LEVERAGING BIOORTHOGONAL TAGGING AND C-GLYCOSIDE CHEMISTRY FOR THE STUDY OF O-MANNOSYLATION

by

ASHLEY CARTER

(Under the Direction of Geert-Jan Boons and Lance Wells)

ABSTRACT

The protein α-dystroglycan (α-DG) is highly glycosylated with *O*-mannose and *O*-GalNAc glycans through serine or threonine within its mucin-like domain. Defects in *O*-mannosylation of α-DG lead to dystroglycanopathies, which are congenital muscular dystrophies involving neurological defects. Furthermore, hypoglycosylation of α-DG leads to defective axonal guidance and neuronal migration as well as mosaic spacing of neurons during retinal development. Interactions between α-DG and its extracellular matrix (ECM) ligands require α-DG to be extended by *O*-mannose structures based on the M3 core structure by the glycosyltransferase POMGNT2. The core M3 can be extended with matriglycan, which is a repeating disaccharide that binds LG-domain containing proteins in the ECM. There is currently no method to identify the M3 core without the presence of matriglycan. Therefore, it remains a possibility that unextended M3 structures exist, but there is currently no tool to identify core M3 glycans without the presence of the repeating disaccharide. Development of antibodies to the core *O*-mannose glycans would help to identify these core glycans and ultimately better understand their function. Generating antibodies against this epitope presents a challenge since the oxygen in the glycosidic linkage is vulnerable to enzymatic degradation. However, carbon-

linked (*C*-linked) glycosides, in which carbon replaces the normal oxygen in the glycosidic linkage, are resistant to enzymatic degradation and provide robust immunogens. In efforts to overcome stability issues of *O*-linked glycoside immunogens, a novel *C*-linked glycoside mimic (*C*-Man-Thr) of *O*-mannose-Threonine was synthesized. This biomimetic has the potential to be used as an antigen to generate antibodies to identify these core *O*-mannose glycans. Extension of the *C*-linked glycoside into *C*-linked glycopeptides to generate antibodies has the potential to provide improved reagents to generate a more comprehensive profile of sites that are modified with *O*-mannose glycans. Another route of enhanced detection includes bioorthogonal tagging with an azido-modified form of UDP-GlcNAc and subsequentially clicking with a bioorthoganol biotin tag to enrich for modified protein sites. Potentially novel *O*-mannosylated proteins were enriched using this methodology.

INDEX WORDS: *O*-mannosylation, bioorthogonal tagging, click chemistry, *C*-glycoside, *C*-glycopeptide, alpha-dystroglycan, *O*-glycopeptide, SEEL

LEVERAGING BIOORTHOGONAL TAGGING AND C-GLYCOSIDE CHEMISTRY FOR THE STUDY OF O-MANNOSYLATION

by

ASHLEY CARTER

BS, Berry College, 2015

MS, Kennesaw State University, 2017

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

2023

LEVERAGING BIOORTHOGONAL TAGGING AND C-GLYCOSIDE CHEMISTRY FOR THE STUDY OF O-MANNOSYLATION

by

ASHLEY CARTER

Major Professors:	Geert-Jan Boons
	Lance Wells
Committee:	David Live

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
December 2023

DEDICATION

I dedicate this dissertation to my grandfather, John Edmond Carter, who was the epitome of perseverance, optimism, and unconditional love. His intentional way of living coupled with his ability to see positive opportunities in every circumstance has inspired me in all aspects of life, especially as a scientist. In that light, this is a dedication to always be an intentional scientist and to forever be willing to expand knowledge by remaining open to expansive views and perspectives.

ACKNOWLEDGEMENTS

I would like to thank my advisors, Dr. Geert-Jan Boons and Dr. Lance Wells, for their mentorship and valuable scientific insight throughout the years. Dr. Geert-Jan Boons' passion for science is felt deeply by all around him, and his enthusiasm and high energy has driven me to stay in a place of high motivation, especially during challenging times during the completion of my doctorate. My pattern of thinking and effervescence as a scientist will forever be shaped in a positive light because of you. I want to thank Dr. Lance Wells for his patience and willingness to train an organic chemist in his biochemistry lab. This has expanded my skillset and scientific experience in ways that will be highly valuable in my future as a scientist. I would like to thank Dr. David Live, for training me initially and for his valuable scientific insights over the years as a member of my committee, even after retirement.

I would like to thank Dr. John Glushka for his tremendous help and insights with NMR analysis, and Dr. Rob Woods for his help with molecular modeling in Chimera X of a key intermediate in my synthesis.

I am grateful to the Glycoscience Training Program (GTP) for funding from 2019-2021 and for providing me with the opportunity to rotate in labs to expand my skillset during my first semester.

I would like to thank all members, both past and present, of the Boons and Wells labs for helpful discussions and support throughout my doctorate. I would especially like to thank the following colleagues and friends: Dr. Anthony Prudden and Dr. Pradeep Chopra, who have been an immense help to me since the first day of my PhD and throughout my studies. I am grateful to

Dr. Jeremy Praissman for his numerous valuable insights, his mentorship and patience in training an organic chemist in a biochemistry lab, as well as his help with mass spectrometry data analysis. Thank you to Dr. Saptashwa Chakraborty, Weigang Lu, and Tehai Li for organic synthesis insights, Rachel Bainum and Rob Bridger for help with scheduling and ordering reagents for my projects, David Steen for cell-culture, and Dr. Linda Zhao for mass spectrometry data analysis. Every member of both the Boons and Wells labs have added value and taught me important lessons during my PhD. Thank you to all of my friends and colleagues at the CCRC. It's difficult to capture the immense gratitude for everyone who has helped me in this journey in a just few paragraphs, but I am forever grateful to everyone who has helped me grow during my PhD.

I am grateful to my master's thesis advisor, Dr. Daniela Tapu, for believing in me and urging me to apply to the chemistry PhD program at UGA. Thank you to my fourth grade science teacher, Mrs. Ruth Pinson, for sparking my passion for science at a young age.

I would also like to thank my family and friends for their continuous encouragement over the years. My friends Morgan White and Dr. Tatiana Moro have been more help and support than they will ever know. Thank you especially to my mom and dad, Rob and Lisa Carter, my sister, Taylor Carter, and my grandparents, Johnny and Gladys McKinney, for your love and support that has kept me motivated throughout the years.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE OVERVIEW

1.1 The Role of *O*-mannosylation in Nervous System and Muscle Development

The addition of carbohydrates, or glycans, to extracellular proteins and lipids is critical for physiological processes across all known organisms, and this post-translational process is known as glycosylation.[1] The diverse array of glycan structures on glycoproteins and glycolipids expressed in the nervous system play key roles neural development, synaptic plasticity, and central nervous system (CNS) maintenance, such as cell differentiation and neuronal migration.[2,3] The substantially sized negatively charged, hydrophilic region of polysialic acid (PSA), for example, facilitates neuronal cell motility by modulating the distance between cells. Furthermore, PSA regulates hippocampal synaptic plasticity, cell differentiation, and promotes repair in lesioned CNS tissues.[2,3] Several glycans, including human natural killer-1 glycan (HNK-1), polysialic acid (PSA), and proteoglycan heparan sulfate (HS), have been shown to play key role in long-term potentiation (LTP), the fundamental molecular mechanism driving spatial memory.[2–6] Within the CNS, an abundance of HNK-1 is expressed in regions involved in neogenesis and synaptic plasticity. Present on both glycoproteins and glycolipids, HNK-1 is comprised of 3’ sulfated glucuronic acid attached to lactosamine. In particular, HNK-1 is expressed in the dentate gyrus of the hippocampus, perineuronal nets, and neural stem cells, where this glycan plays a key role in learning and memory, specifically modulating LTP.[4–6] The

HNK-1 epitope on branched *O*-linked mannose (*O*-mannose or *O*-man) glycans modulates brain development and remyelination upon myelin sheath damage.[5,7–9]

Comprising at least 30% of glycans in the brain, *O*-mannose glycans are essential for proper nervous system development through their involvement in axon guidance and neuronal migration.[10–13] *O*-mannose glycans are defined as mannose covalently linked to the hydroxyl of a serine or threonine amino acid residue.[1] Defective *O*-mannosylation is tied to nervous system disorders including cobblestone lissencephaly and multiple sclerosis in which the myelin sheath is compromised.[9] *O*-mannose glycans also play a key role in muscle structure and function. Furthermore, deficient *O*-mannosylation leads to congenital muscular dystrophies (CMDs), many of which exhibit brain and eye developmental abnormalities.[14]

Aberrant *O*-mannosylation of the protein α-dystroglycan (α-DG) results in dystroglycanopathies, a subset of CMDs which involve neurodevelopmental defects along with a range muscular dystrophic phenotypes, ranging from the most severe case of Walker–Warburg syndrome (WWS) to the less severe limb-girdle muscular dystrophies (LGMDs). [13,15–18] α-DG is a peripheral part of the highly glycosylated basement membrane receptor dystroglycan, which plays a key role in several physiological processes, including brain and central nervous system development, neuronal synapse maintenance, and regulation of skeletal muscle integrity. Post-translationally, dystroglycan is cleaved into extracellular α-DG, which is non-covalently linked to the transmembrane subunit β-DG. While hypoglycosylation of the protein dystrophin leads to muscular dystrophies, defects in *O*-mannosylation of α-DG results in both muscular dystrophies and neurodevelopmental abnormalites.[18] The most well-defined role of α-DG is the glycan dependent linkage it serves between the actin cytoskeleton and extracellular matrix (ECM) of

cells.[12] This linkage takes place via a repeating disaccharide (-Xylα1,3-GlcAβ1,3-)$_n$, termed matriglycan, acting as the receptor for the laminin globular-containing ECM proteins. Although α-DG is highly glycosylated with *O*-linked mannosides, only three sites (T317, T319, and T379) bind to the laminin globular domains of the extracellular matrix via the matriglyan linkage. These sites are modified by highly selective glycosyltransferase POMGNT2 and are termed *O*-man core M3 glycans. α-DG is the only known protein to have this M3 modification. The *O*-mannose biosynthetic pathway is further discussed section 1.2.

In addition to α-DG, an increasing number of protein targets for *O*-mannosylation in the brain and additional biosynthetic pathways have been identified.[19] For example, *O*-mannosylation of cadherins was recently discovered.[20] Previous studies in the Wells laboratory have demonstrated that mouse brains lacking α-DG did not show a significant difference in *O*-mannosylation compared to the control group with α-DG.[21] This finding implies that other proteins are *O*-mannosylated in the brain. Glycoproteomic analysis has revealed that many *O*-mannosylated proteins exist, including cadherins and plexins.[20,22–24] The Clausen group demonstrated that cadherins and plexins contain non-extended M0 structures, and these proteins are *O*-mannosylated by an unknown family of novel protein *O*-mannosyltransferases.[19] Cadherins are cell adhesion molecules which are critical for neuronal circuit assembly, because they mediate neuronal self-recognition so that there is no overlap of isoneural branches.[25] *O*-mannosylation of cadherin is distinguished from that of α-DG in that the glycans of cadherin are not further extended.[19]

While muscular dystrophy has thus far been the most widely studied effect of defective glycosyation of α-dystroglycan, limited research has been carried on the defects in *O*-

mannosylation leading to neurodevelopmental abnormalities. Dystroglycan glycosylation plays a critical role in dendritic stratification, axon guidance, and neuronal migration. A recent study showed that defects in glycosylation of dystroglycan lead to defective axon guidance and neuronal migration as well as mosaic spacing of neurons during retinal development.[16] *O*-mannosylation is prevalent in the brain, but not well understood from a neurological standpoint.[13,16] The *O*-mannosylation of the dystroglycan protein is involved in many neurological developmental pathways. Studies have shown that the protein dystroglycan is expressed in neurons and glia, connective tissue of the nervous system, in the developing brain.[12,26–28]

1.2 The Mammalian *O*-Mannosylation Biosynthetic Pathway

The biosynthesis of *O*-mannose is initiated in the ER by a complex of Protein *O*-mannosyltransferase 1 (POMT1) and Protein *O*-mannosyltransferase 2 (POMT2). The POMT1/POMT2 complex catalyzes the transfer of mannose from Dol-P-Man to the hydroxyl of serine or threonine residues in alpha linkage to α-DG. A serine or threonine residue linked with one α-mannose is referred to as core M0. Core M0 can be modified by two enzymes, POMGNT1 and POMGNT2, which mediate the branching point of *O*-mannosylation by facilitating the addition N-acetylglucosamine (GlcNAc) in either a β-1,2 or β-1,4 linkage, respectively. Most commonly, POMGNT1 adds GlcNAc in a β-1,2 linkage to form the core M1.[17,18] However, the function of core M1 glycans remains unknown. Core M1 glycans account for 15% of brain *O*-glycans.[18,21] The core M1 structure can be branched by another glycotransferase, MGAT5B, which adds GlcNAc in a β-1,6 position. In contrast, POMGNT2 adds GlcNAc to the core M0 in

a β-1,4 linkage, giving rise to the rare core M3 glycans (**Figure 1.1a**). Interestingly, *O*-mannose modified α-DG encounters POMGNT2 in the ER before POMGNT1 in the cis golgi.[17,18] This implies that POMNGT2 must exhibit substrate sequence selectivity beyond simply an *O*-Man modified amino acid.[17,18]

M1 and M2 glycan cores can be further extended with galactose, fucose, and sialic acids.[29] The core M3 glycans are extended with β-1,3 linked GalNAc by B3GALNT2 and phosphorylated by POMK, creating the phosphotrisacharide core M3 glycan structure. FKTN and FKRP insert ribitol-1-phosphates onto the phosophotrisacharide.[30] Together, TMEM5 and B4GAT1 prime the ribitol for matriglycan addition. The priming enzyme TMEM5 adds a xylose to the distal ribitol, while B4GAT1 catalyzes the addition of glucuronic acid in a β-1,4 linkage to the xylose. Furthermore, B4GAT1 primes the M3 glycan for matriglycan addition on α-DG. LARGE-1 then catalyzes the addition of matriglycan, a repeating disaccharide (a-1,3-linked-xylose- β-1,3-glucuronic acid) that binds LG domain-containing proteins in the ECM.[18] No reported method currently exists to identify M3 cores without the presence of the repeating disaccharide. Therefore, unextended M3 cores could exist, but there is no current tool to identify the unextended core independently of matriglycan (**Figure 1.1b**).

The development of chemical tools to probe sites of *O*-mannosylation could provide further mechanistic insights and enable opportunities for therapeutic targeting.

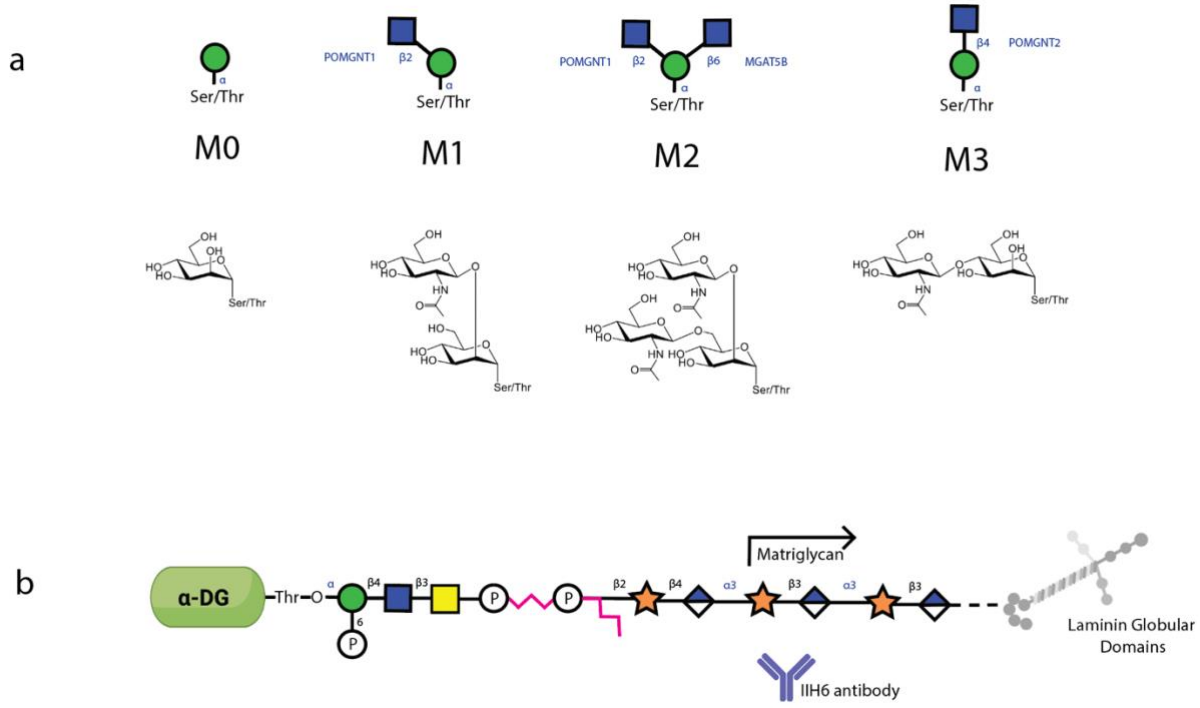


Figure 1.1 *O*-mannosylation biosynthetic pathway. (a) Four primary core *O*-mannose glycan structures are involved in the *O*-mannosylation pathway. Enzymes that catalyze the sugar transfer are indicated in blue. (b) Core M3 structure fully extended by matriglycan. The IIH6 antibody recognizes the disaccharide repeats of matriglycan.

1.3 Chemical Tools for Enhanced Detection of Glycoprotein Sites

The study of cellular glycosylation remains a challenge due to the structural complexity and diversity of glycans. In contrast to the template-driven biosynthesis of proteins and nucleic acids, complex carbohydrates are assembled via stepwise, enzymatic additions of monosaccharide building blocks.[1] While the resulting diversity of these complex oligosaccharides gives rise to a plethora of biological roles for glycans, the vast array of

heterogeneous complex glycans also leads to challenges in the detection of glycosylated protein sites. For example, glycan-binding lectins and antibodies often have the drawback of low affinity and lack of specificity. Additionally, several chemical methods, such as periodic acid for the detection of sialic acid, often cause irreversible damage to the glycan chains, preventing downstream analysis and applications.[1]

As a consequence of these challenges in probing glycosylation, there is a desire for chemical probes that act orthogonally to native cell components and do not disrupt the inherent composition and nature of the cell. Bioorthogonal chemistry, a methodology developed by Carolyn Bertozzi, overcomes this limitation in glycan profiling by employing chemical enrichment tools that are inert to biological environments.[31,32] This field was pioneered through incorporation of the Staudinger reaction.[31] In this approach, azides were installed within the cell surface by metabolism of a synthetic azidosugar. The azide-modified glycan sites were then reacted with a biotinylated triarylphosphine to form stable cell-surface labels. This methodology was shown to be effective both *in vitro* and *in vivo*.[32] However, this Staudinger ligation is limited by its slow reaction at physiological temperatures.

Simultaneous to Bertozzi's pioneering work in bioorthogonal chemistry, Barry Sharpless at the Scripps Research Institute developed a highly selective method to conjugate two complex molecules, regardless of surrounding chemical functionality. Appropriately termed "click chemistry", this methodology would allow for efficient production of complex molecules while circumventing the use of protecting groups.[33] In 2002, Barry Sharpless and Morten Meldal independently reported what is now the epitome of click chemistry – the Cu (I) catalyzed cycloaddition of an azide and terminal alkene (**Figure 1.2a**).[34,35] Employment of the Cu (I)

catalyst not only allowed for formation of the previously unattainable triazole adduct, but also for completion of the reaction within only minutes. Morten Meldal demonstrated the powerful utility of click chemistry by using this methodology to link together peptides and sugar moieties on the solid phase.[35] Click chemistry quickly became the new paradigm for assembly of previously unattainable highly functionalized biomolecules. This efficient reaction inspired Bertozzi to incorporate click chemistry into the bioorthoganol methodology.

At the intersection of bioorthogonal and click chemistry, strain-promoted azide-alkyne cycloaddition (SPAAC) was developed by Bertozzi.[36] This strategy employs ring strain to facilitate formation of the triazole ring, giving rise to substrate-ligand binding without the use of a cytotoxic copper catalyst. Overcoming this limitation of cellular toxicity, SPAAC offers the advantage of probing biological systems with click chemistry methodology while circumventing unwanted side reactions resulting from the Cu(I) catalyst (**Figure 1.2b**).

The application of both Cu (I)-catalyzed and SPAAC click chemistries has proven to be expansive and has found vast utility in imaging, profiling, and discovery of biomolecules, including proteins, glycans, lipids, and ligands for small molecule drugs. Furthermore, bioorthogonal click chemistry has been used to construct antibody-drug conjugates and vaccines. In recognition of the immense impact of bioorthogonal chemistry and click chemistry, the Chemistry Nobel Committee of the Royal Swedish Academy of Sciences awarded the 2022 Nobel Prize in Chemistry to Sharpless, Meldal, and Bertozzi.

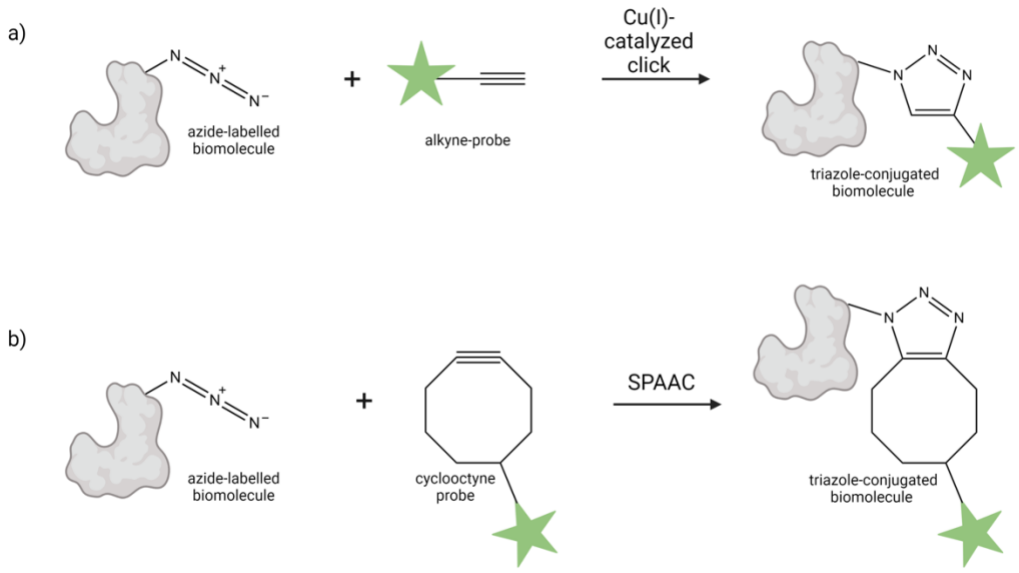


Figure 1.2. Click Chemistry. a) Cu-catalyzed click chemistry; b) SPAAC bioorthogonal labeling induced by ring strain. Created with Biorender.com

1.3.1. Selective Exo-enzymatic Labelling (SEEL)

The SPAAC bioorthogonal labelling strategy gave rise to labeling of cells with metabolic chemical reporters. Metabolic labelling has two primary limitations. First, natural sugar precursors compete with the metabolic chemical reporter. Second, this strategy is not amenable to probing subclasses of glycans modified by particular glycosyltransferases. Capitalizing on the advantages of SPAAC and overcoming the limitations of metabolic labelling, the Boons laboratory developed selective exo-enzymatic labelling (SEEL) in which chemical reporters are installed directly onto endogenous glycans on the cell surface via recombinant or purified glycosyltransferases. [37,38]

SEEL offers the advantage of glycan subclass specificity.[37,38] Through the SEEL methodology in combination with glycoproteomic analysis, selective labeling of both N- and O-glycans has been achieved. For example, CMP-Neu5Ac has been selectively installed on N-glycans with an α-2,6 linkage using sialyltransferase ST6GAL1 and on *O*-glycans with α-2,3 linkage using ST3GAL1. [37,38] Given these successful examples of SEEL enrichment of specific subclasses of glycans, we envisioned using the SEEL methodology to enrich for *O*-mannosylated sites, and we hypothesized that we would uncover previously undetectable, novel M3 sites. This enrichment methodology is detailed in Chapter 2.

1.3.2. *C*-glycosides as Metabolically Stable Therapeutic Agents

Glycopeptide mimetics are another set of tools used to study glycan-protein interactions. Generation of antibodies using these glycopeptides can be used to detect specific glycoprotein sites.[39,40] Compared to traditional methods of antibody enrichment, such as immunoprecipitation, glycopeptide derived antibodies offer higher specificity and design flexibility. For example, oligomannose glycopeptide conjugates elicit antibodies that preferentially recognize the core glycan rather than the extremities.[40] Another drawback of traditional antibodies is that they often recognize *O*-glycan structures with low affinity toward a monovalent epitope. A synthetic glycopeptide library overcame this challenge through the generation of novel anti-mucin 1 antibodies with intentionally designed glycan specificities.[41]

Despite these successful examples of glycopeptide derived antibodies, the utility of glycopeptide conjugates has been limited by their susceptibility to enzymatic degradation and instability to acidic environments. Actually, the aforementioned example of glycopeptide derived

antibodies that specifically recognize the core glycan instead of its extremities is hypothesized to be a direct result of mannosidase trimming.[40] *C*-glycopeptides offer a means to circumvent the fragility of traditional glycopeptides in that these glycoconjugates are robust to enzymatic degradation and acidity. In *C*-glycopeptide conjugates, the typical C-O bond of the glycoside moiety is replaced with a C-C bond. Compared the C-O bond in natural glycosides, the C-C bond in *C*-glycosides is resistant to hydrolysis, offering potential opportunities for the development more robust antigens.

C-glycopeptides have been shown to have similar, and in some cases enhanced, biological activity compared to that of their natural glycopeptide counterparts. A *C*-linked mimic of a high-mannose N-linked glycopeptide was shown to be resistant to N-glycanase-catalyzed hydrolysis and displayed inhibitory activity toward the enzyme.[42] Pro-Xylane™ is a *C*-glycoside employed as a cosmeceutical to improve skin barrier function by increasing glycosaminoglycan synthesis in fibroblasts, and subsequently boosting collagen production.[43–46] *C*-glycoside analogues, such as such as dapaglifozin, canagliflozin, and empaglioflazin, have found therapeutic utility as SGLT2 inhibitors against type II diabetes.[43,47–51] Furthermore, enhanced activity of *C*-glycoside mimetics compared to *O*-glycosides has been demonstrated. For example, the *C*-glycoside analogue of the immunostimulant α-galactosylceramide (KRN7000) was shown to be 100x more effective than the *O*-glycoside in preventing melanoma spread in the lungs of melanoma challenged mice.[52]

1.3.3. Glycopeptide Mimetics as Antigens to Core Glycans

Given the enhanced stability and demonstrated biological activity of *C*-glycopeptides compared to *O*-glycopeptides, we envisioned engineering *C*-glycopeptides antigens to the *O*-mannose core glycans as a means of direct detection to all four of the core *O*-mannose structures. Glycoproteomic analysis has profoundly aided the discovery of M0 and M1 *O*-mannosyated sites.[19–21,29] Direct detection of these sites via intentional glycopeptide antibody design has the potential to further aid in the discovery of novel *O*-mannose sites. Currently, the only means of detection for M3 core *O*-mannose sites is the IIH6 antibody, which does not recognize the core glycan, but rather the extended polymer matriglycan. A *C*-glycopeptide mimetic of the core M3 structure would provide the opportunity for direct detection of the core glycan, potentially leading to discovery of previously unknown M3 sites of *O*-mannosylation. Chapter 3 details the synthetic strategy to create *C*-mannosyl-Threonine that can be used to generate such *C*-glycopeptides.

1.4. Overview of Dissertation

This dissertation describes research with the aim to overcome the limitations in detection of *O*-mannosylated sites, leveraging bioorthoganol tagging and *C*-glycoside chemistry (**Figure 1.2a**). In one approach detailed in Chapter 2, *O*-mannosylated sites were enriched by bioorthogonally labelling living cells with azido-modified UDP-GlcNAc (UDP-GlcNAz) (**Figure 1.2b**). In Approach 2, a carbon-linked (*C*-linked) glycoside was synthesized to develop potential antibodies to the core *O*-mannose glycans. Chapter 3 describes the synthesis of a novel *C*-glycoside, *C*-mannose-Thr (**Figure 1.2c**). Chapter 4 details α-DG-based glycopeptide

synthesis and NMR conformational analysis with the long-term goal of creating antigens to recognize core *O*-mannose glycans, and Chapter 5 provides future directions for these projects with the focus of enhanced detection of *O*-mannosylated sites through further development of specifically designed chemical tools.

