SYNTHESIS OF CELLULOSIC GLYCOLIPIDS USING ENGINEERED ENZYMES

by

Martina Natasha Chambers

B.Sc., University of Alberta, 2010

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Chemistry)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

February 2014

© Martina Natasha Chambers, 2014
Abstract

Cellulose, a linear polymer of D-glucose units connected by β-1,4 glycosidic bonds, adopts a highly-ordered crystalline structure in solution. In cellulose I, the dominant form of cellulose in nature, the polymeric chains are aligned in the same direction. Previous attempts to synthesize cellulose I in vitro have resulted in the synthesis of cellulose II, which has the thermodynamically favored anti-parallel orientation of chains. The synthesis of soluble fragments or defined surfaces of cellulose I would enable more detailed study of carbohydrate binding domains and other proteins that interact with cellulose in nature.

The objective of this thesis is to prepare a crystalline surface of cellulose I in a controlled manner through the alignment of cellulolipids. A major focus of this thesis is the synthesis of cellulolipids with a cellohexaosyl head group. Cellohexaose is the shortest cello-oligosaccharide with cellulosic properties, but is consequently insoluble in aqueous solution. To improve the solubility of cellohexaose, the addition of a removable charged functionality was explored: either a terminal sialic acid or a phosphate group at the 6 position of the non-reducing terminal glucose. Abg2F6 glycosynthase from Agrobacterium sp. was used to synthesize β-1,4 linked cello-oligosaccharide fluorides from DP = 2 to DP = 4. These cello-oligosaccharides were modified with a removable charged functionality and utilized as donor substrates by CelB glycosynthase, a mutant of a β-1,4 endoglucanase from Caldicellulosiruptor saccharolyticus. Through the combination of glycosynthase enzymes and charged functionalities, a variety of soluble cellohexaosyl analogs were synthesized.

Lyso-glycosphingolipids were prepared by transferring cello-oligosaccharyl fluorides to D-erythro-C18-sphingosine using EGCase glycosynthase. CelB glycosynthase used charged glycosyl fluoride donors to extend the lyso-glycosphingolipids, yielding soluble cellulolipids.
The soluble cellulolipids were aligned along an aqueous:organic interface and the charged functionality was removed. Thus, a surface was prepared that appeared to interact with a carbohydrate binding module functionalized with a fluorescent tag. The soluble cellulolipids were successfully incorporated into a nanodisc, as shown through the incorporation of phosphorylated cellohexaosyl sphingosine. Cleavage of the phosphate using alkaline phosphatase yielded a nanodisc containing cellulolipids.
Preface

This thesis is submitted for the Master of Science degree at the University of British Columbia. All research was performed under the supervision of Dr. Stephen G. Withers in the Department of Chemistry at the University of British Columbia, Point Grey Campus.

The research performed for this thesis is original. The author, Martina Natasha Chambers, performed all experiments, data analysis, and interpretation. Assistance from collaborators within the Withers group is listed below.

Chapter 1: Figures 1-2, 1-3, 1-9, and 1-10 are used with permission from applicable sources.

Chapter 2: Dr. Jamie Rich provided compound 2.23 and Cst-I enzyme. Dr. Hong-Ming Chen provided compounds 2.43, 2.44, and 2.51. For studies in section 2.5, James Frank provided BglK enzyme. In section 2.7, Emily Kwan converted wild-type CelB, CelC, and CenD to glycosynthase enzymes, and determined the linkage synthesized by CelC glycosynthase. Emily Kwan also provided cells containing expression vectors for Abg2F6 glycosynthase, CelB glycosynthase, CelC glycosynthase, CenD glycosynthase, and BglK enzyme.

Chapter 3: Dr. Jamie Rich provided compound 3.01. Dr. Hong-Ming Chen provided compounds 3.33 and 3.36. Dr. Guangyu Yang provided SCDase enzyme and compound 3.41. Emily Kwan provided cells containing expression vectors for EGCase glycosynthase and CBM2a. Experiments in section 3.8 were performed with the assistance of Spence MacDonald.
Table of Contents

Abstract ......................................................................................................................................................... ii
Preface ........................................................................................................................................................ iv
Table of Contents ........................................................................................................................................ v
List of Tables ............................................................................................................................................... x
List of Figures ........................................................................................................................................... xi
List of Schemes .......................................................................................................................................... xiv
List of Abbreviations ................................................................................................................................. xviii
Acknowledgements ...................................................................................................................................... xxiii
Dedication ................................................................................................................................................... xxiv

Chapter 1: General Introduction .............................................................................................................. 1

1.1 Cellulose ............................................................................................................................................ 1
  1.1.1 Crystalline Polymorphs of Cellulose ....................................................................................... 2
  1.1.2 Difficulties in Solubilizing Cellulose .................................................................................... 3

1.2 Interaction of Cellulose and Proteins ............................................................................................. 5
  1.2.1 Synthesis of Cellulose ............................................................................................................ 5
  1.2.2 Proteins that Bind to Cellulose ............................................................................................. 7
  1.2.3 Degradation of Cellulose ....................................................................................................... 8

1.3 Engineering of Glycosidase Enzymes .............................................................................................. 12
  1.3.1 Classification of Glycoside Hydrolases (Glycosidases) ......................................................... 14
  1.3.2 Catalytic Mechanism of Retaining β-Glycosidases ............................................................... 16
  1.3.3 Conversion of Glycosidases into Glycosynthases ............................................................... 17
1.4 Structure of Glycolipids in Solution ................................................................. 20

1.4.1 Properties of Glycosphingolipids ................................................................. 23

1.5 Aims of this Thesis .............................................................................................. 25

Chapter 2: Synthesis of Soluble Cellulosic Oligosaccharides Through Introduction of
Charged Functionalities .......................................................................................... 27

2.1 Synthetic Strategy ............................................................................................... 27

2.2 Synthesis of α-Glycosyl Fluorides .................................................................... 28

2.3 Use of Abg2F6 Glycosynthase in Oligosaccharyl Fluoride Synthesis ............... 30

2.3.1 Synthesis of Longer Cellulosic Glycosyl Fluorides ....................................... 31

2.3.2 Strategies to Purify Glycosyl Fluoride Mixtures ............................................ 33

2.3.2.1 Size Exclusion Chromatography ............................................................... 33

2.3.2.2 Acetylation and Flash Column Chromatography ....................................... 35

2.4 Sialic Acid Transfer to Glycosyl Fluorides Using Cst-I ..................................... 37

2.5 Phosphorylation of Glycosyl Fluorides Using BglK ........................................ 41

2.5.1 Cleavage of Phosphate Group ....................................................................... 42

2.6 Exploration of Glycoside Acceptors for Testing Glycosynthase Enzymes .......... 43

2.6.1 Synthesis of DNPC ....................................................................................... 43

2.6.1.1 Cleavage of DNP Group ........................................................................... 45

2.6.2 Synthesis of Alkoxyamino Glycoside Acceptors ......................................... 48

2.6.2.1 Cleavage of Alkoxyamino Functionality .................................................. 51

2.6.3 Comparison of Acceptor Aryl Glycosides with Abg2F6 ............................... 52

2.7 Study of New Glycosynthase Enzymes ............................................................ 53

2.7.1 Transfer of Cello-oligosaccharyl Fluorides Using Endo-glycosynthases ...... 54
### Chapter 2: Chemistry of Cellulosic Oligosaccharides

#### 2.7 Transfer of Functional Groups

- 2.7.2 Transfer of Sialic Acid Functionalized Glycosyl Fluorides ........................................... 59
- 2.7.3 Transfer of Phosphorylated Glycosyl Fluorides .............................................................. 62

#### 2.8 Effectiveness of Charged Functionalities in Improving Cellulosic Oligosaccharide Solubility

- Effectiveness of Charged Functionalities in Improving Cellulosic Oligosaccharide Solubility ................................................................. 64

#### 2.9 Conclusions

- Conclusions ................................................................................................................................ 69

### Chapter 3: Synthesis of Cellulosic Glycolipids and Preparation of a Soluble Cellulose Surface

#### 3.1 Synthesis of Glycosphingolipids Using EGCase Glycosynthase

- Synthesis of Glycosphingolipids Using EGCase Glycosynthase ................................................. 72
  - 3.1.1 EGCase Glycosynthase Reaction Using a Domestic Microwave ........................................ 74

#### 3.2 Synthesis of Glycosylceramides

- Synthesis of Glycosylceramides ................................................................................................. 77
  - 3.2.1 Enzymatic Synthesis Using SCDase .................................................................................... 77
  - 3.2.2 Chemical Synthesis Using Pentafluorophenyl Esters ......................................................... 79
  - 3.2.3 Chemical Synthesis Using Succinimidyl Esters ................................................................. 82

#### 3.3 Preparation of Elongated Glycosphingolipids

- Preparation of Elongated Glycosphingolipids ........................................................................... 84
  - 3.3.1 Ability of CelB Glycosynthase to Utilize Glycosphingolipid Acceptors ............................... 84
  - 3.3.2 Sialic Acid Functionalization Using Cst-I ............................................................................. 88
  - 3.3.3 Phosphorylation Using BglK ............................................................................................... 90
  - 3.3.4 Synthesis of Phosphorylated Cellohexaosyl Sphingosine .................................................. 93

#### 3.4 Preparation of Fluorescent Glycosphingolipids

- Preparation of Fluorescent Glycosphingolipids ....................................................................... 94
  - 3.4.1 Synthesis of 4-Keto Glycosyl Fluoride Donors .................................................................. 95
  - 3.4.2 Synthesis and Testing of 4-Amino Glycosyl Fluoride Donors .......................................... 97
  - 3.4.3 Synthesis of a Bodipy Labeled Cellulolipid ........................................................................ 101

#### 3.5 Preparation of Bodipy-Cellotetraoside

- Preparation of Bodipy-Cellotetraoside ...................................................................................... 103
3.6 Preparation of Fluorescent CBM2a for Cellulose Surface Detection ........................................ 106
  3.6.1 Labeling of CBM2a with Fluorescein Isothiocyanate ...................................................... 106
  3.6.2 Confirmation of FITC-CBM2a Binding to Cellulose Columns ....................................... 108
3.7 Attempted Preparation of a Cellulose Surface Along an Aqueous:Organic Interface.... 109
3.8 Preparation of Nanodiscs Containing Cellulosic Glycolipids ................................. 113
3.9 Conclusions and Future Directions .................................................................................. 118

Chapter 4: Materials and Methods ...................................................................................... 123

4.1 Generous Gifts .................................................................................................................. 123
4.2 General Synthesis and Methods ..................................................................................... 123
  4.2.1 General Procedure for Acetylation of Free Sugars ...................................................... 124
  4.2.2 General Procedure for Synthesis of α-Glycosyl Fluorides ........................................ 125
  4.2.3 General Procedure for Deacetylation .......................................................................... 125
  4.2.4 General Procedure for Synthesis of Oligosaccharides Using Abg2F6 Glycosynthase .................................................................................. 126
  4.2.5 General Procedure for Acetylation of α-Glycosyl Fluoride Products of Glycosynthase Reactions .................................................................................. 126
  4.2.6 General Procedure for Sialoside Synthesis Using Cst-I Enzyme .............................. 127
  4.2.7 General Procedure for Phosphorylation Using BglK Enzyme ................................ 127
  4.2.8 General Procedure for Ganglioside Synthesis Using EGCase Glycosynthase .... 128
4.3 Synthesis and Characterization ........................................................................................ 128
4.4 Enzymology ..................................................................................................................... 178
  4.4.1 Production of Abg2F6 Glycosynthase ....................................................................... 178
  4.4.2 Production of BglK .................................................................................................... 179
4.4.3 Production of CelB Glycosynthase ................................................................. 180
4.4.4 Production of CelC Glycosynthase ................................................................. 181
4.4.5 Production of CenD Glycosynthase ................................................................. 182
4.4.6 Production of EGCase Glycosynthase ............................................................ 183
4.4.7 Production of CBM2a ..................................................................................... 184
   4.4.7.1 Labeling of CBM2a with FITC ................................................................. 185

References ..................................................................................................................... 186
List of Tables

Table 2-1: Results for the synthesis of α–glycosyl fluorides...................................................... 30
Table 2-2: Glycosyl fluoride donor specificity for CelB, CelC, and CenD glycosynthase enzymes........................................................................................................... 59
Table 2-3: Glycosyl fluoride and phosphorylated glycosyl fluoride donor specificity for CelB glycosynthase and CelC glycosynthase. .................................................................................................................. 64
Table 3-1: Cellulolipid synthesis results using EGCase glycosynthase. ..................................... 74
Table 3-2: Synthesis of PFP activated esters of long-chain fatty acids..................................... 81
Table 3-3: MALDI-MS results from reactions using CelB glycosynthase and glycosphingolipid acceptors. ......................................................................................................................... 87
List of Figures

Figure 1-1: Structure of cellulose, a linear polymer of β-(1→4) linked D-glucopyranose units.... 2
Figure 1-2: General mechanism of copper-dependent polysaccharide monooxygenases. .......... 10
Figure 1-3: Cleavage of cellulose chains using polysaccharide monooxygenases.................. 10
Figure 1-4: Glycone and aglycone components of a D-glucoside, along with the carbohydrate numbering scheme used in this thesis.......................................................... 13
Figure 1-5: General glycosidase-catalyzed hydrolysis reaction. ........................................... 13
Figure 1-6: Stereochemical outcome of reactions catalyzed by either retaining or inverting glycosidases. ........................................................................................................ 15
Figure 1-7: General mechanism for retaining β-glycosidases. ............................................. 17
Figure 1-8: General mechanism for glycosynthase enzymes with the catalytic nucleophile mutated to a glycine residue. ................................................................. 19
Figure 1-9: Structure of lipid aggregates in solution................................................................. 22
Figure 1-10: Formation of nanodiscs by mixing MSP with detergent solubilized phospholipids, followed by removal of the detergent....................................................... 23
Figure 1-11: Structure of sphingosine backbone ((2S,3R,4E)-2-amino-4-octadecene-1,3-diol), along with the labeling scheme used in this thesis. ........................................... 24
Figure 2-1: Example of proposed strategy to synthesize cellulolipids................................... 28
Figure 2-2: Structure of sialic acid (α-anomer) and the numbering scheme used in this thesis .. 37
Figure 2-3: Theorized product synthesized from the reaction of DNPC (2.33) and piperidine. 47
Figure 2-4: Crystalline cellohexaose (2.55, A) and α-cellohexaosyl fluoride (2.56, B) synthesized using CelB glycosynthase................................................................. 66
Figure 2-5: Comparison of crystalline α-cellohexaosyl fluoride (2.56, A) and β-1,3 linked hexasaccharide (2.57, B) synthesized using CelC glycosynthase.

Figure 2-6: Ability of phosphorylation or sialic acid functionalization to improve solubility of cellohexaose. A) Control reaction (2.54); B) Cellohexaose (2.55) C) Reaction containing 2.28 and 2.54; D) Reaction containing 2.25 and 2.54.

Figure 3-1: Hydrolysis of glycosylceramide using EGCase. The ceramide product consists of a sphingosine base (highlighted in red) that is N-acylated with a long-chain fatty acid.

Figure 3-2: Result of domestic microwave study using EGCase glycosynthase.

Figure 3-3: Illustration of how bodipy cellotetraoside could act as a probe to detect a surface of cellulose I.

Figure 3-4: Structure of fluorescein isothiocyanate (FITC).

Figure 3-5: (A) Fluorescence of cellulose columns after washing with 500 mL deionized H$_2$O. The three columns on the left were originally loaded with FITC dye. The three columns on the right were loaded with FITC-CBM2a. (B) Fluorescence of cellulose columns originally loaded with FITC-CBM2a after washing with 25 mL of 6 M guanidinium chloride in 50 mM potassium phosphate (pH 7.0).

Figure 3-6: Aqueous:organic interface prepared with 50 mM HEPES (pH 7.5) and chlorinated organic solvent. 34 pmole of 3.32 was added, and the vials were incubated at 4°C with alkaline phosphatase for 2 hours. Left: Chloroform; Middle: Dichloromethane; Right: 1,2-Dichloroethane.

Figure 3-7: Vials containing an aqueous:organic interface prepared with 50 mM HEPES (pH 7.5) and chlorinated organic solvent. The vials have been illuminated with 365 nm UV-light. Vials on the left contained cellotetraosyl sphingosine and FITC-CBM2a. Vials in the middle
contained cellohexaosyl sphingosine and bodipy cellotetraoside (3.43). Vials on the right contained cellohexaosyl sphingosine and FITC-CBM2a. Vial order: Left: Chloroform; Middle: Dichloromethane; Right: 1,2-Dichloroethane................................. 112

Figure 3-8: Native gel of reconstituted nanodiscs prepared with DOPG phospholipid. ........ 115

Figure 3-9: Nanodiscs reconstituted with fluorescent cellulosic sphingolipid (FLS, 3.42). (A) Native PAGE gel stained with coomassie blue (left) and UV-exposed native gel (right). (B) Size exclusion chromatography (SEC) of nanodisc reconstitutions shown in (A).............................. 117

List of Schemes

Scheme 2-1: Synthesis of α–glucosyl fluoride (2.04). i) NaOAc (1.1 eq.), Ac₂O (3 eq./OH), reflux, 1 hour, 65%; ii) HF-pyridine (45 eq.), 4°C, overnight, 95%; iii) NH₃ (g), MeOH, 4°C, overnight, 88%.................................................................................................................................................. 29

Scheme 2-2: Synthesis of 2.17 using Abg2F6 glycosynthase.......................................................... 31

Scheme 2-3: Synthesis of 2.18 and 2.19 using Abg2F6 glycosynthase............................................. 32

Scheme 2-4: Per-O-acetylation of 2.17 using Ac₂O and pyridine....................................................... 36

Scheme 2-5: Synthesis of 2.24 using Cst-I enzyme......................................................................... 38

Scheme 2-6: Synthesis of 2.25 using Cst-I enzyme. i) Abg2F6 glycosynthase (0.10 mg/mL), 100 mM sodium phosphate (pH 7.0), RT, overnight; ii) Cst-I (0.20 mg/mL), 50 mM HEPES (pH 7.5), 10 mM MnCl₂, RT, 2 hours...................................................................................................................... 39

Scheme 2-7: Synthesis of 2.26 using Abg2F6 glycosynthase.......................................................... 40

Scheme 2-8: Synthesis of 6’-O-phospho-cellobiosyl fluoride (2.27) using BglK.......................... 41

Scheme 2-9: Cleavage of phosphate functionality on 2.28 using alkaline phosphatase............. 42

Scheme 2-10: Synthesis of DNPC (2.33). i) Hydrazine acetate (1.2 eq.), DMF, RT, 4 hours, 83%; ii) DABCO (4 eq.), DMF, RT, 20 hours, 67%; iii) AcCl (45 eq.), MeOH, 4°C, 72 hours, 72%. 44

Scheme 2-11: Reaction of GalF (2.08) and DNPC (2.33) using Abg2F6 glycosynthase to synthesize 2.34................................................................................................................................. 45

Scheme 2-12: Cleavage of DNP group using HF-pyridine. ............................................................... 46

Scheme 2-13: Attempted cleavage of DNP group using three different reagents. a) 2-mercaptoethanol (59 eq.), DMF:pyridine (10:1), RT, 42 hours; b) thiophenol (39 eq.), DMF, RT, 42 hours; c) piperidine (20% v/v), DMF, RT, 42 hours. .................................................................................................................................................. 47
Scheme 2-14: Synthesis of alkoxyamino β-D-glucoside (2.40). i) NaOAc (1.2 eq.), 4:1 H₂O:MeOH, RT, 2 hours, 97%; ii) NaBH₃CN (3.9 eq.), EtOH, RT, 30 min, 85%; iii) 75:25 THF:AcOH, RT, 20 hours, 58%.

Scheme 2-15: Synthesis of 2.41 using Abg2F6 glycosynthase.

Scheme 2-16: Synthesis of alkoxyamino cellobioside acceptor (2.42).

Scheme 2-17: Cleavage of alkoxyamino functionality on 2.41.

Scheme 2-18: Ability of new glycosynthase enzymes to transfer 2.12 to 2.33. The synthesized glycosidic linkage is highlighted in red. a) 50 mM sodium phosphate (pH 8.0), CelB glycosynthase (1.25 mg/mL), 30°C; b) 50 mM sodium phosphate (pH 7.0), CelC glycosynthase (0.25 mg/mL), 30°C.

Scheme 2-19: Ability of new glycosynthase enzymes to transfer 2.17 to 2.33. a) 50 mM sodium phosphate (pH 8.0), CelB glycosynthase (1.25 mg/mL), 30°C; b) 50 mM sodium phosphate (pH 7.0), CelC glycosynthase (0.25 mg/mL), 30°C.

Scheme 2-20: Ability of CelB or CelC glycosynthase to transfer 2.25. a) 50 mM sodium phosphate (pH 8.0), CelB glycosynthase (1.0 mg/mL), 30°C; b) 50 mM sodium phosphate (pH 7.0), CelC glycosynthase (0.25 mg/mL), 30°C.

Scheme 2-21: Ability of CelB or CelC glycosynthase to transfer 2.24. a) 50 mM sodium phosphate (pH 8.0), CelB glycosynthase (1.0 mg/mL), 30°C; b) 50 mM sodium phosphate (pH 7.0), CelC glycosynthase (0.25 mg/mL), 30°C.

Scheme 2-22: Synthesis of cellohexaose (2.55). i) 50 mM sodium phosphate (pH 8.0), CelB glycosynthase (0.50 mg/mL), 30°C, 18 hours.

Scheme 2-23: Synthesis of α-cellohexaosyl fluoride (2.56). i) 50 mM sodium phosphate (pH 8.0), CelB glycosynthase (0.50 mg/mL), 30°C, 18 hours.
Scheme 2-24: Synthesis of β-1,3 linked hexasaccharide (2.57). i) 50 mM sodium phosphate (pH 7.0), CelC glycosynthase (0.50 mg/mL), 30°C, 18 hours.......................................................... 67

Scheme 3-1: Synthesis of 3.02 using EGCase glycosynthase. ......................................................... 74

Scheme 3-2: Synthesis of cellobiosyl ceramide (3.07) using SCDase. ............................................. 78

Scheme 3-3: Synthesis of PFP activated esters using pentafluorophenol (3.08) and myristic acid (3.09).......................................................................................................................... 80

Scheme 3-4: Synthesis of PFP 2-hydroxybenzoate (3.16) using TFA-OPFP (3.14).......................... 80

Scheme 3-5: Synthesis of cellobiosyl ceramide (3.07) using PFP-stearate (3.13). ......................... 82

Scheme 3-6: Synthesis of 3.18 using NHS (3.17) and myristic acid (3.09). ................................. 83

Scheme 3-7: Synthesis of 3.21 using NHS-tetradecanoate (3.18) .................................................. 84

Scheme 3-8: Test reaction of 2.16 and 3.03 using CelB glycosynthase. ........................................ 85

Scheme 3-9: Test reaction of 2.17 and 3.03 using CelB glycosynthase. ........................................ 86

Scheme 3-10: Test reaction of 2.17 and 3.05 using CelB glycosynthase. ....................................... 86

Scheme 3-11: Synthesis of sialylated lactosyl sphingosine (3.26) using Cst-I ......................... 89

Scheme 3-12: Synthesis of 3.28 using Abg2F6 glycosynthase and Cst-I ......................................... 90

Scheme 3-13: Synthesis of 6’-O-phospho-cellobiosyl sphingosine (3.29) using BglK .......... 92

Scheme 3-14: Cleavage of phosphate functionality on 3.30 using alkaline phosphatase............. 93

Scheme 3-15: Synthesis of phosphorylated cellohexaosyl sphingosine (3.32). ............................. 94

Scheme 3-16: Synthesis of 3.34 using Abg2F6 glycosynthase. .................................................... 95

Scheme 3-17: Synthesis of 4’’-keto-cellotriosyl fluoride (3.35) using Abg2F6 glycosynthase .. 96


Scheme 3-19: Synthesis of 3.38 using CelB glycosynthase. .......................................................... 99
Scheme 3-20: Conversion of 3.04 to 3.39 using TfN₃. i) NaN₃ (1.5 eq.), Tf₂O, H₂O, toluene, 10°C, 2 hours; ii) TfN₃, CuSO₄·5H₂O (1.3 eq.), NaHCO₃ (2 eq.), H₂O, toluene, RT, overnight, 64%. 

Scheme 3-21: Synthesis of 3.40 using CelB glycosynthase. 


Scheme 3-23: Synthesis of bodipy cellotetraoside (3.43) using CelB glycosynthase. 

Scheme 3-24: Synthesis of a thioacetate-terminated NHS-activated fatty acid (3.46). i) Aqueous HBr (41.2 eq.), AcOH, reflux, 3 days, 99%; ii) AcSK (2.0 eq.), DMF, 50°C, 15 hours, 70%; iii) NHS (4.0 eq.), EDC (4.15 eq.), DCM, RT, overnight, 64%. 

Scheme 3-25: Synthesis of a cellobiosyl ceramide analog with a terminal thioacetate (3.47). i) 3.46, triethylamine (1.65 eq.), DMF, RT, overnight, 84%. 
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>2xYT</td>
<td>Cell media containing double the yeast extract and tryptone compared to LB cell media</td>
</tr>
<tr>
<td>Å</td>
<td>Angstrom</td>
</tr>
<tr>
<td>A&lt;sub&gt;280&lt;/sub&gt;</td>
<td>Absorbance at 280 nm</td>
</tr>
<tr>
<td>A&lt;sub&gt;496&lt;/sub&gt;</td>
<td>Absorbance at 496 nm</td>
</tr>
<tr>
<td>Abg</td>
<td><em>Agrobacterium</em> sp. β-glucosidase</td>
</tr>
<tr>
<td>Abg2F6</td>
<td>Abg with 4 mutations: A19T, E358G, Q248R, and M407V.</td>
</tr>
<tr>
<td>Ac</td>
<td>Acetyl group</td>
</tr>
<tr>
<td>AcOH</td>
<td>Acetic acid</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine-5’-triphosphate</td>
</tr>
<tr>
<td>BCA</td>
<td>Bicinchoninic acid</td>
</tr>
<tr>
<td>BglK</td>
<td>β-glucoside kinase</td>
</tr>
<tr>
<td>Bodipy</td>
<td>Boron-dipyrromethene</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>CAZy</td>
<td>Carbohydrate-active enzyme</td>
</tr>
<tr>
<td>CBD</td>
<td>Cellulose binding domain</td>
</tr>
<tr>
<td>CBM</td>
<td>Carbohydrate binding module</td>
</tr>
<tr>
<td>CBM2a</td>
<td>Carbohydrate binding module family 2a</td>
</tr>
<tr>
<td>CDH</td>
<td>Cellobiose dehydrogenase</td>
</tr>
<tr>
<td>CelB</td>
<td><em>Caldicellulosiruptor saccharolyticus</em> endoglucanase B</td>
</tr>
<tr>
<td>CelC</td>
<td><em>Clostridium thermocellum</em> endoglucanase C</td>
</tr>
<tr>
<td>CenD</td>
<td><em>Cellulomonas fimi</em> endoglucanase D</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>cmc</td>
<td>Critical micelle concentration</td>
</tr>
<tr>
<td>CMP</td>
<td>Cytidine-5’-monophosphate</td>
</tr>
<tr>
<td>Cst-I</td>
<td>Campylobacter jejuni α-2,3-sialyltransferase</td>
</tr>
<tr>
<td>CTP</td>
<td>Cytidine-5’-triphosphate</td>
</tr>
<tr>
<td>δ</td>
<td>Chemical shift</td>
</tr>
<tr>
<td>DABCO</td>
<td>1,4-Diazobicyclo[2.2.2]octane</td>
</tr>
<tr>
<td>DCM</td>
<td>Dichloromethane</td>
</tr>
<tr>
<td>DDM</td>
<td>n-Dodecyl-β-D-maltoside</td>
</tr>
<tr>
<td>DHB</td>
<td>2,5-Dihydroxybenzoic acid</td>
</tr>
<tr>
<td>DMAP</td>
<td>4-Dimethylaminopyridine</td>
</tr>
<tr>
<td>DME</td>
<td>1,2-Dimethoxyethane</td>
</tr>
<tr>
<td>DMF</td>
<td>N,N-Dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNFB</td>
<td>1-Fluoro-2,4-dinitrobenzene</td>
</tr>
<tr>
<td>DNPC</td>
<td>2,4-Dinitrophenyl β-cellobioside</td>
</tr>
<tr>
<td>DOPC</td>
<td>1,2-Dioleoylphosphatidylcholine</td>
</tr>
<tr>
<td>DOPG</td>
<td>1,2-Dioleoylphosphatidylglycerol</td>
</tr>
<tr>
<td>DP</td>
<td>Degree of polymerization</td>
</tr>
<tr>
<td>ε</td>
<td>Molar absorptivity</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>EDC</td>
<td>1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide</td>
</tr>
<tr>
<td>EGCase</td>
<td>Endoglycoceramidase</td>
</tr>
</tbody>
</table>
**Eq.**
Equivalent

**ESI-MS**
Electrospray ionization mass spectrometry

**Et<sub>3</sub>N**
Triethylamine

**EtOAc**
Ethyl acetate

**EtOH**
Ethanol

**FITC**
Fluorescein isothiocyanate

**Fmoc**
9-Fluoroenylmethoxycarbonyl

**GalF**
α-Galactosyl fluoride

**GH**
Glycoside hydrolase

**GlcF**
α-Glucosyl fluoride

**HEPES**
4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid

**HF**
Hydrogen fluoride

**Hz**
Hertz

**IPTG**
Isopropyl β-D-1-thiogalactopyranoside

**IUPAC**
International Union of Pure and Applied Chemistry

**J**
Coupling constant

**kan**
Kanamycin

**kDa**
Kilodalton

**KU**
1000 Enzyme units

**LB**
Lysogeny broth

**M**
Molarity (mol/L)

**MΩ**
Megaohm

**m/z**
Mass-to-charge ratio
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>MALDI-MS</td>
<td>Matrix-assisted laser desorption/ionization mass spectrometry</td>
</tr>
<tr>
<td>MeCN</td>
<td>Acetonitrile</td>
</tr>
<tr>
<td>MeOD</td>
<td>Deuterated methanol</td>
</tr>
<tr>
<td>MeOH</td>
<td>Methanol</td>
</tr>
<tr>
<td>MES</td>
<td>2-(N-morpholino)ethanesulfonic acid</td>
</tr>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>MSP</td>
<td>Membrane scaffold protein</td>
</tr>
<tr>
<td>MW</td>
<td>Molecular weight</td>
</tr>
<tr>
<td>Neu5Ac</td>
<td>N-acetylneuraminic acid</td>
</tr>
<tr>
<td>NHS</td>
<td>N-hydroxysuccinimide</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PE</td>
<td>Petroleum ether</td>
</tr>
<tr>
<td>PFP</td>
<td>Pentafluorophenyl</td>
</tr>
<tr>
<td>PNPC</td>
<td>Para-nitrophenyl β-celllobioside</td>
</tr>
<tr>
<td>PNPG</td>
<td>Para-nitrophenyl β-D-glucopyranoside</td>
</tr>
<tr>
<td>ppm</td>
<td>Parts per million</td>
</tr>
<tr>
<td>R&lt;sub&gt;f&lt;/sub&gt;</td>
<td>Retention factor</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>RT</td>
<td>Room temperature</td>
</tr>
<tr>
<td>SCDase</td>
<td><em>Shewanella alga</em> G8 sphingolipid ceramide N-deacylase</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>TFA</td>
<td>Trifluoroacetic acid</td>
</tr>
<tr>
<td>TFA-OPFP</td>
<td>Pentafluorophenyl trifluoroacetate</td>
</tr>
<tr>
<td>TLC</td>
<td>Thin layer chromatography</td>
</tr>
<tr>
<td>THF</td>
<td>Tetrahydrofuran</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>TYP</td>
<td>Cell media containing tryptone, yeast extract, and dipotassium hydrogen phosphate</td>
</tr>
<tr>
<td>UDP</td>
<td>Uridine 5’-diphosphate</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
</tbody>
</table>
Acknowledgements

I would like to thank my supervisor, Dr. Steve Withers, for all of his helpful advice, guidance, and support throughout this degree. I also want to thank past and present members of the Withers group for all of their assistance in the lab. Special thanks go out to Dr. Jamie Rich and Dr. Michael Fischer for all of their hands-on help, suggestions, and insight regarding organic synthesis. I want to acknowledge Dr. Tina Rasmussen, Dr. Anna Win-Mason, and Dr. Guangyu Yang for many helpful discussions. I want to thank Emily Kwan for her assistance with enzyme expression, and Dr. Hong-Ming Chen for providing many substrates for this project. I especially want to thank the staff in the NMR and mass spectrometry facilities for all of their hard work.

I would like to thank all of the friends I made in the department throughout this degree. I am especially grateful to Leanne Racicot who allowed me to invade her lab space constantly. I am grateful to all the ladies on the chemistry dodgeball team, who gave me something to look forward to on Tuesday evenings. I am also grateful to the ladies in my book club for all of the interesting discussions and book suggestions. I also want to acknowledge the CGSS for giving me the great opportunity to be the council secretary for two years.

I am indebted to all of my friends and family for their support and encouragement throughout this degree. I especially want to thank my parents for all of their support. I also want to thank my parents-in-law, Keith and Brenda Chambers, for always sending “good thoughts” my way.

Last, but certainly not least, I would like to thank my husband, Stu Chambers. Your unconditional love, support, and guidance inspire me constantly. I could not have completed this degree without you by my side.
This thesis is dedicated to my parents, Fay and Jim Morrissey.

Without your support, I would not be where I am today.

Thank you.
Chapter 1: General Introduction

1.1 Cellulose

Cellulose, the most abundant organic polymer on Earth, is a fibrous, rigid, water-insoluble substance.\(^1\) Cellulose is a dominant component of plant cell walls that provides cellular strength and system rigidity. Importantly, cellulose prevents cell swelling, and aids in maintaining a stable osmotic environment in the cell. Although cellulose is primarily found in plants, it is also synthesized by bacteria, algae, and fungi.\(^2\) It is estimated that \(1.5 \times 10^{12}\) tons of cellulose is naturally synthesized every year.\(^3\) In industry, cellulose is heavily utilized in paper, cotton, cellophane, building materials, and as a fiber supplement in food. Recently, there has been significant interest in utilizing cellulose as an alternative energy source by converting it into cellulosic ethanol.\(^4\)

In 1838, Anselme Payen, a French chemist, discovered cellulose.\(^5\) By treating various plant tissues with nitric acid and an alkaline solution, Payen attempted to separate the base components of plants. This treatment yielded an insoluble fibrous substance that was similar to starch, having the same chemical formula: \(\text{C}_6\text{H}_{10}\text{O}_5\). This residue was first referred to as “cellulose” in 1839 by the French Academy after reviewing Payen’s work.\(^6\) The polymeric structure of cellulose was first elucidated by Staudinger in 1920.\(^7\)
1.1.1 Crystalline Polymorphs of Cellulose

Cellulose is a linear, unbranched polymer of repeating D-glucose units connected by β-(1→4) glycosidic bonds (Figure 1-1). The D-glucose units adopt the $^4C_1$ chair confirmation, enabling all equatorial hydroxyl groups to hydrogen bond with neighboring cellulose chains. The inter-chain and intra-chain hydrogen bonding network produces highly ordered, crystalline microfibril bundles of cellulose.\(^8\)

![Image of cellulose structure]

**Figure 1-1:** Structure of cellulose, a linear polymer of β-(1→4) linked D-glucopyranose units.

Six different polymorphs of crystalline cellulose have been confirmed by x-ray diffraction.\(^9\) Cellulose I is the dominant form of cellulose found in nature and consists of parallel polymeric cellulose chains aligned in the same direction. Cellulose I can be further divided into two subtypes: I\(_a\) (triclinic) and I\(_β\) (monoclinic). Cellulose I\(_a\) is commonly found in algae and bacterial cellulose, while cellulose I\(_β\) is typically found in cotton and wood.\(^10\) Cellulose I\(_a\) and I\(_β\) have very similar crystal structures, but I\(_β\) contains two cellulose chains that are conformationally distinct, with different hydrogen bonding patterns and greater thermodynamic stability.\(^11\) Cellulose I\(_a\) can be converted to cellulose I\(_β\) by heating to 220-260°C in the presence of sodium hydroxide.\(^12\)
Cellulose II is the most prevalent form of cellulose utilized in industrial applications. Unlike cellulose I, the polymeric cellulose chains are aligned in an antiparallel orientation. This alignment makes cellulose II the most thermodynamically stable crystalline form of cellulose.\(^{13}\) Cellulose II can be synthesized irreversibly from cellulose I by regeneration. In brief, solubilization of the cellulose fibers in a solvent is followed by reprecipitation through dilution in water, or mercerization by swelling the cellulose fibers in concentrated sodium hydroxide.\(^{14}\) If cellulose II prepared through mercerization is washed with water, cellulose II hydrate is synthesized.\(^{15}\) Cellulose II hydrate has an expanded crystalline lattice due to intercalated water molecules, and can be converted back to cellulose II by drying in air.

Cellulose III\(_I\) and III\(_{II}\) are reversibly synthesized from cellulose I and II respectively through treatment with liquid ammonia or specific amines.\(^{16}\) Cellulose IV\(_I\) and IV\(_{II}\) are synthesized from cellulose III\(_I\) and III\(_{II}\) respectively by heating the fibers in glycerol at 206°C.\(^{14}\) The polymorphs of cellulose can interconvert between each other, with the exception of not being able to interconvert back to cellulose I.\(^{9}\) Even though cellulose I is the most abundant form of cellulose found in nature, it is not the most stable. Crystallizing pre-existing cellulose chains only results in the formation of thermodynamically stable cellulose II, complicating attempts to synthesize cellulose I in vitro.

1.1.2 Difficulties in Solubilizing Cellulose

Even though cellulose is composed of individual water-soluble D-glucose units, polymeric cellulose is insoluble in water and typical organic and inorganic solvents found in the laboratory. This insolubility is due to the strong inter-chain and intra-chain hydrogen bonding
network, along with hydrophobic interactions between the cellulose chains.\textsuperscript{18} Therefore, methods explored to dissolve cellulose focus on solvent systems that can interrupt both the hydrogen bond network and tight packing of the cellulose chains. For example, cellulose has been shown to dissolve in aqueous solutions at extreme pH, such as in solutions of phosphoric acid.\textsuperscript{19} The dissolution of cellulose at extreme pH is aided by the addition of co-solutes, which weaken the hydrophobic interactions between the cellulose chains. This has been demonstrated with the addition of thiourea or polyethylene glycol (1\%) to solutions of sodium hydroxide.\textsuperscript{20}

There is also evidence that cellulose can be dissolved in non-conventional amphiphilic solvents, such as \textit{N}-methylmorpholine \textit{N}-oxide in water at elevated temperatures (>90°C), or tetrabutylammonium fluoride in dimethyl sulfoxide.\textsuperscript{21} There is also evidence that cellulose can be dissolved in ionic liquids, solutions of low-melting-point salts, due to the polar-electrostatic interactions between the solvent ions and hydroxyl groups on the cellulose chains.\textsuperscript{22} Ionic liquids are attractive as they are designated “green solvents” due to their thermal and chemical stability, as well as their low vapor pressure.\textsuperscript{23} Cellulose was successfully dissolved in the ionic liquids 1-\textit{N}-butyl-3-methylimidazolium chloride and benzyl(dimethyl(tetradecyl)ammonium chloride, without any evidence of derivatization or degradation according to \textsuperscript{13}C NMR.\textsuperscript{24}

While the methods mentioned above have all dissolved cellulose, the solvents required are not readily available in most laboratories, and can be highly toxic or reactive.\textsuperscript{25} Additionally, the reaction conditions required to dissolve cellulose are typically too severe to allow studies of the dissolved cellulose with enzymes and other cellulose-interacting proteins of interest.

One method considered to counteract the problems mentioned above involves studying shorter lengths of cellulose, which would have significantly reduced hydrogen bonding and hydrophobic interactions compared to their polymeric counterparts. However, cellohexaose is the
shortest cello-oligosaccharide that has a $^{13}$C NMR spectrum similar to polymeric cellulose, and is consequently insoluble in aqueous solutions. Shorter cello-oligosaccharides, while soluble in aqueous solutions, do not interact with cellulose binding proteins with nearly the same affinity as their insoluble counterparts.

1.2 Interaction of Cellulose and Proteins

Cellulose is heavily utilized in many industrial processes. Coupled with its unique crystalline properties, there is significant interest in researching cellulose for both academic and industrial applications. While there has been significant progress in understanding the characteristics and structure of crystalline cellulose, there are still challenges related to synthesizing and handling cellulose I in vitro. Since cellulose is a key structural component of plant cell walls, studies have focused on proteins that are used in nature to synthesize, bind to, and degrade crystalline cellulose.

1.2.1 Synthesis of Cellulose

In plant cell walls, cellulose I is synthesized by cellulose synthase, an enzyme that utilizes UDP-glucose (uridine 5’-diphospho-α-D-glucose) as a donor to elongate cellulose chains. Cellulose synthase is part of a terminal complex, referred to as a rosette, which is located within the plasma membrane and has six-fold symmetry. The simplest explanation for the mechanism of cellulose synthase is that the enzyme catalyzes the synthesis of cellulose in a one-step process, where glucose is transferred from cytosolic UDP-glucose to a cellulose chain, through inversion
of stereochemistry at the anomeric center.\textsuperscript{28} Cellulose synthase elongates cellulose chains to a degree of polymerization of several thousand, and each rosette subunit synthesizes six cellulose chains simultaneously.\textsuperscript{29} During synthesis, the growing cellulose polymer is transported across the cell membrane.\textsuperscript{30} The cellulose chains synthesized by each subunit are assembled in parallel into bundles of around 36 chains, forming aggregates referred to as microfibrils. There have been many articles published that have proposed cellulose synthase mechanisms and the subsequent formation of crystalline microfibrils, primarily focused on \textit{Arabidopsis thaliana}, but no explanation has been proven correct.\textsuperscript{31} Recently, the Zimmer group determined the crystal structure of a cellulose synthase complex from \textit{Rhodobacter sphaeroides} containing a translocating cellulose chain.\textsuperscript{32} While this complex suggests a mechanism for cellulose synthesis and subsequent cellulose transport across the cell membrane, further evidence is required.

