

DESIGN, SYNTHESIS AND EPITOPE MAPPING OF GROUP A STREPTOCOCCUS
OLIGOSACCHARIDE ANTIGENS FOR THE DEVELOPMENT OF NEW VACCINE CANDIDATES

by

HUZI SUN

(Under the Direction of Geert-Jan Boons)

ABSTRACT

Group A Streptococcus (GAS) is a leading cause of many deadliest diseases in humans, which lead to millions of deaths every year. GAS is the sole species of Lancefield group A and all GAS serotypes express the Lancefield group A carbohydrate (GAC), which is comprised of a polyrhamnose backbone with an immunodominant *N*-acetylglucosamine (GlcNAc) side chain. The hurdle towards a carbohydrate-based vaccine includes the potential of the GAC GlcNAc side chain to provoke cross-reactive antibodies that are relevant to the immunopathogenesis of rheumatic fever. Researchers have reported that antibodies to GAC lacking the GlcNAc side chain and containing only the polyrhamnose promoted opsonophagocytic killing of multiple GAS serotypes and protected against systemic GAS challenge after passive immunization. Several synthetic challenges associated with these oligosaccharide antigens, including sensitivity and reactivity of the substrate, etc. need to be addressed. Besides, lack of structure-activity relationships for the cross-reactive antibody makes it challenging to understand the pathogenesis of the GAS autoimmunity at a molecular level. Moreover, it greatly complicates the design and the development of safe and effective GAS vaccine candidates.

In this dissertation, we have described a synthetic methodology that can rapidly provide a library of well-defined GAC oligosaccharides with different GlcNAc side chain variations and

chain lengths. It is based on the use of a “key disaccharide” to modularly assemble GAC oligosaccharides with different chain variations and lengths. We have discovered two new chemical reactions, i.e., PMB migration and the “4+4” *in situ* bond cleavage polymerization. Having conducted comprehensive studies, we proposed effective mechanisms for the discovered reactions as we successfully synthesized biologically important GAC oligosaccharide antigens. Besides, we have collaborated with biologists at Utrecht University to conduct comprehensive biological investigations of the GAC library and the glycoconjugate vaccine candidate. Different types of GAS antibodies were investigated by the GAC-microarray. The binding study showed that each antibody recognized multiple compounds and exhibited distinct structure–binding relationships. The GAC-microarray data made it possible to investigate and validate the influences of different lengths and side-chain variations on the structure–binding relationships with type-specific antibodies. The array data supports a notion that the side-chain variation and length can have significant influences on the structure–binding relationships and provide valuable insights to the cross-reactivity associated with the autoimmunity of the GAS infections.

INDEX WORDS: Group A Streptococcus, glycoconjugate vaccine, microarray, carbohydrate

DESIGN, SYNTHESIS AND EPITOPE MAPPING OF GROUP A STREPTOCOCCUS
OLIGOSACCHARIDE ANTIGENS FOR THE DEVELOPMENT OF NEW VACCINE CANDIDATES

by

HUZI SUN

B.S., China Agricultural University, China, 2014

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2019

© 2019

Huzi Sun

All Rights Reserved

DESIGN, SYNTHESIS AND EPITOPE MAPPING OF GROUP A STREPTOCOCCUS
OLIGOSACCHARIDE ANTIGENS FOR THE DEVELOPMENT OF NEW VACCINE CANDIDATES

by

HUZI SUN

Major Professor:	Geert-Jan Boons
Committee:	Vladimir Popik
	Eric Ferreira

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2019

DEDICATION

To my loving family

ACKNOWLEDGEMENTS

It is truly an adventure in the broader field of life science, which challenged and inspired me to grow not only professionally but also personally. I want to thank my advisor Dr. Geert-Jan Boons, an inspiring advisor who provides valuable guidance and supports during my Ph.D. career. I also would like to thank my doctoral advisory committee, Dr. Vladimir Popik, and Dr. Eric Ferreira for their suggestions, guidance, and supports. Additionally, I want to thank Dr. Jason Locklin, Dr. Jin Xie, Dr. Ryan Hili, Dr. Huge Waston, James Flannery, Dr. Jonathan Amster, Dr. Gary Douberly, Dr. George Majetich, and Dr. Jeffrey Urbauer for their supports.

My special thanks go to people with whom I have had the privilege to work, especially to Dr. Ning Ding and Dr. Margreet Wolfert for the glycoconjugate vaccine project and the synthetic GAC microarray project. I would also like to thank my lab mates, including, but not limited to, Dr. Xiuru Li, Dr. Lin Liu, Dr. Nitin Supekar, Dr. Robert Chapman, Dr. Chengli Zong, Dr. Tiantian Sun, Dr. Wei Huang, Dr. Tiehai Li, Dr. María J. Moure García, Dr. Apoorva Joshi, Dr. Shailesh Ambre, Dr. Pradeep Chopra, Dr. Anthony Prudden, Dr. Manish Hudlikar, Dr. Josette Wilkes, Apoorva Srivastava, Weiyu Li, Zeshi Li, Mehman Bunyatov, Ingrid 't Hart, Frederik Broszeit, Dushen Chen, Minglong Liu, Xianke Meng, Lifeng Sun, Liangwei Zhang, Ashley Carter, Gerlof Bosman, Weigang Lu, Alex Walker, Ivan Gagarinov for their scientific discussions and companionship. I also want to thank Rachel Bainum and Lauren Bowman for their administrative arrangements. For the wonderful facilities at CCRC, I would like to thank Dr. John Glushka and members of analytic service for their training and maintenance of instruments. Last but not least, I want to attribute this achievement to my family.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF ABBREVIATIONS	xi
CHAPTER	
1 INTRODUCTION AND LITERATURE REVIEW	1
Group A streptococcus infections, disease burden and current status	1
The pathophysiology of invasive GAS infection	3
The host immunities toward GAS infection	6
GAS vaccinology	8
The group A carbohydrate	16
Carbohydrate chemistry	20
An introduction to the discoveries and proposed mechanisms in this work	30
References	35
2 DESIGN AND CHEMICAL SYNTHESIS OF POLYRHAMNOSE OLIGOSACCHARIDES FOR THE DEVELOPMENT OF GROUP A STREPTOCOCCUS VACCINE CANDIDATES	55
Abstract	56
Introduction	57

Discussions and results	59
Conclusion	69
Experimental Section	70
References	150
3 CHEMICAL SYNTHESIS AND EPITOPE MAPPING OF GROUP A STREPTOCOCCUS OLIGOSACCHARIDE ANTIGENS	153
Abstract	154
Introduction	155
Discussions and results	158
Conclusion	170
Experimental Section	171
References	203
4 CONCLUSIONS	206

APPENDICES

A SUPPLEMENTARY DATA OF KEY PRODUCTS IN CHAPTER 2	208
B SUPPLEMENTARY DATA OF KEY PRODUCTS IN CHAPTER 3	239

LIST OF TABLES

	Page
Table 1.1: The development status of current vaccine candidates.....	9
Table 1.2: PMB migration	31
Table 2.1: The optimization of the key disaccharide glycosylation condition	62
Table 2.2: Model test results for the optimization of the disaccharide assemble condition	76
Table 2.3: PMB Migration reactions	78
Table 2.4: A library of polyrhamnose oligosaccharides after LH20 purification.....	85
Table 2.5: Conditions for linker addition reactions	89
Table 2.6: Reaction condition for linker addition with <i>N</i> -Phenyl imidate.....	90
Table 2.7: Reaction condition for N ₃ -linker addition with <i>N</i> -Phenyl imidate	90

LIST OF FIGURES

	Page
Figure 1.1: <i>S. pyogenes</i> bacteria (also called Group A streptococcus) at 900x magnification.....	1
Figure 1.2: Chemical structure and computer model structures of GAC polysaccharide	3
Figure 1.3: Chemical structure and computer model structures of the side chain deficient GAC.....	17
Figure 1.4: Carbohydrate numbering and the ring-closing process.....	21
Figure 1.5: Linkage configuration	22
Figure 1.6: General glycosylation reaction with strategic selection of protecting groups.....	23
Figure 1.7: General activation of the glycosyl donor in the glycosylation process.....	24
Figure 1.8: Various leaving groups of glycosyl donors.....	25
Figure 1.9: Anomeric effects	26
Figure 1.10: Oligosaccharide assembly strategies.....	27
Figure 1.11: General mechanisms of the glycosyltransferases.....	29
Figure 1.12: Reverse hydrolysis process with glycosidase.....	30
Figure 1.13: The “4+4” in situ bond cleavage polymerization.....	33
Figure 2.1: Building blocks.....	73
Figure 2.2: TLC of the ”4+4” polymerization	84
Figure 2.3: MALDI of the “4+4” polymerization reaction mixture	84
Figure 2.4: MALDI of the dodecasaccharide after LH20 purification	84
Figure 2.5: MALDI of the tetradecasaccharide after LH20 purification	84
Figure 2.6: NMR spectra of the deprotected polyrhamnose oligosaccharides	92

Figure 2.7: TLC and MALDI of 4+4-Bn reaction	96
Figure 2.8: TLC and MALDI of the octasaccharide reaction	99
Figure 2.9: NMR spectra of the deprotected octasaccharide 32	101
Figure 2.10: MALDI of the BSA glycoconjugate	103
Figure 2.11: Standard curve of the HPAEC analysis.....	104
Figure 3.1: NMR Spectra of deprotected GAC library	163
Figure 3.2: Glycan Microarray results for the anti-polyrhamnose antibody.	165
Figure 3.3: Glycan Microarray results for the anti-streptococcus pyogenes Group A Carbohydrate antibody.....	166
Figure 3.4: Glycan Microarray results for the anti-Streptococcus Group A antibody.....	167
Figure 3.5: Glycan Microarray results for the rabbit serum derived from immunization of the octa-polyrhamnose KLH-glycoconjugate	168
Figure 3.6: NMR Spectra of deprotected GAC library with GlcNAc variations.....	179
Figure 3.7: NMR Spectra of deprotected polyrhamnose library.....	180

LIST OF ABBREVIATIONS

Ac.....	Acetyl
Ac ₂ O.....	Acetic Anhydride
AcOH.....	Acetic Acid
ARF.....	Acute Rheumatic fever
BF ₃ -Et ₂ O.....	Boron Trifluoride Diethyl Etherate
Bn.....	Benzyl
BnBr.....	Benzyl Bromide
BSA.....	Bovine Serum Albumin
Bz.....	Benzoyl
COSY.....	Correlation Spectroscopy
Cbz.....	Carboxybenzyl
CSA	Camphorsulfonic Acid
DBU.....	1,8-Diazabicycloundec-7-ene
DCC.....	<i>N,N'</i> -Dicyclohexylcarbodiimide
DCE.....	1,2-Dichloroethane
DCM.....	Dichloromethane
DDQ.....	2,3-Dichloro-5,6-dicyano-1,4-benzoquinone
DTT	Dithiothreitol
DHB	2,5-Dihydroxybenzoic Acid
DIC	<i>N,N'</i> -Diisopropylcarbodiimide

DIPEA.....	Diisopropylethylamine
DMAP.....	4- <i>N,N'</i> -Dimethylaminopyridine
DMF.....	<i>N,N'</i> -Dimethylformamide
EDC.....	<i>N</i> -(3-Dimethylaminopropyl)- <i>N'</i> -ethylcarbodiimide hydrochloride
Fmoc	9-Fluorenylmethyloxycarbonyl
Gal	Galactose
GAS.....	Group A Streptococcus
GAC.....	Group A carbohydrate
Glc	Glucose
GlcNAc.....	<i>N</i> -Acetylglucosamine
GAS-PS.....	Group A Streptococcus cell membrane polysaccharide
HMBC.....	Heteronuclear Multiple Bond Correlation Spectroscopy
HSQC.....	Heteronuclear Single Quantum Coherence Spectroscopy
Lev.....	Levulinyl
LG.....	Leaving group
MALDI-TOF	Matrix Assisted Laser Desorption Ionization Time-Of-Flight
Man	Mannose
mCPBA	meta-Chloroperbenzoic Acid
Me	Methyl
MS	Molecular Sieves
Nap	2-Methylnaphthyl
NBS.....	<i>N</i> -Bromosuccinimide

Neu5Ac	<i>N</i> -Acetylneuraminic Acid
NIS	<i>N</i> -Iodosuccinimide
NMR	Nuclear Magnetic Resonance
OMe	Methoxy
PBS	Phosphate Buffered Saline
Ph	Phenyl
PMB.....	<i>p</i> -Methoxybenzyl ether
Py	Pyridine
TBAF	Tetrabutyl Ammonium Fluoride
TBAI.....	Tetrabutyl Ammonium Iodide
TBAB.....	Tetrabutyl Ammonium bromide
TDS	Dimethyl (1,1,2-trimethylpropyl) silyl
TCA	Trichloroacetonitrile
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
TEA	Triethylamine
Tf ₂ O	Trifluoromethanesulfonic Anhydride
TFA	Trifluoroacetic Acid
TfOH	Trifluoromethanesulfonic Acid
THF	Tetrahydrofuran
TLC	Thin Layer Chromatography
TMSOTf	Trimethylsilyl Trifluoromethanesulfonate
TOCSY	Total Correlation Spectroscopy
Troc	2,2,2-Trichloroethyloxycarbonyl

RHD.....Rheumatic heart disease
Rha.....Rhamnose

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

***Group A streptococcus* infections, disease burden, and current status**



Figure 1. *S. pyogenes* bacteria (also called *Group A streptococcus*) at 900x magnification¹

Group A Streptococcus (GAS) is a species of gram-positive bacteria. These bacteria are aero-tolerant, and they are parts of the skin microbiota. *Group A streptococci* are the predominant species that harbor the Lancefield group A antigen. They are round usually pathogenic, beta-hemolysis bacteria, which cause the destruction of red blood cells. The name “*streptococcus pyogenes*” is derived from Greek words; “strepto” means a chain, “coccus” means berries, and “pyogenes” means pus-forming. Streptococcal cells tend to link in chains of round cells, as shown in **Figure 1**.¹

GAS infections are leading healthcare problems throughout the world. They can cause a wide range of diseases, including asymptomatic colonization, uncomplicated pharyngeal, skin infections, and life-threatening invasive illnesses such as sepsis, necrotizing fasciitis, and toxic shock syndrome. Pharyngitis may lead to delayed sequela as rheumatic fever, which is caused by an autoimmune reaction to Group A streptococci that results in valvular damage.² In the developing world, rheumatic fever is the primary cause of acquired heart diseases in children,

adolescents, and young adults, responsible for at least 350,000 premature deaths per year.³ In the western world, it is renowned, but often overlooked, for causing tonsillitis and impetigo amongst teenagers. Despite its continued susceptibility to beta-lactam antibiotics, unexplained resurgences of invasive GAS infections occurred in the UK over the past 30 years.⁴ According to recent research, the resurgences were mainly due to the reappearance and/or increased circulation of a highly invasive clone of M1T1 GAS.¹¹ This highlights the increased need for applying preventive methods against the spread and the resurgence of invasive GAS diseases and close surveillance of GAS in the community. There is currently no safe and efficacious commercial vaccine against GAS infection despite the high demand globally. Therefore, it would prove invaluable to achieve a safe and effective GAS vaccine.

GAS is primarily transmitted by direct person-to-person contact or aerosol spread¹². In a clinical setting, identification of GAS is recorded by the formation of beta-haemolytic, bacitracin sensitive colonies on 5% blood agar plates. GAS can be differentiated from other pathogenic streptococci serologically based on cell surface Lancefield antigens¹³. The “type A” antigen of GAS consists of a highly conserved $\{-2\}[\beta\text{-D-GlcNAc}(1\rightarrow3)]\alpha\text{-L-Rha}(1\rightarrow3)\}\alpha\text{-L-Rha}(1-\}_n$ polysaccharide as shown in **Figure 2**. Based on computer modeling, GAS polysaccharide chains adopt a helical conformation where the immunodominant *N*-acetylglucosamine residues are exposed on the periphery of the polyrhannose backbone.^{5,14,15} Historically, based upon the surface M protein, GAS isolates have been subdivided into serological M types.^{16,17} Contemporarily, based on sequencing of the hypervariable N terminal region of the M protein gene, GAS isolates have genealogically subtyped as emm types. The hypervariable N terminal region of the mature protein has been used for differentiating the classical M serotypes.¹⁸

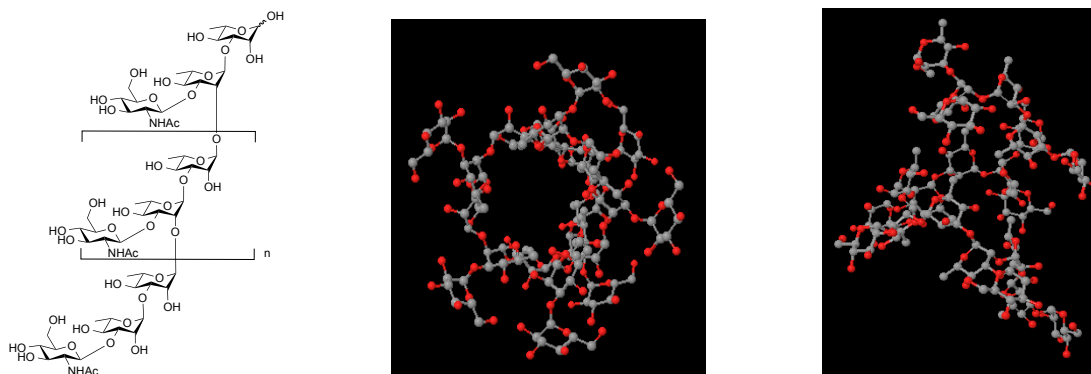


Figure 2. Chemical structure and computer model structures of GAC polysaccharide with GlcNAc side chains

The occurrence and distribution of different GAS serotypes in the developed world are vastly different from those reported in the developing countries. According to a recent global review of GAS epidemiology, 25 distinct emm types accounted for more than 90% of all GAS diseases in Europe, North America, Australia, New Zealand, and Japan.¹⁹ However, a much more diversified sequence types were reported in developing countries, none of which were significantly dominant¹⁹. The ten most prominent emm types in Europe, emm1, 3, 4, 5, 12, 28, 81, 83, 87, and 89, purportedly account for over 70% of all invasive GAS infections.^{20, 21} Because of the enormous diversity of the GAS subtype species, it is very difficult to find a conserved surface M protein coding sequence for the development of a universally effective vaccine.

The pathophysiology of invasive GAS infection

GAS is primarily disseminated via direct person-to-person contact or aerosol spread, and it is an obligate human pathogen with no known environmental reservoir. In the developing countries, the person-to-person spread of GAS infection can be readily facilitated by streptococcal impetigo pustules due to the lack of medical infrastructure and environmental hygiene.^{22, 23,24} In regions where endemic impetigo is less common, studies have shown that

clandestine infection of the oropharynx occurred with 15-20% GAS carriage rates among school-aged children.²⁵ These persistent carriers consist of an important reservoir for the spread of GAS infection in a community setting.^{26, 13}

Invasive GAS infections occur when the organism spreads from the initial nidus of superficial infection into the deeper host tissues, which is usually associated with the soft tissue of the skin. Invasive GAS infections can lead to a variety of clinical conditions, including ecthyma, cellulitis, and necrotizing fasciitis. Oral or parenteral penicillin is usually the initial treatment for invasive GAS skin infection, and an appropriate anti-staphylococcal agent, such as flucloxacillin, is often prescribed in concert with the initial penicillin treatment.^{27,23} In the severe case of necrotizing fasciitis, drainage or surgical debridement of the afflicted region is necessary due to the depth of soft tissue involvement.²⁸ To control the systemic inflammatory symptoms of the underlying invasive GAS infection, it is necessary to treat patients with adjunct antimicrobials, such as clindamycin and adjunct intravenous immunoglobulin G therapy, due to the purported correlation between soft tissue diseases and the onset of streptococcal toxic shock syndrome.^{29,30}

Sepsis is a fatal inflammatory condition, and a wide range of bacterial, viral, and fungal pathogens can trigger sepsis. Sepsis occurs when the immune system overproduces pro-inflammatory cytokines in response to a severe systemic infection. Thus, the symptoms are largely immune-mediated. A series of non-focal symptoms including fever, chills, malaise, nausea, and vomiting occurs in the initial condition, and the associated inflammation progresses rapidly to cause reduced organ perfusion, clotting abnormalities, and respiratory distress if the underlying infection remains untreated.^{31, 32} The condition will eventually progress to a profound drop in arterial blood pressure and multi-organ failure.^{32,33}

Streptococcal toxic shock syndrome is induced through the production of superantigenic toxins during invasive GAS infections.^{32, 33} The streptococcal toxic shock syndrome worsens the systemic inflammation associated with severe sepsis,³⁴ as these superantigens induce systemic inflammation through the non-specific activation of host T-cells. As a result, incongruous release of pro-inflammatory cytokines occurs.³⁵ Further, this condition triggers a feedback loop between cytokine release and immune cell activation, which precipitates the pathophysiological symptoms of streptococcal sepsis.^{34, 35}

Acute pharyngitis is one of the most frequently consulted conditions. Though only about 30% of the cases can be attributed to GAS infection, the economic burden of streptococcal pharyngitis is too heavy.³⁶ Except for the economic implications of GAS pharyngitis, it is also associated with seasonal outbreaks of acute rheumatic fever (ARF), since the typical GAS pharyngitis symptoms are mainly immune-mediated.³⁶ ARF is an immune-mediated sequel of long-term GAS infection and it primarily affects the joints and cardiac tissues.³⁷ Severe ARF can cause aggressive inflammation of the myocardium and heart valves, which can lead to first degree heart block and congestive heart failure.³⁷ Antimicrobial chemotherapy treatment can resolve individual cases of ARF. However, it is not an efficient and cost-effective means for the treatment of a large population, especially in many regions of the developing world where GAS infections remain an endemic disease.³⁸

Though the pathophysiology of ARF is under debate, it has been shown that ARF results from the generation of human cross-reactive antibodies that recognize both GAS organism and host tissues.^{10,37} Studies have shown that the M protein can stimulate a myosin-reactive autoantibody response, as the M protein has similar alpha-helical structure compared with myosin.^{10,39,40} The cross-reactivity of the M protein raises concerns that M protein-based

vaccinations may induce ARF side effects. Therefore, the field of GAS vaccinology is shifting away from historically dominating M protein-based approaches.

In addition to the M protein, the autoreactivity of antibodies, which recognize the native group A carbohydrate (GAC) GlcNAc side chain, against human tissues has been raised by several groups.^{7,8,9} Cunningham groups derived anti-GlcNAc monoclonal antibodies, which cross-react with human cardiac myosin and valvular laminin, from rheumatic carditis patients, further supported the cross-reactivity of native GAC GlcNAc side chain.¹⁰ It has been shown that rheumatic carditis patients are often present with high circulating anti-GAC antibody titers, and the current research results further explained these historical observations.⁴¹

In recent years, Nizet and co-workers identified the genes responsible for GAC biosynthesis. They demonstrated that antisera raised against the polyrhmannose core of GAC, as purified from the Δ GacI mutant, may still provide significant broad-spectrum opsonophagocytic activity. These data provide researchers a great opportunity to develop a novel universal vaccine candidate and to investigate the autoreactive antibody epitope recognition pattern associated with GAC antigens.

The host immunities towards GAS infection

GAS infections usually involve a two-step process. The first step of GAS bacteria colonization is the attachment of the bacteria to pharyngeal and dermal epithelial cells. The second step overcomes the electrostatic repulsion of the host cell surface and forms covalent binding interactions between the bacteria and several ligands to transmit the GAS bacteria into the host tissues.^{42,13} Binding interactions with surface ligands enhance the bacteria resistance to the mechanical removal via the salivary flow and exfoliation, and the primary passive immunity to invasive GAS infections is provided by the epithelial barriers of the skin and oropharynx.¹³

Besides, a plethora of antimicrobial molecules secreted by the epithelial layer is involved in the pathogen colonization process.^{43, 44}

Studies have shown that cathelicidin LL-37 is a vital antimicrobial peptide in the process of killing the bacteria, and it also inhibits the growth of GAS bacteria during the infection.^{44,45,46} Researchers have also utilized cathelicidin deficient mice to support further that this class of molecule plays a vital role in containing the GAS infection. The studies show that cathelicidin-mice are more susceptible to the GAS infection compared with their wild-type counterpart.⁴⁷ It is known that GAS produces SpeB and the streptococcal inhibitor of complement, which are capable of degrading LL-37, to evade the epithelial immunity of the host tissue. Thus, inactivation of cathelicidin is essential for the establishment of invasive GAS infections.^{48, 49}

Once the epithelial barrier has been breached, cell-mediated immunity by the recruitment of phagocytic cells to the sites of infection facilitates the clearance of the infecting GAS population. Early recognition of GAS includes the activation of the Toll-like receptor(TLR)-2 and the intracellular pattern recognition receptor (PRR) NOD₂. The TLR-2 signaling is triggered by GAS cell wall components, lipoteichoic acid and several other bacterial lipoproteins.^{54, 50, 55} The PRR NOD₂ signaling is triggered by peptidoglycan derived muramyl dipeptide moieties.⁵¹ TLR-2/NOD₂ activation releases proinflammatory cytokines and phagocyte chemo-attractants, which drive the early symptoms of sepsis. Early recognition of GAS and the phagocyte recruitment are essential for adequate clearance of the invading organism before incongruous activation of the adaptive immune response by the streptococcal superantigens.^{51,52,53} Failed containment and dissemination of GAS trigger activation of the alternative complement cascade, which ultimately results in the formation of C3 convertase. The plasma protein C3 is split by C3 convertase into two fragments, denoted as C3a and C3b.⁵⁶ The phagocyte recognizes the soluble

peptide opsonin C3b on the cell surface and ultimately destroys the invading organism.⁵⁶ The alternative complement pathway plays a significant role in GAS clearance before the adaptive immune response, since it does not require functional antibody to promote efficient opsonisation.⁵⁶ However, GAS has evolved several exoenzymes that degrade soluble components of the complement cascade to evade the innate immune response. Thus, the adaptive immunity plays an essential role in clearing the invasive GAS infection.

The anti-streptococcal humoral immune response plays a crucial role in the process of clearing invasive GAS infections, and the clearance of superficial GAS infections is primarily mediated by opsonophagocytosis. Recognition and uptake of GAS by the surface antibody, more specifically, via phagocyte Fc receptors recognition, facilitate the clearance of the infection.⁵⁷ Humoral immunity has been shown to prevent ECM protein binding and can neutralize the activity of the secreted streptococcal exotoxins.^{58, 59, 60} Research shows that rabbits infected with GAS had a strong anti-SAg humoral immune response, and direct immunization with SpeA produced toxin neutralizing antibodies which protected against invasive GAS infection.⁶¹ Clinically, it has been reported that sera from invasive GAS patients have lower neutralizing antibody titers than those from uncomplicated pharyngitis cases, which suggests that a strong humoral immune response protects against invasive GAS disease^{62, 63, 64}. Efficient priming of the humoral immune response can effectively prevent invasive GAS infection, and it is a critical factor to be considered during the design process of the vaccine formulation.

GAS vaccinology

GAS vaccines can be generally divided into M protein-based and non-M protein-based vaccines. M protein-based vaccines are well-developed as compared to GAS carbohydrate-based

vaccines whereas research in this area is undergoing extensively.⁴ **Table 1** shows the development status of current vaccine candidates.

Candidate name/Identifier	Stage of development		
	Pre-clinical	Phase I	Phase II
M protein: 6-valent <i>N</i> -terminal	X	X	
M protein: 26-valent <i>N</i> -terminal	X	X	X
M protein: 30-valent <i>N</i> -terminal	X	*	
M protein: minimal epitope J8	X	X	
M protein: minimal epitope J14/p145	X		
M protein: <i>C</i> -repeat epitope (StreptInCor)	X	*	
M protein: <i>C</i> -repeat epitopes	X		
Three conserved antigens(Combo)	X		
GAS carbohydrate	X		
GAS carbohydrate defective for GlcNAc side-chain	X		
GAS C5a peptidase	X		
Fibronectin-binding protein	X		
Streptococcal protective antigen	X		
Serum opacity factor	X		
Streptococcal pyrogenic exotoxin A/B/C	X		
Streptococcal pili (T antigen)	X		
Serine protease (SpyCEP)	X		
Nine common antigens	X		
Identified but untested antigens:			
G-related 2-macroglobulin binding (GRAB) protein			
Metal transporter of streptococcus (MtsA)			
Superoxidase dismutase			
Lipoproteins			

Table 1. The development status of current vaccine candidates (*approaching trials).⁴

Currently, there is no licensed effective and efficient vaccine available for the prevention of GAS infection. In the early anti-GAS vaccine formulations research, the goal was mainly aimed

to prevent scarlet fever, an inflammatory condition of the oropharynx that is associated with superantigen production following tonsillar colonization.⁶⁵ The whole heat-killed GAS or crudely purified “Dick” antigens were used in human administration trails. These preparations can provide little protection against following streptococcal challenges, and often will induce a range of aggressive side effects.^{66, 65, 67, 68}

Crudely purified cell wall extracts and M protein with different purities as vaccinations have been proven ineffective and resulted in significant morbidity in a series of human trials.^{69, 70, 71} In 1978, the administration of streptococcal extracts directly to human subjects was banned by the Federal Drug Administration due to the safety issues related with ARF side effects in response to a partially purified M3 protein vaccine^{72,73}. This legislation changed the landscape of the clinical trial, and further shifted the human subjects of the early clinical trial to the use of an animal model for the development and evaluation of GAS vaccinations. The field of GAS vaccinology is, thus, divided into two broad categories, i.e., the M protein-based vaccines, and the non-M protein-based vaccines. Identification of a region of the M protein sequence that not only can serve as an effective antigen but also can circumvent the cross-reactivity associated with ARF is the main cause in the M protein-based vaccine development. In contrast, the non-M protein-based vaccine development sought to identify and characterize alternative surface antigens and to use those promising antigens for the design of novel vaccine candidates.

1. Contemporary M protein-based vaccines

The M protein from GAS bacteria surface is attracting lots of research interests as it has long been known to contain type0 specific epitopes, which can be utilized to induce opsonophagocytosis and long-lasting immunological memory^{74, 75, 76}. In the early 1980s, researches showed that the N terminus of the M protein, specifically the first 20 amino acids of

the M5 protein, was sufficient to evoke protective, type-specific immunity⁷⁷. It has been recorded that over 100 distinct M types can be explored for the development of vaccine candidates. Researchers had designed multivalent hybrid molecules with several M protein isoforms in order to elicit broad-spectrum immunity against GAS infection^{74, 77, 78}. Over the past decades, many multivalent hybrid M protein-based vaccine candidates containing a variety of N terminal epitopes were developed and tested in the pre-clinical setting. These research results showed that many of the vaccine candidates could successfully evoke the production of type-specific antibody in a rabbit model. The original monovalent fusion molecule has also been shown to provide type-specific protection against systemic GAS infection. All these studies showed that the induced humoral immunity could efficiently promote adaptive immune clearance of the GAS bacteria⁷⁹.

A 26 amino acids multivalent formulation vaccine candidate has been shown to include 85-90% of serotype associated with pharyngitis and other invasive diseases using the data available from a multitude of North American epidemiological studies.^{80, 81} Besides, a novel vaccine candidate Septa B.2 was designed to include the streptococcal protective antigen Spa, and it had been shown to protect against iGAS infection in a murine model.⁸² Also, the growth of 84% of the represented serotypes was inhibited by rabbit antiserum raised against the multivalent formulation with the non-immune human whole blood supplementation.⁸¹ Interestingly, passive immunization of whole blood promoted opsonophagocytosis of several M4 GAS isolates *in vitro*, and no human cross-reactive antibodies were detected.⁸¹ These results suggest that the functional and cross serotype humoral immunity can be induced by the 26-valent formulation. Phase I and phase II human trials mostly repeat the earlier animal study results with strong antibody titers reported for the majority of the represented serotypes.^{81, 83} The 26-valent vaccine

is the only vaccine candidate that has reached phase II clinical trials. Since then the formulation changed to include the most prevalent serotypes in Europe in order to increase the geographical coverage.⁸⁴ The new 30-valent vaccine candidate has been shown to fight against several serotypes that were not included within the formulation in addition to the majority of the represented serotypes.^{84, 85}

In summary, the multivalent vaccine candidate could theoretically facilitate opsonophagocytosis and provide a broad range of protection against invasive GAS infections. This requires the production of several multivalent vaccine candidates with a wide range of *emm* type coverage based on geographical differences in serotype prevalence.⁸⁶ However, there are other problems which need to be addressed to achieve an effective vaccine candidate, such as the difference of *emm* types between the developed world and the developing world^{87, 88}, the mutation propensity of the M protein in response to adaptive immune pressure, and induction of unusual serotypes due to the vaccination. All these problems will exert limitations on the multivalent formulation.

The highly conserved C terminus is a very attractive target for various GAS vaccine formulations, even though the N terminus of the M protein has the most immunogenic epitopes which are expressed on the bacterial cell surface. The C repeat region shares many conserved sequences among different M protein isoforms.^{89, 90} In the late 1980s, it has been shown that the C repeat formulation vaccine candidate can reduce pharyngeal colonization through the generation of mucosal IgA in a mouse model.^{91, 90} However, it could not induce streptococcal phagocytosis and provide routine protection during systemic GAS challenges.^{90,91} It has been shown that the lack of T cell epitopes within the C terminal region led to the incapability of the C formulation to generate a comprehensive functional immune response.⁹² Immunoglobulin class

switching and the generation of memory B cells require the activation of T helper cells during the adaptive immune response.^{93,94} The B cell receptor recognizes and internalizes the protein antigen. The internalized protein antigens will be degraded and returned to the B cell surface as MHC class II associated peptides. Antigen-specific helper T cells will then recognize the peptide : MHC complexes, and provide the co-stimulatory signal required for B cell proliferation.⁹⁴ The induction of B cell proliferation and the production of memory B cells are often not presented with small peptide antigens due to the lack of natural T cell epitopes.^{93,94} Therefore, it is often presented with a weak and short-lived humoral response when vaccinated with small peptides. Usually, the weak immunogenic molecules are often conjugated to “carrier” molecules in order to provide the co-stimulation required for efficient B cell proliferation.^{94,95}

Alternatively, we can select peptides containing naturally occurring B and T cell epitopes so that it will stimulate B cell and T helper cell responses simultaneously.⁹² These formulations can have the added-value and facilitate simultaneous priming of a cognate T cell response because the T cell epitopes present in the vaccine are derived from the antigen itself.^{92,95} It will also facilitate rapid co-stimulation of the humoral immune response following natural infection because it has the same T cell epitope with the challenge serotype.⁹² However, it has been shown that C terminal peptides containing T cell epitopes can generate heart cross-reactive T cells that could potentially stimulate rheumatic carditis *in situ*.⁹² Subsequently, this approach was not considered as an optimal strategy for the development of the M protein vaccine.

The current J8 C-repeat vaccine formulation is comprised of a minimal B cell epitope, which conjugates to a diphtheria toxin carrier.⁹⁶ The 12 aa peptide is contained within a 16 aa sequence in order to maintain the coiled-coil structure of the molecule and avoid the human cross-reactive T cell epitopes.⁹⁶ However, the T helper cell response will come solely from naïve

T cells activation due to the lack of a dominant T cell epitope.⁹⁵ Despite that, J8 can still stimulate a long-lasting memory B cell response against invasive GAS challenges as it has shown in a murine model.^{95,97}

2. Non-M protein-based vaccine strategies

Many alternative vaccine candidate antigens have been investigated for the development of new vaccine formulations, due to the inherent problems associated with M protein-based vaccines. Initially, scientists mainly resorted to applying well-characterized extracellular virulence factors, such as C5a peptidase, Group A Streptococcus carbohydrate antigens, and the fibronectin binding proteins, for the non-M protein vaccine formulations.^{98, 99, 100} With the advent of whole-genome sequencing era, genome sequencing can identify potential targets *in silico* by common structures in their amino acid sequence and produce the target for evaluation by recombinant expression.^{101, 102, 103} The reverse vaccinology can screen a large number of antigens very efficiently. However, it can be very laborious, and the uncertainty of the immunological function of the target can further complicate the process.

C5a peptidase (ScpA), which is an extracellular endopeptidase, can cleave the soluble complement chemotaxin C5a to inhibit phagocyte recruitment.¹⁰⁴ ScpA is not only expressed on the surface of all GAS serotypes, but also in the group B, G and C streptococci.^{105, 106} It has been shown that elimination of scpA in an intradermal airsac model can facilitate the phagocyte influx and clearance of the invasive GAS organism.¹⁰⁷ Besides, ScpA disruption has also been shown to slow down the murine nasopharyngeal colonization. Thus, it can protect against superficial mucosal infection and invasive GAS disease.⁹⁸ In a murine model, intranasal and parenteral immunization with recombinant ScpA can evoke mucosal IgA and serum IgG production and

prevent nasopharyngeal colonization.^{98, 105} In addition to antibody production, the vaccination induces humoral immunity to inhibit cleavage of C5a, which may restore phagocyte recruitment during the infection.¹⁰⁵ This finding suggests that the leukocyte chemotactic activity of C5a may be restored by naturally acquired anti-ScpA antibodies *in vitro*.¹⁰⁸

The *Streptococcus pyogenes* cell envelope protease (SpyCEP), which is a conserved serine protease, can cleave the neutrophil chemotaxin IL-8 and help to facilitate streptococcal immune evasion.^{109, 110} It has been shown that heterologous expression of SpyCEP in a murine model facilitates bacterial dissemination and increase its ability to exist in the upper respiratory tract. These results suggest that SpyCEP may promote bacterial colonization in soft tissues and the oropharynx.¹¹¹ Besides, it has also been reported that the SpyCEP activity and the recorded severity of disease outcome are positively correlated, which suggests that SpyCEP may indicate the severity of the human diseases.¹¹² The previous observations also confirm these results and suggest that the attenuation of disease severity might be influenced by the SpyCEP activity.^{113, 109}

A CEP-5 vaccine candidate comprising the full-length SpyCEP peptide has been developed, and it has been shown that the vaccine can inhibit GAS dissemination and reduce disease severity in a murine model.¹¹² It also has been reported that serum antibody derived from CEP-5 antigen can inhibit the IL-8 cleaving activity, which suggests that the SpyCEP based vaccine and the ScpA based vaccine have a similar mode of action. The protection functionality in both cases is mediated by phagocyte chemotaxis and bacterial opsonization. It has been reported that during average cell turnover, SpyCEP is shed from the GAS cell surface and can inhibit neutrophil chemotaxis.^{110, 114} These results support the protection model described in the previous research. The fibronectin binding proteins can promote bacterial colonization, and it has been reported that a wide range of protection against GAS infection can be provided by

neutralizing the fibronectin-binding proteins' ligand-binding activity.^{115, 116,117} The fibronectin binding proteins are highly conserved which makes them a very attractive vaccine candidate that could offer a broad range of protection against GAS infection.^{118,119} It has been shown that the well-characterized fibronectin binding proteins based vaccine candidates can induce a strong immune response against systemic GAS challenges.^{115, 116, 117,120}

3. The group A carbohydrate

The cell surfaces oligosaccharides of the bacteria, parasites, and viruses are usually very different from those of their hosts. Various pathogens utilize carbohydrates to escape the host immune response and establish an infection. The cell surface oligosaccharides are also shown to prevent complement activation and phagocytosis inhibition.¹²⁸ The functionalities of the cell surface carbohydrates of the pathogens include the provision of protective shields for their conserved viral protein epitopes,^{129,130,131,132} interaction with host cell surface receptors to facilitate viral spread,^{133,134,135} and induction of immunosuppressive responses,¹³⁶ etc. The cell surface carbohydrate, therefore, is a very attractive target for developing promising vaccine candidates.

Group A Streptococcus express the group A carbohydrate antigen (GAC) on the cell surface. GAC comprises polysaccharide repeating units of *N*-acetylglucosamine coupled to a polyrhamnose backbone^{121,122}. It has been shown that GAS polysaccharides have a conserved and constant expression pattern, which makes them an attractive vaccine antigen for the development of new vaccine candidates. Isolated or synthetic GAC antigens have been utilized to formulate various vaccine candidates, which have been shown to confer protection in relevant mouse models. The analysis results support that high anti-GAS-PS antibody titers inversely correlated with GAS infection cases within high-risk populations. Pinto *et al.* investigated a

chemically synthesized core carbohydrate antigen for glycoconjugate GAS vaccine candidate and evaluated the efficiency of synthetic antigens with the comparison of the isolated GAS polysaccharide. The impact of antigen length and terminal residue composition has been explored and investigated in synthetic studies.

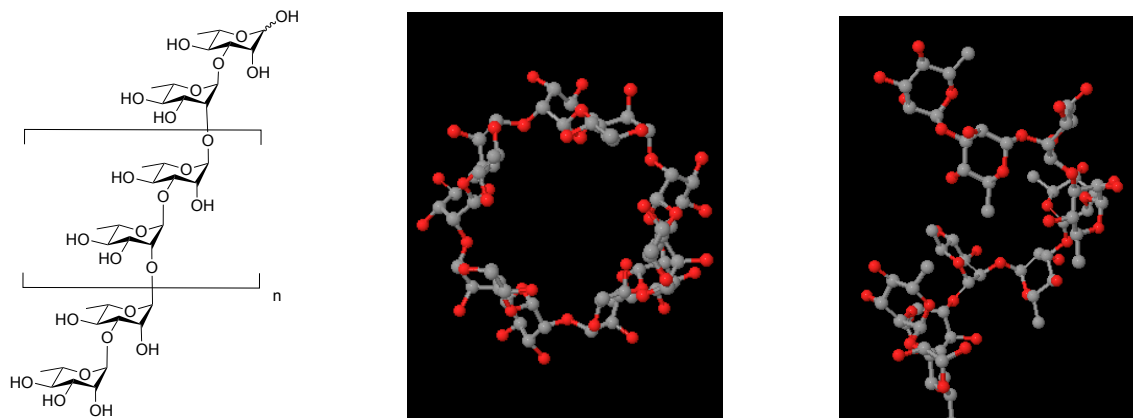


Figure 3. Chemical structure and computer model structures of the side chain deficient GAC

Traditionally, most of the carbohydrate antigens for vaccine development were isolated from biological sources. Powerful analytical instruments have enabled precise elucidation of the bacterial carbohydrate structure. The chemical nature of the native polysaccharides determined the serotype of the organism. The protective immunity against bacteria cell surface oligosaccharides is usually serotype-specific, and multivalent formulations with several types of the polysaccharides are preferred for the vaccine research. In fact, GAS is the sole species of Lancefield group A, and it expresses the conserved pattern of the GAC polysaccharide. The vaccine, therefore, can be formulated to contain the only serotype carbohydrate antigen to confer broad range protection against various GAS variants.

In contrast to the conventional polysaccharide isolation, the synthetic approach could generate pure, homogeneous, and well-defined oligosaccharides, which contains a single reactive group for covalent conjugation. Based on the core antigenic determinants of long polysaccharide

chains, we could design synthetic oligosaccharides to evaluate how particular structural features, such as length, non-reducing end residue, and the composition, influence carbohydrate immunogenicity. Furthermore, we could synthesize well-defined substructures to study structure-activity relationships or to determine a minimal epitope for eliciting immune responses. The major challenge for the development of synthetic carbohydrate vaccines is the chemical synthesis of oligosaccharides, which is highly labor and resource intensive, expensive and laborious process. With the advent of the automated synthesis technology and glycan microarrays, the development of various vaccines will be facilitated dramatically, and the field will be blooming for many efficient and cost-effective vaccines in the following decades.

Carbohydrates are poor immunogens since, with only a few exceptions, they cannot be presented by MHC-antigen complex to the T-helper cells. Thus, most of them cannot induce a long-lived T-cell dependent immune response. It is, therefore, not an ideal vaccine formulation.¹³⁹ This problem can be circumvented by conjugating the carbohydrate antigen to a carrier protein. The glycoconjugate vaccine can induce the T-cell-dependent immune response, and the “class-switching” mechanism will ensure to induce the B-cell response and generate memory B cells and antibodies that not only recognize the carrier protein portion of the vaccine but also recognize the carbohydrate portion of the vaccine. The discovery of the glycoconjugate vaccines originated from the 1930s; it has been shown that the carbohydrate-protein conjugates are capable of inducing an enhanced immune response in adults and children.

Carbohydrate antigens have been shown to conjugate with various carrier proteins, such as tetanus toxoid, diphtheria toxoid and CRM197.^{140,141,142,143} Interestingly, researchers have proposed that short peptides and proteins incorporated with CD4⁺ T cell epitopes have better carrier properties.^{145,146} The glycoconjugates technology appears to be generic and can provide

broad range of protection against various infections. As described before, carbohydrate antigens can only induce short-lived T-cell independent antibody response.¹²³ In fact, anti-GAC IgG can be detected in the sera from healthy, recovered children, which indicates that the class switching occurred during the GAS infection. It is possible that both GAC T cell epitopes and the protein antigens are co-expressed on the cell surface during the infection.¹²⁴ It has been shown that immunization with synthetic CAC-TT glycoconjugate vaccine can offer protection against GAS challenge in a mouse model.^{124,125} However, recent research studies indicated that immunization with the GAC antigens could lead to the generation of human cross-reactive antibodies that also recognize the heart tissue and result in the inflammation associated with rheumatic carditis.¹²⁶ Crude GAC purification methods that have been mentioned above may cause the co-purification of proteinaceous antigens, such as the M protein.¹²⁷ This raised concerns about the validity of the protection data and the purported human cross-reactive antigen due to the potential influence of the M protein that was presented within the GAC preparations.¹²⁷

The synthetic glycoconjugate vaccine formulated by coupling chemically synthesized GAC oligosaccharides to the CRM197 was developed in recent years.⁶ It has been shown that this synthetic glycoconjugate vaccine can provide adequate protection against GAS infection in a mouse model.¹²⁷ The new method for the synthesis of the GAC antigens may resolve the questions about human cross-reactivity antibody, and facilitate the analysis of the glycoconjugate vaccines. Thus, the new technique and theories in carbohydrate chemistry will address the challenges associated with synthesizing these complex molecules. This is the primary work that I have undertaken for my doctoral dissertation. I believe that successful application of cutting-edge technology in carbohydrate chemistry to address the challenging problems for the synthesis

of these complex molecule will facilitate the development of the GAS vaccine and move the research forward into clinical trials.

Carbohydrate chemistry

1. Chemical Synthesis of Complex Carbohydrates

The synthesis of a disaccharide is a representative chemical synthesis of saccharides. The disaccharide can be synthesized by reacting a glycosyl donor equipping with a leaving group with the hydroxyl group of a glycosyl acceptor. This reaction is called glycosylation reaction, which usually can be promoted by an activating system. There are many challenges associated with the synthesis and analysis of complex carbohydrates: 1) The building blocks can form pyranose or furanose rings, therefore, synthetic chemists need to control the form of a specific building block in order to obtain the desired product; 2) The hydroxyl groups of complex carbohydrates intrinsically have similar reactivity, and appropriate protecting groups have to be incorporated into the synthetic strategies for differentiating the reactivity of a specific hydroxyl group at the target position. Therefore, the manipulation of various protecting groups can be a risky and daunting task. It is often required that the introduction and the deprotection of the protecting group have to be highly efficient and orthogonal with each other. 3) The stereo-geometries of the newly formed glycosidic bonds can have two potential epimeric(anomeric) isomers. Thus, it is required to employ an appropriate strategy to assist the desired anomeric configuration formation, especially for certain difficult linkages such as β -mannosidic linkages.

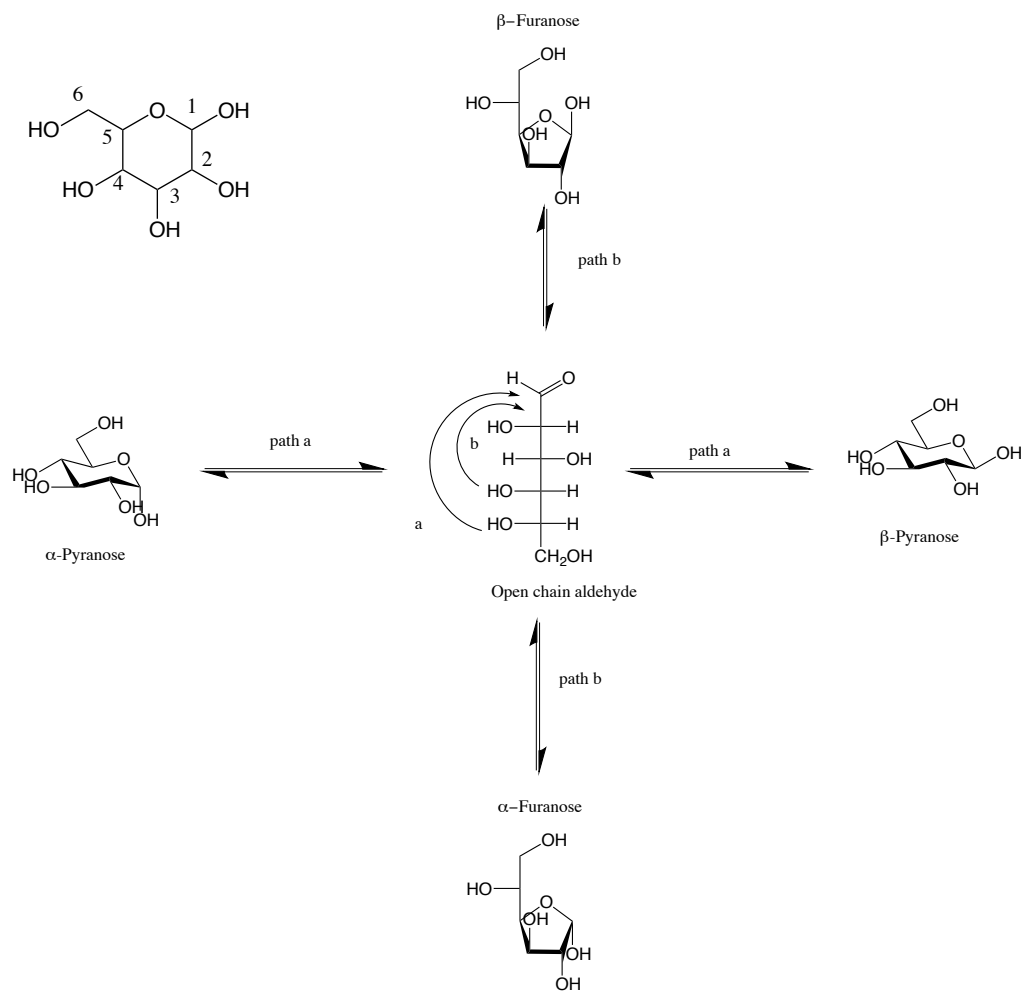


Figure 4. carbohydrate numbering and the ring-closing process.

Chemical synthesis offers many advantages for the potential applications in pharmaceuticals or biomarkers. In contrast to the conventional polysaccharide isolation, the synthetic approach could generate pure, homogeneous, and well-defined oligosaccharides, which contain a single reactive group for covalent conjugation. Based on the core antigenic determinants of long polysaccharide chains, we could design synthetic oligosaccharides to evaluate how particular structural features, such as length, non-reducing end residue, and the composition, influence carbohydrate immunogenicity. Furthermore, we could synthesize well-defined substructures to study structure-activity relationships or to determine a minimal epitope for eliciting immune

responses. Besides, as the technology develops in the chemical industry, sophisticated instruments for the chemically and enzymatically automated synthesis system will become widely available to assist the everyday work of chemists. The automated purification system is facilitating the R&D progress dramatically and can reduce the cost with minimal human supervision. AI and machine learning technologies help scientists to identify novel targets from a vast amount of data that have been generated over the past decades. The industry and academia are collaborating closely, and there will be significant development and synergies across different discipline and fields.

2. A brief review of synthetic strategies for carbohydrate synthesis

Ideally, the desired glycosylation reaction can be performed between two free sugars with complete control of regio/stereochemistry to generate a single oligosaccharide. However, it is rare the case to have such an idealized condition in practice. Thus, scientists have utilized mutated enzymes to facilitate the transformation by using the substrate-specific recognition properties of the enzymes. However, in a practical setting, the availability of the enzyme remains a significant hurdle for the synthesis of various biological active chemicals. Thus, chemistry has its unique advantage and challenges in application.

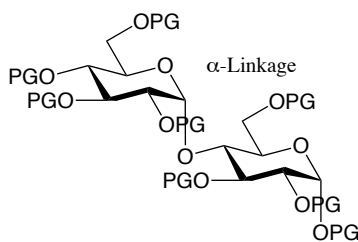


Figure 5. Linkage configuration

In the field of organic chemistry, it has been well established that protecting group strategies can successfully apply to various synthetic tasks. By blocking the functional groups with a set of protecting groups, it ensures that a particular functional group undergoes reaction

without disturbance of other reactants. In particular, carbohydrates present a unique problem as they have many nucleophilic and mildly acidic hydroxyl groups. Different hydroxyl groups interact with each other and are influenced by neighboring functionality. Therefore, modifications such as protecting groups at the hydroxyl can change the reactivity of the other hydroxyl groups. Chemists have devised many strategies to control the reactivity of different hydroxyl groups, and the success in the protecting group strategies leads to numerous chemical synthesis applications not only in the oligosaccharide synthesis but also in the preparation of carbohydrate mimics and natural products.

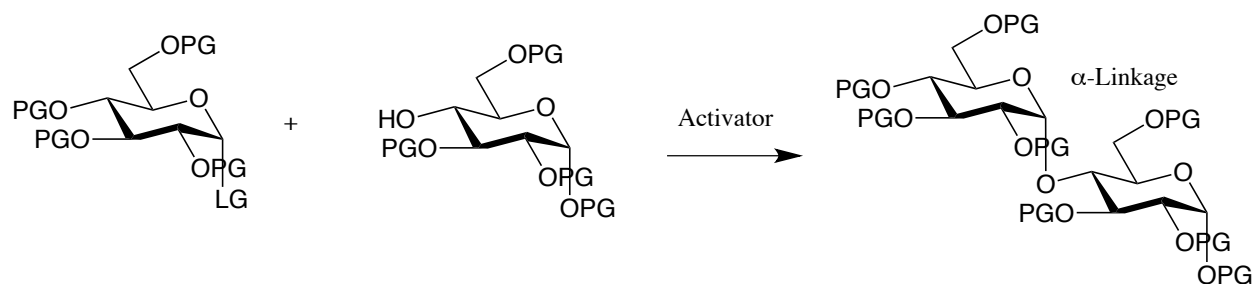


Figure 6. General glycosylation reaction with strategic selection of protecting groups

Judicious selection of protecting groups can reduce the ambiguity in site selectivity during the reactions. The protecting groups can have favorable steric and electronic properties that can change the dynamic of a reaction. Protecting group selection needs to be thoroughly considered based on a comprehensive literature review. The most important factors to consider include 1) a given hydroxyl group needs to be able to protect selectively; 2) the reaction condition needs to be compatible with the selected protecting group; 3) and the protecting group needs to be able to deprotect selectively and conveniently.

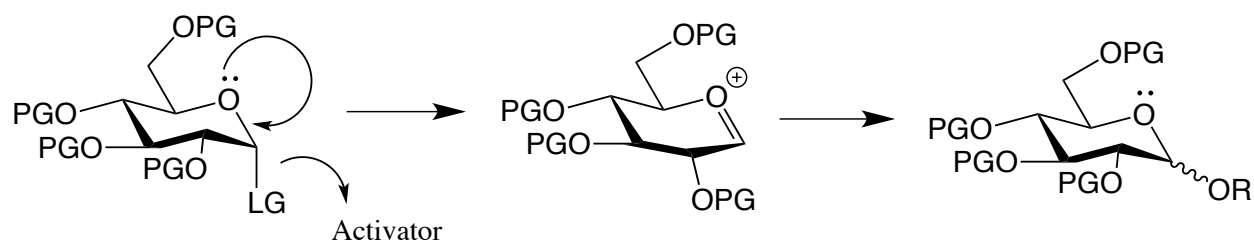


Figure 7. General activation of the glycosyl donor in the glycosylation process

Many leaving groups can be used for glycosylation, and a range of different glycosyl donor systems can be exploited for the synthesis of various glycosides. Commonly used leaving groups to include bromides, chlorides, acetates, and certain less familiar leaving groups, such as fluorides, phenylselenides, or electrophiles, such as 1,2-anhydride/epoxides. Besides those leaving groups, we also need to consider the stability and the amenability of the temporarily masking leaving group at the anomeric center. It will dramatically facilitate the synthesis if the leaving group can perform straightforward chemo-selective transformations with significant stability.¹⁵¹ Therefore, chemists have strong preferences for thioglycosides as the glycosyl donor for oligosaccharide synthesis, as an arylthio group placed at the anomeric center satisfies the requirement. Literature over the past decades have described various methods for activating thioglycosides in detail.¹⁵²⁻¹⁵⁸ More efficient leaving group strategies have been developed over the past decade to facilitate the development of complex oligosaccharide synthesis. Particularly, anomeric acetimidates, *n*-pentenyl groups and vinyl glycosides that contain an unsaturated component at the anomeric center attracted great attention in the chemical strategy of complex carbohydrate synthesis. These leaving groups often have an S_N1 -like fashion to displace the unsaturated component during the glycosylation selectively. Naturally, implementation of these strategies requires a unique condition to maintain the reactivity of the glycosyl donor; on the

other hand, the stability of the donor will be dramatically affected, thus, suffering from hydrolysis, even in the trace moisture environment.

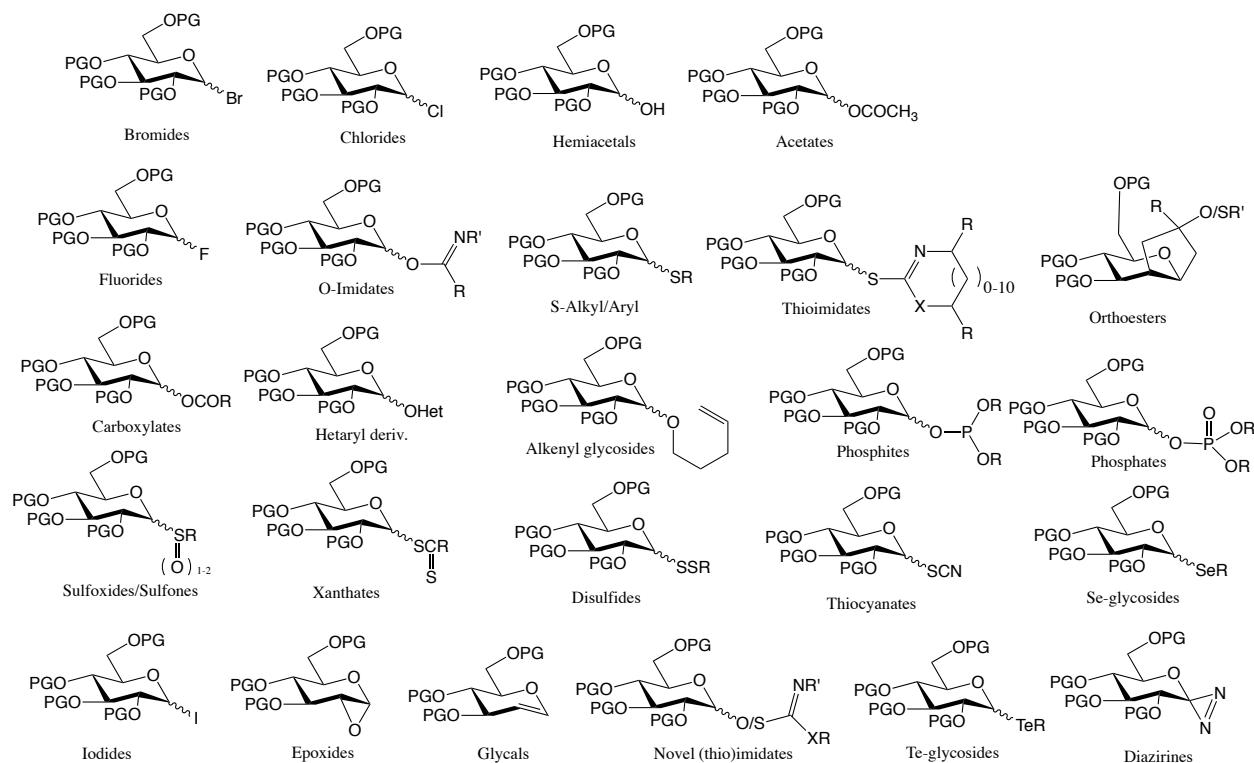


Figure 8. Various leaving groups of glycosyl donors

A successful synthetic strategy design commands the subtle tuning of the reactivity. Chemists developed various strategies to assemble the complex oligosaccharide, such as linear sequence assembly, and more ideally, in a one-pot fashion. Fine-tuning and understanding the reactivity of various glycosyl donor make the one-pot strategy a realizable task. By carefully manipulating the reactivity of the donors, the building blocks can coexist and survive during the sequential glycosylation process. One-pot strategy can generate a fairly complex carbohydrate in a semi-controlled fashion efficiently. However, the main drawbacks of this strategy are the tedious purification process and low yield. The reactivity of the donor is dependent on both the

leaving group and the protecting group. Chemists also need to take the solvent effects into considerations. The first step of the glycosylation forms a cyclic oxonium ion upon activating the leaving group by the promoters. This is the critical determinant step for the glycosylation, and the oxonium ion formation is often influenced by torsional and electronic effects of the protecting groups. Therefore, it is critical to select an efficient leaving group and appropriate reaction conditions to stabilize the oxonium intermediate.

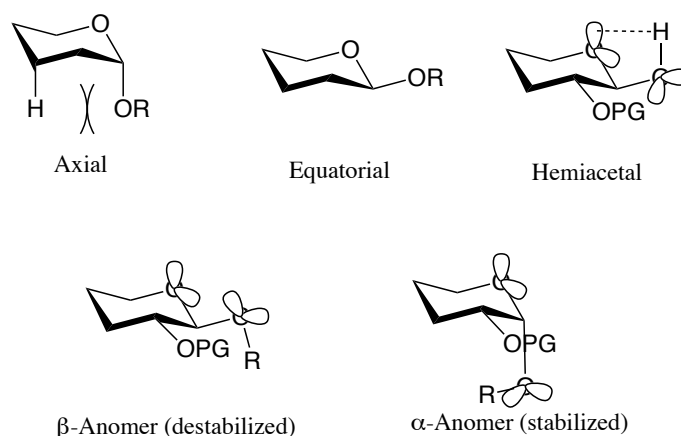


Figure 9. Anomeric effects.

Fraser-Reid's group reported the armed/disarmed strategy for fine-tuning of the reactivity.¹⁵⁹⁻¹⁶¹ He firstly described that oxidative hydrolysis of the n-pentenyl glycoside with NBS in water is much faster when the C-2 protecting group was an ether than an ester. The concept is not limited to the n-pentenyl glycoside only, and it can also apply to thioglycosides and selenoglycosides. Later, Danishefsky and coworkers introduced the glycals for fine-tuning of the reactivity.^{162, 163} Ley *et al.* and Wong *et al.* further investigated glycosyl donor reactivity in various positions of the saccharide and utilized different strategies to design oligosaccharide synthesis.¹⁶⁴⁻¹⁶⁶ Extensive research has been conducted to determine the electronic effects during the glycosylation process. The deactivating power follows this order generally: $N_3 > OAcCl$

(chloroacetate) > NPhth (phthalimido) > OBz > OAc > NHTroc > OBn > OH > OSilyl > H.

Torsional strain can account for another important factor to influence the reactivity of the anomeric leaving groups. The formation of the cyclic oxonium ion flattens the 4C_1 chair conformation in pyranosides, which leads to low reactivity, especially in the bicyclic glycosyl donors.

An alternative approach for the glycosylation includes the concept of orthogonal glycosylation. In this case, glycosyl donors bear different leaving groups, which can be selectively activated. Therefore, the synthetic strategy can be designed in a way that involves minimal glycosylation manipulations. It is crucial to choose the leaving groups strategically, which requires to consider latency-oriented or condition-directed methods.

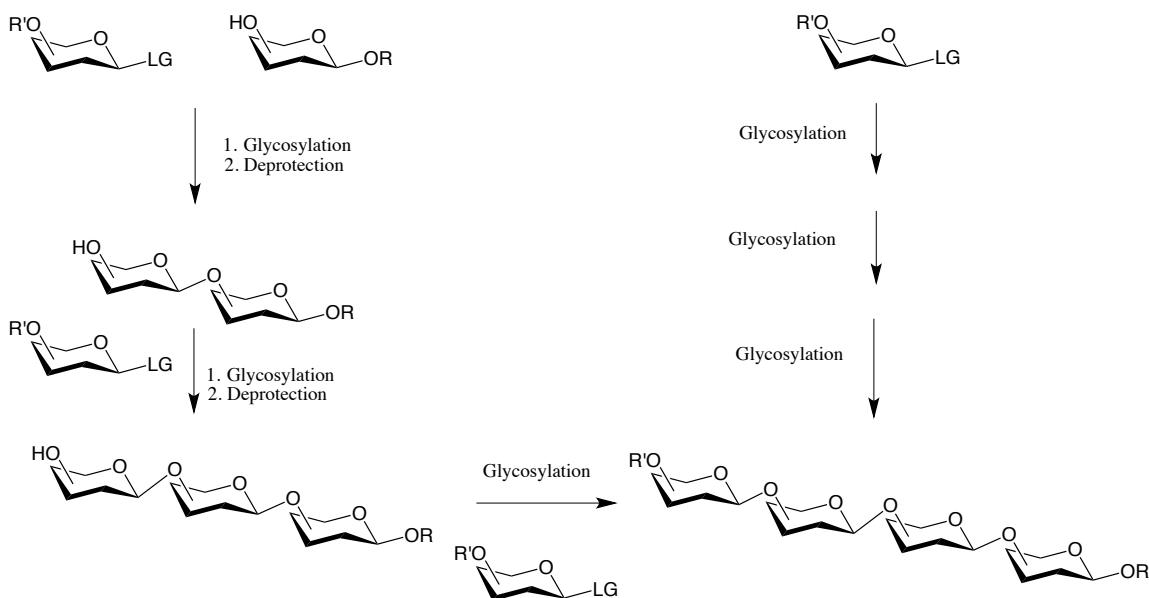


Figure 10. Oligosaccharide assembly strategies.

Various synthetic strategies have been developed and applied to different kinds of complex oligosaccharides synthesis, including β -mannose, Sialic Acid containing carbohydrates, glycosyl

amino acids, C-linked glycosides, etc. These complex carbohydrates play a critical role in many biological processes, such as cell growth, differentiation, and adhesion, cell-mediated immune response and oncogenesis, etc. Therefore, it is critical to employ effective and efficient synthetic strategies for the synthesis of these important complex carbohydrates. Many research groups have focused on developing methodologies for these oligosaccharides synthesis, and it is a trend to incorporate chemical and biological methods to facilitate the development of these fields. Many potential applications of these biologically active materials in pharmaceutical and vaccine have drawn much attention not only in academia but also in the pharmaceutical R&D sector.

The chemical strategy to the synthesis of complex carbohydrate requires chemists to consider the reactivity, regio-selectivity, and the stereo-selectivity. The tedious protection and deprotection process are mandated for the chemical synthesis, and the purification process often generates wastes, and the recovery efficiency needs to be considered during the manufacturing process. In order to overcome the problems associated with the chemical approaches, chemists resort to nature's product, i.e. the uses of glycoprocessing enzymes. There are two main classes of enzymes in glycobiology, namely glycosyltransferases and glycosidases. The glycosyltransferases construct glycosidic bonds, and glycosidases cleave the bonds. Enzymatic synthesis requires an excellent leaving group from the anomeric center. For glycosyltransferases, the leaving groups are either nucleotide phosphate esters or phosphate groups. The enzyme recognizes the reactant or substrate with high precision through the so-called binding pocket. The glycosylation mediated by glycosyltransferases is highly stereo-selective, as the shape of the enzyme active site and the binding pocket have a specific mechanism to form a particular type of bond. The advantages of the enzymatic glycosylation are self-evident, which include high specificity and efficiency, no need of protecting groups, active under mild conditions, and

environmentally friendly. The main limitation of the enzymatic approach is the limited availability of enzymes and the occasional need for careful conditions.

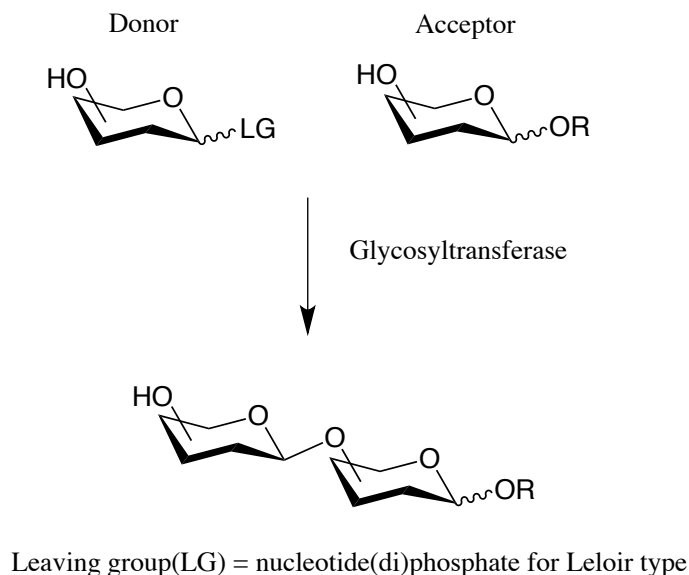


Figure 11. General mechanisms of the glycosyltransferases.

The mechanism of the glycosyltransferases is shown in **Figure 11**. Glycosyltransferases use specific pairs of sugar and nucleotide. The nucleotide diphosphate glycosyl donors are hard to synthesize. The catalytic residues in the active site of the enzyme can attack the hydroxyl group and induce a cascade reaction with the nucleotide glycosyl donor. The intermediate passes through an oxocarbenium-ion-like transition state. Then, donor and acceptor form the new glycosidic bond with certain metal mediation. A potential problem with the glycosyltransferases is product inhibition, and the use of membrane technology may provide a solution. The application of the glycosyltransferases has been shown in several literature reviews.

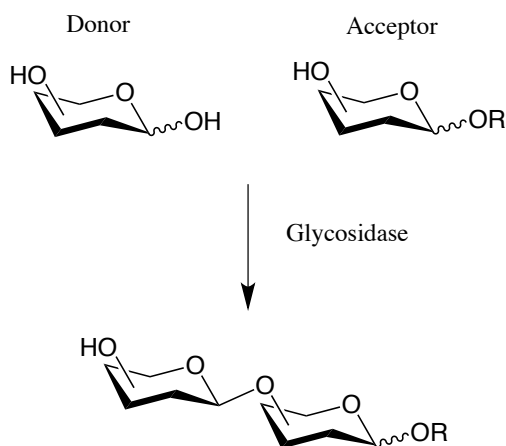


Figure 12. Reverse hydrolysis process with glycosidase.

Reverse hydrolysis process is shown in **Figure 12**. It involves a reaction between a donor and an acceptor to yield a glycoside and H_2O . A high concentration of reactant, high temperature, and the use of organic co-solvents are commonly used to low the thermodynamic activity of water and to facilitate the process. The main limitation of this approach is the stability and the activity of the glycosidase.¹⁶⁷⁻¹⁷⁰

An introduction to the discoveries and proposed mechanisms in this work

We have conducted comprehensive studies to investigate various GAC oligosaccharides with different chain length and side chain GlcNAc variations. We also utilized microarray technology to investigate the epitope recognition of different antibodies toward the GAC library. Over five years of intensive research, we have discovered two new chemical reactions and proposed mechanism for these discoveries that have never been reported in the literature. Here, I will briefly summarize these discoveries and proposed mechanism to give the reader an overview of the significance of this work in terms of chemistry field. The biological evaluations of the synthetic glycoconjugate vaccine candidate and the corresponding glycan microarray results will be described in chapter two and chapter three in detail.

1. PMB Migration.

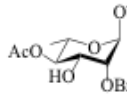
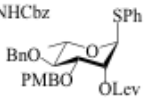
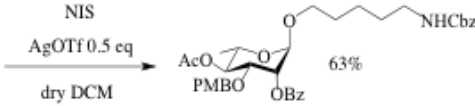
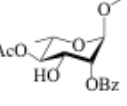
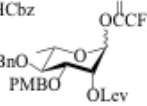
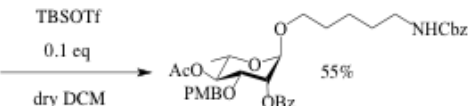
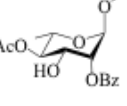
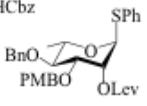
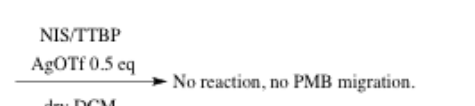
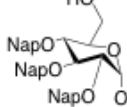
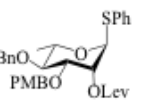
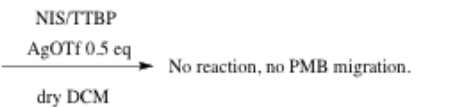
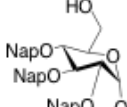
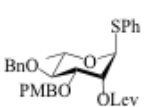
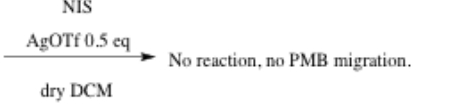
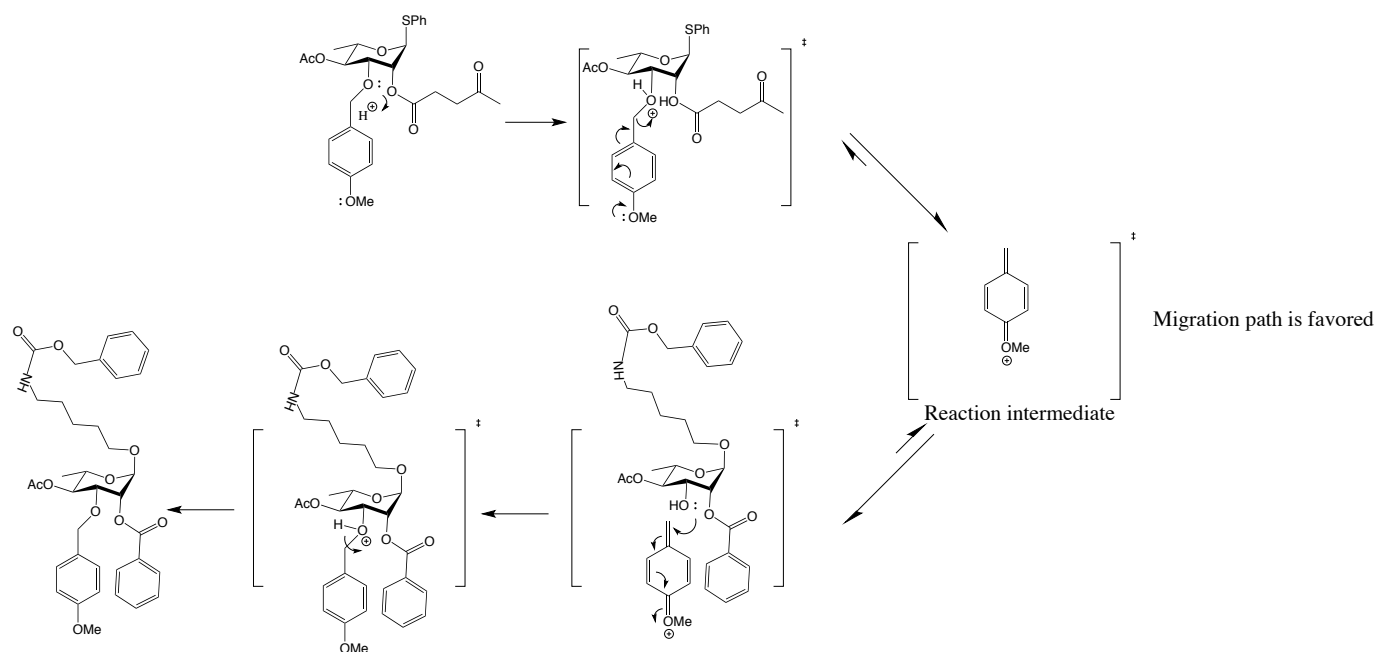
Acceptor 1.0eq	Donor 1.5eq	Reaction condition
 S31	 S16α	 NIS AgOTf 0.5 eq dry DCM -50~0 °C 63% 1
 S31	 S26	 TBSOTf 0.1 eq dry DCM -50~0 °C 55% 1
 S31	 S16α	 NIS/TTBP AgOTf 0.5 eq dry DCM -30~0 °C No reaction, no PMB migration.
 ND	 S16α	 NIS/TTBP AgOTf 0.5 eq dry DCM -25~0 °C No reaction, no PMB migration.
 ND	 S16α	 NIS AgOTf 0.5 eq dry DCM -25~0 °C No reaction, no PMB migration.

Table 2. PMB migration

During the process of examining the optimal protecting group pattern and reaction condition, we discovered the PMB Migration surprisingly. In summary, an *N*-iodosuccinimide (NIS)/silver trifluoromethanesulfonate (AgOTf) promoted glycosylation of **S24** with **S31** and **S14α** with **S31** gave monosaccharide **1** surprisingly. PMB migrates from the donor to the

acceptor, and no desired disaccharide product was formed. It can migrate under very mild conditions, and it is substrate dependent.



Scheme 1. Proposed PMB migration mechanism.

We proposed a mechanism as shown in **Scheme 1**. The PMB protecting group is acid labile due to the p - π conjugation of the lone pair electron on the methoxy oxygen to the π system of the aromatic ring. The migration is probably an equilibrium process, and the migration path is favored assuming that the lower energy of the product is preferred during the process. This requires further support from dynamic molecular simulation and theoretical calculations.

1. The “4+4” *in situ* bond cleavage polymerization.

The second important discovery is the “4+4” *in situ* bond cleavage polymerization. We discovered the new reaction during the assembly process of the polyrhmannose oligosaccharides. In summary, a TBSOTf promoted glycosylation of **9** with **10 α** in DCM at -40 °C~0 °C yielded a library of polyrhmannose oligosaccharides as shown in **Figure 13**. TLC showed a string of new

spots, and high-resolution MALDI-MS of the reaction mixture showed an array of new products. Pure products can be obtained by LH20 and silica gel chromatography purification.

A library of polyramnose oligosaccharides

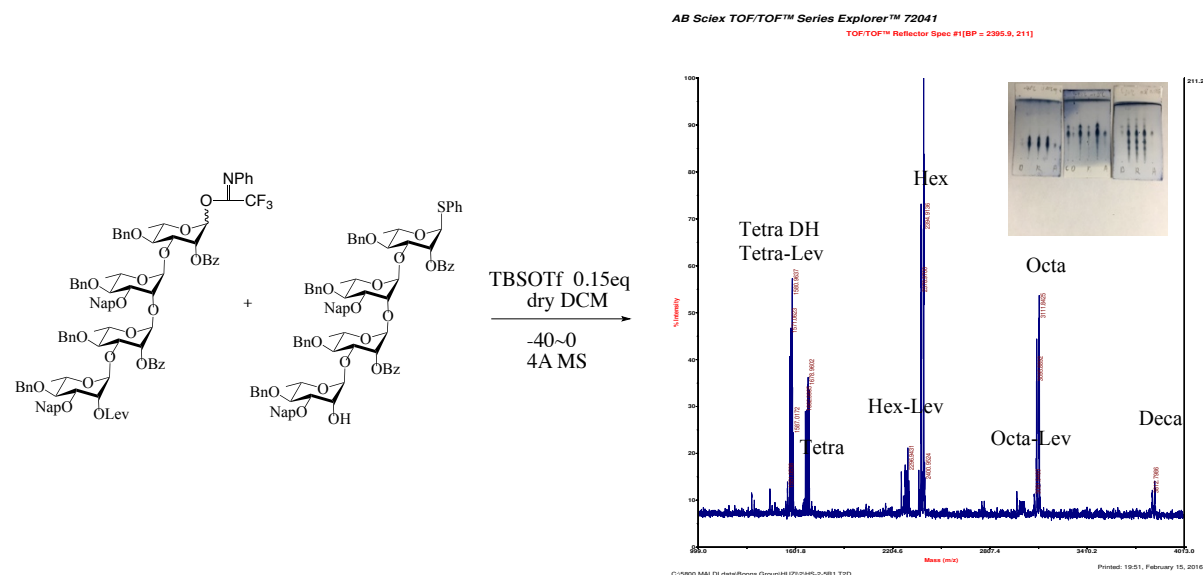
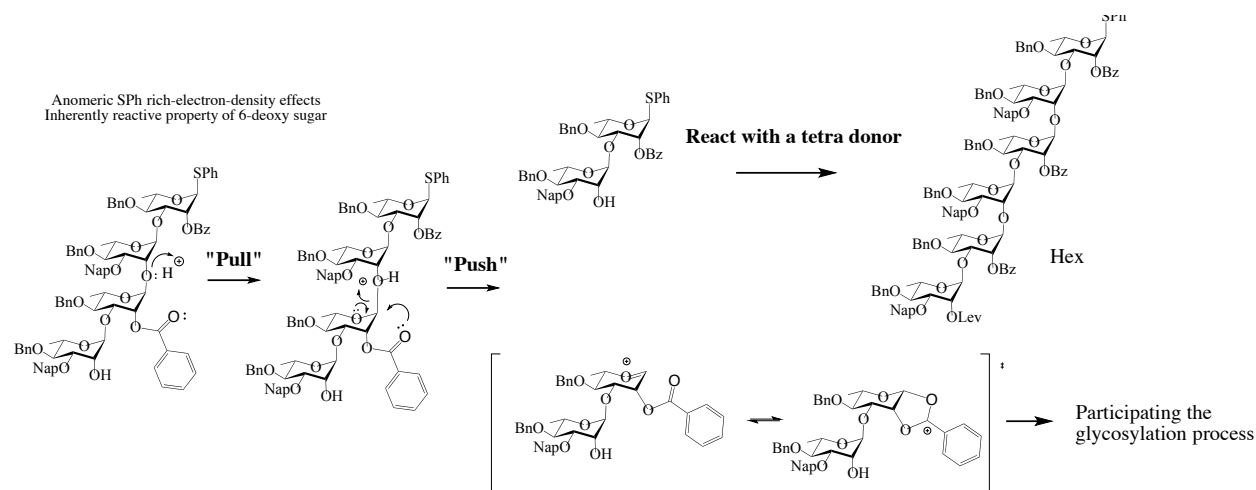


Figure 13. The “4+4” in situ bond cleavage polymerization

We proposed a “Push and Pull” mechanism as shown in **Scheme 2**. The mechanism shows the process of the *in situ* breakage of the acid labile 1,2-linkage to form the new acceptor, which could be further reacted with a tetra donor to yield the new product. Polymerization repeats this process until the reactivity of the substrate is too low to react.



