5-9-2016

Chemical and Chemoenzymatic Synthesis of Outer Core Oligosaccharide of Escherichia Coli R3 and a Library of Human Milk Oligosaccharides & Design and Synthesis of Glycoconjugates

Zhongying Xiao

Follow this and additional works at: http://scholarworks.gsu.edu/chemistry_diss

Recommended Citation
http://scholarworks.gsu.edu/chemistry_diss/122

This Dissertation is brought to you for free and open access by the Department of Chemistry at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Chemistry Dissertations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.
CHEMICAL AND CHEMOENZYMATIC SYNTHESIS OF OUTER CORE
OLIGOSACCHARIDE OF ESCHERICHIA COLI R3 AND A LIBRARY OF HUMAN MILK
OLIGOSACCHARIDES

&

DESIGN AND SYNTHESIS OF GLYCOCONJUGATES

by

ZHONGYING XIAO

Under the Direction of Peng George Wang, Ph.D.

ABSTRACT

Lipopolysaccharides (LPS), major virulence determinants in Gram–negative bacteria, are responsible for many pathophysiological responses and can elicit strong immune responses. In order to better understand the role of LPS in host–pathogen interactions and elucidate the immunogenic properties of LPS outer core oligosaccharide, an all α–linked Escherichia coli R3 outer core pentasaccharide was first synthesized with a propyl amino linker at the reducing end. This oligosaccharide was also covalently conjugated to a carrier protein (CRM₁₉₇) via the reducing end propyl amino linker. An immunological analysis demonstrated
that this glycoconjugate can elicit specific anti-pentasaccharide antibodies with in vitro bactericidal activity. These findings will contribute to further exploring this pentasaccharide antigen as a vaccine candidate.

Human milk oligosaccharides (HMOs) are a family of diverse unconjugated glycans that exist in human milk as one of the major components. Characterization, quantification and biofunctional studies of HMOs remain a big challenge due to their diversity and complexity. The accessibility of homogenous HMOs library is essential to solve these issues which have beset academia for several decades. In this study, an efficient chemoenzymatic strategy, namely Core Synthesis/Enzymatic Extension (CSEE), for rapid production of diverse HMOs was reported. Based on 3 versatile building blocks and 4 robust glycosyltransferases, a library of 31 HMOs were chemoenzymatically synthesized and characterized by MS and NMR. CSEE indeed provides a practical approach to harvest structurally defined HMOs for various applications.

Glycoproteins are extremely important for all life on the planet. Glycoproteins play important roles in various biological processes. Increasing evidences demonstrate that glycosylation of proteins could improve stability of proteins by stabilizing their tertiary structure and protecting them from proteolysis. Besides, glycosylation of proteins could provide targeting effects through glycan-lectin interaction. Furthermore, carbohydrates play crucial roles in humoral immunity in that many sugar epitopes are identified as antigens for antibodies. Glycoprotein could boost strong T cells mediated intercellular immune responses because homogeneous antigens present on the surface of proteins by multivalent bonds. In this study, the three advantages of glycoproteins, namely
stabilizing proteins, targeting effects and eliciting immunological response, were extensively explored by broad collaboration with other groups.

INDEX WORDS: LPS, Outer Core, HMOs, CSEE, Glycosylation of Proteins, Carbohydrate Epitopes.
CHEMICAL AND CHEMOENZYMATIC SYNTHESIS OF OUTER CORE
OLIGOSACCHARIDE OF *ESCHERICHIA COLI* R3 AND A LIBRARY OF HUMAN MILK
OLIGOSACCHARIDES
&
DESIGN AND SYNTHESIS OF GLYCOCONJUGATES

by

ZHONGYING XIAO

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy
in the College of Arts and Sciences
Georgia State University
2016
CHEMICAL AND CHEMOENZYMATIC SYNTHESIS OF OUTER CORE OLIGOSACCHARIDE OF *ESCHERICHIA COLI* R3 AND A LIBRARY OF HUMAN MILK OLIGOSACCHARIDES

&

DESIGN AND SYNTHESIS OF GLYCOCONJUGATES

by

ZHONGYING XIAO

Committee Chair: Peng G. Wang

Committee: Markus W. Germann

Suri Saranathan Iyer

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2016
DEDICATION

Dedicated to my wife Rong Hu, and my parents, for their love, understanding and support.
ACKNOWLEDGEMENTS

I really appreciate my research advisor Dr. Peng George Wang for his financial support and incessant guidance in the past five years. Dr. Wang’s enthusiasm for sciences and intelligence inspired my research careers. Dr. Markus W. Germann and Dr. Suri Saranathan Iyer deserve recognition for their willingness to serve on my Defense Committees. Many thanks to Dr. Jenny. J. Yang, Dr. Brantley R. Herrin and Dr. Eli Gilboa for the collaborative projects.

It is my fortune to work with the Wang Group, deserving to mention here. Especially, I need to thank Dr. Pintu Kumar Mandal who introduced me to the art of organic synthesis during my first year. I also want to thank Dr. Tiehai Li and Dr. Yunpeng Liu for their patient help and guidance for my research. Many thanks to Dr. Lei Li, Dr. Cheng Ma and Dr. Jingyao Qu for providing meaningful suggestions for my projects. Thanks for Yuxi Guo for HMOs project, and other members for my research career.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................ v

LIST OF SCHEMES ................................................................................................................... X

LIST OF FIGURES ..................................................................................................................... XI

LIST OF ABBREVIATIONS ........................................................................................................ XVII

1. Chemical Synthesis of Outer Core Oligosaccharide of *Escherichia coli* R3 and Immunological Evaluation ................................................................. 1

   1.1 Introduction ....................................................................................................................... 1

   1.2 Results and discussion ...................................................................................................... 3

       1.2.1 Synthesis of Fragments of outer core ................................................................. 3

       1.2.2 Initial attempt to synthesize outer core pentasaccharide .................. 7

       1.2.3 Synthesis of Disaccharide 1-7 ................................................................. 9

       1.2.4 Assembly of pentasaccharide 1-1 .......................................................... 10

       1.2.5 Conjugation of the pentasaccharide to carrier protein ..................... 11

       1.2.6 Immunological evaluation of the pentasaccharide-CRM$_{197}$ glycoconjugate ................................................................. 12

       1.2.7 Bactericidal activity ............................................................................................... 14

   1.3 Conclusion ....................................................................................................................... 15

   1.4 Experimental Section ..................................................................................................... 16

   1.5 References ...................................................................................................................... 31
2. Chemoenzymatic Synthesis of a library of Human Milk Oligosaccharides ......................................................................................................................... 35

2.1 Introduction ........................................................................................................... 35

2.1.1 History of HMOs .......................................................................................... 35

2.1.2 Structures of HMOs ...................................................................................... 35

2.1.3 Previous Synthesis of HMOs ........................................................................ 37

2.2 Convergent Core Synthesis ................................................................................. 39

2.3 Synthesis of Building Blocks ............................................................................. 40

2.4 Synthesis of Core Oligosaccharides ................................................................... 41

2.5 Enzymatic Extension of Core Structures ........................................................... 44

2.6 Conclusion ............................................................................................................ 47

2.7 Experimental Section ......................................................................................... 47

2.8 References ............................................................................................................ 80

3. Glycosylation of Protein-based Magnetic Resonance Imaging Contrast Agents 84

3.1 Introduction .......................................................................................................... 84

3.2 Results and Discussion ....................................................................................... 88

3.3 Conclusion ............................................................................................................ 94

3.4 Experimental Section ......................................................................................... 95

3.5 References ............................................................................................................ 96
4. Design and Synthesis of Glycoprotein Antigens for Immunotherapy 98

4.1 Introduction ........................................................................................................... 98

4.2 Carbohydrate antigens based vaccine for activating VLRs in Lampreys 100

4.3 Design and synthesis of glycoproteins ................................................................. 100

4.3.1 Synthesis of α-Gal conjugated proteins ......................................................... 103

4.3.2 Synthesis of H-antigen conjugated proteins .................................................. 105

4.4 Conclusion ............................................................................................................ 108

4.5 References ........................................................................................................... 108

5. Facile Preparation of Amylose-Proteins by Phosphorylase-catalyzed Polymerization ......................................................................................................................... 110

5.1 Introduction ........................................................................................................... 110

5.1.1 PEGylation and Glycosylation of Therapeutic Proteins ......................... 111

5.1.2 Phosphorylase-catalyzed enzymatic polymerization .............................. 113

5.2 Results and Discussion ........................................................................................ 115

5.3 Conclusion ............................................................................................................ 120

5.4 Experimental Section .......................................................................................... 120

5.5 References ........................................................................................................... 122

6. Preparation of α-Gal-aptamer Conjugates for anticancer therapy .... 124

6.1 Introduction ........................................................................................................... 124
6.2 Results and Discussion ......................................................... 127

6.3 Conclusion ............................................................................. 132

6.4 References ............................................................................. 133

APPENDICES: $^1$H, $^{13}$C NMR Spectra and HPLC Profiles of Compounds ... 135
LIST OF SCHEMES

Scheme 1.1. Synthesis of Trisaccharide 1-21 ............................................................... 5
Scheme 1.2. Synthesis of Tetrasaccharide 1-23 .............................................................. 6
Scheme 1.3. Synthesis of Disaccharide 1-7 ................................................................. 9
Scheme 1.4. Synthesis of Pentasaccharide 1-1 ............................................................... 11
Scheme 1.5. Preparation of the pentasaccharide-proteins conjugate .................... 12
Scheme 2.1. Synthesis of Building Block 2-1 ............................................................... 40
Scheme 2.2 Synthesis of Building Block 3 ................................................................. 41
Scheme 2.3 Synthesis of HMO1 ............................................................................... 42
Scheme 2.4. Synthesis of HMO2 ............................................................................... 43
Scheme 2.5 Synthesis of HMO3 ............................................................................... 44
Scheme 3.1. Chemical Synthesis of Lac-rProCA1 ....................................................... 89
Scheme 4.1. Chemical Ligation of Glycoproteins ....................................................... 101
Scheme 4.2. Chemical Synthesis of Glycoprotein ..................................................... 102
Scheme 4.3 Chemoenzymatic Synthesis of α-Gal Linker ....................................... 103
Scheme 4.4 Preparation of α-Gal Proteins ................................................................. 104
Scheme 4.5 Preparation of H Antigen Proteins ......................................................... 106
Scheme 5.1 Application of phosphorylase-catalyzed polymerization .................... 114
Scheme 5.2 Design of enzymatic polymerization of maltoheptaose-proteins ... 117
Scheme 5.3 Synthesis of Maltoheptaose Donors ....................................................... 118
Scheme 5.4 Conjugation and polymerization of Glyco-BSA ................................. 119
Scheme 6.1. ssRNA Conjugation Reaction ................................................................. 130
LIST OF FIGURES

Figure 1.1. (A) The structure of E. coli O157:H7 LPS; (B) The structure of E. coli R3 outer core oligosaccharide with propyl amino linker ......................................................... 3

Figure 1.2. Fragments of Outer Core ................................................................. 4

Figure 1.3. Preparation of Oligosaccharide-CRM_{197} Glycoconjugates ............ 7

Figure 1.4. Retrosynthetic analysis of outer core oligosaccharide of E. coli R3... 8

Figure 1.5. Evaluation of anti–pentasaccharide antibody titers (IgG and IgM) after the third immunization. ........................................................................................................ 13

Figure 1.6. Antibody IgG subclass profiles after the third immunization, (A) IgG1 titers, (B) IgG2a titers, (C) IgG2b titers, (D) IgG3 titers. ................................................................. 14

Figure 1.7. Bactericidal activity of serum from the immunized mice against E. coli O157:H7 ......................................................................................................................... 15

Figure 2.1. Representative fucosylated or sialylated HMOs ......................... 36

Figure 2.2. HMOs synthesized by the Core Synthesis/Enzymatic Extension (CSEE) strategy starting with 3 chemically prepared core structures (boxed) .......... 38

Figure 2.3. The three versatile building blocks for the assembly of the three core structures. ..................................................................................................................... 39

Figure 2.4. Enzymatic extension of human milk oligosaccharides ................. 46

Figure 3.1. Examples of Clinical Gd^{3+}-Based MRI Contrast Agents. ............ 85

Figure 3.2. A model structure of ProCA1 with Gd^{3+} and relaxivity value r1 and r2 of ProCA1 .................................................................................................................. 86
Figure 3.3. Binding model for ASGPR ligands in an optimal conformation to the heterooligomeric receptor................................................................. 88
Figure 3.4. Glycosylation of rProCA1. ......................................................... 91
Figure 3.5. Serum Stability of rProCA1 and Lac-rProCA1................................. 92
Figure 3.6. MR images of CD1 mouse with injection of rProCA1 and Lac-
rProCA1. ................................................................................................. 93
Figure 3.7. MRI contrast of Lac-rProCA1 in mice ............................................ 94
Figure 4.1. Carbohydrate antigen/Antibody-directed vaccine design................. 99
Figure 4.2. Characterization of α-Gal Biotin BSA ......................................... 104
Figure 4.3. Characterization of α-Gal OVA.................................................. 105
Figure 4.4. Characterization of H Antigen Biotin BSA................................. 107
Figure 4.5. Characterization of H Antigen OVA........................................... 107
Figure 5.1. Benefits of PEGylated therapeutic proteins .................................. 112
Figure 5.2. Benefits of glycosylated therapeutic proteins ............................... 113
Figure 5.3. Phosphorylase-catalyzed enzymatic polymerization ................... 113
Figure 5.4. Model Reaction of Polymerization catalyzed by Phosphorylase..... 116
Figure 5.5. SDS gel of Glycosylation and Polymerization of BSA................. 120
Figure 6.1. Screen RNA Aptamer-the SELEX process.................................. 125
Figure 6.2. Anticancer Mechanism of α-Gal-aptamer Conjugates ............... 128
Figure 6.3. Methods of Synthesis of Carbohydrate-Oligonucleotide Conjugates (COCs).............................................................................................. 129
Figure 6.4. MALDI Result for Biotin-ssRNA conjugation............................ 131
Figure 6.5. Modeling Reaction of COCs ...................................................... 132
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>alpha</td>
</tr>
<tr>
<td>Ac</td>
<td>acetyl</td>
</tr>
<tr>
<td>br</td>
<td>broad (IR and NMR)</td>
</tr>
<tr>
<td>β</td>
<td>beta</td>
</tr>
<tr>
<td>n-Bu</td>
<td>normal-butyl</td>
</tr>
<tr>
<td>t-Bu</td>
<td>tert-butyl</td>
</tr>
<tr>
<td>Bz</td>
<td>benzoyle</td>
</tr>
<tr>
<td>°C</td>
<td>degrees Celsius</td>
</tr>
<tr>
<td>calcld</td>
<td>calculated</td>
</tr>
<tr>
<td>δ</td>
<td>chemical shift in parts per million downfield from tetramethylsilane</td>
</tr>
<tr>
<td>d</td>
<td>doublet (spectra); day(s)</td>
</tr>
<tr>
<td>DBU</td>
<td>1,8-diazabicyclo[5.4.0]undec-7-ene</td>
</tr>
<tr>
<td>DMAP</td>
<td>4-(N,N-dimethylamino)pyridine</td>
</tr>
<tr>
<td>DMF</td>
<td>N,N-dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulfoxide</td>
</tr>
<tr>
<td>EDCI</td>
<td>N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride</td>
</tr>
<tr>
<td>eq.</td>
<td>equivalent</td>
</tr>
<tr>
<td>Et</td>
<td>ethyl</td>
</tr>
<tr>
<td>γ</td>
<td>gamma</td>
</tr>
<tr>
<td>g</td>
<td>gram(s)</td>
</tr>
<tr>
<td>h</td>
<td>hour(s)</td>
</tr>
</tbody>
</table>
J  coupling constant in Hz (NMR)
k  kilo
L  liter(s)
m  milli; multiplet (NMR)
μ  micro
M  moles per liter
Me  methyl
MHz  megahertz
min  minute(s)
mol  mole(s)
Ms  methanesulfonyl
MS  mass spectrometry; molecular sieves
m/z  mass to charge ratio (MS)
NMR  nuclear magnetic resonance
p  para
Ph  phenyl
PMB  p-methoxybenzyl
ppm  parts per million
py  pyridine
q  quartet (NMR)
rt  room temperature
s  singlet (NMR); second(s)
t  tertiary (tert)
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Compound Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>triplet (NMR)</td>
</tr>
<tr>
<td>TBAF</td>
<td>tetrabutylammonium fluoride</td>
</tr>
<tr>
<td>TBAI</td>
<td>tetrabutylammonium iodide</td>
</tr>
<tr>
<td>TBS</td>
<td>t-butyl(dimethyl)silyl</td>
</tr>
<tr>
<td>Tf</td>
<td>trifluoromethanesulfonyl</td>
</tr>
<tr>
<td>THF</td>
<td>tetrahydrofuran</td>
</tr>
<tr>
<td>TLC</td>
<td>thin layer chromatography</td>
</tr>
<tr>
<td>TMS</td>
<td>trimethylsilyl</td>
</tr>
<tr>
<td>TSTU</td>
<td>N,N,N',N-tetramethyl (succinimido) uronium tetrafluoroborate</td>
</tr>
</tbody>
</table>
1. Chemical Synthesis of Outer Core Oligosaccharide of *Escherichia coli* R3 and Immunological Evaluation

1.1 Introduction

Lipopolysaccharides (LPS) are important cell envelope components of Gram-negative bacteria, embedded in the outer leaflet of the asymmetric outer membrane and exposed on the cell surface.\(^1\) LPS contribute to the integrity of outer membrane and protect the cells against various environmental stresses including lipophilic antibiotics and host immune system.\(^2\) Moreover, LPS act as strong stimulators of innate and adaptive immunity in diverse eukaryotic species ranging from insects to humans.\(^2-5\) Therefore, the study of LPS has attracted considerable interest with respect to the development of vaccines, therapeutics and diagnostics.\(^6-8\) LPS typically consist of three parts: an endotoxic lipid A comprising a bisphosphorylated and acylated \(\beta-\(1\rightarrow6\)\)–linked glucosamine disaccharide backbone, a core oligosaccharide, and a distal polysaccharide called O–antigen whose composition widely varies in different bacterial species.\(^4, 9, 10\) Structurally, the core oligosaccharide region can be further subdivided into an inner core region typically consisting of 3–deoxy–D–manno–oct–2–ulosonic acid (Kdo) and L–glycero–D–manno–heptopyranonse (Hep) residues, and an outer core region that contains various hexoses and hexosamines.\(^4, 11, 12\)

In *Escherichia coli* (E. coli), the core oligosaccharides have limited variation with only five core structures named R1, R2, R3, R4 and K12.\(^13\) The R3 core type is of biomedical interest because it is found in most verotoxigenic isolates such as O157:H7,\(^14\) which causes approximately 210 million cases of diarrhea and 380,000 deaths in the developing world each year.\(^15\) The structure of *E. coli* O157:H7 LPS
has been determined and reported in previous literatures (Figure 1A).\textsuperscript{14, 16, 17} \textit{E. coli} O157:H7 has a typical LPS consisting of a lipid moiety, a core oligosaccharide (inner core and outer core) and an O–antigen. The development of LPS–based vaccines, therapeutics, and diagnostics requires well-defined oligosaccharides conjugated to carrier proteins for immunological studies to identify the structural motifs that can elicit protective antibody responses.\textsuperscript{8, 18–21} Although oligosaccharides can be isolated from bacteria, the homogeneity cannot be achieved due to the variations of natural oligosaccharides during the processes of isolation and conjugation to carriers.\textsuperscript{22} Chemical synthesis of well-defined oligosaccharides can overcome this problem by utilizing an artificial linker at the reducing end that allows conjugation to carrier proteins.\textsuperscript{20, 23–27} Moreover, it can provide various sub–structures for structure–activity relationships study to determine the minimal epitope that can elicit protective immune response.\textsuperscript{19, 28} Herein, an all α–linked pentasaccharide of \textit{E. coli} R3 outer core was chemically synthesized with a propyl amino linker at the reducing end (Figure 1.1B), and was conjugated to the nontoxic mutant of diphtheria toxin CRM\textsubscript{197} to afford a glycoprotein. Furthermore, levels of IgG and IgM antibodies against the pentasaccharide and \textit{in vitro} bactericidal activity of these antibodies against \textit{E. coli} O157:H7 were evaluated.
1.2 Results and discussion

1.2.1 Synthesis of Fragments of outer core

In the early stage of the project, we planned to develop vaccines based on fragments of outer core of *E. coli* R3 rather than the whole core, which could decrease the difficulties of preparation of vaccines. Furthermore, to explore the possible immunogenicity of inner core, we tried to couple the outer core with the first Hep residue of inner core, which was substituted by mannose (Man) block because Man is structurally similar with Hep. Based on our initial ideas, linear backbones of outer core linked with a propyl amino linker,
Compound 1-17 and 1-18, were designed. These simplified outer cores were conjugated with CRM\textsubscript{197} for immunological tests.

The synthesis of trisaccharide 1-21 was started with glycosylation of donor 1-2 with acceptor 1-8 promoted by TMSOTf under stable -20 °C to furnish disaccharide 1-19, it was separated with its β-isomer by silica gel (α/β = 6:1). Oxidative cleavage of the anomeric 4–methoxyphenyl (MP) moiety of 1-19 by ceric ammonium nitrate gave a lactol, which was converted into the corresponding trichloroacetimidate 1-20 under CCl\textsubscript{3}CN/DBU condition in 73% yield over two steps. A TMSOTf–mediated coupling of trichloroacetimidate 1-20 with the acceptor 1-15 in a diethyl ether/dichloromethane (1:1) solvent system led to desired α-linked trisaccharide 1-17 ($J_{\text{H-H}} = 3.2$ Hz, α/β = 10:1) in 60% yield. The primary amine of trisaccharide 1-17 was reacted with an excess of succinic anhydride in the presence of triethylamine and DMF to afford monoacid linker 1-21 for next step’s conjugation (Scheme 1.1).
The tetrasaccharide 1-23 was synthesized by a convergent [2+2] block coupling strategy (Scheme 1.2). Glycosylation of Donor 1-20 and Acceptor 1-21 in DCM/Et\textsubscript{2}O solvent system was proceeded fast catalyzed by TMSOTf to afford tetrasaccharide 1-22. O-Deacetylation of compound 1-22 was performed under Zemplén conditions, followed by the global deprotection of Benzyl group (Bn) and azido group by catalytic hydrogenolysis with Pd(OH)\textsubscript{2}/H\textsubscript{2} in MeOH. The tetrasaccharide 1-18 was produced in a total yield of 90% over the two steps. The primary amine of tetrasaccharide 1-18 was reacted with an excess of succinic anhydride in the presence of triethylamine to achieve monoacid linker 1-23 for next step’s conjugation (Scheme 1.2).
Scheme 1.2. Synthesis of Tetrasaccharide 1-23

With these two oligosaccharides in hand, we need to explore a good method to couple them onto the carrier protein CRM\textsuperscript{197}. \textit{N}, \textit{N}, \textit{N'}, \textit{N'}-tetramethyl (succinimido) uranium tetrafluoroborate (TSTU) was found to be an ideal reagent for the formation of carboxamidase from acid.\textsuperscript{29} This reagent is compatible with polar solvent such as aqueous solvent and DMF. As a result, NHS ester was accomplished by treating acid 1-21 and 1-23 with TSTU in the presence of Et\textsubscript{3}N in DMF solvent. Both TLC and MS confirmed the completion of the reaction in 1 hour. During this activation step, only minimal amount of TSTU (1.1 eq) and base (1.1 eq) were applied and completion of reaction was confirmed by TLC. The crude product was directly used for next step after concentrated to dryness. The modified oligosaccharide with an activated ester was then conjugated to amino groups of lysine residues of the protein (the molar ratio: carbohydrate/protein 50:1) in PBS buffer (pH 7.4) to afford the glycoproteins. These two glycoproteins were confirmed by SDS–PAGE analysis, displaying a shift toward a higher mass of glycoprotein compared with unconjugated CRM\textsuperscript{197}. Unfortunately, these two glycoproteins did not show strong
immunological activities in mice test and we have to move forward to synthesize the whole pentaccharide *E. coli* R3.

![Figure 1.3. Preparation of Oligosaccharide-CRM$_{197}$ Glycoconjugates](image)

### 1.2.2 Initial attempt to synthesize outer core pentasaccharide

The chemical synthesis of the outer core pentasaccharide of *E. coli* R3 is challenging, because it is highly branched and consists of all α–linked glycosidic bonds. Furthermore, the pentasaccharide contains a number of glycosides that are difficult to install in a stereo–selective fashion. Especially, the introduction of α–glucosides and α–galactosides often leads to the formation of a mixture of anomer, which are difficult to separate and result in lower yields of the desired products. Installing the highly crowded cis–1,2,3–α–linked galactose (B) with glucose (C) and N-acetylglucosamine (D) needs to establish an optimal order of glycosylation. As illustrated in Figure 1.4, the fully protected pentasaccharide was initially envisioned to be synthesized by a convergent and stereo–controlled [2 + 3] approach. Unfortunately, the desired product was not detected by ESI mass spectrometry analysis, and substantial decomposition of the glycosyl donor (trichloroacetimidate) and acceptor were observed by TLC. It was considered that the steric hindrance resulted in inaccessibility of C–2 hydroxy group for the third glycosylation
after C–1 and C–3 hydroxy groups of galactose (B) were glycosylated. Therefore, a [4 + 1] coupling approach was carried out by initially installing C–2 hydroxy group of galactose (B), followed by glycosylation of C–3 and C–1 hydroxy groups. Stereo-controlled glycosylations were assisted by solvent effects and temperature control. Moreover, C–3 and C–1 of the galactose building block (B) were modified by orthogonal protecting groups benzoyl ester (Bz) and 4–methoxyphenyl (MP), respectively. The orthogonal protecting groups made it possible to establish the proper order of glycosylation to assemble the highly branched pentasaccharide.

Figure 1.4. Retrosynthetic analysis of outer core oligosaccharide of *E. coli* R3
1.2.3 Synthesis of Disaccharide 1-7

Having established the optimal order of glycosylation, cis-1,2-α-linked disaccharide donor 1-7 needed to be synthesized (Scheme 1.3). Assisted by solvent effects (diethyl ether as co-solvent with dichloromethane), glycosylation of the donor 1-2 with a nonparticipation benzyl ether at C–2 and the acceptor 1-3 synthesized in our previous report predominantly afforded α-anomer ($^1J_{H,H} = 3.6$ Hz, $\alpha/\beta=10:1$) at $-40 ^\circ C$. The isolated disaccharide 1-4 was contaminated by the rearrangement product of trichloroacetimidate donor 1-2. The contaminated compound 1-4 was directly subjected to removal of the benzyldiene acetal using trifluoroacetic acid (TFA) in a mixture of CH$_2$Cl$_2$ and water, which made it easy to separate the byproduct to give desired diol 1-5. Benzylaion of compound 1-5 with benzyl bromide and NaH in DMF gave the purified benzyl ether 1-6 in 53% yield over three steps. The compound 1-6 was treated with N–bromosuccinimide (NBS) to afford the resulting lactol, which was converted into the corresponding trichloroacetimidate 1-7 using trichloroacetonitrile and 1, 8–diazabicycloundec–7–ene (DBU) in 68% yield over two steps.

