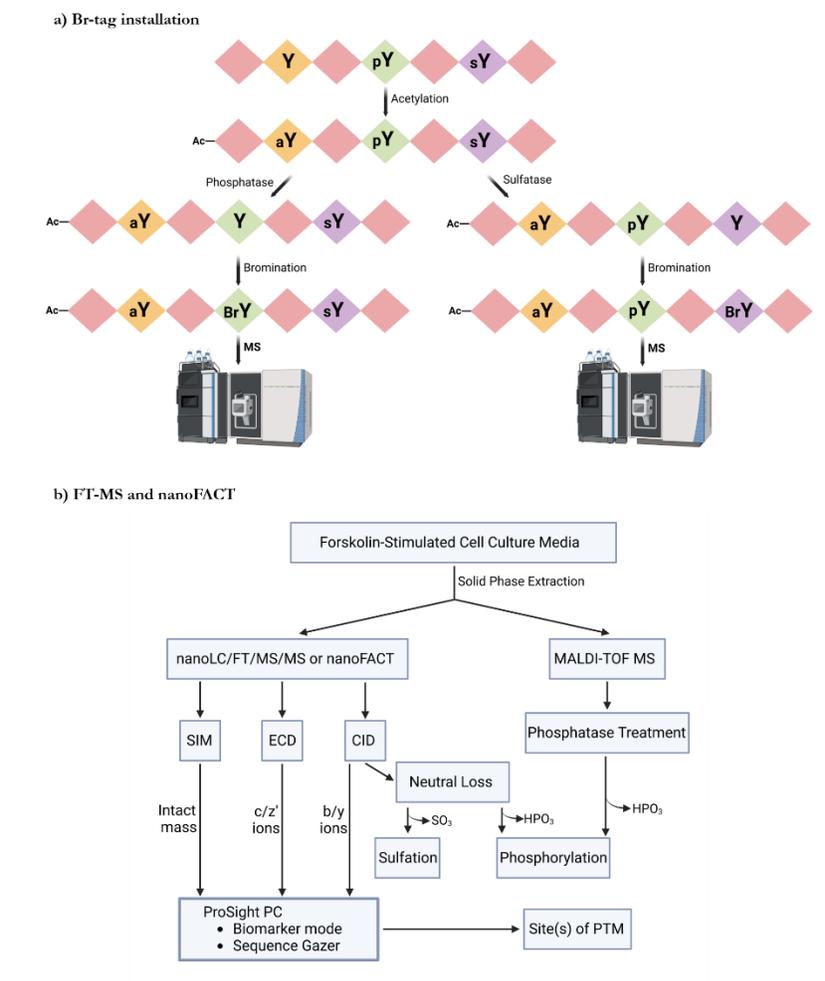


Figure 1-9. Sulfation detection with Br-tag installation or advanced MS

1.2.2 Anti-Sulfotyrosine Antibody

Besides MS proteomics, convenient antibody-based immune assays such as immunoprecipitation and Western blot are also widely utilized for sY-containing macromolecule detection. A significant milestone in this area was achieved by Schwartz *et al.*, who employed direct immunoprecipitation of cell lysate with various antibodies complexes like anti-human DR and –DS. With this method, they achieved the isolation and characterization of sulfated proteoglycan associated with human leukocyte antigen (HLA) class II antigen.³⁹ In order to characterize the isolated molecules, they subjected them to NaDodSO₄/polyacrylamide gel electrophoresis (PAGE) tests. The results indicated the obtained molecules sized between 40,000-70,000 and displayed polyanionic properties. Protease digestion experiments suggested that these molecules were resistant towards proteases but susceptible to chondroitinase ABC, confirming the isolated molecules belonged to the chondroitin sulfate class conjugated proteins.

Interestingly, among the diverse array of sulfated carbohydrates, the presence of a sole tyrosine sulfation form offers a unique opportunity for generating sY-specific antibodies. This specificity can be harnessed to develop targeted assays for the detection and study of sY-containing macromolecules. In 2006, Kevin L. Moore *et al.* reported a novel anti-sulfotyrosine antibody utilizing phage display technology.⁴⁰ Regardless of the surrounding sequence context, the produced monoclonal antibody, named PSG2, exhibited strong and specific binding to sY residues in both peptide and protein substrates. Importantly, PSG2 demonstrated no binding affinity towards pY residues in proteins or sulfated glycans. In their study, PSG2 was successfully employed to identify sY-containing proteins in crude mouse epididymis tissue samples through Western blot analysis, indicating the presence of some undersulfated sperm or epididymal proteins in TPST double knock-out male mice. Recently, Guo *et al.* systematically engineered a pY binding pocket of a ~100 amino acid composed Src Homology 2

(SH2) domain to serve as sY antibody mimic.⁴¹ By evolving the pY-targeting pocket with tailored selection schemes, the authors identified several SH2 mutants (e.g. SH2-1.8 and SH2-3.1) featuring high affinity and specificity to sY residues which were supported by molecular docking simulation. One of the major advantages of these SH2 mutants was their reduced size and cost compared to conventional antibodies, presenting a valuable addition to existing strategies. However, it is worth noting that the retention of these mutants towards pY residues limits their ability to selectively enrich or identify particular sulfoproteins in complex biological samples.

Till now, two commercial anti-sulfotyrosine antibodies are predominantly used in Western blot⁴²⁻⁵⁰ and imaging assays: sulfo-1C-A2 and 7C5^{40, 51-52}. However, there is a clear need for more novel anti-sulfotyrosine antibodies that offer higher specificity, easier accessibility, and increased robustness. The development of such antibodies will undoubtedly have significant benefits for advancing research in related fields.

1.2.3 Isotope Labeling

Metabolic labeling of sulfoproteome with ³⁵S-sulfate represents another valuable method to determine the exact sulfation sites and explore related enzyme activity in various physiological contexts, including cell culture, isolated tissue, and even whole animals. In this protocol, inorganic ³⁵S-sulfate is enzymatically converted to organic ³⁵S-PAPS sulfate donor by PAPSS upon cellular uptake. Subsequently, sulfotransferase transfer ³⁵S-sulfate onto specific residues. However, it's essential to note that the isotope labeling protocol employed in this approach cannot differentiate between sulfation occurring on tyrosine residues and carbohydrates. As a result, the reported cases have primarily focused on sulfation processes exclusive to tyrosine residues. Researchers achieved this specificity by using glycosylation inhibitors, glycosylation-deficient cell lines, site-directed mutagenesis, or glycosidase treatment.⁵³ Moreover, the concern that inorganic ³⁵S-sulfate might be reduced for the

synthesis of methionine or cysteine residues is unwarranted, since mammalian cells lack the ability for sulfate reduction. In 1986, Strauss *et al.* successfully identified the sulfation sites on the α -chain of the fourth component of human complement (C4) using an isotope labeling strategy. They performed *in vivo* labeling of cell cultures with inorganic ^{35}S -sulfate, and the targeted C4 was then directly isolated from the complex through immunoprecipitation. The presence of tyrosine sulfation positions was confirmed through a combination of trypsin digestion, sequential Edman degradation, and chymotrypsin cleavage, providing cross-certification of the results. Further analysis, involving a comparison of the ^3H -tyrosine and [$^{35}\text{S}+^3\text{H}$]-tyrosine sulfate ratio along with Pronase digestion, indicated that the biologically active form of C4 contained an average of 2-3 sulfations. Furthermore, the exact positions of sulfation were verified by overlaying the autoradiogram, adding to the robustness of the findings.⁵⁴

Despite the success of the radioactive tagging method, it is important to acknowledge some drawbacks. The approach involves complex operational procedures, significant costs, and potential radiological hazards, which may limit its widespread application in certain settings.

1.2.4 Fluorescent assay

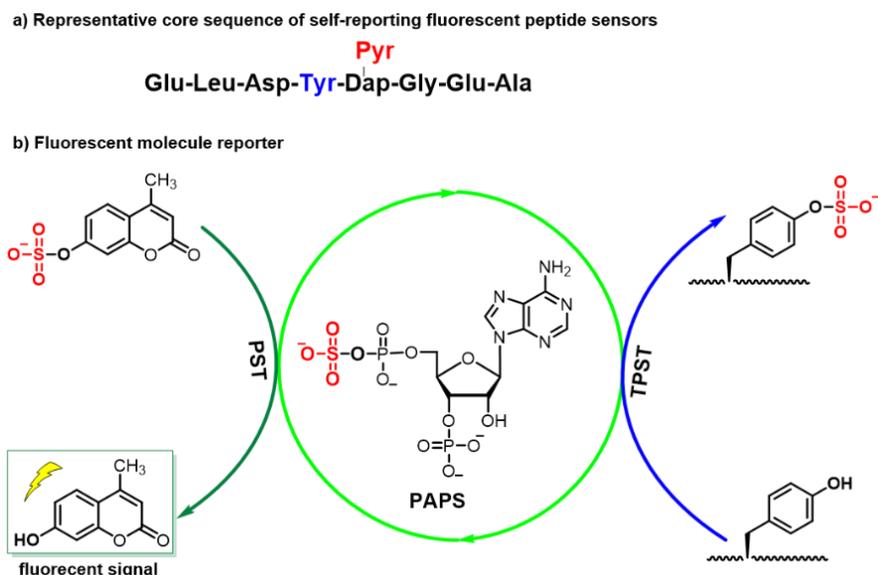


Figure 1-10. Fluorescent assay probing sulfation pathway

To shed light on the kinetics of sulfation occurring in cellular systems, real-time fluorescence monitoring systems were established by hijacking the native sulfation pathways. In 2014, Robert J. Geraghty *et al.* developed a continuous direct assay to investigate the two isoforms of human tyrosylprotein sulfotransferase (hTPST1 and hTPST2) with self-reporting fluorescent peptides (Figure 1-10a).⁵⁵ A core peptide sequence, surrounded by amino acids with acidic side chains, was modified with L-2,3-diaminopropionic acid (L-DAP) conjugated pyrene (Pyr) moiety at various positions. The fluorescence of the pyrene moiety was quenched by the aromatic side chain of tyrosine, depending on their proximity, and the quenching effect was attenuated by tyrosine sulfation. Consequently, an increase in fluorescence emission was captured by microplate reader SpectraMax M5e (Molecular Devices), when tyrosylprotein sulfotransferase (TPST)-mediated tyrosine sulfation occurs at optimal pH = 6. The recombinant hTPSTs were confirmed to be active even without the divalent ions, and the inhibition assay of the enzyme activity was also applicable in this design. Building on the fundamental discovery that TPST only employs PAPS as the sulfate donor, Yang *et al.* proposed a

strategy to determine the activity of target TPST by controlling the PAPS regeneration process (Figure 1-10b).⁵⁶ The experiment was simplified into an *in vitro* setting, where the production of PAPS only originated from 4-methylumbelliferyl sulfate via the aid of phenol sulfotransferase (PST). Subsequently, the activity of a recombinant N utilization substance protein A fused with *Drosophila melanogaster* TPST was immediately detected by the fluorescence of 4-methylumbeliferone (MU). Using this protocol, significant kinetic variation was observed between a small peptide (MW 1541) of P-selectin glycoprotein ligand-1 (PSGL-1) and its large glutathione S-transferase fusion protein (MW 27833).

These fluorescent peptide sensors or reporters provide a convenient tool to interrogate the kinetics of sulfation pathway involved enzymes and show great potential in exploring the functions of related enzyme and protein substrates in cellular contexts. Moreover, they can be applied in high-throughput screening of small molecular inhibitors and modulators in disease models.

In contrast to other heavily investigated PTMs across the proteome, certain well-developed detection methods, such as chemical labeling strategies that take advantage of the intrinsic reactivity of the target function group for covalent bond formation, have NOT been transplanted to the sulfation research. On the other hand, the advent of bioinformatics databases like PredSulSite⁵⁷⁻⁵⁸ and Sulfinator⁵⁹⁻⁶⁰ provide valuable tools for preliminary screening of the potential sulfation sites and binding sites or pockets. The ever-growing techniques and research approaches immensely promote the delineation of the overreaching sites and substrate repertoire of naturally occurring sulfation.

1.3 Access to Sulfated Biomolecules

Since the first discovery of heparin in 1916 from dog liver by Jay Mclean,⁶¹ tremendous progress has been made in accessing the desired sulfoglycan, sulfopeptides, and sulfoproteins. Those advances have subsequently facilitated evaluations of corresponding structure-activity relationships,

enriching our understanding of the structure features that are crucial in molecular interactions and biological functions. Over the century, two major ways are devised to obtain these molecules: (1) enrichment methods for sulfated material from livestock with countless diversities but inherently heterogeneous; (2) chemical, chemoenzymatic, and enzymatic synthesis with well-defined structures and sulfation patterns. These two complementary strategies have been instrumental in driving research in the field of sulfation, providing valuable tools to explore the diverse roles and functions of sulfoglycans, sulfopeptides, and sulfoproteins in various biological processes.

1.3.1 Enrichment from Biomaterials

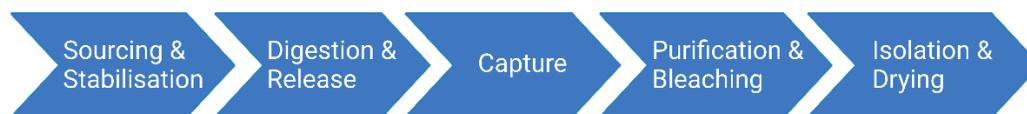


Figure 1-11. Schematic representation of an industrial heparin purification process

To date, manufacturing heparin from animal offal is still the sole source of commercial heparin in industry since 1922, albeit with the fact that the detailed heparin purification protocol in the industry is closely guarded and not readily available to the public (Figure 1-11).⁶² However, exciting progress has been made in microscale sulfated glycopeptide enrichment. Leveraging the high negative charge density of sulfoglycan, Imami *et al.* achieved the separation of sulfated glycopeptide from digested thyroid stimulating hormone (TSH) by capillary electrophoresis (CE) and MS signal enhancement, using basic ion-pair interaction with a tripeptide(Lys-Lys-Lys) (Figure 1-12a).⁶³

In virtue of controlling the net charge of sulfoglycopeptides, Kameyama *et al.* demonstrated the enrichment of target bovine luteinizing hormone (bLH) using sulfate emerging (SE) and strong anion exchange chromatography (SAX) from a complex mixture of human serum spiked with 0.1% bLH (Figure 1-12b).⁶⁴ Starting from limited proteolysis to get rid of basic residues like arginine and

lysine, the residual peptide side chain (acidic) was subsequently treated with acetohydrazide modification. Following the SAX procedure, the characterization of challenging sulfated glycopeptide with low abundance and ionization efficiency by MS was accomplished.

Considering that most reported methods for sulfated glycan enrichment are often limited to small scale and compromised specificity, the vast diversity of obtained molecules from various species under normal or pathological conditions provides invaluable information. These enriched sulfated glycopeptides hold significant potential for advancing our understanding of complex biological systems and contributing to the exploration of various physiological and disease-related processes.

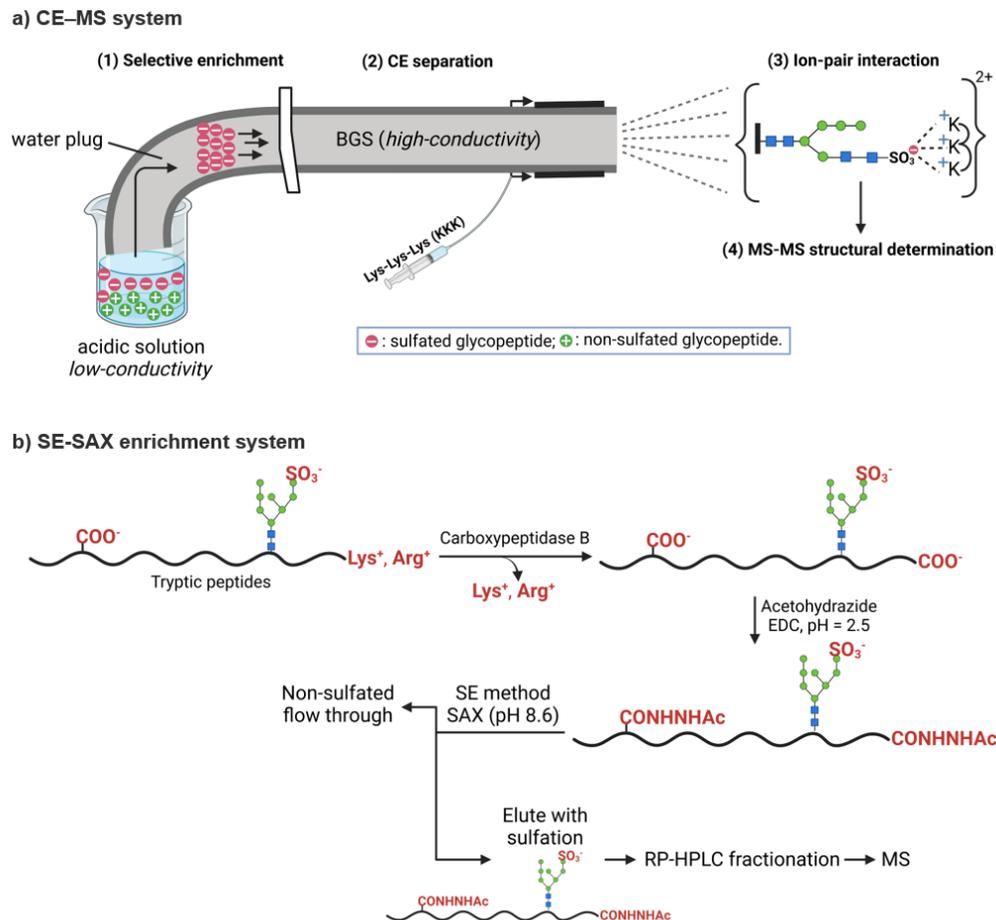


Figure 1-12. Enrichment for analysis systems

With respect to sulfotyrosylprotein enrichment,⁶⁵ the immunoprecipitation method with specific antibody still faces limitations due to the scarcity of available anti-sulfotyrosine antibodies, and its application in native samples remains unproven.⁴⁰ On the other hand, Batista *et al.* successfully enriched sulfopeptides from complex skin secretions of *Pachymedusa dancicolor* frog by immobilized metal ion affinity chromatography (IMAC-Ga). Coupling with high-resolution liquid chromatography/electrospray ionization mass spectrometry (LC/ESI-MS) in an Orbitrap XL and MALDI-TOF in an ABI 4800 instrument, they established a high-throughput sulfopeptide analysis system.⁶⁶ In a separate study, Brodbelt *et al.* employed weak anion exchange (WAX) for the enrichment of sulfotyrosine-containing peptide (Figure 1-13).⁶⁷ By attenuating peptide basicity via carbamylation of all primary amines, improved differentiation of the sulfated and non-sulfated peptides was achieved. Upon elution, electrospray ionization in negative mode, and ultraviolet photodissociation (UVPD) for sequencing, a new tyrosine sulfation site at 1513 was identified on bovine coagulation factor V. These results laid the foundation for sulfotyrosulpeptide enrichment and is promising to be scaled up for complex biological sample analysis.

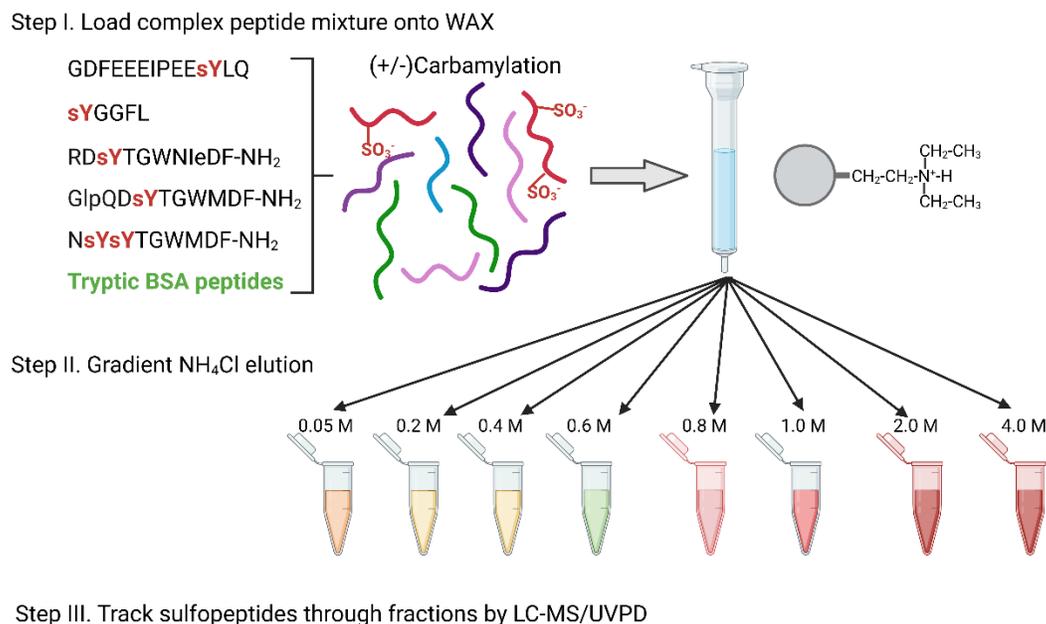


Figure 1-13. WAX system for sulfotyrosylpeptide enrichment

1.3.2 Synthesis

Chemical synthesis

Naturally polysulfated biomolecules are typically synthesized in the Golgi apparatus without a strict template, resulting in a wide diversity of structures and sulfation patterns. In order to comprehensively understand the related biological events, a systematic evaluation of the structure-activity relationship is of key significance, which requires the subjective molecules with well-defined skeletons and sulfated sites. To this end, chemical synthesis is currently the leading and most reliable approach to forging these compounds.

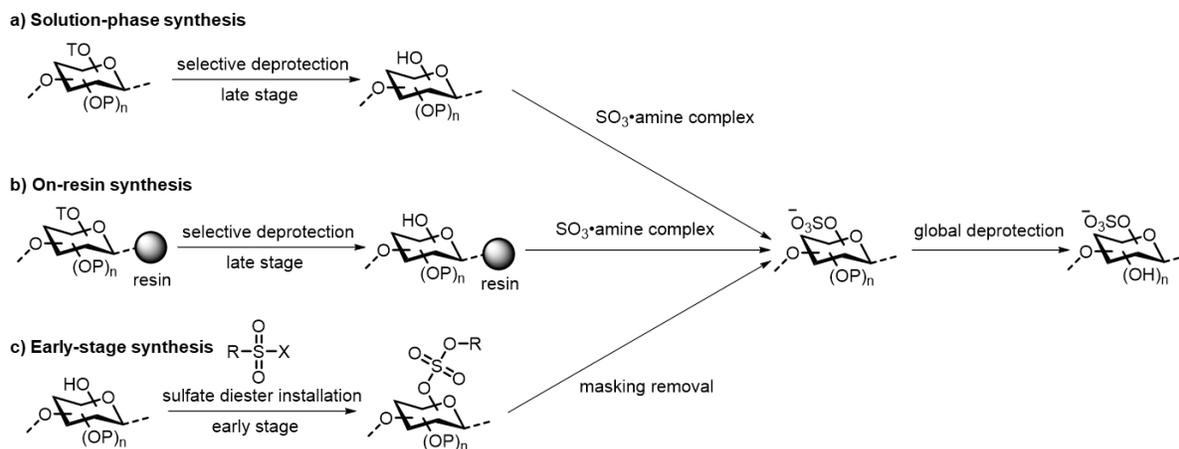


Figure 1-14. Chemical synthesis of sulfated glycan

Since the discovery of sulfated glycan, three major synthetic approaches have been developed so far: solution-phase synthesis, on-resin synthesis, and early-stage sulfation strategy (Figure 1-14).⁶⁸⁻⁶⁹ Among them, the solution-phase strategy is the most sophisticated one, which is synergistically evolved with the advances in glycosylation methodology and site-specific sulfation introduction reagents (Figure 1-14a). In 2019, Wei Zhao and his colleagues reported a one-pot glycosylation method for quick access to fondaparinux (Arixtra) in a preactivated and highly stereoselective manner (Figure 1-15a).⁷⁰ They presented two convergent routes for constructing three glucosamine building blocks using classic thiol donors, significantly simplifying the preparation procedures. Sequential assembling of the glucosamine moieties to oligosaccharides, followed by successive *O*- and *N*-sulfation transformations with $\text{SO}_3 \cdot \text{Et}_3\text{N}$ and $\text{SO}_3 \cdot \text{Pyridine}$, eventually yielded the desired fondaparinux sodium in 27.3% yield at a 38.5 mg scale.

Very recently, Ye group reported a remarkable 1080-mer polyarabinosides using an automated solution-phase glycan synthesizer, which exceedingly surpassed the artificial synthesis of nucleic acids (up to 200-mer) and proteins (up to 472-mer).⁷¹⁻⁷³ Based on their preactivation one-pot multicomponent and continuous multiplicative synthesis strategy, gram-scale (1.06 g) synthesis of a

protected fondaparinux pentasaccharide in 62% isolated yield was achieved, although the final sulfation step was not included (Figure 1-15b).

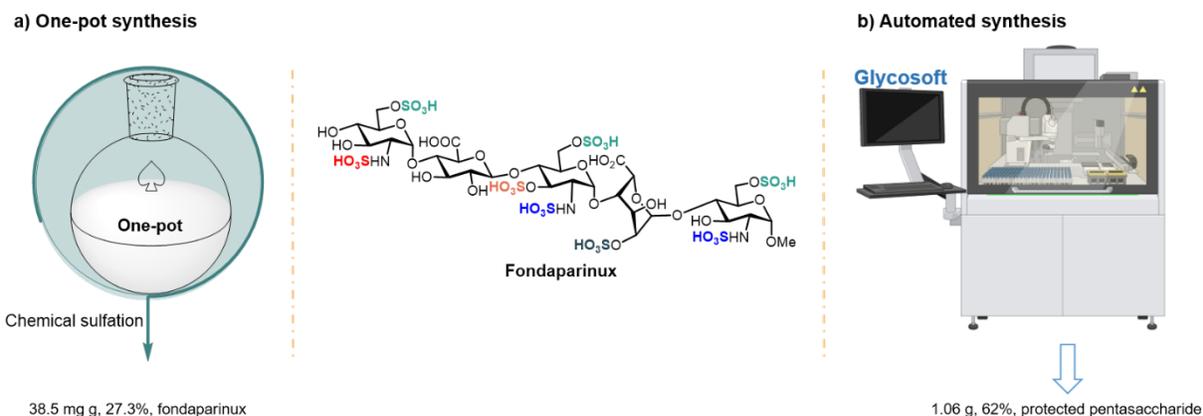
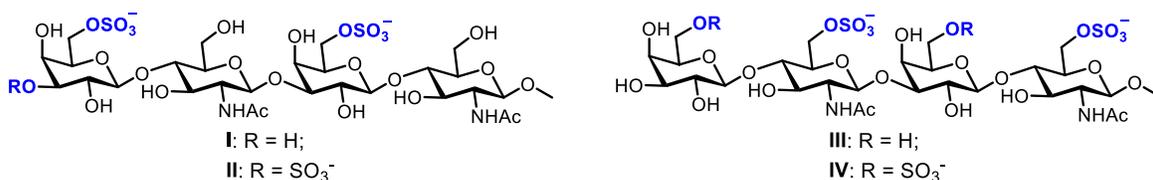


Figure 1-15. Solution-phase synthesis of sulfated glycan

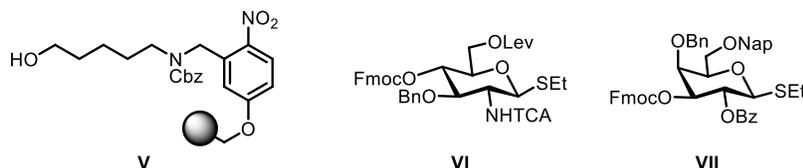
The ease of purification and its readily programmable characteristics render on-resin synthetic approaches a viable alternative to traditional solution-phase one (Figure 1-14b). With a photolabile 2-nitrobenzyl type linker and judiciously chosen protecting groups on the modules, linear glycan equipped with various sulfation patterns was quickly made by automated glycan assembly (AGA) with differentiated procedure applied.⁷⁴ Additionally, the terminal amine tail on the reducing end enables the click reaction conjugation, which is particularly suitable for biological interaction studies like microarrays (Figure 1-16). One noteworthy example among these synthesized glycans is the keratin sulfate (KS) tetrasaccharide III, which exhibited specific binding with adeno-associated virus AAVrh10, representing a potential glycan acceptor for this pathogen.

To tackle the challenges posed by high temperature, excessive amount of sulfation reagent (5 equivalents per sulfonation site), and repetitive SO_3 treatment in multiple sulfation cases, a novel strategy of masking sulfate at the early stage of synthesis has emerged (Figure 1-14c). Huang and his co-workers undertook such an endeavor to human syndecan-4 glycopeptides synthesis with robust

1) KS oligosaccharide target



2) Resin linker and building blocks



3) AGA of sulfated KS oligosaccharides

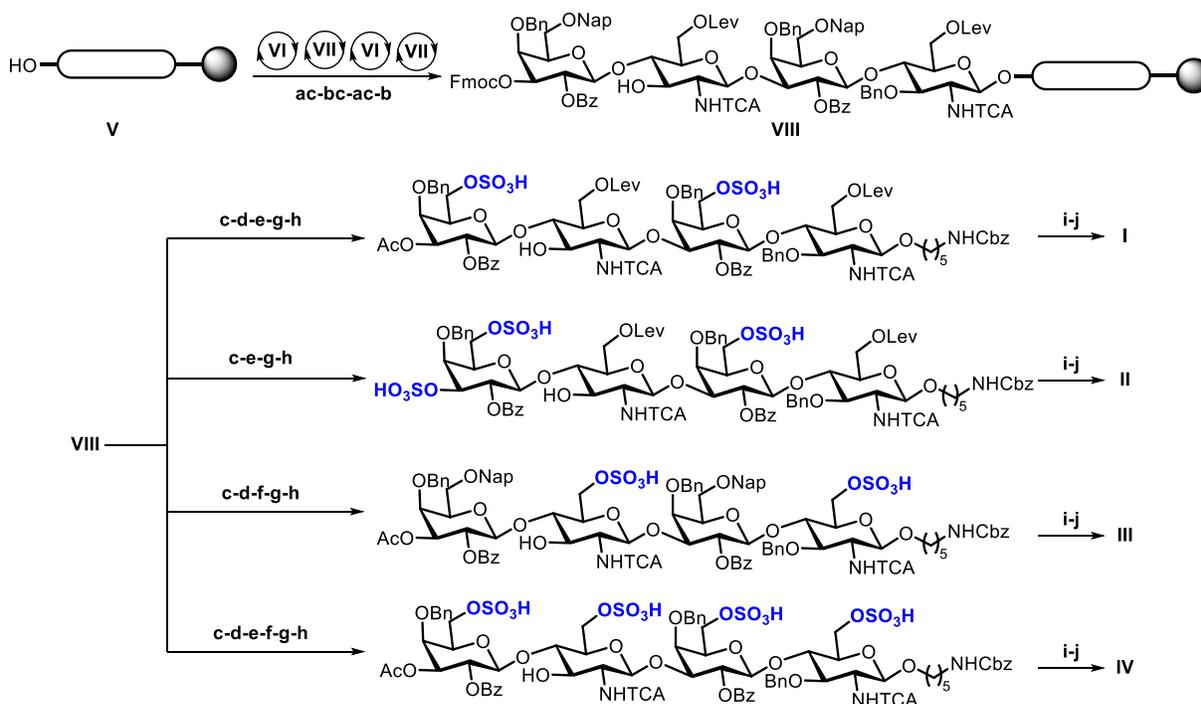


Figure 1-16. On-resin synthesis of KS oligosaccharides. A) Bioactive KS oligosaccharide as target; B) Photolabile linker on beads and modular building block for AGA; C) AGA synthesis. Reaction conditions: (a) Acidic wash and glycosylation 1; (b) acidic wash and glycosylation 2; (c) Fmoc deprotection; (d) capping; (e) Nap deprotection; (f) Lev deprotection; (g) sulfation; (h) UV cleavage (305 nm); (i) NaOMe, MeOH, 40°C; and (j) Pd/C, H₂, MeOH/triple-distilled water/AcOH (15:15:1).

dichlorovinyl (DCV) sulfate ester on both *N*- and *O*-sulfation positions.⁷⁵⁻⁷⁶ The masked sulfate esters were left untouched during peptide backbone assembly, and the reductive sulfate revealing condition was proved to be compatible with multiple aspartic residues. The rapid access to sulfated glycopeptide via this approach serves as a successful example to modulate the functions of heparan sulfate (HS) by conjugation with core peptides or proteins.

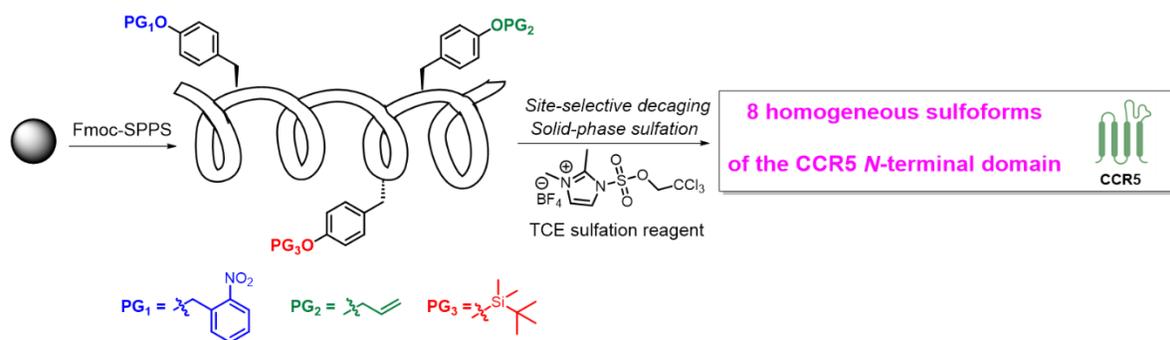


Figure 1-17. Site-selective SPPS of sulfated CCR5 library

Compared to complex carbohydrate structures, the easily-applied solid-phase peptide synthesis (SPPS) protocol dramatically reduced the working load for the synthesis of sY-containing peptides. In the traditional approach, Fmoc-Tyr(SO₃⁻)-OH is coupled to the growing peptide chain. However, the presence of the negative charge often impedes further amino acid coupling and comprises resin swelling. Moreover, the lability of sulfate to acidic treatment renders the final global protecting group removal and resin cleavage not applicable, leading to desulfurized products.⁷⁷

To overcome these challenges, cassette-based suitably masked sulfotyrosine has grown to be the prevailing method for homogeneous peptide synthesis.⁷⁸ Payne's group reported a divergent synthesis of eight distinct forms of *N*-terminal C-C chemokine receptor type 5 (CCR5, 2-22), each bearing discrete sulfation at Tyr10, Tyr14, and Tyr15 with three orthogonally protected Tyr building blocks on a single resin-bound intermediate (Figure 1-17).⁷⁹ With three preinstalled *o*-nitrobenzyl (*o*-NB), Allyl (All), and *tert*-butyl-dimethylsilyl (TBS) which could be differentially removed, sulfation at

the exposed sites with Tylor's 2,2,2-trichloroethyl (TCE) reagent was quickly accomplished.⁸⁰ The importance of the sulfation in the obtained sulfopeptides was demonstrated in an enzyme-linked immunosorbent assay (ELISA)-based competitive binding assay between sulfated CCR5₂₋₂₂ and HIV-1 Env glycoprotein gp120.⁸¹ Among all the sulfation sites, sY at position 14 were found to be critical to complement viral invasion into cells.

Chemoenzymatic synthesis

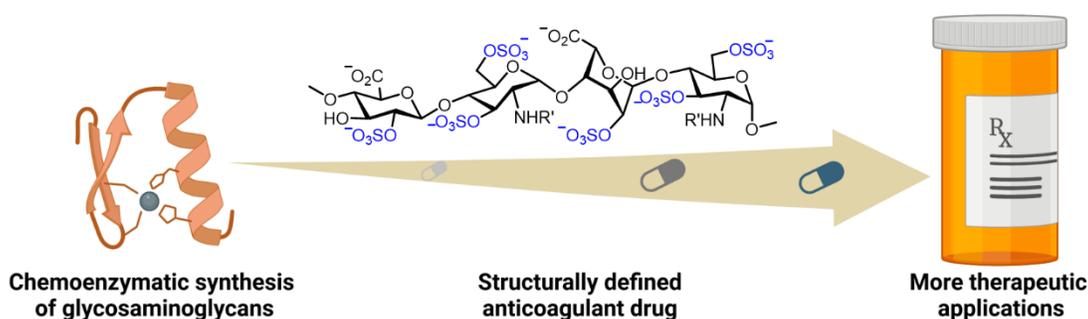


Figure 1-18. Chemoenzymatic synthesis of glycosaminoglycans

While chemical synthesis is widely applied in sulfated glycan synthesis, the lengthy synthetic routes always take a substantial amount of time and are limited in the library and molecular size. In this regard, the chemoenzymatic approach is an ideal complementary alternative with degraded natural polymers, *in vitro* assembled oligosaccharides, and chemically synthesized modules as substrates.⁸²⁻⁸³ Contrary to our intuition, small sulfate molecule targets can be more challenging to prepare because heparosan synthase-2 (pmHS2) has poor performance in elongating glucosamine monosaccharide.

Taking on this challenge, Liu and his collaborators paved the way to efficient construction of heparin pentasaccharide with ultralow molecular weight (ULMW) with minimum artificial protecting group manipulation (Figure 1-19).⁸⁴ Two heparin ULMW heparin constructs **1** (MW = 1778.5) and **2** (MW = 1816.5) with discrete sulfation patterns were forged in 10- and 12-steps with 45% and 37%

overall yields, respectively. On the other hand, chemical synthesis of similar fondaparinux entails about 50 steps with an overall yield of $\sim 0.1\%$, which highlights the remarkable efficiency of the chemoenzymatic approach.⁸⁵ The biosynthesized molecules demonstrated excellent anticoagulant activity in the *in vitro* experiment as its parent structure porcine and bovine heparin.⁸⁶ Moreover, in a rabbit model, these molecules exhibited pharmacokinetic properties similar to fondaparinux. Though only tens of milligrams of the target compounds were made in the reported laboratory setup, the chemoenzymatic approach is a rapidly evolving direction of great potential for scalable and broad substrate production.

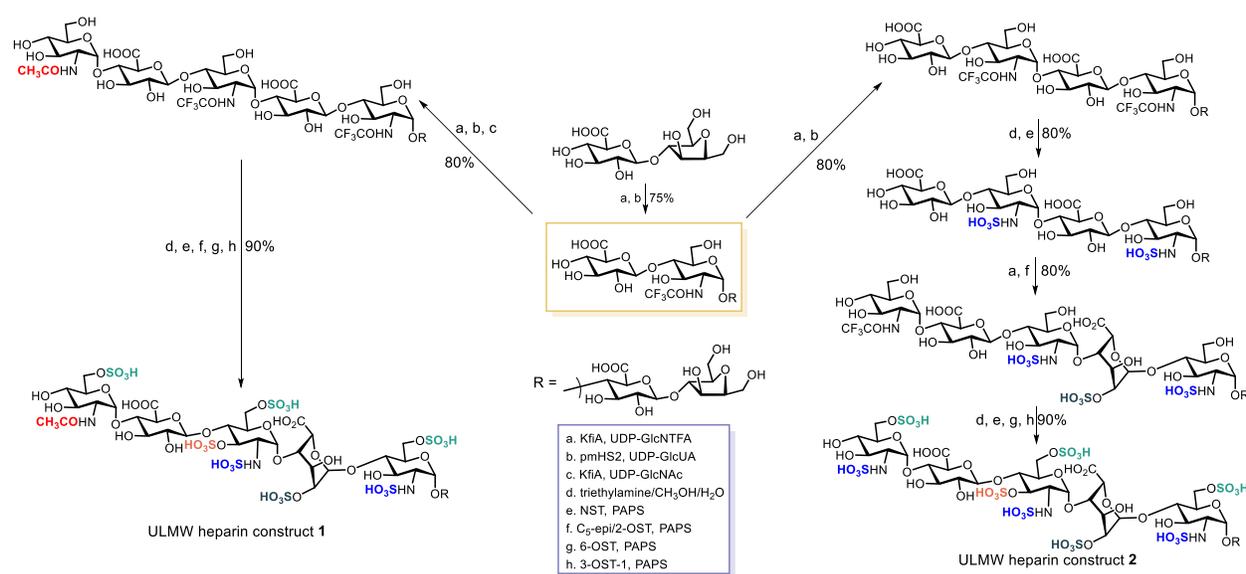


Figure 1-19. Chemoenzymatic synthesis of ULMW heparin

Enzymatic synthesis

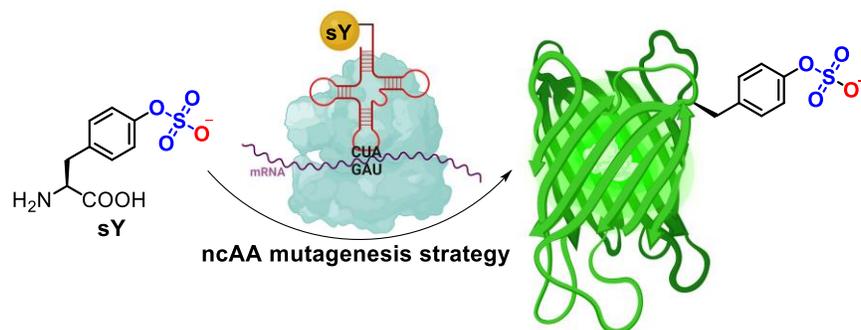


Figure 1-20. Direct sulfoprotein expression via ncAA mutagenesis

Although the active *N*-terminal truncated sequences serve as promising mimics of the parent macromolecules, synthesizing the whole protein target with well-defined sulfation patterns could significantly enhance our understanding of their acting mechanism with their native form in multicellular living systems. In pursuit of this goal, the non-conical amino acid (ncAA) mutagenesis strategy has become a powerful way to access recombinant whole sulfoproteins (Figure 1-20).⁸⁷

Peter G. Schultz and Chang C. Liu were the pioneers in reporting the selective incorporation of sY onto hirudin in *Escherichia coli* (*E. coli*) by genetically encoding the modified amino acid in response to the amber nonsense codon TAG.⁸⁸⁻⁸⁹ Both sulfo-hirudin and desulfo-hirudin were obtained and compared in a thrombin inhibition assay, where the sulfated form displayed over ten-fold higher affinity than the nonsulfated one. Chatterjee and his coworkers further adapted the homogeneous sulfoprotein expression in mammalian cells, allowing for the evaluation of sulfation influence in eukaryotic native environment.^{47, 90} It is noteworthy to mention that *N*-glycosidase (PNGase F) treatment evidently reduced the detected molecular weight of the target proteins indicated other PTM type like *N*-linked glycosylation also occurred in the endogenous environment. This observation

underlined the importance of eukaryotic protein expression in its original circumstances to account for any naturally-installed modifications.

In the same year 2020, Niu *et al.* evolved this chemical biology tool from yeast to mammalian cells and performed an *in vivo* cell-based studies on the role of sulfation sites in chemokine receptor CXCR4.⁹¹ They revealed that sulfation at Tyr21 position is critical for the activation of this receptor by its binding ligand.⁴⁶ Furthermore, the analysis of the crystal structure of the evolved sY-specific tyrosyl-tRNA synthetase (sTyrRS) complexed with sY brought valuable insight into the substrate specificity mechanism and future enzyme engineering. Different from exogenous feeding of the unnatural amino acids in previous genetic expansion methods, Xiao and his colleagues created autonomous prokaryotic and eukaryotic cells for biosynthesizing sY-containing proteins.⁴⁴ This system circumvented the poor membrane permeability of sY and produced the highest sulfoprotein expression yields ever reported.

In summary, the development of a particular area is considerably limited by the accessible amount of target materials and the diversity of obtained libraries using current chemical and biological technologies.⁹² With the advancement of isolation methods and synthetic approaches to the physiologically significant and therapeutically valuable agents, all these works collectively facilitate substantial strides in cellular research of post-translational sulfation, laying the foundation for cost-effective production of sulfated biomolecules.

1.4 Our work

Centering on the sulfation pathway and sulfur (VI) fluoride exchange (SuFEx) reaction, we developed a series of general approaches to caged O-sulfated biomolecules and their selective deprotection. First, an O-sulfation strategy is developed by coupling aromatic fluorosulfate with

silylated target molecules. Scalable synthesis was demonstrated on monosaccharides, disaccharides, amino acids, and steroids. Selective hydrolytic and hydrogenolytic removal of the aryl masking groups yielded the corresponding *O*-sulfated products in excellent yields. Furthermore, a complete knowledge gap was noted in the biocompatible caging of sulfate. With the rational design and systematic optimizations, we discovered that fluorosulfotyrosine in peptides and proteins was an ideal precursor for sulfotyrosine (sY), which can be efficiently converted into the anionic active form by hydroxamic acid activators under physiologically relevant conditions. Photocaging the hydroxamic acid activators further allowed for the light-controlled activation of functional sulfopeptides. This system featuring fast kinetics, high selectivity, excellent robustness, and on-demand release provides a valuable tool for probing the functional roles of sulfation in the peptides and proteins.

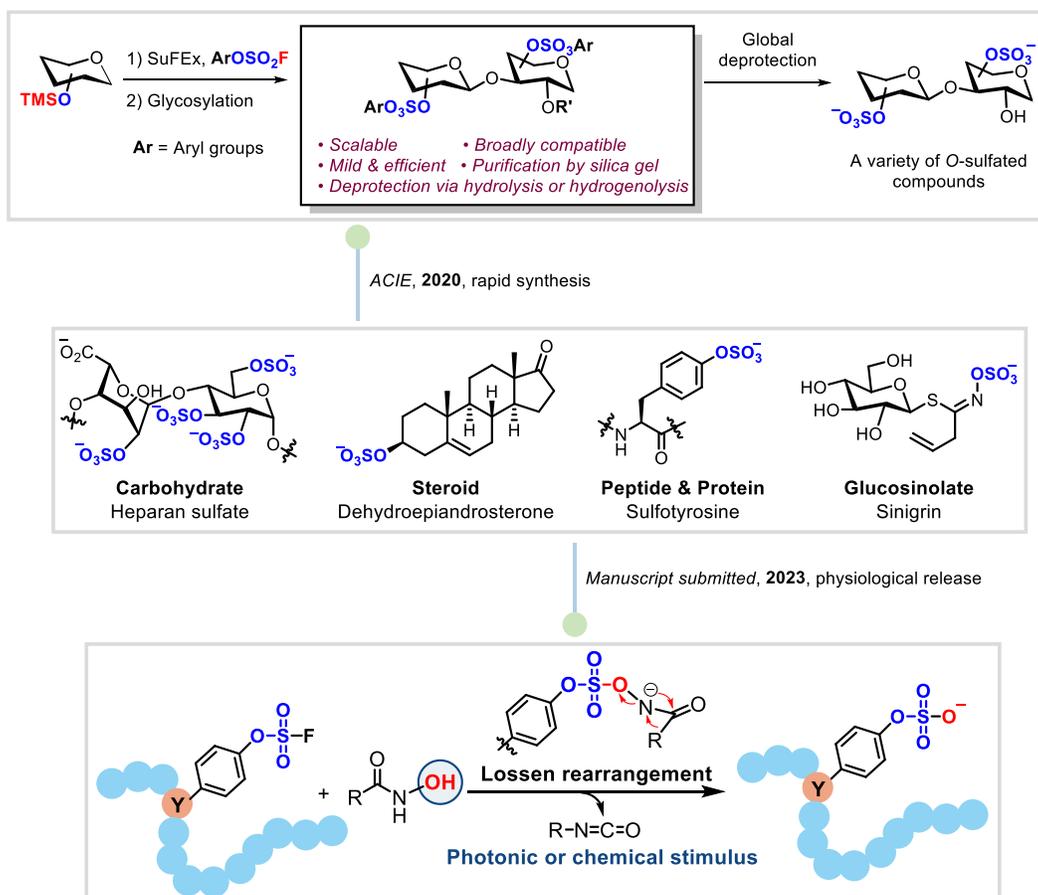


Figure 1-21. Our work

1.5 References

1. International Human Genome Sequencing Consortium, Finishing the euchromatic sequence of the human genome. *Nature* **2004**, *431* (7011), 931-945.
2. International Human Genome Sequencing Consortium, Initial sequencing and analysis of the human genome. *Nature* **2001**, *409* (6822), 860-921.
3. Khoury, G. A.; Baliban, R. C.; Floudas, C. A. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Scientific reports* **2011**, *1*, 1-5.
4. Van der Laan, S.; Maiorano, D., Post-translational modifications in embryonic cell cycle. *Cell Cycle* **2014**, *13*, 1364-1365.
5. Chatterjee, B.; Thakur, S. S., Investigation of post-translational modifications in type 2 diabetes. *Clinical Proteomics* **2018**, *15*, 32.
6. Hirschmann, F.; Krause, F.; Papenbrock, J., The multi-protein family of sulfotransferases in plants: composition, occurrence, substrate specificity, and functions. *Frontiers in plant science* **2014**, *5*, 556.
7. Xu, Z.-H.; Otterness, D. M.; Freimuth, R. R.; Carlini, E. J.; Wood, T. C.; Mitchell, S.; Moon, E.; Kim, U.-J.; Xu, J.-P.; Siciliano, M. J. Human 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPSS1) and PAPSS2: gene cloning, characterization and chromosomal localization. *Biochem. Biophys. Res. Commun.* **2000**, *268*, 437-444.
8. Leung, A. W.; Backstrom, I.; Bally, M. B. Sulfonation, an underexploited area: from skeletal development to infectious diseases and cancer. *Oncotarget* **2016**, *7*, 55811.
9. Leung, A. W.; Veinotte, C. J.; Melong, N.; Oh, M. H.; Chen, K.; Enfield, K. S.; Backstrom, I.; Warburton, C.; Yapp, D.; Berman, J. N. *In vivo* Validation of PAPSS1 (3'-phosphoadenosine 5'-phosphosulfate synthase 1) as a Cisplatin-sensitizing Therapeutic Target. *Clinical Cancer Research* **2017**, *23*, 6555-6566.
10. Oostdijk, W.; Idkowiak, J.; Mueller, J. W.; House, P. J.; Taylor, A. E.; O'reilly, M. W.; Hughes, B. A.; De Vries, M. C.; Kant, S. G.; Santen, G. W. PAPSS2 deficiency causes androgen excess via impaired DHEA sulfation—in *vitro* and *in vivo* studies in a family harboring two novel PAPSS2 mutations. *The Journal of Clinical Endocrinology & Metabolism* **2015**, *100*, E672-E680.
11. Coughtrie, M. W. Function and organization of the human cytosolic sulfotransferase (SULT) family. *Chem. Biol. Interact.* **2016**, *259*, 2-7.
12. Chuzel, L.; Fossa, S. L.; Boisvert, M. L.; Cajic, S.; Hennig, R.; Ganatra, M. B.; Reichl, U.; Rapp, E.; Taron, C. H. Combining functional metagenomics and glycoanalytics to identify enzymes that facilitate structural characterization of sulfated N-glycans. *Microb. Cell Factories* **2021**, *20*, 162.
13. WEISSHAAR, G.; HIYAMA, J.; RENWICK, A. G. Site-specific N-glycosylation of ovine lutropin: Structural analysis by one-and two-dimensional ¹H-NMR spectroscopy. *Eur. J. Biochem.* **1990**, *192*, 741-751.
14. Poulain, F. E.; Yost, H. J. Heparan sulfate proteoglycans: a sugar code for vertebrate development? *Development* **2015**, *142*, 3456-3467.
15. Toida, T.; Chaidedgumjorn, A.; Linhardt, R. J. Structure and bioactivity of sulfated polysaccharides. *Trends in glycoscience and glycotechnology* **2003**, *15*, 29-46.
16. Yayon, A.; Klagsbrun, M.; Esko, J. D.; Leder, P.; Ornitz, D. M. Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. *Cell* **1991**, *64*, 841-848.
17. Folkman, J.; Klagsbrun, M. Angiogenic factors. *Science* **1987**, *235*, 442-447.

18. Li, J.; Sparkenbaugh, E. M.; Su, G.; Zhang, F.; Xu, Y.; Xia, K.; He, P.; Baytas, S.; Pechauer, S.; Padmanabhan, A. Enzymatic synthesis of chondroitin sulfate E to attenuate bacteria lipopolysaccharide-induced organ damage. *ACS Central Science* **2020**, *6*, 1199-1207.
19. Braz-de-Melo, H. A.; Faria, S. S.; Pasquarelli-do-Nascimento, G.; Santos, I. O.; Kobinger, G. P.; Magalhaes, K. G., The Use of the Anticoagulant Heparin and Corticosteroid Dexamethasone as Prominent Treatments for COVID-19. *Front. Med.* **2021**, *8*, 615333.
20. Liu, J.; Thorp, S. C. Cell surface heparan sulfate and its roles in assisting viral infections. *Medicinal research reviews* **2002**, *22*, 1-25.
21. Jevons, F. Tyrosine O-sulphate in fibrinogen and fibrin. *Biochem. J* **1963**, *89*, 621.
22. Moore, K. L. The biology and enzymology of protein tyrosine O-sulfation. *J. Biol. Chem.* **2003**, *278*, 24243-24246.
23. Bunschoten, A.; Feitsma, L. J.; Kruijtzter, J. A.; de Haas, C. J.; Liskamp, R. M.; Kemmink, J. CHIPS binds to the phosphorylated N-terminus of the C5a-receptor. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3338-3340.
24. Mehta, A. Y.; Heimburg-Molinaro, J.; Cummings, R. D.; Goth, C. K. Emerging patterns of tyrosine sulfation and O-glycosylation cross-talk and co-localization. *Curr. Opin. Struct. Biol.* **2020**, *62*, 102-111.
25. Guo, R.-F.; Ward, P. A. Role of C5a in inflammatory responses. *Annu. Rev. Immunol.* **2005**, *23*, 821-852.
26. Ippel, J. H.; de Haas, C. J.; Bunschoten, A.; van Strijp, J. A.; Kruijtzter, J. A.; Liskamp, R. M.; Kemmink, J. Structure of the tyrosine-sulfated C5a receptor N terminus in complex with chemotaxis inhibitory protein of *Staphylococcus aureus*. *J. Biol. Chem.* **2009**, *284*, 12363-12372.
27. Yu, W.; Zhou, R.; Li, N.; Lei, Z.-C.; Guo, D.; Peng, F.; Li, Y.; Bai, X.; Feng, S.; Wang, Y.; He, J.; Yin, S.; Zeng, X.; He, L.; Gao, Y.; Li, M.; Guo, Y. R.; Liu, K.; Wang, Y. Histone tyrosine sulfation by SULT1B1 regulates H4R3me2a and gene transcription. *Nat. Chem. Biol.* **2023**, *19*, 855-864.
28. Zhao, Y.; Jensen, O. N. Modification-specific proteomics: strategies for characterization of post-translational modifications using enrichment techniques. *Proteomics* **2009**, *9*, 4632-4641.
29. Yang, Y. S.; Wang, C. C.; Chen, B. H.; Hou, Y. H.; Hung, K. S.; Mao, Y. C. Tyrosine sulfation as a protein post-translational modification. *Molecules* **2015**, *20*, 2138-64.
30. Pepi, L. E.; Sanderson, P.; Stickney, M.; Amster, I. J. Developments in Mass Spectrometry for Glycosaminoglycan Analysis: A Review. *Mol Cell Proteomics* **2021**, *20*, 100025.
31. Sarbu, M.; Ica, R.; Sharon, E.; Clemmer, D. E.; Zamfir, A. D. Identification and Structural Characterization of Novel Chondroitin/Dermatan Sulfate Hexassacharide Domains in Human Decorin by Ion Mobility Tandem Mass Spectrometry. *Molecules* **2022**, *27*, 6026.
32. Karst, N. A.; Linhardt, R. J. Recent chemical and enzymatic approaches to the synthesis of glycosaminoglycan oligosaccharides. *Current medicinal chemistry* **2003**, *10*, 1993-2031.
33. Salek, M.; Costagliola, S.; Lehmann, W. D. Protein tyrosine-O-sulfation analysis by exhaustive product ion scanning with minimum collision offset in a NanoESI Q-TOF tandem mass spectrometer. *Anal. Chem.* **2004**, *76*, 5136-5142.
34. Seibert, C.; Sakmar, T. P. Toward a framework for sulfoproteomics: synthesis and characterization of sulfotyrosine-containing peptides. *Peptide Science* **2008**, *90*, 459-477.
35. Onnerfjord, P.; Heathfield, T. F.; Heinegård, D. Identification of tyrosine sulfation in extracellular leucine-rich repeat proteins using mass spectrometry. *J. Biol. Chem.* **2004**, *279*, 26-33.
36. Kim, J.-S.; Song, S.-U.; Kim, H.-J. Simultaneous identification of tyrosine phosphorylation and sulfation sites utilizing tyrosine-specific bromination. *J. Am. Soc. Mass. Spectrom.* **2011**, *22*, 1916-1925.
37. Taylor, S. W.; Sun, C.; Hsieh, A.; Andon, N. L.; Ghosh, S. S. A sulfated, phosphorylated 7 kDa secreted peptide characterized by direct analysis of cell culture media. *J. Proteome Res.* **2008**, *7*, 795-802.

38. Medzihradzky, K. F.; Darula, Z.; Perlson, E.; Fainzilber, M.; Chalkley, R. J.; Ball, H.; Greenbaum, D.; Bogyo, M.; Tyson, D. R.; Bradshaw, R. A.; Burlingame, A. L. O-sulfonation of serine and threonine: mass spectrometric detection and characterization of a new posttranslational modification in diverse proteins throughout the eukaryotes. *Mol. Cell. Proteomics* **2004**, *3*, 429-40.
39. Sant, A. J.; Cullen, S. E.; Schwartz, B. D., Identification of a sulfate-bearing molecule associated with HLA class II antigens. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81* (5), 1534-1538.
40. Hoffhines, A. J.; Damoc, E.; Bridges, K. G.; Leary, J. A.; Moore, K. L., Detection and purification of tyrosine-sulfated proteins using a novel anti-sulfotyrosine monoclonal antibody. *J. Biol. Chem.* **2006**, *281* (49), 37877-87.
41. Lawrie, J.; Waldrop, S.; Morozov, A.; Niu, W.; Guo, J., Engineering of a small protein scaffold to recognize sulfotyrosine with high specificity. *ACS chemical biology* **2021**, *16* (8), 1508-1517.
42. Rosell-García, T.; Paradela, A.; Bravo, G.; Dupont, L.; Bekhouche, M.; Colige, A.; Rodriguez-Pascual, F., Differential cleavage of lysyl oxidase by the metalloproteinases BMP1 and ADAMTS2/14 regulates collagen binding through a tyrosine sulfate domain. *J. Biol. Chem.* **2019**, *294* (29), 11087-11100.
43. Schiza, C.; Korbakis, D.; Panteleli, E.; Jarvi, K.; Drabovich, A. P.; Diamandis, E. P., Discovery of a human testis-specific protein complex TEX101-DPEP3 and selection of its disrupting antibodies. *Mol. Cell. Proteomics* **2018**, *17* (12), 2480-2495.
44. Chen, Y.; Jin, S.; Zhang, M.; Hu, Y.; Wu, K. L.; Chung, A.; Wang, S.; Tian, Z.; Wang, Y.; Wolynes, P. G.; Xiao, H., Unleashing the potential of noncanonical amino acid biosynthesis to create cells with precision tyrosine sulfation. *Nat Commun* **2022**, *13* (1), 5434.
45. Cimbri, R.; Gallant, T. R.; Dolan, M. A.; Guzzo, C.; Zhang, P.; Lin, Y.; Miao, H.; Van Ryk, D.; Arthos, J.; Gorshkova, I.; Brown, P. H.; Hurt, D. E.; Lusso, P., Tyrosine sulfation in the second variable loop (V2) of HIV-1 gp120 stabilizes V2-V3 interaction and modulates neutralization sensitivity. *Proc Natl Acad Sci U S A* **2014**, *111* (8), 3152-7.
46. He, X.; Chen, Y.; Beltran, D. G.; Kelly, M.; Ma, B.; Lawrie, J.; Wang, F.; Dodds, E.; Zhang, L.; Guo, J.; Niu, W., Functional genetic encoding of sulfotyrosine in mammalian cells. *Nat Commun* **2020**, *11* (1), 4820.
47. Italia, J. S.; Peeler, J. C.; Hillenbrand, C. M.; Latour, C.; Weerapana, E.; Chatterjee, A., Genetically encoded protein sulfation in mammalian cells. *Nat. Chem. Biol.* **2020**, *16* (4), 379-382.
48. Pejchal, R.; Walker, L. M.; Stanfield, R. L.; Phogat, S. K.; Koff, W. C.; Poignard, P.; Burton, D. R.; Wilson, I. A., Structure and function of broadly reactive antibody PG16 reveal an H3 subdomain that mediates potent neutralization of HIV-1. *Proc Natl Acad Sci U S A* **2010**, *107* (25), 11483-8.
49. Wang, C. C.; Chen, B. H.; Lu, L. Y.; Hung, K. S.; Yang, Y. S., Preparation of Tyrosylprotein Sulfotransferases for *In vitro* One-Pot Enzymatic Synthesis of Sulfated Proteins/Peptides. *ACS Omega* **2018**, *3* (9), 11633-11642.
50. Xu, J.; Deng, X.; Tang, M.; Li, L.; Xiao, L.; Yang, L.; Zhong, J.; Bode, A. M.; Dong, Z.; Tao, Y., Tyrosylprotein sulfotransferase-1 and tyrosine sulfation of chemokine receptor 4 are induced by Epstein-Barr virus encoded latent membrane protein 1 and associated with the metastatic potential of human nasopharyngeal carcinoma. *PLoS one* **2013**, *8* (3), e56114.
51. Lee, R.; Huttner, W., Tyrosine-O-sulfated proteins of PC12 pheochromocytoma cells and their sulfation by a tyrosylprotein sulfotransferase. *J. Biol. Chem.* **1983**, *258* (18), 11326-11334.
52. Kehoe, J. W.; Velappan, N.; Walbolt, M.; Rasmussen, J.; King, D.; Lou, J.; Knopp, K.; Pavlik, P.; Marks, J. D.; Bertozzi, C. R., Using phage display to select antibodies recognizing post-translational modifications independently of sequence context. *Mol. Cell. Proteomics* **2006**, *5* (12), 2350-2363.
53. Huttner, W. B., Determination and occurrence of tyrosine O-sulfate in proteins. In *Methods Enzymol.*, Elsevier: 1984; Vol. 107, pp 200-223.

54. Liu, T.-A.; Yasuda, S.; Williams, F. E.; Liu, M.-Y.; Suiko, M.; Sakakibara, Y.; Yang, Y.-S.; Liu, M.-C., A target-specific approach for the identification of tyrosine-sulfated hemostatic proteins. *Anal. Biochem.* **2009**, *390* (1), 88-90.
55. Zhou, W.; Duckworth, B. P.; Geraghty, R. J., Fluorescent peptide sensors for tyrosylprotein sulfotransferase activity. *Anal. Biochem.* **2014**, *461*, 1-6.
56. Chen, B.-H.; Wang, C.-C.; Lu, L.-Y.; Hung, K.-S.; Yang, Y.-S., Fluorescence assay for protein post-translational tyrosine sulfation. *Analytical and bioanalytical chemistry* **2013**, *405*, 1425-1429.
57. Krogh, A.; Larsson, B.; Von Heijne, G.; Sonnhammer, E. L., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **2001**, *305* (3), 567-580.
58. Huang, S.-Y.; Shi, S.-P.; Qiu, J.-D.; Sun, X.-Y.; Suo, S.-B.; Liang, R.-P., PredSulSite: prediction of protein tyrosine sulfation sites with multiple features and analysis. *Anal. Biochem.* **2012**, *428* (1), 16-23.
59. Monigatti, F.; Gasteiger, E.; Bairoch, A.; Jung, E., The Sulfinator: predicting tyrosine sulfation sites in protein sequences. *Bioinformatics* **2002**, *18* (5), 769-770.
60. Chang, W. C.; Lee, T. Y.; Shien, D. M.; Hsu, J. B. K.; Horng, J. T.; Hsu, P. C.; Wang, T. Y.; Huang, H. D.; Pan, R. L., Incorporating support vector machine for identifying protein tyrosine sulfation sites. *J. Comput. Chem.* **2009**, *30* (15), 2526-2537.
61. McLean, J., The thromboplastic action of cephalin. *Am. J. Physiol.* **1916**, *41* (2), 250-257.
62. van der Meer, J.-Y.; Kellenbach, E.; Van den Bos, L. J., From farm to pharma: an overview of industrial heparin manufacturing methods. *Molecules* **2017**, *22* (6), 1025.
63. Li, J.; Zhang, Y.; Soubias, O.; Khago, D.; Chao, F.-a.; Li, Y.; Shaw, K.; Byrd, R. A., Optimization of sortase A ligation for flexible engineering of complex protein systems. *J. Biol. Chem.* **2020**, *295* (9), 2664-2675.
64. Toyoda, M.; Narimatsu, H.; Kameyama, A., Enrichment method of sulfated glycopeptides by a sulfate emerging and ion exchange chromatography. *Anal. Chem.* **2009**, *81* (15), 6140-6147.
65. Pieroni, L.; Iavarone, F.; Olanas, A.; Greco, V.; Desiderio, C.; Martelli, C.; Manconi, B.; Sanna, M. T.; Messina, I.; Castagnola, M., Enrichments of post-translational modifications in proteomic studies. *J. Sep. Sci.* **2020**, *43* (1), 313-336.
66. Balderrama, G. D.; Meneses, E. P.; Orihuela, L. H.; Hernández, O. V.; Franco, R. C.; Robles, V. P.; Batista, C. V. F., Analysis of sulfated peptides from the skin secretion of the *Pachymedusa dactinolor* frog using IMAC-Ga enrichment and high-resolution mass spectrometry. *Rapid Commun. Mass Spectrom.* **2011**, *25* (8), 1017-1027.
67. Robinson, M. R.; Brodbelt, J. S., Integrating weak anion exchange and ultraviolet photodissociation mass spectrometry with strategic modulation of peptide basicity for the enrichment of sulfopeptides. *Anal. Chem.* **2016**, *88* (22), 11037-11045.
68. Al-Horani, R. A.; Desai, U. R., Chemical sulfation of small molecules—advances and challenges. *Tetrahedron* **2010**, *66* (16), 2907.
69. Zulueta, M. M. L.; Hung, S.-C., Synthesis of Sulfated Glycans. In *Glycoscience: Biology and Medicine*, Taniguchi, N.; Endo, T.; Hart, G. W.; Seeberger, P. H.; Wong, C.-H., Eds. Springer Japan: Tokyo, 2015; pp 365-371.
70. Jin, H.; Chen, Q.; Zhang, Y.-Y.; Hao, K.-F.; Zhang, G.-Q.; Zhao, W., Preactivation-based, iterative one-pot synthesis of anticoagulant pentasaccharide fondaparinux sodium. *Organic Chemistry Frontiers* **2019**, *6* (17), 3116-3120.
71. Yao, W.; Xiong, D.-C.; Yang, Y.; Geng, C.; Cong, Z.; Li, F.; Li, B.-H.; Qin, X.; Wang, L.-N.; Xue, W.-Y.; Yu, N.; Zhang, H.; Wu, X.; Liu, M.; Ye, X.-S., Automated solution-phase multiplicative synthesis of complex glycans up to a 1,080-mer. *Nat. Synth.* **2022**, *1* (11), 854-863.

72. Ma, S.; Tang, N.; Tian, J., DNA synthesis, assembly and applications in synthetic biology. *Curr. Opin. Chem. Biol.* **2012**, *16* (3-4), 260-267.
73. Sun, H.; Brik, A., The journey for the total chemical synthesis of a 53 kDa protein. *Acc. Chem. Res.* **2019**, *52* (12), 3361-3371.
74. Hahm, H. S.; Broecker, F.; Kawasaki, F.; Mietzsch, M.; Heilbronn, R.; Fukuda, M.; Seeberger, P. H., Automated glycan assembly of oligo-N-acetyllactosamine and keratan sulfate probes to study virus-glycan interactions. *Chem* **2017**, *2* (1), 114-124.
75. Yang, W.; Eken, Y.; Zhang, J.; Cole, L. E.; Ramadan, S.; Xu, Y.; Zhang, Z.; Liu, J.; Wilson, A. K.; Huang, X., Chemical synthesis of human syndecan-4 glycopeptide bearing O-, N-sulfation and multiple aspartic acids for probing impacts of the glycan chain and the core peptide on biological functions. *Chem. Sci.* **2020**, *11*, 6393-6404.
76. Ali, A. M.; Taylor, S. D., Efficient solid-phase synthesis of sulfotyrosine peptides using a sulfate protecting-group strategy. *Angew. Chem. Int. Ed.* **2009**, *48* (11), 2024-2026.
77. YAGAMI, T.; SHIWA, S.; FUTAKI, S.; KITAGAWA, K., Evaluation of the final deprotection system for the solid-phase synthesis of Tyr (SO₃H)-containing peptides with 9-fluorenylmethyloxycarbonyl (Fmoc)-strategy and its application to the synthesis of cholecystokinin (CCK)-12. *Chem. Pharm. Bull.* **1993**, *41* (2), 376-380.
78. Stone, M. J.; Payne, R. J., Homogeneous sulfopeptides and sulfoproteins: synthetic approaches and applications to characterize the effects of tyrosine sulfation on biochemical function. *Acc. Chem. Res.* **2015**, *48* (8), 2251-61.
79. Liu, X.; Malins, L. R.; Roche, M.; Sterjovski, J.; Duncan, R.; Garcia, M. L.; Barnes, N. C.; Anderson, D. A.; Stone, M. J.; Gorry, P. R.; Payne, R. J., Site-selective solid-phase synthesis of a CCR5 sulfopeptide library to interrogate HIV binding and entry. *ACS Chem Biol* **2014**, *9* (9), 2074-
80. Ingram, L. J.; Taylor, S. D., Introduction of 2, 2, 2-Trichloroethyl-Protected Sulfates into Monosaccharides with a Sulfuryl Imidazolium Salt and Application to the Synthesis of Sulfated Carbohydrates. *Angew. Chem. Int. Ed.* **2006**, *45* (21), 3503-3506.
81. Farzan, M.; Mirzabekov, T.; Kolchinsky, P.; Wyatt, R.; Cayabyab, M.; Gerard, N. P.; Gerard, C.; Sodroski, J.; Choe, H., Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1 entry. *Cell* **1999**, *96* (5), 667-676.
82. DeAngelis, P. L.; Liu, J.; Linhardt, R. J., Chemoenzymatic synthesis of glycosaminoglycans: Re-creating, re-modeling and re-designing nature's longest or most complex carbohydrate chains. *Glycobiology* **2013**, *23* (7), 764-777.
83. Zhang, X.; Lin, L.; Huang, H.; Linhardt, R. J., Chemoenzymatic synthesis of glycosaminoglycans. *Acc. Chem. Res.* **2019**, *53* (2), 335-346.
84. Xu, Y.; Masuko, S.; Takeddin, M.; Xu, H.; Liu, R.; Jing, J.; Mousa, S. A.; Linhardt, R. J.; Liu, J., Chemoenzymatic synthesis of homogeneous ultralow molecular weight heparins. *Science* **2011**, *334* (6055), 498-501.
85. Petitou, M.; Jacquinet, J.-C.; Sinay, P.; Choay, J.; Lormeau, J.-C.; Nassr, M., Process for the organic synthesis of oligosaccharides and derivatives thereof. US4818816A: 1989.
86. Loganathan, D.; Wang, H. M.; Mallis, L. M.; Linhardt, R. J., Structural variation in the antithrombin III binding site region and its occurrence in heparin from different sources. *Biochemistry* **1990**, *29* (18), 4362-4368.
87. Fleisher, R. C.; Michael, N.; Gonzalez Jr, R. L., Mechanistic studies of non-canonical amino acid mutagenesis. In *Methods Enzymol.*, Elsevier: 2021; Vol. 656, pp 375-428.
88. Liu, C. C.; Schultz, P. G., Recombinant expression of selectively sulfated proteins in *Escherichia coli*. *Nat. Biotechnol.* **2006**, *24* (11), 1436-1440.

89. Liu, C. C.; Cellitti, S. E.; Geierstanger, B. H.; Schultz, P. G., Efficient expression of tyrosine-sulfated proteins in *E. coli* using an expanded genetic code. *Nat. Protoc.* **2009**, *4* (12), 1784-1789.
90. Zheng, L.; Du, G., A new path to tyrosine sulfation. *Nat Chem Biol* **2020**, *16* (4), 364-365.
91. Farzan, M.; Babcock, G. J.; Vasilieva, N.; Wright, P. L.; Kiprilov, E.; Mirzabekov, T.; Choe, H., The role of post-translational modifications of the CXCR4 amino terminus in stromal-derived factor 1 α association and HIV-1 entry. *J. Biol. Chem.* **2002**, *277* (33), 29484-29489.
92. Qiao, M.; Lin, L.; Xia, K.; Li, J.; Zhang, X.; Linhardt, R. J., Recent advances in biotechnology for heparin and heparan sulfate analysis. *Talanta* **2020**, *219*.

Chapter 2

A General Approach to *O*-Sulfation via Sulfur(VI) Fluoride Exchange Reaction

A significant portion of the work described in this chapter has been published in:

Liu, C.‡; Yang, C.‡; Hwang, S.; Ferraro, S. L.; Flynn, J. P.; Niu, J., A General Approach to *O*-Sulfation by a Sulfur(VI) Fluoride Exchange Reaction. *Angew. Chem. Int. Ed.* **2020**, *59*, 18435-18441.

2.1 INTRODUCTION

2.1.1 Overview

O-Sulfation is an important chemical code widely existing in bioactive molecules, but the scalable and facile synthesis of complex bioactive molecules carrying O-sulfation remains challenging. Herein, we report a general approach to O-sulfation via the Sulfur (VI) Fluoride Exchange (SuFEx) reaction between aryl fluorosulfates and silylated hydroxyl groups. Efficient sulfate diester formation was achieved through systematic optimization of the electronic properties of aryl fluorosulfates. The versatility of this O-sulfation strategy was demonstrated in the scalable syntheses of a variety of complex molecules carrying sulfate diesters at various positions, including monosaccharides, disaccharides, amino acid, and steroid. Selective hydrolytic and hydrogenolytic removal of the aryl masking groups from sulfate diesters yielded the corresponding O-sulfated products in excellent yields. This strategy provides a powerful tool for the synthesis of O-sulfated bioactive compounds.

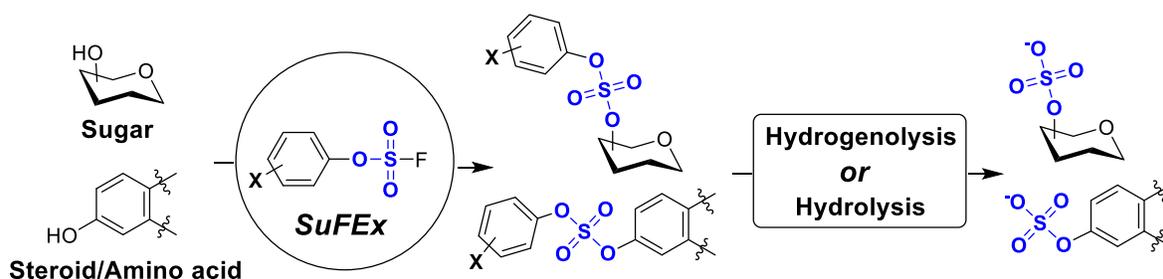


Figure 2-1. SuFEx-based rapid access to sulfated molecules

2.1.2 Significance of sulfation and major challenge

O-Sulfation widely exists in polysaccharides, liposaccharides, peptides, proteins, marine natural products, and drug metabolites in nature (Figure 2-1). The spatiotemporal distribution of the O-sulfation in these molecules plays important roles in a variety of biological activities such as telomerase inhibition,¹ cell signaling,² anticoagulation,³ drug detoxification,⁴ and cancer metastasis.⁵ However, the

lack of synthetic tools for efficient and scalable *O*-sulfation imposes a significant constraint on our abilities to access bioactive molecules carrying complex sulfation patterns and use them to study the structure–function relationships of the sulfate modifications in biology.

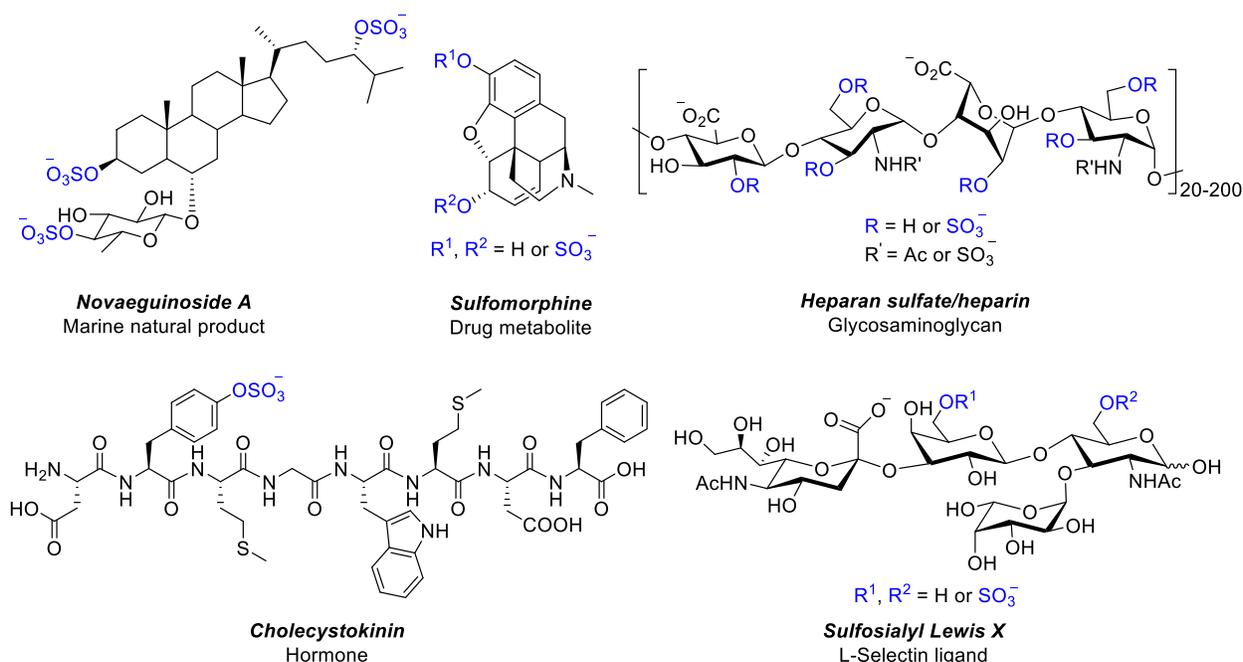


Figure 2-2. Diverse bioactive natural products consist of *O*-sulfation

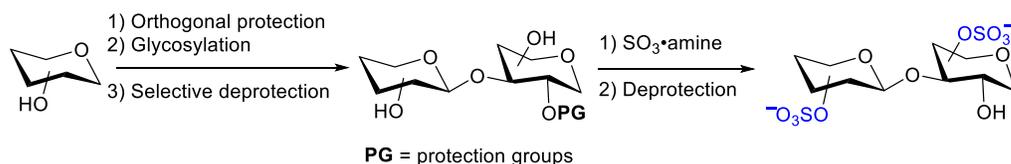
2.1.3 Current approaches

To date, sulfur trioxide–nitrogen base remains the most common chemical reagent to introduce *O*-sulfation to a variety of substrates (Figure 2-3a).^{6,7} *O*-Sulfation using this reagent can only be performed at the late stage of synthesis due to the challenging purification of the highly polar *O*-sulfated products and their chemical instability. These challenges are frequently exacerbated in the synthesis of polysulfated compounds of which *O*-sulfation at multiple residues are targeted.⁸ Efforts to address these deficiencies have led to the development of the so-called “early-stage” sulfation strategies, in which targeted hydroxyl groups were converted to sulfate diesters at the early stage of synthesis that were deprotected to generate *O*-sulfate in the final step (Figure 2-3a). Penney and Perlin

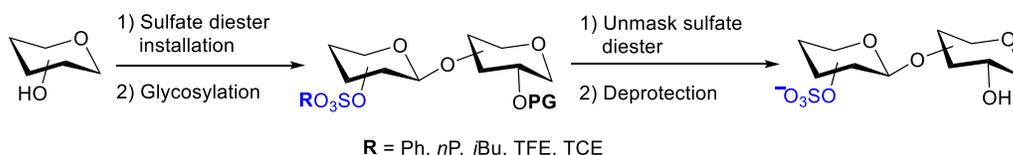
first reported the phenyl sulfate diester as a precursor of *O*-sulfation,⁹ but their method necessitated chemically labile reagents, harsh reaction conditions, and long reaction time. Further optimization of the reactivity and stability of the phenyl sulfate diester was not explored, either. Since this seminal work, other sulfate diesters were employed in early-stage *O*-sulfation approaches, including those with neopentyl (*nP*),¹⁰ isobutyl (*iBu*),¹⁰ trifluoroethylene (TFE),¹¹⁻¹³ and 2, 2, 2-trichloroethylene (TCE)¹⁴⁻¹⁷ groups. However, incompatibility with common reaction conditions and the deactivating effect on the modified carbohydrate substrates in glycosylation limited their utility in the synthesis of *O*-sulfated complex carbohydrates.¹⁸⁻¹⁹ Recently, enzymatic *O*-sulfation by sulfotransferases emerged as a promising strategy with excellent yields and regioselectivity (Figure 2-3b).²⁰⁻²³ However, the stringent specificity of sulfotransferase enzymes has caused the inflexible reaction sequence in polysulfation²¹ and limited substrate scope. The highly polar *O*-sulfated carbohydrates from enzymatic reactions also required high-resolution purification techniques.²³

a) Chemical *O*-sulfation

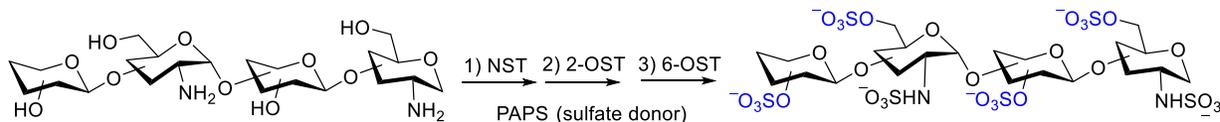
"Late stage" *O*-sulfation



"Early stage" *O*-sulfation



b) Enzymatic *O*-sulfation



NST: *N*-sulfotransferase; 2-OST and 6-OST: 2-*O*- and 6-*O*-sulfotransferase; PAPS: 3'-phosphoadenosine-5'-phosphosulfate

Figure 2-3. Existing approaches to access *O*-sulfation

2.1.4 Our strategy

Reported by Sharpless and coworkers in 2014, the SuFEx reaction has been employed to prepare a myriad of sulfate diesters.²⁴⁻²⁷ We envisioned that the SuFEx reaction between an aliphatic or aromatic silylether and an aryl fluorosulfate could serve as a general early-stage *O*-sulfation approach to site-specifically install sulfate diesters onto carbohydrate and non-carbohydrate substrates (Figure 2-4). Subsequent selective deprotection of the aryl sulfate monoester would lead to the desired *O*-sulfated compounds. Compared to the existing chemical and enzymatic *O*-sulfation strategies, advantages of this approach include the ability to balance stability and reactivity to accommodate different substrates and achieve efficient deprotection.

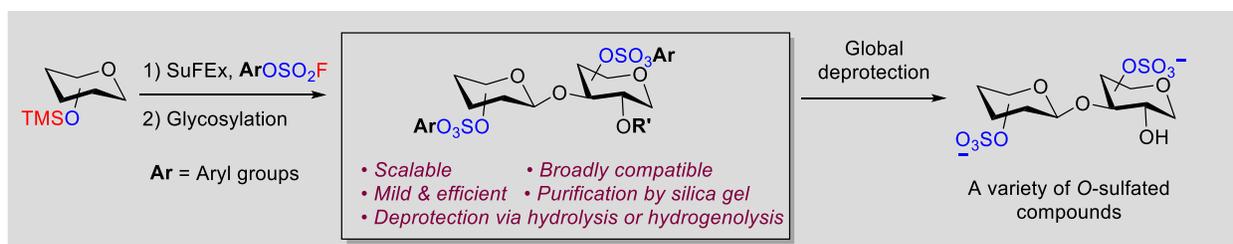


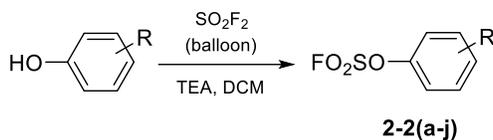
Figure 2-4. Our design: early-stage *O*-sulfation via SuFEx

2.2 RESULTS AND DISCUSSION

2.2.1 Initial screening with template reaction

We first investigated the model reaction between trimethylsilyl (TMS)-protected galactopyranose (**2-1**)²⁸ and substituted aryl fluorosulfates. Aryl fluorosulfates **2-2(a-j)** carrying variable substitutions ranging from strongly electron-withdrawing groups to electron-donating groups were readily prepared from corresponding phenols (Table 2-1).²⁶⁻²⁷

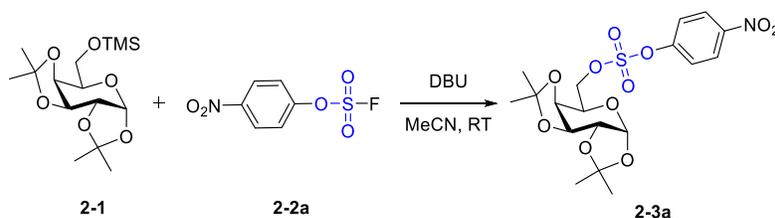
Table 2-1. Preparation of phenyl fluorosulfates with various electron-withdrawing and electron-donating substitutions



| Number | 2-2a | 2-2b | 2-2c | 2-2d | 2-2e | 2-2f | 2-2g | 2-2h | 2-2i | 2-2j |
|-------------|---------------------------|---------------------------|---------------------------|--------------|--------------|--------------|--------------|-------------|--------------|---------------|
| Substituent | <i>p</i> -NO ₂ | <i>m</i> -NO ₂ | <i>p</i> -CF ₃ | <i>p</i> -CN | <i>p</i> -Br | <i>p</i> -Cl | <i>p</i> -Ph | <i>p</i> -H | <i>p</i> -Me | <i>p</i> -OMe |
| Yield | 95% | 99% | 85% | 96% | 96% | 91% | 99% | 82% | 60% | 98% |

A systematic screening identified that 0.2 equivalents of 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) efficiently catalyzed the coupling of **2-1** with **2-2a** at room temperature in acetonitrile with a satisfactory 94% yield (Table 2-2).

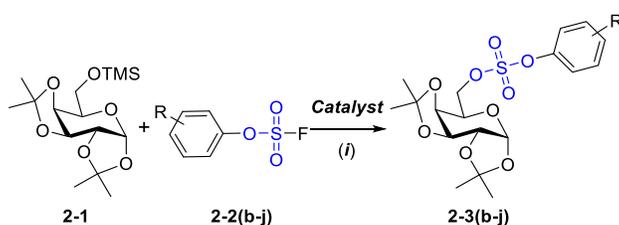
Table 2-2. Screening reaction factors for the coupling of **2-1** with **2-2a**



| 2a (equiv.) | DBU (equiv.) | [M] (mol/L) | Time (h) | Yield |
|--------------------|---------------------|--------------------|-----------------|--------------|
| 1.1 | 0.1 | 0.3 | 2 | 89% |
| 1.1 | 0.1 | 0.3 | 1 | 86% |
| 1.1 | 0.1 | 0.6 | 2 | 88% |
| 2 | 0.1 | 0.3 | 2 | 89% |
| 1.1 | 0.2 | 0.3 | 2 | 94% |
| 2 | 0.2 | 0.3 | 2 | 94% |
| 1.1 | 0.5 | 0.3 | 2 | 89% |

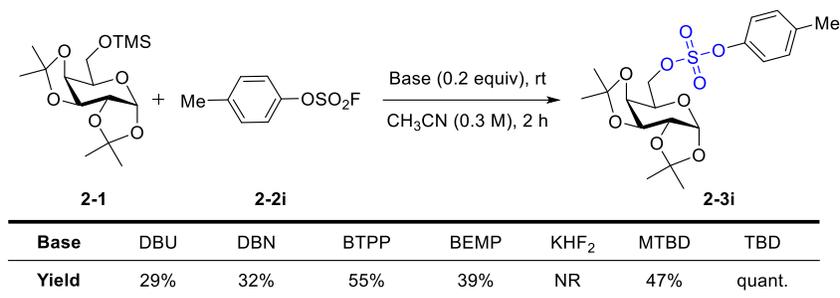
Excellent yields were observed for the preparation of **2-3(b-f)** when the optimized conditions were applied into the reaction of **2-1** with **2-2(b-f)** carrying electron-withdrawing substituents. In contrast, the coupling of **2-1** with **2-2(g-j)** carrying electron-neutral or electron-donating substitutions were sluggish, affording **2-3(g-j)** in modest to low yields from 76% to 29%. These results suggested that the reaction rates positively correlated with the electron deficiency of the aryl fluorosulfates (Table 2-3).

Table 2-3. The coupling of 1 with various fluorosulfate

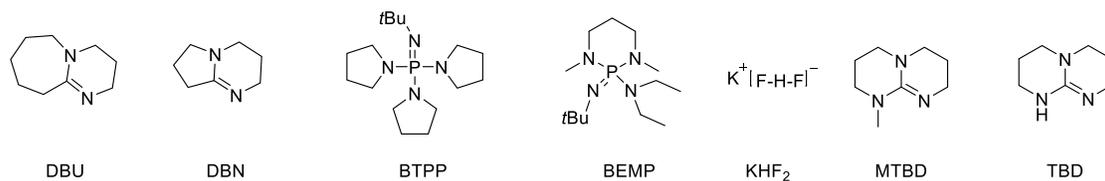


| Number | 2-3b | 2-3c | 2-3d | 2-3e | 2-3f | 2-3g | 2-3h | 2-3i | 2-3j |
|-------------|---------------------------|---------------------------|--------------|--------------|--------------|--------------|-------------|--------------|---------------|
| Substituent | <i>m</i> -NO ₂ | <i>p</i> -CF ₃ | <i>p</i> -CN | <i>p</i> -Br | <i>p</i> -Cl | <i>p</i> -Ph | <i>p</i> -H | <i>p</i> -Me | <i>p</i> -OMe |
| Yield | 92% | 91% | 95% | 93% | 86% | 76% | 43% | 29% | 31% |

Various catalysts were investigated for their efficiencies in catalyzing the SuFEx reaction between *p*-methylphenyl fluorosulfate **2-2i** and O-6 TMS silylated galactose substrate **2-1**, including 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU), 1,5,7-triazabicyclo[4.4.0]dec-5-ene (TBD), 1,5-diazabicyclo[4.3.0]non-5-ene (DBN),²⁹ 1,5,7-triazabicyclo[4.4.0]dec-5-ene (MTBD),³⁰ 2-tert-butylimino-2-diethylamino-1,3-dimethylperhydro-1,3,2-diazaphosphorine (BEMP), tert-Butylimino-tri(pyrrolidino)phosphorene (BTTP), and potassium bifluoride (Table 2-4). To our delight, the SuFEx reaction efficiency of the electron-rich aryl fluorosulfates **2-2i** with **2-1** could be improved to quantitative when 1,5,7-triazabicyclo[4.4.0]dec-5-ene (TBD)³¹ was used as catalyst.

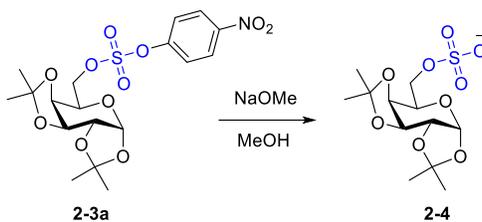
Table 2-4. Catalyst screening for the coupling of **2-2i** with **2-1**

NR: There was no reaction at room temperature or 80 °C for 2 hours.



We then sought to generate the desired 6-*O*-sulfated galactose (**4**) through the selective deprotection of the aryl groups in **3a-j**. Previous studies have shown that the fission rate of the two S–O bonds in a sulfate diester is determined by the pK_a difference between the two leaving groups.³² Consistently, we observed that **2-4** was efficiently generated as the sole product when **2-3a** carrying strongly electron-withdrawing groups were treated with 5 M sodium methoxide in methanol at room temperature (Table 2-5). A real-time ¹H-NMR monitoring of the reaction further confirmed the clean transformation from **2-3a** to **2-4** in 1 hour (Figure 2-5). Similar results were obtained for **2-3(c-d)** when the optimized alkaline hydrolysis conditions applied. On the other hand, the hydrolysis of **2-3(e-j)** carrying weak electron-withdrawing or electron-neutral groups was inefficient (Table 2-6). The trend became more clear by comparing hydrolytic kinetics of several sulfate diesters varying in electronic properties (Figure 2-6). The feasibility of sulfate diester hydrolysis positively correlated to their electron-deficiency.

Table 2-5. Hydrolysis of **2-3a**



| NaOMe (eq) | [NaOMe] (M) | Time (h) | Yield (%) |
|------------|-------------|----------|-----------|
| 50 | 2.5 | 4 | 94 |
| 50 | 2.5 | 1 | 44 |
| 100 | 5 | 1 | 92 |

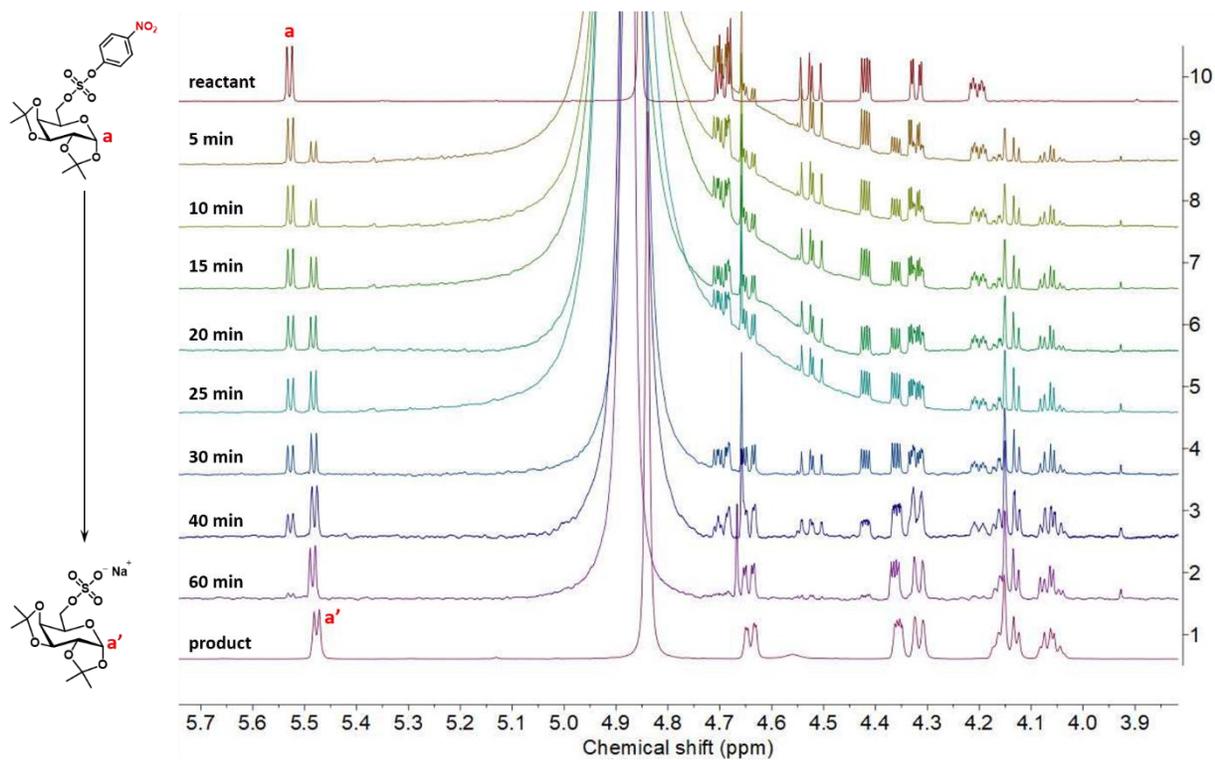
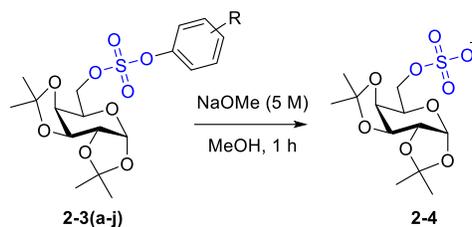


Figure 2-5. Kinetics of the hydrolysis of **2-3a** monitored by ¹H NMR in CD₃OD

Table 2-6. Hydrolysis of a variety of sulfate diester **2-3(a-j)**



| Number | 2-3a | 2-3b | 2-3c | 2-3d | 2-3e | 2-3f | 2-3g | 2-3h | 2-3i | 2-3j |
|-------------|---------------------------|---------------------------|---------------------------|--------------|--------------|--------------|--------------|-------------|--------------|---------------|
| Substituent | <i>p</i> -NO ₂ | <i>m</i> -NO ₂ | <i>p</i> -CF ₃ | <i>p</i> -CN | <i>p</i> -Br | <i>p</i> -Cl | <i>p</i> -Ph | <i>p</i> -H | <i>p</i> -Me | <i>p</i> -OMe |
| Yield | 92% | 96% | 98% | 95% | 38% | 40% | 10% | 5% | 0 | 0 |

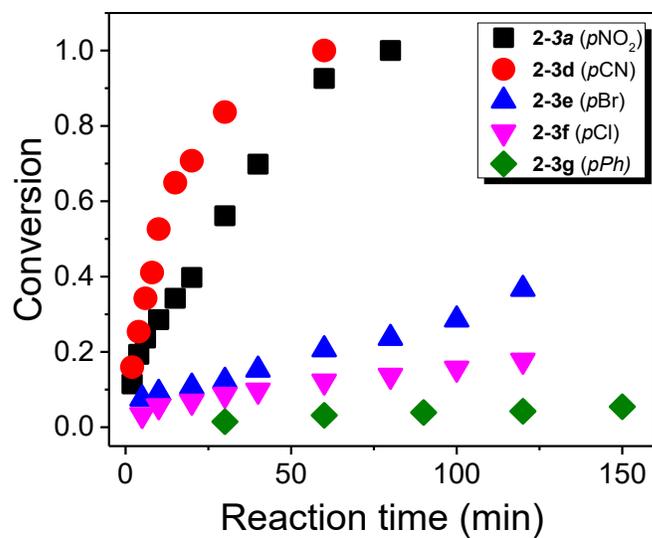


Figure 2-6. Effect of aromatic substitutions on the kinetics of the hydrolysis of sulfate diesters

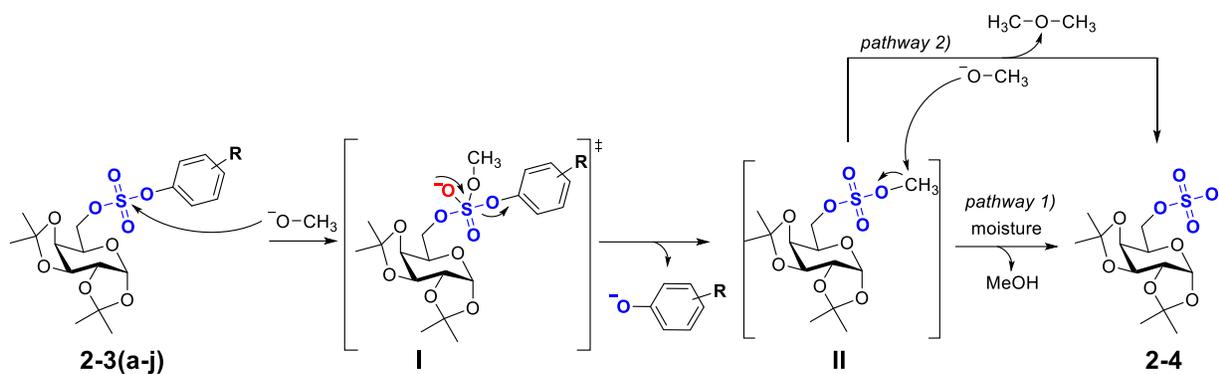


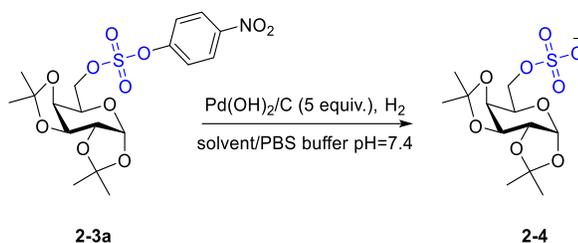
Figure 2-7. Proposed mechanism of hydrolysis

The proposed mechanism of this reaction involved the formation of a glycosyl methyl sulfate diester intermediate, which was subsequently converted into the *O*-sulfated compound through two possible pathways (Figure 2-7): 1) hydrolysis by reacting with the trace amount of water present in the methanolic solution of sodium methoxide (S–O fission);³²⁻³³ 2) nucleophilic substitution at the methyl group (C–O fission).³⁴⁻³⁵ In the presence of NaOMe, the nucleophilic attack of the methoxide anion towards the sulfur (VI) center generates **I**. The release of the phenolic anion leaving group provides **II**. The properties of the leaving groups vary from excellent to poor following the electron-withdrawing effect of the phenyl substituents, which is consistent with the trend shown in Figure S2. The methyl glycosyl sulfate diester underwent two possible pathways to afford final sulfated compound **4**. *Pathway 1*). Intermediate **II** hydrolyzed by reacting with the trace amount of water present in the methanolic solution of sodium methoxide (S–O fission) to release methanol and sulfated compound **4**;³²⁻³³ *Pathway 2*). Intermediate **II** underwent nucleophilic substitution at the methyl group (C–O fission) to release dimethyl ether and sulfated compound **2-4**.³⁴⁻³⁵

Concerns over undesired side reactions caused by strong base treatment³⁶⁻³⁷ prompted us to investigate an alternative deprotection strategy. Penny and Perlin first reported hydrogenolysis of phenyl sulfate monoester using PtO₂ and hydrogen gas (H₂) in modest yields.⁹ Inspired by this work,

we studied the selective hydrogenolysis of the aryl sulfate monoester in glycosyl aryl sulfate diesters by Pd(OH)₂/C/H₂, and observed high deprotection efficiency in an acetonitrile/methanol/water solution of phosphate salts (Table 2-7).

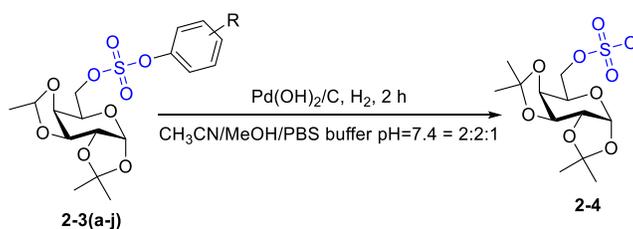
Table 2-7. Optimization of the hydrogenolysis of **2-3a**



| [M] (mM) | Solvent/PBS buffer | Reaction time(h) | Yield |
|----------|---------------------|------------------|-------|
| 30 | MeOH/PBS=5/2 | 1 | 51% |
| 25 | MeOH/MeCN/PBS=2/3/1 | 1 | 86% |
| 20 | MeOH/MeCN/PBS=2/2/1 | 2 | 96% |

Notably, this condition is insensitive to the electronic properties of the aryl substitution, yielding **2-4** from all substrates **2-3(a-j)** in excellent yields at room temperature (Table 2-8).

Table 2-8. Hydrogenolysis of sulfate diester **2-3(a-j)**



| Number | 2-3a | 2-3b | 2-3c | 2-3d | 2-3e | 2-3f | 2-3g | 2-3h | 2-3i | 2-3j |
|-------------|---------------------------|---------------------------|---------------------------|--------------|--------------|--------------|--------------|-------------|--------------|---------------|
| Substituent | <i>p</i> -NO ₂ | <i>m</i> -NO ₂ | <i>p</i> -CF ₃ | <i>p</i> -CN | <i>p</i> -Br | <i>p</i> -Cl | <i>p</i> -Ph | <i>p</i> -H | <i>p</i> -Me | <i>p</i> -OMe |
| Yield | 96% | 94% | 93% | 95% | quant. | 97% | 96% | quant. | quant. | 93% |

quant.: quantitative yield

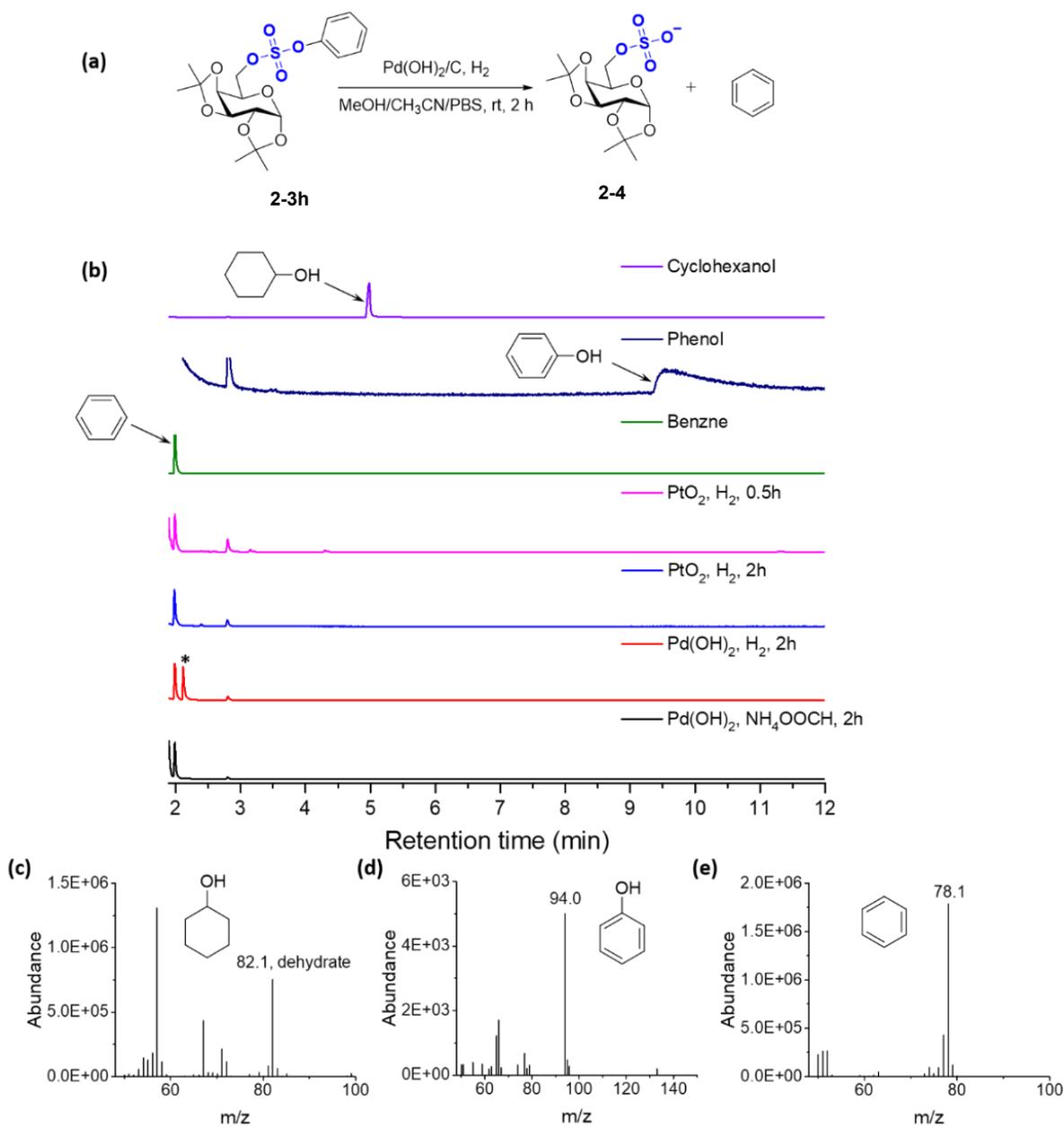


Figure 2-8. GC-MS analysis of the hydrogenolysis of **2-3h**. (a) Scheme of the reaction. (b) From top to bottom: standard GC traces of cyclohexanol, phenol, and benzene, and GC traces of the hydrogenolysis reactions under following conditions: PtO₂/H₂, 0.5 h; PtO₂/H₂, 2 h; Pd(OH)₂/C/H₂, 2h; Pd(OH)₂/C/NH₄OOCH, 2h. A byproduct, trimethylamine (labeled with *), was formed from the hydrogenation of acetonitrile by Pd(OH)₂/C/H₂.³⁸⁻⁴⁰ (c)-(e): GC-MS spectra of cyclohexanol, phenol, and benzene.

In order to gain insight of the hydrogenolytic decaging mechanism, we applied gas chromatography–mass spectrometry (GC-MS) to detect what was released from the reaction mixture. By comparing to the standard GC traces of cyclohexanol, phenol, and benzene, we observed that,

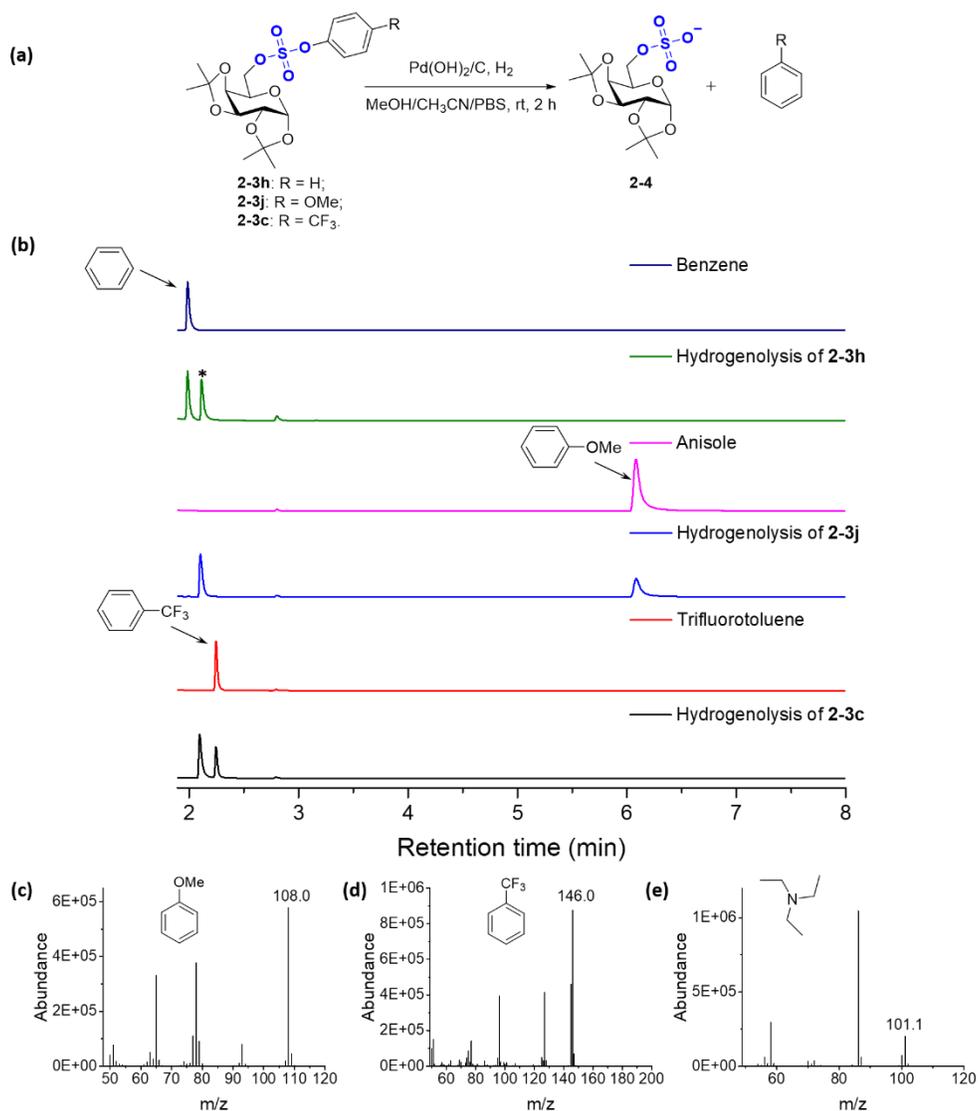


Figure 2-9. GC-MS analysis of the hydrogenolysis of sulfate diesters with various substitutions. (a) Scheme of reaction. (b) From top to bottom: standard GC trace of benzene, the hydrogenolysis of **2-3h**, standard GC trace of anisole, the hydrogenolysis of **2-3f**, standard GC trace of trifluorotoluene, and the hydrogenolysis of **2-3c**. A byproduct, trimethylamine (labeled with *), was formed from the hydrogenation of acetonitrile by $\text{Pd}(\text{OH})_2/\text{C}/\text{H}_2$.³⁸⁻⁴⁰ (c)-(e): GC-MS spectra of anisole, trifluorotoluene, and triethylamine.

besides the sulfated carbohydrate, benzene was generated in the hydrogenolysis process of **2-3h** while no other byproducts were observed (Figure 2-8). The GC profile did not change when the reaction time was extended to two hours. Switching the catalyst from platinum oxide to palladium hydroxide

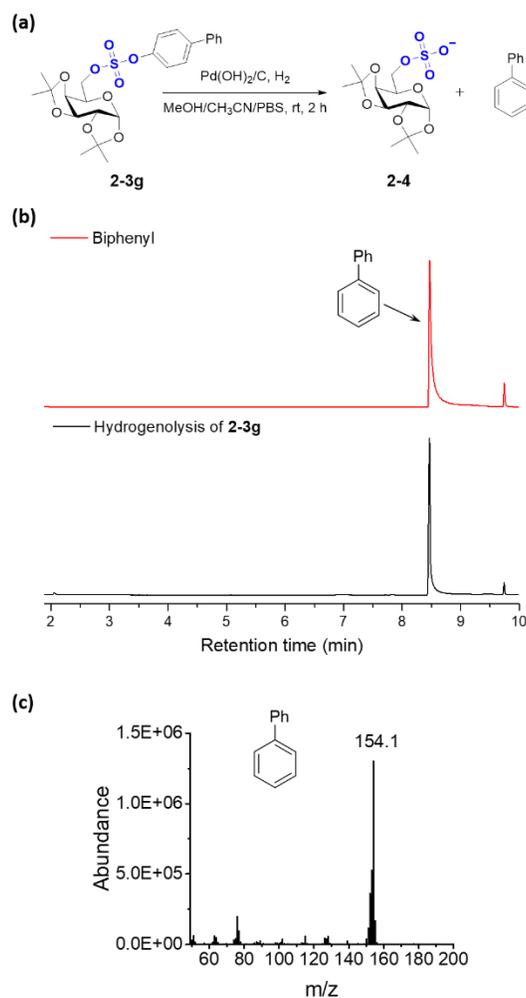


Figure 2-10. GC-MS analysis of the hydrogenolysis of **2-3g**. (a) Scheme of the reaction. (b) From top to bottom: standard GC trace of biphenyl and the hydrogenolysis of **2-3g**. (c)-(e): GC-MS spectra of biphenyl.

produced similar results. Furthermore, various hydride sources including H_2 and ammonium formate can all be used for the hydrogenolysis of the sulfate diester. Moreover, aromatic side products were uniformly detected from the reaction mixture of electron-deficient, electron-rich to electron-neutral sulfate diester (Figure 2-9 and 2-10).

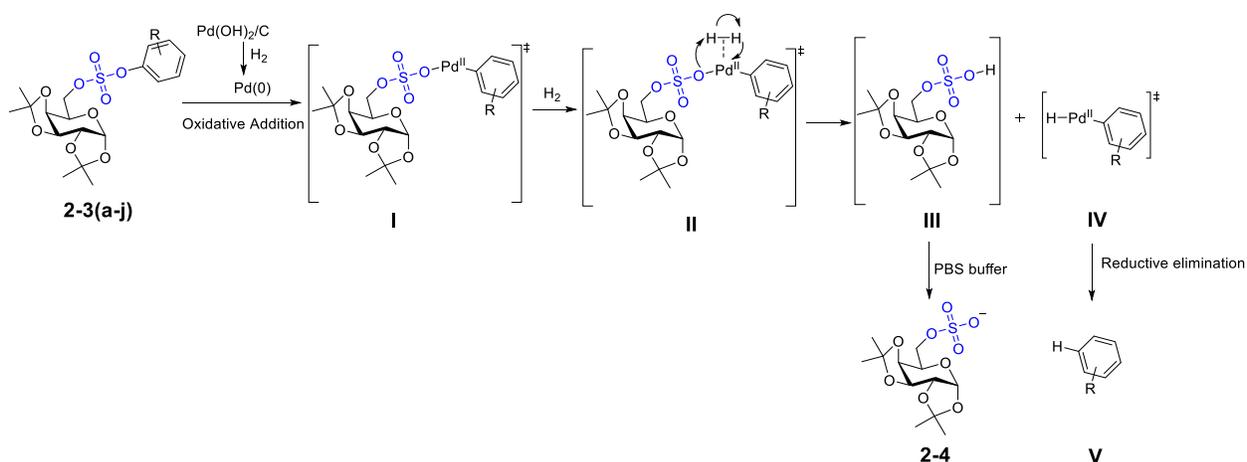


Figure 2-11. Our proposed mechanism of hydrogenolysis

Subsequently, gas chromatography-mass spectrometry (GC-MS) analysis of the hydrogenolysis reaction revealed a mechanism of oxidative addition of palladium into the aryl C–O bond followed by the reductive elimination of arenes, as opposed to the previously proposed mechanism of hydrogenation of the aryl group followed by S–O bond fission to release cyclohexanol (Figure 2-11).⁹ In this process, the catalyst palladium hydroxide on carbon is reduced to palladium (0) [Pd(0)] *in situ*. The oxidative addition of Pd(0) to the C—O bond of the aryl sulfate monoester generates **I** carrying a Pd(II) center. Subsequent coordination of H₂ to the palladium center of the catalyst leads to **II**. The dissociation of **II** yields galactosyl sulfuric acid **III** and Pd(II)-aryl complex **IV**. **III** is immediately quenched by PBS buffer to give O-sulfated galactose **2-4** while **IV** undergoes reductive elimination of arene **V** and regenerates the Pd(0) catalyst. Our observation that hydrogen gas can be replaced by ammonium formate, a far less potent hydrogenation agent, without affecting the hydrogenolysis efficiency, further proved the oxidative addition-reductive elimination mechanism.⁴¹⁻⁴³ Meanwhile, the comparison of proton spectra data indicates that sulfate **2-4** obtained from hydrolysis and hydrogenolysis are identical (Figure 2-12).

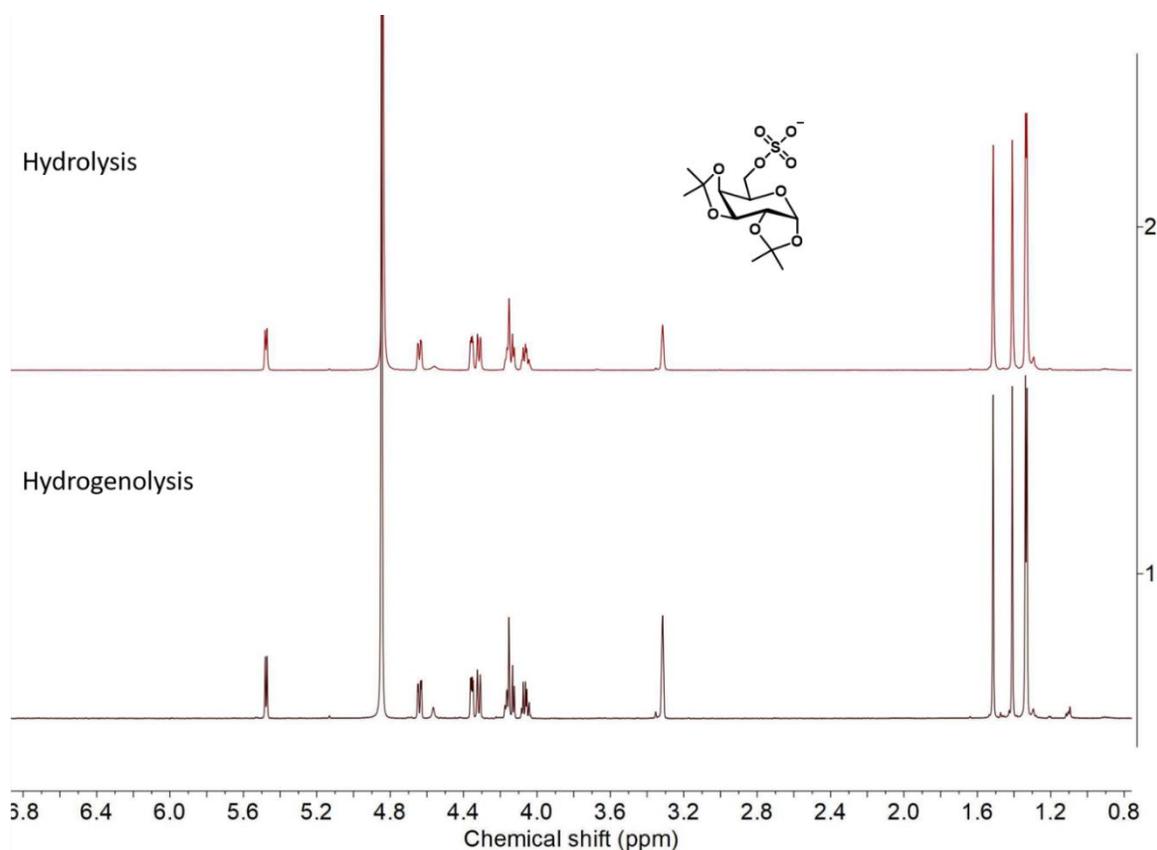
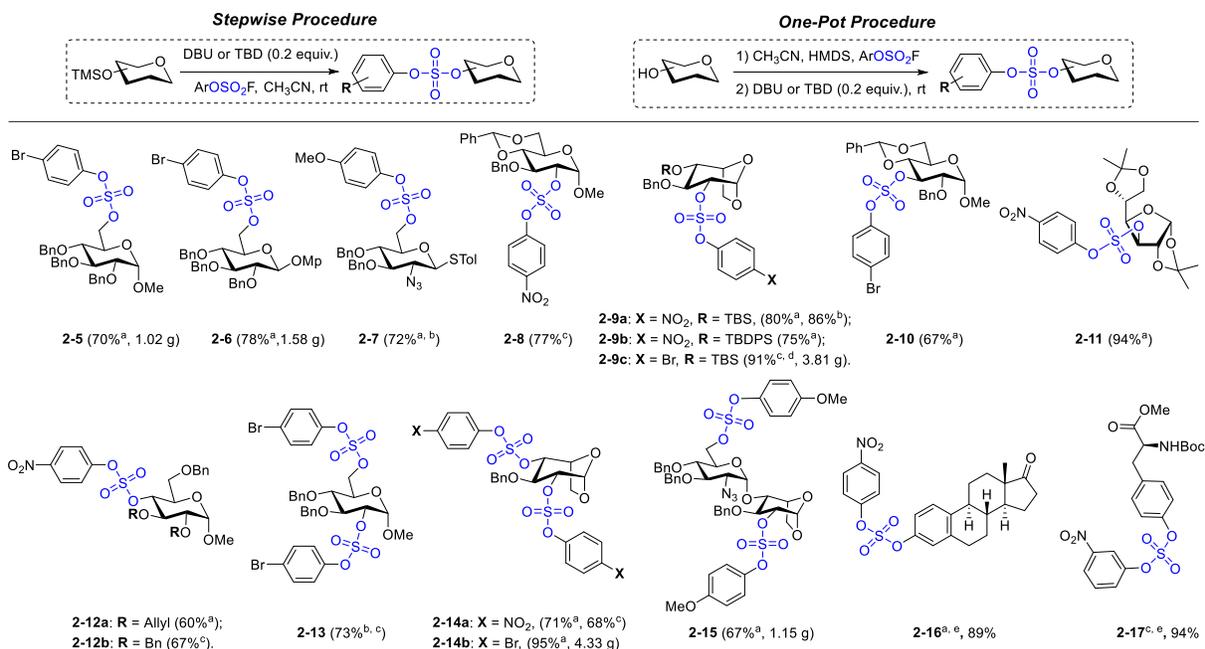


Figure 2-12. $^1\text{H-NMR}$ spectra of **4** from hydrolysis and hydrogenolysis process

Our subsequent investigation was focused on the feasibility of this strategy for a variety of substrates (Figure 2-13). Initial attempts to install *p*-nitrophenyl sulfate diester onto the *O*-6 position of methyl 2,3,4-*tri-O*-benzyl- α -D-glucopyranoside by reacting with **2-2a** generated an ether byproduct methyl 2,3,4-*tri-O*-benzyl-6-*O*-(*p*-nitrophenyl)- α -D-glucopyranoside along with the targeted sulfate diester. We attributed this side reaction to the susceptibility of the electron-deficient sulfate diesters to nucleophiles. Indeed, *O*-6 substitution using **2-2e** provided sulfate diesters **2-5** and **2-6** in 70% and 78% yields, respectively, without ether formation. Furthermore, *O*-6 substitution of the glucosyl thiol donor



^aIsolated yield from stepwise reaction. ^b0.3 equiv. of TBD was used. ^cIsolated yield from one-pot reaction. ^d0.1 equiv. of DBU was used. ^eTBS was used as the *O*-silyl protecting group. Bn: benzyl; Me: methyl; Mp: *para*-methoxyphenyl; Tol: tolyl; Boc: *tert*-butyloxycarbonyl.

Figure 2-13. Substrate scope of the established method

required **2-2j** carrying the strongly electron-donating *p*-methoxy group, affording **2-7** in 72% yield. In these reactions, we also discovered that aryl fluorosulfates remained stable in the presence of silylethers before the introduction of organobase catalysts, making it possible for a one-pot procedure combining *in situ* hydroxyl silylation by hexamethyldisilazane (HMDS)⁴⁴ and the subsequent SuFEx reaction. Following the one-pot procedure, sulfate diesters **2-8** was prepared in 77% yield. No major difference in reaction efficiency was found between the stepwise and one-pot procedures, as shown in the preparation of **2-9a** and **2-14a**. In addition to simplifying the operation, the one-pot procedure could also circumvent the challenge to isolate TMS silylethers that are often unstable on silica gel chromatography. Other silylether protecting groups, such as *tert*-butyldimethylsilyl (TBS) in **9a** and *tert*-butyldiphenylsilyl (TBDPS) in **2-9b**, remained stable in both stepwise and one-pot procedures. Sulfate diesters at different positions like *O*-3 and *O*-4 of glucose were also successfully obtained, giving products **2-10**, **2-11**, **2-12a** and **2-12b** in 67%, 94%, 60%, and 67% yields, respectively.⁴⁴

Moreover, **2-13** and **2-14a**, monosaccharides carrying two sulfate diesters were prepared successfully in 73% and 71% yields, respectively. The efficiency and scale of disubstitution could be further improved to 95% in multi-gram quantities when switching from **2-2a** to **2-2e**. Disubstituted disaccharide **2-15** was also synthesized in gram scale from the corresponding dual-TMS-protected substrate, with **2-2j** being found to enable the highest SuFEx efficiency of all aryl fluorosulfates tested at 67% yield. Besides carbohydrates, the SuFEx reaction was further applied to install sulfate diesters on non-carbohydrate substrates. Estrone sulfate diester **2-16** and tyrosine sulfate diester **2-17** were efficiently prepared in 89% and 94% yields, respectively. It is noteworthy that all compounds mentioned above were prepared in readily scalable procedures and purified by silica gel chromatography. Taken together, these examples highlight the broad substrate scope of our method and its versatility in tuning the electronic properties of the aryl fluorosulfates to optimize their reactivities.