Scheme 2. The proposed “Push and Pull” mechanism for “4+4” polymerization

The proposed mechanism is based on: 1) model acid test suggested 1,3-linkage is acid stable, and 1,2-linkage is acid labile. The peaks corresponding to oligosaccharides without Lev in high-resolution MALDI-MS result from the *in situ* cleavage of the acid labile 1,2-linkage; 2) ester neighbor group participating effects; 3) the electron richness of the thiophenyl protecting group at the anomeric center increases the overall reactivity; 4) the steric hindrance and electron donating properties of the Nap protecting group might increase the breaking tendency of the 1,2-glycosidic linkage.

In summary, two new chemical reactions have been discovered in the process of synthesizing a new efficient GAS vaccine candidate. Detailed descriptions are included in chapter two and chapter three. The details about the biological valuation of the glycoconjugate vaccine and the glycan microarray results will also be covered in the following chapters. The binding study showed that each antibody recognized multiple compounds and exhibited distinct structure–binding relationships. The GAC-microarray data made it possible to investigate and validate the influences of different length and side-chain variations on the structure–binding

relationships with type-specific antibodies. This work will provide important insights into the cross-reactivity associated with the autoimmunity of the GAS infections.

References

- 1) Wikipedia contributors. "Streptococcus pyogenes." *Wikipedia, The Free Encyclopedia*. **2019**. Web. Retrieved 22:06, May 1, 2019.
- 2) Kaplan M.; Bolande R.; Rakita L.; Blair J. Presence of bound immunoglobulins and complement in the myocardium in acute rheumatic fever association with cardiac failure. *The New England Journal of Medicine*. **1964**, 13, 637–45.
- 3) Lozano R.; Naghavi M.; Foreman K.; Lim S.; Shibuya K.; Aboyans V. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. **2012**, 380, 2095–128.
- 4) Andrew S.; Jonathan C.; James D.; John F.; Michael G.; Luiza G.; Nicole M.; Kim M.; Florian S.; Pierre S. Status of research and development of vaccines for *Streptococcus pyogenes*. *Vaccine*. **2016**, 26, 2953–2958.
- 5) Michon F.; Moore L.; Kim J.; Blake S.; Auzanneau I.; Johnston D.; Johnson A.; Pinto M. Doubly branched hexasaccharide epitope on the cell wall polysaccharide of group A streptococci recognized by human and rabbit antisera. *Infectious Immunology*. **2005**, 73, 6383-6389.
- 6) Anna K.; Immaculada M.; Francesco B.; Maria R.; Guido G.; Giuliano B.; Emiliano C.; Daniela P.; Erwin S.; Emilia C.; Paola F.; Daniele C.; Roberto A.; Vittoria P.; David S.; Sabrina C.; Giada B.; Marilena G.; William C.; Stewart C.; John P.; Peter H. S.; Rino R.; Paolo C. Evaluation of a Group A *Streptococcus* synthetic oligosaccharide as vaccine candidate. *Vaccine*. **2010**, 1, 104–114.

- 7) Goldstein I.; Rebeyrotte P.; Parlebas J.; Halpern B. Isolation from heart valves of glycopeptides which share immunological properties with *Streptococcus haemolyticus* group A polysaccharides. *Nature*. **1968**, 219, 866–868.
- 8) Shulman T.; Ayoub M.; Victorica E.; Gessner H.; Tamer F.; Hernandez A. Differences in antibody response to streptococcal antigens in children with rheumatic and non-rheumatic mitral valve disease. *Circulation*. **1974**, 50, 1244–1251.
- 9) Appleton S.; Victorica E.; Tamer D.; Ayoub M. Specificity of persistence of antibody to the streptococcal group A carbohydrate in rheumatic valvular heart disease. *Journal of Laboratory and Clinical Medicine*. **1985**, 105, 114–119.
- 10) Galvin E.; Hemric E.; Ward K.; Cunningham W. Cytotoxic mAb from rheumatic carditis recognizes heart valves and laminin. *The Journal of Clinical Investigation*. **2000**, 106, 217–224.
- 11) Sumby P., Whitney R., Graviss A., DeLeo R., and Musser M. Genome-wide analysis of group a streptococci reveals a mutation that modulates global phenotype and disease specificity. *PLoS Pathogens*. **2006**, 10, 1371.
- 12) Anna R., Marea F., Rosalie S., Dale T., Michelle D., Tom C., Matthew P., Michele C., Malcolm M., Keith E., Jonathan C. and Ross B. Improvement in rheumatic fever and rheumatic heart disease management and prevention using a health centre-based continuous quality improvement approach. *BMC Health Services Research*, **2013**, 13, 525.
- 13) Cunningham W. Pathogenesis of group A streptococcal infections. *Clinical microbiology reviews*, **2000**, 10, 1128.
- 14) Slade H. and Slamp W. Cell-wall composition and the grouping antigens of Streptococci. *Journal of bacteriology*, **1962**, 84, 345-351.

- 15) Krause R. and McCarty M. Studies on the chemical structure of the streptococcal cell wall. *Journal of Experimental Medicine*, **1961**, 114, 127.
- 16) Lancefield R. The antigenic complex of *Streptococcus haemolyticus*. *Journal of Experimental Medicine*, **1928**, 47, 91.
- 17) Facklam R. What happened to the streptococci: overview of taxonomic and nomenclature changes. *Clinical microbiology reviews*, **2002**, 10, 1128.
- 18) Beall B.; Facklam R. and Thompson T. Sequencing emm-specific PCR products for routine and accurate typing of group A streptococci. *Journal of clinical microbiology*, **1996**, 953-958.
- 19) Andrew S.; Irwin L.; Laisiana M.; Bernard B. and Jonathan C. Global emm type distribution of group A streptococci: systematic review and implications for vaccine development. *The Lancet infectious diseases*, **2009**, 10, 611-616.
- 20) Bogdan L.; Jessica D.; Shona N.; Tuula S.; Lenka S.; Asha T.; Roberta C.; Kim E.; Maria K.; Panayotis T.; Mark L.; Monica S.; Jaana V.; Anne B.; Androulla E.; Claes S.; Birgitta H. and the Strep-EURO Study Group. Clinical and microbiological characteristics of severe *Streptococcus pyogenes* disease in Europe. *Journal of clinical microbiology*, **2009**, 1155-1165.
- 21) Bart V.; Ellen M.; Joop S.; Leo S.; Armand P.; Ad F.; Rodger N.; Jan V. and Franz S. Site-specific manifestations of invasive group A streptococcal disease: type distribution and corresponding patterns of virulence determinants. *Journal of clinical microbiology*, **2003**, 4941-4949.
- 22) Anna R.; Marea F.; Rosalie S.; Dale T.; Michelle D.; Tom C.; Matthew P.; Michele C.; Malcolm M.; Keith E.; Jonathan C. and Ross B. Improvement in rheumatic fever and rheumatic heart disease management and prevention using a health centre-based continuous quality improvement approach. *BMC Health Services Research*, **2013**, 13, 525.

- 23) Daniel L.; Stulberg M.; Penrod and Richard B. Common Bacterial Skin Infections. *American Family Physician*, **2012**, 66, 1.
- 24) Oshri W.; Lea V.; Eyal K.; Yael B.; Nadav D.; Nadav O.; Zina K.; Raid K.; Tamar S.; Ruhama A.; Estela D.; Ron D. and Salman Z. A Cluster of Ecthyma Outbreaks Caused by a Single Clone of Invasive and Highly Infective *Streptococcus pyogenes*. *Clinical Infectious Diseases*, **2009**, 48-9, 1213-1219.
- 25) Bessen D. Population biology of the human restricted pathogen, *Streptococcus pyogenes*. *Infection, Genetics and Evolution*, **2009**, 9-4, 581.
- 26) Wannamaker L. The epidemiology of streptococcal infections. *Streptococcal infections*. **1954**, 157–175.
- 27) Gunderson C. and Martinello R. A systematic review of bacteremias in cellulitis and erysipelas. *Journal of Infection*, **2012**, 64-2, 148-155.
- 28) Bisno A. and Stevens D. Streptococcal infections of skin and soft tissues. *New England Journal of Medicine*, **1996**, 334, 240-246.
- 29) Zimbelman J.; Palmer A. and Todd J. Improved outcome of clindamycin compared with beta-lactam antibiotic treatment for invasive *Streptococcus pyogenes* infection. *The Pediatric infectious disease journal*. **1999**, 18-12, 1096-1100.
- 30) Shiranee S.; Melissa F.; Victoria E.; Lee F.; Jonathan C. Human intravenous immunoglobulin for experimental streptococcal toxic shock: bacterial clearance and modulation of inflammation. *Journal of antimicrobial chemotherapy*. **2006**, 58-1, 117–124.
- 31) Reglinski M. and Sriskandan S. The contribution of group A streptococcal virulence determinants to the pathogenesis of sepsis. *Virulence*, **2014**, 127-136.

- 32) Angus D.; Helterbrand J. and Levy H. Drotrecogin alfa (activated) administration across clinically important subgroups of patients with severe sepsis. *Critical care*, **2003**, 31, 12-9.
- 33) Breiman R.; Davis J.; Facklam R.; Gray B. Defining the group A streptococcal toxic shock syndrome: rationale and consensus definition. *Jama*, **1993**, 269, 390-1.
- 34) Lappin E. and Ferguson A. Gram-positive toxic shock syndromes. *The Lancet infectious diseases*, **2009**, 9, 725-6.
- 35) Murphey E.; Fang G.; Sherwood E. Pretreatment with the Gram-positive bacterial cell wall molecule peptidoglycan improves bacterial clearance and decreases inflammation and mortality in mice challenged with *Staphylococcus aureus*. *Critical Care Med*, **2008**, 3067-3073.
- 36) Shulman S.; Bisno A.; Clegg H. Clinical practice guideline for the diagnosis and management of group A streptococcal pharyngitis: 2012 update by the Infectious Diseases Society of America. *Clinical infectious diseases journal*. **2012**, 55, 1279-82.
- 37) Barash J. Rheumatic Fever and post-group a streptococcal arthritis in children. *Current infectious disease reports*, **2013**, 15, 263-8
- 38) Bisno A.; Rubin F.; Cleary P. Prospects for a group A streptococcal vaccine: rationale, feasibility, and obstacles—report of a National Institute of Allergy and Infectious Diseases workshop. *Clinical Infectious diseases journal*. **2005**, 41, 1150-6.
- 39) Quinn A.; Ward K.; Fischetti V.; Hemric M. Immunological relationship between the class I epitope of streptococcal M protein and myosin. *Infection and immunity*. **1998**. 66-9, 4418–4424.
- 40) McNamara C.; Zinkernagel A.; Macheboeuf P. Coiled-coil irregularities and instabilities in group A *Streptococcus* M1 are required for virulence. *Science*, **2008**, 319, 1405-8.
- 41) Dudding B. and Ayoub E. Persistence of streptococcal group A antibody in patients with rheumatic valvular disease. *Journal of Experimental Medicine*, **1968**, 128, 1081-98.

- 42) Courtney H.; Hasty D.; Dale J. Anti-phagocytic mechanisms of *Streptococcus pyogenes*: binding of fibrinogen to M-related protein. *Molecular microbiology*, **2006**, 59, 936-47.
- 43) Dale B. and Fredericks L. Antimicrobial peptides in the oral environment: expression and function in health and disease. *Current issues in molecular biology*, **2005**, 7, 119-33.
- 44) Yamasaki K. and Gallo R. Antimicrobial peptides in human skin disease. *European journal of dermatology*, **2008**, 18, 11-21.
- 45) Dorschner R.; Pestonjamas V. Cutaneous Injury Induces the Release of Cathelicidin Anti-Microbial Peptides Active Against Group A *Streptococcus*. *Journal of Investigative Dermatology*. **2001**, 117, 91-7.
- 46) Dale B.; Fredericks L. Antimicrobial peptides in the oral environment: expression and function in health and disease. *Current issues in molecular biology*, **2005**, 7, 119-33.
- 47) Nizet V.; Ohtake T.; Lauth X.; Trowbridge J. and Rudisill J. Innate antimicrobial peptide protects the skin from invasive bacterial infection. *Nature*, **2001**, 414, 454-7.
- 48) Frick I.; Åkesson P.; Rasmussen M. SIC, a Secreted Protein of *Streptococcus pyogenes* That Inactivates Antibacterial Peptides. *Journal of Biological Chemistry*. **2003**, 278, 16561-6.
- 49) Schmidtchen A.; Frick I.; Andersson E. Proteinases of common pathogenic bacteria degrade and inactivate the antibacterial peptide LL-37. *Molecular Microbiology*. **2002**, 46, 157-68.
- 50) Akira S.; Takeda K. Toll-like receptor signaling. *Nature reviews immunology*, **2004**, 4, 499–511.
- 51) Travassos L.; Girardin S.; Philpott D. Toll-like receptor 2-dependent bacterial sensing does not occur via peptidoglycan recognition. *EMBO Reports*. **2004**, 5, 1000-6.
- 52) Fritz J.; Ferrero R.; Philpott D.; Girardin S. Nod-like proteins in immunity, inflammation and disease. *Nature immunology*, **2006**, 7, 1250-7.

- 53) Armstrong L.; Medford A.; Hunter K. Differential expression of Toll-like receptor (TLR)-2 and TLR-4 on monocytes in human sepsis. *Clinical Experimental Immunology*, **2004**, 136, 312–319.
- 54) Lai Y.; Cogen A.; Radek K.; Jeong P.; Daniel T.; Mac L.; Anke L.; Allen F.; Ryan, A.; Di N.; Richard L. Activation of TLR2 by a Small Molecule Produced by *Staphylococcus epidermidis* Increases Antimicrobial Defense against Bacterial Skin Infections. *Journal of Investigative Dermatology*. **2010**, 2211–2221.
- 55) Nakayama H.; Kurokawa K.; Lee B. Lipoproteins in bacteria: structures and biosynthetic pathways. *The FEBS journal*, **2012**, 279(23):4247-68.
- 56) Lynch A.; Murphy J.; Byers T.; Gibbs R. Alternative complement pathway activation fragment Bb in early pregnancy as a predictor of preeclampsia. *American journal of Obstetrics and Gynecology*. **2008**, 198, 385.
- 57) Aderem A.; Underhill D. Mechanisms of phagocytosis in macrophages. *Annual review of immunology*, **1999**, 17, 593-623.
- 58) Bisno A.; Rubin F.; Cleary P. Prospects for a group A streptococcal vaccine: rationale, feasibility, and obstacles—report of a National Institute of Allergy and Infectious Diseases workshop. *Clinical infectious diseases*. **2005**, 41, 1150-6.
- 59) Kapur V.; Maffei J.; Greer R.; Li L.; Adams G. Vaccination with streptococcal extracellular cysteine protease (interleukin-1 β convertase) protects mice against challenge with heterologous group A streptococci. *Microbial Pathogenesis*. **1994**, 16, 443-50.
- 60) Ulrich R. Vaccine based on a ubiquitous cysteinyl protease and streptococcal pyrogenic exotoxin A protects against *Streptococcus pyogenes* sepsis and toxic shock. *Journal of immune based therapies and Vaccines*. **2008**, 6, 8.

- 61) Köhler W.; Gerlach D.; Knöll H. Streptococcal outbreaks and erythrogenic toxin type A. *Bakteriologie, Mikrobiologie und Hygiene*. **1987**, 266, 1-2, 104-115.
- 62) Lancefield R. Current knowledge of type-specific M antigens of group A streptococci. *The Journal of Immunology*, **1962**, 89, 307-13.
- 63) Stegmayr B.; Björck S.; Holm S. and Nisell J. Septic shock induced by group A streptococcal infection: clinical and therapeutic aspects. *Journal of infectious diseases*. **1992**, 24, 589-97.
- 64) Norrby-Teglund A.; Pauksens K. Relation between low capacity of human sera to inhibit streptococcal mitogens and serious manifestation of disease. *Journal of Infectious diseases*. **1994**, 170-3, 585-591.
- 65) Dick G. The etiology of scarlet fever. *Journal of the American Medical Association*, **1924**, 82, 301-302.
- 66) Rantz L.; Randall E.; Rantz H. Immunization of human beings with group A hemolytic streptococci. *The American journal of medicine*, **1949**, 6, 424-32.
- 67) Jackson R. Active immunisation against the streptococcus scarlatinÆ. *Irish Journal of Medical Science*, **1941**, 4, 9.
- 68) Bloomfield A.; Felty A. Bacteriologic observations on acute tonsillitis with reference to epidemiology and susceptibility. *Archives of Internal Medicine*, **1923**, 32, 483-496
- 69) Polly S.; Waldman R.; High P. Protective studies with a group A streptococcal M protein vaccine. II. Challenge of volunteers after local immunization in the upper respiratory tract. *Journal of Infectious diseases*. **1975**, 131, 3, 217–224.
- 70) Moody M.; Siegel A.; Pittman B. Fluorescent-antibody identification of group A streptococci from throat swabs. *American Journal of Public Health*. **1963**, 53, 1083-92.

- 71) Fox E.; Waldman R.; Wittner K. Protective study with a group A streptococcal M protein vaccine. Infectivity challenge of human volunteers. *The Journal of Clinical Investigation*. **1973**, 52, 1885-92.
- 72) Massell B.; Honikman H.; Amezcua J. Rheumatic fever following streptococcal vaccination: report of three cases. *JAMA*. **1969**, 207, 1115-9.
- 73) Stevens D.; Bisno A.; Chambers H. Practice guidelines for the diagnosis and management of skin and soft-tissue infections. *Clinical Infectious Diseases*. **2005**, 15, 59, 10-52.
- 74) Dale J.; Washburn R.; Marques M. Hyaluronate capsule and surface M protein in resistance to opsonization of group A streptococci. *Infection and Immunity*, **1996**, 64, 1495-501.
- 75) Polly S.; Waldman R.; High P. Protective studies with a group A streptococcal M protein vaccine. II. Challenge of volunteers after local immunization in the upper respiratory tract. *Journal of Infectious Diseases*. **1975**, 131, 3, 217-224.
- 76) Lancefield R. Persistence of type-specific antibodies in man following infection with group A streptococci. *Journal of experimental medicine*, **1959**, 110, 271-92.
- 77) Beachey E.; Seyer J.; Dale J. and Hasty D. Repeating covalent structure and protective immunogenicity of native and synthetic polypeptide fragments of type 24 streptococcal M protein. Mapping of protective and nonprotective epitopes with monoclonal antibodies. *Journal of biological chemistry*, **1983**, 258, 13250-13257.
- 78) Beachey E.; Seyer J. Protective and nonprotective epitopes of chemically synthesized peptides of the NH₂-terminal region of type 6 streptococcal M protein. *The Journal of Immunology*, **1986**, 136, 2287-2292;

- 79) Dale J.; Chiang E. Intranasal Immunization with Recombinant Group A Streptococcal M Protein Fragment Fused to the B Subunit of Escherichia coli Labile Toxin Protects Mice against Systemic Challenge Infections. *Journal of Infectious Diseases*, **1995**, 171, 1038-41.
- 80) Palmer K.; Campbell J.; Reddish M.; Hu M. Safety and immunogenicity of a recombinant multivalent group a streptococcal vaccine in healthy adults: phase 1 trial. *JAMA*. **2004**, 292, 709-15.
- 81) McNeil S.; Halperin S.; Langley J. Safety and Immunogenicity of 26-Valent Group A Streptococcus Vaccine in Healthy Adult Volunteers. *Clinical Infectious Diseases*. **2005**, 41, 1114-22.
- 82) Li Y.; Chiang H.; Thacker J.; Dale J. Serum opacity factor is a major fibronectin-binding protein and a virulence determinant of M type 2 Streptococcus pyogenes. *Molecular Microbiology*. **1999**, 32, 89-98.
- 83) McArthur J.; Walker M. Domains of group A streptococcal M protein that confer resistance to phagocytosis, opsonization and protection: implications for vaccine development. *Molecular microbiology*, **2006**, 59, 1-4.
- 84) Dale J.; Penfound T.; Chiang E.; Walton W. New 30-valent M protein-based vaccine evokes cross-opsonic antibodies against non-vaccine serotypes of group A streptococci. *Vaccine*, **2011**, 29, 8175-8.
- 85) Dale J.; Penfound T.; Tamboura B.; Sow S. Potential coverage of a multivalent M protein-based group A streptococcal vaccine. *Vaccine*, **2013**, 31, 1576-81.
- 86) Dale J.; Simmons M.; Chiang E. Recombinant, octavalent group A streptococcal M protein vaccine. *Vaccine*, **1996**, 14, 944-8.

- 87) Steer A.; Law I.; Matatolu L.; Beall B. Global emm type distribution of group A streptococci: systematic review and implications for vaccine development. *The Lancet infectious diseases*. **2009**, 9, 611-6.
- 88) Steer A.; Batzloff M.; Mulholland K. Group A streptococcal vaccines: facts versus fantasy. *Current opinion in infectious diseases*. **2009**, 22, 544-52.
- 89) Smeesters P.; McMillan D.; Sriprakash K. The streptococcal M protein: a highly versatile molecule. *Trends in microbiology*, **2010**, 18, 275-82.
- 90) Pancholi V.; Fischetti V. A major surface protein on group A streptococci is a glyceraldehyde-3-phosphate-dehydrogenase with multiple binding activity. *Journal of Experimental Medicine*, **1992**, 176, 415-26.
- 91) Bessen D.; Fischetti V. Influence of intranasal immunization with synthetic peptides corresponding to conserved epitopes of M protein on mucosal colonization by group A streptococci. *Infection and immunity*, **1988**, 56, 2666-72.
- 92) Davies M.; Zeng M.; Pruksakorn S. Protection against group A streptococcus by immunization with J8-diphtheria toxoid: contribution of J8-and diphtheria toxoid-specific antibodies to protection. *The Journal of Infectious Diseases*. **2003**, 187, 1598-608.
- 93) Stephenson A.; Wu H.; Novak J. The Fap1 fimbrial adhesin is a glycoprotein: antibodies specific for the glycan moiety block the adhesion of *Streptococcus parasanguis* in an in vitro tooth model. *Molecular microbiology*. **2002**, 43, 147-57.
- 94) Sethi S.; Murphy T. Infection in the pathogenesis and course of chronic obstructive pulmonary disease. *New England Journal of Medicine*, **2008**, 359, 2355-65.

- 95) Pandey M.; Wykes M.; Hartas J.; Good M. Long-term antibody memory induced by synthetic peptide vaccination is protective against *Streptococcus pyogenes* infection and is independent of memory T cell help. *The Journal of Immunology*. **2013**, 190, 2692-2701.
- 96) Hayman W.; Brandt E.; Relf W. Mapping the minimal murine T cell and B cell epitopes within a peptide vaccine candidate from the conserved region of the M protein of group A streptococcus. *International Immunology*. **1997**, 9, 1723-33.
- 97) Pandey M.; Batzloff M.; Good M. Mechanism of protection induced by group A *Streptococcus* vaccine candidate J8-DT: contribution of B and T-cells towards protection. *PLoS One*, **2009**, 4, 4, 5147.
- 98) Ji Y.; Carlson B.; Kondagunta A.; Cleary P. Intranasal immunization with C5a peptidase prevents nasopharyngeal colonization of mice by the group A *Streptococcus*. *Infection and immunity*, **1997**, 65, 6, 2080–2087.
- 99) Sabharwal H.; Michon F.; Nelson D. Group A streptococcus (GAS) carbohydrate as an immunogen for protection against GAS infection. *The Journal of Infectious Diseases*. **2006**, 193, 129-35.
- 100) Guzmán C.; Talay S.; Molinari G. Protective Immune Response against *Streptococcus pyogenes* in Mice after Intranasal Vaccination with the Fibronectin-Binding Protein SfbI. *The Journal of Infectious Diseases*. **1999**, 179, 901-6.
- 101) Rodríguez-Ortega M.; Norais N.; Bensi G. Characterization and identification of vaccine candidate proteins through analysis of the group A *Streptococcus* surface proteome. *Nature*, **2006**, 24, 191-7.

- 102) Bensi G.; Mora M.; Tuscano G.; Biagini M. Multi high-throughput approach for highly selective identification of vaccine candidates: the Group A Streptococcus case. *Molecular & Cellular Proteomics*. **2012**, 11, 6, 15693.
- 103) Smeesters P.; McMillan D.; Sriprakash K. The streptococcal M protein: a highly versatile molecule. *Trends in microbiology*, **2010**, 18, 275-82.
- 104) Wexler D.; Chenoweth D. Mechanism of action of the group A streptococcal C5a inactivator. *Proceedings of the National Academy Sciences*, **1985**, 82, 8144–8148.
- 105) Cleary P.; Matsuka Y.; Huynh T.; Lam H. and Olmsted S. Immunization with C5a peptidase from either group A or B streptococci enhances clearance of group A streptococci from intranasally infected mice. *Vaccine*, **2004**, 22, 4332-41.
- 106) Chmouryguina I.; Suvorov A.; Ferrieri P. Conservation of the C5a peptidase genes in group A and B streptococci. *Infection and Immunity*. **1996**, 64, 7, 2387–2390.
- 107) Ji Y.; McLandsborough L.; Kondagunta A. C5a peptidase alters clearance and trafficking of group A streptococci by infected mice. *Infection and Immunity*. **1996**, 64, 503–510.
- 108) O'Connor S.; Darip D.; Fraley K. The human antibody response to streptococcal C5a peptidase. *Journal of Infectious Diseases*. **1991**, 163, 109-16.
- 109) Hidalgo-Grass C.; Mishalian I.; Dan-Goor M. A streptococcal protease that degrades CXC chemokines and impairs bacterial clearance from infected tissues. *The EMBO Reports*. **2006**, 25, 4628–4637.
- 110) Edwards R.; Taylor G. and Ferguson M. Specific C-Terminal Cleavage and Inactivation of Interleukin-8 by Invasive Disease Isolates of Streptococcus pyogenes. *The Journal of Infectious Diseases*. **2005**, 192, 783-90.

- 111) Kurupati P.; Turner C. and Tziona I. Chemokine-cleaving *Streptococcus pyogenes* protease SpyCEP is necessary and sufficient for bacterial dissemination within soft tissues and the respiratory tract. *Molecular Microbiology*. **2010**, 76, 1387-97.
- 112) Turner C.; Kurupati P.; Wiles S. and Edwards R. Impact of immunization against SpyCEP during invasive disease with two streptococcal species: *Streptococcus pyogenes* and *Streptococcus equi*. *Vaccine*, **2009**, 27, 4923-9.
- 113) Zinkernagel A.; Timmer A.; Pence M. and Locke J. The IL-8 protease SpyCEP/ScpC of group A *Streptococcus* promotes resistance to neutrophil killing. *Cell host & Microbe*. **2008**, 4, 170-8.
- 114) Chiappini N.; Seubert A.; Telford J. and Grandi G. *Streptococcus pyogenes* SpyCEP influences host-pathogen interactions during infection in a murine air pouch model. *PLoS One*, **2012**, 7, 7, 40411.
- 115) Guzmán C.; Talay S. and Molinari G. Protective Immune Response against *Streptococcus pyogenes* in Mice after Intranasal Vaccination with the Fibronectin-Binding Protein SfbI. *The Journal of Infectious Diseases*. **1999**, 179, 901-6.
- 116) Terao Y.; Kawabata S. and Kunitomo E. Fba, a novel fibronectin-binding protein from *Streptococcus pyogenes*, promotes bacterial entry into epithelial cells, and the fba gene is positively transcribed under the Mga regulator. *Molecular Microbiology*. **2001**, 42, 75-86.
- 117) Schulze K.; Olive C.; Ebensen T.; Guzmán C. Intranasal vaccination with SfbI or M protein-derived peptides conjugated to diphtheria toxoid confers protective immunity against a lethal challenge with *Streptococcus pyogenes*. *Vaccine*, **2006**, 24, 6088-95.

- 118) Talay S.; Valentin-Weigand P. and Jerlström P. Fibronectin-binding protein of *Streptococcus pyogenes*: sequence of the binding domain involved in adherence of streptococci to epithelial cells. *Infection and Immunity*. **1992**, 60, 3837–3844.
- 119) Talay S.; Valentin-Weigand P. and Timmis K. Domain structure and conserved epitopes of Sfb protein, the fibronectin-binding adhesin of *Streptococcus pyogenes*. *Molecular Microbiology*. **1994**, 13, 531-9.
- 120) Nair S.; Pallas J. and Williams M. Fibronectin: a multidomain host adhesin targeted by bacterial fibronectin-binding proteins. *FEMS microbiology reviews*. **2010**, 35, 147-200.
- 121) Slade H. and Slamp W. Cell-wall composition and the grouping antigens of *Streptococci*. *Journal of bacteriology*, **1962**, 84, 345.
- 122) Krause R. and McCarty M. Studies on the chemical structure of the streptococcal cell wall: I. The identification of a mucopeptide in the cell walls of groups A and A-variant streptococci. *Journal of Experimental Medicine*, **1961**, 114, 127-40.
- 123) Astronomo R. and Burton D. Carbohydrate vaccines: developing sweet solutions to sticky situations? *Nature reviews Drug discovery*, **2010**, 9, 308-24.
- 124) Salvadori L.; Blake M. and McCarty M. Group A streptococcus-liposome ELISA antibody titers to group A polysaccharide and opsonophagocytic capabilities of the antibodies. *Journal of Infectious Diseases*. **1995**, 171, 593-600.
- 125) Sabharwal H.; Michon F. and Nelson D. Group A streptococcus (GAS) carbohydrate as an immunogen for protection against GAS infection. *The Journal of Infectious Diseases*. **2006**, 193, 129-35.
- 126) Cunningham M. *Streptococcus* and rheumatic fever. *Current opinion in rheumatology*, **2012**, 24, 408-16.

- 127) Kabanova A.; Margarit I.; Berti F.; Romano M. and Grandi G. Evaluation of a Group A Streptococcus synthetic oligosaccharide as vaccine candidate. *Vaccine*, **2010**, 29, 104-14.
- 128) Delves P.; Lund T. and Roitt I. Can epitope-focused vaccines select advantageous immune response? *Molecular medicine today*, **1997**, 3, 55-60.
- 129) Baghian A.; Luftig M.; Black J.; Meng Y. and Pau C. Glycoprotein B of human herpesvirus 8 is a component of the virion in a cleaved form composed of amino-and carboxyl-terminal fragments. *Virology*, **2000**, 269, 18-25.
- 130) Hayman W.; Toth I.; Flinn N.; Scanlon M. Enhancing the immunogenicity and modulating the fine epitope recognition of antisera to a helical group A streptococcal peptide vaccine candidate from the M protein using lipid-core peptide technology. *Immunology and cell Biology*. **2002**, 80, 178-87.
- 131) Kim S.; Sung H.; Han J.; Jackwood D. Protection against very virulent infectious bursal disease virus in chickens immunized with DNA vaccines. *Veterinary microbiology*, **2004**, 101, 39-51.
- 132) Barrientos L.; Gronenborn A. The highly specific carbohydrate-binding protein cyanovirin-N: structure, anti-HIV/Ebola activity and possibilities for therapy. *Mini reviews in medicinal chemistry*. **2005**, 5, 21-31.
- 133) Pöhlmann S.; Soilleux E.; Baribaud F. DC-SIGNR, a DC-SIGN homologue expressed in endothelial cells, binds to human and simian immunodeficiency viruses and activates infection in trans. *Proceedings of the National Academy of Sciences*, **2001**, 98, 2670-5.
- 134) Klimstra W.; Nangle E.; Smith M. DC-SIGN and L-SIGN can act as attachment receptors for alphaviruses and distinguish between mosquito cell-and mammalian cell-derived viruses. *Journal of Virology*. **2003**, 77, 12022-32.

- 135) Lozach P.; Burleigh L.; Staropoli I. Dendritic cell-specific intercellular adhesion molecule 3-grabbing non-integrin (DC-SIGN)-mediated enhancement of dengue virus infection is independent of DC-SIGN Internalization Signals. *Journal of Biological Chemistry*, **2005**, 280, 23698-708.
- 136) Shan M.; Klasse P.; Banerjee K. and Dey A. HIV-1 gp120 mannoses induce immunosuppressive responses from dendritic cells. *PLoS Pathogens*. **2007**, 3, 11, 1637.
- 137) Vliegenthart J. Carbohydrate-based vaccines. *FEBS letters*, **2006**, 580, 12, 2945-2950.
- 138) Bröker M.; Dull P.; Rappuoli R.; Costantino P. Chemistry of a new investigational quadrivalent meningococcal conjugate vaccine that is immunogenic at all ages. *Vaccine*, **2009**, 27, 5574-80.
- 139) Avcı F.; Kasper D. How bacterial carbohydrates influence the adaptive immune system. *Annual review of immunology*, **2009**, 28, 107-30.
- 140) Giannini G.; Rappuoli R.; Ratti G. The amino-acid sequence of two non-toxic mutants of diphtheria toxin: CRM45 and CRM197. *Nucleic Acids Research*, **1984**, 12, 4063–4069.
- 141) Verez-Bencomo V.; Fernandez-Santana V. A synthetic conjugate polysaccharide vaccine against *Haemophilus influenzae* type b. *Science*, **2004**, 305, 522-5.
- 142) Jones C. Vaccines based on the cell surface carbohydrates of pathogenic bacteria. *Anais da Academia Brasileira de Ciencias*, **2005**, 77, 293-324.
- 143) Pace D. MenACWY-CRM, a novel quadrivalent glycoconjugate vaccine against *Neisseria meningitidis* for the prevention of meningococcal infection. *Current opinion in molecular therapeutics*, **2009**, 11, 692-706.
- 144) Lucas A.; Apicella M. and Taylor C. Carbohydrate moieties as vaccine candidates. *Clinical infectious diseases*, **2005**, 41, 5, 705-712.

- 145) Falugi F.; Petracca R.; Mariani M. and Luzzi E. Rationally designed strings of promiscuous CD4⁺ T cell epitopes provide help to Haemophilus influenzae type b oligosaccharide: a model for new conjugate vaccines. *European journal of Immunology*. **2001**, 31, 3816-24.
- 146) Liakatos A.; Kunz H. Synthetic glycopeptides for the development of cancer vaccines. *Current opinion in molecular therapeutics*, **2007**, 9, 35-44.
- 147) Goldblatt D. Recent developments in bacterial conjugate vaccines. *Journal of medical microbiology*, **1998**, 47, 563-567.
- 148) Paoletti L. and Madoff L. Vaccines to prevent neonatal GBS infection. *Seminars in Neonatology*, **2002**, 7, 4, 315–323.
- 149) Sabharwal H.; Michon F.; Nelson D. Group A streptococcus (GAS) carbohydrate as an immunogen for protection against GAS infection. *The Journal of Infectious Diseases*. **2006**, 193, 129-35.
- 150) Pollard A.; Perrett K.; Beverley P. Maintaining protection against invasive bacteria with protein–polysaccharide conjugate vaccines. *Nature Reviews Immunology*, **2009**, 9, 213-20.
- 151) Kartha K. and Jennings H. A facile, one-step procedure for the conversion of 2-(trimethylsilyl) ethyl glycosides to their glycosyl chlorides. *Tetrahedron Lett.*, **1990**, 31, 2537–2540.
- 152) Toshima K. and Tatsuta K. Recent progress in O-glycosylation methods and its application to natural products synthesis. *Chem. Rev.*, **1993**, 93, 1503–1531.
- 153) Norberg T. Modern Methods in Carbohydrate Synthesis. *Harwood Academic Publishers*, The Netherlands, **1996**, 82–106.
- 154) Geert B. Strategies in oligosaccharide synthesis. *Tetrahedron*, **1996**, 52, 1095–1121.

- 155) Geert B. Carbohydrate Chemistry. *Blackie Academic and Professional*, London, **1998**, 175–222.
- 156) Garegg P. Thioglycosides as glycosyl donors in oligosaccharide synthesis. **1997**, 179–205.
- 157) Veeneman G. Carbohydrate Chemistry. *Blackie Academic and Professional*, London, **1998**, 98–174.
- 158) Peter S. and Haase W. *Chem. Rev.*, 2000, 100, 4349–4393.
- 159) Mootoo D.; Konradsson P.; Udodong U.; Fraser-Reid B. Armed and disarmed n-pentenyl glycosides in saccharide couplings leading to oligosaccharides *J.Am.Chem.Soc.* **1988**, 110, 5583 – 5584.
- 160) Fraser-Reid B.; Konradsson P.; Mootoo D. and Udodong U. Direct elaboration of pent-4-enyl glycosides into disaccharides. *J. Chem. Soc., Chem.Comm.*, **1988**, 823–825.
- 161) Fraser-Reid B.; Udodong U.; Wu Z.; Ottosson H.; Merritt J.; Rao C.; Roberts C. and Madsen R. n-Pentenyl glycosides in organic chemistry: a contemporary example of serendipity. *Synlett*, **1992**, 927–942.
- 162) Friesen R. and Danishefsky S. On the controlled oxidative coupling of glycals: a new strategy for the rapid assembly of oligosaccharides. *J. Am. Chem. Soc.*, **1989**, 111, 6656–6660.
- 163) Friesen R. and Danishefsky S. On the use of the haloetherification method to synthesize fully functionalized disaccharides. *Tetrahedron*, **1990**, 46, 103–112.
- 164) Douglas N.; Ley S.; Lu'cking U. and Warriner S. Carbohydrates. *J. Chem. Soc., Perkin Trans. 1*, **1998**, 51–65.
- 165) Zhang Z.; Ollmann I.; Ye X.; Wischnat R.; Baasov T. and Wong C. Programmable one-pot oligosaccharide synthesis. *J.Am.Chem.Soc.*, **1999**, 121, 734–753.

- 166) Ye X. and Wong C. Anomeric reactivity-based one-pot oligosaccharide synthesis: a rapid route to oligosaccharide libraries. *J. Org. Chem.*, **2000**, 65, 2410–2431.
- 167) Wong C.; Halcomb R.; Ichikawa Y. and Kajimoto T. Enzymes in organic synthesis: application to the problems of carbohydrate recognition. *Angew. Chem., Int. Ed. Engl.*, **1995**, 34, 521.
- 168) Crout D. and Vic G. Glycosidases and glycosyl transferases in glycoside and oligosaccharide synthesis. *Curr. Opin. Chem. Biol.*, **1998**, 2, 98.
- 169) Hrlein R. Glycosyltransferase-catalyzed synthesis of non-natural oligosaccharides. *Top. Curr. Chem.*, **1999**, 200, 227.
- 170) Palcic M. Biocatalytic synthesis of oligosaccharides. *Curr. Opin. Biotechnol.*, **1999**, 10, 616.

CHAPTER 2

DESIGN AND CHEMICAL SYNTHESIS OF POLYRHAMNOSE OLIGOSACCHARIDES FOR THE DEVELOPMENT OF *GROUP A STREPTOCOCCUS* VACCINE CANDIDATES

Huzi Sun, Ning Ding, Geert-Jan Boons.* To be submitted to J. Am. Chem. Soc.

Abstract

Group A Streptococcus (GAS) is a leading cause of many lethal human diseases, and millions of people die each year due to severe GAS diseases, e.g., invasive infections, acute rheumatic fever (ARF), and rheumatic heart disease (RHD). It is the sole species of Lancefield group A and all GAS serotypes express the Lancefield group A carbohydrate (GAC), which is comprised of a polyrhamnose backbone with an immunodominant *N*-acetylglucosamine (GlcNAc) side chain. The hurdle toward a carbohydrate-based vaccine includes the potential of the GAC GlcNAc side chain to provoke cross-reactive antibodies relevant to the immunopathogenesis of rheumatic fever. Nizet *et al.* have reported that antibodies to GAC lacking the GlcNAc side chain and containing only polyrhamnose promoted opsonophagocytic killing of multiple GAS serotypes and protected against systemic GAS challenges after passive immunization. Synthetic challenges of these oligosaccharide antigens include: 1) Because of the lacking of the 6-OH as comparing with the common pyranoses, rhamnose is very sensitive and active under acidic conditions; 2) Selection of suitable and optimal protecting groups for this substrate; 3) Screening of the optimal glycosylation method; 4) The introduction of the right type of linker to this substrate and the timing for the addition. In this paper, we reported the chemical synthesis of those promising vaccine candidate antigens and the “4+4” *in situ* bond cleavage polymerization. Through comprehensive studies, we proposed mechanisms for the discovered reactions, and successfully synthesized the desired octasaccharide antigen by a highly convergent approach and a library of polyrhamnose oligosaccharides by the “4+4” *in situ* bond cleavage polymerization. These compounds are utilized for epitope mapping by glycan microarray and for the development of new glycoconjugate vaccine candidates.

Introduction

Group A Streptococcus (GAS) infections are leading healthcare problems throughout the world. GAS causes a wide range of diseases, including asymptomatic colonization, uncomplicated pharyngeal, skin infections, and life-threatening invasive illnesses such as sepsis, necrotizing fasciitis, and toxic shock syndrome. Pharyngitis may lead to delayed sequela as rheumatic fever, which is caused by an autoimmune reaction to Group A streptococci that results in valvular damage.² In the developing world, rheumatic fever is the primary cause of acquired heart disease in children, adolescents and young adults, responsible for at least 350,000 premature deaths per year.³ There is currently no safe and efficacious commercial vaccine against GAS infection despite the high demand globally.

GAS vaccines can be generally divided into M protein-based and non-M protein-based vaccines. GAS carbohydrate-based vaccines are less developed as compared to M protein-based vaccines, and research in GAS carbohydrate-based vaccines is undergoing extensively.⁴ The major challenge for the development of an anti-GAS vaccine is the diversity of GAS strains. All GAS serotypes express the Lancefield group A carbohydrate (GAC), which consists of $\{\rightarrow 2)[\beta\text{-D-GlcNAc}(1\rightarrow 3)]\alpha\text{-L-Rha}(1\rightarrow 3)\}\alpha\text{-L-Rha}(1\text{-})_n$ repeating units, and it could serve as a common antigen for the development of a universal GAS vaccine. Pinto *et al.* reported a core antigenic determinant—a hexamer structure of two repeating units based on the antibody-binding epitope of GAS-PS⁵. Costantino and co-workers evaluated optimal oligosaccharide length and terminal molecule residues by comparing the *in vivo* efficacy of CRM197 glycoconjugates of the native GAS-PS material versus that of the synthetic oligosaccharides.⁶

However, the autoreactivity of antibodies, which recognize the native GAC GlcNAc side chain, against human tissues has been raised by several groups.^{7,8,9} Cunningham groups derived anti-GlcNAc monoclonal antibodies that are cross-reactive for heart or brain tissue from patients with rheumatic fever.¹⁰ In recent years, Nizet and co-workers identified the genes responsible for GAC biosynthesis. They demonstrated that antisera raised against the polyrhamnose core of GAC, as purified from the Δ GacI mutant, may still provide significant broad-spectrum opsonophagocytic activity.¹

Although these antigens can be isolated from Δ GacI mutant, chemical synthesis offers a much more attractive approach. In contrast to the conventional polysaccharide isolation, the synthetic approach could generate pure, homogeneous, and well-defined oligosaccharides, which contain a single reactive group for covalent conjugation. Based on the core antigenic determinants of long polysaccharide chains, we could design synthetic oligosaccharides to evaluate how particular structural features, such as length, non-reducing end residue, and the composition, influence carbohydrate immunogenicity. Furthermore, we could synthesize well-defined substructures to study structure-activity relationships or to determine a minimal epitope for eliciting immune responses.

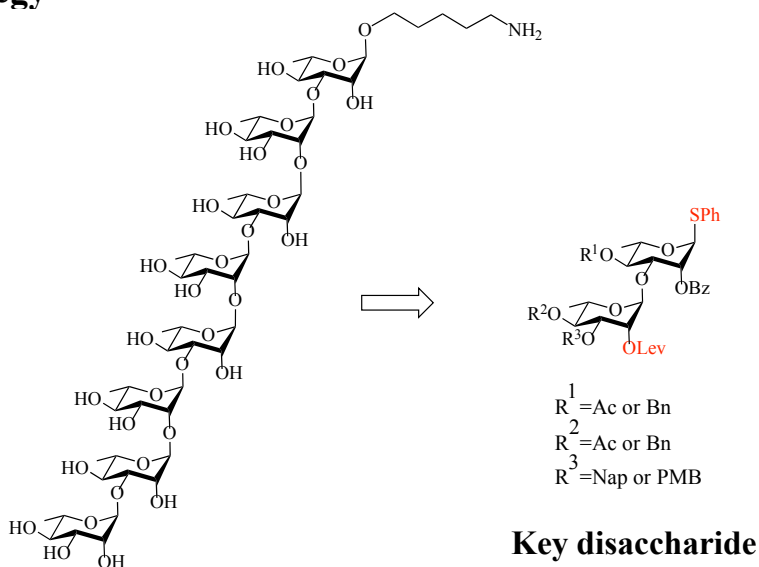
A number of synthetic challenges are included in the synthetic strategies of those oligosaccharide antigens. Firstly, because of the lacking of the 6-OH as comparing with the common pyranoses, rhamnose is very sensitive and active under acidic conditions. Therefore, it requires comprehensive studies of the stabilities of different linkages and subtle handling techniques. Secondly, the selection of the suitable protecting groups and screening of the optimal glycosylation method for this substrate require an exhaustive and challenging examining process.

Lastly, the addition of the right type of linker to this substrate and the timing for the addition are difficult problems that need to be solved.

We reported here a highly convergent approach for the chemical synthesis of those promising vaccine candidate antigens and the “4+4” *in situ* bond cleavage polymerization. We have established a thorough examining process for the designed strategy, conducted comprehensive studies for the discovered reactions, and successfully synthesized these biologically important antigens for developing new GAS vaccine candidates.

Discussions and results

Synthetic strategy

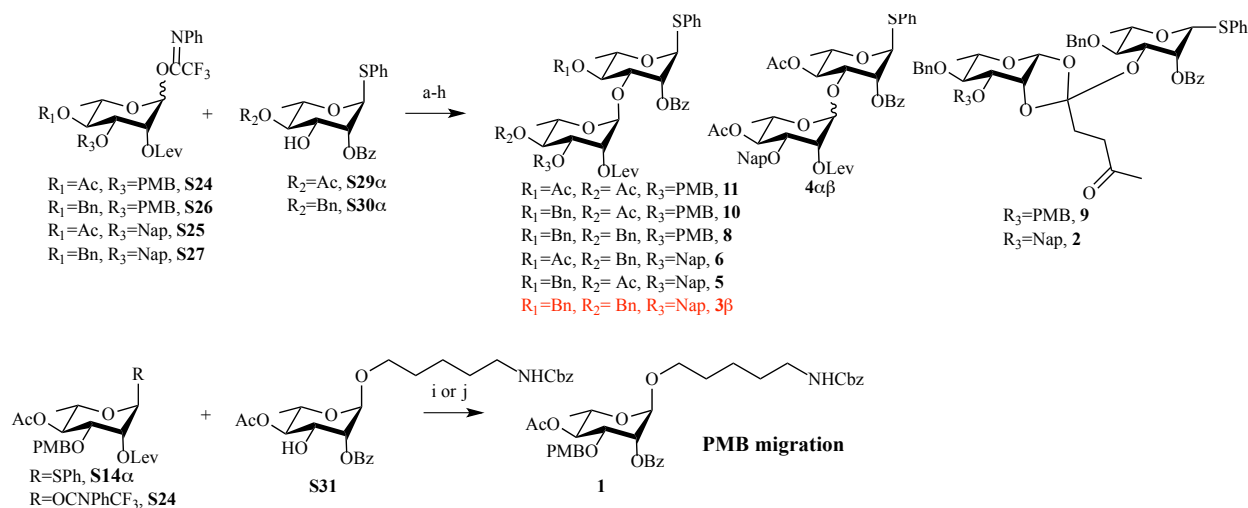


Scheme 1. The designed “key disaccharide” retrosynthetic strategy

The designed retrosynthetic strategy employed a key disaccharide to perform a convergent “2+2” and “4+4” glycosylation to synthesize the desired octasaccharide polyrharnnose antigen. We also equipped ester protecting groups at C-2 to utilize neighboring group participating effects to assist the α -selectivity, and employed the levulinoyl (Lev) ester as an orthogonal protecting group to differentiate the bottom C-2. Furthermore, we deployed an orthogonal protecting group

at C-3 position, such as 4-Methoxybenzyl ether (PMB) or 2-Naphthylmethyl ether (Nap), to further assemble the trisaccharide repeating units. This way, we could synthesize both polyramnose targets and GAS oligosaccharides with different GlcNAc side chain variations from one key disaccharide as shown in **Scheme 1**.

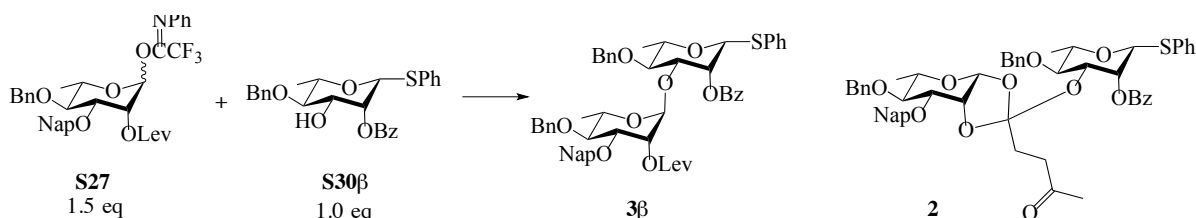
Synthesis of the key disaccharide



Scheme 2. Selection of the suitable protecting group combination. Reagents and conditions: a) **S26**, **S30β**, TBSOTf 0.05eq, DCM, -78°C, **8** 6%, **9** 33%. b) **S26**, **S29α**, TBSOTf 0.05eq, DCM, -78°C, No reaction. c) **S24**, **S30β**, TBSOTf 0.05eq, DCM, -20°C, **10** 43%. d) **S24**, **S29α**, TBSOTf 0.05eq, DCM, -20°C, **11** 27%. e) **S25**, **S29α**, TBSOTf 0.05eq, DCM, -50°C-0°C, **4α** 49%, **4β** 7%, α:β 7:1. f) **S25**, **S30β**, TBSOTf 0.05eq, DCM, -78°C-0°C, **5** 20%. g) **S27**, **S30β**, TBSOTf 0.05eq, DCM, -78°C, **3β** 10%, **2** 70%. h) **S27**, **S29α**, TBSOTf 0.05eq, DCM, -78°C-0°C, **6** 14%. i) **S14α**, NIS, AgOTf 0.5eq, DCM, -50°C-0°C, **1** 63%. j) **S24**, TBSOTf 0.1eq, DCM, -50°C-0°C, **1** 55%.

The assembly of the key disaccharide requires thorough screening of the suitable protecting group combination and the optimal glycosylation condition as shown in **Scheme 2**. We investigated 1) the electronic effects on C-4 position by employing standard electron donating protecting group or electron withdrawing protecting group; 2) the efficiency of the leaving group during the glycosylation; 3) the influence of the linker on the anomeric center. Firstly, we examined the electronic effects of C-4 position on both donor and acceptor with PMB at the C-3 position on the donor. Benzyl ether (Bn) and acetyl ester (Ac) were selected to be the standard electron donating and electron withdrawing groups. We started with low temperature and mild promoting catalyst, as 6-Deoxy rhamnose substrate is very reactive. Tert-Butyldimethylsilyl

trifluoromethanesulfonate (TBSOTf)-promoted coupling of **S26** with **S30 β** yielded disaccharide **8** and orthoester **9** in moderate reactivity because the low temperature froze the glycosylation at the orthoester stage. TBSOTf-promoted coupling of **S24** with **S29 α** gave disaccharide **11** in low reactivity, suggesting C4-Ac electron withdrawing effects could reduce the reactivity. TBSOTf-promoted coupling of **S26** with **S29 α** and **S24** with **S30 β** resulted in low yield or no reaction, as the reactivity of the donor and the acceptor did not match. Secondly, we examined the efficiency of the leaving group during the glycosylation and the influence of the linker on the anomeric center for this substrate. An *N*-iodosuccinimide (NIS)/silver trifluoromethanesulfonate (AgOTf) promoted glycosylation of **S24** with **S31** and **S14 α** with **S31** yielded monosaccharide **1** surprisingly. PMB migrated from the donor to the acceptor, and no desired disaccharide product was formed (see **Table 2, SI**). It can migrate under very mild conditions, and it is substrate dependent. Thus, PMB protecting group is not suitable in this case, and we turned our attention towards Nap protecting group since it is more acid stable comparing to PMB. Then, we examined the electronic effects of C-4 position on both donor and acceptor with Nap at the C-3 position on the donor. TBSOTf-promoted coupling of **S27** with **S29 α** and **S25** with **S30 β** resulted in low yield, as the reactivity of the donor and the acceptor did not match. Because of the long-range participation of the C4-Ac on the donor, TBSOTf-promoted coupling of **S25** with **S29 α** gave disaccharide **4** as α/β mixture, even in the presence of the Lev neighbor group participation. This indicates that long-range participation competes with the neighbor group participation effect, and it has a dramatic influence on the stereo-configuration of the anomeric center.



Scheme 3. The optimal protecting group pattern for the key disaccharide glycosylation

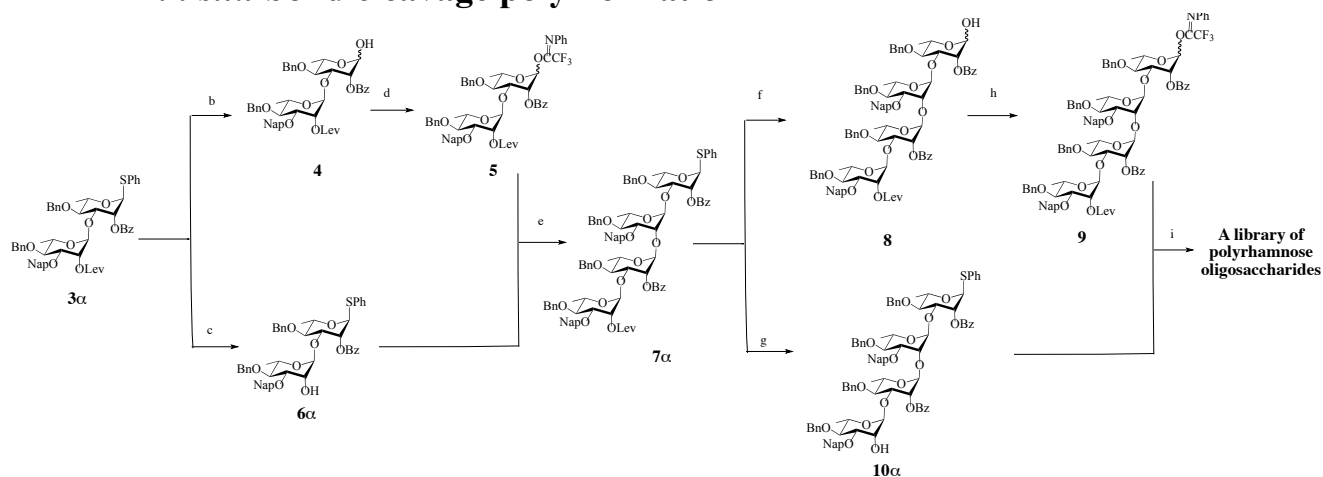
Entry	Temperature/°C	Catalyst and solvent	Yield
1	-78	0.05 eq TBSOTf, DCM	3β 10%; 2 70%
2	-78~rt	0.1 eq TBSOTf, DCM	3β 75% only
3	-30~rt	0.1 eq TBSOTf, DCM	3β 57% only
4	-30~rt	0.1 eq TMSOTf, DCM	3β 81% only
5	-30~rt	0.25 eq TMSOTf, DCM	3β 43% only
6	-30~rt	0.1 eq HOTf, DCM	3β 46% only
7	0~rt	0.1 eq TMSOTf, DCM	3β Trace
8	-30~rt	0.1 eq TMSOTf, CH ₃ CN	No reaction

Table 1. The optimization of the key disaccharide glycosylation condition

The optimal protecting group pattern with the best reactivity was achieved by glycosylation of **S27** with **S30β**, as shown in **Table 1**, as it yielded high amount of disaccharide **3β** and orthoester **2**, even under very low temperature condition. The optimal protecting group combination is C3-Nap on the donor and C4-Bn on both donor and acceptor. The armed donor and acceptor with C-4 electron donating ether protecting group increase the reactivity of the substrate, and the overall electronic effects of both donor and the acceptor need to match with each other in order to achieve the best glycosylation reactivity. Further investigation of different temperature, catalysts, and solvents for the glycosylation of **S27** with **S30β**, as shown in **Table 1**, revealed the optimal glycosylation condition for the key disaccharide synthesis. The low temperature will freeze the glycosylation process at the orthoester stage. Use of too much acid or use of weak/strong acid will decrease the reaction yield. The glycosylation process mainly takes place from -30 °C to 0 °C for this substrate. The key disaccharide **3β** could be obtained in a yield of 87% through trimethylsilyl trifluoromethanesulfonate (TMSOTf) mediated glycosylation of

S27 with **S30 β** in dichloromethane (DCM) in the presence of 4A molecular sieve under -30 °C~0 °C (See **Scheme 5, SI**).

“4+4” *in situ* bond cleavage polymerization



Scheme 4. Synthesis of the tetrasaccharide **7 α** . Reagents and conditions: a) TMSOTf, DCM; 72% b) NIS, AgOTf, TTBP, wet DCM; 97%; c) Hydrazine acetate, DCM:MeOH(1:1), 93%; d) ClCNPhCF₃, Cs₂CO₃, DCM, 87%; e) TBSOTf, DCM, 85%. f) NBS, Acetone:PBS Buffer(6:1); 87% g) Hydrazine acetate, DCM:MeOH(1:1), 99%; h) ClCNPhCF₃, Cs₂CO₃, DCM, 87%; i) TBSOTf, DCM.

The assembly of the polyrhannose oligosaccharides and the “4+4” *in situ* bond cleavage polymerization are depicted in **Scheme 4**. Disaccharide assembly was performed by the optimal glycosylation method as described before. Glycosylation yield are very good for both α and β acceptors (See **Scheme 5, SI**). The anomeric thioglycoside of **3 α** was removed by NIS and a catalytic amount of AgOTf in the presence of the acid scavenger 2,4,6-Tri-tert-butylpyrimidine (TTBP) in wet DCM to afford **4** in 97% excellent yield. It was subsequently converted into the *N*-phenyltrifluoroacetimidate **5** by using standard conditions (72%, 97% respectively). The Lev ester of **3 α** was removed by hydrazine acetate in DCM:MeOH (1:1) to yield the acceptor **6 α** (93% yield). In a TMSOTf promoted glycosylation of **5** with **6 α** in DCM at -30 °C~0 °C, we observed the product decomposition over the process. The 1,2-linkage of the tetrasaccharide is acid labile and can be cleaved under the acid glycosylation condition to yield the decomposed disaccharide products. Thus, a milder catalyst TBSOTf was used, the amount of acid was reduced to 0.075 eq, and the reaction temperature was decreased to -40 °C~0°C. With the

modified method, the glycosylation worked satisfyingly, and yield were very good for both α and β acceptors (See **Scheme 6, SI**). Then, the tetrasaccharide **7 α** was converted to the acceptor **10 α** by the standard manipulation (99% yield). The anomeric thioglycoside of **7 α** was removed by NBS in acetone:PBS buffer (6:1) to afford **8** in 87% good yield, which was subsequently converted into the *N*-phenyltrifluoroacetimidate **9** by using standard conditions (87% yield). Surprisingly, a TBSOTf promoted glycosylation of **9** with **10 α** in DCM at -40 °C~0 °C yielded a library of polyrhannose oligosaccharides (See **SI**). TLC showed a string of new spots, and high-resolution MALDI-MS of the reaction mixture showed an array of new products. Pure products can be obtained by LH20 and silica gel chromatography purification.

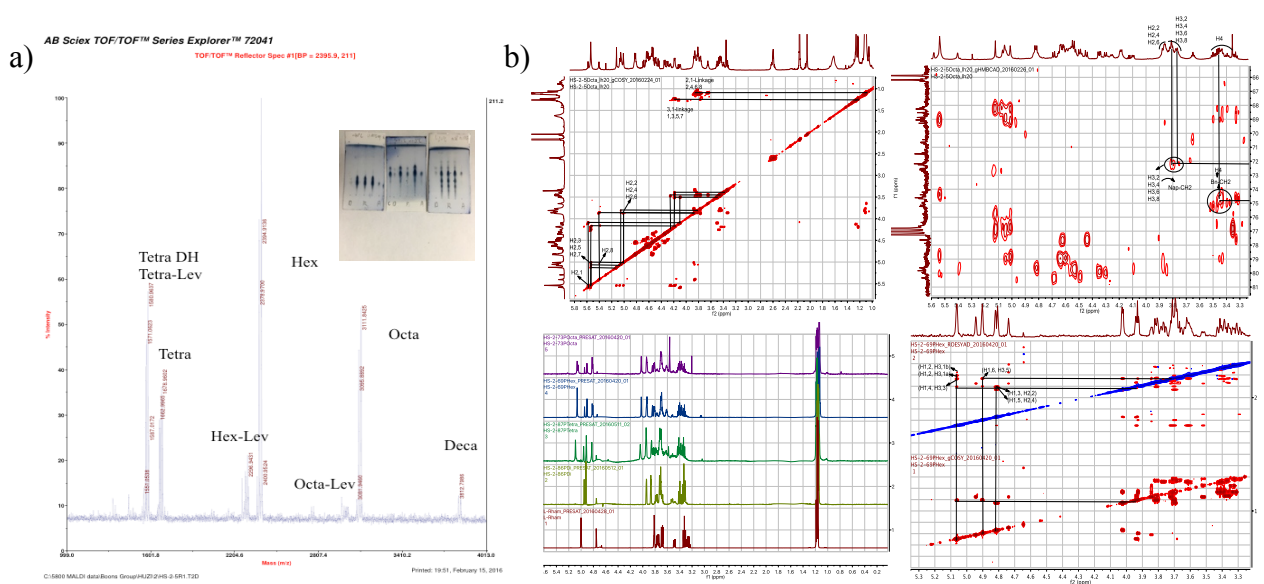
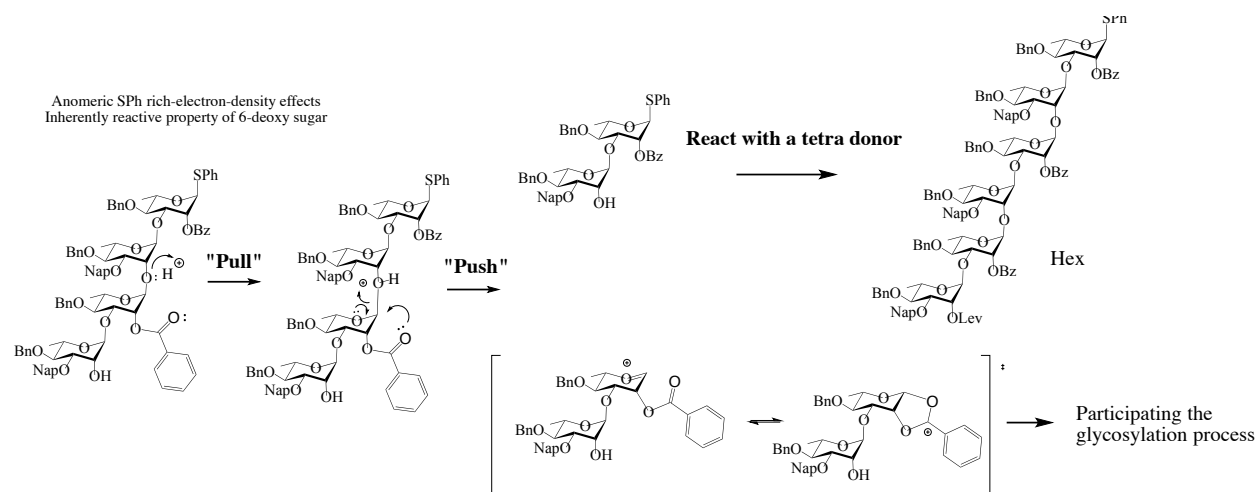


Figure 1. a) MALDI-MS and TLC of the “4+4” polymerization reaction mixture. b) H-H 2D COSY and HMBC confirmation of the structure integrity of the octasaccharide c) Proton spectrum of the deprotected polyrhannose library and H-H 2D ROESY and H-H 2D

Next, the purified library of polyrhannose oligosaccharides was deprotected to yield the final compounds, and the deprotection sequence is crucial for these compounds. Firstly, deprotection of the esters by sodium methoxide in methanol must be conducted in the first step to avoid β -elimination. Secondly, the anomeric thioglycoside was removed by NBS in acetone:10

mM PBS buffer (6:1). Lastly, deprotection of the benzyl ether by 10% Pd/C in THF/H₂O, and subsequent purification by P2 Bio Gel yielded pure final compounds. The three-steps deprotection yielded the desired products in good yield. The pure compounds could be used for the structure-activity analysis, the glycan microarray, and ELISA antibody binding studies. Two-dimensional NMR experiments combined with high-resolution MALDI-MS confirmed the structural integrity of the compound. The J_{H-H} coupling constants of non-reducing end anomeric protons were all between 169 and 170 Hz, indicating all α -glycosidic linkages. For protected compounds, all H2 connected with ester protecting group were down shifted (~ 5.6 ppm) compared with all H2 without ester protecting group connection (~ 3.8 ppm) as shown in H-H 2D COSY. All H3 connected with Nap protecting group were cross-talking with C_{Nap-CH₂}, and all H3 connected with Bn protecting group were cross-talking with C_{Bn-CH₂} as shown in HMBC. After deprotection, H-H 2D ROESY and H-H 2D COSY provided a second confirmation of the correct glycosidic linkage.

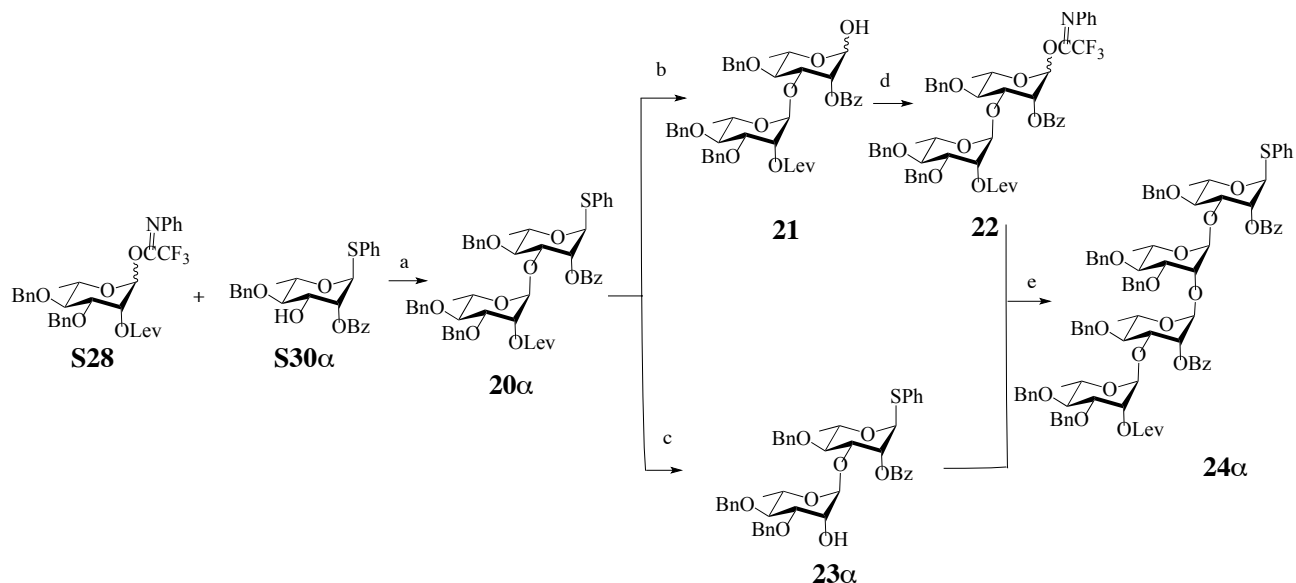
We proposed a “Push and Pull” mechanism as shown in **Scheme 5**. The mechanism shows the process of the *in situ* breakage of the acid labile 1,2-linkage to form the new acceptor, which could be further reacted with a tetra donor to yield the new product. Polymerization repeats this process until the reactivity of the substrate is too low to react. The proposed mechanism is based on: 1) model acid test suggested that 1,3-linkage is acid stable, and 1,2-linkage is acid labile, and the peaks corresponding to oligosaccharides without Lev in high-resolution MALDI-MS were resulted from the *in situ* cleavage of the acid labile 1,2-linkage; 2) ester neighbor group participating effects; 3) the electron richness of the thiophenyl protecting group at the anomeric center increases the overall reactivity; 4) the steric hindrance and electron donating properties of the Nap protecting group might increase the breaking tendency of the 1,2-glycosidic linkage.



Scheme 5. The proposed “Push and Pull” mechanism for “4+4” polymerization

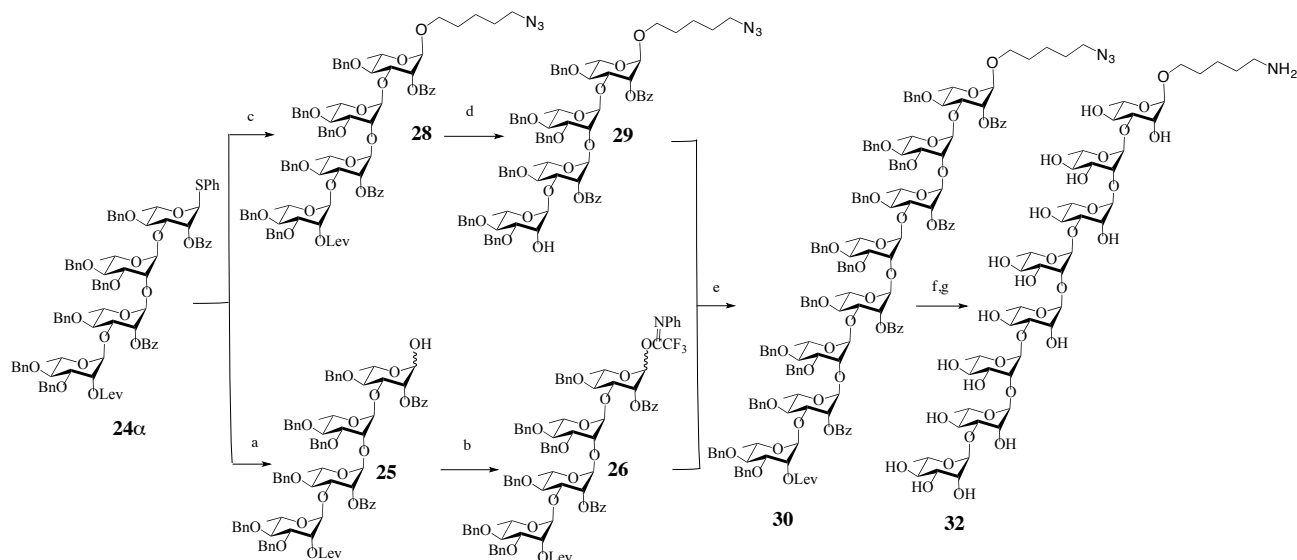
Then, we investigated two possible strategies for the linker introduction: 1) introducing the linker at the final octasaccharide stage; 2) introducing the linker at the tetrasaccharide stage. Several failed attempts excluded the first strategy. Because the efficiency of the glycosylation and the reactivity of the substrate are low, no products were observed in the case of the octasaccharide linker introduction. The second strategy required installation of the linker on the tetrasaccharide first, then perform “4+4-linker” glycosylation. We explored the azido and the Cbz-protected amino linkers, and disappointingly, both “4+4-linker” reactions did not work (See **Scheme 14, SI**). Interestingly, glycosylation of the tetrasaccharide **7α** with the Cbz-protected amino linker yielded the “1,2-linkage bond cleavage” disaccharide as the major product, and the yield of the desired tetrasaccharide were very low (**Scheme 11, SI**). However, glycosylation of the tetrasaccharide **9** with the Cbz-protected amino linker yielded the desired tetrasaccharide product only. The yield improved dramatically when using the azido linker as the acceptor (**Scheme 12 and 13, SI**), as azido group on the linker is a small-sized protecting group, and it does not bear any charge, thus, it has minimal effects during the glycosylation process. These observations suggest that the thiophenyl at the anomeric center increases the overall reactivity of the substrate.

Synthesis of the octasaccharide: replace Nap by Bn



Scheme 6. Synthesis of the tetrasaccharide **24α**. Reagents and conditions: a) TMSOTf, DCM; 79% b) NBS, Acetone:PBS Buffer(6:1),84%; c) Hydrazine acetate, DCM:MeOH(1:1), 80%; d) ClCNPfCF₃, Cs₂CO₃, DCM, 80%; e) TBSOTf, DCM, 90%

Based on the failed experiment of “4+4-linker” glycosylation, we hypothesized that the steric hindrance of C3-Nap might decrease the accessibility of the adjacent hydroxyl in the glycosylation process. In order to test this hypothesis, we decided to replace Nap by Bn and reassembled the building blocks as shown in **Scheme 6**. The assembly of the tetrasaccharide **24α** by the optimal glycosylation methods described before was done in good yield (see **Scheme 17, SI**). Interestingly, a TBSOTf promoted “4+4-Bn” glycosylation of **26** with **27α** in DCM at -40 °C~0 °C yielded the octasaccharide with low reactivity and the substrate was not polymerized. The reactivity of the acceptor was quite low, and most of the donors were hydrolyzed or unreacted. We also observed the bond-cleavage product hexasaccharide from high-resolution MALDI-MS and TLC. However, the polymerization could not occur due to the low reactivity of the substrate (see **Scheme 19, SI**).



Installation of the azido linker was tested on donor **24α** and **26**. Interestingly, a *N*-iodosuccinimide (NIS)/triflic acid (HOTf)-mediated glycosylation of the tetrasaccharide **24α** with the azido linker yielded the desired tetrasaccharide **28** in 78% good yield, and the “1,2-linkage bond cleavage” disaccharide, in this case, was the minor product. TMSOTf-mediated glycosylation of the tetrasaccharide **26** with the azido linker yielded only the desired tetrasaccharide **28** in 60% acceptable yield (see **Scheme 21, SI**). These observations suggested that substrates with C3-Nap have higher reactivity compared to those with C3-Bn, and C3-Nap increases the breaking tendency of the 1,2-glycosidic linkage. Then, the tetrasaccharide **28** was converted to the acceptor **29** by the standard manipulation (94% yield). The anomeric thioglycoside of **24α** was removed by NBS in acetone:PBS buffer (6:1) to afford **25** in 87% yield, and it was subsequently converted into the *N*-phenyltrifluoroacetimidate **26** by using standard conditions (87% yield). Finally, a TBSOTf promoted glycosylation of **26** with **29** in DCM at -40 °C~0 °C yielded the octasaccharide **30** in 33% low yield. The reactivity was low, and we observed the bond cleavage product hexasaccharide after adding an excessive amount of

acid, which indicated that the octasaccharide was still quite acid sensitive. However, the polymerization could not occur due to the low reactivity of the substrate. Fortunately, the yield improved to 62% when the reaction was done at -30 °C~0 °C by using toluene as the solvent, as the polarity of the substrate is more similar to that of toluene according to the theory of "similarity and intermiscibility". Purification of the octasaccharide **30** by LH20 yielded pure products due to the large mass differences. Deprotection of the octasaccharide **30** by standard two-step procedure gave the final product **32** in 94% excellent yield.

Conclusion

In conclusion, a highly convergent synthetic strategy has been developed for the chemical synthesis of those promising vaccine antigens. We have described the thorough examining process of the synthesis of the “key disaccharide”, and the unexpected discoveries of “PMB migration” and the “4+4” *in situ* bond cleavage polymerization. We have established a comprehensive investigation process to address the challenges associated with those oligosaccharides, and proposed a “push and pull” mechanism for the “4+4” *in situ* bond cleavage polymerization. The synthetic oligosaccharides can be used for epitope mapping of various GAS antibody and patient serum by glycan microarray, which will provide important insights for the autoimmune cross-reactivity associated with GAC antigens and lay the foundation for the development of new GAS vaccine candidates.

General procedures

NMR spectra were recorded in the NMR facility of Complex Carbohydrate Research Center, UGA, on a Varian Mercury 300 (300 MHz for ^1H , 75 MHz for ^{13}C), Varian Inova 500 (500 MHz for ^1H , 125 MHz for ^{13}C), Varian Inova 600 with cryoprobe (600 MHz for ^1H , 150 MHz for ^{13}C), Varian VNMRs 600 with cryoprobe (600 MHz for ^1H , 150 MHz for ^{13}C) or Varian Inova 800 with cryoprobe (800 MHz for ^1H , 200 MHz for ^{13}C). Chemical shifts are reported in parts per million (ppm) relative to tetramethylsilane (TMS) as the internal standard. NMR data is presented as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, dd = doublet of doublet, m = multiplet and/or multiple resonances), integration, coupling constant in Hertz (Hz). The assignments of ^1H NMR peaks were made from 2D ^1H - ^1H COSY, ^1H - ^{13}C HSQC, ^1H - ^{13}C HMBC, ^1H - ^1H ROESY and ^1H - ^1H TOCSY spectra. Mass spectra were recorded on an ABISciex 5800 MALDI-TOF-TOF, Bruker Microflex MALDI-TOF or Shimadzu LCMS-IT-TOF mass spectrometer. The matrix was used was 2, 5-dihydroxy-benzoic acid (DHB). TLC-analysis performed on Silica gel 60 F254 (EMD Chemicals inc.) with detection by UV-absorption (254 nm) when applicable, and by spraying with a solution of $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24}\cdot\text{H}_2\text{O}$ (25 g/L) in 5% sulfuric acid in ethanol followed by charring. Acid washed molecular sieves (4Å) were flame activated *in vacuo*. All moisture sensitive reactions were carried out under an argon atmosphere. Ambient temperature in the laboratory was usually 20 °C. CH_2Cl_2 and CH_3CN were distilled freshly from CaH_2 . Other commercially available reagents were obtained from Aldrich, Fisher or TCI and used as received.

The optimal glycosylation method

According to the study of the selection of the protecting group combination and the optimization of the reaction condition, the optimal glycosylation method for this substrate is generally summarized as follows:

Donor (1.2 eq) and acceptor (1.0 eq) were co-evaporated with toluene (3×3 mL) and dried under high vacuo overnight. Then, they were dissolved under argon in anhydrous DCM to maintain the acceptor concentration of 0.1 M. Freshly activated powdered 4Å acid washed molecular sieves were added at -40 °C, then the solution was argon purged 3 times and stirred under argon for 20 min. The appropriate type and amount of acid was added at -40 °C or -30 °C, and the temperature was maintained for additional 20 min before raising it to 0 °C naturally. The reaction mixture was quenched by the addition of a drop of Et_3N (3 μL) at 0 °C. The mixture was diluted with DCM and filtered. Then the filtrate was concentrated under reduced pressure and the residue was purified by silica gel column chromatography or LH20.

General procedure for the synthesis of (*N*-phenyl)-trifluoroacetimidate donor

To a solution of hydroxyl precursor in CH_2Cl_2 at 0 °C was added trifluoro-*N*-phenylacetimidoyl chloride (2.0 eq) and Cs_2CO_3 (2.0 eq). The reaction mixture was stirred for 3 h and filtered through a pad of Celite, then concentrated in vacuo. The concentrates were purified by flash chromatography over silica gel (hexane/EtOAc/1% NEt_3 , 6/1 to 2:1, v/v) to give the (*N*-phenyl)-trifluoroacetimidate donor as an anomeric mixture and was used directly.

Global deprotection

Deprotection sequence is crucial for the compound with thio-glycoside. Generally, deprotection of the esters by sodium methoxide in methanol must be conducted in the first step to avoid β -elimination. Secondly, deprotection of the thio-glycoside by NBS in acetone:10 mM

PBS buffer (6:1). Lastly, deprotection of the benzyl ether by 10% Pd/C hydrogenation. For compounds with GlcNAc side chain, firstly, deprotection of Troc by Zn and acetylation, and the rest of deprotection sequence was the same as described before. The final compound was purified by P2 or P4 column by DI water.

In the case of the compound with an azide linker, deprotection of the esters by sodium methoxide in methanol should be conducted in the first step. Then, deprotection of the benzyl ether and reduction of the azide by 10% Pd/C hydrogenation yielded the final compound smoothly. The final compound was purified by P2 or P4 column by 0.1 M ammonium bicarbonate buffer as the eluent only; it will decompose on the column if using DI water.