Using activated $\beta$-cellobiosyl fluoride as a substrate, crystalline cellulose I was supposedly assembled \textit{in vitro} using a cellulase enzyme from \textit{Trichoderma viride}.\textsuperscript{33} The activated fluoride substrate was transferred to another $\beta$-cellobiosyl fluoride substrate or a lengthened cellulose chain by transglycosylation, which will be further discussed in section 1.3.3. The assembly reaction was performed in a 2:1 organic:aqueous solvent mixture consisting of acetonitrile and 10 mM sodium acetate buffer (pH 5.0). The optimized solvent mixture was postulated to favor the alignment of the synthesized cellulose chains based on their polarity. This resulted in the formation of thin cellulose aggregates. However, it seems suspicious that cellulose I was synthesized, given that cellulose II is the more thermodynamically stable conformation of cellulose.

Cellulose II can be synthesized enzymatically \textit{in vitro} using cellodextrin phosphorylase, an enzyme from \textit{Clostridium thermocellum} that reversibly phosphorolyzes cello-
oligosaccharides.\textsuperscript{34} Cellodextrin phosphorylase was able to successfully utilize cellobiose as an acceptor, synthesizing a mixture of soluble cello-oligosaccharides and insoluble cellulose II. Hiraishi et al. demonstrated that by utilizing high concentrations of \( \alpha \)-glucose-1-phosphate as a donor and \( \text{D} \)-glucose as an acceptor, cellodextrin phosphorylase was able to solely synthesize insoluble cello-oligosaccharides with an average degree of polymerization (DP) of 9.\textsuperscript{35} Electron and x-ray diffraction studies confirmed the insoluble cello-oligosaccharides were highly crystalline with an orientation similar to cellulose II.

1.2.2 Proteins that Bind to Cellulose

Many carbohydrate-active enzymes have two or more separate protein domains: a catalytic domain containing the enzyme active site, and a non-catalytic domain that facilities binding to the enzyme substrate. The first substrate-binding domain discovered was bound to cellulose, and the binding domain was referred to as a cellulose binding domain (CBD).\textsuperscript{36} CBGs can recognize and bind to both crystalline and amorphous regions of cellulose.\textsuperscript{37} Since the publication detailing the first CBD, many other substrate-binding modules have been discovered that have displayed binding specificity for polysaccharides other than cellulose. This facilitated a name change from CBD to carbohydrate binding module (CBM). CBMs can be located at the N or C terminus of the catalytic enzyme, or within the enzyme itself.

CBMs are classified into 68 different families in the CAZy database, based on their amino acid sequence (further described in section 1.3.1). The CBM families are grouped into three different subtypes that have been recently refined, based on their substrate specificity, structure, and function.\textsuperscript{38} Type A CBMs bind to the crystalline surfaces of insoluble
polysaccharides, including cellulose, and do not display a significant binding affinity for soluble cello-oligosaccharides. Type A CBMs are characterized by a hydrophobic planar surface containing aromatic amino acids that helps facilitate binding to crystalline cellulose. Type B CBMs (endo-type) bind internally to glycan chains. Type C CBMs (exo-type) bind to the non-reducing end of glycan chains.

The major role of CBMs is to increase the concentration of enzyme on the substrate polysaccharides, facilitating the enzymatic reaction and increasing the reaction rate. If a CBM is removed from the enzyme, the enzymatic reaction rate is significantly decreased on insoluble polysaccharides. The CBMs are still able to bind to their substrate with high specificity regardless of whether or not the catalytic enzyme is attached. This enables the use of CBMs as affinity tags for protein purification, and molecular probes to visualize crystalline polysaccharides through functionalization of the CBM with a fluorescent tag.

1.2.3 Degradation of Cellulose

Cellulases are a group of enzymes that work collaboratively to hydrolyze cellulose. Many different microorganisms, including *Humicola insolens* and *Trichoderma reesei*, produce cellulases that are highly specific for crystalline cellulose. In order to hydrolyze cellulose, the cellulose chains must first be separated from the crystalline surface in order to enable the chains to access the cellulase active site, which is the rate-limiting step. This is complicated by the fact that cellulose contains both crystalline and amorphous regions, which can significantly impact the reaction rate of the cellulase enzymes depending on their specificity. Additionally, CBMs
present on the cellulase enzyme can bind non-productively to the cellulose chains, further affecting the reaction rate.

In order to degrade cellulose completely from their polymeric chains to monomeric glucose units, three different types of cellulases work collaboratively. Endocellulases cleave internal linkages in the amorphous regions of the cellulose chain, breaking the chain down to smaller units. Exocellulases (also referred to as cellobiohydrolases) cleave cellulose from either the reducing end (cellobiohydrolase I from \textit{T. reesei}) or non-reducing end (cellobiohydrolase II from \textit{T. reesei}) of the cellulose chains, resulting in the formation of cellobiose. Cellobiosidases (also referred to as $\beta$-glucosidases) degrade cellobiose to $\beta$-glucose units. The specificity and mechanism of $\beta$-glucosidases will be further discussed in section 1.3.

In addition to cellulases, polysaccharide monooxygenases are a class of copper-dependent metalloenzymes that oxidatively cleave glycosidic linkages on the crystalline cellulose surface.\textsuperscript{43} This process avoids the rate-limiting step of having to separate the cellulose chains from the crystalline microfibrils. In order to function, polysaccharide monooxygenases require the presence of cellobiose dehydrogenase (CDH) and molecular oxygen (Figure 1-2). Cellobiose dehydrogenase is an oxidase that catalyzes the two-electron oxidation of cellobiose to cellobiono-1,5-lactone, subsequently reducing oxygen to hydrogen peroxide.\textsuperscript{44} Beeson \textit{et al.} theorized that the combination of both enzymes allows for the insertion of oxygen at either carbon 1 or 4 of glucose, destabilizing the glycosidic linkage between the two glucose units.\textsuperscript{45} This destabilized linkage is susceptible to an irreversible elimination reaction catalyzed by water, resulting in cleavage of the cellulose chain (Figure 1-3).
**Figure 1-2:** General mechanism of copper-dependent polysaccharide monooxygenases.


**Figure 1-3:** Cleavage of cellulose chains using polysaccharide monooxygenases.

Cellodextrin phosphorylase, an enzyme mentioned in section 1.2.1 for its ability to synthesize crystalline cellulose II, is also a candidate for degrading cellulose. Ye et al. demonstrated that through the addition of a family 9 CBM, the activity of cellodextrin phosphorylase could be significantly increased on insoluble amorphous cellulose and celloheptaose. This research may eventually lead to an evolved cellodextrin phosphorylase enzyme that could degrade crystalline cellulose.

Phosphorylation has been shown in nature to assist in the degradation of carbohydrate polymers, such as starch. Starch is an insoluble polymer of repeating D-glucose units connected by α-(1→4) and α-(1→6) glycosidic bonds. In nature, starch is found as granules that consist of 75% amylopectin and 25% amylose. Amylopectin is a semi-crystalline, highly branched polymer of D-glucose units connected by both glycosidic linkages, while amylose exists as an amorphous linear polymer of α-(1→4) glycosidic bonds, similar to cellulose. In starch, two enzymes modulate the phosphate content of amylopectin in starch granules. Glucan water dikinase first phosphorylates the C6 position of the crystalline glucose residues, followed by phosphoglucan water dikinase, which phosphorylates the C3 position. This combined phosphorylation significantly enhances the solubility of the starch granules in vivo, since it disrupts the crystallinity of amylopectin. Disruption of the amylopectin structure enables the hydrolysis enzymes to access the glucose chains. The subsequent dephosphorylation and degradation by hydrolysis enzymes allows for the complete degradation of starch.
1.3 Engineering of Glycosidase Enzymes

Carbohydrates, the most abundant biomolecules on Earth, have a number of essential roles in biological systems. The primary function of carbohydrates is energy storage. Polymeric carbohydrates, such as glycogen in animal cells and starch in plant cells, can be broken down to their monosaccharide constituents for energy. Carbohydrate polymers also have significant structural roles in nature. Cellulose gives plant cell walls their strong, water-insoluble properties, and chitin is the primary component of the hard exoskeletons of arthropods. Carbohydrates are also important in terms of cell recognition and messaging, since they can be $O$-linked or $N$-linked to proteins and other biological molecules, forming glycoconjugates. Carbohydrates also play an important role in protein folding, as they help prevent misfolding, degradation, and aggregation in cells.

Carbohydrates are covalently linked by glycosidic bonds, which are formed between the anomeric carbon of one glycoside (glycone), and a hydroxyl group on the glycoside counterpart (aglycone), as illustrated in Figure 1-4. According to IUPAC, the aglycone is the non-sugar compound remaining if the glycone is replaced by a hydrogen atom. The glycosidic bond is one of the most stable linkages in nature, with the half-life of spontaneous hydrolysis for cellulose and starch estimated to be around five million years.
In nature, carbohydrates are synthesized, catabolized, transferred, and modified by enzymes. One class of enzyme that plays an active role in carbohydrate catabolism are glycosidases. Glycosidases catalyze the hydrolysis of the carbon-oxygen glycosidic bond between the glycone and aglycone (Figure 1-5). Glycosidases are able to speed up the hydrolysis of glycosidic bonds by up to 17 orders of magnitude.

**Figure 1-4**: Glycone and aglycone components of a D-glucoside, along with the carbohydrate numbering scheme used in this thesis.

**Figure 1-5**: General glycosidase-catalyzed hydrolysis reaction. Sugar hydroxyls are omitted for clarity.
1.3.1 Classification of Glycoside Hydrolases (Glycosidases)

Glycosidases can be classified in several different ways:

1. **Substrate Specificity**: Glycosidase active sites have evolved to recognize and catalyze the cleavage of specific glycosidic linkages. Glycosidases can be promiscuous in terms of the glycosidic linkages they hydrolyze, but this activity generally occurs at a lower rate compared to the preferred substrate.

2. **Anomeric Specificity**: Glycosidases exclusively catalyze glycosidic bond cleavage of either α-linked or β-linked glycosides in essentially all cases. GH4 glycosidases are able to hydrolyze natural substrates with different anomeric configurations.56

3. **Stereochemical Outcome of Catalyzed Reaction**: Glycosidases will catalyze glycoside bond hydrolysis with either retention or inversion of stereochemistry at the anomeric center compared to the starting substrate (Figure 1-6).
4. **Cleavage Position of Enzyme**: Glycosidases that cleave the terminal residue of polymeric chains are called exo-glycosidases. Glycosidases that cleave polymeric chains between residues that are not at the terminus are called endo-glycosidases. Endo-glycosidases tend to have long cleft-shaped active sites that allow for interactions with multiple glycosyl residues.

5. **Sequence Homology**: The CAZy (carbohydrate-active enzyme) database (www.cazy.org), developed by BernardHenrissat, has classified glycosidases into 133 families based on amino acid sequence homology. The CAZy classification reflects the evolutionary and structural relationships between glycosidase families. However, glycosidases in the same family can have significantly different substrate specificity.\(^5\)
1.3.2 Catalytic Mechanism of Retaining β-Glycosidases

The catalytic mechanism for retaining β-glycosidases, first proposed by Koshland in 1953, is illustrated in Figure 1-7. The glycosidic bond is cleaved through a double-displacement mechanism, which utilizes two key carboxylic acid residues spaced approximately 5.5 Å apart in the active site. One carboxylic acid functions as the catalytic nucleophile, while the other residue acts as a general acid/base catalyst. In the first step of the mechanism (glycosylation), the catalytic nucleophile attacks the anomeric center of the glycone. The aglycone is concurrently protonated by the acid/base residue. This step proceeds through an oxocarbenium ion-like transition state, resulting in the formation of a covalently bound glycosyl-enzyme intermediate. In the second step (deglycosylation), the acid/base residue deprotonates an incoming water molecule, facilitating its attack on the anomeric center of the glycone. This step also proceeds through an oxocarbenium ion-like transition state, resulting in a product with the same anomeric configuration as the substrate.
1.3.3 Conversion of Glycosidases into Glycosynthases

Due to their importance in biological processes, there is significant interest in synthesizing both natural and unnatural carbohydrates. However, carbohydrates are structurally complex, complicating their synthesis. Major areas of concern include controlling the anomeric configuration of the product, and ensuring that glycosidic linkages are formed between the correct hydroxyls. There are many synthetic procedures that have successfully synthesized complex carbohydrates, including the Koenigs-Knorr reaction, trichloroacetimidate donors,
thioglycoside donors, and the n-pentenyl glycoside method.\textsuperscript{62} However, these methods involve laborious installation and removal of protecting groups, use of toxic solvents and catalysts, and the formation of unwanted side products.\textsuperscript{63}

Due to the challenges in using chemical synthetic strategies, there has been significant interest in using enzymes to synthesize complex carbohydrates. Enzymes allow for control of regioselectivity and stereoselectivity, single-step reactions, faster reaction rates, and environmentally friendly reactions. Retaining β-glycosidases have been used in chemo-enzymatic carbohydrate synthesis under controlled conditions that promote “reversal” of the hydrolysis reaction, yielding a product with the same anomeric configuration as the starting substrate.

There are two approaches that can be used for promoting carbohydrate synthesis using glycosidases.\textsuperscript{64} One approach involves controlling the equilibrium of the reaction by utilizing Le Chatelier’s principle. This method involves using a large excess of glycosyl donor and acceptor, lowering the water concentration through the use of organic co-solvents or salt, or continuously removing the desired product by adsorption or crystallization \textit{in situ}. However, the high concentration of water (55.5 M) favors hydrolysis.

The second approach involves forming products under kinetic control. Activated sugar donors, such as glycosyl fluorides or sugar oxazolines,\textsuperscript{65} form a high steady state concentration of glycosyl-enzyme intermediate, facilitating reaction with the acceptor substrate. This approach, termed transglycosylation, leads to higher yields compared to the equilibrium controlling strategy. However, the transglycosylation products are vulnerable to hydrolysis by the enzyme, limiting the yield of the reaction.
In order to address these limitations, a new class of enzyme was invented, called glycosynthases. Glycosynthases are catalytically inactive mutants of glycoside hydrolases. Through mutation of the catalytic nucleophile to a non-nucleophilic residue (alanine, glycine, or serine), the enzyme is unable to hydrolyze transglycosylation products, significantly increasing the reaction yield. Glycosynthases utilize an activated glycosyl donor, such as an α-fluoride, as illustrated in Figure 1-8. The fluoride anion is a good leaving group, and α-glycosyl fluorides are sufficiently stable in aqueous buffers. This activated donor mimics the glycosyl-enzyme intermediate formed during glycosylation, and is subsequently attacked by the acceptor substrate. This attack is facilitated by deprotonation of the attacking hydroxyl by the acid/base residue in the active site. This results in the synthesis of a glycosidic linkage with the opposite anomeric configuration of the donor glycosyl fluoride.

![Figure 1-8: General mechanism for glycosynthase enzymes with the catalytic nucleophile mutated to a glycine residue.](image)

The first glycosynthase was produced from a family GH1 β-glucosidase from *Agrobacterium sp.* (Abg). The catalytic nucleophile was replaced with an alanine residue
(E358A), producing a catalytically inactive enzyme. When provided with an α-glucosyl fluoride donor and an acceptor substrate, the enzyme was able to catalyze the transglycosylation reaction without detectable hydrolysis of the product. X-ray crystallographic studies confirmed that enzymes with a mutated catalytic nucleophile retained the same active site structure and folding pattern as their wild-type counterparts.68

Glycosynthase enzymes have many advantages and disadvantages. One advantage of glycosynthase enzymes is that the transglycosylation product can also act as an acceptor substrate, increasing the degree of polymerization of synthesized products. This enables the synthesis of carbohydrate polymers such as chitin, xylan, and cellulose.69 However, it is challenging to control the extent of polymerization. Another advantage of glycosynthase enzymes is that they utilize activated glycosyl donors that are relatively inexpensive compared to UDP-sugars, which are utilized by glycosyl transferases.70 For reasons that are not yet apparent, not every glycosidase can be successfully mutated to form a glycosynthase. However, directed evolution approaches can be used to improve weak activities. Glycosynthase enzymes have been successfully used to synthesize a wide variety of carbohydrates.71

1.4 Structure of Glycolipids in Solution

Glycolipids are amphipathic molecules consisting of a hydrophilic head group and a hydrophobic tail. The hydrophilic head groups can be neutral, cationic, anionic, or zwitterionic, depending on the carbohydrate composition. In aqueous solutions, glycolipids spontaneously aggregate together based on the hydrophobic effect. In this, the hydrophobic lipid tails cluster together, while the polar head groups interact with the surrounding water molecules. This
interaction provides the driving force to form different lipid aggregates in solution, depending on the conditions in solution and the nature of the lipids (Figure 1-9).72

Micelles are spherical aggregates of lipids. The hydrophilic head groups are positioned on the surface of the sphere, while the hydrophobic tails are sequestered in the center. Micelles are formed when the lipid concentration in an aqueous solution reaches the critical micelle concentration (cmc), which is defined by a sudden change in surface tension. The cmc varies depending on the pH, ionic strength, temperature, and properties of the glycolipid head group and hydrophobic tail.73 The formation of micelles is favored when the cross-sectional area of the hydrophilic head group is larger than that of the hydrophobic tail. In solution, micelles display a significant degree of polydispersity in terms of their size and aggregation number.

When the cross-sectional area of the hydrophilic head group is similar to that of the hydrophobic tail, the formation of a lipid bilayer is favored. Lipid bilayers consist of two sheets of lipids, where the hydrophilic head groups face the water at each surface, while the hydrophobic tails are sequestered together in the center. However, at the edges of the bilayer, the hydrophobic tails are in contact with water. This causes the lipid bilayers to be unstable, leading to the spontaneous formation of vesicles. A vesicle has the same spherical shape and hydrophilic surface as a micelle, but the center of the sphere is hollow, where the hydrophilic head groups from the bilayer form an aqueous cavity. Vesicles can also be prepared in vitro, as first demonstrated by Bangham in the 1960’s.74 Vesicles prepared in this manner are referred to as liposomes. Liposomes can occur as multilamellar vesicles containing several concentric lipid bilayers, or can be converted to unilamellar vesicles by sonication, extrusion through large pore size filters, ethanol injection, or adaptive focused ultrasound.75
In order to better model lipid bilayers and biological membranes, nanoparticulate phospholipid bilayer discs (referred to as nanodiscs) were created. The nanodiscs consist of a segment of phospholipid bilayer that is held together by a membrane scaffold protein (MSP). The MSP controls the size and dimensions of the lipid bilayer, making the nanodiscs much more monodisperse compared to micelles and liposomes. Nanodiscs are self-assembled by adding MSP to a stabilized mixture of detergent and phospholipids, followed by removal of the detergent through dialysis or addition of a hydrophobic adsorbent (Figure 1-10). While
nanodiscs have been primarily assembled to study membrane proteins, they have also been used to incorporate cell surface receptors, such as glycosphingolipids, to study receptor-protein interactions.\textsuperscript{79}

![Formation of nanodiscs by mixing MSP with detergent solubilized phospholipids, followed by removal of the detergent.](image)

\textbf{Figure 1-10}: Formation of nanodiscs by mixing MSP with detergent solubilized phospholipids, followed by removal of the detergent. Reproduced from Ref. 78 with permission from The Royal Society of Chemistry.

\subsection{1.4.1 Properties of Glycosphingolipids}

Sphingolipids were first discovered by Johan Thudichum in ethanolic brain extracts in 1884, and were named after the mythological Sphinx due to their enigmatic nature.\textsuperscript{80} The two key roles of sphingolipids are to act as messengers in a multitude of signaling pathways, and to define the physical properties and composition of cell membranes.\textsuperscript{81} Sphingolipids are distinguished by their amphiphilic sphingosine backbone, illustrated in Figure 1-11. The
sphingosine backbone contains an amino group (pKa = 6.6) that improves sphingolipid solubility in aqueous solutions, and leads to the formation of micelles and larger aggregates depending on the pH and degree of intermolecular hydrogen bonding. The amino group can also be N-acylated through the formation of an amide linkage, forming ceramides that enhance membrane stability.

Figure 1-11: Structure of sphingosine backbone ((2S,3R,4E)-2-amino-4-octadecene-1,3-diol), along with the labeling scheme used in this thesis.

Sphingolipids can also be glycosylated, forming glycosphingolipids. Glycosphingolipids containing sialic acid residues are referred to as gangliosides. Both glycosphingolipids and gangliosides are predominantly located in cell membranes and have many critical physiological roles, including biochemical signaling, lipid raft organization, and pathogen engagement. Glycosphingolipids show promise in promoting the regeneration of nerve tissue, which may help alleviate the symptoms of Alzheimer’s disease, Parkinson’s disease, cancer, diabetes, and autoimmunity in transplant cases. Therefore, there is significant interest in the isolation and purification of natural glycosphingolipids. However, glycosphingolipids are very difficult to isolate and purify from natural sources, and are typically only available in small quantities. Further, especially in the wake of mad cow disease and associated neuropathies, there is concern
about working with such materials, as their extraction may expose the researcher to infectious diseases. Therefore, there is significant interest in synthesizing glycosphingolipids, either chemically or chemo-enzymatically.\(^{86}\)

### 1.5 Aims of this Thesis

As mentioned in sections 1.1.1 and 1.1.2, polymeric cellulose I is challenging to work with due to its insolubility in water and common solvents, along with the inability to crystallize cellulose I \textit{in vitro} without forming cellulose II. Therefore, there is significant interest in finding methods to synthesize cellulose I \textit{in vitro}. The synthesis of cellulose I would enable the study of carbohydrate binding domains and other proteins that interact with cellulose in nature. Cellulose I naturally exists in crystalline microfibrils. Therefore, in order to study the interactions of proteins and cellulose, crystalline cellulose I must be utilized. Previous chemoenzymatic attempts to synthesize cellulose I did not produce cellulose with the desired crystalline properties (section 1.2.1). Therefore, a strategy must be found to synthesize cellulose I in a controlled manner.

In aqueous solutions, the preferential arrangement of lipids is a lipid bilayer (section 1.4). The hydrophilic head groups are aligned along a planar surface, while the hydrophobic tails are sequestered in the center. From this, if a cello-oligosaccharide is attached to a lipid, such as sphingosine, the cellulolipid should spontaneously form lipid aggregates in solution. Due to the hydrophobicity of the lipid compared to cellulose, the cellulolipids could be aligned along an aqueous:organic interface. The lipid tail would preferentially be soluble in the organic layer, while the cellulose head group would be soluble in the aqueous layer. The cellulolipids could
also be incorporated into a nanodisc, which models a lipid bilayer. Either method would result in the formation of cellulose I, since the cellulose chains would be aligned in a parallel orientation. This surface of cellulose I would be free to interact with proteins and cellulose binding domains.

The challenge involved with cellulolipid synthesis and alignment is that cello-oligosaccharides with a DP of 6 or higher are insoluble in aqueous solution. One method that can be used to improve cello-oligosaccharide solubility is the addition of a functionality that is charged at neutral pH. By improving the solubility, the synthesis of longer cello-oligosaccharides and cellulolipids can be accomplished.

The focus of this thesis will be synthesizing cellulolipids with a cellohexaose head group. Cellohexaose is the shortest cello-oligosaccharide that has the same structural properties as cellulose, and consequently should be able to interact with cellulose binding proteins of interest. Cellohexaose analogs will be synthesized that will be functionalized with either a terminal sialic acid residue, or a phosphate on the terminal non-reducing D-glucose residue. Once the charged cellulolipids are synthesized chemoenzymatically, the charged cellulolipids will be aligned along an aqueous:organic interface, or incorporated into a nanodisc. Once the charged cellulolipids are aligned, the charged functionalities will be removed enzymatically, producing a surface of crystalline cellulose I that can be used to study proteins of interest.
Chapter 2: Synthesis of Soluble Cellulosic Oligosaccharides Through
Introduction of Charged Functionalities

2.1 Synthetic Strategy

The synthetic strategy for this thesis is outlined in Figure 2-1. The focus of chapter 2 is to
explore methods to synthesize longer cello-oligosaccharides using glycosynthase enzymes. First,
the glycosyl fluoride donors required by glycosynthase enzymes will be synthesized chemically
(section 2.2). Once synthesized, Abg2F6 glycosynthase will be used to extend the length of the
cello-oligosaccharyl fluorides (section 2.3). Cello-oligosaccharides longer than five glucose units
are insoluble in aqueous solution. Therefore, Abg2F6 glycosynthase should be able to synthesize
soluble cello-oligosaccharyl fluorides from DP = 2 to DP = 5.

To improve cello-oligosaccharyl fluoride solubility, a removable charged group will be
added onto the non-reducing end. Two different charged functionalities will be presented: a
terminal sialic acid (section 2.4), and a phosphate at the 6 position of the non-reducing terminal
glucose (section 2.5). The ability of the charged cello-oligosaccharyl fluorides to be transferred
to aryl glycoside acceptors by endo-glycosynthase enzymes will be explored (section 2.7). Once
optimized, the synthesis of soluble cellohexaose analogs will be investigated (section 2.8).

In chapter 3, cello-oligosaccharide fluorides will be transferred to d-erythro-sphingosine
using EGCase glycosynthase, yielding cellulolipids (section 3.1). Once the cellulolipids are
synthesized, endo-glycosynthase enzymes will be used to synthesize soluble cellulolipids with a
DP of 6 (section 3.3). Once the soluble cellulolipids are prepared, they will be aligned along an
aqueous:organic interface (section 3.7), or incorporated into a nanodisc (section 3.8). The terminal charged groups of the cellulolipids will be removed enzymatically, which should yield a surface of cellulose I.

In chapter 4, experimental methods, characterization, and glycosynthase enzyme expression details will be fully described.

$$
\text{GlcF} + \text{Glc}_2\text{F} \xrightarrow{\text{Abg2F6 Glycosynthase}} \text{Glc}_3\text{F} \xrightarrow{\text{BglK}} \text{P-Glc}_3\text{F} \\
\text{P-Glc}_3\text{F} + \text{Glc}_3\text{Sph} \xrightarrow{\text{Endo Glycosynthase}} \text{P-Glc}_6\text{Sph} \xrightarrow{\text{Alkaline Phosphatase}} \text{Glc}_6\text{Sph}
$$

**Figure 2-1:** Example of proposed strategy to synthesize cellulolipids.

### 2.2 Synthesis of α-Glycosyl Fluorides

α-Glucosyl fluoride (2.04) was prepared according to Scheme 2-1. Commercially available d-glucose (2.01) was acetylated using acetic anhydride (Ac$_2$O) in the presence of sodium acetate (NaOAc), yielding d-glucose penta-acetate (2.02, 65%). Sodium acetate catalyzes rapid anomerization of the free sugar, leading to acetylation of the kinetically favored β-
anomer. The per-\(O\)-acetylated sugar was reacted with hydrogen fluoride pyridine (HF-pyridine) overnight at 4°C, producing per-\(O\)-acetylated \(\alpha\)-glucosyl fluoride (\textbf{2.03}, 95%), which is the thermodynamically favored anomer. In order to remove the acetate groups, ammonia gas was slowly bubbled into a suspension of \textbf{2.03} in dry MeOH over 10 minutes. The reaction was left overnight at 4°C, yielding \textbf{2.04} (88% yield).

\[ \textbf{2.01} \xrightarrow{i} \textbf{2.02} \xrightarrow{\text{ii}} \textbf{2.03} \xrightarrow{\text{iii}} \textbf{2.04} \]

\textbf{Scheme 2-1}: Synthesis of \(\alpha\)-glucosyl fluoride (\textbf{2.04}). i) NaOAc (1.1 eq.), Ac\(_2\)O (3 eq./OH), reflux, 1 hour, 65%; ii) HF-pyridine (45 eq.), 4°C, overnight, 95%; iii) NH\(_3\) (g), MeOH, 4°C, overnight, 88%.

Scheme 2-1 was repeated for commercially available starting materials D-galactose, D-cellobiose octaacetate, and D-lactose, with the results listed in Table 2-1. Full characterization of all listed compounds is found in section 4.3. For storage, \(\alpha\)-glycosyl fluorides are stable for a few months in the freezer as crystalline solids. However, it is best to store them in their per-\(O\)-acetylated form to minimize hydrolysis of the \(\alpha\)-fluoride.
<table>
<thead>
<tr>
<th>Starting Material</th>
<th>i</th>
<th>ii</th>
<th>iii</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-Glucose (2.01)</td>
<td>65%</td>
<td>95%</td>
<td>88%</td>
</tr>
<tr>
<td></td>
<td>(2.02)</td>
<td>(2.03)</td>
<td>(2.04)</td>
</tr>
<tr>
<td>D-Galactose (2.05)</td>
<td>42%</td>
<td>59%</td>
<td>66%</td>
</tr>
<tr>
<td></td>
<td>(2.06)</td>
<td>(2.07)</td>
<td>(2.08)</td>
</tr>
<tr>
<td>D-Cellobiose (2.09)</td>
<td>N/A</td>
<td>87%</td>
<td>90%</td>
</tr>
<tr>
<td></td>
<td>(2.10)</td>
<td>(2.11)</td>
<td>(2.12)</td>
</tr>
<tr>
<td>D-Lactose (2.13)</td>
<td>61%</td>
<td>73%</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>(2.14)</td>
<td>(2.15)</td>
<td>(2.16)</td>
</tr>
</tbody>
</table>

**Table 2-1:** Results for the synthesis of α–glycosyl fluorides.

### 2.3 Use of Abg2F6 Glycosynthase in Oligosaccharyl Fluoride Synthesis

One method used to synthesize cello-oligosaccharides is to utilize a glycosynthase enzyme and an α–glycosyl fluoride donor. *Agrobacterium* sp. β-glucosidase is a family 1 retaining glycosidase that hydrolyzes the terminal, non-reducing end of cello-oligosaccharides and simple β-glucosides. Wild-type Abg also functions as a β-galactosidase, and displays transglycosylation activity at higher substrate concentrations. The first glycosynthase mutant of Abg (E358A) successfully transferred donor substrates α–glucosyl fluoride (GlcF, 2.04) and α–galactosyl fluoride (GalF, 2.08) to a variety of aryl glycoside acceptors, with yields ranging from 64-92%, and no evidence of product hydrolysis.

In order to find improved glycosynthase mutants of Abg, random mutagenesis was used. By randomly mutating the catalytic nucleophile of Abg, three additional glycosynthases were discovered (E358C, E358G, and E358S). E358G demonstrated 50-fold higher activity compared
to the original E358A mutant. Two subsequent rounds of random mutagenesis on Abg E358G yielded the glycosynthase mutant Abg2F6. Abg2F6 has four mutations; A19T, E358G, Q248R, and M407V. Mutations A19T and M407V are postulated to change the conformation of the enzyme active site, increasing the library of acceptor substrates that can be utilized by Abg2F6.

Using Abg2F6, GalF was transferred to α–celllobiosyl fluoride (2.12). GalF has an axial 4-hydroxyl, which is not an acceptor substrate for Abg2F6. Using an acceptor concentration of 20 mM and 1.5 equivalents of donor GalF, 2.17 was synthesized overnight (Scheme 2-2). TLC and mass spectrometry (529.2 m/z, [M+Na]+) confirmed the synthesis of 2.17.

![Scheme 2-2: Synthesis of 2.17 using Abg2F6 glycosynthase.](image)

### 2.3.1 Synthesis of Longer Cellulosic Glycosyl Fluorides

In order to synthesize longer cello-oligosaccharides, Abg2F6 was used to transfer GlcF to α–celllobiosyl fluoride (2.12). The product of a single GlcF addition is an excellent acceptor substrate for Abg2F6, allowing the transfer of further GlcF molecules (Scheme 2-3).
of $\alpha$–cellotriosyl fluoride (2.18) and $\alpha$–cellotetraosyl fluoride (2.19) was confirmed by TLC and mass spectrometry (529.2 m/z ([M+Na]$^+$) and 691.2 m/z ([M+Na]$^+$), respectively).

![Scheme 2-3: Synthesis of 2.18 and 2.19 using Abg2F6 glycosynthase.](image)

Changing the number of equivalents of donor GlcF used in the reaction can control the ratio of synthesized trisaccharide and tetrasaccharide products. However, even when 5 equivalents of GlcF are added to the Abg2F6 reaction, $\alpha$–cellobiosyl fluoride is never completely consumed. Also, while there is evidence of $\alpha$-cellopentaosyl fluoride in the reaction mixture according to mass spectrometry (853.5 m/z, [M+Na]$^+$), it was not produced in isolable quantities. Addition of 10 equivalents of GlcF resulted in the synthesis of insoluble cello-oligosaccharyl fluorides. The precipitate was marginally soluble in $N,N$-dimethylformamide (DMF). To maximize the yield of 2.19, 5 eq. of GlcF was used for the synthesis of cello-oligosaccharide fluorides using Abg2F6 glycosynthase.
2.3.2 Strategies to Purify Glycosyl Fluoride Mixtures

In order to utilize $\alpha$-cellotriosyl fluoride and $\alpha$-cellotetraosyl fluoride in further glycosynthase reactions, they have to be isolated from contaminating $\alpha$-cellobiosyl fluoride and other impurities. Glycosyl fluorides are stable for a short time in buffered aqueous solutions, but hydrolyze rapidly in strongly acidic and basic solutions. Therefore, any purification strategy used cannot expose the glycosyl fluorides to extreme conditions.

2.3.2.1 Size Exclusion Chromatography

Size exclusion chromatography is a technique that separates molecules and polymers according to their size. Separation is achieved through the migration of molecules through the pores of a column packed with cross-linked polymer beads. Large molecules typically cannot enter the pores and will migrate around the polymer beads, shortening their path through the column. Small molecules are able to enter the pores between the polymer beads, increasing their travel time through the column compared to large molecules. The polymer beads are non-adsorptive, and do not modify or react with the molecules during the separation.

To purify the glycosyl fluorides, a size exclusion column packed with Sephadex™ LH-20 (GE Healthcare®) was used. Sephadex LH-20 consists of a beaded, cross-linked hydroxypropylated dextran, which has both hydrophilic and lipophilic character. Sephadex LH-20 is used extensively in the separation of natural products, as it has been demonstrated to separate closely related molecules, such as epimers of budesonide, a glucocorticoid steroid.
The ability of Sephadex LH-20 to separate α–glycosyl fluorides was first demonstrated on crude 2.17 that was synthesized using Abg2F6 glycosynthase. According to TLC and mass spectrometry, the only compounds present in the reaction mixture were 2.17, leftover GalF (2.08), and sodium phosphate buffer salts. To remove the insoluble buffer salts, the reaction was suspended in 0.5 mL MeOH, sonicated for 10 minutes, and centrifuged to remove any precipitates. The solution was added to the column (50 mm x 600 mm), and the separation was performed at a flow rate of 0.5 mL/min using MeOH as an eluent. After 100 mL, 1 mL fractions were collected. Compound 2.17 started eluting from the column at 117 mL, while GalF started eluting at 135 mL. 35 mg of 2.17 was collected after three columns, with purity confirmed by NMR (section 4.3), TLC, and mass spectrometry.

Based on this success, the next step was to attempt purification of crude 2.18 and 2.19 synthesized using Abg2F6 glycosynthase. Using the same procedure to remove buffer salts and the same column conditions, elution of 2.19 started at 116 mL, while elution of 2.18 started at 124 mL. However, the elution of each compound occurred over multiple fractions. Therefore, a mixture of 2.18 and 2.19 was collected after 124 mL. Contaminating 2.12 began to elute at 138 mL. The fractions containing a mixture of 2.18 and 2.19 were concentrated in vacuo and the separation was attempted again using a smaller sample. However, a similar result occurred. While the Sephadex LH-20 column is able to remove cellobiosyl fluoride, only 8 mg of 2.19 was collected after two columns, with purity confirmed by NMR (section 4.3), TLC, and mass spectrometry.

Even though the Sephadex LH-20 column was able to separate the glycosyl fluorides, the amount of material collected after each separation was not sufficient for further studies in this
thesis. Therefore, other methods were examined to isolate the glycosyl fluorides in larger amounts.

2.3.2.2 Acetylation and Flash Column Chromatography

Per-O-acetylation is a common method used to protect carbohydrate hydroxyl groups. Per-O-acetylation lowers the polarity of the carbohydrates, increasing their solubility in organic solvents and enabling their purification by flash column chromatography. Removal of the acetate groups is easily achieved under basic conditions (see section 2.2).

One of the most common reagents used to acetylate carbohydrates is acetic anhydride, with pyridine as the base catalyst. Under these conditions, the reaction is faster than carbohydrate mutarotation, producing the same anomeric distribution as the starting material. If the reaction is slow, 0.01 mol% of 4-dimethylaminopyridine (DMAP) can be added as a catalyst to speed up or force the acetylation.

This separation strategy was first demonstrated on crude 2.17 synthesized using Abg2F6 glycosynthase. The insoluble buffer salts were removed (see section 2.3.2.1), and the carbohydrates were suspended in dry pyridine. Three equivalents of Ac₂O were added for every hydroxyl on the molecule, and the reaction mixture was left to stir at room temperature overnight (Scheme 2-4). Once complete, the reaction was quenched with MeOH and worked up to remove excess pyridine and pyridinium acetate. Crude 2.20 was purified using a flash chromatography column packed with silica gel, and an eluent of 1.5:1 EtOAc:PE. TLC, NMR (section 4.3), and mass spectrometry confirmed purity of 2.20 collected from the column. Compound 2.20 was deacetylated using ammonia gas (section 2.2).
This strategy was repeated for the purification of 2.18 and 2.19 from the Abg2F6 reaction. The insoluble buffer salts were removed, and the carbohydrates were acetylated using the same reaction conditions, yielding per-O-acetylated α-cellotriosyl fluoride (2.21) and per-O-acetylated α-cellotetraosyl fluoride (2.22), along with contaminating 2.11. The mixture of per-O-acetylated glycosyl fluorides was purified using a flash column chromatography column packed with silica gel, and an eluent of 1:1 EtOAc:PE. Once 2.11 started to elute from the column, the eluent strength was increased to 1.5:1 EtOAc:PE. These conditions enabled complete separation of 2.21 and 2.22, which was confirmed by TLC, NMR (section 4.3), and mass spectrometry. The acetate groups were removed from both compounds separately using ammonia gas (section 2.2). Purification of the α-glycosyl fluorides by flash column chromatography involves laborious acetylation and deacetylation steps. However, both the reactions and column size can be scaled up considerably, enabling purification on the gram scale. It is unfeasible to scale up the size exclusion column detailed in section 2.3.2.1, due to the cost that would be associated with purchasing additional Sephadex LH-20, and the size of column required. The acetylation and flash chromatography column purification strategy was used for the remainder of this thesis for α-glycosyl fluorides.
2.4 Sialic Acid Transfer to Glycosyl Fluorides Using Cst-I

Sialic acid, also referred to as N-acetylneuraminic acid (Neu5Ac), is an acidic sugar that is a major component of many glycolipids and glycoproteins (Figure 2-2). Glycoconjugates containing sialic acid are involved in cell-cell recognition, microbial infection, and cell differentiation. Sialic acids are commonly located on the non-reducing terminus of cell surface glycoconjugates, such as gangliosides.

![Figure 2-2: Structure of sialic acid (α-anomer) and the numbering scheme used in this thesis.](image)

Sialic acid is transferred from an activated sugar donor to an acceptor substrate by sialyltransferase enzymes. The activated sugar donor is synthesized using CMP-sialic acid synthetase, an enzyme found in Neisseria meningitidis. CMP-sialic acid synthetase utilizes cytidine triphosphate (CTP) and Neu5Ac to synthesize CMP-Neu5Ac and pyrophosphate.

Cst-I from Campylobacter jejuni is an α-2,3-sialyltransferase that solely utilizes acceptor substrates with a terminal galactose residue. Therefore, in order to transfer sialic acid onto cello-oligosaccharyl fluorides, the carbohydrate chain first has to be “capped” with a galactose...
residue using Abg2F6 glycosynthase. The transfer of negatively charged sialic acid onto the carbohydrates will potentially improve their solubility in aqueous solution.

To demonstrate the ability to synthesize sialic acid terminated glycosyl fluorides, Cst-I (0.20 mg/mL) was used to transfer sialic acid from CMP-Neu5Ac (1.3 eq., 2.23) to lactosyl fluoride (2.16, Scheme 2-5). The reaction was performed in the presence of HEPES buffer (50 mM, pH 7.5) and manganese(II) chloride (10 mM, MnCl₂). Alkaline phosphatase (0.1 µL, 10 KU solution) was added to drive the reaction equilibrium towards product formation, through the conversion of CMP to cytidine. This also minimizes potential product inhibition by CMP-Neu5Ac. The reaction was complete after 2 hours, yielding sialylated lactosyl fluoride (2.24, 634.2 m/z, [M]).

Scheme 2-5: Synthesis of 2.24 using Cst-I enzyme.
In order to synthesize a longer sialylated glycosyl fluoride, Abg2F6 glycosynthase was first used to transfer GalF to $\alpha$-cellobiosyl fluoride (section 2.3). The reaction was complete overnight and lyophilized, yielding crude 2.17, which was purified by flash column chromatography (section 2.3.2.2). Next, Cst-I was used to transfer sialic acid from CMP-Neu5Ac to 2.17, using the same reaction conditions mentioned above (Scheme 2-6). The reaction was complete in 2 hours, yielding 2.25 (796.3 m/z, [M$^{+}$]).

