

CHEMICAL APPROACHES TO UNDERSTANDING GLYCOBIOLOGY

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Wen Yi, B.S.

The Ohio State University
2008

Dissertation Committee:

Professor Peng G. Wang, Advisor

Professor Ross E. Dalbey

Professor Donald H. Dean

Professor F. Robert Tabita

Approved by

Advisor

the Ohio State Biochemistry Graduate Program

ABSTRACT

Glycans are vital biopolymers found in organisms across all domains of life. They play critical roles in numerous complex biological processes, ranging from mediation of protein folding, stability and trafficking, to molecular recognition by immune cells, viruses and bacteria. Traditional genetic approach has yielded significant insight into glycan functions, but also has limitations. Glycans are not primary gene products; their biosynthesis is not under direct transcriptional control, making their genetic manipulation far more complicated than for nucleic acids or proteins. Disruption of genes encoding glycan biosynthetic enzymes oftentimes causes the impaired synthesis of multiple glycan structures, making the interpretation of mutant phenotypes very difficult or impossible. Recently, chemical approaches are emerging to be powerful alternatives to genetics in the study of glycobiology. In this thesis, the work has been focusing on using chemical approaches to gain a better understanding of cell surface glycan biosynthesis, investigate pathway mechanism and modulate cell surface glycan presentation.

Chapter 1 provides a general introduction of glycan functions, bacterial cell surface polysaccharide biosynthesis, and an overview of chemical approaches used in glycobiology. Two most widely used chemical approaches (also the ones used in this thesis) are highlighted: synthesis of structurally defined oligosaccharides/glycoconjugates

to dissect glycan biosynthesis, and use of unnatural sugar analogs to modulate the cell surface glycosylation.

Chapter 2 describes the effort to reconstitute bacterial polysaccharide biosynthesis in a cell free system. Structurally defined oligosaccharide substrates were obtained using chemo-enzymatic synthesis. These substrates were used *in vitro* to demonstrate sugar polymerization. Polysaccharide chain length modality was generated in this chemical reconstituted system.

In Chapter 3, a sugar nucleotide synthetic pathway was engineered, and its promiscuity was exploited to *in vivo* incorporate a panel of unnatural sugar substrates into bacterial polysaccharides to generate structurally novel polysaccharides. Selective chemical reactions were carried out *in vitro* to further elaborate the functional groups appended to the polysaccharides.

In Chapter 4, two novel microbial glycosyltransferases were characterized and applied in the facile synthesis of two types of important human oligosaccharide antigens: T-antigen mimics and human blood group antigens.

Finally, Chapter 5 summarizes the main results of studies in this thesis and also provides some directions for future studies in each project.

To my parents

ACKNOWLEDGMENTS

The past six years has become one of the most important and memorable chapters in my life. This, in many ways, is the beginning, not the end, of a journey. Looking back, there are so many people that deserve my heartfelt gratitude. Let me start by thanking my advisor, Dr. Peng G. Wang. He always encourages and challenges me scientifically, and constantly pushes me to do my best. His energy, enthusiasm and passion for science have been an inspiration for me in the past six years, and will always be in the years to come. I cannot think of a better advisor who can give me such freedom to pursue science and foster me to become an independent thinker.

Secondly, I want to thank my committee, Dr. Ross Dalbey, Dr. Donald Dean, Dr. F. Robert Tabita, for their help and insightful discussion during my research.

I also want to thank my lab-mates. I am so fortunate to be able to work with such a wonderful group of people. Our joy, frustration and struggle over research will always remain a memorable experience for me. Especially, I want to thank Dr. Chengfeng Xia (Charles) for his help in chemical synthesis. He is also a wonderful mentor for my organic synthetic skill. I thank Dr. Jie Shen for his help in NMR analysis of oligosaccharide and polysaccharide structures. I also thank Bob Woodward for synthesizing sugar lipid conjugates.

I am also thankful for my former lab-mates, who helped me one way or another during my graduate study. Dr. Mei Li and Dr. Jun Shao taught me various techniques in molecular biology and biochemistry. Without their help and encouragement, it would take me much longer time to figure out what I was supposed to do in research. Dr. Hongjie Guo generously taught me the strategies for prokaryotic gene knockout and LPS analysis, which have become essential techniques in my research. Dr. Hironobu Eguchi has inspired me with his dedication, patience and attentiveness in research and his work ethic.

It goes without saying that friendship has been an important part in my life over the past years. I am thankful for all that they have done for me. I thank them for their company and support through every step of the way. They are: Jianjun Yuan, Wenpeng Zhang, Yalong Zhang, Bin Li, Zhun Ma, Renhua Zhang, Wenlan Chen, Edwin Motari, Wenzhe Lu, Jing Song, Xinyi Chen (Lulu), Mamata Singh, Nga Peng, and Lily Peng.

Special thanks go to our lab managers: Cindy Coleman and Juliana Pernik. They have done a fantastic job keeping the lab running smoothly.

Also, my work wouldn't be possible without the help from our collaborators: Dr. Jianjun Li from National Research Council of Canada, Dr. C. Allen Bush from University of Maryland, and Dr. Xi Chen from University of California-Davis.

Finally, I want to thank my parents, who have always been there for me. I wouldn't be who I am today if it wasn't for them. There are no better words to express my heartfelt gratitude for their unconditional love, gracious understanding and lifetime support.

VITA

November 13, 1979 Born – Maoming, Guangdong, China

July, 2002..... Bachelor of Science, Polymer Science
Fudan University, Shanghai, China

2002 – 2003..... Graduate Teaching Assistant
Wayne State University, Detroit, MI

2003 – 2005..... Graduate Teaching Assistant
The Ohio State University, Columbus, OH

2006 – 2008..... Graduate Research Associate
The Ohio State University, Columbus, OH

PUBLICATIONS

Research Publications

1. Faridmoayer, A., Fentabil, M.A., Haurat, M.F., **Yi, W.**, Woodward, R., Wang, P.G., and Feldman, M.F. “Substrate Promiscuity of a Bacterial Oligosaccharyltransferase Involved in O-Linked Protein Glycosylation” *J. Biol. Chem.*, **2008**, In press.
2. Li, M., Shen, J., Liu, X., Shao, J., **Yi, W.**, Chow, C.S., and Wang, P.G. “Identification of a New Fucosyltransferase (WbwK) Involved in O-Antigen Biosynthesis of Escherichia coli O86:B7 and Formation of H-type III Blood Group Antigen” *Biochemistry*, **2008**, In press.

3. **Yi, W.**, Shen, J., Li, J. and Wang, P. G. “Bacterial Homologue of Human Blood Group A Transferase” *J. Am. Chem. Soc.* **2008**, In press.
4. **Yi, W.**, Liu, X., Li, Y., Zhou, G., Xia, C., Li, J., Chen, X. and Wang, P.G., “Remodeling Bacterial Polysaccharides by Metabolic Pathway Engineering” Submitted., **2008**.
5. **Yi, W.**, Woodward, R., Eguchi, H., Perali, R.S. and Wang, P.G., “In vitro Reconstitution of Bacterial Polysaccharide Biosynthetic Pathway” Submitted, **2008**.
6. Su, D., Eguchi, H., **Yi, W.**, Li, L., Wang, P.G., and Xia, C., “Enzymatic Synthesis of Tumor-Associated Carbohydrate Antigen Globo-H Hexasaccharide” *Org. Lett.*, **2008**, 10, 1009-1012.
7. **Yi, W.**, Perali, R., Eguchi, H., Motari, E., Woodward, R., and Wang, P.G., “Characterization of a Bacterial β 1,3-galactosyltransferase With Application in the Synthesis of Tumor-Associated T-Antigen Mimics” *Biochemistry*, **2008**, 47, 1241-1248.
8. Guo, H., **Yi, W.**; Song, J. and Wang, P.G. “Current Understanding on Biosynthesis of Microbial Polysaccharides” (Invited Review) *Curr. Top. Med. Chem.*, **2008**, 8, 141-151.
9. Li, M., Liu, X., Shao, J., Jia, Q., Shen, J., **Yi, W.**, Chow, C., and Wang, P.G., “Characterization of a Novel α 1,2-Fucosyltransferase of Escherichia coli O128:B12 and Functional Investigation of its Common Motif” *Biochemistry*, **2008**, 47(1), 378-387.
10. Nagy, J.O., Zhang, Y., **Yi, W.**, Liu, X., Motari, E., Song, J., LeJeune, J.T., and Wang, P.G., “Glycopolydiacetylene nanoparticles as a biosensor to detect Shiga-like toxin producing Escherichia coli O157:H7” *Bioorg. Med. Chem. Lett.*, **2008**, 18, 700-703.
11. Tang, K-H., Guo, H., **Yi, W.**, Tsai, M-D., and Wang, P.G., “Investigation of the Conformational States of the Wzz and Wzz-O-Antigen Complex Under Near Physiological Conditions” *Biochemistry*, **2007**, 46(42), 11744-11752.
12. Guo, H.; **Yi, W.**; Li, L., and Wang, P.G. “Three UDP-hexose 4-epimerases with overlapping substrate specificity coexist in E. coli O86:B7” *Biochem. Biophys. Res. Comm.*, **2007**, 256(3), 604-609.
13. **Yi, W.**; Zhu, L.; Guo, H.; Li, M.; Shao, J.; Li, J.; and Wang, P. G. “Formation of a new O-polysaccharide in Escherichia coli O86 via disruption of

- glycosyltransferase gene involved in O-unit assembly” *Carbohydr. Res.*, **2006**, 341, 2254-2260.
14. **Yi, W.**; Yao, Q.; Zhang, Y.; Motari, E.; Lin, S. and Wang, P.G. “The wbnH gene of *E. coli* O86:H2 encodes an α -1,3-N-acetylgalactosaminyl transferase involved in the O-repeating unit biosynthesis” *Biochem. Biophys. Res. Comm.*, **2006**, 344, 631-639.
 15. **Yi, W.**; Bystricky, P; Yao, Q.; Guo, H.; Zhu, L.; Li, H.; Shen, J.; Ryu, K.; Li, M.; Ganguly, S.; Bush A. C. and Wang, P. G. “Two different O-polysaccharides from *Escherichia coli* O86 are produced by different polymerization from the same O-unit” *Carbohydr. Res.*, **2006**, 341, 100-108.
 16. Guo, H.; Lokko, K.; Zhang, Y.; **Yi, W.** and Wang, P.G. “Overexpression and characterization of Wzz of *Escherichia coli* O86:H2” *Protein Expr. Purif.*, **2006**, 48, 49-55.
 17. **Yi, W.**; Shao, J.; Zhu, L.; Li, M.; Singh, M.; Lu, Y.; Lin, S.; Li, H.; Ryu, K.; Shen, J.; Guo, H.; Yao, Q.; Bush, C.A.; and Wang P. G., “*Escherichia coli* O86 O-antigen biosynthetic gene cluster and stepwise enzymatic synthesis of human blood group B antigen tetrasaccharide” *J. Am. Chem. Soc.*, **2005**, 127, 2040-2041.
 18. Guo, H.; **Yi, W.**; Shao, J.; Lu, Y.; Zhang, W.; Song, J. and Wang, P.G. “Molecular analysis of the O-antigen gene cluster of *Escherichia coli* O86:B7 and characterization of the chain length determinant gene (wzz).” *Appl. Environ. Microbiol.*, **2005**, 71(12), 7995-8001.
 19. Ryu, K., Lin, S., Shao, J., Song, J., Chen, M., Wang, W., Li, H., **Yi, W.**, and Wang P.G., “Synthesis of complex carbohydrates and glyconjugates: enzymatic synthesis of globotetraose using alpha-1,3-N-acetylgalactosamin-yltransferase LgtD from *Haemophilus influenzae* strain Rd” *Methods Mol. Biol.*, **2005**, 310, 93-105.
 20. Shao, J., Li, M., Jia, Q., **Yi, W.**, and Wang, P.G., “Genetic basis of *E. coli* O128 polysaccharide biosynthesis gene cluster” *Polymer Preprints* (American Chemical Society, Division of Polymer Chemistry), **2003**, 44, 562-563.

FIELDS OF STUDY

Major Field: Biochemistry

TABLE OF CONTENTS

	<u>Page</u>
Abstract.....	ii
Dedication.....	iv
Acknowledgments.....	v
Vita.....	vii
List of Schemes.....	xv
List of Tables	xvi
List of Figures.....	xvii
 CHAPTERS:	
1. INTRODUCTION	1
1.1. Glycan Antigens	2
1.1.1. Tumor-Associated Antigens	2
1.1.2. ABH Blood Group Antigens.....	4
1.2. Bacterial Cell Surface Polysaccharides	5
1.3. Biosynthesis of Polysaccharides	8
1.4. Current Understanding of <i>wzy</i> -dependent Mechanism.....	12
1.5. Chemical Approaches in Glycobiology.....	13

1.5.1. Synthesis of Structurally Defined Glycans	15
1.5.1.1. Chemical Synthesis	16
1.5.1.2. Enzymatic Synthesis	19
1.5.2. Unnatural Sugars to Modulate Glycan Biosynthesis	22
1.5.3. Chemistry in Living Systems	24
1.6. Outline of the Work in this Thesis	29
2. RECONSTITUTION OF POLYSACCHARIDE BIOSYNTHETIC PATHWAY	30
2.1. Introduction	30
2.2. Structural Analysis of <i>Escherichia coli</i> O86 O-Polysaccharides	32
2.2.1. Experimental Methods	33
2.2.2. Results	36
2.2.3. Discussion	43
2.3. Sequencing of Biosynthetic Gene Cluster	44
2.3.1. Experimental Methods	45
2.3.2. Results and Discussion	46
2.4. Biochemical Characterization of Glycosyltransferase Genes	50
2.4.1. Experimental Methods	50
2.4.2. Results and Discussion	53
2.5. Reconstitution of Repeating Unit Biosynthesis In Vitro	63
2.5.1. Experimental Methods	63
2.5.1.1. Chemical Synthesis of GalNAc-PP-Und	63
2.5.1.2. Enzymatic Sequential Assembly of Repeating Unit	69
2.5.2. Results and Discussion	70

2.6. Over-expression and Detection of Wzy Membrane Protein.....	73
2.6.1. Experimental Methods	74
2.6.2. Results and Discussion	75
2.7. In Vitro Polymerization of Repeating Unit by Wzy.....	77
2.7.1. Experimental Methods	77
2.7.2. Results and Discussion	78
2.8. In Vitro Reconstitution of Polysaccharide Chain Length Modality	80
2.8.1. Experimental Methods	80
2.8.2. Results and Discussion	81
3. ENGINEERING POLYSACCHARIDES BY METABOLIC PATHWAY	
REMODELING	84
3.1. Introduction	84
3.2. Experimental Methods.....	88
3.3. Results and Discussion	93
3.3.1. Disruption of GDP-Fucose De Novo Pathway in <i>E. coli</i> O86.....	93
3.3.2. Functional Expression of GDP-Fucose Salvage Pathway in <i>E. coli</i> O86	
mutant	94
3.3.3. Promiscuity of fucose salvage pathway towards fucose analogs.....	96
3.3.4. Incorporation of fucose analogs into polysaccharides	98
3.3.5. Fluorescence labeling of azido containing LPS by click chemistry	101
3.3.6. Live cell labeling by click chemistry	103
4. ENZYMATIC SYNTHESIS OF BIOMEDICALLY IMPORTANT	
OLIGOSACCHARIDES	106

4.1. A bacterial β 1,3-Galactosyltransferase and Synthesis of Tumor Associated T-antigen Mimics	107
4.1.1. Introduction.....	107
4.1.2. Experimental Methods.....	109
4.1.3. Results.....	115
4.1.4. Discussion.....	124
4.2. Bacterial Homologue of Human Blood Group A Transferase	126
4.2.1. Introduction.....	126
4.2.2. Experimental Methods.....	128
4.2.3. Results and Discussion	132
5. CONCLUSIONS AND PERSPECTIVES.....	140
REFERENCES	142
APPENDIX: NMR and MS Spectra.....	156

LIST OF SCHEMES

<u>Scheme</u>	<u>Page</u>
2.1 In vitro enzymatic reaction of WbnH	55
2.2 In vitro enzymatic reaction of WbnJ	59
2.3 In vitro enzymatic reaction of WbnK	60
2.4 In vitro enzymatic reaction of WbnI	62
2.5 Synthesis of the sugar phosphate, peracetylated GalNAc-phosphate	64
2.6 Synthesis of the lipid phosphate moiety, Und-P	65
2.7 Synthesis of GalNAc-PP-Und	65
3.1 Two pathways for GDP-fucose biosynthesis	87
4.1 One-pot two enzyme system for synthesis of blood group A antigens	134
4.2 In vivo conversion of B-antigen to A-antigen	135

LIST OF TABLES

<u>Table</u>		<u>Page</u>
2.1	Chemical shift assignments for the O-PS of <i>E. coli</i> O86 :K61:B7 strain	39
2.2	HMBC peaks indicating glycosidic linkages.....	39
2.3	Summary of <i>E. coli</i> O86 O-antigen biosynthesis genes	48
2.4	Substrate specificity of WbnH.....	57
2.5	Substrate specificity of WbnI and WbnK.....	63
4.1	WbiP mutants and primers used in the site-directed mutagenesis.....	107
4.2	Substrate specificity of WbiP.....	112
4.3	Kinetic parameters of WbiP and mutants	117
4.4	Substrate specificity, kinetics and yields of BgtA-catalyzed reactions	127

LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
1.1	Interactions between glycans and receptors mediate biological processes.....	2
1.2	Examples of tumor associated glycan antigens	3
1.3	Biosynthesis of blood group A/B antigens	5
1.4	LPS molecules on Gram-negative bacterial cell surfaces.....	6
1.5	CPS on Gram-positive bacterial cell surfaces.....	7
1.6	Biosynthesis of bacterial polysaccharides	9
1.7	Formation of desired stereoisomers with chemical glycosylation.....	16
1.8	Programmable one-pot synthesis of Globo H hexasaccharide	17
1.9	Automatic solid-phase synthesis of GPI anchor	18
1.10	Glycosyltransferase (A) and glycosidase (B) catalyzed glycosylation.....	19
1.11	Glycosynthase catalyzed glycosylation	21
1.12	Unnatural substrate metabolic intervention of glycan synthesis.....	24
1.13	Amine reactive conjugation reaction	25
1.14	Selective chemical reactions with ketone (aldehyde)	26
1.15	Azide mediated bio-orthogonal chemical reactions.....	27
2.1	<i>E. coli</i> O86 O-antigen repeating unit and O-PS structures	33

2.2	ELISA end-point titers of anti-blood group B antibody with LPS	40
2.3	Molecular modeling of tetrasaccharide fragments of O-PS from <i>E. coli</i> O86:B7 and O86:H2 strains	42
2.4	O-antigen gene cluster of <i>E. coli</i> O86:B7	46
2.5	Expression and purification of WbnH	54
2.6	Identification of disaccharide product catalyzed by WbnH using LC-MS.....	54
2.7	Determination of divalent metal ion requirement for WbnH activity.....	58
2.8	Expression and purification of GST-WbnJ.....	59
2.9	Expression and purification of GST-WbnK	60
2.10	Expression and purification of WbnI.....	61
2.11	In vitro reconstitution of <i>E. coli</i> O86 repeating unit oligosaccharide.....	71
2.12	In vivo disruption and complementation of <i>wzy</i> gene	75
2.13	Overexpression and detection of Wzy	76
2.14	Radioactive based paper chromatography assay of Wzy polymerization	78
2.15	Analysis of Wzy polymerization reaction by SDS-PAGE and visualized with auto-radiography	79
2.16	Pull down assay for probing Wzy-Wzz interactions.....	82
2.17	In vitro generation of chain length modality.....	83
3.1	Strategy for in vivo generation of structurally modified polysaccharides.....	85
3.2	Incorporation of fucose analogs into <i>E. coli</i> O86 polysaccharides.....	86
3.3	Disruption of GDP-fucose <i>de novo</i> pathway and complementation with salvage pathway	94
3.4	Expression and purification of Fkp.....	95

3.5	In vitro Fkp reaction with fucose as the substrate.....	96
3.6	Fucose analogs used in the PS engineering	97
3.7	<i>In vitro</i> analysis of the specificity of Fkp toward fucose analogs.....	97
3.8	Silver staining analysis of LPSs containing different fucose analogs	98
3.9	Analysis of structurally modified polysaccharides by mass spectrometry	100
3.10	Fluorescence labeling of LPS using in vitro click chemistry	101
3.11	Labeling azido containing LPS.....	102
3.12	Control experiments demonstrates specific labeling	102
3.13	Biotin linked reagents for live cell labeling.....	104
3.14	Live cell labeling flow cytometry data	104
4.1	O-polysaccharide structure of <i>E. coli</i> O127 and its biosynthetic gene cluster	109
4.2	WbiP catalyzed reaction and HPLC identification of product.....	116
4.3	Metal ion requirement of WbiP	120
4.4	Amino acid sequence alignment of segments from different β 1,3- galactosyltransferases	120
4.5	Enzymatic activity of WbiP mutants	122
4.6	UDP-beads binding assay	124
4.7	LPS analysis of different <i>E. coli</i> strains.....	136
4.8	LPS ELISA with blood group antibodies	137
4.9	MS analysis of polysaccharide structures	138

CHAPTER 1

INTRODUCTION

Like nucleic acids and proteins, glycans in the form of oligosaccharides and glycoconjugates (glycoproteins and glycolipids) are vital biopolymers found in organisms across all domains of life. They play critical roles in mediation of numerous complex biological processes [1] (Figure 1.1). For example, in eukaryotes, glycans decorate cell surface proteins and secreted proteins to participate in molecular recognition events, such as attachment of immune cells, bacteria and viruses [2, 3]. They also serve as determinants of protein folding, stability, and pharmacokinetics [4]. In addition, specific changes in glycan profiles have been associated with certain disease states such as cancer and inflammation, illustrating the potential of using glycans in clinical diagnosis and perhaps as targets to develop therapeutics [1, 5, 6]. Moreover, the discovery and intense research efforts on O-linked β -N-acetylglucosamine protein modification in the past couple of decades highlight the importance of glycans involved in signal transduction, transcription regulation, stress response, apoptosis and immune cell maturation [7]. The past several decades has witnessed an explosion of interest in glycobiology. A number of experimental tools such as synthetic chemistry, molecular biology, peptide/protein

chemistry, genetic manipulation, structural biology, small molecule perturbation and systems biology have begun to shed light on the function of glycans in molecular levels.

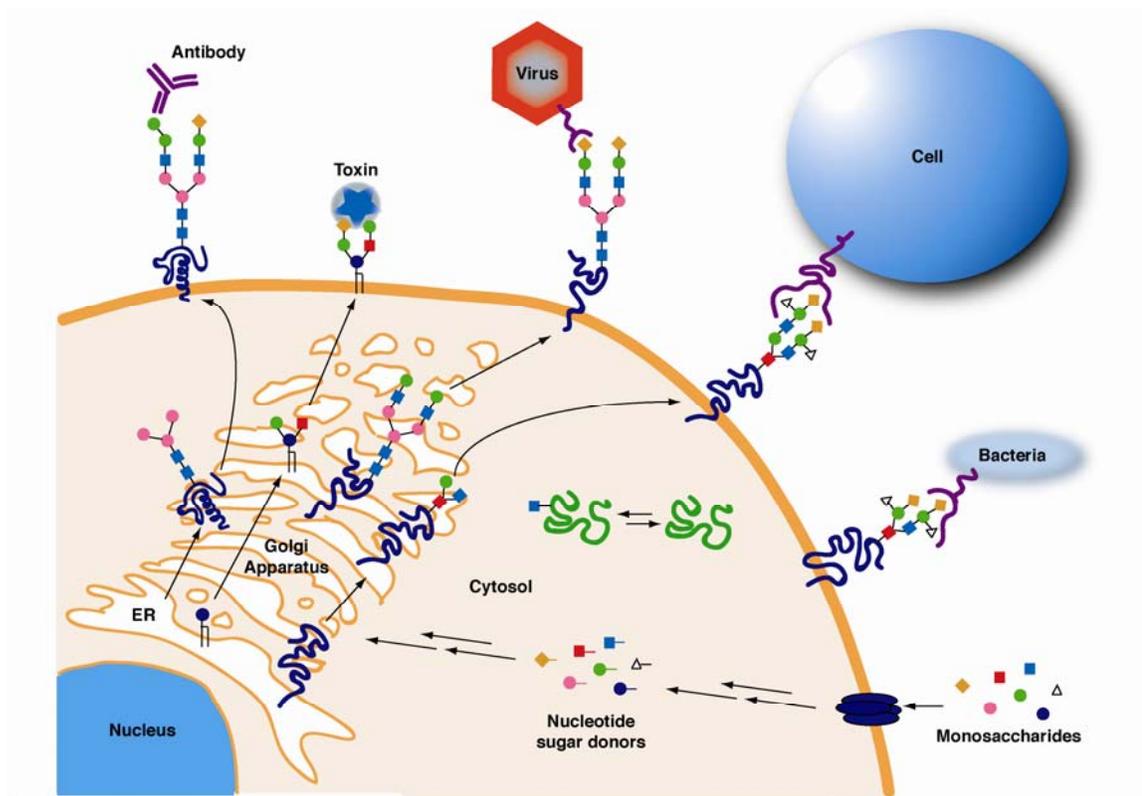


Figure 1.1. Interactions between glycans and receptors mediate biological processes.

1.1 Glycan antigens

1.1.1 Tumor-associated antigens

Altered expression of glycans constitutes a hallmark of the tumor phenotype [8]. These changes include both the under- and over-expression of naturally occurring glycans, as well as novel glycan structures primarily restricted to embryonic tissues [5]. As a result, these tumor-associated glycan structures can serve as diagnostic markers for cell malignancy [5]. Since cancer cell glycans differ from those found in the healthy

tissues, it is possible to elicit the immune system to target cancer cells based on the altered glycan structures [8]. This notion has been put to test by the development of glycan-based anti-cancer vaccines, and has shown some exciting promise in early clinical trials [9-12].

Several glycan structures are commonly found on malignant cells. They include Tn antigen, T antigen, sialyl Lewis x, Lewis y, gangliosides, Globo-H and polysialic acid (PSA) (Figure 1.2). Differential expression of these glycans is associated with different types of cancer. For example, T antigen has been shown to express in >90% of primary human carcinomas, and high expression levels of T antigen are correlated with increased metastasis and aggressiveness of breast and colon cancers [13]. Complex gangliosides (for example, GD2, GD3 and fucosyl GM1) are found at elevated levels in small-cell lung cancer, neuroblastomas and melanomas [14]. Globo-H is normally associated with advanced breast, prostate, lung and ovary cancers [15, 16].

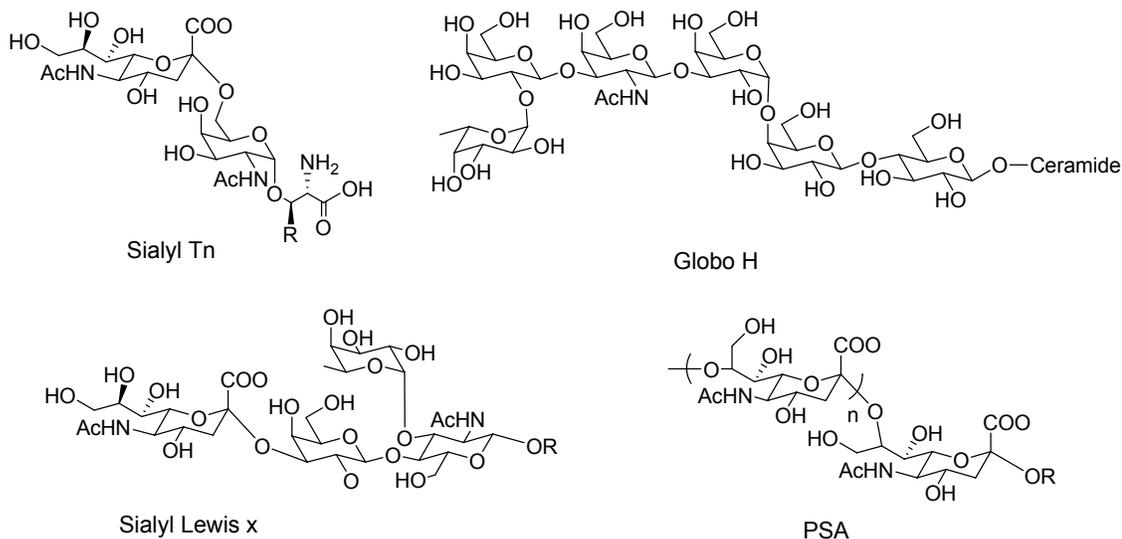


Figure 1.2. Examples of tumor associated glycan antigens.

These altered glycan structures have been exploited as targets for cancer immunotherapy via the development of glycan-based vaccines[12]. Several of these vaccines are presently undergoing clinical evaluations and have shown some promise. For example, administration of the sTn-KLH conjugate combined with hormone therapy had improved survival rates with a time-to-disease progression of 8.3 months compared with 5.8 months for those undergone hormone therapy alone[9]. Globo-H conjugates were administered with the immunologic adjuvant QS-21 as a vaccine for patients with prostate cancer in clinical trial I and were shown to be safe and capable of inducing specific high-titer IgM antibodies against Globo-H [17]. In addition, a tri-antigenic vaccine containing Globo-H, Lewis y and Tn, developed by Danishefsky and coworkers has been shown to elicit IgG antibodies against each glycan antigen in a preclinical study [18]. These promising results open up an exciting opportunity to pursue a nontoxic immuno treatment of human cancers.

1.1.2 ABH Blood group antigens

ABH system is one of the most common and important blood group systems in transfusion medicine. The structures of the ABH glycan epitopes are defined as GalNAc- α 1,3(Fuc- α 1,2)-Gal β (A antigen), Gal- α 1,3(Fuc- α 1,2)-Gal β (B antigen) and Fuc- α 1,2-Gal β (H antigen) [19]. A and B antigens are formed by the transfer of GalNAc and Gal to the H antigen disaccharide by human blood group transferase GTA and GTB, respectively [20, 21] (Figure 1.3). The ABH antigens are expressed on the surface of red blood cells and many other tissues, including vascular endothelium and a variety of epithelia. The naturally occurring anti-A and anti-B antibodies exist in individuals who

do not express the corresponding antigens, the principle of Landsteiner's Law that states that only matched donor blood can be safely transfused. ABH matching is not critical not only in blood transfusion, but in tissue and organ transplantation.

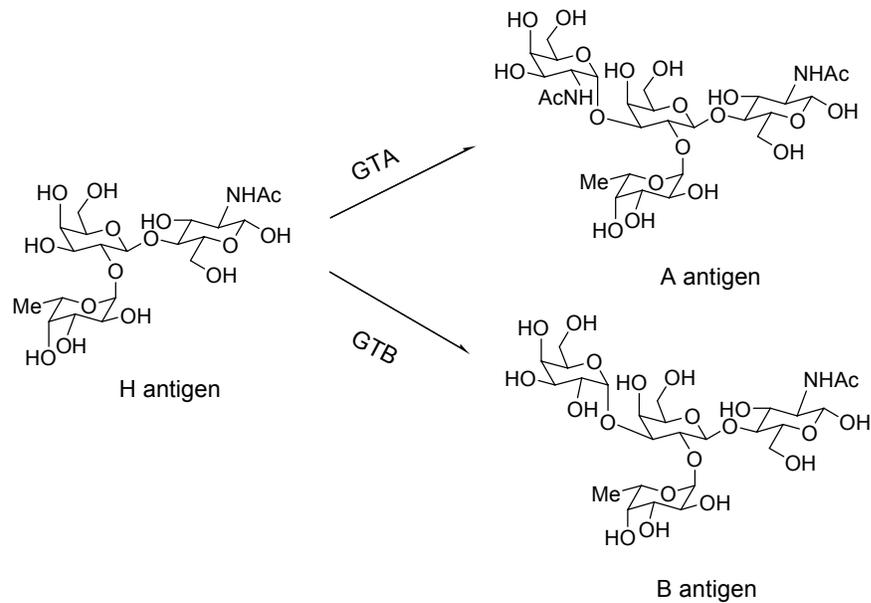


Figure 1.3. Biosynthesis of blood group A/B antigens.

1.2 Bacterial Cell Surface Polysaccharides

The bacterial cell surface is decorated with remarkable variations of polysaccharide structures [22]. Based on the way in which they are associated with the cell surface, bacterial polysaccharides can be classified into three major groups [23]: O-polysaccharides (O-PS), capsular polysaccharides (CPS), and exopolysaccharides (EPS). These glycans mediate the direct interactions between the bacterium and its host, and have been implicated as an important factor in the virulence of many plant and animal pathogens [23].

O-polysaccharide (O-PS): O-polysaccharide (also called O-antigen) is a major component of the bacterial cell surface lipopolysaccharide (LPS) (Figure 1.4). It is composed of multiple copies (as many as 100) of an oligosaccharide repeating unit (O-unit). O-PS are presented on the cell surface via the covalent attachment to the core oligosaccharide and the lipid-A, both of which are essential components of LPS. O-PS is the most structurally varied among different bacterial species [24]. For example, there are more than 180 O-PS structures found in *E. coli*, and more than 90 in *Samonella*. The variations arise from different sugar compositions, positions, and stereochemistries of the glycosidic linkage within the O-unit, as well as the linkages between them. These extensive structural variations of the O-PS allow further classification of the bacterial species into different O-serological groups such as *E. coli* O86 and O157. An increasing body of evidence suggests that the O-PS plays an important role in mediating bacteria-host interactions such as the effective colonization of the host and the resistance to complement-mediated immune responses [24].

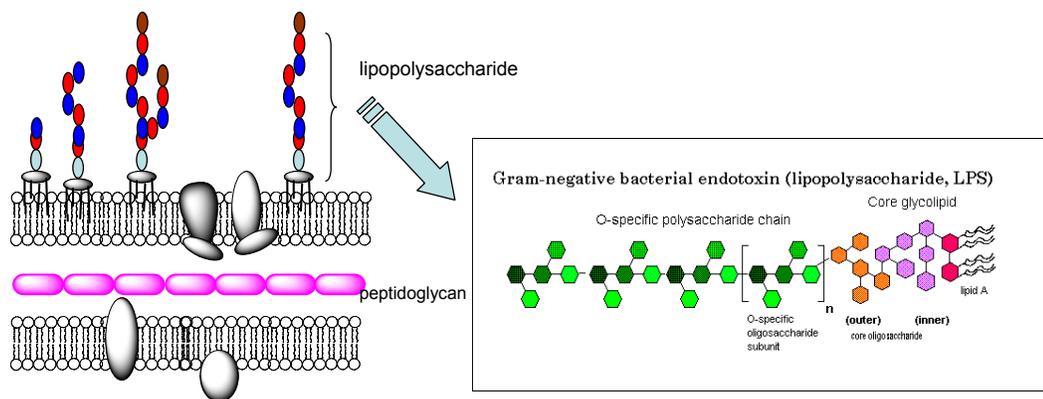


Figure 1.4. LPS molecules on gram-negative bacterial cell surface.

Capsular Polysaccharide (CPS): Capsular polysaccharides are linked to the cell surface via covalent attachments to either phospholipids or lipid-A molecules [25] (Figure 1.5). They are linear polymers and also consist of multiple copies of repeating units. They are highly hydrated and can be substituted by non-sugar residues such as acetyl, acyl, phosphate and sulfate. Capsular polysaccharides also display enormous structural variations. *E. coli* can produce over 80 chemically and serologically different capsular polysaccharides, which are further divided into four major classes based on genetic criteria [22]. *S. pneumoniae*, a gram-positive pathogen that causes pneumonia, meningitis and blood stream infections, can express more than 90 structures [26]. Furthermore, in infectious bacterial diseases, CPSs have been shown to be essential virulence factors. This property has been exploited to develop polysaccharide-based vaccines against infectious diseases such as pneumonia, meningitis and sepsis [27]. Current pneumococcal vaccines contain CPS purified from 23 of the 90 recognized serotypes and are proven to exert serotype specific protection against 90% of the disease causing serotypes.

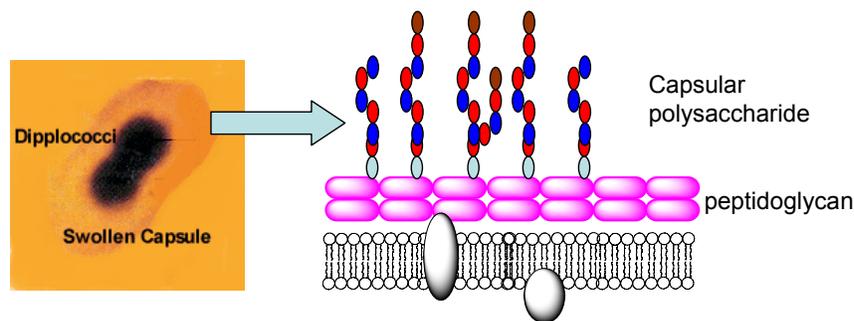


Figure 1.5. CPS on gram-positive bacterial cell surfaces.

Exopolysaccharide (EPS): Many microorganisms are capable of producing exopolysaccharides, which are secreted outside the cells [23]. EPSs are released onto the cell with no visible means of attachment and are often sloughed off to form slime. EPSs have a much higher molecular weight compared to both O-PSs and CPSs, possessing more than 1000 repeating units. Expression of EPS has important applications in both biomedical research and the food industry [28, 29]. For example, EPSs are critical for biofilm formation, and are key components of the biofilm matrix in many biofilms [30, 31]. Furthermore, EPSs also play an important role in immune evasion and tolerance toward antibacterial agents. The food industry uses EPSs for their unique properties to create viscosifiers, stabilizers, emulsifiers, or gelling agents.

1.3 Biosynthesis of Polysaccharides

In spite of the remarkable structural diversity of the O-PS, CPS and EPS, only three major elongation/polymerization mechanisms (discussed below) are utilized in nature, as shown by extensive genetic studies [23]. This implies that nature has come up with only a few solutions to the ubiquitous problem of polysaccharide biosynthesis. In this proposed research program, we will mainly focus on the synthesis of O-PS since O-PS has been used as a model system to study polysaccharide biosynthesis in the past decades. Usually, O-PSs are synthesized separately before being transferred to the core-Lipid A to form LPS. The genes involved in the synthesis of O-PS are clustered at the locus historically known as *rfb* [32]. This locus encodes enzymes involved in biosynthesis of sugar nucleotide precursors, glycosyltransferases that synthesize O-unit or

nascent O-PS, and enzymes required for polysaccharide processing. There are three enzymatic stages in the biosynthesis (Figure 1.6).

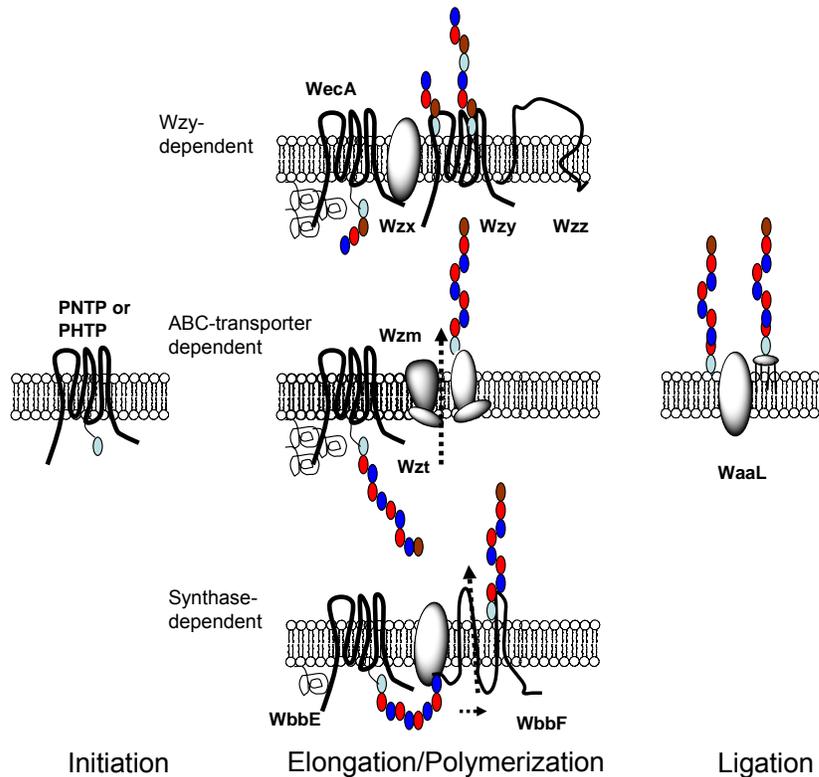


Figure 1.6. Biosynthesis of bacterial polysaccharides.

Initiation: The initiation reaction has conserved features of all three pathways and involves the addition of a sugar phosphate to an undecaprenyl phosphate (Und-*P*) to form an und-PP-linked sugar [33, 34]. This reaction takes place at the cytosolic face of the plasma membrane where the sugar nucleotides are available. Depending on the identity of the sugar residue (usually an N-acetyl hexosamine or hexose), the reaction is catalyzed by WecA or WbaP which belong to the PNTP and PHTP protein families, respectively.

Elongation/translocation/polymerization: Depending on the respective assembly and translocation mechanisms during the elongation step, O-PS biosynthesis is further subdivided into three pathways: *Wzy*-dependent, ABC-transporter dependent and synthase-dependent (Figure 1.6).

Wzy-dependent: Following the initiation reaction, the subsequent steps involve the sequential addition of different sugar residues onto the Und-PP-sugar intermediate via catalysis by specific glycosyltransferases. These glycosyltransferases are predicted to be soluble or peripheral membrane bound proteins. Upon completion of the Und-PP-O-units, they are translocated to the periplasmic side of the inner membrane by the flippase **Wzx** [35]. In the periplasm, the Und-PP-O-units are polymerized into a sugar polymer by the integral membrane protein **Wzy** [36, 37]. The polymerization reaction involves the transfer of the nascent polymer from its Und-PP carrier to the non-reducing end of the new Und-PP-O-units. The final component of the *wzy*-dependent pathway is the **Wzz** protein, which works as a chain length regulator to generate strain specific polysaccharide chain lengths [38, 39].

ABC transporter-dependent: In this pathway, the nascent polysaccharide chain is synthesized at the cytoplasmic face of the inner membrane. Specifically, following the chain initiation by *WecA*, a specific glycosyltransferase adds a sugar residue as an adaptor between Und-PP-GlcNAc and a repeating unit domain. The subsequent chain extension is accomplished by the processive transfer of residues to the non-reducing terminus of the Und-PP-linked acceptor, catalyzed by either monofunctional or multifunctional glycosyltransferases. Following the polymerization, the nascent polysaccharide is translocated across the inner membrane by an ABC-2 subfamily of

ABC transporters [40, 41]. However, the exact mode of action and the spatial organization of the transporters still remain elusive.

Synthase-dependent: This pathway occurs the least in polysaccharide synthesis [24]. It involves the participation of the WbbF of the synthase family, which is comprised of integral membrane proteins. These proteins are characterized by two separate but conserved domains. One domain is responsible for the processive glycosyl transfer of the sugar residues resulting in chain elongation, and the other is implicated in the mediation of the nascent polysaccharide translocation across the inner membrane.

Termination and Ligation: The ligation step is common to all three pathways, and involves the transfer of the O-polysaccharide to the nascent core-Lipid A. The ligation reaction occurs at the periplasmic face of the cytoplasmic membrane, but its mechanism is currently unclear. WaaL is currently the only known protein required for the ligation [42, 43]. WaaL proteins exhibit low levels of similarity in their amino acid sequences, and are predicted to be integral membrane proteins with 8 or more transmembrane segments. It has been reported that WaaL proteins lack discrimination for donor structures, which suggests that they recognize the Und-P carrier rather than the saccharide attached to it [43]. Translocation of complete LPS to the outer membrane involves the protein complex Imp/RlpB [44].

1.4 Current Understanding of *wzy*-dependent Mechanism

The *wzy*-dependent mechanism is arguably the most widely used system in polysaccharide biosynthesis [45]. Its characteristic features are the presence of the O-unit flippase (Wzx), the polymerase (Wzy), and the chain length regulator (Wzz) in the gene cluster.

Based on the predicted hydrophobicity and topology, **Wzx** proteins with 11 or 12 transmembrane segments were classified into a family of integral membrane proteins. They exhibit little nucleotide sequence homology, and are considered as specific genes to distinguish different O-serogroups. However, they do share some distant structural features with bacterial permease and transporters of major facilitators. The *wzx* deficient strain accumulated O-unit intermediates on the cytoplasmic side of the inner membrane, which appears to support the role of Wzx as a flippase [46]. It is proposed that Wzx facilitates the transit of Und-PP-O-units to the periplasmic side using a proton or electrochemical gradient as the source of energy [24].

Wzy is proposed to function as polymerase. Wzy proteins are also integral membrane proteins with 11-13 predicted transmembrane helices, with little primary sequence homology. Wzy was previously proposed to have strict substrate specificity for cognate O-units, however, recent studies from our lab and from other groups showed that Wzy has a relatively relaxed substrate specificity [47, 48]. Wzy is the key enzyme in the elucidation of the biosynthetic mechanism of O-PS; however, its function has not yet been studied *in vitro* due to the following difficulties: (1) The structurally defined substrates are difficult to obtain; (2) It is difficult to isolate or express Wzy in practical

amounts; and (3) a sensitive polymerization assay is yet to be developed. In the past two years, we have made significant progress in overcoming these obstacles.

Wzz proteins belong to a super family of Polysaccharide Co-Polymerases (PCP), which function in determining the modality of a variety of polysaccharides [49]. The secondary structural features of Wzz proteins are well conserved. Wzz homologs are anchored in the cytoplasmic membrane by two transmembrane helices flanking a large hydrophilic loop located in the periplasm [50]. No clear domain or active site has yet been identified to impart the modality. Several mechanisms have been proposed for the role of Wzz in chain length determination. Bastin *et al.* [51] suggested that Wzz functions as a molecular timer that modulates the Wzy polymerase activity between two states that favor either elongation or the transfer to the ligase. Morona *et al.*, however, suggested that the Wzz protein functions as a molecular chaperone to assemble a complex consisting of Wzy and WaaL, and that the specific modality was a result of different ratios of Wzy to WaaL [50]. Each of these models suggested the interaction of Wzz with Wzy or WaaL to form a functional complex; however, neither model possesses definitive experimental evidence. Since the superfamily of PCP represents a new mechanism of action, elucidation of the Wzz function will have broad implications in polysaccharide biosynthesis.

1.5 Chemical Approaches in Glycobiology

In biology, genetic and traditional biochemical approaches have contributed substantial insight into the function of biopolymers and the underlying molecular mechanisms of a number of physiological processes. For example, molecular biology

techniques, such as gene cloning, overexpression, disruption and silencing, allow the perturbation of specific proteins and thus their subsequent functional assignment. Fluorescence imaging technique via using GFP genetic tag has been demonstrated in numerous studies of protein localization, trafficking, and association in both cells and living organisms [52, 53]. Moreover, DNA microarray technologies have served as a powerful tool in the system-wide profiling of transcripts in the field of genomics.

Unfortunately, these powerful technologies cannot be directly applied to the study of glycans. There are several reasons: firstly, glycans are not primary gene products; their biosynthesis is neither under direct transcriptional control nor template-driven, making their genetic manipulation far more complicated than for nucleic acids or proteins. Disruption of genes encoding glycan biosynthetic enzymes oftentimes causes the impaired synthesis of multiple glycan structures or, in some cases, embryonic lethality, making the interpretation of mutant phenotypes very difficult or impossible [54]. Secondly, glycans from natural sources exhibit enormous heterogeneity, making the functional assignment of specific glycan structures rather elusive. Finally, visualization of glycans is not trivial in the cellular level. The traditional GFP fusion tag, which has been widely used in monitoring protein trafficking, cannot be used for glycan imaging. Therefore, new sets of approaches that are complementary to the traditional genetic, molecular biology approaches need to be developed to define glycan functions at the molecular level.

Chemical approaches are emerging to be powerful alternatives to genetics and biochemistry in the study of glycobiology [55]. For example, small molecule inhibitors can disrupt or modulate biosynthesis of glycans in a spatial and temporal specific manner

[56, 57]. Synthesis of chemically defined oligosaccharides and glycoconjugates offers an unprecedented opportunity to unravel their specific functions [58, 59]. Glycomimetics modulate or interrogate cell surface glycan-receptor binding events [60, 61]. Unnatural sugars that substitute for native monosaccharides can intercept biosynthetic pathways, modulate glycosylation events and endow novel functionalities to glycoconjugates [62, 63]. Collectively, chemical tools have opened up exciting revenue to understand the intricacy of glycan structure and functions. In this chapter, instead of providing a comprehensive review of the current myriad of chemical tools used in glycobiology, I have chosen to highlight two of the most fundamental and widely employed approaches: synthesis of structurally defined oligosaccharides/glycoconjugates to dissect glycan biosynthesis, and use of unnatural sugar analogs to modulate the cell surface glycosylation. These are also the tools I have used in my research.

1.5.1 Synthesis of structurally defined glycans

Access to structurally defined glycans is a prerequisite for deconvoluting the biological roles of this class of molecules. A number of elegant methods/methodologies have been developed to synthesize complex glycan structures. Advances in this front are providing valuable materials for the assignment of glycan function, elucidation of structure-function relationship, dissection of biosynthetic pathways, and the production of glycan based anti-bacterial and anti-cancer vaccines. Two general strategies are used for in vitro synthesis of glycans: chemical synthesis and enzymatic (including chemo-enzymatic) synthesis.