1.2.4 Assembly of pentasaccharide 1-1

A successful synthetic strategy of compound 1-1 was illustrated in Scheme 1.4. With the disaccharide donor 1-7 in hand, coupling of the trichloroacetimidate 1-7 with the acceptor 1-8 in the presence of catalytic amount of TMSOTf as the promoter in diethyl ether/dichloromethane (1:1) solvent system provided a trisaccharide at –20 °C, which was purified to afford 1-9 as main α–anomer ($^{1}J_{H-H} = 3.6$ Hz, $\alpha/\beta \geq 20:1$) in 55% yield. Removal of benzoyl group (Bz) was easily accomplished by treatment with base (NaOMe) to afford trisaccharide acceptor 1-10 in 96% yield. The partially benzylated donor 1-11 was much more stable and easier to use compared to the corresponding extremely reactive fully benzylated donor. The trisaccharide 1-10 was coupled with glycosyl donor 1-11 using a catalytic amount of TMSOTf in diethyl ether/dichloromethane (1:1) solvent system to afford the desired α-linked tetrasaccharide 1-12 ($^{1}J_{H-H} = 3.2$ Hz) in 76% yield, no β–anomer byproduct was detected. The azide moiety of 1-12 was reduced with NaBH$_4$ and NiCl$_2$H$_2$O to give amine, which was immediately acetylated to afford the resulting 1-13 in excellent yield (94%). Oxidative cleavage of the anemic MP (4–methoxyphenyl) moiety of 1-13 using ceric ammonium nitrate gave a lactol, which was converted into the corresponding trichloroacetimidate 1-14 using trichloroacetonitrile and DBU in 53% yield over two steps. A TMSOTf–mediated coupling of trichloroacetimidate 1-14 with the acceptor 1-15 in a diethyl ether/dichloromethane (1:1) solvent system led to desired α-linked pentasaccharide 1-16 ($^{1}J_{H-H} = 3.2$ Hz, $\alpha/\beta \geq 20:1$) in 42% yield. Finally, the deprotection of 1-16 started with removal of the acetyl groups using base (NaOMe), followed by using Pearlman's catalyst (Pd(OH)$_2$/C) and H$_2$ to afford the desired target compound 1-1 in 72% yield.
1.2.5 Conjugation of the pentasaccharide to carrier protein

To perform immunological experiments, the outer core pentasaccharide of *E. coli* R3 was conjugated to the nontoxic mutant of diphtheria toxin CRM<sub>197</sub>.<sup>29</sup> CRM<sub>197</sub> is an immunogenic carrier protein, which can improve the immunogenicity of oligosaccharide antigens and induce a T-cell dependent immune response.<sup>45</sup> This particular carrier protein has been used as a constituent of licensed conjugate vaccines against *Neisseria meningitidis*, *Streptococcus pneumoniae* and *Haemophilus influenzae* type B.<sup>46</sup> The outer core pentasaccharide–CRM<sub>197</sub> glycoconjugate was prepared as illustrated in Scheme 3. The primary amine in the propyl amino linker of pentasaccharide 1 was treated with an excess of di–N–hydroxysuccinimidyl suberate in the presence of triethylamine and DMF to afford activated monoester.<sup>47</sup> The modified pentasaccharide with an activated ester was then conjugated to amino groups of the protein (the molar ratio: carbohydrate/protein 50:1) in PBS buffer (pH 7.4) to give the glycoprotein.<sup>48</sup> This glycoprotein was confirmed by SDS–PAGE analysis, displaying a shift toward a higher mass of glycoprotein.
compared with unconjugated CRM$_{197}$. The corresponding saccharide loading ratio was 17.5 by MALDI–TOF mass spectrometry analysis, and the conjugation efficiency was 35%. Based on the same procedure, another glycoprotein pentasaccharide–BSA was also prepared as a coating antigen to install the plates for reflecting specific binding between penatsaccharide moiety and the induced anti–pentasaccharide antibody in following ELISA assay.

Scheme 1.5. Preparation of the pentasaccharide-proteins conjugate

1.2.6 Immunological evaluation of the pentasaccharide-CRM$_{197}$ glycoconjugate

The resulting pentasaccharide–CRM$_{197}$ glycoconjugate was evaluated for its ability to elicit antibody responses to pentasaccharide 1-1 in a mouse model. Freund’s adjuvant was selected, because it is an effective adjuvant in mice that has been used to improve antibody responses to a synthetic oligosaccharide antigen. The initial immunization was performed with Freund’s complete adjuvant (FCA), and the second and third immunizations were performed with Freund’s incomplete adjuvant (FIA). Female BALB/c mice were subcutaneously immunized three times at biweekly intervals with 2.5 μg carbohydrate based doses of pentasaccharide–CRM$_{197}$ conjugate formulated with
adjuvant (FCA/FIA). The same dose of CRM$_{197}$ formulated with adjuvant (FCA/FIA), adjuvant (FCA/FIA), and only PBS were used as negative controls. The serum antibodies of immunized mice against pentasaccharide were measured by ELISA assay (Figure 1.5). The total IgG titers were significantly increased in the serum of mice immunized with pentasaccharide–CRM$_{197}$ compared with the other immunized groups (P<0.001). Moreover, higher IgM titers were also detected in the serum of mice immunized with pentasaccharide–CRM$_{197}$ compared with the other immunized groups (P<0.01). These results indicated that the pentasaccharide–CRM$_{197}$ glycoconjugate effectively elicited humoral response in this mouse model.

![Figure 1.5. Evaluation of anti–pentasaccharide antibody titers (IgG and IgM) after the third immunization.](image)

In order to illuminate the nature of immune response produced in mice immunized with pentasaccharide–CRM$_{197}$ and its potential relevance to overall T cell phenotype, IgG subclass profiles were evaluated by ELISA assay (Figure 1.6). In pentasaccharide–CRM$_{197}$ immunized mice, the antibody titers of IgG1 and IgG2b were dramatically increased (Figure 1.6A & C), indicating a Th2–type response. Moreover, predominantly higher titers of IgG2a were produced in serum of the immunized mice with
pentasaccharide–CRM197 compared with other groups (Figure 4B), which indicated that a Th1–type response was evoked in mice immunized with pentasaccharide–CRM197. In addition, an increase in IgG3 titers in group of pentasaccharide–CRM197 was also observed (Figure 1.6D), which is correlated with a Th1–type response. The IgG subclass profiles indicated that not only a Th2–type response but also a Th1–type response was induced in mice immunized with pentasaccharide–CRM197.

![Figure 1.6. Antibody IgG subclass profiles after the third immunization, (A) IgG1 titers, (B) IgG2a titers, (C) IgG2b titers, (D) IgG3 titers.]

### 1.2.7 Bactericidal activity

To evaluate the bactericidal activity induced by pentasaccharide–CRM197 conjugate against an *E. coli* strain containing R3 outer core structure, diluted mouse serum samples were incubated with *E. coli* O157:H7 in rabbit sera and then
developed by Cell Counting Kit–8 (CCK–8), which allows sensitive colorimetric assays for the determination of cell viability. An approximate 50% killing of *E. coli* O157:H7 was observed with a 160–fold dilution of the serum from mice immunized with pentasaccharide–CRM197, while 50% bactericidal activity was not achieved even with only a 20–fold dilution of the serum from the other groups. When the serum was diluted 80–fold from mice immunized with pentasaccharide–CRM197, about 80% of bacteria were killed (Figure 1.7). These results indicated that the serum from mice immunized with pentasaccharide–CRM197 showed remarkable bactericidal activity against *E. coli* O157:H7.

![Figure 1.7. Bactericidal activity of serum from the immunized mice against *E. coli* O157:H7.](image)

1.3 Conclusion

In summary, we described the first total synthesis of the outer core pentasaccharide of *E. coli* R3 using a [4 + 1] coupling strategy. The orthogonal protecting groups modified galactose building block (B) made it possible to establish the optimal order of glycosylation for synthesis of the highly crowded 1,2,3–*cis* configured oligosaccharide. Furthermore, solvent effects and temperature were exploited to control the anomic
selectivity of glycosylation. These strategies will be significant for preparation of other highly branched oligosaccharides. The immunological evaluation of the pentasaccharide–CRM\textsubscript{197} glycoconjugate indicated that this glycoconjugate was able to elicit specific anti-pentasaccharide antibodies with \textit{in vitro} bactericidal activity against \textit{E. coli} O157:H7. Overall, this work represented a new perspective in the design and synthesis of carbohydrate antigens to be explored as vaccine candidates.

1.4 Experimental Section

All reagents were purchased from commercially sources and were used without further purification. All solvents were available with commercially dried or freshly dried and distilled prior to use. Reactions were monitored by thin layer chromatography (TLC) using silica gel GF\textsubscript{254} plates with detection by short wave UV light ($\lambda = 254$ nm) and staining with 10\% phosphomolybdic acid in EtOH or $p$–anisaldehyde solution (ethanol/$p$–anisaldehyde/acetic acid/sulfuric acid 135:5:4:1.5), followed by heating on a hot plate. Column chromatography was conducted by silica gel (200–300 mesh) with ethyl acetate and hexane as eluent. Optical rotation values were measured using a JASCO DIP-360 polarimeter at the ambient temperature in specified solvents. $^1$H NMR and $^{13}$C NMR were recorded with Bruker AV 400 spectrometer at 400 MHz ($^1$H NMR), 100 MHz ($^{13}$C NMR) using CDCl\textsubscript{3} and CD\textsubscript{3}OD as solvents. Chemical shifts were reported in $\delta$ (ppm) from CDCl\textsubscript{3} (7.26 ppm for $^1$H NMR, 77.00 ppm for $^{13}$C NMR), CD\textsubscript{3}OD (3.31 ppm for $^1$H NMR, 49.00 ppm for $^{13}$C NMR). Coupling constants were reported in hertz. High-resolution mass spectra (HRMS) were obtained on a Varian QFT–ESI mass spectrometer, and glycoproteins were analyzed by Bruker ultrafleXtreme MALDI TOF/TOF mass spectrometer.

A mixture of donor 1-2 (1.39 g, 2.03 mmol) and acceptor 1-3 (610 mg, 1.36 mmol) and 4 Å molecular sieves (2.0 g) in 10 mL dry Et₂O/CH₂Cl₂ (1:1) was stirred at room temperature under Ar for 30 min. Then this mixture was cooled –40 °C, and TMSOTf (35 µL, 0.20 mmol) was added. The reaction was slowly warmed to room temperature in 1 h, TLC analysis showed complete conversion of starting material to a major product (hexane/ethyl acetate 5:1, \( R_f = 0.36 \)). The reaction was quenched by the addition of 0.1 mL triethylamine and filtered. The filtrate was concentrated in vacuum and purified by silica gel chromatography (hexane/ethyl acetate 10:1) to afford a syrup 1-4 (965 mg) contaminated by the rearrangement product of trichloroacetimidate donor 1-2. ESI HRMS: m/z calcd for C₆₀H₆₄NO₁₀S [M +NH₄]⁺ 990.4251, found 990.4254.

This disaccharide syrup 1-4 (965 mg, 0.99 mmol) was dissolved in 10 mL CH₂Cl₂, 1 mL trifluoroacetic acid and 0.1 mL water was added. The mixture was stirred for 30 min at room temperature. TLC analysis showed complete conversion of starting material to a major product (hexane/ethyl acetate 1:1, \( R_f = 0.42 \)). The reaction was quenched by the addition of 1.0 mL triethylamine. The mixture was washed with water, the organic layer was dried (Na₂SO₄) and filtered. The filtrate was concentrated in vacuum and purified by silica gel chromatography (hexane/ethyl acetate 2:1) to afford corresponding diol 1-5 as syrup (698 mg). ESI HRMS: m/z calcd for C₅₃H₅₆O₁₀SNa [M +Na]⁺ 907.3492, found 907.3478.

The above isolated diol (698 mg, 0.79 mmol) was dissolved in 10 mL DMF, and benzyl bromide (0.28 mL, 2.36 mmol) was added. Sodium hydride (60% dispersion in mineral
oil, 94 mg, 2.36 mmol) was then added slowly at 0 °C. The reaction mixture was stirred at room temperature overnight. TLC analysis showed complete conversion of starting material to a major product (hexane/ethyl acetate 8:1, $R_f = 0.17$). The reaction was quenched with MeOH and concentrated in vacuum. The residue was dissolved with CH$_2$Cl$_2$, and the organic layer was washed with water, followed by drying and filtration. The filtrate was concentrated in vacuum and purified by silica gel chromatography (hexane/ethyl acetate 8:1) to afford a purified syrup 1-6 (768 mg, 53% from 1-3, three steps). [α]$_D^{20} = +61.5$ (c = 0.2, CHCl$_3$); $^1$H NMR (CDCl$_3$, 400 MHz): δ 3.14 (dd, $J = 2.0$ Hz, $J = 10.4$ Hz, 1 H), 3.25 (d, $J = 10.0$ Hz, 1 H), 3.53–3.55 (m, 1 H), 3.64–3.78 (m, 6 H), 3.92–4.01 (m, 2 H), 4.09 (d, $J = 10.4$ Hz, 1 H), 4.24 (d, $J = 12.0$ Hz, 1 H), 4.36 (d, $J = 11.2$ Hz, 1 H), 4.48–4.60 (m, 4 H), 4.73 (d, $J = 11.6$ Hz, 1 H), 4.76–4.83 (m, 4 H), 4.87 (d, $J = 9.6$ Hz, 1 H, H–1$^C$), 4.92 (d, $J = 11.6$ Hz, 1 H), 4.76–4.83 (m, 2 H), 5.94 (d, $J = 3.6$ Hz, 1 H, H–1$^E$), 7.00–7.05 (m, 4 H), 7.10–7.14 (m, 3 H), 7.17–7.32 (m, 31 H), 7.50–7.52 (m, 2 H); $^{13}$C NMR (CDCl$_3$, 100 MHz): δ 67.75, 68.73, 70.08, 73.05, 73.21, 73.30, 74.51, 74.65, 74.75, 75.43, 75.52 77.45, 78.52, 78.77, 79.61, 81.67, 84.62, 86.86 (C–1$^C$), 95.61(C–1$^E$), 127.12, 127.18, 127.22, 127.33, 127.39, 127.42, 127.46, 127.57, 127.67, 127.76, 127.87, 127.94, 127.99, 128.09, 128.17, 128.26, 128.81, 131.11, 133.34, 137.66, 137.70, 137.74, 137.87, 138.02, 138.61, 138.72 ESI HRMS: m/z calcd for C$_{67}$H$_{72}$NO$_{10}$S [M +NH$_4$]$^+$ 1082.4877, found 1082.4865.


benzyl–β–D–glucopyranosyl trichloroacetimidate (1-7)
To a solution of \textbf{1-6} (750 mg, 0.7 mmol) in acetone (10 mL) and water (1.0 mL) was added NBS (250 mg, 1.4 mmol) at 0 °C. The mixture was stirred for 3 h at room temperature. TLC analysis showed complete conversion of starting material to a major product (hexane/ethyl acetate 3:1, \( R_f = 0.25 \)). The reaction was quenched by the addition of 10 mL Na\(_2\)S\(_2\)O\(_3\) (sat.) solution. The acetone was removed in vacuum, the crude product was diluted with CH\(_2\)Cl\(_2\) (30 mL), washed with H\(_2\)O (10 mL). And the aqueous phase was extracted by CH\(_2\)Cl\(_2\) (20 mL), the organic fractions were combined, dried over Na\(_2\)SO\(_4\) and filtered. The filtrate was concentrated in vacuum and purified by silica gel chromatography (hexane/ethyl acetate 3:1) to afford corresponding hemiacetal as syrup (540 mg). ESI HRMS: m/z calcd for C\(_{61}\)H\(_{68}\)NO\(_{11}\) [M +NH\(_4\)]\(^+\) 990.4792, found 990.4805.

A mixture of the above isolated hemiacetal (540 mg, 0.55 mmol), CCl\(_3\)CN (1 mL) and DBU (0.1 mL) in anhydrous CH\(_2\)Cl\(_2\) (8 mL) was stirred for 2 h at room temperature. TLC analysis showed complete conversion of starting material to a major product (hexane/ethyl acetate 3:1, \( R_f = 0.70 \)). The mixture was concentrated in vacuum and purified by silica gel chromatography (hexane/ethyl acetate 5:1) to afford a syrup \textbf{1-7} (532 mg, 68%, two steps) for next coupling step quickly.

\textbf{4-Methoxyphenyl 2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl–(1→2)–O-3,4,6-tri-O-benzyl-α-D-glucopyranosyl–(1→2)–O-benzoyl–4,6-O-benzylidene–β-D-galactopyranoside (1-9)}

A mixture of donor \textbf{1-7} (503 mg, 0.45 mmol) and acceptor \textbf{1-8} (258 mg, 0.54 mmol) and 4 Å molecular sieves (1.0 g) in 8 mL anhydrous Et\(_2\)O/CH\(_2\)Cl\(_2\) (1:1) was stirred at room temperature under Ar for 30 min. Then this mixture was cooled \(-20 \ ^\circ \text{C}\), TMSOTf (12 µL,
0.068 mmol) was added. The reaction was warmed slowly room temperature in 1 h, TLC analysis showed complete conversion of starting material to a major product (hexane/ethyl acetate 3:1, \( R_f = 0.25 \)). The reaction was quenched by the addition of 0.1 mL triethylamine and filtered. The filtrate was concentrated in vacuum and purified by silica gel chromatography (hexane/ethyl acetate 5:1) to afford a syrup trisaccharide 1-9 (354 mg, 55%). \([\alpha]_D^{20} = +83.5 \) (c = 0.2, CHCl\(_3\)); \(^1\)H NMR (CDCl\(_3\), 400 MHz): \( \delta \) 3.26 (dd, \( J = 3.6 \) Hz, \( J = 9.6 \) Hz, 1 H), 3.46 (dd, \( J = 3.2 \) Hz, \( J = 10.0 \) Hz, 1 H), 3.54–3.64 (m, 3 H), 3.67–3.70 (m, 2 H), 3.74 (s, 3 H, PhOCH\(_3\)), 3.86–3.87 (m, 2 H), 3.95–4.08 (m, 4 H), 4.18 (d, \( J = 12.4 \) Hz, 1 H), 4.23 (d, \( J = 11.6 \) Hz, 1 H), 4.31 (d, \( J = 11.6 \) Hz, 1 H), 4.40–4.43 (m, 3 H), 4.46–4.52 (m, 2 H), 4.57–4.66 (m, 5 H), 4.67–4.72 (m, 2 H, H–1\(^B\), PhCH\(_2\)), 4.74–4.86 (m, 4 H), 5.06 (d, \( J = 3.6 \) Hz, 1 H, H–1\(^C\)), 5.28 (dd, \( J = 3.6 \) Hz, \( J = 10.0 \) Hz, 1 H), 5.48 (s, 1 H, PhCH\(_3\)), 5.92 (d, \( J = 2.0 \) Hz, 1 H, H–1\(^E\)), 6.75–6.85 (m, 6 H), 6.94–6.98 (m, 2 H), 7.08–7.13 (m, 3 H), 7.17 (m, 5 H), 7.19–7.25 (m, 10 H), 7.27–7.36 (m, 14 H), 7.40–7.48 (m, 7 H), 8.01–8.03 (m, 2 H); \(^{13}\)C NMR (CDCl\(_3\), 100 MHz): \( \delta \) 55.55 (PhOCH\(_3\)), 65.76, 68.14, 68.34, 68.64, 69.10, 70.21, 70.33, 71.43, 73.07, 73.27, 73.31, 73.44, 73.73, 74.26, 74.78, 75.33, 76.13, 77.64, 79.45, 80.32, 81.69, 92.88 (C–1\(^C\)), 93.58 (C–1\(^E\)), 100.69 (PhCH\(_3\)), 101.22 (C–1\(^B\)), 114.49, 117.00, 126.17, 127.02, 127.08, 127.29, 127.38, 127.49, 127.51, 127.67, 127.75, 127.86, 127.95, 128.00, 128.11, 128.18, 128.28, 128.31, 128.63, 128.80, 129.55, 129.78, 133.18, 137.63, 137.99, 138.15, 138.21, 138.53, 138.68, 138.72, 139.32, 150.37, 154.99, 166.22 (PhCO\(_2\)); ESI HRMS: m/z calcd for C\(_{88}H_{88}O_{18}\)Na [M +Na]\(^+\) 1455.5868, found 1455.5809.
4-Methoxyphenyl 2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→2)-O-3,4,6-tri-O-benzyl-α-D-glucopyranosyl-(1→2)-O-4,6-benzylidene-β-D-galactopyranoside (1-10)

A solution of trisaccharide (215 mg, 0.15 mmol) in dry methanol (4 mL) was added sodium methoxide (10 mg, 0.18 mmol). The solution was stirred at 45 °C overnight. TLC analysis showed complete conversion of starting material to a major product (hexane/ethyl acetate 2:1, RF = 0.23). Then the acid resin (Dowex® 50WX2 H+–form) was added and stirred to adjust pH 7, followed by filtration. The filtrate was concentrated in vacuum and purified by silica gel chromatography (hexane/ethyl acetate 2:1) to afford a syrup 1-10 (191 mg, 96%). [α]_D⁰₂₀ = +75.0 (c = 0.2, CHCl₃); ¹H NMR (CDCl₃, 400 MHz): δ 3.13 (s, br, 1 H, OH), 3.23–3.26 (m, 2 H), 3.42–3.49 (m, 2 H), 3.63–3.78 (m, 9 H), 3.95 (dd, J = 4.0 Hz, J = 9.6 Hz, 1 H), 4.01–4.08 (m, 3 H), 4.11 (dd, J = 3.2 Hz, J = 9.2 Hz, 1 H), 4.16 (dd, J = 7.6 Hz, J = 9.6 Hz, 1 H), 4.23–4.28 (m, 3 H), 4.36 (d, J = 11.6 Hz, 1 H), 4.43–4.49 (m, 3 H), 4.54–4.61 (m, 2 H), 4.62–4.65 (m, 2 H, H–¹B, PhCH₂), 4.70 (d, J = 11.6 Hz), 4.75–4.79 (m, 2 H), 4.81–4.86 (m, 2 H), 4.98 (d, J = 10.4 Hz, 1 H), 5.22 (d, J = 3.2 Hz, 1 H, H–¹C), 5.55 (s, 1 H, PhCH), 5.79 (d, J = 4.0 Hz, 1 H, H–¹E), 6.75–6.77 (m, 2 H), 6.85–6.87 (m, 2 H), 7.01–7.05 (m, 2 H), 7.13–7.17 (m, 10 H) 7.23–7.34 (m, 26 H), 7.48–7.56 (m, 2 H); ¹³C NMR (CDCl₃, 100 MHz): δ 55.57 (PhOCH₃), 66.21, 68.60, 68.79, 70.02, 70.51, 71.38, 73.15, 73.23, 73.95, 74.00, 74.77, 75.34, 75.73, 76.24, 77.71, 78.15, 79.79, 81.22, 81.58, 92.57 (C–¹C), 94.41 (C–¹E), 100.61 (PhCH), 101.37 (C–¹B), 114.51, 116.92, 126.32, 127.09, 127.30, 127.38, 127.49, 127.61, 127.67, 127.92, 128.08, 128.13, 128.15, 128.24, 128.30, 128.61, 129.04, 137.49, 137.87, 138.25, 138.35, 138.37, 138.55,
138.75, 139.15, 150.62, 154.94; ESI HRMS: m/z calcd for C₅₁H₄₁O₁₇Na [M +Na]^+ 1351.5606, found 1351.5599.

4--Methoxyphenyl 2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→2)O-3,4,6-tri-O-benzyl-α-D-glucopyranosyl-(1→2)[6-O-acetyl-2-azido-3,4-di-O-benzyl-2-deoxy-α-D-glucopyranosyl-(1→3)]-4,6-O-benzylidene-β-D-galactopyranoside (1-12)

A mixture of donor 1-11 (103 mg, 0.18 mmol) and acceptor 1-10 (160 mg, 0.12 mmol) and 4 Å molecular sieves (200 mg) in 8 mL anhydrous Et₂O/CH₂Cl₂ (1:1) was stirred at room temperature under Ar for 30 min. Then this mixture was cooled −20 °C, TMSOTf (3.3 µL, 0.018 mmol) was added. The reaction was slowly warmed room temperature in 2 h, TLC analysis showed complete conversion of starting material to a major product (hexane/ethyl acetate 2:1, Rᵢ = 0.18). The reaction was quenched by the addition of 0.1 mL triethylamine and filtered. The filtrate was concentrated in vacuum and purified by silica gel chromatography (hexane/ethyl acetate 2:1) to afford as a syrup tetrasaccharide 1-12 (158 mg, 76%). [α]D²⁰ = +86.0 (c = 0.2, CHCl₃); ¹H NMR (CDCl₃, 400 MHz): δ 1.95 (s, 3 H, CH₃CO), 3.12–3.35 (m, 2 H), 3.49 (dd, J = 3.2 Hz, J = 10.4 Hz, 1 H), 3.57 (dd, J = 2.4 Hz, J = 9.2 Hz, 1 H), 3.63–3.69 (m, 2 H), 3.74–3.76 (m, 4 H), 3.81–3.87 (m, 2 H), 3.88–3.93 (m, 2 H), 3.98–3.99 (m, 2 H), 4.02–4.08 (m, 3 H), 4.12 (dd, J = 4.4 Hz, J = 10.4 Hz, 1 H), 4.25–4.28 (m, 2 H), 4.31–4.35 (m, 2 H), 4.38–4.44 (m, 2 H), 4.45–4.50 (m, 3 H), 4.53–4.57 (m, 5 H), 4.59–4.64(m, 4 H, H–1B, PhCH₂), 4.75–4.81 (m, 6 H), 4.84–4.89 (m, 2 H), 5.02 (d, J = 3.2 Hz, 1 H, H–1D), 5.25 (d, J = 3.6 Hz, 1 H, H–1C), 5.56 (s, 1 H, PhCH), 6.06 (s, 1 H, H–1E), 6.76–6.78 (m, 2 H), 6.83–6.85 (m, 2 H), 6.97–
7.00 (m, 6 H), 7.08–7.13 (m, 10 H), 7.15–7.25 (m, 15 H), 7.27–7.33 (m, 17 H), 7.54–7.56 (m, 2 H); $^{13}$C NMR (CDCl$_3$, 100 MHz): δ 20.79 (CH$_3$CO), 55.56 (PhOCH$_3$), 62.19, 63.20, 65.85, 68.31, 68.47, 68.96, 69.66, 70.07, 70.27, 70.43, 71.68, 73.27, 73.49, 74.43, 74.80, 75.05, 75.11, 75.35, 75.94, 77.22, 77.74, 78.10, 79.33, 80.17, 81.19, 81.83, 93.01 (C–1$^E$), 93.15 (C–1$^C$), 93.26 (C–1$^D$), 100.58 (PhCH), 101.71 (C–1$^B$), 114.55, 117.16, 126.01, 127.09, 127.18, 127.27, 127.31, 127.37, 127.43, 127.47, 127.62, 127.73, 127.82, 127.90, 127.96, 128.09, 128.17, 128.24, 128.36, 128.49, 128.75, 137.43, 137.49, 137.84, 138.16, 138.24, 138.44, 138.75, 138.77, 139.05, 150.46, 155.08, 170.20 (CH$_3$CO); ESI HRMS: m/z calcd for C$_{103}$H$_{107}$N$_3$O$_{22}$Na [M + Na]$^+$ 1760.7244, found 1760.7251.


To a stirred solution of tetrasaccharide 1-12 (140 mg, 0.08 mmol) in 5 mL DCM/MeOH (1:1.5) was cooled at 0 ºC, NiCl$_2$·6H$_2$O (114 mg, 0.48 mmol) and NaBH$_4$ (30 mg, 0.8 mmol) was added. The color of the solution changed from green to black. The reaction mixture was maintained at 10 ºC for 2 h, TLC analysis showed the starting material disappeared, followed by adding Ac$_2$O (0.1 mL) at 10 ºC. The reaction mixture was stirred at 10 ºC for 1 h, TLC analysis showed a major product (hexane/ethyl acetate 1:2, $R_f$ = 0.42), then the reaction was quenched with 0.1 mL triethylamine. The mixture was concentrated in vacuo to give a residue, which was purified by silica gel chromatography (hexane/ethyl acetate 2:3) to afford a syrup 1-13 (132 mg, 94%). [$\alpha$]$_{D}^{20}$ = +83.8 (c = 0.16, CHCl$_3$); $^1$H NMR (CDCl$_3$, 400 MHz): δ 1.33 (s, 3 H, NHCOCH$_3$), 2.01
(s, 3 H, $CH_3CO$), 3.23 (dd, $J = 3.6$ Hz, $J = 9.6$ Hz, 1 H), 3.29 (dd, $J = 3.6$ Hz, $J = 10.8$ Hz, 1 H), 3.56–3.66 (m, 3 H), 3.73 (s, 3 H, PhOCH$_3$), 3.76–3.82 (m, 2 H), 3.89–4.03 (m, 7 H), 4.09–4.15 (m, 2 H), 4.22 (d, $J = 12.4$ Hz, 1 H), 4.32–4.33 (m, 1 H), 4.38–4.48 (m, 6 H), 4.51–4.58 (m, 6 H), 4.60–4.66 (m, 5 H, H–$1^B$, PhCH$_2$), 4.71–4.88 (m, 8 H), 5.07 (d, $J = 3.6$ Hz, 1 H, H–$1^D$), 5.09 (d, $J = 4.0$ Hz, 1 H, H–$1^C$), 5.42 (s, 1 H, PhCH), 5.99 (d, $J = 2.8$ Hz, 1 H, H–$1^E$), 6.54 (d, $J = 10.0$ Hz, 1 H, AcNH), 6.75–6.83 (m, 4 H), 6.95–7.07 (m, 9 H), 7.09–7.21 (m, 20 H), 7.25–7.30 (m, 17 H), 7.37–7.43 (m, 4 H); $^{13}$C NMR (CDCl$_3$, 100 MHz): $\delta$ 20.76 (NHCO$CH_3$), 22.20 ($CH_3CO$), 52.12, 55.42, 62.32, 65.71, 68.21, 68.41, 68.86, 69.79, 69.92, 71.67, 72.22, 72.36, 72.90, 73.02, 73.58, 74.75, 74.84, 74.87, 75.12, 76.07, 76.88, 77.00, 77.55, 77.81, 79.54, 81.29, 81.44, 81.50, 92.31 (C–$1^C$), 92.65 (C–$1^D$), 93.01 (C–$1^E$), 101.19 (C–$1^B$), 101.25 (PhCH), 114.44, 116.72, 126.26, 126.75, 127.19, 127.28, 127.37, 127.44, 127.52, 127.64, 127.89, 127.93, 128.00, 128.06, 128.11, 128.16, 128.24, 128.29, 129.02, 137.39, 137.61, 137.68, 137.76, 138.02, 138.11, 138.31, 138.34, 138.47, 139.01, 150.20, 154.96, 170.36 (NHCOCH$_3$), 170.46 (CH$_3$CO); ESI HRMS: m/z calcd for C$_{105}$H$_{111}$NO$_{23}$Na [M + Na]$^+$ 1776.7445, found 1776.7452.


