UNIVERSITY OF CALIFORNIA SAN DIEGO

Development of platforms for manipulation and interrogation of glycan interactions and assessment of their effects on cellular signaling

A dissertation submitted in partial satisfaction of the requirements

for the degree of Doctor of Philosophy

in

Chemistry

by

Austen Larson Michalak

Committee in charge:

Professor Kamil Godula, Chair
Professor Daniel Donoghue
Professor Valerie Schmidt
Professor Yitzak Tor
Professor Dong Wang

2022

The Dissertation of Austen Larson Michalak is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

DEDICATION

I dedicate this dissertation to my recently departed father, Gerald Michalak, who taught me so much before he passed early in my graduate studies. Fair winds, Dad.

EPIGRAPH

Fame or integrity: which is more important?
Money or happiness: which is more valuable?
Success or failure: which is more destructive?

If you look to others for fulfillment,
you will never truly be fulfilled.
If your happiness depends on money,
you will never be happy with yourself.

Be content with what you have;
rejoice in the way things are.
When you realize there is nothing lacking,
the whole world belongs to you.

Tao Te Ching

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SCHEMA

ACKNOWLEDGEMENTS

For their unfaltering support, I would like to thank my mother Yvonne as well as my stepfather Mark. Your encouragement and patience reinforced me as I spent many long years studying for my Ph.D. Whenever I could venture back to the east coast, a homecooked meal, a thoughtful bike ride through Umstead, and some good ol' family time went a long way to recharge my batteries.

My brother, Logan Michalak was especially supportive through my Ph.D., supporting eachother as we balanced the rigors of university while going through the tumultuous challenge of losing our father too early. Whether going backpacking through the Canadian Rockies, or traversing the Alps, or just hopping online after work to play a game and vent our daily struggles, my life is always better with the support of my brother. To many more adventures, Logan!

Although he passed away early in my graduate studies, I wouldn't have pursued my chemistry Ph.D. without the guidance and upbringing of my father, Gerald Michalak. Some of my earliest memories were him reading me from lying in our driveway to look at the stars, to reading a 6-year-old me "Scientific American" as a bedtime story. He inspired curiosity in the world at an early age, and encouraged me to pursue things that were weird and interesting, which led me to chemistry. He always set the example as someone successful enough to author over ten patents, yet have time teach a young me to sail, to

read the surface of the water and the stars of the night sky, and how to make the finest pancakes on Sunday mornings.

The daily life of a research scientist can be quite punctilious, and I would like to thank my loving girlfriend for her support and adding a brightly burning romance to my life. Without you, grad school would have been inconceivably more difficult, and life bland. Thank you for being my first person to talk to, my best friend, and for the burst of flavor and spice in our intertwined lives.

During my time and work in graduate studies, I have been fortunate enough to have some of the brightest and best burgeoning scientists as colleagues, which also doubled as a close knit “family” of friends, spanning across several labs at UCSD. Thank you for many long lab nights, many lab clean-up days, and after work hangouts at apartments or bars, and long enjoyable tournaments of Civ. Working in the trenches of science has forged rock-hard friendships, and some of the greatest gifts from my graduate studies are the lifelong friends I’ve gained. There are too many names to name; if you know, you know.

Throughout my tenure in graduate studies, a bright and jovial post-doctoral scholar, Yinan Wang, particularly influenced me in graduate school. In our time together, we went on many adventures from snowboarding local mountains to exploring the deserts and Joshua trees. Yinan, you were taken from us too soon, and your intelligence, your

brightness and vigor for life are deeply missed. We had planned future tour of China together; I will complete this trip and will meet your spirit in Jiuzhai Valley, friend.

I've been fortunate enough to have many mentors, in all shapes and sizes throughout my graduate studies. As an eager but unguided student, I approached my PI Kamil Godula eager to get into research, where he accepted me into his lab and hosted me for my Ph.D, for which I am grateful. In addition to my PI, I've been fortunate enough to work with many talented post-docs, grad students, and undergrads in my lab. I'm thankful for the mentorship I've received in graduate school and I'm eager to pass my lessons forward to future generation.

***

Chapter 1 is an introductory background chapter written by Austen Larson Michalak. Figure 1.5 was partially modified from 1 as it appears in *Advanced Healthcare Materials* 2021. Michalak AL, Trieger GW, Trieger KA, Godula K., Stem Cell Microarrays for Assessing Growth Factor Signaling in Engineered Glycan Microenvironments.

Chapter 2 is an adaptation and reprint of the material as it appears in *Advanced Healthcare Materials* 2021. Michalak AL, Trieger GW, Trieger KA, Godula K., Stem Cell Microarrays for Assessing Growth Factor Signaling in Engineered Glycan Microenvironments. Austen Larson Michalak was the primary author on this publication, and shared equal contribution of work as Greg W Trieger.

Chapter 3 is an adaptation and reprint of the material as it appears in *Stem Cells* 2018. Huang ML, Michalak AL, Fisher CJ, Christy M, Smith RAA, Godula K. Small

Molecule Antagonist of Cell Surface Glycosaminoglycans Restricts Mouse Embryonic Stem Cells in a Pluripotent State. Austen Larson Michalak made contributions as a primary author on this publication.

Chapter 4 was an adaptation and reprint of the material as it appears in *RSC Chemical Science* 2022. Purcell SC, Zhang MH, Honigfort DJ, Ng HFC, Michalak AL, Godula K. Cell surface photoengineering enables modeling of glycocalyx shedding dynamics. Austen Larson Michalak was a primary author of this paper.

Chapter 5 was original unpublished work for conjugation of oligonucleotide barcodes to lectins which served as at the basis for future macromolecular conjugation approaches and was conducted by Austen Larson Michalak.

Chapter 6 was original work primarily conducted by Austen Larson Michalak. Ethan J. Armand developed “glycode extractor” algorithm to process sequencing data sets.

.

VITA

2003-2007 Needham B. Broughton High school, Raleigh NC

2007-2011 Bachelor of science in Chemistry with biochemistry emphasis, biology minor, *cum laude*

University of North Carolina Wilmington, Honors thesis under Dr. Antje Pokorny Almeida

2012-2013 Assistant Scientist, AAI Pharma

2013-2014 Production Chemist, Hologic

2014-2016 Master of Science, Chemistry, University of California San Diego

2016-2022 Doctor of Philosophy, Chemistry, University of California San Diego

PUBLICATIONS

"Cell surface photoengineering enables modeling of glycocalyx shedding dynamics" S. Purcell, M. Zhang, D. Honigfort, H.J. Ng, **A.L. Michalak**, K. Godula. *Chem. Sci.* **2022**

"Stem Cell Microarrays for Assessing Growth Factor Signaling in Engineered Glycan Microenvironments" **A.L. Michalak**, G.W. Trieger, K.A. Trieger, K. Godula. *Adv. Healthcare. Mat.* **2021**,12,32

"Silencing glycosaminoglycan functions in mouse embryonic stem cells with small molecule antagonists" S. Chatterjee, T.N. Stephenson, **A.L. Michalak**, K. Godula, M.L. Huang. *Methods in Enzymology.* **2019**,626:249-270

"Small Molecule Antagonist of Cell Surface Glycosaminoglycans Restricts Mouse Embryonic Stem Cells in a Pluripotent State" M.L. Huang, **A.L. Michalak**, C.J. Fisher, M. Christy, R.A.A. Smith, and K. Godula. *Stem Cells*, Oct **2017**, 36,1

"Lysylated Phospholipids Stabilize Models of Bacterial Lipid Bilayers and Protect Against Antimicrobial Peptides" E. Cox, **A.L. Michalak**, S. Pagentine, P. Seaton, and A. Pokorny. *Biochimica et Biophysica Acta-Biomembranes*, Sep **2014**; 1838(9):2198-2204.

FIELDS OF STUDY

Chemistry, Biochemistry, Chemical biology, stem cell biology, and bioconjugation

ABSTRACT OF THE DISSERTATION

by

Austen Larson Michalak

Doctor of Philosophy in Chemistry

University of California San Diego 2022

Professor Kamil Godula, Chair

The glycocalyx consists of a dense layer of carbohydrates which coat the surface of virtually all living cells, and plays pivotal roles in development, disease progression, and cellular signaling. Although glycans are ubiquitous and central to biology, our understanding of them is still rapidly evolving due to their multivalent properties, heterogenous composition, and genetically non-templated nature – which complicate

their study. This dissertation is dedicated to the conception, development, and application of tools to both measure and manipulate glycans in biology.

Glycosaminoglycans are linear, charged polysaccharides which harbor binding sites for both cytokines and their receptors, and thus are key regulators of cell signaling. The small molecule aminoquinoline, surfen, is a reversible antagonist of interactions between heparan sulfate glycosaminoglycans and growth factors. In Chapter 2, we show that it can be used as a molecular tool to drug the glycome. Surfen is a potent molecule that can reversibly inhibit differentiation of stem cells while promoting maintenance of pluripotency, and as such provides a powerful alternative to genetic methods to control stem cell fate.

While glycosaminoglycans are of paramount importance in development, the inherent structural heterogeneity makes elucidating structure function relationships difficult. In Chapter 3, we engineer glycan microenvironments by conjugating chemically modified heparan sulfate to the gelatin matrix surrounding stem cells, where it influences growth factor binding and downstream cell signaling. By carefully controlling the glycan microenvironment around stem cells, we enable assessment of the contributions of extracellular heparan sulfate to growth factor binding and cell signaling. In Chapter 4, we developed a new method to remodel the cellular glycocalyx using photocleavable glycopolymers to enable a photopatterning approach to engineer the glycocalyx with spatial resolution.

Human epithelial surfaces contain hydrophilic mucin proteins which are richly colonized by bacteria in the human microbiome. Altered bacteria-glycan interactions are associated with inflammatory disease, and new methods to assess interactions would be

of great use. In Chapters 5 and 6, we develop a new platform for assessment of glycan interactions in whole cell bacteria. By implementing a DNA-barcoding system to encode the identity of the glycans in conjugation with a mucin mimetic platform, we assessed the glycan interactions of whole cell *E. coli* using next generation sequencing techniques to provide a rapid readout of binding analysis.

Understanding the intricacies of glycobiology will pave the way for technological breakthroughs in science and medicine – however scientists need tools to rapidly interrogate and control the glycocalyx. The work described in this dissertation addresses these needs by expanding the modern scientists' toolkit to probe and manipulate the glycome.

1. General introduction to Glycoscience

The topic of this dissertation is the study of carbohydrates, or glycans, which cover the surface of virtually all living cells. Glycans are potent biological regulators which facilitate an incredible diversity of processes. The dense collection of glycans on the cell surface, collectively termed the glycocalyx, serves as the interface between the cell and the outside world. The glycocalyx may extend for >1000 of nanometers away from the cell surface, and is often conceptualized as a “dense forest” of coating the cell surface.[1] (Figure 1.1) As a pathogen, protein, or microorganism approaches a cell, the glycocalyx represents the first interaction with the cellular surface, a molecular “handshake” between two entities. The glycocalyx transmits information between the cell and the outside world primarily in molecular recognition events between the cell and pathogens, proteins, bacteria and growth factors, and have profound influences in biology. At least 50% of proteins in the human proteome are glycosylated, with some estimates being as high as 70%, highlighting the importance of glycans to protein function and biological interactions.[2] Thus, glycans are essential for biological functions on both the cell surface, as well as protein interactions inside the cell.

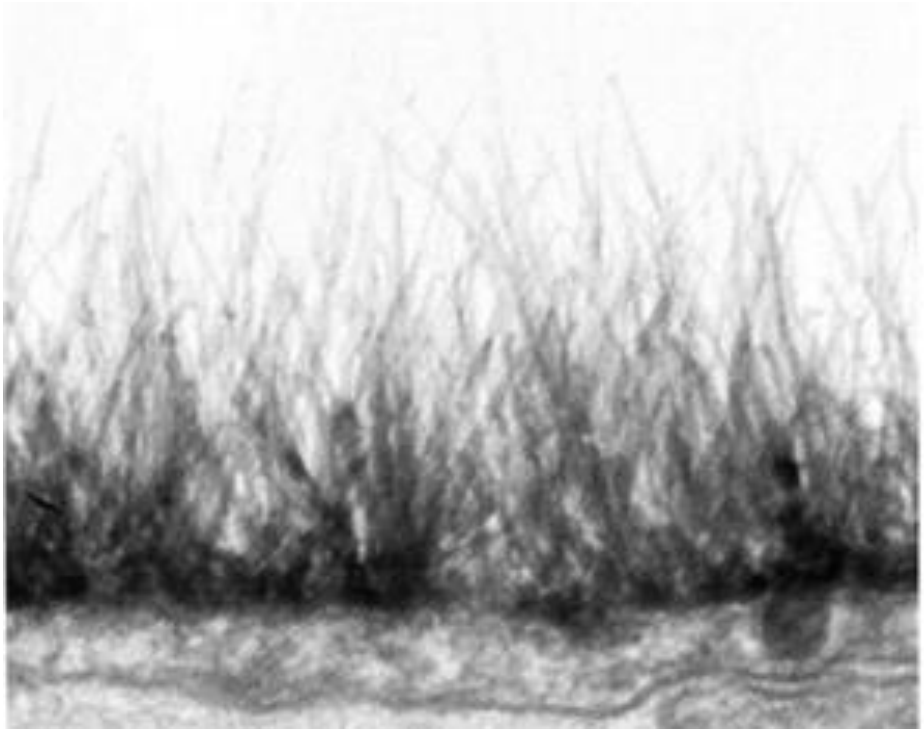


Figure 1.1 The endothelial glycocalyx. The glycocalyx of myocardial tissue extends for hundreds of nanometers. The hair-like strands represent a dense covering of glycoconjugates, collectively termed the glycocalyx. Figure reproduced from [3].

1.1 Functions of glycans, and challenges to researching them

Glycans serve diverse functions across biology. As the first layer of contact between the cellular surface and the outside world, their influence spans development, pathogenesis, host immune response, and also serve as structural elements in many tissues[4]. In development, glycans coordinate complex gradients of growth factors, hormones, and other transformative cytokines. During embryogenesis, the transformative cytokines in the FGF, BMP, and Wnt families bind glycosaminoglycans on the cellular surface, which pattern cellular maturation and differentiation. Glycans are indispensable for development.[5] Proteins critical for cellular signaling are also dependent on glycans - for example, lectins, which are proteins that recognize a particular glycan structure via

carbohydrate recognition domains influence of galectin-3 on VEGF and FGF signaling in development as well.[6] In addition to coordinating growth factor binding, the glycocalyx also serves as a storage depot for growth factors adjacent to the glycocalyx, which may be liberated during glycocalyx shedding events to bind GAGs adjacent to cells to direct signaling in instances of development, and wound healing.[7,8] Glycans also play key roles outside of mammalian life, as they central biological roles in bacterial cells, and most viruses.[9]

However, when approaching glycans, the researcher is confronted with complexity on nearly all levels. There is an incredible diversity of length, structure, and function of glycoconjugates on the cell surface. For example, mucins may extend up to 1500 nanometers away from the cell surface and provide a physical barrier on the outside of the cell[10], while glycolipids and other glycoconjugates are directly adjacent to the cell surface.[11] In contrast to linear biopolymers, like nucleic acids or peptides, glycans can be branched, into two or more branches. For a typical reducing hexasaccarhide, there are 10^{12}, or one trillion, possible structures due to linkage variations, branched structures, or stereochemical isomers.[12] Nature takes advantage of these possibilities and there is an incredible diversity of glyconjugates found across all domains of life on earth.

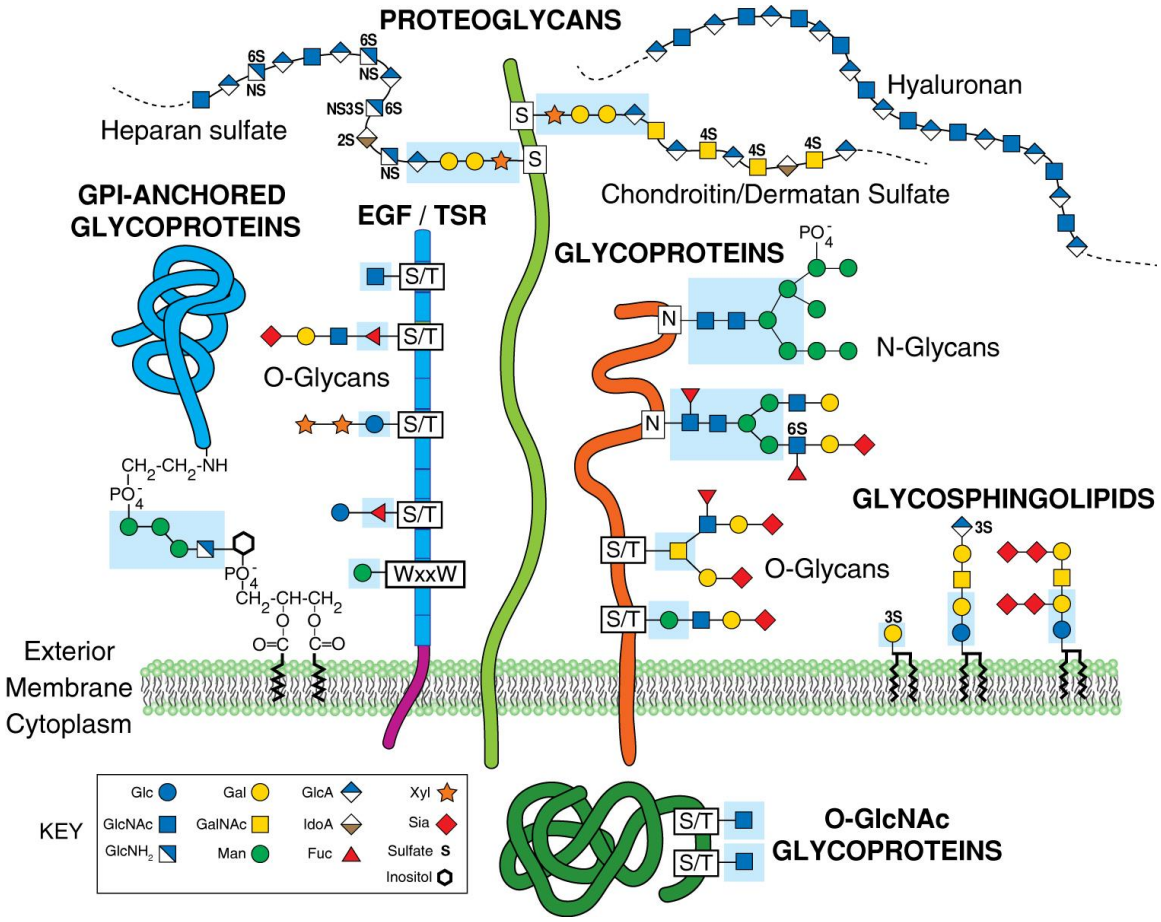


Figure 1.2 Common classes of animal glycans. (Modified and updated from Varki A. 1997. *FASEB J* **11**: 248–255; Fuster M, Esko JD. 2005. *Nat Rev Can* **7**: 526–542, with permission from Macmillan; and Stanley P. 2011. *Cold Spring Harb. Perspect. Biol.* **3**: a005199. Original art has been adapted and redrawn by R.D. Cummings.)

Although glycans are of tremendous importance in biology, our understanding of glycans lags significantly behind those of other essential biomolecule classes, like proteins, nucleic acids, and lipids.[13] There is much about glycans which remains to be uncovered, and a detailed picture of the glycome is just coming into focus at the time of this writing. This gap in our understanding of glycans is due to many factors, such as the complexity, non-templated synthesis, and inherent heterogeneity. Perhaps most of all, this gap in our understanding of glycans in biology arises from a lack of easily accessible

methods to rapidly study and manipulate glycans – a shortcoming addressed in this dissertation.

1.2 Glycan structure

Glycans cover an incredible breadth of structures, however they can be divided into two main classes of glycans: those which are N-linked to the nitrogen on an asparagine residue, or O-linked to an oxygen molecule in an serine or threonine protein reside, or the hydroxyl group on the head of a ceramide.[14] A sequence of amino acids typically N-glycosylated, termed sequon, is where N-glycans are attached via an asparagine residue in the amino acid sequence (Asn-X-Ser/Thr) motif, where X cannot be proline.[15]

Complex polysaccharides arise from the sequential addition of monosaccharide building blocks through glycosidic linkages between the anomeric carbon of one residue and one of the hydroxyl groups of another.[16] The complexity and number of possible polysaccharide structures is much greater than other prominent biopolymers such as DNA or proteins for several reasons. Due to the diversity of monosaccharides, and the variety of glycosidic linkages between monosaccharides, and branched confirmations covers vast chemical space.[12,] Monosaccharide residues can be linked together in an extended linear confirmation (proteoglycans) or highly branched configurations, as seen in tri- or tetra- antennary N-glycans (figure 1.2), while DNA and proteins are restricted to linear confirmations. Additionally, monosaccharides can be modified with sulfation,

phosphorylation, or are commonly N-acetylated. To add an additional level of complexity, glycosidic linkages can differ in stereochemistry, with both alpha and beta confirmations occurring in biology.

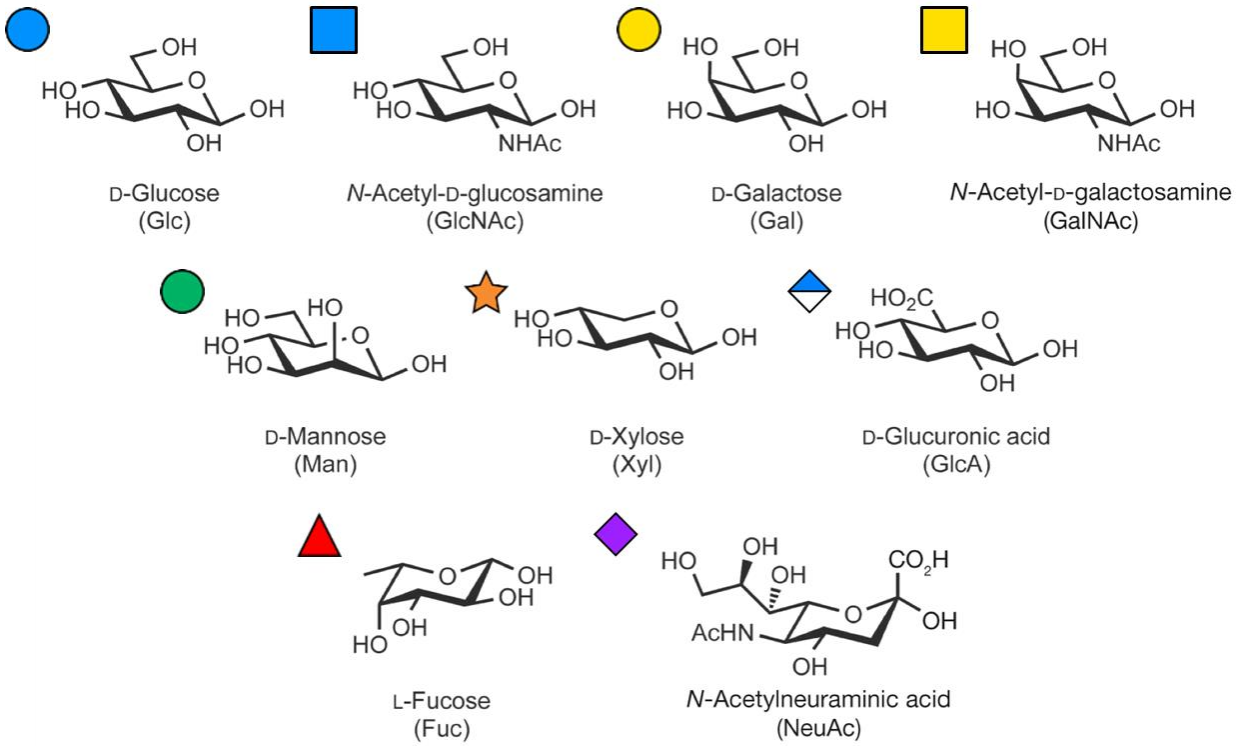


Figure 1.3 Common monosaccharide building blocks found in vertebrates. These 12 monosaccharides are the primary glycans found in vertebrates. The C6 isomer of glucuronic acid, iduronic acids are additionally found in glycosaminoglycans polysaccharides.

1.3 Glycan biosynthesis

The majority of glycosylation occurs in an assembly line fashion within the ER and golgi apparatus. For the O- and N-glycosylation of proteins, and the O-glycosylation of glycolipids the process begins in the ER and continues as nascent proteins migrate through the golgi apparatus. Glycans are elaborated and extended through a symphony of compartmentalized enzymes, which elaborate glycans based on the substrate availability.[15] For O-glycosylation of proteins bearing glycans through an O-GalNAc residue attached to serine or threonine, such as mucins for example. N-glycan biosynthesis is initiated in the ER through via of a highly conserved $Glc_3Man_9GlcNAc_2$ glycan linked to dolichol anchor to the asparagine residues of a nascent N-glycoprotein. The trimming and subsequent elaboration of this core structure then occurs when the protein is moved through the cis and trans goligi.[23] The glycosyltransferase machinery responsible for synthesis of N-glycans may share some also act on elaboration of O-glycans, but may have a preference for O-glycans as acceptor substrates.[17] Enzymes involved in glycan biosynthesis may have multiple activities, such as the glycosaminoglycan modifying enzyme N-deacetylation N-sulfotransferase (NDST), which contains two catalytic domains for the N-deacetylation of GlcNAc and subsequent N-sulfation, within the same enzyme.[23]

However, some glycan modifications occur outside of the conventional ER/golgi sub compartments, such as O-glcNac, which is a widespread the modification of protein residues with glcNAc, and occurs primarily in the nuclear, cytosolic, and mitochondrial compartments of the cell. O-GlcNacylation is added via a single enzyme, O-GlcNAc

transferase (OGT) and is more similar to protein modifications such as phosphorylation, due the fact that the O-GlcNAc moiety isn't elongated into more complex glycan structures, and it may be added and removed multiple times during the lifetime of a protein.[18] Another notable exception is the biosynthesis of the extremely high molecular weight glycan, hyaluronan, which is secreted from the cell membrane, in part due to its large size of 10^4 disaccharides with end to end lengths reaching close to 10 µM.[19]

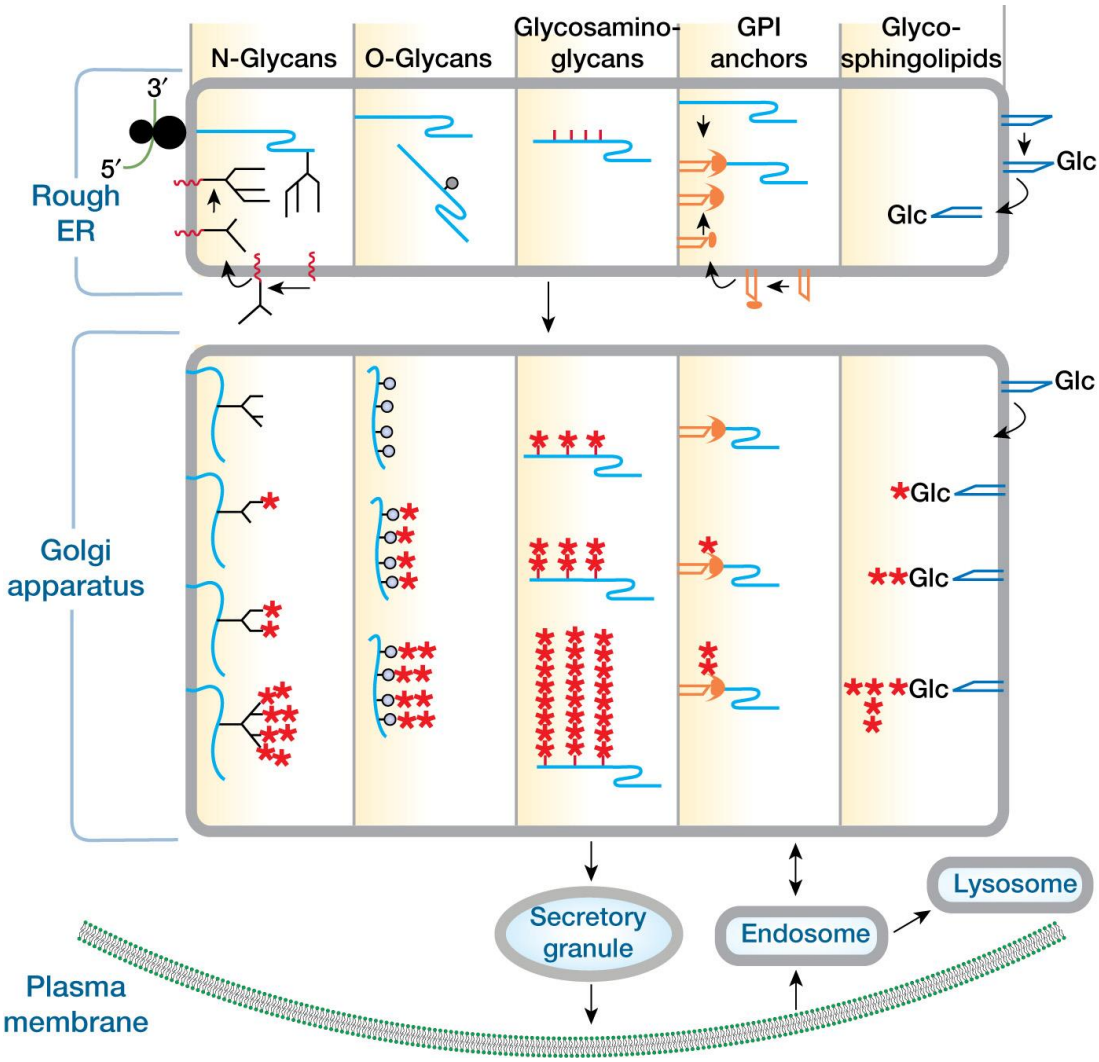


Figure 1.4 Glycoconjugate biosynthesis primarily occurs the ER and Golgi apparatus. For N-glycans, the dolichol anchored precursor $Glc_3Man_9GlcNAc_2$ is transferred to the Golgi where it is trimmed and elaborated. Red stars represent addition of monosaccharide units to growing glycan residues. Figured reproduced from[20].

Glycosylation is a non-templated biological process, meaning the exact structure of glycoconjugates are inaccessible through sequencing of DNA or RNA transcripts. This inherent heterogeneity is a hallmark of glycans, and multiple copies of the same protein will have slightly different glycosylation profiles based on enzyme expression and availability of nucleoside donor sugars.

1.4 Glycosaminoglycans as biological modulators

One group of glycans of particular importance, which function as modulators of growth factors and cytokine mediated signal transduction are the glycosaminoglycans (GAGs). GAGs are extended, unbranched polysaccharides which are appended to a serine residue proteoglycan (PG) protein core, and a major ECM component in all mammalian tissues.[21] One of the primary functions of GAGs is to facilitate cell signaling – as they are an essential component of growth factor/receptor complexes on the cell surface. In development, GAGs direct tissue specification by creating growth factor gradients.[22,23] In mature organisms, they mediate cell-cell and cell-matrix interactions by functioning as a potent co-receptor for growth factors and cytokines.[23] GAGs may be tethered to the cell surface via their parent proteoglycan, may be secreted into the surrounding environment, or localized to the adjacent ECM, which acts as a depot for storage or future secretion.[24]

There are several major classes of GAGs, which are characterized by their repeating disaccharide units, as well as glycan modifications, such as acetylation, epimerization or sulfation. The main classes of GAGs are Heparan sulfate/heparin, chondroitin sulfate/dermatan sulfate, keratan sulfate, or the non-sulfated hyaluronic acid. Heparan sulfate (HS)/heparin, Chondroitin sulfate (CS) and dermatan sulfate (DS) GAG classes are synthesized in the Golgi and differ in their disaccharide composition, glycan modification, expression and function. In contrast, hyaluronic acid (HA) is an extremely high weight GAG with molecular weights ranging up to 10 million daltons.[25] Due to its extreme size, HA is synthesized from the cytoplasm and extruded from the cell membrane. The unique size and charge of HA gives it an ability to form a hydrogel for lubrication and shock absorption in joints. Heparin sulfate (HS), and its secreted highly sulfated analog, heparin, are major players in regulating signaling. The primary core structure of HS consists of repeating units of N-acetylglucosamine (GlcNAc) and glucuronic acid (GlcA) in the repeating disaccharide motif GlcNAcβ1-4GlcA1-3.

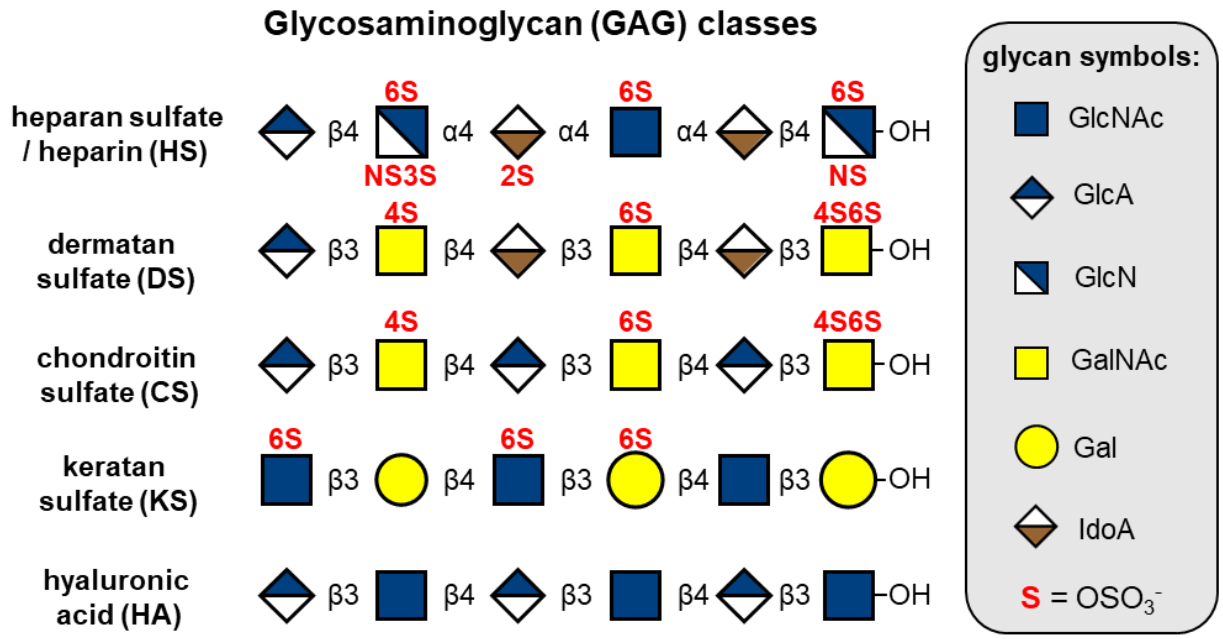


Figure 1.5 Glycosaminoglycan structures. GAGs consist of repeating disaccharide units composed of an N-acetylated or N-sulfated hexosamine and either a uronic acid (GlcA or IdoA) or galactose. Hyaluronan lacks sulfate groups, but the rest of the glycosaminoglycans contain sulfates at various positions. Dermatan sulfate is distinguished from chondroitin sulfate by the presence of IdoA. Heparan sulfate is the only glycosaminoglycan that contains an N-sulfated hexosamine. Keratan sulfates lack uronic acids and instead consist of sulfated galactose and GlcNAc residues. Reducing termini are to the right in all sequences. Figure reproduced from[26].

GAGs are essential for development and growth of mammalian organisms. Cells with reduced or abolished GAGs on the cell surface cannot form high affinity complexes between receptors and growth factors. For example, the essential signaling molecules Fibroblast growth factor (FGF), and Bone Morphogenic Protein 4 (BMP4) show inability for robust signal transduction without the GAG component of their receptor complex.[27,28] Thus, GAGs are a requirement to assemble high affinity signaling complexes for signal transduction to the interior of the cell. For example, one of the primary enzymes responsible for HS chain elongation, *EXT1*, can be deleted from embryonic cells for yield an *EXT1*$^{-/-}$ cell line which doesn't display HS on the cell surface, while other GAGs like

CS remain intact. These *EXT1*$^{-/-}$ cells are unable to differentiate, and the removal of *EXT1* is lethal by day 8.5 in embryonic mice.[29] GAGs are indispensable for development in the early stages of life.

1.5 Stem Cell development and glycans in signaling and differentiation.

Since the isolation of pluripotent cell lines from mice in 1981, stem cells have increased our understanding of development and are burgeoning tools for regenerative medicine.[30] The primary attributes of ESCs are continuous self-renewal, and the ability to differentiate into multiple cell lines (pluripotency). The potential of stem cells in regenerative medicine is enormous, but in order to effectively and safely implement stem cell-based therapies, differentiation processes must be completely controlled. While much is known about the integration of many complex signals to carefully control cell fate and development, significant strides must be made before widespread regenerative therapies controlling stem cell differentiation. Significant hurdles still remain in controlling differentiation processes for desired fate outcomes.

In development, the nascent blastocyst differentiates into three primary germ layers - the outer ectoderm layer, the inner endoderm layer, and with the middle mesoderm layer in-between. From these three germ layers, all organs and tissues in the mature orgasm develop. The ectoderm gives rise to nervous tissues and skin, mesoderm develops into muscle and cardiac tissues, and the endoderm forms gut, pancreas, and liver.[31]

The complex organization in the process of development arises from carefully choreographed gradients and patterns of signaling proteins or cytokines. A delicate balance of signaling pathways determines the fate of differentiation. Mouse embryonic stem cells (mESCs) are similar to human embryonic stem cells in their overall strategy to regulate self-renewal and pluripotency, however the signaling proteins and gene expression to maintain pluripotent or differentiated states differs considerably. Mouse stem cells remain an attractive alternative for understanding differentiation and development due to their small size, and short doubling time. One of the notable differences between murine and human stem cells is that human stem cells do not express LIF and maintain their pluripotency via other protein signaling cascades. As such, mESCs represent a more undifferentiated state of cell development, whereas hESCs are in a more progressed state of differentiation, resembling the state of murine epiblast stem cells.[32] These properties make mESCs an attractive tool for studying processes early in development.

When the first mESC lines were established in-vitro, a requirement was co-culture a layer of mitotically quiescent mouse embryonic fibroblast cells (MEFs). These MEF feeder cells secrete a host of soluble growth factors into solution to maintain mESC pluripotency. The cytokine responsible for the pluripotency of mESCs has been identified as LIF, and MEFs lacking a functional gene for LIF are unable to maintain a pluripotent population of mESCs.[30,33] In mESCs, the pluripotent state of self-renewal is maintained in cell culture by a cytokine of the IL-6 cytokine, leukemia inhibitory factor (LIF), first

named for its ability to hinder leukemia growth. LIF binds to the cell surface receptor LIFR to recruit glycoprotein 130 (gp130) to form a heterodimeric active signaling complex.[34] This signaling event activates the LIF/JAK/STAT3 pathway, which then results in a phosphorylation events activation of the “Signal transducer and activator of transcription 3” protein (STAT3) which enters the nucleus and upregulates the core transcription factors responsible for maintenance of pluripotency, NANOG, Sox2, and Oct4. Another signaling pathway activated through the LIF/LIFR/gp130 complex is PI3K/Akt signaling, which similarly results in an upregulation of the trio of core pluripotency transcription factors, Sox2, Nanog, and Oct4.[35]

In both murine and human primitive embryonic cells, a ubiquitous signaling protein is essential in triggering the departure from a pluripotent condition into an undifferentiated state is triggered by a ubiquitous signaling protein family of Fibroblast Growth Factors (FGFs). In mESCs, FGF signaling initiates the departure from pluripotency into differentiated states. This key signaling pathway which is regulated by HS GAGs, and is an attractive target to modulate biological outcomes by controlling interactions between glycan and protein components.

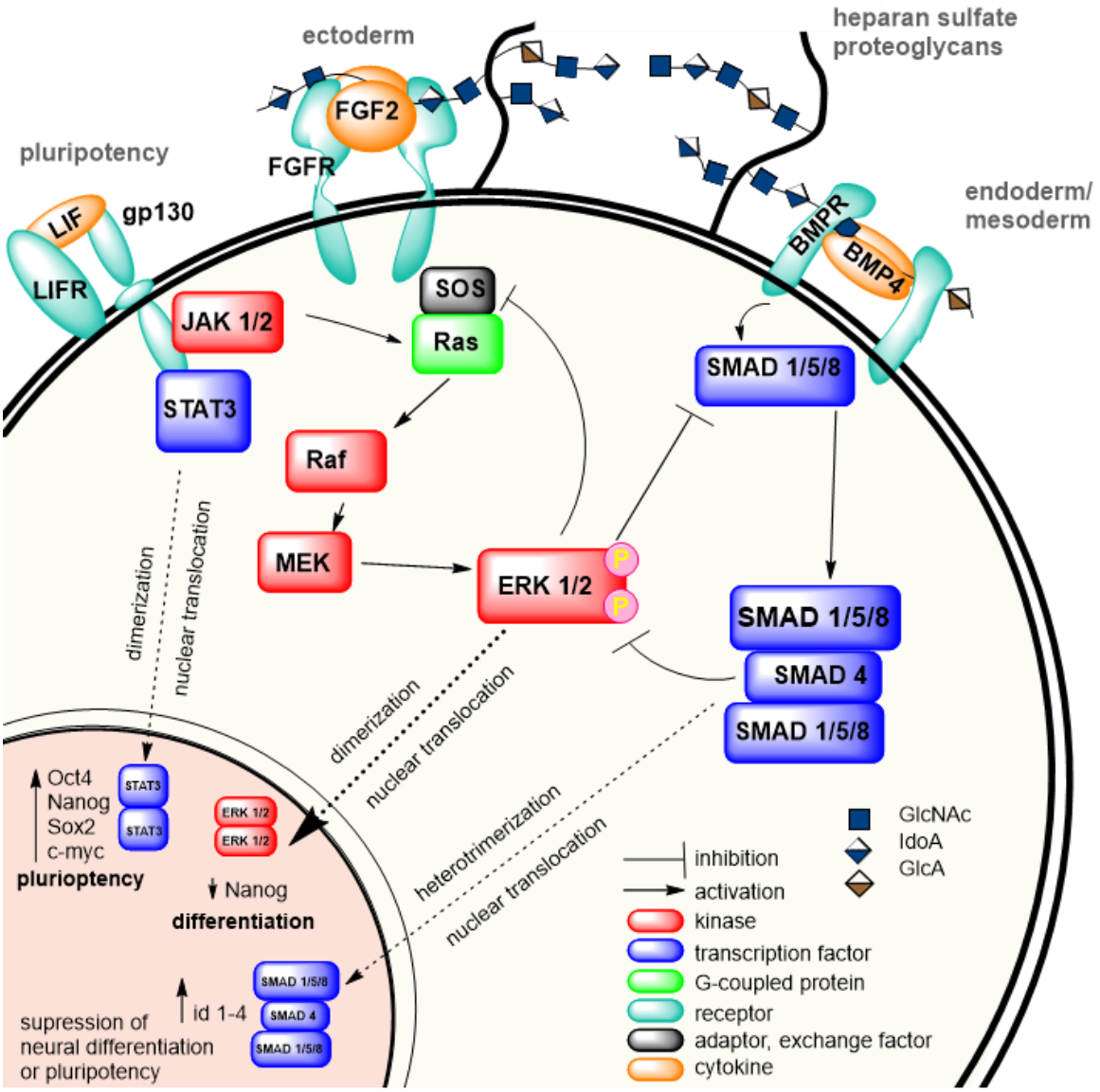


Figure 1.6 Pathways affecting pluripotency and differentiation in mESCs. Several critical signaling pathways affect mESC fate. LIF is the primary signaling cascade which maintains cells in a pluripotent state, while FGF2 initiates the MAPK cascade to initiate differentiation. FGF2 will lead cells into neuroectodermal lineages, while BMP4 driven signaling promotes endo/mesodermal fates.

1.6 Tools to manipulate and study the glycome

Another reason our understanding the glycans needs development is the lack of tools for their study. Glycans are often enzymatically released from proteins prior to crystallographic studies, and structures with glycans intact and underrepresented in structural protein data bases.[36,37] Furthermore, glycans evade genomic as well at proteomic surveys due to their non-templated property of not being the direct product from genes, but rather glycan epitopes present are dependent on many factors[38]. The field of glycomics lags behind that of proteomics and genomics but is a rapidly growing field showing great promise. The United States national institute of health (NIH) has identified glycomics as field with great potential with need for advancement in the U.S.[39]

However, at the time of this writing, the field of glycomics is rapidly expanding, and coming into focus. One of the most transformative tools for the field of glycomics has been the glycan microarray, a platform for rapidly interrogating glycan interactions.[40] Glycan microarrays typically consist of a solid glass substrate, which is a then robotically printed with a grid-like array individual glycans, and can subsequently be used for assessing individual glycan binding specificity between purified protein or virus.[41] While the microarray has significantly advanced glycoscience, the platform still suffers considerable limitations in that the three dimensional presentation of glycan structures must be precisely controlled for reproducible results, and care must be taken when interpretation of glycan binding preferences of purified proteins and application to real world scenarios

in complex environments.[42,43] Another key limitation, which is addressed in this dissertation (chapter VI), is the inaccessibility of profiling glycan binding properties between traditional solid phase microarrays to whole cells. These whole cell glycan binding analyses which have evaded systematic probing, with some notable exceptions.[44,45,46]

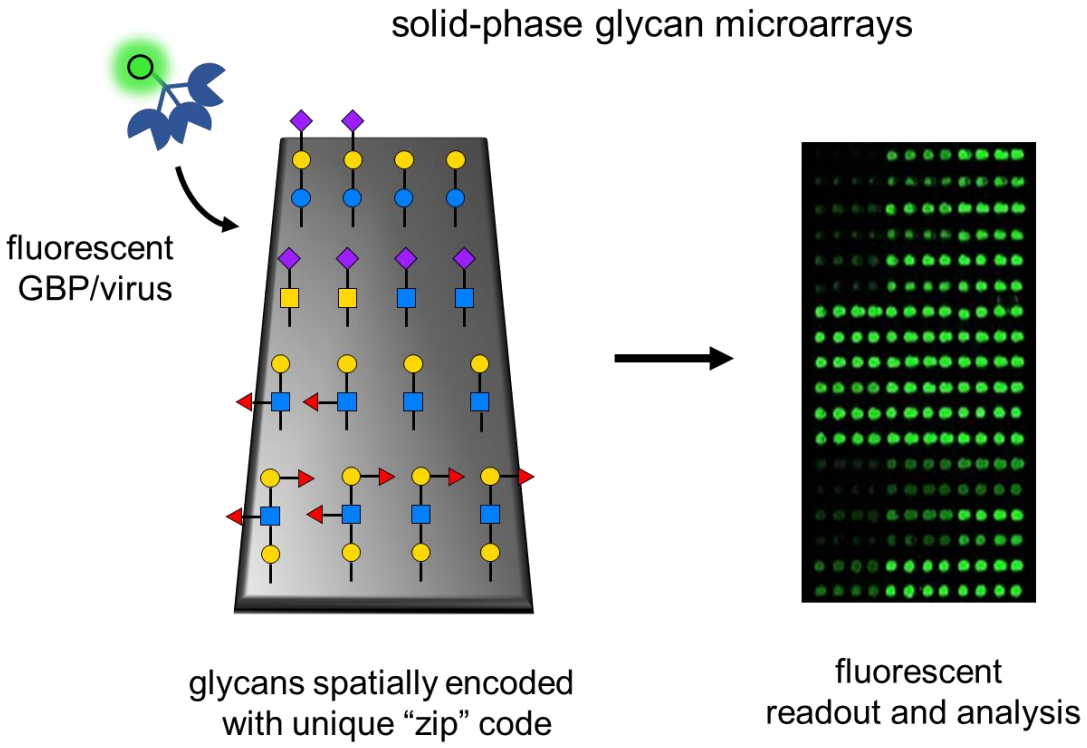


Figure 1.7 Schematic of solid phase microarrays. Glycans are robotically immobilized onto solid glass surface, and glycan binding proteins (GBPs) or virus are fluorescently tagged and washed over the array. Binding specificities are then measured by fluorescent intensity and are spatially decoded into the corresponding glycan structures.