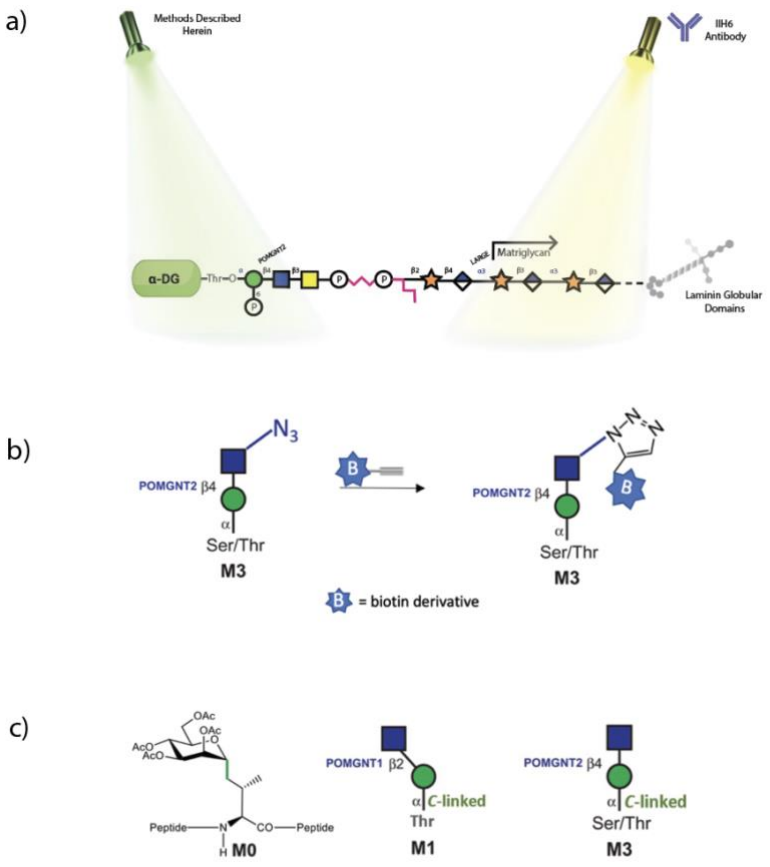


Figure 1.3. Enrichment of *O*-mannosylated sites via a) direct detection methods as opposed to the IIH6 antibody recognizing matriglycan; b) clickable glycans using SEEL; c) *C*-glycopeptides mimetics to core *O*-mannose glycans.

RERFERENCES

(1) *Essentials of Glycobiology*, 4th ed.; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2022.

(2) Kleene, R.; Schachner, M. Glycans and Neural Cell Interactions. *Nat. Rev. Neurosci.* **2004**, *5* (3), 195–208. https://doi.org/10.1038/nrn1349.

(3) Iqbal, S.; Ghanimi Fard, M.; Everest-Dass, A.; Packer, N. H.; Parker, L. M. Understanding Cellular Glycan Surfaces in the Central Nervous System. *Biochem. Soc. Trans.* **2018**, *47* (1), 89–100. https://doi.org/10.1042/BST20180330.

(4) Morise, J.; Takematsu, H.; Oka, S. The Role of Human Natural Killer-1 (HNK-1) Carbohydrate in Neuronal Plasticity and Disease. *Biochim. Biophys. Acta Gen. Subj.* **2017**, *1861* (10), 2455–2461. https://doi.org/10.1016/j.bbagen.2017.06.025.

(5) Yamamoto, S.; Oka, S.; Inoue, M.; Shimuta, M.; Manabe, T.; Takahashi, H.; Miyamoto, M.; Asano, M.; Sakagami, J.; Sudo, K.; Iwakura, Y.; Ono, K.; Kawasaki, T. Mice Deficient in Nervous System-Specific Carbohydrate Epitope HNK-1 Exhibit Impaired Synaptic Plasticity and Spatial Learning. *J. Biol. Chem.* **2002**, *277* (30), 27227–27231. https://doi.org/10.1074/jbc.C200296200.

(6) Rollenhagen, A.; Czaniera, R.; Albert, M.; Wintergerst, E. S.; Schachner, M. Immunocytological Localization of the HNK-1 Carbohydrate in Murine Cerebellum, Hippocampus and Spinal Cord Using Monoclonal Antibodies with Different Epitope

Specificities. *J. Neurocytol.* **2001**, *30* (4), 337–351. https://doi.org/10.1023/A:1014412530722.

(7) Gao, T.; Yan, J.; Liu, C.-C.; Palma, A. S.; Guo, Z.; Xiao, M.; Chen, X.; Liang, X.; Chai, W.; Cao, H. Chemoenzymatic Synthesis of O-Mannose Glycans Containing Sulfated or Nonsulfated HNK-1 Epitope. *J. Am. Chem. Soc.* **2019**, *141* (49), 19351–19359. https://doi.org/10.1021/jacs.9b08964.

(8) Yuen, C. T.; Chai, W.; Loveless, R. W.; Lawson, A. M.; Margolis, R. U.; Feizi, T. Brain Contains HNK-1 Immunoreactive O-Glycans of the Sulfoglucuronyl Lactosamine Series That Terminate in 2-Linked or 2,6-Linked Hexose (Mannose). *J. Biol. Chem.* **1997**, *272* (14), 8924–8931. https://doi.org/10.1074/jbc.272.14.8924.

(9) Kanekiyo, K.; Inamori, K.; Kitazume, S.; Sato, K.; Maeda, J.; Higuchi, M.; Kizuka, Y.; Korekane, H.; Matsuo, I.; Honke, K.; Taniguchi, N. Loss of Branched O-Mannosyl Glycans in Astrocytes Accelerates Remyelination. *J. Neurosci.* **2013**, *33* (24), 10037–10047. https://doi.org/10.1523/JNEUROSCI.3137-12.2013.

(10) Yoshida, A.; Kobayashi, K.; Manya, H.; Taniguchi, K.; Kano, H.; Mizuno, M.; Inazu, T.; Mitsuhashi, H.; Takahashi, S.; Takeuchi, M.; Herrmann, R.; Straub, V.; Talim, B.; Voit, T.; Topaloglu, H.; Toda, T.; Endo, T. Muscular Dystrophy and Neuronal Migration Disorder Caused by Mutations in a Glycosyltransferase, POMGnT1. *Dev. Cell* **2001**, *1* (5), 717–724. https://doi.org/10.1016/S1534-5807(01)00070-3.

(11) Beltrán-Valero de Bernabé, D.; Currier, S.; Steinbrecher, A.; Celli, J.; van Beusekom, E.; van der Zwaag, B.; Kayserili, H.; Merlini, L.; Chitayat, D.; Dobyns, W. B.; Cormand, B.; Lehesjoki, A.-E.; Cruces, J.; Voit, T.; Walsh, C. A.; van Bokhoven, H.; Brunner, H. G.

Mutations in the O-Mannosyltransferase Gene POMT1 Give Rise to the Severe Neuronal Migration Disorder Walker-Warburg Syndrome. *Am. J. Hum. Genet.* **2002**, *71* (5), 1033–1043. https://doi.org/10.1086/342975.

(12) Satz, J. S.; Ostendorf, A. P.; Hou, S.; Turner, A.; Kusano, H.; Lee, J. C.; Turk, R.; Nguyen, H.; Ross-Barta, S. E.; Westra, S.; Hoshi, T.; Moore, S. A.; Campbell, K. P. Distinct Functions of Glial and Neuronal Dystroglycan in the Developing and Adult Mouse Brain. *J. Neurosci.* **2010**, *30* (43), 14560–14572. https://doi.org/10.1523/JNEUROSCI.3247-10.2010.

(13) Wright, K. M.; Lyon, K. A.; Leung, H.; Leahy, D. J.; Ma, L.; Ginty, D. D. Dystroglycan Organizes Axon Guidance Cue Localization and Axonal Pathfinding. *Neuron* **2012**, *76* (5), 931–944. https://doi.org/10.1016/j.neuron.2012.10.009.

(14) Martin, P. T. Congenital Muscular Dystrophies Involving the O-Mannose Pathway. *Curr. Mol. Med.* **2007**, *7* (4), 417–425.

(15) Stalnaker, S. H.; Stuart, R.; Wells, L. Mammalian O-Mannosylation: Unsolved Questions of Structure/Function. *Curr. Opin. Struct. Biol.* **2011**, *21* (5), 603–609. https://doi.org/10.1016/j.sbi.2011.09.001.

(16) Clements, R.; Turk, R.; Campbell, K. P.; Wright, K. M. Dystroglycan Maintains Inner Limiting Membrane Integrity to Coordinate Retinal Development. *J. Neurosci. Off. J. Soc. Neurosci.* **2017**, *37* (35), 8559–8574. https://doi.org/10.1523/JNEUROSCI.0946-17.2017.

(17) Halmo, S. M.; Singh, D.; Patel, S.; Wang, S.; Edlin, M.; Boons, G.-J.; Moremen, K. W.; Live, D.; Wells, L. Protein O-Linked Mannose β-1,4-N-Acetylglucosaminyl-Transferase 2

(POMGNT2) Is a Gatekeeper Enzyme for Functional Glycosylation of α-Dystroglycan. *J. Biol. Chem.* **2017**, *292* (6), 2101–2109. https://doi.org/10.1074/jbc.M116.764712.

(18) Sheikh, M. O.; Halmo, S. M.; Wells, L. Recent Advancements in Understanding Mammalian O-Mannosylation. *Glycobiology* **2017**, *27* (9), 806–819. https://doi.org/10.1093/glycob/cwx062.

(19) Larsen, I. S. B.; Narimatsu, Y.; Joshi, H. J.; Siukstaite, L.; Harrison, O. J.; Brasch, J.; Goodman, K. M.; Hansen, L.; Shapiro, L.; Honig, B.; Vakhrushev, S. Y.; Clausen, H.; Halim, A. Discovery of an O-Mannosylation Pathway Selectively Serving Cadherins and Protocadherins. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (42), 11163–11168. https://doi.org/10.1073/pnas.1708319114.

(20) Vester-Christensen, M. B.; Halim, A.; Joshi, H. J.; Steentoft, C.; Bennett, E. P.; Levery, S. B.; Vakhrushev, S. Y.; Clausen, H. Mining the O-Mannose Glycoproteome Reveals Cadherins as Major O-Mannosylated Glycoproteins. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (52), 21018–21023. https://doi.org/10.1073/pnas.1313446110.

(21) Stalnaker, S. H.; Aoki, K.; Lim, J.-M.; Porterfield, M.; Liu, M.; Satz, J. S.; Buskirk, S.; Xiong, Y.; Zhang, P.; Campbell, K. P.; Hu, H.; Live, D.; Tiemeyer, M.; Wells, L. Glycomic Analyses of Mouse Models of Congenital Muscular Dystrophy. *J. Biol. Chem.* **2011**, *286* (24), 21180–21190. https://doi.org/10.1074/jbc.M110.203281.

(22) Baenziger, J. U. O-Mannosylation of Cadherins. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (52), 20858–20859. https://doi.org/10.1073/pnas.1321827111.

(23) Bartels, M. F.; Winterhalter, P. R.; Yu, J.; Liu, Y.; Lommel, M.; Möhrlen, F.; Hu, H.; Feizi, T.; Westerlind, U.; Ruppert, T.; Strahl, S. Protein O-Mannosylation in the Murine Brain:

Occurrence of Mono-O-Mannosyl Glycans and Identification of New Substrates. *PloS One* **2016**, *11* (11), e0166119. https://doi.org/10.1371/journal.pone.0166119.

(24) Winterhalter, P. R.; Lommel, M.; Ruppert, T.; Strahl, S. O-Glycosylation of the Non-Canonical T-Cadherin from Rabbit Skeletal Muscle by Single Mannose Residues. *FEBS Lett.* **2013**, *587* (22), 3715–3721. https://doi.org/10.1016/j.febslet.2013.09.041.

(25) Rubinstein, R.; Thu, C. A.; Goodman, K. M.; Wolcott, H. N.; Bahna, F.; Mannepalli, S.; Ahlsen, G.; Chevee, M.; Halim, A.; Clausen, H.; Maniatis, T.; Shapiro, L.; Honig, B. Molecular Logic of Neuronal Self-Recognition through Protocadherin Domain Interactions. *Cell* **2015**, *163* (3), 629–642. https://doi.org/10.1016/j.cell.2015.09.026.

(26) Zaccaria, M. L.; Di Tommaso, F.; Brancaccio, A.; Paggi, P.; Petrucci, T. C. Dystroglycan Distribution in Adult Mouse Brain: A Light and Electron Microscopy Study. *Neuroscience* **2001**, *104* (2), 311–324. https://doi.org/10.1016/s0306-4522(01)00092-6.

(27) Henion, T. R.; Qu, Q.; Smith, F. I. Expression of Dystroglycan, Fukutin and POMGnT1 during Mouse Cerebellar Development. *Brain Res. Mol. Brain Res.* **2003**, *112* (1–2), 177–181. https://doi.org/10.1016/s0169-328x(03)00055-x.

(28) Ohtsuka-Tsurumi, E.; Saito, Y.; Yamamoto, T.; Voit, T.; Kobayashi, M.; Osawa, M. Co-Localization of Fukutin and Alpha-Dystroglycan in the Mouse Central Nervous System. *Brain Res. Dev. Brain Res.* **2004**, *152* (2), 121–127. https://doi.org/10.1016/j.devbrainres.2004.06.006.

(29) Stalnaker, S. H.; Hashmi, S.; Lim, J.-M.; Aoki, K.; Porterfield, M.; Gutierrez-Sanchez, G.; Wheeler, J.; Ervasti, J. M.; Bergmann, C.; Tiemeyer, M.; Wells, L. Site Mapping and Characterization of O-Glycan Structures on Alpha-Dystroglycan Isolated from Rabbit

Skeletal Muscle. *J. Biol. Chem.* **2010**, *285* (32), 24882–24891. https://doi.org/10.1074/jbc.M110.126474.

(30) Praissman, J. L.; Willer, T.; Sheikh, M. O.; Toi, A.; Chitayat, D.; Lin, Y.-Y.; Lee, H.; Stalnaker, S. H.; Wang, S.; Prabhakar, P. K.; Nelson, S. F.; Stemple, D. L.; Moore, S. A.; Moremen, K. W.; Campbell, K. P.; Wells, L. The Functional O-Mannose Glycan on α-Dystroglycan Contains a Phospho-Ribitol Primed for Matriglycan Addition. *eLife* **2016**, *5*, e14473. https://doi.org/10.7554/eLife.14473.

(31) Saxon, E.; Bertozzi, C. R. Cell Surface Engineering by a Modified Staudinger Reaction. *Science* **2000**, *287* (5460), 2007–2010. https://doi.org/10.1126/science.287.5460.2007.

(32) Prescher, J. A.; Dube, D. H.; Bertozzi, C. R. Chemical Remodelling of Cell Surfaces in Living Animals. *Nature* **2004**, *430* (7002), 873–877. https://doi.org/10.1038/nature02791.

(33) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chem. Int. Ed.* **2001**, *40* (11), 2004–2021. https://doi.org/10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.0.CO;2-5.

(34) *A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes - Rostovtsev - 2002 - Angewandte Chemie International Edition - Wiley Online Library*. https://onlinelibrary.wiley.com/doi/full/10.1002/1521-3773%2820020715%2941%3A14%3C2596%3A%3AAID-ANIE2596%3E3.0.CO%3B2-4 (accessed 2023-06-16).

(35) Tornøe, C. W.; Christensen, C.; Meldal, M. Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal

Alkynes to Azides. *J. Org. Chem.* **2002**, *67* (9), 3057–3064. https://doi.org/10.1021/jo011148j.

(36) Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. A Strain-Promoted [3 + 2] Azide−Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *J. Am. Chem. Soc.* **2004**, *126* (46), 15046–15047. https://doi.org/10.1021/ja044996f.

(37) Sun, T.; Yu, S.-H.; Zhao, P.; Meng, L.; Moremen, K. W.; Wells, L.; Steet, R.; Boons, G.-J. One-Step Selective Exoenzymatic Labeling (SEEL) Strategy for the Biotinylation and Identification of Glycoproteins of Living Cells. *J. Am. Chem. Soc.* **2016**, *138* (36), 11575–11582. https://doi.org/10.1021/jacs.6b04049.

(38) Capicciotti, C. J.; Zong, C.; Sheikh, M. O.; Sun, T.; Wells, L.; Boons, G.-J. Cell-Surface Glyco-Engineering by Exogenous Enzymatic Transfer Using a Bifunctional CMP-Neu5Ac Derivative. *J. Am. Chem. Soc.* **2017**, *139* (38), 13342–13348. https://doi.org/10.1021/jacs.7b05358.

(39) Brooks, C. L.; Schietinger, A.; Borisova, S. N.; Kufer, P.; Okon, M.; Hirama, T.; MacKenzie, C. R.; Wang, L.-X.; Schreiber, H.; Evans, S. V. Antibody Recognition of a Unique Tumor-Specific Glycopeptide Antigen. *Proc. Natl. Acad. Sci.* **2010**, *107* (22), 10056–10061. https://doi.org/10.1073/pnas.0915176107.

(40) Nguyen, D. N.; Xu, B.; Stanfield, R. L.; Bailey, J. K.; Horiya, S.; Temme, J. S.; Leon, D. R.; LaBranche, C. C.; Montefiori, D. C.; Costello, C. E.; Wilson, I. A.; Krauss, I. J. Oligomannose Glycopeptide Conjugates Elicit Antibodies Targeting the Glycan Core Rather than Its Extremities. *ACS Cent. Sci.* **2019**, *5* (2), 237–249. https://doi.org/10.1021/acscentsci.8b00588.

(41) Naito, S.; Takahashi, T.; Onoda, J.; Uemura, S.; Ohyabu, N.; Takemoto, H.; Yamane, S.; Fujii, I.; Nishimura, S.-I.; Numata, Y. Generation of Novel Anti-MUC1 Monoclonal Antibodies with Designed Carbohydrate Specificities Using MUC1 Glycopeptide Library. *ACS Omega* **2017**, *2* (11), 7493–7505. https://doi.org/10.1021/acsomega.7b00708.

(42) Michael, K.; Wittmann, V.; König, W.; Sandow, J.; Kessler, H. S- and C-Glycopeptide Derivatives of an LH-RH Agonist. *Int. J. Pept. Protein Res.* **1996**, *48* (1), 59–70. https://doi.org/10.1111/j.1399-3011.1996.tb01107.x.

(43) Yang, Y.; Yu, B. Recent Advances in the Chemical Synthesis of C-Glycosides. *Chem. Rev.* **2017**, *117* (19), 12281–12356. https://doi.org/10.1021/acs.chemrev.7b00234.

(44) Leseurre, L.; Merea, C.; Paule, S. D. de; Pinchart, A. Eco-Footprint: A New Tool for the "Made in Chimex" Considered Approach. *Green Chem.* **2014**, *16* (3), 1139–1148. https://doi.org/10.1039/C3GC42201A.

(45) Cavezza, A.; Boulle, C.; Guéguiniat, A.; Pichaud, P.; Trouille, S.; Ricard, L.; Dalko-Csiba, M. Synthesis of Pro-Xylane: A New Biologically Active C-Glycoside in Aqueous Media. *Bioorg. Med. Chem. Lett.* **2009**, *19* (3), 845–849. https://doi.org/10.1016/j.bmcl.2008.12.037.

(46) Pineau, N.; Carrino, D. A.; Caplan, A. I.; Breton, L. Biological Evaluation of a New C-Xylopyranoside Derivative (C-Xyloside) and Its Role in Glycosaminoglycan Biosynthesis. *Eur. J. Dermatol. EJD* **2011**, *21* (3), 359–370. https://doi.org/10.1684/ejd.2011.1340.

(47) Chao, E. C.; Henry, R. R. SGLT2 Inhibition--a Novel Strategy for Diabetes Treatment. *Nat. Rev. Drug Discov.* **2010**, *9* (7), 551–559. https://doi.org/10.1038/nrd3180.

(48) Wang, X.; Zhang, L.; Byrne, D.; Nummy, L.; Weber, D.; Krishnamurthy, D.; Yee, N.; Senanayake, C. H. Efficient Synthesis of Empagliflozin, an Inhibitor of SGLT-2, Utilizing an AlCl3-Promoted Silane Reduction of a β-Glycopyranoside. *Org. Lett.* **2014**, *16* (16), 4090–4093. https://doi.org/10.1021/ol501755h.

(49) Henschke, J. P.; Lin, C.-W.; Wu, P.-Y.; Tsao, W.-S.; Liao, J.-H.; Chiang, P.-C. β-Selective C-Arylation of Diisobutylaluminum Hydride Modified 1,6-Anhydroglucose: Synthesis of Canagliflozin without Recourse to Conventional Protecting Groups. *J. Org. Chem.* **2015**, *80* (10), 5189–5195. https://doi.org/10.1021/acs.joc.5b00601.

(50) Guo, C.; Hu, M.; DeOrazio, R. J.; Usyatinsky, A.; Fitzpatrick, K.; Zhang, Z.; Maeng, J.-H.; Kitchen, D. B.; Tom, S.; Luche, M.; Khmelnitsky, Y.; Mhyre, A. J.; Guzzo, P. R.; Liu, S. The Design and Synthesis of Novel SGLT2 Inhibitors: C-Glycosides with Benzyltriazolopyridinone and Phenylhydantoin as the Aglycone Moieties. *Bioorg. Med. Chem.* **2014**, *22* (13), 3414–3422. https://doi.org/10.1016/j.bmc.2014.04.036.

(51) Cai, W.; Jiang, L.; Xie, Y.; Liu, Y.; Liu, W.; Zhao, G. Design of SGLT2 Inhibitors for the Treatment of Type 2 Diabetes: A History Driven by Biology to Chemistry. *Med. Chem. Shariqah United Arab Emir.* **2015**, *11* (4), 317–328. https://doi.org/10.2174/1573406411666150105105529.

(52) Yang, G.; Schmieg, J.; Tsuji, M.; Franck, R. W. The C-Glycoside Analogue of the Immunostimulant α-Galactosylceramide (KRN7000): Synthesis and Striking Enhancement of Activity. *Angew. Chem. Int. Ed.* **2004**, *43* (29), 3818–3822. https://doi.org/10.1002/anie.200454215.

CHAPTER 2

PROBING O-MANNOSYLATED SITES VIA SELECTIVE EXOENZYMATIC LABELLING (SEEL) SUGGESTS CLOSE ASSOCIATION OF APMAP WITH α-DYSTROGLYCAN

---

[1] Carter, A., Zhao, L., Steen, D., Praissman, J.L., Boons, G.J. and Wells, L. To be submitted to *Glycobiology.*

ABSTRACT

Bioorthogonal chemistry provides an attractive means to label protein-carbohydrate interactions selectively in biological environments while circumventing undesired side reactions. Using a bioorthogonal labelling strategy, sites of *O*-mannosylation were enriched. Employing selective exoenzymatic labelling (SEEL), the *O*-mannose sites on the surface of living cells were extended by POMGNT1 and POMGNT2, glycosyltransferases that mediated the branching point of the *O*-mannosylation pathway. Interactions between α-DG and its ECM ligands require α-DG to be extended by *O*-mannose structures based on the rare M3 core structure, which is highly selectively formed by POMGNT2 facilitating the addition of UDP-GlcNAc in a β-1,4 linkage. The core M3 can be extended with matriglycan, which is a repeating disaccharide that binds LG-domain containing proteins in the ECM. Previously, no method existed to identify the M3 core in absence of its extension with matriglycan. The methodology described herein enriches for *O*-mannosylated probes independently of matriglycan. *O*-mannosylated sites were probed using the azido-form of UDP-GlcNAc and a bioorthogonal biotin tag. Along with known *O*-mannnosylated proteins α-DG and KIAA1549, APMAP and LAMB1 were also enriched. Immunoprecipitation of APMAP revealed that α-DG co-precipitates with APMAP, suggesting a close association of APMAP with α-DG.

INTRODUCTION

Glycosylation of certain proteins by *O*-mannose is essential for growth and development across all higher organisms. Defects in *O*-mannosylation of α-dystroglycan (α-DG), the most well-studied *O*-mannosylated protein, lead to dystroglycanopathies, a subcategory of congenital

muscular dystrophies (CMDs) that involve neurological and ocular defects.[1–5] α-DG is an extracellular component of the dystrophin-glycoprotein complex (DGC) that binds proteins with the laminin-globular (LG) domains, and thus serves as a link between the cytoskeleton of cells and the extracellular matrix (ECM). The most well-characterized role of α-DG in the DGC is the glycan-dependent link that α-DG serves between the actin cytoskeleton and the ECM.[6] Interactions between α-DG and its ECM ligands require α-DG to be extended by *O*-mannose structures. While hypoglycosylation of the protein dystrophin leads to muscular dystrophies, defects in *O*-mannosylation of α-DG results in both muscular dystrophies and neurological defects.[4]

The biosynthesis of *O*-mannose is initiated in the ER by a complex of Protein *O*-mannosyltransferase 1 (POMT1) and Protein *O*-mannosyltransferase 2 (POMT2). The POMT1/POMT2 complex catalyzes the transfer of mannose from Dol-P-Man to the hydroxyl of serine or threonine residues in alpha linkage to α-DG. A serine or threonine residue linked with one α-mannose is referred to as core M0. Core M0 can be modified by two enzymes, POMGNT1 and POMGNT2, which mediate the branching point of *O*-mannosylation by facilitating the addition N-acetylglucosamine (GlcNAc) in either a β-1,2 or β-1,4 linkage, respectively. Most commonly, POMGNT1 adds GlcNAc in a β-1,2 linkage to form the core M1.[3,4] However, the function of core M1 glycans remains unknown. Core M1 glycans account for 15% of brain *O*-glycans.[4,7] The core M1 structure can be branched by another glycotransferase, MGAT5B, which adds GlcNAc in a β-1,6 position. In contrast, POMGNT2 adds GlcNAc to the core M0 in a β-1,4 linkage, giving rise to the rare core M3 glycans (**Figure 1a**). Interestingly, *O*-mannose modified α-DG encounters POMGNT2 in the ER before POMGNT1 in the cis golgi.[3,4] This implies that

POMNGT2 must exhibit substrate sequence selectivity beyond simply an *O*-Man modified amino acid.[3,4] B4GAT1 primes the M3 glycan for matriglycan addition on α-DG. LARGE-1 then catalyzes the addition of matriglycan, a repeating disaccharide (a-1,3-linked-xylose- β-1,3-glucuronic acid) that binds LG domain-containing proteins in the ECM (**Figure 1b**).[4] No method currently exists to identify M3 cores without the presence of the repeating disaccharide. Therefore, unextended M3 cores could exist, but there is no current tool to identify the unextended core independently of matriglycan.

The analysis of the distribution of core M3 glycans has been limited by current detection methods. Specifically, the only method to detect the M3 core is by the IIH6 antibody that recognizes the multiple disaccharide repeats in matriglycan. Therefore, the M3 core would not be detected if it is not fully extended. Currently, no reagents exist to identify the M3 core itself without the repeating disaccharide of matriglycan present. This begs the question- Are there M3 core structures independent of matriglycan and LARGE? The only protein shown to be modified by M3 glycans is α-DG. However, a recent study demonstrated that LARGE2 has the ability to modify proteoglycans with matriglycan in absence of an M3 core.[8] We therefore hypothesized that there exist POMGNT2 sites that are not completely elaborated with matriglycan and cannot be detected with the current chemical toolbox, and we set out to develop an enrichment method to unbiasedly label M3 sites independently of matriglycan.

To address this limitation, we used an azido-modified biomimetic of UDP-GlcNAc (UDP-GlcNAz) to enzymatically label the surface of cells using the SEEL (selective exoenzymatic labeling) (**Figure 1c**). Azide-functionalized nucleotide sugar analogues are commonly employed as bioorthoganol chemical reporters due to their small size, inertness to cellular environments,

and virtual absence in biological systems. Furthermore, azides offer versatility for enrichment strategies in that they can be tagged by Staudinger ligation using modified phosphines,[9] copper(I)-catalyzed cycloaddition with terminal alkynes (CuAAC),[10] or strain-promoted alkyne−azide cycloaddition (SPAAC).[11] In particular, the SPAAC method offers the advantage of circumventing cellular toxicity associated with copper ions while maintaining alkyne reactivity.