To a solution of 1-13 (123 mg, 0.07 mmol) in 1:1.5:1 toluene–MeCN–water (7 ml) was added ceric ammonium nitrate (192 mg, 0.35 mmol) at 0 °C. The mixture was stirred vigorous and was warmed room temperature in 30 min. TLC analysis showed complete
conversion of starting material to a major product (hexane/ethyl acetate 1:2, \( R_f = 0.40 \)), then the mixture was diluted with ethyl acetate, washed with water, saturated NaHCO\(_3\) (aq), and brine, dried (Na\(_2\)SO\(_4\)) and filtered. The filtrate was concentrated in vacuum and purified by silica gel chromatography (hexane/ethyl acetate 2:3) to afford the corresponding hemiacetal (83 mg). ESI HRMS: m/z calcd for C\(_{98}\)H\(_{105}\)NO\(_{22}\)Na [M +Na]\(^+\) 1670.7026, found 1670.7032.

To a solution of the above isolated hemiacetal (83 mg, 0.05 mmol) in dry DCM was added 0.5 mL CCl\(_3\)CN and 0.05 mL DBU. The mixture solution was stirred at room temperature for 3 h under the Ar atmosphere. TLC analysis showed complete conversion of starting material to a major product (hexane/ethyl acetate 1:2, \( R_f = 0.74 \)). The mixture was concentrated in vacuum to give a residue, which was purified by silica gel chromatography (hexane/ethyl acetate 1:1) to afford a yellow syrup 1-14 (66 mg, 53% from 1-13) for next coupling step quickly.


A mixture of donor 1-14 (65 mg, 0.036 mmol) and acceptor 1-15 (28.8 mg, 0.054 mmol) and 4 Å molecular sieves (100 mg) in dry 6 mL Et\(_2\)O/CH\(_2\)Cl\(_2\) (1:1) was stirred at room temperature under Ar for 30 min. Then this mixture was cooled –25 °C, TMSOTf (1 \( \mu \)L, 0.0054 mmol) was added. The reaction was slowly warmed room temperature in 2 h, TLC analysis showed complete conversion of starting material to a major product (hexane/ethyl acetate 1:1, \( R_f = 0.37 \)). The reaction was quenched by the addition of 0.1
mL triethylamine and filtered. The filtrate was concentrated in vacuum and purified by silica gel chromatography (hexane/ethyl acetate 3:2) to afford a syrup 1-16 (33 mg, 42%). 

\[ \alpha \]$_{20}^{20} = +87.4 \ (c = 0.8, \text{CHCl}_3); \ \ \ ^1\text{H NMR } (\text{CDCl}_3, 400 \text{ MHz}): \ \delta \ 1.33 \ (s, 3 \text{ H}, \text{NHCOCH}_3), 1.72 \ (s, 3 \text{ H}, \text{CH}_3\text{CO}), 1.92 \ (m, 2 \text{ H}, \text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3), 3.32 \ (m, 2 \text{ H}), 3.43 \ (m, 3 \text{ H}), 3.62–3.71 \ (m, 8 \text{ H}), 3.76–3.80 \ (m, 6 \text{ H}), 3.91–4.08 \ (m, 7 \text{ H}), 4.10–4.14 \ (m, 2 \text{ H}), 4.23 \ (m, 3 \text{ H}), 4.33–4.38 \ (m, 5 \text{ H}), 4.43–4.90 \ (m, 3 \text{ H}), 4.54–4.60 \ (m, 5 \text{ H}), 4.66–4.76 \ (m, 6 \text{ H}), 4.83–4.87 \ (m, 5 \text{ H}), 4.89–4.98 \ (m, 4 \text{ H}), 5.16 \ (d, J = 3.6 \text{ Hz}, 1 \text{ H}, \text{H–1}^\text{D}), 5.22 \ (s, 1 \text{ H}, \text{PhCH}), 5.31 \ (d, J = 3.2 \text{ Hz}, 1 \text{ H}, \text{H–1}^\text{B}), 5.51–5.58 \ (m, 2 \text{ H}, \text{H–1}^\text{C} \text{ and AcNH}), 5.91 \ (s, 1 \text{ H}, \text{H–1}^\text{E}), 6.97–7.00 \ (m, 2 \text{ H}), 7.11–7.22 \ (m, 24 \text{ H}), 7.29–7.42 \ (m, 37 \text{ H}), 7.57 \ (m, 2 \text{ H}); \ \ ^{13}\text{C NMR } (\text{CDCl}_3, 100 \text{ MHz}): \ \delta 20.69 \ (\text{NHCOCH}_3), 22.58 \ (\text{CH}_3\text{CO}), 28.79 \ (\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3), 48.20 \ (\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3), 52.30, 61.43, 62.02, 64.57, 68.48, 68.88, 69.27, 69.57, 70.00, 70.12, 70.47, 71.73, 72.62, 73.00, 73.44, 74.93, 75.00, 75.20, 75.47, 76.15, 77.26, 77.73, 78.89, 79.09, 80.60, 81.37, 81.82, 93.58 \ (\text{C–1}^\text{C}), 93.71 \ (\text{C–1}^\text{B}), 93.78 \ (\text{C–1}^\text{D}), 95.87 \ (\text{C–1}^\text{E}), 95.88 \ (\text{C–1}^\text{A}), 100.98 \ (\text{PhCH}), 126.23, 127.19, 127.31, 127.35, 127.44, 127.48, 127.53, 127.64, 127.75, 127.86, 128.08, 128.12, 128.17, 128.23, 128.26, 128.31, 128.37, 128.44, 128.57, 128.73, 129.07, 137.72, 137.77, 137.88, 138.13, 138.23, 138.29, 138.40, 138.84, 139.22, 170.04 \ (\text{NHCOCH}_3), 170.08 \ (\text{CH}_3\text{CO}); \ \text{ESI HRMS: m/z } \text{calcd for } \text{C}_{128}\text{H}_{139}\text{N}_4\text{O}_{27} \ [\text{M +H}]^+ 2163.9627, \text{found } 2163.9648.


A solution of 16 (22 mg, 0.01 mmol) in dry methanol (2 mL) was added catalytic amount of sodium methoxide (pH 9–10). The solution was stirred at room temperature for
2 h. TLC analysis showed complete conversion of starting material to a major product (hexane/ethyl acetate 1:2 \( R_f = 0.33 \)). Then the acid resin (Dowex® 50WX2 H\(^+\)-form) was added and stirred to adjust pH 7, followed by filtration. The filtrate was concentrated in vacuum to afford a syrup (16 mg). ESI HRMS: m/z calcd for C\(_{126}\)H\(_{137}\)N\(_4\)O\(_{26}\) [M +H\(^+\)]\(^+\) 2121.9521, found 2121.9507.

To a stirred solution of the above syrup (16 mg, 0.0075 mmol) in 2 mL methanol was added Pd(OH)\(_2\)/C (20 mg) at 50 psi H\(_2\) atmosphere for 3 days. The catalyst was filtered off, the filtrate was concentrated to afford a white solid \( \textbf{1-16} \) (6.7 mg, 72% from \( \textbf{1-16} \)). \( \text{[\(\alpha\)]}_D^{20} = +178.0 \) (c = 0.2, H\(_2\)O); \(^1\)H NMR (CD\(_3\)OD, 400 MHz): \( \delta \) 1.97–2.00 (m, 2 H, OCH\(_2\)CH\(_2\)CH\(_2\)N\(_3\)), 2.02 (s, 3 H, NHCOCH\(_3\)), 3.15–3.20 (m, 2 H), 3.35–3.44 (m, 6 H), 3.58–3.61 (m, 3 H), 3.68–3.77 (m, 10 H), 3.85–3.87 (m, 2 H), 3.91–3.99 (m, 6 H), 4.09 (m, 2 H), 4.25–4.30 (m, 3 H), 5.04–5.05 (m, 2 H, H–1\(^A\) and H–1\(^D\)), 5.17 (d, \( J = 3.6 \) Hz, 1 H, H–1\(^B\)), 5.47 (d, \( J = 3.2 \) Hz, 1 H, H–1\(^C\)), 5.79 (s, H–1\(^E\)), \(^{13}\)C NMR (CD\(_3\)OD, 100 MHz): \( \delta \) 20.43 (CH\(_3\)CO), 28.34 (OCH\(_2\)CH\(_2\)CH\(_2\)N\(_3\)), 37.91 (OCH\(_2\)CH\(_2\)CH\(_2\)N\(_3\)), 52.91, 59.42, 60.10, 60.56, 60.73, 64.29, 66.30, 66.83, 68.93, 69.01, 69.18, 69.60, 70.23, 70.40, 70.55, 71.03, 71.21, 71.28, 71.49, 71.66, 72.18, 72.57, 75.15, 76.84, 90.77 (C–1\(^C\)), 91.96 (C–1\(^B\)), 93.67 (C–1\(^D\)), 95.98 (C–1\(^E\)), 98.76 (C–1\(^A\)), 167.49 (CH\(_3\)CO); ESI HRMS: m/z calcd for C\(_{35}\)H\(_{62}\)N\(_2\)O\(_{26}\) Na [M +Na\(^+\)] \(^+\) 949.3489, found 949.3484.

**Preparation of NHS activated pentasaccharide**

A solution of the propyl amino–linked pentasaccharide \( \textbf{1-1} \) (2.0 mg, 0.002 mmol) in DMF (0.5 mL) containing triethylamine (50 µL), was added dropwise to a stirred solution of di–N–hydroxysuccinimidy suberate (7.4 mg, 0.02 mmol) in DMF (0.5 mL). The reaction was kept under gentle stirring at room temperature for 3 h. TLC analysis showed complete
conversion of starting material to a major product (ethyl acetate/methanol/water/acetic acid 5:1.5:0.75:0.15, \( R_f = 0.50 \)). The reaction was then concentrated, the residue was added 0.5 mL water, extracted with EtOAc (1 mL x 4). The aqueous layer was lyophilized to give NHS activated pentasaccharide as a white power (2.0 mg). ESI HRMS: m/z calcd for \( C_{47}H_{77}N_3O_{31}Na \) [M +Na]+ 1202.4439, found 1202.4453.

**Conjugate NHS activated pentasaccharide to CRM\textsubscript{197} and BSA**

The NHS activated pentasaccharide was conjugated to CMR\textsubscript{197} (or BSA) at a molar ratio 50:1 in 3 x PBS buffer (pH 7.4). The solution was incubated overnight at room temperature. Then the resultant solution was ultrafiltrated and washed with 1 x PBS buffer using Amicon Centrifugal Filter Devices (Ultrace 10, 000). The glycoproteins solution was lyophilized to give a white solid. Glycoproteins were analyzed by Bruker ultrafleXtreme MALDI TOF/TOF mass spectrometer.

**SDS–PAGE**

The glycoproteins (5 µg) and CRM\textsubscript{197} (or BSA) (5 µg) were suspended in 10 µL of sample buffer (5% (w/v) SDS, 10% (v/v) glycerol, 25 mM Tris–HCl, pH 6.8, 10 mM DTT, 0.01% (w/v) bromophenol blue), loaded into different lanes of a 1.5–mm–thick, 12% (w/v) SDS–PAGE gel, and visualized by Coomassie Brilliant Blue R–250 staining.

**Immunization of mice**

The mice were purchased from Vital River Laboratories. Groups of 5 female BALB/c mice (6–8 week old) were subcutaneously (several different sites with a total of
150 µL) immunized on days 0, 14, and 28 with 2.5 µg carbohydrate antigen based doses of pentasaccharide–CRM<sub>197</sub> conjugate formulated with Freund’s adjuvant (mixing equal volume of conjugate and Freund’s adjuvant, v/v, 1:1). PBS, Freund’s adjuvant and CRM<sub>197</sub>/Freund’s adjuvant (v/v, 1:1) were used as controlled groups. For immunizations with Freund’s adjuvant, the first immunization was performed with Freund’s complete adjuvant (FCA) (Sigma–Aldrich, F5881), and the second and third immunizations with Freund’s incomplete adjuvant (FIA) (Sigma–Aldrich, F5506). Seven days after the third immunization, the blood was taken via lateral saphenous, and then centrifuged for 20 min at 4,000 RPM to collect the sera (without anticoagulants). These collected sera were ready to the following ELISA assay and bactericidal activity assay.

**ELISA assay**

The ELISA 96–well plates (Corning®, #3590) were coated with 1 µg/mL pentasaccharide–BSA in 1 x PBS buffer (pH 7.4) overnight at 4 °C. The coated plates were washed three times with PBS buffer containing 0.05% (v/v) Tween 20 (PBST) (pH 7.4), and then blocked for 2 hours at room temperature with 2% BSA (w/v) in PBST. The coated plates were incubated with 100 µL of sera diluted with dilution buffer (1% BSA, PBST) 2–fold from 1:200 to 204,800 for 2 hours at room temperature after being washed three times with PBST. Then quintuple PBST washing was performed, and 100 µL/well of 1:3000–1:20000 HRP–conjugated goat anti-mouse IgG (Invitrogen, USA), IgG1, IgG2a, IgG2b, IgG3 and IgM (Abcam) were added respectively, and incubated 1 hour at room temperature. After the plates were washed with PBST, 100 µL/well of enzyme substrate tetramethylbenzidine (TMB) solution was added and incubated for 15 min (a blue color
developed for the sera with antibodies) before the enzymatic reaction was terminated by adding 1 M HCl, and the OD of each sample was measured at 450 nm with a microplate spectrophotometer (Biorad). Antibody titers were expressed by evaluating the logarithm of the reciprocal of sera dilution based two. The statistical and graphical analysis was performed using GraphPad Prism 5 software.

Bactericidal activity

Complement-dependent bactericidal activity was measured in serum of day 35 (one week after the third immunization). *E. coli* O157:H7 containing R3 outer core was used as the test strain. 104 bacteria per mL suspended in 1% peptone supplemented with 5% 3-week-old rabbit sera as the complement source were added to 96-well plates (Costar®, # 3590), and mixed with serum diluted with 2-folds PBS (pH 7.4) from 1:20 to 1:12,800. The mixture was incubated at 37°C for 1 h. 10 μL/well of the CCK8 was added, and plates were incubated for 6 h. The absorbance of plates at 450 nm was read by a microplate spectrophotometer (Biorad). Percent bactericidal activity was calculated as the proportion of the dead bacteria exposed to immunized–serum compared with the number of bacteria in control culture condition.

Statistical analysis

All the statistical analyses and figures were generated by GRAPHPAD PRISM software version 5.0. Data were shown as means ± standard deviation (SD). The difference between two groups was compared by *t* test. For multiple comparisons, one–
way ANOVA was used. A probability (P) value ≤ 0.05 was considered statistically significant.


1.5 References


34. Demchenko, A.; Stauch, T.; Boons, G. J., Solvent and other effects on the stereoselectivity of thioglycoside glycosidations. *Synlett* 1997, 818-.


2. Chemoenzymatic Synthesis of a library of Human Milk Oligosaccharides

2.1 Introduction

2.1.1 History of HMOs

At the end of the 19th century, the overall infant first-year mortality rates were as high as 30%. Surprisingly, it was observed that breast-fed infants had a much lower chance of death and had lower incidences of infectious diarrhea and many other diseases than “bottle-fed” infants. At that time, Escherich, an Austrian pediatrician and microbiologist, had discovered a relationship between intestinal bacteria and human milk, which components remained unknown until more than half a century later. In 1926, Schönfeld had reported that the whey fraction of human milk contains a growth-promoting factor for *Lactobacillus bifidus* (later reclassified as *Bifidobacterium bifidus*).\(^1\) In 1954, Kuhn and György confirmed that the chemical nature of “bifidus factor” is oligosaccharides.\(^2,3\) During the following fifty years, more than 100 HMOs were isolated and characterized by scientists.

2.1.2 Structures of HMOs

Human milk oligosaccharides (HMOs) are the third major component in human milk, only after lactose and lipid.\(^4,6\) Concentrations and components of HMOs vary depending on the stages of lactation.\(^1,7\) Particularly, one liter mature human milk contains about 12-14 grams HMOs.\(^8,14\) The structures of about 130 discovered HMOs have been elucidated.\(^15-17\) The major building blocks of HMOs are 5 monosaccharides, including D-glucose (Glc), D-galactose (Gal), N-acetylglucosamine (GlcNAc), L-Fucose (Fuc), and N-acetyl neuraminic acid (Neu5Ac). All HMOs contain lactose (Galβ1,4Glc) at their reducing
end, which can be extended by the addition of β1-3 or β1-6-linked lacto-N-biose (Galβ1-3GlcNAc) or N-acetyllactosamine (Galβ1-4GlcNAc). A β1-3 linkage generates linear HMOs. A β1-6 linkage between two disaccharide units introduces chain branching. Lactose or the elongated oligosaccharide chain can be sialylated in α 2-3 or α 2-6 linkage and/or fucosylated in α1-2, α1-3 or α1-4 linkage. (Figure 2.1)

**Figure 2.1. Representative fucosylated or sialylated HMOs**

Even though HMOs were first discovered and confirmed in 1950s, a comprehensive understanding of their functions is still out of reach, due to their inherit diversity and complexity. Increasing evidences show that HMOs can provide significant beneficial effects to the health of breast-fed infants through several mechanisms. For instance, HMOs could serve as prebiotics to promote the growth of desired bacteria in infant’s intestine.\(^1, 6, 18-21\) Besides, HMOs are antiadhesive antimicrobials by serving as a receptor to prevent pathogen attachment to infant mucosal surfaces.\(^22-26\) In addition, evidence has
demonstrated that HMOs can modulate epithelial and immune cell responses, and reduce excessive mucosal leukocyte infiltration and activation, which in turn decreases the risk of necrotizing enterocolitis (NEC), one of the most common fatal disorders in preterm infants.\textsuperscript{27, 28} Furthermore, sialylated HMOs may also provide necessary nutrients for the development of brain and cognition of infants.\textsuperscript{29-32}

2.1.3 Previous Synthesis of HMOs

Even though the general functions of HMOs has been explored and discovered, the functional roles of individual HMOs are far less clear because of very limited accesses towards sufficient amount of structurally defined HMOs. To date, only a handful of short-chain HMOs can be produced in large scale and the supply of more complicated and branched HMOs is highly demanded.

Until now, only a few approaches has been developed for the synthesis of a small number of well-defined HMOs.\textsuperscript{1, 33-42} For example, Schmidt developed a convergent strategy to synthesize highly branched structure of human milk octasaccharide 1 by solution-phase.\textsuperscript{36} In this study, target compound was assembled by six building blocks with consistent using trichloroacetimidate as glycosylation donors. The coupling efficiency is achieved due to the convergency as well as to several regioselective and stereoselective glycosylation steps. Based on their previous work, Schmidt developed more versatile methodologies to produce complex HMOs based on solid-phase synthesizer.\textsuperscript{39} First, they designed a cleavable linker attached to resin. Then sugar moiety and linker was coupled by trichloroacetimidate/TMSOTf condition, followed by sequential
adding of other building blocks to furnish the targeted hexasaccharide. Enzymatic methods have also been employed to achieve relatively simple structures.\textsuperscript{35}

One of the biggest roadblocks in previous synthesis remains to be small quantity and limited variety of HMOs needed for biofunctional studies. Recently, we have developed an efficient Core Synthesis/Enzymatic Extension (CSEE) strategy for rapid preparation of N-glycan libraries.\textsuperscript{43} In this study, similar strategy was successfully applied for HMOs synthesis. Briefly, 3 core oligosaccharides with one or two GlcNAc residue(s) at the non-reducing end were first synthesized by convergent assembly of 3 simple building blocks followed by extension of the cores by 4 robust glycosyltransferases to produce a library of 31 HMOs (Figure 2.2).

![Figure 2.2](image.png)

**Figure 2.2.** HMOs synthesized by the Core Synthesis/Enzymatic Extension (CSEE) strategy starting with 3 chemically prepared core structures (boxed).
2.2 Convergent Core Synthesis

Previous studies highlighted the complexity and challenges associated with synthesizing HMOs via block coupling strategy. Schmidt developed sequential synthesis of lactose-containing oligosaccharides, including HMO lacto-N-tetraoside based on Solid-Phase Synthesis concept.38 Madsen used one-pot glycosylations to achieve several human milk oligosaccharides.37 Both of the methods can synthesize linear and simple oligosaccharides with obvious limitations in achieving more complex HMOs, especially highly branched ones. In this study, we developed an efficient and versatile methodology that utilized oligosaccharyl thioethers and oligosaccharyl bromides as convergent donors for glycosylation, enabling branching assembly in one or two steps of glycosylations with excellent stereoselectivity and yields.

We envisaged that protected lactose 1 (Figure 2.3) would be a versatile precursor for the synthesis of core structures, including symmetric and asymmetric ones, as C4, C6-hydroxyl groups (OH) on Gal are protected by benzylidene and the C3-OH is unprotected for chemical glycosylation. In order to achieve the selective protection of building block 1, C3-OH group was selectively protected by 4-Methoxybenzyl ether (PMB) using standard condition, followed by C4, C6-OH protection with benzylidene group. In order to furnish the target cores, two oligosaccharyl thioethers and oligosaccharyl bromide were prepared (Figure 2.3).

![Figure 2.3](image_url). The three versatile building blocks for the assembly of the three core structures.
2.3 Synthesis of Building Blocks

The synthesis of precursor 2-1 began with lactose peracetate, which was converted to the β-lactoside 2-4 by reaction with benzyl alcohol in the presence of BF$_3$·Et$_2$O, followed by deacetylation with NaOMe/MeOH condition to furnish compound 2-5. Then 2-5 was treated with dibutyltin oxide, followed by the reaction with 4-Methoxybenzyl chloride to provide selective 3’-O-PMB protected lactoside with fair yield, which has been extensively studied$^{41, 44}$ The following Benzylidene protection on 4’, 6’-OH was conducted with Benzaldehyde dimethyl acetal, catalyzed by Camphorsulfonic acid (CSA), to get compound 2-6. The perbenzylation of the remaining hydroxyls of 2-6 was performed by using sodium hydride and benzyl bromide in anhydrous DMF to give compound 2-7. After removing PMB protecting group of 2-7 treated with 2, 3-Dichloro-5, 6-dicyano-1, 4-benzoquinone (DDQ), building block 2-1 was achieved (Scheme 2.1).

Lactosamine building block 2-3 was initially envisioned to be synthesized using a straightforward fashion by coupling monosaccharide 2-9 and 2-10 (Scheme 2.2A). Unfortunately, only minor desired product was isolated by silica gel chromatography, and substantial amount of byproduct 2-11 was generated through a thioether migration
Therefore, a Koenigs–Knorr reaction coupling approach was carried out by installing C4 hydroxy group of glucosamine with high yield (Scheme 2.2B). Building block 2-2 would be conveniently used in the synthesis of two asymmetric core oligosaccharides.

Scheme 2.2 Synthesis of Building Block 3.

2.4 Synthesis of Core Oligosaccharides

With all building blocks in hand, we began to assemble the three core oligosaccharides using our convergent strategy. The synthesis of HMO1 was depicted in Scheme 3. Disaccharide 2-13 was obtained in a good yield (85%) by selective opening of the benzylidene ring at C6 of building block 2-1 using Et3SiH/PhBCl2. Oligosaccharyl bromide 2-2 was freshly prepared using HBr/AcOH condition. Silver triflate (AgOTf) promoted glycosylation of 2-13 with donor 2-2 in dichloromethane resulted in the formation of tetrasaccharide 2-14 in an excellent yield of 85%. Deprotection of 2-14 with ethylene diamine, followed by treatment with acetic anhydride, furnished peracetylated tetrasaccharide 2-15 in 75% yield. O-Deacetylation of compound 2-15 was performed
under Zemplén conditions, followed by the global deprotection of Benzyl group (Bn) by catalytic hydrogenolysis with Pd(OH)$_2$/H$_2$ in MeOH/H$_2$O (10 : 1). The core structure HMO1 was produced in a total yield of 80% over the two steps.

Scheme 2.3 Synthesis of HMO1.

Glycosylation of 3'-O-unprotected acceptor 2-1 with donor 2-3 proceeded at -20°C under AgOTf/NIS conditions to furnish the desired tetrasaccharide 2-16 in 85% yield. Then selective opening of the benzylidene ring at C6 of 2-16 using Et$_3$SiH/PhBCl$_2$ condition provided 6'-O-unprotected acceptor 2-17 in 80% yield. The fully protected pentasaccharide was initially attempted to be synthesized by a convergent glycosylation of acceptor 2-17 and thiol donor 2-18. Unfortunately, the desired product was not detected by TLC and ESI mass spectrometry analysis (Scheme 2.4A). Several other donors including oligosaccharyl trichloroacetimidate and thioethers donors were tried to install the pentasaccharide, but no product was detected. The big challenge should be caused by the bulky benzyl group on 4'-OH position, which has very large steric hindrance and stops the glycosylation on 6'-OH position, even though primary alcohol is very nucleophilic. Therefore, 4', 6'-O-unprotected lacto-N-tetraose 2-19 was proposed as acceptor for the glycosylation. Removal of the 4', 6'-O-benzylidene of 2-16 by treatment
with ethanethiol in the presence of TsOH afforded acceptor $2-19$. Glycosylation of acceptor $2-19$ with glycosyl bromide $2-2$ to achieve the protected target pentasaccharide $20$ proceeded smoothly and regioselectively by use of AgOTf as Lewis acid at -20 °C in an excellent yield (85%). The two phthalimides of $2-20$ were then converted into acetamides $2-21$, followed by the global deprotection of Ac and Bn groups. The core oligosaccharide HMO2 was produced in a total yield of 53% over the four steps (Scheme 2.4B).

Scheme 2.4. Synthesis of HMO2.
AgOTf promoted Glycosylation of 3’-O-unprotected acceptor **2-1** with donor **2-2** proceeded at -20°C to furnish the trisaccharide **2-22** in a good yield of 85%. Then deprotection of the 4’, 6’-O-benzylidene of **2-22** by treatment of EtSH/TsOH provided the dialcohol **2-23**. Glycosylation of acceptor **2-23** with thiol donor **2-3** by treatment of AgOTf/NIS condition at -20°C to achieve the protected target pentasaccharide **2-24** in 70% yield. The two phthalimides of **2-24** were then converted into acetamides **2-25**. Complete deprotection of **2-25** was achieved by hydrogenolytic debenzylolation (Pd(OH)$_2$/C, H$_2$) and complete de-O-acetylation using sodium methoxide in methanol, resulting in core oligosaccharide **HMO3** in a total yield of 67% over the four steps (Scheme 2.5).

![Scheme 2.5 Synthesis of HMO3.](image)

**2.5 Enzymatic Extension of Core Structures**

A total of 31 HMOs were enzymatically synthesized starting from the 3 core structures (**HMO1**, **HMO2**, **HMO3**) via an enzymatic extension approach using 4 robust GTs: β1,4 galactosyl-transferase from *Neisseria meningitidis* (LgtB), $^4$6 α2,6-
sialyltransferase from *Photobacterium damselae* (Pd2,6ST), C-terminal 66 amino acids truncated α1,3-fucosyltransferase from Helicobacter pylori (Hpα1,3FT), and α1,2-fucosyltransferase from *Helicobacter mustelae* (Hmα1,2FT). All GTs were from bacteria and had high expression levels in *Escherichia coli*, high activity, and relatively relaxed substrate tolerance. As shown in Figure 2.4A, glycans HMO11-HMO16 were prepared starting with the chemically prepared core HMO1. Briefly, in a 2 mL reaction system, 30 mg of HMO1 (20 mM) was incubated with Gal (20 mM), MgCl₂ (20 mM), ATP (20mM), UTP (20mM) and variant amount of BiGalK, BiUSP and LgtB (Figure 2.4D). After overnight reaction, the mixture was terminated by boiling for 10min and analyzed by MALDI-MS which shows a single peak at m/z 1095.748, corresponding to HMO11 [M+Na]⁺. Meanwhile, on the HPLC-ELSD (Evaporative Light Scattering Detector) profile, a new peak (TR= 11.946 min) was observed. The reaction mixture was purified by HPLC using a water/acetonitrile gradient elution, yielding 40mg of HMO11 (93% yield). The purified HMO11 (99% pure) was then utilized for the syntheses of HMO12-HMO16 (Figure 2.4A) catalyzed by Pd2,6ST, and α1,2FT, α1,2FT respectively (see ESI for details). Interestingly α1,3FT can specifically distinguish the GlcNAc from terminal Galβ1,4GlcNAc and α1,2FT preferentially attach to terminal Gal. We basically use this feature to biosynthesize Lewis X (Le⁺) and Lewis Y (Le⁻). In addition, difucosylated LacNAc motif [Fucα1, 2-Gal-β1, 4-(Fucα1, 3-)-GlcNAc] was also generated while using Hmα1,2FT. The synthesis of asymmetric bi-antennary HMO2x and HMO3x (Figure 2.4B&C) were carried out by enzymatic extension of core HMO2 and HMO3. Asymmetric Core HMO2 and HMO3 can more efficiently take advantage of different substrate specificities of GTs over symmetric HMO1 via coupling a Gal to the terminal GlcNAc of one antennary. For
example, to obtain HMO311, the Gal from β6GlcNAc branch was sequentially extended by Hmα1,2FT, Hpα1,3FT and LgtB. (Figure 2.4C) In contrast, HMO310 were sequentially synthesized by Hmα1,2FT, LgtB and Hpα1,3FT (Figure 2.4C). Such synthetic routes were designed according to the substrate specificities of GTs to avoid undesirable glycosylation.