General procedure for conjugation

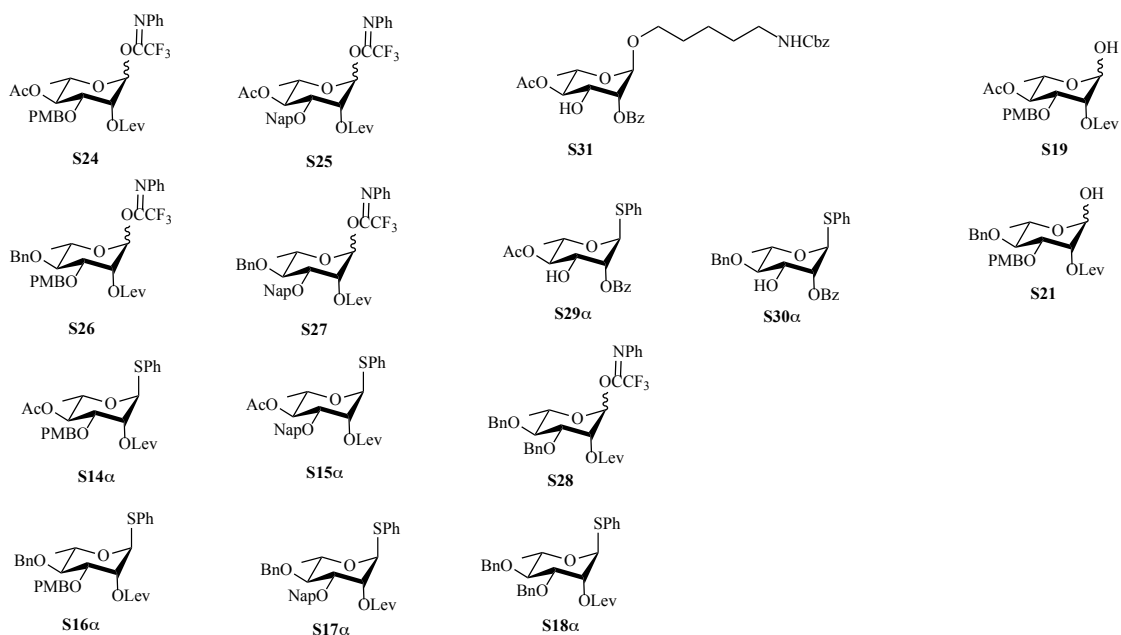
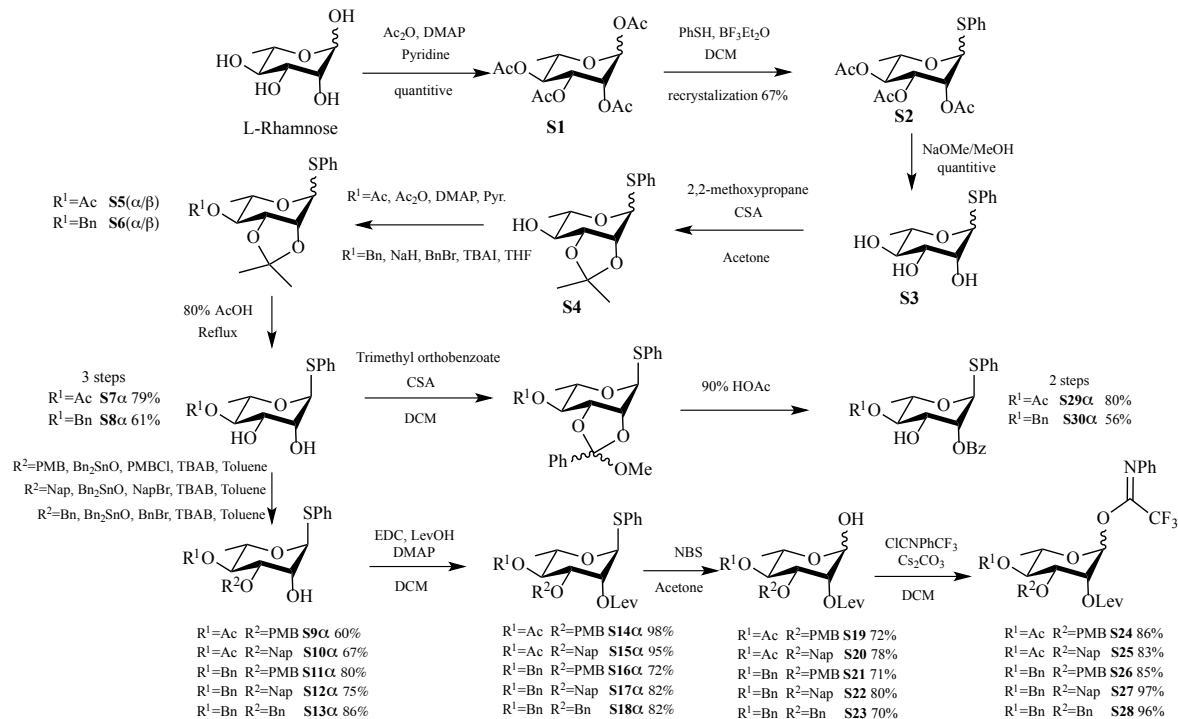
Maleimide modified BSA or KLH was purchased from Thermo Scientific (Product 77116) and dissolved in DI water (10 mg/mL). The thiol-activated carbohydrate (1:1 by mass) was dissolved in 200 μ L conjugation buffer and mixed with the protein solution. The mixture was gently stirred for 16 h at r.t. and then was purified by ultrafiltration using the filter with 10K NMWL five times (BSA) or spin column (KLH).

MALDI-TOF analysis of conjugates

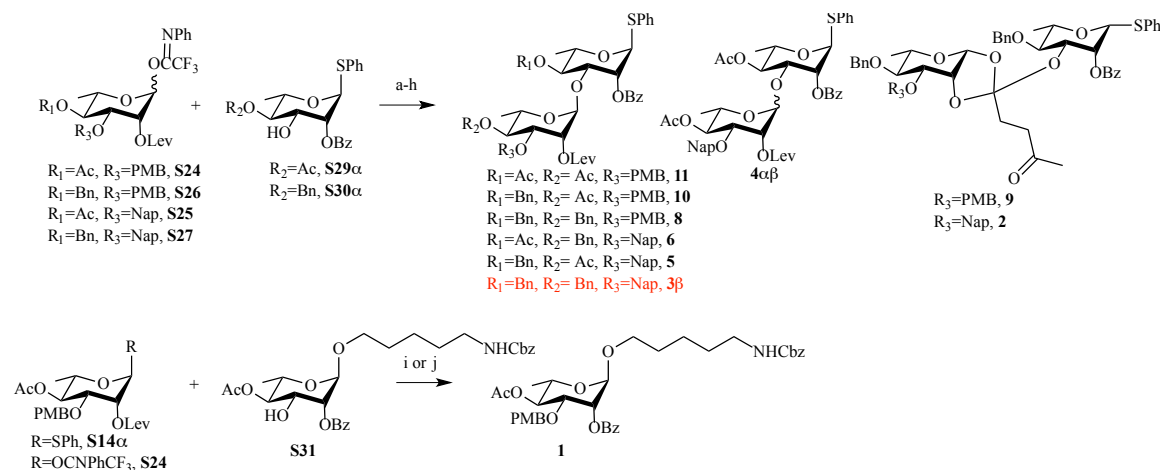
The glycoconjugates were desalted using C18 ZipTip tips following the manufacturer's protocol. The final concentration was $\sim 1 \mu\text{g } \mu\text{L}^{-1}$. Sinapinic acid matrix was freshly prepared using the following protocol: 5 mg sinapinic acid was added to 500 μ L 30% $\text{CH}_3\text{CN}/70\%$ $\text{H}_2\text{O}/0.1\%\text{TFA}$. The mixture was vortexed and centrifuged. The supernatant was used as a matrix. 1 μ L of conjugates sample was mixed with 2 μ L matrix solution. 1 μ L of the mixed sample was pipetted onto the MALDI plate. The spot was air dried before analysis using linear

positive mode using a Bruker Microflex MALDI-TOF spectrometer or an ABISciex 5800 MALDI-TOF-TOF spectrometer.

Synthesis of building blocks



Selection of the suitable protecting group



Scheme 2. Selection of the suitable protecting group combination. Reagents and conditions: a) **S26**, **S30β**, TBSOTf 0.05eq, DCM, -78°C, **8** 6%, **9** 33%. b) **S26**, **S29α**, TBSOTf 0.05eq, DCM, -78°C, No reaction. c) **S24**, **S30β**, TBSOTf 0.05eq, DCM, -20°C, **10** 43%. d) **S24**, **S29α**, TBSOTf 0.05eq, DCM, -20°C, **11** 27%. e) **S25**, **S29α**, TBSOTf 0.05eq, DCM, -50°C-0°C, **4α** 49%, **4β** 7%, α:β 7:1. f) **S25**, **S30β**, TBSOTf 0.05eq, DCM, -78°C-0°C, **5** 20%. g) **S27**, **S30β**, TBSOTf 0.05eq, DCM, -78°C, **3β** 10%, **2** 70%. h) **S27**, **S29α**, TBSOTf 0.05eq, DCM, -78°C-0°C, **6** 14%. i) **S14α**, NIS, AgOTf 0.5eq, DCM, -50°C-0°C, **1** 63%. j) **S24**, TBSOTf 0.1eq, DCM, -50°C-0°C, **1** 55%.

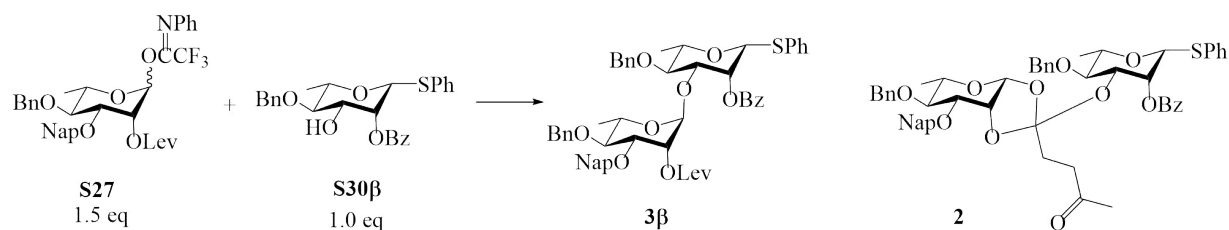
Firstly, we examined the electronic effects of C-4 position on both donor and acceptor with PMB at the C-3 position on the donor. Bn and Ac were selected to be the standard electron donating and electron withdrawing groups, and donor **S26** or **S24** reacted with acceptor **S29α** or **S30β** respectively. We started with low temperature and mild promoting catalyst since the 6-Deoxy rhamnose substrate is very reactive. For the reaction of **S26** with **S30β**, the reactivity of the substrate was moderate, and low temperature froze the glycosylation at the orthoester stage. For the reaction of **S24** with **S29α**, the reactivity of the substrate was low and only the desired α linkage was formed. For the C4-Ac and C4-Bn mixed combination, the reactivity of the donor and the acceptor did not match which led to low yield or no reaction.

Secondly, we examined the efficiency of the leaving group during the glycosylation and the influence of the linker on the anomeric center for this substrate. When donor **S24** or **S14α** reacted with acceptor **S31**, surprisingly PMB migrated from the donor to the acceptor and

no desired disaccharide product was formed (see **Table 1**, Supporting information). It can migrate under very mild conditions, and it is substrate dependent. Thus, PMB protecting group is not suitable in this case, and we turned our attention towards Nap protecting group, as it is more acid stable comparing to PMB.

Then, we examined the electronic effects of C-4 position on both donor and acceptor with Nap at the C-3 position on the donor. Bn and Ac were selected to be the standard electron donating and electron withdrawing groups, and donor **S25** or **S27** reacted with acceptor **S29 α** or **S30 β** respectively. For the reaction of **S25** with **S29 α** , C4-Ac on the donor led to α/β isomers. This is due to the long-range participation of the C4-Ac on the donor, even if in the presence of the Lev neighbor group participation. For the C4-Ac and C4-Bn mixed combination, the reactivity of the donor and the acceptor did not match, which led to the low yield of the reaction. The best reactivity was achieved by reacting donor **S27** with **S30 β** . Therefore, the optimal protecting group combination is C3-Nap on the donor and C4-Bn on both donor and acceptor.

Screening of the optimal reaction condition



Scheme 3. Optimization of the disaccharide assemble condition

Entry	Temperature/°C	Catalyst and solvent	Yield
1	-78	0.05 eq TBSOTf, DCM	3β 10%; 2 70%
2	-78~rt	0.1 eq TBSOTf, DCM	3β 75% only
3	-30~rt	0.1 eq TBSOTf, DCM	3β 57% only
4	-30~rt	0.1 eq TMSOTf, DCM	3β 81% only
5	-30~rt	0.25 eq TMSOTf, DCM	3β 43% only
6	-30~rt	0.1 eq HOTf, DCM	3β 46% only
7	0~rt	0.1 eq TMSOTf, DCM	3β Trace
8	-30~rt	0.1 eq TMSOTf, CH ₃ CN	No reaction

Table 2. Model test results for the optimization of the disaccharide assemble condition

After we selected the protecting group combination, we conducted a thorough reaction condition screening as it shown in **Table 2**. Firstly, we raised the temperature, and the yield improved dramatically, as the low temperature would freeze the glycosylation process at the orthoester stage. However, the glycosylation process for this substrate majorly happened below 0 °C, since Entry 7 shows a trace amount of product forming under this condition. Secondly, we increased the strength of the promoting catalyst by using less hindered TMSOTf instead of TBSOTf, which improved the efficiency of the glycosylation. However, the use of too much acid

or too strong acid will decrease the reaction yield. Thirdly, DCM is generally the best solvent for this substrate. In summary, the optimal reaction condition for this substrate is $-30\text{ }^{\circ}\text{C}\sim 0\text{ }^{\circ}\text{C}$, 0.1 eq TMSOTf as the promoting catalyst, DCM as the solvent.

PMB migration

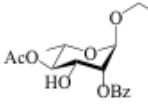
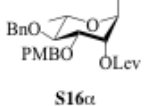
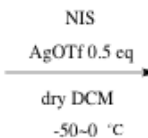
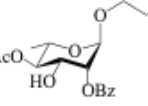
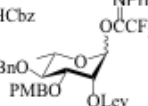
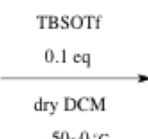
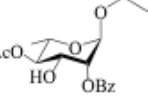
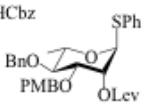
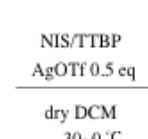
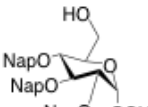
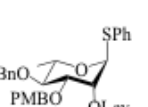
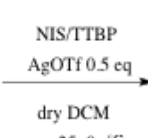
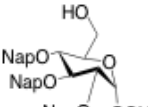
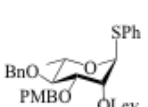
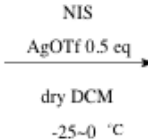
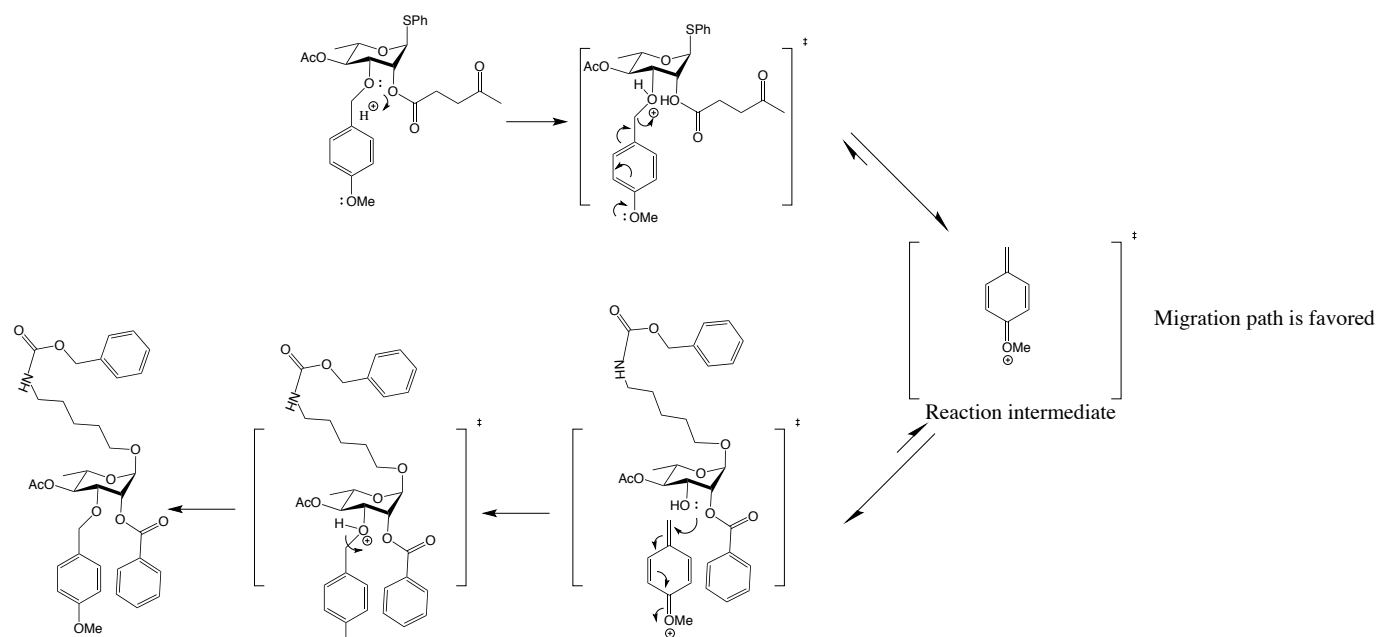
Acceptor 1.0eq	Donor 1.5eq	Reaction condition
 S31	 S16α	 NIS AgOTf 0.5 eq dry DCM -50-0 °C
 S31	 S26	 TBSOTf 0.1 eq dry DCM -50-0 °C
 S31	 S16α	 NIS/TTBP AgOTf 0.5 eq dry DCM -30-0 °C No reaction, no PMB migration.
 ND	 S16α	 NIS/TTBP AgOTf 0.5 eq dry DCM -25-0 °C No reaction, no PMB migration.
 ND	 S16α	 NIS AgOTf 0.5 eq dry DCM -25-0 °C No reaction, no PMB migration.

Table 3. PMB Migration reactions

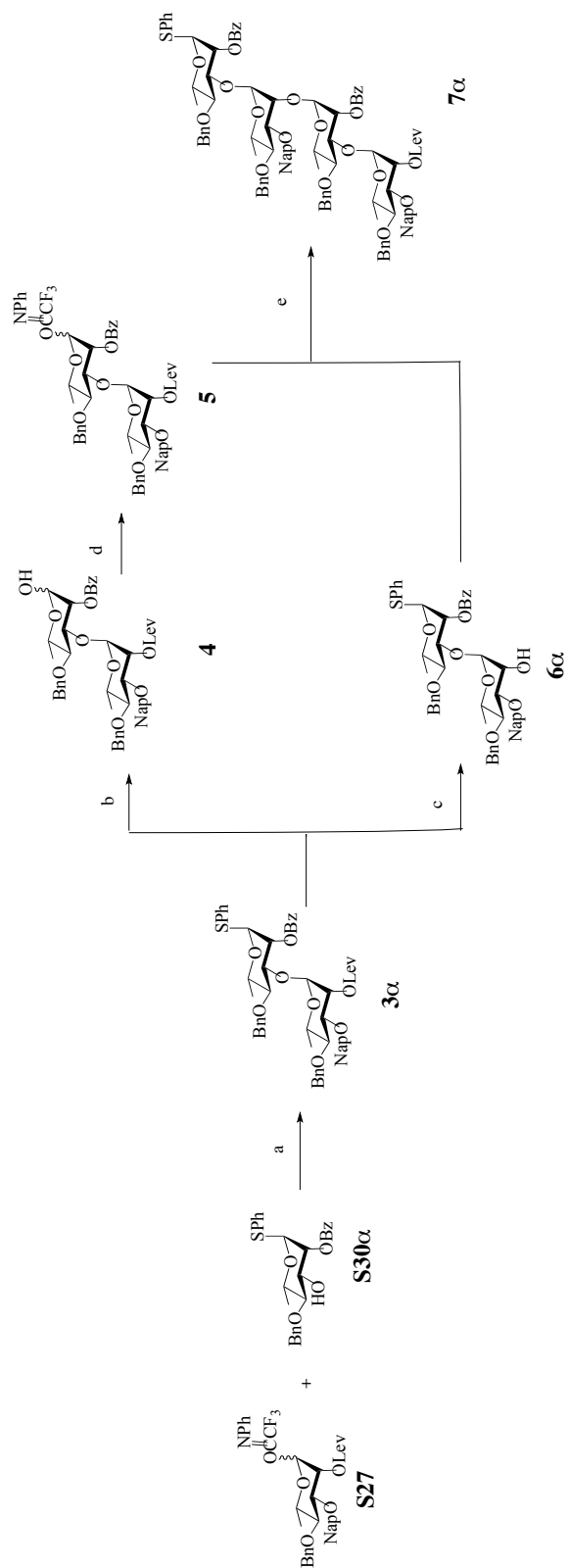
It can migrate under very mild conditions, and it is substrate dependent. Thus, PMB protecting group is not suitable in this case, and we turned our attention towards Nap protecting group, as it is more acid stable comparing to PMB.

Proposed PMB migration mechanism



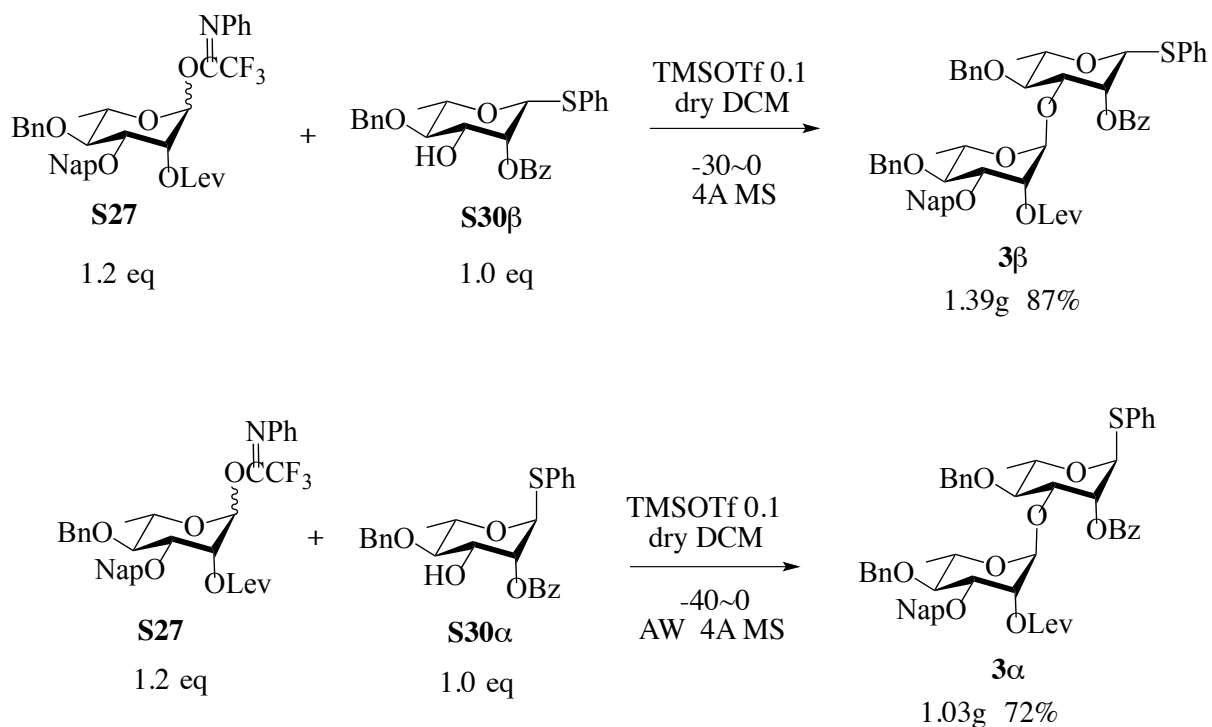
Scheme 4. Proposed PMB migration mechanism

The PMB protecting group is acid labile due to the p- π conjugation of the lone pair electron on the methoxy oxygen to the π system of the aromatic ring. The migration is probably an equilibrium process, and the migration path is favored because of the lower energy of the product. *This requires further supports from dynamic molecular simulation and theoretical calculations.



Scheme 4. Synthesis of the tetrasaccharide **7α**. Reagents and conditions: a) TMSOTf, DCM; 72% b) NIS, AgOTf, TTBP, wet DCM; 97%; c) Hydrazine acetate, DCM:MeOH(1:1), 93%; d) ClCNPhCF₃, Cs₂CO₃, DCM, 87%; e) TBSOTf, DCM, 85%.

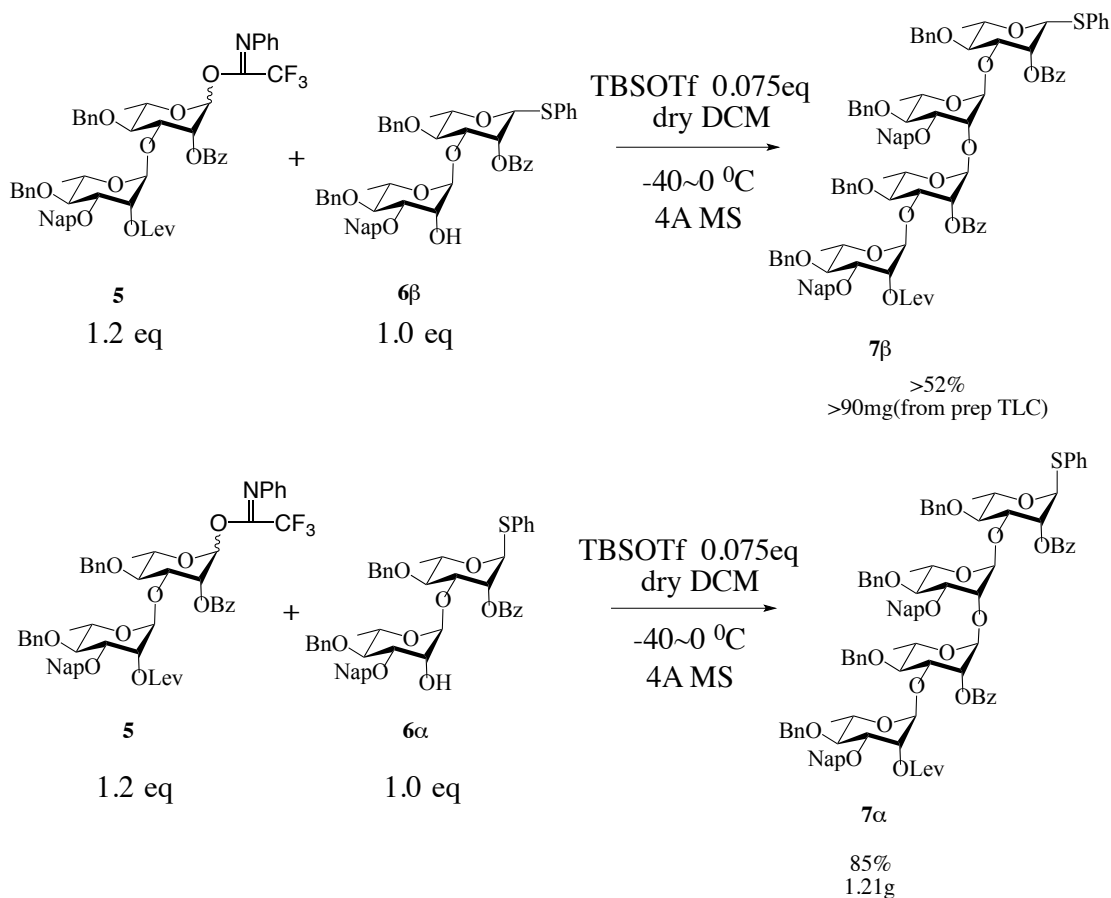
Disaccharide assembly



Scheme 5. Synthesis of the disaccharides

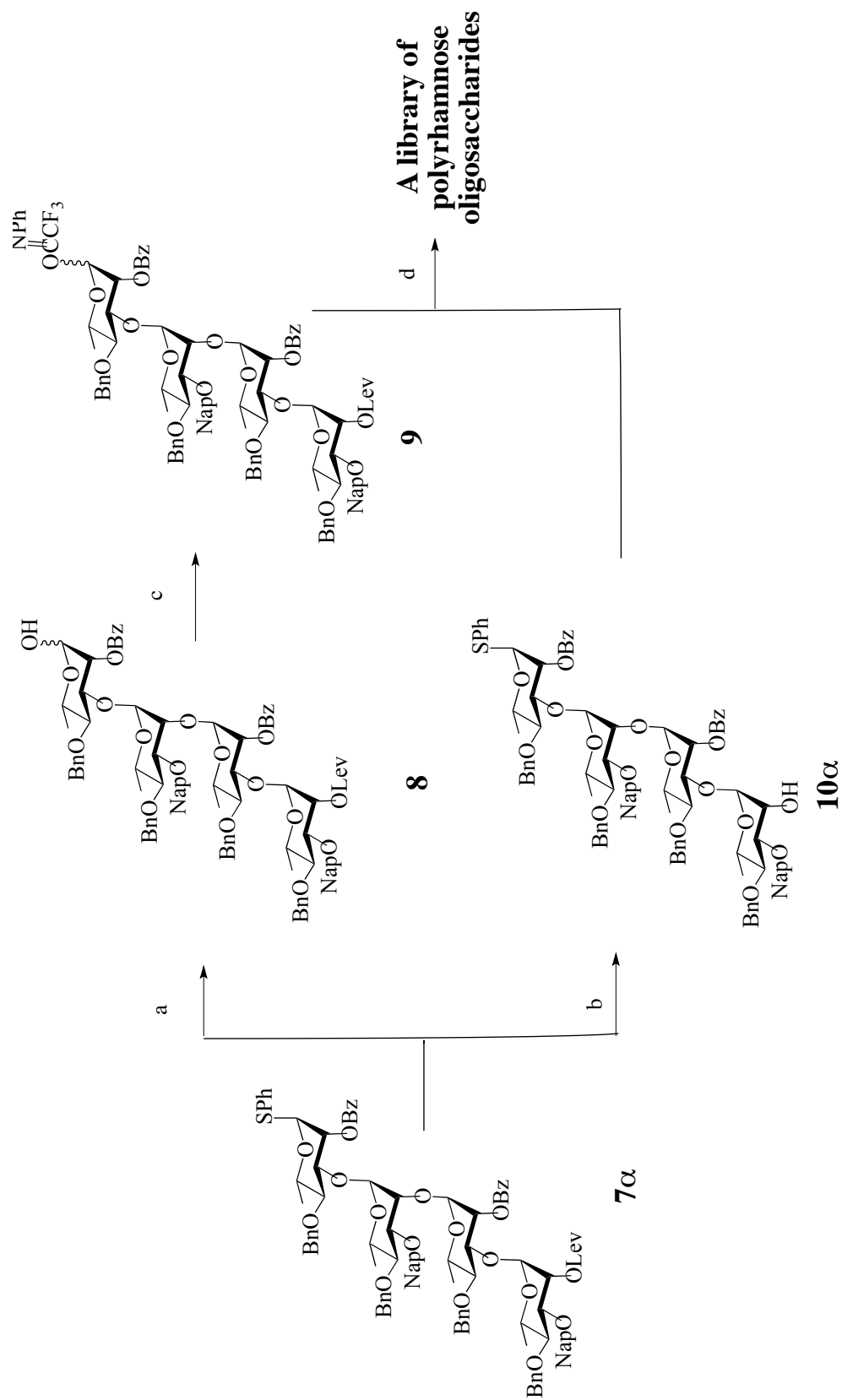
Disaccharide assembly was performed by the optimal glycosylation method as described before. Glycosylation processes are very smooth for both α and β acceptors. The configuration of the thioglycoside does not affect the glycosylation. The acid test shows that disaccharides are stable in acidic condition, and 1,3-linkage is acid stable. The acid test is conducted by adding a drop of TMSOTf (an excessive amount) into a concentrated solution (~ 0.1 M) of the disaccharide in DCM, and monitored the reaction by TLC and MALDI immediately.

Tetrasaccharide assembly

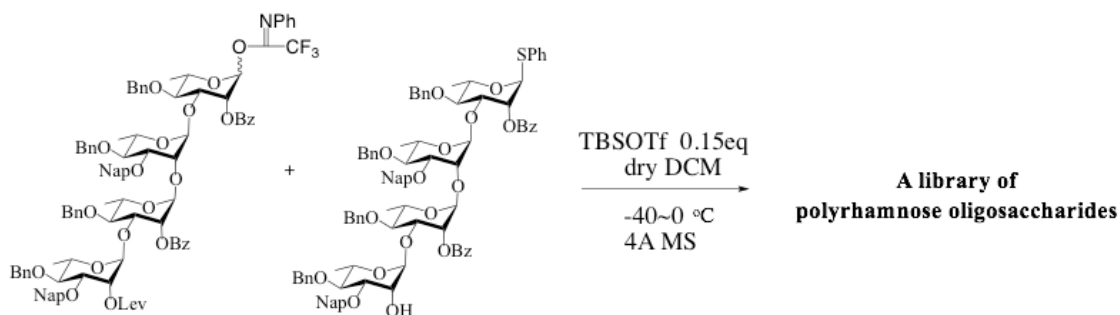


Scheme 6. Synthesis of the tetrasaccharides

In the case of tetrasaccharide assembly, milder catalyst TBSOTf was used, and the amount of acid was reduced, due to the observance of the product decomposition when using TMSOTf as the catalyst. The acid test shows that the tetrasaccharide is acid labile, which is decomposed to hydrolyzed disaccharide donor and disaccharide acceptor. Thus 1,2-linkage is acid labile in this case. The acid test is conducted by adding a drop of TBSOTf (an excessive amount) into a concentrated solution (~ 0.1 M) of the tetrasaccharide in DCM, and monitored the reaction by TLC and MALDI immediately.



“4+4” Polymerization



Scheme 8. “4+4”

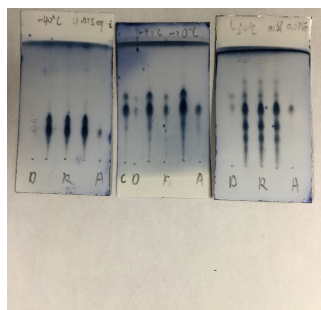


Figure 2. TLC of the “4+4” polymerization

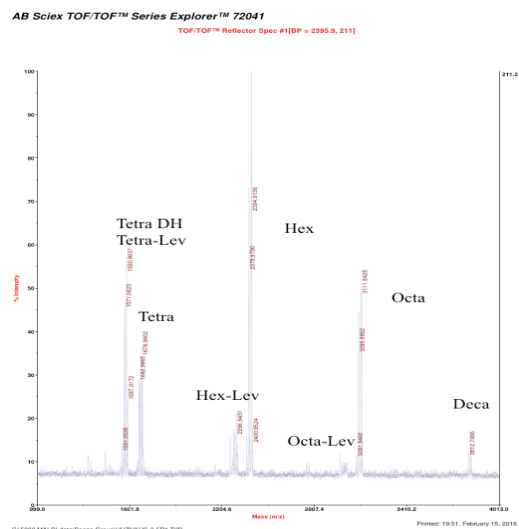


Figure 3. MALDI of the “4+4” polymerization reaction mixture

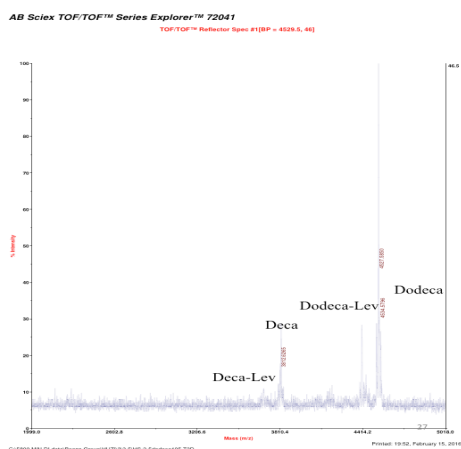


Figure 4. MALDI of the dodecasaccharide after LH20 purification

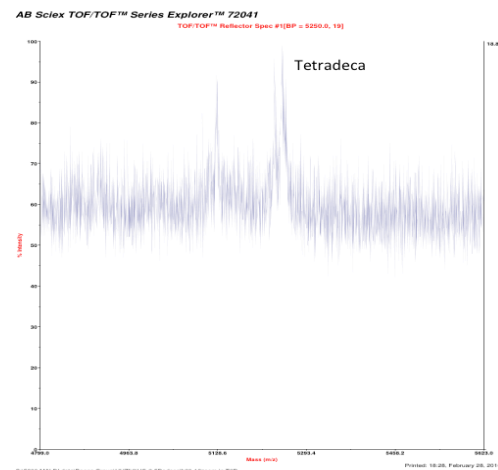


Figure 5. MALDI of the tetradecasaccharide after LH20 purification

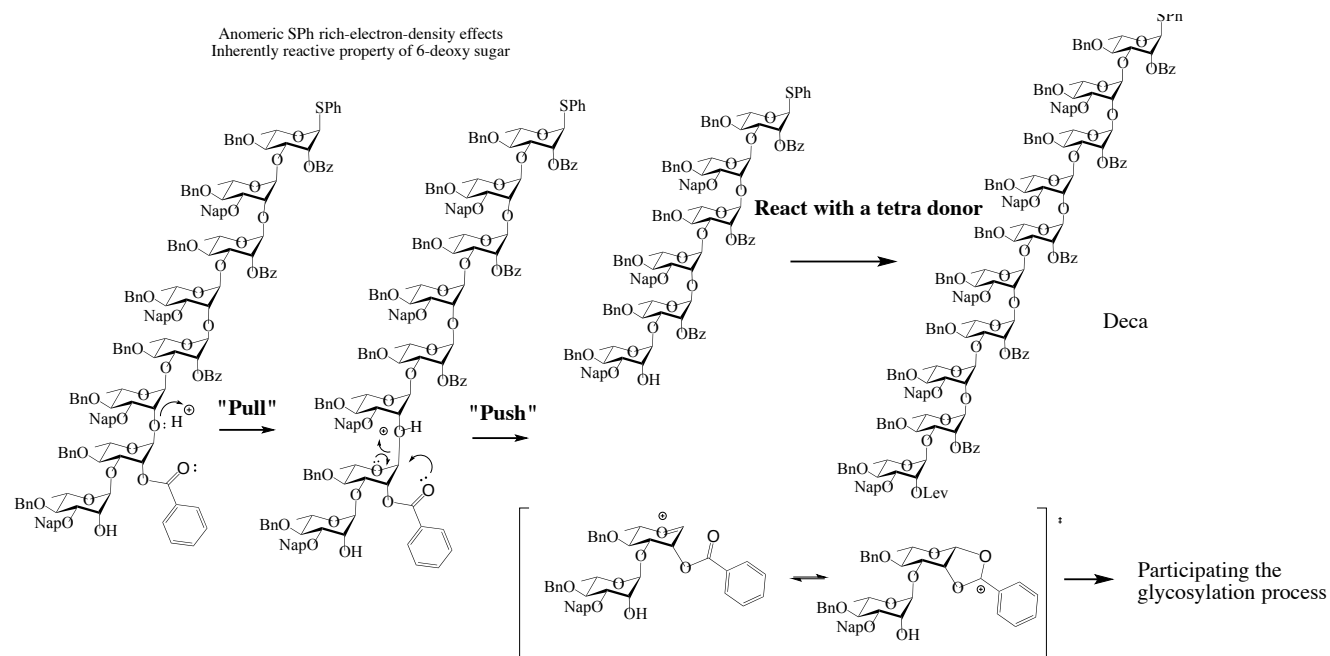
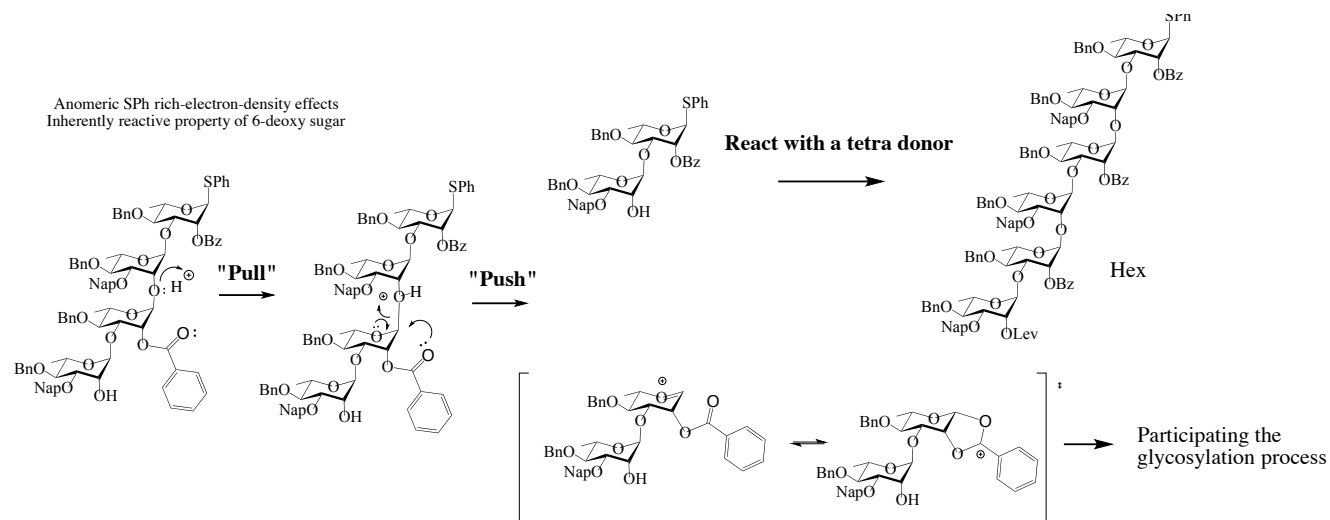
Di(DH+A)	Tetra(DH+Agly)	Hex(Hex+Hex-Lev)	Octa(Octa+Octa-Lev)	Deca(Deca+Deca-Lev)	Dodeca(Dodeca+Dodeca-Lev)	Tetradeca
71mg	135mg	116mg	101mg	23mg	9mg	Detected by MALDI

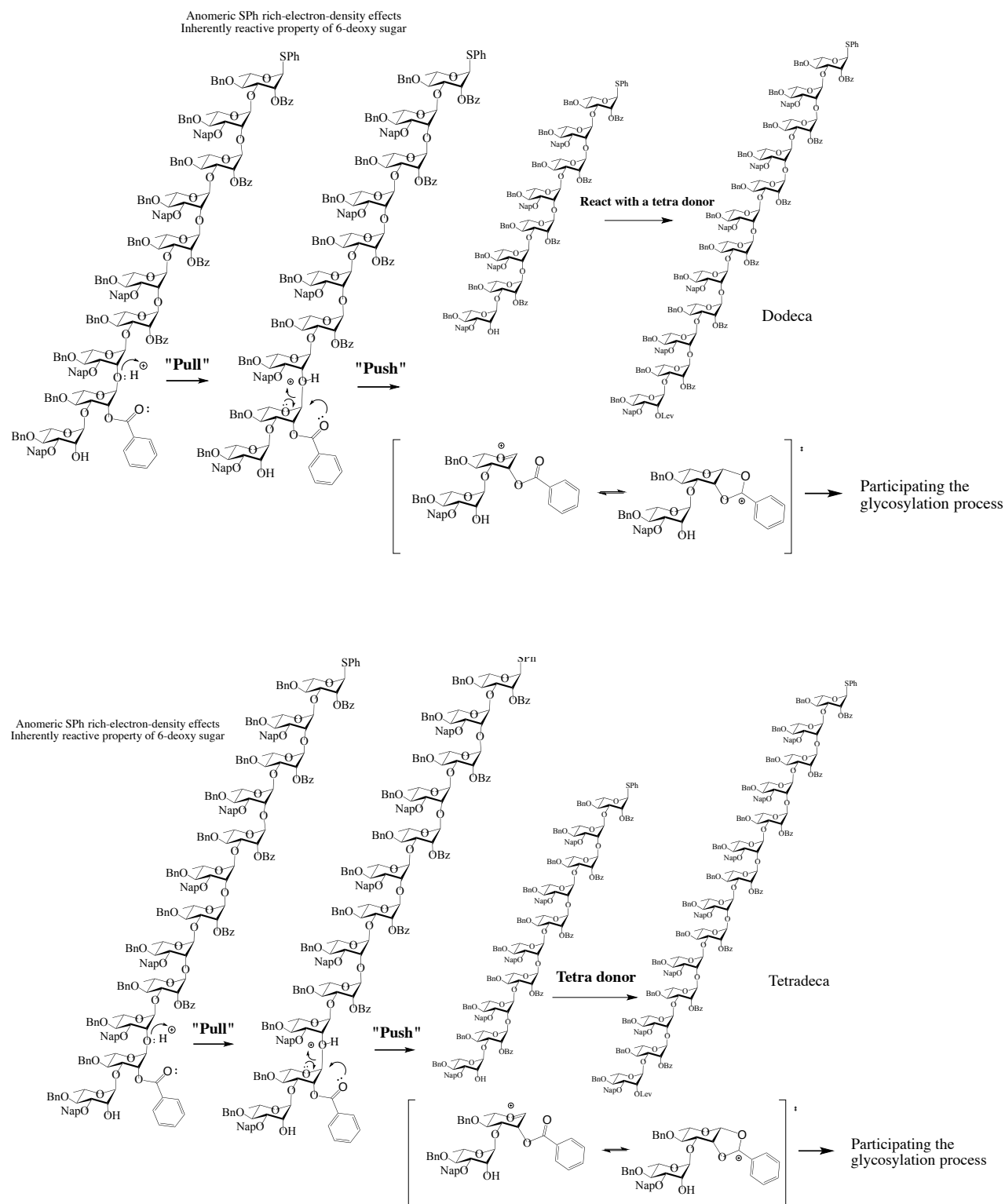
Table 4. A library of polyrhamnose oligosaccharides after LH20 purification

In the case of “4+4”, instead of yielding the octasaccharide specifically, it yielded a library of polyrhamnose oligosaccharides unexpectedly due to the *in situ* bond-cleavage polymerization. The products can be first purified by LH20, and further purification by silica gel is needed. The peaks corresponding to oligosaccharides without Lev are very important for deciphering the mechanism of this “4+4” polymerization, and we proposed a “Push and Pull” mechanism:

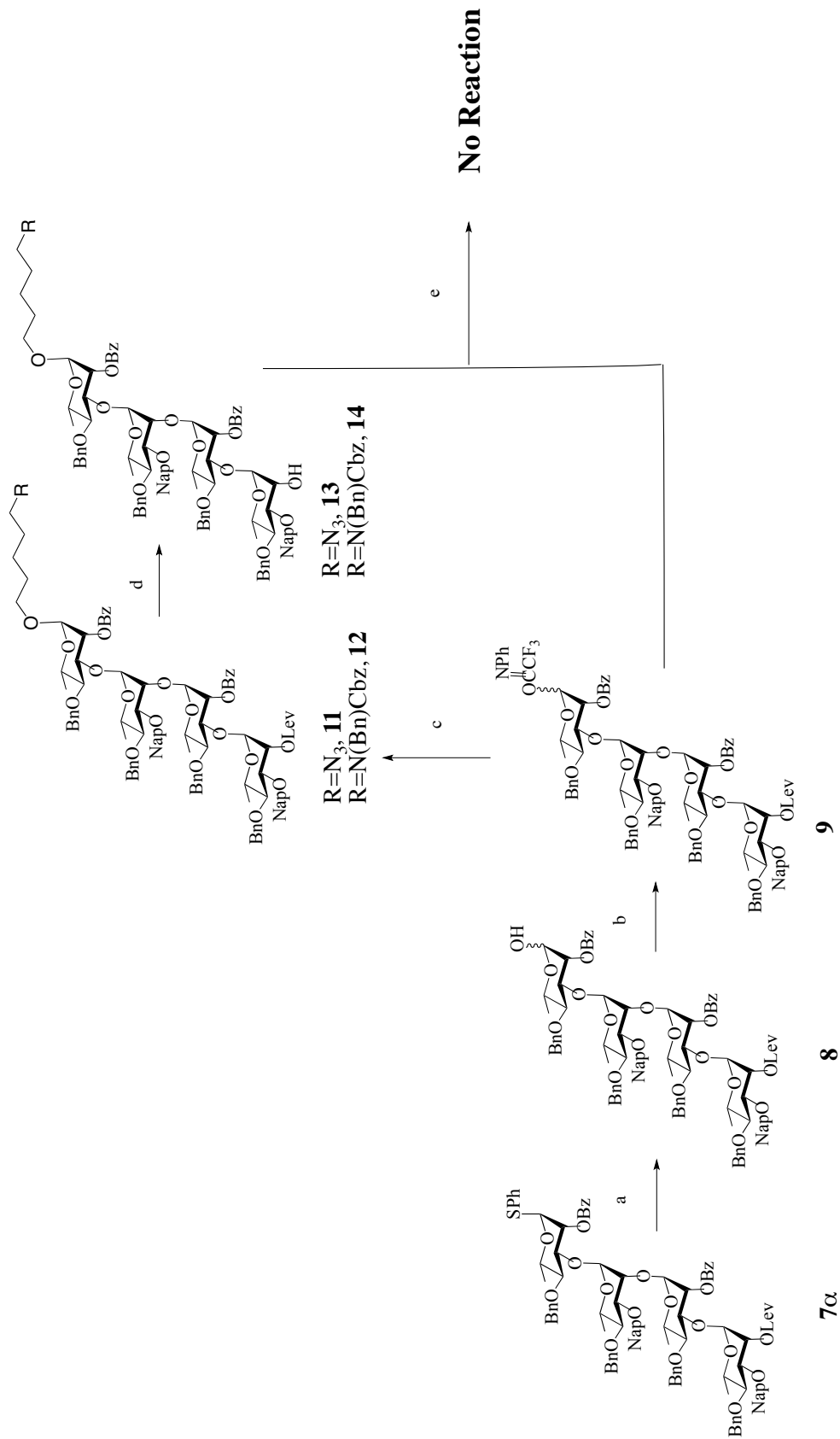
- 1) 1,3-linkage is acid stable, but 1,2-linkage is acid labile;
- 2) Rhamnose is very reactive and sensitive under acidic condition comparing with common pyranoses, due to the lacking of 6-OH;
- 3) Ester neighbor group participating effects give all α linkage;
- 4) The electron richness of the thiophenyl protecting group at the anomeric center increases the overall reactivity;
- 5) The steric hindrance and electron donating properties of the Nap protecting group plays a role.

Proposed “Push and Pull” mechanism for “4+4” polymerization



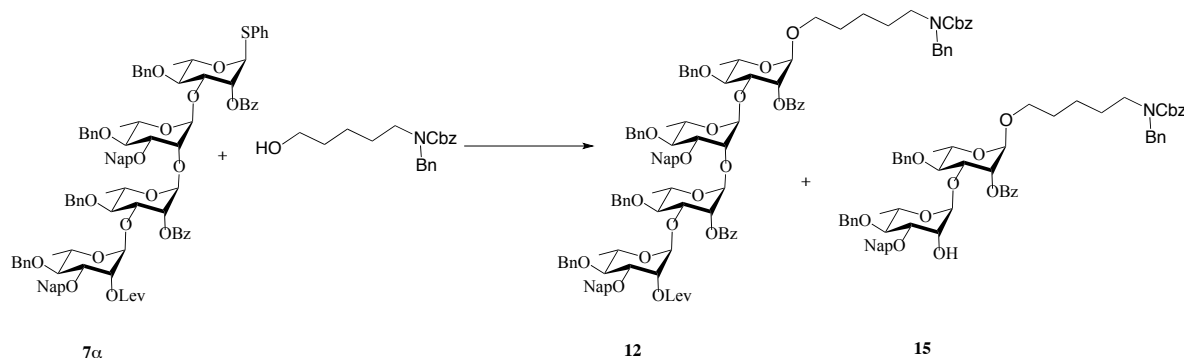


Scheme 9. Proposed "Push and Pull" mechanism



Scheme 10. Unsuccessful trial synthesis of the octasaccharide. Reagents and conditions: a) NBS, Acetone:PBS Buffer(6:1); 87%; b) CICNPhCF_3 , Cs_2CO_3 , DCM, 87%; c) $\text{R}=\text{N}_3$, 5-azido-1-pentanol, TBSOTf, DCM, 92%. $\text{R}=\text{N}(\text{Bn})\text{Cbz}$, N,N-Bn , Cbz -5-amino-1-pentanol, TBSOTf, DCM, 36% d) Hydrazine acetate, DCM:MeOH(1:1), $\text{R}=\text{N}_3$, 93%. $\text{R}=\text{N}(\text{Bn})\text{Cbz}$, 85%; e) TBSOTf, DCM

Linker problems

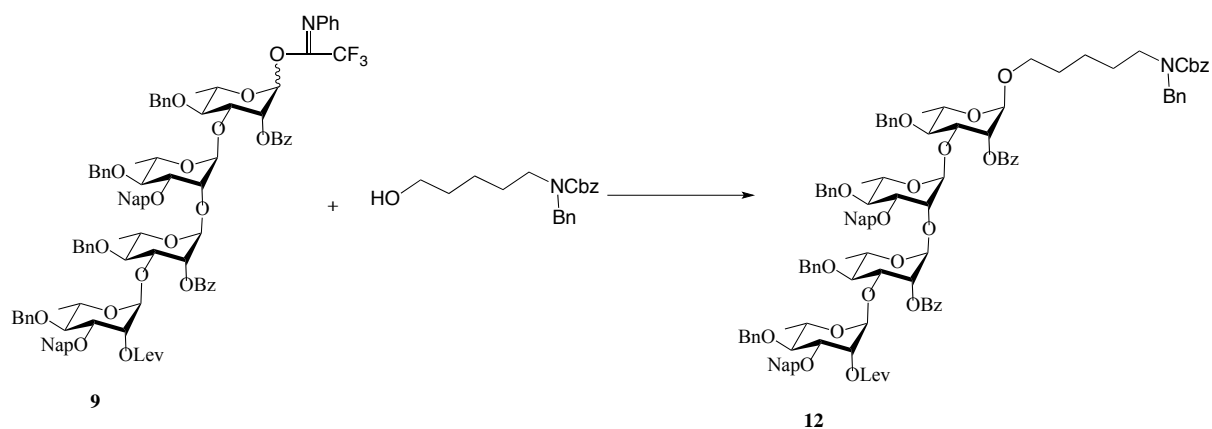


Scheme 11. Linker addition-SPh

Solvent	Catalyst	Temp/°C	Product	Byproduct	Yield
DCM	NIS/AgOTf	-40-0	Minor	Major	13%
DCM	BSP/TTBP/Tf ₂ O	-78	No	No	No
DCM	NIS/TBSOTf	-40-0	Minor	Major	Low
toluene:dioxane (1:3)	NIS/AgOTf	0	Minor	Major	Very low

Table 5. Conditions for linker addition reactions

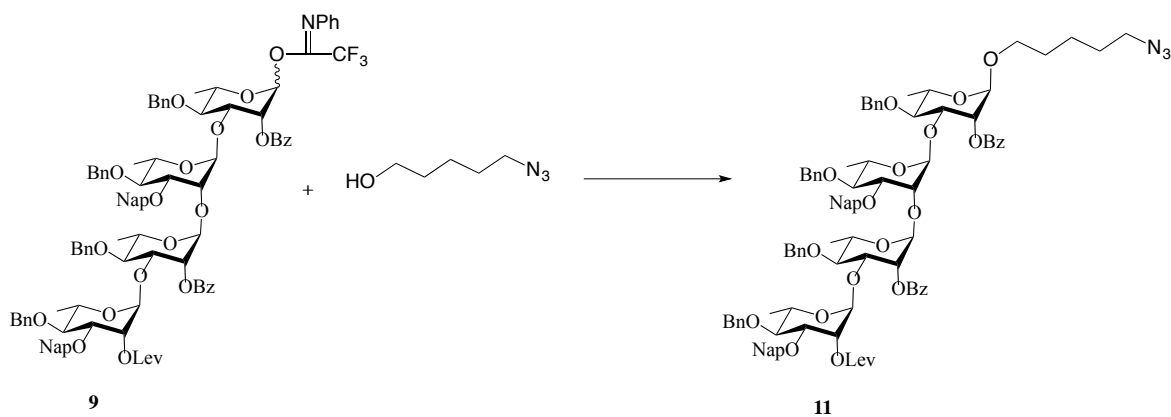
In this case, instead of yielding the tetrasaccharide specifically, it yielded the disaccharide from the breakage of the acid labile 1,2-linkage as the major product, and the yield of the desired tetrasaccharide were very low. At that time, we thought the low yield might come from the low efficiency of the thio glycoside and the influence of the lone pair of electrons from the linker nitrogen, so we decided to try different linker and different leaving group.



Scheme 12. Linker addition with *N*-Phenyl imidate

Solvent	Catalyst	Temp/°C	Product	Byproduct	Yield
DCM	TBSOTf 0.15 eq	-40-0	Main	No	36%

Table 6. Reaction condition for linker addition with *N*-Phenyl imidate

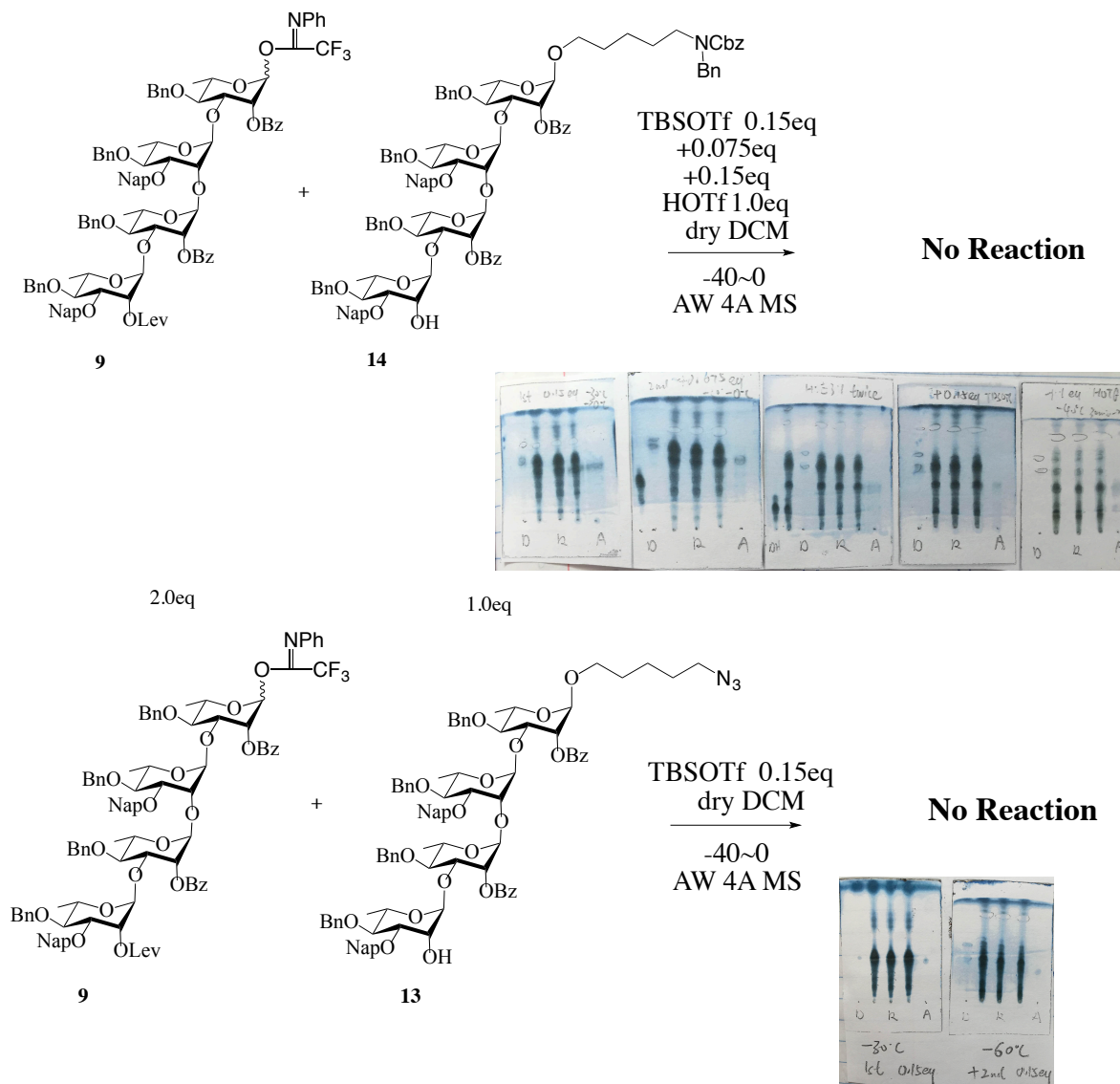


Scheme 13. Linker addition with *N*-Phenyl imidate

Solvent	Catalyst	Temp/°C	Product	Byproduct	Yield
DCM	TBSOTf 0.15 eq	-40-0	Main	No	92%

Table 7. Reaction condition for N₃-linker addition with *N*-Phenyl imidate

With *N*-phenyl imidates as the leaving group, both reactions yield specifically the tetrasaccharide product without the disaccharide byproduct. The yield indeed improved dramatically when using the azide linker as the acceptor. The thio phenyl at the anomeric center plays a significant role in those reactions.

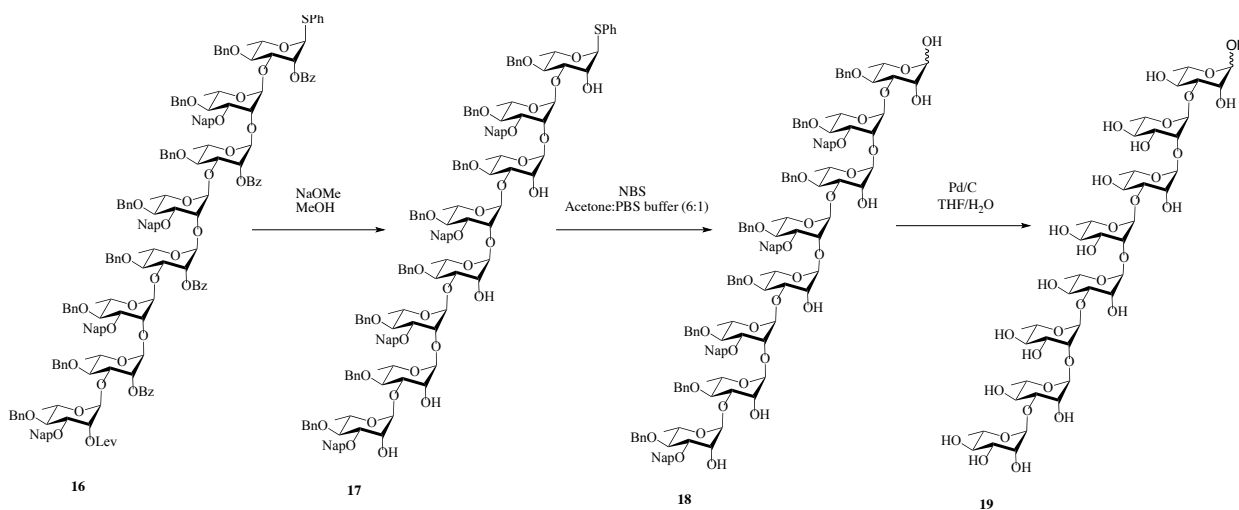


Scheme 14. Unsuccessful trail synthesis of the octasaccharide

Disappointedly, both reactions did not work. Donors are totally hydrolyzed in the end, and acceptors are left unreacted. The reactions were monitored by both TLC and MALDI, and no

product was detected. On the other hand, polymerization did not occur, which suggests that thio phenyl plays a very important role in the polymerization process. At this stage, we thought that the steric hindrance of the Nap at C-3 of the acceptor might cause the trouble in the glycosylation process. Therefore, we decided to replace Nap by Bn.

Deprotection



Scheme 15. Deprotection of the polyrhamnose oligosaccharide

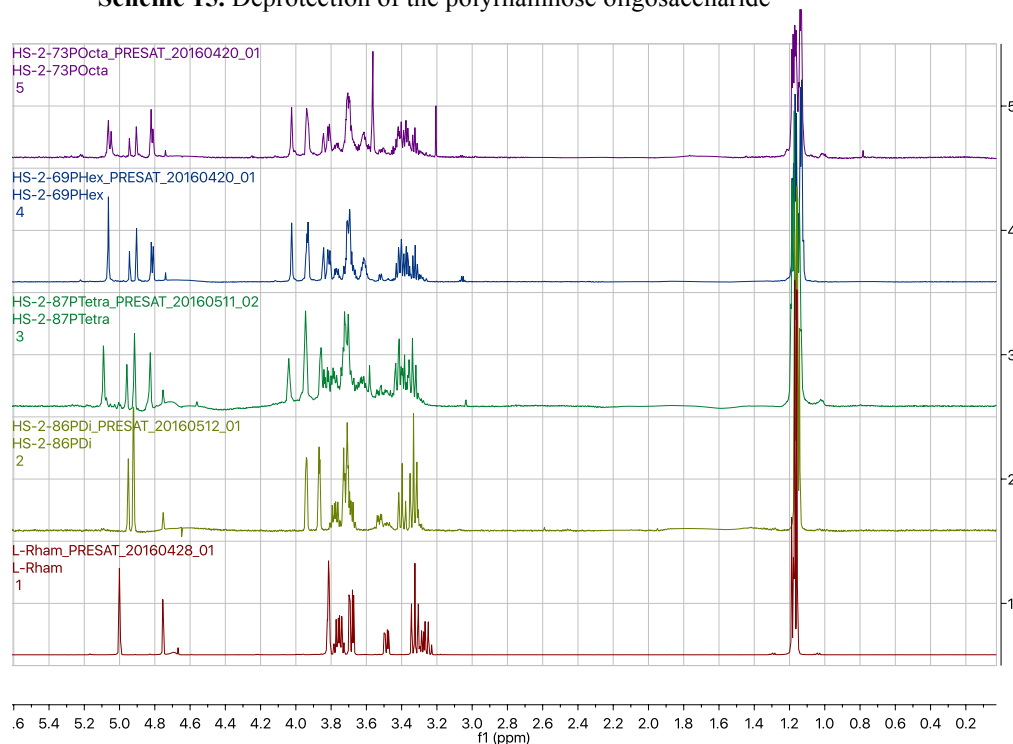
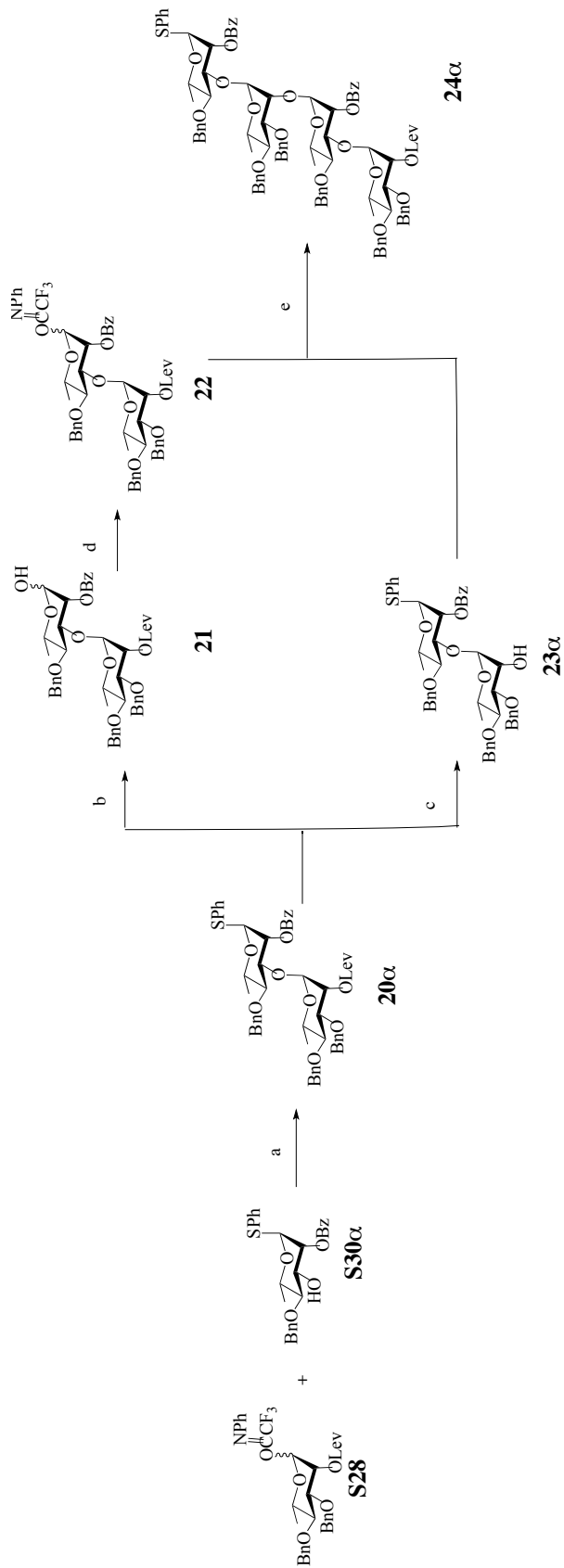


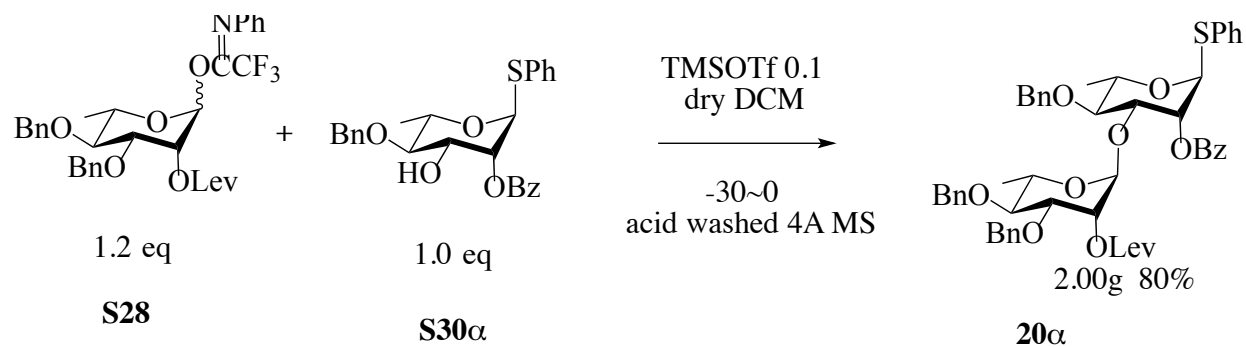
Figure 6. NMR spectra of the deprotected polyrhamnose oligosaccharides

Modular synthesis: replace Nap by Bn

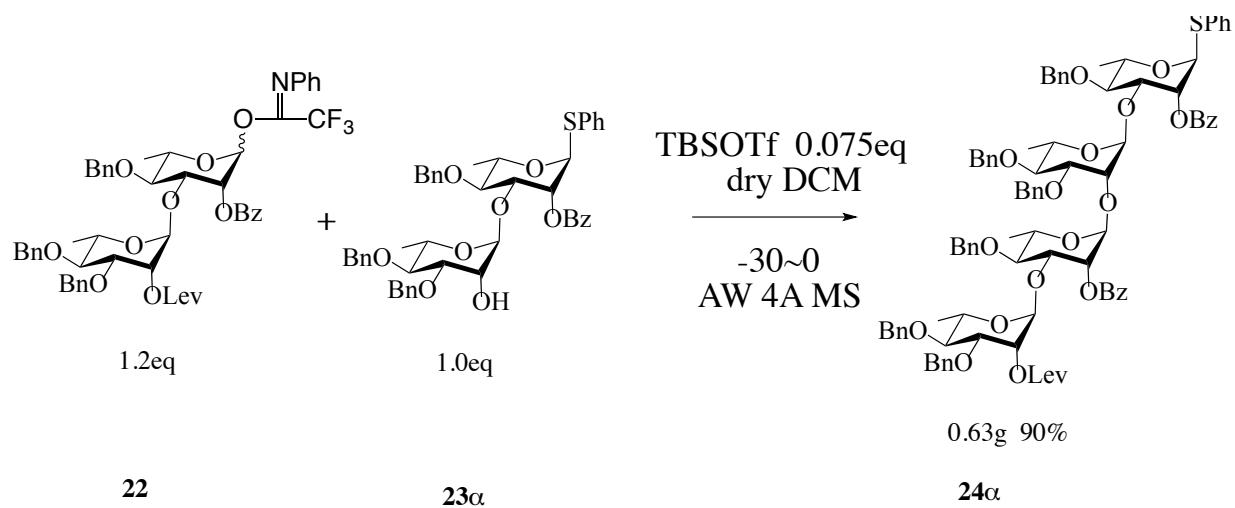


Scheme 16. Synthesis of the tetrasaccharide **24α**. Reagents and conditions: a) TMSOTf, DCM; 79% b) NBS, Acetone:PBS Buffer(6:1), 84%; c) Hydrazine acetate, DCM:MeOH(1:1), 80%; d) ClCNPhCF₃, Cs₂CO₃, DCM, 80%; e) TBSOTf, DCM, 90%.

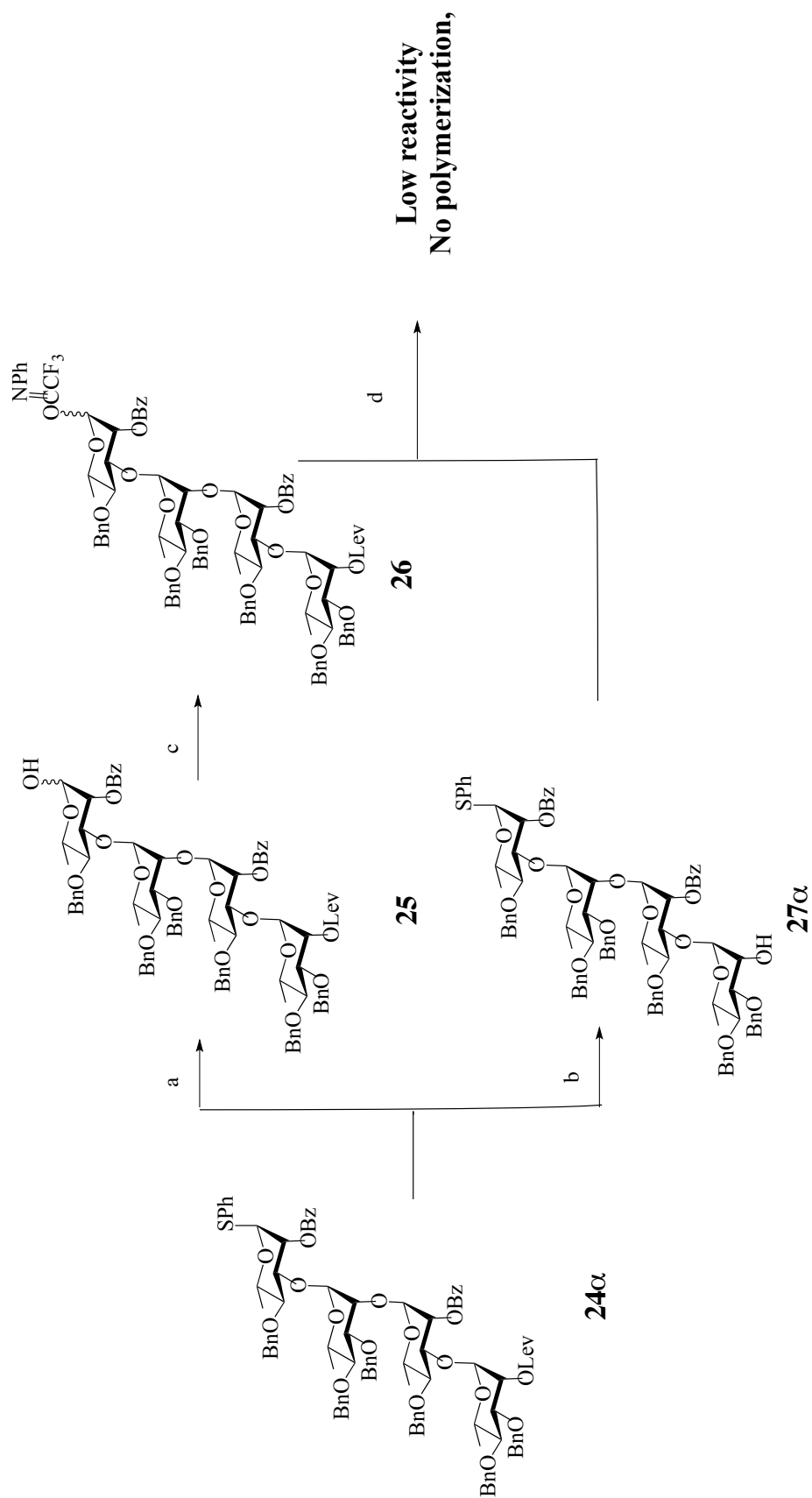
Disaccharide assembly



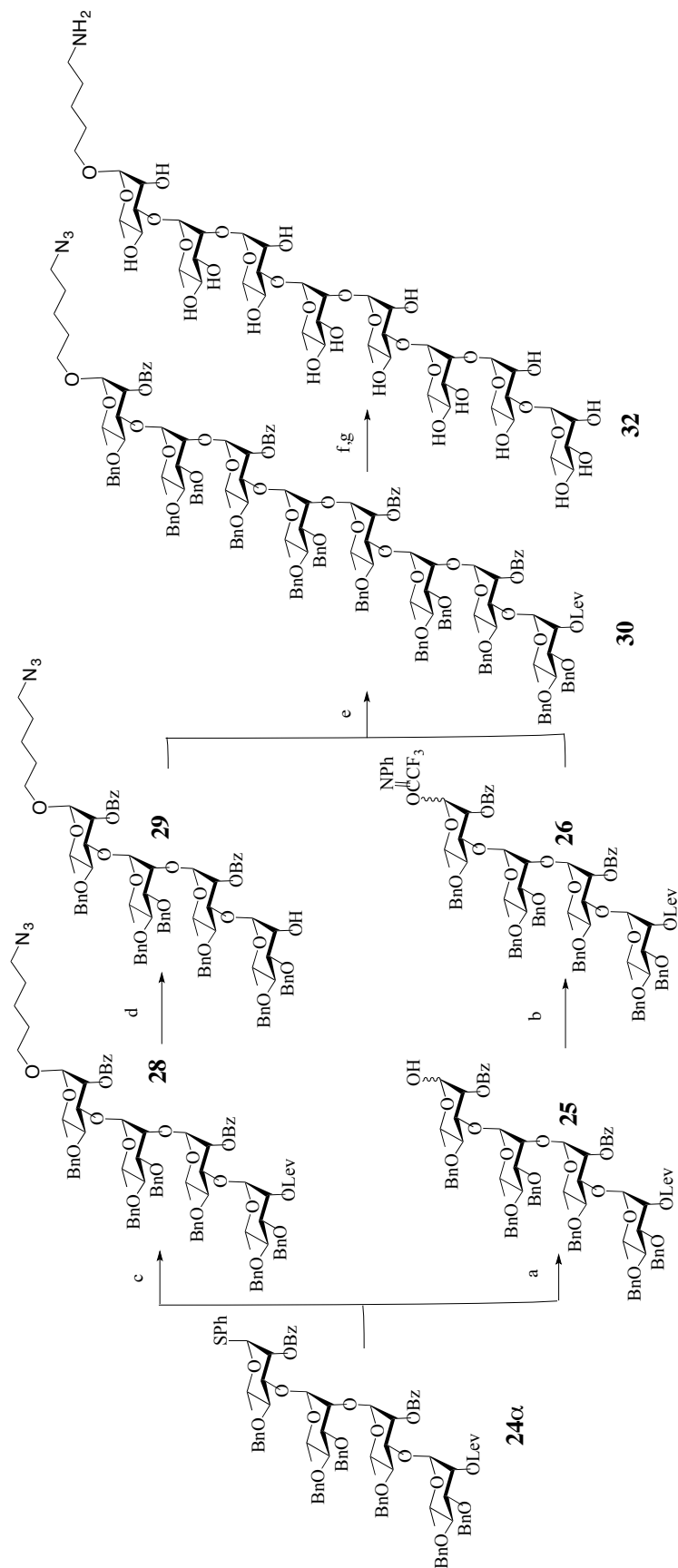
Tetrasaccharide assembly



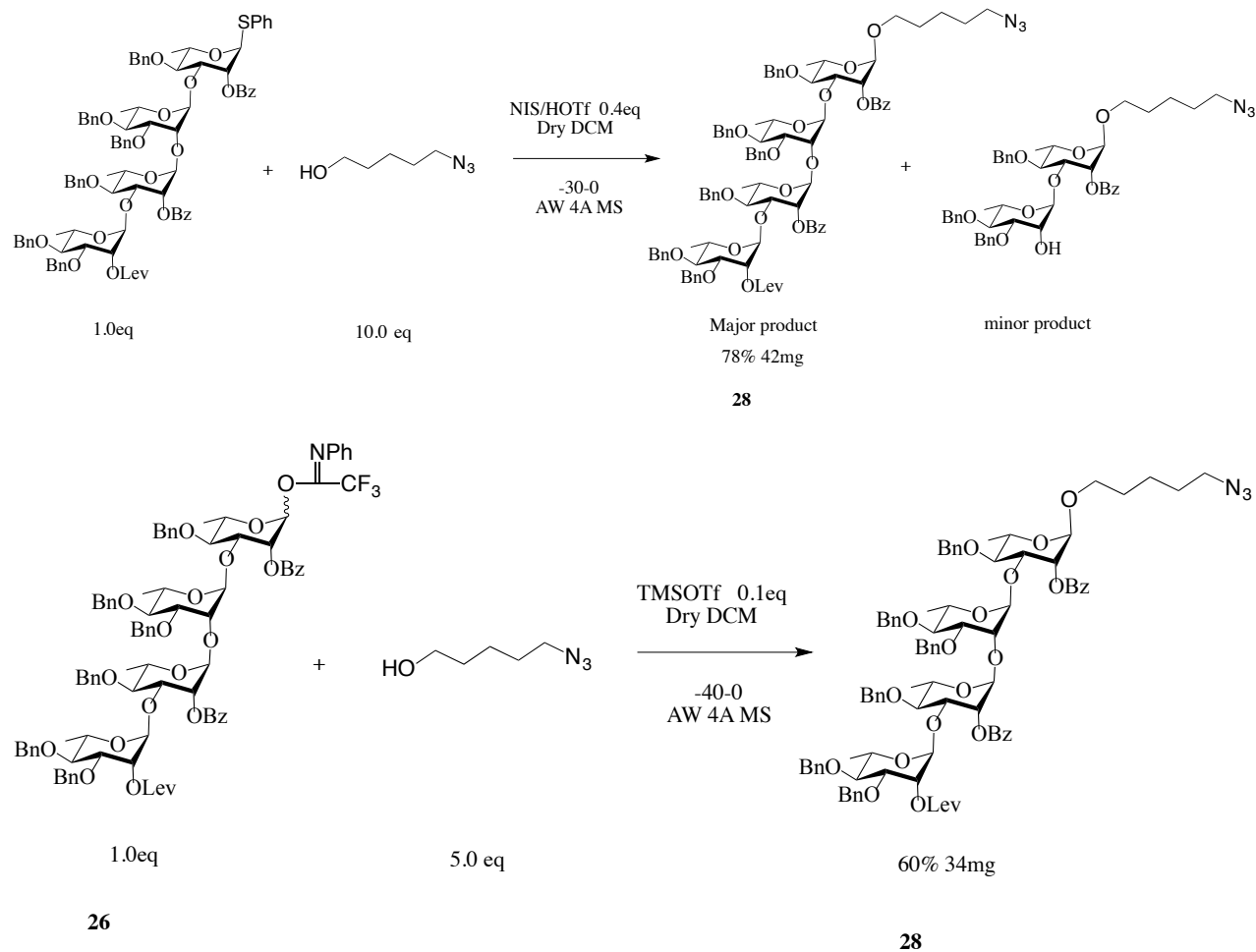
Scheme 17. Assembly of tetrasaccharide **24α**



Scheme 18. 4+4-Bn. Reagents and conditions: a) NBS, Acetone:PBS Buffer(6:1); 92%
b) Hydrazine acetate, DCM:MeOH(1:1), 81%; c) ClCNPhCF₃, Cs₂CO₃, DCM, 94%; d) TBSOTf, DCM.

