UNIVERSITY OF CALIFORNIA SAN DIEGO

Synthesis and application of glycopolymers to probe cell surface interactions in the presence of a mucinous glycocalyx model

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Chemistry

by

Daniel Jason Honigfort

Committee in charge:

Professor Kamil Godula, Chair
Professor Ada Almutairi
Professor Nathan Gianneschi
Professor Simpson Joseph
Professor Joe O'Connor

2021

The Dissertation of Daniel Jason Honigfort is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

DEDICATION

To my family who has been so supportive, especially when it was hard. Without them none of this is possible.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Boc	*tert*-butyloxycarbonyl
^{13}C NMR	Carbon Nuclear Magnetic Resonance
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DCM	Dichloromethane
DIBAL	Diisobutylaluminum Hydride
DIPEA	Diisopropylethylamine
DMAP	4-dimethylaminopyridine
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DLS	Dynamic Light Scattering
EA	Ethyl Acetate
EDC	1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide
EtOH	Ethanol
FLIM	Fluorescence Lifetime Imaging Microscopy
FRET	Förster Resonance Energy Transfer
GAG	Glycosaminoglycan

GP	GlycoPolymer
GPC	Gel Permeation Chromatography
HCl	Hydrochloric Acid
Hex	Hexanes
^{1}H NMR	Proton Nuclear Magnetic Resonance
HPLC	High Performance Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
IAV	Influenza A Virus
IPA	2-Propanol
IR	Infrared Spectroscopy
K_2CO_3	Potassium Carbonate
LiCl	Lithium Chloride
MeOH	Methanol
4-MU NANA	2'-(4-methylumbelliferyl)-α-D-N-acetylneuraminic acid
NaOH	Sodium Hydroxide
NH_4OH	Ammonium Hydroxide
PBS	Phosphate Buffered Saline

RBC	Red Blood Cell
r.t.	Room Temperature
TBAB	Tetrabutylammonium bromide
TBAN	Tetrabutylammonium Azide
t-BuOH	*tert*-butyl alcohol
TBSCl	*tert*-butyldimethylsilyl chloride
TFA	Trifluoroacetic Acid
TEA	Triethylamine
TLC	Thin Layer Chromatography
UV/Vis	UltraViolet/Visible light

LIST OF FIGURES

LIST OF SCHEMES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Dr. Kamil Godula for accepting me into his lab, and for support, funding, and guidance given along the way. I would like to thank Dr. Mia Huang for her wisdom and guidance, as well as Dr. Meghan Altman for her help developing RBC and IAV assays. I would like to extend special thanks to Dr Steven Verespy and especially Dr. Greg Trieger for their assistance with my job search. I would like to acknowledge Michelle Zhang for her support for my projects as an undergraduate, and for taking charge of her own project spinoff as a masters student. I would also like to thank all of my friends in the Godula lab for their support and encouragement throughout grad school. Good friends make grad school much more bearable!

I would like to thank Kiana Miyamoto for her support and patience during my PhD, and for pushing me to finish. Also thanks to Kona for being the best boy, and forcing me to get out of the house to take walks during quarantine and writing.

Finally, I would like to thank my family for their love and encouragement throughout my life. I would not be where I am today without their belief in me, and their pushing me to find my passion. Your support means the world to me.

Chapter 1, in full, is a reprint of the material as it appears: Huang, M.L, Honigfort, D.; Godula, K. Carbohydrates - The Glycocalyx, and its Biological Roles. Cell Physiology Sourcebook (Elsevier 2018, Edna Kaneshiro (Eds), in press). The dissertation author is a co-author of this manuscript.

Chapter 2, in full, is a reprint of the material as it appears: Honigfort, D; Zhang M. H; Verespy, S; and Godula, K. Engineering of spectator glycocalyx structures to evaluate

molecular interactions at crowded cellular boundaries. *Faraday Discussion*, 2019, DOI 10.1039/C9FD00024K. The dissertation author is the primary co-author of this manuscript.

Chapter 3, in full, is work in preparation for publication. The manuscript has been deposited in its current form on Bioarxiv and can be accessed at https://doi.org/10.1101/2021.05.07.443169. This work was contributed to by Meghan Altman, Pascal Gagneux, and Kamil Godula. Daniel Honigfort is the primary author of this material.

Chapter 4, in part, is unpublished work, and was contributed to by Kamil Godula, and Michelle Zhang through synthesis and scale up of reaction intermediates and products including propargylated glycans and the amino-oxy to alkyne linker. Daniel Honigfort is the primary author of this material.

VITA

2012 The University of Notre Dame, Notre Dame, IN

B.S. in Chemistry. Research Advisor: Richard Taylor

2014 The University of California San Diego, La Jolla, CA

M.S. in Chemistry. Research Advisor: Kamil Godula

2021 The University of California San Diego, La Jolla, CA

Ph.D. in Chemistry. Research Advisor: Kamil Godula

PUBLICATIONS

1. Zhang, M. H; Honigfort, D; Godula, K. Photopatterning the glycocalyx for spatial and temporal control over glycan interactions. *In Preparation*

2. Honigfort, D; Altman, M; Gagneux, P; and Godula, K. Glycocalyx crowding with synthetic mucin mimetics strengthens interactions between soluble and virus-associated lectins and cell surface glycan receptors. *Submitted*. Available on Bioe: https://doi.org/10.1101/2021.05.07.443169

3. Naticchia, M; Laubach, L; Honigfort, D; Purcell, S; Godula, K. Spatially controlled glycocalyx engineering for growth factor patterning in embryoid bodies. *Biomaterial Sci.*, 2021, DOI 10.1039/d0bm01434f.

4. Honigfort, D; Zhang M. H; Verespy, S; and Godula, K. Engineering of spectator glycocalyx structures to evaluate molecular interactions at crowded cellular boundaries. *Faraday Discussion*, 2019, DOI 10.1039/C9FD00024K

5. Huang, M.L, Honigfort, D.; Godula, K. Carbohydrates - The Glycocalyx, and its Biological Roles. Cell Physiology Sourcebook (Elsevier 2018, Edna Kaneshiro (Eds), in press)

FIELDS OF STUDY

Major Field: Polymer Chemistry, Chemical Biology

Studies in Polymer Chemistry and Assay Development
Professor Kamil Godula

ABSTRACT OF THE DISSERTATION

Synthesis and application of glycopolymers to probe cell surface interactions in the presence of a mucinous glycocalyx

by

Daniel Jason Honigfort

Doctor of Philosophy in Chemistry

University of California San Diego, 2021

Professor Kamil Godula, Chair

The mucosal glycocalyx dictates cellular interactions with extracellular factors by providing a physical protective barrier around the host cell. This barrier, due to its heterogenous structure and prevalent non-templated glycosylation, is difficult to study through traditional biochemical means. Of particular challenge is the subclass of cell surface glycoconjugates known as mucins, which impart physiological function through their glycans

and cytosolic signaling components but are also unique due to their massive size and extended architecture. The steric and exclusionary roles this enables are poorly understood due to a lack of chemical and biological tools with which to study these types of non-interacting "spectator" interactions. Especially relevant in light of the current COVID-19 pandemic are interactions between viral and epithelial surfaces mediated by the host's defensive mucosal barrier. Though much attention has been paid to the viral subversion of innate defenses and co-opting of native sialyl receptors to aid infection, the role of the physical structure and bulk of mucins at the cellular interface is largely unexplored. In this thesis, I will describe the synthesis of PEG-based glycopolymers, appended with a diversity of pendant glycan moieties that approximate the structural roles of mucins, as a chemical biology tool to enable precise molecular control of these physical cell surface properties. Through incorporation onto red blood cells (RBCs), these glycopolymers generated an artificial mucosal barrier model system that reduced the association of multivalent lectins (SNA, ConA) with their underlying cell surface glycan receptors, and hinted at differences in how glycan binding proteins that target disparate regions of the glycocalyx are impacted by steric crowding. Next, I will show that increasing the density and extension of this artificial barrier reduced the association of both SNA and Influenza A virus to the cell surface, but counterintuitively led to enhanced receptor-ligand stability by inducing formation of high avidity Velcro-like patches. This approach provided new insights into the role of membrane anchored mucins in infection, and a new example of viral mechanisms to overcome host defenses. Finally, I will discuss the development of complementary sidechain modification chemistries to append larger and more complex glycans to these glycopolymers in an effort to further expand their functional utility.

Chapter 1. The Glycocalyx and its biological roles

Section 1.1 An introduction to Glycans

Carbohydrates, or glycans, are a large class of essential biopolymers intimately linked to cellular physiology. Although the public mostly knows carbohydrates as macronutrient energy sources, they are also intriguing cellular messengers, molecules that control and report on a cell's physiological state. Their central importance to life is reflected by the fact that evolution has produced no single organism devoid of glycans. Glycans are prominently found conjugated to lipids or proteins populating the outer leaflet of cell membranes or integrated as components of bacterial cell walls (**Fig. 1.1**). The entire collection of cell surface glycoconjugates is typically referred to as the glycocalyx. The chemical diversity of glycan structures found in the glycocalyx is rather vast, with additional complexity arising from the spatial organization of glycans on glycoconjugate scaffolds. For instance, glycans can be presented in close proximity to the cell membrane when attached to small glycolipids and glycopeptides, or they can extend hundreds of nanometers away from the cell surface when displayed on large proteins, such as mucins and proteoglycans. This chapter focuses on the biology of the glycocalyx, with a special emphasis on their macromolecular mucin glycoprotein and proteoglycan components.

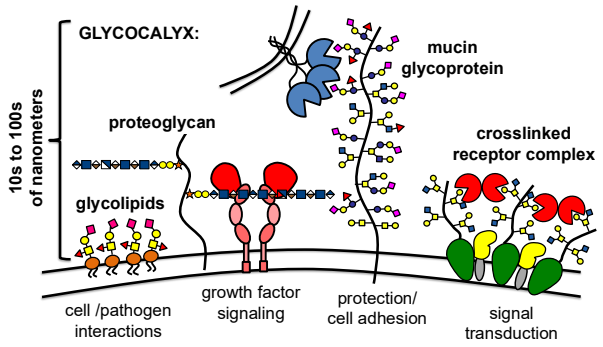


Figure 1.1. The nanoscale architecture of the glycocalyx. Cell surface glycans and glycoconjugates that decorate cell membranes and bacterial cell walls serve important physiological functions. The glycocalyx can extend tens to hundreds of nanometers away from the cellular membrane. (*Reprinted with permission from Huang, M. L. and Godula, K. (2016). Nanoscale materials for probing the biological functions of the glycocalyx. Glycobiology. 26, 797-803. Copyright 2016 Oxford University Press.)*

IA. Structural complexity of glycans

Glycans are composed of monosaccharide building blocks with the general formula $C_n(H_2O)_n$. (**Fig. 1.2**) One must also take note of the number of building blocks for glycans, in contrast to other types of biopolymer molecules. Only twenty-two amino acid building blocks give rise to proteins, whereas DNA and RNA are made from five nucleobases. On the other hand, glycans can be made from upwards of 50 different monosaccharides, a number that greatly outnumbers the English alphabet; thus significant recent efforts have been made to standardize how glycans should be catalogued and represented graphically. The glycoscience community is coalescing around an agreement that each monosaccharide should be represented by a specific shape and color. For example, glucose is represented as a blue circle, whereas a yellow circle indicates galactose. The study of glycans is complex; it begins with determining the chemical structures of each monosaccharide building bock and extends toward the understanding of the biosynthesis and function of large macromolecular glycoconjugates. The entirety of glycans, free or in the form of glycoconjugates, within an organism is referred to as the Glycome. The Glycome is involved in virtually every aspect of

cellular biology and it is fascinating to think of how it works in concert with the Proteome and the Genome to control the physiological state of each cell.

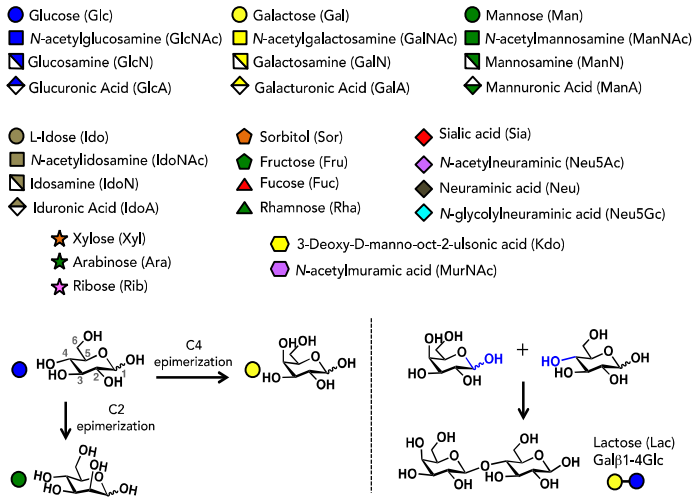


Figure 1.2. Common monosaccharide building blocks and their symbol representation. The simple monosaccharide glucose can generate epimers at carbon-2 to yield mannose (Man) or at carbon-4 to generate galactose (Gal). Glycans contain carbonyl functionality at the anomeric carbon-1 position to form covalent glycosidic linkages with hydroxyl groups of another monosaccharide, to yield a disaccharide.

Before we turn our attention to some of the complex glycan structures dominating the cellular glycocalyx, we should consider perhaps one of the simpler monosaccharides, glucose. Glucose is the most widely used energy source, which animals derive by the hydrolysis of carbohydrates such as lactose from milk or sucrose from sugar cane. The released glucose can be directly metabolized or it can be stored in the form of glycogen, a protein decorated with densely branched glucose polysaccharides. In addition to serving as am energy molecule, glucose is also a key building block used for the construction of more complex glycans. It is a hexose with six carbon atoms on its backbone, and four out of the six are chiral centers. (**Fig. 1.2**) Because the chemical groups around each chiral center can be arranged in two different configurations, glucose can exist as (4^2) 16 aldohexose stereoisomers. Interestingly, D-glucose is a naturally occurring isomer, whereas L-glucose is not (but other L-monosaccharides can

be). Epimerization, a change in the stereochemical configuration of a single carbon atom of a sugar molecule, also leads to a completely different monosaccharide. Glucose can epimerize into two other glycans, epimerization of carbon 2 yields mannose (Man), whereas epimerization of carbon 4 generates galactose (Gal). Glucose is a vital component of human life, whereas Man is not considered an essential nutrient. In water, a monosaccharide exists in equilibrium as the open and closed form.

The individual monosaccharides containing carbonyl functionality at the anomeric carbon-1 position and several hydroxyl groups are strung together into larger structures through covalent glycosidic linkages formed between a hydroxyl group on one monosaccharide and the carbonyl group of another. Since most monosaccharides contain multiple hydroxyl groups, branched structures can be obtained, which distinguishes glycans from other biopolymers, such as DNA, RNA or proteins, which only exist in the form of linear chains. Each glycosidic linkage can also exist in α or ß configuration. Thus, the disaccharide lactose (a milk glycan) is properly notated as ß-D-galactopyranosyl-(1–>4)-D-glucopyranose or Gal-ß-1-4Glc.

IB. Glycosylation and glycan biosynthesis

The process of stringing glycans together or attaching them to other biomolecules is called "glycosylation." In eukaryotes, glycans are biosynthesized in the endoplasmic reticulum (ER) and Golgi apparatus as post-translational modifications on peptides, proteins, or lipids to form glycoconjugates. Various glycosyltransferases and glycosidase enzymes work in concert to add or remove individual monosaccharides to a growing chain.

At any given point in time, the abundance and diversity of glycans is a tightly regulated process. (Fig. 1.3) Glycosylation has been known to be a highly dynamic process that changes with respect to cell cycle, developmental stage, and disease states. A number of factors can dictate glycan levels, including the (1) metabolic flux of monosaccharide building blocks and nucleotide sugars, as well as the (2) expression levels of glycosyltransferases and glycosidases. For glycans present on the glycocalyx, (3) extracellular glycan remodeling, as well as (4) internalization and degradation, can also affect glycan abundance at the cell surface. Glycosylation is extremely sensitive to intracellular pools of monosaccharide building blocks and nucleotide sugars (e.g. UDP-GlcNAc, CMP-sialic acid), the activated forms of monosaccharides that act as donors in glycosylation reactions. Some monosaccharides can be directly transported into the cellular membrane by transporters, whereas others can be biosynthesized from other precursors. For example, the transmembrane glucose transporter brings extracellular glucose and mannose into the cell, and sialic acid can be biosynthesized from glucose. Catabolism of glycoproteins can also feed monosaccharides into the cellular pool. Upon synthesis in the cytoplasm or nucleus, the nucleotide sugars are transferred to the Golgi or ER lumen by nucleotide sugar transporters. There, the levels of glycosyltransferases and glycosidases affect the extent of glycan biosynthesis. Some cancers exhibit aberrant glycosylation due to altered levels of glycosylation. Upon completion of biosynthesis, cell surface glycans are exported to the cell surface, where primary glycoforms can be altered by extracellular glycan remodeling enzymes, such as glycosidases and sulfatases. Sialidase enzymes can cleave terminal sialic acids off glycan chains to expose underlying glycans, whereas sulfatases can remove sulfate groups already present on the glycan chain. Cell surface glycans can be removed from the glycocalyx by clathrin-mediated or endocytic mechanisms. These can also be recycled back to

the cell surface or targeted for degradation into the lysosomes. Although the enzymes responsible for elongating and decorating the glycan chains are encoded by the genome, there is no template for the exact manner, pattern, and length, in which an individual glycan is constructed. Moreover, a single glycosylation site in a protein is not always associated with a unique glycan structure, providing considerable heterogeneity in glycoconjugates and their glycoforms.

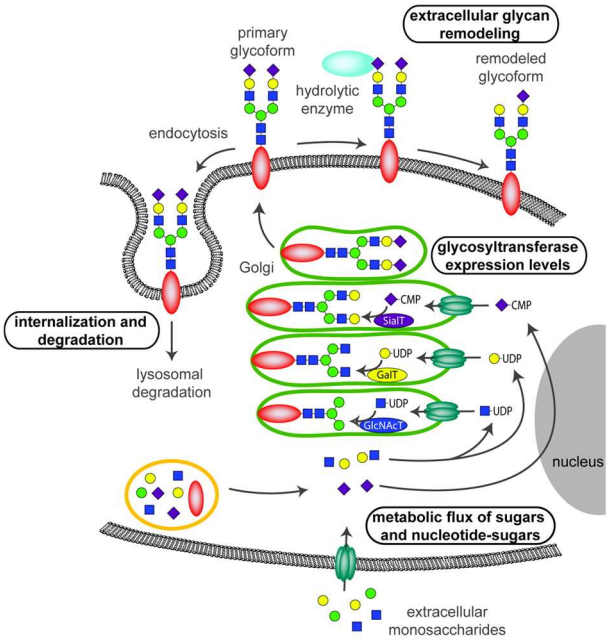


Figure 1.3. Glycan abundance is dictated by multiple steps in its biosynthesis. Intracellular concentrations of monosaccharides and nucleotide sugar precursors can influence the extent of biosynthesis, as well as glycosyltransferases expression levels. For cell surface glycans that are exported to the cell surface upon completion of biosynthesis, extracellular glycan remodeling enzymes, such as sialidases or sulfatases, can further alter the presented glycans. Internalization and degradation mechanisms also reduce the expression of cell surface glycans.

Glycans can be classified based on their connectivity. Glycans attached to lipids and proteins via a nitrogen atom are referred to as *N*-glycans, whereas those that are attached through an oxygen atom are called *O*-linked glycans. The *O*-linked glycans can be further divided into two distinct classes: branched mucin type glycans and linear glycosaminoglycan

polysaccharides of proteoglycans. While the two families of mucin and proteoglycan-type *O*-glycans will be discussed in considerable detail later in this chapter, a short introduction and comparison of the two uniquely distinct biosynthetic pathways to *N*-linked and *O*-linked glycans is warranted.

IC. *N*-linked glycans.

The assembly of *N*-linked glycans occurs co-translationally, or simultaneous with protein translation **(Fig. 1.4).** This timing creates a logistical challenge, where the growing peptide chain arising from a ribosome must meet the ER- or Golgi-resident enzymes responsible for attaching glycans. *N*-glycosylation begins in the cytoplasmic side of the ER where a dolichol phosphate lipid-linked GlcNac2-Man5 oligosaccharide is generated. This lipid-linked precursor is flipped into the ER lumen, where glycosyltransferases further elaborate the glycan chain into GlcNac2Man9Glc3. This completed glycan is then transferred to an asparagine residue of a nascent polypeptide chain by an oligosaccharyltransferase (OST) enzyme. The resulting glycoconjugate then undergoes extensive processing, where the three terminal glucose residues are removed by ß-glucosidases as part of a quality control check for proper folding of the protein. Correctly folded proteins are exported to the Golgi apparatus, whereas improperly folded proteins bind to the chaperone calnexin/calreticulin protein to be re-folded. At the Golgi, mannosidases edit some of the Man residues and glycosyltransferases further add various monosaccharides. It is at this step that three major types of *N*-glycans are generated: high-mannose, hybrid, or complex.

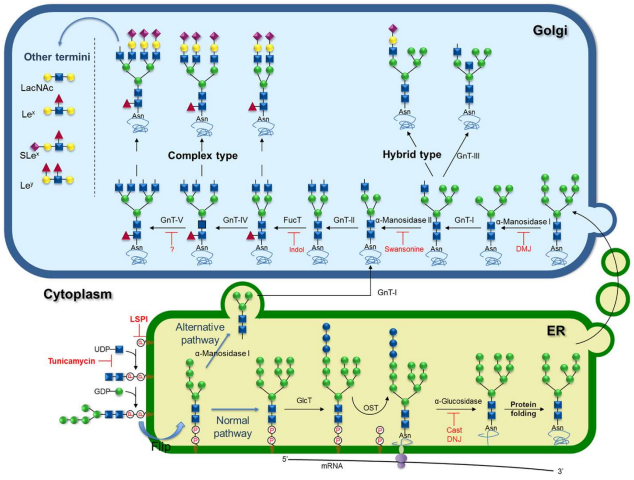


Figure 1.4. Biosynthesis of N-linked glycans and known chemical inhibitors. A dolichol phosphate lipid is initially glycosylated at the cytoplasmic side of the ER (bottom left). Upon flipping and assembly of the Glc2Man9Glc3 precursor in the ER lumen, the glycan is transferred to the side chain of an asparagine residue. Glucosidases trim the terminal Glc residues, and after checking for proper protein folding, the glycoconjugate is transported to the Golgi, where further processing occurs. This process results in complex type and hybrid type glycans. Molecules that disrupt this pathway can exert anticancer activity. Used under the Creative Commons Attribution License. Vasconcelos-dos-Santos, A.; Oliveira I. A.; Lucena, M. C.; Mantuano, N. R.; Whelan, S. A.; Dias, W. B.; Todeschini, A. R. Biosynthetic machinery involved in aberrant glycosylation: promising targets for developing of drugs against cancer. *Front. Oncol.* (2015) 5:138.

N-glycans play important roles in cellular physiology and immunology. Inhibition of N-glycan biosynthesis is costly for cellular function, and molecules that disrupt *N*-glycan biosynthesis have been implicated for chemotherapeutic purposes. For example, the natural product tunicamycin, which inhibits GlcNAc phosphotransferase to prevent the formation of GlcNAc-PP-dolichol, has been shown to induce the unfolded protein response and arrest cells in the G1 phase. *N*-linked glycans also play important roles in regulating the immune system, and *N*-glycosylation of recombinant therapeutic proteins, such as antibodies, have recently come into focus as problems in biotechnology. In order to produce glycoprotein antibodies, eukaryotic expression systems must be used, as prokaryotes cannot efficiently generate post-translational modifications. However, these eukaryotic expression systems, such as the yeast

Saccharomyces cerevisiae, generate different glycosylation patterns compared to humans (e.g. high mannose) that can cause severe undesired immunogenic reactions.

ID. *O*-linked glycosylation

In this short section, we will tackle *O*-linked glycans that only contain one or a few glycans. The first attachment of an *O*-linked monosaccharide to serine/threonine protein residues occurs in the Golgi through the action of specific glycosyltransferases. GalNAc is transferred by UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase (ppGalNAcT), whereas fucose is transferred by fucosyltransferases. Further extension of these monosaccharides can occur to generate antigenic markers. For example, the simple GalNAcαSer/Thr modification is known as the Tn antigen, and is commonly found on the cancer cell glycocalyx. Subsequent sialylation or galactosylation generates other types of antigens, but further elongation generates *O*-glycan core structures found on the surface of healthy cells. *O*-linked glycans also carry the basis for ABO blood group determinants. The O blood group is characterized by a Fucα1-2Gal modification, the A group by a GalNAcα1-3(Fucα1-2)Gal, and the B blood group by a Galα1-3(Fucα1-2)Gal modification.

IE. *O*-GlcNAc

Whereas most mucin and proteoglycan-type O-linked glycans are heterogeneous and complex, even very simple glycosylation modifications can produce significant functional effects. For example, the simple GlcNAc modification of proteins serves an important regulatory function in signaling and gene expression, similar to protein phosphorylation. Indeed, some competitive phosphorylation or O-GlcNAc modification of serine/threonine residues has been observed. Two enzymes, O-GlcNAc Transferase (OGT) and O-GlcNAcase

(OGA) control the installation and removal of α-linked GlcNAc to serine/threonine. The *O*-GlcNAc modification is highly dynamic, and is known to modify histone proteins to regulate transcription, ubiquitination, and other processes critical to cellular homeostasis. *O*-GlcNAc can also regulate autophagy (lysosomal protein degradation), subcellular localization of target proteins, and other nuclear functions. *O*-GlcNAcylation has also been implicated as an epigenetic modification that regulates stem cell reprograming. UDP-GlcNAc is considered a metabolic sensor, and it has been estimated its intracellular concentrations range from 0.1 to 1 mM, a concentration only second to that of ATP.

IF. Glycan binding

Glycans serve as ligands for many proteins, including cell adhesion molecules or signaling receptors, and thus regulate important biological functions. For example, in mammals, the Ashwell-Morell Receptor (AMR), found in hepatocytes, is responsible for detecting asialoglycoproteins and targeting them for clearance. These asialoglycoproteins, glycoproteins that have lost their terminal sialic acid residues, now expose a terminal Gal that serves as the recognition site for the AMR. In another example, ricin, a highly lethal lectin isolated from the castor bean, binds to GalNAc or Gal residues on cell-surface glycoproteins, and upon internalization inhibits protein biosynthesis. The binding interaction between an individual glycan molecule and a lectin is often low in affinity (K_D ~ mM) and is too weak for most biological functions. However, lectins often possess multiple carbohydrate recognition domains (CRDs) to bind several glycans. This "multivalent" interaction is a hallmark of glycan-protein interactions, and results in avidity enhancements (K_D ~ □M to nM), stronger binding and more selective recognition events. Thermodynamic studies on glycan-lectin interactions have proposed a few binding models to describe that ability of

lectins to cross-link glycoconjugates. (**Fig. 1.5**) Because glycan-protein interactions often trigger biological events, significant effort has been devoted to developing tools to study glycan binding and function. This area of research is referred to as Functional Glycomics. Among the most powerful analytical methods that have emerged in the field in the last decade is the glycan array. Similar to the DNA or protein array, it presents libraries of glycan structures organized in micrometer-sized spots on a chip. Taking the multivalency of glycan interactions into consideration, each spot consists of multiple copies of a unique glycan attached to the surface in close proximity to each other, where they can be interrogated for binding by any protein receptor of interest or even whole cells. Such high-throughput techniques have enabled rapid survey of the structural and functional complexity of the Glycome.

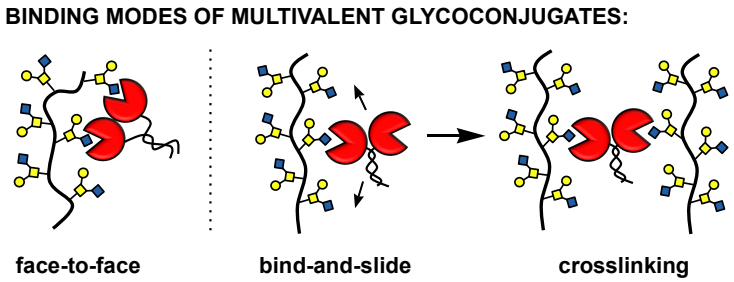


Figure 1.5. Lectins, or carbohydrate-binding proteins engage in multivalent interactions with glycans to generate high *avidity* interactions. Lectins may engage in a face-to-face, or a bind-and-slide interaction. The latter provides an explanation for the ability of lectins to oligomerize cell surface glycoconjugates. (*Reprinted with permission from Huang, M. L. and Godula, K. (2016). Nanoscale materials for probing the biological functions of the glycocalyx. Glycobiology. 26, 797-803. Copyright 2016 Oxford University Press.)*

IG. Consequences of glycosylation defects.

There is increasing evidence that proves glycans can play critical roles for proper biological function. Much of this knowledge results from the observation of phenotypes resulting from mutations in glycosyltransferases and glycan-binding receptors. Many of these

mutations lead to abnormalities that span from embryonic development to aging and disease. Some of these mutations in humans can cause severe diseases, classified collectively as Congenital Disorders of Glycosylation (CDGs). CDGs are autosomal recessive, inborn diseases that cause serious (sometimes lethal) malfunctions of the nervous, muscular, and gastrointestinal systems. The incidence of CDG worldwide is still unclear, although minimal estimates report 1 to every 5000 births. To date, there are over 1000 patients, but many still remain undiagnosed. Historically described as underglycosylation of transferrin, a plasma protein that is responsible for binding free iron to maintain homeostasis, significant progress has now been made in understanding the molecular basis of CDGs. There are now over 50 genes known to cause CDGs. In the most common sub-type of CDG, CDG-1a, a genetic mutation that affects phosphomannulase, an isomerase enzyme that converts mannose-6-phosphate, a marker for lysosomal degradation, to mannose-1-phosphate, leads to a functional loss. Patients with CDG-1a can exhibit different severities in abnormalities, but affected infants exhibit developmental delays and underdeveloped cerebellums. There are currently no treatments available for CDGs, although dietary glycan supplementation (e.g. fucose, mannose) exhibits partial effects on some forms of CDG.

While the functional contributions of individual glycan structures within the cellular glycocalyx are coming into focus, their collective properties are much less understood. The cellular glycocalyx (**Fig. 1.1**) at biological barriers is often dominated by large glycoconjugates, including mucins and proteoglycans, performing unique functions in directing the passage of molecular nutrients or cargoes, while delineating self from non-self and keeping opportunistic organisms at bay. For instance, mucosal barriers are composed predominantly of mucin glycoproteins, whereas the endothelial tissues are populated by proteoglycans. The following sections will discuss these two glycocalyx-dominant classes of

glycoproteins in detail and highlight how the unique structural features of their glycans give rise to the distinct properties and functions of these glycoconjugates.

Section 1.2 Mucins

Mucins are a family of large, heavily *O*-glycosylated proteins present primarily on epithelial luminal surfaces, where they serve lubricating, hydrating, and protective functions. This section will describe the structure, biosynthesis and properties of mucins, as well as their roles in host-pathogen interactions, cancer, and cell surface biophysics.

IIA. Mucin structure

Mucins are distinct from other glycoconjugates, because of their large size, dense glycosylation, extended structure, and ability to form cross-linked gels. The characteristic "mucin domain" is composed of a Variable Number of Tandem Repeat (VNTR) region (**Fig. 1.6**), which provides the structural features that give mucins their unique functions. The VNTR is composed of repeating sequences rich in serine and threonine sites for glycan attachment, and is also enriched for proline residues to prevent protein folding and facilitate glycosylation. Depending on the mucin, these VNTR regions can contain >100 repeat sequences, each of which can harbor potential glycosylation sites. The composition of the VNTR and the number of repeats are mucin-specific. For example, the VNTR of human MUC1 (Mucin1) contains 20-100 repeats of the sequence GSTAPPAHGVTSAPDTRPAP. This 20 amino acid polypeptide contains five serine (S) or threonine (T) residues, and roughly 50-100% of these are glycosylated, creating a glycan-rich domain. In addition to the VNTR module, some mucins have Irregular Repeat (IR) domains. IR domains, like the VNTR modules, are densely *O*-glycosylated through S/T residues, but lack a specific repeating amino acid sequence. In some cases (e.g., in MUC 16), this domain can be larger than the VNTR.

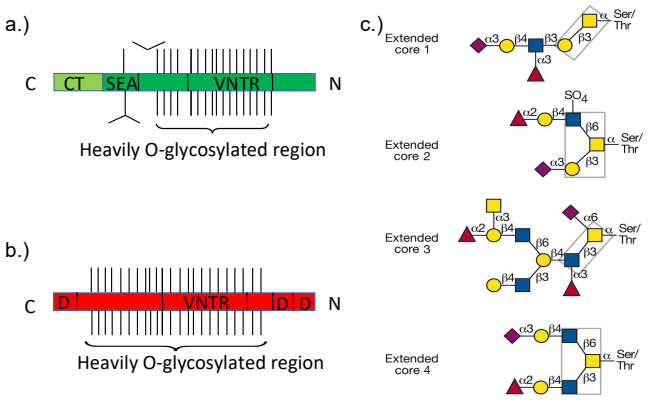


Figure 1.6. Structure of mucin glycoproteins. a.) Cell surface bound mucins contain a cytosolic tail (CT) at their C-terminus, a SEA domain to facilitate cleavage and re-association, and a heavily glycosylated region comprised of VNTR regions as well as irregular repeats b.) Secreted mucins contain D domains rich in cysteine at both the N and C terminus to facilitate crosslinking and mucus gel formation in addition to their heavily o-glycosylated region c.) Four common core structures present in O-glycans.

O-glycosylation of the VNTR and IR domains occurs in the Golgi. The enzyme ppGalNacT initiates this process by transferring an α-*O*-GalNAc to the mucin chain (see ID above). Subsequent elaboration of this core structure is achieved with mucin-specific glycosyltransferases. Just like other glycans, the structure and chemical composition of these complex *O*-glycans is highly variable. In general, there are 4 common core structures of *O*-glycans (and an additional 4 cores have been identified) (**Fig. 1.6**), each of which can be extended to produce a broad range of structures. These glycans typically comprise 50-80% of the glycoprotein by mass.

Such dense glycosylation is often necessary to exert the protective function of mucins. The steric constraints of packing so many glycans into this region force the protein to adopt a more linear secondary structure extended hundreds of nanometers in length above cell surface. An additional consequence of such heavy glycosylation is that the VNTR and IR regions of mucins are often quite resistant to proteolysis, because the presence of the glycan chains can shield the protein backbone from enzymatic degradation. In order for these

domains to be proteolysed, glycosidases must first remove or reduce the glycan chains to allow access to the core protein. Although all mucins contain the characteristic highly glycosylated VNTR and IR regions, differences are often observed in the sparsely glycosylated regions outside this domain between secreted and cell surface bound mucins.

IIA.1. Cell Surface Mucins.

Cell surface mucins are confined to the apical surfaces of cells via a non-covalent linkage between its core protein and a separate transmembrane domain. This class of mucins includes MUC1, MUC3A, MUC3B, MUC4, MUC12, MUC13, MUC15, MUC16, MUC17, MUC20, and MUC21, each of which vary in size, composition, distribution, and function. These mucins are variably expressed throughout the body, where they fill a range of functions. The single-pass transmembrane domains contain a cytoplasmic tail for anchoring onto the cell surface. Although the extracellular and cytoplasmic domains of the protein are co-translated, they quickly undergo autoproteolysis into a heterodimer within the ER to yield a large extracellular region that contains a short cell surface linker (**Fig. 1.7**). These cleaved subunits then associate non-covalently, often through a SEA (Sea urchin sperm protein, Enterokinase, and Agrin) module. Additionally in the ER, the protein is *N*-glycosylated before it is transported to the Golgi, where *O*-glycosylation takes place before transport to the cell membrane. On the cell surface, the extracellular portion of these mucins can be cleaved through a variety of mechanisms, ranging from proteolysis of the protein backbone near the surface, to decoupling of the two subunits. This release is especially significant in mucins function as a barrier to pathogens (section IID).

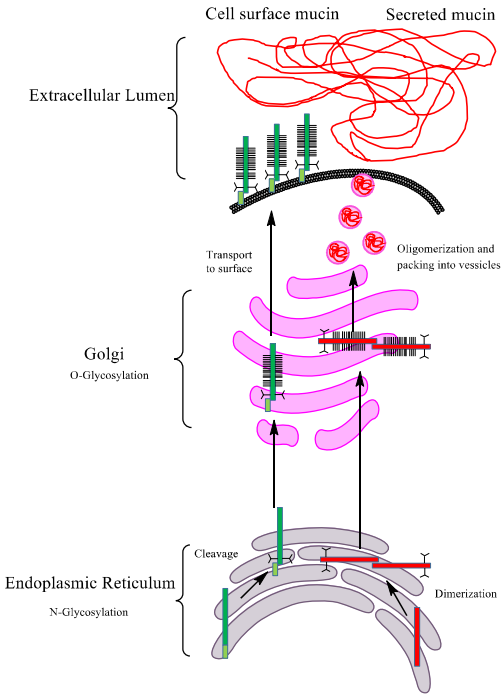


Figure 1.7. Biosynthesis of mucins. Transmembrane mucins are transcribed in the ER, where they are N-glycosylated and cleaved at their SEA domain. The two subunits re-associate before transport to the golgi for O-glycosylation prior to surface presentation. Secreted mucins are transcribed in the ER, where they are N-glycosylated and dimerize before transport to the golgi for O-glycosylation. The mucins are packaged into vesicles where they oligomerize and are secreted into the extracellular space.

IIA.2. Secreted Mucins.

Secreted mucins produced by specialized goblet cells in the epithelium lack the transmembrane and cytoplasmic domains of their surface bound counterparts, and are instead outfitted with domains tailored to extracellular function and mucus formation. This comes primarily in the form of "D" domains on either side of the VNTR/IR region, which are cysteine rich and allow for oligomerization and generation of viscous gels through formation of disulfide bonds. The family of secreted mucins includes MUC2, MUC5AC, MUC5B, MUC6, MUC7, MUC8, MUC9 and MUC19. Similar to cell surface mucins, secreted mucins first

receive *N*-glycan modifications in the ER. However, rather than cleavage of the mucin into discretely functional subunits, dimerization occurs via their C-terminal D domains. This dimer is then transported to the Golgi for *O*-glycosylation, and is packed into secretory granules. Within these granules, oligomerization of the dimers into networks occurs through the D-domains on their N-terminus. Upon secretion, hydration of the mucin layer occurs, swelling in size by 100-1000 fold into a gel-like mucus layer.

IIB. Mucins and mucosal barriers

Epithelial cell surfaces are the physical barriers protecting underlying tissues from exposure to pathogens, and must often present a selectively permeable barrier to allow transport of beneficial materials while excluding harmful particles. Mucosal barriers are found in tissues that directly interface with the outside world and include the gastrointestinal (GI), respiratory and reproductive tracts, as well on the ocular surface or in the middle ear. The chemical and physical properties of mucins provide key contributions to the defense mechanism of mucosal barriers. In this section we will examine the structure and function of mucins as they relate to physiological function as well as barrier formation against pathogenic advancement. The opposite side of the same proverbial coin is of course the continuous evolution of pathogens to take advantage of host glycans to circumvent these defense mechanisms.

IIB.1. Gastrointestinal tract.

From the moment food reaches the stomach to its excretion from the colon, a single layer of columnar epithelial cells acts as a barrier between the underlying host tissue and the intestinal lumen. This barrier balances a variety of competing protective and harmful

processes. For instance, the mucus layer must serve as a harbor for commensal bacteria (microorganisms that have evolved a mutually beneficial relationship with their human host) and must allow food to be digested, absorbed and processed, but cannot allow pernicious pathogens to penetrate the cell surface. The mucosal barrier is composed of a secreted and a cell surface-bound layer. Each layer has its own mucin composition, and provides specialized barrier functions to the underlying epithelium. Goblet cells are the primary constituent of this mucosal lining.

MUC5AC is the most highly expressed secreted mucin in the stomach, whereas MUC2 is the primary component of mucus in the small and large intestines. Although mucins are the main contributors, the extracellular barrier also contains a wide variety of proteins, most of which do not yet have defined roles. It is thought that many of these proteins, such as trefoil peptides, interact non-covalently with mucins to fine-tune the rheological properties of the mucus. Antimicrobial molecules and secretory antibodies are also present in the mucosal barrier to provide defense against invasive pathogens. These molecules are contained within the mucin matrix, where they can be better retained and localized at increased concentrations. Without the mucus barrier, these secreted molecules would be lost quickly to diffusion in the lumen. The mucus matrix also allows for these components to be present in a gradient, which promotes robust pathogen defense close to the cell, while still enabling colonization of the outer layer by beneficial commensal bacteria.

The mucosal barrier is composed of two continuous layers of mucin matrix, termed the inner (closer to the epithelial surface) and outer layer (directly in contact with the GI tract). The inner mucus layer is formed from mucins of the secretory vesicles released by goblet cells and is primarily defined by its viscoelastic and antibacterial properties. It is extremely viscous as a result of the cross-linked high molecular weight mucins. The

compounded effects of viscosity and the presence of antimicrobial molecules provide resistance to colonization of the inner mucus layer of healthy epithelium by both commensal and pathogens. The outer layer of mucus is formed from the partial enzymatic degradation of the inner mucus, making it less viscous and allowing for greater diffusion of its contents. This decrease in antimicrobial peptide concentration allows commensal bacteria to colonize the outer mucus. This microenvironment provides protective conditions and nutrients for commensals. The continuous turnover of the mucosal barrier generates a constant flux of secreted mucins passing from the inner mucus layer to the outer, which is in turn degraded and removed through mucociliary motion, a cycle which takes ~24 hours to complete under normal conditions.

A variety of cell surface-bound mucins are also produced throughout the GI tract, including MUC1, MUC3A, MUC3B, MUC4, and MUCs 12-17. MUC1 is the most common and well characterized. In the gut epithelium, transmembrane mucins (section IIB) localize only to the apical (outward facing, into the GI tract) surface of epithelia cells. Due to the increased size of mucins relative to other glycocalyx components (>100 nm compared to ~10 nm), localization of mucins to the lateral or basal surface of the cell can interfere with adhesion of the cell to the surrounding tissue. As such, the primary functions of these glycoprotein structures are localized on the apical surface, where they maintain protective functions (see section IIC), and can serve as “outside-in” signaling molecules.

IIB.2. Respiratory tract.

Similar to the GI tract, the respiratory tract (RT) serves as an interface with the outside environment to facilitate nutrient exchange (in this case, gases like oxygen). As with the GI tract, the air passing through the RT also contains a variety of physical and pathogenic

components. The mucin barrier serves as a critical first line of defense against these airborne particulates and pathogens, and both cell surface and secreted mucins are present to provide two layers of protection.

The composition of secreted mucus in the RT is different compared to the GI tract, and is formed by a variety of mucins, of which MUC5B and MUC5AC are the primary gel formers. The secretion and crosslinking of these mucins provide the scaffold, which then absorbs additional secreted globular proteins that further modulate the viscoelastic properties of the mucus gel. Secreted mucus also contains a variety of RT-specific defensive molecules (e.g. antimicrobial peptides, reactive oxygen species) that directly attack foreign bodies and pathogens. Particulates caught in the secreted mucus matrix can also be mechanically removed through mucociliary action or coughing. The common over-the counter cough drug, Guaifenesin (sold as Robitussin or Mucinex in the US), assists in relieving the symptoms of a respiratory infection by facilitating the clearance of pathogens from the mucus. In experiments involving differentiated human airway epithelial cells, researchers have shown that guaifenesin can reduce the expression of mucins and decrease mucus viscosity.

Transmembrane mucins form a final protective barrier on the epithelial cells lining the RT. Like in the GI tract, MUC1, MUC4, and MUC16 have been identified as the primary mucins present in the airway epithelium, with MUC1 again being the best characterized. As in the gut, MUC1 is capable of providing outside-in feedback from the extracellular environment, although the mechanisms are unclear. Overexpression of MUC1 in response to the pro-inflammatory factor TNF-α has been shown to decrease inflammation during infection.

IIC. Mucins in innate immune defense against pathogens.

Because pathogens often contain glycan-binding proteins on their surface, various mucin glycans can serve as decoy ligands for these receptors, trapping pathogens in the mucus before they can reach underlying cells. Pathogen binding to decoy cell surface mucins often triggers shedding of the extracellular region of the glycoprotein. Ideally, the pathogen remains bound to the glycans as the mucin diffuses away, removing it from the cell surface. Research suggests that cleavage of the extracellular subunit could alert the cell to the pathogen's presence and stimulates signaling cascades to mount a defense. As part of this mechanism, the detection of pathogenic infection results in increased flux of mucins to the cell membrane, allowing for the untethered decoy to be carried away into the secreted mucus layer for clearance. However, more evidence is still needed to fully corroborate this hypothesis.

Aside from direct binding interactions, the viscous, cross-linked mucin network can mechanically exclude larger particles, such as viruses and bacteria, from penetrating. Cell surface mucins, because they tower over the rest of the glycocalyx, can also sterically hinder incoming pathogens from reaching their receptors closer to the cell surface. Even if the relevant mucin does not contain the appropriate ligand to act as a decoy, its sheer steric bulk can act as a shield to prevent the pathogen from engaging the cell surface.

These defense mechanisms have been beautifully demonstrated by the interactions of the bacterium, *H. pylori*. In recent *in vitro* experiments, gastric epithelial cells exhibited resistance to the bacterium depending on the presence of the mucin binding adhesins SabA (Sialic acid binding adhesin) and BabA (blood group binding adhesin). In the case that the bacterium contains these adhesins and can bind to mucins, infection is prevented through shedding of the mucins. However, infection is still not observed for bacterium without the

adhesins. In this case, the mucin is not shed, but may be physically blocking *H. pylori* from binding to the cell surface, preventing infection.

IID. Pathogen subversion of mucosal barriers.

Despite the innate defenses presented by the mucus layer, pathogens are at times able to breach this barrier and infect the host cells. (**Fig. 1.8**) To overcome its protective features, pathogens have evolved strategies to circumvent or destroy the cell's natural defenses. Broadly speaking, pathogen countermeasures are designed to either degrade mucins on their way to the cell, or secrete factors that interfere with their production or release.

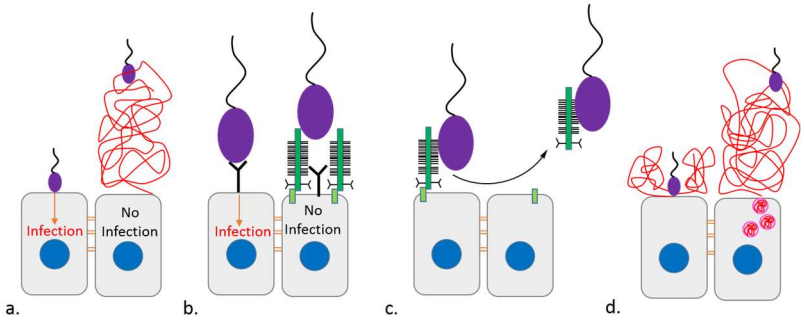


Figure 1.8. Mucin defenses against pathogen invasion. a.) Secreted mucins can protect the cell from pathogens through steric hindrance or through glycan binding interactions b.) Transmembrane mucins can sterically prevent pathogens from reaching receptors on the cell surface c.) Upon binding of pathogens to transmembrane mucins, the extracellular domain can be released from the cell surface d.) During pathogen invasion mucin production is upregulated, increasing the flux of mucus and amplifying its protective effects.

The degradation of mucins is a challenging task, as the densely glycosylated VNTR/IR regions are inaccessible to peptidases. Therefore, bacterial enzymes must act on the non-glycosylated regions (such as the cysteine rich crosslinking domains) or be able to digest the glycans prior to deploying proteolytic enzymes. Indeed, bacteria often express glycosidases that trim decoy glycans to expose the protein backbone. The breakdown of the cross-linked mucin network results in reduced viscosity, which disrupts the steric exclusionary effects of

the mucus barrier, and release of antimicrobial agents. Some pathogens can also secrete small molecule toxins that pass through the mucus layer and cell membrane to disrupt normal cellular protective functions, such as mucin biosynthesis. The mucin barrier is constantly being degraded even during normal cell function, so without the secretion of new mucus layers from goblet cells, the barrier will lose its protective functions rather quickly.

A classic example of pathogenic invasion comes from the influenza A virus, which is notable for its interaction with the mucin barrier. In a healthy mucosal barrier, influenza A virions can be sterically excluded from the cell. Influenza has two sialic-acid recognizing proteins that allow it to cut its way through the mucus: the hemagglutinin (HA) binding receptor and neuraminidase (NA), which cleaves sialic acids. There are hundreds of copies of HA on the viral envelope acting as receptors for sialic acids present on mucins. Although the affinity of each HA receptor to individual sialic acids is relatively low, multivalent binding of the collective HA on the viral envelope to ubiquitous sialic acids present on mucin glycoproteins leads to high avidity binding. NA then cleaves the sialic acids off of the glycoprotein, removing the carbohydrate ligand for HA to promote viral release and passage through the mucus layer.

IIE. Mucins in cancer

Although the primary role of mucins is to protect cells, they can also be co-opted by cancer cells to evade immune responses and, by altering cell adhesion, to aid in metastasis and tumor spreading. Mucins are aberrantly expressed in roughly 65% of all cancers and more than 90% of epithelial cell adenocarcinomas. Both over-production and anomalous glycosylation of mucins are common in cancer cells. Altered glycosylation stems from differential expression of glycotransferases in cancer cells and the resulting cancer-specific glycans can be used as biomarkers for cancer detection and prognosis. While the transmembrane MUC1 was the first mucin recognized as a cancer biomarker, other mucins have been implicated in various carcinomas, including MUC2, MUC4 (pancreatic, ovary, breast, lung, etc.), MUC13 (colon), MUC16 (ovarian), and MUC17 (pancreatic). In this section we will review some of the mechanisms through which mucins can aid in cancer cell survival and tumorigenesis (**Fig 1.9**).

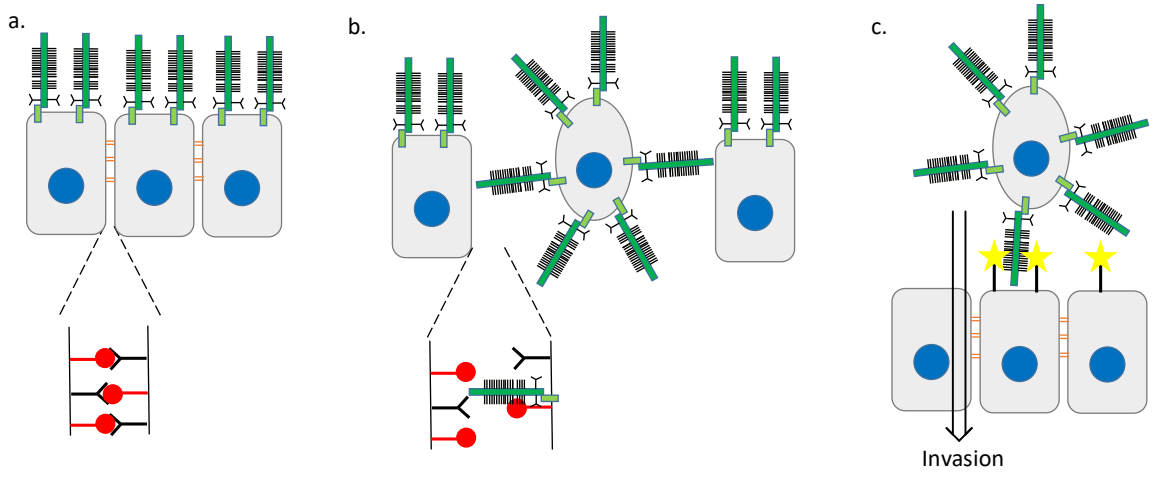


Figure 1.9. Mucins in metastasis. a.) In normal epithelial tissue, tight junctions are maintained between cells b.) Loss of polarity results in loss of tight junctions through steric exclusion and cis-binding to receptors on the cell surface c.) Glycans on mucins can bind to receptors on new tissues, and can trigger invasion and metastasis.

IIE.1. Disruption of cell-cell adhesion.

Often in cancerous states, the normal basal-apical polarity of the cell is lost, allowing mucins to localize to non-apical surfaces. Because these mucins extend hundreds of nanometers from the cell surface, they cover a much larger distance than the gap between cells in the epithelium, which are sealed with tight junctions (to maintain polarity) and adherens junctions (to maintain epithelial morphogenesis). In cancer cell progression, mucins can promote detachment and metastasis through disruption of these connections (**Fig. 1.9b**).

Tight junctions serve as a partition between the lateral and apical surfaces of the cell, a function maintained through PAR complexes with atypical protein kinase C (aPKC). Formation of the PAR complex and establishment of polarity is mediated by ERBB2 receptor tyrosine kinase, which disrupts the PAR complex association and leads to loss of polarity. Mucins (specifically MUC1 and MUC4) can bind to ERBB2, which leads to promotion of ERBB2 signaling and disruption of the tight junctions. The resulting loss of polarity allows mucins to localize to the basal and lateral surfaces of the cell, and paves the way for loss of cell-cell adhesion through disruption of adherens junctions.

Adherens junctions maintain the spatial integrity of cells by forming intercellular homotypic cadherin interactions between adjacent cells. In addition to cell-cell binding, the intracellular domain of cadherins directs the actin cytoskeleton network through interactions with cytosolic ß-catenin. MUC1 is able to doubly disrupt this interaction on both the extracellular and cytosolic fronts. When MUC1 is overexpressed on the lateral region of the cell, the steric bulk of the extended mucins can physically separate cadherins on adjacent cells. In addition, the glycans on the mucins are capable of creating cis-interactions with adhesive receptors on the cell surface, preventing them from forming trans interactions with

other cells. Additionally, the intracellular CT domain of MUC1 it is capable of disrupting the ß-catenin/cadherin complex. Upon phosphorylation, the CT domain competes with cadherin for binding of ß-catenin, sequestering it and interfering with the cytosolic components of the adherens junctions. Additionally, the ß□catenin/MUC1 complex can traffic to the nucleus and influence signaling, and can modulate Wnt genes involved in proliferation and differentiation. This disruption of the adherens junctions is one part of the complex process of detachment and metastasis of cancer cells. In short, mucins appear to be crucial in the loss of polarity and intercellular adhesion in cancer, and have emerged as key players in motility and metastasis.

IIE.2. Contributions to adhesion.

Paradoxically, while MUC1 is heavily involved in detachment of cancerous cells, it is also implicated in promoting their subsequent adhesion in their target tissues. To fully metastasize, circulating cancer cells must re-adhere and re-colonize to create a new tumor. To do so, cancer cells often borrow a strategy from the immune system. During inflammation, epithelial cells are known to express E-selectins, which are sialic acid binding proteins. This binding event pulls the circulating immune cells into the tissue so they can perform their function. Mucins in cancer cells also contain sialic acid residues that serve as E-selectin ligands on epithelial cells, in addition to other adhesion markers, such as CD-44 (which itself is a non-mucin glycoprotein). Through this interaction, rolling of the tumor cells along the metastatic site is initiated, but full adhesion to the tissue is dependent on secondary anchoring mechanisms to prevent release back into circulation. Integrin-mediated interactions with ECM components (such as ICAM-1) then occur to complete metastasis.

Mucins are likely involved in this process as well, as certain MUC1 isoforms are involved with binding to ICAM-1 (**Fig 1.9c**)

Recent studies also point to the biophysical properties of mucins, which may play a role in tumor cell adhesion through cell-surface integrin organization. In experiments utilizing synthetic mucin-mimetic glycopolymers, researchers discovered that the presence of sterically demanding glycopolymers causes integrins to cluster in "kinetic traps", which resulted in enlarged focal adhesions and integrin-dependent activation of proliferation and cell survival.

IIE.3. Inhibition of immune response.

Mucins also play a role in cancer cell survival by protecting it from immune surveillance. Similar to the shielding action of epithelial mucins towards pathogens, cancer cells also utilize mucins to protect themselves from the body's natural defense mechanisms. For example, when healthy epithelial cells lose their adhesive properties and detach from their local environment, a specialized apoptotic process called anoikis is initiated to remove the cell from circulation. In the case of metastasizing cancer cells, MUC1 shields the associated death receptors on the cell surface to inhibit anoikis. More generally, when cells become cancerous, they can present factors that are recognized by various immune responses to target for cell death and removal. Although research in the field is ongoing, it is plausible that, through their roles in barrier function, secreted or cell surface mucins could prevent the approach of anti-tumor molecules designed to detect and remove cancerous cells. Significantly, the glycosylated domains of mucins are capable of interacting with leukocytes, and at high levels of soluble mucin secretion, may trap immune cells before they reach cancer

cells. Similar mechanisms are possible for sequestering other anti-tumor immune effectors, such as natural killer cells.

In summary, mucins are formidable biomolecules that provide robust protective barriers at the interface of organisms and hostile outside environments. Their structure and glycosylation provides mucins with unique physical and biological properties to control the passage of nutrients into cells while restricting access from particulates and pathogens. Mucins not only provide passive protection but also report on the cell's physiological state, modulate its interactions with the surrounding matrix and mediate cellular communication. Mucins can also contribute to the pathophysiology of cells, such as carcinogenesis, by facilitating detachment and adhesion in metastasis or by protecting aberrant cells from innate defenses and immune surveillance. Although the exact mechanisms cells use to control these seemingly contradictory functions are not yet fully known, mucins are increasingly recognized as biomarkers for cancer and other diseases and will continue to be a focus of intensive biomedical research.

Section 1.3 PROTEOGLYCANS

Proteoglycans are glycoconjugates modified with extended *O*-linked glycosaminoglycan (GAG) polysaccharide chains. Proteoglycans are a major component of the endothelial glycocalyx and are abundantly present in the ECM. Like mucins, proteoglycans serve both as a structural element as well as a relay for biological information. Structural PGs can serve as scaffolds for tissues, and are found in joints and cartilage. As information molecules, PGs can function as co-receptors for adhesion molecules, growth factors and cytokines at the cell surface and facilitate cellular signaling. At the ECM, proteoglycans can sequester growth factors (GFs) away from their receptors to inhibit stimulation, or they can transport and stabilize gradients for GFs and morphogens. A plethora of proteoglycans are present in various tissue to regulate cellular physiology (**Table 1.1**).

Table 1.1. Membrane-bound and secreted proteoglycans, their locations, and functions.

Proteoglycan	number/ type GAG chains	Presence	Functions
Membrane-bound Proteoglycans			
Syndecans 1-4	1-3 HS, 0-2 CS	nucleated cells	regulates cell adhesion, migration, GF signaling, ligand clearance
Betaglycan	0-1 HS, 0-1 CS	fibroblasts	regulates TGFβ signaling
Glypicans 1-6	1-3 HS	epithelial & mesenchymal cells	GF co-receptors to regulate signaling
Phosphacan	2-5 CS/DS	brain	regulates migration of cortical neurons
Thrombomodulin	1 CS/DS	endothelial cells	anticoagulant
CD44	0-2 CS/DS/ HS	lymphocytes, among many others	receptor for hyaluronan
NG2/CSPG4	2-3 CS/DS	stem cells, glial progenitors,	modulates cell-ECM interactions, GF interaction, integrin activation
CD74	1 CS	antigen-processing cells	chaperones MHC II; enhances T-cell activation
SV2	1-3 KSI	synaptic vesicles	regulates exocytosis of neurotransmitters
Secreted proteoglycans			
Aggrecan	~100 CS, ~20 KSII	cartilage, brain	hydration barrier
Versican	0-15 CS/ DS	ECM, blood vessels, brain, leukocytes	regulates inflammation, cell migration
Neurocan	1-2 CS/DS	brain	regulates neurite outgrowth
Brevican	0-4 CS/DS	brain	regulates synaptic plasticity
Decorin	1 CS/DS	connective tissue	regulates collagen fibrillogenesis
Biglycan	0-2 CS/DS	macrophages	activates innate immunity via TLR2/4
Lumican	3-4 KS I	widely distributed	collagen matrix assembly
Keratocan	3-4 KS I	cornea	corneal transparency
Fibromodulin	2-4 KS I	widely distributed	collagen matrix assembly
Osteoglycan	2-3 KS I	cornea	corneal transparency
Perlecan	1-3 HS, 0-2 CS	ECM	regulates cell migration
Agrin	1-3 HS	brain, NMJ	regulates NMJ maturation
Collagen type IX	1 CS/DS	cartilage	stabilizes cartilage-collagen fibrils
Collagen type XVIII	2-3 HS	basement membranes	regulates stability of basement membranes

IIIA. Proteoglycan biosynthesis, structure, and the "sulfation code."

The biosynthesis of proteoglycans begins with the genome-templated expression of the core protein in the ribosomes (**Fig. 1.10**). The protein transcript is then processed in the Golgi apparatus, where the attachment of glycosaminoglycan (GAG) chains occurs within Ser-Gly consensus sequences. This process begins with the priming of protein Ser/Thr residues with a tetrasaccharide (GlcA-Gal-Gal-Xyl-) linker. During this process, xylosyltransferase enzymes (XYLT1, XYL2) attach a xylose monosaccharide to the Ser residue of the core protein. The attached xylose residue then serves as a substrate for galactosyltransferases (B4GALT7, B4GALT6, and B4GALT3, respectively) that successively attach the Gal, Gal,

and GlcA residues of the linker. Following linker synthesis, elongation of the GAG chain occurs. GAGs are composed of a repeating disaccharide motif, composed by an uronic acid linked to an amino sugar. For example, heparan sulfate (HS) chains are composed of GlcAα(1-4)GlcNAc disaccharides, whereas chondroitin sulfate (CS) GAGs are composed of GlcAα(1-3)GalNAc. The resulting long, linear, and highly negatively charged GAG chains are considered the most abundant heteropolysaccharides in the body. Individual GAG chains can be as long as 100 disaccharide units long. Given the large size of GAGs, they tend to dominate the chemical and physical properties of proteoglycans.

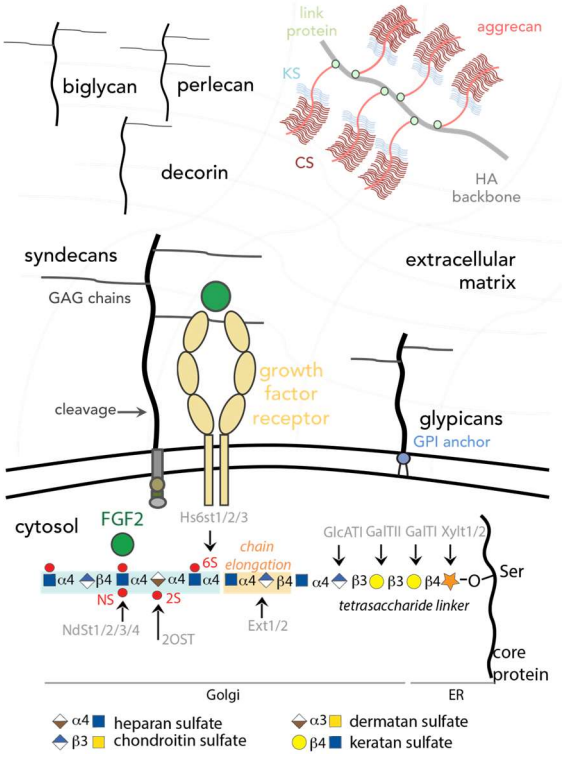


Figure 1.10. Proteoglycan examples, structure, and biosynthesis. Proteoglycans can be present as transmembrane or soluble forms. Proteoglycan biosynthesis occurs in the ER and Golgi, where a tetrasaccharide GlcA-Gal-Gal-Xyl linker is first attached to Ser/Thr residues of the core protein. Subsequent elongation and modification by exostosin and sulfotransferase enzymes, respectively generate linear long-chain sulfated glycosaminoglycan chains that encode binding sites for growth factors. At the glycocalyx, proteoglycans can serve as co-receptors for growth factors and their receptors.

Sulfate groups can decorate the GAG chains, and proteoglycans are known to possess domains of high, medium, and low sulfation. These modifications are the result of sulfotransferases that transfer sulfate groups from activated PAPS (3'-phosphoadenosine-5'-phosphosulfate) substrates during chain polymerization. A suite of sulfotransferases work in concert to transfer sulfate groups to various positions of the sugar backbone. For example, NDST1/2 (*N*-deacetylase/*N*-sulfotransferase) is a bifunctional enzyme that can add or remove sulfate groups to or from the amino group of the glucosamine residues of HS. There are also 6-O-sulfotransferases and 2-O-sulfotranferases that catalyze the transfer of sulfate groups to the 6- and 2- hydroxyl position groups of HS, respectively. For CS, chondroitin-4-O-sulfotranseferase (C4ST) and chondroitin-6-O-sulfotransferase (C6ST) catalyze the sulfation of sulfate groups at 4-O or 6-O positions of the CS GalNAc residue, respectively, whereas 2-O sulfation of the GlcA residue is achieved with uronyl 2-O-sulfotransferase (UST). A few naming schemes have emerged throughout the years for classifying the disaccharide repeating units present in GAGs. **Table 1.2** outlines the structures and naming schemes for the major naturally-occurring disaccharide motifs.

Table 1.2. GAG nomenclature. Disaccharide structure code (DSC) of HS and CS repeating units.

DSC	Roman numeral	Structure
Heparan sulfate disaccharides		
D0H0	IV-H	ΔUA-GlcNH$_2$
D0H6	II-H	ΔUA-GlcNH$_2$-6S
D2H0	III-H	ΔUA-GlcNH$_2$
D2H6	I-H	ΔUA2S-GlcNH$_2$-6S
D0A0	IV-A	ΔUA-GlcNAc
D0A6	II-A	ΔUA-GlcNAc6S
D2A0	III-A	ΔUA2S-GlcNAc
D2A6	I-A	ΔUA2S-GlcNAc6S
D0S0	IV-S	ΔUA-GlcNS
D0S6	II-S	ΔUA-GlcNS6S
D0S3		ΔUA-GlcNS3S
D2S0	III-S	ΔUA2S-GlcNS
D2S6	I-S	ΔUA2S-GlcNS6S
D2S3		ΔUA2S-GlcNS3S
D0S9		ΔUA-GlcNS3S6S
D2S9		ΔUA2S-GlcNS3S6S
Chondroitin sulfate disaccharides		
D0a0		ΔUA-GalNAc
D0a4		ΔUA-GalNAc4S
D0a6		ΔUA-GalNAc6S
D2a0		ΔUA2S-GalNAc
D2a4		ΔUA2S-GalNAc4S
D2a6		ΔUA2S-GalNAc6S
D0a10		ΔUA-GalNAc4S6S
D2a10		ΔUA2S-GalNA4Sc6S

In contrast to the other major biopolymers (e.g. DNA/RNA, proteins), the acknowledgment that GAGs (or even glycans in general) possess some type of specificity and that it harbors information in a sequence-specific manner, has historically been slower to come. Given the complexity of glycans, and the heterogeneity of sulfation in GAGs, it is not surprising that the "sulfation code" has taken years to decipher and still remains an ongoing enterprise. Through painstaking work, we now know that specific sulfation patterns and three-dimensional structures of HS and CS chains can encode for growth factor binding sequences to affect various biological processes.

Proteoglycans can be present as membrane-anchored or secreted proteins (**Fig. 1.10**). In some cases, alternative mRNA splicing heavily dictates these two forms. Proteoglycans can be anchored onto the plasma membrane via transmembrane protein domains, such as those found in syndecans. On the other hand, smaller proteoglycans, such as glypicans, use

the glycophosphatidylinositol (GPI) glycolipid anchor. Transmembrane proteoglycans, such as syndecans, are known to also exist in the soluble form via an event called "ectodomain shedding," in which the extracellular portions can be clipped off the membrane by "sheddases" that recognize a region of the core protein proximal to the cell membrane. Ectodomain shedding is a highly regulated event, and it can rapidly alter the resulting phenotype of affected cells by reducing the amount of ligands on the cell surface or competing with membrane-bound ligand. On the other hand, glypicans can be released from the cell surface by the action of phospholipases that cleave the ester bonds of the GPI anchor.

IIIB. Other major classes of GAGs

IIIB.1. Hyaluronic acid.

In contrast to other GAGs that are sulfated and anchored onto a core protein, hyaluronic acid (HA) or hyaluronan, is non-sulfated and soluble. Named after the Greek word for vitreous, *hyaloid*, this glass-like substance, was first isolated in the early 1930s. This linear polysaccharide is larger than other types of GAGs (10^7 Da), and can be composed of up to 10^4 repeating GlcA-GlcNAc disaccharide units. Rather than being produced in the Golgi, HA is formed in the plasma membrane. Hyaluronan synthases are responsible for the biosynthesis of HA. These integral membrane proteins extend HA by the repeated addition of GlcA and GlcNAc to a growing polysaccharide. The resulting HA is extruded via the transmembrane ABC (ATP-binding cassette) transporter into the extracellular space. The carboxylic acid groups on the GlcA residues of HA are anionic at physiological pH, and its high anionicity and large size confers its high hydration volume. HA is a major component of the ECM, as well as the connective, epithelial, and neural tissues. HA has been known to

exist in multiple sizes, low (LMWHA, <500 kDa), and high-molecular weight HA (>500 kDa). Peculiarly, there is feedback inhibition that occurs among different HA sizes. For example, HA synthesis can be inhibited by LMWHA, but is stimulated by HMWHA in human synovial fibroblasts. Moreover, it is thought that fragments of HA that occur as a result of injury can be pro-inflammatory, whereas the native HMWHA are not. Due to its ubiquitous presence in the ECM, HA is a popular scaffold for tissue engineering and biomaterials applications.

IIIB.2 Keratan sulfate (KS).

KS is a type of linear sulfated GAG molecule found enriched in the corneal and skeletal tissue. KSPGs include lumican, aggrecan, mimecan, among others. KS is composed of a Galß1-4GlcNAc repeating units that can be sulfated at carbon 6 of either or both of the Gal or GlcNAc residues. Interestingly, KS can be mannosylated at their linkage region, and some regions along the chain can be fucosylated, and a large majority of KS chains are terminally sialylated. There are two major classes of KS, originally classified based on where they were isolated, but have now been assigned based on their linkage to the core protein. KSI is *N*-linked to Asn via GlcNAc, whereas KSII is *O*-linked to Ser/Thr via GalNAc residues. At the cornea, KS is produced by specialized fibroblasts called keratocytes. KS retains the hydration content of the cornea. Lumican specifically is known to organize collagen fibrils in the cornea to achieve optical transparency.

IIIC. Glycocalyx proteoglycans and glycosaminoglycans in the vasculature.

GAGs are prominent players in the vasculature. HS alone comprises 50-90% of the total endothelial GAGs. Heparin is perhaps the most well studied GAG in the context of the

circulatory system. In contrast to the easily interchangeable "heparan sulfate," heparin is a soluble form of HS. Whereas domains of sulfation occur in HS, heparin is isolated mostly as a trisulfated product, and is often considered an over-sulfated variant of HS, with ~2.3 sulfate groups per disaccharide versus 0.8 for HS. In its natural form, heparin has a molecular weight ranging from 3 to 30 kDa. Heparin has significant anticoagulant activity and is often used in the clinic as a blood thinner in coronary syndromes and kidney dialysis. The anticoagulant activity of heparin, lies in its unique pentasaccharide sequence that includes a 3-*O*-sulfated glucosamine residue that is critical for binding to the enzyme inhibitor antithrombin III (AT). This rare modification occurs in 1 out of 20 residues in heparin, and 1 out of 200 in HS. Seven mammalian 3-*O*-sulfotransferase (Hs3st) enzymes that install these sulfate groups on HS, constitute the largest group of sulfotransferases involved in HS biosynthesis. Despite this large number, there is still a limited array of proteins known to be modulated by 3-*O*-sulfation.

Until recently, pharmaceutical heparin was isolated from porcine mucosa or bovine lungs. There is substantial chemical heterogeneity in these samples, and improperly purified and characterized sources can lead to the presence of deleterious contaminants. In 2008, oversulfated CS was found to be a major contaminant of heparin that led to a number of adverse reactions in patients. To circumvent problems associated with insufficiently pure heparin, low-molecular-weight heparins (LMWHs) that have a molecular mass of <8 kDa are now being used in the clinic. LMWHs such as Arixtra (Fondaparinux) are now being used in the clinic for deep vein thrombosis and pulmonary embolisms.

IIID. Glycocalyx proteoglycans in development

During development, GFs and morphogens provide the cues for directing cell fate. Proteoglycans can interact with a number of GFs through their GAG chains. The secreted morphogens, FGF, Wnt, Hedgehog, and BMP emanate from cells and can trigger various signaling events in a concentration-dependent manner. The ability of proteoglycans and morphogens to work in concert and create these concentration gradients has been a key issue in understanding mammalian development.

IIID.1. Embryonic development.

Studies involving genetic knockouts of the HS elongating enzyme, exostosin-1 (Ext1), have revealed the importance of proteoglycans in embryonic development. Ext1 deficiency in mouse embryos is considered embryonic lethal, as these embryos fail to gastrulate or develop. The mouse embryonic stem cells isolated from these embryos revealed that HS at the stem cell glycocalyx functions as a co-receptor to organize FGF2 and FGFR ternary complexes. The formation of this complex triggers the activation of signaling pathways that provide cues for the stem cell to exit from pluripotency and differentiate. Void of Ext1, these mutant stem cells cannot different and remain fully pluripotent.

Random mutagenesis experiments have also identified that mutations in the *lazy mesoderm* gene in mice (*sugarless* in *Drosophila)*, involved in proteoglycan synthesis causes improper mesoderm migration during gastrulation, and is also embryonic lethal. *Lazy mesoderm* encodes for UDP-glucose dehydrogenase, the enzyme that oxidizes UDP-Glc into UDP-GlcA. Tthese defects in cell migration resulted from the loss of FGF signaling. Further

confirming the importance of proteoglycans in FGF signaling, mutations in the HS domains required for FGF binding have also been explored. Genetic knockouts of 2-O-sulfotranferases and glucuronyl C5-epimerase leading to the loss of 2-O-sulfate or GlcA residues of HS, respectively, also result in embryonic lethality due to abnormal development, but such mice progress further than those with complete loss of GAGs. In addition to HS, mutation of *Hyal2* has also been shown to be embryonic lethal. It is worthwhile to note, however, that not all single mutations or knockouts lead to embryonic lethality, as the function of some enzymes can be compensated by others.