Scheme 2-6: Synthesis of 2.25 using Cst-I enzyme. i) Abg2F6 glycosynthase (0.10 mg/mL), 100 mM sodium phosphate (pH 7.0), RT, overnight; ii) Cst-I (0.20 mg/mL), 50 mM HEPES (pH 7.5), 10 mM MnCl$_2$, RT, 2 hours.
Next, an attempt was made to synthesize the sialylated pentaosyl fluoride. Abg2F6 glycosynthase was used to transfer GalF (34.0 mg, 1.2 eq., 2.08) to α-cellootriosyl fluoride (76.5 mg, 2.18), yielding the Gal-terminated tetrasaccharide fluoride (2.26, Scheme 2-7). TLC and mass spectrometry (691.4 m/z, [M+Na]⁺) confirmed the synthesis of 2.26. The crude tetrasaccharide was acetylated and purified using the same strategy detailed in 2.3.2.2, yielding 60 mg of per-O-acetylated 2.26 as a clear oil. Unfortunately, upon removal of the acetate groups, only 10 mg of 2.26 remained, which was contaminated with 2.18. The synthesis and purification was repeated, but similar results were obtained. Therefore, a different strategy to improve the solubility of longer glycosyl fluorides was required.

Scheme 2-7: Synthesis of 2.26 using Abg2F6 glycosynthase.
2.5 Phosphorylation of Glycosyl Fluorides Using BglK

As mentioned in section 1.2.3, phosphorylation of crystalline amylopectin by the combined activity of glucan water dikinase and phosphoglucan water dikinase significantly increases the solubility of both amylopectin and starch granules.\(^{48}\) From this observation, phosphorylation could be used to significantly improve the solubility of longer celloligosaccharides, particularly cellobiose.

An ATP-dependent β-glucoside kinase (BglK) from *Klebsiella pneumoniae* was able to phosphorylate the 6’ position of cellobiose and other disaccharide substrates in 2 hours at room temperature.\(^{103}\) To demonstrate whether BglK can phosphorylate glucosyl fluorides, α-cellobiosyl fluoride (2.12) was used as a test substrate. 2.12 was dissolved in 25 mM HEPES (pH 7.5) containing 2 mM MgSO\(_4\), and 1.2 eq. of ATP was added (pH 7.5, adjusted with 3 M NH\(_4\)OH). BglK (0.50 mg/mL) was added, and the reaction was incubated at room temperature (Scheme 2-8). The reaction was complete after 2.5 hours according to TLC (4:5:1.5 CHCl\(_3\):MeOH:0.2% CaCl\(_2\), Rf = 0.36) and mass spectrometry (423.1 m/z, [M+H]+).

![Scheme 2-8: Synthesis of 6’-O-phospho-cellobiosyl fluoride (2.27) using BglK.](image)
BglK successfully phosphorylated 2.12. As a result, its ability to phosphorylate longer cello-oligosaccharide fluorides, such as α-cellotriosyl fluoride (2.18) and α-cellotetraosyl fluoride (2.19), was tested. Using an increased amount of BglK (0.75 mg/mL), and the same reaction conditions used to synthesize 2.27, 6''-O-phospho-cellotriosyl fluoride (2.28) and 6''''-O-phospho-cellotetraosyl fluoride (2.29) were synthesized overnight. Mass spectrometry confirmed the synthesis of 2.28 (585.1 m/z, [M+H]+) and 2.29 (747.2 m/z, [M+H]+). It is evident that phosphorylation is a strategy to pursue further.

2.5.1 Cleavage of Phosphate Group

To confirm that the phosphate could be removed from the cello-oligosaccharide fluorides, dephosphorylation of 2.27, 2.28, and 2.29 was performed to see if the starting material would be generated. By generating 2.12, 2.18, and 2.19, this would solidify phosphorylation as a strategy to improve the solubility of cello-oligosaccharides in aqueous solution. Solutions containing 2.27, 2.28, or 2.29 were diluted with 200 mM Tris (pH 8.6) and 50 mM MgCl₂ to a final substrate concentration of 10 mM. Alkaline phosphatase (1 µL, 10 KU solution) was added, and the reaction was incubated at room temperature (Scheme 2-9).

![Scheme 2-9: Cleavage of phosphate functionality on 2.28 using alkaline phosphatase.](image-url)
After 15 minutes, all three reactions were complete (determined by TLC, using an eluent of 4:5:1.5 CHCl₃:MeOH:0.2% CaCl₂). Success of the reaction containing 2.28 was measured by an increase in the Rf value from 0.30 to 0.62, indicating cleavage of phosphate generated the starting material (2.18). This change in Rf was also observed in the reactions containing 2.27 (0.36 to 0.75), and 2.29 (0.28 to 0.58). This experiment confirms that the phosphorylated cello-oligosaccharide fluorides can be easily dephosphorylated in one step.

2.6 Exploration of Glycoside Acceptors for Testing Glycosynthase Enzymes

One strategy to synthesize longer cello-oligosaccharides enzymatically involves using a glycoside acceptor that has an anomeric protecting group. Through anomeric protection, there would be more control over the reaction, since the glycoside acceptor does not function as a donor. If the anomeric protecting group could be directly replaced by an α-fluoride, or cleaved to yield the cello-oligosaccharide, this would be advantageous in synthesizing longer cello-oligosaccharides.

2.6.1 Synthesis of DNPC

Aryl glycoside acceptors are known to bind well in the active site of glycosidases, without subsequent hydrolysis by glycosynthase mutants. Aryl glycoside acceptors previously studied contain an anomeric para-nitrophenyl or 2,4-dinitrophenyl group. The synthesis of 2,4-dinitrophenyl β-cellobioside (DNPC) was performed using the method reported by Namchuk et al. (Scheme 2-10). The anomeric position of cellobiose
octaacetate (2.10) was selectively deprotected using hydrazine acetate, yielding 2.30 (97% yield). Next, 1-fluoro-2,4-dinitrobenzene (DNFB, 2.31) was coupled to 2.30 using 1,4-diazabicyclo[2.2.2]octane (DABCO) as a catalyst. The reaction was complete overnight, yielding per-O-acetylated DNPC (2.32, 85%). The acetate groups were removed using acetyl chloride (AcCl) in MeOH over 3 days, yielding DNPC (2.33) with 72% yield. Full characterization of 2.30, 2.32, and 2.33 is found in section 4.3.

Scheme 2-10: Synthesis of DNPC (2.33). i) Hydrazine acetate (1.2 eq.), DMF, RT, 4 hours, 83%; ii) DABCO (4 eq.), DMF, RT, 20 hours, 67%; iii) AcCl (45 eq.), MeOH, 4°C, 72 hours, 72%.

GalF (2.08, 1.5 eq.) was mixed with DNPC (2.33), and dissolved in 100 mM sodium phosphate (pH 7.0) to an acceptor concentration of 20 mM. Abg2F6 (0.10 mg/mL) was added, and the reaction was incubated at room temperature. The reaction was complete overnight.
according to TLC (7:2:1 EtOAc:MeOH:H$_2$O). The new UV-active spot (Rf = 0.23) confirmed formation of the trisaccharide (2.34).

![Chemical Structures](image)

**Scheme 2-11**: Reaction of GalF (2.08) and DNPC (2.33) using Abg2F6 glycosynthase to synthesize 2.34.

### 2.6.1.1 Cleavage of DNP Group

DNPC is a good aryl glycoside acceptor for Abg2F6 glycosynthase. There is the possibility that after extension with Abg2F6 glycosynthase, the anomeric dinitrophenyl group can be cleaved from the product, yielding an extended cello-oligosaccharide. HF-pyridine was used to see if the dinitrophenyl group on per-O-acetylated DNPC (2.32) could be replaced with an α-fluoride (Scheme 2-12).
Scheme 2-12: Cleavage of DNP group using HF-pyridine.

On the first attempt, the reaction was run at 4°C overnight, which are the same conditions used in section 2.2. After 24 hours, the only spot visible on the TLC plate was the starting material (2.32). The same reaction was incubated at room temperature overnight, but the starting material decomposed.

Since the DNP group is stable in the presence of HF, two other strategies were considered. One strategy involves thiolysis, where a thiol is used to cleave the DNP group in a mechanism similar to hydrolysis. A second strategy involved cleavage using piperidine (20% v/v in DMF), which is the same reagent used to remove 9-fluorenylethoxycarbonyl (Fmoc) amino acid protecting groups. Three test reactions with DNPC were set up on the 25 mg scale in 1 mL DMF with the following reagents: 2-mercaptoethanol, thiophenol, and piperidine (Scheme 2-13). The reactions were incubated at room temperature for 42 hours.
Scheme 2-13: Attempted cleavage of DNP group using three different reagents. a) 2-mercaptoethanol (59 eq.), DMF:pyridine (10:1), RT, 42 hours; b) thiophenol (39 eq.), DMF, RT, 42 hours; c) piperidine (20% v/v), DMF, RT, 42 hours.

After 42 hours, the reactions were checked by TLC (7:2.5:1 EtOAc:MeOH:H₂O). Reaction (a) with 2-mercaptoethanol was the most successful, with all starting material converted to 2.09 (Rf = 0.19). Reaction (b) with thiophenol also produced 2.09 (Rf = 0.19), but there was a significant amount of 2.33 remaining in the reaction (Rf = 0.63). In reaction (c), neither 2.09 or 2.33 were detected. However, there was a new spot with a Rf value of 0.42. The TLC spot was theorized to belong to a disaccharide with the DNP group replaced by piperidine (2.35, Figure 2-3).

Figure 2-3: Theorized product synthesized from the reaction of DNPC (2.33) and piperidine.
2-Mercaptoethanol was able to cleave the DNP group from 2.33 after 42 hours. This was longer than anticipated, given that 59 equivalents of thiol were used. While DNPC is still a great acceptor for Abg2F6 glycosynthase, other acceptor substrates were explored.

### 2.6.2 Synthesis of Alkoxyamino Glycoside Acceptors

As mentioned in section 2.6.1, glycosynthase enzymes are able to utilize aryl glycoside acceptors. An aryl group can be transferred to the anomeric position of a glycoside using an oxime. Oximes are readily synthesized from the condensation of hydroxylamine and an aldehyde, and selectively react with the anomeric position of glycosides, without the use of protecting groups. The oxime functionality can also be removed from the deprotected glycosides. If the alkoxyamino glycosides are good acceptors for glycosynthase enzymes, they would be good candidates for extending cello-oligosaccharides.

The alkoxyamino glucoside was synthesized according to the method of Teze et al. (Scheme 2-14). Commercially available O-benzylhydroxylamine (2.36) was reacted with 3 eq. of propionaldehyde (2.37) in the presence of NaOAc, yielding the O-benzyl-oxime (2.38) after 2 hours (97% yield). Next, 2.38 was reduced with 3.9 eq. of sodium cyanoborohydride (NaBH₃CN) acidified to pH 3.0 with concentrated HCl, yielding 85% of 2.39 after 30 minutes. 2.39 was coupled to D-glucose (2.01) in 3:1 THF:AcOH over 20 hours at room temperature, yielding 2.40 (58% yield). Full characterization of 2.38, 2.39, and 2.40 can be found in section 4.3.

The ability of the alkoxyamino glucoside to act as an acceptor for glycosynthase enzymes was demonstrated using Abg2F6 glycosynthase. GalF (2.08, 1.2 eq.) and 2.40 were dissolved in
100 mM sodium phosphate (pH 7.0) to a final acceptor concentration of 10 mM, and Abg2F6 glycosynthase (0.50 mg/mL) was added (Scheme 2-15). The reaction was incubated at room temperature for 3 days. TLC (7:2:1 EtOAc:MeOH:H₂O) confirmed that the reaction contained excess GalF (Rf = 0.60), and disaccharide product (2.41, Rf = 0.20). Formation of 2.41 was also confirmed by mass spectrometry (512.3 m/z, [M+Na]⁺).

Scheme 2-14: Synthesis of alkoxyamino β-D-glucoside (2.40). i) NaOAc (1.2 eq.), 4:1 H₂O:MeOH, RT, 2 hours, 97%; ii) NaBH₃CN (3.9 eq.), EtOH, RT, 30 min, 85%; iii) 75:25 THF:AcOH, RT, 20 hours, 58%.
Scheme 2-15: Synthesis of 2.41 using Abg2F6 glycosynthase.

The alkoxyamino β-D-glucoside acceptor (2.40) was utilized by Abg2F6 glycosynthase. Based on this result, the alkoxyamino cellobioside acceptor (2.42) was synthesized (Scheme 2-16). D-Cellobiose (2.09) was dissolved in 10 mL of 100 mM NaOAc (pH 4.5), and 4 eq. of 2.39 was added. The reaction was incubated at room temperature for 72 hours, and monitored by TLC (7:2:1 EtOAc:MeOH:H₂O). While formation of 2.42 was evident by TLC (Rf = 0.63) and mass spectrometry (490.2 m/z, [M+H]⁺), significant amounts of 2.09 and 2.39 remained in the reaction. Addition of excess 2.39 and NaOAc buffer did not improve the yield of 2.42 (14%). Attempts to perform the reaction using the same conditions as step iii in the synthesis of 2.40 were unsuccessful, due to the limited solubility of 2.09 in THF.
Scheme 2-16: Synthesis of alkoxyamino celllobioside acceptor (2.42).

2.6.2.1 Cleavage of Alkoxyamino Functionality

According to Teze et al., the anomic alkoxyamino functionality can be removed in aqueous solution that is acidified to pH 2, yielding the deprotected glycoside. To test this claim, both 2.41 and 2.42 were dissolved in water, and the reactions were acidified to pH 2 by adding 10 drops of 6 M HCl (Scheme 2-17). The reactions were incubated at room temperature, and monitored every 5 minutes by TLC (7:2:1 EtOAc:MeOH:H2O). After 5 minutes, both 2.41 and 2.42 were converted to their corresponding hemiacetals (2.13 and 2.09, respectively). The reactions were analyzed by mass spectrometry, confirming formation of the deprotected glycoside (365.3 m/z, [M+Na]$^+$) and 2.39 (166.4, [M+H]$^+$). The alkoxyamino functionality can be removed in 5 minutes, which is much more efficient than cleavage of the DNP functionality (section 2.6.1.1).
2.6.3 Comparison of Acceptor Aryl Glycosides with Abg2F6

In order to determine which aryl glycoside acceptor worked best with Abg2F6 glycosynthase, five different acceptors were tested: DNPC (2.33), alkoxyamino β-D-glucoside (2.40), alkoxyamino β-D-cellobioside (2.42), para-nitrophenyl β-D-glucoside (PNPG, 2.43), and para-nitrophenyl β-D-cellobioside (PNPC, 2.44). GalF (2.08, 1.2 eq.) and the acceptor (5 mg) were dissolved in 100 mM sodium phosphate (pH 7.0) to an acceptor concentration of 10 mM. Abg2F6 (0.2 mg/mL) was added, and the reactions were incubated at room temperature.

The reactions were analyzed by TLC (7:2:1 EtOAc:MeOH:H2O) after 1, 2, and 24 hours. Since the acceptor glycosides are UV-active, product formation was confirmed by visualizing the TLC plates under UV-light prior to staining with ammonium molybdate. 2.33 and 2.44 were the best acceptors, with the reactions reaching completion in 1 hour (Rf = 0.36 and 0.33, respectively). The reaction utilizing the alkoxyamino cellobioside acceptor (2.42) was complete after 24 hours (Rf = 0.35). Disaccharide product was detected after 24 hours using either 2.40 or 2.43 glucoside acceptors (Rf = 0.53 and 0.51, respectively). However, there was a significant amount of glucoside acceptor remaining in both reactions.
Based on these results, the best acceptors for Abg2F6 glycosynthase are cellobiosides with para-nitrophenyl or dinitrophenyl functionalities. However, Abg2F6 glycosynthase has significantly less affinity for glucoside acceptors. Both alkoxyamino glycoside acceptors are utilized by Abg2F6, but will not be discussed further in this thesis.

2.7 Study of New Glycosynthase Enzymes

While Abg2F6 glycosynthase is useful in the synthesis of α-cellotriosyl fluoride (2.18) and α-cellotetraosyl fluoride (2.19), attempts to synthesize longer cello-oligosaccharide fluorides were unsuccessful, likely due to potent product inhibition. In order to synthesize cellohexaose analogs, efforts were focused on finding new glycosynthase enzymes that would be able to transfer α-cellobiosyl fluoride (2.12) or larger cello-oligosaccharide donors onto pre-existing cello-oligosaccharide acceptors. As noted earlier, hopefully these enzymes would also accommodate a charged substituent at the non-reducing end of the glycosyl donor. Cellulase enzymes were the most promising candidates for making endo-glycosynthases. Enzymes derived from three different cellulase families (B, C, and D)\textsuperscript{111} were examined in close collaboration with Emily Kwan, who had created these mutants previously and run a number of prior tests.

Family 5 endo-β-1,4-glucanases have a conserved glutamate residue as the catalytic nucleophile.\textsuperscript{112} Through sequence alignment with the catalytic glutamate residue identified in other family 5 endoglucanases, the relevant residue can be identified and mutated to a non-nucleophilic residue. This can lead to the formation of new glycosynthase enzymes that would have the potential to synthesize β-1,4 glycosidic linkages using longer α-glycosyl fluoride donors. CelC, a family 5 retaining endoglucanase C from \textit{Clostridium thermocellum},
nonspecifically hydrolyzes β-1,3 and β-1,4 glycosidic linkages from the non-reducing end of barley β-glucan and cello dextrans. The enzymatic hydrolysis of these carbohydrates produces a mixture of cellobiose, cellotriose, and cellotetraose. Wild-type CelC was converted to a glycosynthase by mutating the catalytic nucleophile to a glycine residue (E280G) (Emily Kwan, unpublished results).

CelB, a family 5 retaining endo-β-1,4-glucanase B from *Caldicellulosiruptor saccharolyticus*, is also able to degrade cellulose. CelB is associated with a family 3 CBM, which is known to bind to crystalline and amorphous cellulose. The catalytic nucleophile of CelB was mutated to a glycine residue (E1322G), and the enzyme was cloned with its CBM intact, producing CelB glycosynthase. Similarly, CenD, a family 5 retaining endo-β-1,4-glucanase D from *Cellulomonas fimi*, is able to degrade crystalline cellulose I. The catalytic nucleophile was mutated to glycine (E349G), yielding CenD glycosynthase.

### 2.7.1 Transfer of Cello-oligosaccharyl Fluorides Using Endo-glycosynthases

As a first test, the ability to synthesize longer cello-oligosaccharyl fluorides was tested using celllobiosyl fluoride as both donor and acceptor. Celllobiosyl fluoride (2.12) was dissolved in 50 mM sodium phosphate buffer to a final concentration of 40 mM, and a different glycosynthase enzyme was added to each reaction. The reactions were incubated at 30°C, and monitored by mass spectrometry after 24 hours. CelB glycosynthase was able to synthesize α-cellotetraosyl fluoride (2.19, 689.4 m/z, [M+Na]+). There was no visible peak in the mass spectrum for 2.19 in the reaction utilizing CenD glycosynthase, but hydrolyzed celllobiosyl fluoride was detected (2.09, 365.2 m/z, [M+Na]+). CelC glycosynthase was able to synthesize a
tetrasaccharide fluoride (689.2 m/z, [M+Na]\textsuperscript{+}), and a hexasaccharide fluoride (1015.3 m/z, [M+Na]\textsuperscript{+}), with no evidence of \textbf{2.12} remaining in the reaction.

It was interesting that a hexasaccharide fluoride was detected by mass spectrometry in the synthesis using CelC glycosynthase, since cellohexaose is insoluble in aqueous solution (section 1.1.2). As mentioned in section 2.7, wild-type CelC can hydrolyze both β-1,3 and β-1,4 glycosidic linkages. Therefore, α-lactosyl fluoride (\textbf{2.16}) was used as a test substrate to determine whether the glycosynthase enzymes could synthesize β-1,3 linkages. Lactosyl fluoride has a 4’ axial hydroxyl, which does not form β-1,4 glycosidic linkages with donor glycosyl fluorides. However, the 3’ hydroxyl is available to form β-1,3 glycosidic linkages.

All three glycosynthase enzymes were incubated with \textbf{2.16} using the same reaction conditions mentioned above. After 24 hours, the reactions were checked using mass spectrometry. The only peaks visible in the mass spectrum for CelB glycosynthase and CenD glycosynthase were \textbf{2.16} and hydrolyzed \textbf{2.16} (\textbf{2.13}, 365.3 m/z, [M+Na]\textsuperscript{+}). In the mass spectrum for CelC glycosynthase, both the tetrasaccharide fluoride (691.3 m/z, [M+Na]\textsuperscript{+}) and hexasaccharide fluoride (1013.5 m/z, [M+Na]\textsuperscript{+}) were detected. This implies that CelC glycosynthase synthesizes β-1,3 glycosidic linkages.

To confirm this, the oligosaccharide fluorides synthesized by the condensation of \textbf{2.12} using CelC glycosynthase were incubated with Cel6A, a β-1,4 endoglucanase from \textit{Cellulomonas fimi}.\textsuperscript{117} This cellulase did not hydrolyze the synthesized oligosaccharide fluorides, according to TLC or mass spectrometry. However, the oligosaccharide fluorides were hydrolyzed by wild-type CelC, confirming that CelC glycosynthase synthesizes β-1,3 glycosidic linkages (Emily Kwan, unpublished results). The β-1,3 glycosidic linkage was also confirmed by \textsuperscript{1}H NMR.\textsuperscript{118} The \textsuperscript{1}H NMR spectra of an acetylated tetrasaccharide synthesized by CelC
glycosynthase was compared to the spectra of an acetylated β-1,4 linked cello-oligosaccharide with the same carbohydrate composition. The coupling constants were very similar, but chemical shifts corresponding to the protons located next the 1,3 linkage differed by approximately 0.2 ppm (data not shown).

The ability of these new potential glycosynthase enzymes to transfer longer α-glycosyl fluoride donors to acceptor substrates was examined. DNPC (2.33, 10 mg, 20 µmol) and α-cellobiosyl fluoride (2.12, 7.2 mg, 1.05 eq.) were dissolved in 50 mM sodium phosphate to provide a final acceptor concentration of 20 mM. To each reaction, CelC glycosynthase, CelB glycosynthase, or CenD glycosynthase was added, and the reactions were incubated at 30°C (Scheme 2-18). The reactions were monitored by TLC (7:2:1 EtOAc:MeOH:H₂O) and mass spectrometry after 24 hours. Synthesis of 2.45 was confirmed for CelB glycosynthase (855.6 m/z, [M+Na]+). CelC glycosynthase synthesized the DNP tetrasaccharide (2.46, 855.6 m/z, [M+Na]+). Also, evidence of DNP hexasaccharide formation was shown by mass spectrometry (1180.1 m/z, [M+Na]+). CenD glycosynthase did not appear to synthesize any products.
Scheme 2-18: Ability of new glycosynthase enzymes to transfer 2.12 to 2.33. The synthesized glycosidic linkage is highlighted in red. a) 50 mM sodium phosphate (pH 8.0), CelB glycosynthase (1.25 mg/mL), 30°C; b) 50 mM sodium phosphate (pH 7.0), CelC glycosynthase (0.25 mg/mL), 30°C.

To test the ability of the glycosynthase enzymes to transfer trisaccharide fluoride donors, 2.17 (1.05 eq.) was incubated with DNPC (2.33) at 30°C (Scheme 2-19). After 24 hours, the pentasaccharide product (2.47) was detected using mass spectrometry in the reactions utilizing CelB glycosynthase (1017.7 m/z, [M+Na]^+) and CelC glycosynthase (2.48, 1017.8 m/z, [M+Na]^+). CenD glycosynthase did not synthesize any products.
Scheme 2-19: Ability of new glycosynthase enzymes to transfer 2.17 to 2.33. The synthesized glycosidic linkage is highlighted in red. a) 50 mM sodium phosphate (pH 8.0), CelB glycosynthase (1.25 mg/mL), 30°C; b) 50 mM sodium phosphate (pH 7.0), CelC glycosynthase (0.25 mg/mL), 30°C.

In summary, CelB glycosynthase is able to synthesize longer cello-oligosaccharide fluorides using both disaccharide and trisaccharide donor fluorides. CelC glycosynthase is able to synthesize longer oligosaccharides in which the donor and acceptor substrates are connected by β-1,3 glycosidic bonds. CenD glycosynthase does not appear to function as a true glycosynthase, and will not be discussed further in this thesis. A summary of these results is provided in Table 2-2.
<table>
<thead>
<tr>
<th>Glycosyl Fluoride Donor</th>
<th>CelB Glycosynthase</th>
<th>CelC Glycosynthase</th>
<th>CenD Glycosynthase</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.12</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>2.16</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>2.17</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
</tbody>
</table>

*Table 2-2: Glycosyl fluoride donor specificity for CelB, CelC, and CenD glycosynthase enzymes. Y = utilized by the enzyme. N = no product detected.*

2.7.2 Transfer of Sialic Acid Functionalized Glycosyl Fluorides

CelB glycosynthase and CelC glycosynthase were both able to utilize disaccharide and trisaccharide glycosyl fluoride donors. To demonstrate their ability to transfer charged species to acceptor substrates, sialic acid-functionalized glycosyl fluorides were explored as donors.

The ability of the glycosynthase enzymes to utilize sialylated tetrasaccharide fluoride (2.25) was examined. The tetrasaccharide was chosen instead of the sialylated trisaccharide (2.24) because the negative charge is further away from the anomeric position. This increases the chance that the charge will not interfere with substrate binding in the active site. Also, 2.25 is an extension of 2.17, which was successfully transferred by both enzymes. DNPC (2.33, 0.25 µmol) and 2.25 (2 eq.) were dissolved in 50 mM sodium phosphate to a final acceptor concentration of 5 mM. Either CelB glycosynthase or CelC glycosynthase was added, and the reactions were incubated at 30°C (Scheme 2-20). The reactions were monitored by mass spectrometry. After 18 hours, product formation was confirmed using both glycosynthase enzymes (1284.6 m/z (CelB, 2.49), 1284.8 m/z (CelC, 2.50), [M]+).
Scheme 2-20: Ability of CelB or CelC glycosynthase to transfer 2.25. Synthesized glycosidic linkage is highlighted in red. a) 50 mM sodium phosphate (pH 8.0), CelB glycosynthase (1.0 mg/mL), 30°C; b) 50 mM sodium phosphate (pH 7.0), CelC glycosynthase (0.25 mg/mL), 30°C.

This experiment demonstrated that sialylated tetrasaccharide fluoride (2.25) is utilized as a donor by both glycosynthase enzymes. To confirm this finding, a TLC assay was performed using bodipy cellobioside (2.51) as an acceptor. Bodipy cellobioside is highly fluorescent, enabling detection of products by visualizing the TLC plate under UV light (365 nm). The TLC assay was performed using the sialylated trisaccharide fluoride (2.24) and the sialylated tetrasaccharide fluoride (2.25).
Bodipy cellobioside (2.51, 0.40 µmol) and sialylated lactosyl fluoride (2.24, 5 eq.) were dissolved in 50 mM sodium phosphate to a final acceptor concentration of 2 mM. CelB glycosynthase or CelC glycosynthase was added, and the reaction was incubated at 30°C (Scheme 2-21). The reactions were monitored by TLC (5:4:1 CHCl₃:MeOH:0.2% CaCl₂). After 23 hours, a new fluorescent spot was formed from the reactions utilizing either glycosynthase enzyme (Rf = 0.37 (CelB, 2.52), Rf = 0.41 (CelC, 2.53)).

![Scheme 2-21: Ability of CelB or CelC glycosynthase to transfer 2.24. The synthesized glycosidic linkage is highlighted in red. a) 50 mM sodium phosphate (pH 8.0), CelB glycosynthase (1.0 mg/mL), 30°C; b) 50 mM sodium phosphate (pH 7.0), CelC glycosynthase (0.25 mg/mL), 30°C.](image)

Using the same conditions as scheme 2-21, the ability of both glycosynthase enzymes to transfer sialylated tetrasaccharide fluoride (2.25) to bodipy cellobioside (2.51) was examined. The reactions were monitored by TLC (4:5:1 CHCl₃:MeOH:0.2% CaCl₂). After 18 hours, a new
fluorescent spot in reactions using either glycosynthase enzyme (Rf = 0.63 (CelB), Rf = 0.67 (CelC)) was visible.

CelB glycosynthase and CelC glycosynthase are both able to transfer sialylated glycosyl fluorides to acceptor substrates. This will enable the transfer of the sialylated donors to cello-oligosaccharides.

2.7.3 Transfer of Phosphorylated Glycosyl Fluorides

The glycosynthase enzymes were successfully able to utilize sialylated glycosyl fluorides as donor substrates. The ability of the glycosynthase enzymes to utilize phosphorylated glycosyl fluoride donors would enable synthesis of longer cello-oligosaccharides and would simplify the process. Bodipy cellobioside (2.51) was used as an acceptor in tests to determine if the negative charge from the terminal phosphate hindered the glycosynthase reaction.

Control reactions were set up on the 20 µL scale with α-cellotriosyl fluoride (2.18) or α-cellotetraosyl fluoride (2.19) to determine if the glycosynthase enzymes transfer the cello-oligosaccharide fluorides. Bodipy cellobioside (2.51, 0.40 µmol) and glycosyl fluoride donor (5 eq.) were dissolved in 50 mM sodium phosphate to a final acceptor concentration of 2 mM. CelB glycosynthase or CelC glycosynthase was added, and the reactions were incubated at 30°C. The reactions were monitored by TLC (5:4:1 CHCl₃:MeOH:0.2% CaCl₂).

After 19 hours, new fluorescent spots were visible on TLC plates of both reactions using 2.18 (Rf = 0.33 (CelB) and Rf = 0.38 (CelC)). Also, a second fluorescent spot was visible in the reaction using CelC glycosynthase (Rf = 0.15), implying that a second addition of 2.18 occurred. In the reaction utilizing 2.19 donor, there was evidence of product formation using CelB
glycosynthase (Rf = 0.25). Two new fluorescent spots were visible in the reaction using CelC glycosynthase (Rf = 0.31, 0.08), suggesting one and two additions of 2.19 occurred.

Next, the ability of the enzymes to transfer phosphorylated cello-oligosaccharide fluorides was examined. Bodipy cellobioside (2.51, 0.40 μmol) and phosphorylated glycosyl fluoride donor (5 eq.) were dissolved in 50 mM sodium phosphate to a final acceptor concentration of 2 mM. CelB glycosynthase or CelC glycosynthase was added, and the reactions were incubated at 30°C. The reactions were monitored by TLC (4:5:2 CHCl₃:MeOH:0.2% CaCl₂).

After 19 hours, product formation was visible using phosphorylated cellobiosyl fluoride (2.27) donor (Rf = 0.17 for both enzymes). Product formation was also confirmed when phosphorylated cellotriosyl fluoride (2.28) was used as a donor (Rf = 0.16 (CelB), Rf = 0.21 (CelC)). However, when phosphorylated cellotetraosyl fluoride (2.29) was used as a donor, product was only synthesized using CelC glycosynthase (Rf = 0.12). CelB glycosynthase did not utilize 2.29 as a donor substrate.

The results of these experiments are summarized in Table 2-3. Phosphorylation does not interfere with either enzyme’s ability to synthesize new glycosidic linkages. CelB glycosynthase is able to transfer both 2.18 and 2.28, and can therefore be used to synthesize longer cello-oligosaccharides.
<table>
<thead>
<tr>
<th>Glycosyl Fluoride Donor</th>
<th>CelB Glycosynthase</th>
<th>CelC Glycosynthase</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.12</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>2.18</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>2.19</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>2.27</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>2.28</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>2.29</td>
<td>N</td>
<td>Y</td>
</tr>
</tbody>
</table>

Table 2-3: Glycosyl fluoride and phosphorylated glycosyl fluoride donor specificity for CelB glycosynthase and CelC glycosynthase. Y = utilized by the enzyme. N = no product detected.

2.8 Effectiveness of Charged Functionalities in Improving Cellulosic Oligosaccharide

Solubility

Cellohexaose is the shortest cello-oligosaccharide that is insoluble in aqueous solutions. To demonstrate the insolubility of cellohexaose (2.55), CelB glycosynthase was used to transfer α-cellotriosyl fluoride (2.18) to d-cellotriose (2.54). 2.18 (2.84 mg, 5.61 µmol) and 2.54 (1.5 eq.) were dissolved in 50 mM sodium phosphate (pH 8.0) to a final donor concentration of 20 mM. CelB glycosynthase (0.50 mg/mL) was added, and the reaction was incubated at 30°C overnight (Scheme 2-22). After 18 hours, a white precipitate was visible in the reaction, which is believed to contain a mixture of cellohexaose and cellohexaosyl fluoride (Figure 2-4A).

To confirm α-cellohexaosyl fluoride (2.56) is insoluble, 2.18 (4.87 mg, 9.62 µmol) was dissolved in 50 mM sodium phosphate (pH 8.0) to a final concentration of 40 mM. CelB
glycosynthase was added (0.50 mg/mL), and the reaction was incubated overnight at 30°C (Scheme 2-23). After 18 hours, a white precipitate was visible (Figure 2-4B), implying that the anomeric fluorine does not improve the solubility of cellohexaose.

Scheme 2-22: Synthesis of cellohexaose (2.55). i) 50 mM sodium phosphate (pH 8.0), CelB glycosynthase (0.50 mg/mL), 30°C, 18 hours.

Scheme 2-23: Synthesis of α-cellohexaosyl fluoride (2.56). i) 50 mM sodium phosphate (pH 8.0), CelB glycosynthase (0.50 mg/mL), 30°C, 18 hours.
In order to confirm that CelC glycosynthase synthesizes soluble hexasaccharides (first demonstrated in section 2.7.1), α-cellotriosyl fluoride (2.18) and α-cellotriose (2.54) were used as substrates. 2.18 (2.25 mg, 4.45 µmol) and 2.50 (1.5 eq.) were dissolved in 50 mM sodium phosphate (pH 7.0) to a final donor concentration of 20 mM. CelC glycosynthase (0.5 mg/mL) was added, and the reaction was incubated at 30°C overnight (Scheme 2-24). After 18 hours, the solution remained clear (Figure 2-5B), and a TLC of the reaction confirmed that both donor and acceptor substrates were consumed by CelC glycosynthase. Therefore, CelC glycosynthase
synthesizes oligosaccharides connected by β-1,3 glycosidic linkages that lack cellulosic character.

Scheme 2-24: Synthesis of β-1,3 linked hexasaccharide (2.57). i) 50 mM sodium phosphate (pH 7.0), CelC glycosynthase (0.50 mg/mL), 30°C, 18 hours.

To demonstrate that sialic acid and phosphorylation improve the solubility of cellohexaose, reactions were set up on the 200 µL scale using D-cellotriose (2.54) as an acceptor. As a control, 2.54 (3.0 mg, 6.0 µmol) was dissolved in 200 µL of 50 mM sodium phosphate (pH 8.0). In a separate vial, phosphorylated cellotriosyl fluoride (2.28, 4.00 mg, 5.36 µmol) and 2.54 (1.5 eq.) were dissolved in 50 mM sodium phosphate (pH 8.0), and CelB glycosynthase (0.5 mg/mL) was added. In another vial, sialylated tetrasaccharide (2.25, 6.76 mg, 8.49 µmol) and 2.54 (1.5 eq.) were dissolved in 50 mM sodium phosphate (pH 8.0) to a donor concentration of 20 mM, and CelB glycosynthase (0.50 mg/mL) was added. All reactions were incubated at 30°C overnight.
Figure 2-5: Comparison of crystalline α-cellohexaosyl fluoride (2.56, A) and β-1,3 linked hexasaccharide (2.57, B) synthesized using CelC glycosynthase.

D-Cellotriose (2.54) is soluble in aqueous solution (Figure 2-6A). Cellohexaose (2.55) is insoluble in aqueous solution, which is confirmed by the formation of a white precipitate (Figure 2-6B). The phosphorylated hexasaccharide is soluble in solution (Figure 2-6C). The sialic acid terminated heptasaccharide also remains soluble in solution (Figure 2-6D). TLC of the reactions was performed to confirm that both donor and acceptor substrates were consumed by CelB glycosynthase. Overall, the addition of a charged functionality enables the cellohexaose analogs to remain soluble in solution.
Figure 2-6: Ability of phosphorylation or sialic acid functionalization to improve solubility of cellohexaose. A) Control reaction (2.54); B) Cellohexaose (2.55) C) Reaction containing 2.28 and 2.54; D) Reaction containing 2.25 and 2.54.

2.9 Conclusions

Cellohexaose, the shortest unit of cellulose I, is insoluble in aqueous solution (Figure 2-4A). Therefore, there is potential that cellohexaose can be used to study type A CBM’s and other proteins that interact with crystalline cellulose. In order to perform these studies, cellohexaose must be carefully assembled in aqueous solution, which is only possible if it is soluble.
Using Abg2F6 glycosynthase, the longest cello-oligosaccharide that was synthesized in isolable quantities was α-cellotetraosyl fluoride (2.19). In order to synthesize longer cello-oligosaccharides, CelB glycosynthase (E1322G) was prepared. CelB glycosynthase is able to transfer α-cellotriosyl fluoride (2.18) and 2.19 to bodipy celllobioside (2.48), demonstrating its potential to synthesize cello-oligosaccharides. Using 2.18 as a donor and d-cellotriose (2.50) as an acceptor, insoluble cellohexaose was synthesized using CelB glycosynthase. A second glycosynthase, CelC glycosynthase, was discovered to synthesize β-1,3 glycosidic linkages. When incubated with 2.18 and 2.54, a soluble hexasaccharide was formed. The β-1,3 linkage disrupts hydrogen bonding between the hexasaccharide chains, removing the cellulosic character. Therefore, the β-1,3 linked carbohydrates cannot be used to study proteins that interact with cellulose, though it serves as a useful control.

In order to synthesize soluble cellohexaose analogs, donor glycosyl fluorides that were permanently charged at pH 7.0 were explored. The permanent charge will improve solubility in aqueous solution, and disrupt the hydrogen bonding that favors crystallization. The first strategy examined was the transfer of sialic acid onto a terminal galactose residue using CMP-Neu5Ac (2.23) and Cst-I enzyme. Using this strategy, sialylated tetrasaccharide fluoride (2.25) was synthesized, and successfully transferred to 2.51 using CelB glycosynthase. The second strategy involved phosphorylation of the terminal glucose residue at the 6 position by BglK enzyme. Using this strategy, both 6’’-O-phospho-cellotriosyl fluoride (2.28) and 6’’’-O-phospho-cellotetraosyl fluoride (2.29) were synthesized. 2.28 was successfully transferred to 2.51 using CelB glycosynthase.

CelB glycosynthase was able to transfer 2.25 and 2.28 separately to 2.54, synthesizing cellohexaose analogs that remained soluble in solution. Therefore, both sialic acid addition and
phosphorylation can keep cellohexaose soluble in solution. However, in order to synthesize an organized surface of cellulose I, the charged functionalities have to be removed. Removal of the sialic acid functionality would have to be performed in two steps. First, a sialidase would remove the terminal sialic acid. Next, the terminal galactose residue would have to be removed using a β-galactosidase. In comparison, the terminal phosphate is easily removed using alkaline phosphatase in 15 minutes. Therefore, phosphorylation is the preferred strategy to synthesize soluble cellohexaose analogs.

The next stage in this research involves using the charged glycosyl fluoride donors to synthesize cellulolipids. This will be further discussed in chapter 3.
Chapter 3: Synthesis of Cellulosic Glycolipids and Preparation of a Soluble Cellulose Surface

3.1 Synthesis of Glycosphingolipids Using EGCase Glycosynthase

The glycosynthase approach is not limited to the synthesis of simple oligosaccharides. The approach has also been extended to the synthesis of glycoconjugates, including glycosphingolipids and gangliosides. Endoglycoceramidase II (EGCase) is a family 5 retaining endo-glycosidase from Rhodococcus sp. strain M-777.\(^{119}\) EGCase directly hydrolyzes the glycosidic linkage between the ceramide and oligosaccharide (Figure 3-1). The EGCase active site is built up from a polar, flexible oligosaccharide binding site, and a narrow, hydrophobic tunnel with strict lipid acceptor specificity.\(^ {120}\)

**Figure 3-1:** Hydrolysis of glycosylceramide using EGCase. The ceramide product consists of a sphingosine base (highlighted in red) that is N-acylated with a long-chain fatty acid.
The first glycosynthase mutant of EGCase (E351S) was able to efficiently transfer a variety of glycosyl fluorides to $d$-erythro-sphingosine, with yields in excess of 94%. Through directed evolution, a second glycosynthase mutant was obtained (E351S, D314Y) that utilized a wider variety of sphingolipid acceptors. The double mutant will be referred to as EGCase glycosynthase for the remainder of this thesis.

Using EGCase glycosynthase, lactosyl sphingosine (3.02) was synthesized using the method of Rich et al. (Scheme 3-1). $d$-erythro-C18-sphingosine hydrochloride (3.01) was used as an acceptor due to its solubility in aqueous buffer. Lactosyl fluoride (1.5 eq., 2.16) and 3.01 were weighed out in a glass vial. The solids were dissolved in 25 mM sodium acetate (pH 5.3) to give a final sphingolipid concentration of 10 mM. 10% v/v 1,2-dimethoxyethane (DME) was added to improve the sphingolipid solubility. The solution was sonicated until it became clear, EGCase glycosynthase (0.5 mg/mL) was added, and the reaction was incubated in a water bath at 37°C. The reaction was monitored by TLC (5:4:1 CHCl$_3$:MeOH:10% AcOH) and mass spectrometry (624.4 m/z, [M+H]$^+$). Once complete, the reaction mixture was loaded onto a Waters Sep-Pak® tC18 cartridge, and the separation was performed using a gradient of acetonitrile:water (MeCN:H$_2$O) with 0.05% TFA as an ion-pairing agent. The product eluted at 50% acetonitrile, and fractions containing 3.02 were lyophilized, yielding pure 3.02 (80%).
Scheme 3-1: Synthesis of 3.02 using EGCase glycosynthase.