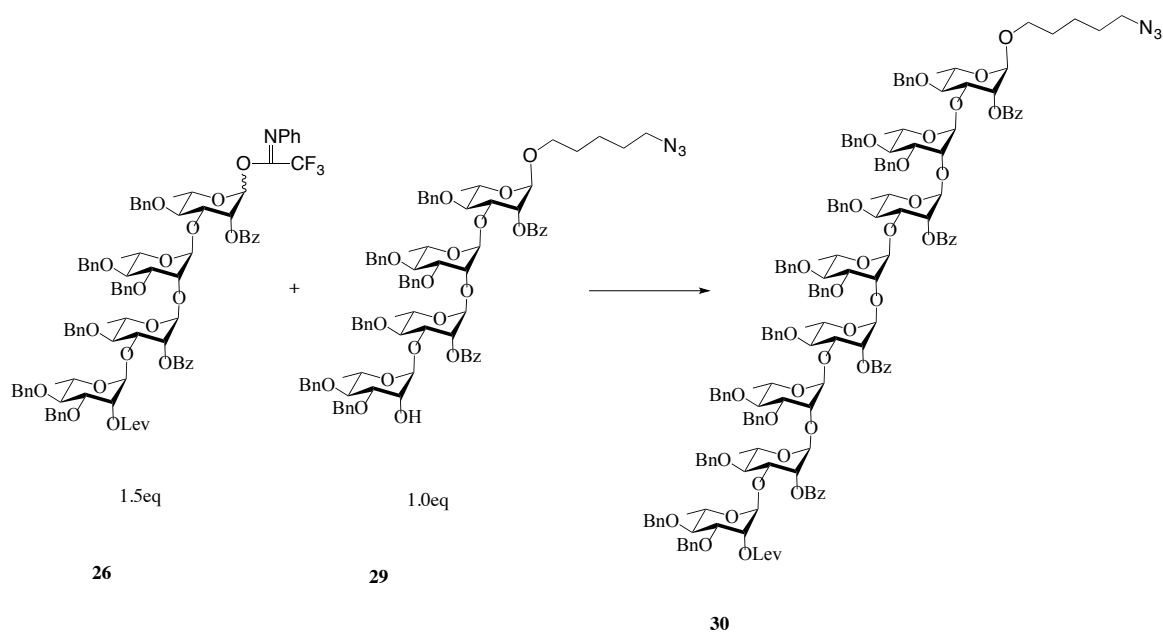


Scheme 20. Synthesis of the octasaccharide **32**. Reagents and conditions: a) NBS, Acetone:PBS Buffer(6:1); 87%; b) ClCNPhCF₃, Cs₂CO₃, DCM, 87%; c) 5-azidopentanol, NIS, HOTf, DCM, 78%. d)Hydrazine acetate, DCM:MeOH(1:1), 94%; e) TBSOTf, Toluene, -30 °C, 62%. f)NaOMe, MeOH; g) 10% Pd/C, THF/H₂O, 94% 2steps.



Scheme 21. Addition of linker

The reactions of coupling the linker to the tetrasaccharide were tested by two different leaving group strategies that we have used before. In the case of the thio glycoside, the yield of the desired tetrasaccharide were good, but it still yielded two products because of the bond-cleavage. However, the desired tetrasaccharide, in this case, is the major product, and the bond-cleavage disaccharide is the minor product. This is the result of the lower reactivity of the substrate, and it suggests that Nap plays a role in the glycosylation process again. In the case of the *N*-phenyl imidate, it specifically yielded the desired tetrasaccharide product in good yield.



Entry	Solvent	Catalyst	Temperature/°C	Yield
1	DCM	0.3 eq TBSOTf	-40~0	33%
		Acid washed 4A MS		
2	toluene	0.3 eq TBSOTf	-30~0	62%
		Acid washed 4A MS		

Scheme 22. Synthesis of octasaccharide **30**

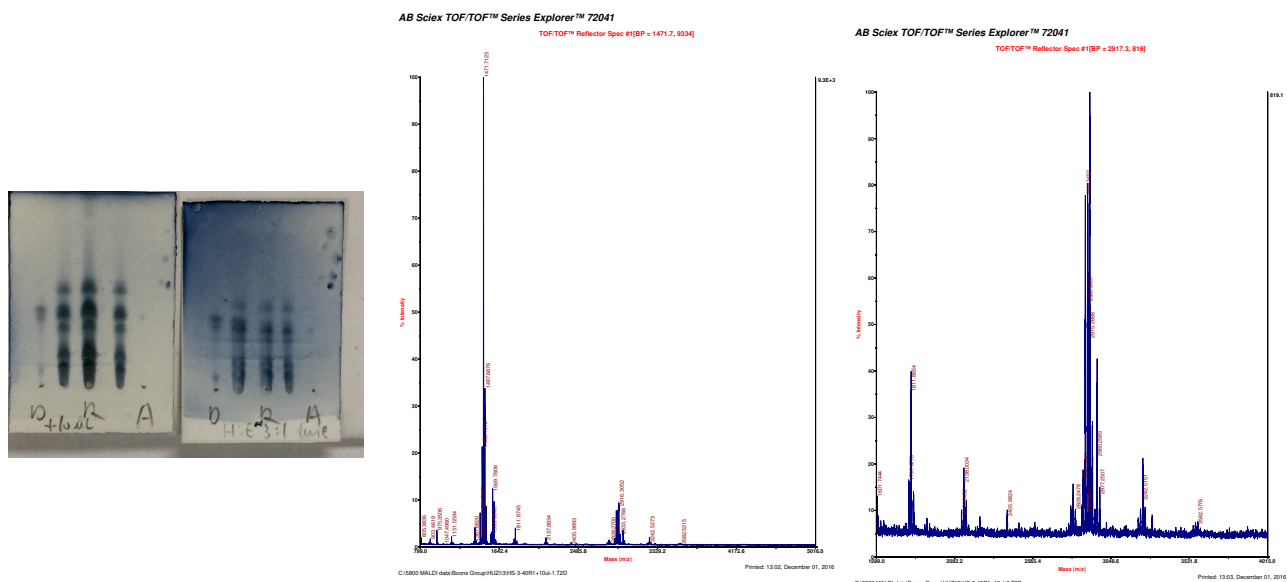
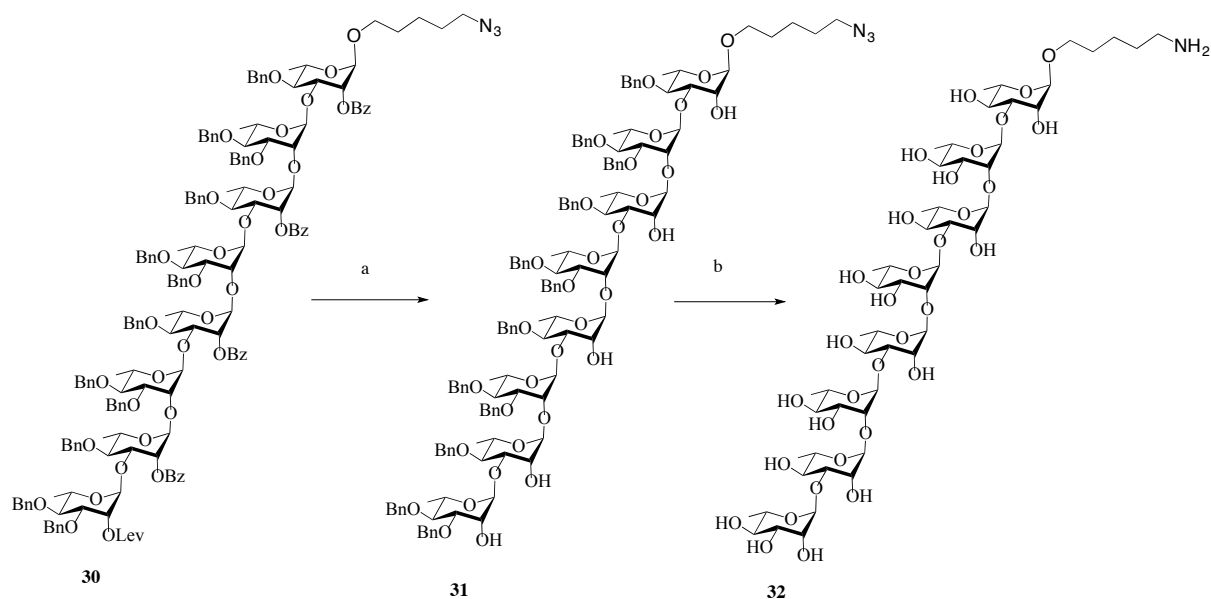


Figure 8. TLC and MALDI of the octasaccharide reaction

Deprotection



Scheme 23. Deprotection of Octasaccharide **30**. Reagent and conditions: a) NaOMe, MeOH; b) 10% Pd/C, THF/H₂O, 94% 2steps.

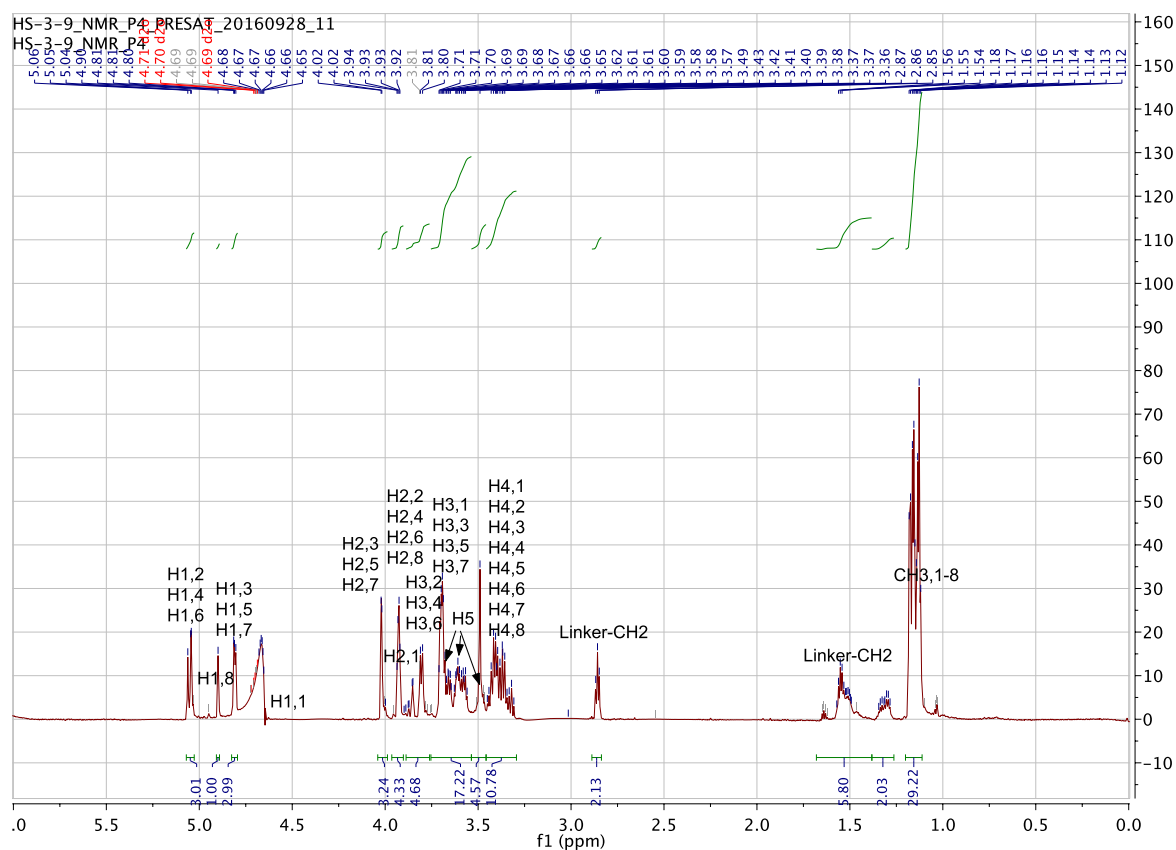
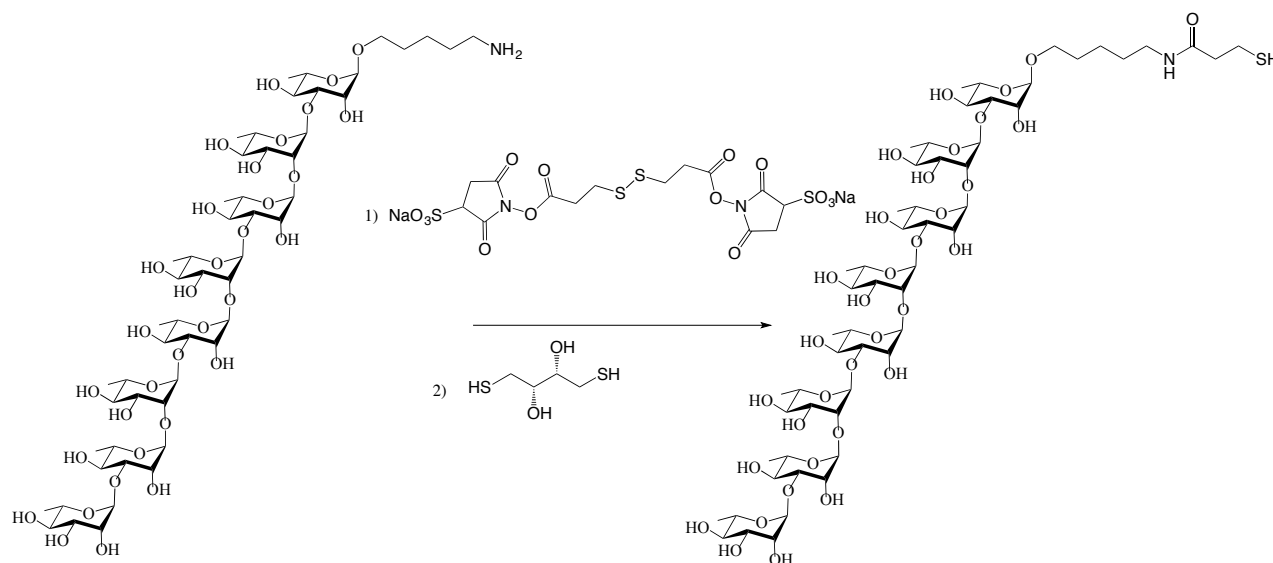


Figure 9. NMR spectra of the deprotected octasaccharide **32**

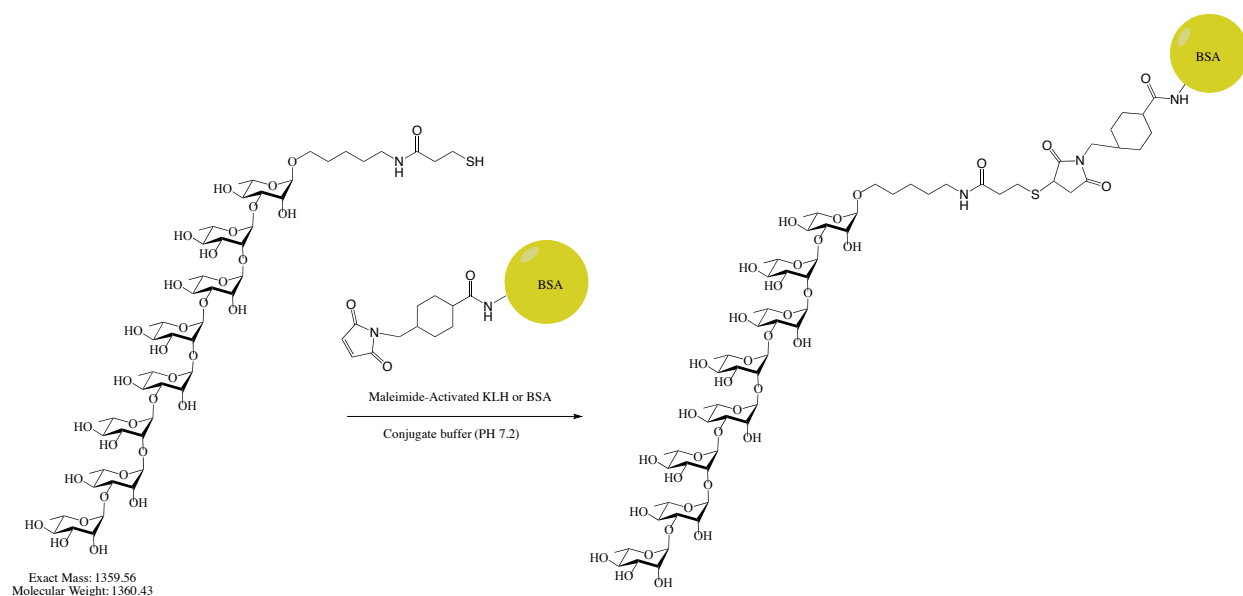
Conjugation of the final product to carrier proteins



Scheme 24. Modification of the final octasaccharide

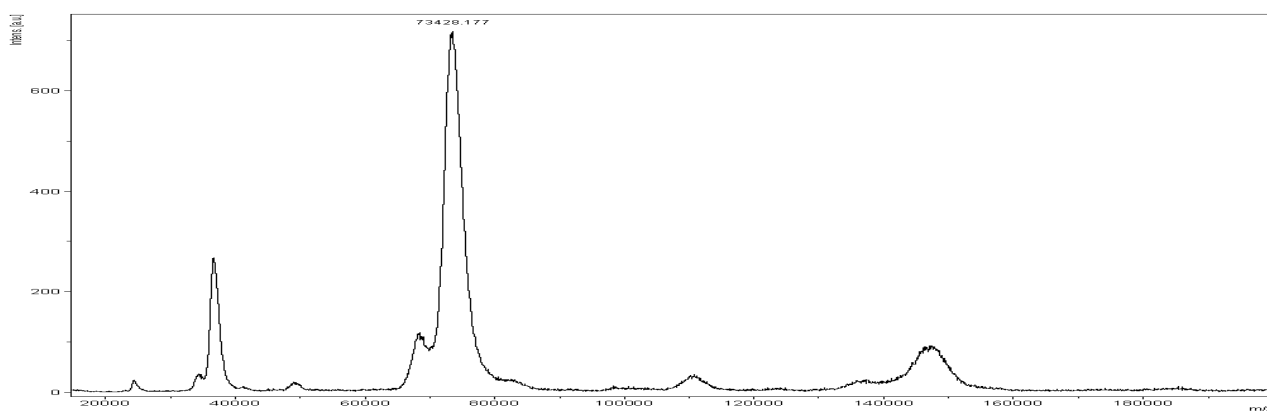
Oligosaccharide modification:

The final compound was dissolved in DI water (120 μL), then 3,3'-dithiobis(sulfosuccinimidylpropionate) (DTSSP) (4.0 mg, 6.6×10^{-3} mmol) was added. The pH of the reaction mixture was adjusted to mild basic by monitoring the pH and adding NaOH solution (0.1 M). The reaction mixture was stirred for 4 h, and lyophilized. The residue was dissolved in DI water (100 μL). Dithiothreitol (DTT) (6 mg, 3.9×10^{-2} mmol) was added, the reaction mixture was stirred at 37 $^{\circ}\text{C}$ for 4 h and monitored by MALDI. The crude reaction mixture was loaded on a P-2 column directly using DI water as eluent to afford the thiol-derivatized compound.



Scheme 25. Conjugation of the octasaccharide to BSA

MALDI of standard BSA before conjugation



MALDI of glyco-BSA conjugate after conjugation(5copies/BSA)

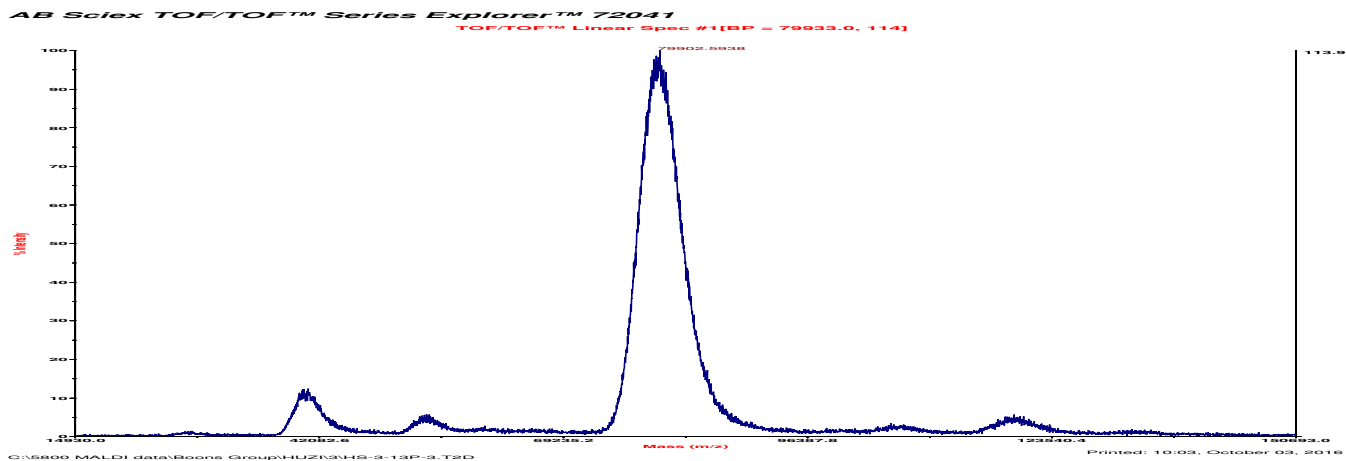
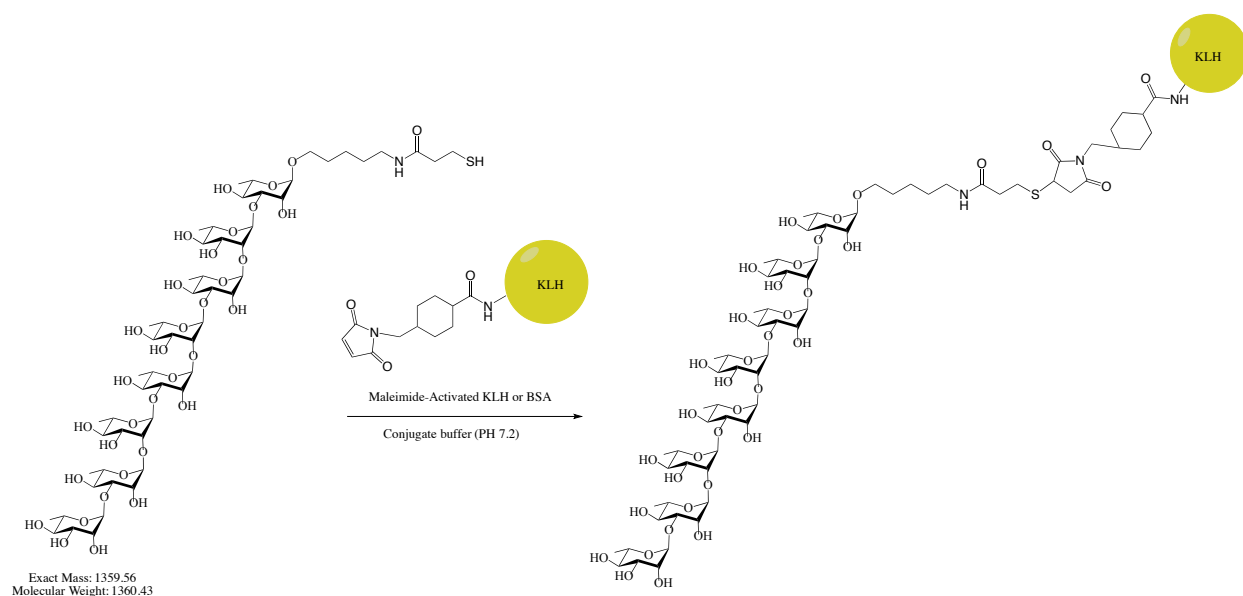


Figure 10. MALDI of the BSA glycoconjugate



Scheme 26. Conjugation of the octasaccharide to KLH

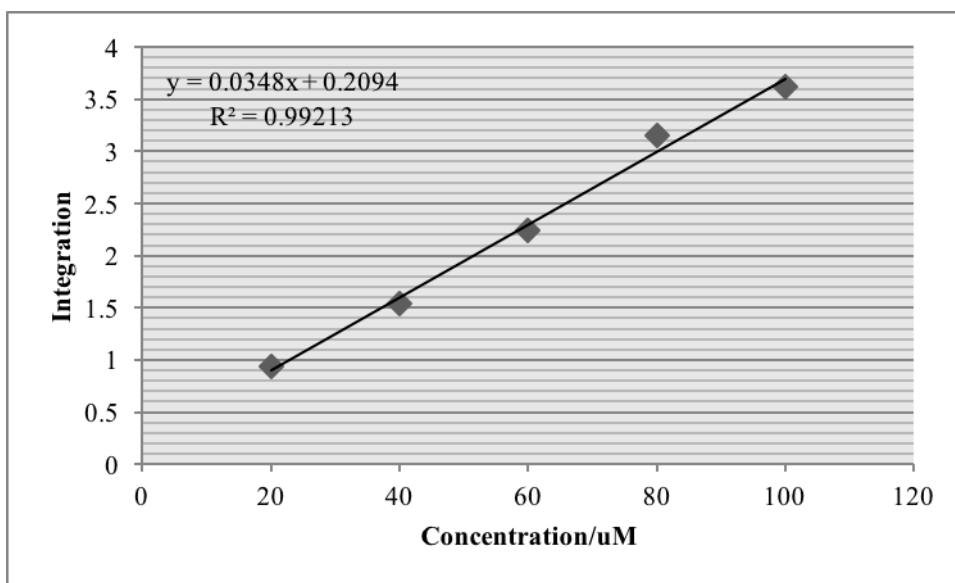
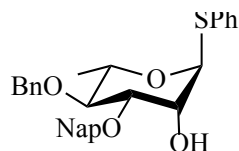


Figure 11. Standard curve of the HPAEC analysis

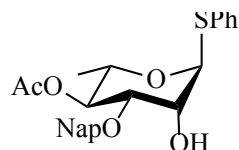
The calculated sugar concentration for KLH-glycoconjugate is 309.2 μM . The final glycoconjugate sugar containing concentration is 53 $\mu\text{g/mL}$. Using the average molecular weight of the KLH 8×10^6 Da for calculation, the result is approximately 215 copies/KLH.

Experimental procedure



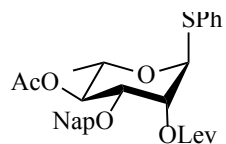
S12 α

Thiophenyl 3-O-naphthylmethyl-4-O-benzyl- α -L-rhamnopyranoside (S12 α). A suspension of **S8 α** (1.92 g, 5.54 mmol) and dibutyltin oxide (1.51 g, 6.10 mmol) in dry toluene (60 mL) was refluxed for 3 h under argon, then cool it to r.t. TBAB (1.96 g, 6.10 mmol), NapBr (1.35 g, 6.10 mmol) were added and the reaction was refluxed overnight. The solution then diluted with ethyl acetate, washed by water, brine, dried with anhydrous MgSO₄, and concentrated under reduced pressure. The resulting crude product was purified by flash chromatography over silica gel (hexane/EtOAc, 3:1 to 1:1, v/v) to give **S12 α** (2.02 g, 75%) as a syrup. ¹H NMR (300 MHz, chloroform-*d*) δ 7.90 – 7.76 (m, 5H, arom. H, Bn), 7.56 – 7.40 (m, 5H, arom. H, SPh), 7.38 – 7.21 (m, 7H, arom. H, Nap), 5.55 (d, J = 1.5 Hz, 1H, H-1), 4.94 (d, J = 11.0 Hz, 1H, CH, Nap), 4.86 (m, 2H, CH₂, Bn), 4.69 (d, J = 10.9 Hz, 1H, CH, Nap), 4.31 (dd, J = 3.3, 1.6 Hz, 1H, H-2), 4.27 – 4.19 (m, 1H, H-5), 3.95 (dd, J = 9.1, 3.2 Hz, 1H, H-3), 3.59 (m, 1H, H-4), 1.35 (d, J = 6.2 Hz, 3H, Rha, CH₃). ¹³C NMR (75 MHz, chloroform-*d*) (α/β) δ 138.30, 135.07, 134.12, 133.12, 133.10, 131.42, 129.03, 128.44, 128.29, 127.98, 127.96, 127.89, 127.79, 127.74, 127.71, 127.34, 126.87, 126.85, 126.29, 126.15, 125.86, 125.84, 125.81, 125.17, 87.10, 80.17, 80.11, 75.47, 72.27, 70.17, 68.91, 65.35, 60.43, 26.93, 26.65, 26.27, 21.05, 17.83, 14.21, 13.51. HR-MALDI-TOF/MS (*m/z*): calcd for C₃₀H₃₀NaO₄S⁺ [M+Na]⁺: 509.1912; found, 509.1913.



S10α

Thiophenyl 3-O-naphthylmethyl-4-O-acetyl-α-L-rhamnopyranoside (S10α). A suspension of **S7α** (4.44 g, 14.88 mmol) and dibutyltin oxide (4.04 g, 16.37 mmol) in dry toluene (100 mL) was refluxed for 3 h under argon, then cool it to r.t. TBAI (5.26 g, 16.37 mmol), NapBr (3.62 g, 16.37 mmol) were added and the reaction was refluxed overnight. The solution then diluted with ethyl acetate, washed by water, brine, dried with anhydrous MgSO₄, and concentrated under reduced pressure. The resulting crude product was purified by flash chromatography over silica gel (hexane/EtOAc, 3:1 to 1:1, v/v) to give **S10α** (4.37 g, 67%) as a syrup. ¹H NMR (300 MHz, chloroform-*d*) δ 7.96 – 7.67 (m, 2H, arom. H, Nap), 7.57 – 7.38 (m, 7H, arom. H, Nap, SPh), 7.39 – 7.15 (m, 3H, arom. H, SPh), 5.56 (d, 1H, H-1), 5.14 (m, 1H, H-4), 4.95 – 4.66 (m, 2H, CH₂, Nap), 4.29 (m, 1H, H-2), 4.24 – 4.11 (m, 1H, H-5), 3.82 (dd, *J* = 9.5, 3.3 Hz, 1H, H-3), 2.03 (s, 3H, CH₃, Ac), 1.19 (d, *J* = 6.3 Hz, 3H, Rha, CH₃). ¹³C NMR (75 MHz, chloroform-*d*) δ 170.05, 134.73, 133.19, 133.11, 131.30, 129.07, 128.46, 127.92, 127.73, 127.45, 126.78, 126.36, 126.21, 125.59, 86.76, 77.19, 76.92, 72.46, 71.92, 69.74, 67.58, 21.00, 17.28. HR-MALDI-TOF/MS (*m/z*): calcd for C₂₅H₂₆NaO₅S⁺ [*M*+Na]⁺: 461.1513; found, 461.1514.

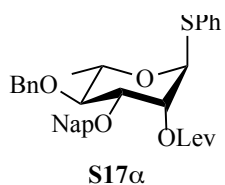


S15α

Thiophenyl 2-O-levulinoyl-3-O-naphthylmethyl-4-O-acetyl-α-L-rhamnopyranoside (S15α).

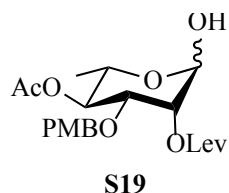
To a solution of **S10α** (4.23 g, 9.67 mmol) in CH₂Cl₂ (60 mL) were added LevOH (1.97 mL,

19.3 mmol), EDC (5.56 g, 29.01 mmol), and DMAP (catalytic amount, adjusted to pH ~8). After stirring overnight at r.t., the reaction mixture was added to a solution of NaHCO₃ (0.1 M) to quench the reaction. Next, the mixture was extracted with DCM (3 × 50 mL), and the combined organic extracts were concentrated *in vacuo*. The resulting crude product was purified by flash chromatography over silica gel (hexane:EtOAc, 4:1 to 2:1, v/v) to give **S15α** (4.00 g, 95%) as a syrup. ¹H NMR (300 MHz, chloroform-*d*) δ 7.84 – 7.61 (m, 3H, arom. H, Nap), 7.48 – 7.10 (m, 9H, arom. H, SPh, Nap), 5.56 (dd, *J* = 3.3, 1.7 Hz, 1H, H-2), 5.38 (d, *J* = 1.6 Hz, 1H, H-1), 5.02 (m, 1H, H-4), 4.76 – 4.46 (m, 2H, CH₂, Nap), 4.15 (dq, *J* = 9.8, 6.2 Hz, 1H, H-5), 3.76 (dd, *J* = 9.7, 3.3 Hz, 1H, H-3), 2.83 – 2.42 (m, 4H, CH₂-CH₂, Lev), 2.06 (s, 3H, CH₃, Ac), 1.97 (s, 3H, CH₃, Lev), 1.13 (d, *J* = 6.2 Hz, 3H, Rha, CH₃). ¹³C NMR (75 MHz, chloroform-*d*) δ 206.25, 171.86, 169.95, 135.00, 133.42, 133.20, 133.03, 131.67, 129.13, 128.16, 128.08, 127.91, 127.76, 127.70, 126.74, 126.20, 126.01, 125.81, 85.87, 74.74, 72.46, 71.23, 70.15, 67.78, 37.99, 29.72, 28.19, 20.99, 17.39. HR-MALDI-TOF/MS (*m/z*): calcd for C₃₀H₃₂NaO₇S⁺ [M+Na]⁺: 559.1912; found, 559.1913.



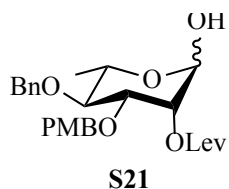
Thiophenyl 2-O-levulinoyl-3-O-naphthylmethyl-4-O-benzyl-α-L-rhamnopyranoside (S17α). To a solution of **S12α** (1.62 g, 3.33 mmol) in CH₂Cl₂ (50 mL) were added LevOH (0.68 ml, 6.66 mmol), EDC (1.92 g, 10.00 mmol) and DMAP (catalytic amount, adjusted to pH ~8). After stirring overnight at r.t., the reaction mixture was added to a solution of NaHCO₃ (0.1 M) to quench the reaction. Next, the mixture was extracted with DCM (3 × 50 mL), and the combined organic extracts were concentrated *in vacuo*. The resulting crude product was purified

by flash chromatography over silica gel (hexane:EtOAc, 4:1 to 2:1, v/v) to give **S17 α** (1.60 g, 82%) as a syrup. ^1H NMR (300 MHz, chloroform-*d*) δ 7.80 – 7.64 (m, 5H, arom. H, Bn), 7.37 (m, 5H, arom. H, SPh), 7.30 – 7.14 (m, 7H, arom. H, Nap), 5.58 (dd, J = 3.3, 1.7 Hz, 1H, H-2), 5.35 (d, J = 1.6 Hz, 1H, H-1), 4.84 (m, 2H, CH, Nap, CH, Bn), 4.61 (m, 2H, CH, Nap, CH, Bn), 4.15 (dq, J = 9.5, 6.2 Hz, 1H, H-5), 3.89 (dd, J = 9.3, 3.2 Hz, 1H, H-3), 3.43 (m, 1H, H-4), 2.62 (m, 4H, CH₂-CH₂, Lev), 2.06 (s, 3H, CH₃, Lev), 1.27 (d, J = 6.2 Hz, 3H, Rha, CH₃). ^{13}C NMR (75 MHz, chloroform-*d*) δ 171.95, 138.33, 135.18, 133.79, 133.27, 133.02, 131.73, 129.05, 128.38, 128.14, 127.94, 127.72, 127.66, 127.61, 126.96, 126.13, 126.03, 125.91, 85.96, 80.13, 78.24, 75.48, 71.74, 70.77, 69.04, 37.98, 29.78, 29.70, 28.12, 17.87. HR-MALDI-TOF/MS (*m/z*): calcd for C₃₅H₃₆NaO₆S⁺ [M+Na]⁺: 607.2215; found, 607.2214.

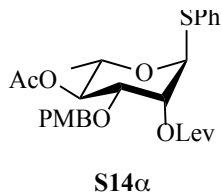


2-O-levulinoyl-3-paramethoxybenzyl-4-O-acetyl-L-rhamnopyranoside (S19). Compound **S14 α** (3.25 g, 6.29 mmol) was dissolved in acetone (50 mL), and NBS (1.12 g, 6.29 mmol) was added at 0 °C. The solution was stirred for 5 h performed in the dark. The reaction was monitored by TLC. Upon completion of the reaction, the reaction mixture was added to an aqueous solution of NaHCO₃ (0.1 M) to quench the reaction. The reaction mixture was extracted by DCM (3 \times 50 mL). The organic extracts were dried with anhydrous MgSO₄, and concentrated under reduced pressure. The crude product was purified by silica gel chromatography (hexane:EtOAc, 2:1 to 1:1, v/v) to provide pure compound **S19** (2.63 g, 72%). ^1H NMR (300 MHz, chloroform-*d*) δ 7.18 – 7.05 (m, 2H, arom. H, PMB), 6.86 – 6.71 (m, 2H, arom. H, PMB), 5.27 (m, 1H, H-2), 5.09 (m, 1H, H-1), 4.88 (m, 1H, H-4), 4.47 (m, 1H, CH, PMB), 4.25 (m, 1H,

CH, PMB), 3.93 (m, 1H, H-5), 3.81 (m, 1H, H-3), 3.77 – 3.69 (s, 3H, CH₃, PMB), 2.67 (m, 4H, CH₂-CH₂, Lev), 2.10 (s, 3H, CH₃, Ac), 1.95 (s, 3H, CH₃, Lev), 1.12 (m, 3H, Rha, CH₃). ¹³C NMR (75 MHz, chloroform-*d*) δ 172.90, 170.87, 160.06, 133.63, 130.85, 130.32, 130.21, 128.91, 126.84, 114.63, 114.54, 112.43, 93.18, 76.73, 74.80, 74.43, 73.34, 73.18, 72.90, 71.65, 71.33, 70.71, 69.87, 69.61, 67.36, 57.12, 56.12, 38.93, 30.62, 30.42, 29.30, 29.08, 21.83, 18.42. HR-MALDI-TOF/MS (m/z): calcd for C₂₁H₂₈NaO₉⁺ [M+Na]⁺: 447.1713; found, 447.1712.

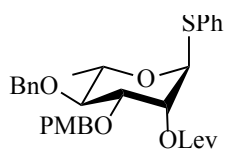


2-O-levulinoyl-3-paramethoxybenzyl-4-O-benzyl-L-rhamnopyranoside (S21). Compound **S16α** (3.10 g, 5.49 mmol) was dissolved in acetone (50 mL), and NBS (0.98 g, 5.49 mmol) was added at 0 °C. The solution was stirred for 5 h performed in the dark. The reaction was monitored by TLC. Upon completion of the reaction, the reaction mixture was added to an aqueous solution of NaHCO₃ (0.1 M) to quench the reaction. The reaction mixture was extracted by DCM (3 × 50 mL). The organic extracts were dried with anhydrous MgSO₄, and concentrated under reduced pressure. The crude product was purified by silica gel chromatography (hexane:EtOAc, 2:1 to 1:1, v/v) to provide pure compound **S21** (1.84 g, 71%). The spectrums match with reported literature.¹¹ HR-MALDI-TOF/MS (m/z): calcd for C₂₆H₃₂NaO₈⁺ [M+Na]⁺: 495.2114; found, 495.2113.



Thiophenyl 2-O-levulinoyl-3-O-paramethoxybenzyl-4-O-acetyl- α -L-rhamnopyranoside

(S14 α). To a solution of **S9 α** (4.67 g, 11.16 mmol) in CH₂Cl₂ (60 mL) was added LevOH (2.28 ml, 22.32 mmol), EDC (6.42 g, 33.48 mmol) and DMAP (catalytic amount, adjusted to pH ~8). After stirring overnight at r.t., the reaction mixture was added to a solution of NaHCO₃ (0.1 M) to quench the reaction. Next, the mixture was extracted with DCM (3 \times 50 mL), and the combined organic extracts were concentrated *in vacuo*. The resulting crude product was purified by flash chromatography over silica gel (hexane:EtOAc, 4:1 to 2:1, v/v) to give **S14 α** (5.65 g, 98%) as a syrup. ¹H NMR (300 MHz, chloroform-*d*) δ 7.46 – 7.40 (m, 2H, arom. H, SPh), 7.35 – 7.17 (m, 5H, arom. H, PMB, SPh), 6.92 – 6.85 (m, 2H, arom. H, PMB), 5.55 (dd, *J* = 3.3, 1.8 Hz, 1H, H-2), 5.43 (d, *J* = 1.7 Hz, 1H, H-1), 5.03 (m, 1H, H-4), 4.56 (d, *J* = 11.8 Hz, 1H, CH, PMB), 4.35 (d, *J* = 11.8 Hz, 1H, CH, PMB), 4.22 (m, 1H, H-5), 4.12 (m, 1H, H-3), 3.81 (s, 3H, OMe, PMB), 3.74 (dd, *J* = 9.7, 3.3 Hz, 1H, H-3), 2.84 – 2.60 (m, 4H, CH₂-CH₂, Lev), 2.16 (s, 3H, CH₃, Lev), 2.04 (s, 3H, CH₃, Ac), 1.20 (d, *J* = 6.2 Hz, 3H, Rha, CH₃). The spectrums match with reported literature.¹¹ HR-MALDI-TOF/MS (*m/z*): calcd for C₂₇H₃₂NaO₈S⁺ [M+Na]⁺: 539.1811; found, 539.1812.

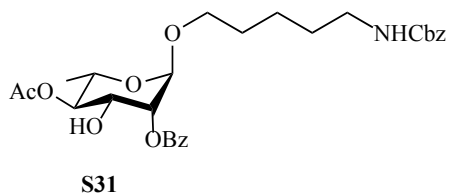


S16 α

Thiophenyl 2-O-levulinoyl-3-O-paramethoxybenzyl-4-O-benzyl- α -L-rhamnopyranoside

(S16 α). To a solution of **S10 α** (2.06 g, 4.42 mmol) in CH₂Cl₂ (50 mL) was added LevOH (0.91 ml, 8.80 mmol), EDC (2.54g, 13.25 mmol) and DMAP (catalytic amount, adjusted to pH ~8). After stirring overnight at r.t., the reaction mixture was added to a solution of NaHCO₃ (0.1 M) to quench the reaction. Next, the mixture was extracted with DCM (3 \times 50 mL), and the

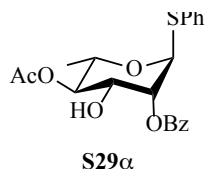
combined organic extracts were concentrated *in vacuo*. The resulting crude product was purified by flash chromatography over silica gel (hexane:EtOAc, 4:1 to 2:1, v/v) to give **S16 α** (1.80 g, 72%) as a syrup. ^1H NMR (300 MHz, chloroform-*d*) δ 7.36 (m, 2H, arom. H, Bn), 7.34 – 7.13 (m, 10H, arom. H, PMB, SPh, Bn), 6.78 (m, 2H, arom. H, PMB), 5.50 (dd, 1H, H-2), 5.32 (d, 1H, H-1), 4.84 (d, J = 10.9 Hz, 1H, CH, Bn), 4.60 – 4.49 (m, 2H, CH, Bn, CH, PMB), 4.39 (d, J = 10.9 Hz, 1H, CH, PMB), 4.13 (dq, J = 11.9, 6.4 Hz, 1H, H-5), 3.80 (dd, J = 9.2, 3.4 Hz, 1H, H-3), 3.38 (m, 1H, H-4), 2.76 – 2.42 (m, 4H, CH₂-CH₂, Lev), 2.10 (d, J = 4.6 Hz, 3H, CH₃, Lev), 1.25 (d, J = 6.2 Hz, 3H, Rha, CH₃). ^{13}C NMR (75 MHz, chloroform-*d*) δ 206.74, 172.49, 159.88, 138.93, 132.28, 130.59, 130.39, 129.60, 128.91, 128.52, 128.26, 128.14, 114.49, 114.35, 86.52, 80.62, 78.38, 75.96, 71.87, 71.35, 69.56, 55.80, 53.96, 38.56, 38.47, 30.37, 30.25, 28.68, 18.41. HR-MALDI-TOF/MS (m/z): calcd for C₃₂H₃₆NaO₇S⁺ [M+Na]⁺: 587.2213; found, 587.2212.



***N*-benzyloxycarbonyl-5-aminopentyl O-2-benzoyl-4-acetyl- α -L-rhamonopyranoside (S31).**

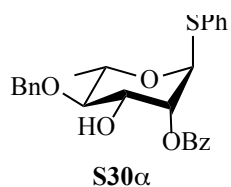
A mixture of compounds **S29 α** (5.00 g, 12.42 mmol) and *N*-benzyloxycarbonyl-5-aminopentanol (5.89 g, 24.84 mmol) was co-evaporated with toluene (3 \times 3 mL), and dried under high vacuo overnight. Next, the compounds were dissolved in dry DCM (120 mL). The solution was cooled to -50 °C, and freshly activated acid washed 4Å MS (1.00 g) was added. The solution was argon-purged 3 times, and stirred at -50 °C for 20 min. NIS (4.19 g, 18.64 mmol) and AgOTf (1.59 g, 6.21 mmol) were added at -50 °C, and the temperature was kept at -50 °C for 20 min. Then, the temperature was raised to 0 °C, and the reaction was monitored by TLC. The reaction was quenched by adding a saturated NaS₂O₃ solution, and the organic phase was washed by NaHCO₃

(0.1 M), and the aqueous phase was extracted with DCM (3×50 mL). The combined organic extracts were dried with anhydrous MgSO_4 , filtered and the filtrate was concentrated under reduced pressure. The resulting crude product was purified by flash chromatography over silica gel (hexane:EtOAc, 4:1 to 2:1, v/v) to give **S31** (6.00 g, 91.2%) as a syrup. ^1H NMR (300 MHz, chloroform-*d*) δ 8.05 (m, 2H, arom. H, Cbz), 7.62 – 7.15 (m, 8H, arom. H, Cbz, Bz), 5.46 – 5.17 (m, 2H, H-1, H-2), 5.13 – 4.74 (m, 5H, H-4, CH_2 , Cbz), 4.29 – 4.04 (m, 1H, H-3), 3.98 – 3.75 (m, 1H, H-5), 3.67 (m, 1H, CH, Linker), 3.47 – 3.28 (m, 1H, CH, Linker), 3.17 (m, 2H, CH_2 , Linker), 2.22 – 1.74 (m, 3H, CH_3 , Ac), 1.27 – 1.21 (m, 3H, Rha, CH_3). ^{13}C NMR (75 MHz, chloroform-*d*) δ 171.53, 166.06, 133.50, 129.87, 129.81, 129.77, 129.34, 128.77, 128.53, 128.08, 99.39, 97.19, 77.19, 74.96, 73.35, 72.74, 72.48, 71.00, 69.58, 68.87, 68.14, 67.89, 66.62, 66.03, 40.94, 38.72, 29.69, 29.01, 28.92, 28.82, 23.35, 21.03, 17.54, 17.45. HR-MALDI-TOF/MS (m/z): calcd for $\text{C}_{28}\text{H}_{35}\text{NNaO}_9^+$ $[\text{M}+\text{Na}]^+$: 552.2302; found, 552.2303.

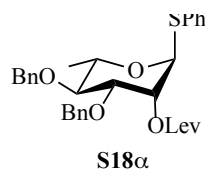


Thiophenyl 2-O-benzoyl-4-O-acetyl- α -L-rhamnopyranoside (S29 α). Compound **S7 α** (10.79 g, 36.1 mmol) was dissolved in DCM (250 mL). Then, trimethyl orthobenzoate (7.45 mL, 43.4 mmol) and a catalytic amount of CSA were added at r.t. under argon. The reaction was stirred at r.t. for 2 h. The solution was concentrated under reduced pressure, and the resulting product was dissolved in 90% HOAc at r.t. The reaction was stirred at r.t. overnight. The solution was washed by NaHCO_3 (0.1 M) to neutral, and the aqueous phase was extracted with DCM (3×50 mL). The combined organic extracts were dried with anhydrous MgSO_4 , filtered and the filtrate was concentrated under reduced pressure. The resulting crude product was purified by flash

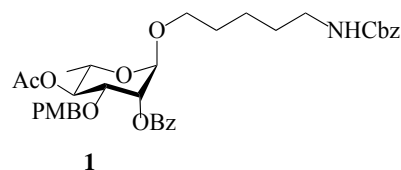
chromatography over silica gel (hexane:EtOAc, 5:1 to 2:1, v/v) to give **S29 α** (11.64 g, 80%) as a syrup. ^1H NMR (300 MHz, chloroform-*d*) δ 8.19 – 7.90 (m, 2H, arom. H, Bz), 7.66 – 7.17 (m, 8H, arom. H, Bz, SPh), 5.61 (m, 2H, H-1, H-2), 5.10 (m, 1H, H-4), 4.38 (dq, J = 9.7, 6.2 Hz, 1H, H-5), 4.23 – 4.03 (m, 1H, H-3), 2.10 (m, 3H, CH₃, Ac), 1.28 (d, J = 6.3 Hz, 3H, Rha, CH₃). ^{13}C NMR (75 MHz, chloroform-*d*) δ 133.58, 131.75, 131.44, 129.86, 129.14, 128.56, 127.79, 87.47, 85.76, 75.08, 74.77, 70.90, 69.58, 67.38, 21.03, 17.44. HR-MALDI-TOF/MS (m/z): calcd for C₂₁H₂₂NaO₆S⁺ [M+Na]⁺: 425.1111; found, 425.1112.



Thiophenyl 2-O-benzoyl-4-O-benzyl- α -L-rhamnopyranoside (S30 α). Compound **S8 α** (2.07 g, 5.98 mmol) was dissolved in DCM (70 mL). Then, trimethyl orthobenzoate (1.23 mL, 7.2 mmol) and a catalytic amount of CSA were added at r.t. under argon. The reaction was stirred at r.t. for 2 h. The solution was concentrated under reduced pressure, and the resulting product was dissolved in 90% HOAc at r.t. The reaction was stirred at r.t. overnight. The solution was washed by NaHCO₃ (0.1 M) to neutral, and the aqueous phase was extracted with DCM (3 \times 50 mL). The combined organic extracts were dried with anhydrous MgSO₄, filtered and the filtrate was concentrated under reduced pressure. The resulting crude product was purified by flash chromatography over silica gel (hexane:EtOAc, 7:1 to 5:1, v/v) to give **S30 α** (1.48 g, 56%) as a syrup. The spectrums match with reported literature.¹³ HR-MALDI-TOF/MS (m/z): calcd for C₂₆H₂₆NaO₅S⁺ [M+Na]⁺: 473.1512; found, 473.1513.

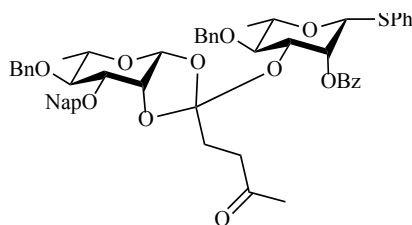


Thiophenyl 2-O-levulinoyl-3-O-benzyl-4-O-benzyl-α-L-rhamnopyranoside (S18α). To a solution of **S13α** (2.14 g, 4.91 mmol) in CH₂Cl₂ (30 mL) were added LevOH (1.01 mL, 9.82 mmol), EDC (2.82 g, 14.73 mmol) and DMAP (catalytic amount, adjusted to pH ~8). After stirring overnight at r.t., the reaction mixture was added to a solution of NaHCO₃ (0.1 M) to quench the reaction. Next, the mixture was extracted with DCM (3 × 50 mL), and the combined organic extracts were concentrated *in vacuo*. The resulting crude product was purified by flash chromatography over silica gel (hexane:EtOAc, 4:1 to 2:1, v/v) to give **S18α** (2.15 g, 82%) as syrup. The spectrums match with reported literature.¹³ HR-MALDI-TOF/MS (m/z): calcd for C₃₁H₃₄NaO₆S⁺ [M+Na]⁺: 557.2102; found, 557.2103.



N-benzyloxycarbonyl-5-aminopentyl O-2-benzoyl-3-paramethoxybenzyl-4-acetyl-α-L-rhamonopyranoside (1). A mixture of compound **S16α** (69.5 mg, 0.1346 mmol) and compound **S31** (59.4 mg, 0.1122 mmol) was co-evaporated with toluene (3 × 3 mL), and dried under high vacuo overnight. Next, the compounds were dissolved in dry DCM (4 mL). The solution was cooled to -50 °C, and freshly activated acid washed 4Å molecular sieves (121.1 mg) was added. The solution was argon-purged 3 times, and stirred at -50 °C for 20 min. NIS (37.9 mg, 0.1683 mmol) and AgOTf (14.4 mg, 0.0561 mmol) were added at -50 °C, after which the temperature was raised to 0 °C. The reaction was quenched by adding a saturated aqueous NaS₂O₃ solution, and

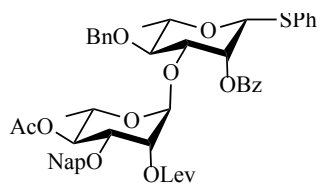
the organic phase was washed with NaHCO₃ (0.1 M). The aqueous phase was extracted with DCM (3 × 50 mL). The combined organic extracts were dried with anhydrous MgSO₄, filtered and the filtrate was concentrated under reduced pressure. The resulting crude product was purified by flash chromatography over silica gel (toluene/acetone, 4:1, v/v) to give **1** (46 mg, 63%) as a syrup. ¹H NMR (300 MHz, chloroform-*d*) δ 8.12 – 8.04 (m, 2H, arom. H, Bz), 7.69 – 6.97 (m, 10H, arom. H, Bz, Cbz, PMB), 6.84 (m, 2H, arom. H, Cbz), 5.32 – 5.27 (m, 1H, H-2), 4.99 (m, 1H, H-4), 4.87 (d, 1H, H-1), 4.44 (m, 2H, CH₂, PMB), 4.18 – 4.08 (m, 1H, H-3), 3.83 (m, 1H, CH, Linker), 3.79 (s, 3H, OMe, PMB), 3.59 (m, 2H, H-5, CH, Linker), 3.22 (m, 3H, CH, Linker), 1.25 (d, *J* = 6.2 Hz, 3H, Rha, CH₃). ¹³C NMR (75 MHz, chloroform-*d*) δ 171.50, 158.90, 133.48, 129.86, 129.76, 129.71, 128.53, 128.44, 127.91, 127.84, 113.90, 113.64, 97.18, 77.19, 74.96, 73.34, 68.86, 67.88, 67.12, 66.01, 55.26, 53.40, 52.65, 30.91, 29.68, 29.08, 23.38, 21.02, 17.54, 7.67. HR-MALDI-TOF/MS (*m/z*): calcd for C₃₆H₄₃NNaO₁₀⁺ [*M*+Na]⁺: 672.2913; found, 672.2912.



2

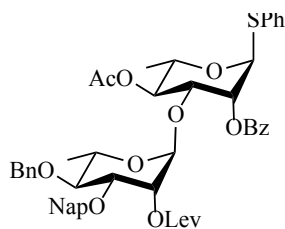
4-O-benzyl-3-O-naphthylmethyl-1,2-(thiophenyl 4-O-benzy-2-O-benzoyl-3-yl-β-L-rhamnopyranosid-2-yloxy-1-levulinoylidene)-β-L-rhamnopyranose (2). A mixture of compounds **S27** (107 mg, 0.1613 mmol) and **S30β** (48 mg, 0.1075 mmol) was co-evaporated with toluene (3 × 3 mL) and dried under vacuo overnight. Next, the mixture was dissolved in anhydrous DCM (2 mL) and placed under argon atmosphere. The freshly activated powdered 4Å

acid washed MS (201.3 mg) was added at $-78\text{ }^{\circ}\text{C}$, after which the solution was argon-purged 3 times and stirred under argon for 20 min. TBSOTf (diluted 1:20 in DCM, 25 μL , 0.005441 mmol) was added at $-78\text{ }^{\circ}\text{C}$ and the temperature was maintained at $-78\text{ }^{\circ}\text{C}$ for 30 min. The reaction mixture was quenched by the addition of Et_3N (3 μL) at $-78\text{ }^{\circ}\text{C}$. The mixture was diluted with DCM (50 mL) and filtered. The filtrate was concentrated under reduced pressure and the residue was purified by silica gel column chromatography (hexane:EtOAc, 3:1, v/v) to provide pure orthoester product **2** (70 mg, 70%) and pure product **3 β** (10 mg, 10%). ^1H NMR (500 MHz, chloroform-*d*) δ 8.11 (m, 2H, arom. H, Bz), 7.82 – 7.13 (m, 25H, arom. H, Bz, Bn, SPh, Nap), 5.89 (m, 1H, H-2,A), 5.27 (d, $J = 2.5\text{ Hz}$, 1H, H-1,B), 4.89 (d, $J = 2.4\text{ Hz}$, 1H, H-1,A), 4.87 (m, 1H, CH, Bn), 4.83 (m, 1H, CH, Bn), 4.64 (m, 1H, CH, Nap), 4.58 (m, 2H, CH, Nap, CH, Bn), 4.53 (m, 1H, H-2,B), 4.46 (m, 1H, CH, Bn), 3.98 (m, 1H, H-3,A), 3.67 (m, 1H, H-3,B), 3.49 (m, 2H, H-4,A, H-5,A), 3.35 (m, 1H, H-4,B), 3.26 (m, 1H, H-5,B), 2.72 – 2.59 (m, 1H, CH, Lev), 2.54 – 2.46 (m, 1H, CH, Lev), 2.32 (m, 1H, CH, Lev), 2.12 – 2.04 (m, 1H, CH, Lev), 1.90 (s, 3H, CH_3 , Lev), 1.23 (m, 3H, Rha, CH_3 , A), 1.20 – 1.15 (m, 3H, Rha, CH_3 , B). ^{13}C NMR (226 MHz, chloroform-*d*) δ 168.81, 140.51, 136.09, 135.81, 133.89, 132.75, 131.66, 131.24, 131.04, 131.03, 130.78, 130.66, 130.61, 130.52, 130.43, 130.40, 130.29, 130.18, 129.39, 128.72, 128.57, 128.53, 127.24, 99.89, 88.06, 81.69, 81.39, 80.88, 79.23, 78.21, 78.08, 78.06, 78.01, 77.28, 73.64, 72.86, 40.50, 32.40, 32.29, 21.14, 20.59. HR-MALDI-TOF/MS (m/z): calcd for $\text{C}_{55}\text{H}_{56}\text{NaO}_{11}\text{S}^+$ $[\text{M}+\text{Na}]^+$: 947.3517; found, 947.3511.



5 β

Thiophenyl O-(4-O-acetyl-3-O-naphthylmethyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- β -L-rhamnopyranoside. (5 β) A mixture of compound **S25** (41 mg, 0.0673 mmol) and compound **S30 β** (20 mg, 0.0449 mmol) was co-evaporated with toluene (3 \times 3 mL) and dried under vacuo overnight. Next, the mixture was dissolved in anhydrous DCM (2 mL) and placed under argon atmosphere. The freshly activated powdered 4Å acid washed MS (51.3 mg) was added at -78 °C, after which the solution was argon-purged 3 times and stirred under argon for 20 min. TBSOTf (diluted 1:20 in DCM, 6 μ L, 0.00130584 mmol) was added at -78 °C and the temperature was maintained at -78 °C for 30 min. The reaction mixture was quenched by the addition of Et₃N (3 μ L) at -78 °C. The mixture was diluted with DCM (50 mL) and filtered. The filtrate was concentrated under reduced pressure and the residue was purified by silica gel column chromatography (hexane:EtOAc, 3:1, v:v) to provide the pure product **5 β** (8 mg, 20%). ¹H NMR (600 MHz, chloroform-*d*) δ 8.09 – 7.60 (m, 4H, arom. H, Bz, Nap), 7.60 – 7.00 (m, 18H, arom. H, Bn, Nap, SPh), 5.60 (m, 1H, H-2,A), 5.57 (d, *J* = 1.7 Hz, 1H, H-1,A), 5.28 (m, 1H, H-4,A), 5.21 (m, 1H, H-2,B), 4.93 (d, *J* = 1.7 Hz, 1H, H-1,B), 4.71 (m, 2H, CH, Bn, CH, Nap), 4.56 – 4.41 (m, 2H, CH, Bn, CH, Nap), 4.33 (dq, *J* = 9.9, 6.3 Hz, 1H, H-5,A), 4.17 (dd, *J* = 9.8, 3.3 Hz, 1H, H-3,A), 3.87 – 3.59 (m, 2H, H-3,B, H-5,B), 3.35 (m, 1H, H-4,B), 2.88 – 2.54 (m, 4H, CH₂-CH₂, Lev), 2.12 (m, 6H, CH₃, Lev, CH₃, Ac), 1.31 – 1.17 (m, 6H, Rha, CH₃,A,B). ¹³C NMR (151 MHz, chloroform-*d*) δ 172.00, 133.38, 132.90, 131.78, 129.85, 129.13, 128.49, 128.15, 127.95, 127.92, 127.79, 127.58, 127.46, 127.35, 126.63, 126.01, 125.91, 125.76, 99.48, 85.66, 79.45, 77.35, 75.53, 74.40, 73.82, 72.83, 71.66, 69.28, 68.62, 67.93, 37.96, 30.16, 29.69, 28.10, 20.80, 17.76, 17.47. HR-MALDI-TOF/MS (*m/z*): calcd for C₅₀H₅₂NaO₁₂S⁺ [*M*+Na]⁺: 899.3219; found, 899.3216



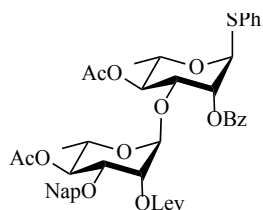
6

Thiophenyl O-(4-O-benzyl-3-O-naphthylmethyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-

(1 \rightarrow 3)-4-O-acetyl-2-O-benzoyl- α -L-rhamnopyranoside. (6)

A mixture of compound **S27** (55 mg, 0.0827 mmol) and compound **S29 α** (22 mg, 0.0552 mmol) was co-evaporated with toluene (3 \times 3 mL) and dried under vacuo overnight. Next, the mixture was dissolved in anhydrous DCM (2 mL) and placed under argon atmosphere. The freshly activated powdered 4Å acid washed molecular sieves (51.3 mg) was added at -78 °C, after which the solution was argon-purged 3 times and stirred under argon for 20 min. TBSOTf (diluted 1:20 in DCM, 12 μ L, 0.002611 mmol) was added at -78 °C and the temperature was maintained at -78 °C for 30 min. The reaction mixture was quenched by the addition of Et₃N (3 μ L) at -78 °C. The mixture was diluted with DCM (50 mL) and filtered. The filtrate was concentrated under reduced pressure and the residue was purified by silica gel column chromatography (hexane:EtOAc, 3:1, v:v) to provide the pure product **6** (7 mg, 14%). ¹H NMR (600 MHz, chloroform-*d*) δ 8.15 – 7.98 (m, 2H, arom. H, Bz), 7.84 – 7.13 (m, 20H, arom. H, Bn, Nap, Bz, SPh), 5.77 (dd, J = 3.5, 1.1 Hz, 1H, H-2,A), 5.22 (dd, J = 3.3, 1.7 Hz, 1H, H-2,B), 5.10 (d, J = 1.7 Hz, 1H, H-1,B), 5.01 (m, 1H, H-4,B), 4.96 (d, J = 1.1 Hz, 1H, H-1,A), 4.72 – 4.27 (m, 4H, CH₂, Bn, CH₂, Nap), 4.20 – 4.07 (m, 1H, H-5,B), 3.97 (dd, J = 9.3, 3.5 Hz, 1H, H-3,A), 3.72 (dd, J = 9.8, 3.3 Hz, 1H, H-3,B), 3.56 (m, 1H, H-4,A), 3.48 (dq, J = 9.4, 6.1 Hz, 1H, H-5,A), 2.70 – 2.53 (m, 4H, CH₂-CH₂, Lev), 2.09 – 2.01 (m, 6H, CH₃, Lev, CH₃, Ac), 1.26 – 1.19 (m, 6H, Rha, CH₃,A,B). ¹³C NMR (151 MHz, chloroform-

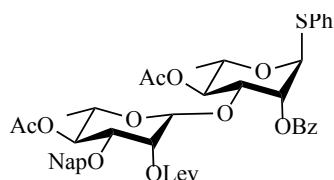
d) δ 171.68, 170.22, 165.81, 137.40, 135.50, 133.96, 133.35, 133.16, 132.81, 131.15, 129.95, 129.63, 129.01, 128.51, 128.40, 128.11, 127.91, 127.88, 127.80, 127.63, 127.59, 126.00, 125.82, 125.77, 125.40, 98.84, 85.60, 80.82, 76.85, 76.31, 75.49, 74.99, 73.88, 72.25, 71.71, 69.05, 67.53, 38.00, 29.68, 28.11, 22.68, 21.06, 18.31, 17.60. HR-MALDI-TOF/MS (m/z): calcd for $C_{50}H_{52}NaO_{12}S^+$ [M+Na] $^+$: 899.3219; found, 899.3217.



4 α - α

Thiophenyl O-(4-O-acetyl-3-O-naphthylmethyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-acetyl-2-O-benzoyl- α -L-rhamnopyranoside. (4 α - α) A mixture of compound **S25** (132 mg, 0.2143 mmol) and compound **S29 α** (52 mg, 0.1429 mmol) was co-evaporated with toluene (3 \times 3 mL) and dried under vacuo overnight. Next, the mixture was dissolved in anhydrous DCM (2 mL) and placed under argon atmosphere. The freshly activated powdered 4Å acid washed molecular sieves (51.3 mg) was added at -50 °C, after which the solution was argon-purged 3 times and stirred under argon for 20 min. TBSOTf (diluted 1:20 in DCM, 30 μ L, 0.0065292 mmol) was added at -50 °C and the temperature was maintained at -50 °C for 30 min. The reaction mixture was quenched by the addition of Et_3N (3 μ L) at -50 °C. The mixture was diluted with DCM (50 mL) and filtered. The filtrate was concentrated under reduced pressure and the residue was purified by silica gel column chromatography (toluene:EtOAc, 5:1, v:v) to provide the pure products **4 α - α** (52 mg, 49%) and **4 β - α** (7 mg, 7%). 1H NMR (500 MHz, chloroform-*d*) δ 8.02 (m, 2H, arom. H, Bz), 7.90 – 7.18 (m, 15H, arom. H, Nap, Bz, SPh), 5.67 – 5.54 (m, 2H, H-1,A, H-2,A), 5.31 (m, 1H, H-4,A), 5.02 – 4.90 (m, 2H, H-2,B, H-1,B), 4.64 (m,

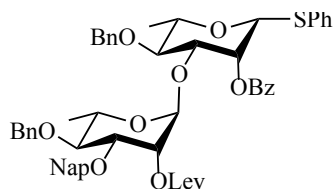
1H, H-4,B), 4.51 – 4.31 (m, 2H, CH, Nap, H-5,A), 4.22 (m, 1H, H-3,A), 3.85 – 3.63 (m, 2H, H-3,B, H-5,B), 2.89 – 2.47 (m, 4H, CH₂-CH₂, Lev), 2.15 (m, 6H, CH₃, Lev, CH₃, Ac), 1.88 (s, 3H, CH₃, Ac), 1.29 (m, 3H, Rha, CH₃,A), 1.10 (m, 3H, Rha, CH₃,B). ¹³C NMR (75 MHz, chloroform-*d*) δ 206.29, 171.87, 170.24, 169.98, 165.74, 135.26, 133.42, 133.27, 133.21, 132.89, 131.80, 129.77, 129.30, 129.16, 128.48, 127.92, 127.89, 127.85, 127.59, 126.30, 126.04, 125.82, 125.61, 99.36, 85.69, 75.50, 74.43, 73.83, 72.81, 71.98, 71.34, 68.79, 67.94, 67.42, 37.96, 29.69, 28.17, 20.85, 20.82, 17.45, 17.29. HR-MALDI-TOF/MS (m/z): calcd for C₄₅H₄₈NaO₁₃S⁺ [M+Na]⁺: 851.2821; found, 851.2827.



4β-α

Thiophenyl O-(4-O-acetyl-3-O-naphthylmethyl-2-O-levulinoyl-β-L-rhamnopyranosyl)-(1→3)-4-O-acetyl-2-O-benzoyl-α-L-rhamnopyranoside. (4β-α) A mixture of compound **S25** (132 mg, 0.2143 mmol) and compound **S29α** (52 mg, 0.1429 mmol) was co-evaporated with toluene (3 × 3 mL) and dried under vacuo overnight. Next, the mixture was dissolved in anhydrous DCM (2 mL) and placed under argon atmosphere. The freshly activated powdered 4Å acid washed molecular sieves (51.3 mg) was added at -50 °C, after which the solution was argon-purged 3 times and stirred under argon for 20 min. TBSOTf (diluted 1:20 in DCM, 30 μL, 0.0065292 mmol) was added at -50 °C and the temperature was maintained at -50 °C for 30 min. The reaction mixture was quenched by the addition of Et₃N (3 μL) at -50 °C. The mixture was diluted with DCM (50 mL) and filtered. The filtrate was concentrated under reduced pressure and the residue was purified by silica gel column chromatography (toluene:EtOAc, 5:1, v:v) to

provide the pure products **4 α - α** (52mg, 49%) and **4 β - α** (7mg, 7%). ^1H NMR (500 MHz, chloroform-*d*) δ 8.09 (m, 2H, arom. H, Bz), 7.92 – 7.21 (m, 15H, arom. H, Nap, Bz, SPh), 5.71 (d, J = 2.9 Hz, 1H, H-2,A), 5.12 (m, 1H, H-4,A), 4.78 (d, J = 12.5 Hz, 1H, CH, Nap), 4.63 (d, 1H, H-1,B), 4.50 (d, J = 12.5 Hz, 1H, CH, Nap), 4.39 (m, 1H, H-5,A), 4.15 (m, 1H, H-3,A), 3.57 – 3.48 (m, 1H, H-3,B), 3.37 (m, 1H, H-5,B), 2.76 – 2.39 (m, 4H, CH₂-CH₂, Lev), 2.13 (s, 3H, CH₃, Lev), 2.09 – 1.99 (m, 6H, 2 \times CH₃, Ac), 1.30 – 1.23 (m, 6H, Rha, CH₃,A,B). ^{13}C NMR (75 MHz, chloroform-*d*) δ 171.63, 131.74, 130.02, 129.20, 128.42, 127.90, 126.77, 126.10, 125.84, 110.00, 72.11, 71.00, 67.74, 67.27, 38.15, 30.91, 29.68, 27.94, 21.03. HR-MALDI-TOF/MS (*m/z*): calcd for C₄₅H₄₈NaO₁₃S⁺ [*M*+Na]⁺: 851.2821; found, 851.2823.



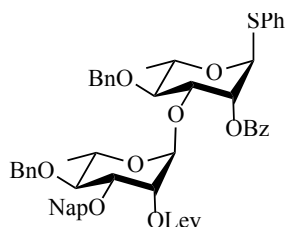
3 β

Thiophenyl O-(4-O-benzyl-3-O-naphthylmethyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-

(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- β -L-rhamnopyranoside. (3 β) Compound **3 β** was obtained by

the optimal glycosylation method as described in the General Procedures. Yield: 87%. ^1H NMR (500 MHz, chloroform-*d*) δ 8.26 – 8.05 (m, 2H, arom. H, Bz), 7.83 – 7.09 (m, 25H, arom. H, Bn, Nap, Bz, SPh), 5.87 (dd, J = 3.5, 1.1 Hz, 1H, H-2,A), 5.31 (m, 1H, H-2,B), 5.10 (d, J = 1.7 Hz, 1H, H-1,B), 5.00 (d, J = 1.1 Hz, 1H, H-1,A), 4.94 (d, J = 11.5 Hz, 1H, CH, Nap), 4.71 (m, 2H, CH₂, Bn), 4.60 – 4.53 (m, 2H, CH₂, Bn), 4.43 (d, J = 11.4 Hz, 1H, CH, Nap), 4.19 – 4.06 (m, 1H, H-5,B), 4.00 (dd, J = 9.3, 3.5 Hz, 1H, H-3,A), 3.85 (dd, J = 9.4, 3.3 Hz, 1H, H-3,B), 3.64 – 3.49 (m, 2H, H-4,A, H-5,A), 3.43 (m, 1H, H-4,B), 2.62 (m, 4H, CH₂-CH₂, Lev), 2.08 (m, 3H, CH₃, Lev), 1.47 (d, J = 6.1 Hz, 3H, Rha, CH₃,A), 1.37 – 1.25 (m, 3H, Rha, CH₃,B). ^{13}C NMR (75

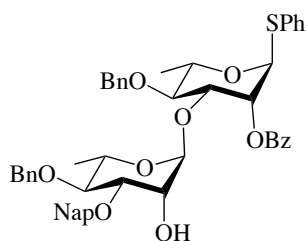
MHz, chloroform-*d*) δ 206.03, 171.80, 165.65, 138.86, 137.58, 135.57, 134.10, 133.23, 132.86, 131.22, 130.05, 129.73, 129.00, 128.52, 128.45, 128.39, 128.22, 128.06, 127.91, 127.87, 127.85, 127.64, 127.59, 127.54, 127.37, 126.42, 125.93, 125.88, 125.73, 98.97, 85.65, 80.76, 79.45, 77.72, 77.56, 76.30, 75.44, 74.26, 73.71, 71.83, 69.50, 68.76, 37.98, 29.75, 28.08, 18.35, 18.08. HR-MALDI-TOF/MS (*m/z*): calcd for C₅₅H₅₆NaO₁₁S⁺ [M+Na]⁺: 947.3513; found, 947.3517.



3 α

Thiophenyl O-(4-O-benzyl-3-O-naphthylmethyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (3 α). Compound **3 α** was obtained by the optimal glycosylation method as described in the General Procedures. Yield: 72%. ¹H NMR (500 MHz, chloroform-*d*) δ 8.07 – 8.02 (m, 2H, arom. H, Bz), 7.87 – 7.13 (m, 25H, arom. H, Bn, Bz, Nap, SPh), 5.66 (dd, *J* = 3.2, 1.8 Hz, 1H, H-2,A), 5.59 (d, *J* = 1.7 Hz, 1H, H-1,A), 5.46 (dd, *J* = 3.3, 1.9 Hz, 1H, H-2,B), 5.12 (d, *J* = 1.8 Hz, 1H, H-1,B), 4.84 - 4.47 (m, 6H, 2xCH₂, Bn, CH₂, Nap), 4.34 – 4.22 (m, 2H, H-3,A, H-5,A), 3.93 – 3.77 (m, 2H, H-3,B, H-5,B), 3.65 (m, 1H, H-4,A), 3.41 (m, 1H, H-4,B), 2.78 – 2.63 (m, 4H, CH₂-CH₂, Lev), 2.15 (s, 3H, CH₃, Lev), 1.36 (d, *J* = 6.2 Hz, 3H, Rha, CH₃, A), 1.21 (d, *J* = 6.2 Hz, 3H, Rha, CH₃, B). ¹³C NMR (75 MHz, chloroform-*d*) δ 138.57, 137.79, 135.47, 133.34, 131.86, 129.76, 129.09, 128.51, 128.28, 128.24, 128.11, 128.00, 127.94, 127.90, 127.67, 127.64, 127.60, 127.47, 126.58, 125.98, 125.96, 125.81, 99.45, 85.68, 80.42, 79.69, 77.90, 75.49, 74.73, 74.30, 71.62, 69.39, 69.20, 68.81, 38.03, 29.81,

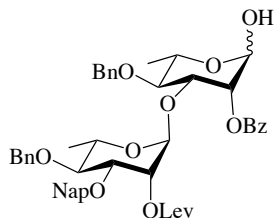
28.15, 17.99, 17.82. HR-MALDI-TOF/MS (m/z): calcd for $C_{55}H_{56}NaO_{11}S^+$ $[M+Na]^+$: 947.3513; found, 947.3518.



6α

Thiophenyl O-(4-O-benzyl-3-O-naphthylmethyl-α-L-rhamnopyranosyl)-(1→3)-4-O-benzyl-2-O-benzoyl-α-L-rhamnopyranoside (6α). Hydrazine acetate (390 mg, 4.232 mmol) was added to a solution of **3α** (783 mg, 0.846 mmol) in a mixture of DCM and MeOH (1/1, v/v, 40 mL). Stirring was continued until MALDI indicated the disappearance of the starting material (~3 h). The reaction mixture was diluted with DCM (30 mL), washed with water (3 × 25 mL) and brine (25 mL), dried with anhydrous $MgSO_4$, and filtered. The crude product was purified by silica gel chromatography (hexane:EtOAc, 5:1 to 3:1) to provide pure compound **6α** (651 mg, 93%). 1H NMR (500 MHz, chloroform- d) δ 8.02 – 7.91 (m, 2H, arom. H, Bz), 7.81 – 6.96 (m, 25H, arom. H, Bn, Nap, Bz, SPh), 5.59 – 5.54 (m, 1H, H-2,A), 5.48 (d, J = 1.7 Hz, 1H, H-1,A), 5.05 (d, J = 1.7 Hz, 1H, H-1,B), 4.88 – 4.30 (m, 6H, 2xCH₂, Bn, CH₂, Nap), 4.19 (m, 1H, H-5,A), 4.13 (m, 1H, H-3,A), 3.89 – 3.86 (m, 1H, H-2,B), 3.69 (m, 2H, H-3,B, H-5,B), 3.52 (m, 1H, H-4,A), 3.36 (m, 1H, H-4,B), 1.25 (m, 4H, CH₂-CH₂, Lev), 1.10 (d, J = 6.2 Hz, 3H, Rha, CH₃,B). ^{13}C NMR (75 MHz, chloroform- d) δ 138.49, 137.81, 135.31, 133.80, 133.29, 131.86, 129.87, 129.77, 129.06, 128.60, 128.52, 128.48, 128.36, 128.28, 128.26, 128.03, 127.95, 127.92, 127.88, 127.84, 127.73, 127.68, 127.63, 127.60, 127.50, 126.53, 126.18, 126.02, 125.67, 101.20, 85.65, 80.43, 79.70, 79.50, 78.08, 76.80, 75.35, 74.77, 74.40, 72.21, 69.17, 69.13, 69.07, 68.45,

17.94, 17.72. HR-MALDI-TOF/MS (m/z): calcd for $C_{50}H_{50}NaO_9S^+$ $[M+Na]^+$: 849.3212; found, 849.3218.

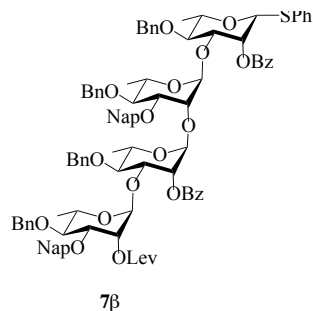


4

2-O-levulinoyl-3-O-Naphthylmethyl-4-O-benzyl- α -L-rhamnopyranosyl-(1 \rightarrow 3)-4-O-benzyl-

2-O-benzoyl- α -L-rhamnopyranose (4). To a mixture of compound **3a** (940 mg, 1.016 mmol), 2,4,6-tri-*tert*-butylpyrimidine (1.01 g, 4.062 mmol), and *N*-iodosuccinimide (457 mg, 2.031 mmol) in wet DCM (50 ml) at 0 °C was added silver triflate (520 mg, 2.031 mmol). After stirring for 30 min, saturated aqueous $Na_2S_2O_3$ was added and the mixture was stirred for an additional 30 min while warming to r.t. The reaction was extracted by DCM (3 \times 50 mL), dried with anhydrous $MgSO_4$, and the organic extracts were concentrated under reduced pressure. The crude product was purified by silica gel chromatography (hexane:EtOAc, 2:1 to 1:1) to provide pure compound **4** (821 mg, 97%). 1H NMR (500 MHz, chloroform-*d*) δ 8.00 – 7.92 (m, 2H, arom. H, Bz), 7.71 – 7.05 (m, 20H, arom. H, Bn, Nap), 5.32 (m, 2H, H-2,A, H-2,B), 5.19 (d, 1H, H-1,A), 4.99 (d, 1H, H-1,B), 4.79 – 4.35 (m, 6H, 2xCH₂, Bn, CH₂, Nap), 4.23 (m, 1H, H-3,A), 3.98 – 3.90 (m, 1H, H-5,A), 3.78 (m, 1H, H-3,B), 3.76 – 3.68 (m, 1H, H-5,B), 3.46 (m, 1H, H-4,A), 3.28 (m, 1H, H-4,B), 2.57 (m, 4H, CH₂-CH₂, Lev), 2.03 (s, 3H, CH₃, Lev), 1.21 (d, J = 6.2 Hz, 3H, Rha, CH₃,A), 1.08 (d, J = 6.2 Hz, 3H, Rha, CH₃,B). ^{13}C NMR (75 MHz, chloroform-*d*) δ 138.57, 137.79, 135.47, 133.34, 131.86, 129.76, 129.09, 128.51, 128.28, 128.24, 128.11, 128.00, 127.94, 127.90, 127.67, 127.64, 99.45, 92.01, 80.42, 79.69, 77.90, 75.49, 74.73, 74.30, 71.62,

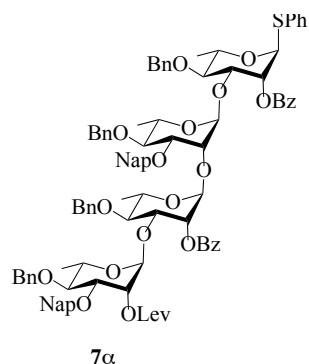
69.39, 69.20, 68.81, 38.03, 29.81, 28.15, 17.99, 17.82. HR-MALDI-TOF/MS (m/z): calcd for $C_{49}H_{52}NaO_{12}^+$ $[M+Na]^+$: 855.3513; found, 855.3519.