In murine embryonic stem cell differentiation, the sulfation patterns on HS are known to be dynamic, and can change depending on the day of differentiation. In studies highlighting the importance of sulfation on differentiation, pluripotent embryonic stem cells synthetically engineered to display high degrees of sulfated HS displayed accelerated exit from self-renewal and differentiation through increased activation of the FGF/Erk mediated signaling pathways.

IIID.2. Myogenesis.

Lipid rafts are important constituents of the plasma membrane. Through these localized structures, signals arising from cell surface events can be efficiently localized to increase binding and enhance cell sensitivity to propagate to downstream events. Glypican-1, has been found to be an important component of lipid rafts in order to allow the development of mature myofibers from precursor myoblast cells. Here, glypican-1 is localized into membrane microdomains, where they sequester FGF2 away from their receptors, which are diffused on the membrane. Through this sequestration, FGF2 signaling is inhibited and myogenic differentiation occurs. Other proteoglcans have also been known to modulate FGF2

signaling by congregating into lipid rafts. For example, syndecan-4, have been known to recluster into lipid rafts upon activation by FGF2, although the significance of this event is less known.

IIID.3. Neuromuscular junction.

The neuromuscular junction (NMJ) is the synapse connecting a motoneuron and a muscle fiber. The development of the NMJ is protracted and extends beyond the embryonic stage and continues after birth. The NMJ harbors clusters of acetylcholine receptors that allow the proper propagation of the action potential. To form the NMJ, motoneurons must secrete an important proteoglycan called agrin. Agrin is an "extracellular synaptic organizing molecule," the connecting piece between the motoneuron and the myofiber, and agrin deficiency halts NMJ development. Agrin expression is highly regulated by alternative mRNA splicing. Through its various forms, agrin is present in the NMJ, in neuronal synapses, as well as in immune synapses. Agrin at the NMJ is a secreted multidomain 600 kDa proteoglycan, composed of a 200 kDa core protein, and GAG chains.

IIID.4. Neuronal development and regeneration.

CS proteoglycans (CSPGs) are critical regulators of the developing and regenerating nervous system. CSPGs have been known to serve dual – both enhancing or inhibitory roles in neuronal growth. CSPGs can enhance neurite outgrowth through interactions with bone-derived neurotrophic factor (BDNF), midkine (MK), or pleiotrophin (PTN). During embryonic development, CS glycans can guide the path of axons as they elongate and form synapses with neurons. Interestingly, growing axons avoid CSPGs in the developing growth cone,

indicating that CSPGs act as negative cues to guide axonal growth. The expression of CSPGs, such as neurocan, brevican, phosphaca, and versican, is significantly upregulated during adult injury, and CSPGs are also known to inhibit axon regeneration after an injury to the spinal cord. CSPGs also contribute to glial scarring post injury, and this scar acts as a barrier for new axons trying to reach the injury site. CSPGs also play crucial roles for the inability of the spinal cord for regeneration after injury, and some research has shown that chondroitinase enzymes that digest CS chains can restore some of this regeneration function. The phosphacan or receptor-type-protein-tyrosine phosphatase (PTP/RPTP) family of proteins is expressed in neurons and astrocytes, and is known to interact with CS chains to exert its inhibitory functions. Some sulfation patterns are highly preferred for binding, such as the disulfated CS-D motif, which strongly interact with PTNs.

IIIE. Glycocalyx proteoglycans and glycosaminoglycans in infection and immunity

Similar to mucins, proteoglycans at the glycocalyx can also serve dual functions in the progression of an infection. The GAG chains at the cell surface can serve as a point of entry or interaction with foreign agents to promote infection. On the other hand, soluble GAGs at the ECM can also serve a protective role, inhibiting the interaction of the foreign agents with their receptors.

IIIE.1. Sexually transmitted infections.

HS has a prime role in interactions with sexually transmitted infections. For instance, the herpes simplex virus type 1 (HSV1), or herpes, which commonly causes “cold sores,” but can also cause genital herpes, interacts with mammalian cells via glycocalyx HS chains.

Interestingly, an additional interaction of the HSV-1 glycoprotein gD with the glycocalyx 3-*O*-sulfate HS chains is required for viral entry. In HIV-1, treatment of T-cells with heparinase enzymes or a metabolic sulfation inhibitor (sodium chlorate), has been shown to inhibit HIV-1 binding and infection. These results indicate that glycocalyx HS chains participate in HIV-1 infection by facilitating the binding of HIV-1 virions to T cells, and that sulfation, in general, is important for this binding event. Subsequent studies revealed that it is the HIV-1 viral envelope glycoprotein gp120 that binds GAGs, and that although this binding event is dependent on general sulfation, gp120 prefers to bind to heparin and HS, compared to CS and dermatan. Similar relations with other surface receptor proteins have also been revealed in HPV, Chlamydia, Gonorrhea, Syphilis, and yeast infections.

IIIE.2. Other pathogens.

HA is a component of some pathogenic bacteria, including the extracellular capsule of group A *Streptococcus*. Capsular HA can promote the adhesion and colonization to the host tissue. Similar to other capsular polysaccharides, it can also aid in pathogen virulence by mimicking host HA, and thus shielding the pathogens from the action of host defenses (e.g. inhibiting phagocytosis and complement). Some bacteria (e.g. *Staphylococcus aureus*, *Clostridium perfringens*) can also synthesize hyaluronidase enzymes to utilize the host HA as a carbon nutrient source.

IIIE.3. The immune response.

The HSPG biglycan has recently been implicated as an endogenous key mediator of inflammation. During an inflammatory challenge, circulating macrophages have been found to engage in the *de novo* synthesis of biglycan. Biglycan acts as an agonist of the innate

immune Toll-like receptors TLR2 and 4 (TLR2/4), where it promotes receptor clustering with the purinergic P2X7/P2X4 receptors to rapidly activate the NLRP3 inflammasome. Thus, biglycan is a new DAMP (Damage-associated molecular pattern) that links the soluble matrix and the innate system. Biglycan and its analog, decorin, are master regulators of tumorigenesis, perhaps by binding other receptors such as LRP6. Several cancers are known to overexpress biglycan. This increased expression of GAG chains at the glycocalyx could point to enhanced metastasis and drug resistance.

In the lungs, GAGs are distributed along the bronchial walls and airway secretions. In this role, they and maintain the hydrated state and structure of the epithelial ECM. GAGs also modulate interactions with respiratory pathogens, as well as influence the resulting inflammatory response. Mast cells that release heparin are heavily enriched in the lungs. Heparin mostly resides within secretory granules in mast cells, but is released into the vasculature during injury. At these sites, heparin serves a defensive rather than an anticoagulant function, inhibiting the entrance of invading foreign pathogens.

IIIF. Proteoglycan Disorders

Given the importance of proteoglycans and GAGs in physiology, it is not surprising that abnormalities and pathophysiological conditions arise from mutations in genes involved in the biosynthesis of breakdown of proteoglycans. As described above, many mutations that result in a lack of GAGs are embryonic lethal. On the other hand, mutations involved in the catabolism of GAGs most often result in lysosomal storage disorders (LSDs), a group of 50 or so rare inherited disorders, mostly recessive, that are characterized by dysfunctional lysosomal breakdown. Mucopolysaccharidoses (MPS) are a group of lysosomal storage

disorders characterized by a lack of one of more eight GAG-degrading enzymes, resulting in an accumulation of pathological GAGs in the lysosome. The effects of MPS are devastating. Patients suffering from MPS exhibit brain and development abnormalities. Hunter syndrome is caused by mutations in MPS II, and MPSIIIA Sanfilippo A is specific for its lack of sulfamidase enzyme. Mutations in *HYAL1* encoding for a lysosomal hyaluronidase, also result in mucopolysaccharidosis IX. There are currently no therapeutic options for this group of diseases, but one promising method lies in enzyme replacement therapy (ERT), the delivery of enzymes that aid in the breakdown of GAGs to treat somatic symptoms.

Section 1.4 Acknowledgements

Chapter 1, in full, is a reprint of the material as it appears: Huang, M.L, Honigfort, D.; Godula, K. Carbohydrates - The Glycocalyx, and its Biological Roles. Cell Physiology Sourcebook (Elsevier 2018, Edna Kaneshiro (Eds), in press). The dissertation author is a co-author of this manuscript.

Section 1.5 References

Bond, M. R., & Hanover, J. A. A little sugar goes a long way: The cell biology of O-GlcNAc. *Journal of Cell Biology* (2015) 208:869.

Corfield, A. P. Mucins: A biologically relevant glycan barrier in mucosal protection. *Biochim. Biophys. Acta BBA - Gen. Subj.* (2015) 1850:236–252.

Derrien, M., van Passel, M.W.J., van de Bovenkamp, J. H. B., Schipper, R. G., de Vos, W. M., Dekker, J. Mucin-bacterial interactions in the human oral cavity and digestive tract. *Gut Microbes* (2010) 1:254–268.

Gama, C. I., Tully, S. E.; Sotogaku, N.; Clark, P. M.; Rawat, M., Vaidehi, N., Goddard, W. A. III, Nishi, A., Hsieh-Wilson, L. C. Sulfation patterns of glycosaminoglycans encode molecular recognition and activity. *Nature Chemical Biology* (2006) 2:467-473

Guerrini, M., Beccati, D., Shriver, Z., Naggi, A. VIswanathan, K., Basio, A., Capila, I., Lansing, J. C., Guglieri, S., Fraser, B., Al-Hakim, A., Gunay, N. S., Zhang, Z., Robinson, L., Buhse, L., Nasr, M., Woodcock, J., Langer, R., Venkataraman, G., Linhardt, R. J., Casu, B., Torri, G., Sasisekharan, R. Oversulfated chondroitin sulfate is a contaminant in heparin associated with adverse clinical events. *Nature Biotechnology* (2008) 26:669-675.

Hattrup, C. L. & Gendler, S. J. Structure and Function of the Cell Surface (Tethered) Mucins. *Annu. Rev. Physiol.* (2008) 70:431–457.

Hollingsworth, M. A. & Swanson, B. J. Mucins in cancer: protection and control of the cell surface. *Nat. Rev. Cancer* (2004) 4:45–60.

Kim, K. C. Role of epithelial mucins during airway infection. *Pulm. Pharmacol. Ther.* (2012) 25: 415–419.

Kufe, D. W. Mucins in cancer: function, prognosis and therapy. *Nat. Rev. Cancer* (2009) 9:874–885.

Lawrence, R.; Lu, H.; Rosenberg, R. D.; Esko, J. D.; Zhang, L. Disaccharide structure code for the easy representation of constituent oligosaccharides from glycosaminoglycans. *Nat. Methods* (2008), 5:291-292.

McGuckin, M. A., Lindén, S. K., Sutton, P. & Florin, T. H. Mucin dynamics and enteric pathogens. *Nat. Rev. Microbiol.* (2011) 9:265–278.

Miller, G. M., & Hsieh-Wilson, L. C. Sugar-dependent modulation of neuronal development, regeneration, and plasticity by chondroitin sulfate proteoglycans. *Exp. Neurol (2015)* 274:115-125.

Neill, T., Schaefer, L., Iozzo, R. V. Decoding the matrix: Instructuve roles of proteoglycan receptors. *Biochemistry* 2015, 54:4583.

Paszek, M. J., DuFort, C. C., Rossier, O., Bainer, R., Mouw, J. K., Godula, K., Hudak, J. E., Laking, J. N., Wijekoon, A. C., Cassereau, L., Rubashkin, M. G., Magbanua, M. J., Thorn, K. S., Davidson, M. W., Rugo, H. S., Park, J. W., Hammer, D. A., Giannone, G., Bertozzi, C. R., Weaver, V. M. The cancer glycocalyx mechanically primes integrin-mediated growth and survival. *Nature* (2014) 511:319–325.

Shukla, D., Liu, J., Blaiklock, P., Shworak, N. W., Bai, X., Esko, J. D.; Cohen, G. H., Eisernberg, R. J.; Rosenberg, R. D., Spear, P. G. A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry. *Cell* 1999, 99:13-22.

Taylor, M. E., & Drickamer, K. Introduction to Glycobiology, Third Ed. Oxford University Press. New York, 2011.

Thacker, B. E., Xu, D., Lawrence, R., Esko, J. D. Heparan sulfate 3-O-sulfation: a rare modification in search of a function. *Matrix Biol* (2014) 35:60-72.

Tiwari, V., Maus, E., Sigar, I. M., Ramsey, K. H., Shukla, D. Role of heparan sulfate in sexually transmitted infection.

Yan, D. & Lin, X. Shaping morphogen gradients by proteoglycans. *Cold Spring Harb Perspect Biol.* 2009, 1:a002493.

Chapter 2: Engineering of spectator glycocalyx structures to evaluate molecular interactions at crowded cellular boundaries.

In the mucosal epithelium, the cellular glycocalyx can project tens to hundreds of nanometers into the extracellular space, erecting a physical barrier that provides protective functions, mediates the exchange of nutrients and regulates cellular interactions. Little is understood about how the physical properties of the mucosal glycocalyx influence molecular recognition at the cellular boundary. Here, we report the synthesis of PEG-based glycopolymers with tunable glycan composition, which approximate the extended architecture of mucin glycoproteins, and tethered them to the plasma membranes of red blood cells (RBC) to construct an artificial mucin brush-like glycocalyx. We evaluated the association of two lectins, ConA and SNA, with their endogenous glycan ligands on the surface of the remodeled cells. The extended glycocalyx provided protection against agglutination of RBCs by both lectins; however, the rate and magnitude of ConA binding were attenuated to a greater degree in the presence of the glycopolymer spectators than those measured for SNA. The different sensitivity of ConA and SNA to glycocalyx crowding likely arises from the distinct presentation of their mannoside and sialoside receptors, respectively, within the native RBC glycocalyx.

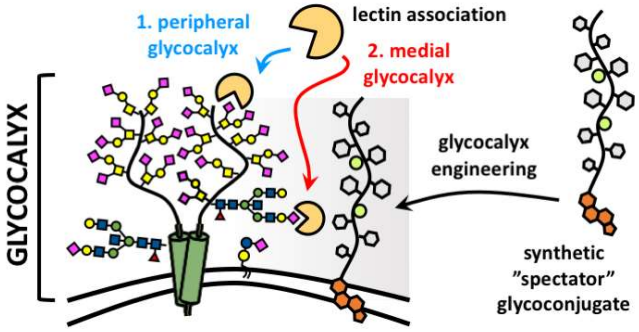


Figure 2.1. Membrane engineering with bystander glycocalyx structures reveals altered protein-receptor association in crowded cell surface environments.

Section 2.1 Introduction

The cellular glycocalyx is a dynamic carbohydrate-rich macromolecular system populating the boundary between a cell and its surroundings.[1] It is composed of glycolipids and glycoproteins, which can extend tens to hundreds of nanometers above the plasma membrane, providing cells with a physical protective barrier while also facilitating cellular interactions with its environment (**Fig 2.1**). The glycan structures distributed throughout the glycocalyx are recognized by a wide array of glycan binding proteins (GBPs), including adhesion molecules found in the extracellular matrix and on neighboring cells, lectins and antibodies of the immune system, or signaling proteins such as growth factors and cytokines.[2] The organization of glycans within glycoconjugates into multivalent ensembles overcomes the characteristically weak binding between individual glycans and their protein partners and produces molecular patterns within the glycocalyx that elicit high-avidity recognition by GBPs.[3]

Macromolecular glycoconjugates isolated from natural sources[4] or generated synthetically[5] have provided important tools for investigating and formulating the paradigms that define our current understanding of how avidity in glycan-protein interactions is achieved.[6] Employed as soluble ligands or integrated into glycan arrays, surface plasmon resonance or biolayer interferometry platforms, synthetic materials (e.g., glycoclusters, spherical glycodendrimers, or linear glycopolymers) have provided key insights into how ligand parameters, such as glycan valency and spacing or scaffold architecture, influence GBP binding and higher-order association (e.g., oligomerization or crosslinking).[7,8] More recently, cell surface engineering approaches have begun to emerge that allow for the integration of these materials directly into the glycocalyx of cells[8,9] to further evaluate how

the compositional heterogeneity, dynamics and nanoscale organization of the glycocalyx influence biological interactions at the cell surface.

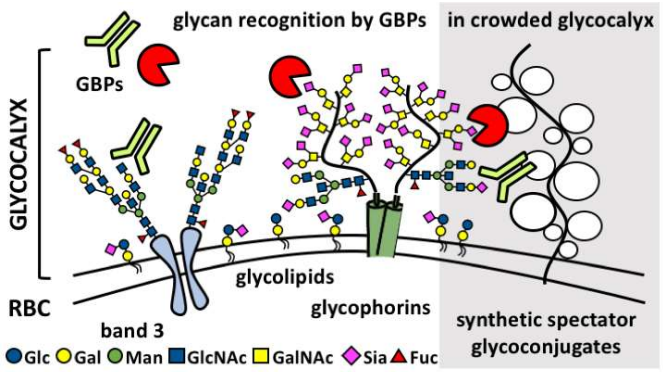


Figure 2.2. Recognition of glycans by protein receptors occurs in the dynamic and compositionally heterogeneous environment of the cellular glycocalyx. These interactions can be influenced by the presence and physical properties of non-binding, spectator, glycoconjugates. Synthetic mimetics of membrane associated glycoproteins, such as mucins, can be generated and installed into the native glycocalyx to evaluate how changes in glycocalyx crowding and physical properties impact ligand-receptor interactions.

While most work in this area has focused primarily on introducing constellations of specific glycan structures within the glycocalyx and evaluating their recognition by cognate receptors, the effects of the bulk properties of the glycocalyx on cellular interactions and functions are also becoming increasingly pursued (**Fig 2.2**). For instance, the remodelling of plasma membranes of mammary epithelial cells with lipidated glycopolymer mimetics of mucin glycoproteins was found to exert biomechanical forces that promoted the assembly of integrin adhesion complexes and enhanced signaling associated with cell proliferation, thus identifying a glycocalyx-dependent mechanism for enhancing the survival and metastatic potential of circulating tumor cells.[10] Further studies using more minimalistic models of bulky glycocalyx components (i.e, linear polyethylene glycol and hyperbranched polyglycerol polymers) covalently grafted to membrane proteins of red blood cells (RBCs) afforded

shielding of ABO antigens and protection from complement-mediated cell lysis according to polymer size[11,12] and localization within the glycocalyx.[13]

Despite these advances, systematic surveys of the influence of the "spectator" components of the cellular glycocalyx, with respect to the composition, physical properties and organization of glycans, on molecular recognition events at the cell surface are still lacking. Here, we report the synthesis of well-defined linear poly(ethylene glycol) (PEG) backbones that approximate the architecture and properties of mucin glycoproteins and their use for the scaffolding of spectator glycocalyx structures with varied monosaccharide composition at the surfaces of RBCs. The resulting mucinous glycocalyx models exerted differential effects on lectin association with imbedded endogenous glycan receptors according to the structure of the spectator glycoconjugate and lectin identity.

Section 2.2 Results and Discussion

Synthesis of mucin-mimetic glycocalyx building blocks. We initiated our study by establishing a synthetic strategy to generate glycopolymer structures that exhibit key architectural features of mucin glycoproteins while allowing for the facile introduction of a range of glycan structures. Mucins are large, heavily glycosylated proteins populating the surfaces of epithelial cells, and are composed of characteristic variable number tandem repeat regions with a high frequency of *O*-glycosylation sites.[14] The close packing of glycans along the polypeptide chain forces the protein into an extended conformation with high persistence lengths.[15] This property of mucins can be recapitulated in synthetic linear glycopolymers by controlling the positioning of glycan appendages in sufficient density and close proximity to the polymer backbone.

The Bertozzi group has developed two mucin-mimetic polymer architectures based on glycosylated poly(methylvinyl ketone) (pMVK) and poly(serine) scaffolds generated by the controlled reversible addition-fragmentation chain transfer (RAFT)[16] and the ring-opening *N*-carboxyanhydride (NCA)[17] polymerization techniques, respectively. Both approaches afforded functional mucin-like polymeric structures with distinct advantages and drawbacks. While the pMVK scaffold allows for post-polymerization modification with a wide array of glycan structures, its utility is somewhat limited due to the hydrophobicity of the core polyhydrocarbon backbone, which at low glycan content may give rise to amphiphilic behaviour. The polypeptide-based approach obviates this problem; however, it necessitates the synthesis and purification of glycosylated NCA monomers, which can be challenging.

We decided to address these challenges by building mucin mimetics on hydrophilic PEG scaffolds suitably modified to allow for later stage introduction of glycans, as well as additional functional elements such as surface anchors and optical probes for characterization. We decided to target a set of glycopolymers displaying the monosaccharides glucose (Glc), galactose (Gal), fucose (Fuc), and glucuronic acid (GlcA) (**Fig 2A**). The sugars were selected based on their common presence in mammalian glycans and their distinct chemical and physical properties. For instance, Glc and Gal share the same molecular composition, but differ in the orientation of the C4 hydroxyl group, which may impact their packing along the polymer chain. In addition, the increased hydrophobicity of Fuc compared to Glc and Gal (logP values of -2.02, -3.24 and -3.38, respectively) and the negative charge of GlcA at physiological pH (pKa = 3.2), may differentially affect the persistence length and physical properties of the resulting glycopolymers.

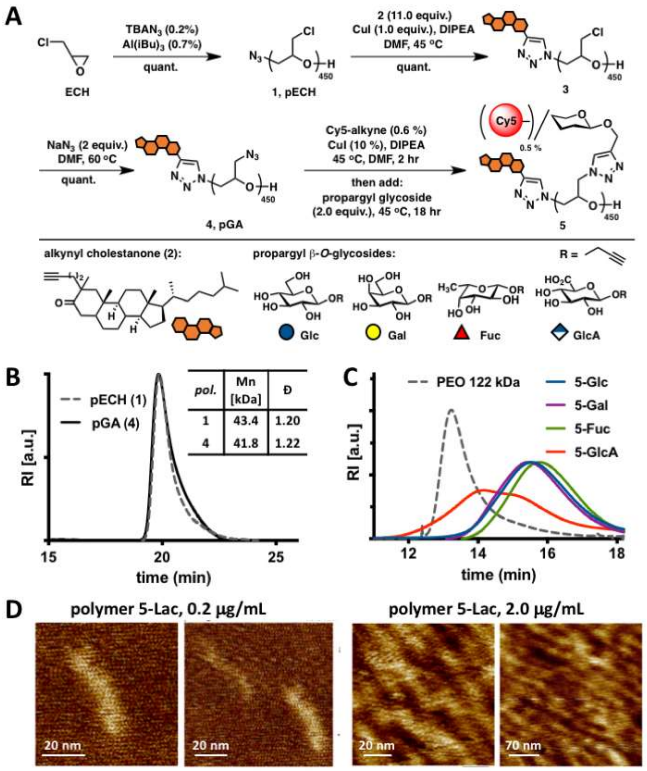


Figure 2.3. Synthesis and characterization of mucin mimetics 5. (A) Glycopolymers **5** were generated from poly(epichlorohydrin) **1** via chloride-to-azide sidechain substitution followed by the CuAAC conjugation of propargyl β-*O*-glycosides (Glc, Gal, Fuc, and GlcA). The polymers were furnished with a hydrophobic cholestanone moiety for anchoring into cell membranes and a Cy5 fluorescent tag for imaging and quantification. (B) SEC analysis indicated narrow molecular weight and chain-length distributions before and after the chloride-to-azide exchange. (C) Glycopolymers **5** exhibited increased retention in aqueous SEC compared to a PEO standard of similar molecular weight (122 kDa), a characteristic behavior of glycopolymers with extended molecular conformations. (D) AFM imaging of lactose-modified glycopolymers confirmed elongated, mucin-like morphology of the PEG-based glycopolymers

Using a monomer-activated anionic ring opening polymerization developed by Carlotti and colleagues,[18] we have synthesized an azide-terminated poly(epichlorohydrin) (pECH) polymer (**1**, **Fig 2A**) with well-defined size (Mn = 43,400 Da, DP ~450) and narrow chain-length distribution (Đ = 1.20). Although the use of a racemic epichlorohydrin monomer produces an atactic pECH polymer **1**, we reasoned that the stereochemical relationships of the pendant sidechains would not significantly impact the overall architecture and physical

properties of the desired mucin mimetic glycopolymers. The polymeric precursor **1** was elaborated into fluorescent cell surface-targeting mucin mimetics in a three-step synthetic sequence (**Fig 2A**). First, we introduced a hydrophobic anchor for plasma membrane targeting via the copper-catalysed alkyne-azide cycloaddition (CuAAC)[19] between the chain-end azide group in **1** and an alkynyl cholestanone derivative **2**[20] (11.0 equiv. per chain-end azide). Treatment of the chloromethyl sidechains in the cholestanone modified pECH polymer **3** with sodium azide[21] (2.0 equiv. per Cl) generated a reactive polymer intermediate **4** primed for a sequential CuAAC conjugation of alkynyl fluorophores (Cy5, ~ 0.5 % sidechains) and propargyl glycosides to complete the desired glycopolymers **5** (**Fig 2A**). ^{1}H NMR spectroscopy and size exclusion chromatography (SEC) analyses of polymeric intermediate **4**, confirmed quantitative azide-chloride replacement without any observable increase in polymer dispersity (**Fig 2B** and **Fig S8**). The levels of fluorophore labelling in glycopolymers **5** were assessed based on UV-VIS absorption profiles and matched values predicted based on reaction stoichiometry. The efficiency of glycan conjugation to the polymer was difficult to establish accurately based on ^{1}H NMR spectroscopy alone, due to the overlap between glycan and polymer proton signals; however, IR spectroscopy analysis of glycopolymers **5** revealed the disappearance of the characteristic azide stretching absorption frequency at $\upsilon = 2100$ cm^{-1} (**Fig S17**) after conjugation, indicating quantitative side-chain modification.

We anticipated that the close positioning of the pendant glycan residues with respect to the polymer backbone would force the glycopolymers into an extended conformation. This behaviour is supported by aqueous SEC analysis of polymers **5** exhibiting significantly increased retention on the stationary phase compared to a PEO standard of similar molecular weight (**Fig 2C**, $M_{w,PEO} = 122$ kDa). Atomic force microscopy (AFM) of glycopolymers carrying larger lactose disaccharide residues to facilitate visualization provided direct evidence for

extended polymer structures (**Fig 2D** and **Fig S18**). We were able to clearly observe individual glycopolymer molecules with length distributions measured at 41 ± 6 nm (**Fig 2D**) confirming that the new glycopolymer architecture can mimic the nanoscale topology of native mucins.

Generation of spectator glycocalyx structures on RBCs. Molecular recognition events at the cell surface occur in the context of the glycocalyx, which often defines the properties of the surrounding microenvironment, such as polarity, pH, permittivity, and crowding, that may influence the energetics of binding as well as molecular diffusion and transport. However, these properties of the glycocalyx environment are rarely considered when evaluating molecular interactions at the cell surface, despite observations that, for instance, the concentrations of the negatively charged sialic acids in the glycocalyx of B cells can reach ~ 100 mM[22] or that transmembrane epithelial mucins produce dense glycoprotein brushes spanning lengths up to ~1,500 nm.[23]

Some of these conditions can be modelled in a controlled fashion at the surfaces of cells by augmenting their endogenous glycocalyx with synthetic glycomaterials. For instance, glycopolymers affixed to the outer leaflet of the plasma membrane via a lipid anchor and exhibit dynamic behaviour, such as lateral membrane diffusion and spatial reorganization in response to cell adhesion forces,[9,10] and to display molecular recognition motifs to promote cell-cell interaction and intra-cellular signalling.[24] We set out to construct a synthetic glycocalyx on the surfaces of RBCs to model the barrier functions of the mucinous epithelial glycocalyx using glycopolymers **5** (**Fig 3A**). RBCs offer a practical choice for this purpose, due to their relatively modest glycocalyx, which extends less than 10 nm above the plasma membrane, the lack of active endocytosis preventing the uptake and clearance of the exogenously added materials, and their amenability to a range of protein-binding assays.

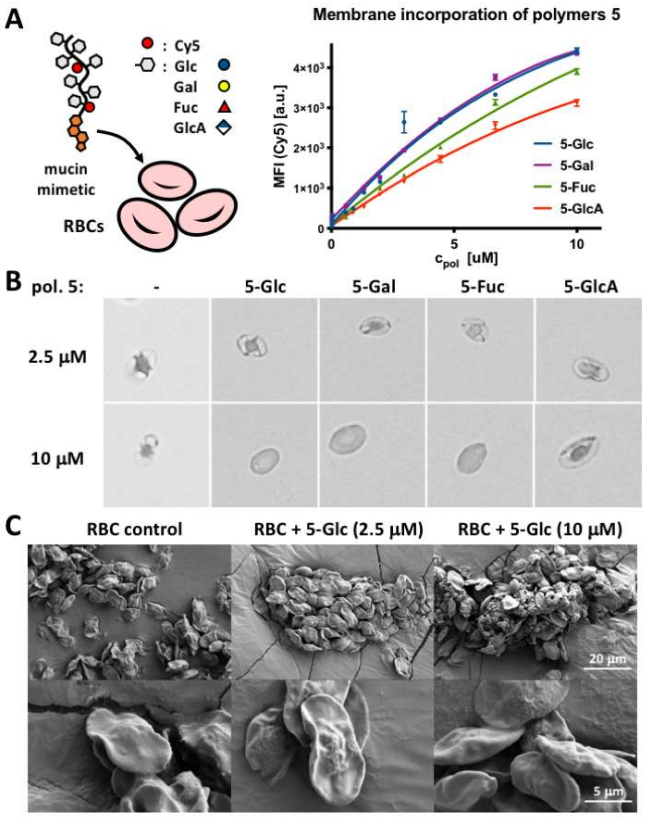


Figure 2.4. RBC glycocalyx remodeling with glycopolymers 5. (A) Polymers **5** incorporate dose-dependently into RBC membranes according to their pendant glycans. Increasing glycopolymer density in RBC membranes induces cell rounding, swelling, and lysis at polymer concentrations above 2.5 µM via brightfield microscopy (B) or SEM (C).

The treatment of RBCs with Cy5-labelled glycopolymers **5** (c_{pol} = 115 nM - 10.0 µM), resulted in a dose-dependent membrane incorporation of the polymers based on flow cytometry analysis (**Fig 3A**). Incubation of RBCs with **5-Glc** lacking the cholestanone anchor gave no appreciable cell-surface incorporation of the glycopolymer (**Fig S20**), indicating that the synthetic introduction of the anchor into the glycopolymers via precursor **3** is successful and necessary for their plasma membrane insertion. As anticipated, the polyanionic GlcA-modified polymer exhibited lower membrane grafting efficiency (~65%) at the highest polymer concentration (10.0 µM) compared to **5**-Glc and **5**-Gal polymers. Interestingly, we

also observed somewhat decreased (~12%) grafting for the polymer bearing fucose (**5**-Fuc), which is a more hydrophobic C6-deoxy analogue of L-galactose (**Fig 3B**). Visual inspection of the remodelled RBCs in a 96 well plate indicated a decreased ability of cells treated with the non-ionic Glc, Gal and Fuc glycopolymers to settle to the bottom of the wells starting at polymer concentrations of ~ 1.8, 3.2, and 5.6 μM, respectively (**Fig S22**). We saw no changes in sedimentation properties for RBCs treated with the polyanionic **5**-GlcA polymer over the entire range of concentrations (115 nM - 10.0 μM). These observations are consistent with prior findings showing that increasing the size and surface density of non-ionic polymers, such as methoxypolyethylene glycol (mPEG), covalently grafted to proteins on RBCs resulted in the shielding of cell surface charge and prevented cell sedimentation and plasma protein-induced cell stacking (Rouleaux formation).11

Increasing the surface density of the mucin-mimetic glycopolymers on RBCs may introduce membrane deformations due to entropic pressures resulting from mushroom to brush transitions in the compressing polymer brush.[25] Indeed, we observed visible rounding of RBCs at polymer concentrations above 2.5 μM with signs of membrane disintegration in cells treated with 10.0 μM polymers using optical microscopy (**Fig 3B** and **Fig S21**). To exclude contributions from altered membrane morphologies to protein binding in the engineered glycocalyx structures, we imaged the remodelled cells using scanning electron microscopy (SEM, **Fig 3C**). We observed no significant effects on cell shape or membrane structures for polymer treatments at concentrations up to 2.5 μM and confirmed the breakdown of RBCs at 10.0 μM polymer concentration (**Fig 3C**).

Our analyses indicate that RBCs can be effectively remodelled with mucin mimetics **5** to introduce synthetic glycocalyx structures to the cell surface. The efficiency of polymer incorporation and the resulting physical properties of the remodelled RBCs were strongly

influenced by the structure of the monosaccharide component in **5,** and too high polymer density at the plasma membrane (c_{pol} > 2.5 μM) led to observable rounding of cells and an eventual cell lysis. Based on these findings, we identified an optimal polymer concentration of 2.5 μM for RBC remodelling, which provides maximal polymer incorporation without adverse effects on membrane morphology.

Effects of spectator glycocalyx structures on lectin interactions with RBCs. The influence of the compositional heterogeneity and organization of the glycocalyx microenvironment on ligand-receptor interactions at the cell surface have recently been drawing attention, as glycans are increasingly considered for therapeutic targeting.[26] For instance, glycan array screens have identified avidity enhancements for the binding of anti-glycan antibodies in the presence of neighboring non-binding glycan epitopes.[27,28] Similarly, ABH blood group antigens presented on the extracellular domains of band 3 proteins on RBCs have been found to modulate the binding of proteins, such as the *Sambucus nigra* agglutinin (SNA) or the human Siglec-2, to sialic acid receptors primarily carried by glycophorins (**Fig 1**).[29] The localization of the A and B antigens either to the periphery or the center of sialylated glycoprotein clusters in the RBC membranes, respectively, resulted in a differential stabilization of sialoglycan patches and changes in their binding avidity toward SNA and Siglec-2. The phenomenon of glycan clustering is not limited to RBCs and has been found to contribute to a range of biological processes, including host pathogen interactions, tumor antigen recognition and T-cell activation via a mucinous glycosynapse.[30] The mucin-rich glycocalyx covering epithelial cells is believed to provide a semi-permeable physical barrier limiting the diffusion of proteins and particulates to the cell surface and protecting membrane protein structures from chemical and enzymatic degradation.14 Mucins

contribute directly to the organization of glycocalyx structures and drive the clustering of cell adhesion molecules by exerting forces against the surrounding extracellular environment.10 We set out to evaluate how the presence of a mucin-like glycocalyx brush at the cell surface would influence the interactions of lectins with endogenous glycan receptors. We selected two lectins, Concanavalin A (ConA) and SNA, for their respective specificity in the targeting of two distinct glycan classes prominently represented in the RBC glycocalyx, the oligomannose core of *N*-linked glycans found primarily on band 3, and to a smaller extent, on glycophorins, and the α2,6-linked sialic acids displayed prominently at terminal positions of *O*-linked glycans on glycophorins and glycosphingolipids.

First, we investigated the ability of glycopolymers **5** to protect against RBC agglutination with both lectins. Using agglutination assays with non-remodeled RBCs, we identified the minimal lectin concentrations (c_{agg}) required to induce cell crosslinking and aggregation (**Fig S23** and **Fig S24**). SNA showed ~ 17-fold greater agglutination capacity ($c_{agg,SNA}$ = 33.0 ng/mL) compared to ConA ($c_{agg,ConA}$ = 0.57 µg/mL). This may, presumably, be due to the greater number and accessibility of sialic acid modifications within the glycocalyx compared to the *N*-linked core mannose structures targeted by ConA.

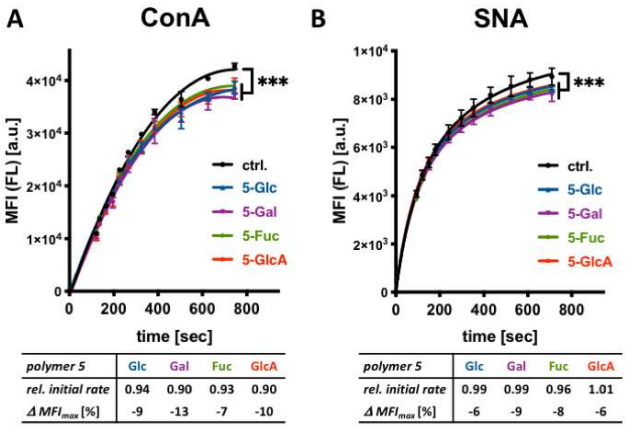
We next examined the ability of polymers **5** to shield RBCs from agglutination by SNA and ConA by determining the changes in the minimal agglutination concentration, (c_{agg}), required for each lectin to agglutinate RBCs remodeled with polymers **5** (c_{pol} = 2.5 µM, (Table 1, **Fig S23** and **Fig S24**). All glycopolymers **5** provided protection against agglutination by both lectins in the following order: Glc < Gal ~ Fuc < GlcA (Table 1). It should be noted that the polyanionic glycopolymer, **5**-GlcA, afforded the most shielding effect (3- to 4-fold increase in c_{agg}), despite its lower density in the glycocalyx (~65% surface grafting efficiency of the non-ionic polymers, **5**-Glc and **5**-Gal, Fig S19). Each polymer within the set showed similar

level of shielding against agglutination by both lectins, as measured by the relative change in c_{agg} with respect to non-remodeled cells. This suggests that the bulk properties of the spectator glycocalyx rather than lectin identity provides the dominant contribution toward protection against agglutination in these experiments. Thus, all polymers established a physical barrier against crosslinking with additional repulsive forces arising from increased negative charge density at the surface of cells presenting the glucuronic acid polymer, **5**-GlcA.

Table 2.1. Agglutination of RBCs remodeled with glycopolymers **5** (2.5 μM) by ConA (0 – 1.8 μg/mL) and SNA (0-105 ng/mL) lectins.

ConA						
C_{ConA} [μg/mL]	0 - 0.43	0.57	0.76	1.00	1.30	1.80
glycopolymer 5 -						
glycopolymer 5 Glc						
glycopolymer 5 Gal						
glycopolymer 5 Fuc						
glycopolymer 5 GlcA						

SNA									
C_{SNA} [ng/mL]	0 - 25	33	45	60	80	105	140	187	250
glycopolymer 5 -									
glycopolymer 5 Glc									
glycopolymer 5 Gal									
glycopolymer 5 Fuc									
glycopolymer 5 GlcA									



polymer 5	Glc	Gal	Fuc	GlcA
rel. initial rate	0.94	0.90	0.93	0.90
Δ MFI_{max} [%]	-9	-13	-7	-10

polymer 5	Glc	Gal	Fuc	GlcA
rel. initial rate	0.99	0.99	0.96	1.01
Δ MFI_{max} [%]	-6	-9	-8	-6

Figure 2.5. Association of ConA and SNA lectins with glycocalyx-remodeled RBCs. (A) Binding of fluorescein (FL)-labeled lectins to remodeled cells were assessed via flow cytometry. The presence of spectator glycopolymers **5** at the surface of RBCs attenuates both the initial rate as well as saturation binding of ConA to cell surface glycans. (B) Glycopolymers **5** have no effect on the initial rates of SNA binding but inhibit lectin association near saturation. Relative initial rates were calculated from the linear regions of lectin binding curves and normalized to control cells without polymer treatment. Δ MFI_{max} corresponds to the change in median fluorescence intensity of cells at saturation lectin binding. (ANOVA, Tukey's multiple comparisons test; $p^{***} < 0.001$).

These observations led us to investigate whether the spectator glycoconjugates might exert an influence on lectin association within the glycocalyx at sub-agglutination concentrations. Using a flow-cytometry assay, we measured the binding of ConA and SNA to the remodeled RBCs (**Fig 4**). At ConA and SNA concentrations of 8 µg/mL and 1 ug/mL, respectively, saturation binding was attained for both lectins in ~ 600 sec without any apparent cell aggregation. We observed 7-13% and 6-9% decrease in the total amount of bound ConA and SNA, respectively, at saturation in the presence of polymers **5** compared to untreated control cells (**Fig 4**). Distinct from the results obtained from hemagglutination experiment, the structure of the glycan appendages in glycopolymers **5** had no significant impact on the magnitude of inhibition of lectin binding; however, all polymers appeared to inhibit ConA binding to the RBC surface to a greater extent compared to SNA.

Analysis of the linear regions of the binding curves revealed ~ 6-10% decrease in initial rates for the association of ConA in the presence of the polymers compared the non-remodelled RBCs (**Fig 4** and **Fig S26**). In contrast, the initial rates for SNA association in the presence of all polymers **5** were indistinguishable from those observed in control cells (**Fig 4** and Fig **S26**), with differences beginning to appear at a later time closer to the point of saturation binding.

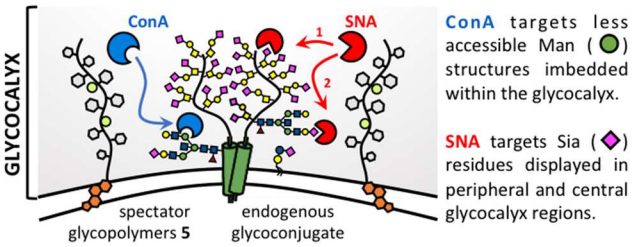


Figure 2.6. Glycocalyx crowding with spectator glycoconjugates differentially affects the association of ConA and SNA with endogenous glycans depending on their distribution throughout the glycocalyx. The glycopolymer spectators inhibit both the initial rates and saturation binding of ConA to the less accessible mannose structures. Glycocalyx crowding influences the association of SNA with sialic acid residues only at later time points near saturation binding, once the more available peripheral glycans have become occupied.

Collectively, these observations indicate distinct influence of the spectator glycocalyx on the association of the two lectins (**Fig 5**). The ConA lectin targets the less accessible mannoside residues of *N*-linked glycans placed in close proximity to the polypeptide chains of glycoproteins and generally positioned deeper within the glycocalyx. As such, the lectin diffusion into the glycocalyx and binding to its glycans is likely to be more sensitive to the presence of crowding spectator glycoconjugates. This is reflected in the initial rates of ConA binding to RBCs in the presence of glycopolymers **5,** and the greater decrease in the total amount of ConA bound at saturation. In contrast, the SNA lectin targets the outermost sialic acid residues in both *N*- and *O*-linked glycans, which are prominently displayed at the periphery of the glycocalyx and, thus, more accessible. As such, the spectator glycoconjugates would not be expected to interfere with SNA binding until all available peripheral glycans are occupied and the diffusion of the lectin deeper into the glycocalyx is required for additional binding. The observed rapid onset of SNA binding with initial rates indistinguishable between polymer-remodeled RBCs and untreated controls followed by spectator-induced inhibition of binding at a later timepoint supports this model. As the binding of the lectins reaches saturation, the effects of the spectator glycocalyx would be

expected to become exacerbated due to the increasing crowding from the newly introduced proteins.

Section 2.3 Conclusions

In this paper, we have described a new synthetic route to generate glycopolymer mimetics of mucin glycoproteins with well-defined architectures and tuneable glycosylation. Using a non-covalent cell membrane engineering approach, we have delivered these materials to the surfaces of red blood cells to augment the physical properties of their glycocalyx and evaluated the effect of increased molecular crowding on lectin recognition of endogenous glycan receptors. While the polymers attenuated the binding of both lectins near saturation, only the initial rates of ConA association were affected due to increased glycocalyx crowding in the presence of the synthetic glycoconjugate spectators. The unequal sensitivity of ConA and SNA can be rationalized based on the differential accessibility of their respective glycan targets within the glycocalyx.

Section 2.4 Methods

Materials. All chemicals, unless stated otherwise, were purchased from Sigma Aldrich and used as received. Reaction progress was checked by analytical thin-layer chromatography (TLC, Merck silica gel 60 F-254 plates) monitored either with UV illumination, or by staining with iodine, ninhydrin, or CAM stain. Solvent compositions are reported on a volume/volume (v/v) basis unless otherwise noted. Alkynyl cholestanone **2**,20 and Glc,[31] Gal,31 Fuc,[32] GlcA,[33] and Lac[34] propargyl glycosides were prepared according to published procedures. Turkey Red Blood Cells as a 10% solution were obtained from Lampire Biological Laboratories (cat # 724908). *Sambucus nigra* agglutinin (SNA) and Concanavalin A (ConA) lectin were purchased from Vector Labs. Fluorescein (FL)-NHS for lectin labeling was purchased from

Thermo Scientific and Cyanine 5 (Cy5)-alkyne for labeling of polymers was obtained from Sigma Aldrich (fluorophore structures are shown in SI on page S4). Lectins were labeled according to manufacturer protocols and the extent of labeling and lectin concentration was adetermined by UV-VIS spectroscopy and BCA assay, respectively. A detailed list of analytical instruments and general procedures used for the purification and structural characterization of synthetic materials and for polymer and cell imaging (optical microscopy, AFM, and SEM) can be found in the *Supporting Information*.

Synthesis of azide-terminated poly(epichlorohydrin), pECH (1). Epichlorohydrin was polymerized according to the procedure developed by Carlotti.[18] Briefly, a flame-dried Schlenk flask (10mL) equipped with a magnetic stirrer and fitted with PTFE stopcock was charged with tetrabutylammonium azide ($TBAN_3$, 20 mg, 0.037 mmol,0.002 equiv.) under argon. A solution of freshly distilled epichlorohydrin (1.29 mL, 16.5 mmol) in anhydrous toluene (4 mL). A solution of triisobutylaluminum in toluene (1.07 M, 104 µL, 0.111 mmol, 0.007 equiv.) was added via a syringe under argon at −30 °C. The reaction was stirred for 4 hours and then stopped by the addition of ethanol. The resulting pECH polymer **1** was precipitated into hexanes and dried under vacuum to yield a clear viscous oil (1500 mg, 99 % yield). The polymer was analyzed by SEC (0.2% LiBr in DMF): M_w = 52,000, M_n = 43,400, Đ = 1.20).

Synthesis of cholestanone-terminated poly(epichlorohydrin) (3). In a flame-dried Schlenk flask (10 mL), pECH polymer **1** (15 mg, 0.3 µmol) was dissolved in degassed anhydrous DMSO (200 µL). Alkynyl cholestanone **2** (1.7 mg, 3.8 µmol, 11.0 equiv.) was added, followed by CuI (~0.05mg, 0.3 µmol, 1.0 equiv.) and one drop diisopropylethyl amine (DIPEA, ~ 5 µL). The reaction was stirred at 40°C for 12 hr, at which time it was quenched by the addition of water

to precipitate the polymer. The resultant polymer was triturated with hexanes to remove unreacted **2** and dried on vacuum to yield a clear viscous oil in (16 mg, 100% yield). The polymer was analyzed by SEC (0.2% LiBr in DMF): M_w =52,000, M_n =43,000, Đ = 1.20.

Synthesis of cholestanone-terminated poly(glycidyl azide), pGA (4). The chloride to azide exchange in pECH polymer **3** was accomplished according to previously published procedure.21 Briefly, in a flame-dried Schlenk flask (10 mL), polymer **3** (15 mg, 0.16 mmol) was dissolved in dry DMF (300 µL). To the solution was added NaN_3 (21 mg, 0.32 mmol, 2.0 equiv.), and the reaction was stirred at 60 °C for 3 days under argon to allow complete conversion. Polymer solution was filtered and precipitated in ethanol to yield a clear viscous oil (16 mg, 100 % yield). The polymer was analyzed by SEC (0.2% LiBr in DMF): M_w = 51,000, M_n = 41,800, Đ = 1.22.

Synthesis of Cy5-labeled glycopolymers 5. In a flame-dried Schlenk flask (10 mL), polymer **4** (9.00 mg, 0.09 mmol) was dissolved in degassed dry DMSO (250 µL). To the solution was added Cy5-alkyne (0.50 mg, 0.50 µmol) in DMSO (50 µL), followed by CuI (2.00 mg, 9.00 µmol) and DIPEA (16 uL, 0.09 mmol). The reaction was stirred in dark under Ar at 40 °C for 2 hrs. After this time, the reaction mixture was aliquoted (50 µL) into separate vials containing □-propargyl glucoside, galactoside, fucoside, and glucuronoside (0.03 mmol, 2.00 equiv. per azide side-chain) in degassed anhydrous DMSO (50 µL). The reactions were stirred in dark at 40°C overnight. After this time, the reactions were diluted with DI water and treated with Cuprisorb beads (SeaChem labs) for 18 hrs to sequester copper. The resulting copper-free solutions were filtered through celite to remove the resin and lyophilized. The dry residues were triturated 3x with methanol with monitoring by TLC to remove excess unreacted glycosides. The resulting Cy5-labeled glycopolymers **5** were dissolved in D_2O and

lyophilized to give a blue solid in a quantitative yield for each polymer (*note*: the blue color of the glycopolymers arises from the presence of the Cy5 label and not residual copper contamination. Glycopolymers lacking the Cy5 label were isolated as white solids). The polymers **5** were characterized using ^{1}H NMR ($CDCl_3$, 300 MHz) and UV-Vis (λ_{max} = 633 nm) spectroscopy and analyzed by aqueous SEC. Absorbance readings at known concentrations of glycopolymers **5** indicating the presence of ~ 2 Cy5 molecules per polymer chain (0.5% sidechain occupancy). The resulting glycopolymers were analyzed by aqueous SEC (0.2M $NaNO_3$ in 0.01M Na_2HPO_4, pH = 7.0). **5 Glc**: $M_{n,calc.}$ = 143 kDa, $M_{n,SEC}$ = 20,578, Đ = 1.369; **5-Gal**: $M_{n,calc.}$ = 143 kDa, $M_{n,SEC}$ = 21,273, Đ = 1.374; **5-Fuc**: $M_{n,calc.}$ = 136 kDa, $M_{n,SEC}$ = 18,562, Đ = 1.306; **5-GlcA**: $M_{n,calc.}$ = 149 kDa, $M_{n,SEC}$ = 31,725, Đ = 2.099.

Remodeling of RBC glycocalyx with glycopolymers 5. RBCs (4% w/v in PBS) were incubated with Cy5-labeled glycopolymers **5** at increasing concentrations (c_{pol} = 0.1-10.0 µM) for 1 h at 37 °C. The cells washed 1x with PBS, then were probed for the presence of Cy5 fluorescence using flow cytometry. The data were analyzed on Cytobank online software. Cells were gated to exclude debris, and the median fluorescence intensities (MFI) of cells are reported. Means and standard deviations for each condition were calculated from three independent biological replicates.

Determination of Cell morphology by Electron Microscopy. RBCs treated with glycopolymers **5** at concentrations of 2.5 µM and 10.0 µM, as well as untreated cells, were prepared fixed in glutaraldehyde solution in PBS (2.5%) at 4 °C overnight. The cells were gradually transferred into EtOH by washes with DI water containing gradually increasing concentrations of EtOH

(0-100%, 10% increments). The samples were dried using Tousimis AutoSamdri 815A critical point dryer and sputter coated with Iridium for 8 seconds using Emitech K575X Iridium Sputter coater. SEM imaging was done with ETD detector at HV 4.00KV with 0.1 nA current.

Agglutination of glycocalyx-remodeled RBCs in the presence of ConA and SNA lectins. In a round-bottom 96 well plates containing RBCs (25µL, 1% in PBS) treated with glycopolymers **5** (2.5 µM) or alkynyl cholestanone **2** (2.5 µM), or untreated cells were added fluorescein-labeled ConA and SNA lectins (25 uL) at increasing concentrations (c_{ConA} = 0 - 1.8 µg/mL and c_{SNA} = 0 – 0.25 µg/mL). The cells were mixed gently but thoroughly using a pipette tip and, then, allowed to agglutinate for 45 min. After this time, the plates were scanned on an EPSON Perfection V700 Photo scanner (Digital ICE technologies), and the lowest lectin concentrations required to induce RBC agglutination were determined. The settling of RBCs to the bottom of a well to form a solid dot shape indicated a lack of agglutination. Each condition was evaluated in three independent biological replicates.

Determination of lectin association with glycocalyx-remodeled RBCs by flow cytometry. In a 96 well round bottom plate, to RBCs (0.33% in PBS) treated with glycopolymers **5** (2.5 µM) or alkynyl cholestanone **2** (2.5 µM), or to untreated cells, were added fluorescein-labeled ConA and SNA lectins at sub-agglutination concentrations (c_{ConA} = 8 µg/mL and c_{SNA} = 1 µg/mL). The cells were vortexed vigorously for ~ 10 s and then analyzed by flow cytometry (Canto II) for the presence of fluorescein signal at discrete time points until saturation lectin binding was observed. The data were analyzed on Cytobank software. Cells were gated to exclude debris, and median fluorescence intensities (MFI) of cells are reported. Means and standard

deviations were calculated from two independent biological experiments, and *p*-values corresponding to each condition vs. untreated RBC control were calculated using 2-way ANNOVA tests with PRISM software. The linear regions of the lectin binding curves were determined (t = 0 – 200 s) and fitted using a linear regression in PRISM software. The slopes designating the initial rates of lectin association and the R^2 values for the linear fits were extracted for each condition and their significance with respect to untreated RBC controls was assessed based on *p*-values calculated using 1-way ANNOVA tests.

Conflicts of interest

There are no conflicts to declare.

Section 2.5 Acknowledgements

We thank the UCSD Microscopy Core facility (via p30 grant NS047101 from NINDS) for assistance with fluorescence microscopy imaging, and the Glycobiology Research and Training Center for access to tissue culture facilities and analytical instrumentation. SEM and AFM work was performed at the San Diego Nanotechnology Infrastructure (SDNI) of UCSD, a member of the National Nanotechnology Coordinated Infra-structure (NNCI), which is supported by the National Science Foundation (Grant ECCS-1542148). We also wish to thank Dr. Meghan Altman for help for her expertise in and valuable comments on RBC assay development. This work was supported by the NIH Director's New Innovator Award (NICHD: 1DP2HD087954-01). K. G. is supported by the Alfred P. Sloan Foundation (FG-2017-9094) and the Research Corporation for Science Advancement via the Cottrell Scholar Award (grant # 24119).

Chapter 2, in full, is a reprint of the material as it appears: Honigfort, D; Zhang M. H; Verespy, S; and Godula, K. Engineering of spectator glycocalyx structures to evaluate

molecular interactions at crowded cellular boundaries. *Faraday Discussion*, 2019, DOI 10.1039/C9FD00024K. The dissertation author is the primary co-author of this manuscript.

Section 2.6 Supporting Information

Engineering of spectator glycocalyx structures to evaluate molecular interactions at crowded cellular boundaries.

Daniel J. Honigfort,[a] Michelle H. Zhang,[a] Stephen Verespy III,[a,] and Kamil Godula*[,a]

[a]*Department of Chemistry and Biochemistry, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0358, USA*

**corresponding author (kgodula@ucsd.edu)*

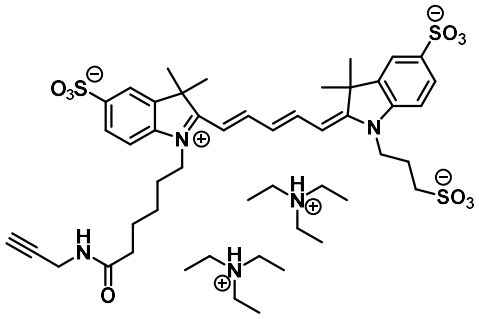
Supporting Information

Instrumentation and general procedures.

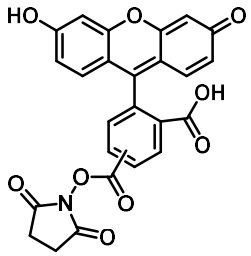
Column chromatography was performed on a Biotage Isolera One automated flash chromatography system. Nuclear magnetic resonance (NMR) spectra were collected on a Bruker 300 MHz and a Jeol 500 MHz NMR spectrometers. Spectra were recorded in $CDCl_3$ or D_2O solutions at 293K and are reported in parts per million (ppm) on the δ scale relative to the residual solvent as an internal standard (for ^1H NMR: $CDCl_3$ = 7.26 ppm, D_2O = 4.79 ppm, for ^{13}C NMR: $CDCl_3$ = 77.0 ppm). HRMS (high-resolution mass spectrometry) analysis was performed on an Agilent 6230 ESI-TOFMS in positive ion mode. UV-Vis spectra were collected with a quartz cuvette using a Thermo Scientific Nanodrop2000c spectrophotometer. IR spectroscopy was performed on a Nicolet 6700 FT-IR spectrophotometer (Thermo Scientific). Size exclusion chromatography (SEC) was performed on a Hitachi Chromaster system equipped with an RI detector and two 5 μm, mixed bed, 7.8 mm I.D. x 30 cm TSKgel columns in series (Tosoh Bioscience). Organic soluble polymers were analyzed using an isocratic method with a flow rate of 0.7 mL/min in DMF (0.2% LiBr, 70 °C). For aqueous SEC, two 8 μm, mixed-M bed, 7.5 mm I.D. x 30 cm PL aquagel-OH columns in series (Agilent Technologies) were run in sequence using an isocratic method with a flow rate of 1.0mL/min in water (0.2M $NaNO_3$ in 0.01M Na_2HPO_4, pH = 7.0). AFM imaging of lactose glycopolymer **5** was performed on Nanoscope IV Scanning Probe Microscope (Veeco) in tapping mode using an 9991EVLR scanner and Tap150AL-G silicon AFM probes with aluminum reflex coating (Ted Pella). For imaging, polymers in aqueous solution were deposited on PELCO mica discs (Ted Pella). After 1 min, excess water was removed, and discs were allowed to dry at ambient pressure before imaging. Flow cytometry analysis was performed on live RBCs using a FACS Canto II cytometer (BD Biosciences). Red blood cells samples for bright field microscopy were prepared by diluting live cells in PBS 1:1 with Vectamount AQ aqueous mounting medium

(Vector Labs H-5501) and imaged Keyence Fluorescent microscope. Scanning Electron Microscopy (SEM) was performed on an FEI Apreo Scanning Electron Microscope. RBCs containing 0nM, 2500nM, and 10000nM were prepared by fixing overnight in 2.5% glutaraldehyde solution at 4 °C, then were gradually dehydrated by washing with 0, 10, 20, 30, 50, 70, 80, 90, 100, 100% EtOH solutions. Prepared samples were then dried using Tousimis AutoSamdri 815A critical point dryer, and sputter coated with Iridium for 8 seconds using Emitech K575X Iridium Sputter coater. SEM imaging was done with ETD detector at HV 4.00KV, 0.1nA current.

Structure of Cy5-alkyne:

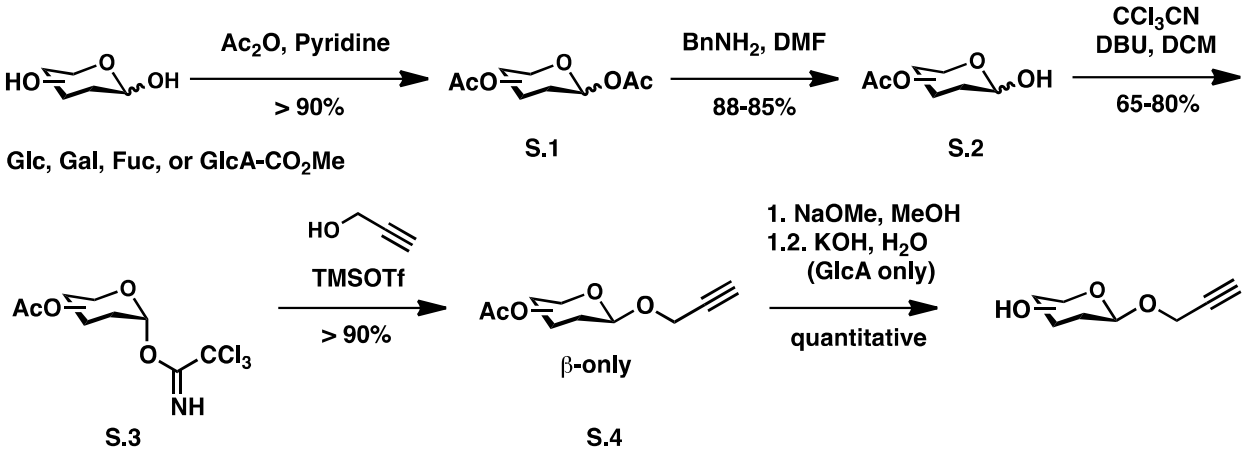


Structure of 5/6 Fluorescein-NHS:



Synthesis of propargyl glycosides.

ß-propargyl glucoside (Glc),[35] galactoside (Gal),[35] fucoside (Fuc)[32] and glucuronopyranoside (GlcA)[33] were prepared according to published procedures. General procedures for the preparation of the glycosides according to Scheme S1 and ^{1}H NMR spectra for all intermediates and final glycosides are listed below.



Scheme 2.1. Preparation of propargyl glycosides via the Schmidt glycosylation.

General procedure for the preparation of peracetylated glycans (S.1). To glucose (Glc), galactose (Gal), fucose (Fuc), or glucopyranuronic acid methyl ester (GlcA-CO_2Me) (1.0 mmol) was added acetic anhydride (1.5 ml) and pyridine (1.5 ml). The reaction was cooled in an ice bath, and catalytic DMAP (50.0 mg, 0.1 mmol) was added. The mixture was brought to room temperature and stirred for 12 hrs. After this time, the reaction mixture was diluted with dichloromethane, washed with a saturated solutions of $NaHCO_3$ and NH_4Cl, dried with anhydrous Na_2SO_4, and concentrated under reduced pressure to give peracetylated glycosides **S.1** in >90% yield. The crude products were used without further purification.

General procedure for anomeric hydroxyl deprotection in peracetylated glycans (S.2). To a solution of peracetylated glycans **S.1** (1.0 mmol) in dry DMF (4.0 mL) was added benzylamine (1.2 mmol). The mixture was stirred overnight. The reaction mixture was diluted with ethyl acetate and washed with a saturated solution of NH_4Cl, followed by brine. The organic phase was dried with anhydrous Na_2SO_4 and concentrated. The products were purified via column chromatography using 1:2 Hexanes/EtOAc to yield anomeric-OH deprotected glycans **S.2** as clear oils: S.2-Glc (95%) S.2-Gal (88%) S.2-Fuc (95%) S.2-GlcA (94%).

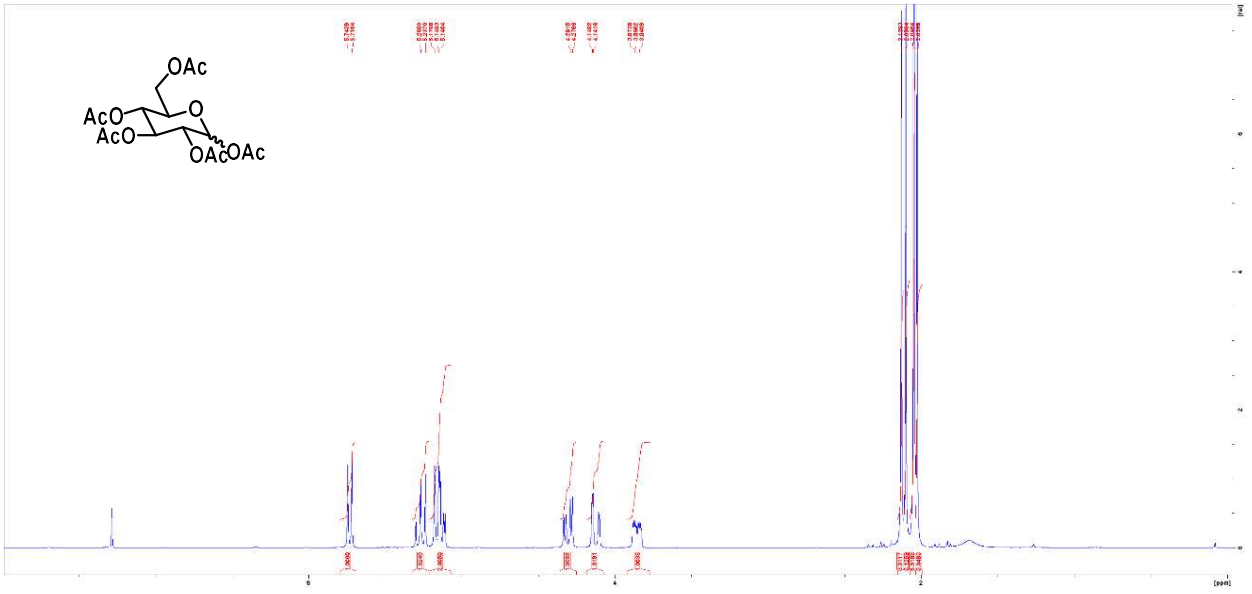
General procedure for synthesis of tricholoroacetimidate (TCA) glycosides (S.3). To a chilled (ice bath) solution of glycans **S.2** (1.0 mmol) in dry DCM (5.0 mL) were added trichloroacetonitrile (5.0 mmol) and DBU (0.1 mmol) under N2 atmosphere. After one hour, the reaction was concentrated under reduced pressure and the crude products were purified via column chromatography using 3:2 Hexanes/EtOAc to yield TCA glycosides **S.3** as light yellow oils: S.3-Glc (75%) S.3-Gal (80%) S.3-Fuc (65%) S.3-GlcA (70%).

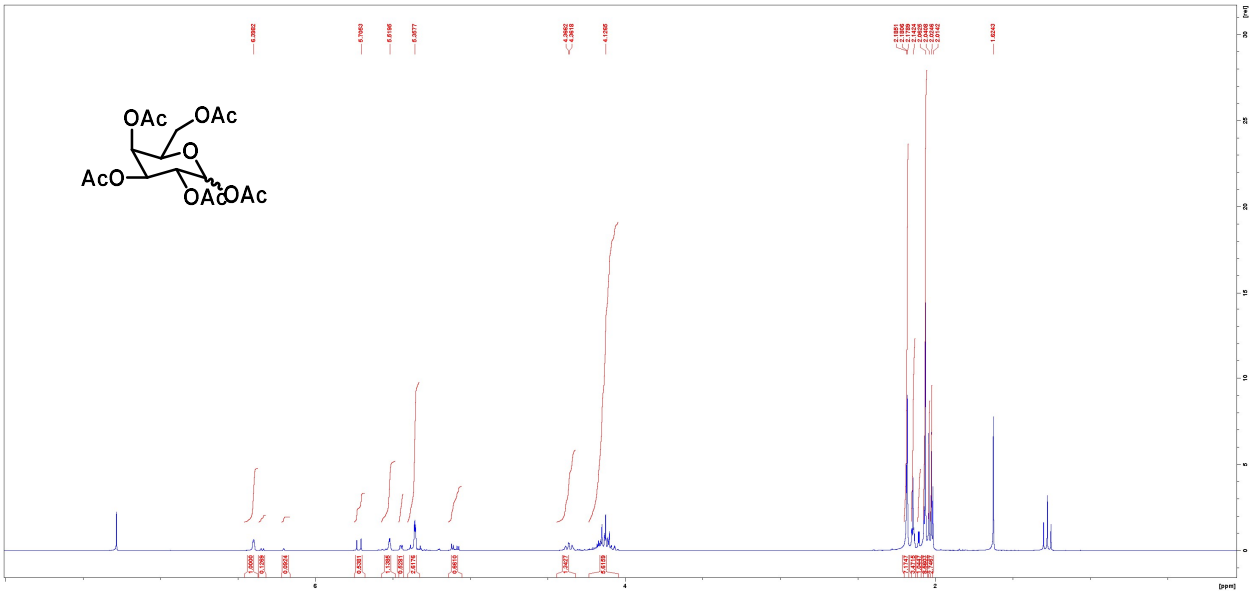
General procedure for the preparation of protected β -propargyl glycosides (S.4). To a solution of TCA glycan donors **S.3** (1.0 mmol) in dry DCM (5.0 mL) was added propargyl alcohol (5.0 mmol). The resulting solutions were cooled on ice, followed by the addition of TMSOTf catalyst (0.2 mmol). After 1h, the reaction was quenched with a saturated solution of $NaHCO_3$. The organic layer was separated, dried with anhydrous Na_2SO_4, and concentrated under reduced pressure. The crude products were purified by column chromatography using 3:2 Hexanes/EtOAc on silica gel affording protected β -propargyl glycosides **S.4** as clear oils in >90% yield after purification.

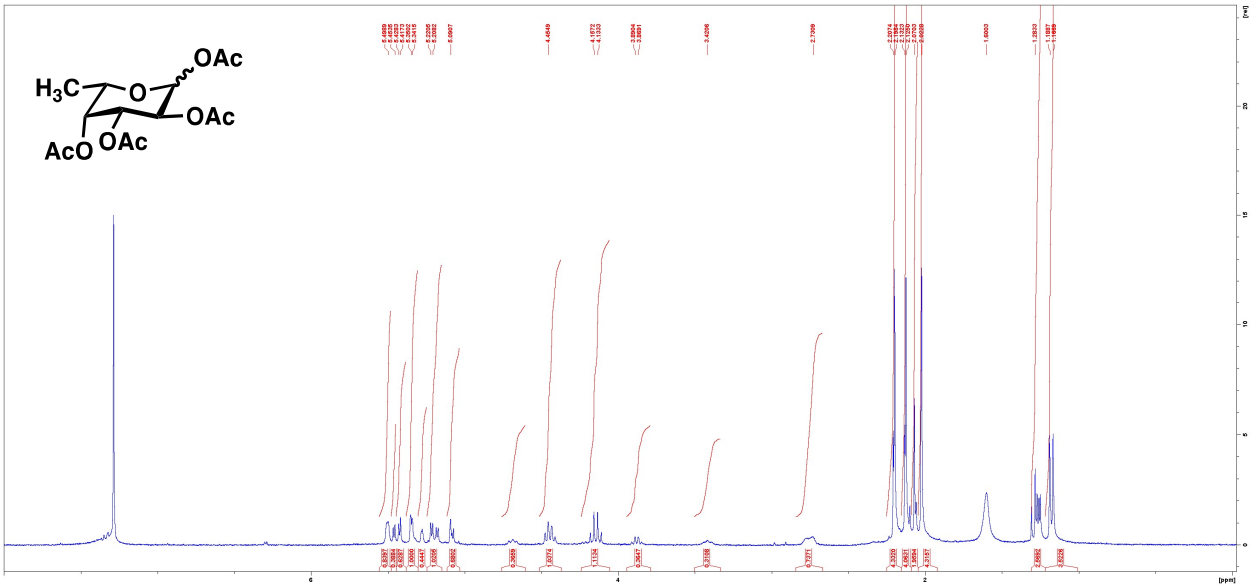
General procedure for the preparation of deprotected β-propargyl glycosides. To a solution of glycosides **S.4** (1.0 mmol) in dry methanol (5.0 mL) was added NaOMe (0.1 mmol) in methanol (0.1 mL). The reaction was stirred at room temperature for 2hrs, then was quenched by the

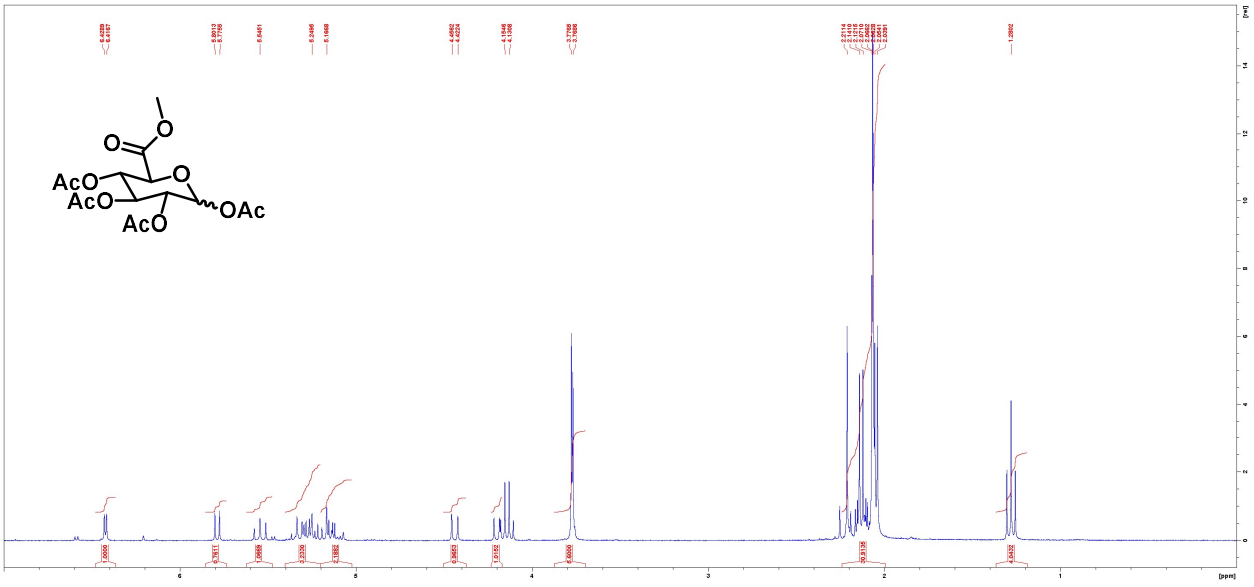
addition of Dowex 50WX8-200 beads. The resin was removed by filtration and the filtrates were concentrated to yield pure deacetylated β-propargyl glycosides. The β-propargyl glucoronopyranoside methyl ester (1.0 mmol) was dissolved in MeOH and an aqueous solution of KOH (1.5 mmol, 10.0 mL) was added. After 1 hour, the reaction was quenched with the addition of Dowex 50WX8-200 ion-exchange resin. The resin was removed by filtration and the filtrate was concentrated to yield a pure fully deprotected β-propargyl glucoronopyranoside.