Figure 2.4. Enzymatic extension of human milk oligosaccharides. Conditions: (a) LgtB, BiGalK, AtUSP, Gal, ATP, UTP, Mg^{2+}; (b) Pd2,6ST, NmCSS, Neu5Ac, CTP, Mg^{2+}; (c) Pd2,6ST, NmCSS, Neu5Gc, CTP, Mg^{2+}; (d) Hpα1,3FT, GDP-Fuc, Mn^{2+}; (e) Hmα1,2FT, GDP-Fuc, Mn^{2+}. LgtB, Neisseria meningitidis β1,4 galactosyltransferase; BiGalk, Bifidobacterium infantis galactokinase; AtUSP, Arabidopsis thaliana pyrophosphorylase; Pd2,6ST, Photobacterium damselae α2,6-sialyltransferase; NmCSS, Neisseria meningitidis CMP-sialic acid synthetase; Hpα1,3FT, C-terminal 66 amino acids truncated Helicobacter pylori α1,3-fucosyltransferase; Hmα1,2FT, Helicobacter mustelae α1,2-fucosyltransferase.
2.6 Conclusion

In summary, we have utilized our well-developed CSEE strategy for efficient synthesis of a library of structure-defined HMOs, which was assisted with rapid HPLC purification. The combination of CSEE and HPLC purification allows us to deliver diverse and high purity of 31 homogenous HMOs. These HMOs are valuable materials for bioactivity evaluation as well as glycan analysis. In this work, oligosaccharyl thioethers and oligosaccharyl bromide were consistently utilized as chemical glycosylation donors for the convergent installation of branched lactose-terminated antennae. This general and efficient method furnished 3 core oligosaccharides with high stereoselectivity and excellent yields. This work further confirmed that any GlcNAc-terminated glycans could be extended to 5 or more glycans, including Le$^X$ and SLe$^X$, which are very important epitopes in glycobiology. The CSEE demonstrated a practical way to harvest diverse and complex HMOs with defined structures for various applications. The “mass” production of more homogenous HMOs and bioactivity evaluation are underway.

2.7 Experimental Section

**General Methods**: All chemicals were purchased as reagent grade and used without further purification. Anhydrous dichloromethane (CH$_2$Cl$_2$), acetonitrile (CH$_3$CN), tetrahydrofuran (THF), N, N-dimethyl formamide (DMF), toluene, and methanol (MeOH) were purchased from a commercial source without further distillation. Pulverized Molecular Sieves MS-4 Å (Aldrich) for glycosylation was activated by heating at 350 °C for 3 h. All reactions were performed with dry solvents and under nitrogen unless otherwise stated. Thin-layer chromatography (TLC) with 60 F$_{254}$ silica gel plastic plates was detected visualized under UV (254 nm) and/or by staining with a solution of 10 ml
anisaldehyde and 10 ml 95% H$_2$SO$_4$ in 400 mL Ethanol, followed by heating on a hot plate. Column chromatography was carried out on silica gel (EMD 230-400 mesh ASTM) and P2 gel (Biorad). Optical rotation values were measured using a PerkinElmer Model 343 polarimeter at the ambient temperature in specified solvents. $^1$H NMR spectra were recorded on a Bruker AVANCE 400 (400 MHz) or Bruker AVANCE 500 (500 MHz) spectrometer at 25 °C. All $^1$H Chemical shifts (in ppm) were assigned according to CDCl$_3$ ($\delta$ = 7.24 ppm) or D$_2$O ($\delta$ = 4.79 ppm). $^{13}$C NMR spectra were obtained with Bruker AVANCE 400 spectrometer and calibrated with CDCl$_3$ ($\delta$ = 77.00 ppm). Coupling constants (J) are reported in hertz (Hz). Splitting patterns are described using the following abbreviations: s, singlet; brs, broad singlet; d, doublet; t, triplet; q, quartet; dd, doublet of doublet; m, multiplet. 1H NMR spectra are reported in the following order: chemical shift, multiplicity, coupling constant(s), and number(s) of protons. All NMR signals were assigned on the basis of $^1$H NMR and $^{13}$C NMR experiments. High resolution MALDI mass spectra were recorded on a Bruker Ultraflextreme spectrometer.

Neu5Ac and Neu5Gc were purchased from Carbosynth Limited. ATP and CTP were purchased from Sigma. Thermosensitive Alkaline Phosphatase from shrimp (FastAP) was purchased from Thermo Scientific. Other enzymes including *Neisseria meningitidis* $\beta$1,4-galactosyltransferase (NmLgtB), $\alpha$2,6-sialyltransferase from *Photobacterium damselae* (Pd2,6ST), C-terminal 66 amino acids truncated *Helicobacter pylori* $\alpha$1,3-fucosyltransferase (Hp$\alpha$1,3FT), *Helicobacter mustelae* $\alpha$1,2-fucosyltransferase (Hm$\alpha$1,2FT), CMP-sialic acid synthetase from *N. meningitides* (NmCSS) were expressed and purified as previously described. Enzymes were then desalted against 50 mM Tris-HCl, 100 mM NaCl, and 50% glycerol, and stored at -20 °C.
for long term use. Sugar nucleotides uridine 5'-diphospho-galactose (UDP-Gal), cytidine 5'-monophospho- N-acetylneuraminic acid (CMP-Neu5Ac), cytidine 5'-monophospho- N-Glycolylneuraminic acid (CMP-Neu5Gc) and guanoside 5'-diphospho-L-fucose (GDP-Fuc) were prepared as described below.

**General Procedures**

**A) Transformation of N-Phth to NHAc procedure:** A mixture of N-Phth protected oligosaccharide was dissolved in n-BuOH at room temperature, followed by addition of ethylenediamine (n-BuOH: ethylenediamine = 2:1). After being stirred at 90 °C for 12 h, the mixture was evaporated in vacuo to give a residue for the next step without further purification. To a solution of the residue in pyridine was added Ac₂O. After being stirred at room temperature for 12 h, the solution was diluted with EtOAc and washed with aqueous 1 M HCl, saturated aqueous NaHCO₃, and brine solution. The organic layer was dried over Na₂SO₄, filtered, and evaporated in vacuo to give a residue, which was purified by silica gel column chromatography to give NHAc compound.

**B) Deacetylation procedure:** Ac-protected oligosaccharide was dissolved in MeOH, and NaOMe in MeOH was added until pH was about 10. After stirring at room temperature for 12 h, the solution was neutralized with ion-exchange resin (H⁺), and then filtered. The residue was concentrated under vacuo to afford the desired deacetylated product.

**C) Deprotection of benzyl group:** Pd(OH)₂ on carbon was added to a solution of protected oligosaccharide in MeOH/H₂O (10/1). The mixture was stirred under 1 atmosphere of hydrogen. After being stirred for 24 h, the mixture was filtered through a PTFE syringe filter and concentrated in vacuo. The residue was purified by Bio-Gel P-2
(BIO-RAD) column chromatography using water as eluent. The product was then lyophilized to get target compound as white powder.

**D) Production of oligosaccharyl bromide:** Add peracetylated oligosaccharide portionwise to a stirred solution of HBr (33%) in glacial acetic acid (20.0 ml) at 0 °C. After all the sugar has been added, the reaction mixture was stirred at room temperature for 45 min. TLC analysis (Hexanes: Ethyl acetate=1:1) indicates formation of product and consumption of starting material. Then the reaction was quenched by ice water (200 ml) and then extract the product with DCM (2 × 200 ml). Wash the combined organic extracts with a solution of NaHCO$_3$ (aq., sat., 2 × 200 ml), dry with Na$_2$SO$_4$, filter and then concentrate in vacuo. The crude product was used without further purification.

**E) General methods for enzyme treatment and HPLC purification** In general, 31 HMOs were enzymatically synthesized by 4 glycosyltransferases (NmLgtB, Pd2,6ST, Hp$\alpha$ 1,3FT, Hm$\alpha$ 1,2FT) in the nearly same reaction condition. Reactions contain 50 mM Tris-HCl (pH 8.0), 10 mM of acceptor HMOs, 12 mM of sugar nucleotide (or its corresponding synthetase), 10 mM of MnCl$_2$, and varying amounts of Glycotransferases. FastAP (1 U/200 µL) was also added to digest the reaction byproduct UDP to drive reaction forward. Reactions incubated at 37 °C for overnight, and monitored by HILIC-ELSD (Waters XBridge BEH amide column, 130 Å, 4.6 mm × 250 mm under a gradient running condition (solvent A: 100 mM ammonium formate, pH 3.4; solvent B: acetonitrile; flow rate: 1 mL/min; B%: 65-50% within 25 min)). The desired products were detected by a high-efficient ELSD (evaporative light scattering detector) which increases the sample concentration to minimize the noise and deliver higher sensitivity. After over 90% acceptor converted, the reaction was quenched by boiling for 10 min, followed concentration by
rotary evaporator. HPLC-A210nm was then used to purify target HMOs using a semi-preparative column (Waters XBridge BEH amide column, 130 Å, 5 μm, 10 mm × 250 mm) under a gradient running condition which are solvent A: 100 mM ammonium formate, pH 3.4; solvent B: acetonitrile; flow rate: 4 mL/min; B%: 65-50% within 25 min.\textsuperscript{43} MS data for purified HMOs were obtained by ESI-MS and MALDI-MS.

**Benzyl O-(4, 6-O-benzylidene-3-O-(4-methoxybenzyl)-β-D-galactopyranosyl)-(1→4)-β-D-glucopyranoside (2-6)**

A suspension of benzylactose 2-5 (12.0 g, 27.78 mmol) and Bu\textsubscript{2}SnO (7.6 g, 30.54 mmol) in anhydrous MeOH (100 mL) was heated to reflux and stirred for 8 h. The solvent was removed in vacuo. Then the residue was dissolved in dry toluene (100 mL). \textit{p}-Methoxybenzyl chloride (3.76 mL, 20.37 mmol), tetrabutylammonium iodide (2.05 g, 11.10 mmol), and 4 Å molecular sieves (5 g) were added. The resulting mixture was heated to reflux for another 8 h and then cooled to room temperature. The suspension was filtered through a Celite pad and the filtrate was concentrated and chromatographed (dichloromethane-methanol=6:1) to afford 9.2 g of crude product (60% yield).

Benzaldehyde dimethyl acetal (2.75 mL, 18.33 mmol) was added to a solution of the above crude product (7.8 g, 14.10 mmol) in anhydrous Dimethylformamide (100 ml), then Camphorsulfonic acid was added to adjust the PH to about 2.0-3.0. The reaction mixture was stirred overnight and then quenched with triethylamine. The mixture was concentrated in vacuum. The residue was purified by flash column chromatography (dichloromethane-methanol=10:1) to give 2-6 as a white solid (8.47 g, 87.0%). [α]\textsubscript{D}\textsuperscript{20} = +6.7 (c 1.0, CH\textsubscript{2}Cl\textsubscript{2}). \textsuperscript{1}H NMR (CDCl\textsubscript{3}, 400 MHz): δ 7.48-7.51 (dd, 2H), 7.28-7.38 (m, 10 H), 6.85-6.87 (m, 2 H), 5.34 (s, 1 H), 4.87 (d, \textit{J} = 11.95 Hz, 1 H), 4.59-4.63 (m, 4 H), 4.45
(d, J = 7.8 Hz, 1 H), 4.36 (d, J = 8.1 Hz, 1 H), 4.21 (d, J = 14.0 Hz, 1 H), 4.10 (s, 1 H),
3.92-4.00 (m, 3 H), 3.79-3.89 (m, 2 H), 3.78 (s, 3 H), 3.60-3.69 (m, 3 H), 3.40-3.50 (m, 2 H), 3.27-3.40 (m, 4 H); \textsuperscript{13}C NMR (CDCl\textsubscript{3}, 100 MHz): δ 137.6, 137.2, 130.0, 129.5, 128.5, 128.3, 128.2, 127.9, 126.3, 113.9, 103.6, 101.8, 101.1, 78.8, 74.9, 74.7, 73.5, 72.7, 71.3, 71.2, 69.1, 69.0, 66.9, 61.9, 55.3. HRMS: [M + Na]\textsuperscript{+} C\textsubscript{34}H\textsubscript{40}NaO\textsubscript{12} calcd for 663.2417, found 663.2420.

**Benzyl O-(2-O-Benzyl-4, 6-O-benzylidene-3-O-(4-methoxybenzyl)-β-D-galactopyranosyl)-(1→4)-2, 3, 6-tri-O-benzyl-β-D-glucopyranoside (2-7)**

60% NaH (2.25 g, 56.25 mmol) and BnBr (6.66 ml, 56.25 mmol) were added to a stirred solution of 2-6 (6.0g, 9.38 mmol) in DMF (60 mL) cooled at 0°C. The color of the solution changed to light yellow. The reaction mixture was maintained at room temperature for 4 h. Then the solution was quenched with MeOH. The mixture was diluted with EtOAc and washed with water. The organic layer was dried with Na\textsubscript{2}SO\textsubscript{4} and concentrated. The residue was purified on a silica gel column (Hexanes: EtOAc = 6:1) to afford the product 2-7 (1.85 g, 92.5%) as white powder. [α]\textsubscript{D}\textsuperscript{20} = +7.4 (c 1.0, CH\textsubscript{2}Cl\textsubscript{2}). \textsuperscript{1}H NMR (CDCl\textsubscript{3}, 400 MHz): δ 7.57-7.59 (m, 2 H), 7.51-7.54 (m, 2 H), 7.42-7.45 (m, 2 H), 7.31-7.38 (m, 23 H), 7.23-7.24 (m, 2 H), 6.90-6.92 (m, 2 H), 5.51 (s, 1 H), 5.25 (d, J = 11.0 Hz, 1 H), 4.98-5.03 (m, 2 H), 4.90 (d, J = 11.1 Hz, 1 H), 4.79-4.85 (m, 3 H), 4.71-4.74 (m, 3 H), 4.63 (d, J = 12.0 Hz, 1 H), 4.52-4.55 (m, 2 H), 4.43 (d, J = 12.0 Hz, 1 H), 4.26 (dd, J = 1.4 Hz, 12.4 Hz, 1 H), 4.04-4.08 (m, 2 H), 3.96 (dd, J = 4.2, 11.3 Hz, 1 H), 3.90 (dd, J = 1.8, 12.5 Hz, 1 H), 3.85 (s, 3 H), 3.79-3.83 (m, 2 H), 3.69 (t, J = 8.8 Hz, 1 H), 3.56-3.60 (m, 1 H), 3.40-3.46 (m, 2 H); \textsuperscript{13}C NMR (CDCl\textsubscript{3}, 100 MHz): δ 159.3, 139.0, 139.0, 138.7, 138.6, 138.2, 137.6, 130.5, 129.4, 128.9, 128.7, 128.4, 128.3, 128.1, 128.0, 127.8, 127.8,
Benzyl O-(2-O-Benzyl-4, 6-O-benzylidene- β-D-galactopyranosyl)-(1→4)-2, 3, 6-tri-O-benzyl- β-D-glucopyranoside (2-1)

2, 3-Dichloro-5, 6-dicyano-1, 4-benzoquinone (2.75 g, 12.10 mmol) was added to a solution of 2-7 (6.0 g, 6.05 mmol) in 9:1 CH₂Cl₂– Phosphate-buffered saline (200 mL) was added. The solution was stirred for 1.5h at room temperature and diluted with CH₂Cl₂. The solution was washed with aqueous NaHCO₃, brine, dried over Na₂SO₄, and concentrated. The residue was was purified on a silica gel column (Hexanes: EtOAc = 5:1) to afford the product 2-1 (5.02 g, 95%) as white powder. [α]D²⁰ = +15.4 (c 1.0, CH₂Cl₂). ¹H NMR (CDCl₃, 400 MHz): δ 7.52-7.55 (m, 4 H), 7.20-7.45 (m, 26 H), 5.56 (s, 1 H), 5.21 (d, J = 8.0 Hz 1 H), 4.97-5.02 (m, 2 H), 4.78-4.83 (m, 4 H), 4.65-4.73 (m, 2 H), 4.51-4.54 (m, 2 H), 4.46 (d, J = 12.3 Hz, 1 H), 4.30 (d, J = 12.3 Hz, 1 H), 4.06-4.13 (m, 2 H), 3.93-3.97 (m, 2 H), 3.79 (dd, J = 1.3, 11.0 Hz, 1 H), 3.68 (m, 1 H), 3.55-3.60 (m, 3 H), 3.38-3.41 (m, 1 H), 3.14 (s, 1 H); ¹³C NMR (CDCl₃, 100 MHz): δ 138.9, 138.7, 138.6, 138.5, 137.9, 137.5, 129.2, 128.8, 128.4, 128.2, 128.0, 127.8, 127.6, 127.4, 126.5, 102.8, 102.6, 101.5, 83.1, 81.9, 80.2, 77.6, 75.9, 75.7, 75.2, 75.1, 73.1, 72.9, 71.0, 68.9, 68.2, 66.5. HRMS: [M + Na]⁺ C₆₂H₆₄NaO₁₂ calcd for 1023.4295, found 1023.4285.

Ethyl 3, 6-di-O-benzyl-2-deoxy-2-phthalimido-1-thio-β-D-glucopyranoside (2-9)

A mixture of compound 2-8 (2 g, 4.09 mmol) and 4 Å molecular sieves (2 g) in dry CH₂Cl₂ was stirred at room temperature under nitrogen for 2 h. Triethylsilane (2.1 mL,
13.1 mmol) and TfOH (1.05 mL, 11.9 mmol) were sequentially added at -78 °C. The reaction mixture was stirred at -78 °C for 2 h and then quenched with MeOH (2 mL) and Et3N (2 mL). The resulting mixture was filtered. The filtrate was diluted with CH2Cl2 and washed with aqueous NaHCO3, brine, dried over Na2SO4, and concentrated. The residue was purified on a silica gel column (Hexanes: EtOAc = 5:1) to afford the product 2-9 (1.85 g, 92.5%) as white powder. [α]D20 = +91.0 (c 1.0, CHCl3). 1H NMR (CDCl3, 400 MHz): δ 7.82 (d, J = 5.9 Hz, 1 H), 7.66-7.72 (m, 3 H), 7.28-7.42 (m, 5 H), 7.03-7.10 (m, 2 H), 6.92-7.00 (m, 3 H), 5.32 (d, J = 9.9 Hz, 1 H), 4.80 (d, J = 12.0 Hz, 1 H), 4.54-4.70 (m, 3 H), 4.25-4.34 (m, 2 H), 3.80-3.89 (m, 3 H), 3.70-3.73 (m, 1 H), 2.59-2.72 (m, 2 H), 1.19 (t, J = 7.3 Hz, 3 H); 13C NMR (CDCl3, 100 MHz): δ 168.1, 167.6, 138.2, 137.7, 134.0, 133.9, 131.6, 128.5, 127.9, 127.9, 127.8, 127.5, 123.5, 123.3, 81.2, 79.7, 78.0, 74.5, 74.2, 73.78, 70.7, 54.5, 24.0, 15.0. HRMS: [M + Na]+ C30H31NNaO6S calcd for 556.1770, found 556.1760.

Ethyl 2, 3, 4, 6-tetra-O-acetyl- β-D-glucopyranosyl-(1→4)-3, 6-di-O-benzyl-2-deoxy-2-phthalimido-1-thio-β-D-glucopyranoside (2-3)

2, 3, 4, 6-Tetra-O-acetyl-β-D-galactosyl bromide (3.35 g, 8.17 mmol) was prepared by following general procedure (D). Then the bromide donor (3.35 g, 8.17 mmol) and 3, 6-di-O-benzyl-2-deoxy-2-phthalimido-1-thio-β-D-glucopyranoside 2-9 (2.64 g, 5.43 mmol) were dissolved in a mixture of dry toluene and CH2Cl2 (1:1, 30 mL). Powdered molecular sieves (4 Å) were added and the mixture was stirred under nitrogen for 1 h. The flask was wrapped in aluminum foil and cooled to −45°C. AgOTf (2.79 g, 10.86 mmol) dissolved in dry toluene (20 mL) was added during 1 h under the exclusion of light. After additional stirring for 30 min at −45°C, the reaction mixture was quenched
by aqueous Na$_2$S$_2$O$_3$. The mixture transferred to a separatory funnel via a Celite-packed glass filter funnel. The organic phase was separated, dried with Na$_2$SO$_4$, filtered, and concentrated. Purification of the residue by silica gel column chromatography (Hexanes:EtOAc=4:1) gave compound 2-3 (5.81 g, 80%). $[\alpha]_D^{20}$ = +31.0 (c 1.0, CHCl$_3$).

$^1$H NMR (CDCl$_3$, 400 MHz): $\delta$ 7.77 (d, $J$ = 6.3 Hz, 1 H), 7.62-7.66 (m, 3 H), 7.29-7.43 (m, 5 H), 7.01 (d, $J$ = 7.0 Hz), 6.82-6.92 (m, 3 H), 5.13-5.28 (m, 3 H), 4.78-4.87 (m, 3 H), 4.62 (d, $J$ = 7.9 Hz, 1 H), 4.43-4.51 (m, 2 H), 4.20-4.28 (m, 2 H), 4.09 (t, $J$ = 9.5 Hz, 1 H), 3.89-4.02 (m, 2 H), 3.79 (s, 2 H), 3.64 (t, $J$ = 6.8 Hz, 1 H), 3.56 (d, $J$ = 10.0 Hz, 1 H), 2.55-2.75 (m, 2 H), 2.06 (s, 3 H), 2.02 (s, 6 H), 1.97 (s, 3 H); $^{13}$C NMR (CDCl$_3$, 100 MHz): $\delta$ 170.3, 170.2, 170.0, 169.2, 167.9, 167.4, 138.5, 137.9, 133.9, 133.7, 131.6, 128.6, 128.0, 127.9, 127.1, 123.4, 123.3, 100.3, 81.1, 79.1, 77.8, 77.6, 74.5, 73.6, 70.4, 69.5, 67.7, 66.9, 60.7, 54.7, 23.9, 20.8, 20.7, 20.6, 20.6, 14.9. HRMS: [M + Na]$^+$ C$_{44}$H$_{49}$NNaO$_{15}$S calcd for 886.2721, found 886.2729.

**Benzyl O-(2, 4-di-O-Benzyl-β-D-galactopyranosyl)-(1 → 4)-2, 3, 6-tri-O-benzyl- β-D-glucopyranoside (2-13)**

A mixture of compound 2-1 (2 g, 2.27 mmol) and 4 Å molecular sieves (2 g) in dry CH$_2$Cl$_2$ was stirred at room temperature under nitrogen for 2 h. Triethylsilane (0.69 mL, 4.34 mmol) and PhBCl$_2$ (0.56 mL, 4.34 mmol) were sequentially added at -78 °C. The reaction mixture was stirred at -78 °C for 2 h and then quenched by the addition of MeOH (2 mL) and Et$_3$N (2 mL). The resulting mixture was filtered. The filtrate was diluted with CH$_2$Cl$_2$ and washed with aqueous NaHCO$_3$, brine, dried over Na$_2$SO$_4$, and concentrated. The residue was purified on a silica gel column (Hexanes: EtOAc = 5:1) to afford the product 2-13 (1.70 g, 85%) as white powder. $[\alpha]_D^{20}$ = +8.9 (c 1.0, CH$_2$Cl$_2$). $^1$H
NMR (CDCl$_3$, 400 MHz): $\delta$ 7.35-7.52 (m, 27 H), 7.26-7.34 (m, 3 H), 5.03-5.16 (m, 3 H), 4.78-4.97 (t, $J = 8.7$ Hz, 1 H), 4.70-4.75 (t, $J = 8.7$ Hz, 1 H), 4.52 (d, $J = 7.0$ Hz, 1 H), 4.08 (t, $J = 8.7$ Hz, 1 H), 3.92 (d, $J = 2.6$ Hz, 2 H); $^{13}$C NMR (CDCl$_3$, 100 MHz): $\delta$ 138.9, 138.7 (2 C), 138.5, 137.6, 128.6, 128.6, 128.5, 128.5, 128.1, 127.9, 127.8, 127.7, 127.6, 102.8, 102.6, 82.8, 81.9, 80.5, 77.0, 75.8, 75.5, 75.3, 75.2 (2 C), 75.1, 74.4, 73.4, 71.1, 68.4, 61.7. HRMS: [M + Na]$^+$ C$_{54}$H$_{58}$NaO$_{11}$ calcd for 905.3877, found 905.3867.

Benzyl O-(3, 4, 6-tri-O-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→3)-[3, 4, 6-tri-O-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl]-(1→6)-(2, 4-di-O-benzyl-β-D-galactopyranosyl)-(1→4)-2, 3, 6-tri-O-benzyl-β-D-glucopyranoside (2-14)

3, 4, 6-Tri-O-acetyl-2-deoxy-2-phthalimido-β-D-glucosyl bromide was prepared by following the general procedure (D). Powdered molecular sieves (4 Å) (3.0 g) was added to a solution of above bromide donor (4.80 g, 5.44 mmol) and 2-13 (800 mg, 0.907 mmol) in anhydrous dichloromethane (20 mL). The suspension was stirred under nitrogen for 1.5 h at room temperature and then cooled to -30° C. Then 2, 4, 6-collidine (0.72 mL, 5.44 mmol), and freshly dried AgOTf (1.40 g, 5.44 mmol) was sequentially added to the reaction mixture. After stirring for 2 h at -30° C, the mixture was allowed to warm up to rt overnight, diluted with CH$_2$Cl$_2$, filtered through Celite. The filtrate was diluted with CH$_2$Cl$_2$ and washed with aqueous NaHCO$_3$, brine, dried over Na$_2$SO$_4$, and concentrated. The residue was purified on a silica gel column (Hexanes: EtOAc = 5:1) to afford the product 2-14 (1.33 g, 85%) as white powder. $[\alpha]_{20}^D = +6.4$ (c 1.0, CH$_2$Cl$_2$). $^1$H NMR (CDCl$_3$, 400 MHz): $\delta$ 7.85 (t, $J = 4.01$ Hz, 1 H), 7.72-7.74 (m, 2 H), 7.43-7.56 (m, 2 H), 7.18-7.40 (m, 27 H), 7.08-7.16 (m, 3 H), 6.97 (m, 2 H), 5.76-5.87 (m, 2 H), 5.62 (d, $J = 8.24$ Hz, 1 H), 4.70-4.75 (t, $J = 8.7$ Hz, 1 H), 4.52 (d, $J = 7.0$ Hz, 1 H), 4.08 (t, $J = 8.7$ Hz, 1 H), 3.92 (d, $J = 2.6$ Hz, 2 H).
5.39 (d, J = 8.43, 1H), 5.16-5.24 (m, 2 H), 4.85-4.93 (m, 4 H), 4.75 (d, J = 10.6 Hz, 1 H), 4.55-4.66 (m, 3 H), 4.35-4.44 (m, 5 H), 4.19-4.33 (m, 5 H), 4.12 (d, J = 12.7 Hz, 1 H), 3.91 (t, J = 5.9 Hz, 1 H), 3.71-3.86 (m, 5 H), 3.47-3.59 (m, 5 H), 3.33-3.47 (m, 2 H), 3.21 (t, J = 5.9 Hz, 1 H), 3.05 (dd, J = 3.2, 9.7 Hz, 1 H), 2.09 (s, 3 H), 2.09 (s, 3 H), 2.02 (s, 3 H), 1.98 (s, 3 H), 1.87 (d, 6 H); 13C NMR (CDCl₃, 100 MHz): δ 170.7, 170.6, 170.1, 170.0, 169.5, 169.5, 139.1, 139.0, 138.6, 138.5, 138.4, 137.7, 134.4, 134.0, 128.4, 128.3, 128.2, 128.1, 128.0, 127.8, 127.6, 127.5, 127.4, 127.4, 127.1, 127.0, 126.9, 126.5, 102.4, 102.3, 99.0, 97.5, 83.1, 81.7, 81.7, 78.6, 76.2, 76.1, 75.6, 75.2, 75.0, 74.9, 74.1, 73.1, 72.8 (2 C), 71.7, 71.6, 70.8, 70.8 (2 C), 70.5, 68.9, 68.8, 68.0, 66.8, 61.7, 61.6, 55.1, 54.7, 20.8, 20.7, 20.6, 20.5, 20.4. HRMS: [M + Na]⁺ C₉₄H₉₆N₂NaO₂₉ calcld for 1739.5996, found 1739.5980.