Thiophenyl O-(4-O-benzyl-3-O-naphthylmethyl-2-O-levulinoyl-α-L-rhamnopyranosyl)-(1→3)-(4-O-benzyl-2-O-benzoyl-α-L-rhamnopyranosyl)-(1→2)-(4-O-benzyl-3-O-naphthylmethyl-α-L-rhamnopyranosyl)-(1→3)-4-O-benzyl-2-O-benzoyl-β-L-

rhamnopyranoside (7β). Compound **7β** was obtained by the optimal glycosylation method as described in the General Procedures. Yield: >52% (From Prep TLC). 1H NMR (500 MHz, chloroform- d) δ 8.22 – 8.16 (m, 2H, arom. H, Bz), 8.07 – 8.00 (m, 2H, arom. H, Bz), 7.87 – 6.95 (m, 45H, arom. H, Bn, Nap, Bz, SPh), 5.85 (d, J = 3.4 Hz, 1H, H-2,A), 5.14 – 5.01 (m, 2H, H-1,D, H-1,B), 4.97 (d, 1H, H-1,A), 4.93 (d, J = 1.9 Hz, 1H, H-1,C), 4.91 – 4.33 (m, 11H, 7xCH, Bn, 2xCH₂, Nap), 4.26 (m, 1H, H-3,C), 4.20 (d, J = 11.4 Hz, 1H, CH, Bn), 4.02 (m, 1H, H-5,C), 3.93 – 3.84 (m, 3H, H-3,D, H-3,A, H-5,A), 3.83 – 3.71 (m, 2H, H-5,B, H-3,B), 3.55 – 3.41 (m, 4H, H-4,C, H-4,B, H-4,A, H-4,D), 3.38 (m, 1H, H-5,D), 2.66 (m, 4H, CH₂-CH₂, Lev), 2.11 (s, 3H, CH₃, Lev), 1.37 – 1.26 (m, 6H, Rha, CH₃,C,D), 1.16 (d, J = 6.2 Hz, 3H, Rha, CH₃,A), 1.05 (d, J = 6.1 Hz, 3H, Rha, CH₃,B). ^{13}C NMR (151 MHz, chloroform- d) δ 138.60, 137.90, 137.49, 135.98, 135.52, 133.25, 133.10, 132.85, 131.27, 130.12, 129.75, 128.95, 128.48, 128.42, 128.39, 128.25, 128.15, 127.97, 127.95, 127.89, 127.87, 127.78, 127.76, 127.72, 127.70, 127.64, 127.61, 127.58, 127.52, 127.49, 127.38, 127.22, 126.50, 126.38, 126.10, 125.97, 125.87, 125.72, 125.67,

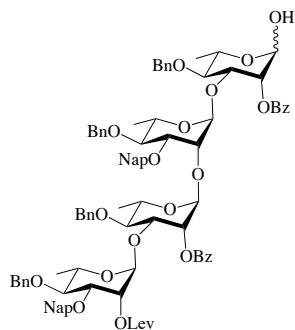
125.56, 100.54, 99.09, 85.56, 80.45, 80.28, 79.67, 78.51, 77.66, 77.07, 76.24, 75.30, 74.95, 74.67, 74.29, 73.91, 72.85, 72.47, 71.51, 69.34, 69.11, 68.58, 68.18, 38.01, 29.75, 28.09, 18.24, 18.12, 17.87, 17.72. HR-MALDI-TOF/MS (m/z): calcd for $C_{99}H_{100}NaO_{20}S^+$ $[M+Na]^+$: 1663.6523; found, 1663.6527.



Thiophenyl O-(4-O-benzyl-3-O-naphthylmethyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-naphthylmethyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-

rhamnopyranoside (7 α). Compound **7 α** was obtained by the optimal glycosylation method as described in the General Procedures. Yield: 85%. 1H NMR (500 MHz, chloroform- d) δ 7.96 (m, 4H, arom. H, Bz), 7.78 – 6.94 (m, 45H, arom. H, Bn, Nap, Bz, SPh), 5.53 (m, 1H, H-2,A), 5.50 – 5.48 (m, 1H, H-2,C), 5.47 (d, J = 1.7 Hz, 1H, H-1,A), 5.01 (m, 2H, H-1,D, H-2,D), 4.98 (d, J = 1.7 Hz, 1H, H-1,C), 4.81 – 4.26 (m, 12H, 4xCH₂, Bn, 2xCH₂, Nap), 4.22 (dd, J = 9.5, 3.2 Hz, 1H, H-3,C), 4.17 – 4.10 (m, 1H, H-5,A), 4.05 (m, 1H, H-3,A), 3.84 – 3.77 (m, 4H, H-3,D, H-5,D, H-5,C, H-2,B), 3.50 – 3.35 (m, 3H, H-4,C, H-4,A, H-4,B), 3.31 – 3.25 (m, 1H, H-4,D), 2.56 (m, 4H, CH₂-CH₂, Lev), 2.01 (s, 3H, CH₃, Lev), 1.21 – 1.14 (m, 6H, Rha, CH₃,A,C), 1.06 (m, 6H, Rha, CH₃,B,D). ^{13}C NMR (151 MHz, chloroform- d) δ 165.65, 137.74, 135.82, 135.52, 133.27, 133.18, 133.14, 131.82, 129.81, 129.77, 129.74, 129.02, 128.46, 128.45, 128.42, 128.40, 128.18, 128.14, 128.13, 128.02, 128.00, 127.93, 127.90, 127.88, 127.83, 127.78, 127.61, 127.59,

127.57, 127.53, 127.38, 127.28, 126.51, 126.32, 125.97, 125.88, 125.74, 125.72, 125.58, 100.97, 99.12, 98.93, 85.55, 80.41, 79.96, 79.86, 79.67, 78.81, 77.66, 76.36, 75.34, 75.17, 74.68, 74.41, 72.53, 72.26, 71.52, 69.33, 69.11, 69.09, 68.59, 68.39, 60.38, 38.01, 29.75, 28.09, 18.07, 17.88, 17.73, 14.19. HR-MALDI-TOF/MS (m/z): calcd for $C_{99}H_{100}NaO_{20}S^+$ $[M+Na]^+$: 1663.6523; found, 1663.6528.

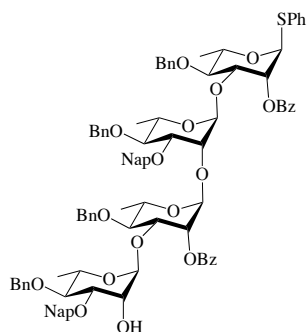


8

4-O-benzyl-3-O-naphthylmethyl-2-O-levulinoyl- α -L-rhamnopyranosyl-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-naphthylmethyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranose (8). Compound

7 α (202 mg, 0.123 mmol) was dissolved in acetone:PBS buffer (6:1, 14 mL), and NBS (109 mg, 0.615 mmol) was added at 0 °C. The solution was stirred for 5 h performed in the dark. The reaction was monitored by TLC. Upon completion of the reaction, the reaction mixture was added to an aqueous solution of $NaHCO_3$ (0.1 M) to quench the reaction. The reaction mixture was extracted by DCM (3 \times 50 mL). The organic extracts were dried with anhydrous $MgSO_4$, and concentrated under reduced pressure. The crude product was purified by silica gel chromatography (hexane:EtOAc, 2:1 to 1:1, v/v) to provide pure compound **8** (191 mg, 87%). 1H NMR (500 MHz, chloroform- d) δ 8.21 – 7.99 (m, 4H, arom. H, Bz), 7.90 – 6.96 (m, 40H, arom. H, Bn, Nap, Bz), 5.60 – 5.54 (m, 1H, H-2,A), 5.45 (m, 1H, H-2,C), 5.33 – 5.24 (d, 1H, H-1,A),

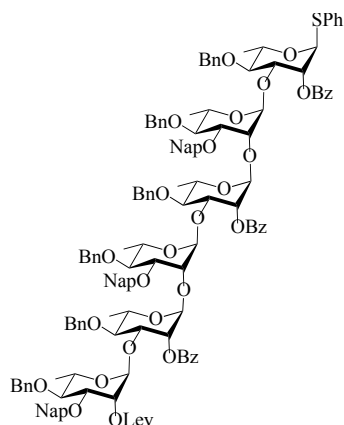
5.10 (d, $J=1.9\text{Hz}$, 1H, H-1,D), 5.08 (d, $J=1.8\text{Hz}$, 1H, H-1,C), 5.06 (m, 1H, H-2,D), 4.93 – 4.42 (m, 12H, 4xCH₂, Bn, CH₂, Nap), 4.37 (m, 1H, H-3,C), 4.31 (m, 1H, H-5,A), 4.24 (m, 1H, H-3,A), 4.14 (m, 1H, H-3,B), 4.00 (m, 1H, H-3,D), 3.96 – 3.84 (m, 2H, H-5,C, H-5,B), 3.84 – 3.65 (m, 3H, H-4,C, H-4,B, H-2,A), 3.60 – 3.51 (m, 1H, H-4,D), 3.46 (m, 2H, H-4,A, H-5,D), 2.70 – 2.57 (m, 4H, CH₂-CH₂, Lev), 2.10 (s, 3H, CH₃, Lev), 1.31 – 1.20 (m, 6H, Rha, CH₃,A,C), 1.16 (m, 6H, Rha, CH₃,B,D). ¹³C NMR (151 MHz, chloroform-*d*) δ 206.20, 171.87, 165.87, 165.46, 138.77, 138.62, 137.91, 135.88, 135.54, 133.23, 133.22, 133.20, 133.16, 132.89, 132.81, 130.15, 130.09, 130.04, 129.94, 129.81, 129.80, 128.57, 128.48, 128.46, 128.44, 128.43, 128.38, 128.37, 128.33, 128.31, 128.26, 128.21, 128.17, 128.15, 128.13, 128.07, 128.05, 128.03, 128.02, 127.96, 127.94, 127.91, 127.89, 127.86, 127.81, 127.78, 127.75, 127.72, 127.64, 127.62, 127.60, 127.56, 127.52, 127.42, 127.29, 126.55, 126.51, 126.34, 126.01, 125.98, 125.93, 125.92, 125.85, 125.79, 125.76, 125.73, 125.60, 125.60, 100.84, 99.12, 98.91, 91.71, 80.45, 80.00, 79.88, 79.69, 78.90, 77.67, 76.73, 76.48, 75.35, 75.04, 74.70, 74.66, 74.47, 73.27, 72.56, 72.34, 72.26, 71.55, 69.37, 69.36, 68.95, 68.67, 68.61, 68.37, 67.69, 60.46, 38.03, 30.95, 29.82, 29.78, 28.11, 18.08, 18.07, 18.00, 17.92, 17.79, 17.75. HR-MALDI-TOF/MS (*m/z*): calcd for C₉₃H₉₆NaO₂₁⁺ [M+Na]⁺: 1571.6423; found, 1571.6427.



10α

Thiophenyl O-(4-O-benzyl-3-O-naphthylmethyl-2- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-naphthylmethyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (10 α).

Hydrazine acetate (86 mg, 0.932 mmol) was added to a solution of **7 α** (306 mg, 0.186 mmol) in a mixture of DCM and MeOH (1/1, v/v, 20 mL). Stirring was continued until MALDI indicated the disappearance of the starting material (~3 h). The reaction mixture was diluted with DCM (30 mL), washed with water (3 \times 25 mL) and brine (25 mL), dried with anhydrous MgSO₄, and filtered. The crude product was purified by silica gel chromatography (hexane:EtOAc, 5:1 to 3:1, v/v) to provide pure compound **10 α** (284 mg, 99%). ¹H NMR (500 MHz, chloroform-*d*) δ 8.04 – 7.90 (m, 4H, arom. H, Bz), 7.80 – 6.94 (m, 45H, arom. H, Bn, Nap, Bz, SPh), 5.53 (m, 1H, H-2,A), 5.51 (m, 1H, H-2,C), 5.47 (d, *J* = 1.8 Hz, 1H, H-1,A), 5.05 – 5.04 (m, 2H, H-1,D, H-2,D), 5.03 (d, *J* = 2.0 Hz, 1H, H-1,C), 4.97 (d, *J* = 1.9 Hz, 1H, H-1,B), 4.80 – 4.25 (m, 12H, 4xCH₂, Bn, 2xCH₂, Nap), 4.19 (m, 1H, H-3,C), 4.16 – 4.10 (m, 1H, H-5,A), 4.09 – 4.02 (m, 1H, H-3,A), 3.86 – 3.74 (m, 4H, H-3,D, H-5,D, H-5,C, H-2,B), 3.71 – 3.31 (m, 4H, H-4,C, H-4,A, H-4,B, H-4,D), 1.23 – 1.13 (m, 6H, Rha, CH₃,A,C), 1.05 (m, 6H, Rha, CH₃,B,D). ¹³C NMR (151 MHz, chloroform-*d*) δ 165.65, 165.42, 138.65, 138.52, 137.95, 137.73, 135.83, 135.35, 133.85, 133.28, 133.17, 133.15, 133.12, 132.92, 132.79, 131.81, 130.12, 129.80, 129.78, 129.74, 129.05, 129.02, 128.45, 128.43, 128.40, 128.33, 128.23, 128.15, 128.14, 128.01, 127.93, 127.88, 127.86, 127.84, 127.82, 127.79, 127.76, 127.65, 127.63, 127.59, 127.57, 127.53, 127.47, 127.29, 126.47, 126.30, 126.13, 125.97, 125.87, 125.73, 125.67, 125.65, 125.57, 100.97, 100.85, 98.89, 85.55, 80.36, 80.33, 79.92, 79.83, 79.73, 79.66, 79.63, 78.77, 76.30, 75.20, 74.77, 74.72, 74.39, 72.82, 72.63, 72.27, 72.11, 69.13, 69.07, 68.40, 68.26, 29.69, 18.07, 17.93, 17.88, 17.74, 17.70, 17.66. HR-MALDI-TOF/MS (*m/z*): calcd for C₉₄H₉₄NaO₁₈S⁺ [*M*+Na]⁺: 1565.6212; found, 1565.6217.

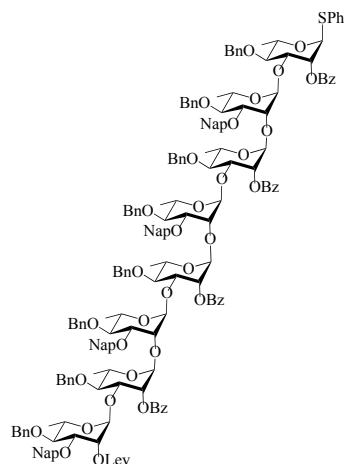


Hex

Thiophenyl O-(4-O-benzyl-3-O-naphthylmethyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-naphthylmethyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-naphthylmethyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (Hex). The Hexasaccharide is obtained from

the “4+4” polymerization, and details are in the “4+4” polymerization section. ^1H NMR (600 MHz, chloroform-*d*) δ 8.04 – 7.90 (m, 6H, arom. H, Bz), 7.80 – 6.84 (m, 75H, arom. H, Bn, Bz, Nap, SPh), 5.51 (m, 1H, H-2,A), 5.47 (m, 3H, H-1,A, H-2,C, H-2,E), 5.37 – 5.30 (m, 1H, H-2,F), 5.05 (m, 1H, H-1,C), 5.01 – 4.88 (m, 4H, H-1,E, H-1,F, H-1,B, H-1,D), 4.80 – 4.25 (m, 16H, 6xCH₂, Bn, 2xCH₂, Nap), 3.85 – 3.64 (m, 6H, H-2,B, H-2,D, H-2,F, H-3,B, H-3,D, H-3,F), 2.53 (m, 4H, CH₂-CH₂, Lev), 1.98 (m, 3H, CH₃, Lev), 1.21 – 1.14 (m, 9H, Rha, CH₃,A,C,E), 1.09 – 1.01 (m, 9H, Rha, CH₃,B,D,F). ^{13}C NMR (151 MHz, chloroform-*d*) δ 171.76, 165.39, 138.66, 137.86, 135.53, 133.87, 133.25, 133.18, 133.08, 132.85, 132.77, 131.80, 130.11, 130.03, 129.80, 129.76, 129.73, 129.00, 128.45, 128.42, 128.39, 128.37, 128.35, 128.32, 128.25, 128.16, 128.11, 128.09, 127.98, 127.95, 127.93, 127.87, 127.85, 127.83, 127.78, 127.76, 127.73, 127.70, 127.67, 127.63, 127.61, 127.59, 127.57, 127.52, 127.35, 127.27, 127.23, 126.49, 126.29, 126.20, 125.97,

125.92, 125.85, 125.79, 125.69, 125.53, 85.52, 80.38, 79.86, 79.65, 77.66, 75.26, 75.18, 74.62, 72.69, 72.54, 72.25, 72.19, 71.49, 69.31, 69.06, 68.55, 68.30, 38.00, 29.73, 29.69, 28.08, 18.02, 17.87, 17.76, 17.71, 14.11. HR-MALDI-TOF/MS (m/z): calcd for $C_{143}H_{144}NaO_{29}S^+$ $[M+Na]^+$: 2379.9523; found, 2379.9528.

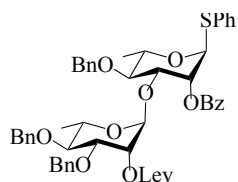


Octa

Thiophenyl O-(4-O-benzyl-3-O-naphthylmethyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-naphthylmethyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-naphthylmethyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-naphthylmethyl- α -L-rhamnopyranosyl) (1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (Octa). The

Octasaccharide is obtained from the “4+4” polymerization, and details are in the “4+4” polymerization section. 1H NMR (800 MHz, chloroform- d) δ 8.14 – 7.93 (m, 8H, arom. H, Bz), 7.81 – 7.01 (m, 85H, arom. H, Bn, Bz, Nap, SPh), 5.57 – 5.51 (m, 4H, H-1,A, H-2,C, H-2,E, H-2,G), 5.12 (d, J = 2.6 Hz, 1H, H-1,C), 5.10 – 4.99 (m, 6H, H-1,E, H-1,G, H-1,H, H-1,B, H-1,D, H-1,F), 4.88 – 4.32 (m, 24H, 8xCH₂, Bn, 4xCH₂, Nap), 3.91 – 3.73 (m, 8H, H-2,B, H-2,D, H-

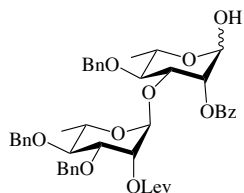
2,F, H-2,H, H-3,B, H-3,D, H-3,F, H-3,H), 2.64 – 2.56 (m, 4H, CH₂-CH₂, Lev), 2.05 (m, 3H, CH₃, Lev), 1.29 – 1.20 (m, 12H, Rha, CH₃,A,C,E,G), 1.15 – 0.99 (m, 12H, Rha, CH₃, B,D,F,G). ¹³C NMR (151 MHz, chloroform-*d*) δ 167.67, 165.40, 133.17, 133.07, 132.76, 132.28, 131.79, 130.90, 130.88, 129.78, 129.73, 129.00, 128.79, 128.77, 128.45, 128.41, 128.38, 128.37, 128.32, 128.25, 128.22, 128.15, 128.10, 127.94, 127.87, 127.85, 127.78, 127.73, 127.68, 127.65, 127.61, 127.58, 127.56, 127.52, 127.21, 126.49, 126.19, 125.97, 125.92, 125.85, 125.68, 125.53, 74.61, 72.18, 71.48, 68.87, 68.52, 68.26, 66.18, 65.89, 65.87, 65.01, 64.85, 64.71, 64.31, 64.26, 62.07, 46.58, 44.54, 44.22, 43.06, 39.33, 38.41, 37.98, 37.08, 36.90, 36.45, 36.22, 36.12, 36.06, 35.40, 35.21, 34.64, 34.32, 33.97, 33.72, 33.56, 33.50, 33.42, 33.33, 33.07, 31.96, 31.91, 31.61, 31.48, 31.46, 31.03, 30.91, 30.57, 30.26, 30.15, 30.13, 30.04, 29.68, 29.65, 29.61, 29.47, 29.39, 29.36, 29.30, 29.23, 29.11, 29.05, 28.98, 28.88, 28.85, 28.26, 28.06, 27.92, 27.44, 27.29, 27.27, 26.73, 26.70, 26.60, 26.13, 26.11, 26.03, 25.95, 25.44, 25.29, 25.20, 25.17, 24.79, 23.91, 23.41, 22.91, 22.78, 22.75, 22.68, 22.65, 22.61, 22.56, 22.53, 22.19, 20.70, 20.66, 20.20, 20.15, 20.08, 19.58, 19.55, 19.33, 19.16, 18.88, 17.93, 17.86, 17.75, 17.70, 16.11, 15.91, 15.25, 14.38, 14.30, 14.24, 14.13, 13.89, 12.17, 11.39, 11.18, 10.43. HR-MALDI-TOF/MS (m/z): calcd for C₁₈₇H₁₈₈NaO₃₈S⁺ [M+Na]⁺: 3096.2523; found, 3096.2527.



20α

Thiophenyl O-(4-O-benzyl-3-O-benzyl-2-O-levulinoyl-α-L-rhamnopyranosyl)-(1→3)-4-O-benzyl-2-O-benzoyl-α-L-rhamnopyranoside (20α). Compound **20α** was obtained by the optimal glycosylation method as described in the General Procedures. Yield: 80%. ¹H NMR

(600 MHz, chloroform-*d*) δ 8.02 (m, 2H, arom. H, Bz), 7.74 – 7.04 (m, 23H, arom. H, Bn, Bz, SPh), 5.60 (m, 1H, H-2,A), 5.35 (m, 1H, H-1,A), 5.13 – 5.01 (m, 1H, H-1,B), 4.92 – 4.12 (m, 8H, 3xCH₂, Bn, H-3,A, H-5,A), 3.81 – 3.69 (m, 2H, H-3,B, H-5,B), 3.63 (m, 1H, H-4,A), 3.37 – 3.30 (m, 1H, H-4,B), 2.73 – 2.59 (m, 4H, CH₂-CH₂, Lev), 2.14 (s, 3H, CH₃, Lev), 1.15 (d, *J* = 6.2 Hz, 3H, Rha, CH₃,A), 0.90 (m, 3H, Rha, CH₃,B). ¹³C NMR (151 MHz, chloroform-*d*) δ 181.94, 171.82, 165.61, 138.48, 133.76, 133.31, 131.81, 130.85, 129.75, 129.34, 129.04, 128.77, 128.48, 128.47, 128.20, 128.17, 128.04, 128.01, 127.90, 127.88, 127.62, 127.60, 127.55, 127.42, 99.31, 85.63, 80.49, 79.60, 77.61, 77.39, 75.49, 74.66, 74.23, 71.57, 69.34, 69.15, 68.71, 68.13, 38.70, 38.01, 30.34, 29.81, 29.68, 28.90, 28.09, 22.97, 17.94, 17.75, 10.94. HR-MALDI-TOF/MS (*m/z*): calcd for C₅₁H₅₄NaO₁₁S⁺ [M+Na]⁺: 897.3412; found, 897.3418.

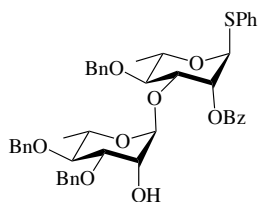


21

2-O-levulinoyl-3-O-benzyl-4-O-benzyl- α -L-rhamnopyranosyl-(1 \rightarrow 3)-4-O-benzyl-2-O-

benzoyl- α -L-rhamnopyranose (21). Compound **20 α** (1.00 g, 1.14 mmol) was dissolved in acetone:PBS buffer (6:1, 35 mL), and NBS (1.02 g, 5.72 mmol) was added at 0 °C. The solution was stirred for 5 h performed in the dark. The reaction was monitored by TLC. Upon completion of the reaction, the reaction mixture was added to an aqueous solution of NaHCO₃ (0.1 M) to quench the reaction. The reaction mixture was extracted by DCM (3 \times 50 mL). The organic extracts were dried with anhydrous MgSO₄, and concentrated under reduced pressure. The crude product was purified by silica gel chromatography (hexane:EtOAc, 2:1 to 1:1) to provide pure compound **21** (750 mg, 84%). ¹H NMR (600 MHz, chloroform-*d*) δ 8.17 – 7.93 (m, 2H, arom. H,

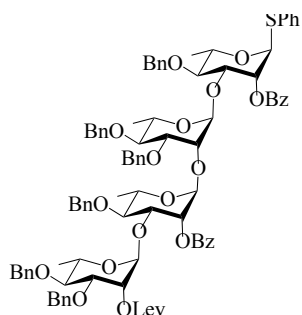
Bz), 7.70 – 7.00 (m, 18H, arom. H, Bn, Bz), 5.38 (m, 1H, H-2,A), 5.34 (m, 1H, H-1,A), 5.31 – 5.26 (m, 1H, H-2,B), 5.04 (d, $J = 2.0$ Hz, 1H, H-1,B), 4.84 – 4.39 (m, 6H, 3xCH₂, Bn), 4.34 – 4.25 (m, 2H, H-3,A, H-5,A), 3.80 – 3.69 (m, 2H, H-3,B, H-5,B), 3.57 (m, 1H, H-4,A), 3.32 (m, 1H, H-4,B), 2.71 – 2.57 (m, 4H, CH₂-CH₂, Lev), 2.13 (s, 3H, CH₃, Lev), 1.32 – 1.28 (m, 3H, Rha, CH₃,A), 1.13 (m, 3H, Rha, CH₃,B). ¹³C NMR (151 MHz, chloroform-*d*) δ 216.25, 208.91, 171.81, 164.20, 133.25, 129.78, 128.46, 128.43, 128.19, 128.15, 127.98, 127.89, 127.60, 127.52, 99.22, 91.83, 80.33, 79.62, 77.42, 75.35, 74.60, 72.88, 71.53, 69.40, 68.56, 67.85, 38.01, 29.80, 28.09, 18.11, 17.74. HR-MALDI-TOF/MS (*m/z*): calcd for C₄₅H₅₀NaO₁₂⁺ [M+Na]⁺: 805.3312; found, 805.3314.



23α

Thiophenyl O-(4-O-benzyl-3-O-benzyl-α-L-rhamnopyranosyl)-(1→3)-4-O-benzyl-2-O-benzoyl-α-L-rhamnopyranoside (23α). Hydrazine acetate (263 mg, 2.86 mmol) was added to a solution of **23α** (501 mg, 0.57 mmol) in a mixture of DCM and MeOH (1/1, v/v, 40 mL). Stirring was continued until MALDI indicated the disappearance of the starting material (~3 h). The reaction mixture was diluted with DCM (30 mL), washed with water (3 × 25 mL) and brine (25 mL), dried with anhydrous MgSO₄, and filtered. The crude product was purified by silica gel chromatography (hexane:EtOAc, 5:1 to 3:1) to provide pure compound **23α** (353 mg, 80%). ¹H NMR (600 MHz, chloroform-*d*) δ 8.10 – 7.95 (m, 2H, arom. H, Bz), 7.64 – 7.13 (m, 23H, arom. H, Bn, Bz, SPh), 5.62 (m, 1H, H-2,A), 5.54 (d, $J = 1.8$ Hz, 1H, H-1,A), 5.11 (d, $J = 1.8$ Hz, 1H,

H-1,B), 4.84 – 4.42 (m, 6H, 3xCH₂, Bn), 4.26 (m, 1H, H-5,A), 4.20 (m, 1H, H-3,A), 3.90 (m, 1H, H-2,B), 3.79 – 3.72 (m, 1H, H-5,B), 3.68 (m, 1H, H-3,B), 3.62 (m, 1H, H-4,B), 3.38 (m, 1H, H-4,A), 1.34 (d, $J = 6.2$ Hz, 3H, Rha, CH₃,A), 1.15 (d, $J = 6.2$ Hz, 3H, Rha, CH₃,B). ¹³C NMR (151 MHz, chloroform-*d*) δ 165.59, 137.83, 137.81, 133.77, 133.27, 131.82, 129.76, 129.04, 128.53, 128.46, 128.42, 128.22, 127.95, 127.89, 127.86, 127.75, 127.63, 127.61, 127.49, 101.11, 85.63, 80.53, 79.63, 79.44, 77.89, 75.44, 74.74, 74.37, 72.11, 69.15, 69.06, 68.37, 17.94, 17.68. HR-MALDI-TOF/MS (*m/z*): calcd for C₄₆H₄₈NaO₉S⁺ [M+Na]⁺: 799.3012; found, 799.3018.

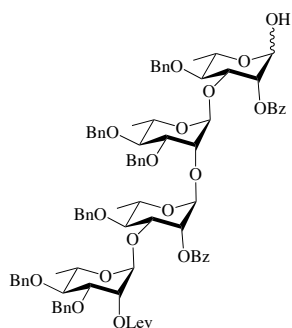


24α

Thiophenyl O-(4-O-benzyl-3-O-benzyl-2-O-levulinoyl-α-L-rhamnopyranosyl)-(1→3)-(4-O-benzyl-2-O-benzoyl-α-L-rhamnopyranosyl)-(1→2)-(4-O-benzyl-3-O-benzyl-α-L-rhamnopyranosyl)-(1→3)-4-O-benzyl-2-O-benzoyl-α-L-rhamnopyranoside (24α).

Compound **24α** was obtained by the optimal glycosylation method as described in the General Procedures. Yield: 90%. ¹H NMR (600 MHz, chloroform-*d*) δ 8.05 (m, 4H, arom. H, Bz), 7.69 – 6.93 (m, 41H, arom. H, Bn, Bz, SPh), 5.61 – 5.57 (m, 1H, H-2,A), 5.54 (d, $J = 1.8$ Hz, 1H, H-1,A), 5.50 (m, 1H, H-2,C), 5.36 (m, 1H, H-2,D), 5.06 (m, 2H, H-1,B, H-1,D), 5.02 (d, $J = 2.0$ Hz, 1H, H-1,C), 4.84 – 4.29 (m, 12H, 6xCH₂, Bn), 4.28 (m, 1H, H-3,C), 4.24 – 4.19 (m, 1H, H-5,C), 4.14 (m, 1H, H-3,A), 3.88 (m, 1H, H-5,A), 3.85 – 3.75 (m, 3H, H-5,D, H-3,D, H-2,B), 3.71 (m, 1H, H-3,B), 3.65 (m, 1H, H-5,B), 3.56 (m, 2H, H-4,A, H-4,C), 3.32 (m, 1H, H-4,D), 2.73 –

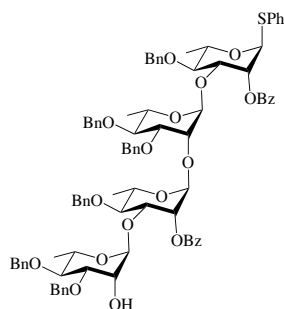
2.53 (m, 4H, CH₂-CH₂, Lev), 2.11 (s, 3H, CH₃, Lev), 1.25 (m, 6H, Rha, CH₃,A,C), 1.12 (m, 6H, Rha, CH₃,B,D). ¹³C NMR (151 MHz, chloroform-*d*) δ 138.68, 138.59, 133.27, 133.14, 131.81, 129.79, 129.75, 129.02, 128.46, 128.41, 128.20, 128.18, 128.15, 128.12, 127.96, 127.90, 127.83, 127.77, 127.69, 127.67, 127.63, 127.57, 127.49, 127.39, 127.37, 127.27, 100.92, 98.99, 98.93, 85.55, 80.55, 80.06, 79.85, 79.65, 78.78, 78.56, 77.58, 76.35, 76.29, 75.33, 74.64, 74.42, 72.48, 72.12, 71.55, 69.36, 69.11, 69.06, 68.53, 68.40, 38.02, 28.09, 18.10, 17.91, 17.73. HR-MALDI-TOF/MS (*m/z*): calcd for C₉₁H₉₆NaO₂₀S⁺ [M+Na]⁺: 1563.6212; found, 1563.6217.



25

4-O-benzyl-3-O-benzyl-2-O-levulinoyl- α -L-rhamnopyranosyl-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-benzyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranose (25). Compound **24 α** (330 mg, 0.214 mmol) was dissolved in acetone:PBS buffer (6:1, 21 mL), and NBS (191 mg, 1.071 mmol) was added at 0 °C. The solution was stirred for 5 h performed in the dark. The reaction was monitored by TLC. Upon completion of the reaction, the reaction mixture was added to an aqueous solution of NaHCO₃ (0.1 M) to quench the reaction. The reaction mixture was extracted with DCM (3 \times 50 mL). The organic extracts were dried with anhydrous MgSO₄, and concentrated under reduced pressure. The crude product was purified by silica gel chromatography (hexane:EtOAc, 2:1 to 1:1) to provide pure compound **25** (285 mg, 92%). ¹H

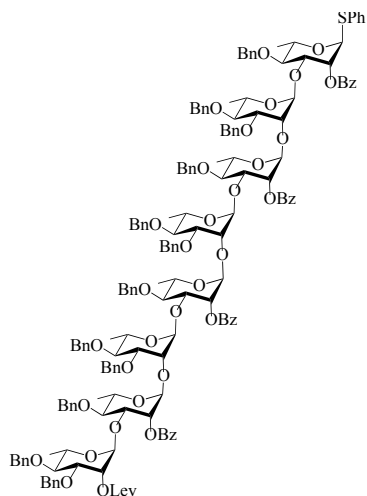
NMR (600 MHz, chloroform-*d*) δ 8.05 (m, 4H, arom. H, Bz), 7.64 – 7.00 (m, 36H, arom. H, Bn, Bz), 5.48 (m, 1H, H-2,A), 5.36 (m, 2H, H-1,A, H-2,C), 5.29 – 5.25 (m, 1H, H-2,D), 5.04 (m, 2H, H-1,B, H-1,D), 4.81 – 4.28 (m, 12H, 6xCH₂, Bn), 4.24 (m, 2H, H-3,C, H-5,C), 3.89 – 3.76 (m, 5H, H-3,A, H-5,A, H-5,D, H-3,D, H-2,B), 3.71 – 3.60 (m, 2H, H-3,B, H-5,B), 3.52 (m, 2H, H-4,A, H-4,C), 3.39 (m, 1H, H-4,D), 2.67 – 2.53 (m, 4H, CH₂-CH₂, Lev), 2.10 (s, 3H, CH₃, Lev), 1.23 (m, 6H, Rha, CH₃,A,C), 1.11 (m, 6H, Rha, CH₃,B,D). ¹³C NMR (151 MHz, chloroform-*d*) δ 138.72, 138.33, 137.97, 133.21, 133.12, 129.88, 129.78, 128.45, 128.43, 128.41, 128.39, 128.17, 128.14, 128.09, 127.94, 127.89, 127.78, 127.75, 127.67, 127.60, 127.48, 127.38, 127.33, 127.23, 100.83, 91.75, 80.57, 79.94, 79.64, 77.57, 75.20, 74.63, 73.12, 72.47, 72.08, 71.54, 69.35, 68.89, 68.51, 68.35, 67.81, 38.02, 29.78, 28.08, 18.07, 17.72, 17.68. HR-MALDI-TOF/MS (*m/z*): calcd for C₈₅H₉₂NaO₂₁⁺ [M+Na]⁺: 1471.6123; found, 1471.6128.



27α

Thiophenyl O-(4-O-benzyl-3-O-benzyl-α-L-rhamnopyranosyl)-(1→3)-(4-O-benzyl-2-O-benzoyl-α-L-rhamnopyranosyl)-(1→2)-(4-O-benzyl-3-O-benzyl-α-L-rhamnopyranosyl)-(1→3)-4-O-benzyl-2-O-benzoyl-α-L-rhamnopyranoside (27α). Hydrazine acetate (44 mg, 0.476 mmol) was added to a solution of **24α** (147 mg, 0.095 mmol) in a mixture of DCM and MeOH (1/1, v/v, 20 mL). Stirring was continued until MALDI indicated the disappearance of the starting material (~3 h). The reaction mixture was diluted with DCM (30 mL), washed with

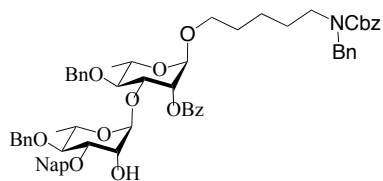
water (3 × 25 mL) and brine (25 mL), dried with anhydrous MgSO₄, and filtered. The crude product was purified by silica gel chromatography (hexane:EtOAc, 5:1 to 3:1) to provide pure compound **27a** (111 mg, 81%). ¹H NMR (600 MHz, chloroform-*d*) δ 8.06 (m, 4H, arom. H, Bz), 7.82 – 6.99 (m, 41H, arom. H, Bn, Bz, SPh), 5.58 – 5.49 (m, 2H, H-2,A, H-1,A), 5.11 (d, 1H, H-1,B), 5.07 (d, *J* = 2.0 Hz, 1H, H-1,D), 5.04 (d, *J* = 2.0 Hz, 1H, H-1,C), 4.83 – 4.39 (m, 12H, 6xCH₂, Bn), 4.31 – 4.08 (m, 4H, H-3,C, H-5,C, H-3,A, H-5,A), 3.90 (m, 2H, H-5,D, H-3,D), 3.86 – 3.78 (m, 2H, H-5,B, H-3,B), 3.71 (m, 2H, H-2,B, H-2,D), 3.69 – 3.49 (m, 4H, H-4,C, H-4,A, H-4,B, H-4,D), 1.32 – 1.23 (m, 6H, Rha, CH₃,A,C), 1.12 (m, 6H, Rha, CH₃,B,D). ¹³C NMR (151 MHz, chloroform-*d*) δ 138.54, 138.36, 137.99, 137.92, 137.82, 133.87, 133.27, 133.09, 131.81, 130.17, 129.82, 129.80, 129.76, 129.02, 128.48, 128.46, 128.43, 128.41, 128.21, 128.12, 127.86, 127.85, 127.81, 127.74, 127.68, 127.66, 127.65, 127.57, 127.46, 127.34, 127.28, 100.95, 100.74, 98.89, 85.56, 80.49, 80.05, 79.82, 79.74, 79.60, 78.77, 78.58, 76.86, 76.29, 75.36, 75.26, 74.73, 74.69, 74.42, 72.60, 72.14, 72.04, 69.11, 69.05, 68.41, 68.19, 18.11, 17.92, 17.74, 17.64. HR-MALDI-TOF/MS (*m/z*): calcd for C₈₆H₉₀NaO₁₈S⁺ [M+Na]⁺: 1465.5823; found, 1465.5829.



Octa-Bn

Thiophenyl O-(4-O-benzyl-3-O-benzyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-

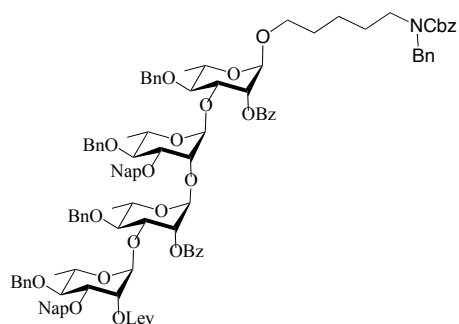
benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-benzyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-benzyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-benzyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (**Octa-Bn**). Compound **Octa-Bn** was obtained by the glycosylation method as described in the “4+4-Bn” section. Yield: 26%. ^1H NMR (600 MHz, chloroform-*d*) δ 8.14 – 7.92 (m, 8H, arom. H, Bz), 7.73 – 6.85 (m, 77H, arom. H, Bn, SPh), 5.51 – 5.43 (m, 4H, H-1,A, H-2,C, H-2,E, H-2,G), 5.10 – 4.91 (m, 7H, H-1,C, H-1,E, H-1,G, H-1,H, H-1,B, H-1,D, H-1,F), 4.82 – 4.33 (m, 24H, 12xCH₂, Bn), 3.93 – 3.58 (m, 4H, H-2,B, H-2,D, H-2,F, H-2,H), 3.58 – 3.24 (m, 4H, H-3,B, H-3,D, H-3,F, H-3,H), 2.65 – 2.55 (m, 4H, CH₂-CH₂, Lev), 1.28 – 1.22 (m, 12H, Rha, CH₃,A,C,E,G), 1.13 – 1.05 (m, 12H, Rha, CH₃,B,D,F,H). ^{13}C NMR (151 MHz, chloroform-*d*) δ 165.31, 131.79, 129.78, 129.00, 128.44, 128.32, 128.16, 128.08, 127.90, 127.84, 127.77, 127.65, 127.29, 74.57, 71.52, 38.02, 17.72. HR-MALDI-TOF/MS (*m/z*): calcd for C₁₇₁H₁₈₀NaO₃₈S⁺ [M+Na]⁺: 2896.1923; found, 2896.1929.



15

N-Benzyl-N-benzyloxycarbonyl-5-aminopentyl O-(4-O-benzyl-3-O-naphthylmethyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (15). Compound **15** was derived from the cleavage of 1,2-glycosyl linkage during the glycosylation of the tetrasaccharide with the linker. Details were described in the linker problem section. ^1H NMR (600 MHz, chloroform-*d*) δ 8.06 (m, 2H, arom. H, Bz), 7.89 – 7.01 (m, 30H, arom. H, Bz, Bn,

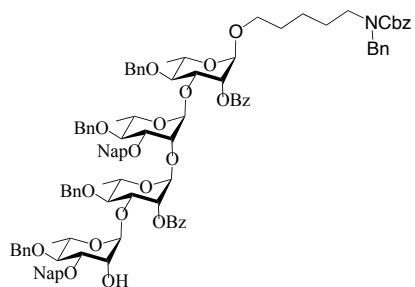
Nap), 5.36 (s, 1H, H-2,A), 5.18 (m, 2H, CH₂, Linker), 5.09 (d, 1H, H-1,B), 4.89 – 4.42 (m, 8H, CH₂, Bn, CH₂, Nap), 4.19 (m, 1H, H-3,A), 3.97 – 3.89 (m, 1H, H-2,B), 3.80 – 3.68 (m, 3H, H-3,B, H-5,A, H-5,B), 3.65 – 3.14 (m, 7H, 3×CH₂, Linker, H-4,B), 1.30 – 1.23 (m, 6H, Rha, CH₃,A,B). ¹³C NMR (151 MHz, chloroform-*d*) δ 165.82, 138.52, 137.90, 135.34, 133.20, 133.15, 132.94, 130.00, 129.77, 128.52, 128.49, 128.45, 128.40, 128.25, 128.23, 127.91, 127.88, 127.85, 127.82, 127.76, 127.66, 127.54, 127.46, 127.31, 127.23, 126.50, 126.15, 125.99, 125.67, 101.19, 97.00, 80.33, 79.72, 79.48, 78.06, 75.33, 74.69, 72.93, 72.12, 69.16, 68.27, 67.79, 67.71, 67.62, 67.19, 67.13, 50.56, 50.25, 47.11, 46.16, 29.70, 29.15, 27.92, 27.51, 23.38, 18.06, 17.70. HR-MALDI-TOF/MS (*m/z*): calcd for C₆₄H₆₉NNaO₁₂⁺ [M+Na]⁺: 1066.4821; found, 1066.4828.



12

***N*-Benzyl-*N*-benzyloxycarbonyl-5-aminopentyl O-(4-O-benzyl-3-O-naphthylmethyl-2-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-naphthylmethyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (12).** Compound **12** was derived from the optimal glycosylation method as described in the General Procedures. Yield: 36%. ¹H NMR (500 MHz, chloroform-*d*) δ 8.16 – 7.92 (m, 4H, arom. H, Bz), 7.92 – 7.01 (m, 50H, arom. H, Bn, Nap, SPh), 5.17 (m, 2H, H-1,B, H-1,D), 5.08 (m, 2H, H-1,C, H-1,A), 4.94 – 4.39 (m, 14H, 5×CH₂, Bn, 2×CH₂, Nap), 3.19 (m, 4H, H-4,C, H-4,A, H-4,B, H-4,D), 2.69 – 2.53 (m, 4H, CH₂-CH₂, Lev),

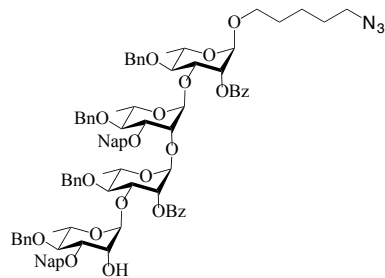
2.08 (s, 3H, CH₃, Lev), 1.27 (m, 6H, Rha, CH₃,A,C), 1.19 – 1.09 (m, 6H, Rha, CH₃,B,D). ¹³C NMR (151 MHz, chloroform-*d*) δ 206.96, 171.80, 167.73, 137.87, 133.17, 132.84, 132.40, 130.85, 129.99, 129.74, 128.76, 128.49, 128.43, 128.39, 128.36, 128.16, 128.10, 127.98, 127.93, 127.90, 127.88, 127.85, 127.82, 127.76, 127.59, 127.57, 127.52, 127.50, 127.37, 127.21, 126.48, 126.29, 125.95, 125.87, 125.71, 125.55, 99.07, 96.91, 79.99, 79.64, 77.63, 77.06, 75.32, 72.48, 72.22, 71.48, 69.29, 68.57, 68.12, 67.13, 52.75, 45.75, 38.69, 37.98, 30.92, 30.32, 29.74, 29.68, 29.08, 28.90, 28.06, 23.71, 22.96, 18.02, 17.98, 17.71, 14.04, 10.94, 8.54, 7.74. HR-MALDI-TOF/MS (m/z): calcd for C₁₁₃H₁₁₉NNaO₂₃⁺ [M+Na]⁺: 1880.8222; found, 1880.8228.



14

***N*-Benzyl-*N*-benzyloxycarbonyl-5-aminopentyl O-(4-O-benzyl-3-O-naphthylmethyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-naphthylmethyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (14).** Hydrazine acetate (15 mg, 0.163 mmol) was added to a solution of **12** (61 mg, 0.033 mmol) in a mixture of DCM and MeOH (1/1, v/v, 10 mL). Stirring was continued until MALDI indicated the disappearance of the starting material (~3 h). The reaction mixture was diluted with DCM (30 mL), washed with water (3 \times 25 mL) and brine (25 mL), dried with anhydrous MgSO₄, and filtered. The crude product was purified by silica gel chromatography (hexane:EtOAc, 5:1 to 3:1) to provide pure compound **14** (49 mg, 85%). ¹H NMR (600 MHz, chloroform-*d*) δ 8.04 (m, 4H, arom. H, Bz), 7.91 – 6.92 (m, 50H, arom. H, Bn, Bz, Nap, Cbz),

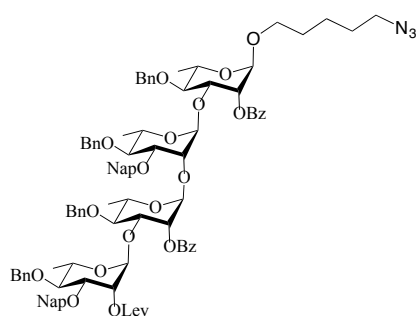
5.15 (m, 3H, H-1,C, H-1,D, H-1,A), 4.87 – 4.28 (m, 14H, 5xCH₂, Bn, 2xCH₂, Nap), 4.21 (m, 5H, H-3,C, H-5,A, H-3,A, H-2,D, H-2,B), 4.15 – 3.11 (m, 10H, H-3,D, H-5,D, H-5,C, H-2,B, H-3,B, H-5,B, H-4,C, H-4,A, H-4,B, H-4,D), 1.20 – 1.17 (m, 6H, Rha, CH₃,A,C), 1.10 (m, 6H, Rha, CH₃,B,D). ¹³C NMR (151 MHz, chloroform-*d*) δ 165.37, 138.53, 137.82, 133.18, 133.15, 133.08, 132.92, 130.13, 129.78, 129.74, 128.44, 128.42, 128.39, 128.24, 128.21, 128.11, 128.03, 127.98, 127.92, 127.87, 127.85, 127.82, 127.80, 127.75, 127.71, 127.64, 127.61, 127.53, 127.52, 127.49, 127.44, 127.22, 126.45, 126.26, 126.11, 125.96, 125.85, 125.70, 125.65, 125.53, 100.98, 100.84, 98.76, 96.94, 80.33, 79.87, 79.73, 79.62, 76.12, 75.21, 74.75, 74.62, 72.94, 72.61, 72.22, 72.09, 69.07, 68.97, 68.34, 68.24, 67.62, 67.57, 51.25, 30.92, 28.96, 28.63, 23.36, 18.01, 17.99, 17.88, 17.70, 17.65. HR-MALDI-TOF/MS (m/z): calcd for C₉₃H₉₉N₃NaO₁₉⁺ [M+Na]⁺: 1584.6923; found, 1584.6928.



13

5-azidopentyl O-(4-O-benzyl-3-O-naphthylmethyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-naphthylmethyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (13). Hydrazine acetate (18 mg, 0.191 mmol) was added to a solution of **11** (64 mg, 0.038 mmol) in a mixture of DCM and MeOH (1/1, v/v, 10 mL). Stirring was continued until MALDI indicated the disappearance of the starting material (~3 h). The reaction mixture was diluted with DCM (30 mL), washed with water (3 \times 25 mL) and brine (25 mL), dried with anhydrous MgSO₄, and

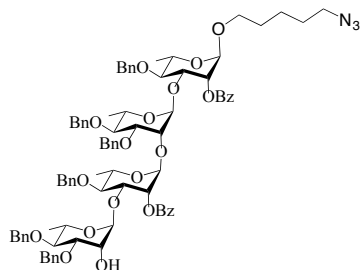
filtered. The crude product was purified by silica gel chromatography (hexane:EtOAc, 5:1 to 3:1) to provide pure compound **13** (55 mg, 92%). ^1H NMR (600 MHz, chloroform-*d*) δ 8.17 – 7.93 (m, 4H, arom. H, Bz), 7.91 – 6.97 (m, 40H, arom. H, Bn, Bz, Nap, Cbz), 5.59 – 5.48 (m, 1H, H-2,C), 5.35 – 5.24 (m, 1H, H-2,A), 5.10 (m, 2H, H-1,C, H-1,D), 4.99 (d, 1H, H-1,B), 4.87 – 4.26 (m, 12H, 4CH₂, Bn, 2xCH₂, Nap), 4.28 – 4.20 (m, 1H, H-3,C), 4.16 – 4.05 (m, 1H, H-3,A), 4.00 – 3.16 (m, 13H, H-5,A, H-2,D, H-2,B, H-3,D, H-5,D, H-5,C, H-2,B, H-3,B, H-5,B, H-4,C, H-4,A, H-4,B, H-4,D), 1.26 – 1.16 (m, 6H, Rha, CH₃,A,C), 1.12 (m, 6H, Rha, CH₃,B,D). ^{13}C NMR (151 MHz, chloroform-*d*) δ 165.37, 138.53, 137.82, 133.18, 133.15, 133.08, 132.92, 130.13, 129.78, 129.74, 128.44, 128.42, 128.39, 128.24, 128.21, 128.11, 128.03, 127.98, 127.92, 127.87, 127.85, 127.82, 127.80, 127.75, 127.71, 127.64, 127.61, 127.53, 127.52, 127.49, 127.44, 127.22, 126.45, 126.26, 126.11, 125.96, 125.85, 125.70, 125.65, 125.53, 100.98, 100.84, 98.76, 96.94, 80.33, 79.87, 79.73, 79.62, 76.12, 75.21, 74.75, 74.62, 72.94, 72.61, 72.22, 72.09, 69.07, 68.97, 68.34, 68.24, 67.62, 67.57, 51.25, 30.92, 28.96, 28.63, 23.36, 18.01, 17.99, 17.88, 17.70, 17.65. HR-MALDI-TOF/MS (m/z): calcd for C₉₃H₉₉N₃NaO₁₉⁺ [M+Na]⁺: 1584.6923; found, 1584.6928.



11

5-azidopentyl O-(4-O-benzyl-3-O-naphthylmethyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-naphthylmethyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-

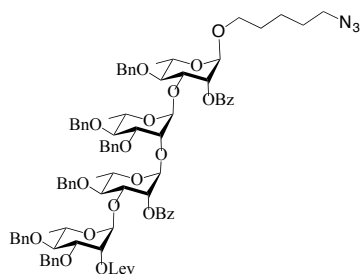
rhamnopyranoside (11). Compound **11** was derived from the optimal glycosylation method as described in the General Procedures. Yield: 92%. ^1H NMR (600 MHz, chloroform-*d*) δ 8.03 (m, 4H, arom. H, Bz), 7.90 – 6.89 (m, 40H, arom. H, Bn, Nap, Bz), 5.55 – 5.52 (m, 1H, H-2,C), 5.41 (m, 1H, H-2,D), 5.35 – 5.24 (m, 1H, H-2,A), 5.08 (m, 2H, H-1,C, H-1,D), 4.87 – 4.30 (m, 15H, H-1,A, 4xCH₂, Bn, 2xCH₂, Nap), 4.27 (m, 2H, H-3,C, H-3,A), 4.14 – 4.07 (m, 1H, H-5,A), 3.94 – 3.82 (m, 6H, H-3,D, H-5,D, H-5,C, H-2,B, H-3,B, H-2,C), 3.73 – 3.57 (m, 6H, H-5,B, CH₂, Linker), 3.56 – 3.20 (m, 14H, H-4,C, H-4,A, H-4,B, H-4,D, CH₂, Linker), 2.66 – 2.56 (m, 4H, CH₂-CH₂, Lev), 2.09 – 2.05 (m, 3H, CH₃, Lev), 1.24 – 1.16 (m, 6H, Rha, CH₃,A,C), 1.11 (m, 6H, Rha, CH₃,B,D). ^{13}C NMR (151 MHz, chloroform-*d*) δ 206.95, 165.85, 133.18, 132.84, 129.75, 128.41, 128.38, 128.17, 128.11, 127.97, 127.90, 127.86, 127.83, 127.59, 127.57, 127.52, 127.37, 126.49, 125.96, 125.86, 125.71, 125.55, 96.94, 79.65, 77.63, 74.67, 72.95, 71.49, 69.31, 68.57, 68.33, 67.62, 67.57, 51.26, 37.99, 30.92, 29.75, 29.69, 28.96, 28.62, 28.07, 23.36, 17.99, 17.71. HR-MALDI-TOF/MS (*m/z*): calcd for C₉₈H₁₀₅N₃NaO₂₁⁺ [M+Na]⁺: 1682.7232; found, 1682.7238.



29

5-azidopentyl O-(4-O-benzyl-3-O-benzyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-benzyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (29). Hydrazine acetate (19 mg, 0.207 mmol) was added to a solution of **28** (65 mg, 0.041 mmol) in a mixture of DCM and

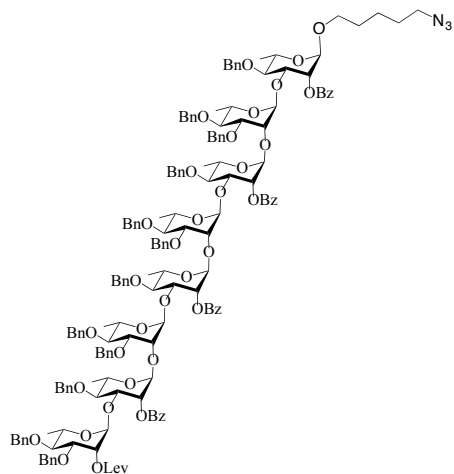
MeOH (1/1, v/v, 10 mL). Stirring was continued until MALDI indicated the disappearance of the starting material (~3 h). The reaction mixture was diluted with DCM (30 mL), washed with water (3 × 25 mL) and brine (25 mL), dried with anhydrous MgSO₄, and filtered. The crude product was purified by silica gel chromatography (hexane:EtOAc, 5:1 to 3:1, v/v) to provide pure compound **29** (57 mg, 94%). ¹H NMR (600 MHz, chloroform-*d*) δ 8.14 – 7.95 (m, 4H, arom. H, Bz), 7.73 – 6.92 (m, 36H, arom. H, Bn, Bz), 5.35 – 5.24 (m, 1H, H-2,A), 5.02 (m, 3H, H-1,B, H-1,D, H-1,C), 4.88 – 4.35 (m, 13H, H-1,A, 6xCH₂, Bn), 4.18 – 4.05 (m, 2H, H-3,A, H-3,C), 3.86 (m, 6H, H-2,B, H-2,D, H-3,B, H-3,D, CH₂, Linker), 3.77 – 3.58 (m, 6H, H-5,C, H-5,A, H-5,B, H-5,D, CH₂, Linker), 3.59 – 3.19 (m, 9H, CH₂, Linker, H-4,A, H-4,C, H-4,B, H-4,D), 1.25 – 1.19 (m, 6H, Rha, CH₃,A,C), 1.09 (m, 6H, Rha, CH₃,B,D). ¹³C NMR (151 MHz, chloroform-*d*) δ 206.96, 165.28, 129.78, 129.75, 128.46, 128.44, 128.42, 128.39, 128.19, 128.16, 128.09, 127.86, 127.83, 127.80, 127.72, 127.66, 127.64, 127.57, 127.45, 96.94, 79.56, 74.72, 72.02, 69.02, 68.17, 67.62, 51.25, 30.92, 28.95, 28.62, 18.03, 17.69. HR-MALDI-TOF/MS (m/z): calcd for C₈₅H₉₅N₃NaO₁₉⁺ [M+Na]⁺: 1484.6623; found, 1484.6628.



28

5-azidopentyl O-(4-O-benzyl-3-O-benzyl-2-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-benzyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (28). Compound **28** was derived from the optimal glycosylation method as described in the General Procedures.

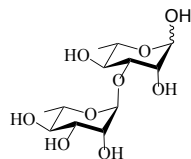
Yield: 78%. ^1H NMR (600 MHz, chloroform-*d*) δ 8.12 – 7.94 (m, 4H, arom. H, Bz), 7.67 – 7.00 (m, 36H, arom. H, Bn, Bz), 5.49 (m, 1H, H-2,C), 5.37 – 5.26 (m, 2H, H-2,D, H-2,A), 5.09 – 4.96 (m, 3H, H-1,B, H-1,D, H-1,C), 4.87 – 4.27 (m, 13H, H-1,A, 6xCH₂, Bn), 4.26 (m, 1H, H-3,C), 4.18 – 4.08 (m, 1H, H-3,A), 3.93 – 3.65 (m, 8H, H-2,B, H-5,C, H-5,A, H-3,B, H-3,D, CH₂, Linker), 3.66 – 3.46 (m, 4H, H-5,B, H-5,D, CH₂, Linker), 3.44 – 3.18 (m, 6H, H-4,A, H-4,C, H-4,B, H-4,D, CH₂, Linker), 2.69 – 2.54 (m, 4H, CH₂-CH₂, Lev), 2.10 (s, 3H, CH₃, Lev), 1.26 – 1.18 (m, 6H, Rha, CH₃,A,C), 1.14 – 1.06 (m, 6H, Rha, CH₃,B,D). ^{13}C NMR (151 MHz, chloroform-*d*) δ 133.18, 133.12, 129.94, 129.77, 129.75, 128.44, 128.42, 128.40, 128.28, 128.17, 128.13, 128.08, 127.94, 127.88, 127.83, 127.79, 127.77, 127.65, 127.54, 127.48, 127.38, 127.32, 127.22, 100.93, 98.96, 98.81, 96.95, 80.54, 80.00, 79.86, 79.63, 78.83, 78.43, 77.55, 77.10, 76.27, 76.16, 75.36, 74.63, 74.57, 72.96, 72.44, 72.07, 71.52, 69.34, 68.89, 68.52, 68.35, 67.63, 67.61, 51.25, 38.01, 28.95, 28.62, 28.07, 23.36, 18.03, 17.68. HR-MALDI-TOF/MS (*m/z*): calcd for $\text{C}_{90}\text{H}_{101}\text{N}_3\text{NaO}_{21}^+$ [$\text{M}+\text{Na}$] $^+$: 1582.6923; found, 1582.6928.



30

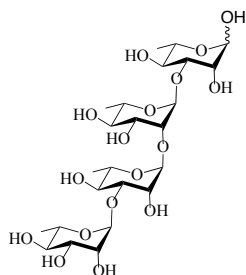
5-azidopentyl O-(4-O-benzyl-3-O-benzyl-2-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-benzyl- α -L-

rhamnopyranosyl)-(1→3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1→2)-(4-O-benzyl-3-O-benzyl- α -L-rhamnopyranosyl)-(1→3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1→2)-(4-O-benzyl-3-O-benzyl- α -L-rhamnopyranosyl)-(1→3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (**30**). Compound **30** was derived from the optimal glycosylation methods as described in detail in “4+4-Bn” section. Yield: 63%. ^1H NMR (600 MHz, chloroform-*d*) δ 8.14 – 7.91 (m, 8H, arom. H, Bz), 7.69 – 6.86 (m, 72H, arom. H, Bn, Bz), 5.48 (m, 3H, H-2,C, H-2,E, H-2,G), 5.41 – 5.20 (m, 2H, H-2,H, H-2,A), 5.10 – 4.90 (m, 8H, H-1,B-H), 4.85 – 4.26 (m, 26H, 12xCH₂, Bn, CH₂, Linker), 4.18 (m, 4H, H-3,C, H-3,E, H-3,G, H-1,C), 3.90 – 3.57 (m, 16H, H-5,H, H-2,B, H-2,D, H-2,F, H-3,B, H-3,D, H-3,F, H-3,H, H-5,B, H-5,D, H-5,F, H-5,C, H-5,E, H-5,G, CH₂, Linker), 3.52 – 3.19 (m, 10H, H-4,C, H-4,E, H-4,G, H-4,A, H-4,B, H-4,D, H-4,F, H-4,H, CH₂, Linker), 2.60 (m, 4H, CH₂-CH₂, Lev), 2.12 – 2.01 (m, 3H, CH₃, Lev), 1.29 – 1.04 (m, 24H, Rha, CH₃,A-H). HR-MALDI-TOF/MS (*m/z*): calcd for C₁₇₈H₁₈₉N₃NaO₃₉⁺ [M+Na]⁺: 3015.2923; found, 3015.2927.



33

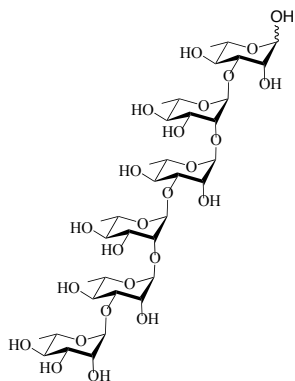
α -L-rhamnosyl-(1→3)- α -L-rhamnopyranose (**33**). ^1H NMR (500 MHz, deuterium oxide) δ 4.95 (d, *J* = 1.9 Hz, 1H, H-1,A), 4.92 (d, *J* = 1.7 Hz, 1H, H-1,B), 3.94 (m, 1H, H-2,B), 3.87 (m, 1H, H-2,A), 3.82 – 3.74 (m, 1H, H-5,A), 3.74 – 3.65 (m, 3H, H-5,B, H-3,B, H-3,A), 3.36 (m, 2H, H-4,A, H-4,B), 1.18 – 1.13 (m, 6H, Rha, CH₃,A,B). HR-MALDI-TOF/MS (*m/z*): calcd for C₃₆H₆₂NaO₂₅⁺ [M+Na]⁺: 917.3612; found, 917.3617.



34

α -L-rhamnosyl-(1 \rightarrow 3)- α -L-rhamnosyl-(1 \rightarrow 2)- α -L-rhamnosyl-(1 \rightarrow 3) α -L-rhamnopyranose

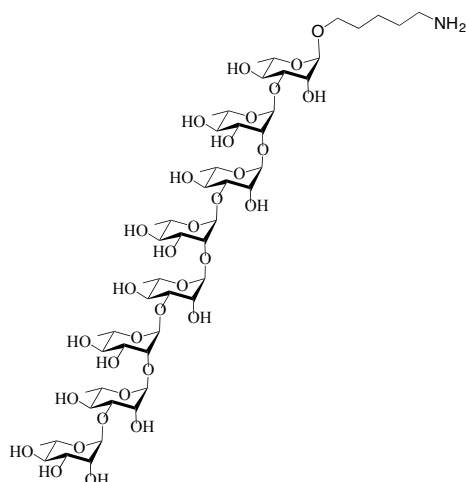
(34). ^1H NMR (500 MHz, deuterium oxide) δ 5.09 (d, $J = 1.7$ Hz, 1H, H-1,B), 4.96 (d, $J = 1.8$ Hz, 1H, H-1,A), 4.92 (d, $J = 1.7$ Hz, 1H, H-1,D), 4.83 (d, $J = 1.9$ Hz, 1H, H-1,C), 4.04 (m, 1H, H-2,C), 3.96 – 3.92 (m, 2H, H-2,B, H-2,D), 3.86 (m, 1H, H-2,A), 3.84 – 3.75 (m, 3H, H-5,B, H-5,D, H-3,C), 3.75 – 3.67 (m, 3H, H-3,A, H-3,B, H-3,C), 3.66 – 3.59 (m, 1H, H-5,A), 3.44 – 3.25 (m, 5H, H-4,A, H-4,B, H-4,C, H-4,D, H-5,C), 1.22 – 1.11 (m, 12H, Rha, CH_3 ,A-D). HR-MALDI-TOF/MS (m/z): calcd for $\text{C}_{24}\text{H}_{42}\text{NaO}_{17}^+$ [$\text{M}+\text{Na}$] $^+$: 625.2423; found, 625.2428.



35

α -L-rhamnosyl-(1 \rightarrow 3)- α -L-rhamnosyl-(1 \rightarrow 2)- α -L-rhamnosyl-(1 \rightarrow 3)- α -L-rhamnosyl-

(1 \rightarrow 2)- α -L-rhamnosyl-(1 \rightarrow 3) α -L-rhamnopyranose (35). ^1H NMR (800 MHz, deuterium oxide) δ 5.06 (m, 2H, H-1,B, H-1,D), 4.90 (m, 2H, H-1,C, H-1,E), 4.02 (m, 2H, H-2,C, H-2,E), 3.93 (m, 3H, H-2,B, H-2,D, H-2,F), 3.84 (m, 1H, H-2,A), 3.81 (m, 3H, H-3,A, H-3,C, H-3,E),



32

5-aminopentyl *O*- α -L-rhamnosyl-(1 \rightarrow 3)- α -L-rhamnosyl-(1 \rightarrow 2)- α -L-rhamnosyl-(1 \rightarrow 3)- α -L-rhamnosyl-(1 \rightarrow 2)- α -L-rhamnosyl-(1 \rightarrow 3) α -L-rhamnopyranoside (32).

¹H NMR (800 MHz, deuterium oxide) δ 5.07 – 5.03 (m, 3H, H-1,B, H-1,D, H-1,F), 4.90 (d, 1H, H-1,H), 4.83 – 4.79 (m, 3H, H-1,C, H-1,E, H-1,G), 4.02 (m, 3H, H-2,C, H-2,E, H-2,G), 3.93 (m, 4H, H-2,B, H-2,D, H-2,F, H-2,H), 3.83 (m, 4H, H-2,A, H-3,B, H-3,D, H-3,F), 3.75 – 3.49 (m, 10H, H-5,A-H, CH₂, Linker), 3.46 – 3.29 (m, 8H, H-4,A-H), 2.86 (m, 2H, CH₂, Linker), 1.53 (m, 4H, 2 \times CH₂, Linker), 1.31 (m, 2H, CH₂, Linker), 1.15 (m, 24H, Rha, CH₃,A-H). HR-MALDI-TOF/MS (m/z): calcd for C₅₃H₉₃NNaO₃₃⁺ [M+Na]⁺: 1294.5612; found, 1294.5617.

References

1) Nina S.; Jason C.; Kirsten K.; Anna H.; Ramy A.; Ana K.; Leo L.; Evelien B.; Mark D.; Gordon D.; Fan Z.; Samira D.; Laura S.; Jennifer G.; Madeleine C.; Joseph M.; Julia H.; Bernd L.; Suzan R.; Richard M.; Mark W.; Sanford S.; Patrick S.; Biswa C.; Victor N. The classical lancefield antigen of Group A Streptococcus is a virulence determinant with implications for vaccine design. *Cell Host & Microbe*. **2014**, 15, 729–740.

- 2) Kaplan M.; Bolande R.; Rakita L.; Blair J. Presence of bound immunoglobulins and complement in the myocardium in acute rheumatic fever association with cardiac failure. *The New England Journal of Medicine*. **1964**, 13, 637–45.
- 3) Lozano R.; Naghavi M.; Foreman K.; Lim S.; Shibuya K.; Aboyans V. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. **2012**, 380, 2095–128.
- 4) Andrew S.; Jonathan C.; James D.; John F.; Michael G.; Luiza G.; Nicole M.; Kim M.; Florian S.; Pierre S. Status of research and development of vaccines for *Streptococcus pyogenes*. *Vaccine*. **2016**, 26, 2953–2958.
- 5) Michon F.; Moore L.; Kim J.; Blake S.; Auzanneau I.; Johnston D.; Johnson A.; Pinto M. Doubly branched hexasaccharide epitope on the cell wall polysaccharide of group A streptococci recognized by human and rabbit antisera. *Infectious Immunology*. **2005**, 73, 6383–6389.
- 6) Anna K.; Immaculada M.; Francesco B.; Maria R.; Guido G.; Giuliano B.; Emiliano C.; Daniela P.; Erwin S.; Emilia C.; Paola F.; Daniele C.; Roberto A.; Vittoria P.; David S.; Sabrina C.; Giada B.; Marilena G.; William C.; Stewart C.; John P.; Peter H. S.; Rino R.; Paolo C. Evaluation of a Group A *Streptococcus* synthetic oligosaccharide as vaccine candidate. *Vaccine*. **2010**, 1, 104–114.
- 7) Goldstein I.; Rebeyrotte P.; Parlebas J.; Halpern B. Isolation from heart valves of glycopeptides which share immunological properties with *Streptococcus haemolyticus* group A polysaccharides. *Nature*. **1968**, 219, 866–868.
- 8) Shulman T.; Ayoub M.; Victorica E.; Gessner H.; Tamer F.; Hernandez A. Differences in antibody response to streptococcal antigens in children with rheumatic and non-rheumatic mitral valve disease. *Circulation*. **1974**, 50, 1244–1251.

- 9) Appleton S.; Victorica E.; Tamer D.; Ayoub M. Specificity of persistence of antibody to the streptococcal group A carbohydrate in rheumatic valvular heart disease. *Journal of Laboratory and Clinical Medicine*. **1985**, 105, 114–119.
- 10) Galvin E.; Hemric E.; Ward K.; Cunningham W. Cytotoxic mAb from rheumatic carditis recognizes heart valves and laminin. *The Journal of Clinical Investigation*. **2000**, 106, 217–224.
- 11) Zaihong Z; Peng W; Ning D; Gaopeng S; Yingxia L. Total synthesis of cleistetroside-2, partially acetylated dodecanyl tetra-rhamnoside derivative isolated from *Cleistopholis patens* and *Cleistopholis glauca*. *Carbohydrate Research*. **2007**, 342, 9, 1159-1168.
- 12) Vince P. Synthesis of Glycoconjugate Vaccines against *Shigella dysenteriae* Type 1. *Journal of Organic Chemistry*. **1998**, 63, 5983–5999.
- 13) Anne V.; Niels R.; Herman O.; Gijs V.; Jeroen C. The Cyanopivaloyl Ester: A New Protecting Group in the Assembly of Oligorhamnans. *European Journal of Organic Chemistry*. **2016**, 31, 5282-5293.

CHAPTER 3

CHEMICAL SYNTHESIS AND EPITOPE MAPPING OF *GROUP A STREPTOCOCCUS* OLIGOSACCHARIDE ANTIGENS

Huzi Sun, Margreet A. Wolfert, Geert-Jan Boons.* To be submitted to J. Am. Chem. Soc.

Abstract

Group A *Streptococcus* (GAS) infections are a global healthcare burden and cause millions of deaths each year. GAS is the sole species of Lancefield group A and all GAS serotypes express the Lancefield group A carbohydrate (GAC), which is comprised of a polyrhamnose backbone with an immunodominant *N*-acetylglucosamine (GlcNAc) side chain. The major challenge for the development of a safe and efficacious carbohydrate-based vaccine includes the potential of the GAC GlcNAc side chain to provoke cross-reactive antibodies relevant to the immunopathogenesis of rheumatic fever. Lack of structure-activity relationships for the cross-reactive antibody makes it challenging to understand the pathogenesis of the GAS autoimmunity at a molecular level and dramatically complicates the design and the development of safe and effective GAS vaccine candidates. We describe here a synthetic methodology that can rapidly provide a library of well-defined GAC oligosaccharides with different GlcNAc side chain variations and chain length. It is based on the use of a “key disaccharide” to assemble GAC oligosaccharides with different chain variations and length modularly. Different types of GAS antibodies, including the anti-polyrhamnose antibody, the anti-*Streptococcus* Group A antibody, the anti-*Streptococcus pyogenes* GAC antibody, and the rabbit serum derived from immunization of the octa-polyrhamnose KLH-glycoconjugate were investigated by the GAC-microarray. The binding study showed that each antibody recognized multiple compounds and exhibited distinct structure–binding relationships. The GAC-microarray data made it possible to investigate and validate the influences of different length and side-chain variations on the structure–binding relationships with type-specific antibodies. Although the library does not cover larger oligosaccharides (>15 mono components), the array data support a notion that side-chain variation and length can have a significant influence on the structure–binding relationships and

provide valuable insights to the cross-reactivity associated with the autoimmunity of the GAS infections.

Introduction

Group A Streptococcus (GAS) infections cause at least 517,000 deaths each year according to recent population-based estimates, and it is a leading healthcare problem throughout the world.¹ GAS causes a wide range of diseases, ranging from asymptomatic colonization, uncomplicated pharyngeal, skin infections, some forms of pneumonia, sepsis, necrotizing fasciitis, and toxic shock syndrome.²⁻⁵ GAS infections may lead to delayed sequela as rheumatic fever, rheumatic carditis, heart valve disease, and acute glomerulonephritis,⁶⁻¹⁰ which are caused by an autoimmune reaction to *Group A streptococci*.¹¹ Rheumatic fever is the primary cause of acquired heart disease in children, adolescents and young adults, which causes at least 350,000 deaths each year.¹² Even though the global demands for an effective GAS vaccine remains high, there is currently no safe and efficacious commercial vaccine against GAS infection.

GAS vaccines can be broadly divided into M protein-based and non-M protein-based vaccines. M protein-based vaccines are well developed. However, GAS carbohydrate-based vaccines are less developed compared to M protein-based vaccines, and research groups from all over the world are working on this issue to create an efficacious and safe GAS vaccine.¹³ All GAS serotypes express the Lancefield group A carbohydrate (GAC), which consists of $\{\rightarrow 2\}[\beta\text{-D-GlcNAc}(1\rightarrow 3)]\alpha\text{-L-Rha}(1\rightarrow 3)\}\alpha\text{-L-Rha}(1\text{-})_n$ repeating units. This structure could serve as a common antigen for the development of a universal GAS vaccine. Pinto and Costantino's groups have evaluated the hexamer core antigenic GAS glycoconjugate vaccine candidates and various length of synthetic GAC compounds as the glycoconjugate vaccine antigens.^{14,15} They demonstrated that the synthetic glycoconjugate vaccine candidates could provide a broad range

of protection against systemic GAS challenges. However, the autoreactivity of the GlcNAc-recognizing antibodies against human tissues has been raised by several groups.^{16,17,18} Cunningham *et al.* derived anti-GlcNAc monoclonal antibodies cross-reactive for heart or brain tissue from patients with rheumatic fever.¹⁹ Nizet and co-workers identified the genes responsible for GAC biosynthesis.²⁰ However, the recognition motif of the cross-reactive antibodies remains unclear. Lack of structure-activity relationships for the cross-reactive antibody makes it challenging to understand the pathogenesis of the GAS autoimmunity at a molecular level and greatly complicates the design and the development of a safe and effective GAS vaccine candidates. The major challenge for addressing this issue lies in the difficulties to obtain structurally well-defined GAC compounds with different chain length and side-chain variations. The problems associated with the compounds isolated from Δ GacI mutant can be addressed by the powerful toolbox of synthetic chemistry, which can be utilized to design and generate synthetic pure, homogeneous, and well-defined oligosaccharides to evaluate how particular structural features, such as lengths and side-chain variations, influence carbohydrate immunogenicity, and structure-activity relationships.

Glycoarrays are emerging as a key glycomics technology. It requires only small amounts of oligosaccharides for fast and systematic evaluation of carbohydrate-protein interactions.^{21,22} However, the difficulties of preparing a library of oligosaccharide antigens hamper the development of the glycan microarray. Here, we describe a synthetic methodology that can provide a library of the GAC oligosaccharide antigens with different GlcNAc side chain variations by a modular approach. Those synthetic oligosaccharides can be utilized for glycan-microarray to investigate the recognition motif of the oligosaccharide antigens with various antibodies. The GAC-microarray was employed to investigate different types of GAS antibodies,

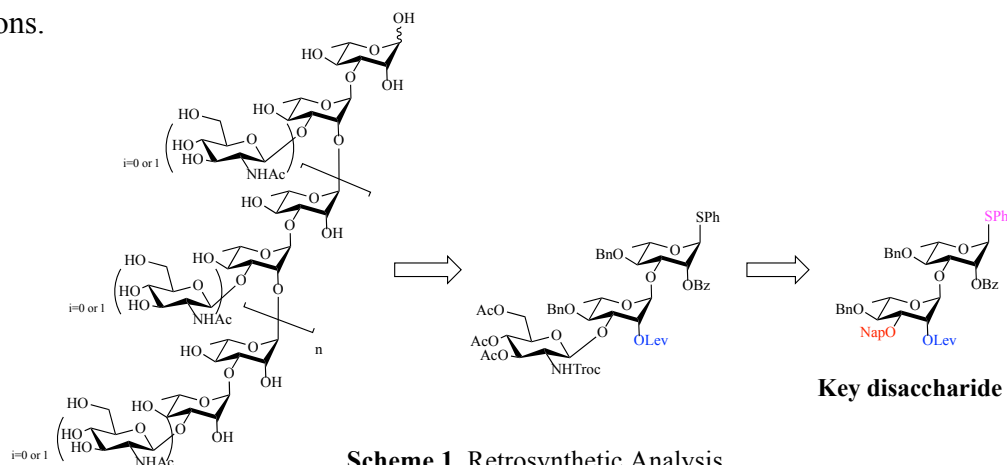
including the anti-polyrhamnose antibody, the anti-Streptococcus Group A antibody, the anti-Streptococcus pyogenes GAC antibody, and the rabbit serum derived from immunization of the octa-polyrhamnose KLH-glycoconjugate. The binding study showed that each antibody recognized multiple compounds and exhibited distinct structure–binding relationships. The GAC-microarray data made it possible to investigate and validate the influences of different length and side-chain variations on the structure–binding relationships with type-specific antibodies.

An attractive approach to determine structure–binding relationships between the antigen and antibodies can be realized by a customized library of GAC oligosaccharides. An important motivation to prepare the GAC oligosaccharides library with different backbone lengths and GlcNAc side chain variations is that the potential structural diversity can provide essential insights into the antibody recognition and the issue related with the cross-reactivity of the autoimmune antibody due to GAS infections. Notably, in the natural GAS bacteria system, the expressed GAC oligosaccharides may not contain the GlcNAc side chains on its polyrhamnose backbone consistently, resulting in the “patch” structure of the GAC oligosaccharide. This structure could have distinct impacts on the immune response of the host immune system, which may lead to mismatched immune responses when vaccinated with the full-GlcNAc GAS oligosaccharide formulation. We envisage that a customized library of GAC oligosaccharides library with different backbone length and GlcNAc side chain variations is expected to provide essential insights for structure–binding relationships of a variety of type-specific antibodies. The GAC microarray binding results will provide significantly important answers to the questions that will be of great interest for the GAS vaccine community.

Discussions and results

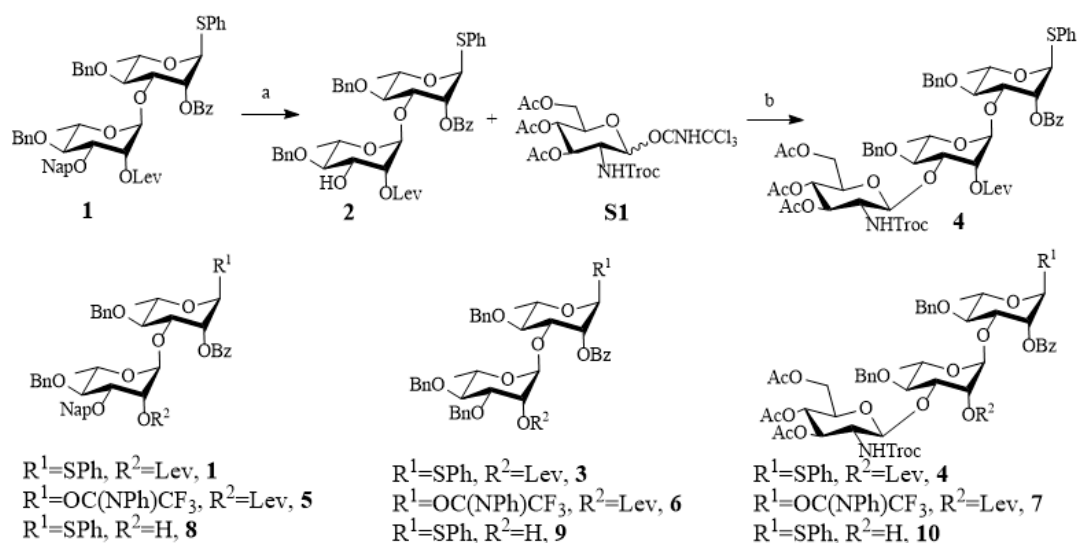
Chemical synthesis of GAC oligosaccharides library with different GlcNAc side chain variations

We envisaged that the designed retrosynthetic strategy employing a selectively protected key disaccharide can provide a highly convergent and customized approach to access a variety of GAC oligosaccharides as shown in **Scheme 1**. A convenient set of orthogonal protecting groups, including Levulinic ester (Lev), 2-Naphthylmethyl ether (Nap) and thiophenyl ether (SPh), were employed for the glycosyl donor/acceptor formation and GlcNAc side chain installation. We have previously described an *in situ* “4+4” polymerization by which we can obtain a library of GAC polyrhannose oligosaccharides, and we have shown the screening and optimization of the key disaccharide in the previous work. Here we describe the chemical synthesis of GAC oligosaccharides with different GlcNAc side chain variations. We will apply ester protecting groups at C-2 to assist the α -selectivity by utilizing neighboring group participating effects, and use the Lev ester as an orthogonal protecting group to differentiate the bottom C-2. We will employ the Nap orthogonal protecting group at C-3 position to further assemble the trisaccharide repeating units and the library GAC oligosaccharides with different GlcNAc side chain variations.