As the field of glycomics grows, and the importance of the glycocalyx comes into focus, the glycome emerges as an attractive target for therapies and manipulating cellular outcomes. To provide a modern toolkit to scientists, progress necessitates methods to both influence the native function of glycans, as well as evaluate the biological outcomes

of glycan-based therapies. Altering glycan interactions transiently can be an attractive alternative to genetic manipulation of cell lines, which may lead to unwanted artifacts, and is unsuitable for regenerative therapies in humans.[29] To systematically evaluate the effects of glyco-therapies, screening tools will be required to investigate outcomes on the cellular level (see chapter 2). The main theme of this dissertation the development of tools to affect glycan interactions (chapter 3), methods to determine glycan binding interactions on whole cells (chapter 6), and methods to assess the resulting biological outcomes of glycan manipulation on stem cells (chapter 2).

1.7 Acknowledgements

Chapter 1 is an introductory background chapter written by Austen Larson Michalak. Figure 1.5 was partially modified from 1 as it appears in *Advanced Healthcare Materials* 2021. Michalak AL, Trieger GW, Trieger KA, Godula K., Stem Cell Microarrays for Assessing Growth Factor Signaling in Engineered Glycan Microenvironments.

1.8 References

1 Leonhard M. The Emerging Role of the Mammalian Glycocalyx in Functional Membrane Organization. *Front. In cell and developmental biology.* **2020**;8(253)

2 An HJ, Froehlich JW, Lebrilla CB. Determination of glycosylation sites and site-specific heterogeneity in glycoproteins. *Curr Opin Chem Biol*. **2009**;13(4):421-426

3 Van den Berg BM, Vink H, Spaan JAE. The Endothelial Glycocalyx Protects Against Myocardial Edema. *Circ. Res.* **2003**; 92(6):592-594

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

5 Gulati K, Poluri KM Mechanistic and therapeutic overview of glycosaminoglycans: the unsung heroes of biomolecular signaling. *Glycoconj J., **2016***;33:1–17

6 Markowska AI, Liu FT, Panjwani N. Galectin-3 is an important mediator of VEGF- and bFGF-mediated angiogenic response. *J Exp Med*. **2010**;207(9):1981-1993

7 Ding K, Lopez-Burks M, Sánchez-Duran JA, Korc M, Lander AD. Growth factor-induced shedding of syndecan-1 confers glypican-1 dependence on mitogenic responses of cancer cells. *J Cell Biol.* **2005**;171(4):729-38.

8 Li Q, Park PW, Wilson CLO, Parks WC. Matrilysin Shedding of Syndecan-1 Regulates Chemokine Mobilization and Transepithelial Efflux of Neutrophils in Acute Lung Injury. *Cell,* **2002**;111(5):635-646

9 Reily C, Stewart TJ, Renfrow MB, Novak J. Glycosylation in health and disease. *Nat Rev Nephrol*, **2019**;15:346–366

10 Kesimer M, Ehre C, Burns KA, Davis CW, Sheehan JK, Pickles RJ. Molecular organization of the mucins and glycocalyx underlying mucus transport over mucosal surfaces of the airways. *Mucosal Immunol,* **2013**;6:379–392

11 Evans SV, Roger MacKenzie C. Characterization of protein-glycolipid recognition at the membrane bilayer. *J. Mol. Recognit*. **1999**;12(3):155-168.

12 Laine, RA. A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05×10^{12} structures for a reducing hexasaccharide: the *Isomer Barrier* to development of single-method saccharide sequencing or synthesis systems, *Glycobiology*, **1994**;4(6): 759–767

13 Varki A, Kornfeld S. Historical Background and Overview. **2017**. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 1.

14 Schnaar RL, Kinoshita T. Glycosphingolipids. **2017**. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 11.

15 Stanley P, Taniguchi N, Aebi M. N-Glycans. **2017**. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 9.

16 Seeberger PH. Monosaccharide Diversity. **2017**. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 2.

17 Brockhausen I, Stanley P. O-GalNAc Glycans. **2017**. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 10.

18 Zachara N, Akimoto Y, Hart GW. The O-GlcNAc Modification. **2017**. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 19.

19 Hascall V, Esko JD. Hyaluronan. **2017**. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 16.

20 Colley KJ, Varki A, Kinoshita T. Cellular Organization of Glycosylation. **2017**. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 4.

21 Gandhi NS, Mancera RL. The Structure of Glycosaminoglycans and their Interactions with Proteins. *Chem. Biol. Drug Des*. **2008**;72: 455-482

22 Kraushaar D, Dalton S, Wang L. Heparan sulfate: a key regulator of embryonic stem cell fate. *Biol. Chem.*, **2013**;394(6), 741-751

23 Lindahl U, Couchman J, Kimata K, et al. Proteoglycans and Sulfated Glycosaminoglycans. **2017**. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 17.

24 Proudfoot AEI, Johnson Z, Bonvin P, Handel TM. Glycosamnioglycan Interactions with Chemokines Add Complexity to a Complex System. *Pharmaceuticals*, **2017**;10(3)

25 Kogan G, Soltes Ladislav, Stern R, Gemeiner P. Hyaluronic acid: a natural biopolymer with a broad range of biomedical and industrial applications. *Biotechnol. Letters*, **2006**;29,17-25

26 Michalak AL, Trieger GW, Trieger KA, Godula K. Stem Cell Microarrays for Assessing Growth Factor Signaling in Engineered Glycan Microenvironments. *Adv. Funct. Mat.*, **2021**;11(4)

27 Mohammadi M, Oslen S, Goetz R. A protein canyon in the FGF-FGFR receptor dimer selects from an á la carte menu of heparin sulfate motifs. *Current opin. In Structural Biology.,* **2005**,15:506-516

28 Kraushaar D, Rai S, Condac E, Narin A, Zhang S, Yamaguchi Y, Moremen K, Dalton S, Wang L. Heparan Sulfate Facilitates FGF and BMP Signaling to Drive Mesoderm Differentiation of Mouse Embryonic Stem Cells. *J. of Biol. Chemistry.*, **2012**, 287,27:22691-22700

29 Österholm C, Barczyk MM, Busse M, Grønning M, Reed RK, Kusche-Gullberg M. Mutation in the Heparan Sulfate Biosynthesis Enzyme EXT1 influences Growth Factor Signaling and Fibroblast Interactions with the Extracellular Matrix. *The Jou. of Biolog. Chem.*, **2009,** 284(50): 34925-34943

30 Smith, A. Embryo-derived stem cells: Of Mice and Men. *Annual Review of Cell Biology*, **2001**;17:435-462

31 Weissman, I. Stem Cells: Units of Development, Units of Regeneration, and Units in Evolution. *Cell*; **2000**;100:157-168

32 Hanna J, Cheng AW, Saha K, Kim J, Lengner CJ, Soldner F, Cassady JP, Muffat J, Carey BW, Jaenisch R. Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *PNAS*, **2010**,107(20):9222-9227

33 Ginis I, Luo Y, Miura T, Brandenberger R, Gerecht-Nir S, Amit M, Hoke A, Carpenter MK, Itskovitz-Eldor J, Rao MS. Differences between human and mouse embryonic stem cells. *Developmental biol.*, **2004**; 269(2):15, 360-380

34 Hirai H, Karian P, Kikyo N. Regulation of embryonic stem cell self-renewal and pluripotency by leukemia inhibitory factor. *Biochem J.*, **2011**;438(1):11-23

35 Niwa H, Ogawa K, Shimosato D, Adachi K. A parallel circuit of LIF signaling pathways maintains pluripotency of mouse ES cells. *Nature Letters*. **2009**;460(2):118-122

36 Prestegard JH. A perspective on the PDB's impact on the field of glycobiology. *J Biol Chem*. **2021**;296:100556.

37 Nagae M, Yamaguchi Y. Function and 3D structure of the N-glycans on glycoproteins. *Int J Mol Sci*. **2012**;13(7):8398-8429

38 Pilobello KT, Mahal LK. Deciphering the glycode: the complexity and analytical challenge of glycomics. *Curr. Opin. In Chem. Biol.* **2007**;11:300-305

39 National Research Council (US) Committee on Assessing the Importance and Impact of Glycomics and Glycosciences. *Transforming Glycoscience: A Roadmap for the Future*. Washington (DC): National Academies Press (US); **2012**.

40 Park S, Gildersleeve J, Blixt O. Carbohydrate Microarrays. *RSC Chem. Soc. Rev.*, **2013** 42:4310-4326

41 Smith D, Song X, Cummings R. Use of Glycan Microarrays to Explore Specificity of Glycan-Binding Proteins. *Meth. In enzymol.* **2010** 480:417-444

42 Wang L, Cummings RD, Smith DF, Huflejt M, Campbell CT, Gildersleeve JC, Gerlach JQ, Kilcoyne M, Joshi L, Serna S, Reichardt NC, Parera PN, Pieters RJ, Eng W, Mahal LK. Cross-platform comparison of glycan microarray formats. *Glycobiol.* **2014**, 24(6):507-517

43 Pochechueva T, Jacob F, Goldstein D, Huflejt M, Chinarev A, Caduff R, Fink D, Hacker N, Bovin NV, Heinzelmann-Schwartz V. Comparison of printed glycan array, suspension array, and ELISA in the detection of human anti-glycan antibodies. *Glycoconj. J.* **2011**, 28:507-517

44 Jonczyk R, Kurth T, Lavrentieva A, Walter JG, Scheper T, Stahl F. Living Cell Microarrays: An Overview of Concepts. *Microarrays*. **2016**, 5:1-29.

45 Purohit S, Li T, Guan W, Song X, Song J, Tian T, Li L, Sharma A, Dun B, Mysona D, Ghamande S, Rungruang B, Cummings RD, Wang PG, She J. Multiplex glycan bead array for high throughput and high content analyses of glycan binding proteins. *Nature comm.* **2018**. 9:258

46 Disney M, Seeberger P. The Use of Carbohydrate Microarrays to Study Carbohydrate-Cell Interactions and to Detect Pathogens. *Chem. & Biol.* **2004**.11:1701-1707

2. Stem Cell Microarrays for Assessing Growth Factor Signaling in Engineered Glycan Microenvironments

2.1 Introduction

The development of biologically active materials that support cell adhesion and proliferation, while also providing signaling cues to guide cellular differentiation, has enabled the translation of the regenerative capacity of stem cells into clinical applications.[1,2] The integration of various components of the native extracellular matrix into hydrogels has emerged as a major strategy for generating responsive materials for organoid and tissue engineering.[3,4] Comprised of hydrated synthetic or biological polymer networks, hydrogels are commonly decorated with peptides or proteins for cell adhesion and supplemented with signaling molecules, such as growth factors (GFs), to promote signaling and differentiation toward desirable cell types.[5,6,7]

Stem cell arrays, which allow for high-throughput analysis of cellular responses to their environment and culture conditions, have enabled the discovery and optimization of new biomaterials for cell-based applications.[8] Such platforms have been particularly useful for examining the ability of various protein components of the ECM to enhance cell interactions and functions when introduced into hydrogels.[9,10] Extracellular glycans, which also provide important biological functions in the ECM but are difficult to access in pure form synthetically or through isolation, have been comparatively less explored as components for biomaterials.[11,12] For example, extracellular heparan sulfate (HS)

polysaccharides, which belong to the family of glycosaminoglycans (GAGs) (Figure 2.1), are essential regulators of GF signaling and are being pursued as biologically active components of hydrogels for stem cell culture and tissue engineering. [13] HS polysaccharides comprise chains of alternating N-acetylglucosamine and glucuronic acid (GlcA) residues, which undergo sequential enzymatic modifications to introduce N-sulfation and to partially epimerize GlcA into iduronic acid (IdoA).[14] Additional O-sulfation is then introduced to produce sulfated domains harboring protein binding motifs. The compositional complexity of HS has made systematic structure-function analysis needed for their integration into biomaterials challenging. Recent advances in chemical[15,16] and chemoenzymatic[17,18] HS oligosaccharide synthesis as well as genetic engineering[19] of HS biosynthetic pathways have produced increasingly large numbers of chemically well-defined HS structures available for examination in the context of biomaterial design.

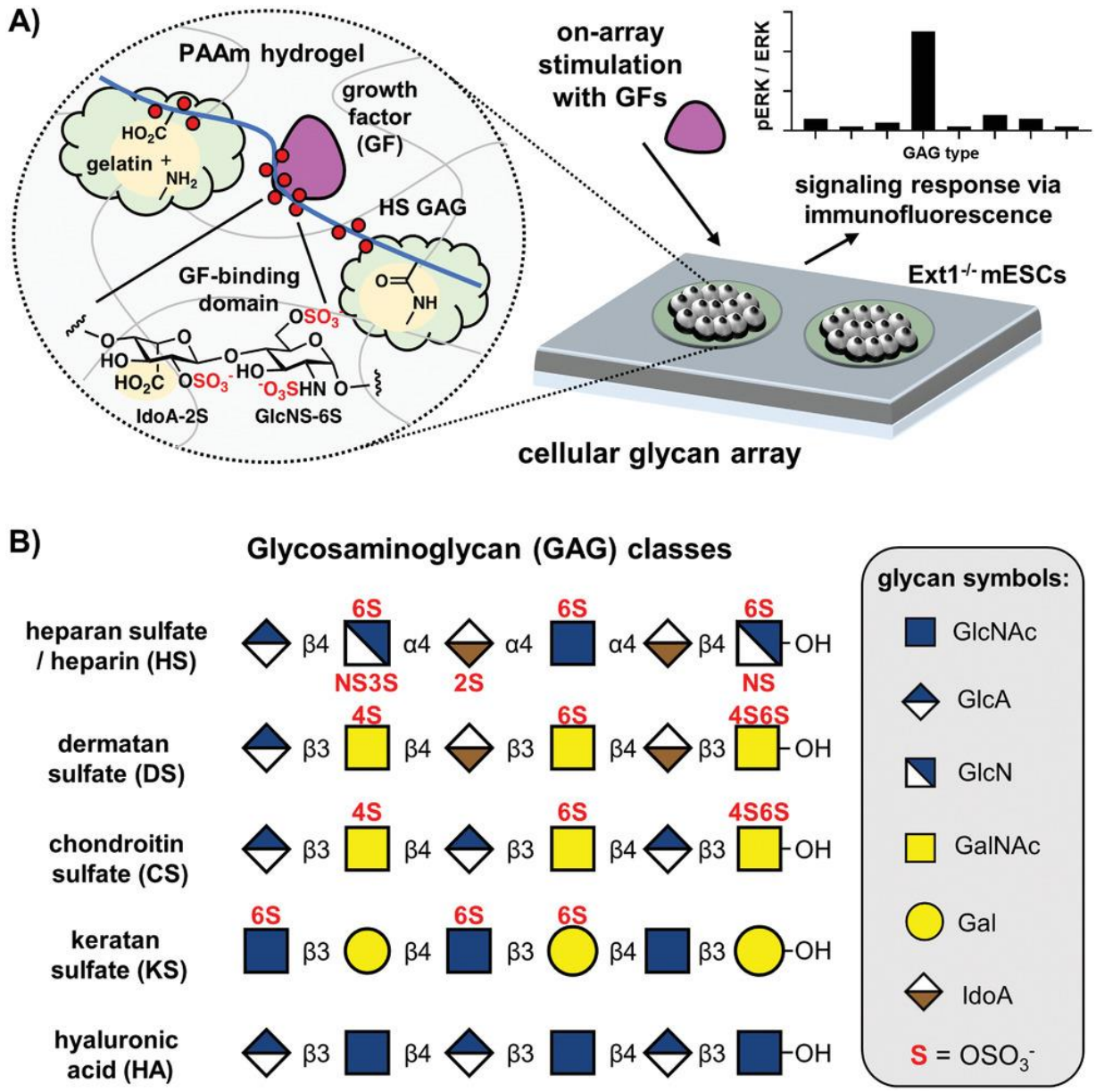


Figure 2.1 Stem cell array for rapid analysis of growth factor signaling in engineered glycosaminoglycan (GAG) microenvironments. A) GAGs covalently crosslinked to gelatin are arrayed and immobilized within a polyacrylamide (PAAm) hydrogel substrate. Signaling responses of embryonic stem cells grown on the GAG array after GF stimulation are assayed directly by immunofluorescence. B) Structures representing the five main families of GAG polysaccharides depicted using the symbol nomenclature for glycans (SNFG) notation.

Arrays comprising isolated or synthetic HS structures immobilized on glass surfaces are routinely used to profile the specificity of HS-binding proteins.[20] Platforms that enable multiplexed, on-array analysis of HS-dependent cellular signaling could significantly streamline the discovery of biomaterials that capitalize on the regulatory functions of ECM glycans. An early example of arrays being used to evaluate the effects of HS structures on cellular responses came from Linhardt and co-workers, who studied proliferation of hydrogel encapsulated non-adherent Ba/F3 cells in the presence of chemically defined HS polysaccharides and fibroblast growth factors (FGFs).[21] The cell-laden hydrogel droplets were printed on glass and exposed to combinations of HS and FGFs as soluble media supplements. Turnbull and his co-workers were able to directly observe activation of the mitogen activated protein kinase (MAPK) signaling pathway after FGF2 stimulation in Swiss 3T3 cells grown on arrays of oligosaccharides derived by partial heparin digestion. The cells were grown as a monolayer on HS oligosaccharides of increasing length (degree of polymerization, DP = 2-18) spotted and covalently immobilized on amine-functionalized glass via reductive amination.[22] MAPK activation was quantified by immunostaining for phosphorylation of Erk1/2 kinases and the magnitude of the observed signal scaled with oligosaccharide length.

To fully harness the multiplexing potential of these the array platform, strategies are needed to present HS structures to progenitor cells in a spatially isolated, yet addressable, format. Here, we present a method for the generation of hydrogel-based GAG microarrays for analysis of growth factor-mediated signaling in murine embryonic

stem cells (ESCs). By arraying HS-protein conjugates on polyacrylamide hydrogels, we were able to generate stable ECM-mimetic microenvironments with the capacity to bind FGF2 and influence ESC signaling. The binding and activity of FGF2 in these cellular microenvironments was defined by the chemical composition of the HS polysaccharides.

2.2 Results

To develop an array for assessing stem cell signaling responses to engineered glycan ECM environments, we sought to present the glycans together with cell adhesion factors in microscopic islands separated by a non-adhesive surface. This would enable multiplexed analysis of a range of glycan structures while minimizing cellular crosstalk. After screening several common surface passivation strategies used for array construction (Figure 2.2), we found that a thin poly(acrylamide) hydrogel deposited on glass according to a method by Brafman et al.[23] and spotted with a solution of gelatin (500 μg mL^{-1}) in PBS best supported ESC growth in well-separated colonies over 6 days in culture. The gelatin, a commonly used substrate for murine ESC culture, was loaded into the hydrogel in its dehydrated form, which allowed the protein to enter and become entrapped within the crosslinked polymer network upon rehydration.

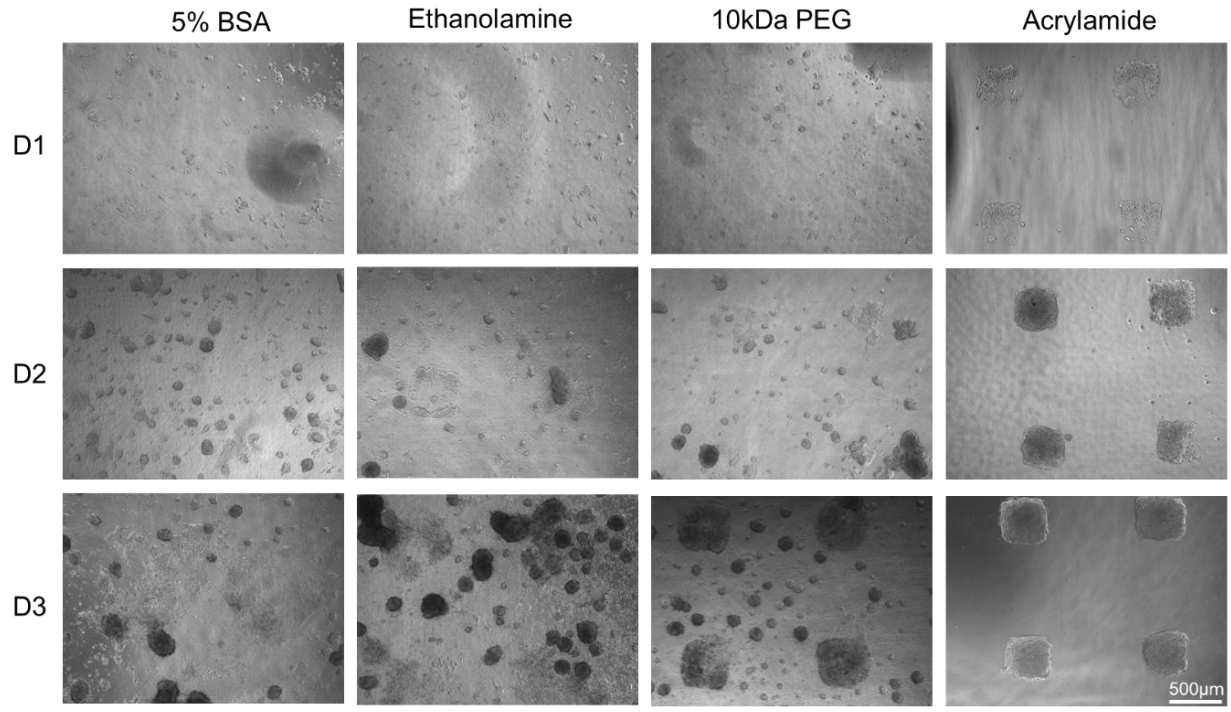


Figure 2.2 Microarray surface passivation To generate a spatially segregated cellular microarray capable of multiplexed assays, several passivation strategies were tested. Epoxide functionalized slides (Thermofisher) were passivated using established passivation protocols[23],24, printed with 0.5 mg/mL gelatin in the absence of crosslinked GAGs, and tested by seeding $4x10^4$ $Ext1^{-/-}$ ESCs and visually monitoring growth over three days using. For passivation, Slides were incubated at room temperature overnight with either 5 % Bovine Serum Albumin (BSA) solution, 100 mM ethanolamine in pH 8.5 borate buffer. For 10kDa PEG, the slides were subjected to 10 mg/mL of 10 kDa amino-peg in PBS containing 62 mM K_2SO_4 overnight at 37 °C. Acrylamide slides were prepared as described, using established procedures. Acrylamide surface passivation afforded the most complete passivation of the surface with little unrestricted cellular growth. (D1 = Day 1, scale bar = 500 µm)

To test whether GAG polysaccharides may similarly be arrayed and retained within the hydrogel, the dry acrylamide substrates were spotted with gelatin solutions (500 µg mL^{-1}) in PBS buffer (10% glycerol, 0.003% triton X-100) supplemented with increasing

concentrations (50–750 μg mL^{-1}) of heparin (12 kDa) as a model HS glycan (Figure 2.3). Anticipating that the polysaccharide may diffuse out of the hydrogel network under cell culture conditions, we also included conditions where the heparin was crosslinked via its carboxylic acid groups activated in the form of N-hydroxysuccinimide (NHS) esters to the solvent exposed lysine residues in gelatin (Figure 2.1A). The heparin was activated by treatment with NHS in HEPES buffer (100 × 10−3 m, pH = 7.4) in the presence of the coupling reagent, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC), at 4 °C for 18 h and purified by size exclusion on a PD10 column to remove small molecule reagents and byproducts. We targeted low levels of crosslinking (≈15% carboxylic acid crosslinks per chain) to promote the retention of the GAG in the hydrogel network without compromising its GF-binding ability; however, we anticipate the actual frequency of crosslinks to be below the target value due to competing hydrolysis of the activated NHS ester groups in the heparin chains under the reaction conditions.

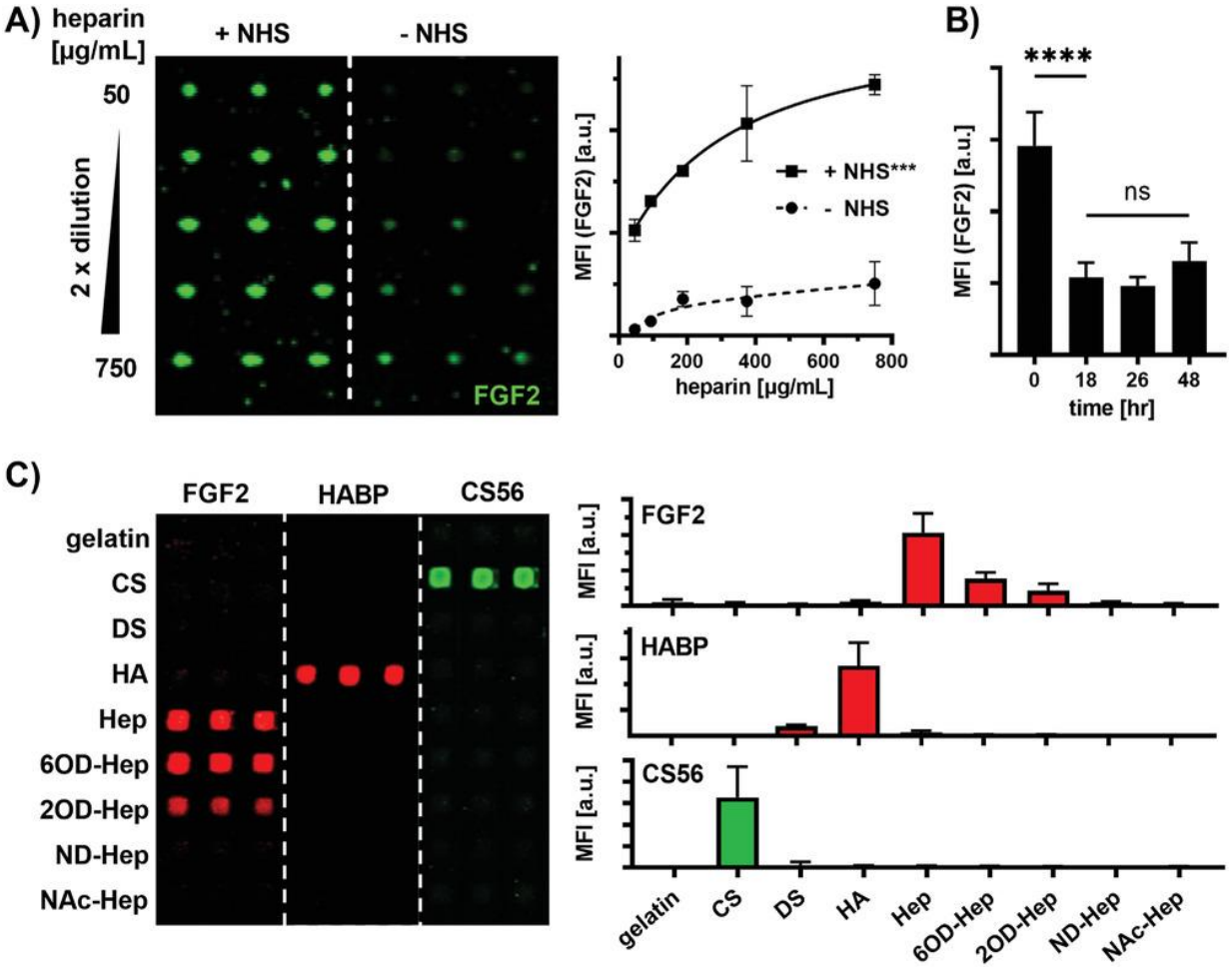


Figure 2.3 Generation and characterization of polyacrylamide-gelatin GAG array substrates for stem cell culture. A) Fluorescence micrographs and graph representations of FGF2 binding to heparin (Hep) on polyacrylamide-gelatin arrays printed at increasing Hep concentrations (cHep = 50–750 mg mL^{-1}) with or without N-hydroxy succinimide (NHS) crosslinking to gelatin. Assessment of FGF2 binding after 3-h wash in PBS buffer indicates ≈five to tenfold increase in Hep retention after crosslinking. B) FGF2-binding to crosslinked Hep arrays was assessed over 48 h under cell culture conditions. After initial decrease over the first 18 h, the arrays retain ≈40% of FGF2 binding activity for up to 48 h. C) FGF2, HA-binding protein (HABP), and the anti-CS antibody, CS56, bind selectively to CS, HA, and chemically desulfated heparinoids on the array. FGF2 bound to Hep and chemically desulfated heparinoids in the following order: Hep > 6OD-Hep > 2OD-Hep >> ND-Hep ≈ NAc-Hep. Data represent the mean and standard deviation, n = 3 (and C) or 10 (B). Differences between + NHS and – NHS conditions (A) were assessed using a two-sided t-test, df = 8. Statistical analysis in (B) was performed using one-way ANOVA, p-values were determined by Tukey's multiple comparisons test. (*p < 0.05, **p < 0.01, ***p < 0.001 ****p < 0.0001, ns = not significant).

The resulting arrays were washed for 3 h and probed with the heparin-binding FGF2 protein labeled with AlexaFluor 647 (FGF2-AF647) to detect the immobilized heparin and to assess the effect of crosslinking on its retention in the hydrogel (Figure 2.3A). Both immobilization strategies resulted in a dose-dependent FGF2 binding, with the NHS-crosslinked heparin providing ≈five- to tenfold higher signal. To further test the stability of the arrays, acrylamide substrates spotted with the NHS-activated heparin (500 μg mL^{-1}) in gelatin (500 μg mL^{-1}) were subjected to ESC culture conditions for 48 h. Over this time, the arrays were probed with FGF2-AF647 to assess heparin retention (Figure 2.3B) After an initial decrease in binding activity during the first 18 h, the heparin arrays remained stable with ≈40% of FGF2 binding activity being retained after 48 h.

With a suitable method for heparin immobilization on the hydrogel substrates in hand, we aimed to test that the protein binding specificities of the crosslinked GAG structures within the hydrogel matrix are preserved (Figure 2C). Using the NHS-crosslinking strategy, we arrayed a panel of CS, DS, and HA GAGs as well as heparin polysaccharides chemically treated to selectively remove their 6-O-, 2-O-, and N-sulfates (6OD-Hep, 2OD-Hep, and ND-Hep, respectively). N-desulfated heparin, in which the exposed amino groups were capped as acetamides (NAc-Hep) to better represent native HS structures, was also included. The array was then probed with FGF2, a CS-specific antibody (CS56), and the hyaluronic acid binding protein (HABP). As shown in Figure 2C, FGF2 bound most strongly to the fully sulfated heparin. Removal of 6-O-, 2-O- and N-sulfates resulted in progressive loss of activity, which is in agreement with the known

requirements of 2-O- and N-sulfation for FGF2 binding to HS.[25,26,27] Likewise, CS56 and HABP proteins exhibited high specificity for CS and HA, respectively.

Having confirmed that NHS-crosslinking to the gelatin matrix enhances the stability of the GAG displays without altering the protein binding specificity of the polysaccharides, we set to evaluate the ability of these arrays to support ESC culture. For our cell model, we chose murine ESCs lacking the expression of Exostosin 1 (Ext1), which is a glycosyl transferase responsible for the assembly of HS chains.[28] In the absence of this enzyme, the Ext1$^{-/-}$ ESCs lack cell surface HS structures and are unable to engage a range of HS-dependent GFs, including FGF2.[29] As such, these mutant ECS are ideally suited to isolate the effects of the arrayed GAGs on FGF2 signaling from those of endogenous HS structures.

The envisioned on-array GF signaling assay would require that the cells formed monolayer colonies on the printed heparin-gelatin spots over at least 2 days in culture. In order to suppress endogenous GF production and establish signaling activity baseline, the last 24 h should be carried out under serum-free conditions. To optimize cell density and colony growth on the array, we seeded increasing number of Ext1$^{-/-}$ ESCs on the substrates and grew them for 24 h in embryonic culture media supplemented with leukemia inhibitor factor (LIF) and fetal bovine serum. The arrays were then washed, and the remaining bound cells were cultured for additional 24 h in the absence of serum (Figure 2.4A). While seeding the stem cells too sparsely (20,000 cells cm^{-2}) resulted in

slow growth and irregular colony formation, too high seeding density (100,000 cells cm^{-2}) led to rapid proliferation resulting in spot overgrowth and cell detachment. The intermediary seeding density (40,000 cells cm^{-2}) produced consistent monolayers of $Ext1^{-/-}$ ESCs (Figure 2.4A), which retained high levels of expression of the embryonic marker, Oct4, and showed no obvious signs of differentiation (via the neural marker, Nestin, Figure 2.4B). The optimized seeding conditions were further tested in the presence of immobilized heparin printed at 500 µg mL^{-1} concentration and under serum-free starvation conditions to ensure no negative effects of these conditions on cell adhesion and growth (Figure 2.4C).

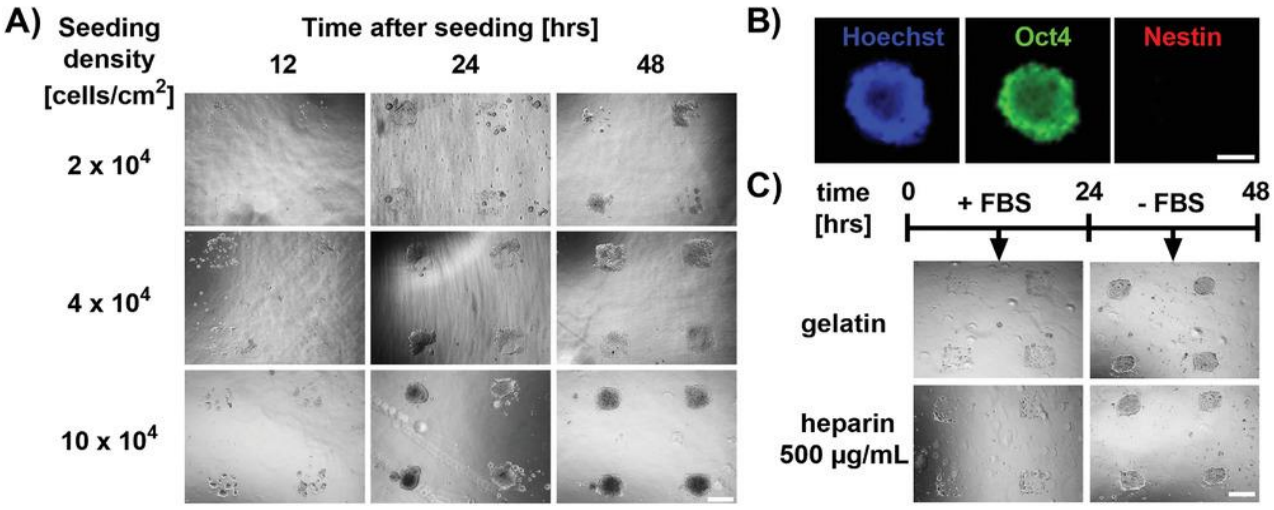


Figure 2.4 Optimization of conditions for embryonic stem cell (ESC) culture on GAG microarrays. A) $Ext1^{-/-}$ ESCs were seeded at densities of 2 × 10^4, 4 × 10^4, 10 × 10^4 cells cm^{-2} on poly(acrylamide) substrates printed with gelatin (0.5 mg mL^{-1}). The cells were assessed for growth and colony morphology over 48 h by optical microscopy (scale bar = 500 µm). B) $Ext1^{-/-}$ ESCs were cultured on gelatin arrays in embryonic media containing LIF for 48 h. The cells retained high levels of pluripotency (Oct4) with no significant spontaneous neural differentiation (Nestin) (scale bar = 200 µm). C) Immobilized heparin (500 µg mL^{-1}) does not significantly alter $Ext1^{-/-}$ ESCs adhesion, growth, or colony formation on the array. Cells were seeded at 4 × 10^4 cells cm^{-2} and cultured for 48 h under embryonic conditions (+ LIF), with the last 24 h under serum-free conditions to minimize autocrine GF signaling activity. (scale bar = 500 µm).

To establish whether the array format is suitable for directly assessing changes in stem cell signaling in the engineered glycan microenvironments, we chose to examine the activation of the MAPK pathway in response to stimulation with exogenous FGF2. The requirement for HS in the formation of a signaling complex between FGF2 and its receptor, FGFR, has been well established and the signaling response is accompanied by well-characterized changes in the phosphorylation status of downstream kinases (i.e., Extracellular regulated kinase 1 and 2, Erk1/2).[30,31,32]

For the on-array FGF2 signaling assay, Ext1$^{-/-}$ ESCs (40,000 cells cm^{-2}) were seeded on spots printed with gelatin (500 μg mL^{-1}) with or without NHS-crosslinked heparin (500 μg mL^{-1}) and grown for 48 h under the optimized embryonic culture and starvation conditions (Figure 2.5A). The cells were then placed in a fresh serum-free media containing FGF2 (0.5 ng mL^{-1}) and stimulated for 15 min at 37 °C. The cells were fixed, permeabilized, and immunoassayed for MAPK activity using antibodies against Erk1/2 proteins and their phosphorylated forms (pErk1/2). Fluorescence from the arrayed cells was detected using a microarray scanner (Figure 2.5B) and validated via fluorescence microscopy (Figure 2.5C). The ratios of fluorescent signals corresponding to the phosphorylated-ERK1/2 (pERK, green) and total ERK1/2 (ERK, red) proteins was used to quantify the signaling response (Figure 2.5B). We used soluble heparin (s-Hep, 5 μg mL^{-1}), which is known to restore MAPK activity in Ext1$^{-/-}$ ESCs, as a positive control and a benchmark in our assay.[19] While ERK1/2 protein levels were similar across all conditions, only Ext1$^{-/-}$ ESCs stimulated with FGF2 in the presence of immobilized or soluble heparin showed significant increase in Erk1/2 phosphorylation (Figure 2.5B). We

performed fluorescence microscopy imaging (Figure 2.5C) and image J analysis to confirm the co-localization of the pERK and ERK signals and to validate our quantification scheme based on signal detection via microarray scanner. We observed somewhat lower levels of MAPK activity on the arrayed heparin compared to its soluble form (Figure 2.5B). This may be due to a more limited accessibility of the immobilized heparin to only a subset of FGFRs localized to the point of cell contact with the array. FGF2 stimulation of Ext1$^{-/-}$ ESCs grown on gelatin spots containing increasing amounts of immobilized heparin (i-Hep, 0–500 μg mL^{-1}) showed a heparin dose-responsive ERK1/2 phosphorylation (Figure 2.5D).

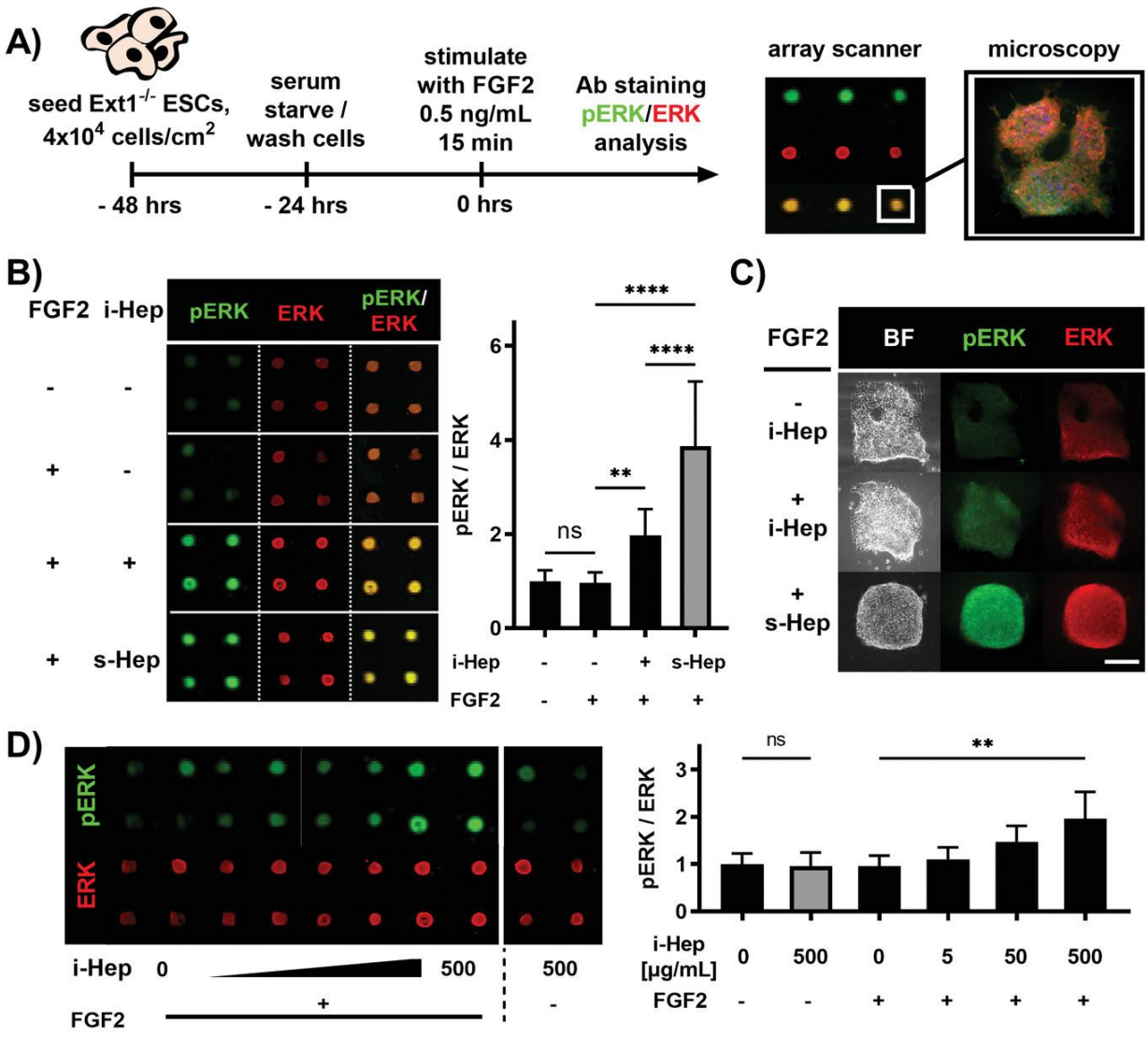


Figure 2.5 Analysis of MAPK signaling in Ext1−/− ESCs on heparin array. A) *Ext1*$^{-/-}$ ESCs were seeded (4 × 10^4 cells cm^{-2}) and grown under embryonic conditions (+ LIF) for 48 h, with the last 24 h under serum starvation. The cells were then stimulated with FGF2 (0.5 ng mL^{-1}) for 15 min. Levels of pERK and total ERK were assessed via immunofluorescence using an array scanner or microscopy. B) Fluorescence images and bar graph representations of arrays stained with anti-pERK (green) and anti-ERK (red) antibodies. Enhanced ERK phosphorylation in response to FGF2 stimulation was observed only in the presence of immobilized or soluble heparin (i-Hep or s-Hep). C) Fluorescence micrographs of ESC colonies after FGF2 stimulation analyzed by microscopy. (Scale bar = 250 µm). D) Dose response in FGF2 stimulated ESCs grown on arrays printed at increasing concentrations of heparin (c_{Hep} = 0–500 mg mL^{-1}). No significant increase of MAPK signaling was observed in colonies cultured on immobilized heparin in the absence of FGF2 (500 mg mL^{-1}, gray bar). Data and error bars represent mean values and standard deviations for at least 11 replicate colonies per condition. Statistical analysis was performed using one-way ANOVA, *p*-values were determined by Tukey's multiple comparisons analysis. ($^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$, ns = not significant).

2.3 Conclusions

We have developed an array platform for direct, rapid and multiplexed profiling of extracellular glycosaminoglycan activity on growth factor signaling in live embryonic stem cells. The arrays were generated by printing and physisorption of GAG polysaccharides chemically crosslinked with extracellular matrix proteins onto polyacrylamide hydrogel substrates. The crosslinking facilitated the retention of the polysaccharides on the array during cell culture without altering the protein-binding specificity of the polysaccharides. The immobilized GAG structures were able to facilitate GF-mediated activation of signaling events in live embryonic stem cells which could be detected and quantified using immunofluorescence. This array offers a convenient platform to systematically analyze biological activities of extracellular GAGs in stem cell signaling and to accelerate the development of new bioactive glycomaterials for stem cell-based therapeutic applications by capitalizing on the rapidly expanding repertoire of available synthetic[16], chemoenzymatic[17,18] and recombinant glycosaminoglycan structures.[19]

2.4 Experimental

General Chemistry Procedures

All chemicals, unless stated otherwise, were purchased from Sigma Aldrich. Purchased starting materials were used as received unless otherwise noted. Heparin and desulfated heparinoids were purchased from Iduron (Manchester, UK). The selectively desulfated heparinoids originated from the unmodified heparin used in this study. Iduron reported the average molecular weight of the parent heparin as 12,000 g mol^{-1}. Solvent compositions are reported on a volume/volume (v/v) basis unless otherwise noted.

Instrumentation

Nuclear magnetic resonance (NMR) spectra were collected on a Bruker 300 MHz NMR spectrometer. Spectra were reported in parts per million (ppm) on the δ scale relative to the residual solvent as an internal standard. Brightfield images of live cells were taken using ZEISS Axio Observer microscope. Fixed cells were fluorescently imaged using a Keyence BZX-700 fluorescent microscope. Microarray slides used for protein binding assays were assessed using an Axon GenePix 4000B microarray scanner (Molecular Devices). All microarray experiments were performed using steel spring ProPlate gaskets (Gracebiolabs), which were attached to the array slide.

Preparation of NHS-Activated Glycosaminoglycans

A HEPES buffer (100×10^{-3} M, 15.0 mL, pH 7.4) was prepared. Then, GAG (0.17 mmol) was dissolved in HEPES buffer (200 µL), 4 equiv. per GAG chain (0.68 mmol) of NHS was added via an aliquot from a stock NHS solution in HEPES and stirred overnight at 4 °C to afford heparin NHS-ester. The activated NHS ester solution was diluted in MQ water, loaded onto a PD-10 column, and eluted with 2.5 mL Milli-Q water (MQ H_2O). The solution was lyophilized to afford intermediate heparin NHS-ester (2.0 mg). The resulting heparin-NHS product was dissolved in a PBS solution containing gelatin (10.0 mg mL^{-1}, 200 µL). The reaction was allowed to proceed overnight. The crosslinked glycosaminoglycan product was purified through a PD-10 column (2.5 mL loading volume, 2.5 mL elution volume). The resulting solution was lyophilized to afford purified gelatin containing crosslinked glycosaminoglycans as a white solid. The same stoichiometry was used for all glycosaminoglycan conjugates.

Glass Slide Cleaning

Untreated 25 × 75 mm glass microscope slides were loaded into a steel slide rack and submerged in a crystallization dish filled with MQ H_2O. The slide rack was washed five times with water, allowing the slides to remain in the last water wash for 30 min on a rocker. After 30 min, the water was removed and replaced with acetone. This solution rocked for 30 min, covered. The acetone was then removed and replaced with methanol, was once more rocked for 30 min, covered. The methanol was replaced with a solution of NaOH (0.05 M), and rocked for 2 h. Slides were then rinsed three times in MQ H_2O

and subsequently spin-dried. The slides were then lightly blow-dried using 0.22 µm filtered air. Once dried, slides can then be placed into a vacuum oven to dry at 70 °C and safely stored for up to a month.

Glass Slide Silanization

Dried, etched slides in a steel rack were placed into a solution of 3-(trimethoxysilyl) propyl methacrylate in toluene (2 %), and rocked for 1 h. The solution was then removed and the slides were washed three times in fresh toluene to remove residual 3-(trimethoxysilyl) propyl methacrylate. The slides were spin-dried, then blow-dried with 0.22 µm filtered air and placed in a desiccator overnight. For glutaraldehyde activation, slides were rocked for 2 h in a solution of glutaraldehyde H_2O (0.05 %). The slides were spin-dried (500 rpm, 5 min), blow-dried with 0.22 µm filtered air, and placed in a desiccator overnight.

Deposition of Acrylamide Hydrogel on Glass Slides

An aqueous 30 % acrylamide solution was prepared by addition of acrylamide (2.85 g), to acrylamide/bisacrylamide (0.150 g, 19:1) to H2O (10 mL). A separate solution of ammonium persulfate (APS) solution (10 % w/v) was prepared in H_2O. The polymerization solution was then prepared by combining MQ H2O (985 µL), Acrylamide solution (500 µL), APS solution (15 µL), and tetramethylethylenediamine (0.6 µL). Immediately reagent addition, aliquots (110 µL) of the polymerization solution were placed in the center of each glutaraldehyde-activated methacrylate slide, and cover slip

was placed on each slide. After 2 h, and the slides were loaded into a steel slide rack with the coverslips still on, and allowed to sit in H_2O for 15 min, causing the coverslip to loosen on the slide as the hydrogel expands. Slides were removed from water, and using a razorblade, the coverslips were gently removed. The hydrogel exposed slides were then carefully reloaded into the steel slide racks and submerged in a crystallization dish filled with H_2O. The H_2O was replaced every 24 h for a total of 48 h of washing. After the 48-h wash, the slides were spin-dried and placed hydrogel-side-up onto a slide warmer (50 °C, 10 min) to partially dehydrate slides for storage.