Benzyl O-(3, 4, 6-tri-O-acetyl-2-deoxy-acetamido-β-D-glucopyranosyl)-(1→3)-[ 3, 4, 6-tri-O-acetyl-2-deoxy-acetamido-β-D-glucopyranosyl]-(1→6)-(2,4-di-O-benzyl-β-D-galactopyranosyl)-(1→4)-2, 3, 6-tri-O-benzyl- β-D-glucopyranoside (2-15)

Following the general procedure (A) compound 2-14 (1.16 g, 0.66 mmol) yielded the compound 2-15 (762mg, 75% over two steps). [α]D²⁰ = -1.9 (c 0.4, CH₂Cl₂). 1H NMR (CDCl₃, 400 MHz): δ 7.25-7.48 (m, 27 H), 7.10-7.18 (m, 3 H), 5.78 (d, J = 9.5 Hz, 1 H), 5.01-5.16 (m, 6 H), 4.90-5.00 (m, 3 H), 4.75-4.91 (m, 4 H), 4.60-4.71 (m, 3 H), 4.41-4.55 (m, 4 H), 4.29 (d, J = 3.5 Hz, 1 H), 3.96-4.15 (m, 5 H), 3.64-3.86 (m, 8 H), 3.52-3.64 (m, 4 H), 3.45-3.51 (m, 1 H), 3.23 (d, J = 9.7 Hz, 1 H), 2.07 (s, 3 H), 2.06 (s, 3 H), 2.05 (s, 3 H), 2.00 (s, 3 H), 1.99 (s, 3 H), 1.98 (s, 3 H), 1.51 (s, 3 H); 13C NMR (CDCl₃, 100 MHz): δ 170.9, 170.8, 170.6, 170.5, 170.3, 169.8, 169.3, 169.2, 139.4, 138.9, 138.8, 138.6, 137.9,
Following the general procedure (B) and (C), compound 2-15 (400 mg, 0.26 mmol) yielded the compound HMO1 (152 mg, 80% over two steps). \(^1\)H NMR (D\(_2\)O, 400 MHz): \(\delta \) 5.15 (d, \(J = 3.6 \) Hz, 0.55 H, Glc-1 H-1 of \(\alpha\) form), 4.53-4.62 (m, overlap with D\(_2\)O, 2.45 H, GlcNAc-1 H-1, GlcNAc-2 H-1 and Glc-1 H-1 of \(\beta\) form), 4.36 (d, \(J = 7.7 \) Hz, 1 H, Gal-1 H-1), 4.07 (d, \(J = 2.4 \) Hz, 1 H), 3.58-3.95 (m, 13 H), 3.45-3.58 (m, 5 H), 3.32-3.45 (m, 4 H), 3.16-3.26 (m, 1 H), 1.99 (s, 3 H), 1.96 (s, 3 H); \(^{13}\)C NMR (D\(_2\)O, 100 MHz): \(\delta \) 174.9, 174.6, 102.9, 102.8 (GlcNAc-1, GlcNAc-2, C-1), 101.1 (Gal-1, C-1), 95.7 (Glc-1, C-1 of \(\beta\) form), 91.8 (Glc-1, C-1 of \(\alpha\) form), 81.7, 78.8, 75.8 (2 C), 75.6, 73.8, 74.7, 74.3, 73.8, 73.5, 73.4, 69.86, 69.8, 69.6, 68.7, 68.4, 60.7, 60.5, 55.6, 55.5, 22.4, 22.2. HRMS: [M + Na]\(^+\) C\(_{28}\)H\(_{48}\)N\(_2\)NaO\(_{21}\) calcd for 771.2647, found 771.2520.

Benzyl O-(2, 3, 4, 6-tetra-O-acetyl-\(\beta\)-D-galactopyranosyl)-(1\(\rightarrow\)4)-3, 6-di-O-benzyl-2-deoxy-2-phthalimido-\(\beta\)-D-glucopyranosyl-(1 \(\rightarrow\) 3)-(2-O-benzyl-4, 6-O-benzylidene-\(\beta\)-D-galactopyranosyl)-(1 \(\rightarrow\) 4)-2, 3, 6-tri-O-benzyl- \(\beta\)-D-glucopyranoside (2-16)

Powdered molecular sieves (4 Å) (3.0 g) was added to a solution of Compound 2-3 (871 mg, 0.96 mmol) and 2-1 (652 mg, 0.74 mmol) in anhydrous dichloromethane (20
mL). The suspension was stirred under nitrogen for 1.5 h at room temperature and then cooled to -30°C. Then NIS (260mg, 1.15mmol), and TMSOTf (35 µl, 0.19 mmol) was sequentially added to the reaction mixture. After stirring for 2 h at -30°C, the mixture was allowed to warm up to rt. overnight, diluted with CH₂Cl₂, filtered through Celite. The filtrate was diluted with CH₂Cl₂ and washed with aqueous NaHCO₃, brine, dried over Na₂SO₄, and concentrated. The residue was purified on a silica gel column (Hexanes: EtOAc = 5:1) to afford the product **2-16** (1.04 g, 86%) as white powder. [α]D²⁰ = +12.9 (c 1.0, CH₂Cl₂). ¹H NMR (CDCl₃, 400 MHz): δ 7.53-7.55 (m, 2 H), 7.45-7.50 (m, 2 H), 7.27-7.45 (m, 27 H), 7.19 (t, J = 3.6 Hz, 2 H), 7.09-7.16 (m, 3 H), 7.02-7.06 (m, 2 H), 6.84-6.94 (m, 5 H), 5.49 (s, 1 H), 5.47 (d, J = 7.4 Hz, 1 H), 5.35 (d, J = 3.2 Hz, 1 H), 5.25 (dd, J = 7.9, 10.4 Hz, 1 H), 5.09 (d, J = 10.7 Hz, 1 H), 4.88-4.99 (m, 3 H), 4.85 (d, J = 12.0 Hz, 1 H), 4.65-4.78 (m, 4 H), 4.61 (d, J = 12.0 Hz, 1 H), 4.44-4.55 (m, 3 H), 4.30-4.39 (m, 4 H), 4.18-4.28 (m, 5 H), 4.09 (t, J = 9.3 Hz, 1 H), 3.98-4.04 (m, 2 H), 3.89-3.96 (m, 1 H), 3.80-3.89 (m 2 H), 3.66-3.79 (m, 3 H), 3.53-3.64 (m, 2 H), 3.44-3.51 (m, 3 H), 3.38 (d, J = 10.1 Hz, 1 H), 2.93-3.02 (m, 2 H), 2.12 (s, 3 H), 2.10 (s, 3 H), 2.05 (s, 3 H), 2.03 (s, 3 H); ¹³C NMR (CDCl₃, 100 MHz): δ 170.3, 170.2, 170.1, 169.2, 139.0, 138.6 (2 C), 138.5, 138.3, 137.8, 137.6, 133.5, 131.3, 128.7, 128.6 (2 C), 128.4, 128.3 (2 C), 128.2, 128.1 (2 C), 128.0, 127.9, 127.8, 127.7, 127.6, 127.5, 127.2, 127.1, 126.7, 126.5, 126.3, 123.1, 102.4 (2 C), 100.7, 100.6, 99.6, 98.1, 81.8, 80.9, 78.4, 77.7, 76.0, 75.7, 75.0, 74.8, 74.7, 74.5, 74.3, 73.8 (2 C), 73.0 (2 C), 71.1, 70.9 (2 C), 70.6, 69.7, 68.8, 68.7, 67.9, 67.0, 66.4, 60.8, 55.8, 20.9, 20.7 (2 C), 20.6. HRMS: [M + Na]⁺ C₉₆H₉₉NaO₂₆ calcd for 1704.6353, found 1704.6383.
**Benzyl O-(2, 3, 4, 6-tetra-O-acetyl-β-D-galactopyranosyl)-(1→4)-3, 6-di-O-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl-(1→3)-(2-O-benzyl-β-D-galactopyranosyl)-(1→4)-2, 3, 6-tri-O-benzyl-β-D-glucopyranoside (2-19)**

To a solution of compound 2-16 (800 mg, 0.49 mmol) in anhydrous MeOH (10 mL) was added TsOH (8.4 mg) and EtSH (0.21 ml, 2.93 mmol). The reaction mixture was stirred at rt. for 6 h and then quenched with triethylamine and evaporated under reduced pressure. The mixture was purified with silica column (Hexanes: Acetone = 5:1) to get white power compound 2-19 (702 mg, 90%). $[\alpha]_{20}^D = +18.4$ (c 0.5, CH$_2$Cl$_2$). $^1$H NMR (CDCl$_3$, 400 MHz): $\delta$ 7.38-7.52 (m, 11 H), 7.25-7.36 (m, 16 H), 7.22-7.26 (m, 2 H), 7.03-7.14 (m, 3 H), 6.99-7.03 (m, 2 H), 6.85-6.92 (m, 2 H), 6.78 (d, $J = 7.1$ Hz, 2 H), 5.29-5.31 (m, 2 H), 5.19-5.25 (m, 1 H), 4.86-4.98 (m, 4 H), 4.83 (d, $J = 12.2$ Hz, 1 H), 4.68-4.77 (m, 3 H), 4.65 (d, $J = 8.3$ Hz, 1 H), 4.57-4.62 (m, 2 H), 4.46 (t, $J = 12.4$ Hz, 2 H), 4.34-4.40 (m, 1 H), 4.26-4.34 (m, 3 H), 4.17-4.26 (m, 3 H), 3.95-4.08 (m, 3 H), 3.81-3.88 (m, 1 H), 3.72-3.79 (m, 3 H), 3.64-3.72 (m, 2 H), 3.47-3.55 (m, 2 H), 3.34-3.46 (m, 5 H), 3.17-3.23 (m, 1 H), 3.00-3.06 (m, 1 H), 2.87 (s, 1 H), 2.11 (s, 3 H), 2.08 (s, 3 H), 2.05 (s, 3 H), 2.02 (s, 3 H); $^{13}$C NMR (CDCl$_3$, 100 MHz): $\delta$ 170.3, 170.2, 170.1, 169.2, 138.9, 138.4, 138.3 (2 C), 137.6, 137.5, 133.6, 131.1, 128.8, 128.4, 128.3 (3 C), 128.2 (2 C), 128.1, 127.9, 127.8, 127.7, 127.6, 127.5, 127.4, 127.2, 126.7, 126.4, 123.2, 102.4, 102.1, 100.6, 99.0, 83.8, 82.8, 81.5, 78.3, 77.9, 76.6, 76.4, 75.6, 74.9, 74.6, 74.3, 73.8, 73.7, 73.1, 71.0, 70.9, 70.7, 69.6, 68.3, 67.9, 67.7, 66.9, 62.2, 60.8, 55.6, 20.8, 20.68, 20.6, 20.6. HRMS: [M + Na]$^+$ C$_{69}$H$_{95}$NNaO$_{26}$ calcd for 1616.6040, found 1616.6065.

**Benzyl O-(2, 3, 4, 6-tetra-O-acetyl-β-D-galactopyranosyl)-(1→4)-3, 6-di-O-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl-(1→3) 2-19**
deoxy-2-phthalimido-β-D-glucopyranosyl](1 → 6)-(2-O-benzyl-β-D-galactopyranosyl)-(1→4)-2, 3, 6-tri-O-benzyl-β-D-glucopyranoside (2-20)

3, 4, 6-Tri-O-acetyl-2-deoxy-2-phthalimido-β-D-glucosyl bromide was prepared by following the general procedure (D). Powdered molecular sieves (4 Å) (3.0 g) was added to a solution of 2-19 (460 mg, 0.29 mmol), 2, 4, 6-collidine (76 µL, 0.58 mmol), and freshly dried AgOTf (150 mg, 0.58 mmol) in anhydrous dichloromethane (20 mL). The suspension was stirred under nitrogen for 1.5 h at room temperature and then cooled to -30°C. Then a solution of above bromide donor (460 mg, 0.29 mmol) in dichloromethane (5.0 mL) was added dropwise during 30 min to the reaction mixture. After stirring for 2 h at -30°C, the mixture was allowed to warm up to rt. overnight, diluted with CH₂Cl₂, filtered through Celite. The filtrate was diluted with CH₂Cl₂ and washed with aqueous NaHCO₃, brine, dried over Na₂SO₄, and concentrated. The residue was purified on a silica gel column (Hexanes: EtOAc = 5:1) to afford the product 2-20 (495 mg, 85%) as white powder. [α]D²⁰ = +11.7 (c 1.0, CH₂Cl₂). ¹H NMR (CDCl₃, 400 MHz): δ 7.76-7.91 (m, 4 H), 7.49-7.56 (m, 2 H), 7.41-7.49 (m, 8 H), 7.14-7.20 (m, 2 H), 7.05-7.11 (m, 1 H), 6.97-7.04 (m, 4 H), 6.85-6.94 (m, 3 H), 6.75 (d, J = 7.4 Hz, 2 H), 5.73 (dd, J = 9.1, 10.6 Hz, 1 H), 5.43 (d, J = 8.3 Hz, 1 H), 5.33 (d, J = 3.5 Hz, 1 H), 5.12-5.25 (m, 2 H), 4.99 (d, J = 8.9 Hz, 1 H), 4.77-4.95 (m, 3 H), 4.72-4.84 (m, 3 H), 4.61-4.71 (m, 2 H), 4.48-4.58 (m, 2 H), 4.33-4.47 (m, 3 H), 4.09-4.31 (m, 8 H), 3.80-4.06 (m, 6 H), 3.65-3.80 (m, 5 H), 3.62 (d, J = 10.0 Hz, 1 H), 3.50-3.56 (m, 2 H), 3.36-3.45 (m, 3 H), 3.20-3.27 (m, 2 H), 3.09-3.17 (m, 2 H), 2.87 (s, 1 H), 2.14 (s, 3 H), 2.11 (s, 3 H), 2.09 (s, 3 H), 2.04 (s, 3 H), 2.02 (s, 3 H), 1.99 (s, 3 H), 1.87 (s, 3 H); ¹³C NMR (CDCl₃, 100 MHz): δ 170.7, 170.3, 170.2, 170.0 (2 C), 169.5, 169.5, 139.1, 138.7, 138.4, 138.3 (2 C), 137.7, 137.6, 131.2, 128.8, 128.4, 128.3
(3 C), 128.2, 128.1, 127.9 (2 C), 127.8, 127.7, 127.6 (2 C), 127.4, 127.2, 126.5, 123.2, 102.4, 101.9, 100.7, 98.5, 97.6, 83.8, 83.1, 81.9, 78.4, 77.8, 76.5, 76.3, 75.7, 75.1, 74.9, 74.5, 74.4, 73.9, 72.9 (2 C), 72.0, 71.7, 71.1, 71.0, 70.8, 70.6, 69.5, 68.8, 68.1, 67.6, 66.9, 66.6, 61.7, 60.8, 55.4, 54.7, 20.9, 20.8, 20.6, 20.6 (3 C), 20.5. HRMS: [M + Na]+

C109H114N2NaO35 calcd for 2033.7100, found 2033.7089.

**Benzyl O-(2, 3, 4, 6-tetra-O-acetyl-β-D-galactopyranosyl)-(1→4)- 3, 6-di-O-benzyl-2-deoxy-acetamido-β-D-glucopyranosyl-(1→3)-[3, 4, 6-tri-O-acetyl-2-deoxy-acetamido-β-D-glucopyranosyl]-(1→6)-(4-O-acetyl-2-O-benzyl-β-D-galactopyranosyl)-(1→4)-2, 3, 6-tri-O-benzyl- β-D-glucopyranoside (2-21)**

Following the general procedure (A), compound 2-20 (400 mg, 0.20 mmol) yielded the compound 2-21 (232 mg, 62% over two steps). [α]D20 = -1.0 (c 1.2, CH2Cl2).

1H NMR (CDCl3, 400 MHz): δ 7.51 (m, 2H), 7.21-7.43 (m, 33 H), 5.89 (d, J = 9.5 Hz , 1 H), 5.43 (d, J = 8.4 Hz, 1 H), 5.33 (d, J = 2.8 Hz, 1 H), 4.33-5.24 (m, 23 H), 4.24 (d, J = 8.2 Hz, 1 H), 3.94-4.10 (m, 6 H), 3.45-3.86 (m, 15 H), 3.23 (d, J = 9.7 Hz, 1 H), 2.12 (s, 3 H), 2.11 (s, 3 H), 2.05 (s, 3 H), 2.04 (s, 3 H), 2.01 (s, 3 H), 2.00 (s, 3 H), 1.99 (s, 3 H), 1.97 (s, 3 H), 1.94 (s, 3 H), 1.54 (s, 3 H); 13C NMR (CDCl3, 100 MHz): δ 170.8, 170.7, 170.4, 170.3, 170.2, 170.0, 170.0 (2 C), 169.7, 169.2, 139.4, 138.7, 138.6 (2 C), 138.0, 137.7, 137.4, 128.7, 128.6, 128.4, 128.4 (2 C), 128.3, 128.2, 128.1, 128.0, 127.9 (2 C), 127.8, 127.7, 127.6, 127.0, 102.6, 102.3, 101.1, 100.7, 100.0, 83.8, 82.4, 79.5, 78.3, 77.9, 76.2 (2 C), 75.6, 75.0, 74.9, 74.6, 74.4, 73.9, 73.5, 73.3, 73.2 (2 C), 73.2, 70.9, 70.8, 70.7, 70.6, 70.0, 69.4, 68.3, 68.0, 67.9, 67.9, 67.3, 66.9, 61.5, 60.8, 54.5, 53.6, 20.9, 20.8, 20.7 (3 C), 20.6 (3 C), 20.5 (2 C). HRMS: [M + Na]+ C99H116N2NaO34 calcd for 1899.7307, found 1899.7287.
\[ \beta\text{-D-galactopyranosyl-(1→4)-2-deoxy-acetamido-\beta-D-glucopyranosyl-(1→3)-[ 2-deoxy-acetamido-\beta-D-glucopyranosyl]-(1→6)- \beta-D-galactopyranosyl-(1→4)-\alpha,\beta-D-glucopyranose (HMO2) } \]

Following the general procedure (B) and (C), compound 2-21 (200 mg, 0.11 mmol) yielded the compound HMO2 (87 mg, 90%). \(^1\)H NMR (D\(_2\)O, 400 MHz): \(\delta\) 5.11 (d, \(J = 3.6\) Hz, 0.42 H, Glc-1 H-1 of \(\alpha\) form), 4.53-4.62 (m, 1.58 H, GlcNAc-1 H-1, Glc-1 H-1 of \(\beta\) form), 4.50 (d, \(J = 8.4\) Hz, 1 H, GlcNHAC-2 H-1), 4.36 (d, \(J = 7.9\) Hz, 1 H, Gal-1 H-1), 4.31 (d, \(J = 8.0\) Hz, 1 H, Gal-2 H-1), 4.03 (d, \(J = 3.0\) Hz, 1 H), 3.77-3.92 (m, 5 H), 3.40-3.77 (m, 21 H), 3.30-3.38 (m, 2 H), 3.17 (t, \(J = 8.9\) Hz, 1 H), 1.95 (s, 3 H), 1.92 (s, 3 H); \(^{13}\)C NMR (D\(_2\)O, 100 MHz): \(\delta\) 174.9, 174.5, 103.0, 102.8, 102.7 (GlcNAc-1, GlcNAc-2, Gal-2, C-1), 101.0 (Gal-1, C-1), 95.7 (Glc-1, C-1 of \(\beta\) form), 91.8 (Glc-1, C-1 of \(\alpha\) form), 81.8, 78.9, 78.8, 78.1, 75.8, 75.3, 74.7, 74.5, 74.3, 73.8, 73.4, 72.5, 72.1, 71.4, 71.2, 70.9, 69.6, 68.5, 68.4, 61.0, 60.1, 60.0, 59.8, 55.5, 55.2, 22.4, 22.2. HRMS: [M + Na]\(^+\) C\(_{34}\)H\(_{58}\)N\(_2\)NaO\(_{26}\) calcd for 933.3175, found 933.3195.

**Benzyl O-(3, 4, 6-tri-O-acetyl-2-deoxy-2-phthalimido-\beta-D-glucopyranosyl)-(1→3)-(2-O-benzyl-4, 6-O-benzyldiene-\beta-D-galactopyranosyl)-(1→4)-2, 3, 6-tri-O-benzyl- \beta-D-glucopyranoside (2-22)**

3, 4, 6-Tri-O-acetyl-2-deoxy-2-phthalimido-\(\beta\)-D-glucosyl bromide was prepared by following the general procedure (D). Powdered molecular sieves (4 Å) (2.0 g) was added to a solution of above bromide donor 2-2 (1.2 g, 2.43 mmol) and 2-1 (500 mg, 0.61 mmol) in anhydrous dichloromethane (20 mL). The suspension was stirred under nitrogen for 1.5 h at room temperature and then cooled to -30°C. Then 2, 4, 6-collidine (0.32 mL, 2.43 mmol), and freshly dried AgOTf (624 mg, 2.43 mmol) was sequentially added to the
reaction mixture. After stirring for 2 h at -30°C, the mixture was allowed to warm up to rt overnight, diluted with CH₂Cl₂, filtered through Celite. The filtrate was diluted with CH₂Cl₂ and washed with aqueous NaHCO₃, brine, dried over Na₂SO₄, and concentrated. The residue was purified on a silica gel column (Hexanes: Acetone = 6:1) to afford the product **2-22** (626 mg, 85%) as white powder. [α]D²⁰ = +12.6 (c 1.0, CH₂Cl₂). ¹H NMR (CDCl₃, 400 MHz): δ 7.30-7.63 (m, 26 H), 7.22-7.27 (m, 3 H), 7.16-7.21 (m, 3 H), 6.89-6.96 (m, 2 H), 5.90 (dd, J = 9.4, 10.65 Hz, 1 H), 5.76 (d, J = 8.3 Hz, 1 H), 5.59 (s, 1 H), 5.29 (t, J = 9.7 Hz, 1 H), 5.17 (d, J = 10.6 Hz, 1 H), 4.92-4.99 (m, 2 H), 4.72-4.82 (m, 2 H), 4.63 (t, J = 11.6 Hz, 1 H), 4.54 (dd, J = 8.3, 10.8 Hz, 1 H), 4.25-4.47 (m, 9 H), 3.93-4.05 (m, 3 H), 3.61-3.73 (m, 2 H), 3.49-3.60 (m, 3 H), 3.44 (d, J = 11.3 Hz, 1 H), 3.14 (s, 1 H), 2.98-3.04 (m, 1 H), 2.13 (s, 3 H), 2.12 (s, 3 H), 1.90 (s, 3 H); ¹³C NMR (D₂O, 100 MHz): δ 170.5, 170.2, 169.6, 139.0, 138.7, 138.6, 138.5, 138.3, 137.6, 134.2, 128.7, 128.4, 128.3, 128.2, 128.2, 128.1, 128.0, 127.8, 127.7, 127.6, 127.3, 126.9, 126.4, 123.4, 102.4 (2 C), 100.9, 99.3, 83.1, 81.8, 81.3, 76.9, 75.9, 75.8, 75.1, 74.8, 74.3, 73.1, 71.9, 71.0, 70.8, 69.0, 68.9, 67.9, 66.4, 62.2, 54.8, 53.7, 20.9, 20.8, 20.5. HRMS: [M + Na]⁺ C₇₄H₇₅NNaO₂₀ calcd for 1320.4780, found 1320.4806.

Benzyl O-(3, 4, 6-tri-O-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1 → 3)-(2-O-benzyl-β-D-galactopyranosyl)-(1 → 4)-2, 3, 6-tri-O-benzyl-β-D-glucopyranoside (2-23)

To a solution of compound **2-22** (602 mg, 0.46 mmol) in anhydrous MeOH (10 mL) was added TsOH (16.0 mg) and EtSH (0.21 mL, 2.93 mmol). The reaction mixture was stirred at rt. for 6 hours then quenched with triethylamine and evaporated under reduced pressure. The mixture was purified with silica column (Hexanes: Acetone=5:1) to get
white power compound 2-23 (445 mg, 80%). $[\alpha]_D^{20} = +17.7$ (c 1.0, CH$_2$Cl$_2$). $^1$H NMR (CDCl$_3$, 400 MHz): $\delta$ 7.51-7.66 (m, 4 H), 7.41-7.45 (m, 2 H), 7.25-7.40 (m, 18 H), 7.06-7.17 (m, 3 H), 6.83 (d, $J = 6.7$ Hz, 2 H), 5.86 (dd, $J = 9.1$, 10.6 Hz, 1 H), 5.65 (d, $J = 8.5$ Hz, 1 H), 5.18 (t, $J = 9.4$ Hz, 1 H), 4.88-4.96 (m, 3 H), 4.69-4.76 (m, 2 H), 4.62 (d, $J = 12.0$ Hz, 1 H), 4.43-4.56 (m, 2 H), 4.37-4.41 (m, 1 H), 4.23-4.32 (m, 9 H), 3.95-4.03 (m, 3 H), 3.87-3.93 (m, 1 H), 3.68-3.75 (m, 1 H), 3.52-3.66 (m, 2 H), 3.38-3.50 (m, 5 H), 3.23-3.29 (m, 1 H), 3.03-3.08 (m, 1 H), 2.83 (d, $J=1.5$ Hz, 1 H), 2.15 (s, 3 H), 2.09 (s, 3 H), 1.88 (s, 3 H); $^{13}$C NMR (D$_2$O, 100 MHz): $\delta$ 170.8, 170.1, 169.5, 138.8, 138.6, 138.4, 138.3, 134.3, 128.4 (2 C), 128.3, 128.2, 128.2, 128.0, 127.9, 127.8, 127.7 (2 C), 127.6, 127.5, 126.8, 126.3, 123.5, 102.4, 102.1, 98.7, 84.3, 82.8, 81.7, 77.8, 76.3, 75.7, 75.0, 74.7, 74.3, 73.9, 73.1, 72.1, 70.9, 70.5, 69.0, 68.0, 67.7, 62.2, 61.9, 54.6, 20.8, 20.7, 20.4. HRMS: [M + Na]$^+$ C$_{67}$H$_{71}$NaO$_{20}$ calcd for 1232.4467, found 1232.4490.

**Benzyl O-(3, 4, 6-tri-O-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→3)-[ 2, 3, 4, 6-tetra-O-acetyl-β-D-galactopyranosyl)-(1→4)-3, 6-di-O-benzyl-2-deoxy-acetamido-β-D-glucopyranosyl]- (1→6)-(2-O-benzyl-β-D-galactopyranosyl)-(1→4)-2, 3, 6-tri-O-benzyl- β-D-glu-copyranoside (2-24)**

Powdered molecular sieves (4 Å) (1.0 g) was added to a solution of Compound 2-3 (230 mg, 0.27mmol) and 2-23 (261 mg, 0.21 mmol) in anhydrous dichloromethane (10 mL). The suspension was stirred under nitrogen for 1.5 hours at room temperature and then cooled to -30°C. Then NIS (61 mg, 0.27 mmol), and AgOTf (35 mg, 0.105 mmol) was sequentially added to the reaction mixture. After stirring for 2 h at -30°C, the mixture was allowed to warm up to rt. overnight, diluted with CH$_2$Cl$_2$, filtered through Celite. The filtrate was diluted with CH$_2$Cl$_2$ and washed with aqueous NaHCO$_3$, brine, dried over
Na$_2$SO$_4$, and concentrated. The residue was purified on a silica gel column (Hexanes: Acetone = 5:1) to afford the product 2-24 (295 mg, 70%) as white powder. $[\alpha]_b^{20} = +13.0$ (c 1.0, CH$_2$Cl$_2$). $^1$H NMR (CDCl$_3$, 400 MHz): δ 7.76-8.08 (m, 5 H), 7.49-7.72 (m, 3 H), 7.17-7.48 (m, 25 H), 6.96-7.11 (m, 5 H), 6.87-6.95 (m, 3 H), 6.74 (d, $J = 7.3$ Hz, 2 H), 5.55-5.61 (dd, $J = 9.3$, 10.37 Hz, 1 H), 5.30 (d, $J = 3.3$ Hz, 1 H), 5.11-5.20 (m, 2 H), 5.04 (d, $J = 8.5$ Hz, 1 H), 4.69-4.95 (m, 9 H), 4.42-4.66 (m, 6 H), 4.21-4.39 (m, 6 H), 4.03-4.15 (m, 5 H), 3.91-4.02 (m, 3 H), 3.70-3.88 (m, 4 H), 3.50-3.68 (m, 4 H), 3.23-3.46 (m, 8 H), 3.00-3.07 (m, 1 H), 2.60-2.69 (m, 1 H), 2.11 (s, 3 H), 2.09 (s, 3 H), 2.08 (s, 3 H), 2.07 (s, 3 H), 2.03 (s, 3 H), 2.00 (s, 3 H), 1.85 (s, 3 H); $^{13}$C NMR (D$_2$O, 100 MHz): δ 171.2, 170.4, 170.3, 170.1, 170.0, 169.2, 169.2, 138.8, 138.7, 138.5, 138.3, 138.2, 137.7, 137.6, 128.7, 128.5, 128.4, 128.3, 128.2 (2 C), 128.0 (2 C), 127.9, 127.8, 127.7, 127.6, 127.5, 127.4, 127.3, 126.6, 126.2, 102.5, 101.8, 100.3, 97.8, 97.4, 84.6, 82.9, 81.8, 77.9 (2 C), 77.3, 76.5, 76.1, 75.7, 75.0, 74.8, 74.7, 74.4, 74.3, 73.1, 71.4, 71.0, 70.9, 70.4, 70.1, 70.0, 69.5, 69.4, 69.2, 67.8, 67.2, 66.9, 65.5, 64.5, 62.2, 60.7, 55.6, 54.1, 20.9, 20.7, 20.7, 20.6 (3 C), 20.4. HRMS: [M + Na]$^+$ C$_{109}$H$_{114}$N$_2$NaO$_{35}$ calcd for 2033.7100, found 2033.7129.