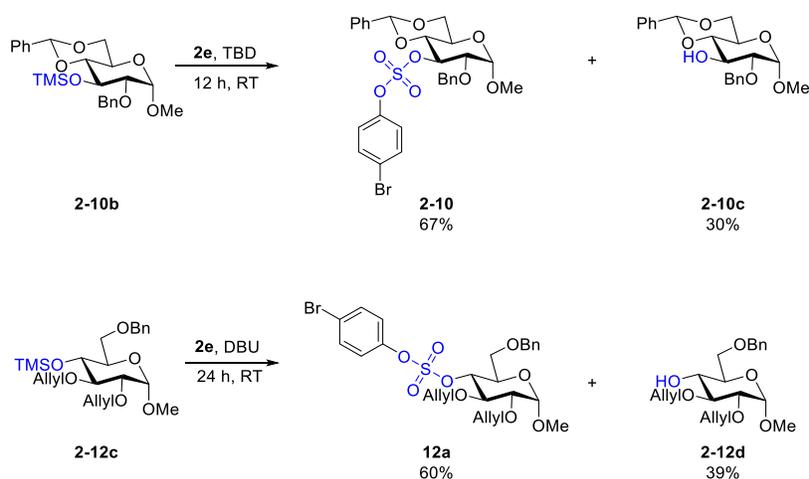


Figure 2-14. Side reactions in the SuFEx coupling to synthesize **2-10** and **2-12a**

It is noteworthy to mention that the side products isolated from the SuFEx coupling reaction were characterized to be the desilyl compounds like **2-10c** and **2-12d** which were in consistency with the previous report (Figure 2-14).⁴⁵⁻⁴⁶ We hypothesize that such byproducts could be generated from

two possible sources: (1) the carbohydrate alkoxide anions that failed to react; (2) transesterification of sulfate diesters in the presence of organobase catalysts and alkoxide anions.⁴⁷ This analysis suggested that precious starting materials used in this protocol were easily recovered from a practical perspective.

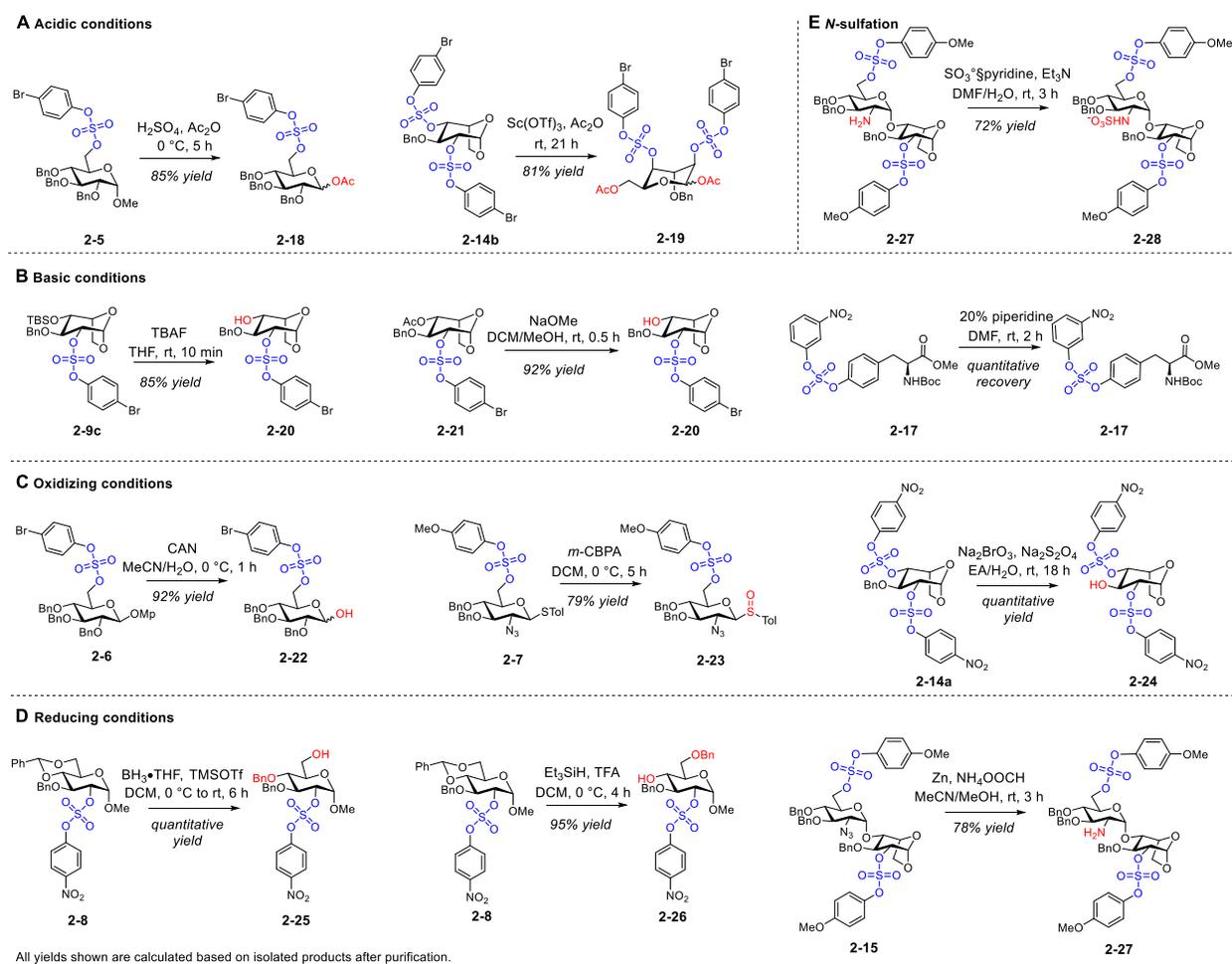
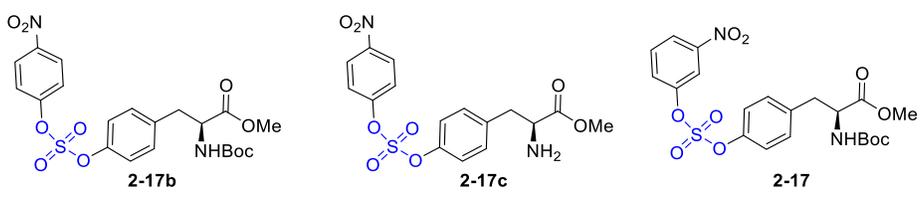


Figure 2-15. Compatibility of aryl sulfate diesters to common reagents used in carbohydrate and peptide chemistries

Next, the compatibility of our approach with frequently used reagents and reaction conditions in carbohydrate and peptide chemistries was examined (Figure 2-15). First, the aryl sulfate diesters as the protected sulfates could overcome the acid sensitivity of the nonprotected sulfates (Figure 2-15A). The anomeric *O*-methyl group of **5** was converted to the *O*-acetyl group in **2-18** by sulfuric acid in 85%

yield. Disulfated 1,6-anhydrosucrose **2-14b** was readily converted into the ring-opening product **2-19** in 81% yield by strong Lewis acid scandium triflate.⁴⁸⁻⁴⁹ Basic conditions were tolerated (Figure 2-15B). From **2-9c** and **2-21**, removal of the TBS and acetyl protecting groups by tetrabutylammonium fluoride (TBAF) and sodium methoxide afforded **2-20** in excellent yields. Notably, TBAF compatibility was a challenge for the state-of-the-art TCE sulfate diesters due to the tendency of the undesired fluorination.¹⁸ While tyrosine *p*-nitrophenylsulfate diester demonstrated limited stability in 20% piperidine in DMF (Table 2-9), the *m*-nitrophenyl sulfate diester **2-17** was perfectly stable under the same condition.

Table 2-9. Robustness screening for tyrosine sulfate diester



| Conditions | Compound | Conditions | Stability |
|------------|--------------|--|---|
| a | 2-17b | Benzylamine (2.0 equiv) + Et ₃ N (1.0 equiv) | 24 h, 92% recovery |
| b | 2-17b | Benzylamine (2.0 equiv) N,N-Diisopropylethylamine (1.0 equiv) | 24 h, 86% recovery |
| c | 2-17b | Trifluoroacetic acid | 4 days, 2-17c (de-Boc, quantitative) |
| d | 2-17b | 20% (v/v) Piperidine in DMF | 1 h, 54% recovery |
| e | 2-17 | 20% (v/v) Piperidine in DMF | 2 h, quantitative recovery |

Oxidizing conditions were also well tolerated (Figure 2-15C). The anomeric *p*-methoxyphenyl group of **2-6** was removed to form **2-22** in 92% yield by cerium ammonium nitrate (CAN). Oxidation of the thiol donor **7** by *m*-chloroperoxybenzoic acid (*m*-CPBA) afforded sulfoxide **2-23** in 79% yield. Oxidative debenylation of **2-14a** by sodium bromate and sodium dithionite quantitatively yielded glycosyl acceptors **2-24**.⁵⁰ Reducing conditions were tolerated (Figure 2-15D). Acid-promoted reductive benzylidene opening of **2-8** exposed *O*-6 or *O*-4 positions as free hydroxyls to afford **2-25** and **2-26** in quantitative and 95% yields, respectively. Zinc-catalyzed azide reduction of **2-15** afforded

the free amine **2-27** in 78% yield. Finally, the sulfate diesters are also compatible with the sulfur trioxide-nitrogen base reagent (Figure 2-15E). *N*-Sulfation of **2-27** by SO₃:pyridine generated **2-28** in 72% yield. Overall, these results indicated excellent compatibilities of the aryl sulfate diesters with acidic, basic, oxidizing, and reducing conditions commonly used in carbohydrate and peptide chemistries.

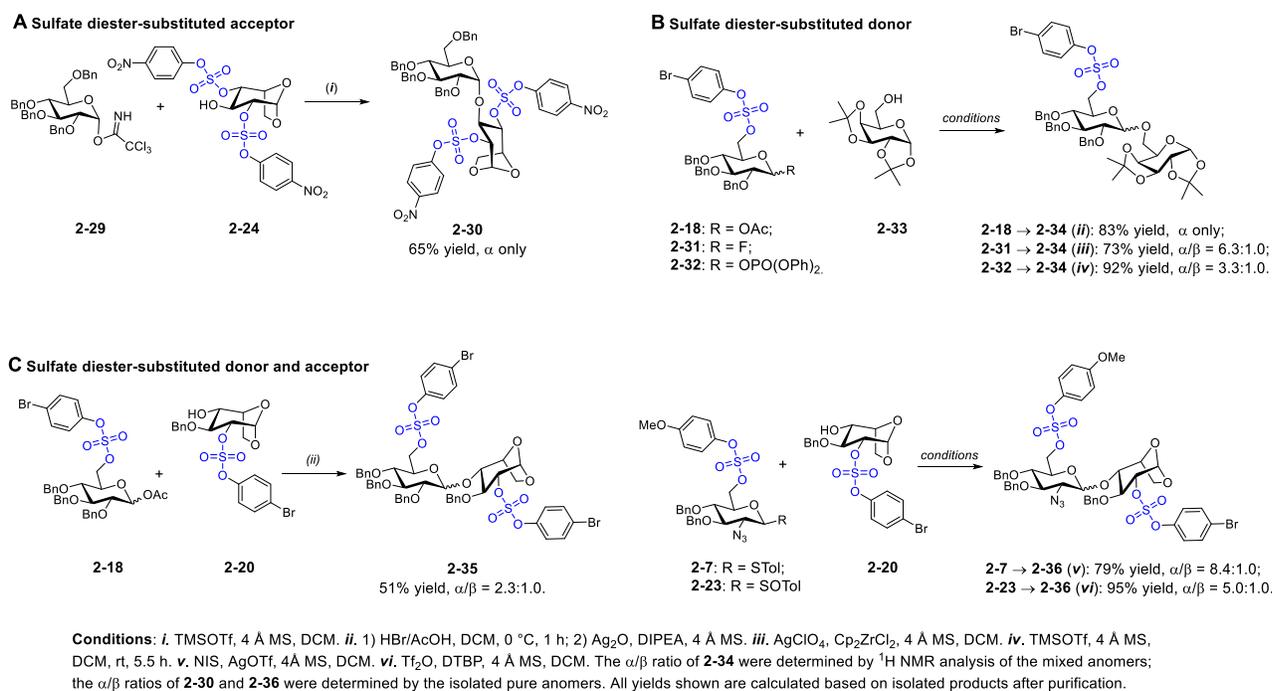


Figure 2-16. Glycosylation coupling with the sulfate masks on

We then examined the glycosylation reactions of both donors and acceptors modified by glycosyl aryl sulfate diesters (Figure 2-16). First, the disubstituted acceptor **2-24** was coupled with imidate donor **2-29**, successfully yielding disaccharides **2-30** in 65% yield (Figure 2-16A). Next, glycosylation of acceptor **2-33** using *O*-6 sulfate diester-substituted glucosyl donors including the bromide donor (*in situ* generated from 1-*O*-acetyl-glucopyranoside **2-18**), fluoride donor (**2-31**), and the phosphate donor (**2-32**) all proceeded efficiently, forming disaccharide **2-34** in excellent (73-92%) yields (Figure 2-16B). Because sulfate diester modifications could deactivate carbohydrate building

blocks in glycosylation reactions, coupling the glycosyl donor and acceptor became challenging when both were modified by sulfate diesters in previous early stage sulfation strategies.^{13, 18} Our initial attempt to couple the bromide donor and the 1,6-anhydrosyl acceptor **2-20** generated the disaccharide product **2-35** in 51% yield (Figure 2-16C). The yield of the disaccharide was improved to 79% when thiol donor **2-7** was employed (Figure 2-16C). However, it was found that **2-7** was partially decomposed after three days at 4 °C, likely attributed to the intramolecular nucleophilic attack at the O-6 sulfate diester by the anomeric thiol group. This challenge was overcome by oxidizing **2-7** into the sulfoxide donor **2-23**, which can be stored in long term without any signs of decomposition. Glycosylation efficiency was further improved when the sulfoxide donor **2-23** was employed, affording **2-36** in 95% yield even when both the donor and the acceptor carried sulfate diester substitutions (Figure 2-16C).

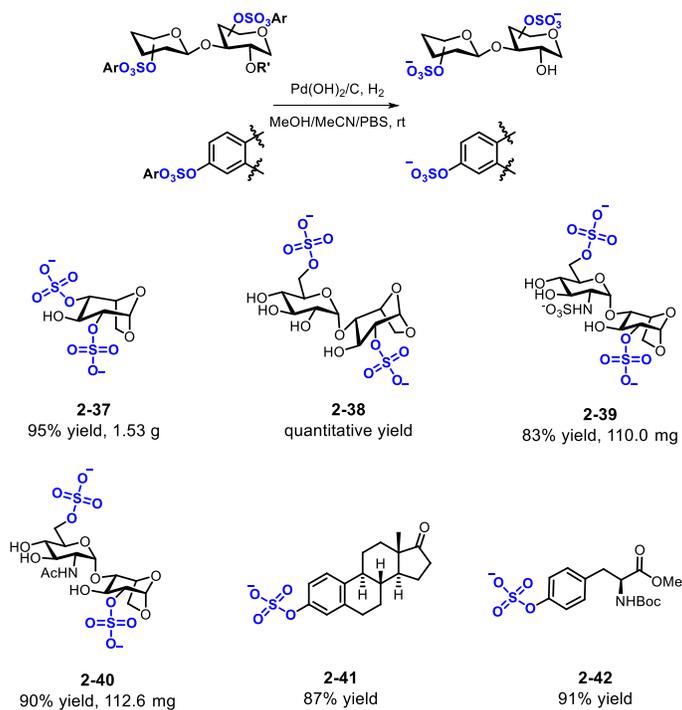


Figure 2-17. Decaging via hydrogenolysis

Because hydrogenolysis was milder and less prone to side reactions than hydrolysis, it became the method of choice for sulfate diester deprotection (Figure 2-17). Indeed, global deprotection of multigram amounts of **2-14b** by Pd(OH)₂/C/H₂ successfully yielded *O*-2, *O*-4 disulfated idose **2-37** in 95% yield. It is of note that the hydrogenolysis reaction was performed in a methoanolic PBS aqueous solution, the inorganic salts cannot be completely separated from **2-37** in spite of extensive purification. ³¹P-NMR was employed to quantify the amount of **2-37** in the hydrogenolysis product after purification which contained ¹H-NMR-inactive inorganic salts (*e.g.*, inorganic phosphates, chlorides, etc.). β-Glycerophosphate was used as an internal standard for quantification. 43.0 mg of compound **2-37** (sodium salt, MW = 366.2185 g/mol) and 17.5 mg of β-glycerophosphate tetrahydrate (MW = 288.0963 g/mol) were mixed together in NMR tube for proton and phosphorus NMR characterization (Figure 2-18). The amount of **2-37** in the purified mixture was determined using ¹H-NMR as follows.

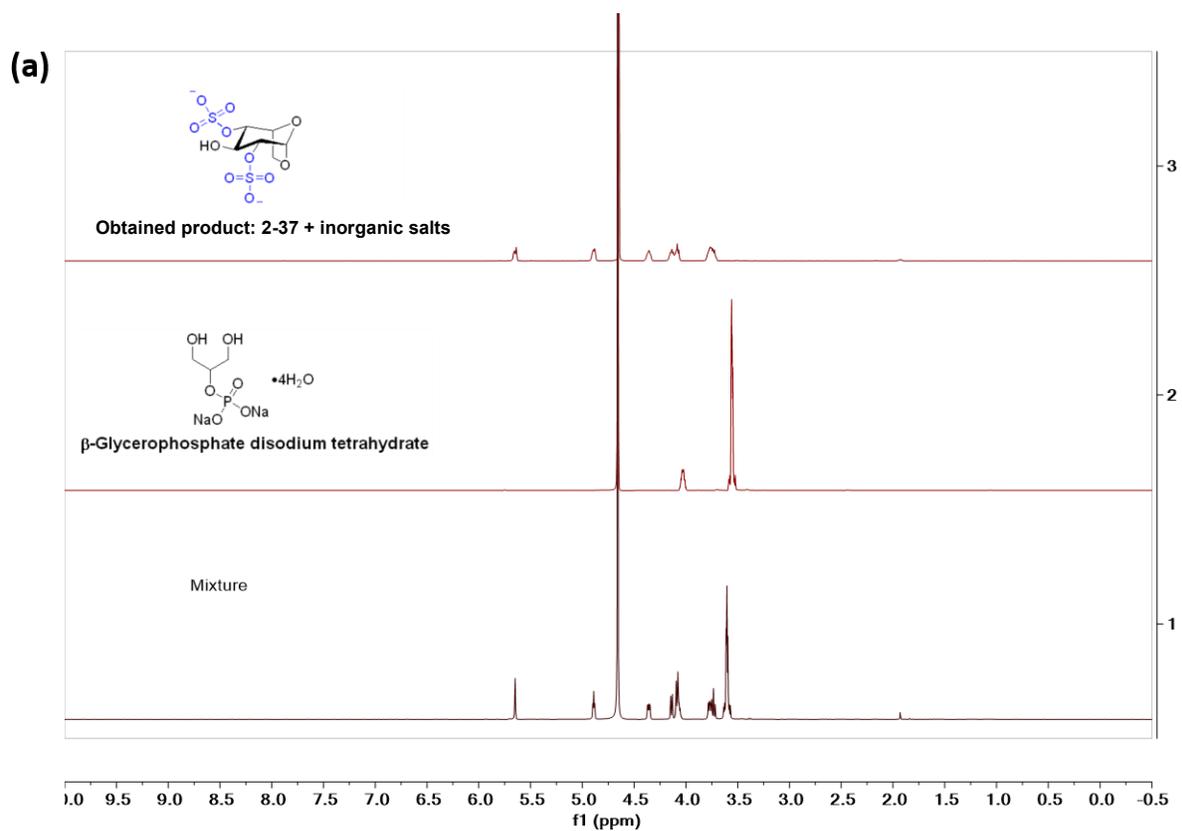
$$\text{Molar ratio of compound } \mathbf{2-37}/\beta\text{-glycerophosphate } \delta = \frac{1.00}{5.58/4} = 1.00/1.395;$$

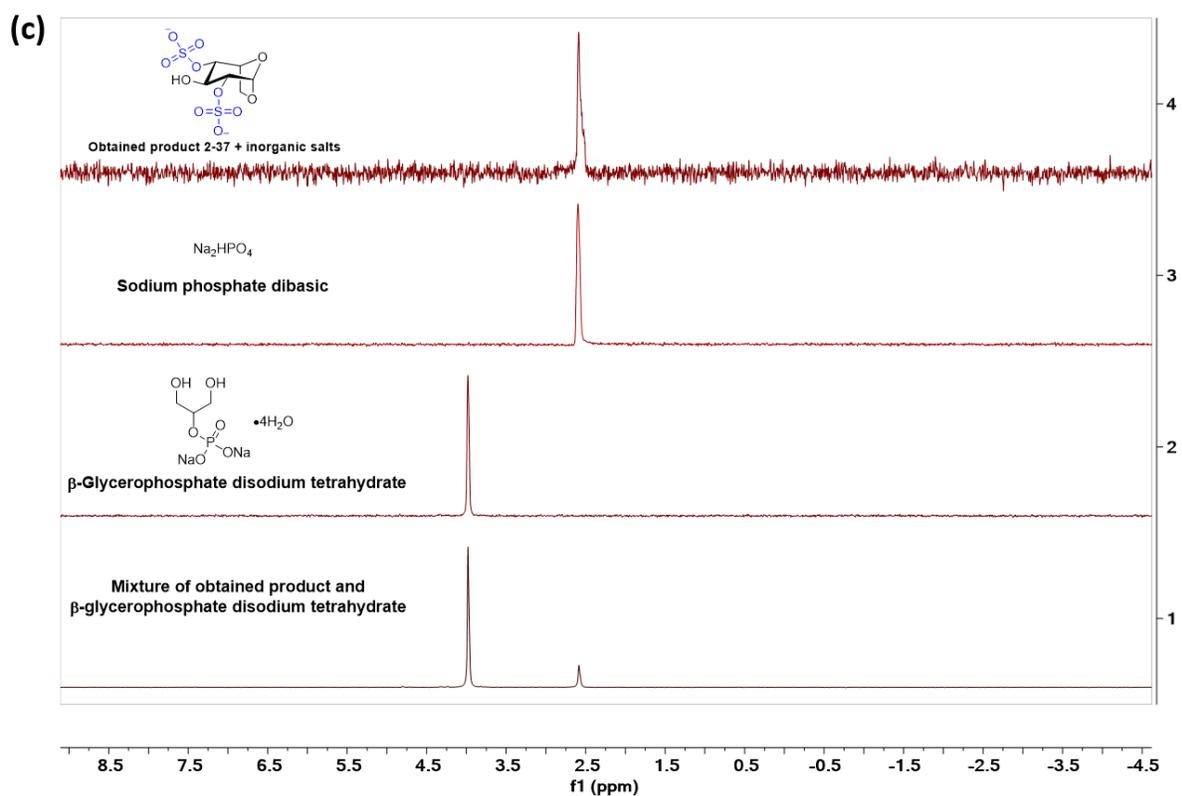
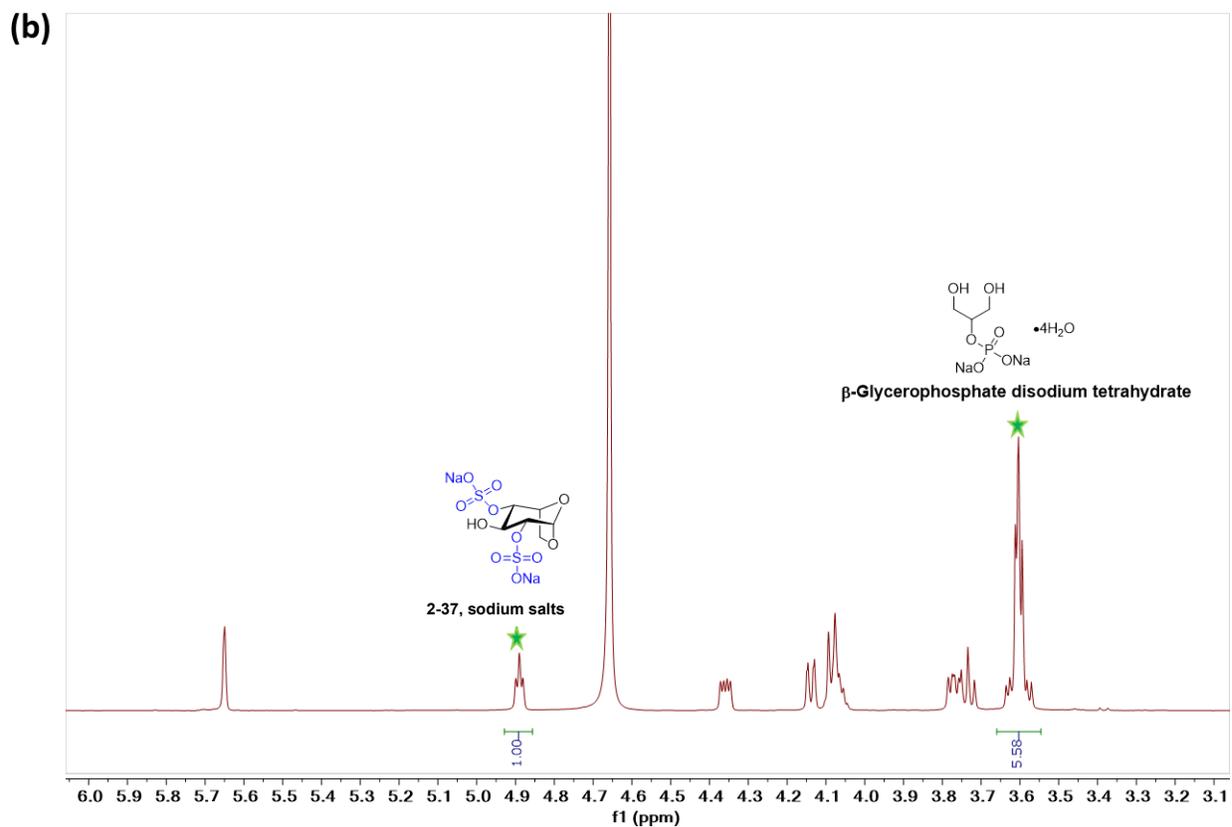
$$\text{Molar amount of } \beta\text{-glycerophosphate } n_1 = \frac{17.5 \text{ mg}}{288.0963 \text{ g/mol}} = 0.0607 \text{ mmol};$$

$$\text{Molar amount of compound } \mathbf{2-37} \ n_2 = \frac{0.0607 \text{ mmol}}{1.395} = 0.0435 \text{ mmol}.$$

Therefore, the mass of **2-37** in the mixture was determined to be $m_1 = 0.0435 \text{ mmol} \times 366.2185 \text{ g/mol} = 15.9 \text{ mg}$. Finally, the weight fraction of **2-37** in the mixture was calculated to be $\eta_1 = \frac{15.9 \text{ mg}}{43.0 \text{ mg}} \times 100\% = 37.0\%$. As the total mass of the product obtained after purification was 4.13 g, the mass of product was determined to be $m_2 = 4.13 \text{ g} \times 37\% = 1.53 \text{ g}$, giving the yield of the reaction $\eta_2 = 95\%$.

We also determined the amount of inorganic phosphate salts in this mixture by ^{31}P -NMR. Using the internal standard β -glycerophosphate, the weight fraction of inorganic phosphate anion in the mixture after purification is $\eta_3 = 2.7\%$. Clearly, the mixture also contains other ^1H -NMR- and ^{31}P -NMR-inactive inorganic salts such as chloride salts.





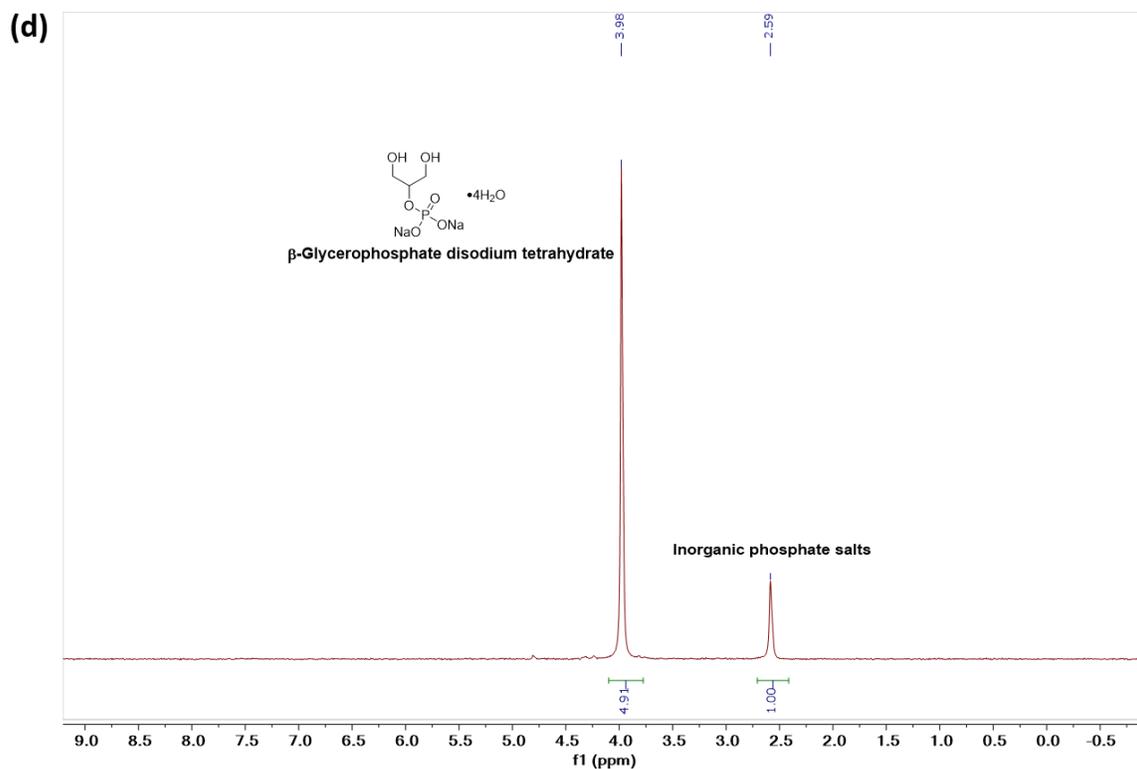


Figure 2-18. NMR quantification of disulfated idose **2-37** with β -glycerophosphate disodium tetrahydrate. **a)** Comparison of proton spectra. From top to bottom: ^1H -NMR spectra of disulfated compound **2-37** plus inorganic salts, β -glycerophosphate disodium tetrahydrate, and the mixture of the above two, respectively. **b)** ^1H -NMR signals for the determination of the molar ratio between **2-37** and reference (label with star), where signal at 4.9 ppm corresponds to the proton of CH from compound **2-37** (containing sodium salt) and signal at 3.6 ppm corresponds to the proton of CH_2 from glycerophosphate. **c)** Comparison of ^{31}P -NMR spectra. From top to bottom: ^{31}P -NMR spectra of the purified compound, sodium phosphate dibasic, β -glycerophosphate disodium tetrahydrate, and the mixture of the purified compound and β -glycerophosphate disodium tetrahydrate, respectively. The ^{31}P -NMR signal of the product aligns with the sodium phosphate dibasic standard, indicating that it contains inorganic phosphates. **d)** ^{31}P -NMR signals for the determination of the ratio of phosphate salts in the mixture.

Furthermore, disulfated **2-38**, trisulfated **2-39**, and disulfated **2-40** were also efficiently prepared in quantitative, 83%, and 90% yields, respectively. The deprotection of non-carbohydrate

sulfate diesters was also efficient. Hydrogenolysis of **2-16** and **2-17** exclusively furnished sulfoestrone **2-41** and sulfotyrosine **2-42** in high yields, without any desulfated byproducts being observed. The excellent selectivity of hydrogenolysis is attributed to the strong electron-withdrawing *p*- or *m*-NO₂ substitution that weakens the nearby aryl C–O bond and makes it susceptible to oxidative addition by palladium.

2.3 CONCLUSION

In summary, a facile and scalable approach for early stage *O*-sulfation via SuFEx reaction was developed for both carbohydrates and non-carbohydrate compounds. The SuFEx coupling reactions were optimized to efficiently generate sulfate diesters. The sulfate diesters demonstrated excellent compatibility with a broad range of reaction conditions and can be efficiently deprotected to yield *O*-sulfated products. This strategy provides a powerful tool for the synthesis of complex *O*-sulfated small molecules and macromolecules including carbohydrates, glycomimetic polymers, and peptides.

2.4 EXPERIMENTAL

2.4.1 Materials

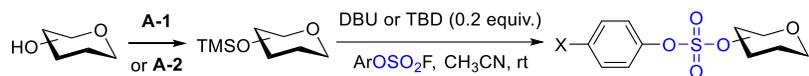
Chemicals were purchased from Alfa Aesar, Sigma-Aldrich, Acros, Fisher Scientific, or TCI chemical companies and used as received. Deuterated solvents were purchased from either Cambridge Isotope Laboratories, Inc. or Acros. Sulfuryl fluoride (Vikane) was purchased from SynQuest Lab. Inc. The organic solvents such as acetonitrile (MeCN), tetrahydrofuran (THF), dichloromethane (DCM), and dimethylformamide (DMF) were purchased from Fisher Scientific and used after the purification by a dry solvent system (Pure Process Technology). Thin layer chromatography was performed on Merck TLC plates (silica gel 60 F254) and visualized by UV irradiation (254 nm) and by charring with sulfuric acid in ethanol. Silica gel chromatography was carried out using an automated flash chromatography (Biotage). The reaction Schlenk bottles were flame dried.

2.4.2 Characterization

$^1\text{H-NMR}$, $^{13}\text{C-NMR}$, $^{19}\text{F-NMR}$, $^{31}\text{P-NMR}$, DOSY, NOESY, and HSQC measurements were conducted in CDCl_3 , D_2O , CD_3COCD_3 , or CD_3OD using a Varian Gemini-600 (600 MHz) or Varian Inova-500 (500 MHz) NMR spectrometer. Chemical shifts are in ppm calibrated using the resonances of the carbon and the residual proton of the deuterated solvent. GC measurements were carried out on a Shimadzu GCMS system (QP2010S) equipped with Chiral GTA and Dex-CB column by using helium as a carrier gas. High-resolution mass spectrometry was performed on JEOL AccuTOF DART Micromass LCT ESI-MS and an Agilent 6220 Time-of-Flight LC/MS instruments.

2.4.3 General Procedures

General Procedure A. Stepwise installation of sulfate diesters via SuFEx



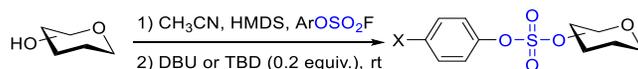
(A-1) Silylation by TMSCl: In a flask, the hydroxyl-exposed substrate (1 mmol of OH group, 1 eq) was dissolved in pyridine (10 mL). After cooling down to $0\text{ }^\circ\text{C}$, chloro(trimethyl)silane (TMSCl, 1.2 mmol, 1.2 eq) was added under the protection of N_2 . The mixture was warmed to room temperature and stirred for 2~24 h. The solution was diluted with 20 mL of ethyl acetate and washed by water for three times. The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The residue was purified by column chromatography to give the TMS-protected product.

(A-2) Silylation by HMDS: A stirred solution of the hydroxyl-exposed substrate (2.19 mmol) in dry acetonitrile (7.30 mL) was added hexamethyldisilazane (HMDS, 2.19 mmol, 1.0 equivalent per free hydroxyl) at room temperature under the protection of nitrogen. The resulting mixture was allowed to stir for 3 h until the TLC showed a full conversion of the starting materials. The whole reaction

mixture was diluted with DCM and concentrated for automated flash chromatography. The product was used directly in the next SuFEx coupling reaction without further characterization.

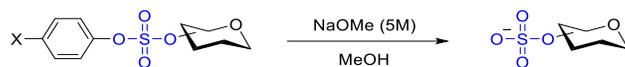
(A-3) SuFEx reaction: The TMS-protected substrate (0.3 mmol of O-TMS, 1 eq) and aryl fluorosulfate (0.33 mmol, 1.1 eq) were dissolved in anhydrous acetonitrile (1 mL). Catalyst (DBU or TBD, 0.06 mmol, 0.2 eq) was added under a positive nitrogen flow. The flask was sealed and stirred at room temperature for 2~24 h depending on the substrate. The solvent was removed immediately, and the residue was purified by column chromatography to give the sulfate diester-protected product. When TBS ether substrate was involved, the temperature of the reaction was elevated to 80 °C while the procedure of purification was similar with that mentioned above.

Procedure B. One-pot installation of sulfate diesters via SuFEx



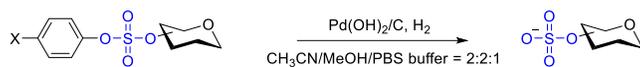
A stirred solution of the hydroxyl-exposed substrate (6.34 mmol) and aryl fluorosulfate (7.02 mmol, 1.1 equivalent per hydroxyl) in dry acetonitrile (21.13 mL) was added hexamethyldisilazane (HMDS, 3.80 mmol, 0.6 equivalent per free hydroxyl) at room temperature under nitrogen atmosphere. 10 min later, DBU or TBD (0.63 mmol, 0.1 equivalent per TMS ether unless otherwise noted) was added at the same temperature. When TLC showed no TMS ether intermediate left in the reaction mixture, the reaction was diluted with acetone and concentrated directly under reduced pressure for purification. When TBS ether substrate was involved, the reaction mixture was diluted with EA and successively washed by phosphate buffer, brine, dried over anhydrous Na₂SO₄. After filtration, the filtrate was concentrated under reduced pressure and purified by automated flash chromatography. Note: Workup by aqueous solution can help to avoid the removal of TBS group by the residue of strong organic base.

Procedure C. Deprotection of sulfate diester via hydrolysis



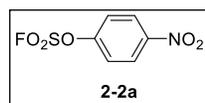
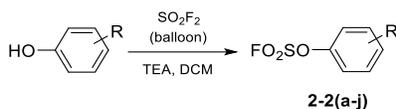
Sulfate diester masked reactant (125 μmol , 1 eq) was dissolved in 0.2 mL of acetonitrile in a flask. NaOMe solution was prepared by dissolving NaOMe powder (12.5 mmol, 100 eq) in 2.3 mL of methanol. Upon vigorously stirring, NaOMe solution was added in the mixture. After 1~2 h, the solution was neutralized by adding DOWEX 50WX8 resin until pH = 8.0. The resin was removed by filtration. The filtrate was concentrated and purified by column chromatography to give the sulfated product.

Procedure D. Deprotection of sulfate diester via hydrogenolysis

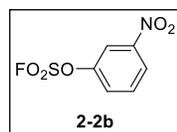


Sulfate diester (200 μmol , 1 eq) and 20% Pd(OH)₂ on carbon (1 mmol, 5 eq) was dispersed in the mixture of MeCN/MeOH/PBS buffer (v/v/v, 2/2/1, 10 mL, pH=7.4). The solution was equipped with a hydrogen balloon and stirred at room temperature for 2 h. The mixture was filtered through celite and the filtrate was concentrated under reduced pressure. The residue could be sequentially purified by column chromatography and passed through a column of Amberlite™ IR-120 Na ion-exchange resin to give the sulfated product. Size exclusion chromatography (LH-20) using methanol or MeOH/H₂O = 1:1 as eluent were necessary for *O*-sulfated carbohydrate compounds. The collected fraction was concentrated and the product was lyophilized over 20 h. Acetonitrile is a good solvent for the reactant to improve its solubility in polar environment. It can be replaced by other good solvents such as THF, DMF etc.

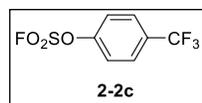
2.4.4 Detailed Experimental Procedure



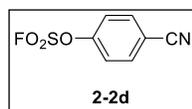
A 25 mL of single-neck round-bottom flask was charged with 4-nitrophenol (2.00 g, 14.38 mmol, 1 eq), dichloromethane (40.0 mL), and triethylamine (4.36 g, 6.01 mL, 43.13 mmol, 3.0 eq) and was then sealed with a septum. The atmosphere above the solution was removed with gentle vacuum, and SO_2F_2 gas (sulfuryl fluoride, Vikane) was introduced by a needle from a balloon filled with the gas. The reaction mixture was vigorously stirred at room temperature overnight. The solvent was evaporated, and the residual was purified by column chromatography using hexane/EA (10/1, v/v) as eluent to afford compound **2a** (3.01 g, 95%) as a light yellow oil. ^1H NMR (500 MHz, CDCl_3) δ 8.43 – 8.37 (m, 2H), 7.60 – 7.52 (m, 2H); ^{19}F NMR (470 MHz, Chloroform-*d*) δ 39.49. Spectral data matched those previously reported.⁵¹ Other arylfluorosulfate compounds **2-2(b-j)** were prepared by following the similar procedure.



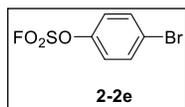
Compound **2-2b** was isolated as a colorless oil in 99% yield. ^1H NMR (400 MHz, CDCl_3) δ 8.38 – 8.30 (m, 1H), 8.26–8.25 (m, 1H), 7.79 – 7.68 (m, 2H); ^{19}F NMR (376 MHz, CDCl_3) δ 39.03 (d, $J = 4.6$ Hz); ^{13}C NMR (126 MHz, CDCl_3) δ 149.79, 149.27, 131.54, 127.30, 123.82, 117.13; HRMS (DART) m/z $[\text{M}+\text{H}]^+$ Calcd. For $\text{C}_6\text{H}_5\text{NO}_5\text{SF}$: 221.9867, found 221.9860. Spectral data matched those previously reported.⁵²



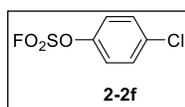
Compound **2-2c** was isolated as a colorless oil in 85% yield. ^1H NMR (500 MHz, CDCl_3) δ 7.78 (d, $J = 8.4$ Hz, 2H), 7.49 (d, $J = 8.4$ Hz, 2H). ^{19}F NMR (470 MHz, CDCl_3) δ /ppm: 38.62, -62.74. Spectral data matched those previously reported.⁵³



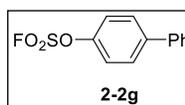
Compound **2-2d** was isolated as a white crystal in 96% yield. ^1H NMR (500 MHz, CDCl_3) δ 7.83 (d, $J = 8.3$ Hz, 2H), 7.50 (d, $J = 8.4$ Hz, 2H). ^{19}F NMR (470 MHz, CDCl_3) δ /ppm: 39.39. Spectral data matched those previously reported.⁵³



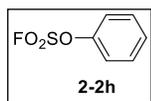
Compound **2-2e** was isolated as a colorless oil in 96% yield. ^1H NMR (500 MHz, CDCl_3) δ 7.61 (d, $J = 7.8$ Hz, 2H), 7.24 (d, $J = 8.5$ Hz, 2H). ^{19}F NMR (470 MHz, CDCl_3) δ 37.78. Spectral data matched those previously reported.⁵⁴



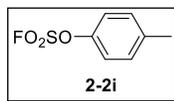
Compound **2-2f** was isolated as a colorless oil in 91% yield. ^1H NMR (500 MHz, CDCl_3) δ 7.46 (d, $J = 8.4$ Hz, 2H), 7.30 (d, $J = 8.5$ Hz, 2H). ^{19}F NMR (564 MHz, CDCl_3) δ 37.66. Spectral data matched those previously reported.⁵³



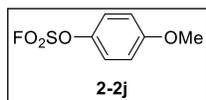
Compound **2-2g** was isolated as a white crystal in 99% yield. ^1H NMR (600 MHz, CDCl_3) δ 7.69 – 7.64 (m, 2H), 7.58 – 7.53 (m, 2H), 7.50 – 7.44 (m, 2H), 7.44 – 7.37 (m, 3H). ^{19}F NMR (564 MHz, CDCl_3) δ 37.61. Spectral data matched those previously reported.⁵³



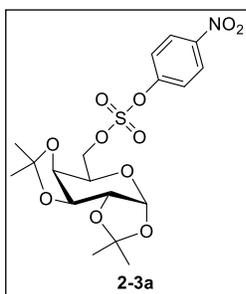
Compound **2-2h** was isolated as a colorless oil in 82% yield. ^1H NMR (500 MHz, CDCl_3) δ 7.52 – 7.46 (m, 2H), 7.45 – 7.40 (m, 1H), 7.38 – 7.32 (m, 2H). ^{19}F NMR (470 MHz, CDCl_3) δ 37.54. Spectral data matched those previously reported.⁵³



Compound **2-2i** was isolated as a colorless oil in 60% yield. ^1H NMR (500 MHz, CDCl_3) δ 7.26 (d, $J = 8.5$ Hz, 1H), 7.22 (d, $J = 8.4$ Hz, 1H), 2.39 (s, 1H). ^{19}F NMR (470 MHz, CDCl_3) δ 37.03. Spectral data matched those previously reported.⁵³

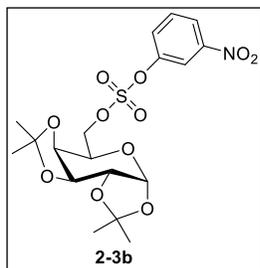


Compound **2-2j** was isolated as a colorless oil in 98% yield. ^1H NMR (500 MHz, CDCl_3) δ 7.23 (d, $J = 9.2$ Hz, 2H), 6.92 (d, $J = 9.3$ Hz, 2H), 3.79 (s, 3H). ^{19}F NMR (470 MHz, CDCl_3) δ 36.26. Spectral data matched those previously reported.⁵³



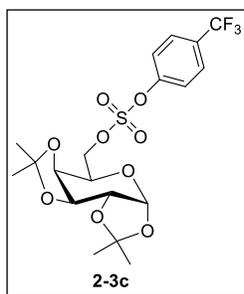
In a 10 mL flask, **2-1** (102.5 mg, 0.31 mmol, 1 eq) and **2-2a** (75 mg, 0.34 mmol) were dissolved in anhydrous acetonitrile (1 mL). DBU (9.4 mg, 9.2 μ L, 0.06 mmol) was added under a positive nitrogen flow. The flask was sealed and stirred at RT for 2 h. The solvent was removed immediately, and the crude product was purified by column chromatography using hexane/ethyl acetate

(7:1) as eluent. The product **2-3a** (134 mg) was obtained as a white powder in 94% yield. $[\alpha]_D^{20} = -32.53$ (*c* 1.00, CHCl₃). ¹H NMR (600 MHz, CDCl₃) δ 8.30 (m, 2H), 7.59 (m, 2H), 5.53 (d, *J* = 4.9 Hz, 1H), 4.66 (dd, *J* = 7.8, 2.6 Hz, 1H), 4.61 (dd, *J* = 10.9, 4.0 Hz, 1H), 4.57 (dd, *J* = 10.9, 7.7 Hz, 1H), 4.36 (dd, *J* = 5.0, 2.6 Hz, 1H), 4.25 (dd, *J* = 7.8, 2.0 Hz, 1H), 4.18 (ddd, *J* = 7.7, 4.0, 2.0 Hz, 1H), 1.49 (s, 3H), 1.45 (s, 3H), 1.35 – 1.32 (m, 6H). ¹³C NMR (151 MHz, CDCl₃) δ 154.32, 146.38, 125.59, 122.44, 110.03, 109.13, 96.14, 73.62, 70.63, 70.58, 70.19, 65.79, 25.93, 25.90, 24.84, 24.37. HRMS (DART) *m/z* [M+NH₄]⁺ Calcd. For C₁₈H₂₇N₂O₁₁S: 479.1336, found 479.1333.



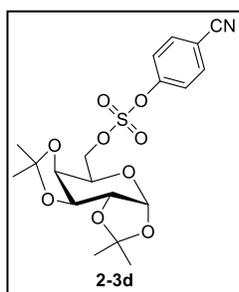
Compound **2-1** (101.5 mg, 0.3 mmol) was reacted with **2-2b** (74 mg, 0.33 mmol) following **General Procedure A-3**. The product was further purified by silica gel chromatography (DCM~hexane/ethyl acetate 3:1) to obtain **2-3b** as a colorless liquid (129 mg, 92%). $[\alpha]_D^{20} = -36.26$ (*c* 1.00, CHCl₃). ¹H NMR (600

MHz, CDCl₃) δ 8.30 (t, *J* = 2.3 Hz, 1H), 8.21 (ddd, *J* = 8.3, 2.1, 1.0 Hz, 1H), 7.76 (ddd, *J* = 8.3, 2.4, 1.0 Hz, 1H), 7.62 (t, *J* = 8.2 Hz, 1H), 5.56 (d, *J* = 5.0 Hz, 1H), 4.66 (dd, *J* = 7.8, 2.6 Hz, 1H), 4.63 – 4.57 (m, 2H), 4.36 (dd, *J* = 5.0, 2.6 Hz, 1H), 4.26 (dd, *J* = 7.8, 2.0 Hz, 1H), 4.19 (ddd, *J* = 7.0, 4.5, 2.0 Hz, 1H), 1.48 (s, 3H), 1.46 (s, 3H), 1.34 (s, 3H), 1.32 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 150.27, 148.96, 130.53, 128.00, 122.29, 117.54, 110.07, 109.07, 96.20, 73.50, 70.69, 70.61, 70.16, 65.72, 25.90, 25.86, 24.81, 24.39. HRMS (DART) *m/z* [M+NH₄]⁺ Calcd. For C₁₈H₂₇N₂O₁₁S: 479.1336, found 479.1318.



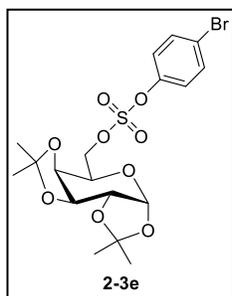
Compounds **2-1** (101.5 mg, 0.31 mmol) and **2-2c** (82 mg, 0.34 mmol) were reacted following **General Procedure A-3**. The product was further purified by silica gel chromatography (hexane/ethyl acetate 8:1) to obtain **2-3c** as a white powder (134 mg, 91%). $[\alpha]_D^{20} = -35.39$ (*c* 1.00 in CHCl₃). ¹H NMR (500 MHz, CDCl₃) δ 7.69 (d, *J* = 8.4 Hz, 2H), 7.52 (d, *J* = 8.4 Hz, 2H), 5.54 (d, *J* = 4.9 Hz,

1H), 4.65 (dd, *J* = 7.9, 2.6 Hz, 1H), 4.59 (dd, *J* = 10.9, 4.2 Hz, 1H), 4.54 (dd, *J* = 10.8, 7.4 Hz, 1H), 4.36 (dd, *J* = 5.0, 2.6 Hz, 1H), 4.25 (dd, *J* = 7.8, 2.1 Hz, 1H), 4.17 (ddd, *J* = 7.2, 4.3, 2.1 Hz, 1H), 1.49 (s, 3H), 1.45 (s, 3H), 1.34 (s, 6H). ¹³C NMR (151 MHz, CDCl₃) δ 152.49, 127.31, 127.29, 127.26, 121.98, 109.97, 109.10, 96.15, 73.09, 70.62, 70.57, 70.23, 65.76, 25.90, 25.88, 24.84, 24.35. ¹⁹F NMR (470 MHz, CDCl₃) δ -62.47. HRMS (DART) *m/z* [M+H]⁺ Calcd. For C₁₉H₂₄F₃O₉S: 485.1093, found 485.1082.

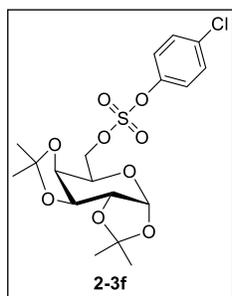


Compounds **2-1** (99.5 mg, 0.3 mmol) and **2-2d** (66.2 mg, 0.33 mmol) were reacted following **General Procedure A-3**. The product was further purified by silica gel chromatography (hexane/ethyl acetate 8:1) to obtain **2-3d** as a white powder (125 mg, 95%). $[\alpha]_D^{20} = -28.33$ (*c* 1.00 in CHCl₃). ¹H NMR (600 MHz, CDCl₃) δ 7.73 (m, 2H), 7.53 (m, 2H), 5.53 (d, *J* = 5.0 Hz, 1H), 4.65 (dd, *J* = 7.9, 2.5 Hz, 1H), 4.60 (dd,

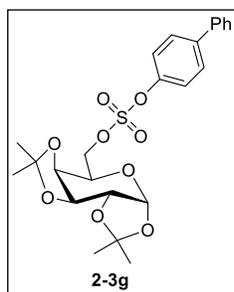
J = 10.9, 4.1 Hz, 1H), 4.55 (dd, *J* = 10.8, 7.6 Hz, 1H), 4.36 (dd, *J* = 4.9, 2.6 Hz, 1H), 4.24 (dd, *J* = 7.8, 2.0 Hz, 1H), 4.17 (ddd, *J* = 7.7, 4.1, 2.0 Hz, 1H), 1.49 (s, 3H), 1.45 (s, 3H), 1.34 (s, 6H). ¹³C NMR (151 MHz, CDCl₃) δ 153.06, 134.09, 122.61, 117.67, 111.51, 110.01, 109.13, 96.13, 73.47, 70.62, 70.58, 70.20, 65.78, 25.93, 25.90, 24.85, 24.37. HRMS (DART) *m/z* [M+H]⁺ Calcd. For C₁₉H₂₄NO₉S: 442.1172, found 442.1167.



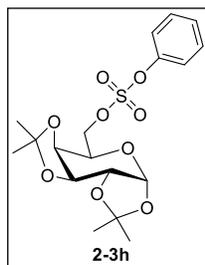
Compounds **2-1** (101.7 mg, 0.3 mmol) and **2-2e** (85.8 mg, 0.33 mmol) were reacted following **General Procedure A-3**. The product was further purified by silica gel chromatography (hexane/ethyl acetate 6:1) to obtain **2-3e** as a white powder (140.8 mg, 93%). $[\alpha]_D^{20} = -37.66$ (*c* 1.00 in CHCl_3). $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.52 (d, *J* = 8.9 Hz, 2H), 7.28 (d, *J* = 9.0 Hz, 2H), 5.54 (d, *J* = 4.9 Hz, 1H), 4.65 (dd, *J* = 7.9, 2.6 Hz, 1H), 4.57 (dd, *J* = 10.7, 4.5 Hz, 1H), 4.51 (dd, *J* = 10.7, 7.4 Hz, 1H), 4.35 (dd, *J* = 4.9, 2.6 Hz, 1H), 4.25 (dd, *J* = 7.8, 2.0 Hz, 1H), 4.17 (ddd, *J* = 6.9, 4.5, 2.0 Hz, 1H), 1.50 (s, 3H), 1.45 (s, 3H), 1.34 (s, 3H), 1.33 (s, 3H). $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ 149.29, 132.94, 123.32, 120.84, 109.93, 109.09, 96.16, 72.82, 70.64, 70.57, 70.27, 65.76, 25.94, 25.91, 24.88, 24.39. HRMS (DART) m/z $[M+H]^+$ Calcd. For $\text{C}_{18}\text{H}_{24}\text{BrO}_9\text{S}$: 495.0324, found 495.0308.



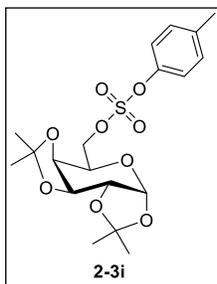
Compounds **2-1** (100.2 mg, 0.3 mmol) and **2-2f** (69.8 mg, 0.33 mmol) were reacted following **General Procedure A-3**. The product was further purified by silica gel chromatography (hexane/ethyl acetate 6:1) to obtain **2-3f** as a white powder (117.5 mg, 86%). $[\alpha]_D^{20} = -36.86$ (*c* 1.00 in CHCl_3). $^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.37 (d, *J* = 9.0 Hz, 1H), 7.34 (d, *J* = 8.7 Hz, 1H), 5.54 (d, *J* = 4.9 Hz, 1H), 4.65 (dd, *J* = 7.9, 2.5 Hz, 1H), 4.57 (dd, *J* = 10.7, 4.5 Hz, 1H), 4.51 (dd, *J* = 10.8, 7.4 Hz, 1H), 4.36 (dd, *J* = 5.0, 2.6 Hz, 1H), 4.25 (dd, *J* = 7.8, 2.0 Hz, 1H), 4.17 (ddd, *J* = 7.0, 4.4, 1.9 Hz, 1H), 1.50 (s, 2H), 1.45 (s, 2H), 1.34 (s, 2H), 1.34 (s, 2H). $^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 148.69, 133.08, 129.93, 122.96, 109.93, 109.09, 96.15, 72.79, 70.62, 70.56, 70.25, 65.74, 25.93, 25.90, 24.87, 24.38. HRMS (DART) m/z $[M+H]^+$ Calcd. For $\text{C}_{18}\text{H}_{24}\text{ClO}_9\text{S}$: 451.0830, found 451.0826.



Compounds **2-1** (101.3 mg, 0.3 mmol) and **2-2g** (84.6 mg, 0.34 mmol) were reacted following **General Procedure A-3**. The product was further purified by silica gel chromatography (hexane/ethyl acetate 8:1) to obtain **2-3g** as a white powder (113.9 mg, 76%). $[\alpha]_D^{20} = -35.13$ (c 1.00 in CHCl₃). ¹H NMR (600 MHz, CDCl₃) δ 7.59 (m, 2H), 7.54 (m, 2H), 7.47 – 7.42 (m, 4H), 7.37 (m, 1H), 5.56 (d, $J = 4.9$ Hz, 1H), 4.65 (dd, $J = 7.9, 2.5$ Hz, 1H), 4.60 (dd, $J = 10.6, 4.8$ Hz, 1H), 4.54 (dd, $J = 10.7, 7.3$ Hz, 1H), 4.35 (dd, $J = 5.0, 2.5$ Hz, 1H), 4.27 (dd, $J = 7.9, 2.0$ Hz, 1H), 4.20 (ddd, $J = 7.0, 4.8, 2.0$ Hz, 1H), 1.51 (s, 3H), 1.46 (s, 3H), 1.34 (s, 3H), 1.33 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 149.69, 140.63, 139.82, 128.83, 128.55, 127.72, 127.16, 121.70, 109.91, 109.09, 96.18, 72.49, 70.63, 70.57, 70.31, 65.77, 25.97, 25.93, 24.90, 24.41. HRMS (DART) m/z [M+H]⁺ Calcd. For C₂₄H₂₉O₉S: 493.1532, found 493.1525.

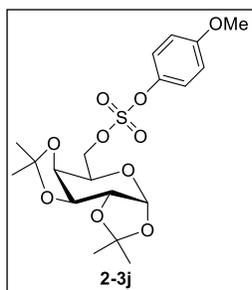


Compounds **2-1** (102.4 mg, 0.3 mmol) and **2-2h** (59.7 mg, 0.33 mmol) were reacted following **General Procedure A-3**. The product was further purified by silica gel chromatography (hexane/ethyl acetate 6:1) to obtain **2-3h** as a colorless liquid (55.5 mg, 43%). $[\alpha]_D^{20} = -51.99$ (c 1.00 in CHCl₃). ¹H NMR (600 MHz, CDCl₃) δ 7.43 – 7.39 (m, 2H), 7.39 – 7.36 (m, 2H), 7.35 – 7.29 (m, 1H), 5.55 (d, $J = 5.0$ Hz, 1H), 4.64 (dd, $J = 7.9, 2.5$ Hz, 1H), 4.57 (dd, $J = 10.6, 4.9$ Hz, 1H), 4.51 (dd, $J = 10.6, 7.2$ Hz, 1H), 4.35 (dd, $J = 5.0, 2.5$ Hz, 1H), 4.26 (dd, $J = 7.9, 2.0$ Hz, 1H), 4.18 (ddd, $J = 7.0, 4.9, 2.0$ Hz, 1H), 1.51 (s, 3H), 1.46 (s, 3H), 1.34 (s, 3H), 1.33 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 150.33, 129.86, 127.38, 121.41, 109.88, 109.07, 96.16, 72.33, 70.61, 70.53, 70.30, 65.72, 25.95, 25.91, 24.89, 24.38. HRMS (DART) m/z [M+NH₄]⁺ Calcd. For C₁₈H₂₄NO₉S: 434.1485, found 434.1478. Spectral data matched those previously reported.⁹



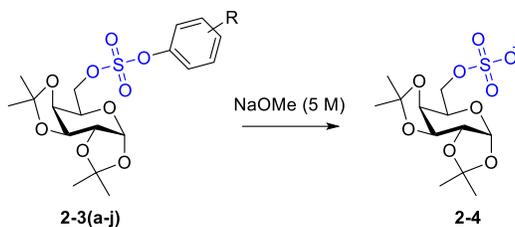
Compounds **2-1** (102.9 mg, 0.31 mmol) and **2-2i** (64.8 mg, 0.34 mmol) were reacted following **General Procedure A-3**. The product was further purified by silica gel chromatography (hexane/ethyl acetate 6:1) to obtain **2-3i** as a white solid (38.7 mg, 29%). $[\alpha]_D^{20} = -49.73$ (c 1.00 in CHCl_3). $^1\text{H NMR}$ (600 MHz, CDCl_3) δ

7.25 (d, $J = 8.5$ Hz, 2H), 7.18 (d, $J = 8.5$ Hz, 2H), 5.54 (d, $J = 4.9$ Hz, 1H), 4.64 (dd, $J = 7.9, 2.6$ Hz, 1H), 4.56 (dd, $J = 10.6, 4.9$ Hz, 1H), 4.49 (dd, $J = 10.6, 7.2$ Hz, 1H), 4.35 (dd, $J = 5.0, 2.5$ Hz, 1H), 4.26 (dd, $J = 7.9, 2.0$ Hz, 1H), 4.17 (ddd, $J = 7.1, 5.0, 2.0$ Hz, 1H), 2.35 (s, 3H), 1.51 (s, 3H), 1.46 (s, 3H), 1.34 (s, 3H), 1.33 (s, 3H). $^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 148.20, 137.30, 130.31, 121.13, 109.85, 109.05, 96.16, 72.20, 70.60, 70.53, 70.31, 65.71, 25.95, 25.91, 24.90, 24.38, 20.87. HRMS (DART) m/z $[\text{M}+\text{H}]^+$ Calcd. For $\text{C}_{19}\text{H}_{27}\text{O}_9\text{S}$: 431.1376, found 431.1369.

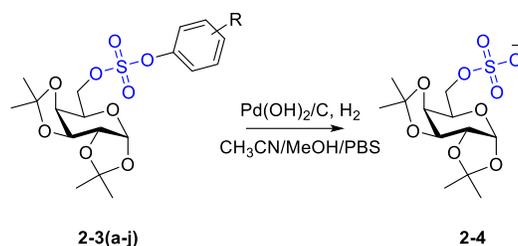


Compounds **2-1** (99.1 mg, 0.3 mmol) and **2-2j** (67.6 mg, 0.33 mmol) were reacted following **General Procedure A-3**. The product was further purified by silica gel chromatography (hexane/ethyl acetate 8:1) to obtain **2-3j** as a white solid (41.5 mg, 31%). $[\alpha]_D^{20} = -35.46$ (c 1.00 in CHCl_3). $^1\text{H NMR}$ (600 MHz,

CDCl_3) δ 7.29 (d, $J = 9.1$ Hz, 2H), 6.88 (d, $J = 9.2$ Hz, 2H), 5.55 (d, $J = 4.9$ Hz, 1H), 4.64 (dd, $J = 7.8, 2.6$ Hz, 1H), 4.55 (dd, $J = 10.6, 4.8$ Hz, 1H), 4.49 (dd, $J = 10.6, 7.3$ Hz, 1H), 4.35 (dd, $J = 5.0, 2.5$ Hz, 1H), 4.26 (dd, $J = 7.8, 2.0$ Hz, 1H), 4.17 (ddd, $J = 7.0, 4.8, 2.0$ Hz, 1H), 3.80 (s, 3H), 1.51 (s, 3H), 1.46 (s, 3H), 1.34 (s, 3H), 1.33 (s, 3H). $^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 158.44, 143.79, 122.54, 114.70, 109.85, 109.06, 96.16, 72.29, 70.61, 70.55, 70.30, 65.74, 55.62, 25.95, 25.91, 24.90, 24.39. HRMS (DART) m/z $[\text{M}+\text{H}]^+$ Calcd. For $\text{C}_{19}\text{H}_{27}\text{O}_{10}\text{S}$: 447.1325, found 447.1312.



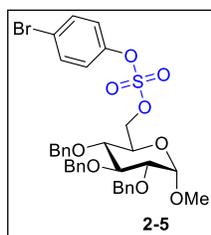
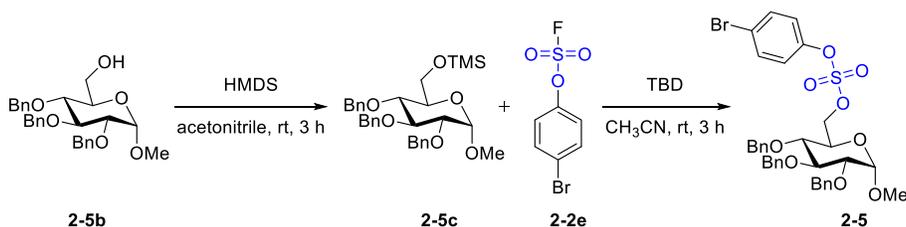
In a 10 mL vial, **2-3a** (57.7 mg, 125 μ mol) was dissolved in 0.2 mL of acetonitrile. NaOMe solution was prepared by dissolving NaOMe powder (675.2 mg, 12.5 mmol) in 2.3 mL of methanol. Upon vigorously stirring, NaOMe solution was added in the vial. After 1 h, the solution was neutralized by adding DOWEX 50WX8 resin until pH = 8.0. The resin was removed by filtration. The filtrate was concentrated and purified by column chromatography using DCM/methanol (5:1) as eluent. The product **2-4** (39.1 mg) was obtained as a white solid in 92% yield. To monitor the kinetics of reaction, methanol- d_4 was used as solvent instead of methanol. The aliquots were obtained, diluted with a fixed volume of MeOD and neutralized with DOWEX 50WX8 resin directly for ^1H NMR test. The progress of the reaction was monitored at 5 min, 10 min, 15 min, 20 min, 25 min, 30 min, 40 min, 60 min. The preparation of **2-4** from **2-3(b-j)** by hydrolysis was carried out by the same procedure to give different yields as shown in Table 1 in the main text. $[\alpha]_{\text{D}}^{20} = -40.39$ (c 1.00 in MeOH). ^1H NMR (500 MHz, CD_3OD) δ 5.47 (d, $J = 4.9$ Hz, 1H), 4.64 (dd, $J = 8.0, 2.4$ Hz, 1H), 4.35 (dd, $J = 5.0, 2.4$ Hz, 1H), 4.32 (dd, $J = 7.8, 1.6$ Hz, 1H), 4.20 – 4.10 (m, 2H), 4.06 (dd, $J = 9.9, 6.3$ Hz, 1H), 1.51 (s, 3H), 1.41 (s, 3H), 1.34 (s, 3H), 1.33 (s, 3H). ^{13}C NMR (151 MHz, CD_3OD) δ 109.06, 108.58, 96.26, 70.89, 70.59, 70.49, 66.41, 66.40, 24.91, 24.89, 23.79, 23.14. HRMS (ESI) m/z $[\text{M}-\text{H}]^-$ Calcd. For $\text{C}_{12}\text{H}_{19}\text{O}_9\text{S}$: 339.0755, found 339.0760.



In a 10 mL vial, **2-3a** (93.2 mg, 200 μmol) and 20% Pd(OH)₂ on carbon (535.2 mg) was dispersed in the mixture of MeCN/MeOH/PBS buffer (v/v/v=2/2/1, 10 mL, pH=7.4). The solution was equipped with a hydrogen balloon and stirred at room temperature for 2 h. The mixture was filtered through celite. The filtrate was concentrated and purified by column chromatography using DCM/methanol (5:1) as eluent, and further passed through a column of Amberlite™ IR-120 Na ion-exchange resin. The product **2-4** (69.9 mg) was obtained as a white solid in 96% yield. The preparation of **2-4** from **2-3(b-j)** by hydrogenolysis was carried out by the same procedure to give comparably high yields as shown in Table 1 in the main text. The spectral data is consistent with that by hydrolysis as shown in Figure. S8.

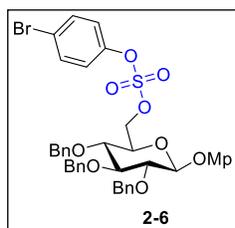
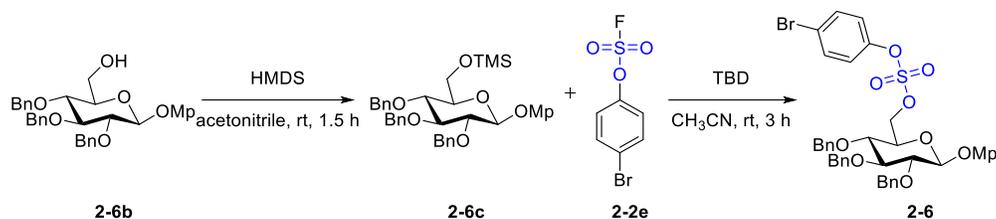
GC-MS detection for hydrogenolysis process: a) PtO₂/H₂ condition. To a stirred suspension of **2-3h** (33.5 mg, 0.08 mmol) in a mixed solvent (4.02 mL, ethanol/H₂O = 9:1) was added potassium carbonate (33.5 mg, 0.24 mmol) and platinum oxide (167.5 mg, 0.74 mmol) at room temperature. The reaction bottle was charged with a gentle vacuum, followed by filling hydrogen gas through a syringe attached balloon. The resulting reaction mixture was allowed to stir at room temperature. After 0.5 h, a small portion of the reaction mixture was collected and filtered for GC-MS analysis. At 2 h of reaction time, another portion of the reaction mixture was collected, filtered and used for GC-MS analysis. b) Pd(OH)₂/H₂ condition. The test was conducted by following the **General Procedure D**. After 2 h, a portion of the reaction mixture was collected for GC-MS analysis. c) Pd(OH)₂/NH₄OOCH condition. A stirred solution of **2-3h** (33.3 mg, 0.08 mmol) in the mixed solvent

(4.0 mL, MeOH/THF = 1:1) was added ammonium formate (30.3 mg, 0.48 mmol) and Pd(OH)₂ on activated carbon (20% Pd, wet, 166.5 mg, 5 g per gram of substrate) at room temperature. After 2 h, a small portion of the reaction mixture was collected and filtered for GC-MS analysis.



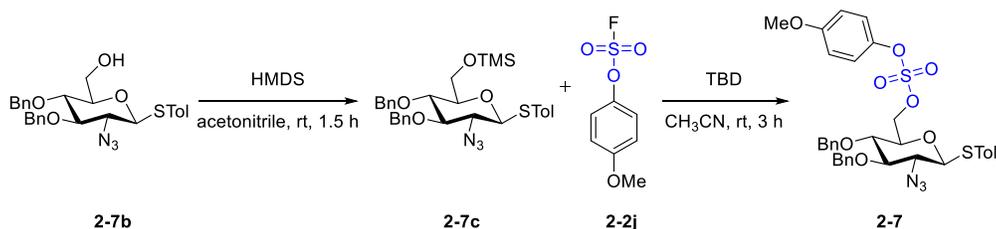
Compound **2-5b**⁵⁶ (1.02 g, 2.19 mmol) was converted to trimethylsilyl ether **2-5c** (1.13 g, 96%) by following **General Procedure A-2**. **2-5c** was used directly in the next step without further characterization.

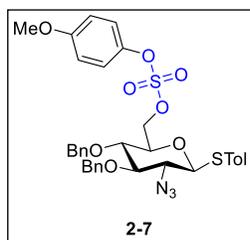
Compound **2-5c** (1.13 g, 2.10 mmol) coupled with arylfluorosulfate **2-2e** (588.0 mg, 2.31 mmol) in the presence of TBD (58.4 mg, 0.42 mmol) providing compound **2-5** (1.02 g, 70%) according to **General Procedure A-3**. $[\alpha]_D^{20} = +62.9$ (*c* 1.00, CHCl₃); ¹H NMR (500 MHz, CDCl₃) δ 7.47 – 7.42 (m, 2H), 7.40 – 7.26 (m, 13H), 7.26 – 7.22 (m, 2H), 7.18 – 7.11 (m, 2H), 5.00 (d, *J* = 10.9 Hz, 1H), 4.90 (d, *J* = 11.0 Hz, 1H), 4.84 – 4.76 (m, 2H), 4.65 (d, *J* = 12.1 Hz, 1H), 4.55 (dd, *J* = 7.3, 3.7 Hz, 2H), 4.50 (dd, *J* = 10.5, 2.1 Hz, 1H), 4.45 (dd, *J* = 10.5, 4.8 Hz, 1H), 4.01 (t, *J* = 9.2 Hz, 1H), 3.87 (ddd, *J* = 10.2, 4.8, 2.0 Hz, 1H), 3.53 – 3.47 (m, 1H), 3.44 (m, 1H), 3.35 (s, 3H); ¹³C NMR (126 MHz, CDCl₃) δ 149.28, 138.59, 138.04, 137.73, 133.08, 128.72, 128.70, 128.59, 128.29, 128.24, 128.20, 128.13, 128.04, 127.86, 123.16, 120.94, 98.30, 81.89, 79.83, 76.85, 75.89, 75.22, 73.65, 73.04, 68.49, 55.65; HRMS (ESI) *m/z* [M+Na]⁺ calcd for C₃₄H₃₅BrO₉SNa 721.1077, found 721.1095.



Compound **2-6b**⁵⁷⁻⁵⁸ (3.14 g, 5.64 mmol) was converted to trimethylsilyl ether **2-6c** (3.22 g, 91%) by following **General Procedure A-2**. **2-6c** was used directly in the next step without further characterization.

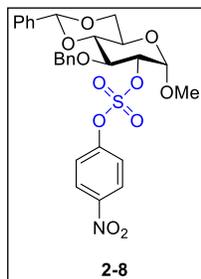
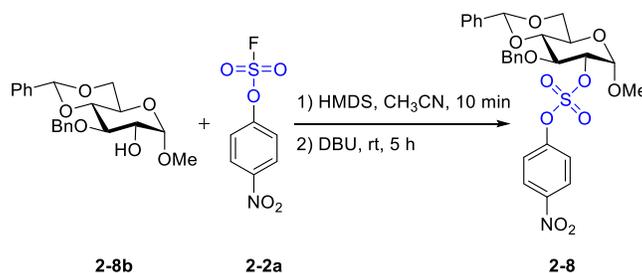
Compound **2-6c** (1.62 g, 2.58 mmol) coupled with aryl fluorosulfate **2-2e** (722.8 mg, 2.83 mmol) in the presence of TBD (71.7 mg, 0.52 mmol), providing compound **2-6** (1.58 g, 78%) according to **General Procedure A-3**. $[\alpha]_D^{20} = -16.3$ (c 1.01, CHCl_3); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.37 – 7.27 (m, 11H), 7.27 – 7.20 (m, 4H), 7.14 – 7.08 (m, 2H), 6.99 – 6.91 (m, 2H), 6.85 – 6.79 (m, 2H), 5.04 (d, $J = 10.9$ Hz, 1H), 4.98 (d, $J = 11.0$ Hz, 1H), 4.94 (d, $J = 7.6$ Hz, 1H), 4.90 (d, $J = 11.0$ Hz, 1H), 4.83 (d, $J = 3.6$ Hz, 1H), 4.81 (d, $J = 3.6$ Hz, 1H), 4.61 – 4.55 (m, 2H), 4.41 (dd, $J = 10.5, 5.7$ Hz, 1H), 3.77 (s, 3H), 3.74 (d, $J = 8.8$ Hz, 1H), 3.72 – 3.67 (m, 2H), 3.55 (dd, $J = 9.9, 8.6$ Hz, 1H); $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ 155.55, 150.94, 149.07, 138.14, 137.95, 137.23, 132.90, 128.66, 128.51, 128.45, 128.25, 128.21, 127.89, 127.82, 127.79, 123.18, 120.82, 118.00, 114.74, 102.16, 84.28, 81.71, 76.30, 75.75, 75.16, 75.09, 72.50, 72.36, 55.66; HRMS (ESI) m/z $[\text{M}+\text{Na}]^+$ calcd for $\text{C}_{40}\text{H}_{39}\text{BrO}_{10}\text{SNa}$ 813.1340, found 813.1357.





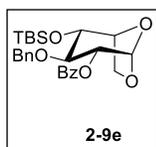
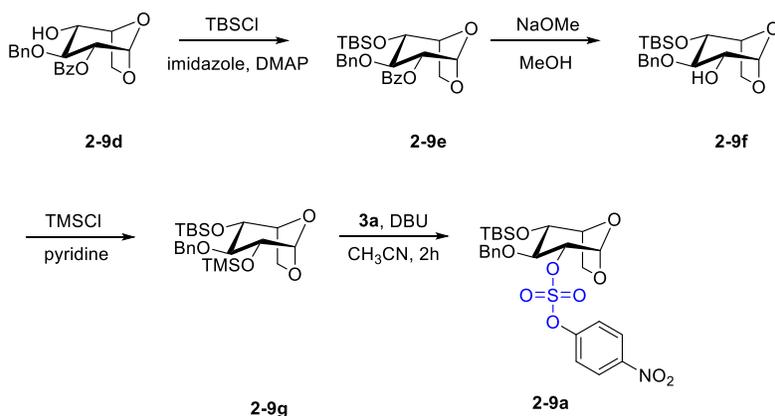
Compound **2-7b**⁵⁹ (1.44 g, 2.93 mmol) was converted to trimethylsilyl ether **2-7c** (1.37 g, 83%) by following **General Procedure A-2**. **S6** was used directly in the next step without further characterization.

Compound **2-7c** (219.2 mg, 0.39 mmol) coupled with aryl fluorosulfate **2-2j** (89.7 mg, 0.44 mmol) in the presence of TBD (18.6 mg, 0.13 mmol) providing compound **2-7** (189.3 mg, 72%) according to **General Procedure A-3**. $[\alpha]_{\text{D}}^{20} = -23.2$ (c 1.00, CHCl_3); $^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.47 – 7.42 (m, 2H), 7.38 – 7.27 (m, 8H), 7.26 (m, 2H), 7.14 – 7.08 (m, 2H), 6.88 – 6.82 (m, 2H), 4.90 (d, $J = 10.6$ Hz, 1H), 4.85 (d, $J = 3.9$ Hz, 1H), 4.83 (d, $J = 3.6$ Hz, 1H), 4.59 (dd, $J = 10.6, 1.9$ Hz, 1H), 4.56 (d, $J = 10.9$ Hz, 1H), 4.46 (dd, $J = 10.6, 4.8$ Hz, 1H), 4.38 (d, $J = 10.1$ Hz, 1H), 3.74 (s, 3H), 3.60 – 3.56 (m, 1H), 3.52 (t, $J = 9.1$ Hz, 1H), 3.44 (dd, $J = 9.9, 8.9$ Hz, 1H), 3.29 (dd, $J = 10.2, 9.3$ Hz, 1H), 2.34 (s, 3H); $^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 158.53, 143.70, 139.08, 137.30, 137.10, 134.27, 129.93, 128.69, 128.58, 128.28, 128.17, 128.15, 128.03, 126.54, 122.50, 114.86, 86.16, 84.91, 76.44, 76.39, 75.95, 75.23, 71.67, 64.73, 55.60, 21.17; HRMS (ESI) m/z $[\text{M}+\text{Na}]^+$ calcd for $\text{C}_{34}\text{H}_{35}\text{O}_8\text{N}_3\text{S}_2\text{Na}$ 700.1758, found 700.1752.



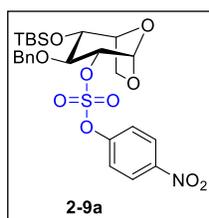
Compound **2-8b**⁶⁰⁻⁶¹ (358.1 mg, 0.96 mmol) was converted to trimethylsilyl ether in situ by HMDS (108.6 mg, 140.4 μL , 0.67 mmol) and the latter coupled with aryl fluorosulfate **2-2a** (360.2 mg, 1.63 mmol) via the following addition of DBU (22.0 mg, 0.14 mmol) to provide sulfoglucufuranose **2-8** (424.4 mg, 77%) as white solid

according to the **General Procedure B**. $[\alpha]_{\text{D}}^{20} = +53.1$ (c 1.01, CHCl_3); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 8.04 – 7.98 (m, 2H), 7.49 – 7.44 (m, 2H), 7.42 – 7.39 (m, 2H), 7.38-7.37 (m, 3H), 7.29 – 7.22 (m, 5H), 5.59 (s, 1H), 5.16 (d, $J = 3.7$ Hz, 1H), 4.94 (d, $J = 11.2$ Hz, 1H), 4.71 (dd, $J = 9.5, 3.7$ Hz, 1H), 4.63 (d, $J = 11.2$ Hz, 1H), 4.33 (dd, $J = 10.3, 4.8$ Hz, 1H), 4.15 (t, $J = 9.4$ Hz, 1H), 3.90 (td, $J = 9.8, 4.7$ Hz, 1H), 3.82 – 3.69 (m, 2H), 3.43 (s, 3H); $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ 154.37, 146.31, 137.67, 137.04, 129.31, 128.53, 128.47, 128.10, 128.04, 126.13, 125.61, 121.95, 101.71, 97.48, 82.71, 82.38, 75.51, 75.29, 68.89, 62.33, 55.81; HRMS (ESI) m/z $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{27}\text{H}_{28}\text{O}_{11}\text{NS}$ 574.1378, found 574.1351.

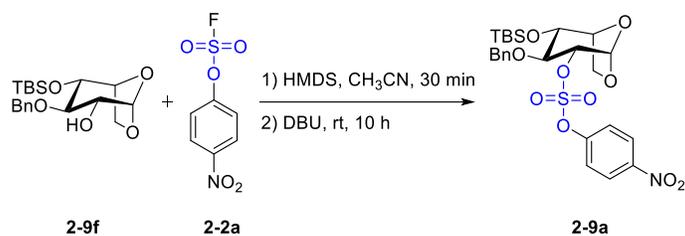


In a 500 mL round-bottom flask, **2-9d**⁶²⁻⁶³ (1.03 g, 2.89 mmol), imidazole (295.2 mg, 4.34 mmol), 4-dimethylaminopyridine (423.7 mg, 3.47 mmol) and *tert*-butyldimethylchlorosilane (522.8 mg, 3.47 mmol) were dissolved in 28 mL of acetonitrile. The mixture was stirred at 90 °C for 5 days. The solvent was removed under reduced pressure and the residue was diluted with ethyl acetate (30 mL). Water (50 mL) was added to the solution for extraction. The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by column chromatography using hexane/ethyl acetate (10: 1) as eluent, yielding the **2-9e** as a colorless oil (840 mg, 62%). $[\alpha]_{\text{D}}^{20} = +120.70$ (c 1.00 in CHCl_3). $^1\text{H NMR}$ (600 MHz, CDCl_3) δ 8.02 – 7.97 (m, 2H), 7.56 (ddt, $J = 8.7, 7.3, 1.3$ Hz, 1H), 7.46 – 7.38 (m, 2H), 7.22 – 7.12 (m, 5H), 5.51 (d, $J = 1.8$ Hz,

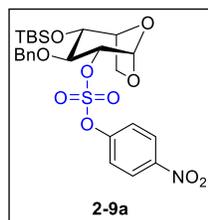
1H), 5.03 (dd, $J = 8.4, 1.9$ Hz, 1H), 4.75 (d, $J = 11.3$ Hz, 1H), 4.70 (d, $J = 11.3$ Hz, 1H), 4.35 (t, $J = 4.7$ Hz, 1H), 4.22 (dd, $J = 7.5, 0.8$ Hz, 1H), 4.01 (ddd, $J = 7.8, 4.4, 1.1$ Hz, 1H), 3.85 (t, $J = 8.1$ Hz, 1H), 3.75 (ddd, $J = 7.6, 5.1, 1.2$ Hz, 1H), 0.92 (s, 9H), 0.11 (s, 6H). ^{13}C NMR (151 MHz, CDCl_3) δ 165.78, 138.06, 133.23, 129.84, 129.53, 128.36, 128.21, 127.71, 127.52, 99.37, 80.70, 76.72, 75.83, 75.14, 72.62, 65.15, 25.71, 18.60, 17.85, 11.18. HRMS (ESI) m/z $[\text{M}+\text{H}]^+$ Calcd. For $\text{C}_{26}\text{H}_{35}\text{O}_6\text{Si}$: 471.2203, found 471.2199.



2-9e (488 mg, 1.04 mmol) was dissolved in methanol (5 mL), followed by the addition of NaOMe-methanol solution (231 μL , 1.04 mmol, 25% purity). The mixture was stirred for 1 h and neutralized by DOWEX 50WX8 resin. After filtration, the filtrate was concentrated under the reduced pressure. **General Procedure A-1** was employed to react with above crude product **2-9f**. The resulting product was further purified by column chromatography with hexane/ethyl acetate (10:1) as eluent to obtain **2-9g** (265.6 mg, 61%). **2-9g** (100 mg, 0.3 mmol) and **2-2a** (55.5 mg, 0.33 mmol) were coupled by following **General Procedure A-3**. The product was further purified by silica gel chromatography hexane/ethyl acetate (10:1) to obtain **2-9a** as a colorless liquid (102.9 mg, 80%). $[\alpha]_{\text{D}}^{20} = +49.79$ (c 1.00 in CHCl_3). ^1H NMR (600 MHz, CDCl_3) δ 8.09 (d, $J = 9.0$ Hz, 2H), 7.42 (d, $J = 9.0$ Hz, 2H), 7.43-7.25 (m, 5H), 5.68 (d, $J = 1.7$ Hz, 1H), 4.81 (d, $J = 10.9$ Hz, 1H), 4.69 – 4.63 (m, 2H), 4.35 (t, $J = 4.7$ Hz, 1H), 4.21 (d, $J = 7.7$ Hz, 1H), 3.99 (dd, $J = 7.8, 4.4$ Hz, 1H), 3.80 – 3.76 (m, 1H), 3.74 (t, $J = 8.0$ Hz, 1H), 0.89 (s, 9H), 0.10 (s, 3H), 0.09 (s, 3H); ^{13}C NMR (151 MHz, CDCl_3) δ 154.22, 146.41, 137.53, 128.62, 128.41, 127.66, 127.47, 125.83, 125.77, 125.68, 122.20, 122.05, 98.53, 98.41, 86.72, 86.64, 80.30, 80.17, 75.91, 75.82, 73.14, 73.00, 65.59, 25.80, 25.77, 17.94; HRMS (ESI) m/z $[\text{M}+\text{H}]^+$ Calcd. For $\text{C}_{25}\text{H}_{34}\text{O}_{10}\text{SiNS}$: 568.1667, found 568.1626.

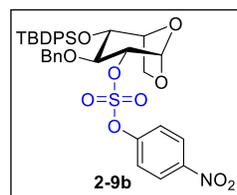
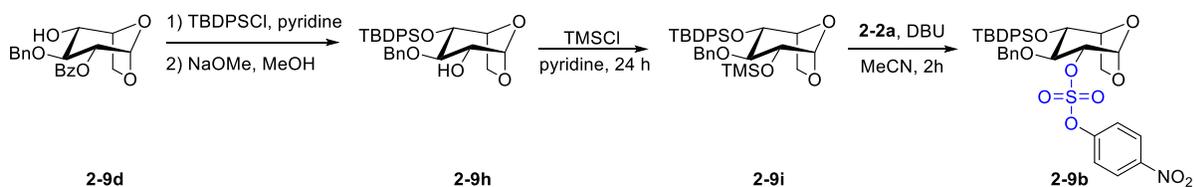


One-pot synthesis for compound **2-9a**:



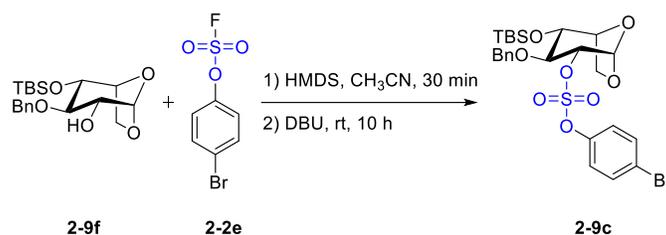
Compound **2-9f** (330.6 mg, 0.90 mmol) was converted to the corresponding trimethylsilyl ether **2-9f** *in situ* by HMDS (101.9 mg, 131.7 μ L, 0.63 mmol) and coupled with aryl fluorosulfate **2-2a** (302.0 mg, 1.63 mmol) via the following addition of DBU (22.0 mg, 0.14 mmol) to provide idosyl aryl sulfate diester **2-9a**

(437.8 mg, 86%) according to the **General Procedure B**. Note: The whole reaction mixture was washed by aqueous solution.

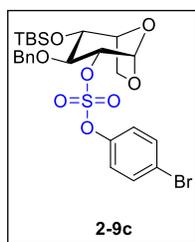


To a solution of compound **2-9d**⁶²⁻⁶³ (437.5 mg, 1.23 mmol) in mixed dry solvent DMF/THF (9mL/4.5 mL) was consecutively added imidazole (167.2 mg, 2.46 mmol), *tert*-butyl(chloro)diphenylsilane (474.9 mg, 2.46 mmol) at rt. After 80 h, the mixture was concentrated under reduced pressure and purified for the usage in next step. The obtained product (731.6 mg, 1.23 mmol) was dissolved in mixed solvent DCM/MeOH (13.3 mL/13.3 mL), followed by the addition of NaOMe (225.9 mg, 4.18 mmol) at room temperature. After 22 h, the mixture was neutralized with H-form resin to pH = 7, concentrated and purified on silica gel column, providing compound **2-9h** (603.5 mg, 89%). **2-9h** (403 mg, 0.82 mmol) was directly used to

prepare **2-9i** by following **General Procedure A-1**. The product was further purified by silica gel column chromatography (hexane/ethyl acetate 10:1) to yield **2-9i** as a colorless liquid (361.5 mg, 78%). Compound **2-9b** was prepared by following **General Procedure A-1** from **2-9i** (104 mg, 0.18 mmol). The product was further purified by silica gel chromatography (hexane/ethyl acetate 12:1) to obtain **2-9b** as a colorless liquid (95.9 mg, 75%). $[\alpha]_{\text{D}}^{20} = +74.48$ (c 2.00 in CHCl_3). $^1\text{H NMR}$ (600 MHz, CDCl_3) δ 8.16 – 8.10 (m, 2H), 7.70 – 7.65 (m, 2H), 7.64 – 7.60 (m, 2H), 7.47 – 7.38 (m, 4H), 7.36 (t, $J = 7.4$ Hz, 2H), 7.32 – 7.24 (m, 5H), 7.20 (dd, $J = 7.1, 2.5$ Hz, 2H), 5.58 (d, $J = 1.7$ Hz, 1H), 4.87 (d, $J = 10.6$ Hz, 1H), 4.76 (d, $J = 10.6$ Hz, 1H), 4.62 (dd, $J = 8.1, 1.8$ Hz, 1H), 4.28 (dd, $J = 7.6, 0.8$ Hz, 1H), 4.14 (ddd, $J = 7.9, 4.4, 1.1$ Hz, 1H), 3.92 (t, $J = 7.9$ Hz, 1H), 3.85 (t, $J = 4.6$ Hz, 1H), 3.59 (dd, $J = 7.8, 5.0$ Hz, 1H), 1.08 (s, 9H). $^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 154.06, 146.30, 137.29, 135.78, 135.74, 133.43, 131.97, 130.44, 130.13, 128.32, 128.13, 127.85, 127.83, 127.68, 125.62, 122.00, 98.22, 86.76, 80.35, 75.61, 74.71, 73.39, 65.46, 26.95, 19.16. HRMS (ESI) m/z $[\text{M}+\text{NH}_4]^+$ Calcd. For $\text{C}_{35}\text{H}_{41}\text{N}_2\text{O}_{10}\text{SSi}$: 709.2251, found 709.2242.



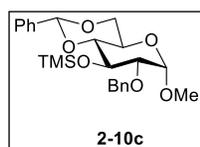
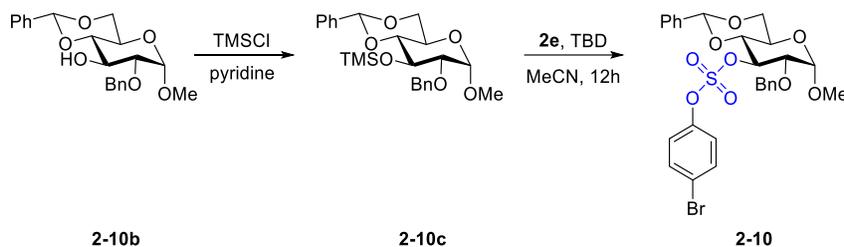
One-pot synthesis for compound **2-9c**:



Compound **2-9f** (2.32 g, 6.34 mmol) was converted to trimethylsilyl ether in situ by HMDS (614.0 mg, 793.2 μL , 3.80 mmol) and coupled with fluoride sulfate **2-2e** (1.79 g, 7.02 mmol) via the following addition of DBU (96.5 mg, 0.63 mmol) to provide sulfoglucofuranose **2-9c** (3.47 g, 91%) as colorless syrup according to the

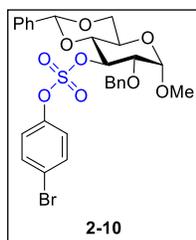
General Procedure B. Note: The whole reaction mixture was washed by aqueous solution. $[\alpha]_{\text{D}}^{20} =$

+41.2 (c 1.01, CHCl_3); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.41 – 7.37 (m, 2H), 7.35 – 7.25 (m, 5H), 7.22 – 7.17 (m, 2H), 5.65 (d, J = 1.7 Hz, 1H), 4.76 (d, J = 10.8 Hz, 1H), 4.71 (d, J = 11.1 Hz, 1H), 4.63 (dd, J = 8.1, 1.7 Hz, 1H), 4.34 (t, J = 4.7 Hz, 1H), 4.20 (d, J = 7.7 Hz, 1H), 3.96 (ddd, J = 7.9, 4.4, 1.1 Hz, 1H), 3.78 (ddd, J = 7.8, 5.1, 1.1 Hz, 1H), 3.72 (t, J = 7.9 Hz, 1H), 0.89 (s, 9H), 0.09 (m, 6H); $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ 149.28, 137.65, 133.12, 128.47, 127.91, 127.89, 123.27, 121.07, 98.59, 86.30, 75.87, 75.73, 72.96, 65.53, 25.81, 17.96, -4.42, -4.55; HRMS (ESI) m/z $[\text{M}+\text{NH}_4]^+$ calcd for $\text{C}_{25}\text{H}_{37}\text{O}_8\text{NSi}$ 618.1187, found 618.1174.



2-10c was prepared by following **General Procedure A-1** with **2-10b** (1.32 g, 3.55 mmol). **2-10b** was obtained according to the literature.⁴⁵ The product was purified

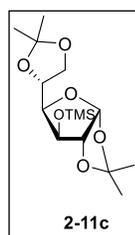
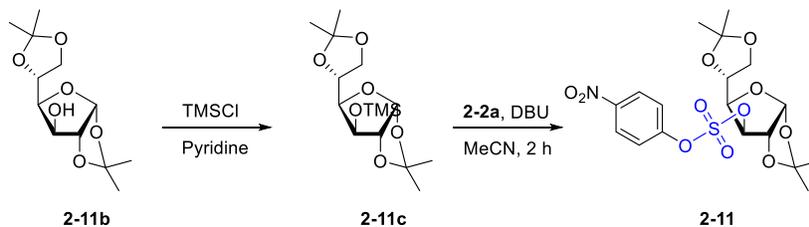
by silica gel chromatography (hexane/ethyl acetate 6:1) to obtain **2-10c** (1.53 g, 97%). $[\alpha]_{\text{D}}^{20}$ = -4.67 (c 1.00 in CHCl_3). $^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.48 (dd, J = 7.5, 1.9 Hz, 2H), 7.41 – 7.31 (m, 7H), 7.30 (t, J = 7.1 Hz, 1H), 5.49 (s, 1H), 4.84 (d, J = 12.3 Hz, 1H), 4.64 (d, J = 12.3 Hz, 1H), 4.52 (d, J = 3.7 Hz, 1H), 4.23 (dd, J = 10.1, 4.8 Hz, 1H), 4.12 (t, J = 9.0 Hz, 1H), 3.77 (ddd, J = 10.6, 9.8, 4.8 Hz, 1H), 3.67 (t, J = 10.3 Hz, 1H), 3.43 – 3.39 (m, 2H), 3.38 (s, 3H), 0.14 (s, 9H). $^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 138.30, 137.39, 128.88, 128.42, 128.16, 128.12, 127.89, 126.14, 101.63, 99.44, 82.07, 79.74, 73.91, 71.70, 69.03, 62.23, 55.30, 0.54. HRMS (DART) m/z $[\text{M}+\text{H}]^+$ Calcd. For $\text{C}_{24}\text{H}_{33}\text{O}_6\text{Si}$: 445.2046, found 445.2035.



Compound **2-10** was prepared by following **General Procedure A-3** from **2-10c** (120 mg, 0.27 mmol) and 12 h of reaction time. The product was purified by silica gel chromatography (hexane/ethyl acetate 6:1) to obtain **2-10** (109.0 mg, 67%) and biproduct **2-10b** (30.0 mg, 30%).

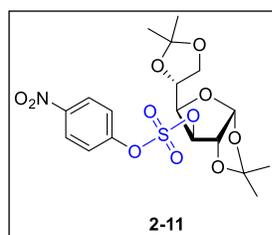
2-10: $[\alpha]_{\text{D}}^{20} = -24.66$ (c 1.00 in CHCl_3). $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.45 – 7.27 (m, 10H), 7.21 – 7.15 (m, 2H), 7.07 – 7.00 (m, 2H), 5.48 (s, 1H), 5.21 (t, $J = 9.5$ Hz, 1H), 4.81 (d, $J = 12.1$ Hz, 1H), 4.66 – 4.57 (m, 2H), 4.28 (dd, $J = 10.4, 4.8$ Hz, 1H), 3.87 (td, $J = 9.9, 4.8$ Hz, 1H), 3.71 (td, $J = 10.0, 9.5, 8.6$ Hz, 2H), 3.65 (dd, $J = 9.4, 3.6$ Hz, 1H), 3.36 (s, 3H). $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ 149.40, 137.24, 136.53, 132.54, 129.32, 128.58, 128.32, 128.25, 126.25, 122.85, 120.17, 101.85, 98.89, 83.76, 78.89, 77.39, 73.64, 68.83, 62.31, 55.59. HRMS (ESI) m/z $[\text{M}+\text{NH}_4]^+$ Calcd. For $\text{C}_{27}\text{H}_{31}\text{BrNO}_9\text{S}$: 624.0903, found 624.0893.

2-10b: $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.57 – 7.27 (m, 9H), 5.51 (s, 1H), 4.81 – 4.65 (m, 2H), 4.61 (d, $J = 3.7$ Hz, 1H), 4.25 (dd, $J = 10.2, 4.8$ Hz, 1H), 4.14 (t, $J = 9.3$ Hz, 1H), 3.80 (td, $J = 9.9, 4.8$ Hz, 1H), 3.69 (t, $J = 10.3$ Hz, 1H), 3.53 – 3.41 (m, 2H), 3.36 (s, 3H), 2.62 (s, 1H). Spectral data matched those previously reported.⁴⁵



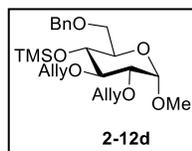
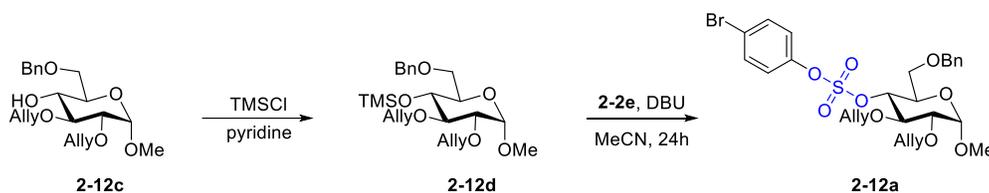
2-11c was prepared by following **General Procedure A-1** from commercially available compound **2-11b** (300 mg, 1.15 mmol). The product was purified by silica gel chromatography (hexane/ethyl acetate 10:1) to obtain **2-11c** (358.1 mg, 93%). $[\alpha]_{\text{D}}^{20} = -18.60$ (c 1.00 in CHCl_3). $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 5.90 (d, $J = 3.6$ Hz, 1H), 4.35 (d,

$J = 3.6$ Hz, 1H), 4.25 – 4.15 (m, 2H), 4.08 (dd, $J = 8.5, 6.2$ Hz, 1H), 4.05 (dd, $J = 7.8, 2.7$ Hz, 1H), 3.97 (dd, $J = 8.5, 6.1$ Hz, 1H), 1.50 (s, 3H), 1.41 (s, 3H), 1.33 (s, 3H), 1.31 (s, 3H), 0.17 (s, 9H). ^{13}C NMR (151 MHz, CDCl_3) δ 111.76, 108.82, 105.32, 85.58, 81.85, 75.25, 72.38, 67.42, 26.88, 26.74, 26.31, 25.34, -0.15. HRMS (ESI) m/z $[\text{M}+\text{H}]^+$ Calcd. For $\text{C}_{15}\text{H}_{29}\text{O}_6\text{Si}$: 333.1733, found 333.1727.



Compound **2-11** was prepared by following **General Procedure A-3** from **2-11c** (100 mg, 0.3 mmol). The product was purified by silica gel chromatography (hexane/ethyl acetate 10:1) to obtain **2-11** (118.0 mg, 94%). $[\alpha]_{\text{D}}^{20} = -89.44$ (c 1.00 in CHCl_3). ^1H NMR (600 MHz, CDCl_3) δ 8.31 (d, $J = 9.2$ Hz, 2H), 7.59

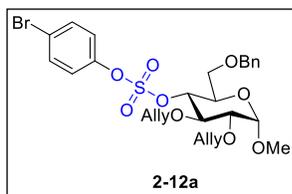
(d, $J = 9.2$ Hz, 2H), 5.96 (d, $J = 3.7$ Hz, 1H), 5.19 (d, $J = 2.6$ Hz, 1H), 4.94 (d, $J = 3.7$ Hz, 1H), 4.22 (ddd, $J = 9.6, 5.9, 4.0$ Hz, 1H), 4.18 (dd, $J = 8.8, 2.6$ Hz, 1H), 4.12 (dd, $J = 8.9, 5.9$ Hz, 1H), 4.05 (dd, $J = 8.9, 4.0$ Hz, 1H), 1.53 (s, 3H), 1.42 (s, 3H), 1.34 (s, 3H), 1.27 (s, 3H). ^{13}C NMR (151 MHz, CDCl_3) δ 154.26, 146.37, 125.63, 121.94, 112.93, 109.83, 104.81, 86.73, 82.53, 79.64, 71.45, 67.34, 26.94, 26.52, 26.20, 25.08. HRMS (ESI) m/z $[\text{M}+\text{H}]^+$ Calcd. For $\text{C}_{18}\text{H}_{23}\text{NO}_{11}\text{S}$: 462.1070, found 462.1064.



2-12d was prepared by following **General Procedure A-1** from **2-12c** (440 mg, 1.31 mmol). **2-12c** was obtained according to the literature.⁴⁶ The product was purified by silica gel chromatography (hexane/ethyl acetate 10:1) to obtain **2-12c** (291.0 mg, 54%).

$[\alpha]_{\text{D}}^{20} = +68.59$ (c 1.00 in CHCl_3). ^1H NMR (600 MHz, CDCl_3) δ 7.37 – 7.25 (m, 5H), 6.00 – 5.85 (m, 2H), 5.26 (ddq, $J = 17.2, 6.8, 1.7$ Hz, 2H), 5.15 (ddq, $J = 25.4, 10.5, 1.4$ Hz, 2H), 4.78 (d, $J = 3.5$ Hz, 1H), 4.60 (d, $J = 12.1$ Hz, 1H), 4.56 (d, $J = 12.1$ Hz, 1H), 4.39 (ddt, $J = 12.4, 5.5, 1.6$ Hz, 1H),

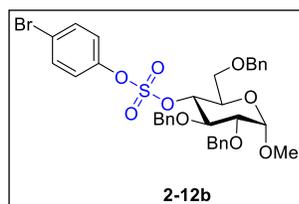
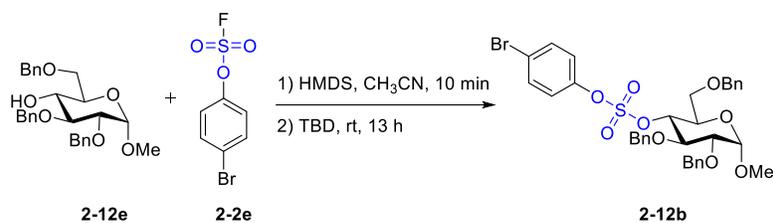
4.21 – 4.10 (m, 3H), 3.70 – 3.59 (m, 4H), 3.56 (dd, $J = 9.5, 8.4$ Hz, 1H), 3.41 (s, 4H), 0.09 (s, 9H). ^{13}C NMR (151 MHz, CDCl_3) δ 138.07, 135.32, 134.87, 128.30, 127.66, 127.55, 117.57, 116.13, 98.11, 81.20, 79.97, 74.18, 73.46, 72.46, 71.16, 70.88, 68.92, 54.99, 0.49. HRMS (ESI) m/z $[\text{M}+\text{NH}_4]^+$ Calcd. For $\text{C}_{23}\text{H}_{40}\text{NO}_6\text{Si}$: 454.2625, found 454.2624.



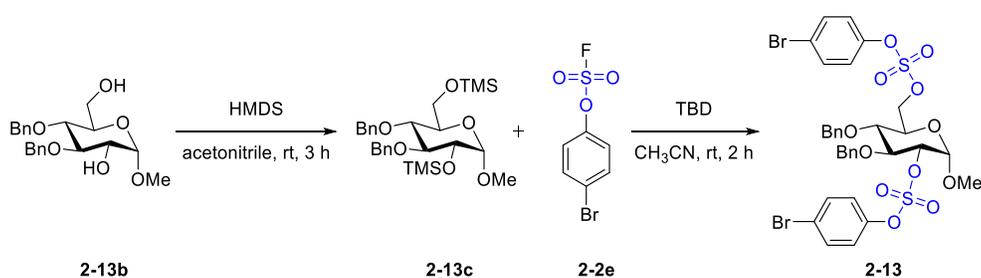
Compound **2-12a** was prepared by following **General Procedure A-3** from **2-12d** (109 mg, 0.25 mmol). DBU was used as the catalyst. The product was purified by silica gel chromatography (hexane/ethyl acetate 6:1) to obtain **2-12a** (89.3 mg, 60%) and biproduct **2-12c** (36 mg, 39%).

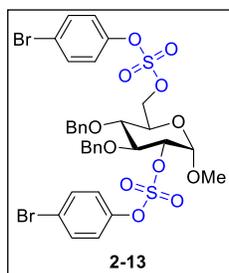
2-12a: $[\alpha]_{\text{D}}^{20} = +57.58$ (c 1 in CHCl_3). ^1H NMR (500 MHz, CDCl_3) δ 7.52 – 7.45 (m, 2H), 7.37 – 7.22 (m, 5H), 7.19 – 7.12 (m, 2H), 5.97 – 5.84 (m, 2H), 5.32 – 5.18 (m, 3H), 5.11 (dq, $J = 10.3, 1.3$ Hz, 1H), 4.91 (dd, $J = 10.1, 9.1$ Hz, 1H), 4.80 (d, $J = 3.5$ Hz, 1H), 4.56 – 4.45 (m, 2H), 4.34 (ddt, $J = 11.8, 6.0, 1.4$ Hz, 1H), 4.24 – 4.11 (m, 3H), 3.92 – 3.85 (m, 2H), 3.69 (dd, $J = 10.9, 2.2$ Hz, 1H), 3.63 (dd, $J = 10.9, 3.9$ Hz, 1H), 3.50 (dd, $J = 9.6, 3.6$ Hz, 1H), 3.44 (s, 3H). ^{13}C NMR (126 MHz, CDCl_3) δ 149.20, 137.66, 134.46, 134.32, 132.96, 128.31, 127.80, 127.67, 123.17, 120.75, 118.21, 117.30, 97.99, 82.14, 79.29, 78.15, 74.41, 73.67, 72.73, 68.20, 67.90, 55.58. HRMS (ESI) m/z $[\text{M}+\text{NH}_4]^+$ Calcd. For $\text{C}_{26}\text{H}_{35}\text{BrNO}_9\text{S}$: 616.1216, found 616.1213.

2-12c: ^1H NMR (500 MHz, CDCl_3) δ 7.39 – 7.24 (m, 5H), 5.94 (m, 2H), 5.33 – 5.24 (m, 2H), 5.18 (ddd, $J = 10.3, 6.5, 1.6$ Hz, 2H), 4.80 (d, $J = 3.5$ Hz, 1H), 4.66 – 4.55 (m, 2H), 4.42 (ddt, $J = 12.6, 5.5, 1.5$ Hz, 1H), 4.28 – 4.09 (m, 3H), 3.76 – 3.69 (m, 3H), 3.67 – 3.56 (m, 2H), 3.42 (s, 4H). Spectral data matched those previously reported.⁴⁶



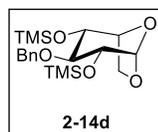
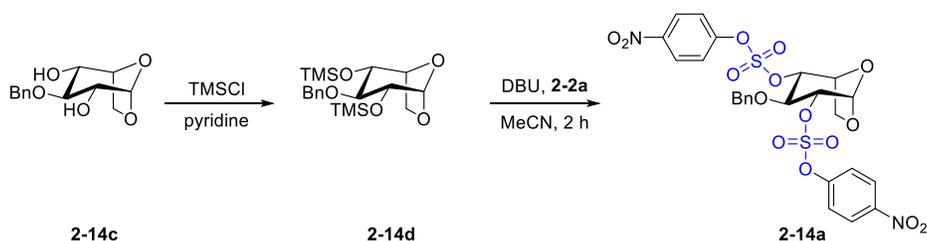
Compound **2-12e**⁴⁵ (515.2 mg, 1.11 mmol) was converted to trimethylsilyl ether in situ by HMDS (107.4 mg, 138.8 μ L, 0.67 mmol) and coupled with fluoride sulfate **2-2e** (311.2 mg, 1.22 mmol) via the following addition of TBD (30.9 mg, 0.22 mmol) to provide sulfoglucofuranose **2-12b** (519.0 mg, 67%) as colorless syrup according to the **General Procedure B**. $[\alpha]_{\text{D}}^{20} = +10.398$ (c 1.00 in CHCl_3) $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 8.12 (d, $J = 9.2$ Hz, 2H), 7.35 – 7.21 (m, 17H), 5.03 (t, $J = 9.6$ Hz, 1H), 4.95 (d, $J = 10.5$ Hz, 1H), 4.77 (d, $J = 12.1$ Hz, 1H), 4.70 (d, $J = 10.5$ Hz, 1H), 4.64 – 4.58 (m, 2H), 4.55 (d, $J = 11.6$ Hz, 1H), 4.51 (d, $J = 11.7$ Hz, 1H), 4.08 (t, $J = 9.3$ Hz, 1H), 3.95 (d, $J = 10.0$ Hz, 1H), 3.77 – 3.67 (m, 2H), 3.62 (dd, $J = 9.6, 3.6$ Hz, 1H), 3.40 (s, 3H). $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ 154.08, 146.09, 137.80, 137.55, 137.53, 128.57, 128.35, 128.25, 128.18, 128.15, 127.83, 127.76, 127.69, 127.62, 125.45, 121.82, 97.90, 82.84, 79.72, 78.39, 75.35, 73.77, 73.61, 68.11, 67.90, 55.71. HRMS (ESI) m/z $[\text{M}+\text{Na}]^+$ Calcd. For $\text{C}_{26}\text{H}_{35}\text{BrNaO}_9\text{S}$: 721.1083, found 721.1054.





Compound **2-13b**⁶⁴ (104.1 mg, 0.28 mmol) was converted to trimethylsilyl ether **2-13c** (137.4 mg, 95%) by following **General Procedure A-2**. **2-13c** was used directly in the next step without further characterization.

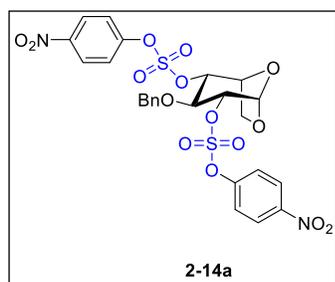
Compound **2-13c** (105.8 mg, 0.20 mmol) coupled with fluoride sulfate **2-2e** (110.7 mg, 0.44 mmol) in the presence of TBD (8.4 mg, 0.06 mmol) providing compound **2-13** (125.2 mg, 73%) according to **General Procedure A-3**. $[\alpha]_D^{20} = +51.6$ (c 1.00, CHCl_3); $^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.53 – 7.46 (m, 2H), 7.40 – 7.36 (m, 2H), 7.36 – 7.27 (m, 8H), 7.25 – 7.22 (m, 2H), 7.22 – 7.17 (m, 2H), 7.17 – 7.14 (m, 2H), 5.04 (d, $J = 3.6$ Hz, 1H), 4.89 (d, $J = 11.0$ Hz, 1H), 4.83 (d, $J = 10.8$ Hz, 1H), 4.77 (d, $J = 10.8$ Hz, 1H), 4.65 (dd, $J = 9.8, 3.6$ Hz, 1H), 4.57 (d, $J = 11.0$ Hz, 1H), 4.50 (dd, $J = 10.7, 2.0$ Hz, 1H), 4.45 (dd, $J = 10.7, 4.8$ Hz, 1H), 4.11 (dd, $J = 9.7, 8.7$ Hz, 1H), 3.93 (ddd, $J = 10.1, 4.9, 1.9$ Hz, 1H), 3.56 (dd, $J = 10.2, 8.7$ Hz, 1H), 3.36 (s, 3H); $^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 149.30, 149.21, 137.40, 137.16, 133.21, 133.14, 128.83, 128.67, 128.46, 128.29, 128.18, 128.09, 123.10, 123.04, 121.13, 121.08, 96.78, 82.64, 79.18, 77.02, 75.99, 75.53, 72.29, 68.54, 55.88; HRMS (ESI) m/z $[M+\text{NH}_4]^+$ calcd for $\text{C}_{33}\text{H}_{36}\text{O}_{12}\text{Br}_2\text{S}_2\text{N}$ 860.0040, found 860.0006.



2-14d was prepared by following **General Procedure A-1** from **2-14c** (200 mg, 0.79 mmol). **2-14c** was prepared according to the literature⁶²⁻⁶³. The product was purified by

silica gel chromatography (hexane/ethyl acetate 10:1) to obtain **2-14d** (313.8 mg, quantitative). $[\alpha]_D^{20} = +37.16$ (c 0.33, CHCl_3) $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.37 – 7.22 (m, 5H), 5.16 (d, $J = 1.8$ Hz, 1H), 4.82 – 4.74 (m, 2H), 4.23 (t, $J = 4.7$ Hz, 1H), 4.15 (dd, $J = 7.6, 0.8$ Hz, 1H), 3.84 (ddd, $J = 8.0, 4.3, 1.1$

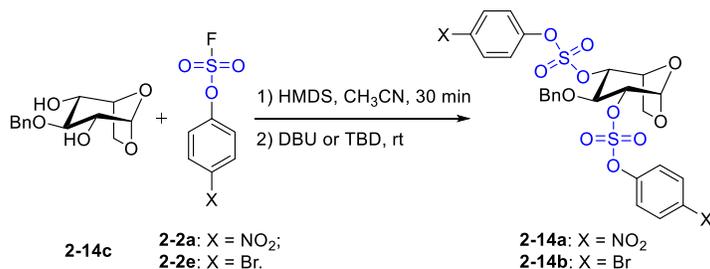
Hz, 1H), 3.71 (ddd, $J = 7.6, 5.1, 1.1$ Hz, 1H), 3.63 (dd, $J = 7.8, 1.8$ Hz, 1H), 3.50 (t, $J = 7.9$ Hz, 1H), 0.14 (s, 9H), 0.11 (s, 9H). ^{13}C NMR (151 MHz, CDCl_3) δ 138.79, 128.14, 127.54, 127.31, 102.17, 83.78, 76.02, 75.89, 75.48, 72.40, 65.16, 0.30, 0.24. HRMS (DART) m/z $[\text{M}+\text{NH}_4]^+$ calcd for $\text{C}_{19}\text{H}_{33}\text{O}_5\text{Si}_2$ 397.1867, found 397.1858.

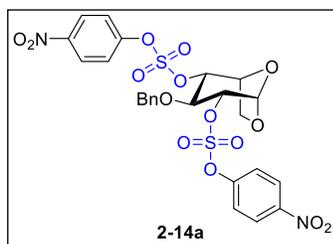


Compound **2-14a** was prepared by following **General Procedure A-3** in two batches: **2-14d** (100 mg, 0.25 mmol), DBU (15.4 mg, 0.1 mmol); **2-14d** (85 mg, 0.21 mmol), TBD (11.9 mg, 0.086 mmol). The product was purified by silica gel chromatography (hexane/ethyl acetate 8:1~4:1) to

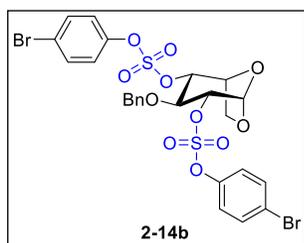
obtain **2-14a** (DBU: 117 mg, 71%; TBD: 100 mg, 71%). $[\alpha]_{\text{D}}^{20} = +10.20$ (c 1.00 in CHCl_3). ^1H NMR (600 MHz, CDCl_3) δ 8.21 – 8.15 (m, 2H), 8.09 – 8.03 (m, 2H), 7.49 – 7.43 (m, 2H), 7.38 – 7.32 (m, 2H), 7.32 – 7.24 (m, 3H), 7.20 – 7.16 (m, 2H), 5.75 (d, $J = 1.7$ Hz, 1H), 5.01 (t, $J = 4.6$ Hz, 1H), 4.97 (ddd, $J = 8.3, 4.3, 1.2$ Hz, 1H), 4.80 – 4.74 (m, 2H), 4.65 (d, $J = 10.9$ Hz, 1H), 4.22 (dd, $J = 8.7, 0.7$ Hz, 1H), 4.03 (t, $J = 8.1$ Hz, 1H), 3.94 (ddd, $J = 8.7, 4.9, 1.3$ Hz, 1H). ^{13}C NMR (151 MHz, CDCl_3) δ 153.81, 153.59, 146.50, 146.41, 136.12, 128.56, 128.37, 127.34, 127.32, 125.74, 125.72, 121.97, 121.61, 98.37, 85.72, 81.97, 76.10, 75.40, 72.56, 65.85. HRMS (ESI) m/z $[\text{M}+\text{NH}_4]^+$ Calcd. For $\text{C}_{25}\text{H}_{26}\text{N}_3\text{O}_{15}\text{S}_2$: 672.0805, found 672.0787.

One-pot preparation of compound 2-14a & 2-14b:

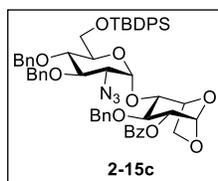




Compound **2-14c**⁶²⁻⁶³ (204.6 mg, 0.81 mmol) was converted to trimethylsilyl ether in situ by HMDS (144.0 mg, 186.0 μ L, 0.89 mmol) and coupled with aryl fluorosulfate **2-2a** (425.6 mg, 1.92 mmol) via the following addition of DBU (24.5 mg, 0.16 mmol) to provide sulfidopyranose **2-14a** (360.2 mg, 68%) in 5 h according to the **General Procedure B**.

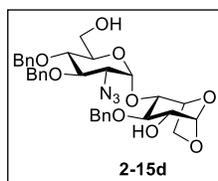


Compound **2-14c**⁶²⁻⁶³ (1.60 g, 6.34 mmol) was converted to trimethylsilyl ether in situ by HMDS (1.13 g, 1.45 mL, 6.98 mmol) and coupled with aryl fluorosulfate **2-2e** (3.40 g, 13.32 mmol) via the following addition of TBD (176.6 mg, 1.27 mmol) to provide sulfidopyranose **2-14b** (4.32 g, 94%) as white solid in 13 h according to the **General Procedure B**. $[\alpha]_{\text{D}}^{20} = +7.8$ (c 1.01, CHCl_3); ^1H NMR (500 MHz, CDCl_3) δ 7.46 – 7.41 (m, 2H), 7.39 – 7.34 (m, 2H), 7.32 – 7.28 (m, 3H), 2.24-7.23 (m, 2H), 7.21 – 7.17 (m, 2H), 7.14 – 7.09 (m, 2H), 5.69 (d, $J = 1.8$ Hz, 1H), 4.93 – 4.87 (m, 2H), 4.74 (d, $J = 10.8$ Hz, 1H), 4.71 (dd, $J = 7.9, 1.8$ Hz, 1H), 4.66 (d, $J = 10.7$ Hz, 1H), 4.15 (d, $J = 8.5$ Hz, 1H), 4.00 – 3.94 (m, 1H), 3.88-3.85 (m, 1H); ^{13}C NMR (126 MHz, CDCl_3) δ 149.15, 148.96, 136.48, 133.30, 133.26, 128.65, 128.64, 128.61, 128.38, 128.08, 123.17, 122.90, 121.41, 121.38, 98.62, 85.42, 81.80, 76.25, 75.56, 72.73, 65.90; HRMS (ESI) m/z $[M+\text{NH}_4]^+$ calcd for $\text{C}_{25}\text{H}_{26}\text{O}_{11}\text{NBr}_2\text{S}_2$ 737.9309, found 737.9297.



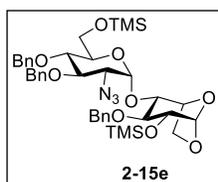
2-9d was prepared according to literature.⁶²⁻⁶³ A mixture of **2-15b** (267.0 mg, 365.8 μmol) and **2-9d** (156.4 mg, 438.9 μmol) was co-evaporated with toluene (3×2 mL) in Schlenk tube (10 mL) and placed under vacuum for 1 h. Under N_2 protection,

the reactants were dissolved in dry DCM (7 mL), followed by the addition of freshly dried 4 Å molecular sieves (700 mg). The mixture was stirred at room temperature for 0.5 h, and then cooled to -48 °C (acetonitrile/dry ice bath). N-Iodosuccinimide (118.5 mg, 526.7 μmol) and trifluoromethanesulfonic acid (TfOH, 11.0 mg, 73.2 μmol , 6.4 μL) were added to the reaction flask. The solution was warmed up by removing the acetonitrile/dry ice bath and stirred for 2 h. Et_3N (1 mL) was added to quench the reaction. The whole mixture was filtered through celite and the filtrate was sequentially washed with aqueous $\text{Na}_2\text{S}_2\text{O}_3$ (10%, 10 mL) and brine. The organic layer was dried over anhydrous MgSO_4 and concentrated. The crude product was purified by column chromatography (hexane/EA = 15:1 ~10:1 v/v) to get **2-15c** (270 mg, 77%). $[\alpha]_{\text{D}}^{20} +90.64$ (*c* 0.50 in CHCl_3). ^1H NMR (500 MHz, CDCl_3) δ 8.07 – 8.02 (m, 2H), 7.67 (dt, $J = 8.2, 1.4$ Hz, 4H), 7.62 – 7.54 (m, 1H), 7.47 – 7.27 (m, 18H), 7.25 – 7.18 (m, 3H), 7.16-7.12 (m, 2H), 5.50 (d, $J = 1.8$ Hz, 1H), 5.27 (d, $J = 3.8$ Hz, 1H), 5.06 (dd, $J = 8.0, 1.9$ Hz, 1H), 4.96 – 4.88 (m, 3H), 4.85 (d, $J = 10.5$ Hz, 1H), 4.76 (d, $J = 10.8$ Hz, 1H), 4.61 (d, $J = 10.5$ Hz, 1H), 4.57 (t, $J = 4.6$ Hz, 1H), 4.12 (d, $J = 7.7$ Hz, 1H), 4.05 (t, $J = 8.1$ Hz, 1H), 4.02 – 3.95 (m, 2H), 3.90 – 3.81 (m, 2H), 3.72 – 3.61 (m, 3H), 3.40 (dd, $J = 10.3, 3.8$ Hz, 1H), 1.06 (s, 9H). ^{13}C NMR (101 MHz, CDCl_3) δ 165.72, 137.92, 137.64, 137.44, 135.75, 135.61, 133.36, 133.13, 132.82, 129.86, 129.82, 129.78, 129.41, 128.57, 128.44, 128.36, 128.14, 128.08, 128.06, 127.88, 127.77, 127.72, 127.69, 99.63, 99.22, 79.92, 79.33, 78.85, 78.24, 77.20, 75.60, 75.46, 74.96, 74.20, 72.79, 65.66, 63.59, 62.51, 26.78, 19.19. HRMS (ESI) m/z $[\text{M}+\text{NH}_4]^+$ Calcd. For $\text{C}_{56}\text{H}_{63}\text{N}_4\text{O}_{10}\text{Si}$: 979.4313, found 979.4293.



2-15c (258.3 mg, 268.5 μmol) was dissolved in the mixture of THF/methanol (v/v, 3 mL/3 mL), followed by the addition of NaOMe-methanol solution (299.34 μL , 1.34 mmol, 25% purity). The mixture was stirred for 1 h and neutralized by

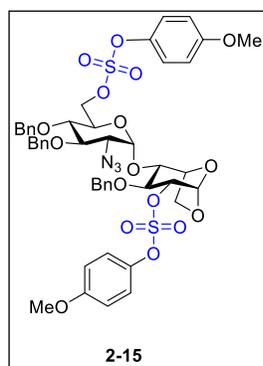
DOWEX 50WX8 resin. After filtration, the filtrate was concentrated under the reduced pressure. The crude product was directly dissolved in THF (5 mL). To the solution was added tetrabutylammonium fluoride (TBAF, 1 M, 1.34 mL) and the mixture was stirred overnight. The solvent was removed under reduced pressure, and the residue was purified by silica column chromatography (ethyl acetate/hexane=1/2~1/0, v/v) to give the **2-15d** (154.6 mg, 93%). $[\alpha]_{\text{D}}^{20} = +12.66$ (c 0.10 in CHCl_3). ^1H NMR (500 MHz, CDCl_3) δ 7.43 – 7.27 (m, 15H), 5.31 – 5.26 (m, 2H), 4.94 – 4.88 (m, 4H), 4.87 (d, $J = 10.7$ Hz, 1H), 4.66 (d, $J = 10.8$ Hz, 1H), 4.50 (t, $J = 4.7$ Hz, 1H), 4.04 (d, $J = 7.8$ Hz, 1H), 3.95 (dt, $J = 10.4, 4.2$ Hz, 1H), 3.88 (dd, $J = 8.0, 4.3$ Hz, 1H), 3.82-3.70 (m, 3H), 3.68 – 3.57 (m, 4H), 3.33 (dd, $J = 10.4, 3.8$ Hz, 1H), 1.92 (d, $J = 8.6$ Hz, 1H), 1.66 (t, $J = 6.6$ Hz, 1H). ^{13}C NMR (101 MHz, CDCl_3) δ 138.27, 137.61, 137.43, 128.65, 128.52, 128.22, 128.13, 128.03, 128.01, 127.95, 127.85, 101.52, 99.42, 83.40, 79.63, 78.11, 77.79, 77.20, 75.84, 75.44, 75.39, 74.83, 74.17, 72.02, 65.41, 63.41, 61.42. HRMS (ESI) m/z $[M+\text{NH}_4]^+$ Calcd. For $\text{C}_{33}\text{H}_{41}\text{N}_4\text{O}_9$: 637.2874, found 637.2858.



2-15e was prepared by following **General Procedure A-1** from **2-15d** (148.5 mg, 0.24 mmol). The product was purified by silica gel chromatography (hexane/ethyl acetate 6:1) to obtain **2-15e** (150 mg, 82%). $[\alpha]_{\text{D}}^{20} = +31.99$ (c 0.40 in CHCl_3). ^1H

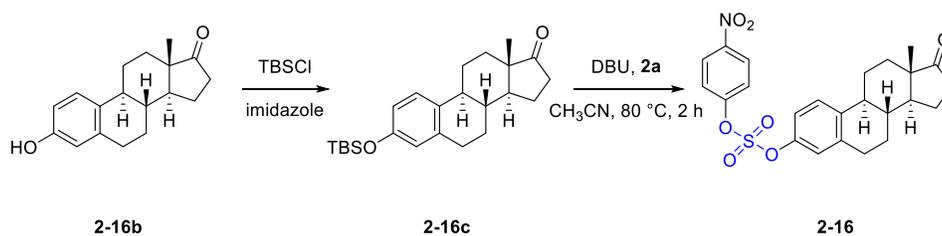
NMR (500 MHz, CDCl_3) δ 7.41 – 7.26 (m, 15H), 5.18 (d, $J = 4.0$ Hz, 1H), 5.18 (d, $J = 1.7$ Hz, 1H), 4.95-4.80 (m, 5H), 4.65 – 4.59 (m, 2H), 4.10 (d, $J = 7.4$ Hz, 1H), 3.93 (dd, $J = 10.3, 8.0$ Hz, 1H), 3.85 – 3.80 (m, 1H), 3.76 – 3.65 (m, 5H), 3.59 – 3.50 (m, 2H), 3.35 (dd, $J = 10.3, 3.8$ Hz, 1H), 0.17 (s, 9H), 0.09 (s, 9H). ^{13}C NMR (101 MHz, CDCl_3) δ 138.56, 137.67, 128.58, 128.49, 128.30, 128.08, 127.98, 127.75, 127.51, 102.10, 99.73, 82.33, 79.97, 79.40, 78.12, 76.45, 75.55, 75.47, 75.23, 74.24, 72.63, 65.76,

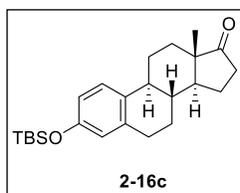
63.58, 61.41, 0.34, -0.50. HRMS (ESI) m/z $[M+NH_4]^+$ Calcd. For $C_{39}H_{57}N_4O_9Si_2$: 781.3664, found 781.3657.