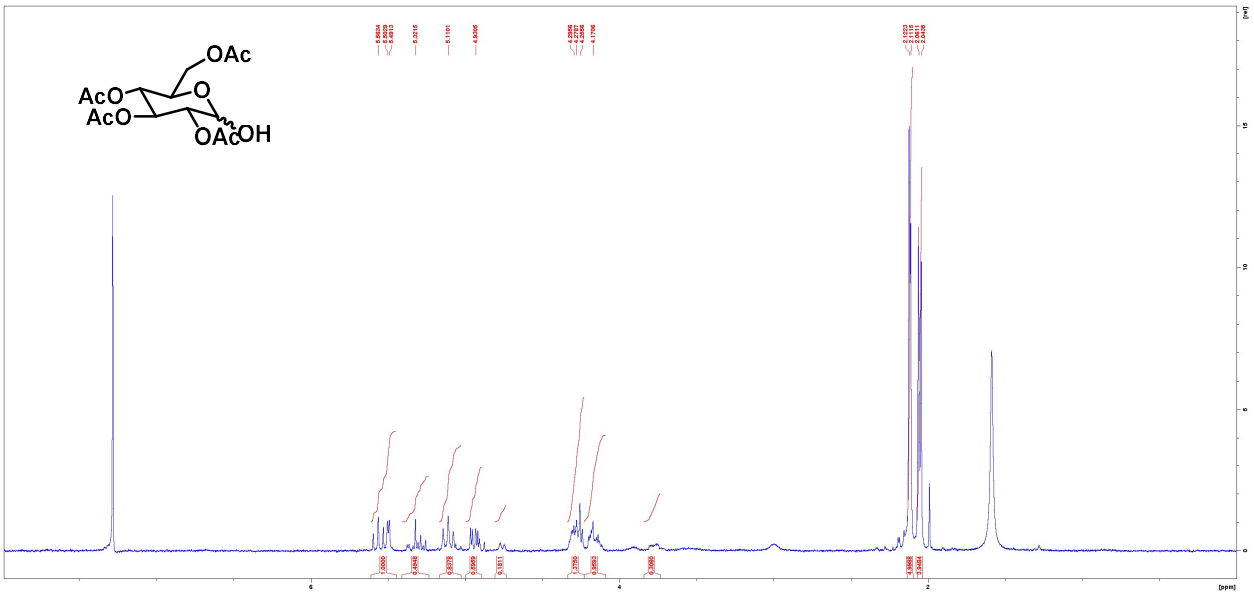
^{1}H NMR spectra of β -propargyl glycosides and synthetic intermediates.

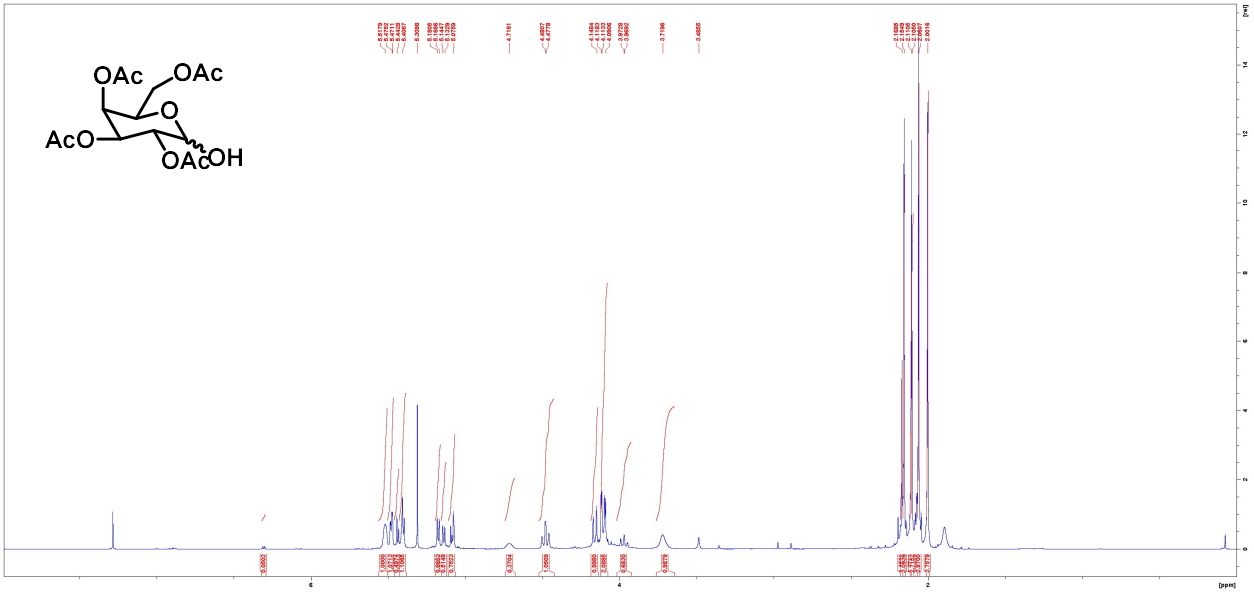


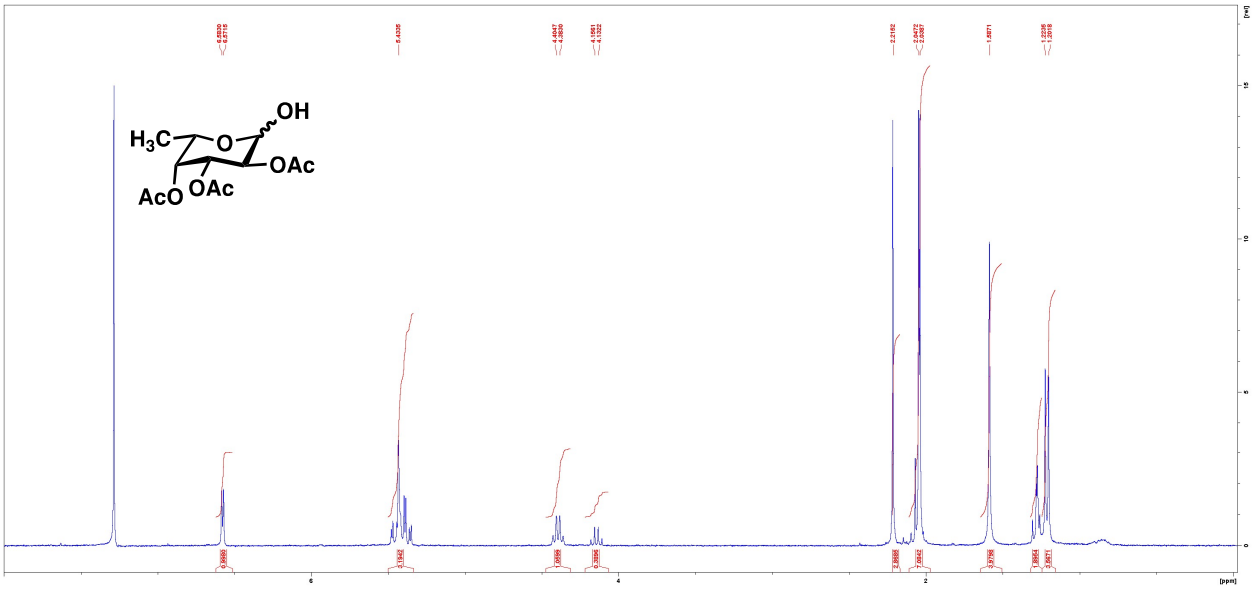


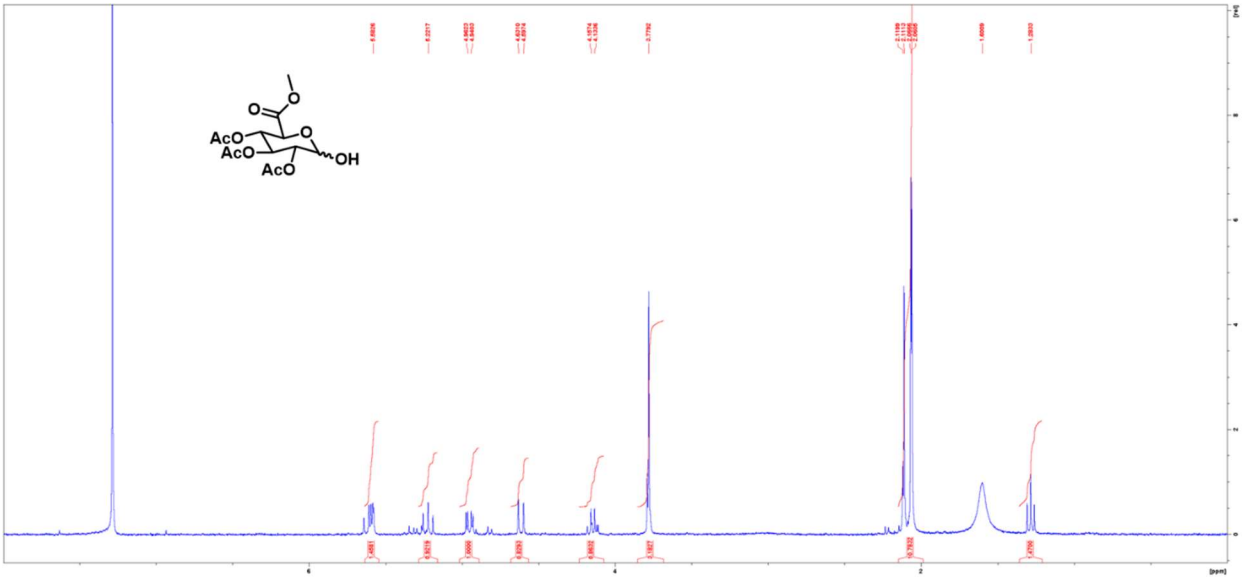

H3C
O
OAc
OAc
AcO
OAc

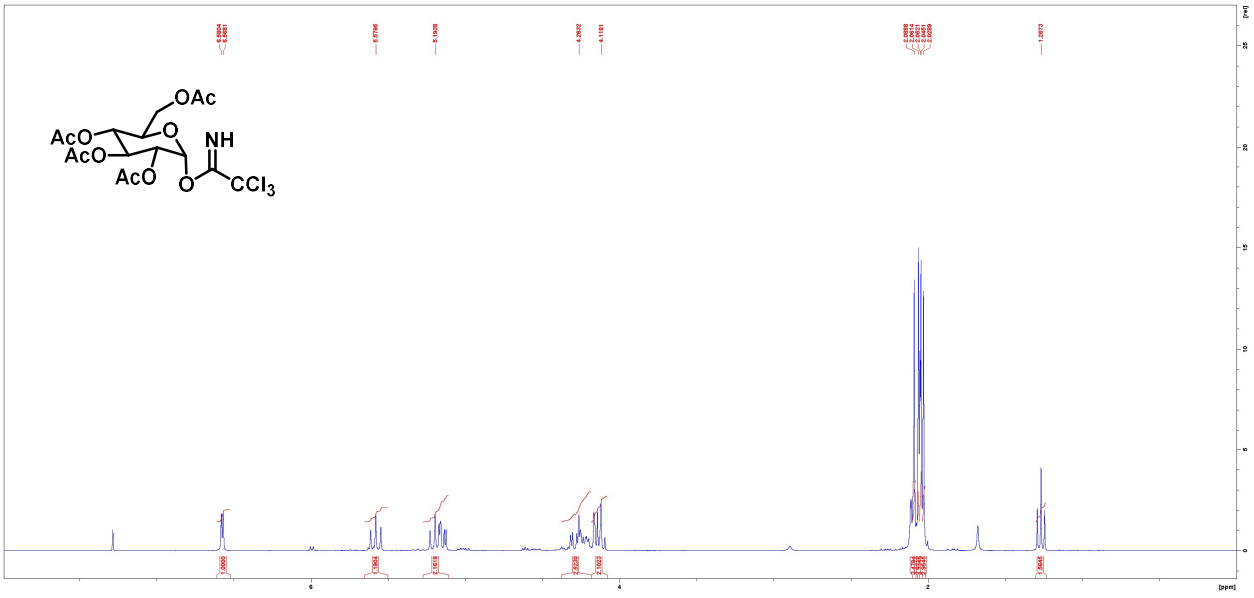

O
O
AcO
AcO
AcO
OAc

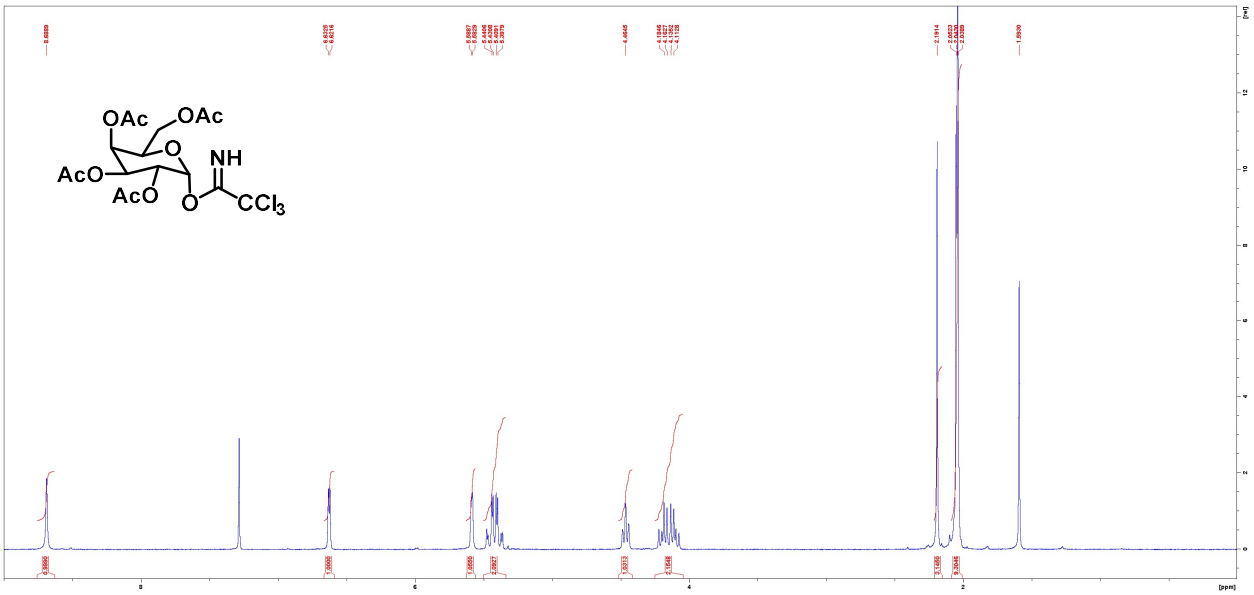

OAc
AcO
O
AcO
OAc
OH

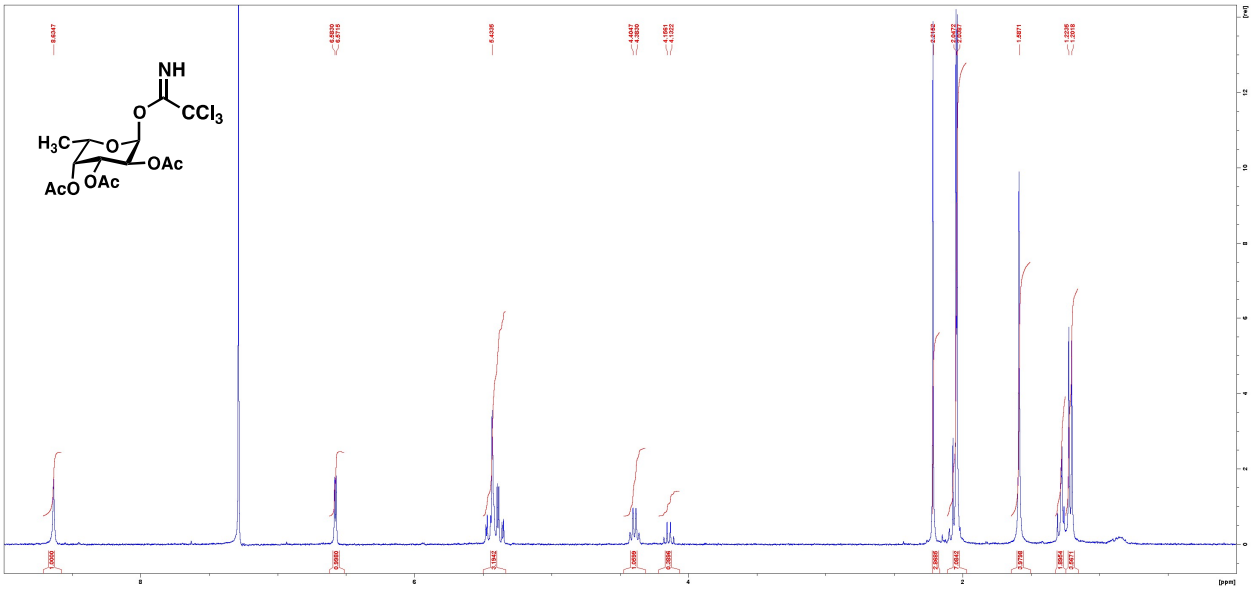

OAc
OAc
O
AcO
OAc
OH

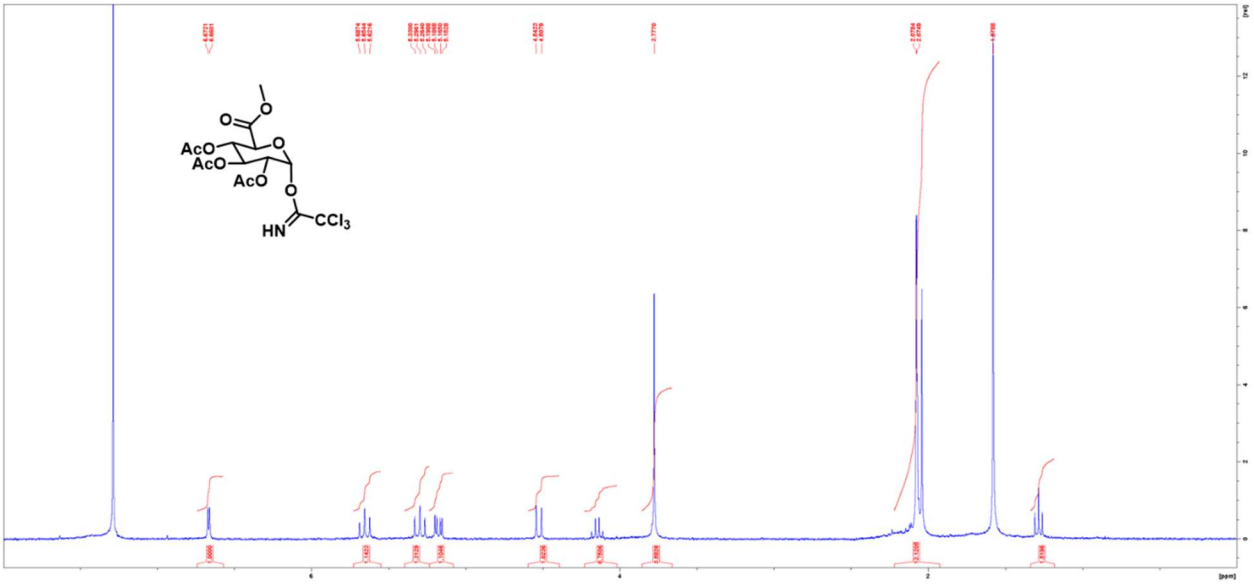

OH
H3C
O
OAc
AcO
OAc

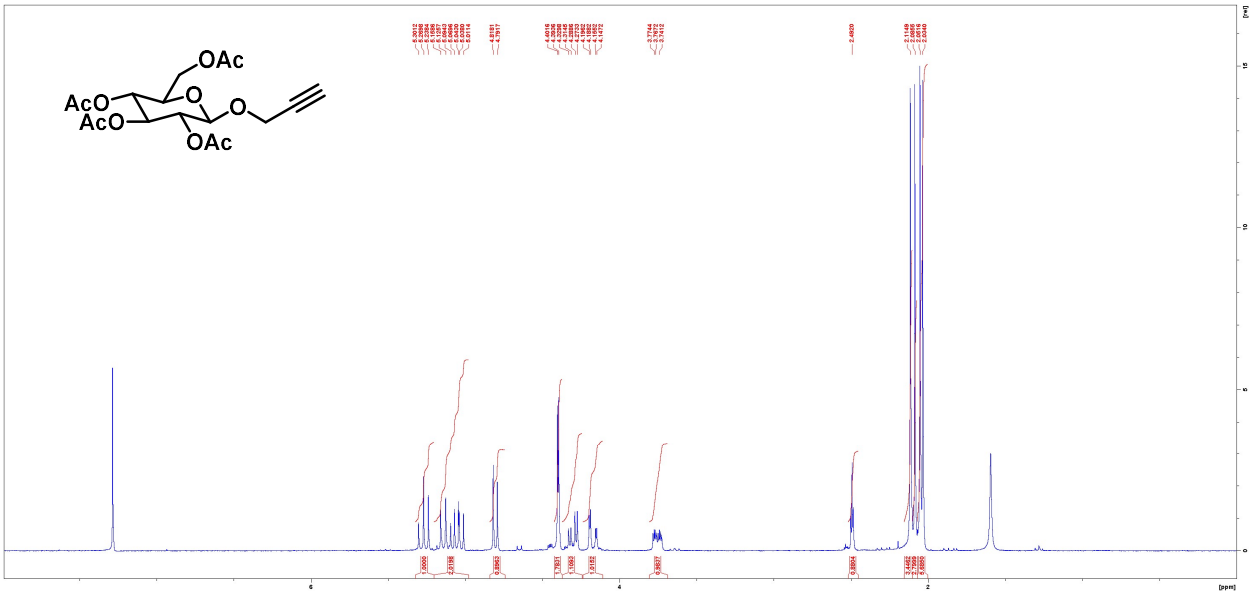

O
O
AcO
AcO
AcO
OH

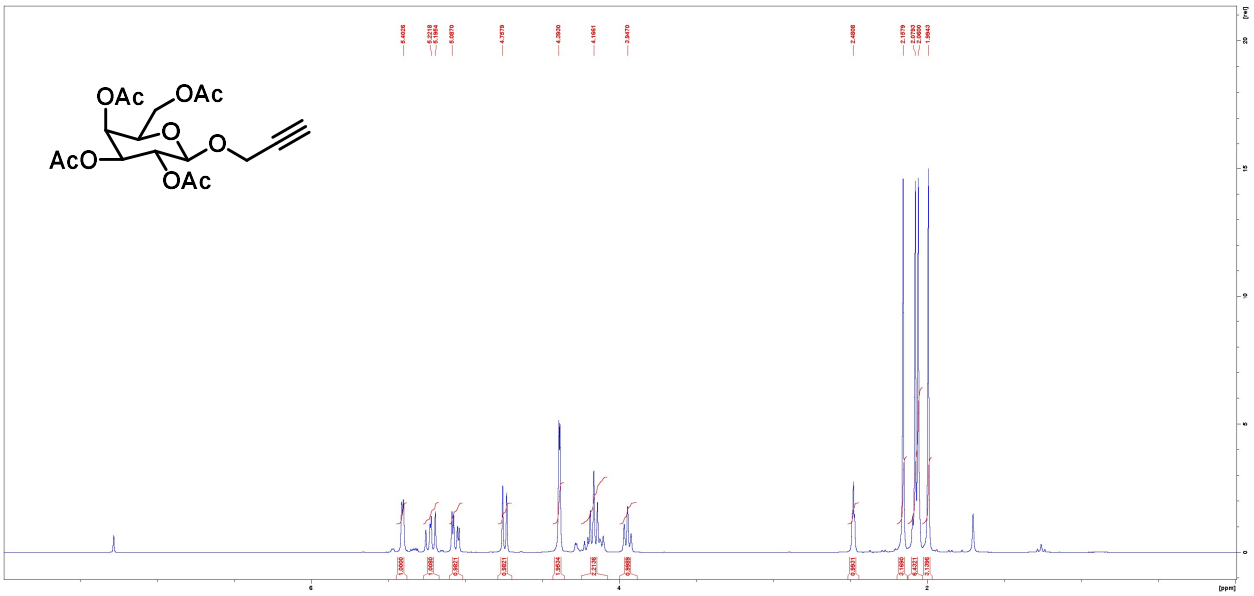

OAc
AcO
AcO
AcO
O
NH
O
CCl3

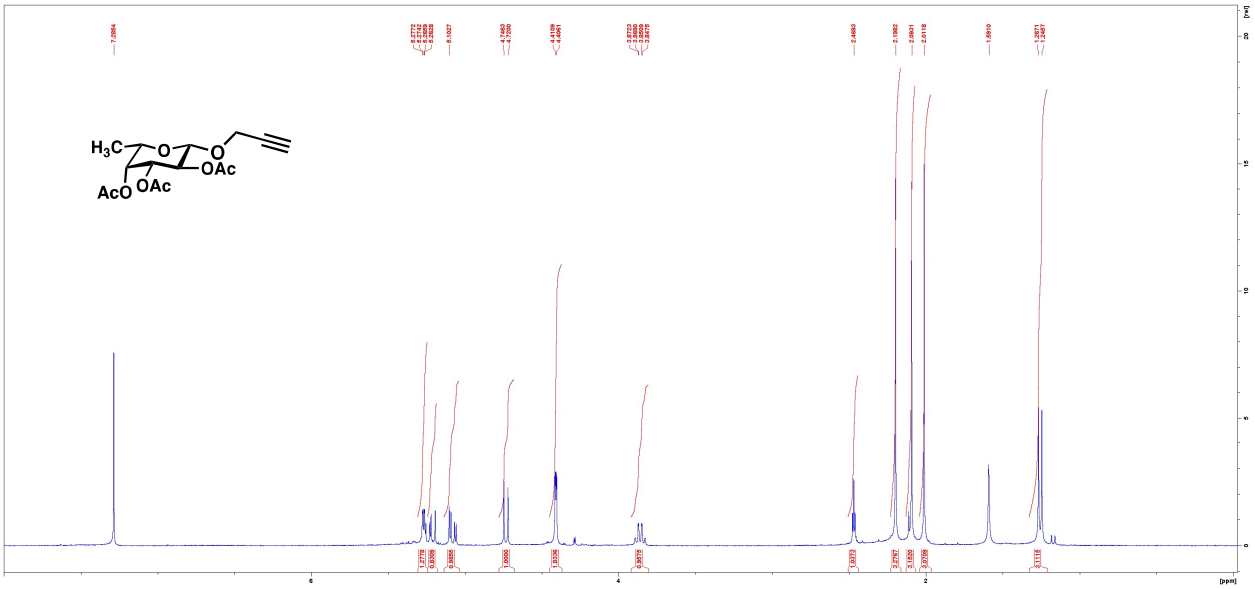

OAc
OAc
AcO
AcO
O
NH
O
CCl3

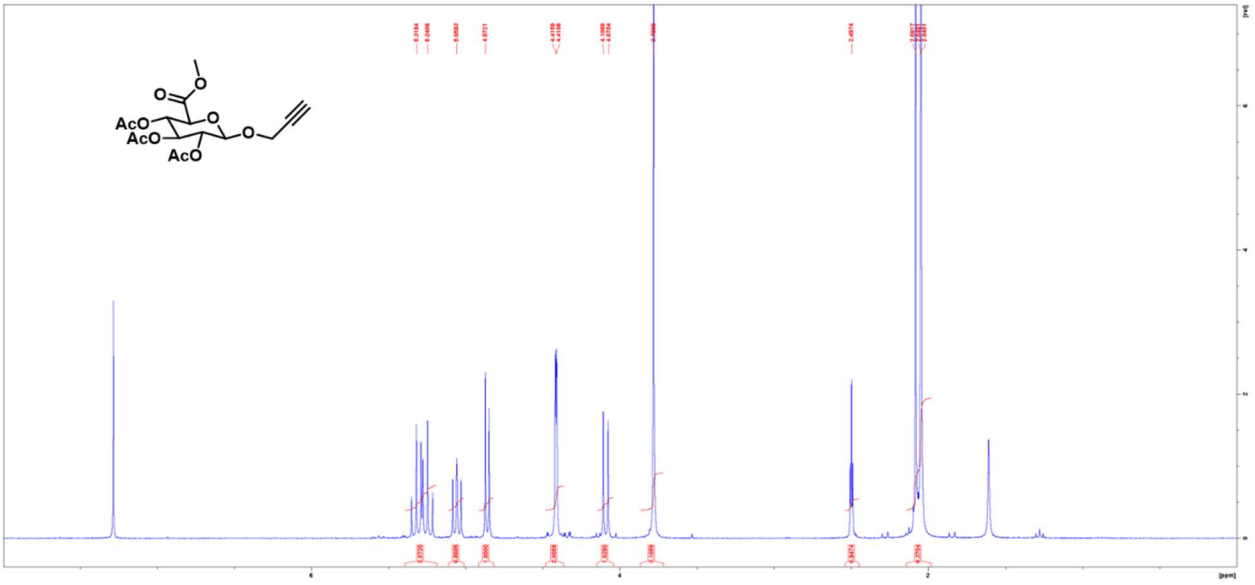

NH
O
CCl3
H3C
O
OAc
AcO
OAc

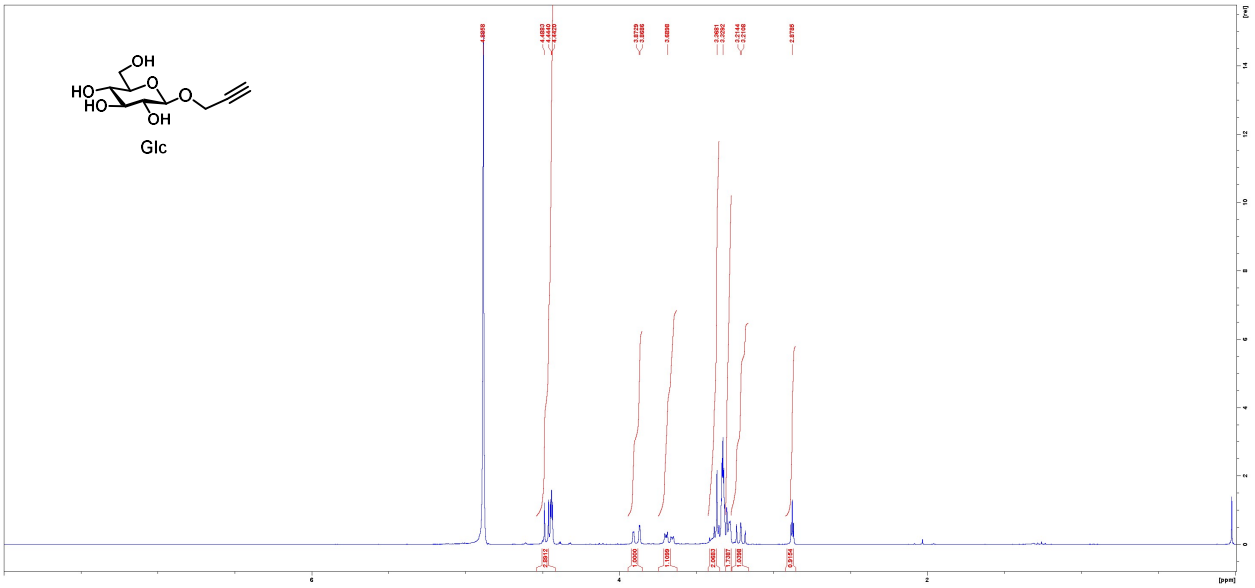

O
O
AcO
AcO
AcO
O
HN
CCl3

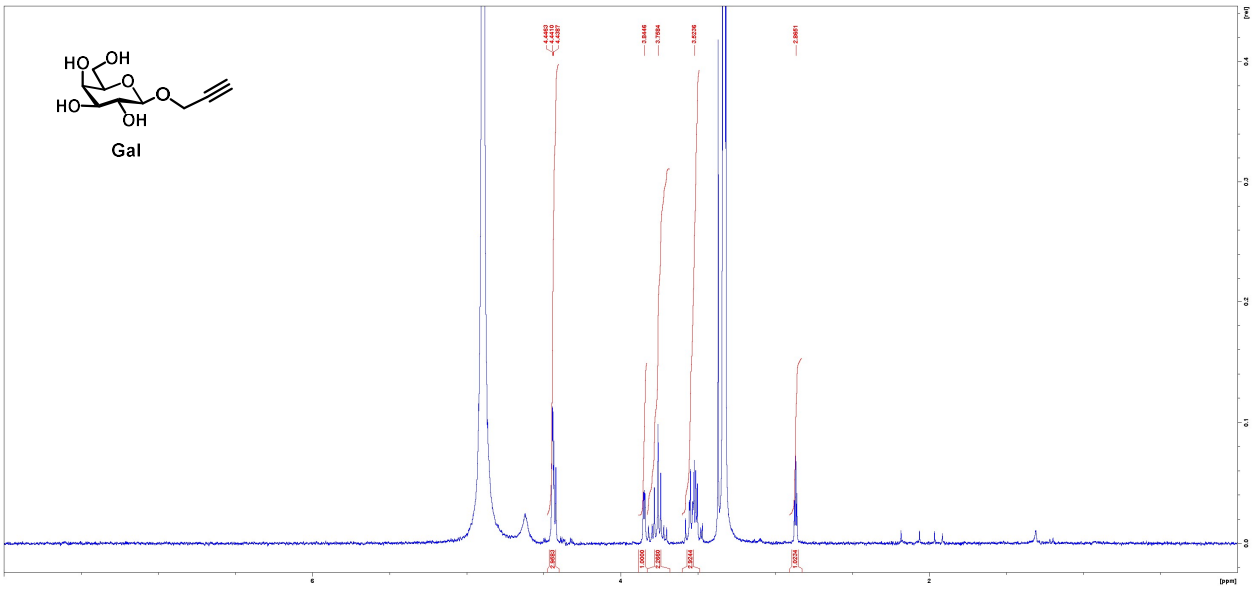

OAc
AcO
AcO
O
O
OAc

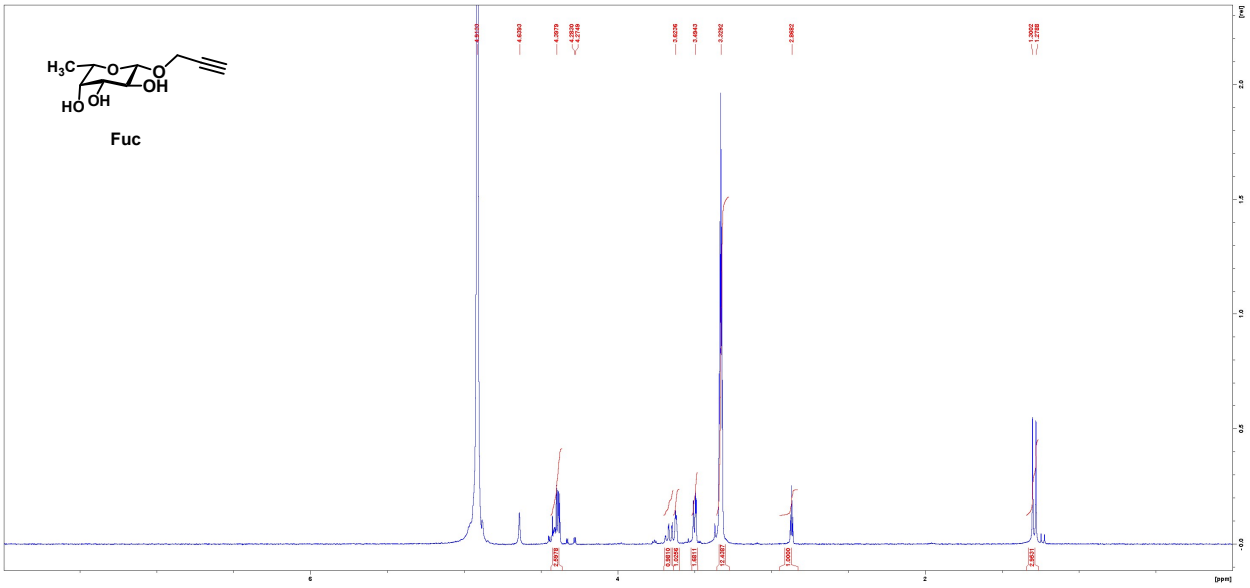

OAc
OAc
AcO
O
O
OAc

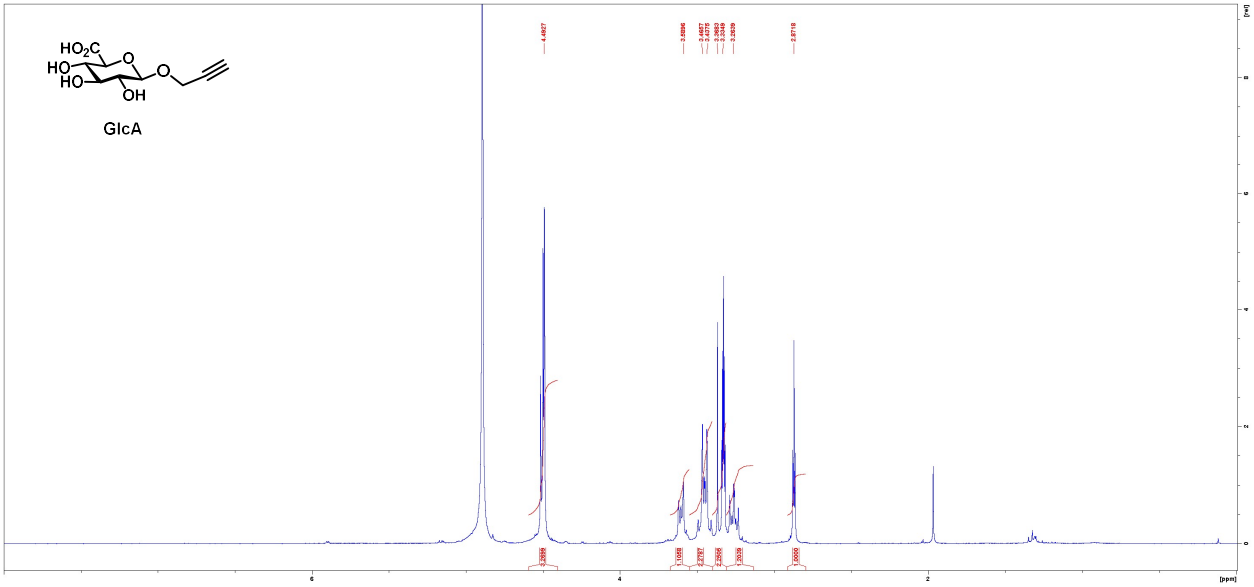

H3C
O
O
OAc
AcO
OAc


O
O
AcO
AcO
O
O
AcO

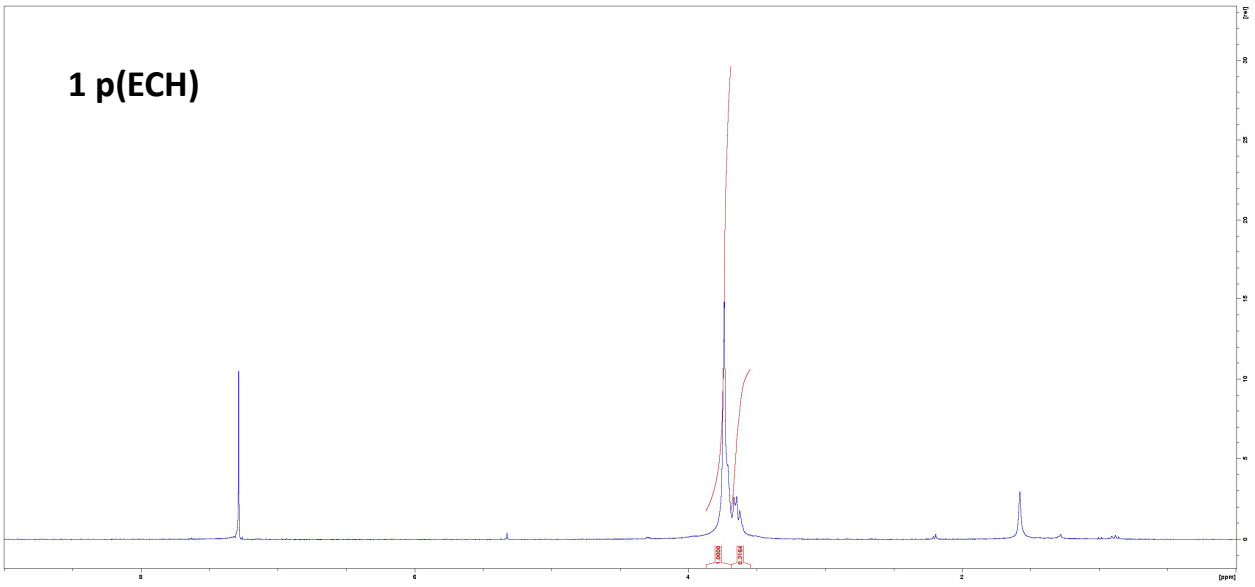

OH
HO
HO
O
O
OH
Glc

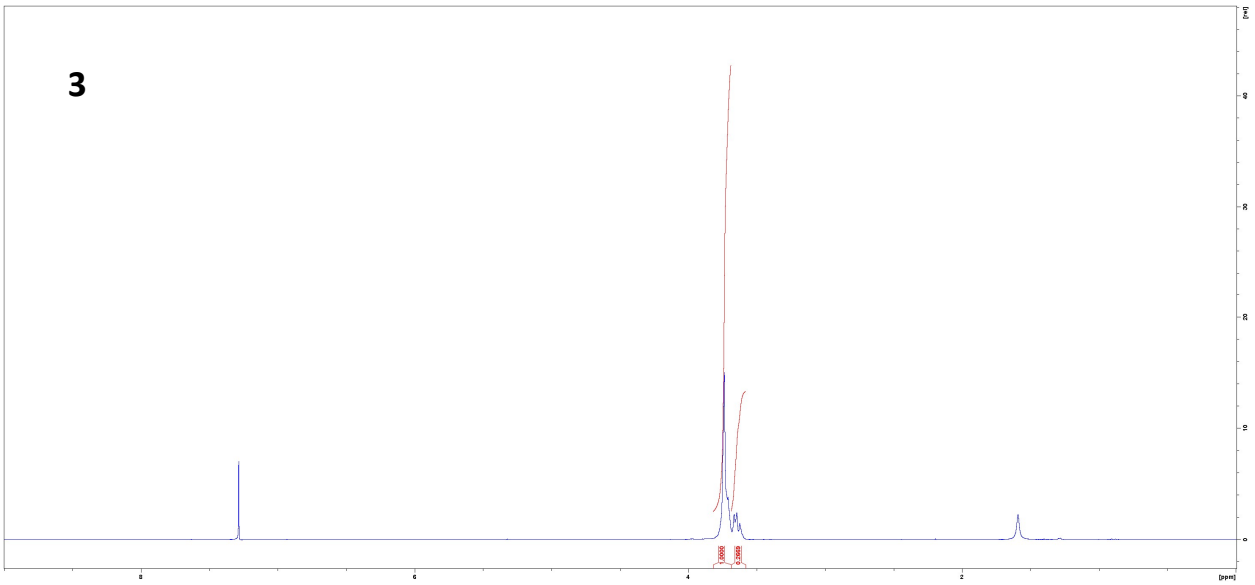

HO
OH
O
HO
O
OH
Gal

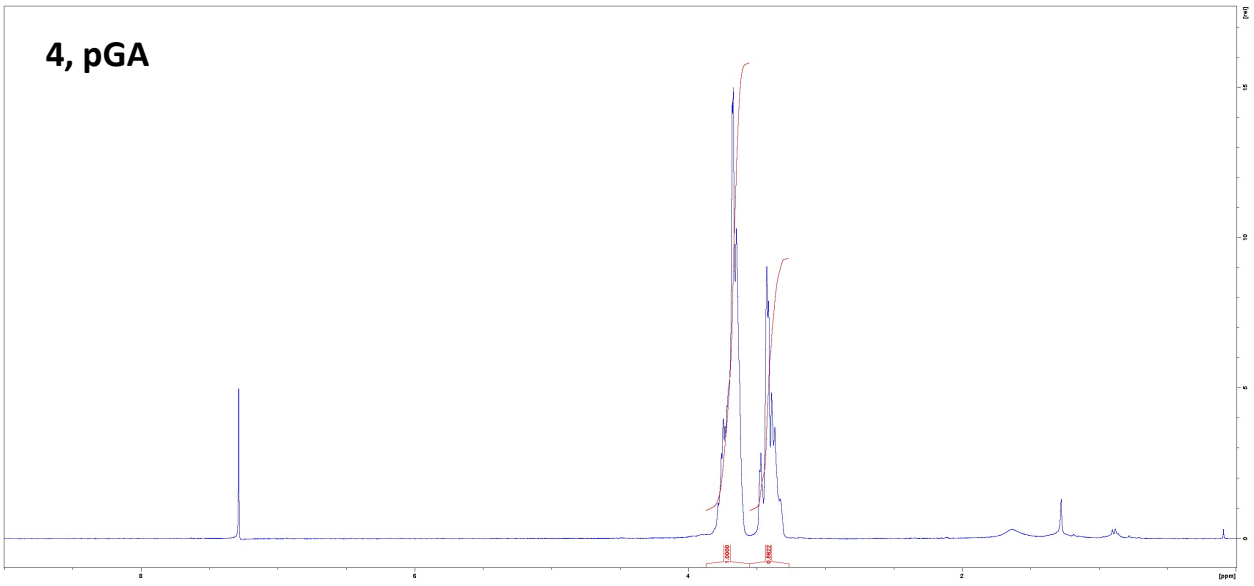

Fuc

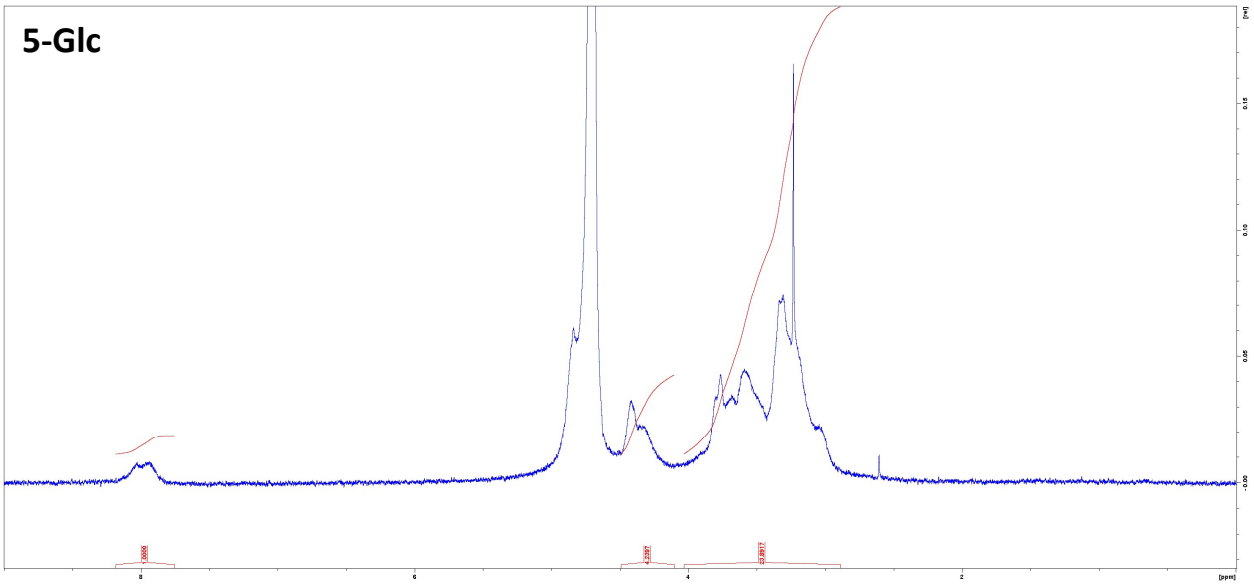

GlcA

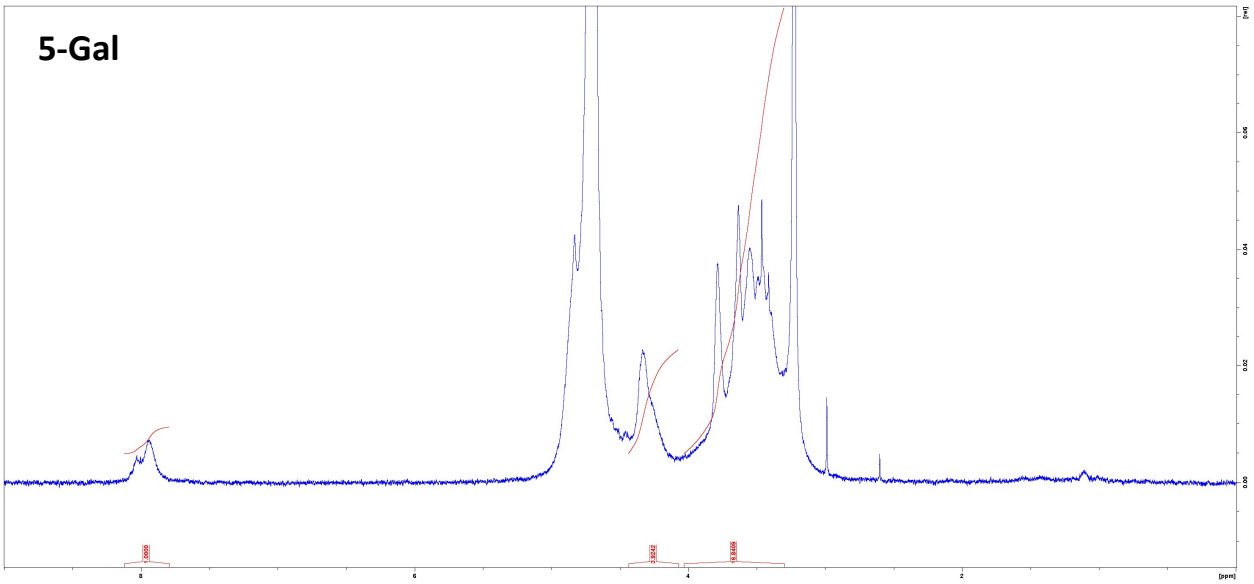
^{1}H NMR spectra of polymers 1, 3, 4, and 5.

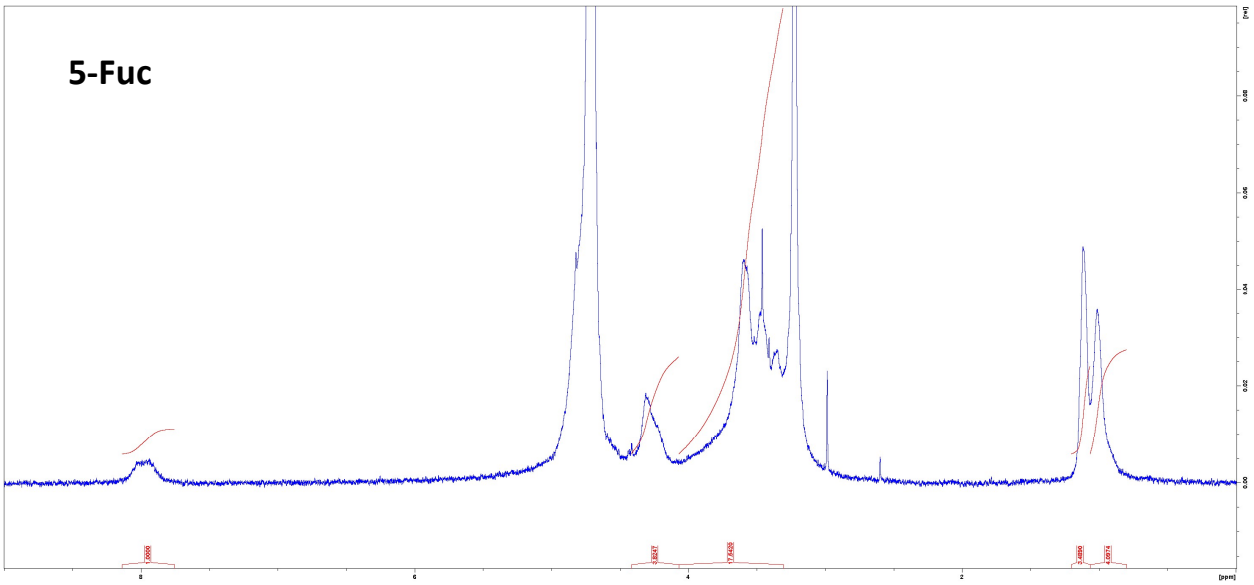


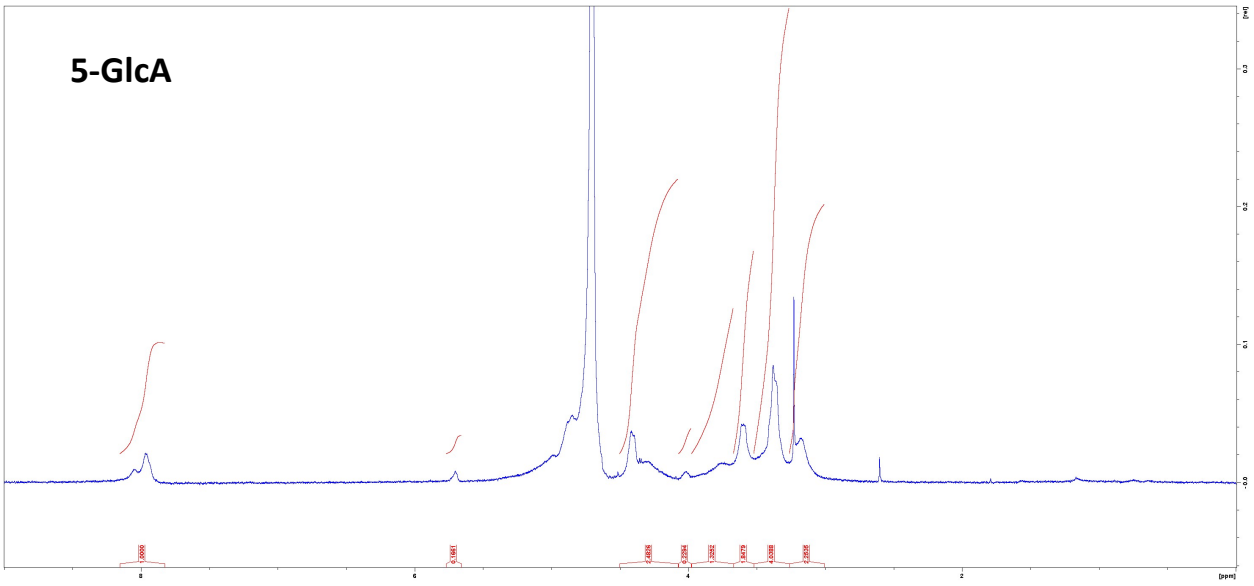


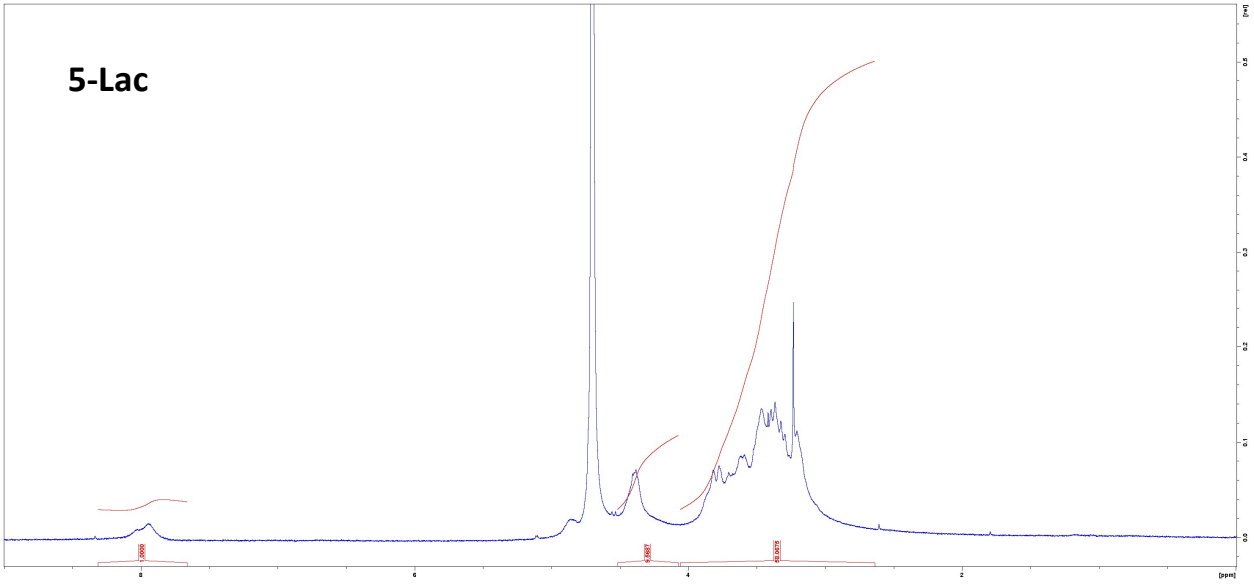

4, pGA


5-Glc


5-Gal


5-Fuc


5-GlcA


5-Lac

IR spectra of polymers 1, 3, 4, and 5.

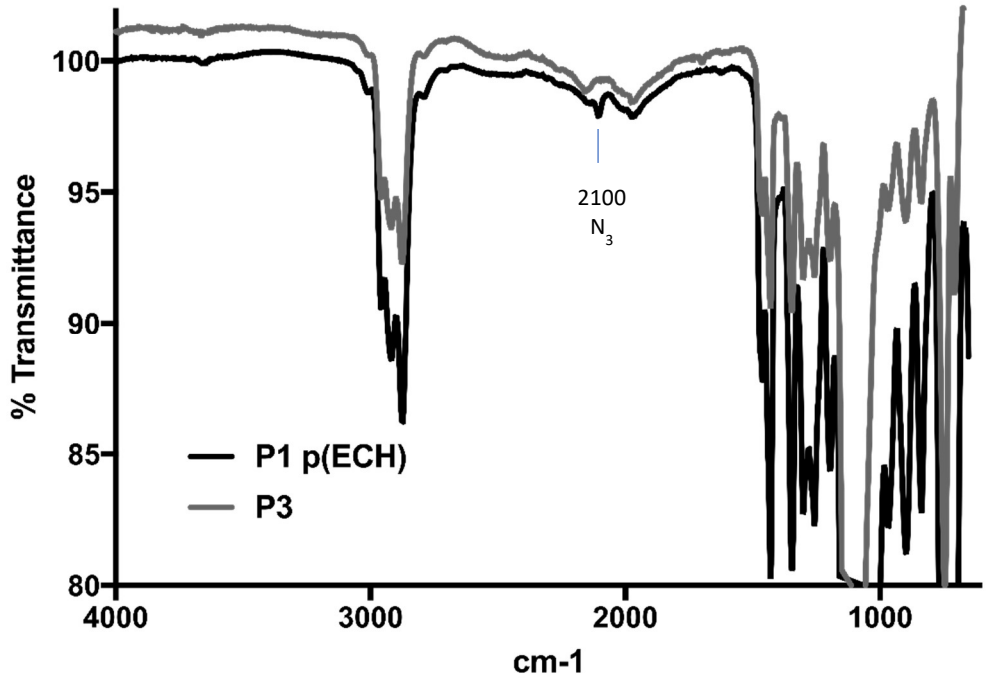


Figure 2.7. IR spectra of pECH polymers 1 and 3.

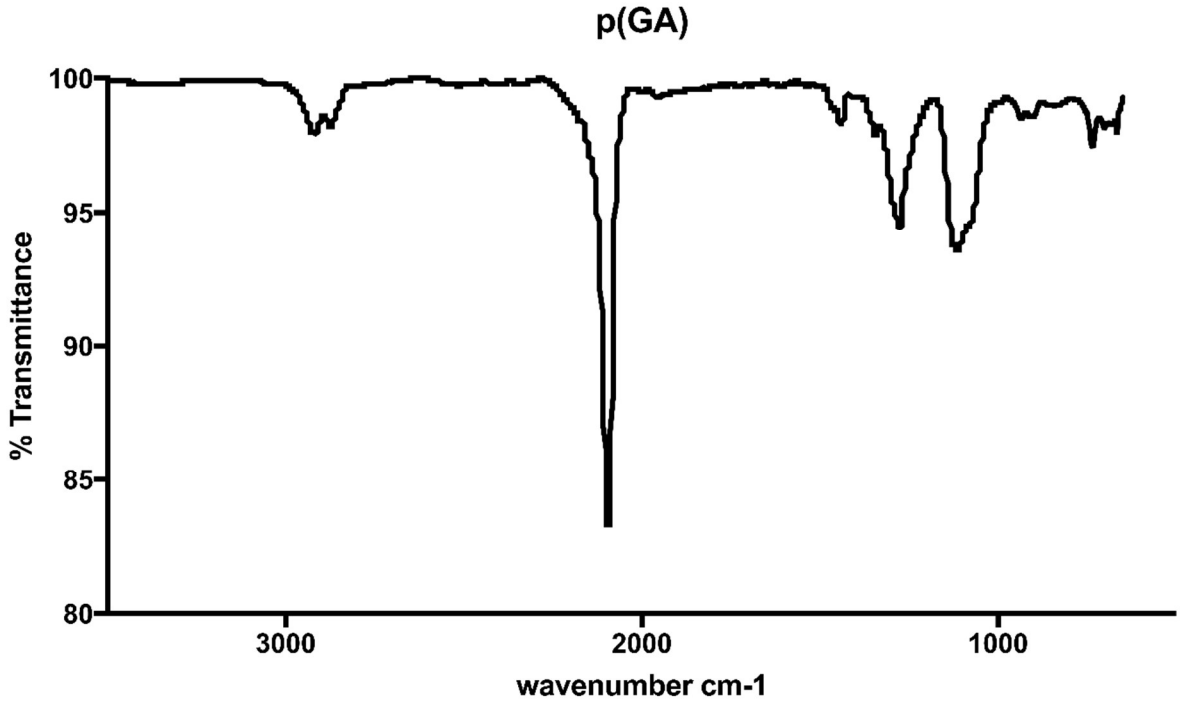
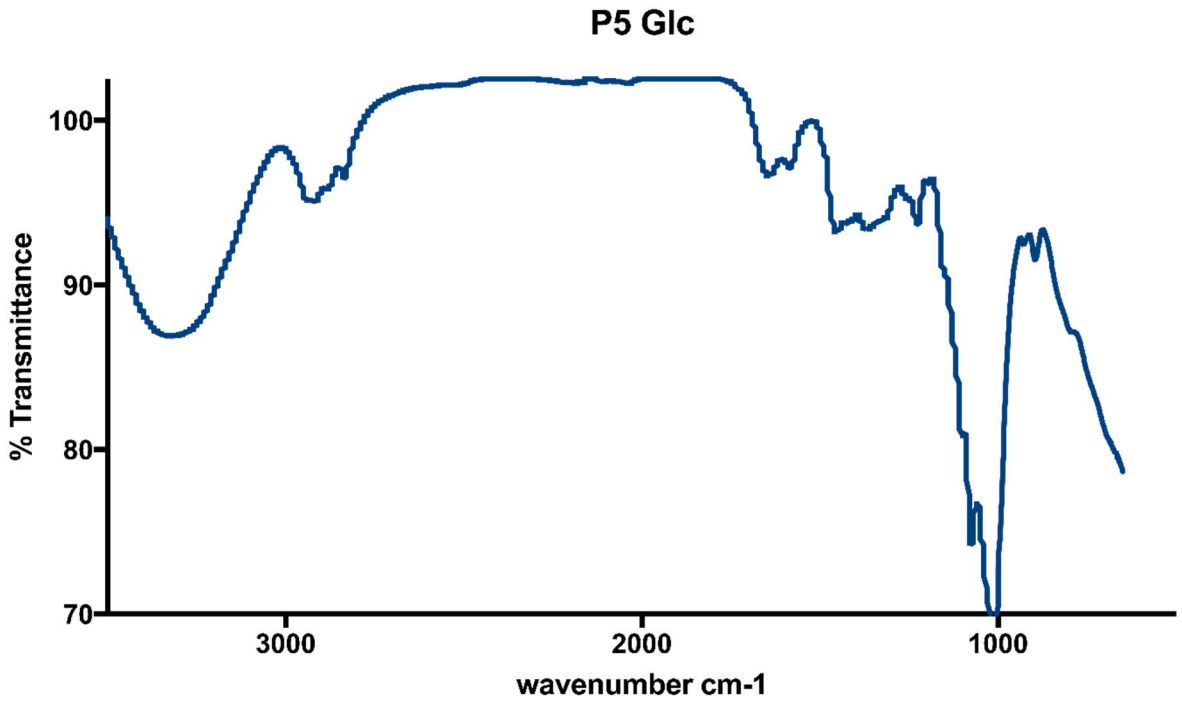
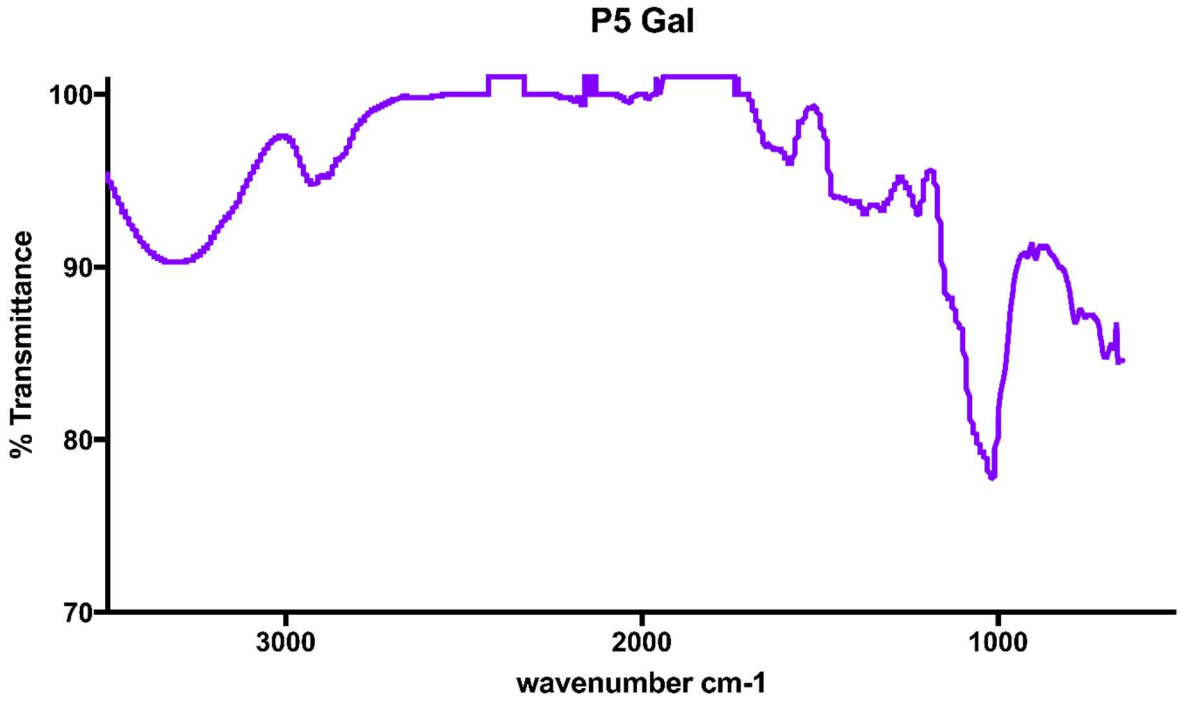


Figure 2.8. IR spectrum of pGA polymer 4.


P5 Glc
100
90
80
70
% Transmittance
3000
2000
1000
wavenumber cm-1


P5 Gal
100
90
80
70
% Transmittance
3000
2000
1000
wavenumber cm-1

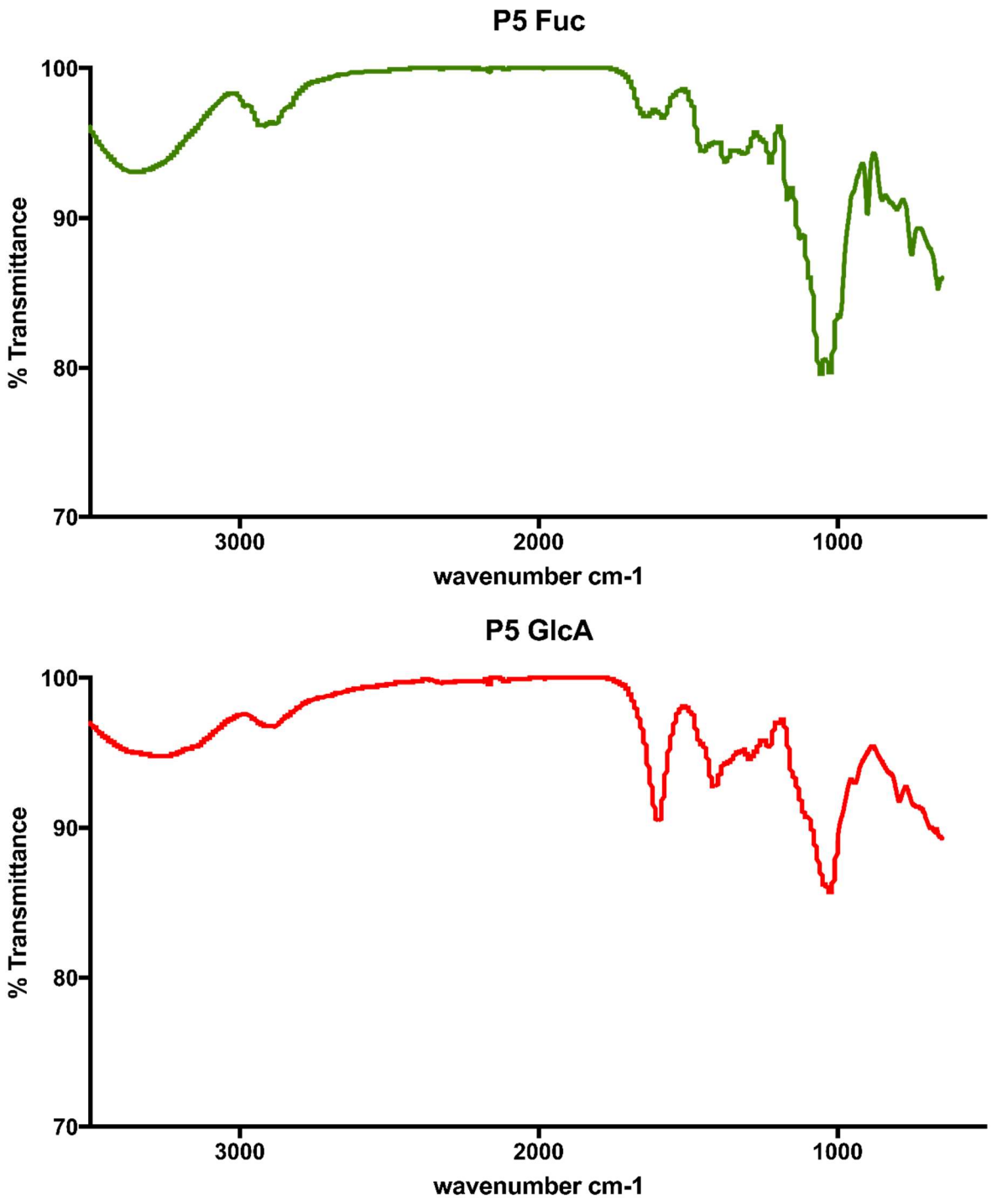


Figure 2.9. IR spectra of glycopolymers **5**

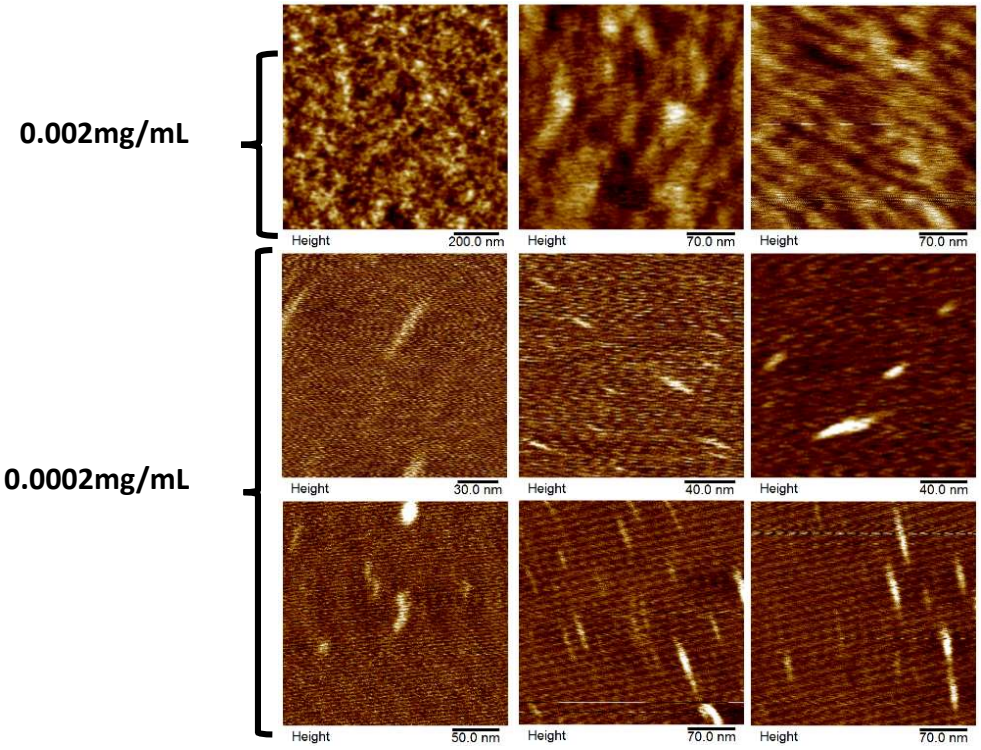
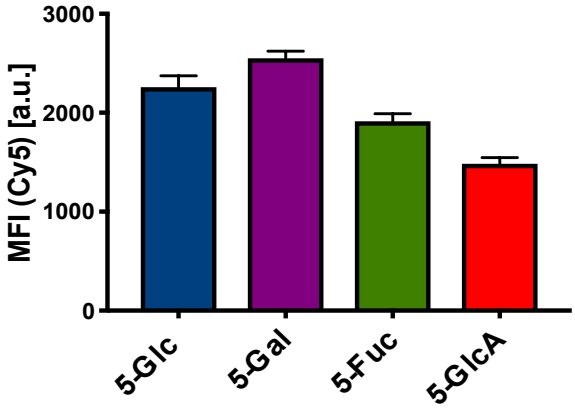


Figure 2.10. AFM images of **5-Lac** (expanded data for Figure 2.2D).



polymer (2.5 ∝M)	
polymer	relative surface incorporation (%)
5-Glc	100
5-Gal	113
5-Fuc	85
5-GlcA	65

Figure 2.11. Relative levels of cell surface incorporation of glycopolymers **5** (2.5 μM).