Most biorthogonal chemical reporter strategies employ metabolic labeling as the means to incorporate azide-modified sugars into polysaccharides during biosynthesis. This methodology is limited by natural sugar precursors competing with the metabolic chemical reporter. Furthermore, this approach lacks the ability to selectively install chemical reporters to probe a particular subclass of glycans and would not be amenable to identifying sites of *O*-mannosylation. As an alternative approach, chemical reporters are installed directly onto endogenous glycans via recombinant or purified glycosyltransferases. This strategy, termed selective exo-enzymatic labeling (SEEL), offers the additional advantage of solely labeling a

specific class of cell-surface molecules (i.e. N- vs. O-glycans) and provides an attractive methodology to identify sites of *O*-mannosylation.[12,13]

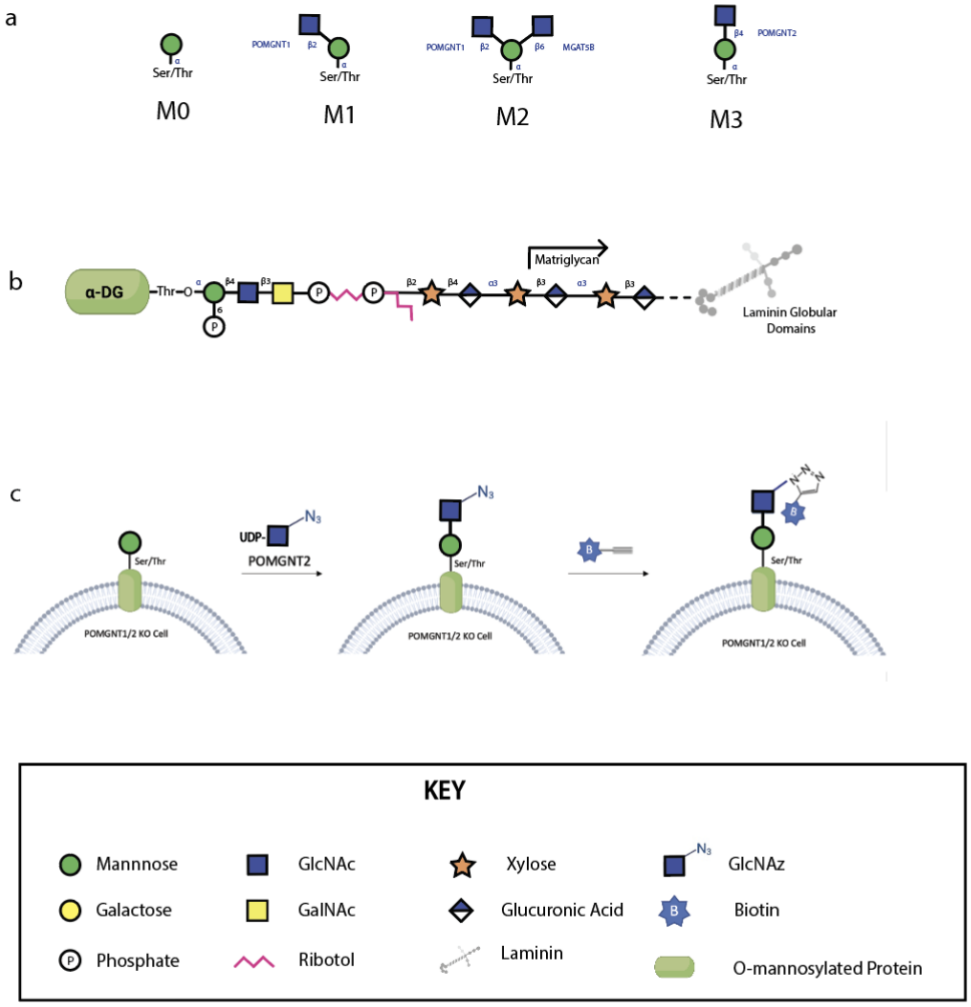


Figure 1. The O-mannosylation pathway and enrichment of O-mannosylated proteins via SEEL.
(a) Four primary core O-mannose glycan structures are involved in the O-mannosylation pathway. Enzymes that catalyze the sugar transfer are indicated in blue. (b) Core M3 structure fully extended by matriglycan. The IIH6 antibody recognizes the disaccharide repeats of matriglycan. (c) SEEL methodolgy for enriching M3 sites (POMGNT2 modified) independently of matriglycan.

Herein, we report enrichment and proteomic identification of *O*-mannosylated protein sites via SEEL with two glycosyltransferases, POMGNT1 and POMGNT2, involved in the *O*-

mannosylation pathway (**Figure 1c**). Azido-modified UDP-GlcNAc (UDP-GlcNAz) was utilized as the chemical reporter. Following SEEL, SPAAC was then employed with a biotin-derivatized alkyne to tag the azido-modified sites of *O*-mannosylation. The biotin-tagged glycoproteins were then enriched via neutravidin pulldown and identified via tandem mass spectrometry.

RESULTS

Kinetic Parameters of POMGNT1 and POMGNT2 with UDP-GlcNAz as Sugar Donor: Glycosyltransferases POMGNT1 and POMGNT2 Tolerate UDP-GlcNAz as a Sugar Donor for a Synthetic α-DG Glycopeptide

Initial efforts to determine the viability of the SEEL protocol applied to enrichment of *O*-mannosylated sites involved both MALDI analysis and enzymatic kinetic assays to study the tolerance of POMGNT1 and POMGNT2 to UDP-GlcNAz. Glycosyltransferase reaction kinetics were performed to test the ability of POMGNT1 and POMGNT2 to extend a synthetic α-DG glycopeptide with UDP-GlcNAz as the sugar donor. Solid-phase peptide synthesis was used to generate synthetic glycopeptide ShortMan 379 (GAIIQT*PTLGPIQPTR). The sequence of this peptide is modified from the known *O*-mannosylated regions of human α-DG, and it has been previously shown to be an acceptor substrate for UDP-GlcNAc.[3] MALDI analysis revealed that glycosyltransferase POMGNT1 tolerates the azido modification of UDP-GlcNAz with ShortMan379 as the acceptor substrate (**Figure 2a**). Enzyme kinetics for POMGNT1 and POMGNT2 with ShortMan379 were investigated by UDP-Glo^{TM} assays. For POMGNT1, the K_M was 5-fold higher with UDP-GlcNAz versus UDP-GlcNAc. For POMGNT2, the K_M was 40

fold higher with UDP-GlcNAz compared to the natural sugar donor UDP-GlcNAc with no appreciable influence on V_{max} (**Figure 2b**).Therefore, POMGNT2 will tolerate the azido-modified biomimetic UDP-GlcNAz, albeit with far less (catalytic) efficiency than UDP-GlcNAc.

Labelling O-mannosylated Sites via Click Chemistry: POMGNT1/2 SEEL Enriches for O-mannosylated Sites

Following confirmation of the tolerance of both POMGNT1 and POMGNT2 for UDP-GlcNAz, the utility of this azido sugar to label *O*-mannosylated sites was tested with POMGNT1/2 knockout cell lysates using click chemistry, namely SPAAC, with a biotin tag (**Figure 3a**). To our delight, tandem mass spectrometry revealed that the built-in positive control, α-DG, was enriched when using both POMGNT1 and POMGNT2 as glycosyltransferases and was absent in the negative control with no glycosyltransferase. However, both western blot analysis and tandem mass spectrometry revealed that click chemistry enrichment with the cell lysate resulted in non-specific binding and led to significant background interference. Although background was an issue, enrichment of α-DG provided confirmation of enrichment of *O*-mannosylated sites via click chemistry using chemical tools UDP-GlcNAz and DIBO biotin alkyne.

Given the undesirable background noise encountered via labelling of the cell lysate, we then inquired if SEEL could be applied to this method of *O*-mannosylation enrichment to label the surface of living cells. POMGNT1/2 knockout HEK293 cells were incubated with UDP-GlcNAz and DIBO in the presence of POMGNT1, POMGNT2, or no enzyme (as a negative control) for 2 h at 37 °C. After washing, the cells were treated with DIBO biotin alkyne, enriched

via magnetic neutravidin beads, and trypsin digested on an S-trap column (**Figure 3b**). The efficiency of the surface-labelling of *O*-mannosylated proteins via SEEL was first inspected by Western blotting using an anti-biotin antibody (**Figure 3c**). This experiment demonstrated enrichment of biotinylated sites with SEEL compared to the negative control with no

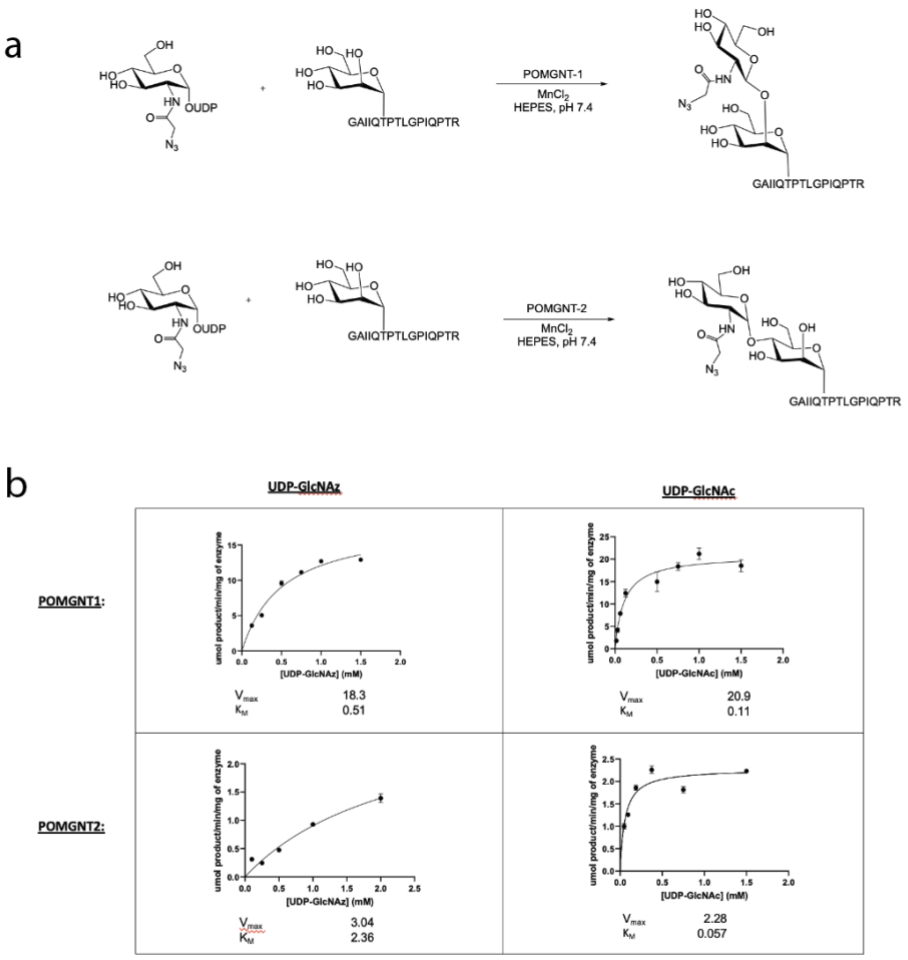


Figure 2. POMGNT1 and POMGNT2 Tolerate the Azido Modification of UDP-GlcNAz.
(a) UDP-GlcNAz extention of a model α-DG glycopeptide with glycosyltransferases POMGNT1 and POMGNT2. (b) POMGNT1 and POMGNT2 enzyme kinetics with ShortMan379 comparing natural sugar donor UDP-GlcNAc to azido modified sugar UDP-GlcNAz measured by UDP-Glo assays. Error bars represent S.E. across three replicates.

glycosyltransferase. Following confirmation of enrichment via SEEL, the biotin enriched samples were analyzed via tandem mass spectrometry.

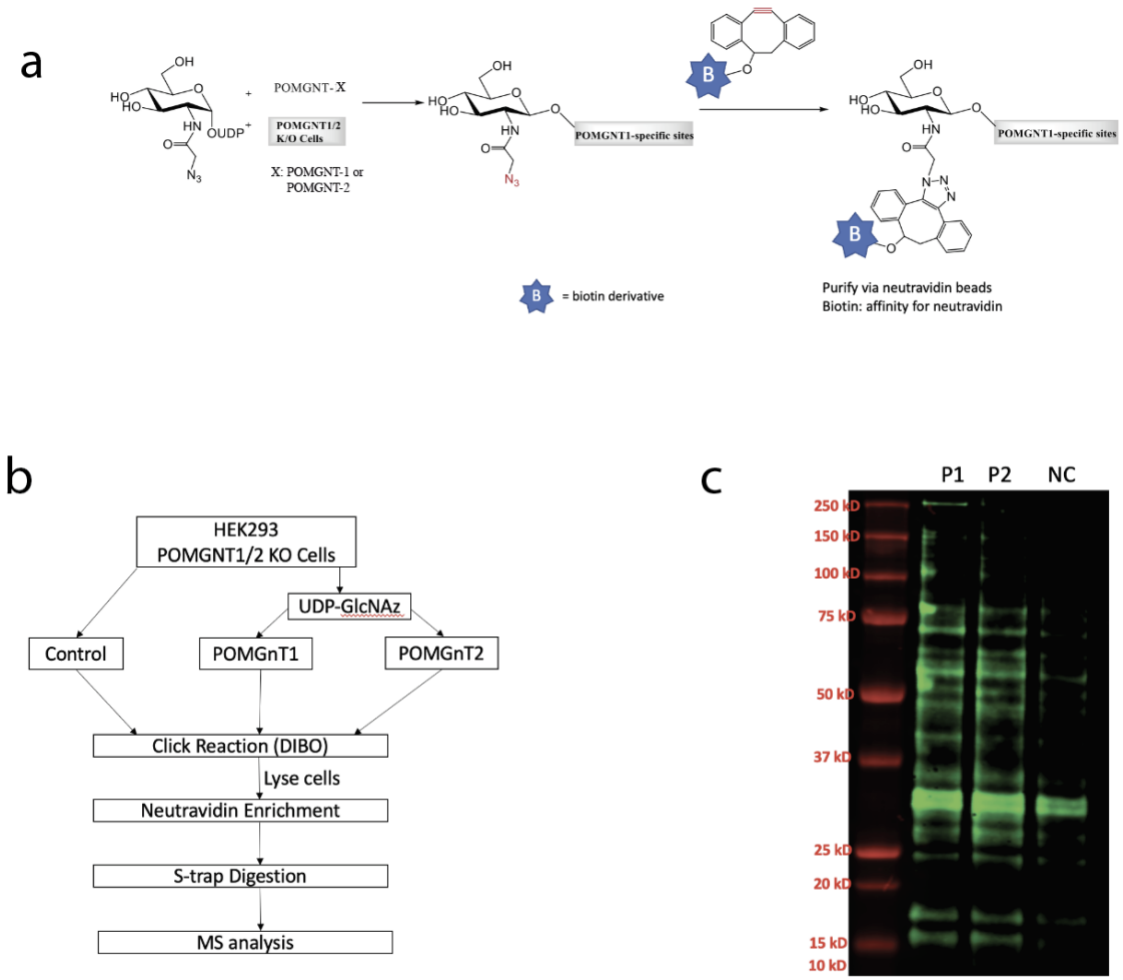


Figure 3. SEEL enriches for POMGNT1/2 O-mannosylated Sites. (a) SPAAC click chemistry strategy to enrich for O-mannosylated sites. (b) SEEL protocol workflow. (c) Anti-biotin western blot analysis of SEEL enriched samples.

Proteomic Analysis of Labeled Glycoproteins Following Neutravidin Enrichment: APMAP and LAMB1 are Enriched Via POMGNT2 SEEL

Proteomic analysis revealed enrichment of the experimentally built-in positive control, α-dystroglyan (DAG1), via cell surface labelling in POMGNT1 SEEL enriched samples across three separate experiments. A known POMGNT1-modified *O*-mannosylated protein, UPF0606 protein KIAA1549, was also enriched in triplicate in the POMGNT1 samples. Enrichment of known *O*-mannosylated proteins DAG1 and KIAA1549 provided evidence of the efficacy of POMGNT1 SEEL. A previously unknown *O*-mannosylated protein, adipocyte plasma membrane-associated protein (APMAP), was enriched in triplicate across both the POMGNT1-modified samples (**Table 1 and Figure 4**).

POMGNT2 SEEL also resulted in enrichment of DAG1 in three independent experiments. The same protein, APMAP, identified via POMGNT1 SEEL was also identified via POMGNT2 SEEL. This suggests that APMAP has both M1 and M3 structures. APMAP was enriched with greater than two times higher abundance (with spectral counts of 1574) than known POMGNT2 modified protein α-DG (with spectral counts of 664). Another previously unidentified *O*-mannosylated protein, Laminin subunit beta-1 (LAMB1), was enriched in triplicate in the POMGNT2 samples, but only enriched in one replicate of the POMGNT1 samples (**Table 1 and Figure 4**). Novelly enriched APMAP and LAMB1 were both assigned at <1% false-discovery rate.

Detection in						UniProt	Protein	Gene	Cell surface protein (Known)	Abundances												# Peptides					
										#Spectra (unnoramlized)						NSAF (normalized)											
P1			P2							P1			P2			P1			P2			P1			P2		
Rep1	Rep2	Rep3	Rep1	Rep2	Rep3					Rep1	Rep2	Rep3	Rep1	Rep2	Rep3	Rep1	Rep2	Rep3	Rep1	Rep2	Rep3	Rep1	Rep2	Rep3	Rep1	Rep2	Rep3
						Q9HDC9	Adipocyte plasma membrane-associated protein	APMAP	Cell surface	53	18	19	54	30	24	1804	1082	1024	1895	1541	1285	6	3	5	6	5	5
						P07942	Laminin subunit beta-1	LAMB1	Cell surface			31	52	20	27			389	425	239	337			4	4	4	4
						Q9HCM3	UPF0606 protein KIAA1549	KIAA1549	Cell surface	18	7	25				131	90	288				4	4	6			
						Q14118	Dystroglycan 1	DAG1	Cell surface	302	115	167	25	33	32	4777	3214	4185	408	788	796	16	13	13	4	8	7

Table 1. Proteomic Analysis of POMGNT1 (P1) and POMGNT2 (P2) SEEL

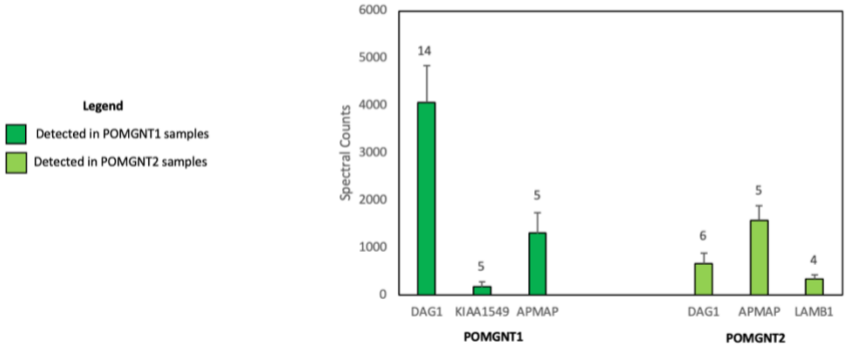


Figure 4. **APMAP and LAMB1 are Enriched via POMGNT2 SEEL.** Average spectral counts of proteins enriched in triplicate via POMGNT1 (dark green) and POMGNT2 (light green) SEEL. Numbers above bars indicate the number of unique peptides for a given protein. Error bars represent SE across three experiments.

Immunoprecipitation assays reveal that alpha-DG co-precipitates with APMAP and LAMB1

Following enrichment of APMAP via POMGNT1 and POMGNT2 SEEL, immunoprecipitation (IP) of APMAP was carried out in brain and liver tissue lysates. Interestingly, Western blot analysis suggested that that α-DG co-precipitates with APMAP (**Figure 5**).

Immunoprecipitation of LAMB1 in brain and liver tissue lysates also resulted in co-precipitation of α-DG.

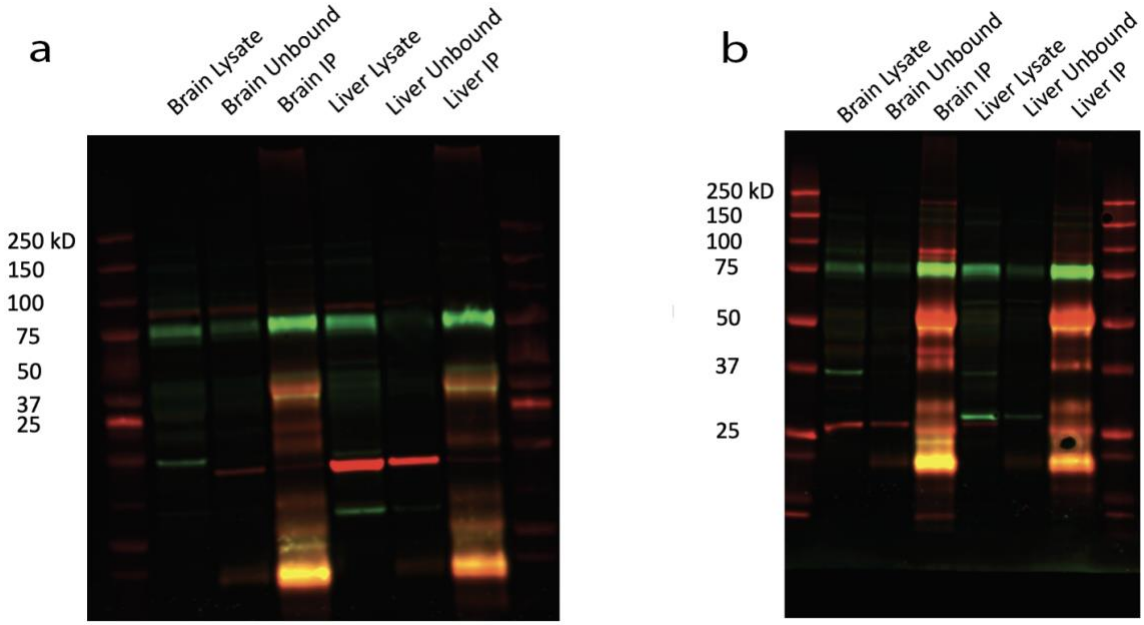


Figure 5. Immunoprecipitation of APMAP and LAMB1. (a) APMAP immunoprecipitation. Red: APMAP (40-50 kD); primary: APMAP 1:200, secondary: donkey anti-rabbit 1: 20,000. Green: α-DG (75 kD); primary: IIH6 hybridoma 1:100, secondary: mouse IgM 1:4000. (b) LAMB1 immunoprecipitation. Red: LAMB1 (200-250 kD); primary: LAMB1 1:500, secondary: donkey anti-rabbit 1:20,000. Green: α-DG (75 kD); primary: IIH6 hybridoma 1:100, secondary: mouse IgM 1:4000. Unbound indicates IP flowthrough.

DISCUSSION

The findings reported herein demonstrate the powerful utility of the SEEL methodology to specifically probe particular subclasses of glycans. We have previously demonstrated the

usefulness of SEEL for enriching N-glycans.[17,18] Here, we show the expansion of SEEL for enriching sites of *O*-mannosylation modified by specific glycosyltransferases, POMGNT1 and POMGNT2. SEEL with glycosyltransferases POMGNT1 and POMGNT2 involved in the *O*-mannosylation pathway enriched previously known *O*-mannosylated sites, including DAG1 (α-DG) for both POMGNT1 and POMGNT2 and KIAA1549 for POMGNT1, as well as novelly enriched APMAP and LAMB1. Cadherins are another class of known *O*-mannosylated proteins, which are M0 structures and are not extended by POMGNT1 or POMGNT2. Therefore, we did not expect nor did we observe enrichment of cadherins.

Following enrichment of APMAP with POMGNT1/POMGNT2 SEEL methodology, immunoprecipitation of APMAP suggested that APMAP and α-DG are closely associated based on the observation of co-precipitation of α-DG along with APMAP. Furthermore, knockdown of APMAP has been shown to affect expression of cadherins, which are M0 *O*-mannosylated proteins.[14] Although initial studies identified the primary function of APMAP as adipocyte differentiation, APMAP is also involved in tumor metastasis.[14,15] Additionally, APMAP is a negative regulator of amyloid-β production in the brain.[16] Given these functions of APMAP, further investigation of the role of APMAP in the *O*-mannosylation pathway and its interaction with α-DG are underway. Future analysis will also involve immunoprecipitation of IGF-R, which is not *O*-mannosylated or associated with α-DG, as a negative control to further support this observation of co-precipitation of APMAP.

Upon initial consideration, one might hypothesize that LAMB1 was potentially enriched along with DAG1 as a nonspecific interacting partner since matriglycan connects DAG1 to the laminin globular domains. However, no matriglycan would be present in the POMGNT1/2

knockout cell line used for the SEEL enrichments since POMGNT2 is required for this extension. Therefore, we feel confident that enrichment of LAMB1 is a direct result of SEEL and is of interest for further studies. The enrichment of LAMB1 gives rise to two possibilities: 1) LAMB1 is *O*-mannosylated, or 2) LAMB1 interacts with α-DG independently of matriglycan. To further investigate this interaction, carrying out immunoprecipitation of LAMB1 in a POMGNT2 knockout cell-line would be of interest. Immunoprecipitation of LAMB1 co-precipitated α-DG in both liver and brain tissue, which is expected given the presence of matriglycan in these tissues. Interrogating this interaction in the knockout cell-line will provide valuable insight into the mechanism of this interaction and could expand our current understanding.

Comparing phenotypic similarities across loss of gene function, both APMAP and DAG1 dysfunctionality lead to retinopathy. Additionally, loss of function of DAG1 and LAMB1 both lead to lissencephaly. Hypoglycosylation of α-DG leads to CMDs, and APMAP knockout mice have been shown to have reduced grip strength. These phenotypic overlaps offer valuable insights for further investigation of the mechanism of hypo-*O*-mannosylation in regards to muscular dystrophy, ocular defects, and neurodevelopmental abnormalities of the associated diseases.

We envision future studies employing POMGNT2 SEEL with neuronal cell lines since a limitation of the current study is potentially overlooking of POMGNT2 extended proteins that are not expressed HEK293 cell lines but could be present in neuronal cell lines. Mealer and colleagues recently reported that POMGNT2 is highly expressed in the cerebellum of the brain compared to other regions of the brain and other tissues.[17] To that end, future studies of the

analysis *O*-man glycans in the cerebellum would be an exciting avenue of pursuit that could potentially shed light on the role of *O*-mannosylation in the neurodevelopmental abnormalities observed among CMD patients. Continued study and development of methods toward enrichment of *O*-mannosylated glycoprotein sites have the potential to identify novel targets, such as APMAP described here, for further investigation that could uncover new biomarkers for diagnosis and accelerate development of therapeutics for CMDs.

METHODS

Cell Lines and Cell Culture

HEK293T cells were cultured in DMEM medium supplemented with 10% fetal bovine serum (FBS, BenchMark) and penicillin (100 IU/mL)/streptomycin (100 μg/mL, MediaTech) in a 5% CO2 atmosphere, 37 °C humid incubator.

Two-step SEEL

Two-step SEEL was done with HEK293 POMGNT1/2 KO cells in 1.5 mL Eppendorf tubes. Cells were lifted with 1X PBS/ 10 mM ETDA (2 mL). After cells were washed with PBS three times, cells were incubated in SEEL reaction mixture in 37 °C for 2 h. For the POMGNT2 samples, the SEEL reaction mixture (300 μL) was prepared using serum free DMEM with POMGNT2 (100 ng), UDP-GlcNAz (7 mM), 2 μL of BSA (2 mg/mL), alkaline phosphatase (2 μL), and 5mM $MnCl_2$. For the POMGNT1 samples, the SEEL reaction mixture (300 μL) was prepared using serum free DMEM with POMGNT1 (50 ng), UDP-GlcNAz (2.5 mM), 2 μL of

BSA (2 mg/mL), alkaline phosphatase (2 μL), and 5mM MnCl2. For the negative control, the reaction conditions were kept constant, but DMEM was added in place of the enzyme and sugar. Every 30 minutes, the Eppendorf tubes were gently inverted and opened for 5 seconds. After the SEEL reaction, cells were washed with PBS three times. Then the cells were further labeled with DIBO-Biotin (300 μM) in 2% FBS containing PBS for 30 min at room temperature (including the negative control). The cells were washed three times with PBS and placed at -80 °C overnight. The cells were lysed with 1 mL of RIPA buffer with 1X protease inhibitor cocktail. Sodium chloride was added to the RIPA lysates to increase the concentration of salt to 250 mM sodium chloride. The lysate was centrifuged at 20,000 x g for 15 minutes at 4 °C, and the supernatant was collected and bound to magnetic neutravidin beads (Sera-Mag SpeedBeads) by end-over-end rotation overnight at 4 °C.