**Benzyl O-(3, 4, 6-tri-O-acetyl-2-deoxy-acetamido-β-D-glucopyranosyl)-(1→3)-[ 2, 3, 4, 6-tetra-O-acetyl-β-D-galactopyranosyl]-(1→4)-3, 6-di-O-benzyl-2-deoxy-acetamido-β-D-glucopyranosyl]-(1→6)-(4-O-acetyl-2-O-benzyl-β-D-galactopyranosyl)-(1→4)-2, 3, 6-tri-O-benzyl- β-D-glucopyranoside (2-25)**

Following the general procedure (A), compound 2-24 (400 mg, 0.20 mmol) yielded the compound 2-25 (183 mg, 70% over two steps). $[\alpha]_b^{20} = -2.0$ (c 1.0, CH$_2$Cl$_2$). $^1$H NMR (CDCl$_3$, 400 MHz): δ 7.16-7.51 (m, 35 H), 5.20-5.36 (m, 2 H), 4.97-5.20 (m, 6 H), 4.89-4.97 (m, 2 H), 4.69-4.89 (m, 8 H), 4.56-4.69 (m, 4 H), 4.29-4.56 (m, 7 H), 4.08-
4.26 (m, 2 H), 3.41-4.04 (m, 20 H), 3.32-3.40 (m, 1 H), 3.23-3.31 (m, 1 H), 3.08-3.19 (d, 
$J = 8.6$ Hz, 1 H), 2.10 (s, 3 H), 2.09 (s, 3 H), 2.08 (s, 3 H), 2.04 (s, 3 H), 2.01 (s, 3 H), 
1.98 (s, 9 H), 1.92 (s, 3 H), 1.57 (s, 3 H); $^{13}$C NMR (D$_2$O, 100 MHz): $\delta$ 171.0, 170.9, 170.6 
(2 C), 170.2, 170.0, 169.8, 169.7, 169.3 (2 C), 139.1, 139.0, 138.7, 138.5, 138.2, 137.9, 
137.4, 128.7, 128.6, 128.5, 128.4, 128.3, 128.2, 128.1, 128.0, 127.9, 127.8, 127.7 (2 C), 
127.6, 127.5, 126.6, 102.4, 102.0, 101.1, 100.0, 99.8, 82.9, 82.0, 79.9, 78.4, 78.2, 77.4, 
76.4, 75.5, 75.0, 74.8, 74.1, 73.7, 73.6, 73.3, 72.7, 72.3, 71.7, 71.0, 70.9, 70.4, 69.5, 
68.3, 67.9, 66.9, 66.7, 61.5, 60.8, 55.2, 54.3, 53.9, 20.9, 20.8 (2 C), 20.7 (2 C), 20.6 (3 
C), 20.50 (2 C).

HRMS: [M + Na]$^+$$^{C_{99}}H_{116}N_{2}NaO_{34}$ calcd for 1899.7307, found 1899.7357.

2-deoxy-acetamido-β-D-glucopyranosyl (1→3)-[β-D-galactopyranosyl-(1→4)-
2-deoxy-acetamido-β-D-glucopyranosyl]- (1 → 6)-β-D-galactopyranosyl-(1 → 4)-
α,β-D-glucopyranose (HMO3)

Following the general procedure (B) and (C), compound 2-25 (183 mg, 0.098 
mmol) yielded the compound HMO3 (85 mg, 95%). $^1$H NMR (D$_2$O, 400 MHz): $\delta$ 5.14 (d, 
$J = 3.8$ Hz, 0.54 H, Glc-1 H-1 of α form), 4.54-4.63 (m, overlap with D$_2$O, 2.46 H, GlcNAc-
1, GlcNAc-2, H-1, Glc-1 H-1 of β form), 4.39 (d, $J = 7.8$ Hz, Gal-1, 1 H), 4.35 (d, $J = 7.8$
Hz, Gal-1, 1 H), 4.06 (d, $J = 3.1$ Hz, 1 H), 3.82-3.95 (m, 4 H), 3.42-3.82 (m, 22 H), 3.33-
3.42 (m, 2 H), 3.21 (t, $J = 8.39$ Hz, 1 H), 1.98 (s, 3 H), 1.96 (s, 3 H); $^{13}$C NMR (D$_2$O, 100 
MHz): $\delta$ 174.9, 174.5, 103.0, 102.9, 102.8 (GlcNAc-1, GlcNAc-2, Gal-2, C-1), 101.0 (Gal-
1, C-1), 95.7 (Glc-1, C-1 of β form), 91.8 (Glc-1, C-1 of α form), 81.7, 79.0, 78.9, 78.4, 
75.7, 75.3, 74.7, 74.7, 74.4, 73.9, 73.6, 73.5, 72.5, 72.4, 71.0, 69.9, 69.7, 68.7, 68.6, 68.4, 
61.0, 60.5, 60.0, 55.7, 55.0, 22.4, 22.2. HRMS: [M + Na]$^+$$^{C_{34}}H_{58}N_{2}NaO_{26}$ calcd for 
933.3175, found 933.3161.
^1^H NMR (D_2O, 500 MHz): δ 5.10 (d, J = 3.5 Hz, 0.61 H, Glc-1 H-1 of α form), 4.50-4.60 (m, 2.39 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of β form), 4.33-4.37 (m, 2 H, Gal-2 H-1, Gal-3 H-1), 4.31 (d, J = 7.9 Hz, 1 H, Gal-1 H-1), 4.02 (d, J = 3.1 Hz, 1 H), 3.78-3.79 (m, 6 H), 3.57-3.76 (m, 18 H), 3.53-3.57 (m, 2 H), 3.39-3.51 (m, 7 H), 3.17 (t, J = 8.3 Hz, 1 H), 1.94 (s, 3 H), 1.91 (s, 3 H). MALDI-MS: [M + Na]^+ C_{40}H_{68}N_{2}O_{31} calcd for 1095.370, found 1095.390.

^1^H NMR (D_2O, 500 MHz): δ 5.10 (d, J = 3.5 Hz, 0.51 H, Glc-1 H-1 of α form), 4.50-4.58 (m, 2.49 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of β form), 4.28-4.36 (m, 3 H, Gal-1 H-1, Gal-2 H-1, Gal-3 H-1), 4.03 (d, J = 3.0 Hz, 1 H), 3.38-3.91 (m, 47H), 3.17 (t, J = 8.3 Hz, 1 H), 2.52-2.57 (m, 2 H), 1.97 (s, 3 H), 1.93 (s, 3 H), 1.91 (s, 6 H), 1.55-1.63 (m, 2 H). ESI-MS: [M – 2H]^2- C_{62}H_{100}N_{4}O_{47} calcd for 826.2784, found 826.2760.
$^1$H NMR (D$_2$O, 500 MHz): $\delta$ 5.09 (d, $J = 3.7$ Hz, 0.29 H, Glc-1 H-1 of $\alpha$ form), 4.50-4.63 (m, 2.71 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of $\beta$ form), 4.28-4.36 (m, 3 H, Gal-1 H-1, Gal-2 H-1, Gal-3 H-1), 4.02 (d, $J = 3.0$ Hz, 1 H), 3.99 (s, 4 H), 3.38-3.91 (m, 48 H), 3.18 (t, $J = 8.3$ Hz, 1 H), 2.52-2.60 (m, 2 H), 1.97 (s, 3 H), 1.93 (s, 3 H), 1.55-1.65 (m, 2 H). ESI-MS: [M – 2H]$^2^-$ C$_{62}$H$_{100}$N$_4$O$_{49}$ calcd for 842.2734, found 842.2760.

HMO14

$^1$H NMR (D$_2$O, 500 MHz): $\delta$ 5.10 (d, $J = 3.5$ Hz, 0.33 H, Glc-1 H-1 of $\alpha$ form), 4.96-5.02 (m, 2 H, Fuc-1 H-1, Fuc-2 H-1), 4.50-4.61 (m, 2.67 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of $\beta$ form), 4.28-4.36 (m, 3 H, Gal-1 H-1, Gal-2 H-1, Gal-3 H-1), 4.01 (d, $J = 3.1$ Hz, 1 H), 3.63-3.91 (m, 25 H), 3.40-3.63 (m, 21 H), 3.35-3.39 (m, 2 H), 3.17 (t, $J = 8.4$ Hz, 1 H), 1.90-1.93 (d, $J = 6.5$ Hz, 6 H). MALDI-MS: [M + Na]$^+$ C$_{52}$H$_{88}$N$_2$NaO$_{39}$ calcd for 1387.487, found 1387.486.

HMO15

$^1$H NMR (D$_2$O, 500 MHz): $\delta$ 5.15-5.22 (m, 2 H, Fuc-1 H-1, Fuc-2 H-1), 5.10(d, $J = 3.5$ Hz, 0.33 H, Glc-1 H-1 of $\alpha$ form), 4.45-4.61 (m, 2.67 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of $\beta$ form), 4.39-4.45 (m, 2 H, Gal-2 H-1, Gal-3 H-1), 4.30 (d, $J = 7.9$ Hz, 1
H, Gal-1 H-1), 4.06-4.13 (m, 2 H), 4.01 (d, $J = 3.1$ Hz, 1 H), 3.41-3.91 (m, 44 H), 3.30-3.39 (m, 2 H), 3.16 (t, $J = 8.4$ Hz, 1 H), 1.93 (s, 3 H), 1.91 (s, 3 H), 1.09-1.11 (d, $J = 6.7$ Hz, 6 H). MALDI-MS: $[\text{M + Na}^+]$ $C_{52}H_{88}N_2NaO_{39}$ calcd for 1387.487, found 1387.490.

$^1$H NMR ($D_2O$, 500 MHz): δ 5.12-5.18 (m, 2 H, Fuc-2 H-1, Fuc-4 H-1), 5.10 (d, $J = 3.5$ Hz, 0.40 H, Glc-1 H-1 of α form), 4.93-5.01 (m, 2 H, Fuc-1 H-1, Fuc-3 H-1), 4.70-4.78 (m, 2 H), 4.45-4.61 (m, 2.60 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of β form), 4.35-4.41 (m, 2 H, Gal-2 H-1, Gal-3 H-1), 4.31 (d, $J = 7.9$ Hz, 1 H, Gal-1 H-1), 4.09-4.16 (m, 2 H), 4.00 (d, $J = 3.1$ Hz, 1 H), 3.41-3.91 (m, 44 H), 3.28-3.38 (m, 2 H), 3.16 (t, $J = 8.4$ Hz, 1 H), 1.93 (s, 3 H), 1.90 (s, 3 H), 1.10-1.15 (m, 12 H). MALDI-MS: $[\text{M + Na}^+]$ $C_{64}H_{108}N_2NaO_{47}$ calcd for 1679.602, found 1679.607.

$^1$H NMR ($D_2O$, 500 MHz): δ 5.09 (d, $J = 3.6$ Hz, 0.22 H, Glc-1 H-1 of α form), 4.45-4.61 (m, 2.78 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of β form), 4.26-4.36 (m, 2 H, Gal-1 H-1, Gal-2 H-1), 4.17-4.22 (m, 1 H), 4.09-4.16 (m, 2 H), 4.00 (d, $J = 3.1$ Hz, 1 H), 3.36-3.90 (m, 32 H), 3.26-3.36 (m, 2 H), 3.16 (t, $J = 8.4$ Hz, 1 H), 2.50-2.56 (m, 1 H),
1.92 (s, 3 H), 1.91 (s, 3 H), 1.89 (s, 3 H), 1.58 (t, \( J = 12.3 \) Hz, 1 H). ESI-MS: \([M - H]\) 

C_{45}H_{74}N_{3}O_{34} calcd for 1200.4159, found 1200.4190.

\( ^1H \) NMR (D$_2$O, 500 MHz): \( \delta \) 5.09 (d, \( J = 3.6 \) Hz, 0.27 H, Glc\(-1\) H\(-1\) of \( \alpha \) form), 4.47-4.62 (m, 2.73 H, GlcNAc\(-1\) H\(-1\), GlcNAc\(-2\) H\(-1\), Glc\(-1\) H\(-1\) of \( \beta \) form), 4.28-4.35 (m, 2 H, Gal\(-1\) H\(-1\), Gal\(-2\) H\(-1\)), 4.06-4.16 (m, 1 H), 4.01 (d, \( J = 3.1 \) Hz, 1 H), 3.99 (s, 2 H), 3.38-3.90 (m, 33 H), 3.29-3.37 (m, 2 H), 3.16 (t, \( J = 8.4 \) Hz, 1 H), 2.53-2.58 (m, 1 H), 1.93 (s, 3 H), 1.92 (s, 3 H), 1.60 (t, \( J = 12.3 \) Hz, 1 H). ESI-MS: \([M - H]\) 

C_{45}H_{74}N_{3}O_{35} calcd for 1216.4108, found 1216.4129.

\( ^1H \) NMR (D$_2$O, 500 MHz): \( \delta \) 5.09 (d, \( J = 3.6 \) Hz, 0.36 H, Glc\(-1\) H\(-1\) of \( \alpha \) form), 5.00 (d, \( J = 3.9 \) Hz, 1 H, Fuc\(-1\) H\(-1\)), 4.46-4.60 (m, 2.64 H, GlcNAc\(-1\) H\(-1\), GlcNAc\(-2\) H\(-1\), Glc\(-1\) H\(-1\) of \( \beta \) form), 4.28-4.36 (m, 2 H, Gal\(-1\) H\(-1\), Gal\(-2\) H\(-1\)), 4.02 (d, \( J = 3.1 \) Hz, 1 H), 3.40-3.90 (m, 31 H), 3.29-3.40 (m, 2 H), 3.16 (t, \( J = 8.4 \) Hz, 1 H), 2.50-2.56 (m, 1 H), 1.94 (s, 3 H), 1.90 (s, 3 H), 1.05 (d, \( J = 6.7 \) Hz, 3 H). MALDI-MS: \([M + Na]^+\) 

C_{40}H_{68}N_{2}NaO_{30} calcd for 1079.376, found 1079.371.
$^1$H NMR (D$_2$O, 500 MHz): δ 5.19 (d, J = 2.6 Hz, 1 H, Fuc-1 H-1), 5.09 (d, J = 3.6 Hz, 0.50 H, Glc-1 H-1 of α form), 4.46-4.60 (m, 2.50 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of β form), 4.42 (d, J = 7.8 Hz, 1 H, Gal-2 H-1), 4.31 (d, J = 8.0 Hz, 1 H, Gal-1 H-1), 4.07-4.12 (m, 1 H), 4.01 (d, J = 3.1 Hz, 1 H), 3.41-3.90 (m, 28 H), 3.29-3.38 (m, 3 H), 3.16 (t, J = 8.4 Hz, 1 H), 1.94 (s, 3 H), 1.91 (s, 3 H), 1.10 (d, J = 6.5 Hz, 3 H). MALDI-MS: [M + Na]$^+$ C$_{40}$H$_{68}$N$_2$O$_{30}$ calcd for 1079.376, found 1079.379.

$^1$H NMR (D$_2$O, 500 MHz): δ 5.09 (d, J = 3.9 Hz, 0.28 H, Glc-1 H-1 of α form), 4.48-4.62 (m, 2.72 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of β form), 4.28-4.37 (m, 3 H, Gal-1 H-1, Gal-2 H-1, Gal-3 H-1), 4.01 (d, J = 3.1 Hz, 1 H), 3.38-3.90 (m, 41 H), 3.17 (t, J = 8.4 Hz, 1 H), 2.50-2.57 (m, 1 H), 1.93 (s, 3 H), 1.92 (s, 3 H), 1.90 (s, 3 H), 1.59 (t, J = 11.9 Hz, 1 H). ESI-MS: [M − H]$^-$ C$_{51}$H$_{84}$N$_3$O$_{39}$ calcd for 1362.4687, found 1362.4759.
$^1$H NMR (D$_2$O, 500 MHz): $\delta$ 5.10 (d, $J = 3.7$ Hz, 0.51 H, Glc-1 H-1 of $\alpha$ form), 4.50-4.63 (m, 2.49 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of $\beta$ form), 4.29-4.37 (m, 3 H, Gal-1 H-1, Gal-2 H-1, Gal-3 H-1), 4.02 (d, $J = 2.9$ Hz, 1 H), 3.99 (s, 2 H), 3.39-3.90 (m, 41 H), 3.17 (t, $J = 8.3$ Hz, 1 H), 2.54-2.59 (m, 1 H), 1.94 (s, 3 H), 1.93 (s, 3 H), 1.61 (t, $J = 12.2$ Hz, 1 H). ESI-MS: [M – H]$^-$ C$_{51}$H$_{84}$N$_3$O$_{39}$ calcd for 1378.4637, found 1378.4660.

$^1$H NMR (D$_2$O, 500 MHz): $\delta$ 5.10 (d, $J = 3.7$ Hz, 0.55 H, Glc-1 H-1 of $\alpha$ form), 5.00 (d, $J = 4.0$ Hz, 1 H, Fuc-1 H-1), 4.69-4.74 (m, 1 H), 4.50-4.61 (m, 2.45 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of $\beta$ form), 4.28-4.38 (m, 3 H, Gal-1 H-1, Gal-2 H-1, Gal-3 H-1), 4.02 (d, $J = 2.9$ Hz, 1 H), 3.34-3.91 (m, 32 H), 3.17 (t, $J = 8.3$ Hz, 1 H), 2.54-2.59 (m, 1 H), 1.94 (s, 3 H), 1.90 (s, 3 H), 1.05 (d, $J = 6.6$ Hz, 3 H). MALDI-MS: [M + Na]$^+$ C$_{46}$H$_{78}$N$_2$NaO$_{39}$ calcd for 1241.428, found 1241.432.

$^1$H NMR (D$_2$O, 500 MHz): $\delta$ 5.19 (d, $J = 2.4$ Hz, 1H, Fuc-1 H-1), 5.10 (d, $J = 3.7$ Hz, 0.42 H, Glc-1 H-1 of $\alpha$ form), 4.48-4.60 (m, 2.58 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-
1 H-1 of β form), 4.28-4.45 (m, 3 H, Gal-1 H-1, Gal-2 H-1, Gal-3 H-1), 4.06-4.15 (m, 1 H), 4.01 (d, \(J = 2.9\) Hz, 1 H), 3.38-3.90 (m, 31 H), 3.31-3.37 (m, 1 H), 3.17 (t, \(J = 8.8\) Hz, 1 H), 1.93 (s, 3 H), 1.91 (s, 3 H), 1.10 (d, \(J = 6.5\) Hz, 3 H). MALDI-MS: [M + Na]\(^+\) \(C_{46}H_{78}N_2NaO_{35}\) calcd for 1241.428, found 1241.423.

\(\text{HMO29}\)

1\(^1\)H NMR (D\(_2\)O, 500 MHz): \(\delta\) 5.15 (d, \(J = 2.6\) Hz, 1 H, Fuc-2 H-1), 5.09 (d, \(J = 3.7\) Hz, 0.36 H, Glc-1 H-1 of α form), 4.99 (d, \(J = 3.7\) Hz, 1 H, Fuc-1 H-1), 4.72-4.78 (m, 1 H), 4.60-4.61 (m, 2.64 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of β form), 4.42 (d, \(J = 7.8\) Hz, 1 H, Gal-2 H-1), 4.31 (d, \(J = 8.0\) Hz, 1 H, Gal-1 H-1), 4.10-4.16 (m, 1 H), 4.01 (d, \(J = 2.9\) Hz, 1 H), 3.40-3.92 (m, 33 H), 3.32-3.37 (m, 3 H), 3.16 (t, \(J = 8.4\) Hz, 1 H), 1.94 (s, 3 H), 1.90 (s, 3 H), 1.09-1.16 (m, 6 H). MALDI-MS: [M + Na]\(^+\) \(C_{46}H_{78}N_2NaO_{34}\) calcd for 1225.433, found 1225.438.

\(\text{HMO210}\)

1\(^1\)H NMR (D\(_2\)O, 500 MHz): \(\delta\) 5.16 (d, \(J = 2.6\) Hz, 1 H, Fuc-2 H-1), 5.10 (d, \(J = 3.5\) Hz, 0.40 H, Glc-1 H-1 of α form), 4.93-5.01 (m, 2 H, Fuc-1 H-1, Fuc-3 H-1), 4.67-4.79 (m, 2 H), 4.45-4.61 (m, 2.60 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of β form), 4.26-4.41 (m, 3 H, Gal-1 H-1, Gal-2 H-1, Gal-3 H-1), 4.09-4.16 (m, 1 H), 4.00 (d, \(J = 3.1\) Hz, 1 H),
3.41-3.91 (m, 44 H), 3.28-3.40 (m, 3 H), 3.16 (t, J = 8.4 Hz, 1 H), 1.93 (s, 3 H), 1.90 (s, 3 H), 1.02-1.17 (m, 9 H). MALDI-MS: [M + Na]$^+$ C$_{46}$H$_{78}$N$_2$O$_{34}$ calcd for 1533.544, found 1533.538.

$^1$H NMR (D$_2$O, 500 MHz):  δ 5.16 (d, J = 2.7 Hz, 1 H, Fuc-2 H-1), 5.10 (d, J = 3.7 Hz, 0.47 H, Glc-1 H-1 of α form), 4.99 (d, J = 3.7 Hz, 1 H, Fuc-1 H-1), 4.72-4.78 (m, 1 H), 4.49-4.62 (m, 2.53 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of β form), 4.29-4.41 (m, 3 H, Gal-1 H-1, Gal-2 H-1, Gal-3 H-1), 4.10-4.16 (m, 1 H), 4.01 (d, J = 3.2 Hz, 1 H), 3.77-3.91 (m, 8 H), 3.39-3.77 (m, 35 H), 3.30-3.36 (m, 1 H), 3.17 (t, J = 8.4 Hz, 1 H), 1.94 (s, 3 H), 1.90 (s, 3 H), 1.09-1.16 (m, 6 H). MALDI-MS: [M + Na]$^+$ C$_{52}$H$_{88}$N$_2$O$_{39}$ calcd for 1387.486, found 1387.490.

$^1$H NMR (D$_2$O, 500 MHz):  δ 5.09(d, J = 3.7 Hz, 0.31 H, Glc-1 H-1 of α form), 4.47-4.58 (m, 2.69 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of β form), 4.28-4.34 (m, 2 H, Gal-1 H-1, Gal-2 H-1), 4.06-4.16 (m, 1 H), 4.02 (d, J = 3.1 Hz, 1 H), 3.28-3.91 (m, 35 H), 3.26-3.36 (m, 2 H), 3.18 (t, J = 8.4 Hz, 1 H), 2.51-2.57 (m, 1 H), 1.96 (s, 3 H), 1.92 (s, 3 H), 1.91 (s, 3 H), 1.59 (t, J = 12.2 Hz, 1 H). ESI-MS: [M − H]$^-$ C$_{45}$H$_{74}$N$_3$O$_{34}$ calcd for 1200.4159, found 1200.4122.
$^1$H NMR (D$_2$O, 500 MHz): $\delta$ 5.09 (d, $J = 3.9$ Hz, 0.28 H, Glc-1 H-1 of $\alpha$ form), 4.50-4.58 (m, 2.72 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of $\beta$ form), 4.28-4.34 (m, 2 H, Gal-1 H-1, Gal-2 H-1), 4.02 (d, $J = 3.2$ Hz, 1 H), 3.28-3.91 (m, 35 H), 3.18 (t, $J = 8.3$ Hz, 1 H), 2.53-2.58 (m, 1 H), 1.96 (s, 3 H), 1.91 (s, 3 H), 1.59 (t, $J = 12.2$ Hz, 1 H). ESI-MS: [M – H]· $C_{45}H_{74}N_3O_{35}$ calcd for 1216.4108, found 1216.4139.

$^1$H NMR (D$_2$O, 500 MHz): $\delta$ 5.09 (d, $J = 3.7$ Hz, 0.19 H, Glc-1 H-1 of $\alpha$ form), 4.98 (d, $J = 3.9$ Hz, 1 H, Fuc-1 H-1), 4.66-4.73 (m, 1 H), 4.48-4.58 (m, 2.81 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of $\beta$ form), 4.27-4.35 (m, 2 H, Gal-1 H-1, Gal-2 H-1), 4.02 (d, $J = 3.1$ Hz, 1 H), 3.28-3.90 (m, 38 H), 3.16 (t, $J = 8.4$ Hz, 1 H), 1.92 (s, 3 H), 1.90 (s, 3 H), 1.04 (d, $J = 6.6$ Hz, 3 H). MALDI-MS: [M + Na]$^+$ $C_{40}H_{68}N_2NaO_{30}$ calcd for 1079.376, found 1079.380.
$^1$H NMR (D$_2$O, 500 MHz): $\delta$ 5.18 (d, $J = 2.6$ Hz, 1 H, Fuc-1 H-1), 5.09 (d, $J = 3.9$ Hz, 0.41 H, Glc-1 H-1 of $\alpha$ form), 4.66-4.73 (m, 1 H), 4.45-4.58 (m, 2.59 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of $\beta$ form), 4.39-4.44 (m, 2 H, Gal-1 H-1, Gal-2 H-1), 4.31 (d, $J = 7.5$ Hz, 1 H), 4.06-4.12 (m, 2 H), 4.02 (d, $J = 3.1$ Hz, 1 H), 3.39-3.90 (m, 28 H), 3.28-3.38 (m, 3 H), 3.16 (t, $J = 8.4$ Hz, 1 H), 1.94 (s, 3 H), 1.91 (s, 3 H), 1.11 (d, $J = 6.7$ Hz, 3 H). MALDI-MS: [M + Na]$^+$ C$_{40}$H$_{68}$N$_2$NaO$_{30}$ calcd for 1079.376, found 1079.372.

$^1$H NMR (D$_2$O, 500 MHz): $\delta$ 5.09 (d, $J = 3.8$ Hz, 0.53 H, Glc-1 H-1 of $\alpha$ form), 4.51-4.60 (m, 2.47 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of $\beta$ form), 4.28-4.37 (m, 3 H, Gal-1 H-1, Gal-2 H-1, Gal-3 H-1), 4.02 (d, $J = 3.0$ Hz, 1 H), 3.38-3.91 (m, 41 H), 3.18 (t, $J = 8.4$ Hz, 1 H), 2.50-2.57 (m, 1 H), 1.96 (s, 3 H), 1.91 (s, 6 H), 1.59 (t, $J = 12.0$ Hz, 1 H). ESI-MS: [M – H]$^-$ C$_{51}$H$_{84}$N$_3$O$_{39}$ calcd for 1362.4687, found 1362.4799.

$^1$H NMR (D$_2$O, 500 MHz): $\delta$ 5.09 (d, $J = 3.8$ Hz, 0.52 H, Glc-1 H-1 of $\alpha$ form), 4.50-4.60 (m, 2.48 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of $\beta$ form), 4.28-4.38 (m, 3 H, Gal-1 H-1, Gal-2 H-1, Gal-3 H-1), 4.02 (d, $J = 3.0$ Hz, 1 H), 3.99 (s, 2 H), 3.38-3.91 (m, 41 H), 3.18 (t, $J = 8.4$ Hz, 1 H), 2.53-2.59 (m, 1 H), 1.96 (s, 3 H), 1.91 (s, 3 H), 1.59 (t, $J = 12.0$ Hz, 1 H). ESI-MS: [M – H]$^-$ C$_{51}$H$_{84}$N$_3$O$_{39}$ calcd for 1378.4637, found 1378.4609.
$^1$H NMR ($\text{D}_2\text{O}$, 500 MHz): $\delta$ 5.09 (d, $J = 3.8$ Hz, 0.34 H, Glc-1 H-1 of $\alpha$ form), 4.98 (d, $J = 4.0$ Hz, 1 H, Fuc-1 H-1), 4.67-4.74 (m, 1 H), 4.50-4.61 (m, 2.66 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of $\beta$ form), 4.27-4.38 (m, 3 H, Gal-1 H-1, Gal-2 H-1, Gal-3 H-1), 4.01 (d, $J = 3.1$ Hz, 1 H), 3.30-3.90 (m, 37 H), 3.16 (t, $J = 8.6$ Hz, 1 H), 1.92 (s, 3 H), 1.91 (s, 3 H), 1.05 (d, $J = 6.6$ Hz, 3 H). MALDI-MS: [M + Na]$^+$ C$_{46}$H$_{78}$N$_2$NaO$_{35}$ calcd for 1241.428, found 1241.436.