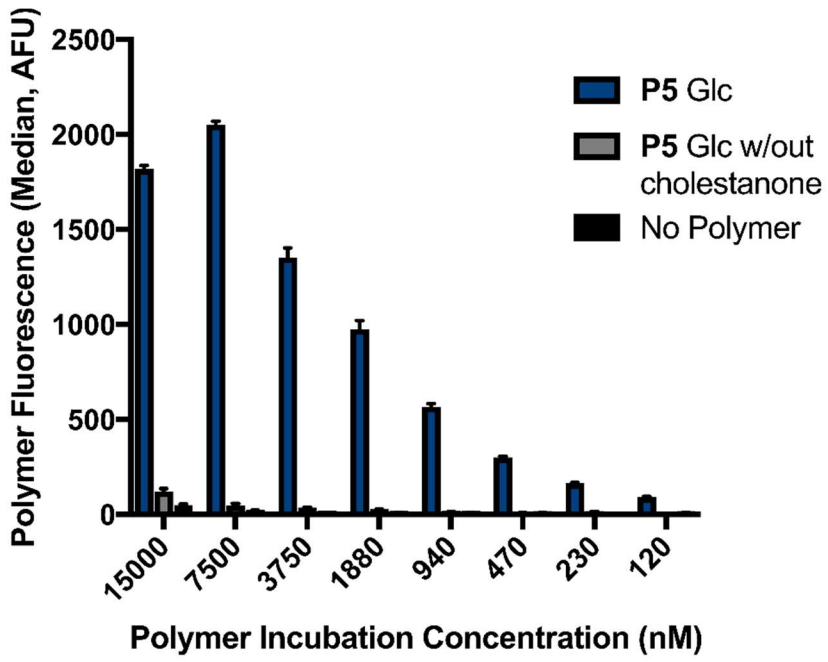


Figure 2.12. Relative incorporation of **P5** polymers vs equivalent polymer without cholestanone

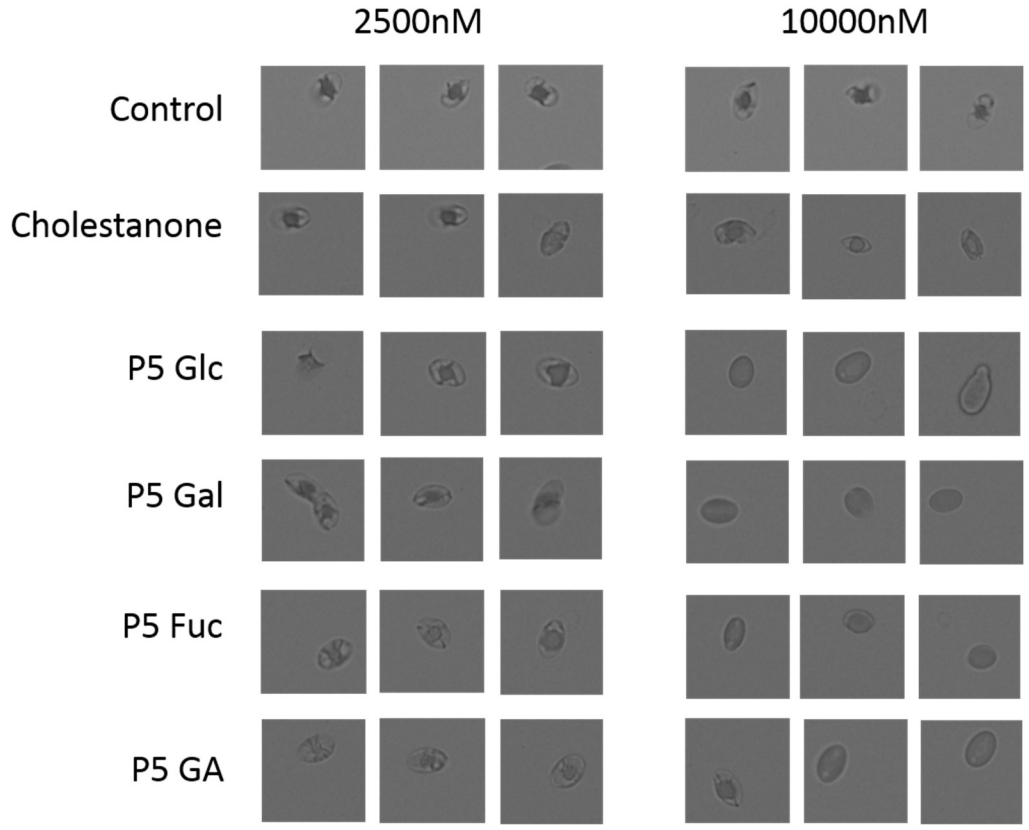
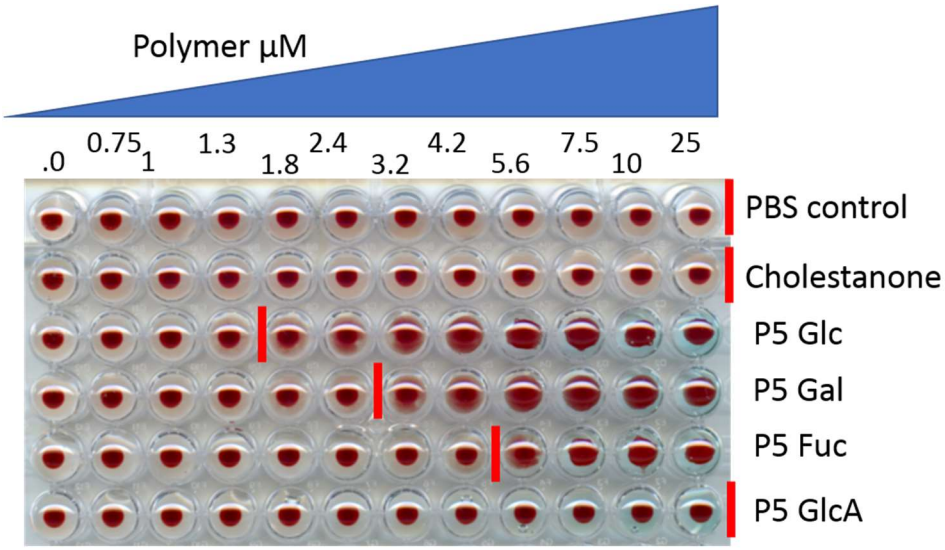


Figure 2.13. Bright field optical microscopy images of remodeled RBCs (associated with Figure 2.3C).



Lectin agglutination of RBCs remodeled with glycopolymers 5.

Figure 2.14. Sedimentation properties of RBCs remodeled with glycopolymers **5.**

*red line indicates point of settling inhibition.

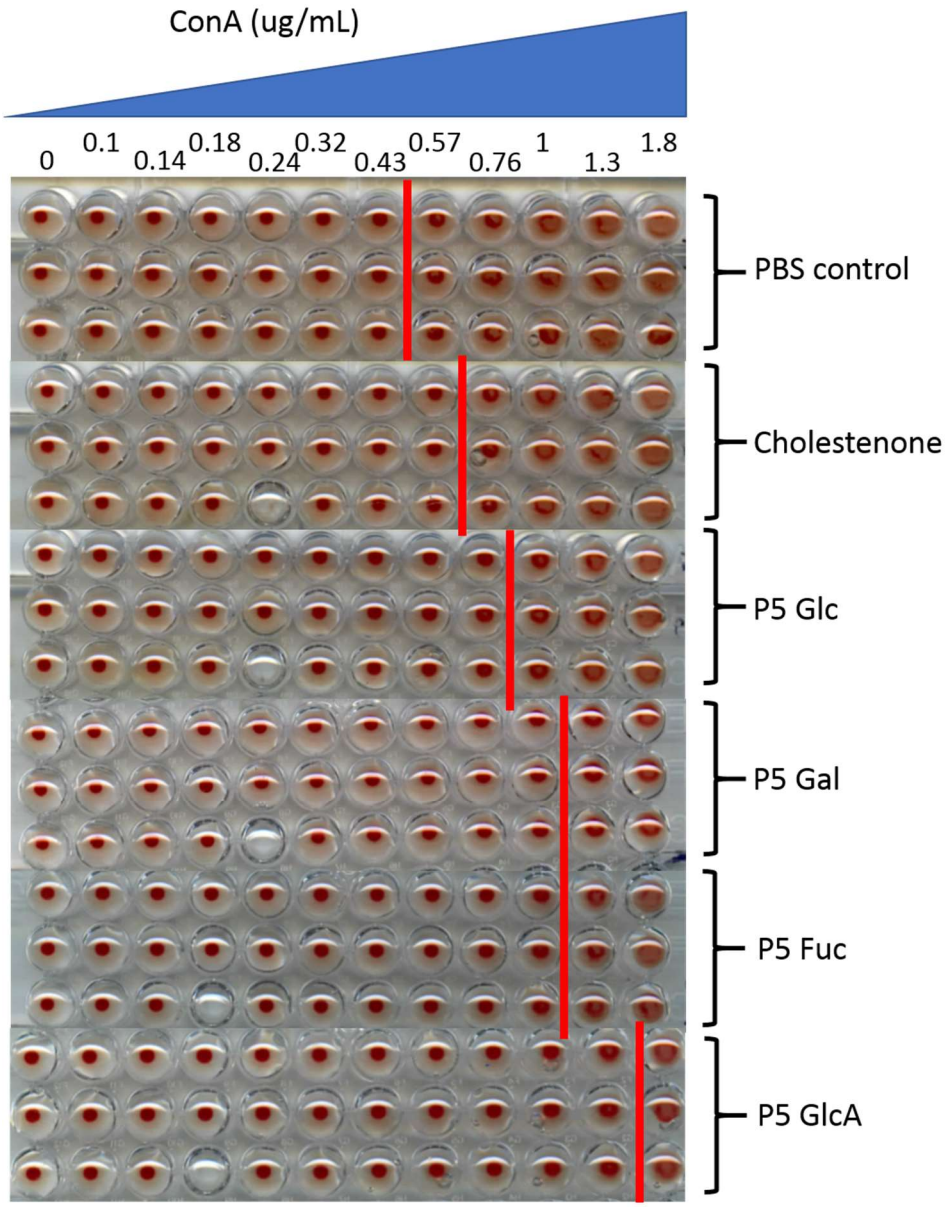


Figure 2.15. ConA agglutination of RBCs treated with alkynyl cholestanone **2** or polymers **5** (2.5 μM).

*Red line indicates point of agglutination.

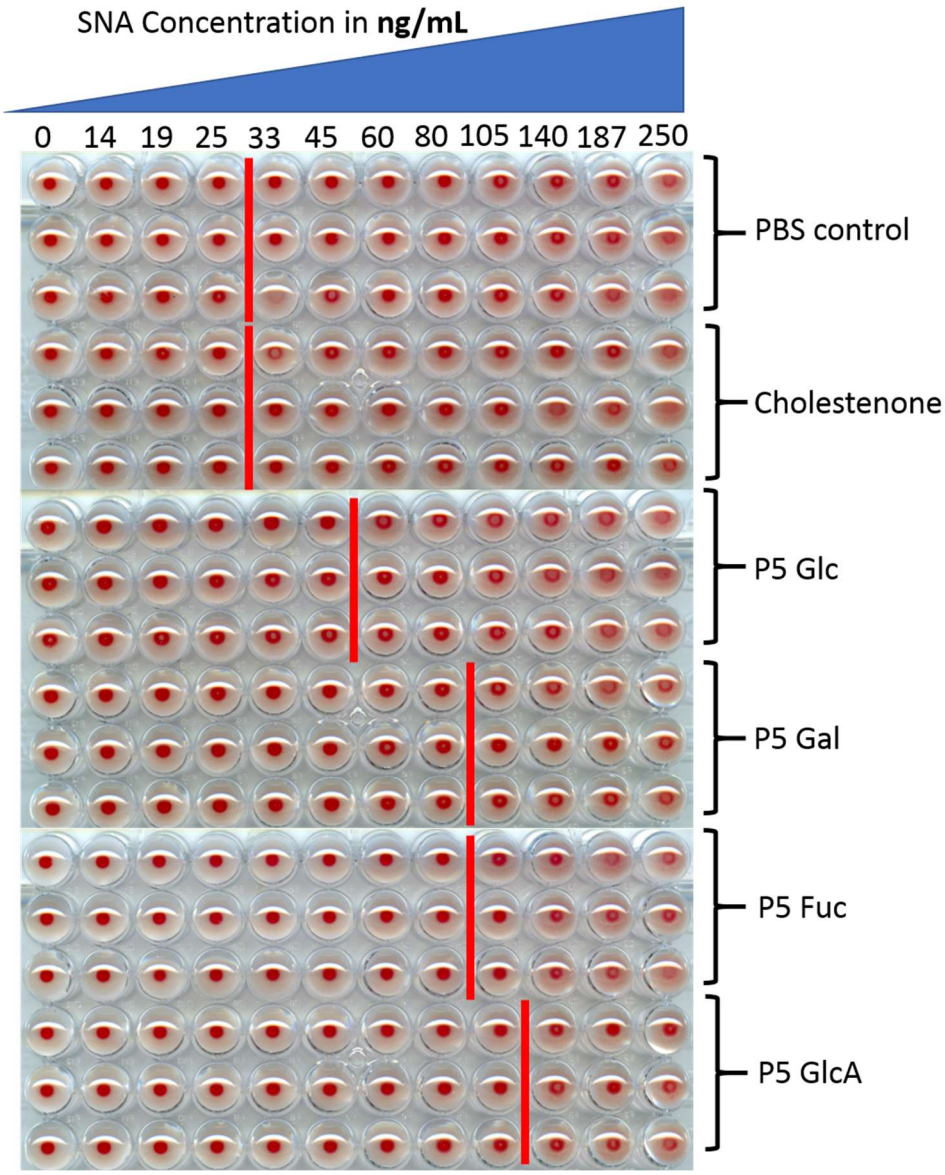


Figure 2.16. SNA agglutination of RBCs treated with alkynyl cholestanone **2** or polymers **5** (2.5 μM).

*Red line indicates point of agglutination.

Lectin binding to RBCs remodeled with glycopolymers 5 via flow cytometry.

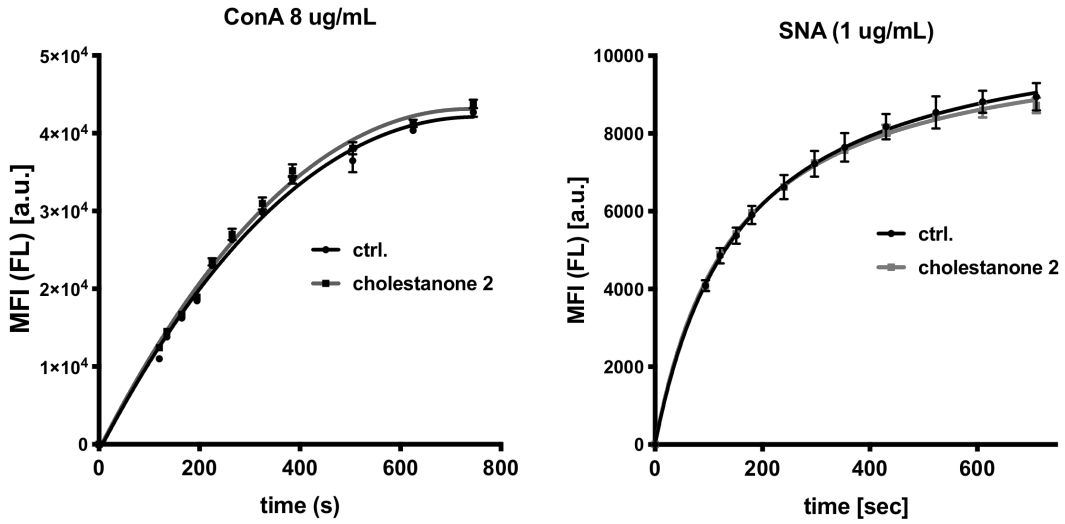
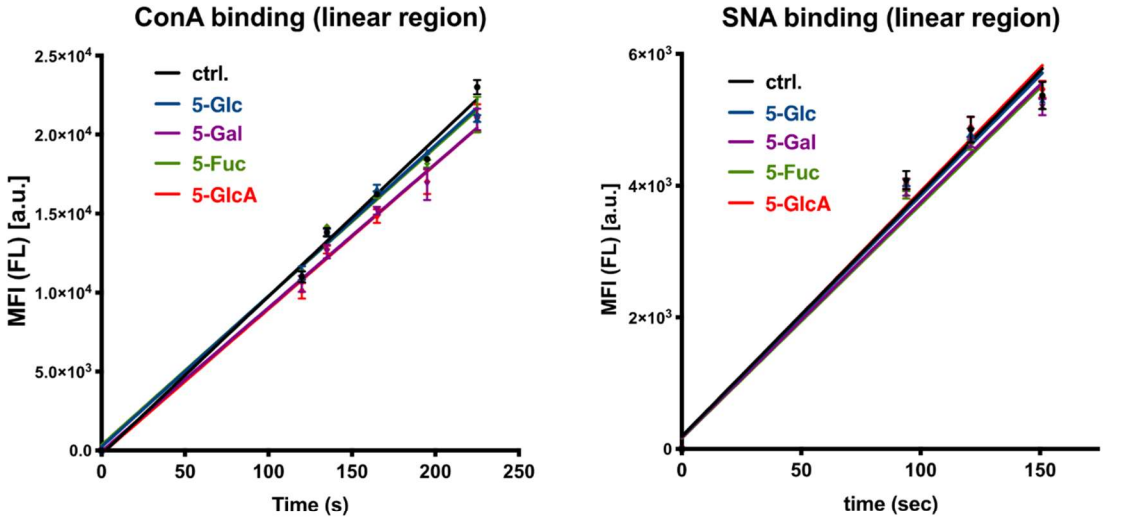


Figure 2.17. Binding of ConA and SNA to RBCs with pre-treatment with alkynyl cholestanone **2** (associated with Figure 4).



ConA (8 μg/mL)			
polymer	slope [s^{-1}]	R^2	P value (vs. ctrl)
-/ctrl.	99.90 ± 2.781	0.9923	-
5-Glc	95.45 ± 2.395	0.9937	0.730 (ns)
5-Gal	91.07 ± 2.699	0.9913	0.0023 (**)
5-Fuc	94.29 ± 2.930	0.9904	0.0287 (*)
5-GlcA	91.65 ± 2.899	0.9901	0.0036 (**)

SNA (1 μg/mL)			
polymer	slope [s^{-1}]	R^2	P value (vs. ctrl)
-/ctrl.	36.98 ± 4.005	0.9771	-
5-Glc	36.56 ± 4.016	0.9764	0.9999 (ns)
5-Gal	35.69 ± 3.695	0.9790	0.9956 (ns)
5-Fuc	35.49 ± 3.448	0.9815	0.9916 (ns)
5-GlcA	37.49 ± 3.488	0.9830	0.9998 (ns)

Figure 2.18. Initial rates of ConA and SNA binding to RBCs (associated with Figure 2.4).

Section 2.7. References

[1]. Weinbaum S, Tarbell JM, Damiano ER. *Annu. Rev. Bio. Eng.* 2007, **9**, 121–167. The structure and function of the endothelial glycocalyx layer.

[2]. Varki A, Gagneux P. *Biological Functions of Glycans*. In: Varki A, Cummings RD, Esko JD, editors. Essentials of Glycobiology 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, Chapter 7, 2015-2017.

[3]. Cummings, R. D., Schnaar, R. L., Esko, J. D., Drickamer, K. & Taylor, M. E. *Principles of Glycan Recognition.* In: Varki A, Cummings RD, Esko JD, editors. Essentials of Glycobiology 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, Chapter 29, 2015-2017.

[4]. Brewer, C. F., Miceli, M. C. & Baum, L. G. Clusters, bundles, arrays and lattices: novel mechanisms for lectin–saccharide-mediated cellular interactions. *Curr. Opin. Struct. Biol.* 2002, **12**, 616–623.

[5]. Huang, M. L. & Godula, K. Nanoscale materials for probing the biological functions of the glycocalyx. *Glycobiology* 2016, **26**, 797–803.

[6]. Bhatia, S., Camacho, L. C. & Haag, R. Pathogen Inhibition by Multivalent Ligand Architectures. *J. Am. Chem. Soc.* 2016, **138**, 8654–8666; Mammen, M., Choi, S.-K. & Whitesides, G. M. Polyvalent Interactions in Biological Systems: Implications for Design and Use of Multivalent Ligands and Inhibitors. *Angew. Chem. Int. Ed.* 1998, **37**, 2754–2794; Gestwicki, J. E., Cairo, C. W., Strong, L. E., Oetjen, K. A. & Kiessling, L. L. Influencing Receptor−Ligand Binding Mechanisms with Multivalent Ligand Architecture. *J. Am. Chem. Soc.* 2002, **124**, 14922–14933.

[7]. Kiessling, L. L., Gestwicki, J. E. & Strong, L. E. Synthetic Multivalent Ligands as Probes of Signal Transduction. *Angew. Chem. Int. Ed Engl.* 2006, **45**, 2348–2368.

[8]. Purcell, S. C. & Godula, K. Synthetic glycoscapes: addressing the structural and functional complexity of the glycocalyx. Interface Focus 2019, 9: 20180080.

[9]. Rabuka, D., Forstner, M. B., Groves, J. T., Bertozzi, C. R. Noncovalent Cell Surface Engineering: Incorporation of Bioactive Synthetic Glycopolymers into Cellular Membranes. *J. Am. Chem. Soc.* 2008, **130**, 5947–5953.

[10]. Paszek, M. J. DuFort, C.C., Rossier, O., Bainer, R., Mouw, J.K., Godula, K., Hudak, J.E., Lakins, J.N., Wijekoon, A.C., Cassereau, L., Rubashkin, M.G., Magbanua, M.J., Thorn, K.S., Davidson, M.W., Rugo, H.S., Park, J.W., Hammer, D.A., Giannone, G., Bertozzi, C.R., Weaver, V.M. The cancer glycocalyx mechanically primes integrin-mediated growth and survival. *Nature,* 2014, **511**, 319–325

[11]. Bradley, A. J., Murad, K. L., Regan, K. L. & Scott, M. D. Biophysical consequences of linker chemistry and polymer size on stealth erythrocytes: size does matter. *Biochim. Biophys. Acta BBA - Biomembr.* 2002, **1561**, 147–158.

[12]. Chapanian, R. Constantinescu, I., Rossi, N.A.A., Medvedev, N., Brooks, D.E., Scott, M.D., Kizhakkedathu, J.N. Influence of polymer architecture on antigens camouflage, CD47 protection and complement mediated lysis of surface grafted red blood cells. *Biomaterials,* 2012, **33**, 7871–7883.

[13]. Siren, E. M. J., Chapanian, R., Constantinescu, I., Brooks, D. E. & Kizhakkedathu, J. N. Oncotically Driven Control over Glycocalyx Dimension for Cell Surface Engineering and Protein Binding in the Longitudinal Direction. *Sci. Rep.* 2018, **8**, 7581.

[14]. Hattrup, C. L. & Gendler, S. J. Structure and Function of the Cell Surface (Tethered) Mucins. *Annu. Rev. Physiol.* 2008, **70**, 431–457.

[15]. Bansil, R. & Turner, B. S. Mucin structure, aggregation, physiological functions and biomedical applications. *Curr. Opin. Colloid Interface Sci.* 2006, **11**, 164–170.

[16]. Godula, K., Rabuka, D., Nam, K. T. & Bertozzi, C. R. Synthesis and Microcontact Printing of Dual End-Functionalized Mucin-like Glycopolymers for Microarray Applications. *Angew. Chem. Int. Ed Engl.* 2009, **48**, 4973–4976.

[17]. Kramer, J. R., Onoa, B., Bustamante, C. & Bertozzi, C. R. Chemically tunable mucin chimeras assembled on living cells. *Proc. Natl. Acad. Sci.* 2015, **112**, 12574–12579.

[18]. Gervais, M., Labbé, A., Carlotti, S. & Deffieux, A. Direct Synthesis of α-Azido,ω-hydroxypolyethers by Monomer-Activated Anionic Polymerization. *Macromolecules*, 2009, **42**, 2395–2400.

[19]. 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**, 3057–3064..

[20]. Alarcón-Manjarrez, C., Arcos-Ramos, R., Álamo, M. F. & Iglesias-Arteaga, M. A. Synthesis, NMR and crystal characterization of dimeric terephthalates derived from epimeric 4,5-seco-cholest-3-yn-5-ols. *Steroids,* 2016, **109**, 66–72

[21]. Meyer, J., Keul, H. & Möller, M. Poly(glycidyl amine) and Copolymers with Glycidol and Glycidyl Amine Repeating Units: Synthesis and Characterization. *Macromolecules,* 2011, **44**, 4082–4091.

[22]. Collins, B. E. Blixt, O., DeSieno, A.R., Bovin, N., Marth, J.D., Paulson, J.C. Masking of CD22 by cis ligands does not prevent redistribution of CD22 to sites of cell contact. *Proc. Natl. Acad. Sci.* 2004, **101**, 6104–6109.

[23]. Kesimer, M. Ehre, C., Burns, K.A., Davis, C.W., Sheehan, J.K., Pickles, R.J. Molecular organization of the mucins and glycocalyx underlying mucus transport over mucosal surfaces of the airways. *Mucosal Immunol.* 2013, **6**, 379–392.

[24]. Hudak, J. E., Canham, S. M., Bertozzi, C. R. Glycocalyx engineering reveals a Siglec-based mechanism for NK cell immunoevasion. *Nat. Chem. Biol.* 2013, **10**, 69–75.

[25]. Gandhi, J. G., Koch, D. L. & Paszek, M. J. Equilibrium Modeling of the Mechanics and Structure of the Cancer Glycocalyx. *Biophys. J.* 2019, **116**, 694–708.

[26]. Hudak, J. E. & Bertozzi, C. R. Glycotherapy: New Advances Inspire a Reemergence of Glycans in Medicine. *Chem. Biol.* 2014, **21**, 16–37.

[27]. Liang C-H, Wang S-K, Lin C-W, Wang C-C, Wong C-H, Wu C-Y. Effects of Neighboring Glycans on Antibody-Carbohydrate Interaction. *Angew. Chem. Int. Ed.* 2011, **50**, 1608–1612.

[28]. Shivatare VS, Shivatare SS, Lee C-CD, Liang CH, Liao K-S, Cheng Y-Y, Saidachary G, Wu C-Y, Lin N-H, Kwong PD, Burton DR, Wu C-Y, Wong C-H. Unprecedented Role of Hybrid N-Glycans as Ligands for HIV-1 Broadly Neutralizing Antibodies. *J. Am. Chem. Soc.* 2018, **140**, 5202–5210

[29]. Cohen, M., Hurtado-Ziola, N. & Varki, A. ABO blood group glycans modulate sialic acid recognition on erythrocytes. *Blood,* 2009, **114**, 3668–3676.

[30]. Cohen, M. and Varki, A. Modulation of Glycan Recognition by Clustered Saccharide Patches. In *International Review of Cell and Molecular Biology*, Vol. *308*, Burlington: Academic Press, **2014**, pp. 75-125.

[31]. Morotti, A. L. M., Lang, K. L., Carvalho, I., Schenkel, E. P. & Bernardes, L. S. C. Semi-Synthesis of new glycosidic triazole derivatives of dihydrocucurbitacin B. *Tetrahedron Lett.* 2015, **56**, 303–307.

[32]. Manabe, Y. Kasahara, S., Takakura, Y., Yang, X., Takamatsu, S., Kamada, Y., Miyoshi, E., Yoshidome, D., Fukase, K. Development of α1,6-fucosyltransferase inhibitors through the diversity-oriented syntheses of GDP-fucose mimics using the coupling between alkyne and sulfonyl azide. *Bioorg. Med. Chem.* 2017, **25**, 2844–2850.

[33]. Sharma, D. K., Lambu, M.R., Sidiq, T., Khajuria, A., Tripathi, A.K., Yousuf, S.K., Mukherjee, D. Ammonium chloride mediated synthesis of alkyl glycosides and evaluation of their immunomodulatory activity. *RSC Adv.* 2013, **3**, 11450–11455.

[34]. Izawa, K. & Hasegawa, T. Tosylated and azidated inulins as key substrates for further chemical modifications to access inulin-based advanced materials: An inulin-based glycocluster. *Bioorg. Med. Chem. Lett.* 2012, **22**, 1189–1193.

[35]. Morotti, A. L. M., Lang, K. L., Carvalho, I., Schenkel, E. P. & Bernardes, L. S. C. Semi-Synthesis of new glycosidic triazole derivatives of dihydrocucurbitacin B. *Tetrahedron Lett.* 2015, **56**, 303–307.

Chapter 3. Glycocalyx crowding with synthetic mucin mimetics strengthens interactions between soluble and virus-associated lectins and cell surface glycan receptors.

Membrane-associated mucins protect epithelial cell surfaces against pathogenic threats by serving as non-productive decoys that capture infectious agents and clear them from the cell surface and by erecting a physical barrier that restricts their access to target receptors on host cells. However, the mechanisms through which mucins function are still poorly defined due to a limited repertoire of tools available for tailoring their structure and composition in living cells with molecular precision. Using synthetic glycopolymer mimetics of mucins, we modeled the mucosal glycocalyx on red blood cells (RBC) and evaluated its influence on lectin (SNA) and virus (H1N1) adhesion to endogenous sialic acid receptors. The glycocalyx inhibited the rate of SNA and H1N1 adhesion in a size- and density-dependent manner, consistent with current view of the mucins as providing a protective shield against pathogens. Counterintuitively, increasing density of the mucin mimetics enhanced the retention of bound lectins and viruses. Careful characterization of SNA behavior at the RBC surface using a range of biophysical and imaging techniques revealed lectin-induced crowding and reorganization of the glycocalyx with concomitant enhancement in lectin clustering, presumably through the formation of a more extensive glycan receptor patch at the cell surface. Our findings indicate that glycan-targeting pathogens may exploit the biophysical and biomechanical properties of mucins to overcome the mucosal glycocalyx barrier.

Significance

Like other animal hosts, humans are constantly challenged by pathogens. This has led to an evolution of physical barriers coating all mucosal tissues, which are most vulnerable to infection. An important part of this defense is a dense brush of large proteins, called mucins, which are heavily decorated with sugars and keep pathogens at bay. Deciphering how pathogens overcome the mucin barrier is necessary to understand early stages of infection and to develop more effective treatments. By artificially installing the mucin-like shield on the surfaces of cells using synthetic sugar-bearing polymers, we have discovered a new physical mechanism by which proteins and viruses can exploit this barrier to more strongly adhere to their targets.

Section 3.1 Introduction

Many viral and bacterial pathogens that circulate in the human population, and continuously pose a risk of disease outbreaks and pandemics, utilize glycans on host cells to facilitate their adhesion and initiate infection.[1] Sialic acids (Sias) represent one such cell surface glycan receptor frequently targeted by pathogens, such as the Influenza A virus (IAV) among others (**Fig 3.1a**). The Sia monosaccharide is typically added as a terminal modification on more complex glycans linked to membrane-associated proteins or glycolipids and as a result is prominently displayed at high concentrations on mammalian cells.[2] Pathogen-associated proteins that recognize Sias, such as hemagglutinins (HA) in the case of IAVs, often share a low affinity binding site for the individual receptor structures but achieve high avidity binding to Sias displayed on cells through oligomerization and multivalency.[3,4,5]

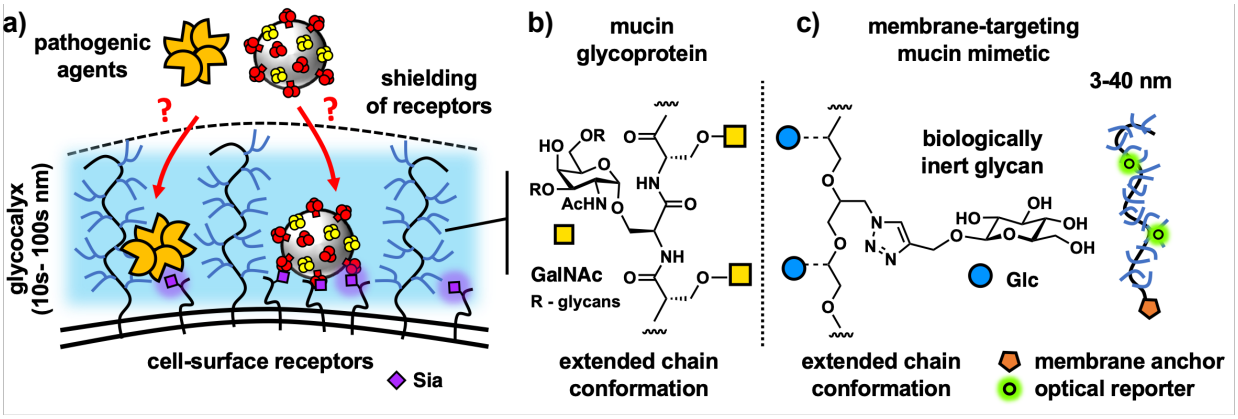


Figure 3.1. The mucosal glycocalyx provides a physical barrier against pathogen association with host cells. a) A brush of membrane-associated mucin glycoproteins restricts the ingress of pathogens toward cell surface glycan receptors, such as sialic acids (Sias). b) The extended conformation of mucins and their physical properties arise from dense glycosylation of their core peptide backbone. c) Cell surface targeting glycopolymers based on a linear poly(ethylene oxide) (PEO) scaffold, which mimic the architecture and physical properties of mucins, are used to investigate mechanisms through which pathogens can breach the protective mucosal glycocalyx barrier.

In response to the constant pathogen challenge, the cells of mucosal epithelial tissues, which are most susceptible to infection, express soluble and membrane bound glycoproteins called mucins to limit pathogen adhesion and spread (**Fig 3.1a**).[6,7] Comprised of long polypeptide chains densely decorated with glycans (**Fig 3.1b**), membrane-associated mucins provide a dual protective function against infection by forming a dense extended biopolymer brush, restricting access to the cell surface and by presenting decoy receptors to capture and clear the pathogen.[8,9] Both the biophysical barrier[10] and the receptor decoy[11] functions of mucins have been shown to limit bacterial and viral infection, including by IAVs. While the interactions between pathogens and mucins based on glycan recognition have received substantial attention, the biophysical mechanisms allowing mucins to limit infection are yet to be examined experimentally in a systematic way. This is primarily due to the lack of tools for manipulating the architecture and glycosylation status of mucins with molecular precision in living cells. Membrane engineering with synthetic glycopolymers, which mimic the architecture and properties of mucins and can be chemically defined at the molecular level,[12,13] has emerged as a convenient approach to study phenomena associated with the mucin glycocalyx, including its interactions with sialic acid-binding proteins[14] and the effects of its biophysical properties on cellular interactions and functions.[15,16] We have recently observed that introducing mucin mimetics to the surfaces of cells resulted in increased glycocalyx crowding, which limited lectin binding to membrane-associated glycoconjugates.[17] Bertozzi and co-workers further showed that mucin mimetics based on synthetic glycopeptides imbedded in supported lipid bilayers could shield sialylated glycosphingolipids from H1N1.[18] Studies with synthetic mucin analogs can reveal mechanistic insights into how mucins may limit pathogen adhesion through biophysical means.

To counteract the protective mucin shield, many pathogens express enzymes that degrade mucins[19,20] or change their glycosylation status.[21] While the complete removal of mucins by microbial proteases disrupts the integrity of the glycocalyx and exposes the cell surface to pathogen attack, the removal of only their receptor decoys with glycosidases, such as the IAV neuraminidase (NA), leaves the physical mucin barrier intact. The mechanism through which pathogens that do not break down the mucin shield overcome this physical barrier are yet to be fully investigated.

In this study, we have deployed synthetic glycopolymers with tunable lengths (**Fig 3.1c**) to model membrane-associate mucin displays on the surfaces of avian (turkey) red blood cells (RBCs) and examined how the mucin-mimetic glycocalyx size and density influenced the kinetics and thermodynamics of lectin (SNA) and viral (H1N1) binding to endogenous Sia receptors. We have observed polymer length-dependent inhibition of SNA and H1N1 adhesion to their cell surface receptors in the presence of the membrane-tethered mucin mimetic brush, providing further support for the mucin shield model. We have also identified a biophysical mechanism by which lectins and viruses exploit the crowded glycocalyx to strengthen their association with the cell.

Section 3.2 Results

Glycocalyx engineering with mucin mimetics alters RBC morphology and properties. To model key physical features of the epithelial glycocalyx; i.e., its dense, extended brush-like organization of mucin glycoproteins, we have generated synthetic mucin-mimetic glycopolymers (**GPs**) with defined length and glycosylation (**Fig 3.1c**). The mucin-mimetic glycopolymers (**Fig 3.2a**) were assembled on linear hydrophilic poly(ethylene oxide) (PEO) scaffolds of increasing length (DP = 30, 140, 440) and with narrow dispersity ($Đ < 1.3$). The

precursor backbones were generated via controlled anionic ring-opening polymerization of epichlorohydrin[22] followed by functionalization with azide side-chains for the attachment of propargyl glycosides via copper-catalyzed azide-alkyne cycloaddition (CuAAc)[23] (**Scheme 3.S1**). The efficiency of this "click" process resulted in a high frequency of glycosylated side-chains in close proximity to the PEO backbone, forcing the glycopolymers into extended conformations characteristic of mucins.[17,24] Seeking to isolate the biophysical properties of mucins from their glycan interactions, we modified our glycopolymers with glucose, a monosaccharide typically not found in mucin glycans (**Fig 3.1**). This produced mucin mimetic "spectators" lacking Sia receptors for IAVs but able to tune the size and density of the glycocalyx. The mucin mimetics were also endowed with a hydrophobic 4,5-*seco*-cholesten-5-one end-group for anchoring of the glycopolymers in cell membranes and a small fraction (< 1%) of their sidechains were labeled with AlexaFluor 488 (AF488) or Cyanine 3 (Cy3) dyes for detection via fluorescence (**Fig 3.1**). To examine the impact of glycocalyx size on molecular recognition events at the cell surface, we generated short (**GP-S**, DP = 30), medium (**GP-M**, DP=140), and long (**GP-L**, DP=440) mucin mimetics (**Fig 3.2A and 3.2B**).[17]

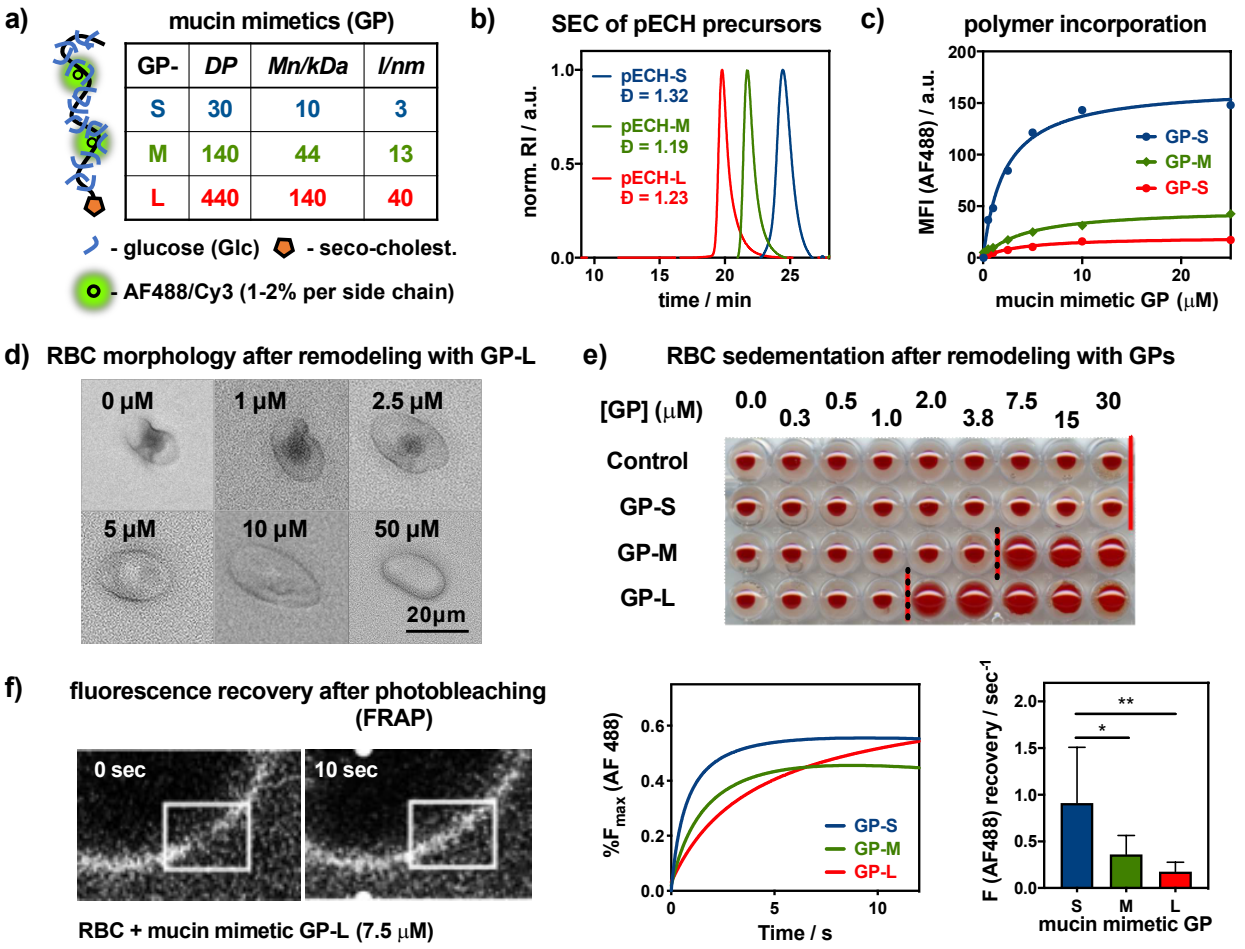


Figure 3.2. Construction and characterization of a mucosal glycocalyx model. **a)** Short (**S**), medium (**M**), and long (**L**) mucin mimetic glycopolymers **GP** ranging in size from ~ 3 to 40 nm were generated from poly(epichlorohydrin) (pECH) precursors (DP = degree of polymerization, *Mn* = number average molecular weight, *l* = estimated end to end length). **b)** Size exclusion chromatography (SEC) of pECH precursors indicates narrow molecular weight distributions of the polymer scaffolds. The polymers were modified with biologically inert glucose sidechains, fluorescent probes (AF488 or Cy3) for visualization, and chain end-terminated with hydrophobic cholestenone membrane anchor. **c)** Incorporation of AF488-labeled mucin mimetics **GPs** into RBC membranes was analyzed by flow cytometry. d) Optical microscopy and light scattering reveal increased cell rounding with increasing glycocalyx density (shown for **GP-L**). **e)** Sedimentation of RBCs remodeled with increasing concentration of all three glycopolymers (c_{GP} = 0.25 to 30 μM) **f)** Fluorescence recovery after photobleaching (FRAP) analysis shows length-dependent diffusion of AF488-labeled mucin mimetics **GP** in RBC membranes. Lines represent the average signal from n = 6 cells, c_{GP} = 7.5 uM), *p-values were determined by one-way ANOVA with multiple comparisons (* < 0.05, ** < 0.01)*.

RBCs provide a useful model for examining the effects of glycocalyx properties on lectin and IAV interactions with sialoglycan receptors. RBCs present endogenous Sia modifications in both α2,3 and α2,6 linkages and are commonly used to assess IAV binding activity via cell hemagglutination. Even though this cell type is not the target of IAV infection in mammals, RBCs also offer a useful platform to study mucin-related phenomena by providing an intermediary between simple laboratory membrane models based on supported lipid bilayers and lipid vesicles containing chemically defined glycoconjugates and the dynamic and heterogeneous environment of the epithelial cell glycocalyx. The RBC present compositionally complex but compact native glycocalyx (~ 5-10 nm),[25] which can be augmented with glycopolymers to introduce extended mucin-like features. At the same time, RBCs lack endocytic activity, which allows for installing mucin-mimetic brushes at high-density and compositional stability.

To introduce mucin-like features into the RBC glycocalyx, the cells were treated with the lipid-terminated glycopolymers **GP** of all three lengths at increasing concentrations (c_{GP} = 0.05 – 50.00 μM) for 1 hour. Membrane incorporation of the fluorescently labeled polymers was measured by flow cytometry (**Fig 3.2C**). The efficiency of membrane insertion was inversely proportional to polymer length, resulting in increasingly extended but less dense glycocalyx structures when the cells were treated at equal polymer concentrations. Polymers lacking the seco-cholestenone hydrophobic anchor did not insert into the cell membranes (**Fig 3.S12**).

Examination of the cells with light microscopy revealed pronounced rounding and loss of membrane features with increasing incorporation of the medium and long polymers (**Fig 3.2D**, shown for **GP-L**). Flow cytometry analysis of RBCs after treatment with mucin

mimetics of all three lengths (c_{GP} = 50 μM) showed negative correlation between **GP** length and side scatter (SSC) with little change in forward scatter (FSC) intensity (**Fig 3.S13**). This supports the visual observation of a polymer size-dependent decrease in cell complexity without significant alterations of the overall cell size (**Fig 3.2D**). The changes in cell morphology likely arise from entropic pressures exerted on the cell membrane upon a mushroom-to-brush conformational transition of the tethered mucin-mimetics in response to glycocalyx crowding, as proposed recently by Paszek and co-workers in their analysis of the role of mucins in cell-shape regulation.[26] The transition in cell shape was accompanied by changes in sedimentation properties of the remodeled RBCs. We identified polymer concentration thresholds of ~ 7.5 μM and 2.0 μM for **GP-M** and **GP-L**, respectively, above which cell sedimentation becomes less efficient (**Fig 3.2E**). No such effect was detected for the shortest polymers **GP-S** over the tested range of concentrations. RBC aggregation and sedimentation are processes facilitated by the cells' discoid shape and surface charge, which are both features affected by the introduction of the extended mucin-like glycocalyx.[27]

The morphological changes in the cells induced by the mucin mimetics indicate an increase in glycocalyx crowding, which should likewise influence the lateral membrane diffusion of the glycopolymers. Taking advantage of the fluorophore labels present in the polymers, we performed fluorescence recovery after photobleaching (FRAP) analysis on cells treated with **GPs** of all three lengths at the threshold concentration, c_{GP} = 7.5 μM, at which the medium-size mucin mimetic **GP-M** induced alterations in cell shape and sedimentation properties (**Fig 3.2F**). This concentration is above the transition concentration for the long polymer **GP-L** and close to the surface saturation level for the short mimetic, **GP-S**, which did not alter sedimentation. The FRAP analysis revealed that all three glycopolymers showed lateral diffusion within the endogenous glycocalyx, with fluorescence recovery rates

decreasing from 0.91 ± 0.24 s^{-1} for **GP-S**, to 0.36 ± 0.08 s^{-1} for **GP-M**, to 0.18 ± 0.04 s^{-1} for the longest polymer, **GP-L**. Considering that the cell surface grafting efficiency for **GP-S** is significantly higher than for **GP-M** and **GP-L** (**Fig 3.S11**), the mobility of the mucin mimetics in the glycocalyx is, thus, influenced more strongly by their size rather than their density in the membrane.

Mucin mimetics slow the rate of lectin association while enhancing binding complex stability. We examined the effects of the mucin-mimetic glycocalyx size and density on the binding of *Sambucus nigra* agglutinin (SNA) to Sias on RBCs. Cells treated with **GPs** of all three lengths near the threshold concentration for inducing cell morphology change (c_{GP} = 7.5 μM) were incubated with a fluorescently labeled (AF647) SNA lectin at sub-agglutination concentration (c_{SNA} = 0.2 μg/mL). The lectin binding was assessed via flow cytometry at regular time points until signal saturation (~ 13 min) and the resulting data were fitted and benchmarked to untreated control cells to determine the relative rates (r_{rel}) of binding for each condition (**Fig 3.3a**). The medium and long glycopolymers **GP-M** and **GP-L** reduced the rates of SNA binding to a similar extent (r_{rel} = 0.84 +/- 0.05 and 0.88 +/- 0.05 s^{-1}, respectively), while the shortest mucin mimetic, **GP-S**, had no effect on the extent or rate of lectin association (r_{rel} = 1.02 +/- 0.05). These observations are consistent with glycopolymer shielding of endogenous glycan receptors[18] according to their size and is expected based on the dimensions of the RBC glycocalyx (~ 10 nm) and the lengths of the different mucin mimetics (~ 3, 13 and 40 nm, **Fig 3.2a**).

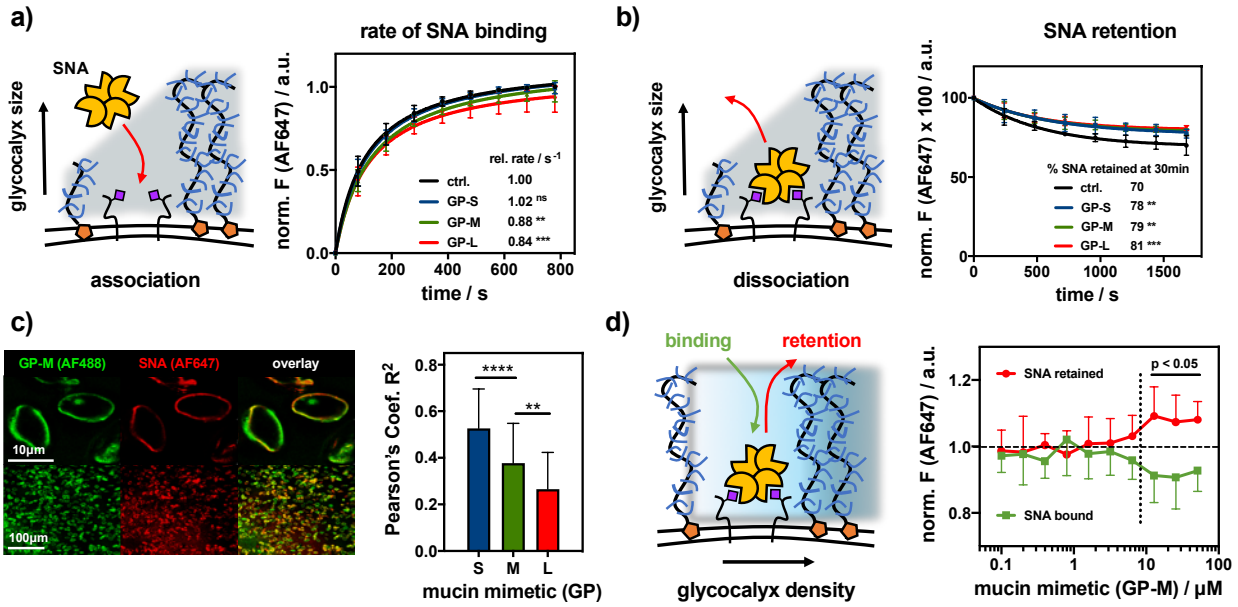


Figure 3.3. Spectator glycocalyx size and density regulate SNA interactions with cell surface receptors. **a)** Extended glycocalyx reduces the rate of SNA binding. Red blood cells (RBCs) were remodeled with short (**S**), medium (**M**) and long (**L**) mucin mimetics **GP** (c_{GP} = 7.5 µM) and the binding for AF647-SNA (c_{SNA} = 0.2 µg/mL) was measured by flow cytometry and normalized to untreated cells. (n=6 experimental replicates). **b)** Extended glycocalyx reduces SNA dissociation from RBCs. The remodeled cells equilibrated AF647-SNA were resuspended in pure buffer and lectin retention was measured by flow cytometry. (n=6 experimental replicates). **c)** Extended mucin mimetic glycocalyx drives segregation of SNA-receptor adhesion complexes. Florescence micrographs show representative confocal images of RBCs remodeled with mucin mimetic **GP-M** (green, c_{GP} = 7.5 µM) and stained with AF647-SNA (red, c_{SNA}= 0.2 µg/mL). Bar graph represents Pearson's correlation coefficient (R^2) analysis of images for SNA binding to RBCs remodeled with polymers of all three lengths (c_{GP} = 7.5 µM). (n > 35 individual cell images per polymer condition). **d)** Glycocalyx crowding limits SNA association with cell surface receptors while stabilizing the resulting adhesion complex. Association (green) and retention (red) of AF647-SNA at the surface of RBCs remodeled with **GP-M** at increasing concentration were measured by flow cytometry and normalized to untreated cells. (n=12 experimental replicates). *Values represent averages and standard deviations, p-values were determined by ANOVA (a-c) or student test (d), *<0.05, **<0.01, ***<0.001, ****<0.0001.*

To assess the effects of the mucin mimetics on lectin retention at the cell surface, the remodeled RBC were allowed to reach saturation binding with SNA (15 min), washed, and allowed to re-equilibrate in fresh PBS buffer (30 min). Flow cytometry was used to monitor the retention of the fluorescent SNA by RBCs over the course of the experiment and the

percentage of SNA remaining at equilibrium was calculated (**Fig 3.3b**). Polymers of all three lengths enhanced SNA retention by ~ 8-11% over untreated control cells. In contrast to lectin association, the impact of polymer length on SNA retention was much less pronounced, with the shortest polymer **GP-S** also exerting significant effect. This points to a change in the avidity of the SNA-receptor binding complex driven by polymer density rather than size. Fluorescence confocal microscopy showed increased exclusion of the mucin mimetics from SNA (**Fig 3c,** c_{GP} = 7.5 μM, c_{SNA} = 0.2 μg/mL), suggesting extensive lectin clustering in the presence of the polymers. Pearson's correlation analysis (R^2 values) showed a decrease in glycopolymer colocalization with SNA according to size (**GP-S**, $R^2 = 0.53 \pm 0.17$; **GP-M**, $R^2 = 0.38 \pm 0.17$; **GP-L**, $R^2 = 0.27 \pm 0.16$).

Our experiments indicate that increasing mucin mimetic size inhibits lectin association, their membrane crowding promotes the formation of more stable adhesion complexes, presumably through lectin clustering. In the mucin glycocalyx, which can be described as a surface-tethered polymer brush,[26] the extension of the glycoproteins away from the membrane can be driven by crowding. Similar behavior was also recently confirmed for synthetic glycopolymers in supported lipid bilayers.[18] To evaluate the effect of mucin-mimetic density in the RBC glycocalyx on SNA interactions, we determined equilibrium binding and retention of SNA at increasing membrane densities for the medium-sized polymer **GP-M** ($c_{GP\text{-}M}$ = 0.1 – 50.0 μM). We observed an emergence of SNA shielding at polymer concentration of 6.4 μM, which also coincided with the transition of the RBCs toward a more rounded phenotype with altered sedimentation properties (**Fig 3.2E**). Likewise, SNA dissociation became significantly inhibited at this threshold concentration (**Fig 3.3b**). Together, these observations support a density-dependent transition of the glycopolymers from coiled to extended conformations in the glycocalyx brush, as predicted by the current biophysical

model for the behavior of membrane-associated mucins.[26] The shielding ability of the mucin-mimetics was further enhanced at higher concentrations of SNA (**Fig 3.S16**), indicating that the lectin also contributes to glycocalyx crowding once bound to its target receptors.

Mucin mimetics drive lectin clustering by increasing glycocalyx crowding. The low levels of colocalization between the mucin mimetics and SNA in the RBC glycocalyx suggest increased clustering of the oligomeric lectin, which may stabilize the binding complex through crosslinking of neighboring sialylated glycoconjugates in the plasma membrane. To confirm enhanced lectin clustering by the glycopolymers, we measured Förster resonance energy transfer (FRET) between SNA probes labeled separately with AF594 (donor) or AF647 (acceptor) fluorophores by fluorescence lifetime imaging microscopy (FLIM). A shortening of the donor excited state lifetime (τ) indicates increased FRET efficiency, and thus reduced distance, between the donor and acceptor fluorophores.

First, RBCs were treated with the SNA labeled with AF594 only for 15 min and imaged by FLIM to establish the fluorescence lifetime of the donor (τ = 3.15 ± 0.02 ns) (**Fig 3.4a,** c_{SNA} = 0.2 μg/mL). Incubation of the cells with equimolar amounts of the donor AF594-SNA and acceptor AF647-SNA probes reduced the donor excited state lifetime (τ = 2.53 ± 0.09 ns), setting the baseline for SNA proximity in the RBC glycocalyx in the absence of the mucin mimetics. Repeating this experiment in the presence of the long glycopolymer **GP-L** (c_{GP} = 7.5 μM) further reduced the donor excited state lifetime (τ = 1.80 ± 0.07 ns), revealing closer spacing of the bound lectins. Control experiments with cells treated with the *seco*-cholestenone anchor only (c_{chol} = 7.5 μM) did not reduce lectin spacing (τ = 2.65 ± 0.21 ns) compared to RBC controls without mucin mimetic treatment, confirming that the glycopolymer domains were responsible for driving lectin clustering.

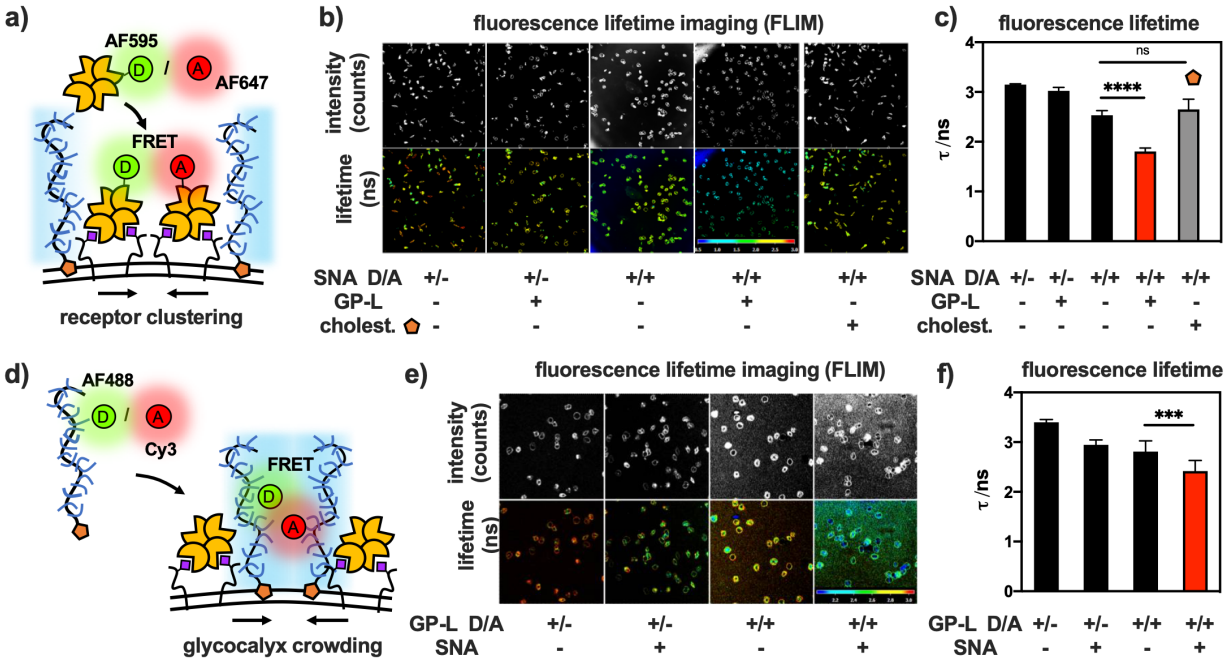


Figure 3.4. Glycocalyx crowding drives clustering of lectin-receptor adhesion complexes. **a**) Changes in SNA clustering in response to glycocalyx crowding were assessed based on Förster resonance energy transfer (FRET) between lectins labeled with donor (D = AF595) and acceptor (A = AF647) probes. **b**) Representative fluorescence lifetime imaging microscopy (FLIM) images and **c**) bar graph representations for the binding of SNA probes to RBCs before and after remodeling with **GP-L** (c_{GP} = 7.5 μM) or the hydrophobic anchor **S.5** alone ($c_{S.5}$ = 7.5 □M). Decrease in fluorescence lifetime (t) indicates closer SNA proximity. **d)** Mucin mimetics labeled with donor (D = AF488) an acceptor (A = Cy3) probes were employed to measure changes in glycocalyx crowding after SNA binding via FRET. **e**) Representative FLIM images and **f**) bar graph representations of RBCs remodeled with mucin-mimetic probes ($c_{GP\text{-}D/A}$ = 15 μM) before and after incubation with SNA. Decrease in fluorescence lifetime (t) indicates enhanced glycocalyx crowding. Color scales represent fluorescence lifetime (t/ns). *Values represent averages and standard deviations for representative image frames containing >10 cells, p-values were determined by student test, ***<0.001, ****<0.0001.*

In the absence of ligand-dependent molecular interactions between the glycopolymers and SNA, the observed clustering of the lectin is likely induced by the biophysical properties and membrane diffusion of the mucin mimetics. Enhanced expression of mucins has been shown to increase glycocalyx crowding, which can manifest through changes in membrane shape and the clustering of cell adhesion complexes.[26,16] We used FRET/FLIM to analyze the crowding of mucin mimetics in the RBC glycocalyx in response to SNA binding by measuring

changes in their proximity (**Fig 3.4b**). Long glycopolymers **GP-L** with identical size and glycan composition were synthesized to display donor (AF488) or acceptor (Cy3) fluorophores with overlapping emission and absorption profiles to enable FRET. RBCs were treated either with polymers containing the donor fluorophore only or with an equimolar mixture of polymers presenting both the donor and acceptor dyes (total c_{GP} = 15μM). FLIM imaging of RBCs remodeled with the glycopolymer FRET pair showed significantly reduced donor fluorophore lifetime from $\tau = 2.95 \pm 0.10$ ns to $\tau = 2.42 \pm 0.21$ ns after incubation with SNA (c_{SNA} = 0.2 μg/mL) for 15 min. The higher FRET efficiency due to shorter distances between the mucin mimetic probes provides experimental support for increased crowding of the glycocalyx environment in response to SNA binding.

Collectively, the FRET/FLIM experiments showed that the lectin clustering and mucin mimetic crowding phenomena are interconnected and occur simultaneously. If the mucin mimetics forced changes in receptor organization prior to lectin binding, we would not except to observe significant changes in FRET efficiency from the polymers after introduction of SNA. FRAP analysis shows that, once binding equilibrium is reached, the lectins become immobile in the RBC glycocalyx (**Fig 3.S17**), providing further support for the stabilization of the SNA binding complexes through clustering and receptor crosslinking.

Glycocalyx crowding with mucin mimetics enhances H1N1 virus retention on RBCs. We sought to evaluate whether the paradox observed for the effect of mucin mimetics on SNA binding, whereby the extended glycopolymers slowed the kinetics of lectin association while stabilizing the resulting binding complex, would also apply to viruses that utilize multivalent receptor interactions for adhesion to host cells.

First, we measured the effects of mucin mimetic length and membrane density on the adhesion of the H1N1 virus to RBCs. Cells remodeled with mucin mimetic glycopolymers **GPs** of all three lengths (c_{GP} = 7.5 μM) alongside untreated controls were incubated with H1N1 (33 HAU) for 15 min. The binding experiments were performed at room temperature to limit neuraminidase (NA) activity and viral release from the cell surface. After incubation, H1N1 binding was detected using a NA enzymatic assay with the fluorogenic reporter substrate, 4-methylumbelliferyl *N*-acetyl-α-neuraminic acid (4-MU-NANA) at 37°C.[28] The rate of fluorescence signal generation by neuraminidase activity was used to determine the concentration of RBC-bound viruses in the sample (**Fig 3.S18**). All three mucin mimetics reduced the overall amount of virus bound to the RBCs compared to non-treated cells (**Fig 3.5a**). The efficiency of viral binding inhibition correlated with polymer length, increasing from ~ 20% for the short mimetic **GP-S** to ~ 50% for the long polymer **GP-L**, consistent with enhanced shielding of cell-surface receptors by the progressively extended glycocalyx. The shielding capacity of the mucin mimetic glycocalyx against viral adhesion increased with membrane density of the polymers and became most effective at the threshold polymer concentration associated with changes in RBC shape and sedimentation properties (**Fig 3.5b**, $c_{GP\text{-}M}$ = 6.3 μM). As observed for SNA, increasing viral titers reinforced the protective properties of the mucin mimetics, presumably, by further increasing local crowding in the glycocalyx (**Fig 3.S19**).

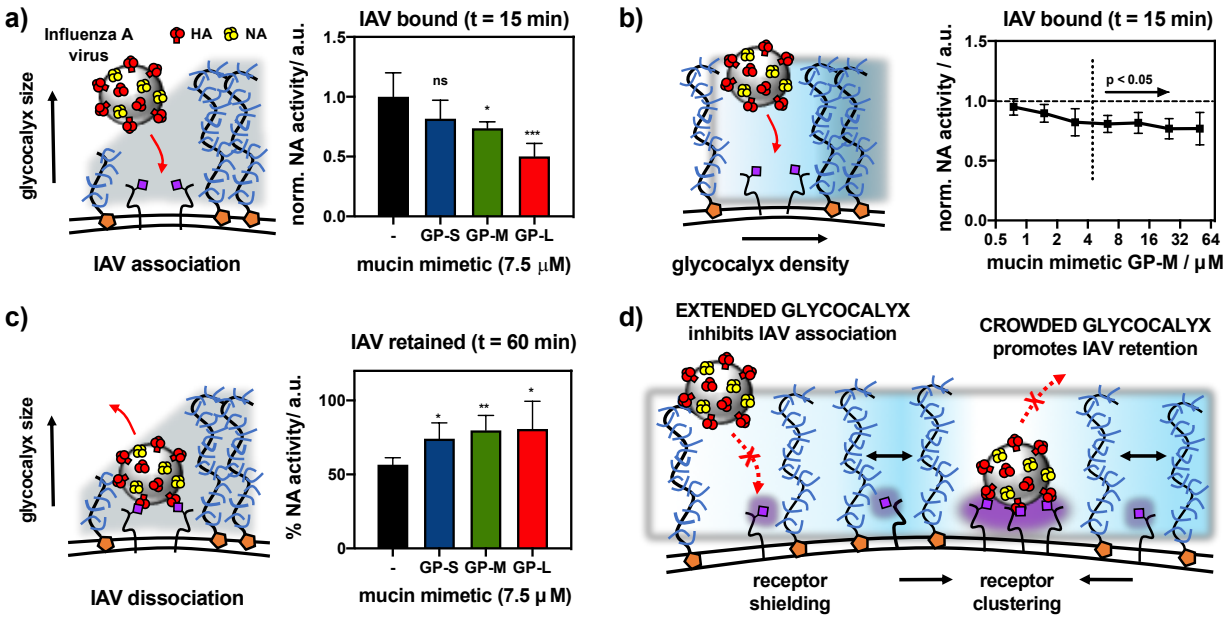


Figure 3.5. Influence of glycocalyx size and density on the binding of H1N1 viruses to sialic acid receptors on RBCs. **a)** Extended glycocalyx shields endogenous sialic acid receptors from H1N1 binding. Saturation binding of H1N1 (t = 15 min) to RBCs remodeled with short (**S**), medium (**M**) and long (**L**) mucin mimetics **GP** (c_{GP} = 7.5 μM) was assessed via their neuraminidase (NA) activity toward a fluorogenic substrate 4MU-NANA. **b)** Increased glycocalyx crowding limits viral adhesion. Saturation binding of H1N1 (t = 15 min) to RBCs remodeled with increasing concentrations of **GP-M** (c_{GP} = 0.8 - 50.0 μM) was evaluated via NA activity assay with 4MU-NANA. **c)** Extended glycocalyx enhances retention of viruses bound to RBC receptors. Retention of H1N1 viruses by RBCs remodeled with mucin mimetics **GP** of increasing length (c_{GP} = 7.5 μM) was measured as a fraction of NA activity before and after equilibration in fresh buffer (t = 60 min). **d)** Proposed model for the influence of a spectator mucin-type glycocalyx on cell-pathogen interactions. Extended and dense glycocalyx shields cell surface receptors from viral adhesion. Glycocalyx crowding drives receptor clustering and enhances retention of bound viruses. *Values represent averages and standard deviations for n = 6 experimental replicates; p-values were determined by t-test (a and c) and ANOVA (b): *<0.05, **<0.01, ***<0.001, ****<0.0001.*

Analysis of H1N1 retention on RBCs confirmed the ability of the mucin mimetic glycocalyx to stabilize the adhesion complex between the virus and its sialoglycan receptors at the cell surface (**Fig 3.5c**). RBCs remodeled with glycopolymers of all three lengths (c_{GP} =

7.5 μM) as well as untreated control cells were first allowed to equilibrate with H1N1 (33 HAU) for 30 minutes, then they were washed to remove any free virus and resuspended in free buffer for 60 min to allow for dissociation. The experiments were carried out at room temperature to minimize NA activity. The levels of virus bound to the RBC surface before and after dissociation were measured via the 4-MU-NANA fluorescence NA activity assay as described above. Mucin mimetic polymers of all three lengths increased the amount of H1N1 that was retained at the cell surface (75-80%) compared to untreated RBC controls (57 %). As observed for SNA, the effect of polymer size on enhancing H1N1 retention was less prominent compared to the inhibition of viral binding (**Fig 3.3 and Fig 3.5a**), pointing to a polymer density-driven stabilization of the adhesion complex.

Section 3.3 Discussion

Membrane-associated mucins contribute to the protection of epithelial surfaces against pathogen invasion in two distinct ways. Mucins present decoy receptors for capturing pathogens and shedding them from the cell surface. Their organization into a dense extended brush also presents a physical barrier blocking the pathogen from accessing functional receptors at the cell surface. Examining how each mechanism contributes to the protection of host cells against pathogen invasion may guide the development of new therapeutics targeting early stages of infection.

While a range of techniques exist to profile the glycan-binding specificity of pathogen-associated adhesion proteins, the repertoire of tools to study the biophysical properties of the mucin glycocalyx remains limited. Recently, genetic tools have begun to emerge that enable expression of recombinant mucins with precisely defined lengths and membrane densities in cells. This has provided new insights into how the composition and physical properties of the

mucin glycocalyx contribute to membrane shape generation[26] and modulation of cell adhesion.[29] Although the structure of mucins can now be tailored, tuning their glycosylation state (e.g., the removal or addition of sialic acids) without affecting other glycoproteins and glycolipids at the cell surface remains a challenge. This makes decoupling the decoy receptor and physical barrier functions of membrane-associated mucins in protecting cells against pathogens, particularly in instances, when both protective mechanisms are deployed simultaneously.

One example where membrane mucins serve both purposes is in preventing IAV adhesion to host cells. The mucins are heavily sialylated and capture IAVs by binding to the viral HA proteins. And while the Sia decoys are rapidly removed by viral NAs, the mucins remain on the cell surface and continue to pose a physical barrier against the virus. Genetic studies identified mucin MUC1 as an important factor in limiting H1N1 infection by shedding the bound virus from the cell surface after the mucin is proteolytically cleaved.[11] Although the inhibition of IAV binding by physically shielding non-mucin Sia receptors on the cell surface is yet to be demonstrated, studies with *H. pylori* lacking the mucin adhesins, BabA and SabA, showed inhibition of bacterial infection with increased MUC1 expression.[10]

To examine the biophysical principles through which membrane-associated mucins limit pathogen association with host cells, we have generated artificial glycocalyx models of the mucin like brush on the surfaces of RBCs, which present a biologically relevant environment with respect to receptor complexity and heterogeneity. Using synthetic glycopolymers with tunable size displaying non-participating model glycans and endowed with hydrophobic plasma membrane anchors and fluorescent optical probes, we were able to tailor the features of the mucin-mimetic brush with respect to size and density and

characterize its dynamics. Using our RBC-based mucin glycocalyx model, we systematically evaluated how the length and membrane density of the mucin mimetics influences both binding and dissociation of SNA and H1N1 at the cell surface. By selecting glucose as the side chain glycan residue for our mimetics, we ensured that the probes would not interact with SNA, H1N1, or any other endogenous glycan-binding protein on the RBC surface. Therefore, any observed changes in lectin or viral binding could be ascribed solely to the biophysical properties of the glycocalyx component.

Installation of the mucin mimetics on the RBC surface introduced significant glycocalyx crowding, which manifested by changes in cell shape toward more rounded morphology, consistent with recent studies showing that increased expression of mucins by cells enforced membrane curvature.[26] The change in morphology was accompanied by altered sedimentation properties of the cells, indicating changes in the physical properties of the glycocalyx. FRAP experiments confirmed that the polymers retained their mobility in the glycocalyx through lateral diffusion in the membrane and the rate of diffusion was inversely correlated to polymer length, indicating physical interactions of the extra-membrane glycopolymer domain with endogenous glycocalyx components. Analysis of SNA and H1N1 binding near the threshold polymer density associated with morphological and sedimentation changes in the RBCs slowed the association rate of the lectin by ~ 12-16 % and reduced the overall amount of virus captured at the cell surface by 20-50%. The protective effects of the mucin mimetics were polymer length- and membrane density-dependent, in agreement with the proposed mucin brush shielding model and the recent study by Bertozzi and coworkers reporting inhibition of H1N1 binding to gangliosides by mucin-mimetic brushes based on lactosylated poly-L-serine polymers anchored in supported lipid bilayers (SLBs).[18]

Unexpectedly, the mucin mimetics enhanced the retention of both SNA (8-11%) and H1N1 (18-23%) at the RBC surface compared to untreated cell controls. This effect was independent of glycopolymer size and appeared to be driven primarily by glycocalyx crowding. This suggests that glycan-binding proteins and pathogens may exploit the crowded glycocalyx interface of the epithelial mucosa to strengthen their association with host cells. Confocal fluorescence microscopy of RBC stained with SNA revealed increased exclusion of the mucin mimetics from the lectin adhesion sites based on their length. FLIM/FRET imaging further showed more extensive clustering of SNA in the presence of the mucin mimetics as well as increased crowding in the excluded mucin brush upon lectin binding.