Printing of GAG Arrays

Microarrays were printed using a SpotBot Extreme microarrayer (ArrayIt). Arrays were printed in 65% humidity using 500 µm spot pins. While the number of spots varies from array to array, spacing between spots was consistently 1400 µm in arrays used for cellular culture. For protein binding assays, spots were spaced 750 µm apart. When designing the spot layout, the print parameter option MAUI4 was selected, and the lateral and vertical offset were 1 and 3 mm respectively. Arrays were printed with porcine gelatin (0.5 mg mL^{-1}, bloom 180) in PBS supplemented with glycerol (10 %) and triton X-100 (0.03 %). When concentration gradients were printed, the lowest concentration was always printed last. The gelatin was printed at 500 µg mL^{-1}.

After printing, slides were placed into a slide holder and allowed to dry overnight at 4 °C. Prior to use, slides were washed in MQ H_2O for 2 min by loading slides into a

steel rack and rapidly and repeatedly dipping slides into a crystallization dish full of MQ H_2O. After this, slides were washed in PBS (3×, 15 min) and then spin-dried. Following this, the slides were snap dried by placing the cells array-side-up onto a slide warmer (50 °C, 10 min).

Growth Factor Binding on Arrays

Microarrays used for protein binding assays were equipped with a 4-well gasket chamber and then blocked for 45 min with a filtered PBS solution containing bovine serum albumin (BSA) (250 µL, 1 % BSA, 0.5 % tween-20). After blocking, protein binding incubations were performed at 4 °C for 90 min in blocking solution with AF647-FGF2 (10×10^{-9} M). Between protein incubations, wells were washed four times with blocking solution. After all incubations, a final series of PBS washes (3×, 15 min) were performed, the slide was spin-dried, and scanned using a microarray scanner.

Preparation of AF647-FGF2

Human FGF2 (100 µg) was dissolved in of 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer (200 µL, 200×10^{-3} M, pH 8.4). Then, a heparin solution (20 µL, 20 mg mL^{-1}) in MQ H_2O was added to the FGF2 solution and incubated for 10 min, after which a NHS-AF647 solution (2 µL, 10 mg mL^{-1}) in DMF was added to the solution. The reaction was gently rocked for 3 h, and was quenched by the addition of glycine solution (80 µL, 20 mg mL^{-1}) in MQ H_2O. To purify the reaction, a heparin sepharose column (1 mL) was prepared and used using a wash buffer (0.5 M

NaCl, 0.2 % BSA, 20 × 10^{-3} M HEPES, pH 7.4) followed by an elution buffer (3 M NaCl, 2% BSA, 20 × 10^{-3} M HEPES, pH 7.4). The column was equilibrated with wash buffer, loaded and rinsed with 5 column volumes of wash buffer and eluted to yield purified FGF2-AF647.

ESC Culture

Ext1$^{-/-}$ mouse embryonic stem cells were a gift from Dr. Cathy Merry, University of Nottingham, UK. ESCs were cultured feeder free in treated plastic well plates (5 % CO_2, 37 °C). Cells were cultured in ESC maintenance media consisting of Knockout-Dulbecco's modified eagle medium (KO-DMEM) supplemented with fetal bovine serum (10 %), non-essential amino acids, L-glutamine, 2-mercaptoethanol and LIF. Serum free media was of identical composition to ESC maintenance media except for the exclusion of FBS. Cells were passaged every other day and spilt at a ratio of 1:10 (1 x 10^5 cells per well).

Sterilization of Arrays

Arrays were sterilized for cell culture in a laminar flow tissue culture hood by placing arrays and autoclaved gaskets in ethanol for 5 min, followed by a sterile PBS wash. The gasket was then assembled onto the slide, and the slide was washed with PBS and left under UV light for at least 15 min, this was repeated twice, each time with fresh PBS.

Seeding Arrays

At least 20 min before seeding, LIF containing ESC maintenance media (500 µL) was added to each well of a 4 well gasket attached to an acrylamide slide with an array printed upon it. Cells were then seeded onto the array in a volume of 1 mL, bringing the final volume to 1.5 mL. At 24 h after seeding, the outlines of gelatin spots became noticeable due to cells growing upon the spots, and the slide was washed once with KO-DMEM and the media appropriate for the desired experiment was added onto the plate in a 1 mL volume.

Growth Factor Stimulation

Cells were seeded onto microarray wells at a density of 4×10^4 cells cm^{-2} and allowed to adhere for 24 h in ESC maintenance media. After this time, cells were washed with PBS once to remove unbound cells, and media was switched to serum free ESC media for the next day. Following serum starvation, cells were washed with DPBS and treated with serum free media containing various amounts of FGF2 with or without 5 µg mL^{-1} heparin. Immediately following starvation, cells were returned to the incubator for 15 min. After this incubation period, cells were placed directly onto ice for subsequent immunocytochemistry.

Immunocytochemistry

After stimulation, cells were immediately washed with cold DPBS and fixed for 10 min at room temperature in 4% paraformaldehyde. Then cells were then washed 3× with

cold PBS and cellular membranes were permeabilized using cold methanol for 20 min. Cells were then washed 3× with PBS and blocked for 1 h at room temperature with immunocytochemistry (ICC) blocking buffer (3 % (w/v) BSA, 2 % goat serum). The appropriate primary antibody was applied overnight in ICC blocking buffer at 4 °C. Cells were washed 3× with PBS and corresponding secondary antibodies were applied for 1 h at room temperature. Cells were washed 3× with PBS and nuclei were stained with hoescht for 15 min at room temperature. Cells were then washed 3× with PBS and mounted overnight at room temperature using ProLong Gold antifade (Cell Signaling Technology). The next day, cells were subjected to fluorescent microscopy imaging or scanner analysis using an Axon GenePix 4000B microarray scanner (molecular devices), equipped with a Cy3 and Cy5 filter.

Statistical Analysis

All mathematical analyses were performed using GraphPad Prism 9.0. The statistical significance between three or more groups was performed using the built-in analysis (one-way ANOVA), and Tukey's multiple comparison test was used to compare means of two groups. Stars and comparisons relate Tukey's values from ANOVA results, $\alpha = 0.05$. An unpaired two-tailed t-test was used to compare differences between two groups. Bar graph values represent mean ± SD. Thresholds for significance for all tests are set as $*p < .05$; $**p < .01$; $***p < .001$; $****p < .0001$. For detailed statistics for each experiment, with relevant p values, see Table 2.1.

2.5 Supplementary information

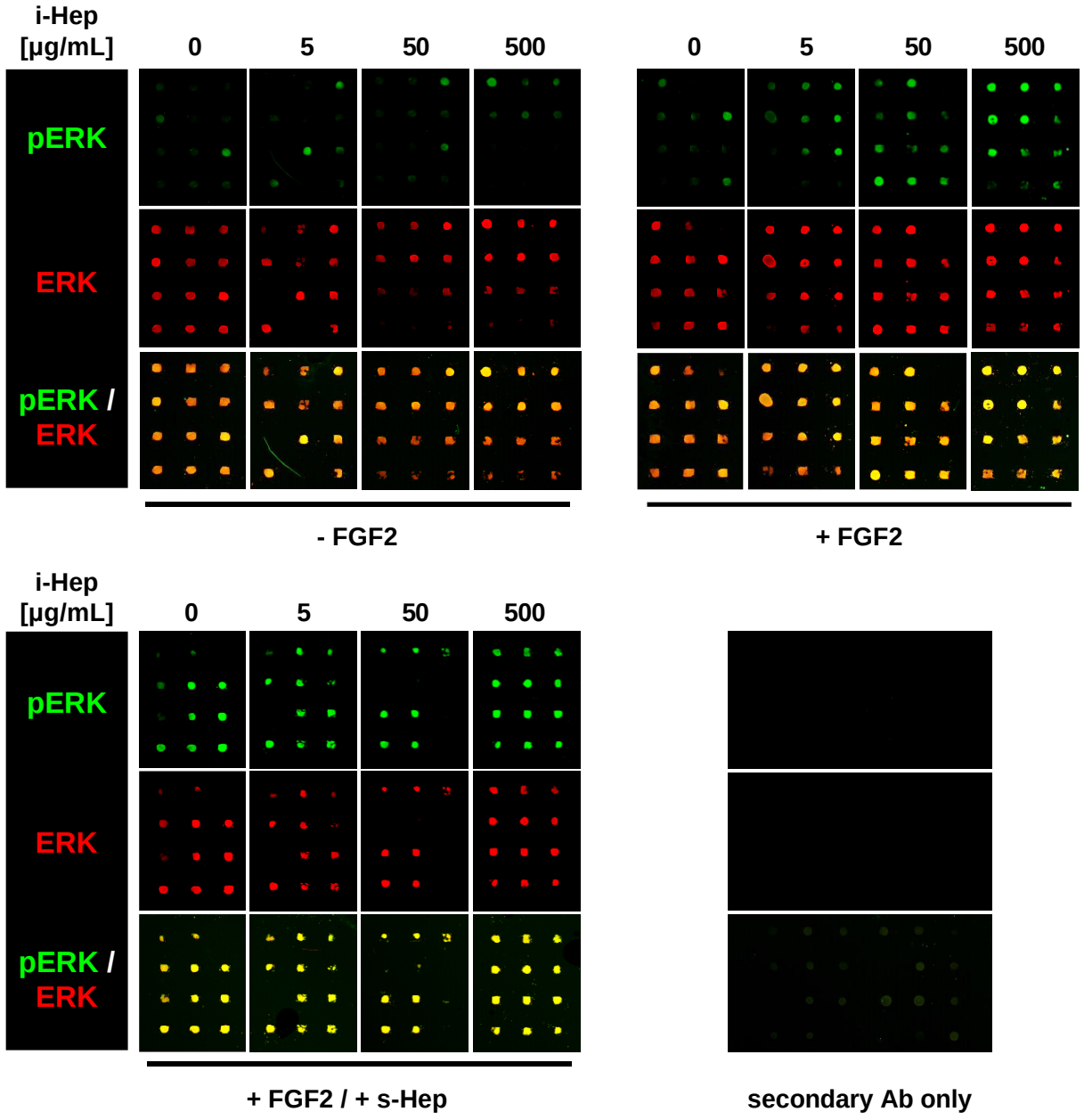


Figure 2.6 Images of microarrays containing Ext1-/- ESC colonies. Colonies were stimulated with FGF2 and immunoassayed for pERK and ERK (Extended data for Figure 2.5D).

Table 2.1 Statistical t-test, ANOVA analysis, and Tukey's multiple comparison tests. A complied statistical analysis from figures 2.2A, 2.2B, 2.4B, and 2.4D.

Unpaired t-test			One or two-tailed?	t value	df	Summary	P value
Figure 2.2A							
+ NHS vs. -NHS			two	5.268	8	***	0.0008
Tukey's multiple comparisons test			Mean Diff.	95.00% CI of diff.	Below threshold ?	Summary	Adjusted P Value
Figure 2.2B							
0 hours vs. 18 hours			915.7	739.3 to 1092	Yes	****	<0.0001
18 hours vs. 48 hours			-113.4	-289.8 to 63.03	No	ns	0.3230
Figure 2.4B and 2.4D. Tukey's multiple comparisons test							
(+/-) FGF2	µg/mL i-hep	(+/-) s-hep	Mean Diff.	95.00% CI of diff.	Below threshold ?	Summary	Adjusted P Value
-	0	-	0.04467	-0.6878 to 0.7772	No	ns	>0.9999
-	500	-					
-	0	-	0.03647	-0.6960 to 0.7690	No	ns	>0.9999
+	0	-					
-	0	-	-0.9681	-1.701 to -0.2356	Yes	**	0.0027
+	500	-					
-	0	-	-2.871	-3.620 to -2.122	Yes	****	<0.0001
+	0	+					
+	500	-	-1.903	-2.652 to -1.154	Yes	****	<0.0001
+	0	+					

2.6 Author Contributions

A.L.M.: conception and design, collection and assembly of data, data analysis and interpretation, manuscript writing. GWT: conception and design, collection and assembly of data, data analysis and interpretation. K.A.T.: Collection and assembly of data, data analysis and interpretation. K.G.: conception and design, manuscript writing, and final approval of manuscript. A.L.M. and G.W.T. contributed equally to this work.

2.7 Acknowledgements

Chapter 2 is an adaptation and reprint of the material as it appears in *Advanced Healthcare Materials* 2021. Michalak AL, Trieger GW, Trieger KA, Godula K., Stem Cell Microarrays for Assessing Growth Factor Signaling in Engineered Glycan Microenvironments. Austen Larson Michalak was the primary author on this publication, and shared equal contribution of work as Greg W Trieger.

2.8 References

1 Nishihara S, Glycans in stem cell regulation: From *Drosophilia* tissue stem cells to mammalian pluripotent stem cells. *FEBS letters* **2018**, 592, 3773-3790

2 Ooi HW, Hafeez S, van Blitterswijk CA, Molsroni L, Baker MS. Hydrogels that listen to cells: A review of cell-responsive strategies in biomaterials design for tissue regeneration. *Materials Horizons* **2017**, 4, 1020-1040

3 McKee C, Chaudhry GR. Advances and challenges in stem cell culture. *Colloids and Surfaces B: Biointerfaces* **2017**, 159, 62, 62-77

4 Choe G, Park J, Park H, Lee JY. Hydrogel Biomaterials for Stem Cell Microencapsulation. *Polymers* **2018**, 10, 9, 997

5 Hudalla GA, Eng TS, Murphy WL. An approach to modulate degradation and mesenchymal stem cell behavior in poly(ethylene glycol) networks. *Biomacromolecules* **2008**, 9, 3, 842-849

6 Caballero Aguilar LM, S.M. Silva, E.M. Simon. Growth factor delivery: Defining the next generation platforms for tissue engineering. *Journal of controlled Release* **2019,** 306, 40-58

7 D. Seliktar, Zisch AH, Lutolf MP, Wrana JL, Hubbell JA. MMp-2 sensitive, VEGF-bearing bioactive hydrogels for promotion of vascular healing. *J. of biomed mat. Research* **2004,** 68, 704-716

8 Seo J, Shin J, Leijten J, Jeon O, Camci-Unal G, Dikina AD, Brinegar K, Ghaemmaghami AM, Alsberg E, Khademhosseini A. High-Throughput Approaches for Screening and Analysis of Cell Behaviors. *Biomaterials* **2018,** 153, 85-101

9 Lutolf MP, Hubbel JA. Synthetic biomaterials as instructive extracellular microenvironments for morphogenesis in tissue engineering. *Nature Biotechnology* **2005**, 23, 47-55

10 Moraes C, Wang GH, Sun Y, Simmons CA. A microfabricated platform for high-throughput unconfined compression of micropatterned biomaterial arrays. *Biomaterials* **2010**, 31, 3, 577-584

11 Restuccia A, Seroski DT, Kelley KL, O'Bryan CS, Kurian JJ, Knox KR, Farhadi SA, Angelini TE, Hudalla GA. Hierarchical self-assembly and emergent function of densely glycosylated peptide nanofibers. *Nat. Comm. Chem.* **2019**, 2, 53

12 Freudenberg U, Liang Y, Kiick KL, Werner C. Glycosaminoglycan-Based Biohybrid Hydrogels: A Sweet and Smart Choice for Multifunctional Biomaterials. *Adv. Mater.* **2016**, 28, 40, 8861-8891

13 Sodhi H, Panitch A. Glycosaminoglycans in Tissue Engineering: A Review. *Biomolecules* **2021**, 11, 9, 29

14 Smith R, Meade K, Pickford C, Holley R, Merry C. Glycosaminoglycans as regulators of stem cell differentiation. *Biochem. Soc. Trans.* **2011**, 39, 383-387

15 Yin J, Seeberger PH. Applications of Heparin and Heparan Sulfate Microarrays. *Meth in Enzymology* **2010**, 9, 478,197-218

16 Pardo-Vargas A, Delbianco M, Seeberger PH. Automated glycan assembly as an enabling technology. *Cur. Opinion in Chem. Biol.* **2018**, 46, 48-55

17 Liu L, Prudden AR, Capicciotti CJ, Bosman GP, Yang JY, Chapla DG, Moremen KW, Janboons G. Streamlining the chemoenzymatic synthesis of complex *N*-glycans by a stop and go strategy. *Nature chemistry* **2019**, 11, 161-169

18 Zhang X, Lin L, Huang H, Linhardt RJ. Chemoenzymatic Synthesis of Glycosaminoglycans. *Acc. Chem. Res.* **2019**, 53, 2, 335-346

19 Chen YH, Narimatsu Y, Calusen T, Gomes C, Karlsson R, Steetoft C, Spliid CB, Gustavsson T, Salanti A, Persson A, Malmström A, Willén D, Ellervik U, Bennett EP, Mao Y, Clausen H, Yang Z. The GAGOme: a cell-based library of displayed glycosaminoglycans. *Nat. methods* **2018**, 15, 881-888

20 Zong C, Venot A, Li X, Li W, Xiao W, Wilkes JL, Salanga CL, Handel TM, Wang L, Wolfert MA, Boons G. Heparan Sulfate Micorarray Reveals That Heparan Sulfate-Protein Binding Exhibits Different Ligand Requirements. *J. Am. Chem. Soc.* **2017**, 139, 28, 9534-9543

21 Sterner E, Meli L, Kwon SJ, Dordick JS, Linhardt RJ. FGF-FGFR Signaling Mediated through Glycosaminoglycans in Microtiter Plate and Cell-Based Microarray Platforms. *Biochemistry* **2013**, 52, 50, 9009-9019

22 T.M. Puvirajjesinghe, Y.A. Ahmed, A.K. Powell, D.G. Fernig, S.E. Guimond, J.E. Turnbull. Array-Based Functional Screening of Heparin Glycans. *Cell Press Chem & Biol.* **2012**, 19, 553-558

23 D.A. Brafman, S. Chien, K. Willert. Arrayed cellular microenvironments for identifying culture and differentiation conditions for stem, primary and rare cell populations. *Nat. Protocols* **2012**, 7, 703-717

24 Ghaemi SR, Harding F, Delalat B, Vasani R, Voelcker NH. Surface Engineering for Long-Term Culturing of Mesenchymal Stem Cell Microarrays. *Biomacromolecules* **2013**, 14, 81 2675-2683

25 Guimond S, Maccarana M, Olwin BB, Lindahl U, Rapraeger AC. Activating and inhibitory heparin sequences for FGF-2 (basic FGF). Distinct requirements for FGF-1, FGF-2, and FGF-4. *J Biol. Chem.* **1993** 268, 32, 23906-14

26 Lundin L, Larsson H, Kreuger J, Kanda S, Lindahl U, Salmivirta M, Claesson-Welsh L. Selectively desulfated heparin inhibits fibroblast growth factor-induced mitogenicity and angiogenesis. *J. Biol. Chem.* **2000**, 275, 32, 24653-60

27 Qiu H, Shi S, Yue J, Xin M, Narin AV, Lin L, Liu X, Li F, Archer-Hartmann SA, Dela Rosa M, Galizzi M, Wang S, Zhang F, Azadi P, van Kuppevelt TH, Cardoso WV, Kimata K, Ai X,

Moremen KW, J.D. Esko JD, Linhardt RJ, Wang L. A mutant-cell library for systematic analysis heparan sulfate structure-function relationships. *Nat. Methods* **2018**, 15, 889-899

28 Presto J, Thuveson M, Carlsson P, Busse M, Wilén M, Eriksson I, Kusche-Gullberg M, Kjellén L. Heparan sulfate biosynthesis enzymes EXT1 and EXT2 affect NDST1 expression and heparan sulfate sulfation. *PNAS* **2008**, 105, 12, 4751-4756

29 Kraushaar DC, Yamaguchi Y, Wang L. Heparan Sulfate is Required for Embryonic Stem Cells to Exit from Self-renewal. *J Biol Chem* **2010**, 285, 8, 5907-5916

30 Zhang F, Zhang Z, Lin X, Beenken A, Eliseenkova AV, Mohammadi M, Linhardt RJ. Compositional analysis on heparin/heparan sulfate interacting with FGF•FGFR complexes. *Biochemistry* **2009**, 48, 35, 8879-8386

31 Roux P, Blenis J. ERK and p38 MAPK-Activated Protein Kinases: A Family of Protein Kinases with Diverse Biological Functions. *Microbiol. Reviews* **2004**, 68, 2, 320-344

32 Mossahebi-Mohammadi M, Quan M, Jin-San Z, Li X. FGF Signaling Pathway: A Key Regulator of Stem Cell Pluripotency. *Front Cell Dev Biol.* **2020**, 8, 79

3. Small Molecule Antagonist of Cell Surface Glycosaminoglycans Restricts Mouse Embryonic Stem Cells in a Pluripotent State

3.1 Introduction

The defining traits of pluripotent stem cells (PSCs), which include embryonic stem cells (ESCs) isolated from the inner cell mass of blastocysts and induced pluripotent stem cells (iPSCs) derived from somatic cells through cellular reprograming, are their high self-renewal capacity and their ability to produce cell types of all three germ layers.[1,2] Chemical approaches for the generation and maintenance of PSCs are attractive[3,4] as they offer important advantages over methods relying on gene manipulation[5], or the use of cytokines[6] and growth factors[7] to confine cells in their pluripotent state. The use of small molecules alleviates safety concerns regarding permanent modifications to the genome of the target cells, while offering increased chemical stability, low cost of production, and better pharmacokinetic profiles for in vivo applications compared to biologics. Whereas the procedures for assessing the toxicity of small molecules are well established as part of the FDA approval process, genetic targeting as a therapeutic strategy is not. Chemical modulation can offer the ability to fine-tune the resulting effects based on reversibility or dosage. Genetic instability has been reported for cells following genetic targeting, especially in iPSCs.[8] Compounds have now been discovered that promote cellular reprograming toward pluripotent state or regulate germ layer specification.[9] Whereas the field of small molecule-based modulators of cellular differentiation has focused primarily on targeting the activity of receptor and intracellular

kinases or epigenetic enzymes controlling gene expression[10], glycan interactions at the cell-matrix interface regulating the activation of receptors upstream of these pathways have received much less attention. Cell surface glycans are essential for proper embryonic development; and stem cell differentiation is often accompanied by alterations in glycosylation patterns.[11] (Fig. 1) In fact, unique glycan structures, such as the stage specific embryonic antigens (SSEAs) 1 and 3 can be used as indicators of pluripotency in murine and human ESCs, respectively.[12] At the same time, elimination of specific cell surface glycan structures caused by mutations in glycan biosynthesis genes results in embryonic lethality or severe congenital disorders.[13] Proteoglycans (PGs) – membrane-associated proteins modified with long chains of sulfated polysaccharides, called glycosaminoglycans (GAGs) – are representatives of one such family of essential glycoconjugate regulators of cellular differentiation. Through their GAG chains, PGs help mediate the association of growth factors with cell-surface receptors into ternary complexes that initiate intracellular signaling cascades and gene expression (Fig. 1).[14] Murine ESC (mESC) mutants missing the exostosin 1 (Ext1) gene that encodes a glycosyltransferase involved in the elongation of heparan sulfate (HS) GAG chains, lack the ability to signal through growth factors of the FGF, TGF or Wnt families and, as a result, are restricted in their capacity to exit from embryonic state.[15,16,17] As master modulators of key intracellular signaling pathways, GAGs constitute unique targets for interfering with extracellular differentiation cues and influencing stem cell state.[18] It is well-known that mouse and human ESCs employ different signaling pathways to maintain pluripotency.[19] For instance, while human ESCs require active FGF2 signaling to maintain pluripotency, murine ESCs can employ FGF2 signaling for differentiation.[19] In

this work, we focus our efforts on mESCs, where the role of HS GAG chains is well-studied providing a suitable model system to study how chemical modulators can affect pluripotency via glycan-mediated mechanisms.

Several methods have been developed for transient attenuation of GAG activity without imparting permanent genomic changes to the cell. Most commonly, the target GAG structures are enzymatically removed via depolymerization of the polysaccharide chains with bacterial lyases.[20] Whereas the substrate specificity of the employed enzyme determines which GAG classes will be eliminated (e.g., heparinases and chondroitinases respectively degrade heparan sulfate or chondroitin sulfate chains), this process requires extended treatment of the cells with high concentrations of enzymes and the complete removal of inaccessible GAG structures can be challenging. As a consequence, heparinase treatment can inhibit endothelial differentiation in mESCs but fails to restrict them in a pluripotent state.[21] Alternatively, small molecule inhibitors can be used to interfere with the biosynthesis of GAGs, although they are yet to be tested as inhibitors of ESC differentiation.[22,23,24,25]

For proper function, GAGs need to be modified with negatively charged sulfate groups that define growth factor and receptor binding sites. The inhibition of the general sulfate donor, PAPS (3'-phosphoadenosine-5'-phosphosulfate), with sodium chlorate leads to under sulfation of GAGs at the cell surface and loss of activity.[26,27] While commonly employed, chlorate treatment indiscriminately affects the sulfation of other

glycan classes well as proteins.[28] Such global and poorly characterized perturbations of sulfation may explain the contradictory reports on the effects of chlorate treatment on mESC differentiation as.[29,30]

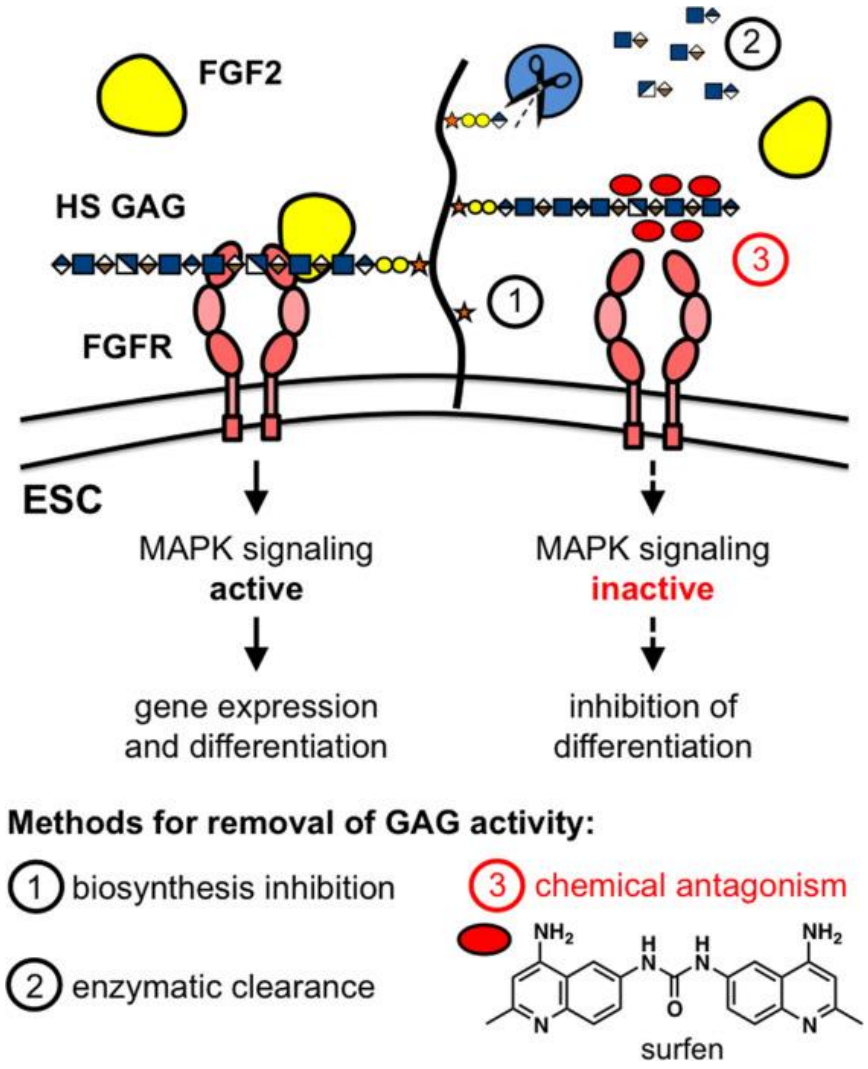


Figure 3.1 Targeting glycosaminoglycans to influence embryonic stem cell fate. Top: HS GAGs are required for FGF2-dependent induction of the MAPK pathway and neural differentiation in murine embryonic stem cells. Deactivation of HS GAGs disrupts MAPK signaling and inhibits differentiation. Bottom: Small molecule antagonists provide an attractive alternative to genetic and enzymatic methods for the attenuation of HS GAG activity.

Direct antagonism of GAGs provides perhaps the most attractive opportunity for modulating signaling activity at the cell surface. Although small molecules can interact

with GAGs, few have been tested for effects in stem cell differentiation.[31] For instance, surfen, an aminoquinoline with heparin-neutralizing properties, was first reported in 1938.[32] Since its discovery, surfen has been shown to inhibit HIV infection[33], vasculogenesis[34], or modulate T-cell activation[35], through the antagonism of HS GAG interactions with signaling and receptor proteins.[36,37] In this study, we report that surfen effectively and reversibly restricts mESCs in their pluripotent state by attenuating the activity of their surface GAG structures in growth factor association and signaling.

3.2 Results

Surfen is a potent, reversible inhibitor of neural differentiation and a promoter of pluripotency in murine ESCs.

Cognizant of the profound effects of the genetic deletion of the Ext1 gene (vide supra), and the ensuing loss of surface HS expression, on the inability of mESCs to exit from the pluripotent state[16], we set out to investigate whether common GAG antagonists may chemically induce similar phenotype in wild type cells. We designed a dual endpoint flow cytometry assay that employs two green fluorescent protein (GFP) knock-in mESC lines to simultaneously evaluate both pluripotency and differentiation in the presence of GAG antagonists.

We used an Oct4-GFP mESC line to monitor the expression of Oct4, which indicates stem cell pluripotency and stemness[38], and a Sox1-GFP mESC mutant (46C)[39] to assess neural commitment (Figure 3.2A). Oct4-GFP mESCs express high levels of

GFP when cultured under embryonic conditions in the presence of the leukemia inhibitory factor (LIF) and gradually lose GFP expression upon differentiation. Conversely, Sox1-GFP mESCs do not express GFP in embryonic culture but become GFP-positive upon acquisition of neural identity, which coincides with the expression of the early neuroectodermal marker, Sox1.[39] Both cell lines exhibit a compact colony morphology when pluripotent, and lose this feature upon differentiation. Using a well-established protocol[39], we induced both reporter mESC lines in serum-free N2B27 media toward neuroectodermal differentiation and assessed the levels of Oct4 and Sox1 expression (Figure 3.2B). At day six of differentiation, we observed ~20-50% of Oct4-GFP mESCs remain GFP-positive, while ~40-70% of Sox1-GFP mESCs have acquired GFP expression and neural phenotype in these surfen-untreated controls. Immunostaining of both Oct4-GFP and Sox1-GFP mESC lines, as well as wild-type E14Tg2a mESCs, also confirmed the loss of Oct4 and gain of Sox1 protein expression after six days of differentiation (Figure 3.7).

Whereas the current repertoire of GAG antagonists is rather small, we acquired and tested three commercially available molecules known to engage HS (surfen, adhesamine, and protamine) in our differentiation assay. Surfen and adhesamine have been reported to modulate FGF signaling, as well as cell adhesion and proliferation, respectively, through interaction with HS.[36, 40] Protamine is a high molecular weight cationic lysine and arginine-rich protein employed as a neutralizing agent for the anticoagulant heparin (also a GAG). [41] Initial evaluation of surfen (5.0 μM) via fluorescence microscopy indicated that it inhibited Sox1-GFP expression, while

maintaining the colony morphology of mESCs and high Oct4 expression (Figure 3.2C). To obtain a more quantitative analysis of differentiation in the presence of all three HS-binding molecules, we performed flow cytometry to assess cell populations on day six of differentiation (Figure 3.8). For comparison, we also included PD173074, a small molecule FGFR antagonist shown previously to restrict mESCs in a pluripotent state, as a positive control (Figure 3.2D).[42] Whereas protamine (10 μM) maintained high levels of Oct4-GFP, it did not inhibit Sox1-GFP expression in our assay. Adhesamine (10 μM) showed no effect on either Sox1-GFP or Oct4-GFP expression compared to the untreated control. In contrast, surfen (5.0 μM) effectively inhibited neural differentiation (<10% Sox1-GFP positive cells), while maintaining the mESCs in a pluripotent state (>90% Oct4-GFP positive population) after six days in differentiation (Figure 3.2D). Further increases in the population of pluripotent cells (~99%; not shown) can be achieved at higher concentration of the antagonist; however, changes to the cellular morphology become apparent at concentrations above 10 μM, presumably, due to aggregation of negatively charged culture medium components by surfen (Figure 3.9). Nonetheless, 5.0 μM surfen consistently inhibited differentiation in both Oct4-GFP and Sox1-GFP mESCs (Figure 3.10) at levels similar to the known FGFR inhibitor, PD173074. For comparison, we also evaluated the effects of heparinase and chlorate treatment on mESC differentiation, since both are commonly used to attenuate HS GAG activity at the cell surface (vide supra). In agreement with prior work, chlorate treatment led to an initial acceleration of neural differentiation of mESCs toward Sox1-GFP+ populations[29], with eventual return to high levels of Oct4-GFP expression after day 2 (Figure 3.11). Although removal of surface HS with bacterial heparinase provided some inhibition of mESC differentiation, it required

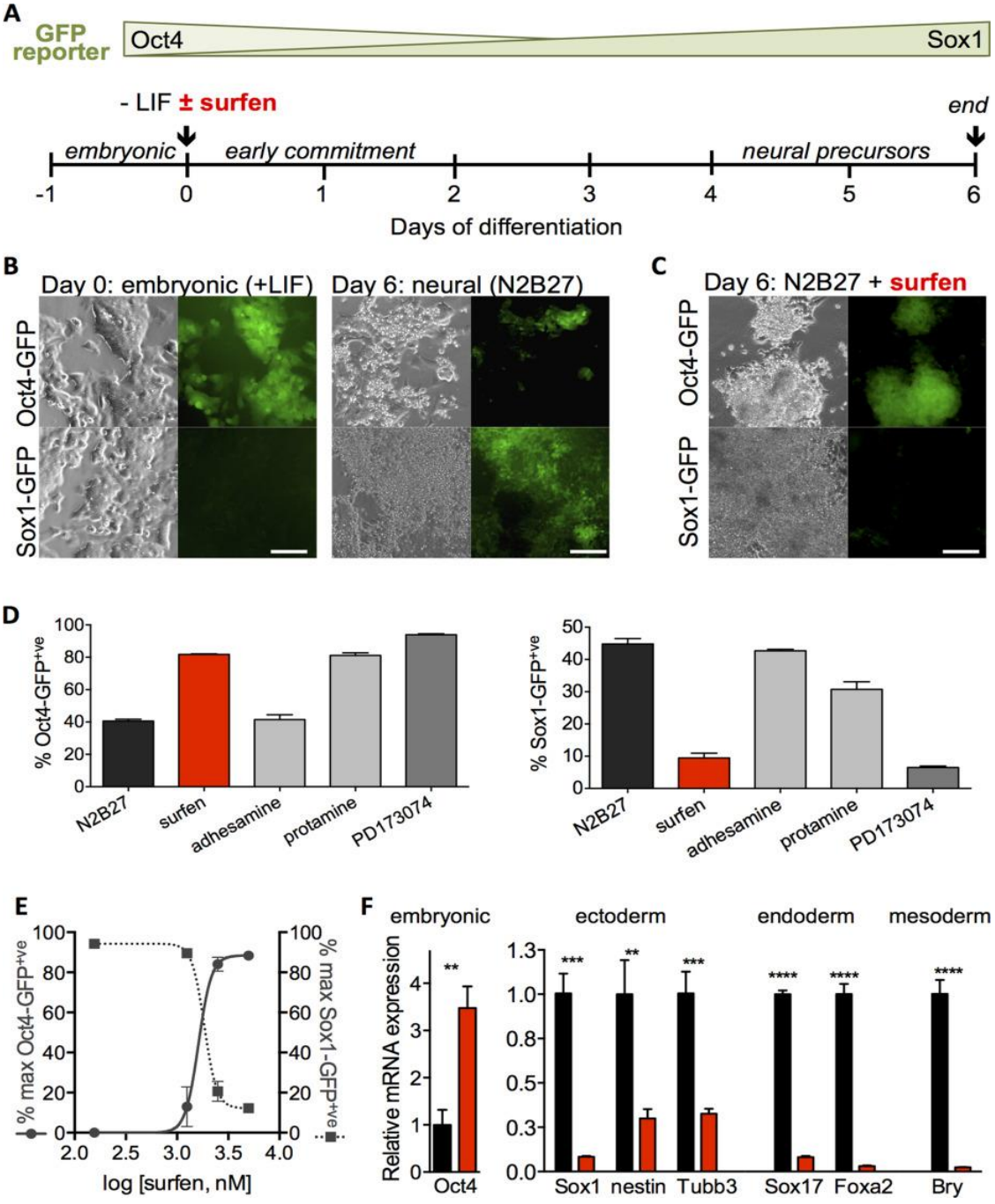
large amounts of costly enzyme (~ 500-fold difference in cost per experiment) and suffered from a high degree of variability and low efficiency compared to surfen treatment (Figure 3.12).

Surfen inhibited Sox1 and promoted the maintenance of Oct4 expression under the neural differentiation conditions in a dose-dependent manner. Inhibition curves established using the GFP fluorescence of the reporter cell lines yielded similar IC50 values of ~2.0 μM, (Figures 3.2E and 3.13) indicating that surfen equally inhibits neural differentiation and promotes pluripotency, or, in other words, that the reduction in neural induction is not caused by mESC differentiation into non-neural lineages, but rather by placing limits on their ability to exit from a pluripotent state.

Our observation that surfen maintains mESCs in a pluripotent state, is further supported by our flow cytometry data indicating high levels of expression of SSEA-1 (Figure 3.14), and a low abundance of the neural marker, Sox1, on day 6 of differentiation in N2B27 in the presence of the inhibitor (Figure 3.7). Quantitative RT-qPCR analysis (see Supporting Information Table 1 for primer sequences) showing high expression of Oct4 and low levels of neuroectodermal (Sox1, Tubb3, nestin), mesodermal (Bry), and endodermal (Foxa2, Sox17) markers provides further evidence for inhibition of differentiation (Figures 3.2F and 3.15). We further characterized the pluripotent nature of the surfen-treated cells, and determined by qRT-PCR analysis that these cells are indeed embryonic and not in the epiblast state, as evidenced by high expression of the embryonic

marker Rex1, but low levels of the epiblast markers FGF5 and Nodal (Figure 3.16).[43] It is important to point out that surfen had no significant effect on cell proliferation rates, excluding the possibility that inhibition of differentiation is caused by a decrease in cell viability (Figure 3.17).

Figure 3.2 Dual endpoint GFP-reporter assay for evaluating heparan sulfate glycosaminoglycan (HS GAG) antagonists as inhibitors of neural differentiation in murine embryonic stem cells (mESCs). (A): Oct4-GFP and Sox1-GFP mESC reporter lines were used to evaluate the ability of HS GAG antagonists to inhibit neural differentiation and promote pluripotency over 6 days in N2B27 culture. (B): Live cell fluorescence micrographs show loss of pluripotency (Oct4) and acquisition of neural phenotype (Sox1) after 6 days of differentiation. Scale bar: 100 μm. (C): In the presence of surfen (5 μM), Oct4 and Sox1 expression profiles indicate that mESCs continue to maintain high levels pluripotency even after 6 days in differentiation. (D): Surfen shows enhanced ability to promote pluripotency and inhibit differentiation compared to other known HS modulators, protamine (10 μM), and adhesamine (5 μM). Surfen activity is comparable to that of PD173074 (1 μM), an ATP-competitive FGF receptor inhibitor. (E): Surfen inhibits differentiation in both Oct4-GFP and Sox1-GFP cell lines with IC50 ~2.0 μM. Each point represents technical duplicates (mean ± SD), representative of two biological replicates. (F): Analysis of mRNA expression in Oct4-GFP mESCs on day 6 of treatment with surfen in N2B27 (5.0 μM) indicates cell arrest in the embryonic state. Relative mRNA expression is calculated from technical triplicates (mean ± SD) normalized to untreated (N2B27) controls (defined as 1.0). This experiment is representative of two biological replicates. **, $p < .0037$; ***, $p < .002$; ****, $p < .0001$. Abbreviations: GFP, green fluorescent protein; LIF, leukemia inhibitory factor.


A
GFP reporter
Oct4
Sox1
- LIF ± surfen
end
embryonic
early commitment
neural precursors
-1
0
1
2
3
4
5
6
Days of differentiation
B
Day 0: embryonic (+LIF)
Day 6: neural (N2B27)
Oct4-GFP
Sox1-GFP
C
Day 6: N2B27 + surfen
Oct4-GFP
Sox1-GFP
D
% Oct4-GFP+ve
N2B27
surfen
adhesamine
protamine
PD173074
% Sox1-GFP+ve
N2B27
surfen
adhesamine
protamine
PD173074
E
% max Oct4-GFP+ve
% max Sox1-GFP+ve
log [surfen, nM]
F
embryonic
ectoderm
endoderm
mesoderm
Relative mRNA expression
Oct4
Sox1
nestin
Tubb3
Sox17
Foxa2
Bry

A key advantage of using small molecule modulators is the possibility to achieve transient and reversible change in cellular activity. We tested whether inhibition of differentiation by surfen can be reversed after its withdrawal. Sox1-GFP mESCs maintained in a pluripotent state during a 6-day challenge under neural differentiation conditions were replated in surfen-free N2B27 media (Figure 3.3). After an additional 7 days in differentiation (D13), the cells underwent successful neural differentiation, as evidenced by loss of Oct4-GFP and increased Sox1- GFP fluorescence. (Figure 3.3A) Cells immunostained at D6 of differentiation express high levels of Oct4 and low levels of nestin compared to non-treated cells (Figure 3.3B), and the withdrawal of surfen allows increased nestin expression at D13 (Figure 3.3C), whereas cells continuously treated with surfen for 13 days still display colony morphology, high levels of Oct4 and low levels of nestin (Figures 3.18 and 3.19).

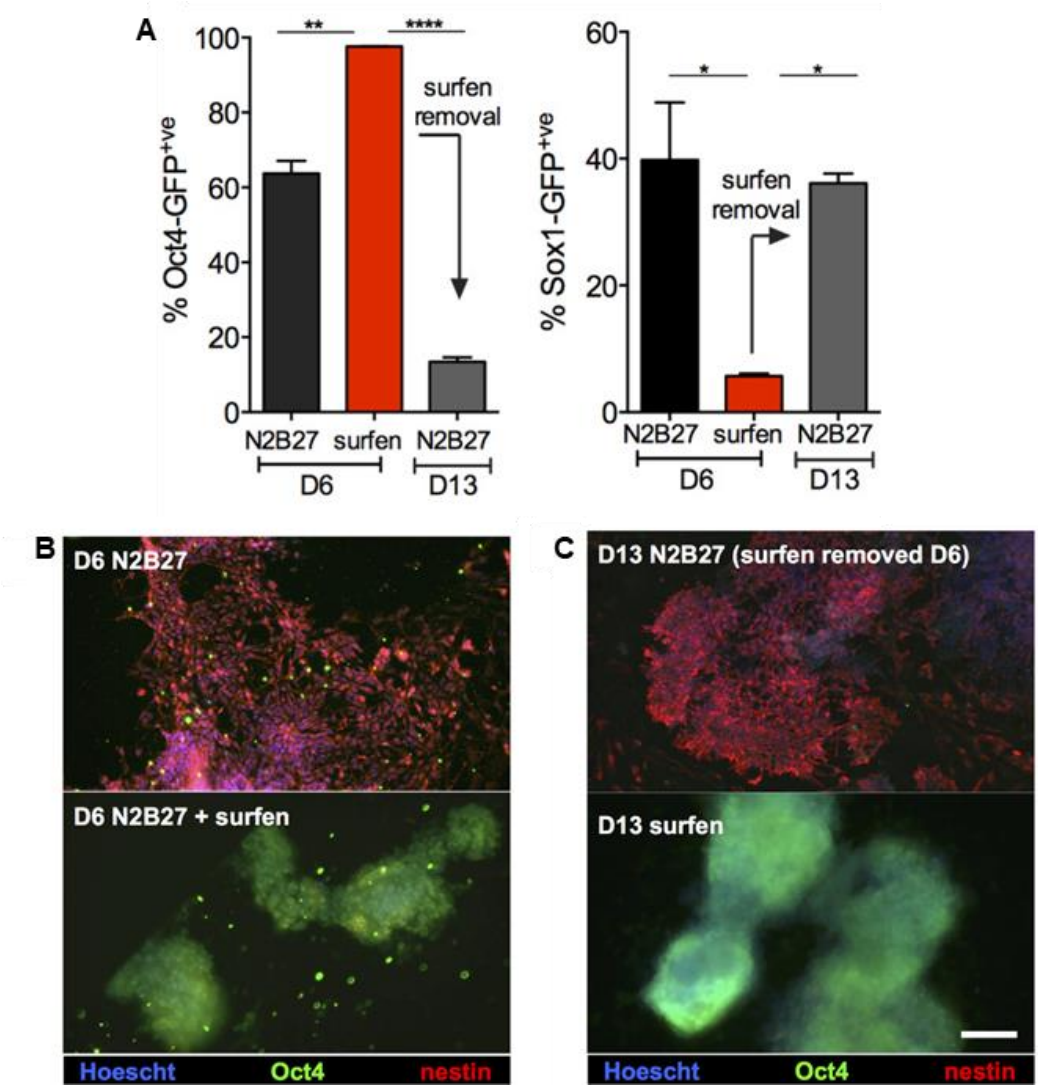


Figure 3.3 Surfen is a reversible inhibitor of differentiation. (A): After 6 days in neural differentiation in the presence of surfen (5 μM) cells remain pluripotent, as evidenced by high Oct4-GFP and low Sox1-GFP expression levels via flow cytometry. Differentiation resumes after removal of surfen at day 6 producing high levels of Sox1-GFP expression with concurrent loss of Oct4-GFP by day 13. %GFP^{+ve} values are provided as technical duplicates (mean ± SD), representative of three biological replicates. *, $p < .0335$; **, $p < .0048$; ***, $p < .0001$. (B): Immunostaining for neural markers, nestin, and β-III-tubulin in Oct4-GFP murine embryonic stem cells (mESCs) after 6 days of neural differentiation in N2B27 with or without surfen. (C): Nestin and β-III-tubulin expression in Oct4-GFP mESCs on day 13, 6 days after surfen removal. Scale bar: 100 μm. Abbreviation: GFP, green fluorescent protein.

Surfen acts as a broad-spectrum inhibitor of signaling.

Neuroectodermal commitment of mESCs to differentiation is known to initiate via autocrine activation of the mitogen- activated protein kinase (MAPK) signaling pathway by fibro- blast growth factors 2 and 4 (FGF2/4).[16] HS GAGs are required for proper FGF function, acting as co-receptors responsible for the recruitment of FGFs to the cell surface and organization of the active growth factor-receptor complex.[27,44] Auto-phosphorylation of the FGFR ensues, triggering downstream signaling events, including the phosphorylation of the Erk1/2 kinases, that ultimately result in gene expression and differentiation.

Western blot analysis of mESCs stimulated by FGF2 show robust levels of phospho-Erk1/2; however, this response is attenuated by surfen in a dose-dependent manner (Figure 3.4A and 3.20). Surfen effectively inhibits binding of FGF2 to heparin in an enzyme-linked immunosorbent assay (ELISA) (Figure 3.21) and MAPK signaling can be restored by the addition of exogenous heparin, which competes for cell surface HS-bound surfen molecules (Figure 3.4B and 3.22). These observations provide evidence that surfen is likely inhibiting MAPK signaling by blocking FGF2 binding sites on cell surface HS. Consistent with Western blot data, flow cytometry analysis also indicates that heparin can restore differentiation following surfen treatment. Oct4-GFP mESCs treated with surfen (5.0 mM) and further titrated with increasing amounts of heparin showed that 5.0 mg/ml is sufficient to abrogate the effects of surfen, as evidenced by the low Oct4-

GFP levels (Figure 3.4C). The exit from pluripotency promoted by heparin in surfen-treated cells was observed to proceed to neural differentiation, as observed in enhanced Sox1-GFP levels (Figure 3.4C).