Using this strategy, celllobiosyl sphingosine (3.03), cellotriosyl sphingosine (3.04), and cellotetraosyl sphingosine (3.05) were synthesized, with the reaction yields listed in Table 3-1. Full characterization of the synthesized lipids is found in section 4.3.

<table>
<thead>
<tr>
<th>Glycosyl Fluoride Donor</th>
<th>Product</th>
<th>Mass [M]+</th>
<th>Reaction Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.12</td>
<td>3.03</td>
<td>624.4 m/z</td>
<td>72</td>
</tr>
<tr>
<td>2.18</td>
<td>3.04</td>
<td>786.4 m/z</td>
<td>80</td>
</tr>
<tr>
<td>2.19</td>
<td>3.05</td>
<td>948.5 m/z</td>
<td>82</td>
</tr>
</tbody>
</table>

Table 3-1: Cellulolipid synthesis results using EGCase glycosynthase.

3.1.1 EGCase Glycosynthase Reaction Using a Domestic Microwave

Glycosphingolipid synthesis utilizing EGCase glycosynthase requires incubation times of 4-6 days at 37°C, depending on the donor substrate used. Therefore, there is significant interest in speeding up the enzymatic reaction, without denaturing the enzyme or hydrolyzing the
glycosyl fluoride donor. Various methods have been used to increase the rate of enzymatic reactions, such as high-pressure cycling\textsuperscript{122} and the use of scientific microwave reactors\textsuperscript{123}. However, the technology required can be expensive, and may not be readily available at most institutions. As a cost-effective alternative, Zhou et al. demonstrated that a standard domestic microwave could be used to significantly decrease the incubation time required for complete de-\(N\)-glycosylation using Peptide-\(N\)-Glycosidase F from overnight to 20 minutes.\textsuperscript{124}

Experiments to determine the effectiveness of Zhou’s domestic microwave protocol for glycosynthase reactions were performed using \(d\)-\textit{erythro}-C18-sphingosine hydrochloride (4.0 mg, 3.01) as an acceptor and 1.5 eq. of \(\alpha\)-cellobiosyl fluoride (2.12) as a donor. Four separate reactions were set up according to the conditions outlined in section 3.1.1 in 1 mL glass vials. Reactions designated DMV(-) and WB(-) contained both substrates but no EGCase glycosynthase. Reactions designated DMV(+) and WB(+) contained both substrates and EGCase glycosynthase (0.50 mg/mL).

The domestic microwave used for these experiments has a maximum output power of 1400 W and a frequency of 2.45 GHz. The glass vials were placed in a Styrofoam holder in a plastic beaker filled with 1.5 L of room temperature tap water (22°C). The reactions DMV(+) and DMV(-) were microwaved for 20 minutes at power level 3, which was approximated to be 20% of the maximum power of the microwave. After 20 minutes, the reactions were immediately placed on ice to halt the reaction. The temperature of the water was measured to be 61.5°C. As a control, reactions WB(+) and WB(-) were placed in a 37°C water bath for 20 minutes, then immediately placed on ice. The reactions were analyzed by TLC (Figure 3-2).
Figure 3-2: Result of domestic microwave study using EGCase glycosynthase.

Eluent: 6:3.5:0.8 CHCl₃:MeOH:10% AcOH. The TLC plate on the left was stained with 10% ammonium molybdate in 2M H₂SO₄, followed by charring. The TLC plate on the right was stained with 0.3% ninhydrin in EtOH, followed by charring. TLC plate spot order: Starting substrates (no enzyme), DMV(-) after 20 minutes, DMV(+) after 20 minutes, WB(-) after 20 minutes, and WB(+) after 20 minutes.

Cellobiosyl sphingosine (3.03) is synthesized in both DMV(+) and WB(+), without significant hydrolysis of cellobiosyl fluoride after 20 minutes. While the reactions appeared to proceed at similar rates, it was not clear whether EGCase glycosynthase survived the temperature reached in the domestic microwave. As a follow-up experiment, DMV(+) and WB(+) were placed in a 37°C water bath over the weekend. After the weekend, the reactions were analyzed by TLC. WB(+) proceeded at a similar rate to the unmodified EGCase glycosynthase reaction, while no further product formation was visible in DMV(+). This confirms that EGCase
glycosynthase denatured due to the high temperature reached in the microwave. While there is the possibility that the rate of EGCase glycosynthase reactions can be enhanced with better temperature controls or more even heating, further experiments were not performed in this thesis.

3.2 Synthesis of Glycosylceramides

Lyso-glycosphingolipids contain a sphingosine base and a carbohydrate headgroup, but lack an N-acyl substituent. Through the formation of an amide linkage between sphingosine and a long-chain fatty acid, glycosylceramides can be synthesized. While lyso-glycosphingolipids have an amine that can be positively charged, the corresponding N-acylated glycosylceramide will be neutral. Therefore, N-acylation may enable tighter packing during celluolipid alignment.

3.2.1 Enzymatic Synthesis Using SCDase

One method that can be used to N-acylate the 2-amino group on sphingosine is enzymatic coupling with a ceramidase run in reverse. A ceramidase is an enzyme that hydrolyzes the amide bond between the sphingosine base and the fatty acid in a ceramide. Glycosylceramides are significantly more hydrophobic than lyso-glycosphingolipids. Therefore, glycosylceramides will have reduced solubility in aqueous solution, and will precipitate during an enzymatic synthesis, driving the reaction equilibrium towards glycosylceramide synthesis.

Sphingolipid ceramide N-deacetylase (SCDase) from Shewanella alga G8 was able to hydrolyze a variety of glycosylceramides, including lactosyl ceramide, galactosyl ceramide, and
glucosyl ceramide, with yields around 28%. From these substrate specificities, SCDase can theoretically be used to synthesize celluloceramides.

Cellobiosyl sphingosine (17.1 mg, 27.4 µmol, 3.03) and stearic acid (8.50 mg, 29.9 µmol, 1.1 eq., 3.06) were dissolved in 50 mM HEPES (pH 7.5) and DME (10% v/v) in a glass vial. SCDase (1% v/v, 0.341 mg/mL stock) was added, and the reaction was incubated at 37°C overnight (Scheme 3-2). A white precipitate was visible in the vial, and TLC (50:30:3 CHCl₃:MeOH:1% CaCl₂) confirmed formation of a new product (Rf = 0.71). The reaction was lyophilized, and the product was purified by flash column chromatography using the same TLC eluent. The isolated product was confirmed to be cellobiosyl ceramide (3.07) according to mass spectrometry (890.7 m/z, [M+H]+) and NMR (section 4.3). However, only 7 mg of 3.07 was isolated (29% yield). Therefore, due to the low reaction yield, other methods were explored for glycosylceramide synthesis.

Scheme 3-2: Synthesis of cellobiosyl ceramide (3.07) using SCDase.
3.2.2 Chemical Synthesis Using Pentafluorophenyl Esters

Activated pentafluorophenyl (PFP) esters are able to synthesize amide linkages, particularly between amino acids.\textsuperscript{128} Therefore, the preparation of PFP esters of long-chain fatty acids was explored.

Adamczyk \textit{et al.} synthesized PFP active esters of 5(6)-carboxyfluorescein derivatives using a mixture of pentafluorophenol and the coupling reagent 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) (3.08).\textsuperscript{129} Using this strategy, myristic acid (114 mg, 0.500 mmol, 3.09) was dissolved in DCM. Pentafluorophenol (1.3 eq., 3.08) and EDC hydrochloride (1.3 eq.) were added, and the reaction mixture was stirred at RT overnight (Scheme 3-3). The crude solid was purified by flash column chromatography using PE as an eluent, yielding the PFP ester as a white solid (3.11, 21%). Using the same reaction conditions, palmitic acid (3.10) and stearic acid (3.06) were activated with pentafluorophenol, yielding PFP-palmitate (3.12, 15%) and PFP-stearate (3.13, 17%), respectively. While this strategy worked, the reaction yields were significantly lower than expected.

As an alternative strategy, pentafluorophenyl trifluoroacetate (TFA-OPFP, 3.14) was explored as a reagent in the synthesis of PFP esters. Using the method of Green \textit{et al.},\textsuperscript{130} a test reaction was run using salicylic acid (3.15), with pyridine as a base catalyst. Salicylic acid (3.15) was dissolved in DCM. Pyridine (1.15 eq.) and TFA-OPFP (1.2 eq., 3.14) were added, and the reaction was stirred at RT overnight (Scheme 3-4). The crude solid was purified by flash column chromatography using PE as an eluent, yielding 3.16 as a while solid (73%). The synthesis of 3.16 was confirmed by NMR (section 4.3).
Using these reaction conditions, PFP activated esters of myristic acid (3.09), palmitic acid (3.10), and stearic acid (3.06) were synthesized, with the results listed in Table 3-2. PFP activated ester synthesis was confirmed by NMR (section 4.3). The reaction yields are much higher when TFA-OPFP is used instead of pentafluorophenol. Therefore, TFA-OPFP (3.14) was used for the synthesis of all PFP activated esters.

**Scheme 3-3:** Synthesis of PFP activated esters using pentafluorophenol (3.08) and myristic acid (3.09).

**Scheme 3-4:** Synthesis of PFP 2-hydroxybenzoate (3.16) using TFA-OPFP (3.14).
<table>
<thead>
<tr>
<th>Fatty Acid</th>
<th>Product</th>
<th>Reaction Yield Pentafluorophenol (%)</th>
<th>Reaction Yield TFA-OPFP (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C14:0 (3.09)</td>
<td>3.11</td>
<td>21</td>
<td>76</td>
</tr>
<tr>
<td>C16:0 (3.10)</td>
<td>3.12</td>
<td>15</td>
<td>79</td>
</tr>
<tr>
<td>C18:0 (3.06)</td>
<td>3.13</td>
<td>17</td>
<td>72</td>
</tr>
</tbody>
</table>

**Table 3-2:** Synthesis of PFP activated esters of long-chain fatty acids.

Using the strategy from Joullie *et al.*, the synthesis of cellobiosyl ceramide (3.07) was performed. Cellobiosyl sphingosine (10.2 mg, 3.03) and PFP-stearate (1.3 eq., 3.13) were dissolved in 1:1 DCM:DMF. Triethylamine (1.7 eq.) was added, and the reaction was stirred at RT overnight (Scheme 3-5). The reaction mixture was concentrated *in vacuo* and the resulting crude solid was purified by flash column chromatography with an eluent of 30:10:1 CHCl₃:MeOH:H₂O, yielding 3.07 (97%). The formation of 3.07 was confirmed by mass spectrometry (890.7 m/z, [M+H]+) and NMR (section 4.3).
3.2.3 Chemical Synthesis Using Succinimidyl Esters

PFP activated esters of long-chain fatty acids were synthesized with yields between 72-79%, and were successfully used to synthesize cellobiosyl ceramide (3.07). However, the PFP activated esters were not stable when stored in the freezer. Therefore, another strategy to synthesize activated fatty acid esters was required.

Fatty acid N-hydroxysuccinimide esters have been used to synthesize N-acylated amino acid esters. Therefore, N-hydroxysuccinimide (NHS) esters were synthesized using the strategy by Adamczyk et al. Myristic acid (3.09) was dissolved in DCM. NHS (4 eq., 3.17) and EDC hydrochloride (4 eq.) were added, and the reaction mixture was stirred at RT overnight (Scheme 3-6). The crude solid was purified by flash column chromatography using 4:1 hexanes:EtOAc as an eluent, yielding 3.18 as a while solid (64%). Synthesis of 3.18 was confirmed by mass spectrometry (348.2 m/z, [M+Na]^+). Using the same reaction conditions, palmitic acid (3.10) and stearic acid (3.06) were activated with NHS, yielding NHS-palmitate
(3.19, 78%) and NHS-stearate (3.20, 60%), respectively. Full characterization of all compounds is found in section 4.3.

Using the same conditions as 3.2.2, glycosylceramides were synthesized using the activated NHS esters. Cellobiosyl sphingosine (3.03) and 3.18 (1.4 eq.) were dissolved in 1:1 DCM:DMF. Triethylamine (1.7 eq.) was added, the reaction stirred at RT overnight (Scheme 3-7) then concentrated in vacuo. The crude solid was purified by flash column chromatography using an eluent of 30:10:1 CHCl₃:MeOH:H₂O, yielding 3.21 (100% yield). Synthesis of 3.21 was confirmed by mass spectrometry (834.6 m/z, [M+H]⁺). The reaction was repeated with NHS-palmitate (3.19) and NHS-stearate (3.20), yielding 3.22 and 3.07, respectively. All three glycosylceramides were synthesized quantitatively. Therefore, activated NHS esters were used for all further N-acylations in this thesis.

\[
\text{Scheme 3-6: Synthesis of 3.18 using NHS (3.17) and myristic acid (3.09).}
\]
3.3 Preparation of Elongated Glycosphingolipids

While EGCase glycosynthase successfully synthesizes cellulolipids using cello-oligosaccharide fluoride donors, donor solubility may place a limit on the length of sugar chain that can be added. An alternative strategy for the synthesis of longer cellulolipids would be to use CelB glycosynthase to couple soluble cello-oligosaccharyl fluorides to previously assembled cellulolipid acceptors.

3.3.1 Ability of CelB Glycosynthase to Utilize Glycosphingolipid Acceptors

Sphingolipids are not very soluble in sodium phosphate buffer. To improve sphingolipid solubility for glycosynthase reactions, 2-(N-morpholino)ethanesulfonic acid (MES, pH 7.0) was used as the reaction buffer.
The reactions were set up in glass vials, with 2 eq. of donor glycosyl fluoride. The glycosyl fluoride donors tested were lactosyl fluoride (2.16), and galactosyl-cellulobiosyl fluoride (2.17). These donor fluorides were chosen because the product would contain a terminal galactose residue, which cannot act as a subsequent acceptor for the glycosynthase enzyme. The glycosynthase acceptors tested were cellulobiosyl sphingosine (3.03) and cellotetraosyl sphingosine (3.05). MES buffer was used to dissolve the substrates to a final acceptor concentration of 5 mM, and CelB glycosynthase (1.0 mg/mL) was added (Schemes 3-8, 3-9, and 3-10). The reaction mixtures were incubated at 30°C for 5 days.

Scheme 3-8: Test reaction of 2.16 and 3.03 using CelB glycosynthase.
ESI-MS (electrospray ionization mass spectrometry), the method used for all previous mass spectrometry measurements in this thesis, was not sufficient for monitoring the glycosphingolipid reactions, due to the low signal-to-noise ratio obtained for longer products. Therefore, matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) was used. The MALDI matrix used was 2,5-dihydroxybenzoic acid (DHB, 20 mg/mL stock). Reaction mixture aliquots were diluted to an acceptor concentration of 100 µM using a solution
of 2:1 H₂O:MeCN containing 0.1% TFA. 5 μL of the diluted reaction was added to 5 μL of DHB, and 1 μL of that mixture was spotted on the MALDI plate. MALDI spectra were collected in positive ion linear mode.

The MALDI-MS results are listed in Table 3-3. Lactosyl fluoride (2.16) was transferred to 3.03, but not 3.05. Galactosyl-cellobiosyl fluoride (2.17) was transferred to both 3.03 and 3.05. A significant amount of both starting substrates remained in all reactions according to TLC, revealing that while glycosphingolipids do work as glycosynthase acceptors, they are not as effective as aryl glycosides.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2.16</td>
<td>3.03</td>
<td>647.6 m/z</td>
<td>3.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>971.3 m/z</td>
<td>3.23</td>
</tr>
<tr>
<td>2.17</td>
<td>3.03</td>
<td>647.5 m/z</td>
<td>3.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1135.1 m/z</td>
<td>3.24</td>
</tr>
<tr>
<td>2.16</td>
<td>3.05</td>
<td>971.9 m/z</td>
<td>3.05</td>
</tr>
<tr>
<td>2.17</td>
<td>3.05</td>
<td>972.1 m/z</td>
<td>3.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1458.4 m/z</td>
<td>3.25</td>
</tr>
</tbody>
</table>

Table 3-3: MALDI-MS results from reactions using CelB glycosynthase and glycosphingolipid acceptors.
3.3.2 Sialic Acid Functionalization Using Cst-I

The first strategy used to synthesize sialylated glycosphingolipids was the direct coupling of sialylated glycosyl fluorides to \(d\)-erythro-C18-sphingosine using EGCase glycosynthase. The sialylated tetrasaccharide fluoride (20.6 mg, 1.1 eq., 2.25) and \(d\)-erythro-C18-sphingosine hydrochloride (7.91 mg, 3.01) were dissolved in 25 mM NaOAc (pH 5.3) to a final acceptor concentration of 10 mM, and 10% v/v DME was added. EGCase glycosynthase (0.5 mg/mL) was added, and the reaction was incubated at 37°C. However, after 5 days, no product formation was visible according to TLC (4:5:1.5 CHCl\(_3\):MeOH:0.2% CaCl\(_2\)). This was surprising, since sialylated lactosyl fluoride (2.24) was previously utilized as a glycosyl fluoride donor using EGCase glycosynthase.\(^85\)

The next strategy used to synthesize sialylated glycosphingolipids was the direct transfer of sialic acid using Cst-I. Using the method of Rich \textit{et al.},\(^86\) lactosyl sphingosine (3.02) and CMP-Neu5Ac (1.1 eq., 2.23) were mixed in a glass vial and dissolved in 50 mM HEPES (pH 7.5) and 10 mM MnCl\(_2\). Cst-I (0.2 mg/mL) and alkaline phosphatase (5 \(\mu\)L, 10 KU) were added, and the reaction was incubated at RT (Scheme 3-11). The reaction was monitored by TLC (4:5:1 CHCl\(_3\):MeOH:0.2% CaCl\(_2\)), and product formation was confirmed by mass spectrometry (915.5 m/z, [M+H]\(^+\)). The reaction was complete after 2 hours, and purified on a Waters Sep-Pak\(^\text{®} \) tC18 cartridge using the same conditions as section 3.1, yielding sialylated lactosyl sphingosine (3.26, 80% yield). Full characterization of 3.26 is found in section 4.3.
Scheme 3-11: Synthesis of sialylated lactosyl sphingosine (3.26) using Cst-I.

Sialylated galactosyl-cellobiosyl sphingosine (3.28) was synthesized in two steps (Scheme 3-12). First, galactosyl-cellobiosyl sphingosine (3.27) was synthesized using Abg2F6 glycosynthase. Cellobiosyl sphingosine (3.03) and galactosyl fluoride (3 eq., 2.08) were dissolved in 100 mM MES (pH 7.0) to a final acceptor concentration of 10 mM. Abg2F6 glycosynthase (1.0 mg/mL) was added, and the reaction was incubated at room temperature. However, the reaction stopped after 5 days, even when extra 2.08 and Abg2F6 were added. Formation of 3.27 was confirmed by mass spectrometry (786.4 m/z, [M]+). The reaction mixture was loaded onto a Waters Sep-Pak® tC18 cartridge, but leftover 3.03 and 3.27 co-eluted at 50% acetonitrile. Further purification was not attempted because Cst-I only transfers CMP-Neu5Ac (2.23) to a terminal galactose residue. Therefore, the enzyme will not utilize 3.03 as an acceptor substrate, thus the separation should be simpler after treatment with Cst-I.

Substrates 3.27 and 2.23 (1.3 eq.) were mixed in a glass vial and dissolved in 50 mM HEPES (pH 7.5) and 10 mM MnCl₂. Cst-I (0.2 mg/mL) and alkaline phosphatase (5 µL, 10 KU)
were added, and the reaction mixture was incubated at RT. The reaction was complete after 2 hours, as confirmed by mass spectrometry (1077.5 m/z, [M+H]+), and the reaction mixture was loaded onto a Waters Sep-Pak® tC18 cartridge using the same conditions as in section 3.1, yielding 3.28 (29% yield). Full characterization of 3.28 is found in section 4.3.

Scheme 3-12: Synthesis of 3.28 using Abg2F6 glycosynthase and Cst-I.

3.3.3 Phosphorylation Using BglK

To synthesize phosphorylated glycosphingolipids, the first strategy used was the direct transfer of phosphorylated glycosyl fluorides to sphingosine using EGCase glycosynthase. Phosphorylated cellotetraosyl fluoride (26.9 mg, 1.1 eq., 2.29) and d-erythro-C18-sphingosine hydrochloride (4.92 mg, 3.01) were dissolved in 25 mM NaOAc (pH 5.3) to a final acceptor
concentration of 10 mM, and 10% v/v DME was added. EGCase glycosynthase (0.5 mg/mL) was added, and the reaction was incubated at 37°C. After 5 days, no product formation was visible according to TLC (4:5:2 CHCl₃:MeOH:10% AcOH). The reaction was repeated with phosphorylated cellotriosyl fluoride (2.28) and phosphorylated cellobiosyl fluoride (2.27) donors, but no product was visible. While EGCase has previously been shown to transfer sialylated lactosyl fluoride (2.24), it did not transfer the sialylated tetrasaccharide fluoride (section 3.3.2). Therefore, the active site of EGCase glycosynthase does not appear to accommodate glucosyl fluoride donors with a terminal phosphate in the 6 position.

The next reaction tried was the direct phosphorylation of glycosphingolipids using BglK. Cellobiosyl sphingosine (3.03) was dissolved in 25 mM HEPES (pH 7.5) containing 2 mM MgSO₄. Once dissolved, ATP (1.2 eq.) and BglK (0.25 mg/mL) were added, and the reaction was incubated at room temperature (Scheme 3-13). The reaction was monitored by TLC (4:5:1 CHCl₃:MeOH:0.2% CaCl₂), and was complete overnight. Product formation was confirmed by mass spectrometry (704.4 m/z, [M+2H]⁺). The reaction was purified on a Waters Sep-Pack® tC18 cartridge using a gradient of acetonitrile:water. The product eluted at 50% acetonitrile and the fractions were lyophilized, yielding 6’-O-phospho-cellobiosyl sphingosine (3.29, 54% yield). The reaction was repeated using the same conditions with cellotriosyl sphingosine (3.04) and cellotetraosyl sphingosine (3.05), yielding 6’’-O-phospho-cellotriosyl sphingosine (3.30, 75%) and 6’’’-O-cellotetraosyl sphingosine (3.31, 70%), respectively. Full characterization of all three compounds is found in section 4.3.
To confirm that the phosphate could be removed from the phosphorylated cellulolipids, dephosphorylation was tested using 6’’-O-phospho-cellotriosyl sphingosine (3.30). 3.30 was diluted with 200 mM Tris (pH 8.6) and 50 mM MgCl₂ to a final substrate concentration of 10 mM. Alkaline phosphatase (2 µL, 10 KU solution) was added, and the reaction mixture was incubated at room temperature (Scheme 3-14). After 15 minutes, the reaction was complete according to TLC (4:5:2 CHCl₃:MeOH:0.2% CaCl₂). The Rf increased from 0.13 to 0.74, indicating that removal of phosphate generated cellotriosyl sphingosine (3.04). This experiment confirms that the phosphorylated cellulolipids can be easily dephosphorylated in one step, which was also seen with the phosphorylated glycosyl fluorides (section 2.5.1).
3.3.4 Synthesis of Phosphorylated Cellohexaosyl Sphingosine

Using all of the previous results obtained in this section, the synthesis of phosphorylated cellohexaosyl sphingosine (3.32) was performed (Scheme 3-15). The phosphate functionality should keep the cellulolipid soluble, and removal of the phosphate would yield cellohexaosyl sphingosine. Cellohexaosyl sphingosine is the shortest cellulolipid that could be used to produce a surface of cellulose I.

Crude phosphorylated cellotriosyl fluoride (1.2 eq., 2.28) and cellotriosyl sphingosine (29.1 mg, 3.04) were dissolved in 100 mM MES (pH 7.0) to a final acceptor concentration of 20 mM. CelB glycosynthase (1.0 mg/mL) was added, and the reaction was incubated at 30°C. The reaction mixture remained clear, and product formation was monitored by TLC (4:5:2 CHCl₃:MeOH:10% AcOH, Rf = 0.30). Product formation was also confirmed by mass spectrometry (1352.6 m/z, [M+2H]^+). After 5 days, the reaction mixture was directly loaded onto a Waters Sep-Pak® tC18 cartridge. The separation was performed using a gradient of acetonitrile:water. 3.32 eluted at 50% acetonitrile along with unreacted 3.04, and the fractions
were lyophilized to yield the crude white solid. This solid could be redissolved in aqueous buffer. Therefore, 3.32 is a good candidate for the preparation of a cellulose surface.

Scheme 3-15: Synthesis of phosphorylated cellohexaosyl sphingosine (3.32).

3.4 Preparation of Fluorescent Glycosphingolipids

In order to monitor the incorporation of the cellulolipids along the aqueous:organic interface or in a nanodisc, the synthesis of a fluorescent cellulolipid was explored. To maintain the cellulosic character of the lipid, a fluorescent label would be attached either at the non-reducing end of the carbohydrate head group, or through \( N \)-acylation of sphingosine.\(^{133} \) Since there is a high chance that \( N \)-acylation could interfere with cellulolipid alignment, non-reducing end fluorescent labeling was explored.
3.4.1 Synthesis of 4-Keto Glycosyl Fluoride Donors

The first method explored was the incorporation of a ketone in the 4 position on the terminal glucose residue. A fluorescent aminooxy moiety can be conjugated to the ketone by oxime ligation, using aniline or \( \sigma \)-phenylenediamine as a catalyst, yielding a fluorescent molecule.\(^{134}\)

First, the ability of Abg2F6 glycosynthase to transfer 4-keto-glucosyl fluoride (3.33) was explored. The position of the 4-ketone is in between the configuration of glucose and galactose. Also, in water, the ketone is hydrated to form a gemdiol. The gemdiol has both axial and equatorial hydroxyls. Therefore, since Abg2F6 utilizes both GalF and GlcF donors, it should be able to transfer 4-keto-GlcF. A reaction was set up using DNPC (2.33) as the aryl glycoside acceptor, and 2 eq. of 4-keto-GlcF (3.33) as the glycosyl fluoride donor (Scheme 3-16). After 21 hours, product formation was confirmed by TLC (7:2:1 EtOAc:MeOH:H\(_2\)O, Rf = 0.53) and mass spectrometry (691.3 m/z, [M+Na]\(^+\)).

Scheme 3-16: Synthesis of 3.34 using Abg2F6 glycosynthase.
From this result, synthesis of 4''-keto-celiotriosyl fluoride (3.35) was performed (Scheme 3-17). 4-keto-GlcF (3.33, 7.8 eq.) and α-cellobiosyl fluoride (2.12) were dissolved in 100 mM sodium phosphate (pH 7.0), and Abg2F6 (1.0 mg/mL) was added. The reaction was incubated at RT overnight. Product formation was confirmed by TLC (7:2:1 EtOAc:MeOH:H2O, Rf = 0.17). The reaction was lyophilized after 18 hours to minimize formation of the gemdiol, even though a significant amount of 2.12 remained in the reaction. The reaction mixture was purified by flash column chromatography using an eluent of 7:2:1 EtOAc:MeOH:H2O, yielding a mixture of 3.35 and the hydrated ketone product. The hydrate forms in the presence of acid or base, which is difficult to prevent, even if the reaction is buffered at pH 7. Even though 3.35 could be prepared, it was not used in further reactions.

Scheme 3-17: Synthesis of 4''-keto-celiotriosyl fluoride (3.35) using Abg2F6 glycosynthase.
3.4.2 Synthesis and Testing of 4-Amino Glycosyl Fluoride Donors

The next strategy explored was the inclusion of a 4-amino group on the terminal glucose residue. NHS-activated carboxylic acids are able to synthesize amide linkages with amines, as demonstrated in section 3.2.3. If Abg2F6 glycosynthase transfers 4-amino-glucosyl fluoride (3.36) to a glycosyl fluoride acceptor, the resulting product could be utilized by CelB glycosynthase to synthesize a cellulolipid (similar to section 3.3.4). Further, the amine has a positive charge at pH 7.0, therefore it has the potential to keep the resulting cellulolipid soluble in aqueous solution.

4-Amino-GlcF (3.36, 2.7 eq.) and α-cellobiosyl fluoride (2.12) were dissolved in 100 mM sodium phosphate (pH 7.0) buffer, and Abg2F6 (1.0 mg/mL) was added (Scheme 3-18). The reaction was monitored by TLC (7:2:1 EtOAc:MeOH:H₂O). After 18 hours, 2.12 was consumed, and a new product spot was visible on the baseline. Mass spectrometry also confirmed product formation (506.2 m/z, [M+H]⁺, 3.37). The reaction mixture was lyophilized, and the crude solid was used without further purification since CelB glycosynthase does not utilize monosaccharide donors. A small amount of crude solid was purified using flash column chromatography and an eluent of 7:2:1 EtOAc:MeOH:H₂O to obtain NMR data for 3.37 (section 4.3).

To test whether CelB glycosynthase can utilize 3.37 as a glycosyl fluoride donor, a reaction was set up with cellotriosyl sphingosine (3.04) as an acceptor. Cellotriosyl sphingosine (3.04) and 4”-amino-cellotriosyl fluoride (theoretical maximum: 1.5 eq., 3.37) were dissolved in 100 mM MES (pH 7.0) to a final acceptor concentration of 10 mM. CelB glycosynthase (1.0 mg/mL) was added (Scheme 3-19), and the reaction was incubated at 30°C. A beige precipitate was visible after 18 hours. TLC (4:5:2 CHCl₃:MeOH:10% AcOH) confirmed that both 3.04 and 3.37 were being consumed, but no new product spot was visible. The insoluble product (3.38) was isolated using centrifugation, and lyophilized to remove residual water. The solid was dissolved in DMF, and product formation was confirmed by mass spectrometry (1271.6 m/z, [M+H]+).
Scheme 3-19: Synthesis of 3.38 using CelB glycosynthase.

While 3.38 was successfully synthesized, the NHS-activated fluorescent label can react with either amine present on the molecule, which could significantly interfere with cellulolipid alignment. One solution is to specifically N-acylate sphingosine with a long-chain fatty acid using SCDase (section 3.2.1). However, due to the low reaction yields obtained previously, another strategy to modify the amine on sphingosine was explored.

Azides have been demonstrated to be effective amine “protecting groups”, with many benefits compared to the traditional choices of amides or carbamates. The amine on cellotriosyl sphingosine (3.04) was converted to an azide using the method of Santana et al. (Scheme 3-20). Sodium azide (NaN₃, 1.5 eq.) was dissolved in 1:1 H₂O:toluene, and the mixture was cooled to 0°C with vigorous stirring. Trifluoromethanesulfonic anhydride (Tf₂O) was added dropwise, and the mixture was vigorously stirred at 10°C for 2 hours. The reaction was adjusted to pH 7 using saturated NaHCO₃, and the two phases were separated, yielding a
solution of trifluoromethanesulfonyl azide (TfN₃) in toluene. The reaction of TfN₃ and primary amines occurs under mild conditions with complete retention of stereochemistry. Cellotriosyl sphingosine (3.04), copper sulfate pentahydrate (1.3 eq.), and NaHCO₃ (2 eq.) were dissolved in H₂O. The TfN₃ solution was immediately added, and the reaction was stirred vigorously overnight at RT. The azide product was purified using a Waters Sep-Pak® tC18 cartridge. The product eluted at 75% acetonitrile, and the fractions were lyophilized, yielding 3.39 as a white solid (834.4 m/z, [M+Na]⁺, 64%).

**Scheme 3-20:** Conversion of 3.04 to 3.39 using TfN₃. i) NaN₃ (1.5 eq.), Tf₂O, H₂O, toluene, 10°C, 2 hours; ii) TfN₃, CuSO₄·5H₂O (1.3 eq.), NaHCO₃ (2 eq.), H₂O, toluene, RT, overnight, 64%.

Next, CelB glycosynthase (1.0 mg/mL) was used to transfer 4′′-amino-cellotriosyl fluoride (3.37, 3.0 eq.) to the azide-protected cellotriosyl sphingosine (3.39) in 100 mM MES (pH 7.0) (Scheme 3-21). The reaction was incubated at 30°C, and a precipitate was visible after 18 hours. TLC (4:5:2 CHCl₃:MeOH:10% AcOH) confirmed that both 3.37 and 3.39 were being
consumed. The insoluble product (3.40) was isolated using centrifugation and lyophilized. The solid was dissolved in DMF, and product formation was confirmed by mass spectrometry (1297.6 m/z, [M+H]+). Therefore, CelB glycosynthase was able to accommodate the 2-azido group on 3.39, with a reaction rate comparable to that of the unmodified substrate (3.04).

![Scheme 3-21: Synthesis of 3.40 using CelB glycosynthase.](image)

3.4.3 Synthesis of a Bodipy Labeled Cellulolipid

The cellulolipid containing a terminal amine (3.40) was thus successfully synthesized using a combination of Abg2F6 glycosynthase and CelB glycosynthase, with the use of an azide “protecting group” on sphingosine. The next step was to couple a fluorophore to the cellulolipid. NHS-activated bodipy (3.41) was chosen for this study, as bodipy is highly fluorescent at low concentrations (section 2.7.2).
The bodipy labeled cellulolipid was synthesized using the method of Yang et al.\textsuperscript{138} Crude 3.40 (theoretical maximum: 18.1 µmol) and NHS-activated bodipy (3.41, 1.15 eq.) were dissolved in DMF in a 25 mL round bottom flask. Triethylamine (1.4 eq.) was added, and the reaction mixture was stirred in the dark at RT overnight, then concentrated in vacuo. The synthesized red solid was dissolved in 25% acetonitrile and loaded onto a Waters Sep-Pak\textsuperscript{®} tC18 cartridge to remove unreacted 3.41 and any contaminating bodipy-labeled 3.37. Fractions containing material not attributed to 3.41 or labeled 3.37, which eluted at 75% acetonitrile, were combined and lyophilized to yield a dark red solid. TLC revealed the presence of a fluorescent spot on the baseline, which was not attributed to starting material. However, only 3.40 was detected by mass spectrometry (1297.6 m/z, [M+H]\textsuperscript{+}). Most likely, 3.42 was insoluble in the samples prepared for mass spectrometry. Therefore, the crude solid was used without further purification.
Scheme 3-22: Synthesis of fluorescent 3.42 using NHS activated bodipy (3.41). i)
Triethylamine (1.4 eq.), DMF, RT, overnight.

3.5 Preparation of Bodipy-Cellotetraoside

A method that can be used to detect a surface of cellulose I is the use of a fluorescent marker. A cellulosic glycoside with a fluorescent group in the anomeric position can potentially intercalate between the cellulolipid head groups. This intercalation would mimic cellulose II, and be stabilized by hydrogen bonding interactions. In order to try this idea, bodipy cellotetraoside (3.43) was synthesized. The cellotetraose head group, while not long enough to have cellulosic properties, was chosen to act as a probe. It was theorized that the cellotetraose head group would be long enough to intercalate into an organized cellulose I structure (Figure 3-3).

The synthesis of bodipy cellotetraoside (3.43) was accomplished by the coupling of celllobiosyl fluoride (2.12, 3.3 eq.) and bodipy celllobioside (2.51) using CelB glycosynthase.
(Scheme 3-23). The reaction mixture was incubated at 30°C for 7 days, lyophilized, and loaded onto a Waters Sep-Pak® tC18 cartridge to remove any cello-oligosaccharide fluorides. The bodipy compounds eluted at 25% acetonitrile, and were lyophilized. Synthesis of 3.43 was confirmed by mass spectrometry (1006.4 m/z, [M+Na]⁺), and crude 3.43 was used without further purification.

Scheme 3-23: Synthesis of bodipy cellotetraoside (3.43) using CelB glycosynthase.
**Figure 3-3**: Illustration of how bodipy cellotetraoside could act as a probe to detect a surface of cellulose I. Fluorescent bodipy is highlighted in red.
3.6 Preparation of Fluorescent CBM2a for Cellulose Surface Detection

Carbohydrate binding modules (CBM) present on carbohydrate active enzymes are able to bind to their specific substrate, even if the catalytic domain is removed. Therefore, in order to detect a crystalline cellulose surface, the CBM can be cloned separately from the catalytic enzyme.

A family 2a CBM from *Cellulomonas fimi* xylanase 10A (CBM2a) binds to crystalline regions of cellulose. Three solvent-exposed tryptophan residues mediate the binding of CBM2a to crystalline cellulose. When provided with bacterial microcrystalline cellulose (BMCC), binding of CBM2a was irreversible over several weeks, even when CBM2a present in the solution phase was removed. While binding is irreversible, CBM2a diffuses along the cellulose surface, as was demonstrated by fluorescence recovery techniques. For studies in this thesis, CBM2a was expressed in *E. coli* BL21 (DE3) cells with an N-terminal polyhistidine tag to enable its purification by nickel affinity chromatography.

3.6.1 Labeling of CBM2a with Fluorescein Isothiocyanate

In order to detect CBM2a binding to a crystalline surface of cellulose, a fluorescent tag was conjugated to CBM2a following literature procedures. CBM2a has two amino groups that can react with a fluorescent tag: the N terminus, and a surface-exposed lysine residue. Neither amine is located near the hydrophobic binding surface.

Using the method of Jervis et al., CBM2a was functionalized with fluorescein isothiocyanate (FITC, Figure 3-4). CBM2a was dialyzed against 100 mM sodium bicarbonate
(pH 9.0) to a final protein concentration around 1 mg/mL. Dialysis was also performed to minimize the amount of Tris buffer remaining after expression and purification. 50 µL of freshly prepared fluorescein-5(6)-isothiocyanate (3 mg/mL in DMSO) was added for every 1 mL of CBM2a. The reaction vessel was wrapped in tin foil, and gently mixed in the dark at RT for 45 minutes. Once complete, the reaction mixture was added to a PD-10 size exclusion column (GE Healthcare®) equilibrated with PBS to separate unbound FITC from the labeled protein. After purification, absorbance measurements were read at 280 nm and 496 nm to determine the protein concentration and number of FITC molecules bound per protein. The extinction coefficient of bound FITC at pH 9.0 is 70 000 M$^{-1}$cm$^{-1}$. From these absorbance measurements, 1.58 mol of FITC were bound for every mol of CBM2a. The formulas used for these calculations are described in section 4.4.7.1. The labeled protein (FITC-CBM2a) was stored in the dark at 4°C.

![Figure 3-4: Structure of fluorescein isothiocyanate (FITC).](image)
3.6.2 Confirmation of FITC-CBM2a Binding to Cellulose Columns

To confirm that CBM2a was labeled with FITC, and that expression of CBM2a was successful, its ability to bind to cellulose was examined. Cellulose columns were prepared using three different types of cellulose: microcrystalline cellulose, medium cellulose fibers, and long cellulose fibers. The columns were packed using 50 mg of cellulose and an eluent of 50 mM potassium phosphate (pH 7.0). Once the columns were packed, 100 µL of FITC-CBM2a (0.263 mg/mL) was loaded onto each cellulose column. As a control, 100 µL of FITC dye (0.3 mg/mL in 10% DMSO) was loaded onto a separate set of cellulose columns.

Once the columns were loaded with FITC or FITC-CBM2a, the columns were washed with 500 mL of deionized water. FITC completely eluted from all three cellulose columns after 5 mL. FITC-CBM2a remained bound to the column due to the high affinity of CBM2a for crystalline cellulose (Figure 3-5A). In order to elute FITC-CBM2a, the cellulose columns were washed with 25 mL of 6 M guanidinium chloride in 50 mM potassium phosphate (pH 7.0). Guanidinium chloride is a strong chaotropich agent that is used to denature proteins. The fluorescence of the cellulose columns was reduced considerably, implying that FITC-CBM2a was denatured and subsequently eluted from the cellulose columns (Figure 3-5B).

Based on these experiments, and as shown previously, CBM2a was successfully labeled with FITC, and is able to bind to crystalline cellulose. Therefore, this fluorescent derivative may bind to the cellulose I surfaces prepared in sections 3.7 and 3.8, and could be used to detect cellulose formation.
Figure 3-5: (A) Fluorescence of cellulose columns after washing with 500 mL deionized H$_2$O. The three columns on the left were originally loaded with FITC dye. The three columns on the right were loaded with FITC-CBM2a. (B) Fluorescence of cellulose columns originally loaded with FITC-CBM2a after washing with 25 mL of 6 M guanidinium chloride in 50 mM potassium phosphate (pH 7.0).

3.7 Attempted Preparation of a Cellulose Surface Along an Aqueous:Organic Interface

The first step required to prepare an aqueous:organic interface was the solvent choice. The organic solvent had to be completely immiscible with water. Also, the organic solvent had to have a higher density than water to ensure the organic layer was on the bottom, allowing easy addition of reagents to the aqueous phase. The only solvents that met these criteria were chlorinated solvents: chloroform, dichloromethane, and 1,2-dichloroethane.

One problem mentioned in section 3.2 was that the amine present on sphingosine, with its pKa of around 6.7, would be charged in aqueous solution. This positive charge might well
interfere with alignment of the cellulolipids along the interface. Therefore, experiments were performed using 50 mM HEPES (pH 7.5) as the aqueous layer, in an attempt to minimize the positive charge on the cellulolipid.

Experiments were performed in glass vials with a diameter of 0.7 cm. The interface was prepared by adding 200 µL of chlorinated organic solvent to the vial, followed by 200 µL of 50 mM HEPES (pH 7.5). Using an estimated area of 55 Å² for every lipid molecule, approximately 34 pmol of lipid was required to cover the aqueous:organic interface. The addition of larger amounts of cellulolipid could result in the formation of micelles and other lipid aggregates in the aqueous layer, complicating detection at the interface.

A solution containing 0.2 mg/mL phosphorylated cellohexaosyl sphingosine (3.32) in 50 mM HEPES (pH 7.5) was prepared. 34 pmol of 3.32 was added to each vial. The vials were incubated at 4°C for 30 minutes. Next, 2 µL of alkaline phosphatase (1 KU solution) was added, and the vials were incubated at 4°C for 2 hours. Alkaline phosphatase was added to cleave the phosphate, enabling alignment of the cellulolipids. As a control, 34 pmol of phosphorylated cellotetraosyl sphingosine (3.31) in 50 mM HEPES (pH 7.5) was added to a separate set of vials, followed by 2 µL of alkaline phosphatase (1 KU solution) using the same incubation times mentioned above. FITC-CBM2a is not expected to bind to cellotetraosyl sphingosine (3.05) because the cellotetraose head group is too short to have crystalline cellulosic properties.