1.5.1.1 Chemical synthesis

Complex glycans were and still remain a daunting class of targets for synthetic chemistry. Part of the challenge stems from the fact that each sugar molecule contains several hydroxyl groups of similar reactivity. The construction of oligosaccharides, especially branched structures, requires exquisite choice of protecting groups to differentiate different hydroxyl groups. In addition, it also requires regiochemical and stereochemical control in glycosidic linkage formation. The development of orthogonal hydroxyl group protecting strategies greatly aids the synthesis of regioselective glycosidic linkages (Figure 1.7). The principle is that protecting groups can be installed to block reaction at one site and later removed to unmask specific hydroxyl groups for glycosylation. To obtain desired stereochemistry of a glycosidic bond, more sophisticated choice of protecting groups is required. It usually involves using protecting groups to direct the orientation of the glycosidic bond via intermolecular or intramolecular participation, altering the steric environment of the anomeric position to bias the desired outcome. Nowadays, two main strategies have developed to illustrate the abovementioned principle: one-pot solution-based synthesis and solid-phase synthesis.

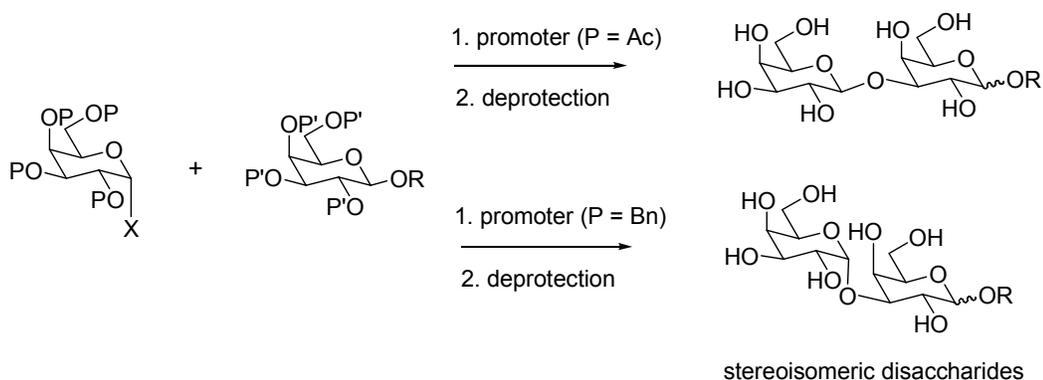


Figure 1.7. Formation of desired stereoisomers with chemical glycosylation.

In one-pot synthesis, glycosylation occurs sequentially in a single reaction vessel, based on the reactivity of protected monosaccharides. A key concept underlying this process is that there are “armed” (reactive) and “disarmed” (less reactive) glycosyl donors [64]. The most reactive glycosyl donor undergoes the first coupling reaction, followed by the less reactive one, and terminated by the least reactive donor. A number of groups have exploited this concept to assemble fairly complex glycan structures in an elegant fashion. One of these efforts is Wong’s reactivity-based and programmable approach [65].

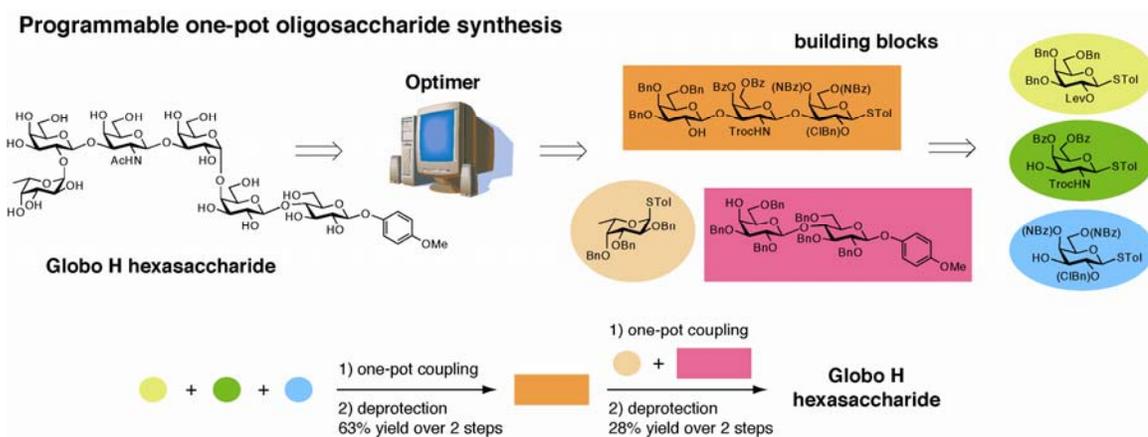


Figure 1.8. Programmable one-pot synthesis of globo H hexasaccharide.

This approach (Figure 1.8) involves sequential assembling of sugar building blocks with decreasing reactivity predicted by computer (programmable). The reactions are then performed in solution such that the oligosaccharide chain is extended from the non-reducing end to the reducing end. This approach eliminates the intermediate steps and tedious protecting group manipulation, providing relatively rapid access to complex oligosaccharides. The power of this strategy has been demonstrated in the synthesis a

variety of tumor associated glycan antigens such as Lewis y, fucosyl GM1 and Globo-H [66].

Solid-phase synthesis offers powerful advantages over the solution phase synthesis, since it circumvents multiple purification steps. One of the elegant examples is Seeberger's automatic solid phase synthesis [66] (Figure 1.9). This strategy uses glycosyl phosphates as building blocks with reducing end sugar attached to the solid support at the anomeric position, and allows the stepwise incorporation of one mono- or disaccharide building blocks at a time. The UV active protecting groups are used to monitor the reaction. The productivity of this strategy is evident from the diversity and complexity of oligosaccharides that have been synthesized. However, not all the oligosaccharides can be assembled onto the solid support. Moreover, the most consuming step is the synthesis of sufficient quantities of the building blocks.

Automated solid-phase oligosaccharide synthesis

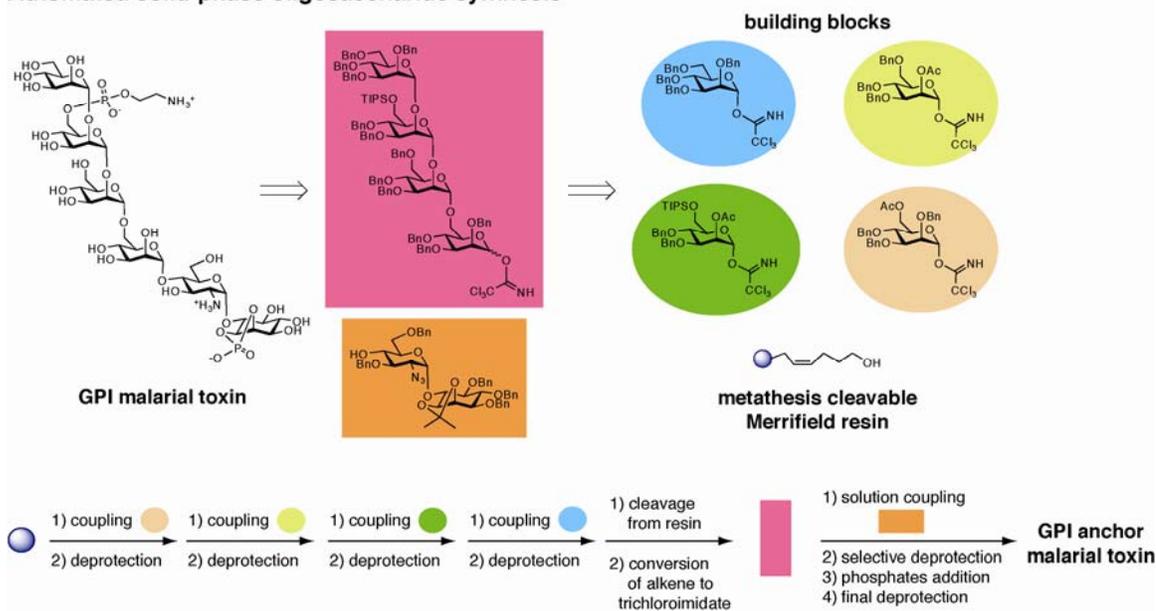


Figure 1.9. Automatic solid-phase synthesis of GPI anchor.

1.5.1.2 Enzymatic synthesis

Enzymatic synthesis is an attractive and powerful alternative to the chemical synthesis [67, 68]. Compared to the chemical approach, enzymatic approach has obvious advantages of achieving high regio- and stereochemistry of glycosidic bonds, often in a high efficient manner. Enzymes used for glycan synthesis are typically grouped into two major categories: glycosyltransferases and glycosidases.

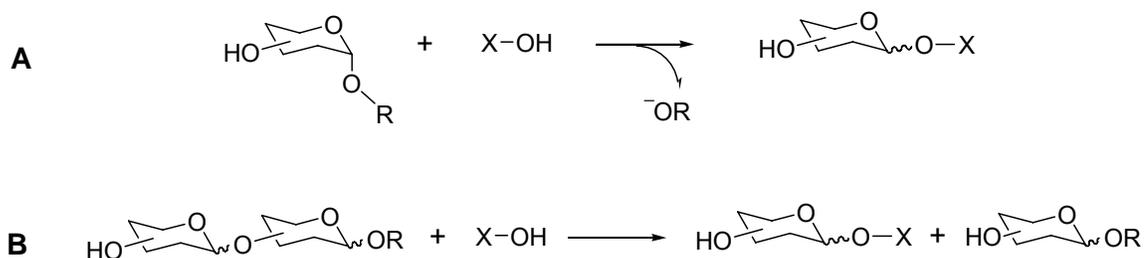


Figure 1.10. Glycosyltransferase (**A**) and glycosidase (**B**) catalyzed glycosylation.

Glycosyltransferases: This class of enzymes is most widely used in glycan synthesis [69, 70]. They transfer a given monosaccharide from the corresponding sugar nucleotide (sugar donor) to a specific hydroxyl group of a sugar acceptor (Figure 10). Glycosyltransferases exhibit very strict stereo- and regiospecificity. Moreover, they can catalyze the formation of either retention or inversion of configuration at the anomeric carbon of the sugar residue. A large number of eukaryotic glycosyltransferases have been cloned to date and used in large-scale syntheses of oligosaccharides and, in general, exhibit exquisite linkage and substrate specificity. It should be noticed that such a considerable number of mammalian enzymes have converged on only nine common

sugar nucleotides as donor substrates. Glucosyl-, galactosyl-, and xylosyltransferases employ substrates activated with uridine diphosphate as the anomeric leaving group (UDP-Glc, UDP-GlcNAc, UDP-GlcA, UDP-Gal, UDP-GalNAc, and UDP-Xyl), whereas fucosyl- and mannosyltransferases utilize guanosine diphosphate (GDP-Fuc and GDP-Man). Sialyltransferases are unique in that the glycosyl donor is activated by cytidine monophosphate (CMP-Neu5Ac).

It is found that mammalian glycosyltransferases are relatively hard to express in large quantities and are rigid in terms of substrate specificity. On the other hand, the rapid growing number of sequenced bacterial genomes and polysaccharide biosynthetic gene clusters provides a rich pool for the identification of bacterial glycosyltransferases with a wide spectrum of acceptor specificity. In contrast to their mammalian counterparts, bacterial glycosyltransferases are more easily overexpressed as soluble and active forms without complicated gene manipulation techniques. Furthermore, they appear to have broader substrate specificity, thereby offering tremendous advantages in the enzymatic synthesis of oligosaccharides and their analogs. In addition, various bacteria exhibit structural mimicry of mammalian glycans on cell surfaces [71-74]. Therefore, corresponding glycosyltransferases can be explored for the synthesis of human-like glycans. The drawbacks of using glycosyltransferases for glycan synthesis are that this class of enzymes is delicate and sugar nucleotides are expensive. Reactions can be affected by the feedback inhibition of resulting by-product nucleoside phosphates. Cost-effective strategies to facilitate large scale synthesis have been developed using recombinant enzymes, relatively inexpensive alternative sugar nucleotides, and in situ regeneration and recycling of sugar nucleotides [75, 76].

Glycosidases: Glycosidases are a class of enzymes responsible for glycan processing reactions during glycoprotein synthesis. The physiological function of these enzymes is the cleavage of glycosidic linkages. However, under controlled conditions, glycosidases can be used to synthesize glycosidic bonds rather than cleavage (Figure 1.10). Therefore, these enzymes have been employed as catalysts in oligosaccharide synthesis [77-80]. Glycosidases are widely available, robust, and require only inexpensive donor substrates. Although glycosidases are generally stereospecific, they only have weak regiospecificity, which may result in the formation of multiple products. In addition, this class of enzymes has two key limitations: challenge of driving the reaction in a thermodynamically disfavored direction and the enzymatic degradation of glycosidic linkages.

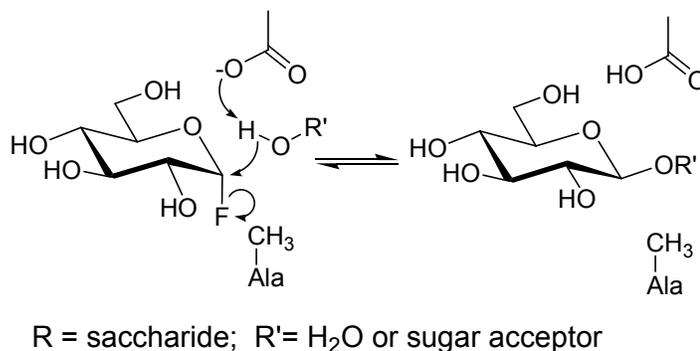


Figure 1.11. Glycosynthase catalyzed glycosylation.

In the past decade or so, efforts have been put forth to engineer glycosidases to overcome those intrinsic limitations. The new class of mutant enzymes, termed glycosynthases, have the ability to catalyze the formation of glycosidic linkages but do

not hydrolyze the newly formed bonds, thereby driving the reaction in the synthetic direction [81]. Glycosynthases are rendered hydrolytically incompetent via the mutation of the nucleophilic residue to another one unable to perform the same hydrolytic function. When combined with the glycosyl fluoride substrates of the opposite anomeric configuration to that of the natural substrate, the enzyme is able to transfer this activated glycosyl donor to a suitable acceptor (Figure 1.11). The catalytic activities of glycosynthases have been further improved and altered by random mutagenesis coupled with high throughput screening of mutants [82, 83]. For example, a recent work has expanded the synthetic scope of glycosynthases from oligosaccharides to glycosphingolipids [84]. In addition, a new class of glycosynthases that catalyzes the efficient synthesis of thioglycosidic linkages was reported [85]. These enzymes have double mutations in which both the catalytic nucleophile and acid/base residues are replaced by residues incapable of carrying out these functions. Thioglycosidic linkage can be synthesized using a combination of glycosyl fluoride donors and nucleophilic thiosugars.

1.5.2 Unnatural sugars to modulate glycan biosynthesis

The ability to alter glycan structures expressed on cell surfaces represents a powerful tool to understand their biological functions. Two existing strategies have been developed to modulate the cell surface glycan structures: small molecule inhibition and unnatural metabolic substrate intervention [62, 63]. Due to the limited scope and relevance of this thesis, only the latter strategy will be discussed.

The concept of unnatural metabolic substrate intervention has been exploited in three different areas (Figure 1.12). 1) Hijacking endogenous glycan synthesis. In this strategy, cell-permeable simple sugars (typically comprising hydrophobic glycosides) will act as decoys for glycosyltransferases and thus compete with endogenous substrates for their activity [86]. When applied to cells, these exogenous substrate primers suppress the elaboration of endogenous glycans, thereby perturbing the structures destined for the cell surface. 2) Termination of glycan elongation. This strategy involves the design of an unnatural substrate that serves as a chain terminator for glycan synthesis [87]. Similar to dideoxy nucleotides in a DNA polymerase reaction, deoxy or other similarly modified monosaccharides can be incorporated into glycans via the cell's metabolic machinery, resulting in truncation of glycans. 3) Incorporation of novel functional groups. Monosaccharide substrates appended with novel functional groups can be incorporated into glycans by the cell's biosynthetic pathway [88]. The result is the presentation of an unnatural epitope that might display different receptor binding properties than its native counterpart, resulting in differential cell-cell interaction profile and the subsequent modulation of communication signals [89, 90]. In addition, incorporation of sugars with bio-orthogonal functional groups, such as keto-, alkyl-, azide-, provides a scenario in which the glycan structure can be further altered by selective chemical reactions at the cell surface [62]. Such technique, also called bio-orthogonal chemical reporter strategy, has been recently widely used in probing glycosylation, glycan imaging and glycoproteomics study. Central to the concept of unnatural metabolic substrate intervention is the substrate promiscuity of glycan biosynthetic enzymes. Even though the detailed studies on specificity of glycan biosynthetic pathway enzymes are still lacking at

this point, fortunately, however, most of the glycan biosynthetic pathways do not appear to possess strict specificity towards unnatural substrates.

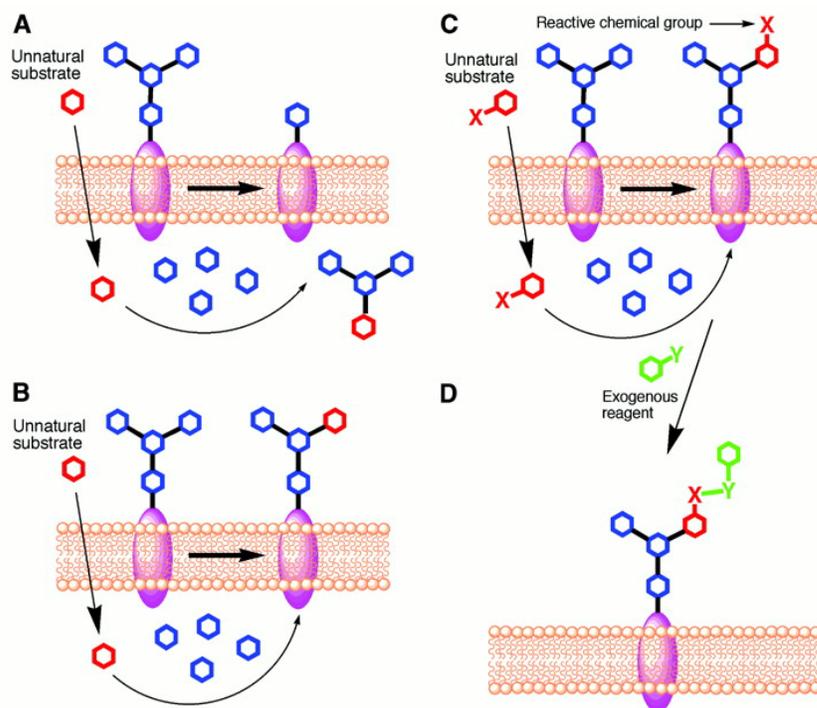


Figure 1.12. Unnatural substrate metabolic intervention of glycan synthesis (from Ref 62).

1.5.3 Chemistry in living systems

Living cells are a factory of chemical reactions. They contain enormous abundance of nucleophiles, reducing agents and functionalities. Therefore, the choice of suitable components for selective chemical transformations in cells is far from obvious. First of all, the chemical reporter molecules (unnatural metabolic substrates) must possess

adequate metabolic stability and permeability for use in cells. Secondly, the unnatural substrates and their reactive partners should be mutually reactive in a physiological environment (37 °C, pH 6-8) and, at the same time, remain inert to the surrounding biological environment. Finally, the reaction between the substrate and its partner must be highly selective, with fast reaction kinetics and capability of forming stable covalent linkages. In consistent with these criteria, discussed in the following are the most commonly used chemical reactions in living systems.

Amine and activated ester: Primary amines can selectively react with activated ester (for example, N-hydroxyl succinimide, NHS) in pH 7-9 buffers to form stable amide bonds (Figure 1.13). Proteins, including antibodies, generally have abundant primary amines in the side chain of lysine residues and the N-terminus of each polypeptide that can react with NHS-activated reagents. This reaction has been widely used in protein labeling with biotin and fluorescent reagents for visualization, protein PEGylation, and addition of other novel functional groups.

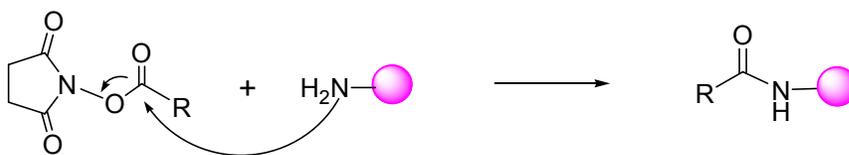


Figure 1.13. Amine reactive conjugation reaction.

Oxime chemistry with ketones and aldehydes: Ketones and aldehydes are normally absent in native cellular environment. The small size of these functional groups renders

them not only suitable for tagging proteins, but also glycans and other metabolites [91]. These mild electrophiles are attractive choices for modifying biomolecules due to their readiness to be incorporated into diverse molecular scaffolds. In addition, they are inert to the reactive moieties normally found in proteins, lipids and glycans. Although these groups can form reversible Schiff bases with primary amines, the equilibrium favors the carbonyl in water. In chemical reactions carried out *in vitro*, stable Schiff bases can be obtained by reacting with hydrozide and aminoxy groups to form hydrazones and oximes, respectively (Figure 1.14). Although suitable for chemical modifications in cells, the application of this reaction is somewhat limited. The optimal pH is 5-6, values that cannot be achieved inside the cell. Therefore, ketones and aldehydes are best used in environs devoid of carbonyl electrophiles, namely, on cell surfaces or in the extracellular environment.

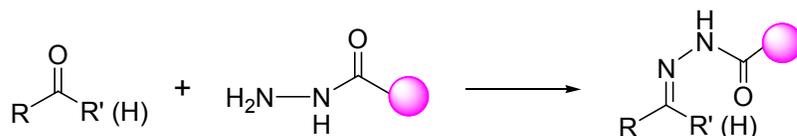


Figure 1.14. Selective chemical reactions with ketone (aldehyde).

Azides: Azides are the most viable and versatile chemical reporters for labeling all classes of bio-molecules in biological environment. This azide functional group is abiotic in animals and absent from nearly all naturally occurring species. Azides are metabolically stable in aqueous environment and are resistant to oxidation. Azides are susceptible to reduction by free thiols, such as glutathione in the cell context. However,

reduction of azides typically requires stringent conditions (100°C, several hours) or auxiliary catalysts [92], which would not be a concern in physiological environments. The bio-orthogonality has been exploited to develop selective chemical reactions, including the Staudinger ligation with functionalized phosphines [93] and the [3+2] cycloaddition with activated alkynes (“Click Chemistry”) [94] (Figure 1.15).

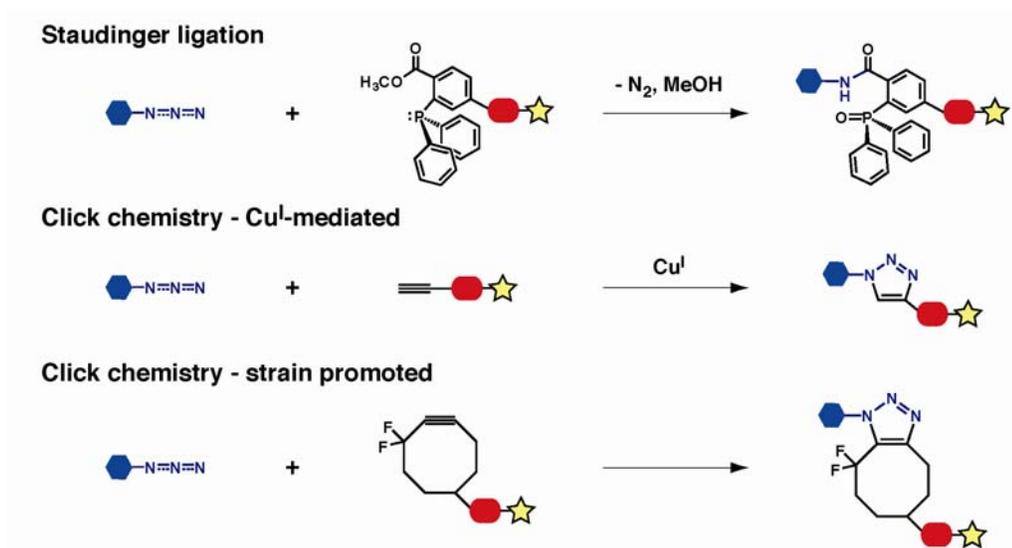


Figure 1.15. Azide mediated bio-orthogonal chemical reactions.

Staudinger ligation. Discovered in 1919 by Hermann Staudinger, and recently modified by Carolyn Bertozzi, this transformation produces a covalent linkage between one nitrogen atom of the azide and the triarylphosphine scaffold. Phosphines do not react appreciably with biological functional groups and are therefore bio-orthogonal. The reaction proceeds at pH 7 with no apparent cellular toxicity. The Staudinger ligation has been widely used in labeling of proteins, glycans, live cells and even live organisms [95, 96].

Click Chemistry. Azides can also undergo reactions with dipolarophiles such as activated alkynes. Described by Huisgen more than four decades ago, this [3+2] cycloaddition between azides and terminal alkynes produces stable triazole adducts. The reaction is thermodynamically favorable. On the other hand, in order to achieve fast kinetics, Cu(I) catalyst is often required. Sharpless and coworkers demonstrated that the rate of this reaction can be accelerated $\sim 10^6$ -fold using catalytic amount of Cu(I) [97]. This copper-mediated reaction has been witnessing an increasing application in labeling nucleic acids, proteins, glycans, viral particles, and live cells [94, 98, 99]. The primary advantage of click chemistry over the Staudinger ligation is its faster kinetics, a 25-fold increase in cell lysates [100]. Moreover, its superior sensitivity provides the most efficient detection of low abundance species. The disadvantage of click chemistry is the cellular toxicity conferred by the Cu catalyst. More recently, to overcome the metal toxicity issue, the Bertozzi lab developed a copper-free strain-promoted cycloaddition [101]. The reaction is activated not by metal catalyst, but by ring strain. The rate of the reaction can be increased by appending electron-withdrawing groups (such as -F) to the octyne ring. This copper-free reaction holds the promise of integrating the low toxicity of the Staudinger ligation and fast kinetics, superior sensitivity of the traditional click chemistry.

1.6 Outline of the work described in the thesis

The work in this thesis mainly focuses on using chemical approaches to understand and modulate glycan structure, synthesis and function. More specifically, in Chapter 2, the most widely occurring polysaccharide biosynthetic pathway was investigated. Using *E. coli* O86 O-polysaccharide as a model, we first elucidated its chemical structure and the corresponding synthetic gene cluster. Glycosyltransferase genes in the gene cluster were then biochemically characterized and applied in the sequential chemo-enzymatic synthesis of repeating unit substrates. With structurally defined substrates, polymerization was demonstrated *in vitro*. This study represents a chemical reconstitution of polysaccharide biosynthesis in a cell free system. In Chapter 3, a sugar nucleotide synthetic pathway was engineered, and its intrinsic promiscuity was exploited to *in vivo* incorporate a panel of unnatural sugar substrates into bacterial polysaccharides in a highly homogenous manner, resulting in the generation of structurally novel polysaccharides. Furthermore, selective chemical reactions were carried out *in vitro* to further elaborate the functional groups appended to the polysaccharide. In Chapter 4, two novel microbial glycosyltransferases were characterized and applied in the facile synthesis of two types of important human oligosaccharide antigens: T-antigen mimics and human blood group antigens. This study represents an ongoing effort in our lab to synthesize biomedically important oligosaccharides with microbial glycosyltransferases.

CHAPTER 2

RECONSTITUTION OF POLYSACCHARIDE BIOSYNTHETIC PATHWAY

2.1 Introduction

Polysaccharides play an important role in bacterial survival, mediation of bacterium-host interaction, regulation of host immune system and bacterial virulence [25, 102]. Therefore, understanding polysaccharide biosynthesis represents a critical step towards identification of viable targets for antibiotic development and facilitation of anti-bacterial vaccine design. Genetic studies, based on polysaccharide synthetic gene cluster sequencing, putative enzyme annotation, and gene knockout/complementation approaches, have provided remarkable insight and greatly enriched the current understanding of polysaccharide synthesis. Despite the extensive genetic studies, the current synthetic model has never been biochemically reconstituted in a cell free *in vitro* system. Chemical approaches using homogenously synthesized oligosaccharide substrates and purified enzymes offer a powerful complement to genetic studies, and can provide unambiguous biochemical evidence to delineate the molecular details of the pathway. Several main difficulties hamper the efforts to reconstitute polysaccharide biosynthesis *in vitro*: 1) structurally defined repeating unit substrates are difficult to

obtain; 2) Over-expression and isolation of Wzy membrane protein has been very challenging; and 3) a sensitive *in vitro* polymerization assay has yet to be developed.

In this study, using *E. coli* O86 O-PS as a model system, we investigated polysaccharide biosynthesis (particularly the *wzy*-dependent pathway) with a combination of chemical and biochemical approaches. Firstly, *E. coli* O86 O-PS chemical structure was characterized by NMR, MS and methylation analysis. Its immuno reactivity with respect to the polysaccharide structure was also investigated. Secondly, the polysaccharide synthetic gene cluster was elucidated using shot-gun sequencing strategy. The possible function of each gene was predicted based on the sequence comparative analysis of databases. Thirdly, biochemical studies were carried out to unambiguously identify specific functions of putative glycosyltransferase genes in the cluster. Fourthly, repeating unit oligosaccharide was synthesized *in vitro* by sequential employment of identified glycosyltransferases from a diphospho-lipid linked precursor. This step represents the reconstitution of repeating unit assembly pathway. Next, the sugar polymerase Wzy was over-expressed using a novel chaperone co-expression system, a significant advance towards isolation of homogenous Wzy protein. With the synthetic repeating unit substrates and partially purified Wzy, we demonstrated *in vitro* formation of polysaccharides using a radio-active based assay, providing definitive evidence for Wzy polymerization activity. Finally, effect of lipid structure on polymerization was investigated using a series of repeating unit analogs with different lipids varied in length, saturation degree and double bond geometry. Collectively, this study demonstrates the first detailed biochemical investigation of *wzy*-dependent polysaccharide biosynthesis. This *in vitro* system provides the groundwork necessary for further probing the molecular

details of sugar polymerization, also represents one step forward towards understanding polysaccharide biosynthesis in general.

2.2 Structural Analysis of *E. coli* O86 O-polysaccharides

The choice of *E. coli* O86 is due to its role as a popular bacterial strain in the study of bacterium-host interaction, especially its strong association with the stimulation of human anti-blood group antibodies. Springer and coworker [103, 104] during 1960s raised an intriguing hypothesis that human naturally occurring anti-A and B antibodies were produced in response to the constant enteric bacterial stimulation within the intestinal flora. He demonstrated that a variety of gram-negative bacteria showed blood group activities, among which *E. coli* O86 has high B blood group activity [105]. By using *E. coli* O86 as a model bacterial strain, he studied the immunological response of humans and germ-free chicken with the inoculation of *E. coli* O86 [103, 104]. The resulting high titer of anti-blood group B antibody supported the hypothesis that bacteria stimulate human natural blood group antibody production. It was shown later that the strong blood group B activity of *E. coli* O86 resulted from the O-polysaccharide structure that resembles the blood group B antigen epitope [106]. The complete structure (Figure 2.1) of O-polysaccharide from a serogroup O86:H2 was elucidated and confirmed the presence of blood group B antigen epitope in O-repeating unit. In this work, we report a new O-polysaccharide structure obtained from a different serotype *E. coli* O86:K61:B7 (B7 for short), which is different from the published structure. Interestingly, comparison of these two structures showed that they could be derived from the same repeating unit via different polymerization sites from different sugar residues (Figure 2.1). The

immuno-activity of these two forms of LPS towards anti-B antibody was then studied and compared. The different antibody binding affinity is explained via the molecular modeling of the tetrasaccharide fragment that contains blood group B epitope.

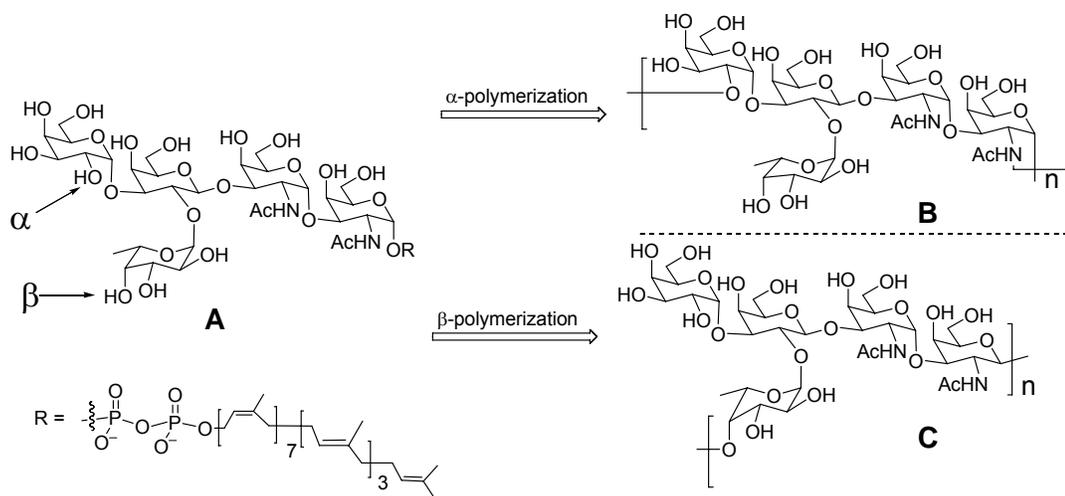


Figure 2.1. *E. coli* O86 O-antigen repeating unit and O-PS structures. A, repeating unit; B, O-PS from *E. coli* O86:K61:B7; C, O-PS from *E. coli* O86:K2:H2.

2.2.1 Experimental Methods

Bacterial strains and antibody

E. coli O86:K61:B7 (ATCC 12701) was obtained from American Type Culture Collection (Rockville, MD). *E. coli* O86:H2 was kindly provided by Dr. Willie Vann of FDA, Bethesda. Commercially available IgM monoclonal anti-B antibody (obtained from clone HEB-29) was purchased from Abcam Ltd (Cambridge, UK).

Isolation of LPS and O-polysaccharide

Dried bacterial cells (20 g) were extracted with 50% (w/v) aqueous phenol (200 mL) at 65°C for 15 min and cooled to 4°C, followed by low speed centrifugation (10,000 g). The top aqueous solution was separated and dialyzed against distilled water until free from phenol. The clear solution was then lyophilized. The products were then dissolved in 0.02 M sodium acetate (pH = 7.0, 20 mL) and treated sequentially for 2 hr each with DNase, RNase and proteinase K at 37°C. After removal of the precipitated materials, the solution was subjected to ultracentrifugation (110,000 g, 4°C, 12 hr). The precipitated gels were then dissolved in distilled water and lyophilized to produce 1.2 g pure LPS. The LPS was degraded with aqueous 1% acetic acid at 100°C for 1.5 hr. The precipitate was removed by centrifugation (12,000 g, 30 min). The carbohydrate portion was purified using Bio-Gel P2 column, and yielded 160 mg purified O-polysaccharide.

Structural determination of O-polysaccharides

High performance anion exchange chromatography (HPAE) was carried out on a Dionex CarboPack PA-10 (4 mm) anion exchange HPLC (HPAE) column with electrochemical detection on ED40 pulsed amperometric detector (PAD) followed by a PDR chiral optical rotation detector (ORD) connected in-line. Then 2 mg of polysaccharide was hydrolyzed with trifluoroacetic acid (2 M) for 2 hrs at 100°C. The retention times were compared to those of the standard monosaccharides: fucose, 2-amino-2-deoxy-galactose, 2-amino-2-deoxy-glucose, galactose, glucose, rhamnose, 2-amino-2-deoxy-mannose, mannose and fructose. The optical rotation chromatogram was used for determination of absolute configuration of a given monosaccharide.

For NMR experiments, 30 mg of O-polysaccharide was dissolved in 0.6 mL of 99.96 % D₂O after being lyophilized against 99.9 % D₂O. For the structure determination of the O86 O-polysaccharide repeating unit, a suite of standard NMR experiments was employed including ge-DQF COSY, ge-TOCSY, NOESY, ge-HSQC and ge-HMBC all in phase sensitive modes acquired on Bruker DRX AVANCE spectrometer operating at 500 MHz ¹H frequency with cryoprobe detection. NOESY spectra were taken with mixing times of 80, 50, 30 and 15 ms and cross-peaks scored according to their integrated volumes as strong (s), medium (m) and weak (w). All spectra were taken at elevated temperature of 50°C due to dynamic properties of the polysaccharide. Data were processed with NMRPipe and NMRDraw under Linux. The observed chemical shifts were referenced to internal acetone at 2.225 and 31.07 ppm for ¹H and ¹³C, respectively. Methylation analysis was done by means of GC-MS of partially methylated alditol acetates performed at Complex Carbohydrate Research Center at University of Georgia.

ELISA of LPS with anti-B antibody:

The LPSs from *E. coli* strains (O86:B7, O86:H2, K12) were dissolved in PBS buffer pH=7.5, and were coated onto flat-bottomed microtiter plates for 48hrs at 4°C at concentration of 10 µg/mL. Then 2% Polyvinylpyrrolidone (PVP) in PBS buffer was used to block the nonspecific binding for 3 hr. The anti-B antibody described above was used in series dilution. ELISA end-point titers were determined as described, employing secondary antibody goat anti-mouse IgM conjugated to HRP (1:1000) (Sigma). The peroxidase substrate (3,3',5,5'-tetramethylbenzidine, Sigma) was used to develop the signal, which was monitored at 405 nm.

Molecular modeling of tetrasaccharide fragments

Molecular modeling of the O86 tetrasaccharide fragments was carried out to interpret the NOESY data. The structures were generated using Macromodel Version 7.0 and minimized in 10000 cycles of Monte Carlo steps with MM2 force field with a continuum solvent model for water. 6700 structure models with energy under 50 kJ/mole above the global minimum were kept for the MC chain. The NOE data were then reconciled with the molecular models and scored as follows. Strong NOEs were assumed to be spaced at 2.4 Å, medium NOEs at 2.8 Å and weak NOEs at 3.1 Å. The sum of the squared deviations for all NOEs (χ NOE) was considered as the NOE score for each model.

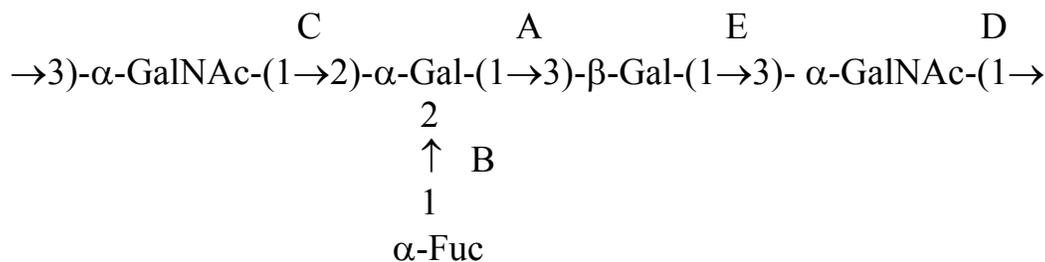
2.2.2 Results

Structure determination of O-PS from E. coli O86:K61:B7

Initial inspection of NMR spectra of B7 O-polysaccharide indicated disagreement of both the ^1H and ^{13}C spectra with the published *E. coli* O86:H2 structure by Andersson et al. Since we have no reason to question either the NMR data or the structure determination reported by Andersson et al. [106], we concluded that the O-polysaccharide we have isolated from B7 strain differs from that reported in the earlier study which utilized a clinical isolate. Therefore we treated this structure as unknown and carried out a full structural determination.

Representative HPAE chromatograms (Appendix) of hydrolyzed O86 polysaccharide shows Gal, Fuc and GalN in a ratio of 2:1:2, suggesting a carbohydrate composition identical to that reported structure. The optical rotation chromatogram indicates the absolute configurations of these residues are L-Fuc, two D-Gal and two D-

GalN. The HSQC spectrum reveals five peaks in the anomeric chemical shift region between 5.5-4.5 ppm (^1H) and 104-89 ppm (^{13}C) suggesting there are five monosaccharide residues in the repeating subunit. We annotated the anomeric proton signals by letters A through E (Appendix). Both the ^1H and ^{13}C chemical shifts as well as the ^1H coupling constants indicate that residues A-D are α -anomers while residue E is a β -anomer. The complete ^1H and ^{13}C spectrum assignment was determined by standard methods and the results at 50°C are given in Table 2.1. Examination of the ^1H coupling constants was used to assign the stereochemistry of each of the five ring systems and the 6-deoxy sugar was recognized by the characteristic chemical shifts of the methyl group. The amino sugars were identified by the characteristic chemical shifts of C2. The methyl group signals at 2.04 and 2.12 ppm indicate that both amino sugars are N-acetylated. We therefore identify the sugars residues as α -Gal (A), α -Fuc (B), α -GalNAc (C and D) and β -Gal (E) in agreement with the HPLC data. HMBC spectra were used to confirm the assignment given in Table 2.1, which accounts for all the magnetization observed in the HSQC spectrum (Appendix).



O86	H-1	H-2	H-3	H-4	H-5	H-6	NHCOCH ₃	
	C-1	C-2	C-3	C-4	C-5	C-6		
A α-D-Gal	5.447	3.99	4.04	4.01	4.19	3.76		
	90.1	72.3	69.3	70.3	71.9	62.3		
B α-L-Fuc	5.308	3.76	3.605	3.66	4.27	1.15		
	99.8	69.0	71.1	73.0	67.7	16.4		
C α-D-GalNAc	5.232	4.43	4.03	4.13	4.02	3.76		2.12
	94.8	48.7	73.0	65.6	71.9	62.3		23.7
D α-D-GalNAc	5.066	4.215	3.94	4.22	3.91	3.76	2.04	
	94.0	49.8	75.6	69.9	71.9	62.1	23.1	
E β-D-Gal	4.693	3.86	3.925	4.27	3.65	3.76		
	103.5	74.0	76.5	63.9	75.8	61.9		

Table 2.1. Chemical shift assignments for the O-PS of *E. coli* O86 :K61:B7 strain.

The methylation analysis shows the presence of five components, namely, 4,6-di-*O*-methyl-galactose, 3,4,6-tri-*O*-methyl-galactose, 2,3,4-tri-*O*-methyl-fucose, 2-deoxy-2-*N*-methylacetamido-4,6-di-*O*-methyl-galactose and 2-deoxy-2-*N*-methylacetamido-3,4,6-tri-*O*-methyl-galactose in the ratio of 1.0:1.0:1.0:1.4:0.4.

	¹ H and ¹³ C chemical shifts	proton – carbon pair
anomeric proton region	5.45, 76.5 (weak)	AH1 → EC3
	5.31, 74.0	BH1 → EC2
	(5.23, 72.3) ^a	(CH1 → AC2) ^a
	5.07, 73.0	DH1 → CC3
anomeric carbon region	3.94, 103.5	EC1 → DH3
	3.86, 99.8	BC1 → EH2
	4.03, 94.0 (weak)	DC1 → CH3

Table 2.2. HMBC peaks indicating glycosidic linkages.

The glycosidic linkage positions of residues B, D and E are readily apparent from the HMBC data (Table 2.2). The NOE data (Appendix) are likewise in accordance with the assignments of B-E as α -Fuc-(1 \rightarrow 2)- β -Gal, of E-D as β -Gal-(1 \rightarrow 3)- α -GalNAc and D-C as α -GalNAc(D)-(1 \rightarrow 3)- α -GalNAc(C). The anomeric proton of residue A shows a weak HMBC cross-peak to C3 of residue E as well as strong NOE to H4 and weaker NOE to H3 of residue E. This pattern of strong NOE to an equatorial proton adjacent to the linkage position is characteristic of this stereochemical configuration and has been reported in previous studies of NOE in blood group B oligosaccharides^[107]. We can therefore assign the linkage of A-E as α -Gal-(1 \rightarrow 3)- β -Gal. This interpretation of the NOE data is confirmed in our molecular modeling study described below.

The HMBC as well as long-range single quantum correlation data do not give unambiguous evidence of the linkage of residue C. The anomeric proton of C shows cross-peaks only with C3 and C5 of that residue and no other signals. The anomeric carbon signal shows no cross-peaks. Strong NOE peaks are seen between CH1 and AH1, AH2, EH2 and EH4 suggesting that residue C is linked to either residue A or E. The clearest indication of the exact position of this linkage comes from methylation analysis which is consistent with the linkages B1-E2, C1-D3, D1-E3 and A1-E3. These assignments account for all the methylation analysis except for 2-linked galactose. Since the 2,3 substituted galactose is accounted for by residue E, we assigned the 2-substituted galactose as residue A. The linkage C-A as α -GalNAc-(1 \rightarrow 2)- α -Gal is further supported by the strong NOE between CH1 and AH1, a characteristic pattern of NOE to an

equatorial proton adjacent to the linkage position. Less prominent NOE is observed between CH1 and AH2, the linkage position. This assignment is also confirmed by our molecular modeling analysis below.

The structure reported in this work for *E. coli* O86:B7 strain differs from the reported structure in the linkage position and anomeric configuration of a single residue GalNAc (residue C in our designation). In our structure (Table 2.1) that GalNAc is linked by α 1 \rightarrow 2 linkage to α -Gal whereas in the published structure, that GalNAc is linked β 1 \rightarrow 4 to α -Fuc.

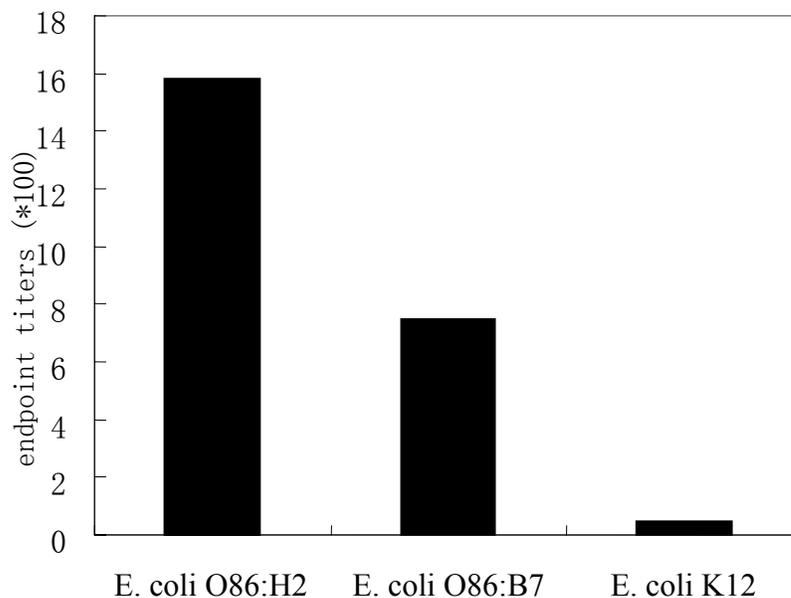


Figure 2.2. ELISA end-point titers of anti-blood group B antibody with LPS.

Two LPSs display different binding affinity with anti-blood group B antibody

In the B7 O-PS, the α -linked galactose residue is part of the polysaccharide backbone, in which its C2-OH has been substituted with GalNAc residue. This C2-OH position was demonstrated by Lemieux et al. [108] to be critical for the binding of anti-B antibody to the blood group B antigen. Therefore, it is envisaged that the LPS produced from these two *E. coli* O86 strains would display different binding affinity with anti-B antibody. To confirm this idea, LPS ELISA assay was carried out. *E. coli* K12 that has no reported blood group activity was used as a negative control. Antibody titration (Figure 2.2) showed that antiserum had an end-point titer of ~ 1600 against *E. coli* O86:H2 LPS, which is consistent with the previous study. Whereas the end-point titer against B7 LPS was two fold lower, *i.e.*, ~ 800 . These results showed that H2 strain possessed higher blood group B reactivity compared to B7 strain, and it further confirmed the critical role of C2-OH of the α -linked galactose residue in the blood group B epitope.

Molecular modeling of tetrasaccharide fragments from two O-polysaccharides

Molecular modeling of a tetrasaccharide fragment (residues A, B, C and E) from the proposed B7 O-PS structure was carried out (Figure 2.3) and the NOE predicted for each model was compared with that measured for the polysaccharide at 30 msec mixing time given in Appendix Table 3. Each molecular model in the Monte Carlo chain was scored following the rules given in the Experimental Part and the penalty scores plotted as a function of the glycosidic dihedral angles along with the calculated energies (Appendix). The models with the lowest values of NOE penalty are those with the lowest energy, especially for the residues of the blood group B trisaccharide composed of

residues A-E-B. This result is in agreement with a relatively rigid single conformation for the blood group B trisaccharide which has been proposed by NMR, x-ray crystallography and molecular modeling studies [107, 109]. The conformation of α -GalNAc-(1 \rightarrow 2)- α -Gal linkage between residues C-A has not been previously studied and, although our data indicate it may not be a single rigid conformer, the results do confirm the strong NOE between CH1 and AH1 which is caused by a very short distance between these two protons in the lowest energy model (2.167 Å). While the distance between CH1 and AH2 (2.723 Å) matched the NOE data well, some other NOE values are not well predicted by the single low energy model suggesting that some other conformers must contribute to the conformation of this linkage.

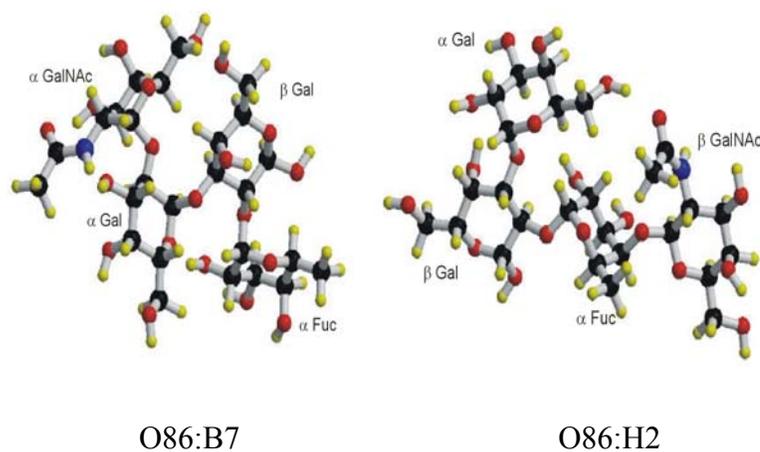


Figure 2.3. Molecular modeling of tetrasaccharide fragments of O-PS from *E. coli* O86:B7 and O86:H2 strains.