Based on our observations, we propose a biophysical mechanism for strengthening of interactions between soluble and virus-associated oligomeric lectins and cell surface glycan receptors (**Fig 3.5d**). Therein, the binding of the protein to its receptors induces exclusion of the mucin-mimetics from the point of contact and induces local crowding in the glycocalyx. To alleviate the crowding developing in the mucin mimetic brush,[26] smaller glycoconjugates may diffuse back toward the adhesion complex. If these glycoproteins and glycolipids contain Sia modifications, this would generate a higher-valency adhesion patch increasing lectin clustering and overall avidity of the binding event. The limited effect of polymer size on SNA and H1N1 retention supports the proposed mechanism, as the density of mucins at the cell surface, not their length, has been previously identified as the primary driver of entropic pressure on the cell membrane influencing its shape and stability.[26] The rounding of RBCs with increasing mucin mimetic density at their membranes provide further evidence for glycocalyx crowding. The proposed mechanism does not require direct binding of the pathogenic adhesins to mucins but assumes multivalency of their interactions with glycan receptors in the glycocalyx. Since most glycan binding proteins utilize multivalency to achieve

avidity, we anticipate that this mechanism will be general and exploited by other pathogens that target glycan receptors in mucosal tissues.

Section 3.4 Conclusion

In this study, we have applied cell surface engineering with synthetic mimetics of mucin glycoproteins to model the mucosal glycocalyx. By tuning the nanoscale dimension and membrane density of the materials, we assessed the effectiveness of the physical barrier to restrict lectin and virus adhesion to target receptors. The extended and dense cell surface displays of mucin mimetics inhibited the rate of SNA and H1N1 binding, consistent with the current view of mucins acting as a physical shield protecting the cell surface. We discovered that the both the soluble and virus-associated oligomeric lectins exploited the crowded dynamic glycocalyx interface to strengthen their interactions with cell-surface receptors. The mucin mimetics were designed to be refractory to the tested glycan-targeting proteins, as such the observed behavior was driven by the biophysical properties of the glycocalyx and may be a general feature of the mucosal barrier.

Conflicts of interest

There are no conflicts to declare.

Section 3.5 Acknowledgements

We thank the UCSD Microscopy Core facility (via p30 grant NS047101 from NINDS) for assistance with fluorescence microscopy imaging, and the Glycobiology Research and Training Center for access to tissue culture facilities and analytical instrumentation. The authors acknowledge the use of facilities and instrumentation supported by the NSF through

the UC San Diego Materials Research Science and Engineering Center (UCSD MRSEC), DMR-2011924. We also wish to thank Dr. Christopher Fisher for help with H1N1 production and Taryn Lucas for her assistance with neuraminidase assays. This work was supported by the NIH Director's New Innovator Award (NICHD: 1DP2HD087954-01). K. G. is supported by the Alfred P. Sloan Foundation (FG-2017-9094) and the Research Corporation for Science Advancement via the Cottrell Scholar Award (grant # 24119).

Chapter 3, in full, is work in preparation for publication. The manuscript has been deposited in its current form on Bioarxiv and can be accessed at https://doi.org/10.1101/2021.05.07.443169. This work was contributed to by Meghan Altman, Pascal Gagneux, and Kamil Godula. Daniel Honigfort is the primary author of this material.

Section 3.6 Supporting Information

Glycocalyx crowding with synthetic mucin mimetics strengthens interactions between soluble and virus-associated lectins and cell surface glycan receptors

Authors: [a]Daniel J. Honigfort, [b,†] Meghan O. Altman, [b]Pascal Gagneux, and Kamil Godula*[,a]

Author Affiliation:

[a]*Department of Chemistry and Biochemistry, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0358, USA*

[b]*Department of Pathology, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0687, USA*

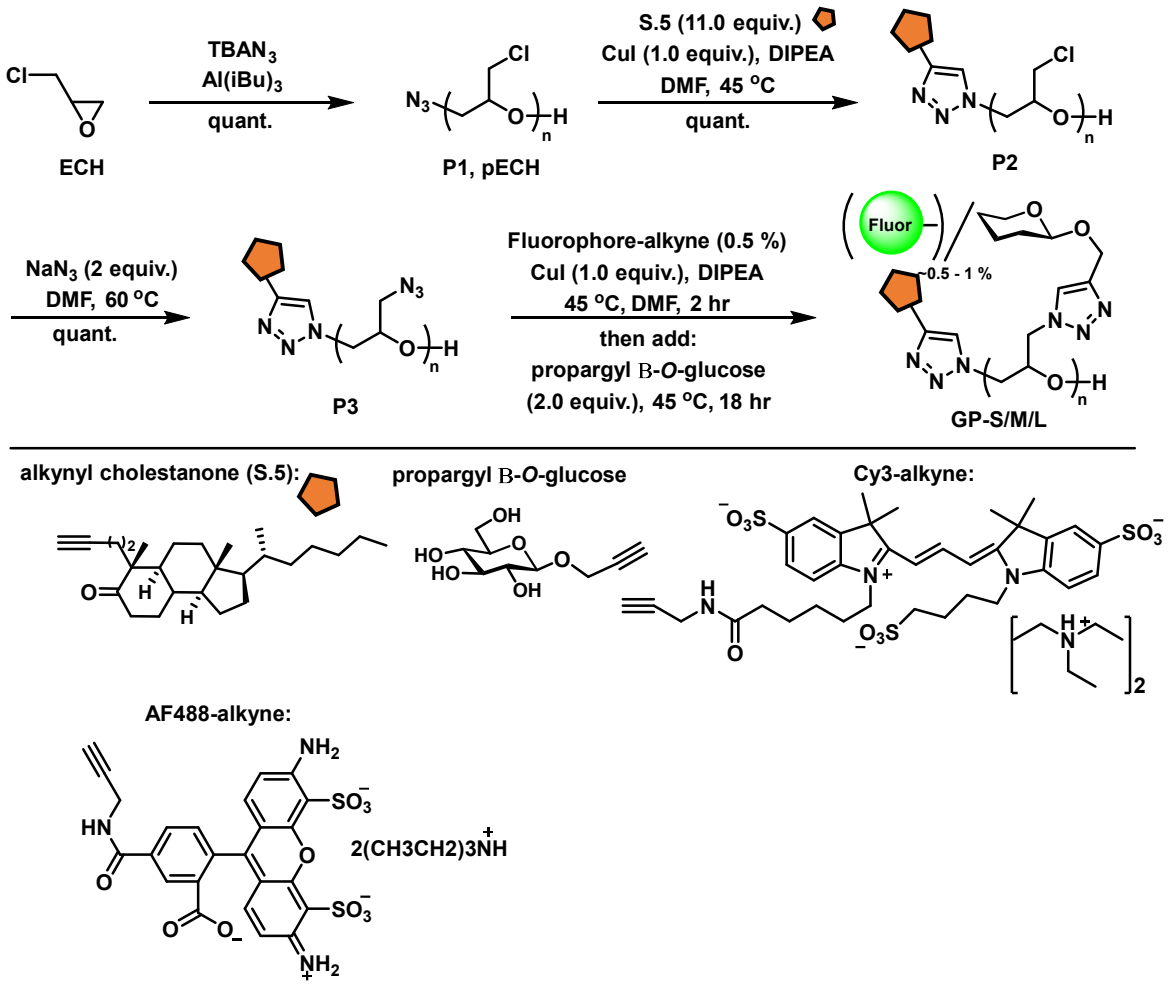
Instrumentation and reagents.

Instrumentation. Column chromatography was performed on a Biotage Isolera One automated flash chromatography system. Nuclear magnetic resonance (NMR) spectra were collected on a Bruker 300 MHz and a Jeol 500 MHz NMR spectrometers. Spectra were recorded in $CDCl_3$ or D_2O solutions at 293K and are reported in parts per million (ppm) on the δ scale relative to the residual solvent as an internal standard (for ^1H NMR: $CDCl_3$ = 7.26 ppm, D_2O = 4.79 ppm, for ^{13}C NMR: $CDCl_3$ = 77.0 ppm). HRMS (high-resolution mass spectrometry) analysis was performed on an Agilent 6230 ESI-TOFMS in positive ion mode. UV-Vis spectra for polymer fluorophore content quantification were collected using a quartz cuvette using a Thermo Scientific Nanodrop2000c spectrophotometer. UV-Vis spectra for kinetic measurement of 4MU-NANA fluorescence turn on was collected in 96 well plate format using a SpectraMax i3x (Molecular Devices). IR spectroscopy was performed on a Nicolet 6700 FT-IR spectrophotometer (Thermo Scientific). Size exclusion chromatography (SEC) was performed on a Hitachi Chromaster system equipped with an RI detector and two 5 μm, mixed bed, 7.8 mm I.D. x 30 cm TSKgel columns in series (Tosoh Bioscience). Organic soluble polymers were analyzed using an isocratic method with a flow rate of 0.7 mL/min in DMF (0.2% LiBr, 70 °C). For aqueous SEC, two 8 μm, mixed-M bed, 7.5 mm I.D. x 30 cm PL aquagel-OH columns in series (Agilent Technologies) were run in sequence using an isocratic method with a flow rate of 1.0mL/min in water (0.2M $NaNO_3$ in 0.01M Na_2HPO_4, pH = 7.0). Flow cytometry analysis was performed on live RBCs using a FACS Canto II cytometer (BD Biosciences). Microscopy techniques were performed on either a Keyence Fluorescent microscope (brightfield) or Leica SP5 (all fluorescence techniques).

Materials. All chemicals, unless stated otherwise, were purchased from Sigma Aldrich and used as received. Reaction progress was checked by analytical thin-layer chromatography

(TLC, Merck silica gel 60 F-254 plates) monitored either with UV illumination, or by staining with iodine, ninhydrin, or CAM stain. Solvent compositions are reported on a volume/volume (v/v) basis unless otherwise noted. 4,5-seco-cholesten-5-one[30] and Glc[31] propargyl glycosides were prepared according to published procedures. Turkey Red Blood Cells as a 10% solution were obtained from Lampire Biological Laboratories (cat # 724908). *Sambucus nigra* agglutinin (SNA) lectin were purchased from Vector Labs. NHS functionalized AlexaFluor 594 (AF 594) and AlexaFluor 647 (AF 647) for lectin labeling were purchased from Sigma Aldrich, and AlexaFluor 488 (AF 488) alkyne Cyanine 3 (Cy3)-alkyne for labeling of polymers was obtained from Sigma Aldrich. The sialidase reporter molecule 2'-(4-methylumbelliferyl)-α-D-N-acetylneuraminic acid (4MU-NANA) was purchased from Carbosynth. ß-propargyl glucoside (Glc) was prepared according to a previously published procedure.[32]

Synthesis of glycopolymers GP-S/M/L.



Scheme 3.S1. Synthesis of Glycopolymers through iterative CuAAc click strategy.

General Procedure for the preparation of poly(epichlorohydrin) (P1).

Epichlorohydrin was polymerized according to the procedure developed by Carlotti.[33] Briefly, for the longest glycopolymers **GP-L** as a representative example, a flame-dried Schlenk flask (10 mL) equipped with a magnetic stirrer and fitted with a PTFE stopcock was charged with tetrabutylammonium azide (TBAN3, 20 mg, 0.037 mmol, 0.002 equiv.) under argon. A solution of freshly distilled epichlorohydrin (1.29 mL, 16.5 mmol) was prepared in

anhydrous toluene (4 mL). A solution of triisobutylaluminum in toluene (1.07 M, 104 µL, 0.111 mmol, 0.007 equiv.) was added via a syringe under argon at −30 °C. The reaction was stirred for 4 hours and then stopped by the addition of ethanol. The resulting pECH polymer 1 was precipitated into hexanes and dried under vacuum to yield a clear viscous oil (1500 mg, 99% yield). The polymer was analyzed by SEC (0.2% LiBr in DMF), IR, and NMR.

General Procedure for the end-functionalization of poly(epichlorohydrin) (P2).

For the longest glycopolymers **GP-L** as a representative example, in a flame-dried Schlenk flask (10 mL), pECH polymer **P1** (15 mg, 0.3 µmol) was dissolved in degassed anhydrous DMSO (200 µL). 4,5-seco-cholesten-5-one **S.5** (1.7 mg, 3.8 µmol, 11.0 equiv.) was added, followed by CuI (~0.05 mg, 0.3 µmol, 1.0 equiv.) and one drop diisopropylethyl amine (DIPEA, ~5 µL). The reaction was stirred at 40 °C for 12 h, at which time it was quenched by the addition of water to precipitate the polymer. The resultant polymer was triturated with hexanes to remove unreacted **S.5** and dried on vacuum to yield a clear viscous oil in (16 mg, 100% yield). The polymer was analyzed by SEC (0.2% LiBr in DMF).

General Procedure for the preparation of poly(Glycidyl Azide) (P3).

The chloride to azide exchange in pECH polymer **P3** was accomplished according to a previously published procedure.[34] Briefly, in a flame-dried Schlenk flask (10 mL), polymer **P2** (15 mg, 0.16 mmol) was dissolved in dry DMF (300 µL). To the solution was added NaN_3 (21 mg, 0.32 mmol, 2.0 equiv.), and the reaction was stirred at 60 °C for 3 days under argon to allow complete conversion. The polymer solution was filtered and precipitated in ethanol to yield a clear viscous oil (16 mg, 100% yield). The polymer was analyzed by SEC (0.2% LiBr in DMF.

General procedure for the preparation of Glycopolymers (GP-S/M/L).

In a flame-dried Schlenk flask (10 mL), polymer **P3** (9.00 mg, 0.09 mmol) was dissolved in degassed dry DMSO (250 µL). To the solution was added AF488-alkyne or Cy3-alkyne (0.50 mg, 0.50 µmol) in DMSO (50 µL), followed by CuI (2.00 mg, 9.00 µmol) and DIPEA (16 µL, 0.09 mmol). The reaction was stirred in the dark under Ar at 40 °C for 2 h. After this time, ß-propargyl glucoside (0.02 mmol, 1.50 equiv. per azide side-chain) in degassed anhydrous DMSO (50 µL) was added to the reaction mixture. The reactions were stirred in the dark at 40 °C overnight. After this time, the reactions were diluted with DI water and treated with Cuprisorb beads (SeaChem labs) for 18 h to sequester copper. The resulting copper-free solutions were filtered through Celite to remove the resin and lyophilized. The dry residues were triturated 3× with methanol with monitoring by TLC to remove excess unreacted glycosides. The resulting AF488 or Cy3-labeled glycopolymers **GP-S/M/L** were dissolved in D2O and lyophilized to give a blue solid in a quantitative yield for each polymer. The polymers were characterized using 1H NMR (D2O, 300 MHz), IR, and UV-Vis (λmax = 488 or 554 nm) spectroscopy. Absorbance readings at known concentrations of glycopolymers **GP-S/M/L** indicated the presence of 0.3 – 2.7 fluorophores per polymer chain depending upon polymer length (~0.5 - 1% sidechain occupancy). The theoretical Mn of the final glycopolymers **GP** were calculated assuming 100% sidechain substitution with glucose.

Methods for biological assays.

Remodeling of RBC glycocalyx with glycopolymers GP.

RBCs (4% w/v in PBS) were incubated with AF488-labeled glycopolymers **GP-S/M/L** at increasing concentrations (c_{pol} = 0.1-30.0 μM) for 1 h at 37 °C. The cells washed 1x with PBS, then were probed for the presence of AF488 fluorescence using flow cytometry. The data was analyzed on Cytobank online software. Cells were gated to exclude debris, and the median fluorescence intensities (MFI) of cells are reported.

Sedimentation of glycocalyx-remodeled RBCs.

RBCs (50μL, 1% in PBS) treated with increasing concentrations of glycopolymers **GP-S/M/L** (c_{pol} = 0.1-30 μM) or left untreated were transferred to round-bottom 96 well plates. The cells were allowed to sediment over 45min. After this time, the plates were scanned on an EPSON Perfection V700 Photo scanner (Digital ICE technologies), and the lowest polymer concentrations required to induce RBC agglutination were determined.

Preparation of RBCs for microscopy.

RBCs (50uL, 1% in PBS) treated with increasing concentrations of glycopolymers **GP-S/M/L** (c_{pol} = 1-50 μM) or left untreated were diluted 50x and transferred to poly(lysine) slides. The cells were allowed to settle to the slide surface, and excess PBS was removed via pipette prior to application of a cover slip in order to prevent aggregation of cells. Bright field images of RBCs were taken on Keyence fluorescent microscope, and all fluorescence images were taken on a Leica SP5 confocal microscope.

Determination of membrane fluidity (FRAP).

FRAP experiments were performed on a Leica SP5 confocal microscope with a 40× water objective. The membranes of RBCs (50uL, 1% in PBS) treated with glycopolymers **GP-S/M/L** (c_{pol} = 7.5 μM) were bleached with a circular spot of diameter ~ 0.5 μm at 488nm wavelength. A single iteration was used for the bleach pulse, and fluorescence recovery was monitored at low laser intensity in 0.11 s intervals for 12 seconds. FRAP was performed on 6 separate cells and then averaged to generate a single FRAP curve. The rate was calculated by fitting the recovery curve to a hyperbola fit in Prism and solving for the rate in the integrated rate equation ($A = A_o\ e^{(-kt)}$).

Determination of SNA association and dissociation with glycocalyx-remodeled RBCs by flow cytometry.

In a 96 well round bottom plate, to RBCs (0.33% in PBS) treated with glycopolymers **GP** (c_{pol} =7.5 μM) or 4,5-seco-cholesten-5-one (c_{chol} =7.5 μM), or to untreated cells, were added AF647-labeled SNA lectins at sub-agglutination concentrations (c_{SNA} = 0.2 μg/mL). The cells were vortexed vigorously for ~ 10 s and then analyzed by flow cytometry (Canto II, BD Biosciences) for the presence of AF647 signal at discrete time points until saturation lectin binding was observed. The data were analyzed on Cytobank software. Cells were gated to exclude debris, and median fluorescence intensities (MFI) of cells are reported. Means and standard deviations were calculated from six independent biological experiments, and *p*-values corresponding to each condition vs. untreated RBC control were calculated using ANNOVA tests with PRISM software. The slopes designating the initial rates of lectin

association were extracted for each condition and their significance with respect to untreated RBC controls was assessed based on *p*-values calculated using 1-way ANNOVA tests.

Determination of co-localization between polymer and lectin via fluorescence microscopy.

RBCs remodeled with glycopolymer (c_{pol} =7.5 μM) were added to poly(lysine) slides. Once settled on the surface, AF 647 labeled SNA in PBS (c_{SNA} = 0.2 μg/mL) was incubated on the slide for 15 min. Unreacted lectin was washed away prior to imaging on Leica sp5 confocal microscope. Colocalization was quantified from the average of 12 replicates using Velocity software (Quorum Technologies).

Determination of polymer and lectin clustering by FLIM.

RBCs (0.33% in PBS) treated with glycopolymers **GP-L** labeled with AF488, Cy3, or an equimolar mixture of the two (c_{pol} =7.5 μM), or 4,5-seco-cholesten-5-one (c_{chol} =7.5 μM), or untreated cells, were added to poly(lysine) slides. Once settled on the surface, SNA lectins labeled with AF595, AF647, or an equimolar ratio of the two in PBS (c_{SNA} = 0.2 μg/mL) were applied to RBCs on the slide for 15 min. Unreacted lectin was washed away prior to imaging on Leica sp5 confocal microscope. FLIM images were analyzed directly on Leica acquisition software.

Maintenance of viral culture.

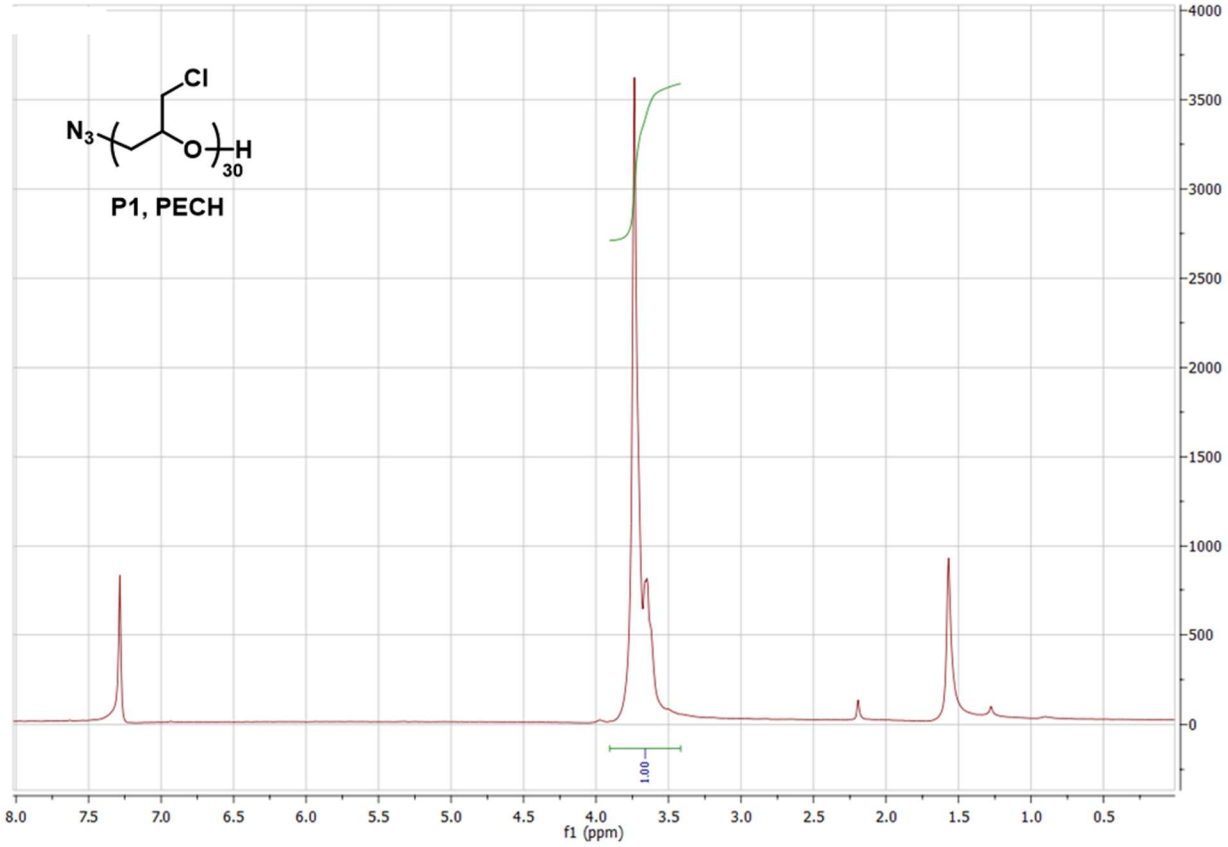
Influenza virus strain A/PR/8/34 (H1N1, ATCC VR-1469) was purchased from ATCC and propagated in MDCK cells that were transferred to DMEM medium supplemented with 0.2% BSA fraction V, 25mM HEPES buffer, 2 μg/ml TPCK-trypsin, and 1%

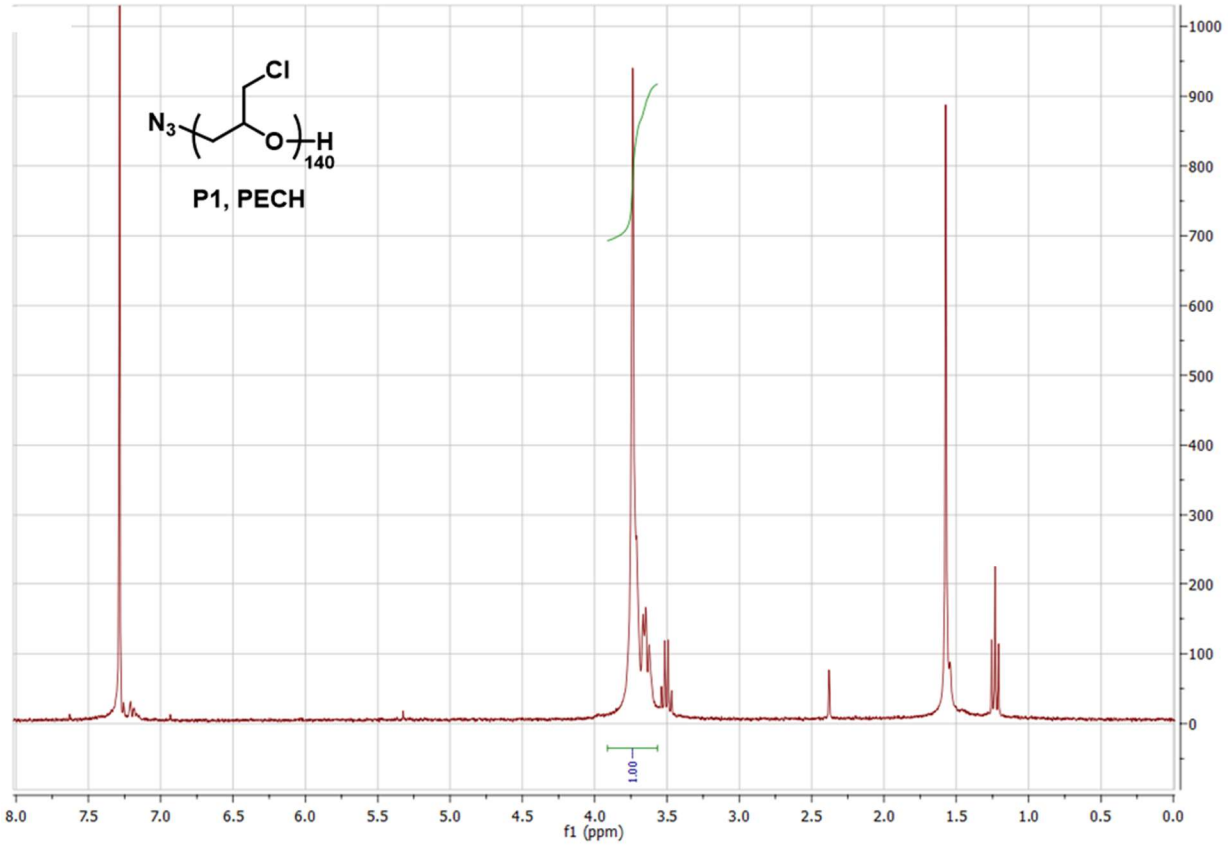
penicillin/streptomycin ("DMEM-TPCK" media). Viral titers were determined via the hemagglutination test (HAU) using a 1% solution of turkey red blood cells purchased from Lampire.

Determination of viral binding to remodeled RBCs (4MU-NANA).

Viral concentration was measured using 4-Methylumbelliferyl N-acetyl-a-D-neuraminic acid (4MU-NANA) fluorescence turn-on reporter for NA activity. Virus (30 HAU) was incubated with 1% RBCs in PBS remodeled with glycopolymers **GP-S/M/L** (c_{pol} = 7.5 µM), or unmodified cells for 15 min. RBCs were centrifuged and unbound virus was removed via pipette. To measure association, cells were then resuspended in 1.2 mM 4MU-NANA and cleavage was monitored by UV/Vis (excitation = 380nm, emission = 450nm) every 60 seconds for one hour at 37 °C. To measure retention, cells were instead resuspended in fresh buffer and allowed to equilibrate for 1hr before they were centrifuged, shed virus was removed by pipette, and RBCs were resuspended in 1.2 mM 4MU-NANA and monitored for cleavage as above.

^{1}H NMR spectra of GP-S/M/L and synthetic intermediates.




Cl
N3
O
H
140
P1, PECH
1000
900
800
700
600
500
400
300
200
100
0
1.00
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

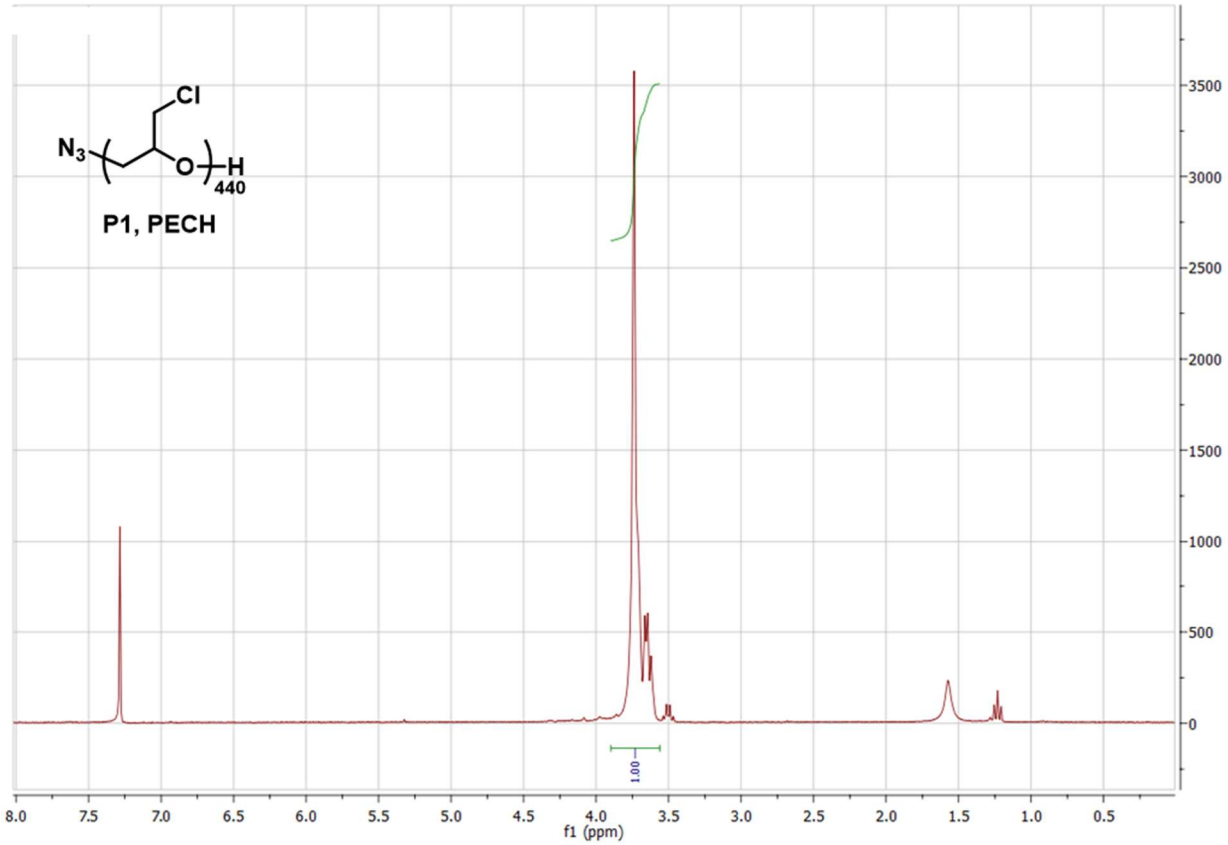
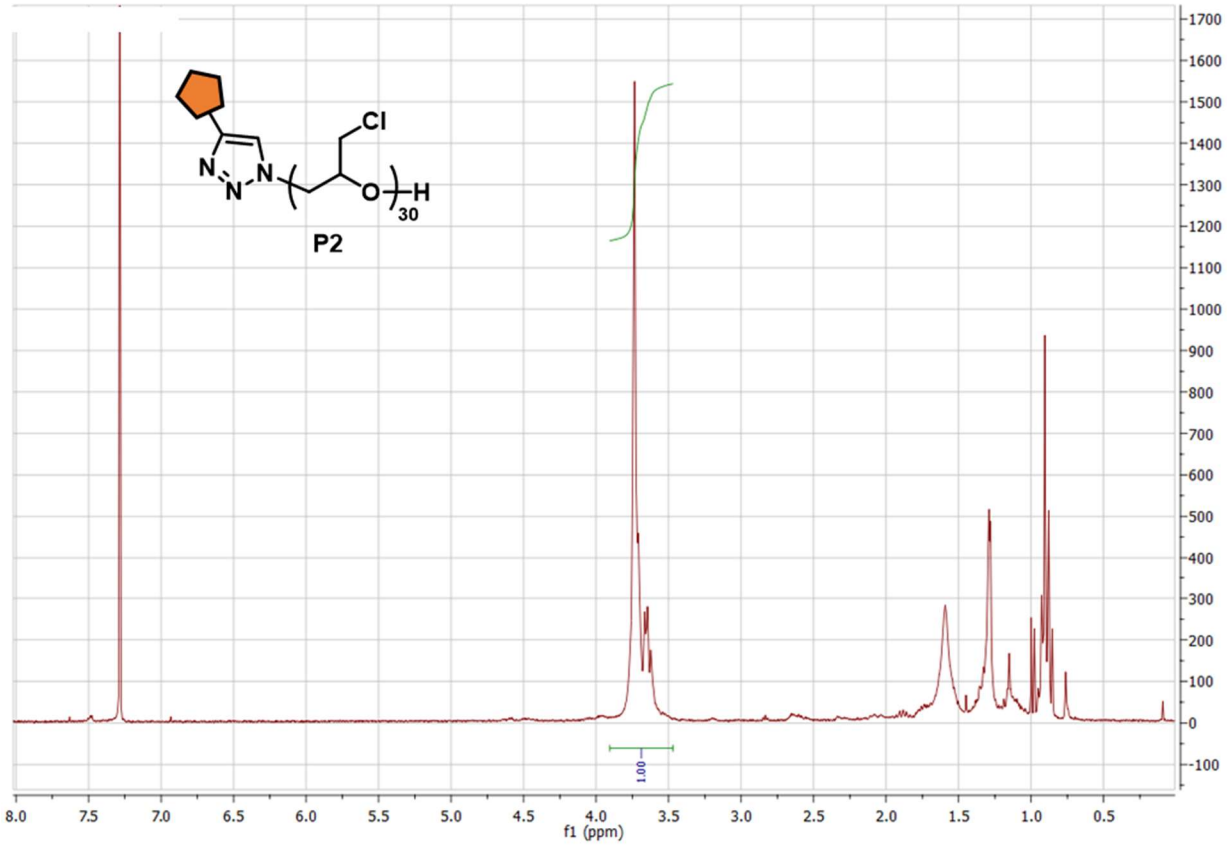
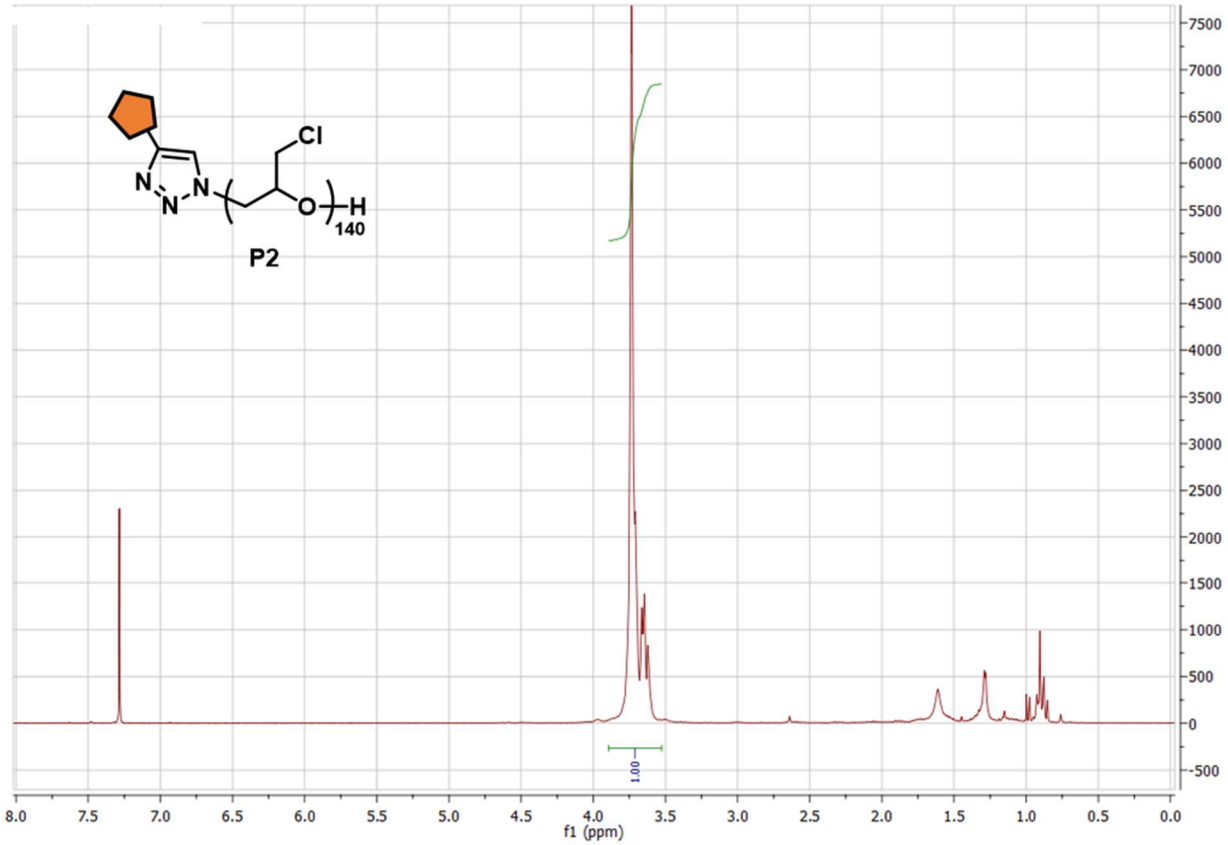


Figure 3.S4. ^{1}H NMR (300 MHz, $CDCl_3$) of pECH polymers **P1-S/M/L**.


Cl
N
N
N
O
H
30
P2
1.00
f1 (ppm)


Cl
N
N
N
O
H
140
P2
7500
7000
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500
1.00
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

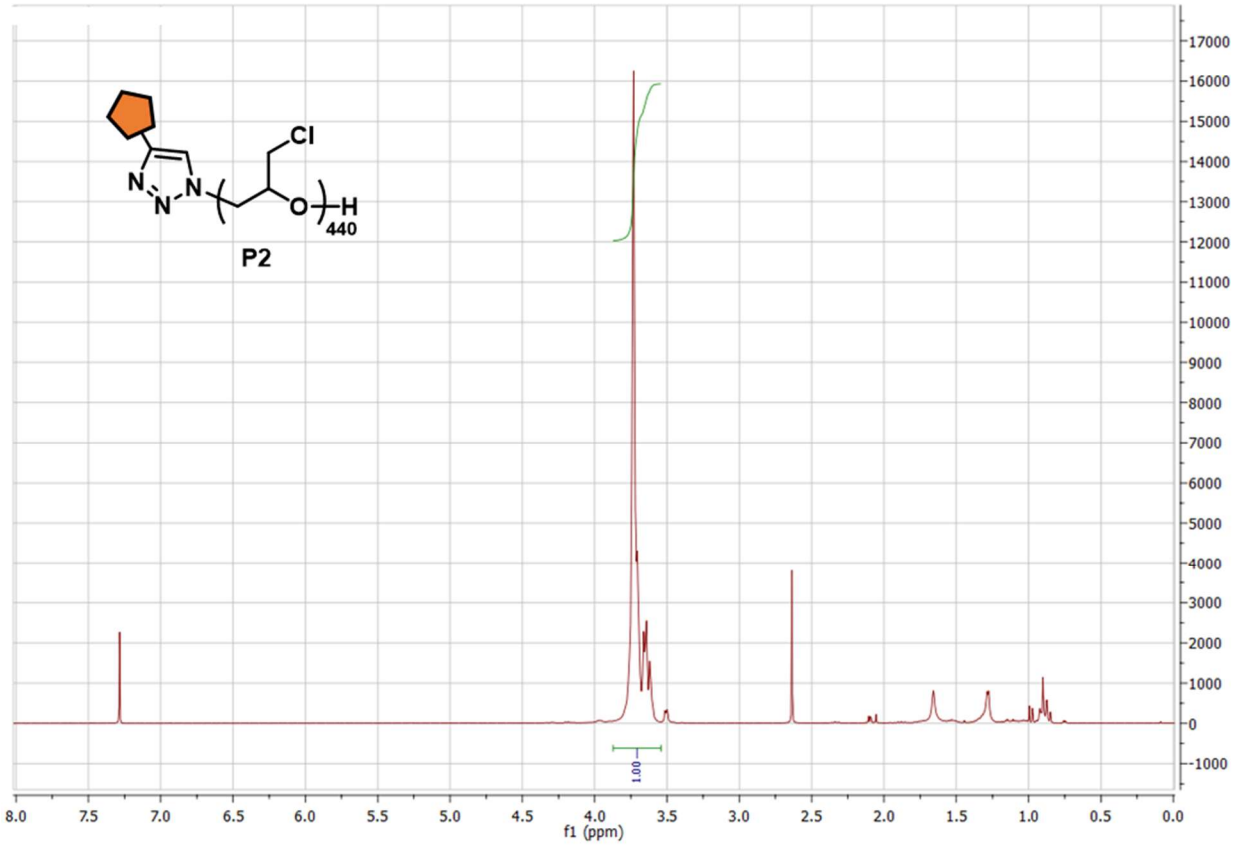
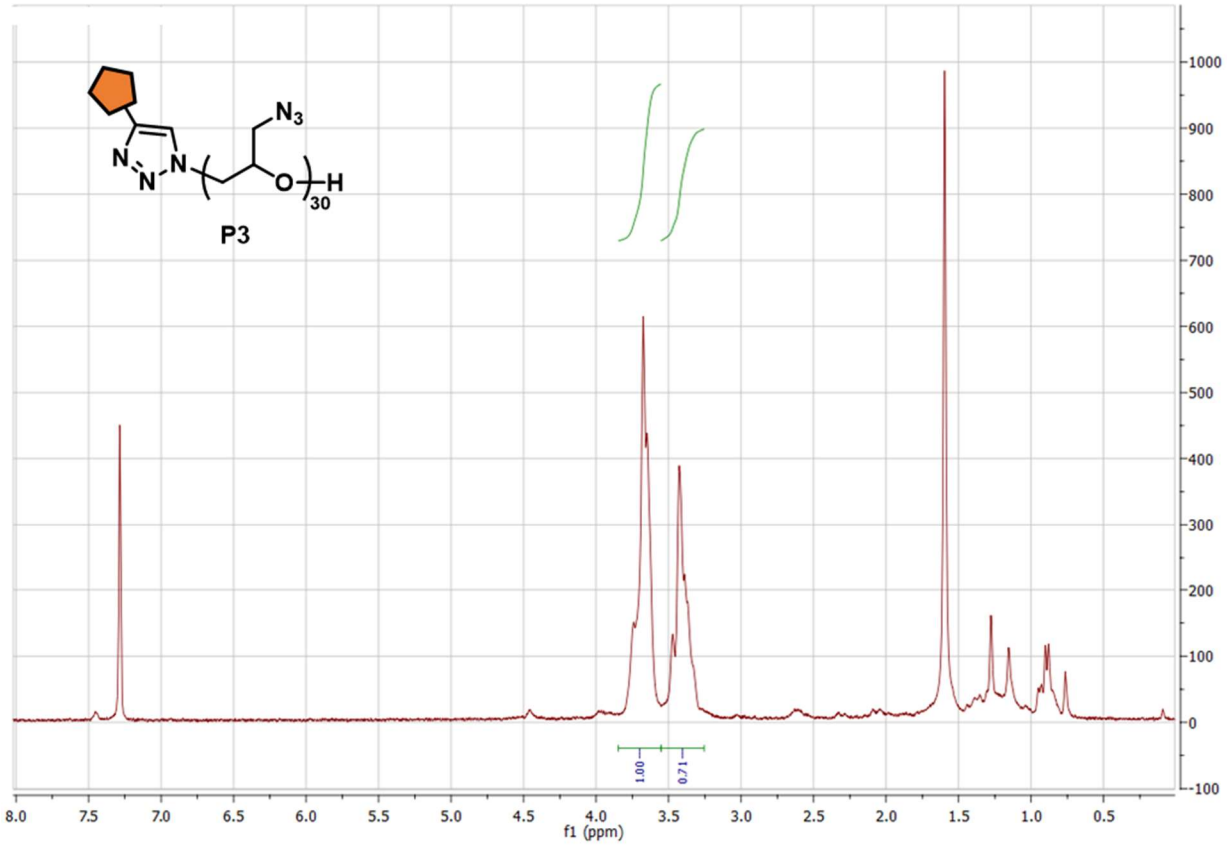
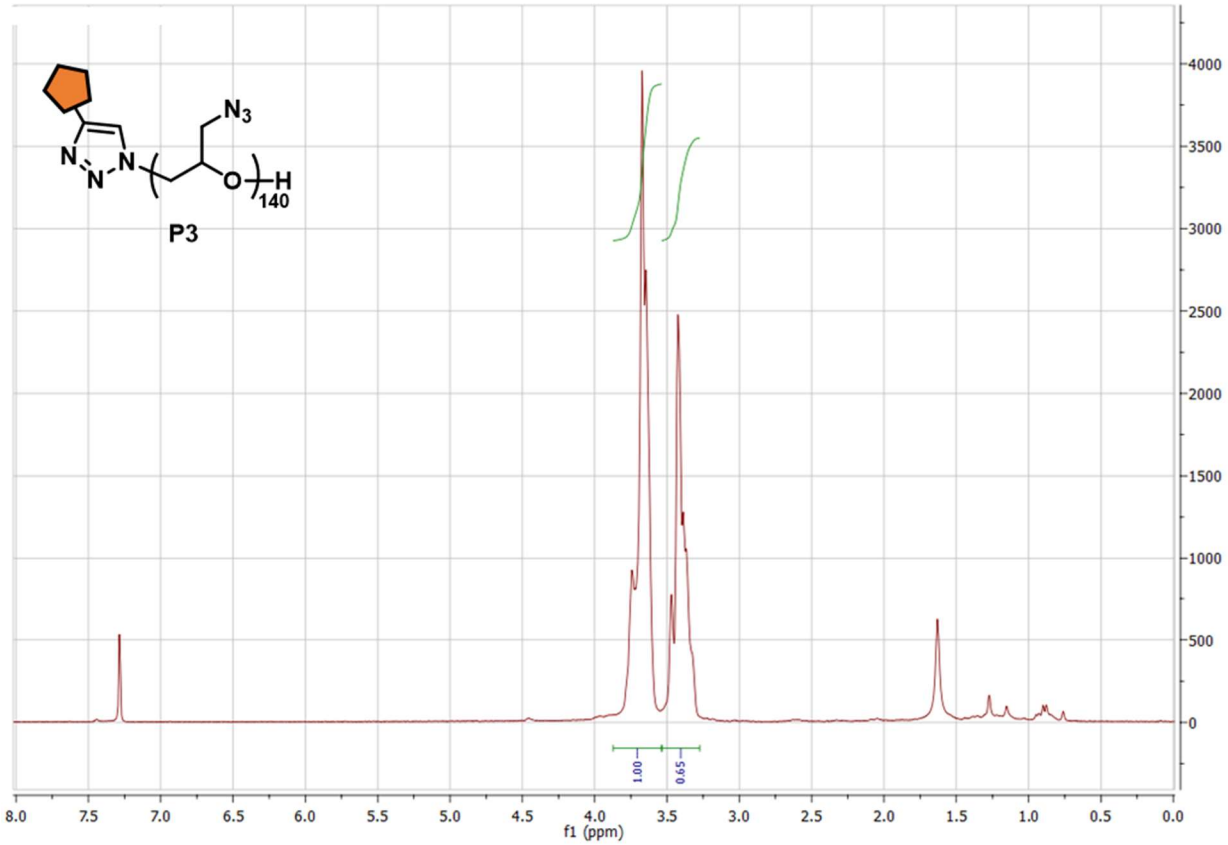


Figure 3.S5. ^{1}H NMR (300 MHz, $CDCl_3$) of cholestanone-terminated pECH polymers **P2-S/M/L**.


N3
N
N
N
O
H
30
P3
1.00
0.71
f1 (ppm)


N3
N
N
N
O
H
140
P3
4000
3500
3000
2500
2000
1500
1000
500
0
1.00
0.65
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

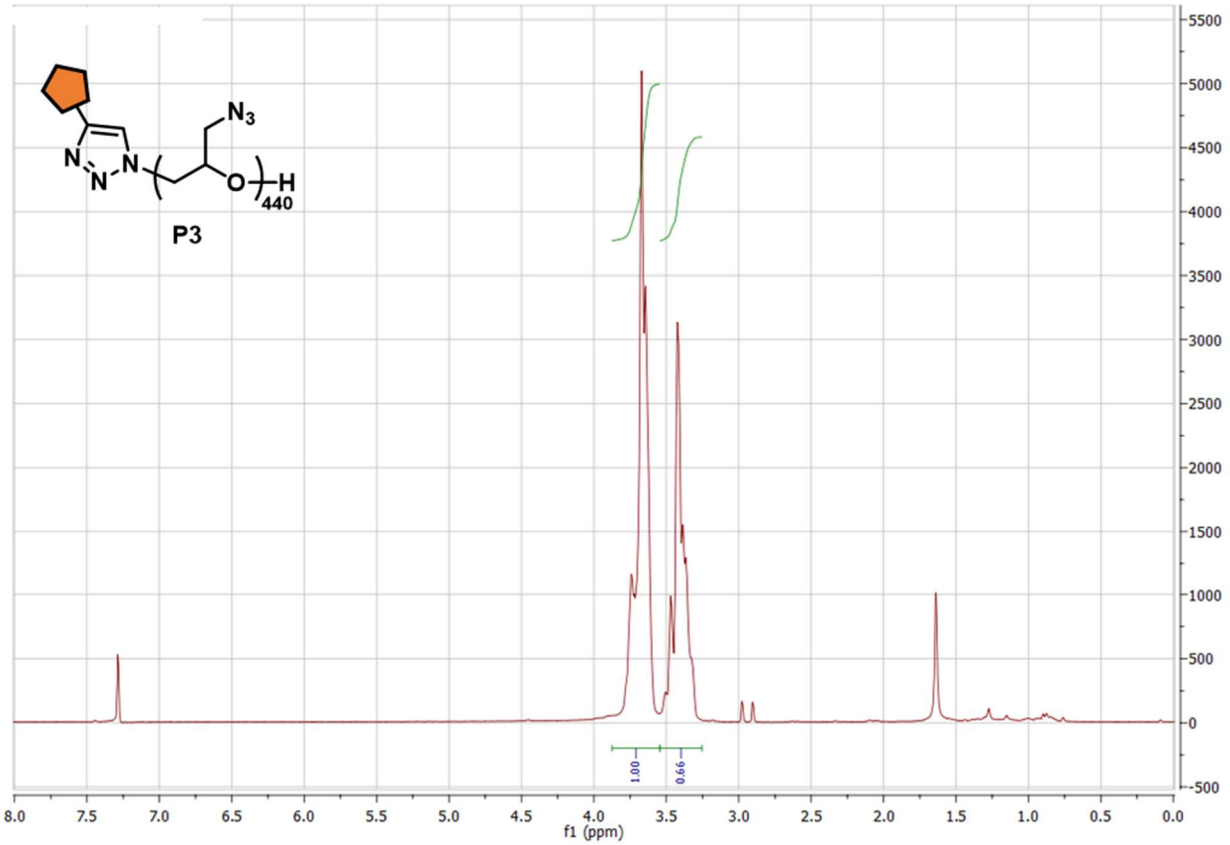
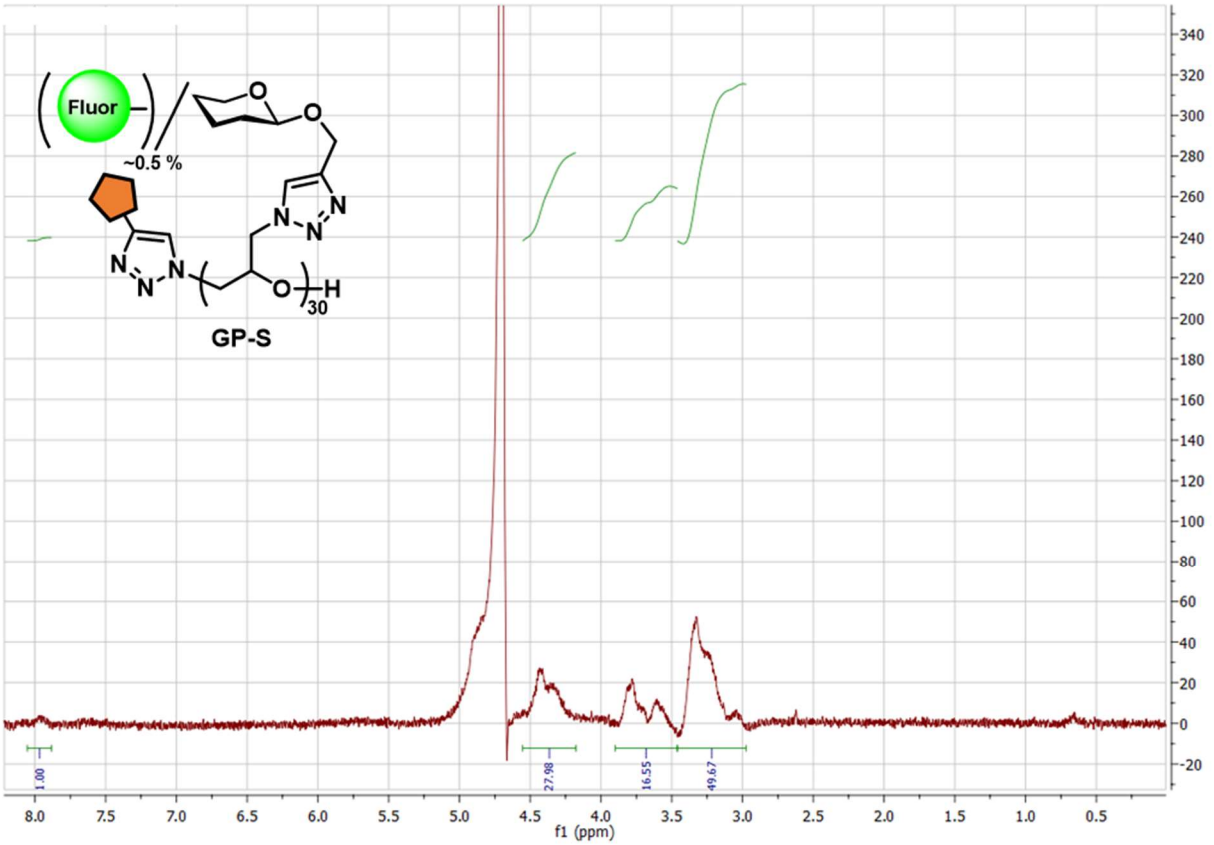
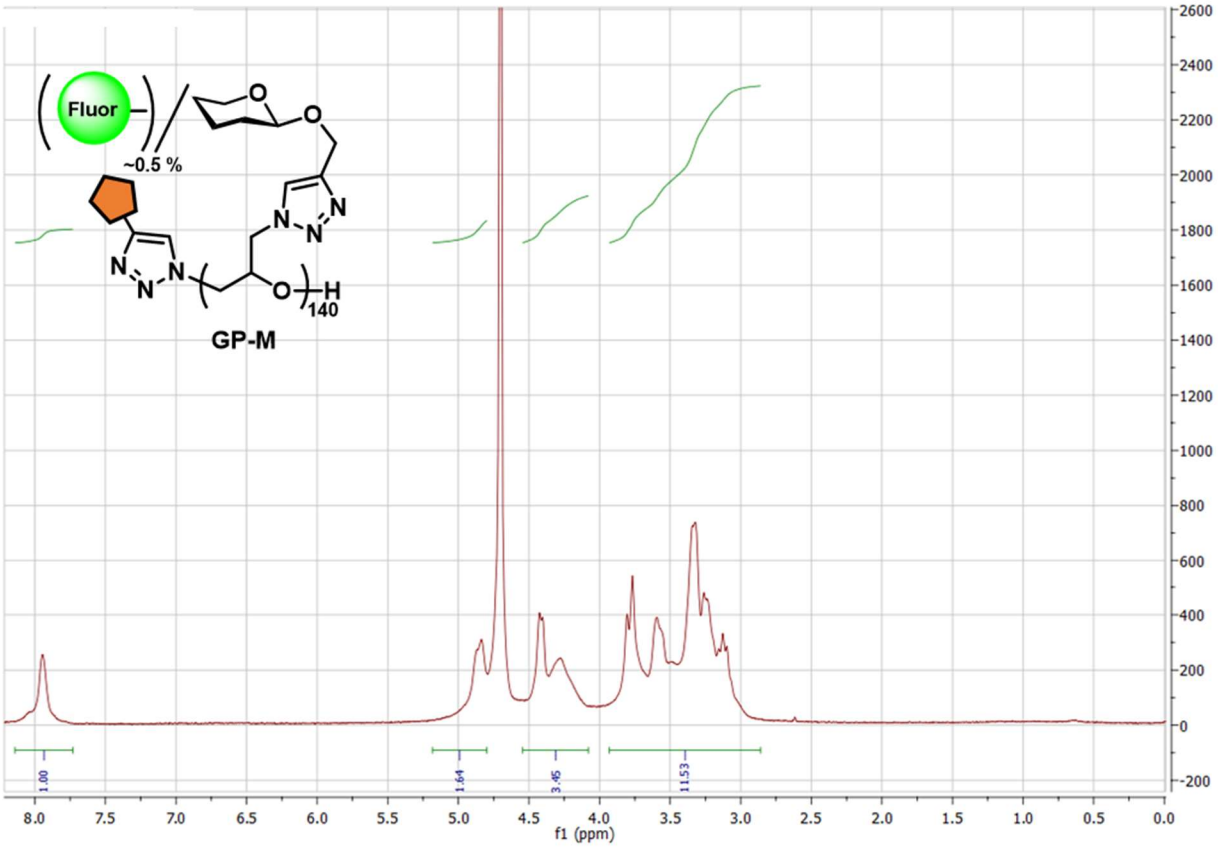


Figure 3.S6. ^{1}H NMR (300 MHz, $CDCl_3$) of cholestanone-terminated pGA polymers **P3-S/M/L**.


Fluor
~0.5 %
O
O
N
N
N
N
N
O
H
30
GP-S
340
320
300
280
260
240
220
200
180
160
140
120
100
80
60
40
20
0
-20
1.00
27.98
16.55
49.67
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)


Fluor
~0.5 %
GP-M
140
1.00
1.64
3.45
11.53
f1 (ppm)

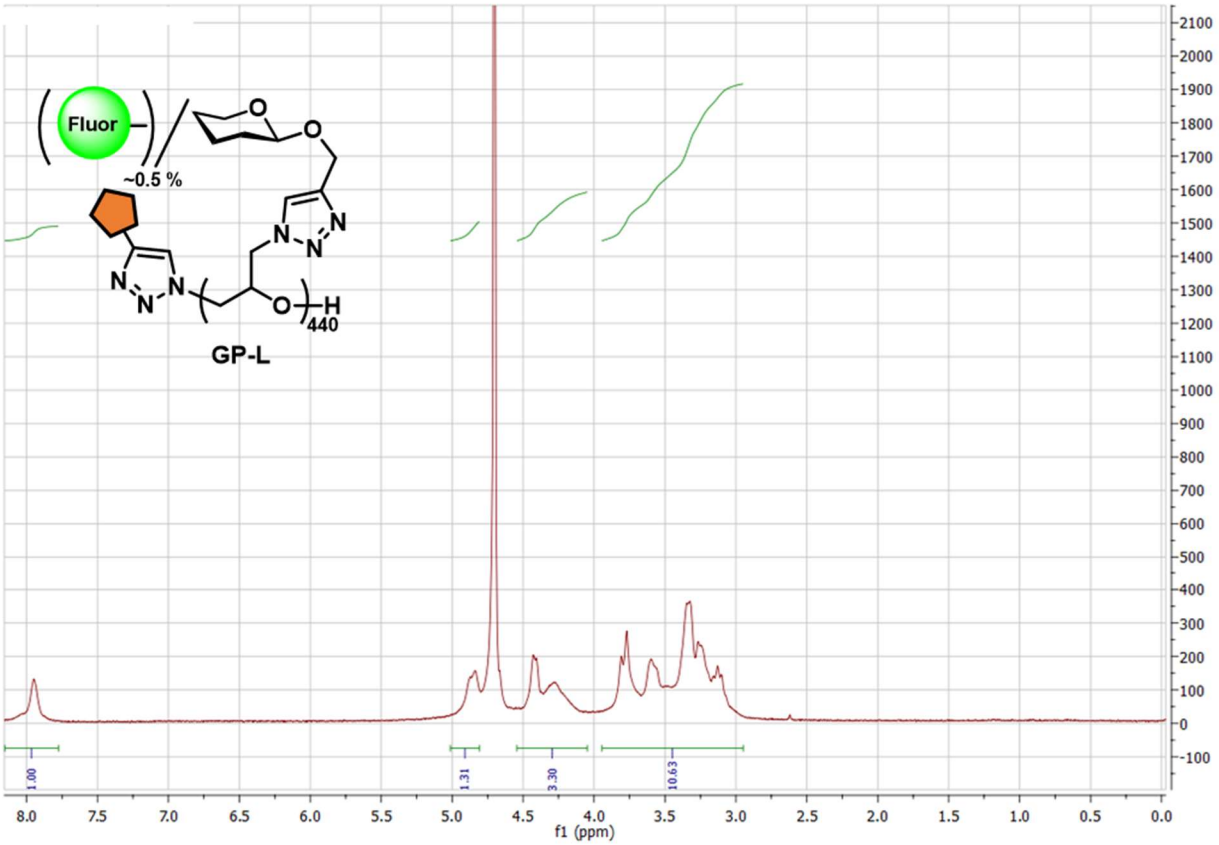


Figure 3.S7. ^{1}H NMR (300 MHz, D_2O) of glycopolymers **GP-S/M/L**.

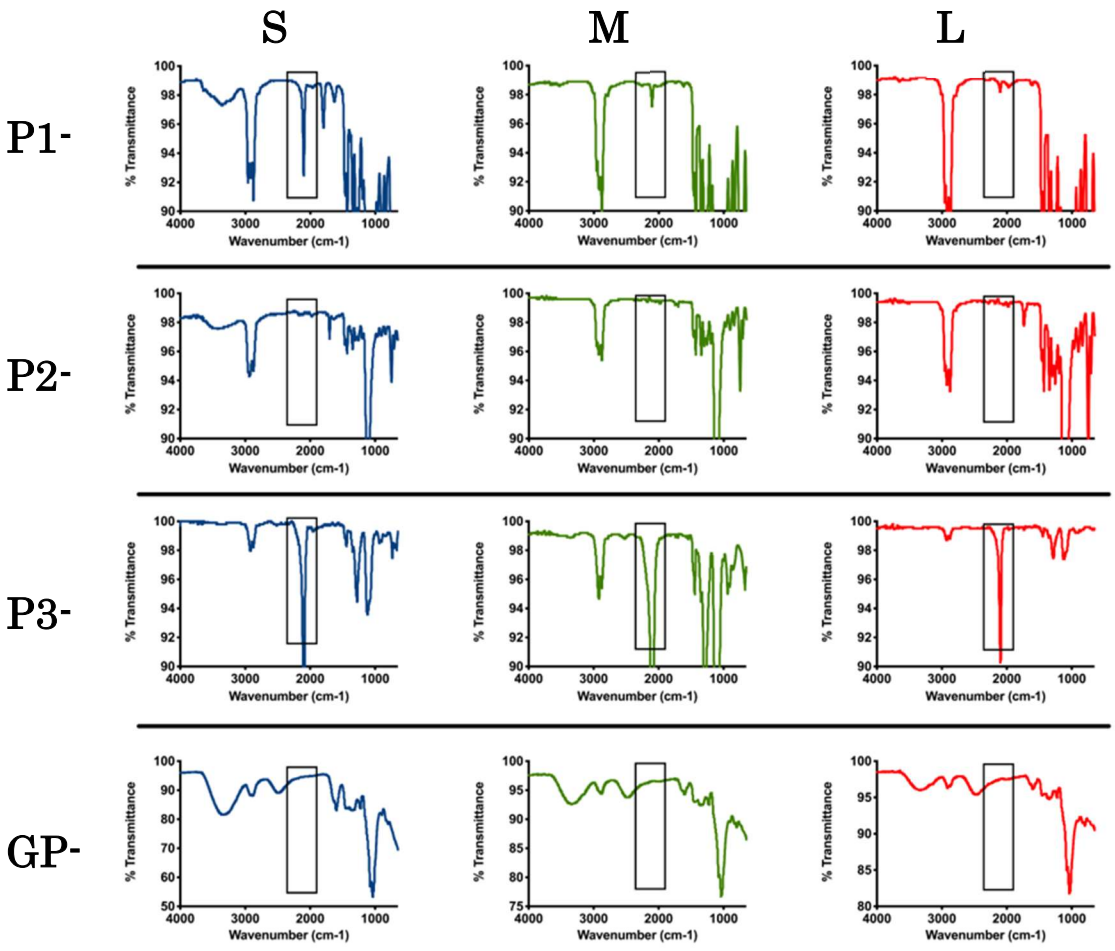
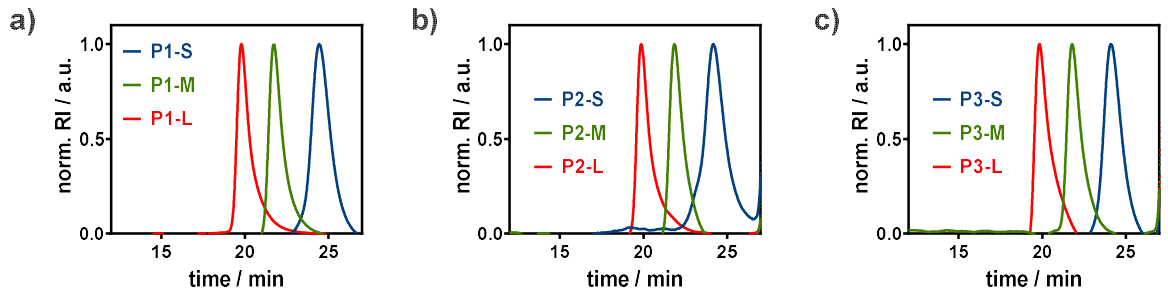


Figure 3.S8. IR spectra of polymers **P1**, **P2**, **P3**, and **GP-S/M/L**.



	Polymerization conditions				p(ECH) P1		P2		P3	
length	[*i*-Bu$_3$Al]/[NBu$_4$N$_3$]	[ECH] (n/L)	Mn$_{th}$ (kDa)	DP	Mn (kDa)	Đ	Mn (kDa)	Đ	Mn (kDa)	Đ
S	1.5	2	3.0	30	3.1	1.32	3.5	1.3	5.8	1.23
M	2	2	10.0	140	12.7	1.19	13.7	1.27	19.9	1.11
L	3	2	45.0	440	40.5	1.23	45.3	1.21	51.1	1.22

Figure 3.S9. GPC traces for polymer intermediates (expanded data for **Fig 3.2A**).

Table 3.S10. Expanded polymer characterization table for final Glycopolymers **GP-S/M/L.**

Polymer (P)	DP	Mn (kDa)	Fl	Fl/P	Fl/DP	Glucose (%)
GP-S	30	9.5	AF488	0.3	1	100
GP-M	140	44.4	AF488	0.86	0.61	100
GP-L	440	139.5	AF488	2.7	0.61	100
GP-L/D	440	139.5	AF488	2.4	0.55	100
GPL-L/A	440	139.5	Cy3	2.7	0.61	100

Mn was calculated assuming 100% sidechain substitution, and using the theoretical DP attained from the parent pECH polymer **P1** (Mn = (Mw glycosylated side chain) x DP). Fluorophore labeling was quantified using the maximum absorbance values from UV/Vis. Glucose attachment was estimated to be quantitative because of the lack of an azide peak at 2100 cm^{-1} in IR (**Fig 3.S8**).

RBC remodeling with glycopolymers GP-S/M/L.

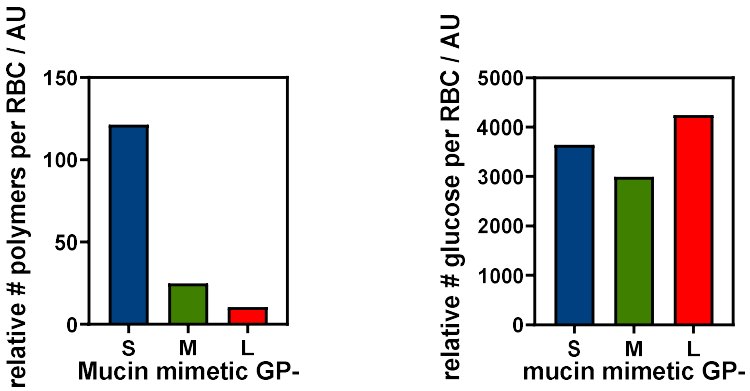


Figure 3.S11. Relative levels of cell surface incorporation of glycopolymers **GP-S/M/L** at 7.5 μM.

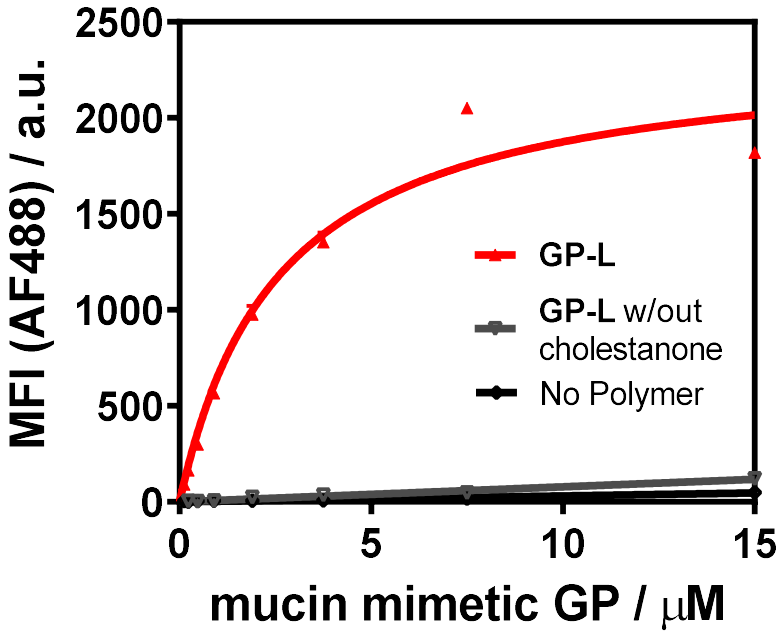


Figure 3.S12. Relative incorporation of **GP-L** polymers vs equivalent polymer without cholestanone

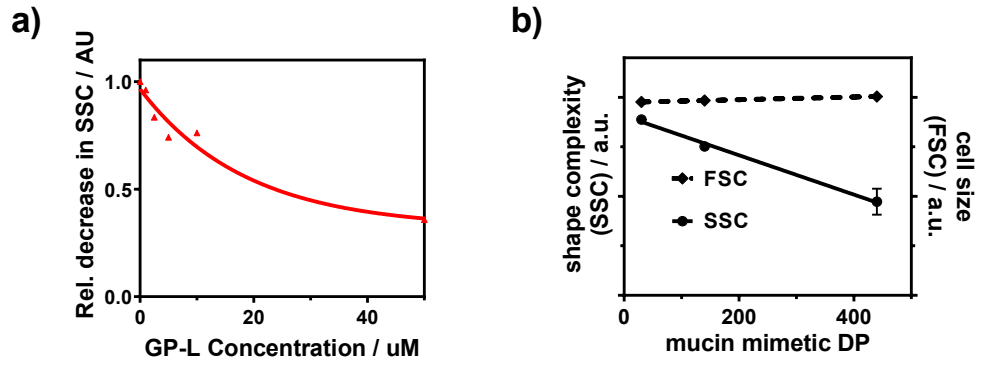


Figure 3.S13. FSC and SSC of RBCs remodeled with glycopolymers **GP-S/M/L.**

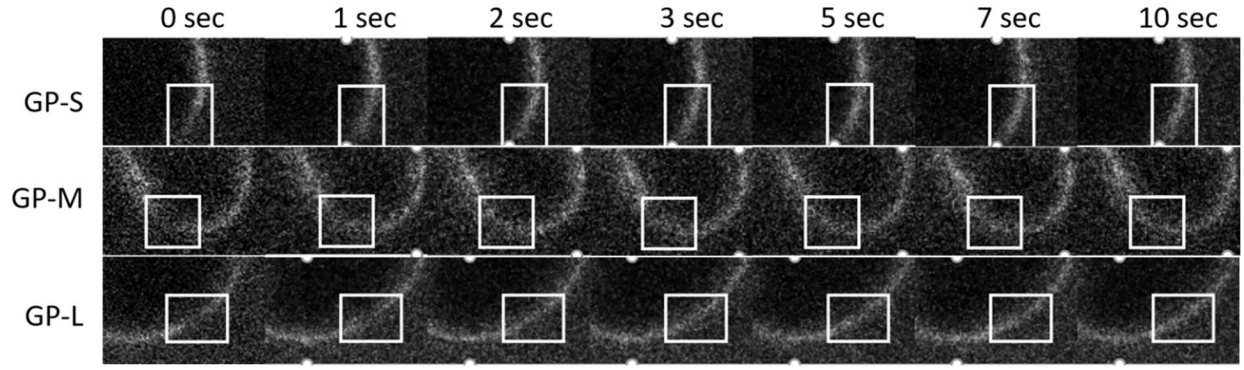


Figure 3.S14. FRAP images of RBCs remodeled with **GP-S/M/L** (associated with **Fig 3.2F**).

Characterization of SNA binding to RBCs remodeled with glycopolymers GP-S/M/L.