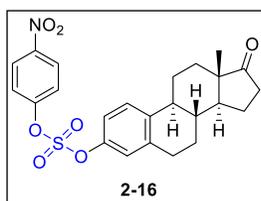
Compound **2-15** was prepared by following **General Procedure A-3** in two batches: **2-15e** (50 mg, 0.065 mmol), DBU (3.99 mg, 0.026 mmol); **2-15e** (1.32 g, 1.73 mmol), TBD (49.0 mg, 0.35 mmol). The product was purified by silica gel chromatography (hexane/ethyl acetate 4:1~2:1) to obtain **2-15** (DBU: 29%, in 2 h; TBD: 67%, in 9 h). $[\alpha]_D^{20} = +44.65$ (c 1.00 in $CHCl_3$). 1H NMR (600

MHz, $CDCl_3$) δ 7.41 – 7.26 (m, 17H), 7.20 – 7.14 (m, 2H), 6.88 – 6.83 (m, 2H), 6.83 – 6.79 (m, 2H), 5.59 (d, $J = 1.7$ Hz, 1H), 5.22 (d, $J = 3.8$ Hz, 1H), 4.94 – 4.87 (m, 3H), 4.85 (d, $J = 10.2$ Hz, 1H), 4.79 (d, $J = 10.2$ Hz, 1H), 4.66 – 4.61 (m, 1H), 4.57 (d, $J = 10.8$ Hz, 1H), 4.49 – 4.43 (m, 2H), 4.35 (dd, $J = 10.7, 6.0$ Hz, 1H), 4.09 – 4.05 (m, 1H), 3.96 – 3.87 (m, 3H), 3.81 (ddd, $J = 10.2, 6.0, 1.8$ Hz, 1H), 3.78 (s, 3H), 3.76 (s, 3H), 3.74 (dd, $J = 7.8, 4.8$ Hz, 1H), 3.46 (dd, $J = 10.2, 8.7$ Hz, 1H), 3.40 (dd, $J = 10.3, 3.7$ Hz, 1H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 158.59, 158.55, 143.61, 143.48, 137.41, 137.18, 136.76, 128.81, 128.61, 128.52, 128.44, 128.23, 128.21, 128.07, 128.01, 127.92, 122.47, 122.21, 114.83, 114.77, 99.40, 98.46, 85.67, 79.87, 79.06, 78.79, 77.17, 75.63, 75.49, 75.43, 74.02, 71.80, 69.95, 65.79, 63.24, 55.66, 55.60. HRMS (ESI) m/z $[M+NH_4]^+$ Calcd. For $C_{47}H_{53}N_4O_{17}S_2$: 1009.2847, found 1009.2816.

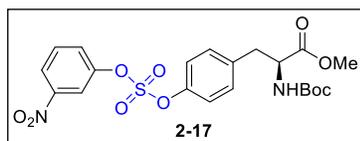
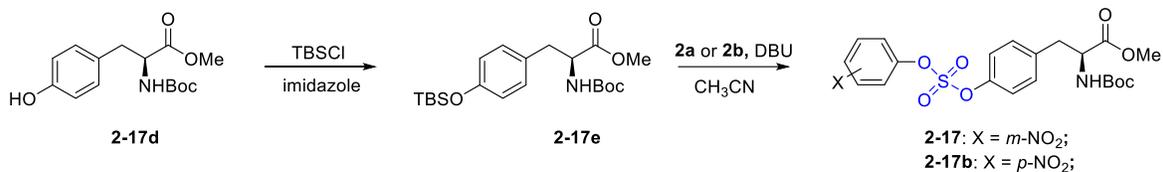




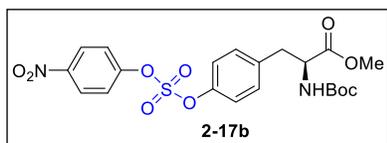
In a 25 mL flask, estrone **2-16b** (0.5 g, 1.85 mmol) and imidazole (188.9 mg, 2.77 mmol) was dissolved in DCM/DMF (5 mL/2 mL). After cooling down to 0 °C, *tert*-butyldimethylsilane chloride (278.7 mg, 1.85 mmol) was added under N₂ atmosphere. After stirred overnight, water (10 mL) was added to the solution, which was further extracted by ethyl acetate (20 mL). The organic layer was concentrated and purified by column chromatography using hexane/ethyl acetate (10: 1) as eluent, yielding the **2-16c** as a white powder (606 mg, 85%). ¹H NMR (500 MHz, CDCl₃) δ 7.12 (d, *J* = 8.5 Hz, 1H), 6.62 (dd, *J* = 8.5, 2.7 Hz, 1H), 6.57 (d, *J* = 2.6 Hz, 1H), 2.86 (dd, *J* = 10.3, 6.3 Hz, 2H), 2.50 (dd, *J* = 19.0, 8.7 Hz, 1H), 2.42 – 2.32 (m, 1H), 2.24 (td, *J* = 10.8, 4.4 Hz, 1H), 2.14 (dt, *J* = 18.6, 8.8 Hz, 1H), 2.09 – 1.91 (m, 3H), 1.68 – 1.37 (m, 6H), 0.98 (s, 9H), 0.91 (s, 3H), 0.19 (s, 6H). Spectral data matched those previously reported²⁸.



In a 25 mL flask, **2-16c** (500 mg, 1.30 mmol) and **2-2a** (316.2 mg, 1.43 mmol) were dissolved in acetonitrile (10 mL). DBU (39.6 mg, 0.26 mmol) was added to the solution under N₂ atmosphere. The mixture was heated to 80 °C and stirred for 2 h. The solvent was removed rapidly under reduced pressure, and the residue was purified by column chromatography using hexane/ethyl acetate (10/1~5/1) as eluent. The product **2-16** was obtained as a white powder (546 mg, 89%). [α]_D²⁰ = +95.44 (c 1.00 in CHCl₃). ¹H NMR (500 MHz, CDCl₃) δ 8.36 – 8.30 (m, 2H), 7.54 – 7.49 (m, 2H), 7.34 (d, *J* = 8.6 Hz, 1H), 7.10 – 7.02 (m, 2H), 2.93 (dd, *J* = 9.2, 4.3 Hz, 2H), 2.52 (ddd, *J* = 18.9, 8.7, 2.2 Hz, 1H), 2.45 – 2.36 (m, 1H), 2.30 (td, *J* = 10.6, 3.7 Hz, 1H), 2.2 - 1.93 (m, 4H), 1.70 – 1.41 (m, 6H), 0.92 (d, *J* = 1.9 Hz, 3H). ¹³C NMR (126 MHz, CDCl₃) δ 220.31, 154.32, 148.20, 146.44, 139.97, 139.26, 127.16, 125.80, 121.88, 120.80, 117.86, 50.39, 47.85, 44.11, 37.81, 35.79, 31.49, 29.42, 26.10, 25.72, 21.56, 13.80. HRMS (DART) *m/z* [M+H]⁺ Calcd. For C₂₄H₂₆NO₇S: 472.1430, found 472.1406.



2-17e was prepared according to the literature.⁶⁵ Compound **2-17** was prepared from **2-17e** (0.5 g, 1.22 mmol) by following the same procedure for the synthesis of **2-16**. The product was purified by silica gel column chromatography (hexane/ethyl acetate 4:1~1:1) to obtain **2-17** (0.57 g, 94%). ¹H NMR (500 MHz, CDCl₃) δ 8.28 – 8.20 (m, 1H), 8.15 (s, 1H), 7.65 (d, *J* = 5.6 Hz, 2H), 7.28 (d, *J* = 8.7 Hz, 2H), 7.23 (d, *J* = 8.7 Hz, 2H), 5.03 (d, *J* = 8.1 Hz, 1H), 4.61 (q, *J* = 6.8 Hz, 1H), 3.73 (s, 3H), 3.19 (dd, *J* = 14.0, 5.7 Hz, 1H), 3.06 (dd, *J* = 14.0, 6.4 Hz, 1H), 1.41 (s, 9H). ¹³C NMR (151 MHz, CDCl₃) δ 171.89, 154.96, 150.24, 149.17, 148.96, 136.54, 131.13, 130.90, 127.39, 122.52, 120.97, 116.95, 80.18, 54.24, 52.39, 37.85, 28.25. HRMS (ESI) *m/z* [M+Na]⁺ Calcd. For C₂₁H₂₄N₂O₁₀SNa: 519.1040, found 519.1045.

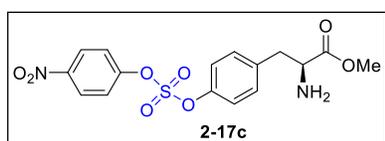
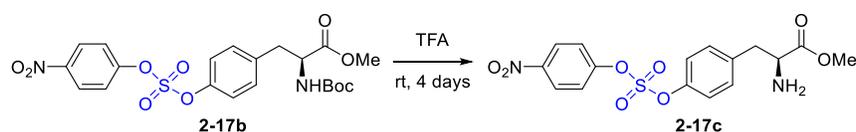


In a 25 mL flask, **2-17e** (200.0 mg, 0.49 mmol) and **2-2a** (160.0 mg, 0.72 mmol) were dissolved in acetonitrile (1.6 mL). DBU (7.4 mg, 0.05 mmol) was added to the solution under N₂ atmosphere. After 12 h, the solvent was removed rapidly under reduced pressure, and the residue was purified by column chromatography using hexane/ethyl acetate (10/1~5/1) as eluent to provide the product **2-17b** (240.2 mg, quantitative). $[\alpha]_D^{20} = +32.89$ (*c* 1.01 in CHCl₃). ¹H NMR (500 MHz, CDCl₃) δ 8.37 – 8.28 (m, 2H), 7.48 – 7.41 (m, 2H), 7.26 – 7.20 (m, 4H), 5.03 (d, *J* = 8.2 Hz, 1H), 4.59 (q, *J* = 6.8 Hz, 1H), 3.72 (s, 3H), 3.18 (dd, *J* = 13.9, 5.7 Hz, 1H), 3.04 (dd, *J* = 13.9, 6.5 Hz, 1H), 1.40 (s, 9H); ¹³C NMR (126 MHz, CDCl₃) δ 172.02, 155.11, 154.39, 149.34, 146.66, 136.74, 131.27, 126.00, 122.05, 121.12, 80.37, 54.43, 52.55, 38.07, 28.41; HRMS (ESI) *m/z* [M+Na]⁺ Calcd. For C₂₁H₂₄N₂O₁₀SNa: 519.1044, found 519.1044.

Compatibility test for compound **2-17b** and **2-17**.

Condition a (Benzylamine + Et₃N): To a solution of compound **2-17b** (23.9 mg, 0.05 mmol) in DMF (0.5 mL) was subsequently added benzylamine (10.3 mg, 10.5 μ L, 0.10 mmol) and Et₃N (4.9 mg, 6.7 μ L, 0.05 mmol) at room temperature. The resulting mixture was allowed to stir at the same temperature for 24 h. The reaction was diluted with EA, washed by water, brine, dried over Na₂SO₄, filtered, and concentrated under reduced pressure. The residue was purified on a silica gel column on Biotage, providing recovered compound **2-17b** (22.0 mg, 92% recovery).

Condition b (Benzylamine + DIPEA): To a solution of compound **2-17b** (19.5 mg, 0.04 mmol) in DMF (0.5 mL) was subsequently added benzylamine (8.4 mg, 8.6 μ L, 0.08 mmol) and DIPEA (5.1 mg, 6.8 μ L, 0.04 mmol) at room temperature. The resulting mixture was allowed to stir at the same temperature for 24 h. The reaction was diluted with EA, washed by water, brine, dried over Na₂SO₄, filtered, and concentrated under reduced pressure. The residue was purified on a silica gel column on Biotage, providing recovered compound **2-17b** (16.7 mg, 86% recovery).

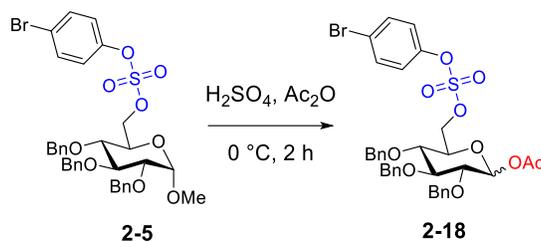


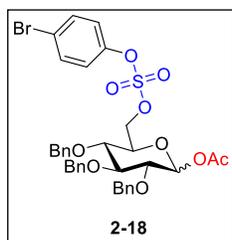
Condition c (Trifluoroacetic acid): To a 20 mL vial equipped with compound **2-17b** (18.8 mg, 0.04 mmol) and a stirring bar was added trifluoroacetic acid (TFA, 740.0 mg, 0.5 mL, 6.49 mmol) at room temperature. The resulting mixture was allowed to stir at the same temperature for 4 days. The mixture was diluted with EA, carefully washed by sat. aq. NaHCO₃, brine, dried over Na₂SO₄, filtered, and concentrated under reduced pressure. The residue was purified on a silica gel column on Biotage, providing de-Boc product **2-17c** (14.0 mg, quantitative). $[\alpha]_D^{20} = +5.63$ (c 1.00 in CHCl₃). ¹H NMR (600 MHz, CDCl₃) δ 8.34 – 8.26 (m,

2H), 7.48 – 7.45 (m, 2H), 7.31 – 7.27 (m, 2H), 7.25 – 7.19 (m, 2H), 3.73-3.71 (m, 1H), 3.70 (s, 3H), 3.08 (dd, $J = 13.7, 5.3$ Hz, 1H), 2.89 (dd, $J = 13.7, 7.7$ Hz, 1H); ^{13}C NMR (151 MHz, CDCl_3) δ 175.15, 154.34, 149.14, 146.57, 137.81, 131.19, 125.97, 125.95, 121.98, 121.09, 55.67, 52.24, 40.27.8 172.02, 155.11, 154.39, 149.34, 146.66, 136.74, 131.27, 126.00, 122.05, 121.12, 80.37, 54.43, 52.55, 38.07, 28.41; HRMS (ESI) m/z $[\text{M}+\text{H}]^+$ Calcd. For $\text{C}_{16}\text{H}_{17}\text{N}_2\text{O}_8\text{S}$: 397.0700, found 397.0702.

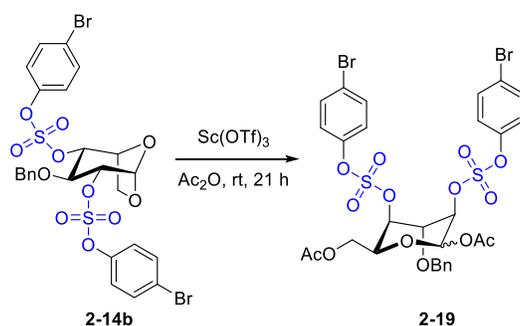
Condition d (Piperidine): To a vial equipped with compound **2-17b** (36.8 mg, 0.07 mmol) and a stirring bar was added 20% piperidine in DMF (0.5 mL) at room temperature. The resulting mixture was allowed to stir at the same temperature for 1 h. The reaction was diluted with EA, washed by water, brine, dried over Na_2SO_4 , filtered, and concentrated under reduced pressure. The residue was purified on a silica gel column on Biotage, providing recovered compound **2-17b** (19.8 mg, 54% recovery).

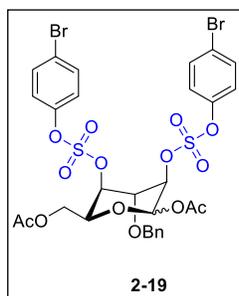
Condition e (Piperidine): To a vial equipped with compound **2-17** (37.0 mg, 0.07 mmol) and a stirring bar was added 20% piperidine in DMF (0.5 mL) at room temperature. The resulting mixture was allowed to stir at the same temperature for 1 h. The reaction was diluted with EA, washed by water (three times), brine, dried over Na_2SO_4 , filtered, and concentrated under reduced pressure. The residue was purified on a silica gel column on Biotage, providing recovered compound **2-17** (36.8 mg, quantitative recovery).





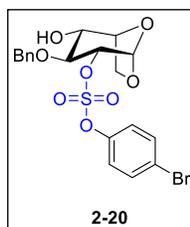
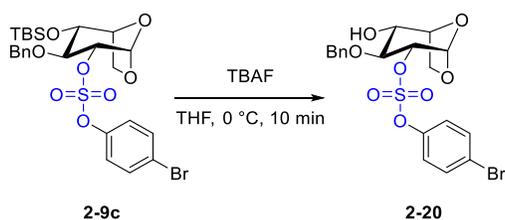
To a solution of compound **2-5** (181.6 mg, 0.26 mmol) in acetic anhydride (865.3 μL) was added sulfuric acid (5.1 mg, 2.8 μL , dissolved in 203.8 μL acetic anhydride) dropwisely at 0 $^{\circ}\text{C}$ under the atmosphere of nitrogen. The resulting mixture was allowed to stir at the same temperature for 2 hours until the NMR detection showed no residue of the starting material. The whole reaction mixture was diluted with EA, washed by sat. aq. NaHCO_3 , brine and dried over anhydrous Na_2SO_4 , filtered. The filtrate was concentrated under reduced pressure for the purification on Biotage to provide a $\alpha/\beta = 11.1:1.0$ mixture of compound **2-18** (159.9 mg, 85%). ^1H NMR (600 MHz, CDCl_3 , major α isomer) δ 7.51 – 7.44 (m, 2H), 7.36 – 7.27 (m, 13H), 7.26 – 7.22 (m, 2H), 7.18 – 7.13 (m, 2H), 6.27 (d, $J = 3.5$ Hz, 1H), 4.99 (d, $J = 10.9$ Hz, 1H), 4.91 (d, $J = 10.7$ Hz, 1H), 4.81 (d, $J = 10.9$ Hz, 1H), 4.68 (d, $J = 11.4$ Hz, 1H), 4.64 (d, $J = 11.4$ Hz, 1H), 4.58 (d, $J = 10.8$ Hz, 1H), 4.51-4.50 (t, $J = 3.0$ Hz, 2H), 4.02 – 3.93 (m, 2H), 3.62 (dd, $J = 9.6, 3.6$ Hz, 1H), 3.54 (dd, $J = 10.2, 8.9$ Hz, 1H), 2.13 (s, 3H); ^{13}C NMR (151 MHz, CDCl_3) δ 169.16, 149.13, 138.25, 137.28, 137.27, 137.21, 133.01, 132.93, 128.64, 128.56, 128.49, 128.46, 128.21, 128.14, 128.12, 128.08, 127.81, 127.79, 127.78, 127.76, 89.40, 89.39, 81.34, 78.58, 75.90, 75.66, 75.39, 73.34, 72.17, 70.74, 20.99; HRMS (ESI) m/z $[\text{M}+\text{Na}]^+$ calcd for $\text{C}_{35}\text{H}_{35}\text{O}_{10}\text{BrSNa}$ 749.1027, found 749.1029.





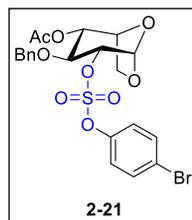
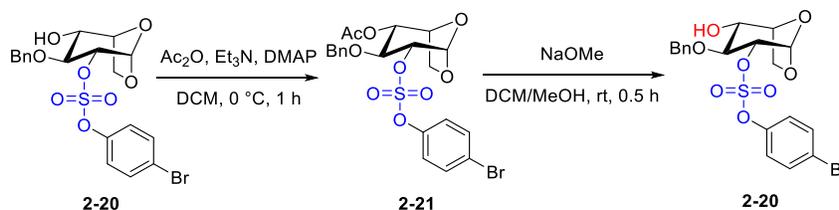
To a solution of compound **2-14b** (1.30 g) in Ac₂O (30.0 mL) was added scandium (III) triflate (Sc(OTf)₃, 578.7 mg, 1.18 mmol) at room temperature under the atmosphere of nitrogen. 21 h later, the whole reaction mixture was diluted with EA, carefully washed by sat. aq. NaHCO₃ twice (a great deal of CO₂ was released), then washed by brine, dried over anhydrous Na₂SO₄, filtered. The

filtrate was concentrated under reduced pressure for the purification on Biotage to provide a $\alpha/\beta = 0.9:1.0$ mixture of compound **2-19** (1.20 g, 81%). ¹H NMR (600 MHz, CDCl₃) δ 7.50 – 7.44 (m, 8H), 7.39 – 7.29 (m, 10H), 7.15 – 7.11 (m, 4H), 7.11 – 7.05 (m, 4H), 6.28 (dt, $J = 1.6, 0.9$ Hz, 1H), 6.08 (d, $J = 1.5$ Hz, 1H), 4.84-4.83 (m, 1H), 4.76-4.75 (m, 2H), 4.72-4.71 (m, 1H), 4.70-4.69 (m, 2H), 4.68-4.67 (m, 1H), 4.61 (td, $J = 6.5, 1.5$ Hz, 1H), 4.45-4.42 (m, 1H), 4.32 (t, $J = 3.0$ Hz, 1H), 4.28 – 4.23 (m, 5H), 4.22-4.19 (m, 1H), 2.12 (s, 3H), 2.06 (s, 3H), 2.05 (s, 6H); ¹³C NMR (151 MHz, CDCl₃) δ 170.34, 170.30, 168.43, 168.39, 149.17, 148.96, 148.95, 148.93, 135.95, 135.53, 133.30, 133.29, 133.16, 129.03, 129.01, 128.88, 128.81, 128.42, 128.06, 122.95, 122.94, 122.78, 122.75, 121.37, 121.24, 121.09, 90.11, 89.57, 75.72, 75.44, 75.30, 74.14, 73.92, 73.31, 72.96, 72.25, 71.23, 65.27, 61.62, 61.58, 20.87, 20.72, 20.67; HRMS (ESI) m/z [M+NH₄]⁺ calcd for C₂₉H₃₂O₁₄NBr₂S₂ 839.9625, found 839.9610.



To a solution of compound **2-9c** (3.47 g, 5.77 mmol) in tetrahydrofuran (THF, 57.7 mL) was added tetrabutylammonium fluoride (TBAF, 1 M solution in THF, 6.34 mL, 6.34 mmol) at 0 °C. 10 min later, TLC showed full consumption of the starting material. The whole reaction mixture was diluted with EA, washed by aq. NH₄Cl,

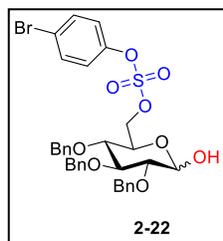
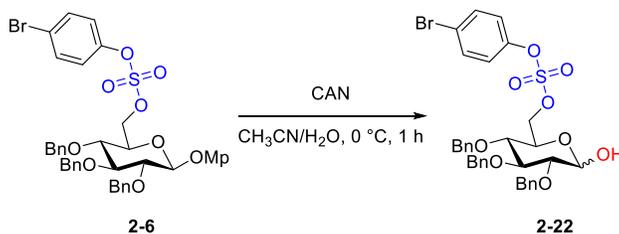
brine, dried over anhydrous Na_2SO_4 , filtered, and concentrated under reduced pressure. The filtrate was purified on a silica gel column on Biotage to provide compound **2-20** (2.40 g, 85%) as white solid. $[\alpha]_{\text{D}}^{20} = +22.0$ (c 1.01, CHCl_3); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.49 – 7.44 (m, 2H), 7.38 – 7.33 (m, 2H), 7.34 – 7.29 (m, 3H), 7.28 – 7.21 (m, 2H), 5.64 (d, $J = 1.7$ Hz, 1H), 4.82 (d, $J = 11.5$ Hz, 1H), 4.65 (dd, $J = 8.1, 1.7$ Hz, 1H), 4.62 (d, $J = 11.5$ Hz, 1H), 4.45 (t, $J = 4.6$ Hz, 1H), 4.11 (d, $J = 7.9$ Hz, 1H), 3.92 (ddd, $J = 8.1, 4.4, 1.1$ Hz, 1H), 3.76 (ddd, $J = 8.0, 4.9, 1.1$ Hz, 1H), 3.71 (t, $J = 8.1$ Hz, 1H), 2.13 (s, 1H); $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ 149.29, 137.75, 133.20, 128.88, 128.41, 128.10, 123.28, 121.20, 98.67, 86.18, 80.00, 75.26, 75.16, 71.74, 65.70; HRMS (DART) m/z $[\text{M}+\text{NH}_4]^+$ calcd for $\text{C}_{19}\text{H}_{23}\text{O}_8\text{NSBr}$ 504.0322, found 504.0318.



Acetylation: To a solution of compound **2-20** (345.0 mg, 0.71 mmol) in dichloromethane (DCM, 3.5 mL) was added triethyl amine (Et_3N , 279.4 mg, 384.8 μL , 2.76 mmol), 4-(Dimethylamino) pyridine (DMAP, 17.3 mg, 0.14 mmol), and acetic anhydride (Ac_2O , 281.9 mg, 261.0 μL , 2.76 mmol) at 0 °C under nitrogen atmosphere. 1 hour later, TLC showed full consumption of the starting material. The whole reaction mixture was diluted with EA, washed by sat. NaHCO_3 , brine, dried over anhydrous Na_2SO_4 , filtered, and concentrated under reduced pressure. The filtrate was purified on a silica gel column on Biotage to provide compound **2-21** (370.0 mg, 99%) as colorless syrup. $[\alpha]_{\text{D}}^{20} = +28.0$ (c 1.03, CHCl_3); $^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.45 – 7.40 (m, 2H), 7.35 – 7.27 (m, 3H), 7.26 – 7.24 (m, 2H), 7.23 – 7.20 (m, 2H), 5.69 (d, $J = 1.7$ Hz, 1H), 5.07 (ddd, $J = 8.5, 4.4, 1.2$ Hz, 1H), 4.72 (d, $J = 11.5$ Hz, 1H), 4.69 (dd, $J = 8.0, 1.7$ Hz, 1H), 4.63 (t, $J = 4.6$ Hz, 1H), 4.60 (d, $J = 11.5$ Hz, 1H), 4.06 (dd, $J = 8.1, 0.7$ Hz, 1H),

3.91 (t, $J = 8.3$ Hz, 1H), 3.78 (ddd, $J = 8.1, 5.0, 1.3$ Hz, 1H), 1.99 (s, 3H); ^{13}C NMR (151 MHz, CDCl_3) δ 169.67, 149.25, 137.44, 133.16, 128.61, 128.14, 127.84, 123.25, 123.24, 121.18, 98.61, 85.73, 85.72, 76.87, 75.11, 72.76, 72.70, 66.02, 20.89; HRMS (ESI) m/z $[\text{M}+\text{Na}]^+$ calcd for $\text{C}_{21}\text{H}_{21}\text{O}_9\text{NaSBr}$ 550.9982, found 550.9989.

Deacetylation: To a solution of compound **2-21** (350.2 mg, 0.66 mmol) in mixed solvent of dichloromethane (DCM, 1.5 mL) and methanol (MeOH, 4.8 mL) was added sodium methoxide (NaOMe, 11.6 mg, 0.21 mmol) at room temperature. 0.5 hour later, TLC showed full consumption of the starting material. H-form resin was used to neutralized the reaction to pH = 7, filtered, and concentrated under reduced pressure. The filtrate was purified on a silica gel column on Biotage to provide compound **2-20** (295.8 mg, 92%) as a white solid.

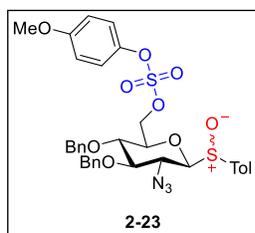
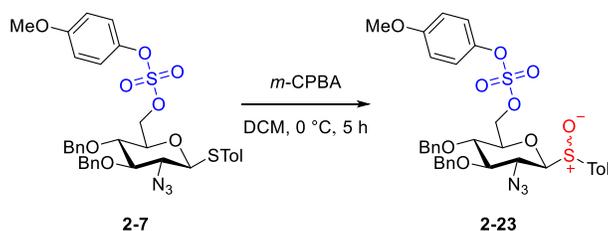


To a suspension of compound **2-6** (2.31 g, 2.92 mmol) in $\text{CH}_3\text{CN}/\text{H}_2\text{O}$ (40.9 mL/17.5 mL) was added ceric ammonium nitrate (CAN, 3.85 g, 17.51 mmol) at 0°C . 1 h later, TLC indicated no starting material left. The whole reaction mixture was poured into ice-water mixture, extracted with DCM twice. The combined

organic layer was washed by sat. aq. NaHCO_3 , brine, dried over anhydrous Na_2SO_4 , filtered. The filtrate was purified on silica gel column to provide lactol **2-22** (1.84 g, 92%) as reddish brown surup.

^1H NMR (600 MHz, CDCl_3) δ (major α -isomer) 7.46 – 7.43 (m, 2H), 7.36 – 7.29 (m, 12H), 7.28 – 7.24

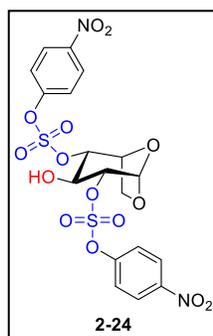
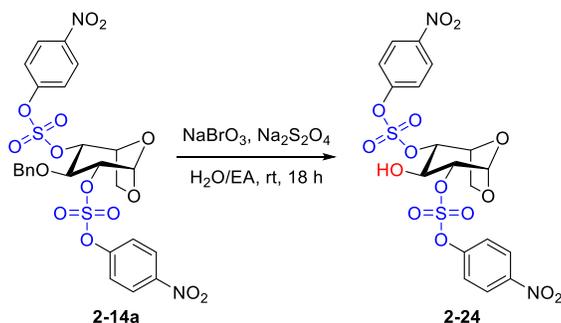
(m, 3H), 7.18 – 7.14 (m, 1H), 5.16 (d, $J = 3.5$ Hz, 1H), 5.00 – 4.97 (m, 1H), 4.91 (d, $J = 10.8$ Hz, 1H), 4.85 (d, $J = 11.0$ Hz, 1H), 4.77 (d, $J = 11.6$ Hz, 1H), 4.68 (d, $J = 11.8$ Hz, 1H), 4.60 – 4.56 (m, 1H), 4.53-4.51 (m, 2H), 4.15 (dt, $J = 10.3, 3.2$ Hz, 1H), 4.00 (t, $J = 9.2$ Hz, 1H), 3.54 (dd, $J = 9.4, 3.5$ Hz, 1H), 3.52-3.48 (m, 1H); ^{13}C NMR (151 MHz, CDCl_3) δ 149.24, 149.22, 138.43, 138.26, 138.17, 137.65, 137.46, 133.13, 133.08, 133.02, 128.79, 128.75, 128.74, 128.71, 128.60, 128.59, 128.58, 123.33, 123.31, 123.18, 121.05, 120.93, 97.48, 91.28, 84.36, 82.84, 81.49, 79.91, 76.62, 76.59, 75.81, 75.79, 75.25, 75.16, 74.92, 73.53, 73.49, 72.89, 72.70, 72.58, 68.68; HRMS (DART) m/z $[\text{M}+\text{NH}_4]^+$ calcd for $\text{C}_{33}\text{H}_{37}\text{O}_9\text{BrSN}$ 702.1367, found 702.1365.



To a solution of compound **2-7** (100.6 mg) in DCM (1.60 mL) was added *m*-CPBA (26.7 mg, 0.16 mmol) at -78 °C under the atmosphere of nitrogen. 2 hour later, the whole reaction mixture was moved to stir for another 3 h at 0 °C. The whole reaction mixture was then diluted with EA, washed by aq. $\text{Na}_2\text{S}_2\text{O}_3$,

sat. aq. NaHCO_3 , brine and dried over anhydrous Na_2SO_4 , filtered. The filtrate was concentrated under reduced pressure for the purification on Biotage to provide a 1.1:1.0 diastereomeric mixture of compound **2-23** (80.9 mg, 79%). ^1H NMR (600 MHz, CDCl_3) δ 7.55 (d, $J = 7.9$ Hz, 2H), 7.51 (d, $J = 7.9$ Hz, 2H), 7.39-7.30 (m, 14H), 7.27-7.24 (m, 3H), 7.19 – 7.17 (m, 2H), 6.95 – 6.87 (m, 1H), 4.94 (d, $J = 10.7$ Hz, 1H), 4.90 (d, $J = 10.8$ Hz, 1H), 4.87 – 4.78 (m, 1H), 4.60 (dd, $J = 10.9, 1.8$ Hz, 1H), 4.55 (dd, $J = 11.0, 4.8$ Hz, 1H), 4.42 (dd, $J = 10.8, 5.0$ Hz, 1H), 4.33 (dd, $J = 10.8, 1.9$ Hz, 1H), 4.27 (dd, J

= 10.8, 5.5 Hz, 1H), 4.17 (d, $J = 9.8$ Hz, 1H), 3.89 (t, $J = 9.8$ Hz, 1H), 3.80-3.74 (m, 7H), 3.67-3.60 (m, 3H), 3.55 – 3.39 (m, 1H), 2.42 (s, 3H), 2.38 (s, 3H); ^{13}C NMR (151 MHz, CDCl_3) δ 158.72, 158.68, 143.73, 143.67, 142.51, 142.23, 137.23, 137.17, 136.95, 136.92, 136.54, 135.47, 130.06, 129.87, 128.83, 128.79, 128.71, 128.69, 128.51, 128.49, 128.33, 128.31, 128.28, 128.26, 128.24, 128.20, 128.17, 125.52, 125.49, 124.54, 122.54, 122.48, 115.05, 115.03, 94.28, 91.40, 84.95, 84.74, 77.28, 76.32, 76.12, 76.06, 75.97, 75.35, 71.33, 71.30, 60.75, 59.82, 55.78, 55.77, 21.58, 21.57; HRMS (DART) m/z $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{34}\text{H}_{36}\text{O}_9\text{S}_2\text{N}_3$ 694.1887, found 694.1906.

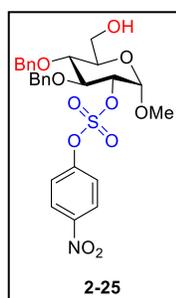
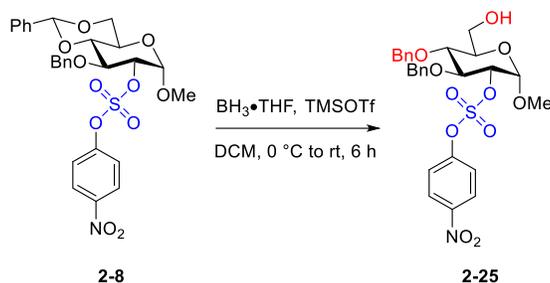


To a solution of compound **2-14a** (360.2 mg, 0.55 mmol) in EA (7.0 mL) was added sodium bromate (NaBrO_3 , 747.3 mg, 4.95 mmol, dissolved in 10.8 mL H_2O). An aqueous solution of sodium dithionite ($\text{Na}_2\text{S}_2\text{O}_4$, 766.5 mg, 4.40 mmol, dissolved in 18.7 mL H_2O) was added dropwisely via an injection pump in 18 min at room temperature. The resulting reaction mixture was allowed to stir for 18 h. TLC

showed a full conversion of the starting material. The reaction was diluted with EA, washed by aq. $\text{Na}_2\text{S}_2\text{O}_3$ four times, brine, dried over anhydrous Na_2SO_4 , filtered, and concentrated under reduced pressure. The filtrate was purified by a silica gel column on Biotage to provide compound **2-24** (310.0 mg, quantitative). $[\alpha]_{\text{D}}^{20} = +14.4$ (c 1.00, CH_3COCH_3); ^1H NMR (600 MHz, CD_3COCD_3) δ 8.43 – 8.37 (m, 4H), 7.83 – 7.74 (m, 4H), 5.79 (d, $J = 1.8$ Hz, 1H), 5.08 – 4.99 (m, 2H), 4.29 – 4.26 (m, 2H), 3.93 (ddd, $J = 8.7, 4.6, 1.5$ Hz, 1H), 2.83 (s, 1H); ^{13}C NMR (151 MHz, CD_3COCD_3) δ 155.15, 155.12,

147.64, 147.61, 126.68, 126.64, 123.49, 123.39, 99.44, 87.65, 84.17, 73.60, 69.40, 66.57; HRMS (ESI)

m/z $[M+Na]^+$ calcd for $C_{18}H_{16}O_{15}N_2S_2Na$ 586.9884, found 586.9870.



To a solution of compound **2-8** (215.5 mg, 0.38 mmol) in DCM (3.8 mL) was added borane tetrahydrofuran complex solution ($BH_3 \cdot THF$, 1 M in THF, 2.3 mL, 2.3 mmol) and trimethylsilyl trifluoromethanesulfonate (TMSOTf, 25.1 mg, 21.7 μL , 0.11 mmol) subsequently at 0 °C under the atmosphere of nitrogen. The resulting mixture was moved to stir at room temperature. 6 h later, TLC showed complete conversion of the starting material. Triethylamine (22.8 mg, 31.4 μL , 0.23 mmol) was added to quench the reaction.

The whole mixture was concentrated under reduced pressure and co-evaporated twice with methanol.

The residue was directly purified with silica gel column on Biotage to provide compound **2-25** (215.0

mg, quantitative) colorless syrup. $[\alpha]_D^{20} = +59.6$ (c 0.54, $CHCl_3$); 1H NMR (500 MHz, $CDCl_3$) δ 8.09 –

8.01 (m, 2H), 7.48 – 7.39 (m, 2H), 7.36 – 7.26 (m, 10H), 5.11 (d, $J = 3.6$ Hz, 1H), 4.88 (d, $J = 2.0$ Hz,

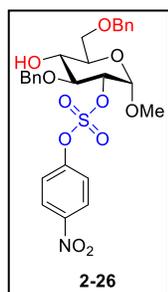
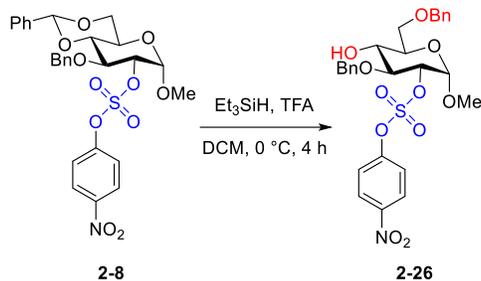
1H), 4.86 (d, $J = 2.0$ Hz, 1H), 4.75 (d, $J = 11.0$ Hz, 1H), 4.71 – 4.65 (m, 2H), 4.15 – 4.07 (m, 1H), 3.83

(ddd, $J = 11.9, 5.2, 1.6$ Hz, 1H), 3.78 – 3.67 (m, 4H), 3.38 (s, 3H), 1.62 (dd, $J = 7.9, 5.1$ Hz, 1H); ^{13}C

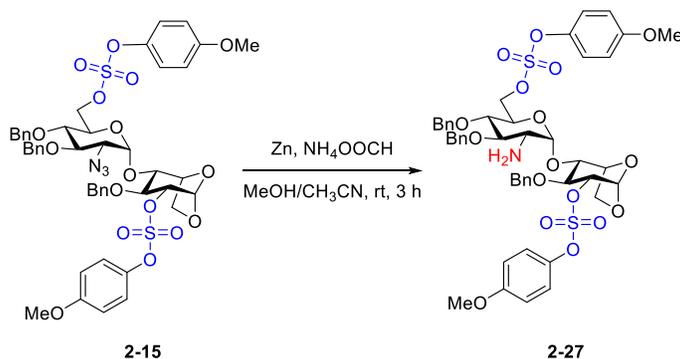
NMR (101 MHz, $CDCl_3$) δ 154.31, 146.32, 137.65, 128.74, 128.62, 128.30, 128.18, 128.10, 127.71,

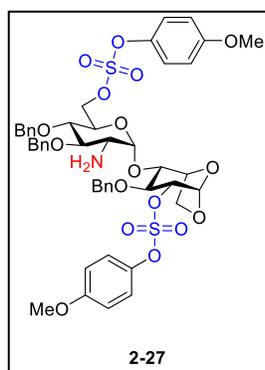
125.70, 121.89, 96.70, 83.52, 79.07, 77.65, 75.80, 75.45, 71.02, 61.38, 55.57; HRMS (DART) m/z

$[M+NH_4]^+$ calcd for $C_{27}H_{33}O_{11}N_2S$ 593.1800, found 593.1785.



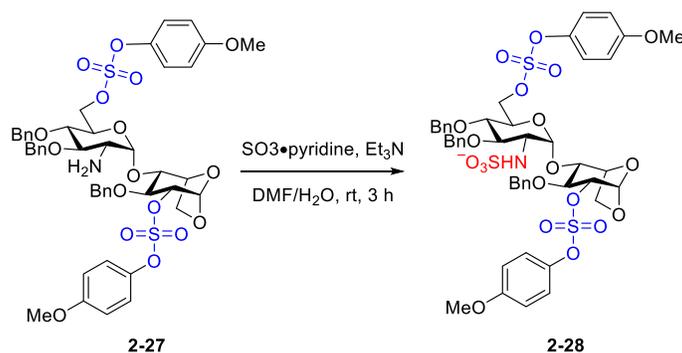
To a solution of compound **2-8** (391.5 mg, 2.46 mmol) in DCM (7.0 mL) was added triethylsilane (Et_3SiH , 400.4 mg, 550.0 μL , 3.44 mmol) and trifluoroacetic acid (TFA, 389.1 mg, 262.9 μL , 3.41 mmol) at 0 $^\circ\text{C}$ under the atmosphere of nitrogen. The resulting mixture was allowed to stir at the same temperature for 4 h. TLC showed complete conversion of the starting material. The whole reaction mixture was diluted with EA, washed by sat. aq. NaHCO_3 , brine, dried over anhydrous Na_2SO_4 , filtered, and concentrated under reduced pressure. The filtrate was purified with silica gel column on Biotage to provide compound **2-26** (374.6 mg, 95%). $[\alpha]_{\text{D}}^{20} = +56.4$ (c 1.01, CHCl_3); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 8.14 – 8.06 (m, 2H), 7.48 – 7.42 (m, 2H), 7.38 – 7.26 (m, 10H), 5.10 (d, $J = 3.6$ Hz, 1H), 4.79 – 4.71 (m, 2H), 4.69 (dd, $J = 9.7, 3.6$ Hz, 1H), 4.61 (d, $J = 12.1$ Hz, 1H), 4.55 (d, $J = 12.1$ Hz, 1H), 3.92 (dd, $J = 9.7, 7.9$ Hz, 1H), 3.82 – 3.72 (m, 3H), 3.71 – 3.65 (m, 1H), 3.39 (s, 3H); $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ 154.35, 146.35, 137.86, 137.71, 128.73, 128.66, 128.24, 128.08, 127.93, 127.89, 125.69, 121.96, 96.76, 83.30, 78.76, 75.60, 73.93, 72.32, 69.71, 69.51, 55.62; HRMS (ESI) m/z $[\text{M}+\text{Na}]^+$ calcd for $\text{C}_{27}\text{H}_{29}\text{O}_{11}\text{NSNa}$ 598.1354, found 598.1349.

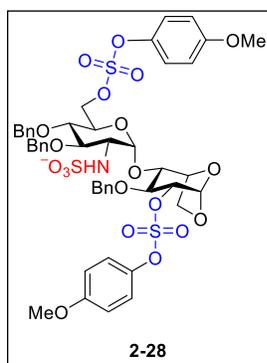




Azido reduction: To a solution of disaccharide **2-15** (789.5 mg, 0.80 mmol) in MeOH/CH₃CN (17.8 mL/3.0 mL) was consecutively added ammonium formate (NH₄OOCH, 1.99 g, 31.6 mmol) and zinc powder (Zn, 520.4 mg, 7.96 mmol) at rt. The resulting mixture was allowed to stir at the same temperature for 3 h. The reaction mixture was filtered, diluted with EA, washed by sat. aq. NaHCO₃, brine, dried over anhydrous Na₂SO₄, filtered. The filtrate was

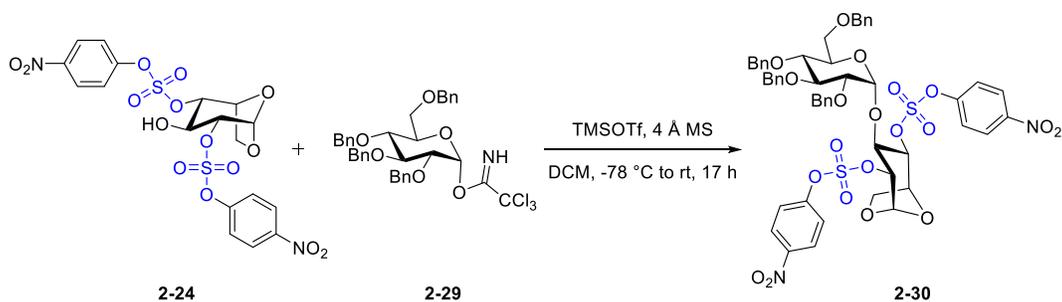
concentrated under reduced pressure for the purification with a silica gel column on Biotage to provide product **2-27** (401.4 mg, 78%). $[\alpha]_D^{20} = +42.6$ (*c* 1.02, CHCl₃); ¹H NMR (500 MHz, CDCl₃) δ 7.40 – 7.23 (m, 17H), 7.21 – 7.12 (m, 2H), 6.90 – 6.84 (m, 2H), 6.84 – 6.77 (m, 2H), 5.58 (d, *J* = 1.7 Hz, 1H), 5.04 (d, *J* = 3.6 Hz, 1H), 4.94 (d, *J* = 11.3 Hz, 1H), 4.88 (d, *J* = 10.9 Hz, 1H), 4.77 (d, *J* = 10.5 Hz, 1H), 4.73 (d, *J* = 11.2 Hz, 1H), 4.68 – 4.61 (m, 2H), 4.57 (d, *J* = 10.9 Hz, 1H), 4.53 (t, *J* = 4.6 Hz, 1H), 4.47 (dd, *J* = 10.5, 1.8 Hz, 1H), 4.34 (dd, *J* = 10.5, 6.5 Hz, 1H), 4.00 (d, *J* = 7.9 Hz, 1H), 3.86 (dd, *J* = 8.3, 4.1 Hz, 1H), 3.84 – 3.79 (m, 2H), 3.78 (s, 3H), 3.76 (s, 3H), 3.75 – 3.69 (m, 1H), 3.49 (t, *J* = 9.4 Hz, 1H), 3.36 (t, *J* = 9.4 Hz, 1H), 2.78 (dd, *J* = 10.1, 3.7 Hz, 1H), 1.38 (broad, 2H; NH₂); ¹³C NMR (151 MHz, CDCl₃) δ 158.73, 158.71, 143.76, 143.66, 138.09, 137.48, 137.18, 128.93, 128.78, 128.60, 128.56, 128.40, 128.27, 128.24, 128.13, 128.04, 122.62, 122.43, 114.96, 114.92, 102.03, 98.55, 86.02, 83.34, 79.72, 78.80, 77.85, 75.90, 75.45, 75.38, 74.17, 72.50, 70.60, 65.83, 56.01, 55.81, 55.75; HRMS (ESI) *m/z* [M+H]⁺ calcd for C₄₇H₅₂O₁₇S₂N 966.2671, found 966.2398.

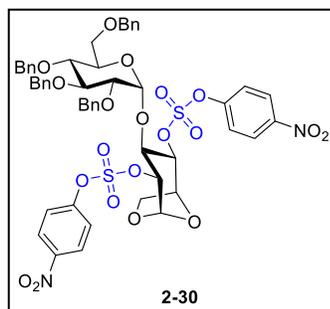




N-Sulfation: To a solution of amine **2-27** (251.3 mg, 0.26 mmol) in DMF/MeOH (1.6 mL/4.9 mL) was consecutively added Et₃N (1.83 g, 18.05 mmol, 2.52 mL) and sulfur trioxide pyridine complex (414.0 mg, 2.60 mmol) at room temperature. 3 h later, aqueous solution of NaHCO₃ (874.2 mg, 10.41 mmol, 404.7 μL, dissolved in 4.9 mL H₂O) was added the reaction. The resulting mixture was allowed to stir for another 30 min, filtered through a pad of Celite, concentrated.

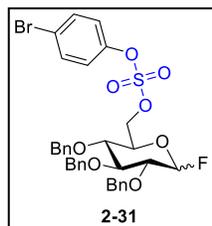
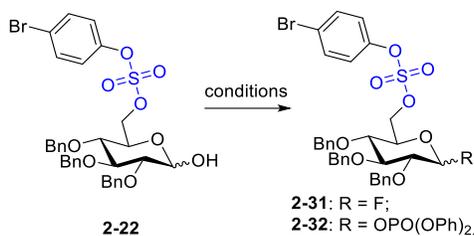
The residue was purified on a silica gel column with DCM/MeOH = 10:1 as eluent on Biotage to provide *N*-sulfated compound **2-28** (199.0 mg, 72%) as a white foam. ¹H NMR (500 MHz, CD₃OD) δ 7.48 – 7.43 (m, 2H), 7.42 – 7.36 (m, 2H), 7.35 – 7.15 (m, 15H), 6.95 – 6.90 (m, 2H), 6.85 – 6.80 (m, 2H), 5.54 (d, *J* = 1.7 Hz, 1H), 5.51 (d, *J* = 3.5 Hz, 1H), 4.98 (d, *J* = 11.0 Hz, 1H), 4.90 (d, *J* = 11.2 Hz, 1H), 4.85 (d, *J* = 11.0 Hz, 1H), 4.69 (d, *J* = 7.3 Hz, 1H), 4.67 (d, *J* = 7.1 Hz, 1H), 4.62 – 4.56 (m, 2H), 4.54 (d, *J* = 11.1 Hz, 1H), 4.50 (dd, *J* = 10.6, 1.8 Hz, 1H), 4.39 (dd, *J* = 10.6, 7.0 Hz, 1H), 4.07 (ddd, *J* = 8.3, 4.2, 1.1 Hz, 1H), 3.93 (d, *J* = 8.1 Hz, 1H), 3.90 – 3.82 (m, 2H), 3.76 (s, 3H), 3.74 (s, 3H), 3.61–3.57 (m, 1H), 3.54–3.51 (m, 2H), 3.43 (dd, *J* = 10.2, 8.4 Hz, 1H); ¹³C NMR (126 MHz, CD₃OD) δ 160.27, 160.16, 145.13, 145.01, 139.89, 139.25, 139.14, 129.70, 129.58, 129.54, 129.42, 129.37, 129.28, 129.24, 129.02, 128.70, 128.69, 128.64, 123.66, 123.63, 115.93, 115.79, 100.28, 99.67, 87.18, 81.14, 80.87, 78.59, 78.45, 76.09, 76.06, 75.70, 75.08, 74.24, 71.41, 66.63, 59.50, 56.26, 56.14; HRMS (ESI) *m/z* [M-H]⁻ calcd for C₄₇H₅₀O₂₀S₃N 1044.2094, found 1044.2010.



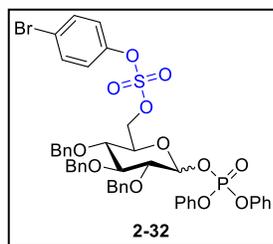


Glycosylation with trichloroacetimidates (Schmidt) donor.

A mixture of acceptor **2-24** (102.9 mg, 0.18 mmol) and Schmidt donor **2-29**⁶⁶⁻⁶⁷ (153.5 mg, 0.22 mmol) was azeotropically dried with toluene three times. To a solution of above mixture in DCM (2.0 mL) was added freshly activated 4 Å MS (200 mg, 100 mg per 1 mL solvent) under the atmosphere of nitrogen. The resulting mixture was allowed to stir at room temperature for 30 min. Then the reaction was cooled down to -78 °C, followed by the addition of TMSOTf (4.6 mg, 4.0 μL, 0.02 mmol). The reaction was allowed to stir at the same temperature for 2 h, then was moved to stir at room temperature for another 15 h. Et₃N (3.6 mg, 0.04 mmol, 5.0 μL) was added to quench the reaction. The whole reaction mixture was filtered through a pad of Celite, diluted with EA, washed by phosphate buffer, brine, dried over anhydrous Na₂SO₄, filtered. The filtrate was concentrated under reduced pressure for the purification on Biotage to provide compound **2-30** (129.0 mg, 65%). $[\alpha]_D^{20} = +22.0$ (*c* 1.01, CHCl₃); ¹H NMR (600 MHz, CDCl₃) δ 8.27 – 8.20 (m, 2H), 7.98 – 7.93 (m, 2H), 7.49 – 7.44 (m, 2H), 7.43 – 7.39 (m, 2H), 7.35 – 7.23 (m, 16H), 7.22-7.20 (m, 2H), 7.12 – 7.05 (m, 2H), 5.82 (d, *J* = 1.8 Hz, 1H), 5.37 (d, *J* = 3.6 Hz, 1H), 5.16 (ddd, *J* = 8.0, 4.5, 1.1 Hz, 1H), 4.98 (t, *J* = 4.7 Hz, 1H), 4.80 (dd, *J* = 8.0, 1.8 Hz, 1H), 4.78 – 4.72 (m, 3H), 4.69 – 4.58 (m, 3H), 4.46 (dd, *J* = 11.2, 5.0 Hz, 2H), 4.38 (t, *J* = 7.9 Hz, 1H), 4.17 (d, *J* = 8.5 Hz, 1H), 4.03 (dt, *J* = 10.1, 2.7 Hz, 1H), 3.93 – 3.86 (m, 2H), 3.79 (dd, *J* = 10.9, 3.2 Hz, 1H), 3.74 – 3.67 (m, 2H), 3.61 (dd, *J* = 9.9, 3.7 Hz, 1H); ¹³C NMR (151 MHz, CDCl₃) δ 153.49, 153.46, 146.84, 146.72, 138.56, 138.15, 137.95, 137.71, 128.60, 128.57, 128.52, 128.44, 128.26, 128.19, 128.08, 128.06, 127.95, 127.77, 127.72, 126.09, 125.78, 122.64, 122.13, 122.08, 98.41, 97.31, 84.78, 83.35, 81.55, 79.35, 75.71, 75.29, 73.68, 73.66, 72.48, 72.08, 71.54, 68.09, 65.77; HRMS (ESI) *m/z* [M+Na]⁺ calcd for C₅₂H₅₀O₂₀N₂S₂Na 1109.2291, found 1109.2271.



To a solution of above lactol **2-22** (596.4 mg, 0.87 mmol) in DCM (13.4 mL) was added (diethylamino)sulfur trifluoride (DAST, 1.40 g, 1.15 mL, 8.70 mmol) at -42 °C under the atmosphere of nitrogen. The whole reaction mixture was moved to stir at 0 °C for 1.5 h. The reaction mixture was cooled down back to -42 °C and sat. aq. NaHCO₃ was added slowly to quench the reaction. The mixture was poured into a two-layer funnel with EA/sat. aq. NaHCO₃. The organic layer was washed by sat. aq. NaHCO₃ twice, brine, dried over anhydrous Na₂SO₄, filtered. The filtrate was concentrated under reduced pressure for the purification on Biotage to provide $\alpha/\beta = 1.0:2.6$ mixture of fluoride donor **2-31** (374.0 mg, 63%) as white solid. Pure β -isomer was obtained from preparative thin layer chromatography: $[\alpha]_{\text{D}}^{20} = +25.5$ (c 0.68, CHCl₃); ¹H NMR (500 MHz, CDCl₃) δ 7.52 – 7.46 (m, 2H), 7.38 – 7.26 (m, 12H), 5.29 (ddd, $J = 52.6, 6.3, 0.9$ Hz, 1H), 4.90-4.85 (m, 2H), 4.82 (d, $J = 11.1$ Hz, 1H), 4.76 (d, $J = 11.1$ Hz, 1H), 4.69 (d, $J = 11.2$ Hz, 1H), 4.61 – 4.53 (m, 2H), 4.47 (dd, $J = 10.6, 4.8$ Hz, 1H), 3.80 – 3.74 (m, 1H), 3.73-3.70 (m, 1H), 3.67 – 3.53 (m, 2H); ¹³C NMR (126 MHz, CDCl₃) δ 149.31, 137.99, 137.46, 137.30, 133.20, 133.16, 128.81, 128.69, 128.65, 128.42, 128.31, 128.29, 128.27, 128.05, 127.96, 123.38, 121.13, 109.25 (d, $J = 218.5$ Hz), 83.20 (d, $J = 9.6$ Hz), 80.90 (d, $J = 22.9$ Hz), 75.58, 75.26, 75.19, 74.41 (d, $J = 2.1$ Hz), 72.55 (d, $J = 5.1$ Hz), 72.09; ¹⁹F NMR (470 MHz, CDCl₃) δ -136.41 (dd, $J = 52.7, 10.8$ Hz); HRMS (ESI) m/z [M+Na]⁺ calcd for C₃₃H₃₂O₈BrFSNa 709.0878, found 709.0892.



Lactol **2-22** (136.2 mg, 0.20 mmol) was azeotropically dried with toluene twice.

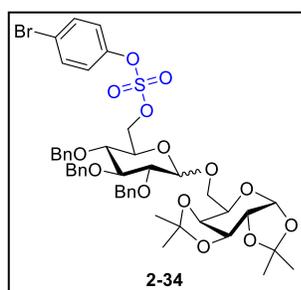
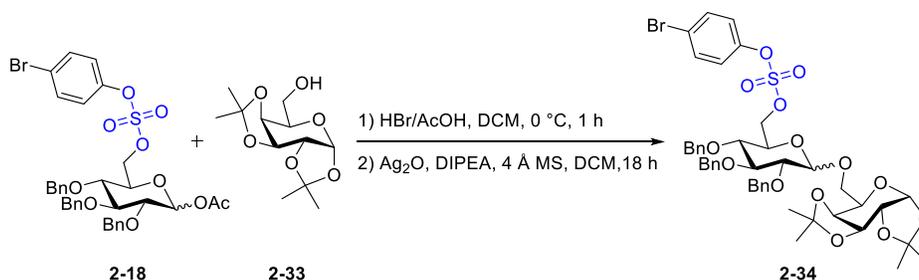
To a solution of above compound in DCM (2.0 mL) was consecutively added diphenyl chlorophosphate (94.0 mg, 0.40 mmol) and 4-(dimethylamino)pyridine (DMAP, 63.4 mg, 0.52 mmol) at -18 °C. The

resulting mixture was allowed to stir at 0 °C for 3 h. The mixture was then diluted with EA, washed by sat. aq. NaHCO₃ twice, brine, dried over anhydrous Na₂SO₄ and filtered. The filtrate was concentrated under reduced pressure for the purification on Biotage to provide $\alpha/\beta = 4.5:1.0$ mixture of phosphate donor **2-32** (81.9 mg, 45%) as white solid.

2-32 α : $[\alpha]_D^{20} = +49.2$ (*c* 0.85, CHCl₃); ¹H NMR (500 MHz, CDCl₃) δ 7.51 – 7.43 (m, 2H), 7.39 – 7.28 (m, 15H), 7.28 – 7.21 (m, 8H), 7.20-7.15 (m, 2H), 7.13 – 7.08 (m, 2H), 5.98 (dd, *J* = 6.7, 3.2 Hz, 1H), 4.95 (d, *J* = 11.0 Hz, 1H), 4.89 (d, *J* = 10.8 Hz, 1H), 4.79 (d, *J* = 11.0 Hz, 1H), 4.75 (d, *J* = 11.4 Hz, 1H), 4.65 (d, *J* = 11.4 Hz, 1H), 4.55 (d, *J* = 10.8 Hz, 1H), 4.40 (dd, *J* = 10.7, 3.8 Hz, 1H), 4.18 (dd, *J* = 10.7, 1.9 Hz, 1H), 3.93 (t, *J* = 9.3 Hz, 1H), 3.88 (ddd, *J* = 10.2, 3.8, 1.9 Hz, 1H), 3.59 (dt, *J* = 9.6, 3.3 Hz, 1H), 3.52 (dd, *J* = 10.3, 9.0 Hz, 1H); ¹³C NMR (126 MHz, CDCl₃) δ 150.57 (d, *J* = 7.7 Hz), 150.45 (d, *J* = 7.1 Hz), 149.22, 133.18, 130.01, 129.82, 128.78, 128.67, 128.64, 128.36, 128.24, 128.03, 127.99, 125.70 (d, *J* = 7.5 Hz), 123.16, 120.56 (d, *J* = 4.7 Hz), 120.24 (d, *J* = 4.8 Hz), 96.37 (d, *J* = 6.2 Hz), 80.85, 78.96 (d, *J* = 7.3 Hz), 75.85, 75.72, 75.54, 73.38, 71.72, 70.79; ³¹P NMR (202 MHz, CDCl₃) δ -13.22; HRMS (ESI) *m/z* [M+Na]⁺ calcd for C₄₅H₄₂O₁₂BrPSNa 939.1210, found 939.1231.

2-32 β : $[\alpha]_D^{20} = +5.8$ (*c* 1.02, CHCl₃); ¹H NMR (500 MHz, CDCl₃) δ 7.43 – 7.37 (m, 2H), 7.36 – 7.09 (m, 27H), 5.45 (t, *J* = 7.2 Hz, 1H), 4.88 (d, *J* = 7.0 Hz, 1H), 4.86 (d, *J* = 6.9 Hz, 1H), 4.77 (d, *J* = 11.0 Hz, 1H), 4.71 (d, *J* = 10.9 Hz, 1H), 4.64 (d, *J* = 11.0 Hz, 1H), 4.57 (d, *J* = 10.9 Hz, 1H), 4.54 (dd, *J* = 10.6, 1.9 Hz, 1H), 4.48 (dd, *J* = 10.7, 4.0 Hz, 1H), 3.75 – 3.69 (m, 2H), 3.65 – 3.55 (m, 2H); ¹³C NMR (151 MHz, CDCl₃) δ 150.47 (d, *J* = 7.4 Hz), 150.41 (d, *J* = 6.9 Hz), 149.16, 138.00, 137.54, 137.29,

133.21, 130.00, 129.96, 128.80, 128.61, 128.50, 128.40, 128.28, 128.09, 127.98, 127.87, 125.79, 125.70, 99.53 (d, $J = 6.5$ Hz), 84.04 (d, $J = 2.3$ Hz), 81.41 (d, $J = 9.0$ Hz), 75.82, 75.70, 75.32, 75.04, 73.31, 71.74; ^{31}P NMR (202 MHz, CDCl_3) δ -13.41; HRMS (ESI) m/z $[\text{M}+\text{NH}_4]^+$ calcd for $\text{C}_{45}\text{H}_{46}\text{O}_{12}\text{NBrPS}$ 934.1656, found 934.1646.



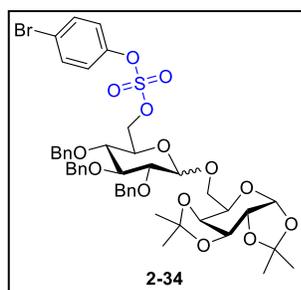
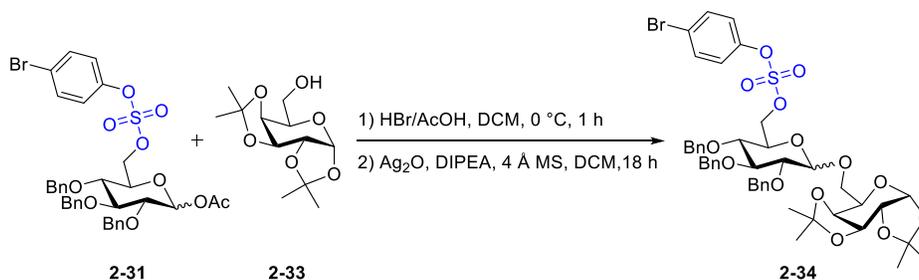
Glycosylation with bromide donor.

1) Bromide donor preparation. To a solution of compound **2-18** (52.7 mg, 0.07 mmol) in DCM (724.3 μL) was added hydrogen bromide solution 33% wt in acetic acid (HBr/AcOH, 108.9 μL) at 0 °C under the atmosphere of nitrogen. 1 h later, the reaction was quenched by dropwise addition of sat. aq. NaHCO_3 at the same temperature. The whole reaction mixture was poured into a two-layer separatory funnel with EA/sat. aq. NaHCO_3 , washed by sat. aq. NaHCO_3 twice, brine, dried over anhydrous Na_2SO_4 , filtered. The filtrate was concentrated under reduced pressure for direct usage in the next step.

2) Glycosylation. A mixture of above newly prepared bromide donor and acceptor **2-33** (25.8 mg, 0.10 mmol) was azeotropically dried over toluene three times. To a solution of above mixture in DCM (2.1 mL) was added freshly activated 4 Å MS (210 mg, 100 mg per mL of solvent). After being stirred for 30 min, the whole reaction was cooled down to 0 °C, followed by consecutive addition of *N,N*-diisopropylethylamine (DIPEA, 14.0 mg, 18.9 μL , 0.12 mmol). The resulting mixture was

allowed to warm up to rt slowly. 18 h later, TLC showed full conversion of the bromide donor. The reaction was filtered through a pad of Celite, concentrated directly under reduced pressure for the purification with a silica gel column on Biotage to solely provide only α product **2-34 α** (55.9 mg, 83%).

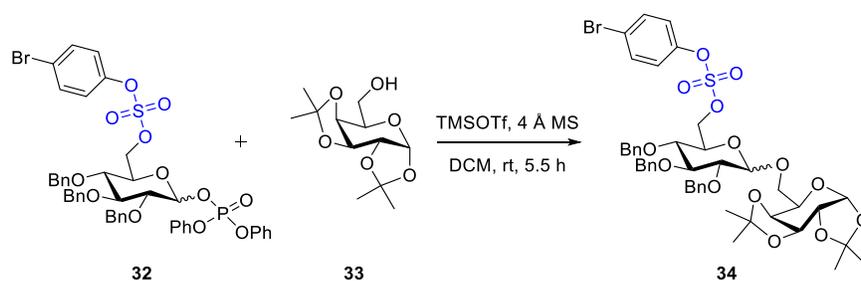
2-34 α : $[\alpha]_D^{20} = +3.6$ (*c* 1.02, CHCl₃); ¹H NMR (400 MHz, CDCl₃) δ 7.50 – 7.42 (m, 2H), 7.39 – 7.22 (m, 15H), 7.20 – 7.13 (m, 2H), 5.50 (d, *J* = 5.0 Hz, 1H), 4.92 – 4.86 (m, 2H), 4.79 (d, *J* = 10.7 Hz, 1H), 4.72 (s, 2H), 4.65 – 4.50 (m, 4H), 4.35 – 4.25 (m, 2H), 4.06 - 3.98 (m, 3H), 3.86 – 3.70 (m, 2H), 3.55 – 3.44 (m, 2H), 1.52 (s, 3H), 1.45 (s, 3H), 1.32 (s, 3H), 1.31 (s, 3H); ¹³C NMR (126 MHz, CDCl₃) δ 149.36, 138.75, 138.18, 137.97, 133.07, 128.65, 128.64, 128.62, 128.54, 128.07, 128.05, 128.03, 128.01, 127.77, 123.23, 123.21, 120.88, 109.52, 108.82, 97.31, 96.45, 81.76, 79.79, 76.73, 75.73, 74.99, 73.03, 72.84, 71.17, 70.85, 70.78, 70.71, 68.50, 67.70, 66.44, 26.27, 26.21, 25.09, 24.74; HRMS (ESI) *m/z* [M+NH₄]⁺ calcd for C₄₅H₅₅O₁₄BrSN 944.2521, found 944.2512.

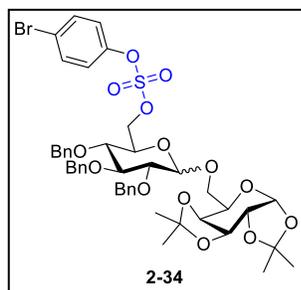


Glycosylation with fluoride donor **2-31**.

A mixture of fluoride donor **2-31** (154.3 mg, 0.22 mmol) and galactose acceptor **2-33** (70.1 mg, 0.27 mmol) was azeotropically dried over toluene three times. To above mixture in DCM (6.6 mL) was added freshly activated 4 Å MS (660 mg, 100 mg per mL of solvent) and zirconocene dichloride (Cp₂ZrCl₂, 98.4 mg, 0.34 mmol) at room temperature under the atmosphere of nitrogen. After stirring for 30 min, silver perchlorate (AgClO₄, 93.1 mg, 0.45 mmol) was added to the reaction. The whole reaction mixture was

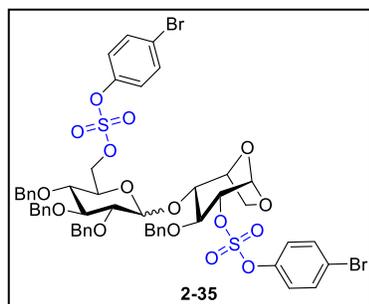
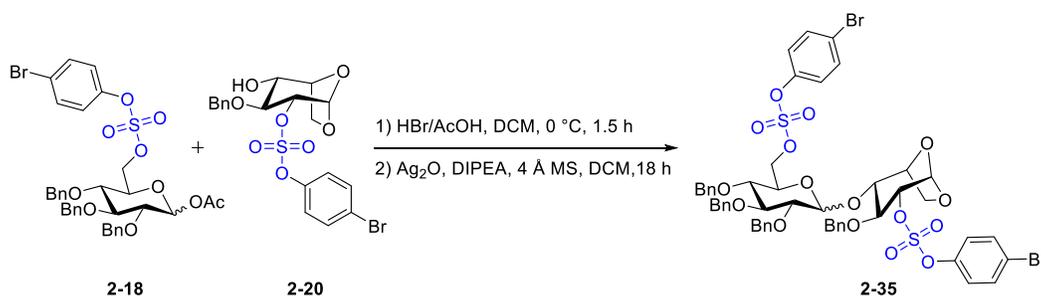
allowed to stir at the same temperature for another 19 h. The whole reaction mixture was filtered through a pad of Celite, concentrated directly for purification with a silica gel column on Biotage to provide $\alpha/\beta = 6.3:1.0$ mixture of product **2-34** (152.7 mg, 73%). **2-34 β** : $[\alpha]_{\text{D}}^{20} = -4.3$ (c 0.49, CHCl_3); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.53 – 7.48 (m, 2H), 7.42-7.40 (m, 2H), 7.35 – 7.26 (m, 11H), 7.24-7.21 (m, 4H), 5.57 (d, $J = 5.0$ Hz, 1H), 5.05 (d, $J = 11.1$ Hz, 1H), 4.98 (d, $J = 11.0$ Hz, 1H), 4.87 (d, $J = 11.0$ Hz, 1H), 4.77 (d, $J = 11.0$ Hz, 1H), 4.71 (d, $J = 11.1$ Hz, 1H), 4.61 (dd, $J = 7.9, 2.4$ Hz, 1H), 4.57 (dd, $J = 10.5, 1.8$ Hz, 1H), 4.54 (d, $J = 10.9$ Hz, 1H), 4.50 (d, $J = 7.8$ Hz, 1H), 4.43 (dd, $J = 10.5, 5.2$ Hz, 1H), 4.32 (dd, $J = 5.0, 2.4$ Hz, 1H), 4.22 (dd, $J = 8.0, 1.6$ Hz, 1H), 4.11 – 4.03 (m, 2H), 3.75 – 3.63 (m, 2H), 3.61 – 3.55 (m, 1H), 3.48-3.41 (m, 2H), 1.50 (s, 3H), 1.45 (s, 3H), 1.33 (s, 3H), 1.31 (s, 3H); $^{13}\text{C NMR}$ (126 MHz, CDCl_3) δ 149.42, 138.55, 138.50, 137.58, 133.16, 128.80, 128.76, 128.55, 128.42, 128.30, 127.97, 127.83, 127.78, 123.35, 121.01, 109.67, 108.76, 104.39, 96.54, 84.33, 81.46, 77.36, 76.57, 75.79, 75.24, 74.55, 72.77, 72.52, 71.50, 70.93, 70.62, 70.05, 67.50, 26.23, 26.15, 25.13, 24.60; HRMS (ESI) m/z $[M+\text{NH}_4]^+$ calcd for $\text{C}_{45}\text{H}_{55}\text{O}_{14}\text{BrSN}$ 944.2521, found 944.2503.





Glycosylation with phosphate donor 2-32.

A mixture of phosphate donor **2-32** (55.2 mg, 0.06 mmol) and acceptor **2-33** (18.7 mg, 0.07 mmol) was azeotropically dried over toluene three times. To a solution of above mixture in DCM (1.8 mL) was added freshly activated 4 Å MS (180 mg, 100 mg per mL of solvent). After being stirred for 30 min, the whole reaction was cooled down to -42 °C, followed by the addition of TMSOTf (14.7 mg, 12.7 μL, 0.07 mmol). 2 h later, the reaction was moved to stir at 0 °C for another 3.5 h until NMR showed no residue of the phosphate donor left. The reaction was quenched with Et₃N (6.1 mg, 8.4 μL, 0.06 mmol) and filtered through a pad of Celite, diluted with EA, washed by sat. aq. NaHCO₃, brine, dried over anhydrous Na₂SO₄, filtered. The filtrate was concentrated under reduced pressure for the purification with a silica gel column on Biotage to provide α/β = 3.3:1.0 mixture of product **2-34** (51.6 mg, 92%).



Glycosylation with bromide donor.

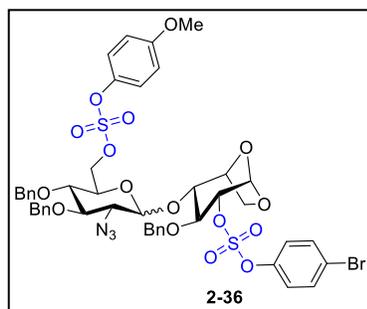
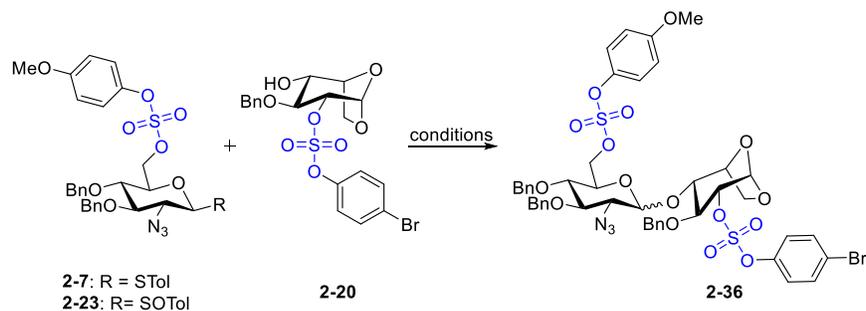
1) Bromide donor preparation. To a solution of compound **2-18** (94.1 mg, 0.13 mmol) in DCM (1.1 mL) was added hydrogen bromide solution 33% wt in acetic acid (HBr/AcOH, 194.5 μL) at 0 °C under the atmosphere of nitrogen. 1.5 h later, the reaction was quenched by dropwise addition of sat. aq. NaHCO₃ at the same temperature. The whole reaction mixture was poured into a two-layer separatory

funnel with EA/sat. aq. NaHCO₃, washed by sat. aq. NaHCO₃ twice, brine, dried over anhydrous Na₂SO₄, filtered. The filtrate was concentrated under reduced pressure for direct usage in the next step.

2) Glycosylation. A mixture of above newly prepared bromide donor and acceptor **2-20** (80.0 mg, 0.16 mmol) was azeotropically dried over toluene three times. To a solution of above mixture in DCM (2.6 mL) was added freshly activated 4 Å MS (260 mg, 100 mg per mL of solvent). After being stirred for 30 min, followed by consecutive addition of DIPEA (25.1 mg, 33.7 μL, 0.19 mmol) at room temperature. The resulting mixture was allowed to stir at the same temperature for 18 h. The reaction was filtered through a pad of Celite, concentrated directly under reduced pressure for the purification with a silica gel column on Biotage to provide product **2-35α** (53.0 mg, 36%) and **2-35β** (22.6 mg, 15%).

2-35α: $[\alpha]_D^{20} = +40.8$ (*c* 0.33, CHCl₃); ¹H NMR (500 MHz, CDCl₃) δ 7.50 – 7.45 (m, 2H), 7.41 – 7.37 (m, 2H), 7.36 – 7.29 (m, 8H), 7.28 – 7.23 (m, 10H), 7.22 – 7.19 (m, 2H), 7.19 – 7.15 (m, 2H), 7.13 – 7.08 (m, 2H), 5.61 (d, *J* = 1.7 Hz, 1H), 5.00 (dd, *J* = 11.1, 1.9 Hz, 1H), 4.92 (d, *J* = 11.0 Hz, 1H), 4.88 (d, *J* = 3.7 Hz, 1H), 4.86 (d, *J* = 10.9 Hz, 1H), 4.72 (d, *J* = 11.2 Hz, 1H), 4.68 – 4.61 (m, 2H), 4.60 – 4.49 (m, 3H), 4.41 (dd, *J* = 10.5, 1.8 Hz, 1H), 4.27 (dd, *J* = 10.5, 6.6 Hz, 1H), 4.03 (d, *J* = 7.9 Hz, 1H), 3.97 (dd, *J* = 9.9, 8.8 Hz, 1H), 3.86 – 3.80 (m, 1H), 3.76 (ddd, *J* = 10.2, 6.6, 1.8 Hz, 1H), 3.69 (dd, *J* = 7.9, 5.0 Hz, 1H), 3.47 (dd, *J* = 9.8, 3.7 Hz, 1H), 3.35 (dd, *J* = 10.2, 8.8 Hz, 1H); ¹³C NMR (126 MHz, CDCl₃) δ 149.29, 149.10, 138.18, 137.79, 137.60, 137.27, 133.25, 133.12, 128.90, 128.74, 128.68, 128.65, 128.55, 128.47, 128.45, 128.41, 128.29, 128.21, 128.12, 128.09, 127.85, 127.77, 123.28, 123.04, 121.20, 121.10, 99.72, 98.42, 85.70, 81.94, 81.28, 79.12, 77.56, 77.36, 76.77, 75.92, 75.53, 75.11, 74.15, 73.65, 72.93, 70.01, 65.77; HRMS (ESI) *m/z* [M+NH₄]⁺ calcd for C₅₂H₅₄O₁₆ Br₂S₂N 1172.1225, found 1172.1197.

2-35 β : $[\alpha]_{\text{D}}^{20} = +26.3$ (*c* 0.53, CHCl₃); ¹H NMR (600 MHz, CDCl₃) δ 7.47 – 7.40 (m, 2H), 7.37-7.23 (m, 22H), 7.26 – 7.23 (m, 2H), 7.19 – 7.15 (m, 2H), 7.11 – 7.02 (m, 2H), 5.63 (d, *J* = 1.8 Hz, 1H), 4.93 – 4.80 (m, 6H), 4.73 (d, *J* = 11.2 Hz, 1H), 4.66 (d, *J* = 11.2 Hz, 1H), 4.61 – 4.56 (m, 2H), 4.52 – 4.47 (m, 3H), 4.38 (dd, *J* = 10.7, 5.0 Hz, 1H), 4.15 (dd, *J* = 8.2, 4.2 Hz, 1H), 3.97 (d, *J* = 8.0 Hz, 1H), 3.79 (t, *J* = 8.1 Hz, 1H), 3.68 (t, *J* = 8.9 Hz, 1H), 3.60 – 3.49 (m, 3H), 3.40 (dd, *J* = 9.2, 7.8 Hz, 1H); ¹³C NMR (151 MHz, CDCl₃) δ 149.12, 148.92, 137.84, 137.62, 137.54, 137.14, 133.10, 132.94, 128.67, 128.59, 128.52, 128.32, 128.28, 128.14, 128.02, 127.90, 127.88, 127.81, 127.71, 127.32, 123.14, 123.12, 122.87, 120.94, 120.91, 101.24, 98.40, 85.34, 84.55, 81.91, 77.65, 76.42, 75.75, 75.24, 75.14, 74.84, 73.00, 72.38, 72.20, 65.52; HRMS (ESI) *m/z* [M+Na]⁺ calcd for C₅₂H₅₀O₁₆ Br₂S₂Na 1175.0799, found 1175.0852.



Glycosylation with thioglycoside 2-7. A mixture of thioglycoside **2-7** (165.2 mg, 0.24 mmol) and acceptor **2-20** (130.7 mg, 0.27 mmol) and was azeotropically dried with toluene three times. To a solution of above mixture in DCM (2.4 mL) was added freshly activated 4 Å MS (240 mg, 100 mg per 1 mL solvent) under the atmosphere of nitrogen.

The resulting mixture was allowed to stir at room temperature for 30 min. Then the reaction was

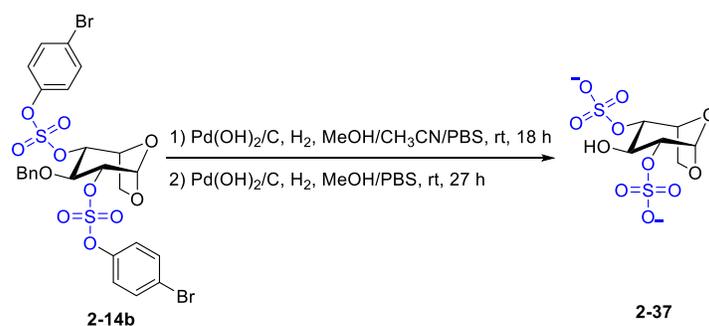
cooled down to $-42\text{ }^{\circ}\text{C}$, followed by the consecutive addition of *N*-iodosuccinimide (NIS, 109.7 mg, 0.49 mmol) and trifluoromethanesulfonic acid (TfOH, 7.3 mg, 4.3 μL , 0.05 mmol). The reaction was allowed to stir at the same temperature for 1.5 h, then Et_3N (14.5 mg, 20.0 μL , 0.14 mmol) was added to quench the reaction. The whole reaction mixture was filtered through a pad of Celite, diluted with EA, washed by a mixture of aq. $\text{Na}_2\text{S}_2\text{O}_3$ and sat. aq. NaHCO_3 (1:1) twice, brine, dried over anhydrous Na_2SO_4 , filtered. The filtrate was concentrated under reduced pressure for the purification on Biotage to provide compound **2-36 α** (179.8 mg, 71%) and **2-36 β** (21.5 mg, 8%).

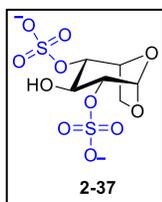
Glycosylation with sulfoxide 2-23. Sulfoxide **2-23** (106.4 mg, 0.15 mmol) was azeotropically dried with toluene three times. To a solution of sulfoxide **2-23** in DCM (3.1 mL) was consecutively added freshly activated 4 \AA MS (460 mg) and 2,6-Di-*tert*-butyl-4-methylpyridine (DTBMP, 78.7 mg, 0.38 mmol) under the atmosphere of nitrogen. The resulting mixture was allowed to stir at room temperature for 30 min. Then the reaction was cooled down to $-42\text{ }^{\circ}\text{C}$, followed by the addition of trifluoromethanesulfonic anhydride (Tf_2O , 42.3 mg, 25.8 μL , 0.38 mmol). The reaction was kept stirring at the same temperature for another 40 min, then a solution of acceptor **20** (89.7 mg, 0.18 mmol) in DCM (1.5 mL) was added. 1 h later, TLC showed completed consumption of donor. Et_3N (15.5 mg, 21.4 μL , 0.15 mmol) was added to quench the reaction. The whole reaction mixture was filtered through a pad of Celite, concentrated directly under reduced pressure for the purification on Biotage to provide compound **2-36 α** (126.3 mg, 79%) and **2-36 β** (25.3 mg, 16%).

2-36 α : $[\alpha]_{\text{D}}^{20} = +38.3$ (c 0.98, CHCl_3); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.45 – 7.40 (m, 2H), 7.40 – 7.26 (m, 15H), 7.24 – 7.19 (m, 2H), 7.19 – 7.13 (m, 2H), 6.88 – 6.83 (m, 2H), 5.61 (d, $J = 1.7$ Hz, 1H), 5.21 (d, $J = 3.8$ Hz, 1H), 4.94 – 4.85 (m, 5H), 4.76 (d, $J = 10.4$ Hz, 1H), 4.66 – 4.62 (m, 1H), 4.57 (d, $J = 10.9$ Hz, 1H), 4.51 – 4.43 (m, 2H), 4.34 (dd, $J = 10.7, 6.1$ Hz, 1H), 4.06 (d, $J = 7.9$ Hz, 1H), 3.96 –

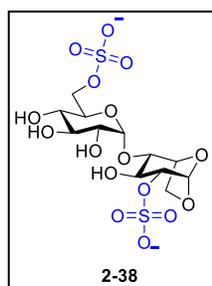
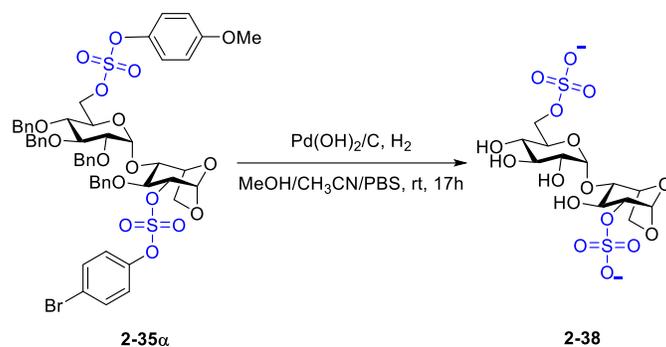
3.86 (m, 3H), 3.81 (ddd, $J = 10.3, 6.1, 1.8$ Hz, 1H), 3.78 (s, 3H), 3.74 (dd, $J = 8.0, 5.0$ Hz, 1H), 3.46 (dd, $J = 10.2, 8.8$ Hz, 1H), 3.42 (dd, $J = 10.3, 3.8$ Hz, 1H); ^{13}C NMR (126 MHz, CDCl_3) δ 158.77, 149.24, 143.66, 137.48, 137.33, 136.92, 133.20, 128.96, 128.76, 128.74, 128.71, 128.68, 128.63, 128.60, 128.47, 128.38, 128.27, 128.23, 128.20, 128.14, 128.06, 123.29, 122.37, 121.23, 115.00, 99.60, 98.49, 86.18, 80.06, 79.48, 78.79, 77.35, 75.80, 75.64, 75.59, 74.17, 71.98, 70.19, 65.97, 63.46, 55.82; HRMS (ESI) m/z $[\text{M}+\text{Na}]^+$ calcd for $\text{C}_{46}\text{H}_{46}\text{O}_{16}\text{N}_3\text{S}_2\text{Na}$ 1062.1395, found 1062.1411.

2-36 β : $[\alpha]_{\text{D}}^{20} = +18.9$ (c 0.80, CHCl_3); ^1H NMR (600 MHz, CDCl_3) δ 7.41 – 7.30 (m, 10H), 7.29 – 7.21 (m, 7H), 7.19 – 7.16 (m, 2H), 7.15 – 7.11 (m, 2H), 6.85 – 6.78 (m, 2H), 5.67 (d, $J = 1.7$ Hz, 1H), 4.95 – 4.78 (m, 5H), 4.67 (d, $J = 11.2$ Hz, 1H), 4.63 – 4.55 (m, 3H), 4.44 (d, $J = 10.8$ Hz, 1H), 4.38 (d, $J = 8.1$ Hz, 1H), 4.37 – 4.32 (m, 1H), 4.21 (d, $J = 8.1$ Hz, 1H), 4.16 (dd, $J = 8.1, 4.2$ Hz, 1H), 3.86-3.83 (m, 2H), 3.73 (s, 3H), 3.52 – 3.49 (m, 2H), 3.47 – 3.42 (m, 1H), 3.34 (dd, $J = 9.7, 8.0$ Hz, 1H); ^{13}C NMR (151 MHz, CDCl_3) δ 158.52, 149.12, 143.48, 137.51, 137.34, 136.94, 133.04, 132.97, 128.80, 128.69, 128.66, 128.65, 128.57, 128.56, 128.53, 128.51, 128.44, 128.35, 128.33, 128.28, 128.25, 128.18, 128.16, 128.14, 128.11, 128.06, 128.04, 127.98, 127.98, 127.84, 127.83, 127.79, 127.70, 123.23, 123.14, 123.12, 122.38, 122.21, 122.17, 122.16, 120.96, 114.86, 114.84, 114.71, 100.20, 98.39, 85.42, 82.69, 77.93, 77.28, 76.25, 75.65, 75.23, 74.97, 73.21, 72.59, 71.34, 66.50, 65.68, 55.61; HRMS (ESI) m/z $[\text{M}+\text{Na}]^+$ calcd for $\text{C}_{46}\text{H}_{46}\text{O}_{16}\text{N}_3\text{S}_2\text{Na}$ 1062.1395, found 1062.1412.

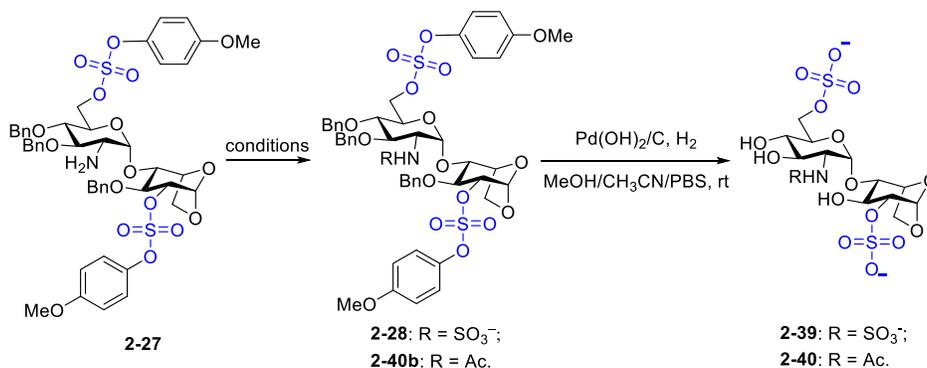


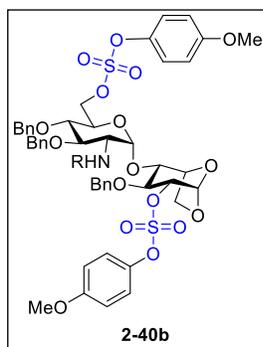


General Procedure D was employed to prepare compound **2-37** with minor modification. Compound **2-14b** (3.17 g, 4.39 mmol) in MeOH/CH₃CN/PBS buffer (97.5 mL/97.5 mL/48.8 mL) was added 20% Pd(OH)₂ on activated carbon (15.9 g, 5 g per gram of substrate). The flask was charged with gentle vacuum, and then quickly filled with hydrogen gas via a syringe attached balloon. 18 h later, NMR showed complete removal of the sulfate diester masks with benzyl group left. The reaction mixture was filtered through a pad of Celite, washed by methanol, concentrated. The reaction residue was dissolved again in MeOH/PBS buffer (97.5 mL/48.8 mL), followed by the addition of 20% Pd(OH)₂ on activated carbon (15.9 g, 5 g per gram of substrate). Same procedure was applied to fill the reaction bottle with hydrogen gas. After 27 h, the ¹H-NMR result of the reaction mixture showed no residual signal in aromatic zone. The whole reaction mixture was filtered through a pad of Celite, concentrated. The residue was passed through a column of Amberlyst IR-120 (Na⁺) resin using water as eluent, and then purified through a Sephadex LH-20 column eluted with methanol/water (1:1). The collected fraction was concentrated under reduced pressure, followed by lyophilization to provide compound **2-37** along with residual ¹H-NMR-inactive inorganic salts (*e.g.*, sodium phosphate, sodium chloride, etc. Total mass: 4.13 g) as white solid. The amount of **2-37** in the hydrogenolysis product was determined to be 1.53 g, giving the yield of 95% according to ¹H-NMR quantification using β-glycerophosphate disodium tetrahydrate as an internal standard (Figure S10). ¹H NMR (600 MHz, D₂O) δ 5.86 – 5.71 (m, 1H), 5.03 (m, 1H), 4.57 – 4.45 (m, 1H), 4.37 – 4.19 (m, 2H), 3.89 (m, 2H); ¹³C NMR (151 MHz, D₂O) δ 99.16, 80.19, 76.98, 73.18, 69.72, 65.35; HRMS (ESI) *m/z* [M-H]⁻ calcd for C₆H₉O₁₁S₂ 320.9592, found 320.9600.



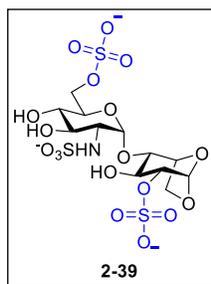
Compound **2-35 α** (54.2 mg, 0.05 mmol) was converted to trisulfated product **2-38** (24.5 mg, 99%) as white solid in 17 h following **General Procedure D**. Purification: Amberlyst IR-120 (Na^+) resin column and Sephadex LH-20 column both with MeOH as eluent. The collected fraction was concentrated under reduced pressure, followed by lyophilization. ^1H NMR (600 MHz, CD_3OD) δ 5.58 (d, $J = 1.8$ Hz, 1H), 5.07 (d, $J = 3.9$ Hz, 1H), 4.86 (t, $J = 4.6$ Hz, 1H), 4.31 (dd, $J = 10.9, 1.9$ Hz, 1H), 4.11 (dd, $J = 8.3, 1.9$ Hz, 1H), 4.05 (d, $J = 7.9$ Hz, 1H), 4.01 (dd, $J = 10.9, 7.2$ Hz, 1H), 3.85 (dd, $J = 7.8, 5.1$ Hz, 1H), 3.82 (t, $J = 8.3$ Hz, 1H), 3.74 – 3.67 (m, 2H), 3.64 – 3.59 (m, 1H), 3.44 (dd, $J = 9.9, 4.0$ Hz, 1H), 3.25 – 3.20 (m, 1H); ^{13}C NMR (126 MHz, CD_3OD) δ 102.44, 101.15, 82.95, 81.42, 75.55, 74.67, 73.43, 72.73, 72.55, 71.78, 68.61, 66.70; HRMS (ESI) m/z $[\text{M}-2\text{H}+\text{Na}]^-$ calcd for $\text{C}_{12}\text{H}_{18}\text{O}_{16}\text{S}_2\text{Na}$ 504.9939, found 504.9939.





N-Acetylation: To a solution of amine **2-27** (288.5 mg, 0.30 mmol) in DCM (6.0 mL) was consecutively added pyridine (472.5 mg, 483.1 μ L, 5.97 mmol), and acetic anhydride (Ac_2O , 304.9 mg, 282.3 μ L, 2.99 mmol) at room temperature under nitrogen atmosphere. The resulting mixture was allowed to stir for 10 h. Then the reaction was quenched with MeOH (237.3 mg, 300.0

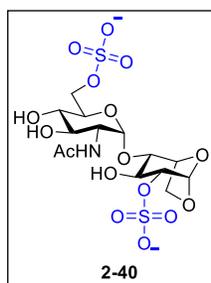
μ L, 7.41 mmol), diluted with EA, washed by sat. aq. NaHCO_3 , brine, dried over anhydrous Na_2SO_4 , filtered, coevaporated with toluene twice to get rid of pyridine. The residue was concentrated under reduced pressure for the purification with a silica gel column on Biotage to provide product **2-40b** (234.7 mg, 78%) as white solid. $[\alpha]_{\text{D}}^{20} = +61.3$ (c 1.01, CHCl_3); $^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.41 – 7.32 (m, 5H), 7.32 – 7.23 (m, 12H), 7.19 – 7.12 (m, 2H), 6.89–6.86 (d, $J = 8.8$ Hz, 1H), 6.83 – 6.79 (m, 2H), 5.14 (d, $J = 9.1$ Hz, 1H), 4.92 (d, $J = 3.7$ Hz, 1H), 4.88 (d, $J = 10.9$ Hz, 1H), 4.84–4.81 (m, 2H), 4.66 – 4.61 (m, 2H), 4.57 (d, $J = 10.9$ Hz, 1H), 4.53 (t, $J = 4.5$ Hz, 1H), 4.48–4.45 (m, 2H), 4.26–4.22 (m, 2H), 3.94 (d, $J = 8.0$ Hz, 1H), 3.86 (dd, $J = 10.1, 7.3$ Hz, 1H), 3.83 – 3.78 (m, 4H), 3.76–3.74 (m, 4H), 3.70 (dd, $J = 8.1, 5.1$ Hz, 1H), 3.65 (dd, $J = 10.6, 8.8$ Hz, 1H), 3.44 (t, $J = 9.4$ Hz, 1H), 1.32 (s, 3H); $^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 170.28, 158.77, 158.75, 143.70, 143.63, 137.83, 137.33, 137.03, 128.96, 128.88, 128.81, 128.65, 128.52, 128.47, 128.33, 128.23, 127.94, 122.56, 122.42, 114.97, 114.94, 100.06, 98.54, 85.78, 80.76, 79.98, 78.41, 75.50, 75.44, 75.22, 73.72, 72.35, 70.75, 65.67, 55.83, 55.76, 52.50, 22.72; HRMS (ESI) m/z $[\text{M}+\text{Na}]^+$ calcd for $\text{C}_{49}\text{H}_{53}\text{O}_{18}\text{NS}_2\text{Na}$ 1030.2596, found 1030.2574.



N-Sulfated compound **2-28** (229.1 mg, 0.21 mmol) was converted to trisulfated product **2-39** (135.0 mg, 83%) as white solid in 60 h following **General Procedure D**. Purification: Amberlyst IR-120 (Na^+) resin column, Sephadex LH-20 column both with MeOH as eluent, and reversed column with MeOH/ H_2O were used as

purification method. The collected fraction was concentrated under reduced pressure, followed by lyophilization. $^1\text{H NMR}$ (500 MHz, D_2O) δ 5.76 (t, $J = 1.5$ Hz, 1H), 5.47 (d, $J = 3.8$ Hz, 1H), 4.92 (t,

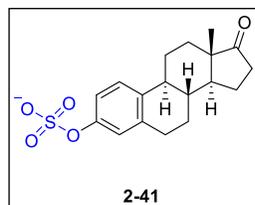
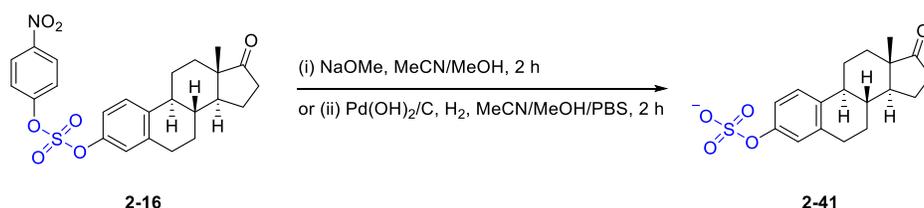
$J = 4.5$ Hz, 1H), 4.44 – 4.32 (m, 1H), 4.26 – 4.18 (m, 3H), 3.97 – 3.90 (m, 3H), 3.85 (ddd, $J = 8.9, 6.6, 1.9$ Hz, 1H), 3.68 (ddd, $J = 10.3, 9.0, 1.1$ Hz, 1H), 3.54 (ddd, $J = 10.2, 9.1, 1.1$ Hz, 1H), 3.37 (ddd, $J = 10.4, 3.8, 1.1$ Hz, 1H); ^{13}C NMR (126 MHz, D_2O) δ 99.52, 99.49, 99.09, 80.26, 80.24, 80.14, 74.06, 71.17, 70.71, 69.79, 67.44, 65.40, 57.75; HRMS (ESI) m/z $[\text{M}-2\text{H}+\text{Na}]^-$ calcd for $\text{C}_{12}\text{H}_{19}\text{O}_{18}\text{NS}_3\text{Na}$ 583.9667, found 583.9663.



Compound **2-40b** (217.3 mg, 215.5 μmol) was converted to disulfated product **2-40** (110.0 mg, 90%) as white solid following **General Procedure D** in 23 h.

Purification: Amberlyst IR-120 (Na^+) resin column, Sephadex LH-20 column both with MeOH as eluent. The collected fraction was concentrated under reduced

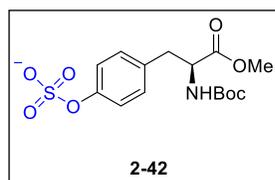
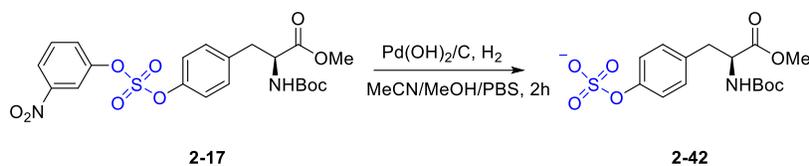
pressure, followed by lyophilization. ^1H NMR (500 MHz, D_2O) δ 5.77 (d, $J = 1.8$ Hz, 1H), 5.21 (d, $J = 3.8$ Hz, 1H), 4.90 (t, $J = 4.0$ Hz, 1H), 4.41 (dd, $J = 11.3, 2.0$ Hz, 1H), 4.30 – 4.20 (m, 3H), 4.04 (dd, $J = 10.8, 3.8$ Hz, 1H), 3.97 – 3.87 (m, 4H), 3.84 (dd, $J = 10.7, 9.0$ Hz, 1H), 3.57 (dd, $J = 10.1, 9.1$ Hz, 1H), 2.12 (s, 3H); ^{13}C NMR (126 MHz, D_2O) δ 174.54, 99.27, 99.23, 99.12, 80.45, 79.86, 74.13, 70.93, 70.89, 70.60, 69.85, 67.43, 65.50, 53.60, 22.06; HRMS (ESI) m/z $[\text{M}-2\text{H}+\text{Na}]^-$ calcd for $\text{C}_{14}\text{H}_{21}\text{O}_{16}\text{NS}_2\text{Na}$ 546.0205, found 546.0199.



Hydrolysis of **2-16** (50 mg, 0.11 μmol) was conducted following **General Procedure C**. The product was purified by silica gel chromatography (dichloromethane/methanol 5:1) to obtain **2-41** (37.4 mg, 95%).

Hydrogenolysis: **General Procedure D** was followed with **2-16** (100 mg, 0.21 μmol). The product

was purified by silica gel column chromatography with dichloromethane/methanol (5:1) as eluent to obtain **2-41** (69 mg, 87%). $[\alpha]_D^{20} = +110.4$ (c 1.00, MeOH). $^1\text{H NMR}$ (500 MHz, CD_3OD) δ 7.25 (d, $J = 8.6$ Hz, 1H), 7.07 – 7.02 (m, 2H), 2.94 – 2.87 (m, 2H), 2.50 (dd, $J = 19.0, 8.5$ Hz, 1H), 2.45 – 2.38 (m, 1H), 2.28 (td, $J = 10.5, 4.3$ Hz, 1H), 2.20 – 2.00 (m, 3H), 1.96 – 1.87 (m, 1H), 1.73 – 1.39 (m, 6H), 0.93 (s, 3H). $^{13}\text{C NMR}$ (126 MHz, CD_3OD) δ 222.30, 150.42, 137.33, 136.12, 125.54, 121.09, 118.41, 50.26, 47.87, 44.04, 38.24, 35.34, 31.40, 29.03, 26.17, 25.60, 21.10, 12.87. HRMS (ESI) m/z $[\text{M-H}]^-$ Calcd. For $\text{C}_{18}\text{H}_{21}\text{O}_5\text{S}$: 349.1115, found 349.1124.



Hydrogenolysis of **2-17** (100 mg, 0.20 mmol) was performed following **General Procedure D**. The product was purified by silica gel column chromatography (dichloromethane/methanol 5:1) to obtain **2-42** (73 mg,

91%). $[\alpha]_D^{20} = -0.4$ (c 1.00 in MeOH). $^1\text{H NMR}$ (500 MHz, CD_3OD) δ 7.23 (d, $J = 8.6$ Hz, 2H), 7.17 (d, $J = 8.6$ Hz, 2H), 4.34 (dd, $J = 8.8, 5.6$ Hz, 1H), 3.69 (s, 3H), 3.08 (dd, $J = 13.9, 5.6$ Hz, 1H), 2.92 (dd, $J = 13.9, 8.8$ Hz, 1H), 1.40 (s, 9H). $^{13}\text{C NMR}$ (126 MHz, CD_3OD) δ 172.79, 156.42, 151.50, 133.42, 129.53, 121.06, 79.26, 55.16, 51.20, 36.49, 27.26. HRMS (ESI) m/z $[\text{M-H}]^-$ Calcd. For $\text{C}_{15}\text{H}_{20}\text{NO}_8\text{S}$: 374.0915, found 374.0924.

2.5 REFERENCES

1. Warabi, K.; Hamada, T.; Nakao, Y.; Matsunaga, S.; Hirota, H.; van Soest, R. W.; Fusetani, N., Axinelloside A, an unprecedented highly sulfated lipopolysaccharide inhibiting telomerase, from the marine sponge, *Axinella infundibula*. *J. Am. Chem. Soc.* **2005**, *127* (38), 13262-13270.
2. Bowman, K. G.; Bertozzi, C. R., Carbohydrate sulfotransferases: mediators of extracellular communication. *Chem. Biol.* **1999**, *6* (1), R9-R22.
3. McLean, J., The thromboplastic action of cephalin. *Am. J. Physiol.* **1916**, *41* (2), 250-257.
4. Ratzka, A.; Vogel, H.; Kliebenstein, D. J.; Mitchell-Olds, T.; Kroymann, J., Disarming the mustard oil bomb. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (17), 11223-11228.
5. Sasisekharan, R.; Shriver, Z.; Venkataraman, G.; Narayanasami, U., Roles of heparan-sulphate glycosaminoglycans in cancer. *Nat. Rev. Cancer* **2002**, *2* (7), 521-528.
6. Lee, J.-C.; Lu, X.-A.; Kulkarni, S. S.; Wen, Y.-S.; Hung, S.-C., Synthesis of heparin oligosaccharides. *J. Am. Chem. Soc.* **2004**, *126* (2), 476-477.
7. Tully, S. E.; Mabon, R.; Gama, C. I.; Tsai, S. M.; Liu, X.; Hsieh-Wilson, L. C., A chondroitin sulfate small molecule that stimulates neuronal growth. *J. Am. Chem. Soc.* **2004**, *126* (25), 7736-7737.
8. Al-Horani, R. A.; Desai, U. R., Chemical sulfation of small molecules—advances and challenges. *Tetrahedron* **2010**, *66* (16), 2907.
9. Penney, C. L.; Perlin, A. S., A method for the sulfation of sugars, employing a stable, aryl sulfate intermediate. *Carbohydr. Res.* **1981**, *93* (2), 241-246.
10. Simpson, L. S.; Widlanski, T. S., A comprehensive approach to the synthesis of sulfate esters. *J. Am. Chem. Soc.* **2006**, *128* (5), 1605-1610.
11. Proud, A. D.; Prodder, J. C.; Flitsch, S. L., Development of a protecting group for sulfate esters. *Tetrahedron Lett.* **1997**, *38* (41), 7243-7246.
12. Karst, N. A.; Islam, T. F.; Linhardt, R. J., Sulfo-protected hexosamine monosaccharides: Potentially versatile building blocks for glycosaminoglycan synthesis. *Org. Lett.* **2003**, *5* (25), 4839-4842.
13. Karst, N. A.; Islam, T. F.; Avci, F. Y.; Linhardt, R. J., Trifluoroethylsulfonate protected monosaccharides in glycosylation reactions. *Tetrahedron Lett.* **2004**, *45* (34), 6433-6437.
14. Liu, Y.; Lien, I.-F. F.; Ruttgaizer, S.; Dove, P.; Taylor, S. D., Synthesis and protection of aryl sulfates using the 2, 2, 2-trichloroethyl moiety. *Org. Lett.* **2004**, *6* (2), 209-212.
15. Ingram, L. J.; Taylor, S. D., Introduction of 2, 2, 2-Trichloroethyl-Protected Sulfates into Monosaccharides with a Sulfuryl Imidazolium Salt and Application to the Synthesis of Sulfated Carbohydrates. *Angew. Chem. Int. Ed.* **2006**, *45* (21), 3503-3506.
16. Ingram, L. J.; Desoky, A.; Ali, A. M.; Taylor, S. D., O- and N-sulfations of carbohydrates using sulfuryl imidazolium salts. *J. Org. Chem.* **2009**, *74* (17), 6479-6485.
17. Desoky, A. Y.; Taylor, S. D., Multiple and regioselective introduction of protected sulfates into carbohydrates using sulfuryl imidazolium salts. *J. Org. Chem.* **2009**, *74* (24), 9406-9412.
18. Tiruchinapally, G.; Yin, Z.; El-Dakdouki, M.; Wang, Z.; Huang, X., Divergent heparin oligosaccharide synthesis with preinstalled sulfate esters. *Chem. Eur. J.* **2011**, *17* (36), 10106-10112.
19. Matsushita, K.; Sato, Y.; Funamoto, S.; Tamura, J.-i., Side reactions with 2, 2, 2-trichloroethoxysulfates during the synthesis of glycans. *Carbohydr. Res.* **2014**, *396*, 14-24.
20. Xu, Y.; Masuko, S.; Takieddin, M.; Xu, H.; Liu, R.; Jing, J.; Mousa, S. A.; Linhardt, R. J.; Liu, J., Chemoenzymatic synthesis of homogeneous ultralow molecular weight heparins. *Science* **2011**, *334* (6055), 498-501.
21. Chen, Y.; Li, Y.; Yu, H.; Sugiarto, G.; Thon, V.; Hwang, J.; Ding, L.; Hie, L.; Chen, X., Tailored design and synthesis of heparan sulfate oligosaccharide analogues using sequential one-pot multienzyme systems. *Angew. Chem. Int. Ed.* **2013**, *52* (45), 11852-11856.

22. Liu, J.; Linhardt, R. J., Chemoenzymatic synthesis of heparan sulfate and heparin. *Nat. Prod. Rep.* **2014**, *31* (12), 1676-1685.
23. Zhang, X.; Pagadala, V.; Jester, H. M.; Lim, A. M.; Pham, T. Q.; Goulas, A. M. P.; Liu, J.; Linhardt, R. J., Chemoenzymatic synthesis of heparan sulfate and heparin oligosaccharides and NMR analysis: paving the way to a diverse library for glycobiochemists. *Chem. Sci.* **2017**, *8* (12), 7932-7940.
24. Dong, J.; Krasnova, L.; Finn, M.; Sharpless, K. B., Sulfur (VI) fluoride exchange (SuFEx): another good reaction for click chemistry. *Angew. Chem. Int. Ed.* **2014**, *53* (36), 9430-9448.
25. Dong, J.; Sharpless, K. B.; Kwisnek, L.; Oakdale, J. S.; Fokin, V. V., SuFEx-Based Synthesis of Polysulfates. *Angew. Chem. Int. Ed.* **2014**, *53* (36), 9466-9470.
26. Gao, B.; Zhang, L.; Zheng, Q.; Zhou, F.; Klivansky, L. M.; Lu, J.; Liu, Y.; Dong, J.; Wu, P.; Sharpless, K. B., Bifluoride-catalysed sulfur (VI) fluoride exchange reaction for the synthesis of polysulfates and polysulfonates. *Nat. Chem.* **2017**, *9* (11), 1083.
27. Yang, C.; Flynn, J. P.; Niu, J., Facile Synthesis of Sequence-Regulated Synthetic Polymers Using Orthogonal SuFEx and CuAAC Click Reactions. *Angew. Chem. Int. Ed.* **2018**, *57* (49), 16194-16199.
28. Gembus, V.; Marsais, F.; Levacher, V., An efficient organocatalyzed interconversion of silyl ethers to tosylates using DBU and p-toluenesulfonyl fluoride. *Synlett* **2008**, (10), 1463-1466.
29. Oediger, H.; Moeller, F.; Eiter, K., Bicyclic amidines as reagents in organic syntheses. *Synthesis* **1972**, *1972* (11), 591-598.
30. Nielsen, M. K.; Ugaz, C. R.; Li, W.; Doyle, A. G., PyFluor: a low-cost, stable, and selective deoxyfluorination reagent. *J. Am. Chem. Soc.* **2015**, *137* (30), 9571-9574.
31. Yatvin, J.; Brooks, K.; Locklin, J., SuFEx on the surface: a flexible platform for postpolymerization modification of polymer brushes. *Angew. Chem. Int. Ed.* **2015**, *54* (45), 13370-13373.
32. Younker, J. M.; Hengge, A. C., A mechanistic study of the alkaline hydrolysis of diaryl sulfate diesters. *J. Org. Chem.* **2004**, *69* (26), 9043-9048.
33. Feng, S.; Bagia, C.; Mpourmpakis, G., Determination of proton affinities and acidity constants of sugars. *J. Phys. Chem. A* **2013**, *117* (24), 5211-5219.
34. Hutchins, R. O.; Milewski, C. A.; Maryanoff, B. E., Selective deoxygenation of ketones and aldehydes including hindered systems with sodium cyanoborohydride. *J. Am. Chem. Soc.* **1973**, *95* (11), 3662-3668.
35. Buncel, E.; Chuaqui, C., Reactivity-selectivity correlations. 2. Reactivity of alkyl aryl sulfates toward oxygen nucleophiles and the reactivity-selectivity principle. *J. Org. Chem.* **1980**, *45* (14), 2825-2830.
36. Herczeg, M.; Lázár, L.; Mándi, A.; Borbás, A.; Komáromi, I.; Lipták, A.; Antus, S., Synthesis of disaccharide fragments of the AT-III binding domain of heparin and their sulfonatomethyl analogues. *Carbohydr. Res.* **2011**, *346* (13), 1827-1836.
37. Lázár, L.; Mező, E.; Herczeg, M.; Lipták, A.; Antus, S.; Borbás, A., Synthesis of the non-reducing end trisaccharide of the antithrombin-binding domain of heparin and its bioisosteric sulfonic acid analogues. *Tetrahedron* **2012**, *68* (36), 7386-7399.
38. Li, Y.; Gong, Y.; Xu, X.; Zhang, P.; Li, H.; Wang, Y., A practical and benign synthesis of amines through Pd@mpg-C₃N₄ catalyzed reduction of nitriles. *Catal. Commun.* **2012**, *28*, 9-12.
39. Ortiz-Cervantes, C.; Iyañez, I.; García, J. J., Facile preparation of ruthenium nanoparticles with activity in hydrogenation of aliphatic and aromatic nitriles to amines. *J. Phys. Org. Chem.* **2012**, *25* (11), 902-907.
40. Saad, F.; Comparot, J.; Brahmi, R.; Bensitel, M.; Pirault-Roy, L., Influence of acid-base properties of the support on the catalytic performances of Pt-based catalysts in a gas-phase hydrogenation of acetonitrile. *Appl. Catal. A-Gen.* **2017**, *544*, 1-9.

41. Guan, B.-T.; Lu, X.-Y.; Zheng, Y.; Yu, D.-G.; Wu, T.; Li, K.-L.; Li, B.-J.; Shi, Z.-J., Biaryl construction through Kumada coupling with diaryl sulfates as one-by-one electrophiles under mild conditions. *Org. Lett.* **2010**, *12* (2), 396-399.
42. Liang, Q.; Xing, P.; Huang, Z.; Dong, J.; Sharpless, K. B.; Li, X.; Jiang, B., Palladium-catalyzed, ligand-free Suzuki reaction in water using aryl fluorosulfates. *Org. Lett.* **2015**, *17* (8), 1942-1945.
43. Ram, S.; Ehrenkauffer, R. E., Ammonium formate in organic synthesis: a versatile agent in catalytic hydrogen transfer reductions. *Synthesis* **1988**, (02), 91-95.
44. Joseph, A. A.; Verma, V. P.; Liu, X. Y.; Wu, C. H.; Dhurandhare, V. M.; Wang, C. C., TMSOTf-Catalyzed silylation: streamlined regioselective one-pot protection and acetylation of carbohydrates. *Eur. J. Org. Chem.* **2012**, *2012* (4), 744-753.
45. Xia, M.-j.; Yao, W.; Meng, X.-b.; Lou, Q.-h.; Li, Z.-j., Co₂ (CO) 6-propargyl cation mediates glycosylation reaction by using thioglycoside. *Tetrahedron Lett.* **2017**, *58* (24), 2389-2392.
46. Miethchen, R.; Fehring, V., Chirale Kronenether mit integriertem 1, 4-verbrückten d-Glucopyranose-Baustein. *Synthesis* **1998**, *1998* (01), 94-98.
47. Cao, Z.; Zhou, F.; Gu, P.-Y.; Chen, D.; He, J.; Cappiello, J. R.; Wu, P.; Xu, Q.; Lu, J., Preparation of aryl polysulfonates via a highly efficient SuFEx click reaction, their controllable degradation and functionalized behavior. *Polym. Chem.* **2020**, *11* (18), 3120-3124.
48. Zulueta, M. M. L.; Lin, S.-Y.; Lin, Y.-T.; Huang, C.-J.; Wang, C.-C.; Ku, C.-C.; Shi, Z.; Chyan, C.-L.; Irene, D.; Lim, L.-H., α -Glycosylation by D-glucosamine-derived donors: synthesis of heparosan and heparin analogues that interact with mycobacterial heparin-binding hemagglutinin. *J. Am. Chem. Soc.* **2012**, *134* (21), 8988-8995.
49. Lee, J.-C.; Tai, C.-A.; Hung, S.-C., Sc (OTf) 3-catalyzed acetolysis of 1, 6-anhydro- β -hexopyranoses and solvent-free per-acetylation of hexoses. *Tetrahedron Lett.* **2002**, *43* (5), 851-855.
50. Niemietz, M.; Perkams, L.; Hoffman, J.; Eller, S.; Unverzagt, C., Selective oxidative debenzoylation of mono- and oligosaccharides in the presence of azides. *Chem. Commun.* **2011**, *47* (37), 10485-10487.
51. J. Dong; Sharpless, K. B. Sulfur (VI) fluoride compounds and methods for the preparation thereof. WO 2015188120 A1, June 5, 2015, 2015.
52. Jiajia, D.; Qian, Y.; Taijie, G.; Xiongjie, Z.; Genyi, M. Fluorine-containing sulfonyl compound as well as intermediate, preparation method and application thereof. CN 107857730 A, March 30, 2018.
53. Ma, C.; Zhao, C.-Q.; Xu, X.-T.; Li, Z.-M.; Wang, X.-Y.; Zhang, K.; Mei, T.-S., Nickel-Catalyzed Carboxylation of Aryl and Heteroaryl Fluorosulfates Using Carbon Dioxide. *Org. Lett.* **2019**, *21* (7), 2464-2467.
54. Gilles, P.; Veryser, C.; Vangrunderbeeck, S.; Ceusters, S.; Van Meervelt, L.; De Borggraeve, W. M., Synthesis of N-acyl sulfamates from fluorosulfates and amides. *J. Org. Chem.* **2018**, *84* (2), 1070-1078.
55. Girard, P.; Kagan, H.; David, S., Spectres de RMN de quelques derives de sucres en presence de chelates de terres rares. *Tetrahedron* **1971**, *27* (23), 5911-5920.
56. Lee, J. C.; Francis, S.; Dutta, D.; Gupta, V.; Yang, Y.; Zhu, J.-Y.; Tash, J. S.; Schönbrunn, E.; Georg, G. I., Synthesis and evaluation of eight- and four-membered iminosugar analogues as inhibitors of testicular ceramide-specific glucosyltransferase, testicular β -glucosidase 2, and other glycosidases. *J. Org. Chem.* **2012**, *77* (7), 3082-3098.
57. Ishiwata, A.; Munemura, Y.; Ito, Y., NAP ether mediated intramolecular aglycon delivery: a unified strategy for 1, 2-cis-glycosylation. *Eur. J. Org. Chem.* **2008**, *2008* (25), 4250-4263.
58. Li, J.; Li, W.; Yu, B., A divergent approach to the synthesis of simplexides and congeners via a late-stage olefin cross-metathesis reaction. *Org. Biomol. Chem.* **2013**, *11* (30), 4971-4974.

59. Ngoje, G.; Li, Z., Study of the stereoselectivity of 2-azido-2-deoxyglucosyl donors: protecting group effects. *Org. Biomol. Chem.* **2013**, *11* (11), 1879-1886.
60. Bauder, C., A convenient synthesis of orthogonally protected 2-deoxystreptamine (2-DOS) as an aminocyclitol scaffold for the development of novel aminoglycoside antibiotic derivatives against bacterial resistance. *Org. Biomol. Chem.* **2008**, *6* (16), 2952-2960.
61. Dasgupta, F.; Garegg, P. J., Monoalkylation of Tributyltin Activated Methyl 4, 6-O-Benzylidene- α -D-glucopyranosides. *Synthesis* **1994**, *1994* (11), 1121-1123.
62. Hung, S.-C.; Lu, X.-A.; Lee, J.-C.; Chang, M. D.-T.; Fang, S.-l.; Fan, T.-c.; Zulueta, M. M. L.; Zhong, Y.-Q., Synthesis of heparin oligosaccharides and their interaction with eosinophil-derived neurotoxin. *Org. Biomol. Chem.* **2012**, *10* (4), 760-772.
63. Lu, L.-D.; Shie, C.-R.; Kulkarni, S. S.; Pan, G.-R.; Lu, X.-A.; Hung, S.-C., Synthesis of 48 disaccharide building blocks for the assembly of a heparin and heparan sulfate oligosaccharide library. *Org. Lett.* **2006**, *8* (26), 5995-5998.
64. Deng, S.; Gangadharmath, U.; Chang, C.-W. T., Sonochemistry: A powerful way of enhancing the efficiency of carbohydrate synthesis. *J. Org. Chem.* **2006**, *71* (14), 5179-5185.
65. Karmakar, A.; Basha, M.; Babu, G. V.; Botlagunta, M.; Malik, N. A.; Rampulla, R.; Mathur, A.; Gupta, A. K., Tertiary-butoxycarbonyl (Boc)—A strategic group for N-protection/deprotection in the synthesis of various natural/unnatural N-unprotected aminoacid cyanomethyl esters. *Tetrahedron Lett.* **2018**, *59* (48), 4267-4271.
66. Bucher, C.; Gilmour, R., Fluorine-Directed Glycosylation. *Angew. Chem. Int. Ed.* **2010**, *49* (46), 8724-8728.
67. Rathore, H.; Hashimoto, T.; Igarashi, K.; Nukaya, H.; Fullerton, D. S., Cardiac glycosides: 5. Stereoselective syntheses of digitoxigenin α -D, β -D, α -L, and β -L-glucosides. *Tetrahedron* **1985**, *41* (23), 5427-5438.

Chapter 3

Fluorosulfate as a Latent Sulfate in Peptides and Proteins

A significant portion of the work described in this chapter has been submitted in a manuscript:

C. Liu, X. Liu, M. Zhou, C. Xia, C. Soni, Y. Lyu, Q. Peng, Z. Zhou, Y. Wu, E. Weerapana, J. Gao, A. Chatterjee, C. Lin, J. Niu, *Manuscript in Revision* **2023**.

3.1 INTRODUCTION

3.1.1 Overview

Tyrosine-*O*-Sulfation widely exists in the eukaryotic proteome. However, understanding of the biological functions of sulfation in peptides and proteins has been hampered by the lack of methods to control its spatial or temporal distribution in the proteome. Herein, we report that fluorosulfotyrosine can serve as a latent precursor of sulfotyrosine in peptides and proteins, which can be efficiently converted into sulfotyrosine residues by hydroxamic acid activators under physiologically relevant conditions. Photocaging the hydroxamic acid activators further allowed for light-controlled activation of functional sulfopeptides. This work provides a valuable tool for probing functional roles of sulfation in the peptides and proteins.

*3.1.2 Significance of tyrosine-*O*-sulfation and current challenge*

O-Sulfation of the tyrosine residue is a post-translational modification (PTM) that widely exists in eukaryotic peptides and proteins (Figure 1a), and has been implicated to regulate a variety of biological functions such as immune response, hemostasis, and pathogen evasion.¹⁻² However, only a small fraction of the sulfoproteome has been annotated.³⁻⁴ A long-standing challenge for studying the sulfoproteome is that sulfation is highly heterogeneous, with various sulfopeptides and sulfoproteins exist in different sulfoforms.⁵ The seminal works of Schultz,⁶ Liu,⁷ Chatterjee,⁸ Niu,⁹ and Xiao¹⁰ that incorporate sulfotyrosine (sY) into proteins as a non-canonical amino acid (ncAA) represent notable examples to address this challenge. Expanding upon these advances, methods that allow researchers to spatiotemporally control sulfation in the proteomic context would be highly valuable for studying their functional roles in biology.¹¹

Caging strategies have been developed for various protein PTMs to probe how these PTMs regulate dynamic cellular events. Although a broad collection of caging groups is available for a variety of PTMs, a caging group that stably protects sulfotyrosine (sY) residues in peptides and

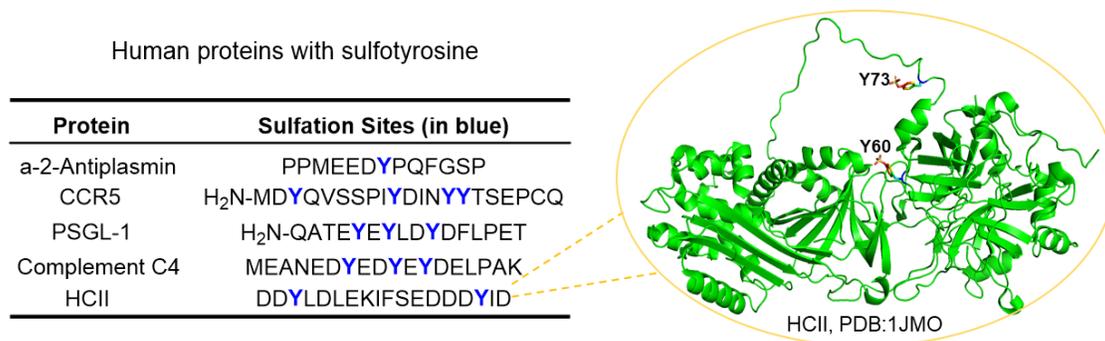


Figure 3-1. Sulfation widely exists in diverse bioactive peptides and proteins

proteins and can be efficiently removed under physiological conditions remains elusive.¹² The reasons for such a knowledge gap includes the high energy barrier for chemically activating the sulfate group for coupling chemistries, the liability of sY to acid, heat, and high-energy ionization, and the strong electron-withdrawing propensity of sulfate that renders commonly used benzylic ester caging groups unstable.¹³⁻¹⁵ On the other hand, while multiple alkyl and aryl esters have been successfully used as protecting groups of sY in solid-phase peptide synthesis,¹⁶⁻¹⁷ such as 2,2,2-trichloroethyl (TCE),¹⁸⁻¹⁹ 2,2-dichlorovinyl (DCV),²⁰⁻²¹ 2,2,2-trifluoroethyl (TFE),²² neopentyl,²³⁻²⁴ and phenyl²⁵ sulfate diesters, their deprotection conditions (e.g., strong base, hydrogenolysis, etc.) are incompatible with biological systems.