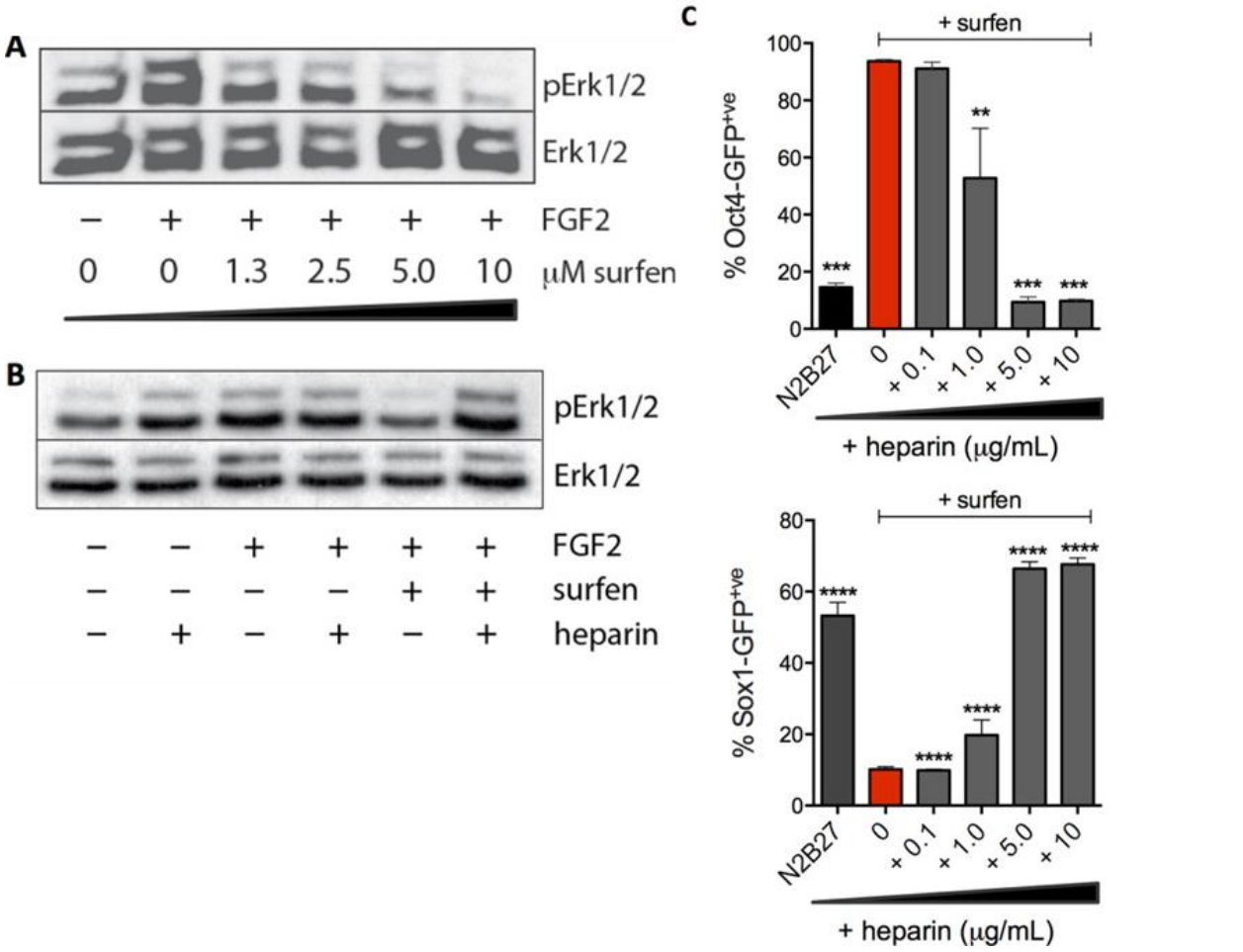


Figure 3.4 Surfen acts by inhibiting FGF2 signaling and its activity can be neutralized with soluble heparin. (A): Stimulation of Oct4-GFP murine embryonic stem cells (mESCs) with FGF2 in the presence of surfen leads to dose-dependent attenuation of Erk1/2 phosphorylation. (B): Erk1/2 phosphorylation in Oct4-GFP mESCs is recovered in the presence of heparin (5 μg/ml), a soluble competitor for cell surface heparan sulfate-bound surfen. (C): Added soluble heparin restores the ability of mESCs to undergo neural differentiation in the presence of surfen (5 μM). Dunnett's multiple comparison test against surfen-treated control, ****, $p < .0001$. Shown are technical duplicates (mean ± SD), repeated with two biological replicates. Abbreviation: GFP, green fluorescent protein.

RTK Array

In addition to the MAPK pathway, HS is suspected to mediate a number of other signaling events associated with mESC differentiation.[44] Therefore, we also evaluated the effects of surfen toward other RTKs using an RTK array. (Figure 3.5A) Oct4-GFP mESCs were stimulated with FGF2 (25 ng/ml) in the presence or absence of surfen (5.0 mM), and the resulting lysates were incubated identically with the capture array, which includes 39 specific RTK antibodies, allowing for the simultaneous detection of receptor tyrosine phosphorylation. (Figures 3.5 and 3.22) Surfen inhibited phosphorylation of numerous RTKs, including platelet-derived growth factor (PDGFRa) [45], ErbB2 [46], epidermal growth factor receptor (EGFR)[47], and macrophage-stimulating protein receptor (MSPR). (Figure 3.5B) Interestingly, this broad inhibitory effect of surfen toward tyrosine kinase phosphorylation is similar to that caused by the Ext1 gene deletion in Sox1-GFP cells, which leads to the truncation of HS chains.[48]

Given the embryonic state of the cells, only low levels of pTrkC, and the vascular endothelial growth factor (VEGF) family receptors pVEGFR1, and pVEGFR2, were detected in the samples with or without FGF2 stimulation, compared to other RTKs. Despite the low abundance of these phospho-RTKs, surfen also caused a reduction in their phosphorylation. VEGF is known to interact with HS[49] and surfen has previously been reported to inhibit VEGF-mediated endothelial sprouting.[37] Quantitative phosphoproteomics studies of *Ext1*$^{-/-}$ endothelial cells similarly revealed that antagonism of HS inhibits phosphorylation of these RTKs.[50]

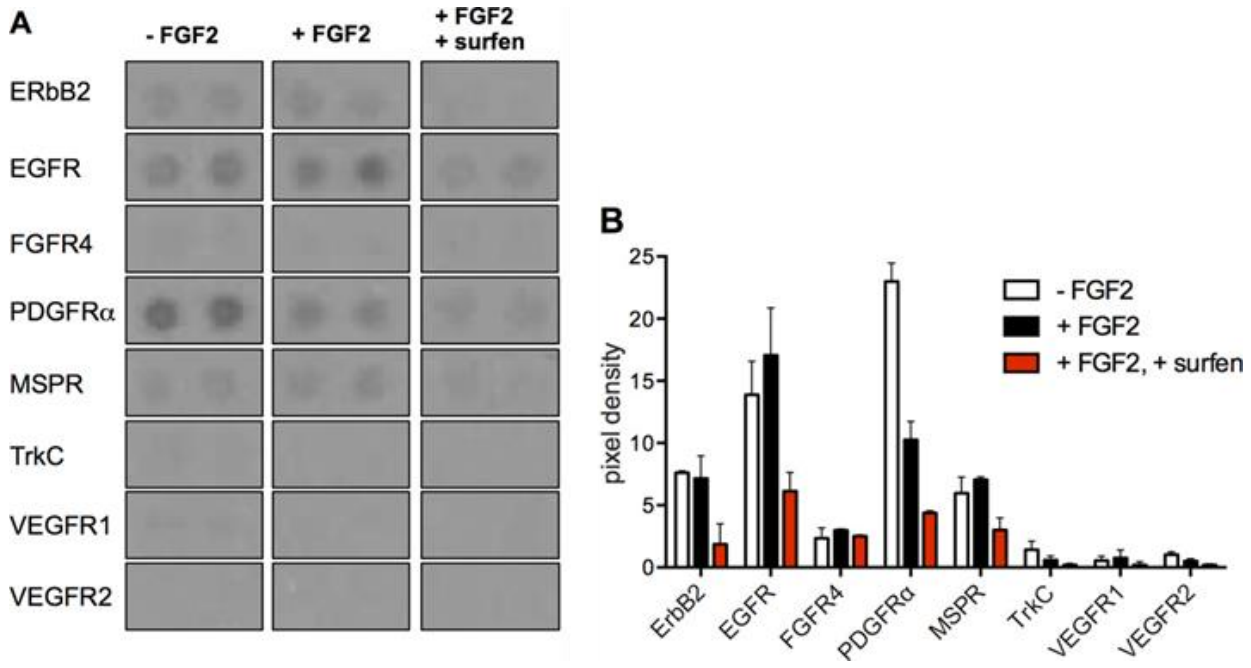


Figure 3.5 Receptor tyrosine kinase (RTK) array analysis of embryonic Oct4-GFP murine embryonic stem cells in response to surfen treatment. Processed blots (A) and bar graphs (B) demonstrate that surfen inhibits phosphorylation of numerous RTKs. Bar graphs were generated using technical duplicates (means ± SD), and the experiment was performed once. Abbreviations: EGFR, epidermal growth factor receptor; GFP, green fluorescent protein; MSPR, macrophage-stimulating protein receptor.

Surfen Is a General Inhibitor of Differentiation

Having established the efficacy of surfen in inhibiting neural differentiation of mESCs in monolayer culture, we sought to evaluate its ability to maintain pluripotency in cells cultured in a three-dimensional format. EBs generated from the reporter cell lines via the hanging drop method were cultured in N2B27 neural induction medium in the presence of surfen for 6 days.[51] Fluorescence micrographs in Figure 6A clearly show suppression of GFP expression in Sox1-GFP mESCs in response to surfen (5 mM) compared to untreated cells, which was mirrored by high levels of fluorescence in the

Oct4-GFP reporter cells. The preliminary visual assessment by microscopy of neural differentiation in EB culture was corroborated by flow cytometry analysis after EB dissociation (Figure 3.6B). In culture, mESCs undergo spontaneous differentiation upon the withdrawal of LIF to generate cell types of all three (i.e., neuroectodermal, mesodermal, and endodermal) germ layer lineages. Accordingly, the reporter mESC lines cultured for 6 days in KSR (15%) differentiation media in the absence of LIF indicate a significant loss of Oct4 and increased levels of Sox1 expression and neural commitment (Figure 3.6C). Inclusion of surfen (5 mM) resulted in ~80% GFP-positive population of Oct4-GFP reporter mESCs, indicating its ability to act as an effective general inhibitor under spontaneous differentiation conditions.

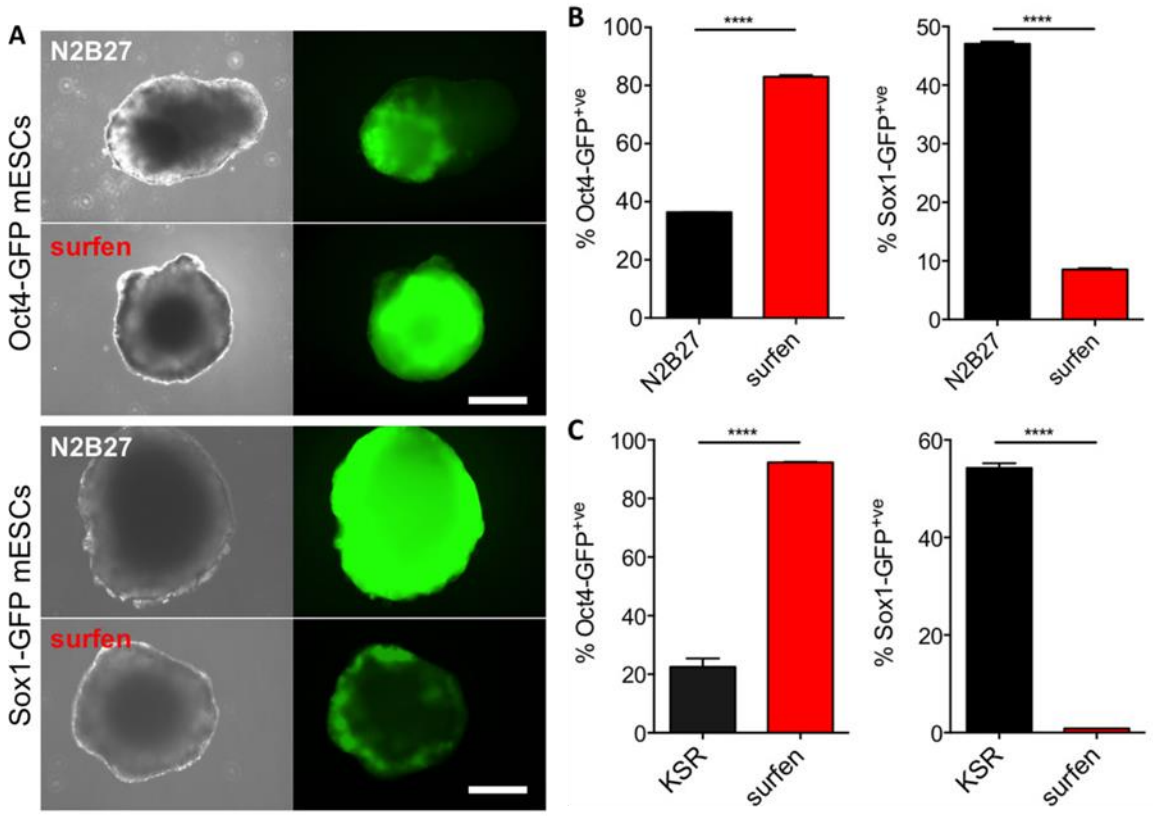


Figure 3.6 Surfen is a general inhibitor of neural and spontaneous differentiation in mESCs. Live cell fluorescence micrograph images (A) and flow cytometry analysis (B) show inhibition of neural differentiation and maintenance of pluripotency in EB culture 6 days after neural induction in N2B27 media. (C): Surfen maintains pluripotency in mESCs under spontaneous differentiation conditions 6 days after withdrawal of the leukemia inhibitory factor. Provided are technical duplicates (means ± SD), as a representative experiment of three biological replicates, ****, $p < .0001$. Scale bar: 100 μm. Abbreviations: GFP, green fluorescent protein; KSR, knockout serum replacement; mESCs, murine embryonic stem cells.

3.3 Discussion

ESCs continue to hold a significant promise for future biomedical applications. Small molecules can be used as reagents to control the pluripotency and differentiation of ESCs. To date, however, small molecules targeting the glycocalyx of stem cells to influence their fate are yet to be identified. Here, we report that surfen, a bis-aminoquinoline molecule, potently maintained pluripotency in mESCs, as shown by flow cytometry, qPCR, and immunostaining. (Figures 3.2 and 3.3)

Surfen acts by binding negatively-charged HS molecules present on the glycocalyx of mESCs via its positively charged aminoquinoline groups.[36] By antagonizing interactions between GAGs and HS-binding cytokines, such as FGF2, surfen can attenuate signaling pathways downstream of FGF2 activation (e.g., Erk phosphorylation) (Figures 3.4 and 3.5). Surfen's effects are dose-dependent and reversible, and the treated mESCs regain the ability to differentiate into neural precursor cells upon the withdrawal of surfen (Figure 3.4). Given surfen's ability to inhibit activation of signaling pathways similar to those observed in *Ext1*$^{-/-}$ mESCs, it presents a chemical means to induce a similar phenotype in mESCs.[48] Interestingly, despite their abilities to bind HS, protamine and adhesamine, failed to maintain pluripotency in mESCs. This observation alludes to the fact that other molecular factors, such as shape complementarity, may be needed for proper antagonism. Indeed, surfen could align itself along the negatively charged sulfate and carboxyl groups of an HS disaccharide via its protonated quinolone rings.[36] This arrangement may be necessary for surfen's observed effects, although more studies are required to properly determine this binding arrangement.

We wish to point out that the extrapolation of the effects of surfen on mESCs, toward human ESCs should be avoided. Human and mouse ESCs use distinct signaling pathways to control pluripotency and differentiation, and express different sets of receptors on their cell surfaces to do so.[19] Nonetheless, the role of GAGs in modulating mESC pluripotency is well documented and these cells provide a good model for establishing small molecules as modulators of stem cell fates through targeting of the cellular glycocalyx. This study thus provides a precedent for further exploration of the glycocalyx in the context human stem cell differentiation.

The cellular glycocalyx plays a defining role in facilitating the transfer of differentiation cues from the extracellular matrix to signaling receptors at the cell surface. As such, it harbors new exciting opportunities for intercepting these signals in order to influence the outcome of differentiation. We have now demonstrated the ability of surfen, a small molecule antagonist of cell-surface GAGs, to arrest ESCs in their pluripotent state by blocking the binding sites for growth factors within their glycocalyx. As a small molecule, surfen inhibits glycan-growth factor interactions in a reversible fashion, providing temporal control over signaling, gene expression, and differentiation. Surfen is a general inhibitor of GAG function; however, we anticipate that ongoing efforts in the molecular design and synthesis of chemical antagonists with enhanced selectivity toward unique subfamilies of GAG structures, or even other prominent classes of cell surface glycans involved in controlling stem cell pluripotency[52], will yield a new, expanded set of selective small molecule modulators of cellular signaling and differentiation.

3.4 Materials and methods

mESC culture

Oct4-GFP and Sox1-GFP (PrimCells) mESCs were maintained in gelatin-coated plastic tissue culture dishes and mESC complete media (0.01% LIF/ 0.1% β-mercaptoethanol/ 1% L-glutamine/ 1% MEMNEAA/10% FBS/ KO-DMEM). Cells were passaged every other day by trypsinization (0.05% trypsin-EDTA) and splitting at 1:10, after washing with DPBS. Unless otherwise indicated, all incubation conditions were conducted at 37°C, 5% CO_2.

mESC differentiation (N2B27, KSR)

To induce differentiation, mESCs were seeded (10,000 cells/cm^2) in 1 mL complete media overnight (D-1). After 24 hours of incubation, cells were washed twice with DPBS, and replenished with N2B27 media with or without compound (D0). N2B27 media was prepared by a 1:1 mixture of Neurobasal media and DMEM/F-12, 1:1000 of B27 and 1:500 N2 supplements, as well as 1% L-glutamine, Penicillin/Streptomycin, and 0.1% β-mercaptoethanol. To induce spontaneous differentiation, cells were seeded similarly, and the differentiation media used was 15% KSR (Knockout Serum Replacement)/1% NEAA/1% L-glutamine/0.1% β-mercaptoethanol/KO-DMEM. Media with or without additive was replenished every 1-2 days.

EB formation and differentiation

mESCs were detached and 20 µL drops of a $3x10^5$ cells/mL suspension in mESC growth media (60,000 cells/EB) was plated using a multi-channel pipet onto the lids of 10 cm low attachment plastic petri dishes (D-2). After 2 days, robust embryoid bodies (EBs) were formed. About 30-40 EBs were used per experimental condition. The EBs were washed twice by re-suspension in DPBS, and differentiation was initiated by re-suspension of the EBs in N2B27 media onto low-attachment 10 cm dishes (D0). The media was replaced every other day by washing once with DPBS and re-suspension in N2B27 with or without compound. After differentiation (D6), EBs were washed 2x, and dissociated into single cells by incubation with Accutase enzyme (RT, 5 min). Live flow cytometry analysis was performed after neutralization with 10% FBS/KO-DMEM, and gentle pipetting.

Flow cytometry and microscopy analysis

For live cell analysis, mESCs or differentiated cells were detached from adherent culture with trypsin-EDTA, diluted in 10% FBS/KO-DMEM and analyzed with a BD Accuri C6 flow cytometer. Cells were gated in FSC vs. SSC plot to exclude cell debris and dead cells, and 15-20,000 events in the relevant gate was collected per sample. Analysis was performed with either Accuri C6 software or FlowJo (TreeStar) version 10.8. The FL-1 channel was used to monitor GFP expression. To determine % GFP positive populations, mESCs cultured in complete media were first used to bisect FL-1 histograms into GFP negative or positive populations. A Zeiss Axiovert epifluorescence microscope equipped

with a black and white AxioCam camera was used to capture brightfield and fluorescence images. Live cell images for differentiation experiments were taken at the indicated day in the corresponding media before washing with DPBS. For immunostaining, cells were washed 2x DPBS, fixed with 4% PFA/PBS (RT, 10 min.), blocked, and stained using primary antibodies (4 °C, overnight) and fluorophore-labeled secondary antibodies (RT, 3 hours).

Western blotting of Erk phosphorylation

Oct4-GFP mESCs were used for all stimulation experiments. Cells were plated ($1x10^5$ cells/cm^2) onto gelatinized six well plates in mESC growth medium. After 8-12 hours, cells were washed with DPBS, and serum starved (~18 hrs) by replacing with FBS-free mESC growth medium. Stimulation was performed by adding FGF2 (25 ng/mL) in FBS-free mESC media to cell monolayers for 10-15 mins at 37°C, 5% CO2. The plate was then immediately incubated on ice, and total protein was extracted using 1x cell lysis buffer and scraping. After 10 mins of incubation on ice, the mixture was centrifuged at 16,000 xg for 10 min (4 °C) to pellet and remove insoluble components. The supernatant was subjected to a BCA assay to quantify total protein levels, and upon normalization, 10 µg of protein was separated by SDS-PAGE (10% Tris-Glycine-SDS) and transferred onto a PVDF membrane. Anti-phosphoErk and anti-Erk antibodies were used to probe for levels of phosphorylated and total Erk protein levels. Densitometry was performed using ImageJ. RTK analysis. The instructions supplied with the Mouse Phospho-RTK Array Kit (R&D Systems Cat. #ARY014) were closely followed as a protocol. Oct4-GFP mESCs were prepared as above (see Western blotting for Erk phosphorylation), except cells were

seeded in a gelatinized T-75 flask. A total of 250 µg of whole cell lysate was used for each individual array, which includes duplicates spots of control and capture antibodies for different RTKs. Pixel density was determined via Adobe Photoshop (v 5.0) as the mean intensity of each capture antibody spot subtracted by the mean intensity of the PBS control spots.

qPCR analysis

Primers were obtained from IDT Technologies. Total RNA was extracted from cells in adherent culture after washing 2x DPBS, and following manufacturer's instructions for subsequent processing (Qiagen RNeasy Mini Kit). RNA purity and levels were assessed by UV analysis (NanoDrop), and lysates were stored at -20 °C until ready for processing. 50 ng of total RNA was used for cDNA synthesis, and gene expression was assessed using SYBR Green as a probe and an Applied Biosystems HT 7900 instrument.

Statistical analysis

All mathematical analyses were performed using GraphPad Prism (v6.0). The statistical significance of a single comparison was performed using the built-in analysis (Student's t-test), and multiple comparisons to a single control were conducted using the Dunnett's test (multiple comparison t-test). In general, each condition was conducted in duplicate in each experiment, and at least two independent biological replicates were used to derive conclusions. Thresholds for significance for all tests is set as $p<0.05$ (*), $p<0.01$ (**), $p<0.001$ (***), $p < 0.0001$ (****).

3.5 Supporting Information

Table 3.1 Reagents, tissue culture materials, antibodies, and sources. The following reagents and materials were used for work with the small molecule surfen.

Product Name	Manufacturer/Distributor	Catalog Number
Cell lines		
E14Tg2a wild-type mESC (E14)	*gift from Cathy Merry (Univ. of Manchester)*	
Oct4-GFP mESCs	PrimCells	PCEMM08
Sox1-GFP mESCs	PrimCells	PCEMM01
Reagents		
Surfen hydrate (cas # 5424-73-3) available at NCI # 12155	Sigma Aldrich	S6951
Sodium chlorate (cas # 7775-09-9)	Acros Organics	223222500
Adhesamine	Calbiochem	362331(discnt)
Protamine sulfate (cas # 9009-65-8)	MP Biomedicals	151971
PD173074 (cas # 219580-11-7)	TSZ Chemicals	RP04
Heparin	Carbosynth	OH03833
Tissue culture reagents		
DPBS without Ca and Mg	Corning	21-031
Non-essential amino acids	Gibco	11140-050
Porcine Gelatin	Sigma	G1890
KO-DMEM	Gibco	10829-018
L-glutamine (200 mM)	Gibco	25030-081
0.05% Trypsin-EDTA	Gibco	25300-054
Heparinase I and III	Sigma Aldrich	H3917-250UN
Heparinase II	Sigma Aldrich	H6512-25UN
Neurobasal medium	Gibco	21103-049
β-mercaptoethanol	Gibco	21985-023
LIF (leukemia inhibitory factor)	Millipore ESGRO	MESG1106
DMEM/F-12 medium	Gibco	11330-032
Knockout Serum Replacement (KSR)	Gibco	10828028
N2 supplement	Gibco	17502048
B27 supplement	Gibco	17504044
Penicillin/Streptomycin (100X)	Sigma	P4333
Western blotting materials		
Cell Lysis Buffer	Cell Signaling Technology	9803
Protease Inhibitor Cocktail	Cell Signaling Technology	5872
PMSF	Cell Signaling Technology	8553
BCA Assay	Thermo Scientific	23225
Luminata HRP substrate	Thermo Scientific	WBLUF0100
Restore Plus Western blot stripping buffer	Thermo scientific	21059
FGF2	Gibco	PHG0264
Bovine serum albumin (BSA)	Spectrum Chemicals	A3611
Tween-20	Fisher Scientific	BP337

Table 3.1 Reagents, tissue culture materials, antibodies, and sources. (continued)
The following reagents and materials were used for work with the small molecule surfen.

Product Name	Manufacturer/Distributor	Catalog Number
ECL Amersham Hyperfilm	GE healthcare	28906839
Protein ladder	Lambda Biotech	G02101
Immobilon FL PVDF membrane	Millipore	IPFL00010
Antibodies		
anti-nestin	Millipore, Clone: 401	MAB353
anti-Oct4	Santa Cruz Biotechnology	sc-25401
chicken anti-Sox1 polyclonal Ab	Millipore	AB5934
anti-β-III-tubulin	Abcam; Clone:EP1569Y	AB52623
anti-HS antibody (F58-10E4)	Amsbio; Clone:F58-10E4	370225
3G10 antibody	Amsbio; Clone: F69-3G10	370260
anti-phosphoErk1/2	Cell Signaling Technology Rbt, Clone: d13.14.4E	4370
anti-Erk	Cell Signaling Technology Rbt, Clone: 137F5	4695
anti-α tubulin	Cell Signaling Technology Mse, Clone: DM1A	3837
anti-SSEA1 (480) Alexa Fluor 647	Santa Cruz Biotechnology	sc-21702
IgM, goat anti-mouse (R-PE)	Molecular Probes	M31504
HRP-goat anti-mouse antibody	Cell Signaling Technology	7076
HRP-goat anti-rabbit antibody	Cell Signaling Technology	7074
AlexaFluor647 goat pAb to chk IgY	Abcam	GR272372-1
Fluoromount-G mounting media	Southern Biotech	0100-01
qPCR materials		
SYBRGreen	Applied Biosystems	4367659
RNAse-Free DNAse	Qiagen	79254
RNeasy Mini Extraction Kit	Qiagen	74104
High capacity cDNA reverse transcription kit	Applied Biosystems	4368814
Heparin ELISA		
TMB substrate	eBioscience	00-4201-56
Heparin coated plates	Bioworld	20140005-3

Stock solutions and storage

Surfen and adhesamine stocks were prepared as 30 mM solutions in DMSO (anhydrous, molecular biology grade), and stored at -20 °C. PD173074 stocks were prepared as 5 mM solutions in DMSO. Protamine Sulfate (source: salmon sperm; 19% sulfate, 24.3% nitrogen; may contain insolubles and histones) stock solutions were prepared as a 20 mg/mL solution, and stored at -20 °C. Protamine has an estimated

molecular weight of 10 KDa (used for the analysis in Fig. 2). Heparin was prepared as a 100 µg/mL solution in doubly-distilled water, stored at 4 °C.

In vitro culture of mESCs

Occasionally, Oct4-GFP mESCs were enriched for GFP-positive populations by treatment with 1 µg/mL puromycin in mESC media for two to three days. Gelatin was prepared as a 1.0 % (w/v) autoclaved stock solution in water, and stored at 4 °C. mESCs are cultured in plastic tissue-culture treated plates that were pre-treated for > 10 mins at RT with 0.1% gelatin/DPBS, prepared by a 1:10 dilution of the stock solution in DPBS.

mESC differentiation (N2B27, KSR)

Stock solutions of compounds were directly dissolved in N2B27 at the highest concentration. In our hands, we found that DMSO concentrations > 0.025% (v/v) consistently affected pluripotency (data not shown).[53] Thus, we ensured that all compounds were dissolved at a significantly lower DMSO composition. Experiments with the addition of soluble heparin were performed by first adding N2B27 + additive into the well containing cells, and then heparin to the desired final concentration.

IC50 determination

To determine IC_{50} values, %GFP+ve values for each cell line was plotted against the logarithm of surfen concentrations in nM. The data points were then fitted using

GraphPad Prism (v6) using a non-linear curve (equation: log(inhibitor) vs. response – variable slope (four parameters)). Oct4-GFP: R2 = 0.9914, IC_{50} = 1619 nM, Hill Slope = 6.808. Sox1- GFP: R2= 0.9978, IC_{50} = 1848 nM, Hill Slope = -7.120.

Differentiation of surfen-treated mESCs

Following a six-day differentiation protocol (in N2B27) as outlined above, cells were washed with DPBS, detached into single cells. Viable cells were counted (via trypan blue exclusion) and re-plated (10,000 cells/cm^2) in N2B27 onto new gelatin-coated twelve-well plates. After an additional six days in N2B27, cells were detached, neutralized, and analyzed by flow cytometry as before. Each experiment was performed with technical duplicates, and repeated successfully with eight biological replicates (Figure 3.9) to account for variations in differentiation efficiencies.

Immunostaining and fluorescence imaging

Live cells were washed twice with DPBS, fixed with 4 % PFA/PBS (RT, 10 min), blocked with immunostaining buffer (1 % BSA, 0.1 % Triton-X 100, 1 % goat serum/DPBS) for 1 hr at RT, then stained with primary antibodies in buffer (4 °C, overnight). Oct4 antibody was used at 1:100 dilution, Sox1 at 1:200 dilution, and nestin at 1:300 dilution. Following washes with immunostaining buffer, secondary antibodies were used at a 1:500 dilution in immunostaining buffer, and incubated with the cells for 2-3 hrs at RT. Hoescht was used at 10 µg/mL final concentration, and incubated with the cells for 10 mins at RT. Following washes, cells were mounted in Fluoromount-G

mounting media. Fluorescence micrographs were obtained either with a Zeiss Axiovert A.1 epifluorescence microscope or a Keyence BZX-700 Fluorescence Microscope.

Viability Assay

Oct4-GFP mESCs were seeded onto gelatin-coated 96-well plates in mESC complete medium overnight. After washing 1x DPBS, cells were incubated in increasing dosage of surfen in mESC media. After 48 hours in culture, cells were washed 2 x DPBS, and re-incubated in 100 µL of KO-DMEM. 20 µL of CellTiter Aqueous reagent (Promega) was then added to each well, and the plate was incubated at 37 °C for 2 hours. Absorbance at 590 nm was then read with a plate reader (Envision Wallac) and viability was quantified. A 100 % death consisting of cells-treated for 2 hours with 0.1 % Triton X-100 was used. This experiment was conducted in triplicate conditions, and two biological replicates. A representative experiment is shown in Figure 3.15.

Staining cell surface HS with antibodies via flow cytometry

Oct4-GFP mESCs cultured on gelatin-coated flasks were re-incubated with complete mESC medium or medium containing 10 % (v/v) of Heparinase I, II, and III (1 U/mL). After 18 hr incubation at 37 °C, 5 % CO_2, cells were washed 2x DPBS, detached with cell dissociation buffer, and fixed in 1% PFA/PBS on ice for 30-60 min. After washing 2x with DPBS, cells were re-suspended to $1x10^7$ cells/mL 0.1 % BSA/PBS, and stored at 4 °C, until ready for staining. To stain for cell surface HS, cells were incubated with either F58-10E4 (1:100) or 3G10 antibody (1:100) for 1 hr on ice, with periodic agitation. After

washing twice with 0.1 % BSA/PBS, cells were stained with PE-anti-mouse IgM (1:1000) or AF555- α-mouse (1:1000), respectively, for 1 hr on ice, with periodic agitation. After two washes, cells were re-suspended in 0.1 % BSA/PBS for flow cytometry analysis.

FGF2 stimulation experiments

Oct4-GFP mESCs (used for all stimulation experiments) were plated ($1x10^5$ cells/cm^2) onto gelatinized six-well plates in mESC growth medium. After 8-12 hours, cells were washed with DPBS, and serum starved overnight (~18 hrs) by replacing with FBS-free mESC growth medium**.** Stimulation was performed by adding FGF2 (25 ng/mL) in FBS-free mESC media to cell monolayers for 15 mins at 37 °C, 5 % CO_2. The plate was then immediately placed on ice, and total protein was extracted (by scraping) using 1x Cell Lysis Buffer with 1x Protease Inhibitor Cocktail and PMSF (1 mM). For experiments with inhibitor (e.g. surfen or PD173074), cells were pre-incubated for ten minutes at 37 °C, 5 % CO_2, prior to the addition of FGF2. For experiments with soluble heparin, heparin (5 µg/mL final concentration) was added after incubation with the inhibitor, and cells were incubated for an additional 10 minutes prior to the addition of FGF2.

Western blot experiments

Cell lysates were quantified for protein content using a standard BCA assay, and 5 µg of total protein was resolved on 10 % SDS-PAGE gels and subsequently blotted onto Immobilon-FL 0.45 µm membranes. Membranes were blocked in 5 % (w/v) BSA in tris-buffered saline containing 0.1 % Tween-20 (TBST) for 1 hour at room temperature.

Primary antibody incubations were performed overnight at 4 °C using anti-phospho-ERK1/2 (1:4000), anti-total ERK1/2 (1:5000), or anti-alpha tubulin (1:25,000). Blotted membranes were then washed 3x with TBST and subsequently incubated with HRP-conjugated secondary antibodies, anti-rabbit HRP (1:10,000) or anti-mouse HRP (1:10,000) for 1 hr at RT. After secondary incubations, membranes were washed 3x with TBST and developed using Luminata Forte HRP Detection Reagent and ECL Amersham Hyperfilm. Membranes were stripped and sequentially stained according to the following procedure. Membranes were incubated with Restore PLUS Western blot stripping buffer for 25 mins at RT, and washed 3x with TBST and blocked with 5 % BSA/TBST for 1 hr at room temperature before additional antibody incubations were conducted. Densitometry was performed using ImageJ analysis software (National Institute of Health), phospho-ERK1/2 and total-ERK1/2 levels were normalized to α-tubulin levels, then phospho-ERK1/2 was normalized to total-ERK1/2.

FGF2 ELISA

Heparin-coated plates (3 µg heparin per well) were first blocked for 1 hr at RT with 2 % BSA/PBS. After washing 3x with 0.05 % (v/v) Tween 20/PBS (PBST), the plate was incubated with dilutions of surfen (0.6 to 40 µM) in 1 % BSA/PBS, in triplicate wells, for 10-15 mins at RT. Without washing the wells, FGF2 was added to each well to a final concentration of 10 nM, and the plate was left to incubate at RT for 2 hours. After washing 3x with PBST, the wells were incubated with mouse anti-FGF2 antibody (1 µg/mL) in 1 % BSA/PBS, for 1 hr at RT. After washing 3x with PBST, the wells were incubated with HRP-conjugated goat anti-mouse antibody (1:1000) in 1 % BSA/PBS for another hour at RT.

After washing 3x with PBST, 1X TMB substrate (100 µL) was added at RT. After 2-5 mins, the reaction was quenched with 2 N sulfuric acid (100 µL), and the absorbance was read at 450 nm.

Real Time-PCR

Primers were obtained from IDT Technologies and stored as 100 µM solutions in doubly-distilled water. RNA extraction was performed according to manufacturer's instructions, using a RNeasy® Mini Kit. Following treatment of cells with RLT lysis buffer (+ 1 % β-mercaptoethanol), cell lysates were homogenized and loaded onto the RNeasy® Mini spin columns. Following column capture of nucleic acids from the lysates, DNase was utilized to remove contaminating genomic DNA. Spin columns were treated with DNase I (27.3 Kunitz units in 80 µL) for 15 minutes as instructed by the kit procedure. Purification of RNA samples was followed by UV-Vis spectroscopy (NanoDrop 2000c, Thermo Scientific) to assess RNA concentration and purity (A260/A280 ~1.8 - 2.1). The RNA samples were then converted to cDNA using a High-Capacity cDNA Reverse Transcription Kit. Finally, real-time PCR (Applied Biosystems 7900 HT) was conducted using a 384-well plate, with each well composed of 5 µL cDNA (10 ng/µL; 50 ng), 1 µL primers (10 µM forward and reverse), 10 µL of 2X SYBRGREEN and 4 µL nuclease-free water.

3.6 Supporting Figures

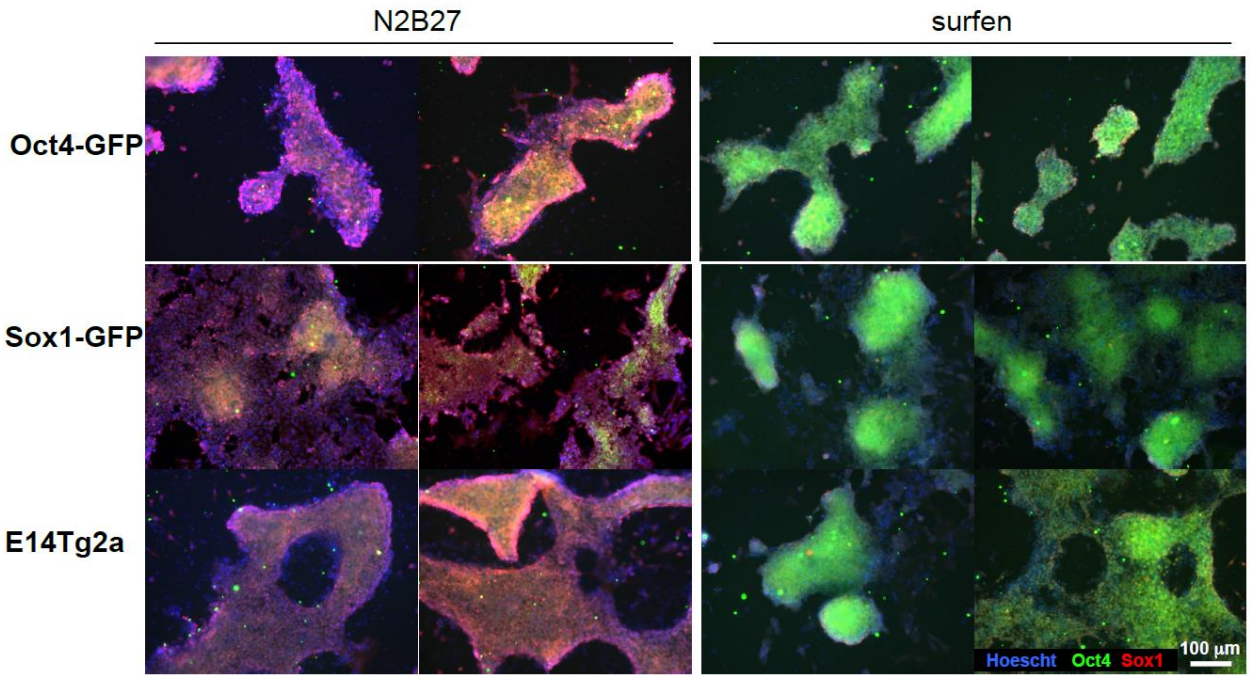


Figure 3.7 Surfen maintains pluripotency in Oct4-GFP, Sox1-GFP, and wild-type E14Tg2a mESCs after six days of N2B27 differentiation. Duplicate fluorescence images of mESCs incubated without (left panels) or with surfen (5 µM; right panels) for six days in N2B27 differentiation media, and immunostained for Oct4 (green), Sox1 (red) and DNA (blue). Surfen inhibits Sox1 and maintains Oct4 expression. Non-GFP channels were used for immunostaining to avoid fluorescence bleed-over.

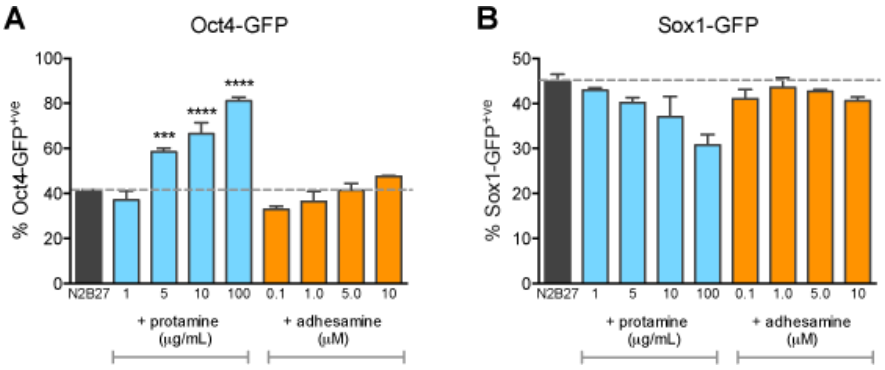


Figure 3.8 Flow cytometry evaluation of the effects of adhesamine (0.1-10 μM) and protamine (1-100 μg/mL) towards the differentiation of Oct4-GFP and Sox1-GFP mESCs. Protamine MW is 10 kDa. Protamine promotes Oct4-GFP expression but does not inhibit Sox1-GFP expression. Adhesamine does not significantly alter Oct4-GFP nor Sox1-GFP expression levels. Dunnett's multiple comparison test against N2B27 untreated control, ****p*<0.0001. Shown are technical duplicates (mean ± SD), repeated with two biological replicates.

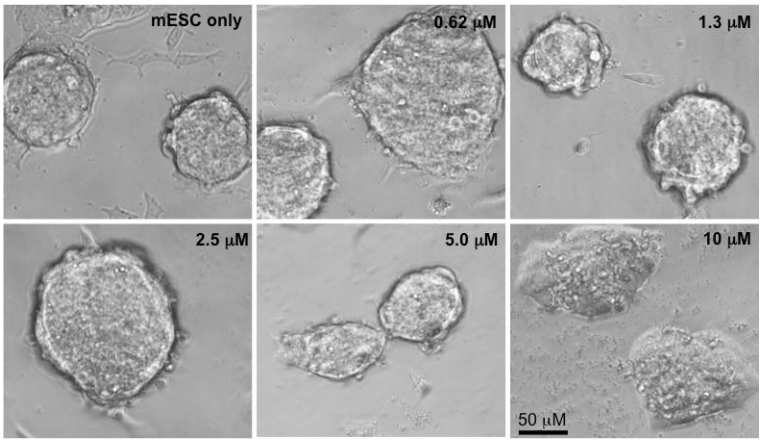


Figure 3.9 Brightfield microscopy images of mESCs following 24 hr incubation with surfen. At higher concentrations of surfen (10 μM), artifacts on the substrate are visible.

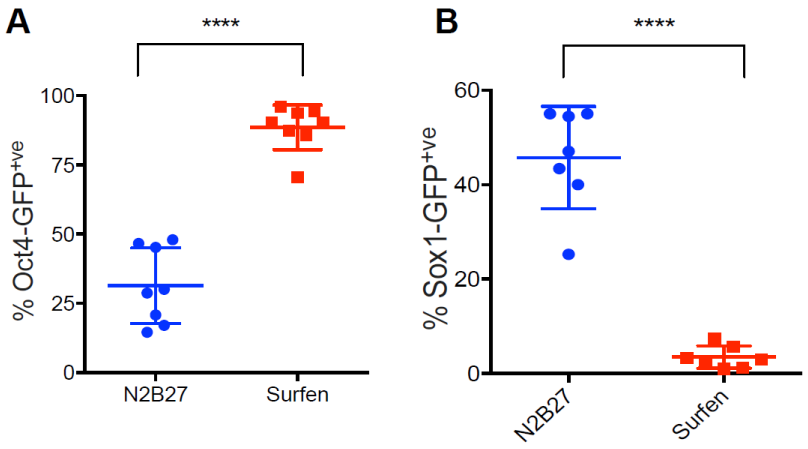


Figure 3.10 Scatter plot of biological replicates of %GFP-positive cells in Oct4-GFP and Sox1- GFP mESCs following six days of N2B27 differentiation with or without surfen. Despite variations in differentiation efficiencies, surfen consistently maintains Oct4 expression and inhibits Sox1 expression after 6 days of differentiation in N2B27. Consistent with reports in literature, Sox1-GFP mESCs can be differentiated into 50 ± 10 % GFP-positive cells.[54,55] Paired t-test against N2B27 untreated control, **** $p < 0.0001$. Shown are eight (Oct4-GFP) or seven (Sox1-GFP) pairs (mean ± SD).

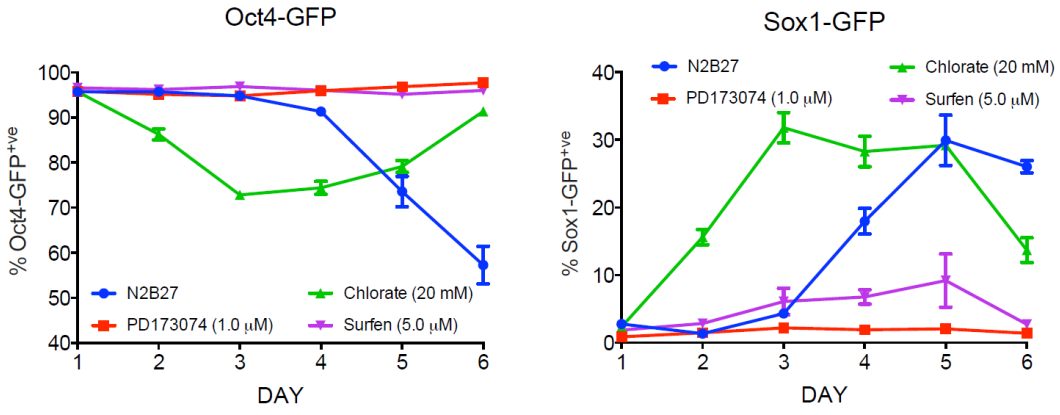


Figure 3.11 Flow cytometry evaluation of % GFP-positive populations each day of N2B27 differentiation. Surfen maintains pluripotency throughout six days of N2B27 differentiation.

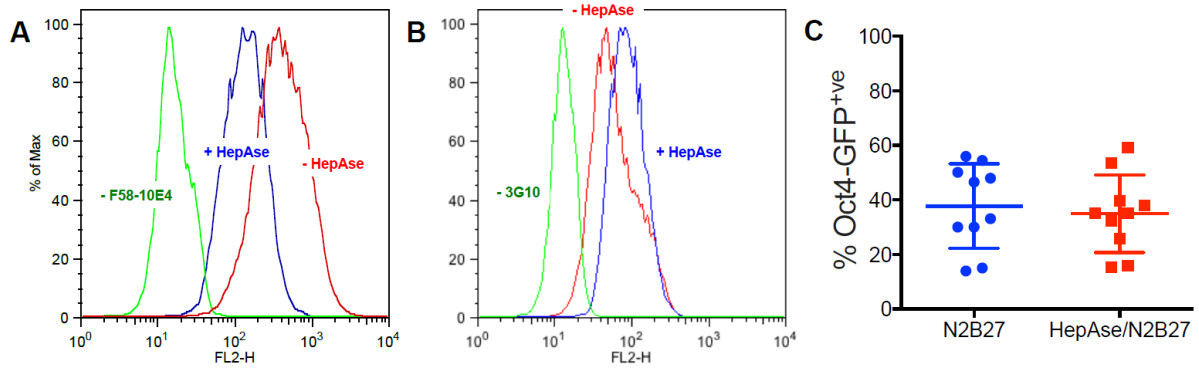


Figure 3.12 Flow cytometry evaluation of cell-surface HS levels following heparinase treatment, and its effect towards the differentiation of Oct4-GFP mESCs. Heparinase treatment causes reduction of cell surface HS chains but does not maintain pluripotency. Oct4-GFP mESCs cultured in complete medium with heparinase and stained with (A) F58-10E4 antibody or (B) 3G10 antibody. Cells cultured in heparinase show a reduction in geometric mean compared to non-treated cells when probed with F58- 10E4, an HS antibody. Cells cultured in heparinase show an increase in 3G10 staining, which detects the enzymatic stub resulting from HS digestion, (C) Heparinase treatment did not significantly affect Oct4 levels compared to untreated control after 6 days in N2B27 differentiation media.