Biotinylated Protein Pulldown

The following day, the neutravidin beads were collected on a magnetic stand, and the supernatant was saved as flow through. The neutravidin beads were washed with high salt RIPA buffer (250 mM NaCl) two times, RIPA buffer two times, followed by one wash with 1X TBS, three washes with high salt (250 mM) TBS, and finally two washes with TBS. The enriched biotinylated proteins were then eluted from the neutravidin with 5% SDS in 50 mM TEAB for 10 minutes at 95 °C. The eluted proteins were then trypsin digested on an S-trap column (Protifi) to simultaneously remove detergent from the samples, according to the manufactures protocol. The resulting peptides were then dried down using speed vacuum.

Proteomic analysis

"The peptides were separated on a 75 μm (i.d.) × 15 cm C18 capillary column (packed in house, YMC GEL ODS-AQ120 ÅS-5, Waters) and eluted into the nano-electrospray ion source of an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific) with a 180 min linear gradient consisting of 0.5–100% solvent B over 150 min at a flow rate of 200 nL/min. The spray voltage was set to 2.2 kV, and the temperature of the heated capillary was set to 280 °C. Full MS scans were acquired from *m*/*z* 300 to 2000 at 120k resolution, and MS2 scans following collision-induced fragmentation were collected in the ion trap for the most intense ions in the Top-Speed mode within a 3 s cycle using Fusion instrument software (v1.1, Thermo Fisher Scientific). The raw spectra were searched against the human protein database (UniProt, Oct. 2014) using SEQUEST (Proteome Discoverer 1.4, Thermo Fisher Scientific) with full MS peptide tolerance of 20 ppm and MS2 peptide fragment tolerance of 0.5 Da, and filtered using ProteoIQ (v2.7, Premier Biosoft) at the protein level to generate a 1% false-discovery rate for protein assignments. Proteins detected at 1% false-discovery rate in the negative controls (no recombinant enzyme added, data not shown) were excluded from the final lists of proteins identified in respective conditions. UniProt was used to define cellular localization. Quantification was performed by normalizing the spectral counts generated by ProteoIQ (v2.7, Premier Biosoft)."[12]

UDP-Glo Glycosyltransferase Assays

UDP-Glo glycosyltransferase assays (Promega) were performed using 10 mM HEPES (pH 7.3), 5 mm$MnCl_2$, 10 μM ShortMan279 glycopeptide acceptor substrate, 40 ng/μL of POMGNT1 or

80 ng/μL of POMGNT2, and varying amounts of UDP-GlcNAz and UDP-GlcNAc at 37 °C for 2 h in a white, flat bottom, 384-well plate. Following completion of the glycosyltransferase reaction, UDP detection reagent (equal volume) was added to convert UDP to ATP and subsequently generate light by a luciferase reaction. A GloMax-Multi + luminometer (Promega) was used to detect luminescence, which was correlated to UDP concentration by using a UDP standard curve. The data was fit to the Michaelis-Menten equation using the non-linear regression fit in GraphPad Prism version 7.1, and the data presented represent the average of three independent experiments.

REFERENCES

(1) Stalnaker, S. H.; Stuart, R.; Wells, L. Mammalian O-Mannosylation: Unsolved Questions of Structure/Function. *Curr. Opin. Struct. Biol.* **2011**, *21* (5), 603–609. https://doi.org/10.1016/j.sbi.2011.09.001.

(2) Clements, R.; Turk, R.; Campbell, K. P.; Wright, K. M. Dystroglycan Maintains Inner Limiting Membrane Integrity to Coordinate Retinal Development. *J. Neurosci. Off. J. Soc. Neurosci.* **2017**, *37* (35), 8559–8574. https://doi.org/10.1523/JNEUROSCI.0946-17.2017.

(3) Halmo, S. M.; Singh, D.; Patel, S.; Wang, S.; Edlin, M.; Boons, G.-J.; Moremen, K. W.; Live, D.; Wells, L. Protein O-Linked Mannose β-1,4-N-Acetylglucosaminyl-Transferase 2 (POMGNT2) Is a Gatekeeper Enzyme for Functional Glycosylation of α-Dystroglycan. *J. Biol. Chem.* **2017**, *292* (6), 2101–2109. https://doi.org/10.1074/jbc.M116.764712.

(4) Sheikh, M. O.; Halmo, S. M.; Wells, L. Recent Advancements in Understanding Mammalian O-Mannosylation. *Glycobiology* **2017**, *27* (9), 806–819. https://doi.org/10.1093/glycob/cwx062.

(5) Wright, K. M.; Lyon, K. A.; Leung, H.; Leahy, D. J.; Ma, L.; Ginty, D. D. Dystroglycan Organizes Axon Guidance Cue Localization and Axonal Pathfinding. *Neuron* **2012**, *76* (5), 931–944. https://doi.org/10.1016/j.neuron.2012.10.009.

(6) Satz, J. S.; Ostendorf, A. P.; Hou, S.; Turner, A.; Kusano, H.; Lee, J. C.; Turk, R.; Nguyen, H.; Ross-Barta, S. E.; Westra, S.; Hoshi, T.; Moore, S. A.; Campbell, K. P. Distinct Functions of Glial and Neuronal Dystroglycan in the Developing and Adult Mouse Brain. *J. Neurosci.* **2010**, *30* (43), 14560–14572. https://doi.org/10.1523/JNEUROSCI.3247-10.2010.

(7) Stalnaker, S. H.; Aoki, K.; Lim, J.-M.; Porterfield, M.; Liu, M.; Satz, J. S.; Buskirk, S.; Xiong, Y.; Zhang, P.; Campbell, K. P.; Hu, H.; Live, D.; Tiemeyer, M.; Wells, L. Glycomic Analyses of Mouse Models of Congenital Muscular Dystrophy. *J. Biol. Chem.* **2011**, *286* (24), 21180–21190. https://doi.org/10.1074/jbc.M110.203281.

(8) Inamori, K.-I.; Beedle, A. M.; de Bernabé, D. B.-V.; Wright, M. E.; Campbell, K. P. LARGE2-Dependent Glycosylation Confers Laminin-Binding Ability on Proteoglycans. *Glycobiology* **2016**, *26* (12), 1284–1296. https://doi.org/10.1093/glycob/cww075.

(9) Köhn, M.; Breinbauer, R. The Staudinger Ligation-a Gift to Chemical Biology. *Angew. Chem. Int. Ed Engl.* **2004**, *43* (24), 3106–3116. https://doi.org/10.1002/anie.200401744.

(10) Meldal, M. Polymer "Clicking" by CuAAC Reactions. *Macromol. Rapid Commun.* **2008**, *29* (12–13), 1016–1051. https://doi.org/10.1002/marc.200800159.

(11) Mbua, N. E.; Guo, J.; Wolfert, M. A.; Steet, R.; Boons, G.-J. Strain-Promoted Alkyne-Azide Cycloadditions (SPAAC) Reveal New Features of Glycoconjugate Biosynthesis. *Chembiochem Eur. J. Chem. Biol.* **2011**, *12* (12), 1912–1921. https://doi.org/10.1002/cbic.201100117.

(12) Sun, T.; Yu, S.-H.; Zhao, P.; Meng, L.; Moremen, K. W.; Wells, L.; Steet, R.; Boons, G.-J. One-Step Selective Exoenzymatic Labeling (SEEL) Strategy for the Biotinylation and Identification of Glycoproteins of Living Cells. *J. Am. Chem. Soc.* **2016**, *138* (36), 11575–11582. https://doi.org/10.1021/jacs.6b04049.

(13) Capicciotti, C. J.; Zong, C.; Sheikh, M. O.; Sun, T.; Wells, L.; Boons, G.-J. Cell-Surface Glyco-Engineering by Exogenous Enzymatic Transfer Using a Bifunctional CMP-Neu5Ac Derivative. *J. Am. Chem. Soc.* **2017**, *139* (38), 13342–13348. https://doi.org/10.1021/jacs.7b05358.

(14) Zhu, X.; Xiang, Z.; Zou, L.; Chen, X.; Peng, X.; Xu, D. APMAP Promotes Epithelial-Mesenchymal Transition and Metastasis of Cervical Cancer Cells by Activating the Wnt/β-Catenin Pathway. *J. Cancer* **2021**, *12* (20), 6265–6273. https://doi.org/10.7150/jca.59595.

(15) Albrektsen, T.; Richter, H. E.; Clausen, J. T.; Fleckner, J. Identification of a Novel Integral Plasma Membrane Protein Induced during Adipocyte Differentiation. *Biochem. J.* **2001**, *359* (Pt 2), 393–402.

(16) *The adipocyte differentiation protein APMAP is an endogenous suppressor of Aβ production in the brain - PubMed*. https://pubmed.ncbi.nlm.nih.gov/25180020/ (accessed 2023-01-25).

(17) Williams, S. E.; Noel, M.; Lehoux, S.; Cetinbas, M.; Xavier, R. J.; Sadreyev, R. I.; Scolnick, E. M.; Smoller, J. W.; Cummings, R. D.; Mealer, R. G. Mammalian Brain Glycoproteins Exhibit Diminished Glycan Complexity Compared to Other Tissues. *Nat. Commun.* **2022**, *13* (1), 275. https://doi.org/10.1038/s41467-021-27781-9.

CHAPTER 3

SYNTHESIS OF A C-LINKED GLYCOSIDE FOR THE STUDY OF O-MANNOSYLATION

[1] Carter, A. and Boons, G.J. To be submitted to *Journal of Biological Chemistry*.

ABSTRACT

Defective *O*-mannosylation of α-dystroglycan (α-DG), the most well studied *O*-mannosylated mammalian protein, leads to congenital muscular dystrophies and neurological defects. The *O*-mannose sites can be elaborated to core M1, M2, and M3 structures. Glycosaminyl transferase POMGNT1 adds *N*-acetylglucosamine (GlcNAc) in a β-1,2 linkage to make the M1 core. Although core M1 glycans are the most abundant in the mucin-like domain of α-DG, the function of these glycans remains unknown. In contrast to the M1 core glycan, M3 is generated very selectively by Protein *O*-linked mannose β-1,4-*N*-acetylglucosaminyltransferase 2 (POMGNT2) adding GlcNAc in a β-1,4 linkage. Thus, POMGNT2 is considered the "branching point" for *O*-mannosylation pathways. Interactions between α-DG and its extracellular matrix (ECM) ligands require α-DG to be extended by *O*-mannose structures based on the M3 core glycan structure. The core M3 can be extended with matriglycan, which is a repeating disaccharide that binds laminin globular domain containing proteins in the ECM. There is currently no method to identify the M3 core without the presence of matriglycan. Therefore, it remains a possibility that unextended M3 structures exist, but there is currently no tool to identify core M3 glycans without the presence of the repeating disaccharide. Development of antibodies to the core *O*-mannose glycans would help to identify these core glycans and ultimately better understand their function. Generating antibodies against this epitope presents a challenge since the oxygen in the glycosidic linkage is vulnerable to enzymatic degradation. However, carbon-linked (*C*-linked) glycosides, in which carbon replaces the normal oxygen in the glycosidic linkage, are resistant to enzymatic degradation and provide robust immunogens. In efforts to overcome stability issues of *O*-linked glycoside immunogens, we synthesized a novel

C-linked glycoside mimic (*C*-Man-Thr) of *O*-mannose-Threonine that has the potential to be used as an antigen to generate antibodies to identify these core *O*-mannose glycans. Extension of the *C*-linked glycoside into *C*-linked glycopeptides to generate antibodies has the potential to provide improved reagents to generate a more comprehensive profile of sites that are modified with *O*-mannose glycans.

INTRODUCTION

C-glycosides as Metabolically Stable Therapeutic Agents

Posttranslational protein glycosylation plays a critical role in several biological processes, including protein-trafficking, intra-cellular recognition, cell-adhesion, and immunogenicity.[1] Carbohydrates that are linked to the peptide via the serine or threonine side-chain oxygen are known as *O*-glycans, and these glycoconjugates have been widely explored for their therapeutic potential. However, utility of these glycoconjugates is limited by their instability to acidic conditions and degradation by glycosidase enzymes. In efforts to overcome these stability issues, synthesis of carbon-linked glycosides (*C*-glycosides) as mimics of the native *O*-glycosides has been employed to generate therapeutic agents. One example of therapeutic *C*-glycosides includes SGLT2 inhibitors against type II diabetes, such as dapaglifozin, canagliflozin, and empaglioflazin.[2–7] Another successful example of a *C*-glycoside therapeutic agent includes Pro-Xylane™ as a skincare ingredient to improve skin barrier function by boosting collagen production via eliciting of an increase in glycosaminoglycan synthesis in fibroblasts.[7–10] Some *C*-glycoside analogues have even shown enhanced biological activity compared to their natural *O*-glycoside counterparts. For example, KRN7000, the *C*-glycoside

mimetic of the immunosilent α-galactosyl ceramide, has been shown to be 100x more effective at preventing melanoma spread compared to the *O*-glycoside.[11]

Generating Antibodies to Core *O*-mannose Glycans

Among *O*-glycans, *O*-mannose structures play an essential role in development of both the musculature and nervous system. The α-linkage to *O*-mannose Ser/Thr forms the infrastructure for the four core glycans of *O*-mannosylated proteins, and this glycan structure is referred to as core M0. The M0 structures can be elaborated by other enzymes in the *O*-mannosylation pathway to form M1, M2, and M3 core glycan structures. Protein *O*-linked mannose β-1,4-*N*-acetylglucosaminyltransferase 1 (POMGNT1) adds *N*-acetylglucosamine (GlcNAc) in a β-1,2 linkage to create the M1 core. In contrast to the M1 core glycan, M3 is generated very selectively by Protein *O*-linked mannose β-1,4-*N*-acetylglucosaminyltransferase 2 (POMGNT2) adding GlcNAc in a β-1,4 linkage. POMGNT2 is considered the "branching point" for *O*-mannosylation pathways. The M3 structures formed by POMGNT2 allow the protein α-dystroglycan (α-DG) to interact with ligands in the extracellular matrix (ECM) via extension with matriglycan, which is a repeating disaccharide that binds laminin globular domain containing proteins in the ECM. No method currently exists to identify the M3 core without the presence of matriglycan. Therefore, the existence of unextended M3 structures remains a possibility, but no tool is currently available to identify core M3 glycans without the presence of the matriyglycan.

In addition to α-DG, an increasing number of protein targets for *O*-mannosylation in the brain and additional biosynthetic pathways have been identified.[12] For example, *O*-

mannosylation of cadherins was recently discovered.[13] Cadherins and plexins contain non-extended M0 structures, and these proteins are *O*-mannosylated by an unknown family of novel protein *O*-mannosyltransferases.[12] Cadherins are cell adhesion molecules which are critical for neuronal circuit assembly, because they mediate neuronal self-recognition so that there is no overlap of isoneural branches.[14] *O*-mannosylation of cadherin is distinguished from that of α-DG in that the glycans of cadherin are not extended.[12] Therefore, the glycosidic linkage of the single sugar residue of these glycans are potentially subject to degradation by α-mannosidase. There is a need for the development of improved reagents to generate a more comprehensive profile of sites that are modified with *O*-Man glycans. Making antibodies to the α-*O*-Man modification would therefore be valuable for future *in vitro* and immunohistochemistry studies that could shed light on the role of this modification. These could also be used for cases where the *O*-Man is normally extended, but may be truncated by cells as well as brain tissue deficient in POMGNT1.[13,15] This will require synthesis of relevant *O*-Man glycopeptide immunogens. Generating antibodies against this epitope presents a challenge since the native α-*O*-Man bond in the unextended glycopeptides that would be used to raise such antibodies is efficiently degraded by α-mannosidase. However, the *C*-glycoside form, where carbon replaces the normal oxygen in the glycosidic linkage, would be resistant to enzymatic degradation and provide a more robust immunogen.

Given the value of generating more stable antigens to both unextended and extended *O*-Man glycans, we set out to synthesize the *C*-glycoside form of *O*-Man threonine (*C*-Man-Thr) (**2**) (**Figure 3.1**). The long-term goal is to use *C*-Man-Thr as a building block in peptide synthesis to create stable antigens to core *O*-manose glycans. Threonine is of particular interest because all

of the *O*-linked glycosylation sites for cadherin are on Threonine.[16] M1, M2, and M3 cores of α-DG are all on serine or threonine sites. Furthermore, all of the M3 cores identified on α-DG are on threonine sites.

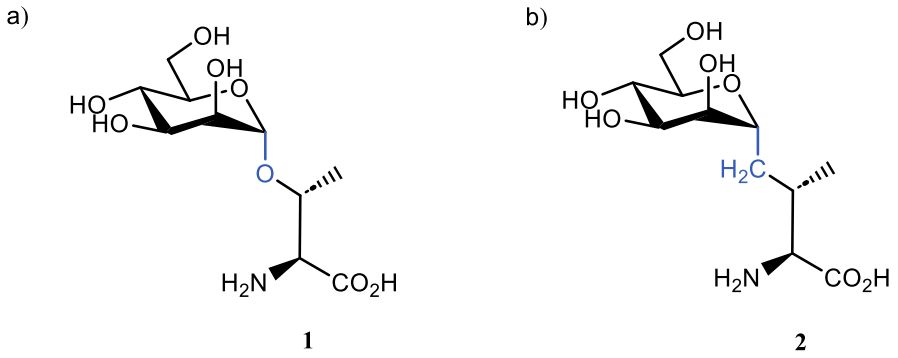


Figure 3.1. *C*-glycoside synthetic target compound. a) natural *O*-mannosyl threonine b) *C*-mannosyl-threonine synthetic target

Synthetic Challenges in the Synthesis of *C*-glycosidic Threonine Derivatives

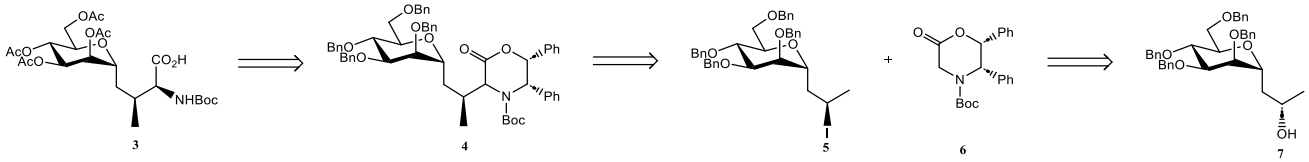
Synthesis of *C*-glycoside threonine derivatives is inherently challenging due to three points stereochemistry that must be addressed. β-branched α-amino acids, namely threonine, have two points of stereochemical relevance. The sugar adduct requires an additional point of stereochemical consideration, in this case alpha. Only one successful synthesis of a Threonine *C*-glycoside has been reported.[17] This synthesis involved 14 steps with 12% overall yield to give Galactosyl-*C*-Threonine. Another report involved synthesis of a Threonine *C*-glycoside, which gave the undesired stereo-configuration as the final product.[18] Mannose-*C*-Serine has been synthesized by Nolen's group.[19] However, the synthesis of Mannose-*C*-Threonine derivative has not been previously reported.

Given the nature of β-branched α-amino acids, synthesis of such a derivative requires creation of a C(sp^3)-C(sp^3) bond. In particular, C(sp^3)-C(sp^3) bond formation has been a challenging pursuit among organic chemists for the last several decades.[20] Cross coupling methodologies have predominately focused on primary alkyl halides. When subjected to cross-coupling reaction conditions, unactivated secondary and tertiary alkyl halides often undergo unwanted side reactions due to kinetic favorability of the eliminated product.[21] Furthermore, successful couplings involving secondary and tertiary alkyl halides often require the presence of aromatic activating groups that would not be amenable to creating *C*-glycoside derivatives.[22]

RESULTS

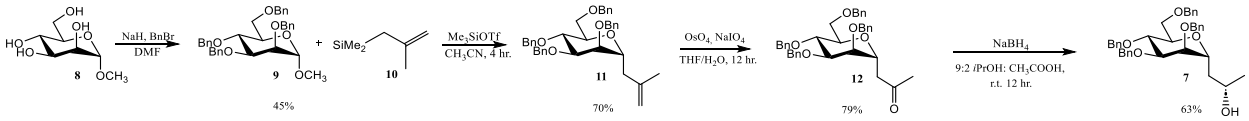
Stereoselective Synthesis of *C*-glycoside Alcohol Intermediate

Inspired by Nolen's *C*-glycosyl serine synthesis, the *C*-glycosyl threonine derivative was planned to be synthesized exploiting the same alkylation strategy with oxazinone **6** as a chiral glycine enolate equivalent, but instead reacting **6** with secondary alkyl halide **5**, as opposed to the primary alkyl halide in Nolen's synthesis, to achieve the *C*-Man-threonine derivative **3**.[19] Retrosynthetic analysis of *C*-Man-Thr revealed *C*-glycosyl secondary alcohol **7** as a key intermediate (**Scheme 3.1**).



Scheme 3.1. Retrosynthetic analysis of *C*-Man-Thr

To synthesize the alcohol derivative, the hydroxyl groups of methyl α-D-mannopyranoside were first protected as benzyl ethers. The α-*C*-glycoside linkage was established by reaction of compound **9** with allyl silane **10** to give **11**.[23] Alkyne derivative **11** was then oxidized to ketone derivative **12**. The stereochemistry of the β-carbon of threonine was synthetically introduced by stereoselective reduction of the ketone to the alcohol derivative. Employing acetic acid as a catalyst, the ketone was stereoselectively reduced to S-alcohol **7** (dr: 9:1) (**Scheme 3.2**).[9] The *C*-glycosyl alcohol skincare ingredient Pro-Xylane™ was synthesized using these reduction reaction conditions to give high diastereoselectivity. A proposed transition state involves coordination of boron to the hydroxyl of C_2 of the sugar and the carbonyl of the aglycon moiety, forcing the hydride attack from the opposite face (**Figure 3.2a**).[9] Another possible transition state is based on a Meerwein-Porndorf type reduction to afford the S-alcohol.[9] NMR analysis in conjugation with rotamer analysis suggested that the major isomer formed had the same configuration as the natural Threonine derivative.



Scheme 3.2. Synthesis of key *C*-glycosyl alcohol intermediate **7**

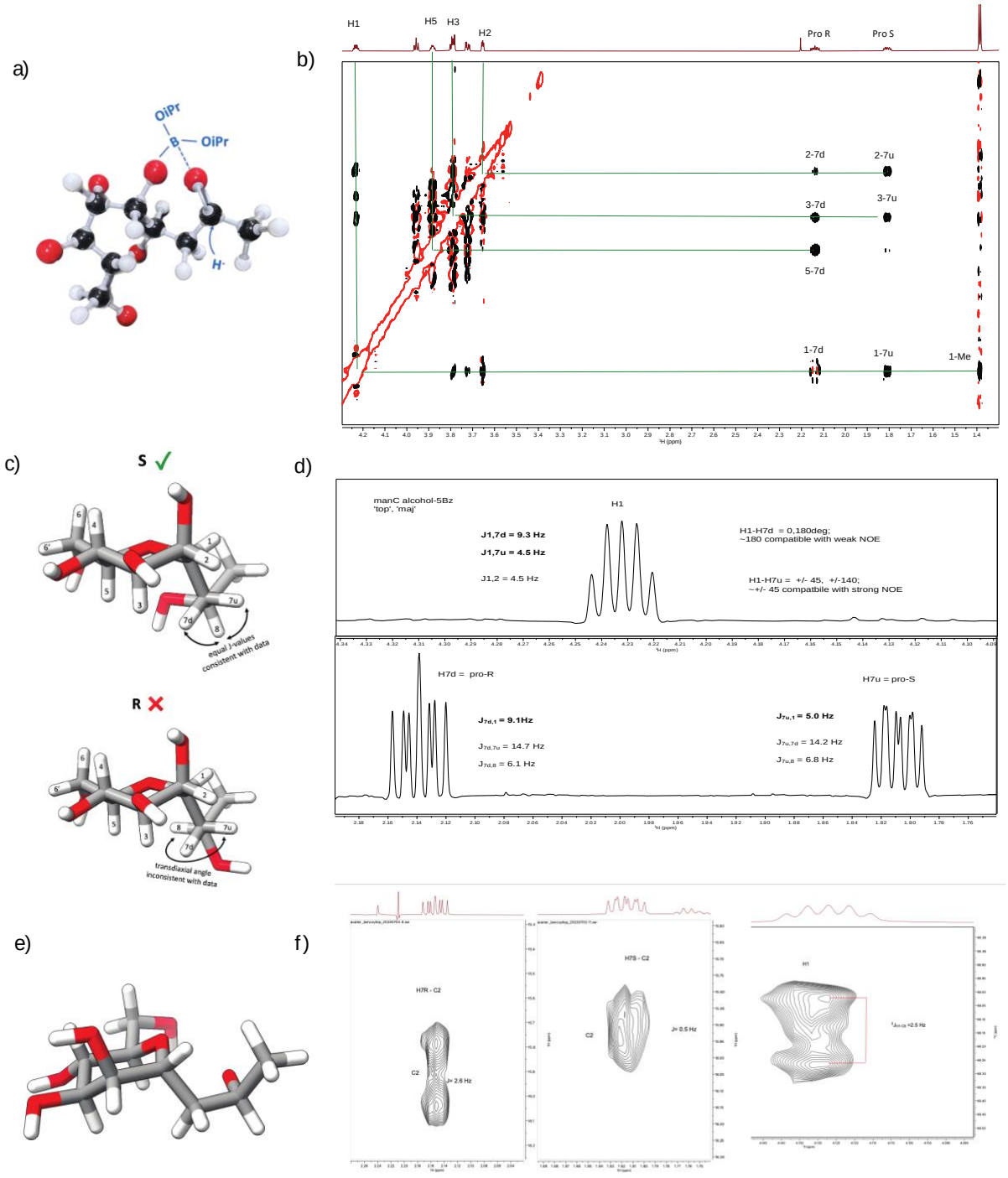


Figure 3.2. Stereochemical Analysis of *C*-Man-Alcohol Derivative **7**. a) Proposed transition state of stereoselective reduction with. b) NOE data of major benzoylated ester of alcohol isomer **7**. c)

(180°, 60°) rotamer of S and R derivative, supporting S-isomer. d) J-coupling of diasterotopic protons. e) conformation of alcohol **7** supported by NMR data, f) J_{CH} couplings between H_1-C_8 and H_{7s}-C_2 and H_{7d}-C_2 consistent with approximately 60° bond angles of the S-isomer.

To establish this stereochemistry, proton-proton, proton-carbon, and NOE data were analyzed of the acetylated ester and benzoylated ester of both the major and minor alcohol derivatives (**Figure 3.2**). First, the assignments for the diastereotopic protons of the methylene carbon were established for the benzoylated major alcohol product, with two distinct chemical shifts. The combination of proton-proton NOEs and J-coupling values indicates the Pro-S proton (H_{7u}) is shifted upfield, while the Pro-R proton (H_{7u}) was shifted downfield (**Figure 3.2 b-d**). The large H_1-H_{7d} J coupling of 9.3 Hz suggests trans-diaxial angle of approximately 180°, compared to the small H_1- H_{7u} J coupling of 4.5 Hz of approximately 60° (**Figure 3.2d**). These coupling values are also consistent with a small H_1-H_{7u} NOE and strong H_1- H_{7d} NOE (**Figure 3.2b**). The assignments for the diastereotopic protons were then used to further identify the stereochemistry around the alcohol derivative (**Figure 3.2)**.