3.1.3 Our Solution

In 2014, Sharpless *et al.* reported the reactivity of fluorosulfate in Sulfur(VI) Fluoride exchange (SuFEx) reaction, which has found broad utilities in organic chemistry, polymer science, and chemical biology.²⁶⁻²⁹ Compared to other halogen-substituted sulfate derivatives, fluorosulfate not only has a

size closest to that of sulfate, but is also far less electrophilic due to the π -donation from fluorine to sulfur.³⁰ As a result, fluorosulfate has demonstrated excellent metabolic stability *in vivo*.³¹ While reactivities of fluorosulfate with cellular nucleophiles have been reported, these examples all require the close spatial proximity through ligand-receptor binding.^{13, 32-33} The chemical inertness of fluorosulfate has allowed its tyrosine derivative, L-fluorosulfotyrosine (fsY), to be incorporated into peptides and proteins via solid-phase peptide synthesis^{17, 34} and non-canonical amino acid (ncAA) mutagenesis.^{13, 32} Herein, we demonstrate that fluorosulfate can serve as a latent sulfate in sulfopeptides and sulfoproteins and can be efficiently converted into sulfate (hereafter denoted as “decaging”) by hydroxamic acid activators under physiologically relevant conditions. Mechanistic studies revealed an unusual Lossen rearrangement pathway of fluorosulfate activation and decaging (Figure 3-2a) that is

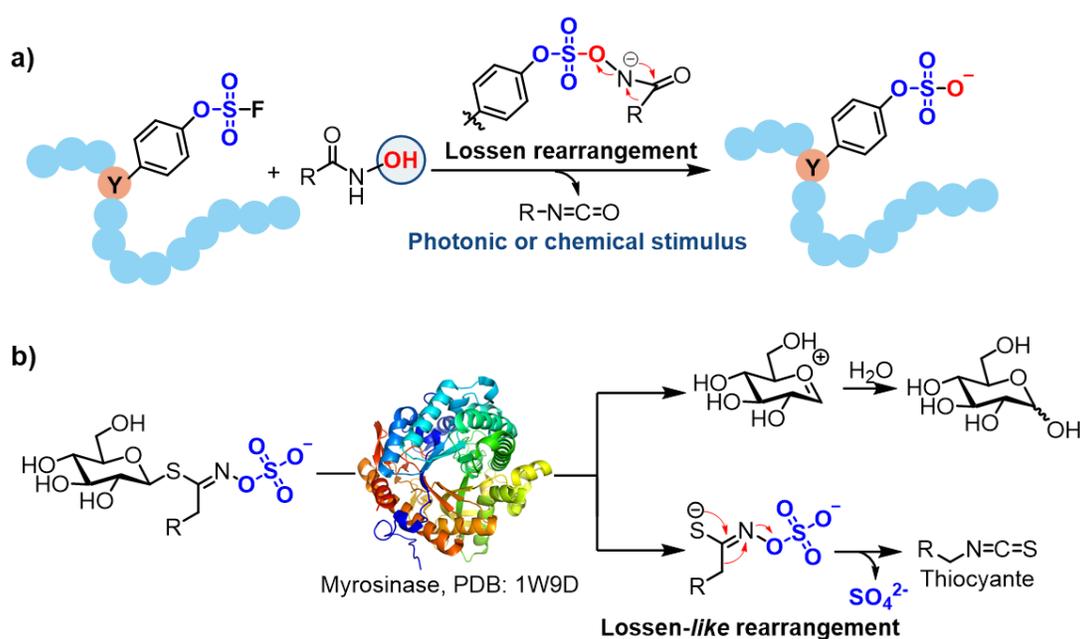


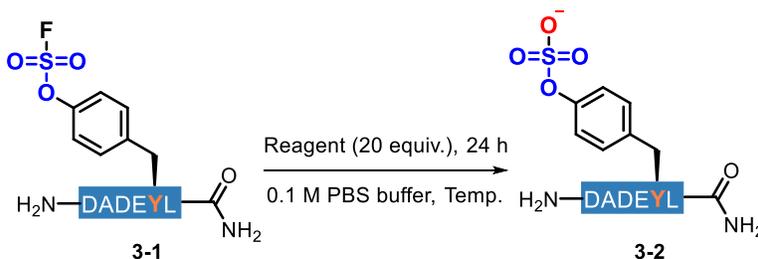
Figure 3-2. Our solution. **a)** In this work, fluorosulfate is incorporated in peptides and proteins as a latent sulfate and can be efficiently converted into sulfate by hydroxamic acid activators under physiologically relevant conditions; **b)** Our approach mirrors the myrosinase-catalyzed Lossen-like rearrangement of glucosinolates in nature.

analogous to the Myrosinase-mediated Lossen-like rearrangement of glucosinolate in nature (Figure 3-2b). It is noteworthy that a recent report by Kelly *et al.* showed that fsY could be converted into sY in synthetic peptides, but the reaction required strongly basic conditions and was incompatible with physiological peptides and proteins.²⁸

3.2 RESULTS AND DISCUSSION

Our investigation began with confirming the stability of fsY in various aqueous environments such as buffer solution, cell lysate, and serum. The fsY-containing hexapeptide **3-1** was found to be stable in pH 7.0 phosphate buffer saline (PBS) or pH 8.0 tris(hydroxymethyl)aminomethane (Tris) Tris buffer with negligible hydrolysis of fluorosulfate after 24 hours (Table 1, entry 1-2).

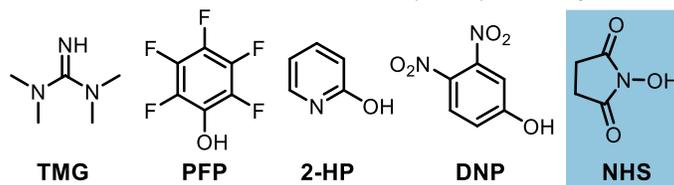
Table 3-1. Initial exploration of the release conditions of fsY



| Entry | Reagent | Temp. | Yield of 2 ^a |
|----------------|----------|-------|-------------------------|
| 1 | - | 37 °C | 4% |
| 2 ^b | - | 37 °C | 2% |
| 3 | TMG | 25 °C | 2% |
| 4 | TMG/PFP | 25 °C | 1% |
| 5 | TMG/2-HP | 25 °C | 1% |
| 6 | TMG/DNP | 25 °C | 1% |
| 7 | NHS | 25 °C | 61% |
| 8 | TMG/NHS | 25 °C | 56% |
| 9 | NHS | 37 °C | quant. |

^aThe yields was calculated by the integral ratio of all emerged peptide peaks.

^b0.1 M Tris buffer was used as the solvent. quant.: quantitative yield.



activator or when it is combined with co-activators such as pentafluorophenol (PFP), 2-hydroxypyridine (2-HP), or 2,4-dinitrophenol (DNP) (Table 3-1, entry 3-6). To our surprise, *N*-hydroxysuccinimide (NHS) alone was able to convert **1** into **2** in 61% yield at 25 °C in 24 hours, and

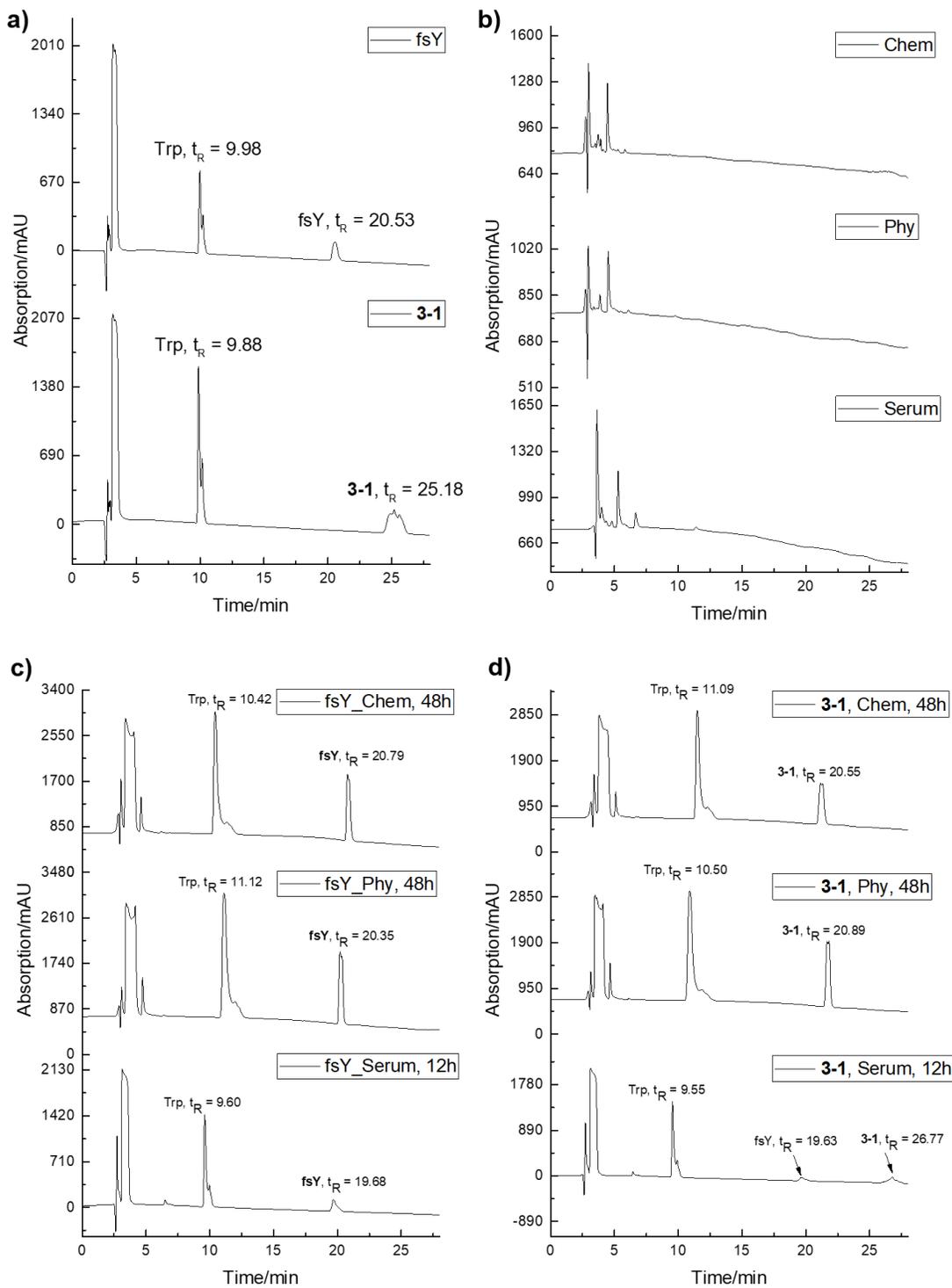


Figure 3-3. Analytical RP-HPLC verified the stability of fsY and 3-1 in cell lysate. See the detailed caption in next page.

Figure 3-3. Analytical RP-HPLC verified the stability of fsY and **3-1** in cell lysate. a) Analytical RP-HPLC traces of fsY or hexpeptide **3-1**. A known amount of tryptophan (Trp) was co-injected as an internal standard. b) Analytical RP-HPLC trace of cell lysate at 37 °C for 48 h after centrifuge filtration. No signal overlap was observed between the lysate and Trp or the substrates. c) Analytical RP-HPLC trace of fsY in cell lysate at 37 °C after centrifuge filtration: in Chem lysate for 48 h (top); in Phy lysate for 48 h (middle); in human serum for 12 h (bottom). A known amount of Trp was co-injected as an internal standard. d) Analytical RP-HPLC trace of hexpeptide **3-1** in cell lysate at 37 °C after centrifuge filtration: in Chem lysate for 48 h (top); in Phy lysate for 48 h (middle); in human serum for 12 h (bottom). A known amount of Trp was co-injected as an internal standard. The stability of the substrate is calculated based on the ratios of the substrate or product peak to Trp peak. The ratio of the substrate to Trp is r_1 , and the ratio of the product to Trp is r_2 . Then the stability is calculated as $\eta = r_2/r_1 * 100\%$. HPLC method was set as: using CH₃CN: 20 mM ammonium acetate buffer as mobile phase, 0-5 min, 7%~10% CH₃CN; 5-26 min, 10%~27% CH₃CN. All analytical RP-HPLC was monitored under 210 nm wavelength with 10.0 μ L injection volume.

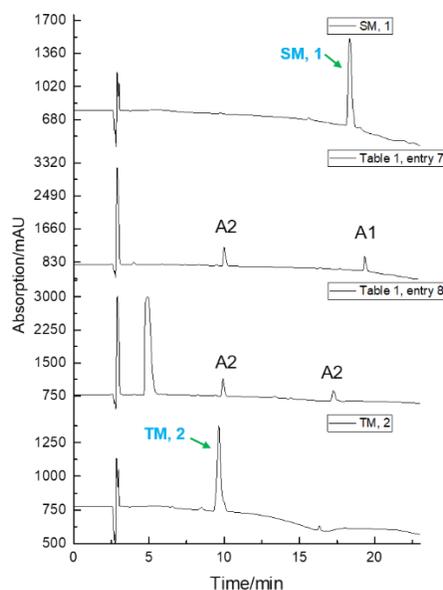
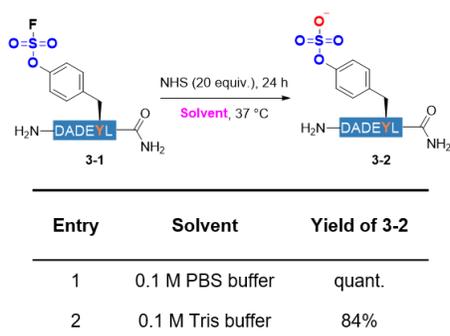


Figure 3-4. HPLC analysis of the yield of the decaging of **1** shown in Table 3-1. From top to bottom: analytical HPLC traces of pure starting material **1**, reaction mixture of Entry 7 (NHS, rt, 24 h), reaction mixture of Entry 8 (TMG+NHS, rt, 24 h), and pure product **2**. Yields were calculated using the ratios of the integrals. Based on the clean reaction mixture results and similar absorption groups in **1** & **2**, we directly utilized the integral ratio as the reaction yield: Yield of the reaction = $A_2/(A_1 + A_2) * 100\%$. The HPLC method was set as: using CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~33% CH₃CN from 0-20 min, 20-21 min 33%~95% CH₃CN, and 95% CH₃CN washing for another 2 min. All analytical RP-HPLC was monitored under 210 nm wavelength.

no other byproducts were detected on HPLC (Table 3-1, entry 7). Adding TMG to the reaction did not further improve the yield of **3-2** (Table 3-1, entry 8). The reaction was accelerated when the temperature was elevated to 37 °C, producing **3-2** quantitatively in 24 hours (Table 3-1, entry 9). HPLC monitoring showed a clean reaction mixture (Figure 3-4).

We also evaluated the influence of different buffer media on this system. In fact, changing the buffering system to Tris buffer system compromised the conversion to 84% (Table 3-3). We hypothesized that the ionization of active NHS species by Tris cation might lead to lower reactivity.

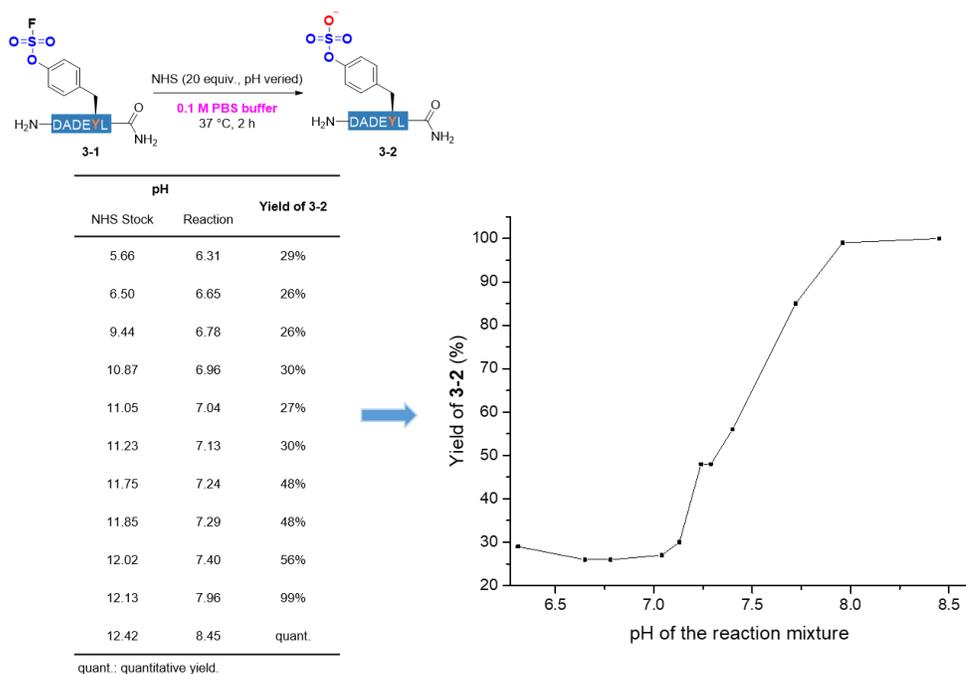
Table 3-3. Decaging reaction mediated by NHS in Tris buffer



As the basicity of the buffer system plays crucial role in the activation of NHS, a pH dependence plot was gained for deeper insight into the mechanism. By adding NHS solution with increasing basicity to the same substrate stock solution, the resulting reaction mixture's initial pH was monitored and allowed to stir at 37 °C for 2 hours (Table 3-4). In contrast to our expectation that the reaction should have proceeded in a gradient pH-dependant fashion, subtle yield changed platformed at certain pH values (6.3-7.2) and sharp jumps of the targeting material **3-2** production in a very narrow pH range were observed.

Taking the side reaction of NHS in SPPS³⁷ into account, we believed that a more reactive species towards fluorosulfate hub was derived from NHS through hydrolysis which was fully characterized by NMR assay (Figure 3-5).

Table 3-4. The pH dependence of the fsY decaging reaction mediated by NHS



We wondered if the hydroxamic acid (HA) motif of NHS is the reactive center that mediated the decaging reaction (Table 3-4 and Figure 3-5). Indeed, we found that the hydrolysis product of NHS, *N*-hydroxylsuccinic acid monoamide (**3-3**), provided **3-2** in 57% yield in one hour, indicating its superior nucleophilicity towards fluorosulfate (Figure 3-6). Encouraged by this finding, we examined other HA derivatives (Figure 3-6 and Figure 3-7). Acetohydroxamic acid (**3-4**) promoted the reaction to 78% over one hour. Good yield (95%) of **3-2** was obtained using aromatic benzohydroxamic acid (**3-5**) as the activator under the same condition. The highest efficiency was observed when the cationic HA **3-6** and heteroaromatic HA **3-7** were used, achieving quantitative conversion in 30 minutes. Similar to **3-7**, heteroaromatic hydroxamic acids consisting of pyridine and imidazole structures, e.g., **3-S13**, **3-S14**, and **3-S17**, were also found to result in efficient decaging, producing **3-2** in 91%, 84%, and 94% yields in 30 minutes, respectively. Other non-HA α -nucleophile activators such as oxime **3-8**,³⁸ 2-aminoxime **3-9**, and 1-hydroxybenzotriazole (**3-10**)³⁹ resulted in lower reaction efficiency.

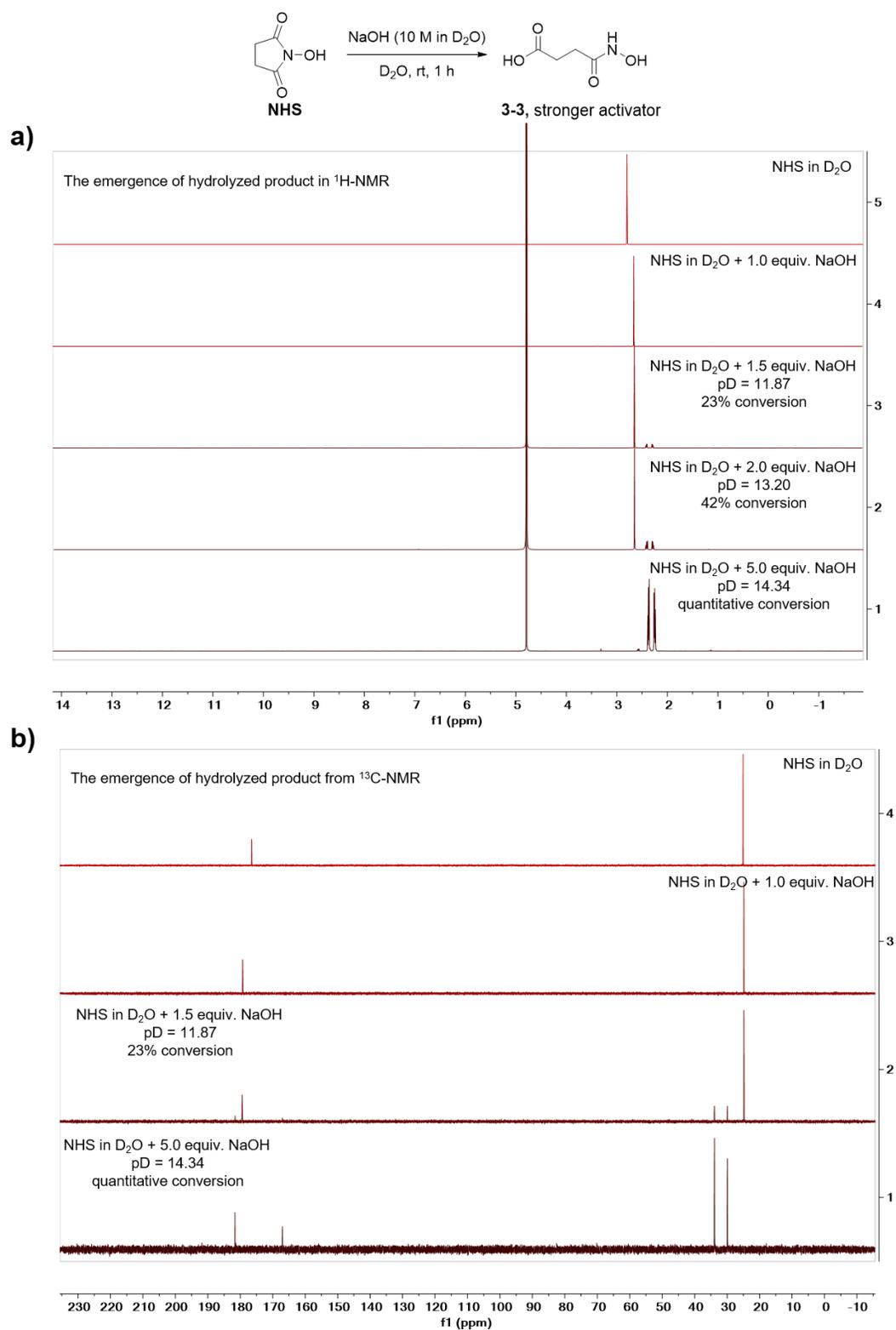


Figure 3-5. Hydrolysis of NHS in deuterated water by titrating NaOH. **a)** $^1\text{H-NMR}$ of the reaction in D_2O . **b)** $^{13}\text{C-NMR}$ of the reaction in D_2O .

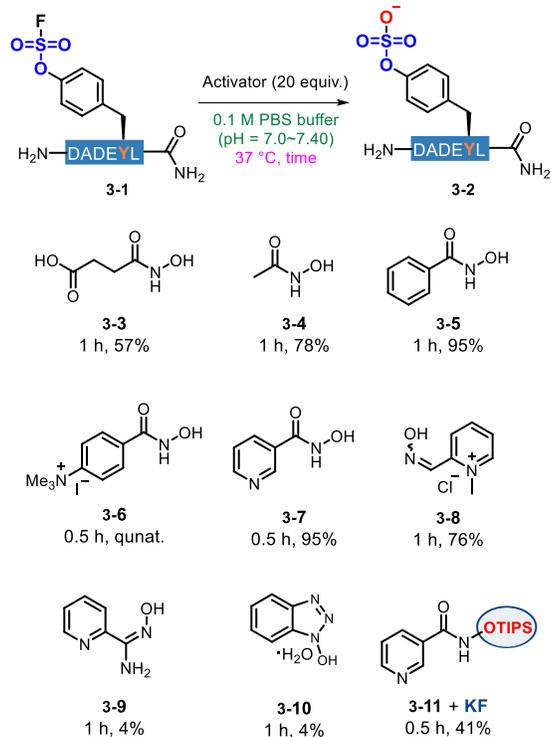


Figure 3-6. A variety of hydroxamic acid activators were investigated for their ability to activate fluorosulfate of model peptide **3-1**

In contrast, triisopropylsilyl (TIPS) ether-masked HA **3-11** showed no reactivity until potassium fluoride (KF) was added to remove the TIPS protecting group (Table 3-5 and Figure 3-8), confirming that HA is the reactive center for fluorosulfate activation.

Table 3-5. Decaging of fsY in **3-1** mediated by silyl-protected activator **3-11**

3-1 $\xrightarrow[\text{30 min, 37 }^\circ\text{C, C = 0.82 mM}]{\text{3-11, additive, 0.1 M PBS buffer}}$ 3-2

| Entry | Additive | Yield of 3-2 |
|-------|----------|--------------|
| 1 | - | 0 |
| 2 | KF | 41% |

3-11: Cc1ccc(Oc2ccccc2)cc1 Potassium fluoride
 KF

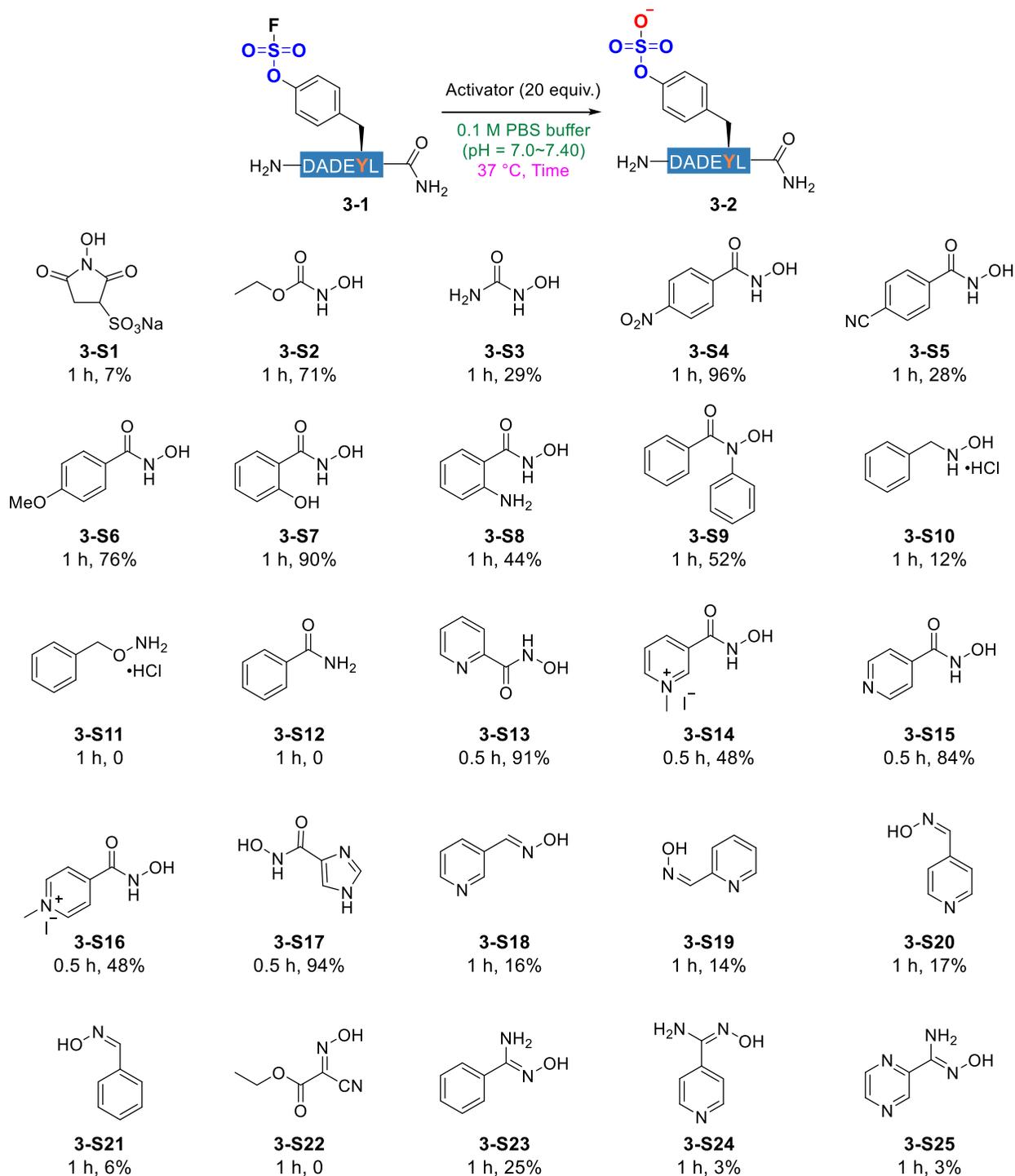


Figure 3-7. Extended activator screening for the fsY decaging using **3-1** as the substrate

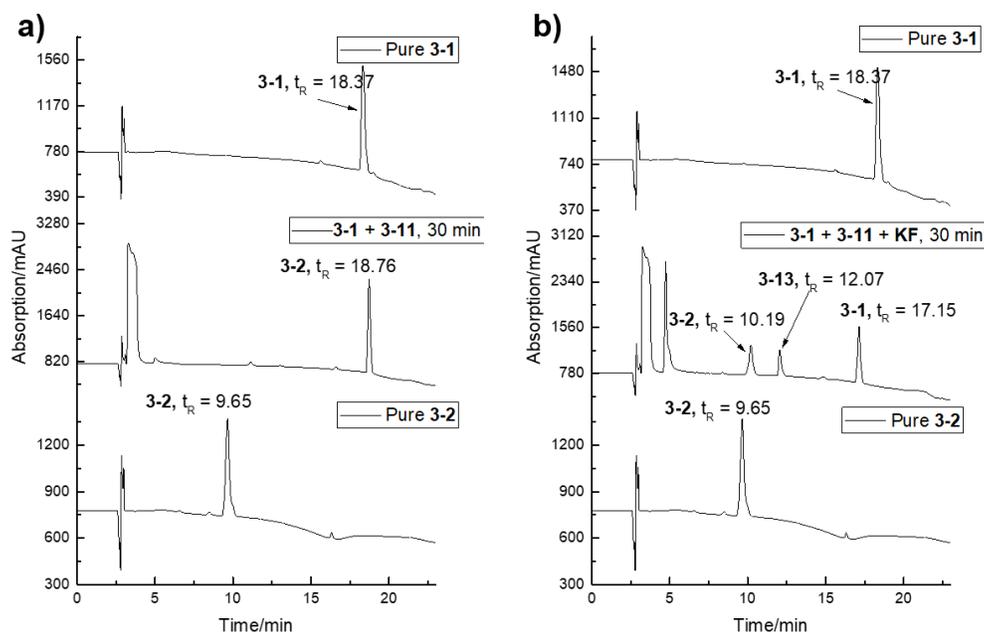


Figure 3-8. Analytical HPLC trace of the fsY decaging mediated by silyl-protected activators. **a)** Substrate 1 remained latent when only silyl caged **3-11** was added. From top to bottom: pure starting material **3-1**, **3-1** with **3-11**, and pure targeting material **3-2**. **b)** The decaging of fsY occurred in the presence of both caged activator **3-11** and potassium fluoride (**KF**). From top to bottom: pure starting material **3-1**, **3-1** with **3-11** and **KF**, and pure targeting material **3-2**. HPLC method was set as CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~33% CH₃CN from 0-20 min, 20-21 min 33%~95% CH₃CN, and 95% CH₃CN washing for another 2 min. All analytical RP-HPLC was monitored under 210 nm wavelength.

It is also noteworthy that the decaging reaction mediated by **3-7** proceeded with no detectable side reaction in the presence of 20 equivalents of nucleophilic amino acids including lysine, histidine, and tyrosine (Table 3-6), highlighting the high selectivity of **3-7** to fluorosulfate.

In order to gain insight into the reaction mechanism, the reaction with **3-1** as the substrate and **3-9** as the activator was monitored using liquid chromatography mass spectrometry (LC-MS) to capture the reaction intermediates (Figure 3-9). An adduct (**3-12**) of **3-7** and **3-1** was detected after the reaction was

Table 3-6. Compatibility test of the decaging reaction of **1** in the presence of free amino acids

| Reagent | pH ^a | Concentration | Yield of 2 | |
|-------------------------------|-----------------|---------------|-------------------|--|
| 3-7 | 7.24 | 0.83 mM | 95% | |
| 3-7 + Ala | 7.16 | 0.71 mM | 80% | |
| 3-7 + Pro | 7.18 | 0.71 mM | 73% | |
| 3-7 + Tyr ^a | 7.50 | 0.71 mM | 78% | |
| 3-7 + Arg | 7.20 | 0.71 mM | 71% | |
| 3-7 + His | 7.15 | 0.71 mM | 71% | |
| 3-7 + His | 7.13 | 0.71 mM | 77% | |
| 3-7 + Cys | 7.21 | 0.71 mM | 84% | |
| 3-7 + Lys | 7.17 | 0.71 mM | 74% | |
| 3-7 + Thr | 7.18 | 0.71 mM | 84% | |
| 3-7 + Ser | 7.23 | 0.71 mM | 82% | |

^aReaction mixture pH was examined after the reaction; ^bTyrosine in DMSO as white suspension.

placed at 25 °C for 24 hours, confirming the nucleophilic coupling between the HA activator and the substrate. Surprisingly, an isocyanate adduct **3-13** was also detected within 10 minutes at 37 °C, suggesting an uncommon intramolecular Lossen rearrangement pathway.

To further verify this possibility, we performed the decaging reaction of **3-1** by **3-7** in the buffer prepared exclusively using H₂¹⁸O. This reaction yielded **3-2** that contained no ¹⁸O isotope (Figure 3-10 and Figure 3-11), suggesting that the conversion of fluorosulfate into sulfate is not through direct hydrolysis. These results further support an exclusive Lossen rearrangement mechanism accounted for the fsY moiety release (Figure 3-12).⁴⁰⁻⁴¹ Such a pathway is similar to the

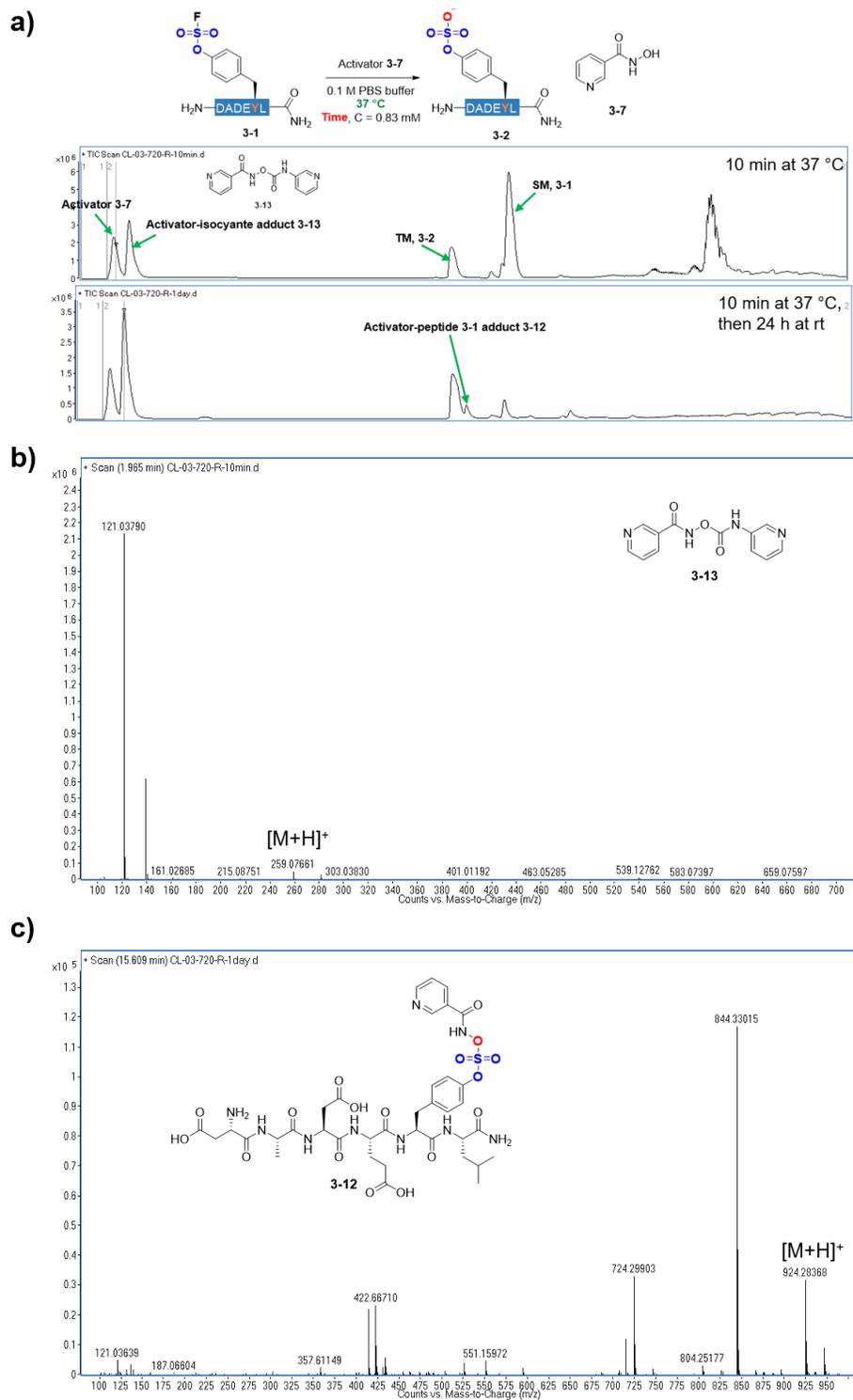


Figure 3-9. Real-time LC-MS analysis to probe reaction mechanism. **a)** Total ion chromatogram (TIC) of the reaction after 10 min at 37 °C and then incubated at room temperature for 24 h. **b)** Mass spectrum of activator (3-7)-isocyanate adduct 3-13. **c)** Mass spectrometry captured activator (3-7)-substrate 3-1 adduct 3-12.

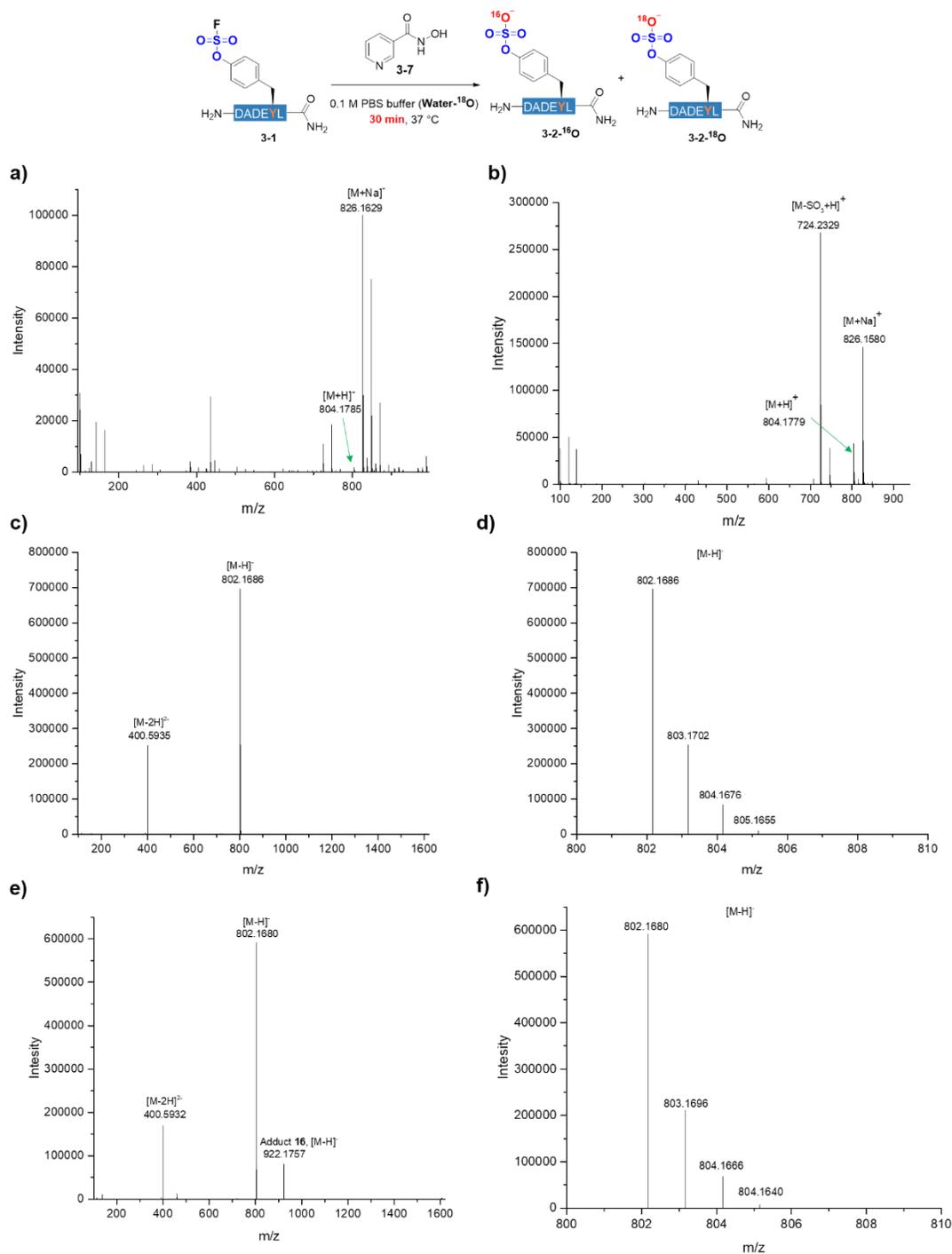


Figure 3-10. Mass spectra of H₂¹⁸O isotope labeling experiment. **a)** Product 2 in H₂¹⁸O buffer, positive mode. **b)** decaying reaction mixture containing starting material 1 in H₂¹⁸O buffer, positive mode. **c)** product 2 in H₂¹⁸O buffer, negative mode. **d)** zoom-in mass spectrum of product 2 in H₂¹⁸O buffer, negative mode. **e)** 1 in H₂¹⁸O buffer, negative mode. **f)** zoom-in mass spectrum of 1 in H₂¹⁸O buffer, negative mode.

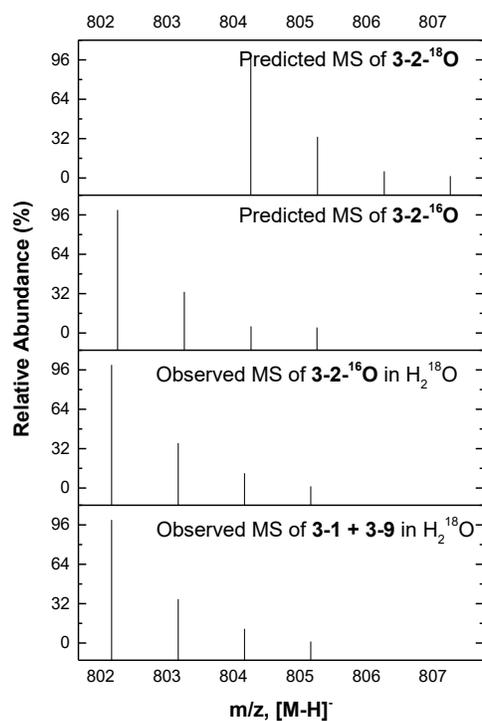


Figure 3-11. No ^{18}O -labeled products were found from the reaction in H_2^{18}O buffer, suggesting that the sulfate product was not generated from direct hydrolysis.

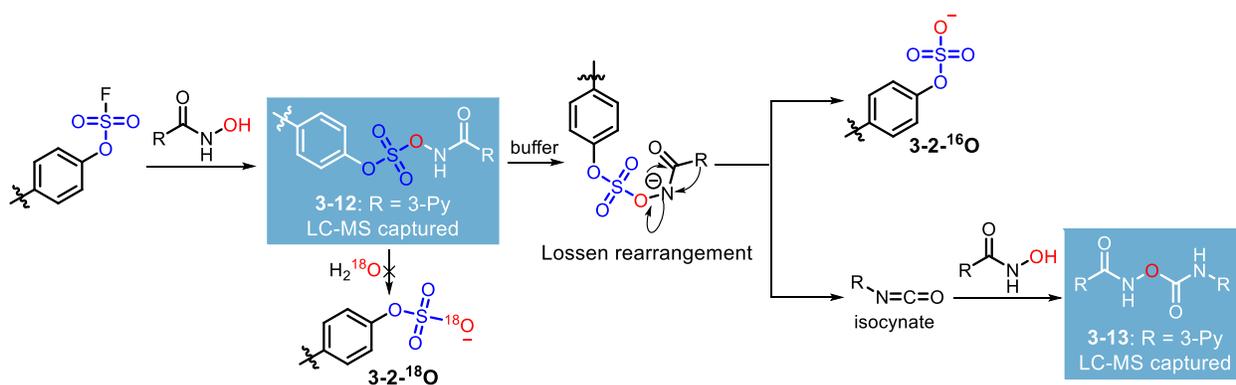


Figure 3-12. Real-time LC-MS reaction monitoring identified two activator **3-7** adducts **3-12** and **3-13**, suggesting a Lossen rearrangement mechanism

myrosinase-catalyzed Lossen-like rearrangement of glucosinolate in *Brassica* plants,⁴² in which an inorganic sulfate and isothiocyanate are generated from a thiohydroximate-*O*-sulfate intermediate (Figure 3-2b).

With the optimized conditions in hand, we examined the decaging of various fluorosulfate-containing peptides mediated by **3-7** under physiological pH. Besides **3-1**, which was decaged to produce **3-2** in 93% yield, an octapeptide **3-14** consisting three fsY residues and a C5a receptor 22-mer⁴³ peptide **3-16** containing two fsY residues were decaged to yield the corresponding sulfopeptides in 69% and 74% isolated yield, respectively (Table 3-7).

Table 3-7. Fluorosulfate activation and decaging in fsY-containing synthetic peptides

| Sequence | Substrate | Product (yield ^a) |
|---|-------------|-------------------------------|
| DADEYL-NH ₂ | 3-1 | 3-2 (93%) |
| YEYLDYDF-NH ₂ | 3-14 | 3-15 (69%) |
| Biotin-TTPDYGHYDDKDTLDLNTVPDK-NH ₂ | 3-16 | 3-17 (74%) |

^aIsolated yields were reported.

In addition to the reagent-mediated decaging, we also wondered if fluorosulfate decaging can be achieved in a light-mediated fashion. To this end, we prepared the 2-nitrobenzyl-caged activator **3-18** and demonstrated that it could activate and decage the fsY residues in **3-16** after exposed to 370 nm UV light irradiation, affording the corresponding sulfopeptide **3-17** in 60% isolated yield over 2.5 hours (Figure 3-13). Semicarbazide was added to scavenge the nitrosobenzaldehyde produced in the light-mediated reaction (Figure 3-14).⁴⁴

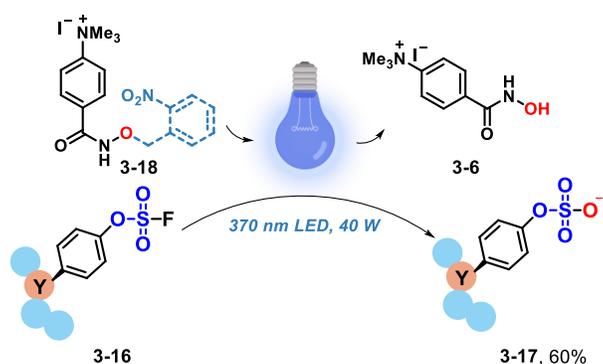


Figure 3-13. Light-mediated release of C5a receptor 22mer peptide from caged activator **3-18**

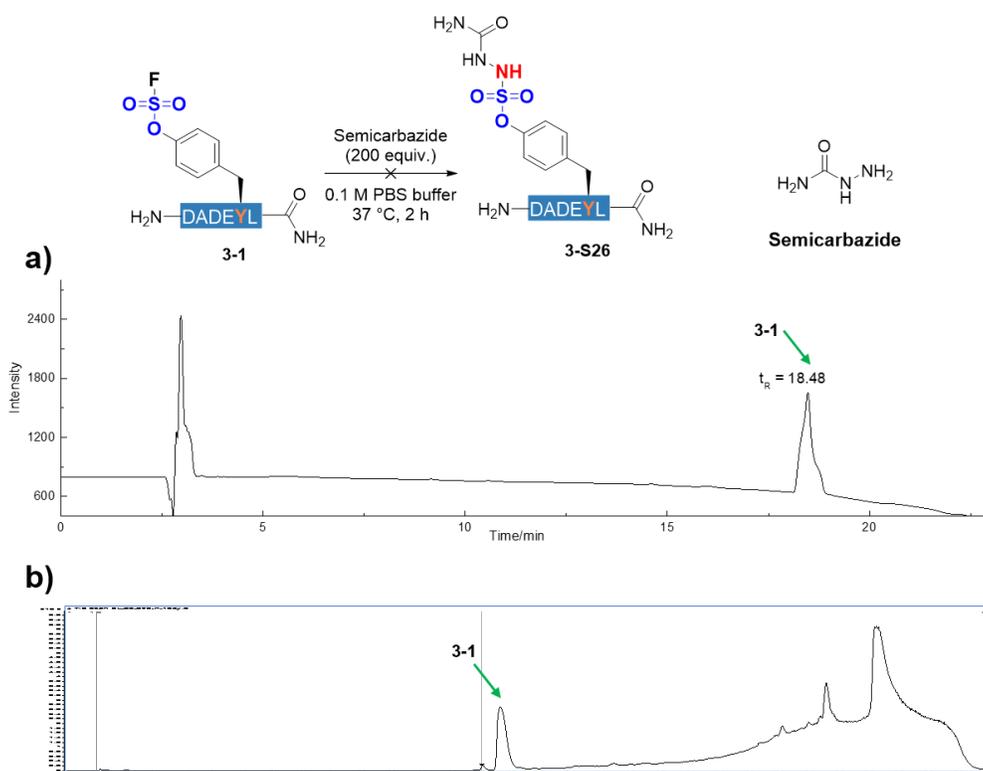


Figure 3-14. Compatibility of semicarbazide with fsY-containing substrates **3-1**. **a)** Analytical HPLC trace after incubating semicarbazide with **3-1** for 2 h. HPLC conditions: CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~33% CH₃CN from 0-20 min, 20-21 min 33%~95% CH₃CN, and 95% CH₃CN washing for another 2 min). All analytical RP-HPLC was monitored under 210 nm wavelength. **b)** Total ion chromatogram (TIC) of the reaction mixture after 1 day.

The sulfation patterns of tsetse fly anticoagulant peptide from tsetse thrombin inhibitor (TTI) have been found to play a critical role in its thrombin inhibitory activities.⁴⁵ We used the TTI peptides as a model system to probe the utility of the HA activators in controlling the bioactivities associated with sulfation under physiologically relevant conditions. First, TTI peptides consisting of fsY residues at position 9 and 12, **TTI02(fsY)**, **TTI03(fsY)**, and **TTI04(fsY)**, were subjected to the standard decaying conditions mediated by **3-7**. These reactions reached quantitative conversion within four hours, and produced the corresponding sulfopeptides **TTI02(sY)**, **TTI03(sY)**, and **TTI04(sY)** in 55%, 57%, and 35% isolated yield after HPLC purification, respectively (Figure 3-15).

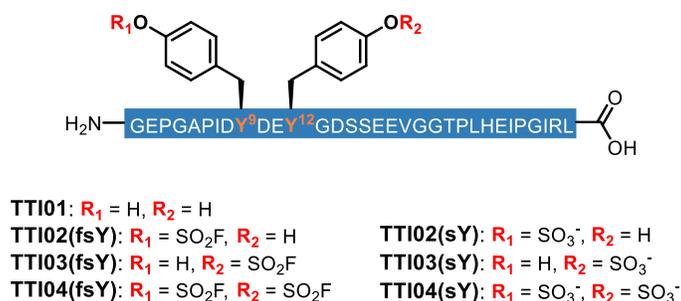


Figure 3-15. TTI peptide sequences and sulfation patterns

Next, we used a standard human α -thrombin activity assay with Chromozym TH as the substrate to determine the inhibitory effects of the fsY-containing TTI peptides (latent) and the sY-containing TTI peptides (active).⁴⁶⁻⁴⁷ Although the latent TTI peptides still exhibited minor inhibitory effects compared to the non-sulfated control **TTI01** ($K_i = 2873$ pM), with K_i values of 700 pM, 1547 pM, and 102 pM for **TTI02(fsY)**, **TTI03(fsY)**, and **TTI04(fsY)**, respectively, the active TTI peptides demonstrated significantly higher potencies, with K_i values of 82 pM, 43 pM, and 0.69 pM for **TTI02(sY)**, **TTI03(sY)**, and **TTI04(sY)**, respectively (Figure 3-16).⁴⁸ The latent TTI peptides that were decaged in situ by adding activator **3-7** into the assay all showed similar inhibitory effects as the purified active TTI peptides. These results confirmed that fluorosulfate can serve as an effective latent sulfate in peptides, and can be facilely decaged in aqueous solution at neutral pH.

The two orders of magnitude difference in the K_i values between the latent and active TTI peptides provided a large window for light-controlled decaging. For example, while the latent **TTI04(fsY)** remains inactive for thrombin inhibition at 3.7 nM in the presence of 2-nitrobenzyl protected activator **3-19** in dark, after irradiation thrombin activity was reduced to 21% (Figure 3-17). Notably, no change in thrombin activity was observed in the absence of the TTI peptide or the activator, suggesting that the reduction of the thrombin activity was caused by the inhibition by the newly decaged latent TTI peptide.

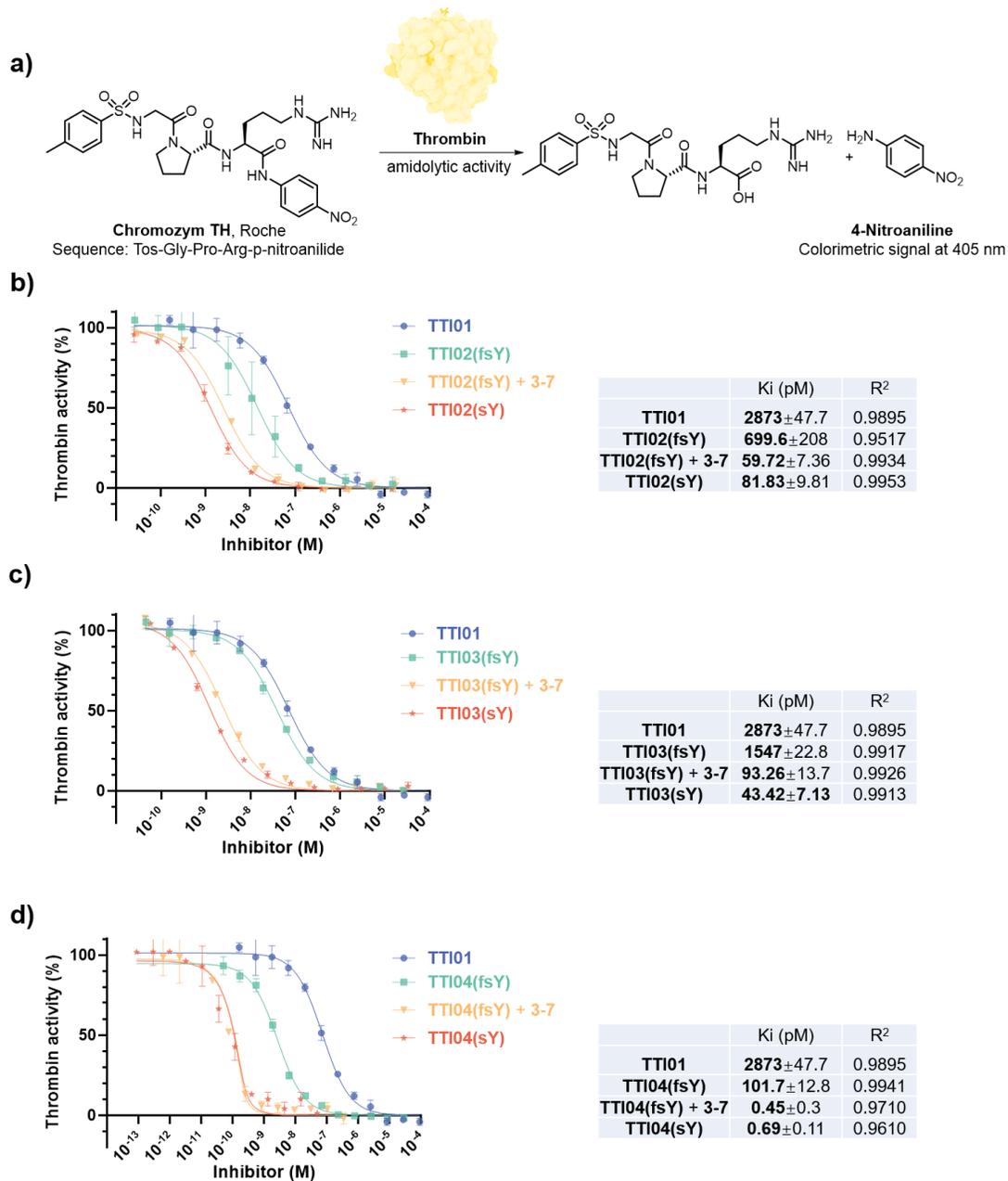


Figure 3-16. Thrombin inhibition by in situ degraded TTI peptides. **a)** Representative scheme of thrombin amidolytic activity with Chromozym TH was the substrate. **b)** Sequences of TTI peptides. **c)** Inhibition of thrombin by **TTI02(fsY)**, **TTI02(sY)**, and in situ degraded **TTI02(fsY)**. **d)** Inhibition of thrombin by **TTI03(fsY)**, **TTI03(sY)**, and in situ degraded **TTI03(fsY)**. **e)** Inhibition of thrombin by **TTI04(fsY)**, **TTI04(sY)**, and in situ degraded **TTI04(fsY)**.

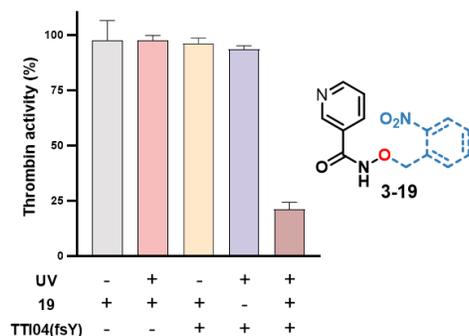


Figure 3-17. Light-mediated activation and decaging of fluorosulfate-containing TTI peptide **TTI04(fsY)** regulates its sulfation-dependent thrombin inhibitory activity

While sY has been successfully incorporated into bacterial⁶ and mammalian⁸ proteins via the ncAA mutagenesis strategy, incorporation of caged sY that enables controlled release of sulfate post-translationally remained elusive. The small size of fluorine atom allows fsY to be facilely incorporated into proteins as a ncAA.^{32,49} Following the procedure established by Wang *et al.*,³² we cloned the fsY-specific aminoacyl tRNA synthetase FsTyrRS and an optimal pyrrolysyl tRNA into plasmids for fsY incorporation into proteins. A sfGFP gene containing a TAG codon at position 151 was co-transformed along with the genes containing the FsTyrRA/tRNA pair into B95 E. coli cells. The targeted sfGFP-151-fsY was successfully expressed in a 12 mg/L yield. Tandem MS results verified the incorporation of fsY at the TAG-specified position-151 (Figure 3-18).^{32,50}

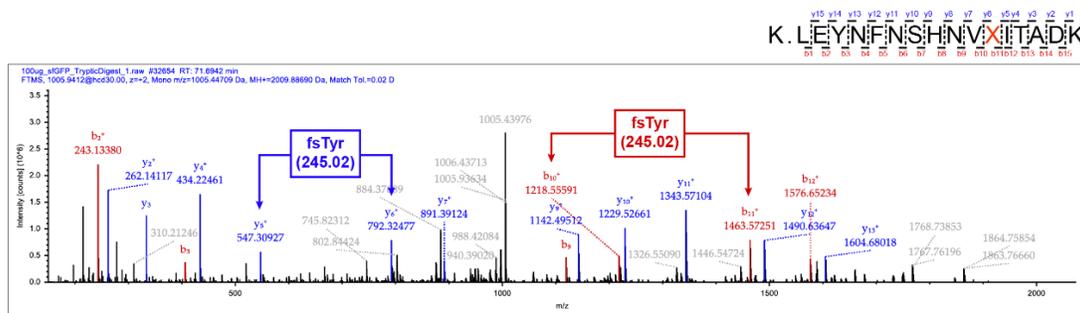


Figure 3-18. Tandem MS analysis of sfGFP-151-fsY after trypsin digestion and alkylation with iodoacetamide. X in the sequence logo represents fsY

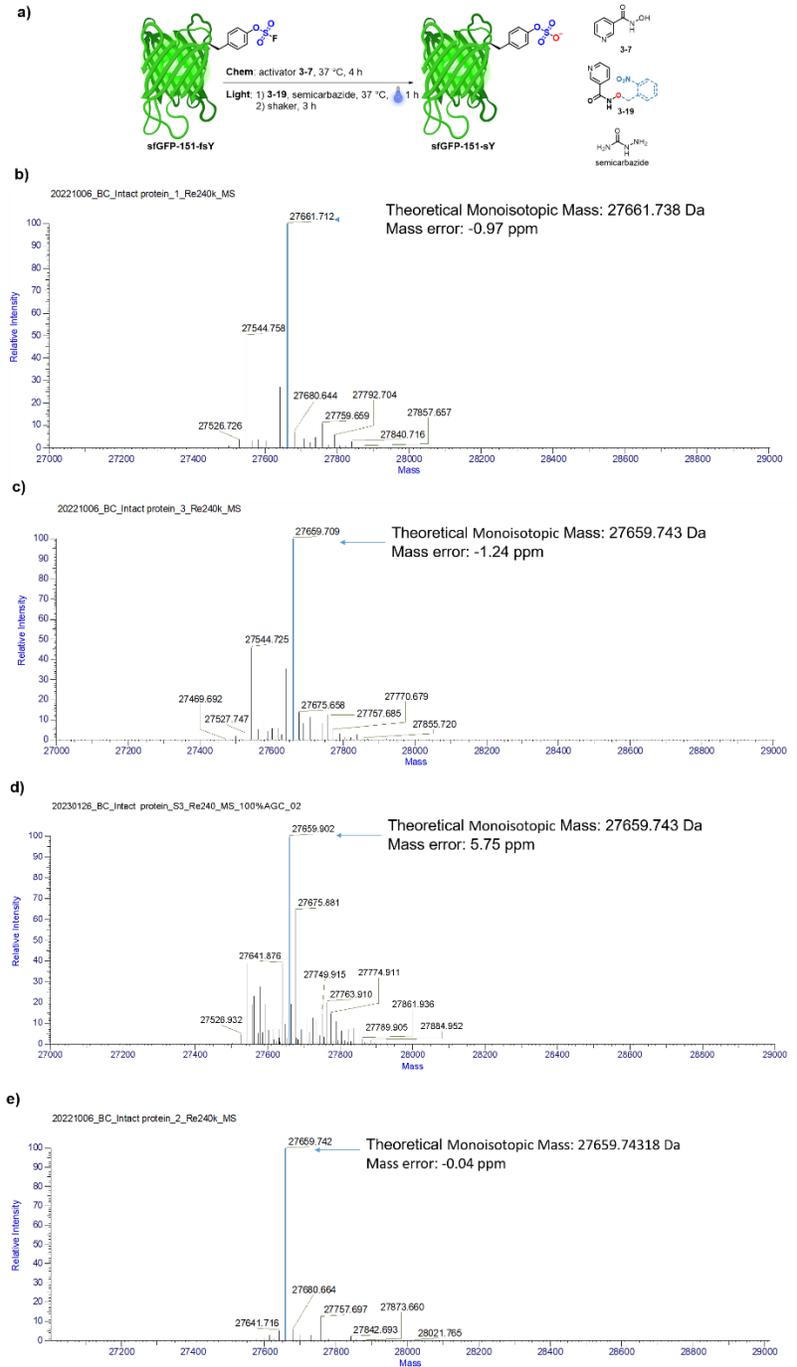


Figure 3-19. High-resolution Orbitrap mass spectrometry analysis of the decaging of sfGFP-151-fsY. **a)** Reaction scheme of sfGFP-151-fsY decaying using free activator **3-7** or photocaged activator **3-19**. **b)** Mass spectrum of purified sfGFP-151-fsY. **c)** Mass spectrum of sfGFP-151-sY, generated by in situ decaging of sfGFP-151-sY by **3-7**. **d)** Mass spectrum of sfGFP-151-sY, generated by in situ decaging of sfGFP-151-sY by **3-19** after it is exposed to 370 nm light. **e)** Mass spectrum of sfGFP-151-sY, expressed with sY incorporated by the ncAA mutagenesis method (expected exact mass 27659.743, observed 27659.742).

Next, to probe the transformation from fluorosulfate to sulfate in sfGFP-151-fsY, as well as the integrity of the resulting sulfoprotein, we performed whole protein intact mass analyses of sfGFP-151-fsY before and after decaging using high-resolution Orbitrap mass spectrometry, which is capable of achieving sub-5 ppm mass accuracy^{s1} and can confidently resolve the 1.996 Da mass shift after decaging (Figure 3-19 and 3-20). Indeed, the Orbitrap mass spectrometry confirmed the decaging of sfGFP-151-fsY (Figure 3-20, left) into sfGFP-151-sY (Figure 3-20, middle) by **3-7**.

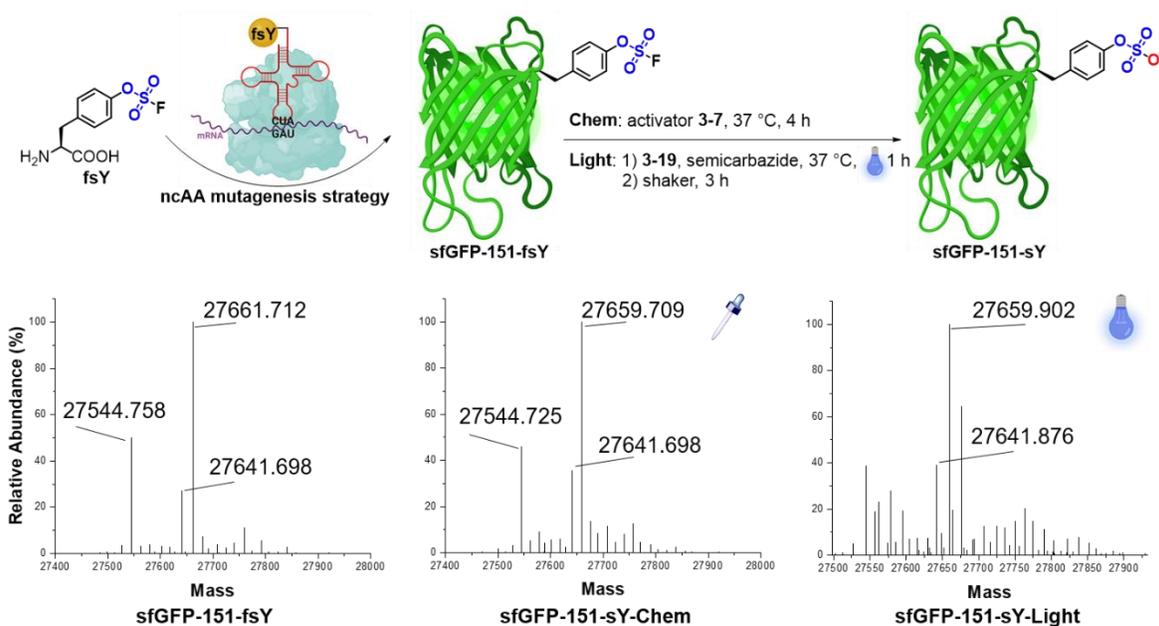


Figure 3-20. Fluorosulfate activation and decaging in fsY-containing protein sfGFP-151-fsY and its corresponding high-resolution mass spectrometry. Left: before the reaction, purified starting material sfGFP-151-fsY (expected exact mass: 27661.739 Da, observed mass: 27661.712 Da) was observed. Middle: after the reaction with free activator **3-7**, the substrate was fully converted and the decaged product sfGFP-151-sY-chem (expected exact mass: 27659.743 Da, observed mass: 27659.709 Da) was observed. Right: after the reaction with caged activator **3-19**, the substrate was fully converted and the decaged product sfGFP-151-sY-light (expected exact mass: 27659.743 Da, observed mass: 27659.902 Da) was observed.

Meanwhile, we also observed that the 27544.758 Da and 27544.725 Da mass peaks were corresponding to the misincorporation of glutamine into sfGFP at position of 151 (expected exact mass: 27544.804 Da), a known byproduct of the ncAA mutagenesis method.³⁶⁻³⁷ The detected 27641.698 Da and 27641.876 Da mass peak corresponds to an intramolecular reaction of fluorosulfotyrosine with nucleophilic residue (expected exact mass: 27641.732). Using this technique, we also confirmed the light-mediated decaging of sfGFP-151-fsY by the photocaged activator **3-19** (Figure 3-20, right), highlighting the potential of our approach for the spatiotemporal release of caged sulfoproteins. The compatibility of the releasing conditions with proteins was also confirmed by Western blot experiment (Figure 3-21).

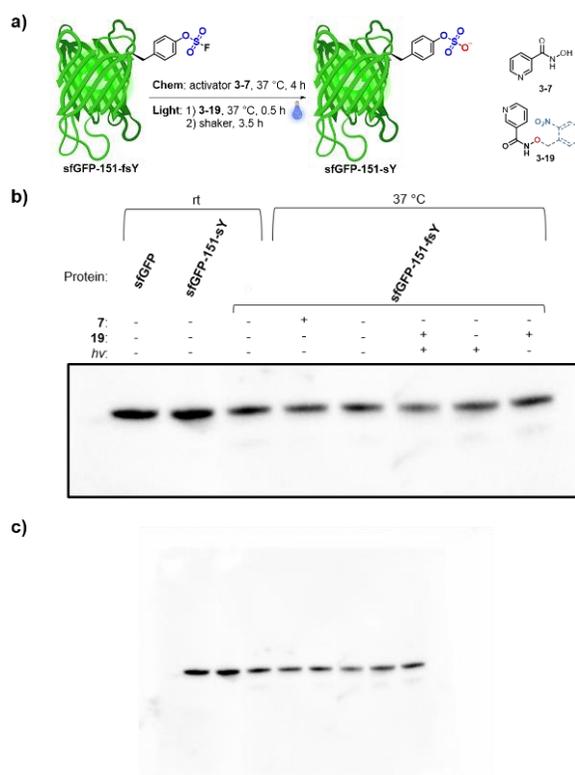


Figure 3-21. sfGFP-151-fsY and sfGFP-151-sY remained intact in decaging conditions by Anti-His western blot. **a)** Reaction setup for decaging fsY-containing sfGFP (**fsGFP-151-fsY**) in which condition **Chem** with free activator **7** and condition **Light** with photocaged activator **19**. **b)** Anti-His western blot of above reaction mixture with purified wild type sfGFP, sfGFP-151-fsY and sfGFP-151-sY as control. **c)** Original Anti-His western blot image.

Last, we tested the cytocompatibility of the fluorosulfate decaging reagents. Previously, cesium carbonate (Cs_2CO_3)/ethylene glycol³³ or 2 M ammonium acetate (NH_4OAc) aqueous solution^{24, 43} was used to remove the protecting groups for sulfate in peptides and small molecules. However, these conditions were found to be strongly denaturing to proteins (Figure 3-22) and highly toxic to live cells (Figure 3-23 and 24). In contrast, our reagents caused no protein denaturation and has low toxicity to cells at various concentrations.

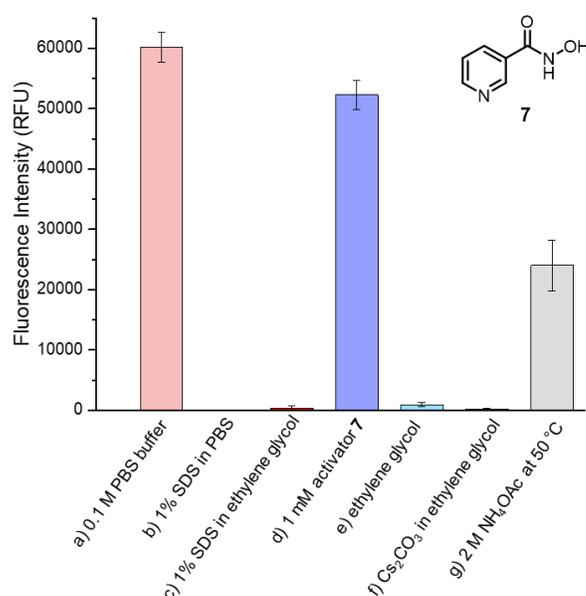


Figure 3-22. Fluorescence intensity of sf-GFP-151-fsY after being treated with various conditions. Either ethylene glycol alone or Cs_2CO_3 in ethylene glycol resulted in complete protein denaturation. 2 M NH_4OAc also caused partial denaturation of sfGFP. In contrast, reagent **7** did not cause significant protein denaturation. Conditions a and d to f: 4 μL of the sfGFP-151-fsY (10 $\mu\text{g}/\mu\text{L}$) was mixed with corresponding reagents (final volume is 50 μL) and placed on the shaker at 37 °C for 1 h under the protection of aluminum foil. Conditions b and c: negative control with complete denaturing conditions. 4 μL of the sfGFP-151-fsY (10 $\mu\text{g}/\mu\text{L}$) was mixed with 1% SDS in 0.1 M PBS buffer or ethylene glycol and placed in the incubator at 95 °C for 5 min. Condition g: the temperature was set as 50 °C, 1 h.

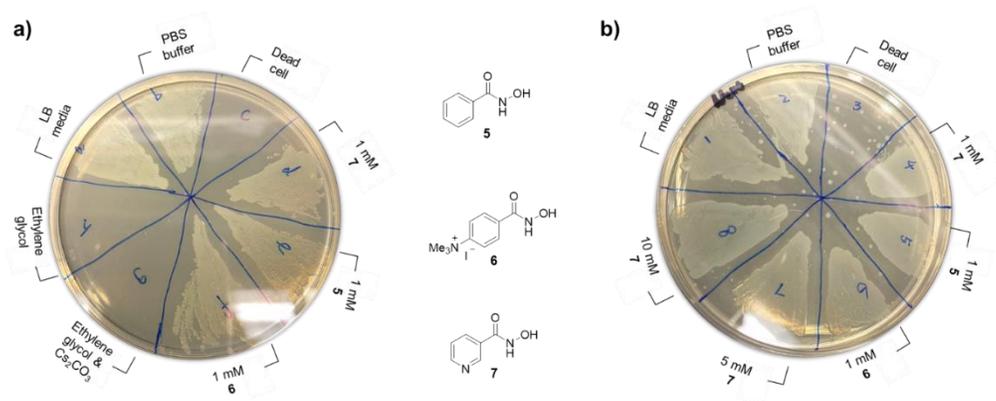


Figure 3-23. Plating assay to determine *E. coli* (BL21) viability under various fluorosulfate decaging conditions. **a)** Either ethylene glycol alone or Cs_2CO_3 in ethylene glycol resulted in complete cell death when used to decaging fluorosulfate. In contrast, HA reagents **5-7** demonstrated low cytotoxicity. **b)** Both 1 mM and 10 mM of HA reagents **5-7** were well tolerated by *E. coli*. 45 μL *E. coli* cell culture (BL21) was mixed with 5 μL corresponding stock solution and incubated at 37 $^\circ\text{C}$ for 1 h. The negative (dead cell) control was generated by treating the cells at 90 $^\circ\text{C}$ for 1 h. Subsequently, 5 μL of each above reaction mixture was streaked on the LB-Agar plate and placed in 37 $^\circ\text{C}$ overnight.

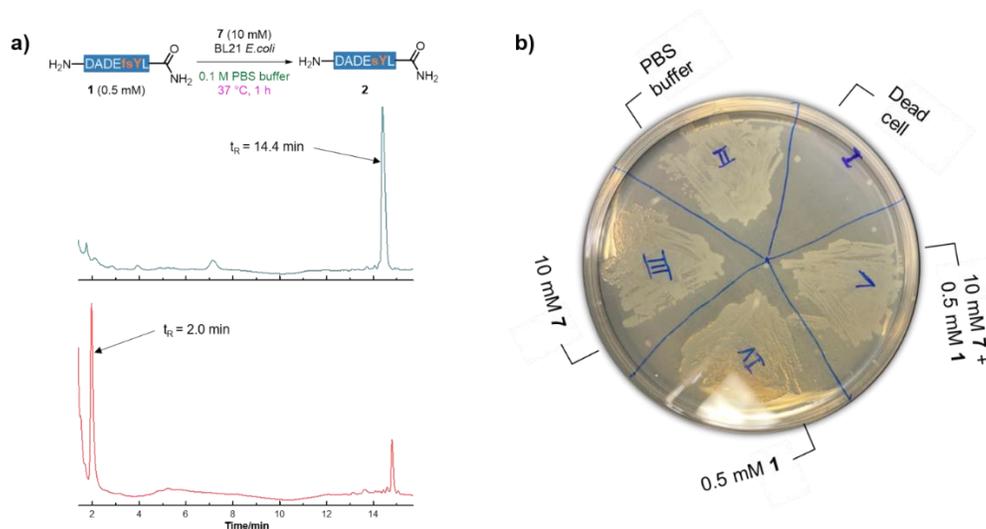


Figure 3-24. Free peptide **1** decaging in the presence of live *E. coli* cells to mimic the soluble sulfated small molecules and peptides.⁴ **a)** The HPLC traces of peptide **1** before (top) and after (bottom) decaging in the presence of live *E. coli* cells, monitored by the UV absorbance at 220 nm. In the presence of *E. coli* cells, **1** was still smoothly converted to the sulfopeptide **2** by reagent **7** in 81% yield. **b)** Plating assay to determine the viability of BL21 *E. coli* cells after the decaging reaction of peptide **1**. After the decaging reaction, the *E. coli* cells showed no sign of reduced viability or growth.

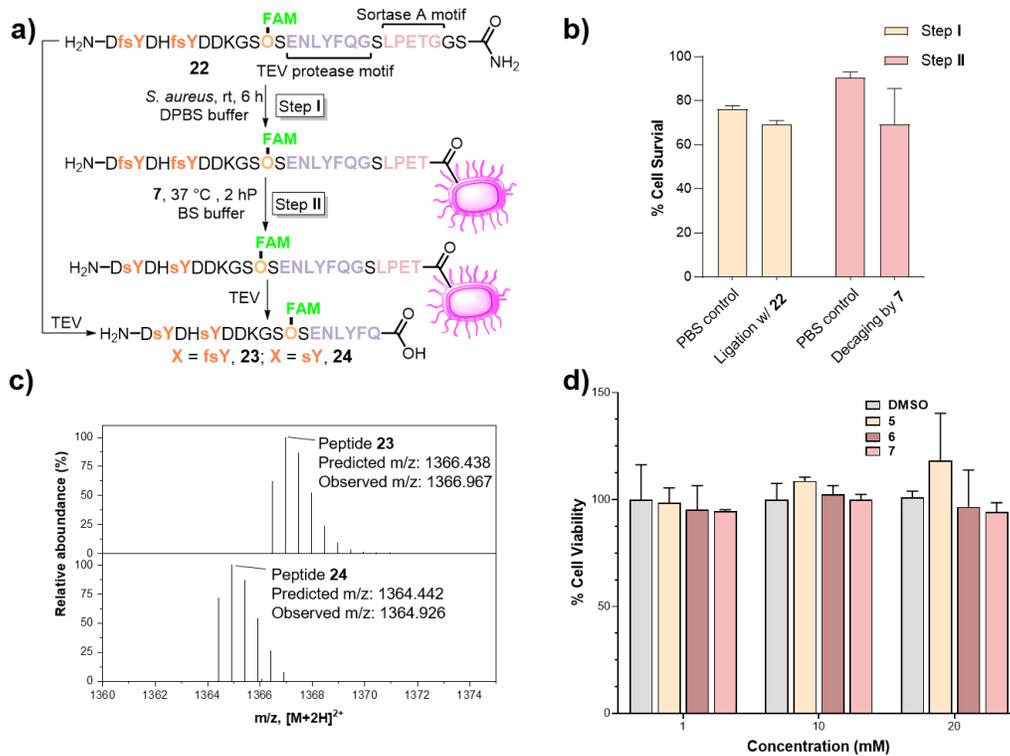


Figure 3-25. Cytopatibility of the reagents. **a)** Sortase A-mediated ligation of peptide 3-20 onto the *S. aureus* cell surface and its decaying followed by the TEV protease cleavage. **b)** Percent of *S. aureus* cell survived after sortase A-mediated ligation of 3-20 (Step I) and after fluorosulfate decaying by 3-7 (Step II) compared to the cells treated with PBS. The average data of two trials were plotted. **c)** LC-MS analysis of samples after the TEV cleavage identified the degraded peptide (3-21, bottom) compared to the cleaved peptide before decaying (3-22, top). **d)** MTT assay of the mammalian HEK-293T cells after incubation with various concentrations of reagents 3-5, 3-6, and 3-7. The average data of three trials were plotted.

To mimic the cell membrane-bound sulfoproteins,⁵² we examined *in situ* fluorosulfate decaying on the surface of live *Staphylococcus aureus* (*S. aureus*) cells (Figure 3-25). *S. aureus* cells were chosen because there are no known endogenous sulfopeptides expressed on their surface, and the endogenous sortase A on their surface can be used to ligate peptides.⁵³⁻⁵⁴ A fluorescently labeled peptide 3-20 consisting of a *Tobacco Etch Virus* (TEV) protease cleavage sequence⁵⁵ and a LPETG sortase A-recognition motif was ligated to the cell surface of *S. aureus* (Figure 3-25a, 3-26 and 3-27).

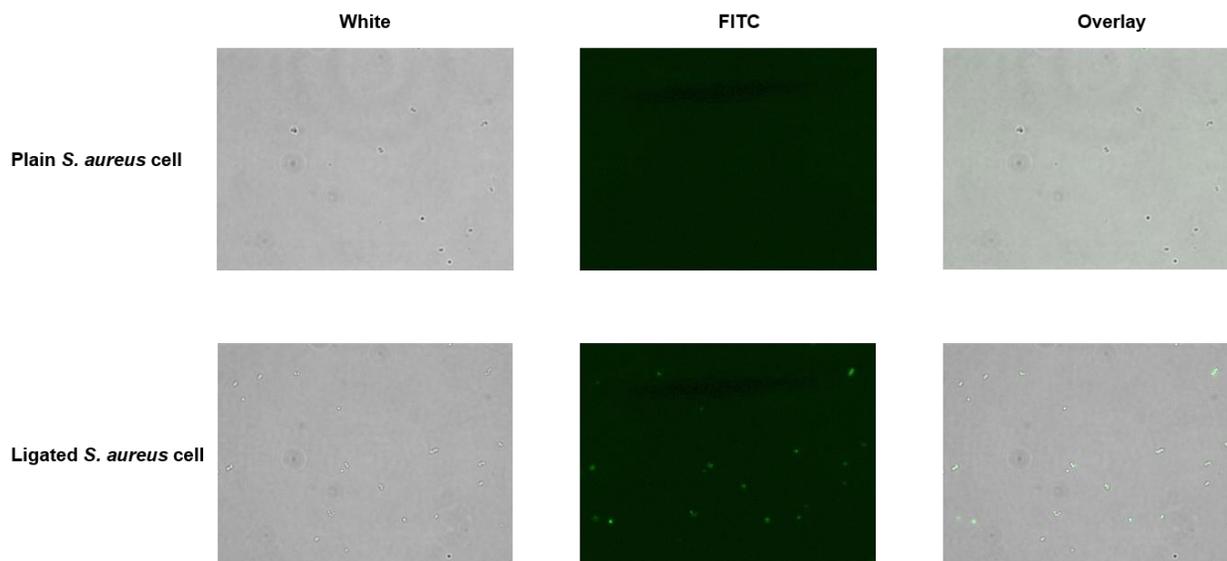
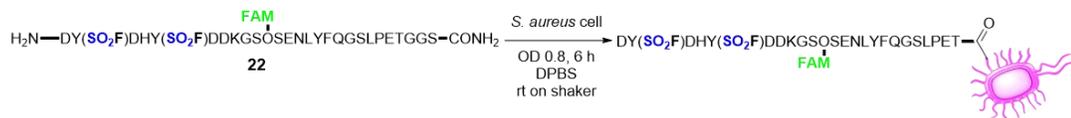


Figure 3-26. Fluorescent micrograph of the *S. aureus* cells before and after ligation of the 5(6)-carboxyfluorescein (FAM)-labeled peptide 3-20.

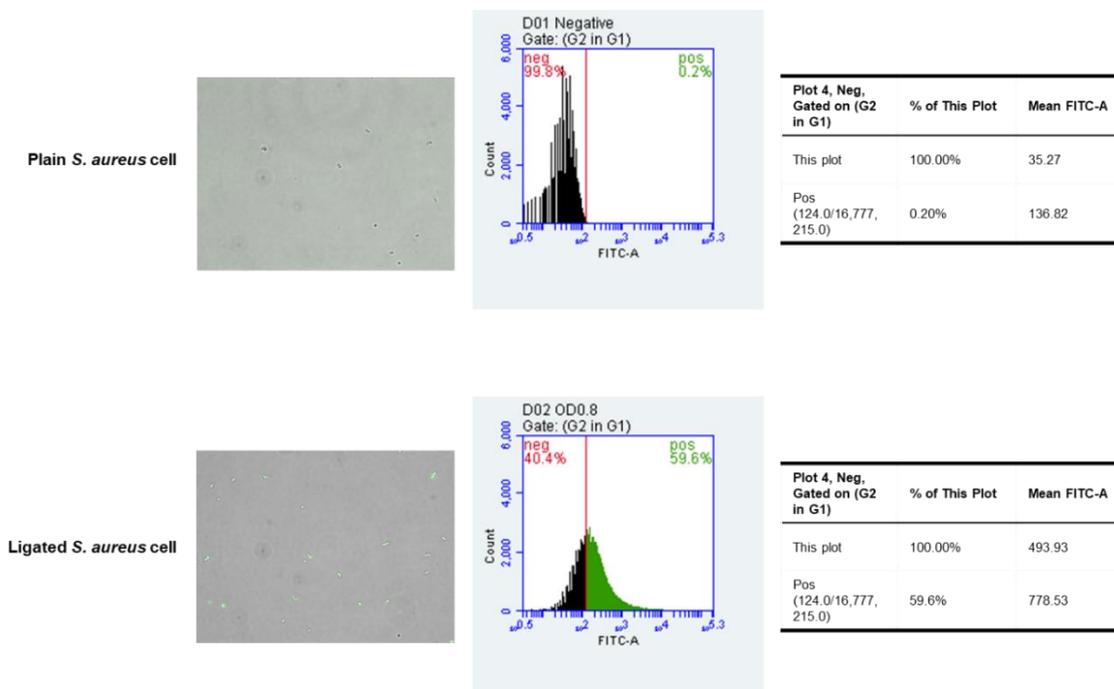


Figure 3-27. Flow cytometry of the *S. aureus* cells before and after ligation of the FAM-labeled peptide 3-20. Fluorescent images of cell populations are shown on the left as reference.

The cell surface-ligated peptide was then decaged by reagent **3-7**. Compared to the phosphate-buffered saline (PBS) buffer control, neither the cell surface ligation nor the fsY decaging experiments caused significant reduction of cell viability (Figure 3-25b). LC-MS analysis of the peptide residues cleaved after the decaging reaction (**3-21**) confirmed that the fsY were successfully converted into sY on live cell surface (Figure 3-25c). Finally, reagents **3-5**, **3-6**, and **3-7** also exhibited low toxicity to mammalian cells even at millimolar concentrations based on the MTT assay (Figure 3-25d).⁵⁶

3.3 CONCLUSION

In conclusion, we demonstrated that fluorosulfate is a physiologically compatible latent sulfate in peptides and proteins. Fluorosulfate is stable in neutral aqueous buffers, cell lysates, and serum, and can be efficiently converted into sulfate by HA-derived activators under physiologically relevant conditions via a Lossen rearrangement pathway. Combined with the facile incorporation of fluorosulfate-containing amino acid fsY via solid-phase peptide synthesis and ncAA mutagenesis approaches, our reported approach can be applied to studying a wide range of sulfopeptides and sulfoproteins in their physiological states. Moreover, the easily modified, readily accessible, and diverse HA derivatives provide a vast playground for future studies on spatiotemporally controlling the functions of sulfated molecules *in vitro* and *in vivo*.

3.4 EXPERIMENTAL

3.4.1 Material and Equipment

Chemicals were purchased from vendors such as MilliporeSigma, Thermo Fisher Scientific, and Chemical Industry Co., Ltd. (TCI) and used as received. Deuterated solvents were purchased from either Cambridge Isotope Laboratories, Inc. or Thermo Fisher Scientific. Sulfuryl fluoride (Vikane) was purchased from SynQuest Lab. Inc. The organic solvents such as acetonitrile (MeCN), tetrahydrofuran (THF), dichloromethane (DCM), and dimethylformamide (DMF) were purchased from Thermo Fisher Scientific and used after the purification by a dry solvent purification system from Pure Process Technology (PPT). 0.1 M phosphate buffered saline (PBS) buffer was prepared from purchased 20X Modified Dulbecco's PBS buffer (thermos scientific, Lot WE325396) with the pH adjusted to 7.60 by 1 M hydrogen chloride and 1 M sodium hydroxide. Human serum (Fisher BioReagents, Lot: 206088) was used as purchased with subtle pH adjustment. Thin layer chromatography was performed on Merck TLC plates (silica gel 60 F254) and visualized by UV irradiation (254 nm) and by charring with a cerium molybdate solution (0.5 g $\text{Ce}(\text{NH}_4)_2(\text{NO}_3)_6$, 24 g $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24}\cdot 4\text{H}_2\text{O}$, 500 mL H_2O , 28 mL H_2SO_4 , filtered if necessary). Silica gel chromatography was carried out using an automated flash chromatography (Biotage Isolera One Flash Chromatography Instrument). The reaction Schlenk bottles were flame dried before use. Labconco FreeZone 2.5 L -84C benchtop freeze dryer was used for lyophilization. Gyros Protein Tribute peptide synthesizer was employed for automated peptide synthesis.

3.4.2 Characterization

^1H NMR, ^{13}C NMR measurements were conducted in deuterated solvents like deuterated chloroform (CDCl_3), deuterated oxide (D_2O), deuterated methanol (CD_3OD), and deuterated dimethyl sulfoxide (DMSO-d_6) using a Varian Gemini-600 (600 MHz), Varian Inova-500 (500 MHz)

NMR, or Bruker Advance Neo 500 MHz (with Helium CryoProbe) spectrometer. Chemical shifts are in ppm calibrated using the resonances of the residual carbon and proton of the deuterated solvent. High-resolution mass spectrometry was performed on a Micromass LCT ESI-MS and JEOL AccuTOF Dart (positive mode) at the Boston College Mass Spectrometry Facility. Mass-spec data for the peptides was generated using an Agilent 6230 LC-TOF mass spectrometer with an Agilent InfinityLab Poroshell 120 column (2.7 μm , 2.1 x 50 mm, for positive mode) and a Phenomenex Luna 3 μm HILIC 200 \AA column (2 x 50 mm, for negative mode). The following methods were used to analyze the pure samples and real-time reaction monitoring in the LC-MS instrument. High resolution mass spectrometry of full protein was performed on a Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Fisher Scientific, San Jose, California, USA) coupled with a TriVersa NanoMate (Advion, Ithaca, NY).

Method A for positive mode, solvent A contains acetonitrile/water/formic acid = 5:95:0.1, solvent B contains acetonitrile/water/formic acid = 95:5:0.1.

| Time (min) | A (%) | B (%) | Flow (mL/min) |
|------------|-------|-------|---------------|
| 0.00 | 95.0 | 5.0 | 0.200 |
| 5.00 | 95.0 | 5.0 | 0.200 |
| 20.00 | 5.0 | 95.0 | 0.200 |
| 25.00 | 5.0 | 95.0 | 0.200 |
| 26.00 | 95.0 | 5.0 | 0.200 |
| 31.00 | 95.0 | 5.0 | 0.200 |

Method B for negative mode, solvent A contains 5 mM ammonium acetate in water, solvent B contains acetonitrile.

| Time (min) | A (%) | B (%) | Flow (mL/min) |
|---------------|----------|----------|------------------|
| 0.00 | 5.0 | 5.0 | 0.200 |
| 2.00 | 5.0 | 95.0 | 0.200 |
| 20.00 | 90.0 | 10.0 | 0.200 |
| 25.00 | 90.0 | 10.0 | 0.200 |
| 26.00 | 10.0 | 90.0 | 0.200 |
| 31.00 | 10.0 | 90.0 | 0.200 |

Analytical RP-HPLC was performed on an Agilent 1100 Series HPLC system with manual injection. Analytical RP-HPLC was done with a Waters XBridge Peptide BEH C18 column (2.5 μm , 4.6 mm x 150 mm) using a 0.5 mL/min flow rate. Semi-preparative RP-HPLC was achieved using a Waters 1525 binary HPLC pump equipped with a Waters 2489 UV/Visible detector. Semi-preparative HPLC was conducted with a Agilent Prep-C18 column (5 μm , 10 mm x 250 mm) using a 5 mL/min flow rate.

3.4.3 General Procedures

General Procedure A. Manual synthesis of fluorosulfotyrosine (fsY)-containing peptides

Fluorenylmethyloxycarbonyl-solid phase peptide synthesis (Fmoc-SPPS) was performed manually using a 50 mL Syntheware peptide synthesis vessel with fritted disc and T-bore polytetrafluoroethylene (PTFE) stopcock employing Rink amide resin (0.49 mmol/g) at normally 100 μ mol scale. The resin was preswollen in DMF for 330 min before use. Amino acids were coupled using the following reagents and reaction time: Fmoc-AA-OH (5 equiv.), HBTU (5 equiv.), DIPEA (5 equiv.)-preactivation was done in a 20 mL vial with 3 mL DMF and a clear reaction mixture would be obtained within 2 min which was directly transferred to the peptide synthesizer vessel. The activated amino acid was added to the resin-bound primary amine with nitrogen gas bubbling for a coupling period of 45 min for natural amino acid or 60 min for fluorosulfotyrosine to generate a new amide bond. Removal of the Fmoc group was achieved using 20% (volume percent) 2-methylpiperidine in DMF (3 x 10 min). The resin was washed with DMF between each coupling and Fmoc-removal step (5 x 3 min). After the final Fmoc deprotection step, the resin was washed with DMF (5 x 3 min), dichloromethane (3 x 3 min), and hexane (3 x 3 min) and dried in air. Cleavage of the crude fsY-containing peptides from the resin and side chain unmasking (except for the fluorosulfate group) was achieved with a cleavage cocktail consisting of TFA:TIPS:H₂O (95:2.5:2.5) for three times and 30 min each. The combined filtrates were concentrated carefully under reduced pressure at 40 °C which were followed by precipitation in cold diethyl ether. The resulting suspension was centrifuged at 5000 rpm at 4 °C for 10 min before dumping the ether. Obtained pellet was redissolved in water and filtered for LC-MS verification. The crude was then subjected for semi-Prep HPLC purification using solvent A (95% CH₃CN, 5% H₂O, and 0.1% TFA): solvent B (5% CH₃CN, 95% H₂O, and 0.1% TFA) as mobile

phase. Appropriate fractions were characterized by LC-MS, combined, concentrated by air stream, and lyophilized. The obtained product was analyzed by HPLC and LC-MS for confirmation.

General Procedure B. Automated synthesis of peptides

Fmoc-SPPS of peptides without the fsY incorporation were performed automatically on the Gyros Protein Tribute peptide synthesizer employing Rink amide resin (0.49 mmol/g) at normally 100 μ mol scale. The resin was loaded in the 10 mL reaction vessel. Subsequently, each amino acid (5 equiv.) mixed with HBTU (5 equiv.) was weighed into the 10 mL vial and sealed.

The working sequence was set as follows:

Swelling with 3 mL DMF for 30 min and then 2 mL DMF washing (6 x 30 sec). 3 mL 20% piperidine solution in DMF for Fmoc removal (3 x 5 min) and 3 mL DMF washing (6 x 30 sec).

Coupling, NMM (0.4 M in DMF, 2.5 mL) was injected into the amino acid and HBTU mixture and then delivered to the reaction vessel. The resulting reaction mixture was shaken for 40 min, followed by the addition of 3 mL 20% piperidine solution in DMF for Fmoc removal (2 x 5 min) and 2 mL DMF washing (6 x 30 sec).

Final Fmoc-removal, 3 mL 20% piperidine solution in DMF for Fmoc removal (8 x 5 min) and 2 mL DMF washing (9 x 30 sec).

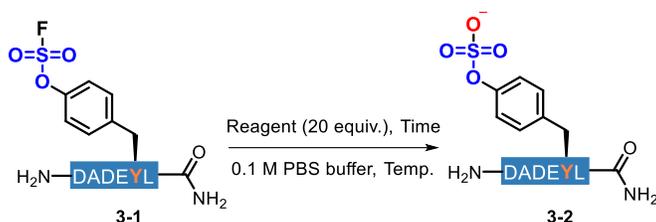
The reaction vessel was next taken out of the machine, washed by 3 mL dichloromethane (3 x 3 min) and hexane (3 x 3 min), and dried in air. The obtained beads with peptides attached on could be used for subsequent fsY incorporation following above procedure. Once the beads were ready for cleavage, TFA:TIPS:H₂O (95:2.5:2.5) cocktail was charged for three times and 30 min each. The combined filtrates were concentrated carefully under reduced pressure at 40 °C which were followed by precipitation in cold diethyl ether. The resulting suspension was centrifuged at 5000 rpm at 4 °C

for 10 min before dumping the ether. Obtained pellet was redissolved in water and filtered for LC-MS verification. The crude was then subjected for semi-Prep HPLC purification using solvent A (95% CH₃CN, 5% H₂O, and 0.1% TFA): solvent B (5% CH₃CN, 95% H₂O, and 0.1% TFA) as mobile phase. Appropriate fractions were characterized by LC-MS, collected, concentrated by air stream, and lyophilized. The obtained product was analyzed by HPLC and LC-MS for confirmation.

Yield Calculation

Yield of purified fsY-containing peptides were calculated from the adjusted resin loading and based on the assumption that all basic side chains, and *N*-terminal amines in the peptides were lyophilized as salts with TFA.

General Procedure C. Activating System Evolvement for Fluorosulfohexpeptide 1 Decaging



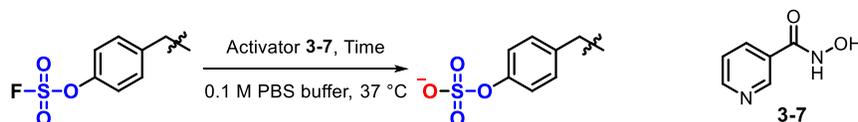
Stock solution preparation and reaction setup:

Fluorosulfohexpeptide **3-1** (TFA complex from Semi-Prep RP-HPLC) was dissolved in 0.1 M PBS buffer. The final concentration of **3-1** was 1 mM in 0.1 M PBS buffer, and pH was tested to be between 7.20~7.40. For reagents like TMG and NHS or activators like **3-3**, **3-5** and **3-7** with good aqueous solubility, they were dissolved in 0.1 M PBS buffer and the final pH was adjusted by 1 M HCl/1 M NaOH to give 0.1 M final aqueous solution with pH maintained between 7.0~7.40. For phenols like PFP and 2-HP or activators like **3-S4** and **3-S17** with poor aqueous solubility, they were

dissolved in DMSO to give 0.1 M corresponding stock solution and used directly in the reaction. The pH of the final reaction mixture was also confirmed to be at the physiological pH range 7.00~7.40.

A 1.5 mL glass vial was equipped with a stirring bar and **3-1** (1 mM, 150.0 μ L, 0.15 μ mol). Meanwhile, the stock solutions of reagents and activators (0.1 M) were also placed at the same temperature (25 $^{\circ}$ C for Table 1, entry 3-7; 37 $^{\circ}$ C for Table 1, entry 8 and Figure 2a, S4) for at least two minutes. Subsequently, necessary reagent and/or activator stock solution (30.0 μ L, 20 equiv.) was added to the stirring vial. The resulting reaction mixture was allowed to stir at the same conditions for certain reaction time (24 h for Table 1; 1 h or 0.5 h for Figure 2a and S4). Then the whole reaction mixture was filtered and subjected for analytical RP-HPLC.

General Procedure D. Decaging of fsY-Containing Peptides



A solution of fsY-containing peptides (1 mM) and 0.1 M 3-pyridinylhydroxamic acid (**7**) both in 0.1 M PBS buffer was warmed up separately in the warm room (37 $^{\circ}$ C constant temperature) for 2 min. Next, 20 equiv. per fluorosulfate moiety of activator **3-7** solution was added to the substrate solution with stirring. The reaction was monitored by RP-HPLC until the reaction was complete. Then the reaction was filtered and loaded directly on the Semi-Prep RP-HPLC for purification with CH₃CN:20mM ammonium acetate buffer as mobile phase. Appropriate fraction was verified by LC-MS positive mode, combined, concentrated with air stream. The resulting residue was redissolved in water and lyophilized. The product was confirmed by LC-MS and RP-HPLC. Note: pH of the stock solutions of substrates and activators in 0.1 M PBS buffer were always adjusted between 7.00~7.40.

Yield Calculation

Yield of purified sY-containing peptides were calculated from the adjusted substrate loading and based on the assumption that all acidic side chains in the peptides were lyophilized as salts with ammonium acetate.

3.4.4 Detailed Procedure for Supplementary Results

Stability test of fluorosulfotyrosine (fsY) and fluorosulfohexpeptide 1 stability in cell lysate

Cell lysate preparation and reaction setup:

HEK-293T cells were split in two different plates (100 mm dish) in adherent culture and grown in 37 °C incubator. the cell monolayer was gently washed by PBS buffer after 48 h. Cell lysate prepared by both chemical and physical methods. The chemical method involves homogenizing cells in RIPA buffer. Cell debris was removed centrifugation to afford a Chem cell lysate (pH = 7.01). In a similar manner, cell lysate can also be prepared by physically disrupting the cell membrane via 4 min sonication (rest for 10 s for every minute) in 4 °C cold room to afford a Phy cell lysate (pH = 7.41). Human serum (1.00 mL) pH was adjusted to 7.40 by 0.1 M PBS buffer (pH = 3.17, 60.0 μ L) and used without any further treatment.

In a 200 μ L Eppendorf tube, freshly prepared stock solution of fsY (50 mM in DMSO, 2.0 μ L) or **1** (50 mM in DMSO, 2.0 μ L) was added to the cell lysis solution (98.0 μ L). The resulting mixture was sealed and placed on the shaker in the 37 °C constant warm room. Meanwhile, a parallel control with only cell lysate was also set up at the same conditions. After 48 hours of incubation, the control cell lysate was directly transferred into the Microcon-10 kDa centrifuge filter and filtered with 15000 rpm for 15 min at 4 °C. For the reaction containing the substrates, a stock solution of tryptophan (Trp, 50 mM in DMSO, 2.0 μ L, well-mixed before use) was added as internal standard and flicked to mix them well. The resulting mixture was transferred for centrifuge filtration with Microcon-10kDa

centrifuge filter with 15000 rpm for 15 min at 4 °C. The filtrate was subjected for analytical RP-HPLC and high-resolution LC-MS (Figure S1).

The pH dependence of the fsY decaging reaction mediated by NHS

Stock solution preparation and reaction setup:

Substrate **3-1** was dissolved in 0.1 M PBS buffer at 1 mM, pH = 6.70. NHS was weighted and dissolved in small amount of 0.1 M PBS buffer, and the final pH and concentration were adjusted to 5.66 to 12.42 by 1 M HCl and NaOH.

In a 1.5 mL vial with a stirring bar, stock solution of **3-1** (1 mM, 150.0 μ L, 0.15 μ mol) and NHS (0.1 M, 30.0 μ L, 30 μ mol) was added subsequently. The resulting reaction mixture was allowed to stir at the 37 °C for 2 h and its pH was measured then filtered, and subjected for analytical RP-HPLC.

Hydrolysis of NHS in deuterated water by titrating NaOH

Reaction set up:

In a 4 mL glass vial equipped with a stirring bar, NHS (10.9 mg, 0.094 mmol) was dissolved in deuterated water (D_2O , 932.9 μ L) at room temperature. A stock solution (10 M in D_2O , 21.4 μ L, 1.0 equivalent) was added and the resulting reaction mixture was allowed to stir at the same temperature for 1 h. Then the reaction was transferred for NMR tests. Same set up for 1.5, 2.0, and 5.0 equivalents of NaOH. Detected pD = pH + 0.4, and pH was obtained from the pH meter.⁵⁷

Decaging of fsY in **1** mediated by silyl-protected activator

Stock solution preparation and reaction setup:

Stock solution of the substrate and the silyl-protected activator were freshly prepared. Fluorosulfohexpeptide **3-1** was dissolved in 0.1 M PBS buffer (1 mM, pH = 7.23). The reagent **3-11** were dissolved in DMSO to afford 0.1 M stock solution of it. Potassium fluoride was dissolved in 0.1 M PBS buffer to give 1.0 M stock solution of it.

In a 1.5 mL vial with a stirring bar, stock solution of **3-1** (0.15 μ mol, 150.0 μ L) was added and warmed up in the 37 °C for 2 min. Simultaneously, stock solutions of the silyl-protected activators **3-11** and additives KF were also warmed up to the same temperature. When only the reagent **3-11** (0.1 M, 30.0 μ L, 20 equiv.) was added, there's no trace of the released product (Table S4, entry 1). When both the reagent **11** (0.1 M, 30.0 μ L, 20 equiv.) and additive potassium fluoride (KF, 1.0 M, 3.0 μ L) were added, a 41% yield of **2** was observed in 30 minutes.

Compatibility test of the decaging reaction of **1** in the presence of free amino acids

Stock solution preparation and reaction set up:

Fluorosulfohexpeptide **3-1** (3.3 mg) was dissolved in 0.1 M PBS buffer (pH = 7.64, 3.60 mL) to provide 1 mM stock solution of **3-1**, pH = 7.31. Activator **3-7** was dissolved in 0.1 M PBS buffer with 1 M HCl and NaOH for pH adjustment to afford 0.1 M stock solution of **3-7**, pH = 7.39. All the free amino acids except tyrosine was dissolved in 0.1 M PBS buffer with 1 M HCl and NaOH solution for pH adjustment to give 0.1 M final concentration of the free amino acids and pH was maintained between 7.0-7.4. Tyrosine showed a very poor solubility even in DMSO. A suspension of tyrosine (0.1 M in DMSO) was made and used directly in the reaction.

To a solution of **3-1** in 0.1 M PBS buffer (1 mM, 150.0 μL) were added activator **3-7** (0.1 M, 30.0 μL , 20 equiv.) and free amino acid stock solution (0.1 M, 30.0 μL , 20 equiv.) in the warm room at 37 °C. The resulting reaction mixture was allowed to stir for 30 min, filtered, and subjected for analytical HPLC.

LC-MS monitoring of H_2^{18}O isotope labeling experiment

Stock solution preparation and reaction setup:

0.1 M PBS buffer (H_2^{18}O) was freshly prepared with 10.2 mg of Na_2HPO_4 and 3.5 mg of NaH_2PO_4 in H_2^{18}O (97 atom %). 0.1 M stock solution of **3-1**, **3-2**, and **3-7** in DMSO was also prepared.

To a 1.5 mL vial equipped with a stirring bar and 0.1 M PBS buffer (H_2^{18}O , 177.1 μL) was added stock solution of **3-1** in DMSO (0.1 M, 1.0 μL) at 37 °C. The resulting reaction mixture was allowed to stir for 30 min, filtered, and subjected for LC-MS analysis.

Similarly, to a 1.5 mL vial equipped with a stirring bar and 0.1 M PBS buffer (H_2^{18}O , 177.1 μL) were added stock solution of **3-2** and **3-7** in DMSO (0.1 M, 1.0 μL) in the warm room at 37 °C. The resulting reaction mixture was allowed to stir for 30 min, filtered, and subjected for LC-MS analysis.

Compatibility of semicarbazide with fsY-containing substrates 3-1

Reaction set up:

To a solution of **3-1** in 0.1 M PBS buffer (1 mM, 0.15 μmol , 150.0 μL) was added the stock solution of semicarbazide (1 M in 0.1 M PBS buffer, pH = 7.38, 30 μmol , 30.0 μL) in the warm room

at 37 °C. The resulting reaction mixture was allowed to stir for 2 h, filtered and subjected for analytical HPLC. After another 22 h standing at room temperature, the sample subjected for LC-MS analysis.

Thrombin inhibition by in situ decaged TTI peptides

Tos-Gly-Pro-Arg-*p*-nitroanilide (Cayman Chemicals) was used as the chromogenic substrate for the inhibition of the amidolytic activity of human- α -thrombin (Invitrogen). The assays were performed in the Tris buffer (50mM Tris-HCl pH 8.0, 50 mM NaCl), which contains 100 mM substrate and various concentrations of the TTI peptides. After the addition of activation buffer (1 mg/mL BSA, 0.2 nM human- α -thrombin), The 96-well microtiter plates were incubated at 37 °C for 45 min and then monitored UV-Vis absorbance at 405 nm on a multi-mode microplate reader. (The data were corrected by subtracting the absorbance value of corresponding background). All measurements were repeated three times independently and then fitted with the Morrison equation (Williams and Morrison, 1979). Inhibition constants (K_i) and standard errors were calculated using this equation by Prism 6 (GraphPad Software).

Thrombin inhibition assay of **TTI04(fsY)** with photocaged **3-19** at specific concentration

To a prewarmed 0.1 M PBS solution of **TTI04(fsY)** (0.05 mg, 1.45×10^{-5} mmol, 50.0 μ L, pH 7.2, 37 °C), a solution of **3-19** (0.1 M, 40.0 equiv., 2.90×10^{-4} mmol, 5.8 μ L) and semicarbazide (1 M, 400 equiv., 2.90×10^{-3} mmol, 5.8 μ L) were added. The resulting mixture was vigorously stirred at 37 °C and 370 nm UV irradiation for 4 h to give **TTI04(sY)**. The peptide substrate **TTI04(fsY)**, **3-19** + UV, **TTI04(fsY)** + **3-19**, **TTI04(fsY)** + UV, **TTI04(fsY)** + **3-19** + UV reaction mixture were diluted to 3.7 nM (peptide concentration) for thrombin inhibition activity test.

Tandem MS analysis of sfGFP-151-fsY

100 µg sfGFP-151-fsY was added in 100.0 µL of 100% solution of trichloroacetic acid in PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, and 1.8 mM KH₂PO₄, pH = 7.50). Vortex the mixture vigorously and then freeze at -80 °C for 1 hour. Then, the thawed sample was centrifuged at 15000 rpm for 10 min. After the removal of supernatant, the pellet was resuspended in 500.0 µL of cool acetone. Afterwards the mixture was centrifuged at 15000 rpm for 10 min. The pellet was saved and dried. Then, the pellet was resuspended in 30.0 µL of 8 M urea in PBS buffer. After the addition of 70.0 µL of 100 mM ammonium bicarbonate and 1.5 µL of 1 M DTT, the sample was incubated at 65 °C for 15 min. Then, 2.5 µL of 500 mM iodoacetamide was added and incubated at room temperature for 30 min before the addition of 120.0 µL PBS buffer. 4.0 µL trypsin solution (0.5 µg/µL in trypsin buffer) and 2.5 µL of 100 mM CaCl₂ were added to digest the protein. After the overnight incubation at 37 °C, the digestion was quenched by 10.0 µL of formic acid. Then the sample was centrifuged at 15000 rpm for 15 min. The supernatant was collected and underwent mass spectra analysis.

LC-MS/MS analysis was performed on an Orbitrap Exploris 240 mass spectrometer with Xcalibur v4.4 (Thermo Scientific) coupled to a Dionex Ultimate 3000 RSLCnano system. The MS sample was desalted on a SepPak C18 cartridge (Waters) and dried on Speedvac. Desalted peptides were resuspended in Buffer A (100% H₂O, 0.1% formic acid) and 5.0 µL of the resulting solution were injected onto a 4 cm Acclaim PepMap 100 C18 column. Then peptides were eluted onto an Acclaim PepMap RSLC and separated with a 1-hour gradient from 5% to 25% of Buffer B (20% H₂O, 80 % CH₃CN, 0.1% formic acid) in Buffer A at a flow rate of 0.3 µL/min. The spray voltage was set to 2.1 kV. One full MS1 scan (120,000 resolution, 350-1800 m/z, RF lens 65%, AGC target 300%, automatic maximum injection time, profile mode) was obtained every 2 secs with dynamic exclusion

(repeat count 2, duration 10 s), isotopic exclusion (assigned), and apex detection (30% desired apex window) enabled. A variable number of MS2 scans (15,000 resolution, AGC 75%, maximum injection time 100 ms, centroid mode) were obtained between each MS1 scan based on the highest precursor masses, filtered for monoisotopic peak determination, theoretical precursor isotopic envelope fit, intensity (5E4), and charge state (2-6). MS2 analysis consisted of the isolation of precursor ions (isolation window 2 m/z) followed by higher-energy collision dissociation (HCD) (collision energy 30%). The MS data was analyzed by Proteome Discoverer V2.4 software package and searched using the SequestHT and Percolator Algorithms. Trypsin was specified as the protease with a maximum of 2 missed cleavages. Peptide precursor mass tolerance was set to 10 ppm with a fragment mass tolerance of 0.02 Da. Fluorosulfated tyrosine (+82.048), oxidation of methionine (+15.995) as well as acetylation (+42.011) and/or methionine-loss (+131.040) of the protein N-terminus were set as dynamic modifications. Cysteine alkylation (+57.021) was set as a static modification. The false discovery rate (FDR) for peptide identification was set to 1%. The mass-spectrometry data were collected as two technical replicates from one biological replicate.

High-resolution Orbitrap mass spectrometry analysis of sfGFP-151-fsY decaging

Decaging with free activator **3-7**: To a 500 μ L Eppendorf tube with fsGFP-151-fsY (299.7 μ L, 1 μ g/ μ L in 0.1 M PBS buffer) was added activator **3-7** (10 mM in 0.1 M PBS buffer, 33.0 μ L, 30 equiv.). The resulting reaction mixture was covered with aluminum foil and placed on 37 °C warm room shaker. 4 h later, the reaction was purified by the following sample preparation for Orbitrap MS analysis procedure.

Decaging with photocaged activator **3-19** and semicarbazide: To a 500 μ L Eppendorf tube with fsGFP-151-fsY (299.7 μ L, 1 μ g/mL in 0.1 M PBS buffer) was added activator **3-19** (10 mM in

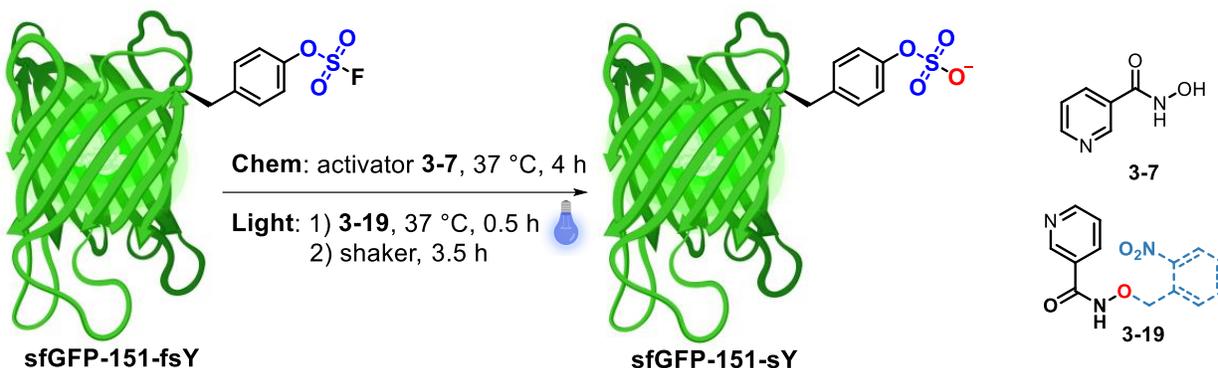
0.1 M PBS buffer, 33.0 μ L, 30 equiv.) and additive semicarbazide (1 M in 0.1 M PBS buffer, 3.3 μ L, 300 equiv.). The resulting reaction mixture was irradiated with 370 nm LED light (40 W, 100%) for 1 h. Subsequently, the whole setup was covered with aluminum foil and placed on 37 °C warm room shaker. After another 3 h, the reaction was purified by the following sample preparation for Orbitrap MS analysis procedure.

Sample preparation for Orbitrap MS analysis: The samples were cleaned up with PD-10 desalting columns packed with Sephadex G-25 resins (Cytiva Life Sciences) followed by applying Amicon Ultra 0.5 mL Centrifugal filter with 10K cut-off (EMD Millipore, Billerica-MA, USA) with water.

Orbitrap MS analysis: The samples were directly detected on an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Fisher Scientific, San Jose, California, USA) coupled with a TriVersa NanoMate (Advion, Ithaca, NY). Data acquisition was performed in positive ion mode. MS scans were acquired from 500-1500 m/z at a resolution of 240,000 @200 m/z with an AGC target of 100% and a maximum injection time of 100 ms. For each scan 10 μ scans were recorded.

Orbitrap MS data processing: MS data were deconvoluted on a Thermo Scientific BioPharma Finder software Version 5.0. Deconvolution Algorithm was set to Xtract (Isotopically Resolved). Output Mass Range was defined from 27,000 to 29,000. The monoisotopic masses of neutral mass (M) were set as Output Mass.

Anti-His western blot analysis of sfGFP-151-fsY decaging



| Expt. | I | II | III | IV | V |
|-----------|------------|---------|-------------|----------|-------------|
| Activator | 3-7 | - (PBS) | 3-19 | - (DMSO) | 3-19 |
| hv | - | - | + | + | - |

Reaction setup for Expt. I and II: To a 200 μL Eppendorf tube with sfGFP-151-fsY (1 $\mu\text{g}/\mu\text{L}$, 30.0 μL) was added activator **3-7** stock solution (10 mM in 0.1 M PBS buffer, 3.3 μL , 30 equiv.) or PBS buffer (0.1 M, 3.3 μL). The resulting reaction mixture was covered with aluminum foil and placed on 37 °C warm room shaker. 4 h later, the reaction was used directly in the western blot experiment.

Reaction setup for Expt. III and IV: To a 200 μL Eppendorf tube with sfGFP-151-fsY (1 $\mu\text{g}/\mu\text{L}$, 30.0 μL) was added activator **3-19** stock solution (10 mM in DMSO, 3.3 μL , 30 equiv.) or DMSO (3.3 μL). The resulting reaction mixture was irradiated with 370 nm LED light (40 W, 100%) for 0.5 h. Subsequently, the whole setup was covered with aluminum foil and placed on 37 °C warm room shaker. After another 3.5 h, the reaction was directly used in the western blot experiment.

Reaction setup for Expt. V: To a 200 μL Eppendorf tube with sfGFP-151-fsY (1 $\mu\text{g}/\mu\text{L}$, 30.0 μL) was added activator **3-19** stock solution (10 mM in DMSO, 3.3 μL , 30 equiv.). The whole setup was covered with aluminum foil and placed in the same photoreactor as Expt. III and IV. After the reaction mixture was irradiated with 370 nm LED light (40 W, 100%) for 0.5 h, the resulting reaction

mixture was kept in dark and placed on 37 °C warm room shaker. 3.5 h later, the reaction was used directly in the western blot experiment.

Protocol for Anti-His western blot:

Western blot was used to confirm the presence of a polyhistidine tag in the reporter protein before and after the releasing reaction to confirm the compatibility of the conditions with protein.⁸ Purified wild-type superfolder green fluorescent protein (sfGFP), sY (sulfotyrosine) or fsY (fluorosulfotyrosine)-incorporated mutant of sfGFP reporter proteins, and reaction mixture of fsY-incorporated sfGFP in different releasing conditions were resolved by LDS-PAGE using a freshly purchased Novex™ Tricine 16% gel (Thermo Fisher Scientific) in Tricine running buffer for 90 min at 120 V. The protein was transferred to a PVDF membrane (Life Technologies) using a Trans-Blot Turbo Transfer System 15 (BioRad) in freshly prepared Towbin transfer buffer (at 12 V for 30 min, twice). After complete transfer, membrane was blocked in 10 mL 5% milk in TBAT at 4 °C with constant agitation. Membrane was subsequently incubated in 1:3000 anti-HisTag mouse mAb (Invitrogen, MA1-21315, in 5% TBST) overnight. Next, the membrane was washed three times (10 min per wash) with TBST at room temperature (rt). Afterwards, the membrane was soaked in a 1:6000 dilution of chicken anti-mouse secondary antibody (Invitrogen, SA1-72021, in 5% milk TBST) mixture for 2 h at room temperature. The membrane was washed and activated using SuperSignal West Dura Kit (Thermo Fisher Scientific). The activated blot was imaged on the ChemiDoc MP imaging system (BioRad).

Note: The Towbin transfer buffer recipe is 14.4 g glycine and 3.0 g Tris base in 100 mL MeOH and 900 mL water. The recipe of TBST (Tris-buffered saline with 0.1% Tween® 20 detergent) is 8.0 g sodium chloride (NaCl), 0.2 g potassium chloride (KCl), and 2.3 g Tris base was added into 1000 mL water with 1 mL Tween® 20 detergent.

Decaging of fsY on the surface of live *S. aureus* cells

- a) Sortase A-mediated ligation of peptide 22 on live *S. aureus* cells.

From an overnight culture of *Staphylococcus aureus*, 500 μ L culture sample was transferred to a 17 \times 100 mm culture tube with 4.5 mL tryptic soy broth (TSB) medium added. After the cells grew to OD 1.0 at 37 $^{\circ}$ C, the cells were washed three times with Dulbecco's phosphate-buffered saline (DPBS, catalog number: 14040117 from gibco) and pelleted by centrifugation (5 min, 4000 rpm). The pellet was resuspended in DPBS buffer (4.95 mL), followed by the addition of peptide **3-20** stock solution (50 μ L, 0.2 mM in DMSO). The resulting mixture was covered with aluminum foil and shaking at 250 rpm for 6 h at room temperature. The above sample was washed by DPBS buffer (3 mL \times 6 times) until the supernatant was transparent and then resuspended in the same buffer (1 mL).

- b) Fluorescence microscopy and flow cytometry characterizations.

For fluorescence microscopy, 2.9 μ L of the above cell suspension was placed on a glass slide. The white field and fluorescent images were obtained on Zeiss microscope equipped with filter set 44 (excitation: BP 475/40, emission: BP 530/50). The images were captured using the 100x oil immersion with 1000 ms exposure time. All the images were further processed using Image J software with same parameters (Figure S18).

The fluorescence from the cells was quantified using flow cytometry analysis (Figure S19). Cells were diluted 10^6 cfu in DPBS and analyzed on Becton Dickinson Accuri C6 Plus (BD Biosciences). To enrich for single cells, a side scatter threshold trigger (SSC-H) was applied. To gate for single bacterial cells, we first selected events that appeared on the center of the FSC-A vs. SSC-A plot, then selected events along the diagonal of the FSC-H vs FSC-A plot. Events that appeared on the edges of the fluorescence histogram were excluded.

- c) LC-MS characterization of the decaging reaction on the live cell surface.

From an overnight culture of *S. aureus*, 50 μ L of culture sample was transferred to a 17 \times 100 mm culture tube with 4.5 mL TSB medium added (two identical batches are performed). After the cells grew to OD 0.1 in the same medium (5 mL) at 37 $^{\circ}$ C, the cells were washed three times with TSB medium and pelleted by centrifugation (5 min, 4000 rpm). The cells were then resuspended in TSB medium (4.95 mL), followed by the addition of 27-FAM peptide stock solution (50 μ L, 0.2 mM in DMSO). The resulting mixture was covered with aluminum foil and shaking for 6 h at 37 $^{\circ}$ C.

The above two batches were combined and washed by 0.1 M PBS buffer (3 mL \times 6 times) until the supernatant was transparent and then resuspended in 0.1 M PBS buffer (1.0 mL). The cells were subsequently used for decaging reaction in 0.1 M PBS buffer (55.6 μ L of 0.1 M reagent **7** stock solution was added to the mixture, 37 $^{\circ}$ C for 2 h on shaker with aluminum foil cover. A sample of 20 μ L of this cell suspension was plated on agar plate for cell viability test, with the heat-killed *S. aureus* (98 $^{\circ}$ C, 20 min) used as a control (Figure 55). The rest of cell suspension was washed by TEV cleavage buffer (50 mM Tris, 0.5 mM EDTA, 1 mM DTT, pH = 8.0) three times before as pelleted by centrifugation (5 min, 8000 rpm). The pellet was then suspended in 1.0 mL of TEV cleavage buffer. 50 μ L of TEV protease (7 μ M) was added to the rest of the cells, and incubate at 30 $^{\circ}$ C overnight or 3 h. The samples were flicked every 30 min to ensure good mixing. The cleaved peptide **24** was filtered using a 50K centrifugal filter. The sample was centrifuged at maximum speed (7830 rpm) for 15 minutes at room temperature and washed with 100 μ L water twice. The flow-through was lyophilized for two days and redissolved in 30 μ L of water. And 27 μ L of this solution was injected into the LC-MS. As a control, peptide **22** in 90 μ L TEV cleavage buffer was subjected to 10 μ L of TEV protease (7 μ M), and incubate at 30 $^{\circ}$ C overnight or 3 h. 10 μ L of the resulting peptide **23** was injected into the LC-MS. LC-MS condition for analyzing **23** and **24** in positive mode, solvent A contains acetonitrile/water/formic acid = 5:95:0.1, solvent B contains acetonitrile/water/formic acid = 95:5:0.1.

| Time (min) | A (%) | B (%) | Flow (mL./min) |
|------------|-------|-------|----------------|
| 0.00 | 95.0 | 5.0 | 0.200 |
| 5.00 | 95.0 | 5.0 | 0.200 |
| 17.00 | 5.0 | 95.0 | 0.200 |
| 22.00 | 5.0 | 95.0 | 0.200 |

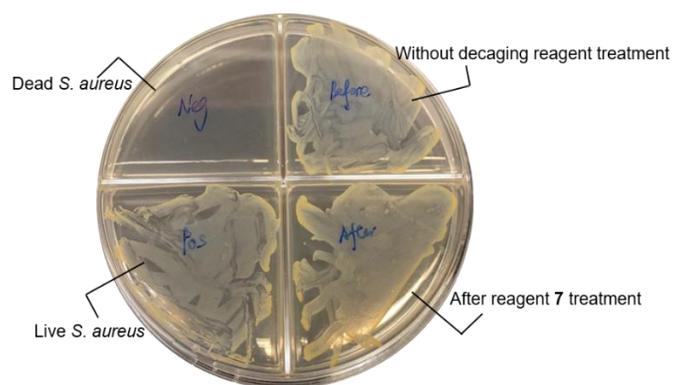
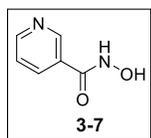
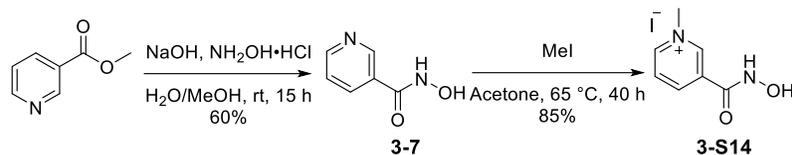
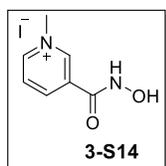


Figure 3-28. Plating assay to determine *E. coli* (BL21) viability under the decaging reaction

3.4.5 Small Molecule Synthesis

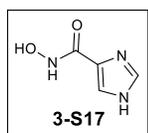
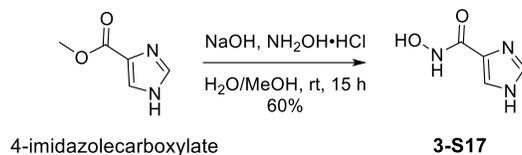


Compound **3-7** was obtained from the previously reported protocol with slight modification.⁵⁸ Hydroxylamine hydrochloride (NH₂OH·HCl, 1.01 g, 14.58 mmol) was added to the solution of sodium hydroxide (NaOH, 1.17 g, 29.17 mmol) in water (7.5 mL) at room temperature. 5 min later, above reaction mixture was added to methyl nicotinate (1.00 g, 7.29 mmol) in methanol (11.1 mL). 16 h later, TLC showed the complete consumption of the starting material and the reaction mixture was acidified by 5% HCl to be pH = 5~6. The solvent was removed under reduced pressure on rotavapor, diluted with methanol, filtered out the sodium chloride (NaCl) salt, and recrystallized from water to yield 604.47 mg solid of **3-7** in 60% yield. The obtained analysis results were consistent with literature report. ¹H NMR (500 MHz, DMSO-*d*₆) δ 11.38 (s, 1H), 9.20 (s, 1H), 8.90 (dd, *J* = 2.2, 0.9 Hz, 1H), 8.69 (dd, *J* = 4.8, 1.7 Hz, 1H), 8.09 (dt, *J* = 7.9, 2.0 Hz, 1H), 7.49 (ddd, *J* = 7.9, 4.9, 0.9 Hz, 1H). HRMS (DART⁺): calculated for [C₆H₇N₂O₂]⁺ (M+H)⁺ 139.0502, found 139.0510.

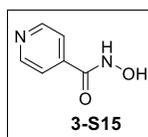
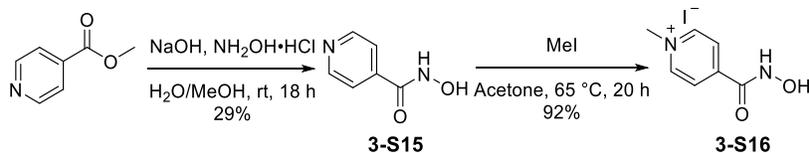


To a flame-dried Schlenk bottle with a stirring bar and activator **3-7** (163.4 mg, 1.18 mmol) were added iodomethane (8.21 g, 3.6 mL, 57.83 mmol) and acetone (4.0 mL) subsequently under the protection of nitrogen. The resulting reaction mixture was heated up to 60 °C in the sealed bottle. 40 h later, the reaction mixture was diluted with ethyl ether and significant amount of precipitation was observed. The reaction mixture was filtered and the solid was washed by cold ether, dried, and weighted to give 280.7 mg light yellow solid of **3-S14** in an 85% yield. ¹H NMR (500 MHz, D₂O) δ 9.25 (s, 1H), 9.01 (d, *J* = 6.1 Hz, 1H), 8.82 (d, *J* = 8.1 Hz, 1H), 8.22

(dd, $J = 8.2, 6.2$ Hz, 1H), 4.51 (s, 3H). ^{13}C NMR (126 MHz, D_2O) δ 161.54, 147.48, 144.66, 143.24, 132.06, 128.27, 48.74. HRMS (EST^+): calculated for $[\text{C}_7\text{H}_9\text{N}_2\text{O}_2]^+$ (M-I) $^+$ 153.0659, found 153.0616.