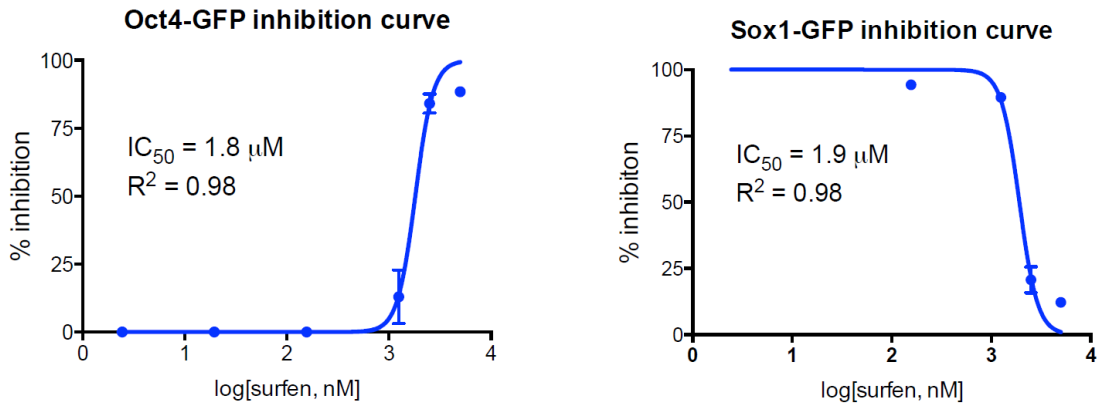


Figure 3.13 Dose-dependent inhibition of differentiation with surfen treatment. Individual curves from figure 3.2E.

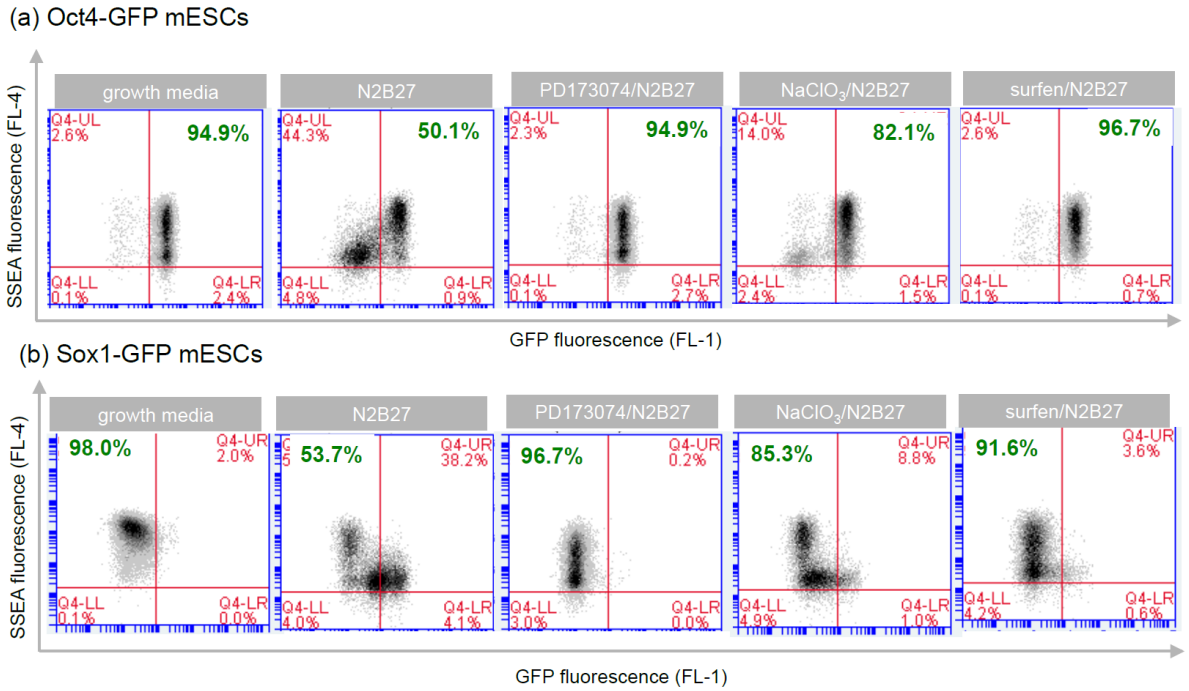


Figure 3.14 Flow cytometry evaluation of cells after 6 days of N2B27 differentiation, analyzed by SSEA-1 immunostaining and GFP fluorescence. Surfen-treated cells display high SSEA expression in both (A) Oct4-GFP and (B) Sox1-GFP mESCs after 6 days of N2B27 differentiation.

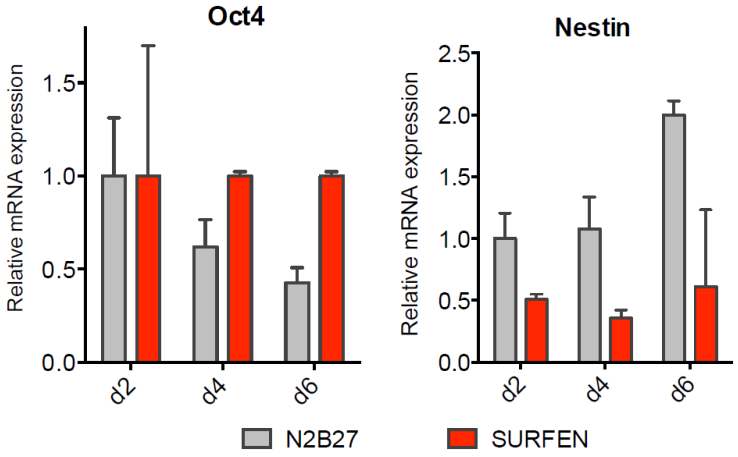


Figure 3.15 qRT-PCR analysis of surfen-treated cells during N2B27 differentiation. Surfen (5.0 μM) maintains the expression of the pluripotency marker, *Oct4*, and inhibits the expression of neuroectodermal marker, *nestin*.

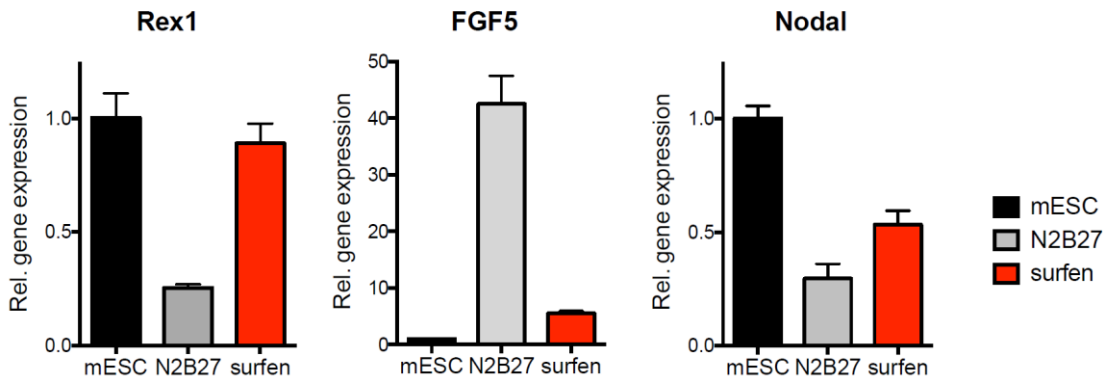


Figure 3.16 Surfen maintains pluripotency and does not cause Oct4-GFP mESCs to enter an epiblast state. qPCR analysis of cells following six-day treatment with or without surfen in N2B27 media show that surfen maintains pluripotency (high *Rex5* expression), but does not cause an epiblast state, asevidenced by low expression of epiblast markers (*FGF5*, *Nodal*).[56]

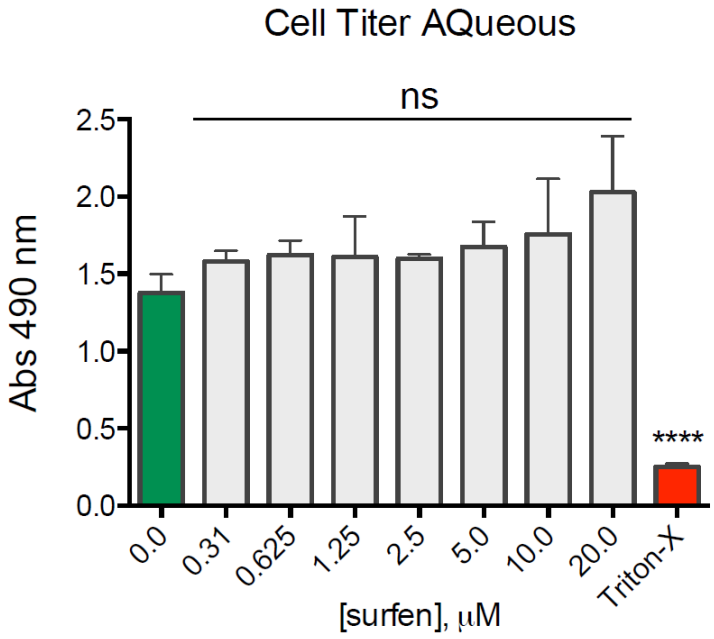


Figure 3.17 Effects of surfen towards Oct4-GFP mESC viability. CellTiter Aqueous experiment shows that surfen does not exhibit decreased cell viability at 5.0 μM. Dunnett's multiple comparison test against untreated control, ****p*<0.0001. Shown are technical triplicates (mean ± SD), repeated with two biological replicates.

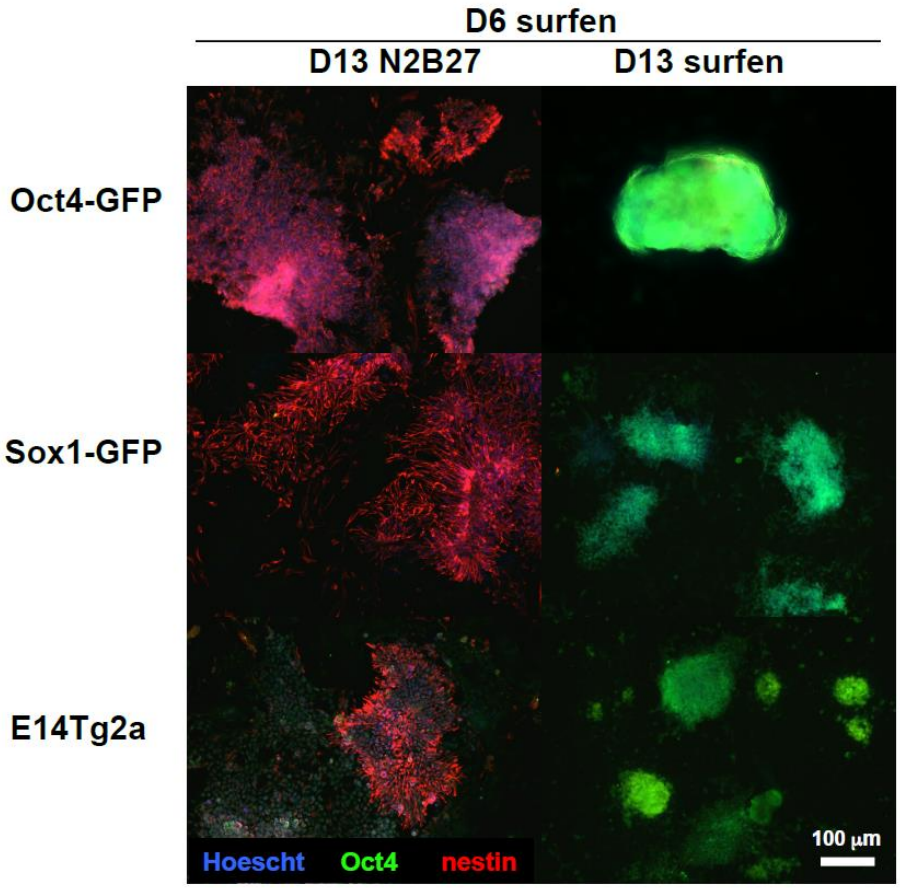


Figure 3.18 Withdrawal of surfen at D6 allows differentiation to proceed, whereas continuous treatment maintains pluripotency. Fluorescence microscopy images of D13 cells, stained for pluripotency marker Oct4 (green) and the neuroectodermal marker nestin (red), following removal of surfen at D6 (left panels) or continuous treatment with surfen (right panels). High nestin and low Oct4 staining was observed in surfen-withdrawn cells, whereas low nestin and high Oct4 staining was observed in cells continuously treated with surfen. These observations are consistent in Oct4-GFP, Sox1-GFP, as well as E14Tg2a wild-type mESCs.

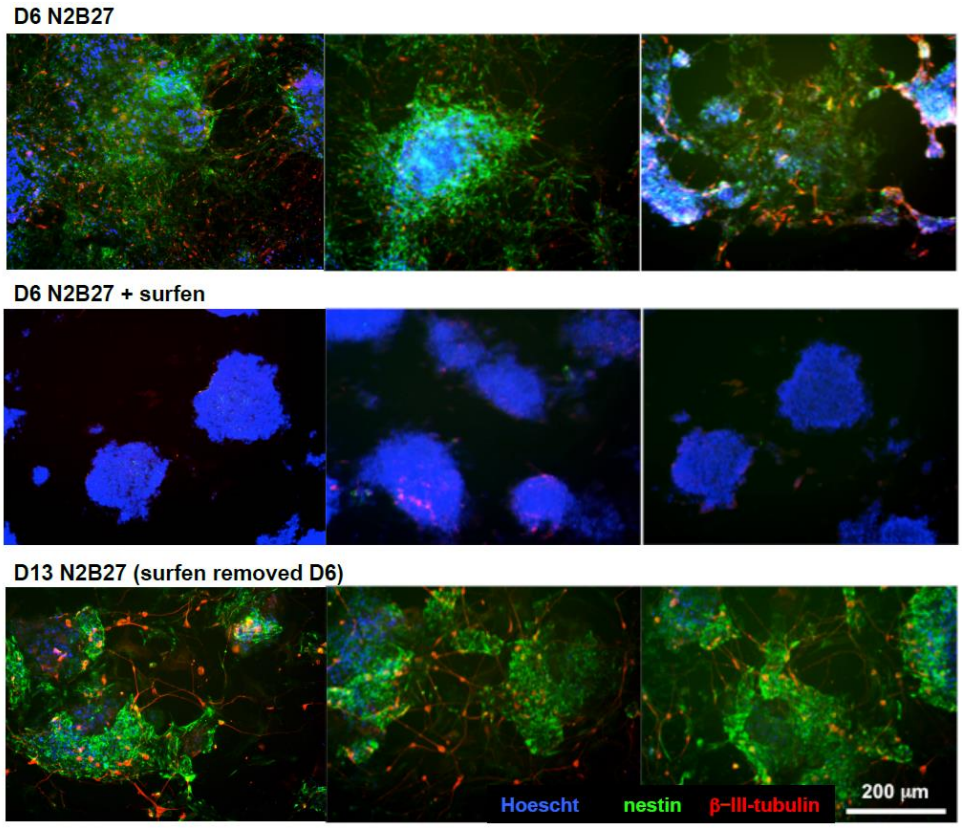


Figure 3.19 Withdrawal of surfen at D6 allows robust neural differentiation to occur. Triplicate fluorescence microscopy images of Oct4-GFP cells, stained for neural differentiation markers nestin (green) and α-III-tubulin (red), six days after differentiation without (top panels) or with surfen (middle panels). Surfen inhibited expression of nestin and α-III tubulin. Following these six days of surfen treatment, surfen was removed and the cells were allowed to differentiate for an additional seven days in N2B27 (bottom panels). Neural differentiation, as evidenced by high nestin and α-III tubulin cells, occurred robustly in these latter cells.

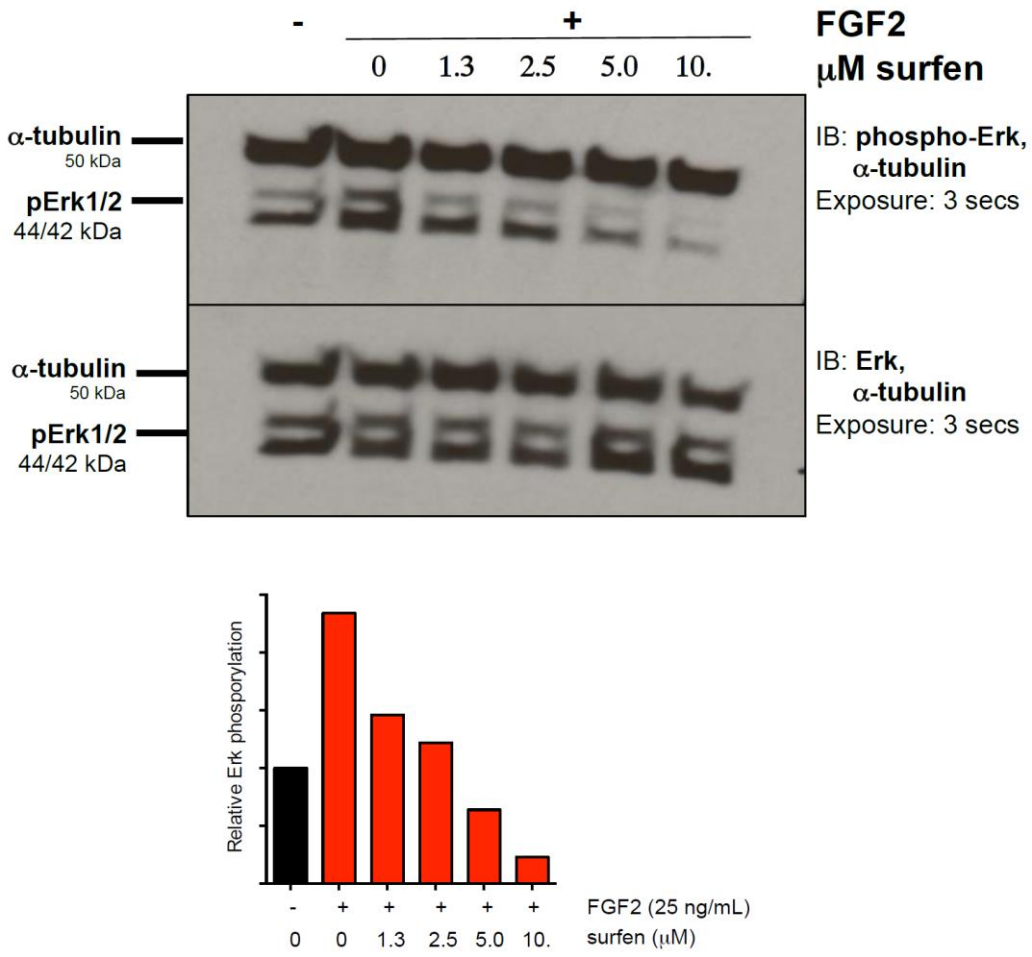


Figure 3.20 Surfen inhibits Erk phosphorylation in a dose-dependent manner. Original Western blot images (top) and corresponding densitometry bar graph (bottom) of Oct4-GFP mESC lysates treated with increasing dosages of surfen. (Figure 3.4A).

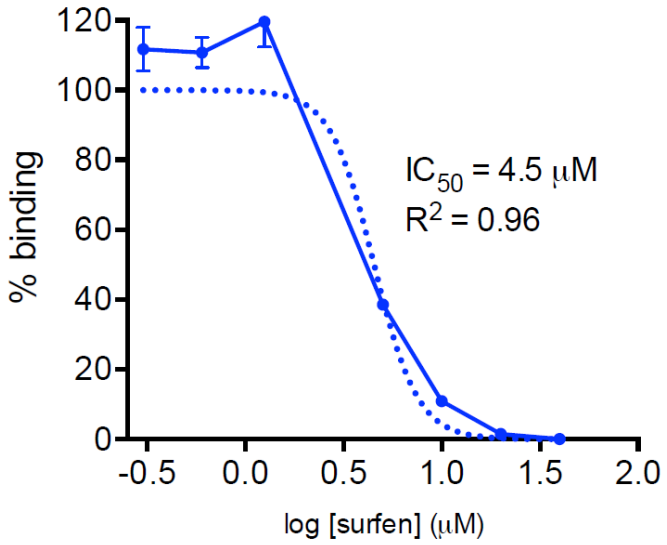


Figure 3.21 Surfen inhibits FGF2 binding to heparin. Heparin-coated ELISA plates were treated with FGF2 and increasing amounts of surfen.

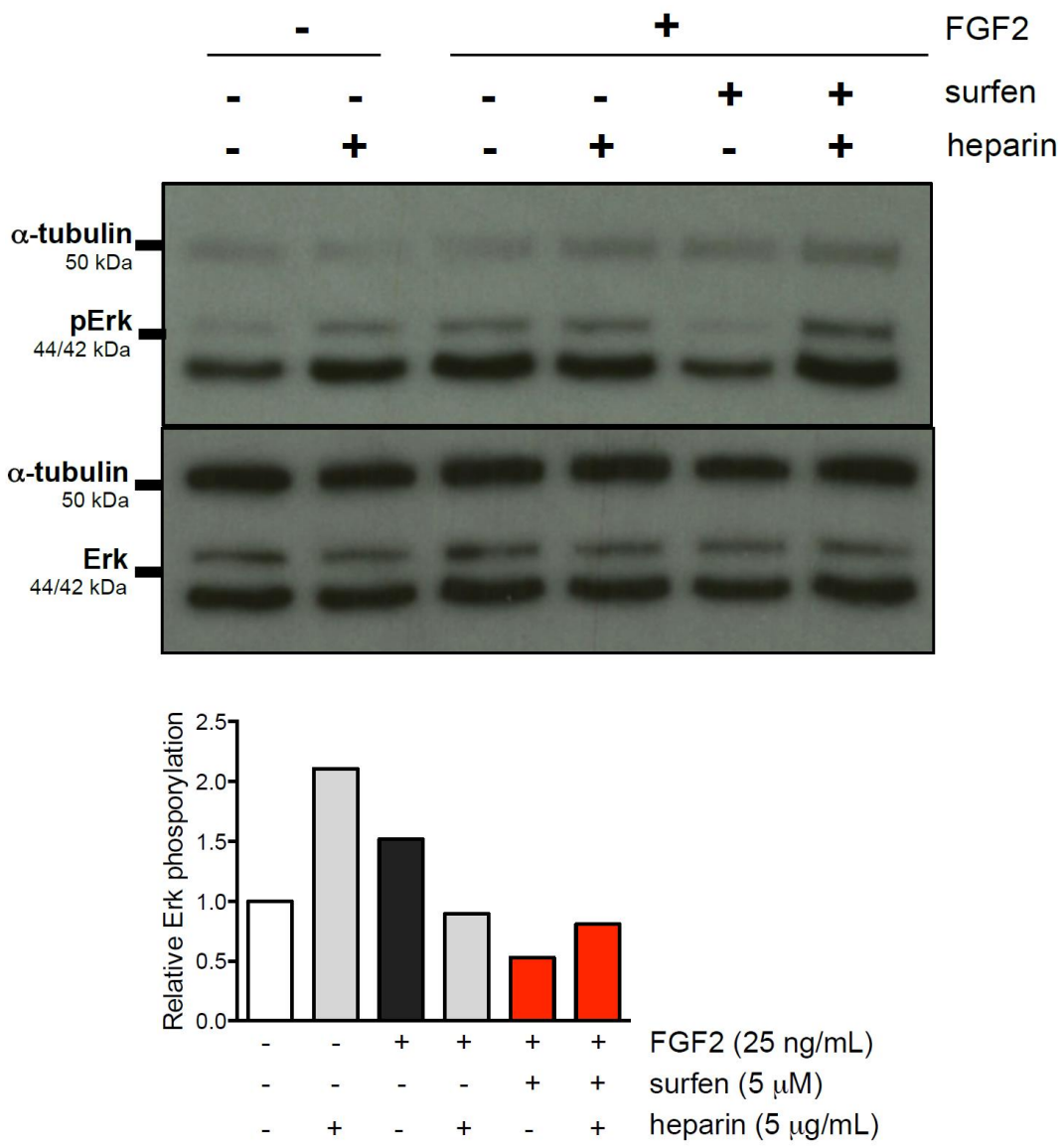
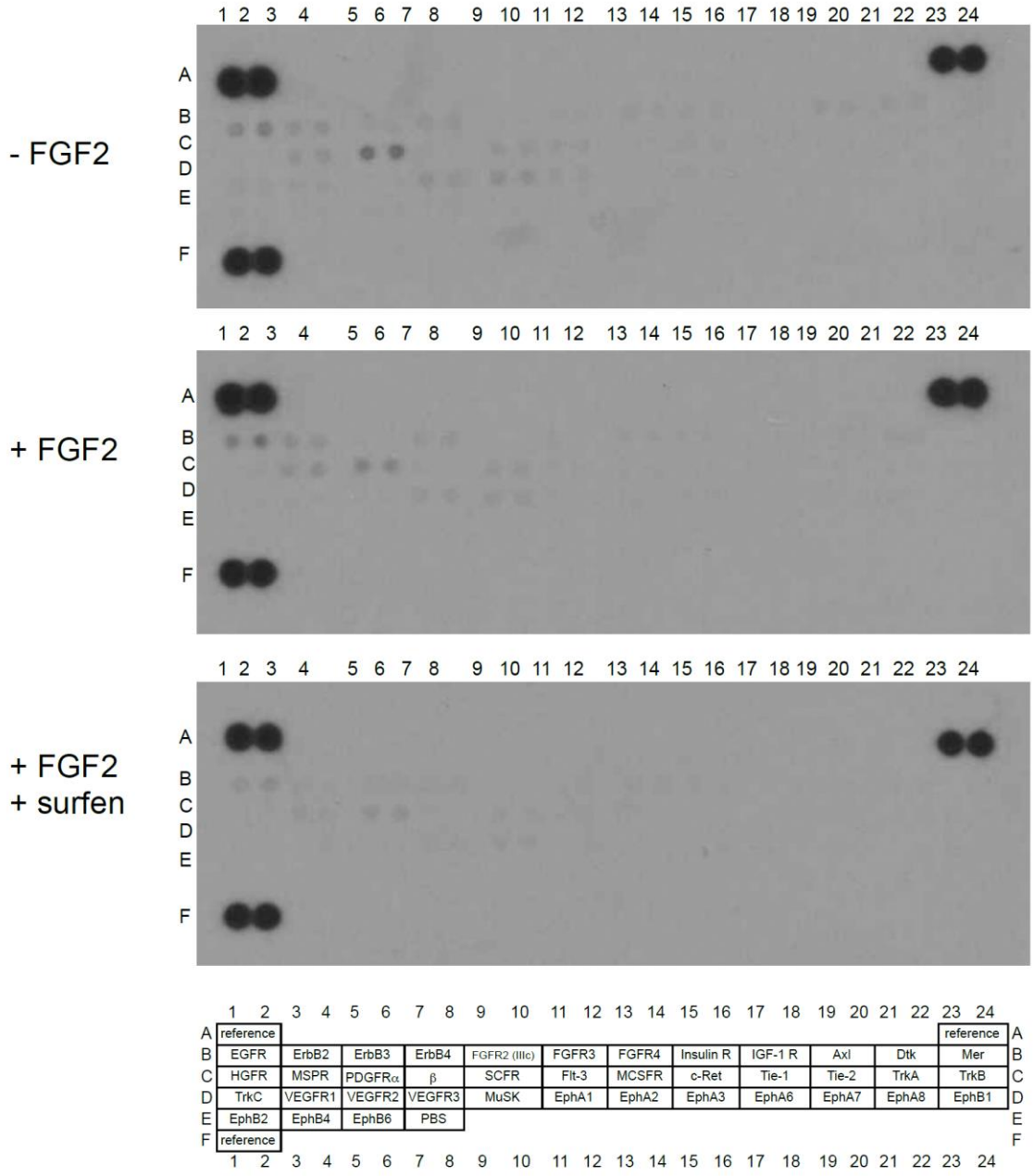


Figure 3.22 Soluble heparin (5 µg/mL) rescues Erk phosphorylation of surfen-treated (5 µM) Oct4- GFP mESCs. Original Western blot images (top) and corresponding densitometry bar graph (bottom) of lysates treated with surfen and rescued with soluble heparin (Figure 3.B).



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
A	reference																						reference		A
B	EGFR	ErbB2	ErbB3	ErbB4	FGFR2 (IIIc)		FGFR3	FGFR4	Insulin R		IGF-1 R		Axl		Dtk		Mer								B
C	HGFR	MSPR	PDGFRα	β	SCFR		Flt-3	MCSFR	c-Ret		Tie-1		Tie-2		TrkA		TrkB								C
D	TrkC	VEGFR1	VEGFR2	VEGFR3	MuSK		EphA1	EphA2	EphA3		EphA6		EphA7		EphA8		EphB1								D
E	EphB2	EphB4	EphB6	PBS																					E
F	reference																								F
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	

Figure 3.23 RTK (Receptor Tyrosine Kinase) array analysis of surfen-treated embryonic Oct4-GFP mESCs following FGF2 activation. Surfen is a broad-spectrum phosphorylation inhibitor of many RTKs (Figure 3 . 5), similar to results observed for Ext1$^{-/-}$mESCs.[57] Duplicate capture antibodies (labeled in the table below) are spotted in these membranes (R&D Systems Cat. # ARY014). PBS control spots are located in E7-8. Increased PDGFRα expression is observed in these mESCs, although expression of this RTK is known to fluctuate within a population of mESCs.[58]

Table 3.2 Primers used for RT-PCR experiments. The following primers were used for DNA amplification in RT-PCR experiments.

AMPLICON NAME	Forward primer (5'-3'')	Reverse primer (5'-3')	Product size (bp)
GAPDH (housekeeping)	TGCCTGCTTCACCACCTTCT	CCAATGTGTCCGTCGTGGAT	83
β-actin (housekeeping	GGGGTGTTGAAGGTCTCAAA	TGTTACCAACTGGGACGACA	168
Oct4 (embryonic)	TTGCCTTGGCTCACAGCATC	TGTTCCCGTCACTGCTCTGG	82
Rex1 (embryonic)	GGCTGCGAGAAGAGCTTTATTCA	AGCATTTCTTCCCGGCCTTT	79
Sox1 (ectoderm)	GGCCGAGTGGAAGGTCATGT	TCCGGGTGTTCCTTCATGTG	93
β -III-tubulin (Tubb3) (ectoderm)	TGATGACGAGGAATCGGAAGC	GGACAGATGCTGCTTGTCTTGG	101
nestin (ectoderm)	CTACCAGGAGCGCGTGGC	TCCACAGCCAGCTGGAACTT	219
Brachyury "T" (mesoderm)	TTGAACTTTCCTCCATGTGCTGA	TCCCAAGAGCCTGCCACTTT	82
Foxa2 (mesoderm)	ACTGGAGCAGCTACTACG	CCCACATAGGATGACATG	152
Sox17 (endoderm)	AGCCATTTCCTCCGTGGTGT	AACACTGCTTCTGGCCCTCAG	104
FGF5 (epiblast)	CCTTGCGACCCAGGAGCTTA	CCGTCTGTGGTTTCTGTTGAGG	98
Nodal (epiblast)	ACTGAGGGCCCACTCACCAT	CGGTGAACGTCTCCATCCAA	103

3.7 Author Contributions

M.L.H.: conception and design, collection and assembly of data, data analysis and interpretation, manuscript writing; A.L.M.: collection and assembly of data, data analysis and interpretation; C.J. F. and M.C.: collection and assembly of data; R.A.A.S.: conception and design, data analysis and interpretation; K.G.: conception and design, manuscript writing, final approval of manuscript.

3.8 Acknowledgements

Chapter 3 is an adaptation and reprint of the material as it appears in *Stem Cells* 2018. Huang ML, Michalak AL, Fisher CJ, Christy M, Smith RAA, Godula K. Small Molecule Antagonist of Cell Surface Glycosaminoglycans Restricts Mouse Embryonic Stem Cells in a Pluripotent State. Austen Larson Michalak made contributions as a primary author on this publication.

3.9 References

1 Hochedlinger K, Jaenisch R. Nuclear transplantation, embryonic stem cells, and the potential for cell therapy. *N. Engl. J. Med.*, **2003**;349:275–286.

2 Wernig M, Meissner A, Foreman R et al. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature*, **2007**; 448:318–324.

3 Zhang Y, Li W, Laurent T, Ding S. Small molecules, big roles- The chemical manipulation of stem cell fate and somatic cell reprogramming. *J. Cell Sci.*, **2012**;125:5609– 5620.

4 Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. *Nature*, **2007**;448:313–317.

5 Ledermann B. Embryonic stem cells and gene targeting. *Exp. Physiol.*, **2000**;85:603–613.

6 Nicola NA, Babon JJ. Leukemia inhibitory factor (LIF). *Cytokine. Growth Factor. Rev.* **2015**;26:533–544.

7 Xu C, Rosler E, Jiang J, Lebkowski JS, Gold JD, O'Sullivan C, Delavan-Boorsma K, Mok M, Bronstein A, Carpenter MK. Basic fibroblast growth factor supports undifferentiated human embryonic stem cell growth without conditioned medium. *Stem Cells*, **2005**;23:315– 323.

8 Yoshihara M, Hayashizaki Y, Murakawa Y. Genomic instability of iPSCs: Challenges towards clinical applications. *Stem Cell Rev.*, **2017**;13:7–16.

9 Chen S, Do JT, Zhang Q, Yao S, Yan F, Peters EC, Schöler HR, Schultz PG, Ding S.. Self-renewal of embryonic stem cells by a small molecule. *Proc. Natl. Acad. Sci. USA*, **2006**;103: 17266–17271.

10 Li W, Li K, Wei W, Ding S. Chemical approaches to stem cell biology and therapeutics. *Cell Stem Cell*, **2013**;13:270–283.

11 Varki A, Freeze HH, Vacquier VD. Glycans in Development and Systemic Physiology. In: Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME, editors. *Essentials of Glycobiology*. 2nd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; **2009**. Chapter 38.

12 Fox N, Damjanov I, Martinez-Hernandez A, Knowles BB, Solter D. Immunohistochemical localization of the early embryonic antigen (SSEA-1) in post implantation mouse embryos and fetal and adult tissues. *Dev. Biol.*, **1981**;83:391–398.

13 Freeze HH. Genetic defects in the human glycome. *Nat. Rev. Genet.,* **2006**;7:537– 551.

14 Izumikawa T, Sato B, Kitagawa H. Chondroitin sulfate is indispensable for pluripotency and differentiation of mouse embryonic stem cells. *Sci. Rep.*, **2014**;4:3701.

15 Lin X, Wei G, Shi Z et al. Disruption of gastrulation and heparan sulfate biosynthesis. *Dev. Biol.*, **2000**;224:299–311.

16 Johnson CE, Crawford BE, Stavridis M, Dam GT, Wat AL, Rushton G, Ward CM, Wilson V, van Kuppevelt TH, Esko JD, Smith A, Gallagher JT, Merry CLR. Essential alterations of heparan sulfate during the differentiation of embryonic stem cells to Sox1-enhanced green fluorescent protein-expressing neural progenitor cells. *Stem Cells*, **2007**;25:1913–1923.

17 Kraushaar DC, Yamaguchi Y, Wang L. Heparan sulfate is required for embryonic stem cells to exit from self-renewal. *J. Biol. Chem*. **2010**;285:5907–5916.

18 Kraushaar DC, Dalton S, Wang L. Heparan sulfate: A key regulator of embryonic stem cell fate. *J. Biol. Chem*., **2013**;394:741–751.

19 Schnerch A, Cerdan C, Bhatia M. Distinguishing between mouse and human pluripotency regulation: The best laid plans of mice and men. *Stem Cells*, **2010**; 28:419– 430.

20 Michaud P, Da Costa A, Courtois B, Courtois J. Polysaccharide lyases: Recent developments as biotechnological tools. *Crit. Rev. Biotechnol.*, **2003**;23:233–266.

21 Harfouche R, Hentschel DM, Piecewicz, Basu S, Print C, Eavarone D, Kiziltepe T, Sasisekharan R, Sengupta S. Glycome and transcriptome regulation of vasculogenesis. *Circulation*, **2009**;120:1883– 1892.

22 Garud DR, Tran VM, Victor XV, Koketsu M, Kuberan B. Inhibition of heparan sulfate and chondroitin sulfate proteoglycan biosynthesis. *J. Biol. Chem,* **2008**;283:28881–28887.

23 Siegbahn A, Manner S, Persson A, Tykesson E, Holmqvist K, Ochocinska A, Rönnols J, Sundin A, Mani K, Westergren-Thorsson G, Widmalm G, Ellervik U. Rules for priming and inhibition of glycosaminoglycan biosynthesis; probing the b4GalT7 active site. Chem. Sci., **2014**;5:3501– 3508.

24 Van Wijk XM, Oosterhof A, van den Broek SAMW, Griffioen A, ten Dam GB, Rutjes FPJT, van Delft FL, van Kuppevelt TH. A 4-deoxy analogue of N-acetyl-D-glucosamine inhibits heparan sulphate expression and growth factor binding in vitro. *Exp. Cell Res*., **2010**;316:2504–2512.

25 Van Wijk XM, Lawrence R, Thijssen VL, van den Broek SA, Troost R, van Scherpenzeel M, Naidu N, Oosterhof A, Griffioen AW, Lefeber DJ, van Delft FL, van Kuppevelt TH. A common sugar-nucleotide-mediated mechanism of inhibition of (glycosamino) glycan biosynthesis, as evidence by 6F-GalNAc (Ac3). *FASEB J*., **2015**;29:2993–3002.

26 Humphries DE, Silbert JE. Chlorate: A reversible inhibitor of proteoglycan sulfation. *Biochem. Biophys. Res. Commun.*, **1988**;154: 365–371.

27 Rapraeger AC, Krufka A, Olwin BB. Requirement of heparan sulfate for bFGFmediated fibroblast growth and myoblast differentiation. *Science*, **1991**;252:1705–1708.

28 Baeuerle PA, Huttner WB. Chlorate: A potent inhibitor of protein sulfation in intact cells. *Biochem. Biophys. Res. Commun*., **1986**; 141:870–877.

29 Sasaki N, Hirano T, Kobayashi K, Toyoda M, Miyakawa Y, Okita H, Kiyokawa N, Akutsu H, Umezawa A, Nishihara S. Chemical inhibition of sulfation accelerates neural differentiation of mouse embryonic stem cells and human induced pluripotent stem cells. *Biochem. Biophys. Res. Commun.*, **2010**;401:480–486.

30 Lanner F, Lee KL, Sohl M, Holmborn K, Yang H, Wilbertz J, Poellinger L, Rossant J, Farnebo F. Heparan sulfation-dependent fibroblast growth factor signaling maintains embryonic stem cells primed for differentiation in a heterogeneous state. *Stem Cells*, **2010**;28:191–200.

31 Choi S, Clements DJ, Pophristic V, Ivanov I, Vemparala S, Bennett JS, Klein ML, Winkler JD, DeGrado WF. The design and evaluation of heparin-binding foldamers. *Angew. Chem. Int. Ed. Engl.*, **2005**;44: 6685–6689.

32 Hunter DT Jr., Hill JM. Surfen: A quinolone with oncogenic and heparin-neutralizing properties. *Nature*, **1961**;191:1378–1379.

33 Roan NR, Sowinski S, Munch J, Kirchoff F, Greene WC. Aminoquinoline surfen inhibits the action of SEVI (semen-derived enhancer of viral infection). *J. Biol. Chem*., **2010**;285:1861–1869.

34 Xu D, Fuster MM, Lawrence R, Esko JD. Heparan sulfate regulates VEGF165- and VEGF-mediated vascular hyperpermeability. *J. Biol. Chem.*, **2011**;286:737–745.

35 Warford J, Doucette CD, Hoskin DW, Easton AS. Murine T-cell activation is regulated by surfen (bis-2-methyl-4-amino-quinolyl-6-carbamide). *Biochem. Biophys. Res. Commun*., **2014**;443:524–530.

36 Schuksz M, Fuster MM, Brown JR, Crawford BE, Ditto DP, Lawrence R, Glass CA, Wang L, Tor Y, Esko JD. Surfen, a small molecule antagonist of heparan sulfate. *Proc. Natl. Acad. Sci. USA*, **2008**; 105:13075–13080.

37 Weiss RJ, Gordts PL, Le D, Xu D, Esko JD, Tor Y. Small molecule antagonists of cell-surface heparan sulfate and heparin-protein interactions. *Chem. Sci*., **2015**;6:5984–5993.

38 Ying QL, Nichols J, Evans EP, Smith AG. Changing potency by spontaneous fusion. *Nature* **2002**;416:545–548.

39 Ying QL, Stavridis M, Griffiths D, Li M, Smith A. Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat. Biotechnol*., **2003**;21:183–186.

40 Yamazoe S, Shimogawa H, Sato S, Esko JD, Uesugi M. A dumbbell-shaped small molecule that promotes cell adhesion and growth. *Chem. Biol.*, **2009**;16:773–782.

41 Wang J, Rabenstein DL. Interaction of heparin with two synthetic peptides that neutralize the anticoagulant activity of heparin. Biochemistry, **2006**;45:15740–15747.

42 Ying QL, Wray J, Nichols J, Batllee-Morera L, Doble B, Woodgett J, Cohen P, Smith A. The ground state of embryonic stem cell self- renewal. *Nature*, **2008**;453:519–523.

43 Brons IGM, Smithers LE, Trotter MWB, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, Vallier L. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature*, **2007**; 448:191–195.

44 Yayon A, Klagsbrun M, Esko JD, P Leder, DM Ornitz. Cell surface, heparin-like molecules are required for binding of fibroblast growth factor to its high affinity receptor. *Cell*, **1991**;64:841–848. 45

45 Feyzi E, Lustig F, Fager G, Spillmann D, Lindahl U, Salmivirta M. Characterization of heparin and heparan sulfate domains binding to the long splice variant of platelet-derived growth factor A chain. *J. Biol. Chem.*, **1997**;272:5518–5524.

46 Pankonin MS, Gallagher JT, Loeb JA. Specific structural features of heparan sulfate proteoglycans potentiate neuregulin-1 signaling. *J. Biol. Chem.,* **2005**;280:383–388.

47 Higashiyama S, Abraham JA, Miller J, Fiddes JC, Klagsbrun M. A heparin-binding growth factor secreted by macrophage-like cells that is related to EGF. *Science*, **1991**;251:936–939.

48 Pickford CE, Holley RJ, Rushton G, Stavridis MP, Ward CM, Merry CLR. Specific glycosaminoglycans modulate neural specification of mouse embryonic stem cells. *Stem Cells*, **2011**;29:629–640.

49 Ono K, Hattori H, Takeshita S, Kurita A, Ishihara M. Structural features in heparin that interact with VEGF165 and modulate its biological activity. *Glycobiology*, **1999**;9:705–711.

50 Qiu H, Jiang JL, Liu M, Huang X, Ding SJ, Wang L. Quantitative phosphoproteomics analysis reveals broad regulatory role of heparan sulfate on endothelial signaling. *Mol. Cell Proteomics,* **2013**;12:2160–2173.

51 Jackson M, Taylor AH, Jones EA, Forrester LM. The culture of mouse embryonic stem cells and formation of embryoid bodies. *Methods Mol. Biol.*, **2010**;633:1–18.

52 Wang YC, Stein JW, Lynch CL et al. Glycosyltransferase ST6GAL1 contributes to the regulation of pluripotency in human pluripotent stem cells. *Sci. Rep*., **2015**;5:13317.

53 Adler S, Pellizer C, Pararella M, Hartung T, Bremer S. The effects of solvents on embryonic stem cell differentiation*. Toxicol. In Vitro.,* **2006**;20:265-271.

54 Johnson CE, Brawford BE, Stavridis M, Dam GT, Wat AL, Rushton G, Ward CM, Wilson V, van Kuppevelt TH, Esko JD, Smith A, Gallagher JT, Merry CLR.. Essential alterations of heparan sulfate duringthe differentiation of embryonic stem cells to Sox1-enhanced green fluorescent protein- expressing neural progenitor cells. *Stem Cells,* **2007**;25:1913-1923.

55 Chen CW, Liu CS, Chiu IM, Shen SC, Pan HC, Lee KH, Lin SZ, Su HL. The signals of FGFs on the neurogenesis of embryonic stem cells. *J. Biomed Sci.,* **2010**;17:33.

56 Brons IGM, Smithers LE, Trotter MWB, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, Vallier L. Derivation of pluripotent epiblast stem cells from mammalian embyros. *Nature,* **2007**;448:191-196.

57 Pickford CE, Holley RJ, Rushton G, Stavridis MP, Ward CM, Merry CLR. Specific glycosaminoglycans modulate neural specification of mouse embryonic stem cells. *Stem Cells,* **2011**; 29:629-640.

58 Lo Nigro A, de Jaime-Soguero A, Khoueiry R, Cho DS, Relazzo GM, Perini I, Escalona VA, Aranguren XL, Chuva de Sousa Lopes SM, Koh KP, Conaldi PG, Hu W, Zwijsen, Lluis F, Verfaillie CM. PDGFRα^{+} cells in embryonic stem cell cultures represent the in vitro equivalent of the pre-implantation primitive endoderm precursors. *Stem Cell Reports,* **2017**; 8:318-333.

4. Glycocalyx photoengineering enables modeling of cell surface mucin shedding dynamics

4.1 Introduction

The epithelial glycocalyx, composed of membrane-associated glycoproteins and glycolipids, is an important functional component of the mucosal barrier that regulates interactions between the epithelium and various components of its external environment.[1,2] A class of extended, highly glycosylated proteins, known as mucins, which are expressed at high levels on mucosal cells and project away from the membrane, form a physical shield that protects cells from pathogenic challenge. Mucins do so by either limiting the access of pathogens to their cell-surface receptors[3,4] or by presenting decoy receptors to capture the pathogens and clear them from the cell surface via shedding.[5,6] In response, pathogens have evolved mechanisms to overcome the barrier functions of the glycocalyx, such as by expressing enzymes that can break down mucins and expose the cell surface.[7] Another possible mechanism through which pathogens may overcome clearance through shedding is by deploying multiple copies of glycan binding proteins (GBPs) to crosslink mucins and inhibit their release from the glycocalyx.[8] Crosslinking of the glycocalyx by extracellular lectins, such as galectins, has been previously shown to contribute to its stabilization and to reduce endocytic turnover of native[9] and synthetic[10,11] glycoconjugates. However, the effects of glycocalyx crosslinking by oligomeric GBPs, including pathogen associated lectins, on mucin shedding are yet to be investigated; likely due to the lack of tools to induce mucin shedding from cells with spatial and temporal control.

Genetic tools to control mucin structure and expression to tailor the physical characteristics of the mucosal glycocalyx are rapidly emerging.[12] While the recent isolation and characterization of a suite of mucinase enzymes[13,14] enables selective removal of mucins from the glycocalyx, this leads to complete digestion of the mucin structure and does not fully recapitulate the process of shedding which leaves the glycosylated mucin ectodomains largely intact and capable of interacting with crosslinking lectins.

Synthetic glycopolymers, which approximate the structure of mucins and can be introduced into cell membranes, have provided a useful tool for the modeling the mucinous glycocalyx to study its biological functions.[15,16] Here we describe cell surface engineering with mucin mimetics bearing photocleavable membrane anchors to model mucosal glycocalyx shedding with spatial and temporal control using light. The synthetic mucin mimetics showed membrane-density dependent crosslinking by the oligomeric lectin, *Ricinus communis agglutinin*, which resulted in increased protection against shedding from the cell surface induced by light. This strategy is poised to enable future investigations into the regulation of mucin shedding by host- and pathogen-associated lectins and provide new insights into the protective functions of the mucosal barrier.

4.2 Results

Generation of mucin mimetics with photo-cleavable membrane anchors

To model mucin glycoprotein shedding from cell surfaces, we designed mucin mimetic glycopolymers that can be presented on the plasma membrane of cells and subsequently released upon application of an external stimulus, such as light (Fig 4.1B). The membrane targeting mucin-mimetic glycopolymers comprised a poly(ethylene oxide) (PEO) backbone glycosylated to produce the mucin-mimetic glycodomain and terminated with a hydrophobic anchor linked through a photocleavable nitrobenzyl group. Additionally, small percentage of the polymer sidechains (~ 1%) were functionalized with a fluorescent reporter (Cy5) for visualization.