$^1$H NMR ($\text{D}_2\text{O}$, 500 MHz): $\delta$ 5.18 (d, $J = 2.6$ Hz, 1 H, Fuc-1 H-1), 5.09 (d, $J = 3.8$ Hz, 0.69 H, Glc-1 H-1 of $\alpha$ form), 4.46-4.60 (m, 2.31 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of $\beta$ form), 4.28-4.43 (m, 3 H, Gal-1 H-1, Gal-2 H-1, Gal-3 H-1), 4.07-4.13 (m, 1 H), 4.02 (d, $J = 3.2$ Hz, 1 H), 3.39-3.89 (m, 36 H), 3.33-3.38 (m, 1 H), 3.17 (t, $J = 8.4$ Hz, 1 H), 1.94 (s, 3 H), 1.91 (s, 3 H), 1.11 (d, $J = 6.5$ Hz, 3 H). MALDI-MS: [M + Na]$^+$ C$_{46}$H$_{78}$N$_2$NaO$_{35}$ calcd for 1241.428, found 1241.435.
\(^1\text{H NMR (D}_2\text{O, 500 MHz):}\) \(\delta\) 5.16 (d, \(J = 2.9\) Hz, 1 H, Fuc-2 H-1), 5.10 (d, \(J = 3.7\) Hz, 0.39 H, Glc-1 H-1 of \(\alpha\) form), 4.97 (d, \(J = 4.0\) Hz, 1 H, Fuc-1 H-1), 4.72-4.78 (m, 1 H), 4.45-4.58 (m, 2.61 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of \(\beta\) form), 4.28-4.40 (m, 2 H, Gal-1 H-1, Gal-2 H-1), 4.10-4.16 (m, 1 H), 4.02 (d, \(J = 3.06\) Hz, 1 H), 3.40-3.93 (m, 33 H), 3.29-3.38 (m, 3 H), 3.16 (t, \(J = 8.4\) Hz, 1 H), 1.93 (s, 3 H), 1.91 (s, 3 H), 1.08-1.17 (m, 6 H). MALDI-MS: \([M + Na]^+\) \(C_{46}H_{78}N_2NaO_{34}\) calcd for 1225.433, found 1225.430.

\[\text{HMO310}\]

\(^1\text{H NMR (D}_2\text{O, 500 MHz):}\) \(\delta\) 5.16 (d, \(J = 3.0\) Hz, 1 H, Fuc-3 H-1), 5.09 (d, \(J = 3.7\) Hz, 0.62 H, Glc-1 H-1 of \(\alpha\) form), 4.96-5.01 (m, 2 H), 4.69-4.78 (m, overlap with D\(_2\)O, 2 H), 4.47-4.61 (m, 2.38 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of \(\beta\) form), 4.28-4.40 (m, 3 H, Gal-1 H-1, Gal-2 H-1, Gal-3 H-1), 4.10-4.16 (m, 1 H), 4.02 (d, \(J = 3.2\) Hz, 1 H), 3.32-3.93 (m, 43 H), 3.16 (t, \(J = 8.4\) Hz, 1 H), 1.93 (s, 3 H), 1.90 (s, 3 H), 1.02-1.17 (m, 9 H). MALDI-MS: \([M + Na]^+\) \(C_{46}H_{78}N_2NaO_{34}\) calcd for 1533.544, found 1533.550.

\[\text{HMO311}\]

\(^1\text{H NMR (D}_2\text{O, 500 MHz):}\) \(\delta\) 5.16 (d, \(J = 3.2\) Hz, 1 H, Fuc-2 H-1), 5.10 (d, \(J = 3.6\) Hz, 0.67 H, Glc-1 H-1 of \(\alpha\) form), 4.97 (d, \(J = 3.9\) Hz, 1 H, Fuc-1 H-1), 4.72-4.77 (m, 1 H), 4.47-4.60 (m, 2.37 H, GlcNAc-1 H-1, GlcNAc-2 H-1, Glc-1 H-1 of \(\beta\) form), 4.29-4.39 (m,
3 H, Gal-1 H-1, Gal-2 H-1, Gal-3 H-1), 4.10-4.16 (m, 1 H), 4.02 (d, \( J = 3.1 \) Hz, 1 H), 3.39-3.93 (m, 41 H), 3.30-3.38 (m, 1 H), 3.16 (t, \( J = 8.4 \) Hz, 1 H), 1.94 (s, 3 H), 1.91 (s, 3 H), 1.09-1.17 (m, 6 H). MALDI-MS: [M + Na]+ \( C_{52}H_{88}N_2NaO_{39} \) calcd for 1387.486, found 1387.478.

This chapter is under review by Journal of Organic Chemistry of American Chemical Society.

2.8 References

26. Ruiz-Palacios, G. M.; Cervantes, L. E.; Ramos, P.; Chavez-Munguia, B.; Newburg, D. S., Campylobacter jejuni binds intestinal H(O) antigen (Fuc alpha 1, 2Gal beta 1, 4GlcNAc), and fucosyloligosaccharides of human milk inhibit its binding and infection. *Journal of Biological Chemistry* 2003, 278, 14112-14120.
3. Glycosylation of Protein-based Magnetic Resonance Imaging Contrast Agents

3.1 Introduction

Magnetic resonance imaging (MRI) is a promising imaging technique that provides three dimensional images of organs and structures inside the body by exploiting the magnetic properties of water hydrogens and their interactions with an external magnetic field. Besides, MRI is a noninvasive technique which has limited side-effects and does not require the use of radioactive probes. Owing to these advantages, MRI has become one of the most prevalent imaging modalities in clinical and preclinical tests. MRI also may show problems that cannot be seen with other imaging methods. It has broad applicability in diagnosing and following the treatment of different types of cancers and other human diseases, including various cardiovascular diseases, liver diseases, and neurodegenerative disorders.¹⁻⁹

In addition to its superior soft tissue contrast, MRI also has high resolution (approximately 50–100 μm) but relatively low sensitivity. To enhance the contrast and sensitivity of imaging, 35% of clinical MRI scans require the injection of Gd³⁺-based (where Gd³⁺ is gadolinium) MRI contrast agents.¹⁰ Current MRI contrast agents are based on paramagnetic, ferromagnetic, or super paramagnetic metal ions.

Currently, clinically approved Gd³⁺-based MRI contrast agents are formed by one Gd³⁺ ion encapsulated by organic chelators, such as diethylene triamine pentaacetic acid (DTPA), 1, 4, 7, 10-tetraazacyclododecane-1, 4, 7, 10-tetraacetic acid (DOTA) and their derivatives.¹¹⁻¹⁶ Figure 3.1 showed some examples of clinical MRI contrast reagents. There are eight nitrogen and oxygen from the chelator binding as Gd³⁺ ion ligands. In
addition, Gd$^{3+}$ ion in clinical MRI contrast agents coordinated with one water molecule through robust Gd–O interactions.

![Chemical structures of contrast agents](image)

**Figure 3.1. Examples of Clinical Gd$^{3+}$-Based MRI Contrast Agents.**

Even if current clinical MRI contrast agents carry many advantages, their shortage is also obvious. Clinical contrast agents have a shorter half-life time for approximately 0.5–3 min in the blood vessels of mice and 1.5 hours in patients, due to their small molecular mass.$^{17, 18}$ Optimizing half-life and pharmacokinetic/pharmacodynamic (PK/PD) properties of these contrast agents would lengthen the time window for MRI data collection and eliminate repeated dose injections. The Food and Drug Administration (FDA) approved contrast agents have poor organ specificity and no targeting capability. Injection of targeted protein-based contrast agent can vastly improve diagnosis capabilities using MRI by increasing the image enhancement of the specified area, increasing relaxivity, rotational correlation time, and having better retention. Jenny developed a new generation of protein-based MRI contrast agents with longer half-life
time, higher resolution and sensitivity. The protein, namely ProCA1, could bind to Gd$^{3+}$ by using amino acid residues as metal coordinating ligands. (Figure 3.2)

Hepatocellular carcinoma (HCC) is one of the most common cancers around the world. Symptoms and signs of HCC include abdominal pain, swelling and weakness, loss of appetite, jaundice, fever, or enlargement of liver, which mainly arise in the later stages of cancer making early detection or screening highly demanded because the five-year survival rate for liver cancer is currently 15%. Early detection could dramatically increase the five-year survival rate to 60-70% with liver transplant. Treatment options for patients include liver resection or liver transplant depending on the level of cirrhosis in the liver. Early detection is key for increasing disease-free survival rates for patients diagnosed with this disease.

Figure 3.2. A model structure of ProCA1 with Gd$^{3+}$ and relaxivity value r1 and r2 of ProCA1

(Left) Gd$^{3+}$ (purple sphere) chelated by oxygen ligand residues E15, D56, D58, D62 and D64. Exchangeable waters (red spheres) are from inner sphere, protein surface and hydration water from added glycosylation chain. (Right) Relaxivity value $r_1$ and $r_2$ of ProCA1 are 14 - 30 fold greater than clinical approved liver contrast agents at 25 °C. The stability constant for Gd of ProCA3 is comparable with DTPA but with 100 to $10^{11}$ fold greater metal selectivity and kinetic stability than DTPA for physiological metal ions such as Zn$^{2+}$, Mg$^{2+}$ and Ca$^{2+}$. 
The asialoglycoprotein receptor (ASGPR) is located on hepatocytes and is a Ca$^{2+}$-dependent carbohydrate binding protein, or C-type lectin. It is expressed on mammalian liver cells. The main function is to maintain serum glycoprotein homeostasis by the recognition, binding, and endocytosis of asialoglycoproteins (ASGPs), which is desialylated glycoproteins with terminal Gal or GalNAc residues. After internalization via clathrin-coated pits and their fusion with endosomes, the ASGPs are released in the acidic environment of the endosome and transported to lysosomes for degradation, while the receptor is recycled back to the cell surface. Baenzinger have shown that the human receptor exhibits specificity for terminal Gal and GalNAc on desialylated glycoproteins. Triantennary ligands displayed a higher affinity than their mono- and diantennary counterparts. Furthermore, the studies led to the conclusion that only the terminal residues are necessary for specific recognition, and that the binding process proceeds through a simultaneous interaction of 2–3 sugar residues with 2–3 binding sites of the heterooligomeric receptor. On the native receptor on the hepatocyte surface these binding sites are 25–30 Å apart.

Lee further reinforced the binding hierarchy of polyvalent ligands: tetraantennary > triantennary >> diantennary >> monoantennary by studies on rabbit hepatocytes using synthetic oligosaccharides. The IC$_{50}$-values from monoantennary to tetraantennary oligosaccharides were found to be approximately $1 \times 10^{-3}$, $1 \times 10^{-6}$, $5 \times 10^{-9}$ and $1 \times 10^{-9}$ M, respectively. In other words, the inhibitory potency increased $1 \times 10^{6}$-fold although the number of Gal residues of ligand increased only 4-fold. Since the fourth Gal moiety on the tetraantennary ligand does not markedly increase the affinity, it was assumed that the
3.2 Results and Discussion

In the collaboration program with Jenny’s group, we have recently developed an asialoglycoprotein receptor (ASGPR)-targeted protein contrast agent lactose-ProCA1 by modifying ProCA1 with a lactose moiety. We have achieved the first observation of significant liver MRI enhancement and retention via tail vein injection of lactose-ProCA1, which is consistent with high liver expression of ASGPR. The goals of this research are to develop targeted ProCAs to enable accurate monitoring of the expression and distribution of lectins and glycans in living animals and cancer models.

Figure 3.3. Binding model for ASGPR ligands in an optimal conformation to the heterooligomeric receptor.
Synthesis and Evaluation of Lactose Conjugated Protein

The glycosylation of ProCA1 was performed with Lactose NHS donor. The major advantages of this method includes: 1. NHS active ester is easy to achieve with TSTU/Et3N condition; 2. The active ester carries high reactivity and it could quickly couple with amine groups on small molecules as well as macromolecules, including protein and DNA; 3. No coupling reagent is needed, which simplifies purification process. In order to achieve lactose-ProCA1 conjugate, the details of synthesis was showed in Scheme 3.1.

Scheme 3.1. Chemical Synthesis of Lac-rProCA1.

To functionalize lactose moiety, an amine group need to be installed at the reducing end. The Kochetkov reaction assisted with microwave irradiation was used to furnish compound 3-2. The lactose 3-1 was mixed with 15 eq. of ammonium carbonate in dry DMSO and mixture was radiated at 30-40 °C for 4 hours. Compound 3-2 was furnished at a yield of 85%. After eliminating excess ammonium carbonate, the solution was used for next step reaction without further purification. Compound 3-2 and 4-tert-Butoxy-4-oxobutanoic acid were coupled to form compound 3-3 smoothly with treatment of HATU and HOBT under basic condition. Then deprotection of t-butyl group of 3-3 using TFA/Anisole condition to provide compound 3-4 in 90% yield. To transform 3-4 to NHS
active ester 3-5, O-(N-Suc-cinimidyl)-1, 1, 3, 3-tetramethyl-uronium tetrafluoroborate (TSTU) was used as a coupling reagent under basic condition to achieve a yield of 80%, with a mild amount of decomposed byproduct, which will not affect the following conjugation with proteins.

With these NHS Lactose linkers ready, we started investigating the coupling reaction with proteins. The glycosylation of ProCA1 was performed in PBS buffer solution with PH=7.0. In the neutral condition, conjugation could proceed smoothly and hydrolysis of active ester could be inhibited well, which extend the coupling time and increase coupling yield. We tested various ratio of Lactose to ProCA1 from 10:1 to 100:1. It was found the coupling efficiency is not very high, and majority of NHS lactose linkers was hydrolyzed in buffer solution. Therefore, we have to set the NHS lactose linkers: ProCA1 as high as 100: 1 in order to get full glycosylated proteins. (Figure 3.4) One of the reasons accounting for this might be that the linker is too short to make it more difficult to interact with amine on the protein. In order to achieve higher coupling efficiency, design and synthesis of longer linkers is necessary.
Glycosylation of ProCA1 for Improved Stability and Targeting

Serum stability studies for rProCA1 and Lac-rProCA1 were conducted. In this study, previously obtained human serum was used and described in Figure 3.5 for the studies of rProCA1 without and with lysine modification. In this figure, ProCA1 without modification (Figure 3.5a) is stable up to 2 days; however, ProCA1 with glycosylation (Figure 3.5b) it has been found to be stable in human serum up to 6 days. These results clearly proved that Glycosylation could significantly improve the stability of proteins in vitro. Glycosylation is a promising way to increase half-life time of many therapeutic proteins.
Figure 3.5. Serum Stability of rProCA1 and Lac-rProCA1

MRI of rProCA1 without modification was performed and the results of this experiment is shown below in Figure 3.6a. In this figure, slight contrast is seen 30 minutes and 3 hours after injection of the contrast agent; however, in comparison to lactose modified ProCA1-PEG, the intensity of the contrast is drastically increased (Figure 3.6b). The biodistribution of Lac-rProCA1 has obviously changed and and uptake of the contrast agent can be seen in the liver region, but not in the kidney. This exciting discovery of Glyco-rProCA1 clearance through the liver provides capability of assigning any possible tumors in the liver.
The comparison of contrast to noise ratio in Figure 3.7a clearly summarizes the impact of lysine modification on ProCA improving relaxivity in both the liver and kidney regions of interest, ROI, for Glyco-rProCA1 which in turn increases signal intensity. Signal is measured as a mean in the ROI and contrast is found by looking at the difference of the signal in two regions. Contrast to noise ratio is created by differences in the signal and noise in the ROI. In the Lac-ProCA1 shows greater uptake in the liver three hours after injection rather than the kidney. Lac-rProCA1 contains not only hydrophilic characteristics, but also a lactose targeting moiety which could bind to ASGPR in the liver. This targeting moiety and its interaction with ASGPR is pertinent for excretion out the liver rather than the kidney region.

ICP-OES can be used to detect gadolinium concentration in specific organs. This is an analytical technique that excites atoms and ions to detect emission at wavelengths that are specific to an element. The detection of gadolinium in organs indicates the distribution of lysine modified and non-modified ProCA1. The results of this are shown in Figure 3.7b.
This results makes it clear that Lac-ProCA1 shows an uptake in the liver in order to confirm the biodistribution to the liver. It is interesting to note the uptake of the contrast agents in other regions of the body indicative of a possible whole body imaging capability to aid in the detection of targeted carcinomas using the modified contrast agent.

**a). MRI Contrast ROI of Lac-rProCA1**

**Figure 3.7. MRI contrast of Lac-rProCA1 in mice**

### 3.3 Conclusion

Overall, these results demonstrated that our strategy has led to the achievement of lactose conjugated ProCA1 with intacted biological activity and will provide us valuable experience for the future glycoprotein synthesis. These results about serum stability clearly proved that glycosylation could significantly improve the stability of proteins *in vitro*. Glycosylation is a promising way to increase half-life time of many therapeutic proteins. Besides, the liver enhancement by Lac-rProCA1 is significantly higher than ProCA1, which proved its strong targeting capability to ASGPR. Lac-rProCA1 is a good candidate for regional contrast agent targeting liver.
3.4 Experimental Section

Compound 3-3

Compound 3-3 (400 mg) and ammonium carbonate (5 g) were dissolved in 20 mL DMSO and stirred with microwave tube at 40 °C for 3 hours. TLC showed major lactose was transformed to compound 3-2. The excess ammonium carbonate was then removed under reduced pressure and the mixture was used for next step’s reaction without further purification.

HATU (889 mg, 1.17 mmol), HOBT (315 mg, 1.17 mmol) and DIPEA (0.81 mL, 4.68 mmol) were sequentially added to the solution of 3-2 in DMSO. The reaction mixture was stirred at room temperature under nitrogen atmosphere for 24 hours. The solvent was removed under reduced pressure. The crude product was purified by C18 reverse phase flash chromatography using 0-100% MeOH in water as eluent to afford compound 3-3 as white solid (441 mg, 76% over two steps). \(^1\)H NMR (D\(_2\)O, 400 MHz): \(\delta\) 4.90 (d, \(J = 9.24\) Hz, 1 H), 4.38 (d, \(J = 7.76\) Hz, 1 H), 3.86-3.83 (m, 2 H), 3.74 (t, 1 H), 3.71 (t, 1 H), 3.69 (m, 2 H), 3.66-3.57 (m, 5 H), 3.47 (dd, 1 H), 3.36 (m, 1 H), 2.73 (s, 1 H), 2.64 (s, 1 H), 2.52 (s, 2 H), 1.36 (s, 9 H). \(^1\)\(^3\)C NMR (100 MHz, D\(_2\)O): \(\delta\) 176.13, 174.22, 102.87, 82.83, 76.26, 75.09, 70.94, 68.56, 61.05, 59.88, 59.87, 38.72, 38.00, 30.33, 30.24, 27.18. HRMS: [M + Na]\(^+\) C\(_{20}\)H\(_{35}\)NNaO\(_{13}\) calcd for 520.2006, found 520.2026.

Compound 3-4

Compound 3-3 (200 mg, 0.40 mmol) was added to TFA : DCM : Anisole=4: 1 : 0.2 with two portions at 0 °C. After stirring for 3 hours at 0 °C, the solution was evaporated under reduced pressure. The residue 3-4 (159 mg, 90%) was used for next step’s reaction
without further purification. $^1$H NMR (D$_2$O, 400 MHz): $\delta$ 4.88 (d, $J$ = 9.20 Hz, 1 H), 4.35 (d, $J$ = 7.76 Hz, 1 H), 3.84-3.81 (m, 2 H), 3.73-3.66 (m, 4 H), 3.63-3.55 (m, 5 H), 3.45 (t, 1 H), 3.34 (t, 1 H), 2.61-2.54 (m, 4 H). $^{13}$C NMR (100 MHz, D$_2$O): $\delta$ 176.90, 176.05, 102.85, 79.09, 75.03, 71.46, 68.53, 61.02, 59.86, 30.10, 28.69. HRMS: [M - H] $^\text{C}_{16}^\text{H}_{26}^\text{NO}_{13}$ calcd for 440.1410, found 440.1423.

3.5 References

1. Cardenas-Rodriguez, J.; Howison, C. M.; Matsunaga, T. O.; Pagel, M. D., A reference agent model for DCE MRI can be used to quantify the relative vascular permeability of two MRI contrast agents. Magnetic resonance imaging 2013, 31, 900-910.
4. Design and Synthesis of Glycoprotein Antigens for Immunotherapy

4.1 Introduction

Antibody, also known as immunoglobulin, is a large, Y-shaped protein that is used by the immune system to identify and neutralize abnormal cells and pathogens through strong interaction with specific antigens, including carbohydrate antigens. Antibody is secreted by B cells of the adaptive immune system and it is essential for humoral immunity. For instance, numerous studies have demonstrated that increasing the production of antibodies can facilitate exterminating cancer cells.\(^1,2\) This attractive effect suggested a potential antibody approach to cancer immunotherapy.

Tumor associated carbohydrate antigens (TACAs) have close relationship with the aberrant glycosylation in essentially all tumor cells. There are many TACAs identified and characterized by antibodies and lectins, including Globo-H, GM2, GD3, Tn, Tf and STn. To date, a few TACA based anti-cancer vaccine have been reported in clinical trials. These include Globo-H for prostate cancer,\(^3\) GM2 and GD3 for melanoma,\(^4,5\) and STn for breast cancer.\(^6\) Clinical studies proved that administration of these carbohydrate based vaccine actively induced their corresponding antibodies, which in turn suppress tumor cells. Based on previous achievements, vaccine development can be applied to any other TACAs. Recent development of novel vaccine, such as multivalent unimolecular and multi-component vaccine, through conjugation with strong immunogenic carrier such as KLH and CRM\(_{197}\) not only increase the antibody response, but also diversifies the antibody production.
Despite these improvements, current tumor associated antigen based vaccines show poor immunogenic effect because of low uptake by antigen-presenting cell (APC). The formation of immune complexes is one of the pre-requisites to design vaccines targeting APC. α-Gal is the most well-known carbohydrate epitope, which could interacts specifically with anti-Gal antibody, contributing 1% of serum IgGs in humans. Previous studies has demonstrated that immunogenicity of any microbial vaccine can be markedly increased by α-gal epitopes due to in vivo formation of immune complexes with anti-Gal antibody existing on APC. 

![Carbohydrate antigen/Antibody-directed vaccine design](image)

**Figure 4.1. Carbohydrate antigen/Antibody-directed vaccine design**

Lampreys are primitive jawless vertebrates capable of eliciting specific immune responses. Compared with T and B cells of jawed vertebrates, different types of lymphocytes could be produced by lampreys, which in turn express somatically diversified antibodies, namely variable lymphocyte receptors (VLRs). Evolutionarily, VLRs are the oldest adaptive immune receptors. There are three distinct antibodies produced by lampreys. T-like lymphocytes express and secrete VLRA; VLRB as a multivalent protein
is expressed on the surface of B-like lymphocytes; VLRC is expressed by a distinct lymphocyte lineage. VLRs are the only natural antigen receptors to use a non-Ig scaffold, which are alternatives to antibodies in applications such as bioimaging, biosensors and biopurification.\textsuperscript{10-12} Site-directed mutagenesis has confirmed the prediction of antigen binding through the concave surface residues of VLRs,\textsuperscript{13} which was also elucidated precisely by the structures of recombinant VLRB antibodies in complex with antigens of both H-antigen trisaccharide and hen egg white lysozyme (HEL).\textsuperscript{14,15} The crystal structure of VLRB in complex with the H-trisaccharide has provided structural insight into how VLRs recognize their antigens and provides a basis for rational design and modification of other antigen-specific VLRs.

4.2 Carbohydrate antigens based vaccine for activating VLRs in Lampreys

As long as H-antigen has been identified by previous study, our collaborator, Max from Emory University, shows strong interest to develop a novel vaccine targeting VLRs. Targeting the vaccine to VLRs through antibody-mediated pathway requires installation of carbohydrate antigen on vaccines. Taking advantage of our expertise in carbohydrate synthesis, especially chemical and chemoenzymatic synthesis of blood group and α-Gal, we are currently developing an efficient strategy to chemically conjugate carbohydrate antigens, including O-antigen and α-Gal to carrier proteins. O-antigen is structurally similar to H-antigen, so we choose O-antigen as a substitute for H-antigen. These synthetic glycoproteins will be subjected to immunogenic studies using lampreys.

4.3 Design and synthesis of glycoproteins

The synthesis of glycoprotein by chemical ligation usually consists of two steps (\textbf{Scheme 5.2}): (a) Installation of a spacer on sugar moiety; (b) Conjugation of
carbohydrate with carrier protein. Several methods have been employed for introducing functional spacer on oligosaccharides (Scheme 3-1). Chemical glycosylation has been widely employed in the first to install the varied spacers, and has successfully provided carbohydrate building blocks for second step. Reductive amination is one of the most efficient ways to introduce linkers to the reducing end of the free carbohydrate. While, this method would produce open-ring sugar moiety at reducing end and requires reducing agent, thus limits its application. Oxime ligation is an alternative way to achieve this goal efficiently by introducing stable oxime bond catalyzed by aniline.\textsuperscript{16} Direct amination by Kochetkov reaction is also a very simple and versatile way to install linkers to reducing end of sugars.

\[ \text{Scheme 4.1. Chemical Ligation of Glycoproteins} \]

In general, there are two approaches to achieve glycoproteins: (1) indiscriminate glycosylation, which targets available amines on the protein, typically those of lysine residues, and (2) site-selective glycosylation, targeting the thiol groups of cysteine residues. Scheme 4.2 demonstrated some most popular chemical conjugation methods of glycoproteins.\textsuperscript{17, 18} Among these methods, the Huisgen 1,3-dipolar cycloaddition (a) and oxime ligation (b) require the installation of artificial groups on the proteins, which limit their implication. Michael addition reaction in (e) and (f) is typically site-selective glycosylation which requires free thiol groups on lysine residue. Normally, there are very limited thiol groups on the surface of proteins as most of them have been reduced to
construct disulfite bonds. Diethyl squarate coupling (c) requires long reaction time and highly basic condition, which also restricts its application. N-hydroxysuccinimide ester (NHS) activation (d) is widely used for specific, efficient and permanent ligation to primary amines on lysine residues. This method is suitable for adding various molecules, including small molecules and PEG, under mild condition. However, the major drawback of this method is that NHS ester could be hydrolyzed quickly in acquis condition.

Scheme 4.2. Chemical Synthesis of Glycoprotein
In this study, we take advantage of chemical glycosylation and direct amination for installation of functionalized spacers (step a), and NHS activation chemistry for glycoprotein synthesis (step b).

### 4.3.1 Synthesis of α-Gal conjugated proteins

The synthesis of α-Gal building block was started from chemical glycosylation of trichloroacetimidate donor \(4-2\) with linker \(4-3\) to harvest compound \(4-4\), which was catalyzed by TMSOTf. \(4-4\) was globally deprotected by NaOMe and \(H_2\) catalyzed by Pd(OH)_2/C condition to furnish compound \(4-5\), which was served as substrate for enzymatic synthesis of α-Gal building block \(4-6\). Following our well-established procedure, one galactose could be specifically added to 3'-OH of \(4-5\) to produce compound \(4-6\), which was catalyzed by α1,3-GalT with an excellent yield. (Scheme 4.3)

![Scheme 4.3 Chemoenzymatic Synthesis of α-Gal Linker](image_url)

With compound \(4-6\) in hand, we could readily convert it to NHS ester by coupling amine group with linker \(4-7\) under basic condition to achieve α-Gal NHS ester \(4-8\), which was subjected to couple with Biotin-BSA or ovalbumin (OVA) in PBS buffer solution (Scheme 4.4). Both linkers and proteins were prepared at concentrations of 10 mg/mL in 1x PBS buffer, and then were mixed at 1:1 (v/v) ratio. The reaction was carried out at room temperature for 24 hours. After ultrafiltration to remove the unreacted linkers with
cut-off at 10 kDa, the collected protein was analyzed by SDS-PAGE. It clearly showed the successful conjugation from SDS gel (Figure 4.2A & 4.3A). MALDI-TOF was also used to analyze the α-Gal-protein conjugates (Figure 4.2B & 4.3B). From the results, we can clearly check the molecular weight of glycoproteins as well as loading ratio of sugar moiety.

Scheme 4.4 Preparation of α-Gal Proteins

Figure 4.2. Characterization of α-Gal Biotin BSA
4.3.2 Synthesis of H-antigen conjugated proteins

After achieving α-Gal-proteins, we attempted to follow the same method to synthesize O-antigen linker, but it did not work well. WbsJ could not transfer efficiently Fucose moiety to 2'-OH of compound 4-5 (yield<10%), even though there was minor product detected by mass spectra (Scheme 4.5A). Therefore, we have to synthesize free H-antigen with WbsJ first, then install the linker to H-antigen (Scheme 4.5B).
Scheme 4.5 Preparation of H Antigen Proteins

To functionalize lactose moiety, an amine group need to be installed at the reducing end. The Kochetkov reaction assisted with microwave irradiation was used to furnish compound 4-2. The lactose 4-1 was mixed with 15 eq. of ammonium carbonate in dry DMSO and mixture was radiated at 30-40 °C for 4 hours with a yield of (70%). After eliminating excess ammonium carbonate, the solution was used for next step reaction without further purification. Compound 4-2 and PEG linker 4-11 were coupled smoothly with treatment of HATU and HOBt under basic condition, followed by deprotection of Fmoc group using piperidine/DMF condition to provide compound 4-12. 4-12 and 4-7 were coupled under mild condition with treatment of trimethylamine to furnish NHS ester 4-13, which was subjected to couple with Biotin-BSA or ovalbumin (OVA) in PBS buffer solution. The purification process is the same with that of α-Gal-proteins. SDS-PAGE and MALDI-TOF were utilized to analyze the H-antigen-proteins (Figure 4.5 & 4.6).
Figure 4.4. Characterization of H Antigen Biotin BSA

Figure 4.5. Characterization of H Antigen OVA
4.4 Conclusion

Carbohydrate antigens based vaccines for activating VLR antibodies in lampreys have been synthesized. The multivalent H antigen glycoproteins are promising vaccine candidates to elicit immune response in lampreys’ unique immune system. Further biological tests are needed to validate this novel vaccine design.