2.2.3 Discussion

The LPS heterogeneity is mostly found in the structural diversity of O-polysaccharide. The variations of O-antigens within one bacterial species can be exerted at different levels, such as non-stoichiometric modifications of polysaccharide backbones with glucosyl and fucosyl residues, change of anomeric linkages of the O-PS repeating unit, and change of major range of polysaccharide chain length [24]. A number of studies have been carried out to reveal the genetic basis of O-polysaccharide variations. Three possible mechanisms have been proposed [110]. They are lateral gene transfer followed by recombination, phage-induced conversion, and regulated expression of O-antigen biosynthetic genes. We report the finding of a new O-polysaccharide structure from *E. coli* O86 serogroup. The new O-polysaccharide is structurally related to the previous reported structure in that they can be derived from the same pentasaccharide repeating unit, but are polymerized from the same GalNAc residue to different terminal residues (Gal and Fucose) via α and β glycosidic linkages, respectively. The genetic basis of this structure difference is currently under investigation in our lab.

The blood group B reactivity of *E. coli* O86 comes from the presence of B antigen epitope in the O-polysaccharide structure. Despite the fact that the two forms of *E. coli* O86 polysaccharide share the same O-repeating unit, they display different binding affinity towards anti-B antibody, resulting from the different polymerization. The structural difference of these two polysaccharides becomes more evident when we inspect the molecular models of two tetrasaccharide fragments which contain the B antigen epitope. In the model of the B7 tetrasaccharide, the α GalNAc substitutes C2-OH of α -Gal residue, blocking the access to the α Gal residue while leaving the Fuc and β Gal

residues relatively free to react with antibody. The model of the tetrasaccharide from strain H2 shows that the β -GalNAc blocks access to the Fuc residue while leaving free access to both the α -Gal and β -Gal residues. Previous studies [108] in mapping the detailed combining sites of anti-B monoclonal antibodies with a panel of synthetic blood group B analogs revealed that C2-OH on α -Gal and C4-OH on β -Gal residues were key elements for antibody recognition, whereas the antibody could either partially or completely tolerate the deoxy form of other hydroxyl groups. More recent characterization of recognition of blood group B trisaccharide derivatives by specific lectin also reached the similar conclusion [108]. The result of LPS ELISA study in this work is consistent with the above findings. It is noted that a number of bacterial species possess blood group reactivity. Structural and immunological studies demonstrated that blood group activity was related to the cell surface carbohydrates that contain epitopes identical or similar to natural blood group antigens. What is remarkable is that any of those cell surface antigens can induce blood group antibodies when the terminal epitope of the blood group antigen is for the most part buried as an internal sequence of an O-antigen [111].

2.3 Sequencing of Biosynthetic Gene Cluster

The differences between various forms of O-polysaccharides are due mostly to genetic variation in the synthetic gene cluster. Nearly all the genes specific for O-polysaccharide biosynthesis are clustered in a locus. That locus is between *galF* and *gnd* genes in *E. coli*. A 39-bp JUMPStart sequence is present upstream of the gene cluster. It

has been shown that expression of all genes in the gene cluster is enhanced by RfaH acting on JUMPStart [112, 113].

The gene cluster includes nucleotide sugar biosynthetic genes, glycosyltransferase genes and processing genes *wzx*, *wzy* and *wzz*. Some O-PSs include O acetyl groups or other modification groups, and the corresponding genes may be in the same cluster. However, genes for the early steps in pathways that are also involved in housekeeping functions are generally not duplicated in the cluster. In general, the genes in the cluster have a relatively low G+C content, mostly ranging from 30% to 45% for individual gene. This is in contrast to the highly conserved *galF* and *gnd* genes adjacent to the clusters, which have G+C contents typical of most genes of the host chromosome.

LPS biosynthesis genes were originally given *rf*** names, but this system cannot copy with the large number of genes recently identified. A new system called Bacterial Polysaccharide Gene Nomenclature (BPGN) system was then set up by researchers in the area (www.microbio.usyd.edu.au/BPGD/default.htm). This system allows each functionally distinctive gene to have a unique name, genes in different clusters but with the same function to have the same name. All the names generally start with letter “w”.

2.3.1 Experimental Methods

Construction of shotgun bank

Chromosomal DNAs were prepared as previously described. The gene cluster was amplified by PCR (MasterAmpTM extra-long PCR kit ; Epicentre, Madison) with primers based on *galF* and *hisI* as previously described. A total of 10 PCR Products were pooled to minimize the effect of PCR errors. After gel purification, the PCR products were

sheared by nebulizers (Invitrogen) and cloned into pCR4Blunt-TOPO vector to generate a bank (280 colonies) according to the instruction manual of the TOPO shotgun cloning kit (Invitrogen).

Sequencing and analysis

Sequencing was carried out using an ABI 3730 automated DNA sequencer. Sequence data were assembled using Phred/Phrap package from the Genome Center of the University of Washington. The program Artemis was used for gene annotation. BLAST and PSI-BLAST were used for searching databases including GenBank and Pfam protein motif databases. Sequence alignment and comparison were done using the program Clustal W.

2.3.2 Results and discussion

A sequence of 18048 bases was obtained from *galF* to *hisI*, which contained fifteen open reading frames excluding *galF* and *hisI*, all have the same transcriptional direction from *galF* to *hisI* (Figure 2.4). The nucleotide and amino acid sequences were used to search the available databases for indication of possible functions (Table 2.3).

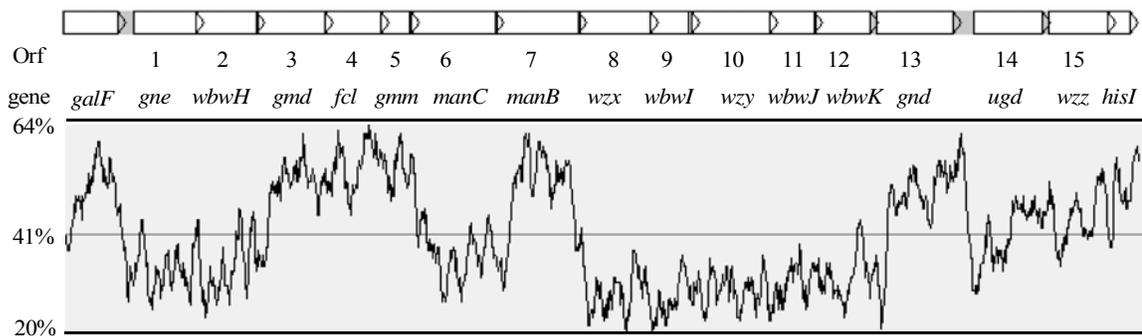


Figure 2.4. O-antigen gene cluster of *E. coli* O86:B7.

Orf	Gene	No. of AA	Similar protein (Genbank accession no.)	Putative function
1	<i>gne</i>	339	Gne of <i>Yersinia enterocolitica</i> O:8 (AAC60777)	UDP-N-acetylglucosamine-4-epimerase
2	<i>wbnH</i>	338	WbcQ of <i>Yersinia enterocolitica</i> (CAA87705)	Glycosyltransferase
3	<i>gmd</i>	373	GDP-mannose 4, 6-dehydratase of <i>E. coli</i> O128 (AA037691)	GDP-D-mannose dehydratase
4	<i>fcl</i>	321	GDP-fucose synthetase of <i>E. coli</i> O128 (AA037692)	GDP-fucose synthetase
5	<i>gmm</i>	167	GDP-mannose mannosyl hydrolase of <i>E. coli</i> O128 (AA037693)	GDP-mannose mannosyl hydrolase
6	<i>manC</i>	482	GDP-mannose pyrophosphorylase of <i>E. coli</i> O128 (AA037694)	GDP-mannose pyrophosphorylase
7	<i>manB</i>	462	Phosphomannomutase of <i>E. coli</i> O128 (AA037695)	Phosphomannomutase
8	<i>wzx</i>	400	O-antigen flippase of <i>Salmonella enterica</i> (AAV34509)	O-antigen flippase
9	<i>wbnI</i>	234	Glycosyltransferase of <i>Homo sapiens</i> (AAQ88542)	Glycosyltransferase
10	<i>wzy</i>	466	O-antigen polymerase of <i>E. coli</i> O111 (AAD46730)	O-antigen polymerase
11	<i>wbnJ</i>	247	WbgO of <i>E. coli</i> O55 (AAL67559)	Glycosyltransferase
12	<i>wbnK</i>	302	Fucosyltransferase of <i>E. coli</i> O128 (AA037698)	Glycosyltransferase
13	<i>gnd</i>	432	Gnd of <i>E. coli</i> K12 (AAC75090)	Gluconate 6-P dehydrogenase
14	<i>ugd</i>	388	UDP-glucose 6-dehydrogenase of <i>E. coli</i> O128 (AA037704)	UDP-glucose dehydrogenase
15	<i>wzz</i>	337	Rol of <i>Shigella flexneri</i> (CAA50783)	O-antigen chain length determinant

Table 2.3. Summary of *E. coli* O86 O-antigen biosynthesis genes.

(i) Nucleotide sugar biosynthesis genes. O86:B7 repeating unit contains five sugar residues: one fucose (Fuc), and two each of N-acetylgalactosamine (GalNAc) and galactose (Gal). As genes for the synthesis of the UDP-Gal are located elsewhere [114], only genes for UDP-GalNAc and GDP-L-Fuc were expected in the gene cluster.

The deduced protein sequence of *orf1* showed 57 and 22% identity to identified Gne proteins of *Yersinia enterocolitica* O:8 and *Pseudomonas aeruginosa* O6, respectively [115, 116]. Gne catalyzes the conversion of UDP-GlcNAc to UDP-GalNAc, so *orf1* was identified as *gne* for the biosynthesis of UDP-GalNAc in *E. coli* O86:B7.

Orf3 to Orf7 shared 95, 99, 94, 85, and 92% identity to the putative GDP-L-Fuc pathway enzymes Gmd, Fcl, Gmm, ManC, and ManB, respectively, of *E. coli* O128 [117]. GDP-L-Fuc is synthesized from GDP-mannose by two enzymes: GDP-mannose 4, 6-dehydratase (Gmd) and GDP-L-Fuc synthetase (Fcl). ManA (phosphomannose isomerase), ManB (phosphomannomutase) and ManC (GDP-mannose pyrophosphorylase) are three additional enzymes needed in the synthesis of GDP-mannose from fructose 6-phosphate. ManA is also a member of mannose metabolism pathway in *E. coli*, and the gene is located elsewhere on the chromosome [118]. Gmm (GDP-mannose mannosyl hydrolase) has been suggested to regulate the cell wall biosynthesis by influencing the concentration of GDP-mannose in the cell [119].

(ii) Sugar transferase genes. Five glycosyltransferases are required to transfer all the five sugar residues in the repeating unit. The first glycosyltransferase (WecA) to transfer UDP-GalNAc is located outside the gene cluster; four glycosyltransferase genes were expected in the gene cluster.

Glycosyltransferases can be classified into 86 distinct sequence-based protein families as described by Campbell et al. (<http://afmb.cnrs-mrs.fr/CAZY/>). WbnH belongs to glycosyltransferase family 1 (PF00534, $E=2.0 \times e^{-16}$), and also shows 68 and 62% similarity to putative glycosyltransferases of *Yersinia enterocolitica* and *Edwardsiella ictaluri*, respectively. WbnI belongs to glycosyltransferase family 6 (PF03414, $E=1.0 \times e^{-11}$), it also shares high similarity with WbgM from *E. coli* O55, which has been proposed to encode α 1,3-galactosyltransferase [120]. Therefore, we tentatively assigned WbnI as α -1,3-galactosyltransferase that makes the Gal α 1,3-Gal linkage. WbnJ belongs to glycosyltransferase family 2 (PF00535, $E=5.0 \times e^{-16}$). WbnK belongs to glycosyltransferase family 11 (PF01531, $E=1.0 \times e^{-15}$), which only consists of several fucosyltransferases from different organisms (<http://afmb.cnrs-mrs.fr/CAZY/>). It also shares 32% identity with the identified fucosyltransferase of *E. coli* O128. It is likely that *wbnK* encodes the fucosyltransferase that links a fucose to the unit.

(iii) processing genes. *Wzx* and *Wzy* are the only two membrane proteins with more than eight transmembrane segments encoded in the gene cluster. The protein encoded by *orf8* has 12 transmembrane segments. When searching in Pfam database, Orf8 belongs to polysaccharide biosynthesis protein family ($E=0.0037$), members of which are involved in the production of polysaccharide including RfbX of the O-PS biosynthesis operon. Therefore, *orf8* was named *wzx*.

Orf10 has 10 transmembrane segments with a large periplasmic loop made up of 75 amino acids, which is a characteristic of polymerase *Wzy*. Although it showed low similarity with other putative *Wzy* proteins, it is one of the only two proteins with multiple transmembrane segments. We tentatively assigned it as *wzy*. The function of *wzy*

gene was further confirmed by comparison of LPS phenotypes between wild type and *wzy* mutant strain in which *wzy* gene was replaced by a CAT gene (discussed later). Therefore, we believed that *orf10* was *wzy* gene and named accordingly.

Orf15 belongs to WzzB family (COG 3765, $E=2.0 \times e^{-73}$), which consists of chain length determinant proteins involved in the LPS biosynthesis. It also shares 98% amino acid identity with identified Wzz protein from *Shigella Flexner* [50]. Analyzed with TMpred program [121], Orf15 exhibited a topology of two transmembrane segments located at the amino-terminal and carboxy-terminal and a large hydrophilic loop in the periplasm, which is the characteristic of Wzz proteins. Accordingly, *orf15* was proposed to encode Wzz protein and named *wzz*.

2.4 Biochemical Identification of Glycosyltransferase Genes

Four putative glycosyltransferase genes were found in the gene cluster whose translation products are responsible for the assembly of repeating unit oligosaccharides. These four genes were cloned and expressed individually. Biochemical assays were carried out to identify their specific functions.

2.4.1 Experimental Methods

Cloning, expression and purification

The *wbnH*, *wbnI*, *wbnJ* and *wbnK* genes were amplified by PCR from the *E. coli* O86 chromosome. The primers with restriction sites underlined for amplification each gene were as follows:

(*wbnH*) 5' GAGATATACCATATGAAAAATGTTGGTTTTATTG (*NdeI*);

Reverse: 5' CCGCCTCGAGTCAACCTAAAATAATGCTTTTATATG (*XhoI*).

(*wbnI*) 5' CCGCCATATGGTTATTAATATATTTT (*NdeI*)

Reverse: 5' GCACACTCGAGTTACTTCTTGATATTACCA (*XhoI*);

(*wbnJ*) 5' GATATACATATGTCATTGAGAATATTAGAT (*NdeI*)

Reverse: 5' CCGCCTCGAGTTATTTTATTAGTGCTTTAC (*XhoI*);

(*wbnK*) 5' CGCGGATCCATGGAAGTTAAAATTATTGGGGGGCT (*BamHI*)

Reverse: 5' CGTTCTCGAGTCATAATTTTACCCACGATTCG (*XhoI*)

The DNA fragments obtained were digested with corresponding restriction enzymes and inserted into pGEX-4T-1 or pET15b vectors linearized by the same restriction enzymes to form pGEX-wbnK, pET15b-wbnH, pET15b-wbnI and pGEX-wbnJ plasmids. The recombinant plasmids were confirmed by restriction mapping and sequencing. The correct constructs were subsequently transformed into *E. coli* BL21 (DE3) for protein expression. *E. coli* BL21 (DE3) harboring the recombinant plasmid was grown in 1L Luria-Bertani (LB) medium at 37 °C. Isopropyl-1-thio-β-D-galactospyranoside (IPTG) was added to a final concentration of 0.8 mM for pGEX-wbnK and 0.4mM for pET15b-wbnI and pGEX-wbnJ. Expression was allowed to proceed for 8 h at 30 °C. Cells were harvested and stored at –80 °C until needed. In protein purification, we used two different affinity columns. In the case of WbnK and WbnJ, the cell pellet was suspended in GST bind buffer (4.3 mM Na₂HPO₄, 1.47 mM KH₂PO₄, 137 mM NaCl, 2.7 mM KCl, pH 7.3) and disrupted by sonication on ice. After centrifugation, the lysate was loaded onto a GST-Bind column (1 mL, Invitrogen). After washing with the same buffer for four times, the protein was eluted with the elution buffer (50 mM Tris-HCl, 10 mM reduced Glutathione, pH 8.0). In the case of WbnI, The

cell pellet was resuspended in chilled lysis buffer (50 mM Tris-HCl, pH 7.0, 0.1 M NaCl, 0.5% (w/v) Triton X-100, 10% glycerin (w/v), and 5 mM 2-mercaptoethanol) and disrupted by brief sonication (Branson Sonifier 450, VWR Scientific) on ice. The lysate was cleared by centrifugation (12,000g, 20 min, 4 °C) and loaded at a low rate of 3 ml/min onto a Ni²⁺-NTA (nickel-nitrilotriacetic acid) agarose affinity column (5 mL, Invitrogen) equilibrated with 50 mM Tris-HCl, pH 7.5, 500 mM NaCl, and 10 mM imidazole. The column was washed with 5 column volumes of 50 mM imidazole in the same buffer, and the protein was eluted with 500 mM imidazole. Protein eluted from the column was analyzed by SDS-PAGE.

Glycosyltransferase activity assay

Standard assay for glycosyltransferase activity is essentially the same with the variation of sugar donor and acceptor in terms of the individual enzymes. Protein concentration is determined by Bradford method. For all the enzymatic assays, 1 unit of enzyme activity was defined as the amount of enzyme required to transform 1 μmol of sugar donor to acceptor per minute at 37°C. Enzyme assays were performed at 37°C for 2 hr in a final volume of 100 μL containing 20 mM Tris-HCl, pH 7.0, 10 mM MnCl₂, 1 mM ATP, 0.3 mM radioactively labeled sugar donor, 10 mM acceptor and variable amount of enzymes. Acceptor was omitted as the blank control. The β1,3-galactosyltransferase (*WbnJ*) was assayed using 0.3 mM radioactively labeled UDP-D-[6-H³]Galactose (10000cpm) as sugar donor, and 10mM monosaccharide (GalNAcα-OMe) as sugar acceptor. The α1,2-fucosyltransferase (*WbnK*) was assayed using 0.3 mM radioactively labeled GDP-L-[¹⁴C]fucose (7000 cpm) as sugar donor, and 10mM

synthesized disaccharide (Gal β 1,3-GalNAc α -OMe) as sugar acceptor. The α 1,3-galactosyltransferase (*WbnI*) was assayed using 0.3 mM radioactively labeled UDP-D-[6- H^3]Galactose(10000cpm) as sugar donor, and 10mM trisaccharide (Fuc α 1,2-Gal β 1,3-GalNAc α -OMe) as sugar acceptor. The reaction was terminated by adding 100ul ice cold 0.1M EDTA. Dowex 1 \times 8-200 chloride anion exchange resin was then added in a water suspension (0.8 mL, v/v = 1/1). After centrifugation, supernatant (0.45 mL) was collected in a 20 mL plastic vial and ScintiVerse BD (10 mL) was added. The vial was vortexed thoroughly before the radioactivity of the mixture was counted in a liquid scintillation counter (Beckmann LS-3801 counter). For *WbnH*, enzymatic assay was carried out using a synthetic acceptor containing diphospho-lipid structure (discussed below) and analyzed by LC-MS, instead of the radioactive based method.

2.4.2 Results and discussion

WbnH: The full open reading frame of *wbnH* gene (1017 bp) was amplified from *E. coli* O86 genomic DNA, and subsequently cloned into pET-15b vector (Novagen). The recombinant plasmid was transformed into *E. coli* BL21 (DE3) strain for induced expression with 0.15 mM IPTG at 18°C. Expression at low temperature significantly increased the solubility portion of the overall protein. *WbnI* protein was produced with His₆ tag at the N-terminus and purified to 90% homogeneity in one step using Ni-column.

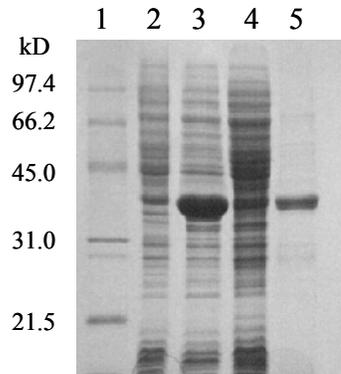


Figure 2.5. Expression and purification of WbnH. Lane 1, low range marker; lane 2, BL21(DE3) with empty plasmid pET15b; lane 3, WbnH whole cell induction; lane 4, WbnH lysate after induction; lane 5, purified WbnH after Ni-affinity column.

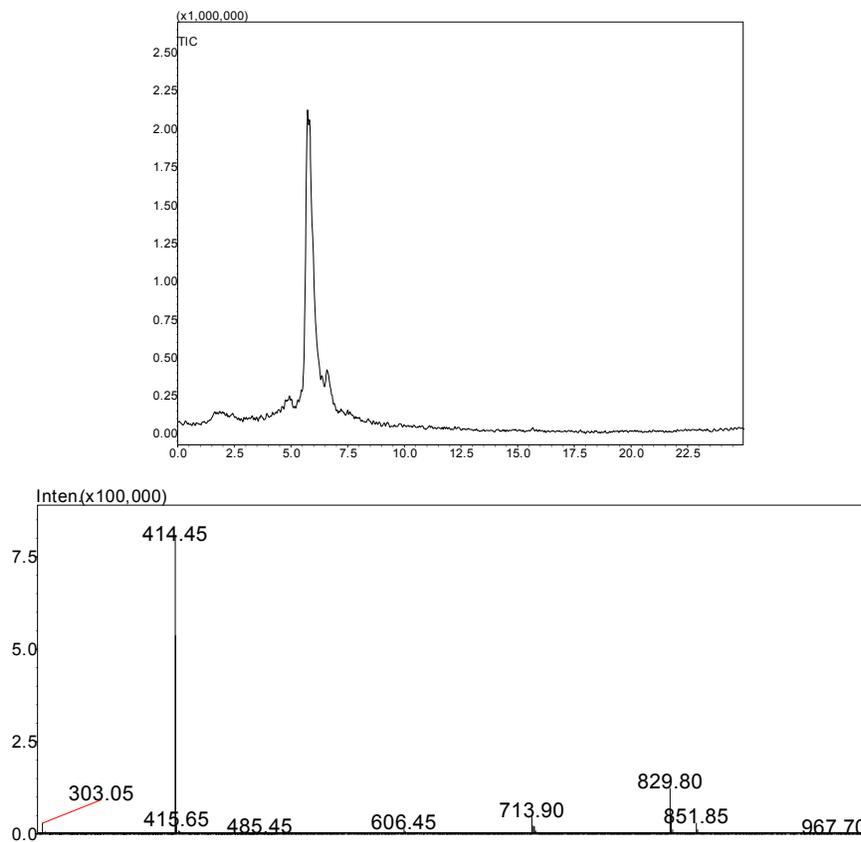
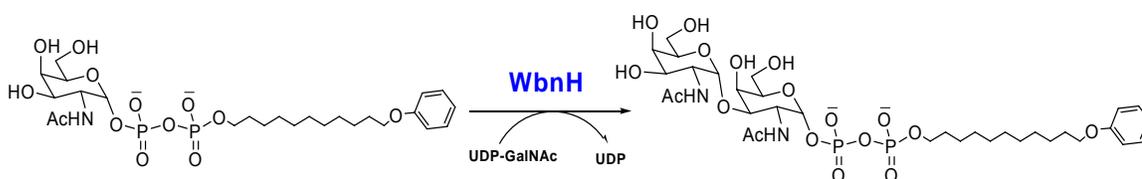


Figure 2.6. Identification of disaccharide product catalyzed by WbnH using LC-MS.

WbnH could be obtained approximately 3 mg per liter of bacterial culture. The recombinant protein has an isoelectric point of 6.4 and an apparent molecular weight of 38 kD estimated by SDS-PAGE (Figure 2.5), similar to the theoretical value (38.148 kD) calculated from its predicted amino acid sequence. Amino acid sequence analysis showed that WbnH has no transmembrane segments, which is consistent with our result that we could be able to purify WbnH without difficulty even though detergent is omitted.



Scheme 1.1. *In vitro* enzymatic reaction of WbnH.

WbnH was proposed to transfer GalNAc from UDP-GalNAc to C3 hydroxyl of GalNAc-PP-Und. The substrate GalNAc-PP-Und is difficult to isolate from natural bacterial sources due to the low abundance. Furthermore, the highly hydrophobic feature of the 55-carbon Und chain makes it hard to handle and is anticipated to cause problems in enzymatic reaction. Since there is no compelling evidence to show that 55-carbon lipid chain is absolutely necessary for WbnH recognition, a GalNAc-PP-lipid acceptor with only 11-carbon chain was prepared using chemical synthesis (see Section 2.5) and used to characterize WbnH activity (Scheme 1.1). The synthetic acceptor was incubated with WbnH enzyme at room temperature for 2 hr. The reaction was then quenched using methanol and analyzed by LC-MS. The mass spectrum (Figure 2.6) showed a prominent

peak with m/z ratio at 414.45 (M^{2-}), and 829.80 (M^+H), consistent with the formation of the GalNAc-GalNAc-PP-O(CH₂)₁₁-OPh product. To further confirm the glycosidic linkage between two GalNAc, the reaction volume was scaled up to 2 mL, from which 7 mg product was isolated using reverse phase chromatography. The 500 MHz ¹H NMR spectrum was collected (Appendix). The ¹Hs of the product were identified with the aid of the ¹H-¹H COSY-DQF. The configuration of the disaccharide was determined as α -linkage by the coupling constant between the H1' and H2' (5.01 ppm, d, $J = 3.5$ Hz). The NOESY spectrum showed the coupling between H1' and the H3, which finally confirmed the α 1-3 linkage between two GalNAc.

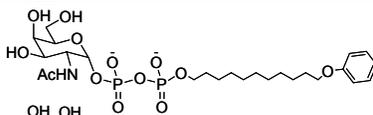
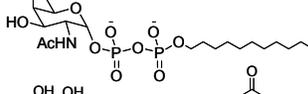
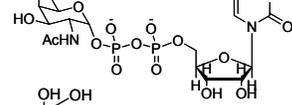
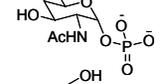
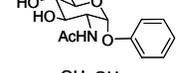
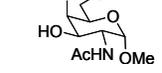
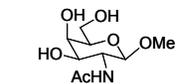
Acceptors	Relative activity %
 (GalNAc- α -PP-C ₁₁ -OPh)	100
 (GalNAc- α -PP-C ₁₁)	95
 (UDP-GalNAc)	2
 (GalNAc- α -P)	N/A
 (GlcNAc- α -OPh)	N/A
 (GalNAc- α -OMe)	N/A
 (GalNAc- β -OMe)	N/A

Table 2.4. Substrate specificity of WbnH.

Purified enzyme was used to investigate the specificity of WbnH towards various sugar acceptors (Table 2.4). It clearly indicates that pyrophosphate and lipid portions are substrate requirements for WbnH activity, since the acceptors without pyrophosphate and lipid substitution do not show any detectable activity. GalNAc with only one phosphate substitution is not a substrate for WbnH either. Comparison of two GalNAc-PP substrates (GalNAc-PP-C₁₁ and GalNAc-PP-C₁₁-OPh) showed that they displayed nearly the same activity, indicating that modification on lipid chain has trivial effect on the activity. It is interesting to note that the sugar donor UDP-GalNAc of WbnH is structurally analogous to its acceptor GalNAc-PP-lipid. The only difference is that lipid portion is substituted by uracil. Thus comparison of the activity of these two substrates would show if WbnH needs hydrophobic lipid portion as part of the substrate requirement. The result showed that UDP-GalNAc only displayed 2% activity relative to GalNAc-PP-C₁₁-OPh, a 50 fold lower in activity. This observation suggests that hydrophobic lipid portion is essential for WbnH recognition. Although the natural substrate contains 55-carbon, we have demonstrated that WbnH is able to use the acceptor with only 11-carbon.

Many glycosyltransferases are found to contain a short conserved amino acid sequence called DXD motif, and exhibit an absolute requirement of divalent metal cation for activity. In order to investigate the metal ion requirements of WbnH, various divalent metal cations as well as EDTA are incubated in a series of reaction mixture. The results showed that WbnH activity does not require metal cation, since it exerts nearly the same level of activity in the absence of cations or even with the presence of EDTA (Figure 2.7). Sequence analysis of WbnH showed that it is absence of the DXD motif, which is consistent with the above observation. Thus WbnH belongs to glycosyltransferase

superfamily B, in which some conserved basic residues play the similar roles as DXD motif does in binding of sugar nucleotide donor.

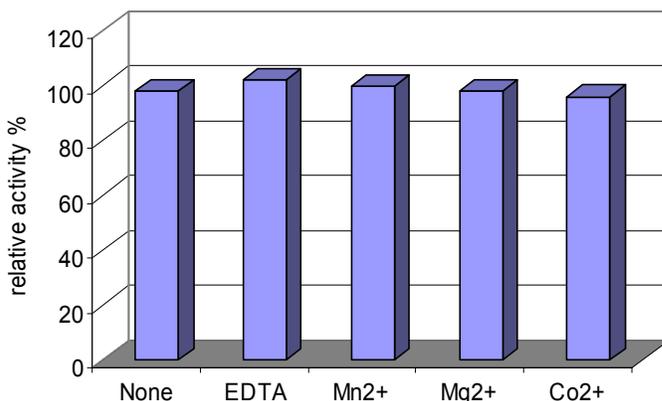


Figure 2.7. Determination of divalent metal ion requirement for WbnH activity. The concentration of EDTA and metal ions are all at 10 mM.

WbnJ: WbnJ was expressed as a fusion protein with an N-terminal GST tag (Figure 2.8). The fusion of GST to WbnJ significantly increased expression of the recombinant protein compared to the one with His₆ tag, as well as the solubility. Radioactivity assay showed that recombinant WbnJ could transfer a galactose residue from UDP-Gal to GalNAc- α -OMe with high efficiency, consistent with its predicted function as a galactosyl-transferase. Acceptor substrate specificity study showed that WbnJ has a preference for GalNAc as the immediate acceptor, and that α -linkage at the anomeric carbon of GalNAc also affects the enzyme recognition. An in vitro enzymatic reaction conducted using GalNAc- α -OMe as acceptor produced 7 mg of disaccharide product (Scheme 2.2). Subsequent NMR and MS analysis confirmed the Gal β 1,3-

GalNAc- α -OMe structure, the same linkage as proposed from the sequence comparison analysis. WbnJ requires divalent cations such as Mn^{2+} and Mg^{2+} as cofactors for full enzymatic activity, suggesting that in contrast to WbnH, WbnJ belongs to superfamily A, where DXD motif plays an essential role in coordinating divalent cations and nucleotide sugar donor in correct orientation within the active site.

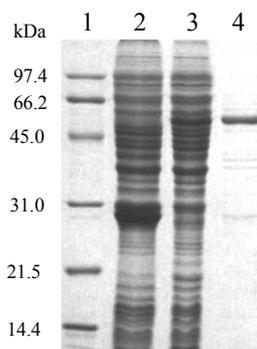
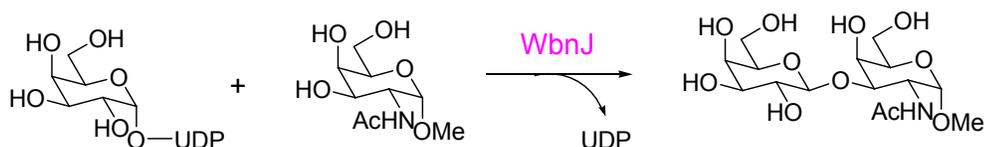


Figure 2.8. Expression and purification of GST-WbnJ.



Scheme 2.2. *In vitro* enzymatic reaction of WbnJ.

WbnK: The C-terminus of WbnK has a conserved domain from glycosyltransferase family 11, which mostly consists of α 1,2-fucosyltransferases from different organisms. WbnK was expressed in *E. coli* as a fusion protein with an N-terminal GST tag (Figure 2.9). Based on the repeating unit structure, a panel of acceptors having the β -

Gal residue was selected to detect the fucosyl-transferase activity of purified GST-WbnK. Radioactive assays showed that WbnK was highly active towards Gal β 1,3-GalNAc α -R structure, but not with any other acceptors tested. The change of regiochemistry of the linkage from 1 \rightarrow 3 to 1 \rightarrow 4, or the stereochemistry from β to α makes very poor substrates for WbnK. Thus, the terminal β 1,3-linked galactose residue is critical in WbnK recognition. Trisaccharide was synthesized from GDP-Fuc and disaccharide Gal β 1,3-GalNAc- α -OMe (Scheme 2.3), and the NMR analysis (Appendix) confirmed the presence of α 1,2-fucose.

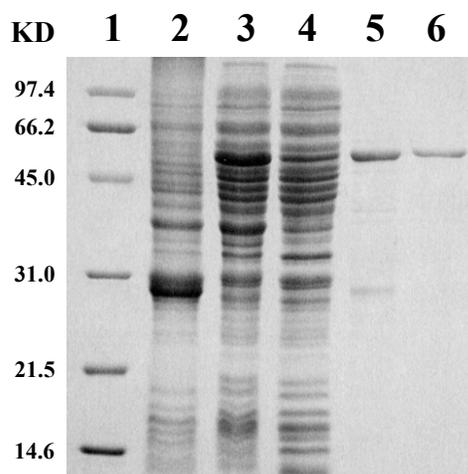
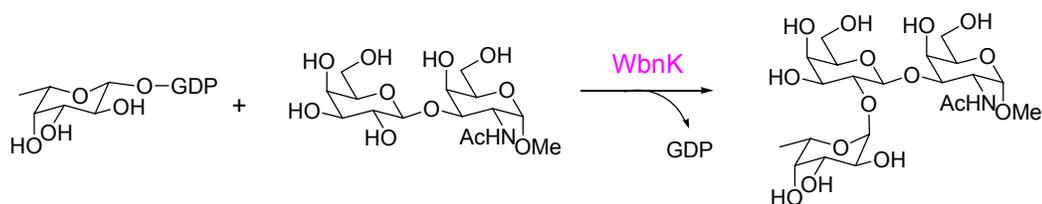


Figure 2.9. Expression and purification of GST-WbnK.



Scheme 2.3. In vitro enzymatic reaction of WbnK.

WbnI: WbnI belongs to family 6 with 45% sequence similarity to human blood group A/B transferase. Thus WbnI is most likely an α 1,3-galactosyltransferase that makes the Gal α 1,3-Gal linkage to form the blood group B structure in the repeating unit. WbnI was expressed as a His₆-tagged fusion protein (Figure 2.10). The acceptor screening showed strict specificity of WbnI (Table 2.5). The trisaccharide Fuc α 1,2-Gal β 1,3-GalNAc α -OMe is the preferred substrate, with nearly seven times more active than the disaccharide Gal β 1,3-GalNAc α -OMe. Thus WbnI prefers terminal fucosylated sugar structures to non-fucosylated ones. Interestingly, since WbnI requires α 1,2-fucose as a sugar side chain in the substrate, the biosynthesis of the blood group B antigen in *E. coli* O86 is essentially the same as that in human. This finding indicates that WbnI is a bacterial homolog of blood group B transferase. Further studies on WbnI can reveal more insights into its substrate specificity and structural relationships with human GTB. The in vitro synthesis of tetrasaccharide using WbnI was also carried out (Scheme 2.4) and the product structure Gal α 1,3(Fuc α 1,2)-Gal β 1,3-GalNAc α -OMe was confirmed by NMR (Appendix).

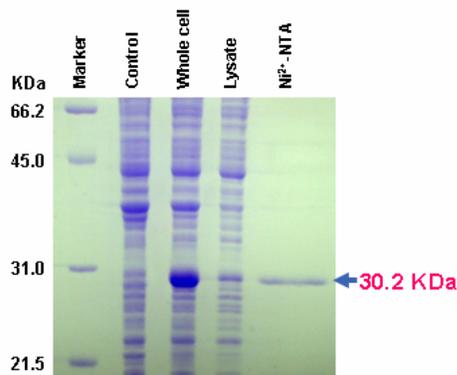
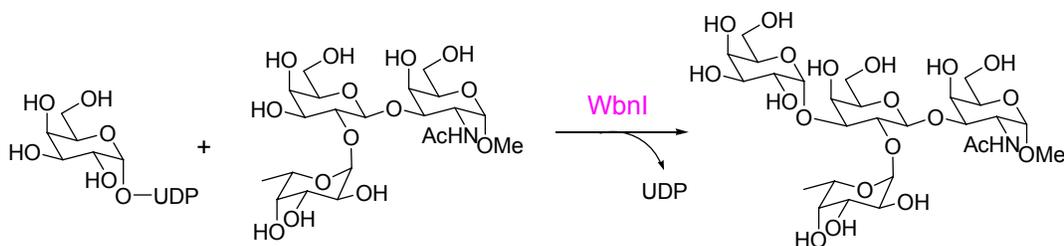


Figure 2.10. Expression and purification of WbnI.

WbnI		WbnK	
Acceptor	R/A*	Acceptor	R/A*
Fuc α 1,2Gal β 1,3GalNAc α -OMe	100%	Gal β 1,3GalNAc α -OMe	100%
Gal β 1,3GalNAc α -OMe	12%	Gal α 1,3Gal β 1,4Glc	N/A
Fuc α 1,2Gal β -OMe	30%	Gal α 1,4Gal β 1,4Glc	N/A
Gal β 1,4Glc β -Oph	N/A	Gal β 1,4Glc	N/A
Gal α 1,3Gal β 1,4Glc	N/A	Gal β -OMe	15%
Gal α 1,4Gal β 1,4Glc	N/A	GalNAc α -OMe	N/A

Table 2.5. Substrate specificity of WbnI and WbnK.



Scheme 2.4. In vitro enzymatic reaction of WbnI.

Collectively, the specific function of each transferase encoded in the gene cluster was biochemically identified. The linkage specificity they catalyzed is in total agreement with the repeating unit structure.

2.5 Reconstitution of repeating unit biosynthesis in vitro

One of the hurdles that hampered the biochemical investigation of polysaccharide biosynthesis in vitro is the difficulty of obtaining chemically defined repeating unit substrates. It is widely accepted that repeating unit is assembled onto a phospho-undecaprenyl (P-Und) carrier by sequential catalysis of several dedicated glycosyltransferases [122, 123]. This P-Und is also considered as a universal lipid carrier for the biosynthesis of various types of bacterial polysaccharides, including O-PS, CPS and EPS. In nature, one of the common precursors, GalNAc/GlcNAc-PP-Und is synthesized from UDP-GalNAc/GlcNAc and P-Und by an inner membrane protein WecA [33]. However, previous attempts to over-express and isolate WecA in homogenous forms have been unsuccessful [33]. It would therefore be very challenging to obtain GalNAc-PP-Und enzymatically even on a mg scale. Accordingly, we turned to chemical synthesis as the methodology used to obtain sugar-PP-Und type compounds is well established. With the synthetic GalNAc-PP-Und precursors, we can proceed to assemble other sugar residues using the four glycosyltransferases previously characterized.

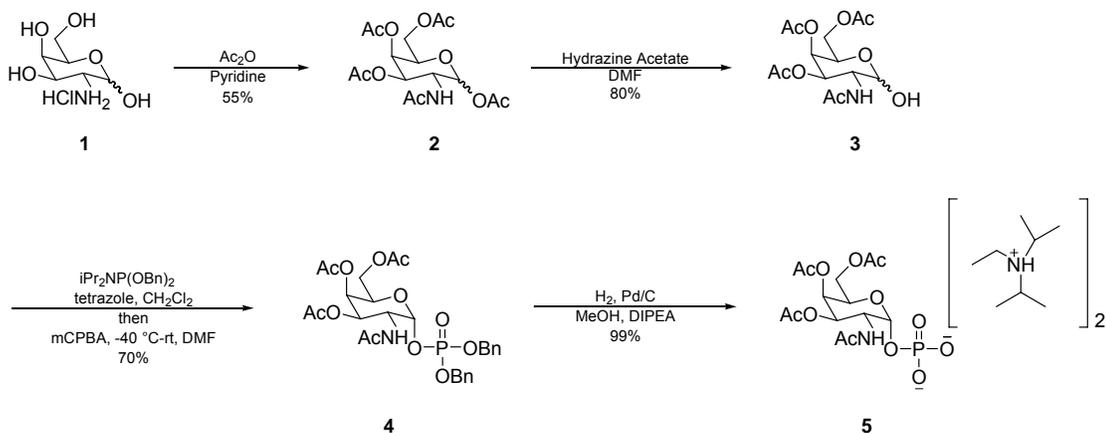
2.5.1 Experimental Methods

2.5.1.1 Chemical Synthesis of GalNAc-PP-Und

The sugar component of the desired target was prepared as outlined in Scheme 2.5. Peracetylation of D-Galactosamine hydrochloride followed by anomeric deprotection with hydrazine acetate gave intermediate **3**. The benzyl-protected phosphate was then accessed through treatment of **3** with bis(benzyloxy)-(diisopropylamino)-

phosphine in the presence of tetrazole followed by oxidation with mCPBA.

Hydrogenolysis of **4** completed the synthesis of the sugar phosphate **5**.

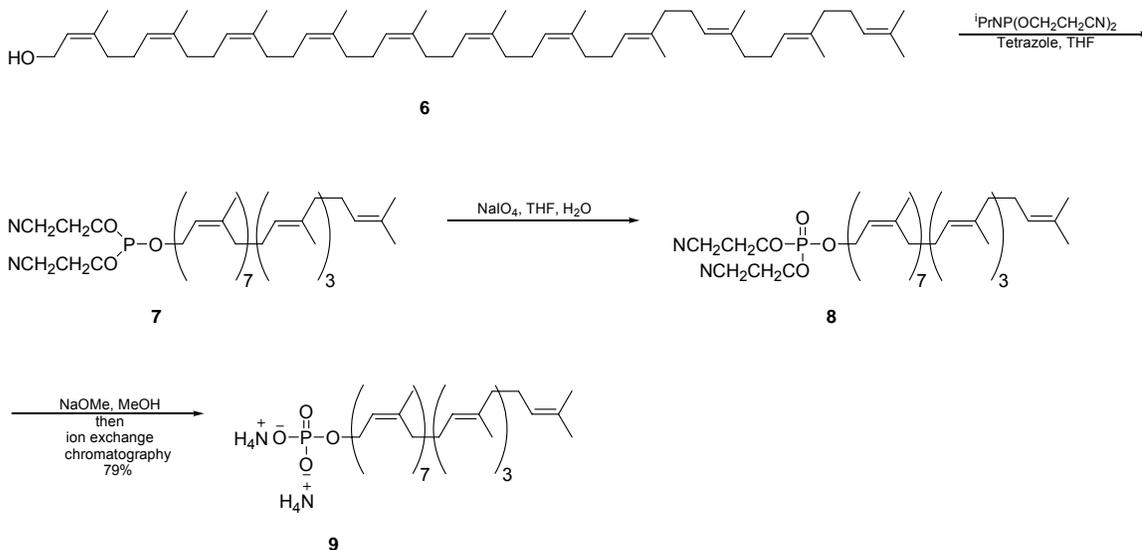


Scheme 2.5. Synthesis of the sugar phosphate, peracetylated GalNAc-Phosphate.

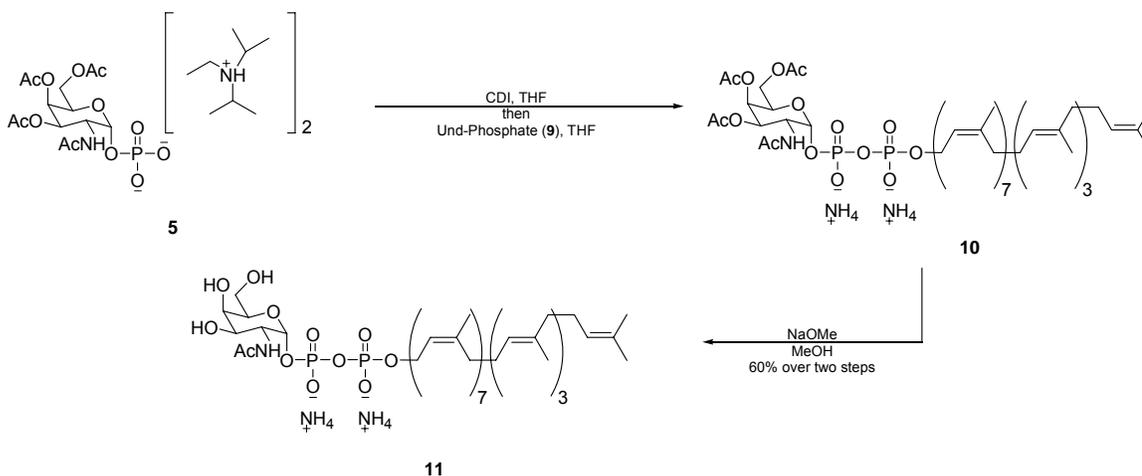
Synthesis of the lipid-phosphate portion of the molecule was subsequently initiated by conversion of commercially available undecaprenol (Und) to the corresponding phosphate **7** via treatment with tetrazole and bis(2-cyanoethoxy)-(diisopropylamino)-phosphine (Scheme 2.6). Subsequent oxidation with sodium periodate and pyridine afforded the protected phosphate **8** which was then deprotected in a 1% solution of sodium methoxide in methanol. Purification via ion exchange chromatography yielded Und phosphate as the ammonium salt.

With each portion of GalNAc-PP-Und in hand, coupling of the fragments was initiated by first treating the sugar phosphate with carbonyldiimidazole to form the phosphoimidazole intermediate (Scheme 2.7). Introduction of the lipid phosphate afforded peracetylated GalNAc-PP-Undecaprenyl after stirring for three days. Sodium methoxide mediated deprotection then gave the desired product, GalNAc-PP-Und.

Formation of the lipid phosphate, peracetylated GalNAc-PP-Und, and GalNAc-PP-Und, was monitored by mass spectrometry. Purity of the compounds following reverse-phase column chromatography was assessed in a similar manner.



Scheme 2.6. Synthesis of the lipid phosphate moiety, Und-P.



Scheme 2.7. Synthesis of GalNAc-PP-Und.

2-acetamido-1,3,4,6-tetra-O-acetyl-2-deoxy-D-galactopyranose (2): A solution of D-galactosamine hydrochloride **1** (1.5 g, 6.94 mmol) in anhydrous pyridine (12 mL) and acetic anhydride (12 mL) was stirred overnight at room temperature. The solution was concentrated *in vacuo*, co-evaporated with toluene, and dried under high vacuum to give a yellow solid. The crude product was purified by flash column chromatography (1:1 pentane: acetone) to give **2** as a white solid (1.2 g, 55%). ¹H NMR (400 MHz, CDCl₃): δ 6.19 (d, 1H, J = 3.6 Hz), 5.40 (dd, 1H, J = 2 Hz, J = 2.8 Hz), 5.34 (bd, 1H, J = 8.8 Hz), 5.20 (dd, 1H, J = 3.2 Hz, J = 11.6 Hz), 4.71 (m, 1H), 4.21 (t, 1H, J = 7.2 Hz), 4.02-4.12 (m, 2H), 2.15 (s, 6H), 2.01 (s, 6H), 1.93 (s, 3H).

2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-D-galactopyranose (3): A solution of **2** (1.44 g, 3.70 mmol) and hydrazine acetate (0.4 g, 4.43 mmol) in anhydrous DMF (20 mL) was stirred for 40 min at room temperature. The mixture was then concentrated *in vacuo* and dried under high vacuum to afford a yellow solid. Flash chromatography (1:1 pentane: acetone) gave the title compound **3** as a white solid (1.02 g, 80 %). ¹H NMR (400 MHz, CDCl₃): δ 6.18 (d, 1H, J = 3.6 Hz), 5.51 (d, 1H, J = 9.2 Hz), 5.37 (d, 1H), 5.28 (dd, 1H, 3.2 Hz), 4.68 (dd, 1H, J = 11.5 Hz), 4.20 (m, 1H), 4.03 (m, 2H), 1.92 – 2.12 (four s, 12H).

2-acetamido-3,4,6-tri-O-acety-1-O-[bis(phenylmethyl)phosphoryl]-2-deoxy-α-D-galactopyranose (4): Compound **3** (0.50 g, 1.16 mmol) and tetrazole (832 mg, 11.6 mmol) were dissolved in dry CH₂Cl₂ (20 mL), after which the solution was cooled to -40 °C. Bis(benzyloxy)(diisopropylamino)phosphine (1.84 g, 5.8 mmol) was then added drop-wise over a 2 min period. The mixture was stirred for 3 h while allowing the temperature to rise to 25 °C. The mixture was cooled to -70 °C and 77% m-CPBA (2.6 g,

11.6 mmol) was added. The solution was stirred for 30 min at 0 °C, then overnight at room temperature. After addition of dichloromethane (100 mL) the solution was washed with saturated Na₂SO₃ (50 mL X 2), water (50 mL X 2), and saturated NaCl (50 mL X 2), dried over Na₂SO₄, and concentrated to yield a yellow oil. The crude product was purified by flash column chromatography (1:1 ethyl acetate-hexane, then 100% ethyl acetate) to give product **4** as a colorless foam (0.6 g, 70 %). ¹H NMR (500 MHz, CDCl₃): δ 7.33-7.37 (m, 10H), 5.71 (dd, 1H, J = 3.5 Hz, J = 6 Hz), 5.49 (d, 1H, J = 9.5 Hz), 5.37 (d, 1H, J = 2 Hz), 5.02-5.13 (m, 5H), 4.60 (m, 1H), 4.24 (t, 1H, J = 7.5 Hz), 4.06 (dd, 1H, J = 7 Hz, J = 11.5 Hz), 3.93 (dd, 1H, J = 7 Hz, J = 11 Hz), 2.13 (s, 3H), 2.03 (s, 3H), 1.98 (s, 3H), 1.75 (s, 3H). ¹³C NMR (62.5 MHz, CDCl₃): δ 170.7, 170.3, 170.2, 170.1, 135.4, 135.3, 135.2 (2), 128.9, 128.8, 128.7, 128.1, 128.0, 97.0, 77.2, 69.9, 68.6, 67.3, 66.7, 61.3, 47.6, 22.8, 20.6 (2), 20.5.