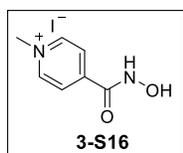


To a solution of sodium hydroxide (NaOH, 1.19 g, 29.74 mmol) in water (14.8 mL) was added hydroxylamine hydrochloride ($\text{NH}_2\text{OH}\cdot\text{HCl}$, 843.1 mg, 12.13 mmol) at room temperature. 5 min later, methyl 4-imidazolecarboxylate (1.00 g, 7.93 mmol) was added at the same temperature. The resulting reaction mixture was allowed to stir overnight. 15 h later, the reaction mixture was neutralized to $\text{pH} = 6\sim 7$ by 1 M HCl. White precipitation was generated from the clear solution. The reaction mixture was directly filtered, washed by cold water (Note: too much water washing could cause targeting material loss) and acetone once to dry. The solid was dried over air and weighted to afford 1.49 g white solid of **3-S17** in a 60% yield. ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 12.55 (broad, 1H), 10.83 (broad, 1H), 8.94 (broad, 1H), 7.73 (d, $J = 1.3$ Hz, 1H), 7.62 (s, 1H). ^{13}C NMR (126 MHz, DMSO) δ 160.60, 136.06, 134.58, 119.32. HRMS (DART^+): calculated for $[\text{C}_4\text{H}_6\text{N}_3\text{O}_2]^+$ ($\text{M}+\text{H}$) $^+$ 128.0455, found 128.0463.

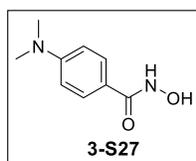
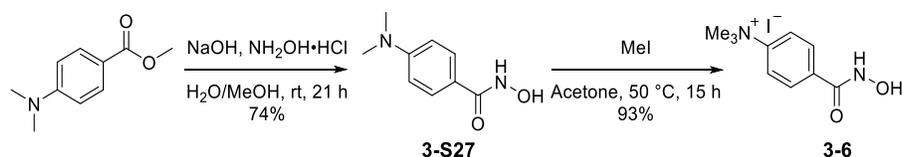


Compound **3-S15** was obtained from the previously reported protocol with slight modification.⁵⁸ Hydroxylamine hydrochloride ($\text{NH}_2\text{OH}\cdot\text{HCl}$, 1.01 g, 14.58 mmol) was added to the solution of sodium hydroxide (NaOH, 1.17 g, 29.17 mmol) in water (7.5 mL) at room temperature. 20 min later, above reaction mixture was added to methyl isonicotinate (1.00 g, 7.29

mmol) in methanol (11.1 mL). 18 h later, TLC showed the complete consumption of the starting material and the reaction mixture was acidified by 5% HCl to be pH = 5~6. The solvent was removed under reduced pressure on rotavapor, diluted with methanol, filtered out the sodium chloride (NaCl) salt, and recrystallized from water to yield 296.6 mg solid of **3-S15** in 29% yield. The obtained was consistent with literature report. ¹H NMR (500 MHz, DMSO-*d*₆) δ 11.51 (s, 1H), 9.29 (s, 0H), 8.73 – 8.67 (m, 1H), 7.70 – 7.60 (m, 1H). HRMS (DART⁺): calculated for [C₆H₇N₂O₂]⁺ (M+H)⁺ 139.0502, found 139.0509.

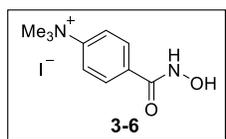


To a flame-dried Schlenk bottle with a stirring bar and activator **3-S15** (52.3 mg, 0.38 mmol) were added iodomethane (2.28 g, 1.0 mL, 16.06 mmol) and acetone (1.0 mL) subsequently under the protection of nitrogen. The resulting reaction mixture was heated up to 60 °C in the sealed bottle. 20 h later, the reaction mixture was diluted with ethyl ether and significant amount of precipitation was observed. The reaction mixture was filtered and the solid was washed by cold ether, dried, and weighted to give 98.0 mg light yellow solid of **3-S16** in an 92% yield. ¹H NMR (500 MHz, D₂O) δ 8.99 (d, *J* = 6.2 Hz, 2H), 8.32 (d, *J* = 6.1 Hz, 2H), 4.48 (s, 3H). ¹³C NMR (126 MHz, D₂O) δ 161.99, 146.48, 146.32, 125.83, 48.52. HRMS (DART⁺): calculated for [C₇H₉N₂O₂]⁺ (M-I)⁺ 153.0659, found 153.0652.

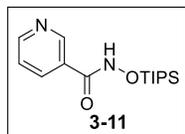


To a solution of hydroxylamine hydrochloride (NH₂OH·HCl, 1.55 g, 22.32 mmol) in water (4.7 mL) was added sodium hydroxide (NaOH, 1.79 g, 44.64 mmol) at room temperature. There's a significant amount of heat was released. After the reaction mixture cooled back to room temperature, a solution of methyl 4-(dimethylamino)benzoate

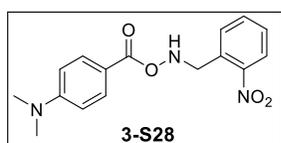
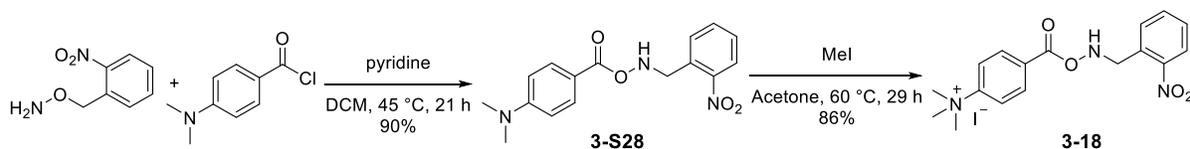
(2.00 g, 11.16 mmol) in methanol (MeOH, 7.5 mL) was added at the same temperature. The resulting suspension was allowed to stir overnight. 21 h later, the reaction mixture was neutralized to pH = 7 by 10 M HCl. The whole reaction became a whole gel-like mixture. Ethyl acetate (EA) was used for dilution and washed by water. The aqueous layer was extracted with EA for another two times. The organic layer was combined, dried over anhydrous sodium sulfate (Na₂SO₄), and concentrated under reduced pressure. The residue was recrystallized from ethanol (20 mL) and the solid was washed by cold ether. The filtrate was concentrated and recrystallized again with ethanol (5 mL) and washed by cold ether. The combined solid was dried over air and weighted to afford 1.49 g brown solid as target compound **3-S27** in a 74% yield. ¹H NMR (600 MHz, CD₃OD) δ 7.68 – 7.53 (m, 2H), 6.76 – 6.64 (m, 2H), 2.99 (s, 6H). ¹³C NMR (151 MHz, CD₃OD) δ 169.03, 154.31, 129.44, 119.66, 112.20, 40.19. HRMS (DART⁺): calculated for [C₉H₁₃N₂O₂]⁺ (M+H)⁺ 181.0972, found 181.0966.



To a flame-dried Schlenk bottle with a stirring bar and **3-S31** (192.4 mg, 1.07 mmol) were added iodomethane (6.43 g, 2.8 mL, 45.29 mmol) and acetone (4.0 mL) subsequently under the protection of nitrogen. The resulting reaction mixture was heated up to 60 °C in the sealed bottle. 15 h later, the reaction mixture was diluted with ethyl ether and significant amount of precipitation was observed. The reaction mixture was filtered and the solid was washed by cold ether, dried, and weighted to give 320.5 mg white solid of **3-6** in an 93% yield. ¹H NMR (600 MHz, D₂O) δ 8.02 – 7.98 (m, 2H), 7.95 (m, 2H), 3.71 (s, 9H). ¹³C NMR (151 MHz, d₂O) δ 166.33, 148.91, 133.21, 129.18, 120.55, 56.95. HRMS (DART⁺): calculated for [C₁₀H₁₅N₂O₂]⁺ (M-I)⁺ 195.1128, found 195.1124.

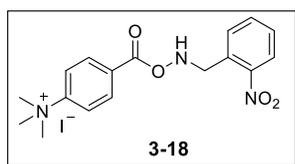


To a solution of activator **3-7** (100.0 mg, 0.73 mmol) in DMF (6.9 mL) were added triisopropylsilyl chloride (TIPSCl, 279.2 mg, 1.45 mmol) and imidazole (98.6 mg, 1.45 mmol) subsequently at room temperature (rt). 48 h later, the reaction was diluted with EA, washed by water, brine, dried over anhydrous Na_2SO_4 , and concentrated under reduced pressure. The residue was loaded on the Biotage for purification, hexane/EA was used as the mobile phase to afford 97.7 mg product **3-11** in a 46% yield. ^1H NMR (500 MHz, CD_3OD) δ 8.87 (s, 1H), 8.67 (s, 1H), 8.14~8.17 (m, 1H), 7.53 (dd, $J = 8.0, 5.0$ Hz, 1H), 1.27~1.35 (m, 3H), 1.17 (d, $J = 7.7$ Hz, 18H). ^{13}C NMR (126 MHz, CD_3OD) δ 167.02, 152.80, 148.74, 136.91, 130.27, 125.20, 18.27, 13.18. HRMS (DART $^+$): calculated for $[\text{C}_{15}\text{H}_{27}\text{N}_2\text{O}_2\text{Si}]^+$ ($\text{M}+\text{H}$) $^+$ 295.1836, found 295.1837.



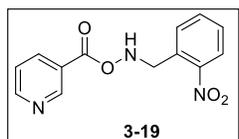
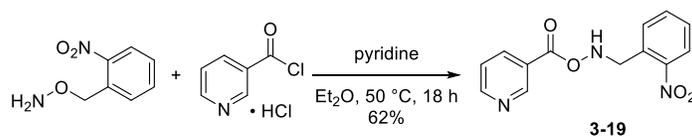
In a flame-dried 25 mL Schlenck bottle with a stirring bar, 4-(dimethylamino)benzoyl chloride (530.8 mg, 2.89 mmol) was dissolved in dichloromethane (DCM, 5.0 mL) which was followed by the addition of dry pyridine (685.9 mg, 701.3 μL , 8.67 mmol) under the protection of argon at room temperature. The reaction mixture turned into an orange-red clear solution quickly. Subsequently, *O*-(2-nitrobenzyl)hydroxylamine was added to the reaction mixture dropwise at the same temperature. The resulting reaction mixture was allowed to be heated up to 45 $^\circ\text{C}$ in the sealed bottle (refluxing occurred). 21 h later, the reaction mixture was diluted with ethyl acetate (EA), washed by water, brine, dried over anhydrous Na_2SO_4 , and concentrated under reduced pressure on rotavapor. The residue

was loaded on the Biotage for purification with hexane/EA as eluent. 42% EA gave 547.6 mg product **3-S28** in a 90% yield. $^1\text{H NMR}$ (600 MHz, CD_3OD) δ 8.08 (dd, $J = 8.2, 1.3$ Hz, 1H), 7.89 (dd, $J = 7.7, 1.4$ Hz, 1H), 7.74 (td, $J = 7.6, 1.3$ Hz, 1H), 7.62 – 7.55 (m, 3H), 6.75 – 6.62 (m, 2H), 5.33 (s, 2H), 3.01 (s, 6H). $^{13}\text{C NMR}$ (151 MHz, CD_3OD) δ 169.25, 154.64, 149.62, 134.62, 133.09, 131.61, 130.23, 129.78, 125.70, 118.92, 112.13, 75.17, 40.16. HRMS (DART $^+$): calculated for $[\text{C}_{16}\text{H}_{18}\text{N}_3\text{O}_4]^+$ (M+H) $^+$ 316.1292, found 316.1292.



To a flame-dried Schlenck bottle with a stirring bar and above compound (219.4 mg, 0.70 mmol) were added iodomethane (4.20 g, 1.84 mL, 29.57 mmol) and acetone (1.6 mL) subsequently under the protection of nitrogen.

The resulting reaction mixture was heated up to 60 °C in the sealed bottle. 29 h later, the reaction mixture was diluted with ethyl ether and significant amount of precipitation was observed. The reaction mixture was filtered and the solid was washed by cold ether, dried, and weighted to give 273.0 mg pale yellow solid of **3-18** in an 86% yield. $^1\text{H NMR}$ (600 MHz, $\text{DMSO}-d_6$) δ 12.08 (s, 1H), 8.11 – 8.04 (m, 3H), 7.91 (d, $J = 8.5$ Hz, 2H), 7.84 (m, 1H), 7.80 (m, 1H), 7.64 (m, 1H), 5.31 (s, 2H), 3.62 (s, 9H). $^{13}\text{C NMR}$ (151 MHz, $\text{DMSO}-d_6$) δ 162.94, 149.41, 148.06, 133.88, 133.37, 131.16, 130.80, 129.67, 128.86, 124.72, 121.00, 73.38, 56.48. HRMS (DART $^+$): calculated for $[\text{C}_{17}\text{H}_{20}\text{N}_3\text{O}_4]^+$ (M-I) $^+$ 330.1448, found 330.1459.



In a flame-dried Schlenck bottle with a stirring bar, *O*-(2-nitrobenzyl)hydroxylamine (245.6 mg, 1.46 mmol) was dissolved in ethyl ether (Et_2O , 6.0 mL) and followed by the addition of dry pyridine (347.6 mg, 4.74

mmol, 383.0 μ L) at room temperature under the protection of nitrogen. Subsequently, nicotinoyl chloride hydrochloride (260.0 mg, 1.46 mmol) was added at the same conditions. The resulting reaction mixture was sealed and heated up to 50 $^{\circ}$ C. 18 h later, the reaction mixture was diluted with dichloromethane (DCM), washed by water and brine, dried over anhydrous Na_2SO_4 , concentrated under reduced pressure, loaded on Biotage for purification using hexane/ethyl acetate as the mobile phase. 80% EA provided the 245.6 mg title compound **3-19** as white solid in a 62% yield. ^1H NMR (600 MHz, CD_3OD) δ 8.85 (d, $J = 2.1$ Hz, 1H), 8.68 (dd, $J = 4.9, 1.7$ Hz, 1H), 8.14 (dt, $J = 8.0, 1.9$ Hz, 1H), 8.08 (dd, $J = 8.3, 1.3$ Hz, 1H), 7.86 (d, $J = 7.7$ Hz, 1H), 7.74 (td, $J = 7.6, 1.3$ Hz, 1H), 7.59 (td, $J = 7.8, 1.4$ Hz, 1H), 7.55 – 7.49 (m, 1H), 5.40 (s, 2H). ^{13}C NMR (151 MHz, CD_3OD) δ 153.06, 149.80, 148.84, 136.94, 134.65, 132.64, 131.83, 130.50, 129.66, 125.77, 125.22, 75.16. HRMS (DART $^+$): calculated for $[\text{C}_{13}\text{H}_{12}\text{N}_3\text{O}_4]^+$ (M+H) $^+$ 274.0822, found 274.0831.

3.4.6 Peptide Synthesis and Decaging

DADEfsYL-NH₂ (3-1)

The synthesis was performed manually following general procedure A with 408.0 mg resin loading (0.2 mmol). Purification of the titled peptide through the Semi-Prep RP-HPLC with eluent A (95% H_2O , 5% CH_3CN , and 0.1% TFA) and B (5% H_2O , 95% CH_3CN , and 0.1% TFA). The running method was set as 0-5 min, 15% B, linear gradient from 15%~40% B over 15 min. The desired peptide was observed at $t_{\text{R}} = 15.00$ min under 210 nm and obtained in a 32% yield after twice purification and lyophilization. Analytical RP-HPLC of this material dissolved in 0.1 M PBS buffer (CH_3CN : 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~33% CH_3CN from 0-20 min, 20-21 min 33%~95% CH_3CN , and 95% CH_3CN washing for another 2 min) showed a single peak at t_{R}

= 18.37 min under 210 nm wavelength (Figure 3-29). HRMS (ESI⁺): calculated for [C₃₁H₄₅N₇O₁₅SF]⁺ (M+H)⁺ 806.2673, found 806.2140 (Figure 3-30).

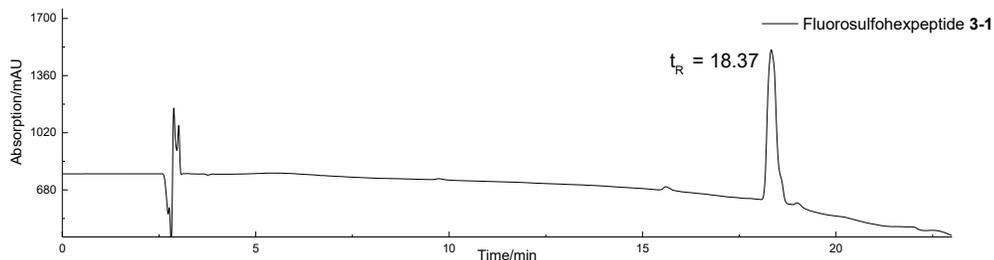


Figure 3-29: Analytical RP-HPLC (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~33% CH₃CN from 0-20 min, 20-21 min 33%~95% CH₃CN, and 95% CH₃CN washing for another 2 min) chromatogram of purified peptide **DADEfsYL-NH₂ (3-1)**

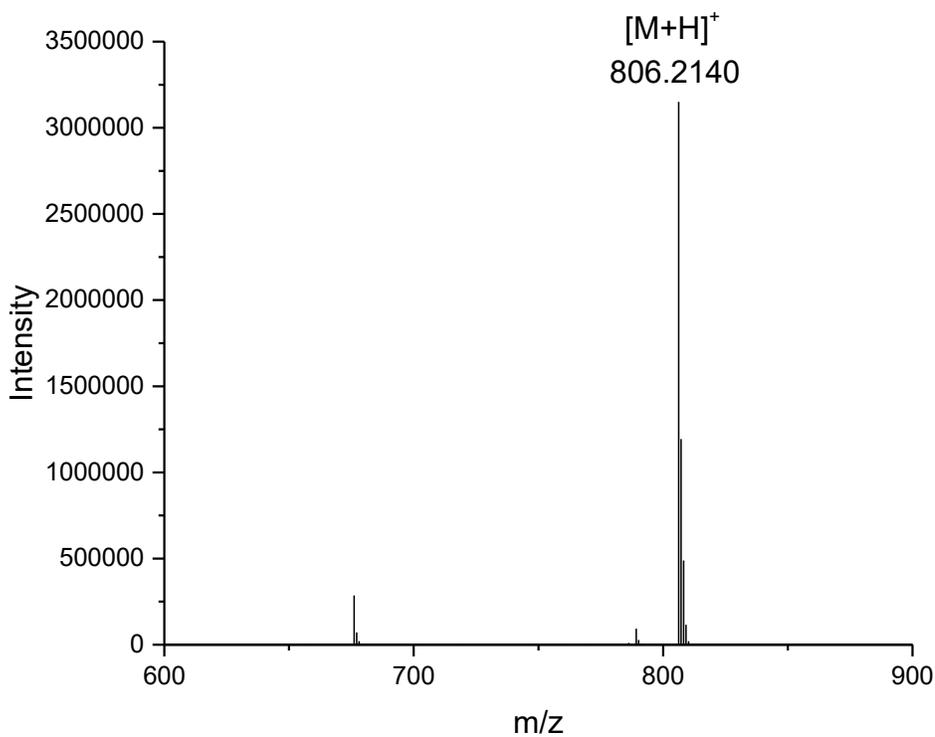


Figure 3-30. ESI-TOF mass spectrum for peptide **DADEfsYL-NH₂ (3-1)**

DADEsYL-NH₂ (3-2)

Decaging of **1** was performed following general procedure C with 5.3 mg substrate dissolved in 0.1 M PBS buffer with 20 equivalents of 3-PHA (0.1 M stock solution in 0.1 M PBS buffer) as activator. 2 h later, the reaction was directly filtered and loaded on the Semi-Prep RP-HPLC for purification. CH₃CN:20 mM ammonium acetate buffer mobile phase was chosen as the mobile phase, and the method was set as 0-5 min, 5% CH₃CN; linear gradient 5-13min, 5%~19.4% CH₃CN. The desired peptide was afforded with a $t_R = 9.50$ min under 210 nm. The combined fractions were air concentrated and the residue was lyophilized to provide 6.4 mg white solid with a 93% yield. Analytical RP-HPLC of this material dissolved in 0.1 M PBS buffer (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~35% CH₃CN from 0-20 min, 20-21 min 35%~95% CH₃CN, and 95% CH₃CN washing for another 2 min) showed a single peak at $t_R = 9.65$ min under 210 nm wavelength (Figure 3-31). HRMS (ESI⁺): calculated for [C₃₁H₄₅N₇O₁₆NaS]⁺ (M+Na)⁺ 826.2536, found 826.1847 (Figure 3-32).

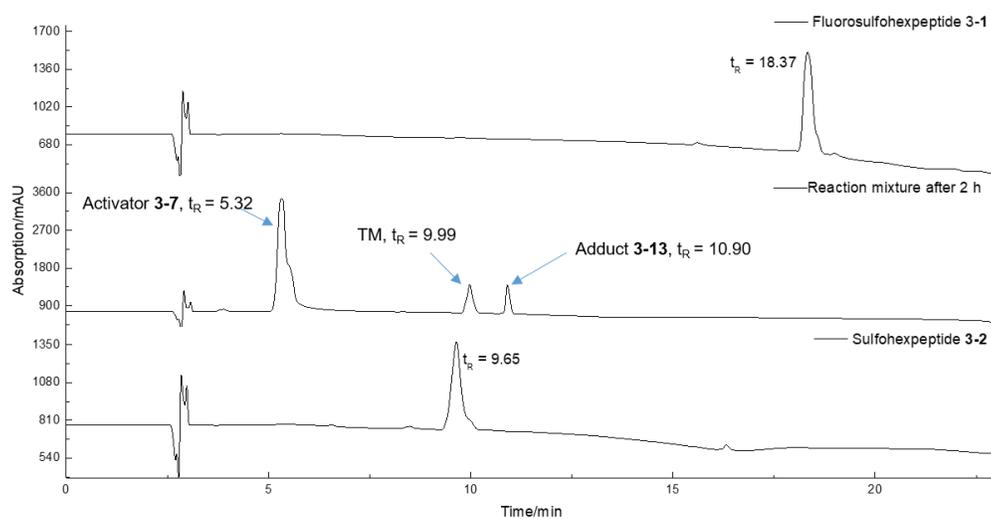


Figure 3-31. Analytical RP-HPLC (CH_3CN : 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~33% CH_3CN from 0-20 min, 20-21 min 33%~95% CH_3CN , and 95% CH_3CN washing for another 2 min) chromatograms of the decaging reaction mixture (middle) and purified peptide **DADEsYL-NH₂** (**3-2**, bottom)

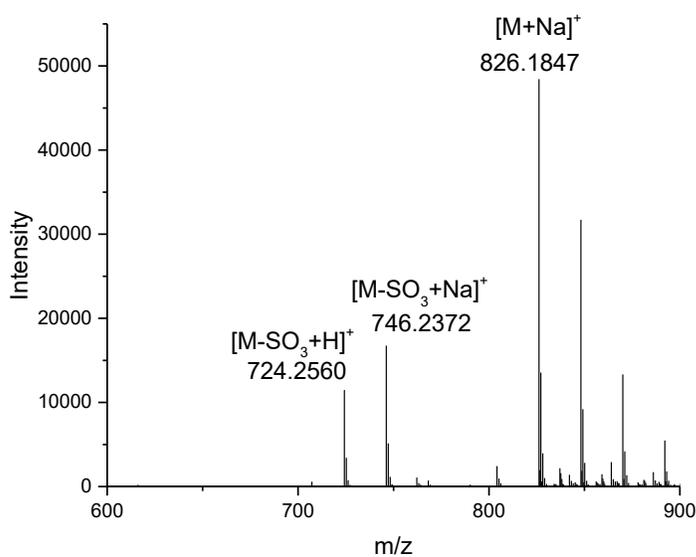


Figure 3-32. ESI-TOF mass spectrum for peptide **DADEsYL-NH₂** (**3-2**)

fsYEfsYLDfsYDF-NH₂ (3-14)

The synthesis of the title octapeptide was performed manually following general procedure A with 102.0 mg resin loading (0.2 mmol). Purification of the titled peptide through the Semi-Prep RP-HPLC with eluent A (95% H₂O, 5% CH₃CN, and 0.1% TFA) and B (5% H₂O, 95% CH₃CN, and 0.1% TFA). The running method was set as linear gradient from 25% to 55% within 30 min. The desired peptide was observed at $t_R = 21.5$ min under 210 nm and 10.2 mg cotton-like white solid was obtained in a 15% yield after twice purification and lyophilization. Analytical RP-HPLC of this material dissolved in DMSO (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~60% CH₃CN over 40 min) showed a single peak at $t_R = 23.35$ min under 210 nm wavelength (Figure 3-33). HRMS (ESI⁺): calculated for [C₅₅H₆₅F₃N₉O₂₃S₃]⁺ (M+2H)⁺ 687.1607, found 687.1832 (Figure 3-34).

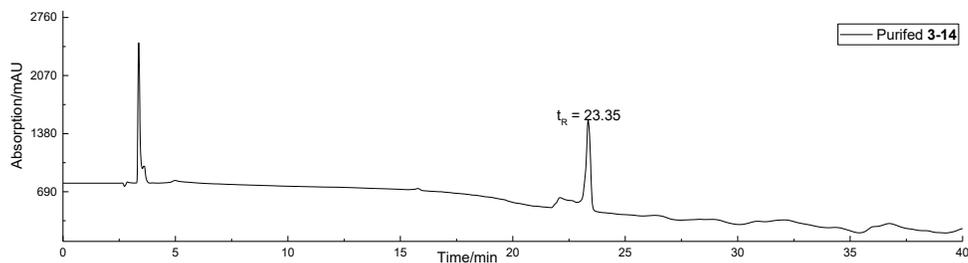


Figure 3-33. Analytical RP-HPLC (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~60% CH₃CN over 40 min) chromatogram of purified peptide **fsYEfsYLDfsYDF-NH₂ (3-14)**

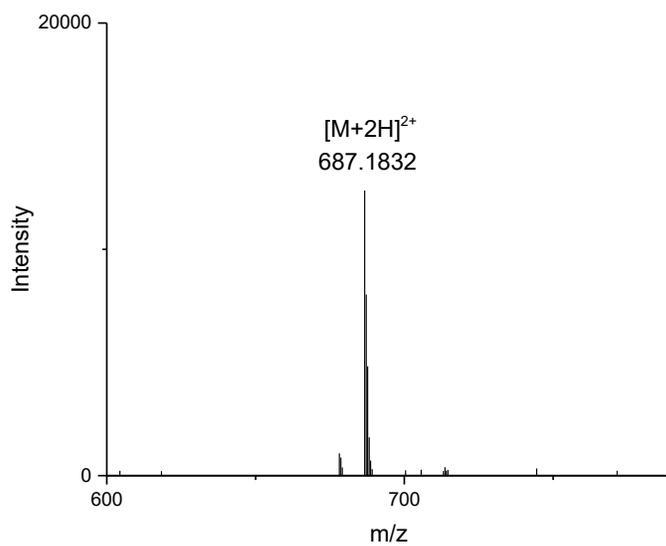


Figure 3-34. ESI-TOF mass spectrum for peptide **fsYEfsYLDfsYDF-NH₂ (3-14)**

sYEsYLDsYDF-NH₂ (3-15)

Decaging of **3-14** was performed following general procedure C with 0.6 mg substrate (suspension in a mixture of 0.1 M PBS buffer/CH₃CN = 7:3) used with 60 equivalents of 3-PHA (0.1 M stock solution in 0.1 M PBS buffer) as activator. 3 h later, RP-HPLC showed the reaction reached completion. The reaction mixture was first passed through a PD-10 column (cytiva, Sephadex™ G-25 M) using pure water as eluent to get rid of the excessive amount of activator and salts. The fraction containing targeting sulfate molecule was verified by LC-MS and combined for subsequent Semi-Prep RP-HPLC purification. CH₃CN: 20 mM ammonium acetate buffer was employed as mobile phases. The running method was set as 0-5 min, 5% CH₃CN, linear gradient from 5%~25% CH₃CN

over 15 min. The desired peptide was observed at $t_R = 12.50$ min under 210 nm and 0.4 mg cotton-like white solid was obtained in a 15% yield after purification and lyophilization. Analytical RP-HPLC of this material dissolved in water (CH_3CN : 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~60% CH_3CN over 40 min) showed a single peak at $t_R = 15.71$ min under 210 nm wavelength (Figure 3-35). HRMS (ESI): calculated for $[\text{C}_{55}\text{H}_{65}\text{N}_9\text{O}_{26}\text{S}_3]^{2-} (\text{M}-2\text{H})^{2-}$ 681.66074, found 681.7419 (Figure 3-36).

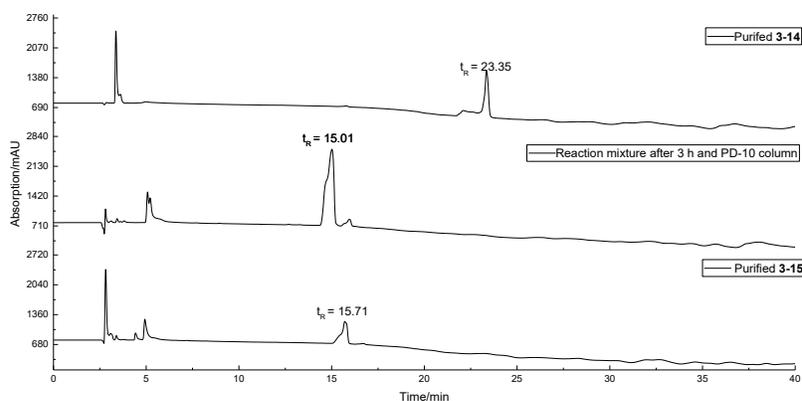


Figure 3-35: Analytical RP-HPLC (CH_3CN : 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~60% CH_3CN over 40 min) chromatogram of reaction mixture after 3 h and passed through a PD-10 column (middle) purified peptide **sYEsYLDsYDF-NH₂** (**3-15**, bottom)

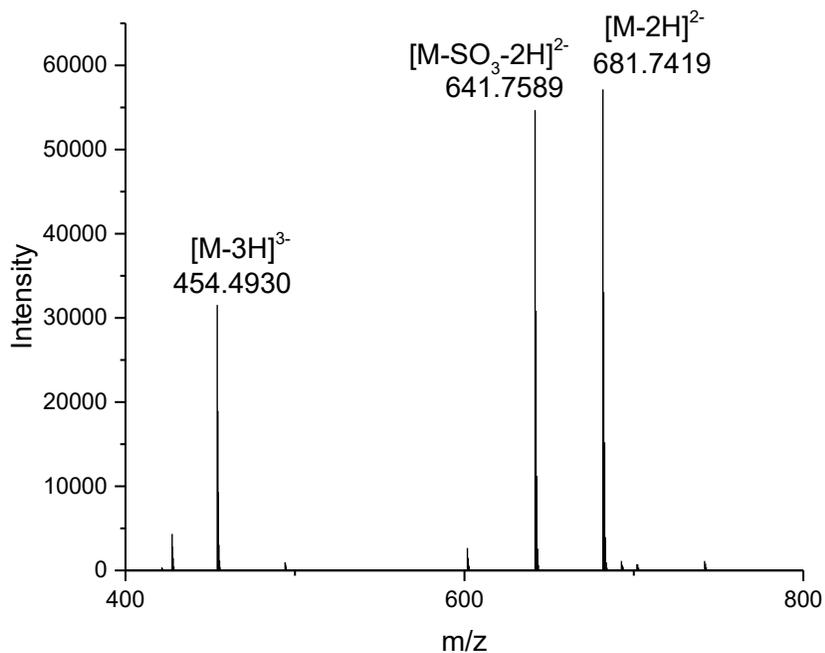


Figure 3-36. ESI-TOF mass spectrum for peptide **sYEsYLDsYDF-NH₂ (3-15)**

Biotin-TTPDfsYGHfsYDDKDTLNLNTPVDK-NH₂ (3-16)

The synthesis of the title 22mer peptide was separated in two steps. The first 14 amino acids (counted from the *C*-terminal) was constructed automatically following general procedure B with 200.0 mg Rink amide resin was loaded After the last coupling step and *N*-terminal Fmoc removal, the resin was washed with DMF, dichloromethane, ethyl ether, hexane subsequently (for each solvent 3 x 3 min), and dried to afford 509.6 mg total resin. A small portion of the resin was cleaved with standard cleavage cocktail, and verified by LC-MS, providing the expected 14mer signal: HRMS (ESI⁺): calculated for [C₆₆H₁₁₂N₁₈O₂₇]²⁺ (M+2H)²⁺ 794.3976, found 794.3710 (Figure 3-37). 112.6 mg (0.022 mmol, calculated based on the previous total resin obtained) resin was used for the subsequent amino acids incorporation following general procedure A. Purification of the titled peptide through the Semi-

Prep RP-HPLC with eluent A (95% H₂O, 5% CH₃CN, and 0.1% TFA) and B (5% H₂O, 95% CH₃CN, and 0.1% TFA). The running method was set as: 0-5 min, 15% B; linear gradient 5-35 min, 15%~35% B. The product peak emerged at t_R = 31.0 min under 210 nm wavelength and appropriate fractions verified by LC-MS were combined. After concentration with air-blowing, the residue was subjected for lyophilization to provide 42.7 mg of the title compound in 60% yield. Analytical RP-HPLC of this material dissolved in water (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~60% CH₃CN over 40 min) showed a single peak at t_R = 18.80 min under 230 nm wavelength (Figure 3-38). HRMS (ESI⁺): calculated for [C₁₁₉H₁₇₉F₂N₃₀O₄₇S₃]³⁺ (M+3H)³⁺ 971.7229, found 971.6719 (Figure 3-39).

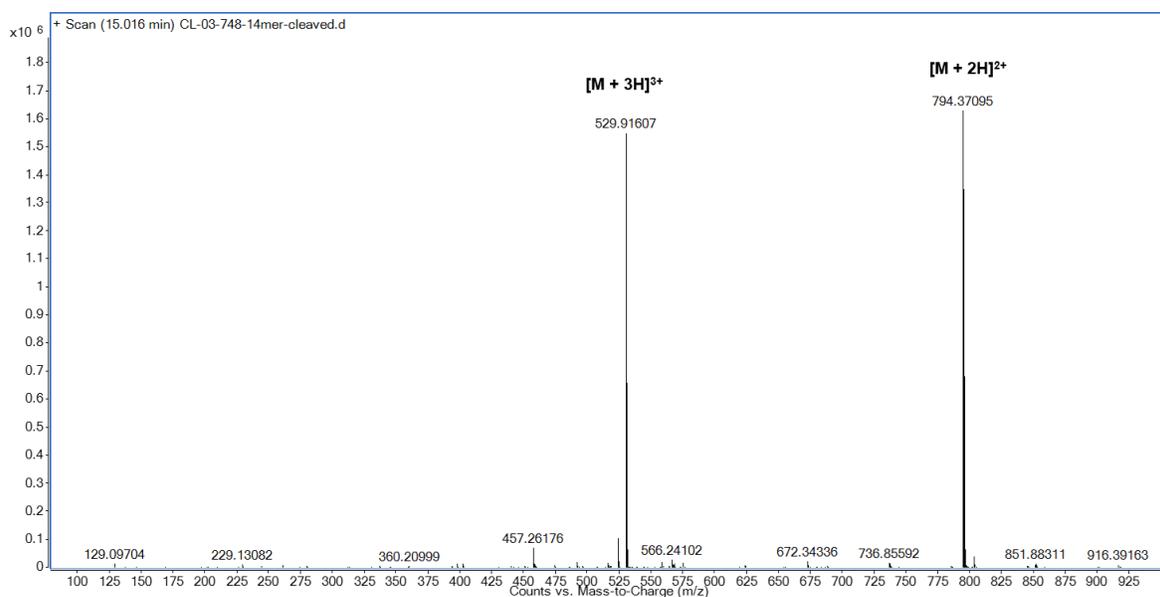


Figure 3-37. ESI-TOF mass spectrum for peptide **DDKDTLDLNTTPVDK-NH₂**

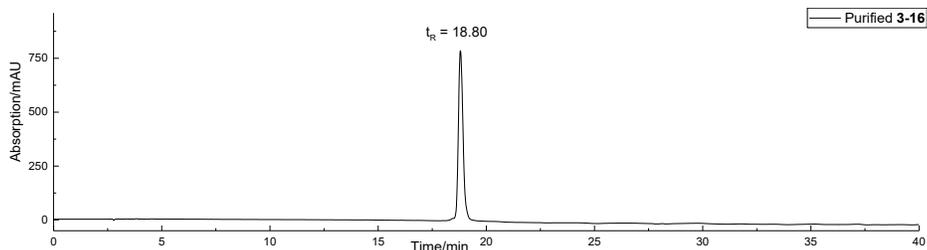


Figure 3-38. Analytical RP-HPLC chromatogram (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~60% CH₃CN over 40 min) of purified peptide **Biotin-TTPDfsYGHfsYDDKDTLNLNTPVVK-NH₂ (3-16)**

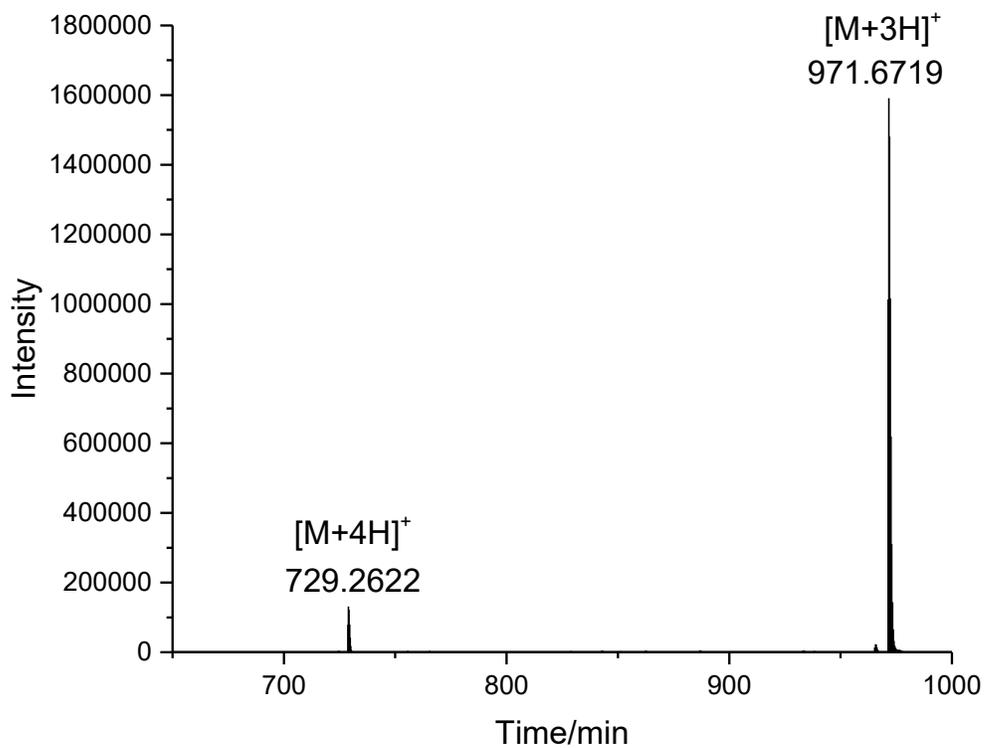


Figure 3-39. ESI-TOF mass spectrum for peptide **Biotin-TTPDfsYGHfsYDDKDTLNLNTPVVK-NH₂ (3-16)**

Biotin-TTPDsYGHsYDDKDTLNLNTPVVK-NH₂ (3-17)

Decaging of **16** was performed following general procedure C with 8.8 mg substrate (dissolved in 0.1 M PBS buffer) and 40 equivalents of 3-PHA (0.1 M stock solution in 0.1 M PBS buffer, pH = 7.35) as activator. 2.5 h later, RP-HPLC showed the reaction reached completion. The reaction mixture was directly filtered and subjected for Semi-Prep RP-HPLC purification. CH₃CN: 20 mM ammonium acetate buffer was employed as mobile phases. The running method was set as 0-5 min, 5% CH₃CN; linear gradient from 5%~25% CH₃CN over 25 min. The desired peptide was observed at t_R = 12.96 min under 230 nm wavelength. Appropriate fractions were verified by LC-MS, combined, dried over air-blowing, lyophilized to give 6.1 mg white solid in a 74% yield. Analytical RP-HPLC of this material dissolved in water (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~60% CH₃CN over 40 min) showed a single peak at t_R = 15.1 min under 230 nm wavelength (Figure 3-40). HRMS (ESI): calculated for [C₁₁₉H₁₈₁N₃₀O₄₉S₃]³⁺ (M+3H)³⁺ 970.3924, found 970.3001 (Figure 3-41).

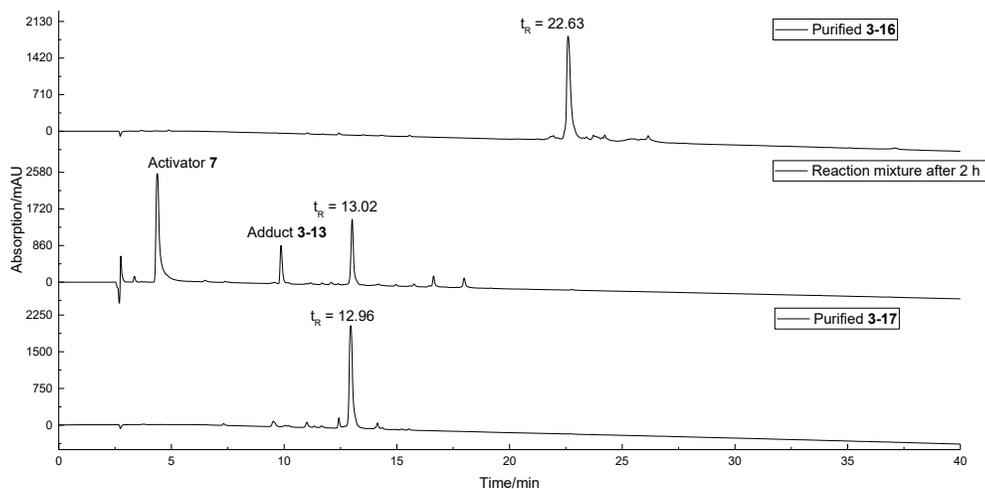


Figure 3-40. Analytical RP-HPLC chromatogram of (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~60% CH₃CN over 40 min) purified reaction mixture (middle) peptide **Biotin-TTPDsYGHsYDDKDTLNLNTPVDK-NH₂** (3-17, bottom)

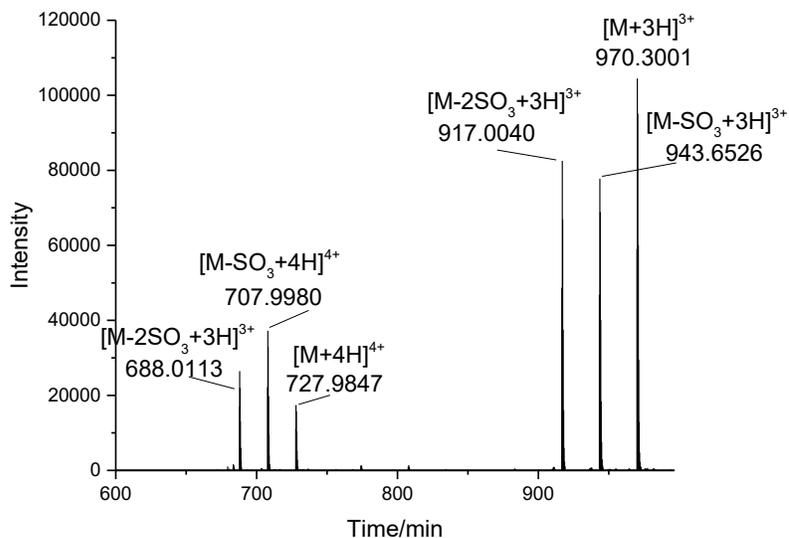
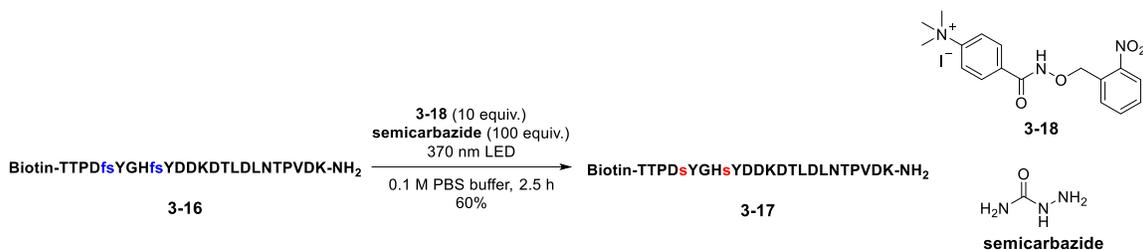


Figure 3-41. ESI-TOF mass spectrum for peptide **Biotin-TTPDsYGHsYDDKDTLNLNTPVDK-NH₂** (3-17)

Light-mediated fluorosulfate decaging of 3-16



A 1.5 mL transparent glass vial was equipped with a stirring bar and 1 mM stock solution of substrate in 0.1 M PBS buffer (3.2 mg, 1 mM in 0.1 M PBS buffer, 983.1 μ L). **3-18** (0.1 M in DMSO, 98.3 μ L) and semicarbazide stock solution (1 M in 0.1 M PBS buffer, 98.3 μ L) were added subsequently at room temperature. Then the whole reaction mixture was placed in the photoreactor to warm up. After two minutes, light irradiation was kept on for another 2.5 hours. The reaction mixture was monitored by analytical RP-HPLC until completion. The reaction mixture was passed through a PD-10 column using 0.1 M PBS buffer as eluent. Appropriate fractions were verified by LC-MS and combined to be dried over air blowing. Lots of salts was observed as residue. Minimum amount of water was used for re-dissolving the mixture, filtered, tested analytical RP-HPLC (Figure 3-42a), and loaded on Semi-Prep RP-HPLC purification. CH₃CN: 20 mM ammonium acetate buffer was employed as mobile phases. The running method was set as 0-5 min, 5% CH₃CN; linear gradient from 5%~20% CH₃CN over 29 min. The desired peptide was observed at about $t_R = 16.00$ min under 210 nm wavelength. Appropriate fractions were verified by LC-MS, combined, dried over air-blowing, lyophilized to give 1.8 mg white solid in a 60% yield. Analytical RP-HPLC of this material dissolved in water (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~60% CH₃CN over 40 min) showed a single peak at $t_R = 14.98$ min under 230 nm wavelength (Figure 3-42b, bottom) which was consistent with previously synthesized pure sulfated 22mer (Figure 3-42b, top). HRMS (ESI): calculated for [C₁₁₉H₁₈₁N₃₀O₄₉S₃]³⁺ (M+3H)³⁺ 970.3924, found 970.4073 (Figure 3-43).

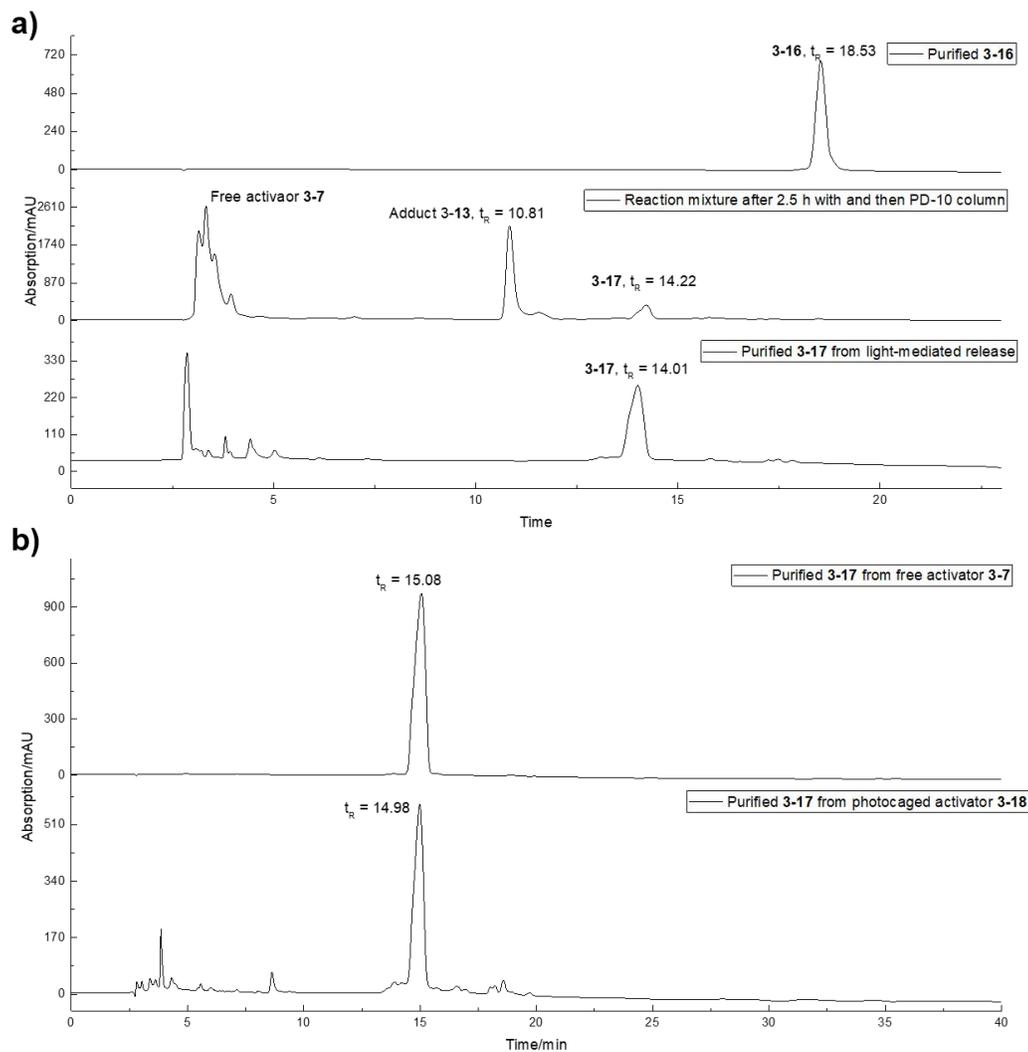


Figure 3-42. Analytical RP-HPLC chromatogram of **a)** photolytic reaction mixture (middle, CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~33% CH₃CN from 0-20 min, 20-21 min 33%~95% CH₃CN, and 95% CH₃CN washing for another 2 min, monitored under 230 nm wavelength) and **b)** comparison of purified **3-17** from free (top) or photocaged (bottom) activator (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~60% CH₃CN over 40 min, monitored under 230 nm wavelength)

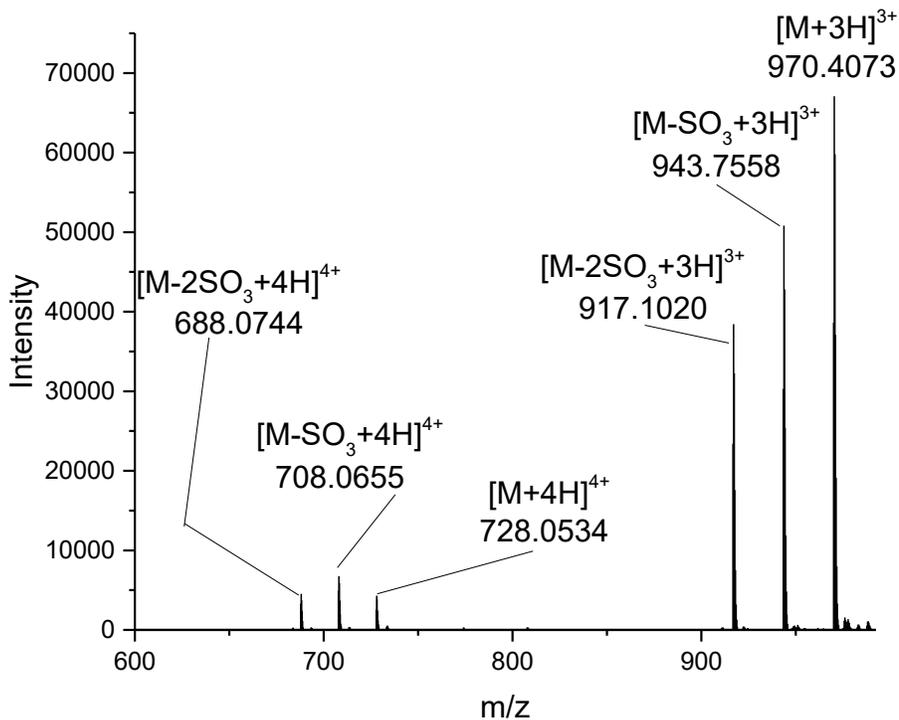


Figure 3-43. ESI-TOF mass spectrum for peptide **Biotin-TTPDsYGHsYDDKDTLDLNTVPDK-NH₂ (3-17)** from light-mediated decaging reaction

GEPGAPIDYDEYGDSSEEVGGTPLHEIPGIRL-OH (TTI01)

The peptide **TTI01** were synthesized according to the Fmoc-based SPPS outlined in the general procedure B using the Fmoc-L-OH loaded Wang resin (0.025 mmol). After cleavage, the crude product was purified by HPLC and lyophilized to afford a white powder (45.0 mg, 0.0134 mmol, 54 % isolated yield). Purification of the **TTI01** through the semi-prep RP-HPLC with eluent A (95% H₂O, 5% CH₃CN, and 0.1% TFA) and B (5% H₂O, 95% CH₃CN, and 0.1% TFA). The running method was set as linear gradient from 15% to 35% within 30 min. The desired peptide was observed at $t_R = 19.6$ min. Analytical HPLC of this material dissolved in water showed a single peak at $t_R = 9.3$ min, monitored by the UV absorbance at 210 nm. Method for analytical HPLC was set as: A (95% H₂O, 5% CH₃CN, and

0.1% formic acid): B (5% H₂O, 95% CH₃CN, and 0.1% formic acid) as mobile phase, linear gradient from 5%~95% B from 0-12 min, 12-16 min with 95% B, and Agilent InfinityLab Poroshell 120 column (2.7 μm, 2.1 x 50 mm) was used (Figure 45). ESI-MS calculated for **TTI01** (C₁₄₇H₂₂₁N₃₇O₅₄): [M + 3H]³⁺ m/z = 1123.8562; found: 1124.4307 (Figure 46).

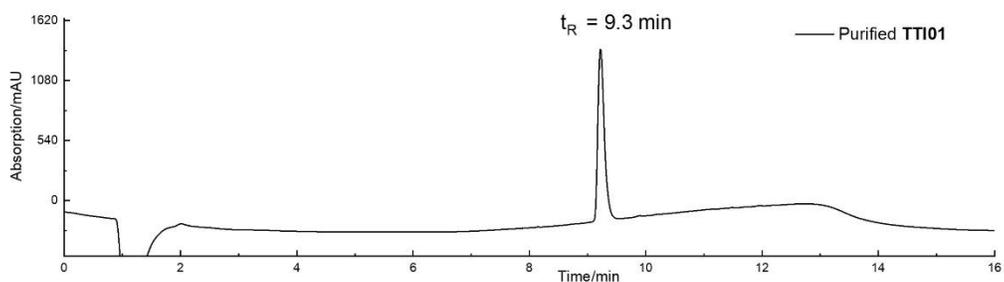


Figure 3-45. Analytical RP-HPLC chromatogram of (A (95% H₂O, 5% CH₃CN, and 0.1% formic acid): B (5% H₂O, 95% CH₃CN, and 0.1% formic acid) as mobile phase, linear gradient from 5%~95% B from 0-12 min, 12-16 min with 95% B) purified peptide **TTI01**.

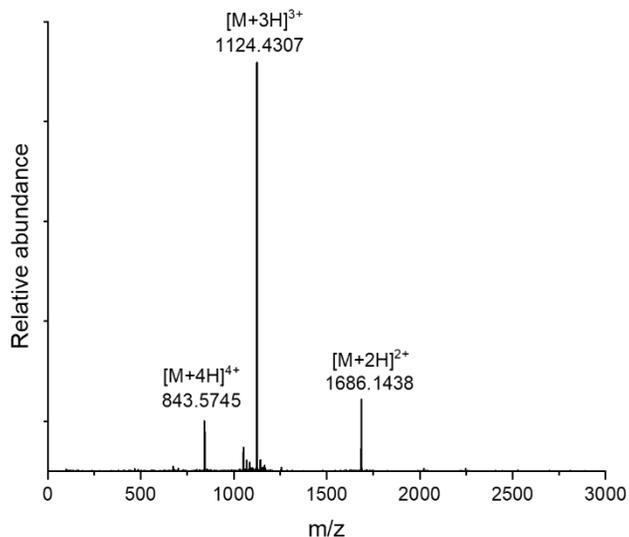


Figure 3-46. ESI-TOF mass spectrum for peptide **TTI01**.

GEPGAPIDfsYDEYGDSSSEEVGGTPLHEIPGIRL-OH [TTI02(fsY)]

The synthesis of peptide **TTI02(fsY)** was separated in two steps. The coupling of first 20 amino acids (counted from the C-terminal) was constructed automatically following general procedure B with Fmoc-L-OH loaded Wang resin (0.025 mmol). After the last coupling step and *N*-terminal Fmoc removal, the resin was washed with DMF and transferred to a plastic vessel to perform the subsequent amino acids incorporation following general procedure A. Purification of the title peptide through the semi-prep RP-HPLC with eluent A (95% H₂O, 5% CH₃CN, and 0.1% TFA) and B (5% H₂O, 95% CH₃CN, and 0.1% TFA). The running method was set as: linear gradient 0-30 min, 20%~35% B. The product peak emerged at $t_R = 26.5$ min, monitored by the UV absorbance at 210 nm and appropriate fractions verified by LC-MS were combined. After concentration with air-blowing, the residue was subjected for lyophilization to provide 18.9 mg of the title compound in 22% yield. Analytical RP-HPLC of this material dissolved in water (mobile phase A: 95% H₂O, 5% CH₃CN, and 0.1% formic acid); mobile phase B: 5% H₂O, 95% CH₃CN, and 0.1% formic acid; linear gradient from 5%~95% B from 0-12 min, 12-16 min with 95% B, and Agilent InfinityLab Poroshell 120 column) showed a single peak at $t_R = 9.0$ min, monitored by the UV absorbance at 210 nm (Figure 3-47). ESI-MS calculated for **TTI02(fsY)** (C₁₄₇H₂₂₀FN₃₇O₅₆S): $[M + 3H]^{3+}$ $m/z = 1151.1737$; found: 1151.7231 (Figure 3-48).

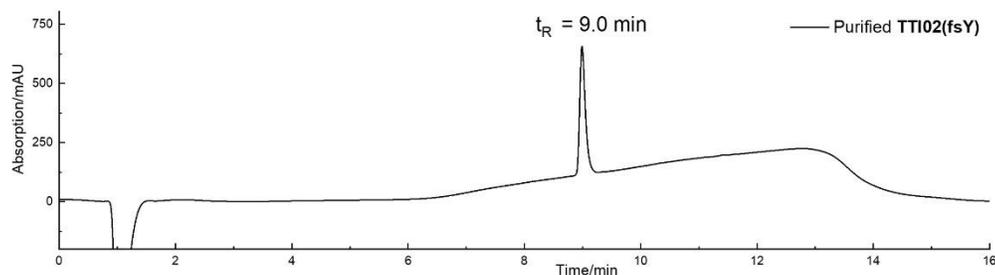


Figure 3-47. Analytical RP-HPLC chromatogram of purified peptide **TTI02(fsY)**. Method: A (95% H₂O, 5% CH₃CN, and 0.1% formic acid); B (5% H₂O, 95% CH₃CN, and 0.1% formic acid) as mobile phase, linear gradient from 5%~95% B from 0-12 min, 12-16 min with 95% B.

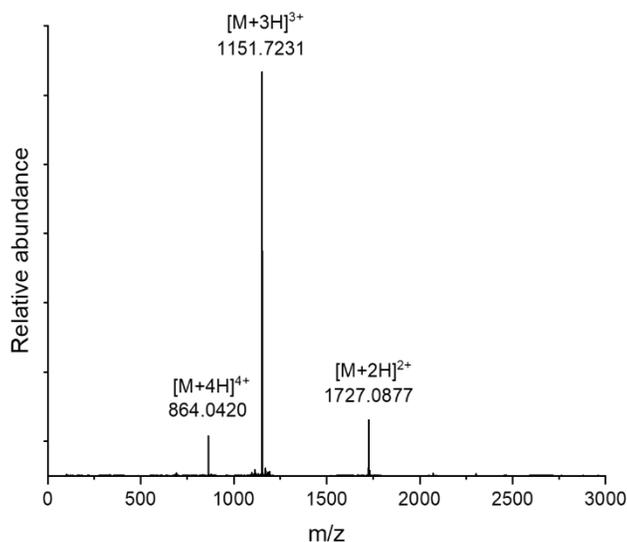


Figure 3-48. ESI-TOF mass spectrum for peptide **TTI02(fsY)**

GEPGAPIDYDEfsYGDSSSEVGGTPLHEIPGIRL-OH [TTI03(fsY)]

The synthesis of peptide **TTI03(fsY)** was separated in two steps. The coupling of first 20 amino acids (counted from the C-terminal) was constructed automatically following general procedure B with Fmoc-L-OH loaded Wang resin (0.025 mmol). After the last coupling step and *N*-terminal Fmoc removal, the resin was washed with DMF and transferred to a plastic vessel to perform the subsequent amino acids incorporation following general procedure A. Purification of the title peptide through the semi-prep RP-HPLC with eluent A (95% H₂O, 5% CH₃CN, and 0.1% TFA) and B (5% H₂O, 95% CH₃CN, and 0.1% TFA). The running method was set as: linear gradient 0-30 min, 20%~35% B. The product peak emerged at *t_R* = 19.7 min, monitored by the UV absorbance at 210 nm and appropriate fractions verified by LC-MS were combined. After concentration with air-blowing, the residue was subjected for lyophilization to provide 22.0 mg of the title compound in 26% yield. Analytical RP-HPLC of this material dissolved in water (mobile phase A: 95% H₂O, 5% CH₃CN, and 0.1% formic acid); mobile phase B: 5% H₂O, 95% CH₃CN, and 0.1% formic acid; linear gradient from 5%~95% B from 0-12 min, 12-16 min with 95% B, and Agilent InfinityLab Poroshell 120 column) showed a single peak at *t_R* = 9.1 min, monitored by the UV absorbance at 210 nm (Figure 3-

49). ESI-MS calculated for **TTI03(fsY)** ($C_{147}H_{220}FN_{37}O_{56}S$): $[M + 3H]^{3+}$ $m/z = 1151.1737$; found: 1151.7433 (Figure 3-50).

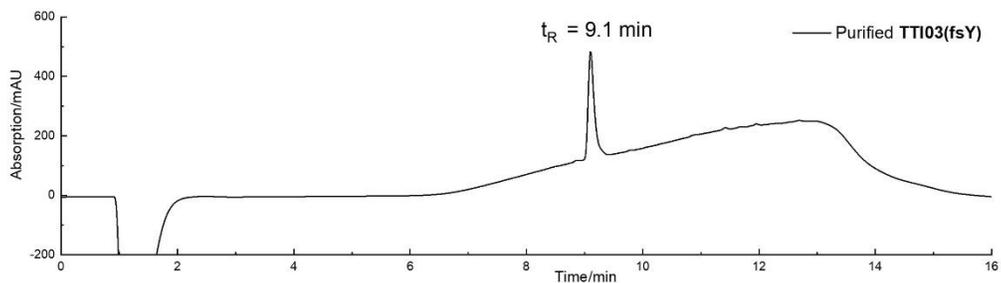


Figure 3-49. Analytical RP-HPLC chromatogram of purified peptide **TTI03(fsY)**. Method: A (95% H₂O, 5% CH₃CN, and 0.1% formic acid): B (5% H₂O, 95% CH₃CN, and 0.1% formic acid) as mobile phase, linear gradient from 5%~95% B from 0-12 min, 12-16 min with 95% B.

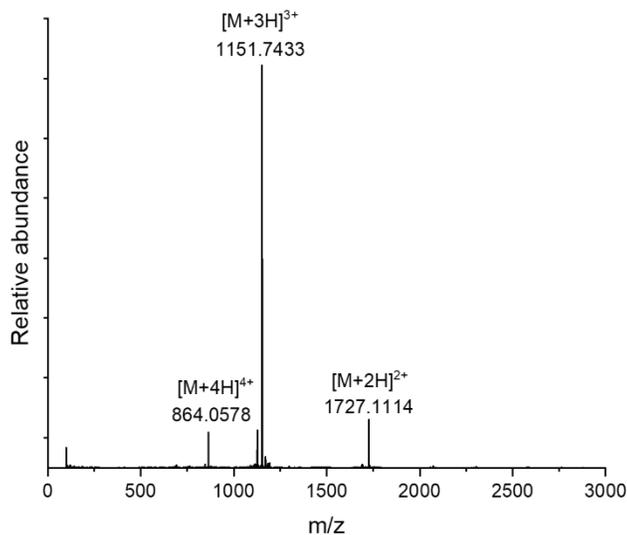


Figure 3-50. ESI-TOF mass spectrum for peptide **TTI03(fsY)**

GEPGAPIDfsYDEfsYGDSSSEVGGTPLHEIPGIRL-OH [TTI04(fsY)]

The synthesis of peptide **TTI04(fsY)** was separated in two steps. The coupling of first 20 amino acids (counted from the C-terminal) was constructed automatically following general procedure B with Fmoc-L-OH loaded Wang resin (0.025 mmol). After the last coupling step and *N*-terminal Fmoc removal, the resin was washed with DMF and transferred to a plastic vessel to perform the subsequent amino acids incorporation following general procedure A. Purification of the title peptide through the semi-prep RP-HPLC with eluent A (95% H₂O, 5% CH₃CN, and 0.1% TFA) and B (5% H₂O, 95% CH₃CN, and 0.1% TFA). The running method was set as: linear gradient 0-30 min, 20%~35% B. The product peak emerged at $t_R = 21.7$ min, monitored by the UV absorbance at 210 nm and appropriate fractions verified by LC-MS were combined. After concentration with air-blowing, the residue was subjected for lyophilization to provide 7.4 mg of the title compound in 8% yield. Analytical RP-HPLC of this material dissolved in water (mobile phase A: 95% H₂O, 5% CH₃CN, and 0.1% formic acid); mobile phase B: 5% H₂O, 95% CH₃CN, and 0.1% formic acid; linear gradient from 5%~95% B from 0-12 min, 12-16 min with 95% B, and Agilent InfinityLab Poroshell 120 column) showed a single peak at $t_R = 9.4$ min under 210 nm wavelength (Figure 3-51). ESI-MS calculated for **TTI04(fsY)** (C₁₄₇H₂₁₉F₂N₃₇O₅₈S₂): [M + 3H]³⁺ $m/z = 1178.4911$; found: 1179.0600 (Figure 3-52).

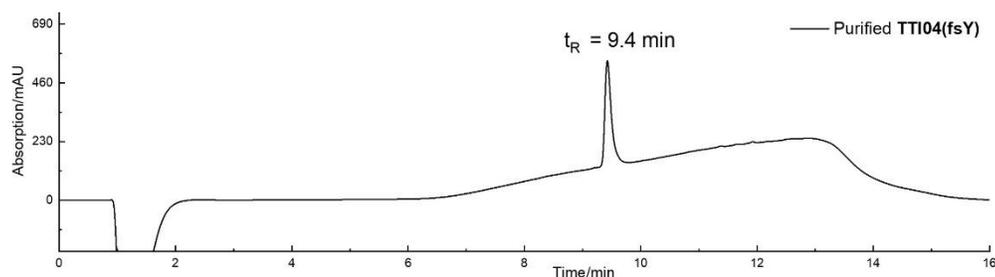


Figure 3-51. Analytical RP-HPLC chromatogram of purified peptide **TTI04(fsY)**. Method: A (95% H₂O, 5% CH₃CN, and 0.1% formic acid): B (5% H₂O, 95% CH₃CN, and 0.1% formic acid) as mobile phase, linear gradient from 5%~95% B from 0-12 min, 12-16 min with 95% B.

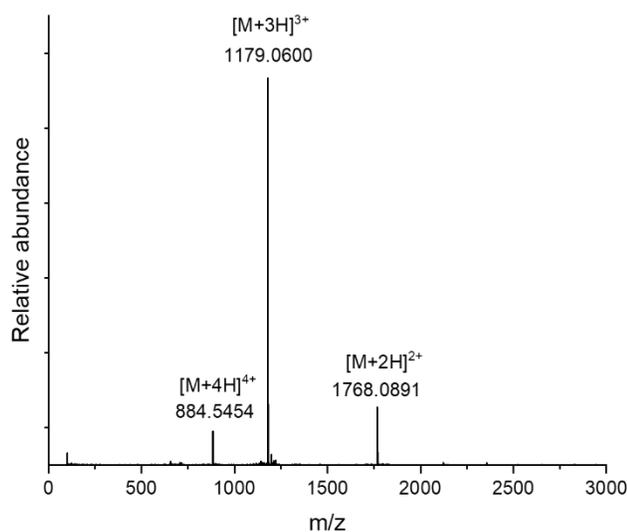


Figure 3-52. ESI-TOF mass spectrum for peptide **TTI04(fsY)**

GEPGAPID_sYDEYGDSSEEVGGTPLHEIPGIRL-OH [TTI02(sY)]

To a prewarmed 0.1 M PBS solution of **TTI02(FsY)** (1.2 mg, 1.0 equiv., 5.78×10^{-4} mmol, 578.0 μ L, pH 7.2, 37 °C), a solution of **7** (0.1 M, 20.0 equiv., 1.16×10^{-3} mmol, 115.6 μ L) was added. The resulting mixture was vigorously stirred at 37 °C for 3 h to give **TTI02(sY)** (Figure 3-53a). The reaction mixture was monitored by RP-HPLC and loaded directly on the semi-prep HPLC for purification with CH₃CN:20mM ammonium acetate buffer as mobile phase. After purification and lyophilization, **TTI02(sY)** white powder (1.1 mg, 3.18×10^{-4} mmol, 55 % isolated yield) was obtained. Purification of the **TTI02(sY)** through the Semi-prep RP-HPLC with CH₃CN and 20 mM ammonium acetate buffer. The running method was set as linear gradient from 15% to 55% CH₃CN within 30 min. The desired peptide was observed at $t_R = 10.6$ min. Analytical HPLC of this material dissolved in water (CH₃CN: H₂O mobile phases, 0.1% TFA, linear gradient from 5%~35% CH₃CN over 30 min) showed a single peak at $t_R = 16.7$ min,

monitored by UV absorbance at 210 nm (Figure 3-53b). ESI-MS calculated for **TTI02(sY)** ($C_{147}H_{221}N_{37}O_{57}S$): $[M - 3H]^3-$ $m/z = 1148.5084$, found: 1149.2087 (Figure 3-54).

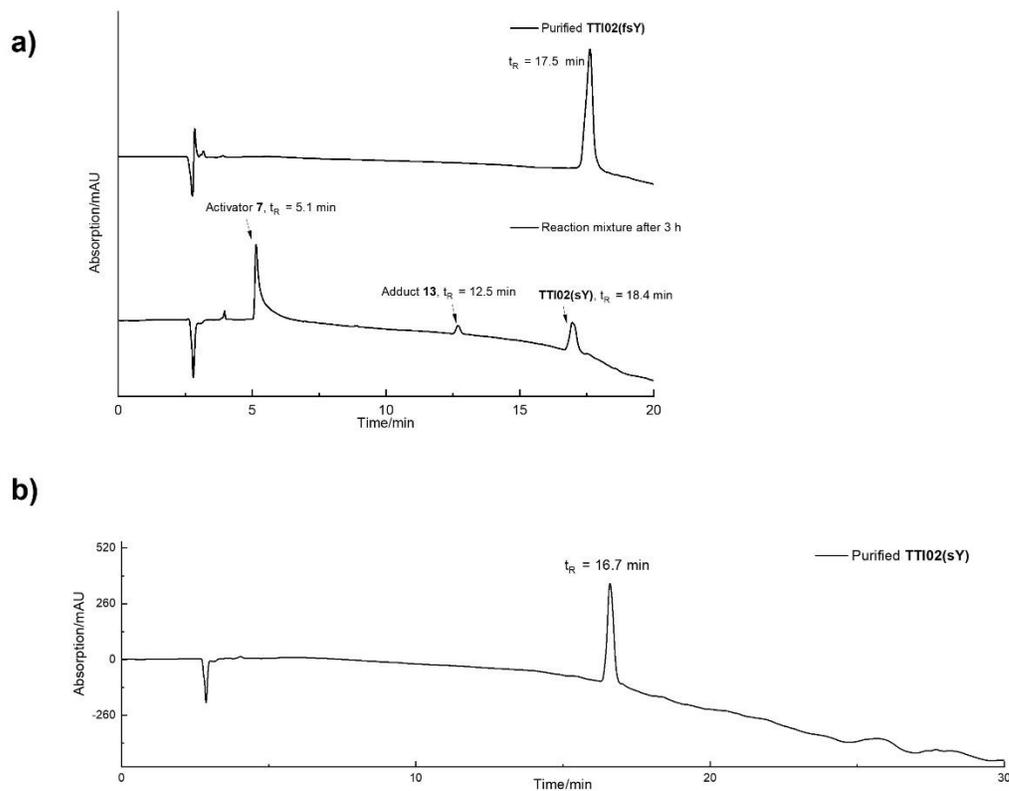


Figure 3-53. a) Analytical RP-HPLC chromatogram of decaging reaction mixture. Method: CH_3CN : 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~35% CH_3CN over 20 min. b) Analytical RP-HPLC chromatogram of purified **TTI02(sY)**. Method: CH_3CN : 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~35% CH_3CN over 30 min.

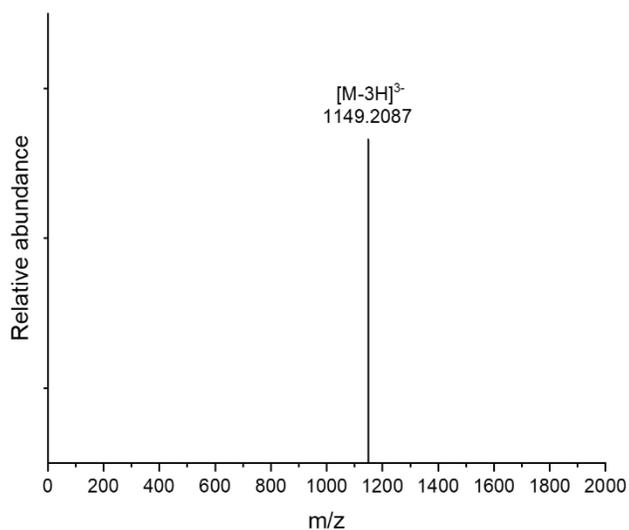


Figure 3-54. ESI-TOF mass spectrum for peptide **TTI02(sY)**

GEPGAPID_sYDE_sYGDSSEEVGGTPLHEIPGIRL-OH [TTI03(sY)]

To a prewarmed 0.1 M PBS solution of **TTI03(fsY)** (1.04 mg, 3.01×10^{-4} mmol, 300.0 μ L, pH 7.2, 37 °C), a solution of **7** (0.1 M, 20.0 equiv., 6.02×10^{-3} mmol, 60.0 μ L) was added. The resulting mixture was vigorously stirred at 37 °C for 3 h to give **TTI03(sY)** (Figure 3-55a). The reaction mixture was monitored by RP-HPLC and loaded directly on the Semi-prep HPLC for purification with CH₃CN:20mM ammonium acetate buffer as mobile phase. After purification and lyophilization, **TTI03(sY)** white powder (0.6 mg, 1.73×10^{-4} mmol, 57 % isolated yield) was obtained. Purification of the **TTI03(sY)** through the Semi-prep RP-HPLC with eluent A (CH₃CN) and B (20 mM ammonium acetate buffer). The running method was set as linear gradient from 15% to 55% within 30 min. The desired peptide was observed at $t_R = 10.7$ min. Analytical HPLC of this material dissolved in water (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~35% CH₃CN over 30 min) showed a single peak at $t_R = 17.0$ min, monitored by UV absorbance at 210 nm (Figure 3-55b). ESI-MS calculated for TTI03(sY) (C₁₄₇H₂₂₁N₃₇O₅₇S): $[M - 3H]^{3-}$ m/z = 1148.5084, found: 1148.8703 (Figure 3-56).

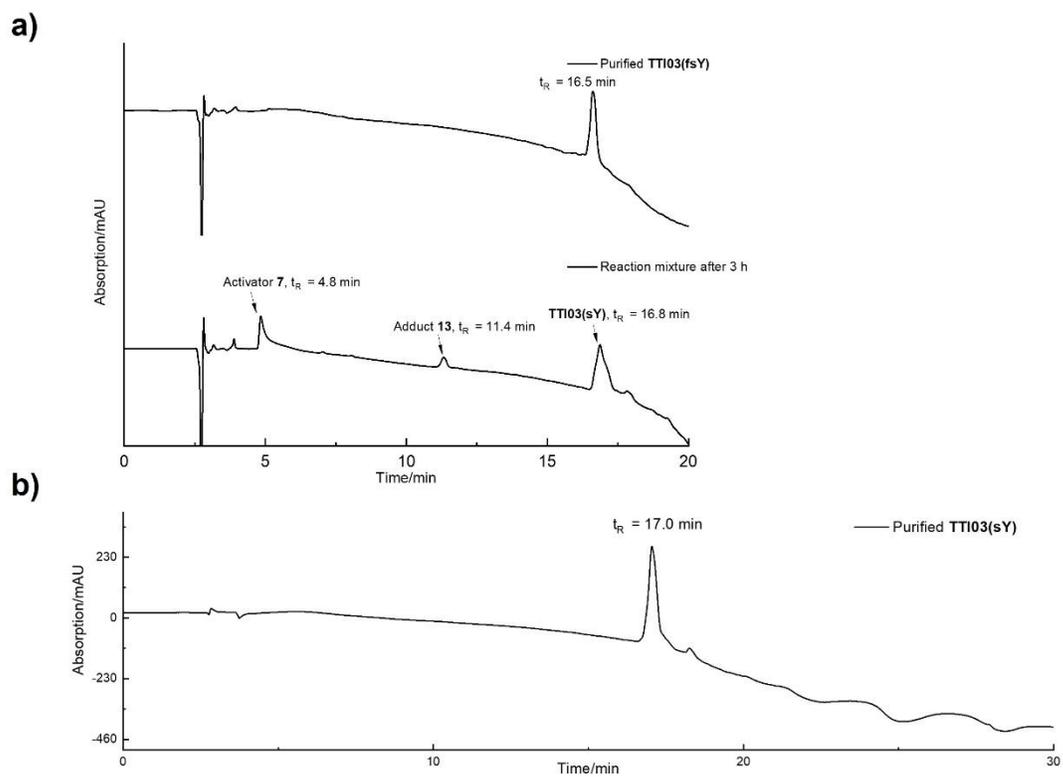


Figure 3-55. a) Analytical RP-HPLC chromatogram of decaging reaction mixture. Method: CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~35% CH₃CN over 20 min. **b)** Analytical RP-HPLC chromatogram of purified **TT103(sY)**. Method: CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~35% CH₃CN over 30 min.

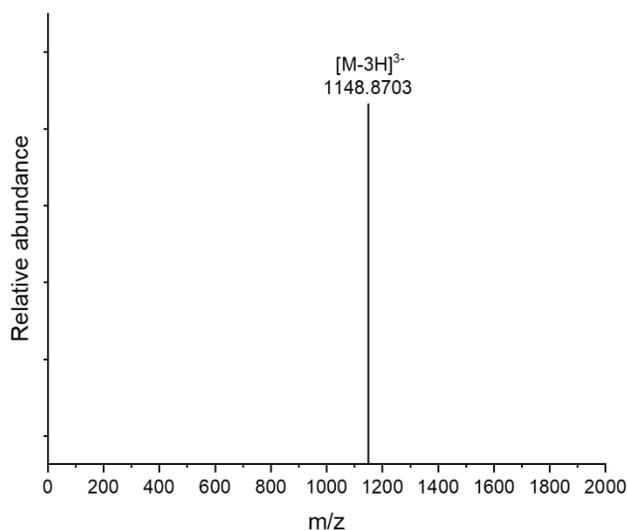


Figure 3-56. ESI-TOF mass spectrum for peptide **TTI03(sY)**

GEPGAPID_sYDE_sYGDSSEEVGGTPLHEIPGIRL-OH [TTI04(sY)]

To a prewarmed 0.1 M PBS solution of **TTI04(fsY)** (1.14 mg, 3.22×10^{-4} mmol, 321.0 μ L, pH 7.2, 37 °C), a solution of **7** (0.1 M, 20.0 equiv., 6.4×10^{-4} mmol, 64.0 μ L) was added. The resulting mixture was vigorously stirred at 37 °C for 4 h to give **TTI04(sY)** (Figure 3-57a). The reaction mixture was monitored by RP-HPLC and loaded directly on the Semi-prep HPLC for purification with CH₃CN:20mM ammonium acetate buffer as mobile phase. After purification and lyophilization, **TTI04(sY)** white powder (0.4 mg, 1.13×10^{-4} mmol, 35 % isolated yield) was obtained. Purification of the TTI04(sY) through the Semi-prep RP-HPLC with eluent A (CH₃CN) and B (20 mM ammonium acetate buffer). The running method was set as linear gradient from 15% to 55% within 30 min. The desired peptide was observed at $t_R = 9.9$ min. Analytical HPLC of this material dissolved in water (CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~35% CH₃CN over 30 min) showed a single peak at $t_R = 17.1$ min under 210 nm wavelength (Figure 3-57b). ESI-MS calculated for **TTI04(sY)** (C₁₄₇H₂₂₁N₃₇O₆₀S₂): [M - 3H]³⁻ m/z = 1175.1607, found: 1175.5287 (Figure 3-58).

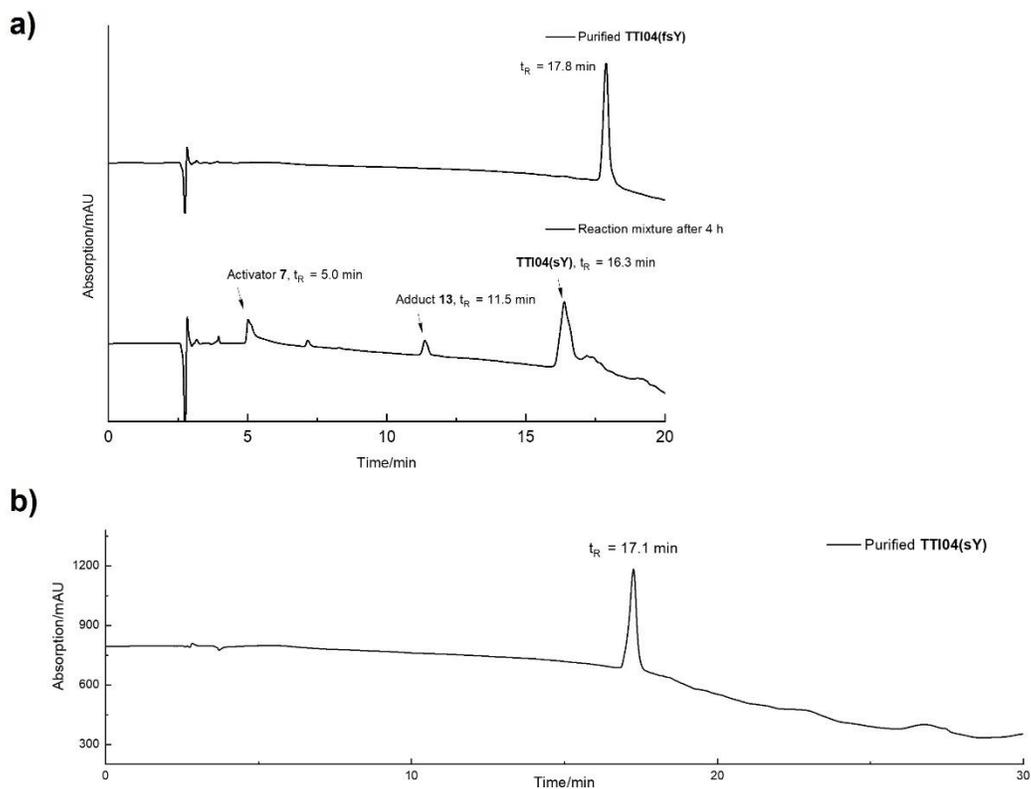


Figure 3-57. a) Analytical RP-HPLC chromatogram of decaying reaction mixture. Method: CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~35% CH₃CN over 20 min. **b)** Analytical RP-HPLC chromatogram of purified **TTI04(sY)**. Method: CH₃CN: 20 mM ammonium acetate buffer mobile phases, linear gradient from 5%~35% CH₃CN over 30 min.

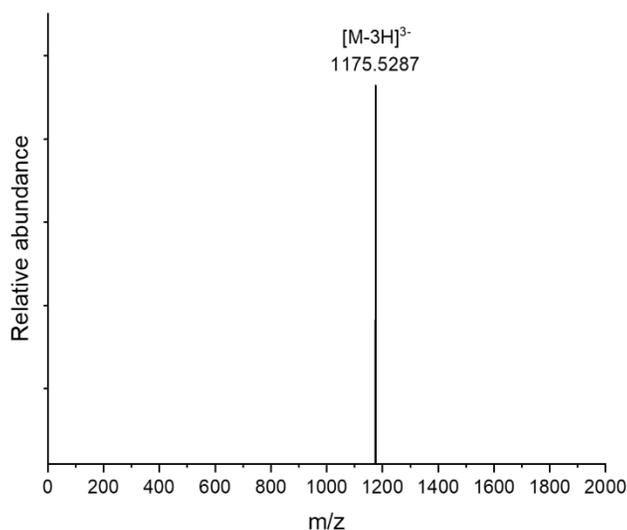


Figure 3-58. ESI-TOF mass spectrum for peptide **TTI04(sY)**

DfsYDHfsYDDKGSO(FAM)SENLVYFQGSLPETGGS-NH₂ (3-20)

The synthesis of the title 27mer peptide **3-20** was separated in two steps. The first 22 amino acids (counted from the *C*-terminal) were constructed automatically following general procedure B with 713.0 mg Rink amide resin (0.3 mmol) was loaded. Among them, easily removed 4-methyltrityl (Mtt) group was used for the side-chain protection of L-ornithine. After the last coupling step and *N*-terminal Fmoc removal, the resin was washed with DMF, dichloromethane, ethyl ether, hexane subsequently (for each solvent 3 mL x 3 min), and dried to afford 3.56 g total resin. 1.78 g dried resin were weighted out and transferred to a Synthware vessel to perform the subsequent amino acids incorporation following general procedure A. Before the last Fmoc-removal step, TFA:TIPS:DCM = 1:2:97 cocktail was applied to the beads 8 mL x 30 min x 3 times to remove the Mtt protection. Then, the resin was washed by DCM (5 mL x 3 min x 2), methanol (5 mL x 3 min x 2), DCM (5 mL x 3 min x 2), 1% DIPEA in DMF (5 mL x 5 min x 2), and DMF (5 mL x 3 min x 2). The fluorescently labeling 5/6-Carboxyfluorescein (FAM, 564.5 mg, 1.50 mmol) was preactivated with *N,N,N',N'*-tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate (HBTU, 568.9 mg, 1.50 mmol) and *N,N*-diisopropylethylamine (DIPEA, 193.9 mg, 261.3 μ L, 1.50 mmol) and then transferred to the resin in the Synthware tube to perform coupling for 2 hours under nitrogen bubbling. Then standard washing

(DMF 3 min x 3), Fmoc removal (20% 2-methylpiperidine 5 min x 3), final deprotection and cleavage (TFA:TIPS:H₂O = 95:2.5:2.5, 30 min x 3), concentration under vacuum, precipitation out of ether. The crude was finally dissolved in a mixture of CH₃CN and water, filtered. Purification of the title peptide through the Prep RP-HPLC with eluent A (95% H₂O, 5% CH₃CN, and 0.1% TFA) and B (5% H₂O, 95% CH₃CN, and 0.1% TFA). The running method was set as: 0-5 min, 5% B; linear gradient 5-30 min, 5%~45% B. The product peak emerged at t_R = 30.5 min, monitored by the UV absorbance at 210 nm and appropriate fractions verified by LC-MS were combined. After concentration with air-blowing, the residue was subjected for lyophilization to provide 89.0 mg of the title compound in 18% yield. Analytical RP-HPLC of this material dissolved in water (CH₃CN: 20 mM ammonium acetate buffer mobile phases, 0-5min, 5% CH₃CN; 5-40 min, linear gradient from 5%~55% CH₃CN over 35 min) showed a single peak at t_R = 24.0 min under 230 nm wavelength (Figure 3-59). HRMS (ESI⁺): calculated for [C₁₅₀H₁₉₄F₂N₃₄O₅₉S₂]²⁺ (M+2H)²⁺ 1759.1329, found 1759.2822 (Figure 3-60).

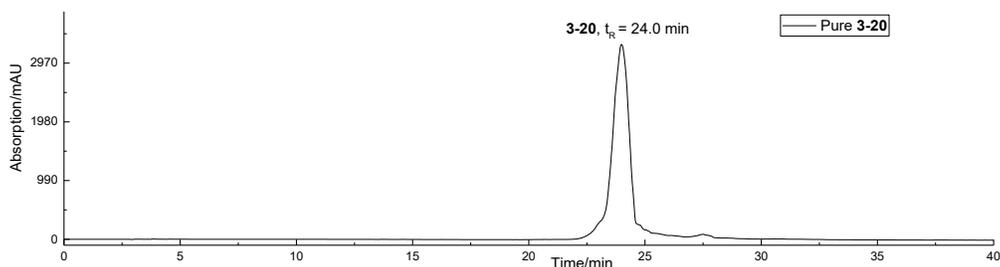


Figure 3-59. Analytical RP-HPLC chromatogram of purified fluorosulfopeptide with fluorescent label **3-20** (CH₃CN: 20 mM ammonium acetate buffer mobile phases, 0-5min, 5% CH₃CN; 5-40 min, linear gradient from 5%~55% CH₃CN over 35 min, monitored by the UV absorbance at 210 nm).

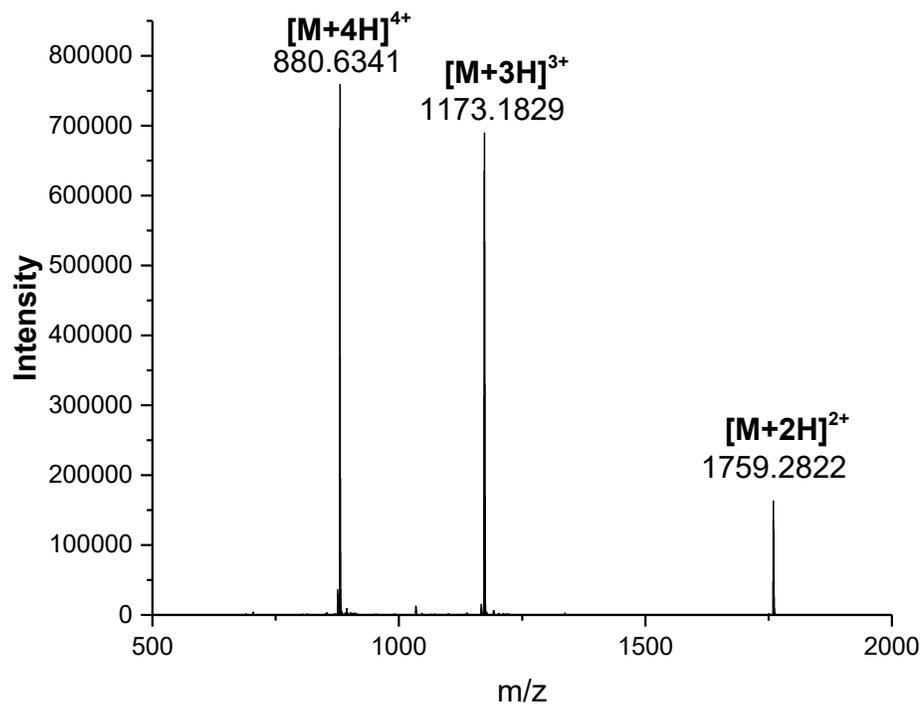


Figure 60. ESI-TOF mass spectrum for peptide **DfsYDHfsYDDKGSO(FAM)SENLYFQGSLPETGGS-NH₂ (22)**

3.4.7 Protein Expression via Unnatural Amino Acid (fsY) Incorporation

Plasmids construction

The pUltra-Opt-MmPylT-TAGc and pEvolT5-EcY-sfGFP151TAG were used as previously reported.⁴⁹ The pUltra-FsTyrRS-OptMmPylT containing FsTyrRS/tRNA pair for fsY incorporation was constructed by insertion of mutated fragment of *Methanosarcina mazei* PylRS (A302I, L305T, N346T, C348I, Y384L and W417K) into the plasmid pUltra-Opt-MmPylT-TAGc via Gibson Assembly.⁵⁹

Fluorescence visualization of cells after sfGFP expression

E. coli B95 cells containing plasmids (1. pET22b-T5.lac-sfGFP; 2. pUltra-FsTyrRS-OptMmPylT and pET22b-T5.lac-sfGFP-151TAG) were grown in LB medium at 37 °C overnight. Then, the cell culture was diluted 100-fold with LB with corresponding antibiotics and cultured in the shaker at 37°C. When the OD600 of the cell culture reached 0.6, protein expression was induced by the addition of 1 mM IPTG and/or non-cannonial amino acids (1. +/-IPTG; 2. +IPTG, +/-fsY). After growth 16 hours at 30 °C, cell pellets from 1 mL cell culture were collected by centrifugation at $4,750 \times g$ for 10 min at 4 °C and resuspended in 1 mL PBS buffer. Then, 100.0 μ L resuspension culture transferred to 96-well black clear bottom assay plates. The sfGFP fluorescence was measured using a bioreader (SpectraMax GeminiEM, Molecular Devices, Excitation: 488/10, Emission: 534/10, cutoff: 515). The normalized fluorescence was obtained by dividing by OD600 value (Figure 3-61).

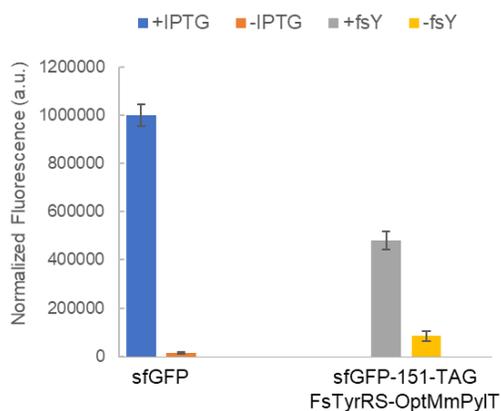


Figure 3-61. Normalized sfGFP fluorescence. Left: sfGFP-wt expression in the presence/absence of IPTG; Right: sfGFP-151-fsY expression in the presence of IPTG and in the presence/absence of fsY

The expression of sfGFP-151-fsY and purification

E. coli B95 cells, cotransformed with pUltra-FsTyrRS-OptMmPylT and pET22b-T5.lac-sfGFP-151TAG were grown in LB medium at 37 °C overnight. Then, the cell culture was diluted 100-fold with 50 mL LB with antibiotics (1 µg/mL Spectinomycin and 3 ug/mL Ampicillin) and cultured in the shaker at 37°C. When the OD600 of the cell culture reached 0.6, protein expression was induced by the addition of 1 mM IPTG and fsY. After growth 16 hours at 30 °C, cell pellets were collected by centrifugation at $4,750 \times g$ for 10 min at 4 °C and resuspended in cool lysis buffer (5 mL B-Per, 0.5 µL universal nuclease and 50.0 µL protease inhibitor). The cool mixture was put in 4 °C shaker for 1 hour and then centrifuged at $4,750 \times g$ for 10 min. The supernatant was collected and dilute with 100 mL equilibrium buffer (20 mM Na₂HPO₄, 300 mM NaCl, 10 mM imidazole, pH 7.4). The equilibrium solution was loaded on chromatography column with pre-equilibrated Ni-NTA agarose resin (1 mL) and washed with 150 mL of wash buffer (20 mM Na₂HPO₄, 300 mM NaCl, 25 mM imidazole, pH

7.4). After the addition of 3 mL elution buffer (20 mM Na₂HPO₄, 300 mM NaCl, 300 mM imidazole, pH 7.4), the fractions of eluates with fluorescence were collected and subjected to concentration and buffer exchange with 0.1M PBS buffer using MilliporeSigma™ Amicon™ Ultra-4 Centrifugal Filter. The protein was confirmed by SDS-PAGE and Mass spectrometric analysis.

Expression and Purification of sfGFP-151-sY

C321-ATMY cells, cotransformed with pBK-sulfo-VGM-CSK and pET22b-T5.lac-sfGFP-151TAG were grown in LB medium at 37 °C for 18 hours.⁶ Then, the cell culture was diluted 100-fold with 50 mL LB with antibiotics (1 µg/mL Chloramphenicol, Kanamycin and Spectinomycin) and cultured in the shaker at 37°C. When the OD600 of the cell culture reached 0.6, protein expression was induced by the addition of 1 mM IPTG and sulfotyrosine (sY). After growth 16 hours at 30 °C, cell pellets were collected by centrifugation at 4,750 × g for 10 min at 4 °C and resuspended in cool lysis buffer (5 mL B-Per, 0.5 µL universal nuclease and 50.0 µL protease inhibitor). The cool mixture was put in 4 °C shaker for 1 hour and then centrifuged at 4,750 × g for 10 min. The supernatant was collected and dilute with 100 mL equilibrium buffer (20 mM Na₂HPO₄, 300 mM NaCl, 10 mM imidazole, pH 7.4). The equilibrium solution was loaded on chromatography column with pre-equilibrated Ni-NTA agarose resin (1 mL) and washed with 150 mL of wash buffer (20 mM Na₂HPO₄, 300 mM NaCl, 25 mM imidazole, pH 7.4). After the addition of 3 mL elution buffer (20 mM Na₂HPO₄, 300 mM NaCl, 300 mM imidazole, pH 7.4), the fractions of eluates with fluorescence were collected and subjected to concentration and buffer exchange with 0.1M PBS buffer using MilliporeSigma™ Amicon™ Ultra-4 Centrifugal Filter. The protein was confirmed by SDS-PAGE and High-resolution Mass spectrometric analysis.

Expression and Purification of sfGFP-wt

E. coli B95 cells, transformed pET22b-T5.lac-sfGFP were grown in LB medium at 37 °C overnight. Then, the cell culture was diluted 100-fold with 50 mL LB with antibiotics (3 ug/mL Ampicillin) and cultured in the shaker at 37°C. When the OD600 of the cell culture reached 0.6, protein expression was induced by the addition of 1 mM IPTG. After growth 16 hours at 30 °C, cell pellets were collected by centrifugation at $4,750 \times g$ for 10 min at 4 °C and resuspended in cool lysis buffer (5 mL B-Per, 0.5 μ L universal nuclease and 50.0 μ L protease inhibitor). The cool mixture was put in 4 °C shaker for 1 hour and then centrifuged at $4,750 \times g$ for 10 min. The supernatant was collected and dilute with 100 mL equilibrium buffer (20 mM Na_2HPO_4 , 300 mM NaCl, 10 mM imidazole, pH 7.4). The equilibrium solution was loaded on chromatography column with pre-equilibrated Ni-NTA agarose resin (1 mL) and washed with 150 mL of wash buffer (20 mM Na_2HPO_4 , 300 mM NaCl, 25 mM imidazole, pH 7.4). After the addition of 3 mL elution buffer (20 mM Na_2HPO_4 , 300 mM NaCl, 300 mM imidazole, pH 7.4), the fractions of eluates with fluorescence were collected and subjected to concentration and buffer exchange with 0.1M PBS buffer using MilliporeSigma™ Amicon™ Ultra-4 Centrifugal Filter. The protein was confirmed by SDS-PAGE and Mass spectrometric analysis (Figure 3-62).

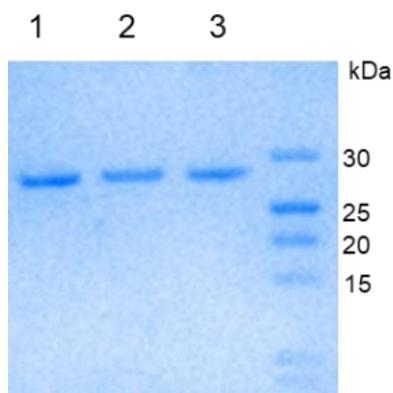


Figure 3-62. SDS-PAGE analysis of purified sfGFP-wt (lane 1), sfGFP-151-sY (lane 2) and sfGFP-151-fsY (lane 3)

3.5 REFERENCES

1. Mehta, A. Y.; Heimbürg-Molinario, J.; Cummings, R. D.; Goth, C. K., Emerging patterns of tyrosine sulfation and O-glycosylation cross-talk and co-localization. *Curr. Opin. Struct. Biol.* **2020**, *62*, 102-111.
2. Huttner, W. B., Sulphation of tyrosine residues—a widespread modification of proteins. *Nature* **1982**, *299* (5880), 273-276.
3. Moore, K. L., The biology and enzymology of protein tyrosine O-sulfation. *J. Biol. Chem.* **2003**, *278* (27), 24243-24246.
4. Baeuerle, P. A.; Huttner, W., Tyrosine sulfation of yolk proteins 1, 2, and 3 in *Drosophila melanogaster*. *J. Biol. Chem.* **1985**, *260* (10), 6434-6439.
5. Ludeman, J. P.; Stone, M. J., The structural role of receptor tyrosine sulfation in chemokine recognition. *Br. J. Pharmacol.* **2014**, *171* (5), 1167-1179.
6. Liu, C. C.; Schultz, P. G., Recombinant expression of selectively sulfated proteins in *Escherichia coli*. *Nat. Biotechnol.* **2006**, *24* (11), 1436-1440.
7. Li, X.; Liu, C. C., Site-Specific Incorporation of Sulfotyrosine Using an Expanded Genetic Code. In *Noncanonical Amino Acids*, Humana Press, New York, NY: 2018; pp 191-200.
8. Italia, J. S.; Peeler, J. C.; Hillenbrand, C. M.; Latour, C.; Weerapana, E.; Chatterjee, A., Genetically encoded protein sulfation in mammalian cells. *Nat. Chem. Biol.* **2020**, *16* (4), 379-382.
9. He, X.; Chen, Y.; Beltran, D. G.; Kelly, M.; Ma, B.; Lawrie, J.; Wang, F.; Dodds, E.; Zhang, L.; Guo, J.; Niu, W., Functional genetic encoding of sulfotyrosine in mammalian cells. *Nat Commun* **2020**, *11* (1), 4820.
10. Chen, Y.; Jin, S.; Zhang, M.; Hu, Y.; Wu, K.-L.; Chung, A.; Wang, S.; Tian, Z.; Wang, Y.; Wolynes, P. G.; Xiao, H., Unleashing the potential of noncanonical amino acid biosynthesis to create cells with precision tyrosine sulfation. *Nat. Commun.* **2022**, *13* (1), 5434.
11. Kehoe, J. W.; Bertozzi, C. R., Tyrosine sulfation: a modulator of extracellular protein-protein interactions. *Chem. Biol.* **2000**, *7* (3), R57-R61.
12. So, W. H.; Wong, C. T. T.; Xia, J., Peptide photocaging: A brief account of the chemistry and biological applications. *Chin. Chem. Lett.* **2018**, *29* (7), 1058-1062.
13. Yang, B.; Wang, N.; Schnier, P. D.; Zheng, F.; Zhu, H.; Polizzi, N. F.; Ittuveetil, A.; Saikam, V.; DeGrado, W. F.; Wang, Q., Genetically introducing biochemically reactive amino acids dehydroalanine and dehydrobutyrine in proteins. *J. Am. Chem. Soc.* **2019**, *141* (19), 7698-7703.
14. Yang, W.; Eken, Y.; Zhang, J.; Cole, L. E.; Ramadan, S.; Xu, Y.; Zhang, Z.; Liu, J.; Wilson, A. K.; Huang, X., Chemical synthesis of human syndecan-4 glycopeptide bearing O-, N-sulfation and multiple aspartic acids for probing impacts of the glycan chain and the core peptide on biological functions. *Chem. Sci.* **2020**, *11*, 6393-6404.
15. Tiruchinapally, G.; Yin, Z.; El-Dakdouki, M.; Wang, Z.; Huang, X., Divergent heparin oligosaccharide synthesis with preinstalled sulfate esters. *Chem. Eur. J.* **2011**, *17* (36), 10106-10112.
16. Al-Horani, R. A.; Desai, U. R., Chemical sulfation of small molecules—advances and challenges. *Tetrahedron* **2010**, *66* (16), 2907.
17. Stone, M. J.; Payne, R. J., Homogeneous sulfopeptides and sulfoproteins: synthetic approaches and applications to characterize the effects of tyrosine sulfation on biochemical function. *Acc. Chem. Res.* **2015**, *48* (8), 2251-61.
18. Ingram, L. J.; Taylor, S. D., Introduction of 2, 2, 2-Trichloroethyl-Protected Sulfates into Monosaccharides with a Sulfuryl Imidazolium Salt and Application to the Synthesis of Sulfated Carbohydrates. *Angew. Chem. Int. Ed.* **2006**, *45* (21), 3503-3506.

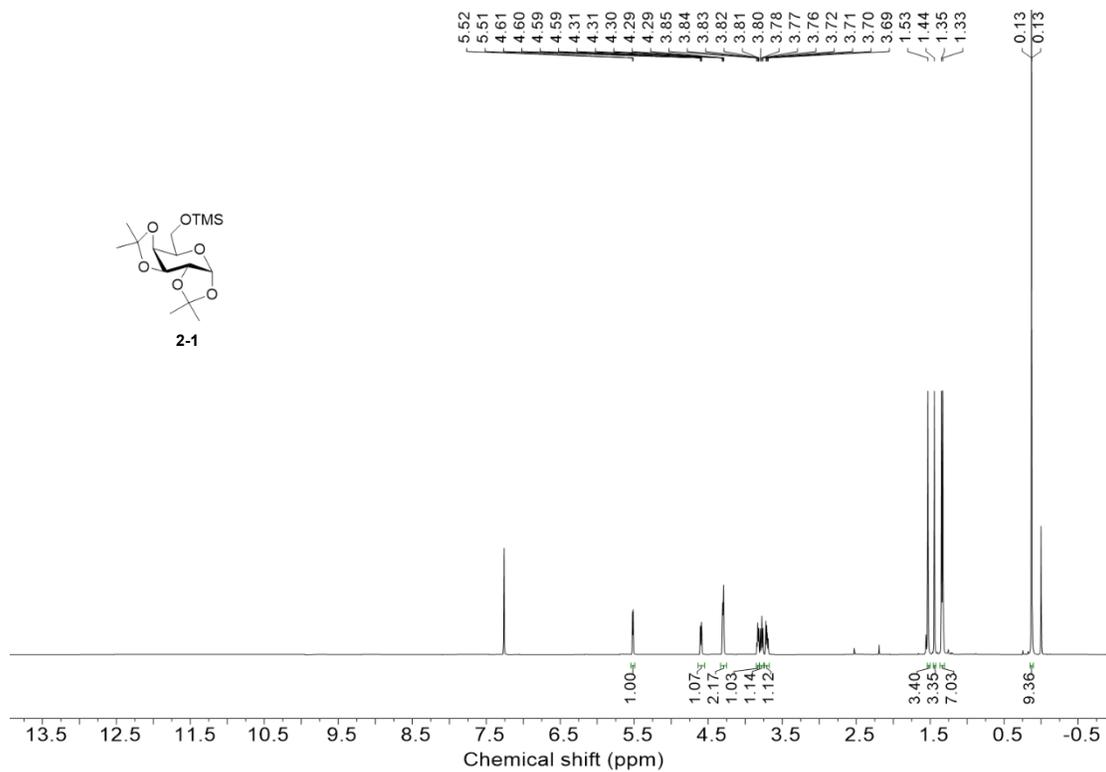
19. Bunschoten, A.; Kruijtzter, J. A.; Ippel, J. H.; de Haas, C. J.; van Strijp, J. A.; Kemmink, J.; Liskamp, R. M., A general sequence independent solid phase method for the site specific synthesis of multiple sulfated-tyrosine containing peptides. *Chem. Commun.* **2009**, (21), 2999-3001.
20. Ali, A. M.; Taylor, S. D., Efficient solid-phase synthesis of sulfotyrosine peptides using a sulfate protecting-group strategy. *Angew. Chem. Int. Ed.* **2009**, *48* (11), 2024-2026.
21. Ali, A. M.; Taylor, S. D., Synthesis of disulfated peptides corresponding to the N-terminus of chemokines receptors CXCR6 (CXCR6 1-20) and DARC (DARC 8-42) using a sulfate-protecting group strategy. *J. Pept. Sci.* **2010**, *16* (4), 190-9.
22. Desoky, A. Y.; Hendel, J.; Ingram, L.; Taylor, S. D., Preparation of trifluoroethyl-and phenyl-protected sulfates using sulfuryl imidazolium salts. *Tetrahedron* **2011**, *67* (6), 1281-1287.
23. Simpson, L. S.; Widlanski, T. S., A comprehensive approach to the synthesis of sulfate esters. *J. Am. Chem. Soc.* **2006**, *128* (5), 1605-1610.
24. Simpson, L. S.; Zhu, J. Z.; Widlanski, T. S.; Stone, M. J., Regulation of chemokine recognition by site-specific tyrosine sulfation of receptor peptides. *Chem. Biol.* **2009**, *16* (2), 153-61.
25. Liu, C.; Yang, C.; Hwang, S.; Ferraro, S. L.; Flynn, J. P.; Niu, J., A General Approach to O-Sulfation by a Sulfur (VI) Fluoride Exchange Reaction. *Angew. Chem. Int. Ed.* **2020**, *132* (42), 18593-18599.
26. Dong, J.; Krasnova, L.; Finn, M.; Sharpless, K. B., Sulfur (VI) fluoride exchange (SuFEx): another good reaction for click chemistry. *Angew. Chem. Int. Ed.* **2014**, *53* (36), 9430-9448.
27. Barrow, A. S.; Smedley, C. J.; Zheng, Q.; Li, S.; Dong, J.; Moses, J. E., The growing applications of SuFEx click chemistry. *Chem. Soc. Rev.* **2019**, *48* (17), 4731-4758.
28. Li, S.; Li, G.; Gao, B.; Pujari, S. P.; Chen, X.; Kim, H.; Zhou, F.; Klivansky, L. M.; Liu, Y.; Driss, H., SuFExable polymers with helical structures derived from thionyl tetrafluoride. *Nat. Chem.* **2021**, *13* (9), 858-867.
29. Miloserdov, F.; Zuilhof, H., Binding S (VI) to alkynes. *Nat. Synth.* **2022**, *1* (6), 415-416.
30. Lee, C.; Cook, A. J.; Elisabeth, J. E.; Friede, N. C.; Sammis, G. M.; Ball, N. D., The Emerging Applications of Sulfur(VI) Fluorides in Catalysis. *ACS Catal.* **2021**, *11*, 6578-6589.
31. Mukherjee, H.; Debreczeni, J.; Breed, J.; Tentarelli, S.; Aquila, B.; Dowling, J.; Whitty, A.; Grimster, N., A study of the reactivity of S (VI)-F containing warheads with nucleophilic amino-acid side chains under physiological conditions. *Org. Biomol. Chem.* **2017**, *15* (45), 9685-9695.
32. Wang, N.; Yang, B.; Fu, C.; Zhu, H.; Zheng, F.; Kobayashi, T.; Liu, J.; Li, S.; Ma, C.; Wang, P. G.; Wang, Q.; Wang, L., Genetically Encoding Fluorosulfate-l-tyrosine To React with Lysine, Histidine, and Tyrosine via SuFEx in Proteins *in vivo*. *J. Am. Chem. Soc.* **2018**, *140* (15), 4995-4999.
33. Chen, W.; Dong, J.; Plate, L.; Mortenson, D. E.; Brighty, G. J.; Li, S.; Liu, Y.; Galmozzi, A.; Lee, P. S.; Hulce, J. J.; Cravatt, B. F.; Saez, E.; Powers, E. T.; Wilson, I. A.; Sharpless, K. B.; Kelly, J. W., Arylfluorosulfates Inactivate Intracellular Lipid Binding Protein(s) through Chemoselective SuFEx Reaction with a Binding Site Tyr Residue. *J. Am. Chem. Soc.* **2016**, *138* (23), 7353-64.
34. Chen, W.; Dong, J.; Li, S.; Liu, Y.; Wang, Y.; Yoon, L.; Wu, P.; Sharpless, K. B.; Kelly, J. W., Synthesis of Sulfotyrosine-Containing Peptides by Incorporating Fluorosulfated Tyrosine Using an Fmoc-Based Solid-Phase Strategy. *Angew. Chem. Int. Ed.* **2016**, *55* (5), 1835-1838.
35. Chen, W.; Dong, J.; Plate, L.; Mortenson, D. E.; Brighty, G. J.; Li, S.; Liu, Y.; Galmozzi, A.; Lee, P. S.; Hulce, J. J., Arylfluorosulfates inactivate intracellular lipid binding protein (s) through chemoselective SuFEx reaction with a binding site Tyr residue. *J. Am. Chem. Soc.* **2016**, *138* (23), 7353-7364.
36. Choi, E. J.; Jung, D.; Kim, J. S.; Lee, Y.; Kim, B. M., Chemoselective Tyrosine Bioconjugation through Sulfate Click Reaction. *Chem. Eur. J.* **2018**, *24* (43), 10948-10952.

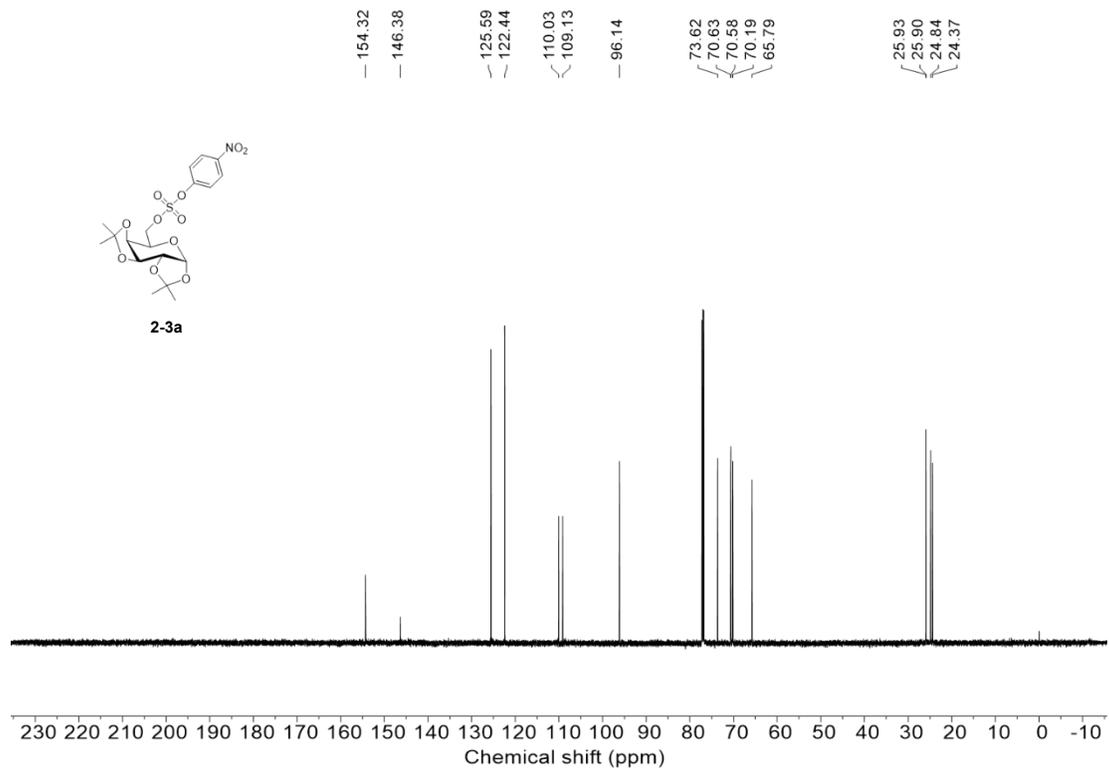
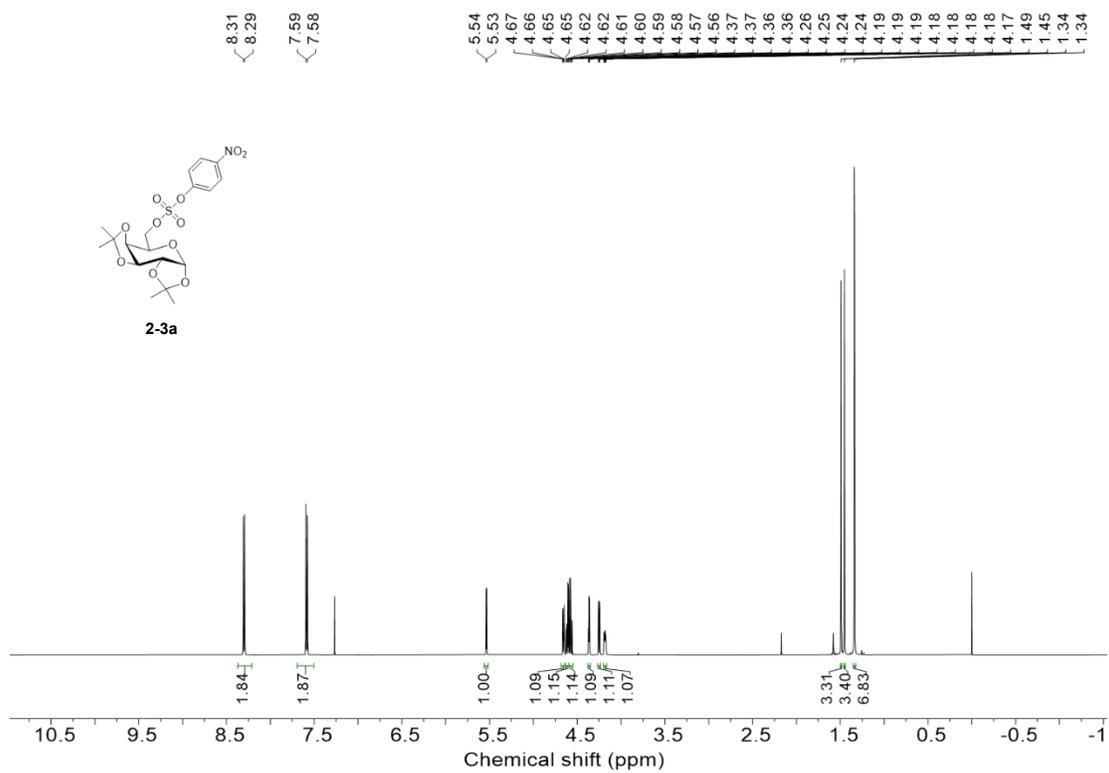
37. El-Faham, A.; Albericio, F., Peptide coupling reagents, more than a letter soup. *Chem. Rev.* **2011**, *111* (11), 6557-602.
38. Eddleston, M.; Buckley, N. A.; Eyer, P.; Dawson, A. H., Management of acute organophosphorus pesticide poisoning. *Lancet* **2008**, *371* (9612), 597-607.
39. Wei, M.; Liang, D.; Cao, X.; Luo, W.; Ma, G.; Liu, Z.; Li, L., A Broad-Spectrum Catalytic Amidation of Sulfonyl Fluorides and Fluorosulfates. *Angew. Chem. Int. Ed.* **2021**, *60* (13), 7397-7404.
40. Hurd, C. D.; Bauer, L., A novel rearrangement of hydroxamic acids using sulfonyl chlorides. *J. Am. Chem. Soc.* **1954**, *76* (10), 2791-2792.
41. Hackley Jr, B.; Plapinger, R.; Stolberg, M.; Wagner-Jauregg, T., Acceleration of the hydrolysis of organic fluorophosphates and fluorophosphonates with hydroxamic acids. *J. Am. Chem. Soc.* **1955**, *77* (13), 3651-3653.
42. Bones, A. M.; Rossiter, J. T., The myrosinase-glucosinolate system, its organisation and biochemistry. *Physiol. Plant.* **1996**, *97* (1), 194-208.
43. Ippel, J. H.; de Haas, C. J.; Bunschoten, A.; van Strijp, J. A.; Kruijtzter, J. A.; Liskamp, R. M.; Kemmink, J., Structure of the tyrosine-sulfated C5a receptor N terminus in complex with chemotaxis inhibitory protein of *Staphylococcus aureus*. *J. Biol. Chem.* **2009**, *284* (18), 12363-12372.
44. Hansen, M. J.; Velema, W. A.; Lerch, M. M.; Szymanski, W.; Feringa, B. L., Wavelength-selective cleavage of photoprotecting groups: strategies and applications in dynamic systems. *Chem. Soc. Rev.* **2015**, *44* (11), 3358-3377.
45. Cappello, M.; Bergum, P. W.; Vlasuk, G. P.; Furmidge, B. A.; Pritchard, D. I.; Aksoy, S., Isolation and characterization of the tsetse thrombin inhibitor: a potent antithrombotic peptide from the saliva of *Glossina morsitans morsitans*. *Am. J. Trop. Med. Hyg.* **1996**, *54* (5), 475-480.
46. Calisto, B. M.; Ripoll-Rozada, J.; Dowman, L. J.; Franck, C.; Agten, S. M.; Parker, B. L.; Veloso, R. C.; Vale, N.; Gomes, P.; de Sanctis, D.; Payne, R. J.; Pereira, P. J. B., Sulfotyrosine-mediated recognition of human thrombin by a tsetse fly anticoagulant mimics physiological substrates. *Cell Chem. Biol.* **2021**, *28* (1), 26-33.
47. Spannagl, M.; Bichler, J.; Birg, A.; Lill, H.; Schramm, W., Development of a chromogenic substrate assay for the determination of hirudin in plasma. *Blood Coagul. Fibrinolysis* **1991**, *2* (1), 121-127.
48. Williams, J. W.; Morrison, J. F., The kinetics of reversible tight-binding inhibition. In *Methods Enzymol.*, Elsevier: 1979; Vol. 63, pp 437-467.
49. Chatterjee, A.; Sun, S. B.; Furman, J. L.; Xiao, H.; Schultz, P. G., A versatile platform for single-and multiple-unnatural amino acid mutagenesis in *Escherichia coli*. *Biochemistry* **2013**, *52* (10), 1828-1837.
50. Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F., Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **2010**, *468* (7325), 790-795.
51. Scheffler, K.; Viner, R.; Damoc, E., High resolution top-down experimental strategies on the Orbitrap platform. *J. Proteomics* **2018**, *175*, 42-55.
52. HILLE, A.; HUTTNER, W. B., Occurrence of tyrosine sulfate in proteins—a balance sheet: 2. Membrane proteins. *Eur. J. Biochem.* **1990**, *188* (3), 587-596.
53. Reja, R. M.; Wang, W.; Lyu, Y.; Haeffner, F.; Gao, J., Lysine-Targeting Reversible Covalent Inhibitors with Long Residence Time. *J. Am. Chem. Soc.* **2022**, *144* (3), 1152-1157.
54. Rentero Rebollo, I.; McCallin, S.; Bertoldo, D.; Entenza, J. M.; Moreillon, P.; Heinis, C., Development of potent and selective *S. aureus* sortase A inhibitors based on peptide macrocycles. *ACS medicinal chemistry letters* **2016**, *7* (6), 606-611.

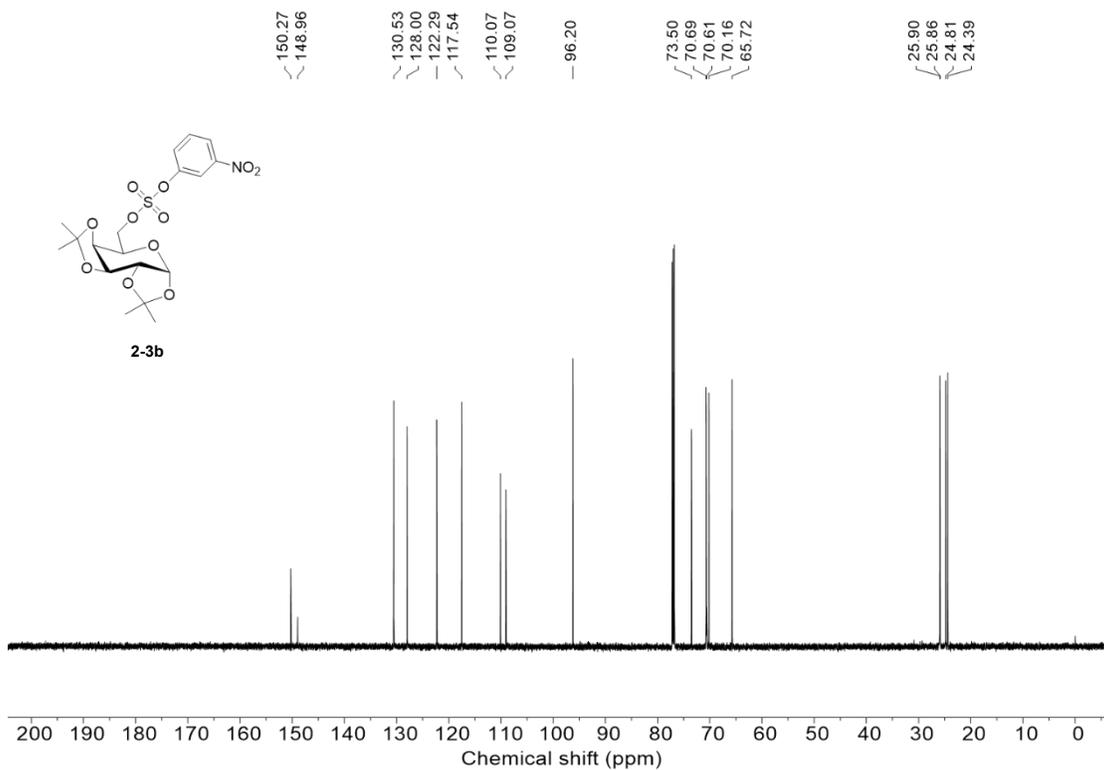
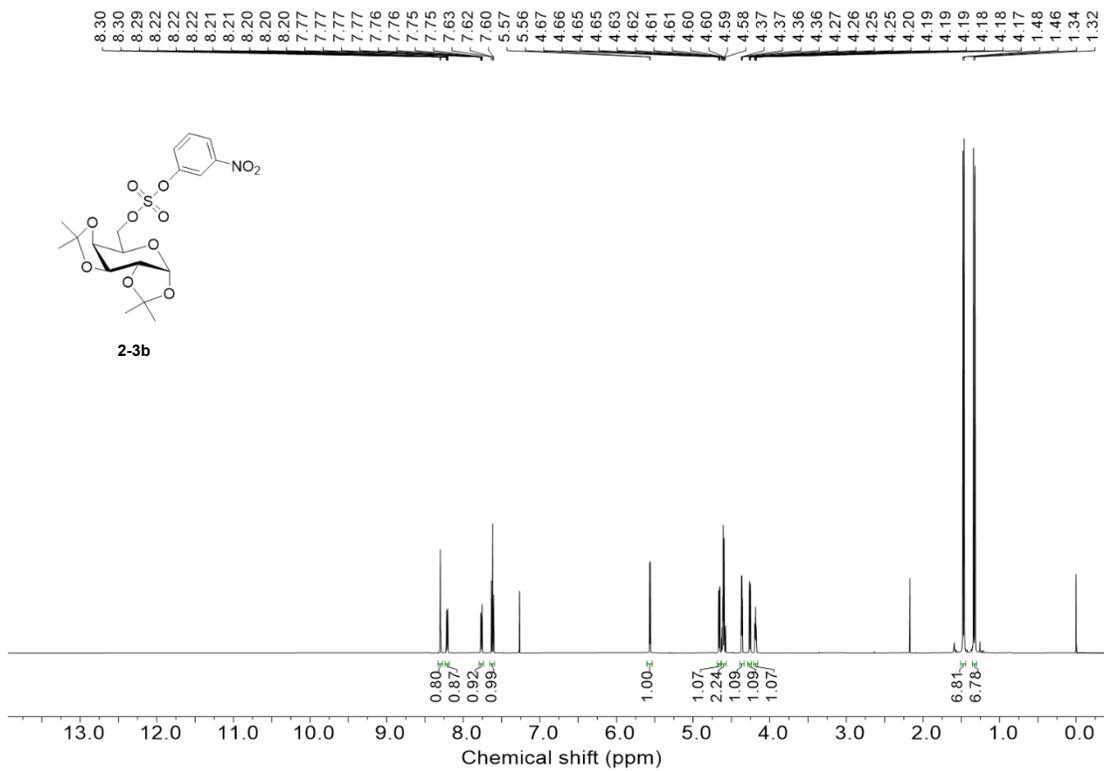
55. Kapust, R. B.; Tözsér, J.; Copeland, T. D.; Waugh, D. S., The P1' specificity of tobacco etch virus protease. *Biochem. Biophys. Res. Commun.* **2002**, *294* (5), 949-955.
56. Mosmann, T., Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J. Immunol. Methods* **1983**, *65* (1-2), 55-63.
57. Glasoe, P. K.; Long, F., Use of glass electrodes to measure acidities in deuterium oxide^{1, 2}. *J. Phys. Chem.* **1960**, *64* (1), 188-190.
58. Mulcahy, C.; Dolgushin, F. M.; Krot, K. A.; Griffith, D.; Marmion, C. J., Synthesis, characterisation and speciation studies of heterobimetallic pyridinehydroxamate-bridged Pt (II)/M (II) complexes (M= Cu, Ni, Zn). Crystal structure of a novel heterobimetallic 3-pyridinehydroxamate-bridged Pt (II)/Cu (II) wave-like coordination polymer. *Dalton Trans.* **2005**, (11), 1993-1998.
59. Wang, N.; Yang, B.; Fu, C.; Zhu, H.; Zheng, F.; Kobayashi, T.; Liu, J.; Li, S.; Ma, C.; Wang, P. G., Genetically encoding fluorosulfate-L-tyrosine to react with lysine, histidine, and tyrosine via SuFEx in proteins *in vivo*. *J. Am. Chem. Soc.* **2018**, *140* (15), 4995-4999.