After 2 hours, a visible cloudiness was present at the aqueous:organic interface in vials originally containing the soluble cellulolipid (3.32, Figure 3-6). There was no cloudiness visible at the interface in vials containing 3.05 (data not shown).
Figure 3-6: Aqueous:organic interface prepared with 50 mM HEPES (pH 7.5) and chlorinated organic solvent. 34 pmole of 3.32 was added, and the vials were incubated at 4°C with alkaline phosphatase for 2 hours. Left: Chloroform; Middle: Dichloromethane; Right: 1,2-Dichloroethane.

To determine whether the cloudiness was attributed to a crystalline surface of cellulose I, two different experiments were performed. Bodipy cellotetraoside (2 µL, 5 µg/µL, 3.43) was added to one set of vials to test if 3.43 could intercalate with the potential cellulose I surface. In another set of vials, FITC-CBM2a (2 µL, 26 µg/mL) was added. FITC-CBM2a (2 µL, 26 µg/mL) was also added to the control vials. The vials were incubated at 4°C overnight (Figure 3-7). There is no evidence of concentrated fluorescence at the interface in the control vials (Figure 3-7, left) or the vials containing 3.43 (Figure 3-7, middle). In the vials containing FITC-CBM2a,
there appears to be concentrated fluorescence in the vial with chloroform as the organic solvent (Figure 3-7, right).

**Figure 3-7:** Vials containing an aqueous:organic interface prepared with 50 mM HEPES (pH 7.5) and chlorinated organic solvent. The vials have been illuminated with 365 nm UV-light. Vials on the left contained cellotetraosyl sphingosine and FITC-CBM2a. Vials in the middle contained cellohexaosyl sphingosine and bodipy cellotetraoside (3.43). Vials on the right contained cellohexaosyl sphingosine and FITC-CBM2a. Additional FITC-CBM2a was added to the middle vial in the right set for clarity. **Vial order:** Left: Chloroform; Middle: Dichloromethane; Right: 1,2-Dichloroethane.

There is a high level of background fluorescence visible in the vials, which could be masking potential fluorescence at the interface. Adding smaller amounts of FITC-CBM2a to the vials could minimize the background fluorescence. However, a repeat of the experiment using smaller amounts of FITC-CBM2a was not performed due to time constraints.
While there is a visible cloudiness at the aqueous:organic interface that appears to interact with FITC-CBM2a, it cannot be concluded that the cloudiness is due to a prepared surface of cellulose I. Therefore, other experiments will have to be performed. One experiment that could be used to test whether the fluorescent interface contains a surface of cellulose I is to add a β-1,4 exoglucanase, such as wild-type Abg or Cex from *Cellulomonas fimi*, to the vial. If cellulose is present, the exoglucanase should degrade it, dispersing the fluorescence. Also, guanidinium chloride could be added to the vial containing the fluorescent interface. If the fluorescence disperses, it could indicate that the fluorescence was due to the interaction of FITC-CBM2a and the cellulosic interface.

### 3.8 Preparation of Nanodiscs Containing Cellulosic Glycolipids

Nanoparticulate phospholipid bilayer discs (nanodiscs) consist of a segment of phospholipid bilayer that is held together by a membrane scaffold protein (MSP). The MSP controls the size and dimensions of the lipid bilayer. In these experiments, MSP1 was used, which is based on the sequence of apolipoprotein A-1 and yields 10 nm diameter nanodiscs containing on average 150 lipid molecules. By changing the ratio of MSP1:phospholipid, the size of the nanodisc is changed. When more negatively charged phospholipids are incorporated, the size of the nanodisc and the negative charge on the nanodisc increases. The size difference is apparent when nanodiscs are analyzed using native polyacrylamide gel electrophoresis (PAGE). Larger nanodiscs will migrate faster on the native gel, due to the increase in negative charge. Size differences are also visible when the nanodiscs are run on a size exclusion column. Larger nanodiscs will elute from the column earlier than smaller nanodiscs. Alternatively, if a
fluorescent lipid is incorporated, measuring the ratio of fluorophore to MSP1 can estimate the amount of fluorescent lipids incorporated into the nanodisc. Nanodiscs were assembled with phospholipids 1,2-dioleoylphosphatidylglycerol (DOPG) and 1,2-dioleoylphosphatidylcholine (DOPC). All nanodisc experiments were performed with the assistance of fellow graduate student Spence MacDonald.

To demonstrate the formation of nanodiscs, different ratios of MSP1 and DOPG were solubilized in a 1% solution of n-dodecyl-β-D-maltoside (DDM) in buffer containing 50 mM Tris, 100 mM NaCl, and 5% glycerol. The detergent was slowly removed by incubating the reconstitutions overnight at 4°C in the presence of BioBeads® SM-2 adsorbents. The nanodiscs were analyzed by running a native PAGE gel, which is run in non-denaturing conditions (Figure 3-8). As the DOPG concentration increases in the nanodisc, the negative charge on the nanodisc increases, which is reflected by the increased migration distance in the gel (Figure 3-8, lanes 4-6). However, when the ratio of DOPG to MSP1 is increased to 80:1, MSP1 can no longer accommodate the lipids. This leads to the formation of lipoparticles, which do not have a uniform size. Therefore, ratios between 1:20 and 1:40 MSP1:DOPG were used in further experiments.
Figure 3-8: Native gel of reconstituted nanodiscs prepared with DOPG phospholipid. Nanodiscs are marked with arrows. The large band is MSP1. 20 µL of nanodisc reconstitutions were loaded onto the top of the gel. All native PAGE gels were run in the direction indicated by the black arrow on the left. **Lane 1:** protein ladder; **Lanes 2 and 10:** MSP1; **Lane 3:** 1:0 MSP1:DOPG; **Lanes 4 and 9:** 1:10 MSP1:DOPG; **Lane 5:** 1:20 MSP1:DOPG; **Lane 6:** 1:40 MSP1:DOPG; **Lane 7:** 1:80 MSP1:DOPG; **Lane 8:** 1:160 MSP1:DOPG.

Next, the fluorescent cellulotic sphingolipid (FLS, 3.42) was incorporated into nanodisc reconstitutions. Nanodisc reconstitutions containing 1:36:24:1 to 1:36:24:20 MSP1:DOPC:DOPG:FLS were prepared overnight. The nanodisc reconstitutions were analyzed using clear native PAGE gels, by either staining with coomassie blue or illuminating with UV light (Figure 3-9A). The only bands visible on the native PAGE gel correspond to nanodiscs. MSP1 is not visible on the gel since it is all incorporated into nanodiscs. Native PAGE confirms
that nanodiscs are being prepared, and that the fluorescent cellulolipids are being incorporated when a ratio of 1:36:24:5 or higher is used.

200 μL of nanodiscs from each reconstitution were injected on a 30 cm Superdex 200 gel filtration column. The size exclusion separation was performed using a buffer containing 50 mM HEPES (pH 7.0), 100 mM NaCl, and 2% glycerol, and a flow rate of 2 mL/min with UV detection (Figure 3-9B). The nanodiscs eluted around 30 mL, as expected. There was no peak visible around 40 mL, which confirms that all MSP1 is incorporated into nanodiscs. However, a second peak is visible around 60 mL of elution. This peak can be attributed to excess FLS that was not incorporated into the nanodisc, or possibly contaminating NHS-activated bodipy (3.41) that was not fully removed during purification.

As a final test, phosphorylated cellohexaosyl sphingosine (pG₆S, 3.32) was incorporated into the nanodisc reconstitutions in a ratio of 1:36:24:20 MSP1:DOPC:DOPG:pG₆S. As a control, empty nanodiscs (1:36:24 MSP1:DOPC:DOPG) were prepared. Alkaline phosphatase (25 U/μL) or buffer (50 mM HEPES (pH 7.0), 100 mM NaCl, 2% glycerol) was added to the nanodisc reconstitutions, and the reactions were incubated at 37°C for 30 minutes. The reactions were run on a native PAGE gel (Figure 3-10). Empty nanodiscs migrate further on the native gel due to their smaller size than nanodiscs containing 3.32. When alkaline phosphatase was added after the nanodisc reconstitution, the migration decreases slightly with 3.32-containing nanodiscs. No such change was visible with the empty nanodiscs. Removal of the negatively charged phosphate would reduce the overall negative charge of the nanodisc. This reduces the migration distance in the native PAGE gel, as seen in Figure 3-8. Therefore, phosphorylated cellohexaosyl sphingosine (3.32) can be incorporated into a nanodisc, and the phosphate can be easily removed without disrupting the nanodisc structure.
Figure 3-9: Nanodiscs reconstituted with fluorescent cellulosic sphingolipid (FLS, 3.42). Nanodiscs were reconstituted with a ratio of 1:36:24 MSP1:DOPC:DOPG. 20 µL of nanodisc reconstitutions were loaded onto the top of the gel. FLS was added to the reconstitutions in 1:1 – 1:20 ratios. (A) Native PAGE gel stained with coomassie blue (left) and UV-exposed native gel (right). (B) Size exclusion chromatography (SEC) of nanodisc reconstitutions shown in (A). FLS X refers to the ratio of 1:X Nanodisc:FLS.

3.9 Conclusions and Future Directions

In order to prepare a surface of cellulose I, the synthesis of cellulolipids was performed. First, EGCase glycosynthase was used to synthesize lyso-glycosphingolipids, using cello-oligosaccharyl fluoride donors prepared using Abg2F6 glycosynthase (section 2.3). Through N-acylation by enzymatic (SCDase) or chemical (NHS-activated esters) methods, the synthesized lyso-glycosphingolipids were easily converted to glycosylceramides.
CelB glycosynthase was able to extend the lyso-glycosphingolipids using either disaccharide or trisaccharide fluoride donors. Cst-I enzyme and BglK enzyme were both able to transfer charged functionalities directly to lyso-glycosphingolipids. From these results, CelB glycosynthase was able to transfer phosphorylated cellotriosyl fluoride (2.28) to cellotriosyl sphingosine (3.04), yielding phosphorylated cellohexaosyl sphingosine (3.32). The phosphorylated cellulolipid remained soluble in aqueous solution. Therefore, it was a great candidate for cellulose I surface preparation.

A fluorescent cellulolipid was prepared using a combination of glycosynthase enzymes. Abg2F6 glycosynthase transferred 4-amino-GlcF (3.36) to cellobiosyl fluoride (2.12), synthesizing 4’’-amino-cellotriosyl fluoride (3.37). Next, CelB glycosynthase transferred 3.37 to a cellotriosyl sphingosine analog (3.39), where the amine on sphingosine was converted to an azide. This yielded a cellulolipid containing a terminal amine (3.40) that could be functionalized with a fluorescent tag. NHS-activated bodipy (3.41) was able to react with the terminal amine on the cellulolipid, resulting in the formation of a fluorescent cellulolipid (3.42).

The synthesis of soluble cellulolipids was a success. Therefore, the preparation of a surface of cellulose I was attempted. To detect the surface, a fluorescent carbohydrate-binding module (FITC-CBM2a) was prepared. FITC-CBM2a irreversibly bound to columns of crystalline cellulose, and only eluted under protein-denaturing conditions.

An aqueous:organic interface was prepared using 34 pmol of 3.32. The phosphate was removed using alkaline phosphatase, and after 2 hours there was visible cloudiness at the interface. However, further experiments will need to be performed in order to prove or disprove the formation of a surface of cellulose I.
Next, the incorporation of cellulolipids into nanodiscs was explored. The fluorescent cellulolipid (3.42) was successfully incorporated into a nanodisc reconstitution, as confirmed by native PAGE and size exclusion chromatography. Phosphorylated cellohexaosyl sphingosine (3.32) was also incorporated into a nanodisc, and the phosphate was easily removed using alkaline phosphatase.

From these results, one experiment that would be interesting to try is the conjugation of cellulolipids to a solid surface. If cellulolipid analogs are synthesized with a functionality that can be anchored to a surface, the cellulolipids can self assemble to create a monolayer. This assembly would mimic a lipid bilayer, in a similar way to an aqueous:organic interface. However, monolayer formation could be monitored using methods other than fluorescence, such as surface plasmon resonance or atomic force microscopy. The anchoring of long-chain alkyl-thiols to gold is a well-published phenomenon, and results in the formation of self-assembled monolayers. Therefore, in order to synthesize a thiol-terminated cellulosic ceramide analog, the method of Ohlsson et al. was used.

16-Hydroxyhexadecanoic acid (3.44) was weighed out in a 50 mL round bottom flask. Aqueous HBr (48% solution, 41.2 eq.) and AcOH (81.3 eq.) were added, and the reaction was refluxed at 120°C for 3 days, yielding 16-bromohexadecanoic acid (99%). 16-bromohexadecanoic acid was dissolved in DMF, and potassium thioacetate (AcSK, 2.0 eq.) was added. The reaction was stirred at 50°C for 15 hours. The product was purified using flash column chromatography, yielding 3.45 (331.4 m/z, [M+H]^+, 70%). To synthesize the NHS activated acid, 3.45 was dissolved in DCM. NHS (4.0 eq., 3.19) and EDC hydrochloride (4.15 eq.) were added, and the reaction mixture was stirred at RT overnight. The crude solid was
purified by flash column chromatography, yielding 3.46 (428.3 m/z, [M+H]^+, 64%). NMR confirmed the formation of 3.45 and 3.46 (section 4.3).

\[
\begin{align*}
\text{HO} & \quad \text{O} \\
\text{O} & \quad \text{S} \\
\text{OH} & \quad \text{O} \\
\end{align*}
\]

**Scheme 3-24**: Synthesis of a thioacetate-terminated NHS-activated fatty acid (3.46). i) Aqueous HBr (41.2 eq.), AcOH, reflux, 3 days, 99%; ii) AcSK (2.0 eq.), DMF, 50°C, 15 hours, 70%; iii) NHS (4.0 eq.), EDC (4.15 eq.), DCM, RT, overnight, 64%.

Next, the thioacetate-terminated NHS-activated fatty acid was used to synthesize an analog of cellobiosyl ceramide. Cellobiosyl sphingosine (3.03) and 3.46 (1.3 eq.) were dissolved in DMF. Triethylamine (1.65 eq.) was added, and the reaction was stirred at RT overnight. Product formation was confirmed by TLC (30:10:1 CHCl_3:MeOH:H_2O, Rf = 0.55) and mass spectrometry (958.6 m/z, [M+Na]^+). The reaction mixture was concentrated *in vacuo*, and the crude solid was purified by flash column chromatography using an eluent of 30:10:1 CHCl_3:MeOH:H_2O, yielding 3.47 (84%). Purity was confirmed by NMR (section 4.3).
Scheme 3-25: Synthesis of a celllobiosyl ceramide analog with a terminal thioacetate (3.47). i) 3.46, triethylamine (1.65 eq.), DMF, RT, overnight, 84%.

A thioacetate-terminated analog of celllobiosyl ceramide was successfully synthesized using this strategy. Therefore, this strategy can be extended, enabling the synthesis of a variety of thioacetate-terminated cellulolipids. The acetate group can be removed using methanolic NaOMe\(^{149}\), yielding a free thiol that can be anchored to gold, which can potentially form a self-assembled monolayer of cellulose I.

In conclusion, the successful synthesis of soluble cellohexaose analogs and soluble cellulolipids was achieved. Phosphorylation was determined to be the best method tested to obtain soluble cellulosic materials. The phosphate functionality can be easily removed using alkaline phosphatase, which regenerates the cellulosic material. The incorporation of the cellulolipids into nanodiscs shows great promise and has laid a solid foundation for preparing a surface of cellulose I. With further optimization of nanodisc reconstitutions containing cellulolipids, as well as developing a protocol for monolayer formation utilizing alkyl-gold interactions, preparation of a cellulose I surface will become a reality.
Chapter 4: Materials and Methods

4.1 Generous Gifts

Dr. Hong-Ming Chen provided substrates PNPG (2.43), PNPC (2.44), Bodipy cellulobioside (2.51), 4-keto-GlcF (3.33), and 4-amino-GlcF (3.36). Dr. Jamie Rich provided Cst-I, along with substrates CMP-Neu5Ac (2.23) and d-erythro-C18-sphingosine hydrochloride (3.01). Emily Kwan provided cells containing expression vectors for Abg2F6, BglK, EGCase glycosynthase, CelB glycosynthase, CelC glycosynthase, CenDcd glycosynthase, and CBM2a. Other generous gifts include SCDase and NHS-activated Bodipy (3.41) from Dr. Guangyu Yang, and BglK for initial studies from James Frank.

4.2 General Synthesis and Methods

All reagents were obtained from commercial suppliers (Fisher Scientific®, Sigma-Aldrich®, Alfa Aesar®, and Carbosynth®) and used without further purification, unless otherwise stated. Solvents used were reagent grade or higher. DMF was dried over 4 Å molecular sieves for a minimum of two days before use. Methanol was distilled in the presence of magnesium. Pyridine was dried over potassium hydroxide. A Millipore Direct-Q™ 5 Ultrapure Water System providing 18 MΩ water was used for all aqueous solutions. TLC was performed on Merck pre-coated 0.2 mm aluminum-backed sheets of Silica Gel 60F254. TLC plates were visualized under UV light (254 or 365 nm), or stained with 10% ammonium molybdate in 2 M H2SO4, 0.3% ninhydrin in EtOH, or KMnO4, followed by charring. Flash column chromatography was
performed under elevated pressure using 230-400 mesh Silicycle silica gel. Reverse-phase column chromatography was performed using Waters Sep-Pak® tC18 cartridges (0.50 – 5.0 g).

$^1$H NMR spectra were acquired on a Bruker AV-400 Inverse (400 MHz), AV-400 Direct (400 MHz), or AV-600 (600 MHz) spectrometer. Chemical shifts were presented on the δ scale in parts per million (ppm) and referenced internally to CDCl$_3$ (7.26 ppm), CD$_3$OD (3.31 ppm), or D$_2$O (4.79 ppm). All $^{13}$C NMR spectra were proton-decoupled, and acquired on a Bruker AV-400 Inverse (100 MHz) or AV-400 Direct (100 MHz) spectrometer. Samples were referenced internally to CDCl$_3$ at 77.16 ppm or CD$_3$OD at 49.00 ppm. Complex $^1$H and $^{13}$C NMR spectra were assigned using COSY, TOCSY, $^{13}$C APT, $^1$H-$^{13}$C HSQC, and $^1$H-$^{13}$C HMBC experiments as required. $^{19}$F NMR spectra were acquired on a Bruker AV-300 (282 MHz) spectrometer. $^{31}$P NMR spectra ($^1$H decoupled) were acquired on a Bruker AV-400 Direct (161 MHz) spectrometer. Mass spectra were obtained using a Waters® ZQ™ Mass Detector equipped with an ESCI ion source.

4.2.1 General Procedure for Acetylation of Free Sugars

A suspension of NaOAc (1.1 eq) and Ac$_2$O (3 eq./OH) was heated with stirring until a temperature of 120°C was reached. The free sugar was added, and the reaction was refluxed for 1-2 hours. Once the reaction was complete, the reaction mixture was poured over ice and left to stir overnight at room temperature. The crude precipitate was isolated by vacuum filtration, dissolved in EtOAc, and washed successively with water, saturated NaHCO$_3$, water, and brine. The EtOAc layer was dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated in vacuo. The resulting white powder was crystallized using EtOH to give the per-O-acetylated glycoside.
4.2.2 General Procedure for Synthesis of α-Glycosyl Fluorides

45 eq. of HF (HF-pyridine, 70% HF, 30% pyridine) was added to the per-O-acetylated glycoside and stirred in a plastic bottle overnight at 4°C. Once the reaction was complete, the reaction mixture was poured over a stirred mixture of ice and saturated NaHCO$_3$ to quench unreacted HF-pyridine. The reaction products were extracted with EtOAc and washed successively with saturated NaHCO$_3$, deionized water, and brine. The organic layer was dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated *in vacuo*. Products were purified by flash column chromatography using an appropriate eluent mixture of EtOAc and petroleum ether.

4.2.3 General Procedure for Deacetylation

The per-O-acetylated glycoside was dissolved or suspended in dry MeOH (1-2 mL per 10 mg sugar). The reaction mixture was cooled down to 0°C under argon. NH$_3$(g) was slowly bubbled into the solution for 10 minutes. The reaction mixture was left at 4°C without stirring for 1-3 days. Once complete, the reaction mixture was concentrated *in vacuo*. If needed, the product was crystallized using MeOH and Et$_2$O, or purified by flash column chromatography using an eluent of 5:1 DCM:MeOH (monosaccharides) or 7:2:1 EtOAc:MeOH:H$_2$O.
4.2.4 General Procedure for Synthesis of Oligosaccharides Using Abg2F6 Glycosynthase

In a glass vial, a mixture of donor $\alpha$-glycosyl fluoride ($\alpha$-GlcF (1.5-5 eq.) or $\alpha$-GalF (1.5 eq.)) and glycosyl acceptor was dissolved in 100 mM sodium phosphate (pH 7) to give an acceptor concentration of 20 mM. Abg2F6 glycosynthase (0.1 mg/mL to 1 mg/mL) was added, and the reaction mixture was incubated at room temperature overnight. The reaction was monitored by TLC (7:2:1 EtOAc:MeOH:H$_2$O or 4:5:1 CHCl$_3$:MeOH:H$_2$O) and mass spectrometry. Upon completion, the reaction mixture was lyophilized.

4.2.5 General Procedure for Acetylation of $\alpha$-Glycosyl Fluoride Products of Glycosynthase Reactions

The lyophilized $\alpha$-glycosyl fluoride mixture was suspended in MeOH, sonicated for 15 minutes, and then centrifuged at 4000 rpm for 10 minutes to remove insoluble buffer salts. The MeOH mixture was concentrated in vacuo. The residue was suspended in a mixture of Ac$_2$O (3 eq./OH) and an equal volume of dry pyridine, and stirred overnight at room temperature. If the reaction was not complete overnight, DMAP (0.01 mol%) was added to push the reaction to completion. Once complete, MeOH was added and the mixture was stirred for 15 minutes to quench any remaining Ac$_2$O. The reaction mixture was then diluted with EtOAc (minimum 10x reaction volume) and this EtOAc solution washed successively with water, 10% HCl, water, saturated NaHCO$_3$, and brine. The organic layer was dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated in vacuo. Products were purified by flash column chromatography using an appropriate eluent mixture of EtOAc and petroleum ether.
4.2.6 General Procedure for Sialoside Synthesis Using Cst-I Enzyme

Acceptor α-glycosyl fluoride or glycosyl sphingosine was weighed out in a glass vial. CMP-Neu5Ac (1.2-2 eq.) was added, and the substrates were dissolved in 50 mM HEPES (pH 7.5) and 10 mM MnCl₂ to give an acceptor concentration of 10 mM. Once the solution was clear, Cst-I (0.20 mg/mL) and 0.1 µL alkaline phosphatase (10 KU solution from bovine intestinal mucosa; Sigma-Aldrich®) was added, and the reaction was left at room temperature. The reaction was monitored by TLC, and lyophilized once complete. Sialylated glycosyl fluoride products were used in reactions without further purification. Sialylated glycosyl sphingosine products were re-dissolved in water before purification on a Waters Sep-Pak® tC18 cartridge using a gradient of acetonitrile:water.

4.2.7 General Procedure for Phosphorylation Using BglK Enzyme

Acceptor α-glycosyl fluoride or glycosyl sphingosine was weighed out in a glass vial and dissolved in 25 mM HEPES (pH 7.5) containing 2 mM MgSO₄ (5 mL HEPES for every 1 mol acceptor). Once dissolved, ATP (1.2 eq, adjusted to pH 7.5 with 3 M NH₄OH) was added to the reaction mixture, followed by BglK (0.25 mg/mL to 1 mg/mL) and the reaction mixture was incubated at room temperature. The reaction was monitored by TLC, and lyophilized once complete. BglK enzyme was removed by filtration through an Amicon® Ultra-15 3K centrifugal filter (EMD Millipore®). Phosphorylated glycosyl fluoride products were used in reactions without further purification. Phosphorylated glycosyl sphingosine products were re-dissolved in
water before purification on a Waters Sep-Pak® tC18 cartridge using a gradient of acetonitrile:water.

4.2.8 General Procedure for Ganglioside Synthesis Using EGCase Glycosynthase

$d$-Erythro-C18-sphingosine hydrochloride and $\alpha$-glycosyl fluoride donor (1.5 eq.) were dissolved in 25 mM sodium acetate (pH 5.3) in a glass vial to give an acceptor concentration of 10 mM. DME (10% v/v) was added, and the reaction mixture was sonicated until the solution was clear. Once clear, EGCase glycosynthase (0.5 mg/mL) was added and the reaction mixture was incubated in a 37°C water bath for 4-6 days. The reaction was monitored by TLC by diluting 5 $\mu$L of the reaction mixture 10-fold in CHCl$_3$:MeOH (1:1) and using a TLC eluent of 5:4:1 CHCl$_3$:MeOH:10% AcOH. Once complete, the reaction was lyophilized and re-dissolved in water before purification on a Waters Sep-Pak® tC18 cartridge using a gradient of acetonitrile/water with 0.05% TFA as an ion-pairing agent. The target glycolipid was lyophilized before use in further reactions.

4.3 Synthesis and Characterization

1,2,3,4,6-Penta-$O$-acetyl-$\beta$-$d$-glucopyranose (2.02)

A suspension of NaOAc (2.74 g, 33.4 mol, 1.0 eq.) and Ac$_2$O (37 mL, 0.391 mol, 2.5 eq./OH) was heated to 120°C, and d-glucose (6.01 g, 33.4 mmol, 2.01) was added. The mixture was
refluxed for 1 hour, and the reaction was worked up according to the **General Procedure for Acetylation of Free Sugars**. Crystallization of the crude material with ethanol yielded white crystals of 2.02 (8.49 g, 21.8 mmol, 65%). **Rf** (1:1 EtOAc:Hexanes): 0.60 **H NMR** (400 MHz, CDCl$_3$): δ ppm 2.00 (s, 3H, Ac), 2.02 (s, 6H, 2 Ac), 2.08 (s, 3H, Ac), 2.11 (s, 3H, Ac), 3.84 (ddd, 1H, J$_{H6'-H5}$ 1H, J$_{H6-H5}$ 12.4 Hz, H6’), 4.28 (dd, 1H, J$_{H5-H6}$ 12.4 Hz, H6), 5.11 (dd, 1H, J$_{H3-H4}$ 9.6 Hz, H4), 5.14 (dd, 1H, J$_{H2-H1}$ 8.4 Hz, H2), 5.24 (dd, 1H, J$_{H4-H3}$ 9.6 Hz, H3), 5.71 (d, 1H, J$_{H2-H1}$ 8.4 Hz, H1) **C NMR** (100 MHz, CDCl$_3$): δ ppm 20.69 (3 CH$_3$), 20.83 (CH$_3$), 20.94 (CH$_3$), 61.57 (C6), 67.87 (C4), 70.35 (C3), 72.85 (C5), 72.92 (C2), 91.83 (C1), 169.08 (C=O), 169.37 (C=O), 169.51 (C=O), 170.22 (C=O), 170.73 (C=O) **ESI MS** calculated for C$_{16}$H$_{22}$O$_{11}$: 390.3, found: 413.2 [M+Na]$^+$.  

2,3,4,6-Tetra-O-acetyl-α-D-glucopyranosyl fluoride (2.03)

![Structure of 2.03](image)

Compound 2.03 was prepared according to the **General Procedure for Synthesis of α-Glycosyl Fluorides**. Per-O-acetylated D-glucose (6.01 g, 15.4 mmol, 2.02) was reacted with HF-pyridine (17.4 mL, 0.692 mol, 45 eq.) overnight, and worked up yielding a crude yellow oil. The oil was crystallized using MeOH and Et$_2$O, yielding 2.03 as white crystals (5.14 g, 14.7 mmol, 95%). Data agreed with that reported previously. **Rf** (1:1 EtOAc:PE): 0.57 **H NMR** (400 MHz, CDCl$_3$): δ ppm 2.02 (s, 3H, Ac), 2.03 (s, 3H, Ac), 2.09 (s, 3H, Ac), 2.10 (s, 3H, Ac), 4.14 (m, 2H, J$_{H5-H6}$ 4.0 J$_{H6-H6'}$ 12.4 Hz, H5, H6), 4.40 (dd, 1H, J$_{H6-H6'}$ 12.4 Hz, H6’), 4.95 (ddd, 1H, J$_{H3-H2}$ 10.0 J$_{H1-H2}$ 2.8 J$_{H2-F}$ 24.4 Hz, H2), 5.14 (dd, 1H, J$_{H4-H3}$ 10.0 J$_{H2-H3}$ 10.0 Hz, H3), 5.48 (dd, 1H,
\( J_{H3-H4} \) 10.0 Hz, H4), 5.74 (dd, 1H, \( J_{H2-H1} \) 2.8 \( J_{H1-F} \) 52.8 Hz, H1) \(^{19}\text{F NMR} \) (282 MHz, CDCl\(_3\)) \( \delta \) ppm -150.45 (dd, \( J \) 24.4 \( J \) 52.8 Hz) \(^{13}\text{C NMR} \) (100 MHz, CDCl\(_3\)): \( \delta \) ppm 20.65 (2 CH\(_3\)), 20.72 (CH\(_3\)), 20.79 (CH\(_3\)), 61.33 (C6), 67.48 (C3), 69.53 (C4), 69.93 (C5), 70.46 (C2), 102.74 (C1), 169.54 (C=O), 170.05 (C=O), 170.09 (C=O), 170.65 (C=O) \( \text{ESI MS} \) calculated for C\(_{14}\)H\(_{19}\)FO\(_9\): 350.1, found: 373.2 [M+Na]\(^+\).

\( \alpha-\text{d-Glucopyranosyl fluoride (2.04)} \)

\[
\begin{align*}
\text{HO} & \quad \text{OH} \\
\text{HO} & \quad \text{O} \\
\text{F} & \quad \text{OAc}
\end{align*}
\]

Compound 2.04 was prepared according to the \textbf{General Procedure for Deacetylation}. Per-\textit{O}-acetylated glucosyl fluoride (6.29 g, 18.0 mmol, 2.03) was deacetylated, and the crude material was purified by flash column chromatography using an eluent of 5:1 DCM:MeOH, yielding 2.04 (2.89 g, 15.9 mmol, 88%). \(^1\text{H NMR} \) (400 MHz, MeOD): \( \delta \) ppm 3.40 (dd, 1H, \( J_{H3-H4} \) 9.6 Hz, H4), 3.43 (ddd, 1H, \( J_{H3-H2} \) 9.6 \( J_{H1-H2} \) 2.8 \( J_{H2-F} \) 26.0 Hz, H2), 3.60-3.75 (m, 4H, H3, H5, H6, H6\('\)), 5.55 (dd, 1H, \( J_{H2-H1} \) 2.8 \( J_{H1-F} \) 53.6 Hz, H1) \(^{19}\text{F NMR} \) (282 MHz, MeOD) \( \delta \) ppm -151.81 (dd, \( J \) 26.0 \( J \) 53.6 Hz) \(^{13}\text{C NMR} \) (100 MHz, MeOD): \( \delta \) ppm 62.11 (C6), 70.63 (C3), 73.09 (C4), 74.44 (C5), 76.09 (C2), 107.87 (C1) \( \text{HRMS} \) calculated for C\(_6\)H\(_{11}\)FO\(_5\): 182.0591, found: 205.0488 [M+Na]\(^+\).

\( \text{1,2,3,4,6-Penta-\textit{O}-acetyl-\textit{\beta}-d-galactopyranose (2.06)} \)

\[
\begin{align*}
\text{AcO} & \quad \text{OAc} \\
\text{AcO} & \quad \text{O} \\
\text{AcO} & \quad \text{OAc}
\end{align*}
\]
A suspension of NaOAc (10.1 g, 0.122 mol, 1.1 eq.) and Ac₂O (140 mL, 1.48 mol, 2.7 eq./OH) was heated to 120°C, and D-galactose (20.0 g, 0.111 mol, 2.05 eq.) was added. The mixture was refluxed for 2 hours, and the reaction was worked up according to the General Procedure for Acetylation of Free Sugars. Crystallization of the crude material with ethanol yielded white crystals of 2.06 (18.4 g, 47.1 mmol, 42%). Rf(1:1 EtOAc:Hexanes): 0.50 \( ^1\text{H NMR} \) (400 MHz, CDCl₃): \( \delta \) ppm 1.98 (s, 3H, Ac), 2.03 (s, 6H, Ac), 2.11 (s, 3H, Ac), 2.15 (s, 3H, Ac), 4.04 (dd, 1H, J\text{H6-H5} 6.8 Hz, H5), 4.13 (m, 2H, J\text{H5-H6} 6.8 J\text{H6-H6} 11.6 Hz, H6, H6'), 5.07 (dd, 1H, J\text{H4-H3} 3.2 J\text{H2-H3} 10.4 Hz, H3), 5.32 (dd, 1H, J\text{H3-H2} 10.4 Hz, H2), 5.41 (d, 1H, J\text{H3-H4} 3.2 Hz, H4), 5.67 (d, 1H, J\text{H2-H1} 8.4 Hz, H1) \( ^{13}\text{C NMR} \) (100 MHz, CDCl₃): \( \delta \) ppm 20.65 (CH₃), 20.76 (CH₃), 20.93 (CH₃), 61.15 (C6), 66.92 (C4), 67.96 (C2), 70.96 (C3), 71.83 (C5), 92.28 (C1), 169.09 (C=O), 169.49 (C=O), 170.07 (C=O), 170.23 (C=O), 170.46 (C=O) \( \text{ESI MS} \) calculated for \( \text{C}_{16}\text{H}_{22}\text{O}_{11} \): 390.3, found: 413.2 [M+Na]+.

2,3,4,6-Tetra-O-acetyl-\( \alpha \)-D-galactopyranosyl fluoride (2.07)

\[
\begin{align*}
\text{AcO} & \quad \text{OAc} \\
\text{AcO} & \quad \text{OAc} \\
\text{AcO} & \quad \text{F}
\end{align*}
\]

Compound 2.07 was prepared according to the General Procedure for Synthesis of \( \alpha \)-Glycosyl Fluorides. Per-O-acetylated D-galactose (5.11 g, 13.1 mmol, 2.06) was reacted with HF-pyridine (14.8 mL, 0.589 mol, 45 eq.) overnight, and worked up yielding a crude yellow oil. The oil was purified using flash column chromatography and an eluent of 1:1 EtOAc:PE, yielding 2.07 as white crystals (2.71 g, 7.74 mmol, 59%). Data agreed with that reported previously.\( ^{89} \) Rf (1:1 EtOAc:PE): 0.57 \( ^1\text{H NMR} \) (400 MHz, CDCl₃): \( \delta \) ppm 2.00 (s, 3H, Ac), 2.04 (s, 3H, Ac), 2.11
(s, 3H, Ac), 2.14 (s, 3H, Ac), 4.13 (m, 2H, \(J_{H5-H6} = 6.8 \) \(J_{H6-H6'} = 10.8\) Hz, H6, H6’), 4.40 (dd, 1H, \(J_{H6-H5} = 6.8\) Hz, H5), 5.18 (dd, 1H, \(J_{H3-H2} = 11.2 \) \(J_{H1-H2} = 2.8 \) \(J_{H2-F} = 24.0\) Hz, H2), 5.35 (dd, 1H, \(J_{H4-H3} = 3.2 \) \(J_{H2-H3} = 11.2\) Hz, H3), 5.52 (m, 1H, H4), 5.80 (dd, 1H, \(J_{H2-H1} = 2.4 \) \(J_{H1-F} = 53.2\) Hz, H1) 19F NMR (282 MHz, CDCl3) \(\delta\) ppm -151.13 (dd, \(J_{24.0} J_{53.2}\) Hz) 13C NMR (100 MHz, CDCl3): \(\delta\) ppm 20.68 (CH3), 20.71 (CH3), 20.77 (CH3), 21.16 (CH3), 61.38 (C6), 67.10 (C3), 67.46 (C4), 67.65 (C2), 69.01 (C5), 105.57 (C1), 169.99 (C=O), 170.10 (C=O), 170.35 (C=O), 170.46 (C=O) ESI MS calculated for C_{14}H_{19}FO_9: 350.1, found: 373.2 [M+Na]^+.

\(\alpha\)-d-Galactopyranosyl fluoride (2.08)

Compound 2.08 was prepared according to the **General Procedure for Deacetylation**. Per-O-acetylated galactosyl fluoride (2.28 g, 6.51 mmol, 2.07) was deacetylated, and the crude material was crystallized using MeOH and Et2O, yielding 2.08 as white crystals (0.785 g, 4.31 mmol, 66%). Data agreed with that reported previously.\(^{89}\) Rf (7:2:1 EtOAc:MeOH:H2O): 0.60 \(^1\)H NMR (400 MHz, MeOD): \(\delta\) ppm 3.67-3.80 (m, 3H, H5, H6, H6’), 3.83, (ddd, 1H, \(J_{H1-H2} = 2.8 \) \(J_{H3-H2} = 10.4 \) \(J_{H2-F} = 26.8\) Hz, H2), 3.91-3.99 (m, 2H, H3, H4), 5.57 (dd, 1H, \(J_{H2-H1} = 2.8 \) \(J_{H1-F} = 55.6\) Hz, H1) 19F NMR (282 MHz, MeOD) \(\delta\) ppm -154.07 (dd, \(J = 26.8 \) \(J = 55.6\) Hz) 13C NMR (100 MHz, MeOD): \(\delta\) ppm 62.42 (C6), 69.66, 70.50, 70.90, 74.80, 110.36 (C1) HRMS calculated for C_{6}H_{11}FO_{5}: 182.0591, found: 205.0488 [M+Na]^+. 

---

132
(2,3,4,6-Tetra-O-acetyl-β-D-glucopyranosyl)-(1→4)-O-2,3,6-tri-O-acetyl-α-D-glucopyranosyl fluoride (2.11)

Compound 2.11 was prepared according to the General Procedure for Synthesis of α-Glycosyl Fluorides. Cellobiose octaacetate (5.10 g, 7.52 mmol, 2.10) was reacted with HF-pyridine (8.48 mL, 0.338 mol, 45 eq.) overnight, and worked up yielding a crude white solid. The white solid was crystallized using EtOAc and hexanes, yielding 2.11 as white crystals (4.16 g, 6.52 mmol, 87%). $^1$H and $^{13}$C NMR data agreed with that reported previously. $^{150}$ Rf (7:2:1

EtOAc:MeOH:H$_2$O: 0.32 $^1$H NMR (400 MHz, CDCl$_3$): δ ppm 1.98 (s, 3H, Ac), 2.01 (s, 3H, Ac), 2.03 (s, 3H, Ac), 2.04 (s, 3H, Ac), 2.09 (s, 3H, Ac), 2.09 (s, 3H, Ac), 2.14 (s, 3H, Ac), 3.67 (ddd, 1H, $^J_{H6''''}$, 9.6 Hz, H5’), 3.83 (dd, 1H, $^J_{H3}$, 9.6 Hz, H4), 4.05 (dd, 1H, $^J_{H5'}$, 12.8 Hz, H6’), 4.12 (m, 1H, H5), 4.15 (dd, 1H, $^J_{H5'}$, 12.0 Hz, H6’), 4.36 (dd, 1H, $^J_{H5'}$, 12.4 Hz, H6’), 4.54 (d, 1H, $^J_{H2'}$, 8.0 Hz, H1’), 4.55 (dd, 1H, $^J_{H5'}$, 1.6 Hz, H6’), 4.87 (dd, 1H, $^J_{H2'}$, 2.8 Hz, H1’), 5.07 (dd, 1H, $^J_{H3'}$, 9.2 Hz, H2’), 5.15 (dd, 1H, $^J_{H3'}$, 9.2 Hz, H2’), 5.47 (dd, 1H, $^J_{H2'}$, 10.0 Hz, H3’), 5.67 (dd, 1H, $^J_{H2'}$, 2.8 Hz, H1’), 19F NMR (282 MHz, CDCl$_3$) δ ppm -149.84 (dd, $^J_{24.4}$ 52.8 Hz, 13C NMR (100 MHz, CDCl$_3$): δ ppm 20.68 (4 CH$_3$), 20.70 (CH$_3$), 20.78 (CH$_3$), 20.96 (CH$_3$), 61.28 (C6), 61.76 (C6’), 67.97 (C4’), 68.91 (C3), 70.82 (C5), 70.86 (C2), 71.75 (C2’), 72.18 (C5’), 73.06 (C3’), 75.53 (C4), 100.69 (C1’), 102.65 (C1), 169.07 (C=O), 169.42
(C=O), 169.60 (C=O), 170.31 (2 C=O), 170.38 (C=O), 170.62 (C=O) ESI MS calculated for C_{26}H_{35}FO_{17}: 638.2, found: 661.4 [M+Na]^+.