2-acetamido-3,4,6-tri-O-acety-2-deoxy-α-D-galactopyranose-1-phosphate (5): To a solution of **4** (60 mg, 0.1 mmol) in methanol (5 mL) was added 15 mg of 20 % Pd(OH)₂. The reaction vessel was filled with hydrogen and stirred at room temperature. A volume of 0.1 mL of diisopropylethylamine was added after 30 min. The solution was diluted with 15 mL of MeOH and stirred for 30 min. The catalyst was filtered off over celite, and the filtrate was concentrated to give **5** (99%) (R_f = 0.2 (CH₃CN/MeOH = 3:2)).

Undecaprenyl phosphate ammonium salt (9): 1-*H* Tetrazole (18 mg, 0.25 mmol) was added to a stirred solution of undecaprenol (50 mg, 0.065 mmol) and bis(2-cyanoethoxy)(diisopropylamino)phosphine (21 mg, 0.08 mmol) in dry tetrahydrofuran (9 mL) at 0 °C under argon. After 5 min. the solution was warmed to room temperature for 1 h and then re-cooled to 0 °C. Pyridine (0.34 mL, 4.2 mmol) was then added, followed

by a solution of sodium periodate (0.53 g, 2.5 mmol) in water (5 mL). The resulting mixture was stirred for 15 min at 0 °C before being stirred at room temperature for 1 h. The mixture was diluted with ethyl acetate (100 mL) and washed successively with 100 mL each of water, 5% sodium sulfite and brine. The organic phase was dried (Na₂SO₄) and concentrated *in vacuo* to give an oil which was stirred with a 1% solution of sodium methoxide in methanol (5 mL, 1.10 mmol) for 3 days at room temperature. The solution was concentrated *in vacuo* to give an oil which was purified via column chromatography using DEAE cellulose and 33 mmol NH₄HCO₃ in methanol as the eluent following addition to the column in 1:4 methanol:chloroform. Product containing fractions were combined and concentrated *in vacuo* to give undecaprenyl phosphate as the ammonium salt **9** (45 mg, 79 %). ESMS calcd for C₅₅H₈₉O₄PH (M-H⁻) 845, found 845 (Appendix).

Ammonium 2-acetamido-3,4,6-tri-O-acety-2-deoxy- α -D-galactopyranose-1-undecaprenyl diphosphate (10): Monophosphate **5** (23 mg, 0.03 mmol) was dissolved in 3 mL of dry THF and treated with carbonyldiimidazole (22 mg, 0.14 mmol). After 2 h, 42 μ L of dry MeOH was added, and the reaction was stirred for 1 h. The solvent was evaporated to give the phosphoimidazole intermediate. This intermediate was dissolved in 2 mL of dry THF and transferred to undecaprenyl phosphate, after which the reaction was stirred at room temperature for 3 days. The solvent was then evaporated and the residue was loaded onto a C18 reverse phase column (8 mm X 80 mm, particle size 40 μ m, pore size 60 \AA , from Aldrich) and eluted with CH₃CN/10 mM NH₄HCO₃ in water. The fractions containing the desired product were combined and concentrated to give **10** as a colorless solid. ESMS calcd for C₆₉H₁₀₉NO₁₅P₂H (M-H⁻) 627, found 627 (Appendix).

Ammonium 2-acetamido-2-deoxy- α -D-galactopyranose-1-undecaprenyl diphosphate

(11): A solution of **10** in 0.034 M NaOMe in MeOH (3 mL) was stirred for 40 min at room temperature. Solvent was then removed *in vacuo* and the residue was loaded onto a C18 reverse phase column (8 mm X 80 mm, particle size 40 μ m, pore size 60 Å, from Aldrich) and eluted with CH₃CN/20 mM NH₄HCO₃ in water. The fractions containing the desired product were combined and concentrated to give product **11** as a colorless solid (18 mg, 60% over two steps). ESMS calcd for C₆₃H₁₀₃NO₁₂P₂H (M-H⁻) 564, found 564 (Appendix).

2.5.1.2 Enzymatic Sequential Assembly of Repeating Unit Substrates

To a buffer containing 20 mM Tris-HCl (pH=7.0) and 5 mM Mn²⁺, 5 mg of GalNAc-PP-Und (solubilized in DMSO) was added followed by 1% (v/v) Triton X-100, 10 mM UDP-GalNAc. The mixture was vortexed vigorously and sonicated for 10 min in a water bath. The reaction was initiated by the addition of 50 μ g of the enzyme WbnH and was incubated at room temperature for 2 hrs or until the reaction was complete. The reaction mixture was purified using reverse phase chromatography with methanol as the eluent. About 0.01 mg of the dried product was labeled with 2-AB using reductive amination. The eluted product mixture was further separated on a normal phase analytic HPLC column using 50 mM ammonium formate (pH 4.5) (solvent A) and acetonitrile (containing 5 mM ammonium formate, solvent B). A gradient of 48-80% solvent B was used at a flow rate of 0.5 mL/min. The peaks were detected with a fluorescence detector (λ_{ex} =330 nm, λ_{em} =420 nm) and collected and characterized by MALDI-MS.

The rest of the dried product was re-solubilized in enzyme reaction buffer as described above and the reaction was initiated by the addition of the next transferase enzyme and the corresponding NDP-sugar. Similar product purification and characterization steps were repeated until the pentasaccharide-PP-Und product was synthesized.

2.5.2 Results and Discussion

Synthesis of disaccharide GalNAc-GalNAc-PP-Und with WbnH

WbnH is an α -1,3-N-acetylgalactosaminyl transferase. We have previously shown that WbnH transfers GalNAc from UDP-GalNAc to a GalNAc-PP-C₁₁Oph acceptor. When WbnH was incubated with GalNAc-PP-Und and UDP-GalNAc, LC/MS analysis of the reaction mixture revealed a peak at $m/z = 665.5$ in the mass spectrum, which corresponded to GalNAc-GalNAc-PP-Und $[M-2H]^{-2}$. To further confirm the product formation, the Und lipid portion was hydrolyzed, after which the free saccharide was fluorescently labeled with 2-AB by reductive amination for HPLC/MS analysis. A peak with a retention time of 16 min was observed (Figure 2.11). When this peak was collected and subjected to ESI-MS analysis, a $m/z = 567.225$ was identified, corresponding to the sodium ion adduct of labeled disaccharide GalNAc-GalNAc-2AB. Therefore, disaccharide GalNAc-GalNAc-PP-Und was reconstituted *in vitro* with WbnH.

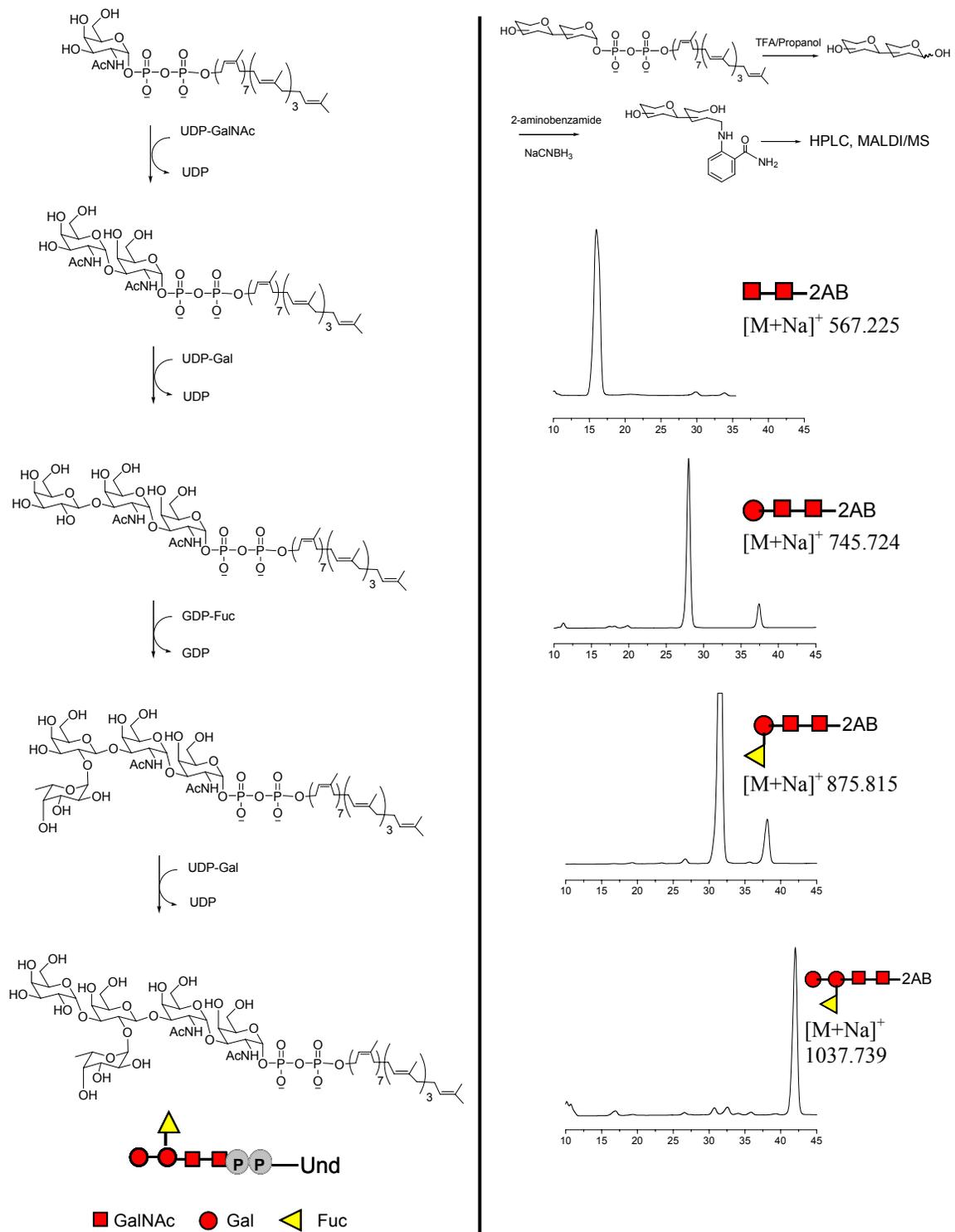


Figure 2.11. In vitro reconstitution of *E. coli* O86 repeating unit oligosaccharide.

Synthesis of trisaccharide Gal-GalNAc-GalNAc-PP-Und with WbnJ

The next enzymatic transformation, the addition of a galactose residue via a β 1,3-linkage, is catalyzed by WbnJ. When the reaction mixture was analyzed by LC-MS, a peak was observed with a m/z of 746.4, corresponding to trisaccharide Gal-GalNAc-GalNAc-PP-Und $[M-2H]^{-2}$. Our previous study showed that WbnJ could readily accept sugar substrates in the absence of the diphospho-lipid moiety, indicating that the diphospho-lipid moiety is not required for WbnJ enzyme recognition. HPLC analysis showed a peak with a retention time of 28 min (Figure 2.11). The MALDI-MS analysis identified a $[M + Na]^+$ $m/z = 745.724$ (Appendix), consistent with the formation of the sodium ion adduct of trisaccharide-2AB. Thus, these experiments demonstrated that Gal-GalNAc-GalNAc-PP-Und was reconstituted with WbnJ.

Synthesis of tetrasaccharide Fuc-Gal-GalNAc-GalNAc-PP-Und with WbnK

WbnK was previously demonstrated to act as a fucosyltransferase, catalyzing the addition of fucose to the Gal-GalNAc-GalNAc trisaccharide substrate. When WbnK and GDP-Fuc were added to the reaction mixture, the LC/MS mass spectrum showed a peak at $m/z = 819.5$, the expected m/z ratio for tetrasaccharide Fuc-Gal-GalNAc-GalNAc-PP-Und $[M-2H]^{-2}$. Consistent with this result, a peak with a retention time of 32 min was observed in HPLC assay (Figure 2.11), and yielded a $[M + Na]^+$ m/z of 875.815 in the MALDI-MS analysis (Appendix). These results verified the fucosyltransferase activity of WbnK and demonstrated the *in vitro* reconstitution of tetrasaccharide Fuc-Gal-GalNAc-GalNAc-PP-Und.

Synthesis of pentasaccharide Gal(Fuc)-Gal-GalNAc-GalNAc-PP-Und with WbnI

The final enzymatic transformation is catalyzed by WbnI through the addition of a galactose residue to the tetrasaccharide to form a complete repeating unit structure. The final pentasaccharide product was observed in the mass spectrum at $m/z = 901.2$ after purification by reverse phase chromatography. The HPLC assay revealed a peak with a retention time of 43 min (Figure 2.11). MALDI-MS of the collected fraction confirmed the formation of pentasaccharide-2AB ($m/z = 1037.739$, $[M + Na]^+$, Appendix).

In summary, the authentic substrate repeating unit pentasaccharide-PP-Und was reconstituted by step-wise enzymatic reactions *in vitro*. This result presented the first biochemical evidence for the sequential assembly model of repeating units in the *wzy*-dependent pathway, and also provided access to homogenous synthetic repeating unit substrates.

2.6 Over-expression and detection of Wzy membrane protein

In the previous *in vivo* genetic study, the *wzy* gene was proposed to encode a sugar polymerase [51]. The mutant strain in which the *wzy* gene had been deleted displayed a semi-rough LPS phenotype, in which only one repeating unit was linked to the Lipid-A-core structure. However, the definitive demonstration of Wzy polymerization activity *in vitro* has been long hampered by the difficulty of expressing Wzy protein [24]. Wzy is a membrane protein with 11 predicted transmembrane segments, which poses a challenge for its over-expression and isolation. The first extensive study for Wzy overexpression so far was carried out by Morona and coworkers [37], where they manipulated the *wzy* gene by mutating the rare codons near the initiation site and

introducing a strong ribosome binding site. Despite these efforts, Wzy could only be detected by [³⁵S]-methionine labeling. We developed a chaperon co-expression system and successfully detected Wzy expression with Western-blotting.

2.6.1 Experimental Methods

The *wzy* gene was amplified by PCR from the *E. coli* O86 chromosome (O-antigen gene cluster accession # AY220982). The primers with restriction sites underlined for amplification of each gene were as follows: Forward: 5' TACGGCCATGGTAATATCAAGAAGTAA (*NcoI*); Reverse: 5' CGCGCCAAGCTTATGACATTTTTTTAT (*HindIII*). The DNA fragments were cloned into the pBAD-myc-His vector. The pBAD-wzy plasmid was co-transformed with the GroEL/GroES vector into *E. coli* C43 (DE3) for protein expression. Expression was induced with 0.1% arabinose, followed by further incubation at 20 °C for 20 hrs. Cells were harvested and washed with 20 mM Tris-HCl (pH=7.0). The membrane fraction was isolated by a sequential ultracentrifugation method, solubilized with 8 M urea and 2% SDS, and then passed through a Ni-affinity column. The bound proteins were eluted with a buffer containing 250 mM imidazole. Different fractions were analyzed by SDS-PAGE and Western-blotting using an anti-myc antibody (Invitrogen).

2.6.2 Results and Discussion

In vivo disruption and trans-complementation of the wzy gene in E. coli O86

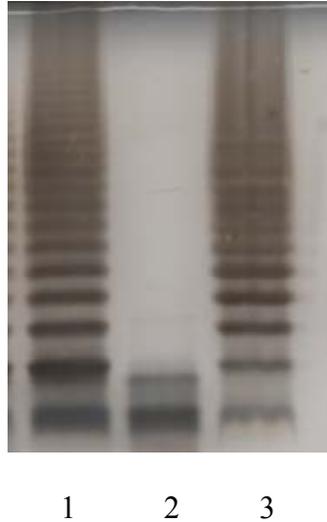


Figure 2.12. In vivo disruption and complementation of *wzy* gene. Lane 1, wild-type LPS from *E. coli* O86:B7; lane 2, LPS from *wzy* deficient strain; lane 3, LPS from *wzy* trans-complemented strain.

The *wzy* gene from *E. coli* O86 was replaced by the chloramphenicol acetyltransferase (CAT) gene using the RED recombination system of phage lambda [39]. The CAT gene was PCR-amplified from plasmid pKK232 using primers binding to the 5' and 3' ends of the gene. Each primer also carries 40 bp gene segments flanking the target genes. The PCR product was electroporated into *E. coli* O86 strain carrying the pKD20 vector. The chloramphenicol resistant transformants were selected after induction of the RED gene according to the protocol described previously. To complement the mutant strain, the pBAD-*wzy* recombinant vector was transformed into the mutant strain, which was subsequently grown in LB media supplemented with 0.1% arabinose. The

LPSs from wild-type, *wzy*-mutant and complemented strains were extracted according to our previous protocol and visualized on SDS-PAGE by silver staining (Figure 2.12). Lane 1, the LPS from wild-type O86 displays typical ladder-like pattern, a smooth LPS phenotype. Lane 2, the LPS from the mutant strain in which the *wzy* gene was deleted shows a semi-rough phenotype, in which only one repeating unit is attached to the Lipid A-core. Lane 3, the LPS from *wzy*-complemented strain shows a smooth phenotype, resembling that of the wild-type strain. This *in vivo* study confirmed Wzy activity as a polymerase and that the recombinant *wzy* gene (with addition of C-terminal myc and His tags) was fully functional.

Expression and detection of Wzy

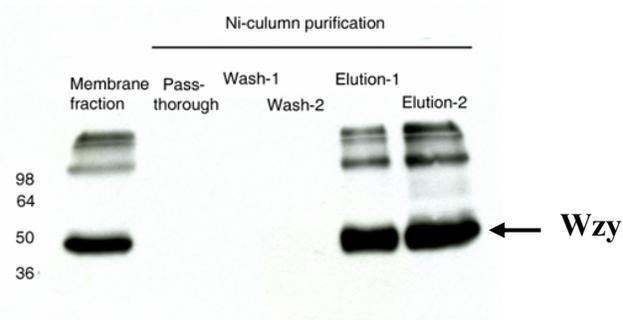


Figure 2.13. Overexpression and detection of Wzy.

We constructed a recombinant plasmid pBAD-*wzy*, encoding Wzy-myc-His₆ with a C-terminal myc epitope for Western-blotting detection and a His₆ epitope for purification by a Ni²⁺ affinity column. The pBAD-*wzy* plasmid was co-transformed with the GroEL/GroES chaperone expression vector into C43 (DE3) cells for co-expression.

Membrane fractions were isolated and analyzed by SDS-PAGE and Western blotting. A band with an apparent molecular weight of 48 kDa was observed in the membrane fraction and during elution (Figure 2.13), consistent with the Wzy molecular weight. Bands at higher molecular weights were also observed, probably due to the aggregation of Wzy resulting from the mild denaturation conditions. Similar observations were also noted in other studies where the expression of integral membrane proteins was attempted. If *wzy* was expressed in the absence of the chaperone system, only high molecular weight protein aggregates were observed (data not shown). Thus, it was speculated that the GroES/GroEL chaperone facilitated the correct folding of Wzy during over-expression.

2.7 In vitro Polymerization of Repeating Units by Wzy

2.7.1 Experimental Methods

The ³H-labeled pentasaccharide-PP-Und was prepared for the final enzymatic reaction using WbnI and 50 μM UDP-[³H]-Gal. For the paper chromatography assay, the reaction mixture (10 μL) containing 0.03 mM radiolabeled pentasaccharide-PP-Und, 0.4% Triton-X100 and 50 μg Wzy membrane fraction was incubated at room temperature for 3 hr in 20 mM Tris-HCl (pH 7.5), 5 mM MnCl₂ and 0.1 M NaCl. The reaction was quenched with an equal volume of cold methanol. After centrifugation to remove the precipitate, the mixture was concentrated to < 10 μL and spotted on Whatman 3MM chromatography paper (1×20 cm). The paper strip was developed in isobutyric acid: 1 M NH₄OH (5:3, v/v). The radioactivity was analyzed as described. For the SDS-PAGE assay, the reaction mixture was quenched with an equal volume of SDS loading buffer

containing 50% glycerol and 0.001% bromophenol blue. The mixture was then electrophoresed in the Bio-Rad Mini-PROTEAN system. The gel was soaked in Amplify Fluorographic Reagent (GE Healthcare) for 45 min, dried and exposed to Hyperfilm MP (GE Healthcare) at -80°C for 6 days before development.

2.7.2 Results and Discussion

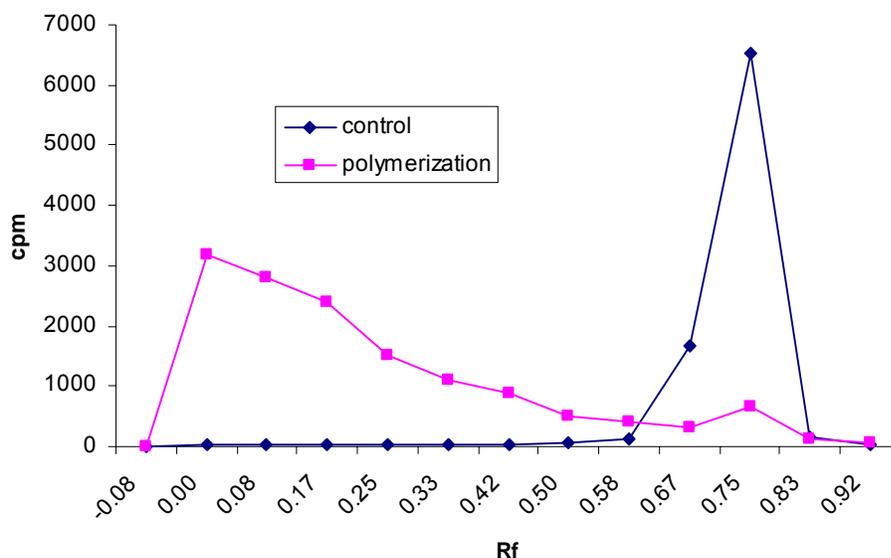


Figure 2.14. Radioactive based paper chromatography assay of Wzy polymerization.

To investigate if Wzy catalyzes the polymerization of repeating units, we carried out an *in vitro* reaction which was monitored using a radioactive-based assay. Repeating unit pentasaccharide-PP-Und was labeled with ^3H in the final enzymatic reaction through use of WbnI in the presence of UDP- ^3H Gal. The radiolabeled substrate was then incubated with a partially purified Wzy membrane fraction for 3 hr at room temperature.

The reaction was quenched by the addition of cold methanol and analyzed using paper chromatography and SDS-PAGE [124, 125]. The reaction mixture containing the membrane which was isolated from expression of empty pBAD vector has the highest radioactivity at an Rf of 0.75 on the paper. However, in the reaction mixture containing Wzy, the radioactivity appeared at areas with low Rf values (Figure 2.14). This result suggested the formation of higher molecular weight compounds in the presence of Wzy. To further verify this result, we subjected the reaction mixture to SDS-PAGE and detection by autoradiography (Figure 2.15). The radioactive signal on the gel gave a ladder-like pattern, consistent with the formation of multiple repeating units. It is evident from the PAGE assay that Wzy catalyzes the polymerization of repeating units, an essential reaction in polysaccharide biosynthesis that had never been unambiguously reconstituted *in vitro* prior to this study.



1 2

Figure 2.15. Analysis of Wzy polymerization reaction by SDS-PAGE and visualized with auto-radiography. Lane 1. repeating unit polymerization by Wzy, ladder-like pattern can

be seen at high molecular weight area; lane 2. radio-labeled repeating unit-PP-Und with membrane fraction in the absence of Wzy expression.

2.8 In Vitro Reconstitution of Polysaccharide Chain Length Modality

Cell surface polysaccharides have a distinct strain-specific pattern of chain length termed modality. In vivo genetic studies suggest that modality is controlled by the concerted action between Wzy and Wzz, a mechanism that still remains a mystery. Several lines of evidence pointed to the possible existence of Wzy-Wzz interaction, a possible mechanism that may contribute to the chain length modality. In this section, we presented the biochemical evidence of Wzy-Wzz interaction using an in vitro pull-down assay, and demonstrated the generation of modality by a two-component (Wzy/Wzz) polymerization system. This result opens up an exciting possibility to further explore chain length control mechanism via in vitro biochemical reconstitution.

2.8.1 Experimental Methods

Purification of O86:B7 Wzy and O86:H2 Wzz

Plasmid pBAD-*wzz* was constructed and transformed into BL21 (DE3). Wzz was expressed and purified as described. Purified Wzz (solubilized in 0.01% DDM) was stored in 4 °C until use.

Wzy-FLAG was expressed and purified using similar procedures as described in section 2.6.

Interaction of Wzy with Wzz

20 uL of purified Wzy and Wzz were incubated with anti-FLAG agarose or Ni-agarose at 4 °C overnight in a reaction buffer containing 20 mM Tris-HCl (pH 7.5) and 0.2 % DDM. After the incubation, the beads were washed with PBS extensively to remove non-specific binding. The bound proteins on the anti-FLAG agarose and Ni-agarose were eluted by 3x FLAG peptide and 250 mM imidazole, respectively. The eluted proteins were analyzed by SDS-PAGE and Western blotting.

2.8.2 Results and Discussion

Interaction of Wzy with Wzz via pull-down assay

Purified Wzy-FLAG was incubated with Wzz-myc-His in the presence of Ni-NTA beads or anti-FLAG beads, and the proteins were eluted by imidazole and FLAG peptide, respectively. Eluted fractions were then detected with anti-FLAG antibody (for the presence of Wzy) and anti-myc antibody (for the presence of Wzz) (Figure 2.16). The elution fraction from Ni-NTA beads contains Wzz as expected, and importantly, Wzy was also co-eluted as indicated by the positive staining of anti-FLAG antibody (lane 2). On the other hand, when protein mixtures were immunoprecipitated with anti-FLAG antibody, Wzz was pulled down together with Wzy, as indicated by positive staining of anti-myc antibody. In a control experiment, an unrelated protein BAP-FLAG was incubated with Wzz under the same condition. The eluted fraction from Ni-NTA beads does not show any positive staining for anti-FLAG antibody (data not shown), indicating that the interaction between Wzy and Wzz is specific.

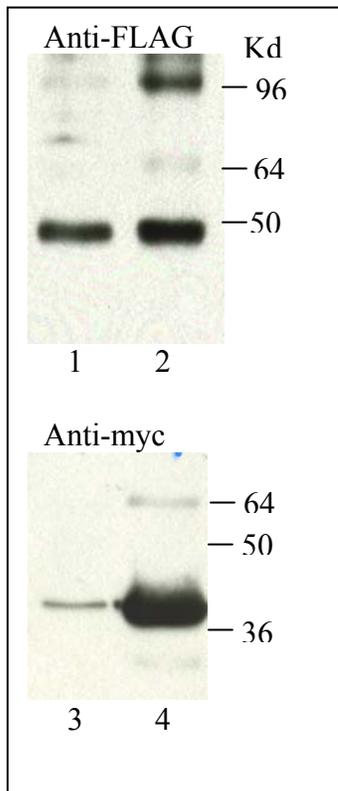
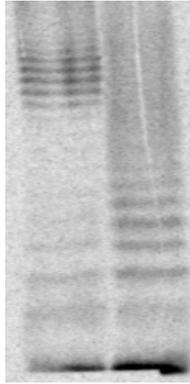


Figure 2.16. Pull-down assay for probing Wzy-Wzz interaction. 1, Wzy precipitated by anti-FLAG beads; 2, Wzy co-eluted with Wzz from Ni-NTA beads; 3, Wzz co-precipitated from anti-FLAG beads; 4, Wzz eluted from Ni-NTA beads.

In vitro generation of chain length modality

The *in vitro* demonstration of Wzy-Wzz interaction poses an intriguing hypothesis about its possible effect on chain length modality. To investigate this possibility, purified Wzz was added to the previous polymerization reaction mixture containing radio-labeled repeating unit substrate and Wzy (membrane fraction). The reaction was allowed to proceed for overnight, quenched and analyzed by SDS-PAGE/autoradiography as described previously in section 2.7. The result (Figure 2.17) showed a different localization of polysaccharide concentration compared to the sample without the addition of Wzz, an observation appearing to show chain length modality. To further confirm that this observation actually reflects the effect from the addition of Wzz, we purified a second Wzz: O86:B7 Wzz, and carried out the similar reaction. Notably, the modality of

polysaccharides appeared to be quite different from the ones with Wzz from O86:H2. Low molecular weight polysaccharides have more pronounced concentration, a modality that is consistent with *E. coli* O86:B7 LPS modality.



1 2

Figure 2.17. In vitro generation of chain length modality. 1. Polymerization with Wzz (O86:H2); 2. Polymerization with Wzz (O86:B7).

CHAPTER 3

ENGINEERING BACTERIAL POLYSACCHARIDES BY METABOLIC PATHWAY REMODELING

3.1 Introduction

The ability to introduce structural modifications into biomolecules presents a powerful tool to dissect their functions and roles in biological processes. For example, introduction of non-natural amino acids into proteins by various chemical and genetic means has permitted the characterization protein folding, enzyme catalytic mechanisms and ligand-receptor interactions [126]. Recently, metabolic introduction of bioorthogonal functional groups such as azide and alkyne into mammalian cell surface oligosaccharides has become a powerful technology for studying glycosylation in native cellular environments [63, 127]. Compared to these fascinating research accomplishments, the area of exploring bacterial polysaccharides with chemical modifications is largely undeveloped, probably due to the complexity of polysaccharide biosynthetic mechanisms and the seemingly unlimited structural diversity of polysaccharides. To date, only a handful of examples have been demonstrated. Jennings and coworkers chemically converted N-acetylated native Group B meningococcal polysaccharides (a homopolymer of α 2,8-linked sialic acids) into N-propionylated

derivatives, which have demonstrated to induce high titers of polysaccharide specific IgG antibodies [128]. The Nishimura group chemically synthesized a series of modified bacterial cell wall precursors tagged with fluoresceine and a keto-group. They were then able to show that these precursors could be metabolized onto the cell wall [129]. In another example, Bertozzi, Gibson and coworkers demonstrated the incorporation of N-acyl modified sialic acid analogs into *Haemophilus ducreyi* cell surface lipooligosaccharides [130]. In these examples, the modifications either rely on the limited chemical functionalities in the native structures that can be utilized in an in vitro chemical reaction, or they can not achieve homogenously modified polysaccharides by simply introducing unnatural substrates into the native biosynthetic pathway. Therefore, there exists a need to develop a strategy for introducing modifications into polysaccharides, which is applicable in general, facile for further derivatization, and capable of achieving high homogeneity.

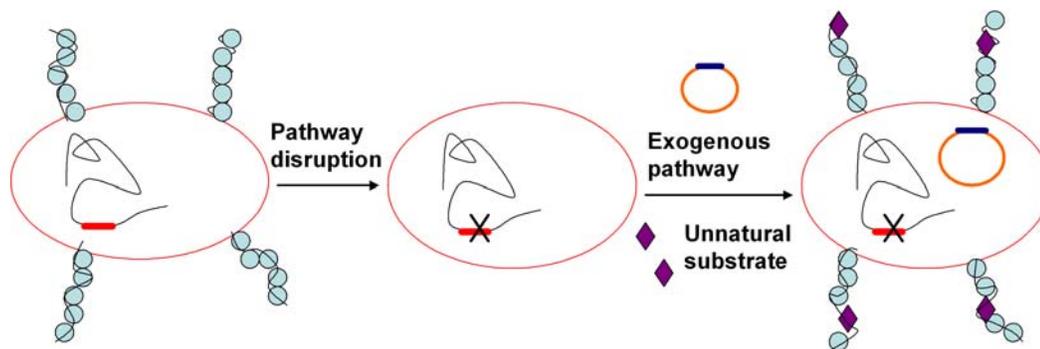


Figure 3.1. Strategy for in vivo generation of structurally modified polysaccharides.

Toward this goal, in this study we demonstrated in *Escherichia coli* that an exogenously introduced sugar nucleotide salvage pathway can not only functionally complement the native de novo pathway, but more importantly, is promiscuous enough to metabolize different sugar analogs and incorporate into polysaccharides, resulting in homogenous modifications (Figure 3.1). Moreover, the bioorthogonal chemical handle incorporated has the potential for further chemical elaborations, increasing the diversity of modifications on bacterial cell surface.

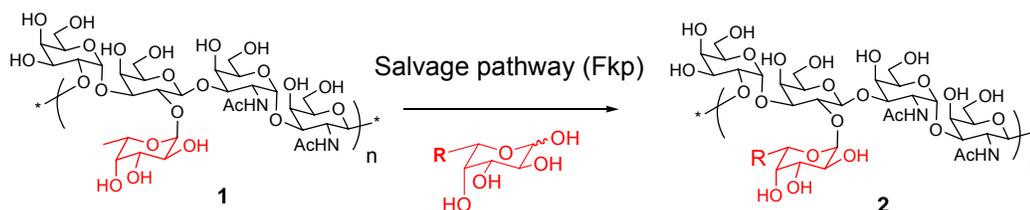
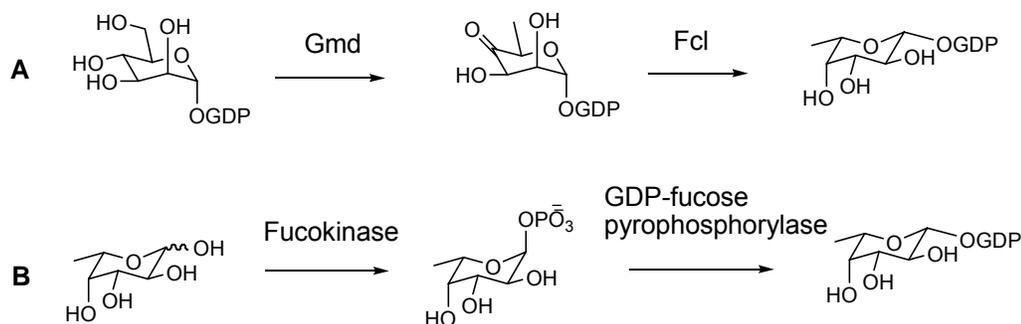


Figure 3.2. Incorporation of fucose analogs into *E. coli* O86 polysaccharides.

We used *E. coli* O86 O-polysaccharide as a model system. Structure characterization of its polysaccharide by our group revealed the presence of a blood group B-type sugar sequence, in which fucose is an important sugar component. We chose to introduce fucose analogs into this polysaccharide to demonstrate our strategy (Figure 3.2), since fucose, as a common monosaccharide component of bacterial polysaccharides, has been implicated as the main determinant for bacterium-induced host immune reactivity and is often the virulence factor [131]. Introducing different analogs can help elucidate the roles of fucose in infection and pathogenicity. There are two pathways to generate GDP-fucose [131], the sugar donor for incorporation of fucose into

polysaccharides (Scheme 3.1). One is called the *de novo* pathway, in which GDP-fucose is formed from GDP-mannose by GDP-mannose dehydratase (Gmd) and fucose synthetase (Fcl). The other one is the salvage pathway, in which free fucose is firstly phosphorylated by fucokinase, and then converted to GDP-fucose by GDP-fucose pyrophosphorylase. The *de novo* pathway is present in both prokaryotic and eukaryotic organisms, while the salvage pathway is thought to be present only in eukaryotic organisms. Recently, the Comstock group [132] reported that the human symbionts *Bacterioides* possess an endogenous fucose salvage pathway, and that the bacteria utilize this pathway to incorporate fucose into surface capsular polysaccharides and glycoproteins. This is the only report of the fucose salvage pathway being present in a bacterial system. In our study, we functionally replaced the *E. coli* native *de novo* pathway with the fucose salvage pathway, with which we were able to introduce a panel of fucose analogs into polysaccharides with high efficiency and fidelity.



Scheme 3.1. Two pathways for GDP-Fucose biosynthesis: **A**, *de novo* pathway; **B**, salvage pathway.

3.2 Experimental Methods

Bacteria strains and plasmids

E. coli O86:B7 (ATCC 12701) was obtained from American Type Culture Collection (Rockville, MD). *E. coli* competent cell DH5 α [*lacZ*Δ*M15 hsdR recA*] was from Gibco-BRL Life Technology. *E. coli* competent cell BL21 (DE3) [*F*⁻ *ompT hsdS_B* (*r_B*⁻ *m_B*⁻) *gal dcm* (DE3)] were from Stratagene (La Jolla, CA). Expression plasmid pET15b was purchased from Novagen (Carlsbad, CA). Plasmids pKD20, pKK232-8 and pTRC99A were kindly provided by L. Wang from Nankai University, China. Chemicals and solvents were from Sigma-Aldrich.

Cloning, expression, purification and characterization of Fkp

The *fkp* gene was PCR amplified from the *B. fragilis* 9343 (ATCC 25285) chromosome DNA using Pfx50 DNA polymerase (Invitrogen) and primers 5'-CTGGACCCATATGCAAAACTACTATCTTTACCGTCC and 5'-ATGGATCCTTATGATCGTGATACTTGGAATCCCTT. The PCR product was digested with NdeI and BamHI and then cloned into pET15b vector (Novagen). The recombinant vector was confirmed by DNA sequencing and transformed into *E. coli* BL21 (DE3) for expression.

For expression, the clones were cultured in LB broth with 100 g/mL of ampicillin at 37 °C until OD600 reached 0.8. Protein expression was induced by the addition of 0.5 mM isopropyl-1-thio-b-D-galactopyranoside (IPTG) and allowed to proceed at 25 °C for 12 h. Then cells were harvested and resuspended in 50 mM Tris-HCl buffer (pH 7.5) with 500 mM NaCl, after which the cells were disrupted by sonication. The cell lysate was centrifuged with 15000 rpm for 20 min at 4 °C. Imidazole was added into the supernatant

to a final concentration of 5 mM. HiTrap Chelating HP column (5 mL, Amersham Biosciences) was used for Ni-affinity purification with 50 mM Tris-HCl buffer (pH 7.5) (containing 500 mM NaCl, 5 mM imidazole for binding buffer, 40 mM imidazole for washing buffer and 300 mM imidazole for elution buffer). Eluted protein were exchanged and concentrated into 50 mM Tris-HCl buffer (pH 7.5) with 10% glycerol by centrifugation. The proteins from different purification steps were analyzed by SDS-PAGE (Figure 3.4). The purified protein concentration in the final solution was determined to be 2.2 mg/mL.

For *in vitro* Fkp activity assay, 100 μ L enzymatic reaction was performed in 50 mM Tris-HCl buffer (pH 7.5) containing 5 mM MgSO₄, 10 mM ATP, GTP and L-fucose, 1 unit of inorganic pyrophosphatase, and 22 μ g purified Fkp. The reaction was incubated at room temperature for 30 min and quenched in boiling water for 1 min. The formation of GDP-fucose was analyzed by thin-layer chromatography (TLC) and further confirmed by LC-MS. TLC was performed on silica gel plates (Sigma-Aldrich). The plates were pre-run with a 1-butanol/acetic acid/H₂O = 2/1/1 (v/v/v) mobile phase and visualized by staining with anisaldehyde/MeOH/H₂SO₄ = 1:15:2 (v/v/v). L-fucose was completely converted into fucose-1-phosphate (Fuc-1-P) but the reaction from Fuc-1-P was not completed (Figure 3.5). The formation of GDP-fucose was further identified in LC-MS: m/z [M⁻+H] = 588.65.

Analysis of substrate specificity of Fkp using capillary electrophoresis

Enzymatic assays were carried out in duplicates in a total volume of 20 μ L in HEPES buffer (200 mM, pH 7.0) containing 5 mM MgCl₂, 1 mM ATP, 1 mM GTP, 1 mM substrate and 0.6 μ g recombinant Fkp. Reactions were allowed to proceed for 15 min at

37°C and quenched by heating at 95°C for 2 min. The reaction mixture was then diluted 10-fold and kept on ice. Samples were analyzed by a P/ACETM MDQ Capillary Electrophoresis System equipped with a UV detector (Beckman Coulter, Fullerton, CA, USA). CE conditions were as follows: 75 µm i.d. capillary, 25 KV/80µÅ, 5 s vacuum injections, monitored at 254 nm, running buffer was 25 mM sodium tetraborate, pH 10.0.

Functional inactivation of fucose *de novo* pathway genes *gmd-fcl*, complementation and LPS analysis

The *gmd-fcl* gene was replaced by a chloramphenicol acetyltransferase (CAT) gene using the RED recombination system of phage lambda. The CAT gene was amplified from plasmid pKK232-8 using primers binding to the 5' and 3' ends of the gene, with each primer carrying 50 bp which flanks *gmd-fcl* gene. The PCR product was then transformed into *E. coli* O86:B7 type strain carrying pKD20, and chloramphenicol-resistant transformants were selected after induction of the RED genes. PCR using primers specific to the CAT gene and *E. coli* O86:B7 DNA flanking the *gmd-fcl* gene was carried out to confirm the replacement. To complement the *gmd-fcl* deficient mutants, the *gmd-fcl* gene was amplified from *E. coli* O86:B7 genomic DNA and then cloned into pTRC99A vector. The recombinant vector was then transformed into the Δ *gmd-fcl* strain for trans-complementation. To transform fucose salvage pathway, recombinant vector pET15b-*fkp* was transformed into the Δ *gmd-fcl* strain. The transformants were grown in LB medium alone or supplemented with 0.1% (final concentration) of various sugars: glucose, fucose, L-galactose, D-arabinose, 6-amino-L-galactose, 6-azido-L-galactose, 6-aldehyde-L-galactose, 6-keto-L-galactose and 6-alkyl-L-galactose. LPSs were extracted

according to the published proteinase K digestion protocol, and analyzed by SDS-PAGE coupled with silver staining.

Analysis of LPS by mass spectrometry:

Following the demonstration on SDS-PAGE/silver staining the incorporation of fucose analogs into *E. coli* O86 O-polysaccharides, we want to do further confirm the structure of the modified polysaccharides by mass spectrometry, which is more informative and definitive.

Small scale isolation of LPSs from wild-type and mutants was carried out as follows: Bacterial cells were collected from 1 mL overnight culture, supplemented with various sugars. The cells were washed with 1 mL H₂O and resuspended in 0.5 mL H₂O in a 1.5 mL centrifuge tube. An equal volume of liquefied phenol was added and the tube was heated to 65°C for 40 min, with vigorous vortexing every 10 min. The mixture was centrifuged at 13,000 rpm for 30 min. The upper aqueous phase was transferred to a new tube and the organic phase was re-extracted with 0.5 mL of H₂O. The combined aqueous phase was concentrated to 100 uL before subjected to precipitation by 150 uL 100% ethanol. The LPS was collected by centrifugation for 45 min. This LPS sample needs to be O-deacetylated, since it is more amenable to MS analysis. To do that, dried LPS sample was dissolved in 100 uL anhydrous hydrazine at 37°C for 45 min. Then chilled acetone (400 uL) was added to remove the excess hydrazine. The solution was kept at -20°C overnight, or -80°C for 1 hr, before centrifugation to pellet the LPS. Finally, the O-deacetylated LPS was dissolved in 20 uL of H₂O and subjected to drop dialysis. The sample after dialysis is ready for CE-MS analysis.

Labeling polysaccharides via *in vitro* Cu(I) mediated [3+2] cycloaddition

Extraction of LPS from bacterial cells was carried out as described. Briefly, 20 mL overnight bacterial culture was centrifuged at 6000 g to harvest the cells. The cells were then washed with 30 mM Tris-HCl (pH = 8.0), and resuspended in buffer A containing 30 mM Tris-HCl, pH 8.0, 20% sucrose, and 10 mM EDTA. The cells were lysed by sonication, followed by centrifugation (10,000 g). The supernatant was collected and subjected to ultracentrifugation (110,000 g, 4 °C, 2 hr). The precipitated gels were resuspended in distilled water for labeling reaction. The cycloaddition conditions were followed as described with TBTA as the triazole ligand. The reaction was allowed to proceed at room temperature for 2 hr. The labeled product was harvested by ultracentrifugation, washed to remove excess reagents. After processed with proteinase K, the label LPSs were electrophoresed and detected with UV light. Various controls showed that azido-containing polysaccharides were specifically labeled under click chemistry reaction conditions.

Labeling live bacterial cells via chemical reactions

Amine and NHS activated ester: 200 µL overnight culture of *E. coli* O86 mutants that express L-galactose-6-NH₂ was collected and washed three times with PBS buffer to remove the media, and subsequently resuspended in 0.8 mL PBS. A 10 mM solution of the biotin reagent ? was prepared by dissolving 2.0 mg reagent in 300 µl of ultrapure water. A 200 µl of the biotin solution was added to cell resuspension and the reaction mixture was incubated at room temperature for 1 hr. The reaction was centrifuged to discard the reaction supernatant. Labeled cells were washed extensively with PBS to remove the unreacted reagent. Cells were then resuspended in 0.5 ml staining buffer, to

which 1 μ l of FITC conjugated streptavidin (0.5 mg/ml) was added. The resuspension was incubated at 4°C with mild agitation for 45 min. Cells were again washed with PBS extensively before subjected to flow cytometry analysis.

***Cu(I)* mediated [3+2] cycloaddition (click chemistry):** Mutants expressing L-galactose-6-N₃ were subjected to cycloaddition reaction as described in the labeling of LPS. Labeled cells were washed extensively with PBS, stained with FITC conjugated streptavidin, and subsequently subjected to flow cytometry analysis.

3.3 Results and Discussion

3.3.1 Disruption of GDP-fucose *de novo* pathway in *E. coli* O86

We disrupted the *de novo* pathway in this *E. coli* O86 by specifically replacing the *gmd-fcl* gene with CAT gene using the RED recombination system as described [39]. Cell surface lipopolysaccharides (LPSs) from mutant Δ *gmd-fcl* and wild-type strains were isolated according to the standard procedure and visualized with silver staining. The wild-type strain, as expected, displayed a typical ladder-like smooth LPS phenotype (Figure 3.3, lane 1). The Δ *gmd-fcl* mutant strain, on the other hand, lacked the ladder-like pattern and displayed a semi-rough phenotype (Figure 3.3, lane 2). Therefore, depletion of the GDP-fucose pool presumably rendered the cell unable to produce a complete repeating unit structure, resulting in the abolishment of polysaccharide formation. The ability of the mutant to produce the smooth LPS phenotype was restored by introduction of a recombinant plasmid pTRC99Af containing the wild-type *gmd-fcl* gene (Figure 3.3, lane 3). This result showed that disruption of the *gmd-fcl* gene has no effect on

downstream biosynthetic genes and that the *gmd-fcl* gene is essential for polysaccharide synthesis.

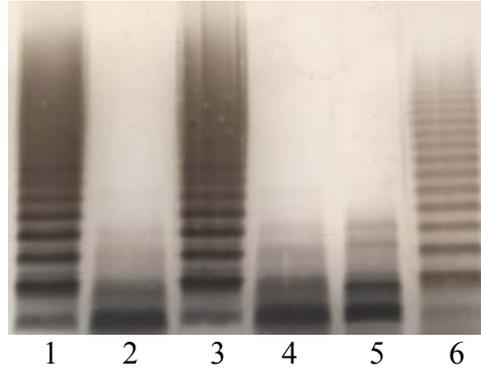


Figure 3.3. Disruption of GDP-fucose *de novo* pathway and complementation with salvage pathway. LPS from wild-type *E. coli* O86; lane 2. Disruption of *de novo* pathway (*gmd-fcl*); lane 3. complementation of *de novo* pathway with pTRC99A-f; lane 4. complemented with salvage pathway (pET15b-fkp), grown in LB without sugar supplement; lane 5. complemented with salvage pathway (pET15b-fkp), grown in LB supplemented with 0.1% glucose; lane 6. complemented with salvage pathway (pET15b-fkp), grown in LB supplemented with 0.1% fucose.

3.3.2 Functional expression of GDP-fucose salvage pathway in *E. coli* O86 mutant

Next, we introduced the salvage pathway into the mutant strain. The salvage pathway gene *fkp* was cloned from *Bacteroides fragilis* strain NCTC9343. The *fkp* gene encodes a bifunctional enzyme with its N-terminal portion possessing pyrophosphorylase activity and its C-terminal portion possessing fucokinase activity. To confirm its function of converting free fucose to GDP-fucose *in vitro*, we constructed *fkp* gene in pET15b

vector and expressed in *E. coli* BL21 (DE3). The proteins were purified by Ni²⁺ affinity chromatography to near homogeneity (Figure 3.4). A 50 µL reaction containing ATP, GTP, fucose and Fkp enzyme was carried out. GDP-fucose was identified in the reaction mixture by TLC and LC-MS after 30 min incubation (Figure 3.5). This result confirmed the in vitro enzymatic function of Fkp. The *fkp* gene was then introduced in trans (pET15b-*fkp*) into Δ *gmd-fcl* strain. When the complemented strain was grown in LB medium supplemented with fucose, the LPS phenotype was restored to that of the wild-type (Figure 3.3, lane 6). However, in the medium supplemented with glucose, the semi-rough LPS phenotype did not change (Figure 3.3, lane 5). This result suggests that the complemented strain can directly use fucose from the medium for incorporation into its polysaccharides, and therefore, the fucose salvage pathway is functionally expressed in *E. coli* and can replace the *de novo* pathway to generate polysaccharides in the presence of exogenous fucose.