Firstly, we investigated a divergent strategy to assemble the GAC oligosaccharides (as shown in SI. **Scheme 1**.) Compound **S2** reacted with DDQ in chloroform:PBS buffer (9:1) yielded compound **S3** in 92% excellent yield. A TMSOTf promoted glycosylation of compound **S3** with compound **S1** yielded expected compounds **S4**, **S5**, **S6** with different GlcNAc side chain variations. However, the overall conversion of the glycosylation is moderate as shown by TLC and High-Resolution MALDI-MS (see **SI**). The purification of these compounds posed a significant problem, as most of the compounds decomposed on silica gel, which suggests that those compounds are quite acid sensitive. Besides, due to the small mass differences of these compounds, it could not be efficiently separated by LH20 column. The scalability of the reaction is another obstacle for this approach. Therefore, this divergent strategy is not suitable for the assembly of the GAC oligosaccharides with different GlcNAc side chain variations.

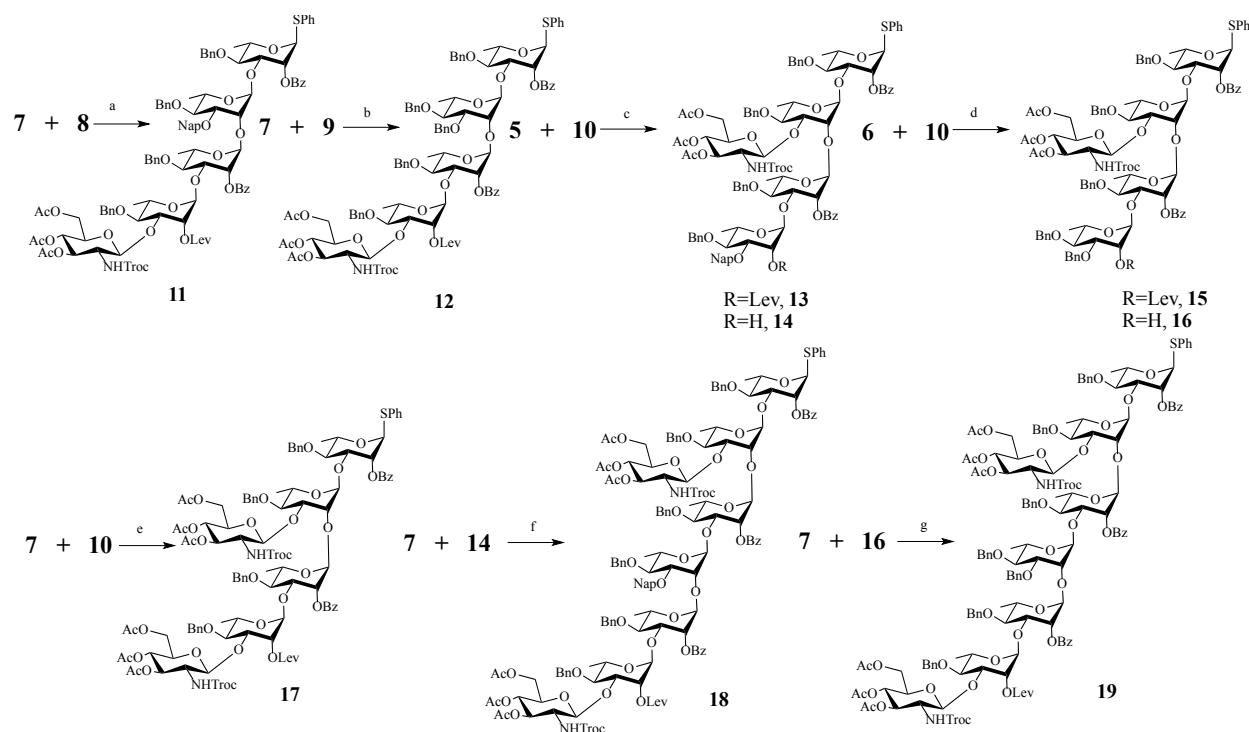
Then, we decided to investigate a modular synthetic strategy to assemble the GAC oligosaccharides with different GlcNAc side chain variations as shown in **Scheme 3**. The building block synthesis is described in **Scheme 2**. The deprotection of the Nap of compound **1** yielded compound **2** in 81% good yield. A TMSOTf promoted glycosylation of compound **2** with compound **S1** yielded the trisaccharide **4** in 95% excellent yield. Modular disaccharides and trisaccharide building blocks **1-10** were prepared by standard manipulations or were described in the previous work.



Scheme 2. Synthesis of the trisaccharide **4** and modular disaccharides and trisaccharide building blocks. Reagents and conditions: a) DDQ 2.3eq, 0°C, Chloroform/PBS buffer(9:1), 81%. b) TMSOTf 0.1eq, -40~0°C, 4A AW MS, Dry DCM, 95%

With these building blocks in hand, we decided to further investigate the influence of the C-3 protecting groups or C-3 GlcNAc on the glycosylation process. A TBSOTf promoted glycosylation of compound **7** with compound **8** yielded the pentasaccharide **11** in 22% poor yield. A TBSOTf promoted glycosylation of compound **7** with compound **9** yielded the pentasaccharide **12** in 31% poor yield. However, a TMSOTf promoted glycosylation of compound **5** with compound **10** yielded the pentasaccharide **13** in 75% good yield. A TBSOTf promoted glycosylation of compound **6** with compound **10** yielded the pentasaccharide **12** in 82% good yield. A TBSOTf promoted glycosylation of compound **7** with compound **10** yielded the hexasaccharide **17** in 96% excellent yield. Comparing these results, it shows that when bulky aromatic protecting groups such as Nap or Bn are present at the C-3 position, the adjacent C-2 hydroxyl will be blocked, and the reactivity of the C-2 hydroxyl in the glycosylation process will be dramatically decreased, as the aromatic ring can rotate around the O-CH₂ bond and block the incoming glycosyl donors. However, when GlcNAc is present at the C-3 position, the adjacent

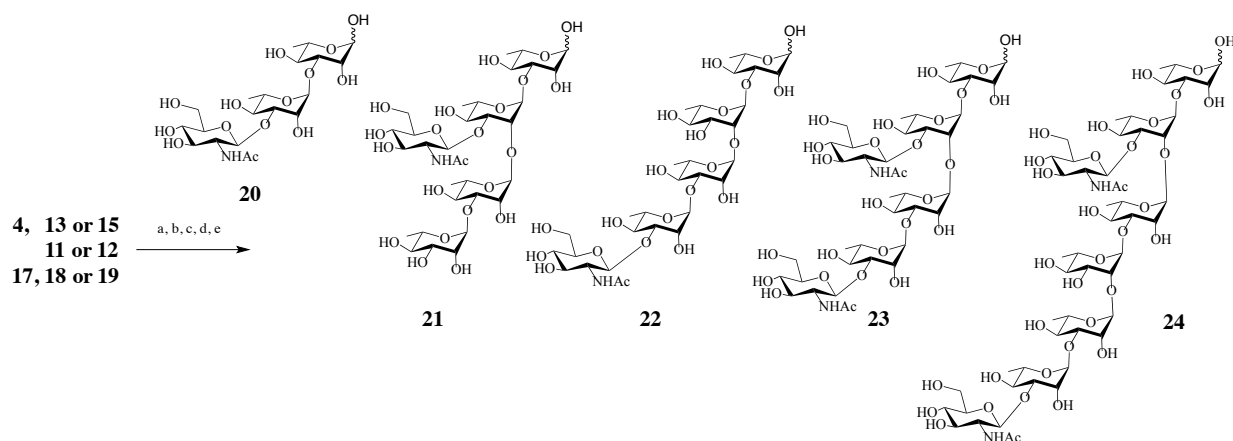
C-2 hydroxyl will not be blocked at all. The reactivity of the C-2 hydroxyl in the glycosylation process is excellent since the glycosidic bond between the GlcNAc and the rhamnose is more rigid, and the GlcNAc cannot rotate around the glycosidic bond. Therefore, the incoming glycosyl donors will not be blocked in the glycosylation process. The overall reactivity of the C-2 hydroxyl is in the order of GlcNAc>Bn>Nap when these protecting groups or carbohydrate are at the adjacent C-3 position. Then, deprotection of the Lev on compounds **13** and **15** by hydrazine acetate yielded compound **14** and **16** in excellent yield. A TBSOTf promoted glycosylation of compound **7** with compound **14** yielded the octasaccharide **18** in 23% poor yield. Finally, a TBSOTf promoted glycosylation of compound **7** with compound **16** yielded the octasaccharide **19** in 43% acceptable yield in DCM and 37% acceptable yield in toluene. The low reactivity of the final octa-assembly glycosylation is due to the presence of the C-2 Nap or Bn bulky aromatic protecting groups, which blocks the incoming glycosyl donors as we have just explained before.



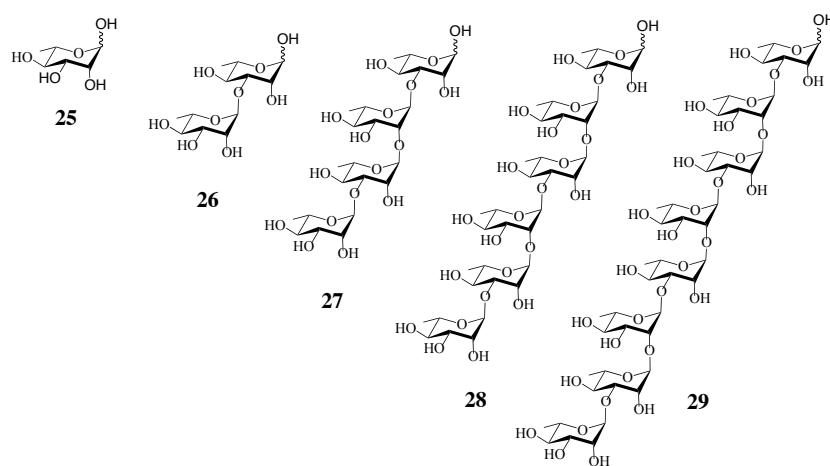
Scheme 3. Modular synthesis of the GAS oligosaccharide antigen library. Reagents and conditions: a) TBSOTf 0.15eq, -40~0°C, 4A AW MS, Dry DCM, 22%. b) TBSOTf 0.15eq, -40~0°C, 4A AW MS, Dry DCM, 31%. c) TMSOTf 0.3eq, -40~0°C, 4A AW MS, Dry DCM, 75%. d) TBSOTf 0.3eq, -40~0°C, 4A AW MS, Dry DCM, 82%. e) TBSOTf 0.15eq, -40~0°C, 4A AW MS, Dry DCM, 96%. f) TBSOTf 0.15eq, -40~0°C, 4A AW MS, Dry DCM, 23%. g) TBSOTf 0.15eq, -40~0°C, 4A AW MS, Dry DCM, 43%; TBSOTf 0.15eq, -40~0°C, 4A AW MS, Dry Toluene, 37%

Deprotection sequence is crucial for these compounds as shown in **Scheme 4**. Firstly, deprotection of the Troc by Zn reduction unmasked the amine group, which can be acetylated by acetyl anhydride subsequently. Secondly, deprotection of the esters by sodium methoxide in methanol must be conducted in this step to avoid β -elimination. Thirdly, the anomeric thioglycoside was removed by NBS in acetone:10 mM PBS buffer (6:1). Lastly, deprotection of the benzyl ether by 10% Pd/C in THF/H₂O, and subsequent purification by P2 Bio Gel yielded pure final compounds in good yield. The pure compounds could be used for the structure-activity analysis, the glycan microarray, and ELISA antibody binding studies. Two-dimensional NMR experiments combined with high-resolution MALDI-MS confirmed the structural integrity of the compound. *Compound 35*, 36* confirmed by HR-MALDI MS only due to limited quantity. The J_{H-H} coupling constants of non-reducing end anomeric protons were all between 169 and 170 Hz, indicating all α -

glycosidic linkages. The J_{H-H} coupling constants of the reducing end anomeric protons show two separated peaks at 160Hz and 170Hz, which indicated β and α isomers respectively.



Scheme 4. Deprotection of the GAS oligosaccharide antigen library. Reagents and conditions: a) Zn, HOAc, DCM. b) Ac₂O, DMAP, pyridine. c) NaOMe/HOMe. d) NBS, Acetone: PBS buffer(6:1). e) Pd/C, THF/H₂O.



Scheme 5. The polyrhmannose antigen library

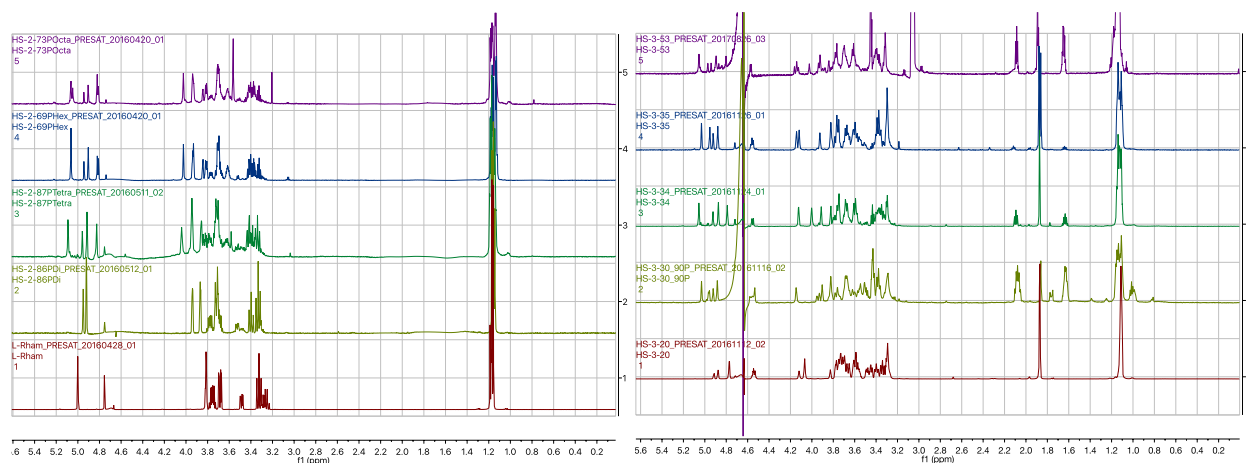
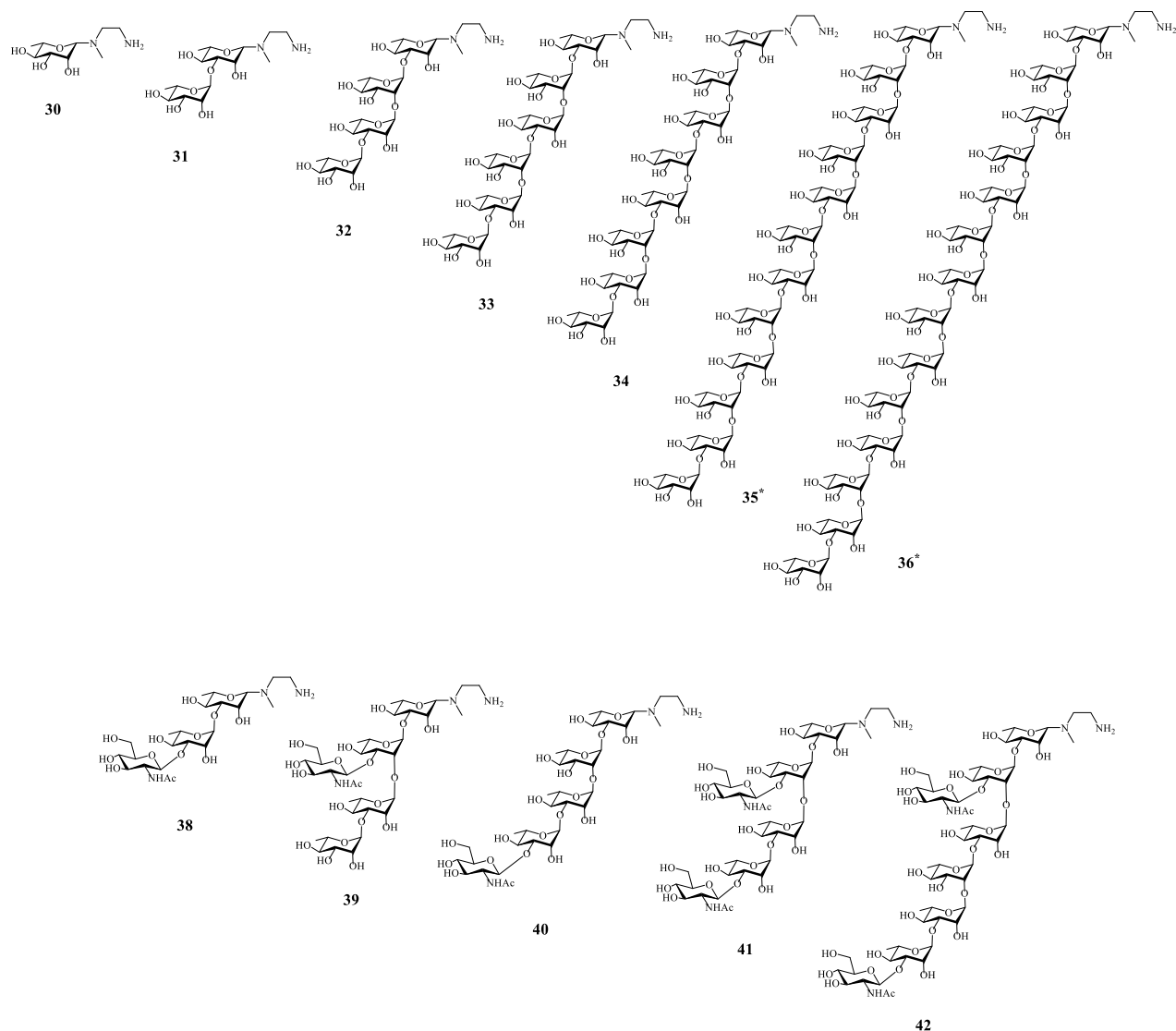


Figure 1. NMR Spectra of deprotected GAC library. Left: Polyrhmannose library; Right: GAC library with GlcNAc variations.

Glycan microarray study



Scheme 6. The customized GAC oligosaccharide library. *Indicates that the structure was confirmed by HR-MALDI MS only due to the limited quantity

Glycan microarrays containing synthetic GAC oligosaccharides with various length and side chain variations can be used to map the glycotopes that are recognized by antibodies. By the treatment with 2-[(methylamino)oxy]ethanamine, the synthetic GAC oligosaccharides were modified with an amino-containing linker, which can be printed on *N*-hydroxysuccinimide (NHS)-activated glass slides. The microarray printing was conducted by printing each sample as replicates of 2 (1 mM in a sodium phosphate (50 mM), pH 9.0 buffer). Then slides were

incubated in a saturated NaCl chamber overnight. The unreacted esters were quenched with ethanolamine (100 mM) and slides are ready for screening experiments.

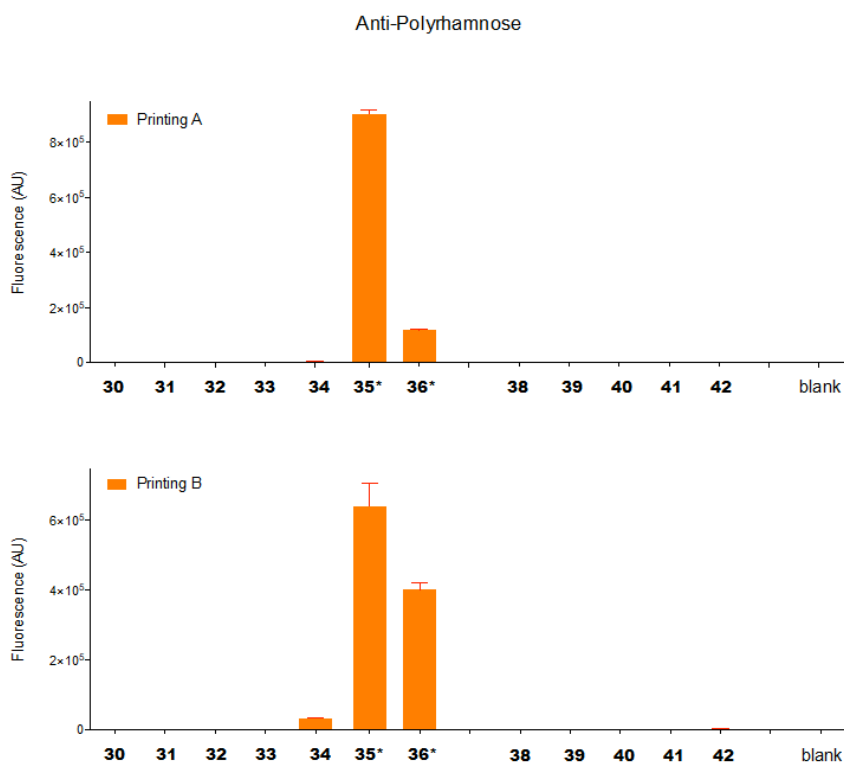


Figure 2. Glycan Microarray results for the anti-polyrhamnose antibody.

We initially screened the synthetic GAC library against the anti-polyrhamnose antibody (**Figure. 2**). The antibody recognizes only long-chain polyrhamnose oligosaccharides, and no recognition was found by synthetic GAC oligosaccharides with GlcNAc side chains (**38-42**). Notably, the anti-polyrhamnose antibody only recognizes **35*** and **36*** with significant binding strength, and smaller length structure, comparatively speaking, shows no or very weak binding towards the anti-polyrhamnose antibody. The binding strength maximized at the dodecasaccharide **35***, and further decreased at the tetradecasaccharide **36***. This result can be explained by the helical structure of the long chain polyrhamnose oligosaccharide, and it suggests that the antibody interacts with the helical backbone and it may require significant

polyrhamnose backbone length (>10, optimal at 12, which is the helical structure with 2 cyclized repetition) to have strong binding strength.

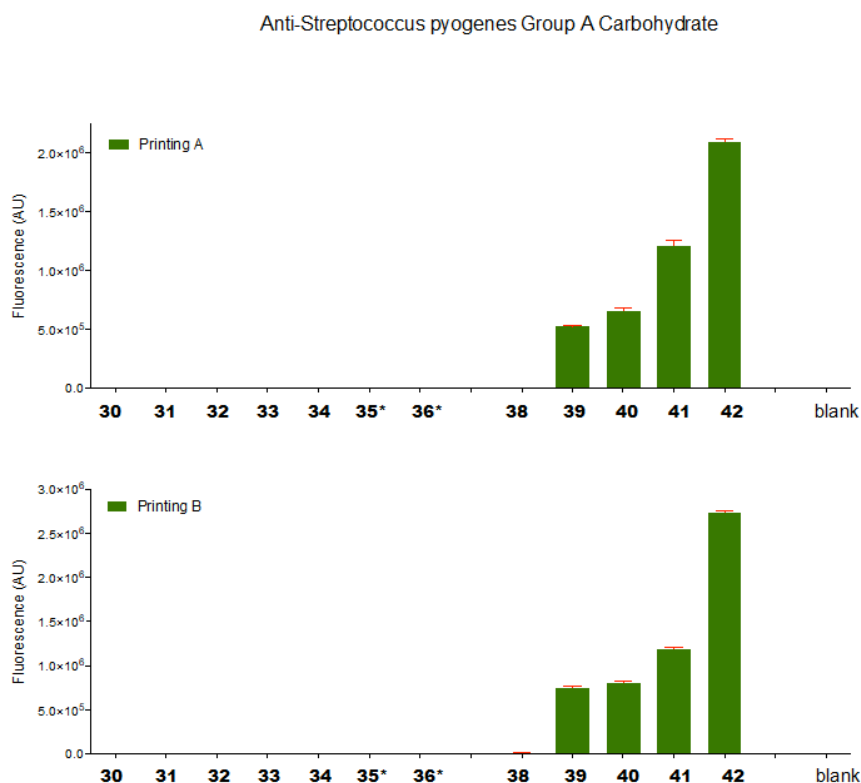


Figure 3. Glycan Microarray results for the anti-streptococcus pyogenes GAC antibody.

Secondly, we used the anti-streptococcus pyogenes GAC antibody to screen the synthetic GAC library (**Figure. 3**). The recognition results showed that the antibody specifically recognizes the synthetic GAC oligosaccharides with the GlcNAc side chains (**39-42**), and no recognition was presented for polyrhamnose oligosaccharides (**30-36***). Interestingly, the antibody has no binding towards the trisaccharide **38**, and similar intermediate levels of binding strength for both pentasaccharides (**39** and **40**). Notably, the antibody shows a strong binding strength towards hexasaccharide **41** and the maximum binding strength towards the octasaccharide **42**. These results suggest that the anti-streptococcus pyogenes GAC antibody

requires the presence of GlcNAc side chains on the periphery of the polyrhamnose backbone for the recognition, and it may require multiple GlcNAc side chains to fit in binding pockets. More specifically, the maximum binding strength was achieved by the octasaccharide **42**, which suggests the distance between the two GlcNAc side chains plays an essential role in the binding process. A longer distance between the two GlcNAc side chains has different angles and different distances from the shorter hexasaccharide **41**, as results of which the octasaccharide **42** may fits better in binding pockets of the antibody.

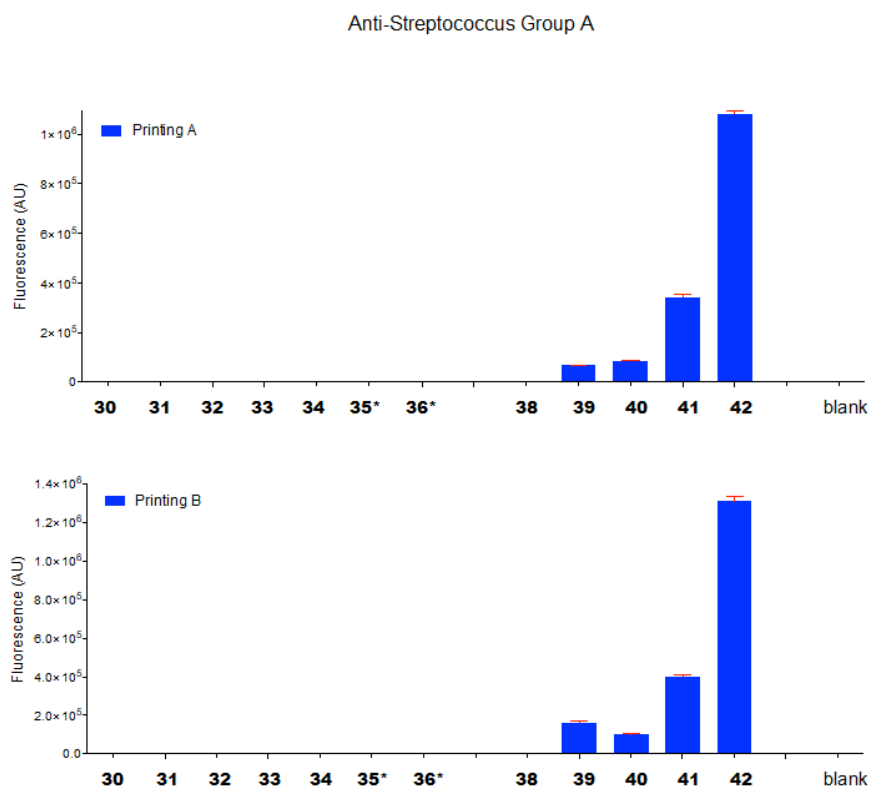


Figure 4. Glycan Microarray results for the anti-Streptococcus Group A antibody

Lastly, we conducted the screening against the anti-Streptococcus Group A antibody with the synthetic GAC library (**Figure. 4**). The binding results showed a similar binding pattern with the binding results of the anti-streptococcus pyogenes GAC antibody. However, the binding

strengths are comparatively 2-4 times less than the previous one. Specifically, the antibody only recognizes the synthetic GAC oligosaccharides with the GlcNAc side chains (**39-42**), and no recognition was presented for polyrhamnose oligosaccharides (**30-36***). The antibody has a no binding towards the trisaccharide **38**, and similar low levels of binding strength for both pentasaccharides (**39** and **40**). The binding strength towards the hexasaccharide **41** is an intermediate level, and the maximum binding strength was achieved by the octasaccharide **42**. Comparing the binding strength of the hexasaccharide **41** and the octasaccharide **42**, the ratio of the binding strength is 1:3, which is less than the previous 1:2 ratio. Comparing the pentasaccharides **39**, **40** and the hexasaccharide **41**, the ratio of the binding strength is about 1:3, which is less than the previous 2:3. These results suggest that the anti-streptococcus pyogene Group A antibody interacts similarly with the synthetic GAC oligosaccharides when comparing with the anti-streptococcus pyogenes GAC antibody, the differences lie in the strength of the interactions. Both antibodies exhibited a strong binding preference for the larger octasaccharide **42**, as the larger distance between the two GlcNAc provides different angles and distance to fit better in the binding pockets of these antibodies.

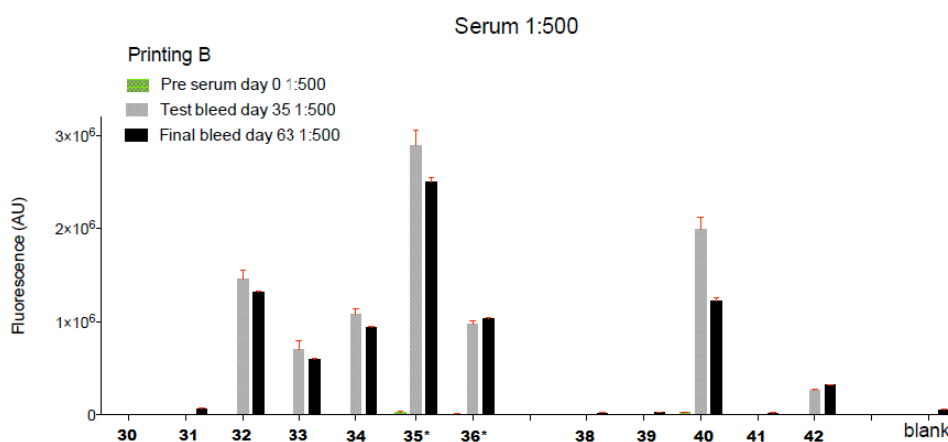


Figure 5. Glycan Microarray results for the rabbit serum derived from immunization of the octa-polyrhamnose KLH-glycoconjugate

The rabbit serum derived from immunization of the octa-polyrhamnose KLH-glycoconjugate recognizes intermediate to long chain polyrhamnose structure with preferences for the long chain polyrhamnose. Interestingly, it also recognizes **40** and **42** structures with GlcNAc chains. It could be explained that the serum recognizes the "Patch" structures, as these structures have a polyrhamnose chain exposure.

Glycan microarray analysis revealed different binding requirements for the anti-polyrhamnose antibody, the anti-streptococcus pyogenes GAC antibody, the anti-Streptococcus Group A antibody, and the rabbit serum derived from immunization of the octa-polyrhamnose KLH-glycoconjugate. The anti-polyrhamnose antibody interacts with the synthetic GAC library differently from the anti-streptococcus pyogenes GAC antibody and the anti-Streptococcus Group A antibody. The latter two require the presence of the GlcNAc side chain on the periphery of the polyrhamnose backbone for the recognition. The anti-polyrhamnose antibody has preferences for the longer chain structure of the polyrhamnose oligosaccharides. The distance between the GlcNAc side chains plays an important role in recognition. This confirmed our initial hypothesis that in the native bacteria system, the expressed polysaccharide might have a "patch-structure", which leads to the longer distance between the GlcNAc side chains. The "patch-structure" has different angles and distances between the two adjacent GlcNAc side chains; Therefore, it can better fit and bind in the binding pocket of the antibodies. The rabbit serum derived from immunization of the octa-polyrhamnose KLH-glycoconjugate recognizes not only polyrhamnose structures but also **40** and **42** structures with GlcNAc side chains. It could be explained by the polyrhamnose chain exposure within these structures. The results remain a great interest for the further detailed crystallography study between the antibody and the synthetic GAC oligosaccharides.

Conclusion

In conclusion, we have established a comprehensive examining process for the key disaccharide modular strategy to assemble a library of GAC oligosaccharides with various length and GlcNAc side chain variations. We employed an orthogonally protected key disaccharide to facilitate the assembly process and built the library with the previously discovered “4+4” in situ polymerization. We utilized the glycan microarray technology to perform glyco-type mapping with different type-specific antibodies, interpreted and developed a multi-GlcNAc “pocket-fitting” binding theory based on the array binding results. The anti-polyrhamnose antibody, the anti-streptococcus pyogenes GAC antibody and the anti-Streptococcus Group A antibody exhibit distinct binding requirements according to the array results. The anti-polyrhamnose requires long-chain structure to have strong binding strength, and the binding strength maximized at double helical cycle dodecasaccharide level. The latter two antibodies require GlcNAc side chains on the periphery of the polyrhamnose backbone, and multiple GlcNAc side chains enhance the binding strength. The “patch” structure of the customized octasaccharide has the maximum binding strength, as the two adjacent GlcNAc side chains have different angles and distances that leads to better fitting and stronger binding to the antibody. The rabbit serum derived from immunization of the octa-polyrhamnose KLH-glycoconjugate recognizes not only polyrhamnose structures but also **40** and **42** structures with GlcNAc side chains. It could be explained by the polyrhamnose chain exposure within these structures. Future directions can focus on conducting further screening against cross-reactive patient serum to reveal the binding requirements for the auto-reactivity of the antibody derived from GAS infection patients with rheumatic fever complications. Another point of interest of the binding results can focus on the

detailed crystallography study between the antibody and the synthetic GAC oligosaccharides to further clarify the antibody binding interaction mechanisms.

General procedures

NMR spectra were recorded in the NMR facility of Complex Carbohydrate Research Center, UGA, on a Varian Mercury 300 (300 MHz for ^1H , 75 MHz for ^{13}C), Varian Inova 500 (500 MHz for ^1H , 125 MHz for ^{13}C), Varian Inova 600 with cryoprobe (600 MHz for ^1H , 150 MHz for ^{13}C), Varian VNMRS 600 with cryoprobe (600 MHz for ^1H , 150 MHz for ^{13}C) or Varian Inova 800 with cryoprobe (800 MHz for ^1H , 200 MHz for ^{13}C). Chemical shifts are reported in parts per million (ppm) relative to tetramethylsilane (TMS) as the internal standard. NMR data is presented as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, dd = doublet of doublet, m = multiplet and/or multiple resonances), integration, coupling constant in Hertz (Hz). The assignments of ^1H NMR peaks were made from 2D ^1H - ^1H COSY, ^1H - ^{13}C HSQC, ^1H - ^{13}C HMBC, ^1H - ^1H ROESY and ^1H - ^1H TOCSY spectra. Mass spectra were recorded on an ABISciex 5800 MALDI-TOF-TOF, Bruker Microflex MALDI-TOF or Shimadzu LCMS-IT-TOF mass spectrometer. The matrix was used was 2, 5-dihydroxy-benzoic acid (DHB). TLC-analysis performed on Silica gel 60 F254 (EMD Chemicals inc.) with detection by UV-absorption (254 nm) when applicable, and by spraying with a solution of $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24}\cdot\text{H}_2\text{O}$ (25 g/L) in 5% sulfuric acid in ethanol followed by charring. Acid washed molecular sieves (4Å) were flame activated *in vacuo*. All moisture sensitive reactions were carried out under an argon atmosphere. Ambient temperature in the laboratory was usually 20 °C. CH_2Cl_2 and CH_3CN were distilled freshly from CaH_2 . Other commercially available reagents were obtained from Aldrich, Fisher or TCI and used as received.

General procedure for the synthesis of (*N*-phenyl)-trifluoroacetimidate donor

To a solution of hydroxyl precursor in CH₂Cl₂ at 0 °C was added trifluoro-*N*-phenylacetimidoyl chloride (2.0 eq) and Cs₂CO₃ (2.0 eq). The reaction mixture was stirred for 3 h and filtered through a pad of Celite, then concentrated in vacuo. The concentrates were purified by flash chromatography over silica gel (hexane/EtOAc/1% NEt₃, 6/1 to 2:1, v/v) to give the (*N*-phenyl)-trifluoroacetimidate donor as an anomeric mixture and was used directly.

The optimal glycosylation method

Donor (1.2 eq) and acceptor (1.0 eq) were co-evaporated with toluene (3 × 3 mL) and dried under high vacuo overnight. Then, dissolved it under argon in anhydrous DCM to maintain the acceptor concentration of 0.1 M. Freshly activated powdered 4Å acid washed molecular sieves were added at -40 °C, then the solution was argon purged 3 times and stirred under argon for 20 min. The appropriate type and amount acid was added at -40°C or -30°C, and the temperature was maintained for additional 20 min before raising it to 0°C naturally. The reaction mixture was quenched by the addition of a drop of Et₃N (3 µL) at 0°C. The mixture was diluted with DCM and filtered. Then the filtrate was concentrated under reduced pressure and the residue was purified by silica gel column chromatography or LH20.

Global deprotection

Deprotection sequence is very important for these compounds. Generally, deprotection of Troc by Zn and acetylation were conducted in the first step. Then, deprotection of the esters by sodium methoxide in methanol must be conducted in the second step to avoid β-elimination. Thirdly, deprotection of the thio-glycoside by NBS in acetone:10 mM PBS buffer (6:1). Lastly, deprotection of the benzyl ether by 10% Pd/C hydrogenation. The final compound was purified by P2 or P4 column by 0.1 M ammonium bicarbonate buffer.

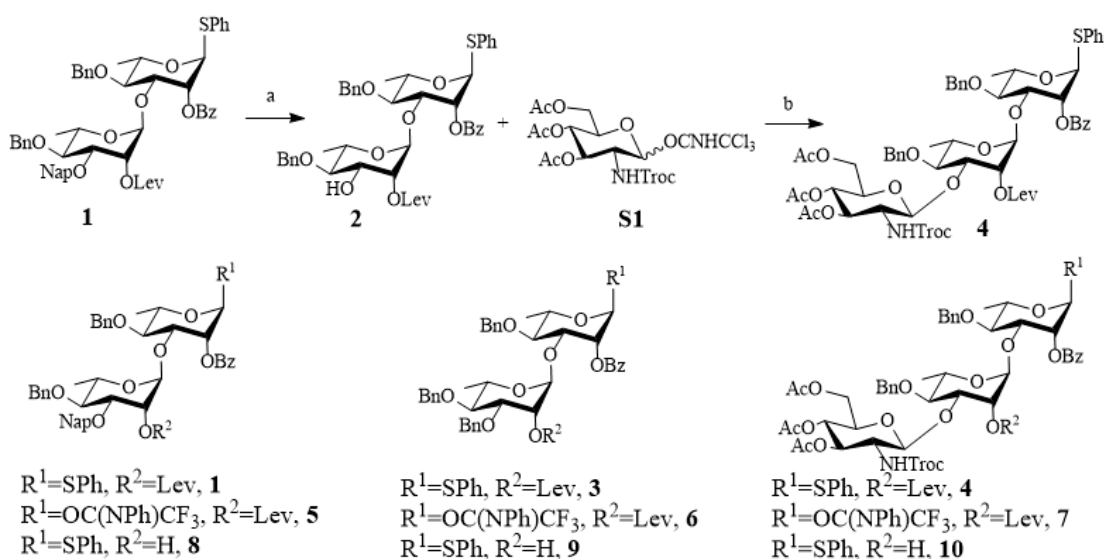
Removal of Lev ester

Hydrazine acetate (5 eq per Lev) was added to a solution of starting material in a mixture of DCM and MeOH (1/1, v/v, 0.02 M) and was stirred until TLC/MALDI indicated the disappearance of starting material (~ 3 h). The reaction mixture was diluted with DCM (30 mL), washed with water (3 × 25 mL) and brine (25 mL), dried over MgSO₄, and filtered. The crude product was purified by silica gel chromatography (hexane:EtOAc, 5:1 to 3:1) to provide the pure product.

Removal of anomeric thiophenol

The reactant was dissolved in acetone, and 2.5 eq NBS was added at 0 °C. The solution was stirred for 5 h performed in the dark. The reaction was monitored by TLC. Upon completion of the reaction, the reaction mixture was added to an aqueous solution of NaHCO₃ (0.1 M) to quench the reaction. The reaction mixture was extracted by DCM (50 mL) three times. The organic extracts were dried with anhydrous MgSO₄, and concentrated under reduced pressure. The crude product was purified by silica gel chromatography (hexane:EtOAc, 2:1 to 1:1) to provide pure products.

Synthesis of building blocks

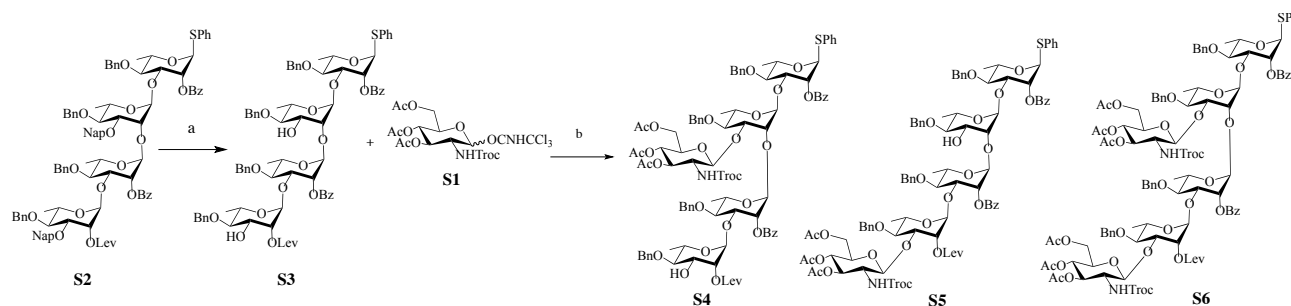


Scheme 2. Synthesis of the trisaccharide **4** and modular disaccharides and trisaccharide building blocks. Reagents and conditions: a) DDQ 2.3eq, 0°C, Chloroform/PBS buffer(9:1), 81%. b) TMSOTf 0.1eq, -40~0°C, 4A AW MS, Dry DCM, 95%

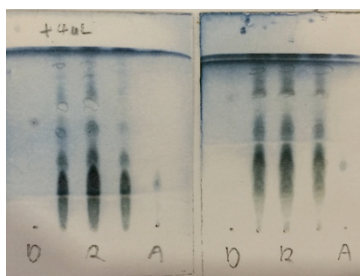
The synthesis of the key disaccharide **1** has been described previously. Start from the key disaccharide, deprotection of Nap by treatment with DDQ in chloroform/PBS buffer (9:1) at 0 °C yielded the acceptor disaccharide **2**. Then a TMSOTf promoted glycosylation between the imidate donor **S1** and the acceptor **2** yielded the trisaccharide building block **4** with 95% excellent yield. With the key disaccharide **1**, the derivatization of **2** can be performed smoothly. The manipulation details have been described fully before, and methods for removing Lev, Nap, and installing *N*-Phenyl imidate have been described in the general method section. Briefly, removal of the anomeric thiophenol with NBS in acetone/H₂O (6:1) yielded the hemiacetal which is further manipulated to obtain the imidate donor **5** by standard conditions. Removal of the Lev on the disaccharide **1** by hydrazine acetate yielded the acceptor **8**. The imidate donor **6** and the acceptor **9** with C3-Bn can be obtained in the same method. Finally, the trisaccharide

building block **4** can be further derivatized to obtain the imidate donor **7** and acceptor **10** with similar methods.

The trial of the divergent strategy



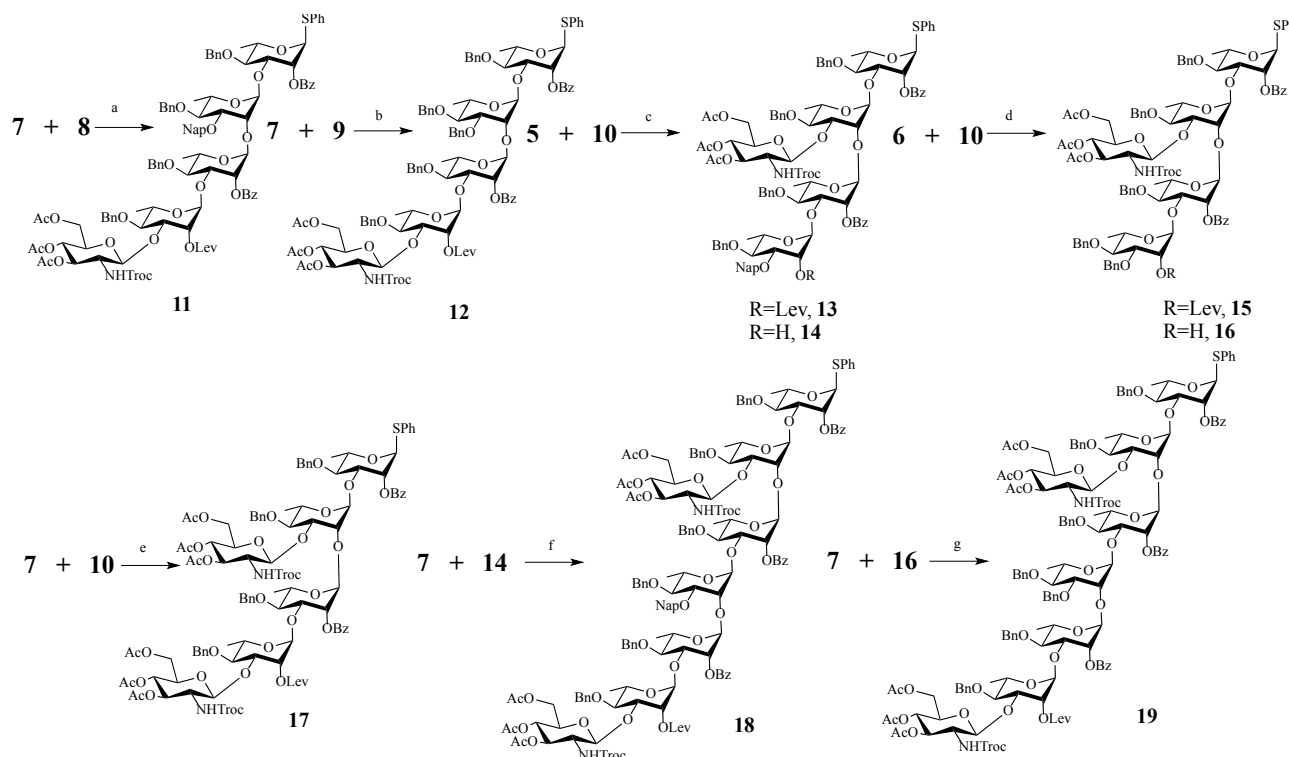
Scheme 1. The divergent strategy of assembly of the GAC oligosaccharides. Reagents and conditions: a) DDQ 5.0eq, Chloroform:PBS buffer(9:1), 92%. b) TMSOTf 0.1 eq, 4A AW MS, dry DCM, -40°C~0°C.



The trial of the divergent strategy is firstly conducted to test the overall reactivity of the tetrasaccharide with two open hydroxyl groups. The initial thought was to employ 3 eq of monosaccharide donor **S1** to push the equilibrium towards the products, with the expectation of yielding a small library of various branched compounds **S4**, **S5**, and the fully branched **S6**. As expected, results indeed showed three closely overlapped spots on the TLC plate. HR-MALDI shows strong signal peaks corresponding to these three branched compounds. However, the purification of these compounds posed a difficult problem. The molecular mass of these compounds is very close to each other, and the compounds can decompose during the silica column purification. The recovered yield is very disappointing. Another issue associated with

this divergent approach is scalability. Therefore, this approach is not a very practical and suitable method for this project.

The Modular approach to assemble the GAC library



Scheme 3. Modular synthesis of the GAS oligosaccharide antigen library. Reagents and conditions: a) TBSOTf 0.15eq, -40~0°C, 4A AW MS, Dry DCM, 22%. b) TBSOTf 0.15eq, -40~0°C, 4A AW MS, Dry DCM, 31%. c) TMSOTf 0.3eq, -40~0°C, 4A AW MS, Dry DCM, 75%. d) TBSOTf 0.3eq, -40~0°C, 4A AW MS, Dry DCM, 82%. e) TBSOTf 0.15eq, -40~0°C, 4A AW MS, Dry DCM, 96%. f) TBSOTf 0.15eq, -40~0°C, 4A AW MS, Dry DCM, 23%. g) TBSOTf 0.15eq, -40~0°C, 4A AW MS, Dry DCM, 43%; TBSOTf 0.15eq, -40~0°C, 4A AW MS, Dry Toluene, 37%

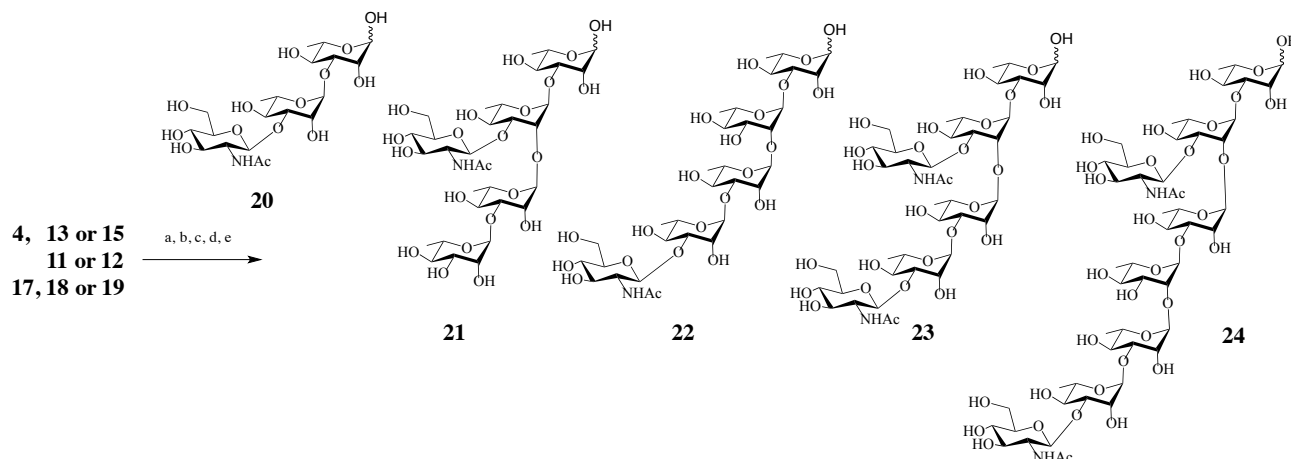
We decided to utilize a modular approach to assemble the GAC library with different side chain variations. This way, we can control the formation of the branch on different position and investigate the reactivity of the glycosylation under the influence of different neighboring protection groups.

A TBSOTf promoted glycosylation of compound **7** with compound **8** yielded the pentasaccharide **11** in 22% poor yield. A TBSOTf promoted glycosylation of compound **7** with compound **9** yielded the pentasaccharide **12** in 31% poor yield. However, a TMSOTf promoted

glycosylation of compound **5** with compound **10** yielded the pentasaccharide **13** in 75% good yield. A TBSOTf promoted glycosylation of compound **6** with compound **10** yielded the pentasaccharide **12** in 82% good yield. A TBSOTf promoted glycosylation of compound **7** with compound **10** yielded the hexasaccharide **17** in 96% excellent yield.

Comparing these results, it shows that when bulky aromatic protecting groups such as Nap or Bn are present at the C-3 position, the adjacent C-2 hydroxyl will be blocked, and the reactivity of the C-2 hydroxyl in the glycosylation process will be dramatically decreased, as the aromatic ring can rotate around the O-CH₂ bond and block the incoming glycosyl donors. However, when GlcNAc is present at the C-3 position, the adjacent C-2 hydroxyl will not be blocked at all. The reactivity of the C-2 hydroxyl in the glycosylation process is excellent, as the glycosidic bond between the GlcNAc and the rhamnose is more rigid and the GlcNAc cannot rotate around the glycosidic bond. Therefore, the incoming glycosyl donors will not be blocked in the glycosylation process. The overall reactivity of the C-2 hydroxyl is in the order of GlcNAc>Bn>Nap when these protecting groups or carbohydrate are at the adjacent C-3 position. Then, deprotection of the Lev on compounds **13** and **15** by hydrazine acetate yielded compound **14** and **16** in excellent yield. A TBSOTf promoted glycosylation of compound **7** with compound **14** yielded the octasaccharide **18** in 23% poor yield. Finally, A TBSOTf promoted glycosylation of compound **7** with compound **16** yielded the octasaccharide **19** in 43% acceptable yield in DCM and 37% acceptable yield in toluene. The low reactivity of the final octa-assembly glycosylation is due to the presence of the C-2 Nap or Bn bulky aromatic protecting groups, which blocks the incoming glycosyl donors as we have just explained before.

Deprotection of GAC oligosaccharides



Scheme 4. Deprotection of the GAS oligosaccharide antigen library. Reagents and conditions: a) Zn, HOAc, DCM. b) Ac₂O, DMAP, pyridine. c) NaOMe/HOMe. d) NBS, Acetone: PBS buffer(6:1). e) Pd/C, THF/H₂O

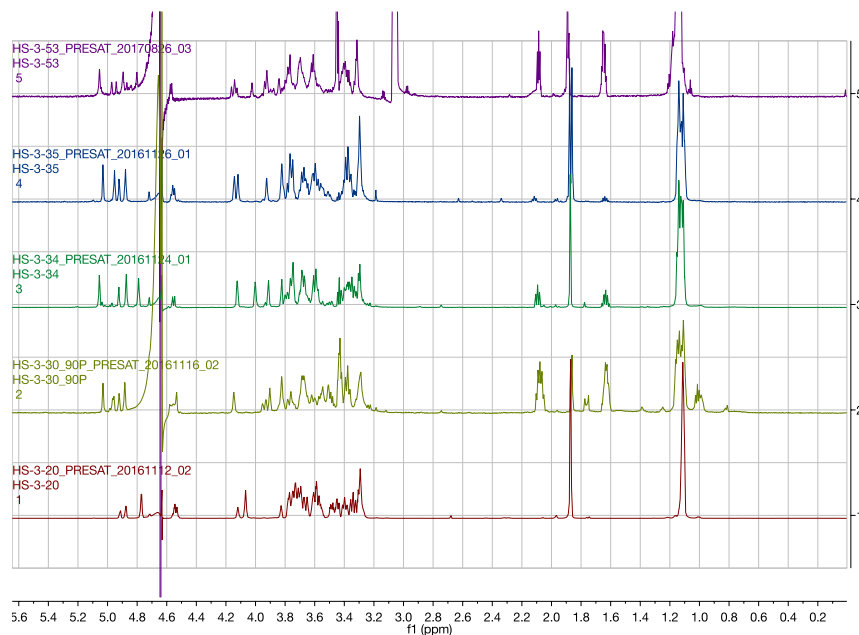
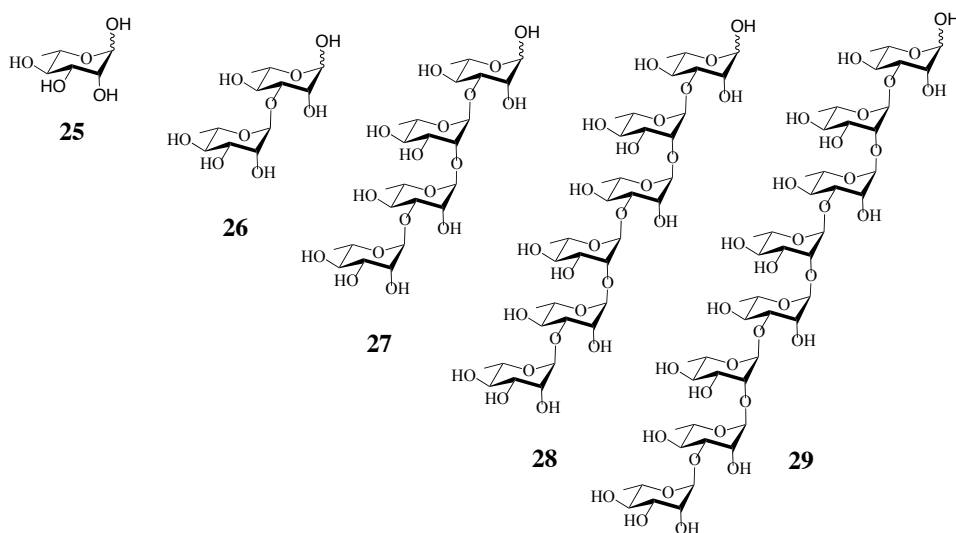


Figure 6. NMR Spectra of deprotected GAC library with GlcNAc variations

Deprotection sequence is crucial for these compounds as shown in **Scheme 4**. Firstly, deprotection of the Troc by Zn reduction to unmask the amine group, then it can be acetylated by acetyl anhydride. Secondly, deprotection of the esters by sodium methoxide in methanol must be conducted in this step to avoid β -elimination. Thirdly, the anomeric thioglycoside was removed by NBS in acetone:10 mM PBS buffer (6:1). Lastly, deprotection of the benzyl ether by 10%

Pd/C in THF/H₂O, and subsequently purification by P2 Bio Gel yielded pure final compounds in good yield. The pure compounds could be used for the structure-activity analysis, the glycan microarray, and ELISA antibody binding studies. Two-dimensional NMR experiments combined with high-resolution MALDI-MS confirmed the structural integrity of the compound. *Compound 35*, 36* confirmed by HR-MALDI-MS only due to limited quantity. The J_{H-H} coupling constants of non-reducing end anomeric protons were all between 169 and 170 Hz, indicating all α -glycosidic linkages. The J_{H-H} coupling constants of the reducing end anomeric protons show two separated peaks at 160Hz and 170Hz, which indicated β and α isomers respectively.



Scheme 5. The polyramnose antigen library

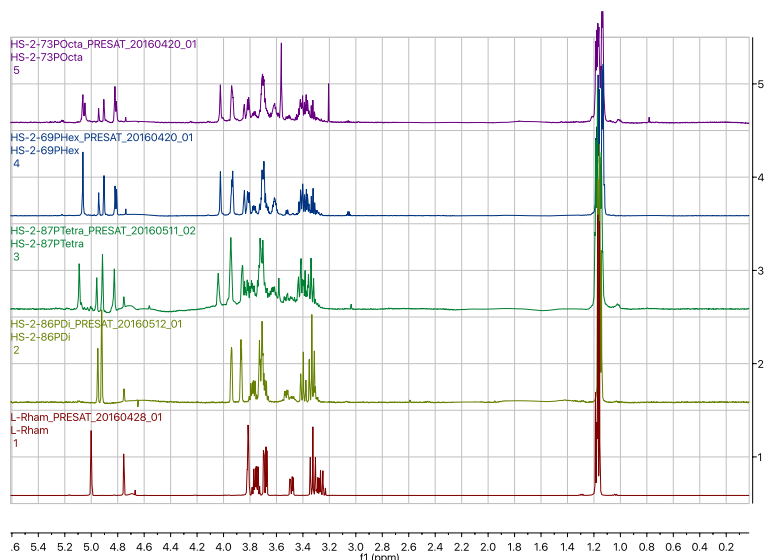
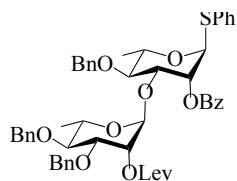


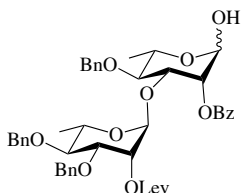
Figure 7. NMR Spectra of deprotected polyramnose library

Experimental procedure



3

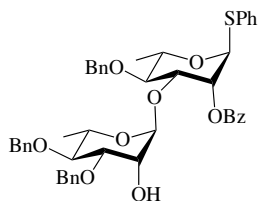
Thiophenyl O-(4-O-benzyl-3-O-benzyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (3). Compound **3** was obtained by the optimal glycosylation method as described in the General Procedures. Yield: 80%. ^1H NMR (600 MHz, chloroform-*d*) δ 8.02 (m, 2H, arom. H, Bz), 7.74 – 7.04 (m, 23H, arom. H, Bn, Bz, SPh), 5.60 (m, 1H, H-2,A), 5.35 (d, 1H, H-1,A), 5.13 – 5.01 (d, 1H, H-1,B), 4.92 – 4.12 (m, 8H, 3xCH₂, Bn, H-3,A, H-5,A), 3.81 – 3.69 (m, 2H, H-3,B, H-5,B), 3.63 (m, 1H, H-4,A), 3.37 – 3.30 (m, 1H, H-4,B), 2.73 – 2.59 (m, 4H, CH₂-CH₂, Lev), 2.14 (s, 3H, CH₃, Lev), 1.15 (d, J = 6.2 Hz, 3H, Rha, CH₃,A), 0.90 (d, J = 7.2 Hz, 3H, Rha, CH₃,B). ^{13}C NMR (151 MHz, chloroform-*d*) δ 181.94, 171.82, 165.61, 138.48, 133.76, 133.31, 131.81, 130.85, 129.75, 129.34, 129.04, 128.77, 128.48, 128.47, 128.20, 128.17, 128.04, 128.01, 127.90, 127.88, 127.62, 127.60, 127.55, 127.42, 99.31, 85.63, 80.49, 79.60, 77.61, 77.39, 75.49, 74.66, 74.23, 71.57, 69.34, 69.15, 68.71, 68.13, 38.70, 38.01, 30.34, 29.81, 29.68, 28.90, 28.09, 22.97, 17.94, 17.75, 10.94. HR-MALDI-TOF/MS (m/z): calcd for C₅₁H₅₄NaO₁₁S⁺ [M+Na]⁺: 897.3312; found, 897.3318.



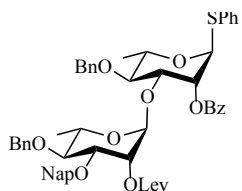
3H

2-O-levulinoyl-3-O-benzyl-4-O-benzyl- α -L-rhamnopyranosyl-(1 \rightarrow 3)-4-O-benzyl-2-O-

benzoyl- α -L-rhamnopyranose (3H). Compound **3** (1.00 g, 1.14 mmol) was dissolved in acetone:PBS buffer (6:1, 35 mL), and NBS (1.02 g, 5.72 mmol) was added at 0 °C. The solution was stirred for 5 h performed in the dark. The reaction was monitored by TLC. Upon completion of the reaction, the reaction mixture was added to an aqueous solution of NaHCO₃ (0.1 M) to quench the reaction. The reaction mixture was extracted by DCM (3 \times 50 mL). The organic extracts were dried with anhydrous MgSO₄, and concentrated under reduced pressure. The crude product was purified by silica gel chromatography (hexane:EtOAc, 2:1 to 1:1) to provide pure compound **3H** (750 mg, 84%). ¹H NMR (600 MHz, chloroform-*d*) δ 8.17 – 7.93 (m, 2H, arom. H, Bz), 7.70 – 7.00 (m, 28H, arom. H, Bn, Bz), 5.38 (m, 1H, H-2,A), 5.34 (m, 1H, H-1,A), 5.31 – 5.26 (m, 1H, H-2,B), 5.04 (d, *J* = 2.0 Hz, 1H, H-1,B), 4.84 – 4.39 (m, 6H, 3xCH₂, Bn), 4.34 – 4.25 (m, 2H, H-3,A, H-5,A), 3.80 – 3.69 (m, 2H, H-3,B, H-5,B), 3.57 (m, 1H, H-4,A), 3.32 (m, 1H, H-4,B), 2.71 – 2.57 (m, 4H, CH₂-CH₂, Lev), 2.13 (s, 3H, CH₃, Lev), 1.32 – 1.28 (m, 3H, Rha, CH₃,A), 1.13 (m, 3H, Rha, CH₃,B). ¹³C NMR (151 MHz, chloroform-*d*) δ 216.25, 208.91, 171.81, 164.20, 133.25, 129.78, 128.46, 128.43, 128.19, 128.15, 127.98, 127.89, 127.60, 127.52, 99.22, 91.83, 80.33, 79.62, 77.42, 75.35, 74.60, 72.88, 71.53, 69.40, 68.56, 67.85, 38.01, 29.80, 28.09, 18.11, 17.74. HR-MALDI-TOF/MS (*m/z*): calcd for C₄₅H₅₀NaO₁₂⁺ [M+Na]⁺: 805.3312; found, 805.3314.



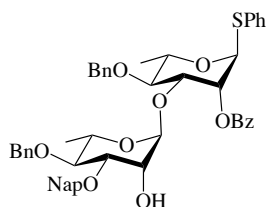
Thiophenyl O-(4-O-benzyl-3-O-benzyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (9). Hydrazine acetate (263 mg, 2.86 mmol) was added to a solution of compound **3** (501 mg, 0.57 mmol) in a mixture of DCM and MeOH (1/1, v/v, 40 mL). Stirring was continued until MALDI indicated the disappearance of the starting material (~3 h). The reaction mixture was diluted with DCM (30 mL), washed with water (3 \times 25 mL) and brine (25 mL), dried with anhydrous MgSO₄, and filtered. The crude product was purified by silica gel chromatography (hexane:EtOAc, 5:1 to 3:1, v/v) to provide pure compound **9** (353 mg, 80%). ¹H NMR (600 MHz, chloroform-*d*) δ 8.10 – 7.95 (m, 2H, arom. H, Bz), 7.64 – 7.13 (m, 23H, arom. H, Bn, Bz, SPh), 5.62 (m, 1H, H-2,A), 5.54 (d, *J* = 1.8 Hz, 1H, H-1,A), 5.11 (d, *J* = 1.8 Hz, 1H, H-1,B), 4.84 – 4.42 (m, 6H, 3xCH₂, Bn), 4.26 (m, 1H, H-5,A), 4.20 (m, 1H, H-3,A), 3.90 (m, 1H, H-2,B), 3.79 – 3.72 (m, 1H, H-5,B), 3.68 (m, 1H, H-3,B), 3.62 (m, 1H, H-4,B), 3.38 (m, 1H, H-4,A), 1.34 (d, *J* = 6.2 Hz, 3H, Rha, CH₃,A), 1.15 (d, *J* = 6.2 Hz, 3H, Rha, CH₃,B). ¹³C NMR (151 MHz, chloroform-*d*) δ 165.59, 137.83, 137.81, 133.77, 133.27, 131.82, 129.76, 129.04, 128.53, 128.46, 128.42, 128.22, 127.95, 127.89, 127.86, 127.75, 127.63, 127.61, 127.49, 101.11, 85.63, 80.53, 79.63, 79.44, 77.89, 75.44, 74.74, 74.37, 72.11, 69.15, 69.06, 68.37, 17.94, 17.68. HR-MALDI-TOF/MS (*m/z*): calcd for C₄₆H₄₈NaO₉S⁺ [M+Na]⁺: 799.2912; found, 799.2918.



1

Thiophenyl O-(4-O-benzyl-3-O-naphthylmethyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (1). Compound **1** was obtained by

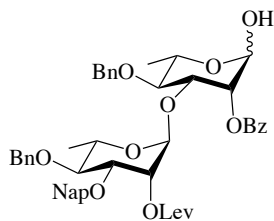
the optimal glycosylation method as described in the General Procedures. Yield: 72%. ^1H NMR (500 MHz, chloroform-*d*) δ 8.07 – 8.02 (m, 2H, arom. H, Bz), 7.87 – 7.13 (m, 25H, arom. H, Bn, Bz, Nap, SPh), 5.66 (m, 1H, H-2,A), 5.59 (d, J = 1.7 Hz, 1H, H-1,A), 5.46 (m, 1H, H-2,B), 5.12 (d, J = 1.8 Hz, 1H, H-1,B), 4.84 - 4.47 (m, 6H, 2xCH₂, Bn, CH₂, Nap), 4.34 – 4.22 (m, 2H, H-3,A, H-5,A), 3.93 – 3.77 (m, 2H, H-3,B, H-5,B), 3.65 (m, 1H, H-4,A), 3.41 (m, 1H, H-4,B), 2.78 – 2.63 (m, 4H, CH₂-CH₂, Lev), 2.15 (s, 3H, CH₃, Lev), 1.36 (d, J = 6.2 Hz, 3H, Rha, CH₃,A), 1.21 (d, J = 6.2 Hz, 3H, Rha, CH₃,B). ^{13}C NMR (75 MHz, chloroform-*d*) δ 138.57, 137.79, 135.47, 133.34, 131.86, 129.76, 129.09, 128.51, 128.28, 128.24, 128.11, 128.00, 127.94, 127.90, 127.67, 127.64, 127.60, 127.47, 126.58, 125.98, 125.96, 125.81, 99.45, 85.68, 80.42, 79.69, 77.90, 75.49, 74.73, 74.30, 71.62, 69.39, 69.20, 68.81, 38.03, 29.81, 28.15, 17.99, 17.82. HR-MALDI-TOF/MS (*m/z*): calcd for C₅₅H₅₆NaO₁₁S⁺ [M+Na]⁺: 947.3412; found, 947.3418.



8

Thiophenyl O-(4-O-benzyl-3-O-naphthylmethyl- α -L-rhamnopyranosyl)-(1→3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (8). Hydrazine acetate (390 mg, 4.232 mmol) was added to a solution of compound **1** (783 mg, 0.846 mmol) in a mixture of DCM and MeOH (1/1, v/v, 40 mL). Stirring was continued until MALDI indicated the disappearance of the starting material (~3 h). The reaction mixture was diluted with DCM (30 mL), washed with water (3 × 25 mL) and brine (25 mL), dried with anhydrous MgSO₄, and filtered. The crude product was purified by silica gel chromatography (hexane:EtOAc, 5:1 to 3:1, v/v) to provide pure compound **8** (651 mg, 93%). ^1H NMR (500 MHz, chloroform-*d*) δ 8.02 – 7.91 (m, 2H, arom. H, Bz), 7.81 – 6.96 (m,

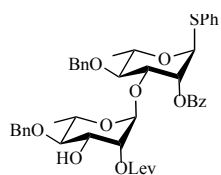
25H, arom. H, Bn, Nap, Bz, SPh), 5.59 – 5.54 (m, 1H, H-2,A), 5.48 (d, $J = 1.7$ Hz, 1H, H-1,A), 5.05 (d, $J = 1.7$ Hz, 1H, H-1,B), 4.88 – 4.30 (m, 6H, 2xCH₂, Bn, CH₂, Nap), 4.19 (m, 1H, H-5,A), 4.13 (m, 1H, H-3,A), 3.89 – 3.86 (m, 1H, H-2,B), 3.69 (m, 2H, H-3,B, H-5,B), 3.52 (m, 1H, H-4,A), 3.36 (m, 1H, H-4,B), 1.25 (m, 3H, Rha, CH₃,A), 1.10 (m, 3H, Rha, CH₃,B). ¹³C NMR (75 MHz, chloroform-*d*) δ 138.49, 137.81, 135.31, 133.80, 133.29, 131.86, 129.87, 129.77, 129.06, 128.60, 128.52, 128.48, 128.36, 128.28, 128.26, 128.03, 127.95, 127.92, 127.88, 127.84, 127.73, 127.68, 127.63, 127.60, 127.50, 126.53, 126.18, 126.02, 125.67, 101.20, 85.65, 80.43, 79.70, 79.50, 78.08, 76.80, 75.35, 74.77, 74.40, 72.21, 69.17, 69.13, 69.07, 68.45, 17.94, 17.72. HR-MALDI-TOF/MS (m/z): calcd for C₅₀H₅₀NaO₉S⁺ [M+Na]⁺: 849.3121; found, 849.3126.



1H

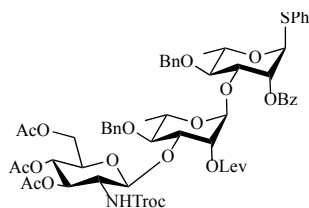
2-O-levulinoyl-3-O-Naphthylmethyl-4-O-benzyl- α -L-rhamnopyranosyl-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranose (1H). To a mixture of compound **1** (940 mg, 1.016 mmol), 2,4,6-tri-*tert*-butylpyrimidine (1.01 g, 4.062 mmol), and *N*-iodosuccinimide (457 mg, 2.031 mmol) in wet DCM (50 mL) at 0 °C was added silver triflate (520 mg, 2.031 mmol). After stirring for 30 min, saturated aqueous Na₂S₂O₃ was added and the mixture was stirred for an additional 30 min while warming to r.t. The reaction was extracted by DCM (3 \times 50 mL), dried with anhydrous MgSO₄, and concentrated under reduced pressure. The crude product was purified by silica gel chromatography (hexane:EtOAc, 2:1 to 1:1) to provide pure compound **1H** (821 mg, 97%). ¹H NMR (500 MHz, chloroform-*d*) δ 8.00 – 7.92 (m, 2H, arom. H, Bz), 7.71 –

7.05 (m, 20H, arom. H, Bn, Nap), 5.32 (m, 2H, H-2,A, H-2,B), 5.19 (d, 1H, H-1,A), 4.99 (d, 1H, H-1,B), 4.79 – 4.35 (m, 6H, 2xCH₂, Bn, CH₂, Nap), 4.23 (m, 1H, H-3,A), 3.98 – 3.90 (m, 1H, H-5,A), 3.78 (m, 1H, H-3,B), 3.76 – 3.68 (m, 1H, H-5,B), 3.46 (m, 1H, H-4,A), 3.28 (m, 1H, H-4,B), 2.57 (m, 4H, CH₂-CH₂, Lev), 2.03 (s, 3H, CH₃, Lev), 1.21 (d, $J = 6.2$ Hz, 3H, Rha, CH₃,A), 1.08 (d, $J = 6.2$ Hz, 3H, Rha, CH₃,B). ¹³C NMR (75 MHz, chloroform-*d*) δ 138.57, 137.79, 135.47, 133.34, 131.86, 129.76, 129.09, 128.51, 128.28, 128.24, 128.11, 128.00, 127.94, 127.90, 127.67, 127.64, 99.45, 92.01, 80.42, 79.69, 77.90, 75.49, 74.73, 74.30, 71.62, 69.39, 69.20, 68.81, 38.03, 29.81, 28.15, 17.99, 17.82. HR-MALDI-TOF/MS (*m/z*): calcd for C₄₉H₅₂NaO₁₂⁺ [M+Na]⁺: 855.3513; found, 855.3519.



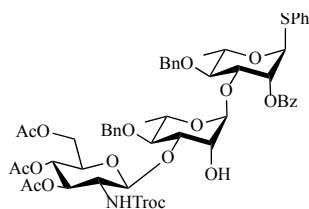
2

Thiophenyl O-(4-O-benzyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (2). Compound **2** was derived from the general method of Nap deprotection as described in the General Procedures. Yield: 81%. ¹H NMR (500 MHz, chloroform-*d*) δ 8.16 – 7.98 (m, 2H, arom. H, Bz), 7.68 – 7.14 (m, 18H, arom. H, Bn, SPh), 5.63 (m, 1H, H-2,A), 5.61 – 5.58 (m, 1H, H-1,A), 5.26 (m, 1H, H-2,B), 5.11 (d, $J = 1.7$ Hz, 1H, H-1,B), 4.98 – 4.50 (m, 4H, 2xCH₂, Bn), 4.34 – 4.22 (m, 2H, H-3,A, H-5,A), 4.02 (m, 1H, H-3,B), 3.92 – 3.67 (m, 2H, H-5,B, H-4,A), 3.38 – 3.29 (m, 1H, H-4,B), 2.87 – 2.52 (m, 4H, CH₂-CH₂, Lev), 2.07 (s, 3H, CH₃, Lev), 1.38 – 1.30 (m, 3H, Rha, CH₃,A), 1.29 – 1.23 (m, 3H, Rha, CH₃,B). HR-MALDI-TOF/MS (*m/z*): calcd for C₄₄H₄₈NaO₁₁S⁺ [M+Na]⁺: 807.2821; found, 807.2827.



4

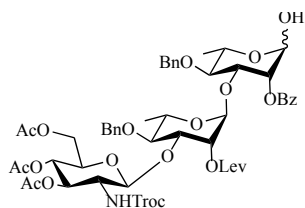
Thiophenyl **O-(2-[[[(2,2,2-trichloroethoxy)carbonyl]amino]-3,4,6-O-acetyl- β -D-glucopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (4).** Compound **4** was prepared by the optimal glycosylation method as described in the General Procedures. Yield: 91%. ^1H NMR (600 MHz, chloroform-*d*) δ 8.18 – 7.94 (m, 2H, arom. H, Bz), 7.75 – 6.99 (m, 18H, arom. H, Bn, Bz, SPh), 5.64 (m, 1H, H-2,A), 5.52 (d, J = 2.2 Hz, 1H, H-1,A), 5.21 (m, 1H, H-2,B), 5.05 (d, 1H, H-1,B), 5.01 – 4.31 (m, 8H, H-3,3G, H-4,3G, 2xCH₂, Bn, CH₂, Troc), 4.31 – 4.18 (m, 2H, H-3,A, H-5,A), 3.94 (m, 3H, H-5,B, H-3,B, H-6,3G), 3.82 – 3.63 (m, 2H, H-6,3G, H-4,A), 3.52 (m, 1H, H-2,3G), 3.40 – 3.33 (m, 1H, H-4,B), 3.13 (m, 1H, H-5,3G), 2.79 – 2.47 (m, 4H, CH₂-CH₂, Lev), 2.19 (s, 3H, CH₃, Lev), 2.03 (s, 3H, CH₃, Ac), 1.97 (s, 3H, CH₃, Ac), 1.94 (s, 3H, CH₃, Ac), 1.35 (m, 3H, Rha, CH₃,A), 1.18 – 1.14 (m, 3H, Rha, CH₃,B). HR-MALDI-TOF/MS (m/z): calcd for C₅₉H₆₆Cl₃NNaO₁₉S⁺ [M+Na]⁺: 1252.2921; found, 1252.2928.



10

Thiophenyl **O-(2-[[[(2,2,2-trichloroethoxy)carbonyl]amino]-3,4,6-O-acetyl- β -D-glucopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (10).** Compound **10** was derived from the general method of

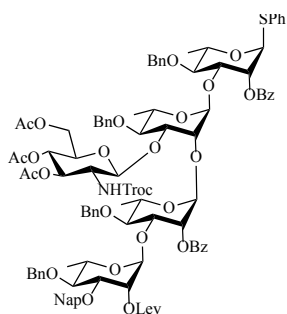
Lev deprotection as described in the General Procedures. Yield: 76%. ^1H NMR (600 MHz, chloroform-*d*) δ 8.15 – 8.01 (m, 2H, arom. H, Bz), 7.71 – 7.12 (m, 18H, arom. H, Bn, SPh, Bz), 5.66 (m, 1H, H-2,A), 5.54 (d, $J = 1.7$ Hz, 1H, H-1,A), 5.12 (d, $J = 1.8$ Hz, 1H, H-1,B), 4.98 – 4.83 (m, 2H, H-3,3G, H-4,3G), 4.79 – 4.38 (m, 7H, 2xCH₂, Bn, CH₂, Troc, H-1,3G), 4.25 – 4.16 (m, 2H, H-3,A, H-5,A), 4.02 (m, 1H, H-6,3G), 3.97 (d, $J = 2.6$ Hz, 1H, H-2,B), 3.82 (m, 2H, H-6,3G, H-3,B), 3.74 (m, 1H, H-5,B), 3.63 (m, 2H, H-4,A, H-6,3G), 3.43 (m, 1H, H-4,B), 3.30 – 3.19 (m, 1H, H-5,3G), 2.00 (s, 3H, CH₃, Ac), 1.96 (m, 6H, 2xCH₃, Ac), 1.36 (d, $J = 6.2$ Hz, 3H, Rha, CH₃,A), 1.14 (d, $J = 6.2$ Hz, 3H, Rha, CH₃,B). ^{13}C NMR (151 MHz, chloroform-*d*) δ 169.22, 153.99, 133.68, 133.37, 131.85, 129.86, 129.75, 129.06, 128.61, 128.53, 128.38, 127.97, 127.67, 127.58, 127.52, 126.90, 101.08, 100.86, 85.69, 80.65, 80.46, 79.83, 75.24, 74.28, 74.22, 74.17, 71.65, 70.40, 69.13, 68.41, 68.06, 61.35, 56.09, 20.58, 17.92, 17.73. HR-MALDI-TOF/MS (*m/z*): calcd for C₅₄H₆₀Cl₃NNaO₁₇S⁺ [M+Na]⁺: 1154.2521; found, 1154.2529.



4H

2-[[[(2,2,2-trichloroethoxy)carbonyl]amino]-3,4,6-O-acetyl- β -D-glucopyranosyl-(1 \rightarrow 3)-(4-O-benzyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (4H). Compound **4H** was derived from the general method of anomeric thiophenol deprotection as described in the General Procedures. Yield: 98%. ^1H NMR (600 MHz, chloroform-*d*) δ 8.22 – 8.04 (m, 2H, arom. H, Bz), 7.71 – 7.13 (m, 13H, arom. H, Bn, Bz), 5.42 (m, 1H, H-2,A), 5.26 (m, 1H, H-2,B), 5.03 (d, $J = 1.8$ Hz, 1H, H-1,A), 4.98 – 4.83 (m, 4H, H-3,3G, H-4,3G, CH₂, Bn), 4.75 – 4.45 (m, 5H, H-1,B, CH₂, Bn, CH₂, Troc), 4.42 (m, 1H, H-

3,A), 4.30 (m, 1H, H-5,A), 4.23 (m, 1H, H-5,B), 4.15 – 3.87 (m, 4H, H-3,B, H-6,3G, H-6,3G, H-4,A), 3.55 (m, 1H, H-2,3G), 3.34 (m, 1H, H-4,B), 3.08 (m, 1H, H-5,3G), 2.76 – 2.52 (m, 4H, CH₂-CH₂, Lev), 2.18 (s, 3H, CH₃, Lev), 1.97 (m, 3H, CH₃, Ac), 1.93 (m, 6H, 2×CH₃, Ac), 1.32 (d, $J = 6.2$ Hz, 3H, Rha, CH₃,A), 1.13 (d, $J = 6.2$ Hz, 3H, Rha, CH₃,B). ¹³C NMR (151 MHz, chloroform-*d*) δ 170.61, 169.28, 165.72, 153.78, 138.52, 138.01, 133.36, 130.04, 129.93, 129.81, 128.66, 128.54, 128.45, 128.39, 128.36, 128.03, 127.83, 127.56, 126.85, 100.87, 98.48, 91.96, 80.41, 79.92, 77.29, 76.29, 75.26, 74.26, 74.07, 72.68, 71.83, 71.39, 68.50, 68.01, 67.82, 61.12, 56.18, 37.96, 29.87, 28.01, 20.60, 20.58, 20.55, 18.07. HR-MALDI-TOF/MS (m/z): calcd for C₅₃H₆₂Cl₃NNaO₂₀⁺ [M+Na]⁺: 1160.2821; found, 1160.2828.