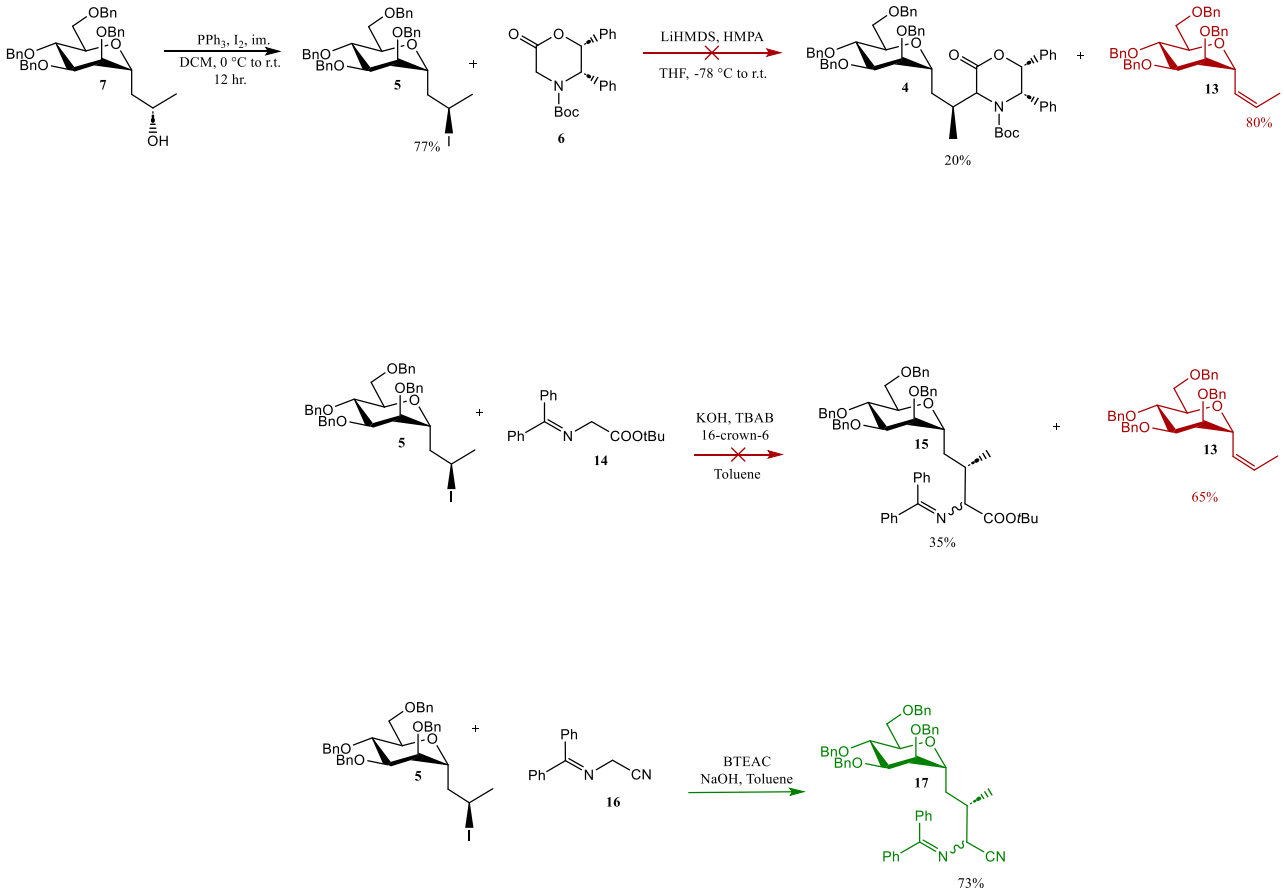
To further aid stereochemical analysis of the *C*-glycosyl alcohol **7**, all nine possible rotamers of both isomers of alcohol **7** were built using Chimera X, with the angles defined as (C_2, C_1, $C_{methylene}$), ($C_{methylene}$, $C_{alcohol}$, C_{methyl}). Comparing both the NOE data and J-values, only one rotamer of the S alcohol (180°, 60°) is consistent with the NOE and J-coupling data of the major alcohol isomer (**Figure 3.2c**). None of the rotamers for the R alcohol are consistent with the major isomer. Additionally, only one rotamer is consistent with the R alcohol for the minor alcohol product (**Supplemental Table 3.2**).

The S (180°,60°) rotamer is consistent with the nearly equal J-coupling values for the diastereotopic protons to the C8 proton (**Figure 3.2c)**. Specifically, the approximately equal H_{7d} to $H_{8\_alcohol}$ and H_{7u} to $H_{8\_alcohol}$ J-coupling values suggests equal distance between H_{7d} to $H_{8\_alcohol}$ and H_{7u} to $H_{8\_alcohol}$ (**Figure 3.2c**). The R (180°,60°) rotamer, however, is not consistent with the equal J-coupling values for the diastereotopic protons (H_{7u} and H_{7d}) to the $H_{8\_alcohol}$, considering that the transdiaxial angle would give a large J-coupling value from H_{7u} to $H_{8\_alcohol}$ and a small J-coupling for the smaller bond angle between H_{7d} and $H_{8\_alcohol}$ (**Figure 3.2c**). Furthermore, the S (180°,60°) rotamer also coincides with the H_1 to diastereomeric proton J-coupling data (H_{7u} and H_{7d}) discussed previously (**Figure 3.2d**). The NOE, J-coupling, and rotamer analysis combined suggests that the major product is the desired S-isomer, as expected from the proposed transition state (**Figure 3.2a,e**). Additionally, the small J_{CH} couplings between H_1-C_8 and H_{7s}-C_2 and H_{7d}-C_2 are consistent with approximately 60° bond angles of the S (180°,60°) rotamer (**Figure 3.2f**). The same conclusions were drawn from the acetylated ester alcohol derivative data.

C(sp^3)-C(Sp^3) Crossing Coupling to Create *C*-glycosyl Threonine Intermediate

With the desired stereoisomer of alcohol **7** in hand, this intermediate was then transformed to secondary alkyl iodide derivative **5** in order to perform the alkylation to create the C(sp^3)-C(Sp^3) bond of β-branched α-amino acid *C*-glycosyl intermediate **4**. The original synthetic plan involved reacting iodide **5** with oxazolidinone **6** to yield β-branched α-amino acid *C*-glycosyl intermediate **4,** based on the synthesis of *C*-Man-Ser reported by Nolan.[19] However, this enolate reaction proceed in poor yields with the di-substituted iodide compound **5**. Instead,

the reaction gave the undesired eliminated product **13** as the major product, as commonly observed among secondary and tertiary alkyl halide in cross coupling reactions. Fortunately, via an alternative route using the O'Donnell amino acid synthetic route, iodide **5** was reacted with benzophenone imine of acetonitrile **16** to give O'Donnell product **17** in 73% yield.[24] A 50:50 mixture of isomers at the α-carbon resulted from the alkylation, and the isomers were successfully separated via column chromatography (**Scheme 3.3, Figure 3.3**). Of note, the reaction of the iodide **5** with benzophenone imine of *tert*-butyl ester also gave eliminated product **13** as the major product (**Scheme 3.3**).



Scheme 3.3. C(sp^3)-C(Sp^3) alkylation to create β-branched α-amino acid *C*-glycosyl intermediate

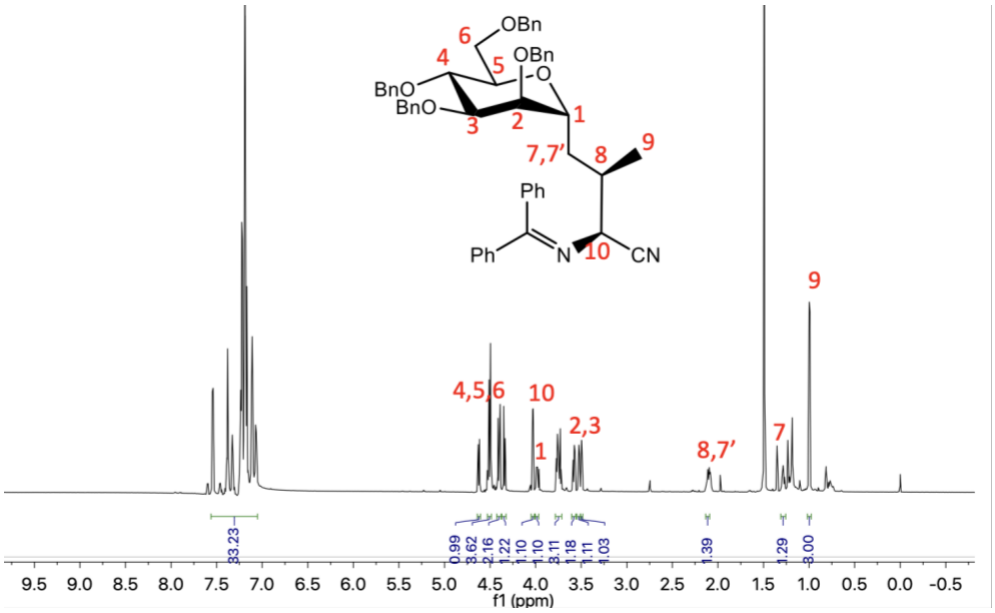
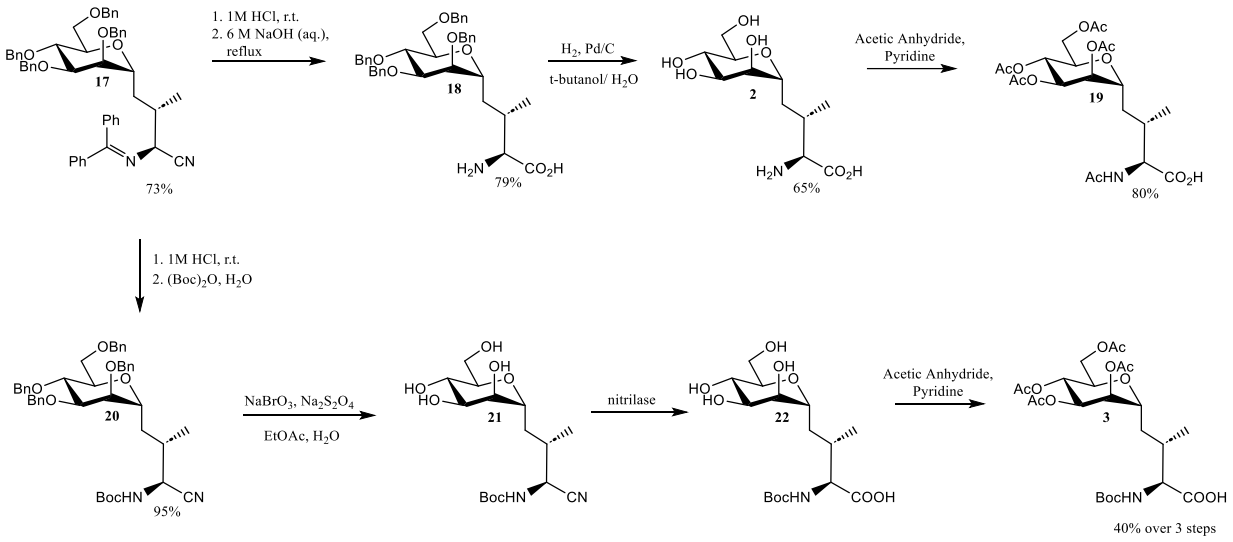


Figure 3.3. ^{1}H-NMR spectrum of O'Donnell *C*-glycoside derivative **17**

Reported methods for O'Donnell amino acid synthesis simultaneously convert the nitrile moiety to the carboxylic acid and the imine to the free amine via acid hydrolysis with 6N HCl. However, such harsh acidic conditions are not viable for sugar derivatives. Therefore, O'Donnell product **17** was transformed to the amino acid derivative with a two-step process. Specifically, compound **17** was treated with 1 M HCl at room temperature for 2 hours to convert the imine to the free amine, followed by refluxing in 6 M NaOH overnight to transform the nitrile to the carboxylic acid moiety of the amino acid derivative. Amphiphilic compound **18** is difficult to purify due to its insolubility in organic solvents. Therefore, the crude mixture underwent hydrogenation, followed by acetylation of the free sugar hydroxyls and free amine to afford acetylated *C*-Man-Thr derivative **19** (**Scheme 3.4**).



Scheme 3.4. *C*-glycoside Synthetic Routes from O'Donnell Intermediates

Optimization of *C*-Mannosyl Threonine Synthesis

Through efforts to optimize the synthesis of Mannose-*C*-Threonine, an alternative route was developed in which O'Donnell product **17** was first treated with 1 M HCl to convert the imine to the free amine, and the amine was subsequently Boc protected to afford compound **20**. Traditional hydrogenation conditions were not amenable for benzyl deprotection of the Boc protected *C*-glycoside due to the presence of the nitrile moiety. Using an alternative radical debenzylation method, the benzyl protecting groups of the hydroxyls were removed to from water-soluble compound **21**.[25] Employing enzymatic transformation with a nitrilase, the nitrile moiety was converted to the free carboxylic acid in PBS.[26] The nitrilase conversion only proceeded with one of the pure isomers, suggesting that this isomer is the natural L-amino acid derivative **22**. The free hydroxyls were then protected as acetyl groups and purified via silica gel chromatography (**Scheme 3.4**).

DISCUSSION

Two synthetic routes to create *C*-Man-Thr, hinging upon alkylation with secondary alkyl iodide derivative **5**, were developed. Limited syntheses to create *C*-glycosyl β-branched α-amino acids moieties have been previously carried out, likely due to the challenges in both stereochemical requirements and inherent challenges with $C(sp^3)$-$C(sp^3)$ alkylation reactions required to obtain such constructs. The syntheses described herein exploit many unconventional chemical transformations that have potential to aid in the synthesis of other unnatural amino acid derivatives with biological relevance. For example, employing the nitrilase enzyme to avoid harsh acidic and basic hydrolysis conditions to transform the nitrile moiety to the carboxylic acid is potentially valuable for future unnatural amino acid derivative syntheses. Furthermore, the radical benzylation strategy offers a means to deprotect the benzyl protected hydroxyls of the sugar without disturbing the nitrile moiety. The robustness of nitrile functional groups to the bromide radical generated in this debenzylation reaction has not been previously explored to our knowledge, and this could be a valuable transformation for future *O*- and *C*-glycoside syntheses.

Cross-coupling reactions of inactivated alkyl halides remains elusive. Several efforts have been made in recent years to overcome the challenges of the undesired β-elimination side reaction occurring during cross coupling with alkyl halides.[20] In particular, cross couplings of secondary alkyl halides are difficult to achieve due to the comparatively rapid elimination to an olefin, as observed with the oxazinone chemistry during the synthesis of *C*-Man-Thr. Interestingly, the O'Donnell amino acid synthesis approach exploited herein employing the benzophenonamine of acetonitrile as the Schiff base provided a viable cross-coupling pathway, while the Schiff Base containing the *tert*-butyl ester in place of the nitrile moiety yielded the

eliminated olefin as a major product. Due to the biphasic nature of the O'Donnell reaction, deprotonation of enolate occurs at interface. The nitrile is strongly electron withdrawing, resulting in positive charge build-up on the sp^2 carbon center. We therefore rationalize that the carbon center becomes highly electrophilic, promoting formation of the enolate, thus providing the O'Donnell amino acid precursor as the kinetically favored product, as opposed to the β-eliminated product. The presence of the nitrile functional group enhanced chemoselectivity in regards to formation of the enolate, and this observation offers exciting new avenues for $C(sp^3)$-$C(sp^3)$ cross-electrophile couplings. It would also be of interest to further explore this transformation to expand the reaction-substrate scope.

A critical intermediate for establishing the stereochemistry of the β-carbon was *C*-glycosyl S-alcohol derivative **7**. NMR data, including J-coupling values and NOEs, along with rotamer analysis together supported the formation of the desired isomer. Specifically, the S (180°, 60°) rotamer coincided with the proton-proton J-coupling and NOE data for the major alcohol. Distinguishing the Pro-R and Pro-S diastereotopic protons (H_{7d} and H_{7u}) with distinct chemical shifts provided the foundation for configurational and stereochemical assignments. A key piece of data to support the S-stereoconfigurational assignment was the approximately equal J-values of H_{7d} to $H_{8\_alcohol}$ and H_{7u} to $H_{8\_alcohol}$ J-coupling values, supporting equal bond angles between H_{7d}-$H_{8\_alcohol}$ and H_{7u}-$H_{8\_alcohol}$, as observed in the S (180°, 60°) rotamer, but not the R (180°, 60°) rotamer. Further studies exploiting Mosher's ester will be carried out to further support this stereochemical assignment.

Future work to further optimize the synthesis *C*-Man-Thr would also be of value. Given the challenges faced with $C(sp^3)$-$C(sp^3)$ alkylations, there is a desire for novel routes to establish

these C-C bonds. Last year, electrochemically driven cross-electrophile coupling of alkyl halides was reported.[20] This synthesis involves a secondary or tertiary alkyl halide undergoing selective cathodic reduction to generate a carbanion, which then is then subjected to bimolecular nucleophilic substitution with a less substituted alkyl halide in a glove box.[20] Significantly, direct electrolysis to create the C-C bond circumvents the undesired β-eliminated product formation by activating alkyl halides. Testing the scope of the electrolysis conditions to create *C*-Man-Thr could potentially be of value. Additionally, chiral phase-transfer catalysts exploited with the O'Donnell amino acid synthesis to develop a route with high stereoselectivity would also be highly valuable in optimization of *C*-glycosyl Threonine syntheses.[9]

We envision extending *C*-Man-Thr into *C*-glycopeptides for immunization to obtain antibodies for detection of core *O*-mannose glycan structures. *C*-glycosyl M0 mimetic will be extended to M1 and M3 core mimetics enzymatically using glycosyl transferases POMGNT1 and POMGNT2. This will aid in understanding if other unextended M3 sites exist beyond the extended M3 sites found in α-DG. Furthermore, M0 antibodies offer potential for detecting novel unextended *O*-man glycans in brain and muscle tissue, beyond those of α-DG, cadherins and plexins. Better understanding *O*-mannosylation by probing with chemical tools, such as *C*-glycopeptides, offers the potential for future therapeutic intervention of CMDs.

METHODS

2-Methyl-3-(2,3,4,6-tetra-O-benzyl-α-D-mannopyranosyl)propene (11)

Benzoylated methyl α-D-mannopyranoside (2.00 g, 3.61 mmol), allylsilane (2.5 mL, 9.63 mmol), and 20 mL of dry acetonitrile were added to a 50 mL round bottom flask. Trimethylsilyl

triflate (0.48 mL, 2.16 mmol) was added via syringe, and the mixture stirred at room temperature for 4 hours at 0 °C. The reaction was quenched with triethylamine (X mL), and the mixture was extracted with ethyl acetate, dried with $MgSO_4$, filtered, and concentrated. Flash silica gel chromatography (9:1 Toluene: EtOAc) gave alkene 1.47 g as a colorless oil in 70% yield. ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.33 (d, *J* = 6.9 Hz, 2H), 7.26 (ddt, *J* = 23.5, 6.5, 4.8 Hz, 10H), 7.19 – 7.14 (m, 2H), 4.75 – 4.69 (m, 2H), 4.60 (d, *J* = 5.7 Hz, 3H), 4.58 – 4.47 (m, 6H), 4.15 (td, *J* = 7.4, 3.8 Hz, 1H), 3.88 (t, *J* = 7.5 Hz, 1H), 3.75 (ddd, *J* = 15.9, 10.4, 5.0 Hz, 3H), 3.69 (dd, *J* = 10.1, 3.1 Hz, 1H), 3.60 (t, *J* = 3.5 Hz, 1H), 2.25 (dq, *J* = 15.2, 7.7 Hz, 2H). ^{13}C NMR (600 MHz, Chloroform-*d*) δ 138.32, 128.39, 128.34, 128.29, 128.05, 127.88, 127.79, 127.71, 127.70, 127.64, 127.47, 112.92, 77.67, 75.20, 75.03, 74.26, 73.49, 73.32, 72.00, 71.82, 71.73, 69.35, 38.19, 22.23.

2-Methyl-3-(2,3,4,6-tetra-O-benzyl-α-D-mannopyranosyl)propone (12)

To a solution of allyl *C*-mannoside (0.100g, 0.172 mmol) in H_2O (3 mL) and THF (3 mL) was added sodium periodate (0.222 g, 1.04 mmol) and 2.5% osmium tetroxide in *t*-butanol (100 uL). The reaction stirred for 16 h. The mixture was then diluted with DCM (20 mL), and the organic layer was washed with water (2x), dried with $MgSO_4$, filtered, and concentrated. Flash silica gel chromatography (10:1 Hexanes: EtOAc) gave ketone **12** as a colorless oil in 79% yield. ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.34 – 7.25 (m, 13H), 7.24 – 7.18 (m, 4H), 4.58 (d, *J* = 11.6 Hz, 1H), 4.55 – 4.49 (m, 5H), 4.51 – 4.43 (m, 3H), 3.89 (q, *J* = 5.3 Hz, 1H), 3.79 (td, *J* = 6.5, 3.1 Hz, 2H), 3.77 – 3.71 (m, 2H), 3.59 (dd, *J* = 6.6, 2.9 Hz, 1H), 2.70 – 2.60 (m, 2H), 2.13 (s, 2H).

2-Methyl-3-(2,3,4,6-tetra-O-benzyl-α-D-mannopyranosyl)propenol-S (7)

Ketone **12** (1.00 g, 1.72 mmol) was dissolved in a mixture of isopropanol (66 mL) and acetic acid (14 mL). Sodium borahydride (325 mg, 8.58 mmol) was added slowly to the solution. The reaction stirred for 2 hours. Column chromatography (gradient 8:1 to 6:1 toluene: ethyl acetate) afforded alcohol **7** in 63% yield. ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.37 – 7.20 (m, 20H), 7.22 – 7.16 (m, 2H), 4.63 – 4.48 (m, 7H), 4.48 (d, *J* = 11.7 Hz, 1H), 4.17 (ddd, *J* = 10.0, 6.2, 3.2 Hz, 1H), 4.00 (h, *J* = 5.0 Hz, 2H), 3.78 – 3.72 (m, 2H), 3.70 (t, *J* = 5.6 Hz, 1H), 3.62 (dd, *J* = 10.3, 4.1 Hz, 1H), 3.55 (dd, *J* = 6.3, 2.8 Hz, 1H), 3.53 (s, 1H), 1.74 – 1.62 (m, 2H), 1.17 (d, *J* = 6.2 Hz, 3H). ^{13}C NMR (600 MHz, Chloroform-*d*) δ 138.03, 127.71, 77.216, 77.18, 76.02, 74.89, 73.28, 71.75, 68.51, 67.98, 38.59.

2-Methyl-3-(2,3,4,6-tetra-O-benzyl-α-D-mannopyranosyl)propyl-Iodide-R (5)

Triphenyl phosphine (1.17 g, 0.635 mmol) and imidazole (0.518 g, 7.62 mmol) were dissolved in dry DCM (75 mL) and cooled to 0 °C and allowed to stir for 20 minutes. Iodine (1.13 g, 4.44 mmol) was added slowly. The reaction stirred for 20 minutes and then alcohol **7** (0.370 g, 0.635 mmol) was added. The reaction stirred at room temperature overnight. The mixture was then diluted with DCM (75 mL) and the organic layer was washed with sodium thiosulfate 2x, followed by water. Column chromatography using 9:1 Hexanes: Ethyl acetate afforded the product in 77% yield. ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.37 – 7.26 (m, 13H), 7.28 – 7.23 (m, 1H), 7.25 – 7.17 (m, 2H), 4.63 – 4.47 (m, 9H), 4.34 (dqd, *J* = 10.3, 6.9, 3.2 Hz, 1H), 4.18 (ddd, *J* = 9.7, 6.4, 2.6 Hz, 1H), 3.82 (tt, *J* = 12.4, 5.3 Hz, 3H), 3.80 – 3.74 (m, 1H), 3.73 (dd, *J* = 5.9, 3.0 Hz, 1H), 3.58 (dd, *J* = 6.3, 2.9 Hz, 1H), 2.03 (ddd, *J* = 14.3, 8.0, 2.7 Hz, 1H), 1.93 (d, *J* =

6.9 Hz, 3H), 1.63 (ddd, *J* = 14.9, 10.2, 3.3 Hz, 1H). ^{13}C NMR (600 MHz, Chloroform-*d*) δ 138.37, 128.39, 128.38, 128.37, 128.31, 128.02, 127.94, 127.83, 127.78, 127.75, 127.72, 127.71, 127.51, 75.87, 75.73, 74.59, 74.22, 73.47, 73.17, 72.20, 71.58, 71.35, 69.02, 43.10, 29.55, 26.35.

2,3,4,6-tetra-O-benzyl-α-D-mannopyranosyl-(2S,3S)-2-((diphenylmethylene)amino)-3-methyl-4-butanenitrile (17)

N-(Diphenylmethylene)aminoacetonitrile **16** (1 g, 4.5 mmol), benzyltriethylammonium chloride (2.4 g, 4.39 mmol), NaOH (0.62 mL as a 50% aqueous solution) and toluene (2 mL) were stirred together at 0 °C in a round bottom flask sealed with a rubber septum. Iodide **5** (130 mg, 0.188 mmol) in 1 mL toluene was added dropwise via syringe over a period of 1.5 hour at 0 °C. The resulting solution was stirred at room temperature overnight. The next day, the mixture was diluted with 5 mL DCM and poured into a separatory funnel containing water (5 mL). The layers were separated, and the aqueous layer was extracted with DCM (3 x 5 mL). The organic layers were combined, washed with water and brine, dried over anhydrous magnesium sulfate, and filtered. The solvent was removed via rotary evaporation. Chromatography [flash silica, with hexanes/ethyl acetate (9:1 v/v)] gave **17** (0.108 g, 73%) as a pale yellow oil. ***2,3,4,6-tetra-O-benzyl-α-D-mannopyranosyl-(2S,3S)-2-((diphenylmethylene)amino)-3-methyl-4-butanenitrile (top isomer)***: ^{1}H NMR (900 MHz, $CDCl_3$) δ 7.57 – 7.53 (m, 2H), 7.41 – 7.35 (m, 3H), 7.37 – 7.31 (m, 1H), 7.31 – 7.22 (m, 4H), 7.24 – 7.16 (m, 12H), 7.16 (dd, *J* = 6.5, 3.1 Hz, 2H), 7.16 – 7.12 (m, 1H), 7.10 (ddd, *J* = 36.2, 6.9, 1.9 Hz, 4H), 4.55 (d, *J* = 11.4 Hz, 1H), 4.53 – 4.39 (m, 7H), 4.06 (d, *J* = 5.2 Hz, 1H), 3.94 (ddd, *J* = 11.4, 5.4, 2.8 Hz, 1H), 3.78 – 3.73 (m, 1H), 3.74 – 3.66 (m, 2H), 3.66 (dd, *J* = 6.3, 3.2 Hz, 1H), 3.59 (dd, *J* = 10.2, 4.1 Hz, 1H), 3.44 (ddd, *J* = 8.4,

5.6, 3.0 Hz, 1H), 2.15 – 2.06 (m, 1H), 1.65 (ddd, *J* = 14.6, 11.3, 3.8 Hz, 1H), 1.37 (ddd, *J* = 13.9, 10.4, 2.8 Hz, 1H), 1.11 (dd, *J* = 17.4, 6.8 Hz, 3H). ***2,3,4,6-tetra-O-benzyl-α-D-mannopyranosyl-(2R,3S)-2-((diphenylmethylene)amino)-3-methyl-4-butanenitrile*** (***bottom isomer***): ^{1}H NMR (900 MHz, $CDCl_3$) δ 7.56 – 7.53 (m, 2H), 7.43 – 7.35 (m, 3H), 7.36 – 7.29 (m, 1H), 7.27 – 7.14 (m, 13H), 7.11 (td, *J* = 7.6, 2.0 Hz, 3H), 7.08 – 7.05 (m, 2H), 4.62 (d, *J* = 11.4 Hz, 1H), 4.54 – 4.46 (m, 3H), 4.40 (dd, *J* = 11.8, 8.7 Hz, 2H), 4.34 (d, *J* = 12.1 Hz, 1H), 4.07 – 4.02 (m, 1H), 4.01 – 3.95 (m, 1H), 3.78 – 3.71 (m, 2H), 3.58 (dd, *J* = 10.4, 6.1 Hz, 1H), 3.54 – 3.48 (m, 2H), 2.14 – 2.06 (m, 2H), 1.50 (s, 1H), 1.31 – 1.25 (m, 1H), 0.99 (d, *J* = 6.7 Hz, 2H). ^{13}C NMR (600 MHz, Chloroform-*d*) δ 15.78 (CH_3-9), 32.21 (CH_2-7), 34.39 (CH-8), 58.74 (CH-10), 69.39 (CH_2-6), 70.04 (CH-1), 73.09 (CH-5), 75.10 (CH-4), 76.64 (CH-2), 77.33 (CH-3), 127.29, 127.24, 128.97, 131.11 (C-aromatic), 175.0 (*C*N).