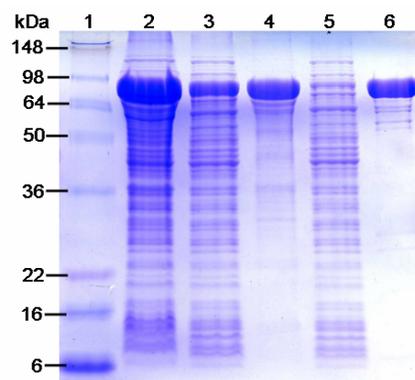


Figure 3.4. Expression and purification of Fkp. 1. Protein Standard; 2. Proteins in whole cell of Fkp-expressing BL21(DE3); 3. Soluble protein fractions; 4. Inclusion body; 5. Flow-through from Ni-affinity purification; 6. Eluted Fkp from Ni-affinity purification.

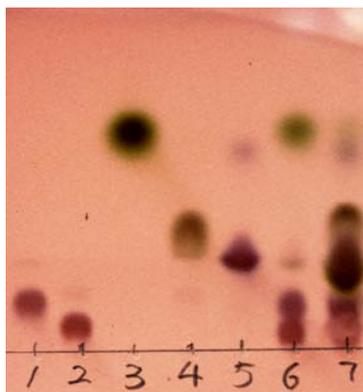


Figure 3.5. In vitro Fkp reaction with fucose as the substrate. 1. ATP; 2. GTP; 3. L-fucose; 4. GDP-fucose; 5. ADP; 6. starting point of GDP-fucose synthesis reaction; 7. GDP-fucose synthesis reaction after 30 min.

3.3.3 Promiscuity of fucose salvage pathway towards different fucose analogs

We next investigate if the salvage pathway is promiscuous for different fucose analogs. Six analogs (Figure 3.6) were used in the study, with modifications at the C-6 position. This position has been demonstrated to be quite tolerant by the mammalian fucose salvage pathway. A volume of 20 μ L reaction for fucose and each analog was carried out, and the formation of the corresponding GDP-sugars was detected and quantified by capillary electrophoresis (CE). Specific activities of Fkp enzyme toward different substrates were derived from conversion of substrate to product at a given reaction time (Figure 3.7). The activity of Fkp toward fucose is 180 mU/mg. The activity toward L-Gal and Fuc-6-N₃ was similar, around 150 mU/mg, approximately 80% of that

compared to the fucose. The activity toward D-Ara, however, was 260 mU/mg, about 1.5 times of that toward fucose. The result demonstrated that Fkp has relaxed specificity toward modifications at C-6 position of fucose *in vitro*.

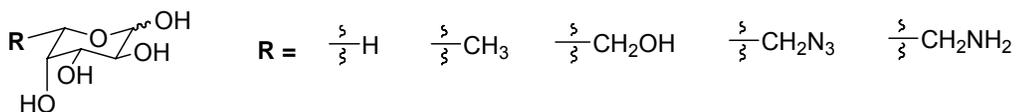


Figure 3.6. Fucose analogs used in the PS engineering.

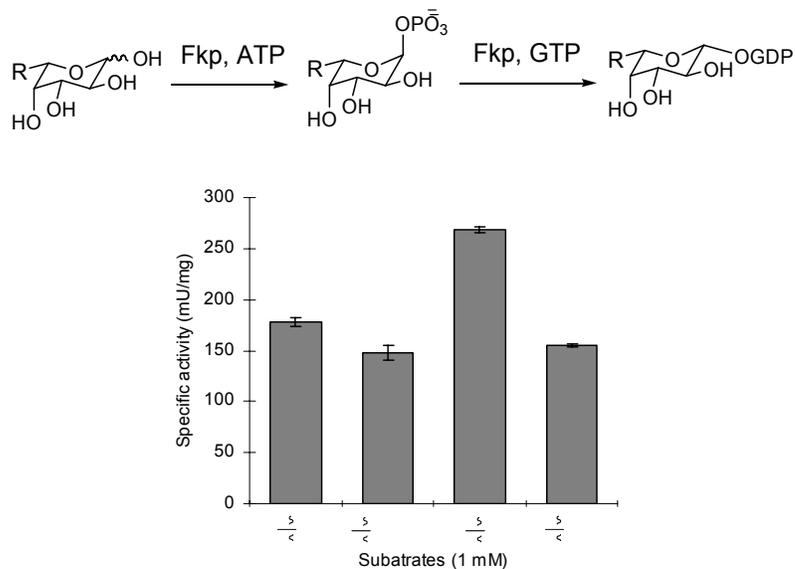


Figure 3.7. *In vitro* analysis of the specificity of Fkp toward fucose analogs.

3.3.4 Incorporation of fucose analogs into polysaccharides

Having demonstrated promiscuity of the fucose salvage pathway, we turn to investigate if these fucose analogs can be metabolized and incorporated into polysaccharides *in vivo*. The $\Delta gmd-fcl(fkp)$ strain was grown in LB medium supplemented with four analogs respectively. These analogs have no toxicity to the cells, since the cell growth rate was comparable to the wild-type strain. The LPSs isolated from these complemented strains showed the ladder-like pattern, suggesting the presence of polysaccharides (Figure 3.8). However, the LPSs have lower chain length distribution (up to 8 repeating units in the case of D-Ara modification) compared to the wild-type, which has more than 20 repeating units. This observation may reflect the specificity of Wzy enzyme, which catalyzes the polymerization of repeating units in the biosynthetic pathway. Nevertheless, it appears that the fucose analogs can restore the ability of the bacterium to produce polysaccharides, though to a less extent compared to fucose.

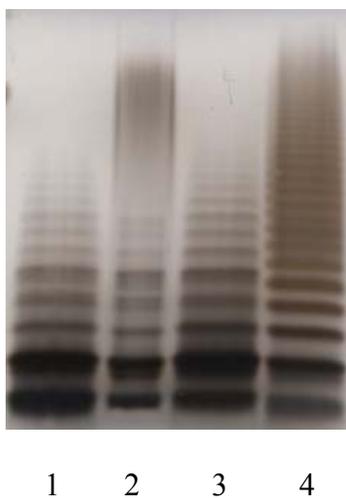


Figure 3.8. Silver staining analysis of LPSs containing different fucose analogs. 1, L-galactose; 2, 6-azido-L-galactose; 3, D-arabinose; 4, L-fucose.

In order to confirm the incorporation of fucose analogs into polysaccharides, we used CE-MS technique to determine the corresponding polysaccharide structures from mutant cells. Figure S6 (Appendix) shows the full spectra of O-deacetylated LPSs from wild-type and corresponding mutant cells. In each spectrum, lipid-A associated ions were detected at m/z 855.9, 468.6 and 388.5, corresponding to the consecutive loss of a H_3PO_4 , a N-acylated glucosamine and the second H_3PO_4 group, respectively. The information on the repeating unit could be easily obtained by selecting the ion at m/z 204.2 as a precursor to perform precursor ion scan experiments, which corresponds to the HexNAc oxonium ion. The precursor-treated spectra showed a close-up at repeating units from each strain and the fragmentation patterns. In the spectra, major peaks were identified and the corresponding sugar sequences were assigned (Figure 3.9). The peaks corresponding to the sugar sequences that contain fucose or fucose analogs are circled in red. In the wild-type strain, two peaks m/z 675.0 and 878.0 represent the sugar sequences Fuc-HexNAc-Hex₂ and Fuc-HexNAc₂-Hex₂, respectively. The presence of a fucose residue (Mw = 146) can be derived from the difference between the ions of m/z 675.0 and m/z 529.0, as well as that between m/z 878.0 and m/z 732.0. Analysis of the LPS from mutant cells with L-galactose substitution showed two characteristic peaks m/z 691.0 and m/z 894.0, corresponding to sugar sequences Gal-HexNAc-Hex₂ and Gal-HexNAc₂-Hex₂, respectively. The presence of this Gal residue (Mw = 162) was also confirmed by the difference between the ions of m/z 691.0 and m/z 529.0, as well as that between m/z 894.0 and m/z 732.0. More importantly, in all the spectra for mutants, the characteristic fucose containing peaks m/z 675.0 and 878.0 were not detected, suggesting that the modifications with fucose analogs were highly homogenous.

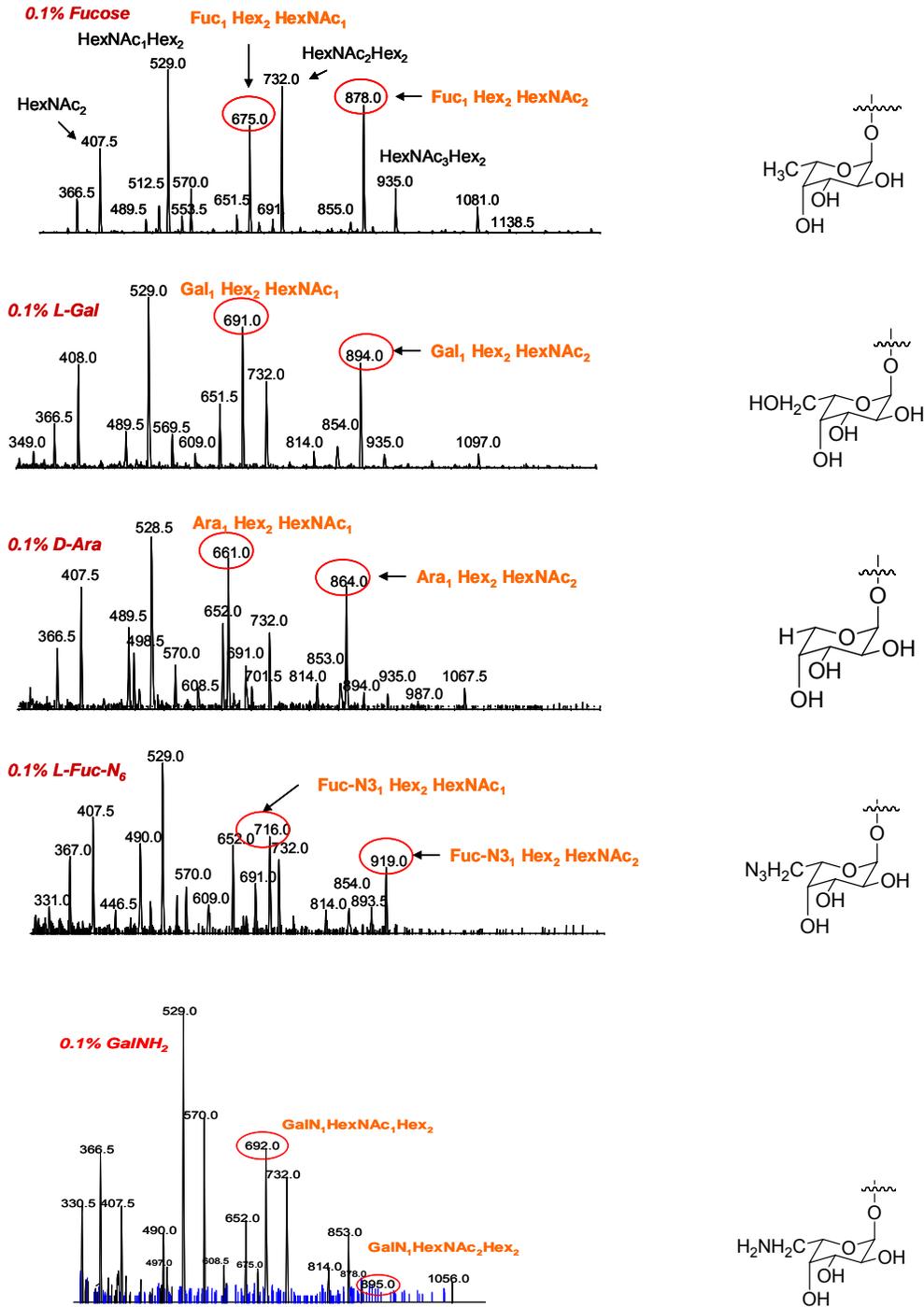


Figure 3.9. Analysis of structurally modified polysaccharides by mass spectrometry.

3.3.5 Fluorescence labeling of azido containing LPS with click chemistry

The incorporation of azido group into the polysaccharides by corresponding fucose analogs can be further derivatized *in vitro* using selective chemical reactions. The azido group, the most versatile chemical handle, is inert in most biological processes. However, it can undergo highly selective chemical reactions [62], including the Staudinger ligation with functionalized phosphines and [3+2] cycloaddition with activated alkynes. These reactions have been exploited for the selective labeling of azido-functionalized biomolecules, such as proteins, sugars and lipids. To illustrate the potential of azido-functionalized polysaccharides for further derivatization, we performed a Cu(I)-catalyzed [3+2] cycloaddition reaction with fluorescein conjugated alkyne reagent (Figure 3.10). The result showed that fluorescence was readily detected in the LPS isolated from the $\Delta gmd-fcl(fkp)$ strain supplemented with azido containing fucose (Figure 3.11). In comparison with various controls (Figure 3.12), it demonstrates that azido-functionalized polysaccharides can be further derivatized to increase the structural diversity by *in vitro* selective chemical ligation.

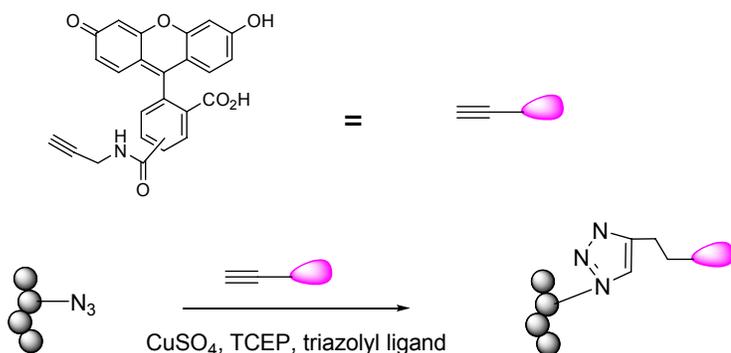


Figure 3.10. Fluorescence labeling of LPS using *in vitro* click chemistry.

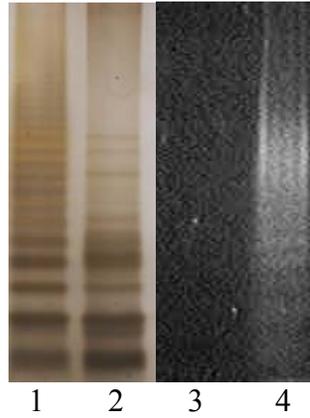


Figure 3.11. Labeling azido containing LPS, silver staining (wild-type LPS, lane 1; azido-containing LPS, lane 2) and fluorescence detection (wild-type LPS, lane 3; azido-containing LPS, lane 4).

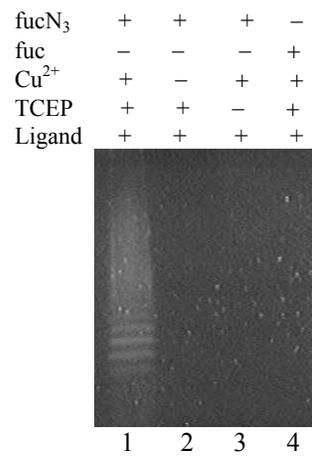


Figure 3.12. Control experiments demonstrate specific labeling.

3.3.6 Live cell labeling by chemical reactions

Specific *in vitro* labeling of azido containing LPS by click chemistry opens up an exciting possibility of direct live cell surface labeling, a technique that can be used in cell imaging, prodrug conjugation and site-specific targeting and delivery applications. Bacterial cells expressing 6-amino-L-galactose and 6-azido-L-galactose were *in vitro* conjugated with sulfo-NHS-LC-LC-biotin and alkyne-biotin, respectively (Figure 3.13). The resulting conjugated cells were then stained with FITC conjugated streptavidin for flow cytometric analysis. As shown in Figure 3.14A, bacterial cells expressing 6-amino-L-galactose showed significantly increased fluorescence intensity compared to control samples. One of the controls, cells expressing fucose that underwent chemical conjugation and FITC staining showed increased fluorescence signal compared to the untreated ones. The possible explanation is that amino group, present in both the core structure of LPS and membrane protein lysine side chain, can react with NHS conjugated biotin, resulting in the positive staining. On the other hand, these cells have lower fluorescence intensity compared to the ones expressing 6-amino-L-galactose, indicating that this modified sugar is incorporated into cell surface LPS where it can be used as for chemical labeling. Cells expressing azido functional group displayed superior labeling efficiency and specificity with extremely low background fluorescence (Figure 3.14B). This result indicates that azido group is more suitable for selective chemical labeling on cell surface.

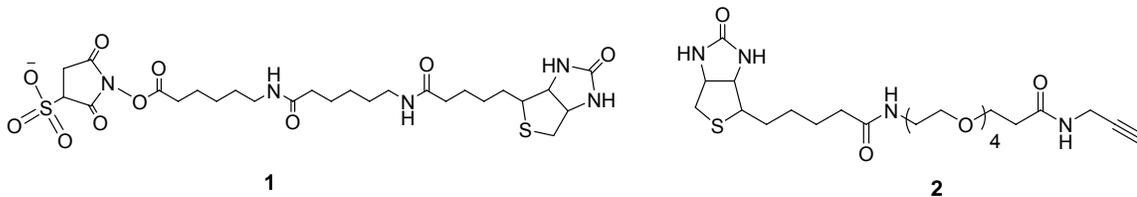


Figure 3.13. Biotin linked reagents for live cell labeling.

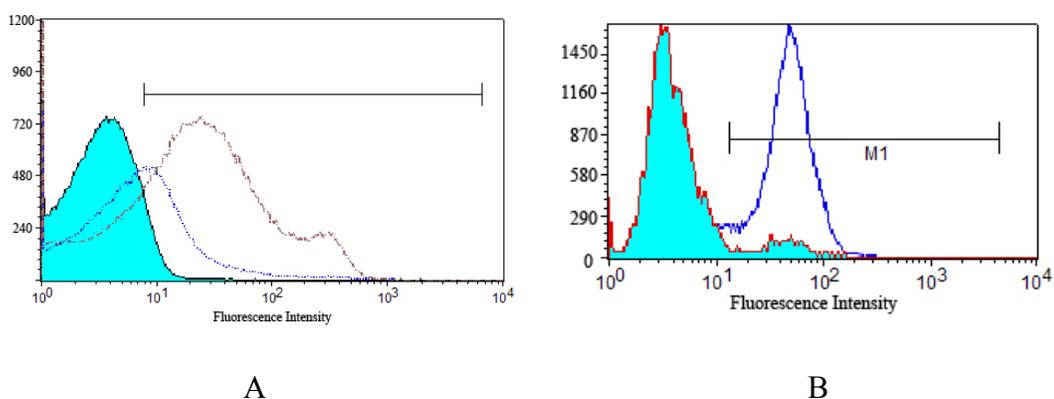


Figure 3.14. Live cell labeling flow cytometry data. (A) amino-expressing cells labeled with biotin compound **1**; (B) azido-expressing cells labeled with biotin compound **2**.

In conclusion, we provide a general, facile and effective means to introduce modifications into polysaccharides *in vivo*. We show in this study that an exogenous promiscuous sugar nucleotide salvage pathway can functionally replace the corresponding *de novo* pathway in the polysaccharide biosynthesis. The promiscuity of the salvage pathway can be exploited to allow incorporation a variety of sugar analogs

into polysaccharides, resulting in highly homogenous modification. In addition, the bioorthogonal chemical handles incorporated can be exploited to further alter the polysaccharide structures via selective chemical reactions in vitro. Even though we only demonstrated the remodeling of polysaccharides with fucose analogs, in principle, this strategy can be applied to other monosaccharide components, such as galactose (Gal), N-acetylgalactosamine (GalNAc), N-acetylglucosamine (GlcNAc) and sialic acid (Neu5Ac). Research is undergoing in our lab to demonstrate the modification of polysaccharides with analogs of monosaccharides other than fucose.

CHAPTER 4

ENZYMATIC SYNTHESIS OF BIOMEDICALLY IMPORTANT

OLIGOSACCHARIDES

Cell surface glycoconjugates contain rich structural information that is central in virtually every aspect of biological processes, ranging from protein folding, cellular development, host-pathogen interaction to immune regulation [1-3]. In recent years, the study of the structure-function relationship of glycoconjugates has been aided tremendously by the *in vitro* synthesis of structurally-defined glycoconjugates and their analogs. Chemical glycosylation, well documented for its synthetic flexibility, has witnessed its success in preparing a wide range of carbohydrate structures. However, the chemical synthesis still suffers from tedious protection/deprotection steps and difficult purification of products. On the other hand, with the rapid development of genomic sequencing, a growing number of glycosyltransferases have been characterized with a wide spectrum of acceptor specificity that has proven invaluable for the facile synthesis of a variety of glycoconjugate structures. It is found that mammalian glycosyltransferases are relatively hard to express in large quantities, and are more rigid in terms of substrate specificity. In contrast, their bacterial counterparts are more easily over-expressed as soluble and active forms without complicated gene manipulation techniques.

Furthermore, bacterial glycosyltransferases seem to have broader acceptor substrate specificities, thereby offering tremendous advantages in the enzymatic synthesis of oligosaccharides and their analogs. In addition, various bacteria exhibit structural mimicry of mammalian glycoconjugates on their cell surface as a survival and infection strategy. Therefore, the corresponding glycosyltransferases might be explored for the *in vitro* synthesis of various glycoconjugates [133]. In fact, a number of bacterial glycosyltransferases have been successfully applied in the synthesis of oligosaccharide moieties of various glycolipids, such as gangliosides [134], blood group antigens [135] and globo-series glycoceramides [73, 76]. In this chapter, I discussed two projects which involved identification of novel bacterial glycosyltransferases and subsequent application in the enzymatic synthesis of biomedically important oligosaccharide sequences.

4.1 A bacterial β 1,3-galactosyltransferase and synthesis of tumor associated T-antigen mimics

4.1.1 Introduction

Mucin-type O-linked glycosylation is a ubiquitous protein post-translational modification in higher eukaryotes, and is involved in a number of fundamental biological processes [136]. For example, mucin-type O-glycans can protect proteins from proteolytic degradation [137], regulate the serum half life of chemokines *in vivo* [138, 139], modulate the intracellular trafficking of proteins [140], mediate cell adhesion events such as sperm-egg fertilization [141], microbe-host interaction and viral infection [142], and serve as tumor-associated antigens [143].

Mucin-type O-glycan is initiated with an α -N-acetylgalactosamine (GalNAc) attached to the hydroxyl group of Ser/Thr side chains, and elaborated by Golgi-resident glycosyltransferases to generate a series of core structures. Core-1 structure (also called T-antigen, Gal- β 1,3-GalNAc- α -O-Ser/Thr) is the precursor of surface antigens in human epithelial and blood group cells [144]. T-antigens are masked in normal tissues, but uncovered during malignancy [13]. Previous research has shown that T-antigens are expressed in >90% of primary human carcinomas, and high expression levels of T-antigens are correlated with increased metastasis and aggressiveness of several types of cancers [145, 146]. Therefore, T-antigens have been used in clinical trials to develop carbohydrate-based cancer vaccines.

E. coli O127 belongs to the O serogroup of enteropathogenic *Escherichia coli* (EPEC) strains, which are important causes of infantile diarrhea in developing countries [147, 148]. *E. coli* O127 strains have been isolated from children with diarrhea worldwide. In addition, *E. coli* O127 was reported to possess high human blood group H (O) activity. The elucidation of the chemical structure of the cell surface O-antigen confirmed that *E. coli* O127 expressed mimicry of human blood group H antigen [149](Figure 4.1). The O-antigen biosynthetic gene cluster (GenBank Accession # AY493508) contains multiple genes needed for the assembly of *E. coli* O127 polysaccharide structures. Among them, three genes (*orf3*, *orf12* and *orf13*) encode putative glycosyltransferases involved in the synthesis of repeating oligosaccharide units. Orf12 (WbiP) contains a conserved domain found in glycosyltransferase family 2. It shows 59% protein sequence identity and 82% similarity to WbnJ from *E. coli* O86, which was previously characterized as a β -1,3-galactosyltransferase [135]. Thus we

propose that WbiP encodes a β 1,3-galactosyltransferase that makes the Gal- β 1,3-GalNAc moiety (T-antigen mimicry) present in the repeating unit structure. In this study, we presented a detailed biochemical characterization of WbiP and a series of mutants. Moreover, we show that this bacterial β -1,3-galactosyltransferase is implicated in the facile synthesis of T-antigen mimics.

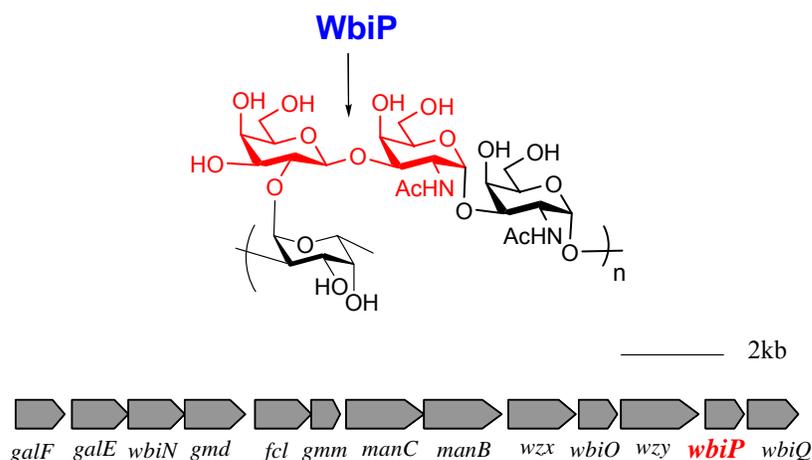


Figure 4.1. O-polysaccharide structure of *E. coli* O127 and its biosynthetic gene cluster. (T-antigen mimic structure catalyzed by WbiP is highlighted.)

4.1.2 Experimental Methods

Bacterial strains, plasmids and materials.

E. coli O127:K63(B8) was obtained from American Type Culture Collection (Rockville, MD, USA). *E. coli* competent cell DH5 α [*lacZ* Δ M15 *hsdR* *recA*] was from Gibco-BRL Life Technology. *E. coli* competent cell BL21 (DE3) [F^- *ompT* *hsdS_B* (*r_B*⁻ *m_B*⁻) *gal* *dcm* (DE3)] were from Novagen Inc. (Madison, WI). Expression plasmid

pET28a was purchased from Novagen (Carlsbad, CA). HiTrap Chelating Ni HP column was obtained from GE Healthcare (Piscataway, NJ). All other chemicals and solvents were from Sigma-Aldrich.

Cloning, expression and purification of WbiP.

The *wbiP* gene was amplified by PCR from the *E. coli* O127 chromosome. The primers with restriction sites underlined for amplification each gene were as follows: Forward: 5' GAGATATACCATATGAAAAATGTTGGTTTTATTG (*NdeI*); Reverse: 5' CCGCCTCGAGTCAACCTAAAATAATGCTTTTATATG (*XhoI*). The DNA fragments obtained were digested with corresponding restriction enzymes and inserted into pET28a vector linearized by the same restriction enzymes to form pET28a-*wbiP* recombinant plasmid. The recombinant plasmid was confirmed by restriction mapping and sequencing. The correct constructs were subsequently transformed into *E. coli* BL21 (DE3) for protein expression. *E. coli* BL21 (DE3) harboring the recombinant plasmid was grown at 37 °C in 1L Luria-Bertani (LB) medium with 35 ug/mL kanamycin antibiotic. When the cells were grown to OD 0.5-0.7, isopropyl-1-thio-β-D-galactopyranoside (IPTG) was added to a final concentration of 0.4 mM. Expression was allowed to proceed for 12 h at 25 °C. Cells were harvested, washed with 20 mM Tris-HCl (pH=7.0) and stored at -80 °C until needed. In protein purification, the cell pellet was resuspended in buffer (20 mM Tris-HCl, pH 7.5, 0.5 M NaCl, 5 mM imidazole, 1 mM PMSF) and disrupted by brief sonication (Branson Sonifier 450, VWR Scientific) on ice. The lysate was cleared by centrifugation (15,000 g, 20 min, 4 °C) and loaded at a low rate of 3 ml/min onto a 5 mL Ni²⁺-NTA chelating column equilibrated with 20 mM Tris-HCl, pH 7.5, 0.5 M NaCl, and 5 mM imidazole. The column was washed with 10 column volumes

of 50 mM imidazole in the same buffer, and the protein was eluted with 250 mM imidazole, followed by analysis using SDS-PAGE.

Radioactive enzymatic assay.

Protein concentration was determined by Bradford method. One unit of enzyme activity was defined as the amount of enzyme required to transform 1 μmol of sugar donor to acceptor per minute at 37 °C. Enzyme assays were performed at 37 °C for 2 hr in a final volume of 100 μL containing 20 mM Tris-HCl, pH 7.0, 10 mM MnCl₂, 0.3 mM radioactively labeled UDP-D-[6-³H]Galactose (specific activity: 4.3 Ci/mmol) as sugar donor, 10mM monosaccharide N-acetylgalactosamine (GalNAc) as sugar acceptor and variable amount of enzymes. Acceptor was omitted as the blank control. The reaction was terminated by adding 100 μl ice cold 0.1M EDTA. Dowex 1× 8-200 chloride anion exchange resin was then added in a water suspension (0.8 mL, v/v = 1/1). After centrifugation, supernatant was collected in a 20 mL plastic vial and ScintiVerse BD (10 mL) was added. The vial was vortexed thoroughly before the radioactivity of the mixture was measured in a liquid scintillation counter (Beckmann LS-3801 counter).

Enzyme assay with HPLC and MALDI-MS.

To a tube containing 0.05 mg GalNAc were added 10 μL of 0.1 M Tris-HCl, pH 7.5, 5 μL of 0.1 M MnCl₂, 5 μL of 50 mM UDP-Gal, and 0.02 mg WbiP (20 μL). The reaction was run at room temperature for 1 hr and quenched by addition of 50 μL cold methanol. The mixture was centrifuged to remove the precipitated proteins. The supernatant was evaporated to dryness. The mixture was then labeled with 2-aminobenzamide (2-AB) as described. Post-labeling cleanup was done by applying the labeling sample onto a Whatman 3MM chromatography paper and performing ascending chromatography in

acetonitrile for 2 hrs. The labeling product was eluted from the paper with 0.5 mL H₂O, and further separated on a normal phase analytic HPLC column as described. The peaks were detected with a fluorescence detector (λ_{ex} =330 nm, λ_{em} =420 nm) and collected and characterized with MALDI-MS.

Site-directed mutagenesis of WbiP.

QuikChange Site-Directed Mutagenesis Kit (Stratagene) was used to construct all the mutants from the parent plasmid, pET28a-WbiP, using the primers shown in the Table 4.1. Expression, purification and enzyme activity assay were carried out following the abovementioned procedures.

Mutants	Primer sequences
D88A	5' CATCGCGAGGATGGCTTCTGATGACATAGC 5' GCTATGTCATCAGAAGCCATCCTCGCGATG
D88E	5' CATCGCGAGGATGGAATCTGATGACATAGC 5' GCTATGTCATCAGATTCCATCCTCGCGATG
D90A	5' CGAGGATGGATTCTGCTGACATAGCCTTACC 5' GGTAAGGCTATGTCAGCAGAATCCATCCTCG
D90E	5' CGAGGATGGATTCTGAAGACATAGCCTTACC 5' GGTAAGGCTATGTCTTCAGAATCCATCCTCG
D174A	5' GTTTTAACTCTGAAGCTTATGATTTGTGGTT 5' AACCACAAATCATAAGCTTCAGAGTTAAAC
D174E	5' GTTTTAACTCTGAAGAATATGATTTGTGGTT 5' CATTCTTAACCACAATTCATAATCTTCAGAG
D176A	5' CTCTGAAGATTATGCTTTGTGGTTAAGAATG 5' CATTCTTAACCACAAAGCATAATCTTCAGAG
D176E	5' CTCTGAAGATTATGAATTGTGGTTAAGAATG 5' CATTCTTAACCACAATTCATAATCTTCAGAG

Table 4.1. WbiP mutants and primers used in the site-directed mutagenesis.

Kinetic studies of wild-type and mutants.

Reactions were performed at 37 °C for 40 mins in a reaction buffer containing 20 mM Tris-HCl (pH 7.5), 10 mM MnCl₂, 0.3 mM of UDP-D-[6-³H]-Gal and 20 µg of enzyme. To determine apparent K_m values, the concentration GalNAc was varied as follows: 10 mM, 20 mM, 40 mM and 80 mM. To determine the K_m value for UDP-Gal, 3 µM of UDP-D-[6-³H]-Gal was supplemented with different amounts of cold UDP-Gal to achieve varying final concentrations (0.08, 0.1, 0.3 and 0.6 mM), with fixed GalNAc concentration at 20 mM. The parameters K_m and V_{max} were obtained by Lineweaver-Burk plots of substrate concentration-initial velocity.

Enzymatic Synthesis of T-antigen mimicry and structural characterization

One pot reaction was conducted for two days at room temperature in a final volume of 1.0 ml containing 20 mM Tris-HCl (pH 7.5), 10mM MnCl₂, 10 mM UDP-Gal, and 15 mM acceptor GalNAc α -OMe, and 20 µg WbiP. The progress of the reaction was monitored by thin-layer chromatography [i-PrOH/H₂O/NH₄OH= 8:2:2 (v/v/v)] conducted on Baker Si250F silica gel TLC plates. Products were visualized by staining solution (anisaldehyde/MeOH/H₂SO₄=1:15:2 (v/v/v)). After complete conversion of donor substrate, protein was removed by brief boiling and centrifugation (12,000 g, 5 min). The supernatant mixture was purified by gel filtration chromatography Bio-Gel P2 (Bio-Rad, CA). The desired fractions were pooled, lyophilized, and stored at -20 °C.

Mass spectra (ESI) were run in the negative model at the mass spectrometry facility at the Ohio state University. Product structure was identified by ¹H and ¹³C NMR spectroscopy

using 500-MHz Varian VXR500 NMR spectrometer. Product Structure was identified through one-dimensional (selective COSY, relay COSY, and NOE) and two-dimensional (COSY, HMQC, NOESY, and HMBC) $^1\text{H}/^{13}\text{C}$ -NMR. The oligosaccharide product was repeatedly dissolved in D_2O and lyophilized before NMR spectra were recorded at 303K in a 5 mm tube. ^1H NMR (500 MHz, D_2O): 1.98 (s, 3H), 3.35 (s, 3H), 3.47 (dd, 1H, $J = 9.9$ Hz, $J = 7.8$ Hz), 3.57 (dd, 1H, $J = 9.9$ Hz, $J = 3.4$ Hz), 3.61 (dd, 1H, $J = 7.7$ Hz, $J = 4.6$ Hz), 3.71 (m, 2H), 3.72 (m, 2H), 3.86 (d, 1H, $J = 3.3$ Hz), 3.91 (dd, 1H, $J = 7.3$ Hz, $J = 5.0$ Hz), 3.96 (dd, 1H, $J = 11.1$ Hz, $J = 3.1$ Hz), 4.18 (d, 1H, $J = 2.9$ Hz), 4.29 (dd, 1H, $J = 11.1$ Hz, $J = 3.7$ Hz), 4.41 (d, 1H, $J = 7.8$ Hz), 4.74 (d, 1H, $J = 3.7$ Hz). δ ^{13}C NMR (125 MHz, D_2O): δ 175.0, 105.1, 98.7, 77.7, 75.3, 72.9, 71.0, 70.8, 69.1, 69.0, 61.6, 61.3, 55.5, 49.0, 22.4. HRMS, m/z : Calculated $\text{C}_{15}\text{H}_{27}\text{NO}_{11} + \text{Na}$: 420.1476; found: 420.1474.

UDP-binding assay.

A 30 μl of aliquot of fresh UDP-beads (CalBiochem) was washed three times with binding buffer (20 mM Tris-HCl pH 7.0, 0.1 M NaCl, 5.0 mM MgCl_2). The beads were incubated on roller at 4°C for 1 h with 40 μl of purified recombinant wild type WbiP and its variant mutants (50 μg proteins) in binding buffer. Negative control was to incubate wild type WbiP with 10 mM UDP to inhibit binding to UDP-beads. The beads were harvested by centrifugation for 1 min at $1000\times g$, followed by washing with binding buffer three times. The bound proteins were treated in SDS loading buffer at 100°C for 5 min before analyzed by SDS-PAGE and western-blotting with anti-His antibody.

4.1.3 Results

Expression, purification and enzymatic assay of WbiP protein

The full open reading frame of *wbiP* gene (750 bp) was amplified from *E. coli* O127 genomic DNA, and subsequently cloned into pET-28a vector. The recombinant plasmid was transformed into *E. coli* BL21(DE3) strain for induced expression with 0.4 mM IPTG at 25°C. WbiP protein has an apparent molecular weight of 28 kDa estimated by SDS-PAGE (data not shown), similar to the theoretical value (28.6 kDa) calculated from its predicted amino acid sequence. Primary sequence analysis shows that WbiP has no transmembrane segments, consistent with the result that we were able to purify WbiP without difficulty even though detergents are omitted.

To investigate the function of WbiP, we tested a panel of radio-labeled sugar donors (UDP-Gal, UDP-GalNAc, GDP-Fuc and UDP-Glc) with GalNAc as an acceptor. The results showed that WbiP displayed high activity with UDP-Gal as the sugar donor, nearly 10 fold reduced efficiency with UDP-GalNAc, and no activity with GDP-Fuc and UDP-Glc. Thus, the donor specificity demonstrated that *wbiP* encodes a galactosyltransferase, consistent with the bioinformatics analysis. To confirm the galactosyltransferase activity of WbiP, we set up a 50 µL scale reaction, which contained 0.05 mg GalNAc, 5 µL of 50 mM cold UDP-Gal, and 0.02 mg WbiP (20 µL). The reaction product was then labeled with 2-aminobenzamide (2-AB), analyzed via a normal phase analytical HPLC column and the molecular weight was characterized by MALDI-MS. The 2-AB labeled starting material GalNAc has a retention volume of 6 mL. When the labeled product of the reaction was analyzed, a new peak with a retention volume of 10 mL was observed (Figure 4.2). When this peak was collected and subjected to

MALDI-MS analysis, a m/z ratio of 526.20 $[M + Na]^+$ was observed (Appendix), consistent with a 2-AB labeled Gal-GalNAc disaccharide. Therefore, the results from both the radioactive assay and the HPLC/MS assay confirm that WbiP encodes a galactosyltransferase.

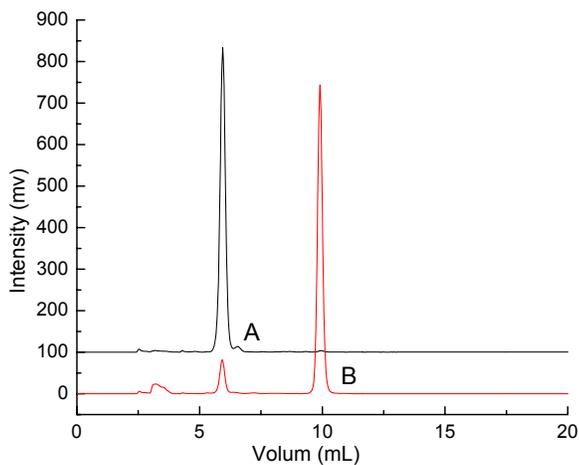
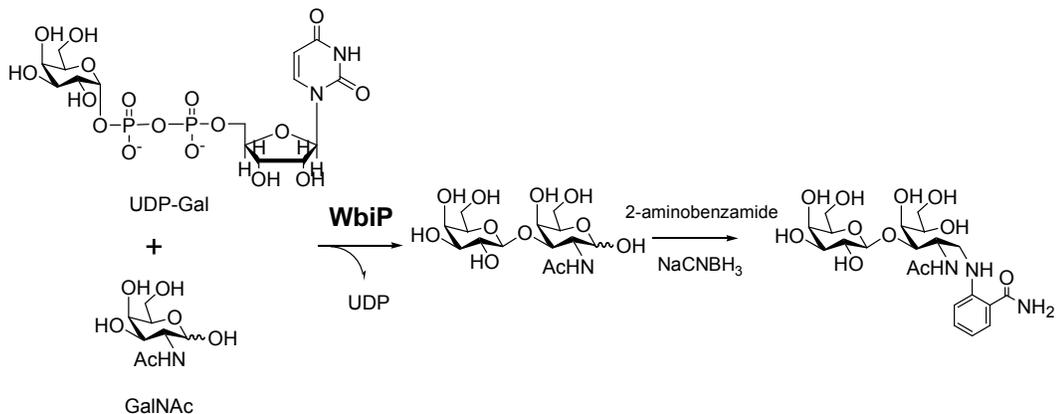


Figure 4.2. WbiP catalyzed reaction and HPLC identification of product.

Characterization of linkage specificity and enzymatic synthesis of T-antigen mimicry

To confirm that WbiP catalyzes the formation of the β -1,3-linkage, a reaction was carried out with GalNAc- α -OMe as an acceptor. GalNAc- α -OMe was used instead of GalNAc due to the fact that GalNAc- α -OMe has the same reactivity as GalNAc (see below) while the fixed configuration of the anomeric carbon by -OMe substitution facilitates the NMR analysis. A milligram scale reaction was conducted using 8 mg GalNAc-OMe and 1.5 equivalents of UDP-Gal. The reaction was incubated at room temperature overnight and monitored by TLC. The disaccharide product was isolated by gel filtration chromatography. A total of 11 mg disaccharide (85% yield) was collected to allow structural analysis by NMR spectroscopy.

Substrate specificity of WbiP

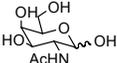
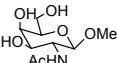
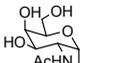
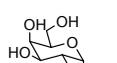
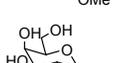
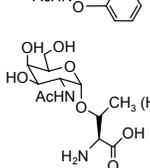
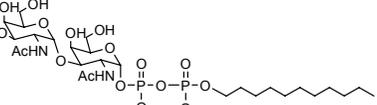
Acceptors		Relative activity %
	(GalNAc)	100
	(GalNAc- β -OMe)	15
	(GalNAc- α -OMe)	105
	(Gal- α -OMe)	N/A
	(GalNAc- α -OPh)	92
	(GalNAc- α -O-Thr(Ser))	95
	(GalNAc- α -1,3-GalNAc- α -PP-C ₁₁)	85

Table 4.2. Substrate specificity of WbiP.

Purified WbiP was used to investigate its specificity towards various sugar acceptors (Table 4.2). The results indicate that the α -configuration at the reducing end of GalNAc is important for enzyme activity since changing the configuration from α to β makes GalNAc a poor substrate for WbiP (GalNAc- β -OMe has only 15% activity compared to GalNAc- α -OMe). The N-acetyl at C2 of GalNAc is essential for enzyme recognition, since no detectable activity was observed when Gal- α -OMe was used as an

acceptor. Furthermore, WbiP can tolerate substitution on the 1-O position of GalNAc, demonstrated by observation that a series of α -substitutions ($-\alpha$ -OMe, $-\alpha$ -OBn, $-\alpha$ -O-Ser/Thr and $-\alpha$ 1,3-GalNAc-PP-C₁₁) have comparable reactivities. It is worth noting that GalNAc- α -O-Ser/Thr are the biosynthetic precursors for T-antigens of mucin-type O-glycans. The fact that these two glycosylated amino acids serve as good substrates indicates that WbiP can be potentially used for the large scale synthesis of T-antigens for biomedical studies. Not surprisingly, WbiP readily accepts α -GalNAc- α 1,3-GalNAc-PP-C₁₁ as a substrate, which is an analog of the immediate synthetic substrate of WbiP based on the *E. coli* O127 O-repeating unit structure. Thus taken together, the acceptor pattern shown in the assay is in agreement with the proposed function of WbiP in the biosynthesis of the *E. coli* O127 O-repeating unit.

Determination of metal ion effect on WbiP

Many glycosyltransferases are found to contain a short conserved amino acid sequence called the DXD motif, and exhibit a requirement for a divalent metal cation for activity. Sequence analysis shows that WbiP contains two DXD motifs (⁸⁸DSD⁹⁰, and ¹⁷⁴DYD¹⁷⁶), suggesting that WbiP activity requires a metal ion cofactor. To further investigate the metal ion requirements, various divalent metal cations as well as EDTA were incubated in a series of separate reactions. As predicted, WbiP activity requires metal ions with Mn²⁺ and Mg²⁺ as the preferred cations (Figure 4.3). The activity was completely abolished in the absence of metal ions or in the presence of 10 mM EDTA. Thus WbiP belongs to glycosyltransferase superfamily A (GTA), in which the DXD motif plays a role in coordinating the binding of sugar nucleotide donors in the active site.

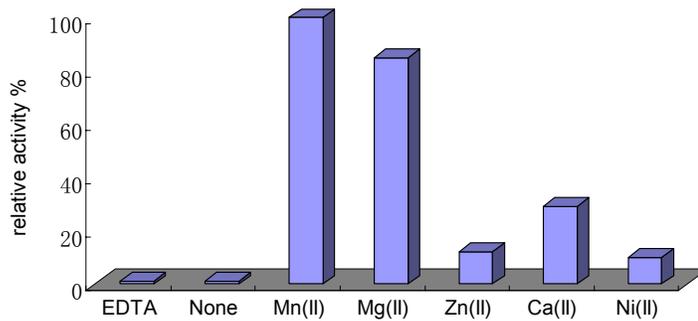


Figure 4.3. Metal ion requirement of WbiP. The concentration of EDTA and metal ions are all at 10 mM.

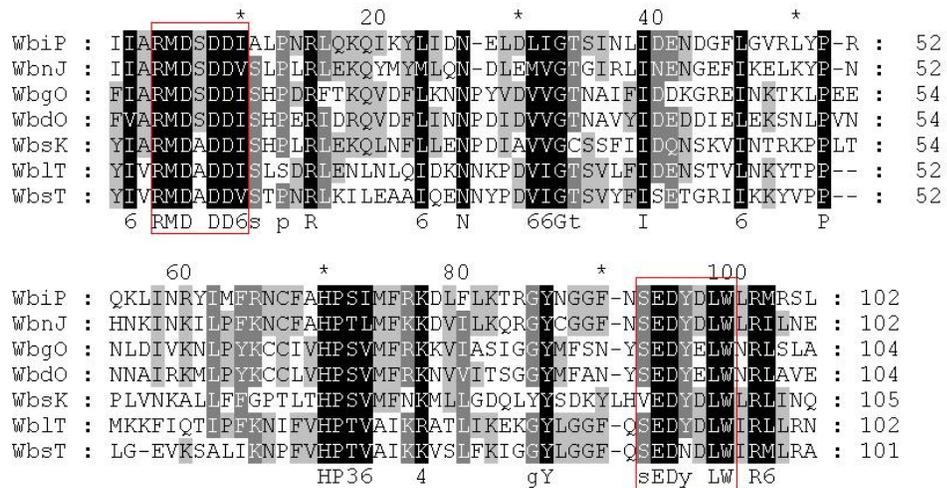


Figure 4.4. Amino acid sequence alignment of segments from different β 1,3-galactosyltransferases. (Conserved DXD motifs are bracketed with red.) WbiP: from *E. coli* O127O127:K63(B8), AAR90893; WbnJ: from *E. coli* O86:H2, AAV80758; WbgO: from *E. coli* O55:H7, AAL67559; WbdO: from *Salmonella enterica* subsp. *salamae*

serovar Greenside 50, AAV34525; WbsK: *E. coli* O128:B12, AAO37699; WblT: from *Photorhabdus luminescens* subsp. *laumondii* TTO1, CAE17191; WbsT: from *E. coli* O126, DQ465248.

Site-directed mutagenesis study of DXD motifs and enzyme kinetics of WbiP and its mutants

Sequence alignment of a group of bacterial β 1,3-galactosyltransferases (putative and characterized) shows that they all possess two DXD or DXD-like motifs (Figure 4.4), suggesting this might reflect a common catalytic mechanism in these galactosyltransferases. In order to investigate the functional role of each DXD motif, we mutated the aspartic acid residues of the DXD motifs in the WbiP sequence. Aspartic acid residues in each DXD motif were mutated to alanine and glutamic acid, respectively. A total of 8 mutations were generated: D88A, D88E, D90A, D90E, D174A, D174E, D176A and D176E. The mutagenic primer pairs are listed in Supporting Information. WbiP mutants were expressed at a similar level compared to the wild-type (data not shown), suggesting that the mutations in the DXD motifs do not compromise protein folding and stability. Enzyme activity of each mutant was assessed using the radioactive assay. The result (Figure 4.5) shows that mutating aspartic acid to alanine at ⁸⁸DSD⁹⁰ completely abolishes the enzyme activity, while the same mutation at ¹⁷⁴DYD¹⁷⁶ still retains some marginal activity (12% and 10% for D174A and D176A, respectively). This observation indicates that ⁸⁸DSD⁹⁰ may play a more essential role in the catalysis than ¹⁷⁴DYD¹⁷⁶. The change of aspartic acid residue to its conserved counterpart, glutamic acid, in ⁸⁸DSD⁹⁰ partially restored enzyme activity (8% for D88E and 10% for D90E),

indicating the importance of charge in these positions for catalysis. The same mutations at ¹⁷⁴DYD¹⁷⁶ restored the enzyme activity to a higher level (23% for D174E and 31% for D176E), further confirming that the requirement of aspartic acid residues in this position is less stringent.

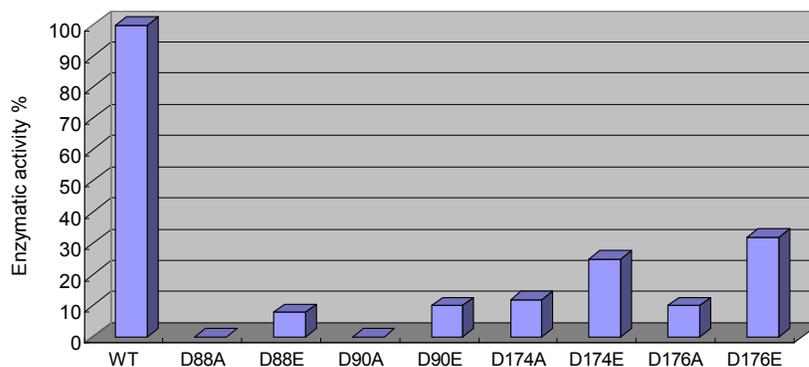


Figure 4.5. Enzymatic activity of WbiP mutants.