The glycopolymer synthesis began by generating an azide-terminated poly(epichlorohydrin) scaffold **P1** primed for copper-catalyzed azide-alkyne cycloaddition (CuAAC) with cholesterol alkyne **1** containing a previously reported photocleavable (**PCL**) nitrobenzyl linker group [17, 18] (Fig 4.2A). Monomer-activated anionic ring opening polymerization of **ECH**[15] in the presence of tetrabutylammonium azide (0.3 mol%) initiator and triisobutyl aluminum activator (0.7 mol%) furnished polymer precursor **P1** near the target molecular weight (Mw = 29 kDa, DP ~ 300) and narrow chain-length distribution (Đ = 1.23). Treatment of **P1** with **1** (10 equiv.) in the presence of a copper(I) iodide catalyst (1 equiv. per end group) and diisopropylethylamine, afforded a photocleavable cholesterol end group-modified **ECH** polymer intermediate **P2-PCL**. Introduction of the cholesterol end-group was difficult to observe directly by ^{1}H NMR spectroscopy; however, it could be confirmed by the disappearance of the IR characteristic, albeit weak, azide group absorption at ν = 2100 cm^{-1} (not shown). Following chain end functionalization, the chloromethyl side chains were primed for glycosylation by reaction with sodium azide to

generate azidomethyl side chain modified polymer **P3-PCL**. Quantitative side-chain conversion was confirmed by ^{1}H NMR and IR spectroscopy. The assembly of the desired mucin mimetic glycopolymer **GP-PCL** was accomplished through a sequential copper-click reaction with sub-stoichiometric (1 mol%) alkynyl-Cy5 to introduce the fluorescent label followed by excess propargyl lactoside (1.5 eq.) as a model glycan. We assessed fluorescent labeling efficiency of **GPs** by UV-Vis spectroscopy to be ~ 2 - 3 fluorophores per **GP**, as expected for a polymer DP = 300. IR spectroscopy confirmed full consumption of the azide side chains following glycan attachment (not shown).

We also generated two analogous mucin-mimetic glycopolymer controls (for details, see SI). The first was glycopolymer **GP-NPCL**, in which the cholesterol anchor was connected to the backbone through a non-photocleavable alkyl chain linker (introduced via 5-hexynoyl cholesteroamide **2**, Fig 4.2A) to serve as a mucin mimetic control resistant to cleavage by UV light. The second was glycopolymer **GP-Ø** lacking the cholesterol end group used to confirm hydrophobic anchor-dependent membrane incorporation of the mucin mimetics.

To characterize photocleavage of the Chol-PCL linker, the UV absorption spectrum was recorded following irradiation of **1** in chloroform (λ = 365 nm) at increasing time intervals (Fig 4.2B). The formation of a new peak at λ = 370 nm is indicative of photocleavage and this absorbance was used to determine the rate of photolysis (k = 6.4 ± 0.3 min^{-1}). Photocleavage of **1** neared completion within 0.75 min of UV exposure. We

observed a similar photocleavage rate for the intermediate **P2-PCL** (k = 5.1 ± 2.3 min^{-1}, not shown). These experiments confirmed the photolysis of the membrane anchor after UV irradiation and suggested that **GP-PCL**, in which the end group photolysis could not be detected directly in aqueous solution, should be suitable for cell-surface engineering.

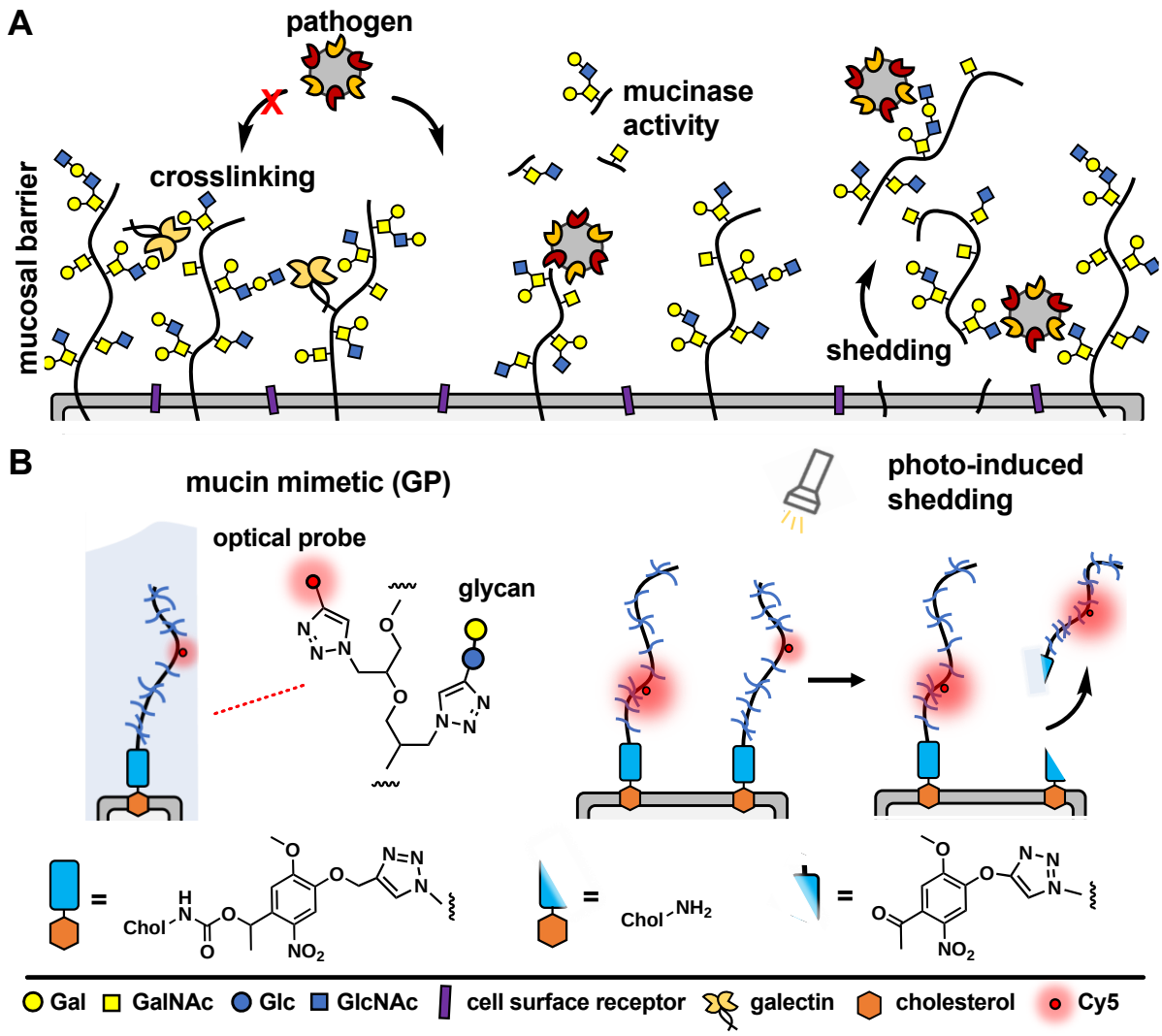


Figure 4.1 Shedding of the mucosal glycocalyx. A) Pathogens disrupt the mucosal glycocalyx barrier by inducing proteolytic mucin shedding or degradation. Crosslinking of mucins by galectins hinders access to the cell surface while crosslinking by pathogenic lectins may reduce ability of host cell to clear pathogens via shedding. B) Synthetic mucin mimetics with photo-cleavable membrane anchors for photo-patterning of the mucinous glycocalyx and the modeling of mucin shedding and mucosal crosslinking dynamics.

Photoengineering of the mucin-mimetic glycocalyx

For the construction of mucin glycocalyx models using our glycopolymer mimetics, we chose mutant Chinese hamster ovary cells, CHO *Lec8*, depleted in Golgi uridine diphosphate galactose (UDP-Gal) pools due to impaired transport of the nucleotide sugar from the cytosol.[19] As a result, these cells do not incorporate Gal into their cell surface glycans, thus providing a suitable cell system for membrane engineering with lactosylated mucin mimetics presenting b1,4-linked Gal residues. To establish optimal concentrations of the glycopolymers for cell-membrane remodeling, suspended CHO Lec8 cells were incubated with the Cy5-labeled **GP-PCL** and **GP-NPCL** at increasing polymer concentrations (0.63 – 20.00 mM, Figure 4.3A) at 4 °C. After one hour, the cells were washed to remove unincorporated polymers and analyzed by flow cytometry based on glycopolymer fluorescence. Both polymers inserted into the cell membrane with similar efficiency, nearing signal saturation at ~ 5 mM. The mucin mimetic **GP-Ø** lacking the cholesterol anchor exhibited no signal above untreated cell background, indicating the requirement for this group for membrane insertion and further confirming successful end-group functionalization of polymer precursors **P2-PCL** and **P2-NPCL**, which was difficult to detect using spectroscopic techniques (Figure 4.2A). We observed slightly greater total fluorescence intensity for cells remodeled with **GP-PCL** compared to **GP-NPCL**, consistent with the ~30% higher fluorophore labeling of the **GPs** (Figure 4.2A).

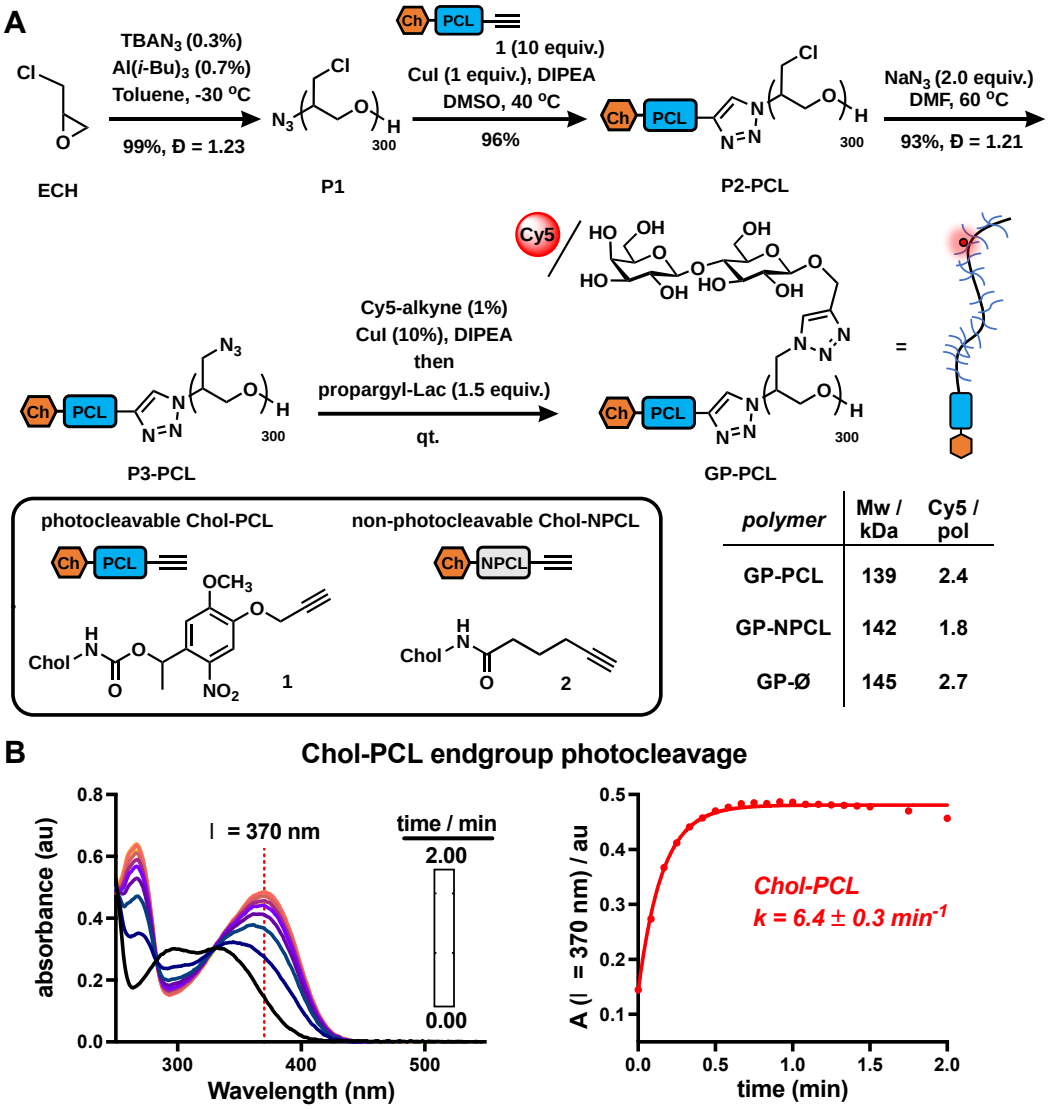


Figure 4.2 Synthesis and characterization of mucin mimetics with photocleavable membrane anchors. A) Mucin mimetic glycopolymers (**GPs**) terminated with photocleavable (PCL) and non-photo cleavable (NPCL) cholesterol anchors were elaborated from a common poly(epichlorohydrin) precursor (**P1**). A sequential end- and side-chain modification via the CuAAC reaction was used to introduce cholesterol anchors **1** and **2** and to construct a mucin mimetic domain comprised of lactosylated side chains and a fluorescent probe for visualization (Cy5, ~ 2-3 per **GP**). B) The photocleavage of cholesterol anchor Chol-PCL (**1**, 10 mg/mL in chloroform) with light at λ = 365 nm was analyzed by UV spectroscopy. The change in absorbance at λ = 370 nm over time was used to determine the rate of photocleavage (k = 6.4 ± 0.3 min^{-1}, n = 3).

Cell surface remodeling was performed at 4 °C to limit endocytosis and polymer internalization. The low temperature can decrease membrane fluidity and limit polymer insertion. Therefore, we assessed the cell-surface incorporation of **GP-PCL** (~ 5 μM) in

CHO Lec8 monolayers at 4, 21, and 37 °C. After a 1-hour incubation with the polymer, the cells were washed and analyzed via fluorescence microscopy. Remodeling at higher temperatures did not significantly improve membrane incorporation but facilitated polymer uptake by the cells, as evidenced by newly visible punctate cytosolic staining (not shown).

We next evaluated light-induced shedding of the mucin mimetics from the plasma membrane by exposing cells remodeled with both **GPs** to ultraviolet light (λ = 365 nm, Figure 4.3B). CHO Lec8 cells remodeled in suspension with **GP-PCL** and **GP-NPCL** (5 mM) on ice were exposed to UV light at increasing time intervals for up to 3 min and the loss of cell fluorescence was measured by flow cytometry (Figure 4.3B). We only observed a light- and time-dependent reduction in Cy5 intensity for **GP-PCL**, indicating clearance of the mucin mimetics from the cell surface via photo-induced cleavage of the nitrobenzyl cholesterol anchor **1**. The UV treatment of cells remodeled with the non-photocleavable polymer **GP-NPCL** resulted in no loss of Cy5 intensity, indicating resistance of the fluorophore to photobleaching under these conditions (Figure 4.3B). Within 2 minutes of UV exposure, mucin mimetic density at the cell surface was reduced by more than 70%, with minimal further photocleavage observed after additional exposure. The small fraction of UV-cleavage resistant mucin mimetics may result from polymer internalization by the cells or, possibly, through crosslinking of the excited state radical intermediate to other membrane components. The rate of **GP-PCL** photocleavage from the cell surface (k = 2.5 ± 0.6 min^{-1}) was similar to that measured for precursor **P2-PCL** in chloroform (k = 5.1 ± 2.3 min^{-1}). The UV light treatment in the presence or absence

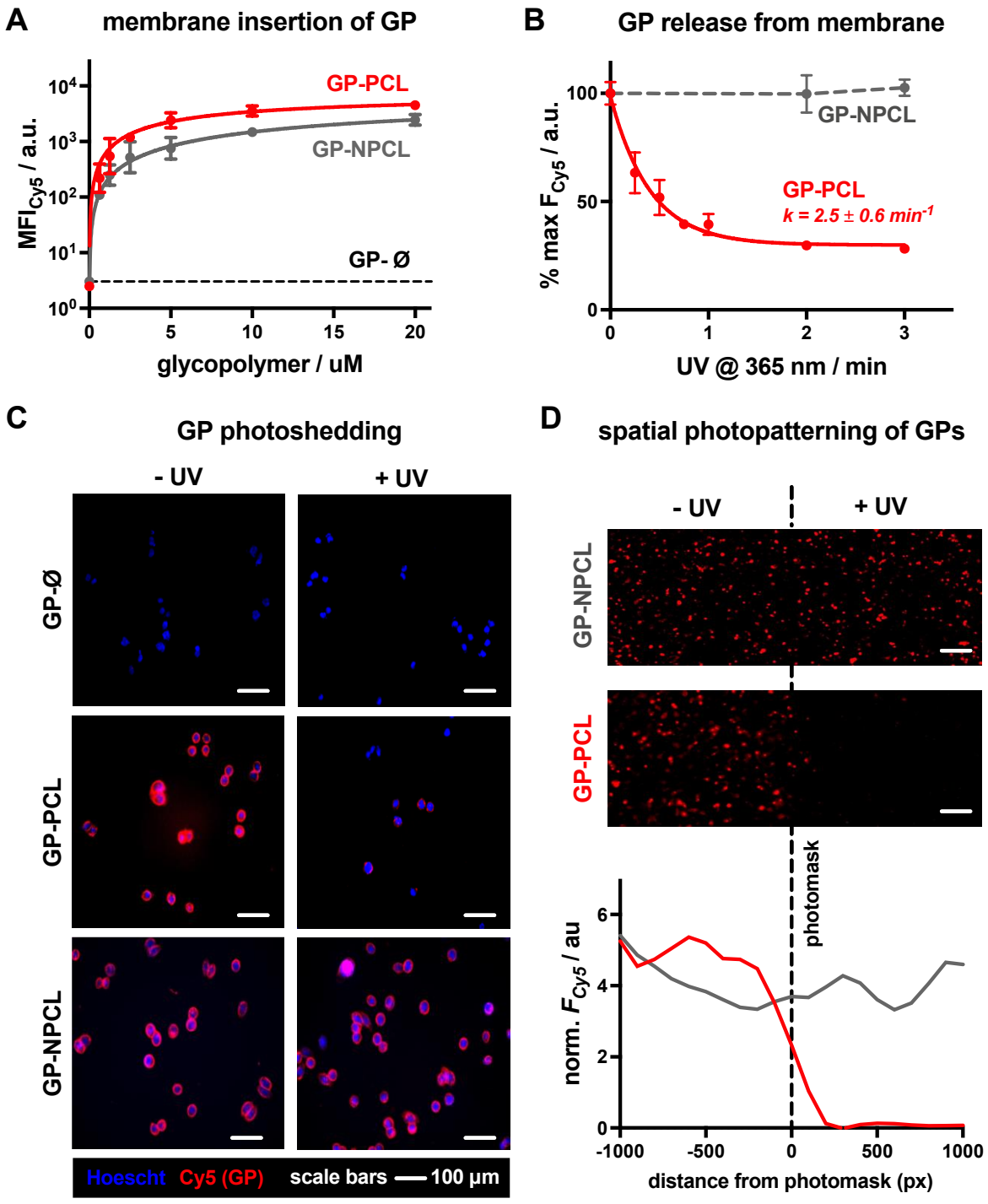
of the **GPs** resulted in little apparent cytotoxicity, as determined by a live-dead staining assay (> 93% cell viability).

Next, we assessed the light-dependent mucin mimetic shedding from CHO *Lec8* cells in adherent culture via microscopy (Figure 4.3C). The cells were incubated with all three **GPs** at sub-saturation conditions (2 mM) at 4 °C for 1 hr. Unincorporated polymers were washed and a subset of the cells was exposed to UV light (λ = 365 nm), after which all cells were washed again and treated with a nuclear stain for imaging. Mucin mimetics **GP-PCL** and **GP-NPCL**, but not **GP-Ø**, showed robust cell surface labeling in the absence of UV light, confirming cholesterol-dependent membrane remodeling (Figure 4.3C). While cells treated with the non-photocleavable mucin mimetic, **GP-NPCL**, retained their fluorescence after UV exposure, most of the **GP-PCL** signal was lost. The cytosolic punctate staining that remained visible was consistent with polymer internalization and the flow cytometry analysis (Figure 4.3B).

The light responsiveness of the mucin mimetic **GP-PCL** enables patterning of the glycocalyx withing a subpopulation of cells. To demonstrate this concept, CHO Lec8 cell monolayers remodeled with either **GP-PCL** or **GP-NPCL** (2 mM) were exposed to UV light (λ = 365 nm) on ice for 3 min in the presence of a photomask. After illumination, the cells were washed, treated with a nuclear stain and imaged (Figure 4.3D). The fluorescence micrographs show a clear drop in Cy5 signal beyond the photomask for **GP-PCL**, which was quantified by plotting fluorescence intensity averaged over 100-pixel bins

extending in both directions from the boundary. Cells remodeled with **GP-NPCL**, which is resistant to photocleavage, showed uniform fluorescence distributions. These experiments demonstrate the applicability of the light-responsive mucin mimetics for tailoring of the glycocalyx composition across a cell population with spatial resolution, which is difficult to achieve using existing glycan engineering techniques.

Figure 4.3 Photo-engineering of the mucin-mimetic glycocalyx in cells. A) Mucin mimetics containing photo-cleavable (**GP-PCL**) and non-photocleavable (**GP-NPCL**) cholesterol anchors incorporate into the plasma membranes of CHO Lec8 cells in a concentration dependent manner. Glycopolymers lacking the cholesterol anchor (**GP-Ø**) showed no association with the cells surface. B) Photo-shedding of the mucin mimetics upon irradiation with UV light (365 nm) was observed only for **GP-PCL** containing the photo-cleavable (PCL) anchor. Flow cytometry was used to determine the rate of photocleavage from the cell surface (k = 2.5 ± 0.6 min^{-1}) and the half-life for mucin mimetic shedding ($t_{1/2}$ = 0.28 ± 0.1 min, n = 3). Loss of fluorescence was not observed for the non-photocleavable mucin mimetic **GP-NPCL**. C) Fluorescence micrographs of CHO cells remodeled with Cy5-labeled **GP-PCL** and **GP-NPCL** (c_{GP} = 5 mM) before and after UV irradiation (l = 365 nm, 3 min). Cell nuclei were stained with Hoechst 33342 dye. (D) Spatial photopatterning of CHO Lec8 cells remodeled with mucin mimetics GP-PCL was accomplished through application of a mask during UV irradiation (l = 365 nm, 3 min). A plot of average fluorescence intensity per cell area with respect to the positioning of the photomask indicates mucin mimetic photo-shedding was specific to the subset of cells carrying the photocleavable **GP-PCL** and exposed to UV light (scale bars = 200 mm).


A
membrane insertion of GP
GP-PCL
GP-NPCL
GP- Ø
MFI Cy5 / a.u.
glycopolymer / uM
B
GP release from membrane
GP-NPCL
GP-PCL
k = 2.5 ± 0.6 min-1
% max F Cy5 / a.u.
UV @ 365 nm / min
C
GP photoshedding
- UV
+ UV
GP-Ø
GP-PCL
GP-NPCL
Hoescht Cy5 (GP) scale bars 100 µm
D
spatial photopatterning of GPs
- UV
+ UV
GP-NPCL
GP-PCL
photomask
norm. F Cy5 / au
distance from photomask (px)

Effects of lectin interactions on mucin-mimetic glycocalyx shedding

The capture and shedding of pathogens by cell surface mucins are important defense mechanisms by which host cells can limit entry and infection.[20,21,22] Pathogens, which often exploit lectin interactions to bind to glycoconjugates on host cells[23,24], may counteract the shedding process by stabilizing the glycocalyx. Better understanding how lectin crosslinking affects mucin shedding from cells may reveal new insights into this important aspect of mucosal barrier function.

To evaluate whether lectin crosslinking can stabilize the mucinous glycocalyx and prevent its shedding, we investigated the interactions and photoshedding of cell-surface displays of **GP-PCL** in the presence or absence of *Ricinus communis* agglutinin (RCA). This lectin, with specificity for terminal b1,4-linked galactosides, is known for its ability to crosslink glycoconjugates and induce cell agglutination.[25,26] In the absence of endogenous Gal on the surfaces of the mutant CHO Lec8 cells, RCA binding after remodeling and photocleavage could be attributed solely to the presence of the lactose-bearing **GP-PCL** mucin mimetics (not shown).

We first established an optimal concentration of RCA for use in binding assays (not shown). Accordingly, a suspension of wild type CHO Pro5 cells was incubated with biotinylated RCA (0-20 mg/mL) on ice, stained with excess AlexaFluor488-streptavidin, and analyzed by flow cytometry. We observed concentration-dependent RCA staining with maximal signal intensity and no evidence of cell aggregation at lectin concentration

of 5 mg/mL (not shown) CHO Pro5 cells in monolayer culture were then stained with RCA at this concentration and analyzed by fluorescence microscopy to confirm robust staining for imaging (not shown).

Using the optimized RCA staining conditions, we evaluated the binding of the lectin to CHO Lec8 cells remodeled with increasing concentrations of the photocleavable mucin mimetic ($c_{GP\text{-}PCL}$ = 0 – 2 mM). As expected, flow cytometry analysis revealed mucin mimetic concentration-dependent RCA binding (Figure 4A). Anticipating that RCA crosslinking may be affected by the membrane-density of the mucin mimetic, we plotted the ratio of **GP-PCL** and RCA fluorescence intensities as a function of polymer concentration (Figure 4B). We observed an increase in the polymer/RCA ratio, indicating more extensive crosslinking with increasing polymer density until saturation of lectin binding sites.

Next, we evaluated the effects of RCA crosslinking on mucin-mimetic shedding. We induced cleavage of **GP-PCL** from the cell surface either before or after crosslinked by the lectin (Figure 4.4C). CHO Lec8 cells in monolayer were remodeled with **GP-PCL** at a concentration sufficient to induce maximal RCA crosslinking (2 mM), stained with RCA (5 mg/mL), and analyzed by fluorescence microscopy (Figure 4C). We observed robust labeling of the remodeled cells by RCA prior to exposure to UV light. Pearson's correlation analysis showed strong colocalization of the RCA and **GP-PCL** signals, confirming association of the lectin with the mucin mimetic in the glycocalyx. When RCA was added to the remodeled cells following UV treatment (3 min), minimal binding was

observed in agreement with the decrease in availability of lectin binding sites after photo-induced shedding of the mucin mimetic. When RCA was added before illumination, a significant portion of **GP-PCL** remained on the cell surface (Figure 4C), which was quantified by measuring the mean Cy5 fluorescence intensity per cell area. This indicates that crosslinking by the lectin prior to shedding prevents clearance of the mucin mimetics from the cell surface, by tethering cleaved polymers to those still anchored to the cell membrane. In the native environment of the mucosal glycocalyx, oligomeric lectins can bridge mucins with other endogenous glycoconjugates present at the cell surface and further decrease the efficiency glycocalyx shedding.

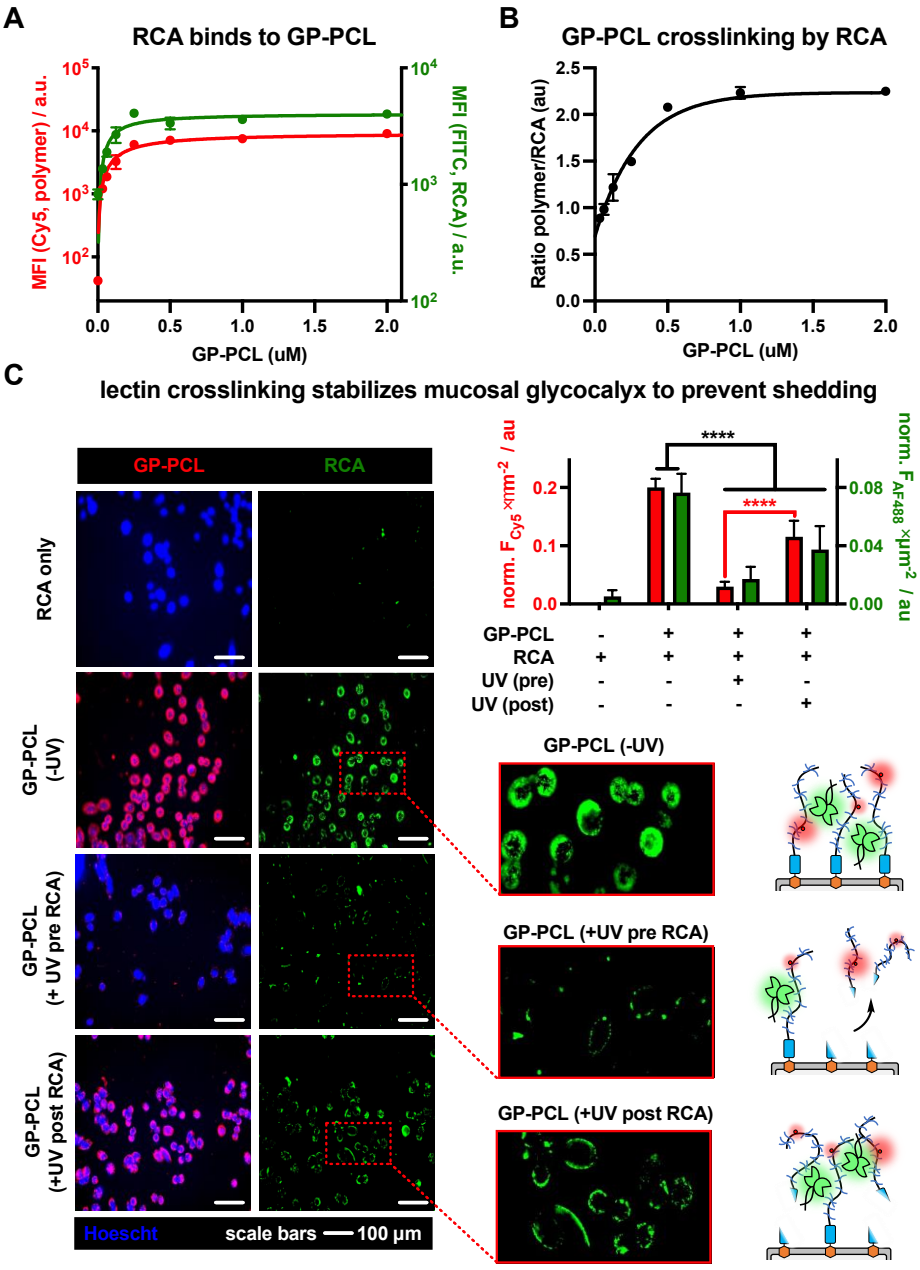


Figure 4.4 Lectin crosslinking limits photo-shedding of mucin mimetic glycocalyx. A) Remodeling of CHO Lec8 cells with mucin mimetic **GP-PCL** (red) introduces galactose binding sites for RCA (green) on the cells surface in a concentration-dependent manner. B) The plot of fluorescence intensity ratios for RCA and **GP-PCL** indicates enhanced lectin crosslinking with increasing polymer density in the membrane. C) Fluorescence micrographs and bar graph representations of CHO Lec8 cells remodeled with **GP-PCL** and irradiated either before (UV pre) or after (UV post) RCA crosslinking. Photo-shedding of the mucin mimetics prior to RCA incubation reduces the number of available binding sites for the lectin. RCA crosslinking of the mucin mimetic stabilizes the glycocalyx and limits glycopolymer photo-shedding of the polymer from the cell surface.

4.3 Conclusions

In this study, we have developed light-responsive glycomimetic materials to model the shedding behavior of mucin glycoproteins in the mucosal barrier. The introduction of a photo-cleavable cholesterol anchor into the mucin mimetics enabled their installation into plasma membranes to augment the glycocalyx of living cells and enable subsequent photo-release. We demonstrated the utility of these materials for the photo-patterning of cell surfaces and their interactions with lectins. We used RCA to model the effects of lectin binding on mucin shedding dynamics and observed that crosslinking can enhance their retention on the cell surface. This glycocalyx engineering strategy may provide new insights into the roles of mucins in regulating host-pathogen interactions and the contributions from endogenous and pathogen-associated lectins to the protective functions of the mucosal barrier.

4.4 Methods

General materials and methods

All chemicals, unless otherwise stated, were purchased from Sigma Aldrich and used as received. Cuprisorb resin was purchased from SeaChem Labs. Reaction progress was monitored by analytical thin-layer chromatography (TLC, Merck silica gel plates) with UV illumination or via CAM, ninhydrin, or $KMnO_4$ staining. Column chromatography was performed on a Biotage Isolera One automated flash chromatography system. Nuclear magnetic resonance (^{1}H and ^{13}C NMR) spectra were recorded on Bruker 300 MHz and Jeol 500 MHz NMR spectrometers. Spectra were

recorded in $CDCl_3$ or D_2O at 293K and are reported in parts per million (ppm) on the δ scale relative to 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, CD_3OD = 49.0 ppm). HRMS (high-resolution mass spectrometry) analysis was performed on an Agilent 6230 ESI-TOFMS in positive ion mode. UV-Vis spectra were collected in a quartz cuvette using a Thermo Scientific Nanodrop2000c spectrophotometer. IR spectroscopy was performed on a Nicolet 6700 FT-IR spectrophotometer. 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 TSK gel columns in series (Tosoh Bioscience) using an isocratic method with a flow rate of 0.7 mL/min in DMF (0.2% LiBr, 70 °C).

Synthesis of azide-terminated poly(epichlorohydrin), P1

Epichlorohydrin was polymerized according to published procedures[27,28]. Briefly, to a 10 mL flame-dried Schlenk flask equipped with a magnetic stirrer was added tetrabutylammonium azide ($TBAN_3$, 30.0 mg, 60.0 mmol) under Ar atmosphere. Distilled epichlorohydrin (1.29 mL, 16.50 mmol) in anhydrous toluene (4.00 mL) and triisobutylaluminum in toluene (1.07 M, 0.10 μL, 0.11 mmol) were then added at −30 °C. The reaction was stirred for 4 hours before quenching with ethanol. The resulting polymer **P1** was precipitated into hexanes and dried under vacuum to yield a clear viscous oil (1500 mg, 99%). The polymer was analyzed by SEC (0.2% LiBr in DMF): Mw = 29,000, Mn = 27,700, Đ = 1.23 and ^1H NMR ($CDCl_3$, 500 MHz).

Synthesis of poly(epichlorohydrin) polymers, P2

To separate 1-dram vials with a magnetic stirrer were added p(ECH) polymer **P1** (7.50 mg, 0.25 μmol, 1 equiv) and anhydrous DMSO (500 μL). Photocleavable cholesterol anchor **1** (1.70 mg, 2.50 μmol, 10 equiv) or non-photocleavable cholesterol anchor **2** (1.20 mg, 2.50 μmol, 10 equiv) was added, followed by CuI (0.05 mg, 0.30 μmol, 1.0 equiv) and one drop diisopropylethyl amine (~ 5 μL). The reactions were stirred at 40 °C for 12 hours before quenching with DCM and mixing with Cuprisorb beads (18 hrs) to sequester copper. The polymers were filtered through celite, concentrated under vacuum, and triturated with chloroform in EtOH (30% v/v) to remove residual **1** or **2**. The resultant polymers **P2** were dried under vacuum to yield **P2-PCL** (7.2 mg, 96%) and **P2-NPCL** (6.7mg, 89%). **P2** were characterized by ^{1}H NMR ($CDCl_3$, 500 MHz), IR spectroscopy, and UV-Vis spectroscopy (not shown).

Synthesis of cholesterol poly(glycidyl azide) polymers, P3

Three separate 1-dram vials were charged with cholesterol-terminated p(ECH) polymers **P2** (6.7 - 7.2 mg, 0.22 - 0.24 μmol, 1 equiv.) in anhydrous DMF (200 μL) and a magnetic stirrer. To the solutions was then added NaN_3 (2.0 mg, ~ 2.0 equiv.) and the reactions was stirred at 60 °C for 72h under Ar. The reaction solutions were filtered, dried, and concentrated from DCM to yield p(GA) polymers **P3-PCL** (7.0 mg, 93%), **P3-NPCL** (6.70 mg, 89%), and **P3-Ø** (7.18 mg, 88%). **P3** were characterized by ^{1}H NMR ($CDCl_3$, 500 MHz), size exclusion chromatography, and IR spectroscopy.

Synthesis of glycopolymers, GP

To three separate 1-dram vials were added p(GA) polymers **P3** (7.50 mg, 0.075 mmol) dissolved in anhydrous DMSO (0.25 mL) and magnetic stirrers. Solutions of Cy5-alkyne (7.50 mg, 0.75 μmol) in DMSO (75 μL) were added, followed by CuI (1.67 mg, 7.50 μmol) and DIPEA (13.3 uL, 0.075 mmol). After 2 hours at 40 °C under Ar, propargyl lactose[16B] (50 μL, 0.113 mmol, 1.5 eq per azide side-chain) in anhydrous DMSO was added to the reactions and stirred at 40 °C overnight. The glycopolymers were diluted in water and treated with Cuprisorb beads for 18 hours to sequester copper before filtration over celite and lyophilization. Methanol was used to remove excess glycoside and **GPs** were again lyophilized to yield the Cy5-labeled glycopolymers as pale blue solids **GP-PCL** (7.50 mg, quant.), **GP-NPCL** (7.50 mg, quant.), and **P2-Ø** (7.50 mg, quant.). **GPs** were characterized by ^{1}H NMR (D_2O, 500 MHz, FigS8) and Cy5 labeling efficiency was quantified via UV-Vis spectroscopy (λ_{max} = 633 nm, ~ 2-3 fluorophores per polymer).

General cell culture

All biological reagents were purchased from Gibco (ThermoFisher) unless otherwise stated. CHO Lec8 and CHO Pro5 cells used were obtained from ATCC (CRL-1737 and CRL-1781, respectively). Biotin-labeled *Ricinus communis agglutinin I* was purchased from Vector Labs (B-1085-5). Cells were cultured at 37 °C and 5% CO_2 following standard tissue culture practices. CHO (Chinese Hamster Ovary cells, Pro5 and Lec8) were cultured in DMEM medium supplemented with 10% FBS, 100 U/mL penicillin, and 100 U/mL streptomycin. Cells were suspended utilizing 0.25% trypsin-EDTA and passaged every 2-4 days to achieve desired confluency on tissue-culture treated lab

plastics. Live cell flow cytometry analysis was performed using a FACSCalibur or FACSCanto II system (BD Biosciences). Microscopy was performed on either a Keyence BZX800 epifluorescent microscope or a ThermoScientific EVOS imaging system and images were analyzed using ImageJ.

GP membrane incorporation

Flow cytometry: CHO Lec8 cells were suspended, washed, and 10^6 cells were pelleted into Eppendorf tubes. GPs prepared in DPBS (0 - 20 mM, 100 mL) were added to the cell pellets, mixed, and incubated on ice for 1 hour. Following two washes with DPBS, cells were resuspended and analyzed by flow cytometry. Microscopy: CHO Lec8 cells grown in 12-well plates were rinsed with DPBS and **GPs** (2 mM) were added. Plates were incubated on ice for 1 hour before three DPBS washes. Nuclei were stained using Hoescht 33342 (10 mg/mL, 10 min) followed by two additional DPBS washes. Fluorescent micrographs were captured on a Keyence epifluorescent microscope.

Photo-induced shedding of GPs from cell surface

Flow cytometry: Remodeled CHO Lec8 cells in clear plastic tubes were subjected to ultraviolet light (l = 365nm, 0 - 3 min) using a handheld 15W lamp. Following irradiation cells were washed twice with DPBS and resuspended for flow cytometry analysis. Microscopy: Remodeled CHO Lec8 cells were irradiated in well plates using a handheld 15W lamp (l = 365nm, 3 min) directly below the plate. Cells were then washed three times

with 1mL of DPBS and nuclei were stained using Hoescht 33342 (10 mg/mL, 10 min). Following two additional DPBS washes fluorescent micrographs were captured.

Photopatterning of GPs in cell membrane

CHO Lec8 cells were grown until confluent in 6-well plates. Standard **GP** membrane incorporation protocols were utilized and fluorescent micrographs were captured within a selected region suitable for stitching. A portion of the well was masked before irradiation using a handheld 15W lamp (l = 365nm, 0 - 3 min). Following mask removal, images were collected within the region, stitched using Keyence BZX Analyzer software, and quantified by ImageJ.

RCA binding to GP-PCL remodeled cells

CHO Lec8 cells were suspended (0.25% trypsin-EDTA), washed, and 10^6 cells were pelleted in Eppendorf tubes. **GP-PCL** prepared in DPBS (0-5uM, 100 mL) were added to the cell pellets and incubated on ice for 1hr. Following two DPBS washes cells were incubated in RCA-biotin (5 ug/mL, 300mL) for 40 min on ice. After two additional washes cells were incubated in excess AlexaFluor488 labeled streptavidin (300mL, 1:750) for 20 min, washed twice with DPBS, and resuspended for flow cytometry analysis.

Lectin crosslinking during GP shedding

CHO Lec8 cells grown in 12-well plates were remodeled with **GP-PCL** (2 mM) on ice for 1 hour. Cells were then washed three times with DPBS and incubated with RCA-

biotin (5 mg/mL) either before (pre RCA) or after (post RCA) UV irradiation using a 15W lamp (l = 365nm, 3 min). After an additional three DPBS washes, cells were incubated on ice for 30 min with an excess of AlexaFluor488 labeled streptavidin (300uL, 1:750) and nuclei were stained with Hoescht 33342 (10 mg/mL, 10 min) for visualization. Following two additional DPBS washes fluorescent micrographs were captured on a Keyence epifluorescent microscope and ImageJ was used to analyze micrographs.

4.6 Author Contributions

K.G. conceived of the research; S.C.P., M.H.Z., and K.G. designed the research; S.C.P., M.H.Z., H.J.C.N., D.J.H., and A.L.M. performed the research; S.C.P., M.H.Z., H.J.C.N, and K.G. analyzed data; S.C.P. and K.G. wrote the manuscript and supporting information; all authors approved the manuscript for publication.

4.7 Acknowledgments

Chapter 4 was an adaptation and reprint of the material as it appears in *RSC Chemical Science* 2022. Purcell SC, Zhang MH, Honigfort DJ, Ng HFC, Michalak AL, Godula K. Cell surface photoengineering enables modeling of glycocalyx shedding dynamics. Austen Larson Michalak was a primary author of this paper.

4.8 References

1 Varki A, Gagneux P. Biological Functions of Glycans. **2017**. In: Varki A, Cummings RD, Esko JD, et al. Essentials of Glycobiology, 3rd edition. Cold Spring Harbor (NY): *Cold Spring Harbor Laboratory Press*; 2015-2017.

2 Linden, SK, Sutton P, Karlsson NG, Korolik V, McGuckin MA. Mucins in the mucosal barrier to infection. *Mucosal Immunol.* **2008**, 1, 183–197.

3 Purcell SC, Godula K, Synthetic glycoscapes: addressing the structural and functional complexity of the glycocalyx. *Interface Focus*. **2018**, 9, 20180080.

4 McGuckin MA, Linden SK, Sutton P, Florin TH. Mucin dynamics and enteric pathogens. *Nat. Rev. Microbiol.* **2011**, 9, 265–278.

5 Honigfort DJ, Altman MO, Gagneux P, Godula K. Glycocalyx crowding with mucin mimetics strengthens binding of soluble and virus-associated lectins to host cell glycan receptors. *Proc. Nat'l. Acad. Sci.* **2021**, 118 (40).

6 Delaveris SC, Webster RE, Banik SM, Boxer SG, Bertozzi CR. Membrane-tethered mucin-like polypeptides sterically inhibit binding and slow fusion kinetics of influenza A virus. *Proc. Nat'l. Acad. Sci*. **2020**, 17 (23) 12643-12650.

7 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*. **2013** 10, 321.

8 Xie Y, Chen S, Li Q, Sheng Y, Alvarez MR, Reyes J, Xu G, Solakyildirim K, Lebrilla C. Glycan–protein cross-linking mass spectrometry reveals sialic acid-mediated protein networks on cell surfaces. *Chem. Sci*. **2021**, 12(25), 8767–8777.

9 Argueso P, Guzman-Aranguez A, Mantelli F, Cao Z, Ricciuto J, Panjwani N. Association of the cell surface mucins with galectin-3 contributes to the ocular surface epithelial barrier. *J. Biol. Chem.* **2009**, 284(34), 23037-23045.

10 Belardi B, O'Donoghue G.P, Smith AW, Groves JT, Bertozzi CR. Investigating cell surface galectin-mediated cross-linking on glycoengineered cells. *J. Am. Chem. Soc.* **2012**, 134, 23, 9549-9552.

11 Belardi B, Bertozzi CR. Chemical lectinology: tools for probing the ligands and dynamics of mammalian lectins in vivo. *J. ChemBiol*. **2015**, 22(8), 983-993.

12 Shurer CR, Colville MJ, Gupta VK, Head SE, Kai F, Lakins JN, Paszek MJ. Genetically encoded toolbox for glycocalyx engineering: tunable control of cell adhesion, survival, and cancer cell behaviors. *ACS Biomat. Sci. Eng*. **2018**, 4(2), 388–399.

13 Shon DJ, Malaker SA, Pedram K, Yang E, Krishnan V, Dorigo O, Bertozzi CR. An enzymatic toolkit for selective proteolysis, detection, and visualization of mucin-domain glycoproteins. *Proc. Nat'l. Acad. Sci*. **2020**, 117(35), 21299–21307.

14 Malaker SA, Pedram K, Ferracane MJ, Bensing BA, Krishnan V, Pett C, Yu J, Woods EC, Kramer JR, Westerlind U, Dorigo O, Bertozzi CR. The mucin-selective protease StcE enables molecular and functional analysis of human cancer-associated mucins. *Proc. Nat'l. Acad. Sci*. **2019**, 116(15), 7278–7287.

15 Kramer JR, Onoa B, Bustamante C, Bertozzi CR. Chemically tunable mucin chimeras assembled on living cells. Proceedings of the National Academy of Sciences, **2015**, 112(41), 12574 – 12579.

16 Pan H, Colville MJ, Supekar NT, Azadi P, Paszek MJ. Sequence-specific mucins for glycocalyx engineering. ACS synthetic biology, **2019**, 8(10), 2315–2326.

17 Gaur P, Kucherak OA, Ermakova YG, Shvadchak VV, Yushchenko DA. Nitrobenzyl-based fluorescent photocages for spatial and temporal control of signaling lipids in cells. *Chem Comm*. **2019**, 55(82), 12288-12291.

18 Kaneko S, Nakayama H, Yoshino Y, Fushimi D, Yamaguchi K, Horiike Y, Nakanishi J. Photocontrol of cell adhesion on amino-bearing surfaces by reversible conjugation of poly(ethylene glycol) via a photocleavable linker. *Phys. Chem. Chem. Phys*. **2011**, 13(9), 4051-4059.

19 Oelman S, Stanley P, Gerardy-Schahn R. Point mutation identified in Lec8 chinese hamster ovary glycosylation mutants that inactivate both the UDP-galactose and cmp-sialic acid transporters. *J. Biol. Chem*. **2001**, 276(28), 26291-26300.

20 Dhar P, McAuley J. The role of the cell surface mucin MUC1 as a barrier to infection and regulator of Inflammation. Frontiers Cell. and Infect. Microbiol. **2019**, 9.

21 McAuley JL, Corcilius L, Tan HX, Payne RJ, McGuckin MA, Brown LE. The cell surface mucin MUC1 limits the severity of influenza A virus infection. *Mucosal Immunol.* **2017**, 10, 1581–1593.

22 Turner, J. Intestinal mucosal barrier function in health and disease. *Nat. Rev. Immunol.* **2009**, 9, 799–809.

23 Corfield AP. Mucins: a biologically relevant glycan barrier in mucosal protection. *Biochim. Biophys. Acta.* **2015**, 1850(1), 236-52.

24 Wagner CE, Wheeler KM, Ribbeck K. Mucins and their role in shaping the functions of mucus barriers. Annu*. Rev. Cell Dev. Biol.* **2018**, 6(34), 189-215.

25 Cummings RD, Schnaar RL. R-Type lectins. **2017**. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 31.

26 Gupta D, Kaltner H, Dong X, Gabius HJ, Brewer CF. Comparative cross-linking activities of lactose-specific plant and animal lectins and a natural lactose-binding immunoglobulin G fraction from human serum with asialofetuin. *Glycobiology*. **1996**, 6(8), 843-849.

27Gervais M, Labbé A, Carlotti S, Deffieux A. Direct synthesis of α-azido,ω-hydroxypolyethers by monomer-activated anionic polymerization. *Macromolecules*, **2009**, 42, 2395–2400.