2,3,4,6-tetra-O-acetyl-α-D-mannopyranosyl-*C*-Threonine (19)

The O'Donnell amino acid derivative **17** (20 mg, 0.025 mmol) was treated with 1M HCl (4 mL) in acetone (2 mL) overnight. The resulting amine was then concentrated and refluxed in 12.5% NaOH overnight. The reaction was neutralized by the addition of 3M HCl and concentrated. The resulting mixture was dissolved in 1:1 t-butanol:water and treated with palladium hydroxide on carbon (10 mol%) under hydrogen gas overnight. The resulting product was filtered, dried down *in vacuo*, and treated with pyridine (1 mL) and acetic anhydride (0.5 mL), and the mixture was allowed to stir overnight. The mixture was concentrated on silica and chromatographed with 50% ethyl acetate: hexanes to afford the product as a white solid in 80% yield. ^{1}H NMR (600 MHz, cd_3od) δ 5.94 (ddd, *J* = 31.0, 9.5, 5.4 Hz, 2H), 5.73 (ddd, *J* = 15.0, 9.5, 5.8 Hz, 2H), 5.46

(q, *J* = 6.2 Hz, 1H), 5.45 – 5.37 (m, 2H), 5.40 – 5.34 (m, 1H), 4.02 (s, 1H), 2.63 – 2.41 (m, 5H), 2.41 – 2.30 (m, 2H), 2.28 – 2.04 (m, 16H), 2.02 – 1.95 (m, 2H), 1.96 (s, 4H), 1.52 (s, 2H), 1.17 (s, 2H). Expected Mass (M+H): 490.47; Observed Mass: 490.25.

2,3,4,6-tetra-O-benzyl-α-D-mannopyranosyl-(2S,3S)-2-((Boc-amino)-3-methyl-4-butanenitrile (20)

The O'Donnell amino acid derivative **17** (20 mg, 0.025 mmol) was treated with 1M HCl (4 mL) in acetone (2 mL) overnight, followed by concentration *in vacuo*. Boc-anhydride (0.011 μL, 0.046 mmol) was added to the residue, and the mixture was sonicated for 20 minutes, after which it was concentrated on silica and purified using 9:1 Hexanes:Ethyl Acetate to afford 15 mg of **20** in 95% yield. ***2,3,4,6-tetra-O-benzyl-α-D-mannopyranosyl-(2S,3S)-2-((Boc-amino)-3-methyl-4-butanenitrile (top isomer)***: ^{1}H NMR (800 MHz, $CDCl_3$) δ 7.22 (dtd, *J* = 20.0, 8.7, 5.9 Hz, 15H), 7.19 – 7.13 (m, 2H), 7.13 – 7.08 (m, 2H), 5.02 (td, *J* = 8.3, 2.9 Hz, 1H), 4.51 – 4.36 (m, 8H), 3.89 (ddd, *J* = 11.3, 6.3, 2.4 Hz, 1H), 3.81 (q, *J* = 5.1 Hz, 1H), 3.74 – 3.63 (m, 3H), 3.59 (ddd, *J* = 19.5, 10.1, 4.1 Hz, 1H), 3.42 (td, *J* = 7.5, 2.4 Hz, 1H), 2.01 (dtd, *J* = 10.7, 7.0, 3.9 Hz, 1H), 1.60 (ddt, *J* = 20.1, 12.8, 6.5 Hz, 1H), 1.48 (m, 1H), 0.99 (d, *J* = 6.9 Hz, 3H). ***2,3,4,6-tetra-O-benzyl-α-D-mannopyranosyl-(2R,3S)-2-((Boc-amino)-3-methyl-4-butanenitrile (bottom isomer)***: ^{1}H NMR (800 MHz, $CDCl_3$) δ 7.31 (tq, *J* = 11.4, 4.8 Hz, 1H), 7.28 – 7.17 (m, 14H), 7.17 – 7.13 (m, 2H), 7.13 – 7.10 (m, 2H), 4.45 (dd, *J* = 12.1, 4.2 Hz, 4H), 4.41 (td, *J* = 10.6, 7.3 Hz, 4H), 3.90 – 3.84 (m, 2H), 3.72 (dd, *J* = 10.2, 7.0 Hz, 1H), 3.68 (d, *J* = 3.1 Hz, 2H), 3.57 (dd, *J* = 9.9, 4.7 Hz, 1H), 3.44 (dd, *J* = 6.9, 1.9 Hz, 1H), 1.98 (dt, *J* = 12.6, 6.4 Hz, 1H), 1.64 (ddd, *J* = 15.7, 10.9, 5.3 Hz, 1H), 1.46 (dd, *J* = 11.8, 6.5 Hz, 1H), 0.98 (d, *J* = 7.1 Hz, 3H).

2,3,4,6-tetra-O-acetyl-α-D-mannopyranosyl-Boc-amino-*C*-Threonine (3)

The Boc-protected O'Donnell derivative (10 mg, 0.021 mmol) was dissolved in 0.27 mL of ethyl acetate. Sodium bromate (38 mg, 0.252 mmol) was dissolved in 0.2 mL of water and added to the reaction stirring solution. Sodium dithionate (44 mg, 0.252mmol) was dissolved in 0.4 mL of water and added dropwise to the stirring solution. After reacting for three hours, the reaction was quenched by the addition of sodium thiosulfate (1 mL), and the product was extracted into the ethyl acetate layer. The product was concentrated onto silica and chromatographed with 10% methanol in dichloromethane, followed by concentration *in vacuo*. The resulting debenzylated product was treated with 10 mol% nitrilase enzyme overnight in PBS buffer at room temperature. The mixture was concentrated, and pyridine (1 mL) and acetic anhydride (0.5 mL) were added. The reaction was left overnight. The following day, it was concentrated onto silica gel and chromatographed (50% ethyl acetate/hexanes) to afford 3 mg of a white solid (40% yield over 3 steps). ^{1}H NMR (800 MHz, $CDCl_3$) δ 5.28 – 5.26 (m, 1H), 5.13, 5.11 (d, *J* = 7.9 Hz, 1H), 5.04 (dd, *J* = 8.2, 4.7 Hz, 2H), 4.69 (d, *J* = 32.8 Hz, 1H), 4.43 – 4.35 (m, 1H), 4.05 (d, *J* = 7.1 Hz, 1H, α-H-Thr), 2.27 (t, *J* = 7.6 Hz, 1H), 2.06 (s, 3H), 2.06 (m, 1H), 2.03 (s, 3H), 2.02 (s, 3H), 2.01 (s, 3H), 1.98 (m, 3H), 1.56 (t, *J* = 7.5 Hz, 1H). ^{13}C NMR (600 MHz, Chloroform-*d*) δ 20.65 (CH_3-9), 20.76 (4 x Ac), 34.57 (CH_2-7), 62.41, 67.72, 70.27, 71.23, 71.55 (CH-10, α-*C*-Thr), 77.18 (CH-1), 77.35.

SUPPLEMENTARY DATA

Proton 1	Proton 2	NOE Signal	
		Minor (R)	Major (S)
H1	H2	Medium	Medium
	H6, H6’	Medium	Medium
	Alcohol	Medium	Medium
	Methyl	Weak	Strong
	(d)Methylene	Strong	Weak
	(u)Methylene	Weak	Strong
Alcohol	(d)Methylene	Medium	Strong
	H1	Medium	Medium
	H2	Weak	Medium
	(u)Methylene	Strong	Medium
	H5	None	Weak
Methylene (d)	H1	Strong	Weak
	H2	Medium	Weak
	H3	Weak	Medium
	H5	Weak	Strong
	Alcohol	Medium	Medium
Methylene (u)	H1	Weak	Strong
	H2	Medium	Strong
	H3	Strong	Medium
	H5	Strong	Weak
	Alcohol	Medium	Medium
Methyl	Alcohol	Strong	Strong
	H1	Weak	Strong
	umethylene	Medium	Strong

	dmethylene	Medium	Medium

Table 3.1S. NOE Data for Minor (R) and Major (S) Acetylated Alcohol Isomer

Rotamer	Stereochemistry	Pass/Fail	Rationale
180, 180	S compared to Major (S)	Fail	Fails strong NOE from H1 to methyl; fails equal J-value for H8_alcohol to methylene protons (HRu and HSd)
	R compared to Major (S)	Fail	Fails equal J-value for H8_alcohol to methylene protons (HRu and HSd); fails strong NOE from H1 to methyl
	R compared to Minor (R)	Fail	Fails transdiaxial angle between methylene (HRu) to H1 (J-value 9.7)
180, 60	S compared to Major (S)	Pass	Passes equal equal J-value for H8_alcohol to methylene protons; passes strong NOE from H1 to methyl; passes HSd (methylene) to H1 transdiaxial (J-value 9.2)
	R compared to Major (S)	Fail	Fails equal J-value for alcohol to methylene protons (HRu and HSd)
	R compared to Minor (R)	Fail	Fails HRu (methylene) to H1 (transdiaxial; J-value: 9.7); Fails HSd (methylene) to H8_alcohol (transdiaxial; J-value 8.9)
180, 300	S compared to Major (S)	Fail	Fails equal J-value for H8_alcohol to methylene protons (HRu and HSd); methyl further away so strong NOE from H1 to methyl not expected

	R compared to Major (S)	Fail	No H5/H6 to methyl present as would be expected in this rotamer; fails H8_alcohol to H2 NOE; methyl further away so strong NOE from H1 to methyl not expected; fails strong NOE form HRu to methyl (while S (180,60) matches)
	R compared to Minor (R)	Fail	Fails HRu (methylene) to H1 (transdiaxial; J-value: 9.7); Fails HSd (methylene) to H8_alcohol (transdiaxial; J-value 8.9)
60, 180	S compared to Major (S)	Fail	Fails strong NOE from H1 to methyl; fails equal J-value for H8_alcohol to methylene protons (HRu and HSd)
	R compared to Major (S)	Fail	Fails strong NOE from H1 to methyl; fails equal J-value for H8_alcohol to methylene protons (HRu and HSd)
	R compared to Minor (R)	Fail	No NOE from H8_alcohol to H3/H5, so not consistent
60, 60	S compared to Major (S)	Fail	Sterics; No NOE observed from methyl to H5
	R compared to Major (S)	Fail	Sterics; No NOE observed from methyl to H5
	R compared to Minor (R)	Fail	Sterics; No NOE observed from methyl to H5
300, 60	S compared to Major (S)	Fail	No NOE from Methyl to H2, so not consistent with rotamer; fails H1 to HSd (methylene)
	R compared to Major (S)	Fail	No NOE from Methyl to H2, so not consistent with rotamer; fails equal J-value for

			H8_alcohol to methylene protons (HRu and HSd)
	R compared to Minor (R)	Fail	No NOE from Methyl to H2, so not consistent with rotamer
300, 180	S compared to Major (S)	Fail	Fails strong NOE from H1 to methyl; fails equal J-value for H8_alcohol to methylene protons (HRu and HSd)
	R compared to Major (S)	Fail	Fails strong NOE from H1 to methyl; fails equal J-value for H8_alcohol to methylene protons (HRu and HSd)
	R compared to Minor (R)	Pass	Passes H8_alcohol to HSd(methylene) transdiaxial (J-value 8.9); passes H1 to HRu(methylene) transdiaxial (J-value 9.7); consistent w/ weak NOE from methyl to H1
60, 300	S compared to Major (S)	Fail	Sterics; No NOE observed from methyl to H3; fails equal J-value for H8_alcohol to methylene protons (HRu and HSd); fails strong NOE from H1 to methyl
	R compared to Major (S)	Fail	Sterics; No NOE observed from methyl to H3
	R compared to Minor (R)	Fail	Sterics; No NOE observed from methyl to H3; fails J-values for methylene protons
300, 300	S compared to Major (S)	Fail	Fails equal J-value for H8_alcohol to methylene protons (HRu and HSd); fails H1 to HSd(methylene) transdiaxial J value: 9.2

	R compared to Major (S)	Fail	Fails H5 to HSd (methylene) strong NOE; fails H1 to HSd (methylene) transdiaxial J value: 9.2
	R compared to Minor (R)	Fail	Fails methylene J-values; not consistent with weak NOE from H1 to methyl; no H5,H3 to methylene, so not consistent

Table 3.2S. Rotamer Analysis

REFERENCES

(1) *Essentials of Glycobiology*, 4th ed.; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2022.

(2) Chao, E. C.; Henry, R. R. SGLT2 Inhibition--a Novel Strategy for Diabetes Treatment. *Nat. Rev. Drug Discov.* **2010**, *9* (7), 551–559. https://doi.org/10.1038/nrd3180.

(3) Wang, X.; Zhang, L.; Byrne, D.; Nummy, L.; Weber, D.; Krishnamurthy, D.; Yee, N.; Senanayake, C. H. Efficient Synthesis of Empagliflozin, an Inhibitor of SGLT-2, Utilizing an AlCl3-Promoted Silane Reduction of a β-Glycopyranoside. *Org. Lett.* **2014**, *16* (16), 4090–4093. https://doi.org/10.1021/ol501755h.

(4) Henschke, J. P.; Lin, C.-W.; Wu, P.-Y.; Tsao, W.-S.; Liao, J.-H.; Chiang, P.-C. β-Selective C-Arylation of Diisobutylaluminum Hydride Modified 1,6-Anhydroglucose: Synthesis of

Canagliflozin without Recourse to Conventional Protecting Groups. *J. Org. Chem.* **2015**, *80* (10), 5189–5195. https://doi.org/10.1021/acs.joc.5b00601.

(5) Guo, C.; Hu, M.; DeOrazio, R. J.; Usyatinsky, A.; Fitzpatrick, K.; Zhang, Z.; Maeng, J.-H.; Kitchen, D. B.; Tom, S.; Luche, M.; Khmelnitsky, Y.; Mhyre, A. J.; Guzzo, P. R.; Liu, S. The Design and Synthesis of Novel SGLT2 Inhibitors: C-Glycosides with Benzyltriazolopyridinone and Phenylhydantoin as the Aglycone Moieties. *Bioorg. Med. Chem.* **2014**, *22* (13), 3414–3422. https://doi.org/10.1016/j.bmc.2014.04.036.

(6) Cai, W.; Jiang, L.; Xie, Y.; Liu, Y.; Liu, W.; Zhao, G. Design of SGLT2 Inhibitors for the Treatment of Type 2 Diabetes: A History Driven by Biology to Chemistry. *Med. Chem. Shariqah United Arab Emir.* **2015**, *11* (4), 317–328. https://doi.org/10.2174/1573406411666150105105529.

(7) Yang, Y.; Yu, B. Recent Advances in the Chemical Synthesis of C-Glycosides. *Chem. Rev.* **2017**, *117* (19), 12281–12356. https://doi.org/10.1021/acs.chemrev.7b00234.

(8) Leseurre, L.; Merea, C.; Paule, S. D. de; Pinchart, A. Eco-Footprint: A New Tool for the "Made in Chimex" Considered Approach. *Green Chem.* **2014**, *16* (3), 1139–1148. https://doi.org/10.1039/C3GC42201A.

(9) Cavezza, A.; Boulle, C.; Guéguiniat, A.; Pichaud, P.; Trouille, S.; Ricard, L.; Dalko-Csiba, M. Synthesis of Pro-Xylane: A New Biologically Active C-Glycoside in Aqueous Media. *Bioorg. Med. Chem. Lett.* **2009**, *19* (3), 845–849. https://doi.org/10.1016/j.bmcl.2008.12.037.

(10) Pineau, N.; Carrino, D. A.; Caplan, A. I.; Breton, L. Biological Evaluation of a New C-Xylopyranoside Derivative (C-Xyloside) and Its Role in Glycosaminoglycan Biosynthesis. *Eur. J. Dermatol. EJD* **2011**, *21* (3), 359–370. https://doi.org/10.1684/ejd.2011.1340.

(11) Yang, G.; Schmieg, J.; Tsuji, M.; Franck, R. W. The C-Glycoside Analogue of the Immunostimulant α-Galactosylceramide (KRN7000): Synthesis and Striking Enhancement of Activity. *Angew. Chem. Int. Ed.* **2004**, *43* (29), 3818–3822. https://doi.org/10.1002/anie.200454215.

(12) Larsen, I. S. B.; Narimatsu, Y.; Joshi, H. J.; Siukstaite, L.; Harrison, O. J.; Brasch, J.; Goodman, K. M.; Hansen, L.; Shapiro, L.; Honig, B.; Vakhrushev, S. Y.; Clausen, H.; Halim, A. Discovery of an O-Mannosylation Pathway Selectively Serving Cadherins and Protocadherins. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (42), 11163–11168. https://doi.org/10.1073/pnas.1708319114.

(13) Vester-Christensen, M. B.; Halim, A.; Joshi, H. J.; Steentoft, C.; Bennett, E. P.; Levery, S. B.; Vakhrushev, S. Y.; Clausen, H. Mining the O-Mannose Glycoproteome Reveals Cadherins as Major O-Mannosylated Glycoproteins. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (52), 21018–21023. https://doi.org/10.1073/pnas.1313446110.

(14) Rubinstein, R.; Thu, C. A.; Goodman, K. M.; Wolcott, H. N.; Bahna, F.; Mannepalli, S.; Ahlsen, G.; Chevee, M.; Halim, A.; Clausen, H.; Maniatis, T.; Shapiro, L.; Honig, B. Molecular Logic of Neuronal Self-Recognition through Protocadherin Domain Interactions. *Cell* **2015**, *163* (3), 629–642. https://doi.org/10.1016/j.cell.2015.09.026.

(15) Stalnaker, S. H.; Stuart, R.; Wells, L. Mammalian O-Mannosylation: Unsolved Questions of Structure/Function. *Curr. Opin. Struct. Biol.* **2011**, *21* (5), 603–609. https://doi.org/10.1016/j.sbi.2011.09.001.

(16) Boggon, T. J.; Murray, J.; Chappuis-Flament, S.; Wong, E.; Gumbiner, B. M.; Shapiro, L. C-Cadherin Ectodomain Structure and Implications for Cell Adhesion Mechanisms. *Science* **2002**, *296* (5571), 1308–1313. https://doi.org/10.1126/science.1071559.

(17) Gustafsson, T.; Saxin, M.; Kihlberg, J. Synthesis of a C-Glycoside Analogue of β-d-Galactosylthreonine. *J. Org. Chem.* **2003**, *68* (6), 2506–2509. https://doi.org/10.1021/jo026758d.

(18) Bragnier, N.; Guillot, R.; Scherrmann, M.-C. Diastereoselective Addition of Sugar Radicals to Camphorsultam Glyoxilic Oxime Ether: A Route toward C-Glycosylthreonine and Allothreonine. *Org. Biomol. Chem.* **2009**, *7* (19), 3918–3921. https://doi.org/10.1039/B910050D.

(19) Nolen, E. G.; Watts, M. M.; Fowler, D. J. Synthesis of C-Linked Glycopyranosyl Serines via a Chiral Glycine Enolate Equivalent. *Org. Lett.* **2002**, *4* (22), 3963–3965. https://doi.org/10.1021/ol026839w.

(20) *Electrochemically driven cross-electrophile coupling of alkyl halides | Nature*. https://www.nature.com/articles/s41586-022-04540-4 (accessed 2023-07-22).

(21) *Cross-Couplings of Unactivated Secondary Alkyl Halides: Room-Temperature Nickel-Catalyzed Negishi Reactions of Alkyl Bromides and Iodides*. https://pubs.acs.org/doi/epdf/10.1021/ja0389366 (accessed 2023-07-22).

(22) Denmark, S. E.; Smith, R. C.; Chang, W.-T. T.; Muhuhi, J. M. Cross-Coupling Reactions of Aromatic and Heteroaromatic Silanolates with Aromatic and Heteroaromatic Halides. *J. Am. Chem. Soc.* **2009**, *131* (8), 3104–3118. https://doi.org/10.1021/ja8091449.

(23) Hosomi, A.; Sakata, Y.; Sakurai, H. Stereoselective Synthesis of 3-(d-Glycopyranosyl)Propenes by Use of Allylsilanes. *Carbohydr. Res.* **1987**, *171* (1), 223–232. https://doi.org/10.1016/S0008-6215(00)90889-9.

(24) O'Donnell, M. J. Benzophenone Schiff Bases of Glycine Derivatives: Versatile Starting Materials for the Synthesis of Amino Acids and Their Derivatives. *Tetrahedron* **2019**, *75* (27), 3667–3696. https://doi.org/10.1016/j.tet.2019.03.029.

(25) Adinolfi, M.; Barone, G.; Guariniello, L.; Iadonisi, A. Facile Cleavage of Carbohydrate Benzyl Ethers and Benzylidene Acetals Using the NaBrO3Na2S2O4 Reagent under Two-Phase Conditions. *Tetrahedron Lett.* **1999**, *40* (48), 8439–8441. https://doi.org/10.1016/S0040-4039(99)01756-6.

(26) Kiziak, C.; Stolz, A. Identification of Amino Acid Residues Responsible for the Enantioselectivity and Amide Formation Capacity of the Arylacetonitrilase from Pseudomonas Fluorescens EBC191. *Appl. Environ. Microbiol.* **2009**, *75* (17), 5592–5599. https://doi.org/10.1128/AEM.00301-09.

CHAPTER 4

GLYCOPEPTIDE CONJUGATES FOR IMMUNIZATION TO GENERATE ANIGENS FOR DETECTECTION OF CORE O-MANNOSE GLYCANS

INTRODUCTION

The protein alpha-dystroglycan (α-DG) is highly glycosylated with *O*-Man and *O*-GalNAc glycans through serine or threonine within its mucin-like domain. Defects in *O*-mannosylation of α-DG lead to dystroglycanopathies, which are congenital muscular dystrophies (CMDs) involving neurodevelopmental abnormalites.[1-4] The *O*-Man sites can be elaborated to core M1, M2, and M3 structures. Glycosaminyl transferase POMGNT1 adds GlcNAc in a β-1,2 linkage to make the M1 core. POMNGT1 displays broad substrate specificity. Another glycosaminyl transferase (MGAT5B) branches M1 with a β-1,6 GlcNAc linkage to form the M2 core. In contrast to the M1 and M2 cores, M3 is generated very selectively by POMGNT2 adding GlcNAc in a β-1,4 linkage. Thus, POMGNT2 is considered the "branching point" for *O*-mannosylation pathways.[3-4] Interestingly, *O*-mannose modified α-DG encounters POMGNT2 in the ER before POMGNT1 in the cis golgi. This implies that POMNGT2 must exhibit substrate sequence selectivity beyond simply an *O*-Man modified amino acid.[3-4]

Currently, the only reported method of detection for POMGNT2 modified M3 core *O*-mannose sites is the IIH6 antibody, which does not recognize the core glycan, but rather the extended polymer matriglycan. We hypothesize that other M3 sites exist, but are not fully

extended by matriglycan and subsequently are not detected by current available methods. The previous chapters have described development of methodologies and chemical tools to study *O*-mannosylation independently of matriglycan. The generation of antibodies to recognize the core *O*-mannose glycans would be valuable in interrogating these *O*-mannosylated sites and ultimately discerning their function. In the previous chapter, a *C*-glycoside was synthesized with the intention of creating a robust *C*-glycopeptide to generate antibodies to the core *O*-mannose glycans. Before creating the *C*-glycopeptide, optimization of the glycopeptide synthesis was required, given that traditional methods required large quantities of the glycoside that would be difficult to attain for the synthetic *C*-glycoside derivative.

When designing the *C*-glycopeptide mimetic, the ability of POMGNT1 and POMGNT2 to extend the equivalent *O*-glycopeptide first needed to be investigated. The peptide sequence was designed based on the M3 modified region of α-dystroglycan. Specifically, the glycopeptide is based on the Thr379 region of α-dystroglycan as an 11-mer sequence. The sequence has added tags on each end, including a cysteine residue (C) for future conjugation to CRM and a lysine residue (K) for future glycan array studies. The glycopeptide sequence is (C)GAIIQT(*O*-Man)PTLGP(K). We hypothesized that this glycopeptide would be extended by POMGNT1 and POMGNT2 and would provide the framework for creating the *C*-glycopeptide mimetics.

RESULTS

First, the *O*-glycopeptide was synthesized, and the synthesis was optimized to use a fraction of material compared to traditional methods. This optimization was necessary, because

the *C*-glycoside material is limited due to the challenging synthesis and difficulty producing the *C*-glycoside on a large-scale. With traditional methods using the Liberty Peptide Synthesizer, the synthesis is carried out on 0.10 mmol scale. Keeping the same instrument settings, the *O*-glycopeptide was synthesized on a 0.05 mmol scale. To further optimize the glycopeptide synthesis, only 1.1 equivalents of the Fmoc-protected glycoside were used compared to 2 equivalents in traditional methods. Lastly, traditional coupling methods involved employing two manual couplings, while the optimized synthesis yielded complete product conversion with just one coupling (**Table 4.1**).

	Excess Glycoside	**Synthesis Scale**	**# of Couplings**
Traditional Method	2 eq.	0.10 mmol	2
Optimized Method	1.1 eq.	0.05 mmol	1

Table 4.1. Optimized Glycopeptide Synthesis

After optimization of the manual coupling of glycopeptides, we next wanted to test the ability of glycosyltransferases to extend the *O*-glycopeptide. To create the M1 and M3 core *O*-mannose mimetics, the glycopeptides were extended with POMGNT1 and POMGNT2. The extended glycopeptides were confirmed via *matrix*-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI- TOF MS). First, glycopeptide (C)GAIIQT(*O*-Man)PTLGP(K) was treated with POMGNT1 and POMGNT2. MALDI data revealed that POMGNT1 modified this glycopeptide. However, POMGNT2 did not GlcNAc extend this peptide. Extended glycopeptide Shortman379 (GAIIQT(*O*-Man)PTLGPIQPTR)[3] was extended

by both POMGNT1 and POMGNT2, revealing the importance of the C-terminal IQPTR portion for POMGNT2 interaction.

Glycopeptide	POMGNT1 extension	POMGNT2 extension
CGAIIQT(*O*-Man)PTLGPK	Yes	No
GAIIQT(*O*-Man)PTLGPIQPTR	Yes	Yes

Table 4.2. POMGNT1 and POMGNT2 Extension of Glycopeptides

We then further characterized the glycopeptides by NMR analysis, ultimately comparing conformation of the peptide backbone of the POMGNT1 and POMGNT2 extended glycopeptides by carrying out NMR analysis of the glycopeptides in 90% 5 mM acetic acid/10% D_2O. A combination of 1D and 2D NMR, namely HSQC, COSY, and HMBC, experiments on samples in a mixture of 90% 5 mM acetic acid/10% D_2O were analyzed to discern the backbone assignments of each glycopeptide. After assigning the peptide backbone amino acid residues, comparison of the COSY spectra revealed differences in chemicals shifts of POMGNT1-extended (C)GAIIQT(*O*-Man)PTLGP(K) (red), but not POMGNT2 extended ShortMan379, compared to the non-extended *O*-Man (M0) mimetic (green) (**Figure 4.1**).