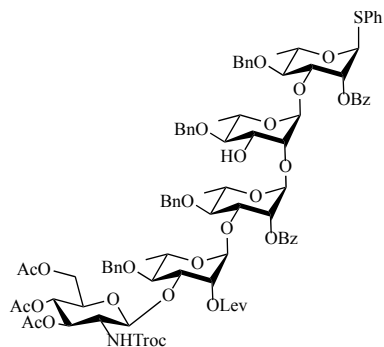


13

Thiophenyl O-(4-O-benzyl-3-O-Naphthylmethyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-{(2-[[2,2,2-trichloroethoxy)carbonyl]amino]-3,4,6-O-acetyl- β -D-glucopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl- α -L-rhamnopyranosyl)}-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (13).

Compound **13** was derived from the optimal glycosylation method as described in the General Procedures. Yield: 75%. ¹H NMR (600 MHz, chloroform-*d*) δ 8.06 (m, 4H, arom. H, Bz), 7.83 – 7.03 (m, 38H, arom. H, Bn, Nap, SPh, Bz), 5.67 (m, 1H, H-2,A), 5.61 (m, 1H, H-3,3G), 5.58 – 5.42 (m, 3H, H-1,A, H-2,D, H-2,C), 5.12 – 5.05 (m, 1H, H-1,C), 5.05 – 4.93 (m, 3H, H-1,3G, H-1,B, H-4,3G), 4.89 – 4.42 (m, 12H, 4×CH₂, Bn, CH₂, Nap, CH₂, Troc), 4.31 – 4.20 (m, 2H, H-

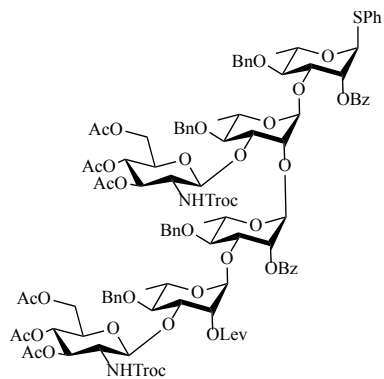
3,D, H-5,A), 4.20 – 4.07 (m, 2H, H-3,A, H-2,B), 3.96 (m, 1H, H-3,B), 3.91 – 3.74 (m, 3H, H-3,C, H-5,D, H-5,C), 3.67 – 3.48 (m, 4H, H-5,B, H-4,A, H-4,D, H-2,3G), 3.44 – 3.27 (m, 3H, H-6,3G, H-4,B, H-5,3G), 2.71 – 2.56 (m, 4H, CH₂-CH₂, Lev), 2.08 (s, 3H, CH₃, Lev), 1.94 (m, 9H, 3×CH₃, Ac), 1.29 – 1.20 (m, 6H, Rha, CH₃,A,C), 1.18 – 1.02 (m, 6H, Rha, CH₃,B,D). ¹³C NMR (151 MHz, chloroform-*d*) δ 206.95, 170.31, 133.69, 133.37, 131.85, 129.79, 129.74, 129.03, 128.55, 128.53, 128.43, 128.25, 128.20, 127.92, 127.83, 127.79, 127.67, 127.62, 127.58, 127.48, 127.36, 127.12, 126.54, 125.94, 125.90, 125.75, 85.60, 80.13, 79.59, 76.90, 75.40, 75.30, 74.24, 73.66, 72.58, 71.46, 71.14, 69.14, 68.62, 56.60, 37.99, 30.91, 29.72, 28.05, 20.61, 20.19, 18.04, 17.85, 17.66. HR-MALDI-TOF/MS (m/z): calcd for C₁₀₃H₁₁₀Cl₃NNaO₂₈S⁺ [M+Na]⁺: 1968.5921; found, 1968.5928.



S5

Thiophenyl O-(2-[[[(2,2,2-trichloroethoxy)carbonyl]amino]-3,4,6-O-acetyl-β-D-glucopyranosyl)-(1→3)-(4-O-benzyl-2-O-levulinoyl-α-L-rhamnopyranosyl)-(1→3)-(4-O-benzyl-2-O-benzoyl-α-L-rhamnopyranosyl)-(1→2)-(4-O-benzyl-3-O-Naphthylmethyl-α-L-rhamnopyranosyl)-(1→3)-4-O-benzyl-2-O-benzoyl-α-L-rhamnopyranoside (S5). Compound **S5** was derived from the trial divergent strategy by the optimal glycosylation method as described in the General Procedures. Yield: Low, less than 5%. ¹H NMR (600 MHz, chloroform-*d*) δ 8.16 – 7.92 (m, 4H, arom. H, Bz), 7.71 – 7.08 (m, 31H, arom. H, Bz, Bn, SPh), 5.60 – 5.53

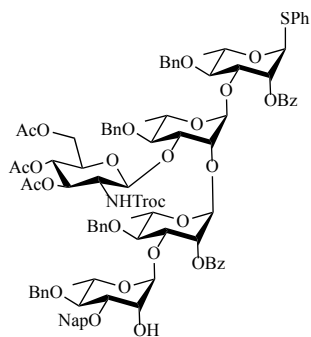
(m, 2H, H-1,A, H-2,A), 5.46 (m, 1H, H-2,D), 5.21 (d, 1H, H-1,B), 5.13 (d, 1H, H-1,D), 5.04 (d, 1H, H-1,C), 4.95 (m, 2H, H-3,5G, H-4,5G), 4.91 – 4.48 (m, 10H, 4xCH₂, Bn, CH₂, Troc), 4.46 (d, $J = 8.7$ Hz, 1H, H-1,5G), 4.24 (m, 3H, H-3,C, H-5,C, H-3,A), 4.18 – 4.06 (m, 1H, H-5,A), 3.98 (m, 1H, H-3,D), 3.95 – 3.84 (m, 4H, H-2,B, H-3,B, H-6,5G, H-5,D), 3.80 (m, 2H, H-5,B, H-4,A), 3.74 – 3.46 (m, 3H, H-4,C, H-2,5G, H-6,5G), 3.27 (m, 1H, H-4,B), 3.12 (d, $J = 6.8$ Hz, 1H, H-5,5G), 2.78 – 2.52 (m, 4H, CH₂-CH₂, Lev), 2.16 (s, 3H, CH₃, Lev), 1.97 (s, 3H, CH₃, Ac), 1.94 (s, 6H, 2×CH₃, Ac), 1.32 – 1.23 (m, 6H, Rha, CH₃,A,C), 1.20 – 1.10 (m, 6H, Rha, CH₃,B,D). ¹³C NMR (151 MHz, chloroform-*d*) δ 133.33, 131.80, 129.80, 129.72, 129.04, 128.58, 128.49, 128.46, 128.37, 128.03, 127.87, 127.82, 127.64, 127.56, 126.84, 100.88, 99.27, 98.45, 85.53, 81.47, 80.34, 80.22, 79.10, 78.77, 77.03, 76.36, 75.50, 75.27, 74.50, 72.19, 71.81, 71.43, 70.62, 69.20, 68.52, 68.38, 68.01, 61.09, 37.96, 28.00, 20.60, 17.99, 17.96, 17.81. HR-MALDI-TOF/MS (*m/z*): calcd for C₉₂H₁₀₂Cl₃NNaO₂₈S⁺ [M+Na]⁺: 1828.5321; found, 1828.5327.



17

Thiophenyl **O-{2-[[[(2,2,2-trichloroethoxy)carbonyl]amino]-3,4,6-O-acetyl- β -D-glucopyranosyl}-(1 \rightarrow 3)-(4-O-benzyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-{(2-[[[(2,2,2-trichloroethoxy)carbonyl]amino]-3,4,6-O-acetyl- β -D-glucopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl- α -L-rhamnopyranosyl}-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (17).**

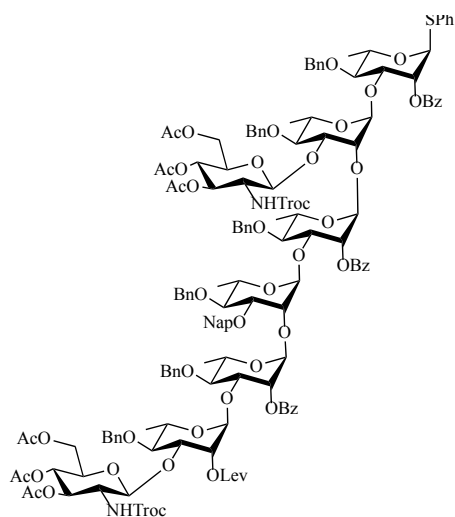
Compound **17** was derived from the optimal glycosylation method as described in the General Procedures. Yield: 93%. ^1H NMR (600 MHz, chloroform-*d*) δ 8.10 (m, 4H, arom. H, Bz), 7.71 – 7.01 (m, 31H, arom. H, Bn, Bz, SPh), 5.63 (m, 1H, H-2,A), 5.56 – 5.50 (m, 2H, H-1,A, H-2,C), 5.44 (m, 2H, H-1,C, H-2,D), 5.27 – 5.18 (m, 2H, H-1,D, H-1,B), 5.03 (m, 2H, H-3,5G, H-4,5G), 4.95 (m, 2H, H-3,3G, H-4,3G), 4.91 – 4.44 (m, 13H, 4xCH₂, Bn, 2xCH₂, Troc, H-1,5G), 4.31 – 4.20 (m, 2H, H-3,C, H-5,C), 4.20 – 4.11 (m, 3H, H-3,A, H-2,B, H-6,3G), 4.06 – 3.77 (m, 7H, H-3,D, H-6,5G, H-6,3G, H-5,A, H-3,B, H-5,D, H-2,3G), 3.76 – 3.44 (m, 6H, H-5,B, H-4,A, H-4,C, H-5,3G, H-2,5G, H-6,5G), 3.38 (m, 2H, H-4,B, H-4,D), 3.28 (m, 1H, H-5,5G), 3.15 (m, 1H, H-2,3G), 2.80 – 2.48 (m, 4H, CH₂-CH₂, Lev), 2.16 (s, 3H, CH₃, Lev), 2.03 – 1.90 (m, 18H, 6xCH₃, Ac), 1.29 – 1.15 (m, 6H, Rha, CH₃,A,C), 1.11 (m, 6H, Rha, CH₃,B,D). ^{13}C NMR (151 MHz, chloroform-*d*) δ 206.64, 170.55, 169.31, 153.83, 153.38, 138.03, 137.73, 133.69, 133.38, 131.85, 129.96, 129.84, 129.75, 129.05, 128.65, 128.56, 128.54, 128.46, 128.36, 128.26, 127.93, 127.89, 127.86, 127.80, 127.64, 127.58, 127.38, 127.08, 126.97, 100.91, 100.85, 100.69, 98.85, 95.46, 85.62, 80.09, 79.84, 77.99, 75.26, 74.41, 74.28, 74.03, 73.71, 72.39, 71.79, 71.41, 71.16, 70.79, 69.14, 68.90, 68.41, 68.03, 61.22, 56.56, 56.21, 37.95, 29.83, 28.00, 20.62, 20.59, 20.56, 17.87, 17.66. HR-MALDI-TOF/MS (*m/z*): calcd for C₁₀₇H₁₂₀Cl₆N₂NaO₃₆S⁺ [M+Na]⁺: 2273.5421; found, 2273.5424.



14

Thiophenyl O-(4-O-benzyl-3-O-Naphthylmethyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-{(2-[(2,2,2-trichloroethoxy)carbonyl]amino]-3,4,6-O-acetyl- β -D-glucopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl- α -L-rhamnopyranosyl)}-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside(14).

Compound **14** was derived from the general method for Lev deprotection as described in the General Procedures. Yield: 76%. ^1H NMR (600 MHz, chloroform-*d*) δ 8.14 – 7.98 (m, 4H, arom. H, Bz), 7.80 – 7.04 (m, 38H, arom. H, Bn, Nap, Bz, SPh), 5.61 (m, 1H, H-2,A), 5.56 – 5.50 (m, 3H, H-1,A, H-3,3G, H-2,C), 5.21 (d, 1H, H-1,C), 5.12 (d, 1H, H-1,D), 5.05 – 4.99 (m, 2H, H-1,B, H-4,3G), 4.96 (m, 1H, H-1,3G), 4.92 – 4.45 (m, 12H, 4xCH₂, Bn, CH₂, Nap, CH₂, Troc), 4.28 – 4.21 (m, 2H, H-3,D, H-5,A), 4.17 (m, 1H, H-2,D), 4.15 – 4.07 (m, 2H, H-3,A, H-2,B), 3.98 (m, 2H, H-3,B, H-5,C), 3.90 – 3.73 (m, 2H, H-3,C, H-5,D), 3.70 (m, 1H, H-2,C), 3.66 – 3.47 (m, 2H, H-4,D, H-5,B), 3.47 – 3.35 (m, 2H, H-4,A, H-4,B), 3.32 (m, 1H, H-5,3G), 1.94 (m, 6H, 2xCH₃, Ac), 1.63 (s, 3H, CH₃, Ac), 1.28 – 1.23 (m, 6H, Rha, CH₃,A,C), 1.15 (m, 6H, Rha, CH₃,B,D).HR-MALDI-TOF/MS (*m/z*): calcd for C₉₈H₁₀₄Cl₃NNaO₂₆S⁺ [M+Na]⁺: 1870.5521; found, 1870.5526.

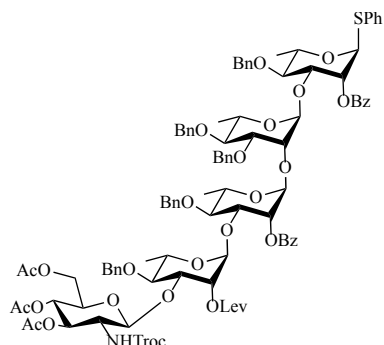


18

Thiophenyl

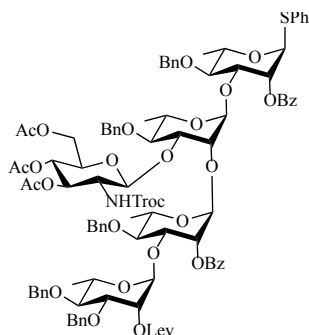
O-{2-[[[(2,2,2-trichloroethoxy)carbonyl]amino]-3,4,6-O-acetyl- β -D-glucopyranosyl}-(1 \rightarrow 3)-(4-O-benzyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(4-O-benzyl-3-O-Naphthylmethyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-{2-[[[(2,2,2-trichloroethoxy)carbonyl]amino]-3,4,6-O-acetyl- β -D-glucopyranosyl}-(1 \rightarrow 3)-(4-O-benzyl- α -L-rhamnopyranosyl)}-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside

(18). Compound 18 was derived from the optimal glycosylation method as described in the General Procedures. Yield: 23%. ^1H NMR (600 MHz, chloroform-*d*) δ 8.07 (m, 6H, arom. H, Bz), 7.68 – 6.92 (m, 51H, arom. H, Bn, Nap, SPh), 5.59 (m, 3H, H-2,A, H-2,E, H-1,A), 5.51 (m, 1H, H-2,F), 5.29 (d, J = 2.6 Hz, 1H, H-1,F), 5.19 (m, 1H, H-2,C), 5.13 – 5.00 (m, 4H, H-3,5G, H-4,5G, H-1,C, H-1,3G), 5.00 – 4.33 (m, 18H, 6xCH₂, Bn, CH₂, Nap, H-1,D, H-3,3G, H-4,3G, H-1,5G), 4.32 – 4.05 (m, 6H, H-3,E, H-5,A, H-3,F, H-3,A, H-4,C, H-5,C), 4.03 – 3.25 (m, 15H, H-3,C, H-5,E, H-3,D, H-3,B, H-5,F, H-6,3G, H-6,5G, H-5,B, H-4,B, H-5,D, H-4,A, H-4,E, H-4,D, H-4,F, H-2,D), 3.09 (m, 1H, H-2,B), 2.19 – 2.10 (m, 4H, CH₂-CH₂, Lev), 2.01 – 1.89 (m, 21H, CH₃, Lev, 6xCH₃, Ac), 1.24 (m, 9H, Rha, CH₃,A,C,E), 1.17 – 1.03 (m, 9H, Rha, CH₃,B,D,F). HR-MALDI-TOF/MS (m/z): calcd for C₁₅₁H₁₆₄Cl₆N₂NaO₄₅S⁺ [M+Na]⁺: 2994.6925; found, 2994.6927.



Thiophenyl

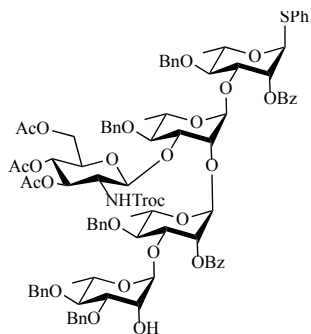
O-(2-[[[(2,2,2-trichloroethoxy)carbonyl]amino]-3,4,6-O-acetyl- β -D-glucopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(3,4-O-benzyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (12). Compound **12** was derived from the optimal glycosylation method as described in the General Procedures. Yield: 31%. ^1H NMR (600 MHz, chloroform-*d*) δ 8.18 – 7.93 (m, 4H, arom. H, Bz), 7.70 – 6.87 (m, 36H, arom. H, Bn, SPh), 5.64 – 5.55 (m, 1H, H-2,A), 5.55 – 5.49 (d, 1H, H-1,A), 5.26 – 5.18 (m, 1H, H-2,C), 5.05 (d, $J = 7.1$ Hz, 1H, H-1,B), 5.00 – 4.79 (m, 4H, H-1,C, H-1,D, H-3,5G, H-4,5G), 4.81 – 4.34 (m, 11H, 5 \times CH₂, Bn, H-1,5G), 4.33 – 4.19 (m, 3H, H-3,C, H-5,C, H-3,A), 4.02 – 3.75 (m, 4H, H-5,B, H-4,A, H-5,A, H-3,D), 3.77 – 3.45 (m, 5H, H-3,B, H-2,B, H-6,5G, H-4,A, H-5,D), 3.43 – 3.27 (m, 2H, H-4,B, H-5,5G), 2.76 – 2.46 (m, 4H, CH₂-CH₂, Lev), 2.15 (s, 3H, CH₃, Lev), 2.06 – 1.89 (m, 9H, 3 \times CH₃, Ac), 1.29 – 1.18 (m, 6H, Rha, CH₃,A,C), 1.19 – 0.97 (m, 6H, Rha, CH₃,B,D). HR-MALDI-TOF/MS (m/z): calcd for C₉₉H₁₀₈Cl₃NNaO₂₈S⁺ [M+Na]⁺: 1918.5721; found, 1918.5725.



15

Thiophenyl O-(3,4-O-benzyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-{2-[[[(2,2,2-trichloroethoxy)carbonyl]amino]-3,4,6-O-acetyl- β -D-glucopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl- α -L-rhamnopyranosyl)}-(1 \rightarrow 3)-4-

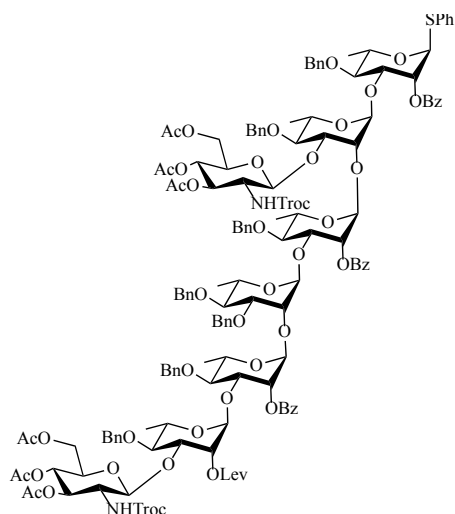
O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (15). Compound **15** was derived from the general method for Lev deprotection as described in the General Procedures. Yield: 82%. ^1H NMR (600 MHz, chloroform-*d*) δ 8.14 – 7.95 (m, 4H, arom. H, Bz), 7.72 – 7.02 (m, 36H, arom. H, Bn, Bz, SPh), 5.62 (m, 1H, H-2,A), 5.58 – 5.51 (m, 2H, H-1,A, H-2,D), 5.51 – 5.46 (m, 1H, H-2,C), 5.39 (m, 1H, H-1,D), 5.25 (m, 1H, H-1,C), 5.08 (d, J = 1.8 Hz, 1H, H-1,B), 5.05 – 4.94 (m, 2H, H-1,3G, H-4,3G), 4.91 – 4.34 (m, 12H, 5xCH₂, Bn, CH₂, Troc), 4.34 – 4.07 (m, 6H, H-3,D, H-5,A, H-3,A, H-2,B, H-3,B, H-5,C), 3.92 – 3.67 (m, 6H, H-5,D, H-3,C, H-5,B, H-4,A, H-4,D, H-2,3G), 3.67 – 3.48 (m, 1H, H-6,3G), 3.40 (m, 1H, H-4,B), 3.37 – 3.27 (m, 1H, H-5,3G), 2.74 – 2.52 (m, 4H, CH₂-CH₂, Lev), 2.12 (s, 3H, CH₃, Lev), 1.95 (s, 6H, 2xCH₃, Ac), 1.62 (s, 3H, CH₃, Ac), 1.33 – 1.18 (m, 6H, Rha, CH₃,A,C), 1.13 (m, 6H, Rha, CH₃,B,D). HR-MALDI-TOF/MS (m/z): calcd for C₉₉H₁₀₈Cl₃NNaO₂₈S⁺ [M+Na]⁺: 1918.5721; found, 1918.5727.



16

Thiophenyl O-(3,4-O-benzyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-{(2-[[2,2,2-trichloroethoxy]carbonyl]amino]-3,4,6-O-acetyl- β -D-glucopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl- α -L-rhamnopyranosyl)}-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (16). Compound **16** was derived from the general method for Lev deprotection as described in the General Procedures. Yield: 94%. ^1H NMR (900 MHz, chloroform-*d*) δ 8.18 – 7.94 (m, 4H, arom. H, Bz), 7.67 – 7.02 (m, 36H, arom. H, Bn, Bz, SPh),

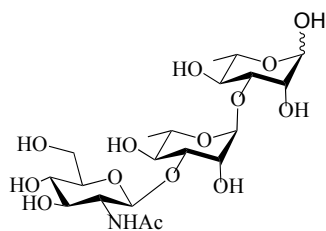
5.63 (m, 1H, H-2,A), 5.59 – 5.51 (m, 3H, H-1,A, H-2,C, H-3,3G), 5.24 (m, 1H, H-3,A), 5.15 – 5.11 (m, 1H, H-4,A), 5.07 – 5.00 (m, 2H, H-1,B, H-1,3G), 4.97 (m, 1H, H-4,3G), 4.91 – 4.43 (m, 12H, 5xCH₂, Bn, CH₂, Troc), 4.30 – 4.10 (m, 5H, H-3,D, H-5,A, H-2,D, H-3,A H-2,B), 4.02 – 3.59 (m, 5H, H-3,B, H-5,C, H-3,C, H-5,D, H-5,B), 3.58 – 3.28 (m, 6H, H-2,3G, H-4,D, H-4,A, H-6,3G, H-4,B, H-5,3G), 2.01 – 1.87 (m, 9H, 3×CH₃, Ac), 1.36 – 1.20 (m, 6H, Rha, CH₃,A,C), 1.22 – 1.07 (m, 6H, Rha, CH₃,B,D). ¹³C NMR (226 MHz, chloroform-*d*) δ 170.35, 170.02, 169.42, 165.84, 165.60, 138.45, 138.33, 138.06, 137.83, 137.69, 133.72, 133.39, 133.35, 131.86, 129.91, 129.86, 129.82, 129.77, 129.76, 129.05, 129.01, 128.57, 128.56, 128.55, 128.52, 128.49, 128.45, 128.41, 128.29, 128.25, 128.21, 127.97, 127.94, 127.87, 127.84, 127.75, 127.74, 127.73, 127.64, 127.57, 127.39, 127.13, 101.10, 100.98, 100.43, 98.71, 85.64, 80.25, 80.17, 79.64, 79.59, 79.43, 79.01, 78.61, 78.11, 77.96, 77.20, 75.38, 75.35, 74.89, 74.33, 73.72, 72.70, 71.98, 71.43, 71.18, 70.68, 69.16, 69.00, 68.92, 68.63, 68.39, 68.31, 61.19, 56.63, 29.69, 20.62, 20.58, 20.24, 18.09, 17.87, 17.77, 17.68, 1.02. HR-MALDI-TOF/MS (m/z): calcd for C₉₄H₁₀₂Cl₃NNaO₂₆S⁺ [M+Na]⁺: 1820.5421; found, 1820.5427.



19

Thiophenyl O-{2-[[[(2,2,2-trichloroethoxy)carbonyl]amino]-3,4,6-O-acetyl- β -D-glucopyranosyl}-(1 \rightarrow 3)-(4-O-benzyl-2-O-levulinoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-(3,4-O-benzyl- α -L-rhamnopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranosyl)-(1 \rightarrow 2)-{2-[[[(2,2,2-trichloroethoxy)carbonyl]amino]-3,4,6-O-acetyl- β -D-glucopyranosyl)-(1 \rightarrow 3)-(4-O-benzyl- α -L-rhamnopyranosyl)}-(1 \rightarrow 3)-4-O-benzyl-2-O-benzoyl- α -L-rhamnopyranoside (19).

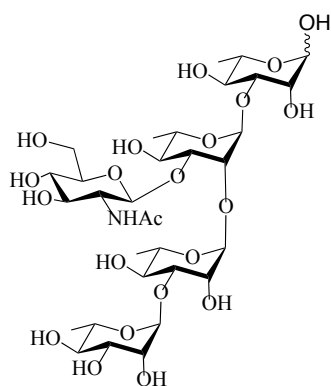
Compound **19** was derived from the optimal glycosylation method as described in the General Procedures. Yield: 43% in DCM, 37% in toluene. ^1H NMR (900 MHz, chloroform-*d*) δ 8.25 – 7.87 (m, 6H, arom. H, Bz), 7.74 – 6.94 (m, 49H, arom. H, Bn, SPh, Bz), 5.53 (m, 2H, H-2,A, H-2,E), 5.21 (d, 1H, H-1,A), 5.16 – 4.82 (m, 10H, H-1,F, H-1,C, H-1,B, H-1,E, H-1,3G, H-3,5G, H-4,5G, H-2,C, H-3,3G, H-4,3G), 4.83 – 4.30 (m, 17H, 7xCH₂, Bn, CH₂, Troc, H-1,5G), 4.34 – 4.09 (m, 2H, H-3,E, H-5,A), 4.07 – 3.01 (m, 23H, H-5,B, H-3,5G, H-4,B, H-4,A, H-5,D, H-3,F, H-3,A, H-4,C, H-5,C, H-6,3G, H-6,5G, H-3,C, H-5,E, H-3,D, H-2,B, H-5,F, H-3,B, H-4,E, H-4,A, H-2,D, H-5,5G, H-4,D, H-4,F), 2.67 (m, 4H, CH₂-CH₂, Lev), 2.26 – 2.05 (s, 3H, CH₃, Lev), 2.03 – 1.84 (m, 18H, 6xCH₃, Ac), 1.52 – 1.16 (m, 9H, Rha, CH₃,A,C,E), 1.20 – 0.78 (m, 9H, Rha, CH₃,B,D,F). HR-MALDI-TOF/MS (m/z): calcd for C₁₄₇H₁₆₂Cl₆N₂NaO₄₅S⁺ [M+Na]⁺: 2944.6321; found, 2944.6327.



20

β -D-(Acetylamino)-2-deoxy-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnosyl-(1 \rightarrow 3)- α -L-

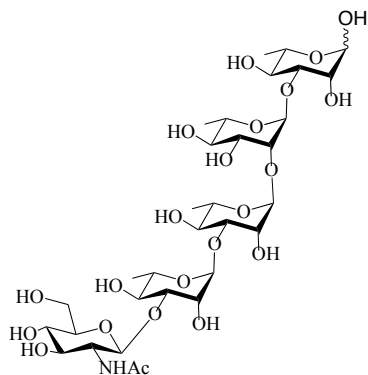
rhamnopyranose (20). Compound **20** was derived from the global deprotection method as described in the General Procedures. Yield: 90%. ^1H NMR (600 MHz, deuterium oxide) δ 4.89 (m, 1H, H-1,A), 4.80 – 4.74 (m, 1H, H-1,B), 4.54 (m, 1H, H-1,C), 4.07 (m, 1H, H-2,B), 3.85 – 3.52 (m, 7H, H-2,A, H-3,A, H-5,A, H-5,3G, H-2,3G, H-5,B, H-5,3G), 3.51 – 3.22 (m, 4H, H-3,3G, H-4,B, H-4,A, H-4,3G), 1.91 – 1.82 (m, 3H, NHAc), 1.11 (m, 6H, Rha, CH₃,A,B). ^{13}C NMR (151 MHz, deuterium oxide) δ 174.81, 157.38, 102.68, 102.59, 101.81, 93.83, 79.79, 78.88, 77.59, 75.52, 75.47, 74.49, 73.53, 72.11, 71.48, 70.75, 70.70, 70.48, 70.07, 69.65, 69.40, 69.17, 68.33, 66.49, 62.65, 60.44, 55.61, 30.11, 22.05, 19.08, 16.66, 16.44, 16.39. HR-MALDI-TOF/MS (m/z): calcd for $\text{C}_{20}\text{H}_{35}\text{NNaO}_{14}^+$ [M+Na] $^+$: 536.1911; found, 536.1917.



21

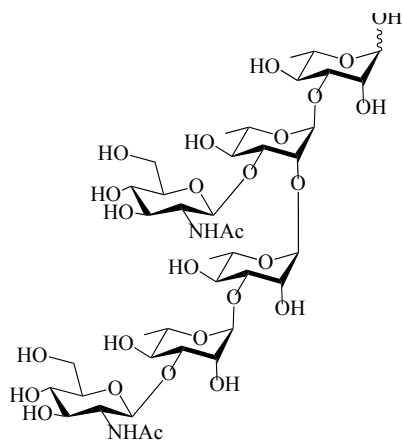
α -L-rhamnosyl-(1 \rightarrow 3)- α -L-rhamnosyl-(1 \rightarrow 2)-[β -D-(Acetylamino)-2-deoxy-glucopyranosyl-(1 \rightarrow 3)]- α -L-rhamnosyl-(1 \rightarrow 3)- α -L-rhamnopyranose (21). Compound **21** was derived from the global deprotection method as described in the General Procedures. Yield: 92% ^1H NMR (600 MHz, deuterium oxide) δ 5.03 (m, 2H, H-1,A, H-1,B), 4.88 (m, 2H, H-1,B, H-1,D), 4.53 (d, 1H, H-1,3G), 4.15 (m, 2H, H-2,B, H-2,D), 3.97 – 3.88 (m, 2H, H-2,A, H-2,C), 3.87 – 3.46 (m, 10H, H-3,A, H-3,B, H-3,C, H-3,D, H-5,A, H-5,3G, H-5,C, H-5,B, H-2,3G, H-5,3G), 3.46 – 3.09 (m, 7H, H-4,A, H-4,C, H-4,D, H-3,2G, H-4,B, H-4,C, H-5,D), 1.20 – 1.05 (m, 12H, Rha, CH₃,A-

D). HR-MALDI-TOF/MS (m/z): calcd for $C_{32}H_{55}NNaO_{22}^+$ [M+Na] $^+$: 828.3131; found, 828.3137.



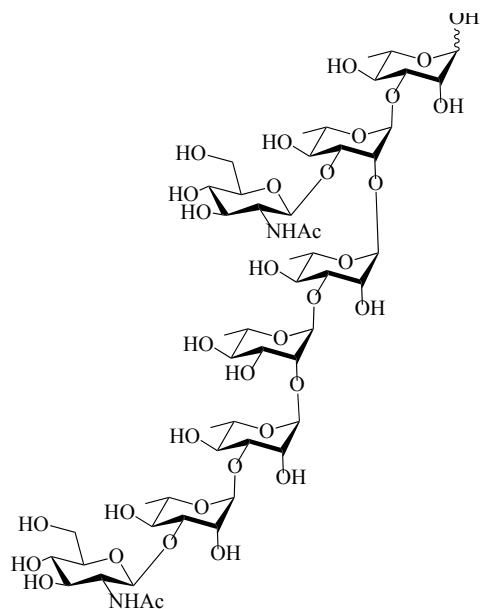
22

β -D-(Acetylamino)-2-deoxy-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnosyl-(1 \rightarrow 3)- α -L-rhamnosyl-(1 \rightarrow 2)- α -L-rhamnosyl-(1 \rightarrow 3)- α -L-rhamnopyranose (22). Compound **22** was derived from the global deprotection method as described in the General Procedures. Yield: 89% 1H NMR (600 MHz, deuterium oxide) δ 5.06 (m, 2H, H-1,A, H-1,C), 4.90 – 4.85 (m, 2H, H-1,B, H-1,D), 4.55 (m, 1H, H-1,5G), 4.12 (m, 2H, H-2,B, H-2,D), 3.95 – 3.89 (m, 2H, H-2,A, H-2,C), 3.85 – 3.46 (m, 11H, H-3,A, H-3,B, H-3,C, H-3,D, H-5,A, H-5,5G, H-5,C, H-2,5G, H-5,B, H-4,A, H-5,5G), 3.46 – 3.16 (m, 6H, H-4,5G, H-4,D, H-3,5G, H-4,B, H-4,C, H-5,D), 1.93 – 1.81 (m, 3H, NHAc), 1.23 – 1.05 (m, 12H, Rha, CH₃,A-D). ^{13}C NMR (151 MHz, deuterium oxide) δ 174.80, 102.60, 101.98, 101.85, 100.63, 100.49, 93.80, 93.25, 79.79, 79.35, 77.98, 77.91, 76.94, 75.52, 73.55, 72.03, 71.84, 71.68, 71.52, 71.21, 71.02, 70.71, 70.52, 69.77, 69.70, 69.68, 69.65, 69.24, 69.09, 68.22, 61.31, 60.45, 55.61, 33.48, 30.11, 28.11, 22.05, 16.58, 16.51. HR-MALDI-TOF/MS (m/z): calcd for $C_{32}H_{55}NNaO_{22}^+$ [M+Na] $^+$: 828.3131; found, 828.3135.



23

β -D-(Acetylamino)-2-deoxy-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnosyl-(1 \rightarrow 3)- α -L-rhamnosyl-(1 \rightarrow 2)-[β -D-(Acetylamino)-2-deoxy-glucopyranosyl-(1 \rightarrow 3)]- α -L-rhamnosyl-(1 \rightarrow 3)- α -L-rhamnopyranose (23). Compound **23** was derived from the global deprotection method as described in the General Procedures. Yield: 94% ^1H NMR (600 MHz, deuterium oxide) δ 5.03 (m, 2H, H-1,A, H-1,C), 4.88 (m, 2H, H-1,B, H-1,D), 4.59 – 4.53 (m, 2H, H-1,3G, H-1,6G), 4.18 – 4.09 (m, 2H, H-2,B, H-2,D), 3.85 – 3.46 (m, 17H, H-3,A, H-3,B, H-3,C, H-3,D, H-5,A, H-5,3G, H-5,6G, H-5,C, H-2,3G, H-2,6G, H-5,B, H-4,A, H-5,3G, H-5,6G, H-4,6G, H-4,3G, H-4,D), 3.47 – 3.17 (m, 5H, H-3,6G, H-3,3G, H-4,B, H-4,C, H-5,D), 1.87 (m, 3H, NHAc), 1.19 – 1.05 (m, 12H, Rha, CH₃,A-D). ^{13}C NMR (151 MHz, deuterium oxide) δ 174.81, 174.37, 126.46, 102.64, 93.82, 79.65, 76.98, 75.59, 75.51, 73.77, 73.55, 71.82, 70.75, 70.50, 69.71, 69.64, 69.26, 68.26, 60.44, 55.75, 55.61, 22.10, 22.06, 16.67, 16.58, 16.52, 16.42. HR-MALDI-TOF/MS (m/z): calcd for $\text{C}_{40}\text{H}_{68}\text{N}_2\text{NaO}_{27}^+$ [M+Na] $^+$: 1031.3931; found, 1031.3937.



24

β -D-(Acetylamino)-2-deoxy-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnosyl-(1 \rightarrow 3)- α -L-rhamnosyl-(1 \rightarrow 2)- α -L-rhamnosyl-(1 \rightarrow 3)- α -L-rhamnosyl-(1 \rightarrow 2)-[β -D-(Acetylamino)-2-deoxy-glucopyranosyl-(1 \rightarrow 3)]- α -L-rhamnosyl-(1 \rightarrow 3)- α -L-rhamnopyranose (24). Compound 24 was derived from the global deprotection method as described in the General Procedures. Yield: 91%. ^1H NMR (900 MHz, deuterium oxide) δ 5.05 (m, 3H, H-1,A, H-1,C, H-1,E), 4.93 – 4.85 (m, 3H, H-1,B, H-1,D, H-1,F), 4.58 – 4.55 (m, 2H, H-1,3G, H-1,6G), 4.14 (m, 3H, H-2,B, H-2,D, H-2,F), 3.98 – 3.54 (m, 13H, H-2,A, H-2,C, H-2,E, H-3,A, H-3,B, H-3,C, H-3,D, H-3,E, H-3,F, H-5,A, H-5,3G, H-5,C, H-5,6G), 3.54 – 3.10 (m, 17H, H-2,3G, H-2,6G, H-5,B, H-4,A, H-5,3G, H-4,C, H-4,D, H-4,E, H-4,F, H-3,3G, H-3,6G, H-4,B, H-4,C, H-5,D, H-5,E, H-5,F, H-4,6G), 1.93 – 1.85 (m, 3H, NHAc), 1.14 (m, 18H, Rha, CH₃,A-F). HR-MALDI-TOF/MS (m/z): calcd for C₅₂H₈₈N₂NaO₃₅⁺ [M+Na]⁺: 1323.5123; found, 1323.5127.

References

- 1) Carapetis R.; Steer C.; Mulholland K.; Weber M. The global burden of group A streptococcal diseases. *Lancet Infectious Diseases*. **2005**, 5-11, 685–94.
- 2) Nowak R. Flesh-eating bacteria: not new but still worrisome. *Science*, **1994**, 264, 1665.
- 3) Read E.; Zabriskie B. Streptococcal Diseases and The Immune Response. *Academic Press*, **1977**.
- 4) Stevens L. The toxins of group A streptococcus, the flesh eating bacteria. *Immunological Investigations*, **1997**, 26, 129.
- 5) Weiss A.; Laverdiere, M. Group A Streptococcus invasive infections: a review. *Canadian Journal of Surgery*, **1997**, 40, 18.
- 6) World Health Organization. Community control of rheumatic heart disease in developing countries. *Weekly Epidemiological Records*, **1981**.
- 7) Mandell L.; Douglas G.; Bennet E. Principles and Practice of Infectious Disease. *Annales de l'Institut Pasteur/Microbiologie*, **1985**, 1133.
- 8) Markowitz M. Observations on the epidemiology and preventability of rheumatic fever in developing countries. *Clinical Therapeutics*, **1981**, 4, 240.
- 9) Wannamaker W. Changes and changing concepts in the biology of group A streptococci and in the epidemiology of streptococcal infections. *Reviews of Infectious Diseases*, **1979**, 1, 967.
- 10) Zabriskie B. Rheumatic fever: the interplay between host, genetics, and microbe. Lewis A. Conner memorial lecture. *Circulation*, **1985**, 71, 1077.
- 11) Kaplan M.; Bolande R.; Rakita L.; Blair J. Presence of bound immunoglobulins and complement in the myocardium in acute rheumatic fever association with cardiac failure. *The New England Journal of Medicine*. **1964**, 13, 637–45.

- 12) Lozano R.; Naghavi M.; Foreman K.; Lim S.; Shibuya K.; Aboyans V. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. **2012**, 380, 2095–128.
- 13) Andrew S.; Jonathan C; James D; John F.; Michael G.; Luiza G.; Nicole M.; Kim M.; Florian S.; Pierre S. Status of research and development of vaccines for *Streptococcus pyogenes*. *Vaccine*. **2016**, 26, 2953–2958.
- 14) Michon F.; Moore L.; Kim J.; Blake S.; Auzanneau I.; Johnston D.; Johnson A.; Pinto M. Doubly branched hexasaccharide epitope on the cell wall polysaccharide of group A streptococci recognized by human and rabbit antisera. *Infectious Immunology*. **2005**, 73, 6383-6389.
- 15) Anna K.; Immaculada M.; Francesco B.; Maria R.; Guido G.; Giuliano B.; Emiliano C.; Daniela P.; Erwin S.; Emilia C.; Paola F.; Daniele C.; Roberto A.; Vittoria P.; David S.; Sabrina C.; Giada B.; Marilena G.; William C.; Stewart C.; John P.; Peter H. S.; Rino R.; Paolo C. Evaluation of a Group A *Streptococcus* synthetic oligosaccharide as vaccine candidate. *Vaccine*. **2010**, 1, 104–114.
- 16) Goldstein I.; Rebeyrotte P.; Parlebas J.; Halpern B. Isolation from heart valves of glycopeptides which share immunological properties with *Streptococcus haemolyticus* group A polysaccharides. *Nature*. **1968**, 219, 866–868.
- 17) Shulman T.; Ayoub M.; Victorica E.; Gessner H.; Tamer F.; Hernandez A. Differences in antibody response to streptococcal antigens in children with rheumatic and non-rheumatic mitral valve disease. *Circulation*. **1974**, 50, 1244–1251.

- 18) Appleton S.; Victorica E.; Tamer D.; Ayoub M. Specificity of persistence of antibody to the streptococcal group A carbohydrate in rheumatic valvular heart disease. *Journal of Laboratory and Clinical Medicine*. **1985**, 105, 114–119.
- 19) Galvin E.; Hemric E.; Ward K.; Cunningham W. Cytotoxic mAb from rheumatic carditis recognizes heart valves and laminin. *The Journal of Clinical Investigation*. **2000**, 106, 217–224.
- 20) Nina S.; Jason C.; Kirsten K.; Anna H.; Ramy A.; Ana K.; Leo L.; Evelien B.; Mark D.; Gordon D.; Fan Z.; Samira D.; Laura S.; Jennifer G.; Madeleine C.; Joseph M.; Julia H.; Bernd L.; Suzan R.; Richard M.; Mark W.; Sanford S.; Patrick S.; Biswa C.; Victor N. The classical lancefield antigen of Group A Streptococcus is a virulence determinant with implications for vaccine design. *Cell Host & Microbe*. **2014**, 15, 729–740.
- 21) Paz J; Horlacher T.; Peter S. Oligosaccharide Microarrays to Map Interactions of Carbohydrates in Biological Systems. *Methods Enzymology*, **2006**, 415, 269–292.
- 22) Oyelaran O.; Gildersleeve C. Glycan arrays: recent advances and future challenges. *Current Opinion in Chemical Biology*, **2009**, 13, 406–413.

CHAPTER 4

CONCLUSIONS

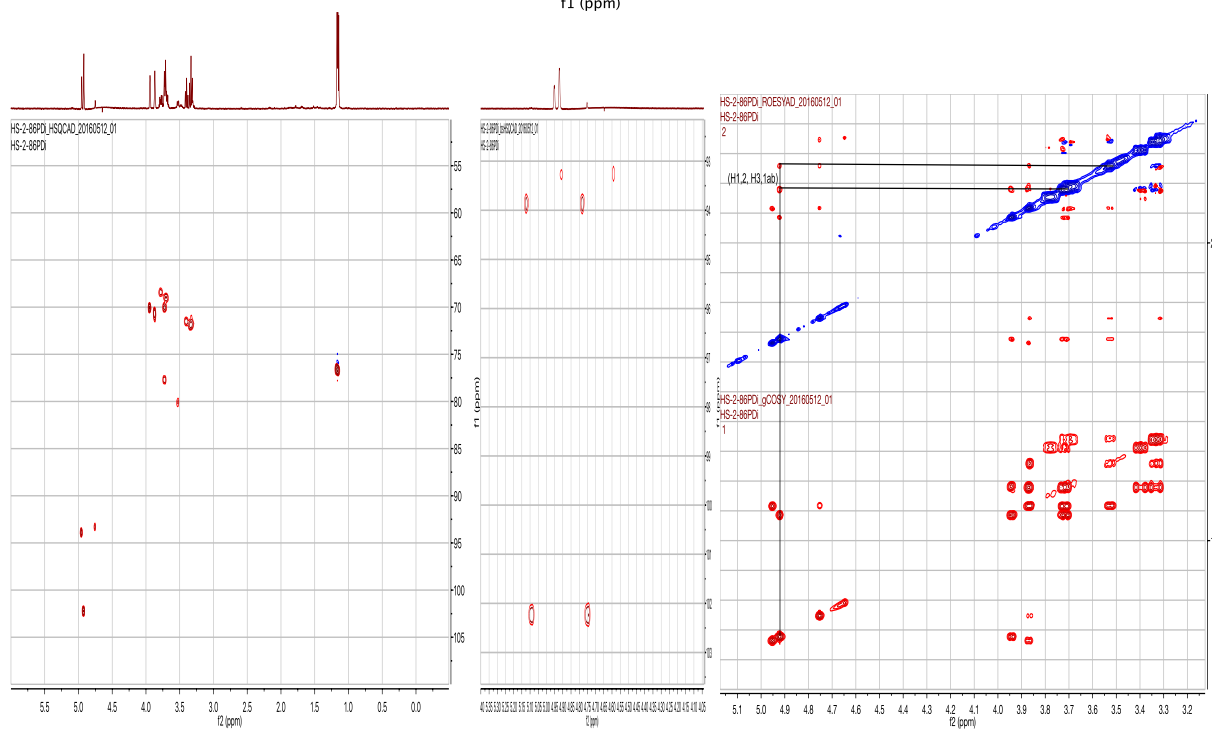
In the second chapter, a highly convergent synthetic strategy has been developed for the chemical synthesis of those promising vaccine antigens. We have described the thorough examining process of the “key disaccharide” synthesis, and the unexpected discoveries of “PMB migration” and the “4+4” *in situ* bond cleavage polymerization. We have established a comprehensive investigation process to address the challenges associated with those oligosaccharides, and proposed a “push and pull” mechanism for the “4+4” *in situ* bond cleavage polymerization. The synthetic oligosaccharides can be used for epitope mapping of different GAS antibodies and patient sera by glycan microarray. The glycan microarray data will provide essential insights for the autoimmunity associated with GAC antigens and lay the foundation for the development of new GAS vaccine candidates.

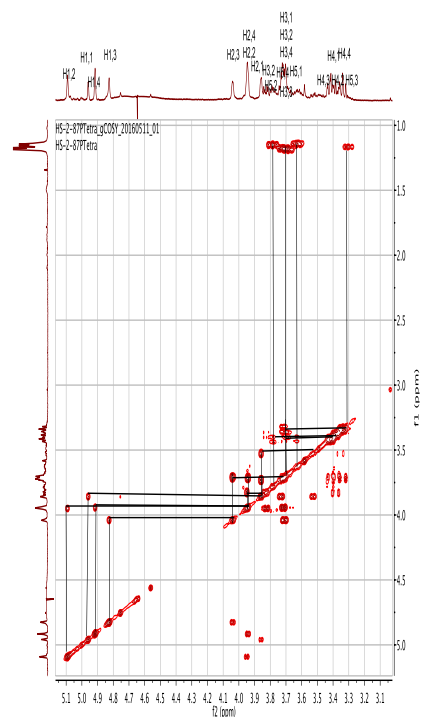
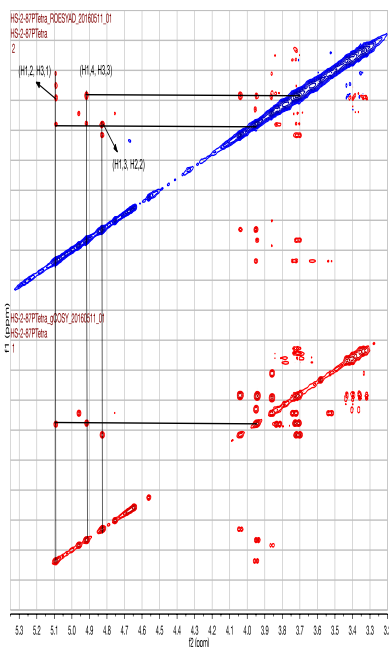
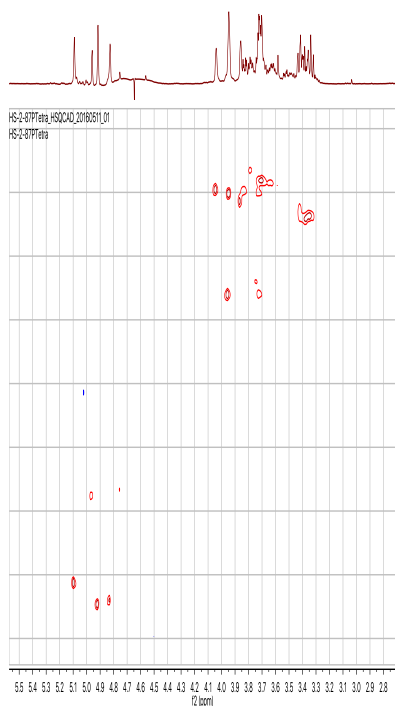
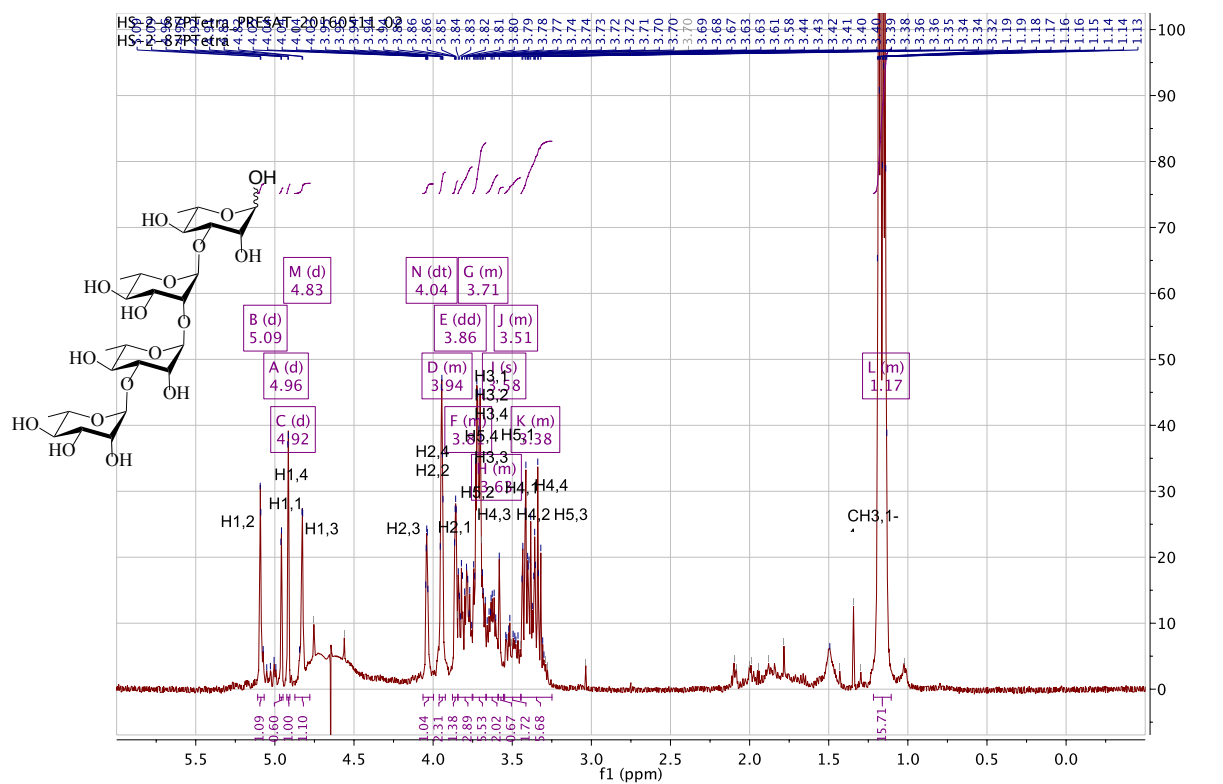
In the third chapter, we have established a comprehensive examining process for the key disaccharide modular strategy to assemble a library of GAC oligosaccharides with various lengths and GlcNAc side chain variations. We employed an orthogonally protected key disaccharide to facilitate the assembly process and built the library with the previously discovered “4+4” *in situ* polymerization. We utilized the glycan microarray technology to perform glycotype mapping with different type-specific antibodies, interpreted and developed a multi-GlcNAc “pocket-fitting” binding theory based on the binding results. The anti-polyrhamnose antibody, the anti-streptococcus pyogenes GAC antibody and the anti-Streptococcus Group A antibody exhibit distinct binding requirements according to the array

results. The anti-polyrhamnose antibody requires long-chain structure to have strong binding strength, and the binding strength maximized at double helical cycle dodecasaccharide level. The latter two antibodies require GlcNAc side chains on the periphery of the polyrhamnose backbone, and multiple GlcNAc side chains enhance the binding strength. The “patch” structure of the customized octasaccharide has the maximum binding strength, as the two adjacent GlcNAc side chains have different angles and distances that lead to better fitting and stronger binding to the antibody. The rabbit serum derived from immunization of the octa-polyrhamnose KLH-glycoconjugate recognizes not only polyrhamnose structures but also 40 and 42 structures with GlcNAc side chains. It could be explained by the polyrhamnose chain exposure within these structures. Future directions can focus on conducting further screening against cross-reactive patient sera to reveal the binding requirements for the auto-reactivity of the antibody derived from GAS infection patients with rheumatic fever complications. Another point of interest of the binding results can focus on the detailed crystallography study between the antibody and the synthetic GAC oligosaccharides to further clarify the antibody binding pocket interaction mechanisms.

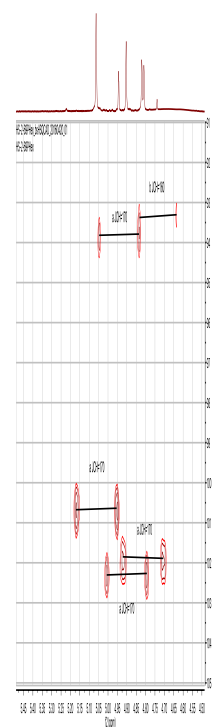
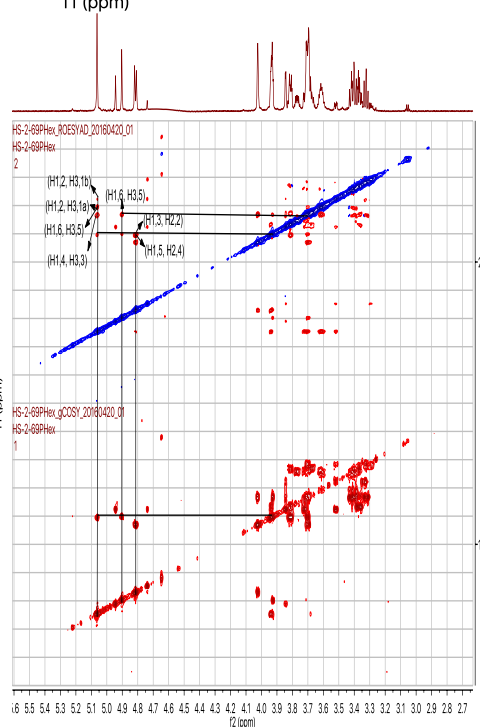
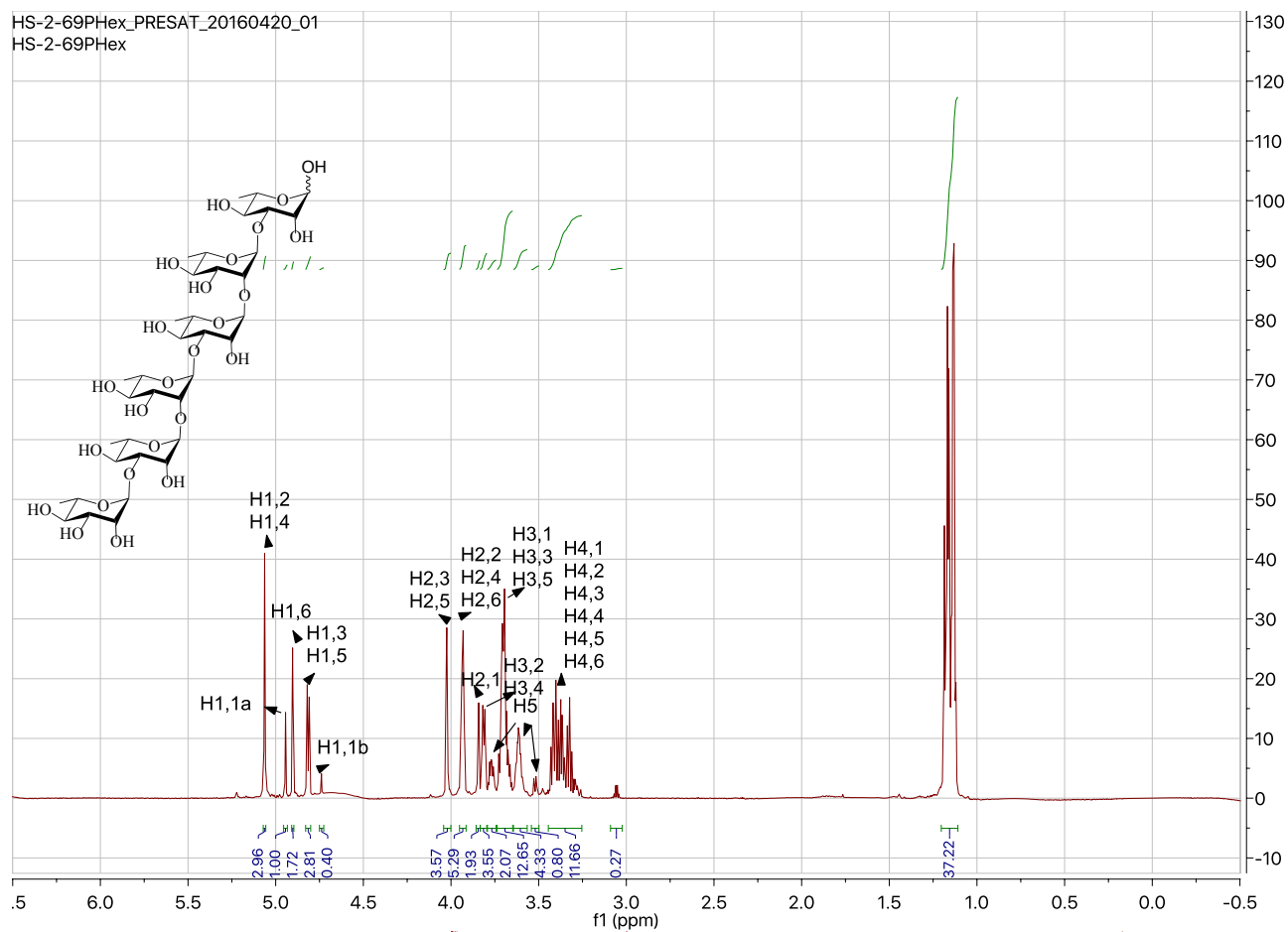
SUPPLEMENTARY DATA OF KEY PRODUCTS IN CHAPTER 2

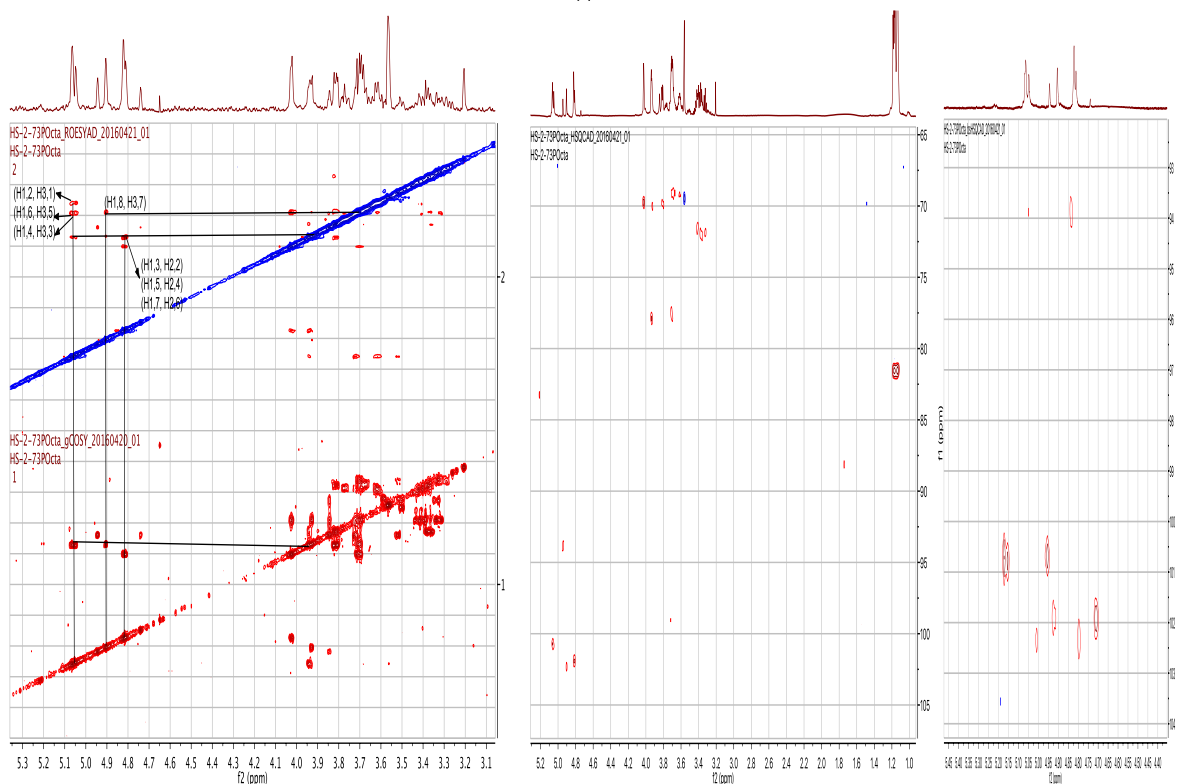
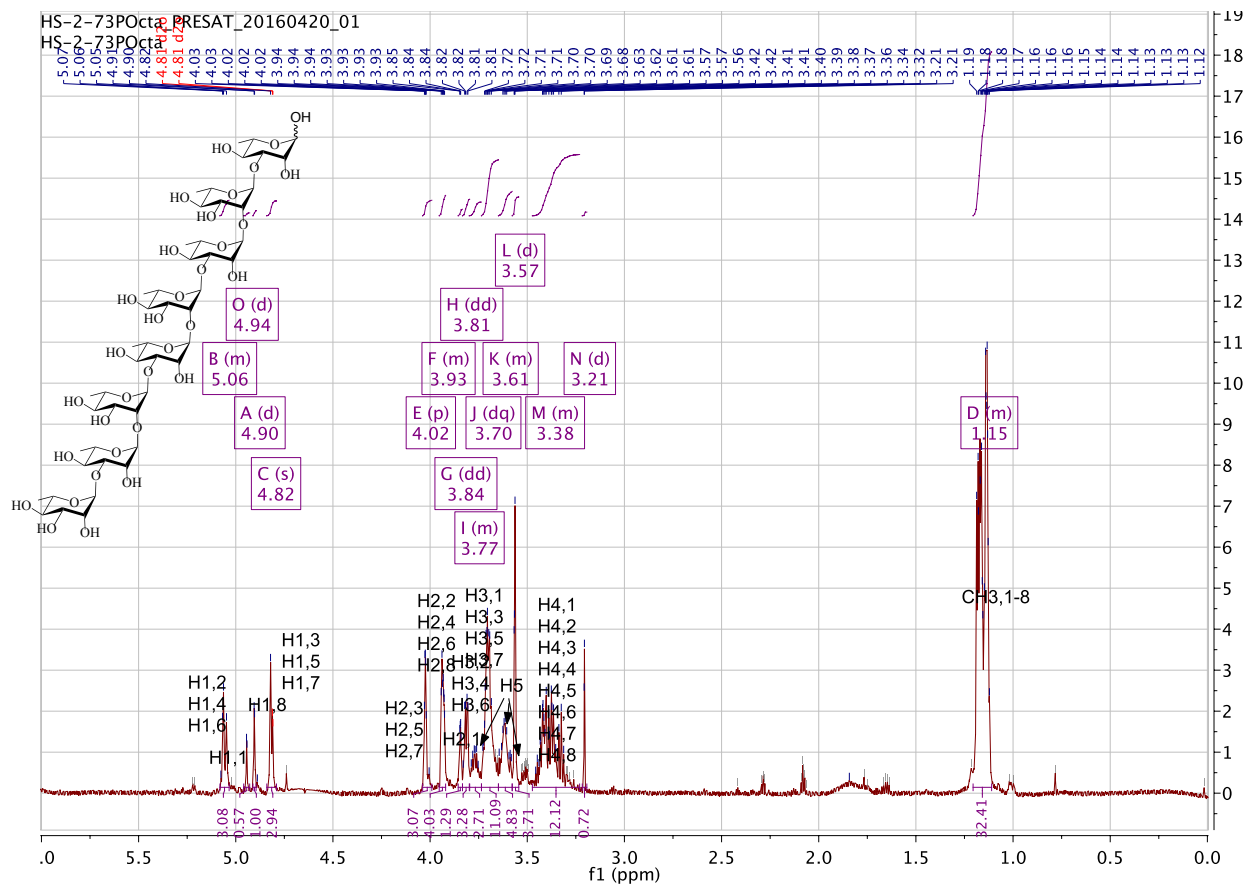


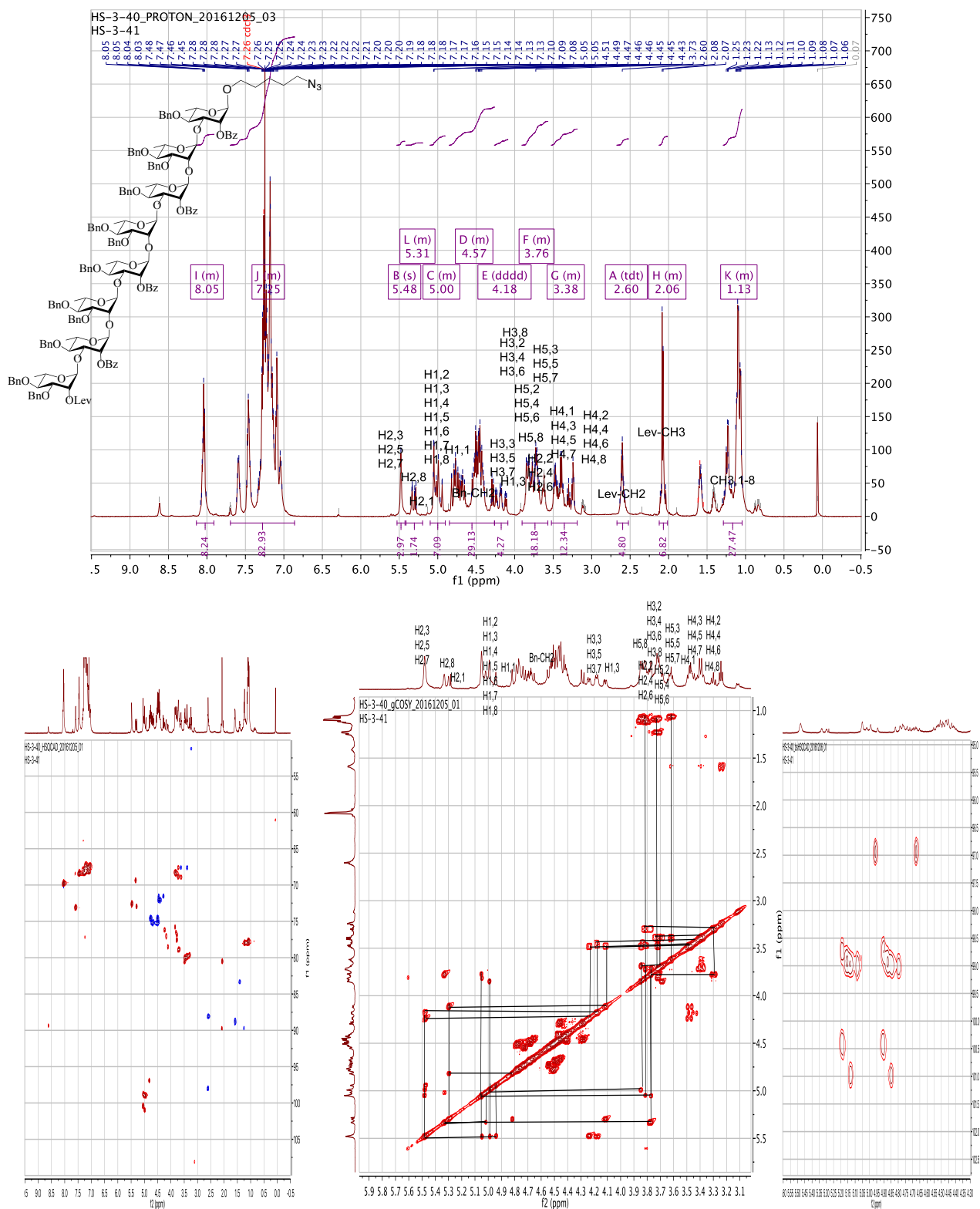


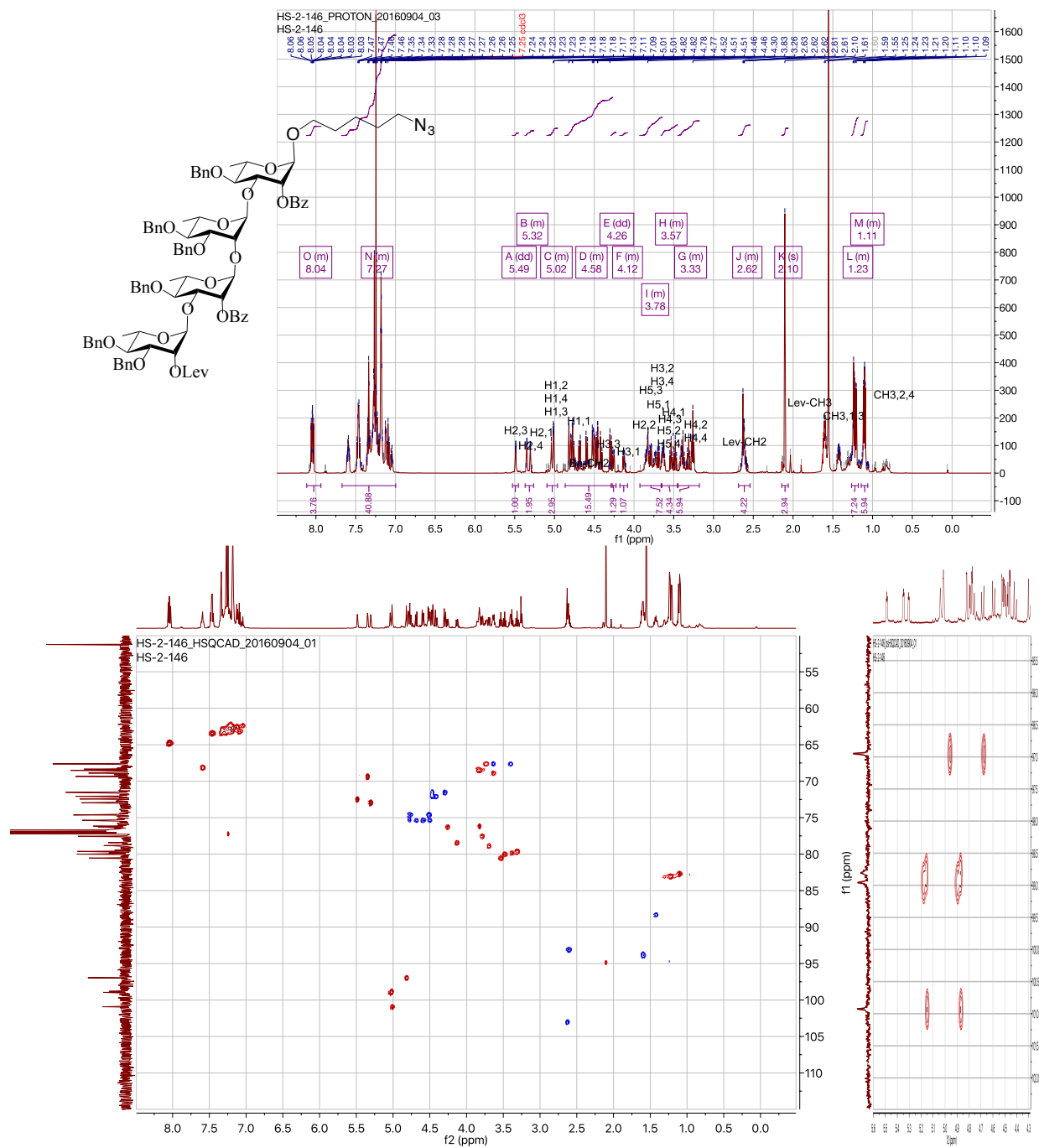


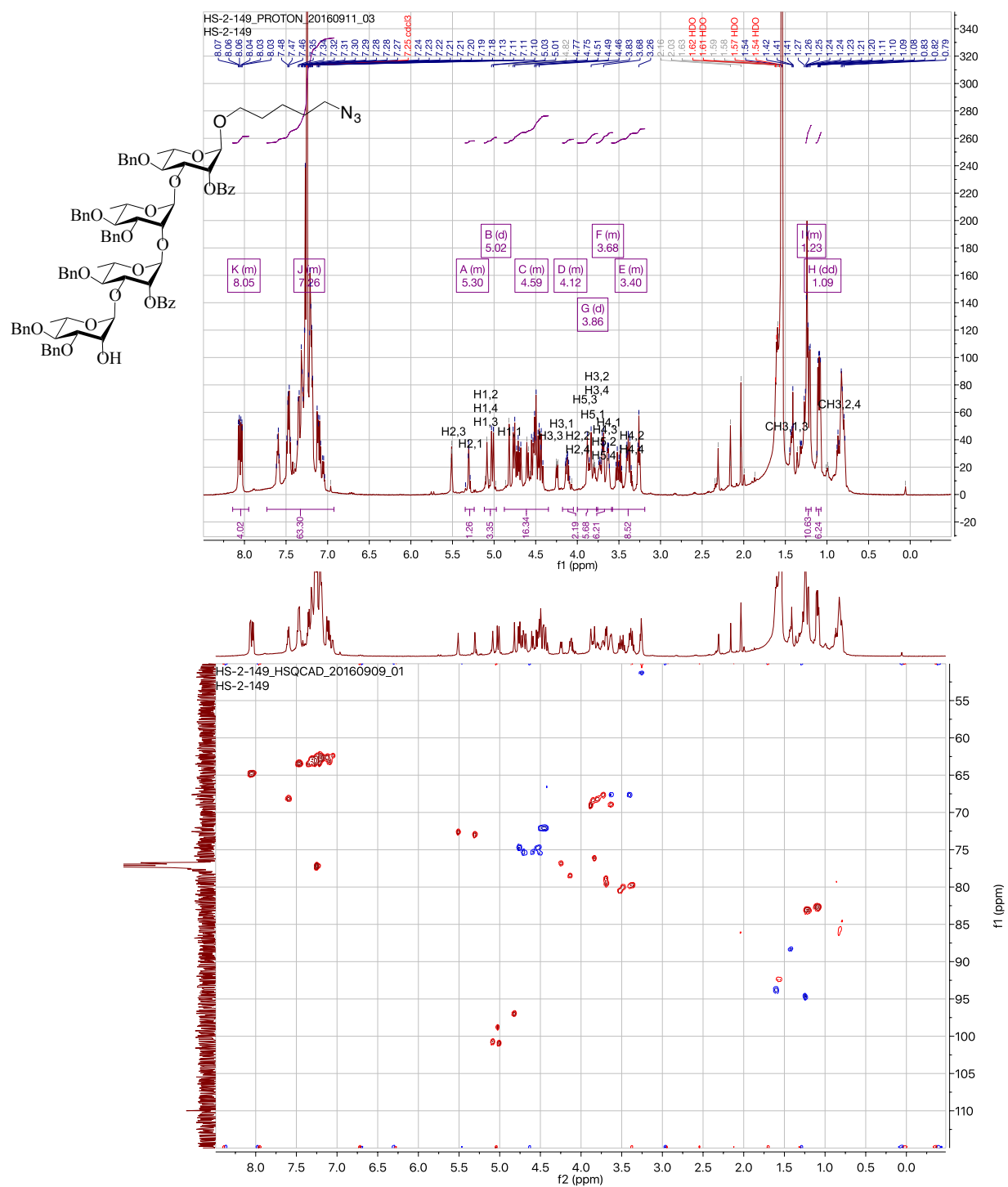
HS-2-69PHeX_PRESAT_20160420_01
HS-2-69PHeX