28 Honigfort DJ, Zhang MZ, Verespy S, Godula K. Engineering of spectator glycocalyx structures to evaluate molecular interactions at crowded cellular boundaries. Faraday Discussions. **2019**, 219, 138-153.

5. Lectin oligonucleotide conjugates for soluble glycan array applications

5.1 Introduction

Both commensal and pathogenic bacteria have many cell-surface glycans which are markers of disease, yet the glycome of bacteria remains poorly understood. Glycans in bacteria may differ greatly from eukaryotes, with key glycans such as lipopolysaccharide, a hallmark of pathogenic bacterial strains.[1] Given the importance of cell surface glycans in pathogenesis, they are an important biomarker for microbial health in the human gut, with great interest in profiling bacterial glycans for therapeutic, diagnostic proposes.[2] Since its inception in 2002, the glycan microarray has been a transformative tool to interrogate glycan interactions in a high throughput manner.[3,4] Glycan microarrays consist of solid substrate (usually a glass slide), which bears immobilized glycans, and a labeled protein is washed over the slide for rapid identification of “punitive” glycan interactions.[5,6]

While significant studies have been made in the profiling the interactions between cell surface glycans in fluorescently labeled bacteria to immobilized lectins on a glass slide[7], we identified the need for a solution-based method to screen interactions between lectins and whole-cell microbes. A solution-based lectin array method would enable the screening of glycan interactions without the need for cells to adhere to the slide, as well as multiplexed detection of bacterial interactions with soluble lectins as a method to profile a “lectin fingerprint” for the microbiome.

However, the caveat of transition to solution-based method is the need to encode the lectin identity in this array, which was previously spatially encoded in a solid-phase lectin microarray. Drawing on inspiration from the nature's method of storing information in nucleic acids, using a DNA encoding method for lectin identity would allow the rapid identification of soluble binders of bacterial, and be compatible with a higher throughput, multiplexed format. By attaching a unique oligonucleotide barcode to each member of a lectin library, lectin binding preferences could be detected using next generation sequencing methods, which is a highly sensitive method due the ease of amplifying and detecting DNA barcodes.

Taking examples from previously established methods of conjugating oligonucleotides to proteins[8, 9, 10], we sought to generate an oligonucleotide-lectin conjugate to be used in downstream assays using a heterobifunctional linker approach based on maleimide and succinic ester chemistries. This project was done in part of a collaboration with lab of Dr. Nathan Lewis, UCSD, and the objective of my work of was develop an efficient conjugation strategy for lectins to oligonucleotide barcodes. The resulting purified lectin-oligonucleotide conjugates would be used to interrogate microbiome samples and be applied as a screening tool to "fingerprint" microbial populations by lectin binding properties. The focus of this chapter will be the application of a method to conjugate oligonucleotides to lectins and purify them.

5.2 Results

A necessity for the construction of a solution-based array of DNA-encoded lectins was a robust and universal method to conjugate DNA to lectin proteins. The requirements of this conjugation strategy were compatibility with aqueous environments, reaction completion on the order of hours at mild pH values, and general compatibility with lectin proteins. Additionally, an affinity purification step would ensure purified lectin-DNA conjugates retain their lectin's binding properties.

Many antibody-DNA conjugation methods exploit the relatively rare thiol functional group on cysteine residues on heavy chain of antibodies, however this approach is limited to proteins with accessible cysteine residues, such as antibodies, and as cysteine is a relatively rare in proteins and thus is not suitable for universal conjugation strategy for glycan binding proteins.[11] Instead, we targeted solvent exposed lysine residues, which appear at a much higher frequency than cysteine. Using a heterobifunctional linker with reactive maleimides on one side and N-hydroxy succinimide on the other, the primary amine will preferentially react with the succinic ester moiety at pH values near 8.0, reserving the thiol-reactive maleimide handle for conjugation with commercially available thiol-modified oligonucleotides.[12] Subsequently, a fluorophore with a reactive succinic ester was conjugated to surface exposed lysine residues on lectins for ease of labeling and detection during preparation and experiments. During conjugation, to ensure the carbohydrate binding pocket remains accessible to glycans and free from modification, the glycan specific to the lectin was included at 200 mM. At pH 8.0 and in the presence

of 200 mM glcNAc, the lectin wheat germ agglutinin (WGA) was labeled with maleimide-PEG_2-NHS linker (**1**) for 30 mins, and then AF-647-NHS (**2**) was added and the protein reaction was incubated for an additional 30 minutes, and then immediately subjected to size exclusion purification (Figure 5.1A and 5.1B). The product labeled WGA was characterized by mass spectrometry, and UV-Vis spectroscopy and was determined to have an average 2.1 maleimides/WGA and 2.9 AF647 fluorophores per WGA (Figure 5.2C).

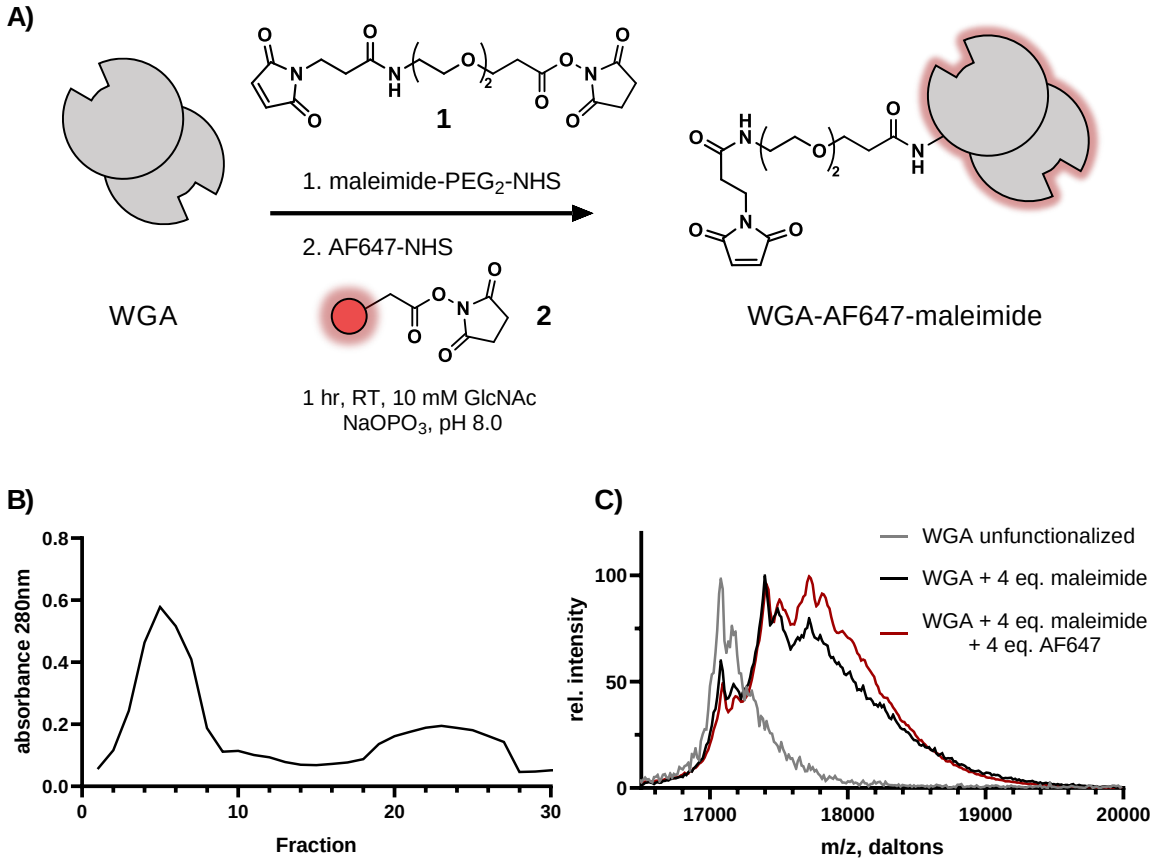
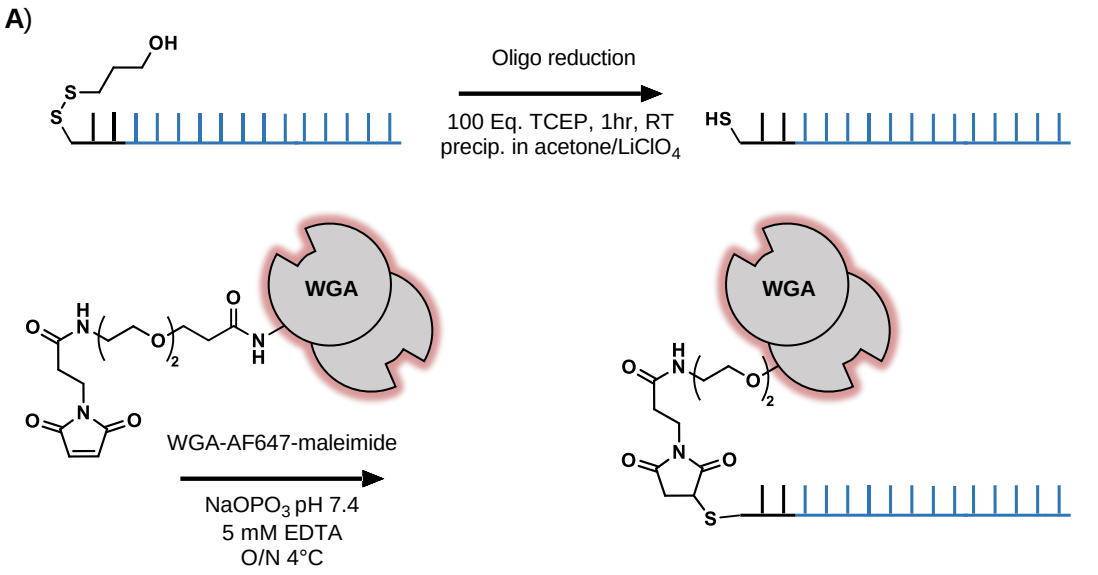
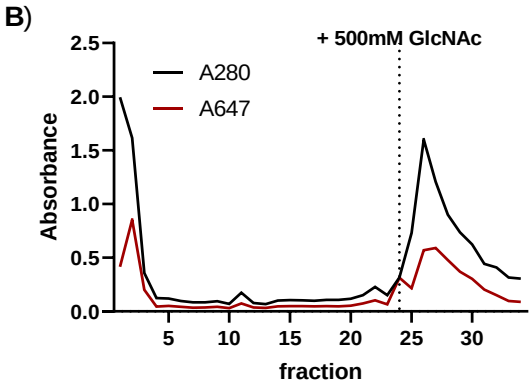


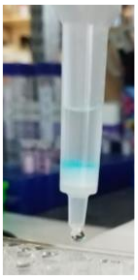
Figure 5.1 Functionalization of WGA with maleimide and fluorophore tag. A) The model lectin WGA was sequentially labeled with 1 and 2 for a fluorescent maleimide bearing conjugate. B) The reaction was purified by size exclusion chromatography and unreacted linker was separated from protein and measured by absorbance at 280nm. C) The product and intermediates were measured by matrix assisted laser desorption ionization mass spectrometry and finished protein.

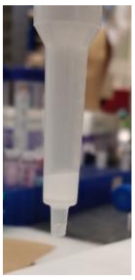
Following purification of maleimide functionalized WGA, an oligonucleotide bearing a 3' was deprotected using tris-carboxyethyl phosphine, precipitated[13] and conjugated to WGA-AF647-maleimide (Figure 5.2A). To purify unreacted oligonucleotide as well as ensure resulting lectin-oligonucleotide conjugate retained binding properties, the crude conjugation reaction was purified on a GlcNAc agarose resin (Figure 5.2B,C,D). The predominant product was WGA conjugated to a single oligonucleotide (Figure 5.2E), while a large portion of the lectin remained unconjugated to oligonucleotide (Figure 5.2F). Due the high sensitivity of DNA detection methods only necessitates a fraction of lectin being labeled with oligonucleotide for efficient detection.

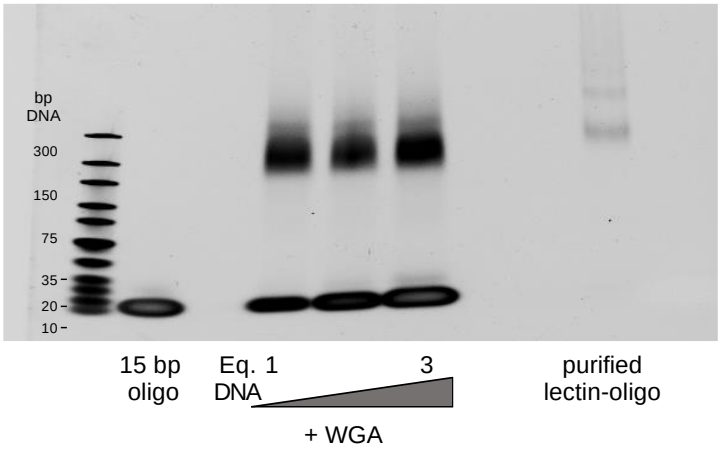
Figure 5.2 Generation of oligonucleotide-lectin conjugates. A) A 15 bp oligonucleotide barcode was deprotected and conjugated to WGA-AF647-maleimide. B) Crude conjugate was then subjected to affinity purification using an GlcNAc resin, and after washing was eluted with the addition of 500 mM GlcNAc. C) The affinity resin bed contains WGA-647-mal. after washing with 30 column volumes, and D) the affinity resin after addition of 500 mM GlcNAc. Using native PAGE, E) crude oligo, crude and purified conjugates were assessed using SYBR, and F) using Coomassie protein stain.

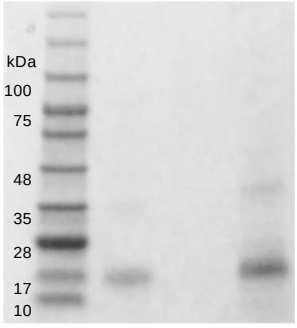

A)
OH
S
S
Oligo reduction
100 Eq. TCEP, 1hr, RT
precip. in acetone/LiClO4
HS
WGA
WGA-AF647-maleimide
NaOPO3 pH 7.4
5 mM EDTA
O/N 4°C
WGA


B)
+ 500mM GlcNAc
A280
A647
Absorbance
2.5
2.0
1.5
1.0
0.5
0.0
5
10
15
20
25
30
fraction


C)


D)


E)
bp
DNA
300
150
75
35
20
10
15 bp
oligo
Eq. 1
DNA
3
+ WGA
purified
lectin-oligo


F)
kDa
100
75
48
35
28
17
10
WGA
WGA +
15 bp
oligo

5.3 Conclusion

The preceding method for efficient conjugation of lectin to oligonucleotides and is compatible with any lectin bearing a solvent exposed lysine residue which is not buried within the carbohydrate binding domain of the lectin. This strategy will pave the wave for future lectin array studies by the Lewis lab at UCSD, which will assess the glycans of microbial communities using lectin-oligonucleotide conjugates.

5.4 Acknowledgments

Chapter 5 was original unpublished work for conjugation of oligonucleotide barcodes to lectins which served as at the basis for future macromolecular conjugation approaches and was conducted by Austen Larson Michalak.

5.5 References

1 Caroff M, Karibian D. Structure of bacterial lipopolysaccharides. *Carbohydrate Research*, **2003**, 338:2431-2447

2 Tra VN, Dube DH. Glycans in pathogenic bacteria – potential for targeted covalent therapeutics and imaging agents. *Chem Comm.* **2014**, 50, 4659-4673

3 Park S, Gildersleeve J, Blixt O. Carbohydrate Microarrays. *RSC Chem. Soc. Rev.*, **2013** 42, 4310-4326

4 Blixt O, Head S, Mondala T, Scanlan C, Huflejt M, Alvarez R, Bryan MC, Fazio F, Calarese D, Stevens J, Razi N, Stevens DJ, Skehel JJ, Die IV, Burton DR, Wilson IA, Cummings R, Bovin N, Wong C, Paulson JC. Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *PNAS* **2004**, 101,49:17033-17038.

5 Liang CH, Wu CY. Glycan array: a powerful tool for glycomics studies. *Expert rev. of proteomics* **2009**, 6(6), 631-645

6 Liang PH, Wang SK, Wong, CH. Quantitative Analysis of Carbohydrate-Proteins Interactions Using Glycan Microarrays: Determination of Surface and Solution Dissociation Constants. *J. Am. Chem. Soc.* **2007** 129,36,11177-11184

7 Hsu K, Mahal LK. A lectin microarray approach for the rapid analysis of bacterial glycans. *Nat. Protocols.* **2006**, 1, 543-549

8 Li G, Moellering RE. A Concise, Modular Antibody-Oligonucleotide Conjugation Strategy Based on Disuccinimidyl Ester Activation Chemistry. *Chembiochem* **2019**, 20, 1599-1605.

9 Kozlov I, Melnyk PC, Stromsborg KE, Chee MS, Barker DL, Zhao C. Efficient strategies for the conjugation of oligonucleotides to antibodies enabling highly sensitive protein detection. *Biopolymers*, **2004**, 73, 621-630

10 Dugal-Tessier J, Thirumalairajan S, Jain N. Antibody-Oligonucleotide Conjugates: A Twist to Antibody-Drug Conjugates. *J. of Clin. Med.* **2021**, 10:838

11 Spicer, C, Davis, B. Selective chemical protein modification. *Nat. Comm.*, **2014**, 5,4740

12 Koniev O, Wagner A. Developments and recent advancements in the field of endogenous amino acid selective bond forming reactions for bioconjugation. *Chem. Soc. Rev.* **2015**, 44:5495-5551

13 Dovgan I, Koniev O, Kolodych S, Wagner A. Antibody-Oligonucleotide Conjugates as Therapeutic, Imaging, and detection agents. *Bioconj. Chem.*, **2009**, 30, 2483-2501

6. DNA-encoded mucin mimetic platform for whole-cell analysis of bacterial glycan binding protein interactions

6.1 Introduction

The mucosal barrier coating the wet epithelial surfaces of various tissues serves a range of vital biological functions, including physical protection, exchange of nutrients and biochemical cues, and mediation of immune responses. The mucosal barrier also provides an environment, which helps establish and maintain microbial communities and facilitates their interactions with host tissues.[1,2] Microbiome-host interactions, in the form of metabolic foraging, signaling, and recognition events, have been increasingly associated with inflammation, disease progression, and overall human health.[3,4] The most abundant macromolecule in mucosal barriers are mucins, heavily glycosylated proteins which serve to protect and lubricate epithelial tissues.[5,6] Truncated mucin glycoforms, and altered microbial interactions with host mucin glycans, such as an increase in bacterial foraging, are associated with inflammatory disease states, irritable bowel disease, ulcerative colitis, and pulmonary disease states.[7,8,9] In a healthy state, inner layers of mucus layers are germ-free, while outer layers of mucus are partially digested by bacterial glycosidases which contribute to mucin turnover and shape host mucin composition.[10,11] Many diseases associated with advanced age correlate with changes in mucin composition, glycosylation, and the resulting glycan interactions and microbiome composition, which is unique to each individual.[12,13] As such, there is value in understanding how bacteria interact with mucins to provide insights into health of host and microbiome.

There is a recent focus on the need to understand glycan interactions between microorganisms and epithelial glycoproteins, yet these interactions remain resistant to systematic interrogation.[14,15] Traditionally, glycan microarrays have been powerful tools for interrogating glycan interactions using immobilized glycans on a glass slide.[16] On the surface of the glass slide, hundreds or thousands of glycans are immobilized and identified by their spatial position on the slide, or "zip code".[17] While glycan microarrays of this variety have been transformative tools for studying purified proteins or virus, the platform's use with whole-cell bacteria has been limited to several early examples and has not been widely adopted.[18,19,20,21,22] More often, recombinant bacterial adhesion proteins are screened in lieu of whole-cells.[23,24]

Recently, there has been increasing interest in the design and implementation of soluble glycan screening using DNA to encode the identity of the glycan. Using DNA encoding in soluble platforms for glycan analysis offers advantages in sensitivity, and DNA enabled technologies are rapidly developing with DNA sequencing methods becoming widely adopted and standardized. Furthermore, the encoding of glycan identity in DNA tags abolishes the requirement of a spatially delineated two-dimensional grid to encode glycan identity. Several groups have included DNA-encoded glycan arrays by covalently tethering glycans to chemically modified DNA[25,26], or by modifying the surface of bacteriophages with hundreds of chemically modified glycans.[27] These methods have enabled glycan screening using DNA sequencing as a readout method, abolishing the need for a solid support to encode glycan identity using a spatial grid.

Mucin mimetics with tunable length, valency, and glycan composition have been invaluable tools to systemically probe glycan interactions.[28] The combination of DNA encoding glycan identity with mucin mimetics to interrogate glycan interactions would provide a useful tool for interrogating glycan interactions with whole-cell bacteria, while abolishing the need for bacterial adhesion to a solid support. The resulting soluble screening platform will be ideal for understanding the glycan binding preferences of whole cells and provide new insight into interactions between bacteria and mucosal barriers.

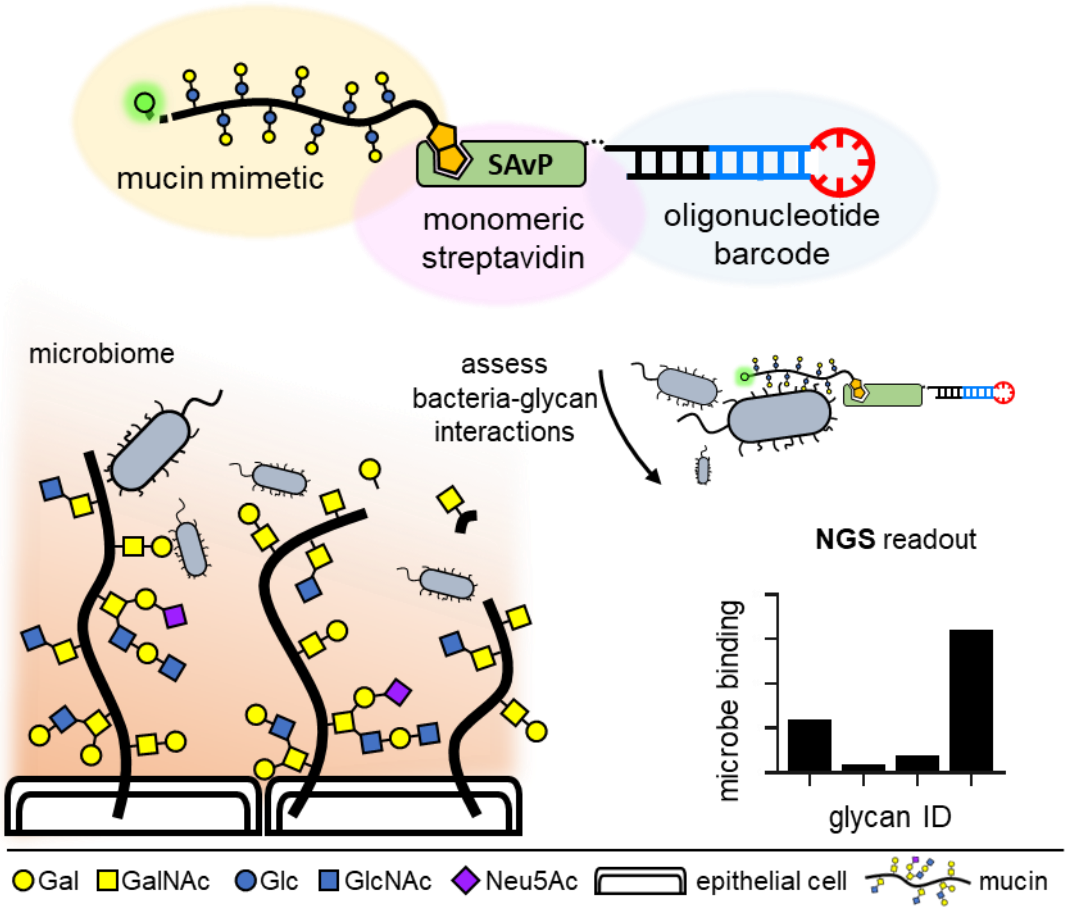


Figure 6.1 Microbial interactions with mucosa are probed with soluble mucin mimetics. The mucosal barrier is a dynamic environment rich with microbes and utilized for microbial adhesion, foraging, which influences mucin turnover and host health. Soluble DNA-encoded mucin mimetics can interrogate glycan interactions using mucin mimetics in conjugation with nucleic acid-based detection methods.

6.2 Results

To generate new technology to assess interactions between whole-cell bacteria and mucosal barriers, we envisioned a platform combining a mucin mimetic domain with DNA component to encode glycan identity. The ideal platform would combine DNA and mucin mimetic domains in a modular fashion to accommodate flexibility in glycan structure and application. For our mucin mimetic domain, we employed a reversible

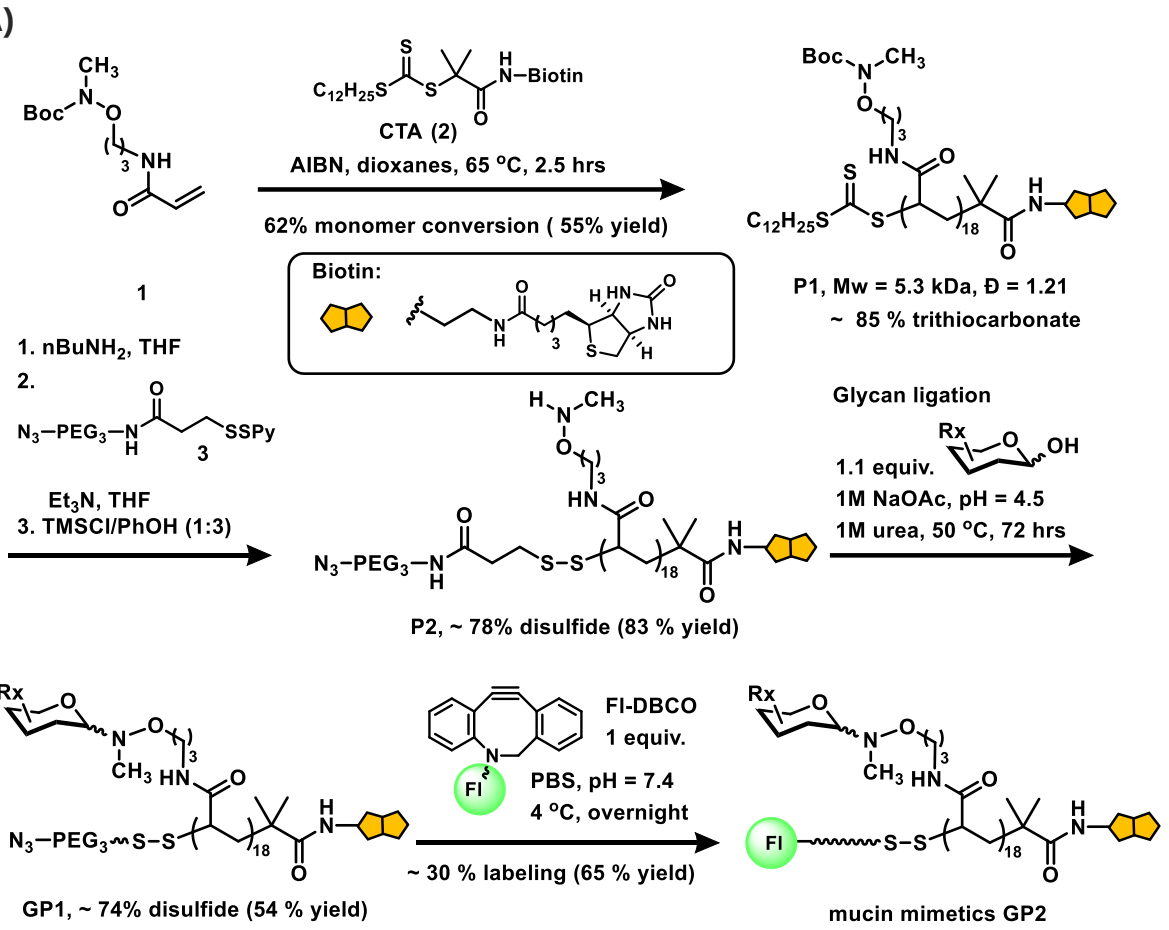
addition fragmentation transfer (RAFT) polymerization technique which affords polymers with control over length and heterobifunctional end group modification. The aminooxy bearing RAFT monomer reacts with chemically-unmodified reducing sugars, affording control over valency and glycan identity. For the DNA encoding element, we directly linked DNA oligonucleotides to a N-terminus of SavPhire monomeric streptavidin variant (SAvP, Millipore), which would be complexed with biotin end groups on mucin mimetic polymers. The resulting DNA, protein, mucin mimetic conjugate would provide the foundation of a solution-based platform to interrogate bacterial glycan interactions. To provide an alternative method to measure binding interactions, and track mucin mimetics throughout synthesis, a fluorophore was introduced to polymer endgroups through a cleavable disulfide bond. Exploiting the thiol endgroup on nascent glycopolymers, we utilized disulfide exchange to introduce an azide for strain promoted click ligation of fluorophores. The fluorophore can be removed by addition of a reducing agent, such as 2-mercaptoethanol, if spectral bandwidth is required for downstream applications or cellular labeling.

To install a biotin end-functionalization into the polymer, a biotin functionalized trithiocarbonate chain transfer agent in conjugation with aminooxy acrylamide monomer to synthesize polymers of controlled dispersity (Figure 6.2B, **P1** Đ = 1.21). [29] Subsequent deprotection of trithiocarbonate end groups yielded a free thiol which was converted using an activated pyridyl disulfide linker to yield polymer **P2** with a cleavable azide for downstream fluorophore ligation. Following side chain deprotection, amino oxy groups of polymer **P2** were functionalized with glycans 6’ sialyllactose (6SL), lactose (Lac), glucose

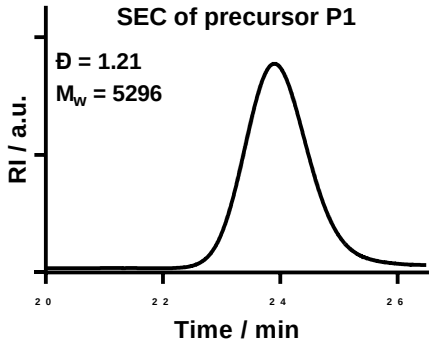
(Glc), and mannose (Man) to generate mucin mimetic library **GP1**. To give each polymer a unique fluorescent signal, the azide was modified by Cu-free click chemistry with AZdye 350, 488, 555, and 647, respectively to give each polymer a unique fluorescent signal, generating mucin mimetic library **GP2** (Figure 6.2D).

Figure 6.2 Synthesis and characterization of mucin mimetic glycopolymers. Synthetic scheme for production of focused glycopolymer library. A) Using an established acrylamide monomer (1) and a biotin functionalized chain transfer agent (2), polymerization was carried out to create DP18 polymer, **P1**. Following trithiocarbonate deprotection, disulfide exchange was carried out to yield **P2** bearing a disulfide linked azide, and then glycosylated under acidic conditions at 50 °C for 72 hrs. Using Cu-free chemistry, fluorophores are ligated to azide groups, generating mucin mimetic library **GP2**. B) Size exclusion analysis of polymer **P1** resulted in Đ = 1.21, showing controlled dispersity of polymerization. C) Table of glycan structures incorporated into mucin mimetics with relevant linkage data. D) Table of glycopolymers, **GP2** library, ligation efficiencies and fluorophore identification tags.

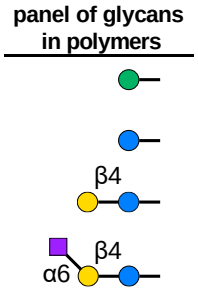
A)



B)



C)



D)

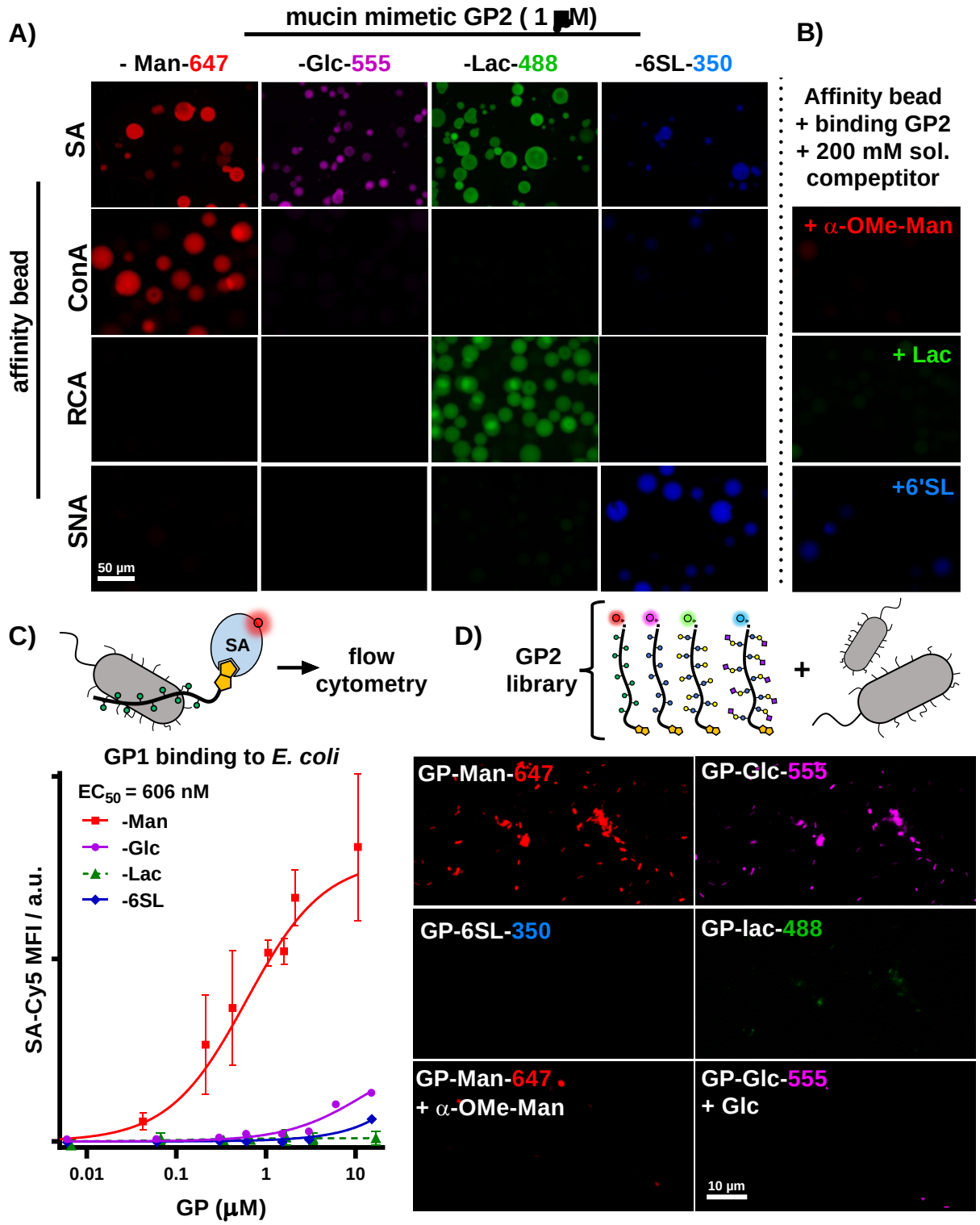
GP2-	Fl ligation %	Valency	AZDye
-man-647	20	5	647
-glc-555	38	12	555
-lac-488	27	11	488
-6SL-350	35	8	350

To validate soluble mucin mimetics, binding studies were conducted using mucin mimetic library GP2 with agarose beads functionalized with either streptavidin or lectins. The streptavidin beads bound mucin mimetics regardless of glycan, while lectins Ricinus Communis Agglutinin (RCA), Sambucus Nigra Lectin (SNA), and Concavilin A (ConA), and observed binding of polymers bearing for lactose (**GP2-lac-488**), 6' sialyllactose (**GP2-6SL-350**), and mannose polymers (**GP2-man-647**), respectively (Figure 6.3A). The inclusion of soluble competing sugars (200 mM) reduced the amount of bound mucin mimetic (Figure 6.3B).

A well-studied glycan-microbe interaction to model our system occurs between urinary tract *E. coli* and glycan binding protein FimH, which lies at the distal end of bacterial fimbriae to bind paucimannose glycans[30], a dominant glycoconjugate within urinary tract epithelia.[31] Exploiting microbiome glycan interactions has enabled targeted therapies, such as ingestion of soluble mannose to clear *E. coli* from urinary tracts by competing for FimH interactions.[32] To evaluate the dose-response of our soluble microarray, wild type MG1655 *E. coli* were probed with a mucin mimetic library **GP1** with a flow cytometry readout to construct a dose panel of *E. coli* and the binding EC_{50} of mannose bearing glycopolymer **GP1-Man** was determined to be the 606 ± 207 nM (Figure 6.3C). To show the ability of the glycopolymer library to profile glycan binding patterns on live, whole-cell bacteria, MG1655 *E. coli* were incubated with mucin mimetic library **GP2** in a multiplexed binding assay. *E. Coli* showed prominent binding when incubated with mannose (**GP2-Man-647)** and glucose (**GP2-Glc-555**) polymers, and a dramatic

reduction in binding intensity was observed with inclusion of 200 mM soluble methyl-mannose and glucose to each polymer incubation, respectively (Figure 6.3D).

Figure 6.3 Fluorescent polymer array for probing glycan interactions A) Mucin mimetic library (**GP2**) was bound to streptavidin beads, as well as lectin beads ConA, SNA, and RCA at 1 µM, 37 °C, 1 hour. B) Reducing binding was shown with the inclusion of addition of 200 mM competing sugar for lectins with binding mimetics. C) Dose curve for library **GP1**, EC_{50} of GP-1 was 606 ± 207 nM. D) Mucin mimetic library **GP2** was incubated with MG1655 WT *E. coli* for 1 hour at 37 °C, 20 µM. A decrease in binding was observed with the inclusion of 200 mM competing soluble glycan for binding mimetics. Images are representative of ≥ 3 images per condition. D) Glycopolymer library was imaged after binding to *E. coli* at 20 µM for one hour at 37 °C with corresponding soluble glycans for GP2-Man-647 and GP2-Glc-555.


A)
mucin mimetic GP2 (1 μM)
- Man-647
-Glc-555
-Lac-488
-6SL-350
affinity bead
SA
ConA
RCA
SNA
50 μm
B)
Affinity bead
+ binding GP2
+ 200 mM sol.
competitor
+ α-OMe-Man
+ Lac
+6'SL
C)
SA
flow
cytometry
D)
GP2
library
+
GP1 binding to E. coli
EC50 = 606 nM
-Man
-Glc
-Lac
-6SL
SA-Cy5 MFI / a.u.
0.01
0.1
1
10
GP (μM)
GP-Man-647
GP-Glc-555
GP-6SL-350
GP-lac-488
GP-Man-647
+ α-OMe-Man
GP-Glc-555
+ Glc
10 μm

Glycan microarrays are capable of interrogating hundreds or thousands of glycan interactions simultaneously in a single slide. However, using fluorophore detection in the solution-based microarray limits the number of fluorophore-labeled glycans which can be examined due to limited spectral bandwidth of fluorophore excitation and emission wavelengths. Inspired by the generation antibody-DNA conjugates, we designed a modular system consisting of an oligonucleotide modified monomeric streptavidin protein core, where each glycan identity is encoded by an oligonucleotide glycan barcode, or "glycode" (Figure 6.4A). The protein core provides stability of conjugate and steric spacing and the monomeric streptavidin variant, SAvP, would prevent higher order conjugates in the multivalent streptavidin. First, site specific introduction of a reactive maleimide group on monomeric streptavidin was conducted using an N-terminal specific 2-pyridylcarboxaldehyde (2-PCA) linker developed by the Francis group (Figure 6.4B).[33] This N-terminal labeling strategy was used in conjunction with reactive maleimide for addition of reactive maleimide which is compatible with commercially available thiol modified oligonucleotides. This N-terminal specific DNA modification provided robust oligonucleotide conjugation on SAvP (Figure 6.4C). Using biotin handle on the end group of mucin mimetics **GP2**, each glycopolymer was ligated to monomeric streptavidin functionalized with oligonucleotide barcode to afford a conjugate of glycopolymer, protein, and oligonucleotide for use in probing microbial glycan interactions (Figure 6.4D).

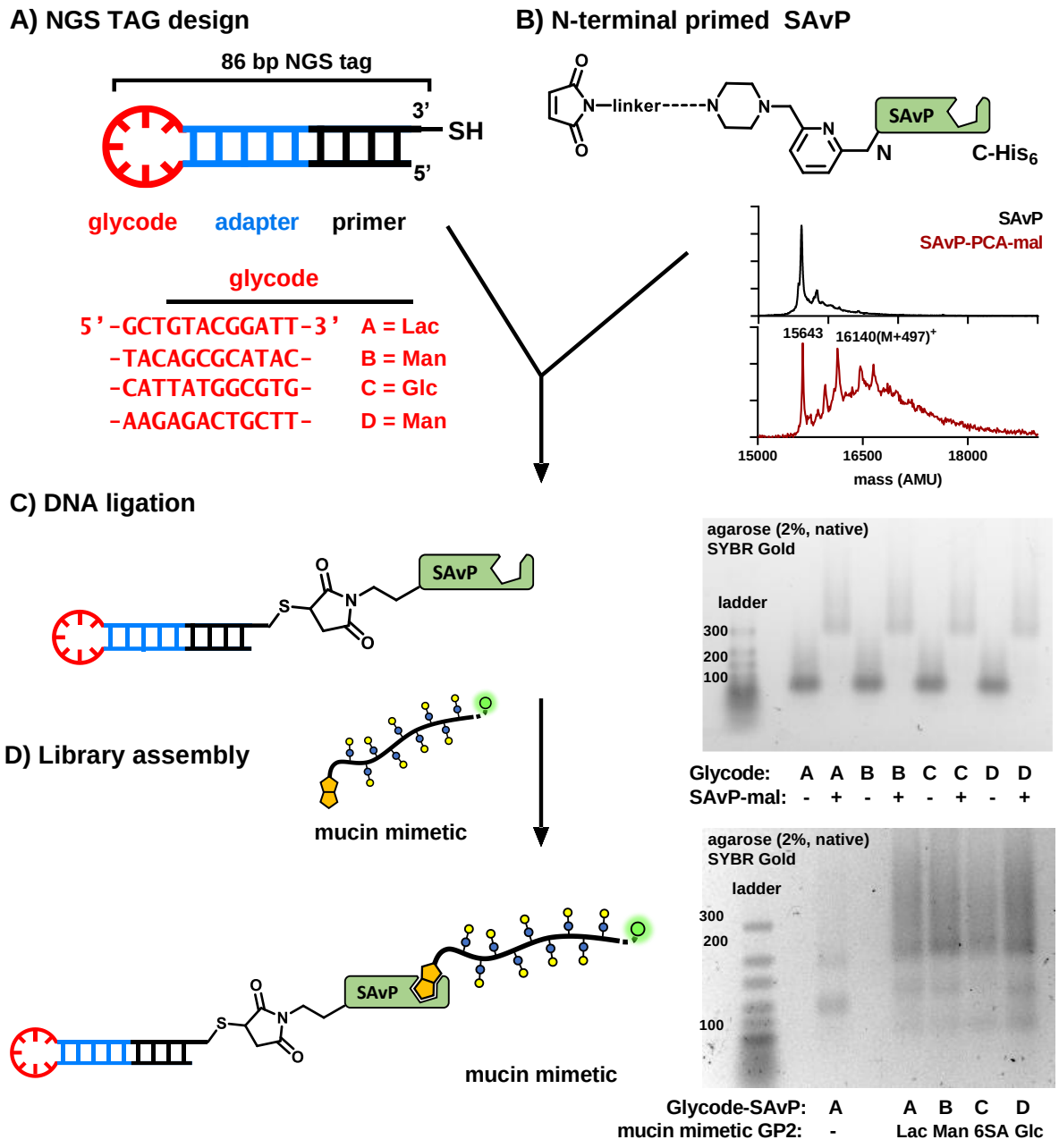


Figure 6.4 Construction of end-functionalized SAvP with and DNA barcode ligation. A) An 86 bp oligonucleotide barcode was designed with encoding regions for glycan, and sequencing primers, and a 3' exposed thiol for facile conjugation B) The N-terminal specific linker, 2-PCA, was functionalized with maleimide and then added to N-terminus of SAvP protein (37 °C, O/N). and attached to the maleimide on the end group of SAvP, and assessed with MALDI mass spectrometry C) DNA was attached to the N-terminus of SAvP, was assessed via native agarose gel to show reduced electrophoretic mobility for protein/DNA conjugates. D) Mucin mimetics were added to DNA protein conjugates via the biotin end group, which show reduced electrophoretic mobility compared to SAvP-DNA conjugates.

Next, the ability of oligonucleotide glycopolymer conjugates to report on whole-cell interactions was assessed using the model gut microorganism, *E*. *coli*. To show the necessity of the fimbrial mannose binding protein, FimH, we included the bacterial strain ORN172, which is a *FimH*$^{-/-}$ mutant of *E. coli*, and thus lacks the affinity for mannose residues. Initially, binding to both bacterial strains was assessed using RT-qPCR assay to determine relative amounts of bound glycopolymer oligonucleotide, and the MG1655 wild type *E. coli* strain showed much greater bound DNA for **GP2-Man-647**, as shown via a lower Cq value. The *FimH*$^{-/-}$ variant ORN172, however, did not show and enrichment for mannose-linked glycode (Figure 6.5A).

We then adopted NGS as our analysis method for a highly multiplexed assay design suited for rapid analysis of larger libraries, which is a hallmark of solid-state microarray glycan analysis. The binding of each *E. coli* strain was assessed against our library at equal concentrations in one pot, and then the bacteria were washed, and an NGS library was amplified from the bound mucin mimetic-oligonucleotide conjugates. Sequencing results showed a robust increase in glycode for mannose, but only in bacteria with functional fimbrial proteins (Figure 6.5B), confirming the compatibility of NGS readout with our platform.

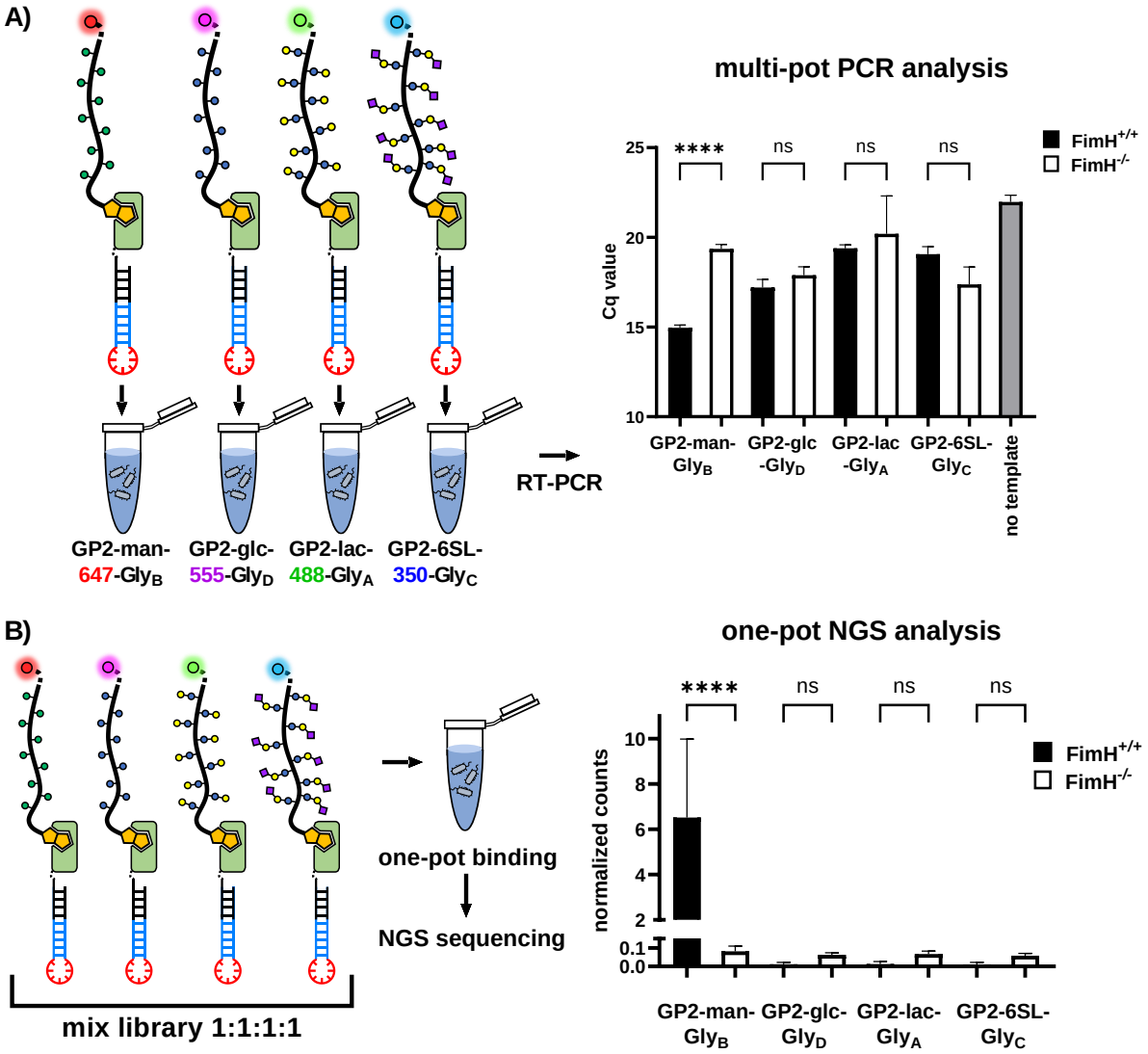


Figure 6.5 DNA based screening of glycan interactions. A) MG1655 and ORN172 *E. coli* were screened with glycode conjugates using a RT-qPCR assay to monitor binding in replication cycles. The average Cq value from 4 replicate qPCR curves plotted and only showed significant difference for mannose between ORN172 and MG1655 strains. B) The sample assay was performed in a one-pot multiplexed manner using NGS to calculate the total number of counts for each barcode from the assay.