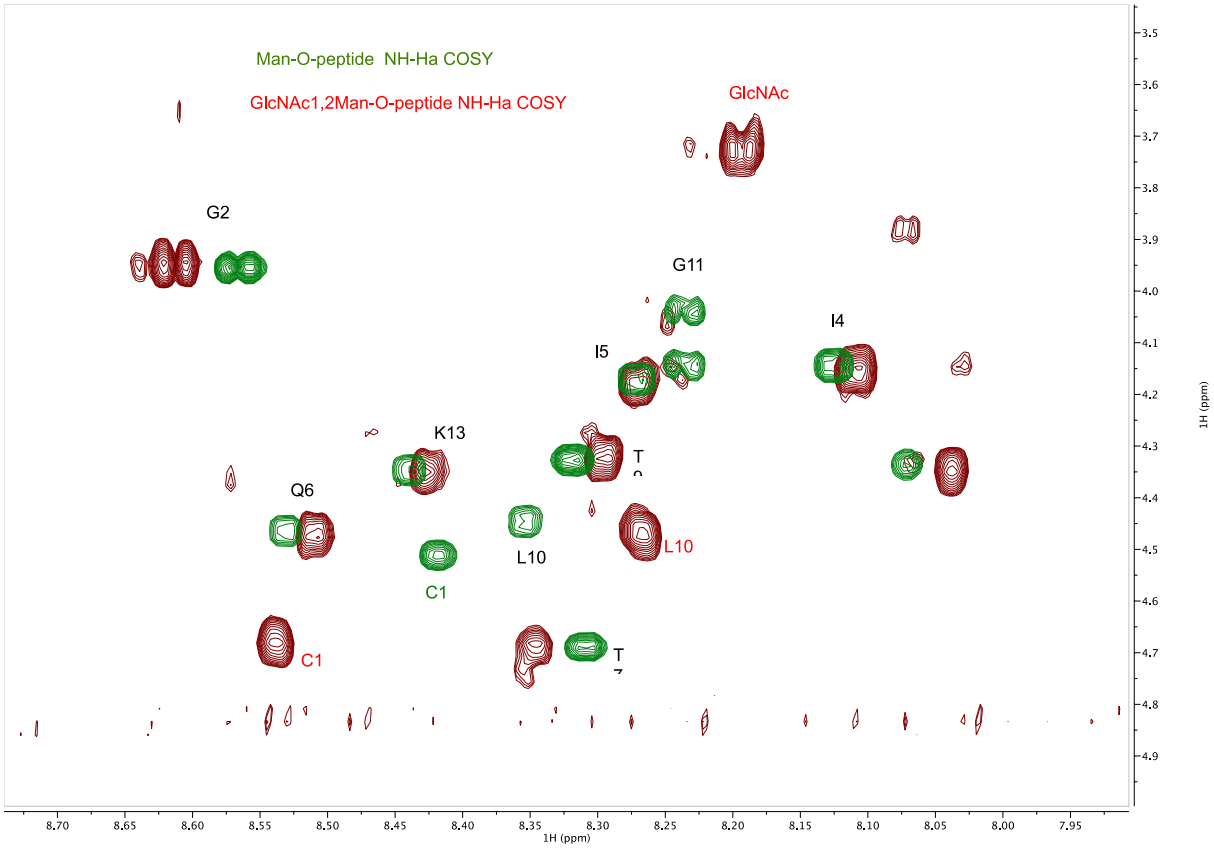


Figure 4.1. COSY NMR comparison of unextended glycopeptide vs POMGNT1 extended glycopeptide

DISCUSSION

The C-terminally truncated glycopeptide (C)GAIIQT(*O*-Man)PTLGP(K) was initially designed with the intention of using the N-terminal cysteine for conjugation to CRM via maleimide conjugation for future immunizations with the glycopeptides to make antibodies to the core *O*-mannose glycans. However, POMGNT2 did not extend this truncated glycopeptide. This finding provides further insights into POMGNT2 selectivity and the importance of the IQPTR C-terminal motif, as in ShortMan379, for POMGNT2 extension. ShortMan379, however,

is extended by POMGNT2. The N-terminal cysteine version of ShortMan379 will also be constructed and enzymatically extended to confirm that the cysteine residue does not affect POMGNT2 extension.

Previous conformational NMR studies on POMGNT1 extension investigated the effect of extension at the Thr381 residue, instead of the T379 residue, which is relevant to the M3 structures.[5] We explored the effect of POMGNT1 and POMGNT2 extension at the T379 residue. POMNGT1 extension resulted in major chemical shifts in both the cysteine residue and the leucine residue (**Figure 4.1**). However, POMGNT2 extension had no significant effect on the chemical shifts. We rationalize this observation based on the proximity of the GlcNAc extension to the glycosidic linkage. Since POMGNT2 adds GlcNAc in the β4 position, compared to the β2 position upon POMGNT1 modification (**Figure 4.2)**, the larger distance between the GlcNAc moiety in the M1 vs M3 mimetic coincides with the minimal shift differences observed in the M3 mimetic and major shifts observed in the leucine residue of the M1 mimetic. Future studies evaluating NOEs of POMGNT1 and POMGNT2 extended structures could offer more insight in regards to conformational changes in the backbone of the peptide upon addition of GlcNAc in the β2 or β4 position. An interesting observation is the chemical shift differences observed in the leucine residue between the M0 and the M1-extended structures of both in the T381 extended glycopeptide[5] and T379 extended peptide described herein. GlcNAc extension at both T379 and T381 gave rise to shift differences in L382, but not the other amino acid residues in proximity to the glycosylated site. Further NMR studies analyzing this chemical shift change in leucine, regardless of the position of glycosylation, could provide more insight to conformational changes of the peptide backbone upon POMGNT1 or POMGNT2 extension.

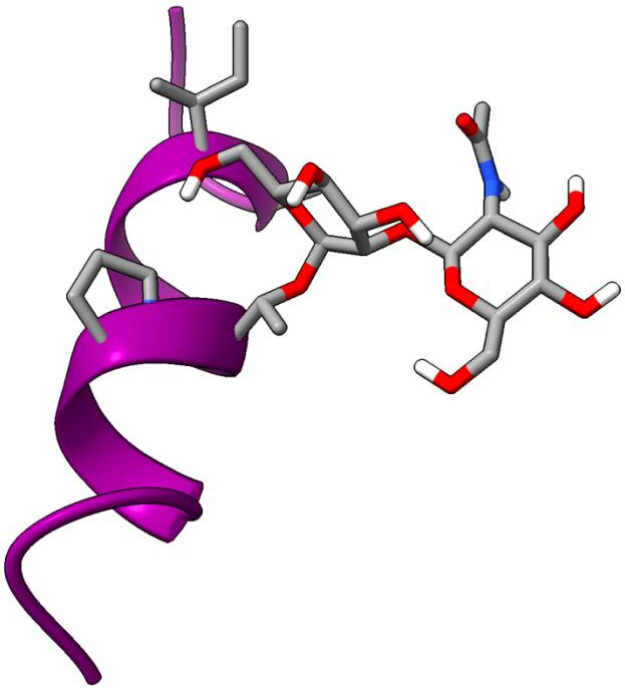


Figure 4.2. Chimera X Model of POMGNT1-extended M1 glycopeptide Mimetic

NMR conformational studies will also be carried out to compare the peptide backbones of the *C*- and *O*-glycopeptide mimetic. NMR analysis will aid in understanding the effect of this chemical difference in the glycosidic linkage. NOESY, HSQC, and HMBC experiments on samples in a mixture of 90% 5 mM acetic acid/10% D_2O will provide the backbone assignments. HMBC will also provide correlations across the peptide bond, specifically the amide proton to the α proton. The chemical shifts, NOEs, and couplings surrounding the modified Thr in the *C*-glycopeptide compared to the *O*-glycoside will provide the basis for conformational analysis.

Optimization of manual coupling reaction decreased the amount of glycoside needed by 80%. This is a highly valuable optimization considering the limited amount *C*-glycoside (Chapter 3) due challenges with scalability of the synthesis. Employing this *C*-glycoside, a *C*-

glycopeptide mimetic synthesis is underway, and the *C*-terminal region of the peptide is coupled to resin. The *C*-glycopeptide will be N-terminally extended, purified by HPLC, and tested for its ability to be extended by POMGNT1 and POMGNT2.

For future studies, it would be interesting to evaluate the ability of other enzymes in the pathway to further extend these constructs. Additionally, future studies will be aimed at determining the ability of these glycoconjugate mimetics to serve as antigens to the core *O*-mannose glycans. This would especially be valuable for detecting M3 sites (POMGNT2 modified) independently of matriglycan. Generating antibodies for *O*-mannose glycans presents a challenge since the oxygen in the glycosidic linkage is vulnerable to enzymatic degradation. However, carbon-linked (*C*-linked) glycopeptides, in which carbon replaces the normal oxygen in the glycosidic linkage, are resistant to enzymatic degradation and have the potential to provide robust immunogens to the core *O*-mannose glycans to aid in their detection.

METHODS

Glycopeptide Synthesis— The glycopeptides were prepared as *C*-terminal carboxamides and acetylated at the N-terminus to emulate the native protein environment. The first six amino acids (PTLGPK) of the *C*-terminal region were assembled on resin support using an automated microwave-assisted solid-phase peptide synthesizer (CEM Corp. Liberty microwave synthesizer) equipped with a UV detector using standard protocols in the instrument software on Rink amide resin (0.5 meq/g; 0.05 mmol, Novabiochem) via an *N*-(9-fluorenyl)methoxycarbonyl (Fmoc)-based approach with *N*,*N*-dimethylformamide (DMF) as the primary solvent. 20% 4-methylpiperidine in DMF was used for Fmoc removal. Ethyl cyano(hydroxyamino)acetate

(Oxyma Pure) in the presence of N,N'-diisopropylcarbodiimide (DIC) was used as the coupling reagent for standard amino acids. The peptide resin was removed from the synthesizer, and coupling of the glycosylated amino acid Fmoc-Thr(-D-Man(Ac)$_4$)-OH (Sussex Research) was performed manually using a CEM Corp. Discover microwave apparatus. 2-(7-Aza-1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) (38 mg, 0.1 mmol) and 1-hydroxy-7-azabenzotriazole (HOAt) (14 mg, 0.1 mmol) in the presence of DIPEA (34 uL) were the activating reagents. One coupling at at 60 °C at 1.1-fold excess of glycosylated amino acid to the resin loading was carried out for this amino acid to optimize reaction conditions and conserve reagent. A small amount of peptide was cleaved from the resin using TFA to confirm completion by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI- TOF MS). The glycopeptide resin was then treated with DMF/acetic anhydride/DIPEA (85:10:5, v/v) for 30 min to cap any the amine of any unreacted peptide. The resin as washed with DMF (5 mL x 3) and DCM (5 mL x 3) and returned to the automated synthesizer to complete assembly. After final *N*-deprotection, the glycopeptide was removed from the automated synthesizer and washed with DMF (5 mL x 3), DCM (5 mL x 3), and MeOH (5 mL x 3). The resin-bound glycopepide was then manually *N*-acetylated by treatment with DMF/acetic anhydride/DIPEA (85:10:5, v/v) for 30 min. *O*-acetyl protecting groups on the mannose sugar moiety were subsequently removed by two treatments with hydrazine/ MeOH (70:20, v/v) for an hour each. Treatment with TFA/triisopropylsilane/H_2 O (95:2.5: 2.5) for 4 h simultaneously cleaved the glycopeptide from the resin as C-terminal carboxamides and deprotected the remaining amino acid side chain protecting groups. Following filtering off the resin, the TFA solution was concentrated on a rotary evaporator to a few milliliters. The resulting

concentrate was added dropwise to cold acetonitrile in the case of (C)GAIIQT(*O*-Man)PTLGP(K) or cold ether (in the case of ShortMan379), and the crude glycopeptides precipitated. After centrifugation and removal of the ether or acetonitrile supernatant, the glycopeptides were redissolved in 50:50 water: acetonitrile with 0.1% TFA and purified via HPLC over an Ultra II 250 10.0-mm 5-m C_{18} column (Restek Corp.) with a 0.1% TFA in water, 0.1% TFA in acetonitrile solvent gradient. The purity of the glycopeptide was verified by both analytical HPLC and MALDI-TOF MS to isolate the glycopeptide in 50% yield.

REFERENCES

1. Stalnaker, S.H., Stuart, R., Wells, L. Mammalian O-mannosylation: unsolved questions of structure/function. *Curr Opin Struct Biol*. 2011;21(5):603-9.
2. Clements, R., Turk, R., Campbell, K. P., Campbell, K.W. Dystroglycan Maintains Inner Limiting Membrane Integrity to Coordinate Retinal Development. Journal of *Neuroscience*. 2017, 37 (35): 8559-8574; DOI: 10.1523/JNEUROSCI.0946-17.2017. http://www.jneurosci.org/content/37/35/8559.long.
3. Halmo, S. M., Singh, D., Patel, S.P., Wang, S., Eldin, M., Boons, G., Moremen, K. W., Live, D., Wells, L. Protein O-Linked Mannose β-1,4-N-Acetylglucosaminyl-transferase 2 (POMGNT2) is a Gatekeeper Enzyme for Functional Glycosylation of α-Dystroglycan. *Biol. Chem*. 2017, 292: 2101. DOI:10.1074/jbc.M116.764712. http://www.jbc.org/content/292/6/2101.long
4. Sheikh, M.S., Halmo, S., Wells, L. Recent advancements in understanding mammalian O-mannosylation. *Glycobiology*. 2017, 27 (9): 806-819.

5. Hinou, H.; Kikuchi, S.; Ochi, R.; Igarashi, K.; Takada, W.; Nishimura, S.-I. Synthetic Glycopeptides Reveal Specific Binding Pattern and Conformational Change at O-Mannosylated Position of α-Dystroglycan by POMGnT1 Catalyzed GlcNAc Modification. *Bioorg. Med. Chem.* **2019**, *27* (13), 2822–2831.

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

The overarching goal of the research described in this dissertation is to develop detection methods, leveraging bioorthogonal tagging and *C*-glycoside chemistry, to probe *O*-mannosylated protein sites. Prior to this research, the study of *O*-mannosylation was limited by available detection methods. The only published method of detection of M3 (POMGNT2 modified) sites was by the IIH6 antibody, which recognizes the repeating disaccharide of matriglycan, as indicated by the yellow flashlight in **Figure 5.1a**. Both the bioorthogonal SEEL tagging methodology (Chapter 2) and the creation of the *O*-mannose-based *C*-glycoside derivative (Chapter 3) offer new opportunities for probing *O*-mannosylated sites independently of matriglycan. Specifically, the bioorthogonal labeling approach using POMGNT1 and POMGNT2 described in Chapter 2 overcomes this limitation in detection by directly labelling the site of *O*-mannosylation, independently of matriglycan, as indicated by the green flashlight in **Figure 5.1a**. In addition to detection of the known α-DG M3 sites, APMAP and LAMB1 were also bioorthogonally labelled and identified as potentially novel POMGNT2-modified proteins in the *O*-mannosylation pathway using the SEEL methodology. Prior to these bioorthogonal tagging experiments, α-DG was the only known protein modified by POMGNT2. These results demonstrated the powerful utility of POMGNT1 and POMGNT2 SEEL as a means to probe *O*-mannosylated sites and identify potentially novel glycoprotein interaction. The synthesis of novel

C-Man-Thr (Chapter 3) also offers new opportunities for probing *O*-mannosylated sites by generation of antibodies to detect the core *O*-mannose glycans.

Expanding the Scope of Bioorthogonal Tagging of *O*-mannosylated Sites

While the bioorthogonal POMGNT1 and POMGNT2 SEEL enrichment studies described herein (**Chapter 2**) probed *O*-mannosylated sites independently of matriglycan, this study looked at *O*-mannosylation solely at the disaccharide level. Probing *O*-mannosylated sites in a B4GAT1 knockout cell line would provide further insight into *O*-mannosylation at the phospho-pentasaccharide level prior to modification by LARGE, which catalyzes the addition of matriglycan. Comparing the resulting enriched protein sites of POMGNT2 and B4GAT1 SEEL would provide further insight into *O*-mannosylation independently of matriglycan. These bioorthogonal tagging studies would aid in addressing the question whether other extended M3 core structures exist? Specifically, are there extended M3 core structures beyond those extended by matriglycan?

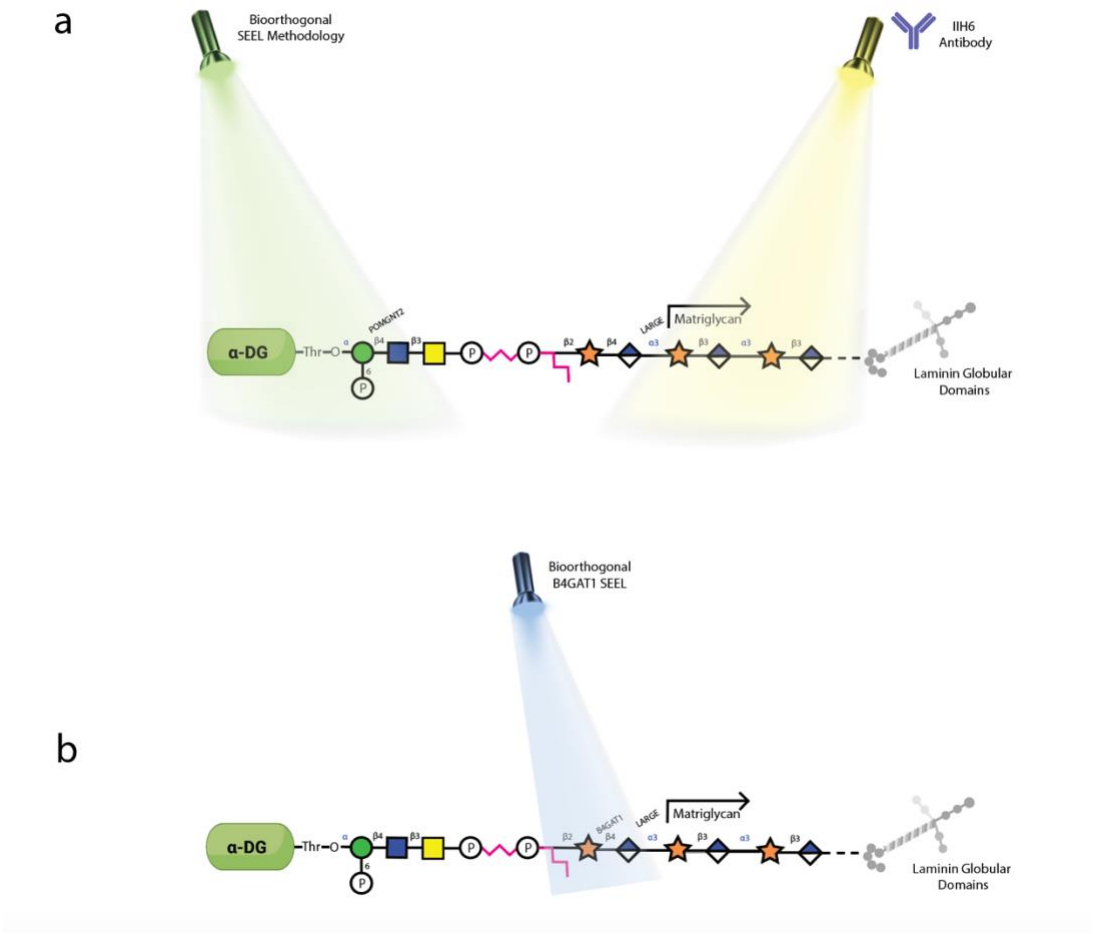
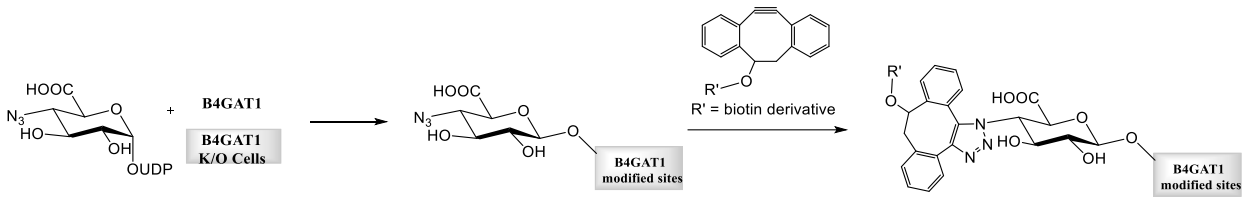


Figure 5.1. Future Studies Employing SEEL Methodology

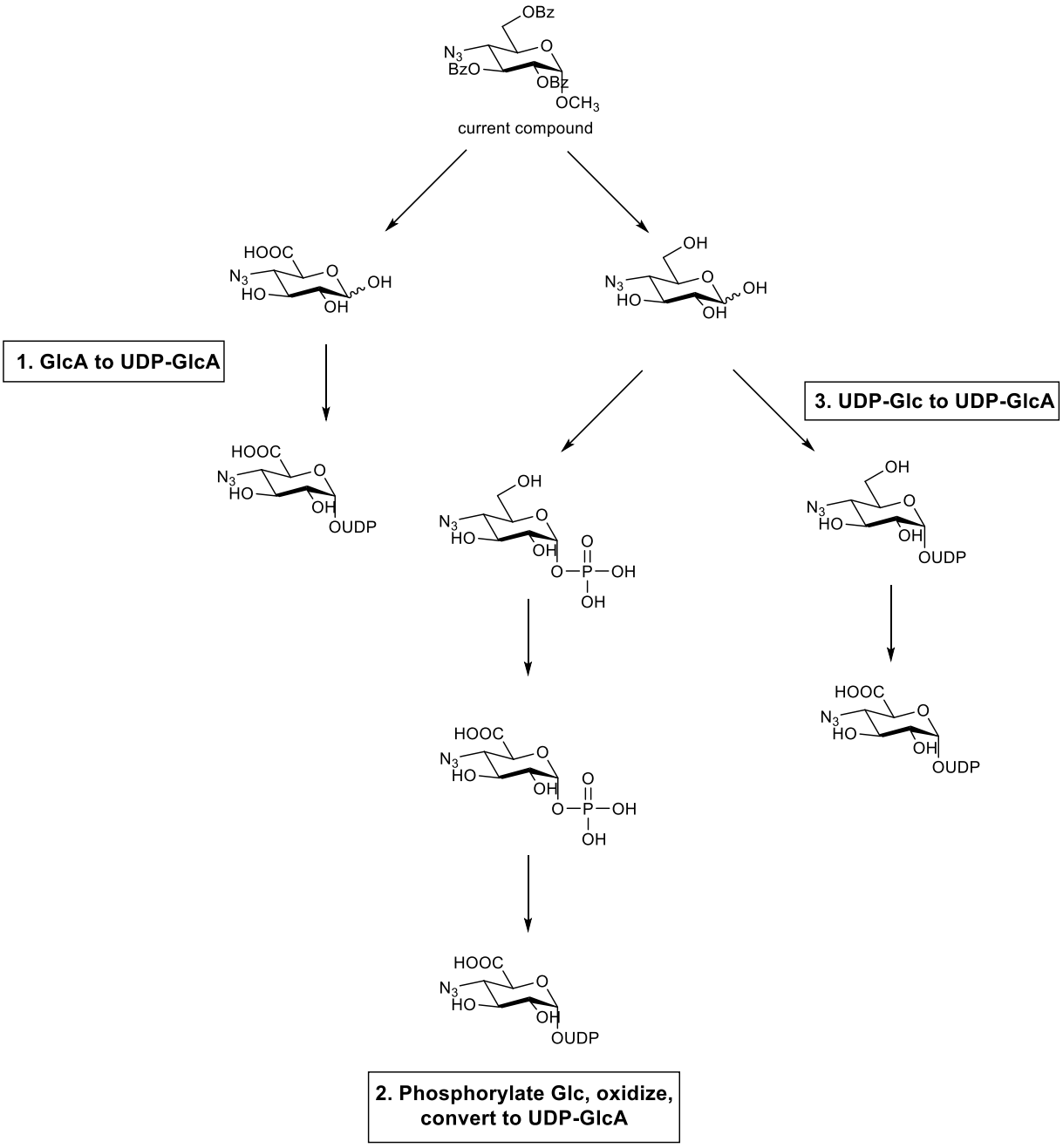
While the bioorthogonal tagging strategy described in Chapter 2 revealed potentially novel POMGNT2-exteneded *O*-mannose sites, it is unclear how far these glycoprotein sites are extended since these experiments were carried out in a POMGNT1/2 double knockout cell-line, halting *O*-mannose glycan synthesis beyond POMGNT1 or POMGNT2 extension. Instead looking at a B4GAT1 knockout cell-line and using an azide-labelled GlcA derivative could further elucidate M3 structures and clarify if any extended M3 structures exist independently of matriglycan. To probe extended M3 sites independently of LARGE, UDP-4-azido-GlcA will be

incubated with B4GAT1 and B4GAT1 knockout cells. The B4GAT1 modified sites will then display an azide functionalized group that can be bioorthogonally tagged using DIBO biotin alkyne. The biotinylated protein sites will then be isolated by neutravidin beads, which have an affinity for biotin, and analyzed by tandem mass spectrometry, as described in Chapter 2 for the SEEL approach using POMGNT1 and POMGNT2 (**Scheme 5.1**). This method will stop glycosylation one step before LARGE catalyzes the addition of matriglycan, which will aid in determining if other extended M3 core structures exist independently of LARGE.



Scheme 5.1. B4GAT1 Bioorthogonal SEEL

The azido moiety of UDP-4-Azido-GlcA will be installed synthetically. Below is a discussion of three proposed synthetic routes to create UDP-4-Azido-GlcA (**Scheme 5.2**).

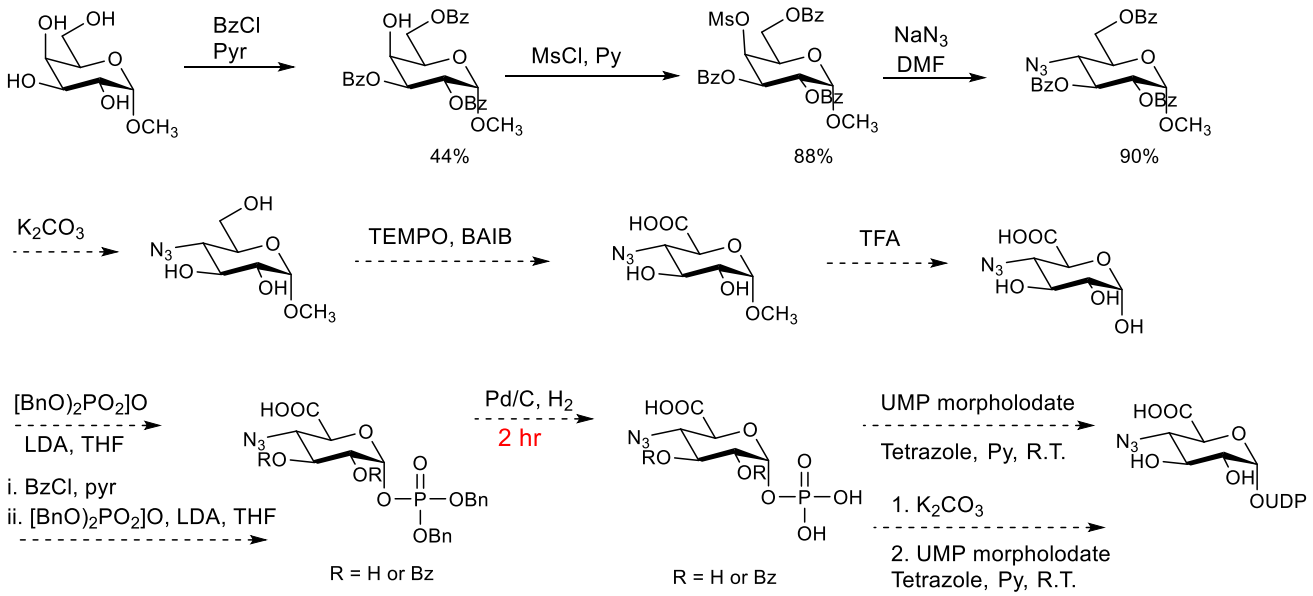


Scheme 5.2. Outline of Proposed Synthetic Routes

Proposed Synthetic Routes:

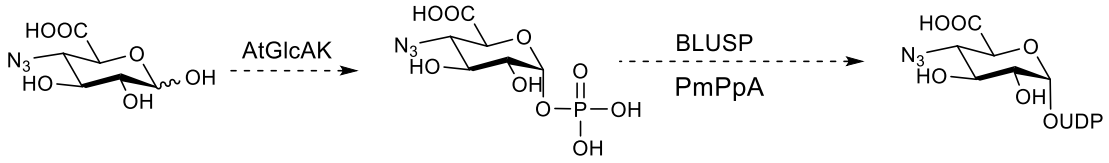
1. **GlcA to UDP-GlcA**

In the first proposed route, azido-GlcA will be synthesized and then converted to UDP-GlcA either chemically or enzymatically. Using a chemical synthetic route, 4-azido-GlcA would be phosphorylated with benzyl protecting groups on the phosphate (**Scheme 5.3**). Zhang and collogues reported the selective reduction of benzyl ethers on the phosphate, without reduction of the azide, via hydrogenation for 2 hours.[1] Protection of the 2 and 3 hydroxyls of azido-GlcA as benzoyl ester groups before phosphorylating the anomeric position could be necessary, as also indicated in Scheme 5.3.