(β-D-Glucopyranosyl)-(1→4)-α-D-glucopyranosyl fluoride (2.12)

![Chemical structure of 2.12](image)

Compound 2.12 was prepared according to the General Procedure for Deacetylation. Per-O-acetylated cellobiosyl fluoride (3.95 g, 6.19 mmol, 2.11) was deacetylated, and the crude material was crystallized using MeOH and Et₂O, yielding 2.12 as white crystals (1.91 g, 5.55 mmol, 90%). Rf (4:5:1.5 CHCl₃:MeOH:0.2% CaCl₂): 0.75

1H NMR (400 MHz, MeOD): δ ppm 3.24 (dd, 1H, J_{H3':H2'} 8.8 J_{H1':H2'} 8.0 Hz, H2'), 3.32 (m, 2H, H4', H5'), 3.38 (dd, 1H, J_{H2':H3'} 8.8 J_{H4':H3'} 9.2 Hz, H3'), 3.50 (ddd, 1H, J_{H11'-H12} 2.4 J_{H3'-H2} 9.6 J_{H2-F} 25.6 Hz, H2), 3.67 (m, 2H, H4, H6'''), 3.77 (dd, 1H, J_{H2-H3} 9.2 Hz J_{H4-H3} 9.2 Hz, H3), 3.82-3.97 (m, 3H, H6, H6', H6''), 3.94 (dd, 1H, J_{H6-H5} 12.4 J_{H6'-H5} 3.6 Hz, H5), 4.42 (d, 1H, J_{H2'-H1'} 8.0 Hz, H1'), 5.54 (dd, 1H, J_{H2-H1} 2.4 J_{H1-F} 53.6 Hz, H1) F NMR (282 MHz, MeOD) δ ppm -152.15 (dd, J 25.6 J 53.6 Hz) C NMR (100 MHz, MeOD): δ ppm 61.14 (C6), 62.44 (C6'), 71.36, 72.93, 73.13, 74.51, 74.91, 77.85, 78.14, 79.46, 104.58 (C1'), 109.83 (C1) HRMS calculated for C_{12}H_{21}FO_{10}: 344.1119, found: 367.1016 [M+Na]^+.
(2,3,4,6-Tetra-\(O\)-acetyl-\(\beta\)-\(D\)-galactopyranosyl)-(1→4)-\(O\)-2,3,6-tri-\(O\)-acetyl-\(\beta\)-\(D\)-glucopyranosyl acetate (2.14)

A suspension of NaOAc (6.00 g, 73.2 mmol) and Ac\(_2\)O (60 mL, 0.64 mol) was heated to 120°C, and d-lactose (6.00 g, 17.5 mol, 2.13) was added. The mixture was refluxed for 1 hour, and the reaction was worked up according to the General Procedure for Acetylation of Free Sugars. Crystallization of the crude material from ethanol yielded a white crystalline solid (2.14, 7.27 g, 10.7 mmol, 61%). \(^1\)H and \(^{13}\)C NMR data agreed with that reported previously. \(^{151}\) Rf (1:1 EtOAc:PE): 0.33 \(^1\)H NMR (400 MHz, CDCl\(_3\)): δ ppm 1.93 (s, 3H, Ac), 2.00 (s, 3H, Ac), 2.01 (s, 3H, Ac), 2.02 (s, 3H, Ac), 2.04 (s, 3H, Ac), 2.07 (s, 3H, Ac), 2.09 (s, 3H, Ac), 2.12 (s, 3H, Ac), 3.74 (ddd, 1H, \(J_{H4-5} 6.8\) \(J_{H6-H5} 10.0\) Hz, H5), 3.84 (m, 2H, \(J_{H3-H4} 9.2\) \(J_{H6'-H5'} 6.8\) Hz, H4, H5’), 4.01-4.14 (m, 3H, H6, H6’, H6”’), 4.44 (m, 2H, \(J_{H6''-H6'''} 12.8\) Hz, H6”, H1’), 4.92 (dd, 1H, \(J_{H4'-H5'} 10.4\) Hz, H3’), 5.01 (dd, 1H, \(J_{H3-H2} 9.6\) Hz, H2), 5.08 (dd, 1H, \(J_{H3'-H2'} 10.4\) Hz, H2’), 5.22 (dd, 1H, \(J_{H4-H3} 9.2\) Hz, H3), 5.32 (d, 1H, \(J_{H5'-H4'} 2.8\) Hz, H4”), 5.65 (d, 1H, \(J_{H2-H1} 8.4\) Hz, H1) \(^{13}\)C NMR (100 MHz, CDCl\(_3\)): δ ppm 20.58 (CH\(_3\)), 20.67 (CH\(_3\)), 20.68 (CH\(_3\)), 20.71 (2 CH\(_3\)), 20.83 (CH\(_3\)), 20.88 (CH\(_3\)), 20.90 (CH\(_3\)), 60.94 (C6’), 61.83 (C6), 66.70 (C4’), 69.09 (C2’), 70.60 (C2), 70.83 (C5’), 71.04 (C3’), 72.71 (C3), 73.58 (C5), 75.75 (C4), 91.62 (C1), 101.03 (C1’), 168.93 (C=O), 169.09 (C=O), 169.63 (C=O), 169.70 (C=O), 170.13 (C=O), 170.22 (C=O), 170.39 (C=O), 170.43 (C=O) ESI MS calculated for C\(_{28}\)H\(_{38}\)O\(_{19}\): 678.6, found: 701.5 [M+Na]\(^+\).
(2,3,4,6-Tetra-\(O\)-acetyl-\(\beta\)-d-galactopyranosyl)-(1→4)-\(O\)-2,3,6-tri-\(O\)-acetyl-\(\alpha\)-d-glucopyranosyl fluoride (**2.15**) 

Compound **2.15** was prepared according to the **General Procedure for Synthesis of \(\alpha\)**-Glycosyl Fluorides. Per-\(O\)-acetylated \(d\)-lactose (4.24 g, 6.25 mmol, **2.14**) was reacted with HF-pyridine (7.05 mL, 0.281 mol, 45 eq.) overnight, and worked up yielding a crude yellow oil. The oil was purified using flash column chromatography and an eluent of 1:1 EtOAc:PE, yielding **2.15** as white crystals (2.93 g, 4.59 mmol, 73%). **Rf** (1:1 EtOAc:PE): 0.26. **\(^1\)H NMR** (400 MHz, CDCl\(_3\)): \(\delta\) ppm 1.92 (s, 3H, Ac), 2.00 (s, 3H, Ac), 2.00 (s, 3H, Ac), 2.02 (s, 3H, Ac), 2.06 (s, 3H, Ac), 2.10 (s, 3H, Ac), 2.12 (s, 3H, Ac), 3.84 (m, 2H, \(J_{H3-H4}\) 9.6 Hz, H4, H5), 4.08 (m, 4H, H5', H6, H6', H6''), 4.49 (m, 2H, H4', H6'''), 4.84 (ddd, 1H, \(J_{H1-H2}\) 2.8 \(J_{H3-H2}\) 10.4 \(J_{H2-F}\) 24.4 Hz, H2), 4.93 (dd, 1H, \(J_{H1':-H2'}\) 3.2 \(J_{H3':-H2'}\) 10.4 Hz, H2'), 5.07 (dd, 1H, \(J_{H2':-H3'}\) 10.4 Hz, H3'), 5.31 (d, 1H, \(J_{H2':-H1'}\) 2.8 Hz, H1'), 5.45 (dd, 1H, \(J_{H2':-H3}\) 10.0 \(J_{H4-H3}\) 9.6 Hz, H3), 5.63 (dd, 1H, \(J_{H2':-H1}\) 2.8 \(J_{H1:F}\) 53.0 Hz, H1'). **\(^{19}\)F NMR** (282 MHz, CDCl\(_3\)) \(\delta\) ppm -149.43 (dd, \(J 24.4 \) J 53.0 Hz) **\(^{13}\)C NMR** (100 MHz, CDCl\(_3\)): \(\delta\) ppm 20.53 (CH\(_3\)), 20.60 (CH\(_3\)), 20.65 (CH\(_3\)), 20.67 (CH\(_3\)), 20.83 (CH\(_3\)), 20.85 (CH\(_3\)), 21.06 (CH\(_3\)), 60.96 (C6), 61.34 (C6’), 69.02 (C3’), 69.11 (C3), 70.37 (C4’), 70.72 (C5’), 70.76 (C5), 70.84 (C2), 71.04 (C2’), 75.16 (C4), 100.84 (C1’), 102.60 (C1), 168.98 (C=O), 169.46 (C=O), 170.09 (C=O), 170.18 (C=O), 170.28 (C=O), 170.38 (C=O), 171.15 (C=O) **ESI MS** calculated for \(C_{26}H_{35}FO_{17}\): 638.2, found: 661.4 [M+Na]\(^+\).
(β-D-Galactopyranosyl)-(1→4)-α-D-glucopyranosyl fluoride (2.16)

Compound 2.16 was prepared according to the General Procedure for Deacetylation. Per-O-acetylated lactosyl fluoride (1.43 g, 2.24 mmol, 2.15) was deacetylated, and the crude material was purified by flash column chromatography using an eluent of 7:2:1 EtOAc:MeOH:H₂O, yielding 2.16 as white crystals (0.383 g, 1.11 mmol, 50%). Rf (7:2:1 EtOAc:MeOH:H₂O): 0.26

**1H NMR** (400 MHz, MeOD): δ ppm 3.44-3.57 (m, 3H, J₁₁−H₂ 2.8 J₁₁−H₂ 7.2 J₁₃−H₂ 7.6 J₁₂−F 26.5 Hz, H₂, H₂’, H₅), 3.58-3.68 (m, 2H, H₃, H₄’), 3.68-3.78 (m, 3H, H₃’, H₅’, H₆’), 3.79-3.87 (m, 3H, J₆₅−H₆ 3.6 J₆₆−H₆ 12.0 Hz, H₄, H₆’, H₆’), 3.90 (dd, 1H, J₆₅−H₁ 2.8 J₆₆−H₁ 54.0 Hz, H₁) **19F NMR** (282 MHz, MeOD) δ ppm -150.64 (dd, J 26.5 J 54.0 Hz) **13C NMR** (100 MHz, MeOD): δ ppm 61.25 (C₆), 62.25 (C₆’), 70.00, 72.35, 72.75, 72.92, 74.53, 74.75, 77.09, 80.13 (C₄), 105.22 (C₁’), 107.79 (C₁) **HRMS** calculated for C₁₂H₂₁FO₁₀: 344.1119, found: 367.1016 [M+Na]⁺.

(β-D-Galactopyranosyl)-(1→4)-β-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl fluoride (2.17, 2.20)

Compound 2.17 was prepared according to the General Procedure for Synthesis of Oligosaccharides using Abg2F6 Glycosynthase. α-Galactosyl fluoride (2.08, 570 mg, 3.13
mmol, 1.6 eq.) and α-cellobiosyl fluoride (2.12, 693 mg, 2.01 mmol) were dissolved in 100 mM sodium phosphate (pH 7.0), and Abg2F6 (0.10 mg/mL) was added. The reaction was monitored by TLC (7:2:1 EtOAc:MeOH:H2O), and was complete overnight. The reaction mixture was lyophilized, and crude 2.17 was extracted using MeOH. Crude 2.17 (theoretical maximum: 1.02 g, 2.01 mmol) was extracted from the lyophilized Abg2F6 reaction mixture using MeOH, and acetylated according to the **General Procedure for Acetylation of α-Glycosyl Fluoride**

**Products of Glycosynthase Reactions.** The crude acetylated carbohydrates were purified by flash column chromatography, using an eluent of 1.5:1 EtOAc:PE, yielding 2.20 as a clear, colorless oil (1.34 g, 1.45 mmol, 72%). **Rf** (1.5:1 EtOAc:PE): 0.36 **1H NMR** (600 MHz, CDCl3): δ ppm 1.95 (s, 3H, Ac), 1.99 (s, 3H, Ac), 2.01 (s, 6H, 2 Ac), 2.03 (s, 3H, Ac), 2.05 (s, 3H, Ac), 2.08 (s, 3H, Ac), 2.12 (s, 3H, Ac), 2.13 (s, 6H, 2 Ac), 3.58-3.65 (m, 1H), 3.75-3.88 (m, 3H), 4.03-4.17 (m, 6H), 4.37 (d, 1H, J 11.7 Hz), 4.43 (d, 1H, J 7.5 Hz), 4.50-4.56 (m, 2H), 4.81-4.90 (m, 2H), 4.93 (dd, 1H, J 3.5 J 10.7 Hz), 5.07 (dd, 1H, J 10.2 Hz), 5.13 (dd, 1H, J 9.0 Hz), 5.32 (d, 1H, J 2.7 Hz), 5.45 (dd, 1H, J 9.9 Hz), 5.64 (dd, 1H, JH2-H1 2.3 JH1-F 53.0 Hz, H1) **19F NMR** (282 MHz, CDCl3) δ ppm -148.83 (dd, J 24.3 J 53.0 Hz) **13C NMR** (100 MHz, CDCl3): δ ppm 20.61 (CH3), 20.67 (CH3), 20.72 (CH3), 20.75 (CH3), 20.85 (CH3), 20.95 (CH3), 60.93, 61.23, 62.45, 66.70, 68.85, 69.19, 70.45, 70.70, 70.79, 70.87, 71.06, 71.98, 72.87, 73.13, 75.49, 76.09, 100.26, 101.21, 102.63 (C1), 169.24 (C=O), 169.36 (C=O), 169.54 (C=O), 169.88 (C=O), 170.16 (C=O), 170.24 (C=O), 170.26 (C=O), 170.32 (C=O), 170.38 (C=O), 170.48 (C=O)

**HRMS** calculated for C38H51FO25: 926.2703, found: 949.2601 [M+Na]+. The acetate groups were removed according to the **General Procedure for Deacetylation**, yielding 388 mg of white crystals (2.17, 0.767 mmol, 53%). **Rf** (7:2:1 EtOAc:MeOH:H2O): 0.15 **1H NMR** (400 MHz, MeOD): δ ppm 3.43-3.67 (m, 8H), 3.66-3.73 (m, 2H), 3.74-3.88 (m, 6H), 3.89-3.98 (m, 2H, J 3.6
$J\ 13.0\ Hz)$, 4.35 (d, 1H, $J\ 7.5\ Hz$), 4.47 (d, 1H, $J\ 7.9\ Hz$), 5.54 (dd, 1H, $J_{H2-H1}\ 2.7\ J_{H1-F}\ 53.6\ Hz$, H1) $^{19}$F NMR (282 MHz, MeOD) $\delta$ ppm -151.85 (dd, $J\ 25.9\ J\ 53.6\ Hz$) $^{13}$C NMR (100 MHz, MeOD): $\delta$ ppm 61.21, 61.70, 62.66, 72.69, 73.03, 73.28, 74.62, 74.65, 74.68, 74.96, 76.33, 76.81, 77.26, 79.48, 80.22, 104.51, 105.24, 107.75 (C1) HRMS calculated for C$_{18}$H$_{31}$FO$_{15}$: 506.1647, found: 529.1545 [M+Na]$^+$. (β-Δ-Glucopyranosyl)-(1→4)-β-Δ-glucopyranosyl-(1→4)-α-Δ-glucopyranosyl fluoride (2.18, 2.21)

Compound 2.18 was prepared according to the General Procedure for Synthesis of Oligosaccharides using Abg2F6 Glycosynthase. α-Glucosyl fluoride (2.04, 2.85 g, 15.6 mmol, 5.3 eq.) and α-cellobiosyl fluoride (2.12, 1.01 g, 2.93 mmol) were dissolved in 100 mM sodium phosphate (pH 7.0), and Abg2F6 (0.10 mg/mL) was added. The reaction was monitored by TLC (7:2:1 EtOAc:MeOH:H$_2$O), and was complete overnight. The reaction mixture was lyophilized, and crude 2.18 was extracted using MeOH. Crude 2.18 (theoretical maximum: 1.48 g, 2.93 mmol) was extracted from the lyophilized Abg2F6 reaction mixture using MeOH, and acetylated according to the General Procedure for Acetylation of α-Glycosyl Fluoride Products of Glycosynthase Reactions. The crude acetylated carbohydrates were purified by flash column chromatography, using an eluent of 1:1-1.5:1 EtOAc:PE, yielding 2.21 as a clear, colorless oil (1.08 g, 1.17 mmol, 40%). Rf (1:1 EtOAc:PE): 0.44 $^1$H NMR (400 MHz, CDCl$_3$): $\delta$ ppm 1.95 (s, 3H, Ac), 1.97 (s, 9H, 3 Ac), 1.99 (s, 3H, Ac), 2.00 (s, 3H, Ac), 2.01 (s, 3H, Ac), 2.06 (s, 6H, 2
Ac, 2.11 (s, 3H, Ac), 3.60 (m, 2H, H5', H5''), 3.77 (m, 2H, H4, H4''), 4.05 (m, 4H, H5, H6, H6''a, H6''''a), 4.32 (dd, 1H, $J_{H2'-H1}$ 4.4 $J_{H6''-H6''''}$ 12.4 Hz, H6''b), 4.38 (dd, 1H, $J_{H6''-H6''''}$ 12.0 Hz, H6''''b), 4.45 (d, 1H, $J_{H2''-H1''}$ 8.0 Hz, H1''), 4.51 (dd, 2H, $J_{H6-H6'}$ 11.6 $J_{H2-H1}$ 8.0 Hz, H1', H6'), 4.85 (m, 3H, H2, H2', H2''), 5.02 (dd, 1H, $J_{H4'}$ 9.6 $J_{H4}$ 9.2 Hz, H3'), 5.09 (dd, 2H, $J_{H3'-H3''}$ 12.0 Hz, H3', H3''), 5.42 (dd, 1H, $J_{H3}-F$ 5.3 Hz, H1), 5.64 (dd, 1H, $J_{H2}-H1$ 2.4 $J_{H1-F}$ 53.0 Hz, H1) 19F NMR (282 MHz, CDCl3) δ ppm -149.45 (dd, $J_{H1-F}$ 24.3 $J_{H1-F}$ 53.0 Hz) 13C NMR (100 MHz, CDCl3): δ ppm 20.55 (CH3), 20.61 (4 CH3), 20.64 (CH3), 20.73 (CH3), 20.80 (CH3), 20.90 (CH3), 61.17 (C6), 61.58 (C6'), 62.27 (C6''), 67.81, 68.76, 70.39, 70.75, 71.65, 71.83, 72.11, 72.74, 72.85, 72.96, 75.48, 76.23, 100.31 (C1''), 100.74 (C1'), 104.86 (C1) 169.15 (C=O), 169.25 (C=O), 169.36 (C=O), 169.49 (C=O), 169.85 (C=O), 169.85 (C=O), 170.19 (C=O), 170.25 (2 C=O), 170.56 (C=O) HRMS calculated for C38H51FO25: 926.2703, found: 949.2601 [M+Na]+.

The acetate groups were removed according to the General Procedure for Deacetylation, yielding 564 mg of white crystals (2.18, 1.11 mmol, 96%). Rf (7:2:1)

EtOAc:MeOH:H2O: 0.21 Rf (4:5:1.5 CHCl3:MeOH:0.2% CaCl2): 0.62 1H NMR (400 MHz, MeOD): δ ppm 3.23 (dd, 1H, $J_{H1}$ 8.1 $J_{H7}$ 8.7 Hz), 3.33-3.40 (m, 2H), 3.44-3.51 (m, 2H), 3.51-3.57 (m, 2H), 3.60 (dd, 1H, $J_{H9}$ 9.0 $J_{H9'}$ 9.2 Hz), 3.63-3.72 (m, 2H), 3.74-3.85 (m, 4H), 3.85-3.91 (m, 3H), 3.94 (dd, 1H, $J_{H3}$ 3.6 $J_{H12}$ 12.8 Hz), 4.40 (d, 1H, $J_{H7}$ 7.8 Hz), 4.46 (d, 1H, $J_{H7'}$ 7.9 Hz), 5.54 (dd, 1H, $J_{H12-H1}$ 2.6 $J_{H1-H1'}$ 53.8 Hz, H1) 19F NMR (282 MHz, MeOD) δ ppm -152.39 (dd, $J_{H1}$ 26.0 $J_{H1}$ 53.8 Hz) 13C NMR (100 MHz, MeOD): δ ppm 61.20, 61.65, 62.58, 71.51, 73.02, 73.27, 74.61, 74.64, 74.75, 75.04, 76.30, 78.00, 78.27, 79.46, 80.30, 104.48, 104.75, 107.74 HRMS calculated for C18H31FO15: 506.1647, found: 529.1545 [M+Na]+.
(β-D-Glucopyranosyl)-(1→4)-β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl fluoride (2.19, 2.22)

Compound 2.19 was prepared according to the General Procedure for Synthesis of Oligosaccharides using Abg2F6 Glycosynthase. α-Glucosyl fluoride (2.04, 2.85 g, 15.6 mmol, 5.3 eq.) and α-cellobiosyl fluoride (2.12, 1.01 g, 2.93 mmol) were dissolved in 100 mM sodium phosphate (pH 7.0), and Abg2F6 (0.10 mg/mL) was added. The reaction was monitored by TLC (7:2:1 EtOAc:MeOH:H₂O), and was complete overnight. The reaction was lyophilized, and crude 2.19 was extracted using MeOH. Crude 2.19 (theoretical maximum: 1.96 g, 2.93 mmol) was extracted from the lyophilized Abg2F6 reaction mixture using MeOH, and acetylated according to the General Procedure for Acetylation of α-Glycosyl Fluoride Products of Glycosynthase Reactions. The crude acetylated carbohydrates were purified by flash column chromatography, using an eluent of 1:1-1.5:1 EtOAc:PE, yielding 2.22 as a clear, colorless oil (0.560 g, 0.461 mmol, 16%). Rf (1:1 EtOAc:PE): 0.37 ¹⁹F NMR (282 MHz, CDCl₃) δ ppm -149.57 (dd, J 24.0, J 53.0 Hz) HRMS calculated for C₅₀H₆₇FO₃₃: 1214.3549, found: 1237.3446 [M+Na]⁺. The acetate groups were removed according to the General Procedure for Deacetylation, yielding 276 mg of a white powder (2.19, 0.413 mmol, 90%). Rf (7:2:1 EtOAc:MeOH:H₂O): 0.076 Rf (4:5:1.5 CHCl₃:MeOH:0.2% CaCl₂): 0.58 ¹H NMR (400 MHz, MeOD): δ ppm 3.21 (dd, 1H, J 5.4, J 6.0 Hz), 3.25-3.30 (m, 3H), 3.32-3.38 (m, 3H), 3.47-3.59 (m, 6H), 3.60-3.70 (m, 3H), 3.75 (dd, 1H, J 6.2 Hz), 3.78-3.83 (m, 2H), 3.83-3.89 (m, 3H), 3.90-3.96 (m, 2H), 4.41 (d, 1H, J 5.3 Hz), 4.47 (d, 1H, J 5.3 Hz), 4.48 (d, 1H, J 5.3 Hz), 4.61 (dd, 1H,
\[ J_{H1-F} 53.9 \text{ Hz, H1} \] \(^{19}\text{F NMR}\) (282 MHz, MeOD) \( \delta \text{ ppm -150.88 (dd, } J 26.5 J 53.9 \text{ Hz)} \)

\(^{13}\text{C NMR}\) (100 MHz, MeOD): \( \delta \text{ ppm 61.18, 61.66, 61.75, 62.58, 70.16, 71.50, 72.89, 74.67, 74.97, 76.21, 76.30, 76.74, 77.99, 78.34, 79.73, 80.53, 80.73, 80.95, 85.89, 88.94, 104.47, 104.79, 108.28, 109.76 } \]

\( \text{HRMS} \) calculated for \( C_{24}H_{41}FO_{20} \): 668.2175, found: 691.2073 \([\text{M+Na}]^+ \).

\((5\text{-Acetamido-3,5-dideoxy-\(\alpha\)-d-glycero-\(\delta\)-galacto-2-nonulopyranosyl})ionic\text{ acid}]-(2\rightarrow3)-O-(\beta\)-d-galactopyranosyl)-(1\rightarrow4)-\(\alpha\)-d-glucopyranosyl fluoride (2.24)

Compound 2.24 was prepared according to the \textbf{General Procedure for Sialoside Synthesis} using \textbf{Cst-I Enzyme}. Lactosyl fluoride (5.15 mg, 15.0 \( \mu \text{mol, 2.16} \)) and CMP-Neu5Ac (12.3 mg, 20.0 \( \mu \text{mol, 1.3 eq., 2.23} \)) were dissolved in 50 mM HEPES (pH 7.5) and 10 mM MnCl\(_2\) to give an acceptor concentration of 10 mM. Cst-I (0.20 mg/mL) and 0.1 \( \mu \text{L alkaline phosphatase (10 KU solution)} \) were added, and the reaction was left at room temperature. The reaction was complete after 2 hours according to TLC (7:2.5:1 EtOAc:MeOH:H\(_2\)O) and lyophilized. Crude 2.24 was used in reactions without further purification. \( \text{Rf} \) (7:2.5:1 EtOAc:MeOH:H\(_2\)O): 0.14

\( \text{HRMS} \) calculated for \( C_{23}H_{37}FNO_{18} \): 634.2000, found: 634.1995 \([\text{M}]^+ \).
[(5-Acetamido-3,5-dideoxy-α-D-glycero-D-galacto-2-nonulopyranosyl)onic acid]-(2→3)-O-(β-D-galactopyranosyl)-(1→4)-β-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl fluoride (2.25)

Compound 2.25 was prepared according to the **General Procedure for Sialoside Synthesis using** Cst-I **Enzyme**. Compound 2.17 (19.9 mg, 39.3 µmol) and CMP-Neu5Ac (30.6 mg, 50.0 µmol, 1.3 eq., 2.23) were dissolved in 50 mM HEPES (pH 7.5) and 10 mM MnCl₂ to give an acceptor concentration of 10 mM. Cst-I (0.20 mg/mL) and 0.1 µL alkaline phosphatase (10 KU solution) were added, and the reaction was left at room temperature. The reaction was complete after 2 hours according to TLC (4:5:1 CHCl₃:MeOH:0.2% CaCl₂) and lyophilized. Crude 2.25 was used in reactions without further purification. Rf (4:5:2 CHCl₃:MeOH:0.2% CaCl₂): 0.32

**HRMS** calculated for C₂₉H₄₇FNO₂₃: 796.2528, found: 796.2523 [M⁺].

(6-O-phospho-β-D-glucopyranosyl)-(1→4)-α-D-glucopyranosyl fluoride (2.27)

Compound 2.27 was prepared according to the **General Procedure for Phosphorylation using** BglK **Enzyme**. α-Cellobiosyl fluoride (68.7 mg, 200 µmol, 2.12) was dissolved in 25 mM HEPES (pH 7.5) containing 2 mM MgSO₄. Once dissolved, ATP (240 µmol, 1.2 eq., adjusted to pH 7.5 with 3 M NH₄OH) was added to the reaction mixture. BglK (0.50 mg/mL) was added, and the reaction mixture was incubated at room temperature. The reaction was complete after 2.5
hours according to TLC. BglK enzyme was removed by filtration through an Amicon® Ultra-15 3K centrifugal filter (EMD Millipore®). The reaction mixture was lyophilized, and 2.27 was used without further purification. Rf (4:5:1.5 CHCl₃:MeOH:0.2% CaCl₂): 0.36 HRMS calculated for C₁₂H₂₀FO₁₃P: 422.0637, found: 423.0704 [M+H].

(6-O-phospho-β-D-glucopyranosyl)-(1→4)-β-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl fluoride (2.28)

Compound 2.28 was prepared according to the General Procedure for Phosphorylation using BglK Enzyme. α-Cellotriosyl fluoride (44.5 mg, 87.9 µmol, 2.18) was dissolved in 25 mM HEPES (pH 7.5) containing 2 mM MgSO₄. Once dissolved, ATP (105 µmol, 1.2 eq., adjusted to pH 7.5 with 3 M NH₄OH) was added to the reaction mixture. BglK (0.75 mg/mL) was added, and the reaction was incubated at room temperature. The reaction was complete overnight according to TLC (4:5:1 CHCl₃:MeOH:0.2% CaCl₂). BglK enzyme was removed by filtration through an Amicon® Ultra-15 3K centrifugal filter (EMD Millipore®). The reaction mixture was lyophilized, and 2.28 was used without further purification. Rf (4:5:1.5 CHCl₃:MeOH:0.2% CaCl₂): 0.30 HRMS calculated for C₁₈H₃₀FO₁₈P: 584.1165, found: 585.1232 [M+H].
(6-O-phospho-β-D-glucopyranosyl)-(1→4)-β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl-
(1→4)-α-D-glucopyranosyl fluoride (2.29)

Compound 2.29 was prepared according to the General Procedure for Phosphorylation using
BglK Enzyme. α-Cellotetraosyl fluoride (66.4 mg, 99.3 µmol, 2.19) was dissolved in 25 mM
HEPES (pH 7.5) containing 2 mM MgSO₄. Once dissolved, ATP (120 µmol, 1.2 eq., adjusted to
pH 7.5 with 3 M NH₄OH) was added to the reaction mixture. BglK (0.75 mg/mL) was added,
and the reaction was incubated at room temperature. The reaction was complete overnight
according to TLC (4:5:1 CHCl₃:MeOH:0.2% CaCl₂). BglK enzyme was removed by filtration
through an Amicon® Ultra-15 3K centrifugal filter (EMD Millipore®). The reaction mixture was
lyophilized, and 2.29 was used without further purification. Rf (4:5:1.5 CHCl₃:MeOH:0.2%
CaCl₂): 0.28 HRMS calculated for C₂₄H₄₀FO₂₃P: 746.1693, found: 747.1760 [M+H]+.

(2,3,4,6-Tetra-O-acetyl-β-D-glucopyranosyl)-(1→4)-O-2,3,6-tri-O-acetyl-D-glucopyranose (2.30)

D-Cellobiose octaacetate (5.99 g, 8.83 mmol, 2.10) and hydrazine acetate (0.982 g, 10.7 mmol,
1.2 eq.) were weighed into a round bottom flask, and 30 mL DMF was added. The mixture was
stirred at 50°C for 10 minutes until the suspension was clear. The reaction was stirred at room
temperature, and monitored every 30 minutes by TLC. The reaction was complete after 4 hours
and was diluted with 500 mL EtOAc. The organic layer was washed successively with 2x200 mL saturated NaHCO₃, 200 mL dH₂O, and 2x200 mL brine. The organic layer was dried over sodium sulfate, filtered, and concentrated in vacuo to give a white solid. The solid was crystallized using ethanol, yielding 2.30 as a white powder (4.69 g, 7.37 mmol, 83%). ¹H data agreed with that reported previously.¹⁵² Rf (2:1 EtOAc:Hexanes): 0.21 ¹H NMR (400 MHz, CDCl₃): δ ppm 1.98 (s, 3H, Ac), 2.00 (s, 3H, Ac), 2.02 (s, 6H, 2 Ac), 2.07 (s, 3H, Ac), 2.08 (s, 3H, Ac), 2.13 (s, 3H, Ac), 3.29 (bs, 1H, OH), 3.67 (m, 2H, H4, H5’), 3.76 (m, 1H, J₉₃₋₉₂ 9.6 Hz, H2), 4.11 (m, 2H, J₉₆₋₉₆’’ 11.6 Hz, H6, H6’’), 4.36 (dd, 1H, J₉₅₋₉₆’ 4.4 J₉₆₋₉₆’’ 12.4 Hz, H6’), 4.52 (m, 2H, J₉₆’’₋₉₆’ 11.6 Hz, H₃’, H₆’’), 4.81 (m, 1H, H5), 4.93 (m, 1H, H2’), 5.02-5.25 (m, 2H, H3, H4’), 5.36 (bm, 1H, H1), 5.50 (dd, 1H, J₉₂’₋₉₁’ 9.6 Hz, H1’) ¹³C NMR (100 MHz, CDCl₃): δ ppm 20.68 (3 CH₃), 20.76 (CH₃), 20.80 (CH₃), 20.88 (CH₃), 21.03 (CH₃), 61.73 (C6), 61.89 (C6’), 67.98, 68.43, 69.42, 71.35, 71.80, 72.06, 73.13, 76.64, 90.25 (C1), 100.80, 169.18 (C=O), 169.46 (C=O), 169.82 (C=O), 170.44 (C=O), 170.58 (C=O), 170.72 (C=O), 171.13 (C=O) ESI MS calculated for C₂₆H₃₆O₁₈: 636.6, found: 659.4 [M+Na]⁺.

2,4-Dinitrophenyl (2,3,4,6-tetra-O-acetyl-β-D-glucopyranosyl)-(1→4)-O-2,3,6-tri-O-acetyl-D-glucopyranoside (2.32)

Compound 2.30 (4.69 g, 7.37 mmol), DABCO (3.30 g, 29.5 mmol, 4 eq.), and DNFB (2.06 g, 11.1 mmol, 1.5 eq., 2.31) were weighed out in a round bottom flask and dissolved in 150 mL
DMF. The clear orange reaction was left to stir at room temperature overnight. Once complete, the reaction was diluted with 550 mL EtOAc, and the organic layer was washed successively with 3x250 mL saturated NaHCO₃ and 2x250 mL brine. The organic layer was dried over sodium sulfate, filtered, and concentrated in vacuo to yield a crude yellow solid. The solid was crystallized using EtOAc and hexanes, yielding 2.32 as fluffy white crystals (3.98 g, 4.96 mmol, 67%). ¹H data agreed with that reported previously.¹² Rf (2:1 EtOAc:Hexanes): 0.59. ¹H NMR (400 MHz, CDCl₃): δ ppm 1.99 (s, 3H, Ac), 2.02 (s, 3H, Ac), 2.03 (s, 3H, Ac), 2.04 (s, 3H, Ac), 2.07 (s, 3H, Ac), 2.10 (s, 3H, Ac), 2.11 (s, 3H, Ac), 3.69 (m, 1H, J_H6'-H5' 5.6 Hz, H5'), 3.90 (m, 1H, H4), 3.99 (dd, 1H, J_H2-H3 8.4 J_H4-H3 9.6 Hz, H3), 4.09 (m, 2H, J_H5-H6' 5.6 J_H6-H6'. 10.0 Hz, H6', H6'''), 4.37 (dd, 1H, J_H5'-H6' 4.4 J_H6''-H6' 12.4 Hz, H6'''), 4.59 (m, 1H, J_H6'-H6 10.0 J_H2'-H1' 8.0 Hz, H1', H6), 4.95 (dd, 1H, J_H1'-H2' 8.0 J_H3'-H2' 9.2 Hz, H2'), 5.08 (dd, 1H, J_H3'-H4' 9.6 Hz, H4'), 5.17 (dd, 1H, J_H2'-H3' 9.2 J_H4'-H3' 9.6 Hz, H3'), 5.23 (m, 2H, J_H1-H2 9.2 Hz, H2, H5), 5.32 (dd, 1H, J_H2-H1 9.2 Hz, H1), 7.43 (d, 1H, J_H1-H2 9.2 Hz, H6(Ar)), 8.41 (dd, 1H, J_H2-H1 2.8 J 9.2 Hz, H5(Ar)), 8.69 (d, 1H, J_H2-H1 2.4 Hz, H3(Ar)) ¹³C NMR (100 MHz, CDCl₃): δ ppm 20.65 (2 CH₃), 20.67 (3 CH₃), 20.81 (2 CH₃), 61.48 (C6), 61.64 (C6'), 67.85 (C4'), 70.71 (C2), 71.71 (C2'), 72.05 (C5), 72.30 (C5'), 72.92 (C3'), 73.48 (C4), 75.75 (C3), 98.84 (C1), 101.00 (C1'), 118.49 (C6(Ar)), 121.49 (C3(Ar)), 128.70 (C5(Ar)) ESI MS calculated for C₃₂H₃₈N₂O₂₂: 802.6, found: 825.5 [M+Na]^+. 
2,4-Dinitrophenyl-(β-D-glucopyranosyl)-(1→4)-β-D-glucopyranoside (2.33)

Per-O-acetylated DNPC (3.98 g, 4.96 mmol, 2.32) was suspended in 200 mL MeOH, and the reaction mixture was stirred in an ice water bath. AcCl (16 mL, 225 mmol, 45 eq.) was added dropwise over 20 minutes, and the suspension was left to stir at 4°C for 72 hours. The mixture was concentrated in vacuo, and the white residue was crystallized using MeOH and Et₂O, yielding 2.33 as a white powder (1.81 g, 3.56 mmol, 72%). ¹H data agreed with that reported previously.¹² Rf (4:4:10 drops CHCl₃:MeOH:H₂O): 0.67 Rf (7:2.5:1 EtOAc:MeOH:H₂O): 0.63

¹H NMR (400 MHz, MeOD): δ ppm 3.24 (dd, 1H, J_H2'-H3: 8.4 J_H4'-H3: 8.4 Hz, H3’), 3.28-3.40 (m, 3H, H2’, H4’, H5), 3.59 (dd, 1H, J_H1-H2: 7.6 J_H3-H2: 8.4 Hz, H2), 3.70 (m, 4H, H3, H4, H5’, H6”), 3.92 (m, 3H, J>H6-H6: 12.0 Hz, H6, H6’, H6”’), 4.47 (d, 1H, J_H2'-H1: 8.0 Hz, H1’), 5.36 (d, 1H, J_H2-H1: 7.6 Hz, H1), 7.68 (d, 1H, J 9.6 Hz, H6(Ar)), 8.51 (dd, 1H, J 2.8 J 9.6 Hz, H5(Ar)), 8.76 (d, 1H, J 2.8 Hz, H3(Ar))

¹³C NMR (100 MHz, MeOD): δ ppm 61.40 (C6), 62.47 (C6’), 71.39 (C4’), 74.22 (C2), 74.88 (C2’), 76.32 (C5), 77.16 (C5’), 77.90 (C3’), 78.19 (C4), 79.83 (C3), 101.80 (C1), 104.57 (C1’), 118.83 (C6(Ar)), 122.24 (C3(Ar)), 129.84 (C5(Ar))

Propionaldehyde-\textit{O}-benzyl oxime (Z+E) (2.38)

\begin{center}
\includegraphics[width=0.3\textwidth]{image.png}
\end{center}

\textit{O}-Benzylhydroxylamine hydrochloride (1.01 g, 6.31 mmol, 2.36) was added to a stirred solution of NaOAc (0.630 g, 7.68 mmol, 1.2 eq.) dissolved in 100 mL of 4:1 H₂O:MeOH.

Propionaldehyde (1.37 mL, 19.1 mmol, 3 eq., 2.37) was added, and the reaction mixture was stirred at room temperature for 2 hours. The product was extracted with DCM (3 x 50 mL), and washed with saturated NaHCO₃. The DCM layer was dried over MgSO₄, filtered, and concentrated \textit{in vacuo}, yielding compound 2.38 as a colorless oil (0.99 g, 6.07 mmol, 97\%). $^1$H and $^{13}$C NMR data agreed with that reported previously. $^1$H NMR (400 MHz, CDCl₃): δ ppm 1.07 (t, 3H, $^\text{J}_{\text{H6-H7}}$ 7.6 Hz, 3 H7), 1.09 (t, 3H, $^\text{J}_{\text{H6-H7}}$ 7.2 Hz, 3 H7), 2.23 (qd, 2H, $^\text{J}_{\text{H5-H6}}$ 6.0 $^\text{J}_{\text{H7-H6}}$ 7.2 Hz, 2 H6), 2.40 (qd, 2H, $^\text{J}_{\text{H5-H6}}$ 5.6 $^\text{J}_{\text{H7-H6}}$ 7.6 Hz, 2 H6), 5.07 (s, 2H, 2 H2), 5.12 (s, 2H, 2 H2), 6.68 (t, 1H, $^\text{J}_{\text{H6-H5}}$ 5.6 Hz, H5), 7.29-7.42 (m, 10H, 2 x 5 H$_{\text{Ar}}$), 7.47 (t, 1H, $^\text{J}_{\text{H6-H5}}$ 6.0 Hz, H5) $^{13}$C NMR (100 MHz, CDCl₃): δ ppm 10.78 (C7), 11.21 (C7), 19.43 (C6), 23.22 (C6), 75.66 (C2), 75.81 (C2), 127.85 (C$_{\text{Ar}}$), 127.94 (C$_{\text{Ar}}$), 128.01 (C$_{\text{Ar}}$), 128.39 (C$_{\text{Ar}}$), 128.52 (C$_{\text{Ar}}$), 137.83 (C1), 138.27 (C1), 152.61 (C5), 153.82 (C5) ESI MS calculated for C$_{10}$H$_{13}$NO: 163.1, found: 164.3 [M+H]$^+$.

\textit{N-nPropyl-\textit{O}-benzylhydroxylamine} (2.39)

\begin{center}
\includegraphics[width=0.3\textwidth]{image.png}
\end{center}
Compound 2.38 (0.99 g, 6.07 mmol) was dissolved in 50 mL of EtOH, and NaBH$_3$CN (1.49 g, 23.7 mmol, 3.9 eq.) was added. The reaction mixture was acidified with concentrated HCl to pH 3, and the reaction was stirred at room temperature for 30 minutes. The reaction mixture was concentrated in vacuo, and the remaining white precipitate was dissolved in 0.5 M NaOH until pH 8 and extracted with DCM (3 x 20 mL). The DCM layer was washed with 60 mL H$_2$O, dried over MgSO$_4$, filtered, and concentrated in vacuo, yielding 2.39 as a colorless oil (0.85 g, 5.15 mmol, 85%). $^1$H and $^{13}$C NMR data agreed with that reported previously.$^{110}$

$^1$H NMR (400 MHz, CDCl$_3$): δ ppm 0.94 (t, 3H, $J_{H6-H7}$ 7.6 Hz, 3 H7), 1.58 (m, 2H, $J_{H5-H6}$ 7.2 $J_{H7-H6}$ 7.6 Hz, 2 H6), 2.93 (t, 1H, $J_{H6-H7}$ 7.2 Hz, 2 H5), 4.76 (s, 2H, 2 H2), 5.22 (s, 1H, H4), 7.28-7.38 (m, 5H, 5 H$_{(Ar)}$)

$^{13}$C NMR (100 MHz, CDCl$_3$): δ ppm 11.71 (C7), 20.48 (C6), 53.94 (C5), 76.35 (C2), 127.94 (C$_{(Ar)}$) 128.52 (4 C$_{(Ar)}$), 137.83 (C1) ESI MS calculated for C$_{10}$H$_{15}$NO: 165.1, found: 166.3 [M+H]$^+$.