4.5 References


5. Facile Preparation of Amylose-Proteins by Phosphorylase-catalyzed Polymerization

5.1 Introduction

Carbohydrates are ubiquitous in nature, and there is little doubt that the elaborate carbohydrate constructs observed on the surface of cells and proteins hold the key to understanding many complex biological processes. In particular, protein folding and stability are deeply influenced by the co- and post-translational glycosylation.\textsuperscript{1-3}

Many proteins could not fold correctly without proper glycosylation. Evidences has showed that incorrect glycosylation on the surface of proteins could fail the trimming process of proteins and proteins would be expelled via the endoplasmic reticulum-associated protein degradation (ERAD) pathway.\textsuperscript{4} Therefore, correct glycosylation is a great indicator of correct protein structure. Besides, carbohydrates could also stabilize tertiary structure of proteins.\textsuperscript{5}

In nature, glycoproteins present as a mixture of glycoforms that has the same peptide sequence but different carbohydrates, which limited the studies about the enhanced proteolytic stability of glycosylation. For instance, RNase B, a glycosylated pancreatic ribonuclease at Asn34, demonstrated higher dynamic stability and functional activity than unglycosylated RNase A.\textsuperscript{6} Chemical glycosylation of proteins can provide practical alternative to achieve pure glycoforms for detailed structural and functional studies, with numerous variation in carbohydrate type and attachment site.\textsuperscript{7,8}

The majority of our work presented herein focuses on developing novel methodology for the glycosylation of proteins to enhance proteins' stabilities against
proteolysis, thermolysis and other forms of degradation, as well as future potential to serve in various medical settings.

5.1.1 PEGylation and Glycosylation of Therapeutic Proteins

Therapeutic proteins, such as insulin and Granulocyte-colony stimulating factor (G-CSF), are becoming rising stars in the pharmaceutical sciences. More than 100 genuine and similar number of modified therapeutic proteins are approved for clinical use in the European Union and the USA with 2010 sales of $108 bln. However, the major drawbacks of therapeutic proteins, including, antigenicity, short circulation half-life time and easy excretion from human body, are waiting for be solved in current stage. So far, the most practical and popular method employed are PEGylation, which involves the conjugation of poly (ethylene glycol) (PEG) with functional groups of proteins. In 1970s, Davies and Abuchowsky first investigated the PEGylation of protein and reported in two papers. Since then, the practice of PEGylation has grown dramatically and several PEGylated therapeutic proteins has been approved by FDA. There is no doubt that PEGylation can increase the efficiency of therapeutic proteins. For instance, PEG could increase the size of proteins, effectively shielding proteins’ surface and limit the recognition by the reticuloehdothelial system. Major benefits of PEGylation include increase solubility of proteins, increase circulation half-life time and limit degradation by metabolic enzymes (Figure 5.1).
Despite of numerous benefits of PEGylation, its shortages are also obvious. In some cases, PEGylation could promote decreased biological activities and accumulate in the human body, giving rise to macromolecular syndrome.\textsuperscript{13} There are some concerns about generation of antibodies against PEG. All of these concerns stem from the fact that PEG is non-degradable and unnatural material.

Therefore, there are needs and necessities to develop novel materials to substitute PEG in terms of modification of therapeutic proteins. Inspired by nature, glycosylation with even small oligosaccharides could play the similar roles in increasing the solubility and half-life time compared with PEGylation (Figure 5.2).\textsuperscript{16, 17} Like PEGylation, glycosylation also provides a large-size hydration shell which increase solubility and protect proteins from proteolysis.\textsuperscript{1} Besides, carbohydrate could increase thermal stability and decrease precipitation and aggregation.\textsuperscript{13, 18} Furthermore, unlike PEG, carbohydrate is natural and bio-degradable material which could be hydrolyzed and removed by glycosidase, making glycosylation an attractive alternative to improve performance of therapeutic proteins.
Figure 5.2. Benefits of glycosylated therapeutic proteins

5.1.2 Phosphorylase-catalyzed enzymatic polymerization

Phosphorylase (EC 2.4.1.1) is an exotype enzyme, which catalyzes in vivo reversible phosphorolysis catalyzes the reversible phosphorolysis of α-(1→4)-glucans, such as amylose and glycogen, at the non-reducing end, giving a glucose 1-phosphate (Glc-1-P). The enzyme can be utilized for glycosylation reaction because the substrate of a glycosylphosphate is activated and possesses the relatively low bond energy, compared to that of a glycosyl-nucleotide (Figure 5.3).

Figure 5.3. Phosphorylase-catalyzed enzymatic polymerization
Indeed, the enzyme catalyzes both reactions, phosphorolysis and glycosylation. In vitro synthesis of polysaccharides was performed via phosphorylase-catalyzed polymerization. The phosphorylase-catalyzed reaction was adapted to various polymerizations. For example, the formation of amylose-polymer inclusion complexes is synthesized by vine-twinning polymerization catalyzed by phosphorylase (Scheme 5.1).

Phosphorylase-catalyzed polymerization was also utilized to synthesize amylose grafted chitin.

Scheme 5.1 Application of phosphorylase-catalyzed polymerization

Previous work in our group demonstrated feasibility of conjugating hemoglobin (Hb) with carbohydrate of variable molecular weights, up to 10,000. Furthermore, biophysical properties of glycosylated Hb are highly conserved compared with native Hb, suggesting glycosylation of Hb is a promising way to increase therapeutic proteins. However, this methodology requires tedious preparation of linker, free cysteine on the surface of proteins and the coupling efficiency of dextran with Hb is low due to bulky size of dextran.
In this chapter, we developed a novel method to generate variable MW carbohydrate-coated proteins based on three key tools: direct amination, NHS-ester chemistry and phosphorylase-catalyzed enzymatic polymerization.

5.2 Results and Discussion

Concerning convergent glycoprotein synthesis, generally two approaches are taken: (1) indiscriminate glycosylation, which targets available amines on the protein, typically those of lysine residues, and (2) site-selective glycosylation, which targets the thiols of cysteine residues or amide of glutamine residues. This report focuses on employing oligosaccharides for the indiscriminate glycosylation of therapeutic proteins. Significantly, the methods developed and presented within hold the potential for glycosylation—with high MW carbohydrates—to be used as a viable alternative to PEGylation.

A substantial fraction of the currently approved protein pharmaceuticals need to be properly glycosylated to exhibit optimum therapeutic efficacy. This is due to the fact that glycosylation can influence a variety of physiological properties at both the cellular (e.g. intracellular targeting) and protein levels (e.g. protein solubility, protein molecular stability). As mentioned before, phosphorylase-catalyzed polymerization is promising way to generate novel amylose-coated materials, including amylose-proteins. Figure 5.4 demonstrated the substrate specificity of phosphorylase. Vial 1 and 3 were maltopentaose and maltoheptaose, respectively. Vial 2 and 4 were maltopentaose mixed with phosphorylase and G-1-P, maltoheptaose mixed with phosphorylase and G-1-P, respectively. After one hour reaction, KI-I$_2$ was added into each vial and only vial 4 changed to blue color, which indicated the formation of amylose. Based on this
experiment, we chose maltoheptaose as primer attached to proteins for further polymerization.

**Figure 5.4. Model Reaction of Polymerization catalyzed by Phosphorylase**

**Scheme 5.2** displayed our design of synthesizing amylose-proteins catalyzed by phosphorylase. We decided to use the microwave-assisted Kochetkov Reaction to introduce amine group at the reducing end of malheptaose. Of the various solvents tested, only DMSO resulted in good yields. Gratifyingly, the reaction was found to be efficient with only 15-fold excess of ammonium carbonate over sugar compared to the 40-50-fold excess needed under thermal conditions. Another key step is to install proper spacer which is ready for conjugating amine groups of lysines on surface of proteins. After conjugating maltoheptaose with proteins, phosphorylase could polymerize the primer on the surface of proteins to afford amylose-proteins.
Scheme 5.2 Design of enzymatic polymerization of maltoheptaose-proteins

Squaric acid can be applied as active reagent to couple two amino-functional compounds. Consecutive coupling of two primary amines could furnish synthesis of asymmetric squaric acid bisamides with either small molecules but also biomolecules or polymers. After the first amidation step of the squaric acid, the resulting ester-amide carries reduced reactivity, which is the key feature of squaric acid mediated coupling. We prepared two squaric acid-conjugated maltoheptaose (Scheme 5.3). Free maltoheptaose 5-3 could add one amine group at the reducing ending with microwave-assisted Kochetkov Reaction to afford compound 5-4, followed by coupling with squaric condition under mild condition to generate compound 5-1. In order to study whether the linker length could affect the coupling and polymerization or not, we decided to synthesize maltoheptaose-PEG-squaric acid 5-2, which has more flexible linker and reduced stereo hindrance to increase coupling efficiency.
Scheme 5.3 Synthesis of Maltoheptaose Donors

Then these two compound were subjected to conjugation of BSA under basic condition (Scheme 5.3). Since these two compounds has lower reactivity as mentioned before, the coupling reaction with BSA requires higher PH value and longer time. PH should increase to 8.5-9.0 and reaction time was extended to 24 hours.
Scheme 5.4 Conjugation and polymerization of Glyco-BSA

Based on the SDS gel (Figure 5.5), compound 5-2 indeed result higher coupling efficiency than compound 5-1, making 5-2 more suitable molecules for future’s experiment. Expect longer chain, high hydrophilicity of PEG chain might also increase the efficiency of compound 5-2. Interestingly, we have difficulties to get the MALDI-MS result from the glycosylated BSA because of the interference of ionization imposed by the large oligosaccharide, a trait commonly seen with proteins bearing large oligosaccharides. The following elongation were promoted by addition of phosphorylase with ratio of G-1-P to glycosylated BSA equal to 1000:1. After 24 hours, the resulting amylose-BSA showed from the SDS gel (add figure). Comparison of lane 3 and lane 5, as well as lane 4 and lane 6, indeed proved the polymerization of the glycol-BSA.
5.3 Conclusion

Overall, phosphorylase has been shown to catalyze the synthesis of amylose-protein conjugates from their corresponding maltoheptaose-protein primer. The maltoheptaose-BSA conjugates has been synthesized in this study. Methods of synthesizing and assaying amylose-BSA with higher molecular weights has been developed. This methodology provides a novel access to achieve polysaccharide-protein conjugates, which could be a promising substitute for PEGylation.

5.4 Experimental Section

General Methods

The synthetic Sugars in solution (10 mg/mL in 3x PBS) were added to the same volume of protein solution (10 mg/mL in 3xPBS) and was stirred at RT for 1 h. Then the resultant solution was ultrafiltrated and washed with 1x PBS using Amicon Centrifugal
Filter Devices (Ultracel 10,000). The collected glycoprotein solution was quantitated by Pierce BCA Protein Assay Kit (Pierce) and stored at 4 degrees.

The SDS-PAGE was performed with standard condition. Protein conjugates were suspended in 12 μL of sample buffer (5% (w/v) SDS, 10% (v/v) glycerol, 25 mM Tris-Cl, pH 6.8, 10 mM DTT, 0.01% (w/v) bromophenol blue), loaded on different lanes of a 1.5-mm-thick, 12% (w/v) SDS-PAGE gel, and visualized by Coomassie Brilliant Blue R-250 staining.

Compound 5-7

Compound 5-3 (400 mg) and ammonium carbonate were dissolved in 20 mL DMSO and stirred with microwave tube at 40 °C for 3 h. TLC plate showed the formation of compound 5-4. The solvent was then removed under reduced pressure without further purification.

To a 25ml round-bottom flask containing a solution of compound 5-4 (90mg, 0.078mmol) in anhydrous DMSO was added compound 5-5 (139mg, 0.31mmol), HOBt (20mg, 0.153mmol) and HATU (58mg, 0.153mmol). The reaction mixture was stirred at room temperature under nitrogen atmosphere for 24hrs. The solvent was removed under reduced pressure. The crude product was purified by C18 reverse phase flash chromatography using 0-100% MeOH in water as an eluent to afford compound 5-6 as a white solid with a yield of 70%.

Compound 5-6 was added to a solution of 20% piperidine in DMF After 2 hours, the TLC plate showed the fully deprotection. The reaction was then concentrated, the residue.
was added 0.5 mL water, extracted with EtOAc (1 mL x 4). The aqueous layer was lyophilized to give compound 5-7 as a white powder.

**Compound 5-2**

The compound was dissolved in EtOH/H2O (1:1). 3, 4-Diethoxy-3-cyclobutene-1, 2-dione (1.0 equiv) was added. Then 5 µL solution of sat. Na₂CO₃ was added in intervals of 5 min, until a pH of 8 was reached. After stirring for 1.5 hours at room temp, the reaction mixture was neutralized by adding HOAc. The organic solvents were removed in vacuum. The residues were purified by C18 reverse phase flash chromatography using 0-100% MeOH in water as an eluent to afford compound 5-2 as a white solid with a yield of 85%.

**5.5 References**


6. Preparation of α-Gal-aptamer Conjugates for anticancer therapy

6.1 Introduction

Nucleic acids are the foundation of life. It is well-known that defined sequences translate certain genetic traits, protein expression and cellular function. Except that, they can be very useful for the recognition of biomolecules. When these nucleic acids are in single-stranded form, they are called aptamers, which have attracted much attention by carrying the ability of nucleic acids to binding certain biological molecules, including tumor-associated proteins, due to their stable three-dimensional structure.

The word “aptamer” derives from the Latin word *aptus*, which means to fit, and the Greek *meros*, meaning region. Ellington and Gold published the first papers on aptamers in 1990.1, 2 The length of nucleic acid aptamers is in the range of 10–100 nucleotides, which are produced from a process called “systematic evolution of ligands by exponential enrichment” (SELEX) (Figure 6.1). SELEX employed enriching for single stranded RNA or DNA sequences based on retention of this sequence by binding a target of interest. SELEX could achieve high-affinity aptamers to be selected against many targets including extracellular domain of prostate-specific membrane antigen (PSMA),3 MUC1 peptides,4 P-selectin,5 and protein tyrosine phosphatase 1B (PTP1B).6 DNA and RNA are not limited in function to storing and translating genetic information. Aptamers have highly stable tertiary structure, which allows them to form solid complexes with different targets, such as proteins.7–11 Aptamers can be referred to as “antibodies”, but aptamers has a number of functional advantages over antibodies. Aptamers are easy to synthesize and modify, inexpensive, long-term stability, have low immunogenicity, and possess greater tissue
penetration than antibodies.\textsuperscript{12, 13} Besides, antibodies need to be developed \textit{in vivo}, but aptamers can be developed \textit{in vitro} making aptamer production much easier and more feasible.

\textbf{Figure 6.1. Screen RNA Aptamer-the SELEX process}

The application of nucleic acid aptamers has greatly expanded since their inception about two decades ago. Even though many of these developments, especially those for biomedical applications such as diagnostics, biomarker discovery, therapeutics, and drug delivery, are still at the research stage, results are promising.\textsuperscript{14} One promising area is therapeutics. Due to their binding affinity and high specificity, aptamers could directly interrupt the activities of target proteins making them promising therapeutic agents for the treatment of diseases. Furthermore, aptamers binding cell surface proteins are well suited for the targeted delivery of other therapeutics, such as conjugated small interfering RNAs (siRNA) that induce RNA interference (RNAi). Thus, aptamer-siRNA
complex may offer dual-functions, in which the siRNA internalizes into the cell to target a specific mRNA, while the aptamer block a receptor function.\textsuperscript{15}

The therapeutic aptamers could modulate downstream signaling pathways through bounding with a specific target molecule on the surface of cell. The major inhibition mechanisms includes inhibiting dimerization to associated molecules, preventing structural changes in the target molecules, or phosphorylation of downstream proteins.\textsuperscript{16, 17} Due to these properties, anti-tumor therapy is one of the most popular application of aptamers. Aptamers are screened and designed to distinguish between normal and tumorigenic cells based on specific binding of aptamer and targeted protein. Tumor-targeting aptamers can be useful and highly specific in tumor cell detection and eliminations, which is needed for cancer treatment, as current chemotherapeutics present significant side effects such as massive death of normal and depletion of immune cells to fight infection. Hence, aptamer therapies could minimize these adverse effects by highly specific on tumorigenic cells and assisting patients to maintain homeostasis.\textsuperscript{10, 18}

Gilboa, one of our collaborators, is focusing on the cancer therapeutics based on highly specific aptamers. His group have developed several promising conjugated aptamers to detect tumor cells and suppress growth of tumor cells in mice models.\textsuperscript{19-21} Our group has rich experience on synthesis of biologically important carbohydrate epitopes such as \(\alpha\)-Gal, SLe\(^x\) and Globo H, which involves in humoral immunity, especially in anti-cancer immunotherapy. However, nobody has ever combined these two powerful tools, aptamer and carbohydrate, together to explore the potentials of developing anti-cancer therapies. This project is to fulfill this blank field to design, synthesize and immunologically evaluate the carbohydrate-aptamer conjugates on anti-cancer therapies.
6.2 Results and Discussion

Monoclonal antibodies such as Rituximab or Trastuzamab mediated their antitumor effects in large part through the recognition of the tumor bound antibodies by complement or macrophages. It is however only partially effective and therefore improving the immune recognition of tumor bound antibodies is paramount to improve this promising platform of immune therapy of cancer. In a recent study, Carmi and colleagues have shown that “coating” tumor cells with polyclonal antibodies present in the patient is remarkably effective in eliminating the said tumor cells.\textsuperscript{22} We aim at developing a clinically applicable methods of coating the patient' resident tumor cells with polyclonal antibodies induced in the cancer patient.

Human individuals have high titers of polyclonal antibodies that recognize a small molecular weight trisaccharide antigen, Gal\textgreek{a}1-3Gal\textgreek{b}1-4Glc (\textgreek{a}-Gal), most likely as a result of postnatal exposure to harmless microbes.\textsuperscript{23} To “coat” tumor cells \textit{in situ} with the anti-\textgreek{a}-Gal antibodies, we will use an oligonucleotide aptamer that binds to a tumor cell specific products such as PSMA or Her2 that will be conjugated to \textgreek{a}-Gal trisaccharide. Oligonucleotide aptamers represent a novel and emerging platform of ligands with desired specificity that offer potentially significant advantages over monoclonal antibodies in terms of development, manufacture, and cost. Upon injection of the aptamer-\textgreek{a}-Gal conjugate into patient or experimental mouse, it will bind specifically to the tumor cells and the \textgreek{a}-Gal moiety will attract the pre-existing anti-\textgreek{a}-Gal antibodies present in the patient, thereby in effect coating the tumors with polyclonal anti-\textgreek{a}-Gal antibodies. To conjugate the \textgreek{a}-Gal moiety to the aptamer we will first conjugate \textgreek{a}-Gal to a short
oligonucleotide that will be hybridized to the aptamer engineered to contain a complementary sequence. (Figure 6.2)

Carbohydrate-oligonucleotide conjugates (COCs) has been extensively studied during last decade since these conjugates not only improve oligonucleotide’s properties but also create new ones. However, preparation of COCs is also challenging because the chemical characteristics of the oligonucleotide and the carbohydrate to be attached are not fully compatible. Several synthetic approaches have been developed for facile preparation of COCs, which can be grouped into several major categories: 1. solid-support synthesis, where carbohydrate phosphoramidite derivatives are coupled to oligonucleotides moiety by automated DNA synthesis method; 24 2. Copper catalyzed azide–alkyne cycloaddition (CuAAC), where ligation of azido-carbohydrates and alkyne-oligonucleotides are catalyzed by Cu(0) to afford triazole rings; 25 3. Reductive amination,
where amine-terminated DNA is coupled with reducing end of carbohydrate in the presence of NaCNBH₃;¹⁻⁶  4. Nucleophilic addition on unsaturated carbon through thiol-oligonucleotide is added an unsaturated carbonyl on carbohydrate via Micheal addition to form thioether bond;²⁷  5. EDC/NHS coupling, carbohydrates bearing carbonic acid coupled with amine-modified oligonucleotides by EDC/NHS condition.²⁸ (Figure 6.3)

Figure 6.3. Methods of Synthesis of Carbohydrate-Oligonucleotide Conjugates (COCs)

Based on our experience of glycoprotein synthesis, we attempted to utilize NHS ester chemistry to couple α-Gal epitope with amine-modified ssRNAs which were purchased from Trilink company. The aforementioned NHS ester is very active due to its electrophilicity and amine group can attach it to form stabile a amide bond at weak base
condition. This method is simple and efficient as no tedious manipulation and harsh condition are needed. But the drawbacks are also obvious because NHS ester are too reactive that water can attack the ester bond to form corresponding acid. **Scheme 6.1** demonstrated our design of synthesis of COCs in PBS buffer. Carbohydrate NHS ester mixed with amine-C6-ssRNA with a ratio of 50:1 in PBS buffer with PH at 7.5. To avoid potential thermolysis of ssRNA, the coupling was performed under 4 °C. After 24 hours, the mixture was detected with MALDI-TOF spectremery. Unfortunately, only a few product was detected for the coupling of α-Gal and 5’P-ssRNA. No product was detected for the coupling of α-Gal and 3’P-ssRNA. There was no improvement of the result even though the ratio of carbohydrate to ssRNA was increased to 100:1. The major reason of this failure might be short carbon linker terminated with amine was shielded by ssRNA backbone and bulky carbohydrates NHS ester has fewer chance to access the amine group.

**Scheme 6.1. ssRNA Conjugation Reaction**

In order to prove this assumption, we performed a model reaction in which small-size biotin-NHS was mixed with the aforementioned ssRNA. It turned out this coupling reaction
is highly efficient and all of ssRNA was conjugated with biotin tag according to MALDI-TOF result (Figure 6.4). The size of biotin-NHS is much smaller than α-Gal that it could penetrate the shields of RNA backbone and access the amine group more easily. Six carbon linker is too short to react with bulky carbohydrate NHS ester.

![Diagram of ssRNA conjugation](image)

**Figure 6.4. MALDI Result for Biotin-ssRNA conjugation**

Based on previous experiments, longer linker should overcome this coupling issue. Another DNA modified with C12 linker terminated with amine group was for modeling reaction. We used lactose NHS and maltoheptaose NHS to couple with the amine-C12-DNA. (Figure 6.5) Carbohydrate NHS ester mixed with amine-C12-DNA with a ratio of 50:1 in PBS buffer with PH at 7.5. The HPLC result showed that the coupling reaction did work well. The retension time of starting material is about 22.58 min. After glycosylation, the retension time of lactose-C12-DNA was shifted to 3.20 min, and the retension time of maltoheptaose-C12-DNA was shifted to 6.73 min (add HPLC fig). However, we have some difficulties to detect the carbohydrates-C12-DNA with MALDI-TOF, even though amine-C12-DNA could be detected very well. The traditional matrix for detecting oligonucleotides are Hydroxypicolinic Acid (HPA) or 2′, 4′, 6′ -
Trihydroxyacetophenone monohydrate (THPA). HPA used in this project is compatible with ssRNA and unconjugated DNA, but has some difficulties to detect COCs prepared by us. Some literatures suggested HPA in acetonitrile/water (1:1, v/v) containing 10% of ammonium citrate and few beads of DOWEX 50W-X8 ammonium sulfonic acid resin were added. Further work is needed to figure out the proper MALDI detection method for our COCs samples.

![Chemical structure diagram]

Figure 6.5. Modeling Reaction of COCs

6.3 Conclusion

A method of synthesis of carbohydrate-oligonucleotide conjugates has been developed based on NHS ester chemistry. Amine terminated linker should be long enough to expose on the surface of oligonucleotides for NHS ester coupling. Short chain will be shielded by oligonucleotides backbones. Further investigation is needed to detect COCs by MALDI spectrometry. We will conjugate α-Gal NHS ester with the coming amine-C12-ssRNA in the future.
6.4 References


APPENDICES: $^1$H, $^{13}$C NMR Spectra and HPLC Profiles of Compounds
ESI HRMS of 1-1 and active ester modified pentasaccharide
MALDI–TOF mass (A) and SDS–PAGE (B) analysis of CRM$_{197}$ and pentasaccharide–CRM$_{197}$

MALDI–TOF mass (A) and SDS–PAGE (B) analysis of BSA and pentasaccharide–BSA
$^1$H NMR of Compound 2-6

$^{13}$C NMR of Compound 2-6
$^1$H NMR of Compound 2-7

$^{13}$C NMR of Compound 2-7
$^1$H NMR of Compound 2-1

$^{13}$C NMR of Compound 2-1
$^1$H NMR of Compound 2-9
$^{13}\text{C NMR}$ of Compound 2-9

$^{1}\text{H NMR}$ of Compound 2-3
\textbf{\textsuperscript{13}C NMR of Compound 2-3}
$^1$H NMR of Compound 2-13

$^{13}$C NMR of Compound 2-13
$^1$H NMR of Compound 2-14

$^{13}$C NMR of Compound 2-14
$^1$H NMR of Compound 2-15
$^{13}\text{C NMR}$ of Compound 2-15

$^{1}\text{H NMR}$ of Compound 2-16
$^{13}$C NMR of Compound 2-16
$^{1}H$ NMR of Compound 2-19

$^{13}C$ NMR of Compound 2-19
$^1$H NMR of Compound 2-20

$^{13}$C NMR of Compound 2-20
\(^1\)H NMR of Compound 2-21

\(^{13}\)C NMR of Compound 2-21
$^1$H NMR of Compound 2-22
$^{13}\text{C NMR}$ of Compound 2-22

$^{1}\text{H NMR}$ of Compound 2-23
$^{13}$C NMR of Compound 2-23
$^1$H NMR of Compound 2-24

$^{13}$C NMR of Compound 2-24
$^{1}H$ NMR of Compound 2-25

$^{13}C$ NMR of Compound 2-25
NMR, and HPLC data of purified HMOs

$^1$H NMR of HMO1

$^{13}$C NMR of HMO1
HMO11

$^1$H NMR

**HILIC-ELSD, $T_R = 11.946$ min**
HMO12

$^1$H NMR

HILIC-ELSD, $T_R = 17.063\text{min}$
HMO13

$^1$H NMR

HILIC-ELSD, $T_R = 19.767 \text{min}$
HMO14

$^1$H NMR

HILIC-ELSD, $T_R = 15.606$ min
HMO15

\(^1\)H NMR

HILIC-ELSD, \(T_R = 13.605\) min
HMO16

$^1$H NMR

HILIC-ELSD, $T_R = 17.718$ min
HMO2

\(^1\)H NMR

\(^{13}\)C NMR of HMO2
HMO21

$^1$H NMR

HILIC-ELSD, $T_R = 12.722$ min
HMO22

$^1$H NMR

HILIC-ELSD, $T_R = 13.927$ min
HMO23

$^1$H NMR

HILIC-ELSD, $T_R = 11.877\text{min}$
HMO24

$^1$H NMR

HILIC-ELSD, $T_R=10.566$ min
HMO25

$^1$H NMR

HILIC-ELSD, $T_R = 14.747$ min
HMO26

$^1$H NMR

HILIC-ELSD, $T_R = 15.796\text{min}$
HMO27

$^1$H NMR

HILIC-ELSD, $T_R = 13.933$ min
HMO28

$^1$H NMR

**HILIC-ELSD**, $T_R = 13.581 \text{min}$
HMO29

$^1$H NMR

HILIC-ELSD, $T_R = 13.993$ min
HMO210

$^1$H NMR

HILIC-ELSD, $T_R = 16.260$ min
HMO211

$^1$H NMR

HILIC-ELSD, $T_R = 15.076$ min
HMO3

$^1$H NMR

HMO3

$^{13}$C NMR
HMO31

$^1$H NMR

HILIC-ELSD, $T_R = 12.548$ min
HMO32

$^1$H NMR

HILIC-ELSD, $T_R = 13.710$ min
HMO33

$^1$H NMR

HILIC-ELSD, $T_R = 11.576$ min
HMO34

$^1$H NMR

HILIC-ELSD, $T_R = 10.512$ min
HMO35

\(^1\)H NMR

HILIC-ELSD, \(T_R = 14.547\) min
HMO36

$^1$H NMR

HILIC-ELSD, $T_R = 15.697$ min
HMO37

$^1$H NMR

HILIC-ELSD, $T_R = 13.611$ min
HMO38

$^1$H NMR

HILIC-ELSD, $T_R = 12.437$ min
HMO39

$^1$H NMR

HILIC-ELSD, $T_R = 12.303$ min
HMO310

$^1$H NMR

HILIC-ELSD, $T_R = 16.506$ min
HMO311

$^1$H NMR

HILIC-ELSD, $T_R = 14.544$ min