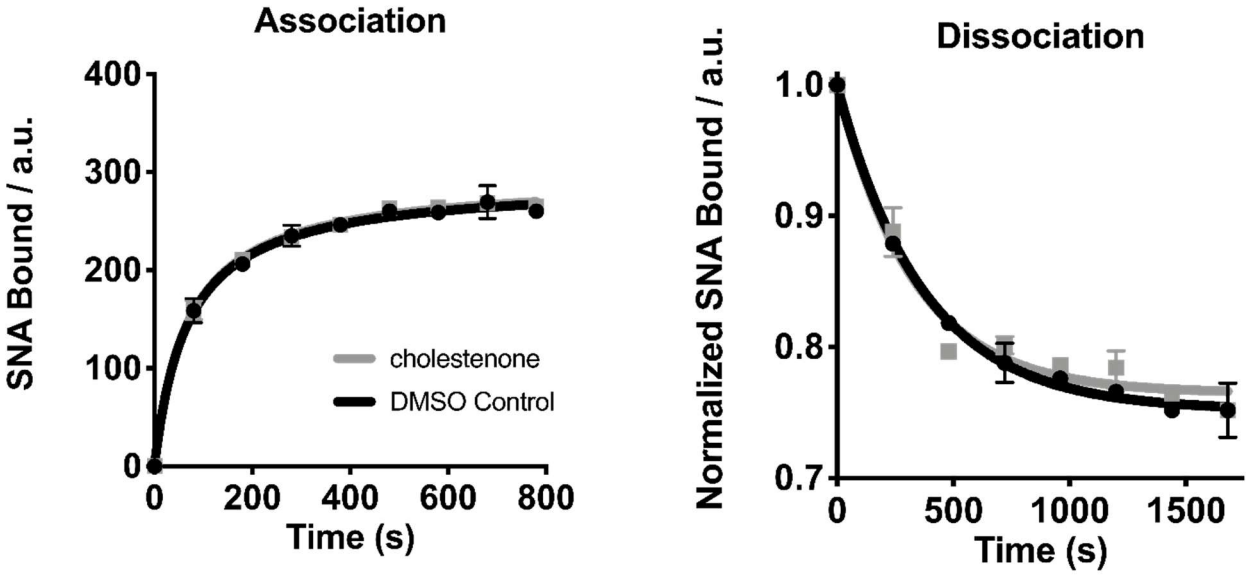
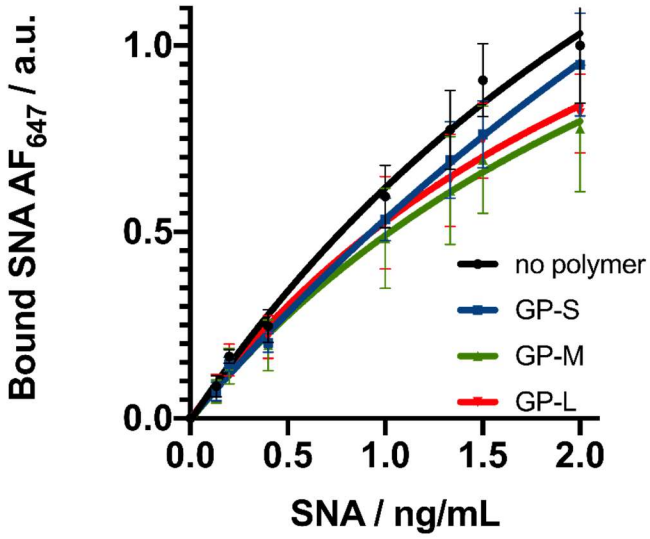


Figure 3.S15. Binding of SNA to RBCs after pre-treatment with alkynyl cholestanone **S5.**



SNA ng/mL	no polymer vs. GP-S		no polymer vs. GP-M		no polymer vs. GP-L	
	significance	p-value	significance	p-value	significance	p-value
0.0	ns	0.9999	ns	0.9999	ns	>0.9999
0.1	ns	0.9774	ns	0.9744	ns	0.9995
0.2	ns	0.9181	ns	0.8719	ns	0.992
0.4	ns	0.5649	ns	0.4702	ns	0.822
1.0	ns	0.3175	*	0.0207	ns	0.2222
1.3	ns	0.1364	***	0.0003	**	0.0037
1.5	**	0.0018	****	<0.0001	***	0.0004
2.0	ns	0.4641	****	<0.0001	****	<0.0001

Figure 3.S16. SNA binding to remodeled RBCs as a function of lectin concentration.

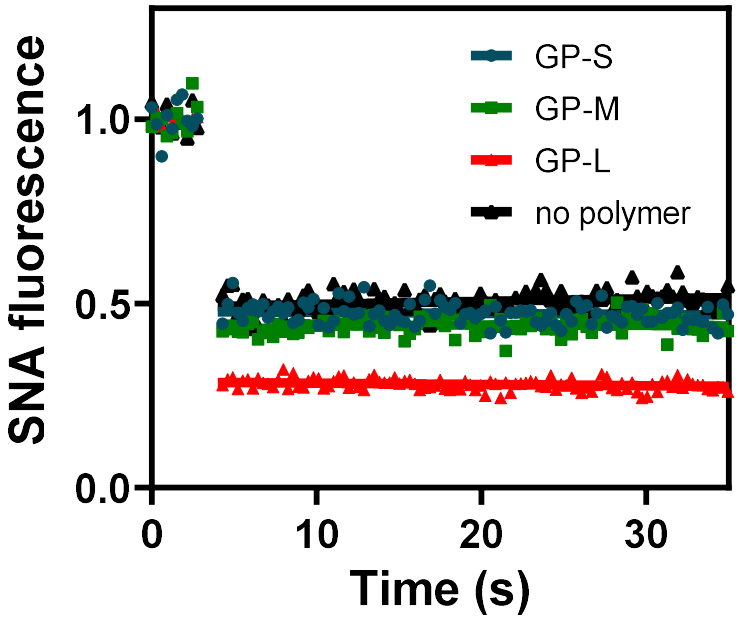


Figure 3.S17. Mobility of SNA bound to RBC membrane (FRAP).

H1N1 binding to RBCs remodeled with glycopolymers GP-S/M/L via enzymatic 4MU-NANA assay.

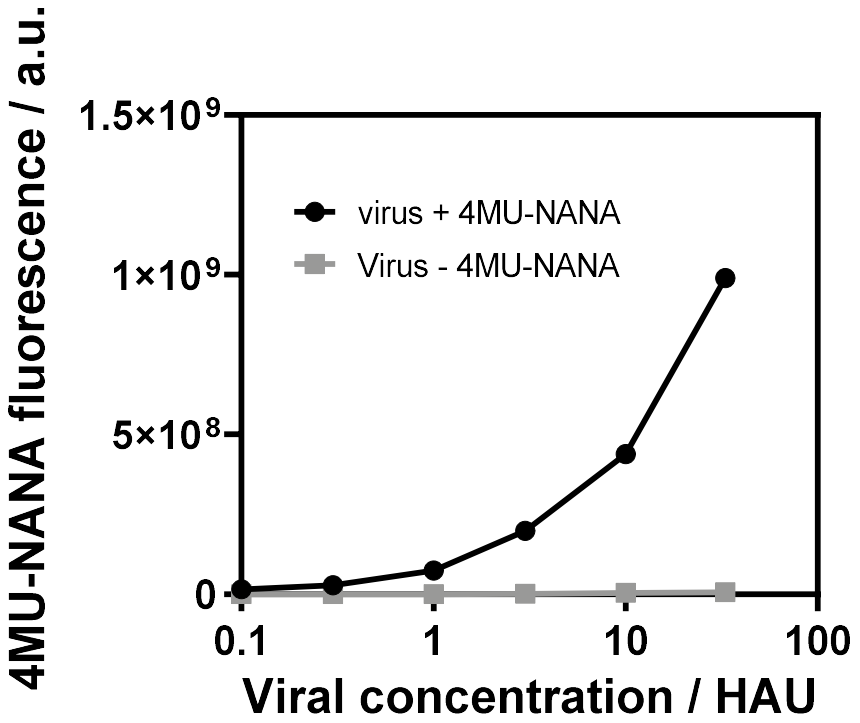


Figure 3.S18. 4MU-NANA fluorescence turn on with increasing viral titer.

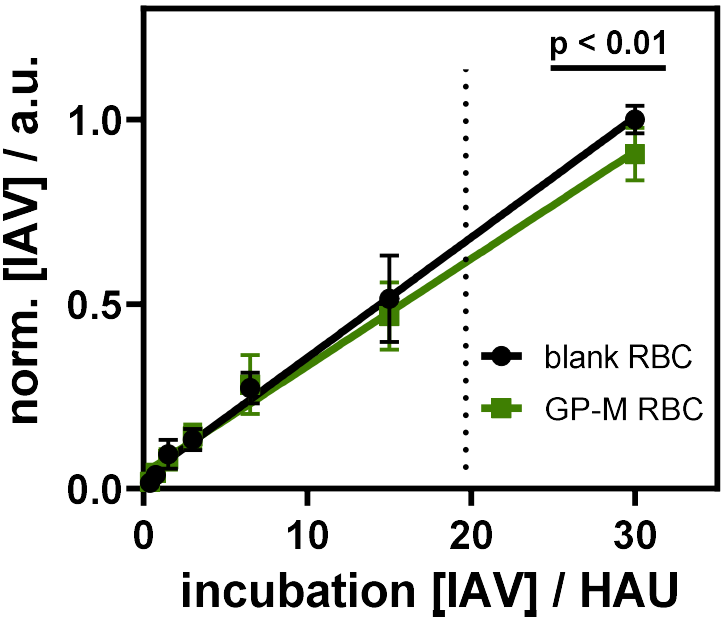


Figure 3.S19. Viral titer dependance on binding to RBCs.

Section 3.7 References

1. Szymanski, C. M., Schnaar, R. L., Aebi, M. *Bacterial and Viral Infections*. In: Varki A, Cummings RD, Esko JD, editors. Essentials of Glycobiology 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, Chapter 42 2015-2017.

2. Varki, A. & Gagneux, P. Multifarious roles of sialic acids in immunity. *Ann. N. Y. Acad. Sci.* **1253**, 16–36 (2012).

3. Cummings, R. D., Schnaar, R. L., Esko, J. D., Drickamer, K. & Taylor, M. E. *Principles of Glycan Recognition.* In: Varki A, Cummings RD, Esko JD, editors. Essentials of Glycobiology 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, Chapter 29, 2015-2017.

4. Cohen, M., Varki, A. The Sialome – Far More Than the Sum of Its Parts. *OMICS: A Journal of Integrative Biology*. **14**, 455-464 (2010).

5. Matrosovich, M., Herrler, G., Klenk, H. D. Sialic Acid Receptors of Viruses. *Top. Curr. Chem.* **367**, 1-28 (2015).

6. Corfield, A. P. Mucins: A biologically relevant glycan barrier in mucosal protection. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1850**, 236–252 (2015).

7. Wagner, C. E., Wheeler, K. M. & Ribbeck, K. Mucins and Their Role in Shaping the Functions of Mucus Barriers. *Annu. Rev. Cell Dev. Biol.* **34**, 189–215 (2018).

8. McGuckin, M. A., Lindén, S. K., Sutton, P., Florin, T. H. Mucin dynamics and enteric pathogens. *Nat. Rev. Microbiol.* **9**, 265–278 (2011).

9. Corfield, A. P. Mucins: A biologically relevant glycan barrier in mucosal protection. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1850**, 236–252 (2015).

10. Lindén, S. K. Sheng, Y.H., Every, A.L., Miles, K.M., Skoog, E.C., Florin, T.H.J., Sutton, P., McGuckin, M.A. MUC1 Limits Helicobacter pylori Infection both by Steric Hindrance and by Acting as a Releasable Decoy. *PLOS Pathog.* **5**, e1000617 (2009).

11. McAuley, J. L. Corcilius, L., Tan, H.-X., Payne, R.J., McGuckin, M.A., Brown, L.E. The cell surface mucin MUC1 limits the severity of influenza A virus infection. *Mucosal Immunol.* **10**, 1581–1593 (2017).

12. Rabuka, D. Parthasarathy, R., Lee, G.S., Chen, X., Groves, J.T., Bertozzi, C.R. Hierarchical assembly of model cell surfaces: synthesis of mucin mimetic polymers and their display on supported bilayers. *J. Am. Chem. Soc.* **129**, 5462–5471 (2007).

13. Purcell, S. C. & Godula, K. Synthetic glycoscapes: addressing the structural and functional complexity of the glycocalyx. *Interface Focus* **9**, 20180080 (2019).

14. Hudak, J. E., Canham, S. M. & Bertozzi, C. R. Glycocalyx engineering reveals a Siglec-based mechanism for NK cell immunoevasion. *Nat. Chem. Biol.* **10**, 69–75 (2014).

15. Rabuka, D., Forstner, M. B., Groves, J. T. & Bertozzi, C. R. Noncovalent Cell Surface Engineering: Incorporation of Bioactive Synthetic Glycopolymers into Cellular Membranes. *J. Am. Chem. Soc.* **130**, 5947–5953 (2008).

16. Paszek, M. J. DuFort, C.C., Rossier, O., Bainer, R., Mouw, J.K., Godula, K., Hudak, J.E., Lakins, J.N., Wijekoon, A.C., Cassereau, L., Rubashkin, M.G., Magbanua, M.J., Thorn, K.S., Davidson, M.W., Rugo, H.S., Park, J.W., Hammer, D.A., Giannone, G., Bertozzi, C.R., Weaver, V.M. The cancer glycocalyx mechanically primes integrin-mediated growth and survival. *Nature* **511**, 319–325 (2014).

17. Honigfort, D., Zhang, M., Verespy III, S., and Godula, K. Engineering of spectator glycocalyx structures to evaluate molecular interactions at crowded cellular boundaries. *Faraday Discuss.*, **219**, 138-153 (2019).

18. Delaveris, C. S., Webster, E. R., Banik, S. M., Boxer, S. G. & Bertozzi, C. R. Membrane-tethered mucin-like polypeptides sterically inhibit binding and slow fusion kinetics of influenza A virus. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 12643–12650 (2020).

19. Valeri, M. Paccani, S.R., Kasendra, M., Nesta, B., Serino, L., Pizza, M., Soriani, M. Pathogenic E. coli Exploits SslE Mucinase Activity to Translocate through the Mucosal Barrier and Get Access to Host Cells. *PLOS ONE* **10**, e0117486 (2015).

20. Wiggins, R., Hicks, S., Soothill, P., Millar, M. & Corfield, A. Mucinases and sialidases: their role in the pathogenesis of sexually transmitted infections in the female genital tract. *Sex. Transm. Infect.* **77**, 402–408 (2001).

21. Cohen, M. Zhang, X.-Q., Senaati, H.P., Chen, H.-W., Varki, N.M., Schooley, R.T., Gagneux, P. Influenza A penetrates host mucus by cleaving sialic acids with neuraminidase. *Virol. J.* **10**, 321 (2013).

22. Gervais, M., Labbé, A., Carlotti, S. & Deffieux, A. Direct Synthesis of α-Azido,ω-hydroxypolyethers by Monomer-Activated Anionic Polymerization. *Macromolecules*, 2009, **42**, 2395–2400.

23. 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**, 3057–3064.

24. Weinbaum, S., Zhang, X., Han, Y., Vink, H. & Cowin, S. C. Mechanotransduction and flow across the endothelial glycocalyx. *Proc. Natl. Acad. Sci.* **100**, 7988–7995 (2003).

25. Rad, S., Meiselman, H. J. & Neu, B. Impact of glycocalyx structure on red cell–red cell affinity in polymer suspensions. *Colloids Surf. B Biointerfaces* **123**, 106–113 (2014).

26. Shurer CR, Kuo JC, Roberts LM, Gandhi JG, Colville MJ, Enoki TA, Pan H, Su J, Noble JM, Hollander MJ, O'Donnell JP, Yin R, Pedram K, Möckl L, Kourkoutis LF, Moerner WE, Bertozzi CR, Feigenson GW, Reesink HL, Paszek MJ. Physical Principles of Membrane Shape Regulation by the Glycocalyx. *Cell* **177**, 1757-1770.e21 (2019).

27. Bradley, A. J., Murad, K. L., Regan, K. L. & Scott, M. D. Biophysical consequences of linker chemistry and polymer size on stealth erythrocytes: size does matter. *Biochim. Biophys. Acta BBA - Biomembr.*, **1561**, 147–158 (2002).

28. Cohen, M., Fisher, C.J., Huang, M.L., Lindsay, L.L., Plancarte, M., Boyce, W.M., Godula, K., Gagneux, P. Capture and characterization of influenza A virus from primary samples using glycan bead arrays. *Virology*. **493**, 128-135 (2016).

29. Shurer, C. R. Colville, M.J., Gupta, V.K., Head, S.E., Kai, F., Lakins, J.N., Paszek, M.J. Genetically Encoded Toolbox for Glycocalyx Engineering: Tunable Control of Cell Adhesion, Survival, and Cancer Cell Behaviors. *ACS Biomater. Sci. Eng.* (2017) doi:10.1021/acsbiomaterials.7b00037.s

30. Alarcón-Manjarrez, C., Arcos-Ramos, R., Álamo, M. F. & Iglesias-Arteaga, M. A. Synthesis, NMR and crystal characterization of dimeric terephthalates derived from epimeric 4,5-seco-cholest-3-yn-5-ols. *Steroids,* **109**, 66–72 (2016).

31. Morotti, A. L. M., Lang, K. L., Carvalho, I., Schenkel, E. P. & Bernardes, L. S. C. Semi-Synthesis of new glycosidic triazole derivatives of dihydrocucurbitacin B. *Tetrahedron Lett*. **56**, 303–307 (2015).

32. Honigfort, D., Zhang, M., Verespy III, S., and Godula,K. Engineering of spectator glycocalyx structures to evaluate molecular interactions at crowded cellular boundaries. *Faraday Discuss.*, **219**, 138-153 (2019).

33. Gervais, M., Labbé, A., Carlotti, S. & Deffieux, A. Direct Synthesis of α-Azido,ω-hydroxypolyethers by Monomer-Activated Anionic Polymerization. *Macromolecules*, **42**, 2395–2400 (2009).

34. Meyer, J., Keul, H. & Möller, M. Poly(glycidyl amine) and Copolymers with Glycidol and Glycidyl Amine Repeating Units: Synthesis and Characterization. *Macromolecules,* **44**, 4082–4091 (2011).

Chapter 4: Expansion of glycopolymer synthesis to introduce glycan diversity.

Section 4.1: Glycan diversity drives biological functions.

The prior chapters of this thesis focused on the steric and biophysical properties of the glycocalyx, exploring the exclusionary and organizational impacts of non-interacting mucin-like structures. For that purpose, we used glycans that are refractory to most glycan binding proteins to isolate the contributions of physical bulk toward cellular binding events. While this work contributed to the understanding of the glycocalyx interface as a physical barrier between an organism and its environment, there are also many cell surface events that deal with curating the exchange of biological information through glycan binding interactions. Within a cellular context, these glycan binding events are difficult to study through traditional biochemical means due to their non-templated biosynthesis and subsequent diversity of glycosylation, as well as complicating cytosolic signaling events originating from the proteins they are appended to. To mitigate these challenges, this chapter will focus on designing chemical approaches to synthesize glycopolymers modified with an expanded repertoire of glycan components. The Godula lab and others have utilized this chemical approach as an alternative to more classical genetic, chemical, and enzymatic techniques, particularly for elucidating the roles of glycans in biological functions,[1,2,3,4] or how nanoscale presentation effects function.[5,6]

Our lab has previously utilized aminooxy chemistry to introduce a library of complex unmodified reducing sugars to RAFT based polymer scaffolds.[7] In order to utilize this well developed ligation strategy to expand the biological scope and experimental utility of the PEG based glycopolymer synthesis that is the central topic of this thesis, this chapter will describe the modification of the polymer architecture (**Fig 4.1a**) to allow for the direct attachment of unmodified glycans. As the polymer backbone and endgroup functionalization has been

described previously in this thesis, this section will focus on the sidechain modifications required to append unmodified glycans in high throughput fashion through use of aminooxy ligation (**Fig 4.1b**).

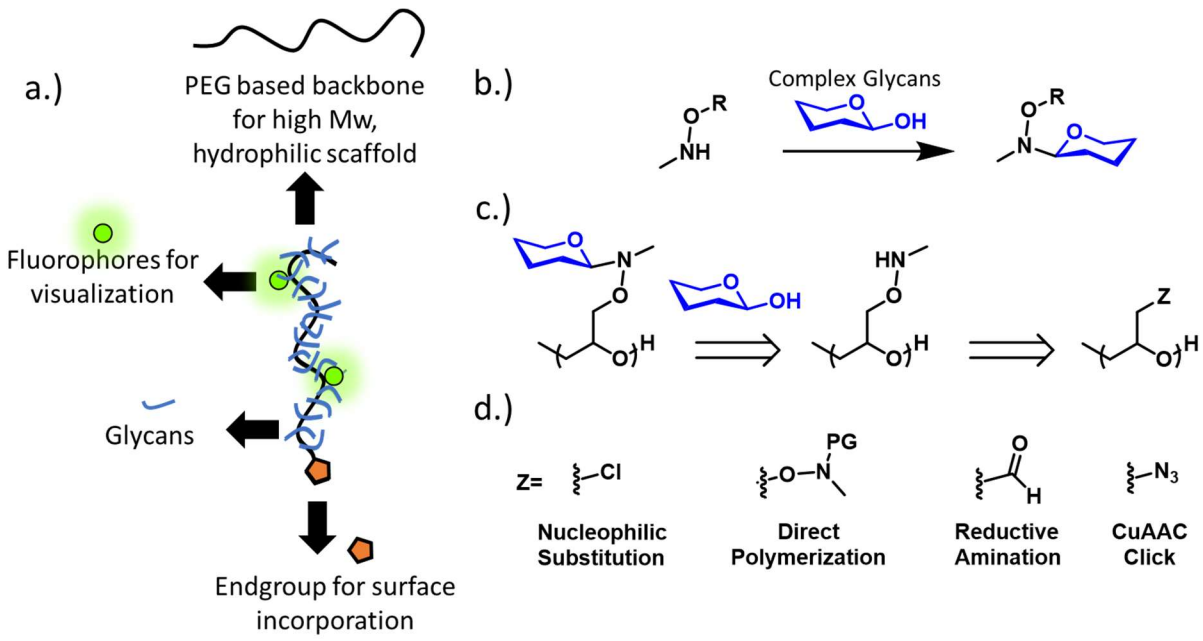


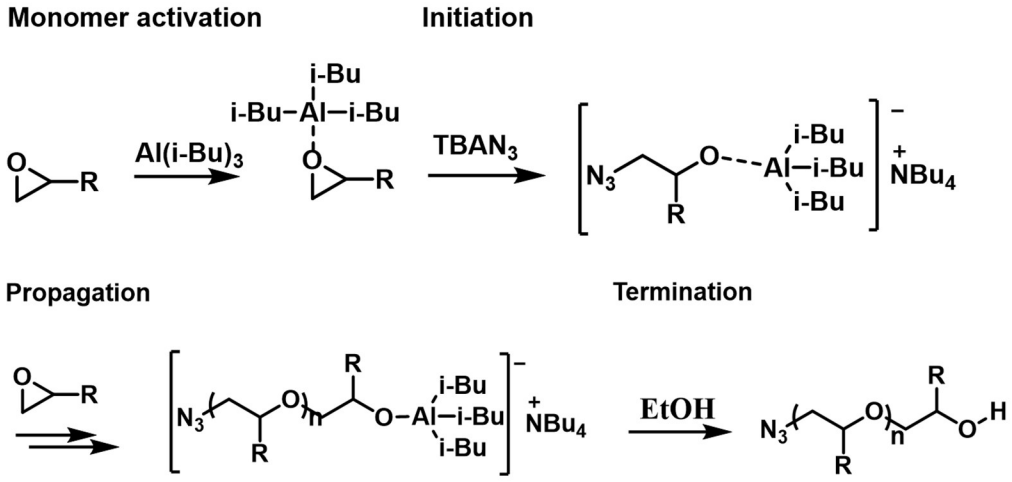
Figure 4.1. Desired glycopolymer features, architecture, and synthesis. a. proposed glycopolymer characteristics. b. Aminooxy moieties react with the reducing end of glycans to yield stable ß - linkages. c. Proposed retrosynthesis of PEG based glycopolymer backbone. d. Potential routes to aminooxy functionalization.

We proposed that this could be accomplished through modification of the PEG backbone to include pendant aminooxy linkages (**Fig 4.1c**), which we envisioned could come from a variety of retrosynthetic sources (**Fig 4.1d**). The following subchapters will focus on substitution of chlorides with aminooxy nucleophiles, direct polymerization of aminooxy containing epoxide monomers, the reductive amination of aldehydes, and the utilization of copper catalyzed azide alkyne cycloaddition (CuAAc) click chemistry to append glycans using and aminooxy-alkyne linker molecule. It is our hope that this molecular architecture can be

used as a chemical biology tool to better understand the interactions that underpin the transfer of biochemical information at the glycocalyx.

Section 4.2: Nucleophilic substitution to append aminooxy functionality to p(ECH).

To build the glycopolymer scaffold, Lewis acid catalyzed anionic polymerization of epoxides was chosen to yield PEG-type polymer backbones with pendant chloride functional groups.[8] As previously described, this procedure developed by Carlotti et al. uses an aluminum Lewis acid catalyst for the activation of epoxies for nucleophilic attack, an azide initiator to introduce endgroup functionality, and living anionic propagation to yield well defined polymers of high Molecular weight and low dispersity (**scheme 4.1**).



Scheme 4.1. Mechanism for polymerization of epoxides by lewis acid catalyzed anionic ring opening polymerization.

Utilization of this technique to polymerize the chloride containing epoxide monomer epichlorohydrin (**1**) generated poly(epichlorohydrin) (pECH, **2**) in good yield with a high degree of control over the final degree of polymerization (DP) with narrow chain length dispersity (Đ=1.15-1.23) (**fig. 4.2**).

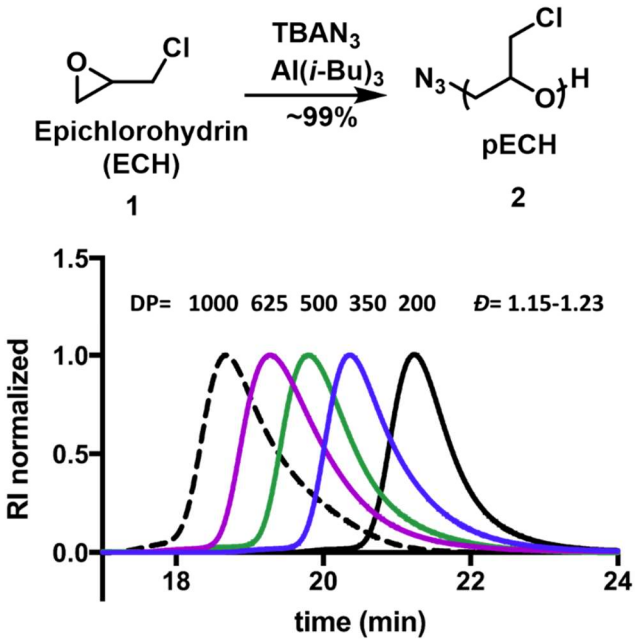


Figure 4.2. polymerization of epichlorohydrin with control over DP and Đ. DP = degree of polymerization, Đ = dispersity, RI = refractive index.

With the parent polymer in hand, we set out to append an aminooxy group through direct nucleophilic substitution of the primary chloride side chains (**fig. 4.3a**). Aminooxy groups N-protected with either a Boc (**3**) or Cbz (**4**) were pretreated with sodium hydride, and the activated alkoxide was allowed to react with pECH **2**. In both cases, the nucleophile successfully substituted the chlorides in the SN2 reaction with ~70% functional group conversion (**5** and **6** respectively). Running the reaction with additional equivalents of base, nucleophile, or additional time did not increase the reaction conversion, indicating that the addition of the sterically constraining protected aminooxy groups hinder further SN2 reaction above this threshold.

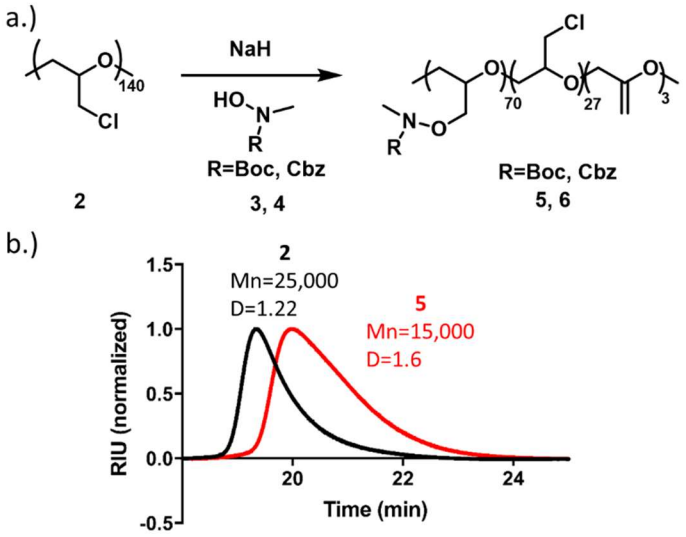
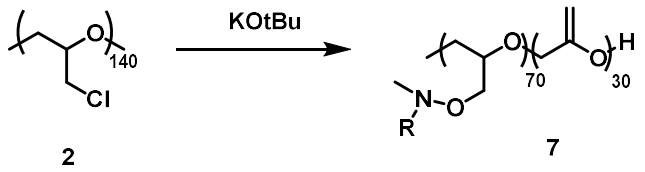


Figure 4.3. Synthesis of aminooxy functionalized polymers. a. synthetic scheme. b. GPC analysis of substitution reactions.

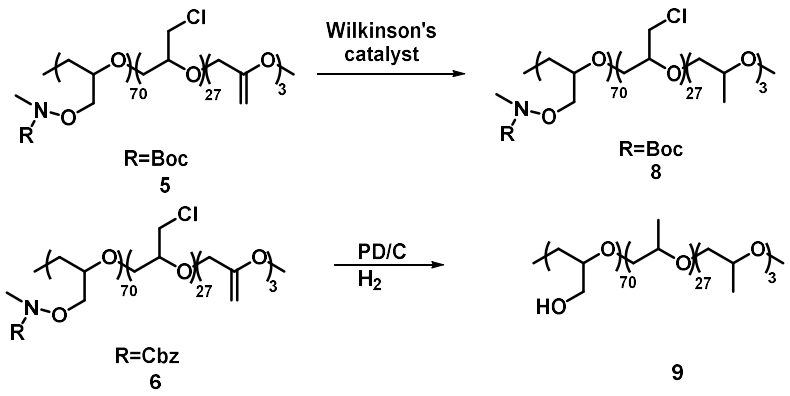
Analysis of the sample by gel permeation chromatography (GPC) indicated that, although the functional group conversion was occurring as intended, the polymer backbone itself was being degraded during the course of the reaction (**fig. 4.3b**) as demonstrated by the reduction in molecular weight and increase in dispersity of the product **5** relative to the parent pECH **2**.

Previous studies from the Hawker lab involving an intentionally biodegradable PEG based backbone suggested that elimination of the chloride from the pECH to yield an enol ether results in a highly degradable polymer architype.[9] Unfortunately, peaks at 4.3 ppm on ^{1}H NMR indicated the formation of a small amount of elimination byproduct in the alkylated product **5** (~3%), indicating that this may be the source of the degradation. To confirm, I intentionally eliminated the Chloride from pECH **2** using potassium tert-butoxide (KOtBu), the product of which (**7**) showed a significant loss of molecular weight and dispersity similar to what was observed with in the formation of **5** (**Scheme 4.2**), indicating that competing elimination reactions are a likely mechanism for the degradation observed.



Scheme 4.2. Determination of p(ECH) degradation pathways.

With this knowledge, we sought to optimize the reaction to avoid the elimination byproduct through utilization of water free conditions, lower temperatures, and different bases, but were unable to completely prevent the formation of the enol ether. As such, we next sought to reduce the enol ether prior to degradation of the backbone.



Scheme 4.3. Attempted synthesis of deprotected aminooxy polymer without backbone degradation.

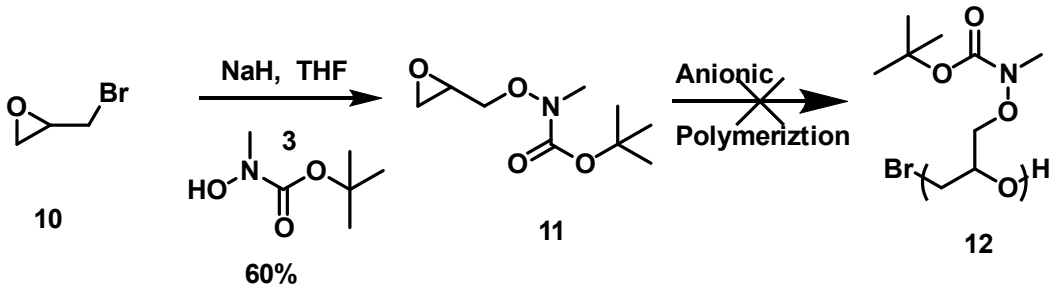
Attempts to reduce the undesired enol ether along with the unreacted chloride and Boc/Cbz functional groups proved challenging (**Scheme 4.3**). Strong reducing agents such as PD/C reduce all the desired functional groups, but also reduce the aminooxy bond to the primary alcohol (**9**). More specialized reducing conditions for the enol ether, such as Wilkinson's catalyst,[10] successfully reduce the enol ether, but not other functional groups (**8**).

Additionally, in all cases some backbone degradation was observed by GPC, indicating that some enol ethers were cleaved prior to reduction even under anhydrous conditions. Due to these difficulties with backbone degradation in the substitution reaction, we next sought to directly polymerize epoxides with an aminooxy group already attached.

Section 4.3: Generation of novel aminooxy epoxide monomers

Generation of epoxide monomers compatible with lewis acid catalyzed conditions for the anionic polymerization employed herein to generate polymers of high molecular weight and low dispersity is a synthetic challenge.[11] The current substrate scope of monomers capable of achieving high degrees of polymerization is limited almost exclusively to alkyl epoxides (such as propylene oxide) or ether containing epoxides (such as ethoxyethyl glycidyl ether). We sought to expand this substate scope in order to achieve our desired aminooxy functionality through the synthesis and development of novel functional monomers (**fig. 4.1.d**).

Our initial attempt at polymerization of an aminooxy containing epoxide came from the nucleophilic substitution of epibromohydrin **10** with a Boc protected aminooxy **3** using sodium hydride as a base in THF (**fig. 4.3**). The resulting epoxide **11** was subjected to the anionic polymerization conditions described above, but resulted in no reaction, with only starting material recovered. This may be due to the large number of chelating groups in the monomer, which could sequester the aluminum catalyst and prevent necessary activation of the epoxide for initiation and propagation.



Scheme 4.4. Synthesis and polymerization of Boc-protected aminooxy containing epoxide.

Next, we sought to generate an epoxide containing an oxime-protected aminooxy functional group, which contains fewer chelating groups than the previous Boc protected version **11**. Freshly distilled[12] glycidol **13** was activated for nucleophilic attack by reaction with MsCl in the presence of triethylamine base catalyst (**fig. 4.4**). The resulting mesylated epoxide **15** was subjected to base catalyzed nucleophilic attack by acetone oxime **16** to yield the final polymerizable monomer **17**.

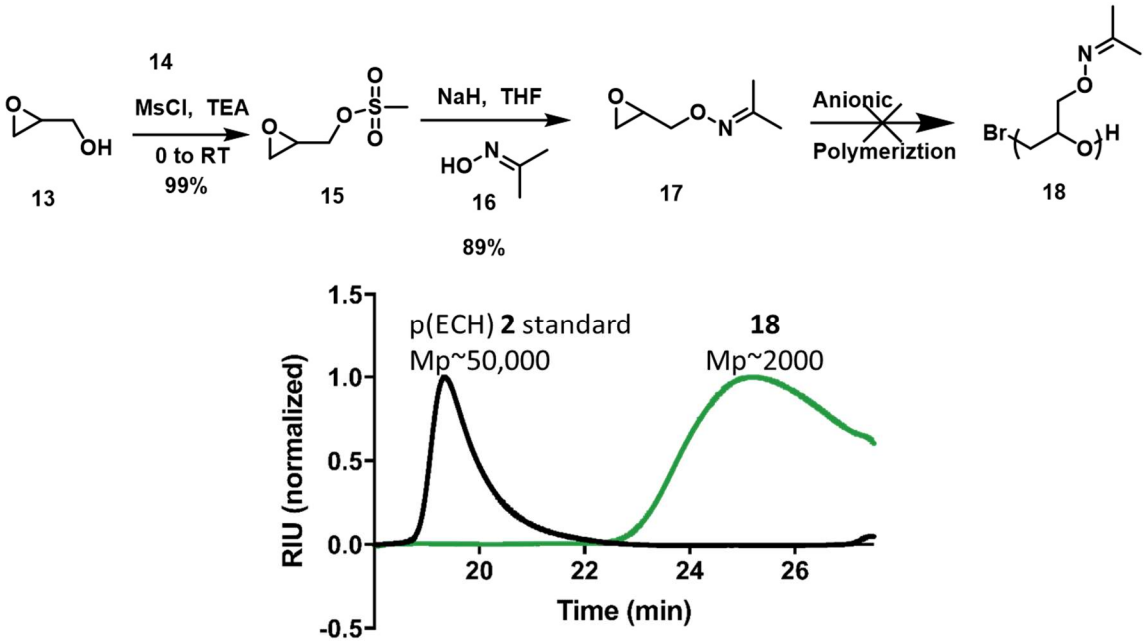
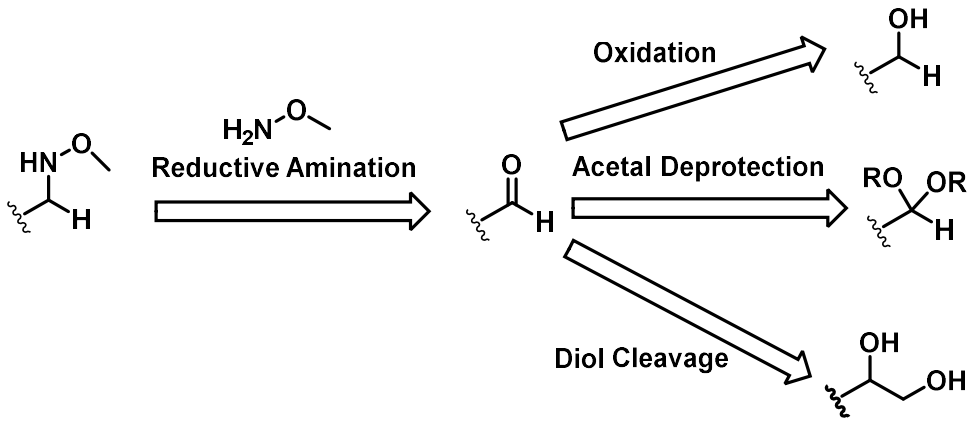


Figure 4.4. Synthesis and polymerization of acetone oxime-protected aminooxy containing epoxide.

Polymerization of **17** to yield a protected aminooxy polymer **18** did successfully generate a polymer as determined by GPC, but attempts to synthesize polymers of high molecular weight or narrow dispersity were unsuccessful. The polymer reached a peak molecular weight of only 2000 Da, with very broad dispersity indicating uncontrolled or non-living reaction mechanisms. Possible rationales include, again, the complexation of aluminum to the lone pairs present in the aminooxy functionality, or the activation of the imine by the Lewis acid catalyst to chain terminating attack by the growing alkoxide chain. Previous approaches to glycopolymers have utilized a RAFT based polymerization of polyacrylamides containing acetone oxime protected aminooxy groups,[13,] but to date oximes have not been successfully polymerized to high molecular weight using ring opening anionic polymerization conditions. Though the monomer does not meet the specifications for our synthesis, it may be suitable for applications in which control over molecular weight and dispersity are less rigorous, or as a low abundance copolymer in a more robust monomer mixture.

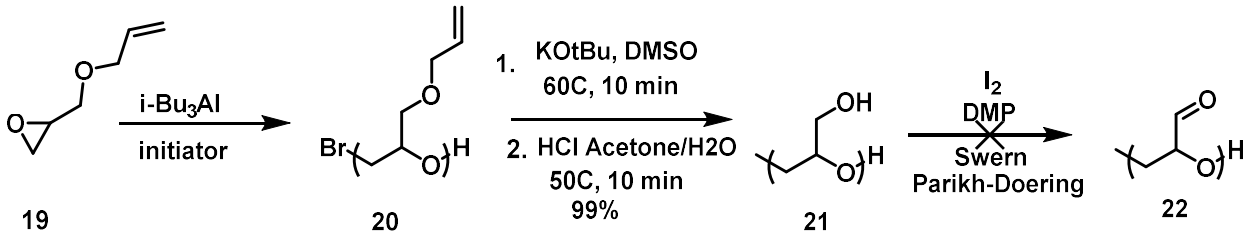
Section 4.4: Appending aminooxy functionality to PEG through reductive amination.

Due to the apparent incompatibility of the aminooxy functionality to the chosen polymerization conditions, we next sought to append the it to the polymer through reductive amination of pendant aldehyde groups (**Scheme 4.5**). Reductive amination to convert aldehydes to aminooxy functionalities using BH3-pyridine is known in the literature,[14] and other groups have synthesized glycopolymers by appending an aminooxy group to their backbone by reductive amination of a ketone. [15,16]



Scheme 4.5. retrosynthesis of aminooxy containing polymer from reductive amination.

Initial attempts to generate a poly(aldehyde) polymer involved the oxidation of a primary alcohol, which was generated through the polymerization of allyl glycidyl ether (AGE) **19**. Poly(AGE) **20** was then deprotected by treating with base to cause migration the alkene from the terminal allylic position to the internal enol ether, which was then cleaved by addition of acid and water to yield poly(glycidol) **21** (**Scheme 4.6**).



Scheme 4.6. Synthesis toward a poly(aldehyde) PEG based polymer.

Attempts to oxidize this alcohol to the poly(aldehyde) **22**, however, proved challenging across a screen of oxidation conditions, including with Iodide, DMP, Swern, and Parikh-Doering. Most commonly, the polymer crashed out of solution as an insoluble precipitate that could not be further analyzed, possibly as a result of inter - and intramolecular reactions of freshly formed aldehydes with unreacted alcohol starting materials to form a highly crosslinked polymer network. In an attempt to circumvent this issue, we next attempted to

synthesize a protected aldehyde monomer that could be deprotected without these side reactions (**fig 4.5**).

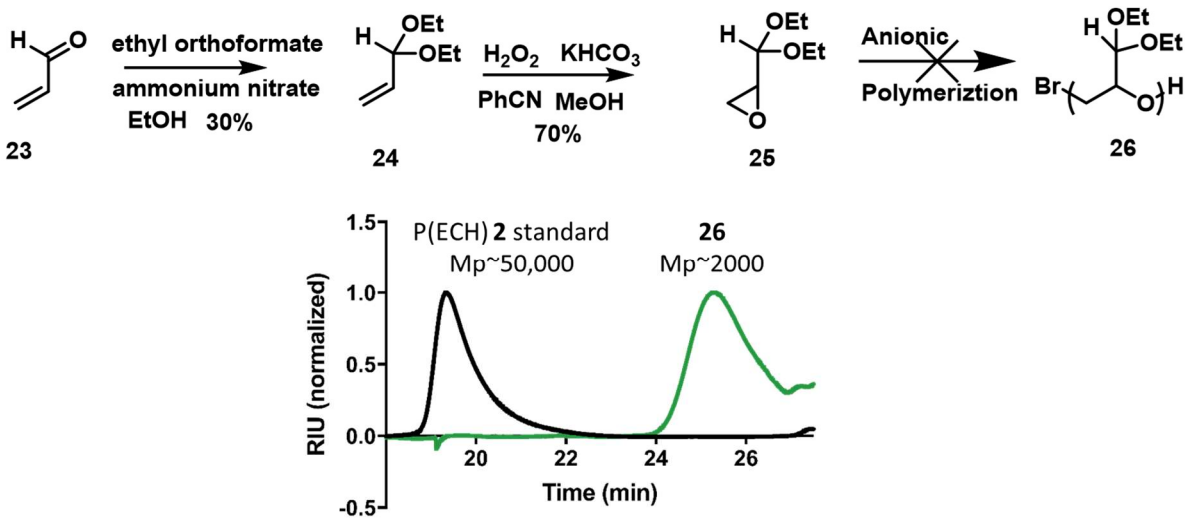


Figure 4.5. Synthesis and polymerization of acetal protected aldehyde containing epoxide.

To generate this structure, we protected acrolein **23** with ethyl orthoformate and ammonium nitrate to yield the acetal **24**. Epoxidation of the alkene with hydrogen peroxide yields the epoxide monomer **25** in good yield.[17,18] However, polymerization to yield the acetal protected polymer **26** again produced a polymer that, though relatively well controlled in dispersity, did not polymerize to high molecular weight (Mp ~ 2000Da). Though this monomer does not appear to be a good candidate for high molecular weight homopolymer applications, it could be suitable for low Mw applications, or as a low abundance copolymer in a more robust monomer mixture.

A final route to an aldehyde containing polymer was though proposed cleavage of a diol sidechain (**fig. 4.6**). In this route, the diol of L-ascrobic acid **27** was protected with acetone to yield **28**, followed by cleavage with potassium carbonate and hydrogen peroxide to yield the acid **29**.[19] Conversion of the carboxylic acid to an ethyl ester **30** using ethyl iodide was

followed by LAH reduction to yield an unprotected diol **31**.[20] The primary alcohol was then selectively tosylated in intermediate **32**, followed by a ring closing reaction under basic conditions to yield the final epoxide **33**.[21]

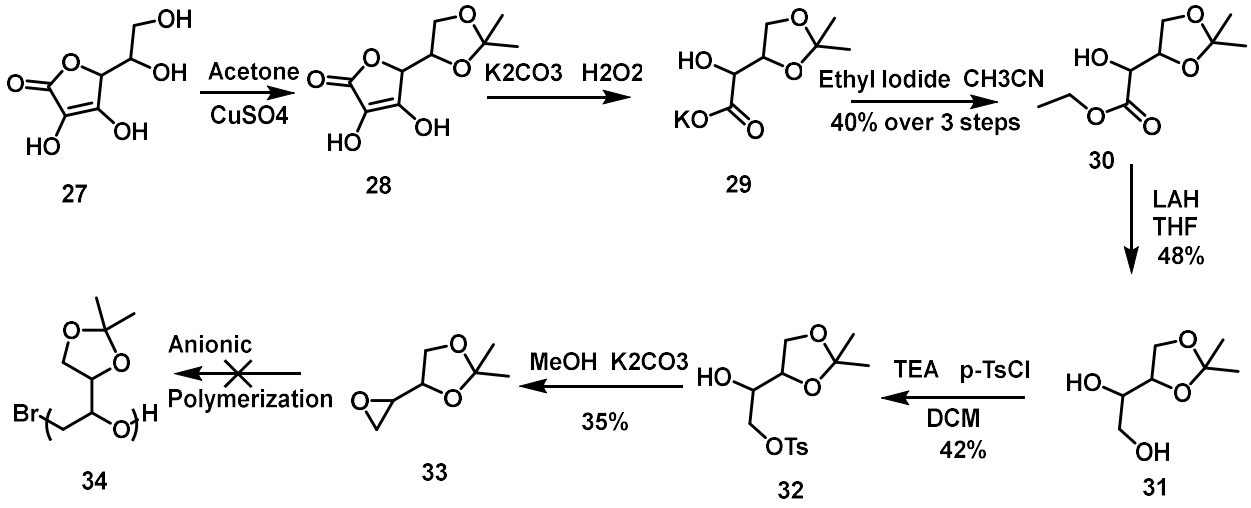


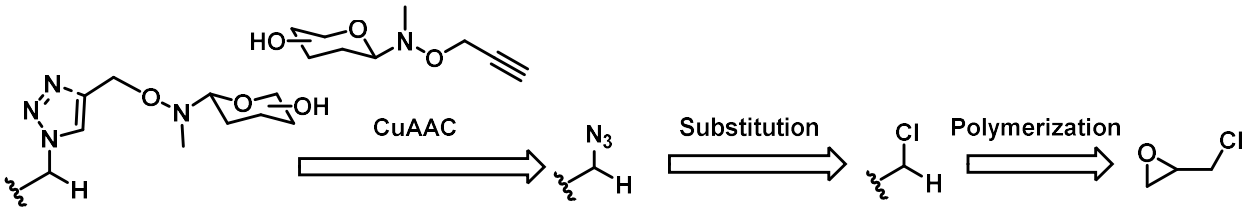
Figure 4.6. Synthesis and polymerization of diol containing epoxide as an aldehyde precursor.

Attempted polymerization of this epoxide resulted in only starting material recovery, with no noticeable polymer formation. This may be due to complexation of the aluminum with the diol, sequestering the aluminum from catalyzing the polymerization reaction at the epoxide.

Overall, we were unable to generate an aminooxy containing polymer of high molecular weight through these methods, and as such turned our attention to more robust chemistry to append complex glycans through a modification of our previous "click" chemistry based method.

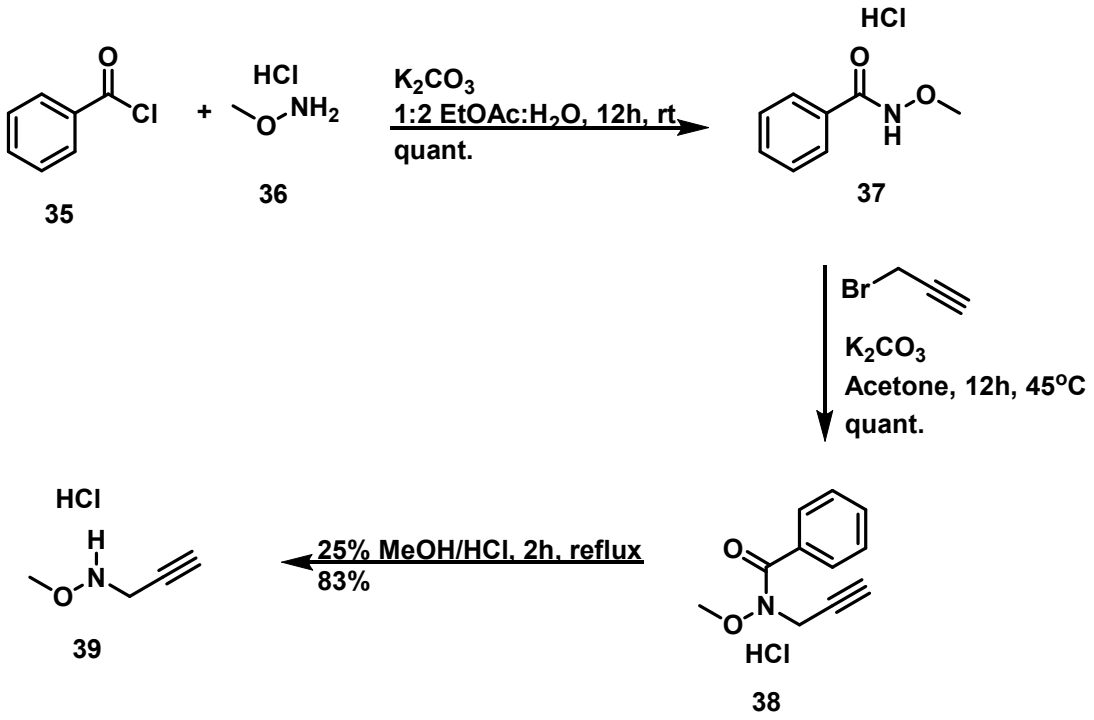
Section 4.5: Appending aminooxy linked glycans through CuAAC reactions.

Seeking a more reliable approach for the addition of complex glycans to the polymer backbone, we sought to reappropriate previously developed methodologies (chapters 1 and 2)[22] utilizing Copper catalyzed Azide-Alkyne cycloaddition.[23] Retrosynthetically, we envisioned appending glycans linked to an alkyne connected to an aminooxy linker by CuAAc chemistry (**scheme 4.7**). This method would still allow for the addition of more complex glycans through reaction at their reducing end, but takes advantage of better developed chemistries for glycopolymer synthesis.



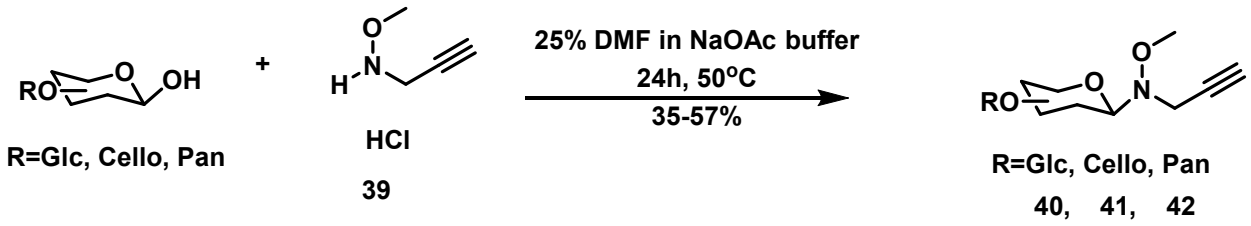
Scheme 4.7. Retrosynthesis of glycopolymers utilizing CuAAC ligation.

In order to utilize this strategy, we needed to develop a functional aminooxy-alkyne linker that reacts with unmodified glycans on one end, and contains an alkyne on the other for attachment to the polymer backbone (**Scheme 4.8**).[24]



Scheme 4.8. Synthesis of aminooxy-alkyne linker.

To make this linker, O-methyl aminooxy **36** was Benzyl protected using benzyl chloride to prevent overalkylation in the subsequent steps, as having multiple alkynyl polymer attachment sites could lead to crosslinking and other undesired outcomes. The protected aminooxy **37** was then alkylated with propargyl bromide to give **38**, and deprotection of the benzyl group in acidic conditions yielded the final product **39**. This short linker will allow one step ligation to any glycan with a reducing end, and will facilitate attachment to the polymer backbone through its terminal alkyne. To prove the viability of this method, we appended this linker to glycans of increasing size – the monosaccharide glucose (Glc, **40**), disaccharide Cellobiose (Cello, **41**), and the trisaccharide Panose (Pan, **42**) (**Scheme 4.9**).



Scheme 4.9. One step synthesis of ß-alkynyl glycans.

With the glycans in hand, we next went about synthesizing polymer scaffolds to which they could be attached (**fig. 4.7**). As described previously, pECH **2** was generated from the epoxide monomer ECH **1**. The resulting azide endgroup from initiation with tetrabutyl ammonium azide was modified using CuAAc chemistry to append the previously described hydrophobic endgroup **43** for cellular incorporation.[25] The end modified polymer **44** was then treated with sodium azide to substitute the chloride sidechains for azides in poly(glycidyl azide) (pGA, **45**).[26] Finally, alkynyl fluorophores and glycans were added to yield the final functional glycopolymer.

This CuAAc catalyzed reaction is done in one pot, with the sub-stoichiometric fluorophore added to an air free solution of the polymer backbone along with copper catalyst. This first step is done to ensure that a desired amount of fluorophore is appended stochastically along the backbone. After 1hr, 1.5 equivalents of glycans **40**, **41**, or **42** were added to the solution, which was then allowed to react to completion overnight, resulting in the successful synthesis of mono- and di- saccharide glycopolymers **GP-46** and **GP-47**. The addition of a trisaccharide pannose in **GP-48** resulted in an insoluble precipitate, indicating that the polymer cannot stably add high quantities of such a large glycan to the backbone. The resulting solid was insoluble in all organic, aqueous, and cosolvents applied, including water, DMSO, DMF, conc. HCl, and conc NaOH.

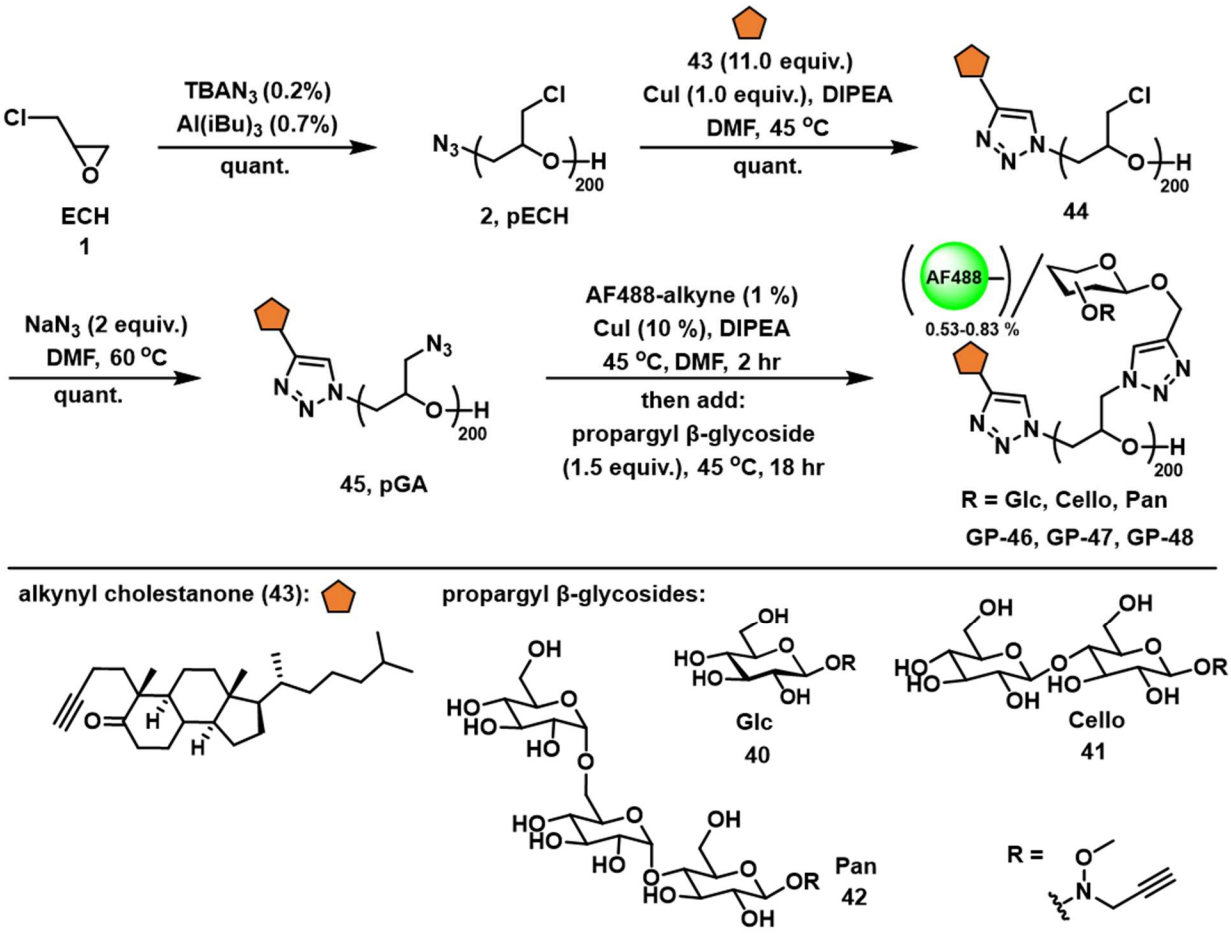


Figure 4.7. Synthesis of glycopolymers from aminooxy functionalized glycans **40** - **42**.

Analysis of the final highly glycosylated polymers **GP-46** and **GP-47** is a challenge due to their likely extended conformation imparted by the bulky sidechain steric interactions.[27,28] All polymer, linker, and glycan peaks in proton NMR congregate between 3-4.5ppm without any distinguishable signals, outside of the triazole proton resulting from the click reaction (NMR appendix pages 220-221). GPC is similarly unreliable, as the polymer's linear conformation causes it to interact with the column differently than globular proteins or polymers. One way to qualitatively measure progress of the reactions are through infrared spectroscopy (IR). The azide moiety utilized for CuAAc reactions throughout the iterative click scheme (**fig. 4.7**) has a strong asymmetric vibration absorption at 2100 cm^{-1}, which is a

distinctive region of the IR spectrum. Monitoring this peak throughout the reaction sequence gives an indication of how much azide is consumed (or generated) at each step, and gives a qualitative readout of reaction completion (**fig. 4.8**).

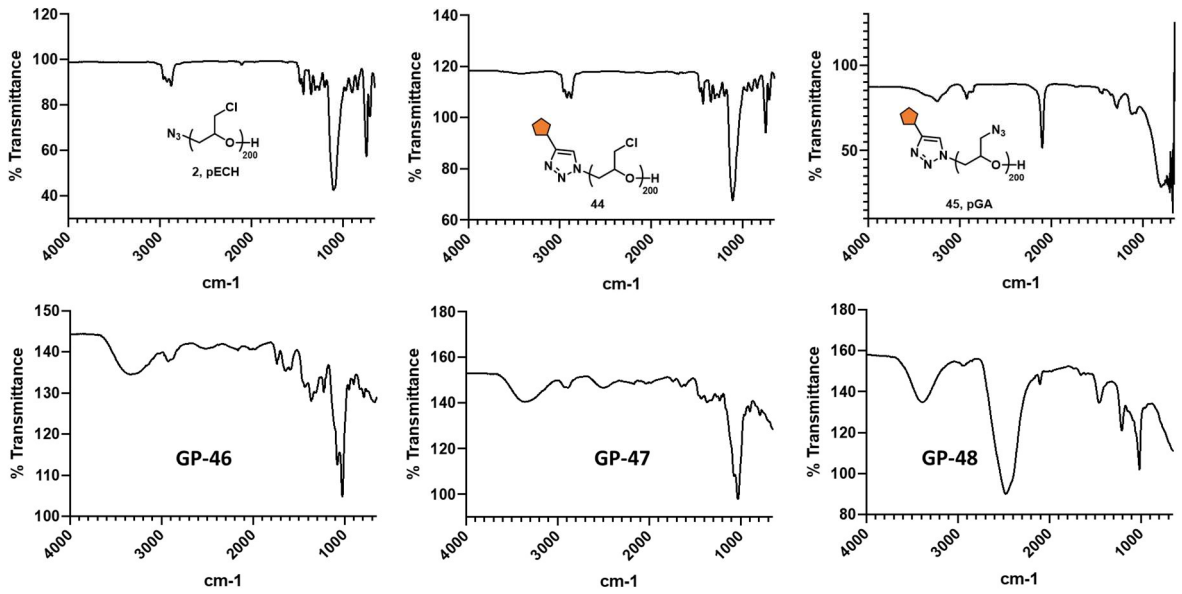


Figure 4.8. Analysis of glycopolymer reaction progress by IR.

The initial pECH **2** has a small peak at 2100 cm^{-1}, as expected due to the presence of only one terminal azide for 200 repeating units. Reaction of that azide to append the hydrophobic endgroup in polymer **44** results in a complete disappearance of the azide peak by IR. Substitution of sidechains with azides in polymer **45** resulted in a significant increase in azide absorbance, and subsequent reaction of the azide to append glycans in **GP-46** and **GP-47** again resulted in a complete disappearance of the peak at 2100 cm^{-1}, indicating quantitative side chain functionalization.[22] Analysis of the solid glycopolymer precipitate from the conversion of polymer **45** to **GP-48** with propargyl pannose glycoside **42** indicates that the trisaccharide glycopolymer crashed out of solution prior to complete substitution of its side chains, as evidenced by the retention of a small residual azide peak at 2100 cm^{-1}. This

is likely due to reduced solubility of the resulting bulky glycopolymer, which led to precipitation as the reaction approached completion. Though in this instance glycans were appended to the trisaccharide polymer backbone until it became insoluble, the fact that the glycosylation reaction proceeded until it was halted by precipitation suggests that larger glycans can be readily appended to the backbone so long as care is taken to maintain the glycan attachment below a critical side chain substitution threshold. In future work, this could be controlled through the stoichiometry of glycan addition, specifically by using < 1 equivalent of larger or more complex glycans.

Section 4.6 Conclusion

This chapter detailed the synthetic routes we explored toward attachment of reducing sugars to previously described PEG based glycopolymers via aminooxy ligation. Despite some success with functional group interconversion, post-polymerization modification to append aminooxy moieties to pECH and pAGE proved non-viable, in many cases due to the fragile nature of the backbone itself to most reaction conditions. The experience handling these polymer reactions informed the design and synthesis of epoxide monomers for anionic polymerization, but these proved incompatible with the Carlotti method. Finally, we took advantage of our expertise in click chemistry functionalization of PEG based polymers to successfully append glycans conjugated to alkynes through aminooxy linkages. Through this method glycopolymers of mono- and di-saccharides can be readily accessed, and larger glycans can be appended to the backbone but require additional optimization to overcome reduced solubility at high ligation efficiencies. Even with this limitation, this architecture will expand the functional capabilities of PEG based glycopolymers, and the methods may be further optimized and expanded to include a larger library of complex glycans.

Section 4.7 Materials and Methods

Experimental Section

Instrumentation. Column chromatography was performed on a Biotage Isolera One automated flash chromatography system. Nuclear magnetic resonance (NMR) spectra were collected on a Bruker 300 MHz and a Jeol 500 MHz NMR spectrometers. Spectra were recorded in CDCl3 or D2O solutions at 293K and are reported in parts per million (ppm) on the δ scale relative to the residual solvent as an internal standard (for 1H NMR: CDCl3 = 7.26 ppm, D2O = 4.79 ppm, for 13C NMR: CDCl3 = 77.0 ppm). HRMS (high-resolution mass spectrometry) analysis was performed on an Agilent 6230 ESI-TOFMS in positive ion mode. UV-Vis spectra for polymer fluorophore content quantification were collected using a quartz cuvette using a Thermo Scientific Nanodrop2000c spectrophotometer. IR spectroscopy was performed on a Nicolet 6700 FT-IR spectrophotometer (Thermo Scientific). Size exclusion chromatography (SEC) was performed on a Hitachi Chromaster system equipped with an RI detector and two 5 μm, mixed bed, 7.8 mm I.D. x 30 cm TSKgel columns in series (Tosoh Bioscience). Organic soluble polymers were analyzed using an isocratic method with a flow rate of 0.7 mL/min in DMF (0.2% LiBr, 70 °C). For aqueous SEC, two 8 μm, mixed-M bed, 7.5 mm I.D. x 30 cm PL aquagel-OH columns in series (Agilent Technologies) were run in sequence using an isocratic method with a flow rate of 1.0mL/min in water (0.2M NaNO3 in 0.01M Na2HPO4, pH = 7.0). **Materials**. All chemicals, unless stated otherwise, were purchased from Sigma Aldrich and used as received. Reaction progress was checked by analytical thin-layer chromatography (TLC, Merck silica gel 60 F-254 plates) monitored either with UV illumination, or by staining with iodine, ninhydrin, or CAM stain. Solvent compositions are reported on a volume/volume (v/v) basis unless otherwise noted.

Synthesis of azide-terminated poly(epichlorohydrin), pECH (2). Epichlorohydrin was polymerized according to the procedure developed by Carlotti. Briefly, a flame-dried Schlenk flask (10mL) equipped with a magnetic stirrer and fitted with PTFE stopcock was charged with tetrabutylammonium azide ($TBAN_3$, 40 mg, 0.074 mmol,0.004 equiv.) under argon. A solution of freshly distilled epichlorohydrin (1.29 mL, 16.5 mmol) in anhydrous toluene (4 mL). A solution of triisobutylaluminum in toluene (1.07 M, 200 µL, 0.222 mmol, 0.014 equiv.) was added via a syringe under argon at −30 °C. The reaction was stirred for 4 hours and then stopped by the addition of ethanol. The resulting pECH polymer **2** was precipitated into hexanes and dried under vacuum to yield a clear viscous oil (1500 mg, 99 % yield). The polymer was analyzed by SEC (0.2% LiBr in DMF): M_w = 24,000, M_n = 20,000, Đ = 1.20).

Synthesis of aminooxy substituted pECH (5-6). Protected aminooxy **3** (Boc) or **4** (Cbz) (0.79 mmol, 1.75 eq) was added to a flame dried Schlenk flask, and was dissolved in anhydrous DMF (2.0 mL). NaH (0.68 mmol, 1.5eq) was added, and the reaction was stirred at 0 °C for 30min. To this solution was added pECH **2** (0.45 mmol, 1.0 eq) in DMF (0.5 mL). The mixture was stirred for 12hrs at 65 °C. The reaction mixture was quenched with DOW-EX resin until pH paper turned neutral, and rotovapped until only a small amount of liquid remained. Residue was diluted with DCM and passed through a cotton plug to remove solids, then precipitated into hexanes to yield protected aminooxy polymers **5-6** as a clear oil with ~90% mass recovery, 80% aminooxy group conversion.

Synthesis of partially eliminated pECH (7). To a solution of pECH **2** (0.1 mmol, 1.0 eq) in dry toluene (1.5 mL) was added potassium *tert*-butoxide (0.03 mmol, 0.3 eq). The reaction mixture was stirred overnight at room temperature, and the reaction mixture was quenched with DOW-EX resin, and rotovapped until only a small amount of liquid remained. Residue was diluted with DCM and passed through a cotton plug to remove solids, then precipitated

into hexanes to yield the partially eliminated polymer **7** as a clear oil with ~90% mass recovery, ~30% side chain elimination.

Synthesis of Boc-deprotected aminooxy polymer (8). To a solution of Boc-protected aminooxy polymer **5** (0.05 mmol, 1.0 eq) in an anhydrous benzene (1.0 mL)/ EtOH (0.15 mL) mixture was added Wilkinson's catalyst (spatula tip). Hydrogen gas was bubbled through the solution, then the vial was capped and the reaction was stirred for 24hrs at room temperature. The mixture was filtered through celite, rotovapped, and the Boc-deprotected aminooxy polymer **8** was isolated as a lightly colored film in >80% mass recovery, 100% of olefins reduced.

Synthesis of fully reduced aminooxy polymer (9). To a solution of Cbz-protected aminooxy polymer **6** (0.05 mmol, 1.0 eq) in anhydrous DMF (1.0 mL) was added Pd/C catalyst (spatula tip). Hydrogen gas was bubbled through the solution, then the vial was capped and the reaction was stirred for 24hrs at room temperature, periodically adding additional anhydrous MeOH until it comprises 75% of reaction volume. The mixture was filtered through celite, rotovapped, and the fully-deprotected aminooxy polymer **9** was isolated as a lightly colored film in >80% mass recovery, 100% of olefins, chlorides, and aminooxy groups reduced.

Synthesis of Boc-protected aminooxy containing epoxide (11). To a solution of Boc-protected aminooxy **3** (13.6 mmol, 1.0 eq) in anhydrous THF (30.0 mL) was added NaH (16.3 mmol, 1.2 eq) in two portions at 0 °C. After 15 min, epibromohydrin **10** was added and the reaction was stirred at room temperature overnight. The product was purified using column chromatography with a 20% EtOAc/Hexanes mixture. The epoxide product was isolated as a clear liquid in 80% yield.

Synthesis of Mesyl activated glycidol (15). To a solution of glycidol **13** (1.5 mmol, 1.0 eq) in anhydrous DCM (2.25 mL) was added anhydrous Triethylamine (3.0 mmol, 2.0 eq) at 0 °C.

After 15 min, Mesyl Chloride **14** was added and the reaction was allowed to warm to room temperature over 45 min. Added 1M HCl, and extracted 3x with DCM. The combined organic extracts were dried with MgSO4 and the product was rotovapped down to a slightly yellow liquid in 99% yield.

Synthesis of acetone oxine-protected aminooxy containing epoxide (17). To a solution of acetone oxime **16** (0.3 mmol, 1.0 eq) in anhydrous THF (30.0 mL) was added NaH (0.3 mmol, 1.0 eq) in two portions at 0 °C. After 15 min, Mesyl Glycidol **15** was added and the reaction was heated to 40 °C and stirred for 3 hours. The reaction mixture was filtered through celite, rotocvapped to dryness, and purified by distillation to yield the epoxide product as a clear liquid in 89% yield.

Synthesis of Bromide-terminated poly(allyl glycidyl ether), pAGE (20). Allyl glycidyl ether was polymerized according to the procedure developed by Carlotti. Briefly, a flame-dried Schlenk flask (10mL) equipped with a magnetic stirrer and fitted with PTFE stopcock was charged with tetraoctylammonium bromide (TOABr, 28 mg, 0.05 mmol,0.004 equiv.) under argon. A solution of freshly distilled Allyl glycidyl ether (1.52 mL, 12.8 mmol) in anhydrous toluene (5 mL). A solution of triisobutylaluminum in toluene (1.07 M, 144 µL, 0.15 mmol, 0.014 equiv.) was added via a syringe under argon at −30 °C. The reaction was stirred for 4 hours and then stopped by the addition of ethanol. The resulting pAGE polymer **20** was precipitated into hexanes and dried under vacuum to yield a clear viscous oil (1400 mg, 96 % yield). The polymer was analyzed by SEC (0.2% LiBr in DMF): M_w = 33,000, M_n = 29,500, Đ = 1.12).

Synthesis of poly(Glycidol) (21). To a solution of pAGE **20** (0.56 mmol, 1.0 eq) in anhydrous DMSO (1.0 mL) was added potassium tert-butoxide (0.6 mmol, 1.1 eq), and the reaction was heated to 60 °C. After 10 min, added diethyl ether and sat. ammonium hydroxide. The ether

layer was extracted 3x, and the combined organic layers were dried using magnesium sulfate. After filtration and evaporation, the resulting polymer film was dissolved in acetone (4.0 mL) plus 1M HCl (0.44 mL) and was stirred for 15 min at 50 °C until the solution is no longer cloudy. The organic layer was removed and evaporated to yield the polymer as a viscous oil in 93% yield.

Synthesis of 3,3-diethoxyprop-1-ene (24). To a solution of freshly distilled acrolein **23** (111.0 mmol, 1.0 eq) in absolute ethanol (7.4 mL) was added ethyl orthoformate (100.0 mmol, 0.9 eq). Ammonium nitrate (5.5 mmol, 0.05 eq) was added and the reaction was heated to 45 °C for 4 hours, during which the reaction durned a dark red. The reaction was filtered through celite to remove precipitates, and was fractionally distilled to yield the target acetal as a clear liquid in 30% yield after distillation.