Kinetic parameters (Table 4.3) were determined for both the wild-type and mutant enzymes. The wild-type enzyme has apparent K_m values of 31 μ M and 16.1 mM for donor UDP-Gal and acceptor GalNAc, respectively. These are similar to the values obtained in other well-characterized glycosyltransferases. Notably, D88E and D90E mutants have considerably higher K_m values for UDP-Gal, but quite similar K_m values for GalNAc. On the contrary, mutations in D174 and D176 resulted in higher K_m values for GalNAc, but similar values for UDP-Gal. This observation suggests that ⁸⁸DSD⁹⁰ is critical in the binding of sugar donor UDP-Gal, while ¹⁷⁴DYD¹⁷⁶ may participate in the binding of the sugar acceptor. These two DXD motifs thus have different functional roles in catalysis.

WbiP	UDP-Gal			GalNAc		
	K_{cat} (s^{-1})	K_m (μM)	K_{cat}/K_m ($s^{-1} \cdot \mu M^{-1}$)	K_{cat} (s^{-1})	K_m (mM)	K_{cat}/K_m ($s^{-1} \cdot mM^{-1}$)
WT	11.2	31	0.36	15.3	16.1	0.95
D88E	0.08	232	0.00034	2.2	25.2	0.087
D90E	0.2	251	0.00079	0.6	32.3	0.018
D174A	2.9	58	0.05	0.4	375	0.0011
D174E	3.7	47	0.079	1.7	175	0.0097
D176A	1.0	67	0.015	1.2	237	0.0051
D176E	3.9	43	0.091	5.1	140	0.036

Table 4.3. Kinetic parameters of WbiP and mutants.

UDP-beads binding assay

To further investigate the involvement of the $^{88}DSD^{90}$ motif in the binding of the sugar donor, we carried out a UDP-beads binding assay with purified wild-type and mutant enzymes (Figure 4.6). Wild-type WbiP binds efficiently to UDP-beads. The binding was substantially decreased in the presence of 10 mM of UDP, indicating that the binding of WbiP to the beads is specific. D174A and D176A mutants have similar binding efficiency as compared to the wild-type. On the other hand, D88A and D90A mutants have remarkably reduced binding, suggesting that D88 and D90 are involved in the binding of sugar donor UDP-Gal. These results are consistent with the kinetic studies.

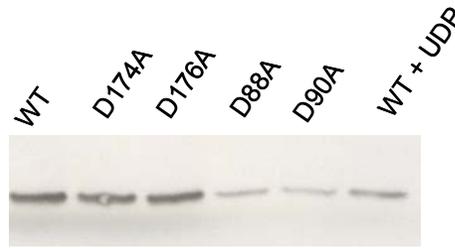


Figure 4.6. UDP-beads binding assay.

4.1.4 Discussion

T-antigen has been recognized as an important tumor antigen that is exploited for the clinical development of carbohydrate-based anti-cancer vaccines. Various chemical strategies have been developed to synthesize T-antigen containing glycopeptides [150]. On the other hand, several eukaryotic glycosyltransferases have been identified to make T-antigens in different species. Ju et al. [151] cloned and identified a human core-1 β 1,3-galactosyltransferase and showed that the enzyme can generate the core-1 structure from a synthetic glycopeptide with GalNAc linked to a Thr residue. Subsequently, a rat core-1 β 1,3-galactosyltransferase was cloned [152], characterized and shown to glycosylate GalNAc- α -O-Ph to provide T-antigen mimicry. More recently, the corresponding enzymes in *Drosophila* and *C. elegans* were also reported [153, 154]. Several bacterial β 1,3-galactosyltransferases have also been studied, all of which have endogenous acceptor specificity, different from the GalNAc α structure present in the T-antigen. We previously reported identification of a β 1,3-galactosyltransferase, WbnJ, involved in the *E. coli* O86 O-polysaccharide biosynthesis, and showed that WbnJ could synthesize a T-antigen analog [135]. However, the detailed biochemical and substrate specificity study

has not been carried out. Nonetheless, it was to our knowledge the first bacterial glycosyltransferase to make a T-antigen structure. In our current study, we characterized a second β 1,3-galactosyltransferase WbiP from *E. coli* O127. A detailed substrate specificity study showed that WbiP can be used for the facile synthesis of a variety of T-antigen mimics.

Most of the glycosyltransferases in the GTA family possess one or more DXD motifs, which have been shown to be involved in the coordination of a divalent metal ion that is required for the binding of sugar nucleotides in the active site [155, 156]. Thus, characterization of the DXD motif is an important step to understanding glycosyltransferase mechanisms. Recently, both structural and biochemical studies revealed that if more than one DXD motif is present, not all the DXD motifs are functionally equivalent. A study in the analysis of two DXD motifs in human xylosyltransferase I [157] showed that the ³¹⁴DED³¹⁶ motif did not participate in catalysis, while the ⁷⁴⁵DWD⁷⁴⁷ motif was clearly involved in the binding of the sugar nucleotide. A well studied galactosyltransferase LgtC [158] from *Neisseria meningitides* contains four DXD motifs. The crystal structure of LgtC revealed that only two DXD motifs were located within the active site. However, of these two, one of them indeed acts to coordinate the Mn²⁺ ion with the phosphate group of UDP, whereas the other is involved in the binding of the sugar acceptor. In our study, WbiP possesses two DXD motifs, ⁸⁸DSD⁹⁰ and ¹⁷⁴DYD¹⁷⁶. Mutation studies on these two motifs indicated that they may play different roles in enzyme catalysis. Mutation of aspartic acid into alanine at the ⁸⁸DSD⁹⁰ position completely abolished the enzyme activity, while marginal activity was still retained with the same mutations at ¹⁷⁴DYD¹⁷⁶ position, suggesting that ⁸⁸DSD⁹⁰ is

more essential and may directly involve in the binding of sugar nucleotides in the active site. In the kinetic studies, substantially reduced affinity for sugar donor UDP-Gal was observed for mutants D88E and D90E, while the affinity for sugar acceptors was similar to the wild-type. This observation further supports the role of ⁸⁸DSD⁹⁰ in the coordination of UDP-Gal binding. On the other hand, similar sugar donor affinity but reduced sugar acceptor affinity were observed for the mutants of ¹⁷⁴DYD¹⁷⁶, suggesting that the ¹⁷⁴DYD¹⁷⁶ motif may be involved in the binding of the sugar acceptor. A third biochemical study, UDP-beads binding assay, adds another piece of experimental evidence to the differential roles of these two DXD motifs, in that reduced binding efficiency was observed in ⁸⁸DSD⁹⁰ mutants but not in ¹⁷⁴DYD¹⁷⁶ mutants. It is worth noting that a number of bacterial β 1,3-galactosyltransferases annotated in the database appear to possess these two conserved DXD motifs, suggesting a common glycosyl transfer mechanism. The elucidation of the functional roles of these two DXD motifs requires determination of the enzyme 3-D structure.

4.2 Bacterial homologue of human blood group A transferase

4.2.1 Introduction

The histo-blood group ABO antigens are defined by carbohydrate sequences mainly found on the surface of red blood cells.[159] These antigens are involved in a variety of important biological processes.[160-163] For example, they are important in blood transfusion, organ transplantation, cell development, differentiation and oncogenesis. Recent research has shown that these antigens are critical recognition

elements of novoviruses to infect human cells.[164, 165] Moreover, the molecular mimicry of ABO antigens found on the bacterial cell surface has been implicated in bacterial pathogenicity and modulation of human immune responses.[166]

The structures of the ABO carbohydrate epitopes are defined as GalNAc- α 1,3(Fuc- α 1,2)-Gal β (A antigen), Gal- α 1,3(Fuc- α 1,2)-Gal β (B antigen) and Fuc- α 1,2-Gal β (O antigen). A and B antigens are formed by the transfer of GalNAc and Gal to O antigen disaccharide by human blood group transferase GTA and GTB, respectively [167]. The Palcic group and the Evans group solved the crystal structures of GTA and GTB, and uncovered the basis of the donor specificity conferred by these two highly homologous enzymes, which differ only in four amino acid residues [168]. A number of bacteria express carbohydrate mimicry of blood group antigens on cell surface as a means to evade host immune clearance [39, 169, 170]. The presence of blood group antigens mimicry on bacteria prompts the speculation of human anti-blood group antibody generation by enteric bacterial stimulation. Previously, our group reported the identification of a bacterial homolog of human blood group B transferase from *Escherichia coli* O86, and used it to synthesize blood group B antigen type 3 carbohydrate structure in vitro [135]. In this study, we identified the first bacterial homolog of human blood group A transferase from *Helicobacter mustelae*. This enzyme possesses flexible substrate specificity and can be used to synthesize 5 different types of blood group A antigens with nearly equal efficiency. The broad specificity of this enzyme was also demonstrated by the *in vivo* conversion of blood group B antigen mimicry to blood group A antigen mimicry on *E. coli* O86 cell surface.

4.2.2 Experimental Methods

Bacteria strains, plasmids and reagents

Helicobacter mustelae type strain ATCC 43772 was purchased from the American Type Culture Collection (Rockville, MD, USA). *E. coli* O86:K62:H2 strain was kindly provided by W. F. Vann from Department of Health and Human Service, Food and Drug Administration. *E. coli* competent cell DH5 α [*lacZ* Δ *M15 hsdR recA*] was from Gibco-BRL Life Technology. *E. coli* competent cell BL21 (DE3) [*F*⁻ *ompT hsdS_B(r_B⁻m_B⁻) gal dcm* (DE3)], and pET28a vector were from Novagen Inc. (Madison, WI). HiTrap Chelating Ni HP column was obtained from GE Healthcare (Piscataway, NJ). Anti-A and anti-B monoclonal antibodies were purchased from Abcam Ltd (Cambridge, UK). All other chemicals and solvents were from Sigma-Aldrich.

Cloning, expression and purification of BgtA

The *bgtA* gene was amplified by PCR from the *Helicobacter mustelae* type strain ATCC 43772 genomic DNA. The primers with restriction sites underlined for amplification were as follows: P1 5' GACTGTCCCATGGGGACAGTCAACCGCACAAAAC (*NcoI*) /P2 5' GACTGTCTCGAGGCGTCCTAGAAAGACGCTT (*XhoI*). The DNA fragments obtained were digested with corresponding restriction enzymes and inserted into pET28a vector linearized by the same restriction enzymes to form recombinant vector pET28a-btgA. The recombinant plasmids were confirmed by restriction mapping and sequencing. The correct constructs were subsequently transformed into *E. coli* BL21 (DE3) for protein expression. *E. coli* BL21 (DE3) harboring the recombinant plasmid was grown in 1L LB medium at 37 °C. Isopropyl-1-thio- β -D-galactospyranoside (IPTG) was added to a

final concentration of 0.5 mM, and the expression was allowed to proceed for 15 h at 20 °C. Cells were harvested and stored at –80 °C until needed. In protein purification, the cell pellet was suspended in binding buffer (20 mM Tris-HCl pH 7.5, 0.5 M NaCl and 10 mM imidazole) and disrupted by sonication on ice. After centrifugation, the lysate was loaded onto a HiTrap Chelating Ni HP column. After extensive washing with the washing buffer (20 mM Tris-HCl pH 7.5, 0.5 M NaCl and 50 mM imidazole), the protein was eluted with the elution buffer (20 mM Tris-HCl pH 8.1, 0.5 M NaCl and 500 mM imidazole). Protein eluted from the column was analyzed by SDS–PAGE and Western-blotting probed with anti-His antibody.

BgtA activity assay

For the enzymatic assays, 1 unit of enzyme activity was defined as the amount of enzyme required to transform 1 μmol of sugar donor to acceptor per minute at 37°C. Protein concentration is determined by Bradford method. Enzyme assays were performed at 37 °C for 1 hr in a final volume of 50 μL containing 20 mM Tris-HCl, pH 7.5, 5 mM MnCl₂, 0.3 mM radioactively labeled sugar donor UDP-D-[6-³H]GalNAc (15,000 cpm), 5 mM trisaccharide acceptor Fucα1,2-Galβ1,3-GlcNAcβ-(CH₂)₂N₃, and variable amount of enzymes. Acceptor was omitted as the blank control. The reaction was terminated by adding 100 μL ice cold 0.1 M EDTA. Dowex 1× 8-200 chloride anion exchange resin was then added in a water suspension (0.8 mL, v/v = 1/1). After centrifugation, supernatant was collected in a 20 mL plastic vial and ScintiVerse BD (8 mL) was added. The vial was vortexed thoroughly before the radioactivity of the mixture was counted in a liquid scintillation counter (Beckmann LS-3801 counter).

Kinetic studies of BgtA

Reactions were performed at 37 °C for 40 min in a reaction buffer containing 20 mM Tris-HCl (pH 7.5), 5 mM MnCl₂, 0.3 mM of UDP-D-[6-³H]-GalNAc and 20 µg of BgtA. To determine apparent K_m values, the concentrations acceptors were varied as follows: 5 mM, 10 mM, 20 mM and 40 mM. To determine the K_m value for UDP-GalNAc, 3 µM of UDP-D-[6-³H]-Gal was supplemented with different amounts of cold UDP-GalNAc to achieve varying final concentrations (0.08, 0.1, 0.3 and 0.6 mM), with fixed GalNAc concentration at 10 mM. The parameters K_m and V_{max} were obtained by Lineweaver-Burk plots of substrate concentration-initial velocity.

Enzymatic Synthesis of Blood group A antigen structures and product characterization

One pot reaction was conducted for two days at room temperature in a final volume of 1.0 mL containing 20 mM Tris-HCl (pH 7.5), 5 mM MnCl₂, 30 mM UDP-GlcNAc, and 10 mM of fucosylated acceptors, 50 µg WbgU and 40 µg BgtA. The progress of the reaction was monitored by thin-layer chromatography [i-PrOH/H₂O/NH₄OH= 7:3:2 (v/v/v)] conducted on Baker Si250F silica gel TLC plates. Products were visualized by staining solution (anisaldehyde/MeOH/-H₂SO₄=1:15:2 (v/v/v)). After complete conversion of acceptors, reactions were quenched by brief boiling and centrifugation (12,000g, 5 min). The supernatant mixture was purified by gel filtration chromatography Bio-Gel P2 (Bio-Rad, CA). The desired fractions were pooled, lyophilized, and stored at -20 °C.

Mass spectra (ESI) were run in the negative mode at the mass spectrometry facility at the Ohio state University. All the NMR data were recorded at 20 °C on a

Brucker 600 M instrument equipped with Cryo probe from CCIC of The Ohio State University. For sample 1, 2 and 3, the recorded spectrum included: ^1H , phase sensitive COSY, HSQC, HMBC, TOCSY. The chemical shifts of all ^1H and ^{13}C atoms, as well as most of the coupling constants, were fully assigned. The regio-chemistry and stereo-chemistry of the glycosylation bonds were confirmed by the coupling constants of anomeric protons and the 3-bond HMBC coupling signals which occurred between two directly tethered mono-saccharide units. See later sections for are the assignment of NMR data and the corresponding spectra.

Remodeling *E. coli* O86 LPS by *bgtA* complementation

The *wbnI* gene that encodes an α 1,3-Gal transferase was functionally disrupted by the replacement with chloramphenicol acetyltransferase (CAT) gene as described. The *bgtA* gene in pET-28a vector was introduced into the Δ *wbnI* mutant and induced expression with 0.1 mM IPTG. Lipopolysaccharide (LPS) was extracted according to the published protocol, and analyzed by SDS-PAGE followed by silver staining.

The LPSs from *E. coli* strains (K12, O86 wild-type, Δ *wbnI*, and Δ *wbnI/bgtA*) were extracted according to the standard phenol-water protocol and dissolved in PBS buffer pH=7.5. The LPSs were subsequently coated onto flat-bottomed microtiter plates for 48 hr at 4°C at the concentration of 10 $\mu\text{g}/\text{mL}$. Then 2% Polyvinylpyrrolidone (PVP) in PBS buffer was used to block the nonspecific binding for 3 hr. The anti-B and anti-A antibodies were used in series dilutions. ELISA end-point titers were determined as described [106], with the secondary antibody as goat anti-mouse IgM conjugated to HRP

(1:1000) (Sigma). The peroxidase substrate (3,3',5,5'-tetramethylbenzidine, Sigma) was used to develop the signal, which was monitored at 405 nm.

CE/IS-CID-MS analysis was done at the Institute of Biological Sciences, National Research Council of Canada. Capillary electrophoresis was performed using a Prince CE system (Prince Technologies, The Netherlands). The CE system was coupled to an API 3000 mass spectrometer (Applied Biosystems/Sciex, Concord, Canada) via a microspray interface. Sample separation and mass spectra acquisition of polysaccharides were carried out as described.

4.2.3 Results and Discussion

Helicobacter mustelae type strains isolated from experimental animals have been employed as animal models for *Helicobacter* infection, ulceration and gastric cancer. Study showed that anti-blood group A antibody reacted strongly with lipopolysaccharide (LPS) of *H. mustelae*. Structural analysis [171] reveals that the outer core of its LPS expresses type 1 blood group A antigen structure, GalNAc- α 1,3(Fuc- α 1,2)-Gal- β 1,3-GlcNAc β . Using the sequence of human blood group A transferase (GTA) as a probe to BLAST against the recently completed *H. mustelae* strain ATCC 43772 genome (http://sanger.ac.uk/Projects/H_mustelae/), we identified an open reading frame with 40% sequence homology with GTA/GTB. We renamed this open reading frame as BgtA (for Bacterial GTA). The *bgtA* gene was amplified from genomic DNA of *H. mustelae* strain ATCC 43772 and cloned into pET28a vector. Expression in *E. coli* BL21 (DE3) and subsequent purification with Ni-affinity column gave homogenous target protein with a C-terminal His₆-tag. Radioactive-based enzymatic assay with UDP-GalNAc[6-³H] and

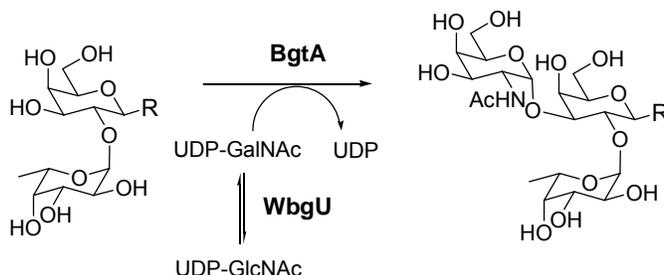
trisaccharide substrate Fuc- α 1,2-Gal- β 1,3-GlcNAc β -(CH₂)N₃ confirmed that bgtA encodes a N-acetyl-galactosaminyl transferase.

Substrate Types	K_M (mM)	k_{cat}/K_M (s ⁻¹ mM ⁻¹)	Product Yield
 Type I	5.5 ± 0.4	4.3	96%
 Type II	2.1 ± 0.2	5.7	98%
 Type III/IV	4.2 ± 0.3	3.9	95%
 Type V	11.3 ± 1.4	1.7	91%

Table 4.4. Substrate specificity, kinetics and yields of BgtA-catalyzed reactions.

Assay of BgtA toward a panel of acceptor substrates revealed that BgtA has dedicated specificity toward Fuc- α 1,2-Gal β -R carbohydrate sequences. On the other hand, BgtA has relaxed specificity (Table 4.4) regardless of the structures appended to the Gal residue, which collectively make up the five types of naturally existing blood group O antigen core structures. This result underscores the merit of BgtA for the

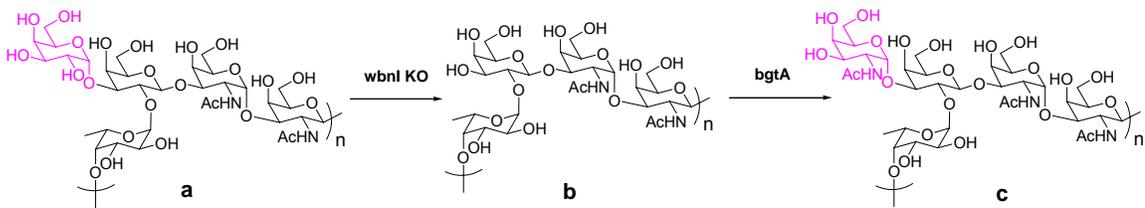
enzymatic synthesis of different A antigen structures. The apparent kinetic parameters (Table 4.4) obtained showed high catalytic efficiency (k_{cat}/K_M) of BgtA toward the five substrates, further confirming the broad specificity of BgtA. In addition, the efficiency and flexibility of BgtA were also demonstrated in the synthesis of different blood group A type structures *in vitro* using a one-pot two-enzyme system (Scheme 4.1). In this system, previously characterized UDP-GlcNAc/GalNAc C-4 epimerase WbgU was used to generate sugar donor UDP-GalNAc from inexpensive alternative donor UDP-GlcNAc. More than 90% yield in each reaction was achieved. The structures of synthesized A antigens were analyzed by NMR and MS (Appendix).



Scheme 4.1. One-pot two enzyme system for synthesis of blood group A antigens.

The enzymatic property of BgtA to synthesize blood group A structures was exploited to remodel the bacterial cell surface blood group B antigen mimicry into A antigen mimicry (Scheme 4.2). *E. coli* O86 has a blood group B-like epitope in the O-polysaccharide of cell surface LPS. Our previous study identified *wbnI* gene encoding a protein homolog of GTB to synthesize blood group B-like structure (type III) in this bacterium[135]. To *in vivo* convert this structure into blood group A structure, we firstly

disrupted *wbnI* gene from *E. coli* O86 chromosome as described before [47]. LPS analysis with SDS-PAGE coupled with silver staining revealed the reduction of the amount of high molecular weight polysaccharides, and the shift of ladder-like band pattern to lower molecular weight region (Figure 4.7). The NMR and MS analysis of the structure of polysaccharides from $\Delta wbnI$ mutant in our previous study[47] revealed a blood group O antigen-like structure, consistent with the *in vitro* activity of WbnI. We then transformed *bgtA* containing pET-28a recombinant vector into $\Delta wbnI$ mutant cells for trans-complementation. LPS isolated from the *bgtA* complemented cells displayed similar spacing and migration distance as seen in the wild-type LPS, but the amount of high molecular weight polysaccharides is similar to that in $\Delta wbnI$ mutant cells. The possible explanation of this observation is that BgtA is able to modify the repeating unit; however, the modified structure is less tolerant by the polymerization step, probably due to the fairly strict substrate specificity of sugar polymerase Wzy.



Scheme 4.2. In vivo conversion of B antigen to A antigen.

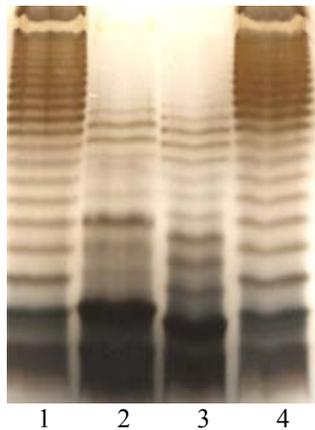


Figure 4.7. LPS analysis of different *E. coli* strains: 1. *E. coli* O86:H2 wild-type; 2. $\Delta wbnI$ mutant with pET28a-bgtA complementation; 3. $\Delta wbnI$ mutants (notice the shift of bands to lower molecular weight compared to the wild-type LPS); 4. $\Delta wbnI$ mutant with pET15b-wbnI complementation.

In order to demonstrate the conversion of blood group B antigen to A antigen by BgtA *in vivo* complementation, we carried out a LPS ELISA assay. As shown in Figure 4.8, the negative control, *E. coli* K12 displays minimal basal reactivity toward both anti-A and anti-B antibodies. Wild-type *E. coli* O86 LPS shows more than 1,000 fold increase of anti-B reactivity, consistent with previous studies [39]. *E. coli* O86 LPS also shows 50 fold increase of anti-A reactivity compared to the background, probably due to the cross-reactivity resulted from the similar structures of B and A antigens. The LPS from $\Delta wbnI$ mutant cells have reactivity slightly higher than the background. The LPS from *bgtA* complemented cells, however, demonstrates strikingly different reactivity toward anti-A and anti-B antibodies. The anti-A activity increases by 800 fold compared to the negative control, while the anti-B activity decreases to the similar level of $\Delta wbnI$ mutant cells.

Thus this result demonstrates that remodeling of cell surface polysaccharides via gene replacement can result in the switching of immuno reactivity of bacterial cells.

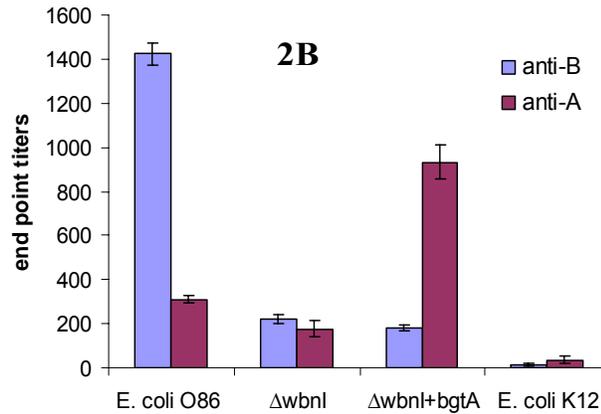


Figure 4.8. LPS ELISA with blood group antibodies.

To further confirm the remodeling of *E. coli* O86 polysaccharides by BtgA, we used CE-MS technique to determine the polysaccharide structures. LPSs from wild-type, *ΔwbnI* mutant and *bgtA* complement cells were extracted and treated to remove O-acetylated fatty acids as previously described before subjected to MS analysis. Figure S2 (Appendix) shows the full spectra of O-deacetylated LPS from wild-type and corresponding mutant cells. In each spectrum, lipid-A associated ions were detected at m/z 855.9, 468.6 and 388.5, corresponding to the consecutive loss of a H_3PO_4 , a N-acylated glucosamine and the second H_3PO_4 group, respectively. The information on the repeating unit could be easily obtained by selecting the ion at m/z 204.2 as a precursor to perform precursor ion scan experiments, which corresponds to the HexNAc oxonium ion. In the precursor-treated spectra, the major peaks were identified and the corresponding

sugar sequences (including different fragmentations) were assigned (Figure 4.9). The peaks showing the full repeating unit structures were circled in red. In the spectrum of *bgtA* complemented mutant (2C-c), the ion of m/z 919.0 represents the sugar sequence $\text{Fuc}_1\text{Hex}_1\text{HexNAC}_3$, consistent with the proposed modified repeating unit structure. Similarly, the repeating units of the wild-type and $\Delta wbnI$ mutant were identified based on m/z 878.0 (2C-a, $\text{Fuc}_1\text{Hex}_2\text{HexNAC}_2$) and 716.0 (2C-b, $\text{Fuc}_1\text{Hex}_1\text{HexNAC}_2$), respectively. In addition, the mass peaks corresponding to the full repeating unit structures were unique in the respective cells, suggesting that the polysaccharide modification is homogenous.

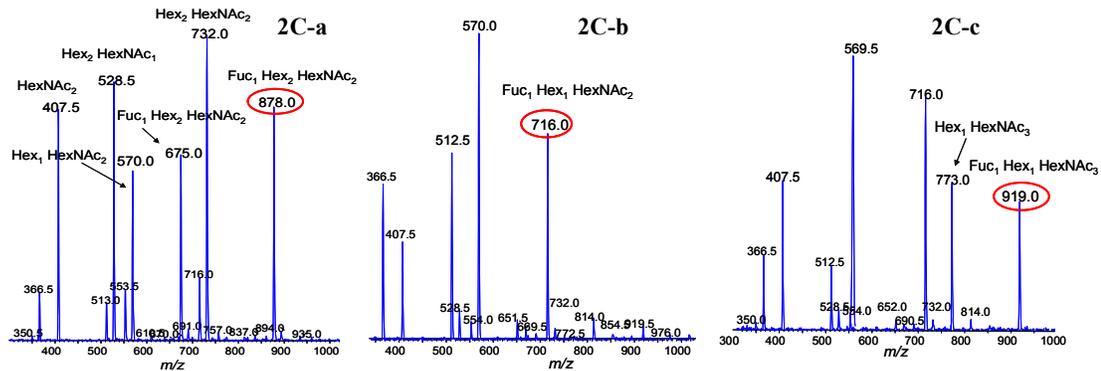


Figure 4.9. MS analysis of polysaccharide structures.

In conclusion, we report in this study the identification of the first bacterial homolog of human blood group A transferase. Biochemical characterization of this enzyme (BgtA) demonstrated its high activity and flexible specificity toward the five types of naturally existing blood group core structures, illustrating its unparalleled versatility compared to its mammalian counterpart GTA to synthesize blood group A

structures. Moreover, the applicability of BgtA was demonstrated in engineering of bacterial cell surface polysaccharides, resulting in the switching the immuno-reactivity of the cell. Further detailed mechanistic and structural study will provide clues to understand bacterium-host interactions through the adaptation and evolution of cell surface polysaccharide mimicry.

CHAPTER 5

CONCLUSIONS AND PERSPECTIVES

Bacterial polysaccharides contain enormously rich structural and functional information. To dissect the role of polysaccharides in bacterial pathogenicity, bacterium-host interactions and human immune regulation, chemical approaches have begun to emerge as powerful tools in these studies. In this thesis, chemo-enzymatic synthesis of oligosaccharide substrates has provided homogenous materials to probe the molecular mechanism of polysaccharide biosynthesis, a daunting challenge for traditional genetic studies. Such a well defined chemical reconstitution system revealed unprecedented insight into different key steps of biosynthesis, and several unexpected new findings. This system opens up an exciting opportunity to chemically synthesis of structurally defined polysaccharides/polysaccharide-based glycoconjugates *in vitro*, a technique that will greatly facilitate current development of polysaccharide-based vaccines. Introduction of modifications into polysaccharides via an engineered promiscuous sugar nucleotide biosynthetic pathway in chapter 3 has permitted *in vivo* generation of structurally novel polysaccharides. This technique has the advantages of obtaining homogenous modifications under native cellular context, a challenge that still remains in the current metabolic labeling of mammalian glycans. Modified polysaccharides provide a unique

structure pool for elucidation of their structure-function relationships. Incorporation of bio-orthogonal functional groups has permitted further derivatization and cell surface labeling. Future directions in this project will be to demonstrate such technique in a more pathologically relevant bacterium, such as *Helicobacter pylori*, in which fucose plays an important role in adhesion and bacterial infection. In addition, as this study only probed the modification of C-6 position of fucose, modifications at other positions can also be systematically investigated with the same technique. Such inventory will provide a complete understanding of the role of fucose in polysaccharides. Furthermore, besides fucose, this technique can be, in principle, applied to other monosaccharides, such as galactose, N-acetyl galactosamine, N-acetyl glucosamine and sialic acid. The work in chapter 4 involves functional identification of two novel bacterial glycosyltransferases and the application in synthesizing human related oligosaccharide structures. Future work will be to investigate the substrate specificity of BgtA in greater details. Comparison of BgtA with its mammalian counterpart GTA will reveal the extricate specificity and subtle differences in terms of substrate recognition. Furthermore, three dimensional structure of BgtA will provide additional dimension for elucidating evolutionary relationship of blood group transferases.

REFERENCES

1. Ohtsubo, K. and J.D. Marth, *Glycosylation in cellular mechanisms of health and disease*. Cell, 2006. **126**(5): p. 855-867.
2. Comstock, L.E. and D.L. Kasper, *Bacterial glycans: key mediators of diverse host immune responses*. Cell, 2006. **126**(5): p. 847-850.
3. Varki, A., *Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins*. Nature, 2007. **446**(7139): p. 1023-1029.
4. Helenius, A. and M. Aebi, *Roles of N-linked glycans in the endoplasmic reticulum*. Annu. Rev. Biochem. 2004. **73**: p. 1019-1049.
5. Dube, D.H. and C.R. Bertozzi, *Glycans in cancer and inflammation - potential for therapeutics and diagnostics*. Nat. Rev. Drug Discovery, 2005. **4**(6): p. 477-488.
6. Freeze, H.H., *Genetic defects in the human glycome*. Nat. Rev. Genet. FIELD Full Journal Title:Nature Reviews Genetics, 2006. **7**(7): p. 537-551.
7. Hart, G.W., M.P. Housley, and C. Slawson, *Cycling of O-linked beta -N-acetylglucosamine on nucleocytoplasmic proteins*. Nature, 2007. **446**(7139): p. 1017-1022.
8. Brooks, S.A., et al., *Altered glycosylation of proteins in cancer: what is the potential for new anti-tumour strategies*. Anti-Cancer Agents Med. Chem. 2008. **8**(1): p. 2-21.
9. Gilewski, T.A., et al., *Immunization of High-Risk Breast Cancer Patients with Clustered sTn-KLH Conjugate plus the Immunologic Adjuvant QS-21*. Clin. Cancer Res. 2007. **13**(10): p. 2977-2985.
10. Krug, L.M., et al., *Vaccination of Patients with Small-Cell Lung Cancer with Synthetic Fucosyl GM-1 Conjugated to Keyhole Limpet Hemocyanin*. Clin. Cancer Res. 2004. **10**(18, Pt. 1): p. 6094-6100.
11. Krug, L.M., et al., *Vaccination of Small Cell Lung Cancer Patients with Polysialic Acid or N-Propionylated Polysialic Acid Conjugated to Keyhole Limpet Hemocyanin*. Clin. Cancer Res. 2004. **10**(3): p. 916-923.

12. Livingston, P.O. and G. Ragupathi, *Cancer vaccines targeting carbohydrate antigens*. Hum. Vaccines 2006. **2**(3): p. 137-143.
13. Springer, G.F., *Immunoreactive T and Tn epitopes in cancer diagnosis, prognosis, and immunotherapy*. J. Mol. Med. 1997. **75**(8): p. 594-602.
14. Livingston, P.O., et al., *Selection of GM2, fucosyl GM1, globo H and polysialic acid as targets on small cell lung cancers for antibody mediated immunotherapy*. Cancer Immunol. Immunother. 2005. **54**(10): p. 1018-1025.
15. Chang, W.-W., et al., *Expression of Globo H and SSEA3 in breast cancer stem cells and the involvement of fucosyl transferases 1 and 2 in Globo H synthesis*. Proc. Natl. Acad. Sci. U. S. A. 2008. **105**(33): p. 11667-11672.
16. Wang, C.-C., et al., *Glycan microarray of Globo H and related structures for quantitative analysis of breast cancer*. Proc. Natl. Acad. Sci. U. S. A. 2008. **105**(33): p. 11661-11666.
17. Gilewski, T., et al., *Immunization of metastatic breast cancer patients with a fully synthetic globo H conjugate: a phase I trial*. Proc. Natl. Acad. Sci. U. S. A. 2001. **98**(6): p. 3270-3275.
18. Ragupathi, G., et al., *Preparation and Evaluation of Unimolecular Pentavalent and Hexavalent Antigenic Constructs Targeting Prostate and Breast Cancer: A Synthetic Route to Anticancer Vaccine Candidates*. J. Am. Chem. Soc. 2006. **128**(8): p. 2715-2725.
19. Watkins, W.M., *Blood group substances*. Science 1966. **152**(3719): p. 172-81.
20. Morgan, W.T.J. and W.M. Watkins, *Unravelling the biochemical basis of blood group ABO and Lewis antigenic specificity*. Glycoconjugate J. 2001. **17**(7/8/9): p. 501-530.
21. Watkins, W.M., *Biosynthesis of glycoproteins and molecular basis of antigenic specificity in the ABO, H and Lewis blood-group systems*. New Compr. Biochem. 1995. **29a**: p. 313-90.
22. Whitfield, C., *Biosynthesis and assembly of capsular polysaccharides in Escherichia coli*. Annu. Rev. Biochem. 2006. **75**: p. 39-68.
23. Guo, H., et al., *Current understanding on biosynthesis of microbial polysaccharides*. Curr. Top. Med. Chem. 2008. **8**(2): p. 141-151.
24. Raetz, C.R.H. and C. Whitfield, *Lipopolysaccharide endotoxins*. Annu. Rev. Biochem. 2002. **71**: p. 635-700.

25. Jann, K. and B. Jann, *Capsules of Escherichia coli, expression and biological significance*. Can J Microbiol 1992. **38**(7): p. 705-10.
26. Paton, J.C. and J.K. Morona, *Pneumococcal capsular polysaccharides: biosynthesis and regulation*. Mol. Biol. Streptococci 2007: p. 119-139.
27. Moreau, M. and D. Schulz, *Polysaccharide based vaccines for the prevention of pneumococcal infections*. J. Carbohydr. Chem. 2000. **19**(4 & 5): p. 419-434.
28. Hallemeersch, I., S. De Baets, and E.J. Vandamme, *Exopolysaccharides of lactic acid bacteria*. Biopolymers 2002. **5**: p. 407-429.
29. Myszka, K. and K. Czaczyk, *The role of microbial exopolysaccharides in food technology*. Żywnosc 2004. **11**(4): p. 18-29.
30. Leathers, T.D. and G.L. Cote, *Biofilm formation by exopolysaccharide mutants of Leuconostoc mesenteroides strain NRRL B-1355*. Appl. Microbiol. Biotechnol. 2008. **78**(6): p. 1025-1031.
31. Sutherland, I.W., *Biofilm exopolysaccharides: a strong and sticky framework*. Microbiology (Reading, U. K.) 2001. **147**(1): p. 3-9.
32. Valvano, M.A., *Export of O-specific lipopolysaccharide*. Front. Biosci. 2003. **8**: p. S452-S471.
33. Lehrer, J., et al., *Functional characterization and membrane topology of Escherichia coli WecA, a sugar-phosphate transferase initiating the biosynthesis of enterobacterial common antigen and O-antigen lipopolysaccharide*. J. Bacteriol. 2007. **189**(7): p. 2618-2628.
34. Price, N.P. and F.A. Momany, *Modeling bacterial UDP-HexNAc:polyprenol-P HexNAc-1-P transferases*. Glycobiology 2005. **15**(9): p. 29R-42R.
35. Marolda, C.L., J. Vicarioli, and M.A. Valvano, *Wzx proteins involved in biosynthesis of O antigen function in association with the first sugar of the O-specific lipopolysaccharide subunit*. Microbiology (Reading, U. K.) 2004. **150**(12): p. 4095-4105.
36. Carter, J.A., et al., *O-antigen model chain length in Shigella flexneri 2a is growth-regulated through RfaH-mediated transcriptional control of the wzy gene*. Microbiology (Reading, U. K.) 2007. **153**(10): p. 3499-3507.
37. Daniels, C., C. Vindurampulle, and R. Morona, *Overexpression and topology of the Shigella flexneri O-antigen polymerase (Rfc/Wzy)*. Mol. Microbiol. 1998. **28**(6): p. 1211-1222.

38. Marolda, C.L., et al., *Functional analysis of predicted coiled-coil regions in the Escherichia coli K-12 O-antigen polysaccharide chain length determinant Wzz*. J. Bacteriol. 2008. **190**(6): p. 2128-2137.
39. Guo, H., et al., *Molecular analysis of the O-antigen gene cluster of Escherichia coli O86:B7 and characterization of the chain length determinant gene (wzz)*. Appl. Environ. Microbiol. 2005. **71**(12): p. 7995-8001.
40. Young, J. and I.B. Holland, *ABC transporters: bacterial exporters-revisited five years on*. Biochim. Biophys. Acta, Biomembr. 1999. **1461**(2): p. 177-200.
41. Fath, M.J. and R. Kolter, *ABC transporters: bacterial exporters*. Microbiol. Rev. 1993. **57**(4): p. 995-1017.
42. Abeyrathne, P.D. and J.S. Lam, *WaaL of Pseudomonas aeruginosa utilizes ATP in in vitro ligation of O antigen onto lipid A-core*. Mol. Microbiol. 2007. **65**(5): p. 1345-1359.
43. Kaniuk, N.A., E. Vinogradov, and C. Whitfield, *Investigation of the Structural Requirements in the Lipopolysaccharide Core Acceptor for Ligation of O Antigens in the Genus Salmonella: WaaL "Ligase" is not the Sole Determinant of Acceptor Specificity*. J. Biol. Chem. 2004. **279**(35): p. 36470-36480.
44. Wu, T., et al., *Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. 2006. **103**(31): p. 11754-11759.
45. Whitfield, C. and K. Larue, *Stop and go: regulation of chain length in the biosynthesis of bacterial polysaccharides*. Nat. Struct. Mol. Biol. 2008. **15**(2): p. 121-123.
46. Liu, D., R.A. Cole, and P.R. Reeves, *An O-antigen processing function for Wzx (RfbX): a promising candidate for O-unit flippase*. J. Bacteriol. 1996. **178**(7): p. 2102-07.
47. Yi, W., et al., *Formation of a new O-polysaccharide in Escherichia coli O86 via disruption of a glycosyltransferase gene involved in O-unit assembly*. Carbohydr. Res. 2006. **341**(13): p. 2254-2260.
48. Yoshida, Y., et al., *Carbohydrate engineering of the recognition motifs in streptococcal co-aggregation receptor polysaccharides*. Mol. Microbiol. 2005. **58**(1): p. 244-256.
49. Morona, R., L. Van Den Bosch, and C. Daniels, *Evaluation of Wzz/MPA1/MPA2 proteins based on the presence of coiled-coil regions*. Microbiology (Reading, U. K.) 2000. **146**(1): p. 1-4.

50. Morona, R., L. Van den Bosch, and P.A. Manning, *Molecular, genetic, and topological characterization of O-antigen chain length regulation in Shigella flexneri*. J. Bacteriol. 1995. **177**(4): p. 1059-68.
51. Bastin, D.A., et al., *Repeat unit polysaccharides of bacteria: A model for polymerization resembling that of ribosomes and fatty acid synthetase, with a novel mechanism for determining chain length*. Mol. Microbiol. 1993. **7**(5): p. 725-34.
52. Shaner, N.C., P.A. Steinbach, and R.Y. Tsien, *A guide to choosing fluorescent proteins*. Nat. Methods 2005. **2**(12): p. 905-909.
53. Miyawaki, A. and R.Y. Tsien, *Monitoring protein conformations and interactions by fluorescence resonance energy transfer between mutants of green fluorescent protein*. Methods Enzymol. 2000. **327**(Applications of Chimeric Genes and Hybrid Proteins, Pt. B): p. 472-500.
54. Lowe, J.B. and J.D. Marth, *A genetic approach to mammalian glycan function*. Annu. Rev. Biochem. 2003. **72**: p. 643-691.
55. Bertozzi, C.R. and L.L. Kiessling, *Chemical glycobiology*. Science 2001. **291**(5512): p. 2357-2364.
56. Gerber-Lemaire, S. and L. Juillerat-Jeanneret, *Glycosylation pathways as drug targets for cancer: glycosidase inhibitors*. Mini-Rev. Med. Chem. 2006. **6**(9): p. 1043-1052.
57. Esko, J.D., *Natural and synthetic inhibitors of glycosylation*. Essent. Glycobiol. 1999: p. 609-623.
58. Rawat, M., et al., *Neuroactive Chondroitin Sulfate Glycomimetics*. J. Am. Chem. Soc. 2008. **130**(10): p. 2959-2961.
59. Gama, C.I., et al., *Sulfation patterns of glycosaminoglycans encode molecular recognition and activity*. Nat. Chem. Biol. 2006. **2**(9): p. 467-473.
60. Epperson, T.K., et al., *Noncovalent association of P-selectin glycoprotein ligand-1 and minimal determinants for binding to P-selectin*. J. Biol. Chem. 2000. **275**(11): p. 7839-7853.
61. Leppanen, A., et al., *A novel glycosulfopeptide binds to P-selectin and inhibits leukocyte adhesion to P-selectin*. J. Biol. Chem. 1999. **274**(35): p. 24838-24848.
62. Prescher, J.A. and C.R. Bertozzi, *Chemistry in living systems*. Nat. Chem. Biol. 2005. **1**(1): p. 13-21.

63. Prescher, J.A. and C.R. Bertozzi, *Chemical technologies for probing glycans*. Cell 2006. **126**(5): p. 851-854.
64. Mootoo, D.R., et al., *Armed and disarmed n-pentenyl glycosides in saccharide couplings leading to oligosaccharides*. J. Am. Chem. Soc. 1988. **110**(16): p. 5583-4.
65. Sears, P. and C.-H. Wong, *Toward automated synthesis of oligosaccharides and glycoproteins*. Science 2001. **291**(5512): p. 2344-2350.
66. Burkhardt, F., et al., *Synthesis of the Globo H hexasaccharide using the programmable reactivity-based one-pot strategy*. Angew. Chem., Int. Ed. 2001. **40**(7): p. 1274-1277.
67. Murata, T. and T. Usui, *Enzymatic synthesis of oligosaccharides and neoglycoconjugates*. Biosci., Biotechnol., Biochem. 2006. **70**(5): p. 1049-1059.
68. Zhang, J., et al., *Enzymatic synthesis of oligosaccharides*. Carbohydr.-Based Drug Discovery 2003. **1**: p. 137-167.
69. Weijers, C.A.G.M., M.C.R. Franssen, and G.M. Visser, *Glycosyltransferase-catalyzed synthesis of bioactive oligosaccharides*. Biotechnol. Adv. 2008. **26**(5): p. 436-456.
70. Ohrlein, R., *Glycosyltransferase-catalyzed synthesis of non-natural oligosaccharides*. Top. Curr. Chem. 1999. **200**(Biocatalysis: From Discovery to Application): p. 227-254.
71. Chan, N.W.C., et al., *The biosynthesis of Lewis X in Helicobacter pylori*. Glycobiology 1995. **5**(7): p. 683-8.
72. Gilbert, M., et al., *Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384. Identification of the glycosyltransferase genes, enzymatic synthesis of model compounds, and characterization of nanomole amounts by 600-MHz H and C NMR analysis*. J. Biol. Chem. 2000. **275**(6): p. 3896-3906.
73. Shao, J., et al., *Overexpression and biochemical characterization of beta -1,3-N-acetylgalactosaminyltransferase LgtD from Haemophilus influenzae strain Rd*. Biochem. Biophys. Res. Commun. 2002. **295**(1): p. 1-8.
74. Wang, G., et al., *Novel Helicobacter pylori alpha 1,2-fucosyltransferase, a key enzyme in the synthesis of Lewis antigens*. Microbiology (Reading, U. K.) 1999. **145**(11): p. 3245-3253.
75. Zhang, J., et al., *Synthesis of galactose-containing oligosaccharides through superbeads and superbug approaches: substrate recognition along different*

- biosynthetic pathways*. Methods Enzymol. 2003. **362**(Recognition of Carbohydrates in Biological Systems, Part A): p. 106-124.
76. Chen, X., et al., *Reassembled biosynthetic pathway for large-scale carbohydrate synthesis: alpha -gal epitope producing "superbug"*. ChemBioChem 2002. **3**(1): p. 47-53.
 77. Shoda, S.-i., M. Kohri, and A. Kobayashi, *Glycosidase-catalyzed synthesis of oligosaccharides through intermediate analogue substrates*. Biocatal.: Chem. Biol. 2005: p. 83-98.
 78. Singh, S., M. Scigelova, and D.H.G. Crout, *Glycosidase-catalyzed synthesis of alpha -galactosyl epitopes important in xenotransplantation and toxin binding using the alpha -galactosidase from Penicillium multicolor*. Chem. Commun. 1999(20): p. 2065-2066.
 79. Nilsson, K.G.I., *Synthesis with glycosidases*. Front. Nat. Prod. Res. 1996. **1**(Modern Methods in Carbohydrate Synthesis): p. 518-547.
 80. Nilsson, K.G.I., *Glycosidase-catalyzed synthesis of di- and trisaccharide derivatives related to antigens involved in the hyperacute rejection of xenotransplants*. Tetrahedron Lett. 1997. **38**(1): p. 133-136.
 81. Hancock, S.M., M.D. Vaughan, and S.G. Withers, *Engineering of glycosidases and glycosyltransferases*. Curr. Opin. Chem. Biol. 2006. **10**(5): p. 509-519.
 82. Shaikh, F.A. and S.G. Withers, *Teaching old enzymes new tricks: engineering and evolution of glycosidases and glycosyl transferases for improved glycoside synthesis*. Biochem. Cell Biol. 2008. **86**(2): p. 169-177.
 83. Nihira, T., *Evolution of the oligosaccharide synthesis by glycosynthases*. Trends Glycosci. Glycotechnol. 2007. **19**(105): p. 49-50.
 84. Vaughan, M.D., et al., *Glycosynthase-mediated synthesis of glycosphingolipids*. J. Am. Chem. Soc. 2006. **128**(19): p. 6300-6301.
 85. Mullegger, J., et al., *Glycosylation of a neoglycoprotein by using glycosynthase and thioglycoligase approaches: the generation of a thioglycoprotein*. Angew. Chem., Int. Ed. 2006. **45**(16): p. 2585-2588.
 86. Brown, J.R., M.M. Fuster, and J.D. Esko, *Glycoside primers and inhibitors of glycosylation*. Carbohydr.-Based Drug Discovery 2003. **2**: p. 883-898.
 87. Murrey, H.E., et al., *Protein fucosylation regulates synapsin Ia/Ib expression and neuronal morphology in primary hippocampal neurons*. Proc. Natl. Acad. Sci. U. S. A. 2006. **103**(1): p. 21-26.