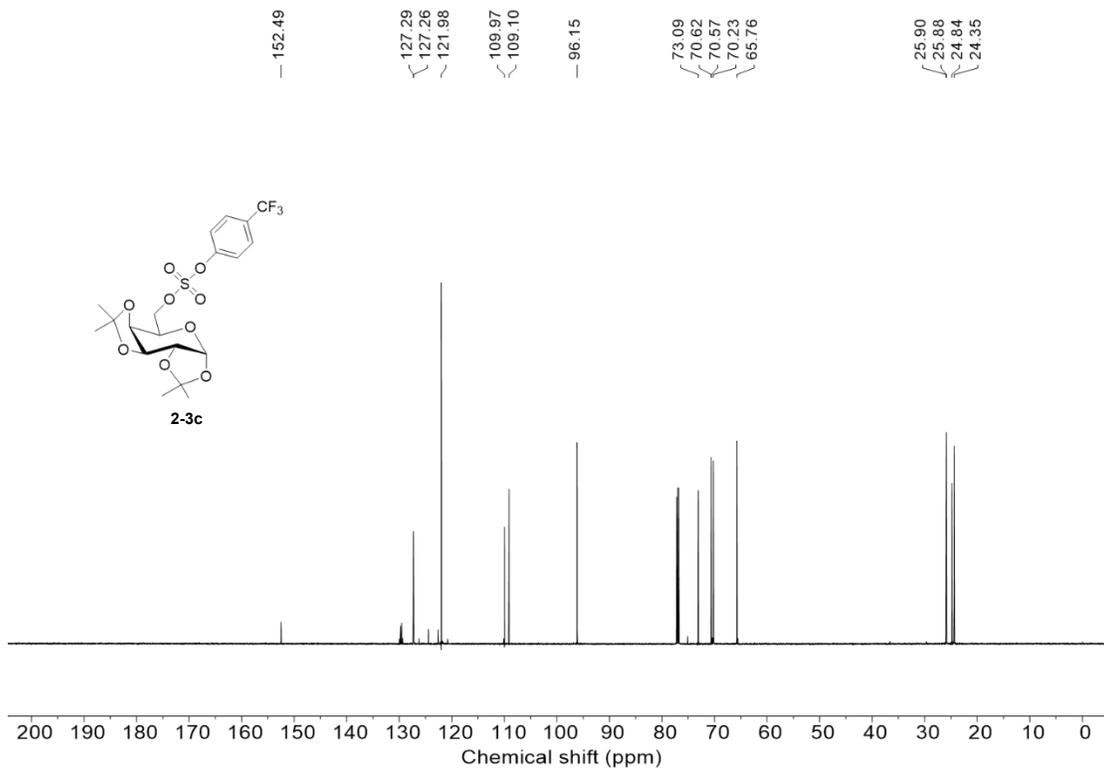
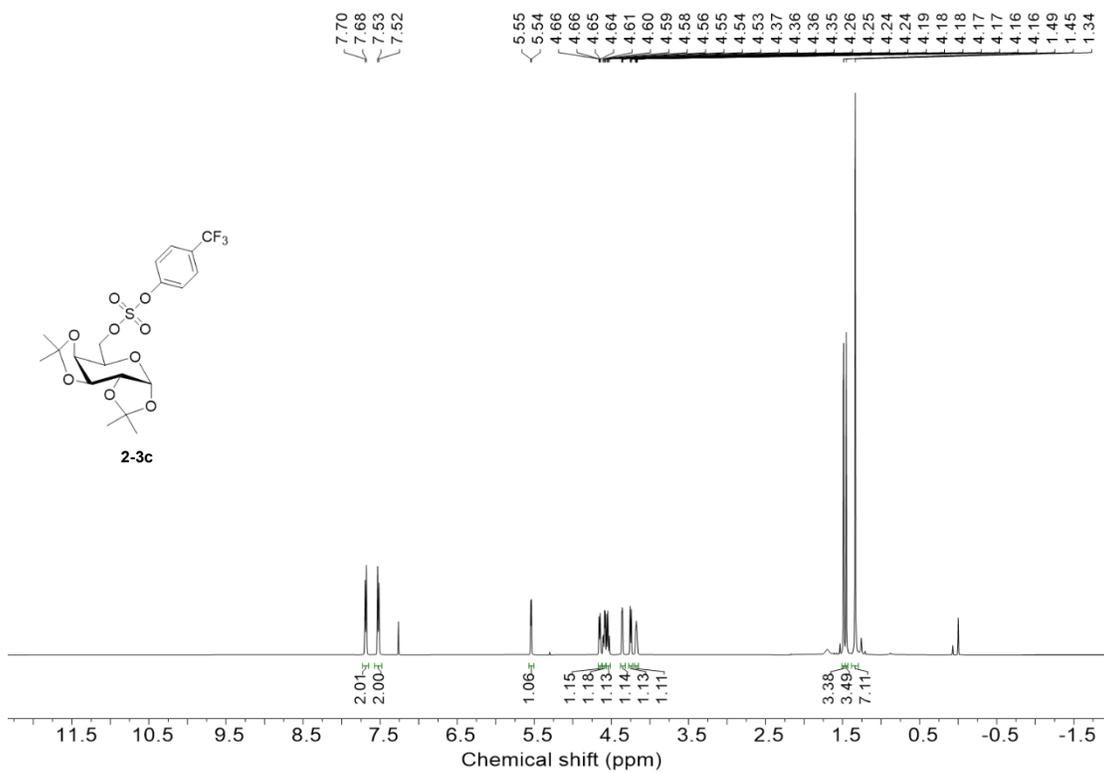
Appendix

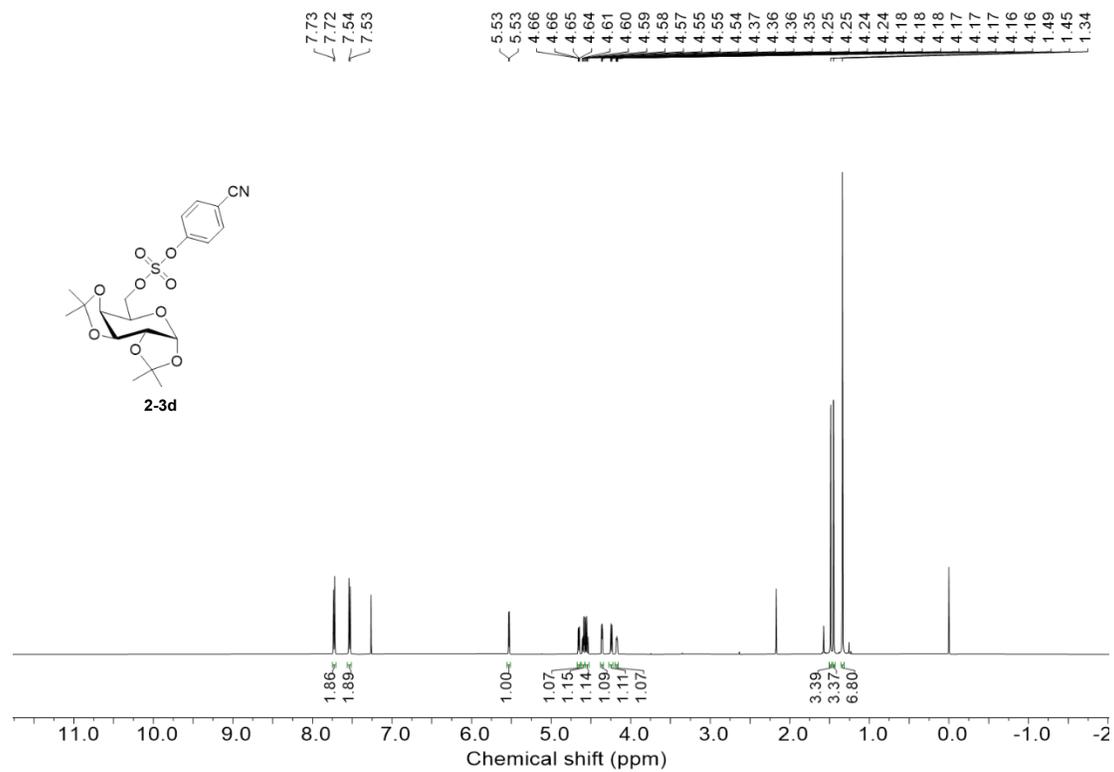
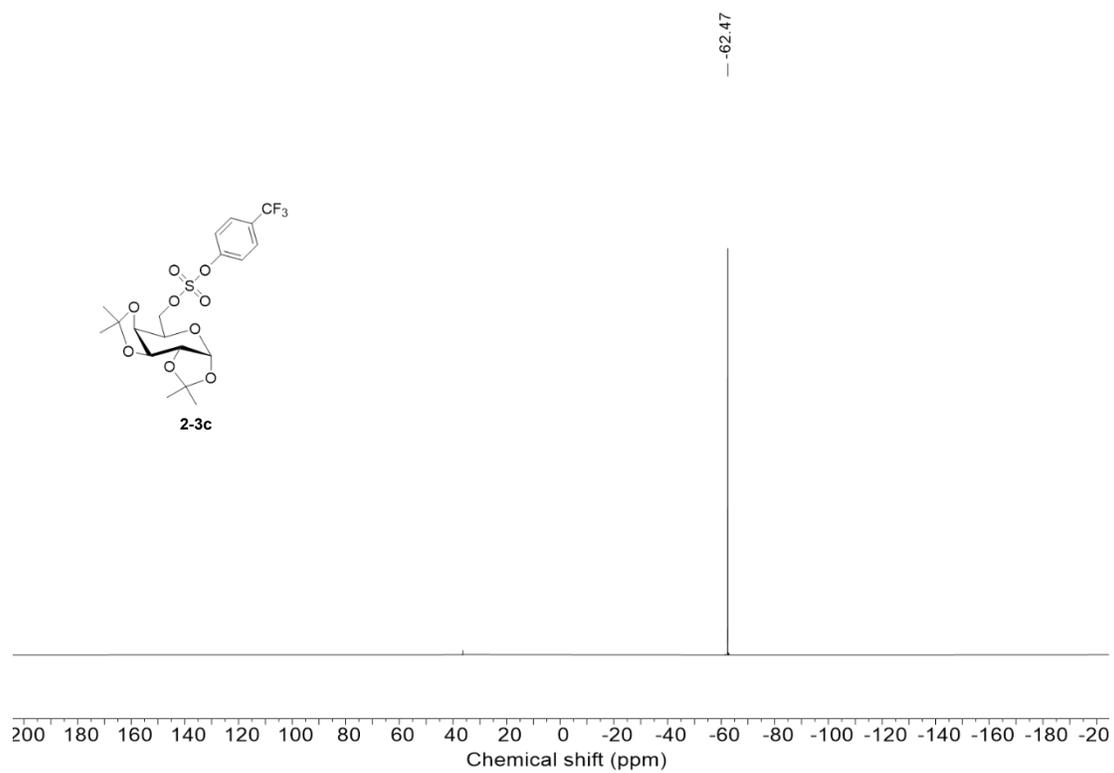
Original NMR spectra of new compounds in Chapter 2

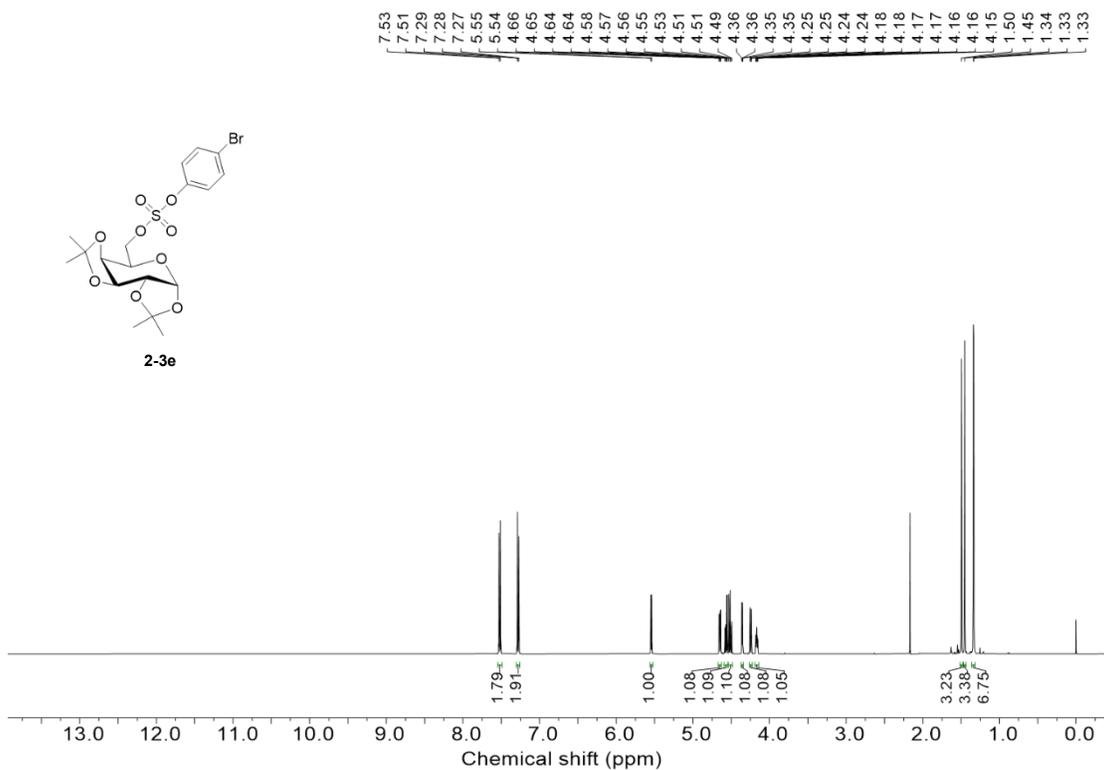
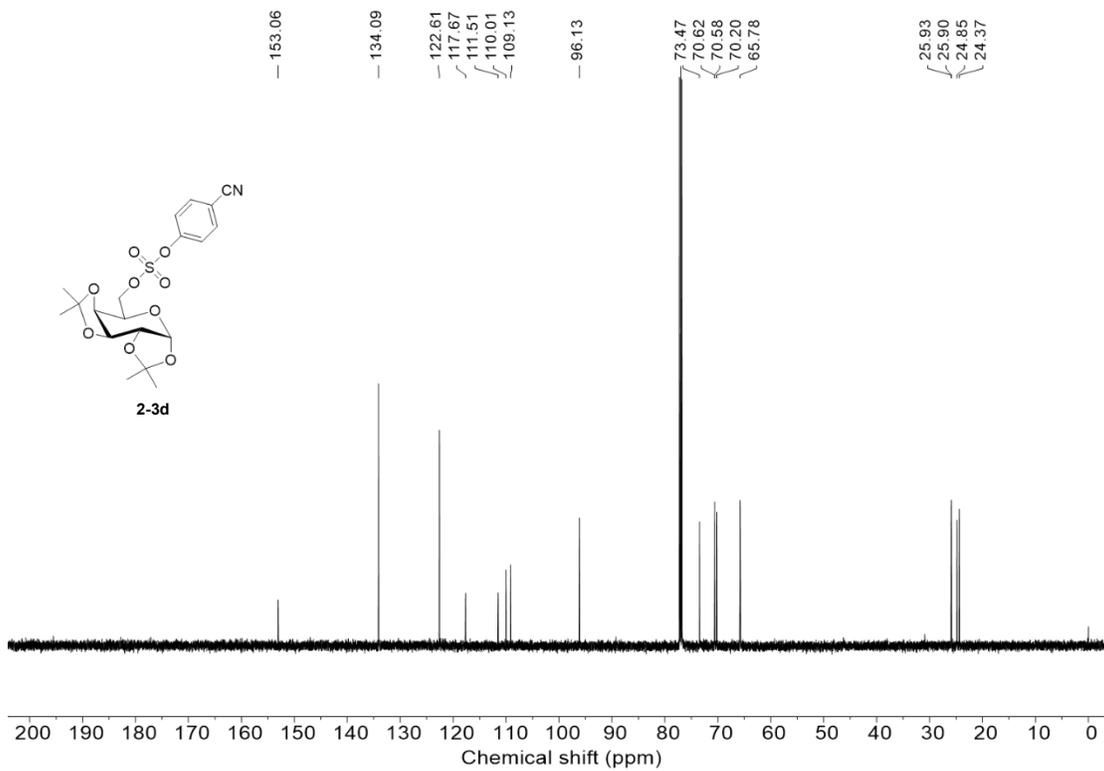


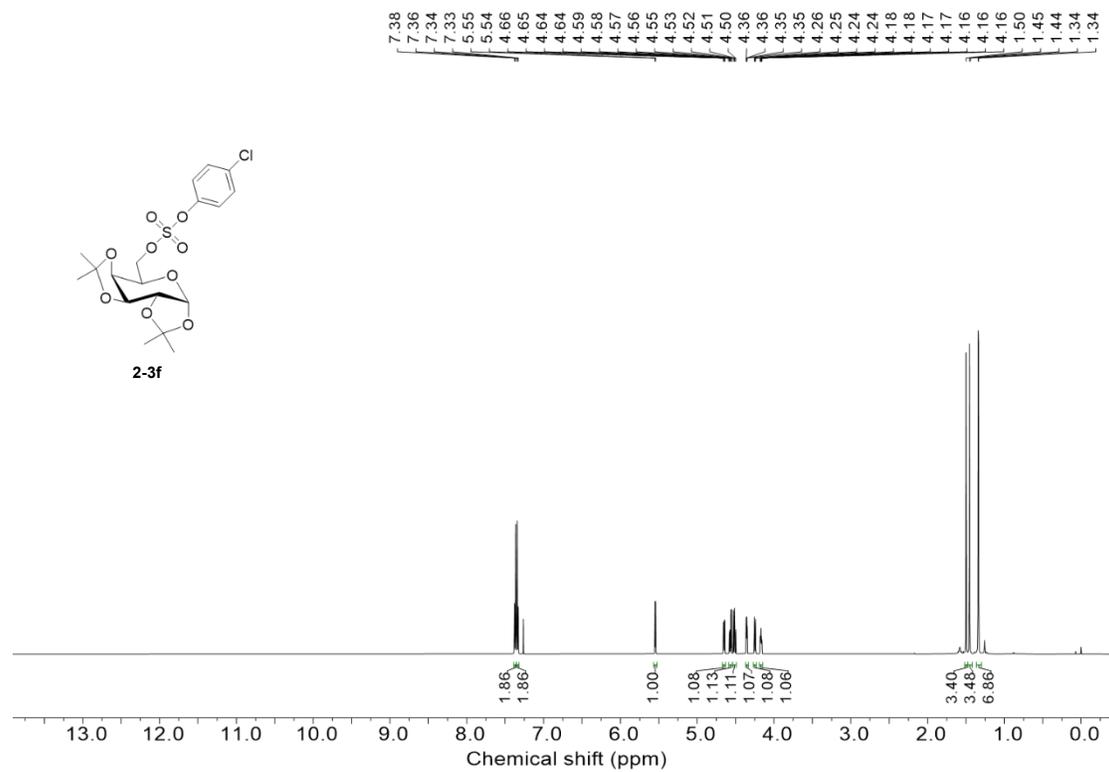
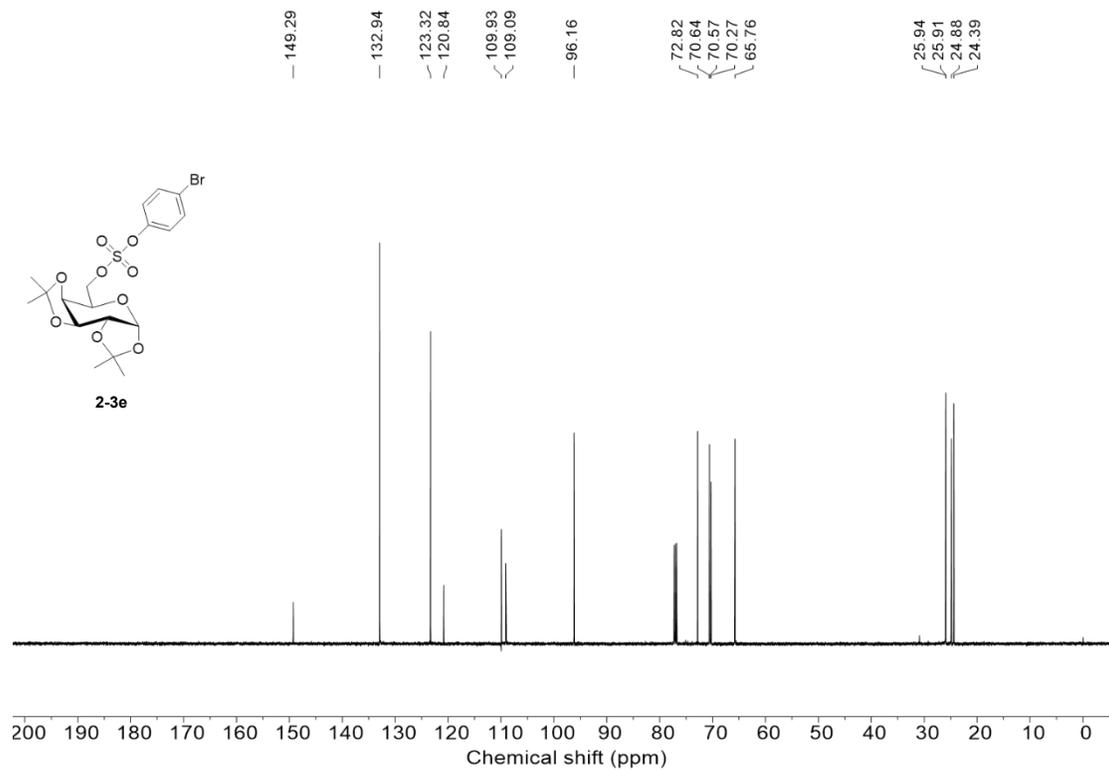


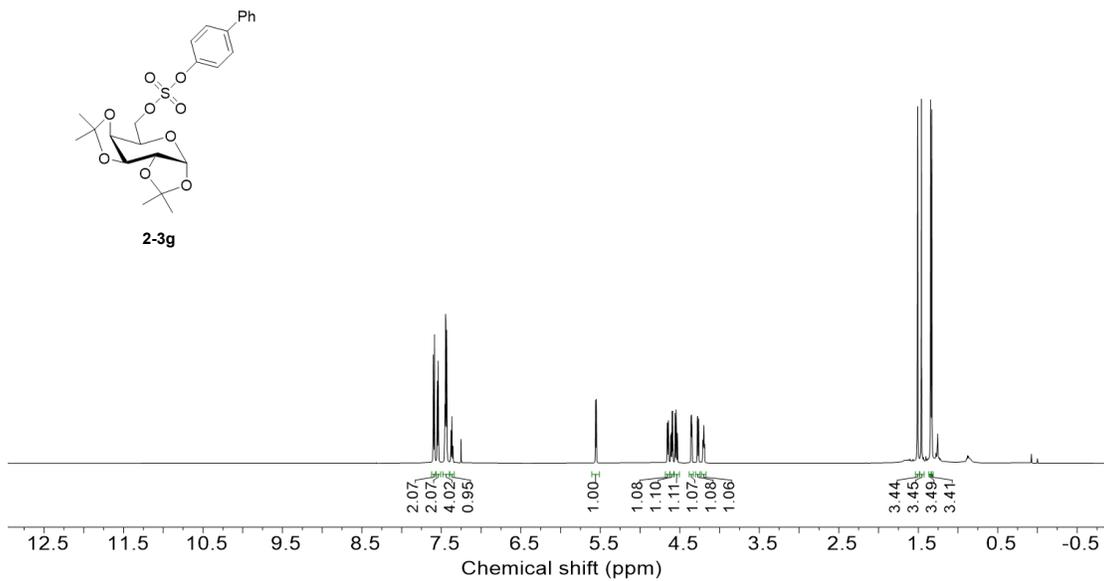
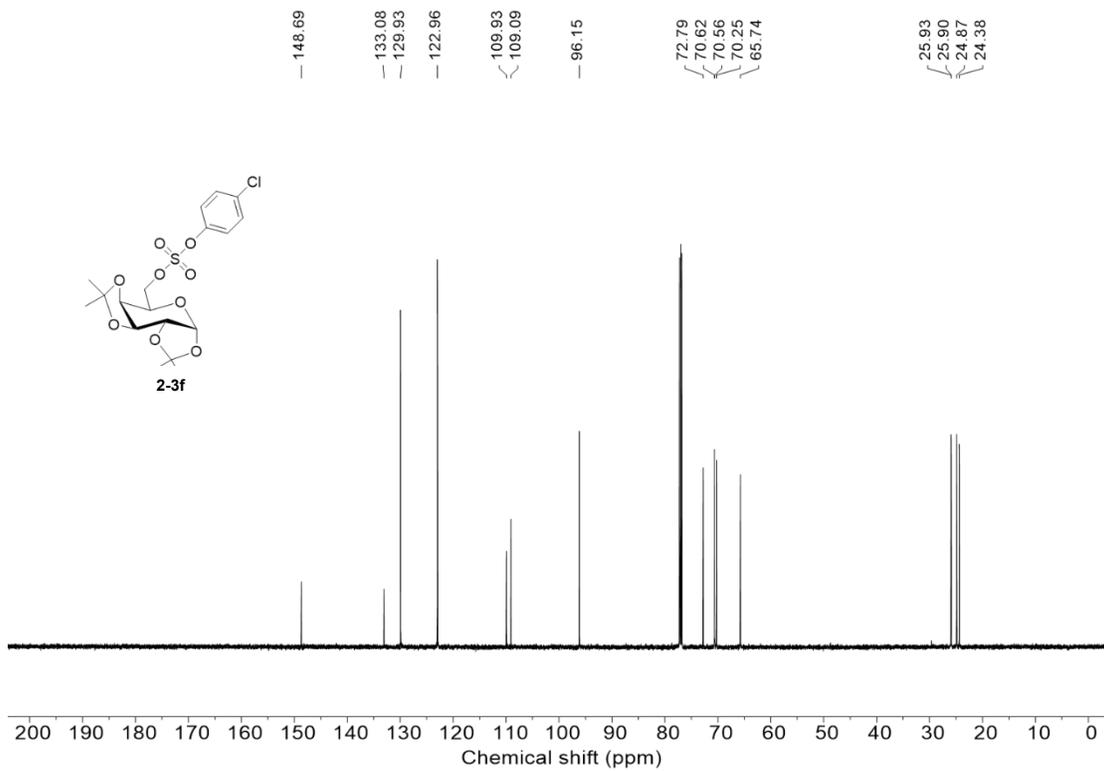


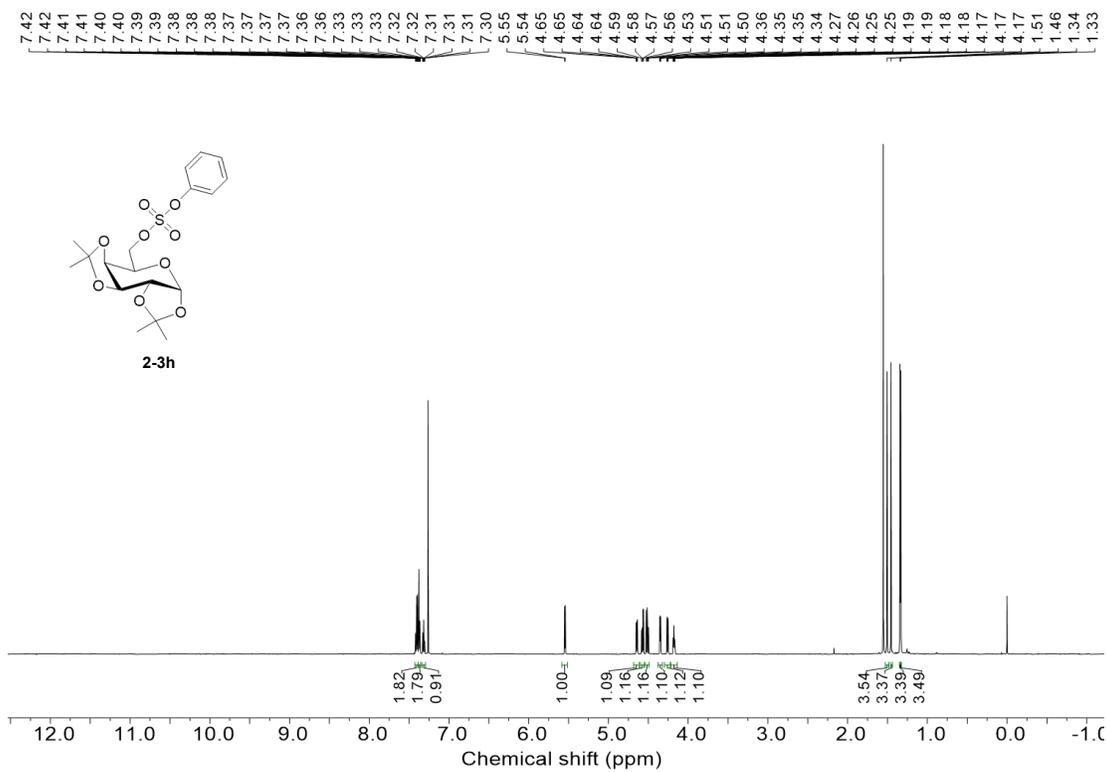
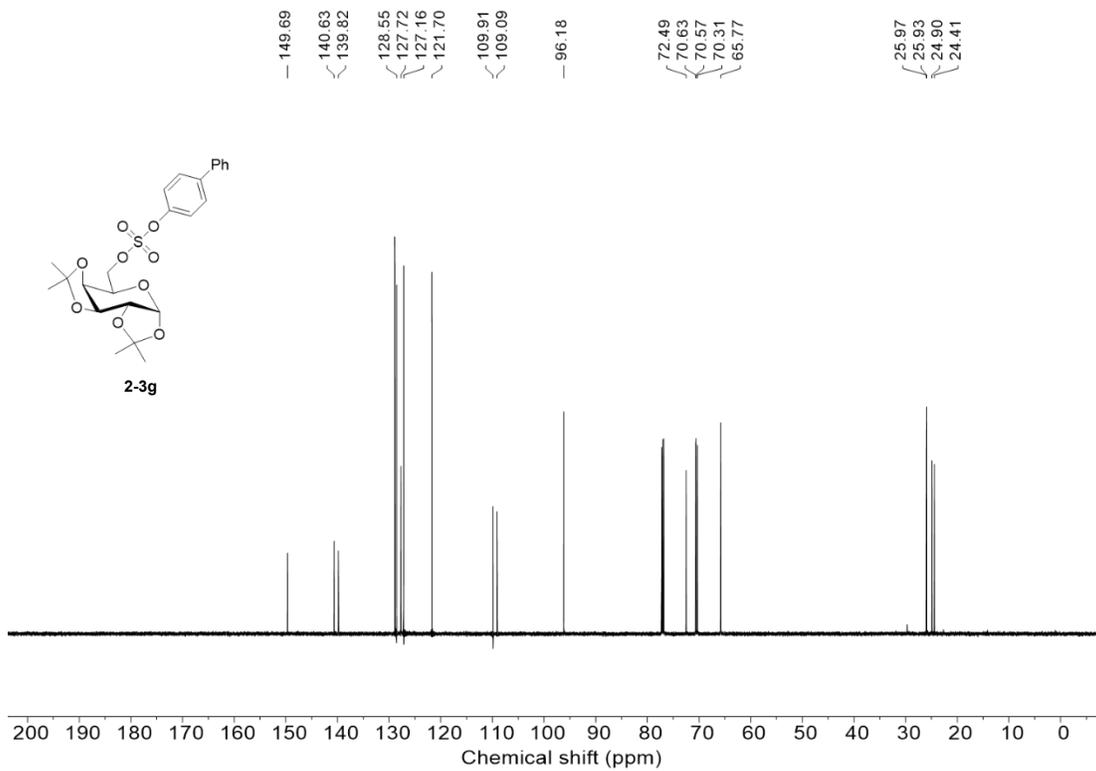


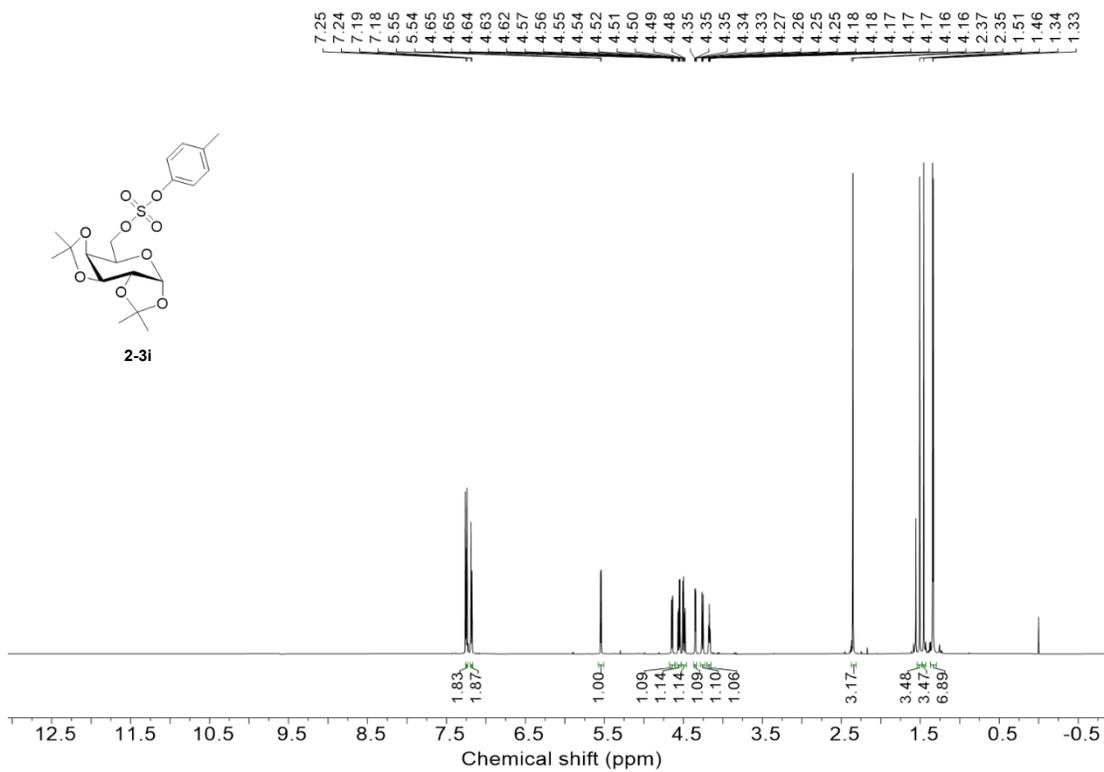
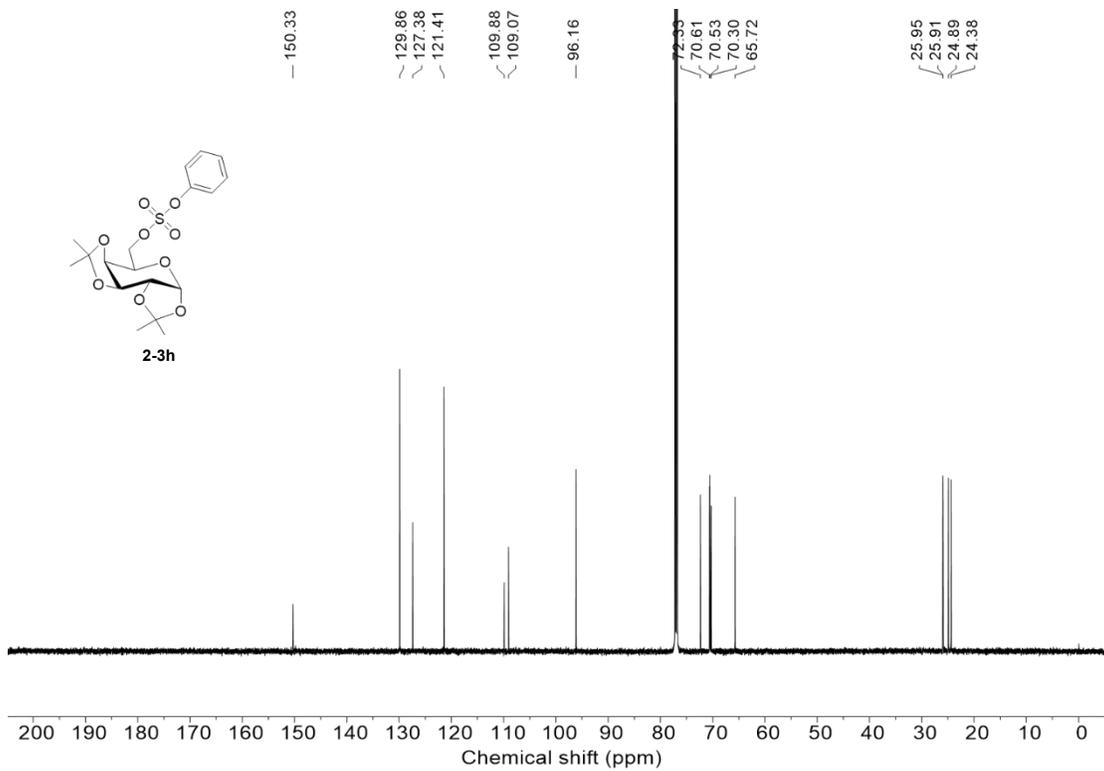


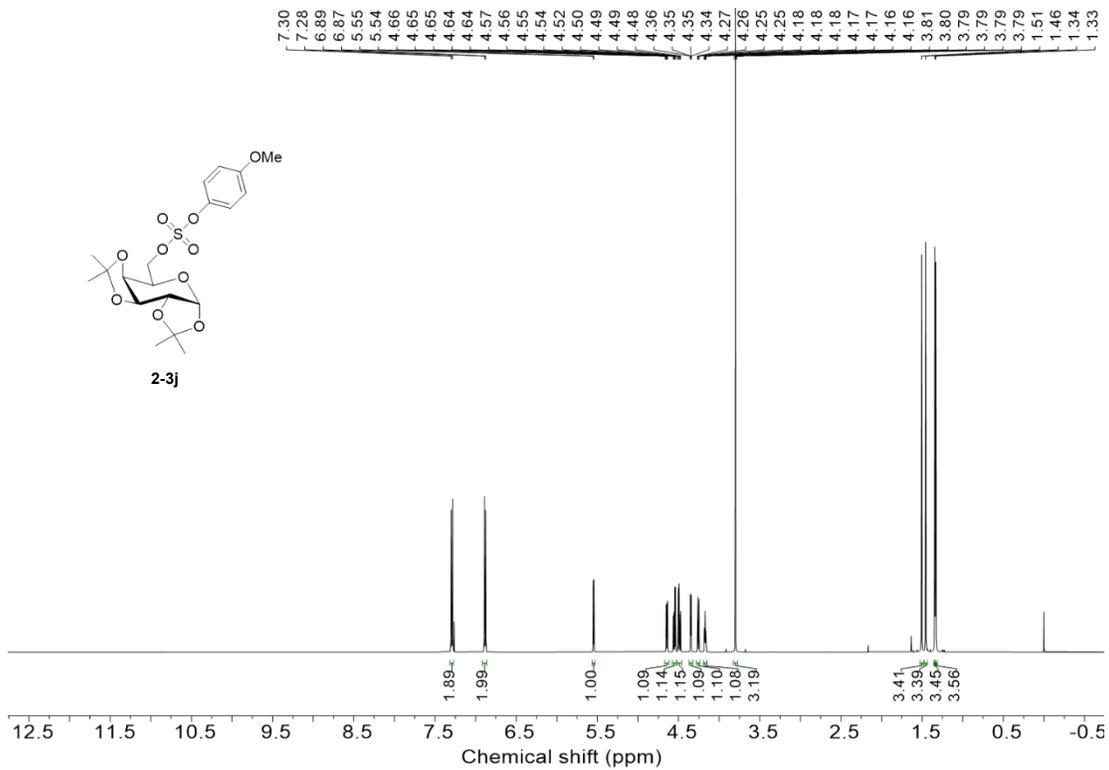
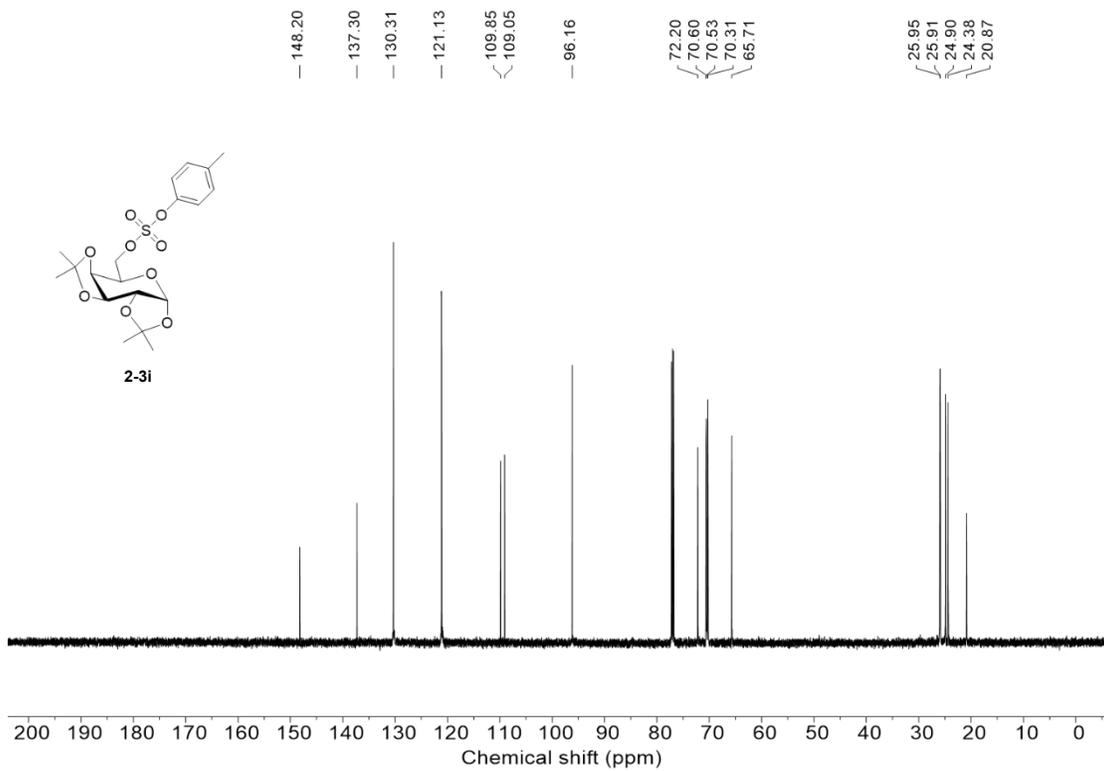


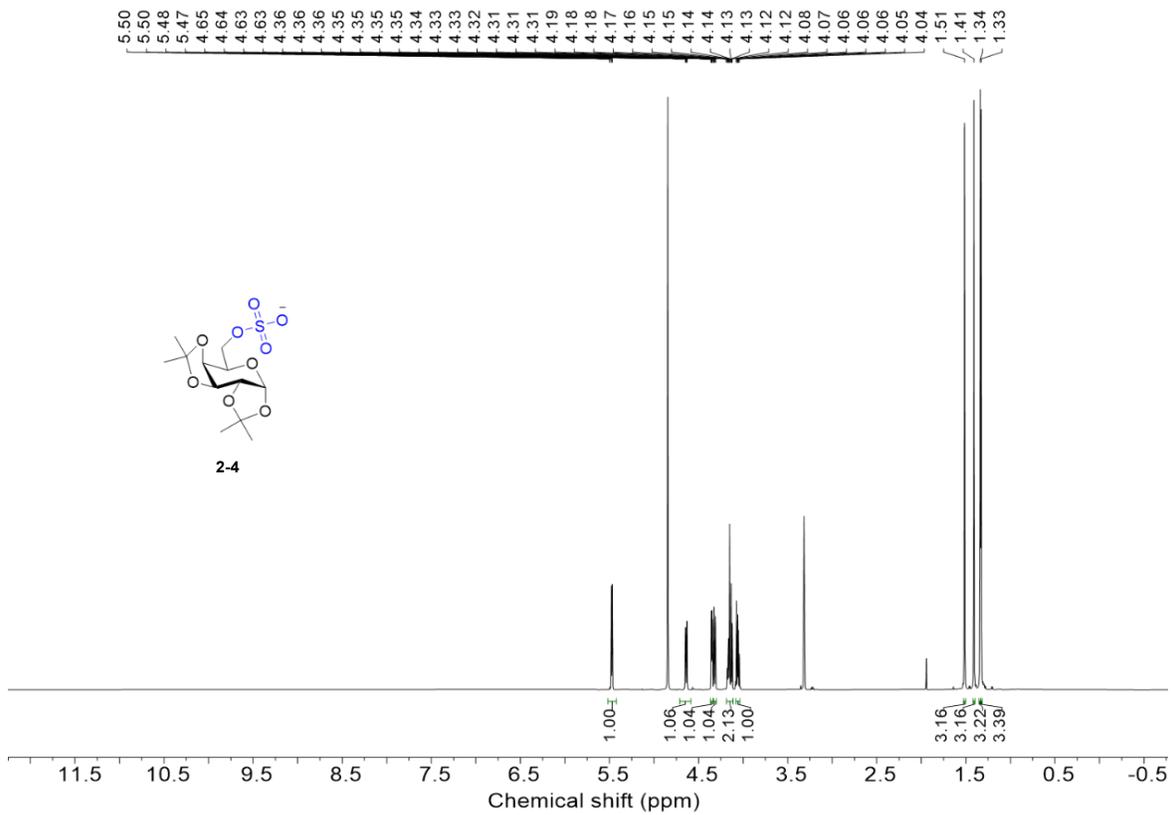
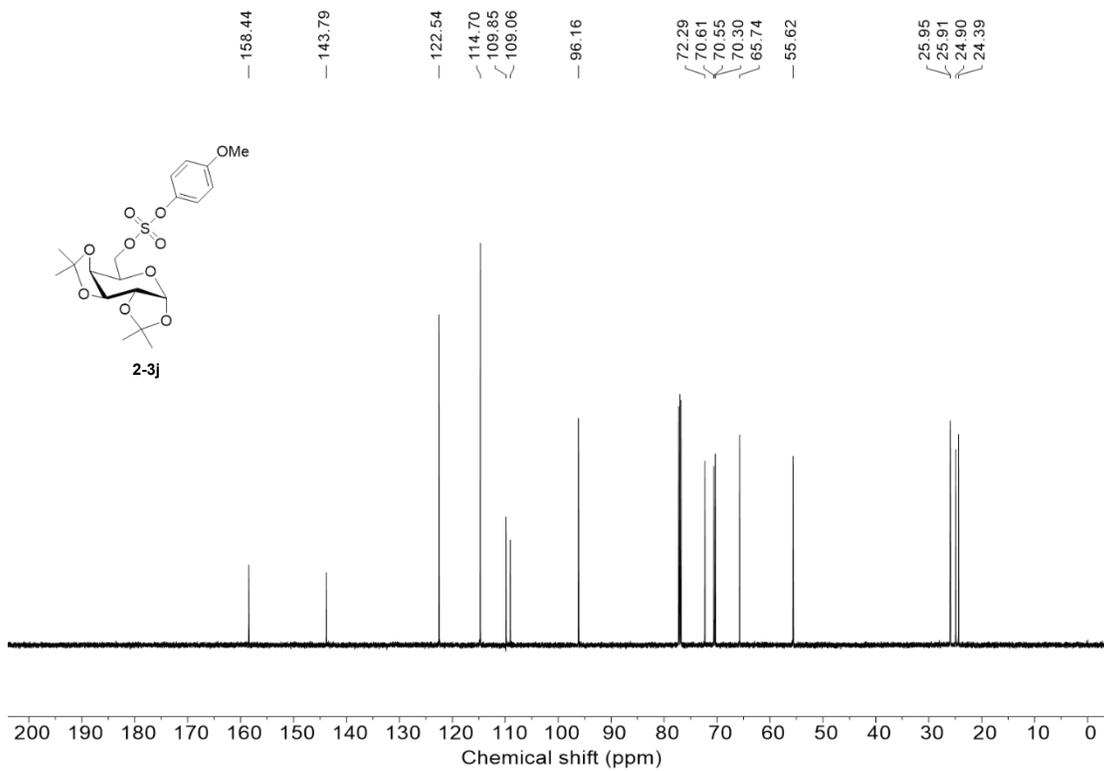


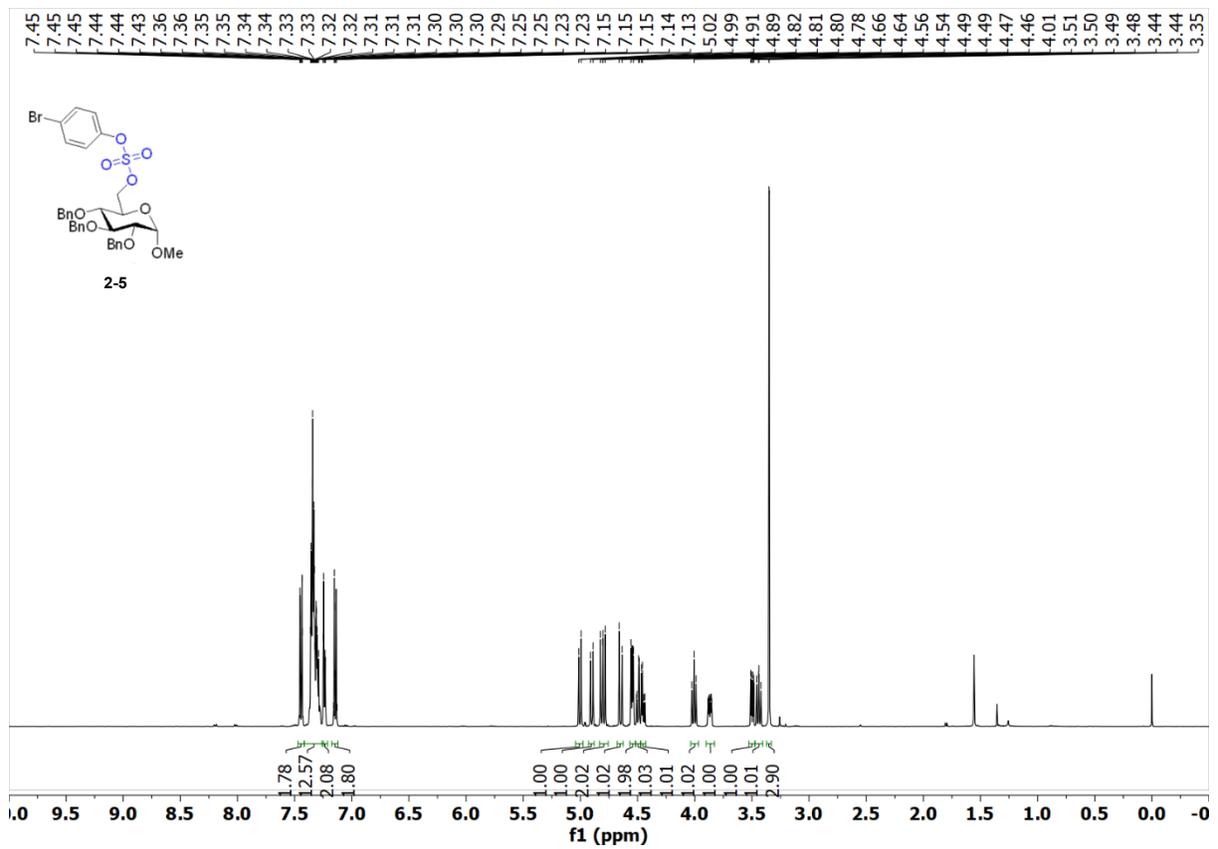
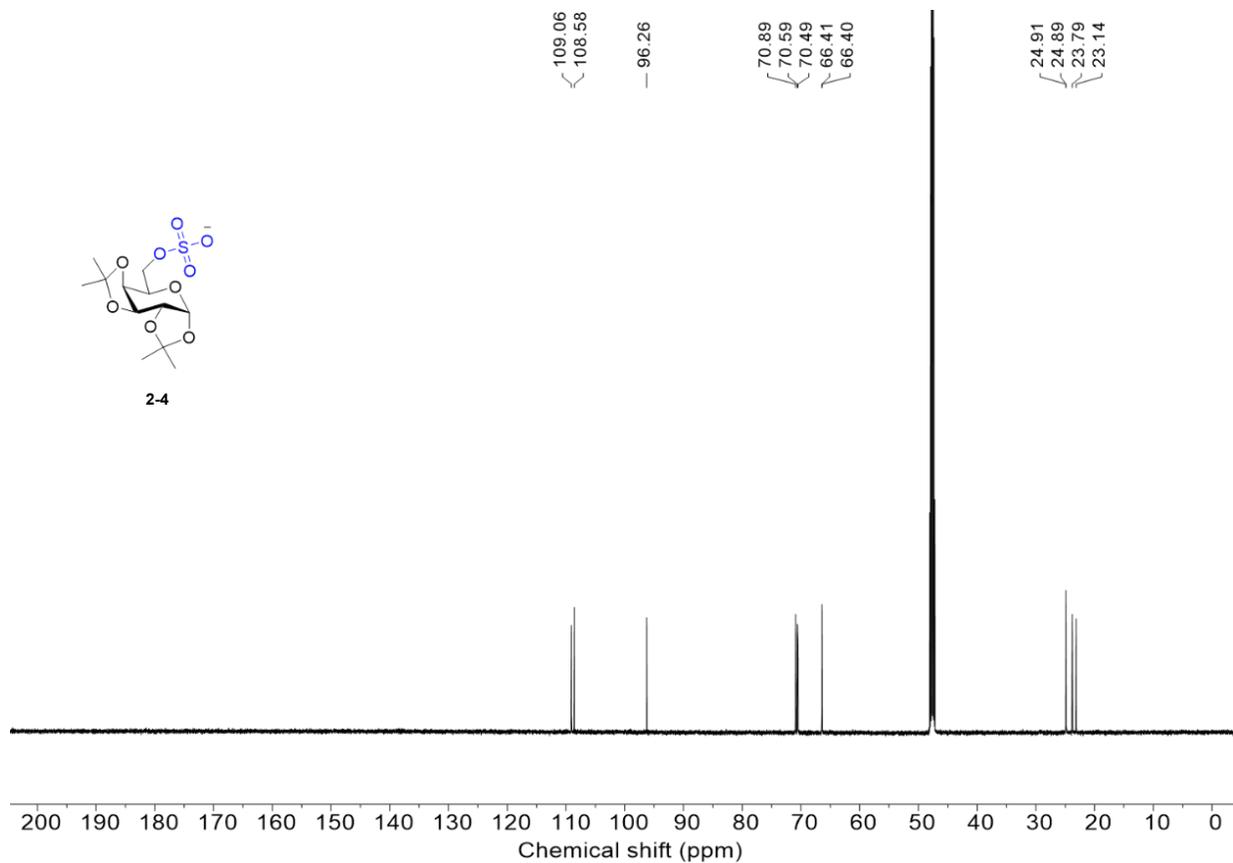


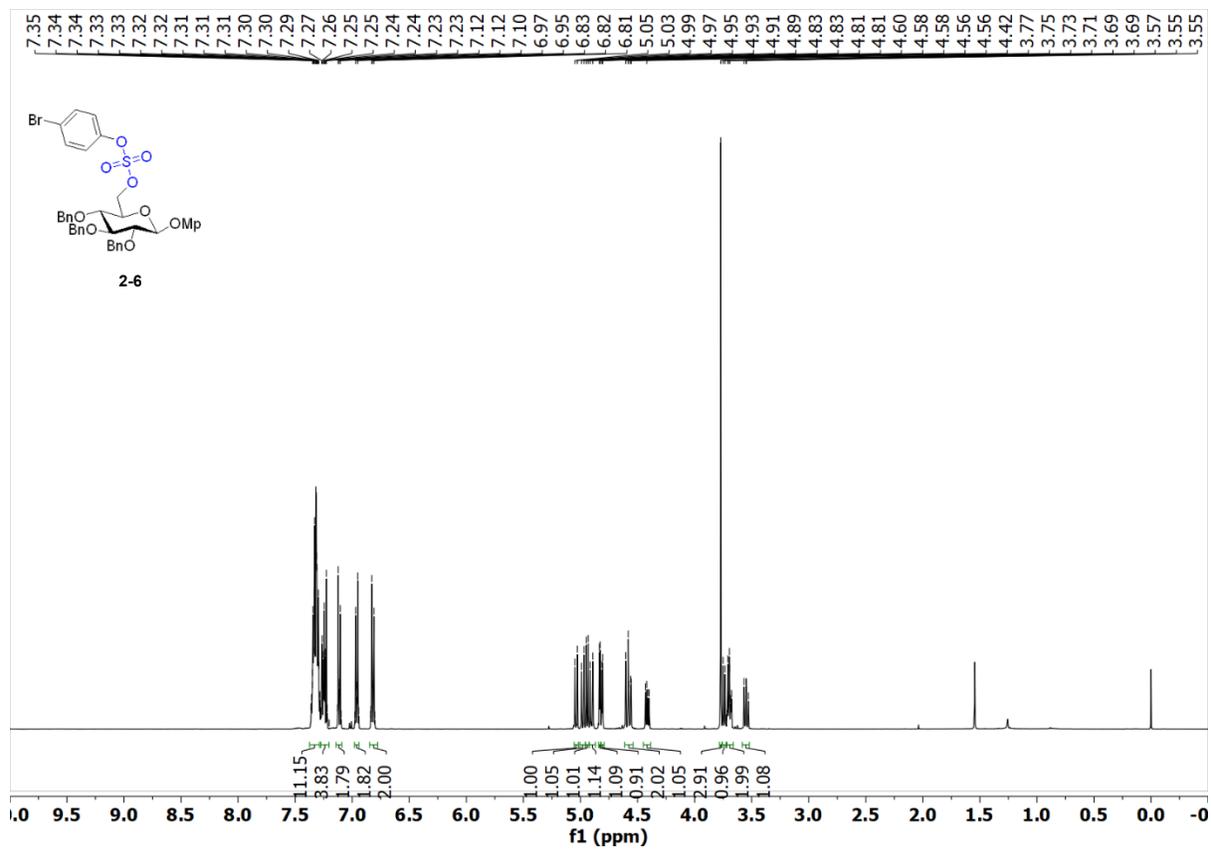
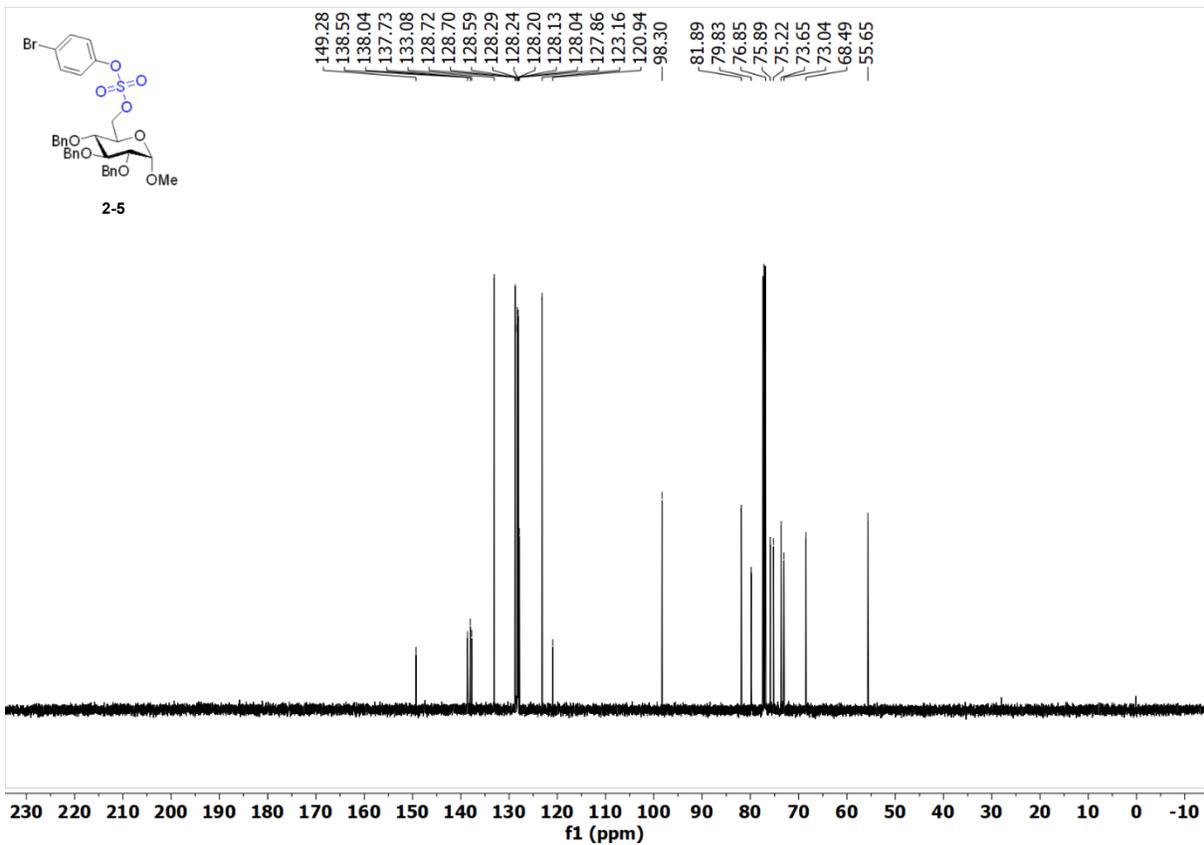


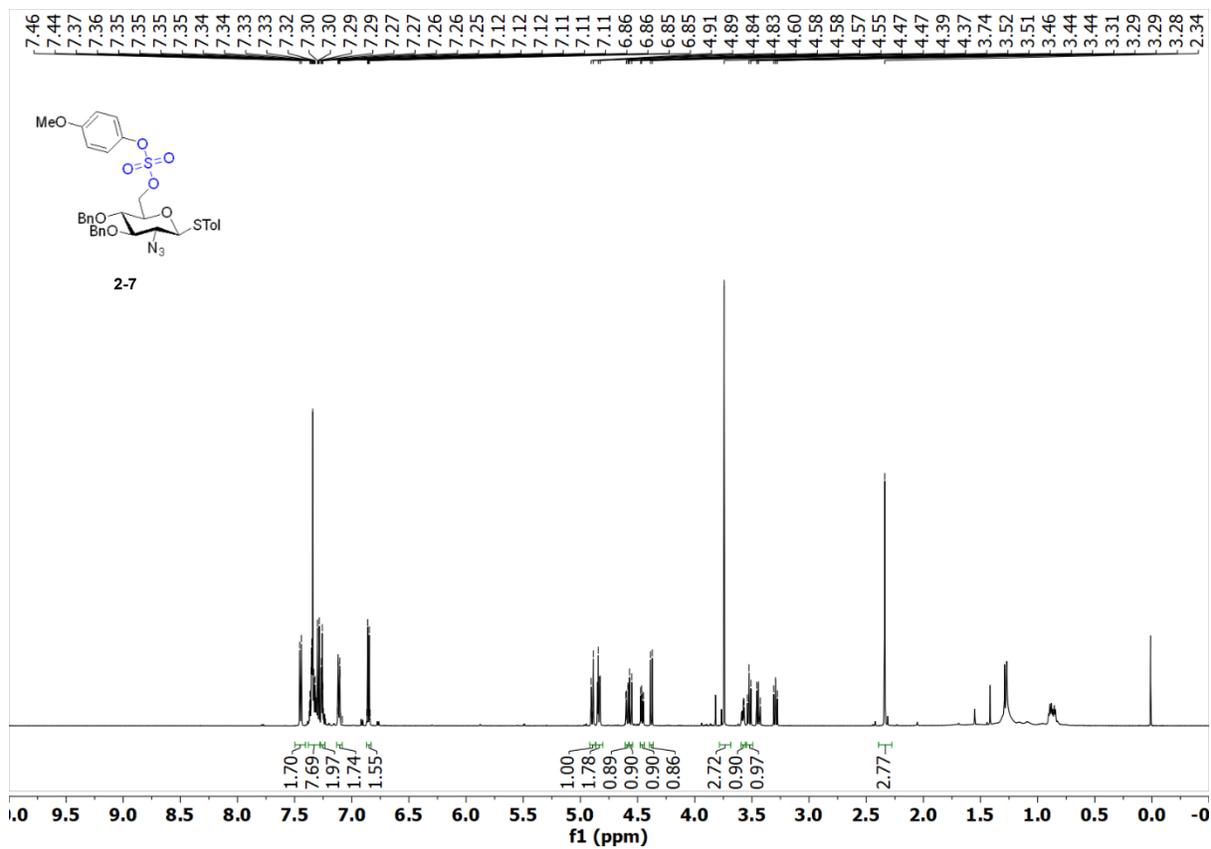
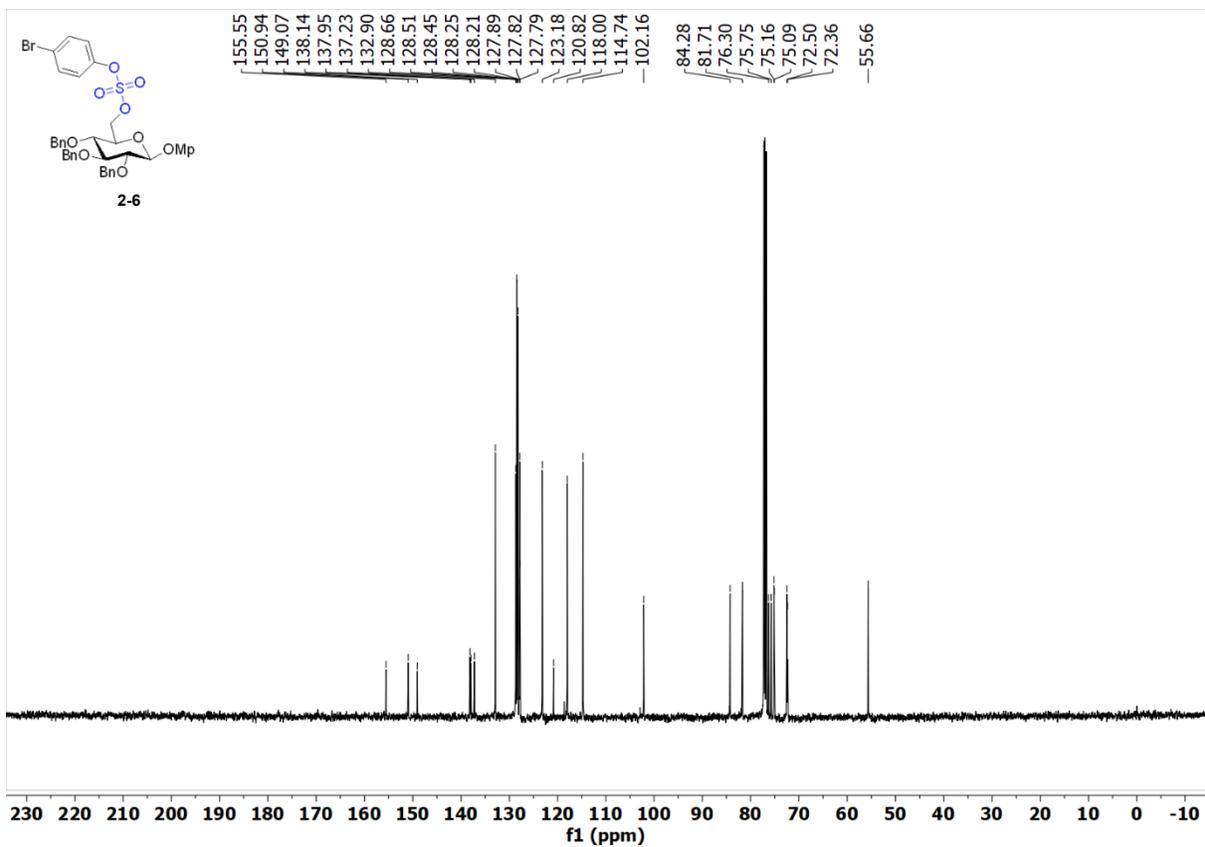


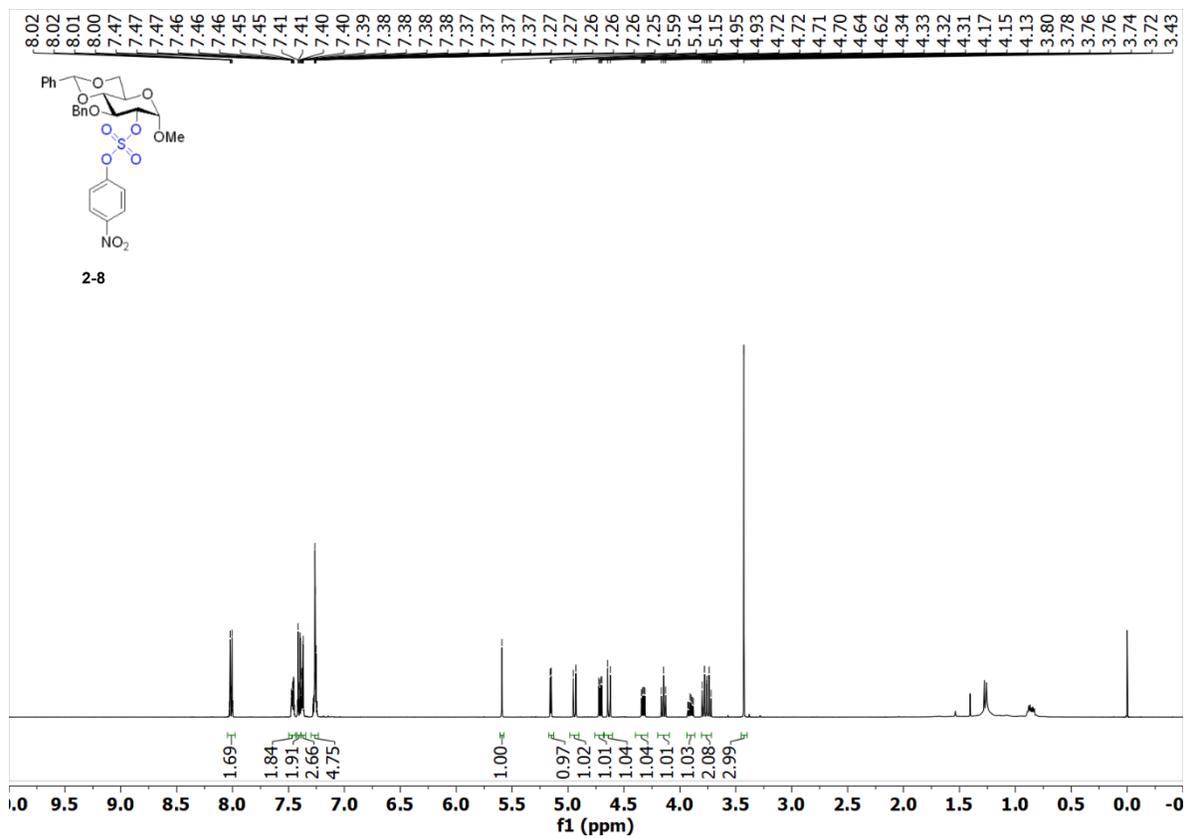
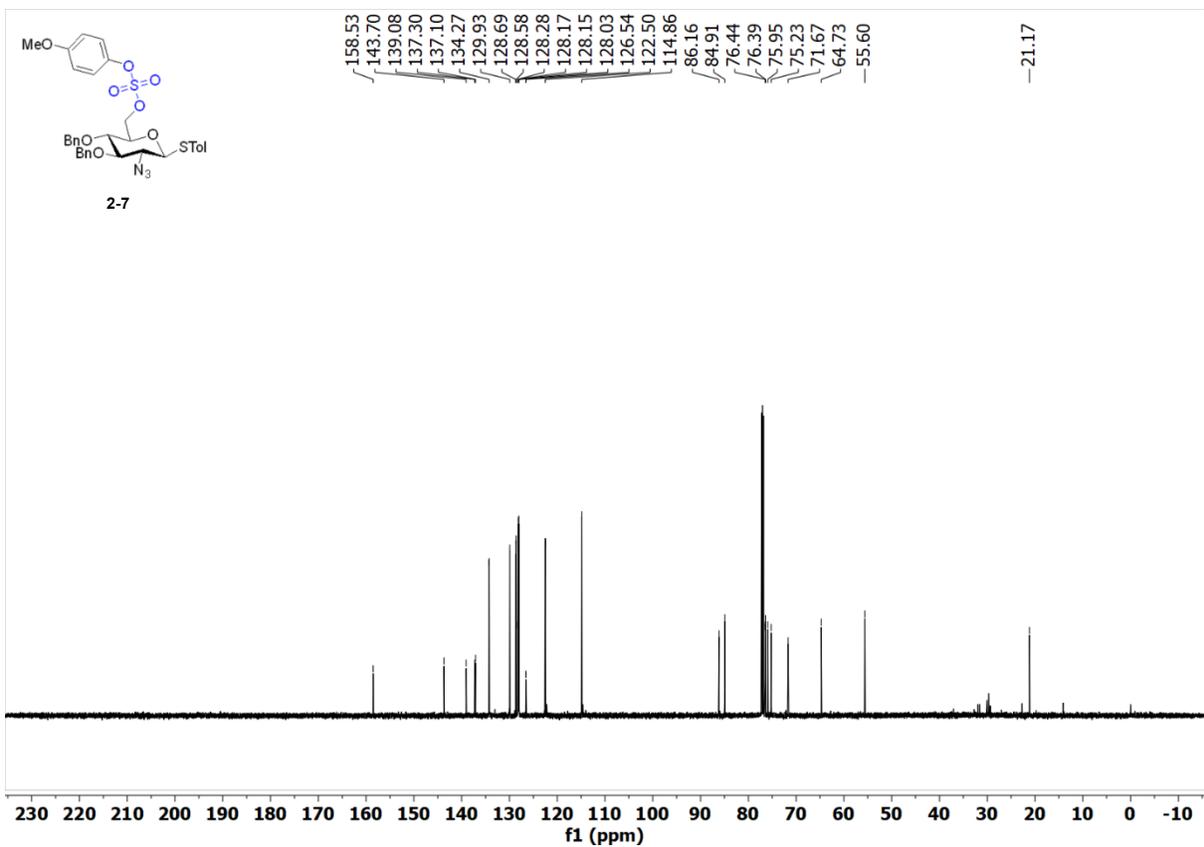


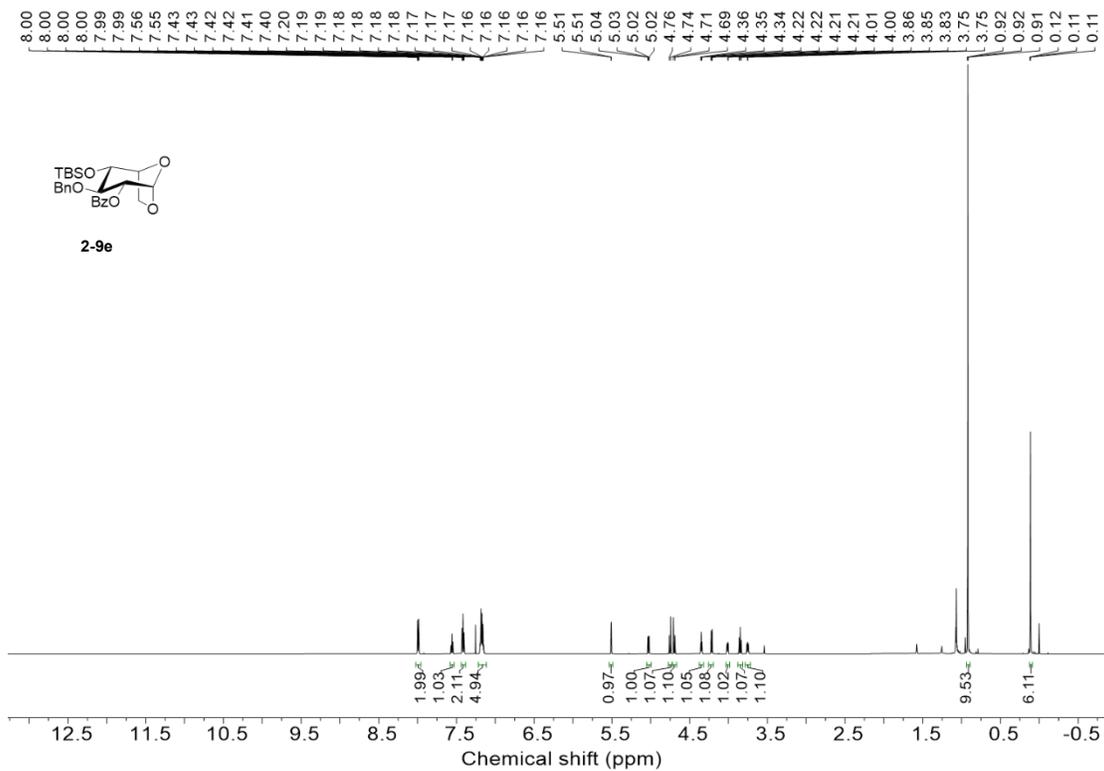
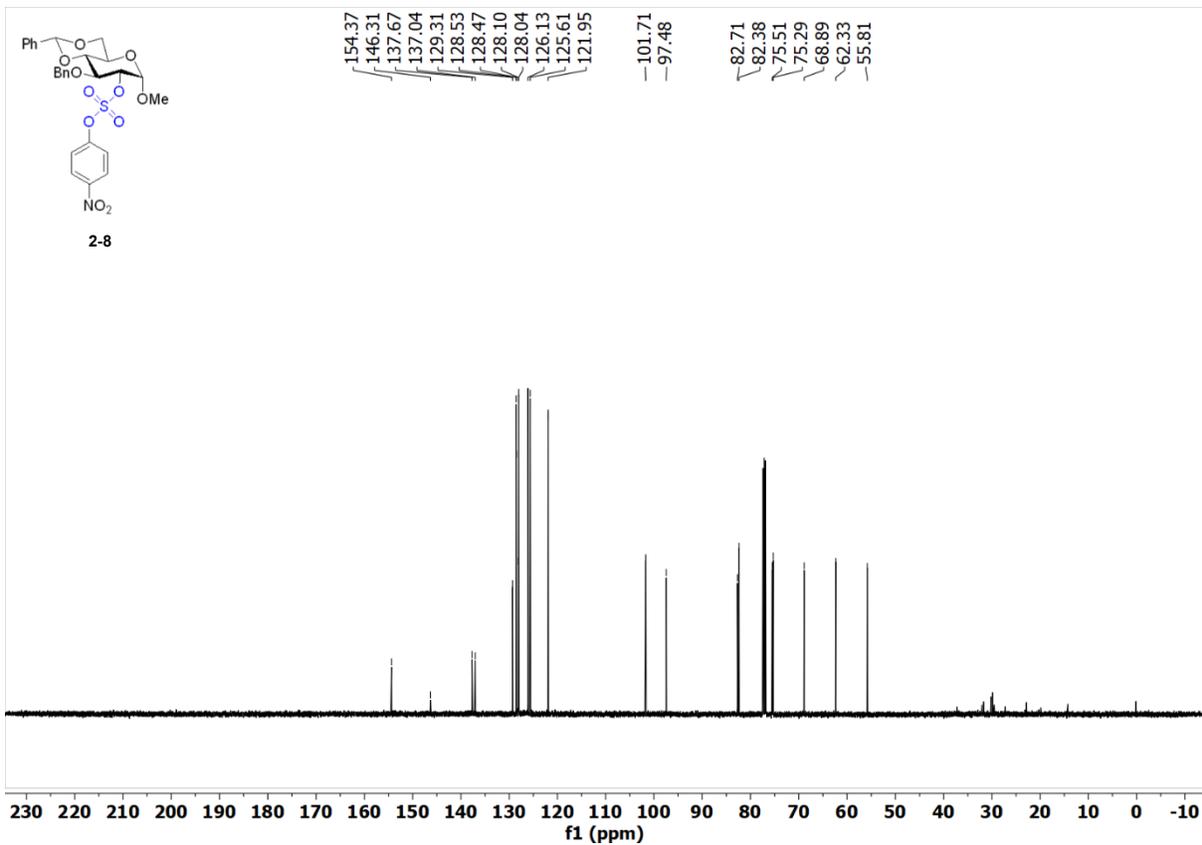


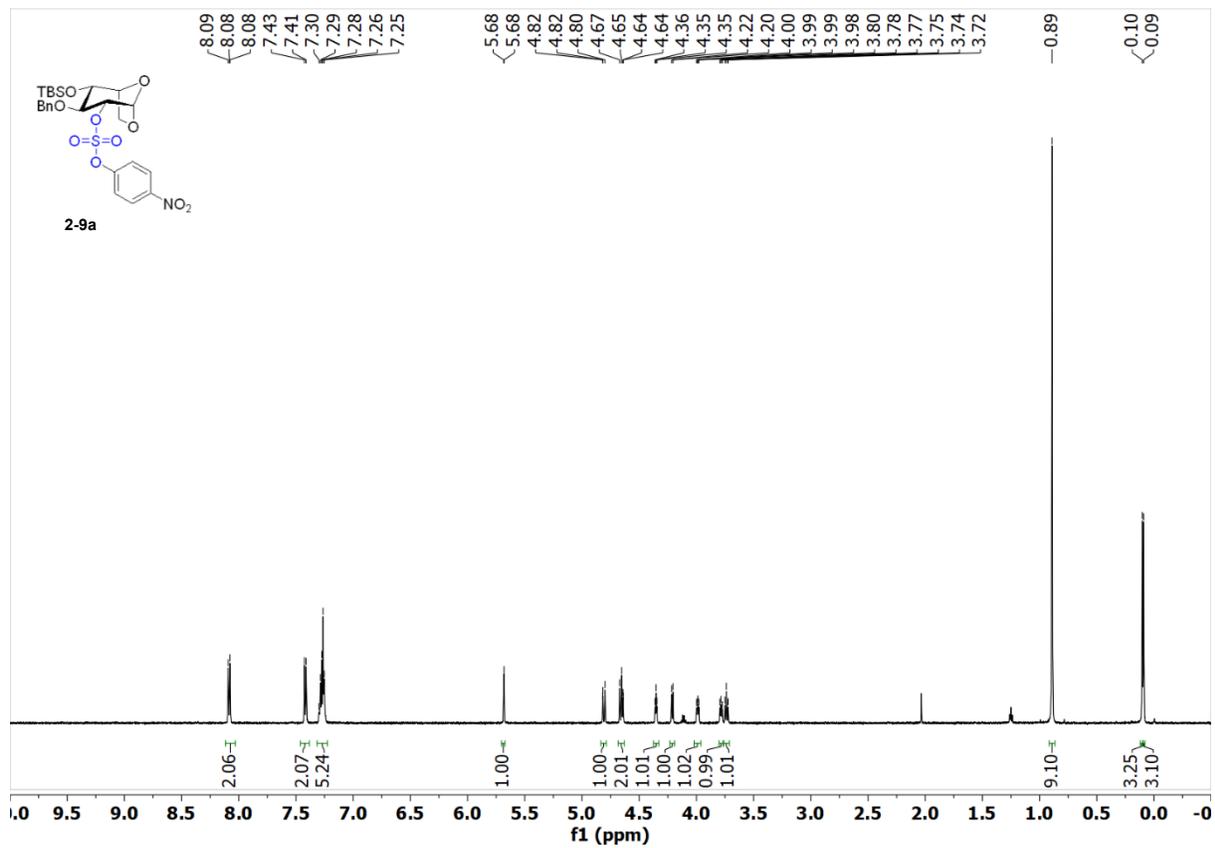
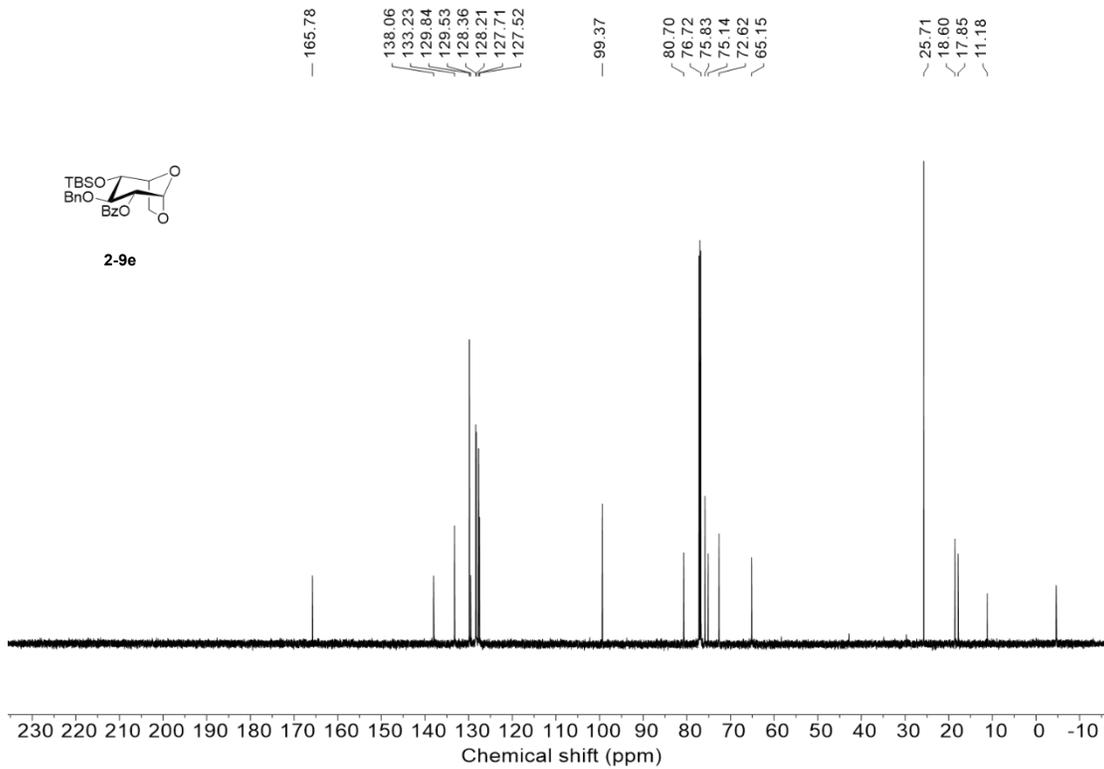


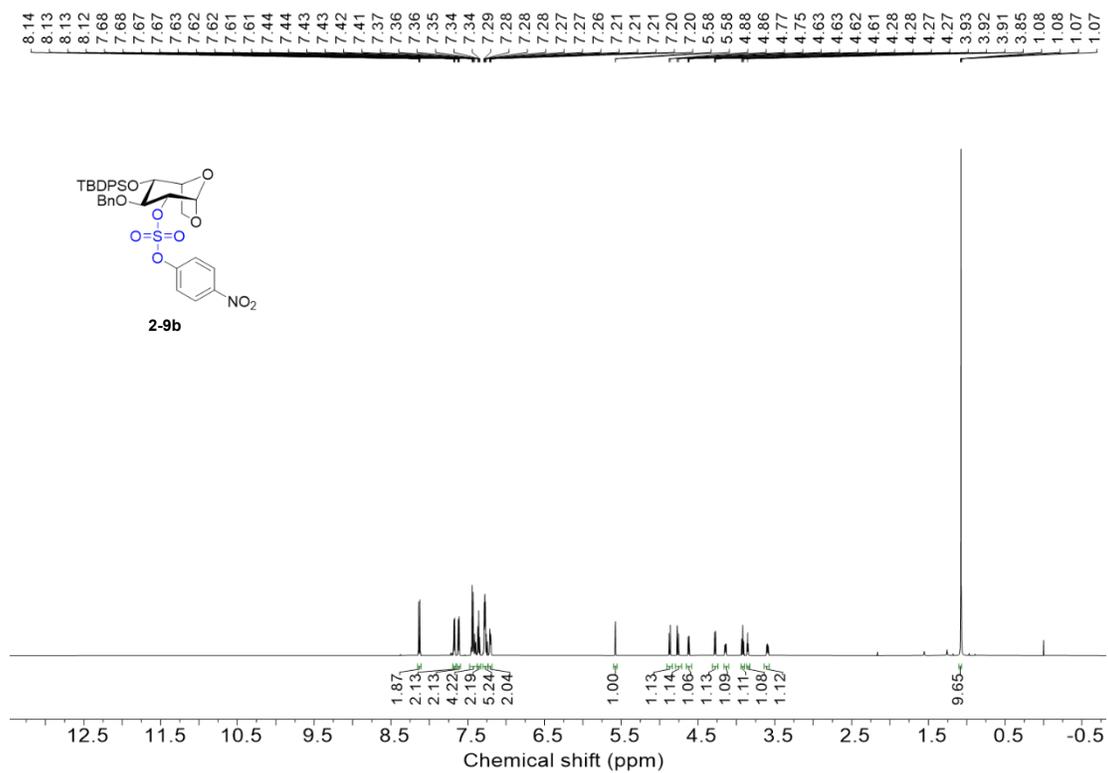
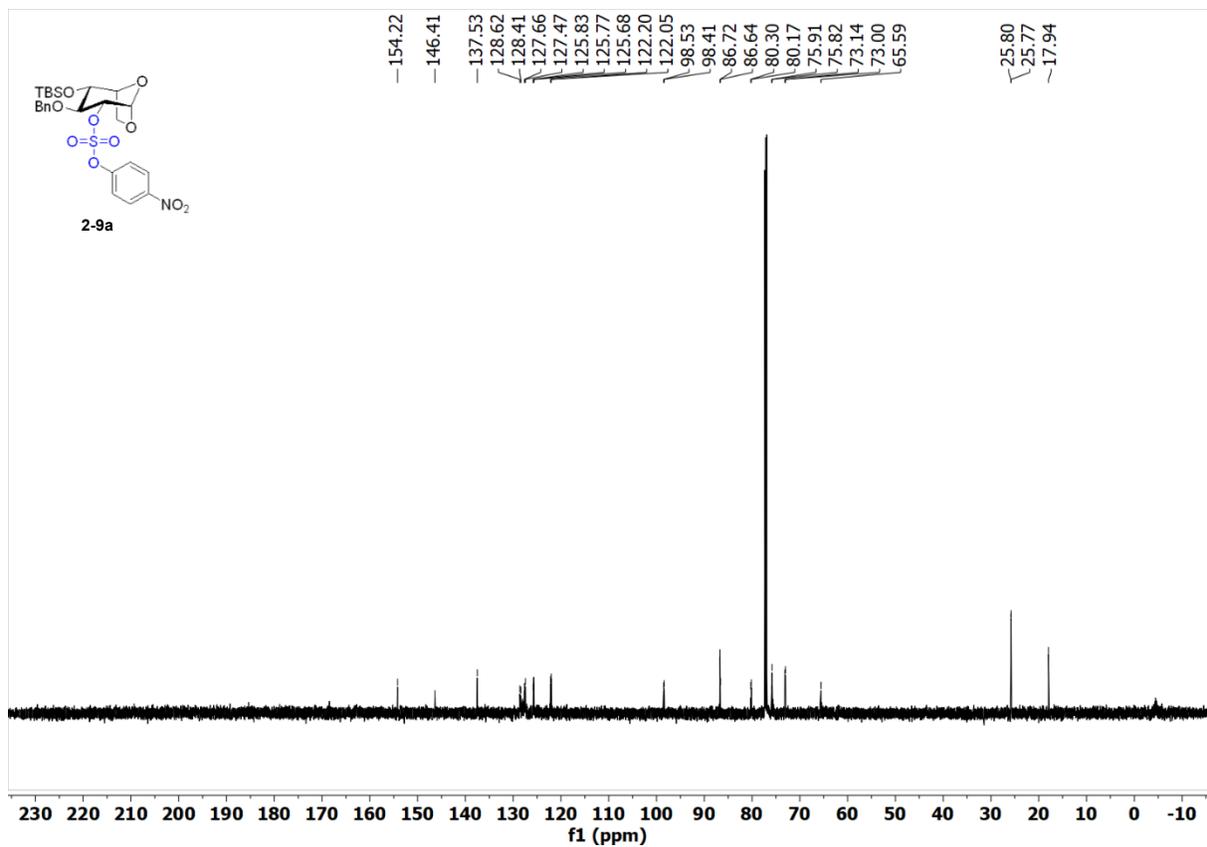


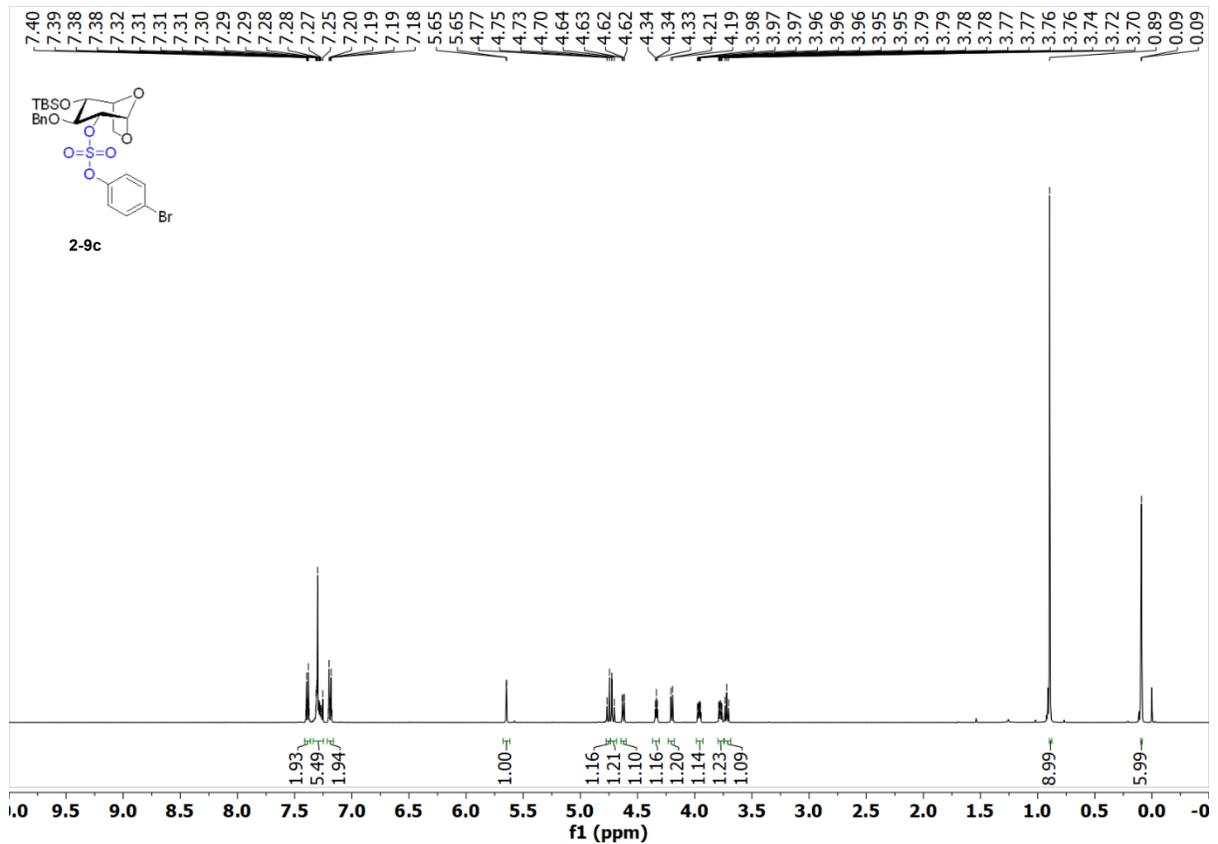
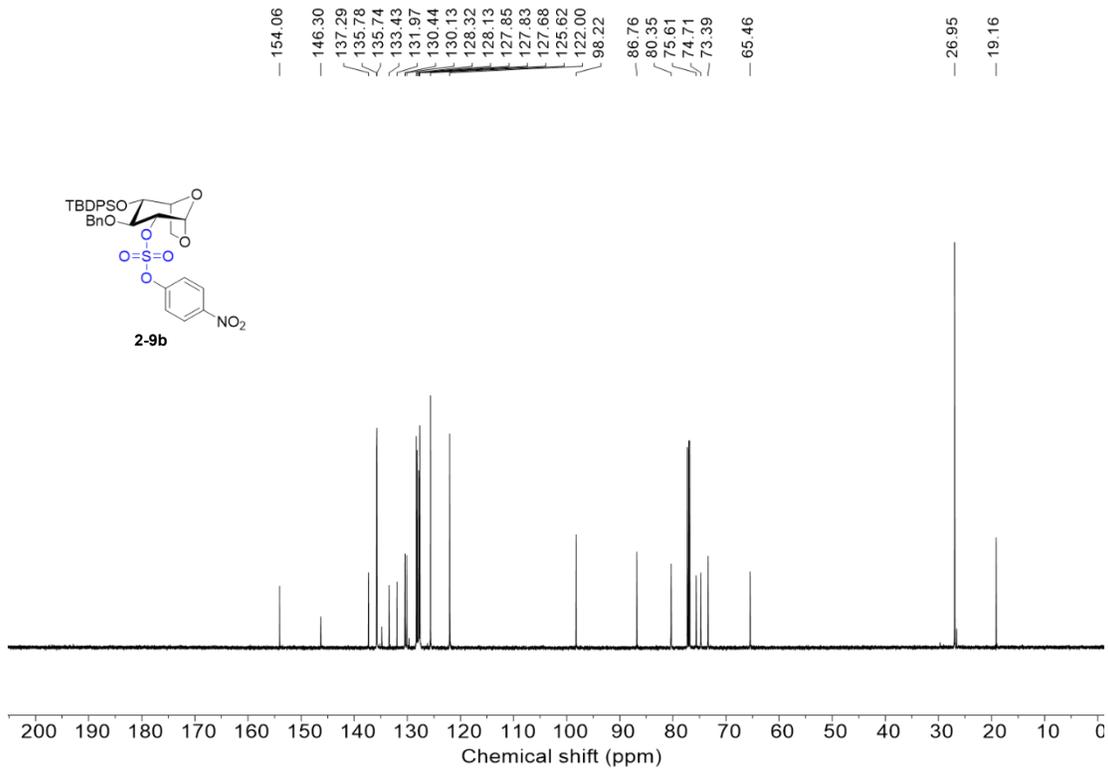


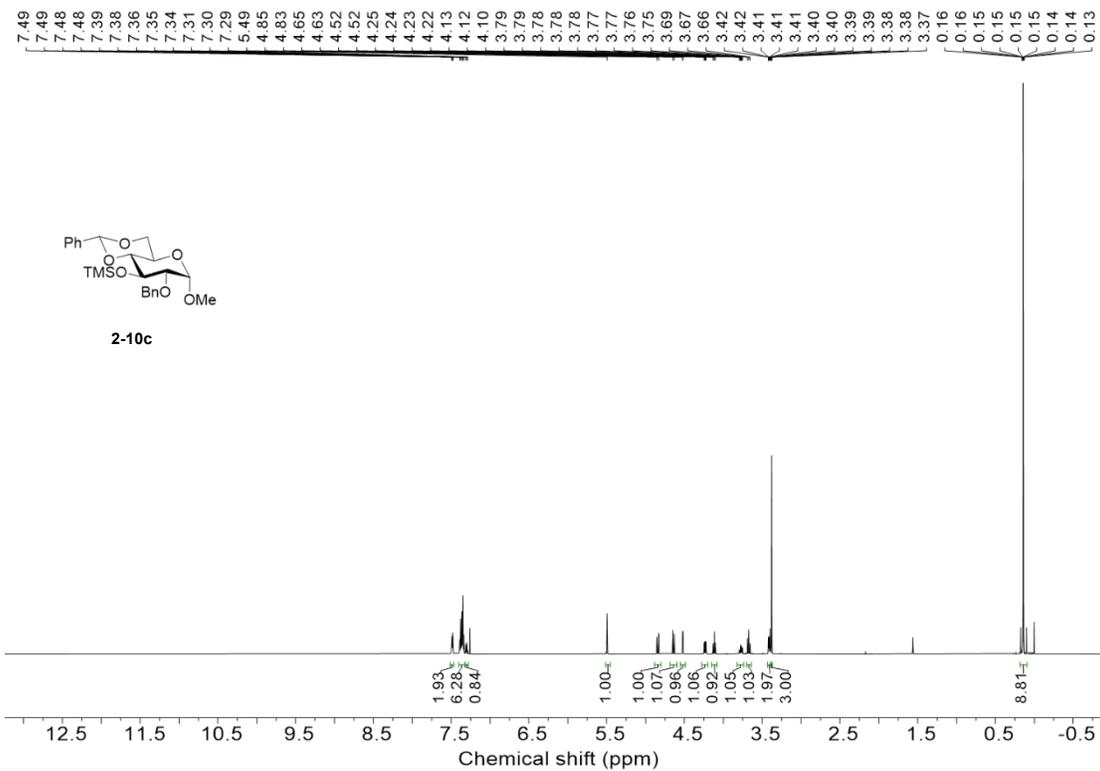
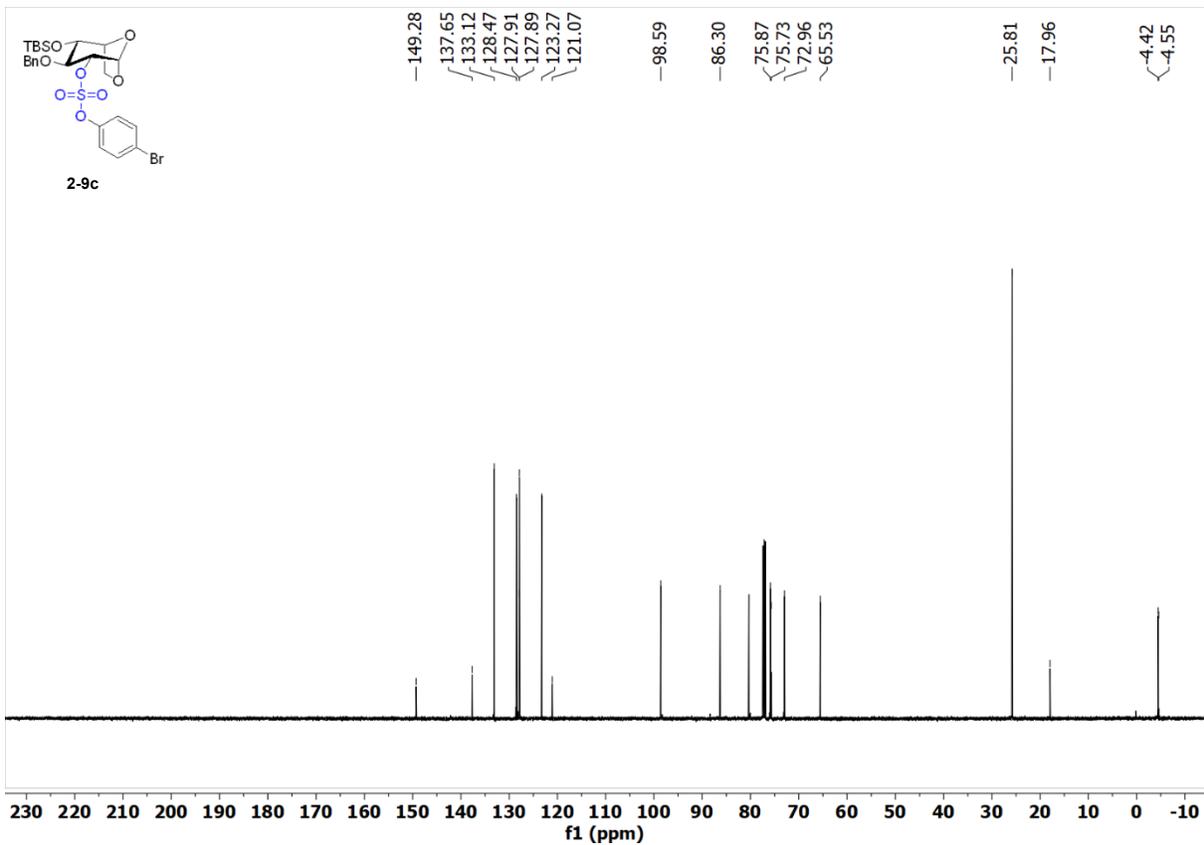


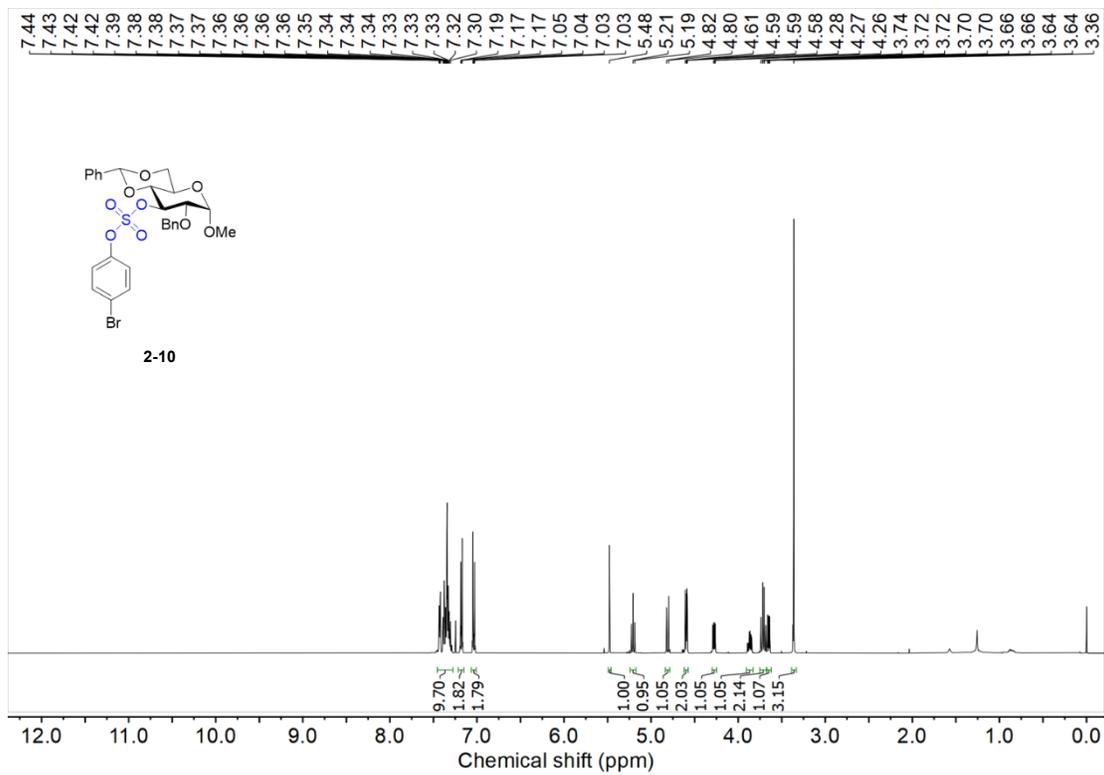
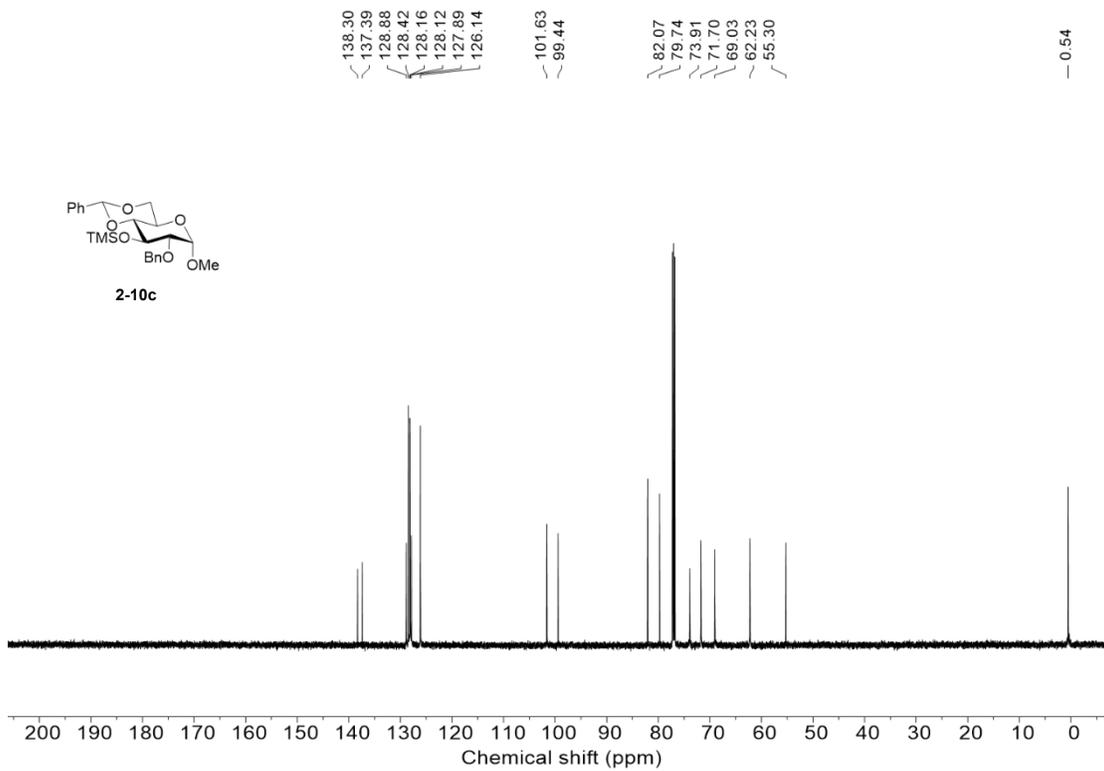


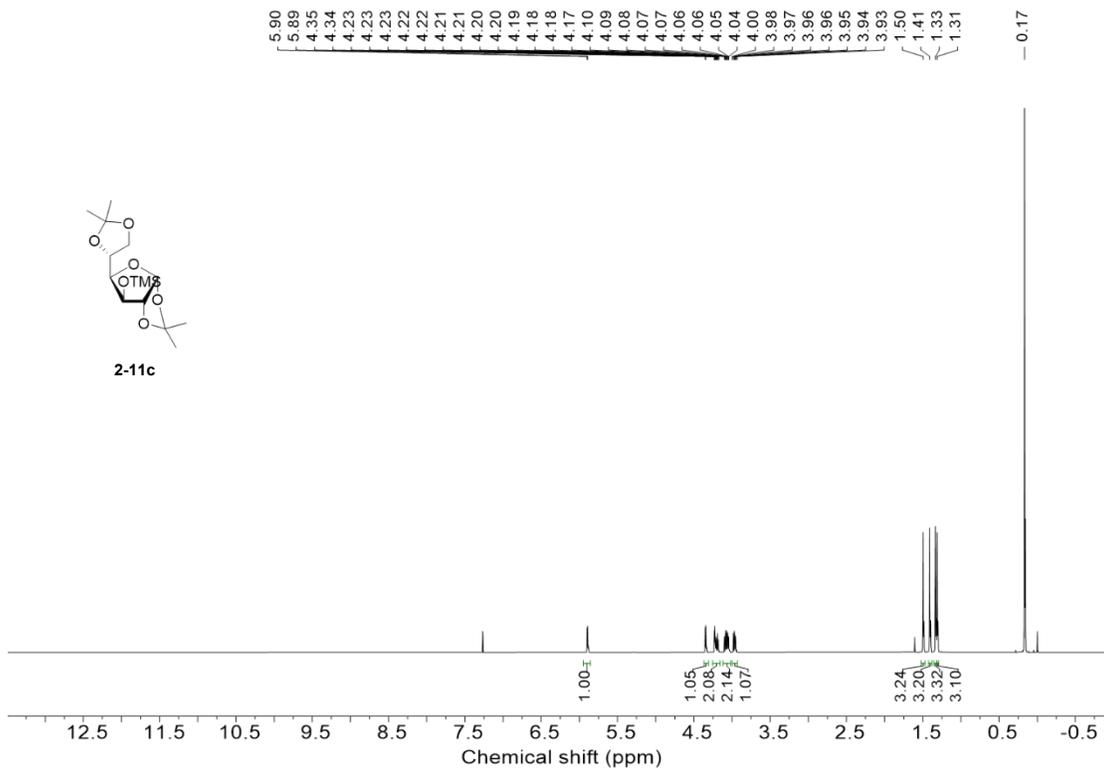
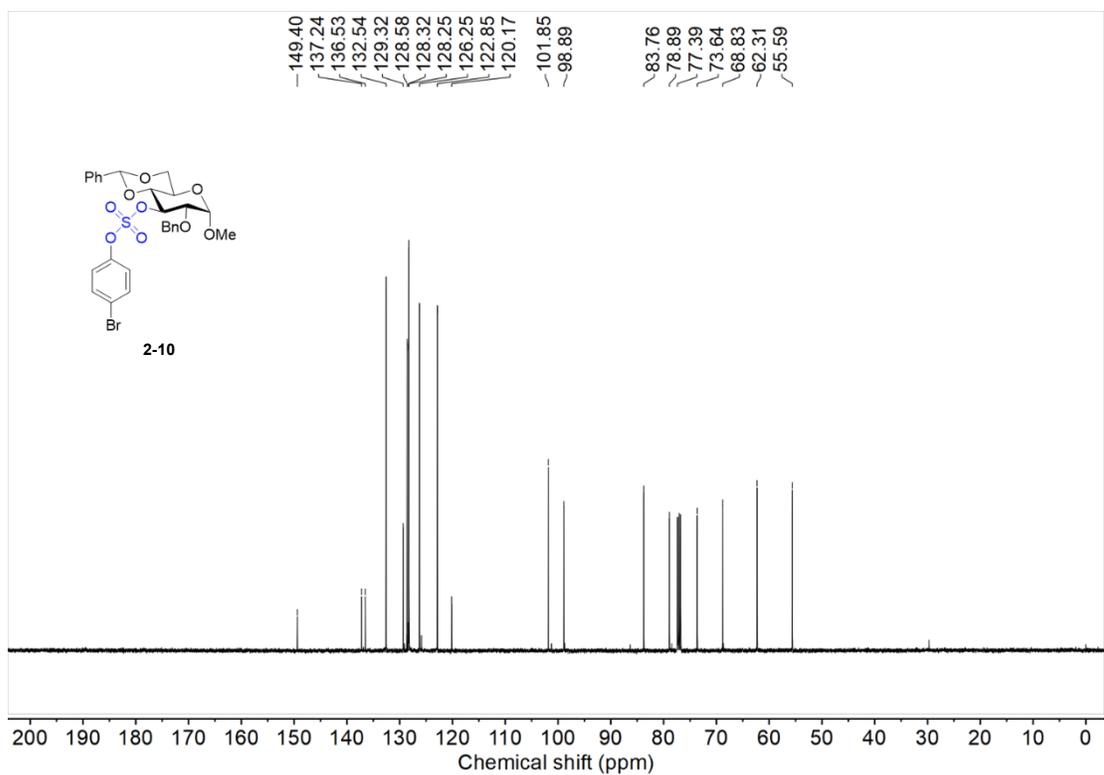


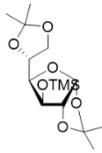




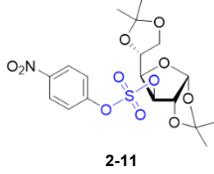
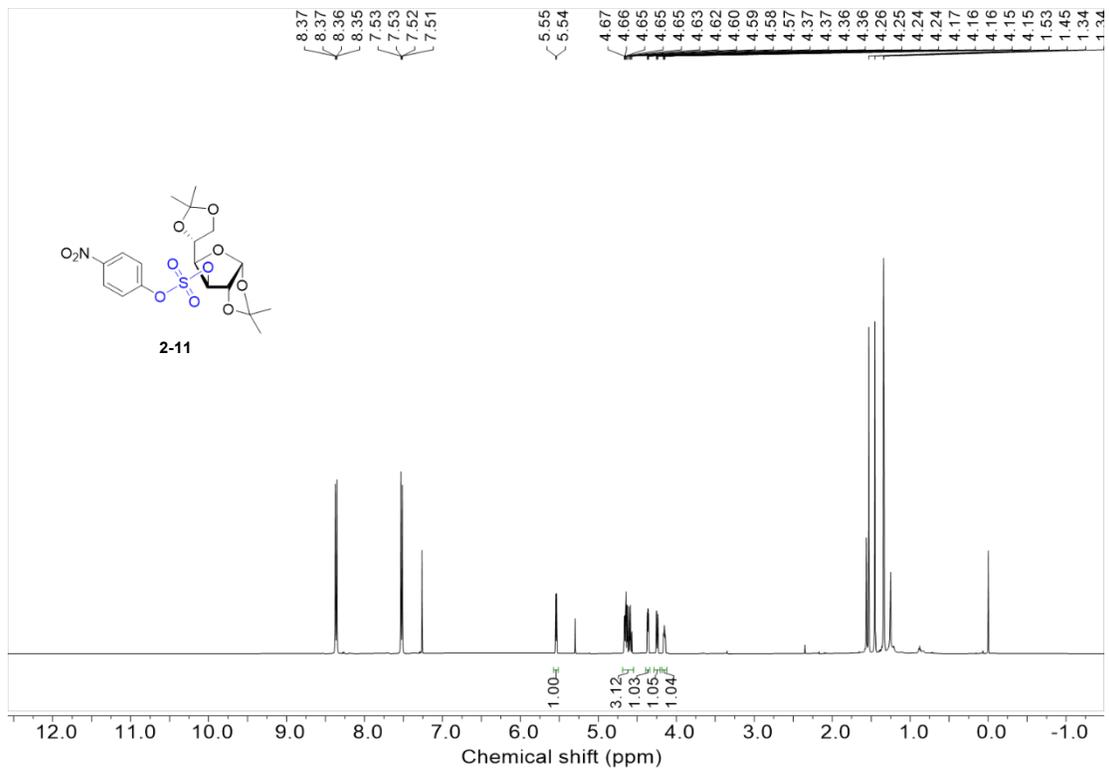
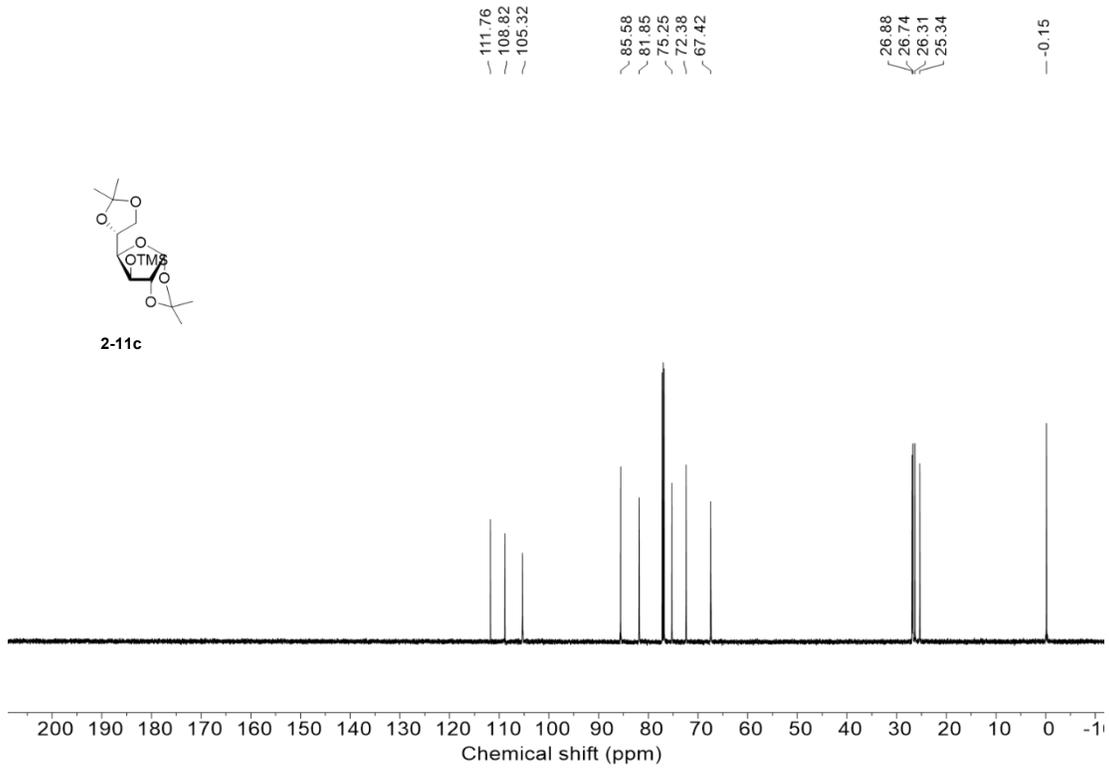




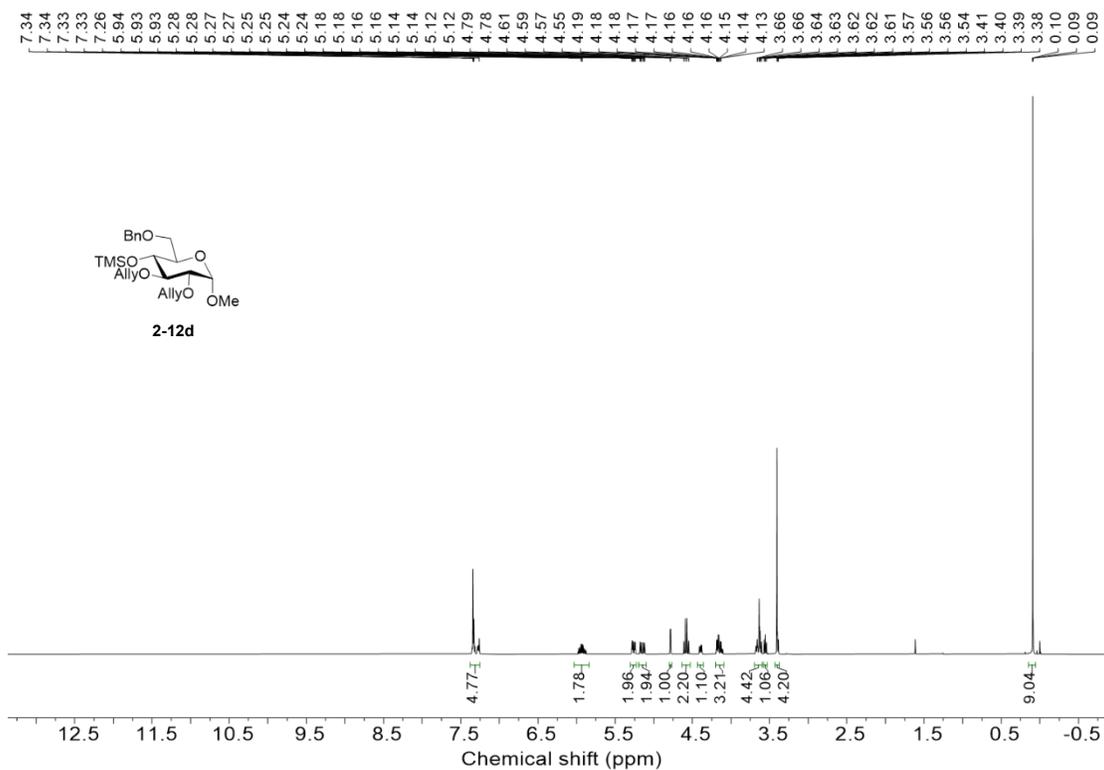
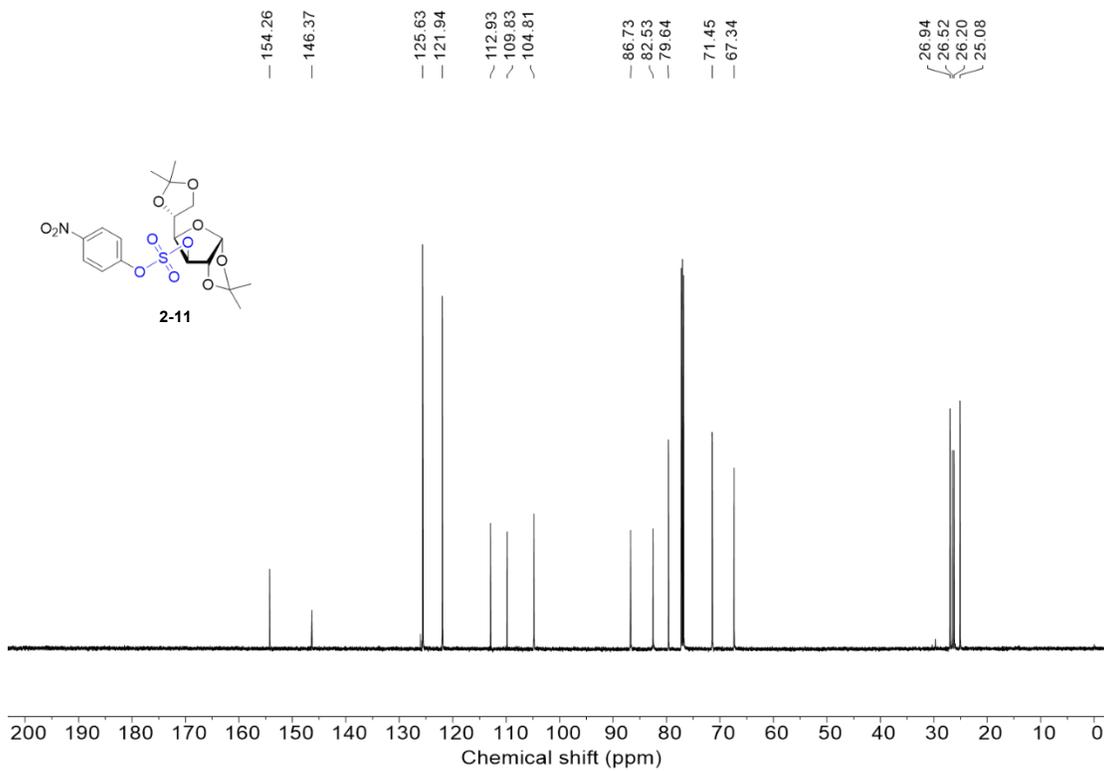