$N$-$n$Propyl-$O$-benzyl-$N$-($β$-$D$-glucopyranosyl)hydroxylamine (2.40)

$\text{\includegraphics[width=0.3\textwidth]{diagram.png}}$

D-Glucose (0.207 g, 1.15 mmol, 2.01) was dissolved in THF:AcOH (32 mL, 3:1 v/v), and compound 2.39 (0.281 g, 1.70 mmol, 1.48 eq.) was added. The reaction mixture was stirred at room temperature for 20 hours, and then concentrated in vacuo. The residue was purified by flash column chromatography (9:1 EtOAc:MeOH) yielding 2.40 as a white solid (0.219 g, 0.669 mmol, 58%). $^1$H and $^{13}$C NMR data agreed with that reported previously.$^{110}$ Rf (7:2:1
EtOAc:MeOH:H$_2$O): 0.74 $^1$H NMR (400 MHz, MeOD): $\delta$ ppm 0.99 (t, 3H, $J_{H8-H9}$ 7.2 Hz, 3 H9), 1.66 (m, 2H, $J_{H9-H8}$ 7.2 $J_{H7-H8}$ 7.2 Hz, 2 H8), 2.92 (dt, 1H, $J_{H7'-H7}$ 13.2 $J_{H8-H7}$ 7.2 Hz, H7), 3.04 (dt, 1H, $J_{H7-H7'}$ 13.2 $J_{H8-H7}$ 7.6 Hz, H7’), 3.15-3.25 (m, 2H, H4, H5), 3.39 (dd, 1H, $J_{H2-H3}$ 8.8 $J_{H3-H2}$ 8.8 Hz, H3), 3.57 (dd, 1H, $J_{H7-H8}$ 8.8 Hz, H2), 3.66 (dd, 1H, $J_{H6-H7}$ 4.6 Hz, H6), 3.80 (dd, 1H, $J_{H6-H6'}$ 11.6 Hz, H6’), 4.09 (d, 1H, $J_{H2-H1}$ 8.8 Hz, H1), 4.72-4.96 (m, 2H, 2 H10), 7.24-7.44 (m, 5H, 5 H$_{(Ar)}$) $^{13}$C NMR (100 MHz, MeOD): $\delta$ ppm 11.03 (C9), 20.53 (C8), 54.72 (C7), 61.38 (C6), 69.90 (C4), 70.39 (C2), 76.73 (C10), 78.26 (C5), 78.38 (C3), 93.12 (C1), 127.99 (C$_{(Ar)}$), 128.19 (2 C$_{(Ar)}$), 129.07 (2 C$_{(Ar)}$), 136.88 (C11) HRMS calculated for C$_{16}$H$_{25}$NO$_6$: 327.1682, found: 328.1760 [M+H].

$N$-$n$Propyl-$O$-benzyl-$N$-($\beta$-$D$-glucopyranosyl(1$\rightarrow$4)$\beta$-$D$-glucopyranosyl)hydroxylamine (2.42)

D-Cellobiose (0.198 g, 0.578 mmol, 2.09) was dissolved in 10 mL of 100 mM NaOAc (pH 4.5), and compound 2.39 (0.386 g, 2.34 mmol, 4 eq.) was added. The reaction mixture was stirred at room temperature for 72 hours, and then concentrated in vacuo. The residue was purified on a Waters Sep-Pak® tC18 cartridge (2 g) using a gradient of acetonitrile/water. The product eluted at 25% acetonitrile, and the fractions were concentrated in vacuo, yielding 2.42 as a white solid (39 mg, 79.7 µmol, 14%). Rf (7:2:1 EtOAc:MeOH:H$_2$O): 0.63 $^1$H NMR (400 MHz, D$_2$O) $\delta$ ppm 0.97 (t, 3H, $J_{H8-H9}$ 7.6 Hz, 3 H9), 1.67 (m, 2H, $J_{H9-H8}$ 7.6 $J_{H7-H8}$ 7.6 Hz, 2 H8), 2.94 (dt, 1H, $J_{H7-H7'}$ 12.8 $J_{H8-H7}$ 7.6 Hz, H7), 3.11 (dt, 1H, $J_{H7-H7'}$ 12.8 $J_{H7-H8}$ 7.6 Hz, H7’), 3.32 (dd, 1H, $J_{H4-H5}$ 8.8
$J_{\text{H3-H4}}$ 8.4 Hz, H4), 3.42 (dd, 1H, $J_{\text{H3-H4}}$ 9.6 $J_{\text{H5-H4}}$ 9.2 Hz, H4”), 3.52 (m, 3H, H2, H2”, H6”), 3.62 (dd, 1H, $J_{\text{H2-H3}}$ 8.4 $J_{\text{H4-H3}}$ 9.6 Hz, H3”), 3.69 (dd, 2H, $J_{\text{H3-H2}}$ 8.4 Hz, $J_{\text{H4-H3}}$ 8.4 Hz, H3), 3.74 (dd, 1H, $J_{\text{H6-H5}}$ 5.6 $J_{\text{H4-H5}}$ 9.2 Hz, H5”), 3.78 (dd, 1H, $J_{\text{H6-H5}}$ 4.4 $J_{\text{H4-H5}}$ 8.8 Hz, H5) 3.80 (dd, 1H, $J_{\text{H2-H1}}$ 8.8 Hz, H1), 4.51 (d, 1H, $J_{\text{H2-H1}}$ 8.0 Hz, H1”), 4.85 (m, 2H, 2 H10), 7.41-7.51 (m, 5H, 5 H(Ar)) $^{13}$C NMR (100 MHz, D$_2$O): δ ppm 11.22 (C9), 20.13 (C8), 55.00 (C7), 60.17 (C6”), 60.70 (C6), 69.57, 69.82, 73.28, 75.60, 75.79, 76.10, 76.18, 76.96, 78.49, 91.94 (C1), 102.65 (C1”), 128.88 (2 C(Ar)), 128.94 (C(Ar)), 129.60 (2 C(Ar)), 135.75 (C11) HRMS calculated for C$_{22}$H$_{35}$NO$_{11}$: 489.2210, found: 490.2288 [M+H]$^+$. 

$O$-(β-D-Galactopyranosyl)-(1→4)-β-D-glucopyranosyl-(1→1)-(2S, 3R, 4E)-2-aminooctadec-4-ene-1,3-diol (3.02)

Lactosyl sphingosine (3.02) was prepared according to the General Procedure for Ganglioside Synthesis using EGCcase Glycosynthase. α-Lactosyl fluoride (66.5 mg, 0.193 mmol, 1.5 eq., 2.16) and d-erythro-C18-sphingosine hydrochloride (44.5 mg, 0.132 mmol, 3.01) were dissolved in a glass vial along with 25 mM NaOAc (pH 5.3) and DME (1.32 mL, 10% v/v). The reaction mixture was sonicated until the solution was clear. EGCcase glycosynthase was added (0.5 mg/mL), and the mixture was incubated in a 37°C water bath for 4 days. The reaction was monitored by TLC (5:4:1 CHCl$_3$:MeOH:10% AcOH) and lyophilized once complete. The product was purified on a Waters Sep-Pak$^\text{®}$ tC18 cartridge (5 g), with product eluting at 50%
acetonitrile. The fractions were lyophilized, yielding 3.02 as a white solid (66 mg, 0.106 mmol, 80%). \( Rf \) (4:5:1 CHCl\(_3\):MeOH:0.2% CaCl\(_2\)): 0.78 \(^1\)H NMR (400 MHz, MeOD): \( \delta \) ppm 0.90 (t, 3H, J 6.8 Hz, \( \text{CH}_3 \)), 1.21-1.38 (m, 21H), 1.38-1.49 (m, 2H, \( \text{Hg}', \text{Hg} \)), 2.06-2.15 (m, 2H, Hf, Hf'), 3.36-3.43 (m, 1H), 3.44-3.52 (m, 2H), 3.52-3.63 (m, 4H), 3.70 (dd, 1H, J 4.4 J 11.4 Hz), 3.75-3.88 (m, 3H), 3.88-4.02 (m, 3H), 4.32 (dd, 1H, J 5.6 J 5.8 Hz), 4.35 (d, 1H, J 7.2 Hz, H1), 4.37 (d, 1H, J 7.2 Hz, H1'), 5.44-5.53 (m, 1H, Hd), 5.83-5.93 (m, 1H, He) \(^{13}\)C NMR (100 MHz, MeOD): \( \delta \) ppm 14.21 (\( \text{CH}_3 \)), 22.98, 29.35, 29.66, 29.67, 29.83, 29.94, 29.97, 30.00, 32.25, 32.66, 55.74, 60.88, 61.92, 66.34, 69.41, 70.04, 71.50, 73.42, 73.87, 75.12, 75.55, 75.99, 79.57, 102.80, 104.06, 126.87 (Cd), 136.30 (Ce) HRMS calculated for \( \text{C}_{30}\text{H}_{58}\text{NO}_{12} \): 624.3954, found: 624.3959 [M+H]\(^+\).

\( \text{O-} \)\( \beta \)-d-Glucopyranosyl-\( (1 \rightarrow 4)\)-\( \beta \)-d-glucopyranosyl-\( (1 \rightarrow 1)\)-(2S, 3R, 4E)-2-aminoctadec-4-ene-1,3-diol (3.03)

Celllobiosyl sphingosine (3.03) was prepared according to the General Procedure for Ganglioside Synthesis using EGCase Glycosynthase. \( \alpha \)-Celllobiosyl fluoride (123.6 mg, 0.359 mmol, 1.5 eq., 2.12) and \( \delta \)-erythro-C18-sphingosine hydrochloride (80.0 mg, 0.238 mmol, 3.01) were dissolved in a glass vial along with 25 mM NaOAc (pH 5.3) and DME (2.38 mL, 10% v/v). The reaction mixture was sonicated until the solution was clear. EGCase glycosynthase was added (0.5 mg/mL), and the reaction was incubated in a 37°C water bath for 4 days. The reaction was monitored by TLC (5:4:1 CHCl\(_3\):MeOH:10% AcOH), lyophilized once complete and
purified on a Waters Sep-Pak® tC18 cartridge (5 g). The product eluted at 50% acetonitrile, and the fractions were lyophilized, yielding 3.03 as a white solid (106 mg, 0.170 mmol, 72%). Rf (4:5:1 CHCl₃:MeOH:0.2% CaCl₂): 0.64 ¹H NMR (400 MHz, MeOD): δ ppm 0.89 (t, 3H, J 6.8 Hz, CH₃), 1.20-1.36 (m, 21H), 1.37-1.47 (m, 2H, Hg, Hg’), 2.06-2.15 (m, 2H, Hf, Hf’), 3.24 (dd, 1H, J 8.2 J 8.6 Hz), 3.29-3.44 (m, 4H), 3.45-3.52 (m, 1H), 3.53-3.61 (m, 2H), 3.66 (dd, 1H, J 6.0 J 12.0 Hz), 3.83-3.97 (m, 4H), 4.01 (dd, 1H, J 8.7 Hz), 4.33 (dd, 1H, J 5.9 Hz), 4.38 (d, 1H, J 7.9 Hz, H1), 4.41 (d, 1H, J 7.9 Hz, H1’), 5.45-5.54 (m, 1H, Hd), 5.83-5.94 (m, 1H, He) ¹³C NMR (100 MHz, MeOD): δ ppm 14.21 (CH₃), 22.97, 29.33, 29.64, 29.82, 30.00, 32.24, 32.64, 55.70, 60.80, 61.57, 66.24, 69.99, 70.22, 73.47, 73.56, 75.14, 75.52, 76.80, 76.98, 79.48, 102.73, 103.60, 126.84 (Cd), 136.31 (Ce) HRMS calculated for C₃₀H₅₈NO₁₂: 624.3954, found: 624.3959 [M]+.

O-(β-d-Glucopyranosyl)-(1→4)-β-d-glucopyranosyl-(1→4)-β-d-glucopyranosyl-(1→1)-(2S, 3R, 4E)-2-aminooctadec-4-ene-1,3-diol (3.04)

Cellotriosyl sphingosine (3.04) was prepared according to the General Procedure for Ganglioside Synthesis using EGCase Glycosynthase. α-Cellotriosyl fluoride (80.1 mg, 0.158 mmol, 1.5 eq., 2.18) and d-erythro-C18-sphingosine hydrochloride (35.1 mg, 0.104 mmol, 3.01) were dissolved in a glass vial along with 25 mM NaOAc (pH 5.3) and DME (1.05 mL, 10% v/v). The reaction mixture was sonicated until the solution was clear. EGCase glycosynthase was added (0.5 mg/mL), and the reaction was incubated in a 37°C water bath for 5 days. The reaction
was monitored by TLC (5:4:1 CHCl₃:MeOH:10% AcOH) and lyophilized once complete. The reaction was purified on a Waters Sep-Pak® tC18 cartridge (5 g). The product eluted at 50% acetonitrile, and the fractions were lyophilized, yielding 3.04 as a white solid (66.0 mg, 0.0839 mmol, 80%). Rf (4:5:2 CHCl₃:MeOH:0.2% CaCl₂): 0.74

**1H NMR** (400 MHz, MeOD): δ ppm 0.90 (t, 3H, J 7.0 Hz, CH₃), 1.20-1.38 (m, 21H), 1.38-1.51 (m, 2H, Hg, Hg’), 2.05-2.16 (m, 2H, Hf, Hf’), 3.23 (dd, 1H, J 8.0 J 8.7 Hz), 3.26-3.43 (m, 11H), 3.43-3.62 (m, 6H), 3.66 (dd, 1H, J 5.5 J 11.7 Hz), 3.79-4.03 (m, 7H), 4.32 (dd, 1H, J 5.3 J 5.6 Hz), 4.37 (d, 1H, J 7.8 Hz), 4.40 (d, 1H, J 7.9 Hz), 4.44 (d, 1H, J 7.9 Hz), 5.43-5.54 (m, 1H, Hd), 5.81-5.93 (m, 1H, He)

**13C NMR** (100 MHz, MeOD): δ ppm 14.20 (CCH₃), 22.97, 29.33, 29.64, 29.81, 29.93, 29.96, 29.98, 32.23, 32.64, 55.70, 60.75, 60.79, 61.56, 70.02, 70.20, 73.24, 73.49, 73.56, 75.19, 75.50, 75.69, 76.78, 76.97, 79.24, 103.37, 103.57, 126.83 (Cd), 136.34 (Ce)

**HRMS** calculated for C₃₆H₆₈NO₁₇: 786.4482, found: 786.4487 [M⁺].

O-(β-D-Glucopyranosyl)-(1→4)-β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl-(1→1)-(2S, 3R, 4E)-2-aminooctadec-4-ene-1,3-diol (3.05)

Cellotetraosyl sphingosine (3.05) was prepared according to the **General Procedure for Ganglioside Synthesis using EGCase Glycosynthase**. α-Cellotetraosyl fluoride (71.9 mg, 0.108 mmol, 1.6 eq., 2.19) and d-erythro-C18-sphingosine hydrochloride (23.0 mg, 0.0684 mmol, 3.01) were dissolved in a glass vial along with 25 mM NaOAc (pH 5.3) and DME (0.335 mL, 10% v/v). The reaction mixture was sonicated until the solution was clear. EGCase
glycosynthase was added (0.5 mg/mL), and the reaction was incubated in a 37°C water bath for 5 days. The reaction was monitored by TLC (5:4:1 CHCl₃:MeOH:0.2% CaCl₂) and lyophilized once complete. The reaction was purified on a Waters Sep-Pak® tC18 cartridge (5 g). The product eluted at 50% acetonitrile, and the fractions were lyophilized, yielding 3.05 as a white solid (50.4 mg, 0.0531 mmol, 82%).

Rf (5:4:1 CHCl₃:MeOH:0.2% CaCl₂): 0.33 ¹H NMR (400 MHz, MeOD): δ ppm 0.94 (t, 3H, J 7.0 Hz, CH₃), 1.24-1.42 (m, 21H), 1.42-1.52 (m, 2H, Hg, Hg’), 2.08-2.18 (m, 2H, Hf, Hf’), 2.58-2.69 (m, 6H), 3.20 (dd, 2H, J 8.0 J 8.8 Hz), 3.24-3.39 (m, 8H), 3.39-3.45 (m, 2H), 3.45-3.60 (m, 10H), 3.60-3.69 (m, 3H), 3.74-3.88 (m, 6H), 3.88-4.01 (m, 6H), 4.33-4.43 (m, 3H), 4.45-4.49 (m, 2H), 5.49-5.58 (m, 1H, Hd), 5.83-5.93 (m, 1H, He)

HRMS calculated for C₄₂H₇₈NO₂₂: 948.5010, found: 948.5015 [M]+.

(2S, 3R, 4E)-3-Hydroxy-2-(octadecanamido)octadec-4-enyl (β-D-glucopyranosyl)-(1→4)-β-D-glucopyranoside (3.07)

Cellobiosyl sphingosine (10.3 mg, 16.5 µmol, 3.03) and 3.20 (8.36 mg, 21.9 µmol, 1.3 eq.) were dissolved in 0.8 mL of 1:1 DCM:DMF. Triethylamine (3.80 µL, 27.2 µmol, 1.6 eq.) was added, and the reaction mixture was stirred at RT overnight. Toluene was added to form a DMF-toluene azeotrope, and the reaction mixture was concentrated in vacuo. The crude solid was purified by flash column chromatography using an eluent of 30:10:1 CHCl₃:MeOH:H₂O, yielding 3.07 (15 mg, 16.5 µmol, 100%).

Rf (30:10:1 CHCl₃:MeOH:H₂O): 0.51 ¹H NMR (600 MHz, CDCl₃): δ ppm 0.12 (dt, 6H, J 7.1 J 7.2 Hz, 2 CH₃), 0.45-0.62 (m, 32H), 0.73-0.80 (m, 2H), 1.19-1.26 (m,
2H), 1.31-1.36 (m, 2H), 1.79-1.82 (m, 3H), 2.39-2.46 (m, 2H), 2.46-2.50 (m, 1H), 2.50-2.59 (m, 22H), 2.67-2.74 (m, 3H), 2.78-2.83 (m, 1H), 3.00-3.05 (m, 3H), 3.11-3.16 (m, 1H), 3.20-3.25 (m, 1H), 3.39 (dd, 1H, J 3.3 J 9.8 Hz), 3.43 (dd, 1H, J 7.8 Hz), 3.58 (dd, 1H, J 7.9 Hz), 3.73 (dt, 1H, J 5.3 Hz), 3.80 (dt, 1H, J 6.4 Hz), 3.98 (dd, 1H, J 2.2 J 5.5 Hz), 4.07 (s, 1H), 4.13 (d, 1H, J 4.0 Hz), 4.41 (d, 1H, J 3.2 Hz), 4.44 (d, 1H, J 4.4 Hz), 4.62-4.68 (m, 1H, Hd), 4.80-4.87 (m, 1H, He), 6.67 (d, 1H, J 9.1 Hz) 13C NMR (150 MHz, CDCl3): δ ppm 12.62 (CH3), 20.93 (CH3), 24.22, 27.57, 27.59, 27.68, 27.84, 27.87, 27.92, 27.95, 27.98, 30.15, 30.66, 34.64, 51.61, 59.37, 59.76, 67.88, 68.53, 69.60, 71.79, 71.88, 73.38, 73.55, 75.15, 75.42, 78.86, 102.06, 102.54, 129.68 (Cd), 130.67 (Ce), 171.12 (C=O) HRMS calculated for C48H91NO13: 889.6490, found: 890.6569 [M+H]+.

Perfluorophenyl tetradecanoate (3.11)

Myristic acid (114 mg, 0.500 mmol, 3.09) was dissolved in 5 mL of DCM. Pyridine (46.5 µL, 0.577 mmol, 1.15 eq.) and TFA-OPFP (103.1 µL, 0.600 mmol, 1.2 eq., 3.14) were added, and the reaction was stirred at RT overnight. The reaction mixture was diluted to 70 mL with DCM. The organic layer was washed with 2x100 mL saturated NaHCO3, dried over sodium sulfate, filtered, and concentrated in vacuo. The crude solid was purified by flash column chromatography using PE as an eluent, yielding 3.11 as a white solid (150 mg, 0.380 mmol, 76%). Rf (PE): 0.26 1H NMR (400 MHz, CDCl3): δ ppm 0.88 (t, 3H, J 6.8 Hz, CH3), 1.20-1.46 (m, 20H, CH2), 1.77 (qn, 2H, J 7.4 Hz, CH2CH2COOPFP), 2.66 (t, 2H, J 7.4 Hz, CH2COOPFP)
\[^{19}\text{F NMR}\ (282 \text{ MHz, CDCl}_3) \delta \text{ ppm} -162.89 \ (\text{dd, } J 21.4 \ J 21.7 \text{ Hz}), -158.68 \ (\text{dd, } J 21.4 \ J 21.7 \text{ Hz}), -153.24 \ (\text{d, } J 17.5 \text{ Hz})\]^\text{13C NMR} (100 \text{ MHz, CDCl}_3): \delta \text{ ppm} 14.26 \ (\text{CH}_3), 22.84 \ (\text{CH}_2\text{CH}_3), 24.93 \ (\text{CH}_2\text{CH}_2\text{COO-PFP}), 29.01, 29.28, 29.51, 29.54, 29.70, 29.77, 29.79, 29.82, 32.07 \ (\text{CH}_2\text{CH}_2\text{CH}_3), 33.51 \ (\text{CH}_2\text{COO-PFP}), 136.74, 139.29, 140.07, 142.65, 169.76 \ (\text{COO-PFP}).

Perfluorophenyl palmitate (3.12)

Palmitic acid (128 mg, 0.500 mmol, 3.10) was dissolved in 5 mL of DCM. Pyridine (46.5 \text{ µL}, 0.577 mmol, 1.15 eq.) and TFA-OPFP (103.1 \text{ µL}, 0.600 mmol, 1.2 eq., 3.14) were added, and the reaction was stirred at RT overnight. The reaction was diluted to 70 mL with DCM. The organic layer was washed with 2x100 mL saturated NaHCO\textsubscript{3}, dried over sodium sulfate, filtered, and concentrated \textit{in vacuo}. The crude solid was purified by flash column chromatography using PE as an eluent, yielding 3.12 as a white solid (167 mg, 0.395 mmol, 79%). \textbf{RF} (PE): 0.42

\[^{1}\text{H NMR}\ (400 \text{ MHz, CDCl}_3): \delta \text{ ppm} 0.88 \ (t, 3\text{H, } J 6.8 \text{ Hz, CH}_3), 1.20-1.46 \ (m, 24\text{H, CH}_2), 1.77 \ (\text{qn, 2H, } J 7.4 \text{ Hz, CH}_2\text{CH}_2\text{COOPFP}), 2.66 \ (t, 2\text{H, } J 7.4 \text{ Hz, CH}_2\text{COOPFP})\]^\text{19F NMR} (282 MHz, CDCl\textsubscript{3}) \delta \text{ ppm} -162.89 \ (\text{dd, } J 21.4 \ J 21.7 \text{ Hz}), -158.68 \ (\text{dd, } J 21.4 \ J 21.7 \text{ Hz}), -153.24 \ (\text{d, } J 17.8 \text{ Hz})\]^\text{13C NMR} (100 \text{ MHz, CDCl}_3): \delta \text{ ppm} 14.26 \ (\text{CH}_3), 22.85 \ (\text{CH}_2\text{CH}_3), 24.93 \ (\text{CH}_2\text{CH}_2\text{COO-PFP}), 29.01, 29.28, 29.51, 29.54, 29.70, 29.77, 29.81, 29.83, 32.08 \ (\text{CH}_2\text{CH}_2\text{CH}_3), 33.51 \ (\text{CH}_2\text{COO-PFP}), 136.76, 139.29, 140.10, 142.53, 169.76 \ (\text{COO-PFP}).
Perfluorophenyl stearate (3.13)

Stearic acid (142 mg, 0.500 mmol, 3.06) was dissolved in 5 mL of DCM. Pyridine (46.5 µL, 0.577 mmol, 1.15 eq.) and TFA-OPFP (103.1 µL, 0.600 mmol, 1.2 eq., 3.14) were added, and the reaction was stirred at RT overnight. The reaction mixture was diluted to 70 mL with DCM. The organic layer was washed with 2x100 mL saturated NaHCO₃, dried over sodium sulfate, filtered, and concentrated in vacuo. The crude solid was purified by flash column chromatography using PE as an eluent, yielding 3.13 as a white solid (162 mg, 0.360 mmol, 72%). Rf (PE): 0.38. ¹H NMR (400 MHz, CDCl₃): δ ppm 0.88 (t, 3H, J 6.8 Hz, CH₃), 1.20-1.46 (m, 28H, CH₂), 1.77 (qn, 2H, J 7.4 Hz, CH₂CH₂COOPFP), 2.66 (t, 2H, J 7.4 Hz, CH₂COOPFP) ¹⁹F NMR (282 MHz, CDCl₃) δ ppm -162.88 (dd, J 21.4 J 21.7 Hz), -158.68 (dd, J 21.4 J 21.7 Hz), -153.24 (d, J 17.8 Hz) ¹³C NMR (100 MHz, CDCl₃): δ ppm 14.26 (CH₃), 22.85 (CH₂CH₃), 24.92 (CH₂CH₂COO-PFP), 29.01, 29.28, 29.52, 29.54, 29.70, 29.77, 29.82, 29.85, 32.08 (CH₂CH₂CH₃), 33.51 (CH₂COO-PFP), 136.76, 139.26, 140.10, 142.56, 169.76 (COO-PFP).

Perfluorophenyl 2-hydroxybenzoate (3.16)
Salicylic acid (69.4 mg, 0.502 mmol, 3.15) was dissolved in 5 mL of DCM. Pyridine (46.5 µL, 0.577 mmol, 1.15 eq.) and TFA-OPFP (103.1 µL, 0.600 mmol, 1.2 eq., 3.14) were added, and the reaction mixture was stirred at RT overnight. The reaction was diluted to 70 mL with DCM. The organic layer was washed with 2x100 mL saturated NaHCO₃, dried over sodium sulfate, filtered, and concentrated in vacuo. The crude solid was purified by flash column chromatography using PE as an eluent, yielding 3.16 as a white solid (112 mg, 0.389 mmol, 73%). ¹H and ¹³C NMR data agreed with that reported previously.¹²⁸ Rf (PE): 0.29 ¹H NMR (400 MHz, CDCl₃): δ ppm 7.00 (dt, 1H, J 0.8 J 7.6 Hz), 7.06 (dd, 1H, J 0.8 J 8.0 Hz), 7.59 (dt, 1H, J 1.6 J 7.6 Hz), 8.05 (dd, 1H, J 1.6 J 8.0 Hz), 9.82 (s, 1H) ¹⁹F NMR (282 MHz, CDCl₃) δ ppm -162.16 (dd, J 19.7 Hz), -157.30 (dd, J 21.7 Hz), -152.54 (d, J 17.5 Hz) ¹³C NMR (100 MHz, CDCl₃): δ ppm 109.97, 118.30, 120.15, 130.80, 137.87, 162.62, 166.26.

2,5-Dioxopyrrolidin-1-yl tetradecanoate (3.18)

Myristic acid (113 mg, 0.496 mmol, 3.09) was dissolved in 4.5 mL of DCM. NHS (230 mg, 2.00 mmol, 4 eq., 3.17) and EDC hydrochloride (383 mg, 2.00 mmol, 4 eq.) were added, and the reaction was stirred at RT overnight. The reaction mixture was diluted to 30 mL DCM and the organic layer washed with 2x25 mL deionized H₂O, dried over sodium sulfate, filtered, and concentrated in vacuo. The crude solid was purified by flash column chromatography using 4:1 Hexanes:EtOAc as an eluent, yielding 3.18 as a white solid (103 mg, 0.317 mmol, 64%). Rf (4:1 Hexanes:EtOAc): 0.18 ¹H NMR (400 MHz, CDCl₃): δ ppm 0.88 (t, 3H, J 6.8 Hz, CH₃), 1.20-
1.44 (m, 20H, CH2), 1.74 (qn, 2H, J 7.6 Hz, CH2CH2COONHS), 2.60 (t, 2H, J 7.4 Hz, 
CH2COOPFP), 2.83 (s, 4H, COCH2CH2CO) 13C NMR (100 MHz, CDCl3): δ ppm 14.26 (CH3), 
22.84 (CH2CH3), 24.72 (CH2CH2COO-NHS), 25.74 (COCH2CH2CO), 28.94, 29.23, 29.49, 
29.69, 29.76, 29.78, 29.80, 31.10 (CH2CH2CH3), 32.06 (CH2COO-NHS), 168.85 (COO-NHS), 
169.32 (NCOC2H4CON) HRMS calculated for C18H31NO4: 325.2253, found: 348.2151 
[M+Na]+.

2,5-Dioxopyrrolidin-1-yl palmitate (3.19)

Palmitic acid (128 mg, 0.499 mmol, 3.19) was dissolved in 4.5 mL of DCM. NHS (231 mg, 2.00 
mmol, 4 eq., 3.17) and EDC hydrochloride (383 mg, 2.00 mmol, 4 eq.) were added, and the 
reaction was stirred at RT overnight. The reaction was diluted to 30 mL DCM. The organic layer 
was washed with 2x25 mL deionized H2O, dried over sodium sulfate, filtered, and concentrated 
in vacuo. The crude solid was purified by flash column chromatography using 4:1 
Hexanes:EtOAc as an eluent, yielding 3.19 as a white solid (137 mg, 0.388 mmol, 78%). Rf (4:1 
Hexanes:EtOAc): 0.16 1H NMR (400 MHz, CDCl3): δ ppm 0.88 (t, 3H, J 6.8 Hz, CH3), 1.20- 
1.44 (m, 24H, CH2), 1.74 (qn, 2H, J 7.6 Hz, CH2CH2COONHS), 2.60 (t, 2H, J 7.4 Hz, 
CH2COOPFP), 2.83 (s, 4H, COCH2CH2CO) 13C NMR (100 MHz, CDCl3): δ ppm 14.26 (CH3), 
22.84 (CH2CH3), 24.72 (CH2CH2COO-NHS), 25.74 (COCH2CH2CO), 28.94, 29.23, 29.50, 
29.70, 29.77, 29.80, 29.83, 31.10 (CH2CH2CH3), 32.07 (CH2COO-NHS), 168.85 (COO-NHS),
169.32 (NCOC$_2$H$_4$CON) **HRMS** calculated for C$_{20}$H$_{35}$NO$_4$: 353.2566, found: 376.2464 [M+Na]$^+$. 

2,5-Dioxopyrrolidin-1-yl stearate (3.20) 

Stearic acid (142 mg, 0.500 mmol, 3.06) was dissolved in 4.5 mL of DCM. NHS (231 mg, 2.00 mmol, 4 eq., 3.17) and EDC hydrochloride (384 mg, 2.00 mmol, 4 eq.) were added, and the reaction was stirred at RT overnight. The reaction was diluted to 30 mL DCM. The organic layer was washed with 2x25 mL deionized H$_2$O, dried over sodium sulfate, filtered, and concentrated *in vacuo*. The crude solid was purified by flash column chromatography using 4:1 Hexanes:EtOAc as an eluent, yielding 3.20 as a white solid (115 mg, 0.302 mmol, 60%). **Rf** (4:1 Hexanes:EtOAc): 0.17 ¹H **NMR** (400 MHz, CDCl$_3$): δ ppm 0.88 (t, 3H, J 6.8 Hz, CH$_3$), 1.20-1.44 (m, 28H, CH$_2$), 1.74 (qn, 2H, J 7.6 Hz, CH$_2$CH$_2$COONHS), 2.60 (t, 2H, J 7.4 Hz, CH$_2$COOPFP), 2.83 (s, 4H, COCH$_2$CH$_2$CO) ¹³C **NMR** (100 MHz, CDCl$_3$): δ ppm 14.26 (CH$_3$), 22.84 (CH$_2$CH$_3$), 24.72 (CH$_2$CH$_2$COO-NHS), 25.74 (COCH$_2$CH$_2$CO), 28.95, 29.24, 29.50, 29.70, 29.77, 29.81, 29.84, 31.10 (CH$_2$CH$_2$CH$_3$), 32.07 (CH$_2$COO-NHS), 168.84 (COO-NHS), 169.31 (NCOC$_2$H$_4$CON) **HRMS** calculated for C$_{22}$H$_{39}$NO$_4$: 381.2879, found: 404.2777 [M+Na]$^+$. 

162
(2S, 3R, 4E)-3-Hydroxy-2-(tetradecanamido)octadec-4-enyl (β-D-glucopyranosyl)-(1→4)-β-D-glucopyranoside (3.21)

Cellobiosyl sphingosine (10.2 mg, 16.3 µmol, 3.03) and 3.18 (7.35 mg, 22.6 µmol, 1.4 eq.) were dissolved in 0.8 mL of 1:1 DCM:DMF. Triethylamine (3.80 µL, 27.2 µmol, 1.7 eq.) was added, and the reaction was stirred at RT overnight. Toluene was added to form a DMF-toluene azeotrope, and the reaction mixture was concentrated in vacuo. The crude solid was purified by flash column chromatography using an eluent of 30:10:1 CHCl₃:MeOH:H₂O, yielding 3.21 (14 mg, 16.3 µmol, 100%). Rf (30:10:1 CHCl₃:MeOH:H₂O): 0.55 HRMS calculated for C₄₄H₈₃NO₁₃: 833.5864, found: 834.5943 [M+H]⁺.

(2S, 3R, 4E)-3-Hydroxy-2-(hexadecanamido)octadec-4-enyl (β-D-glucopyranosyl)-(1→4)-β-D-glucopyranoside (3.22)

Cellobiosyl sphingosine (9.96 mg, 16.0 µmol, 3.03) and 3.19 (7.65 mg, 21.6 µmol, 1.4 eq.) were dissolved in 0.8 mL of 1:1 DCM:DMF. Triethylamine (3.80 µL, 27.2 µmol, 1.7 eq.) was added, and the reaction was stirred at RT overnight. Toluene was added to form a DMF-toluene azeotrope, and the reaction was concentrated in vacuo. The crude solid was purified by flash
column chromatography using an eluent of 30:10:1 CHCl$_3$:MeOH:H$_2$O, yielding 3.22 (11 mg, 12.8 µmol, 80%). Rf (30:10:1 CHCl$_3$:MeOH:H$_2$O): 0.51 HRMS calculated for C$_{46}$H$_{87}$NO$_{13}$: 861.6177, found: 884.6075 [M+Na]$^+$. O-[(5-Acetamido-3,5-dideoxy-$\alpha$-D-glycero-D-galacto-2-nonulopyranosyl)onic acid]-(2→3)-O-(\beta-D-galactopyranosyl)-(1→4)-\beta-D-glucopyranosyl-(1→1)-(2S, 3R, 4E)-2-aminoctadec-4-ene-1,3-diol (3.26)

![Chemical Structure](image_url)

Compound 3.26 was prepared according to the General Procedure for Sialoside Synthesis using Cst-I Enzyme. Lactosyl sphingosine (20.0 mg, 32.1 µmol, 3.02 eq., 2.23) and CMP-Neu5Ac (22.1 mg, 36.0 µmol, 1.1 eq., 2.23) were mixed in a glass vial and dissolved in 50 mM HEPES (pH 7.5) and 10 mM MnCl$_2$. Once dissolved, Cst-I (0.2 mg/mL) and alkaline phosphatase (5 µL, 10 KU) were added, and the reaction was incubated at RT. The reaction was monitored by TLC (4:5:1 CHCl$_3$:MeOH:0.2% CaCl$_2$), and was complete after 2 hours. The reaction was purified on a Waters Sep-Pak$^\text{®}$ tC18 cartridge (2 g). The product eluted at 50% acetonitrile, and the fractions were lyophilized, yielding 3.26 as a white solid (23.5 mg, 25.7 µmol, 80%). $^1$H and $^{13}$C NMR data agreed with that reported previously.$^{86}$ Rf (4:5:1 CHCl$_3$:MeOH:0.2% CaCl$_2$) 0.31 $^1$H NMR (400 MHz, MeOD): δ ppm 0.93 (t, 3H, J 7.0 Hz, CH$_3$), 1.25-1.40 (m, 20H), 1.40-1.49 (m, 2H, Hg, Hg$^\delta$), 1.68 (dd, 1H, $J_{H4''-H3ax'}$ 11.9 Hz, H3ax$'$), 2.01 (s, 3H, NHAc), 2.06-2.16 (m, 2H, Hf, Hf$^\delta$), 2.87 (dd, 1H, $J_{H4''-H3eq'}$ 4.4 $J_{H3ax'-H3eq'}$ 12.1 Hz, H3eq$'$), 3.44 (dd, 1H, $J_{H3-H2}$ 9.2 Hz, H2), 3.48-3.75 (m, 10H, H9b, H2’, H3, H4, H5, H5’, H6’, H7, H8, Hb), 3.74-3.79 (m, 3H, H6’, H6,
H5”), 3.81-3.87 (m, 2H, H9a), 3.87-3.94 (m, 3H, H6, Ha’, H4’), 3.94-4.00 (m, 2H, H6, Ha), 4.05 (dd, 1H, \(J_{H4':H3'}\) 3.1 \(J_{H2':H3'}\) 9.7 Hz, H3’), 4.30 (dd, 1H, \(J_{5.3}\) Hz, Hc), 4.37 (d, 1H, \(J_{7.8}\) Hz, H1), 4.40 (d, 1H, \(J_{7.8}\) Hz, H1’), 5.46-5.55 (m, 1H, Hd), 5.79-5.91 (m, 1H, He) \(^{13}\)C NMR (100 MHz, MeOD): \(\delta\) ppm 14.69 (\(C\)H\(_3\)), 22.96, 23.57 (NHAc), 30.02, 30.19, 30.26, 30.45, 30.56, 30.60, 32.85, 33.19, 53.94, 58.26, 59.20, 62.37, 64.72, 67.10, 68.83, 70.63, 72.68, 74.30, 75.91, 76.42, 77.05, 100.75, 103.82, 105.27, 128.64(Cd), 135.83 (Ce) HRMS calculated for \(C_{41}H_{74}N_2O_{20}\): 914.4835, found: 915.4913 [M+H]+.

\(O\)-(β-\(d\)-Galactopyranosyl)-(1\(→4\))-β-\(d\)-Glucopyranosyl-(1\(→4\))-β-\(d\)-Glucopyranosyl-(1\(→1\))(2S, 3R, 4E)-2-aminooctadec-4-ene-1,3-diol (3.27)

![Chemical structure](image.png)

Compound 3.27 was prepared according to the **General Procedure for Synthesis of Oligosaccharides using Abg2F6 Glycosynthase**. Cellobiosyl sphingosine (7.57 mg, 10.3 mmol, 3.03) and \(\alpha\)-galactosyl fluoride (5.54 mg, 30.4 mmol, 3 eq., 2.08) was dissolved in 100 mM MES buffer (pH 7.0), and Abg2F6 (1.0 mg/mL) was added. The reaction was incubated at RT, and monitored by TLC (5:4:0.8 CHCl\(_3\):MeOH:0.2% CaCl\(_2\)). The reaction was complete after five days, lyophilized, and purified using a Waters Sep-Pak\textsuperscript{®} tC18 cartridge (5 g). 3.27 co-eluted with leftover 3.03, and was used in reactions without further purification. HRMS calculated for \(C_{36}H_{68}NO_{17}\): 786.4482, found: 786.4487 [M]+.
O-[(5-Acetamido-3,5-dideoxy-α-D-glycero-D-galacto-2-nonulopyranosyl)onic acid]-(2→3)-O-(β-D-galactopyranosyl)-(1→4)-β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl-(1→1)-(2S, 3R, 4E)-2-aminoctadec-4-ene-1,3-diol (3.28)

Compound 3.28 was prepared according to the **General Procedure for Sialoside Synthesis** using Cst-I Enzyme. Crude 3.27 (theoretical maximum: 20.7 mg, 26.3 µmol) and CMP-Neu5Ac (20.4 mg, 33.3 µmol, 1.3 eq., 2.23) were mixed in a glass vial and dissolved in 50 mM HEPES (pH 7.5) and 10 mM MnCl₂. Once dissolved, Cst-I (0.2 mg/mL) and alkaline phosphatase (5 µL, 10 KU) were added, and the reaction was incubated at RT. The reaction was monitored by TLC (4:5:1 CHCl₃:MeOH:0.2% CaCl₂), and was complete after 2 hours. The reaction was purified on a Waters Sep-Pak® tC18 cartridge (2 g). The product eluted at 50% acetonitrile, and the fractions were lyophilized yielding 3.28 as a white solid (8.0 mg, 7.44 µmol, 29%) Rf (4:5:1 CHCl₃:MeOH:0.2% CaCl₂) 0.18

**1H NMR** (600 MHz, MeOD): δ ppm 0.90 (t, 3H, J 6.9 Hz, CH₃), 1.25-1.37 (m, 20H), 1.40-1.46 (m, 2H, Hg, Hg’), 1.76-1.82 (m, 1H), 2.01 (s, 3H, NHAc), 2.08-2.14 (m, 3H, Hf, Hf’), 2.83 (dd, 1H, J 4.0 J 13.0 Hz), 3.38-3.42 (m, 1H), 3.44-3.51 (m, 3H), 3.51-3.60 (m, 7H), 3.60-3.68 (m, 3H), 3.71-3.79 (m, 3H), 3.81-3.99 (m, 11H), 4.05 (dd, 1H, J 3.1 J 9.6 Hz, H3’’), 4.31 (dd, 1H, J 5.6 Hz, Hc), 4.36 (d, 1H, J 7.9 Hz, H1), 4.42 (d, 1H, J 7.9 Hz, H1’), 4.43 (d, 1H, J 7.8 Hz, H1’’), 5.46-5.51 (m, 1H, Hd), 5.84-5.90 (m, 1H, He)

**13C NMR** (150 MHz, MeOD): δ ppm 14.45 (CH₃), 22.59 (NHAc), 23.74, 30.18, 30.40, 30.49, 30.66, 30.77, 30.81, 33.09, 33.38, 41.75, 53.84, 56.70, 61.51, 61.59, 62.62, 64.71, 67.06, 69.07, 69.18, 70.08, 70.17, 70.90 (Hc), 72.83, 74.48, 74.59, 75.14, 76.21, 76.25, 76.62, 76.68, 77.02, 77.76, 80.35,
80.40, 103.68 (C1), 104.43 (C1’’), 105.04 (C1’), 128.25 (Cd), 136.79 (Ce), 175.43 (CO₂) HRMS calculated for C₄₇H₈₄N₂O₂₅: 1076.5363, found: 1077.5441 [M+H]+.