6.3 Conclusion

The integration of mucin mimetics with oligonucleotide barcodes provides a new way to systemically inspect glycan interactions within microbial communities at the whole-cell level. As DNA based microarray technologies become more prevalent[25,26,27], new methodologies to integrate DNA based reporting methods into sophisticated glycan mimetic scaffolds will be of great use. While this technology offers many advantages over fluorescence, including sensitivity, endless multiplexing capacity, further development will be required to integrate these advantages into routine glycan binding screens for whole-cell bacteria.

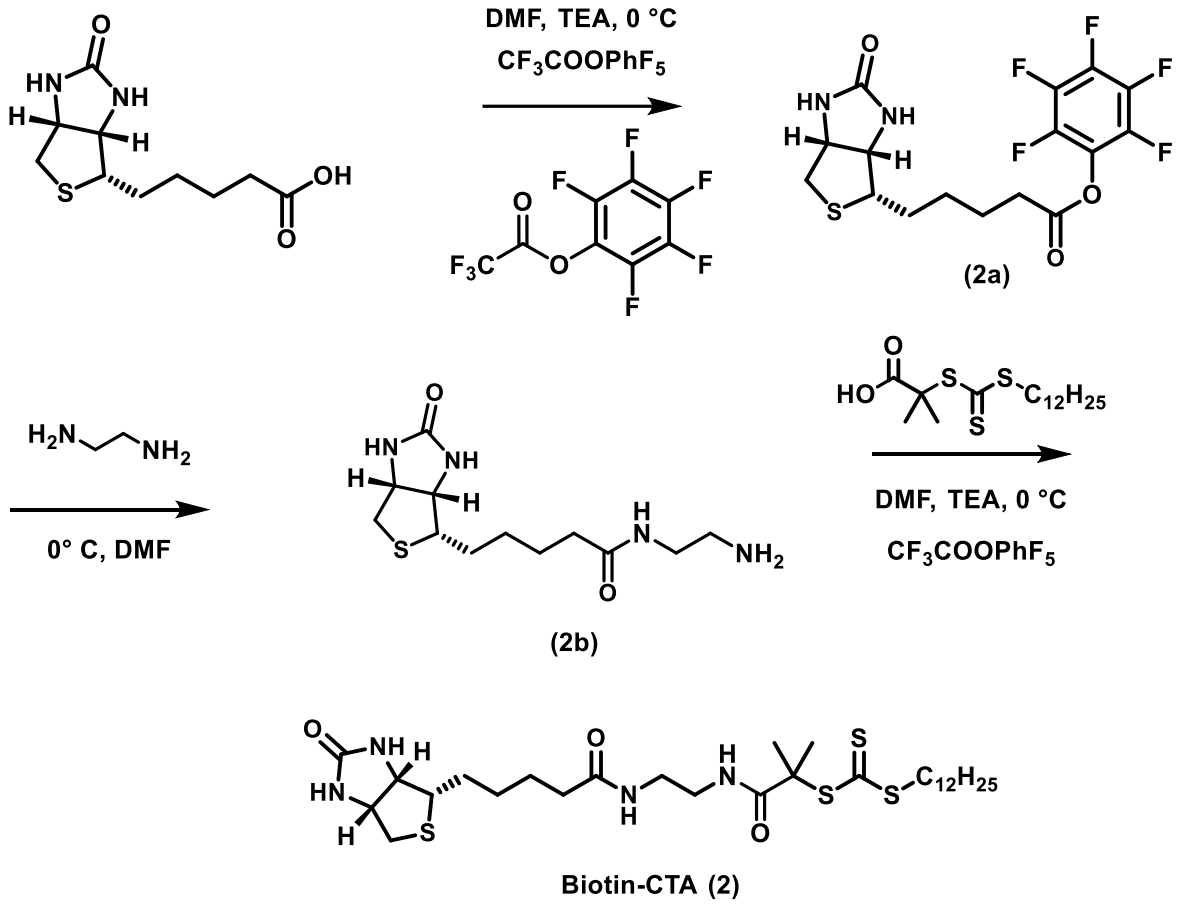
6.4 Materials and Methods

Materials and General Procedures

All chemicals, unless stated otherwise, were purchased from Sigma Aldrich. Purchased starting materials were used as received unless otherwise noted. Glycans were purchased from Iduron (Manchester, UK), V-labs (Covington, LA) or Carbosynth (San Diego, CA). Anhydrous dioxane was generated via filtration through basic alumina. Polymers were isolated by gel filtration on Sephadex G-25 columns (PD-10, GE Healthcare). Lectin beads were purchased from vector labs (CA, USA). 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 or ninhydrin.

Instrumentation

Column chromatography was performed on a Biotage Isolaera One automated flash chromatography system. Nuclear magnetic resonance (NMR) spectra were collected on either a Bruker 300 MHz or Jeol 500 MHz NMR spectrometer. Spectra was recorded in $CDCl_3$ or D_2O solutions at 293 °K 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). Data are reported as follows: chemical shift, multiplicty (s = singlet, d = doublet, dd = doublet of doublets, t = triplet, q = quartet, br = broad, m = multiplet), coupling constants (Hz), and integration. Size exclusion chromatography was performed on a Hitachi Chromaster system equipped with an RI detector and a 5 µm, mixed bed, 7.8 mm I.D. x 30 cm TSKgel column (Tosoh Bioscience). Polymers were analyzed using an isocratic method with a flow rate of 0.7 mL/min in DMF (0.2 % LiBr, 70 °C). UV-Vis spectra were collected using a Thermo Scientific Nanodrop2000c spectrophotometer. Glycan ligation reactions were conducted in a BioRad MyCycler thermocycler (Hercules, CA).



Scheme 6.1 Synthetic scheme for biotin CTA (2).

Biotin pentafluorophenyl ester (2a)

A flame-dried Schlenk flask (100mL) equipped with a magnetic stirring bar was charged with biotin (4.05 g, 16.61 mmol) under a stream of nitrogen. Dry DMF (80 mL) was added and the solution was to approximately 70°C until biotin dissolved, then cooled to room temperature. Then triethylamine (TEA) was added (4μL, 3.04mg, 30mmol), and pentaflorophenyl trifluoroacetate (4.05 mL, 6.86 g, 24.91 mmol) was added slowly, and allowed to react for 1.5 at room temperature, turning the solution pale yellow. Residual

pentafluorophenyl trifluoroacetate and solvents were removed under reduced pressure. The crude product was purified via precipitation in diethylether (4.77g, 70% yield).

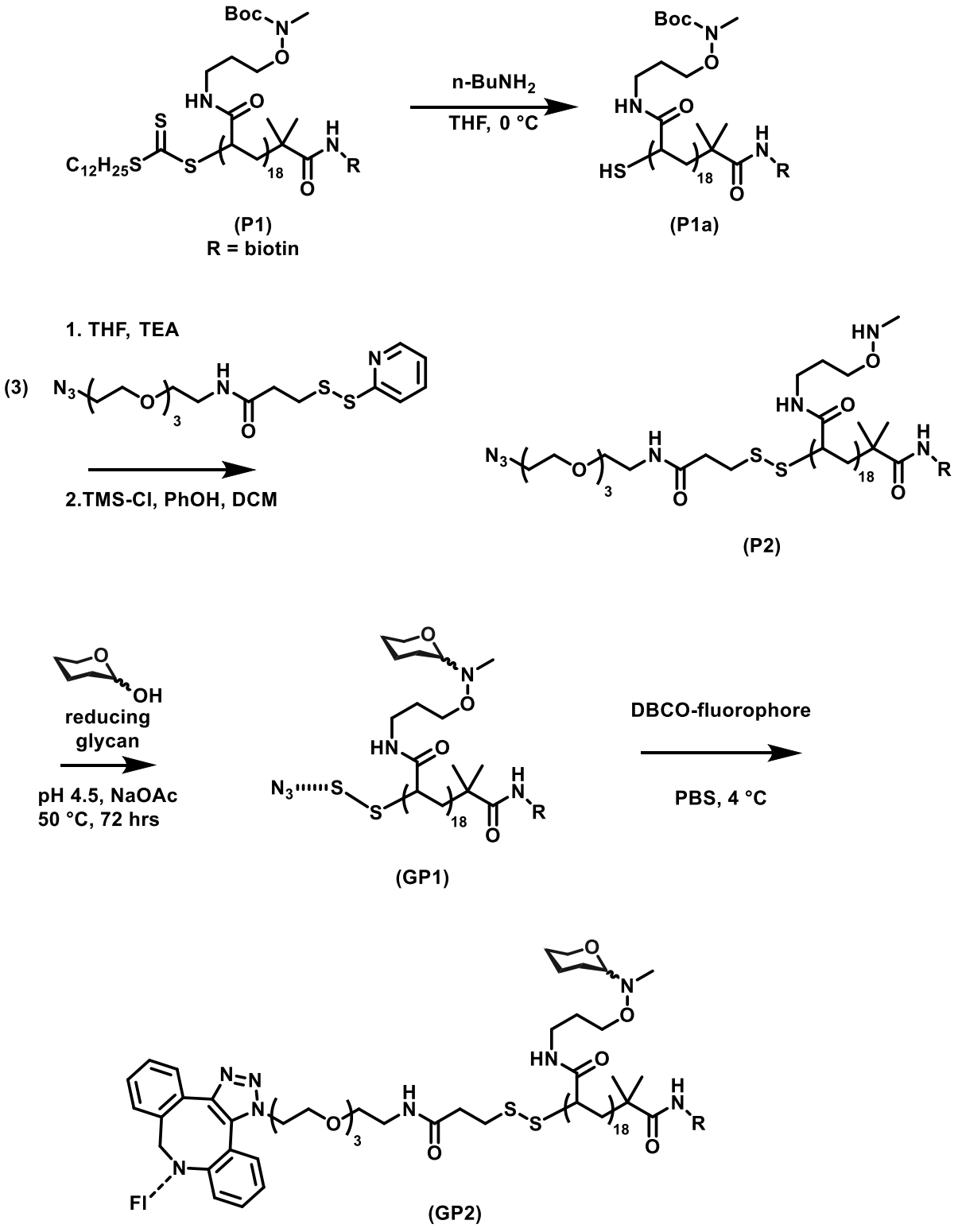
Biotin ethylaminediamine (2b)

Anhydrous DMF (30mL) was added to biotin-pentafluorophenyl ester **2a** (4.154g, 10.13 mmol). Ethylenediamine (6.12g, 6.8mL, 166.1mmol) was added to anhydrous solution dropwise and reaction was allowed to proceed for 4.5 hrs at room temperature while stirring before completion by TLC. The product was precipitated and washed in diethylether (2.25g, 77.6% yield).

Biotin-CTA (2)

A frame dried Schlenk flask (100 mL) was charged with magnetic stirring bar and CTA (500 mg, 1.36 mmol) under a stream of nitrogen. Anhydrous DMF (10 mL) was added, and the solution was cooled to 0 °C. Triethylamine (480 µL, 3.4 mmol, 2.5 eq) was added to solution. Next, trifluoroacetate pentafluorophenyl ester (260 µL,1.5 mmol, 1.5 eq) was added and the mixture was stirred for 1 hr at 0°C. Then, biotin ethylaminediamine (432 mg) was added to the stirring mixture, which was allowed to warm to room temperature and stirred vigorously overnight. Solvents were then removed under reduced pressure and the yellow residue was washed with diethylether, and the product was purified via flash chromatography via dry loading on a KP-sil SNAP cartridge.

The Biotinylated CTA product was isolated (229.1 mg, 25.9% yield). The product was characterized by NMR and mass spectrometry (Expected mass: 633.29, found: 633.27).



Scheme 6.2 Polymerization and modification of mucin mimetic glycopolymers

General procedure for the RAFT polymerization of tert-butyl (3-acrylamidopropoxy) methylcarbamate monomer

Procedure for RAFT polymerization with amino oxy monomer proceeded as previously described[34]. In brief, a flame-dried Schlenk flask (10 mL) equipped with a magnetic stirring bar was charged with biotin-CTA (29.4 mg, 46.5 µmol, 2.0 mol% with respect to monomer), AIBN (1.57 mg, 9.3 µmol, 0.4 mol% with respect to monomer), delivered as 500 µL of a 0.109 mM solution in anhydrous dioxane, amino oxy monomer (600 mg, 2.32 mmol), and anhydrous dioxane (443.3 mg). The flask was equipped with a rubber septum and filled with N_2. The resulting yellow solution was thoroughly degassed by several freeze-pump-thaw cycles. Next, the reaction mixture was warmed to room temperature, and then immersed into an oil bath preheated to 65 °C. After 2.5 hours, the viscous reaction mixture was then diluted in ether and precipitated into excess hexanes. The solid residue was then re-dissolved into ether and precipitated again in hexanes (2x). The yellow polymer was then concentrated from $CHCl_3$ three times to remove residual hexanes and dried under vacuum overnight to yield backbone **P1** as a pale-yellow solid (329.6 mg, 55%). 1H NMR (CDCl3, 500MHz) δ (ppm): 3.90-3.65 (bs, 2H), 3.35-2.80 (bm, 5H), 1.80-1.05 (bm, 16H). GPC (DMF, 0.2% LiBr): Mn = 5296 Da, PDI = 1.21, DP ≈ 18.

General procedure for the synthesis of end-deprotected polymers

A Schlenk flask (10 mL) equipped with a magnetic stir bar was charged with **P1** (1 mg, 3.44 µM) and 1.71 mL of a degassed 20 mM n-butylamine solution in THF. The Schlenk flask was submerged in an ice bath and allowed to react for 2 hours. The reaction

mixture was then diluted in ether and precipitated into excess hexanes with vigorous stirring. The precipitation step was performed three times. The polymer was concentrated from CHCl3 three times to remove residual hexanes and dried under vacuum overnight to give polymer intermediate as a white solid. ^{1}H NMR (CDCl3, 500 MHz) δ (ppm): 3.85 (bs, 2H), 3.27-3.08 (bm, 5H), 2.13-1.46 (bm, 15H).

General procedure for N-deprotection of polymers and disulfide exchange

A flame dried Schlenk flash containing (20 mg, 3.8 µmol) of end-deprotected **P1** was added dissolved into 2 mL of anhydrous THF, and the pyridyl disulfide linker **3** (4.56 µmol, 1.89 mg,1.2 Eq) was added carefully. 2 drops (~60 µL) triethylamine was added and the reaction was stirred vigorously for 48 hrs at RT, protected from light.

Deprotection was performed as previously described.[34] In short, a fresh solution containing TMS-Cl (1M) and phenol (3M) in anhydrous DCM was added to polymer intermediate from previous step, and the reaction was stirred vigorously for 2 hours. The deprotected polymer was precipitated in ether and washed 3x with ether, using centrifugation to isolate polymer between washes. The polymer was then purified using a PD-10 column (GE health sciences) and lyophilized to afford polymer **P2** (9.8 mg, 83 % yield).

Procedure for glycan ligation to polymers

In a PCR tube (0.2 mL), lyophilized polymer from the previous step was added to a sodium acetate buffer (1M NaOAc, 1M urea, pH 4.5) with 1.1 eq of the appropriate glycan, for a final concentration of 200 mM (by side-chain). The reaction mixture was held at 50 °C for 72 hrs, using a thermocycler (BioRad). To purify the glycosylated polymer, reactions were purified through addition to a pre-washed Amicon Ultra Centrifugal Filter (3K MWCO, Millipore), and were then spin dialyzed 4 times using deuterated phosphate buffered saline (100 mM phosphate, 150 mM NaCl, pH 7.4), discarding flow through each rinse. The spin column was inverted into a clean microcentrifuge tube to afford glycopolymer library **GP1.**

Quantification of glycopolymers using HABA biotin assay

Lyophilized 4-Hydroxyazobenzene-2-carboxylic acid (HABA)/avidin reagent (Sigma Aldrich) was dissolved in 10 mL milliQ H_2O, and 180 µL was added to a 96-well plate, and a pre-sample absorbance reading at 500 nm was taken. Then, 20 µL of polymer was added to each well, and the resulting decrease in absorbance at 500 nm was measured. Biotin concentration was calculated using a molar absorptivity of 34,000 M^{-1} cm^{-1}.

Ellman test to determine thiol content of polymer end groups

Using established protocols, a 10 mM stock of 5,5′-dithiobis-(2-nitrobenzoic acid) (Ellman's reagent) was made using fresh sodium bicarbonate, pH 8.4. Using a 96 well plate reader format, 100 µL Ellman's stock was mixed with 100 µL polymer sample containing thiol, and incubated at room temperature for 30 minutes. The resulting solution was measured for absorbance at 412 nm and the thiol content was calculated using a molar absorptivity of 14,160 M^{-1} cm^{-1}.

Bead binding assays

Prior to binding with mucin mimetics, Eppendorf tubes were blocked overnight with 1 % BSA in PBS at 4 °C. The next day, agarose beads (50 µL) functionalized with lectin or streptavidin were rinsed 3x with PBS, and then incubated with 1 % BSA in PBS for 1 hr at room temperature. Then, beads were rinsed and incubated with glycopolymer library **GP2** for 1 hour at 37 °C, before washing 3x with PBS prior to imaging analysis. Beads were pelleted by centrifugation (2000 xg) between washes.

Bacterial binding assays

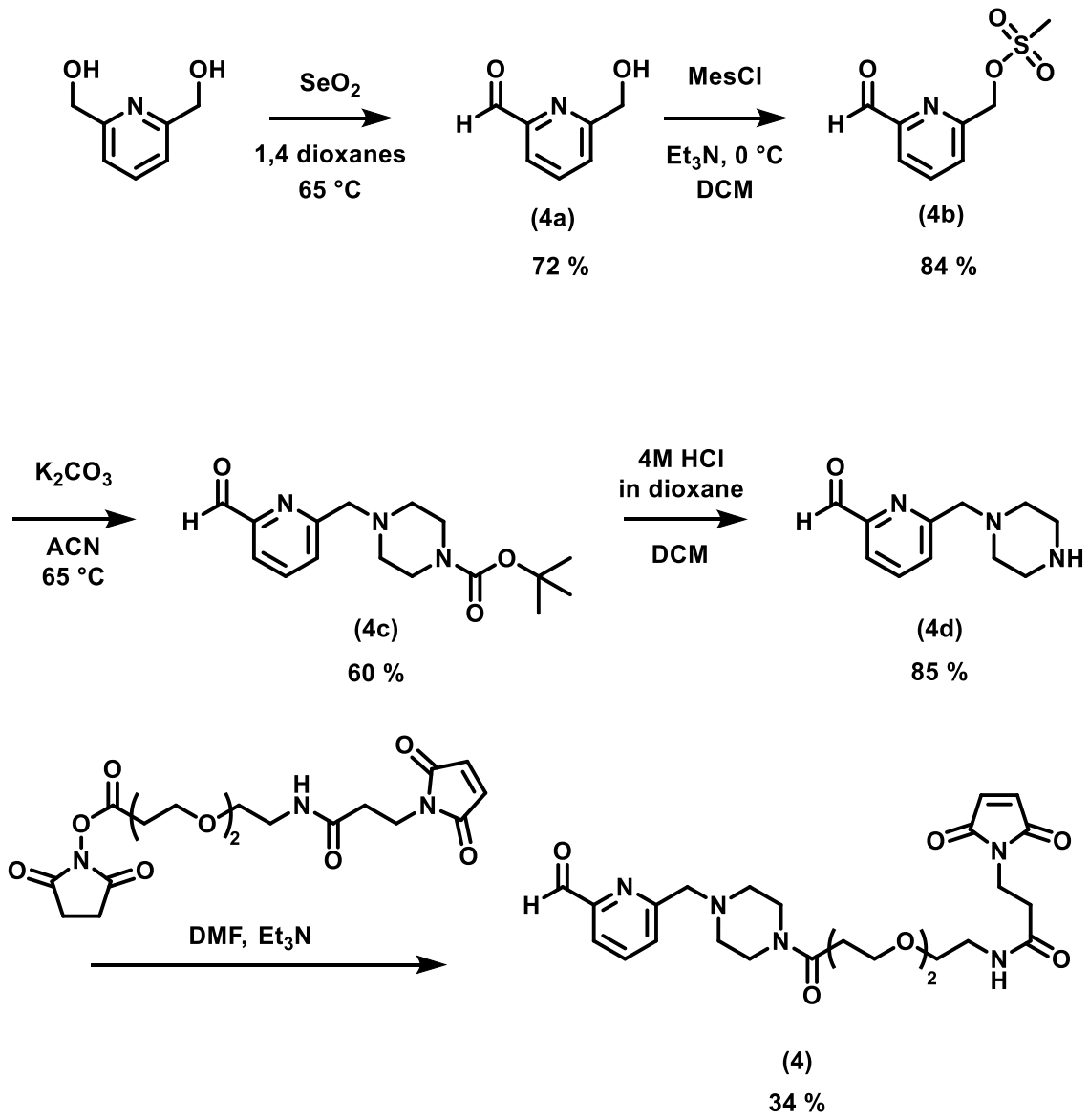
Using Eppendorf tubes blocked overnight with 1% BSA in PBS at 4 °C, bacterial stock (120 µL, OD_{600} = 0.6) of with MG1655 or ORN172 *E. Coli* were added to tubes and washed 3x with PBS, pelleting by centrifugation (1000xg, 3 min) in between washes. Then, mucin mimetics were incubated at the desired concentration for one hr at 37 °C to

bind. Bacteria were then washed 3x with PBS, with pelleting via centrifugation (1000 xg, 3 min) in between washes. Bacteria with bound mimetics were then analyzed by flow cytometry, microscopy, RT-PCR, or submitted for NGS.

To prepare stocks for binding experiments, *E. Coli* strains were grown, and then aliquoted at -80 °C for subsequent binding experiments. In a sterile culture tube, 4 mL of tryptic soy broth was inoculated with 1 µL of bacterial stock. The mixture was allowed to grow for 48 hours, until OD_{600} was 0.6. Then, stock was diluted 1:1 with glycerol and 100 µL aliquots were frozen at -80 °C for subsequent binding experiments.

Flow cytometry

After binding, washing and centrifugation, bacteria were resuspended in 50 µL PBS and incubated with streptavidin-Cy5 (Invitrogen, SA1011) at a 1:1000 dilution. Cells were then washed 3x to remove unbound Cy5-streptavidin, and fixed for 20 mins at room temperature using 4% paraformaldehyde. Fixed bacteria were pelleted and washed with PBS once, before being resuspended in 200 µL PBS for analysis by flow cytometry, using an Accuri BD6 flow cytometer.



Scheme 6.3 Generation of PCA-PEG$_2$-maleimide linker

6-(Hydroxymethyl)-2-pyridinecarboxaldehyde (4a)

In a Schlenk flask, 2,6-pyridinedimethanol (1.0g , 7.20 mmol) was dissolved in 20 mL of 1,4-dioxane and combined with 0.5 Eq selenium dioxide, stirring at 65 °C for 20 hours under N_2. The reaction mixture was then diluted with dichloromethane, filtered and

then purified chromatographically using 10% MeOH in DCM. Solvent was removed under reduced pressure to yield purified product (714.5 mg, 5.21 mmol, 72% yield).

6-(Formylpyridin-2-yl)methyl methanesulfonate (4b)

6-(Hydroxymethyl)-2-pyridinecarboxaldehyde product was added to 35 mL dry chloroform, and sonicated until dissolved. Solution was cooled to 0 °C and triethylamine (3 Eq) was added, and to this stirring solution, methanesulfonyl chloride (490 µL, 6.29 mmol) was added dropwise. The reaction mixture was allowed to warm to room temperature and stirred for one hour before addition of saturated aqueous sodium bicarbonate was added, and this aqueous layer was partitioned with dichloromethane three times, and the combined organic fraction was dried over sodium sulfate and concentrated (941 mg, 4.38 mmol, 84 % yield).

Tert-butyl 4-((6-formylpyridin-2-yl)methyl)piperazine-1-carboxylate (4c)

6-(Formylpyridin-2-yl)methyl methanesulfonate from previous reaction was dissolved in 15 mL dried acetonitrile containing potassium carbonate (1222 mg, 2 Eq) and 1-boc-piperizine (990 mg, 5.28 mmol,1.2 Eq) under N_2 and stirred at 60 °C overnight. The next day, the reaction solvent was removed and the product was extracted in dichloromethane and saturated aqueous sodium bicarbonate. The resulting crude material was purified chromatographically using 33% ethylacetate in hexanes, to afford product (801.5 mg, 2.62 mmol, 60 % yield).

6-(piperazin-1-ylmethyl)-2-pyridinecarboxaldehyde HCl salt (4d)

Tert-butyl 4-((6-formylpyridin-2-yl)methyl)piperazine-1-carboxylate (294.1 mg, 0.96 mmol) from the previous reaction was dissolved in 8.4 mL dichloromethane and combined with 4 M HCl in 1,4-dioxane (2.4 mL, 9.62 mmol, 10 Eq). After 1 hour stirring, solids had formed and the reaction solvents were removed and the hydrgoscopic product was dried under vacuum overnight, and the tan solid was stored in a desiccator. (85 % yield)

Functionalization of 6-(piperazin-1-ylmethyl)-2-pyridinecarboxaldehyde with NHS-PEG$_2$-maleimide (4)

6-(piperazin-1-ylmethyl)-2-pyridinecarboxaldehyde HCl salt (7.5 mg, 0.04 mmol) and NHS-PEG$_2$-maleimide linker (15 mg, 0.04 mmol) were combined and dissolved in 1.0 mL dried DMF. Then, triethylamine (112 µL, 0.80 mmol, 20 Eq) was added dropwise and the reaction was stirred for 1 hr at RT. Reaction was concentrated under reduced pressure, and subjected to flash column chromatography to isolate product as a clear oil (34 % yield).

Generation of modification of SAvP endgroups with DNA barcodes

Purified recombinant SAvP (Thermofisher) was dissolved into PBS (100 mM) at a concentration of 2 mg/mL (129 µM). The maleimide containing PCA derivative **4** was added to the reaction mixture to a concentration of 10 mM, and put on the shaker

overnight, at 37 °C. The resulting modified protein was purified using size exclusion column and lyophilized overnight to afford end-modified proteins (86% recovery).

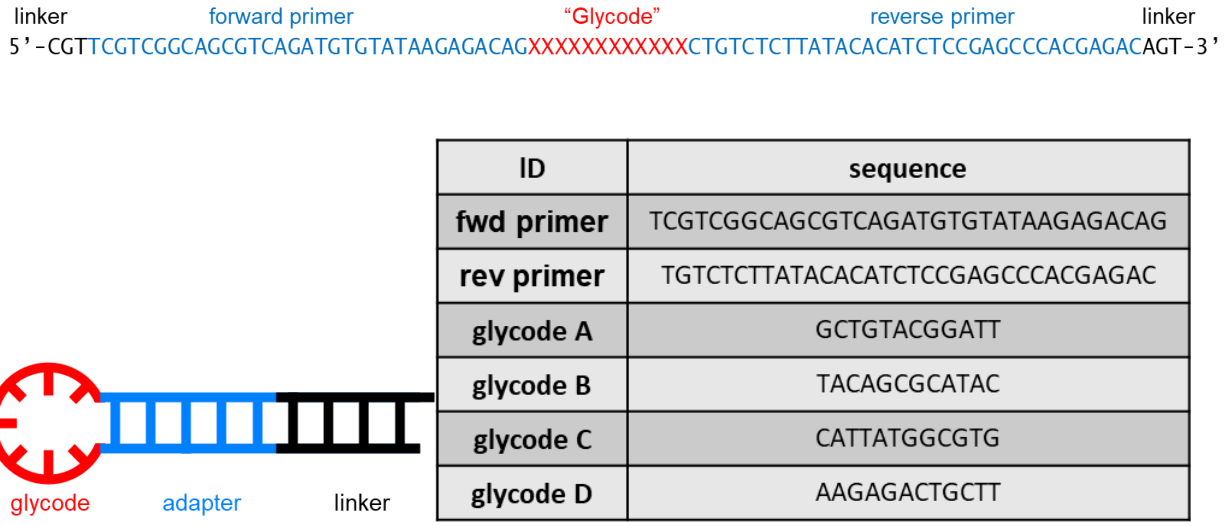
The next day, thiol bearing DNA (integrated DNA technologies) was deprotected using established procedures. In short, 2 µL of DNA stock (2 mM) was incubated with 2 µL of PBS containing 200 mM TCEP for one hour at room temperature. Following deprotection, a solution of 500 µL HPLC grade acetone with 10 µL 3M $LiClO_4$ was added to DNA to precipitate deprotected oligonucleotides, which were pelleted by spinning at 14,000 xg at 4 °C. This process was repeated three times. Purified oligonucleotides were quantitated using absorbance at 260 nm and then immediately reacted with SAvP bearing terminal maleimide groups in PBS using a 1:1 molar ratio, overnight at 4 °C.

DNA barcoded SAvP was purified using NTA resin chromatography. In an Eppendorf tube, 300 µL of HisPur Ni-NTA superflow agarose (Thermo Scientific) was bound with SAvP-DNA on rotisserie for 60 minutes at room temperature. The resin containing bound DNA-SAvP was loaded into a manual column, and washed with 6 mL of a wash buffer (20 mM sodium phosphate, 300 mM NaCl, 10 mM imidazole, pH 7.4). Afterwards, DNA SAvP was eluted with 2 mL of elution buffer (20 mM sodium phosphate, 300 mM NaCl, 300 mM imidazole, pH 7.4). The resulting elutant was then desalted using a disposable PD-10 size exclusion column to remove excess imidazole for downstream applications.

DNA Barcodes

DNA barcodes were designed for compatibility with NGS. An 86 bp oligonucleotide, or “glycode”, was designed with compatibility with Illumina Nextera

adapters for sequencing, by incorporating sequencing primers flanking a 12 bp barcode. The following sequences were used for NGS tags.



ID	sequence
fwd primer	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG
rev primer	TGTCTCTTATACACATCTCCGAGCCCACGAGAC
glycode A	GCTGTACGGATT
glycode B	TACAGCGCATAC
glycode C	CATTATGGCGTG
glycode D	AAGAGACTGCTT

Figure 6.6 Diagram of DNA barcodes for NGS with full sequences. Illumina Nextra adapters were installed flanking a 12 bp "glycode", for compatibility with NGS technology.

RT-PCR readout of binding assays

After washing and binding of mucin mimetic **GP2** library, 1 µL of cell suspension was transferred as template to a PCR mixture containing 1x oneTaq master mix (New England Biolabs M0482), 1x SYBR green (Thermo S7567), and 0.2 µM forward and reference primer. The resulting mixture was subjected to PCR under standard conditions.

General procedure for mucin mimetic binding assays

Eppendorf tubes used for binding assays were blocked with 1% BSA in PBS (100mM) at 4 °C overnight. The next day, tubes were washed thoroughly with PBS, and

either 50 µL beads or frozen *E. coli* stock (OD_{600} = 0.6) were incubated against polymer-SAvP-DNA conjugates for one hour at 37 °C. Bound conjugates was washed three times, spinning at 3000 xg for 4 minutes to pellet between washes. Binding results were then assessed using qPCR (Biorad CFX) or amplified for 16s sequencing.

Release and quantitation of fluorophore

To assess the ability of reducing agents to cleave disulfide-linked fluorophore, mucin mimetic library **GP2** was incubated with streptavidin beads for 1 hour at room temperature, after which the bead pool was split. To one half, β-mercaptoethanol (100 µM) was added and incubated at room temperature for one hour before washing 3x with PBS and imaging afterwards.

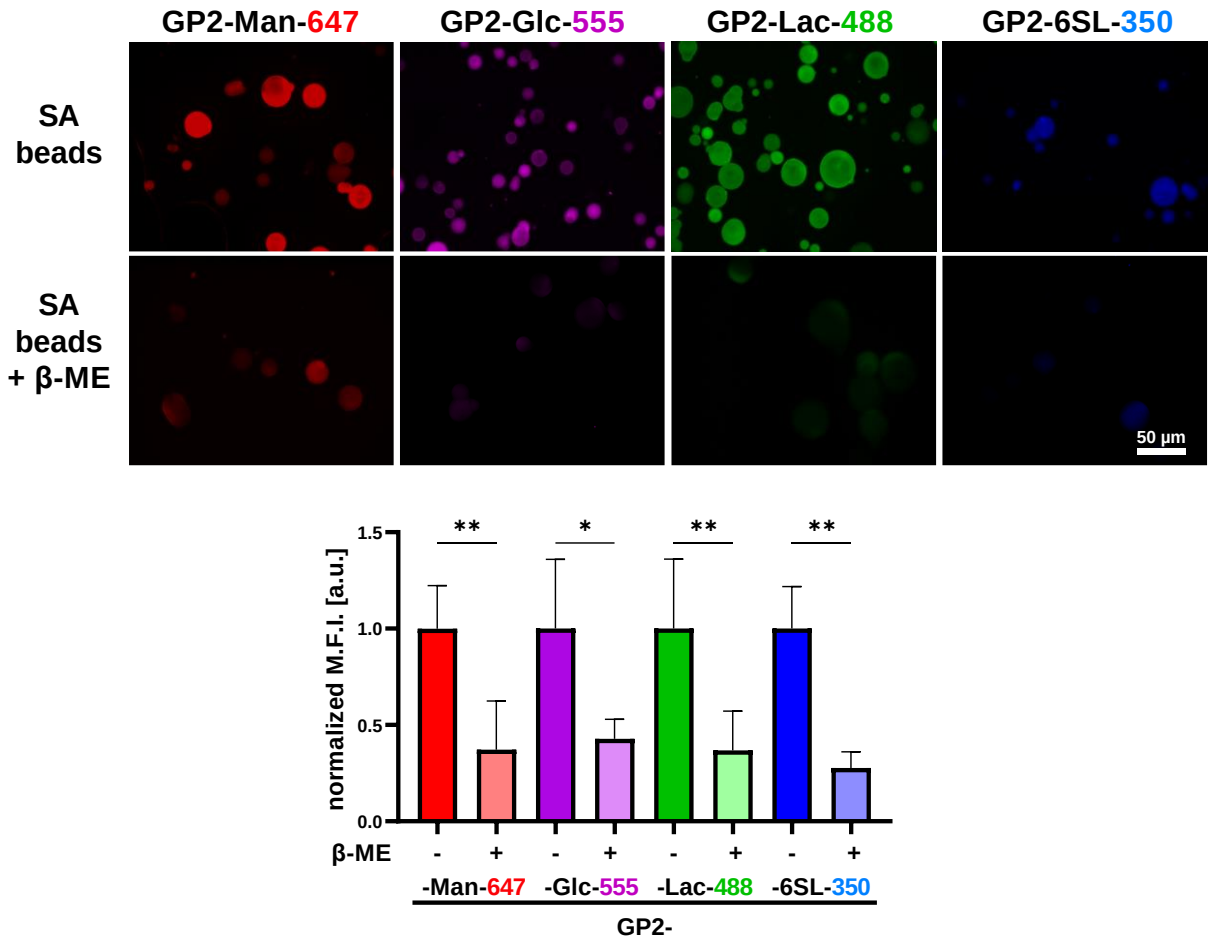


Figure 6.7 Binding to mucin mimetics to streptavidin beads and subsequent fluorophore release. Binding of mucin mimetic library GP2 with streptavidin beads and subsequent release of fluorophore upon incubation with 100 µM β-mercaptoethanol for one hour.

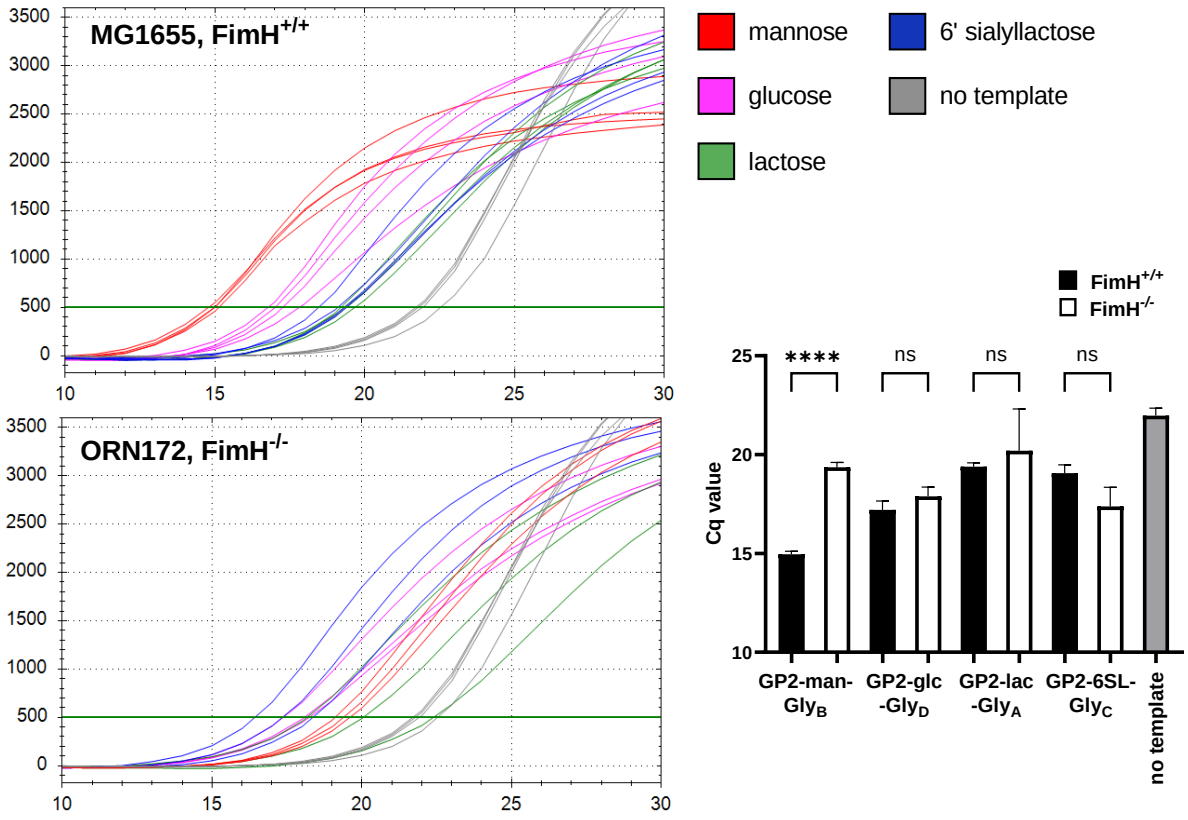


Figure 6.8 qPCR amplification curves for mucin mimetic library GP2 binding to E. coli. Full PCR data for figure 6.5 shows binding between GP2-man-Gly_B mucin mimetics and MG1655 *E. Coli*.

6.5 Acknowledgements

Chapter 6 was original work primarily conducted by Austen Larson Michalak. Ethan J. Armand developed “glycode extractor” algorithm to process sequencing data sets.

6.6 References:

1 Werlang C, Cárcarmo-Oyarce G, Ribbeck K. Engineering mucus to study and influence the microbiome. *Nat. Rev. Mater.*, **2019**, 4**:**134–145.

2 Johansson ME, Larsson JM, Hansson GC. The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. *Proc. Natl. Acad. Sci.*, **2011** 108,1: 4659-65.

3 Martín SP, Seeberger PH, Silva DH. Mucins and Pathogenic Mucin-Like Molecules Are Immunomodulators During Infection and Targets for Diagnostics and Vaccines. *Front. Chem.* **2019**, 7:710

4 Paone P, Cani PD. Mucus barrier, mucins and gut microbiota: the expected slimy partners? *Gut,* **2020**, 69:2232-2243.

5 Rose MC. Mucins: structure, function, and role in pulmonary diseases. *Am. J. Physiol.* **1992**, 263,4: 413-429

6 Hattrup CL, Gendler SJ. Structure and Function of the Cell Surface (Tethered) Mucins. *Ann. Rev. of Physio.*, **2008**, 70:431-457

7 McGuckin MA, Lindén SK, Sutton P, Florin TH. Mucin dynamics and enteric pathogens. *Nat Rev Microbiol*, **2011,** 9, 265–278

8 Kudelka MR, Stowell SR, Cummings RD, Neish AS. Intestinal epithelial glycosylation in homeostasis and gut microbiota interactions in IBD. *Nat Rev Gastroenterol Hepatol*, **2020** 17,597-617

9 Hansson GC, Mucus and mucins in diseases of the intestinal and respiratory tracts. *J. Intern. Med.*, **2019** 285(5):479-490

10 Corfield, AP. Mucins: A biologically relevant glycan barrier in mucosal protection. *Biochimica et Biophysica Acta (BBA) - General Subjects*, **2015** 1850,1: 236-252

11 Tailford LE, Crost EH, Kavanaugh D, Juge N. Mucin glycan foraging in the human gut microbiome. *Front. In Genetics.* **2015** 6,81

12 Paton B, Suarez M, Herrero P, Canela N. Glycosylation Biomarkers Associated with Age-Related Diseases and Current Methods for Glycan Analysis *Int. J. Mol. Sci.* **2021**, 22(11): 5788

13 Elderman M, Sovran B, Hugenholtz F, Graversen K, Huijskes M, Houtsma E, Belzer C, Boekschoten M, Vos P, Dekker J, Wells J, Faas M. The effect of age on the intestinal mucus thickness, microbiota composition and immunity in relation to sex in mice. *Plos One*. **2017**

14 Sicard JF, Bihan GL, Vogeleer P, Jacques M, Harel J. Interactions of Intestinal Bacteria with Components of the Intestinal Mucus. *Front. In Cell. And Infect. Microbiol.* **2017**, 7(387)

15 Corfield AP. The Interaction of the Gut Microbiota with the Mucus Barrier in Health and Disease in Human. *Microorganisms.* **2018** 6(3):78

16 Gao C, Wei M, McKitrick TR, McQuillan AM, Heimburg-Molinaro J, Cummings RD. Glycan Microarrays as Chemical Tools for Identifying Glycan Recognition by Immune Proteins. *Front. Chem.* **2019**, 7(833)

17 Rillahan CD, Paulson JC. Glycan Microarrays for Decoding the Glycome. *Annu. Rev. Biochem.*, **2011**. 80:797–823

18 Zhu X, Wang Y, Wang L, Orndorff P, Guo A. Quantitative Glycomics from Fluidic Glycan Microarrays. *J. Am. Chem. Soc.*, **2009** 131,13646-13650

19 Jonczyk R, Kurth T, Lavrentieva A, Walter JG, Scheper T, Stahl F. Living Cell Microarrays: An Overview of Concepts. *Microarrays* **2016**, 5:1-29.

20 Purohit S, Li T, Guan W, Song X, Song J, Tian T, Li L, Sharma A, Dun B, Mysona D, Ghamande S, Rungruang B, Cummings RD, Wang PG, She J. Multiplex glycan bead array for high throughput and high content analyses of glycan binding proteins. *Nature comm.* **2018**, 9:258

21 Nimrichter L, Gargir A, Gortler M, Altstock R, Shtevi A, Weisshaus O, Fire E, Dotan N, Schnaar R. Intact cell adhesion to glycan microarrays. *Glycobiology,* **2004**, 14,2,197-203

22 Disney M, Seeberger P. The Use of Carbohydrate Microarrays to Study Carbohydrate-Cell Interactions and to Detect Pathogens. *Chem. & Biol.* **2004**, 11, 1701-1707

23 Wang L, Cummings RD, Smith DF, Huflejt M, Campbell CT, Gildersleeve JC, Gerlach JQ, Kilcoyne M, Joshi L, Serna S, Reichardt NC, Parera PN, Pieters RJ, Eng W, Mahal LK. Cross-platform comparison of glycan microarray formats. *Glycobiol.* **2014**, 24(6),507-517

24 Song X, Xia B, Stowell SR, Lasanajak Y, Smith D., Cummings R. Novel Fluorescent Glycan Microarray Strategy Reveals Ligands for Galectins. *Chem. & Biol.* **2009**, 16,36-47

25 Yan M, Zhu Y, Liu X, Lasanajak Y, Xiong J, Lu J, Lin X, Ashline D, Reinhold V, Smith DF, Song X. Next-Generation Glycan Microarray Enabled by DNA-Coded Glycan Library and Next-Generation Sequencing Technology. *Analytical Chemistry **2019***, 91(14), 9221-9228

26 Kondengaden SM, Zhang J, Zhang H, Parameswaran A, Kondengadan SM, Pawar S, Puthengot A, Sunderraman R, Song J, Polizzi SJ, Wen L, Wang PJ. DNA Encoded Glycan Libraries as a next-generation tool for the study of glycan-protein interactions. *bioRxiv* **2020**.03.3 0.017012

27 Sojitra M, Sarkar S, Maghera J, Rodrigues E, Carpenter EJ, Seth S, Vinals DF, Bennett NJ, Reddy R, Khalil A, Xue X, Bell MR, Zheng RB, Zhang P, Nycholat C, Bailey JJ, Ling CC, Lowary TL, Paulson JC, Macauley MS, Derda R. Genetically encoded multivalent liquid glycan array displayed on M13 bacteriophage. *Nat. Chem. Biol.,* **2021**, 17**,** 806–816.

28 Authimoolman SP, Dziubia TD. Biopolymeric Mucin and Synthetic Polymer Analogs: Their Structure, Function, and Role in Biomedical Applications. *Polymers.* **2016**, 8(3), 71

29 Lucas TM, Gupta C, Altman MO, Sanchez E, Naticchia MR, Gagneux P, Singharoy A, Godula K. Mucin-Mimetic Glycan Arrays Integrating Machine Learning for Analyzing Receptor Pattern Recognition by Influenza A Viruses. *Cell Press Chem,7,3393-3411*

30 Weichhart T, Zlabinger GJ, Säemann MD. The multiple functions of Tamm–Horsfall protein in human health and disease: A mystery clears up. *Wien Klin Wochenschr, **2005**,* 117, 316–322

31 Taganna J, de Boer AR, Wuhrer M, Bouckaert J. Glycosylation changes as important factors for the susceptibility to urinary tract infection. *Biochem. Soc. Trans*. **2011**, 39(1):349-54.

32 Lenger SM, Bradley MS, Thomas DA, Bertolet MH, Lowder JL, Sutcliffe S. D-mannose vs other agents for recurrent urinary tract infection prevention in adult women: a systematic review and meta-analysis. *Am. J. Obstet. Gynecol*. **2020**;223(2):265

33 MacDonald JI, Munch HK, Moore T, Francis MB. “One-step site-specific modification of native proteins with 2-pyridinecarboxyaldehydes.” *Nature chemical biology*, **2015**, 11,5:326-31.

34 Huang M, Smith R, Trieger G, Godula K. Glycocalyx Remodeling with Proteoglycan Mimetics Promotes Neural Specification in Embryonic Stem Cells. *J. Am. Chem. Soc.* **2014**, 136 (3), 10565-10568