Scheme 5.3. Chemical Conversion of Azido-GlcA to UDP-Azido-GlcA

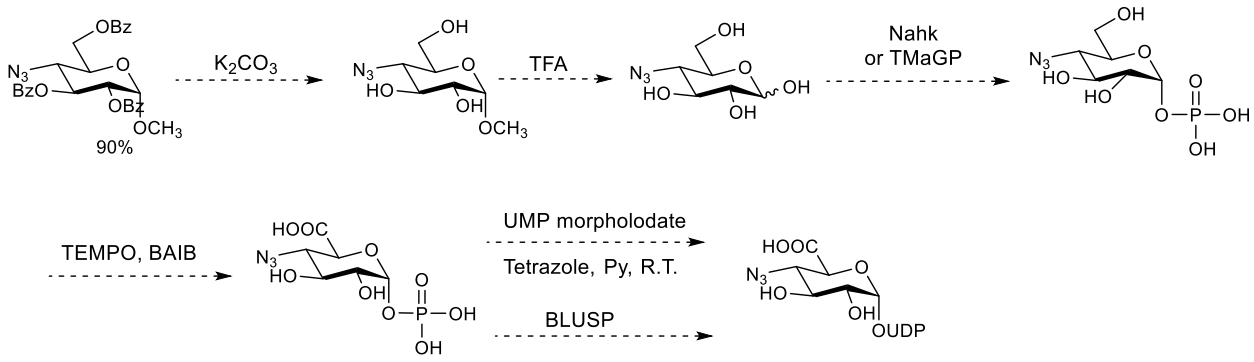
Another strategy is to use a one-pot multienzyme (OPME) strategy for chemoenzymatic conversion of azido-GlcA to UDP-azido GlcA. GlcA has been converted to UDP-GlcA using *Arabidopsis thaliana* glucuronokinase (**AtGlcAK**) and *Bifidobacterium longum* UDP-sugar pyrophosphorylase (**BLUSP**).[2] Addition of an inorganic pyrophosphatase from *Pasteurella multocida* (PmPpA) can be used to break down the pyrophosphate formed in the USP-catalyzed reaction to drive the reaction towards the formation of the UDP-uronic acid (**Scheme 5.4**). One benefit of this method is that it avoids use of expensive NAD+ cofactor to convert UDP-Glc to UDP-GlcA.



Scheme 5.4. Proposed Enzymatic Conversion of Azido-GlcA to UDP-Azido-GlcA[2]

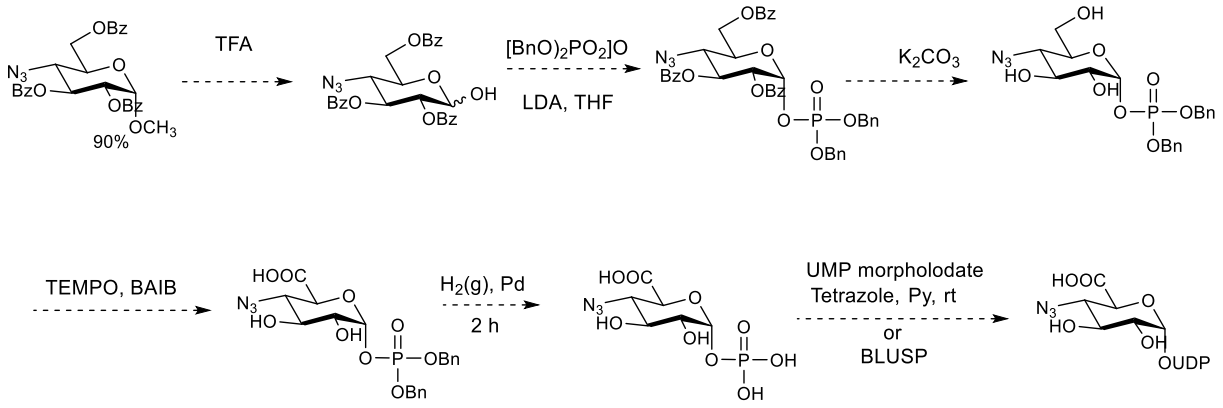
2. <u>Phosphorylate Glc, Oxidize, Convert to UDP-GlcA</u>

Another proposed route involves phosphorylating 4-azido glucose, followed by oxidation with TEMPO and BAIB to obtain 4-azido-GlcA-phosphate. Finally, 4-azido-GlcA-phosphate will be treated with a UDP-sugar pyrophoshorylase (USP) to obtain UDP-4-azido-GlcA (**Scheme 5.5**). UDP-sugar pyrophoshorylase BLUSP has been shown to convert GlcA-1P to UDP-GlcA.[2]



Scheme 5.5. Proposed Synthesis Involving Phosphorylating Glucose, Followed by Oxidation

Alternatively, benzoyl protected 4-azido-Glc will be phosphorylated chemically (as described in Scheme 5.3). The benzoyl groups will be removed, and TEMPO, BAIB will be used to oxidize to the GlcA-1P derivative (**Scheme 5.6**).

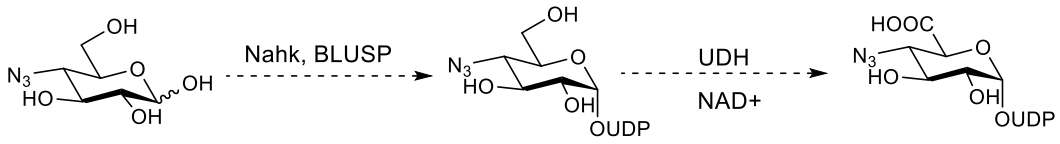


Scheme 5.6. Alternative Method Using Chemical Installation of Phosphate

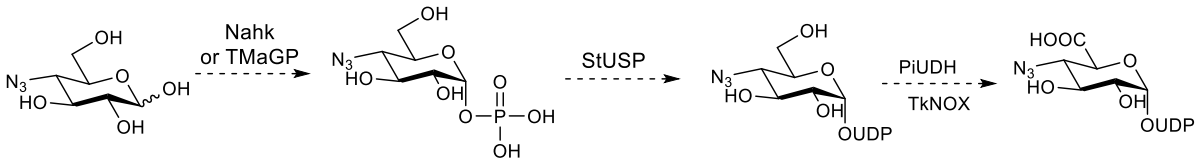
3. UDP-Glc to UDP-GlcA

In this route, 4-azido-glucose would be converted to UDP-4-azido-glucose enzymatically using NahK (N-acetylhexosamine 1-kinase cloned from *Bifidobacterium infantis* strain

ATCC15697) and BLUSP.[3] 2-Azido-Glc was reported to be substrate for these enzymes and was converted to UDP-2-azido-Glc.[3] The UDP-glucose derivative would then be converted to UDP-4 azido-GlcA using UDP-glucose-6-dehydrogenase (UDH), which we have readily available in our lab. The drawback of this method is that it requires the use of the expensive cofactor nicotinamide adenine dinucleotide (NAD^+) (**Scheme 5.7**). However, a recent paper reported a cascade system using whole cells expressing enzymes, and this route does not require exogenous addition of NAD^+ (**Scheme 5.8**).[4]



Scheme 5.7. Enzymatic Conversion of UDP-Glc to UDP-GlcA[3]



Scheme 5.8. Proposed Route Using Cascade Enzymatic Synthesis[4]

In the reported cascade synthesis, Glucose-1-P is obtained by treating starch with α-glucan phosphorylase TMαGP. Glucose-1-P is then converted to UDP-glucose using StUSP from *Sulfolobus tokodaii*. UDP-glucose is converted to UDP-GlcA using PiUDH and TkNOX, which

regenerates NAD+. The main advantage of the cascade synthesis is that it does not require the addition of exogenous NAD+.[4]

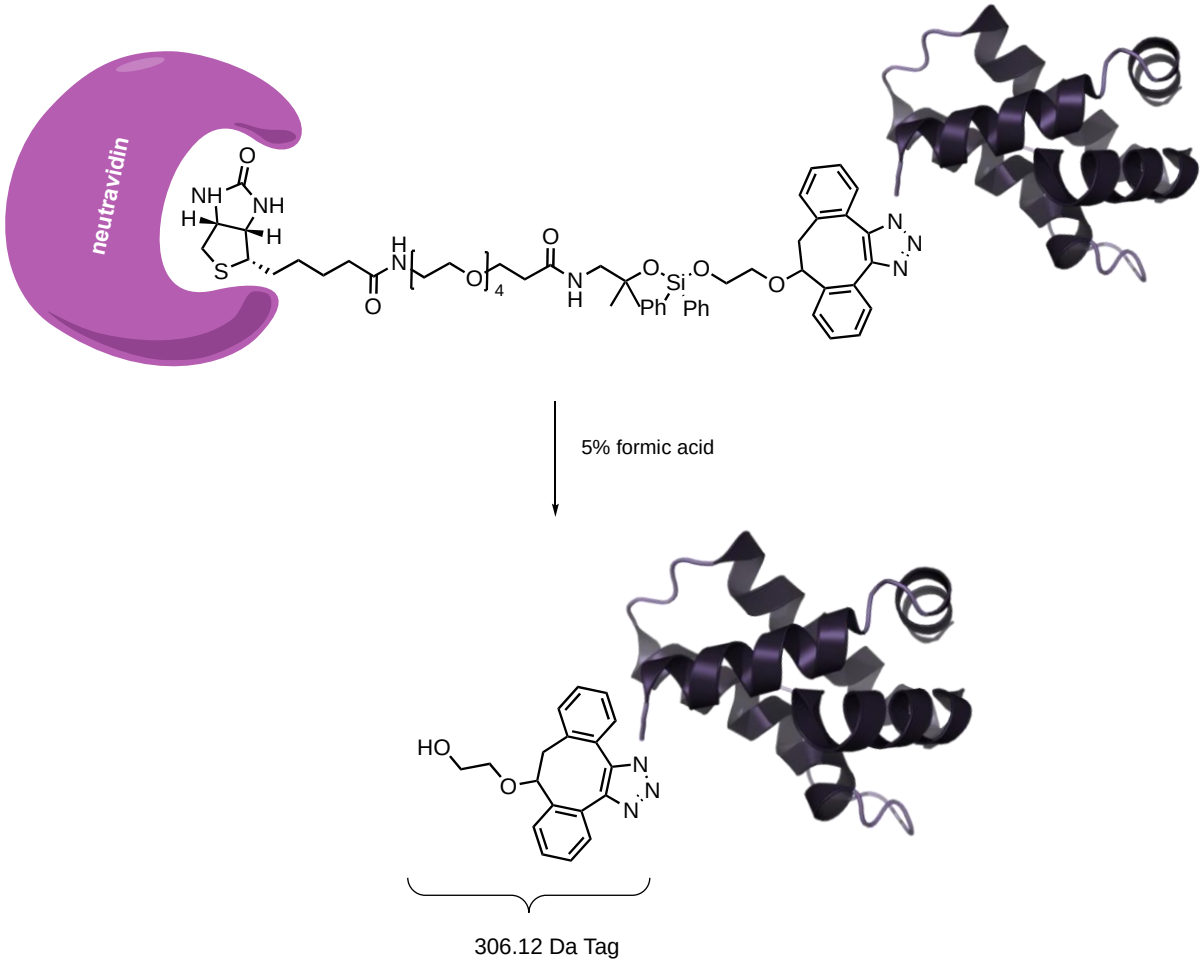
Refining Analysis of SEEL by Employing an Acid-Cleavable DIBO-Biotin Linker

The SEEL methodology developed in Chapter 2 has two primary limitations, including purification and validation of glycan sites, that could be overcome by use of a bioorthoganol tag containing an acid-cleavable linker. Specifically, one drawback of the bioorthogonal SEEL experiments described in Chapter 2 is the glycan moiety bound to the biotinylated PEG linker is not detectable in mass spectroscopy. Therefore, orthogonal validation via immunoprecipitation of enriched proteins followed by sequential mass spectroscopy analysis is required to validate the presence of *O*-mannose in the protein sites enriched via SEEL.

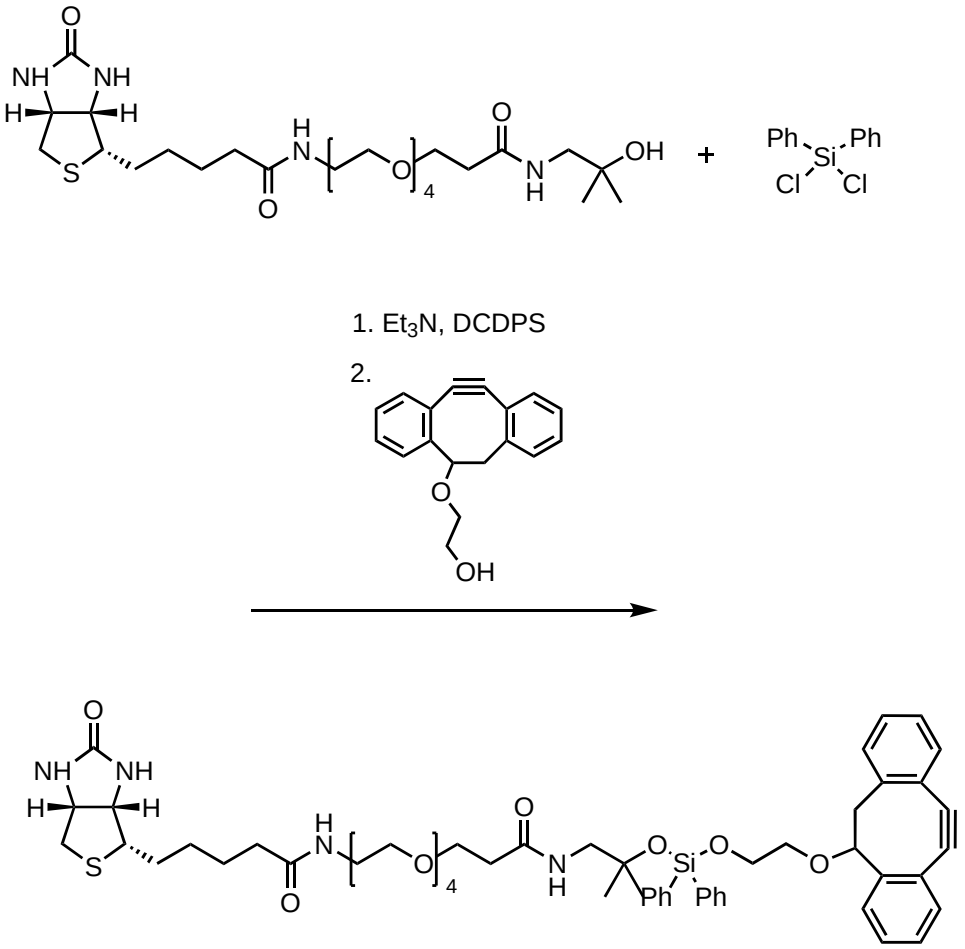
Creation of a biotinylated-DIBO molecule with an acid-cleavable (silicon-oxygen bond) linker has the potential to circumvent this limitation in detection. Cleavage of the biotin linker would enhance the downstream tandem mass-spectrometry analysis of labelled protein sites because a form of validation would be inherently built into the experiment. The sugar moiety of the enriched sites would then be readily detected via tandem mass-spectrometry. We observed the glycan in preliminary data using DADPS-biotin-alkyne, an acid-cleavable copper-catalyzed based click-chemistry probe, on an α-DG based *O*-man peptide. However, the copper-catalyzed approach is not amenable to labelling in a living cell environment. In contrast, DADPS-DIBO-

Biotin would be a valuable probe for our SEEL methodology, circumventing copper-catalysis and enhancing proteomic detection (**Scheme 5.9**).

Employing the acid-cleavable linker could also be advantageous in reducing background by refining the purification process. Cleavage would only require treatment with 5% formic acid, as opposed to boiling in detergent (5% SDS) to cleave the neutravidin-biotin bond. Non-specifically bound proteins will also elute under the harsh condition of boiling in SDS, but we would only expect the proteins attached via the silicon-oxygen linker to elute upon treatment with 5% formic acid at room temperature **(Scheme 5.9)**. For example, initial click experiments were carried out on lysed cells, but non-specific binding and background was a consistent issue. One interesting observation in this case was enrichment of protocadherins, which are currently only known to be unextended M0 structures, in both the POMGNT1 and POMGNT2 modified samples. However, background remained a consistent issue. The acid-cleavable linker could potentially overcome this limitation. The proposed synthetic strategy for the creation of DADPS-DIBO-Biotin is shown in **Scheme 5.10**.



Scheme 5.9. Acid-Cleavable DIBO-Biotin Linker



Scheme 5.10. Acid-Cleavable DIBO Synthesis Strategy

Further Exploring Laminin Subunit Beta 1 (LAMB1) Interaction with *O*-mannosylated Sites

Interestingly, laminin subunit beta-1 (LAMB1) was enriched via bioorthogonal tagging in triplicate with POMGNT2 SEEL, as described in **Chapter 2**. The SEEL experiments were carried out in POMGNT1 and POMGNT2 double knockout cells. Therefore, matriglycan is not present within this cell line, and we would not expect to enrich for laminin. This finding suggests either 1) possible *O*-mannosylation of LAMB1, or 2) the potential of an interaction between

laminin and α-DG independent of matriglycan. Consistent with this observation, Hinou and colleagues observed binding of laminin with M1 modified glycopeptides under acidic conditions via glycan array.[5]

Acidic Extracellular Microenvironments Induce Protein Aggregation

The observation of laminin binding to M1 sites under acidic conditions is interesting in the context of localized acidic extracellular microenvironments inducing protein aggregation along with cell adhesion and migration. Specifically, laminin aggregation increases with an increasing ratio of acidic to neutral phospholipids.[6] Similarly, this observation of polymerization of laminin is repeated with acidification of the solution phase. The consistent pattern across these two separate experiments led to the proposal of self-assembly of laminin evoked by a localized acidic extracellular microenvironment provided by negatively charged lipids. Negatively charged sialic acid residues and sulfate groups of glycosoaminoglycans can reduce the local charge on the cell surface by forming an electrostatic potential. Freire and colleagues hypothesize that the negative surface potential generated by charged carbohydrate species in the glycocalyx triggers laminin aggregation.[6] The extracellular pH can become acidic in various physiological and pathological contexts. For example, although normal physiological pH is 7.4, the pH of the extracellular environment of tumor cells is typically within the range 6.2-6.9. The extracellular pH in the early stages of wound healing is in the range of 5.7-6.1.[7] Furthermore, the microenvironment of a cell can be acidified via the Na^+/H^+ ion exchanger NHE1, which pulls an intracellular H^+ ion to the cell surface in exchange for an extracellular Na^+ ion. NHE1 localizes to adhesion sites and has been proposed to selectively acidify the local extracellular environment surrounding integrin receptors.[7] Provided these examples of microenvironmental pH changes

along with observed M1 glycopeptide/laminin interactions, the potential exists of a transient acidic extracellular microenvironment inducing laminin aggregation that has implications in *O*-mannosylation.

Given the enrichment for LAMB1 via SEEL coupled together with the data from Hinou and colleagues[5] supporting that M1 structures bind to laminin under localized acidic microenvironments, an expanded hypothesis regarding laminin binding is proposed. Specifically, is there laminin binding with α-DG and/or other *O*-mannosylated proteins independent of matriglycan? If this is the case, does pH affect binding as with M1 structures? To address the above questions the following experiments are proposed:

Immunoprecipitation of Laminin subunit beta 1 (LAMB1) in POMGNT1/2 Knockout Cells

To investigate whether LAMB1 interacts with α-DG independently matriglycan, LAMB1 will be immunoprecipitated in a POMGNT1/2 knockout cell-line. Since this cell-line contains no POMGNT2, matriglycan would not be present. If α-DG is immunoprecipitated along with LAMB1, this would provide further support that laminin interacts with α-DG independently of matriglycan. In-gel digestion of LAMB1, followed by tandem mass-spectrometry analysis would aid in studying this interaction. LAMB1 was enriched in triplicate with POMGNT2 SEEL in a POMGNT1/2 knockout cell-line. Immunoprecipitation of LAMB1 also co-precipitated α-DG, as expected, in both liver and brain tissue. Interrogating this interaction in the knockout cell-line will provide valuable insight into the mechanism of this interaction and could expand our current understanding.

Microarray binding studies with POMGNT2 extended *O*-man peptides of varying glycan lengths

Inspired by the observations of Hinou and colleagues in regards to M1 glycopeptide/laminin binding,[5] unextended M3 glycopeptide/laminin will be studied. To investigate whether POMGNT2-modified M3 sites bind laminin independently of LARGE/matriglycan, glycopeptides based on the M3 core *O*-mannose glycans with varying degrees of glycosylation (disaccharide, trisaccharide, pentasaccharide, etc.) before LARGE modification will be synthesized. These glycopeptides will then be immobilized on a glycan array slide and tested for laminin-lectin binding. This will be tested under both acidic pH and neutral pH. Negative controls will include naked glycopeptides with no glycan, as well as glycopeptides which would not be expected to bind laminin with the same terminal glycans as the M3 glycopeptides to ensure laminin binding is not a result of adherence to glycans. Positive controls would be M1 (POMGNT1-modified peptides), which must been shown to bind laminin under acidic pH.

NMR Conformational studies with POMGNT2 extended *O*-Man and *C*-Man glycopeptides

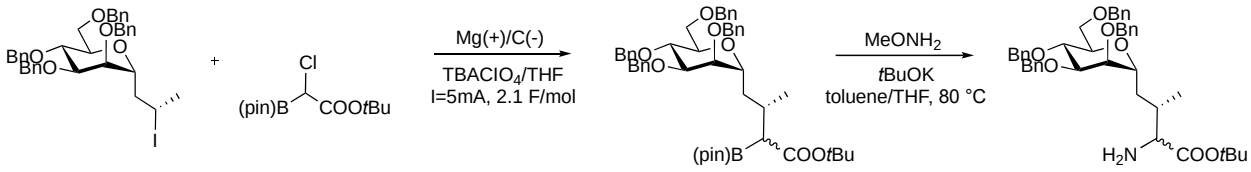
NMR studies to determine if the POMGNT2/M3 modification leads to conformational changes could provide further insights to the M3 modification. Previously reported NMR studies have demonstrated that the POMGNT1/M1 modification induces conformational change in the peptide backbone compared to that of the equivalent naked peptide (no glycan) and the M0 *O*-mannosylated peptide. Carrying out such studies comparing chemical shifts and NOE data with an M3-based glycopeptide in parallel to the naked peptide and the M0 structure could give provide more mechanistic clues to M3 structure and function.

In addition to M0 *vs.* M3 modification conformational studies, comparison of chemical shifts and NOE data of the *O*-glycopeptide *vs,* the *C*-glycopeptide (Chapters 3 and 4) would aid in determining if the *C*-glycopeptide would potentially be a suitable biomimetic for immunization to obtain antibodies for detection of core *O*-mannose sites.

Optimization of *C*-glycoside Synthesis

Cross-coupling reactions of inactivated alkyl halides remains an arduous task due to the undesired β-elimination side reaction occurring during cross coupling with both primary and secondary alkyl halides. We faced this challenge with the oxazinone chemistry during the synthesis of *C*-Man-Thr in Chapter 3. Interestingly, the O'Donnell amino acid synthesis approach exploiting the benzophenonamine of acetonitrile as the Schiff base provided a viable means to perform the alkylation, while the Schiff Base containing the *tert*-butyl ester in place of the nitrile moiety also gave the eliminated olefin as a major product. Deprotonation of enolate occurs at interface due to the biphasic nature of the O'Donnell amino acid alkylation reaction. The strongly electron withdrawing nitrile causes a positive charge build-up on the sp^2 carbon center. We therefore hypothesize that the carbon center thus becomes highly electrophilic, promoting formation of the enolate, thus providing the O'Donnell amino acid precursor as the kinetically favored product, instead of the β-eliminated product. The presence of the nitrile functional group offered enhanced chemoselectivity to allow formation of the enolate, and this observation could be useful for future $C(sp^3)$-$C(sp^3)$ cross-electrophile couplings. Expanding the reaction-substrate scope would also be of interest.

While the enhanced chemoselectivity of the cross-coupling reaction employing the benzophenonamine of acetonitrile was highly useful in the synthesis of *C*-Man-Thr, scaling this route remains a challenge due to the large number of steps and difficulty of purification of intermediates. Efforts to reduce the number of synthetic steps to create the glycoside would be useful. Zhang and colleagues recently reported cross-electrophile coupling of alkyl halides driven by electrochemistry.[8] In this methodology, a more substituted alkyl halide undergoes selective cathodic reduction, which generates a carbanion. This charged intermediate then undergoes bimolecular nucleophilic substitution with a less substituted alkyl halide.[8] Drawbacks of this method include the requirement of a glovebox for inert atmospheric conditions. However, this methodology has been extended to a wide scope of substrates and is potentially worth investigating in regards to simplifying the synthesis of *C*-glycosyl-Threonine derivatives. **Scheme 5.11** depicts a proposed synthesis employing electrolysis, followed by amination of the pinacol boranic ester, to create the C-C bond of *C*-Man-Thr. Direct electrolysis circumvents the undesired β-eliminated product formations by activating alkyl halides via an alternative mechanism compared to transition metal catalysis. Another useful investigation in regards to simplifying this synthesis would be the use of chiral phase-transfer catalysts with the O'Donnell reaction to develop a route with high stereoselectivity.[9] Success in this pursuit would likely provide improvements in yield and ease of purification, and subsequently scalability.



Scheme 5.11. Alternative *C*-glycoside Synthetic Approach Employing Electrolysis

Continued study and development of methods to probe *O*-mannosylated glycoprotein sites have the potential to identify novel targets, such as APMAP described in Chapter 2, for further investigation that could uncover new biomarkers for diagnosis and accelerate development of therapeutics for CMDs. We also envision the *C*-glycoside mimetics to be evaluated as antibodies to detect the core *O*-mannose glycans.

REFERENCES

1. Zhang, X.; Green, D. E.; Schultz, V. L.; Lin, L.; Han, X.; Wang, R.; Yaksic, A.; Kim, S. Y.; DeAngelis, P. L.; Linhardt, R. J. Synthesis of 4- Azido-N-acetylhexosamine Uridine Diphosphate Donors: Clickable Glycosaminoglycans. *J. Org. Chem*. 2017, **82**, 9910−9915. https://pubs.acs.org/doi/pdf/10.1021/acs.joc.7b01787
2. M. M. Muthana, J. Qu, M. Xue, T. Klyuchnik, A. Siu, Y. Li, L. Zhang, H. Yu, L. Li, P. G. Wang and X. Chen. Improved one-pot multienzyme (OPME) systems for synthesizing UDP-uronic acids and glucuronides. *Chem. Commun*., 2015, **51**, 4595–4598. https://pubs.rsc.org/en/content/articlelanding/2015/CC/C4CC10306H#!divCitation

3. M. M. Muthana, J. Qu, Y. Li, L. Zhang, H. Yu, L. Ding, H. Malekan and X. Chen. Efficient one-pot multienzyme synthesis of UDP-sugars using a promiscuous UDP-sugar pyrophosphorylase from Bifidobacterium longum (BLUSP). *Chem. Commun*., 2012, **48**, 2728–2730. https://pubs.rsc.org/en/content/articlelanding/2012/CC/c2cc17577k#!divAbstract
4. Meng, DH., Du, RR., Chen, LZ. *et al.* Cascade synthesis of uridine-5′-diphosphate glucuronic acid by coupling multiple whole cells expressing hyperthermophilic enzymes. *Microb Cell Fact* **18,** 118 (2019). https://doi.org/10.1186/s12934-019-1168-z
5. Hinou, H., Kikuchi, S., Ochi, R., Igarashi, K., Takada, W., Nishimura, S.I. Synthetic glycopeptides reveal specific binding pattern and conformational change at O-mannosylated position of α-dystroglycan by POMGnT1 catalyzed GlcNAc modification. *Bioorg Med Chem*. **27**(13):2822-2831 (2019). doi:10.1016/j.bmc.2019.05.008.
6. Freire, E., Coelho-Sampaio, T. Self-assembly of laminin induced by acidic pH. *J Biol Chem*. **275** (2):817-22 (2000). doi: 10.1074/jbc.275.2.817. PMID: 10625612.
7. Paradise, R. K., Lauffenburger, D. A., & Van Vliet, K. J. Acidic extracellular pH promotes activation of integrin α(v)β(3). *PloS one*, **6**(1), e15746 (2011). https://doi.org/10.1371/journal.pone.0015746.
8. Zhang, W., Lu, L., Zhang, W. *et al.* Electrochemically driven cross-electrophile coupling of alkyl halides. *Nature* **604**, 292–297 (2022). https://doi.org/10.1038/s41586-022-04540-4.

9. O'Donnell M. J. The enantioselective synthesis of alpha-amino acids by phase-transfer catalysis with achiral Schiff base esters. *Accounts of chemical research*, **37**(8), 506–517 (2004). https://doi.org/10.1021/ar0300625