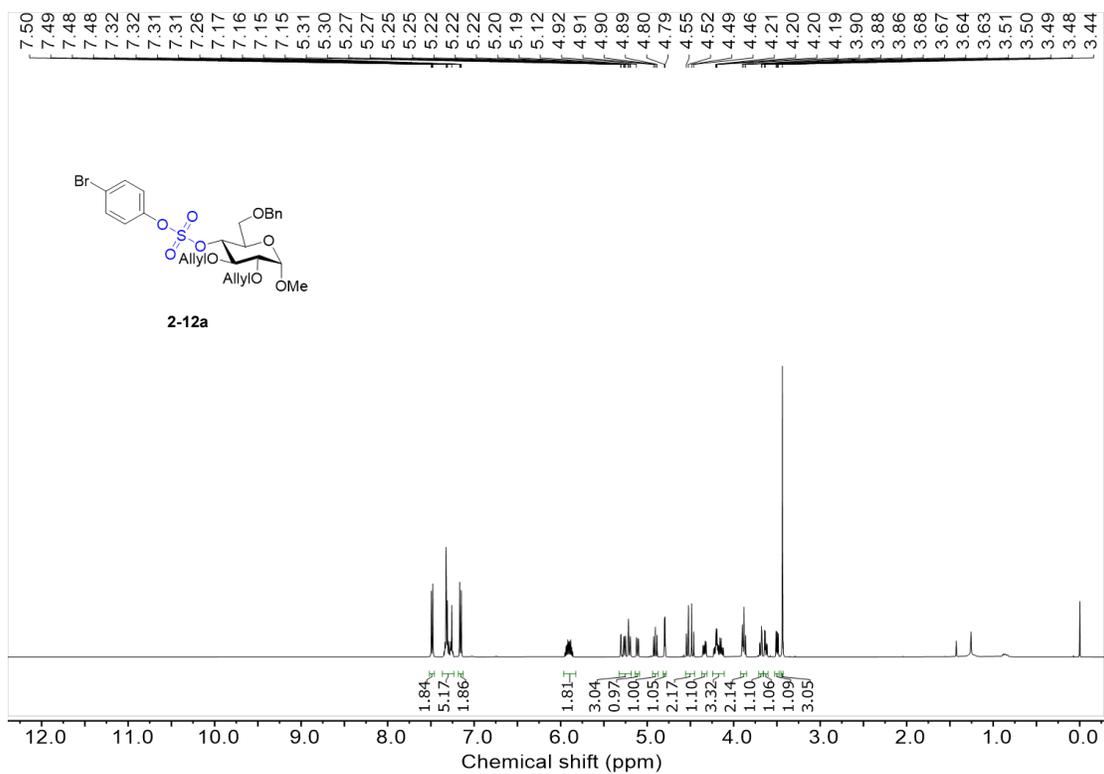
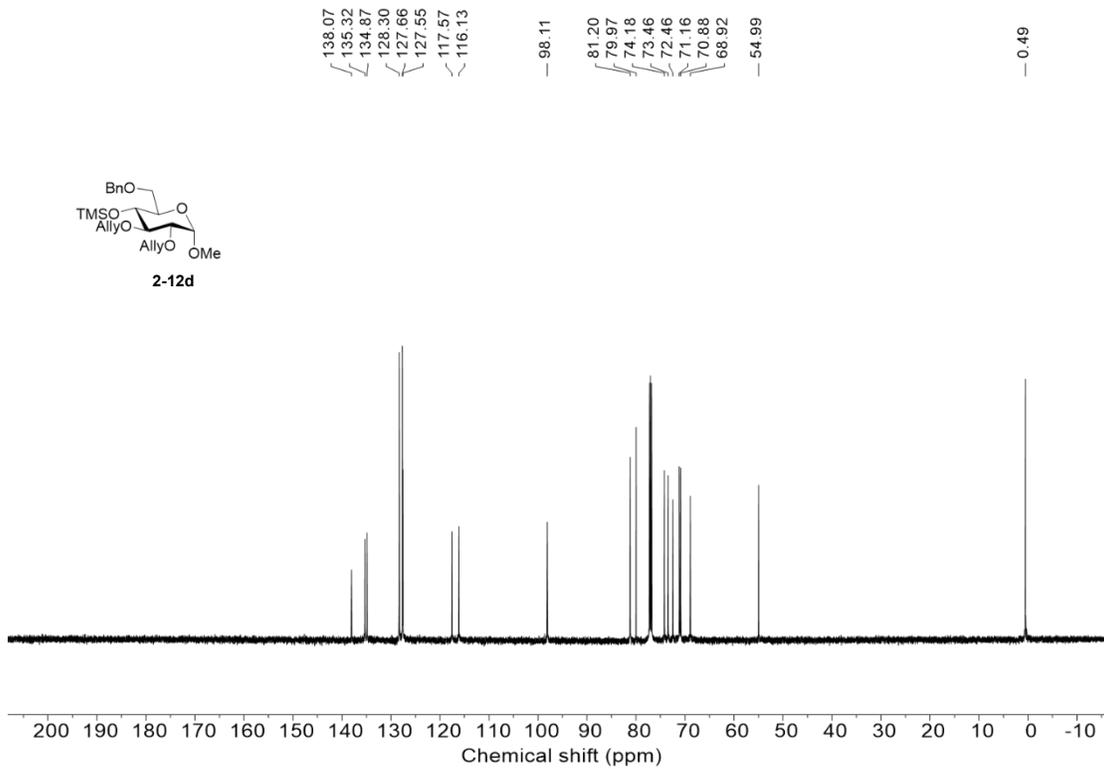


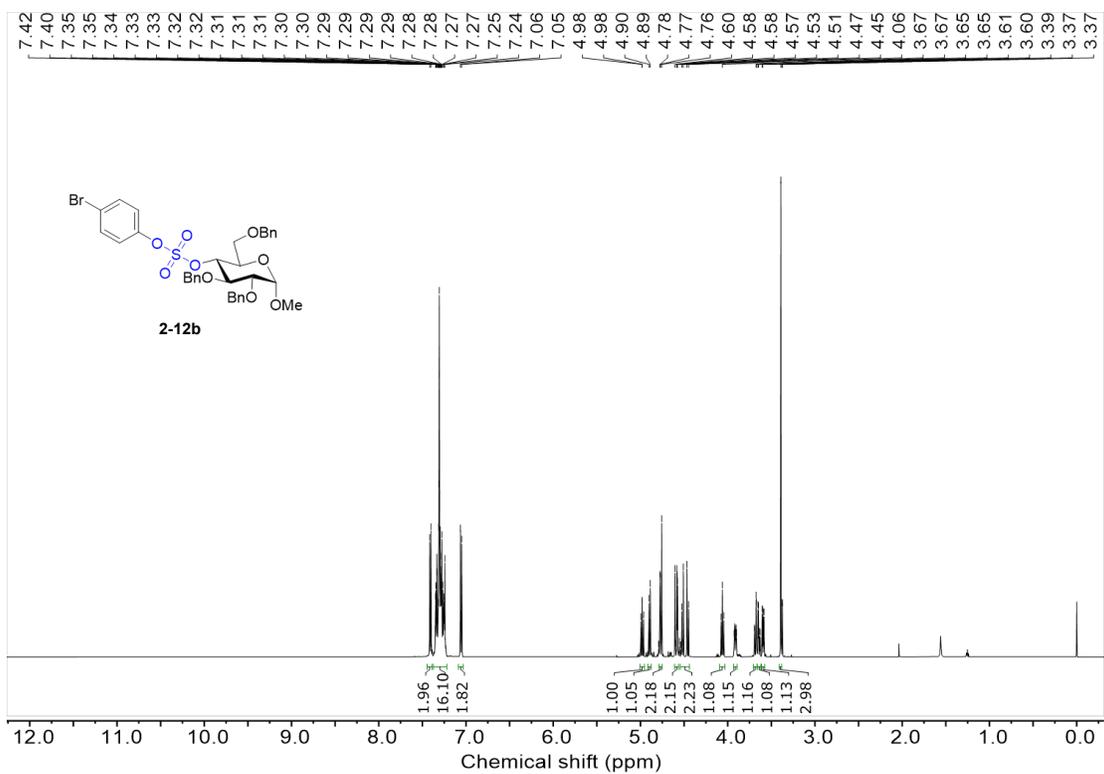
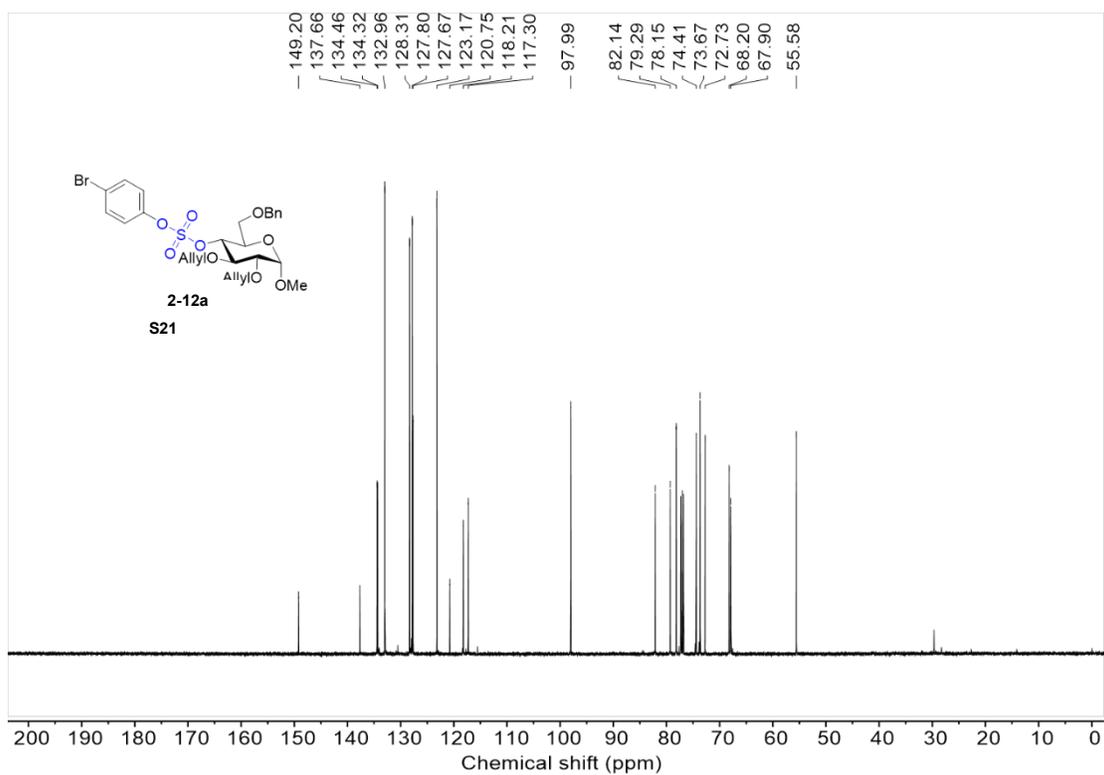
2-11c

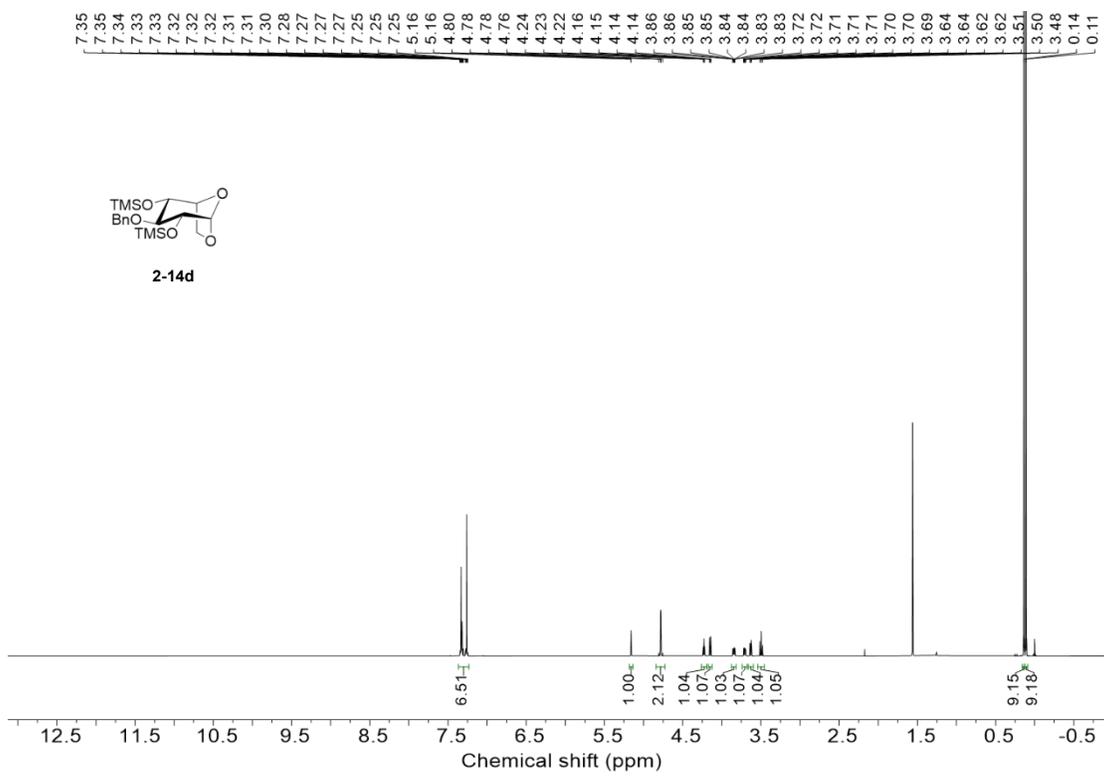
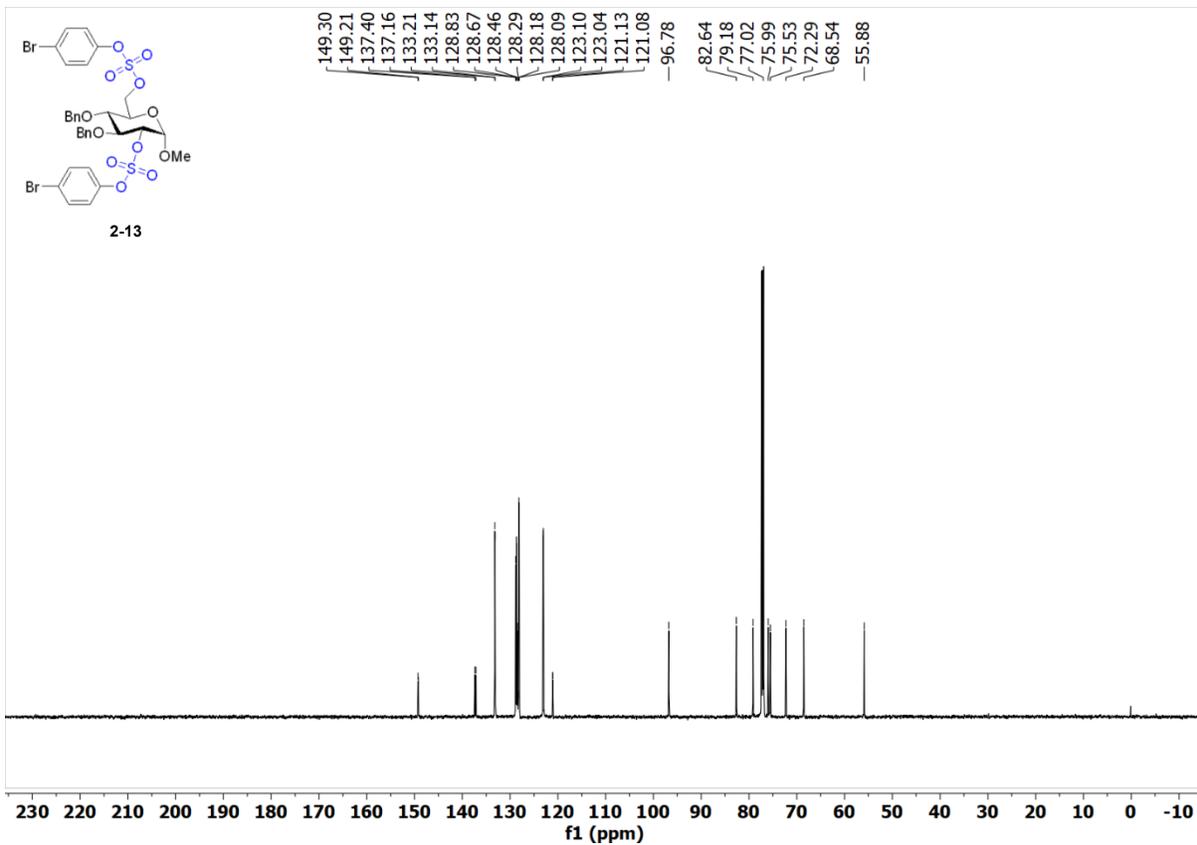


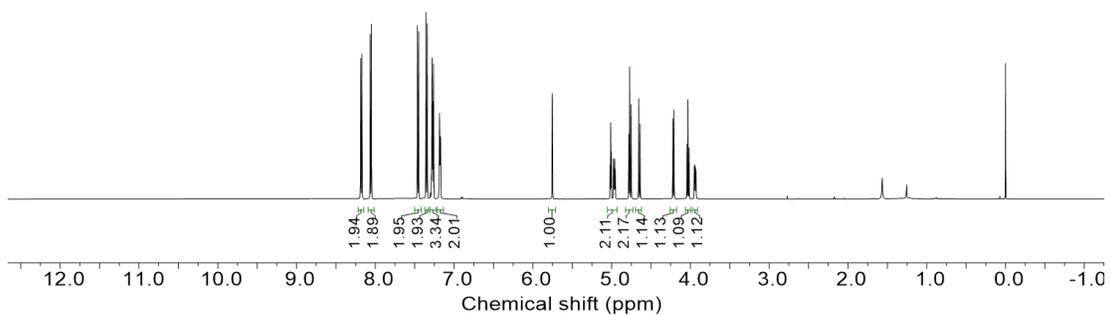
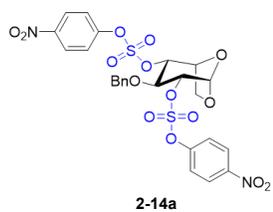
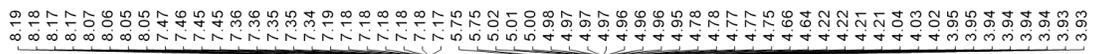
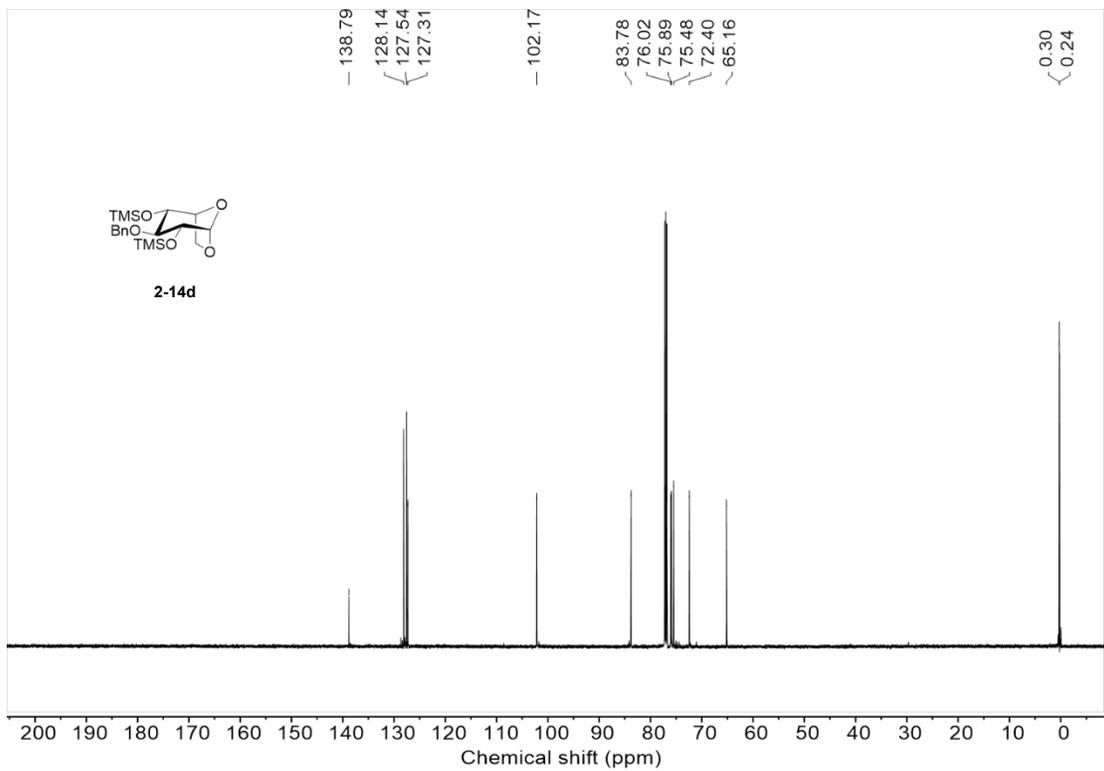
2-11

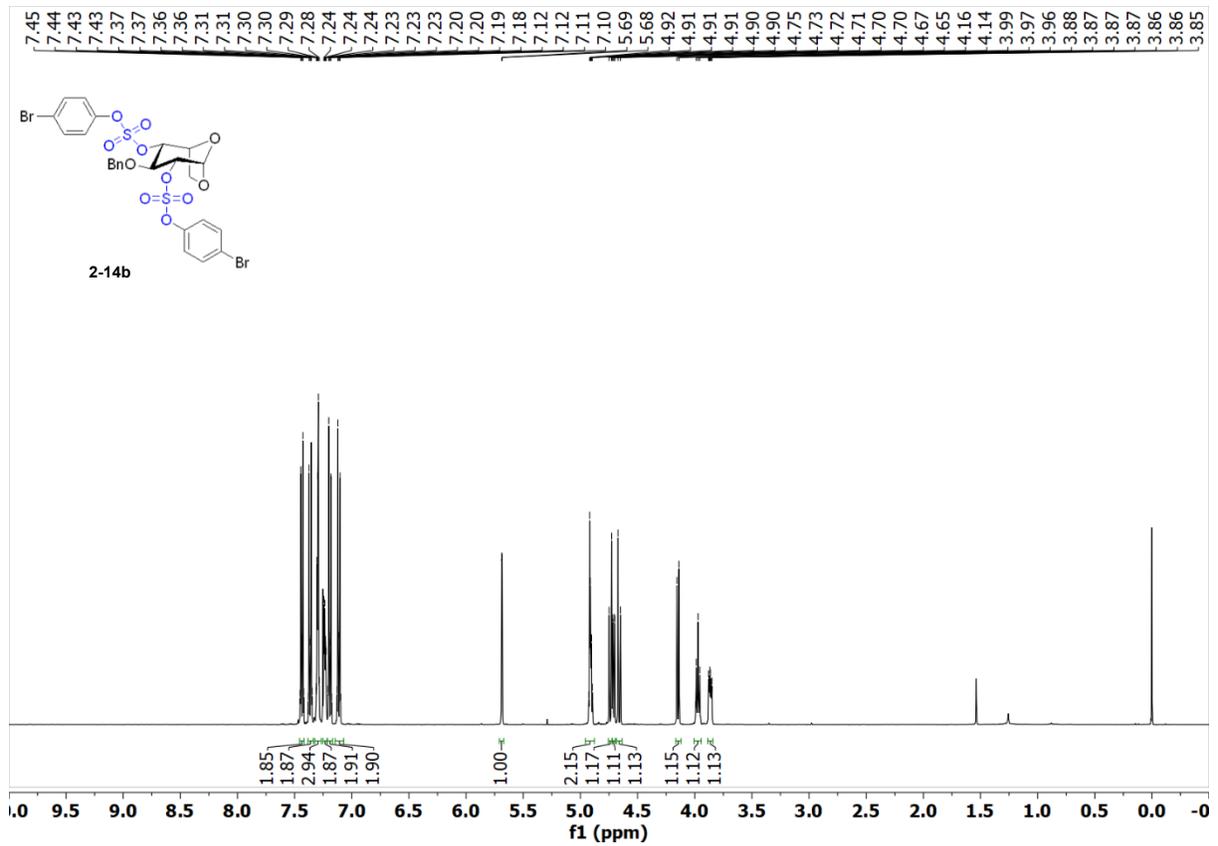
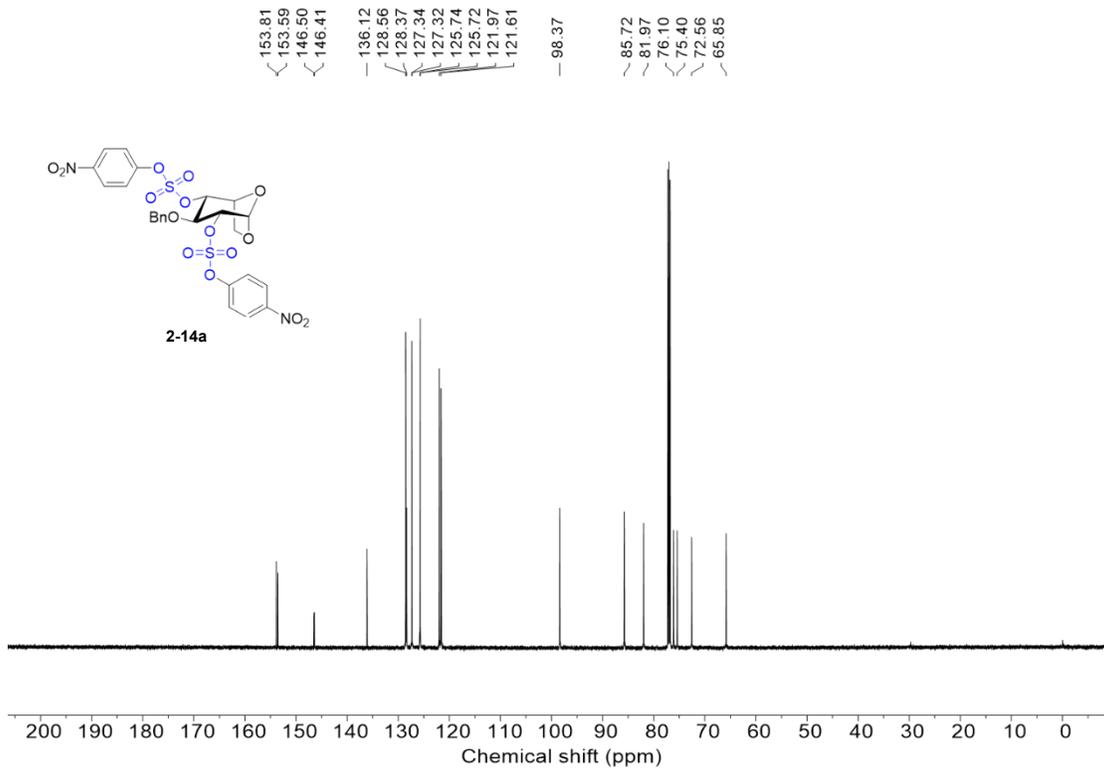


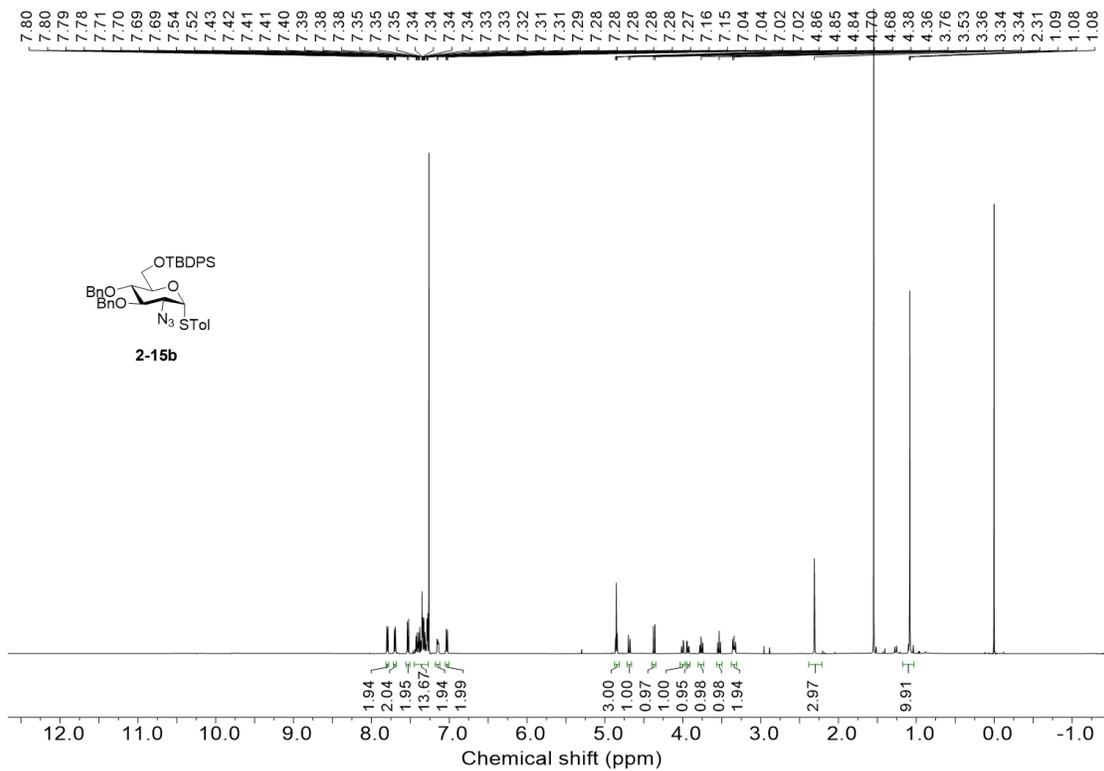
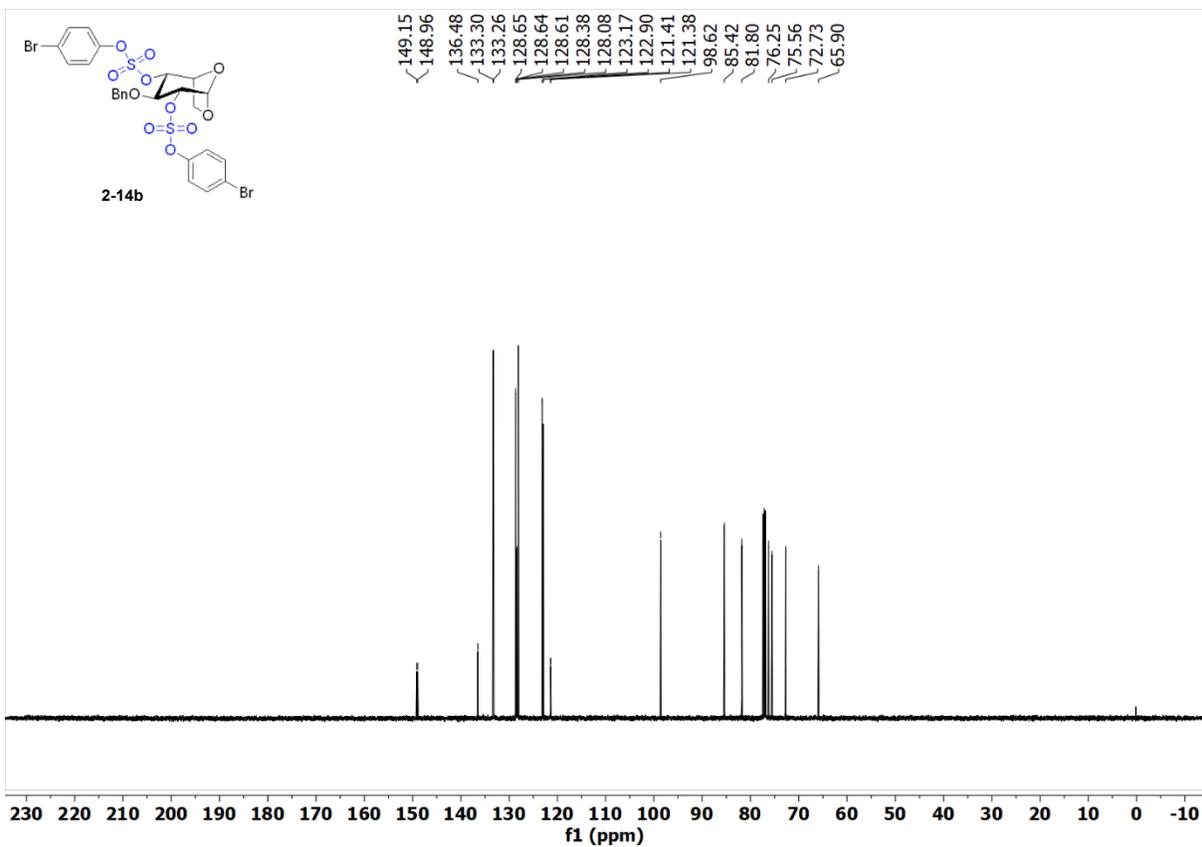


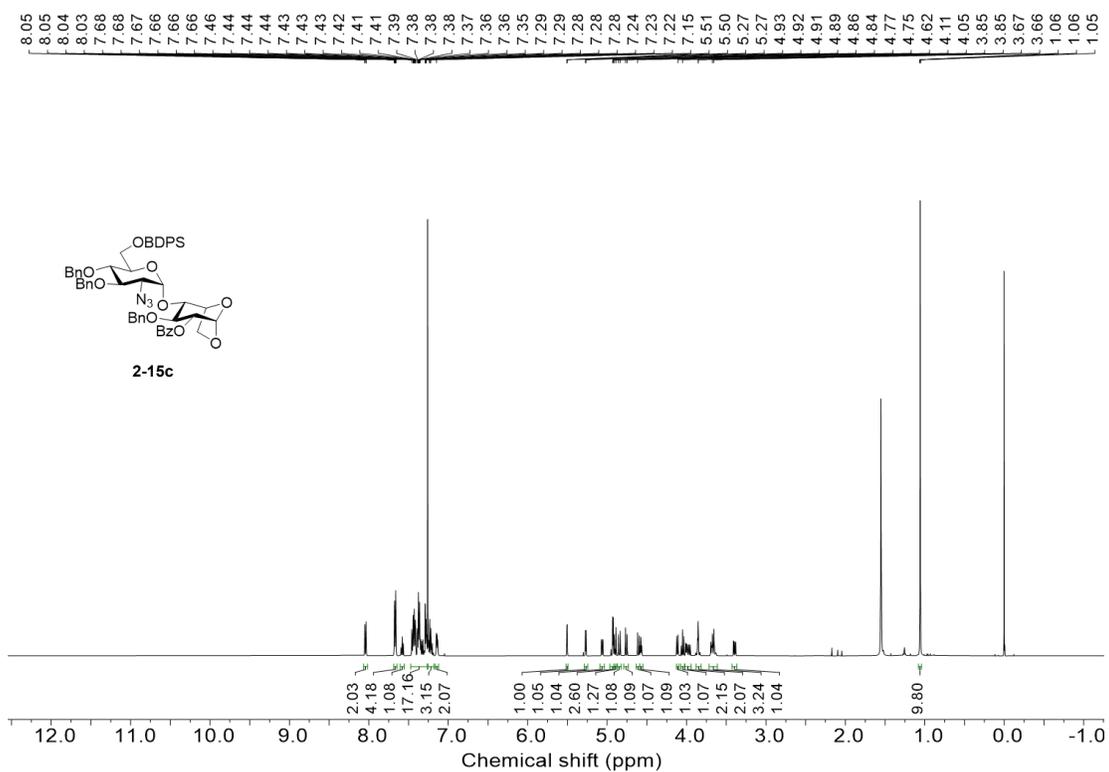
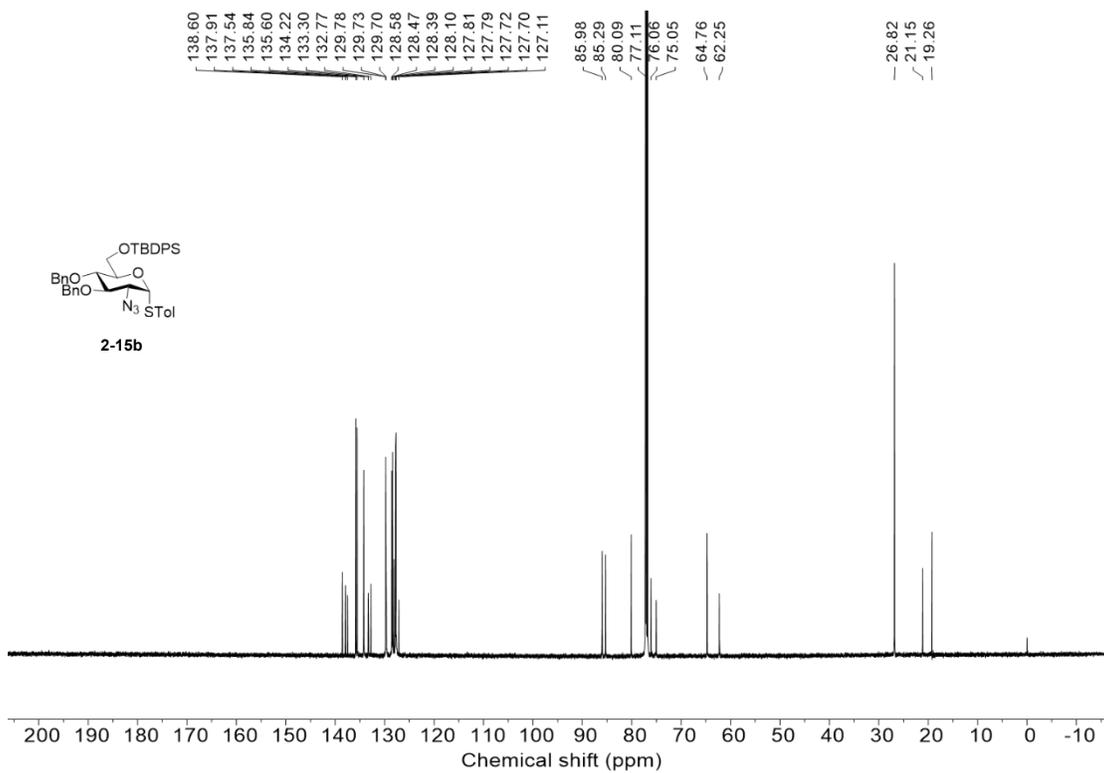


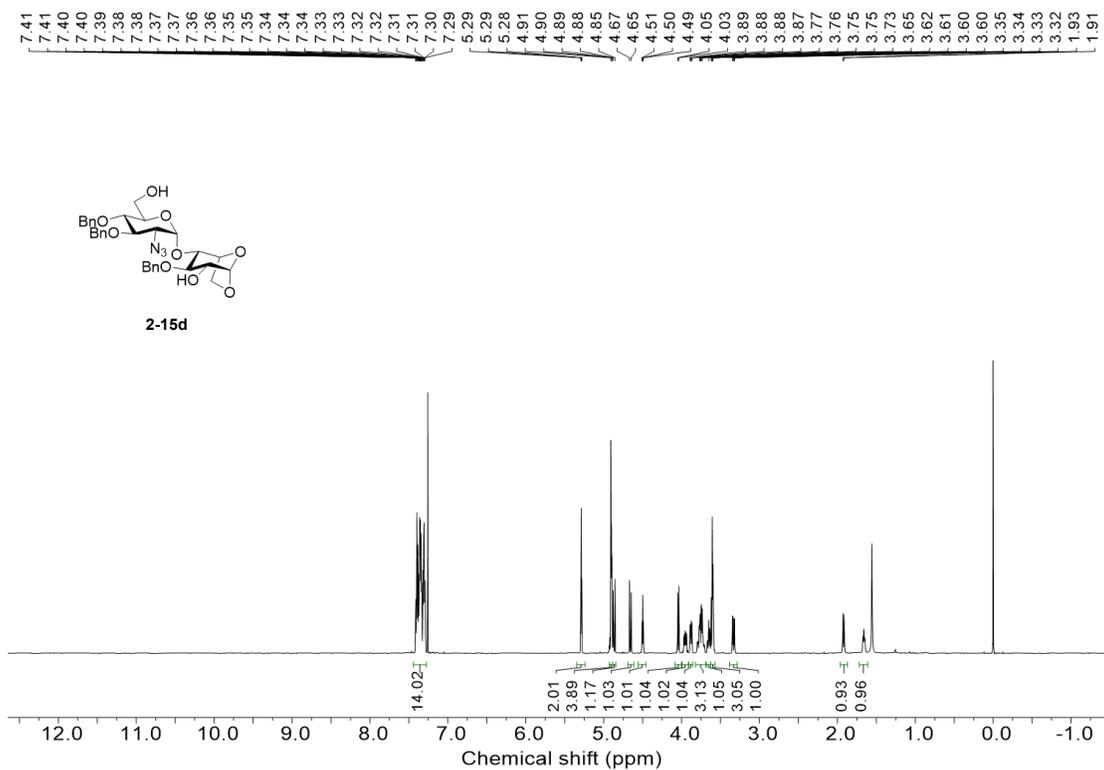
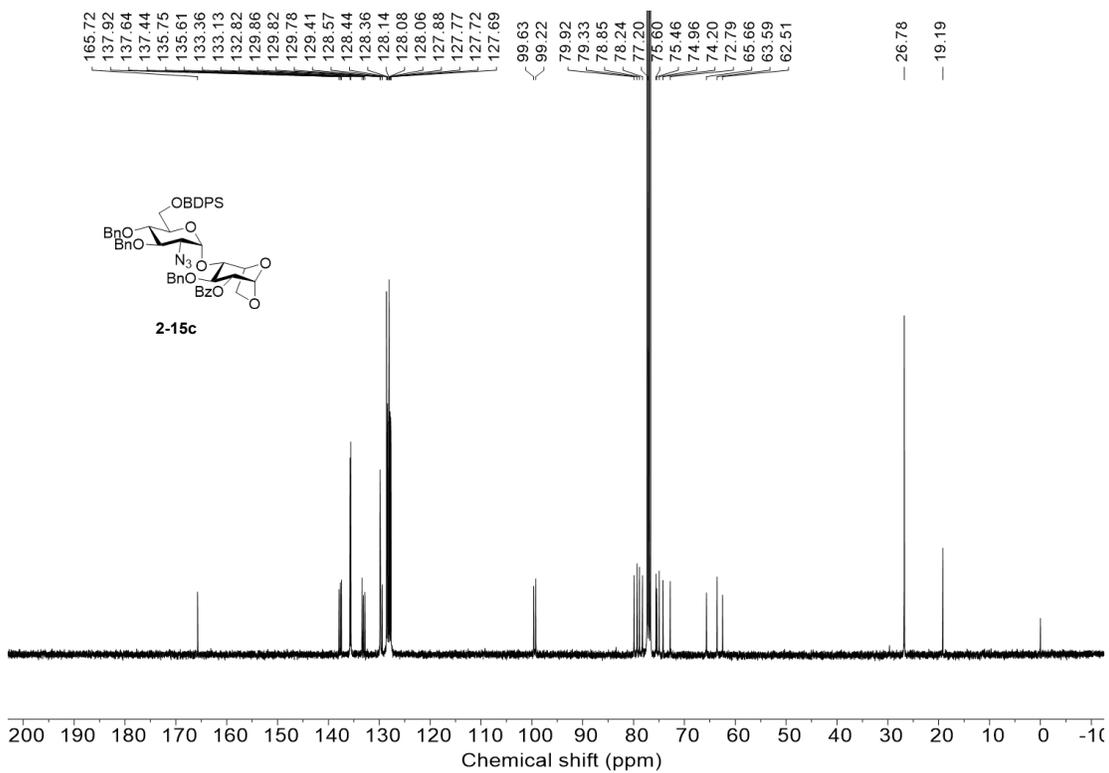


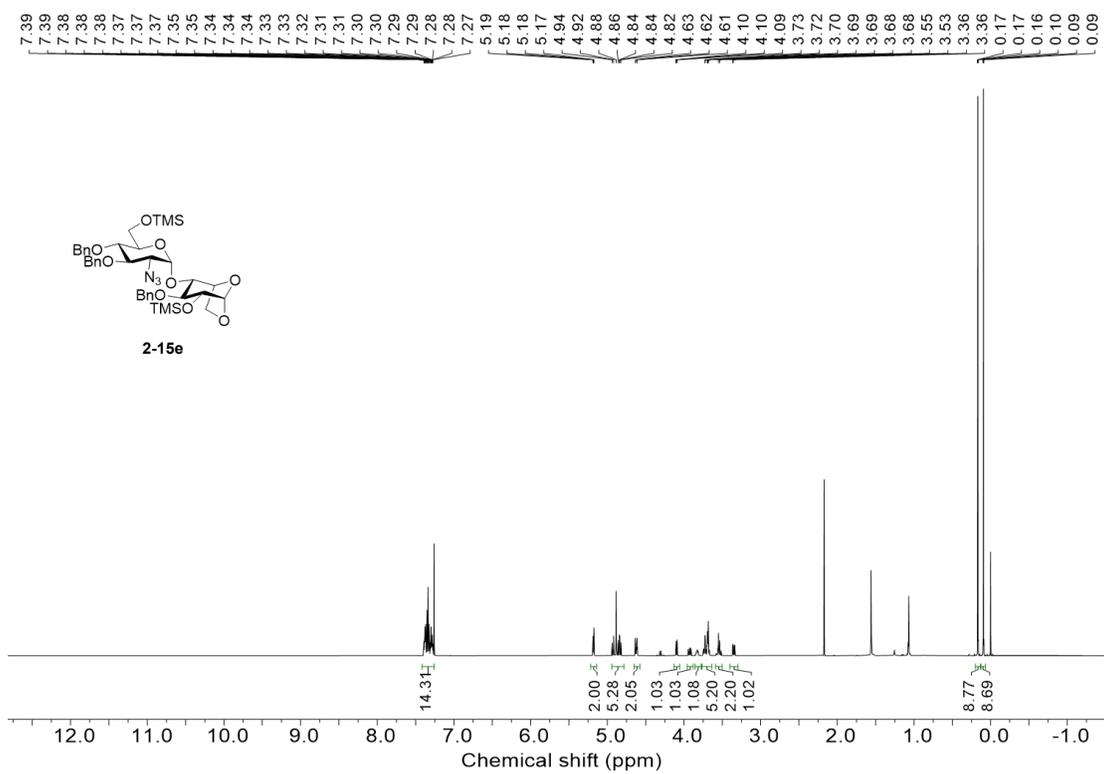
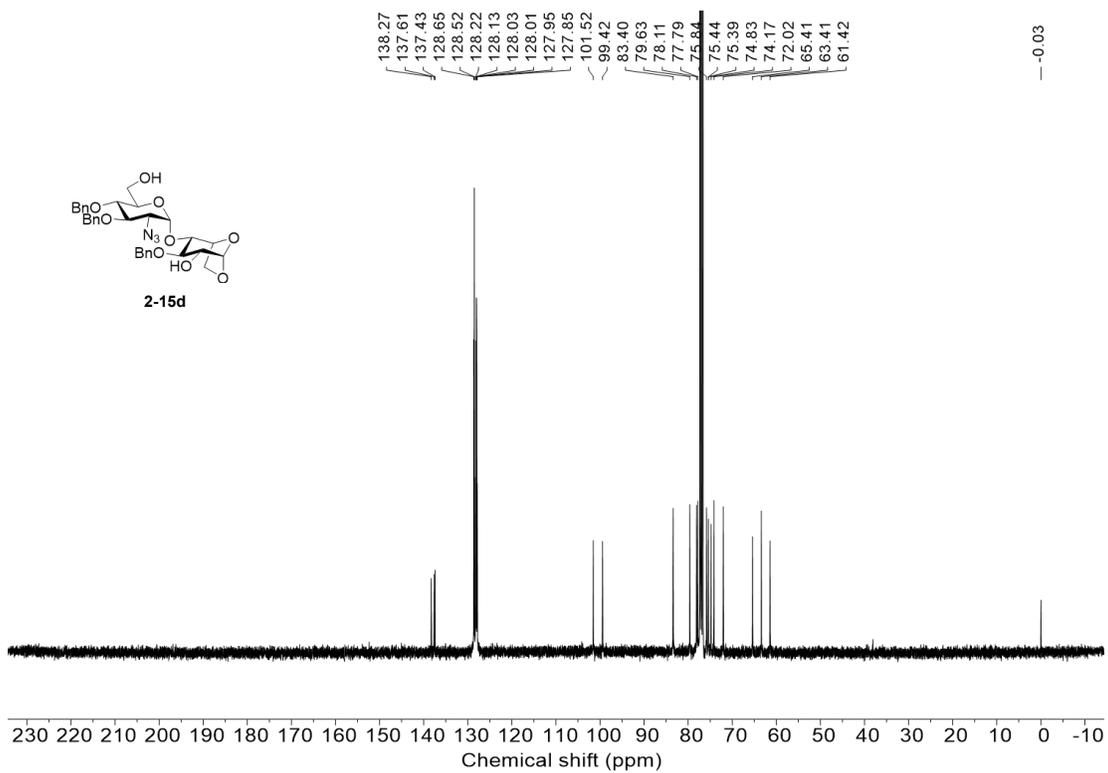


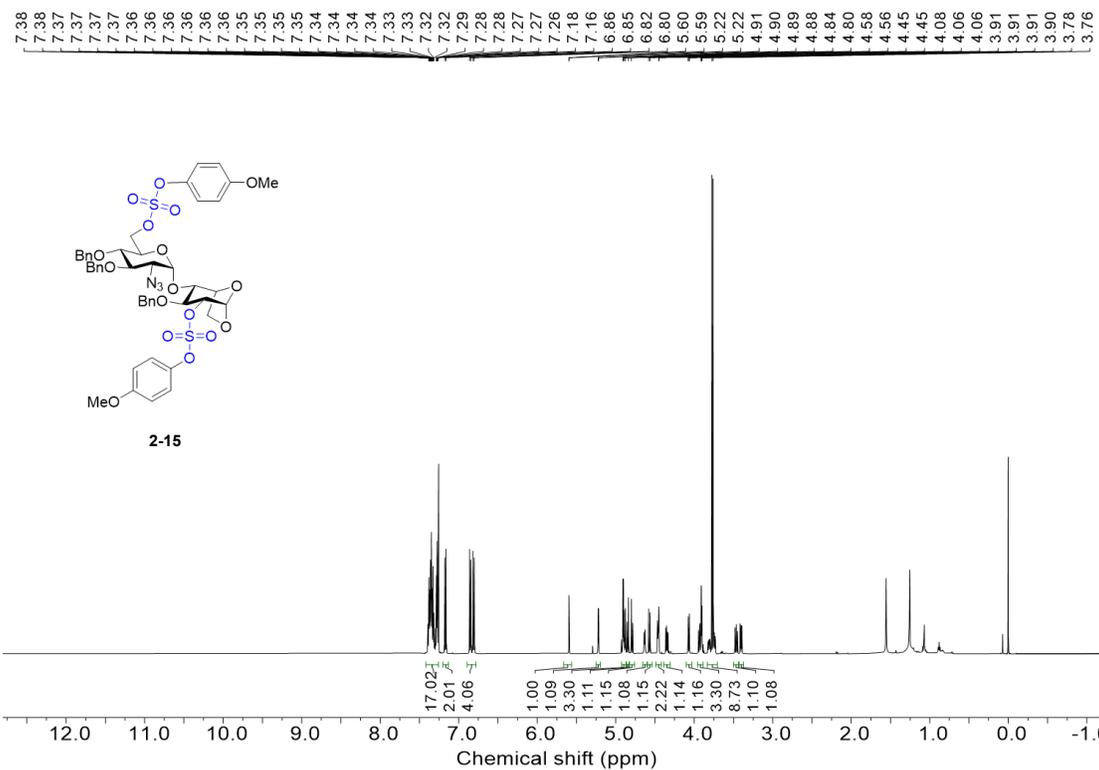
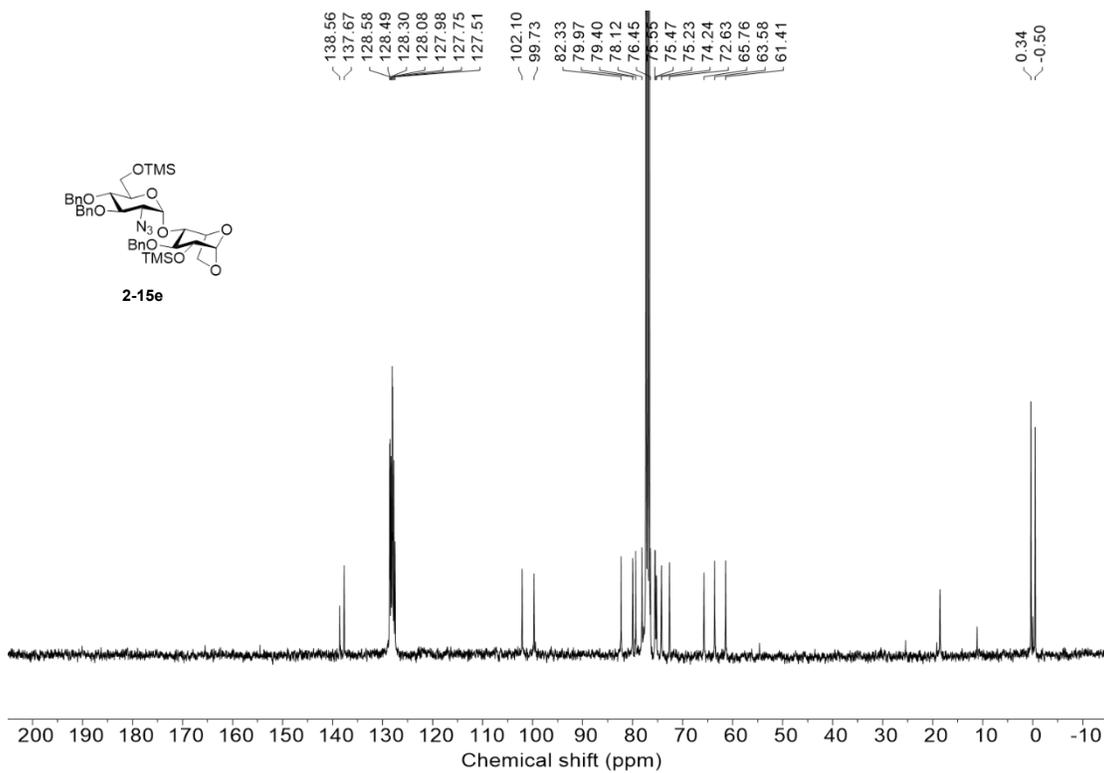


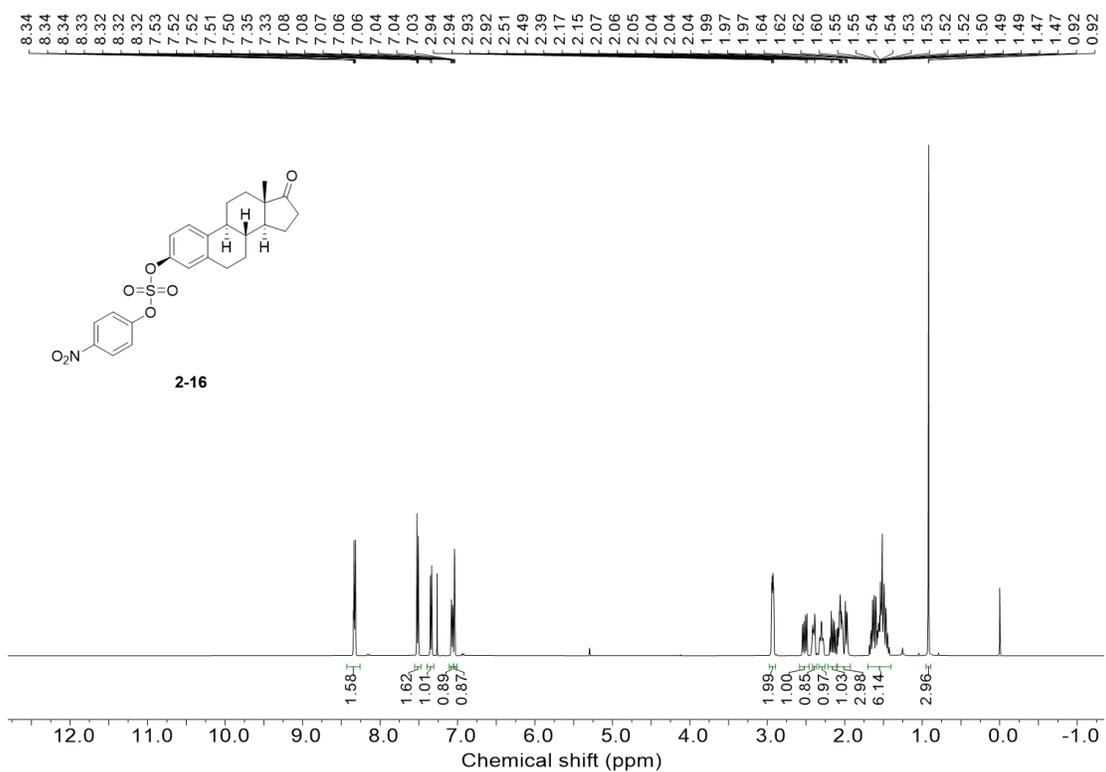
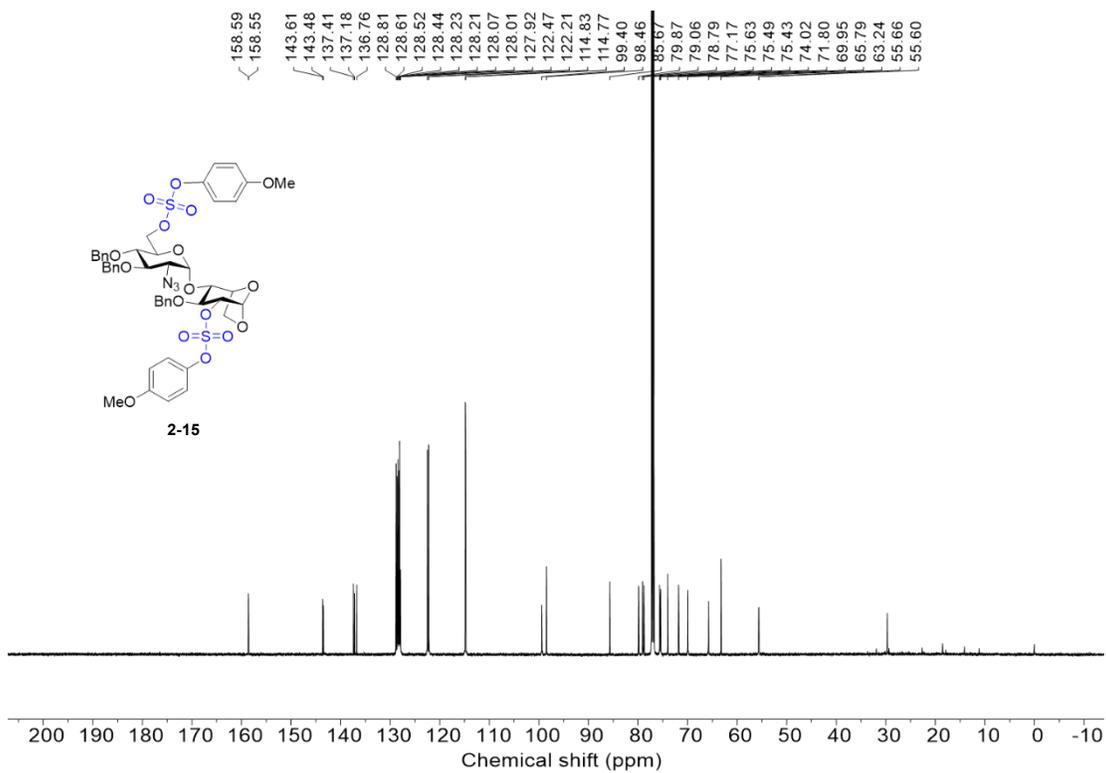


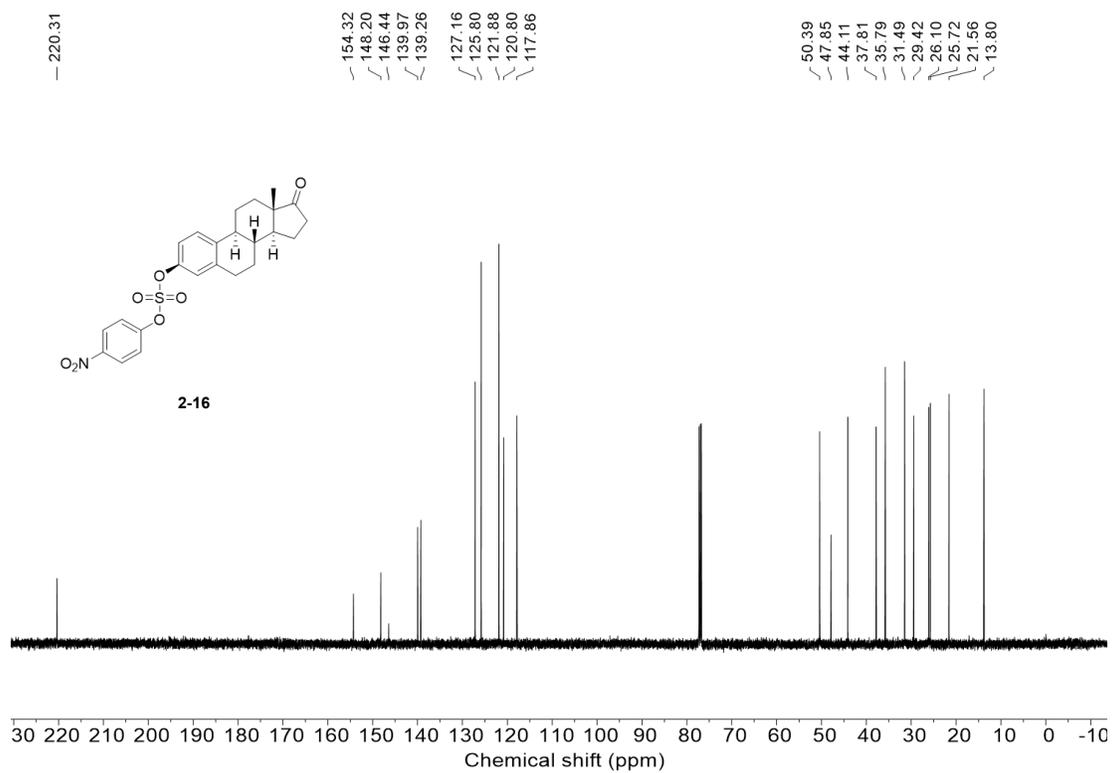


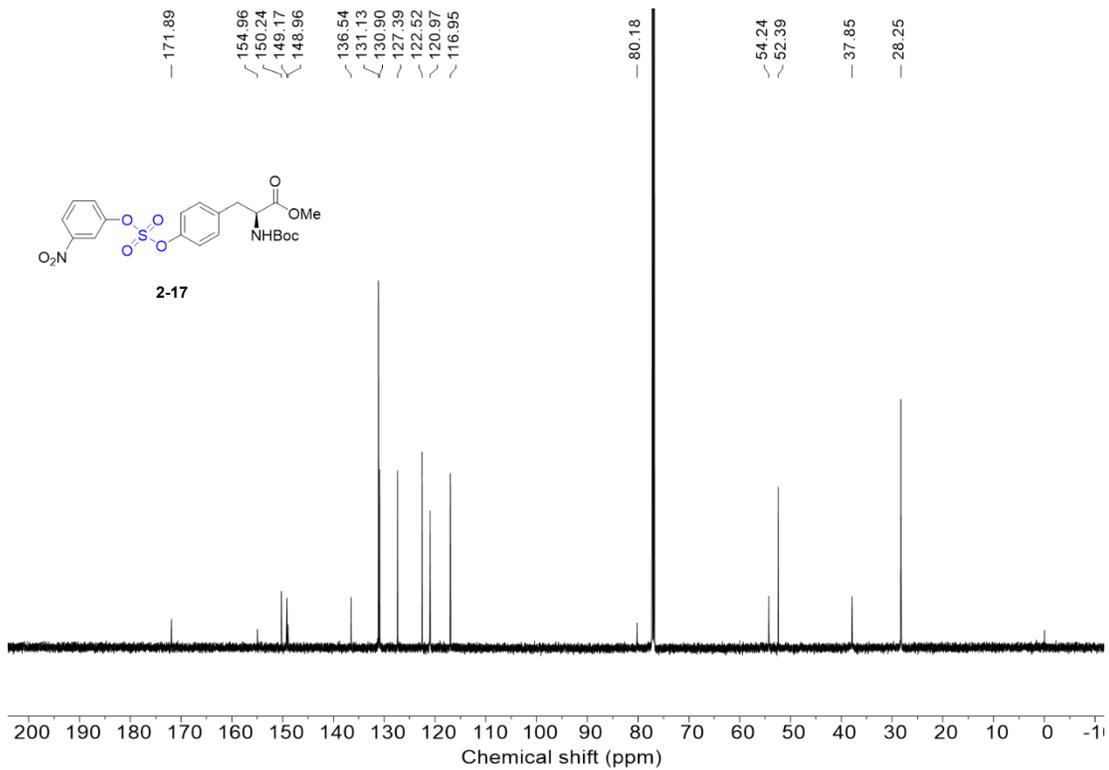
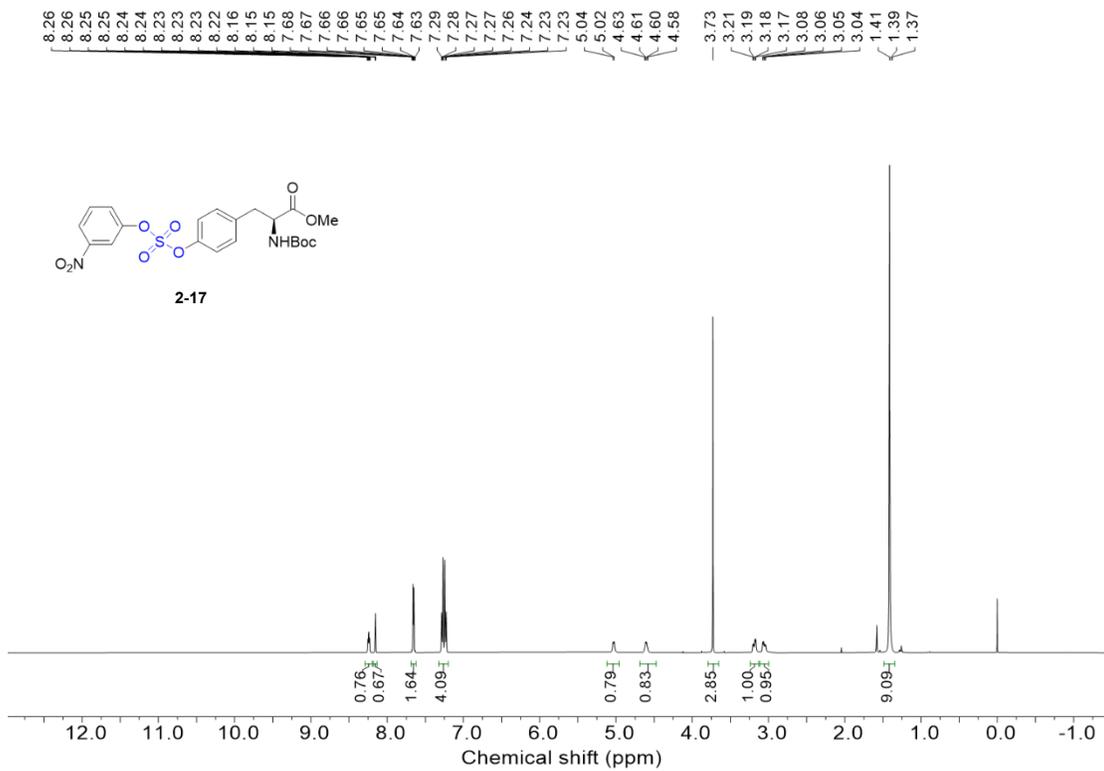


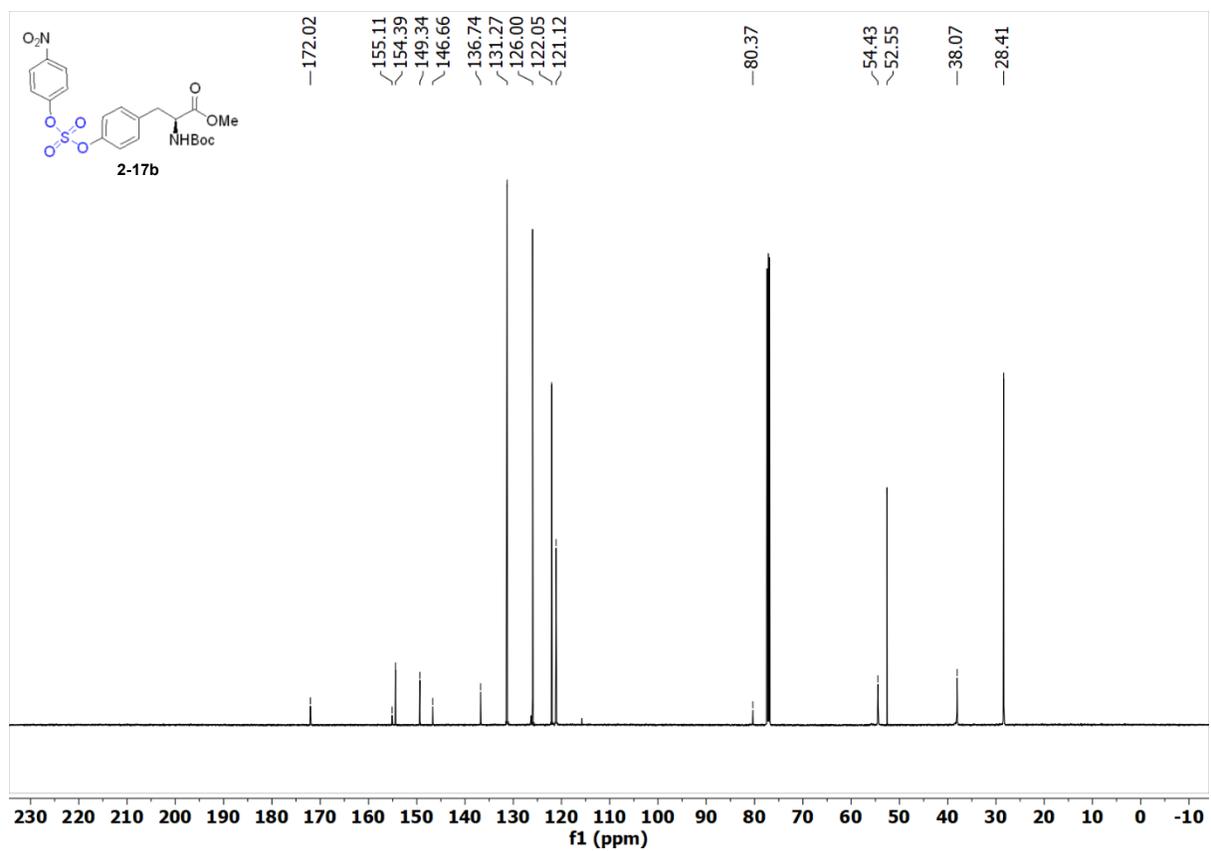
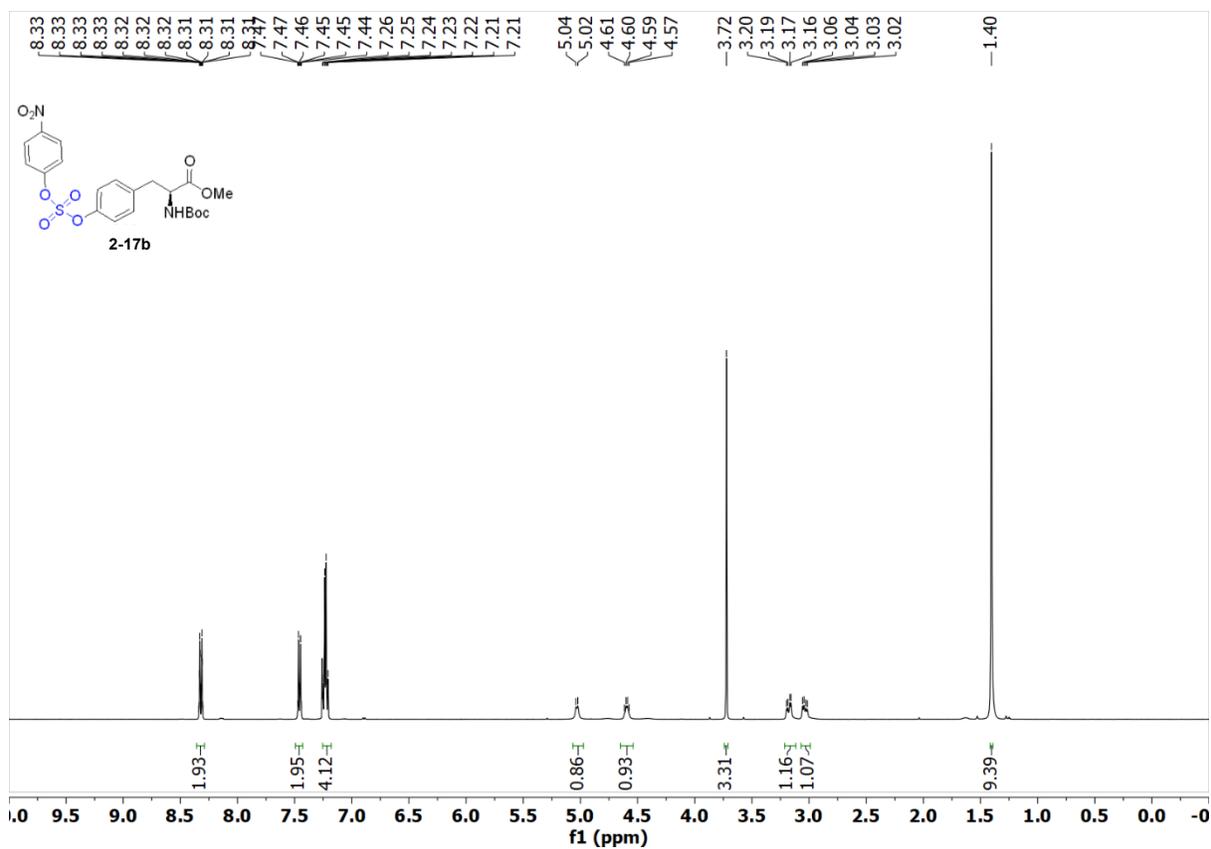


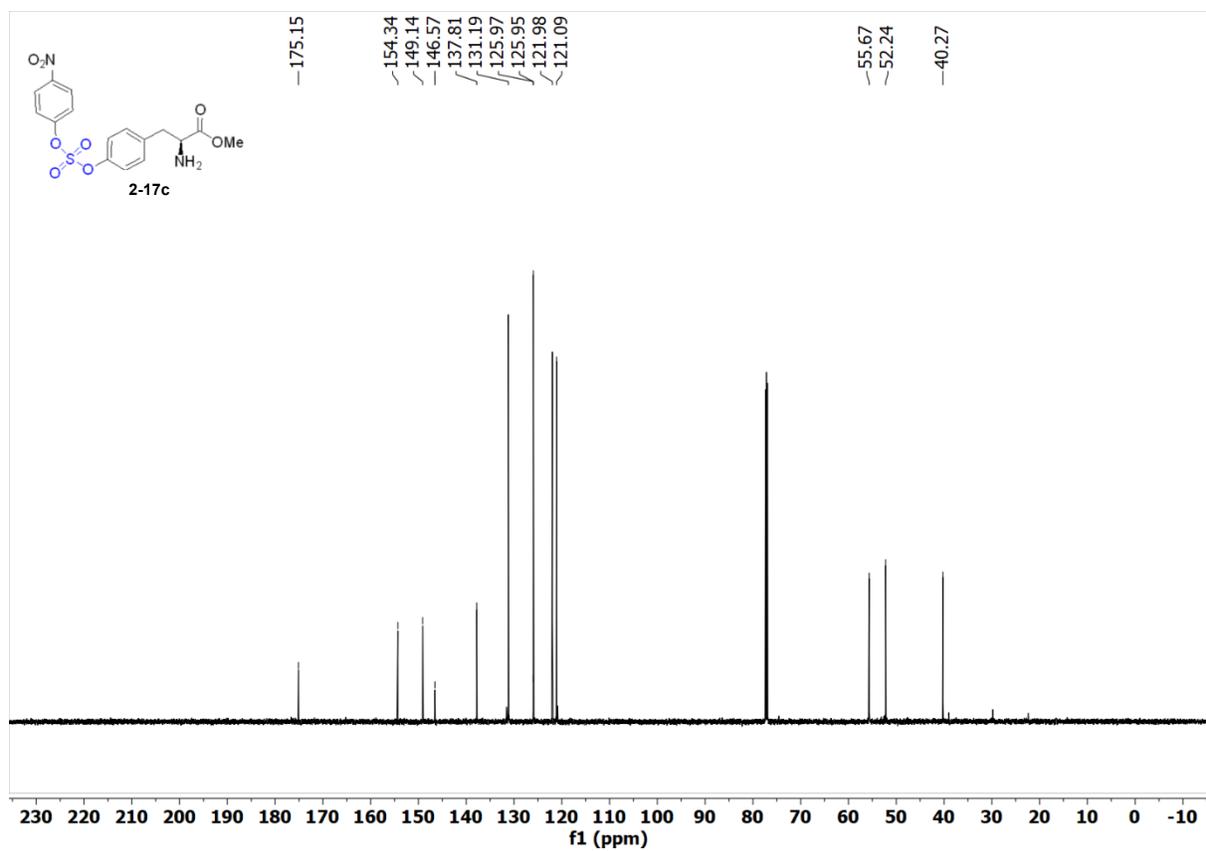
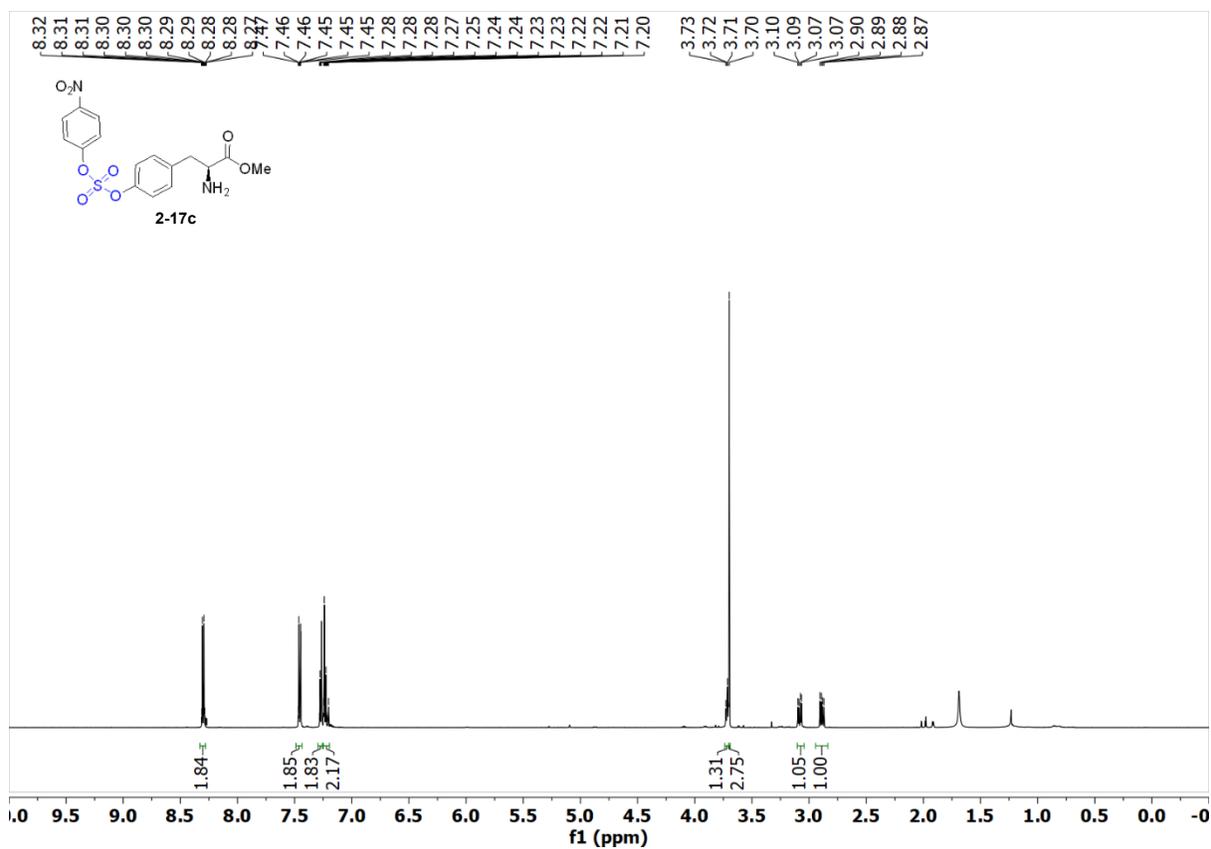


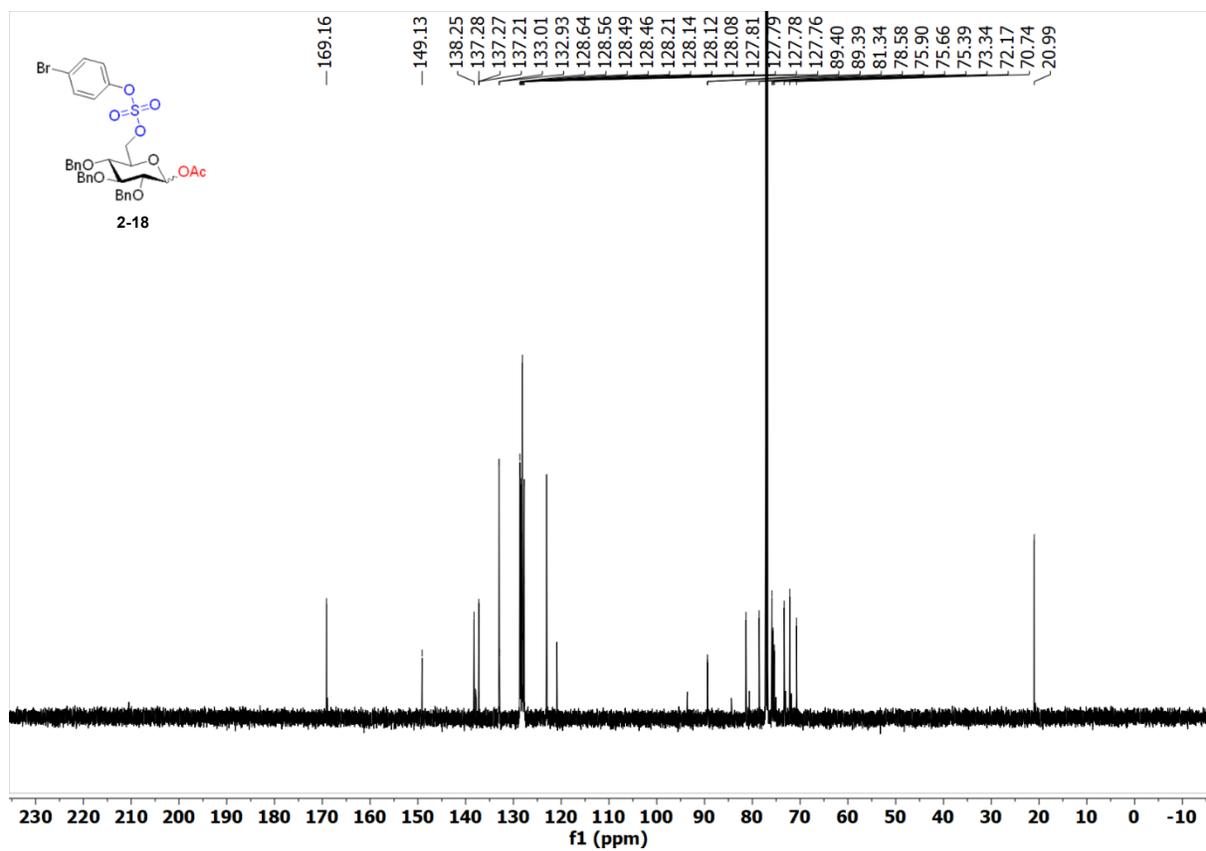
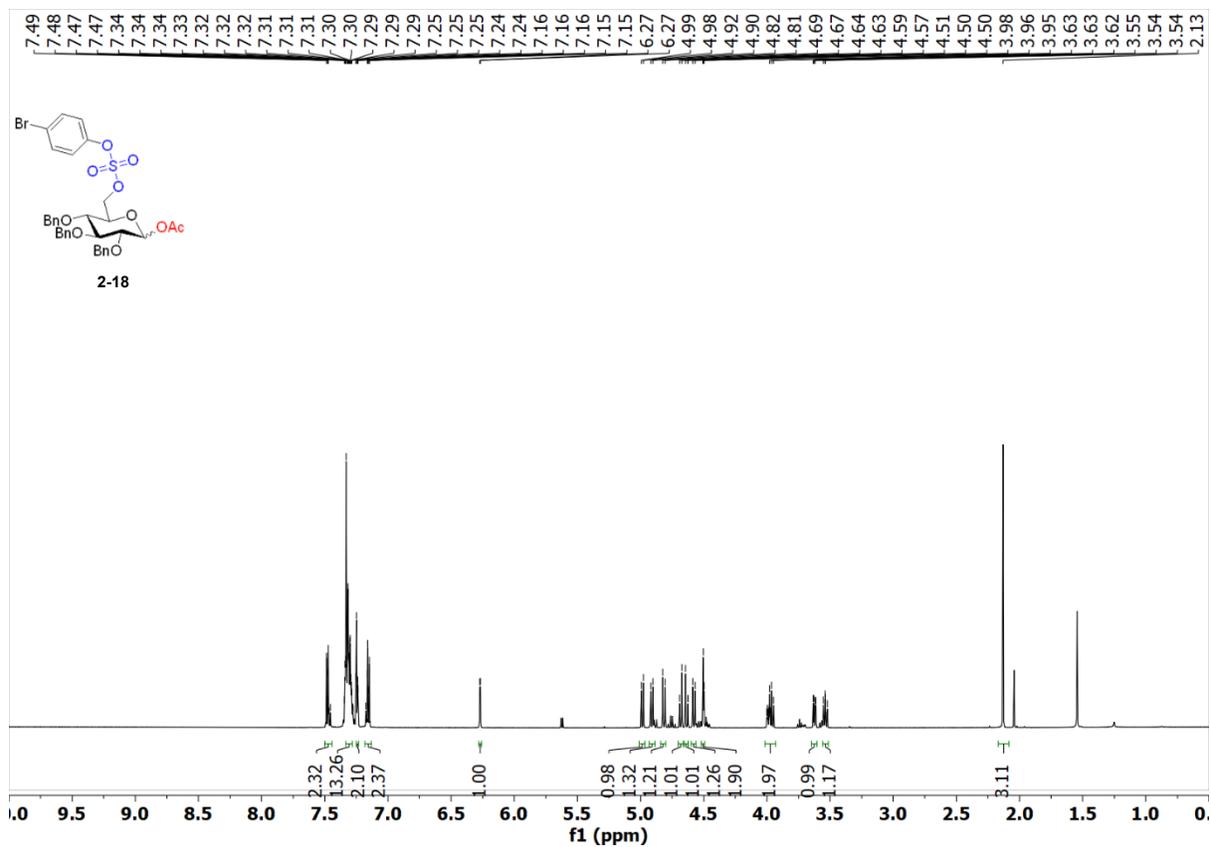


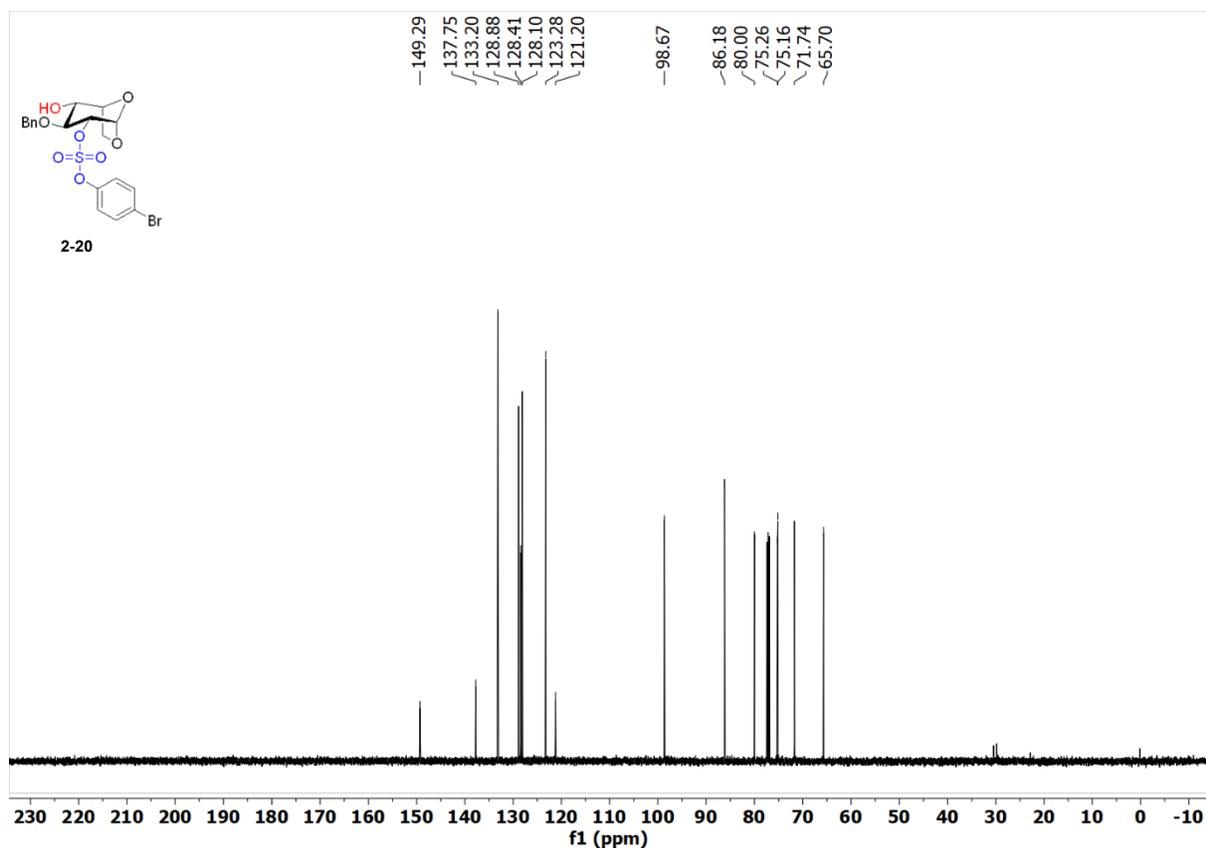
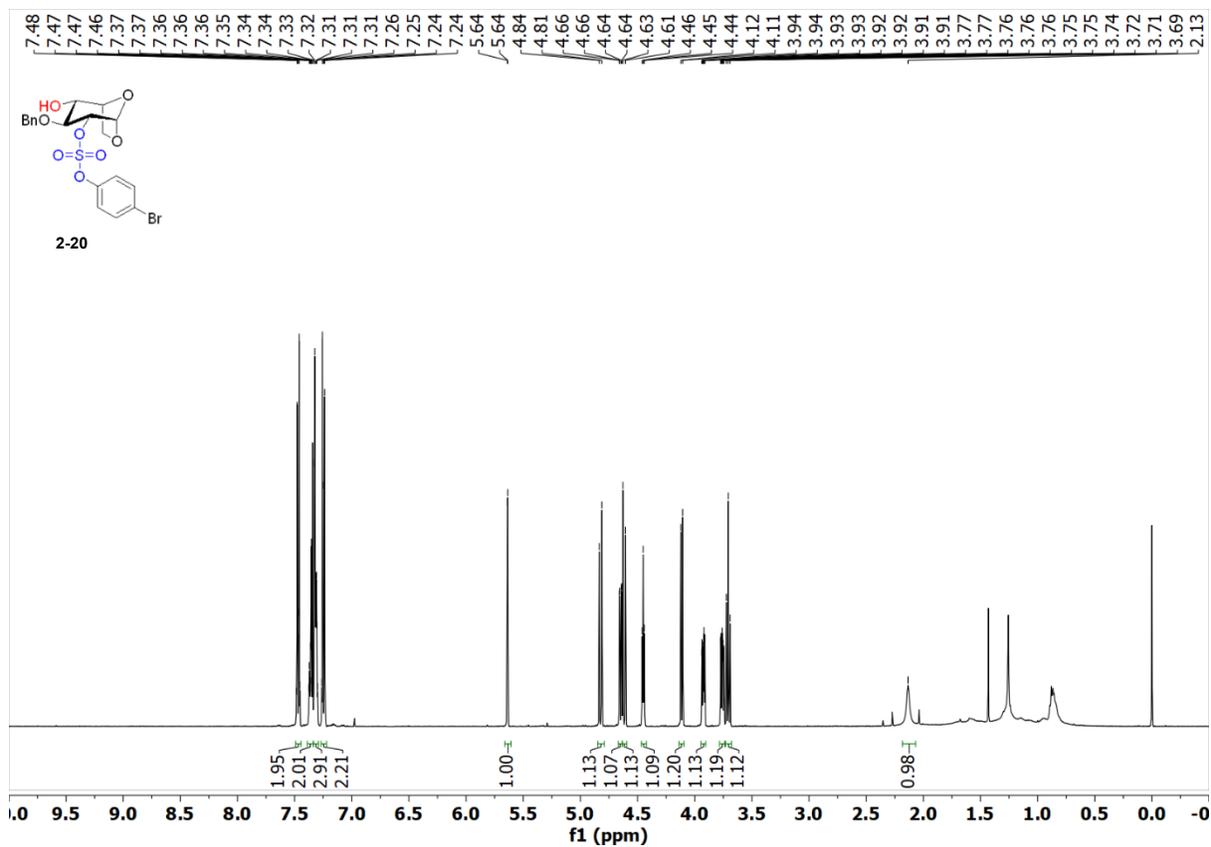


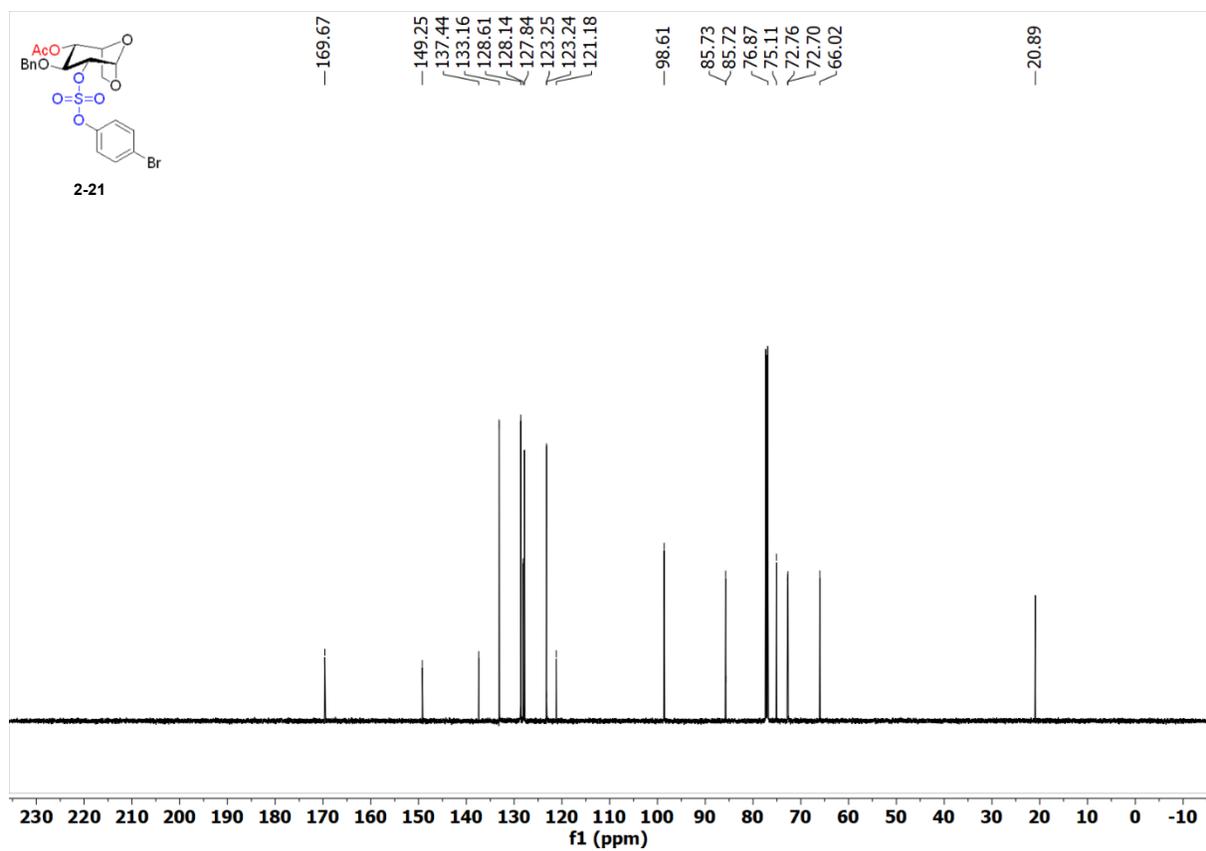
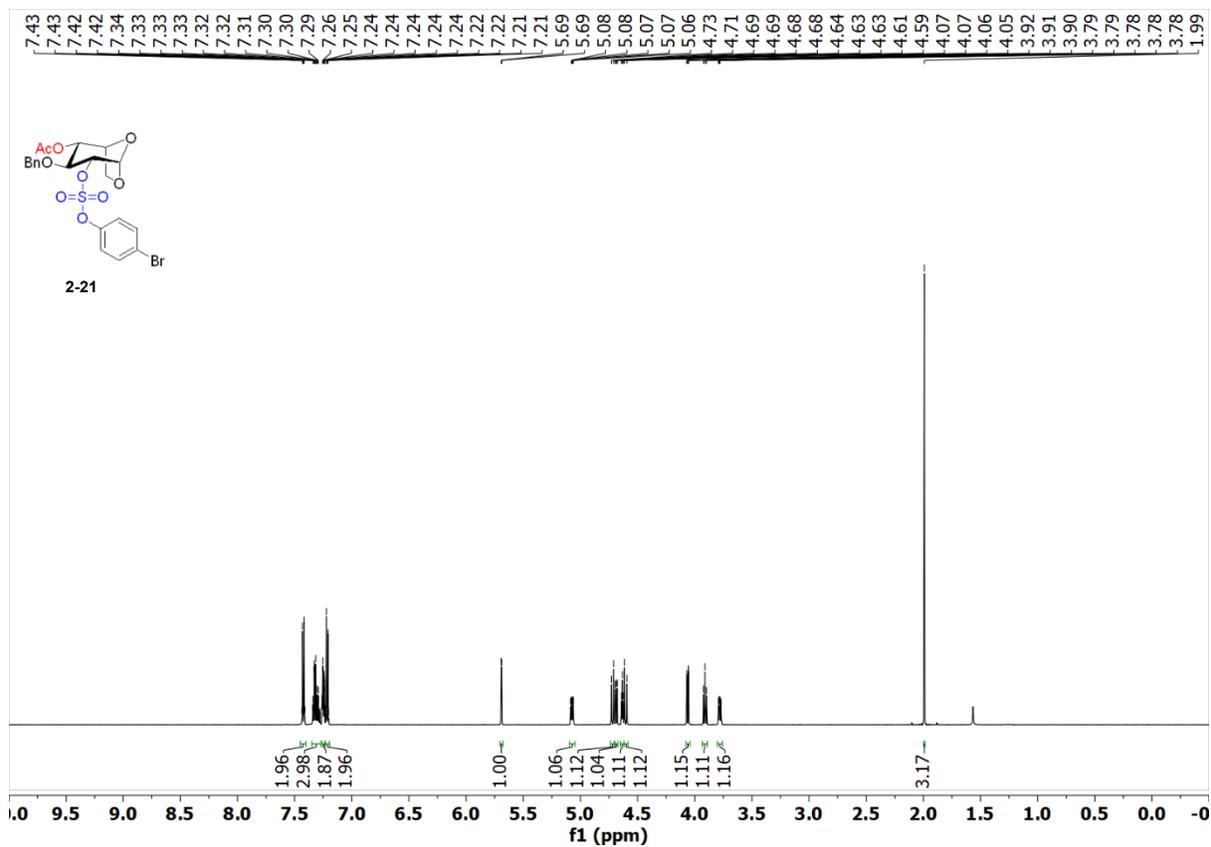


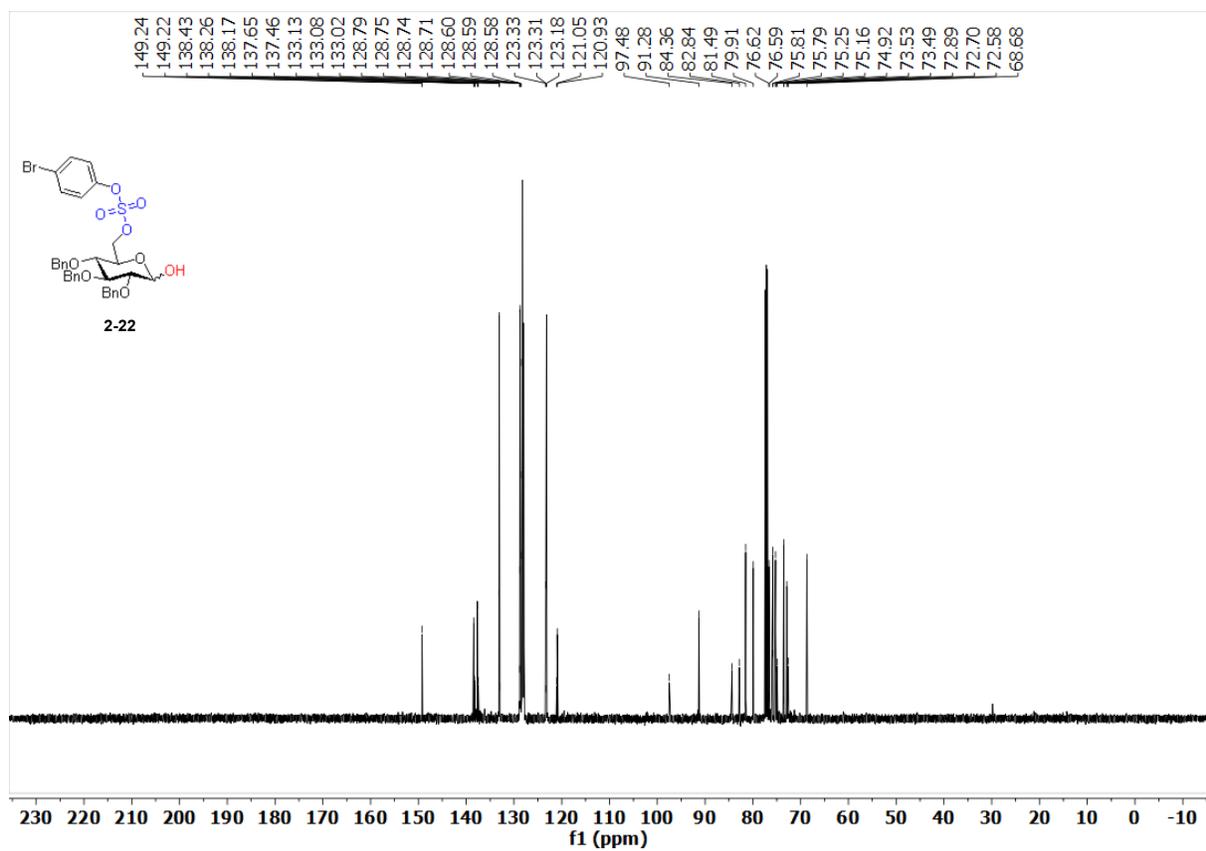
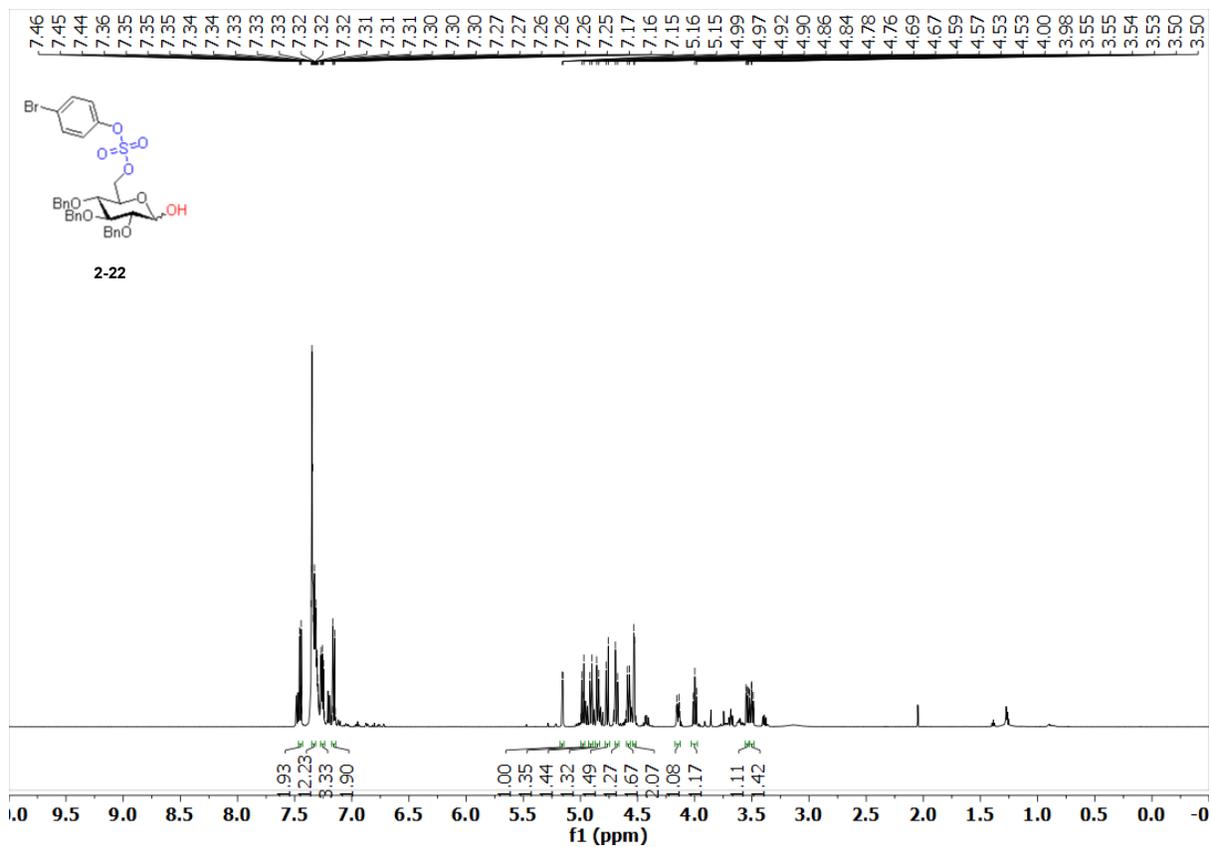


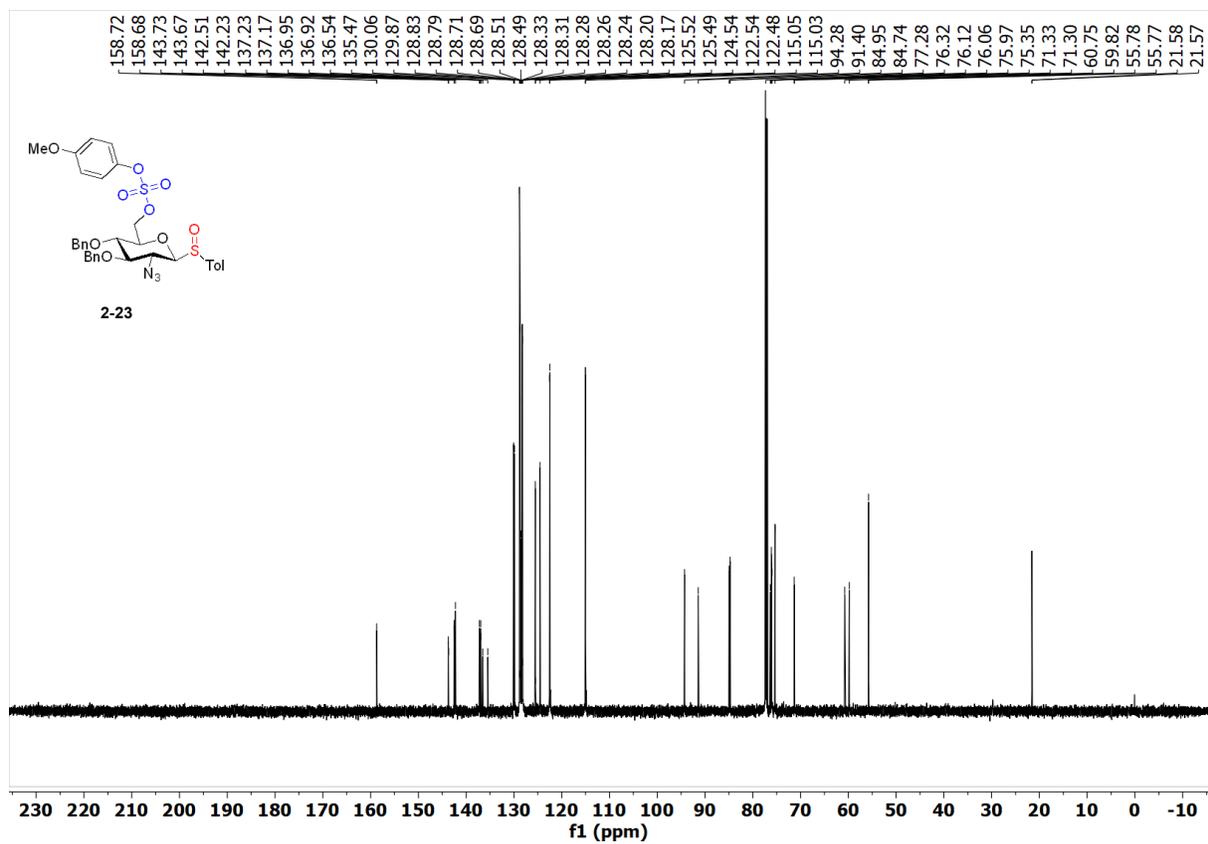
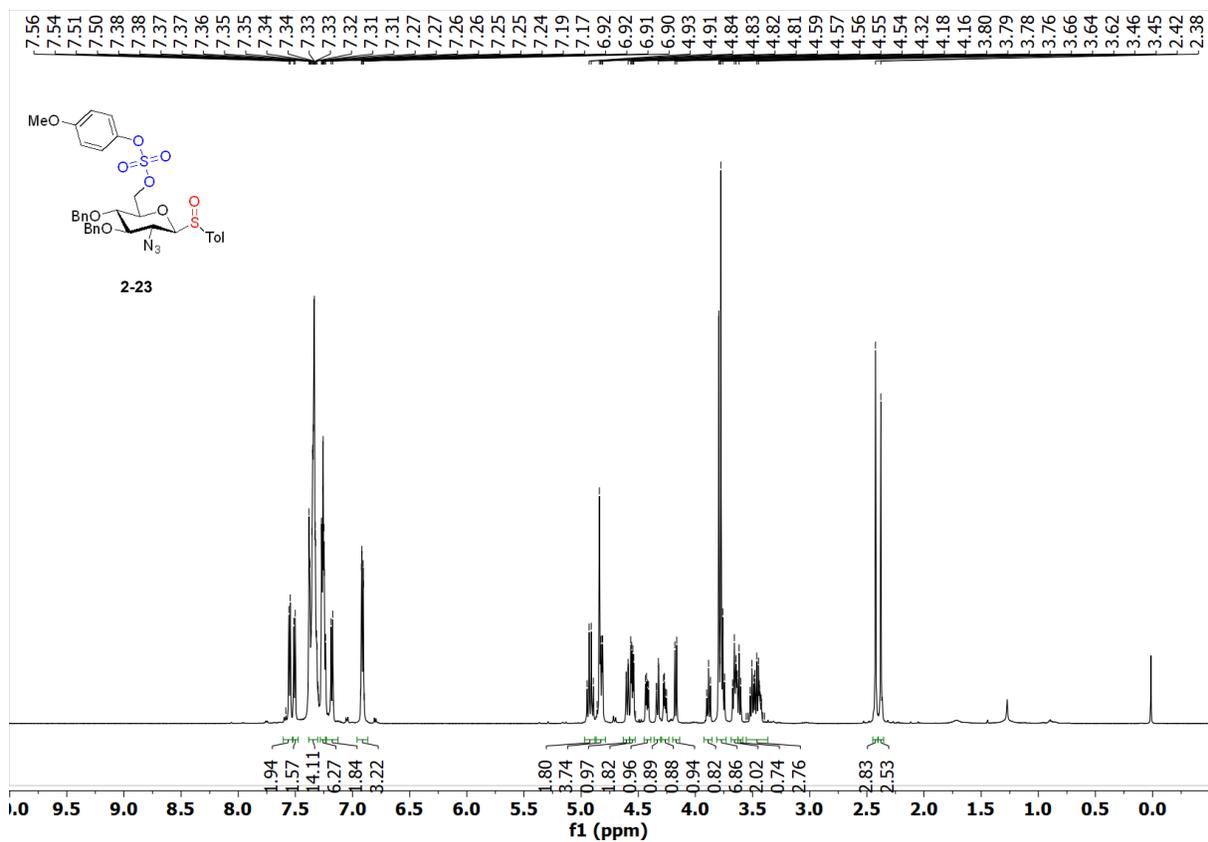


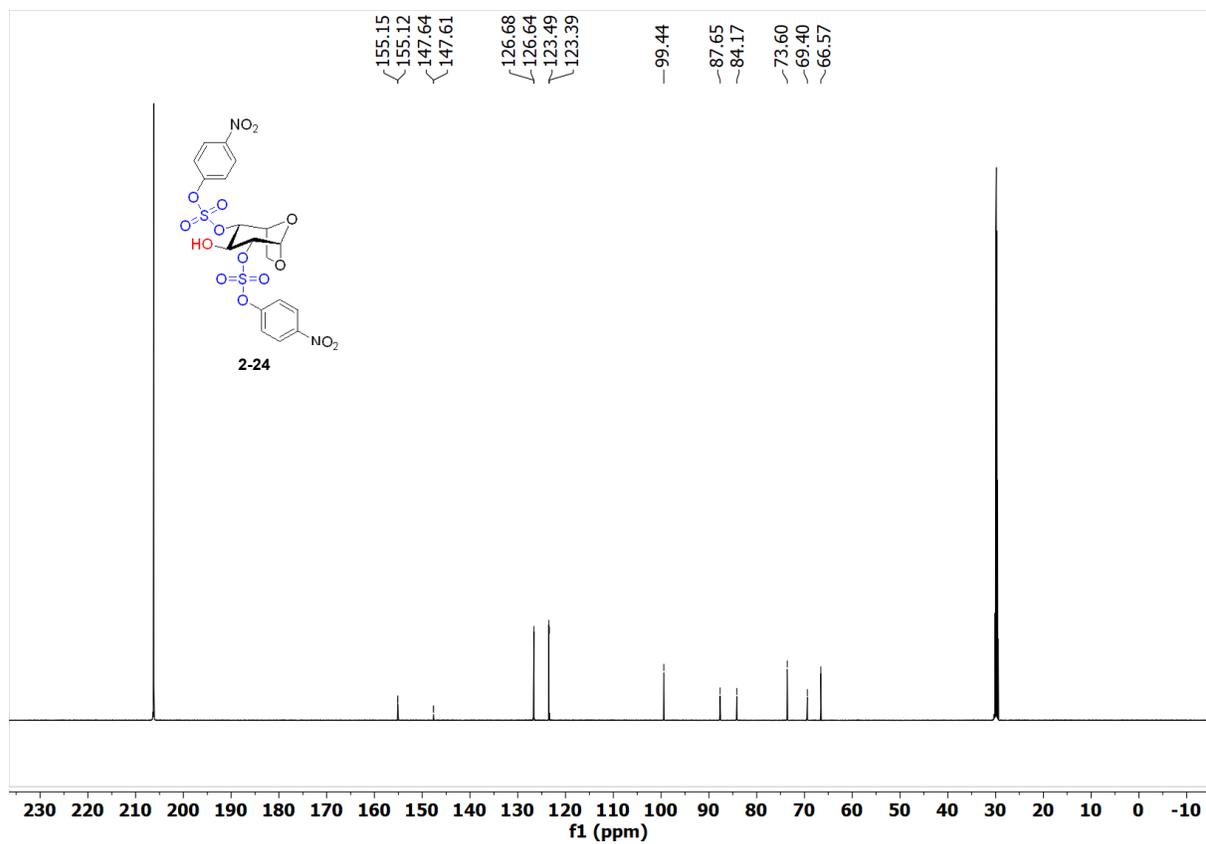
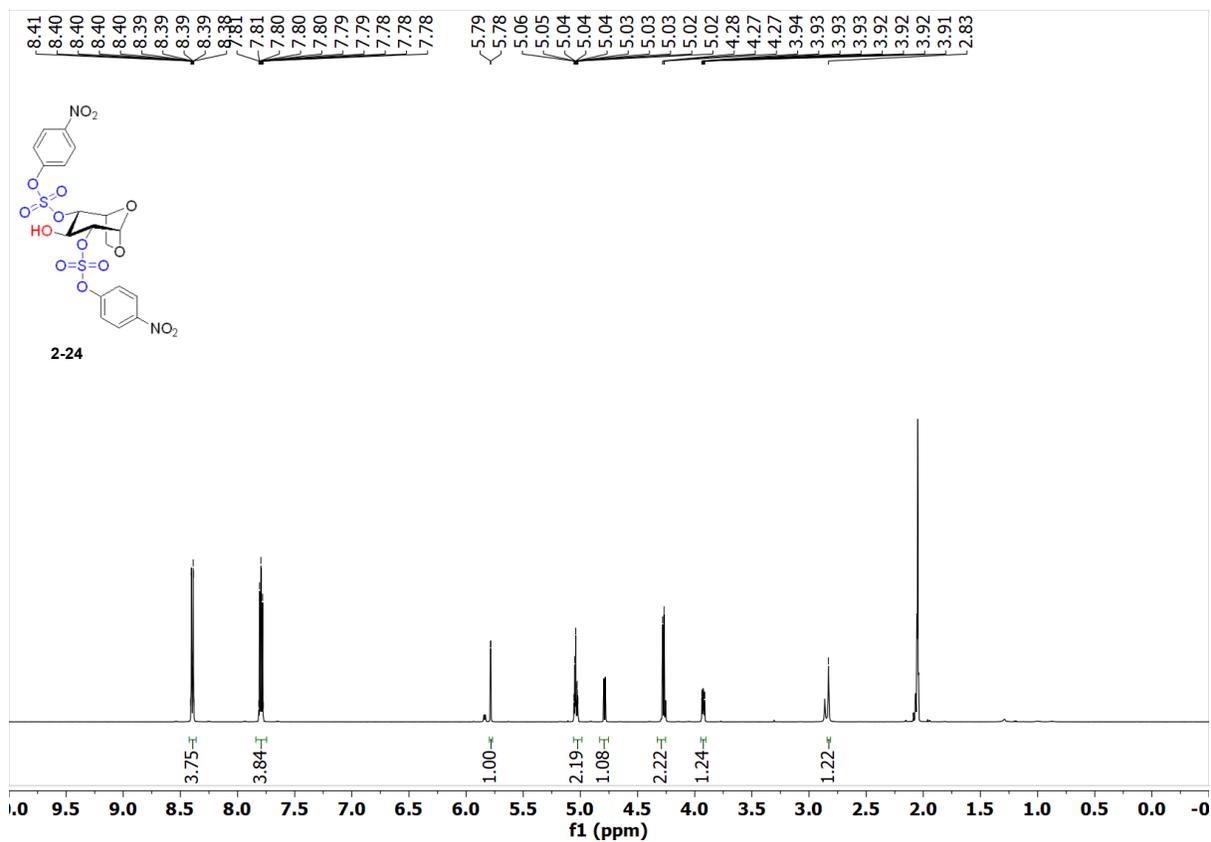


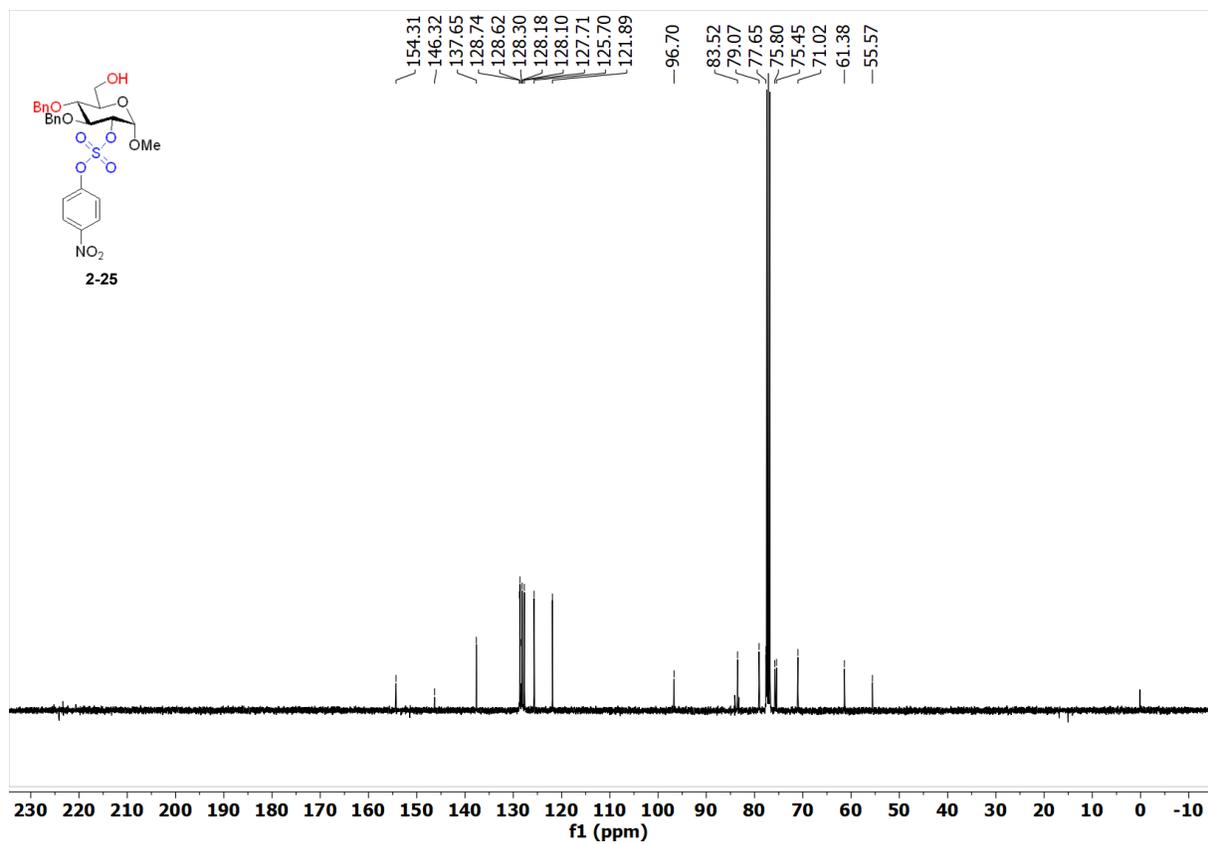
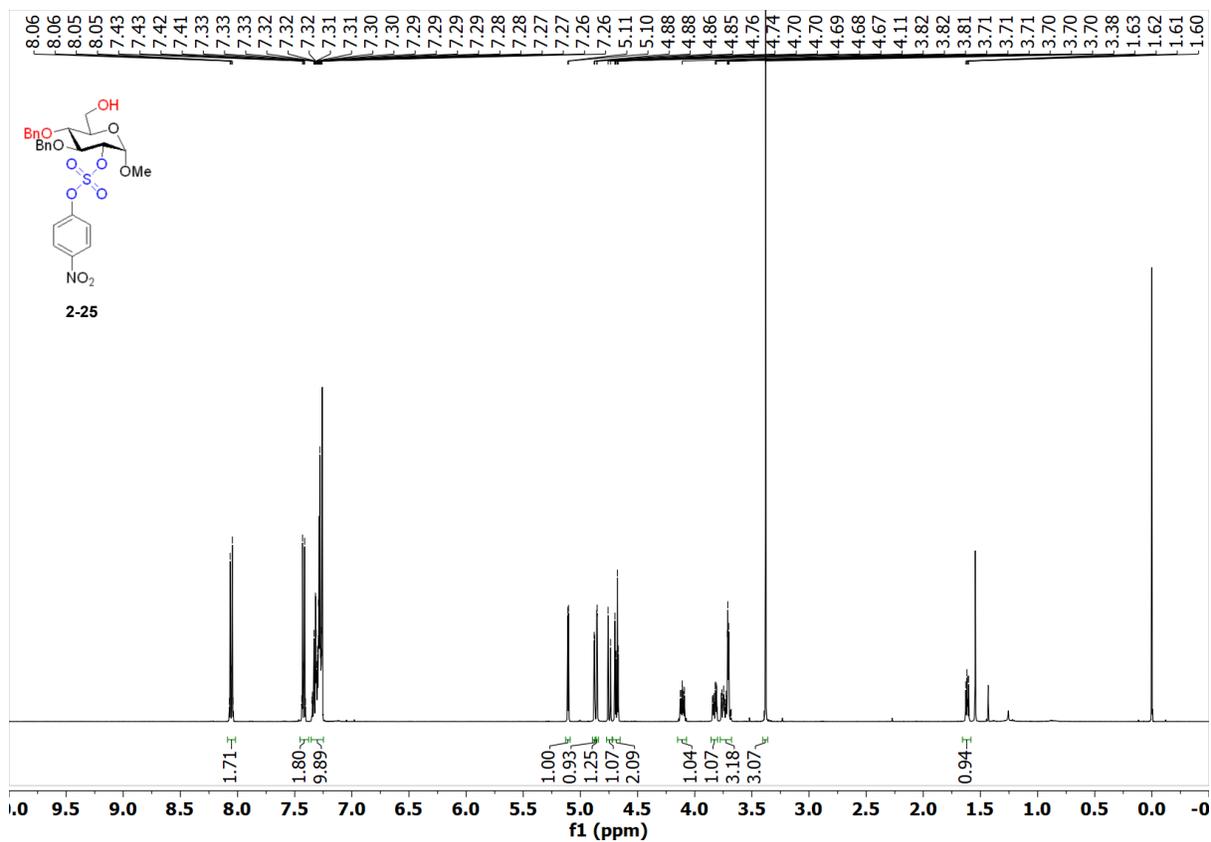