88. Chang, P.V., et al., *Imaging Cell Surface Glycans with Bioorthogonal Chemical Reporters*. J. Am. Chem. Soc. 2007. **129**(27): p. 8400-8401.
89. Campbell, C.T., S.G. Sampathkumar, and K.J. Yarema, *Metabolic oligosaccharide engineering: perspectives, applications, and future directions*. Mol. BioSyst. 2007. **3**(3): p. 187-194.
90. Sampathkumar, S.-G., et al., *Metabolic installation of thiols into sialic acid modulates adhesion and stem cell biology*. Nat. Chem. Biol. 2006. **2**(3): p. 149-152.
91. Luchansky, S.J., S. Goon, and C.R. Bertozzi, *Expanding the diversity of unnatural cell-surface sialic acids*. ChemBioChem 2004. **5**(3): p. 371-374.
92. Griffin, R.J., *The medicinal chemistry of the azido group*. Prog. Med. Chem. 1994. **31**: p. 121-232.
93. Koehn, M. and R. Breinbauer, *The Staudinger ligation - A gift to chemical biology*. Angew. Chem., Int. Ed. 2004. **43**(24): p. 3106-3116.
94. Wang, Q., et al., *Bioconjugation by copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition*. J. Am. Chem. Soc. 2003. **125**(11): p. 3192-3193.
95. Prescher, J.A., D.H. Dube, and C.R. Bertozzi, *Chemical remodelling of cell surfaces in living animals*. Nature 2004. **430**(7002): p. 873-877.
96. Saxon, E. and C.R. Bertozzi, *Cell surface engineering by a modified Staudinger reaction*. Science 2000. **287**(5460): p. 2007-2010.
97. Rostovtsev, V.V., et al., *A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes*. Angew. Chem., Int. Ed. 2002. **41**(14): p. 2596-2599.
98. Jessani, N. and B.F. Cravatt, *Chemical strategies for activity-based proteomics*. Chem. Biol. 2007. **1**: p. 403-426.
99. Seo, T.S., et al., *Photocleavable fluorescent nucleotides for DNA sequencing on a chip constructed by site-specific coupling chemistry*. Proc. Natl. Acad. Sci. U. S. A. 2004. **101**(15): p. 5488-5493.
100. Agard, N.J., et al., *A comparative study of bioorthogonal reactions with azides*. ACS Chem. Biol. 2006. **1**(10): p. 644-648.
101. Baskin, J.M., et al., *Copper-free click chemistry for dynamic in vivo imaging*. Proc. Natl. Acad. Sci. U. S. A. 2007. **104**(43): p. 16793-16797.

102. Westphal, O., K. Jann, and K. Himmelspach, *Chemistry and immunochemistry of bacterial lipopolysaccharides as cell wall antigens and endotoxins*. Prog. Allergy 1983. **33**(Host Parasite Relat. Gram-Negat. Infect.): p. 9-39.
103. Springer, G.F. and R.E. Horton, *Blood group isoantibody stimulation in man by feeding blood group-active bacteria*. J Clin Invest 1969. **48**(7): p. 1280-91.
104. Springer, G.F., R.E. Horton, and M. Forbes, *Origin of antihuman blood group B agglutinins in germfree chicks*. Ann N Y Acad Sci 1959. **78**: p. 272-5.
105. Kochibe, N. and S. Iseki, *Immunochemical studies on bacterial blood group substances. V. Structures of B determinant groups*. Jpn J Microbiol 1968. **12**(4): p. 403-11.
106. Andersson, M., et al., *Structural studies of the O-antigenic polysaccharide of Escherichia coli O86, which possesses blood-group B activity*. Carbohydr Res 1989. **185**(2): p. 211-23.
107. Otter, A., et al., *Crystal state and solution conformation of the B blood group trisaccharide alpha -L-Fucp-(1->2)-[alpha -D-Galp]-(1->3)]-beta -D-Galp-OCH3*. Eur. J. Biochem. 1999. **259**(1/2): p. 295-303.
108. Lemieux, R.U., et al., *Molecular recognition. V. The binding of the B human blood group determinant by hybridoma monoclonal antibodies*. Can. J. Chem. 1985. **63**(10): p. 2664-8.
109. Azurmendi, H.F. and C.A. Bush, *Conformational studies of blood group A and blood group B oligosaccharides using NMR residual dipolar couplings*. Carbohydr. Res. 2002. **337**(10): p. 905-915.
110. Lerouge, I. and J. Vanderleyden, *O-antigen structural variation: mechanisms and possible roles in animal/plant-microbe interactions*. FEMS Microbiol. Rev. 2002. **26**(1): p. 17-47.
111. Perry, M.B. and L.L. MacLean, *Structure of the polysaccharide O-antigen of Salmonella riogrande O:40 (group R) related to blood group A activity*. Carbohydr Res 1992. **232**(1): p. 143-50.
112. Wang, L., et al., *Expression of the O antigen gene cluster is regulated by RfaH through the JUMPstart sequence*. FEMS Microbiol. Lett. 1998. **165**(1): p. 201-206.
113. Hobbs, M. and P.R. Reeves, *The JUMPstart sequence: a 39 bp element common to several polysaccharide gene clusters*. Mol. Microbiol. 1994. **12**(5): p. 855-6.

114. Reeves, P.P. and L. Wang, *Genomic organization of LPS-specific loci*. Curr. Top. Microbiol. Immunol. 2002. **264**(Pathogenicity Islands and the Evolution of Pathogenic Microbes, Volume 1): p. 109-135.
115. Bengoechea, J.A., et al., *Functional characterization of Gne (UDP-N-acetylglucosamine-4-epimerase), Wzz (chain length determinant), and Wzy (O-antigen polymerase) of Yersinia enterocolitica serotype O:8*. J. Bacteriol. 2002. **184**(15): p. 4277-4287.
116. Ishiyama, N., et al., *Crystal Structure of WbpP, a Genuine UDP-N-acetylglucosamine 4-Epimerase from Pseudomonas aeruginosa: Substrate Specificity in UDP-Hexose 4-Epimerases*. J. Biol. Chem. 2004. **279**(21): p. 22635-22642.
117. Shao, J., et al., *Sequence of Escherichia coli O128 antigen biosynthesis cluster and functional identification of an alpha -1,2-fucosyltransferase*. FEBS Lett. 2003. **553**(1-2): p. 99-103.
118. Neidhardt, F.C. and Editor, *Escherichia Coli and Salmonella Typhimurium: Cellular and Molecular Biology, Vol. 1*. 1987. 806 pp.
119. Frick, D.N., B.D. Townsend, and M.J. Bessman, *A novel GDP-mannose mannosyl hydrolase shares homology with the MutT family of enzymes*. J. Biol. Chem. 1995. **270**(41): p. 24086-91.
120. Wang, L., et al., *The O-antigen gene cluster of Escherichia coli O55:H7 and identification of a new UDP-GlcNAc C4 epimerase gene*. J. Bacteriol. 2002. **184**(10): p. 2620-2625.
121. Punta, M., et al., *Membrane protein prediction methods*. Methods 2007. **41**(4): p. 460-474.
122. Valvano, M.A., *Undecaprenyl phosphate recycling comes out of age*. Mol. Microbiol. 2007. **67**(2): p. 232-235.
123. Tatar, L.D., et al., *An Escherichia coli undecaprenyl-pyrophosphate phosphatase implicated in undecaprenyl phosphate recycling*. Microbiology (Reading, U. K.) 2007. **153**(8): p. 2518-2529.
124. Brown, S., Y.-H. Zhang, and S. Walker, *A revised pathway proposed for Staphylococcus aureus wall teichoic acid biosynthesis based on in vitro reconstitution of the intracellular steps*. Chem. Biol. 2008. **15**(1): p. 12-21.
125. Barrett, D., et al., *Analysis of Glycan Polymers Produced by Peptidoglycan Glycosyltransferases*. J. Biol. Chem. 2007. **282**(44): p. 31964-31971.

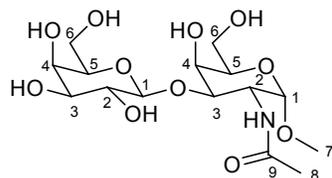
126. Wang, L., J. Xie, and P.G. Schultz, *Expanding the genetic code*. Annu. Rev. Biophys. Biomol. Struct. 2006. **35**: p. 225-249.
127. Hsu, T.-L., et al., *Alkynyl sugar analogs for the labeling and visualization of glycoconjugates in cells*. Proc. Natl. Acad. Sci. U. S. A. 2007. **104**(8): p. 2614-2619.
128. Jennings, H.J., A. Gamian, and F.E. Ashton, *N-Propionylated group B meningococcal polysaccharide mimics a unique epitope on group B Neisseria meningitidis*. J. Exp. Med. 1987. **165**(4): p. 1207-11.
129. Sadamoto, R., et al., *Cell-wall engineering of living bacteria*. J. Am. Chem. Soc. 2002. **124**(31): p. 9018-9019.
130. Goon, S., et al., *Metabolic incorporation of unnatural sialic acids into Haemophilus ducreyi lipooligosaccharides*. Proc. Natl. Acad. Sci. U. S. A. 2003. **100**(6): p. 3089-3094.
131. Ma, B., J.L. Simala-Grant, and D.E. Taylor, *Fucosylation in prokaryotes and eukaryotes*. Glycobiology 2006. **16**(12): p. 158R-184R.
132. Coyne, M.J., et al., *Human symbionts use a host-like pathway for surface fucosylation*. Science 2005. **307**(5716): p. 1778-1781.
133. Blixt, O. and N. Razi, *Chemoenzymatic synthesis of glycan libraries*. Methods Enzymol. 2006. **415**(Glycobiology): p. 137-153.
134. Antoine, T., et al., *Large-scale in vivo synthesis of the carbohydrate moieties of gangliosides GM1 and GM2 by metabolically engineered Escherichia coli*. ChemBioChem 2003. **4**(5): p. 406-412.
135. Yi, W., et al., *Escherichia coli O86 O-antigen biosynthetic gene cluster and stepwise enzymatic synthesis of human blood group B antigen tetrasaccharide*. J. Am. Chem. Soc. 2005. **127**(7): p. 2040-2041.
136. Hang, H.C. and C.R. Bertozzi, *The chemistry and biology of mucin-type O-linked glycosylation*. Bioorg. Med. Chem. 2005. **13**(17): p. 5021-5034.
137. Garner, B., et al., *Structural elucidation of the N- and O-glycans of human apolipoprotein(a): role of O-glycans in conferring protease resistance*. J. Biol. Chem. 2001. **276**(25): p. 22200-22208.
138. Ellies, L.G., et al., *Sialyltransferase ST3Gal-IV operates as a dominant modifier of hemostasis by concealing asialoglycoprotein receptor ligands*. Proc. Natl. Acad. Sci. U. S. A. 2002. **99**(15): p. 10042-10047.

139. Smith, K.A., *Interleukin-2: inception, impact, and implications*. Science 1988. **240**(4856): p. 1169-76.
140. Altschuler, Y., et al., *Clathrin-mediated endocytosis of MUC1 is modulated by its glycosylation state*. Mol. Biol. Cell 2000. **11**(3): p. 819-831.
141. Primakoff, P. and D.G. Myles, *Penetration, adhesion, and fusion in mammalian sperm-egg interaction*. Science 2002. **296**(5576): p. 2183-2185.
142. Hooper, L.V. and J.I. Gordon, *Commensal host-bacterial relationships in the gut*. Science 2001. **292**(5519): p. 1115-1118.
143. Brockhausen, I., *Pathways of O-glycan biosynthesis in cancer cells*. Biochim. Biophys. Acta, Gen. Subj. 1999. **1473**(1): p. 67-95.
144. Sotiriadis, J., et al., *Thomsen-Friedenreich (T) antigen expression increases sensitivity of natural killer cell lysis of cancer cells*. Int. J. Cancer 2004. **111**(3): p. 388-397.
145. Coon, J.S., R.S. Weinstein, and J.L. Summers, *Blood group precursor T-antigen expression in human urinary bladder carcinoma*. Am J Clin Pathol 1982. **77**(6): p. 692-9.
146. Laurent, J.C., P. Noel, and M. Faucon, *Expression of a cryptic cell surface antigen in primary cell cultures from human breast cancer*. Biomedicine 1978. **29**(8): p. 260-1.
147. Barua, D., *An outbreak of infantile diarrhoea due to Escherichia coli O127:B8*. Indian J Med Res 1962. **50**: p. 612.
148. Rivas, M., et al., *Intestinal bleeding and occlusion associated with Shiga toxin-producing Escherichia coli O127:H21*. 2000, Instituto Nacional de Enfermedades Infecciosas, ANLIS Dr. Carlos G. Malbran, Buenos Aires, Argentina.: Argentina. p. 249-52.
149. Widmalm, G. and K. Leontein, *Structural studies of the Escherichia coli O127 O-antigen polysaccharide*. Carbohydr. Res. 1993. **247**: p. 255-62.
150. Brocke, C. and H. Kunz, *Synthesis of tumor-associated glycopeptide antigens*. Bioorg. Med. Chem. 2002. **10**(10): p. 3085-3112.
151. Ju, T., et al., *Cloning and expression of human core 1 beta 1,3-galactosyltransferase*. J. Biol. Chem. 2002. **277**(1): p. 178-186.
152. Ju, T., R.D. Cummings, and W.M. Canfield, *Purification, characterization, and subunit structure of rat core 1 beta 1,3-galactosyltransferase*. J. Biol. Chem. 2002. **277**(1): p. 169-177.

153. Ju, T., Q. Zheng, and R.D. Cummings, *Identification of core 1 O-glycan T-synthase from Caenorhabditis elegans*. *Glycobiology* 2006. **16**(10): p. 947-958.
154. Muller, R., et al., *Characterization of mucin-type core-1 beta 1-3 galactosyltransferase homologous enzymes in Drosophila melanogaster*. *FEBS J.* 2005. **272**(17): p. 4295-4305.
155. Lairson, L.L., et al., *Glycosyltransferases: structures, functions, and mechanisms*. *Annu. Rev. Biochem.* 2008. **77**: p. 521-555.
156. Schuman, B., J.A. Alfaro, and S.V. Evans, *Glycosyltransferase structure and function*. *Top. Curr. Chem.* 2007. **272**(Bioactive Conformation I): p. 217-257.
157. Goetting, C., et al., *Analysis of the DXD Motifs in Human Xylosyltransferase I Required for Enzyme Activity*. *J. Biol. Chem.* 2004. **279**(41): p. 42566-42573.
158. Persson, K., et al., *Crystal structure of the retaining galactosyltransferase LgtC from Neisseria meningitidis in complex with donor and acceptor sugar analogs*. *Nat. Struct. Biol.* 2001. **8**(2): p. 166-175.
159. Milland, J. and M.S. Sandrin, *ABO blood group and related antigens, natural antibodies and transplantation*. *Tissue Antigens* 2006. **68**(6): p. 459-466.
160. Dabelsteen, E., U. Mandel, and H. Clausen, *Blood group antigens as differentiation and tumor-associated markers in oral epithelium*. *Proc. Finn. Dent. Soc.* 1988. **84**(1): p. 19-29.
161. Galili, U., *Immune Response, Accommodation, and Tolerance to Transplantation Carbohydrate Antigens*. *Transplantation* 2004. **78**(8): p. 1093-1098.
162. Kannagi, R., *Carbohydrate blood group antigens and tumor antigens*. *Immunobiol. Carbohydr.* 2003: p. 1-33.
163. King, K.E., et al., *Antibody, complement and accommodation in ABO-incompatible transplants*. *Curr. Opin. Immunol.* 2004. **16**(5): p. 545-549.
164. Gustafsson, K., et al., *The role of ABO histo-blood group antigens in viral infections*. *Trends Glycosci. Glycotechnol.* 2005. **17**(98): p. 285-294.
165. Harrington, P.R., et al., *Norovirus attachment, susceptibility and vaccine design*. *Recent Res. Dev. Virol.* 2003. **5**: p. 19-44.
166. Garratty, G., *Blood group antigens as tumor markers, parasitic/bacterial/viral receptors, and their association with immunologically important proteins*. *Immunol. Invest.* 1995. **24**(1&2): p. 213-32.

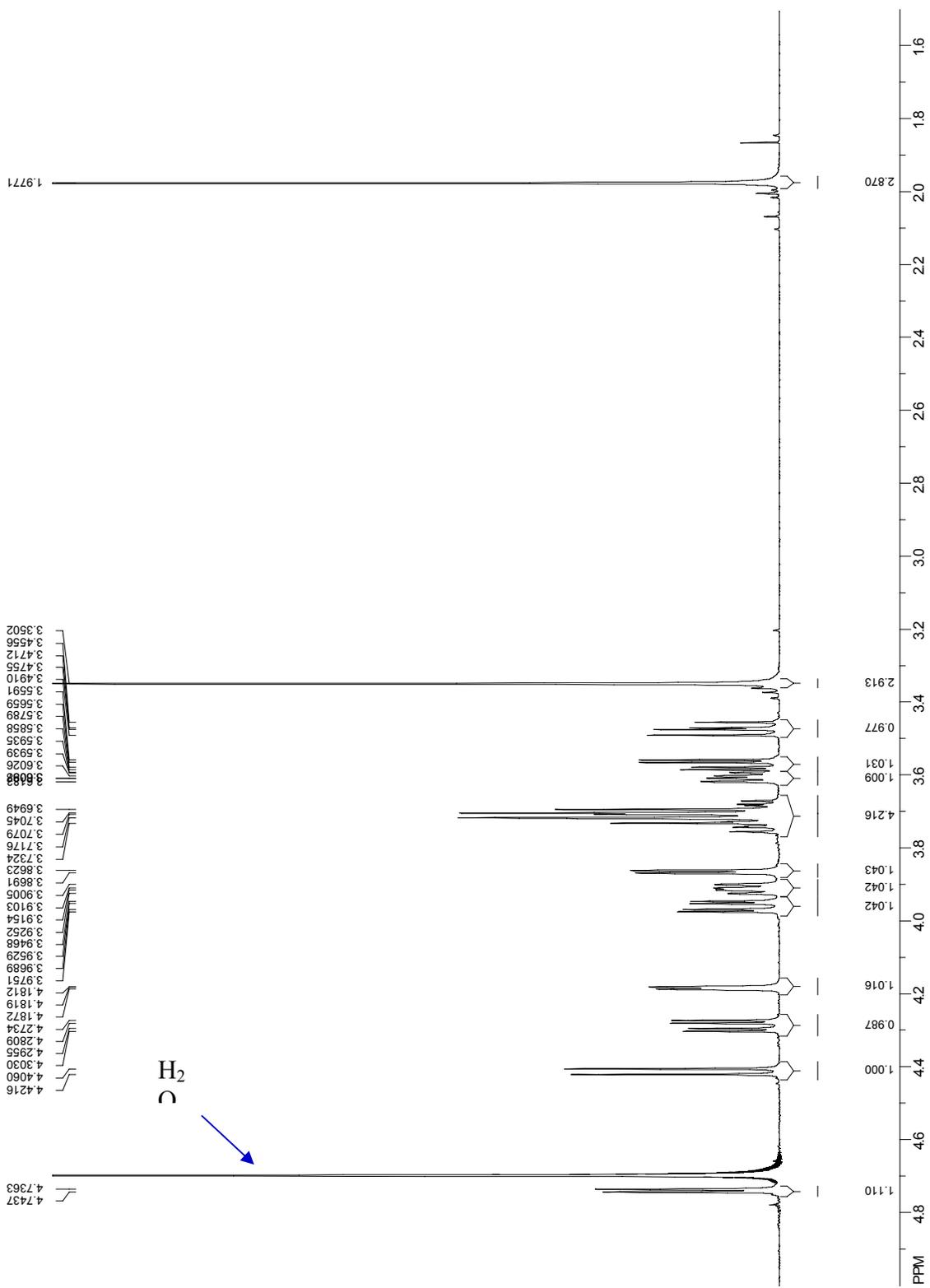
167. Rose, N.L., M.M. Palcic, and S.V. Evans, *Glycosyltransferases A and B: four critical amino acids determine blood type*. J. Chem. Educ. 2005. **82**(12): p. 1846-1853.
168. Patenaude, S.I., et al., *The structural basis for specificity in human ABO(H) blood group biosynthesis*. Nat. Struct. Biol. 2002. **9**(9): p. 685-690.
169. Yi, W., et al., *Characterization of a Bacterial beta -1,3-Galactosyltransferase with Application in the Synthesis of Tumor-Associated T-Antigen Mimics*. Biochemistry 2008. **47**(5): p. 1241-1248.
170. Li, M., et al., *Characterization of a Novel alpha 1,2-Fucosyltransferase of Escherichia coli O128:B12 and Functional Investigation of Its Common Motif*. Biochemistry 2008. **47**(1): p. 378-387.
171. Feldman, M.F., et al., *The activity of a putative polyisoprenol-linked sugar translocase (Wzx) involved in Escherichia coli O antigen assembly is independent of the chemical structure of the O repeat*. J. Biol. Chem. 1999. **274**(49): p. 35129-35138.

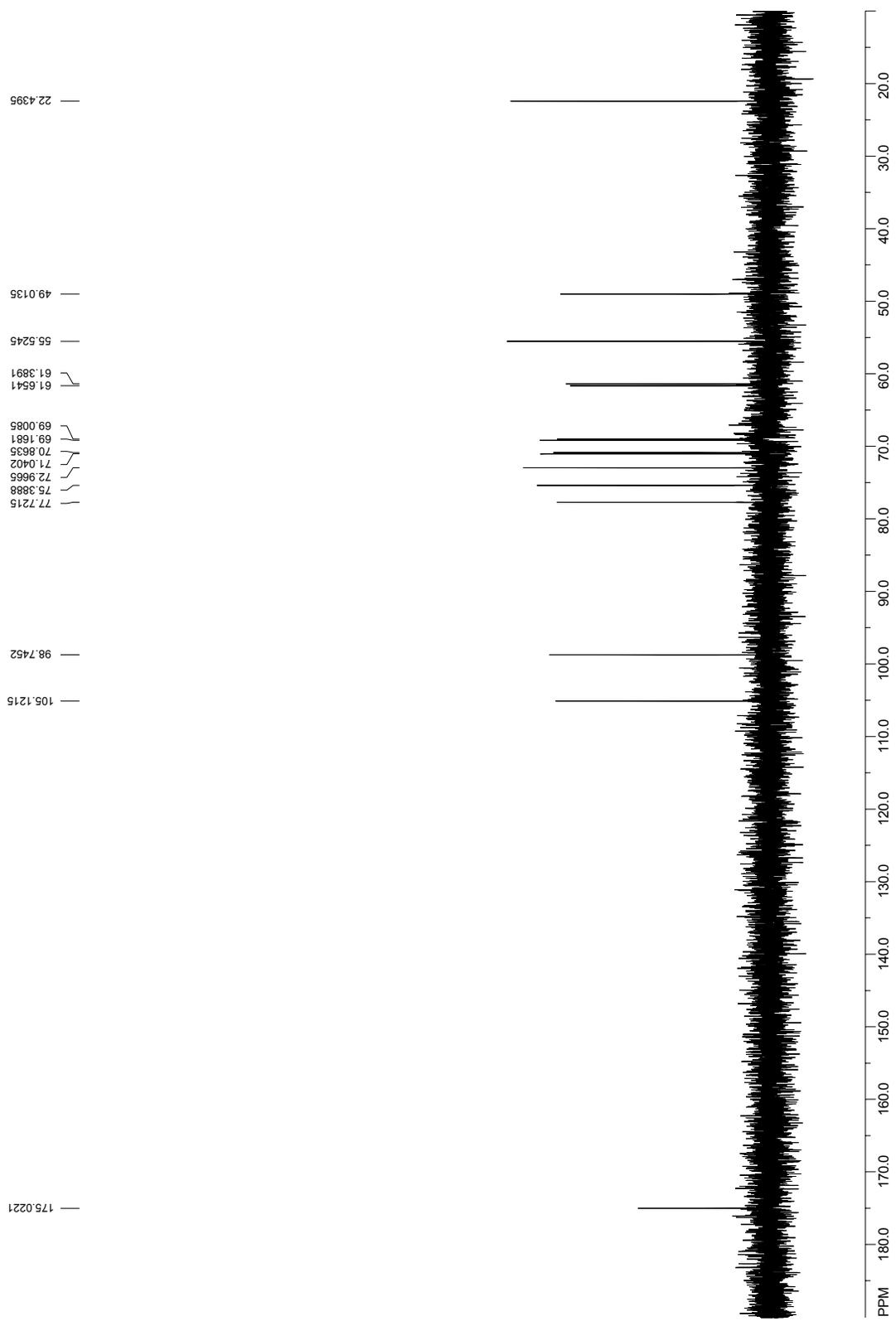
APPENDIX: NMR AND MS SPECTRA



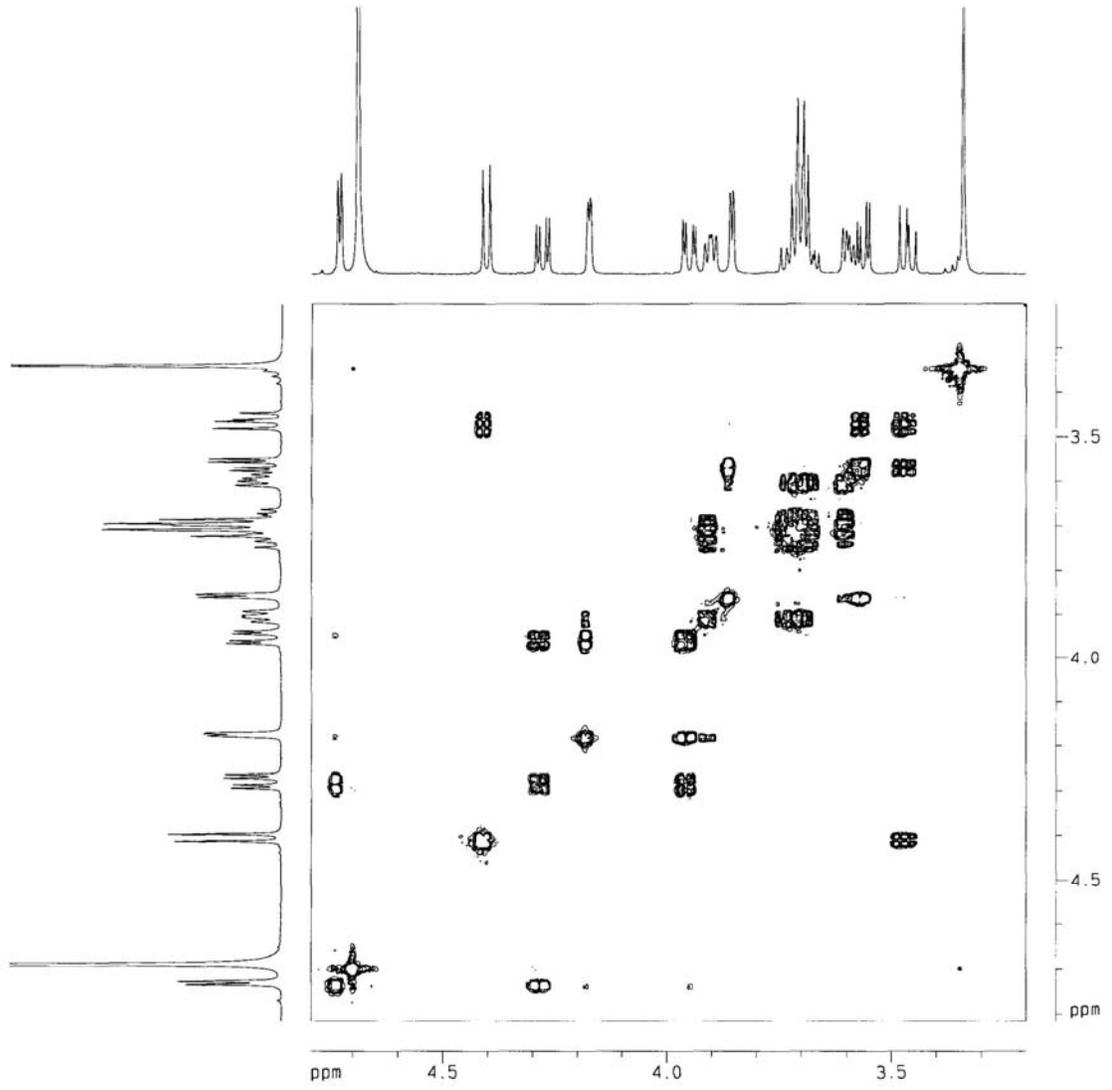
Exact Mass(M+Na)+: 420.1476
 Found: 420.1474

#	Gal		GalNAc	
	¹ H NMR	¹³ C NMR	¹ H NMR	¹³ C NMR
1	4.41, d, J = 7.8 Hz, 1H	105.12	4.74, d, J = 3.7 Hz, 1H	98.74
2	3.47, dd, J = 9.9, 7.8 Hz, 1H	70.86	4.29, dd, J = 11.1, 3.7 Hz, 1H	49.01
3	3.57, dd, J = 9.9, 3.4 Hz, 1H	72.97	3.96, dd, J = 11.1, 3.1 Hz, 1H	77.72
4	3.86, d, J = 3.3 Hz, 1H	69.01	4.18, d, J = 2.9 Hz, 1H	69.17
5	3.61, dd, J = 7.7, 4.6 Hz, 1H	75.39	3.91, dd, J = 7.3, 5.0 Hz, 1H	71.04
6	3.71, m, 2H	61.39	3.72, m, 2H	61.65
7			3.35, s, 3H, (OMe)	55.52
8			1.98, s, 3H (COCH ₃)	22.44
9			C=O	175.02

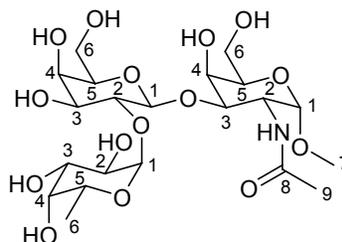




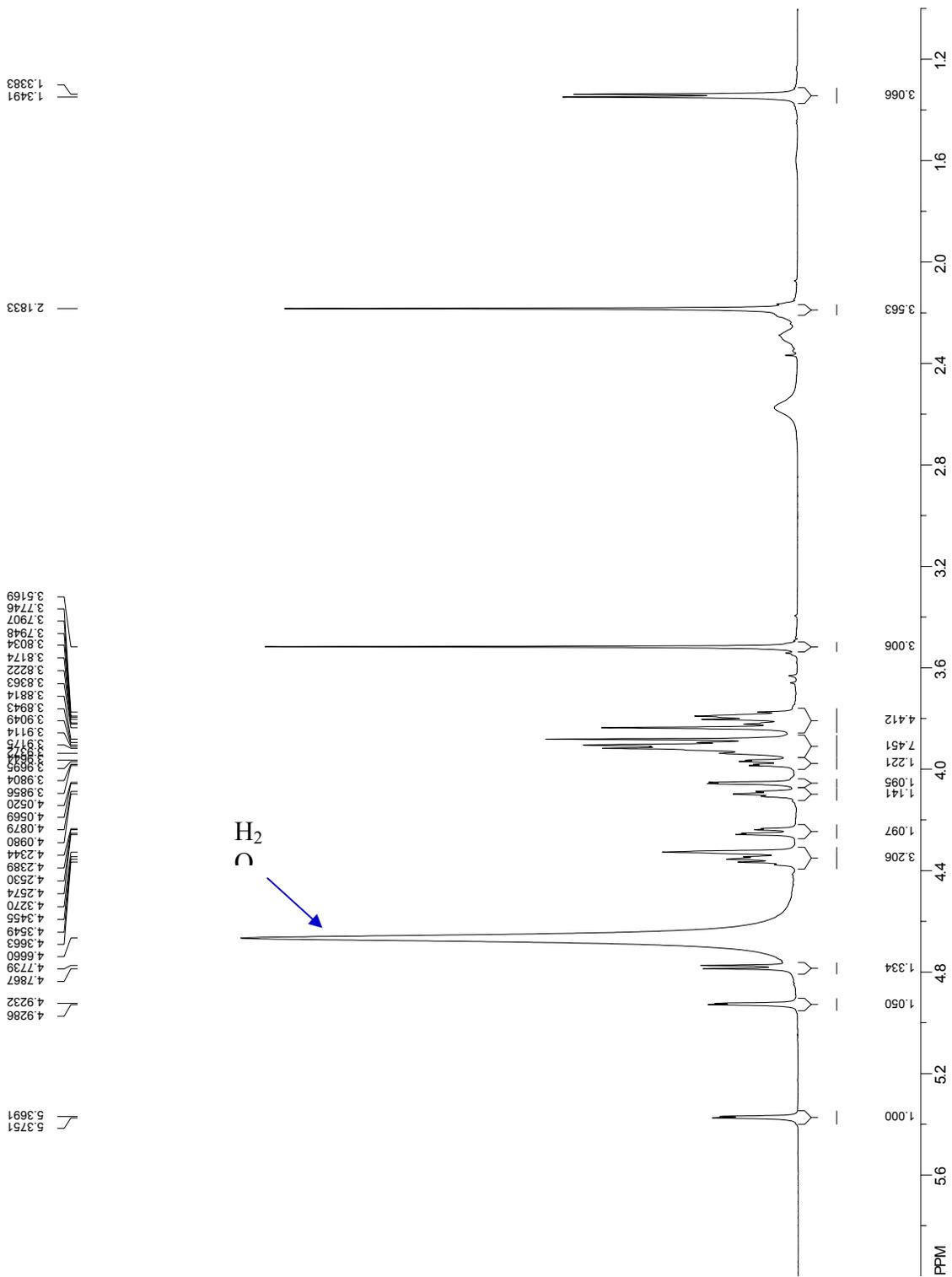
HH-COSYGS of Gal(b1-3)GalNAc in D2O at rt on

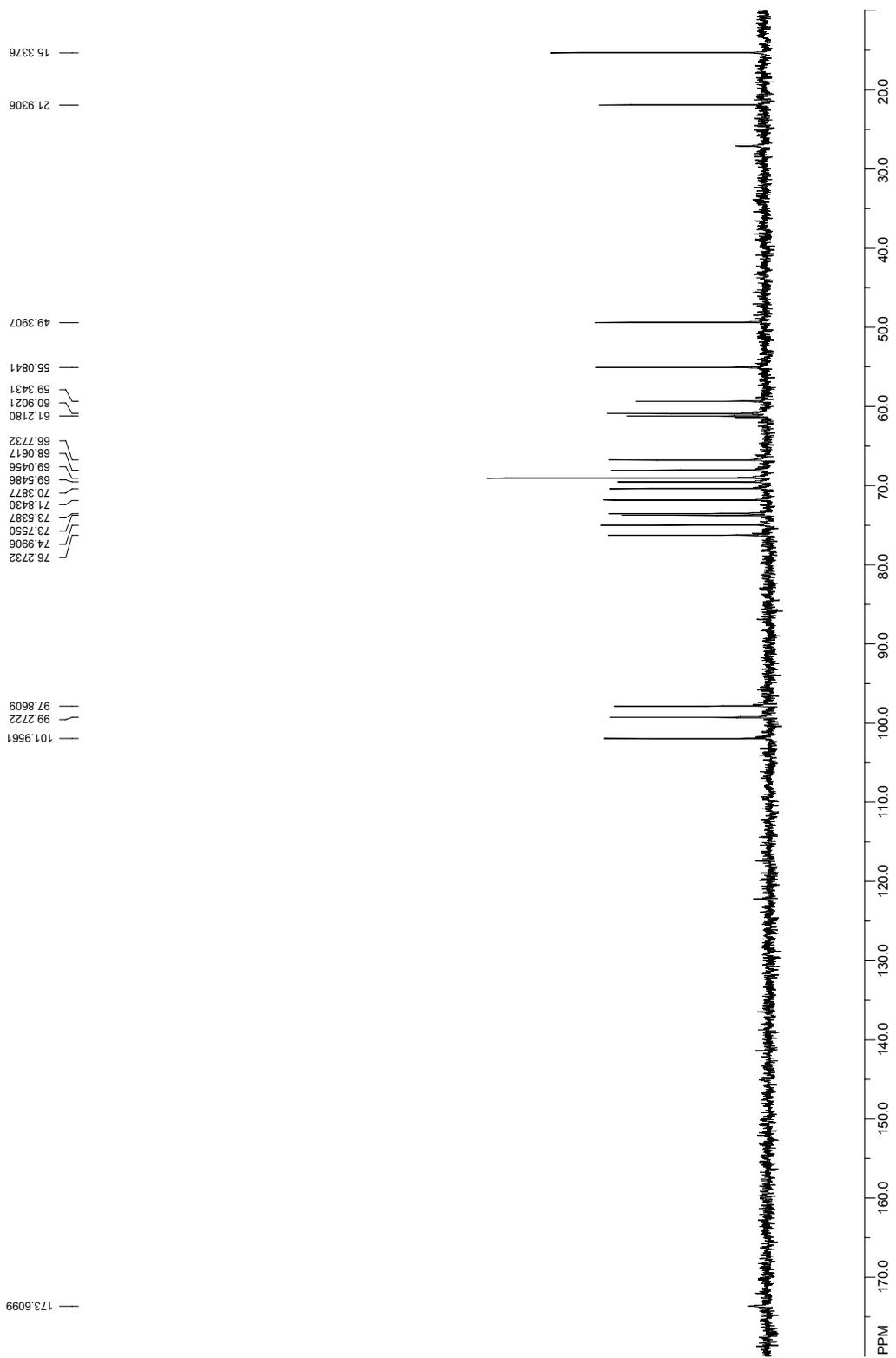


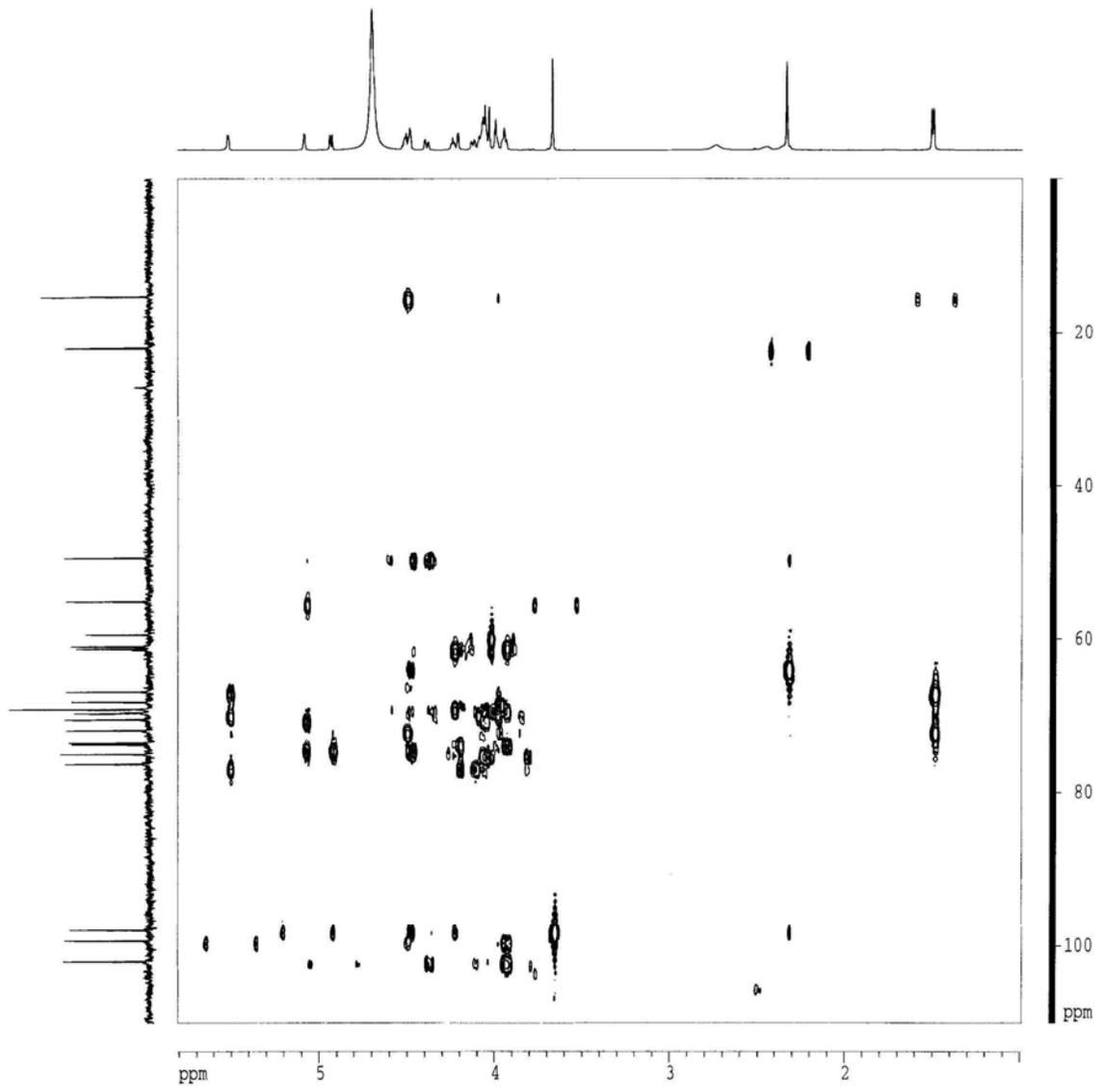
Trisaccharide α -L-Fucp(1 \rightarrow 2)- β -D-Galp-(1 \rightarrow 3)- α -GalpNAc-OMe



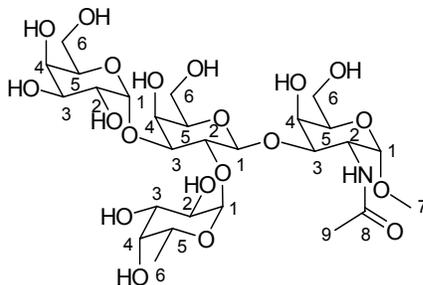
#	GalNAc		Gal		Fuc	
	^1H (δ)	^{13}C (δ)	^1H (δ)	^{13}C (δ)	^1H (δ)	^{13}C (δ)
1	4.96, d, $J = 3.2$ Hz, 1H	97.87	4.81, d, $J = 7.7$ Hz, 1H	101.96	5.41, d, $J = 3.6$ Hz, 1H	99.28
2	4.37, m, 1H	49.39	3.82, m, 1H	74.99	3.95, m, 1H	68.07
3	4.28, dd, $J = 11.1$, 2.7 Hz,	73.76	4.01, dd, $J = 9.7$, 3.1 Hz	73.54	3.86, m, 1H	71.85
4	4.36, m, 1H	69.05	3.84, m, 1H	69.55	3.81, m, 1H	76.28
5	4.13, t, $J = 5.8$ Hz, 1H	70.39	4.09, d, $J = 3.0$ Hz, 1H	69.05	4.39, m, 1H	66.78
6	3.95, m, 2H	61.22	3.96, m, 2H	60.90	1.38, d, $J = 6.5$ Hz, 3H	15.34
7	3.55, s, 3H (OCH ₃)	55.09				
8	C=O	173.66				
9	2.22, s, 3H (COCH ₃)	21.93				





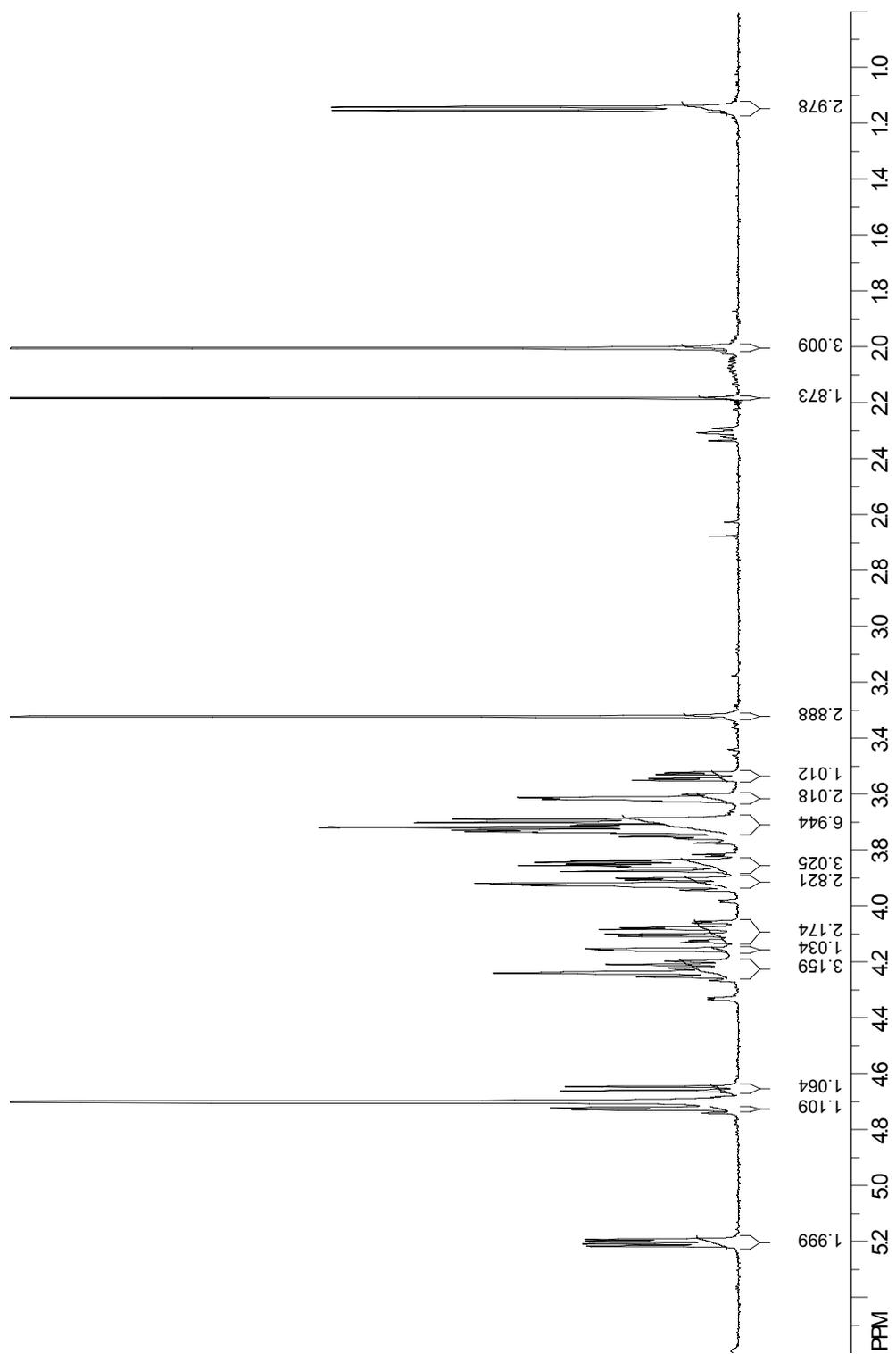


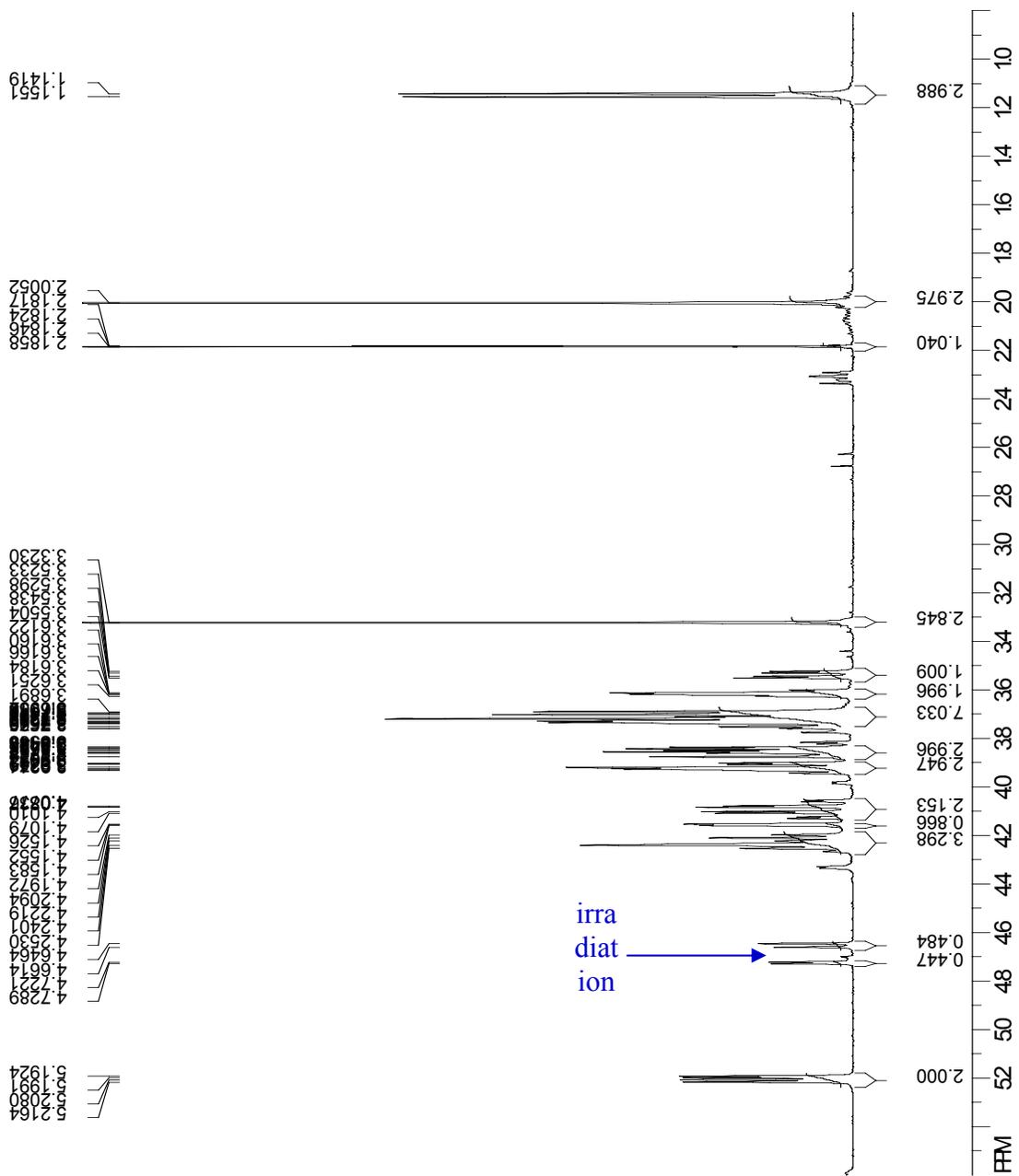
Tetrasaccharide α -L-Fucp-[α -D-Galp(1,3)](1 \rightarrow 2)- β -D-Galp-(1 \rightarrow 3)- α -GalpNAc-OMe