Synthesis of 2-(diethoxymethyl)oxirane (25). To a solution of freshly distilled acrolein acetal **24** (0.77 mmol, 1.0 eq) in methanol (0.5 mL) was added benzonitrile (0.77 mmol, 1.0 eq). Hydrogen peroxide (1.2 mmol, 1.3 eq) was added in two parts and the reaction was stirred at 40 °C for 24 hours. Water was added, and the solution was extracted 3x with DCM. The combined organic layers were dried using sodium sulfate, and solvent was evaporated. The final product was fractionally distilled to yield the target epoxide as a clear liquid in 70% yield after distillation.

Synthesis of ethyl 2-(2,2-dimethyl-1,3-dioxolan-4-yl)-2-hydroxyacetate (30). To a solution of 5,6-isoproylidene-L-Ascorbic acid **27** (92.5 mmol, 1.0 eq) in water (100.0 mL) was added potassium carbonate (231.0 mmol, 2.5 eq). The mixture was cooled to 0 °C and Hydrogen peroxide (324.0 mmol, 3.5 eq) was added in dropwise. The reaction was stirred at room temperature for 12 hours, at which point the solvent was removed on rotovap. The resulting mixture was washed with boiling ethanol, and solids were filtered off. The ethanol was

evaporated to yield the crude intermediate potassium 2-(2,2-dimethyl-1,3-dioxolan-4-yl)-2-hydroxyacetate **29** a yellow oil that was taken immediately to the next step.

The crude intermediate **29** was dissolved in acetonitrile (100.0 mL) and ethyl iodide (111.0 mmol, 1.2 eq) was added. The reaction mixture was stirred at 75 °C for 22 hours, at which point solids were filtered off and solvent evaporated. The product ester **30** was purified by distillation to yield a clear liquid in 40% overall yield.

Synthesis of 1-(2,2-dimethyl-1,3-dioxolan-4-yl)ethane-1,2-diol (31). To a solution of ester **30** (39.2 mmol, 1.0 eq) in anhydrous THF (15 mL) was added Lithium Aluminum Hydride (54.8 mmol, 1.4 eq) dropwise at 0 °C. The reaction was allowed to slowly warm to room temperature, then was heated to 60 °C for one hour. The reaction was cooled to room temperature, then water was added, followed by 5% NaOH solution until foaming stopped. The Aq layer was extracted 3x using chloroform, and the solid aluminum salts were extracted using Soxhlet technique with boiling chloroform for 24hrs. The combined organic layers were dried using magnesium sulfate, and rotovapped to yield the diol product 31 as a clear oil in 48% yield.

Synthesis of 2-(2,2-dimethyl-1,3-dioxolan-4-yl)-2-hydroxyethyl 4-methylbenzenesulfonate (32). To a solution of diol **31** (18.5 mmol, 1.0 eq) in anhydrous DCM (30.1 mL) was added triethylamine (55.5 mmol, 3.0 eq). The solution was cooled to 0 °C and tosyl chloride (20.4 mmol, 1.1 eq) was added dropwise. The reaction mixture was stirred at room temperature for 5 hours, at which point the solvent was evaporated and the mixture was purified via column chromatography using 3:1 Hexanes/EtOAc to yield the mono-mesylated product **32** as a clear oil in 42% yield.

Synthesis of 2,2-dimethyl-4-(oxiran-2-yl)-1,3-dioxolane (33). To a solution of mesylated diol **32** (7.9 mmol, 1.0 eq) in anhydrous methanol (14.0 mL) was added potassium carbonate (19.8 mmol, 2.5 eq). The reaction mixture was stirred at room temperature for 1 hour, at which point saturated ammonium chloride was added. The methanol component was rotovapped off, then the Aq layer was extracted 3x with diethyl ether. The combined organic layers were dried using magnesium sulfate, and solvent was evaporated. The mixture was purified via column chromatography using 3:1 Hexanes/EtOAc to yield the epoxide product containing a protected diol **33** as a clear oil in 35% yield.

Synthesis of Benzyl protected aminooxy (37). To a solution of O-methyl aminooxy (60.0 mmol, 1.5 eq) in 1:2 EtOAc: water (132.0 mL) containing potassium carbonate (90.0mmol, 2eq) was added Benzyl bromide (40 mmol, 1eq). The mixture was stirred for 12hrs at room temperature. The reaction mixture was washed with sat. sodium bicarbonate, and extracted 3x with EtOAc. The products were dried over sodium sulfate and evaporated to give benzyl protected aminooxy **37** as a clear oil in quantitative yield.

Synthesis of mono-propargylated aminooxy (38). To a solution of benzyl protected aminooxy **37** (40.0 mmol, 1eq) in acetone (200 mL) containing potassium carbonate (60.0mmol, 1.5eq) was added propargyl bromide (100 mmol, 2.5eq). The mixture was stirred for 12hrs at 45 °C. The reaction mixture was washed water, and extracted 3x with DCM. The products were purified via column chromatography using 7:1 Hexanes/EtOAc to yield the mono-propargylated aminooxy **38** as a clear oil in quantitative yield.

Synthesis of propargyl-aminooxy linker (39). The mono-propargylated aminooxy **38** (40.0 mmol, 1eq) was dissolved in a 25% solution of MeOH in conc. HCl (110.0 mL), and the mixture was stirred for 2hrs at reflux. The crude mixture was evaporated to dryness and the resulting

solid was washed 3x with ether to remove byproducts. The product was isolated as a white solid the propargyl - aminooxy linker **39** as a clear oil in 83% yield.

General procedure for propargylation of glycans (40-42). To a solution of glycans Glucose, Cellobiose, or Panose (0.35 mmol, 1eq) in 25% DMF in NaOAc buffer pH= 5.5 (1.76 mL) was added aminooxy-propargyl linker **39** (1.23 mmol, 3.5eq). The mixture was stirred for 24hrs at 50 °C. The reaction mixture was rotovaped to dryness, and the products were purified via column chromatography using 5:2:1EtOAc:MeOH:water to yield propargylated glycans **40-42** as white solids: **40**-Glc (35%) **41**-cello (57%) **42**-Pan (95%).

Synthesis of azide-terminated poly(epichlorohydrin), pECH (2). Epichlorohydrin was polymerized according to the procedure developed by Carlotti. Briefly, a flame-dried Schlenk flask (10mL) equipped with a magnetic stirrer and fitted with PTFE stopcock was charged with tetrabutylammonium azide ($TBAN_3$, 40 mg, 0.074 mmol,0.004 equiv.) under argon. A solution of freshly distilled epichlorohydrin (1.29 mL, 16.5 mmol) in anhydrous toluene (4 mL). A solution of triisobutylaluminum in toluene (1.07 M, 200 µL, 0.222 mmol, 0.014 equiv.) was added via a syringe under argon at −30 °C. The reaction was stirred for 4 hours and then stopped by the addition of ethanol. The resulting pECH polymer **2** was precipitated into hexanes and dried under vacuum to yield a clear viscous oil (1500 mg, 99 % yield). The polymer was analyzed by SEC (0.2% LiBr in DMF): M_w = 24,000, M_n = 20,000, Đ = 1.20).

Synthesis of cholestanone-terminated poly(epichlorohydrin) (44). In a flame-dried Schlenk flask (10 mL), pECH polymer **1** (7.5 mg, 0.3 µmol) was dissolved in degassed anhydrous DMSO (200 µL). Alkynyl cholestanone **43** (1.7 mg, 3.8 µmol, 11.0 equiv.) was added, followed by CuI (~0.05mg, 0.3 µmol, 1.0 equiv.) and one drop diisopropylethyl amine (DIPEA, ~ 5 µL). The reaction was stirred at 40°C for 12 hr, at which time it was quenched by the addition of water to precipitate the polymer. The resultant polymer was triturated with hexanes to

remove unreacted **2** and dried on vacuum to yield a clear viscous oil in (8 mg, 100% yield). The polymer was analyzed by SEC (0.2% LiBr in DMF): M_w =24,000, M_n =20,000, Đ = 1.20.

Synthesis of cholestanone-terminated poly(glycidyl azide), pGA (45). The chloride to azide exchange in pECH polymer **44** was accomplished according to previously published procedure.[22] Briefly, in a flame-dried Schlenk flask (10 mL), polymer **44** (15 mg, 0.16 mmol) was dissolved in dry DMF (300 µL). To the solution was added NaN_3 (21 mg, 0.32 mmol, 2.0 equiv.), and the reaction was stirred at 60 °C for 3 days under argon to allow complete conversion. Polymer solution was filtered and precipitated in ethanol to yield a clear viscous oil (16 mg, 100 % yield). The polymer was analyzed by SEC (0.2% LiBr in DMF): M_w = 25,000, M_n = 21,000, Đ = 1.19.

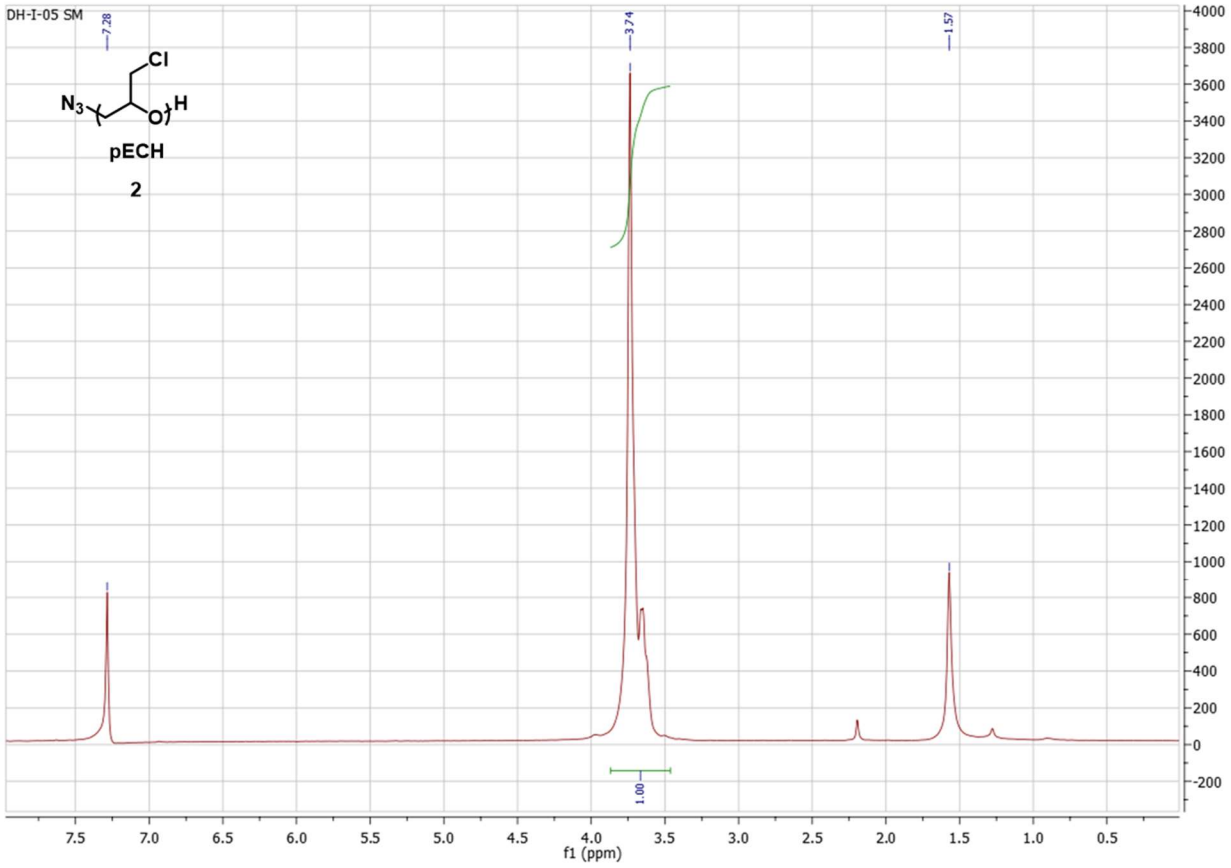
Synthesis of AF488-labeled glycopolymers 46-48. In a flame-dried Schlenk flask (10 mL), polymer **45** (9.00 mg, 0.09 mmol) was dissolved in degassed dry DMSO (250 µL). To the solution was added AF-488-alkyne (0.50 mg, 0.50 µmol) in DMSO (50 µL), followed by CuI (2.00 mg, 9.00 µmol) and DIPEA (16 uL, 0.09 mmol). The reaction was stirred in dark under Argon at 40 °C for 2 hrs. After this time, the reaction mixture was aliquoted (50 µL) into separate vials containing ß-propargyl glucoside, glucose **40**, cellobiose **41**, and panose **41** (0.03 mmol, 2.00 equiv. per azide side-chain) in degassed anhydrous DMSO (50 µL). The reactions were stirred in dark at 40°C overnight. After this time, the reactions were diluted with DI water and treated with Cuprisorb beads (SeaChem labs) for 18 hrs to sequester copper. The resulting copper-free solutions were filtered through celite to remove the resin and lyophilized. The dry residues were triturated 3x with methanol with monitoring by TLC to remove excess unreacted glycosides. The resulting AF-488-labeled glycopolymers **5** were dissolved in D_2O and lyophilized to give a pink solid in a quantitative yield for each. The polymers **5** were characterized using ^1H NMR (D_2O, 300 MHz) and UV-Vis (λ_{max} = 490 nm)

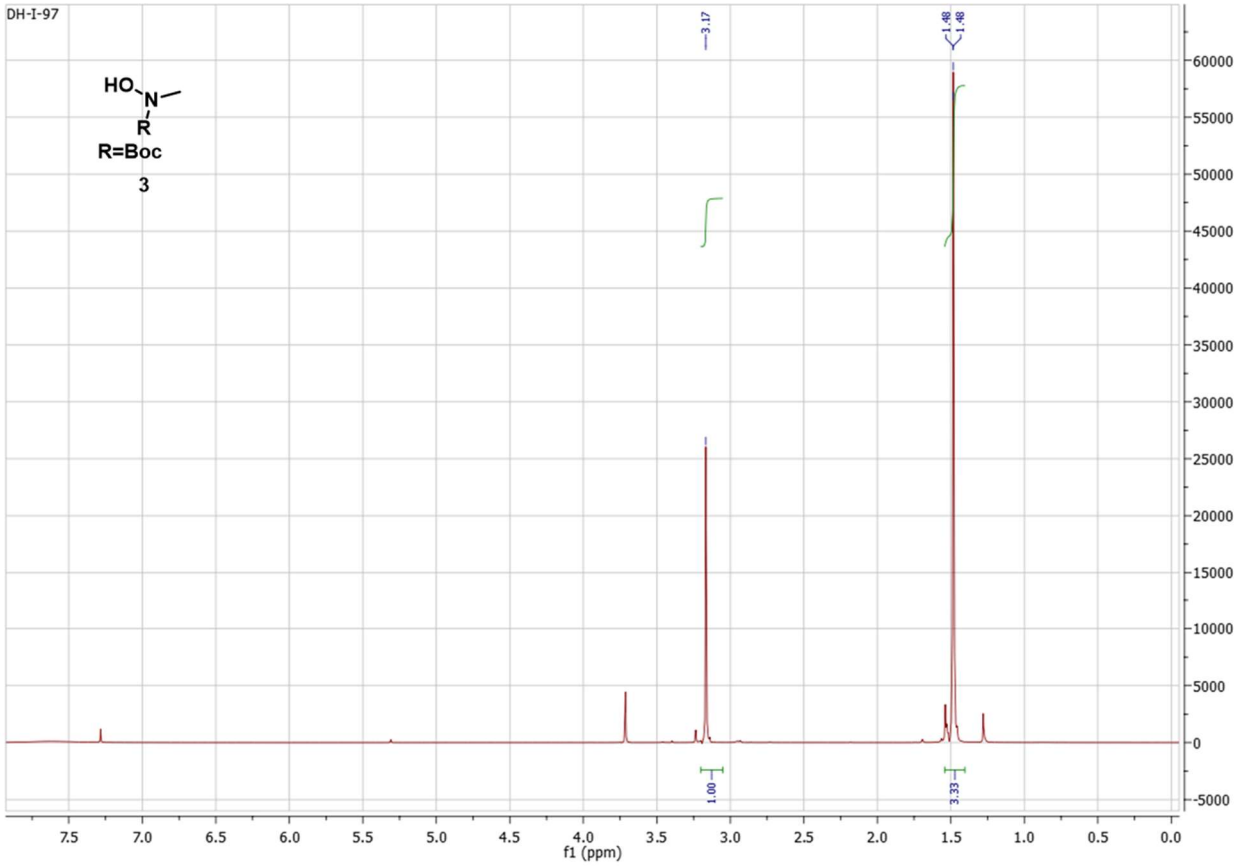
spectroscopy. Absorbance readings at known concentrations of glycopolymers **46-48** indicated the presence of ~ 0.53-0.83 AF-488 molecules per polymer chain (~0.25% sidechain occupancy).

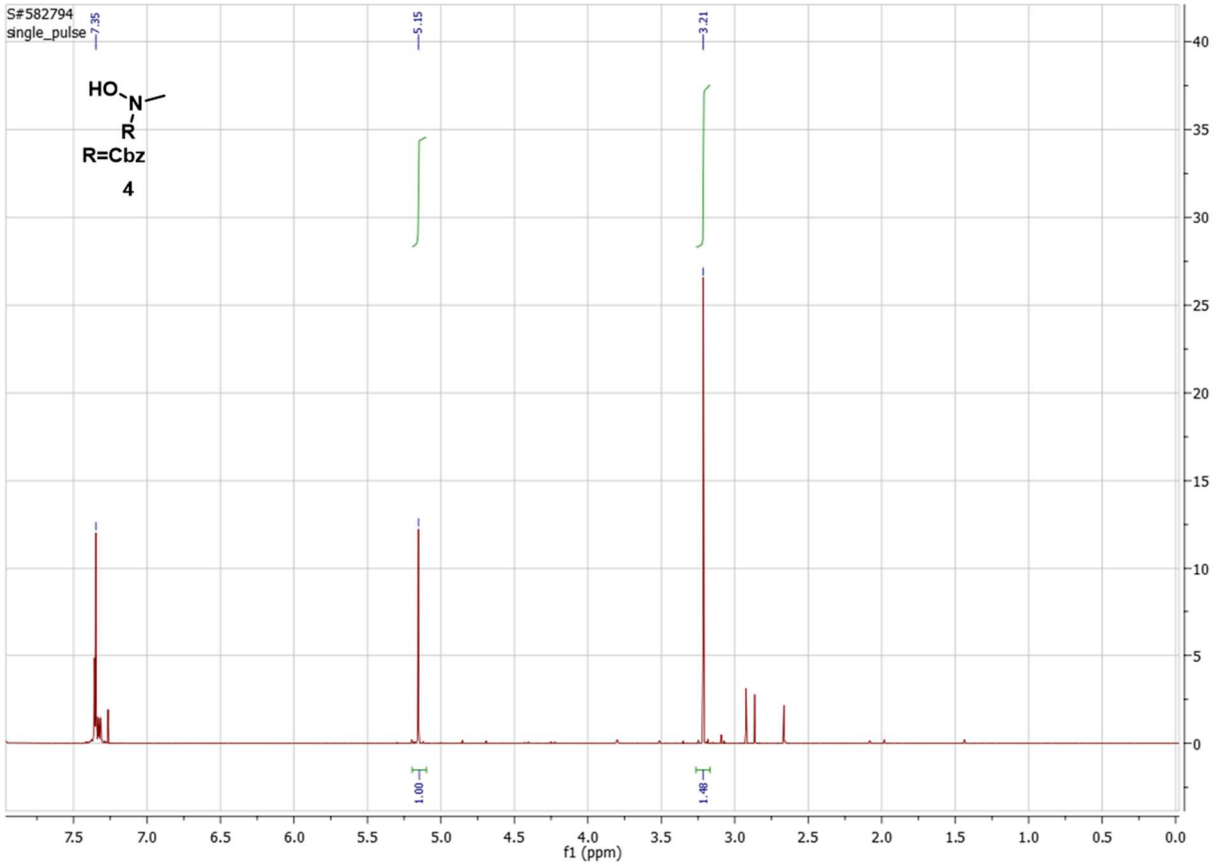
Section 4.8 Acknowledgements

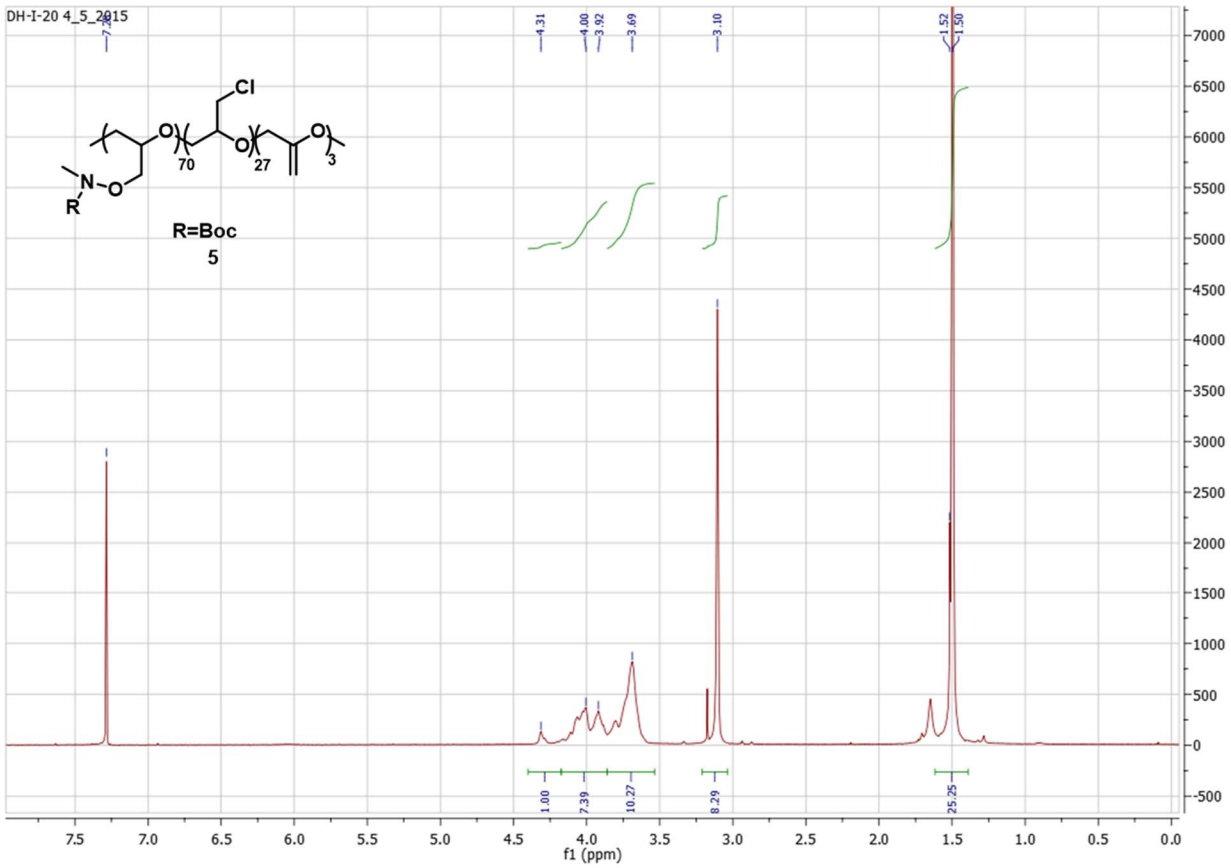
Chapter 4, in part, is unpublished work, and was contributed to by Michelle Zhang through synthesis and scale up of reaction intermediates and products including propargylated glycans and the amino-oxy to alkyne linker. Daniel Honigfort is the primary author of this material.

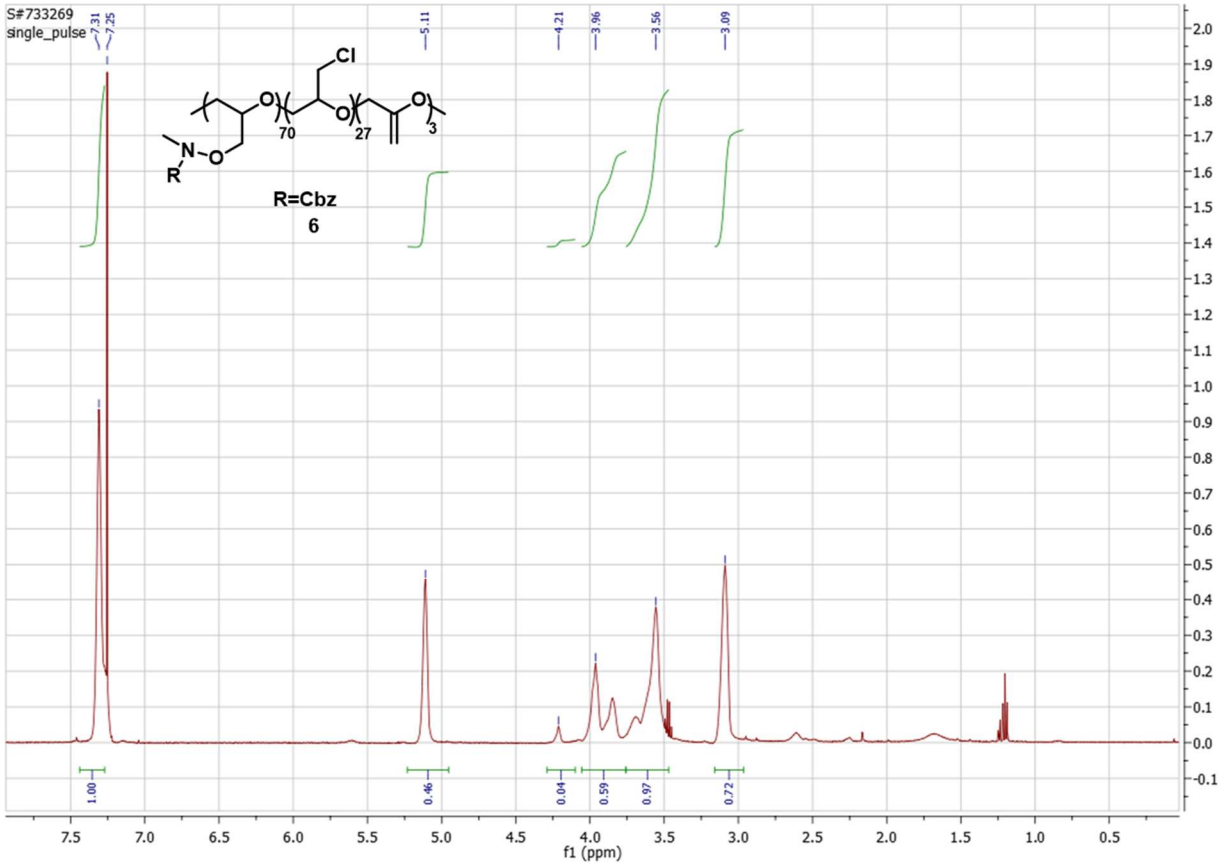
Section 4.9 NMR appendix:

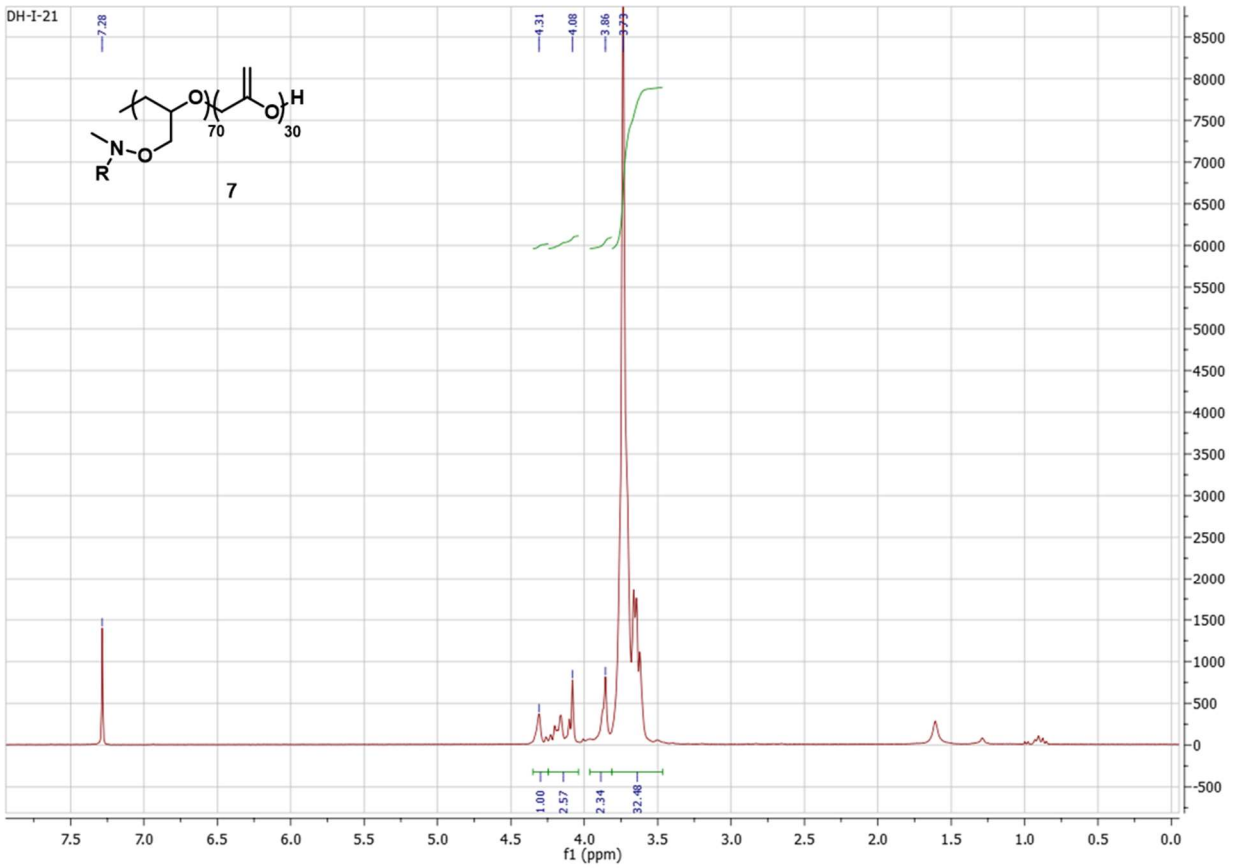


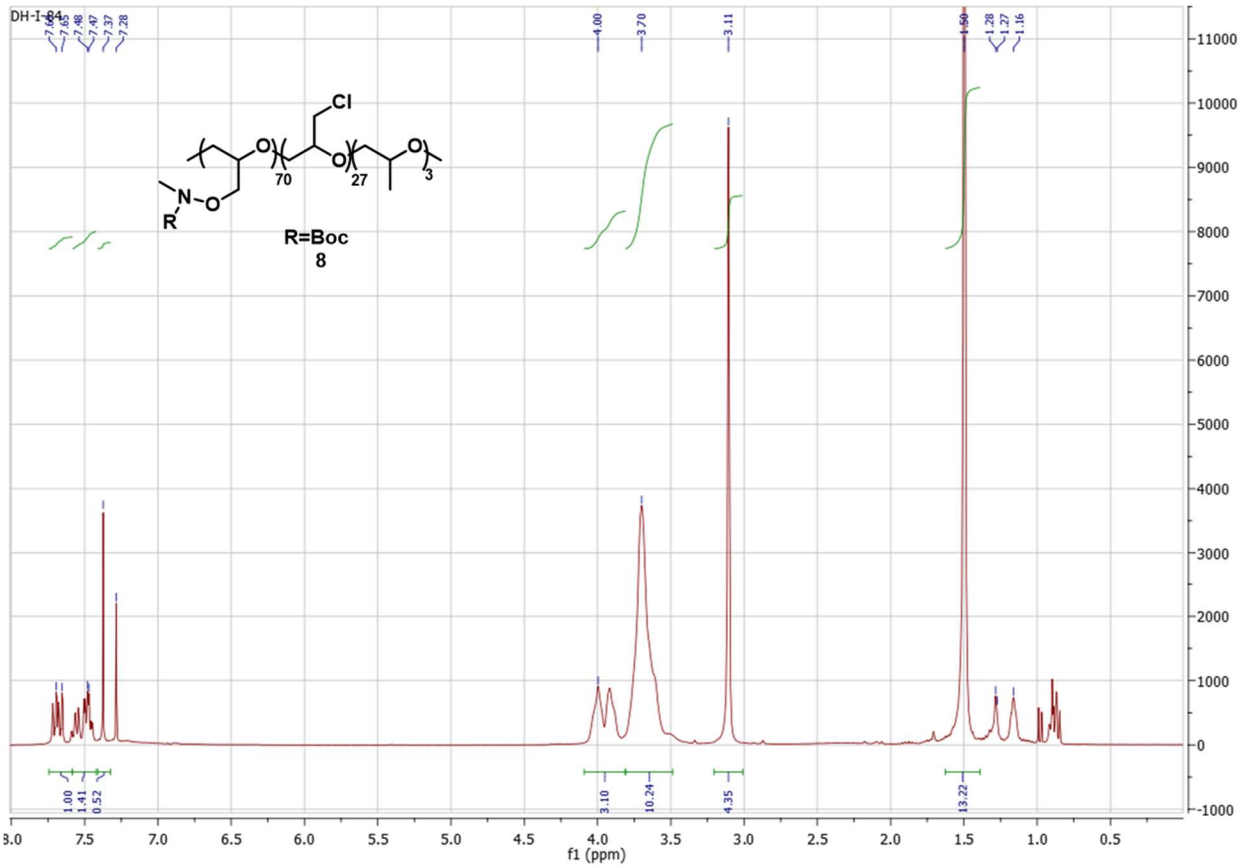

DH-I-97
HO
N
R
R=Boc
3
3.17
1.48
1.48
1.00
3.33
f1 (ppm)

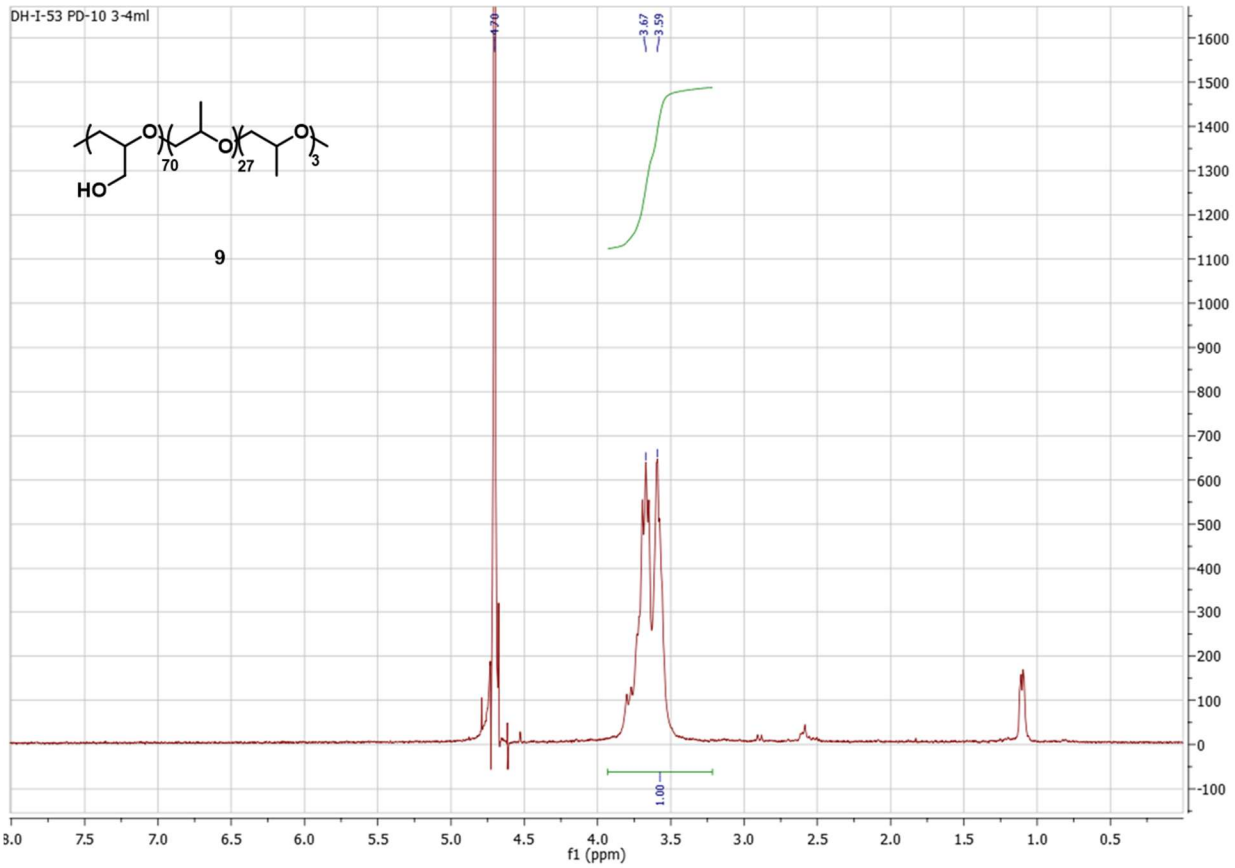

S#582794
single_pulse
7.35
5.15
3.21
HO
N
R
R=Cbz
4
1.00
1.48
f1 (ppm)

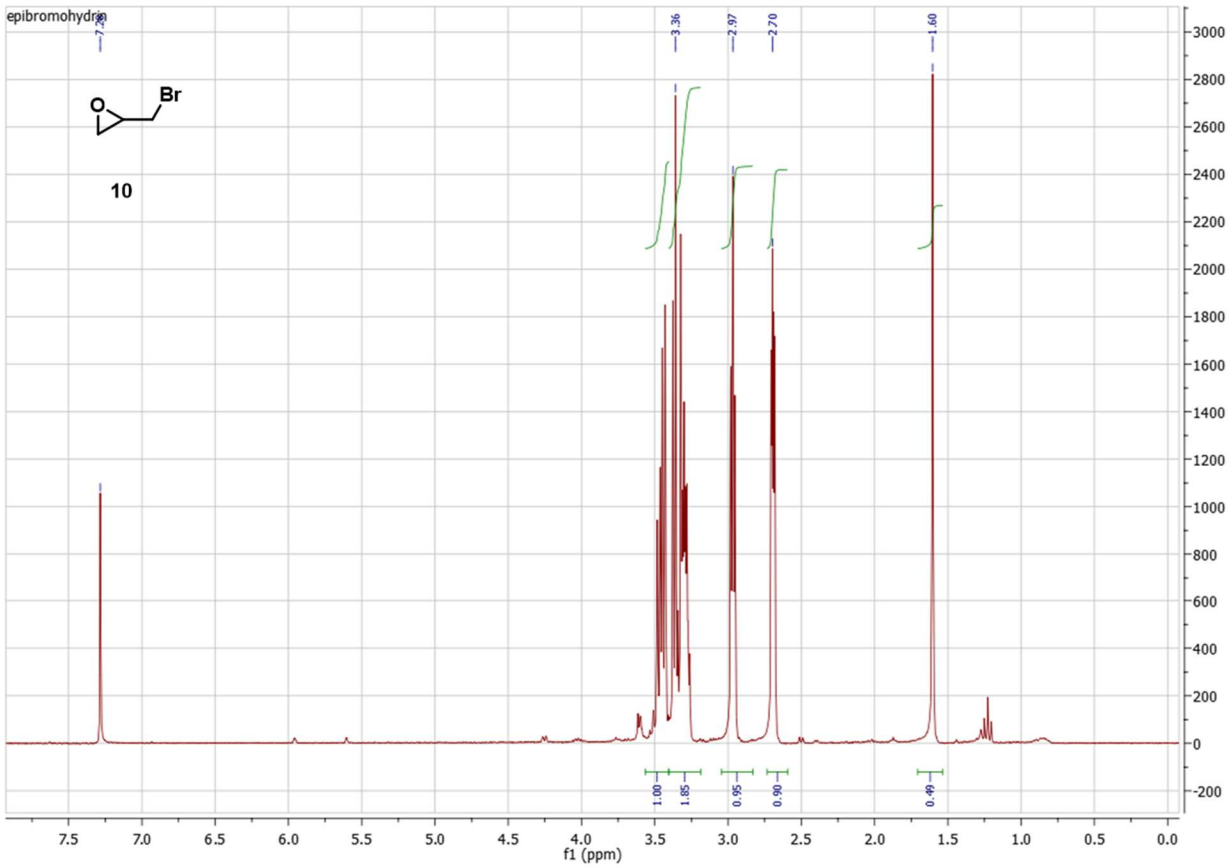

DH-I-20 4_5_2015
Cl
O
70
O
27
O
3
N
O
R
R=Boc
5
4.31
4.00
3.92
3.69
3.10
1.52
1.50
1.00
7.39
10.27
8.29
25.25
f1 (ppm)

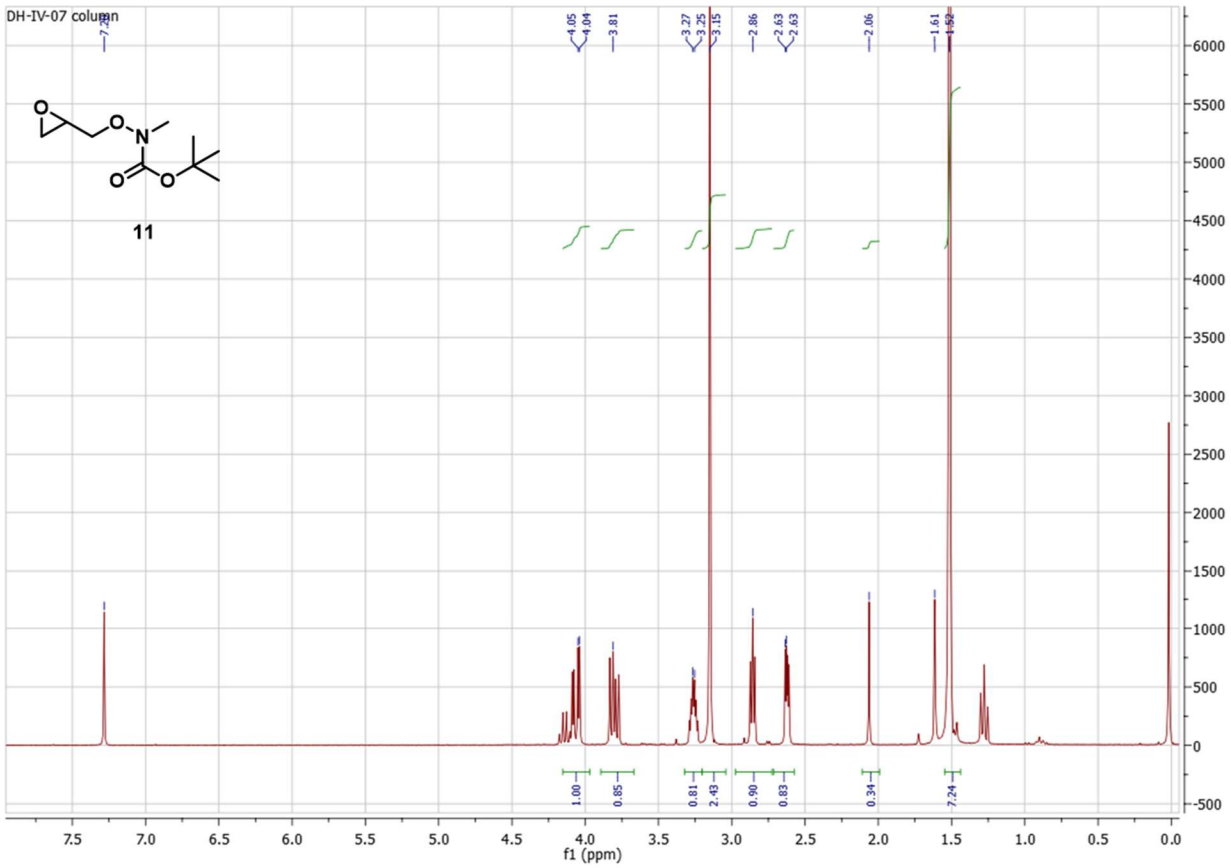

S#733269
single_pulse
7.31
7.25
5.11
4.21
3.96
3.56
3.09
Cl
O
70
O
27
O
3
N
O
R
R=Cbz
6
1.00
0.46
0.04
0.59
0.97
0.72
f1 (ppm)

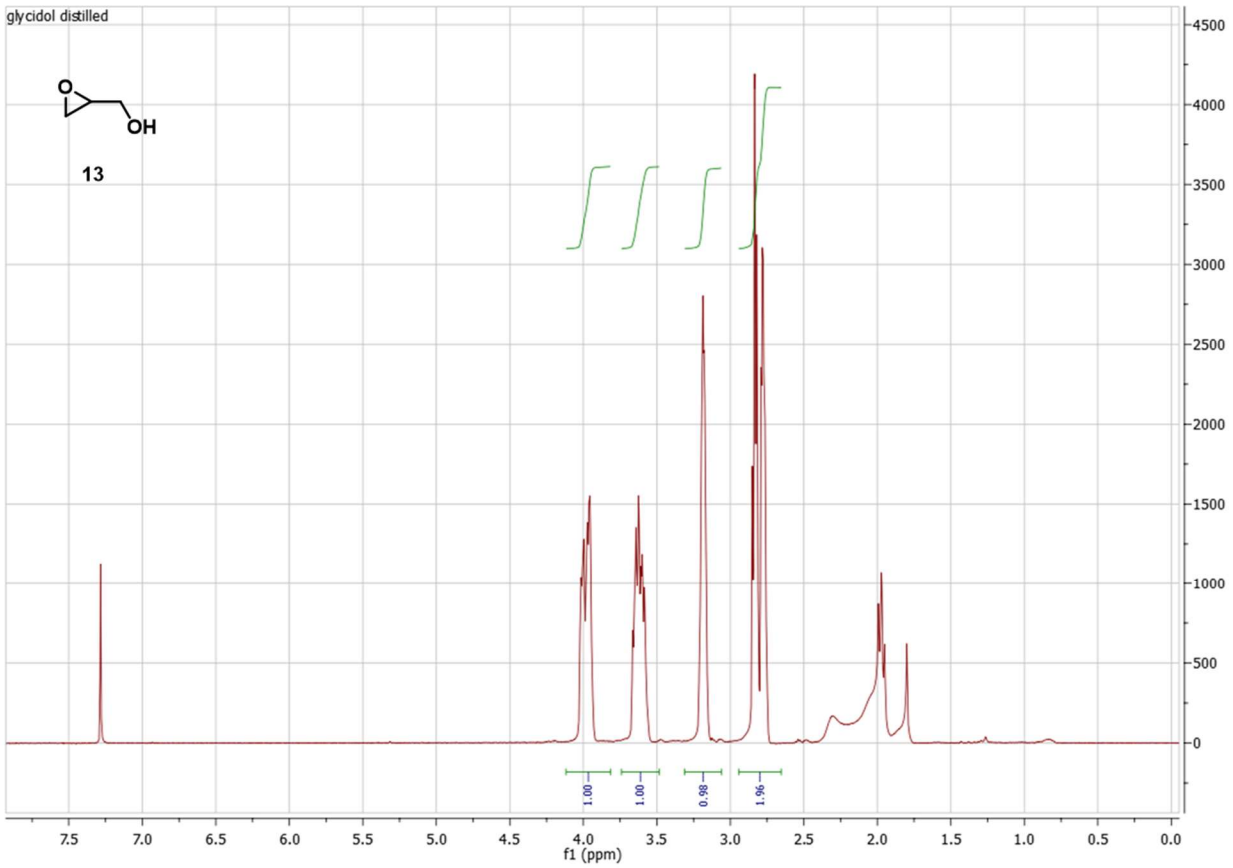

DH-I-21
7.28
4.31
4.08
3.86
3.73
O
70
O
30
H
N
O
R
7
1.00
2.57
2.34
32.48
f1 (ppm)

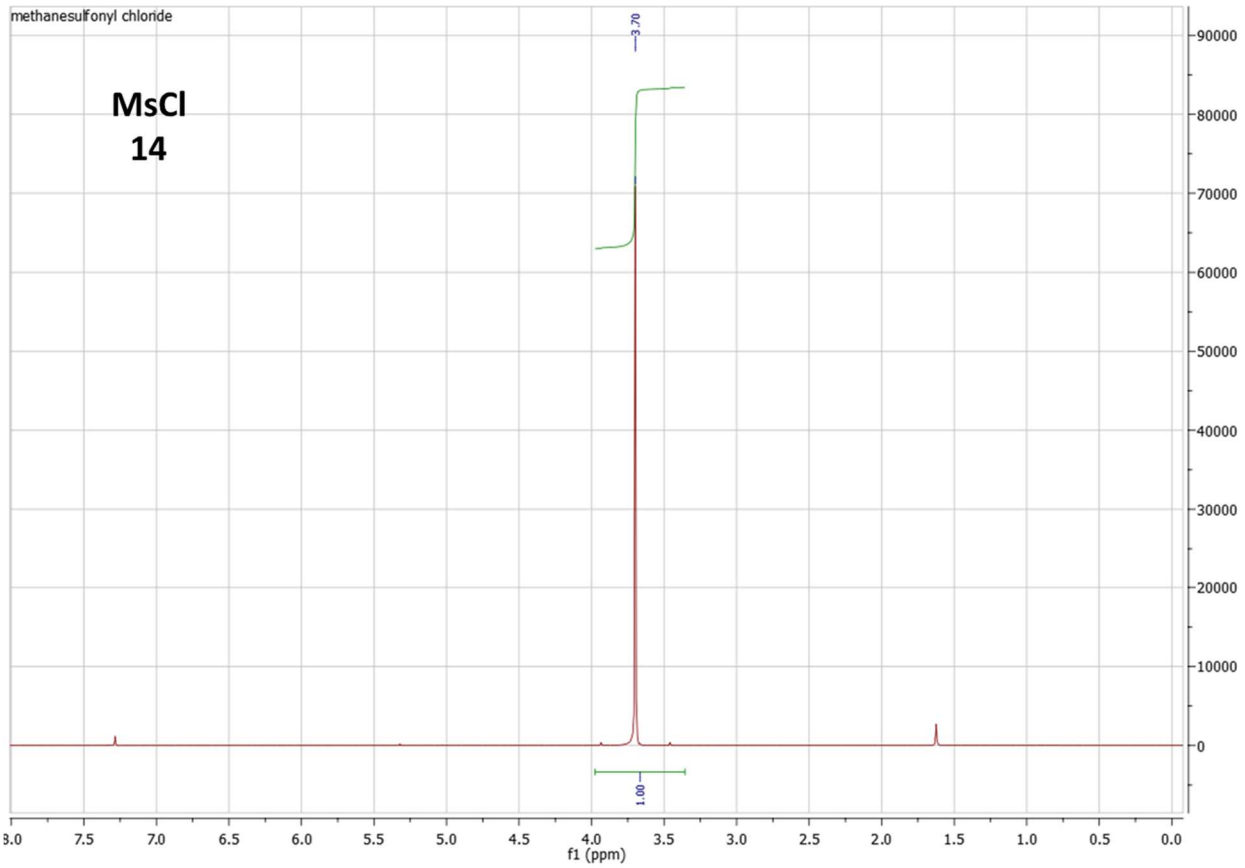

DH-I-84
7.65
7.65
7.48
7.47
7.37
7.28
4.00
3.70
3.11
1.50
1.28
1.27
1.16
Cl
O
70
O
27
O
3
N
O
R
R=Boc
8
1.00
1.41
0.52
3.10
10.24
4.35
13.22
f1 (ppm)

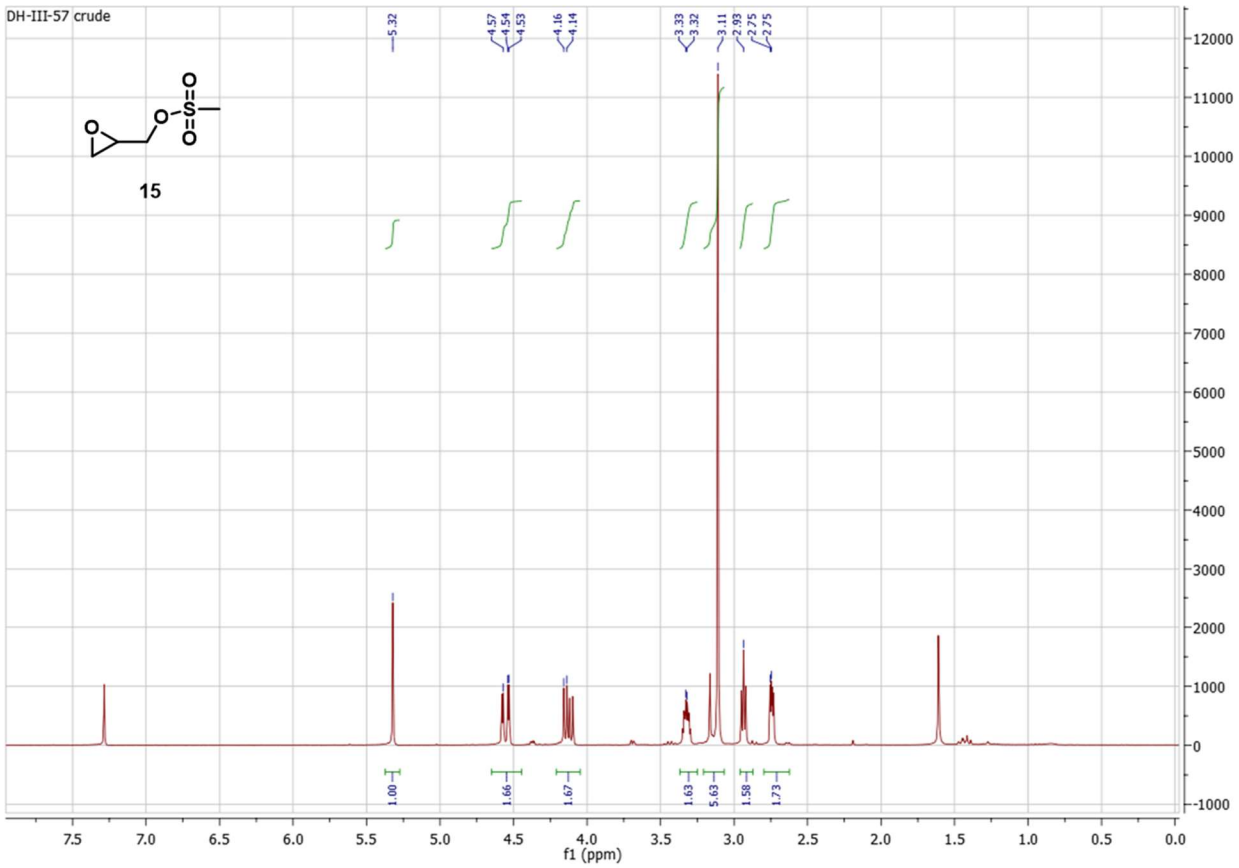

DH-I-53 PD-10 3-4ml
HO
70
27
3
9
4.70
3.67
3.59
1.00
f1 (ppm)

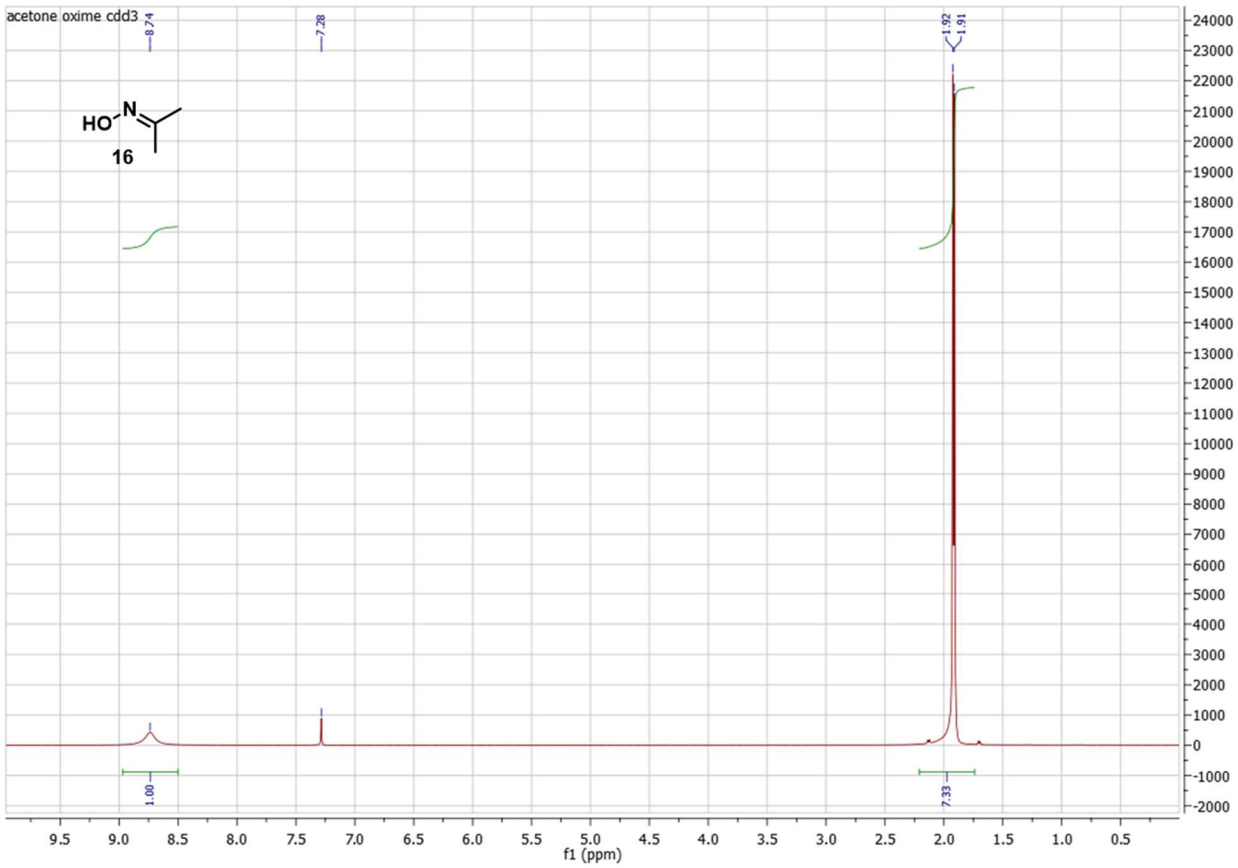

epibromohydrin
7.26
3.36
2.97
2.70
1.60
O
Br
10
1.00
1.85
0.95
0.90
0.49
f1 (ppm)

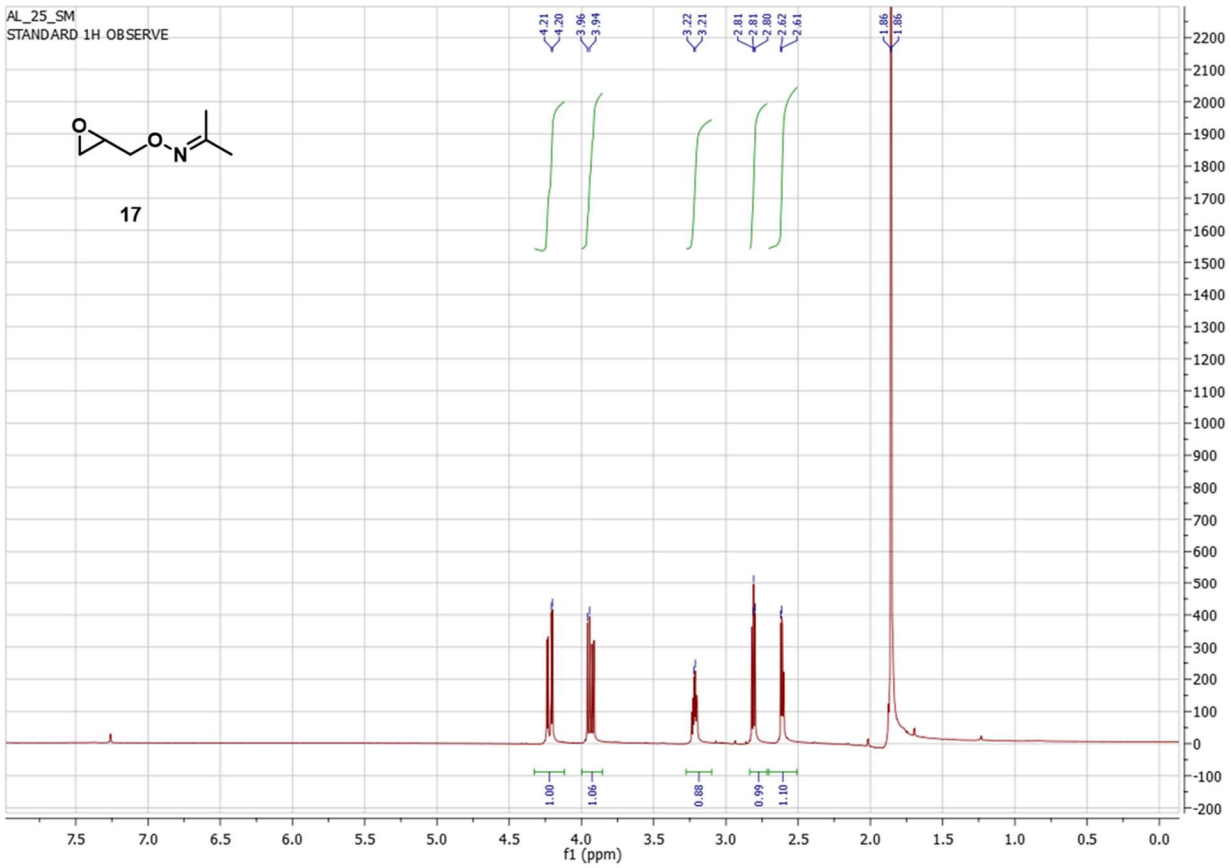

DH-IV-07 column
11
f1 (ppm)

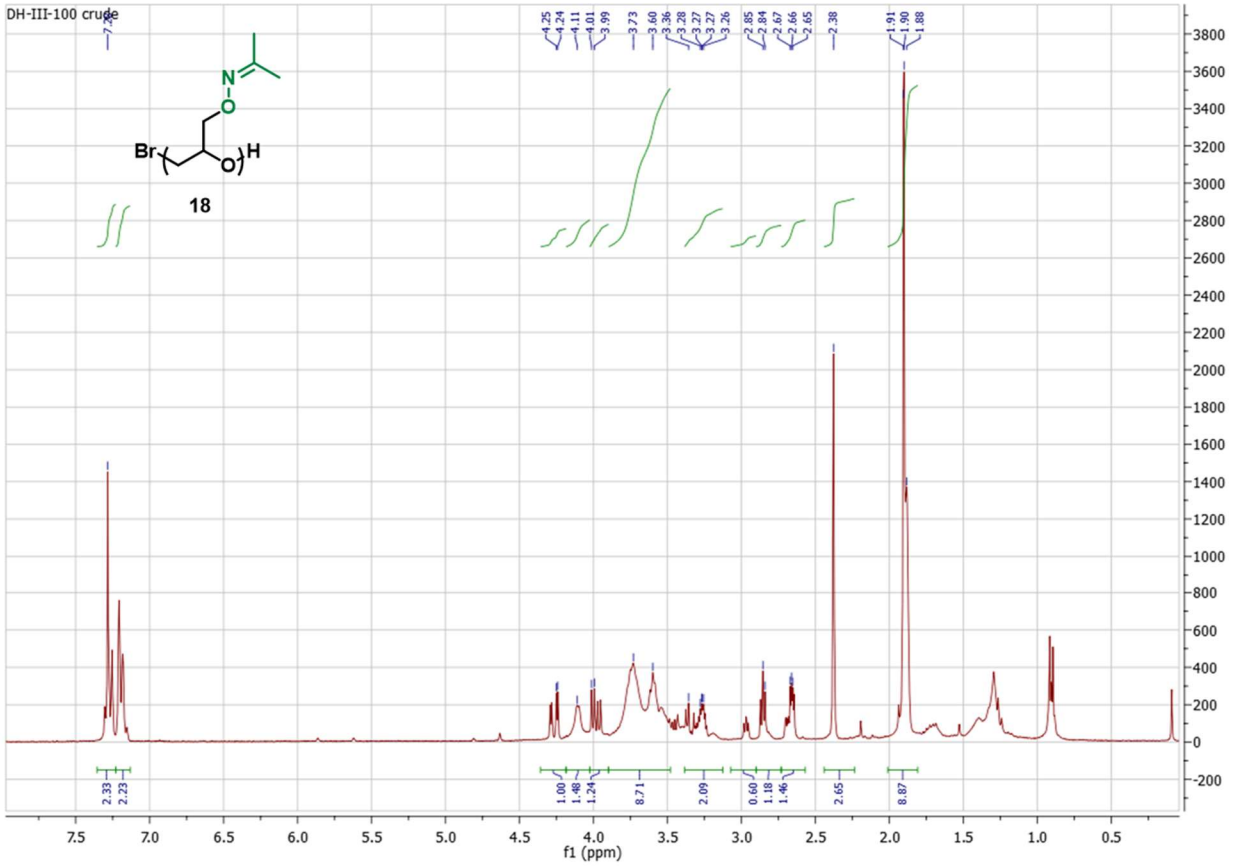

glycidol distilled
O
OH
13
4500
4000
3500
3000
2500
2000
1500
1000
500
0
1.00
1.00
0.98
1.96
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

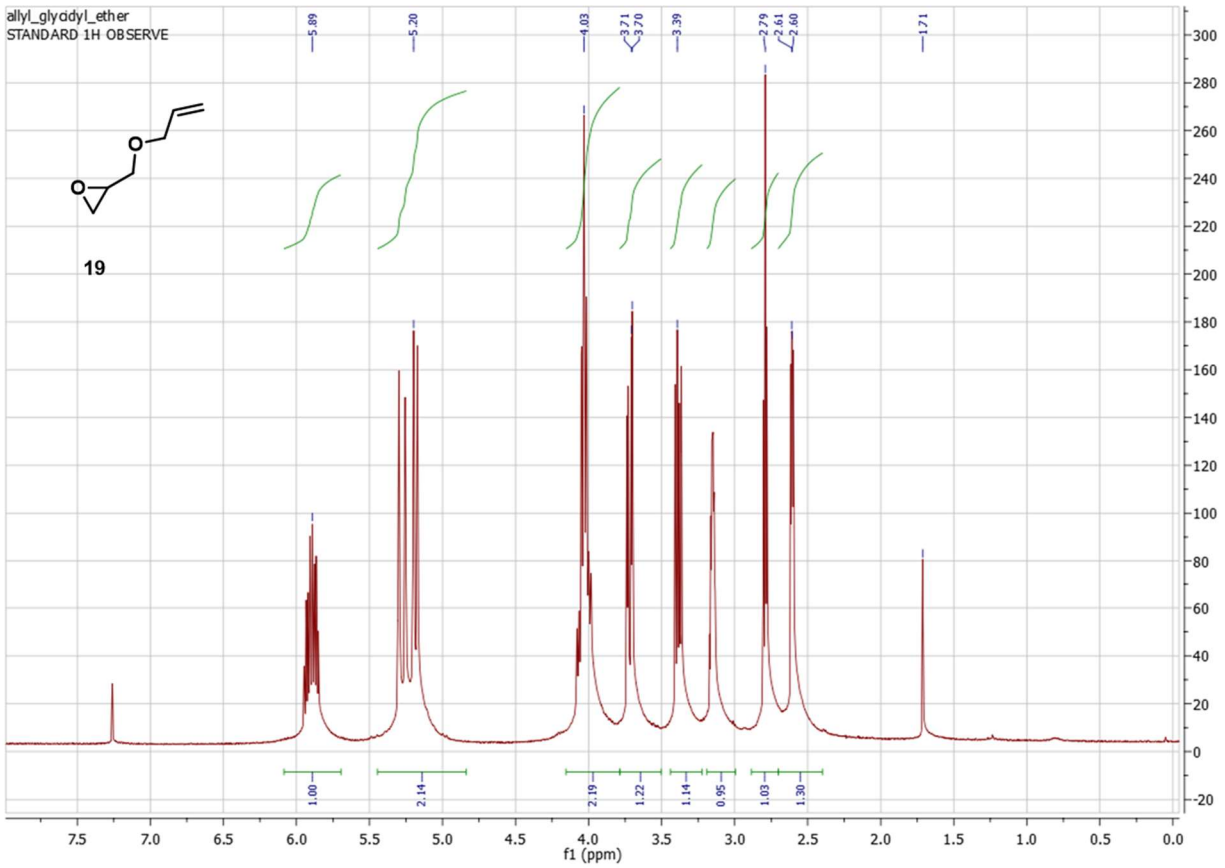

methanesulfonyl chloride
MsCl
14
3.70
1.00
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

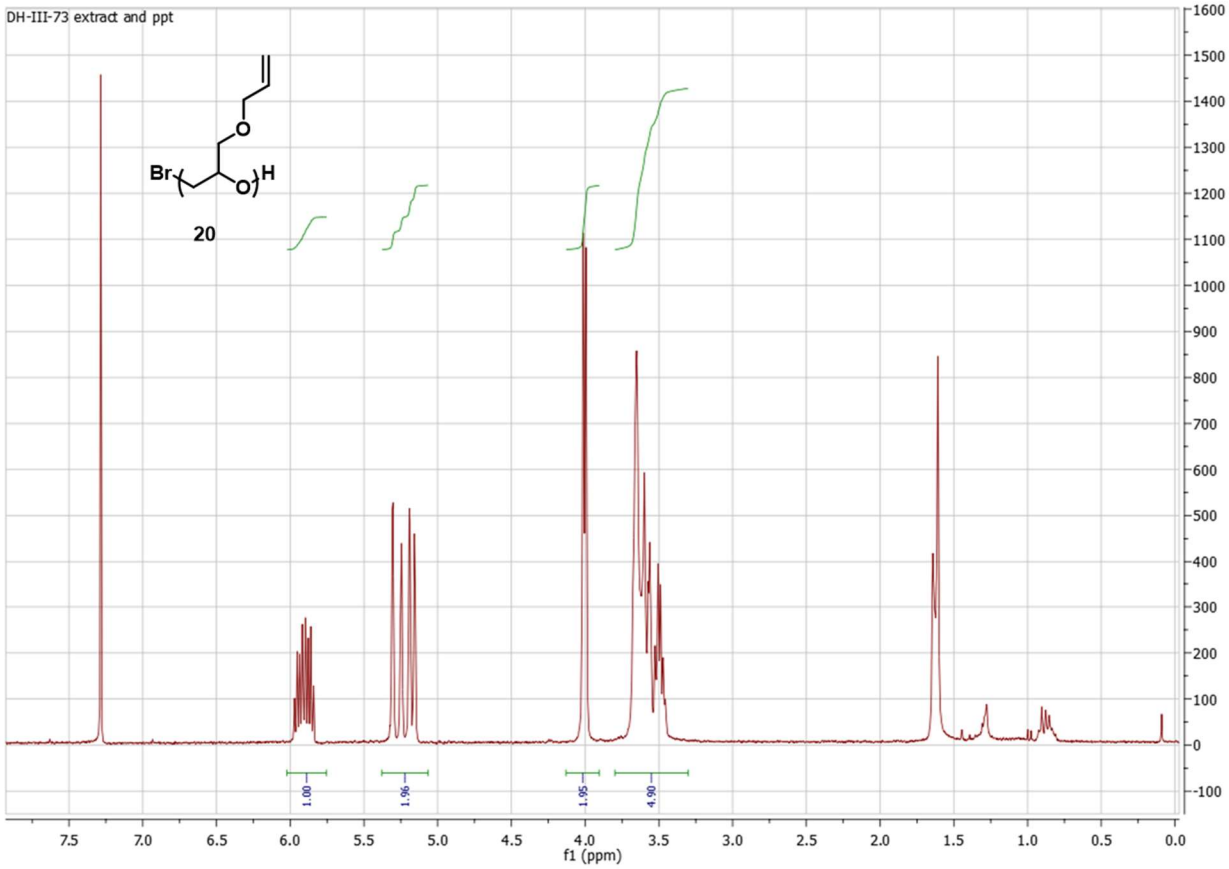

DH-III-57 crude
15
5.32
4.57
4.54
4.53
4.16
4.14
3.33
3.32
3.11
2.93
2.75
2.75
1.00
1.66
1.67
1.63
5.63
1.58
1.73
f1 (ppm)

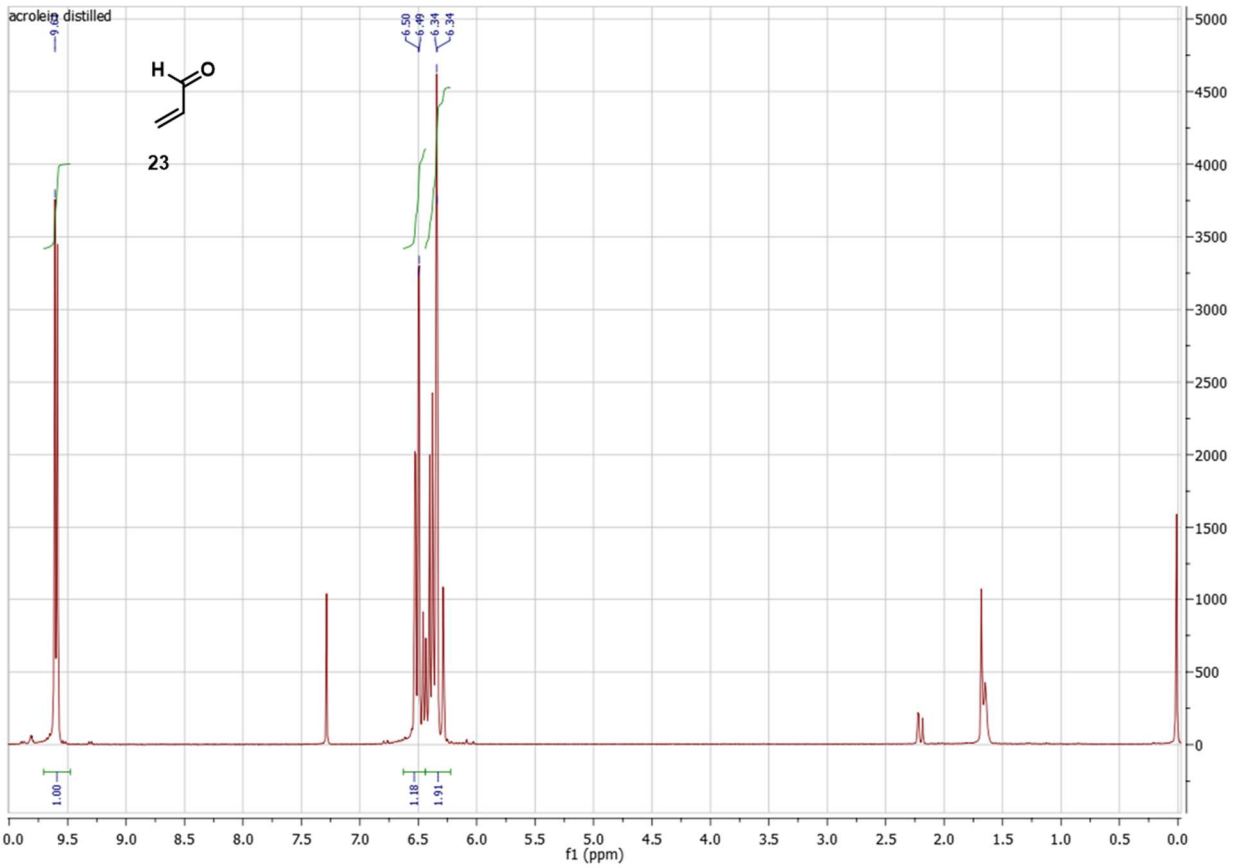

acetone oxime cdd3
8.74
7.28
1.92
1.91
HO
N
16
1.00
7.33
f1 (ppm)