O-(6-O-Phospho-β-d-glucopyranosyl)-(1→4)-β-d-glucopyranosyl-(1→1)-(2S, 3R, 4E)-2-aminooctadec-4-ene-1,3-diol (3.29)

Compound 3.29 was prepared according to the **General Procedure for Phosphorylation using BglK Enzyme**. Cellobiosyl sphingosine (19.4 mg, 31.1 µmol, 3.03) was dissolved in 25 mM HEPES (pH 7.5) containing 2 mM MgSO₄. Once dissolved, ATP (37.3 µmol, 1.2 eq., adjusted to pH 7.5 with 3 M NH₄OH) was added to the reaction mixture. BglK (0.25 mg/mL) was added, and the reaction mixture was incubated at room temperature. The reaction was complete overnight according to TLC (4:5:1 CHCl₃:MeOH:0.2% CaCl₂) and was purified on a Waters Sep-Pak® tC18 cartridge (2 g). The product eluted at 50% acetonitrile and the fractions were lyophilized, yielding 3.29 as a white solid (11.8 mg, 16.8 µmol, 54%). **Rf** (4:5:2 CHCl₃:MeOH:0.2% CaCl₂): 0.20 **¹H NMR** (600 MHz, MeOD): δ ppm 1.63 (t, 3H, J 6.9 Hz, CH₃), 1.96-2.09 (m, 21H), 2.09-2.18 (m, 2H, Hg, Hg’), 2.77-2.85 (m, 2H, Hg, Hg’), 3.85 (dd, 1H, J 8.2 J 8.7 Hz), 3.93-4.10 (m, 8H), 4.13-4.26 (m, 4H), 4.40-4.46 (dd, 1H, J 4.9 J 11.9 Hz), 4.54-4.59 (m, 1H), 4.59-4.71 (m, 4H), 4.81 (dd, 2H, J 9.7 Hz), 5.07 (d, 1H, J 7.9 Hz), 5.10 (d, 1H, J 7.9 Hz), 6.20-6.27 (m, 1H, Hd), 6.51-6.59 (m, 1H, He) **³¹P NMR** (161 MHz, MeOD) δ ppm 0.62 (s) **¹³C NMR** (150 MHz, MeOD): δ ppm 14.99 (CH₃), 23.67, 30.15, 30.31, 30.34, 30.55, 30.63, 30.67, 32.92, 33.26 (Cf), 56.58, 61.93, 65.57, 65.60, 67.38, 70.48, 71.33, 74.35, 74.85,
75.57, 76.49, 77.02, 77.07, 77.50, 81.93, 104.00, 104.54, 129.25 (Cd), 135.26 (Ce) HRMS calculated for C\textsubscript{30}H\textsubscript{57}NO\textsubscript{15}P: 702.3471, found: 704.3622 [M+2H]\textsuperscript{+}.

\[ O-(6-O-\text{Phospho-}\beta-\text{d-glucopyranosyl}-(1\rightarrow4)-\beta-\text{d-glucopyranosyl}-(1\rightarrow4)-\beta-\text{d-glucopyranosyl}-(1\rightarrow1)-(2S, 3R, 4E)-2\text{-amino}o\text{ctadec-4-ene-1,3-diol (3.30)} \]

\[
\begin{array}{c}
\text{OPO}_3^{2-} \\
\text{NH}_3 \\
\text{C}_{13}H_{27}
\end{array}
\]

Compound 3.30 was prepared according to the General Procedure for Phosphorylation using BglK Enzyme. Cellotriosyl sphingosine (20.1 mg, 25.6 \text{\mu}mol, 3.04) was dissolved in 25 mM HEPES (pH 7.5) containing 2 mM MgSO\textsubscript{4}. Once dissolved, ATP (30.7 \text{\mu}mol, 1.2 eq., adjusted to pH 7.5 with 3 M NH\textsubscript{4}OH) was added to the reaction mixture. BglK (0.25 mg/mL) was added, and the reaction was incubated at room temperature. The reaction was complete overnight according to TLC (4:5:1 CHCl\textsubscript{3}:MeOH:0.2% CaCl\textsubscript{2}). The reaction was purified on a Waters Sep-Pak\textsuperscript{®} tC18 cartridge (2 g). The product eluted at 50% acetonitrile and the fractions were lyophilized, yielding 3.30 as a white solid (16.6 mg, 19.2 \text{\mu}mol, 75\%). R\text{f} (4:5:2 CHCl\textsubscript{3}:MeOH:0.2% CaCl\textsubscript{2}): 0.13 \textsuperscript{1}H NMR (600 MHz, MeOD): \( \delta \) ppm 0.93 (t, 3H, J 7.1 Hz, CH\textsubscript{3}), 1.27-1.40 (m, 22H), 1.41-1.49 (m, 2H, Hg, Hg'), 2.10-2.16 (m, 2H, Hf, Hf'), 2.63-2.65 (m, 2H) 3.22 (dd, 1H, J 8.2 Hz), 3.37 (dd, 2H, J 4.4 J 4.0 Hz), 3.40-3.44 (m, 1H), 3.47-3.52 (m, 2H), 3.53-3.58 (m, 5H), 3.80 (bd, 1H, J 9.4 Hz), 3.86 (dd, 1H, J4.2 J 11.9 Hz), 3.90-4.01 (m, 4H), 4.01-4.08 (m, 1H), 4.13-4.19 (m, 1H), 4.36 (dd, 1H, J 5.3 J 4.9 Hz), 4.39 (d, 1H, J 7.9 Hz), 4.42 (d, 1H, J 7.8 Hz), 4.48 (d, 1H, J 8.0 Hz), 5.50-5.56 (m, 1H, Hd), 5.85-5.93 (m, 1H, He) \textsuperscript{31}P NMR (161 MHz, MeOD) \( \delta \) ppm 0.60 (s) \textsuperscript{13}C NMR (150 MHz, MeOD): \( \delta \) ppm 14.67 (CH\textsubscript{3}), 168
23.70, 30.17, 30.37, 30.41, 30.60, 30.72, 30.75, 33.00, 33.33, 56.47, 61.63, 61.81, 65.54, 67.10, 70.66, 71.09, 74.44, 74.51, 74.89, 76.00, 76.04, 76.54, 76.57, 77.05, 77.10, 77.46, 80.57, 81.11, 103.88, 104.22, 104.58, 128.69 (Cd), 136.17 (Ce) **HRMS** calculated for C$_{36}$H$_{67}$NO$_{20}$P: 864.4000, found: 888.3970 [M+H+Na]$^+$. 

\[O-(6-O-\text{Phospho}-\beta-d-\text{glucopyranosyl})-(1\rightarrow4)-\beta-d-\text{glucopyranosyl}-(1\rightarrow4)-\beta-d-\text{glucopyranosyl}-(1\rightarrow4)-\beta-d-\text{glucopyranosyl}-(1\rightarrow1)-(2S, 3R, 4E)-2\text{-aminoocadec-4-ene-1,3-diol (3.31)}\]

Compound 3.31 was prepared according to the General Procedure for Phosphorylation using BglK Enzyme. Cellotetraosyl sphingosine (21.9 mg, 23.1 µmol, 3.05) was dissolved in 25 mM HEPES (pH 7.5) containing 2 mM MgSO$_4$. Once dissolved, ATP (27.7 µmol, 1.2 eq., adjusted to pH 7.5 with 3 M NH$_4$OH) was added to the reaction mixture. BglK (0.25 mg/mL) was added, and the reaction mixture was incubated at room temperature. The reaction was complete overnight according to TLC (4:5:1 CHCl$_3$:MeOH:0.2% CaCl$_2$) and was purified on a Waters Sep-Pak$^\circledR$ tC18 cartridge (2 g). The product eluted at 50% acetonitrile and the fractions were lyophilized, yielding 3.31 as a white solid (16.6 mg, 16.2 µmol, 70%). **Rf** (4:5:2 CHCl$_3$:MeOH:0.2% CaCl$_2$): 0.14 $^1$H NMR (600 MHz, MeOD): $\delta$ ppm 0.93 (t, 3H, J 7.0 Hz, CH$_3$), 1.26-1.40 (m, 22H), 1.41-1.48 (m, 2H, Hg, Hg$'$), 2.09-2.16 (m, 2H, Hf, Hf$'$), 2.62-2.67 (m, 2H), 3.22 (dd, 1H, J 8.0 Hz), 3.27-3.35 (m, 3H), 3.35-3.39 (m, 2H), 3.40-3.45 (m, 2H), 3.45-3.51 (m, 2H), 3.52-3.62 (m, 8H), 3.77-3.90 (m, 3H), 3.90-4.06 (m, 6H), 4.13-4.21 (m, 1H), 4.34-4.44 (m, 3H), 4.44-4.50 (m, 2H), 5.49-5.56 (m, 1H, Hd), 5.85-5.93 (m, 1H, He) $^{31}$P NMR (161 MHz,
MeOD) δ ppm 1.72 (s) $^{13}$C NMR (150 MHz, MeOD): δ ppm 14.63 (CH$_3$), 23.71, 30.19, 30.37, 30.42, 30.61, 30.71, 30.72, 30.75, 33.01, 33.33, 56.56, 61.60, 61.65, 61.78, 65.51, 67.13, 70.60, 71.19, 74.39, 74.43, 74.57, 74.89, 76.02, 76.11, 76.48, 76.54, 77.03, 77.08, 77.47, 80.61, 80.93, 81.42, 103.91, 104.26, 104.48, 104.78, 128.71 (Cd), 136.03 (Ce) HRMS calculated for C$_{42}$H$_{77}$NO$_{25}$P: 1026.4528, found: 1028.4679 [M+2H]$^+$.  

$O$-($6$-$O$-Phospho-$\beta$-$d$-$glucopyranosyl$)-(1$→$4$)$-$\beta$-$d$-$glucopyranosyl$-(1$→$4$)$-$\beta$-$d$-$glucopyranosyl$-(1$→$4$)$-$\beta$-$d$-$glucopyranosyl$-(1$→$4$)$-$\beta$-$d$-$glucopyranosyl$-(1$→$1$)$-(2$S$, 3$R$, 4$E$)-2$-amino$octadec$-4$-ene$-1,3$-diol (3.32)

Crude phosphorylated cellotriosyl fluoride (theoretical maximum: 26.0 mg, 44.5 µmol, 1.2 eq., 2.28) and cellotriosyl sphingosine (29.1 mg, 37.1 µmol, 3.04) were dissolved in 100 mM MES (pH 7.0) to a final acceptor concentration of 20 mM. CelB glycosynthase (1.0 mg/mL) was added, and the reaction was incubated at 30°C and monitored by TLC (4:5:2 CHCl$_3$:MeOH:10% AcOH). After 5 days, the reaction mixture was directly added to a Waters Sep-Pak® tC18 cartridge (2 g). The product eluted at 50% acetonitrile along with unreacted 3.04. The fractions were lyophilized, yielding crude 3.32 as a white solid. RF (4:5:2 CHCl$_3$:MeOH:10% AcOH): 0.30 HRMS calculated for C$_{54}$H$_{97}$NO$_{35}$P: 1350.5584, found: 1352.5735 [M+2H]$^+$.  

170
(4-Keto-β-d-glucopyranosyl)-(1→4)-β-d-glucopyranosyl-(1→4)-α-d-glucopyranosyl fluoride (3.35)

Compound 3.35 was prepared according to the General Procedure for Synthesis of Oligosaccharides using Abg2F6 Glycosynthase. 4-keto-glucosyl fluoride (3.33, 84.7 mg, 470 µmol, 7.8 eq.) and α-cellobiosyl fluoride (2.12, 20.7 mg, 60.1 µmol) were dissolved in 100 mM sodium phosphate (pH 7.0), and Abg2F6 (1.0 mg/mL) was added. The reaction was monitored by TLC (7:2:1 EtOAc:MeOH:H₂O), and was lyophilized after 18 hours. The reaction mixture was purified by flash column chromatography using an eluent of 7:2:1 EtOAc:MeOH:H₂O, yielding a mixture of 3.35 and the hydrate. Rf (7:2:1 EtOAc:MeOH:H₂O): 0.17 19F NMR (282 MHz, MeOD) δ ppm -151.78 (dd, J 26.4 J 53.9 Hz) ESI MS calculated for C₁₈H₂₉F₂O₁₅: 504.4, found: 527.4 [M+Na]+.

(4-Amino-β-d-glucopyranosyl)-(1→4)-β-d-glucopyranosyl-(1→4)-α-d-glucopyranosyl fluoride (3.37)

Compound 3.37 was prepared according to the General Procedure for Synthesis of Oligosaccharides using Abg2F6 Glycosynthase. 4-Amino-glucosyl fluoride (3.36, 29.7 mg, 163 µmol, 2.7 eq.) and α-cellobiosyl fluoride (2.12, 20.6 mg, 59.9 µmol) were dissolved in 100
mM sodium phosphate (pH 7.0), and Abg2F6 (1.0 mg/mL) was added. The reaction was monitored by TLC (7:2:1 EtOAc:MeOH:H$_2$O), and was lyophilized after 18 hours. A small amount of 3.37 was purified using flash column chromatography (4:5:1 CHCl$_3$:MeOH:H$_2$O) for characterization, and the remainder was used without further purification. Rf (4:5:1 CHCl$_3$:MeOH:H$_2$O): 0.10 ¹H NMR (400 MHz, MeOD): $\delta$ ppm 2.61-2.72 (m, 3H, NH$_2$, H4’’), 3.18-3.28 (m, 2H), 3.44-3.61 (m, 4H), 3.63-3.71 (m, 2H), 3.75 (dd, 1H, J 9.2 Hz), 3.79-3.97 (m, 6H), 4.41 (d, 1H, J 7.5 Hz, H1’’’), 4.48 (d, 1H, J 7.8 Hz, H1’), 5.57 (dd, 1H, J 53.9 Hz, H1) ¹⁹F NMR (282 MHz, MeOD) $\delta$ ppm -151.64 (dd, J 26.3 J 53.9 Hz) ¹³C NMR (100 MHz, MeOD): $\delta$ ppm 54.53 (C4’’’), 61.05, 61.57, 62.67, 72.76, 72.98, 74.55, 75.16, 76.10, 76.60, 77.24, 78.04, 79.54, 80.53, 104.31 (C1’), 104.76 (C1’’’), 107.71 (C1) HRMS calculated for C$_{18}$H$_{32}$FNO$_{14}$: 505.1807, found: 506.1885 [M+H]$^+$. 

O-(4-Amino-β-d-glucopyranosyl)-(1→4)-β-d-glucopyranosyl-(1→4)-β-d-glucopyranosyl-(1→4)-β-d-glucopyranosyl-(1→4)-β-d-glucopyranosyl-(1→1)-(2S, 3R, 4E)-2-amino-octadec-4-ene-1,3-diol (3.38)

Cellotriosyl sphingosine (15.6 mg, 19.8 µmol, 3.04) and 4’’-amino-cellotriosyl fluoride (theoretical maximum: 15.0 mg, 29.7 µmol, 1.5 eq., 3.37) were dissolved in 100 mM MES (pH 7.0) to a final acceptor concentration of 10 mM, and CelB glycosynthase (1.0 mg/mL) was added. The reaction mixture was incubated at 30°C. The insoluble beige precipitate (3.38) was
isolated by centrifugation, lyophilized, and used without further purification. **HRMS** calculated for \(\text{C}_{34}\text{H}_{98}\text{N}_2\text{O}_{31}\): 1270.6154, found: 1271.6232 \([\text{M+H}]^+\).

\[
\text{O-} (\beta\text{-}\text{D-Glucopyranosyl})-(1\rightarrow4)-\beta\text{-}\text{D-glucopyranosyl}-(1\rightarrow4)-\beta\text{-}\text{D-glucopyranosyl}-(1\rightarrow1)-(2\text{S}, \, 3\text{R}, \, 4\text{E})-2\text{-azidooc-tadec}-4\text{-ene}-1,3\text{-diol (3.39)}
\]

Sodium azide (155 mg, 2.39 mmol, 1.5 eq.) was dissolved in 0.8 mL of \(\text{H}_2\text{O}\), and 0.8 mL of toluene was added. The mixture was cooled to \(0^\circ\text{C}\) with vigorous stirring. Trifluoromethanesulfonic anhydride (0.268 mL, 1.60 mmol) was added dropwise. The mixture was vigorously stirred for 30 minutes, warmed up to \(10^\circ\text{C}\), and stirred for 2 hours. The reaction mixture was adjusted to pH 7 using saturated \(\text{NaHCO}_3\). The two phases were separated, and the aqueous layer was extracted with toluene, yielding a solution of trifluoromethanesulfonyl azide in toluene. Cellotriosyl sphingosine (21.1 mg, 26.9 µmol, 3.04), copper sulfate pentahydrate (5.68 mg, 35.6 µmol), and \(\text{NaHCO}_3\) (4.46 mg, 53.1 µmol) were dissolved in 0.8 mL \(\text{H}_2\text{O}\). The triflyl azide solution was immediately added, and the reaction was stirred vigorously overnight at RT, during which the reaction mixture changed color from green to yellow. The mixture was concentrated \textit{in vacuo}, and the residue was loaded onto a Waters Sep-Pak® tC18 cartridge (2 g) using a gradient of acetonitrile:water. The product eluted at 75% acetonitrile, and the fractions were lyophilized, yielding \textbf{3.39} as a white solid (14.0 mg, 17.3 µmol, 64%). **Rf** (4:5:2 \(\text{CHCl}_3\):MeOH:10% AcOH): 0.82 **HRMS** calculated for \(\text{C}_{36}\text{H}_{65}\text{N}_3\text{O}_{17}\): 811.4314, found: 834.4212 \([\text{M+Na}]^+\).
O-(4-Amino-β-D-glucopyranosyl)-(1→4)-β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl-(1→1)-(2S, 3R, 4E)-2-azidooctadec-4-ene-1,3-diol (3.40)

3.39 (9.48 mg, 11.7 µmol) and 4-amino-cellotriosyl fluoride (theoretical maximum: 17.7 mg, 35.0 µmol, 3.0 eq., 3.37) were dissolved in 100 mM MES (pH 7.0) to a final acceptor concentration of 10 mM. CelB glycosynthase (1.0 mg/mL) was added, and the reaction mixture was incubated at 30°C. The insoluble product (3.40) was isolated by centrifugation, lyophilized, and used without further purification. HRMS calculated for C_{54}H_{96}N_{4}O_{31}: 1296.6059, found: 1297.6137 [M+H]^+.

2-(4,4-Difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacene-3-propionamido) ethyl β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl-(1→4)-β-D-glucopyranoside (3.43)

Bodipy cellobioside (9.99 mg, 15.1 µmol, 2.51) and α-celllobiosyl fluoride (17.1 mg, 49.7 µmol, 3.3 eq., 2.12) were dissolved in 50 mM sodium phosphate buffer (pH 8.0) to a final acceptor concentration of 20 mM, and CelB glycosynthase (0.50 mg/mL) was added. The reaction was incubated at 30°C, and monitored by TLC (7:2:1 EtOAc:MeOH:H_2O). After 7 days, the reaction
mixture was lyophilized, and loaded onto a Waters Sep-Pak® tC18 cartridge (5 g) to remove any cello-oligosaccharyl fluorides. The bodipy compounds eluted at 25% acetonitrile, and crude 3.43 was used without further purification. HRMS calculated for C_{40}H_{60}BF_{2}N_{3}O_{22}: 983.3730, found: 1006.3627 [M+Na]^+.

16-(Acetylthio)hexadecanoic acid (3.45)

16-Hydroxyhexadecanoic acid (700 mg, 2.57 mmol, 3.44) was weighed out in a 50 mL round bottom flask. Aqueous HBr (48% solution, 12 mL, 106 mmol, 41.2 eq.) and AcOH (12 mL, 209 mmol, 81.3 eq.) were added, and the reaction mixture was heated to 60°C to dissolve 16-hydroxyhexadecanoic acid. Once dissolved, the mixture was refluxed at 120°C for 3 days. The product crystallized upon cooling to RT from the brown solution. The crystals were washed with ice cold water, yielding 16-bromohexadecanoic acid as a white-grey solid (886 mg, 2.64 mmol, 99%). 16-Bromohexadecanoic acid was dissolved in 18 mL DMF, potassium thioacetate (603 mg, 5.28 mmol, 2.0 eq.) was added and the mixture was stirred at 50°C for 15 hours, yielding a dark brown solution. The reaction mixture was poured into 100 mL of 1:1 Et_{2}O:H_{2}O. The pH was adjusted to pH 5 with 1 M HCl, and the product was extracted with 3x150 mL Et_{2}O. The organic layer was dried over magnesium sulfate, filtered, and concentrated in vacuo, yielding a crude orange-brown solid. The product was purified using flash column chromatography and an eluent of 4:1 hexanes:EtOAc, yielding 3.45 as a yellow solid (613 mg, 1.86 mmol, 70%). ^{1}H and ^{13}C NMR data agreed with that reported previously. \[^{148}\] Rf (4:1 Hexanes:EtOAc): 0.15 \[^{1}H\] NMR (400 MHz, CDCl_{3}): δ ppm 1.21-1.40 (m, 22 H, CH_{2}), 1.56 (dt, 2H, J 7.6 Hz, CH_{2}CH_{2}COOH),
1.63 (dt, 2H, $J 7.2$ Hz, $CH_2CH_2SAc$), 2.32 (s, 3H, SAc), 2.34 (t, 2H, $J 7.6$ Hz, $CH_2COOH$), 2.86 (t, 2H, $J 7.2$ Hz, $CH_2SAc$) $^{13}$C NMR (100 MHz, CDCl$_3$): $\delta$ ppm 24.83, 28.97, 29.20, 29.26, 29.32, 29.38, 29.57, 29.62, 29.64, 29.71, 29.76, 30.79, 34.06, 179.40 (COOH), 196.31 (SAc) ESI MS calculated for C$_{18}$H$_{34}$O$_3$S: 330.2, found: 331.4 [M+H]$^+$. 

2,5-Dioxopyrrolidin-1-yl 16-(acetylthio)hexadecanoate (3.46) 

![Chemical Structure](image)

3.45 (600 mg, 1.82 mmol) was dissolved in 16 mL of DCM. NHS (836 mg, 7.26 mmol, 4.0 eq., 3.19) and EDC hydrochloride (1.45 g, 7.56 mmol, 4.15 eq.) were added, and the reaction was stirred at RT overnight. The reaction was diluted with 60 mL DCM. The organic layer was washed with 2x50 mL deionized H$_2$O, dried over sodium sulfate, filtered, and concentrated in vacuo. The crude solid was purified by flash column chromatography using a gradient of 4:1-2:1 Hexanes:EtOAc, yielding 3.46 as a white solid (497 mg, 1.16 mmol, 64%). $^1$H and $^{13}$C NMR data agreed with that reported previously.$^{148}$ Rf (2:1 Hexanes:EtOAc): 0.42 $^1$H NMR (400 MHz, CDCl$_3$): $\delta$ ppm 1.20-1.43 (m, 22 H, CH$_2$), 1.55 (dt, 2H, $J 5.6$ $J 7.6$ Hz, $CH_2CH_2SAc$), 1.73 (dt, 2H, $J 7.2$ $J 7.6$ Hz, $CH_2CH_2COON$), 2.31 (s, 3H, SAc), 2.59 (t, 2H, $J 7.2$ Hz, $CH_2COON$), 2.86 (m, 6H, $CH_2SAc$, COCH$_2$CH$_2$CO) $^{13}$C NMR (100 MHz, CDCl$_3$): $\delta$ ppm 24.69, 25.71, 28.91, 28.94, 29.20, 29.23, 29.28, 29.46, 29.59, 29.61, 29.65, 29.67, 29.73, 30.72, 30.75, 31.06, 168.81, 169.31, 196.18 ESI MS calculated for C$_{22}$H$_{37}$NO$_5$S: 427.2, found: 428.3 [M+H]$^+$. 

176
(2S, 3R, 4E)-3-Hydroxy-2-(S-acetyl-16-mercaptophexadecanamido)octadec-4-enyl (β-D-glucopyranosyl)-(1→4)-β-D-glucopyranoside (3.47)

Cellobiosyl sphingosine (10.3 mg, 16.5 µmol, 3.03) and 3.46 (9.24 mg, 21.6 µmol, 1.3 eq.) were dissolved in 0.8 mL of DMF. Triethylamine (3.80 µL, 27.2 µmol, 1.65 eq.) was added, and the reaction was stirred at RT overnight. Toluene was added to form a DMF-toluene azeotrope, and the mixture was concentrated in vacuo. The crude solid was purified by flash column chromatography using an eluent of 30:10:1 CHCl₃:MeOH:H₂O, yielding 3.47 (13 mg, 13.9 µmol, 84%). Rf (30:10:1 CHCl₃:MeOH:H₂O): 0.55 ¹H NMR (600 MHz, CDCl₃): δ ppm 0.92 (t, 3H, J 6.9 Hz, CH₃), 1.26-1.34 (m, 40H, CH₂), 1.35-1.41 (m, 4H), 1.54-1.61 (m, 4H), 1.99-2.06 (m, 2H), 2.14 (dd, 2H, J 7.4 J 7.7 Hz), 2.34 (s, 3H, SAc), 2.60-2.63 (m, 2H), 2.87 (dd, 2H, J 7.3 J 7.4 Hz), 3.24-3.29 (m, 2H), 3.33-3.41 (m, 4H), 3.50-3.57 (m, 3H), 3.62-3.68 (m, 1H), 3.80-3.88 (m, 3H), 3.94-4.00 (m, 1H), 4.03-4.08 (m, 1H), 4.20 (dd, 1H, J 3.8 J 10.0 Hz), 4.24 (d, 1H, J 7.8 Hz), 4.41 (d, 1H, J 7.9 Hz), 4.47 (dd, 1H, J 5.6 Hz), 4.55 (dd, 1H, J 6.4 Hz), 4.69 (d, 1H, J 5.5 Hz), 4.83-4.91 (m, 3H), 5.18-5.23 (m, 2H), 5.44-5.50 (m, 1H, Hd), 5.62-5.68 (m, 1H, He), 7.93 (s, 1H, NH) ¹³C NMR (150 MHz, CDCl₃): δ ppm 12.87 (CH₃), 21.22, 24.15, 24.49, 27.30, 27.55, 27.64, 27.87, 27.94, 27.96, 28.01, 28.11, 28.14, 28.18, 28.21, 28.22, 28.25, 28.34, 30.44, 30.93, 34.98, 51.87, 59.79, 60.07, 68.08, 68.77, 70.04, 71.99, 72.10, 73.64, 73.75, 75.40, 75.59, 79.10, 102.33, 102.77, 129.69 (Cd), 131.14 (Ce), 171.61 (C=O), 193.91 (RS-C=O) HRMS calculated for C₄₈H₈₉NO₁₄S: 935.6004, found: 958.5901 [M+Na]^+. 
4.4 Enzymology

4.4.1 Production of Abg2F6 Glycosynthase

Abg2F6 glycosynthase was expressed and purified using a previously published procedure.\textsuperscript{92} E. coli BL21 (DE3) cells containing the expression vector for Abg2F6, designated BL21DE3pet29Abg2F6H6, were plated on LB kanamycin agar (50 $\mu$g/mL kanamycin) and incubated at 30°C overnight. Single cell colonies were inoculated into 30 mL TYP medium containing 50 $\mu$g/mL kanamycin and incubated with shaking at 200 rpm overnight at 30°C. 5 mL of the overnight cultures were inoculated into 500 mL TYP-kan media and incubated at 30°C with shaking at 200 rpm overnight. The cells were induced with 0.1 mM IPTG, and grown with shaking at 200 rpm for four hours at 37°C. The cells were harvested by centrifugation at 5000 rpm for 30 minutes at 4°C. The cell pellets were re-suspended in binding buffer (20 mM Tris (pH 8.0), 500 mM NaCl, 5 mM imidazole) and French pressed. One protease inhibitor cocktail tablet (Roche\textsuperscript{®}) was suspended in 1 mL binding buffer and added to the soluble cell extract. The soluble cell extract was clarified by centrifugation at 15 000 rpm for 30 minutes at 4°C. The protein was purified from the soluble cell extract using a 1 mL HisTrap FF Ni\textsuperscript{2+} affinity column (GE Healthcare\textsuperscript{®}). The protein was eluted with a gradient of 20 mM – 250 mM imidazole over 20 mL. The purity of the protein fractions was verified by SDS-PAGE and visualized by staining with Coomassie Blue. The pure protein fractions were concentrated using an Amicon\textsuperscript{®} Ultra-15 30K centrifugal filter (EMD Millipore\textsuperscript{®}) by centrifugation at 4000 rpm at 4°C. The final protein concentration was estimated by absorbance at 280 nm ($\varepsilon_{280} = 2.01$). The protein yield was 74 mg/L cell culture. The protein was stored at 4°C.
4.4.2 Production of BglK

BglK was expressed and purified using a previously published procedure.\textsuperscript{103} \textit{E. coli} Top10 cells containing the expression vector for BglK, designated Topp10pTrcHisBBglK, were inoculated into LB ampicillin media (100 µg/mL ampicillin) and incubated with shaking at 200 rpm at 30°C overnight. 5 mL of the overnight cultures were inoculated into 500 mL LB-amp media and incubated at 30°C with shaking at 200 rpm for four hours. The cells were induced with 0.1 mM IPTG, and grown with shaking at 200 rpm overnight at 30°C. The cells were harvested by centrifugation at 5000 rpm for 30 minutes at 4°C. The cell pellets were re-suspended in binding buffer (20 mM sodium phosphate (pH 7.2), 500 mM NaCl, 5 mM imidazole) and homogenized. One protease inhibitor cocktail tablet (Roche\textsuperscript{®}) was suspended in 1 mL binding buffer and added to the soluble cell extract. The soluble cell extract was clarified by centrifugation at 15 000 rpm for 30 minutes at 4°C. The protein was purified from the soluble cell extract using a 1 mL HisTrap FF Ni\textsuperscript{2+} affinity column (GE Healthcare\textsuperscript{®}). The protein was eluted with a gradient of 20 mM – 250 mM imidazole over 20 mL. The purity of the protein fractions was verified by SDS-PAGE and visualized by staining with Coomassie Blue. The pure protein fractions were concentrated using an Amicon\textsuperscript{®} Ultra-15 10K centrifugal filter (EMD Millipore\textsuperscript{®}) by centrifugation at 4000 rpm at 4°C. The final protein concentration was estimated using a Pierce\textsuperscript{TM} BCA protein assay kit using 100 mM HEPES (pH 7.5) as a compatible buffer. The protein yield was 32 mg/L cell culture. The protein was stored at 4°C.
4.4.3 Production of CelB Glycosynthase

*E. coli* BL21 (DE3) cells containing the expression vector for CelB glycosynthase, designated pet29bCBM3CelB-glyH6, were plated on LB kanamycin agar (50 µg/mL kanamycin) and incubated at 30°C overnight. Single cell colonies were inoculated into 30 mL 2xYT medium containing 50 µg/mL kanamycin and incubated with shaking at 200 rpm overnight at 30°C. 5 mL of the overnight cultures were inoculated into 600 mL 2xYT-kan media and incubated at 30°C with shaking at 200 rpm. After five hours, the cells were induced with 0.1 mM IPTG, and grown with shaking at 200 rpm overnight at 30°C. The cells were harvested by centrifugation at 5000 rpm for 30 minutes at 4°C. The cell pellets were re-suspended in binding buffer (20 mM sodium phosphate (pH 8.0), 500 mM NaCl, 5 mM imidazole) and French pressed. One protease inhibitor cocktail tablet (Roche®) was suspended in 1 mL binding buffer and added to the soluble cell extract. The soluble cell extract was clarified by centrifugation at 15 000 rpm for 30 minutes at 4°C. The protein was purified from the soluble cell extract using a 1 mL HisTrap FF Ni²⁺ affinity column (GE Healthcare®). The protein was eluted with a gradient of 25 mM – 500 mM imidazole over 20 mL. The purity of the protein fractions was verified by SDS-PAGE and visualized by staining with Coomassie Blue. The pure protein fractions were concentrated using an Amicon® Ultra-15 30K centrifugal filter (EMD Millipore®) by centrifugation at 4000 rpm at 4°C. The final protein concentration was estimated by absorbance at 280 nm (ε₂₈₀ = 2.59). The protein yield was 66 mg/L cell culture. The protein was stored at 4°C.
4.4.4 Production of CelC Glycosynthase

_E. coli_ BL21 (DE3) cells containing the expression vector for CelC glycosynthase, designated pTugCelC-glyH6, were plated on LB kanamycin agar (50 µg/mL kanamycin) and incubated at 30°C overnight. Single cell colonies were inoculated into 30 mL LB medium containing 50 µg/mL kanamycin and incubated with shaking at 200 rpm overnight at 30°C. 5 mL of the overnight cultures were inoculated into 600 mL LB-kan media and incubated at 30°C with shaking at 200 rpm. After five hours, the cells were induced with 0.1 mM IPTG, and grown with shaking at 200 rpm overnight at 30°C. The cells were harvested by centrifugation at 5000 rpm for 30 minutes at 4°C. The cell pellets were re-suspended in binding buffer (20 mM sodium phosphate (pH 7.0), 500 mM NaCl, 5 mM imidazole) and French pressed. One protease inhibitor cocktail tablet (Roche®) was suspended in 1 mL binding buffer and added to the soluble cell extract. The soluble cell extract was clarified by centrifugation at 15 000 rpm for 30 minutes at 4°C. The protein was purified from the soluble cell extract using a 1 mL HisTrap FF Ni^{2+} affinity column (GE Healthcare®). The protein was eluted with a gradient of 20 mM – 500 mM imidazole over 20 mL. The purity of the protein fractions was verified by SDS-PAGE and visualized by staining with Coomassie Blue. The pure protein fractions were concentrated using an Amicon® Ultra-15 30K centrifugal filter (EMD Millipore®) by centrifugation at 4000 rpm at 4°C. The final protein concentration was estimated by absorbance at 280 nm (_ε_{280} = 1.622_). The protein yield was 57 mg/L cell culture. The protein was stored at 4°C.
**4.4.5 Production of CenD Glycosynthase**

*E. coli* BL21 (DE3) cells containing the expression vector for CenDcd glycosynthase, designated pet29bCenDcd-glyH6, were plated on LB kanamycin agar (50 µg/mL kanamycin) and incubated at 30°C overnight. Single cell colonies were inoculated into 30 mL 2xYT medium containing 50 µg/mL kanamycin and incubated with shaking at 200 rpm overnight at 30°C. 5 mL of the overnight cultures were inoculated into 600 mL 2xYT-kan media and incubated at 30°C with shaking at 200 rpm. After five hours, the cells were induced with 0.1 mM IPTG, and grown with shaking at 200 rpm overnight at 30°C. The cells were harvested by centrifugation at 5000 rpm for 30 minutes at 4°C. The cell pellets were re-suspended in binding buffer (20 mM sodium phosphate (pH 8.0), 500 mM NaCl, 5 mM imidazole) and French pressed. One protease inhibitor cocktail tablet (Roche®) was suspended in 1 mL binding buffer and added to the soluble cell extract. The soluble cell extract was clarified by centrifugation at 15 000 rpm for 30 minutes at 4°C. The protein was purified from the soluble cell extract using a 1 mL HisTrap FF Ni²⁺ affinity column (GE Healthcare®). The protein was eluted with a gradient of 25 mM – 500 mM imidazole over 20 mL. The purity of the protein fractions was verified by SDS-PAGE and visualized by staining with Coomassie Blue. The pure protein fractions were concentrated using an Amicon® Ultra-15 30K centrifugal filter (EMD Millipore®) by centrifugation at 4000 rpm at 4°C. The final protein concentration was estimated by absorbance at 280 nm (ε₂₈₀ = 3.168). The protein yield was 3.5 mg/L cell culture. The protein was stored at 4°C.
4.4.6 Production of EGCcase Glycosynthase

EGCase glycosynthase was expressed and purified using a previously published procedure.\textsuperscript{120} Tuner\textsuperscript{TM} (DE3) cells containing the expression vector for EGCcase glycosynthase, designated pet28aEGC-E351S/D14Y, were plated on LB kanamycin agar (50 µg/mL kanamycin) and incubated at 30°C overnight. Single cell colonies were inoculated into 30 mL TYP medium containing 50 µg/mL kanamycin and incubated with shaking at 200 rpm overnight at 30°C. 5 mL of the overnight cultures were inoculated into 600 mL TYP-kan media and incubated at 30°C with shaking at 200 rpm. After six hours, the cells were induced with 0.1 mM IPTG, and grown with shaking at 200 rpm overnight at 30°C. The cells were harvested by centrifugation at 5000 rpm for 30 minutes at 4°C. The cell pellets were re-suspended in binding buffer (20 mM sodium phosphate (pH 7.4), 500 mM NaCl, 5 mM imidazole) and French pressed. One protease inhibitor cocktail tablet (Roche\textsuperscript{®}) was suspended in 1 mL binding buffer and added to the soluble cell extract. The soluble cell extract was clarified by centrifugation at 15 000 rpm for 30 minutes at 4°C. The protein was purified from the soluble cell extract using a 1 mL HisTrap FF Ni\textsuperscript{2+} affinity column (GE Healthcare\textsuperscript{®}). The protein was eluted with a gradient of 5 mM – 300 mM imidazole over 25 mL. The purity of the protein fractions was verified by SDS-PAGE and visualized by staining with Coomassie Blue. The pure protein fractions were dialyzed into 25 mM sodium acetate (pH 5.3) buffer overnight at 4°C. The final protein concentration was estimated using a Pierce\textsuperscript{TM} BCA protein assay kit using 200 mM sodium acetate (pH 4.8) as a compatible buffer. The protein yield was 94 mg/L cell culture. The protein was stored in the freezer in 0.5 mL aliquots.
4.4.7 Production of CBM2a

CBM2a was expressed and purified using a previously published procedure.\textsuperscript{140} \textit{E. coli} BL21 (DE3) cells containing the expression vector for CBM2a, designated pTugH6CBM2a, were plated on LB kanamycin agar (50 µg/mL kanamycin) and incubated at 30°C overnight. Single cell colonies were inoculated into 30 mL TYP medium containing 50 µg/mL kanamycin and incubated with shaking at 200 rpm overnight at 30°C. 5 mL of the overnight cultures were inoculated into 500 mL TYP-kan media and incubated at 30°C with shaking at 200 rpm. After four hours, the cells were induced with 0.1 mM IPTG, and grown with shaking at 200 rpm overnight at 30°C. The cells were harvested by centrifugation at 5000 rpm for 30 minutes at 4°C. The cell pellets were re-suspended in binding buffer (20 mM Tris (pH 7.6), 500 mM NaCl, 5 mM imidazole) and French pressed. One protease inhibitor cocktail tablet (Roche\textsuperscript{®}) was suspended in 1 mL binding buffer and added to the soluble cell extract. The soluble cell extract was clarified by centrifugation at 15 000 rpm for 30 minutes at 4°C. The protein was purified from the soluble cell extract using a 1 mL HisTrap FF Ni\textsuperscript{2+} affinity column (GE Healthcare\textsuperscript{®}). The protein was eluted with a gradient of 20 mM – 250 mM imidazole over 20 mL. The purity of the protein fractions was verified by SDS-PAGE and visualized by staining with Coomassie Blue. The pure protein fractions were concentrated using an Amicon\textsuperscript{®} Ultra-15 3K centrifugal filter (EMD Millipore\textsuperscript{®}) by centrifugation at 4000 rpm at 4°C. The final protein concentration was estimated using a BCA protein assay. The protein yield was 30 mg/L cell culture. The protein was stored at 4°C.
4.4.7.1 Labeling of CBM2a with FITC

CBM2a was labeled with FITC using a previously published procedure. CBM2a protein was dialyzed against 100 mM sodium bicarbonate (pH 9.0) using a Vivaspin MW cut-off 3000 sample concentrator (GE Healthcare®) by centrifugation at 4000 rpm at 4°C to a concentration ranging from 1-4 mg/mL. 50 µL of freshly prepared fluorescein-5(6)-isothiocyanate dissolved in DMSO (3 mg/mL) was added for every 1 mL of CBM2a. The reaction vessel was wrapped in tin foil and gently mixed for 45 minutes in the dark at RT. Once complete, the reaction mixture was added to a PD-10 desalting column (GE Healthcare®) equilibrated with PBS. After purification, absorbance measurements were read at 280 nm and 496 nm to determine the number of FITC molecules bound per protein. The protein concentration (mg/mL) was estimated using the formula: $(A_{280} - 0.35A_{496})$ divided by 2.24. The protein concentration was converted from mg/mL to mol/L by dividing the concentration by the molecular weight of the protein (12272 Da). The concentration of FITC was estimated using the formula $A_{496}$ divided by the extinction coefficient of FITC (70 000 M$^{-1}$ cm$^{-1}$ at pH 9.0). The number of FITC molecules bound per protein was calculated by dividing the concentration of FITC (mol/L) by the protein concentration (mol/L). The labeled protein was stored in the dark at 4°C.
References

15. Wada, M.; Ike, M.; Tokuyasu, K., Enzymatic hydrolysis of cellulose I is greatly accelerated via its conversion to the cellulose II hydrate form. Polymer Degradation and Stability 2010, 95 (4), 543-548.


120. Caines, M. E. C.; Vaughan, M. D.; Tarling, C. A.; Hancock, S. M.; Warren, R. A. J.; Withers, S. G.; Strynadka, N. C. J., Structural and mechanistic analyses of endo-


of thiols on gold - variation in the length of the alkyl chain. *Journal of the American Chemical Society* 1989, 111 (18), 7164-7175.