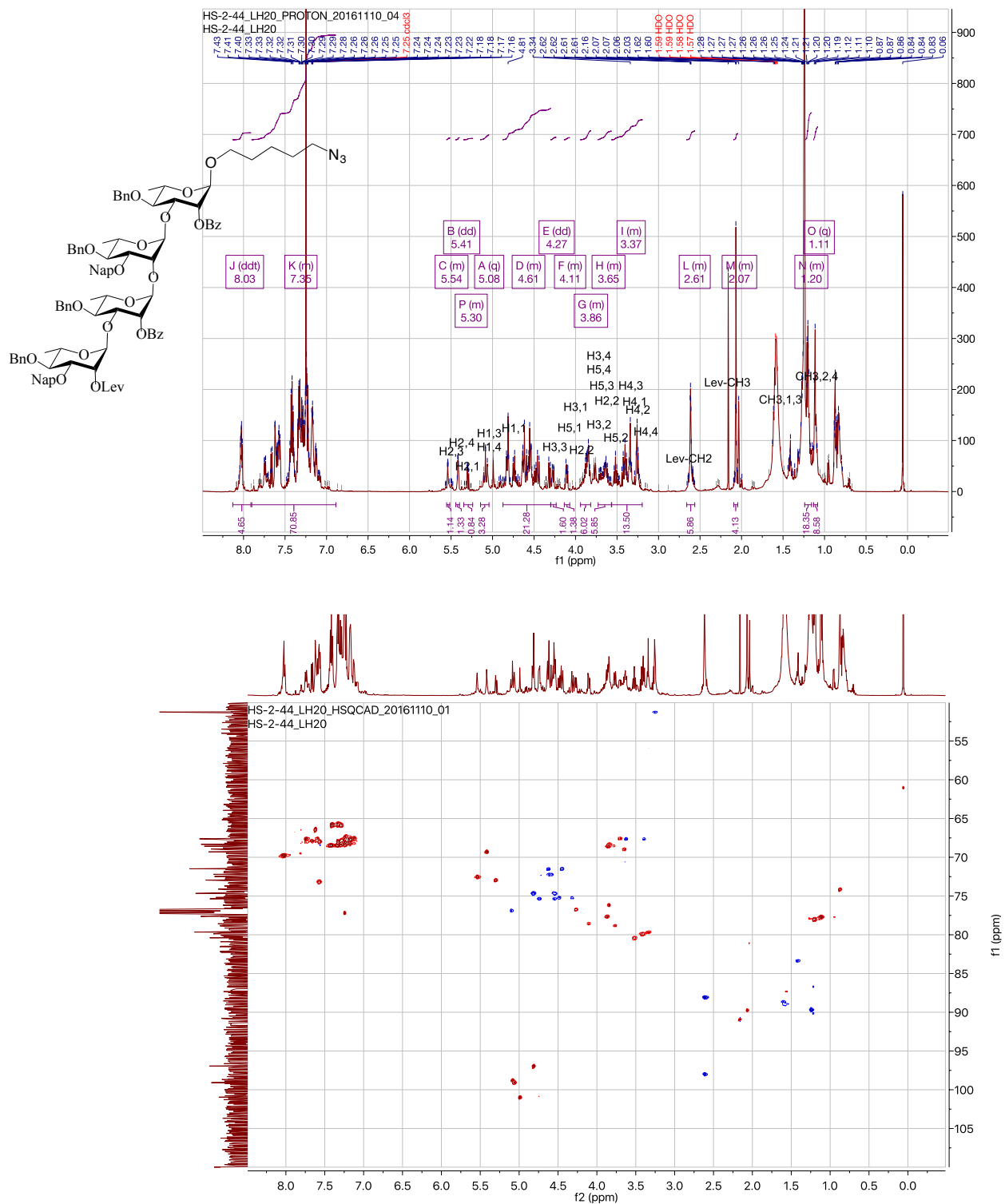


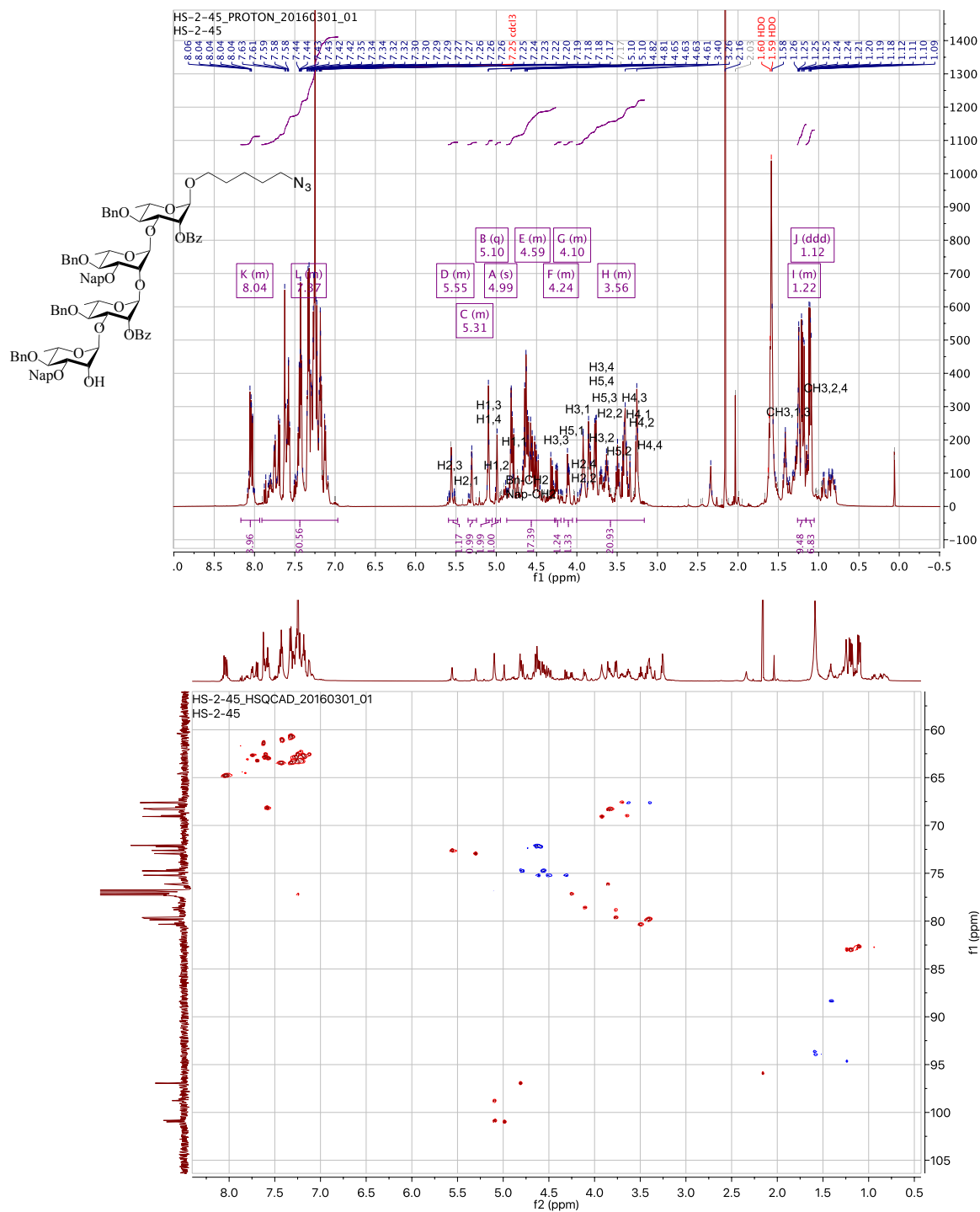


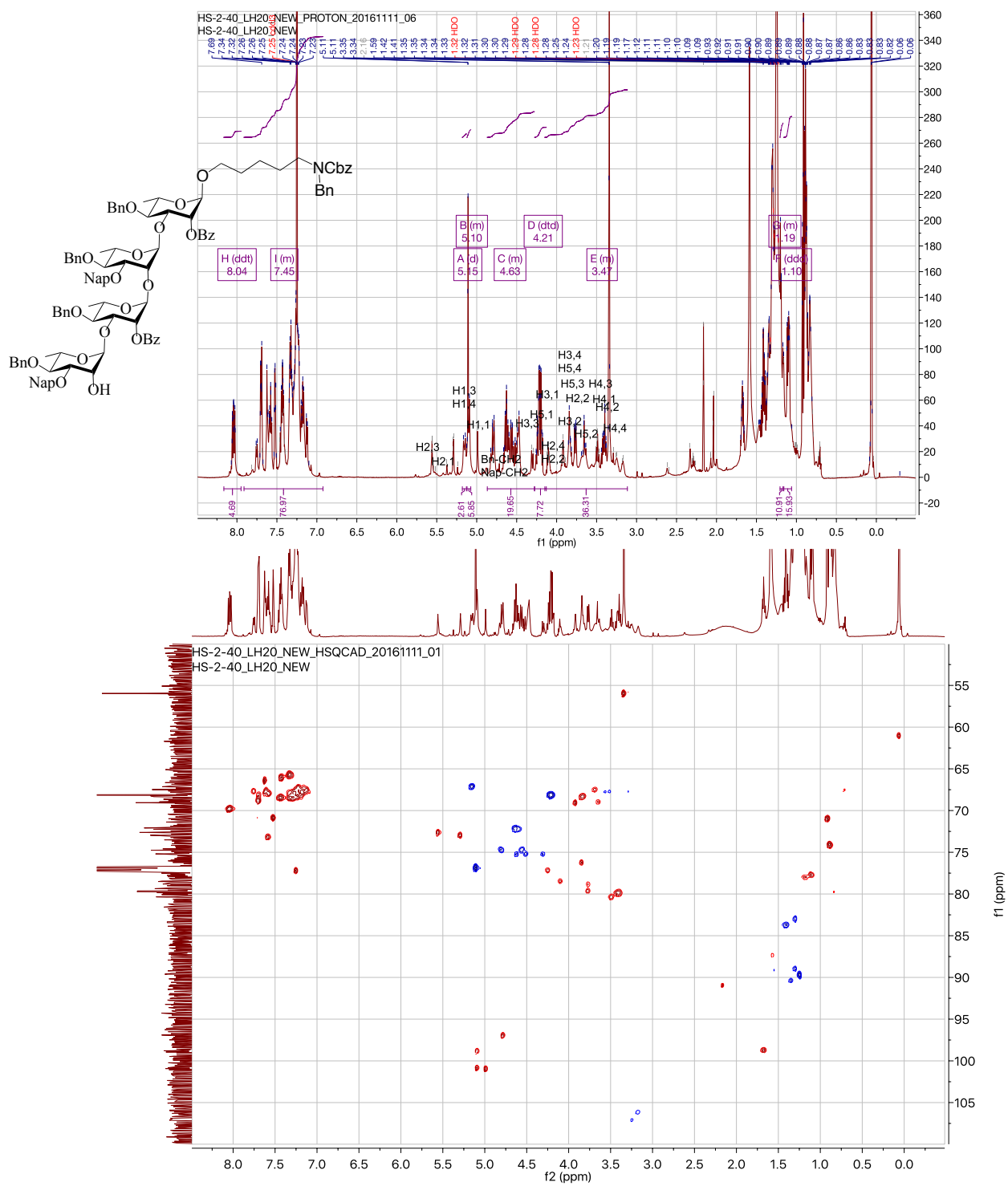






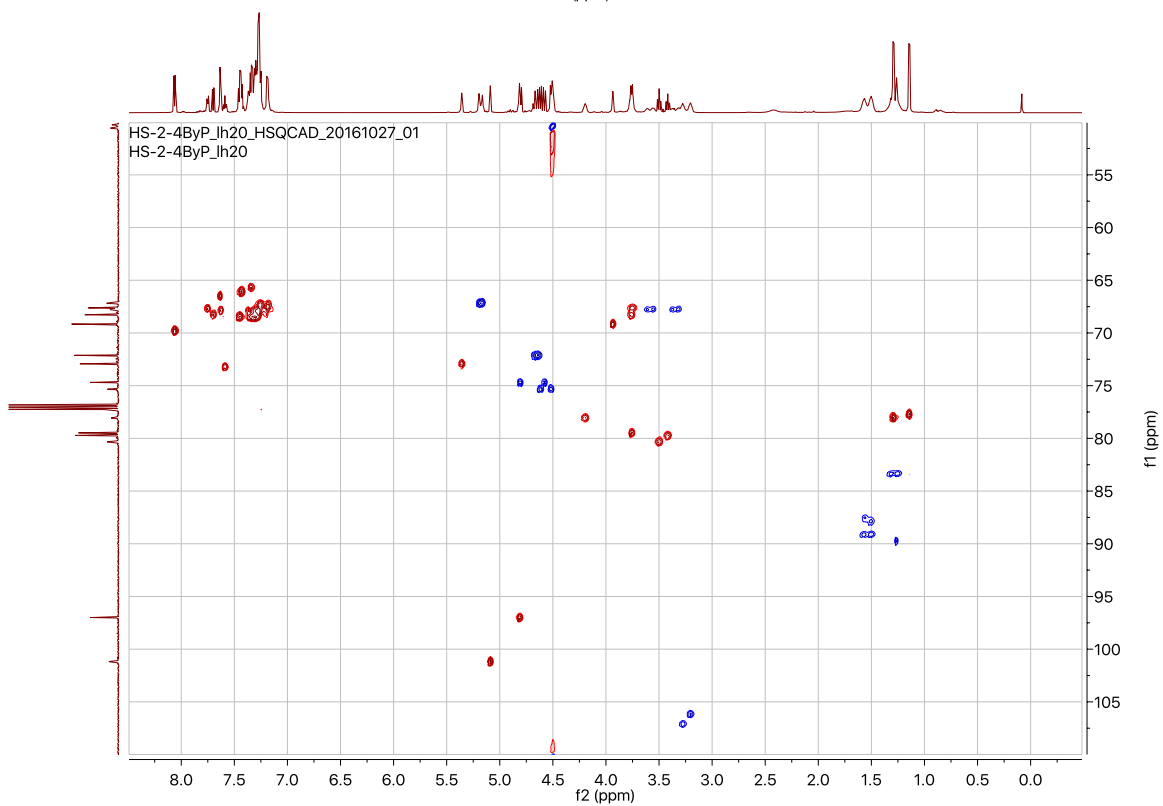
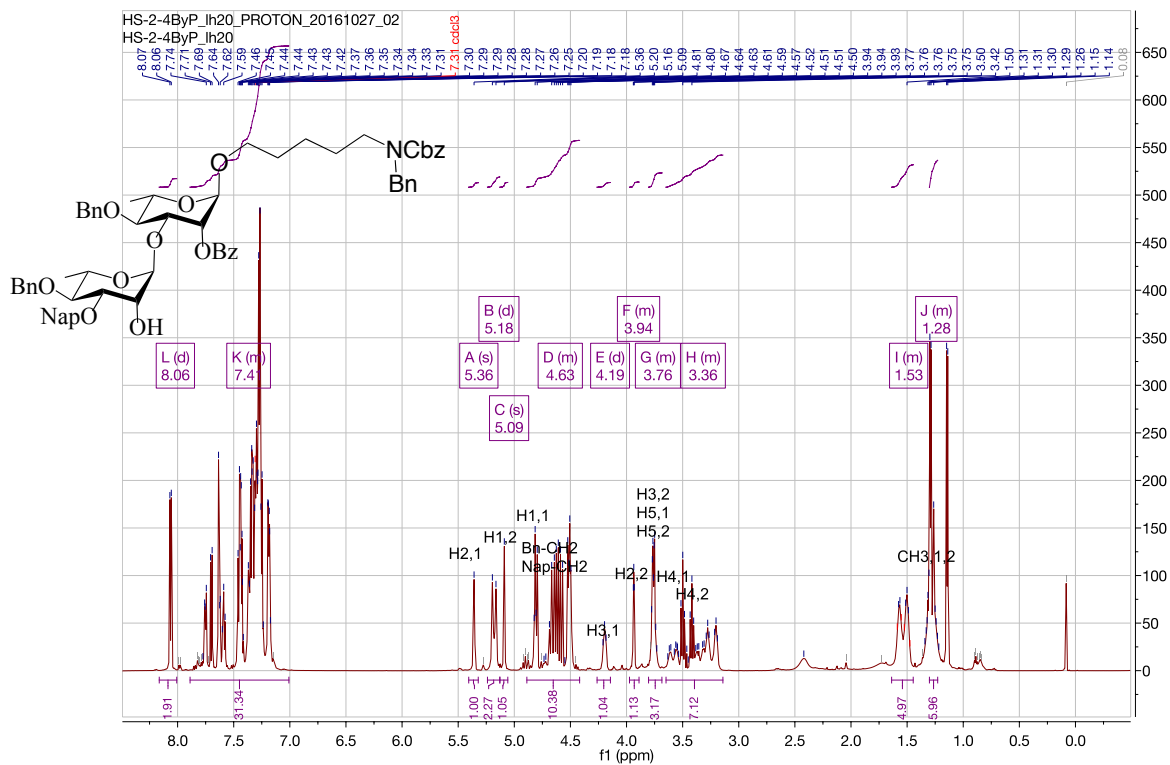




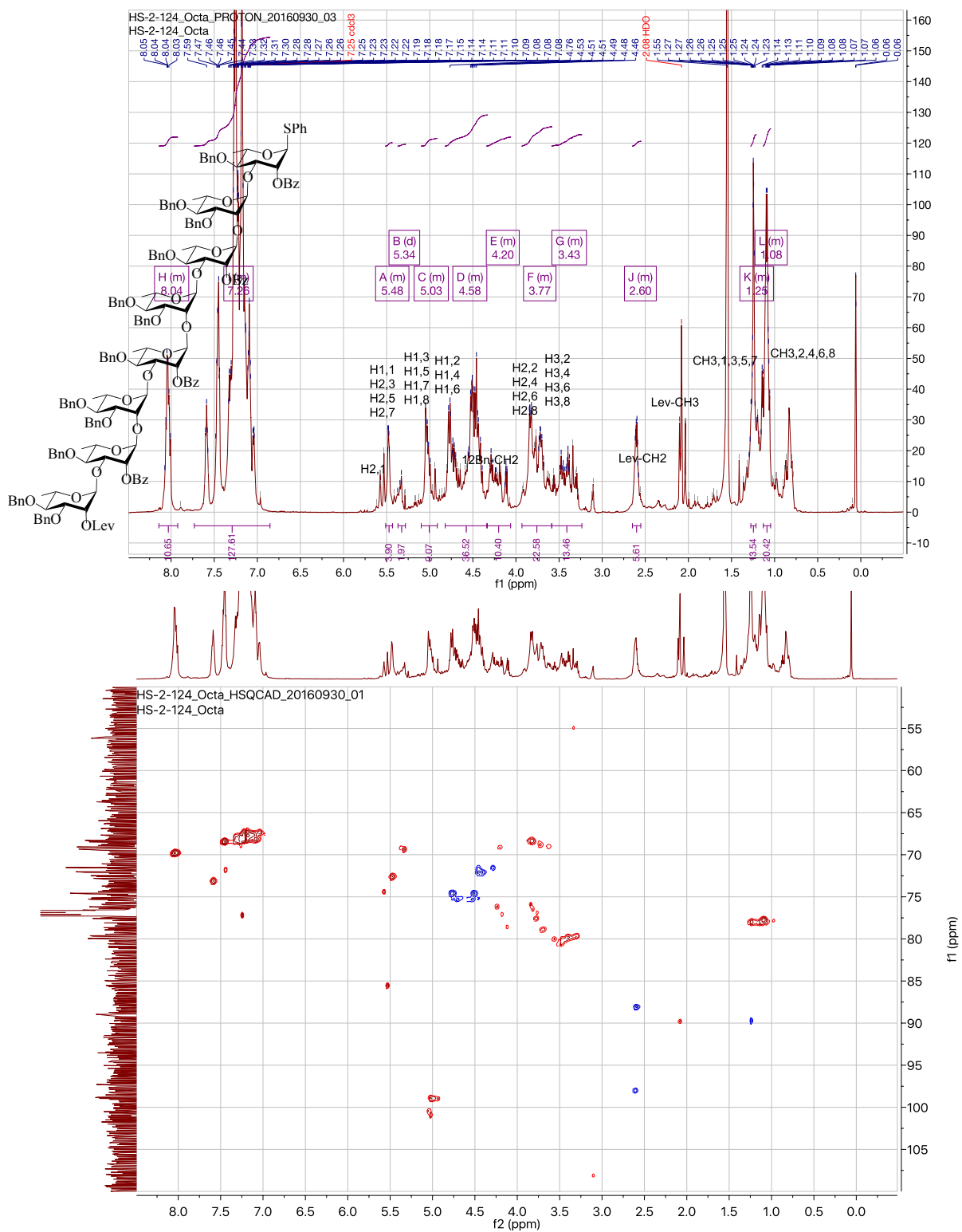




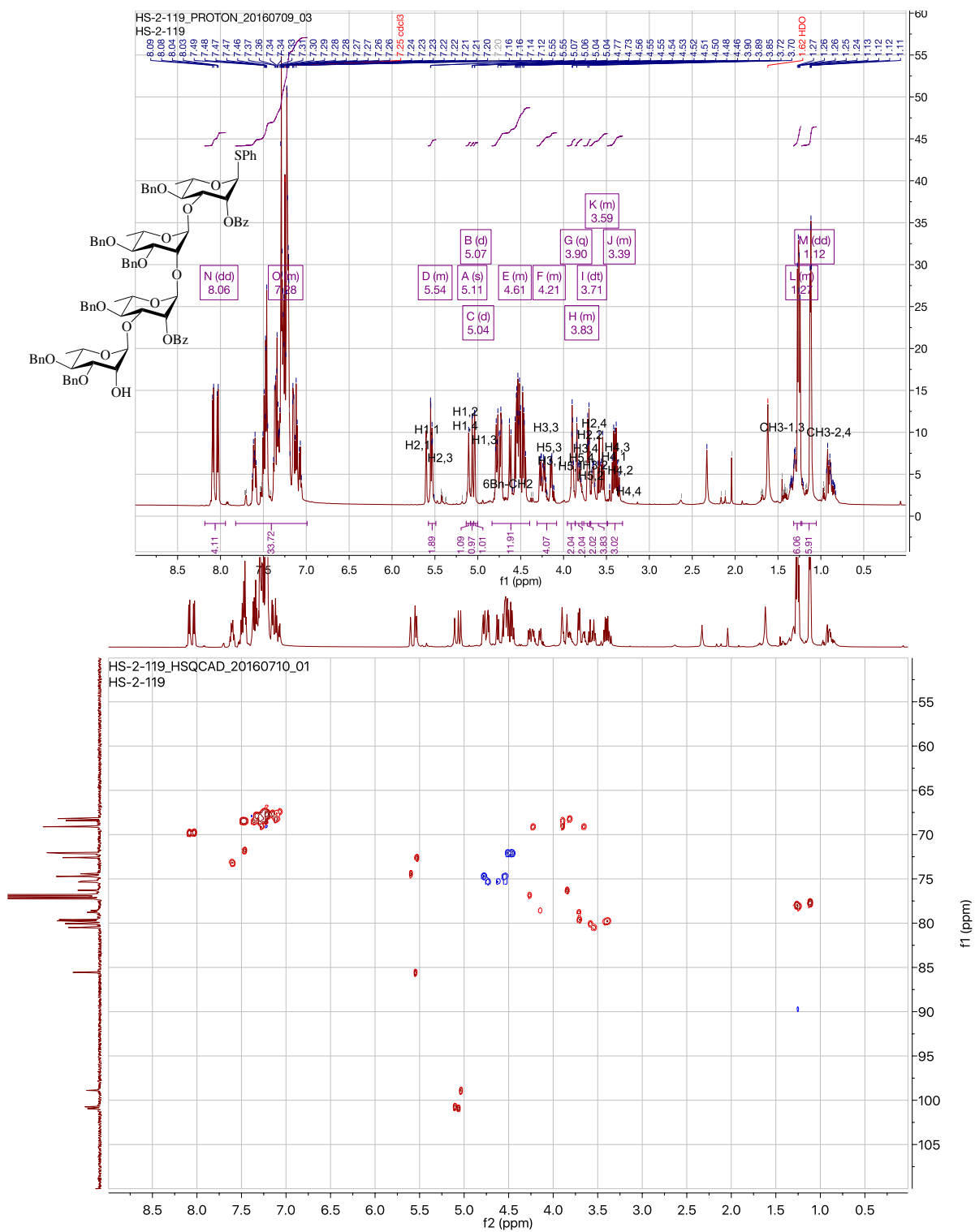
15



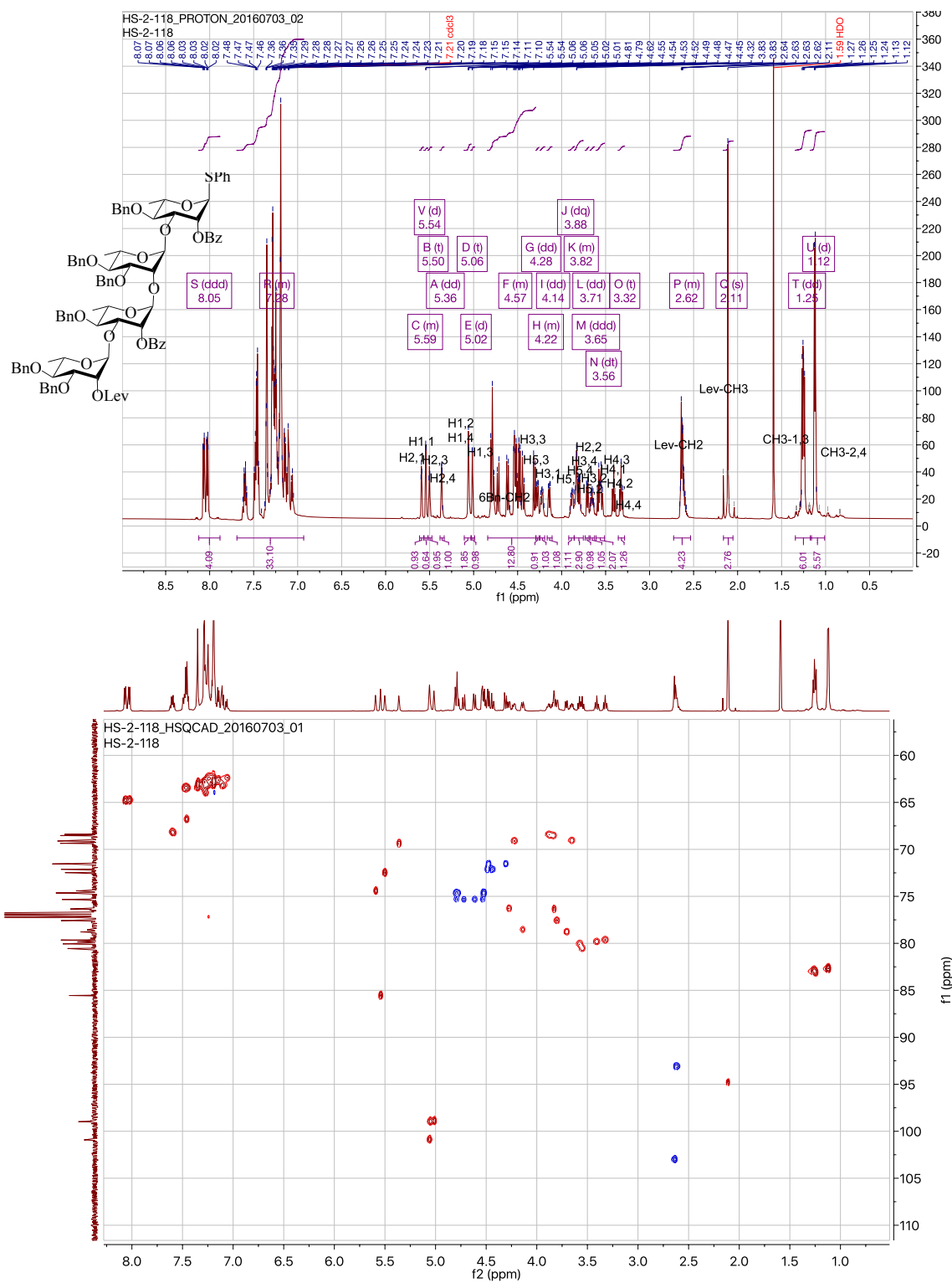
Octa-Bn



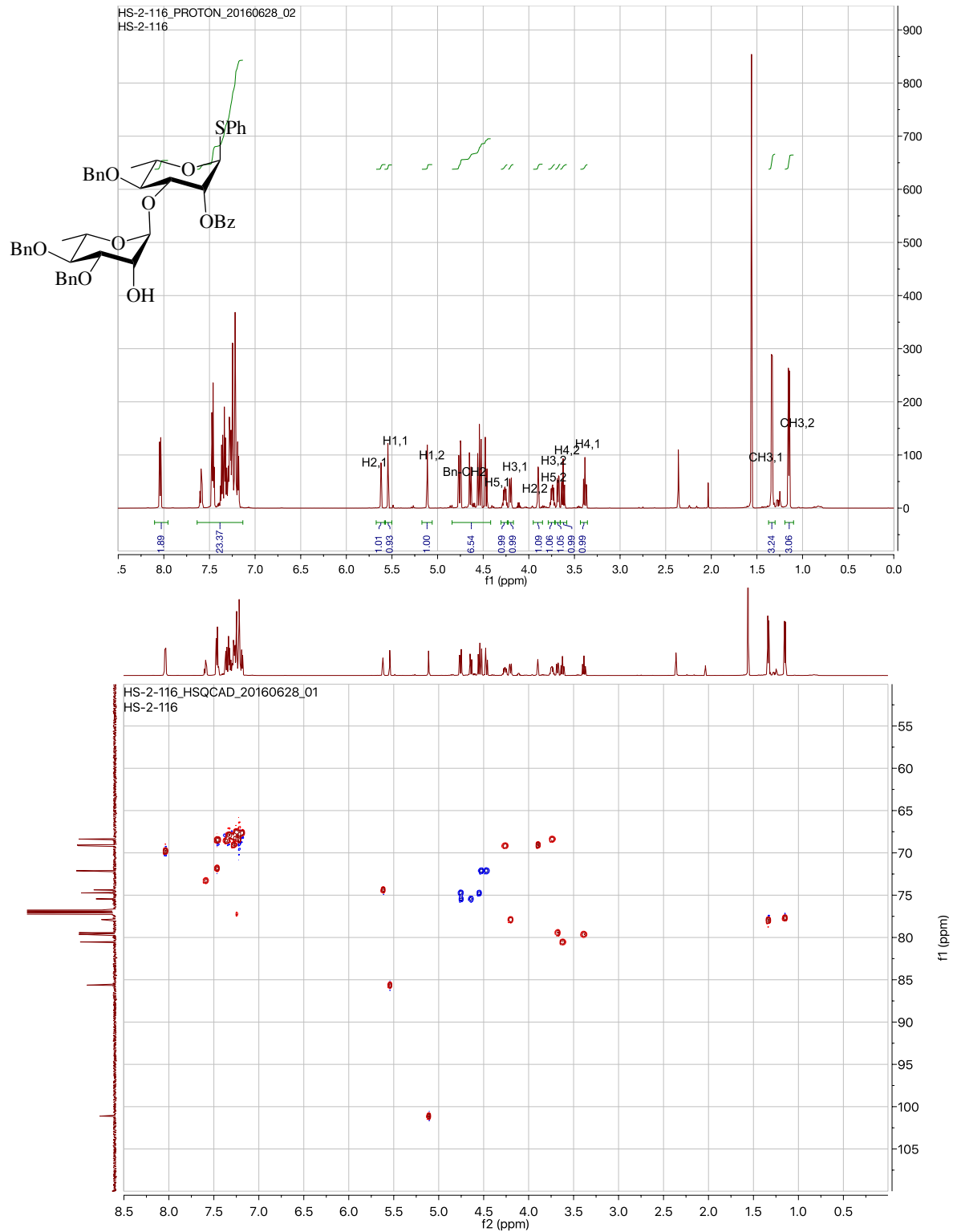
27 α



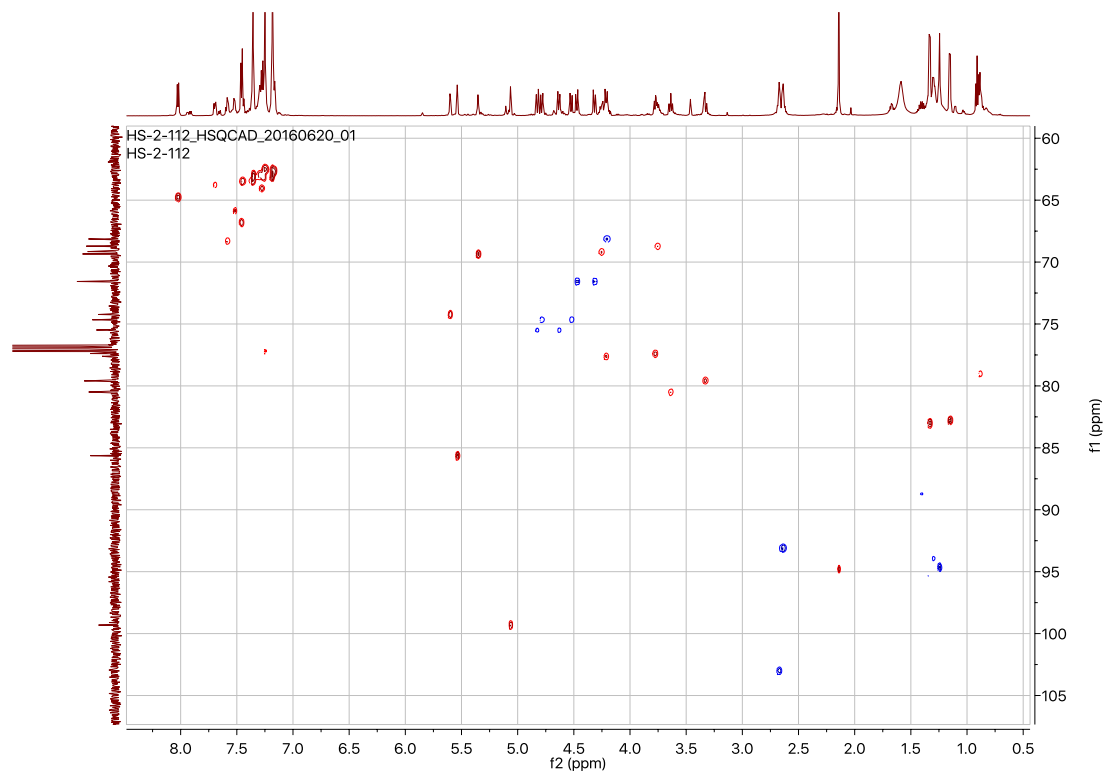
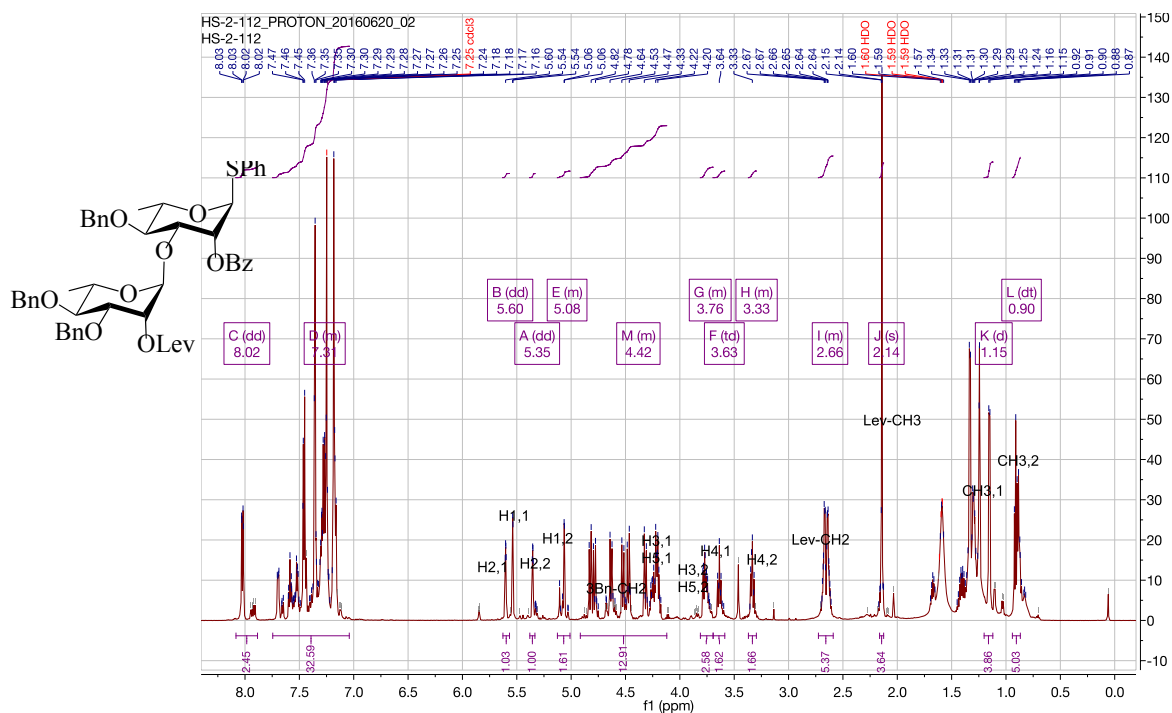
24 α

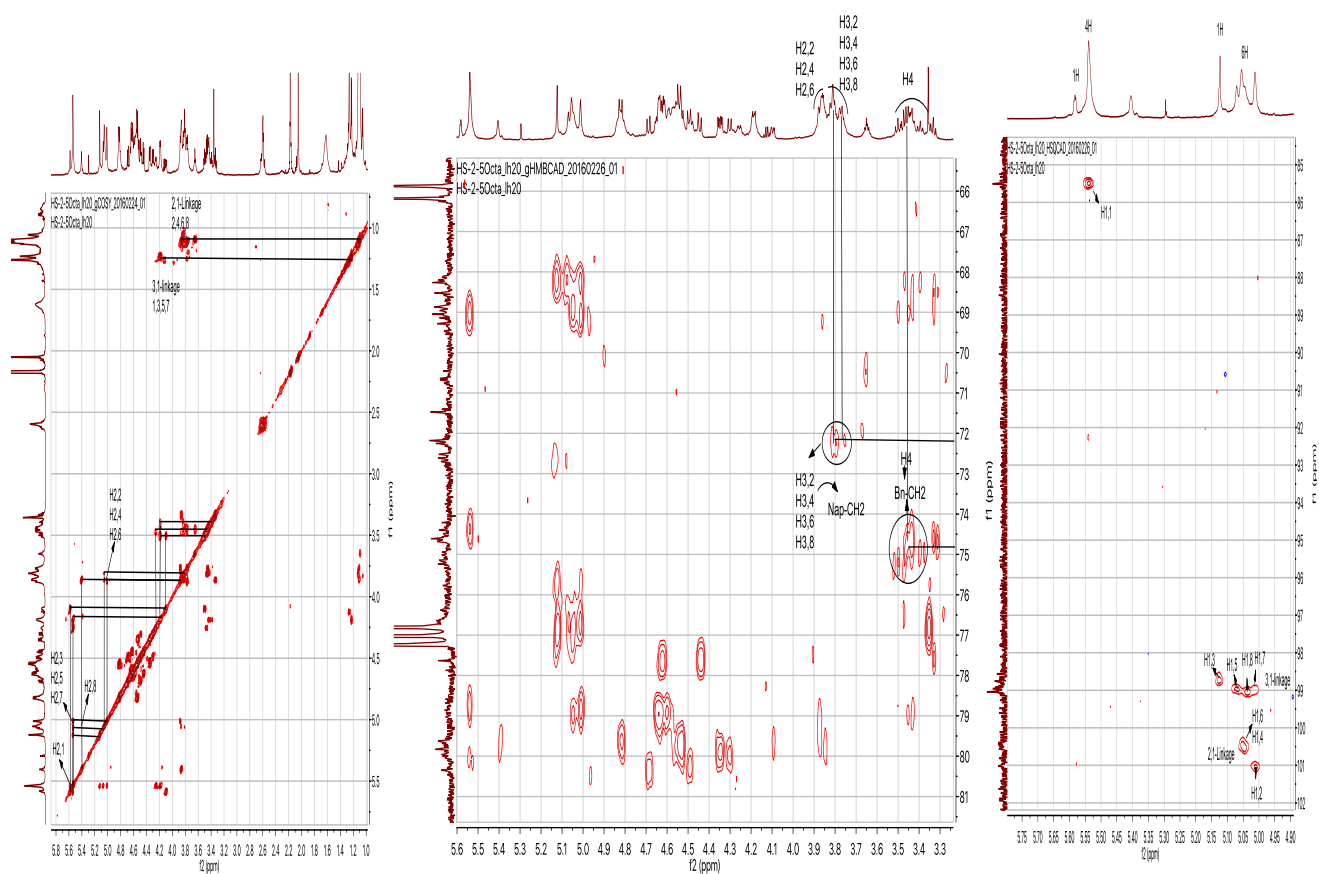


23 α

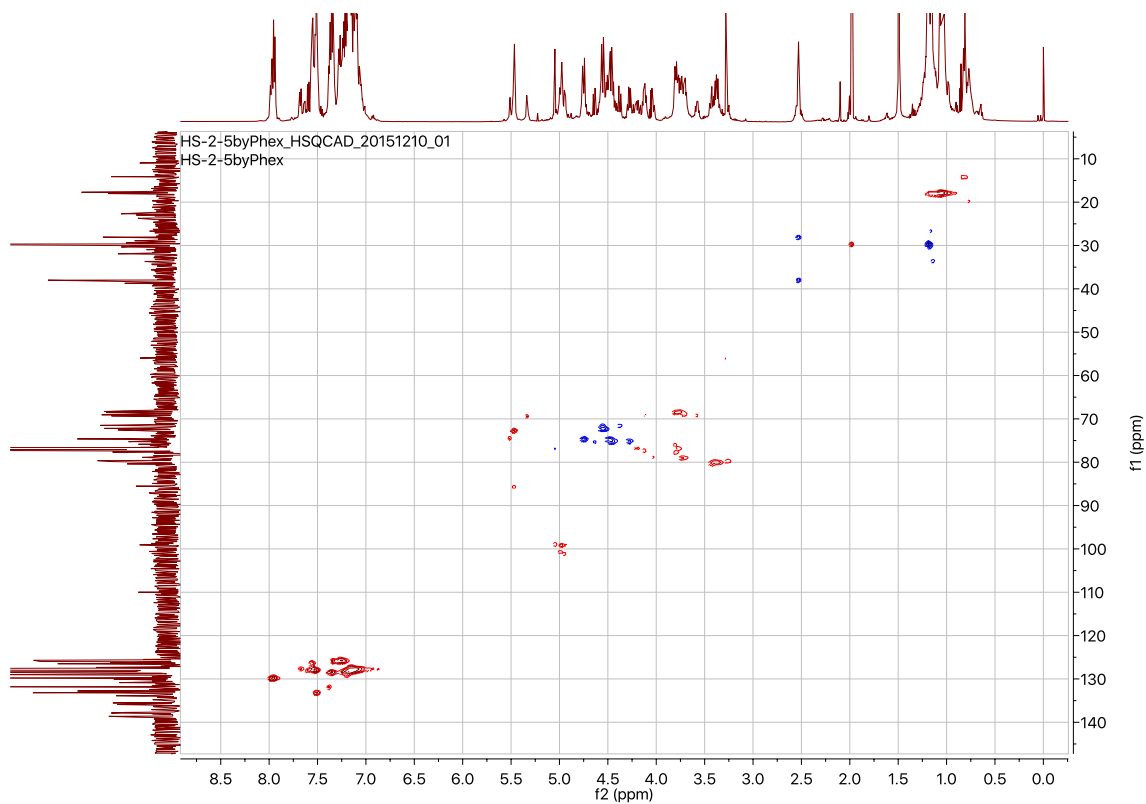
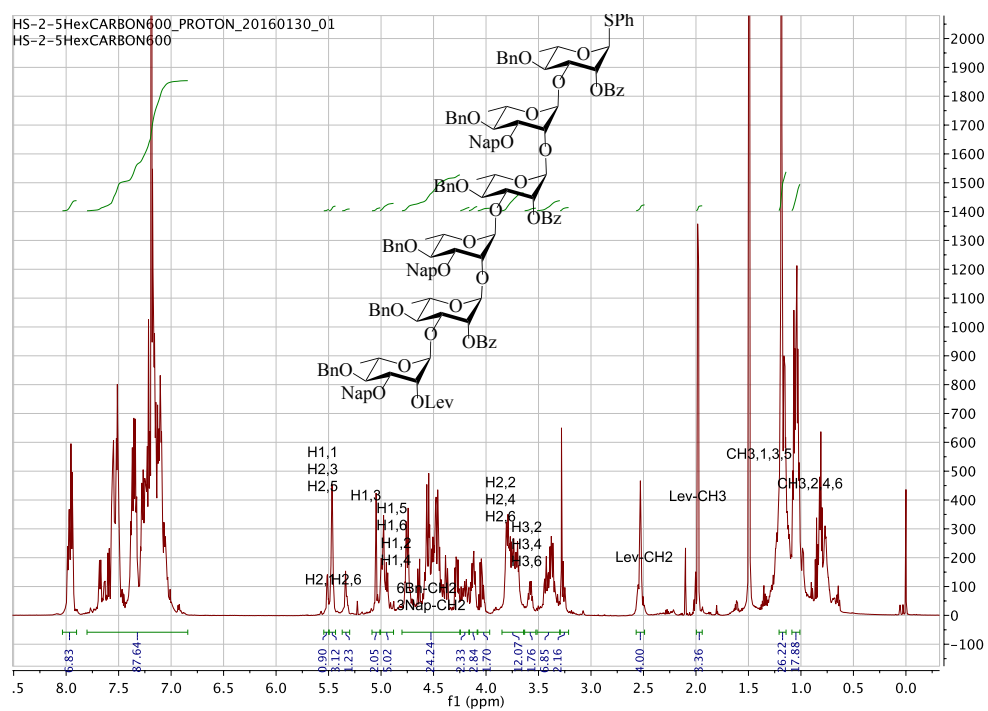


20 α

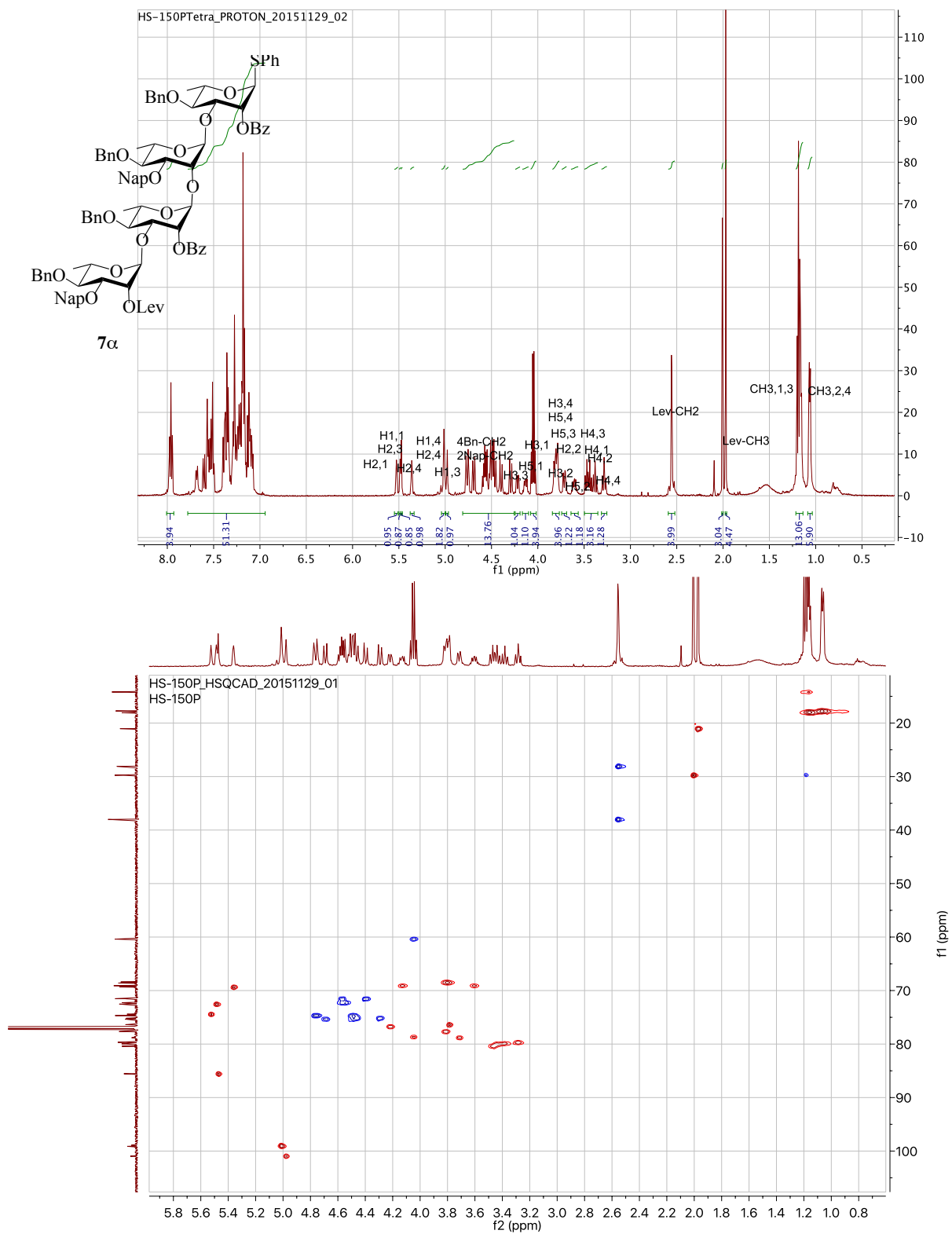


[illegible]

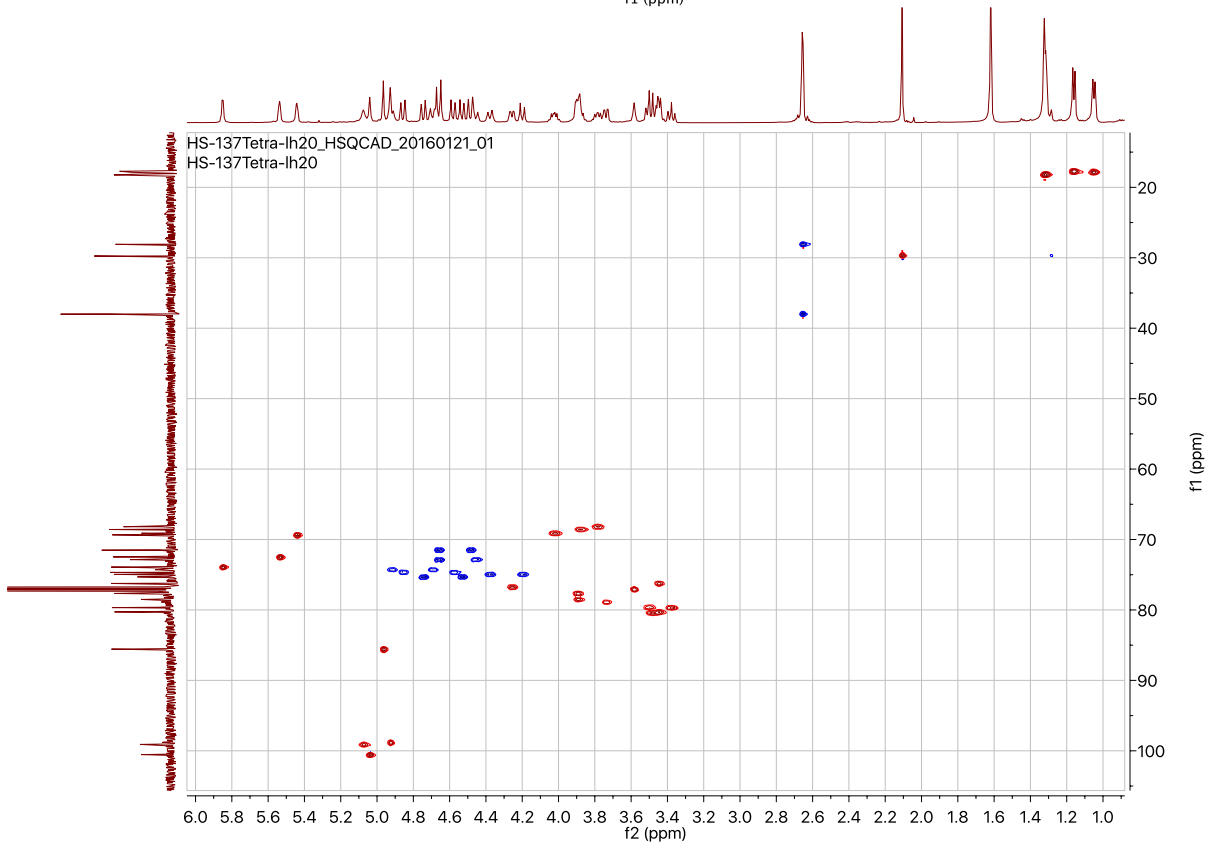
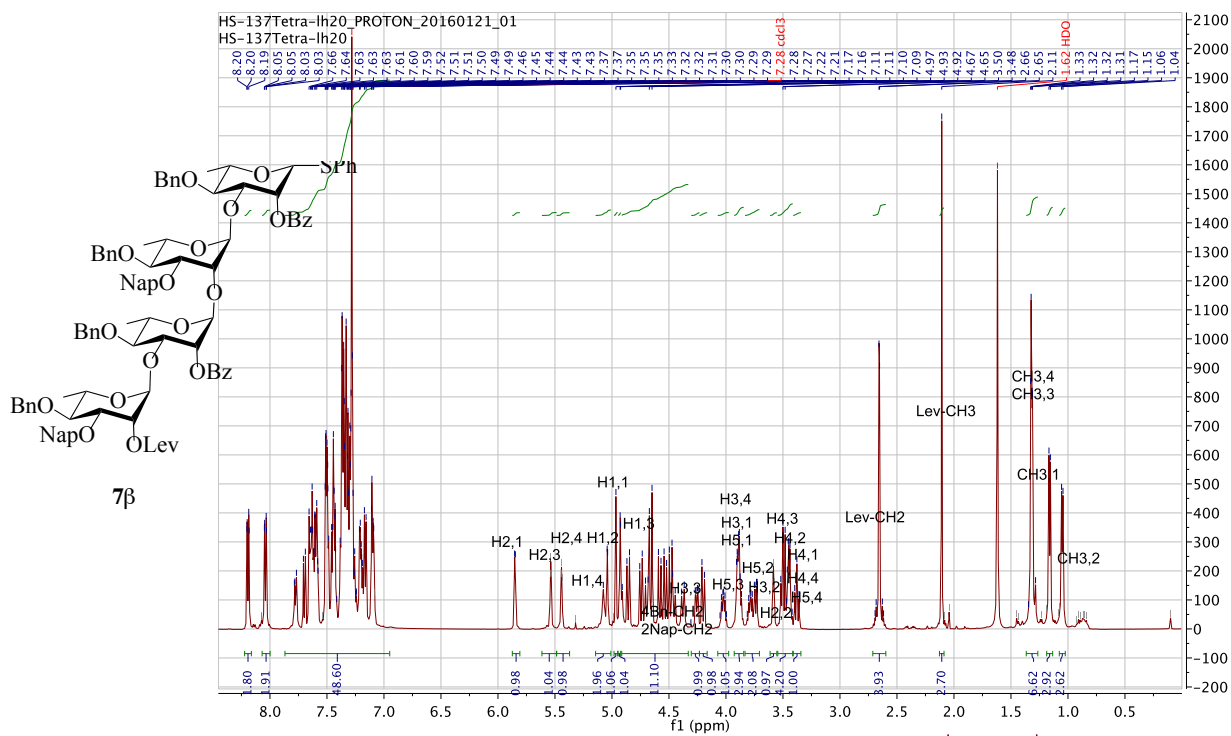
Hex



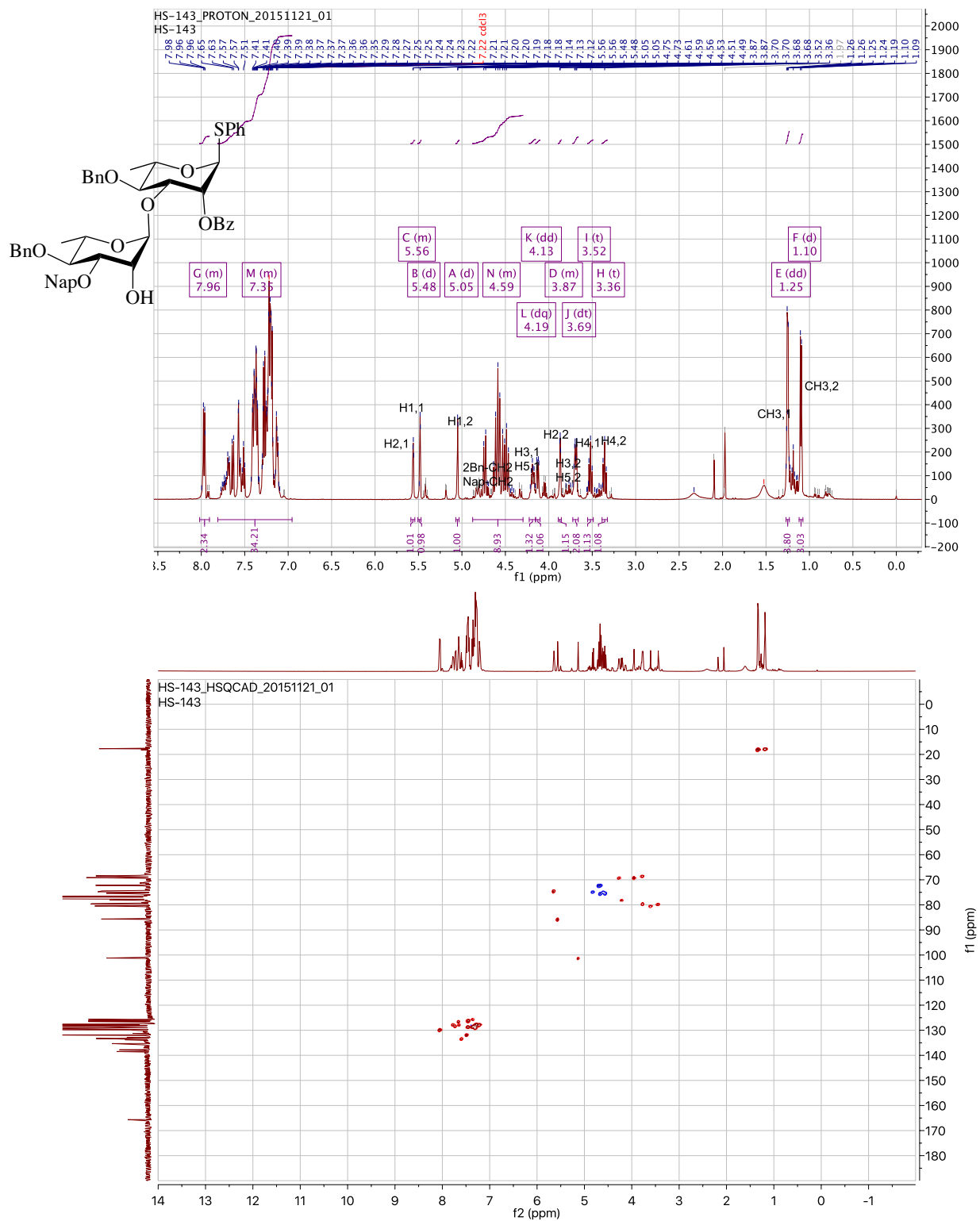
7 α



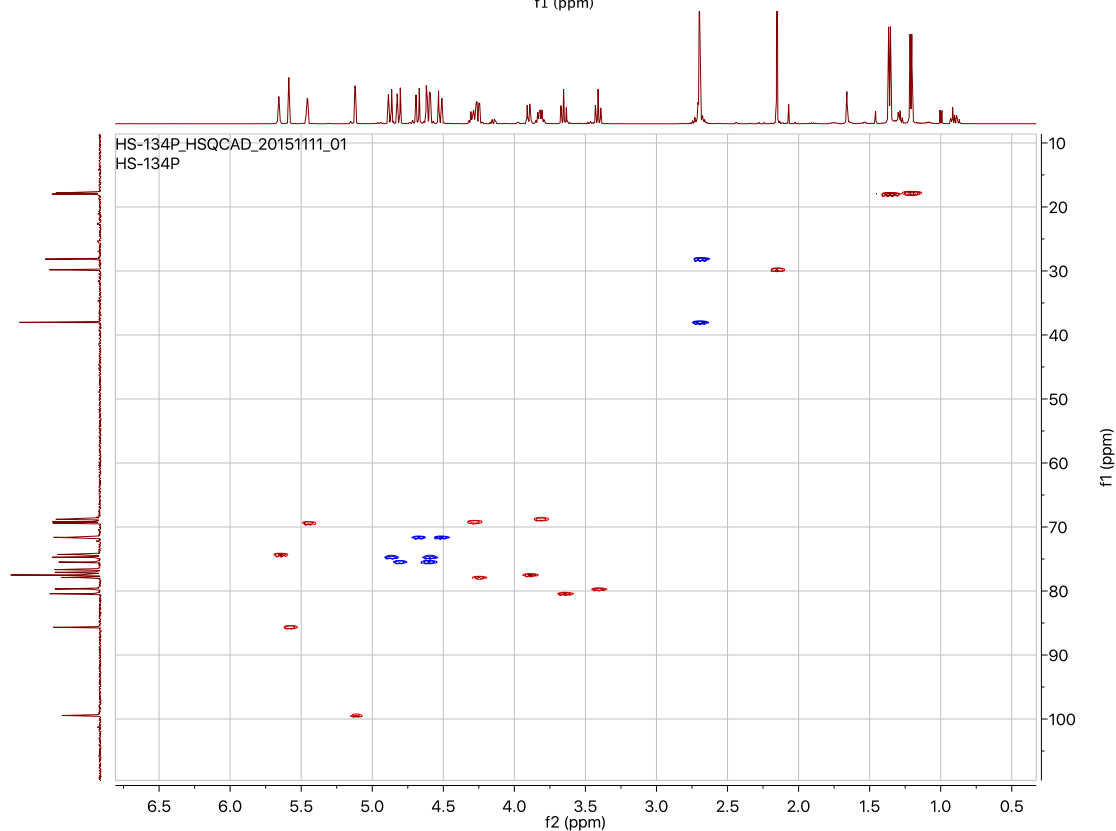
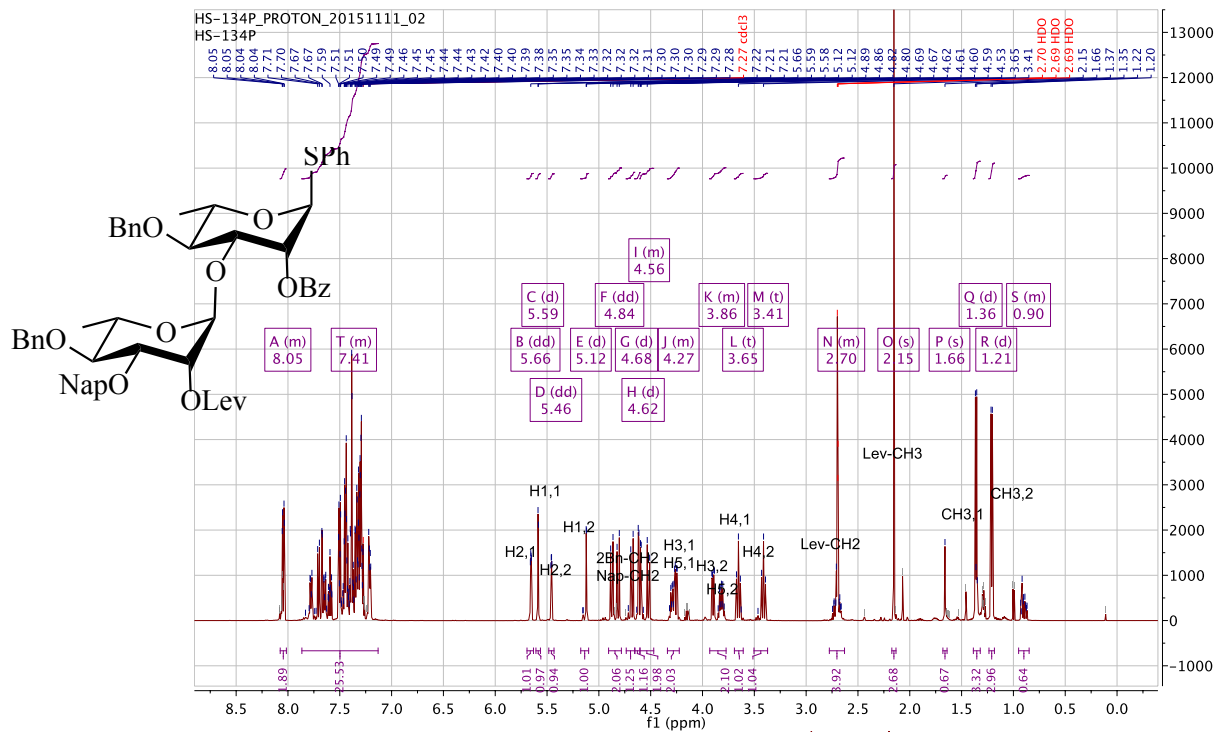
7β



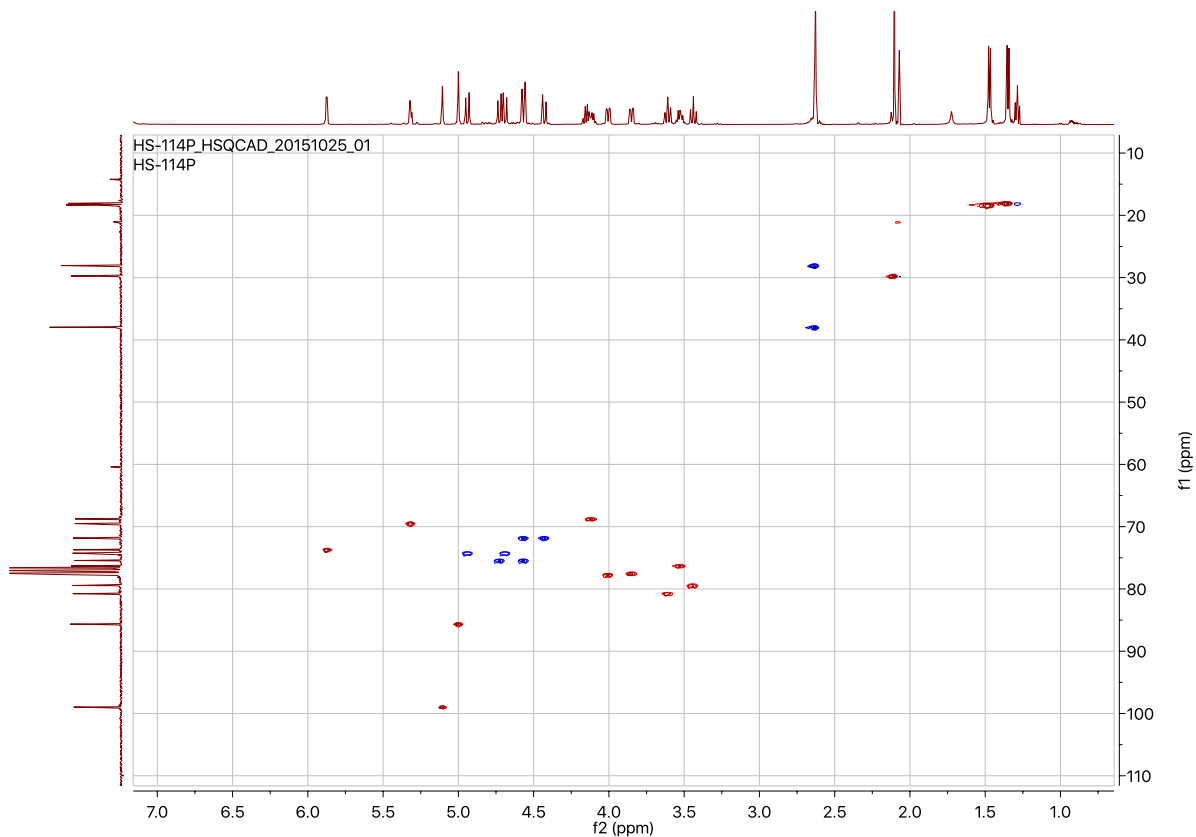
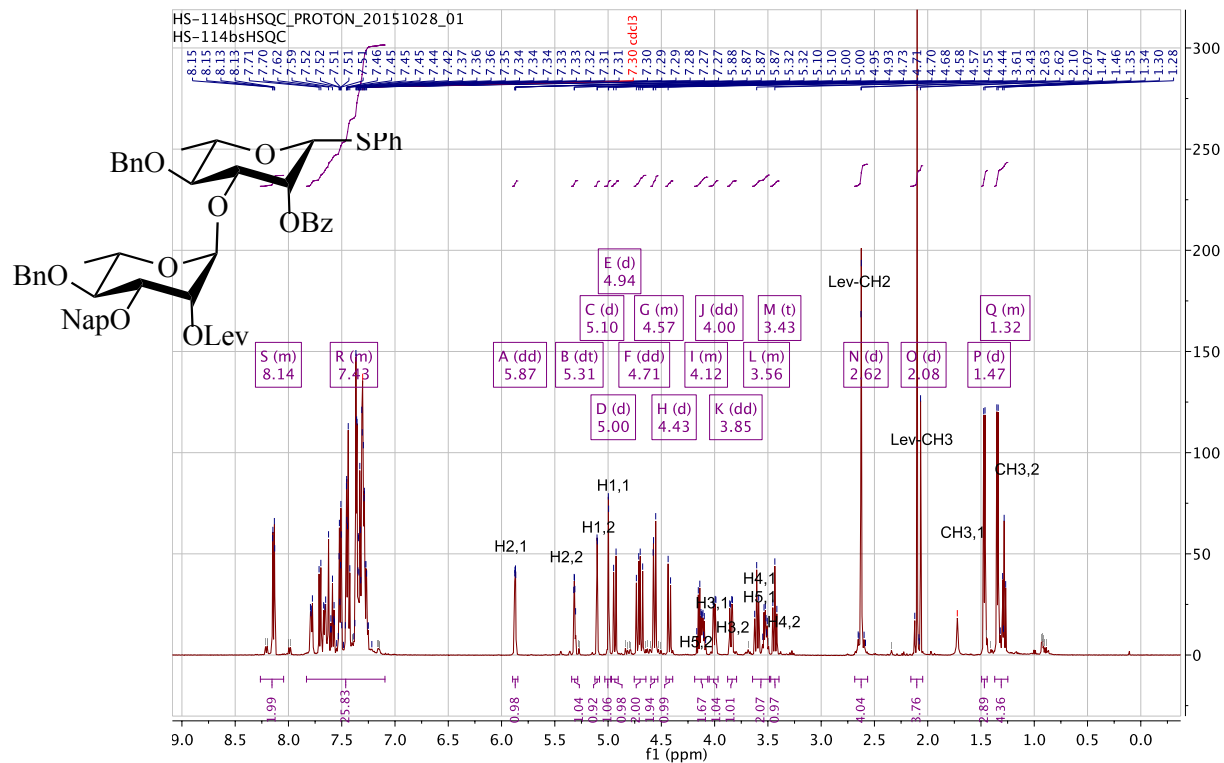
6α



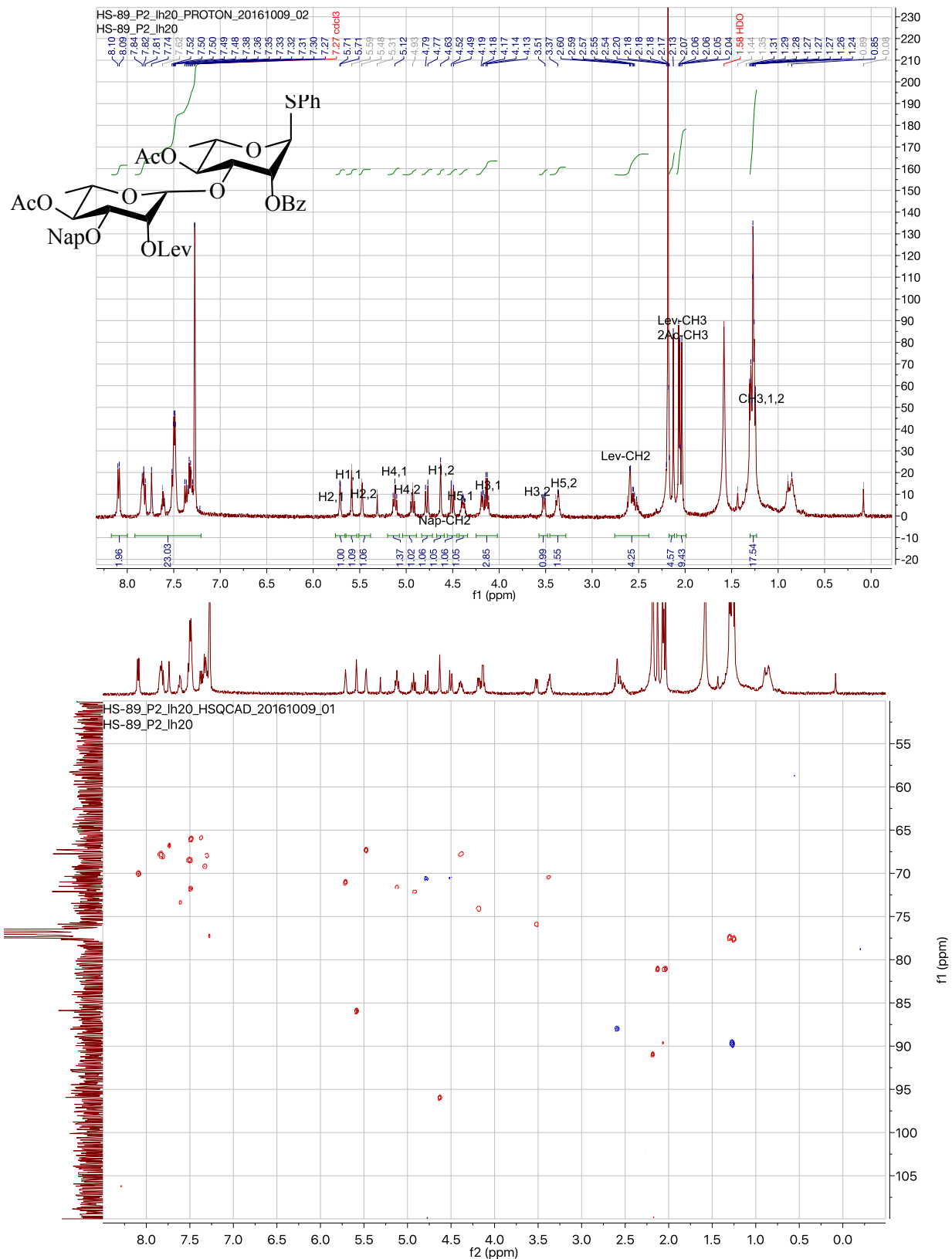
3 α



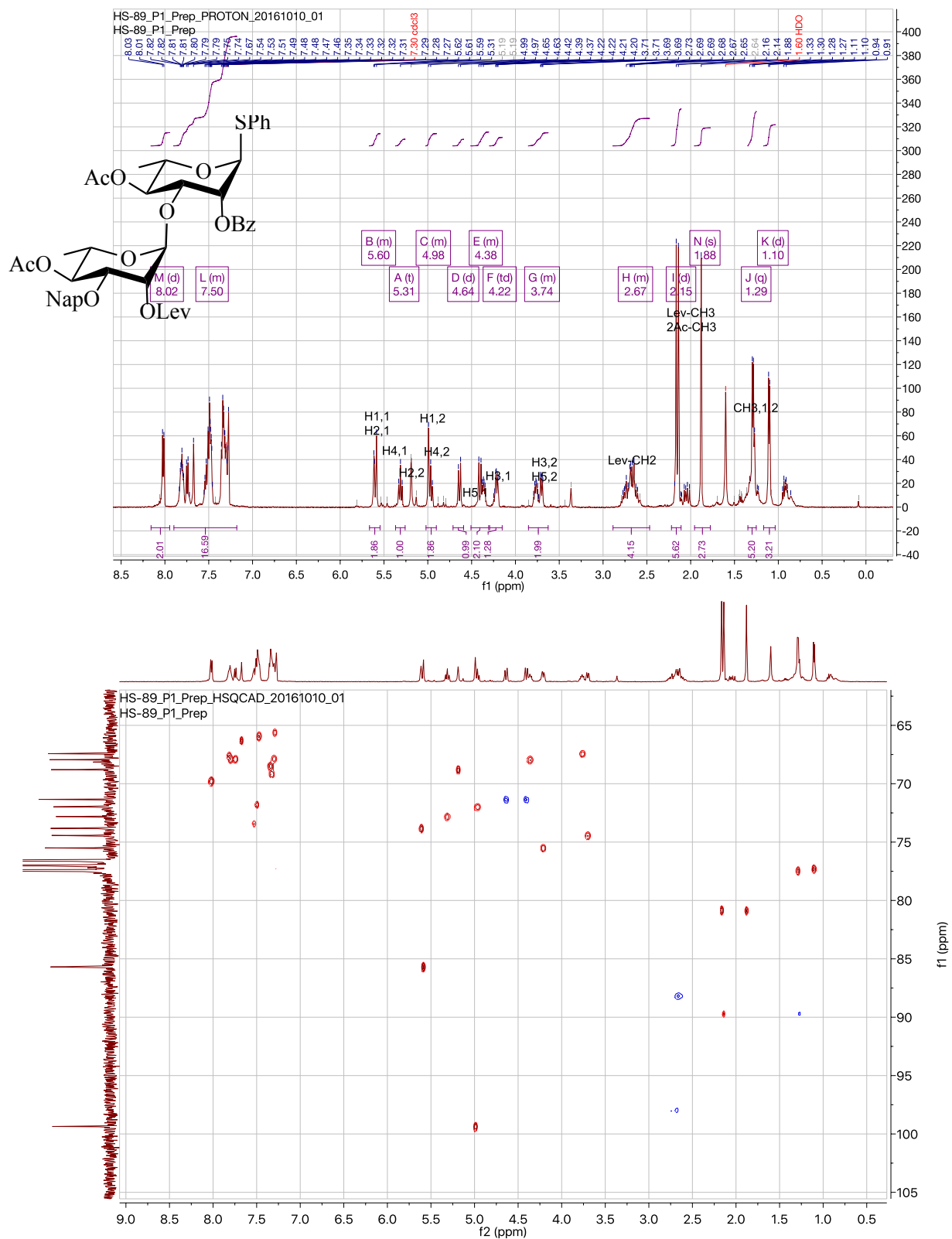
3β



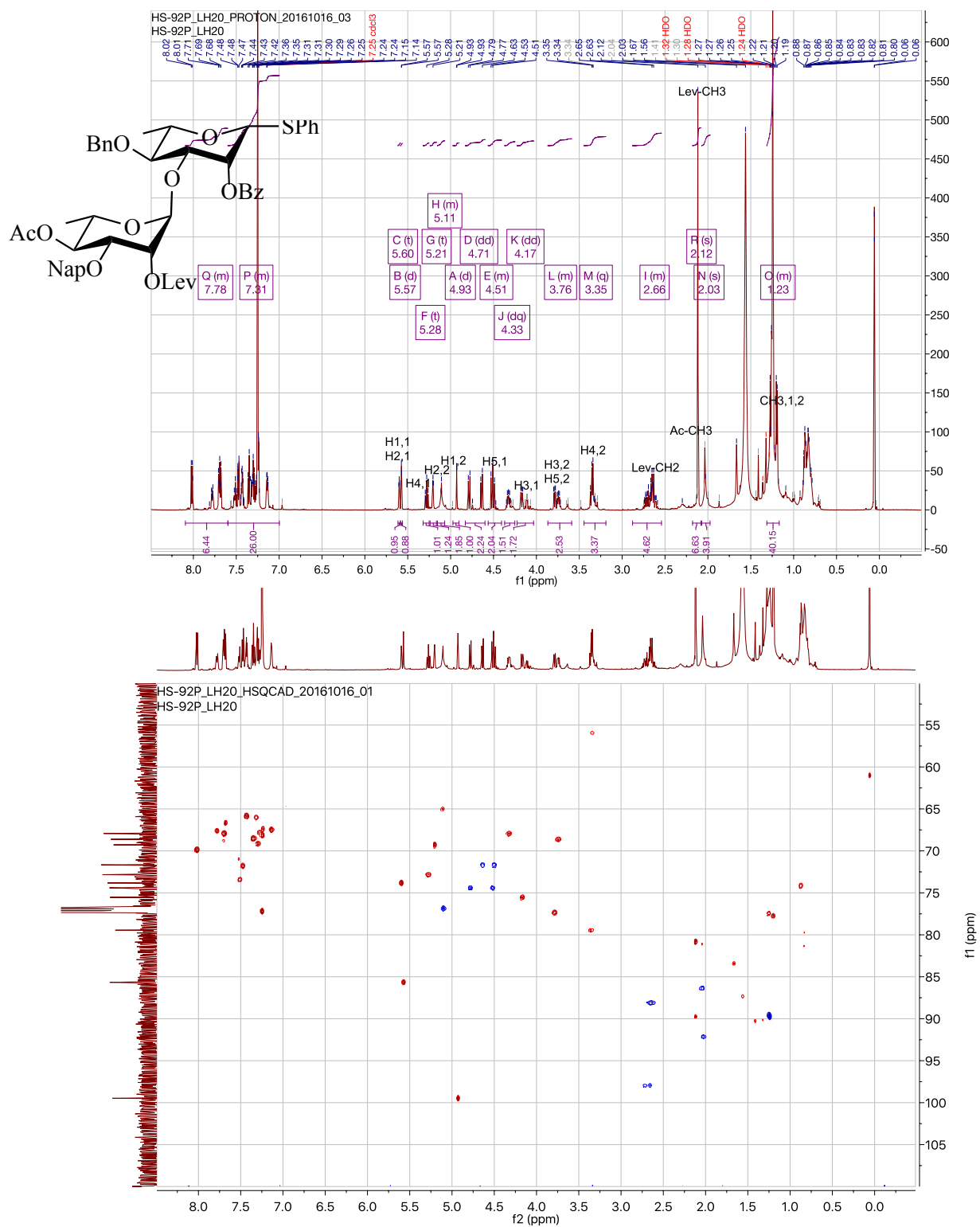
4 β - α



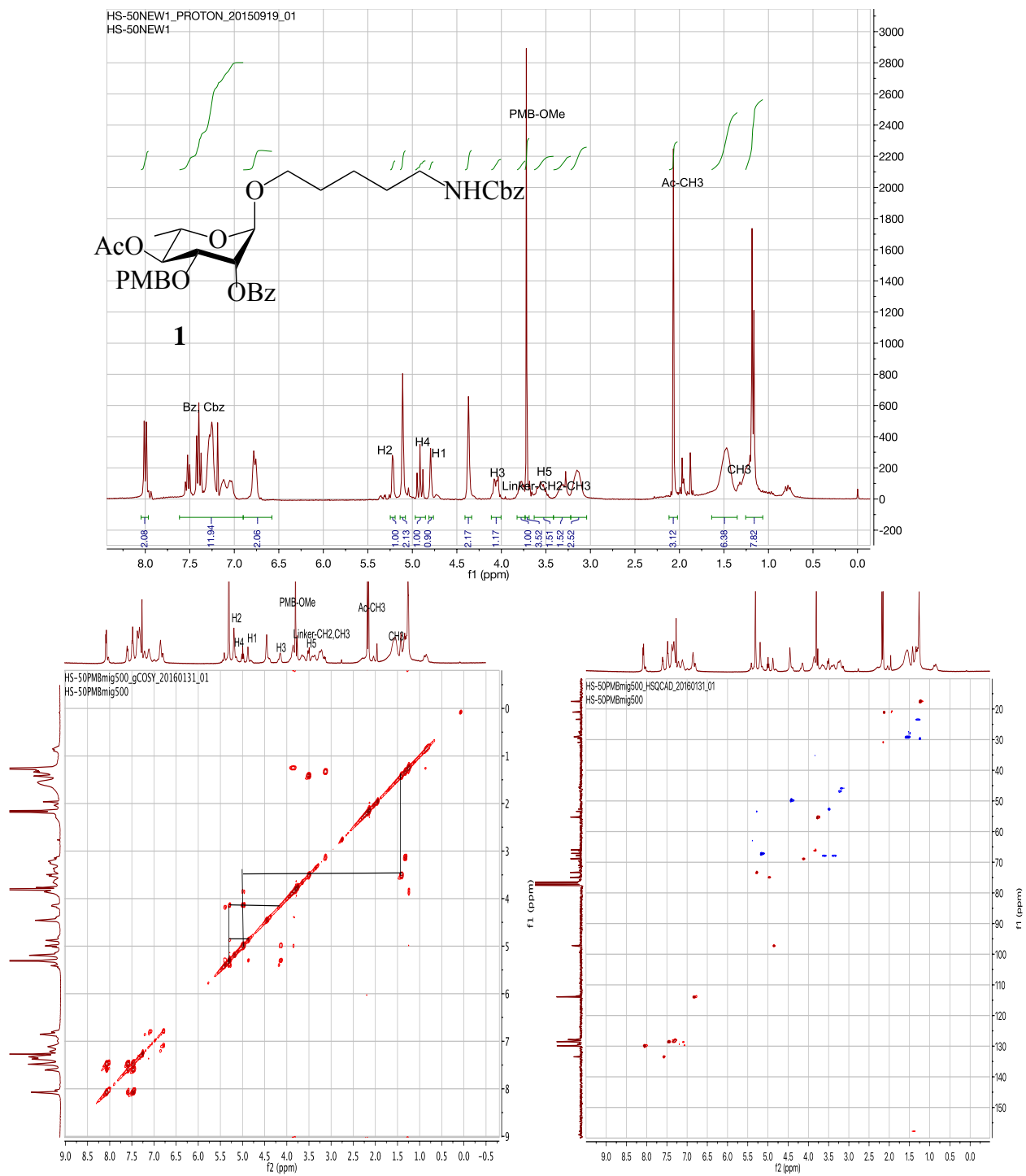
$4\alpha\text{-}\alpha$



5β







APPENDIX B

SUPPLEMENTARY DATA OF KEY PRODUCTS IN CHAPTER 3