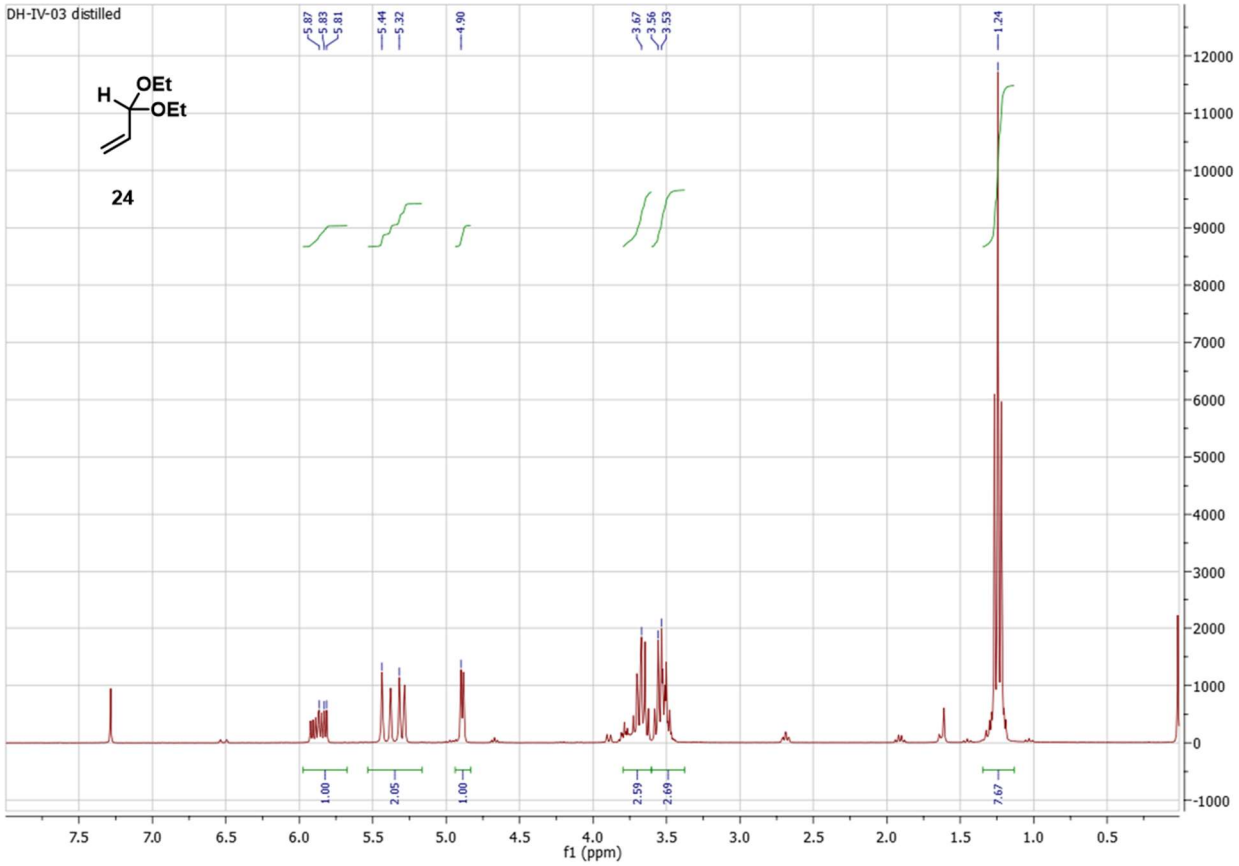

AL_25_SM
STANDARD 1H OBSERVE
17
4.21
4.20
3.96
3.94
3.22
3.21
2.81
2.81
2.80
2.62
2.61
1.86
1.86
1.00
1.06
0.88
0.99
1.10
f1 (ppm)

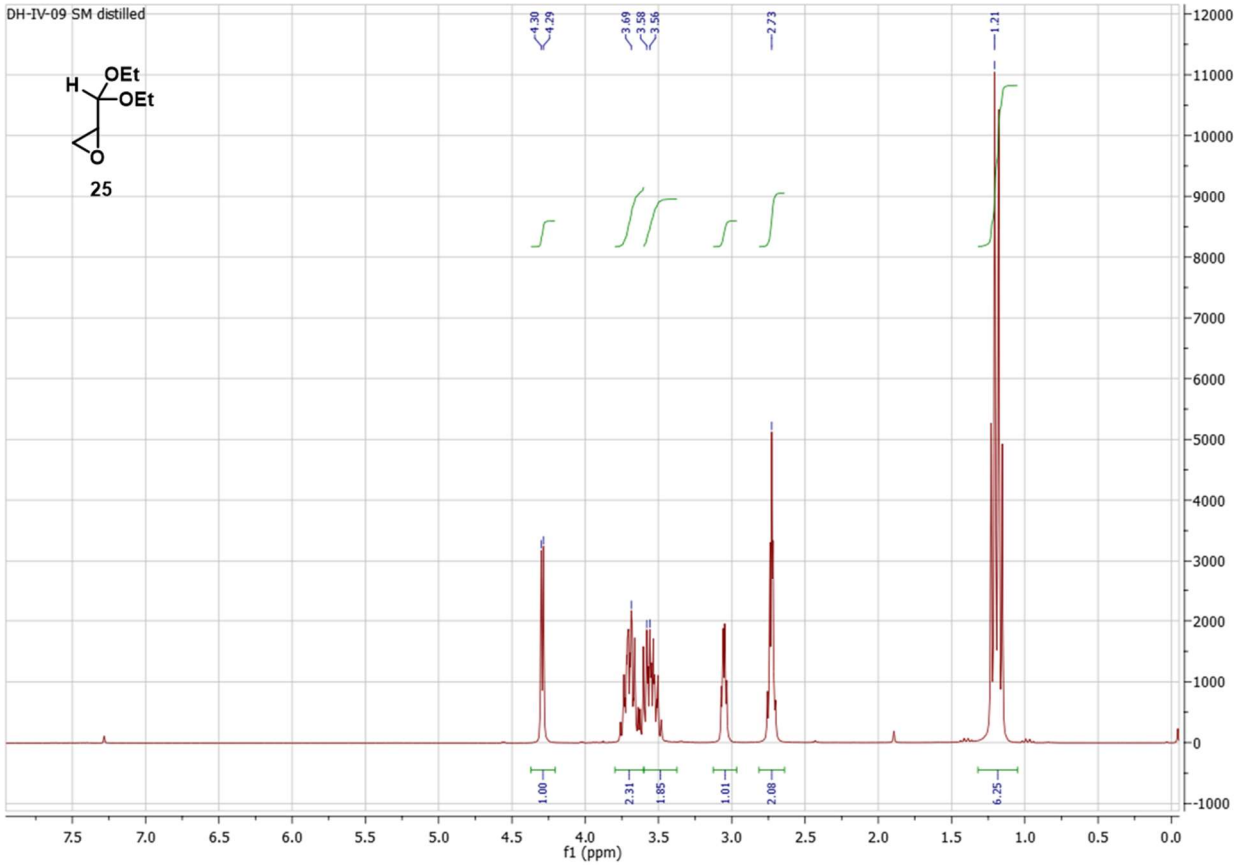

DH-III-100 crude
18
f1 (ppm)

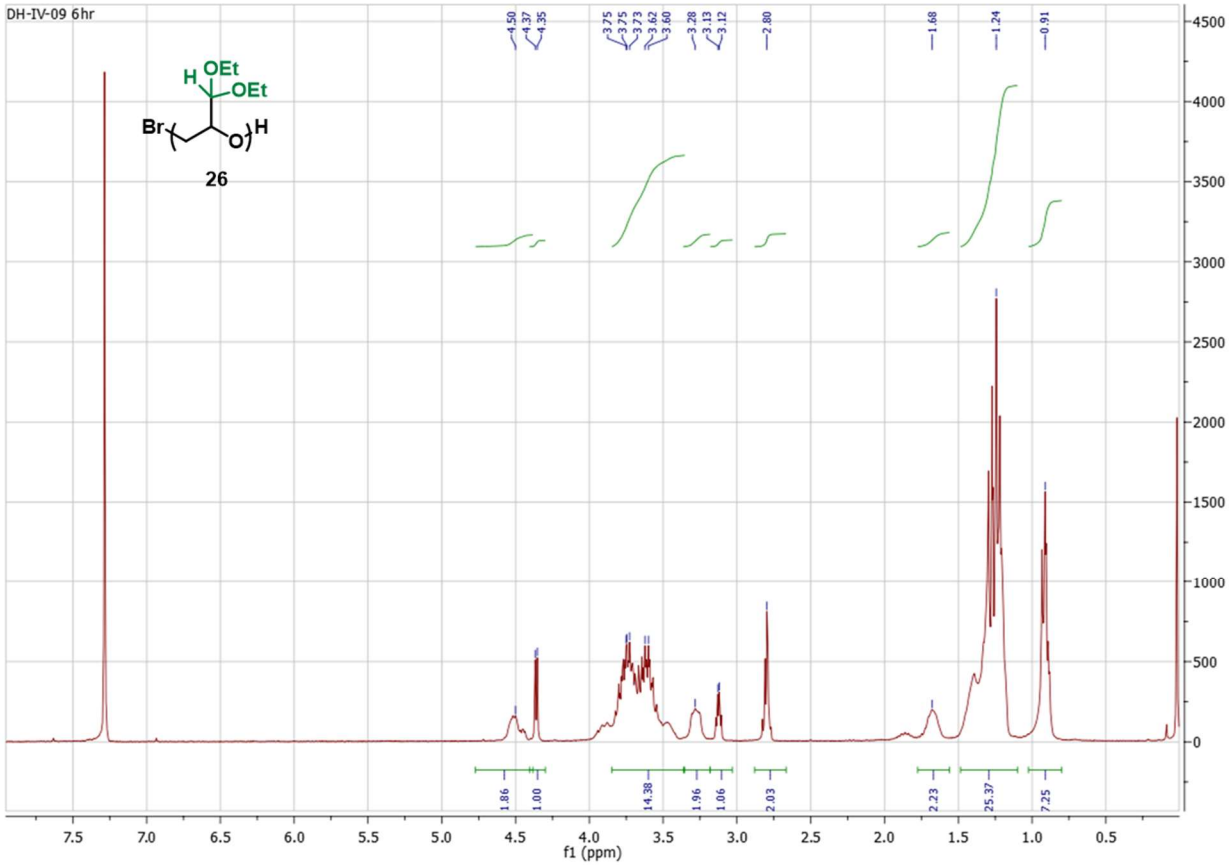

allyl_glycidyl_ether
STANDARD 1H OBSERVE
19
5.89
5.20
4.03
3.71
3.70
3.39
2.79
2.61
2.60
1.71
1.00
2.14
2.19
1.22
1.14
0.95
1.03
1.30
f1 (ppm)

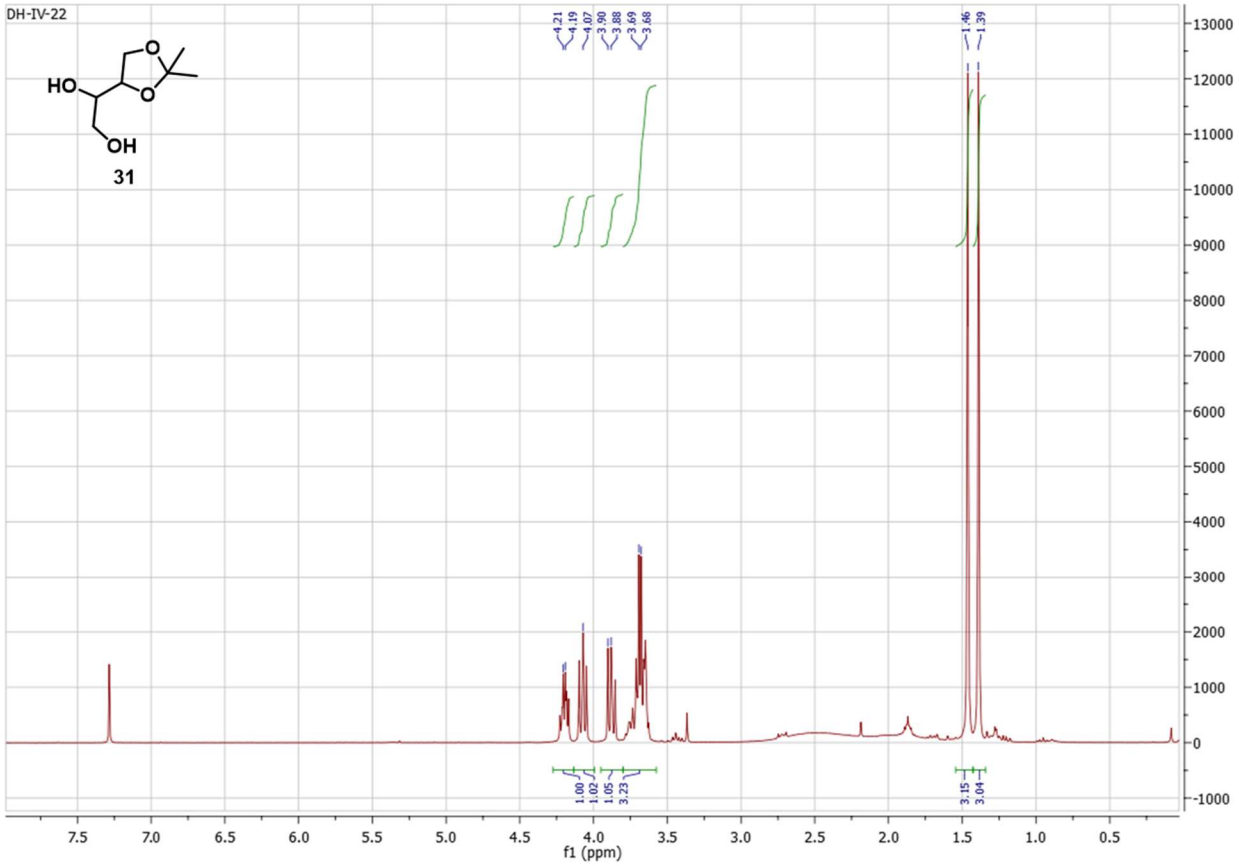

DH-III-73 extract and ppt
Br
O
O
H
20
1.00
1.96
1.95
4.90
f1 (ppm)

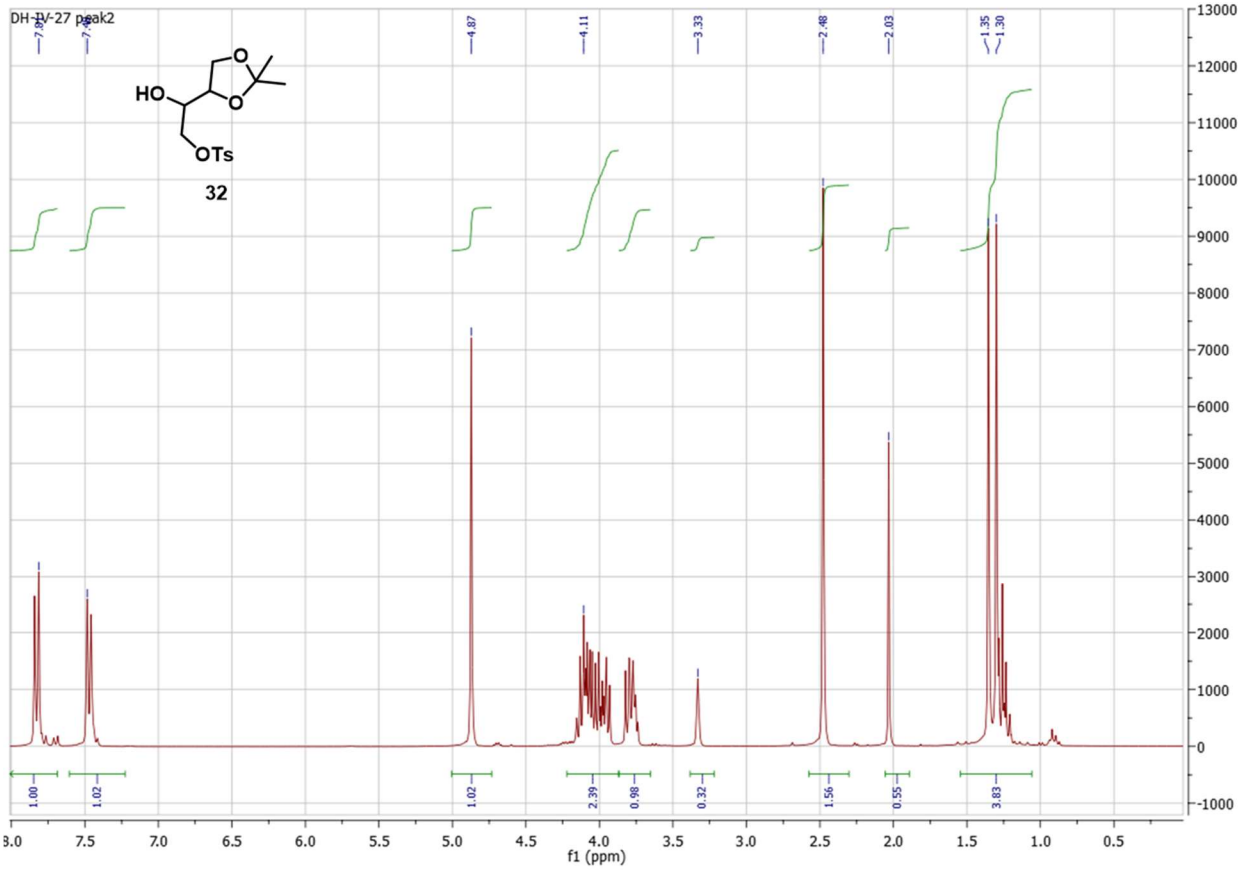

acrolein distilled
H
O
23
9.49
6.50
6.49
6.34
6.34
1.00
1.18
1.91
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
0.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

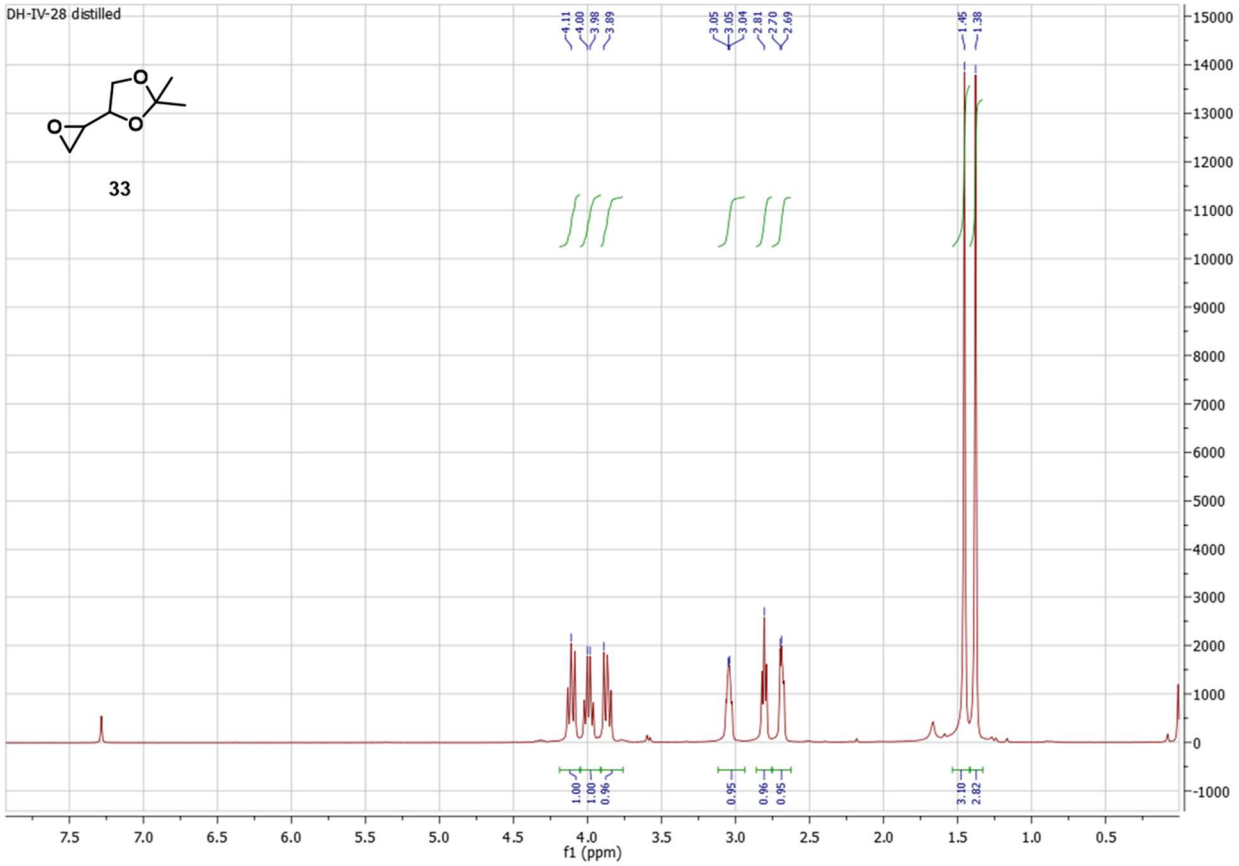

DH-IV-03 distilled
H
OEt
OEt
24
5.87
5.83
5.81
5.44
5.32
4.90
3.67
3.56
3.53
1.24
1.00
2.05
1.00
2.59
2.69
7.67
f1 (ppm)

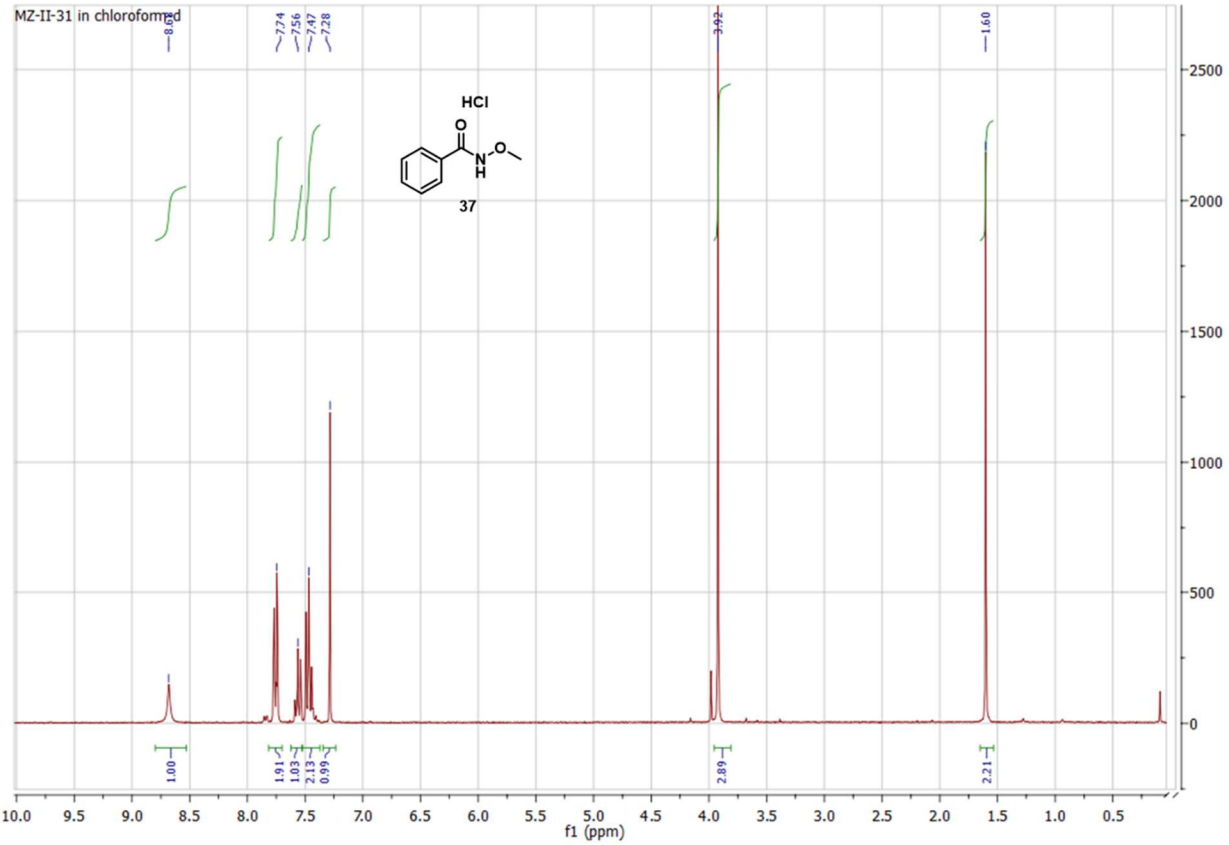

DH-IV-09 SM distilled
H
OEt
OEt
O
25
4.30
4.29
3.69
3.58
3.56
2.73
1.21
1.00
2.31
1.85
1.01
2.08
6.25
f1 (ppm)

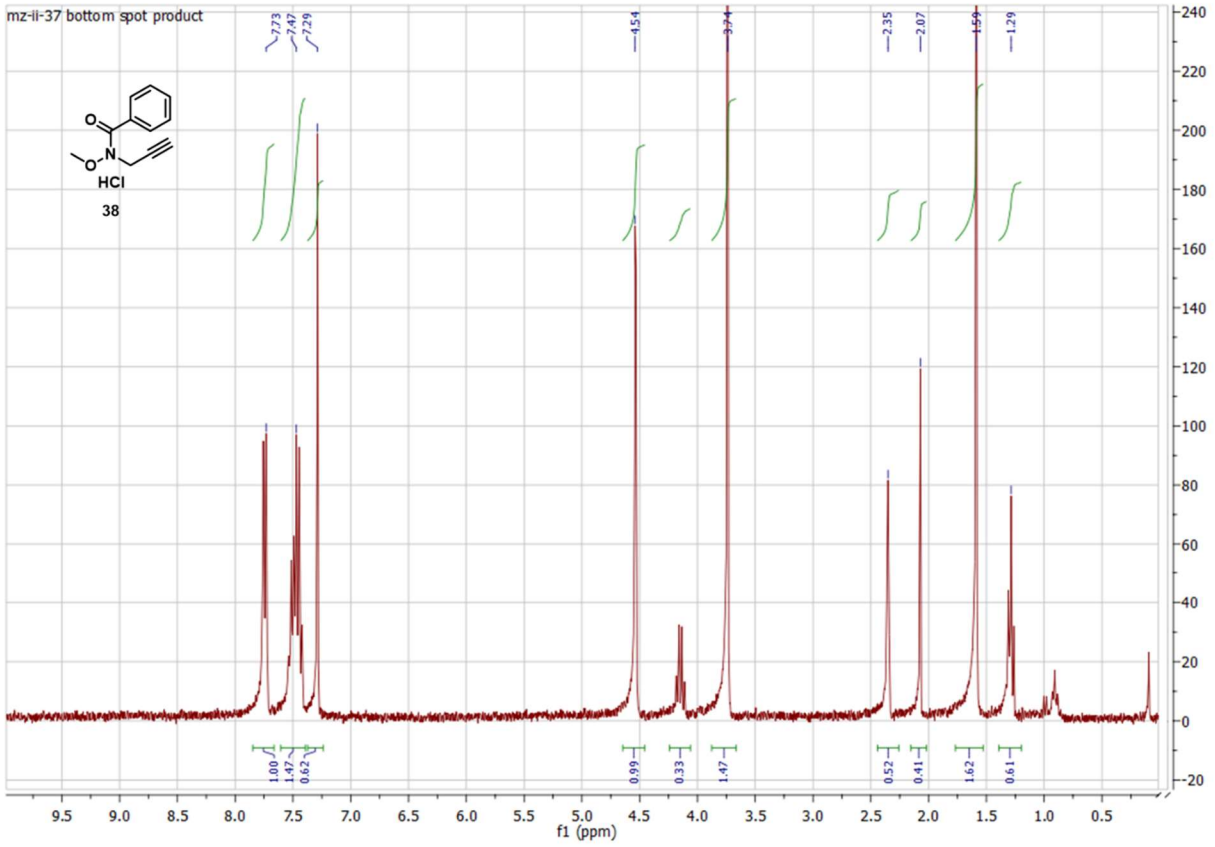

DH-IV-09 6hr
H
OEt
OEt
Br
O
H
26
4.50
4.37
4.35
3.75
3.75
3.73
3.62
3.60
3.28
3.13
3.12
2.80
1.68
1.24
0.91
1.86
1.00
14.38
1.96
1.06
2.03
2.23
25.37
7.25
f1 (ppm)

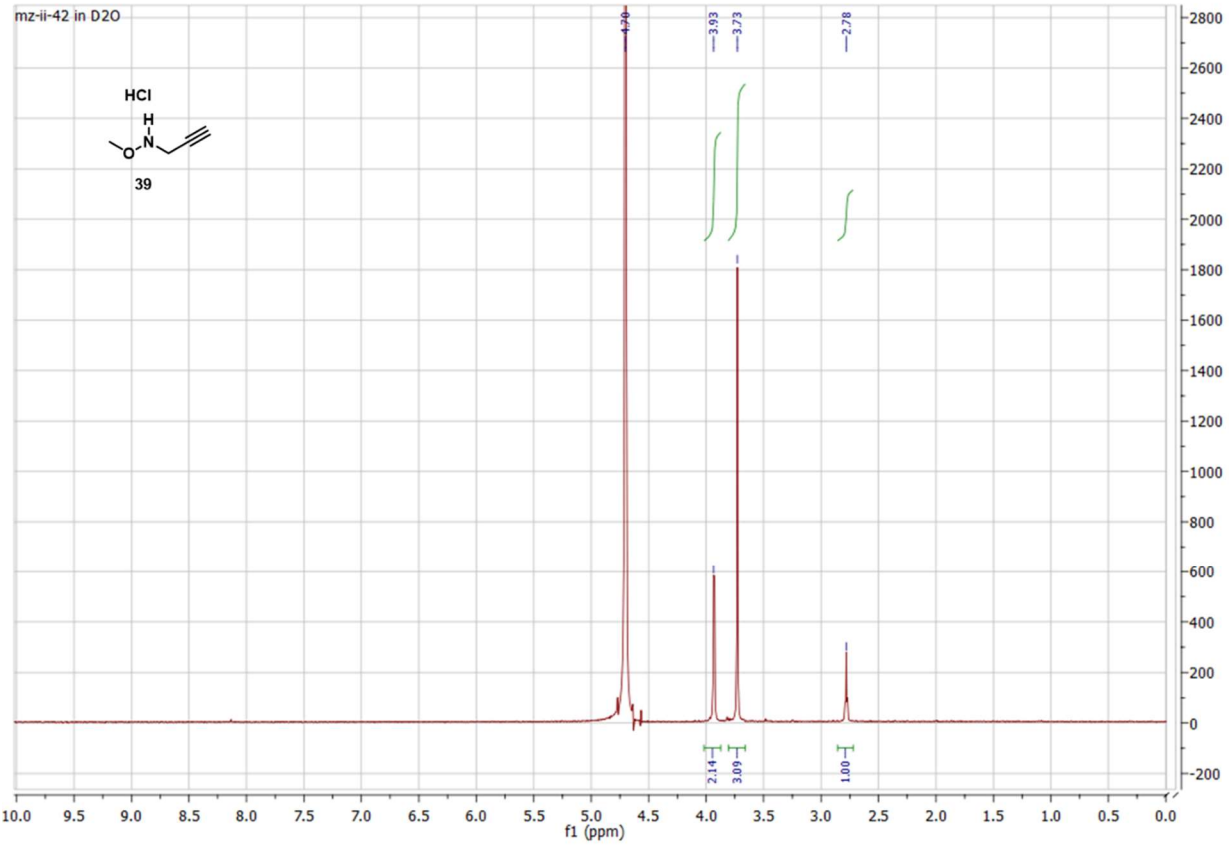

DH-IV-22
HO
O
O
OH
31
4.21
4.19
4.07
3.90
3.88
3.69
3.68
1.46
1.39
1.00
1.02
1.05
3.23
3.15
3.04
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)

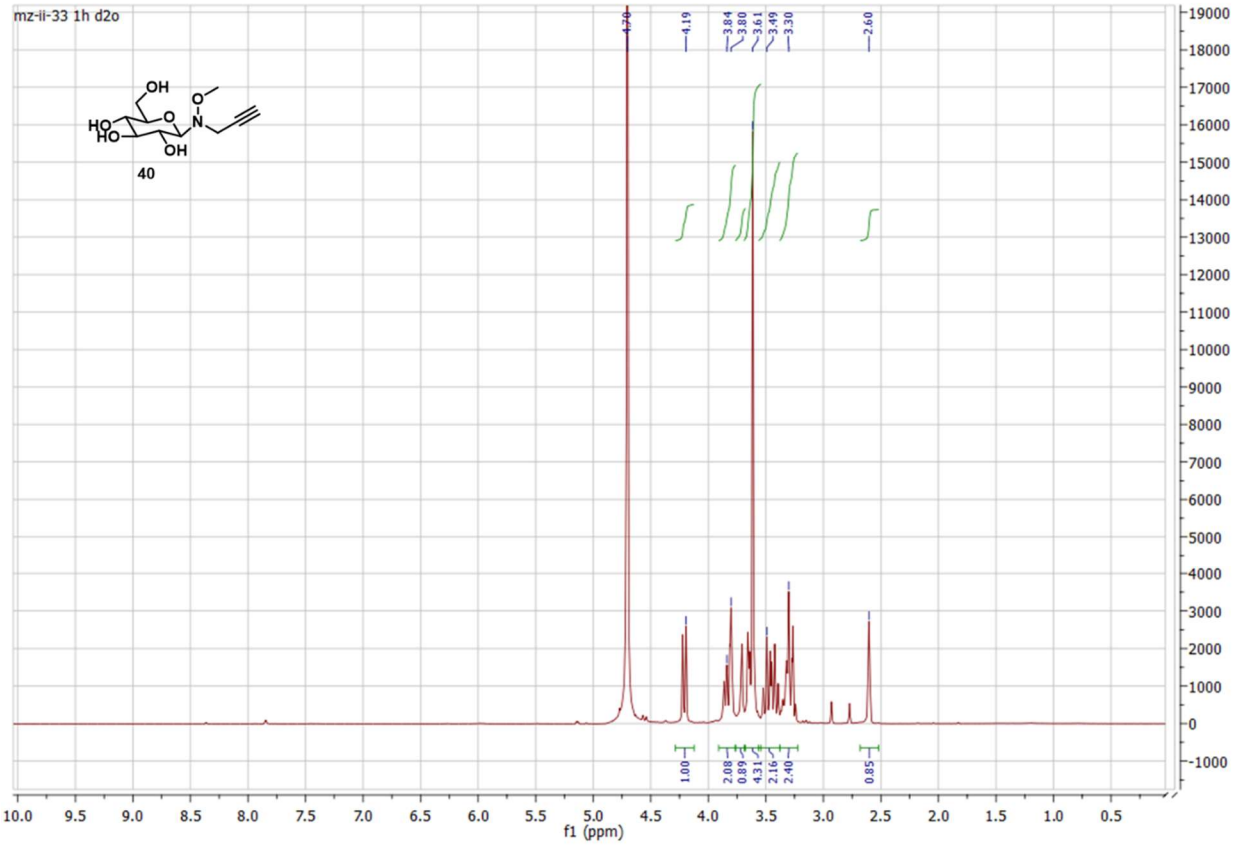

DH-IV-27 peak2
HO
OTs
O
O
32
7.50
7.47
4.87
4.11
3.33
2.48
2.03
1.35
1.30
1.00
1.02
1.02
2.39
0.98
0.32
1.56
0.55
3.83
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)

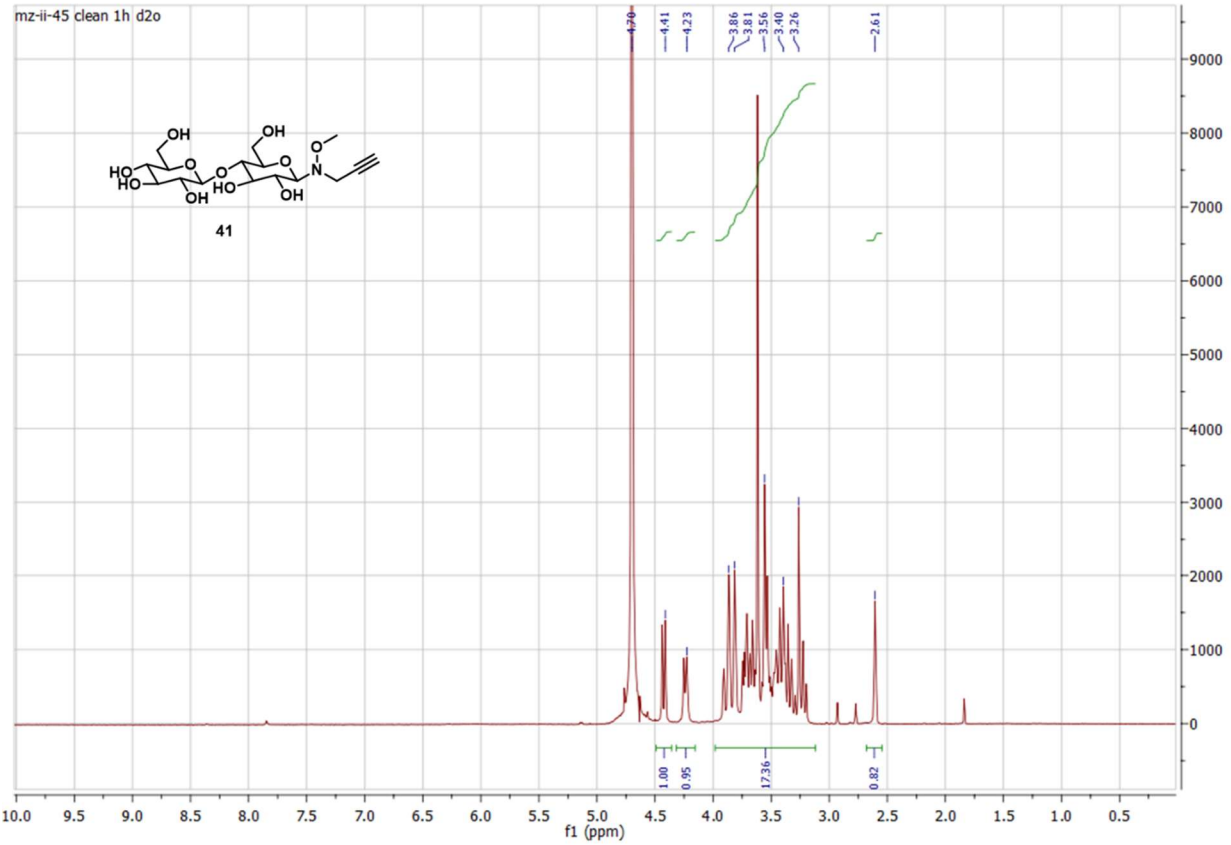

DH-IV-28 distilled
33
4.11
4.00
3.98
3.89
3.05
3.05
3.04
2.81
2.70
2.69
1.45
1.38
1.00
1.00
0.96
0.95
0.96
0.95
3.10
2.82
f1 (ppm)

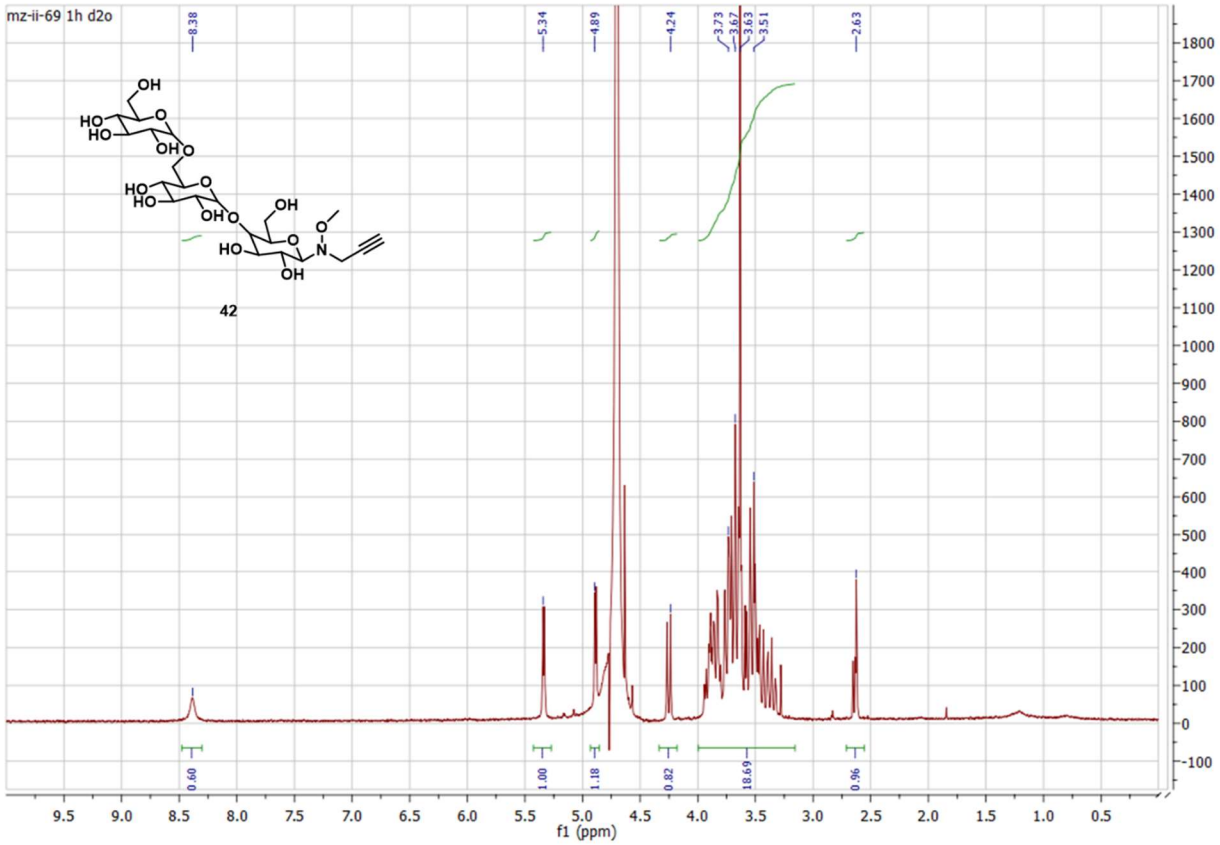

MZ-II-31 in chloroform-d
8.69
7.74
7.56
7.47
7.28
3.92
1.60
HCl
O
N
H
O
37
2500
2000
1500
1000
500
0
1.00
1.91
1.03
2.13
0.99
2.89
2.21
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)


mz-ii-37 bottom spot product
HCl
38
7.73
7.47
7.29
4.54
3.74
2.35
2.07
1.59
1.29
1.00
1.47
0.62
0.99
0.33
1.47
0.52
0.41
1.62
0.61
f1 (ppm)

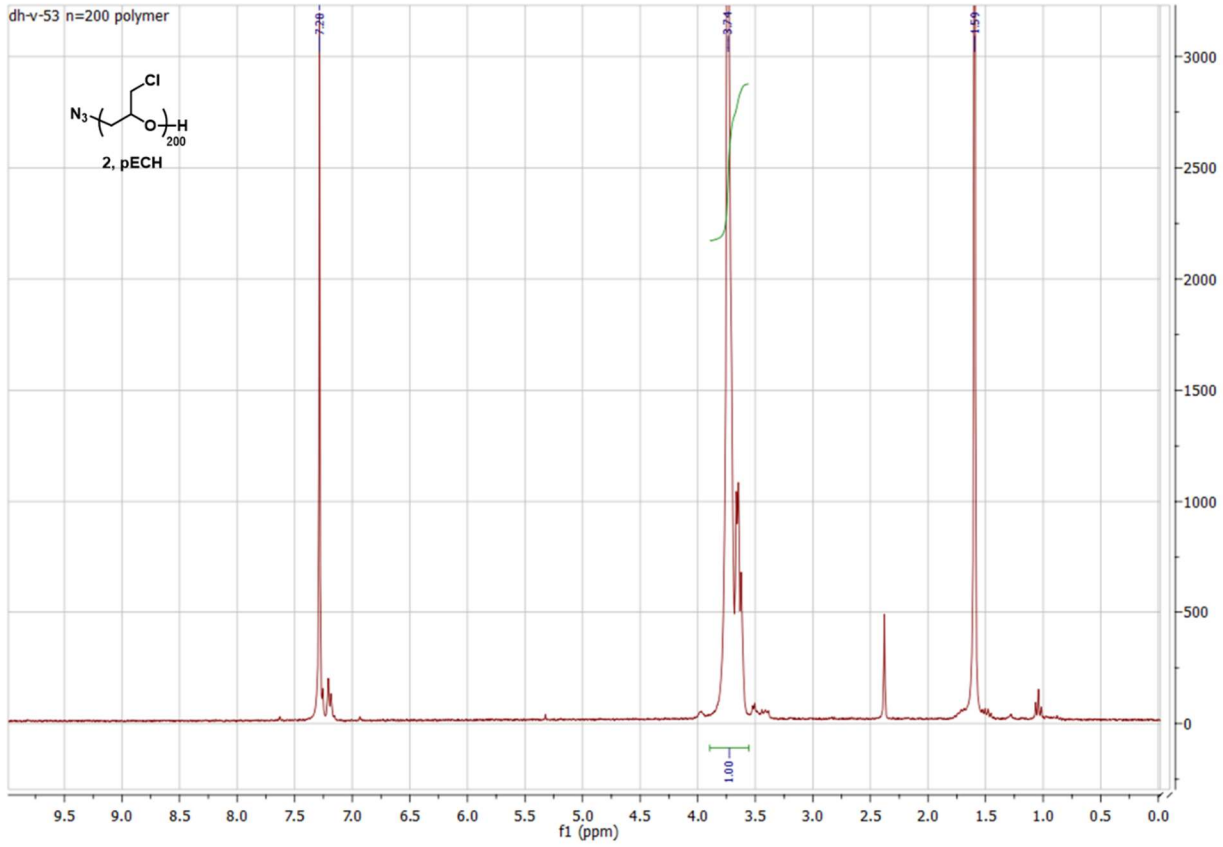

mz-ii-42 in D2O
HCl
H
N
O
39
4.70
3.93
3.73
2.78
2.14
3.09
1.00
f1 (ppm)

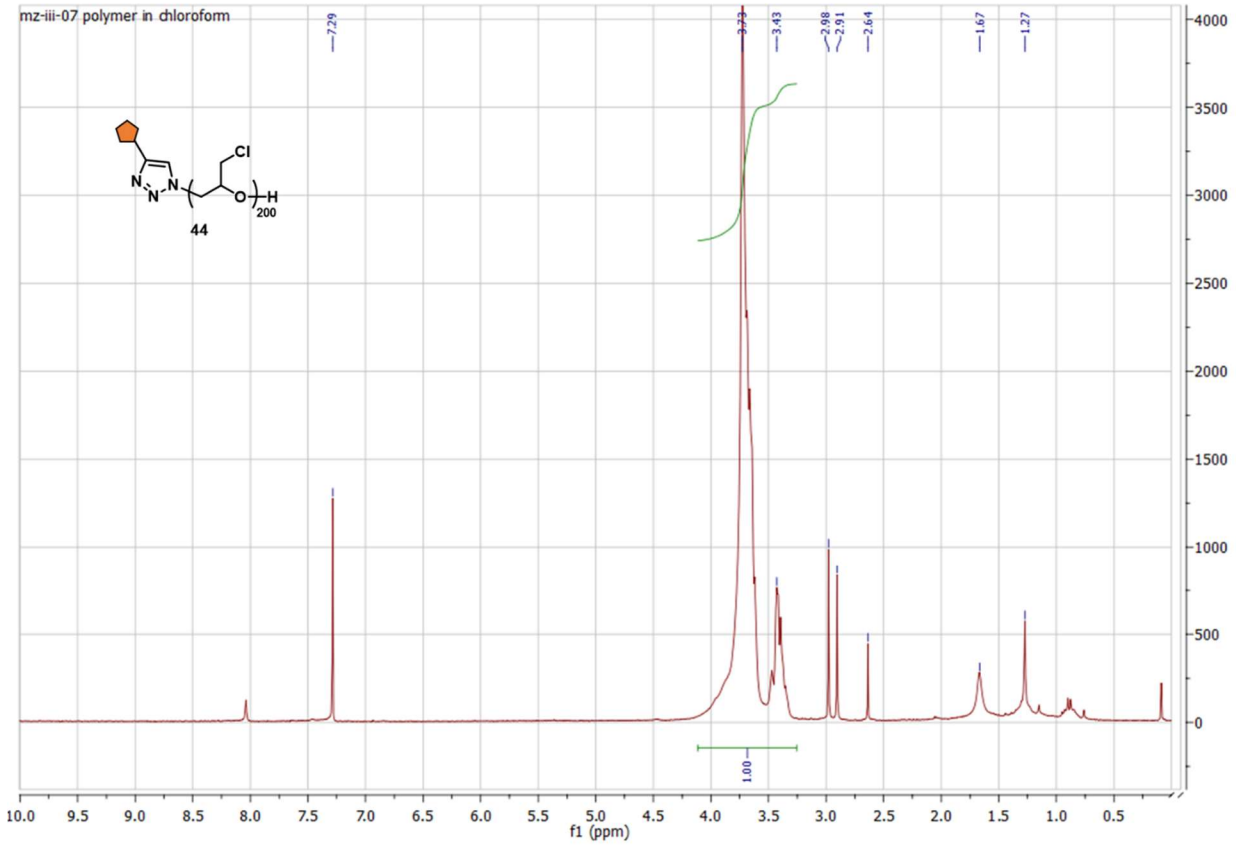

mz-ii-33 1h d2o
OH
HO
HO
OH
O
N
40
4.70
4.19
3.84
3.80
3.61
3.49
3.30
2.60
1.00
2.08
0.89
4.31
2.16
2.40
0.85
f1 (ppm)

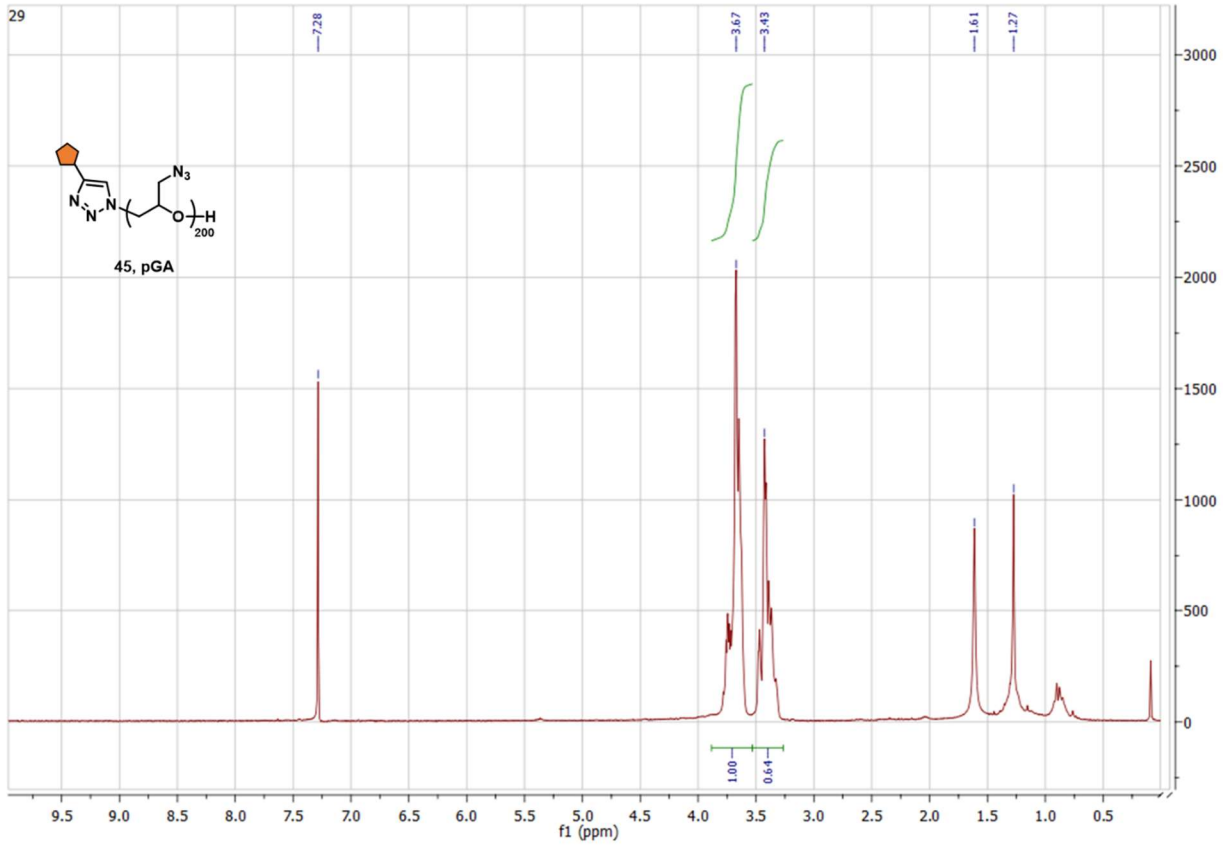

mz-ii-45 clean 1h d2o
41
4.70
4.41
4.23
3.86
3.81
3.56
3.40
3.26
2.61
1.00
0.95
17.36
0.82
f1 (ppm)

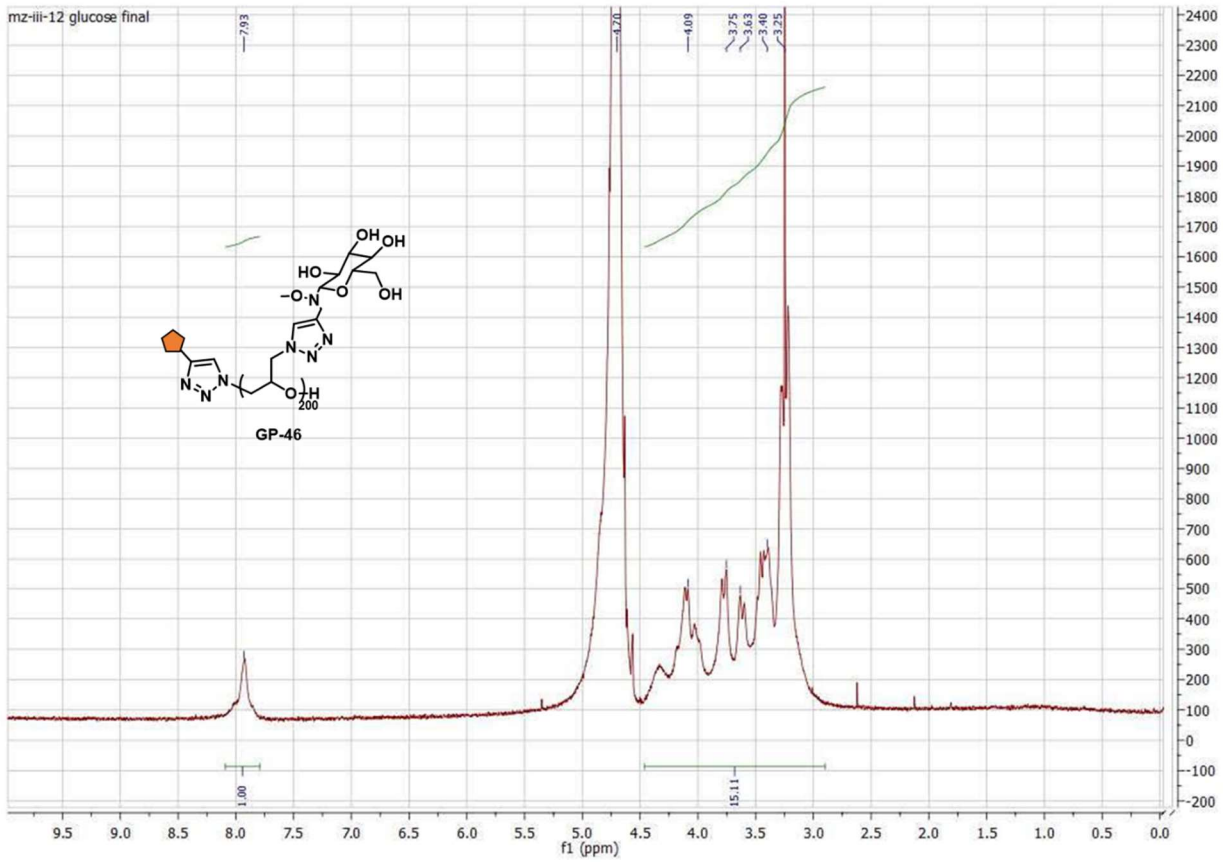

mz-ii-69 1h d2o
8.38
5.34
4.89
4.24
3.73
3.67
3.63
3.51
2.63
0.60
1.00
1.18
0.82
18.69
0.96
f1 (ppm)

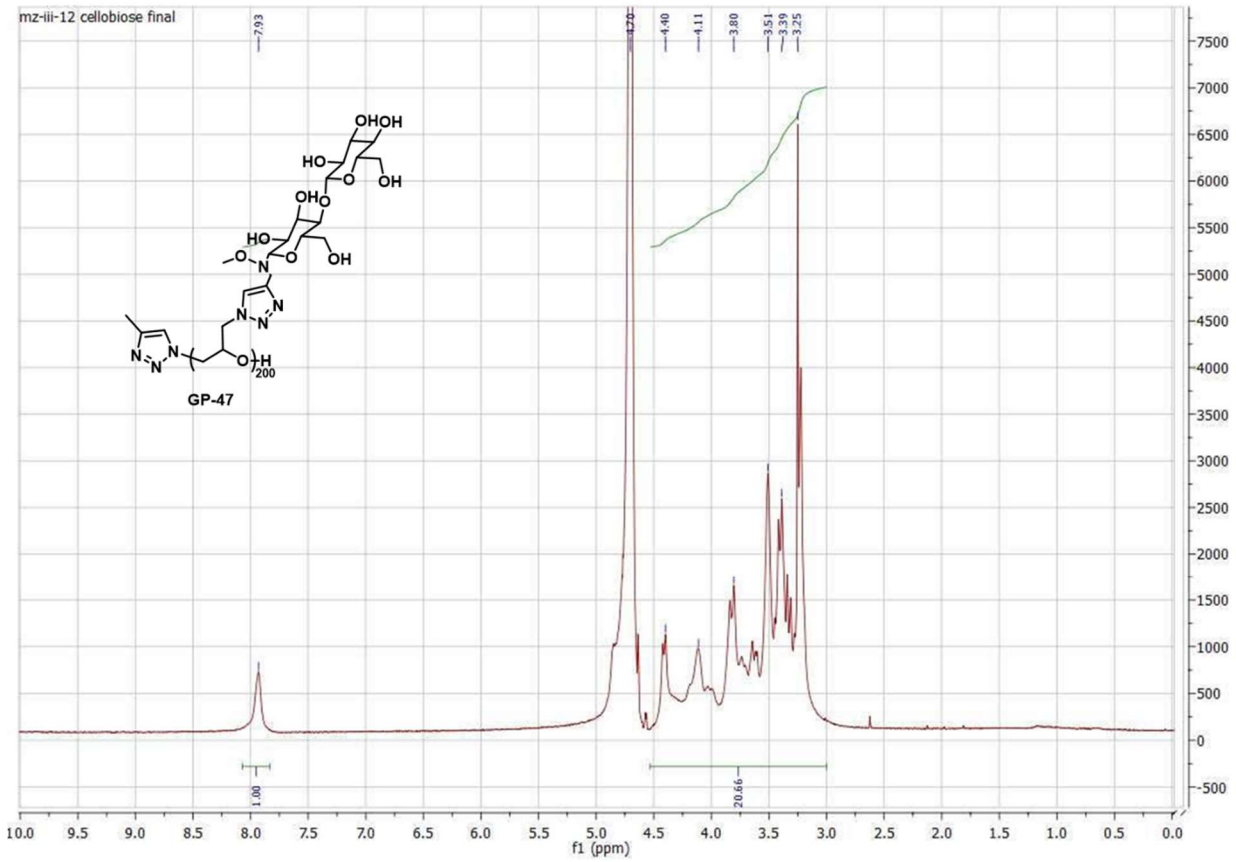
42

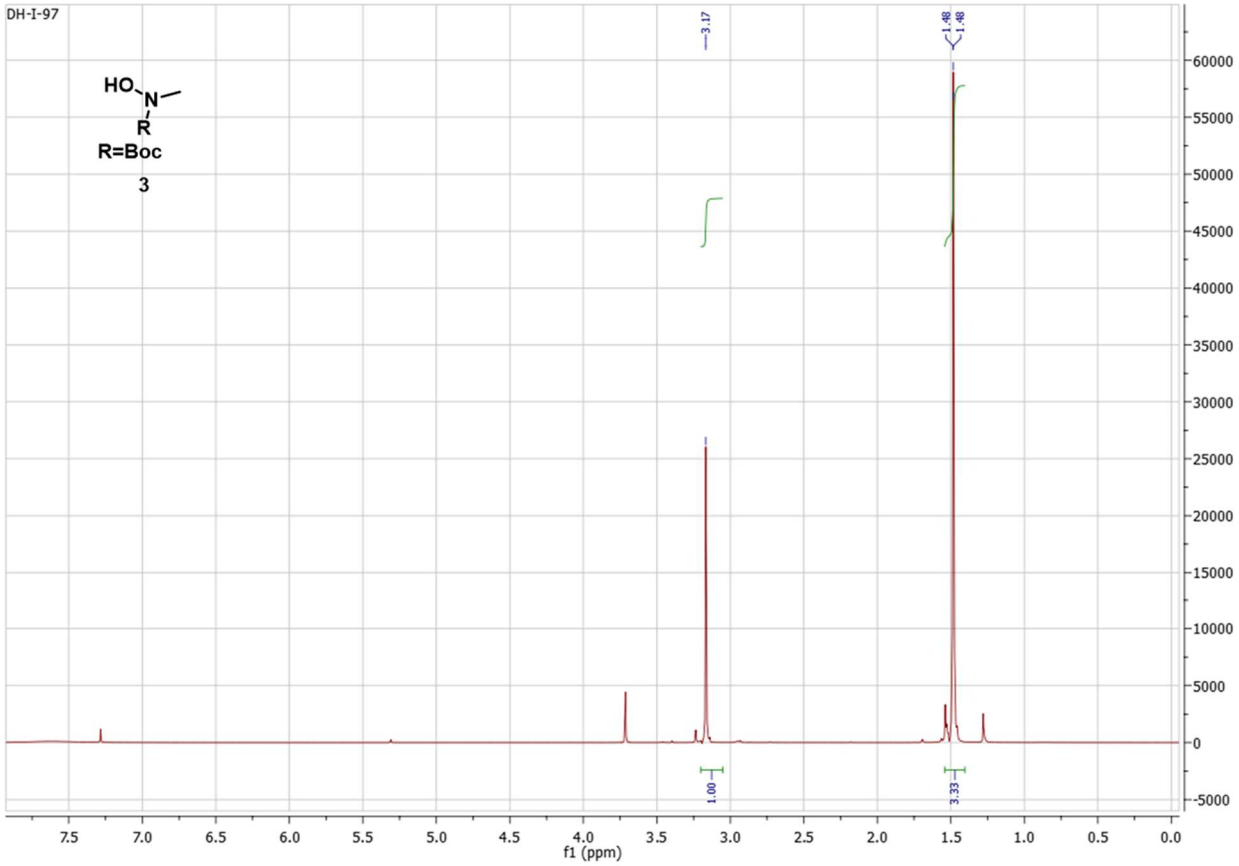

dh-v-53 n=200 polymer
Cl
N3
O
H
200
2, pECH
7.20
3.74
1.59
1.00
3000
2500
2000
1500
1000
500
0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


mz-iii-07 polymer in chloroform
7.29
3.73
3.43
2.98
2.91
2.64
1.67
1.27
Cl
N
N
N
O
H
200
44
1.00
f1 (ppm)


29
7.28
3.67
3.43
1.61
1.27
45, pGA
1.00
0.64
f1 (ppm)


mz-iii-12 glucose final
7.93
4.70
4.09
3.75
3.63
3.40
3.25
OH
OH
HO
O
N
OH
N
N
N
N
O
H
200
GP-46
1.00
15.11
f1 (ppm)


mz-iii-12 cellobiose final
7.93
4.70
4.40
4.11
3.80
3.51
3.39
3.25
GP-47
1.00
20.66
f1 (ppm)


DH-I-97
HO
N
R
R=Boc
3
3.17
1.48
1.48
1.00
3.33
f1 (ppm)

Section 4.10 References

[1] Huang ML, Godula K. Nanoscale materials for probing the biological functions of the glycocalyx. *Glycobiology*. 2016;26(8):797-803. doi:10.1093/glycob/cww022 PMID: 26916883

[2] Pulsipher A, Griffin ME, Stone SE, Brown JM, Hsieh-Wilson LC. Directing neuronal signaling through cell-surface glycan engineering. *J Am Chem Soc.* 2014;136(19):6794-6797. doi:10.1021/ja5005174 PMID: 24746277

[3] Griffin ME, Hsieh-Wilson LC. Synthetic probes of glycosaminoglycan function. *Curr Opin Chem Biol*. 2013;17(6):1014-1022. doi:10.1016/j.cbpa.2013.09.015 PMID: 24148269

[4] Tanaka J, Gleinich AS, Zhang Q, Whitfield R, Kempe K, Haddleton DM, Davis TP, Perrier S, Mitchell DA, Wilson P. Specific and Differential Binding of N-Acetylgalactosamine Glycopolymers to the Human Macrophage Galactose Lectin and Asialoglycoprotein Receptor. *Biomacromolecules.* 2017;18(5):1624-1633. doi:10.1021/acs.biomac.7b00228 PMID: 28418238

[5] Huang ML, Purcell SC, Verespy S III, Wang Y, Godula K. Glycocalyx scaffolding with synthetic nanoscale glycomaterials. *Biomater Sci.* 2017;5(8):1537-1540. doi:10.1039/c7bm00289k PMID: 28616946

[6] Huang H, Rodolis MT, Bhatia SR, Sampson NS. Sugars Require Rigid Multivalent Displays for Activation of Mouse Sperm *Acrosomal Exocytosis. Biochemistry.* 2017;56(22):2779-2786. doi:10.1021/acs.biochem.7b00166 PMID: 28509550

[7] Huang, M. L., Smith, R. A. A., Trieger, G. W. & Godula, K. Glycocalyx Remodeling with Proteoglycan Mimetics Promotes Neural Specification in Embryonic Stem Cells. *J. Am. Chem. Soc.* **136**, 10565–10568 (2014).

[8] Carlotti, S. Labbé, A., Rejsek, V., Doutaz, S., Gervais, M., Deffieux, A. Living/Controlled Anionic Polymerization and Copolymerization of Epichlorohydrin with Tetraoctylammonium Bromide−Triisobutylaluminum Initiating Systems. *Macromolecules* **41**, 7058–7062 (2008).

[9] P, L. Bf, L., Sa, van den B., Ed, P., A, L., Cj, H., Na, L. Poly[(ethylene oxide)-co-(methylene ethylene oxide)]: A hydrolytically-degradable poly(ethylene oxide) platform., Poly[(ethylene oxide)-co-(methylene ethylene oxide)]: A hydrolytically-degradable poly(ethylene oxide) platform. *ACS Macro Lett. ACS Macro Lett.* **1, 1**, 1240, 1240–1243 (2012).

[10] Jourdant, A., González-Zamora, E. & Zhu, J. Wilkinson's Catalyst Catalyzed Selective Hydrogenation of Olefin in the Presence of an Aromatic Nitro Function: A Remarkable Solvent Effect. *J. Org. Chem.* **67**, 3163–3164 (2002).

[11] Herzberger, J. & Frey, H. Epicyanohydrin: Polymerization by Monomer Activation Gives Access to Nitrile-, Amino-, and Carboxyl-Functional Poly(ethylene glycol). *Macromolecules* **48**, 8144–8153 (2015).

[12] Cresswell, A. J. A.J., Davies, S.G., Lee, J.A., Morris, M.J., Roberts, P.M., Thomson, J.E. Ring-Opening Hydrofluorination of 2,3- and 3,4-Epoxy Amines by HBF4·OEt2: Application to the Asymmetric Synthesis of (S,S)-3-Deoxy-3-fluorosafingol. *J. Org. Chem.* **76**, 4617–4627 (2011).

[13] Godula, K., Rabuka, D., Nam, K. T. & Bertozzi, C. R. Synthesis and Microcontact Printing of Dual End-Functionalized Mucin-like Glycopolymers for Microarray Applications. *Angew. Chem. Int. Ed.* **48**, 4973–4976 (2009).

[14] Beak, P. & Selling, G. W. Displacements at the nitrogen of lithioalkoxylamides by organometallic reagents. *J. Org. Chem.* **54**, 5574–5580 (1989).

[15] Godula, K. & Bertozzi, C. R. Synthesis of Glycopolymers for Microarray Applications via Ligation of Reducing Sugars to a Poly(acryloyl hydrazide) Scaffold. *J. Am. Chem. Soc.* **132**, 9963–9965 (2010).

[16] Ec, W., Na, Y., J, S. & Cr, B. Glycocalyx Engineering with a Recycling Glycopolymer that Increases Cell Survival In Vivo. *Angew. Chem. Int. Ed Engl.* **54**, 15782–15788 (2015).

[17] Devulapally, R. & Hon, Y.-S. The first total synthesis of (±)-zenkequinone B. *Tetrahedron Lett.* **52**, 3183–3185 (2011).

[18] Charmantray, F., Dellis, P., Hélaine, V., Samreth, S. & Hecquet, L. Chemoenzymatic Synthesis of 5-Thio-D-xylopyranose. *Eur. J. Org. Chem.* **2006**, 5526–5532 (2006).

[19] André, C., Bolte, J. & Demuynck, C. Syntheses of l-threose and d-erythrose analogues modified at position 2. *Tetrahedron Asymmetry* **9**, 1359–1367 (1998).

[20] Abushanab, E. Vemishetti, P., Leiby, R.W., Singh, H.K., Mikkilineni, A.B., Wu, D.C.J., Saibaba, R., Panzica, R.P. The chemistry of L-ascorbic and D-isoascorbic acids. 1. The preparation of chiral butanetriols and -tetrols. *J. Org. Chem.* **53**, 2598–2602 (1988).

[21] Sharma, G. V. M. & Cherukupalli, G. R. Stereoselective total synthesis of microcarpalide. *Tetrahedron Asymmetry.* **17**, 1081–1088 (2006).

[22] Honigfort, D; Altman, M; Gagneaux, P; and Godula, K. Glycocalyx crowding with synthetic mucin mimetics strengthens interactions between soluble and virus-associated lectins and

cell surface glycan receptors. *Submitted*. Available on Bioachive: https://doi.org/10.1101/2021.05.07.443169

[23] 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.* **67**, 3057–3064 (2002).

[24] El-Boubbou, K. Zhu, D.C., Vasileiou, C., Borhan, B., Prosperi, D., Li, W., Huang, X. Magnetic Glyco-Nanoparticles: A Tool To Detect, Differentiate, and Unlock the Glyco-Codes of Cancer via Magnetic Resonance Imaging. *J. Am. Chem. Soc.* **132**, 4490–4499 (2010).

[25] Alarcón-Manjarrez, C., Arcos-Ramos, R., Álamo, M. F. & Iglesias-Arteaga, M. A. Synthesis, NMR and crystal characterization of dimeric terephthalates derived from epimeric 4,5-seco-cholest-3-yn-5-ols. *Steroids* **109**, 66–72 (2016).

[26] Synthesis and Characterization of Glycidyl Azide Polymers Using Isotactic and Chiral Poly(epichlorohydrin)s | Macromolecules. https://pubs.acs.org/doi/10.1021/ma951839f.

[27] Gerken TA. Biophysical approaches to salivary mucin structure, conformation and dynamics. *Crit Rev Oral Biol Med.* 1993;4(3-4):261-270. doi:10.1177/10454411930040030201 PMID: 8373983

[28] Shogren R, Gerken TA, Jentoft N. Role of glycosylation on the conformation and chain dimensions of O-linked glycoproteins: light-scattering studies of ovine submaxillary mucin. *Biochemistry*. 1989;28(13):5525-5536. doi:10.1021/bi00439a029 PMID: 2775721