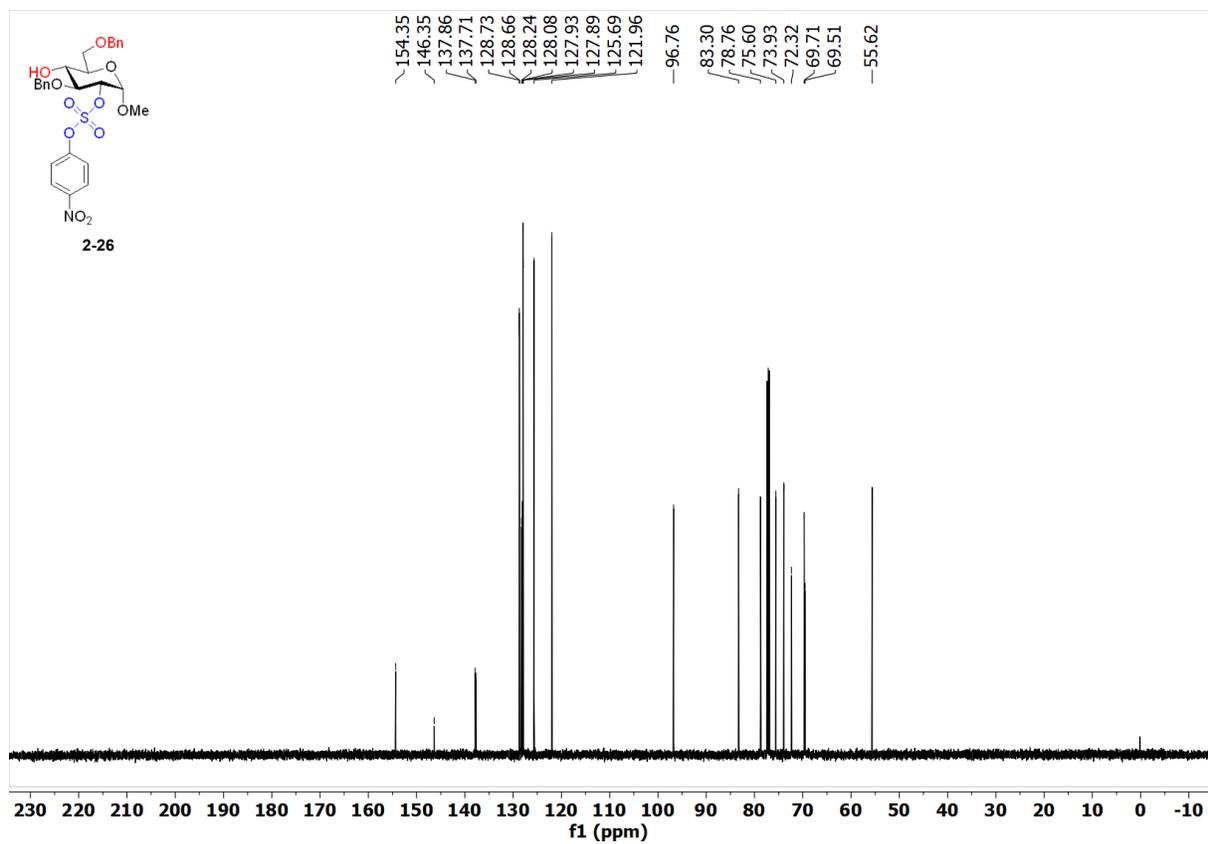
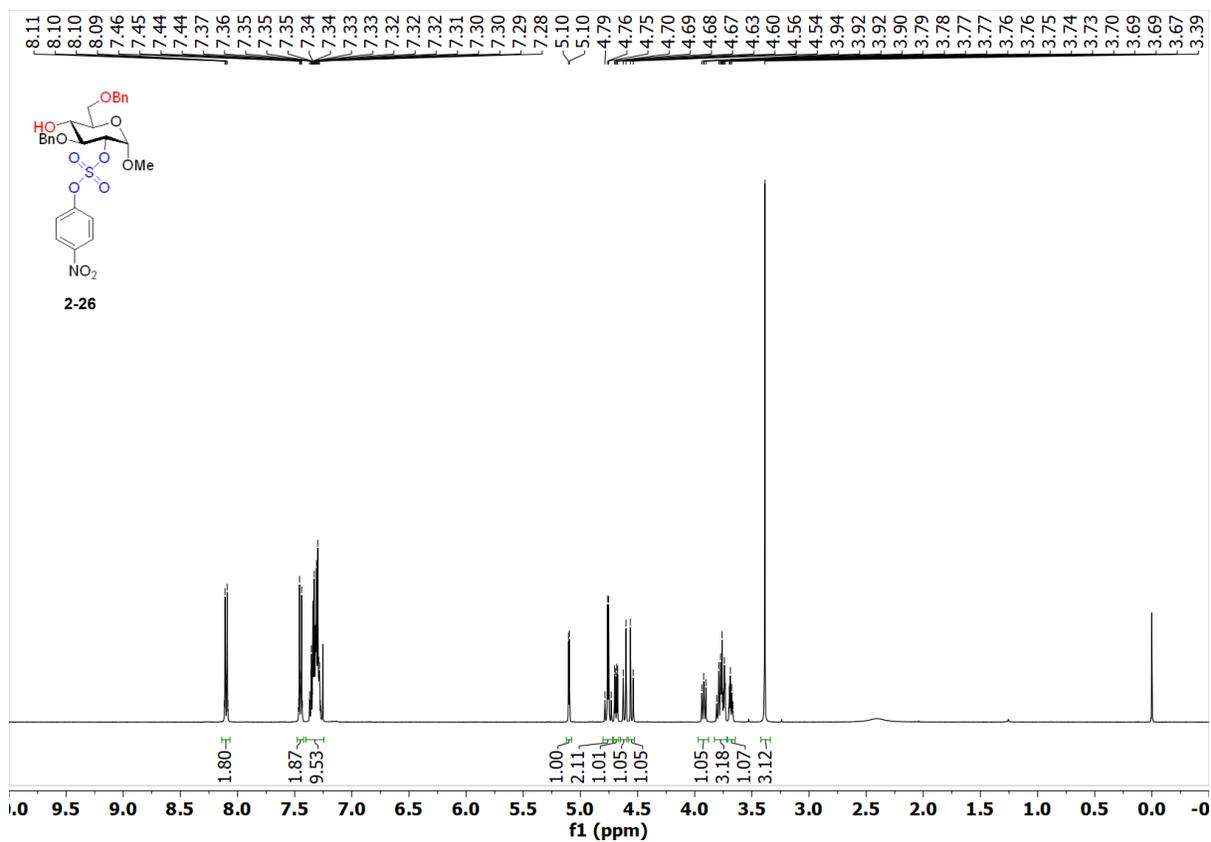


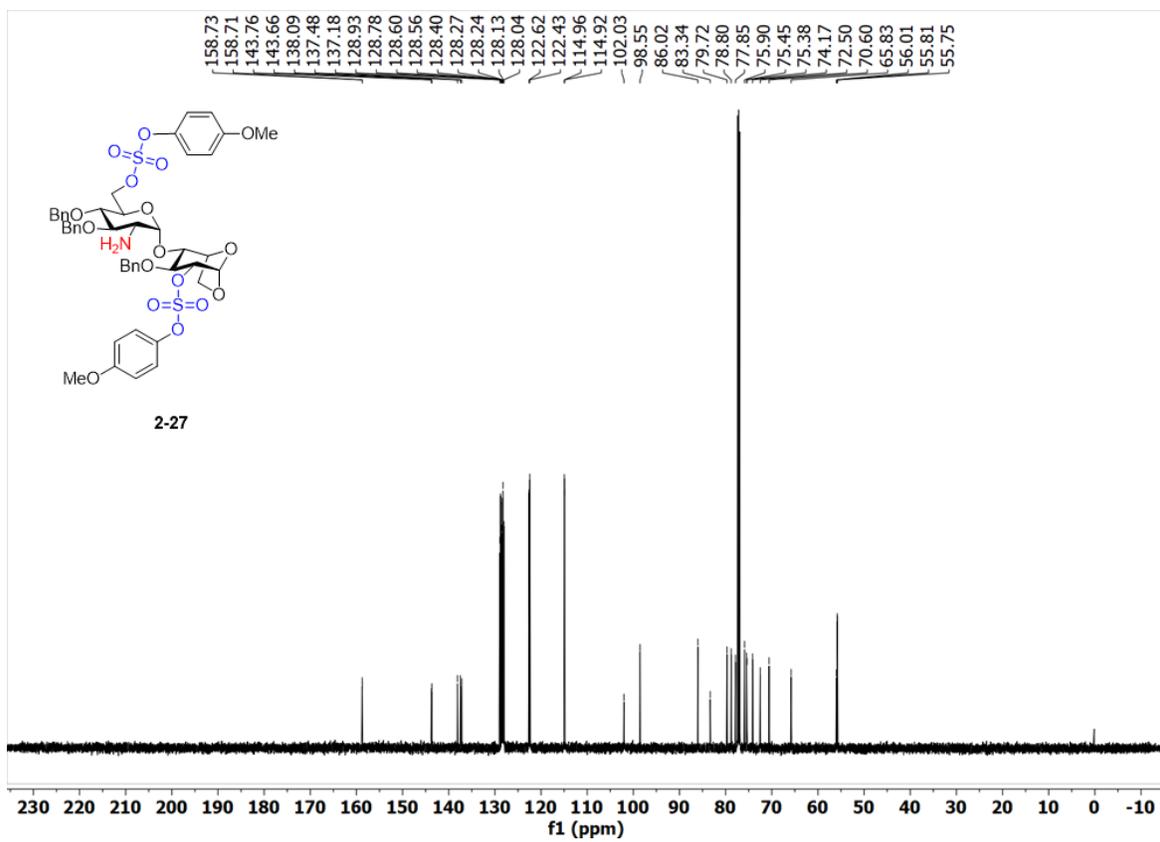
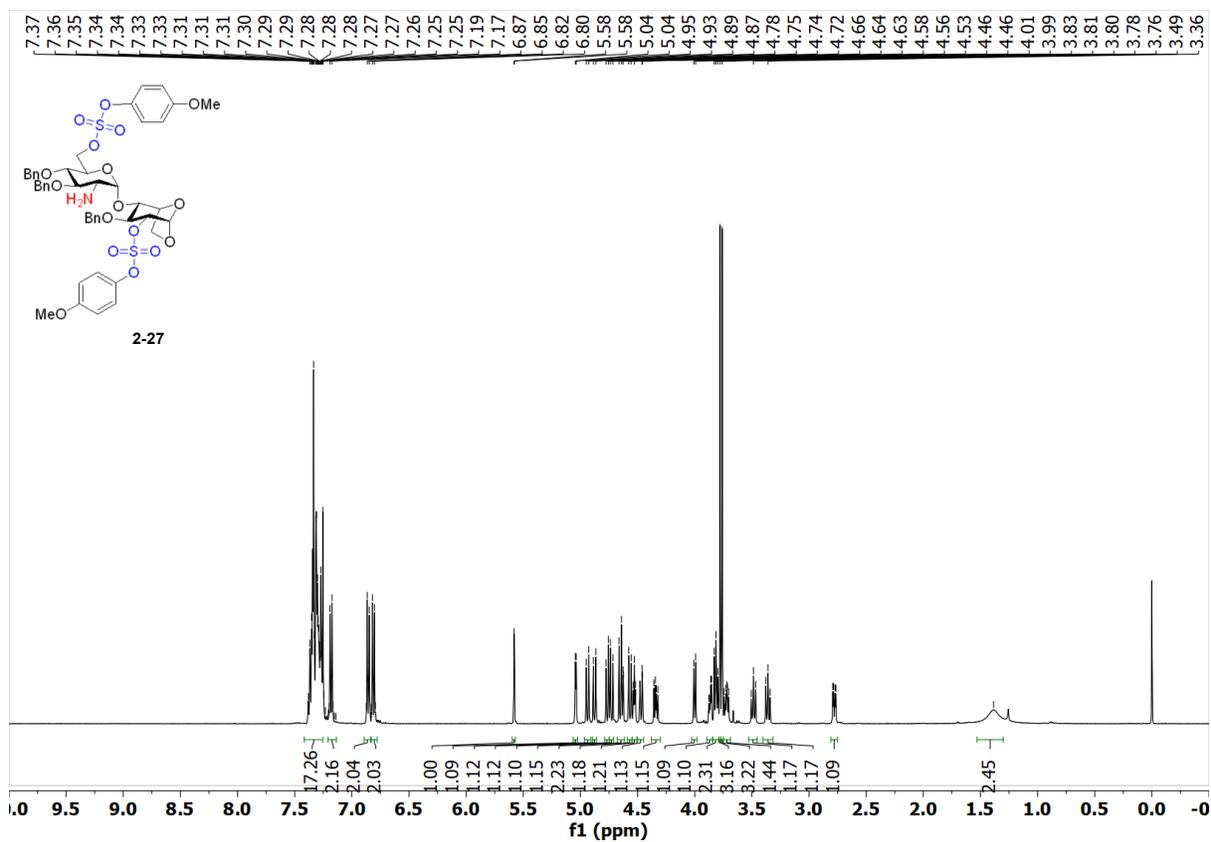


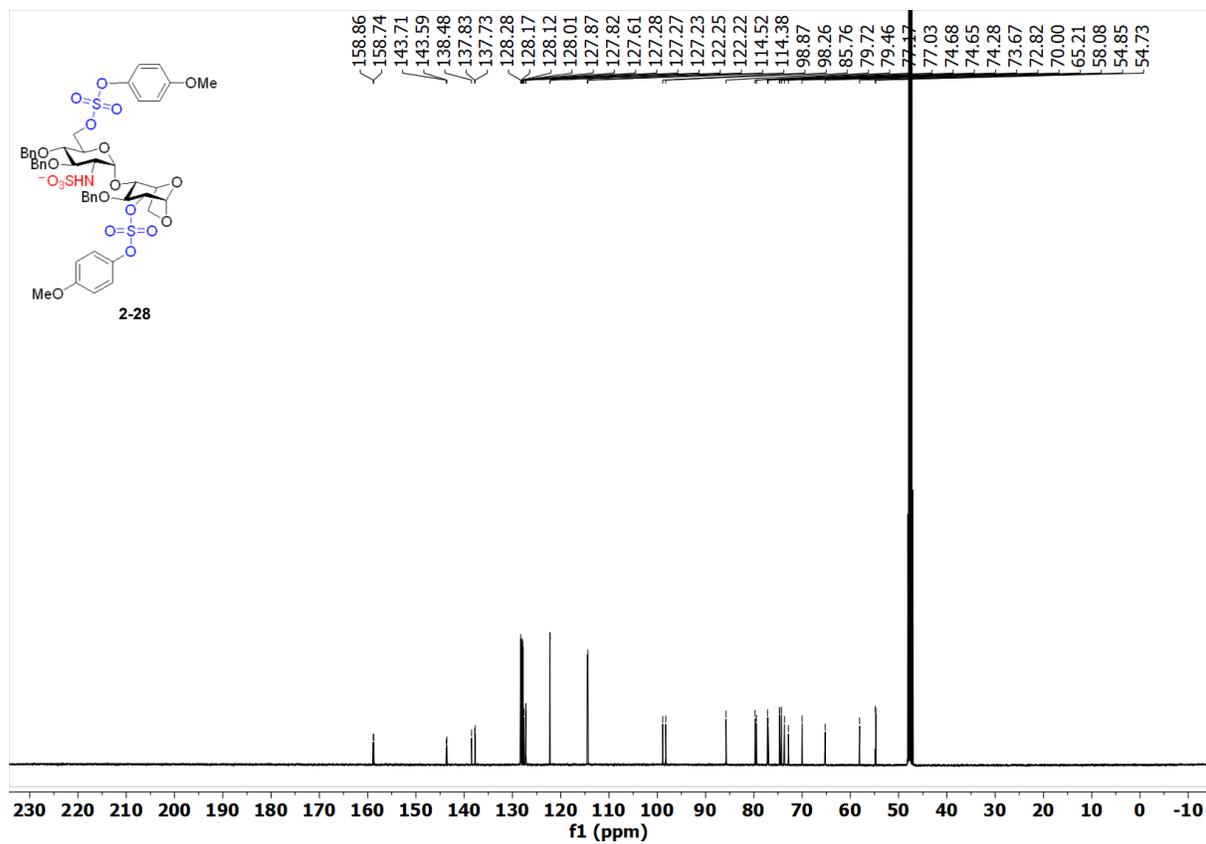
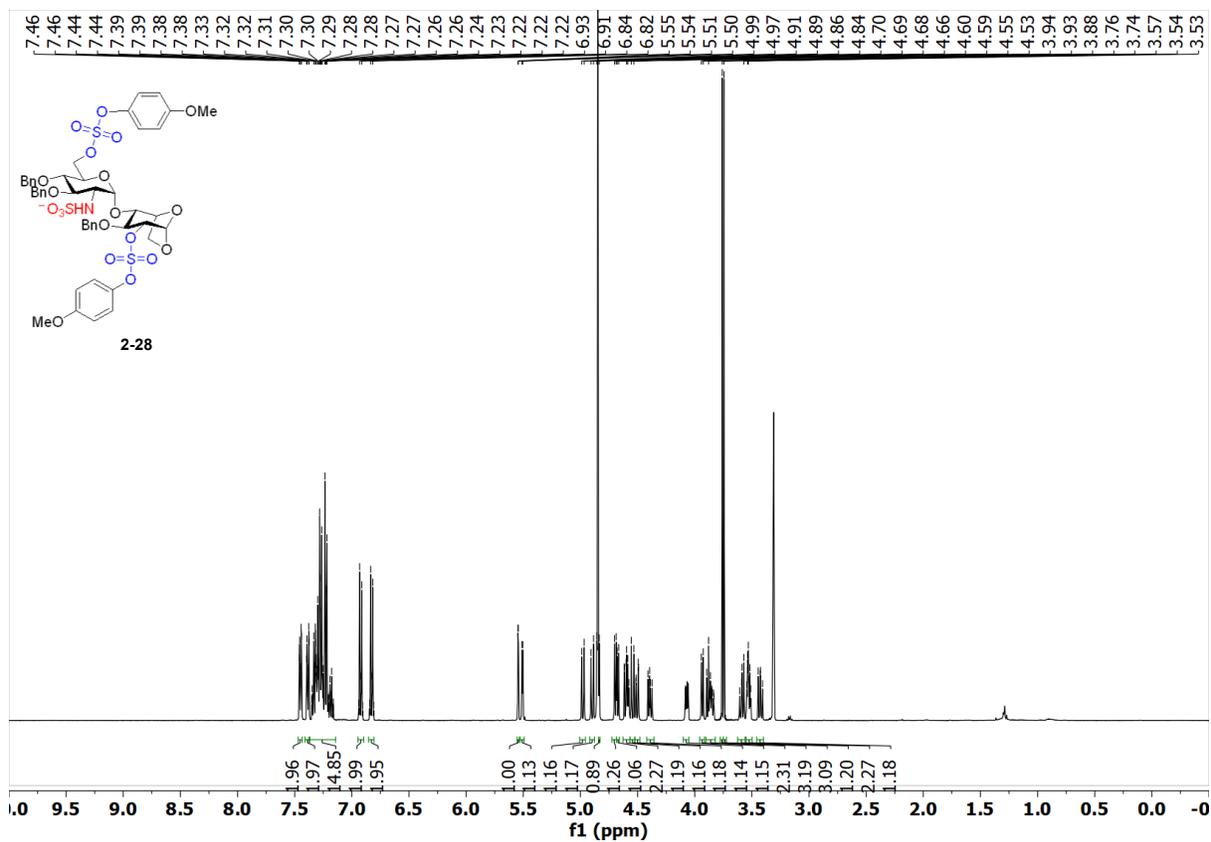


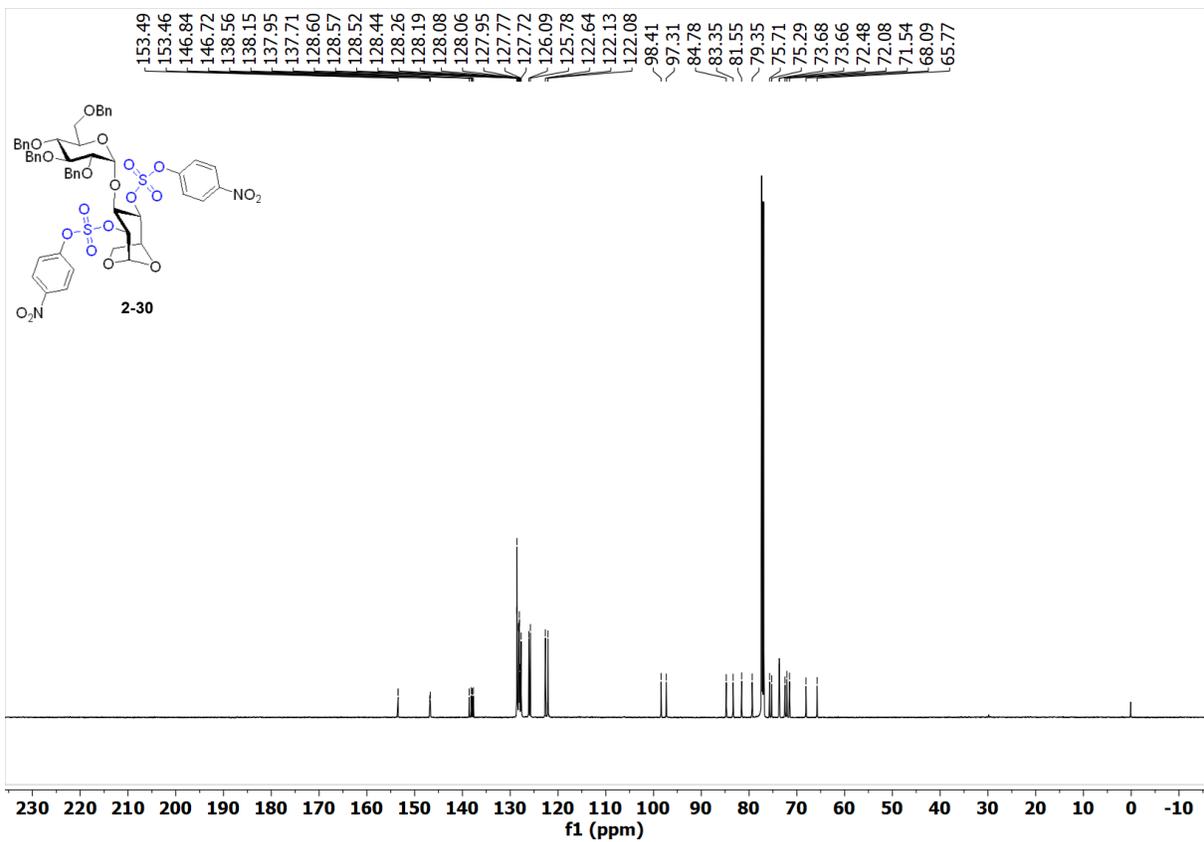
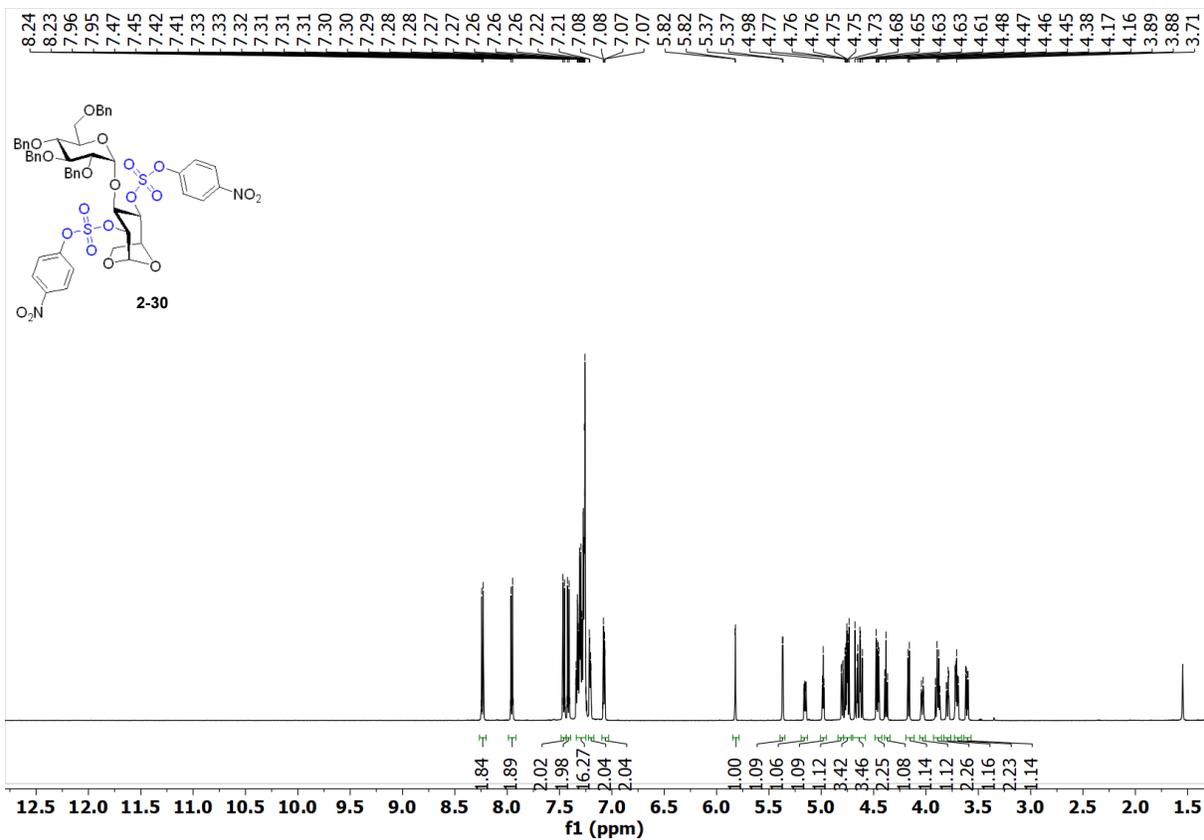


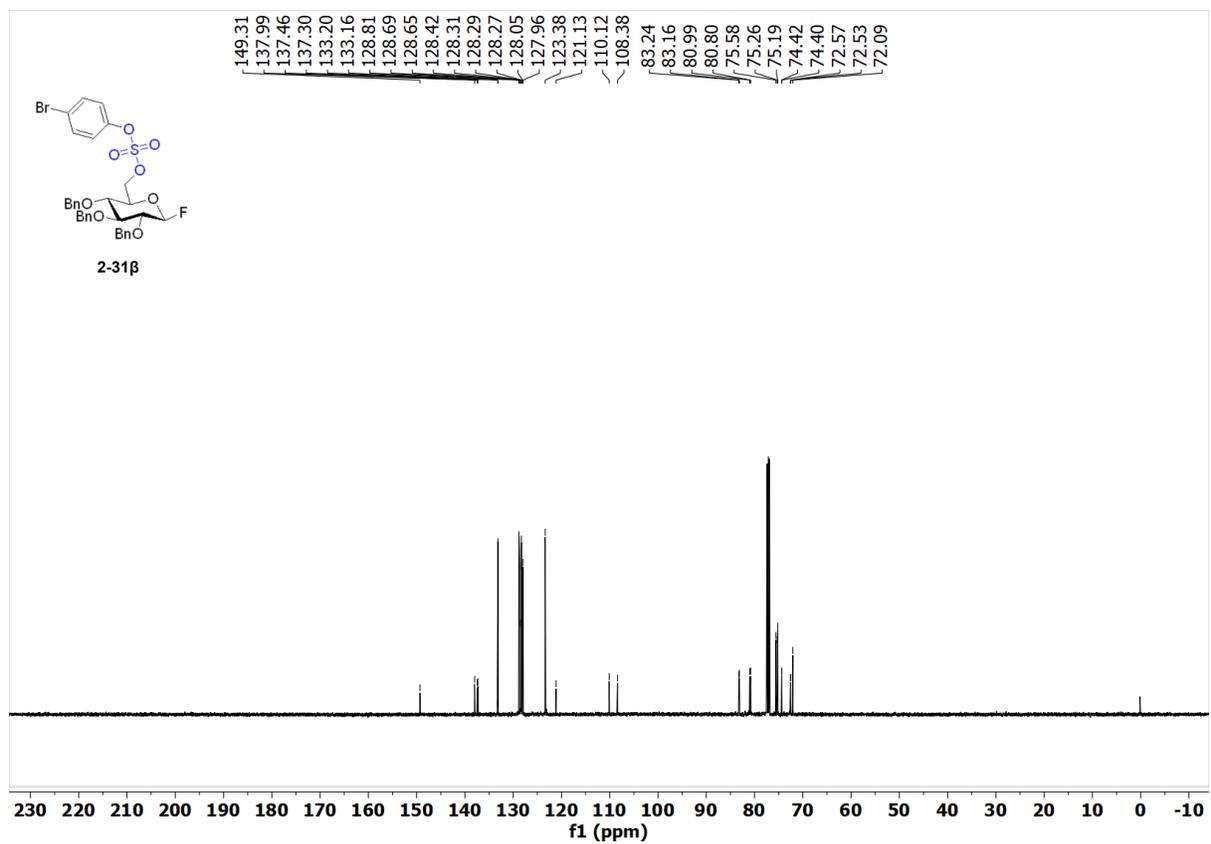
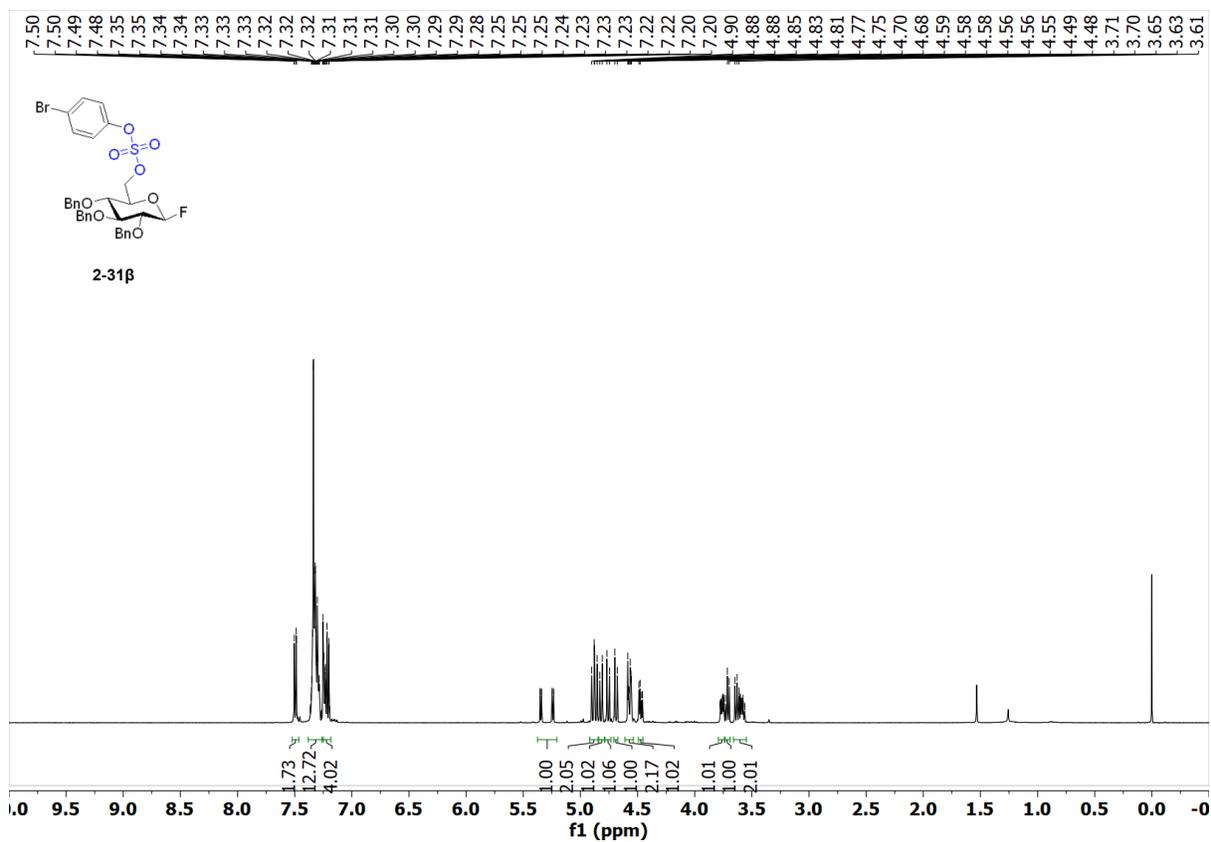


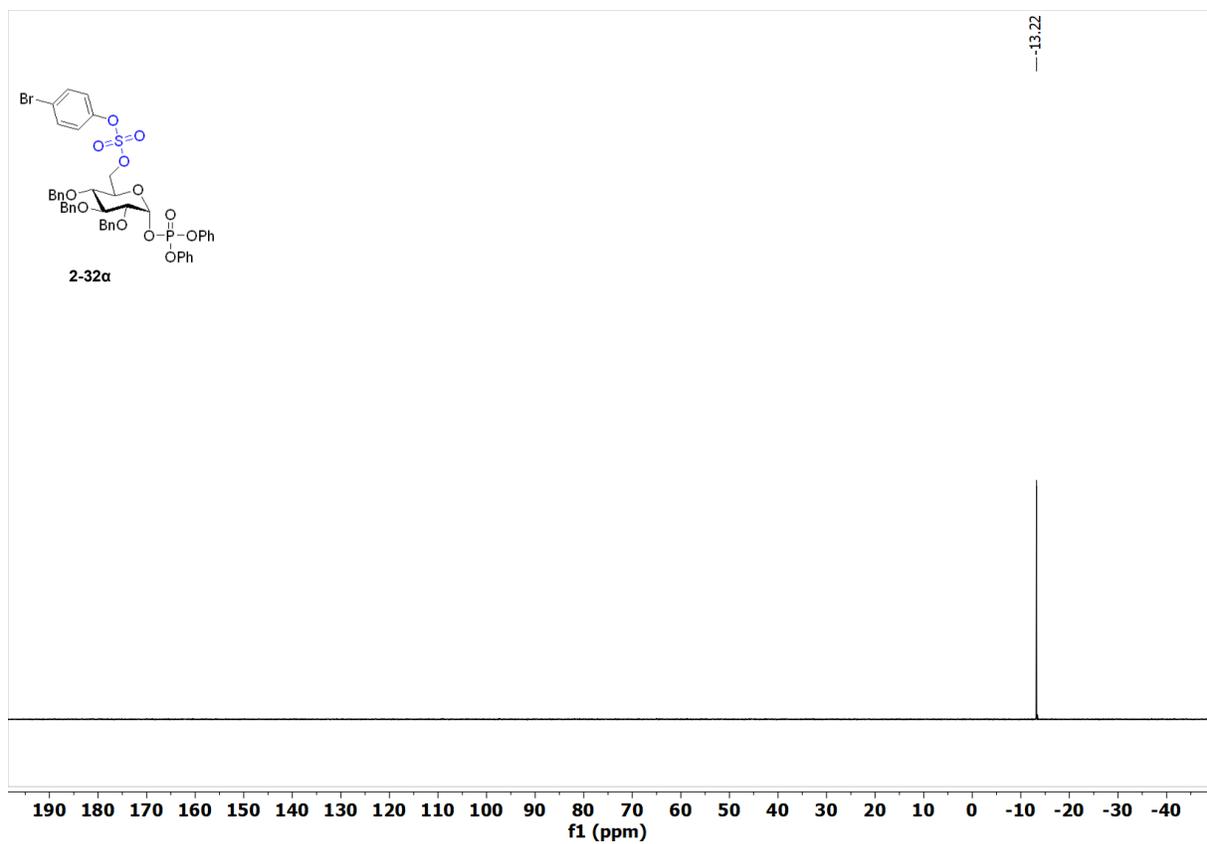
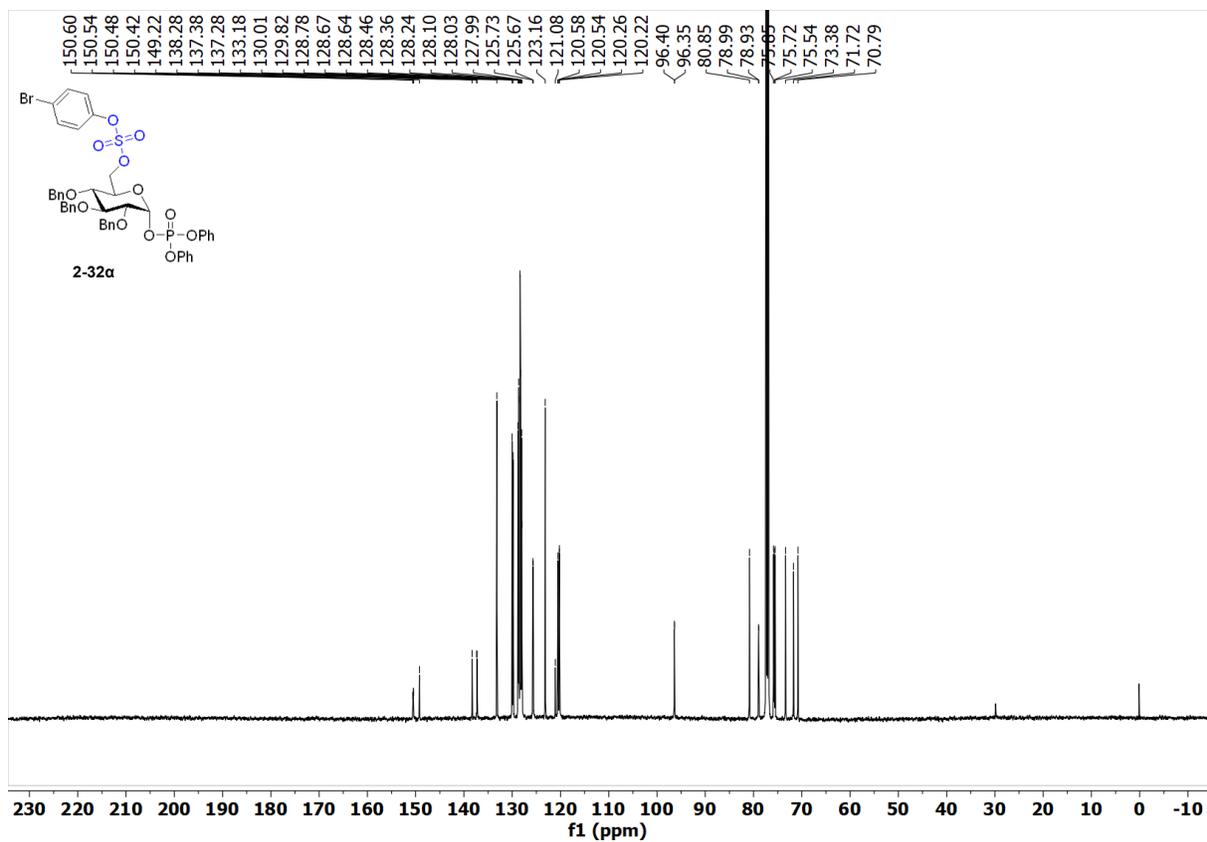


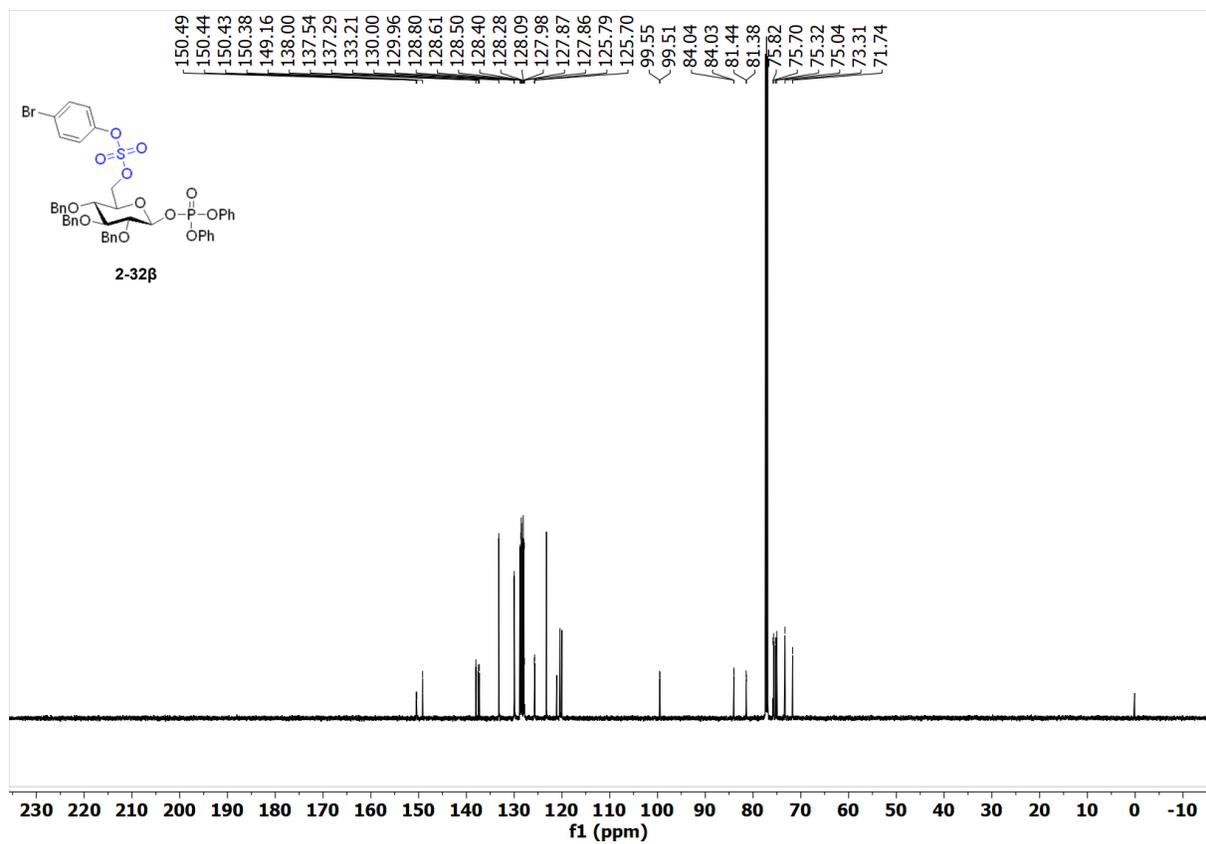
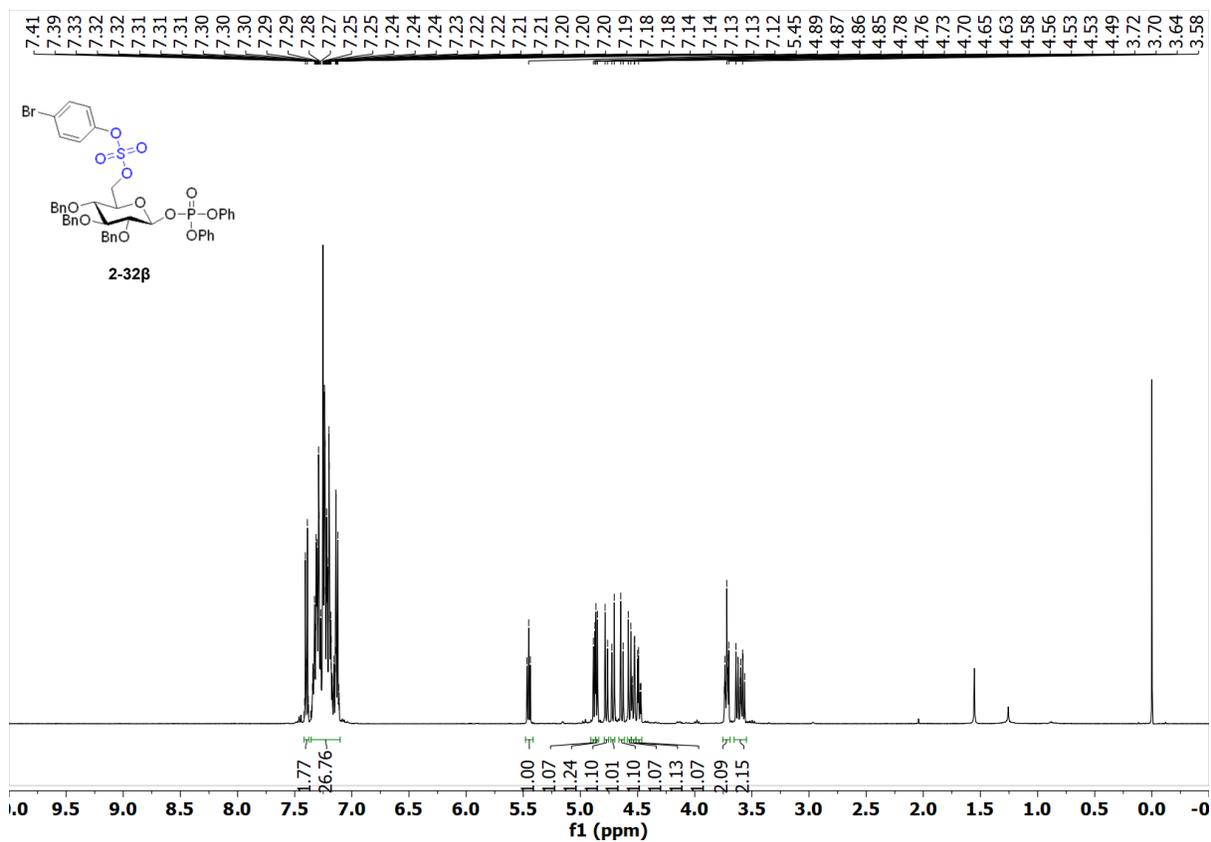


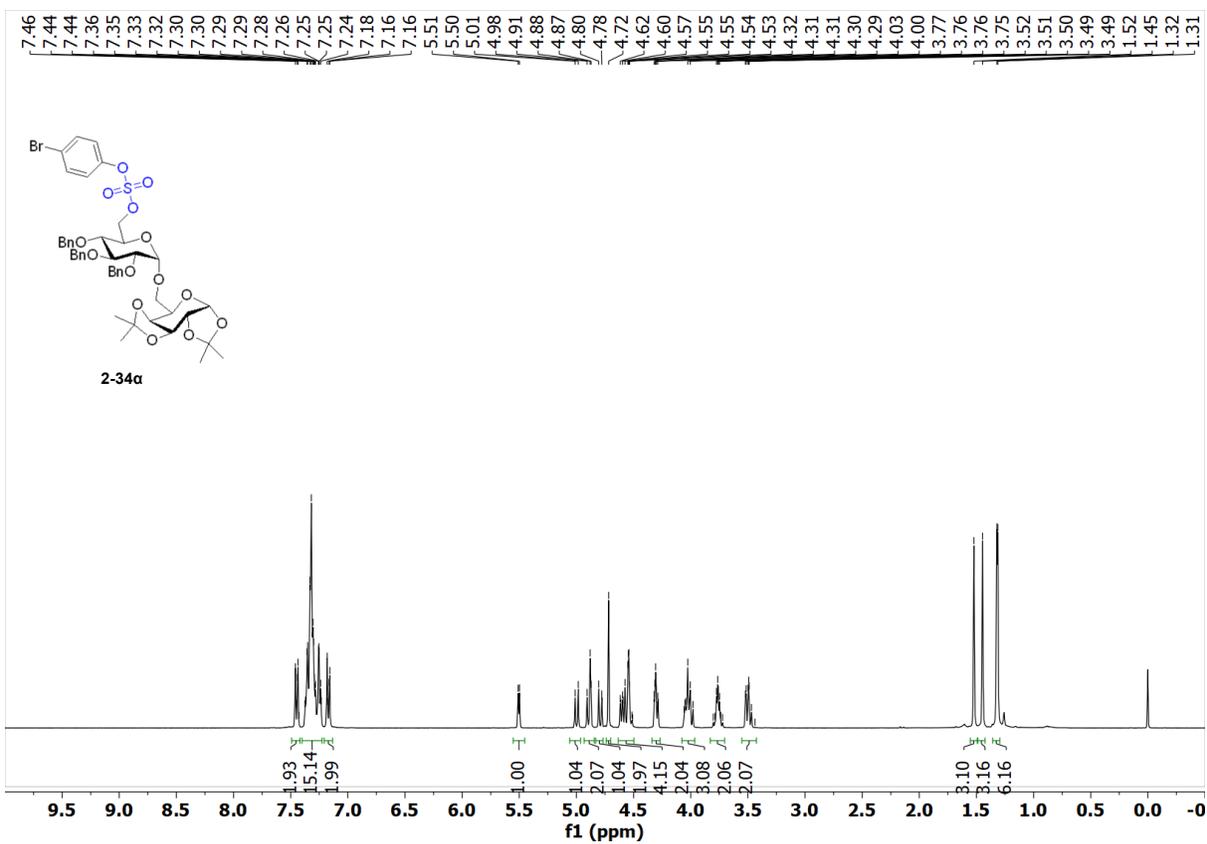
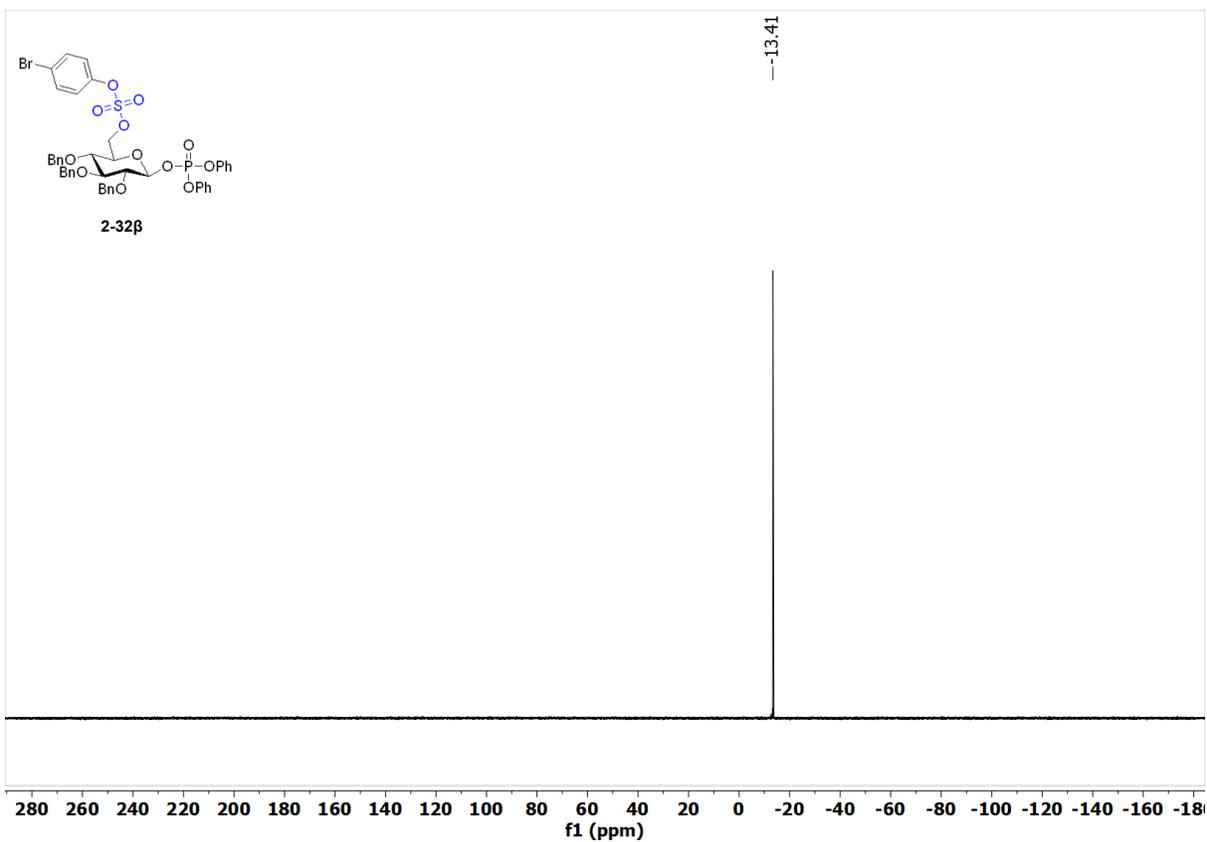


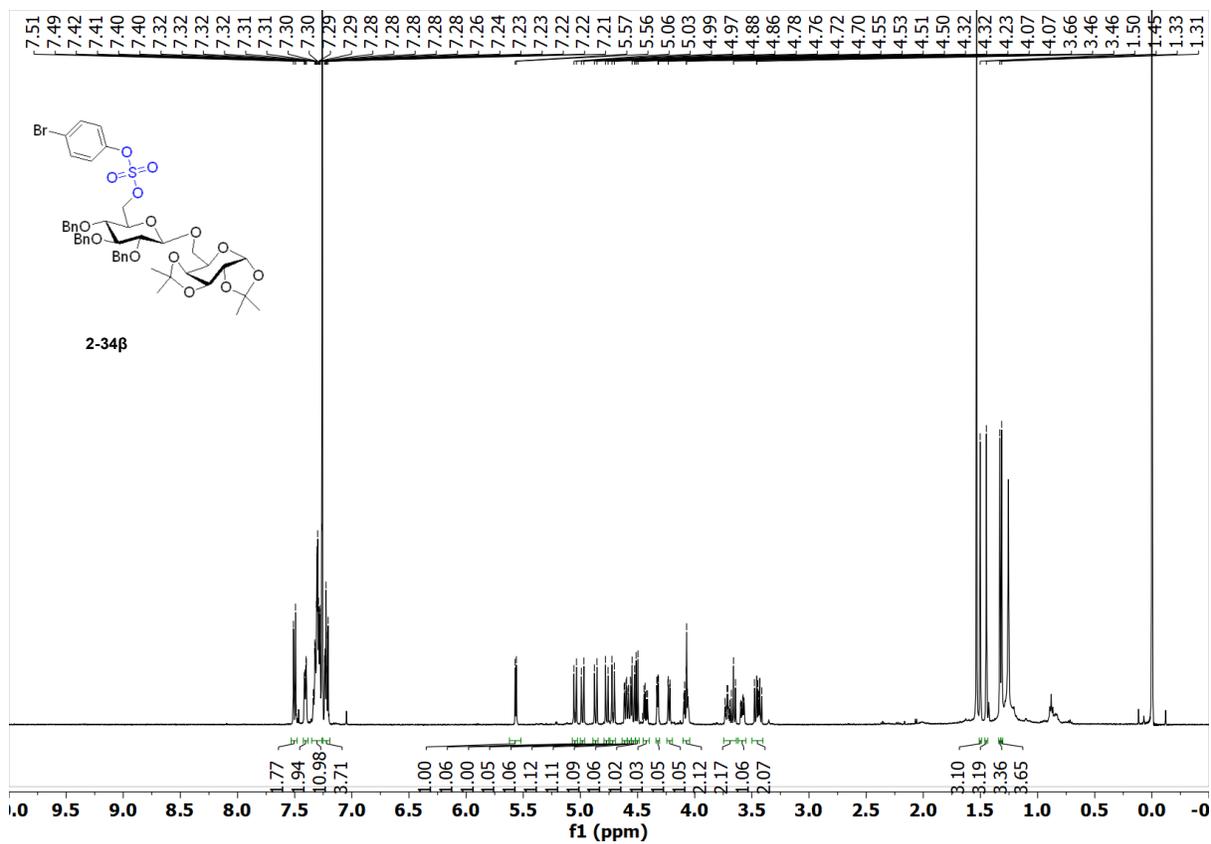
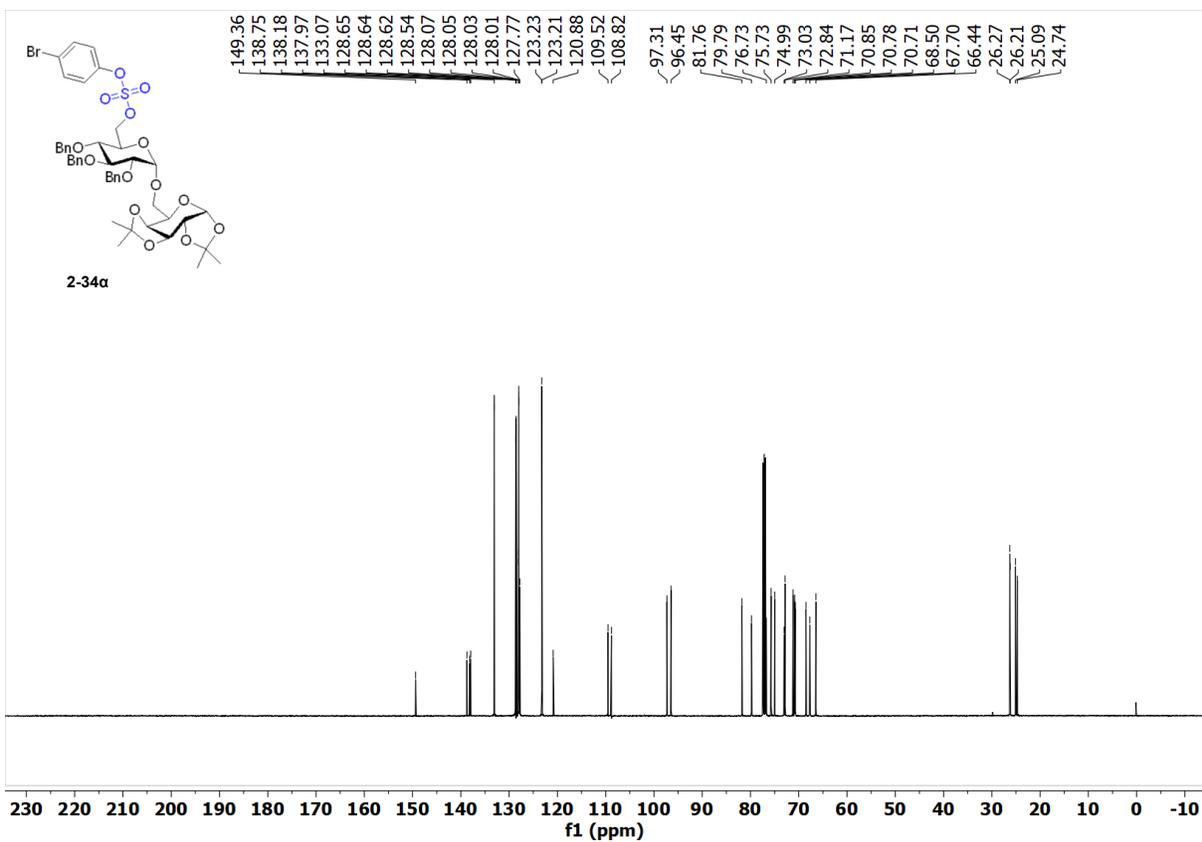


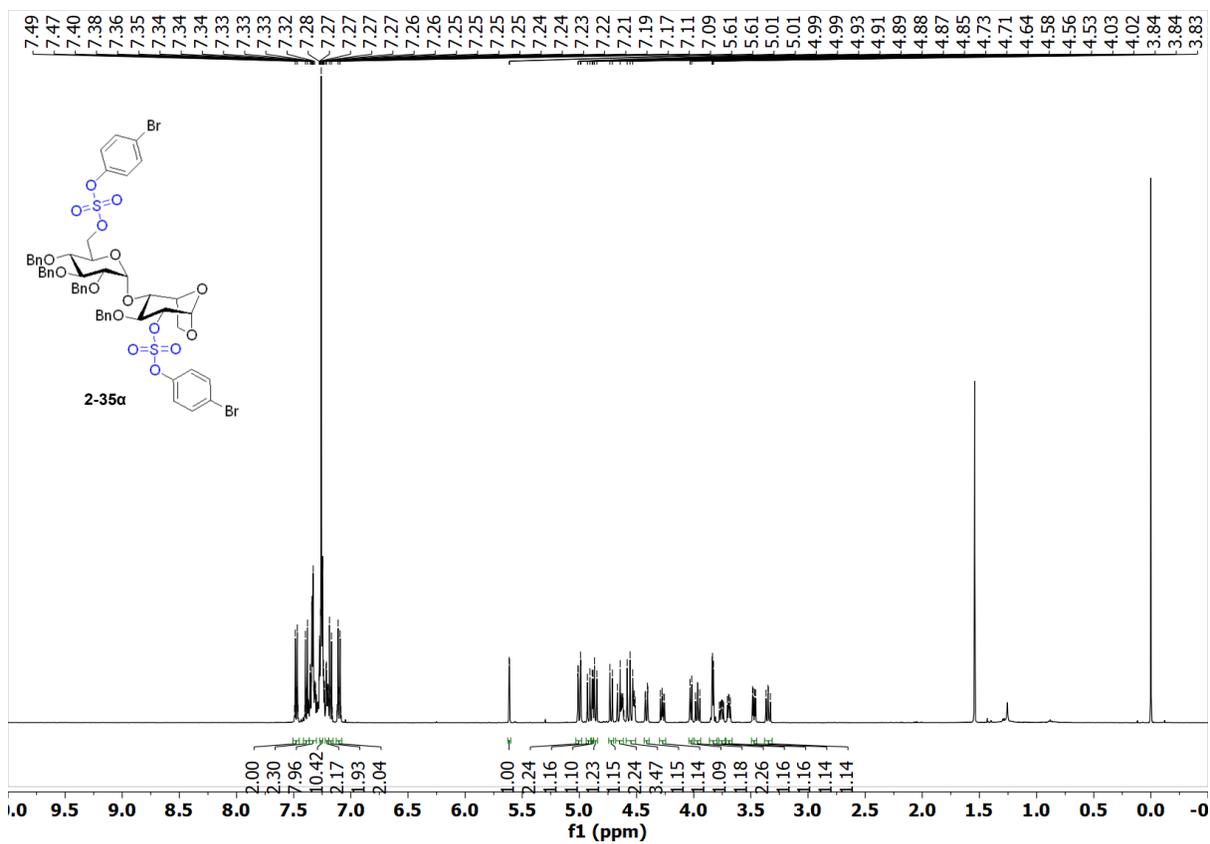
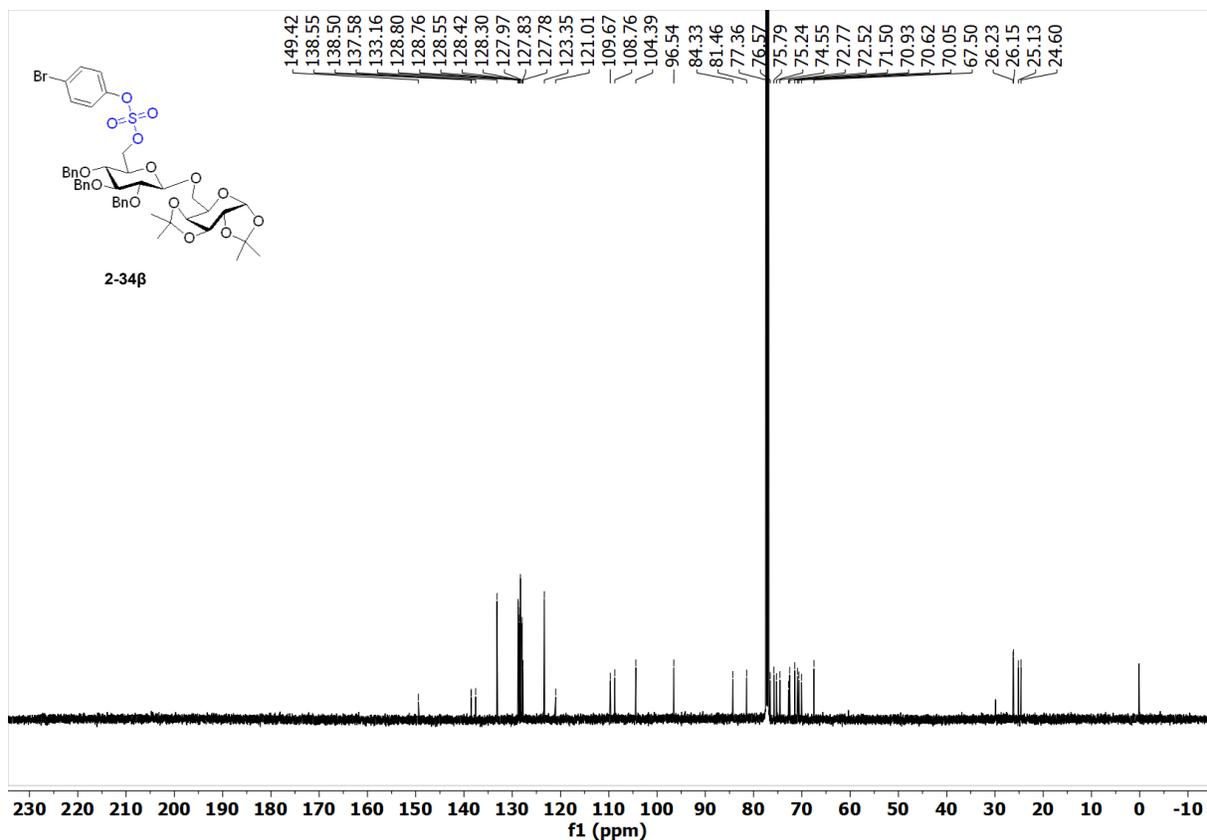


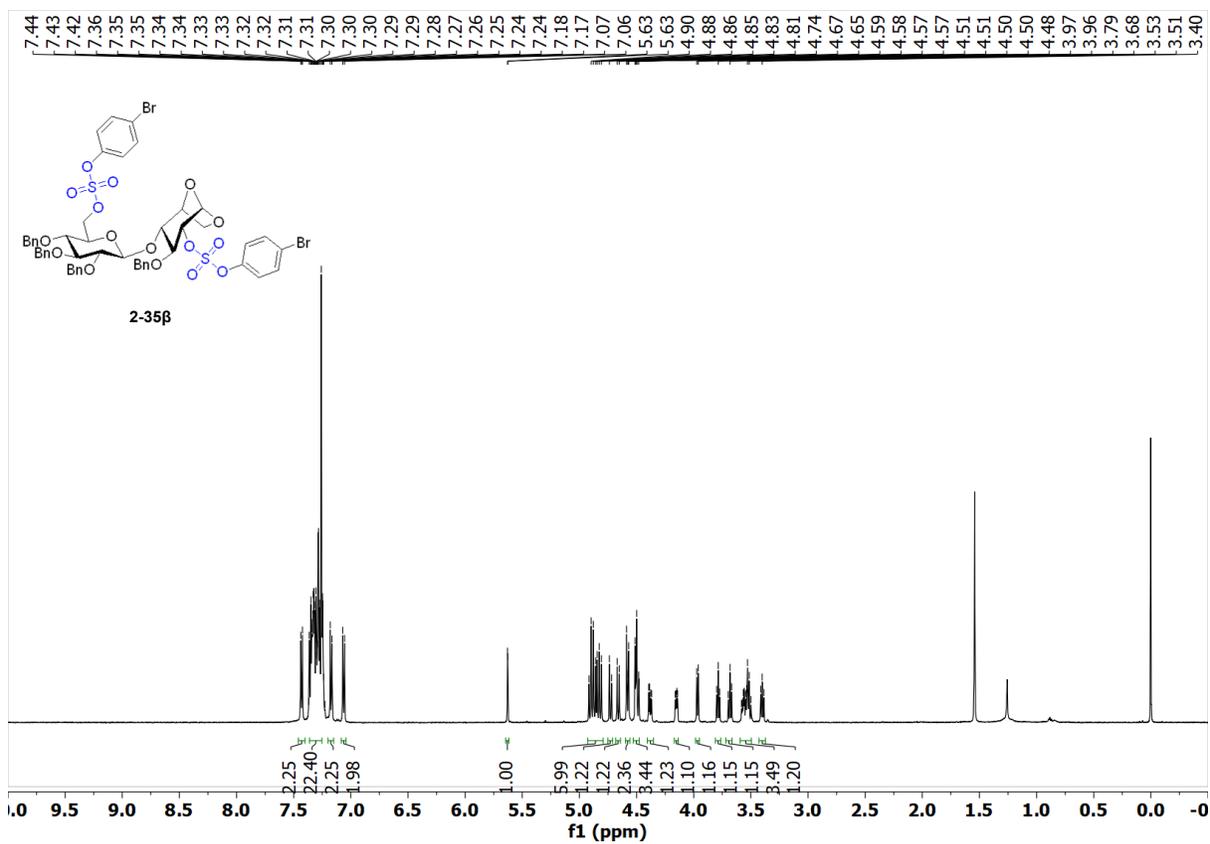
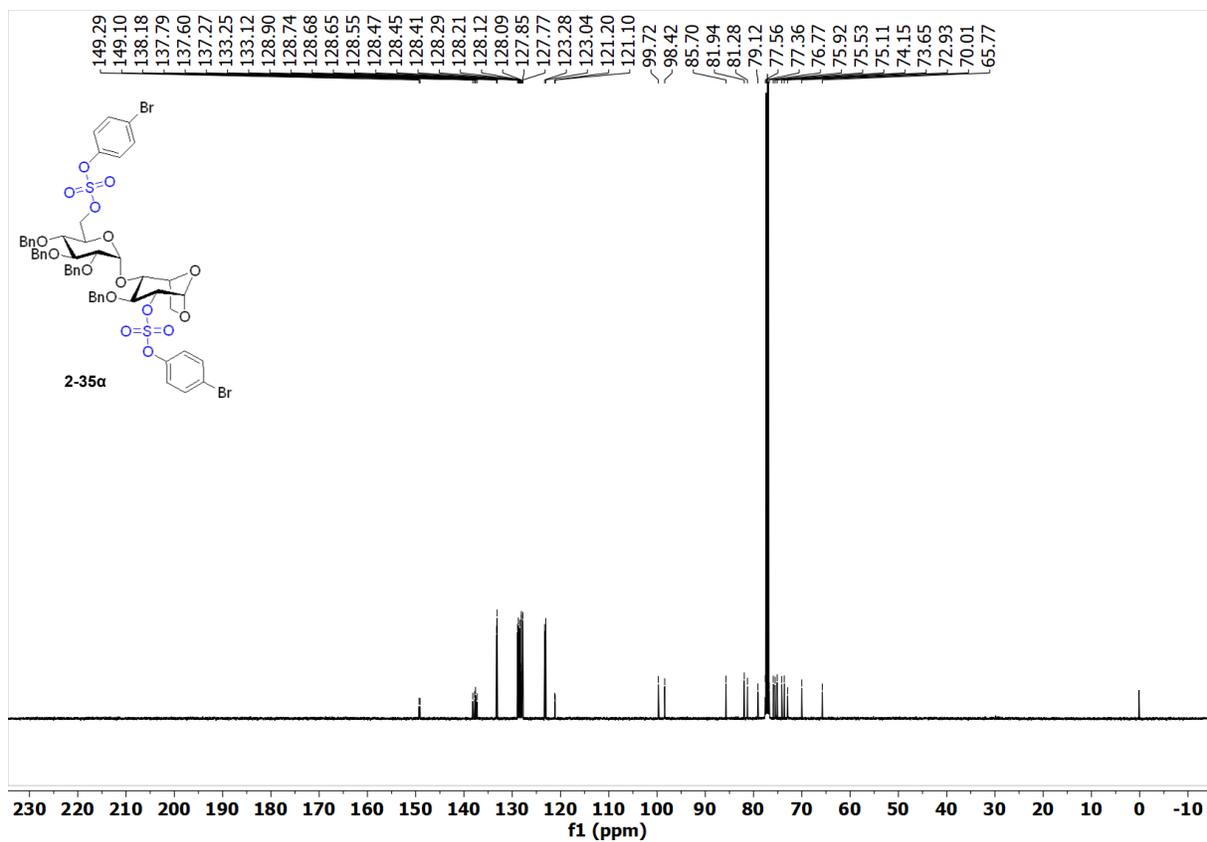


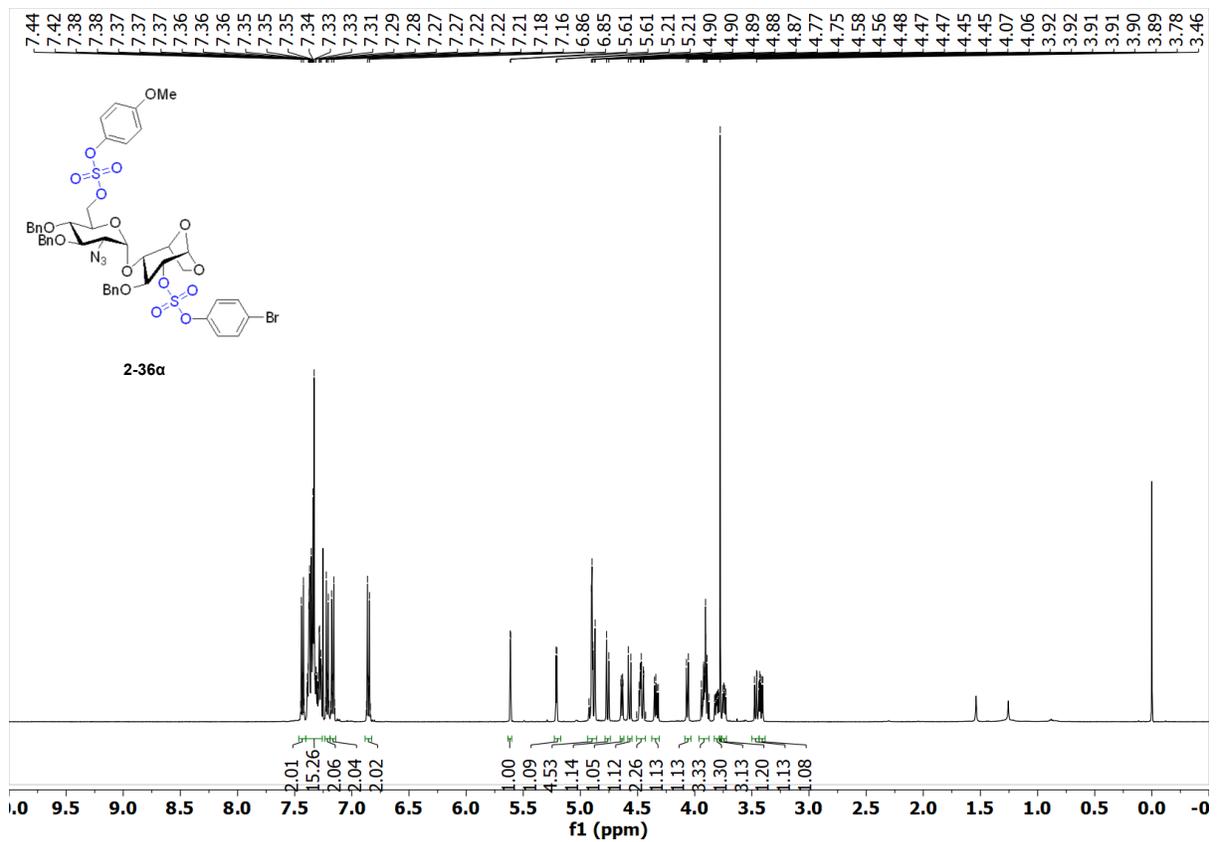
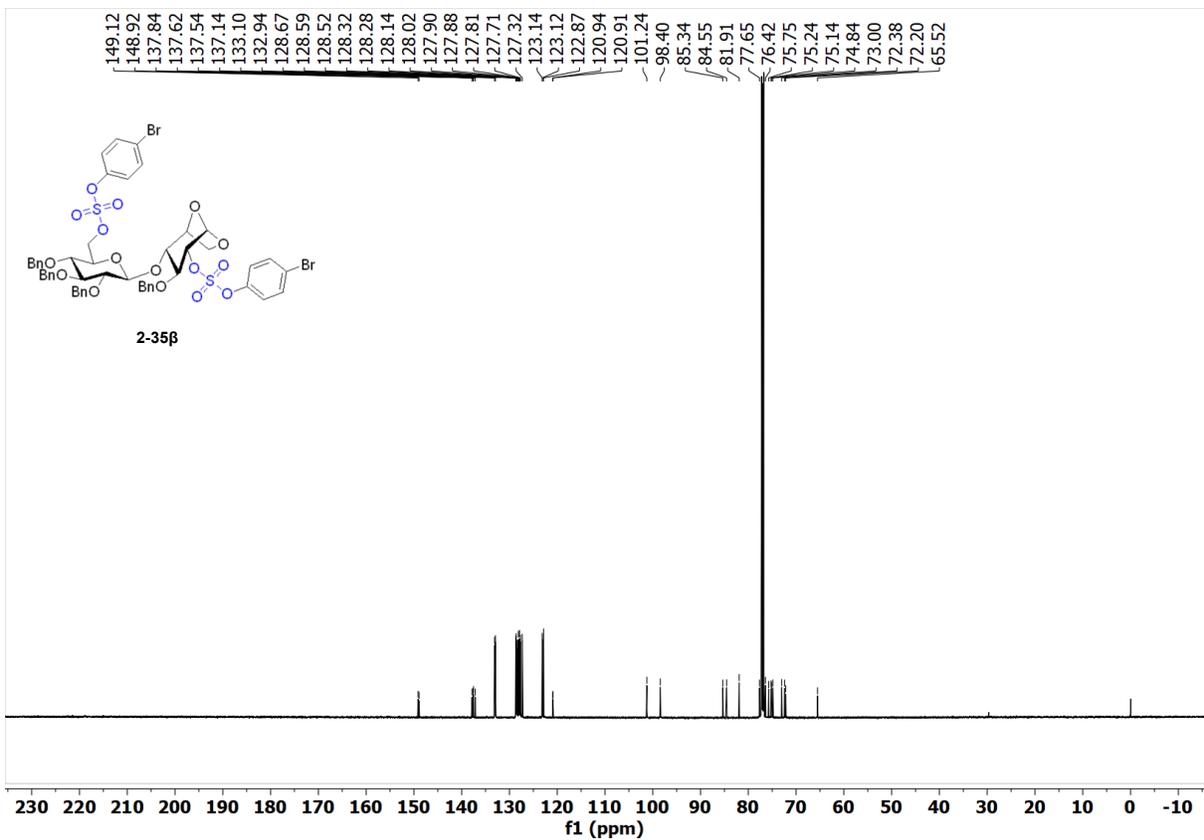


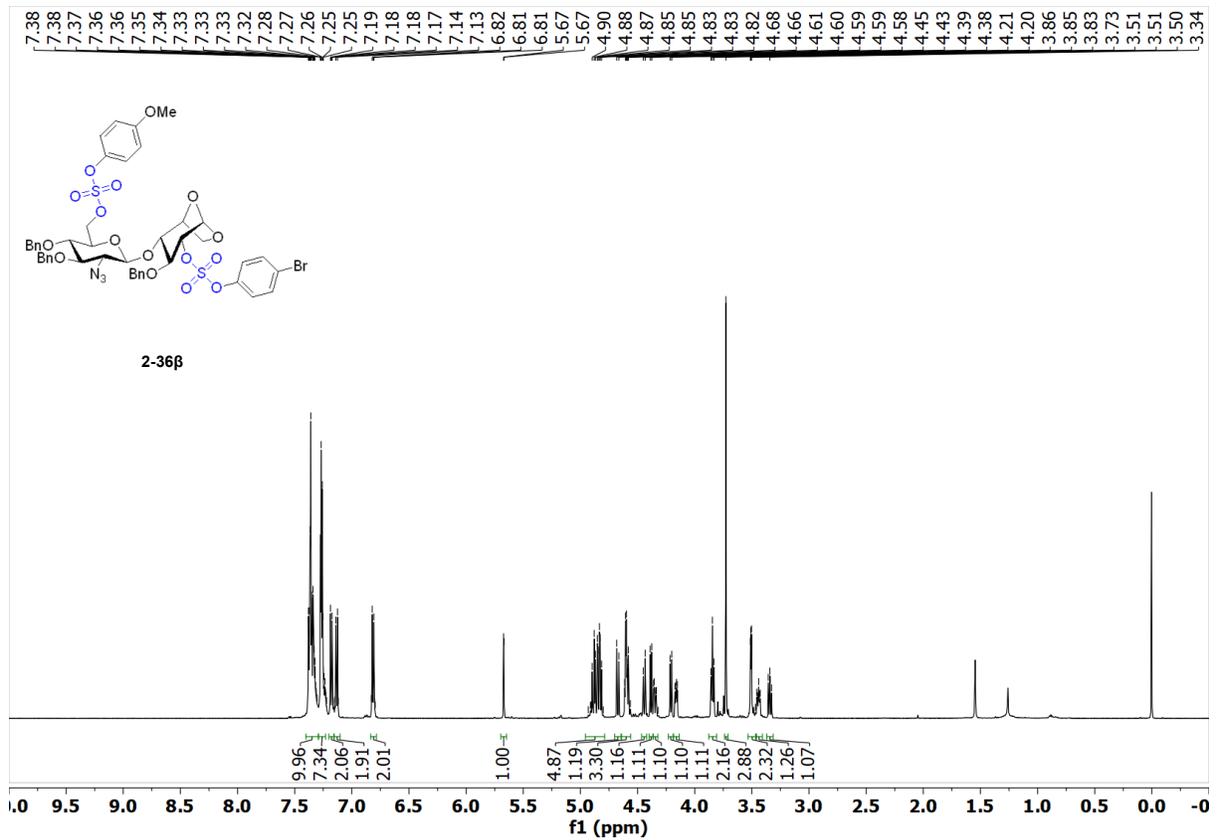
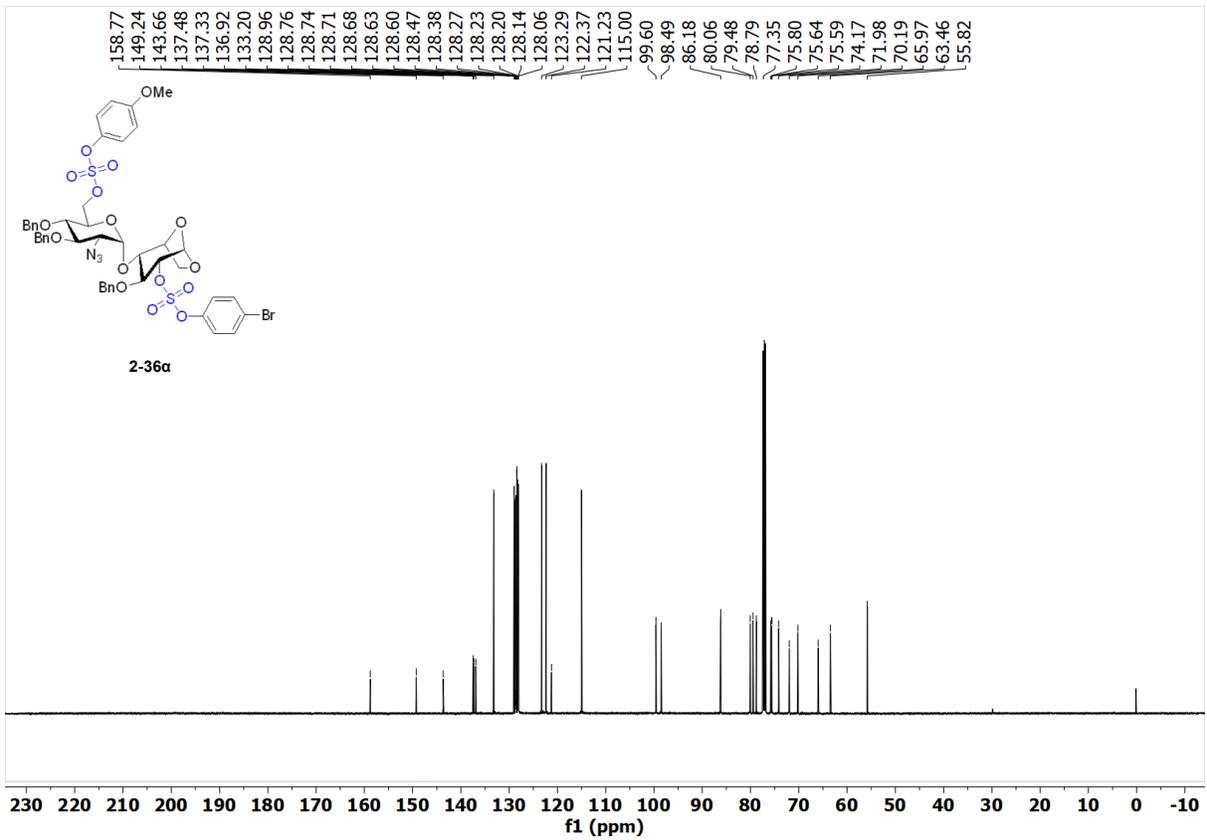


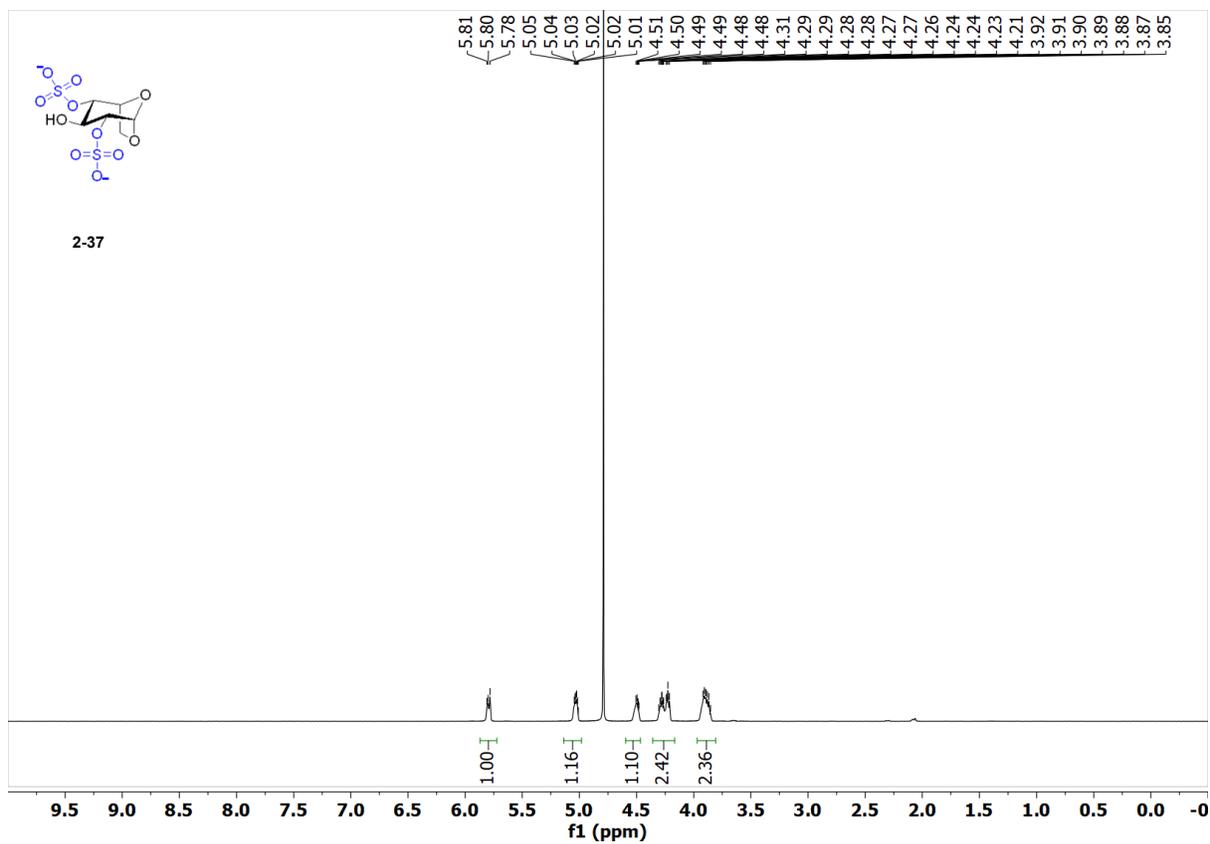
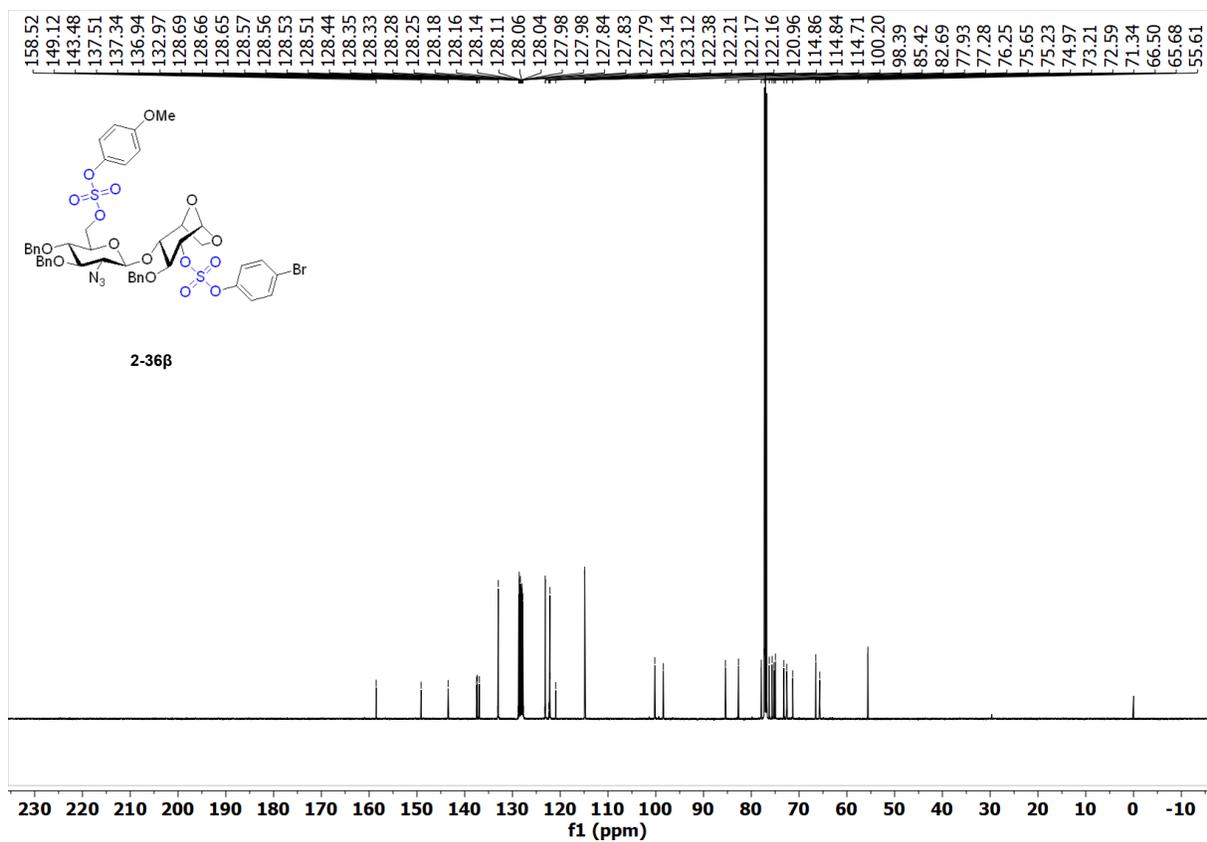


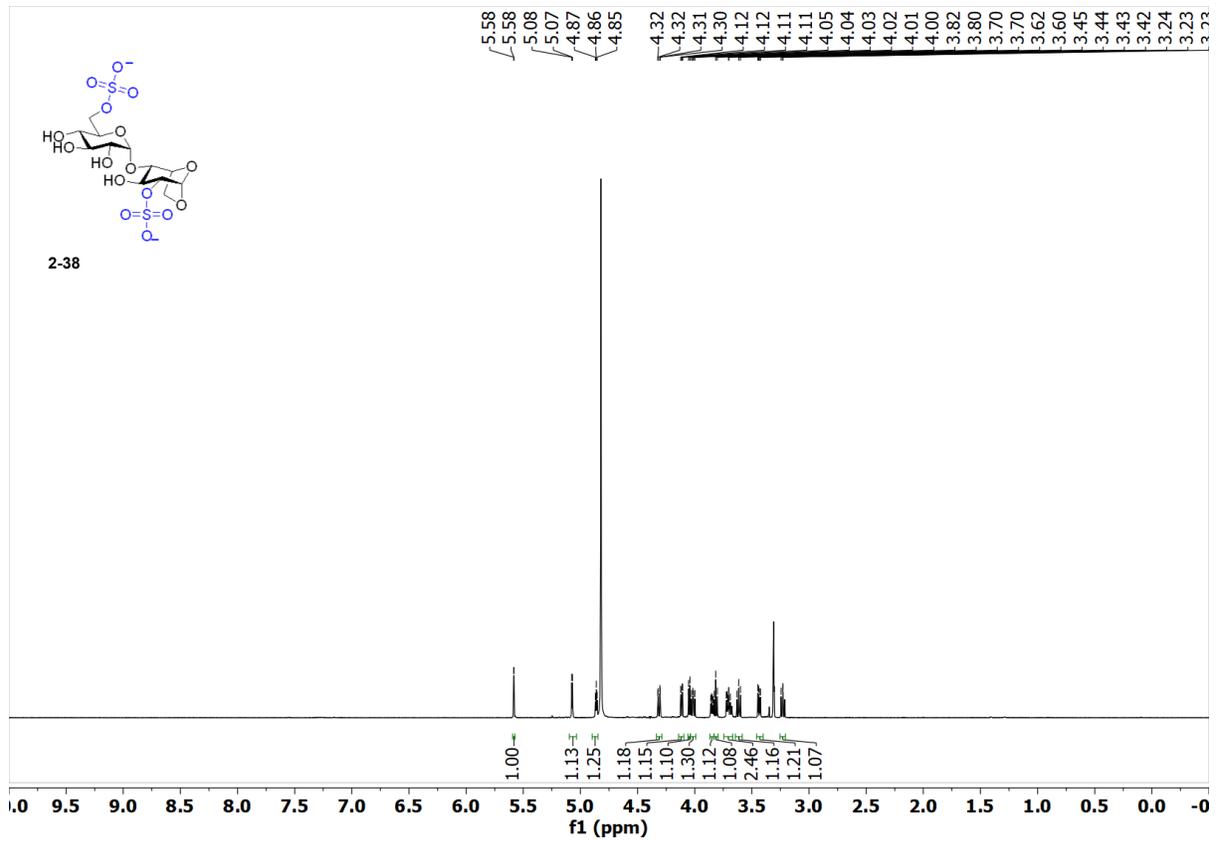
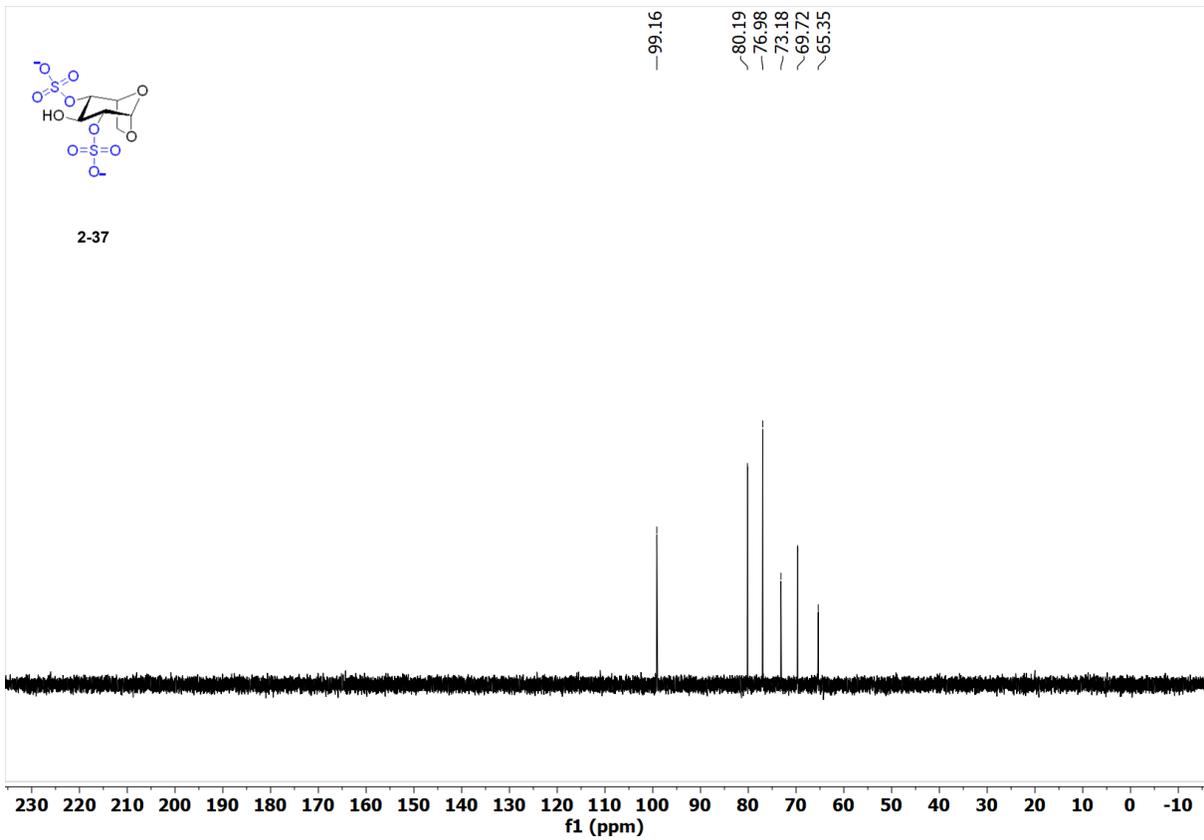


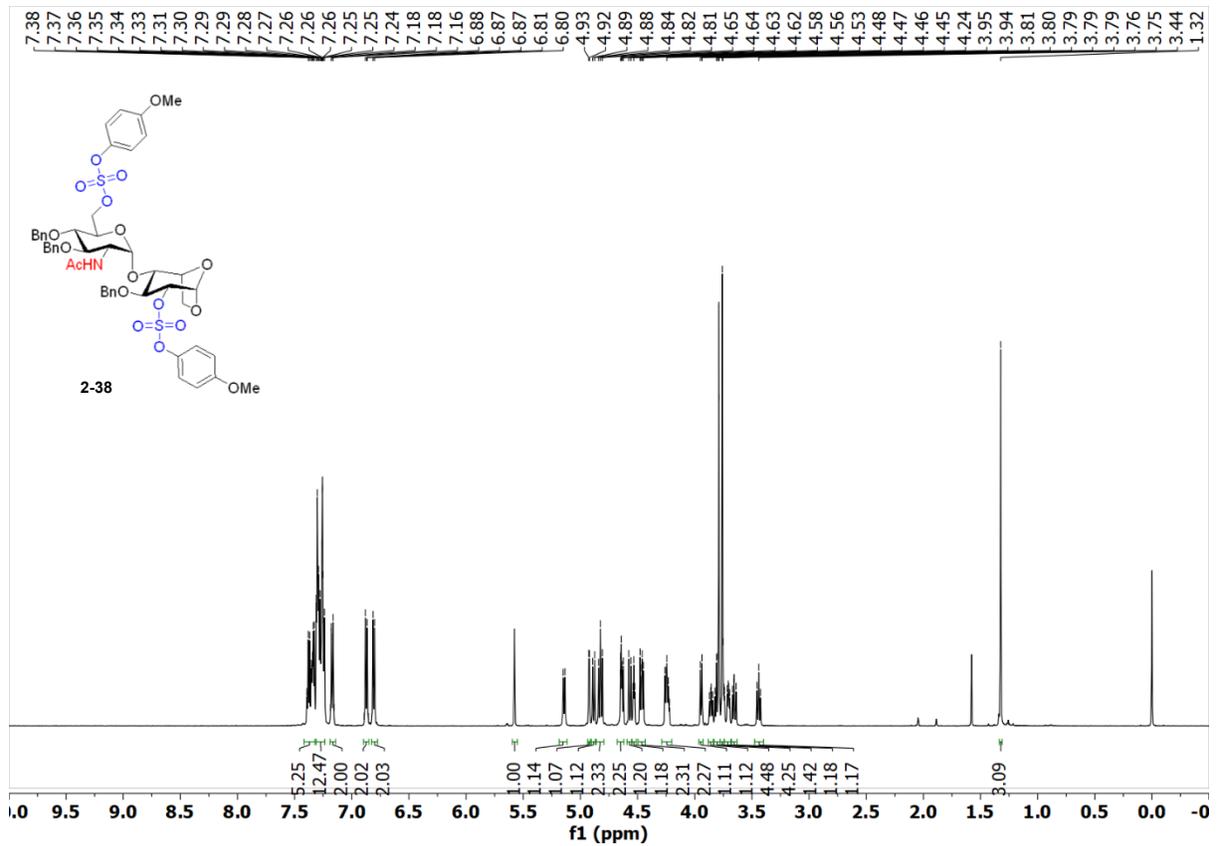
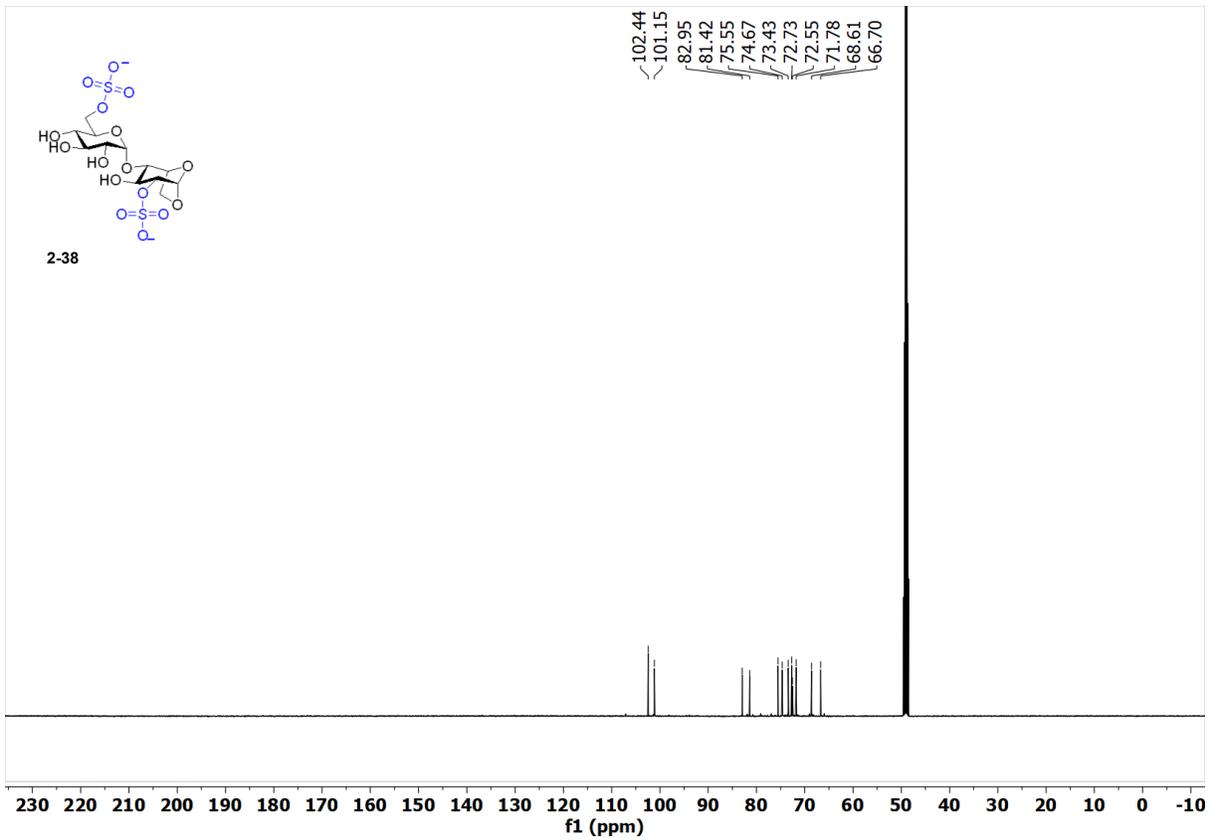


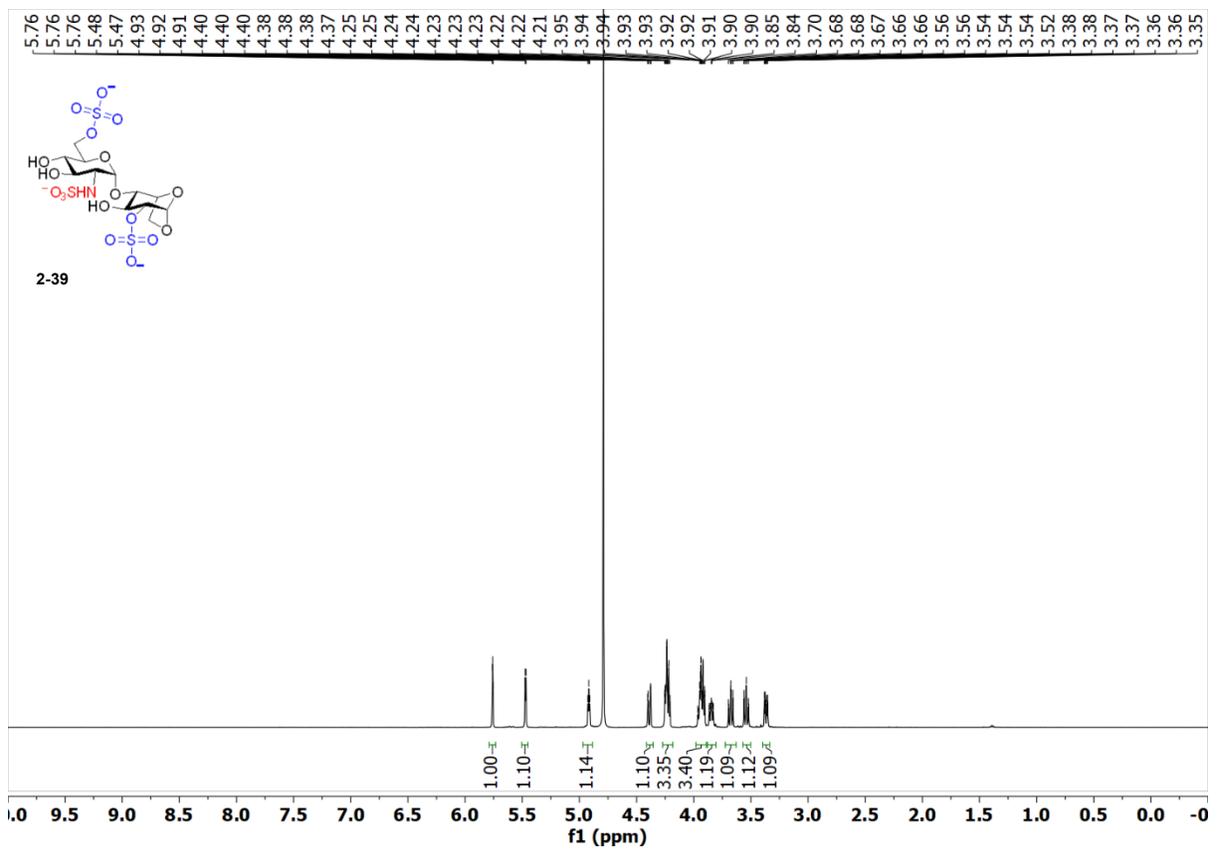
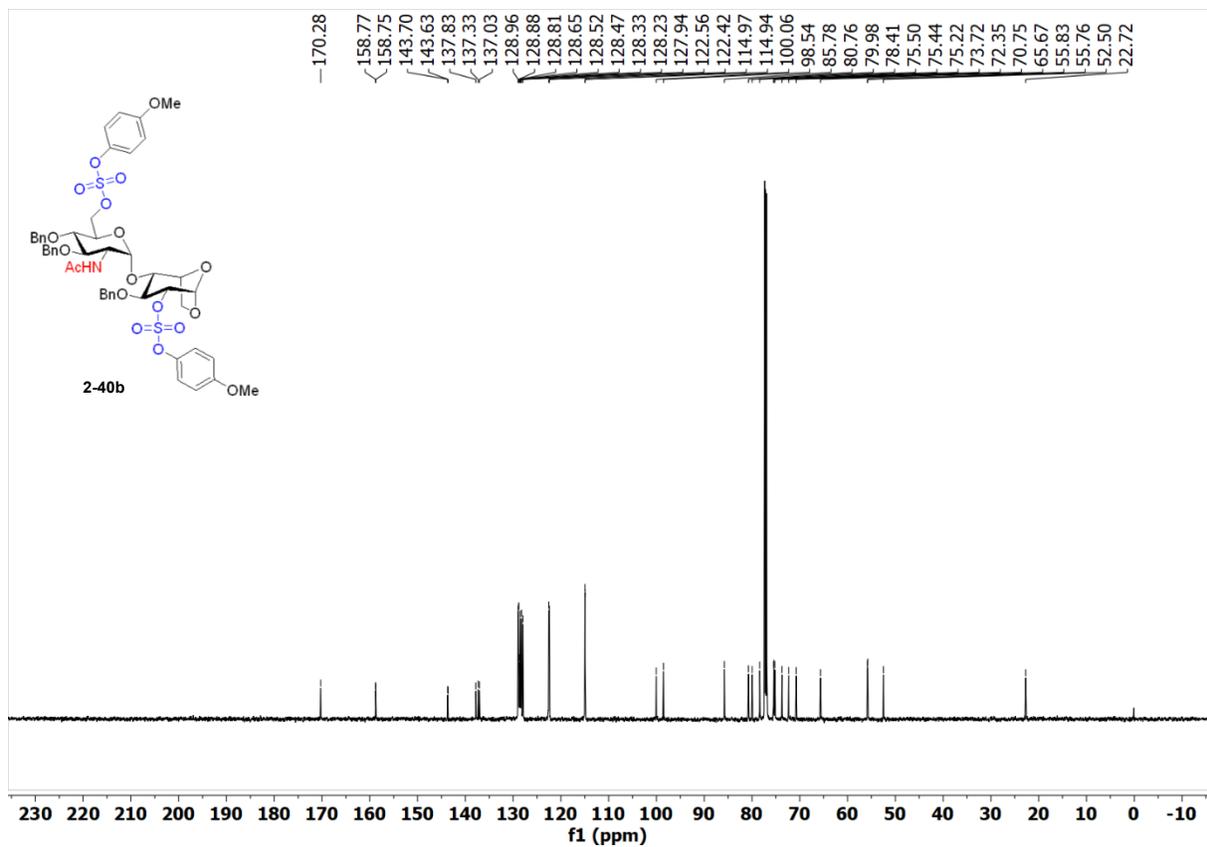


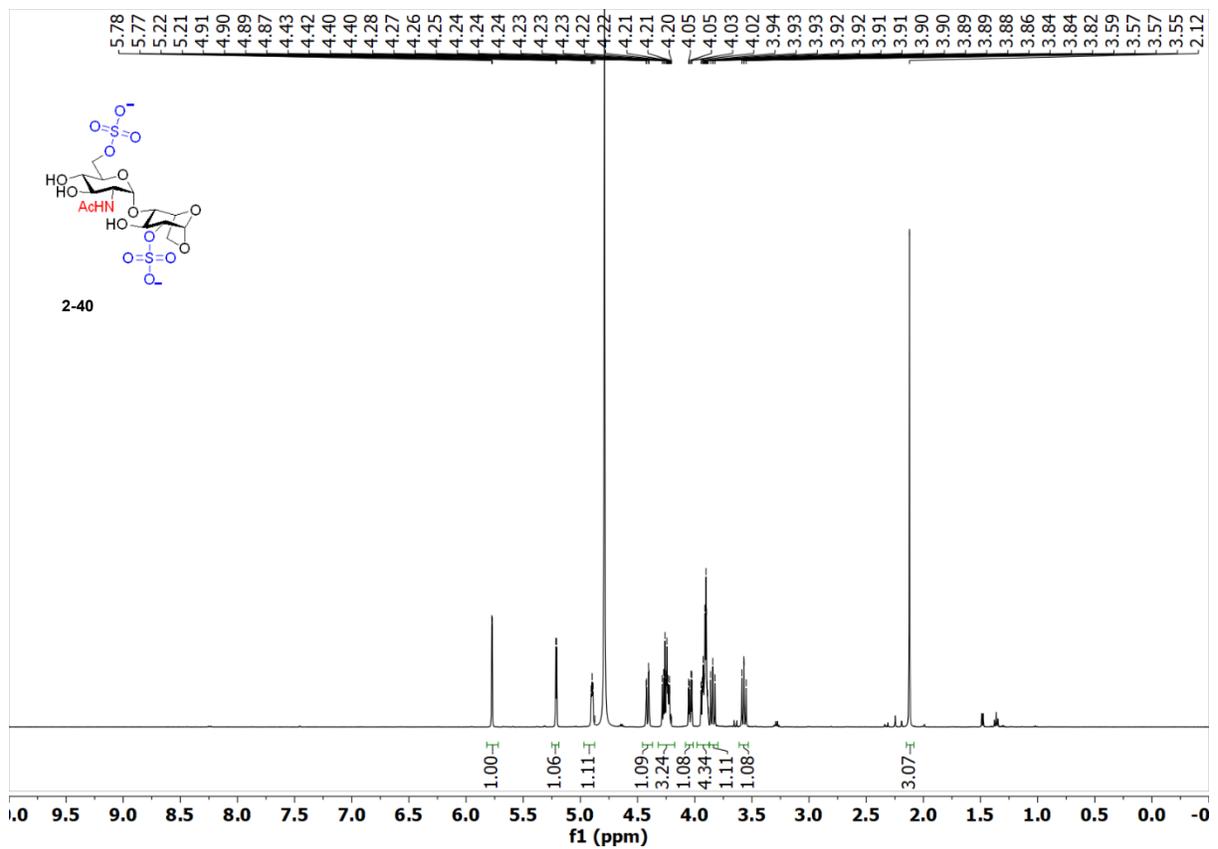
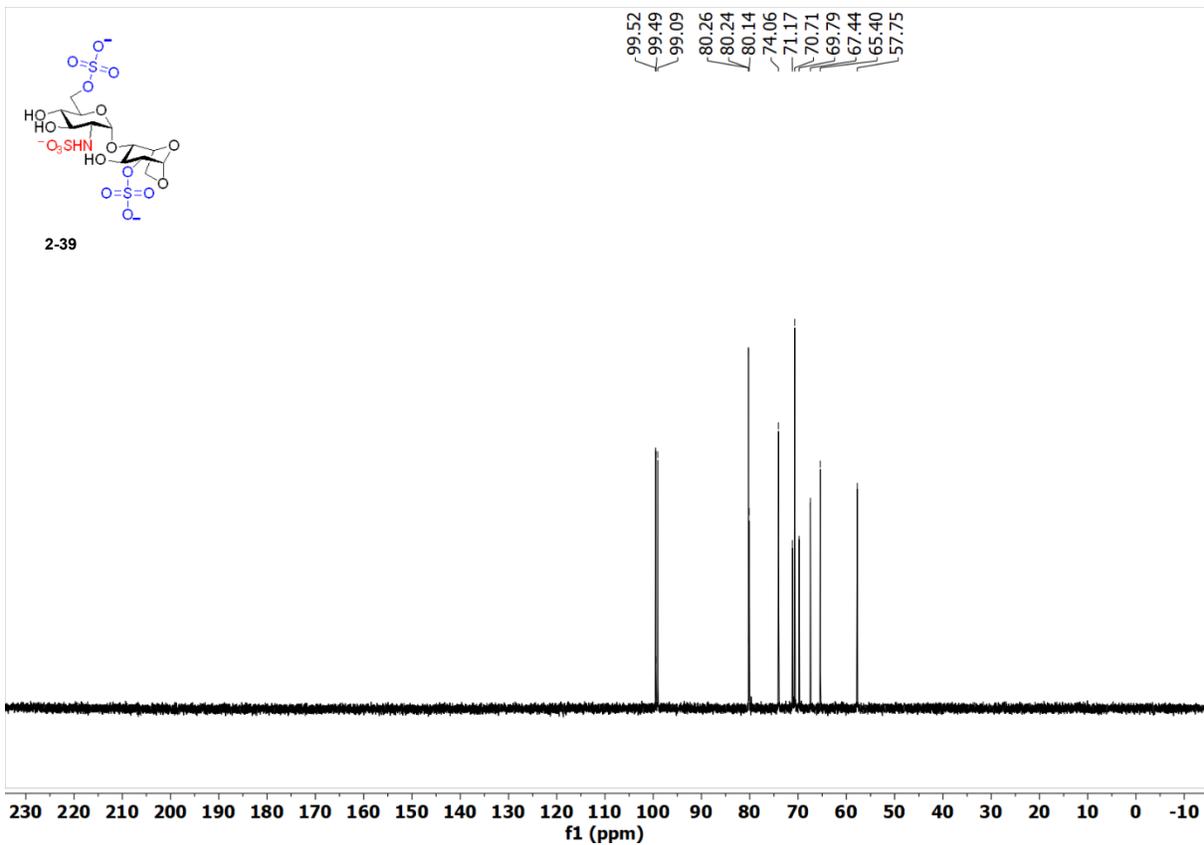


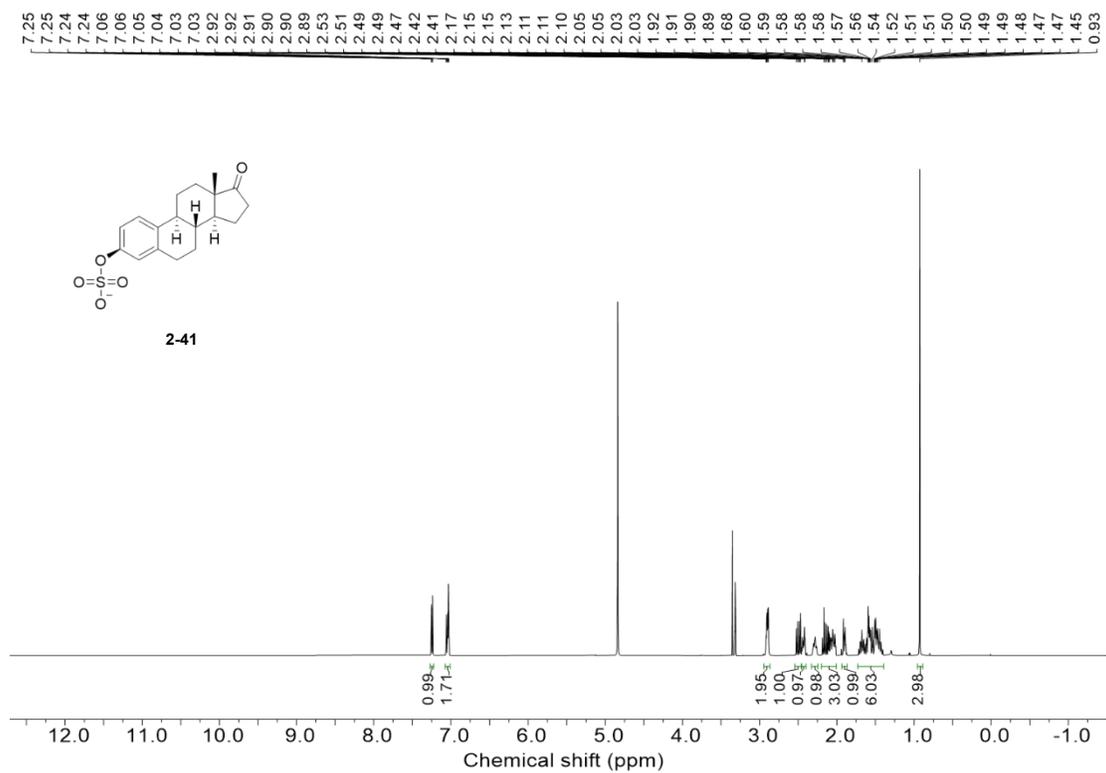
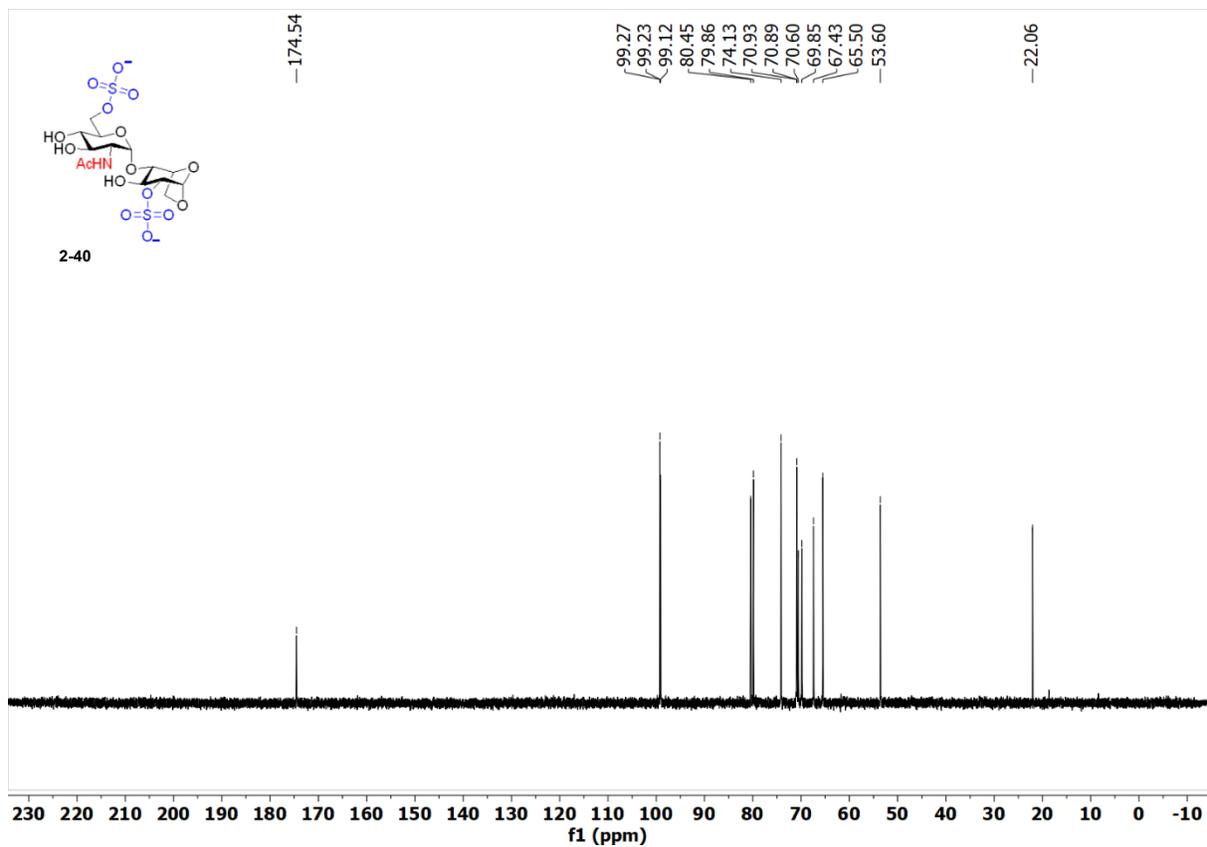


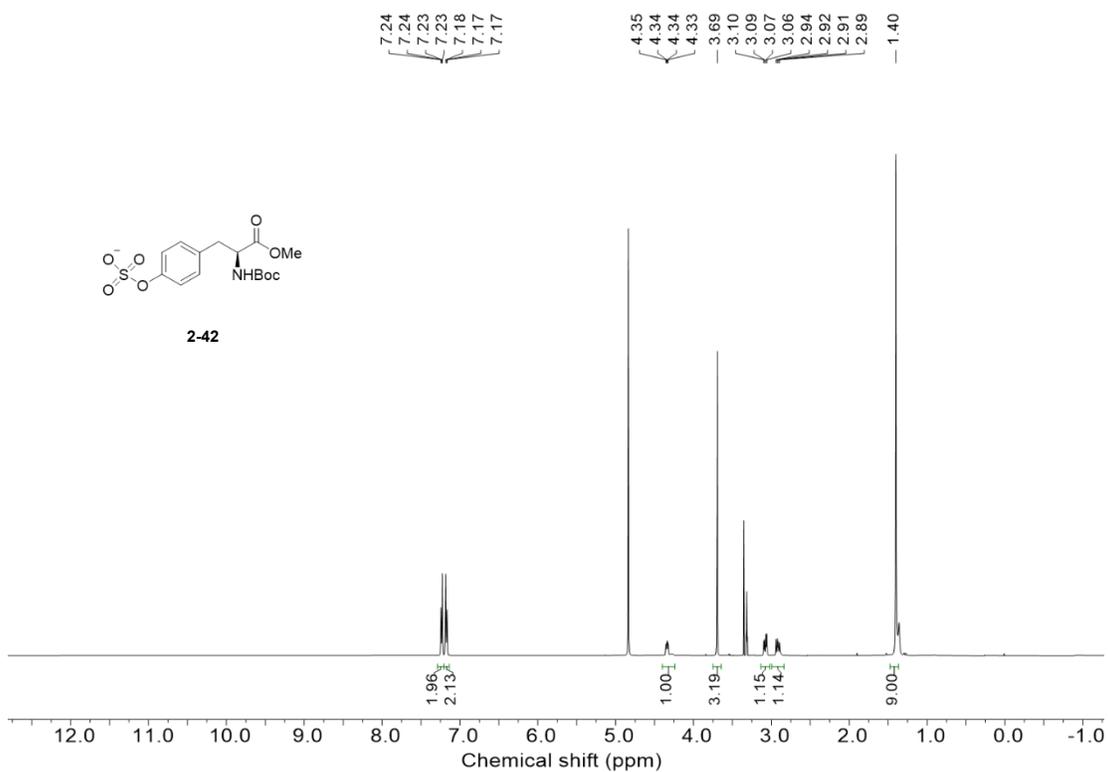
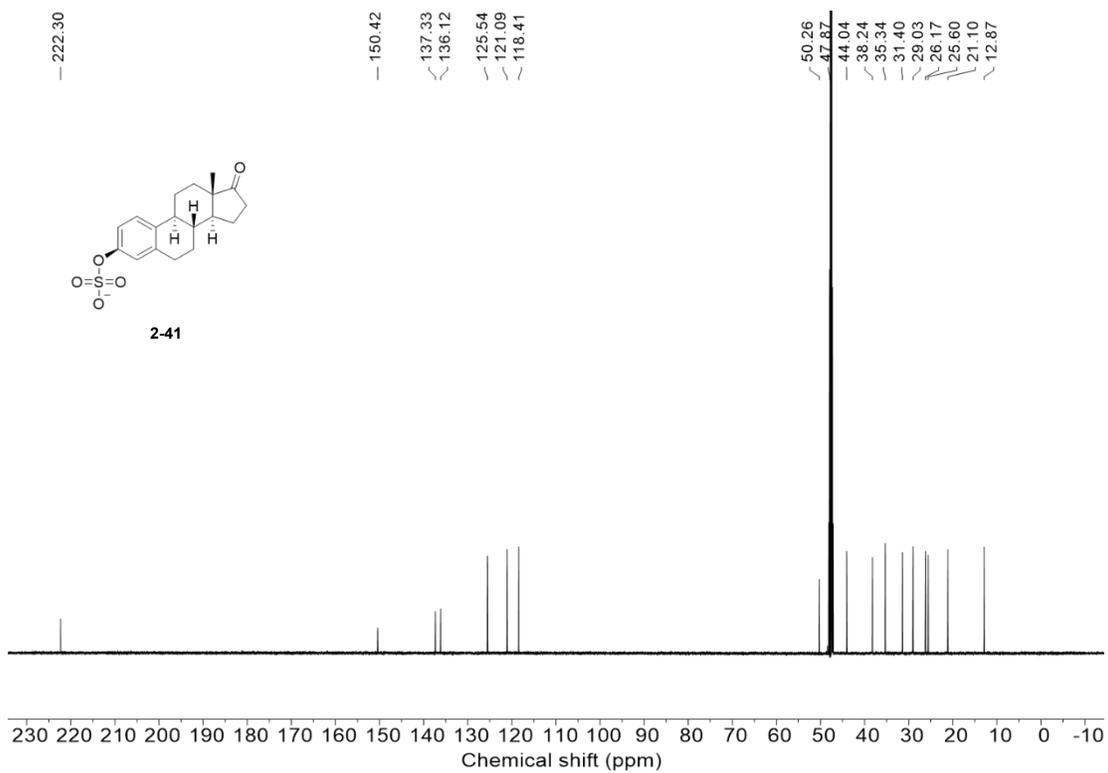


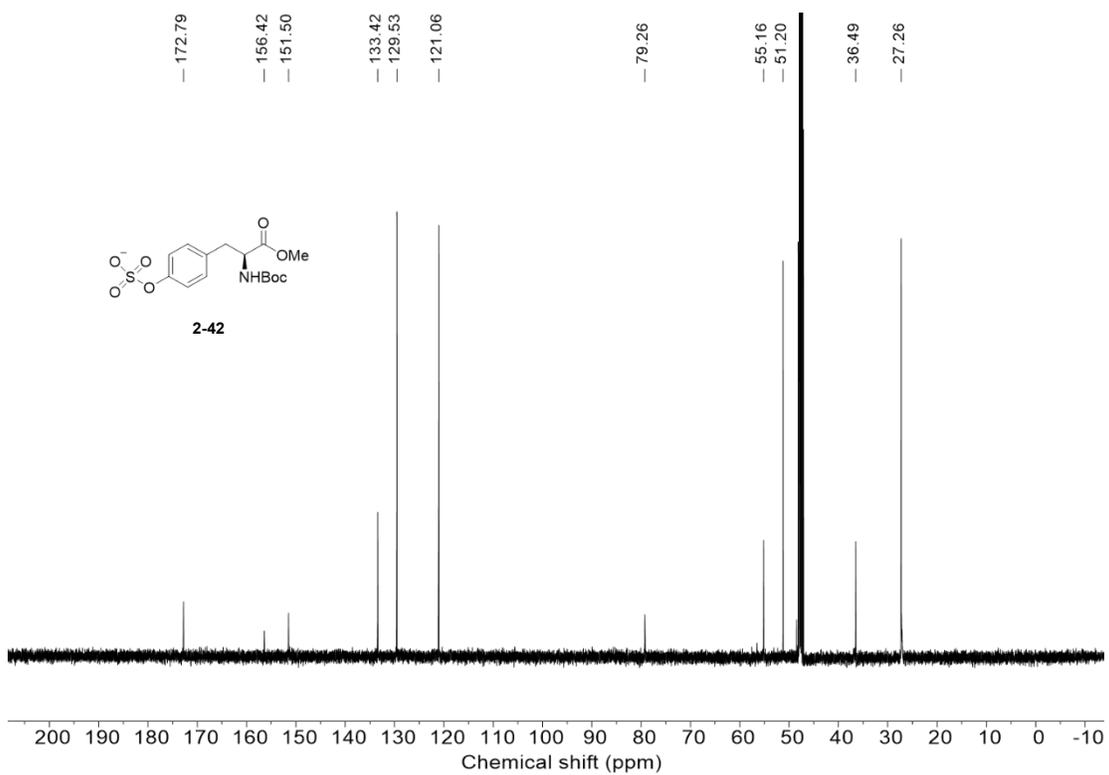












Original NMR spectra of new compounds in Chapter 3