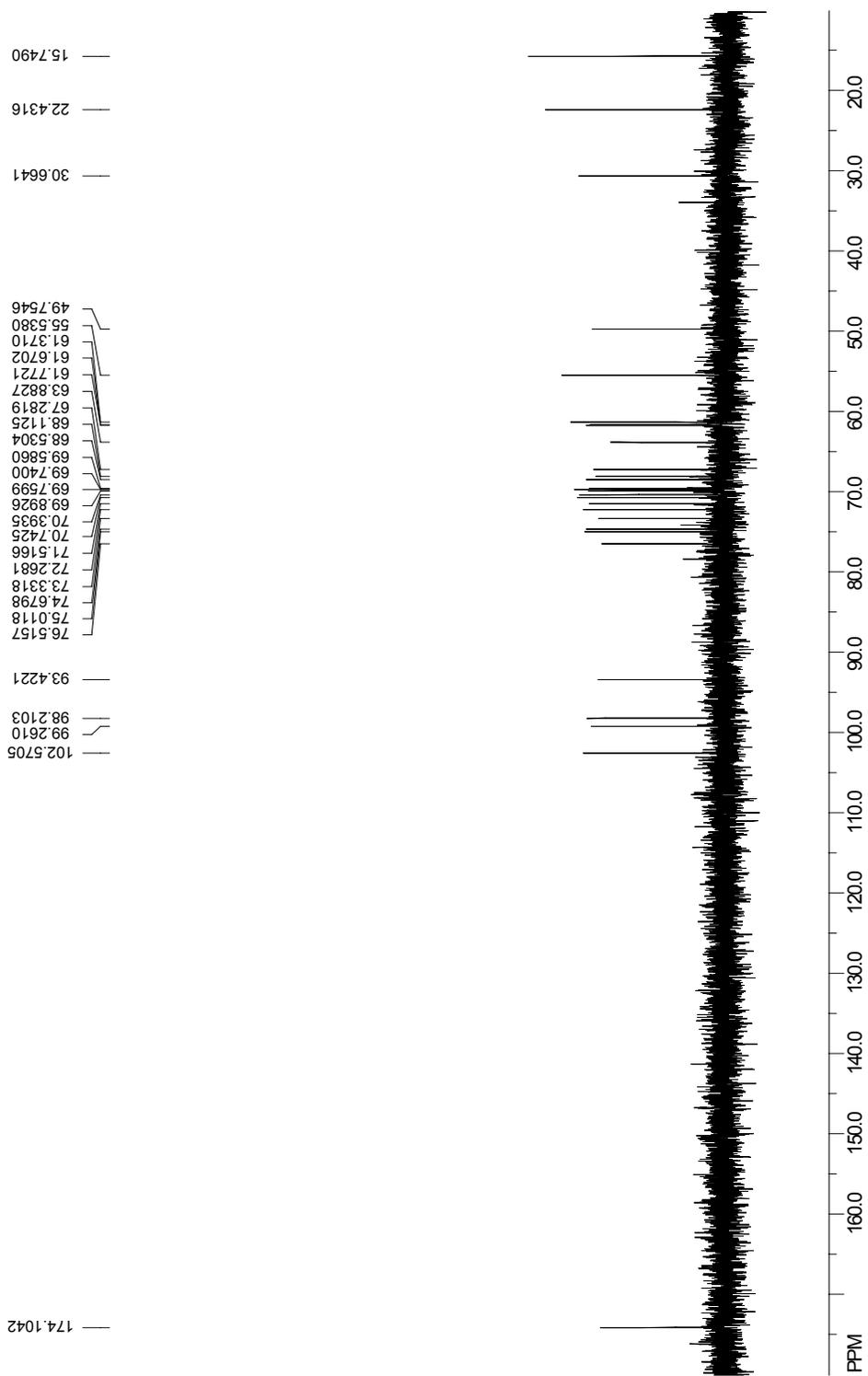


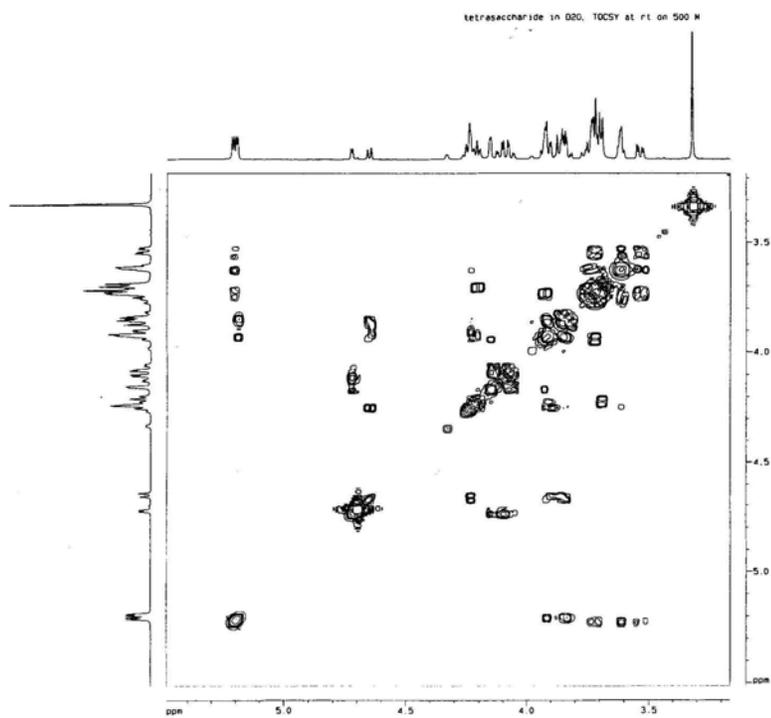
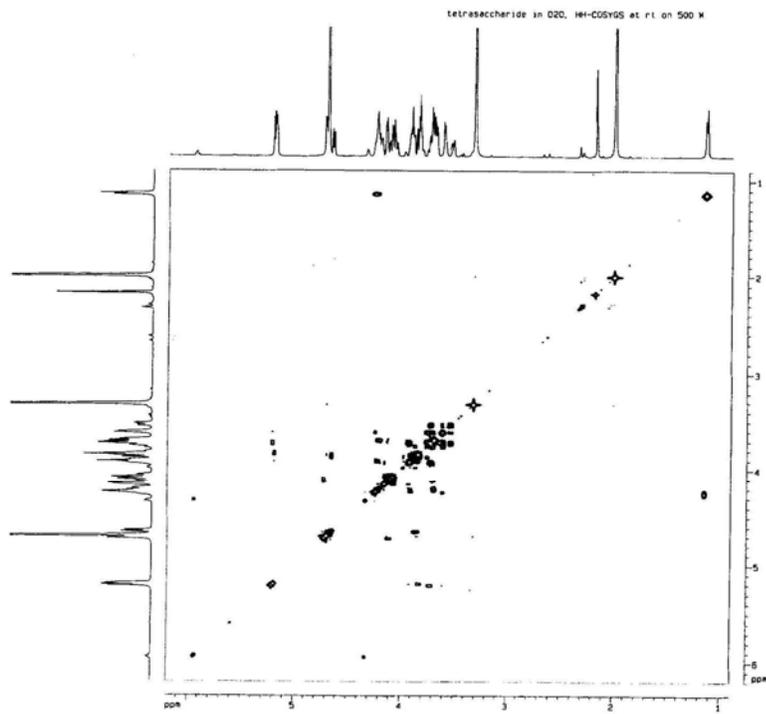
Exact Mass (M+Na)⁺: 728.2583
Found: 728.2584

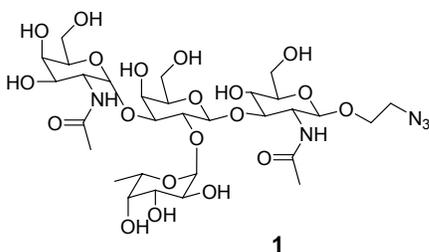
#	GalNAc		β -Gal		α -Gal		Fuc	
	¹ H (δ)	¹³ C (δ)	¹ H (δ)	¹³ C (δ)	¹ H (δ)	¹³ C (δ)	¹ H (δ)	¹³ C (δ)
1	4.73, d, $J = 3.3$ Hz, 1H	98.20	4.65, d, $J = 7.5$ Hz, 1H	102.56	5.20, d, $J = 3.3$ Hz, 1H	93.41	5.21, d, $J = 4.5$ Hz, 1H	99.25
2	4.10, m, 1H	49.74	3.96, m, 1H	73.32	3.84, m, 1H	68.51	3.72, m, 1H	68.10
3	3.70, m, 1H	63.87	3.92, m, 1H	76.50	3.92, m, 1H	69.88	3.55, dd, $J = 7.8, 3.2$ Hz, 1H	70.38
4	3.61, m, 1H	72.25	4.15, m, 1H	69.57	4.08, m, 1H	69.73	3.61, m, 1H	75.00
5	3.83, m, 1H	70.73	4.10, m, 1H	74.66	4.23, m, 1H	71.50	4.22, m, 1H	67.27
6	3.75, m, 2H	61.35	3.74, m, 2H	61.65	3.73, m, 2H	61.76	1.47, d, $J = 6.6$ Hz, 3H	15.73
7	3.32, s, 3H (OCH ₃)	55.52						
8	C=O	174.09						
9	2.01, s, 3H (COCH ₃)	22.42						





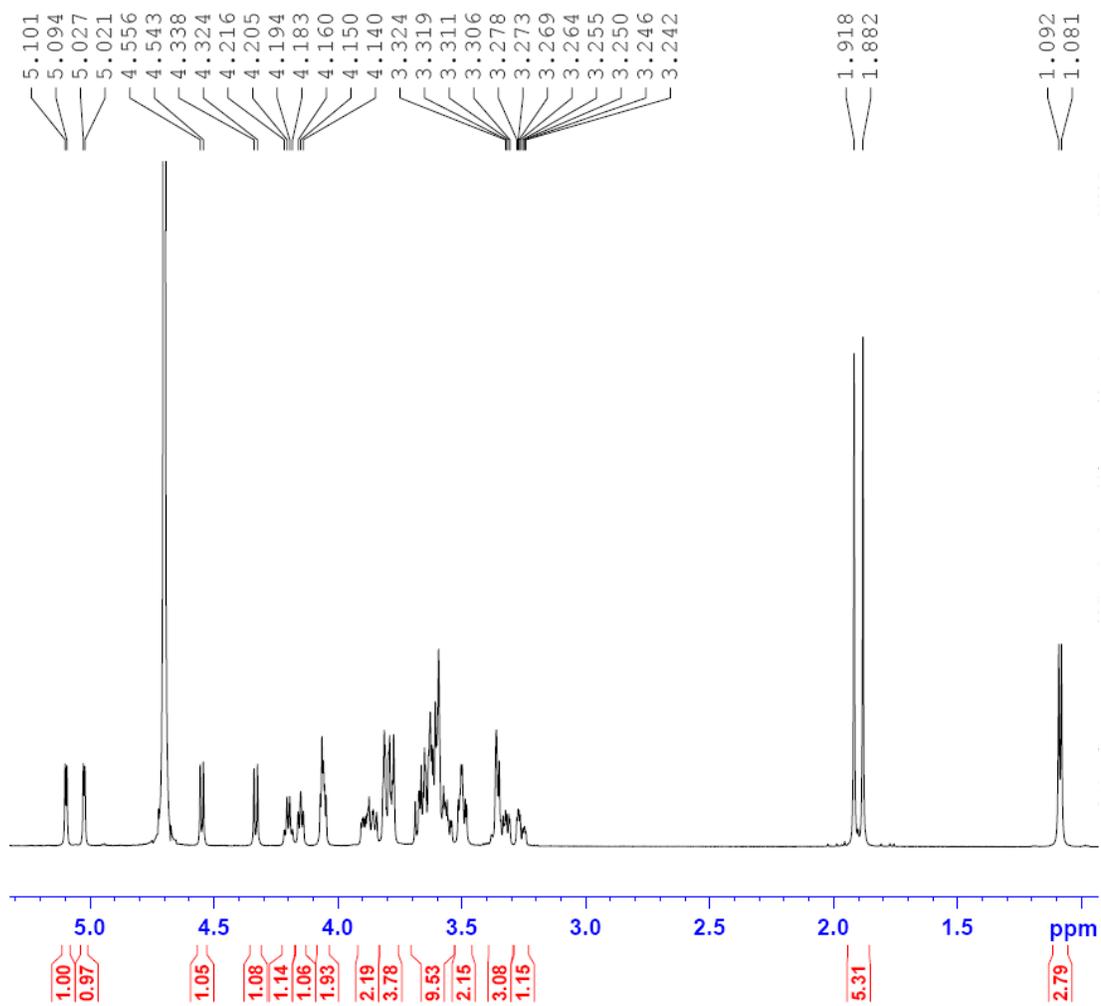


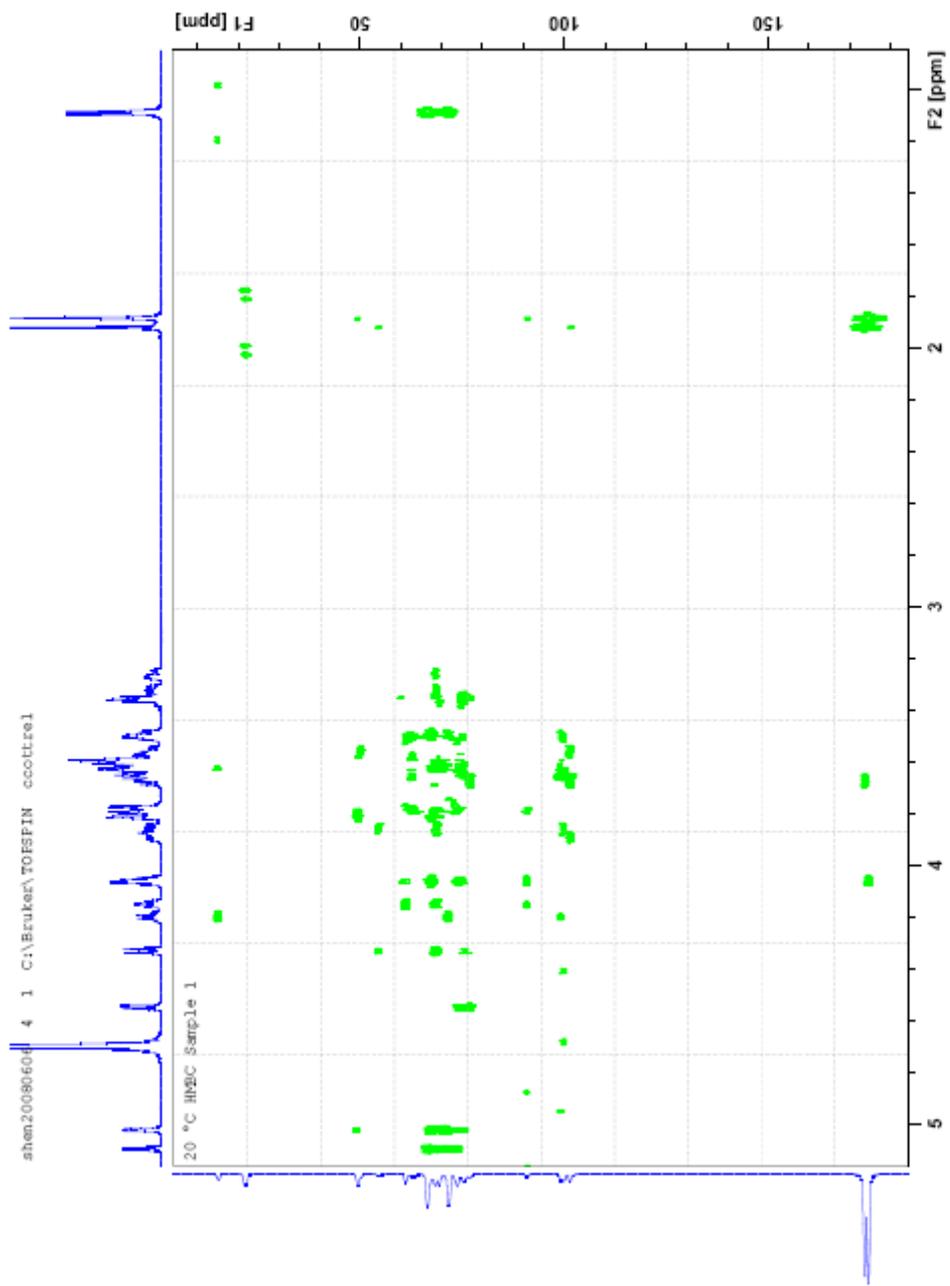


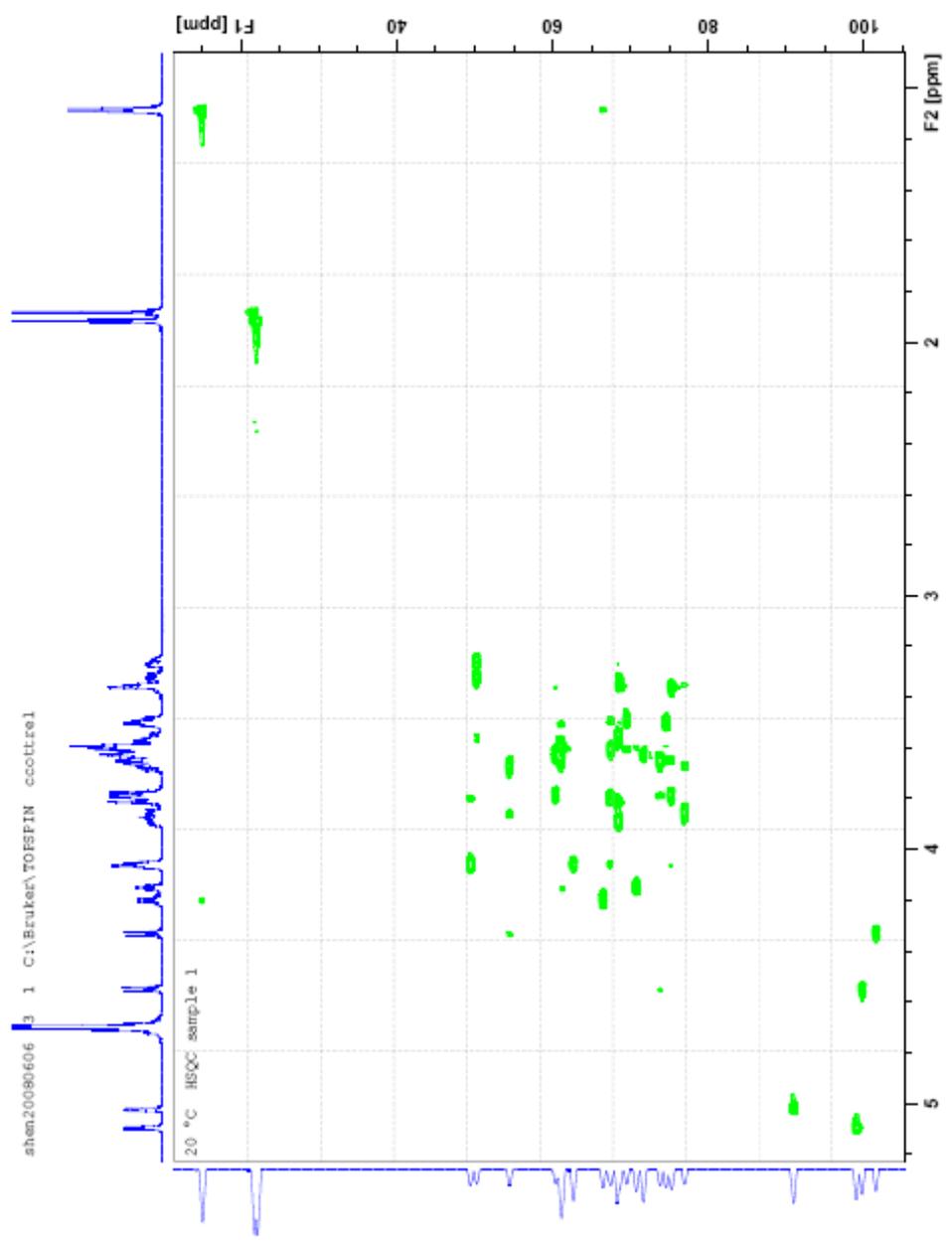


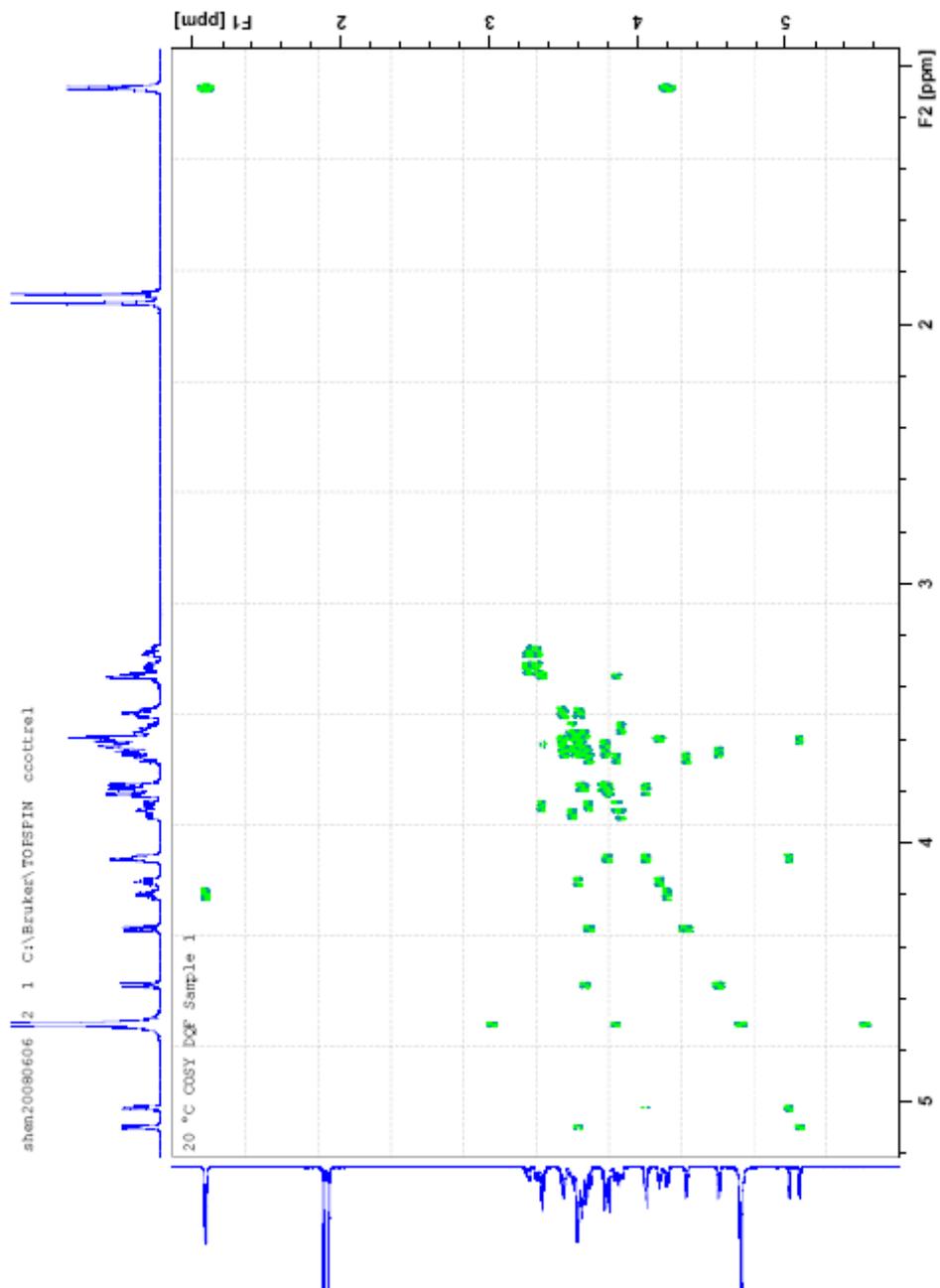
	GalNAc	Gal	GlcNAc	Fuc	others
H-1	5.02 (d) J = 3.7 Hz 91.0	4.55 (d) J = 7.6 Hz 99.8	4.33 (d) J = 8.7 Hz 101.6	5.10 (d) J = 4.0 Hz 99.0	- CH_2N_3
H-2	4.06 (dd) J = 3.7 Hz J = 11.0 Hz 49.5	3.65 (dd) J = 7.6 Hz J = 11.4 Hz 73.9	3.67 (dd) J = 7.8 Hz J = 12.0 Hz 54.5	3.60 (dd) J = 4.0 Hz J = 11.0 Hz 67.5	3.32 (m), 50.3 3.26 (m), 50.3
H-3	3.81 (d) J = 11.4 Hz 68.4	3.80 (dd) J = 11.0 Hz J = 3.8 Hz 67.5	3.86 (dd) J = 7.3 Hz J = 11.3 Hz 77.0	3.49 (d) J = 10.4 Hz 69.6	- OCH_2 - CH_2N_3
H-4	3.36 (s) 68.7	4.06 (*) 62.7	3.36 (dd) J = 6.2 Hz J = 11.5 Hz 68.7	3.62 (*) 71.7	3.89 (m), 68.5 3.56 (m), 68.5
H-5	4.15 (pseudo t) J = J = 6 Hz 70.8	3.49 (*) 74.6	3.78 (*) 75.2	4.20 (q) J = 6.6 Hz 66.5	
H-6	3.60 (pseudo d, 2H)	3.62 (* , 2H)	3.78 (dd) J = 12.2 Hz J = 4.0 Hz 60.4	1.09 (d, 3H) J = 6.6 Hz 15.2	
H-6'	J = 6 Hz 61.2	61.2	3.62 (dd) J = 12.7 Hz J = 5.5 Hz 60.4		
NHCO	174.7	N/A	173.8	N/A	
NHCOMe	1.88 (s, 3H) 21.7	N/A	1.92 (s, 3H) 22.1	N/A	

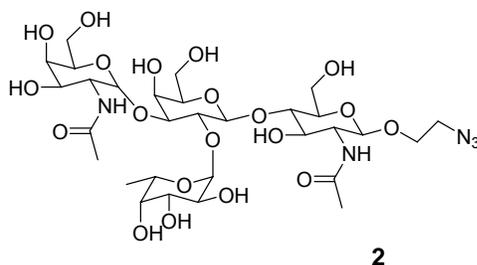
20 °C sample 1



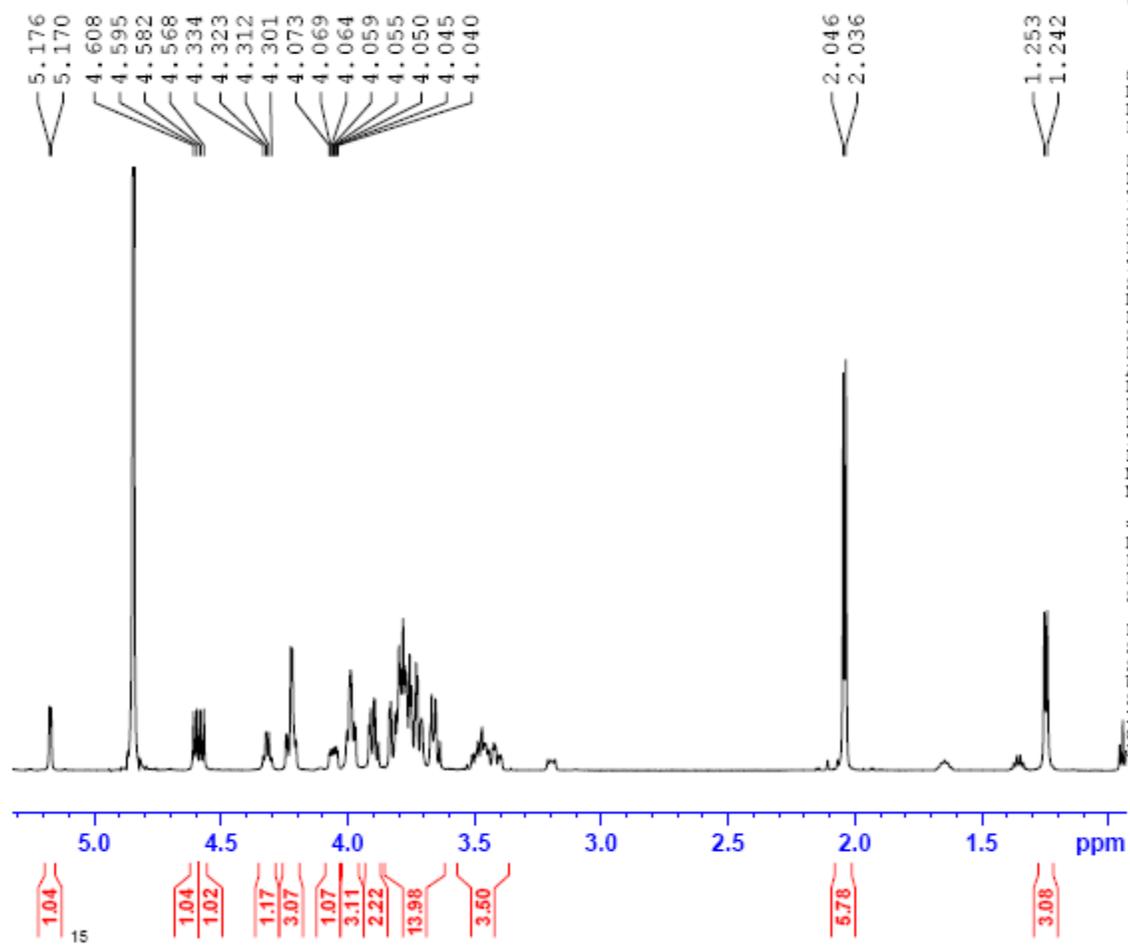


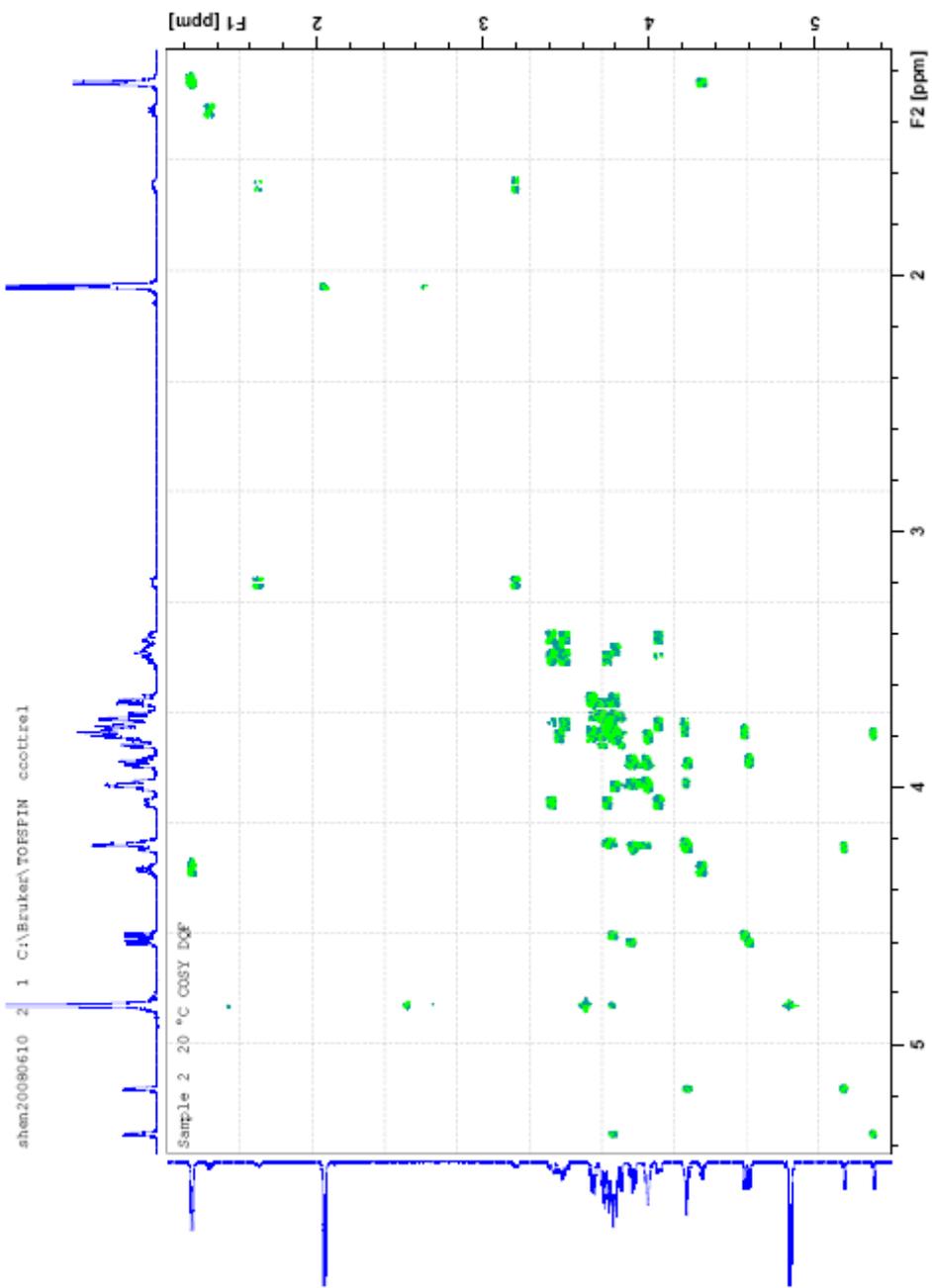


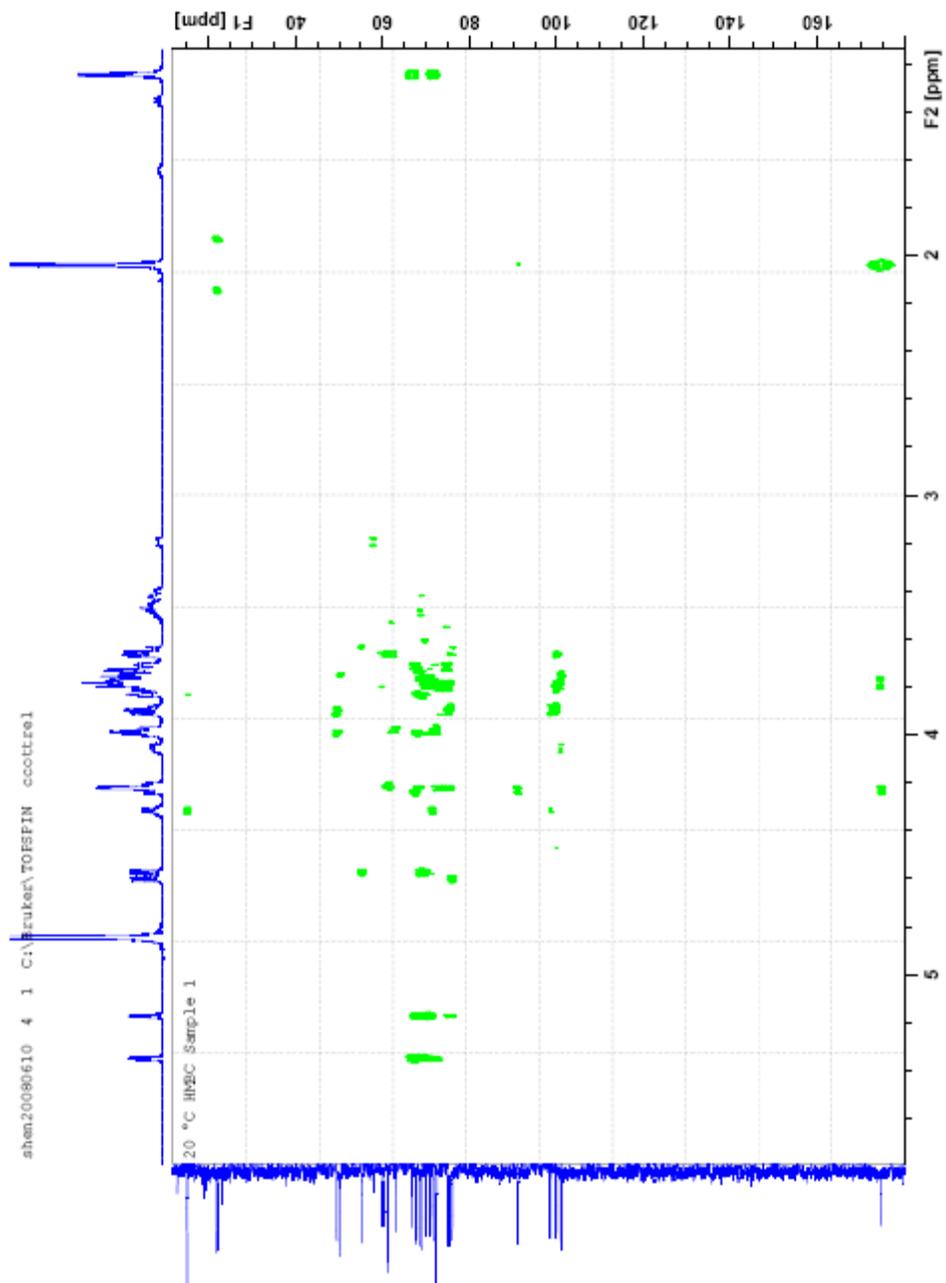


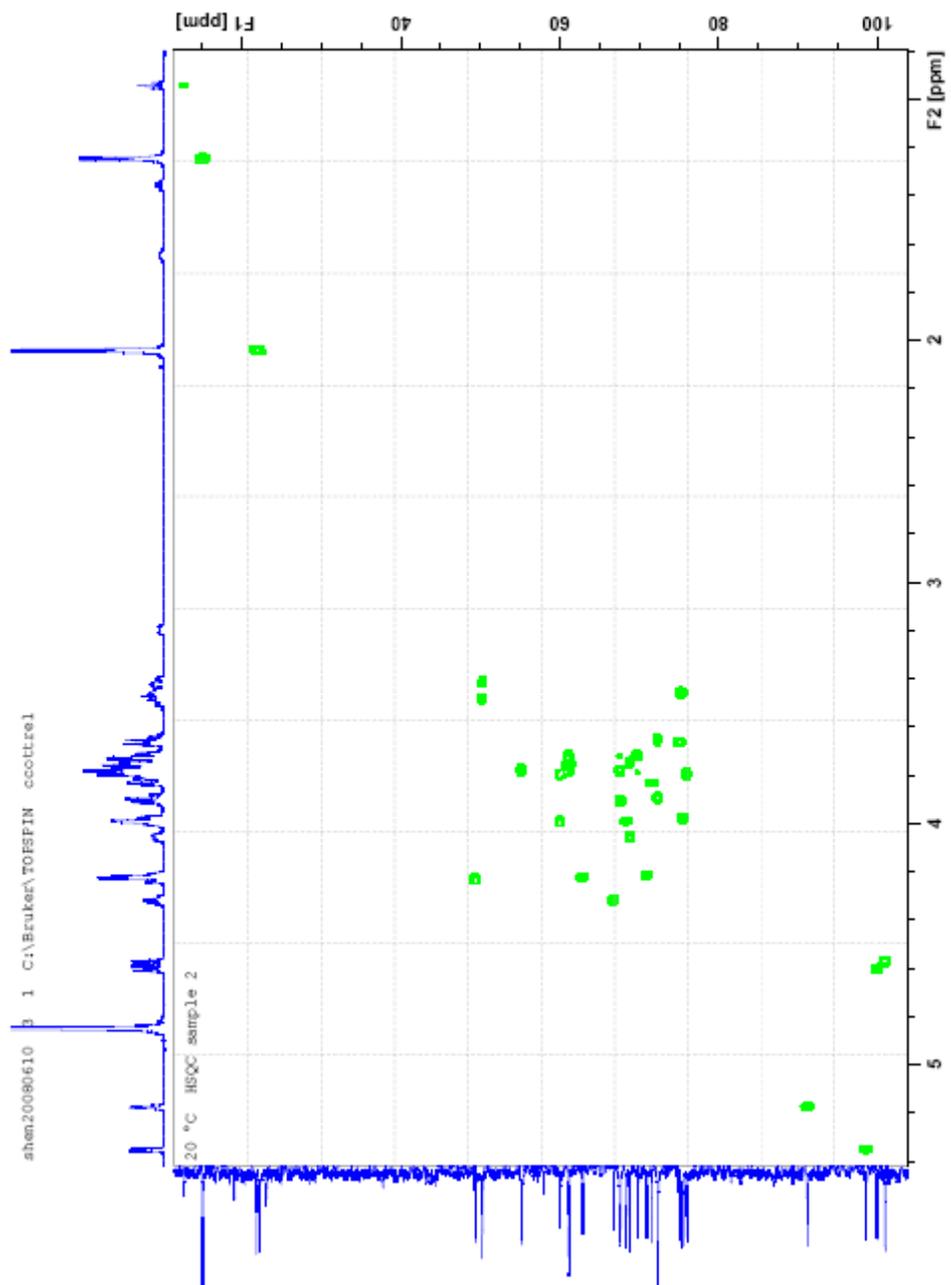


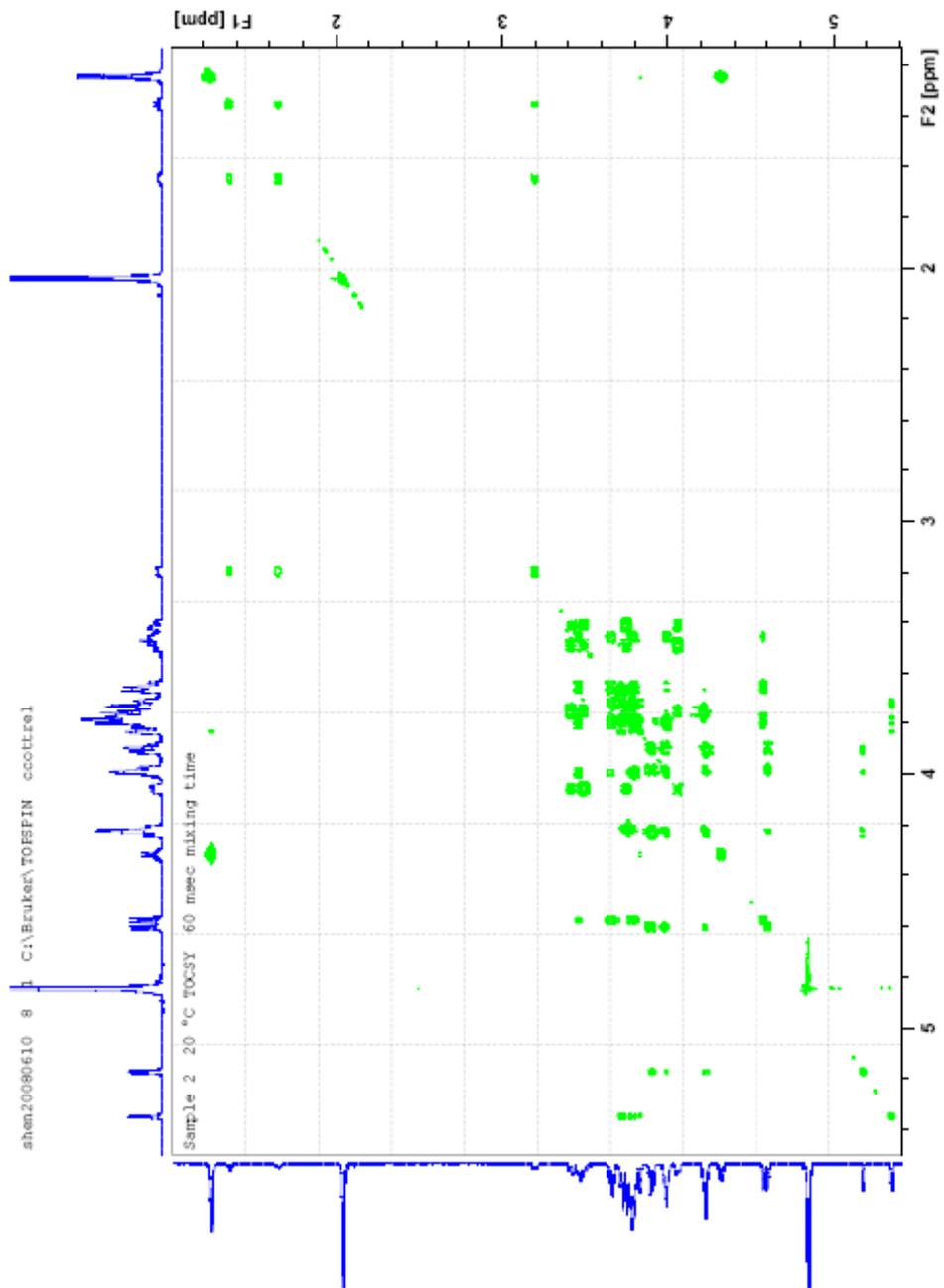
	GalNAc	Gal	GlcNAc	Fuc	others
H-1	5.17 (d) J = 3.6 Hz 91.2	4.60 (d) J = 7.7 Hz 100.0	4.57 (d) J = 8.5 Hz 101.0	5.35 (d) J = 4.0 Hz 98.6	-CH ₂ N ₃
H-2	4.23 (dd) J = 4.6 Hz J = 10.8 Hz 49.4	3.90 (dd) J = 6.9 Hz J = 11.2 Hz 72.4	3.78 (dd) J = 7.4 Hz J = 12.2 Hz 55.2	3.79 (dd) J = 4.8 Hz J = 11.1 Hz 67.6	3.42 (m), 50.2 3.49 (m), 50.2
H-3	3.90 (d) J = 11.4 Hz 67.6	3.98 (dd) J = 3.3 Hz J = 11.5 Hz 75.6	3.79 (*) 76.0	3.72 (d) J = 11.8 Hz 69.8	-OCH ₂ - CH ₂ N ₃ 3.75 (m), 68.8
H-4	3.99 (s) 68.3	4.22 (d) J = 4.8 Hz 62.9	3.66 (dd) J = 13.6 Hz J = 10.3 Hz 75.1	3.83 (s) 71.6	4.06 (m), 68.8
H-5	4.21 (dd) J = 3.4 Hz J = 9.1 Hz 70.9	3.65 (*) 72.4	3.46 (dd) J = 11.2 Hz J = 6.0 Hz 75.3	4.32 (q) J = 6.5 Hz 66.8	
H-6	3.78-3.72 (m, 2H) 61.2	3.78-3.72 (m, 2H) 61.1	3.99 (d) J = 12.1 Hz 60.0	1.25 (d, 3H) J = 6.6 Hz 15.1	
H-6'			3.80 (*) 60.0		
NHCO	174.7	N/A	174.6	N/A	
NHCOMe	2.04 (s, 3H) 21.8	N/A	2.05 (s, 3H) 22.2	N/A	

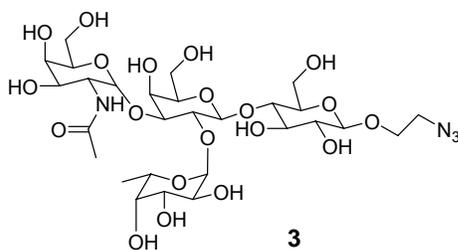












	GalNAc	Gal	Glc	Fuc	others
H-1	5.17 (d) J = 3.7 Hz <u>91.2</u>	4.58 (d) J = 7.7 Hz <u>100.0</u>	4.48 (d) J = 8.0 Hz <u>102.3</u>	5.35 (d) J = 4.0 Hz <u>98.5</u>	-CH ₂ N ₃
H-2	4.23 (dd) J = 12.0 Hz J = 5.0 Hz <u>49.4</u>	3.90 (dd) J = 10.7 Hz J = 6.8 Hz <u>72.3</u>	3.35 (pseudo t) J = 7.3 Hz <u>75.4</u>	3.78 (dd) J = 10.6 Hz J = 5.2 Hz <u>67.5</u>	3.55 (m), <u>50.4</u> 3.56 (m), <u>50.4</u>
H-3	3.90 (d) J = 11.3 Hz <u>67.6</u>	3.97 (dd) J = 10.8 Hz J = 4.9 Hz <u>75.6</u>	3.58 (dd) J = 10.8 Hz J = 6.8 Hz <u>74.2</u>	3.71 (dd) J = 10.9 Hz J = 4.8 Hz <u>69.8</u>	-OCH ₂ - CH ₂ N ₃ 4.05 (m), <u>68.5</u>
H-4	3.99 (s) <u>68.4</u>	4.22 (d) J = 5.0 Hz <u>62.9</u>	3.73 (dd) J = 4.5 Hz J = 10.9 Hz <u>75.8</u>	3.83 (d) J = 4.5 Hz <u>71.6</u>	3.83 (m), <u>68.5</u>
H-5	4.21 (dd) J = 9.0 Hz J = 3.5 Hz <u>71.0</u>	3.45 (dd) J = 6.8 Hz J = 7.6 Hz <u>75.4</u>	3.65 (dd) J = 4.0 Hz J = 8.7 Hz <u>75.0</u>	4.32 (q) J = 6.6 Hz <u>66.7</u>	
H-6	3.79-3.73 (m, 2H) <u>69.2</u>	3.79-3.73 (m, 2H) <u>69.1</u>	3.98 (d) J = 12.3 Hz <u>60.0</u>	1.25 (d, 3H) J = 6.6 Hz <u>15.0</u>	
H-6'			3.78 (dd) J = 7.7 Hz J = 13.1 Hz <u>60.0</u>		
NHCO	<u>174.8</u>	N/A	N/A	N/A	
NHCOMe	2.037 (s, 3H) <u>21.8</u>	N/A	N/A	N/A	

